Methods in Molecular Biology 1928

Springer Protocols

Majda Haznadar *Editor*

Cancer Metabolism

Methods and Protocols

EXTRAS ONLINE

Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor
John M. Walker
School of Life and Medical Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:
http://www.springer.com/series/7651

Cancer Metabolism

Methods and Protocols

Edited by

Majda Haznadar

National Cancer Institute, North Bethesda, MD, USA

Humana Press

Editor
Majda Haznadar
National Cancer Institute
North Bethesda, MD, USA

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-9026-9 ISBN 978-1-4939-9027-6 (eBook)
https://doi.org/10.1007/978-1-4939-9027-6

Library of Congress Control Number: 2018966343

Cover Illustations: Imaging fatty acid synthesis in a mouse model of lung adenocarcinoma with [11C]acetate PET/CT. Image courtesy of David Lewis, Beatson Institute, Glasgow.

This Humana Press imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.
The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

Preface

In 1956, Otto Heinrich Warburg published seminal research wherein he described a novel process by which cancer cells survive, the process of relying on aerobic glycolysis rather than oxidative phosphorylation, used by normal differentiated cells. This novel process has been termed "the Warburg effect." Since then, researchers have been working on discerning how cancer cells adapt some and bypass other metabolic processes, in order to fulfill their high energy needs. Some more recent studies have also illuminated processes by which cancer cells utilize external energy sources to generate building blocks required for their high energy turnover. This volume of *Methods in Molecular Biology* series is designed to provide common experimental approaches in studies designed to illuminate various processes studied in cancer metabolism. Each chapter opens up with the theory behind the method being described. Every laboratory protocol chapter includes chemicals and reagents necessary to carry out a given protocol. Each methods section of laboratory protocol chapters contains detailed step-by-step description for successful completion of that method. The book is divided into three compartments. The first part of the book focuses on specific protocols commonly utilized in cancer metabolism studies, such as protocols comprising stable isotope labeling methods, protocols for studying glycolysis, gluconeogenesis, and mitochondrial metabolism. This portion of the book also includes chapters containing imaging tools to study cellular metabolism, as well as descriptions of tools to study macrophages and autophagy and their relevance in cancer. The second part of the book is geared toward describing methods used for generating hypotheses and identifying cancer markers, such as mass spectrometry- and NMR-based profiling tools, as well as a protocol with a description of tools for studying the human microbiome. This portion of the book also contains two protocols utilizing Seahorse, a commonly used laboratory platform for studying cell metabolism. The last part of the book is designed to describe an overview of vital and actively researched topics in the field of cancer metabolism, as well as computational methodological approaches. We hope that this book will become an essential compilation of protocols utilized in many laboratories that conduct this type of research.

North Bethesda, MD, USA *Majda Haznadar*

Contents

Contributors

Basheer Alas • *Tumor Initiation and Maintenance Program, NCI-Designated Cancer Center, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA*

Shibdas Banerjee • *Indian Institute of Science Education and Research Tirupati, Tirupati, Andhra Pradesh, India*

Oliver F. Bathe • *Arnie Charbonneau Cancer Institute, University of Calgary, Calgary, AB, Canada; Department of Surgery, Tom Baker Cancer Center, University of Calgary, Calgary, AB, Canada; Department of Oncology, Tom Baker Cancer Center, University of Calgary, Calgary, AB, Canada*

Robert B. Bentham • *Department of Cell and Developmental Biology, Consortium for Mitochondrial Research, University College London, London, UK*

Thomas S. Blacker • *Research Department of Cell & Developmental Biology, University College London, London, UK; UCL Consortium for Mitochondrial Research, University College London, London, UK; Department of Physics & Astronomy, University College London, London, UK*

Kevin Bryson • *Department of Computer Sciences, University College London, London, UK*

Alessandro Carrer • *Department of Cancer Biology, Abramson Family Cancer Research Institute, University of Pennsylvania, Philadelphia, PA, USA*

Tanushree Chakraborty • *Biophysics and Structural Genomics Division, Saha Institute of Nuclear Physics (HBNI), Kolkata, India*

Jun Chen • *Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA*

David Kung-Chun Chiu • *Department of Pathology, Queen Mary Hospital, The University of Hong Kong, Pokfulam, Hong Kong*

Cosimo Commisso • *Tumor Initiation and Maintenance Program, NCI-Designated Cancer Center, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA*

Tomer Cooks • *Faculty of Health Sciences, The Shraga Segal Department of Microbiology, Immunology and Genetics, Ben Gurion University of the Negev, Beer Sheva, Israel*

Sarah Creed • *Monash Micro Imaging, Hudson Institute of Medical Research, Clayton, VIC, Australia*

Daniel R. Crooks • *Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA*

Michael R. Duchen • *Research Department of Cell & Developmental Biology, University College London, London, UK; UCL Consortium for Mitochondrial Research, University College London, London, UK*

Teresa W. -M. Fan • *Department of Toxicology and Cancer Biology, Center for Environmental and Systems Biochemistry, Markey Cancer Center, University of Kentucky, Lexington, KY, USA*

Farshad Farshidfar • *Department of Oncology and Arnie Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, Canada*

Koen M. O. Galenkamp • *Tumor Initiation and Maintenance Program, NCI-Designated Cancer Center, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA*

Elyssia S. Gallagher • *Department of Chemistry and Biochemistry, Baylor University, Waco, TX, USA*

Benjamin A. Garcia • *Department of Biochemistry and Biophysics, Perelman School of Medicine, Epigenetics Institute, University of Pennsylvania, Philadelphia, PA, USA*
K. Leigh Greathouse • *Robbins College of Health and Human Sciences, Baylor University, Waco, TX, USA*
Dan He • *Department of Biology, College of Arts and Sciences, Drexel University, Philadelphia, PA, USA*
Myriam Y. Hsu • *Department of Molecular Biotechnology and Health Science, Molecular Biotechnology Center, University of Torino, Torino, Italy*
M. Eric Hyndman • *Division of Urology, Department of Surgery, Southern Alberta Institute of Urology, University of Calgary, Calgary, AB, Canada*
Ramiz Islam Sk • *Biophysics and Structural Genomics Division, Saha Institute of Nuclear Physics (HBNI), Kolkata, India*
Mikko I. Kettunen • *A.I. Virtanen for Molecular Sciences, University of Eastern Finland, Kuopio, Finland*
Christiaan F. Labuschagne • *The Francis Crick Institute, London, UK*
David Y. Lewis • *Molecular Imaging Laboratory, Cancer Research UK Beatson Institute, Glasgow, UK; Institute of Cancer Sciences, University of Glasgow, Glasgow, UK*
W. Marston Linehan • *Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA*
Oliver D. K. Maddocks • *Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, Glasgow, UK*
Soumen Kanti Manna • *Biophysics and Structural Genomics Division, Saha Institute of Nuclear Physics (HBNI), Kolkata, India*
Ewy Mathé • *The Ohio State University College of Medicine, Columbus, OH, USA*
Matthew McKenzie • *Centre for Innate Immunity and Infectious Diseases, Hudson Institute of Medical Research, Clayton, VIC, Australia; Department of Molecular and Translational Science, Monash University, Clayton, VIC, Australia; School of Life and Environmental Sciences, Deakin University, Waurn Ponds, VIC, Australia*
Beata Mickiewicz • *Department of Biological Sciences, Bio-NMR-Centre, University of Calgary, Calgary, AB, Canada*
Elisa Morales • *Robbins College of Health and Human Sciences, Baylor University, Waco, TX, USA*
Alice C. Newman • *Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, Glasgow, UK*
Ioannis S. Pateras • *Molecular Carcinogenesis Group, Department of Histology and Embryology, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece*
Andy Patt • *The Ohio State University College of Medicine, Columbus, OH, USA*
Amanda J. Pearson • *Department of Chemistry and Biochemistry, Baylor University, Waco, TX, USA*
Paolo E. Porporato • *Department of Molecular Biotechnology and Health Science, Molecular Biotechnology Center, University of Torino, Torino, Italy*
Guillem Pratx • *Division of Medical Physics, Department of Radiation Oncology, Stanford University, Stanford, CA, USA*
Shuo Qie • *Department of Biology, College of Arts and Sciences, Drexel University, Philadelphia, PA, USA; Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, SC, USA*

Lake-Ee Quek • *School of Mathematics and Statistics, The University of Sydney, Sydney, NSW, Australia*
Jodi Rattner • *Arnie Charbonneau Cancer Institute, University of Calgary, Calgary, AB, Canada*
Nianli Sang • *Department of Biology, College of Arts and Sciences, Drexel University, Philadelphia, PA, USA; Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA, USA*
Debanti Sengupta • *Division of Medical Physics, Department of Radiation Oncology, Stanford University, Stanford, CA, USA*
Michael D. E. Sewell • *Research Department of Cell & Developmental Biology, University College London, London, UK*
Jalal Siddoqui • *The Ohio State University College of Medicine, Columbus, OH, USA*
Simone Sidoli • *Department of Biochemistry and Biophysics, Perelman School of Medicine, Epigenetics Institute, University of Pennsylvania, Philadelphia, PA, USA*
Gyorgy Szabadkai • *Department of Cell & Developmental Biology, University College London, London, UK; Consortium for Mitochondrial Research, University College London, London, UK; The Francis Crick Institute, London, UK; Department of Biomedical Sciences, University of Padua, Padua, Italy*
Sophie Trefely • *Department of Cancer Biology, Abramson Family Cancer Research Institute, University of Pennsylvania, Philadelphia, PA, USA; A. J. Drexel Autism Institute, Drexel University, Philadelphia, PA, USA*
Aki Pui-Wah Tse • *Department of Pathology, Queen Mary Hospital, The University of Hong Kong, Pokfulam, Hong Kong*
Nigel Turner • *Department of Pharmacology, School of Medical Sciences, UNSW Australia, Sydney, NSW, Australia*
Hans J. Vogel • *Department of Biological Sciences, Bio-NMR-Centre, University of Calgary, Calgary, AB, Canada*
Natalia von Muhlinen • *Laboratory of Human Carcinogenesis, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA*
Karen H. Vousden • *The Francis Crick Institute, London, UK*
Timothy H. Witney • *Imaging Chemistry and Biology, King's College London, London, UK*
Carmen Chak-Lui Wong • *Department of Pathology, Queen Mary Hospital, The University of Hong Kong, Pokfulam, Hong Kong*
Elisabeth Wyart • *Department of Molecular Biotechnology and Health Science, Molecular Biotechnology Center, University of Torino, Torino, Italy*
Bofei Zhang • *The Ohio State University College of Medicine, Columbus, OH, USA*
Jing Zhang • *Department of Pathology and Laboratory Medicine, Lineberger Comprehensive Cancer Center, UNC-Chapel Hill, Chapel Hill, NC, USA*
Misty Shuo Zhang • *Department of Pathology, Queen Mary Hospital, The University of Hong Kong, Pokfulam, Hong Kong*
Qing Zhang • *Department of Pathology and Laboratory Medicine, Lineberger Comprehensive Cancer Center, UNC-Chapel Hill, Chapel Hill, NC, USA; Department of Pharmacology, UNC-Chapel Hill, Chapel Hill, NC, USA*
Tong Zhang • *Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, Garscube Estate, Glasgow, UK*

Chapter 1

Metabolic Labeling of Cultured Mammalian Cells for Stable Isotope-Resolved Metabolomics: Practical Aspects of Tissue Culture and Sample Extraction

Daniel R. Crooks, Teresa W.-M. Fan, and W. Marston Linehan

Abstract

Stable isotope-resolved metabolomics (SIRM) methods are used increasingly by cancer researchers to probe metabolic pathways and identify vulnerabilities in cancer cells. Analytical and computational advances are being made constantly, but tissue culture and sample extraction procedures are often variable and not elaborated in the literature. This chapter discusses basic aspects of tissue culture practices as they relate to the use of stable isotope tracers and provides a detailed metabolic labeling and metabolite extraction procedure designed to maximize the amount of information that can be obtained from a single tracer experiment.

Key words Tissue culture, Stable isotopes, Metabolomics, Metabolite extraction, Glutamine, Glucose

1 Introduction

At present, numerous publications and protocols exist in the literature detailing the theoretical and practical aspects of conducting stable isotope tracer experiments in mammalian cells, 3D organoid cultures, tumor tissue slices, animal models, and even in human patients [1–5]. Recent work has demonstrated that in vivo tumor metabolism can differ significantly from that observed in tumor cell monoculture [5–9]. Nevertheless, two-dimensional cell monoculture stable isotope tracer experiments are still heavily relied upon to draw conclusions about the nature of metabolic pathways in mammalian cells and tissues [10–14]. The continued reliance of researchers on 2D monoculture cell culture systems for stable isotope tracer experiments is due to a variety of factors, including the experimental power conferred by the ready control of the extracellular environment, the feasibility of extensive genetic manipulations, the lower cost and simple infrastructure needed for performing tissue culture experiments relative to animal

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_1,

and/or human patient work, and the ability to rapidly test a large variety of experimental conditions [1].

On a practical basis, access to analytical platforms, data analysis, and biological interpretation typically comprise the bottleneck in metabolomics workflows. Given the enormous amount of time that must be invested for data analysis and interpretation, careful and thorough consideration should be applied to the practical aspects of design and execution of stable isotope tracer experiments. Otherwise, subsequent in-depth analyses and interpretation efforts will be compromised by avoidable flaws in experimental design and execution. The purpose of this chapter is to provide a simple and practical laboratory protocol for stable isotope labeling of cultured mammalian cells in 2D culture in the typical cancer research laboratory setting. The growing number of cancer researchers wishing to conduct targeted stable isotope tracer experiments with their cells often do not possess the analytical instrumentation and expertise to perform mass spectrometry and NMR analyses of metabolites in their own labs. Thus, it is becoming increasingly common for researchers to perform the tracer labeling experiments in their own lab and send the resulting samples or quenched metabolite extracts to a core facility or company to perform the subsequent analyses. Data acquisition and targeted analysis in stable isotope experiments will not be discussed here, but these topics are well covered by other comprehensive literature, e.g., [3, 15–17].

The protocol detailed in Subheading 3.1 covers the practical aspects of the stable isotope labeling procedure for adherent mammalian cells. The subsequent metabolite quenching and extraction procedures tend to vary significantly between different laboratories, academic core facilities, and commercial metabolomics companies. For reference, a robust metabolite extraction protocol is detailed in Subheading 3.2; this procedure was developed and refined by Teresa W.-M. Fan, Andrew N. Lane, and Richard Higashi and represents a careful and comprehensive approach to the quantitative recovery of polar and nonpolar metabolites as well as total cellular protein from mammalian cells. This extraction procedure has several attractive features, including:

1. Initial rapid and complete inactivation of intracellular enzymes using cold 100% acetonitrile.
2. Quantitative and reproducible recovery of a broad complement of polar and nonpolar intracellular metabolites from a single sample using a sequential combination of acetonitrile, water, chloroform, and methanol.
3. Safeguards for minimizing degradation of labile polar and nonpolar metabolites during the extraction and sample reduction steps,

4. Versatile use of the resulting polar and nonpolar metabolite extracts and protein residues, e.g., analysis by both NMR and multiple modalities of mass spectrometry, as well as Western blotting and other proteomic analyses.

Although this metabolite extraction procedure is more laborious than most metabolite extraction protocols, the above features maximize the amount of information that can be obtained from a single stable isotope tracer experiment. Finally, Subheading 3.3 outlines two simple extraction procedures for deproteinization and recovery of culture medium extracts for subsequent metabolomics analysis.

1.1 Experimental Design

Excellent summaries of the theoretical and analytical considerations involved in the design and interpretation of SIRM experiments can be found in the following references [2, 3, 18]. A discussion of preclinical cancer models that extend beyond 2D monoculture of tumor cell lines has also been published recently [1]. Briefly, experimental design considerations include consideration of which metabolites are of greatest interest for measurement, selection of a tracer molecule that is best suited to answer the questions at hand [18], selection of a sample preparation and extraction method best suited to observe the desired metabolites [19], and choice of duration of the labeling period [2, 20].

Many researchers aim to obtain a broad complement of information about the metabolic transformations of a common fuel molecule such as D-glucose or L-glutamine in their cancer cells, which can be achieved by using fully ^{13}C-labeled glucose or glutamine tracers. Doubly labeled $^{13}C_5$-,$^{15}N_2$-glutamine can also be utilized to gain even more information about the transformed products of glutamine; however, use of doubly labeled tracers necessitates the use of advanced analytical methods that can distinguish the two isotope labels, such as NMR and ultrahigh-resolution mass spectrometry, e.g., [21, 22]. Other tracers are particularly useful for evaluating specific pathways such as the pentose phosphate pathway ($^{13}C_2$–1,2-glucose; [18, 23]), fatty acid biosynthesis ($^{13}C_3$-glycerol, $^{13}C_8$-octanoic acid; [12]), and reductive carboxylation of glutamine ($^{13}C_1$–1-glutamine; [24]). In a sophisticated study on the compartmentalization of NAD(P)H-dependent pathways in different cellular compartments, several deuterated (^2H) forms of glucose, serine, and glycine were employed to explore the directionality of the pyridine nucleotide redox cycles [25]. A more comprehensive listing of various isotope tracers and their utility in tracer experiments has recently been published [17].

In addition to evaluating cells in the basal metabolic setting, experiments can be designed to assess perturbations in metabolic pathways caused by addition of a drug, a genetic manipulation, nutrient availability, or altered physiological parameters such as

hypoxia and acidosis. In most cases, such questions can be addressed by aiming for pseudo-steady-state isotopic labeling of the metabolites of interest [2, 17, 20]. If subsequent steady-state metabolic flux modeling analysis is desired, a general rule of thumb of the tracer studies is to allow for cells to be grown in the presence of the tracer molecule for sufficient time for at least one population doubling to occur. This allows for many of the central metabolic pathways of interest to approach pseudo-steady-state isotope enrichment [2, 26]. For example, uridine-diphosphate N-acetyl glucosamine (UDP-GlcNAc), an activated form of N-acetyl glucosamine used for N- and O-linked protein glycosylation, derives carbon from glycolysis, the pentose phosphate pathway, hexosamine and pyrimidine biosynthetic pathways, and the Krebs cycle [26]. In a study utilizing a prostate cancer cell line with a doubling time of ~40 h, it was determined that ^{13}C incorporation in UDP-GlcNAc did not approach steady state until 30 h of growth in the presence of $^{13}C_6$-glucose [26]. However, macromolecules such as proteins and cellular lipids require many population doublings to approach steady-state labeling patterns [20], and this must be taken into account if these macromolecules are evaluated for stable isotope incorporation. In contrast, intermediates of central metabolism such as glycolysis, the pentose phosphate, and most components of the Krebs cycle reach isotopic steady state much more quickly, and for dynamic studies, rather short exposure times can be used [17].

The choice of number of treatment groups and number of experimental replicates per group must be balanced between statistical considerations and practical workflow considerations. In general, a minimum of three replicate samples should be included per treatment group. More biological replicates may be needed to detect more subtle differences in metabolite labeling patterns. Before drawing scientific conclusions, it is always important to repeat the experiment to ensure that the results are independent of cell passage numbers, density, or other subtle differences in growth conditions. It can also be very helpful to include an unlabeled sample for each treatment group, which is grown in the presence of the natural abundance form of the tracer molecule. Analysis of this additional sample allows for qualitative comparisons of the fate of the ^{13}C tracer carbons, e.g., in overlays of ^{1}H-{^{13}C} HSQC NMR spectra [14], and as a technical control to discern incorrectly assigned isotope-labeled peaks during curation of mass spectra. This additional unlabeled sample can also be included as a replicate in measurement of total abundance of metabolites in the treatment group. However, the unlabeled sample should not be used to manually correct for the presence of natural abundance isotopologues of the metabolites of interest in mass spectrometry data [2], as formal algorithms must be applied for such corrections [2, 27].

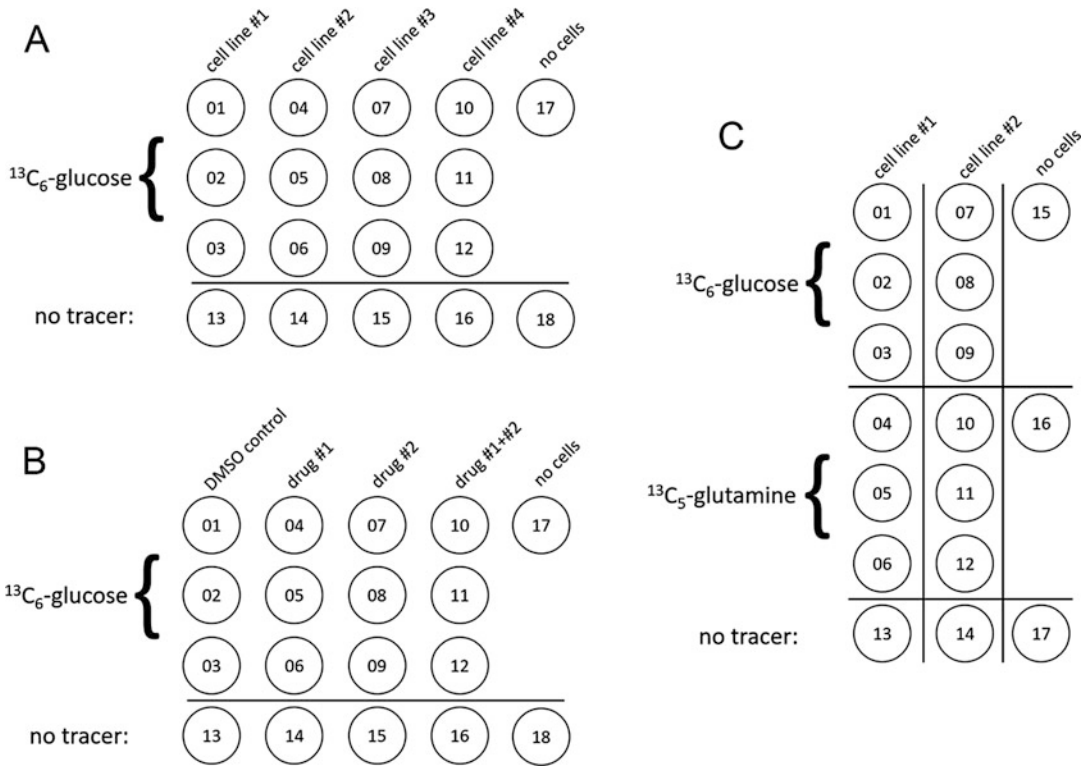


Fig. 1 Examples of experimental setups for stable isotope tracer experiments. (**a**) Sample dish configuration for simultaneous $^{13}C_6$-glucose labeling of four cell lines, including unlabeled replicate plates (#13–16) incubated with ^{12}C-glucose, as well as a blank replicate plate incubated with media but no cells. (**b**) Sample dish configuration for evaluation of metabolic perturbations induced by two drugs separately or in combination. Note that pilot experiments should be performed to determine the IC_{50} of each drug as well as the combination of both drugs, which may act antagonistically or synergistically. (**c**) Sample dish configuration for the comparative labeling of two cell lines with two different tracers, including unlabeled control replicates and the blank plate incubated with media but no cells

A reasonable number of 10 cm dishes with cells to process in 1 day using the extraction method detailed here range from 12 to 18, and the procedure runs much more efficiently with two people working together. More dishes can be processed if desired; however, staggering of the tracer additions may be needed so that the tracer labeling period is consistent across all dishes. Example experimental setups are provided in Fig. 1, which include several configurations for different kinds of experiments, including blank "t = 0" plates with no cells but containing tracer-labeled or unlabeled medium (see Fig. 1).

1.2 Nutrient Requirement, Growth Kinetics, and Tracer Medium Choice

Cultured tumor cells exhibit different nutritional requirements depending on cell type, genotype, tissue of origin, and numerous other factors. For assessing pseudo-steady-state labeling patterns of intracellular metabolites, it is important to design the experiment such that nutrient supplies remain adequate throughout the course

of the experiment. For example, if the culture medium is depleted of $^{13}C_6$-glucose by the cells before the end of the tracing period, the resulting intracellular ^{13}C metabolite labeling patterns will likely reflect a state of altered growth and/or metabolic adaptation to alternative energy sources rather than glucose-driven labeling patterns observed during pseudo-steady-state growth. Similarly, premature depletion of glutamine tracer by cells during a glutamine tracer experiment might result in the diversion of glucose-derived carbon into pathways normally fueled by glutamine in the cultured cells. Significant unintended changes in nutrient levels and the resulting compensatory changes in cellular metabolism will make biological interpretation of metabolite labeling patterns exceedingly difficult.

Population doubling times and metabolic rates can vary greatly between different cell lines even of the same nominal origin (e.g., MDAMB231 cells double in 18 h, whereas ZR75–1 double in 80 h in rich media) [12]. Mouse cells typically double faster than analogous human cells, and the rate of metabolism reflects this difference [28]. In order to minimize the confounding effects of changing nutrient availability during the tracer experiment, preliminary experiments should be conducted in order to assess the growth rate (doubling time) and bulk utilization of the major fuels of cultured cancer cells, i.e., glucose and glutamine. These measurements can be performed using cell counting and relatively inexpensive enzymatic methods, ready-made kits, mass spectrometry, or NMR. We routinely measure the concentrations of glucose, lactate, and glutamine in both freshly prepared culture medium and in spent culture medium supernatants from experiments using immobilized enzyme electrode amperometry (YSI 2950 Bioanalyzer), which allows direct measurement of metabolites in culture medium without the need for extraction/deproteinization. The results of these simple analyses help to determine optimal cell seeding densities for stable isotope tracer experiments and can also aid in the determination of whether initial cell extracts from a stable isotope tracer experiment should be further processed and submitted for NMR and mass spectrometry analyses or discarded due to loss of adequate nutrient supplies. It is important to keep in mind that true steady-state maintenance of the extracellular environment can never be achieved in a standard culture dish [29]. Quantitative metabolite flux experiments require complex control of medium components to maintain steady-state conditions using a chemostat, sampling of multiple time points, and in-depth computational modeling of metabolic network kinetics [2, 30–32].

Another factor to take into consideration is whether a genetic manipulation or drug treatment might result in reduced cell growth rate. For instance, treatment with a cytostatic drug will result in reduced cell numbers at the end of the experiment. If this is the case, cell dishes in the “treated” group may need to be

seeded at a higher density than the control group in order to obtain sufficient cell biomass for metabolite extraction at the end of the experiment. Careful consideration of the effects of decreased cellular proliferation on isotope labeling patterns will need to be included during subsequent data analysis [2].

Although the majority of routine mammalian tumor cell culture work is performed using either DMEM or RPMI 1640 culture medium with the desired tracer nutrient replaced by the corresponding stable isotope-enriched counterpart at the identical concentration, recent studies have underscored the influence of the nonphysiological nutrient concentrations and lack of certain metabolites in these culture media on cellular metabolism in culture [33]. In addition to containing nonphysiological concentrations of vitamins and amino acids [28], culture media including DMEM and RPMI 1640 also lack metabolites that are found in human plasma [34], which have been shown to influence cell metabolism and drug susceptibility [33]. These findings led to the development of a new medium formulation, human plasma-like medium (HPLM), which better recapitulates the in vivo human setting [33]. However, preparation of customized media is both tedious and costly, and to date HPLM is not commercially available. In most cases, cancer researchers aim to incorporate the data obtained from SIRM experiments into a larger dataset encompassing other biological outcomes such as gene and protein expression, drug susceptibility studies, analysis of invasion potential, protein subcellular localization studies, gene knockdown or overexpression, functional analysis of gene mutations, hypoxic response, respiration studies, etc. Thus, if all possible, the stable isotope tracer experiment should be performed under identical culture conditions as those employed for other biological experiments.

1.3 Dialyzed FBS vs Non-dialyzed FBS

Dialyzed FBS is a preferred additive for stable isotope tracer experiments with mammalian cells because it retains most of the growth factors and lipids present in FBS but lacks the small molecule metabolites present in non-dialyzed serum. Depending on the stable isotope tracer and associated pathways being evaluated, the natural abundance metabolites present in non-dialyzed FBS could confound the interpretation of both extracellular and intracellular metabolite labeling patterns. Table 1 lists natural abundance concentrations of some commonly evaluated metabolites found in two separate lots of FBS obtained from two different commercial sources, measured amperometrically using a YSI 2950 Bioanalyzer. When used at the standard supplementation rate of 10%, these non-dialyzed FBS reagents would contribute ~0.4–0.8 mM natural abundance glucose and ~0.8–1.8 mM natural abundance lactate to the tracer medium. The presence of these unlabeled substrates and products of glycolysis would need to be considered in the interpretation of pseudo-steady-state labeling patterns of both intracellular

Table 1
Total metabolite concentrations measured amperometrically (YSI 2950 Bioanalyzer) in non-dialyzed and dialyzed FBS samples as well as laboratory water sources

	Standard heat-inactivated FBS				Dialyzed FBS		Water sources	
	Mfg #1, batch #1	Mfg #1, batch #2	Mfg #2, batch #1	Mfg #2, batch #2	Mfg #1, batch #1	Mfg #1, batch #2	Tap H_2O, Bethesda, MD, USA	ddH2O
D-glucose, mM	7.95	4.74	3.82	2.75	0.05	0.08	0.04	0.00
L-lactate, mM	18.2	8.25	14.44	17.9	0.17	0.18	0.03	0.00
L-glutamine, mM	0.38	0.1	0.65	0.47	0.01	0.01	0.00	0.00
L-glutamate, mM	0.87	0.57	0.67	0.75	0.00	0.00	0.02	0.03
K+, mM	10.7	n/a	n/a	10.5	n/a	n/a	0.2	0.0
NH4+, mM	2.9	n/a	n/a	2.2	n/a	n/a	0.01	0.0

and extracellular metabolites. This issue would be avoided with dialyzed FBS, as the concentrations of natural abundance glucose and lactate in two separate lots were 50–100-fold lower than the non-dialyzed FBS (Table 1). Also notable is the persistence of small amounts of carbohydrate and lactate in the municipal water supply, underscoring the need to utilize double-deionized water for experiments and to avoid the contamination of dishes and sample tubes by stray pieces of ice from the ice tray.

In our own experience with tumor cell lines that are deficient in Krebs cycle enzymes and/or respiratory chain components, we have found that some cell lines exhibit attenuated growth in dialyzed FBS, an effect that is likely due to auxotrophy for one or more small molecule metabolites present in FBS such as uridine, uric acid, etc. Many cell lines require an adaptation period of several days in the presence of dialyzed FBS, presumably to allow for transcriptional changes and metabolic adaptation, before the growth rate is restored to suitable levels. For these reasons, it is important to perform pilot experiments to test the growth characteristics of the cell line of interest in dialyzed FBS. It may also be important to test whether gene expression in the pathway(s) of interest is modulated by growth of the cells in dialyzed FBS, as metabolic compensation might occur in the absence of small molecule nutrients supplied in non-dialyzed FBS. As a rule of thumb, the same batch of dialyzed and non-dialyzed FBS should be utilized for all experiments in a project to maintain consistency of experimental results [28]. Finally, in tracer studies where extracellular vesicles such as exosomes are to be isolated or their influence on cell metabolism studied, exosome-free FBS (either non-dialyzed or dialyzed) should be adopted. Exosome-free FBS can be prepared using ultra-centrifugation or purchased commercially.

1.4 Other Considerations of Tissue Culture Experiments

Other practical factors are also important to consider in minimizing variations in tissue culture conditions. Adequate humidity in the incubator should be maintained by ensuring that ample water is present in the incubator water tray or by acquiring an incubator that has an active humidification system. If possible, the stable isotope tracer experiment should be conducted in a dedicated incubator that is free of vibrations and will not be opened and closed excessively during the experiment. Excessive opening and closing of the incubator door leads to transient decreases in CO_2 concentration, resulting in fluctuations in culture medium pH due to changes in bicarbonate concentration and drying of the internal environment due to injection of dry CO_2 gas to restore CO_2 levels. Fluctuation in humidity is especially pertinent in tri-gas incubators operating at O_2 concentrations that are far below the ambient (~21% O_2), as large quantities of dry nitrogen gas are injected into the incubator to decrease the O_2 concentration. Table 2 shows the results of an experiment in which triplicate 10 cm dishes without cells were filled with 15 mL of DMEM/10% non-dialyzed FBS medium and allowed to incubate at 37 °C unperturbed in a 95% room air/5% CO_2 atmosphere, in either a well-humidified incubator or in an incubator in which the water tray was allowed to become dry. Plates incubated in the non-humidified incubator for 40 h showed an approx. 5% decrease in total mass, indicative of significant culture medium evaporation. In contrast, the dishes incubated in the humidified incubator lost only 1.3% of their mass over the same incubation period. The evaporative losses in the inadequately humidified incubator showed significant effects on the observed concentrations of glucose, lactate, and glutamine in the spent culture medium samples. Namely, glucose and lactate concentrations increased by 16–17%, while glutamine concentrations decreased by 1.3% in samples incubated in the non-humidified

Table 2
Evaporation and changes in metabolite concentrations in culture dishes incubated in humidified versus non-humidified incubators

	$t = 0$	Humidified	Non-humidified
Plate mass, % ctl	100%	98.7%	95.1%
[glucose], mM	11.5	12.1	13.4
[glucose], % $t = 0$	100%	104.9%	115.8%
[lactate], mM	1.76	1.86	2.06
[lactate], % $t = 0$	100%	106.0%	117.0%
[glutamine], mM	2.94	2.67	2.90
[glutamine], % $t = 0$	100%	91.0%	98.7%

incubator. In contrast, glutamine concentration decreased significantly by 9% in the well-humidified incubator, likely reflecting the conversion of glutamine to NH_4^+ and pyroglutamate [35], an effect that was masked in the non-humidified incubator due to evaporative loss of water in the dish. Differences in medium metabolite fluxes in non-humidified versus humidified growth conditions can also be attributed to altered cell physiology (e.g., reduced growth rate; Fan, unpublished data) and are likely to involve increases in culture medium osmolality. Thus, a poorly controlled incubator environment can considerably influence the interpretation of apparent metabolite uptake and secretion rates (Table 2).

Finally, other general good practices in tissue culture are of the utmost importance in metabolomics experiments, in order to justify the significant cost, time, and effort involved in acquisition, analysis, and interpretation of metabolomics data. Cell lines should be regularly tested for mycoplasma, as this cryptic and antibiotic-resistant contaminant microbe can go unnoticed indefinitely in tissue cultures [28], and by nature of its physical attachment to cells in tissue culture, mycoplasma is very likely to significantly alter cellular metabolism [36]. Unless specifically desired for differentiation or other physiological strategies, cells should be seeded evenly across the dishes and should not be allowed to become 100% confluent before quenching and extraction, as some cell lines undergo contact inhibition and thus will not be in log-phase growth at the time of harvest. In some cases, cells may also show dramatic differences in intracellular metabolite levels depending on their state of confluence [37]. Cell size can vary greatly across different cell lines; thus, initial cell seeding numbers and the number of cells present at 100% confluence can be very different. In general, it is a good practice to perform initial experiments to determine the number of cells required for seeding of culture dishes such that the cells are at 70–90% confluence at the time of harvest.

1.5 Sample Preparation for Stable Isotope-Resolved Metabolomics

Numerous protocols exist for metabolite quenching/extraction from mammalian cells [3, 19, 25, 38, 39]. There is no single sample preparation technique that can quantitatively recover all intracellular metabolites, as some metabolites are inherently labile and prone to degradation, and some others are less extractable using a universal solvent system. Notable examples of metabolites that can be degraded during extraction procedures include NADPH and to a lesser extent NADH, both of which show decreased stability in acidic conditions [40]. Glutamine is inherently unstable in culture medium [35], and metabolite extraction procedures can cause marked conversion of glutamine to pyroglutamate [41]. Issues with glutamine degradation can also arise as a direct result of ionization in the mass spectrometer [42]. Direct, in-cell derivatization of labile metabolites has been shown to be useful for detection of unstable carbonyl- and aldehyde-containing metabolites

including oxaloacetate, dihydroxyacetone phosphate, glyceraldehyde 3-phosphate, and acetoacetate [21, 43]. Derivatization has also been shown to be useful for resolving multiple labeled amino acids by direct infusion and ultrahigh-resolution mass spectrometry [22].

Careful studies in *E. coli* have demonstrated that significant metabolite decomposition can occur as an artifact of the extraction procedure [44] and that acetonitrile is a superior solvent for metabolite extraction in *E. coli*, especially in the case of the preservation of nucleoside triphosphates [45]. Such improvements in extraction techniques were employed in the quantitative evaluation of total metabolite concentrations and metabolite enzyme occupancy in *E. coli* [46]. In mammalian cells, acetonitrile-water was also found to be a superior polar metabolite extraction solvent [19]. When acetonitrile-water extraction is combined with a second chloroform-methanol extraction, this single procedure can yield extracts suitable for analysis of polar metabolites, total cellular lipids, and cellular proteins all from one sample, e.g., [7, 14]. This metabolite extraction procedure, termed the Fan extraction method by the first author, is outlined below in Subheading 3.2.

2 Materials

1. Laminar flow biosafety cabinet for tissue culture work.
2. Dual-gas (CO_2) or tri-gas (CO_2, N_2) humidified, temperature-regulated incubator.
3. 37 °C heated water bath
4. 10 cm tissue culture treated culture dishes (*see* **Note 1**)
5. Electric pipettor and serological pipettes (5, 10, 25 mL).
6. Vacuum aspirator fitted with tubing and disposable 2 mL plastic aspirator pipettes.
7. 15 mL polypropylene conical vials (Sarstedt)
8. 50 mL polypropylene conical vials (Sarstedt)
9. 1.5 mL polypropylene Eppendorf tubes
10. 2.0 mL polypropylene screw-top vials
11. 0.5 mL polypropylene screw-top vials
12. 2.0 mL screw-top glass vials
13. Basal DMEM: DMEM without glucose, glutamine, pyruvate, and phenol red (*see* **Note 2**).
14. Dialyzed FBS (*see* **Note 3**) [or exosome-depleted FBS (e.g., Gibco# A2720801) if you are studying exosomes].
15. 0.5% phenol red solution (*see* **Note 4**)

16. Stable isotope tracer (e.g., $^{13}C_6$-D-glucose, $^{13}C_5$-L-glutamine, $^{13}C_5$,$^{15}N_2$-L-glutamine, etc.) (*see* **Notes 5** and **6**).
17. DMEM tracer media (*see* **Notes 7** and **8**).
18. HPLC-grade acetonitrile (*see* **Note 9**).
19. HPLC-grade methanol (*see* **Note 9**).
20. HPLC-grade chloroform (*see* **Note 9**).
21. 2:1 chloroform/methanol-BHT solution (*see* **Note 9**)
22. Disposable plastic cell lifters (corning costar #3008) (*see* **Note 10**).
23. Ice buckets.
24. Rectangular anodized metal plate of at least 14 × 28 cm (e.g., Biocision, Inc. #BCS-123).
25. Phosphate-buffered saline (PBS), pH 7.4: 1.06 mM KH_2PO_4, 2.97 mM Na_2HPO_4, 155 mM NaCl.
26. Fine-tipped transfer pipettes (e.g., Samco scientific #235).
27. Freezer lyophilizer equipped for handling organic solvents (*see* **Note 11**).
28. Vacuum centrifuge (Eppendorf Vacufuge plus, fitted with rotors for both 1.5 mL tubes and for 2.0 mL glass vials) (*see* **Note 12**).

3 Methods

3.1 Metabolic Labeling of Adherent-Cultured Mammalian Cells

1. Perform preliminary experiments using tracer medium prepared with the desired natural abundance ^{12}C tracer counterpart to ensure that the tracer medium composition is suitable to support cell growth and that nutrient levels remain adequate throughout the duration of the experiment. Considerations include (1) assessment of adequate cell growth rate in tracer medium, (2) determination of the target cell seeding density such that cells will reach 70–90% confluence at the time of quenching/harvest, (3) selection of a tracer incubation time that equals or exceeds the cell population doubling time in order to approach pseudo-steady-state labeling of many intracellular metabolites, and (4) determination of bulk metabolite uptake and secretion at the end of the labeling period to ensure that the major nutrients (e.g., glucose and glutamine) are not depleted at the end of the labeling period.
2. Seed cells in 10 cm dishes in 10–25 mL tracer medium (*see* **Note 13**) containing natural abundance tracer compound (i.e., ^{12}C-glucose and ^{12}C-L-glutamine) 1 or 2 days prior to the labeling experiment, and allow them to attach and proliferate

(*see* **Notes 14** and **15**) (*see* **Note 19 for alternative protocol for non-adherent cells**).

3. Prepare tracer media freshly and equilibrate to 37 °C several hours prior to addition to the culture dishes. Also, prewarm a sufficient quantity of basal DMEM (without FBS or any other added compounds) to allow for washing each plate once with 10 mL basal DMEM (*see* **Note 16**).
4. Remove the culture dishes from the incubator, and aspirate the ^{12}C tracer medium using vacuum aspiration. Gently add 10 mL of warm DMEM basal medium, rock the plate by the hand several times, and then aspirate. Finally, add 10–25 mL of prewarmed ^{13}C tracer medium to each plate, remove 0.2 mL of medium ($T = 0$), and weigh plate on a three-place balance before returning plates to the incubator. For each type of tracer medium used, include one blank plate in which no cells are seeded but tracer medium is added. This plate is incubated alongside the plates with cells and serves as the blank plate (*see* Fig. 1**,** Subheading 1.2**,** and **Notes 17** and **18**).
5. On the next day, set up the lab bench for medium harvest and cell quenching by obtaining several trays of ice and pre-labeling all of the tubes that will be needed. In one ice tray, place a slab of rectangular metal of at least 14 × 28 cm (e.g., Biocision, Inc. #BCS-123) horizontally, tilted toward the front of the bench by approximately 35°. Precool a bottle of HPLC-grade acetonitrile to −20 °C.
6. At the end of the incubation period, weigh plates using the same balance, and remove the plates from the incubator in groups of three at a time. Place the dishes on the metal slab, with the pen mark on the side of the dish facing you. Transfer 1 mL of medium from each plate to a snap-cap microcentrifuge tube, and set the tube aside on ice for subsequent processing and extraction. Aspirate the remainder of the tracer medium using a fine-tipped transfer pipette into a 15 mL conical centrifuge tube if the medium is to be harvested or using the vacuum aspirator fitted with a disposable 2 mL plastic aspirator pipette if the medium is to be discarded (**for non-adherent cells,** *see* **Note 19**).
7. Wash the dishes three times with 10 mL ice-cold PBS by gently adding the PBS to the side of the dish where the mark is and aspirating from the dish wall slightly above the same spot each time to minimize cell loss. Make sure that the entire circumference of the dish wall and the entire surface of the plate are rinsed by each of the three washes by gentle rocking of the dish during each wash. After the third wash, allow the dish to remain tilted toward you at ~35° on the metal block with the pen mark facing toward you. Carefully observe the walls of each

dish, and aspirate any residual liquid that might remain on the walls of the circumference of the plate while residual liquid on the plate is collecting near the pen mark. Finally, aspirate again any additional residual fluid from the bottom of the plate (*see* **Notes 20 and 21**).

8. Immediately add 1 mL cold (−20 °C) acetonitrile to quench cell metabolism. Ensure that the solvent covers the entire surface of the dish by gently but rapidly shaking the plate horizontally. Place the culture dishes on a shelf in a −20 °C freezer, and incubate for 5–20 min.
9. Remove the dishes from the freezer, and return to the ice-cold metal block. Add 0.75 mL ice-cold ddH_2O, and gently rock the dish to ensure mixing of the water and acetonitrile (*see* **Note 22**).
10. With the plate angled at ~35° toward you resting on the ice-cold metal, scrape the cells and solution toward the bottom of the dish using a cell lifter with firm circular motions until all material has been scraped to the lower end of the dish. Next, rotate the dish by 90° and scrape the remaining cells and fluid to the bottom of the dish (*see* **Note 23**).
11. Transfer the quenched cells and solvent to a 15 mL conical centrifuge tube, and then repeat **steps 8** and **9** but omit the −20 °C incubation step. The resulting volume in the 15 mL conical vial is typically ≥ 3.2 mL. The final solvent composition is 2:1.5 (acetonitrile-water). This solution can be kept on ice for several hours before proceeding to chloroform extraction or stored at −80 °C for up to 1 week pending chloroform extraction.
12. After all dishes of cells have been quenched and scraped, centrifuge the 1 mL culture medium samples at 3000 × *g* for 10 min at 4 °C in a microcentrifuge to remove any floating cells and large debris.
13. Transfer 900 μL of the resulting culture medium supernatant to a new pre-labeled tube for storage at −80 °C until metabolite extraction. A 100 μL of this culture medium sample can be stored separately for direct acetone extraction. Culture medium metabolite extraction procedures are outlined in Subheading 3.3.

3.2 Cell Metabolite Extraction Using the Fan Extraction Method

1. Pre-label all tubes to be used during the phase separation and extraction:
   (a) Aqueous/polar fraction: 5 mL snap-cap Eppendorf tube, pre-weighed to 0.1 mg precision. Label simply and clearly with sample number; it will be discarded after aliquoting of the polar phase.

(b) Nonpolar/lipid fraction: 2 mL screw-cap glass vial with Teflon-faced liner, wrap with clear scotch tape around the side of the vial to protect labeling.

(c) Protein fraction: 1.5 mL snap-cap Eppendorf tube (pre-weighed to 0.01 mg precision if protein dry mass is desired), taped around the side to protect labeling (*see* **Note 24**).

2. Add 1 mL ice-cold chloroform to the acetonitrile-water cell extracts in their 15 mL conical tubes, using a pre-wetted 1 mL polypropylene micropipette tip. Seal the tubes tightly, and shake the extracts vigorously up and down at least 60 times manually or on a vortexer (*see* **Note 25**).
3. Centrifuge the extracts in a refrigerated swinging-bucket centrifuge at maximum speed (typically 3000–4650 × *g*, depending on the model of the centrifuge), 4 °C, 30 min. The extracts will separate into three phases: An upper aqueous phase containing polar metabolites, a lower nonpolar chloroform phase containing lipids, and a middle phase containing denatured proteins and other macromolecules. Gently transfer the tubes to an ice bucket, taking care not to disrupt the phase separation (*see* **Note 26**).
4. Transfer ~90% of the aqueous phase to pre-tared (±0.1 mg) 5 mL Eppendorf tube using a fine-tipped transfer pipette, taking care not to disrupt or transfer any of the denatured protein layer along with the aqueous phase. The remaining aqueous phase will be recovered during the second extraction. Save the transfer pipette for the second extraction.
5. Transfer ~90% of the lower chloroform phase to a 2 mL glass vial until the vial is filled to the neck, using a gel-loading 200 µL pipette tip; save the tip for a second transfer (*see* **Note 27**).
6. Place the 2 mL glass vial in a vacuum centrifuge, and reduce the volume of the chloroform fraction by vacuum centrifugation for 30 min without heating.
7. Centrifuge the 15 mL conical vial containing the remaining polar and nonpolar fractions as well as the protein fraction for 5 min at maximum speed at 4 °C. if more than ~400 µL of nonpolar fraction remains in the 15 mL vial, transfer more nonpolar fraction to the respective 2 mL glass vial after the volume has been reduced sufficiently by vacuum centrifugation, until <100 µL of nonpolar fraction remains.
8. Using a 1 mL pipettor set to 500 µL, mix the protein fraction along with the polar and nonpolar fractions until the protein fraction is sufficiently dispersed to allow for the transfer of the remaining contents of the 15 mL tube to the 1.5 mL protein fraction tube. Set aside this pipettor with the tip still attached.

Using a separate 1 mL micropipette, add 350–500 μL ice-cold chloroform-methanol/1 mM BHT to the 15 mL vial. Then use the first pipettor and tip to pipette the solution up and down, washing the residual protein residue from both the walls of the tube and the inner surfaces of the pipette tip. Transfer the protein and chloroform-methanol/BHT wash to the protein fraction tube (*see* **Note 28**).

9. Vigorously shake and vortex the protein fraction tube for >1 min and then centrifuge the tube at maximum speed in a refrigerated microcentrifuge for 20 min at 4 °C, with the hinges of the tubes facing away from the center of the rotor.
10. Transfer as much of the upper aqueous layer in the protein fraction tube as possible to the 5 mL polar phase Eppendorf tube using the transfer pipette from **step 4** while avoiding disruption or transfer of the middle protein layer. Weigh the resulting polar fraction, and subtract the initial tube weight to determine the total weight of the polar metabolite fraction (*see* **Note 29**).
11. Transfer all but ~10–20 μL of the lower nonpolar lipid fraction to the 2 mL glass lipids vial, and return the lipid vials to the vacuum centrifuge, and evaporate to dryness. As soon as the lipid fractions are dry, add 300 μL of chloroform-methanol/BHT solution to them, cap the glass tube tightly, vortex gently, and store at −80 °C until analysis (*see* **Note 30**).
12. Add 500 μL of cold methanol to the protein fraction, vortex briefly, and then centrifuge the tube at maximum speed in a refrigerated microcentrifuge for 10 min at 4 °C with the hinges of the tubes facing away from the center of the rotor. Carefully aspirate the supernatant using a gel-loading tip via vacuum suction or a 1 mL pipette, taking care not to disrupt the protein pellet (*see* **Note 31**).
13. Evaporate the protein residues to dryness by vacuum centrifugation (~25–30 min). Subsequently, on the same day, measure and record the mass of the protein fraction tube with the dried protein residue to 0.01 mg and subtract the tare weight of the same tube to obtain the mass of the dry protein residue, if desired. The expected residue mass is 0.5–2 mg (*see* **Note 32**).
14. The polar fraction can be sub-aliquoted into many fractions and stored in polypropylene screw cap or Eppendorf tubes for different analyses. Half or more of the polar fraction in the 5 mL tube is suitable for analysis by high-field NMR spectroscopy [14, 15]. One eighth of the polar fraction in a 2 mL glass vial is suitable for MTBSTFA derivatization for analysis by GC-MS [14, 16] or ion chromatography (IC)-FTMS analysis [7, 14, 47]. One 16th of the polar fraction in a 2 mL screw-cap

tube is suitable for multiple direct infusion FT-MS analyses [16, 22, 48] (*see* **Note 33**).

15. Cap all aliquot tubes with screw caps or snap caps containing 3–6 small perforations on top, and freeze all tubes upright in flexible 9 × 9 cardboard racks by placing them in ~1 cm of liquid nitrogen in a Styrofoam or other cold-resistant tray. After all aliquots reach liquid nitrogen temperature, rapidly transfer the racks to a suitably sized lyophilization can and apply vacuum as quickly as possible. Lyophilize to dryness for 18–24 h; the polar aliquots should remain frozen under the vacuum for the duration of the lyophilization period (*see* **Note 11**). After lyophilization, replace the perforated caps with tightly sealed pre-labeled caps and store at −80 °C until analysis.

3.3 Acetone Extraction of Culture Medium

Proteins in culture medium are chemically precipitated using cold 80% acetone and removed by centrifugation, and the supernatant is collected for downstream analyses.

1. Add 400 μL of ice-cold acetone to a 100 μL aliquot of culture medium in a microcentrifuge tube.
2. Vortex vigorously and then incubate the samples in a −80 °C freezer for >30 min.
3. Remove the precipitated samples from the freezer, and centrifuge them at maximum speed (15–20 K × *g*) for 10 min at 4 °C.
4. Transfer the clarified supernatant to a new tube(s), and remove the solvent using either a vacuum microcentrifuge, or freeze-dry using a lyophilizer fitted with a liquid nitrogen pre-trap (*see* **Notes 34 and 35**).

3.4 Culture Medium Protein Removal Using Ultrafiltration

Proteins are removed by centrifugal ultrafiltration using 10,000 molecular-weight-cutoff dialysis membranes, and the flow-through is collected for downstream analyses.

1. 300–400 μL of culture medium is added to the top of an ultrafiltration spin column (e.g., Millipore Ultracel YM-10) after extensive wash with ddH_2O.
2. Centrifuge the spin column at 5000 × *g* for 30 min at 4 °C in a microcentrifuge.
3. Collect 100 μL of the filtrate flowthrough, and remove the solvent using either a vacuum microcentrifuge, or freeze-dry using a lyophilizer fitted with a liquid nitrogen cold trap (*see* **Notes 35 and 36**).

4 Notes

1. Poly-D-lysine or collagen-coated plates can be used for cells that are poorly adherent.
2. Some cell lines, including mtDNA-deficient rho-0 cells and patient-derived fumarate hydratase-deficient ($FH^{-/-}$) tumor cell lines, grow poorly or not at all in culture medium made with dialyzed FBS without supplementation of 1 mM pyruvate. The necessity for the presence of pyruvate in the culture medium should be evaluated prior to the tracer experiment; we typically do not supplement with exogenous pyruvate unless it is necessary for cell proliferation.
3. The part number given here is for pre-dialyzed FBS; however, dialyzed FBS can be prepared in-house with substantial savings in cost.
4. While not required for tissue culture, phenol red is an invaluable tool that allows the researcher to continuously monitor the pH of the culture medium. An excessively pink color indicates greater alkalinity, usually due to loss of bicarbonate buffering by prolonged exposure of the medium bottle to room air atmosphere. Yellowing of the culture medium is frequent in highly glycolytic tumor cell cultures and is indicative of substantial acidification due to lactate secretion and possible depletion of glucose in the culture medium. Yellowing of the culture medium can also be an early indicator of bacterial contamination. Finally, changes in the color of phenol red sometimes occur after additions of drugs or other chemicals, immediately indicating the need for pH adjustment of the agent before addition to the culture medium. Caution must be taken with cells that are estrogen-sensitive, as phenol red has been reported to be a weak estrogen mimetic [49], though this may be due to lipophilic impurities in certain phenol red preparations rather than the pH indicator itself [50].
5. The choice of the tracer to be used depends on the goals of the project. Some strategies for tracer selection in mammalian cell tracer experiments are presented in the following references [3, 12, 18, 22, 25].
6. Prepare all tracer solutions using double-deionized water (ddH_2O, 18.2 M Ω). Sterile water for injection can be used if ddH_2O water is not available.
   (a) Prepare 46.5% (w/v) $^{13}C_6$-glucose (2.498 M) volumetrically using a volumetric flask. For example, for a 10 mL solution, weigh 4.65 g of $^{13}C_6$-glucose powder into the flask, and bring up to 10 mL volume line with ddH_2O. 45% (w/v) ^{12}C-glucose solution (2.498 M) is

prepared in the same way using 4.5 g of natural abundance D-glucose. The solutions should be sterile filtered using a 0.22 μm syringe filter and stored in a Parafilm-wrapped vial at 4 °C for several months or in aliquots in tightly sealed vials at −80 °C for years.

(b) Prepare 200 mM L-glutamine solutions using the appropriate molecular weight of the tracer compound in ddH_2O (e.g., ^{12}C-L-glutamine MW = 146.14 g/Mol, $^{13}C_5$-L-glutamine MW = 151.11 g/Mol). The solution should be sterile filtered using a 0.22 μm syringe filter and immediately aliquoted in vials and stored at −80 °C for at least 1 year. Repeated freeze and thawing of the stock solution should be avoided.

(c) Dialyzed FBS can be stored in aliquots of 10–45 mL in sealed polypropylene conical vials at −20 °C or −80 °C.

7. For a typical tracer experiment such as the schemes shown in Fig. 1, prepare 300 mL basal tracer medium by combining 30 mL dialyzed (or exosome-depleted) FBS with 270 mL basal DMEM and optional 300 μL of 0.5% phenol red solution. Sterile filter using a disposable 500 mL 0.2 μm filter unit. Dialyzed FBS tends to show flocculence following freeze-thaw cycles; sterile filtration of the final solution ensures that any particulate matter is removed from the tracer medium. Prepare tracer media as required by the experimental design. We typically supplement basal DMEM media with 11 mM D-glucose (0.2%) and 2 mM L-glutamine (e.g., 11 mM $^{13}C_6$-D-glucose plus 2 mM ^{12}C-L-glutamine or 11 mM ^{12}C-D-glucose plus 2 mM $^{13}C_5$-L-glutamine).

8. Unless it is explicitly part of your downstream analysis strategy (e.g., [22]), never include more than one tracer molecule in a given tracer medium preparation. Also, doubly labeled tracers such as $^{13}C_5$- and $^{15}N_2$-L-glutamine should not be utilized unless downstream analyses will include platforms that are able to distinguish the two heavy isotopes (e.g., $^{1}H/^{13}C/^{15}N$ NMR or ultrahigh-resolution Orbitrap or ion cyclotron Fourier-transform mass spectrometry). Be sure to prepare ~3–5 mL excess of each type of tracer medium for sufficient distribution into each plate, and remember to include enough medium for the blank plates with no cells.

9. Store all solvents in a flammable-certified −20° freezer in the dark.

(a) Decant 100 mL of HPLC-grade acetonitrile into an amber bottle and store at −20 °C.

(b) Decant 100 mL of HPLC-grade methanol into an amber bottle and store at −20 °C.

(c) Decant 100 mL of HPLC-grade chloroform into an amber bottle and store at −20 °C.

(d) Prepare 100 mL of chloroform/methanol-BHT solution, prepare a 100 mM butylated hydroxytoluene (BHT) stock solution in HPLC-grade methanol, and then prepare 100 mL of 2:1 chloroform/methanol-1 mM BHT in an amber bottle by combining 66 mL chloroform with 32 mL methanol and 1 mL of 100 mM BHT stock solution in an amber bottle. Seal tightly, and store at −20 °C.

10. The cell lifter is superior to hinged cell scrapers because more force can be applied at a 45° angle to scrape the solvent and quenched cells.

11. This metabolite extraction procedure calls for freeze-drying/lyophilization of the resulting polar fraction aliquots. Freeze-drying allows for the removal of solvents and water from the polar fraction while maintaining the sample at below freezing temperature and minimizing exposure of the sample to oxygen during the drying process by the maintenance of high vacuum [51] and is far superior to vacuum centrifugation for the quantitative preservation of the maximum number of metabolites. However, depending on the stability of the metabolite(s) of interest, other evaporative techniques could be utilized for sample concentration, including vacuum centrifugation or drying under a flow of nitrogen gas.

    The liquid nitrogen pre-trap is important for lyophilizers fitted with oil-containing pumps, as the solvents with very low melting points (e.g., methanol, acetone) derived from the sample will degrade the pump oil over time and lead to premature failure of the pump.

    An alternative currently being evaluated by D.R.C. is to utilize a − 86 °C lyophilizer with a Teflon-coated condenser chamber and coils, attached to an oil-free Edwards scroll pump with silencer kit, with the open ballast venting into a chemical fume hood.

12. A vacuum centrifuge fitted with an oil-free diaphragm pump (e.g., Eppendorf Vacufuge plus™) is desirable for the evaporation of solvents such as methanol and chloroform, as oil-containing vacuum pumps are subject to the same solvent issues as discussed above.

13. The standard quantity of 10 mL of tracer medium per 10 cm dish is used for routine experiments. However, for very rapidly growing cells such as HEK293 or cells exhibiting extraordinarily high levels of aerobic glycolysis (e.g., mtDNA-depleted cells or Krebs cycle enzyme-deficient tumor cells), greater amounts of medium may be needed to sustain adequate

nutrient levels during routine culture and during the tracer labeling period.

14. The metabolite extraction procedure outlined in this protocol allows for obtaining a protein residue for subsequent quantification and normalization of metabolite levels. If you choose a different metabolite extraction protocol that does not allow for recovery of cellular protein, then additional dishes can be included to allow for cell counting and/or cellular protein determination for inferred normalization of metabolite concentrations [3]. However, this practice is imprecise, as each plate will be seeded somewhat differently and should not formally be used for normalization of a separate plate.
15. Many cell lines require 24 h or more to adhere, stabilize, and enter into log-phase growth. Do not seed the cells, and initiate the tracer labeling period simultaneously.
16. If the medium is excessively pink (alkaline) due to loss of bicarbonate to the air as CO_2, the media can be left in the 5% CO_2 incubator for several hours to overnight with the lid loosened to allow for re-equilibration of the bicarbonate buffering system.
17. At the beginning of the experiment, use a permanent marker to mark the lower outer wall of the culture dish indicating where aspiration and additions will occur throughout the subsequent steps of culture and quenching. This, together with aspiration slightly above the mark, will help to minimize cell loss and keep the loss more consistent across the dishes and treatment groups during the course of experiment.
18. If an inverted phase-contrast microscope with a camera is available, then take a picture of one or more regions of the dish from at least one replicate of each treatment group before and after the tracer incubation period. This will reveal any morphological or cell number changes in the tracer experiment.
19. For non-adherent cells or cells that are very poorly adherent, an alternative harvesting procedure is as follows:
    (a) Collect cells and medium from each dish in a conical tube, and centrifuge at 350 × *g*, 5 min, 4 °C. Weakly adherent cells can be collected by gentle trituration of the tracer medium over the plate until cells are detached.
    (b) Collect 1 mL of culture medium supernatant, and set aside for subsequent processing and extraction.
    (c) Vacuum aspirate the remainder of the tracer medium with a fine pipette tip, add 10 m L of ice-cold PBS, and gently triturate the cell pellet. Centrifuge at 350 × *g*, 5 min, 4 °C, and then vacuum aspirate the supernatant.

(d) Repeat **step 3** once, taking care to vacuum aspirate as much PBS as possible during the second aspiration.

(e) Immediately add 1 mL cold (−20 °C) acetonitrile, and vortex the mixture in pulses to resuspend the cells. If clumps of cells are still visible, then triturate up and down with the 1 mL pipette tip to break up the clumps.

(f) Add 0.75 mL ice-cold ddH_2O and vortex, then add an additional 1 mL of acetonitrile and 0.75 mL ddH_2O, and proceed to **step 11** in Subheading 3.1.

20. Work quickly, and minimize the time between which the culture dishes leave the incubator and the cells are quenched with solvent, thus minimizing the time that elapses between removal of the dish from the relative equilibrium of the incubator to the time when cellular enzymatic activities are eliminated by the 100% acetonitrile quench.

21. Prior to metabolite quenching and extraction, it is important to minimize the salt contribution to the final extracts, which can interfere with metabolomics analysis, particularly by mass spectrometry. Residual sodium, chloride, and phosphate ions can interfere with MS analyses via ion suppression and decreased signal-noise ratios in NMR analyses [52]. Be sure to aspirate as much liquid as possible after the last wash.

    Greatly reduced salt interference can be achieved by a brief rinse of the dish with Nanopure water (30 s to 1 min) right after removing the medium components with PBS. We have found that this water rinse significantly improved Fourier-transform mass spectrometry performance without loss of cellular metabolites. The water rinse is aspirated rapidly, and the quenching solvent is added immediately.

22. An internal standard can be included in the ddH20 at this step. 53.3 μM tris, pH 8.0 can be included in the first 0.75 mL ddH20 addition, as a convenient internal standard in ^{1}H NMR analyses.

23. Think of the cell lifter as a squeegee being used to move all liquid and cell material to the bottom of the plate, reminiscent of squeegeeing a window or windshield after washing it.

24. High-quality gloss-finish scotch tape is used to protect the writing on the tube from being smeared by solvents during pipetting, centrifugal evaporation, and lyophilization.

25. This is the commitment step in the metabolite extraction procedure. Once chloroform is added, the fractionation procedure must be completed in the same day, as the chloroform will slowly degrade the polypropylene tubes. Never use polystyrene tubes for metabolite extractions.

26. If the protein interphase looks excessively thick, fuzzy, or uneven, then centrifuge for an additional 20 min.
27. If the protein interphase layer is sufficiently abundant and forms a solid white layer, a 1 mL pipette set to 900 μL with a non-filtered polypropylene tip can be carefully teased along the side of the tube, past the protein layer, and down into the chloroform phase. If the protein layer is less abundant or more fluffy, a 200 μL micropipette fitted with a gel-loading tip must be used, with the tip carefully inserted through the protein layer along the side of the tube. In either case, take care not to pipette the aqueous layer by first depressing the pipette plunger down to its first resistance point before inserting it into the tube, then slowly insert the tip past the aqueous and protein phases into the chloroform phase, and then depress the plunger slightly further to expel any aqueous phase until a small air bubble is passed. Then begin to draw the chloroform phase into the tip. Take care to minimize the amount of protein that sticks to the outer wall of the pipette tip.
28. Loss of some of the protein fraction due to adherence to the walls of the 15 mL conical vial as well as adherence of the protein to the inner and outer walls of the pipette tip is inevitable. The chloroform/methanol-BHT helps to reduce this loss; however, protein recovery is never 100%. Work hard to minimize this fractional loss of protein residue, but most importantly, try to keep the degree of loss consistent across all of the samples.
29. The net mass yield of the polar fraction can be used to assess consistency between samples and to determine split ratios for polar metabolite fraction aliquots. After weighing, the entire polar metabolite fraction can be stored at −80 °C for several days before thawing, sub-aliquoting, and lyophilization.
30. Avoid excessive drying time and exposure of the lipid fraction to heat/air. BHT is a preservative used in the food industry to stabilize fats; it is added here to slow the peroxidation of unsaturated lipids and fatty acids.
31. Make sure that the hinges of the microfuge tubes face outward from the center of the microcentrifuge rotor, so that the resulting protein residue pellet is located at the same position at the bottom of every tube.
32. Loss of protein residue during vacuum centrifugation can be minimized by attaching a separate cap cut from additional Eppendorf tubes with 4–8 perforations made using an 18-gauge needle or other perforation tool. Proteins can subsequently be redissolved in lysis buffer and their concentration measured by BCA or Bradford assay.

33. Record the mass of each polar aliquot as it is pipetted in order to determine the exact fraction represented by the aliquot, by dividing the fraction mass by the total mass of the polar fraction. Aliquoting into the polar fraction tubes can be performed directly on an analytical balance and the polar fraction masses recorded as they are being aliquoted.
34. Approximately 50–75% of the dried culture medium extract is sufficient for analysis by ^{1}H NMR, and 1/8–1/16 quantity of the extract is suitable for mass spectrometry applications.
35. The choice of solvent removal method for culture medium samples depends on the lability of the metabolite(s) of interest. For example, while lactate and glucose are relatively stable to temperature and pH, other metabolites of potential interest such as glutamine and pyruvate will degrade during vacuum centrifugation.
36. Approximately 50–75% of the dried extract is sufficient for analysis by ^{1}H NMR.
37. Centrifugal ultrafiltration membranes contain significant amounts of glycerol that will end up in the sample. If this is an issue, then prerinse and spin the membranes with ddH_2O prior to addition of your medium samples to the membranes.

Acknowledgments

This work was supported by the Intramural Research Program of the National Cancer Institute, NIH, NIH 1U24DK097215-01A1, 1P01CA163223-01A1, and 3R01ES022191-04S1.

References

1. Lane AN, Higashi RM, Fan TW (2016) Preclinical models for interrogating drug action in human cancers using stable isotope resolved metabolomics (SIRM). Metabolomics 12(7). https://doi.org/10.1007/s11306-016-1065-y
2. Buescher JM, Antoniewicz MR, Boros LG, Burgess SC, Brunengraber H, Clish CB, DeBerardinis RJ, Feron O, Frezza C, Ghesquiere B, Gottlieb E, Hiller K, Jones RG, Kamphorst JJ, Kibbey RG, Kimmelman AC, Locasale JW, Lunt SY, Maddocks OD, Malloy C, Metallo CM, Meuillet EJ, Munger J, Noh K, Rabinowitz JD, Ralser M, Sauer U, Stephanopoulos G, St-Pierre J, Tennant DA, Wittmann C, Vander Heiden MG, Vazquez A, Vousden K, Young JD, Zamboni N, Fendt SM (2015) A roadmap for interpreting (13)C metabolite labeling patterns from cells. Curr Opin Biotechnol 34:189–201. https://doi.org/10.1016/j.copbio.2015.02.003
3. Mackay GM, Zheng L, van den Broek NJ, Gottlieb E (2015) Analysis of cell metabolism using LC-MS and isotope tracers. Methods Enzymol 561:171–196. https://doi.org/10.1016/bs.mie.2015.05.016
4. Sellers K, Fox MP, Bousamra M 2nd, Slone SP, Higashi RM, Miller DM, Wang Y, Yan J, Yuneva MO, Deshpande R, Lane AN, Fan TW (2015) Pyruvate carboxylase is critical for non-small-cell lung cancer proliferation. J Clin Invest 125(2):687–698. https://doi.org/10.1172/JCI72873
5. Hensley CT, Faubert B, Yuan Q, Lev-Cohain N, Jin E, Kim J, Jiang L, Ko B, Skelton R, Loudat L, Wodzak M, Klimko C, McMillan E, Butt Y, Ni M, Oliver D,

Torrealba J, Malloy CR, Kernstine K, Lenkinski RE, DeBerardinis RJ (2016) Metabolic heterogeneity in human lung tumors. Cell 164 (4):681–694. https://doi.org/10.1016/j.cell.2015.12.034

6. Hui S, Ghergurovich JM, Morscher RJ, Jang C, Teng X, Lu W, Esparza LA, Reya T, Le Z, Yanxiang Guo J, White E, Rabinowitz JD (2017) Glucose feeds the TCA cycle via circulating lactate. Nature 551(7678):115–118. https://doi.org/10.1038/nature24057
7. Sun RC, Fan TW, Deng P, Higashi RM, Lane AN, Le AT, Scott TL, Sun Q, Warmoes MO, Yang Y (2017) Noninvasive liquid diet delivery of stable isotopes into mouse models for deep metabolic network tracing. Nat Commun 8 (1):1646. https://doi.org/10.1038/s41467-017-01518-z
8. Faubert B, Li KY, Cai L, Hensley CT, Kim J, Zacharias LG, Yang C, Do QN, Doucette S, Burguete D, Li H, Huet G, Yuan Q, Wigal T, Butt Y, Ni M, Torrealba J, Oliver D, Lenkinski RE, Malloy CR, Wachsmann JW, Young JD, Kernstine K, DeBerardinis RJ (2017) Lactate metabolism in human lung tumors. Cell 171 (2):358–371 e359. https://doi.org/10.1016/j.cell.2017.09.019
9. Davidson SM, Papagiannakopoulos T, Olenchock BA, Heyman JE, Keibler MA, Luengo A, Bauer MR, Jha AK, O'Brien JP, Pierce KA, Gui DY, Sullivan LB, Wasylenko TM, Subbaraj L, Chin CR, Stephanopolous G, Mott BT, Jacks T, Clish CB, Vander Heiden MG (2016) Environment impacts the metabolic dependencies of Ras-driven non-small cell lung Cancer. Cell Metab 23(3):517–528. https://doi.org/10.1016/j.cmet.2016.01.007
10. Keibler MA, Wasylenko TM, Kelleher JK, Iliopoulos O, Vander Heiden MG, Stephanopoulos G (2016) Metabolic requirements for cancer cell proliferation. Cancer Metab 4:16. https://doi.org/10.1186/s40170-016-0156-6
11. Kim J, Hu Z, Cai L, Li K, Choi E, Faubert B, Bezwada D, Rodriguez-Canales J, Villalobos P, Lin YF, Ni M, Huffman KE, Girard L, Byers LA, Unsal-Kacmaz K, Pena CG, Heymach JV, Wauters E, Vansteenkiste J, Castrillon DH, Chen BPC, Wistuba I, Lambrechts D, Xu J, Minna JD, DeBerardinis RJ (2017) CPS1 maintains pyrimidine pools and DNA synthesis in KRAS/LKB1-mutant lung cancer cells. Nature 546(7656):168–172. https://doi.org/10.1038/nature22359
12. Lane AN, Tan J, Wang Y, Yan J, Higashi RM, Fan TW (2017) Probing the metabolic phenotype of breast cancer cells by multiple tracer stable isotope resolved metabolomics. Metab Eng 43(Pt B):125–136. https://doi.org/10.1016/j.ymben.2017.01.010
13. Liu L, Shah S, Fan J, Park JO, Wellen KE, Rabinowitz JD (2016) Malic enzyme tracers reveal hypoxia-induced switch in adipocyte NADPH pathway usage. Nat Chem Biol 12 (5):345–352. https://doi.org/10.1038/nchembio.2047
14. Crooks DR, Maio N, Lane AN, Jarnik M, Higashi RM, Haller RG, Yang Y, Fan TWM, Linehan M, Rouault TA (2018) Acute loss of iron-sulfur clusters results in metabolic reprogramming and generation of lipid droplets in mammalian cells. J Biol Chem 293:8297. https://doi.org/10.1074/jbc.RA118.001885
15. Fan TW, Lane AN (2011) NMR-based stable isotope resolved metabolomics in systems biochemistry. J Biomol NMR 49(3–4):267–280. https://doi.org/10.1007/s10858-011-9484-6
16. Higashi RM, Fan TW, Lorkiewicz PK, Moseley HN, Lane AN (2014) Stable isotope-labeled tracers for metabolic pathway elucidation by GC-MS and FT-MS. Methods Mol Biol 1198:147–167. https://doi.org/10.1007/978-1-4939-1258-2_11
17. Jang C, Chen L, Rabinowitz JD (2018) Metabolomics and isotope tracing. Cell 173 (4):822–837. https://doi.org/10.1016/j.cell.2018.03.055
18. Bruntz RC, Lane AN, Higashi RM, Fan TW (2017) Exploring cancer metabolism using stable isotope-resolved metabolomics (SIRM). J Biol Chem 292(28):11601–11609. https://doi.org/10.1074/jbc.R117.776054
19. Fan TW-M (2012) The handbook of metabolomics. Springer, New York
20. Lane AN, Fan TW, Xie Z, Moseley HN, Higashi RM (2009) Isotopomer analysis of lipid biosynthesis by high resolution mass spectrometry and NMR. Anal Chim Acta 651 (2):201–208. https://doi.org/10.1016/j.aca.2009.08.032
21. Lane AN, Arumugam S, Lorkiewicz PK, Higashi RM, Laulhe S, Nantz MH, Moseley HN, Fan TW (2015) Chemoselective detection and discrimination of carbonyl-containing compounds in metabolite mixtures by 1H-detected 15N nuclear magnetic resonance. Magn Reson Chem 53(5):337–343. https://doi.org/10.1002/mrc.4199
22. Yang Y, Fan TW, Lane AN, Higashi RM (2017) Chloroformate derivatization for tracing the fate of amino acids in cells and tissues by multiple stable isotope resolved metabolomics (mSIRM). Anal Chim Acta

976:63–73. https://doi.org/10.1016/j.aca.2017.04.014

23. Lee WN, Boros LG, Puigjaner J, Bassilian S, Lim S, Cascante M (1998) Mass isotopomer study of the nonoxidative pathways of the pentose cycle with [1,2-13C2]glucose. Am J Phys 274(5 Pt 1):E843–E851
24. Saxena N, Maio N, Crooks DR, Ricketts CJ, Yang Y, Wei MH, Fan TW, Lane AN, Sourbier C, Singh A, Killian JK, Meltzer PS, Vocke CD, Rouault TA, Linehan WM (2016) SDHB-deficient cancers: the role of mutations that impair Iron sulfur cluster delivery. J Natl Cancer Inst 108(1):djv287. https://doi.org/10.1093/jnci/djv287
25. Lewis CA, Parker SJ, Fiske BP, McCloskey D, Gui DY, Green CR, Vokes NI, Feist AM, Vander Heiden MG, Metallo CM (2014) Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. Mol Cell 55(2):253–263. https://doi.org/10.1016/j.molcel.2014.05.008
26. Moseley HN, Lane AN, Belshoff AC, Higashi RM, Fan TW (2011) A novel deconvolution method for modeling UDP-N-acetyl-D-glucosamine biosynthetic pathways based on (13) C mass isotopologue profiles under non-steady-state conditions. BMC Biol 9:37. https://doi.org/10.1186/1741-7007-9-37
27. Moseley HN (2010) Correcting for the effects of natural abundance in stable isotope resolved metabolomics experiments involving ultra-high resolution mass spectrometry. BMC Bioinformatics 11:139. https://doi.org/10.1186/1471-2105-11-139
28. Freshney RI (2005) Culture of animal cells: a manual of basic techniques, 5th edn. John Wiley & Sons, Inc., Hoboken, NJ
29. Lane AN, Fan TW, Bousamra M 2nd, Higashi RM, Yan J, Miller DM (2011) Stable isotope-resolved metabolomics (SIRM) in cancer research with clinical application to nonsmall cell lung cancer. OMICS 15(3):173–182. https://doi.org/10.1089/omi.2010.0088
30. Theobald U, Mailinger W, Baltes M, Rizzi M, Reuss M (1997) In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae: I. experimental observations. Biotechnol Bioeng 55(2):305–316. https://doi.org/10.1002/(SICI)1097-0290(19970720)55:2<305::AID-BIT8>3.0.CO;2-M
31. Zhang GF, Sadhukhan S, Tochtrop GP, Brunengraber H (2011) Metabolomics, pathway regulation, and pathway discovery. J Biol Chem 286(27):23631–23635. https://doi.org/10.1074/jbc.R110.171405
32. Vasilakou E, Machado D, Theorell A, Rocha I, Noh K, Oldiges M, Wahl SA (2016) Current state and challenges for dynamic metabolic modeling. Curr Opin Microbiol 33:97–104. https://doi.org/10.1016/j.mib.2016.07.008
33. Cantor JR, Abu-Remaileh M, Kanarek N, Freinkman E, Gao X, Louissaint A Jr, Lewis CA, Sabatini DM (2017) Physiologic medium rewires cellular metabolism and reveals uric acid as an endogenous inhibitor of UMP synthase. Cell 169(2):258–272 e217. https://doi.org/10.1016/j.cell.2017.03.023
34. Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, Bouatra S, Sinelnikov I, Krishnamurthy R, Eisner R, Gautam B, Young N, Xia J, Knox C, Dong E, Huang P, Hollander Z, Pedersen TL, Smith SR, Bamforth F, Greiner R, McManus B, Newman JW, Goodfriend T, Wishart DS (2011) The human serum metabolome. PLoS One 6(2):e16957. https://doi.org/10.1371/journal.pone.0016957
35. Heeneman S, Deutz NE, Buurman WA (1993) The concentrations of glutamine and ammonia in commercially available cell culture media. J Immunol Methods 166(1):85–91
36. Nikfarjam L, Farzaneh P (2012) Prevention and detection of mycoplasma contamination in cell culture. Cell J 13(4):203–212
37. Sanchez JF, Crooks DR, Lee CT, Schoen CJ, Amable R, Zeng X, Florival-Victor T, Morales N, Truckenmiller ME, Smith DR, Freed WJ (2006) GABAergic lineage differentiation of AF5 neural progenitor cells in vitro. Cell Tissue Res 324(1):1–8
38. Liu X, Ser Z, Locasale JW (2014) Development and quantitative evaluation of a high-resolution metabolomics technology. Anal Chem 86(4):2175–2184. https://doi.org/10.1021/ac403845u
39. Sheikh KD, Khanna S, Byers SW, Fornace A Jr, Cheema AK (2011) Small molecule metabolite extraction strategy for improving LC/MS detection of cancer cell metabolome. J Biomol Tech 22(1):1–4
40. Wu JT, Wu LH, Knight JA (1986) Stability of NADPH: effect of various factors on the kinetics of degradation. Clin Chem 32(2):314–319
41. Nagana Gowda GA, Gowda YN, Raftery D (2015) Massive glutamine cyclization to pyroglutamic acid in human serum discovered using NMR spectroscopy. Anal Chem 87(7):3800–3805. https://doi.org/10.1021/ac504435b
42. Purwaha P, Silva LP, Hawke DH, Weinstein JN, Lorenzi PL (2014) An artifact in LC-MS/MS measurement of glutamine and

glutamic acid: in-source cyclization to pyroglutamic acid. Anal Chem 86(12):5633–5637. https://doi.org/10.1021/ac501451v

43. Deng P, Higashi RM, Lane AN, Bruntz RC, Sun RC, Ramakrishnam Raju MV, Nantz MH, Qi Z, Fan TW (2017) Quantitative profiling of carbonyl metabolites directly in crude biological extracts using chemoselective tagging and nanoESI-FTMS. Analyst 143 (1):311–322. https://doi.org/10.1039/c7an01256j
44. Kimball E, Rabinowitz JD (2006) Identifying decomposition products in extracts of cellular metabolites. Anal Biochem 358(2):273–280. https://doi.org/10.1016/j.ab.2006.07.038
45. Rabinowitz JD, Kimball E (2007) Acidic acetonitrile for cellular metabolome extraction from Escherichia coli. Anal Chem 79 (16):6167–6173. https://doi.org/10.1021/ac070470c
46. Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. Nat Chem Biol 5 (8):593–599. https://doi.org/10.1038/nchembio.186
47. Fan TW, Warmoes MO, Sun Q, Song H, Turchan-Cholewo J, Martin JT, Mahan A, Higashi RM, Lane AN (2016) Distinctly perturbed metabolic networks underlie differential tumor tissue damages induced by immune modulator beta-glucan in a two-case ex vivo non-small-cell lung cancer study. Cold Spring Harb Mol Case Stud 2(4):a000893. https://doi.org/10.1101/mcs.a000893
48. Lorkiewicz P, Higashi RM, Lane AN, Fan TW (2012) High information throughput analysis of nucleotides and their isotopically enriched isotopologues by direct-infusion FTICR-MS. Metabolomics 8(5):930–939. https://doi.org/10.1007/s11306-011-0388-y
49. Welshons WV, Wolf MF, Murphy CS, Jordan VC (1988) Estrogenic activity of phenol red. Mol Cell Endocrinol 57(3):169–178
50. Bindal RD, Carlson KE, Katzenellenbogen BS, Katzenellenbogen JA (1988) Lipophilic impurities, not phenolsulfonphthalein, account for the estrogenic activity in commercial preparations of phenol red. J Steroid Biochem 31 (3):287–293
51. Villas-Bôas SG (2007) Metabolome analysis: an introduction. Wiley-Interscience series in mass spectrometry. Wiley-Interscience, Hoboken, NJ
52. Kelly AE, Ou HD, Withers R, Dotsch V (2002) Low-conductivity buffers for high-sensitivity NMR measurements. J Am Chem Soc 124 (40):12013–12019

Chapter 2

Imaging Cancer Metabolism with Positron Emission Tomography (PET)

Timothy H. Witney and David Y. Lewis

Abstract

Positron emission tomography (PET) enables the noninvasive spatiotemporal analysis of cancer metabolism in vivo. Both natural and nonnatural PET tracers have been developed to assess metabolic pathways during tumorigenesis, cancer progression, and metastasis. Here we describe the dynamic in vivo PET/CT imaging of the glucose analogue [^{18}F]fluoro-2-deoxy-D-glucose (FDG), taking into consideration the methodology for alternative metabolic PET substrates.

Key words Fluorodeoxyglucose, FDG, Positron emission tomography, PET, Mouse, Imaging, Carbon-11, Fluorine-18

1 Introduction

Positron emission tomography noninvasively measures the uptake and retention of radiolabeled metabolites with picomolar sensitivity and submillimeter resolution, providing a three-dimensional view of cancer metabolism in living subjects. The most common radiolabeled metabolite is 2-[^{18}F]fluoro-2-deoxy-D-glucose (FDG), which is used as a surrogate for aerobic glycolysis, an often reported metabolic feature in cancer cells. FDG is taken up by glucose transporters, phosphorylated by hexokinase, and trapped intracellularly; it cannot be further metabolized and therefore measures the rate of glucose utilization rather than fluxing through the whole glycolytic pathway. FDG is used in the clinic to diagnose cancer, image invasion (staging), and monitor the effectiveness of cancer therapy [1].

To probe cancer metabolism, numerous other PET radiotracers have been developed in addition to FDG, including tracers that image fatty acid synthesis, fatty acid oxidation, oxidative stress, choline uptake and metabolism, amino acid uptake, and protein synthesis rate [2] (Fig. 1). These radiotracers tend to be, like FDG, either fluorinated analogues which are taken up by specific

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_2,

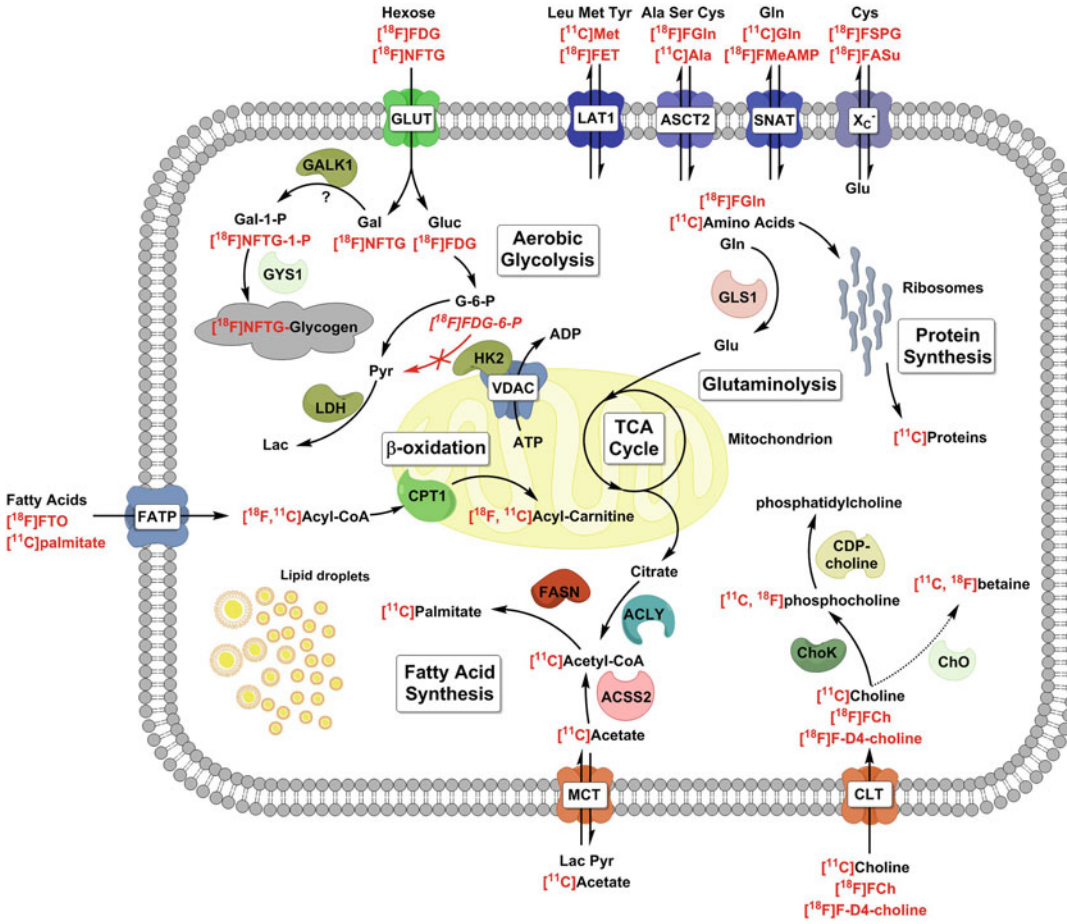


Fig. 1 A schematic of the metabolic pathways and enzymes responsible for the intracellular trapping of key PET substrates for imaging cancer metabolism. Radionuclides and PET substrates are shown in red. *ACLY* ATP citrate lyase, *ACSS2* acetyl CoA synthetase short-chain family member 2 cytosolic, *ASCT2* neutral amino acid transporter (SLC1A5), *Ala* alanine, *CPT1* carnitine palmitoyltransferase I, *CTL* choline transporter-like proteins (SLC44A), *Cys* cysteine, *FASN* fatty acid synthase, *FATP* fatty acid transport protein, *Gal* galactose, *GALK1* galactokinase 1, *GLS1* glutaminase 1, *Gluc* glucose, *Gln* glutamine, *Glu* glutamate, *GLUT* glucose transporter, *GYS1* UDP-glucose-glycogen glucosyltransferase, *HK2* hexokinase 2, *LAT1* L-type amino acid transporter 1 (SLC7A5), *LDH* lactate dehydrogenase, *MCT* monocarboxylate transporter, *SNAT* sodium-coupled neutral amino acid transporter, X_C^- anionic amino acid transporter light chain, system xc- (SLC7A11), *TCA* tricarboxylic acid cycle, *VDAC* voltage-dependent anion channel. Reproduced with permission from [2]

mechanisms but not extensively metabolized or carbon-11-labeled versions of endogenous substrates which are transported into the cell and metabolized via native pathways. There are advantages and disadvantages of each approach, for example, carbon-11 radiotracers require careful metabolic validation to determine what information can be obtained from the PET image, as PET provides no information of the chemical nature of the labeled metabolites. Similarly, great care needs to be taken when considering

non-natural substrates as surrogates as these tracers may not reflect genuine metabolic pathway flux due to their different affinities for cognate transporters or enzymes and retention may reflect atypical metabolism (*see* **Note 1**). The final choice of metabolic PET imaging will also be dictated by tracer availability (*see* **Note 2**).

This chapter takes you through a preclinical FDG PET/CT experiment, with special consideration given to nonstandard PET tracers, where relevant, in the Notes section. We describe the steps required for a successful FDG PET experiment, including PET/CT scanner, animal and radiotracer preparation, animal injection, PET/CT image acquisition, reconstruction, and analysis.

2 Materials

2.1 PET/CT Scanner Preparation

1. Small animal PET/CT scanner (e.g. Mediso NanoPET/CT) with integrated heated bed and animal respiratory and temperature monitoring.
2. Phantoms and calibration sources (Hounsfield, normalization, image quality, sodium-22 point source, 1 mL syringe with 3.7 MBq FDG).

2.2 Animal Preparation

1. Tumor-bearing mice or non-tumor-bearing control mice (*see* **Note 3**).
2. Absorbent paper.
3. Small animal heating chamber (e.g. small warm air system, Vet Tech).
4. Small animal heating plate (e.g. UNO controlled heating system).
5. Isoflurane anesthetic machine with medical air supply via central distribution system or by pressurized tank.
6. 1 mL syringes
7. 30 G needles
8. Polyethylene tubing (0.28 mm Ø).
9. 0.9% NaCl saline solution in water (bag of 100 mL)
10. Heparin sodium 1000 IU/mL solution for injection, 5 mL ampoule.
11. Needle forceps.
12. Zinc oxide tape (1.25 cm wide).
13. Animal balance.
14. Infrared heat lamp.
15. Topical skin adhesive.

2.3 Radiotracer Preparation and Animal Injection

1. FDG (or other metabolic PET radiotracer).
2. 5-cm-thick lead isotope workstation (e.g. Von Gahlen)
3. Contamination monitor (Geiger-Muller or plastic scintillator type).
4. 0.9% NaCl saline solution in water (bag of 100 mL)
5. Heparinized saline solution (bag of 100 mL; 50 IU/mL heparin sodium).
6. 1 mL syringes
7. Luer tip syringe caps.
8. Dose calibrator.
9. Isoflurane anesthetic machine with medical air supply via central distribution system or by pressurized tank.
10. Timer.

2.4 Image Analysis

1. DICOM database manager.
2. Image analysis software (VivoQuant, PMOD, Inveon research workplace, etc.)

3 Methods

3.1 PET/CT Scanner Preparation (Prior to Experiment)

1. PET/CT calibrations including Hounsfield calibration, activity calibration, normalization, and PET/CT alignment should be performed on a regular (minimum biannually) basis in advance of the imaging study (*see* **Note 4**).
2. PET/CT quality control (QC) also needs to be performed regularly (recommended intervals provided in parenthesis): PET detector check (daily), Hounsfield QC (weekly), PET/CT co-registration (weekly), image quality phantom (monthly), PET activity QC with Na22 point source (weekly), and F18 syringe (monthly).

3.2 Animal Preparation (the Day Prior to the Experiment)

1. Fast animals overnight the day prior to imaging (*see* **Note 5**). Water must be provided ad libitum.

3.3 PET/CT Scanner Preparation (on Day of Experiment)

1. X-ray tube conditioning of the CT scanner (or similar, depending on make/model) should be performed on the day of imaging at least 30 min prior to receiving radioactivity.
2. Perform daily PET detector QC using a Na22 point source.
3. The animal bed should also be prepared in advance of receiving the radioactivity.
4. Select the correct bed for single mouse imaging (*see* **Note 6**).

5. Line bed with absorbent paper.
6. Heat the animal bed to 37 °C.
7. Attach and position the animal breathing rate monitor and rectal temperature probe.
8. Turn on the isoflurane scavenger for the PET/CT scanner.
9. Enter the experiment and animal identification information into the study planner section of the PET acquisition software.

3.4 Animal Preparation

1. Switch on and set to 37–40 °C the small animal heating chamber 20–30 min prior to use.
2. Line with absorbent paper.
3. Switch on the small animal heating plate and set to 37 °C.
4. Prepare anesthesia machine by filling the isoflurane vaporizer, switching on the medical air delivery, and lining the induction box with absorbent paper (*see* **Note 7**).
5. Prepare heparinized saline by injecting 5 mL sodium heparin solution (5000 IU) into 100 mL bag of saline (0.9% NaCl).
6. Injection cannulas can be prepared (one cannula per mouse) using two 30G needles and a 10–15 cm length of polyethylene tubing (0.28 mm Ø). break one needle by holding with needle forceps and twisting repeatedly until the needle breaks at the plastic end. The blunt end of this broken needle can then be placed into one end of the polyethylene tubing. The second 30G needle can then be placed needle end first into the other end of the polyethylene tubing. Finally, ensuring that no bubbles are present, fill a 1 mL syringe with 200 μL heparinized saline, and place into the plastic port of the 30G needle (Fig. 2b).
7. Bring mice to the imaging laboratory, and place in the heating chamber for 10–30 min prior to cannulation in order to dilute the tail vein in preparation for cannulation (*see* **Note 8**).
8. Attach the induction chamber to the isoflurane vaporizer and switch on. The isoflurane concentration should be set to 3% for the induction of anesthesia.
9. Once the induction chamber is filled with isoflurane, remove one mouse at a time from the heating box, and place into the chamber.
10. When the mouse is fully anesthetized, remove from the chamber, and place on its side on the small heating plate, ensuring the nose is placed fully in the anesthetic nose cone. Change the concentration of isoflurane to 2%.
11. Weigh and record the mouse weight by briefly transferring to an animal balance before placing back on the anesthesia.

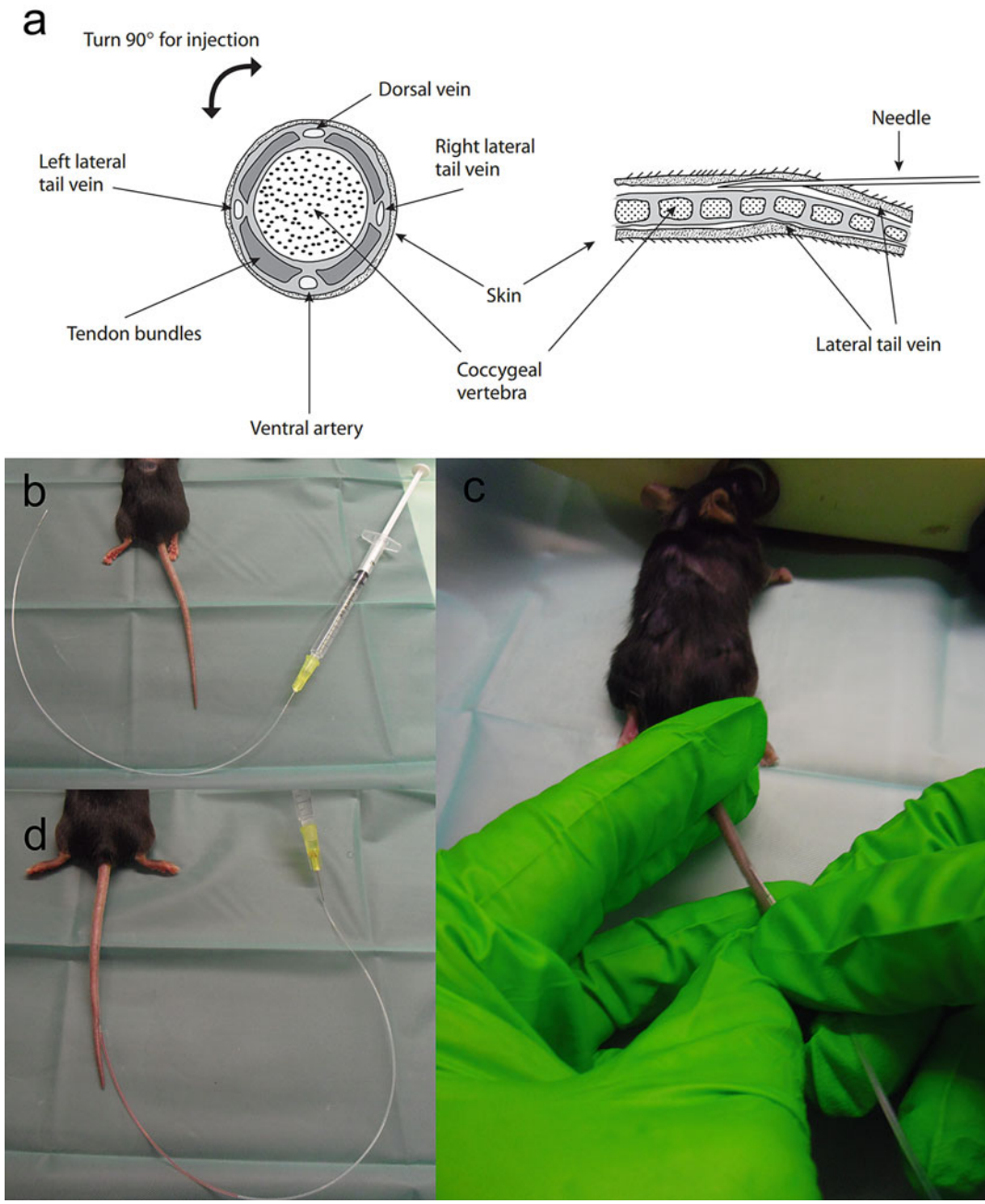


Fig. 2 Mouse tail vein cannulation procedure for i.v. injection of FDG. (**a**) Schematic of mouse tail anatomy showing path of needle insertion. Reproduced with permission from [12]. (**b**) Assembled tail vein cannula; (**c**) cannula is inserted into the lateral tail vein; (**d**) cannula in place secured using tissue adhesive shows blood flashback. Images courtesy of Dr. De-en Hu, University of Cambridge, UK

12. The heat lamp can also be used to ensure the mouse is warm and tail veins are fully dilated prior to inserting the cannula. Care must be taken not to place to heat lamp too close to the tail in order to prevent damage to the tail tissue.
13. Stroke the tail moving from a proximal to distal direction to further dilate the vein. Begin cannulation attempts as distal as possible, allowing more proximal attempts as necessary.
14. To insert the cannula, grasp the mouse tail using the thumb and index finger of the non-dominant hand, and bend the tail slightly so that the needle and the vein are parallel to each other (Fig. 2a). Then holding the needle of the cannula in the dominant hand with the bevel up, insert the needle into the lateral tail vein of the mouse, and advance the needle a few millimeters (Fig. 2c). When the cannula is inserted into the vein, blood may flash back into the cannula (Fig. 2d). Correct placement can be checked by flushing a small amount of saline into the vein; there will be little resistance if the cannula is in the correct place. If there is resistance and the tail around the injection site blanches white, then the needle should be withdrawn, and cannulation can be attempted again.
15. Once the cannula is inserted, secure it to the tail using topical skin adhesive (Fig. 2d).
16. Turn on the scanner's isoflurane vaporizer; set at 2% isoflurane.
17. Transfer the animal onto the PET scanner bed. Care must be taken to support the cannula and associated syringe during transfer so it does not become dislodged.
18. Place the mouse into the scanner headfirst in the prone position, ensuring the nose is placed fully in the anesthetic nose cone. Check that the respiratory pad is positioned on the chest of the mouse.
19. Secure the mouse to the bed with a piece of tape around the midriff, tight enough to put pressure on the breathing pad and to reduce excessive movement but loose enough to allow unrestricted respiration.
20. Secure the cannula tubing to the animal bed with tape.
21. Outstretch all of the limbs and secure these to the bed with tape.
22. Place the temperature probe into mouse rectum and secure with tape.
23. The isoflurane can be reduced to ~1% for maintenance; this will be model- and stain-dependent; however, the depth of anesthesia can be monitored using the respiration pads and monitoring system. The respiration rate should be maintained at about 60–100 breaths/min. If the breathing rate falls outside

these limits, then the concentration of isoflurane can be raised or lowered accordingly.

24. Similarly, the mouse body temperature should be maintained at 35–37 °C by changing the set temperature on the animal bed as necessary.
25. Position the bed in the scanner's center field of view using the CT's scout view by changing the bed's table height and extension into the scanner.
26. Metabolic tracers other than FDG may require different animal handling conditions to those described here for FDG (*see* **Note 9**).

3.5 Radiotracer Preparation

1. Collect FDG directly from the supplier or temporary storage location; as F18 and C11 have short half-lives, there is a necessity to work quickly.
2. Prepare radioactive doses inside a 5-cm-thick lead isotope workstation for radioprotection, and check regularly for personal and laboratory contamination using a monitor.
3. Following FDG collection, the exact concentration of activity in MBq/mL needs to be calculated. This can be achieved by placing 100 μL of the stock solution into the dose calibrator and measuring the total radioactivity. It is important to note the time of this measurement, so all calculations can be decay-corrected.
4. Estimate the time of scan, and using the following equation, calculate the amount of activity at the time of scan:

$$A(t) = A_0 \times e^{-\lambda t} \quad (1)$$

$A(t)$—Radioactive dose at time t (MBq).

$A(0)$—Original radioactive dose at time zero (MBq).

t—Elapsed time from initial radioactivity measurement (min).

λ—Decay constant:

$$\lambda = \frac{\ln(2)}{t_{1/2}} \quad (2)$$

$t_{1/2}$—Radioactive half-life (min); 109.7 min for F18 and 20.4 min for C11.

5. From the measured stock activity, calculate the amount of activity required to make a stock solution of 18.5 MBq/mL at the time of scan. 3.7 MBq is required per animal, and a maximum of 10 mL/kg should be injected per mouse.
6. Once the activity doses have been calculated for 3.7 MBq of injected activity, add the exact dose volume required into a

1 mL syringe, accounting for ~60 μL dead volume, and place a Luer tip syringe cap on the end.

7. Using the dose calibrator, measure the syringe activity noting the time $A_1(t_1)$ (*see* **Note 10**).

3.6 Animal Injection and PET/CT Scanning

1. Under the scout view tab, drag the scan area to cover the whole mouse, and conduct a scout (X-ray) scan of the mouse for positioning.
2. Using the scout view as a guideline, select the PET imaging volume making sure the tumor and any other regions of interest are central to the FOV (*see* **Note 11**).
3. For the PET acquisition parameters, set the radionuclide to F18 and the termination conditions to dynamic data collection over 90 min (*see* **Notes 12** and **13**).
4. Static PET imaging can also be conducted with FDG (*see* **Note 14**).
5. Set a timer for 15 s.
6. Remove the syringe containing heparinized saline from the end of the cannula, remove the Luer tip cap from the activity syringe, and place it to one side. Place the activity syringe on the end of the cannula taking care not to introduce any air into the injection volume.
7. When ready, start the scan (t_0) and the timer simultaneously.
8. After 15 s inject the mouse with radiotracer over a period ~3 s.
9. Remove the activity syringe and replace the Luer tip cap. Put the syringe containing heparinized saline back into the cannula again taking care not to introduce air, and flush the cannula with 100 μL of heparinized saline.
10. Measure the remaining radioactivity within the activity syringe noting the time, $A_2(t_2)$.
11. While the first scan is being performed, the next mouse can be prepared for scanning. During this 90-min period, other mice can be cannulated, and syringe radioactive doses can be drawn up and measured.
12. After the PET scan, acquire the CT over the same FOV as the PET using the following settings: Same scan range as PET acquisition, helical scan, 360 projections, 55 kVp, and 1100 ms exposure time (or similar depending on make and model of scanner) (*see* **Note 15**).
13. Following CT scan, remove the dose cannula, and press absorbent paper onto the mouse tail to stop bleeding.
14. Reduce the isoflurane concentration to zero, and place mouse into the heated chamber for recovery, if required.

15. When mouse is freely moving around, it can be put back into the home cage.
16. It may be necessary to keep the mouse in a radioactive designated area overnight, depending on the local radiation protection guidelines.
17. Alternatively, after the imaging session, place the mouse under terminal anesthesia if tissues are required for ex vivo analysis (*see* **Note 16**).

3.7 Reconstruction

1. Reconstruction parameters are specific to each PET scanner manufacturer and require optimization depending on your application. We recommend dynamic 3D iterative reconstruction with full detector modeling using four iterations and six subsets with 0.4 × 0.4 × 0.4 mm isotropic voxels and with decay (to the start time of the PET scan) and random, attenuation, and scatter correction, with 4 × 15 s, 4 × 1 min, 17 × 5 min dynamic time bins (*see* **Note 17**).

3.8 Image Analysis

1. Reconstructed PET and CT scans can be transferred to image analysis software through a DICOM server connection. Alternatively, DICOM files can be saved on a portable storage device if a DICOM server is not available.
2. PET and CT image datasets are loaded, which should be co-registered if the PET/CT alignment has been performed correctly (*see* Subheading 3.1).
3. Tumor uptake can be quantified by manual drawing of regions of interest (ROI) using structural detail from the CT scanner (*see* **Note 18**). For novel tracers, pharmacokinetics can additionally be extracted for normal tissue by placing a spherical ROI over these regions, as identified from the CT image. Muscle from the hind limb is frequently used as a measure of tracer uptake in background tissue.
4. Averaged tracer uptake across the tumor and/or tissue for each time frame should be extracted for analysis in Microsoft excel or similar tools.
5. Using Eq. (2) above, decay correctly the original syringe radioactivity $A_1(t_1)$ and the radioactivity remaining in the syringe after injection $A_2(t_2)$ to the start time of the PET scan (t_0). The injected dose (ID) at t_0 is then calculated:

$$ID(t_0) = A_1(t_0) - A_2(t_0) \tag{3}$$

6. The most commonly used metric for reporting PET is the standardised uptake value (SUV; *see* **Note 19**) and this can be calculated from the region of interest using the following formula:

$$SUV(t) = \frac{C(\mathrm{t})}{\mathrm{ID}(t_0)/\mathrm{BW}} \quad (4)$$

SUV(t)—Standardized uptake value at time t.

$C(t)$—Tissue radioactive concentration at time t in Bq/mL.

ID (t_0)—Injected dose at PET scan start time (Bq).

BW—Mouse body weight (g).

7. Finally PET/CT images can be displayed, with a scale bar, using appropriate thresholds to visualize the tumor; the skeleton can be surface-rendered using Otsu thresholding to give anatomical context to the PET image (Fig. 3).

4 Notes

1. C11 tracers have the advantage that they are "true tracers" and not analogues; therefore they follow endogenous metabolic pathways. However, this provides the added challenge that due to the lack of chemical resolution, there needs to be a careful consideration of tracer metabolism. The determination of tracer metabolism can be performed ex vivo using radioactive thin-layer chromatography (TLC) or high-performance liquid chromatography (HPLC) with radiodetection to determine the radiochemical purity. Alternatively, the metabolic species can be determined using C13 labeling as a surrogate for C11 and HPLC or gas chromatography (GC)-mass spectrometry [3]. Similarly, metabolism of F18 tracers can be followed using nuclear magnetic resonance with F19 as a surrogate. The caveat here is that C13 and F19 detection requires much higher concentrations than required for analogous detection of C11 and F18 by PET which may perturb the metabolic pathway under investigation.
2. In most situations PET imaging of nonstandard tracers requires access to specialized radiochemistry facilities as generally only FDG (and possibly a few other metabolic tracers) is available commercially. Fluorine-18-labeled tracers have a radioactive half-life of 109.7 min and therefore can to be transported over short (>4 hour) distances. Carbon-11 has a half-life of 20.4 min, so it requires a nearby cyclotron facility.
3. The applicability and relevance of the information gained from metabolic PET imaging is proportional to the quality of the mouse model used. Where available, always use the most advanced cancer models for imaging such as genetically engineered mice, patient-derived xenografts, somatic cell transduction, and orthotopically implanted organoids rather than reliance on subcutaneously implanted cancer cell line models [4].

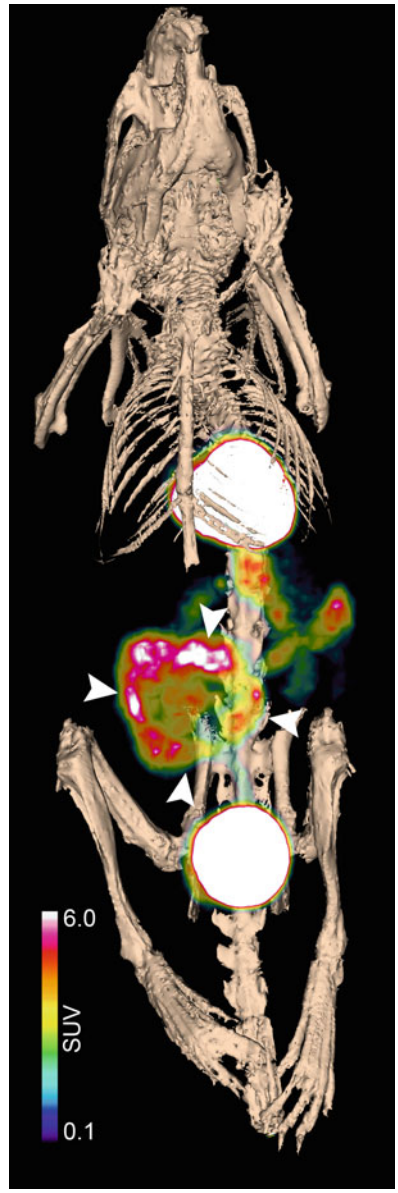


Fig. 3 FDG PET/CT image of a $Kras^{G12D/+}$; $Tpr53^{R172H/+}$; Pdx-1-Cre mouse bearing a pancreatic ductal adenocarcinoma (PDAC) tumor (white arrows) with high glucose uptake. Physiological FDG uptake in the heart, bladder, and kidneys can also be observed

4. PET scanners can only be considered accurate if calibrations and QC procedures are followed. Important calibrations include detector functioning, normalization (to give a uniform image in the field of view despite differing sensitivity), PET/CT alignment, and quantification calibration [5].

5. Fasting mice prior to FDG PET imaging increases the tumor and reduces physiological uptake improving tumor contrast by reducing the plasma blood glucose level [6].
6. Depending on the size of the transaxial field of view, it may be possible to use a multi-mouse bed for imaging up to four mice simultaneously [7]. This can greatly improve the efficiency of PET imaging, especially for C11 studies.
7. Imaging requires pharmacological restraint using general anesthesia. There have been a number of studies looking at the effects of anesthetic and carrier gas on FDG uptake [8, 9], the best combination to maintain high tumor, and low background uptake is 0.5–1.0% isoflurane (or sevoflurane, if available) and medical air. 100% oxygen as a carrier gas may be beneficial to maintain blood oxygen saturation in long PET experiments (>4 h) or where mice are respiratory compromised such as autochthonous lung tumor-bearing mice.
8. Mice should be pre-warmed before a PET scan and during PET imaging as this has a beneficial effect on FDG distribution, reducing physiological uptake in other organs, such as brown fat, and improving the image contrast between the tumor and the background tissues [10]. Under anesthesia mice are unable to regulate their core body temperature; therefore, external support is required to keep the animal temperature at 37 °C. External heat sources should be set to the range of 37–40 °C, but careful temperature monitoring using a rectal probe is required, adjusting the external heat source accordingly. There is less evidence for temperature-dependent uptake of other metabolic PET tracers, but animal heating is important for consistency, animal welfare and to preserve thermoneutral metabolism.
9. There is little data optimizing anesthesia and animal handling regimes for metabolic PET tracers other than FDG, so careful consideration needs to be given. Additionally, there is little data for the benefits of fasting for the uptake of other tracers; it may reduce tumor uptake in some cases. If mice are sick due to high tumor burden, then fasting can push them into a torpor, which could be considered non-relevant. To be sure of the effect of any animal handling procedures (i.e. diet, anesthesia, or temperature), PET tracer uptake should be compared under the different proposed conditions.
10. The time recorded for the injected radioactivity dose needs to be synchronous with the time of the PET/CT scanner as these are used for decay correction and a few minutes' difference can be significant particularly with C11 imaging [11].
11. The center of the PET field of view is the point in the PET scanner with the highest resolution and sensitivity (hence

signal-to-noise ratio); therefore, it is important to position the animal so that the tumor (or other regions of interest) is central.

12. Initial PET studies with novel tracers should always be dynamic, i.e. injection and imaging of the animal throughout the uptake and distribution period. FDG is irreversibly bound inside the cell, so uptake should increase until a plateau, usually 60–90 min after injection. A number of metabolic PET tracers (e.g. FET, FSPG, FACBC) are reversible; therefore they will wash in and out of the tumor. Here the plateau phase is transient and varies between different models; therefore, dynamic imaging is frequently required to determine the optimal imaging time point. Often the peak of the tumor uptake is used, but the time point with maximum tumor-to-background ratio will improve visualization. The optimal time point also depends on the tracer metabolism to ensure that the process of interest is dominant in the PET image [3].
13. Dynamic PET data can be analyzed using quantitative models, ranging in complexity from graphical methods (i.e., Patlak plot) using image-derived input functions to fully quantitative multi-compartment models using metabolite-corrected arterial plasma samples. The advantage of PET pharmacokinetic modeling is that it allows the derivation of in vivo flux measurements, which take into account variable tracer delivery and whole-body tracer metabolism [10].
14. The kinetics of FDG tissue distribution are well-known; therefore, static PET scans (~10 min) are often performed. Here, the animal is injected with radioactivity outside of the scanner, with imaging performed at a pre-determined time point (s) after injection. For FDG this should be a minimum of 60 min after injection. It is important to keep a consistent imaging time, as small variations in time may result in large differences in tissue uptake of the radiotracer. Another point to consider for static scans is whether the animals are allowed to recover from anaesthesia in-between tracer injection and scanning. This will have a large effect on tracer pharmacokinetics due to the altered metabolism, breathing, and heart rates experienced during anesthesia.
15. The purpose of the CT scan is threefold: to provide anatomical context to the PET image, to produce an attenuation map for PET signal attenuation correction and, where possible, to localize the tumor.
16. If tissue needs to be collected immediately after imaging for dissection, gamma counting, and other downstream analysis, CT scans can be performed prior to PET imaging.

17. Most manufacturers will have their own reconstruction algorithms for PET reconstruction; these can generally be separated into 2D or 3D filtered-back projection (FBP) and 2D or 3D iterative reconstruction. Generally, 3D iterative reconstruction such as 3D-ordered subset expectation maximization (OSEM) is preferable as it gives the best image quality. 3D iterative reconstructions have high computational demands and therefore used to have long reconstruction times. Newer reconstruction algorithms run on graphics processing units (GPU) allowing dynamic 3D reconstructions in hours rather than days.
18. There is less selection bias when performing region of interest analysis on the CT rather than the PET image. However, as CT soft tissue contrast is poor, visualizing autochthonous or orthotopic tumors can be difficult. Here, image thresholding of the PET scan may be required (i.e. selecting all the voxels inside the tumor volume above a certain value). Due to the lack of CT soft tissue contrast, MRI is generally a better imaging modality for tumor segmentation.
19. SUV is the radioactivity concentration normalized to the injected dose and the body weight and provides a semiquantitative measure of metabolic activity. An SUV of 1 is equivalent to complete and uniform distribution of radioactive signal throughout the body. Therefore, tumor SUV >1 suggests tumor localization, and SUV is, for the most part, proportional to the glucose utilization rate. In addition, the percentage of injected activity, normalized for the volume of the tissue or organ of interest, is frequently used for data analysis in preclinical imaging studies when body weights are broadly equivalent.

References

1. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, Verzijlbergen FJ, Barrington SF, Pike LC, Weber WA, Stroobants S, Delbeke D, Donohoe KJ, Holbrook S, Graham MM, Testanera G, Hoekstra OS, Zijlstra J, Visser E, Hoekstra CJ, Pruim J, Willemsen A, Arends B, Kotzerke J, Bockisch A, Beyer T, Chiti A, Krause BJ, European Association of Nuclear M (2015) FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging 42(2):328–354. https://doi.org/10.1007/s00259-014-2961-x
2. Lewis DY, Soloviev D, Brindle KM (2015) Imaging tumor metabolism using positron emission tomography. Cancer J 21(2):129–136. https://doi.org/10.1097/PPO.0000000000000105
3. Lewis DY, Boren J, Shaw GL, Bielik R, Ramos-Montoya A, Larkin TJ, Martins CP, Neal DE, Soloviev D, Brindle KM (2014) Late imaging with [1-^{11}C]acetate improves detection of tumor fatty acid synthesis with PET. J Nucl Med 55(7):1144–1149. https://doi.org/10.2967/jnumed.113.134437
4. Gengenbacher N, Singhal M, Augustin HG (2017) Preclinical mouse solid tumour models: status quo, challenges and perspectives. Nat Rev Cancer 17(12):751–765. https://doi.org/10.1038/nrc.2017.92
5. Osborne DR, Kuntner C, Berr S, Stout D (2017) Guidance for efficient small animal imaging quality control. Mol Imaging Biol 19 (4):485–498. https://doi.org/10.1007/s11307-016-1012-3

6. Fueger BJ, Czernin J, Hildebrandt I, Tran C, Halpern BS, Stout D, Phelps ME, Weber WA (2006) Impact of animal handling on the results of ^{18}F-FDG PET studies in mice. J Nucl Med 47(6):999–1006
7. Yagi M, Arentsen L, Shanley RM, Hui SK (2014) High-throughput multiple-mouse imaging with micro-PET/CT for whole-skeleton assessment. Phys Med 30 (7):849–853. https://doi.org/10.1016/j.ejmp.2014.06.003
8. Flores JE, McFarland LM, Vanderbilt A, Ogasawara AK, Williams SP (2008) The effects of anesthetic agent and carrier gas on blood glucose and tissue uptake in mice undergoing dynamic FDG-PET imaging: sevoflurane and isoflurane compared in air and in oxygen. Mol Imag Biol 10(4):192–200. https://doi.org/10.1007/s11307-008-0137-4
9. Woo S-K, Lee TS, Kim KM, Kim J-Y, Jung JH, Kang JH, Cheon GJ, Choi CW, Lim SM (2008) Anesthesia condition for ^{18}F-FDG imaging of lung metastasis tumors using small animal PET. Nucl Med Biol 35(1):143–150. https://doi.org/10.1016/j.nucmedbio.2007.10.003
10. Kuntner C, Stout D (2014) Quantitative preclinical PET imaging: opportunities and challenges. Front Phys 2(12). https://doi.org/10.3389/fphy.2014.00012
11. Boellaard R (2009) Standards for PET image acquisition and quantitative data analysis. J Nucl Med 50(Suppl 1):11S–20S. https://doi.org/10.2967/jnumed.108.057182
12. Hirota J, Shimizu S (2012) Chapter 5.2 - routes of administration A2 - Hedrich, Hans J. In: The laboratory mouse, 2nd edn. Academic Press, Boston, pp 709–725. https://doi.org/10.1016/B978-0-12-382008-2.00030-1

Chapter 3

Radioluminescence Microscopy: A Quantitative Method for Radioisotopic Imaging of Metabolic Fluxes in Living Cancer Cells

Debanti Sengupta and Guillem Pratx

Abstract

Radionuclide imaging with cellular-scale resolution allows characterization of biological processes and metabolic fluxes in single live cells. In this protocol, we describe how to image radiotracer uptake with single-cell resolution and compare the method to conventional bulk-scale gamma counting. We describe the utility of both techniques, give examples where each technique is recommended, and provide detailed side-by-side instructions for both techniques.

Key words Single-cell analysis, Cancer metabolism, Radiotracer quantitation, Microscopy

1 Introduction

Radioluminescence microscopy (RLM) is an emerging modality designed for high-resolution imaging of radionuclides. In particular, this technique can be utilized to image the uptake of biologically relevant radiotracers in single living cells. Applications of this technique range from assessing cancer cell proliferation [1] to characterizing metabolic fluxes [2] and measuring radiolabeled drug uptake [3]. The technique sets itself apart because of its ability to image small molecules quantitatively at cellular resolution. Molecules such as glucose cannot accurately be imaged through conventional microscopy techniques such as fluorescence microscopy but can be radiolabeled effectively through radiochemistry. With the increased use of in vivo positron-emission tomography (PET), there is now an abundance of new radiotracers that can be used to quantify metabolic processes in live cancer cells [4]. Moreover, depending on the specific application, radioluminescence microscopy may also be multiplexed with fluorescence and brightfield microscopy, allowing multi-parametric imaging of single-cell behaviors [5].

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_3,

While other techniques exist to probe radionuclide uptake in biological samples, the singular strength of RLM lies in its ability to characterize radiotracer signal with high resolution. While bulk methods such as gamma counting and liquid scintillation counting have provided useful information regarding radiotracer uptake in biological samples, they typically measure millions of cells at once, which obscures the specific pattern of uptake of single cells. This is particularly important when probing heterogeneity within a mixed cell population or tissue–for instance, a mixture of heterogeneous cancer cells and stromal cells.

This paper lays out techniques for measuring radionuclide uptake in cells using both RLM and gamma counting techniques.

In general, we recommend *RLM* for the following applications:

- When there is significant cell heterogeneity.
- When the biological sample contains two or more cell types (co-culture) with different responses to the radiotracer in question.
- When the number of cells is low, as in the case of biopsy samples, stem cells, circulating tumor cells, or other rare cell types.

We recommend *bulk gamma counting* for the following applications:

- When the cell population is homogeneous, meaning that the average is a good measure of the population as a whole.
- When the biological sample contains multiple populations of cells with similar responses to the radionuclide in question.
- When one is not limited by the number of cells – usually, bulk experiments require 10^4–10^6 cells.

2 Materials

1. Radiotracer (suitable isotopes include ^{18}F, ^{11}C, ^{68}Ga, and ^{64}Cu).
2. Bioluminescence microscope equipped with EMCCD camera or custom-built low-light microscope [5].
3. Microscopy immersion oil.
4. Microscopy objectives: 40× (oil, 1.3 NA) and 20× (air, 0.75 NA).
5. Bulk gamma counter.
6. Centrifuge.
7. Eppendorf tubes.
8. Falcon tubes (15 and 50 mL).

9. Radiation dosimeter badge/ring.
10. Glass-bottom microscopy dishes.
11. Fibronectin.
12. Scintillators (we recommend $CdWO_4$, both sides polished, 0.5 mm thick).
13. Ethanol (70%).
14. Piranha solution (three parts concentrated sulfuric acid to one part 30% hydrogen peroxide solution).
15. Tweezers.
16. Cell culture media (with and without glucose).
17. Camera software (MetaMorph or HCImage).
18. ORBIT software package for the reconstruction of radioluminescence images (found at http://med.stanford.edu/pratxlab/research/downloads.html).
19. MATLAB software.
20. Live mammalian cells.

3 Methods

3.1 Experiment Preparations (Day 1)

1. For nonadherent cells, culture cells as per standard protocol, for instance, in cell culture flasks.
2. For RLM imaging of adherent cells, seed cells on a glass-bottom microscopy dish the night before (Fig. 1). The dish must be large enough for the scintillator to fit. Depending on how quickly cells multiply, the recommended seeding density is

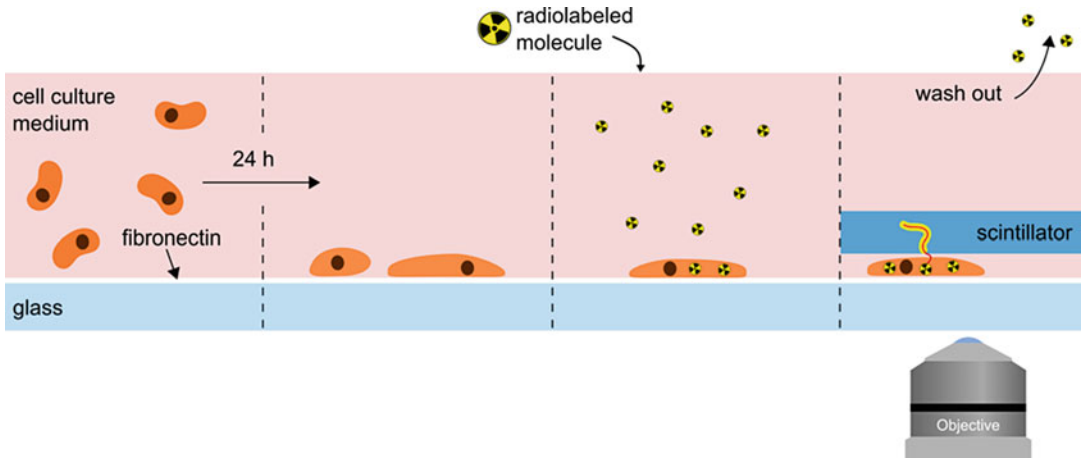


Fig. 1 Overview of the protocol. *From left to right*: Cells are seeded onto a fibronectin-coated glass-bottom dish, then incubated overnight. A radiotracer is introduced and taken up by the cells. After washing, a scintillator is placed above the cells and the radioluminescence resulting from individual ionization tracks is imaged via a low-light microscope

50,000–100,000 cells for a 35 mm dish. To increase cell adherence, one may adsorb a biomaterial coating onto the dish. This can be done prior to cell seeding by adding a fibronectin solution to the dish at a concentration of 10 μg/mL in PBS for 30 min at 37°C and then washing three times prior to seeding (*see* **Note 1**).

3. For RLM imaging, sterilize scintillators by submerging in ethanol solution overnight.
4. For bulk gamma counting of adherent cells, seed a 6-well plate with 10^5 cells per well as per standard cell culture protocol. Seed a minimum of three wells per condition tested.

3.2 Radiotracer Cell Labeling

1. Obtain desired radiotracer dose from radiochemistry facility (*see* **Note 2**). Follow institutional guidelines for handling radioactive substances.
2. Dilute radiotracer to the desired concentration for incubation with cells. Note that the relevant parameter for cell uptake is *radioactivity concentration*, not total radioactivity per well. Therefore, use small volumes when possible to minimize the use of radioactivity.
3. *For RLM imaging*: We recommend a radiotracer concentration of 100–500 μCi (9.3 MBq) per mL. This amount may be adjusted depending on the cells and radiotracers being used. Dilute radionuclide using cell medium or PBS. For [^{18}F]FDG experiments, we recommend diluting the radiotracer with glucose-free medium with 10% FBS to facilitate rapid transport of FDG into cells without competition from unlabeled glucose. Incubate cells with radiotracer for a minimum of 30 min at 37 °C under 5% CO_2. Incubate adherent cells directly in the imaging dish and nonadherent cells in 6-well plates. After incubation, remove medium (noting that it is radioactive and thus must be allowed to decay prior to disposal), and wash cells with PBS three times. Adherent cells should be washed directly in the imaging dishes, and new medium should be added afterward (use at least 1 mL of medium to avoid evaporation during imaging). Nonadherent cells should be washed by centrifugation (3 min, 1300 RPM).
4. *For bulk gamma counting:* We recommend a radiotracer concentration of 10–50 μCi/mL. Incubation with radiotracer should be performed following the same protocol as for RLM, but a lower activity concentration is used due to the larger number of cells. Incubation and washing steps can be performed directly in the 6-well plate. Cells should then be removed from the dish by trypsinization (if adherent), washed three times by centrifugation, and transferred into a vial to be assayed. The radioactivity of each vial should not exceed 1 μCi to avoid saturating the sensitive gamma counter.

3.3 Radioluminescence Microscopy

1. Using tweezers, remove the scintillator from the ethanol solution, and pat dry on a Kimwipe or lens paper. Place scintillator gently on top of the cells in the dish so that the scintillator is submerged in medium (Fig. 1). Avoid moving the dish. It is important for the scintillator to be very close to the cells (i.e., 10 μm or less). For nonadherent cells, put a 10 μL drop containing 10^4 radiolabeled cells on a glass-bottom dish, and then gently place the scintillator on top of the drop, allowing the liquid to spread. Matrigel or other quick-gelling substance can be added to the drop to immobilize the nonadherent cells under the scintillator. Fill the dish with 1 mL medium.
2. Turn microscope on (note that the EMCCD camera requires 30 min to cool down).
3. Use the 40× oil objective for larger magnification or the 20× air objective for larger field of view (*see* **Note 3**).
4. To focus the microscope, switch to brightfield mode. Once the cells are in focus, move the XY stage to an area containing a suitable density of cells (see Fig. 3, e.g.; *see* **Note 4**). If the cells are too sparse, little information will be obtained from the experiment. If they are too dense, it may not be possible to achieve accurate single-cell measurements.
5. Turn the bright-field lamp off, and set pixel binning to 4 × 4, electron-multiplication (EM) gain to maximum value (1200), and exposure time to 100 ms. Then turn on "live mode" to visualize radioactive ionization tracks in the scintillator.
6. For the sharpest RLM images, the microscope must be refocused so that the focal place lies exactly on the interface between the scintillator and the cells (Fig. 2). First, move the

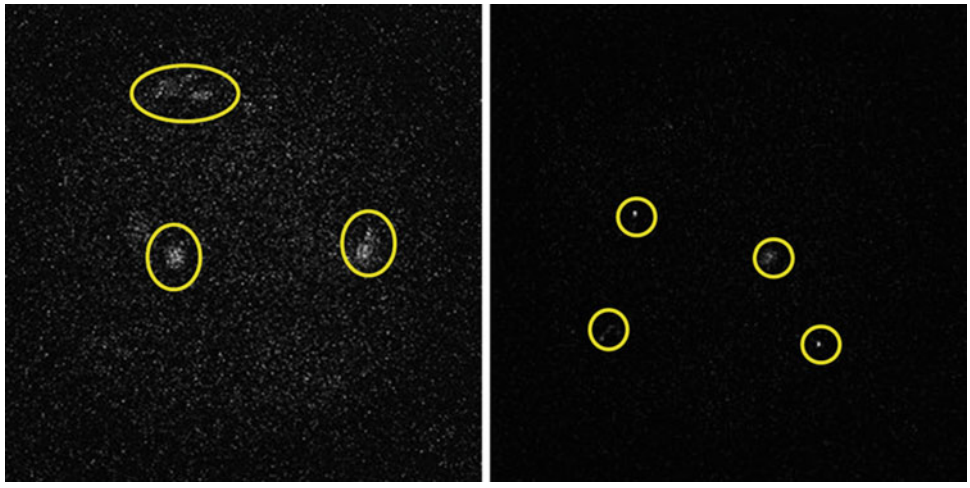


Fig. 2 A comparison of ionization tracks obtained using an out of focus frame (*left*) and in focus frame (*right*). Individual ionization tracks are circled in yellow

objective down, away from the scintillators. The radioactive decay events (called ionization tracks) will slowly become blurry blobs. Then, slowly move the focal plane back toward the scintillator. At the interface, the ionization tracks should transition from blurry back to sharp, as demonstrated in Fig. 2. Set the focal plane to the very first position where the events stop being blurry and become sharp. If the focal plane is moved deeper inside the scintillator, the events will still look sharp, but the reconstructed image will be blurry. This is because the ionizing particles extend deep into the scintillator.

7. Note the number of ionization tracks captured per frame. Ideally, there should be between 5 and 10 tracks per frame. If too many tracks are being captured in each frame, reduce exposure time. If there are too few tracks, increase exposure time.
8. Once you are satisfied with the focusing and number of tracks per frame, begin RLM image acquisition by acquiring a sequence of 10,000 frames (or more, if lower noise is desired). The total acquisition time can be estimated as (number of frames) × (exposure per frame) (*see* **Note 5**). Data should be saved as 16-bit TIFF (*see* **Note 6**). Acquisition of RLM images should take between 20 and 30 min (*see* **Note 7**).
9. Once acquisition is complete, take another bright-field image to confirm that cells are still in focus and in the same location. Fluorescence and/or bioluminescence images can also be acquired if desired.
10. Reconstruct the stack of 10,000 frames using the ORBIT software package in MATLAB (Fig. 3).
11. Quantitation of RLM signal per cell can be performed by drawing circular regions of interest (ROI) around individual cells (Fig. 3). It is highly recommended to use the same ROI size for all the cells (*see* **Note 8**). The number of frames acquired, the decay time, the sensitivity of the microscope, and the duration of the experiment must be considered to estimate radiotracer per cell. Similar regions of interest can be drawn to estimate background radiation levels outside of the cells. These ROI measurements are in unit of counts per minute (cpm) per cell. A calibration is required to convert these measurements into units of radioactivity (Bq per cell). Compartmental kinetic modeling can also be used to estimate the metabolic rate of glucose utilization by the cells.
12. Correct the measured data to account for radioactive decay over the course of the experiment, using the formula:

$$A_{\text{ref}} = A_{\text{mes}} \times 2^{\frac{t-t_0}{T_{\text{HL}}}}$$

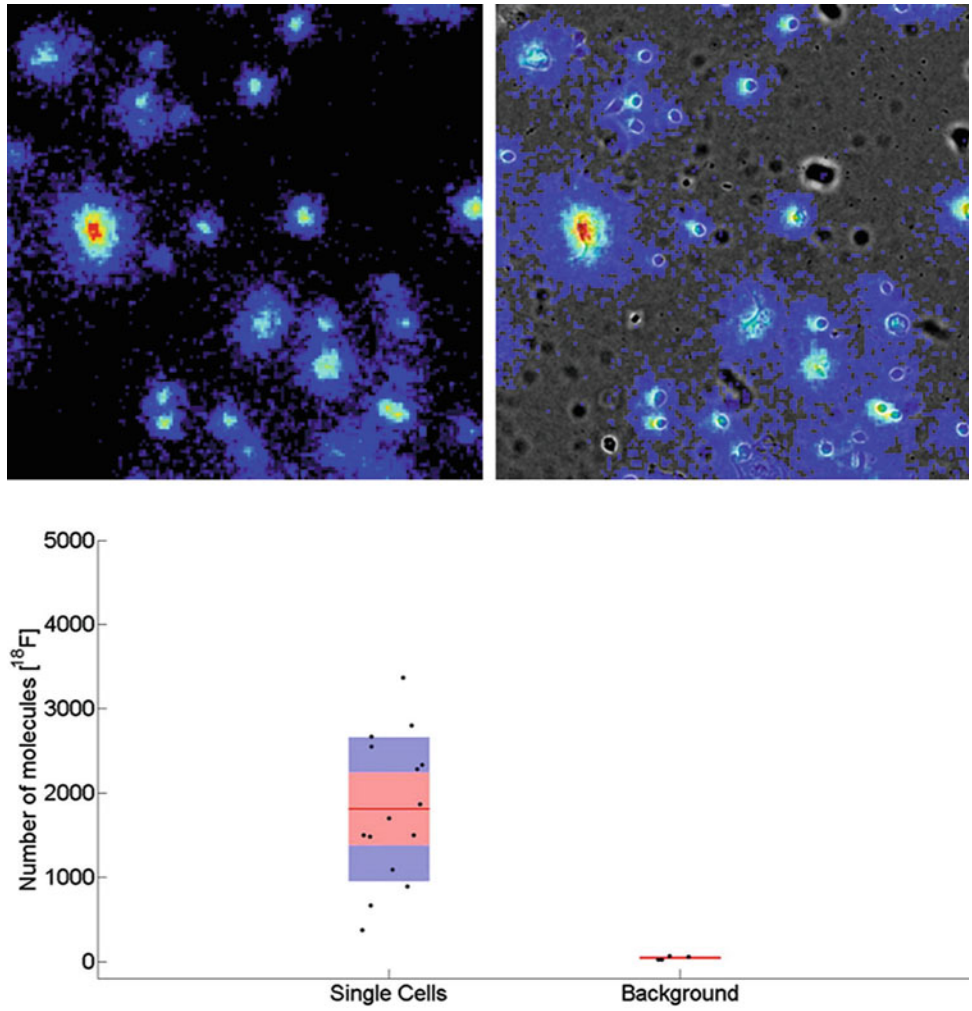


Fig. 3 A sample radioluminescence image reconstructed using ORBIT (*top panel*). Reconstructed radioluminescence image (*top left*) and (*top right*) same image overlaid onto brightfield image showing co-localization of [^{18}F]FDG uptake to individual MDA-MB-231 cells. An example of ROI analysis performed to quantitate FDG concentration in single cells as well as background radiation levels using ORBIT software (*bottom panel*)

where A_{ref} is the activity at reference time, t_0, A_{mes} is the activity measured at time t, and T_{HL} is the half-life of the measured radionuclide (109 min for ^{18}F). You can also use an online decay calculator.

13. If imaging multiple samples, start incubating the $(N + 1)$th sample while acquiring the images for the Nth sample. Each sample should be imaged promptly after the final wash to minimize biological efflux of the tracer from the cells and to ensure that all samples are exposed to the same radiotracer concentration for the same amount of time prior to imaging.

Due to radioactive decay, the volume of the radiotracer solution that is diluted with cell medium for incubation should be adjusted to obtain the same activity concentration.

3.4 Bulk Gamma Counting

1. Prior to imaging, a calibration ladder should be established by preparing a series of solutions of known radioactivity. We recommend radioactivity of 0, 0.25, 0.5, 0.75, and 1 μCi. These solutions should be prepared by diluting a stock solution measured using the radioactive dose calibrator. Each solution should be prepared in triplicate.
2. Follow the instructions of your specific gamma counter to obtain readings for each individual sample. The recommended integration time is 30–60 s per sample.
3. Gamma counters generally correct for radioactive decay between samples but not for decay occurring between sample preparation and sample measurement. Decay correction can be performed as previously explained.
4. Gamma counters measure total radioactivity per vial. Radioactivity measurements should be normalized to sample volume, number of cells per vial (measured by cell counting), or amount of protein per vial (measured using Bradford assay).

3.5 Post Experiment

1. Allow radioactivity to decay for at least 10 half-lives before disposing of contaminated waste. For ^{18}F, all radioactivity decays to background levels within 24 h. Always survey radioactivity levels before disposing of waste. Follow your institution's protocols for safe disposal of any items that have been in contact with radionuclides and/or biological substances.
2. Scintillators can be washed in ethanol if contact with cells was limited. In case of extended contact or organic buildup, scintillators made of $CdWO_4$ can be cleaned using a 10-min incubation in piranha solution.

4 Notes

1. If imaging cells on a plastic or coated surface, be aware that some radiotracers may absorb onto such surfaces, producing a high level of background signal.
2. Always order extra radiotracer prior to running your experiment. Calculate the amount needed prior to the experiment, factoring in decay time. It is preferable to obtain freshly made radiotracer; radiotracer that has undergone significant decay (>1 half-life) has lower radioactivity concentration and lower specific activity. Radioactivity concentration should be at least 1 mCi/mL and preferably higher.

3. When using a bioluminescence microscope, the effective magnification is 4–5 times lower than the nominal magnification indicated by the objective due to the nonstandard tube lens. For instance, images acquired with the 40× objective are only magnified 8×.
4. When imaging cells, avoid the edges of the scintillator, and try to image near the center of the scintillator because light piping at the edge of the scintillator will produce artificial background.
5. We do not recommend imaging samples for longer than one half-life; 50% of the signal is measured during the first half-life, 25% during the second half-life, 12.5% during the third half-life, and so on.
6. Invest in a fast, multiple TB hard drive, and expect raw data files in the 1–10 GB range.
7. Depending on the software and specifications of the computer used, budget in the time needed to save a batch of data (it can take up to 20 min for 10,000 frames).
8. Always order extra radiotracer prior to running your experiment. Calculate the amount needed prior to the experiment, factoring in decay time. It is preferable to obtain freshly made radiotracer; radiotracer that has undergone significant decay (>1 half-life) has lower radioactivity concentration and lower specific activity. Radioactivity concentration should be at least 1 mCi/mL and preferably higher.

References

1. Sengupta D, Pratx G (2016) Single-cell characterization of FLT uptake with radioluminescence microscopy. J Nucl Med 57(7):1136–1140. https://doi.org/10.2967/jnumed.115.167734
2. Pratx G, Chen K, Sun C, Martin L, Carpenter CM, Olcott PD, Xing L (2012) Radioluminescence microscopy: measuring the heterogeneous uptake of radiotracers in single living cells. PLoS One 7(10):e46285
3. Natarajan A, Türkcan S, Gambhir SS, Pratx G (2015) Multiscale framework for imaging radiolabeled therapeutics. Mol Pharm 12 (12):4554–4560
4. Rauscher I, Eiber M, Souvatzoglou M, Schwaiger M, Beer AJ (2014) PET/MR in oncology: non–18F-FDG tracers for routine applications. J Nucl Med 55(Suppl 2):25S–31S
5. Kim TJ, Türkcan S, Pratx G (2017) Modular low-light microscope for imaging cellular bioluminescence and radioluminescence. Nat Protoc 12(5):1055–1076

Chapter 4

Use of $^{13}C_3{}^{15}N_1$-Serine or $^{13}C_5{}^{15}N_1$-Methionine for Studying Methylation Dynamics in Cancer Cell Metabolism and Epigenetics

Alice C. Newman, Christiaan F. Labuschagne, Karen H. Vousden, and Oliver D. K. Maddocks

Abstract

Tracing the fate of carbon-13 (^{13}C) labeled metabolites within cells by liquid chromatography mass spectrometry (LCMS) is a powerful analytical technique used for many years in the study of cell metabolism. Conventional experiments using LCMS and labeled nutrients tend to track the incorporation of ^{13}C from exogenous nutrients (such as amino acids) into other, relatively proximal, cellular metabolites. Several labs have extended this technique to track transfer of ^{13}C from the metabolite pool onto macromolecules, such as DNA, where methylation acts as an important functional modification. Here we describe a complete method that integrates previously established techniques to simultaneously track the use of ^{13}C-serine or ^{13}C-methionine into metabolite pools of the methionine cycle and into methylation of DNA and RNA. Given the ability to track methyl-transfer in a time-dependent way, this technique can provide temporal information about active methyl-transfer as well as quantification of total DNA/RNA methylation levels.

Key words Methylation, Methyl-transfer, DNA, RNA, Flux, Carbon-13, One-carbon metabolism, Methionine, Serine, Liquid chromatography mass spectrometry

1 Introduction

The generation and transfer of single carbons (referred to as one-carbon units/methyl groups) is a vital process in mammalian cell metabolism. One-carbon utilization can be crudely divided into two processes: (1) biosynthetic reactions, such as nucleotide synthesis, where one-carbons are used to build new molecules and one-carbon transfer is primarily mediated by the tetrahydrofolate (THF) cycle, and (2) functional modification reactions, such as DNA methylation, where one-carbons are transferred onto complete molecules in order to modify their function, primarily mediated by the methionine/S-adenosyl methionine (SAM) cycle [1, 2]. Competent one-carbon metabolism is essential for normal

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_4,

embryogenesis, growth, development, and maintenance of the epigenome, and perturbations of one-carbon metabolism are commonly implicated in diseases, including cancer. Major efforts have therefore been invested in understanding the inputs and outputs of the methionine cycle and the potential interaction between THF-mediated one-carbon metabolism and methionine cycle-mediated methylation.

Standard assays using cells fed ^{13}C labeled nutrients allow the interconversion of one metabolite to another to be evaluated by LCMS. Several groups have taken this technique a step further, following the fate of nutrient-derived carbons beyond the metabolome and into DNA [3–5]. By integrating a variety of previously established techniques, including a method using acid hydrolysis [6], we have developed an assay that allows the fate of amino acid-derived carbons (^{13}C) to be tracked through the methionine cycle and into the methylation of DNA and RNA [7] (Fig. 1a, b). Whereas epigenetic methylation was traditionally viewed as a relatively stable modification (encoding heritable epigenetic information), recent work has highlighted that DNA methylation can be highly dynamic in some circumstances [8]. Given the ability to vary the time during which cells are exposed to ^{13}C-labeled amino acids, it is possible to gain insight into the dynamics of methyl-transfer, a major strength of this approach.

Furthermore, the methods used here can be used as a simple assay to quantify total levels of cellular DNA or RNA methylation (without the need for ^{13}C labeling). Compared to other basic techniques for assessing total DNA methylation in cells (e.g., antibodies against 5-methylcytosine, 5mC), this LCMS approach is intrinsically normalized for total cytosine levels. This is a major benefit, as without accurate normalization for total cytosine/DNA levels it is not possible to determine if changes in 5mC staining represent a specific change in DNA methylation or just a change in the total level of DNA present.

2 Materials

2.1 Materials List

1. Cell line(s) of choice, e.g., HCT116
2. Standard cell culture medium, e.g., DMEM (2 mM L-glutamine, 10% FBS)
3. Assay medium; L-methionine/L-serine-free cell culture medium
4. Phosphate buffered saline (PBS)
5. Trypsin cell dissociation solution
6. Mass spectrometry grade water
7. Mass spectrometry grade acetonitrile

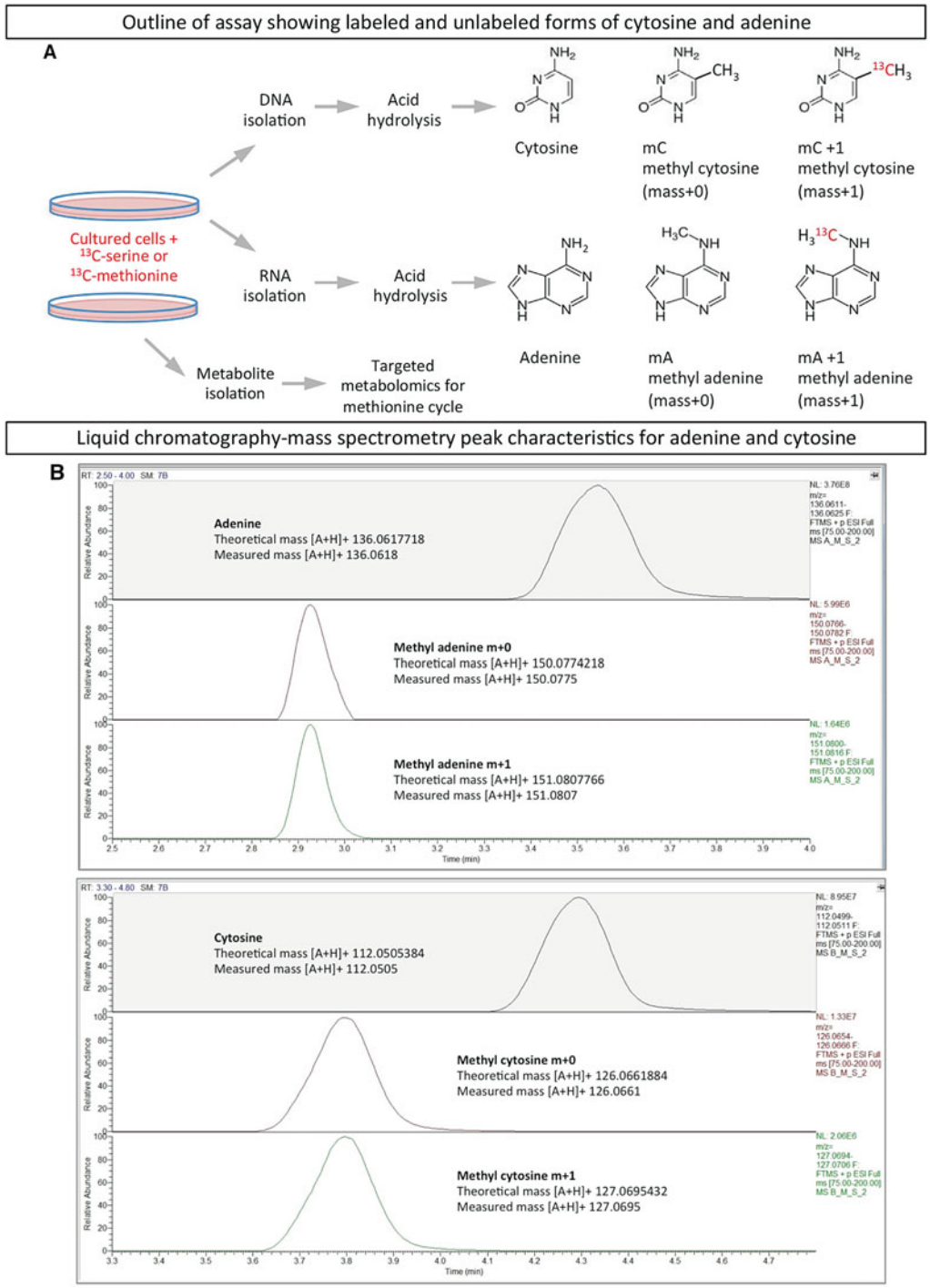


Fig. 1 (**a**) Cells are cultured in the presence of ^{13}C-serine or ^{13}C-methionine. DNA, RNA, and metabolites can be separately extracted and analyzed for labeling. The presence of ^{13}C labeling in cytosine and adenine (within 3 h) indicates methyl-transfer via the methionine cycle. Metabolite abundance and labeling patterns in the methionine cycle can assist in interpretation of changes in DNA/RNA methylation, for example, whether there is a change in levels of the methyl-donor SAM. (**b**) Example chromatogram peaks, including simulated and measured masses, for the cytosine and adenine peaks shown in (**a**). Images adapted from [7]

8. Mass spectrometry grade methanol
9. 6-well cell culture plates
10. Cell scrapers
11. $^{13}C_3{}^{15}N_1$-serine made up as a sterile filtered stock (100 mM) solution in PBS (*see* **Note 1**)
12. $^{13}C_5{}^{15}N_1$-methionine made up as a sterile filtered stock (100 mM) solution in PBS
13. L-serine
14. L-methionine
15. L-homocysteine
16. 5-azacytidine (optional control)
17. Wet ice
18. Standard 1.5 mL sample tubes
19. Mass spectrometry sample vials
20. DNA isolation kit with RNase digest step
21. RNA isolation kit with DNase digest step
22. Formic acid
23. High-temperature heat block with nitrogen blow-down dry evaporator
24. Locking 1.5 mL tubes
25. Personal safety equipment—gloves, goggles, lab coat
26. Fume hood
27. Method to quantify DNA and RNA (e.g., Nanodrop)
28. Hemocytometer or other cell counting methods (e.g., CASY cell counter)
29. Vortex mixer

2.2 Assay Media

1. Start with a basal medium that is analogous to standard complete medium (i.e., the normal medium used to grow your cells) but lacks methionine and serine, "MS medium." This can be purchased or formulated.
2. This example assay requires three types of media, made by supplementing – MS medium with:
   (a) 0.1 mM $^{13}C_5{}^{15}N_1$-methionine + 0.4 mM unlabeled serine
   (b) 0.1 mM unlabeled methionine + 0.4 mM $^{13}C_3{}^{15}N_1$-serine
   (c) 0.1 mM unlabeled methionine + 0.4 mM unlabeled serine

3 Methods

3.1 Cell Culture Setup

1. Plan the experimental conditions and plate layouts carefully and well in advance of the experiment (Fig 2a, b).
2. Each different condition to be assessed requires three sets of replicate wells (*see* Fig 2a):
   (a) One set is for metabolite extraction. This requires three replicate wells per condition, seeded as half of a 6-well plate, with the remaining 3 wells left empty.
   (b) One set (2–3 replicate wells per condition) is used for cell counts to allow normalization of lysis solvent volume.
   (c) One set is for DNA/RNA extraction. This requires 3 replicate wells per condition, seeded in 6-well plates.
3. Cells are seeded in complete medium and left for 24–48 h before the experiment is started.
4. The desirable number of cells per well depends on the cell type and the experimental conditions, aiming for 1–2 million cells per well (at the time of lysis) as a guide (*see* **Note 2**).

3.2 Treatment Period (Optional)

1. If you wish to assess the effect of a treatment (e.g., drug/starvation, etc.) and/or control treatment (e.g., 5-azacytidine, *see* **Note 3**) on methyl-transfer, then cells can be treated at this point. During this period you do not need to include any labeled metabolites.
2. If comparing different cell types or isogenic cells where a gene of interest is manipulated, then a treatment intervention may not be necessary.

3.3 Add Assay Medium

1. Wash all wells with 2 mL PBS.
2. Add pre-prepared assay media (*see* Subheading 3.1, **step 2**), 2 mL per well.
3. Incubate cells for 3 h (*see* **Note 4**).

3.4 Cell Count and Normalization for Metabolite Extraction

1. Approximately 30 min prior to metabolite extraction, begin the cell counting process.
2. For the wells designated for counting, aspirate the media and wash with 2 mL PBS per well.
3. Add trypsin cell dissociation reagent, and incubate at 37 °C until cells are fully detached.
4. Resuspend cells in culture medium to final volume of 1 mL per well.
5. Ensure cells are fully dissociated into single-cell suspension using a P1000 pipette.

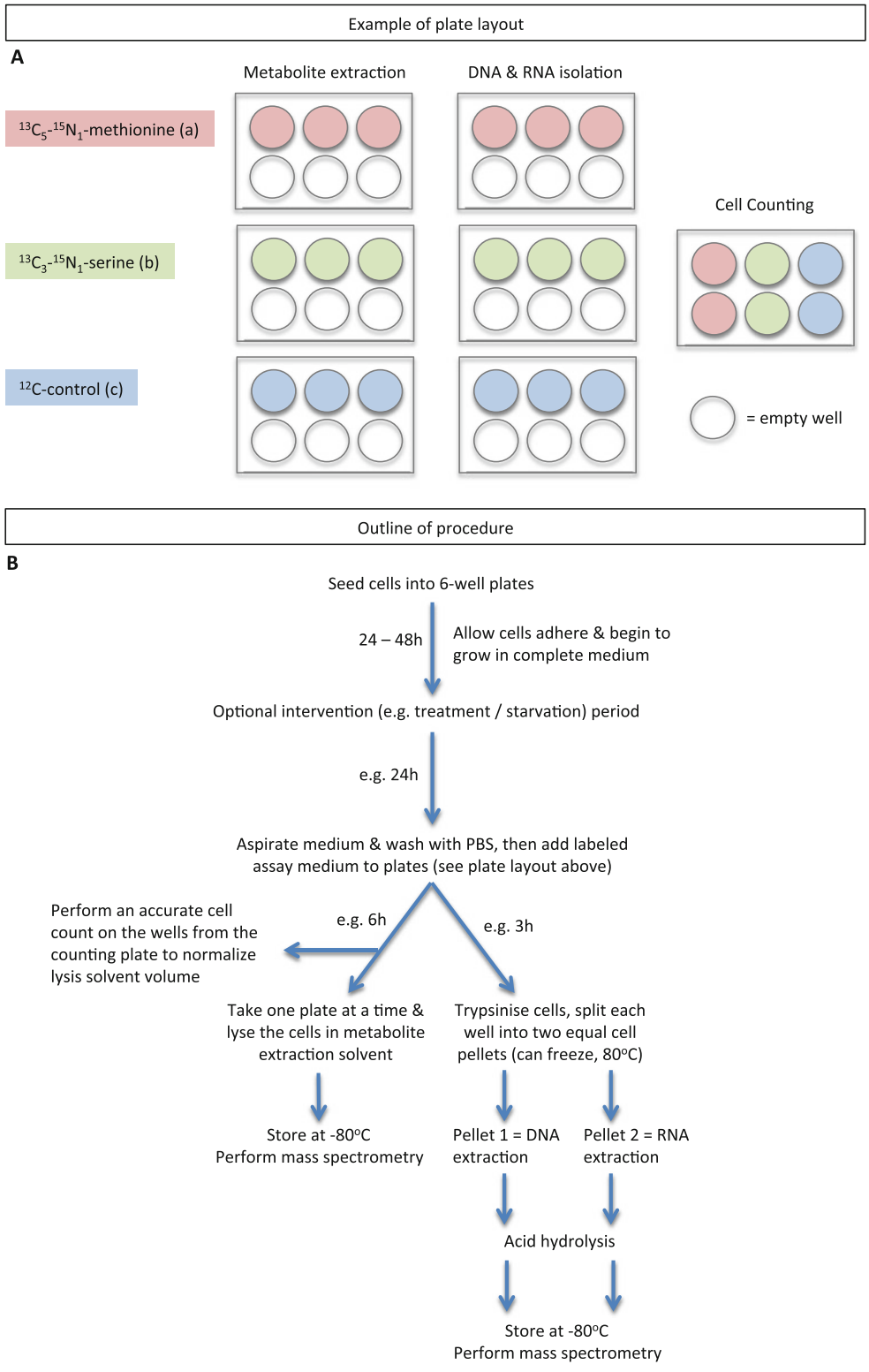


Fig. 2 (**a**) Example of experimental plate layout (using 6-well plates). In this example ^{13}C labeled versions of methionine and serine have been used, but the assay can also be done with ^{13}C-methionine only. An unlabeled (^{12}C) control is essential to establish the background levels of natural ^{13}C present. (**b**) Outline of experimental procedures showing example time courses

6. Perform an accurate cell count (e.g., using a hemocytometer or CASY cell counter) to determine the number of cells per well in each condition.
7. Calculate the volume of lysis solvent required to achieve 2×10^6 cells/mL for each experimental condition, e.g., if cell count indicates 1×10^6 cells per well, lyse cells in 500 μL; if 1.5×10^6 cells per well, lyse cells in 750 μL; etc. (*see* **Note 2**).

3.5 Metabolite Extraction

1. Prepare wet ice, 1.5 mL sample tubes (pre-labeled), cell scrapers, P1000 & tips, PBS, container for waste medium/PBS, list of pre-calculated lysis buffer volumes for each condition, and ensure lysis solvent is pre-cooled (−20 °C freezer).
2. Lyse each plate (i.e., 3 wells) individually; it takes approximately 5 min for each plate.
3. After 6 h labeling, take the plate from the incubator.
4. To remove the medium, tip the plate out into a suitable waste container; do this by hand, as quickly as possible.
5. Pour excess PBS into each well, and then tip the PBS out into a suitable waste container; do this by hand as quickly as possible.
6. Quickly remove residual PBS from each well using an aspirator or P1000 pipette, and place the plate on ice.
7. To each well add the pre-calculated volume of cold lysis solvent using P1000 pipette.
8. Briefly tilt the plate to ensure all cells are submerged in solvent.
9. Use a cell scraper to scrape all of the cells into the solvent; do this for all three wells (*see* **Note 5**).
10. Using a P1000 pipette, quickly homogenize the lysate and transfer the contents of each well into a pre-labeled 1.5 mL tube kept on wet ice.
11. Perform the lysis process until all samples are on wet ice.
12. Vortex all samples for 1 min.
13. Allow samples to sit on wet ice for further 10 min.
14. Centrifuge samples at top speed (e.g., 14,000–18,000 × *g*) at 4 °C for 15 min.
15. Transfer cleared supernatant to clean pre-labeled 1.5 mL tubes, and discard pellets.
16. The samples can be stored at −80 °C until being prepared for mass spectrometry.
17. Any mass spectrometry method that reliably detects the metabolites of interest and their labeled isotopomers can be applied, for example, *see* [7].

3.6 DNA and RNA Isolation

1. After 3 h labeling, aspirate the media and wash with 2 mL PBS per well.
2. Add trypsin cell dissociation reagent, and incubate at 37 °C until cells are fully detached.
3. Resuspend the cells in PBS, and split the cell suspension from each well into two pre-labeled 1.5 mL tubes (one for DNA, one for RNA); depending on cell size and capacity of DNA and RNA protocol, aim for 0.5×10^6 to 1×10^6 cells per tube.
4. Centrifuge the tubes to form cell pellets, and then aspirate the PBS supernatant.
5. The cell pellets can be immediately used for DNA/RNA isolation or stored at −80 °C.
6. Extract DNA and RNA from the cell pellets using standard methods (e.g., column-based kits), for DNA isolation include RNase treatment, and for RNA isolation include DNase treatment.
7. Elute DNA and RNA in nuclease-free water; a DNA/RNA concentration of ~50 ng/μL is ideal. Samples can be used immediately or stored at −80 °C.

3.7 Acid Hydrolysis of DNA and RNA

1. Health and safety considerations are critical when performing acid hydrolysis (*see* **Notes 6** and **7**).
2. Determine the concentration of the DNA and RNA solutions.
3. Aliquot the appropriate volume for 1 μg of each DNA sample and 3 μg of each RNA sample into clean pre-labeled locking 1.5 mL tubes, and spin samples into base of tubes.
4. Place the samples into a heat block (lids off) with blow-down nitrogen dryer, and dry at 40 °C until all moisture is removed.
5. Add 100 μL of formic acid to the dried samples, seal tubes, quickly vortex, and incubate tubes at 130–150 °C for 3.5 h (*see* **Note 6**).
6. Turn down the heat block to 40 °C, and allow tubes to fully cool before handling tubes; there is a risk of tubes popping open, releasing hot acid, if they are not allowed to cool.
7. Once the tubes have cooled to 40 °C, remove from heat block, ensure the lids are tightly secured, quickly vortex, and spin to collect solution in base of tubes.
8. Place the samples back into the heat block (lids off) with blow-down nitrogen dryer, and dry at 40 °C until all moisture is removed.
9. Once dry, add 25 μL of LCMS grade water to each tube, quickly vortex and spin to collect solution in base of tubes, and leave tubes at room temperature for 20 min.

10. Add 100 μL of an ice-cold solution of LCMS grade methanol (62.5%) and acetonitrile (37.5%), quickly vortex samples, and centrifuge at full speed 4 °C for 15 min.
11. Transfer supernatant directly to an LCMS vial, and samples can be stored at −80 °C until LCMS analysis.
12. Any mass spectrometry method that reliably detects the metabolites of interest and their labeled isotopomers can be applied (Fig. 1a), for example, using a Thermo Q Exactive (*see* [7]).

3.8 Data Analysis; DNA and RNA

1. To assess methyl-transfer to DNA, the ratio of m+1 methylcytosine to m+0 methylcytosine should be calculated; for RNA the ratio of m+1 methyladenine to m+0 methyladenine should be calculated (*see* Fig. 3a–c).

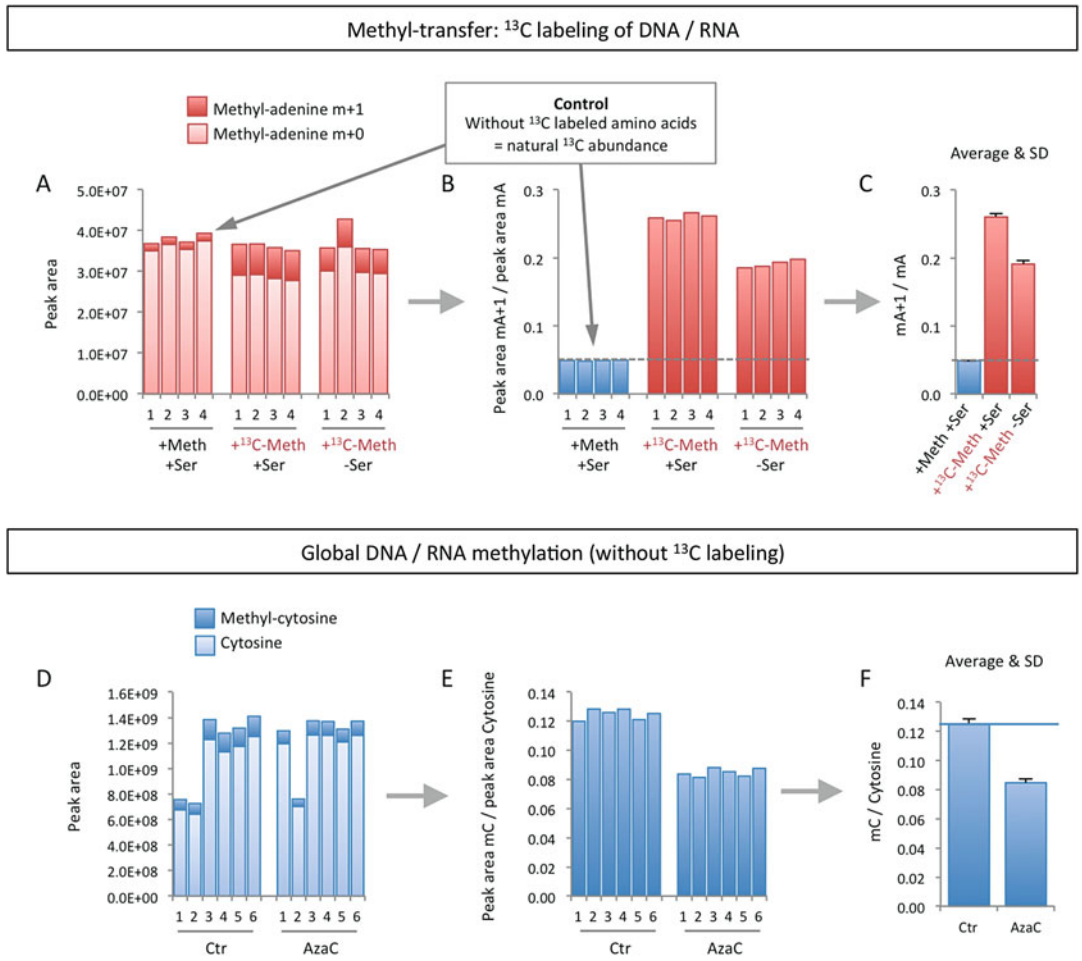


Fig. 3 Methyl-transfer example showing data from ^{13}C labeled RNA; example plots of raw peak areas (**a**), ratios of labeled and unlabeled peaks (**b**), and presentation of final average data (**c**). Total DNA methylation example showing data from unlabeled DNA; example plots of raw peak areas (**d**), ratios of methylcytosine to cytosine (**e**), and presentation of final average data (**f**). Images adapted from [7]

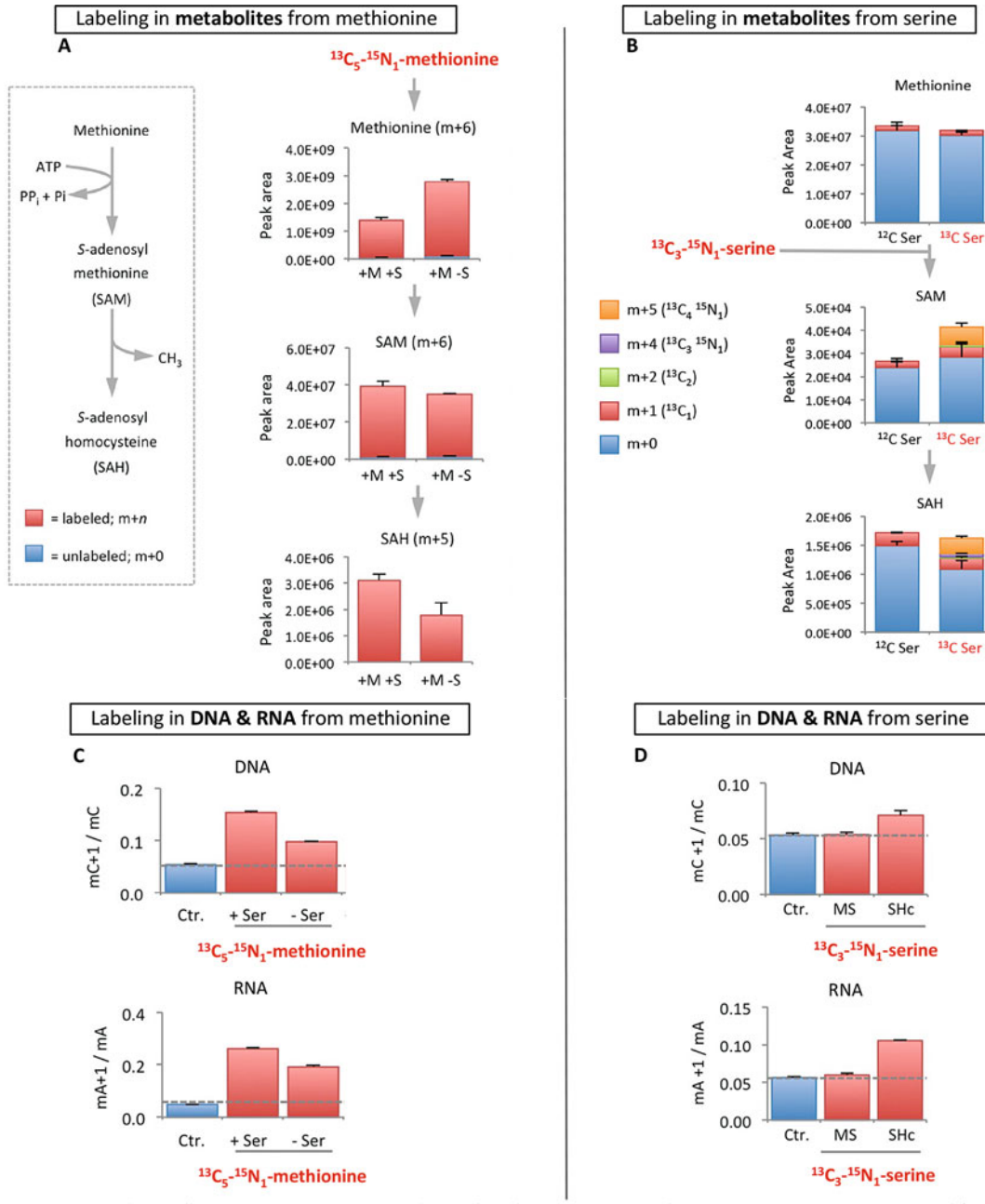
2. The ^{12}C control takes into account the background level of naturally occurring ^{13}C (i.e., an m+1 peak will be seen even under label-free conditions). This background ratio is used as a baseline and should be constant across independent experiments.
3. If wishing to assess total DNA methylation levels, it is not necessary to use labeled metabolites, and in this case the ratio of methylcytosine to cytosine should be calculated (*see* Fig. 3d–f). For total methylation levels in RNA, calculate the ratio of methyladenine to adenine.
4. When interpreting the data, it is important to bear in mind that perturbing cells (e.g., by drug treatment or gene knockout) can influence basic cell properties – such as the rate of growth and proliferation—which may have a generalized effect on DNA/RNA methylation and/or synthesis, rather than an effect that is specific to the treatment itself. For this reason it is important to interpret the results of this assay in the context of wider phenotypic data, and where possible make comparisons between conditions where cell growth/proliferation rates are similar.

3.9 Data Analysis: Metabolites

1. Standard approaches for targeted metabolomics can be applied, for example, *see* [7] and also *see* **Note 8**. Figure 4a, b gives examples of how targeted metabolomics can be reported.

4 Notes

1. We chose to use fully carbon (^{13}C) and nitrogen (^{15}N) labeled serine and methionine. We find these versions give good signals for detection of labeling in metabolites (e.g., of serine labeling in SAM, which can include the nitrogen) and of DNA/RNA methylation. However, other labeled forms can also be used, e.g., those without nitrogen labeling. If analyzing DNA/RNA methylation, it is critical to ensure that the relevant carbon that is transferred via SAM for methylation reactions is labeled.
2. Optimal cell number depends on the size of the cells being used. This has an impact both on the confluence of the cell monolayer and the preferable concentration (cells/mL) at which the cells are lysed. Using cells like HCT116, we usually aim for cell confluence of approximately 50–75% at time of lysis, which translates into an easily manageable volume of metabolite extraction solvent at 2×10^6 cells/mL, e.g., 500–1000 μL. However, if using larger cells, then the lysis concentration can be decreased, e.g., 5×10^5–1×10^6 cells/mL. However, it is *critical* that the same concentration is used



SAH = *S*-adenosylhomocysteine, SAM = *S*-adenosylmethionine, M = methionine, S = serine, Ctr. = control (no labeled metabolites added), Hc = homocysteine, mA = methyl-adenine, mC = methyl-cytosine, '+1' = mass+1 labeled form.

Fig. 4 Examples of real data generated using the complete method. (**a**) Shows targeted metabolomics of the methionine cycle after $^{13}C_5{}^{15}N_1$-methionine labeling. (**b**) Shows targeted metabolomics of the methionine cycle after $^{13}C_3{}^{15}N_1$-serine labeling. (**c**) Shows the corresponding data for DNA and RNA labeling after $^{13}C_5{}^{15}N_1$-methionine labeling (under serine-fed or serine-starved conditions). (**d**) Shows data for DNA and RNA labeling using $^{13}C_3{}^{15}N_1$-serine, which is only revealed under methionine-starved, homocysteine-fed conditions. Images adapted from [7]; for specific experimental conditions, *see* [7]

across all samples in a given experiment. The ideal volume of lysis solvent used should be enough to comfortably cover the base of the well (>300 μL), but not so large as to be difficult to handle with a P1000 pipette (ideally <1 mL). If a volume larger than 1 mL per well is required, then initially add 1 mL per well, scrape the cells, and then top-up with the additional solvent after scraping. This ensures each well is lysed quickly.

3. 5-Azacytidine (e.g., 0.5 μM) treatment is a good control for causing a decrease in methyl-transfer to DNA and a decrease in total DNA methylation levels.
4. Longer or shorter incubation periods with the labeled serine/methionine can be used. We chose 3 h because it is long enough for carbon-13 to be transferred from serine/methionine, but not long enough for one-carbons from serine to enter de novo nucleotide synthesis via the THF-cycle: introduction of these carbons into the purine/pyrimidine structure would confound the ability to detect methylation when using serine. The same issue does not apply to methionine, as, unlike serine, its carbons are not used in de novo nucleotide synthesis.
5. When scraping the cells, a white precipitate will be observed; this is normal, caused by macromolecules such as proteins precipitating. The complete lysate including precipitate should be transferred to the 1.5 mL tube. The precipitate will be removed by the centrifugation step.
6. Full attention to health and safety procedures is very important when performing the acid hydrolysis step. Formic acid heated to 130 °C is a clear safety risk. Use of locking tubes decreases the chance of tubes popping open. The heating block should be placed inside a fume hood with the shield down, and personal protective equipment (gloves, lab coat, and goggles) should be worn at all times. To prevent tubes from opening during the high-temperature phase, we place multiple spare metal heat blocks on top of the tubes, so the weight keeps the tubes closed. It is important to allow the tubes to fully cool before removing from the heat block.
7. As an alternative to acid hydrolysis, enzymatic digestion of DNA has also been used to yield bases that can be analyzed by mass spectrometry (*see* [3]).
8. This protocol provides the possibility for three outputs: targeted metabolomics (focused on the methionine cycle), DNA methylation, and RNA methylation. The assay can be run to provide all three outputs from the same experiment or truncated to provide only one or two of the possible outputs. Splitting the assay (e.g., into one experiment for metabolomics and one for DNA and RNA) can make it easier to carry out.

Comparing the labeling and total metabolite levels within the methionine cycle—in particular methionine and SAM – with DNA and RNA labeling can give a more complete picture of changes in methylation than simply viewing DNA and RNA in isolation.

Acknowledgment

ODKM and TZ are funded by Cancer Research UK Career Development Fellowship C53309/A19702. KHV and CFL are funded by CRUK Grant C596/A10419.

References

1. Newman AC, Maddocks ODK (2017) Serine and functional metabolites in cancer. Trends Cell Biol 27(9):645–657. https://doi.org/10.1016/j.tcb.2017.05.001
2. Newman AC, Maddocks ODK (2017) One-carbon metabolism in cancer. Br J Cancer 116(12):1499–1504. https://doi.org/10.1038/bjc.2017.118
3. Bachman M, Uribe-Lewis S, Yang X, Williams M, Murrell A, Balasubramanian S (2014) 5-Hydroxymethylcytosine is a predominantly stable DNA modification. Nat Chem 6 (12):1049–1055. https://doi.org/10.1038/nchem.2064
4. Field MS, Kamynina E, Watkins D, Rosenblatt DS, Stover PJ (2015) Human mutations in methylenetetrahydrofolate dehydrogenase 1 impair nuclear de novo thymidylate biosynthesis. Proc Natl Acad Sci U S A 112(2):400–405. https://doi.org/10.1073/pnas.1414555112
5. Herbig K, Chiang EP, Lee LR, Hills J, Shane B, Stover PJ (2002) Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide and S-adenosylmethionine biosyntheses. J Biol Chem 277(41):38381–38389. https://doi.org/10.1074/jbc.M205000200
6. Kok RM, Smith DE, Barto R, Spijkerman AM, Teerlink T, Gellekink HJ, Jakobs C, Smulders YM (2007) Global DNA methylation measured by liquid chromatography-tandem mass spectrometry: analytical technique, reference values and determinants in healthy subjects. Clin Chem Lab Med 45(7):903–911. https://doi.org/10.1515/CCLM.2007.137
7. Maddocks OD, Labuschagne CF, Adams PD, Vousden KH (2016) Serine metabolism supports the methionine cycle and DNA/RNA methylation through de novo ATP synthesis in cancer cells. Mol Cell 61(2):210–221. https://doi.org/10.1016/j.molcel.2015.12.014
8. Bhutani N, Burns DM, Blau HM (2011) DNA demethylation dynamics. Cell 146(6):866–872. https://doi.org/10.1016/j.cell.2011.08.042

Chapter 5

Measurement of Mitochondrial Membrane Potential with the Fluorescent Dye Tetramethylrhodamine Methyl Ester (TMRM)

Sarah Creed and Matthew McKenzie

Abstract

The mitochondrial membrane potential ($\Delta\psi_m$) drives the generation of ATP by mitochondria. Interestingly, $\Delta\psi_m$ is higher in many cancer cells comparted to healthy noncancerous cell types, providing a unique metabolic marker. This feature has also been exploited for therapeutic use by utilizing drugs that specifically accumulate in the mitochondria of cancer cells with high $\Delta\psi_m$. As such, the assessment of $\Delta\psi_m$ can provide very useful information as to the metabolic state of a cancer cell, as well as its potential for malignancy. In addition, the measurement of $\Delta\psi_m$ can also be used to test the ability of novel anticancer therapies to disrupt mitochondrial metabolism and cause cell death.

Here, we outline two methods for assessing $\Delta\psi_m$ in cancer cells using confocal microscopy and the potentiometric fluorescent dye tetramethylrhodamine methyl ester (TMRM). In the first protocol, we describe a technique to quantitatively measure $\Delta\psi_m$, which can be used to compare $\Delta\psi_m$ between different cell types. In the second protocol, we describe a technique for assessing changes to $\Delta\psi_m$ over time, which can be used to determine the effectiveness of different therapeutic compounds or drugs in modulating mitochondrial function.

Key words Mitochondria, Membrane potential, TMRM, Cancer cells, Osteosarcoma, Confocal imaging, Fluorescence

1 Introduction

The mitochondrial membrane potential ($\Delta\psi_m$) is the main driving force for the generation of ATP by the mitochondria and consists of a proton gradient generated by the mitochondrial respiratory chain complexes. The reducing equivalents NADH and $FADH_2$, produced by the tricarboxylic acid cycle, donate their electrons to complex I (NADH-ubiquinone oxidoreductase, EC 1.6.5.3) and complex II (succinate-ubiquinone oxidoreductase, EC 1.3.5.1), respectively. These electrons pass through the respiratory chain, causing complexes I, III (ubiquinol-ferrocytochrome-*c* oxidoreductase, EC 1.10.2.2), and IV (ferrocytochrome-*c*: oxygen

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_5,

oxidoreductase, EC 1.9.3.1) to pump protons out of the mitochondrial matrix into the intermembrane space. This creates a potential across the inner membrane ($\Delta\psi_m$) in the order of 180 mV negative to the cytosol. The $\Delta\psi_m$ provides the driving force for proton influx through complex V (F_1F_o-ATP synthase, EC 3.6.3.14), which condenses ADP and inorganic phosphate to generate ATP [1].

A number of different fluorescent dyes can be used to assess $\Delta\psi_m$, all of which have different properties that determine their suitability for the type of analyses to be performed (for review *see* [2]). JC-1 (5,5′,6,6′-tetrachloro-1,1′,3,3′-tetraethylbenzimidazolylcarbocyanine iodide) can be used as a ratiometric $\Delta\psi_m$ indicator, with the accumulation of the dye within mitochondria resulting in aggregates with a different emission spectra to its monomeric form (which is found at low concentrations). However, this makes JC-1 signals very concentration sensitive, with dye fluorescence also affected by other factors including reactive oxygen species production [3]. $DiOC_6$ (3,3′-dihexyloxacarbocyanine iodide) is commonly used to measure $\Delta\psi_m$ during flow cytometry. However, due to its high mitochondrial toxicity, it has to be used at very low (<1 nM) concentrations, making it difficult to accurately measure $\Delta\psi_m$ [4].

Other lipophilic cationic dyes, such as tetramethylrhodamine methyl ester (TMRM), tetramethylrhodamine ethyl ester (TMRE), and rhodamine 123, are also very useful for measuring $\Delta\psi_m$. Due to their positive charge, these dyes accumulate within the mitochondria in an inverse proportion to $\Delta\psi_m$ according to the Nernst equation. While these dyes can be used in a "quenching" mode at high concentrations between ~1 and 20 μM, they can also be used in non-quenching mode at much lower concentrations to ensure mitochondrial function is not altered by the presence of the dye [5]. At 20 nM, no significant mitochondrial binding or electron transport chain inhibition by TMRM is apparent, making it suitable for detailed $\Delta\psi_m$ assessment. Furthermore, TMRM can be used simultaneously with other fluorescent dyes, for example, with Fluo-4 to measure calcium signals within the cell [6] or with chloromethyldihydrodichlorofluorescein diacetate (cmH_2DCF-DA) to assess reactive oxygen species generation [7].

Interestingly, in many cancer cells, $\Delta\psi_m$ is higher than normal and can be correlated with the malignant potential of the cancer [8]. Lung cancer cell lines, including A549, H446, and SPC, as well as MCF-7 breast cancer cells, have been shown to have higher $\Delta\psi_m$ compared to healthy, noncancerous cell types [9, 10]. Furthermore, cancer stem cells (CSCs), which are believed to drive the initiation and recurrence of malignant tumors, have even higher $\Delta\psi_m$ compared to cancer cells lacking stemness [11].

The elevated $\Delta\psi_m$ observed in cancer cells has led to alternative therapeutic strategies that target mitochondrial function.

Compounds such as MKT-077, a cyanine dye analogue that preferentially accumulates in tumor cells with high $\Delta\psi_m$, have been shown to disrupt mitochondrial metabolism and subsequently inhibit tumor growth [12]. Alternatively, some anticancer therapies utilize mitochondrial respiratory chain inhibitors to disrupt mitochondrial metabolism [13]. These inhibitors can diminish $\Delta\psi_m$, resulting in mitochondrial permeability transition and the induction of apoptotic cell death [14]. As such, the assessment of $\Delta\psi_m$ in cancer cells can provide useful information about tumorigenicity, as well as the effects of anticancer therapeutics that target mitochondrial function.

Here, we outline two methods for assessing $\Delta\psi_m$ in cancer cells using TMRM in non-quenching mode. Firstly, we describe a protocol for quantitatively measuring $\Delta\psi_m$ using the acquisition of z-stacks and thresholding the total TMRM signal and, secondly, the assessment of $\Delta\psi_m$ over time in the presence of mitochondrial inhibitors or uncouplers that modulate $\Delta\psi_m$.

2 Materials

1. DMEM medium: DMEM (high glucose, pyruvate, Glutamax) supplemented with 5% fetal bovine serum (FBS) and 1× penicillin/streptomycin.
2. Record Solution (RS): 109 mM NaCl, 50 mM KCl, 2 mM $MgSO_4$, 1.25 mM KH_2PO_4, 10 mM D-glucose, 2 mM $CaCl_2$, 10 mM HEPES. Adjust pH to 7.35 with 1 M NaOH. Store in aliquots at −20 °C (*see* **Note 1**).
3. 1× phosphate buffered saline (PBS): 1.54 mM KH_2PO_4, 155.17 mM NaCl, 2.71 mM Na_2HPO_4. $7H_2O$, pH 7.2.
4. Trypsin–EDTA: 0.25% (w/v) Trypsin/0.25% (w/v) ethylenediaminetetraacetic acid (EDTA) with phenol red.
5. Carbonyl cyanide p-trifluoromethoxyphenylhydrazone (FCCP): stock solution of 1 mM in 100% ethanol.
6. 2 μM tetramethylrhodamine, methyl ester, perchlorate (TMRM): stock solution of 10 mM in 100% methanol. From this, make a working stock of 2 μM in distilled H_2O.
7. 10 mM verapamil: stock solution in 100% ethanol (*see* **Note 2**).
8. 1 mM rotenone: stock solution in 100% ethanol.
9. 500 μg/mL oligomycin: stock solution in 100% ethanol.
10. 1 mg/mL Hoechst 33342: stock solution in distilled H_2O.
11. RS Staining Solution: 20 nM TMRM, 2 μg/mL Hoechst 33342, 10 μM verapamil (*see* **Note 3**).
12. RS Imaging Solution: 20 nM TMRM, 10 μM verapamil.
13. Sterile chambered glass coverslips for cell culture.

3 Methods

3.1 Preparation of Cells

1. Grow cells on 10 cm cell culture dishes or in 75 cm^2 flasks in DMEM at 37 °C/5% CO_2.
2. Harvest cells by aspirating medium and washing with 5 mL PBS. Remove PBS, and dissociate cells with Trypsin–EDTA and incubation at 37 °C/5% CO_2 for 2 min. Tap dish or flask gently to remove cells, and then resuspend in 5 mL culture media.
3. Count cells by adding 12 μL of resuspended cells onto a hemocytometer.
4. Plate cells in suitable dishes or chambered coverslips for confocal imaging so that ~80% confluency is achieved on the day of imaging. For example, 2 to 5 × 10^4 143B osteosarcoma cells can be plated into one well of an 8-well chamber slide one day before imaging.
5. Incubate cells overnight at 37 °C/5% CO_2 to allow cells to attach and recover.

3.2 Staining Cells with TMRM and Hoechst 33342

1. Remove culture media from cells by pipette, and wash with 100 μL PBS.
2. Incubate cells in 200 μL RS Staining Solution for 45 min at 37 °C/5% CO_2.
3. Remove RS Staining Solution, and wash cells twice with 100 μL PBS to remove excess TMRM and Hoechst 33342.
4. Add 200 μL RS Imaging Solution to each well. Cells are now ready for imaging.

3.3 Quantitative Measurement of $\Delta\psi_m$

1. Place dish or chambered coverslip with cells in RS Imaging Solution onto an inverted laser scanning confocal microscope equipped with an environmental chamber set to 37 °C and delivering 5% humidified CO_2. We routinely use a 60×/1.35NA oil immersion objective for the measurement of $\Delta\psi_m$; however, a 40× objective lens can also be used.
2. Set the excitation and emission spectra: excitation of Hoechst 33342 using a 405 nm diode laser and TMRM with a 543 nm He-Ne laser. Set laser power to 5% for both channels to minimize phototoxicity to cells.
3. Adjust the gain and offset settings on the photo multiplier tube (PMT) detectors. To ensure accuracy in intensity measurements between images, do not change these values once set.
4. The mitochondrial network should be clearly visible from the TMRM signal. Acquire a *z*-stack image of the cells. The thickness of each step through the *z*-plane should be kept consistent

for all images. We used a thickness of 0.5 μm for 143B osteosarcoma cells. The *z*-stack images can be used for downstream image processing for quantitative assessment of $\Delta\psi_m$ (*see* Subheading 3.5).

3.4 Assessment of $\Delta\psi_m$ Sensitivity to Mitochondrial Inhibitors or Drugs

1. Place dish or chambered coverslip with cells in RS Imaging Solution onto an inverted laser scanning confocal microscope. For assessing changes to $\Delta\psi_m$, we routinely use a 40× objective lens, which allows the imaging of multiple cells in the same field at a suitable resolution.
2. Set microscope to scan images every 20 s. Scan cells for 2–5 min to establish baseline readings for the TMRM signal.
3. Using a pipette, add inhibitor or drug of interest to the cells. We routinely use 100× stocks that allow simple addition of the compound directly to the cells during imaging. For example, we add a 1:100 dilution of 1 mM rotenone (final concentration of 10 μM) to assess the effects of NADH- ubiquinone oxidoreductase (complex I) inhibition or a 1:100 dilution of 500 μg/mL oligomycin (final concentration of 5 μg/mL) to assess the effects of F_1F_o-ATPase (complex V) inhibition.
4. Continue to collect images every 20 s for 20 min (or longer if required) to monitor any changes in $\Delta\psi_m$ over time.
5. Using a pipette, add a 1:100 dilution of 1 mM FCCP (final concentration of 10 μM) directly to the RS Imaging Solution. Restart imaging every 20 s for 5 min. FCCP is a protonophore that uncouples mitochondrial oxidative phosphorylation, resulting in the complete dissipation of $\Delta\psi_m$. This results in a reduction of the mitochondrial TMRM signal, with the TMRM equilibrating evenly throughout the cell and the extracellular RS Imaging Solution.
6. The image series can now be analyzed for changes to $\Delta\psi_m$ over time (*see* Subheading 3.5).

3.5 Image Analysis

1. Once imaging is complete, determine the intensity of the TMRM signals using image analysis software. We routinely use ImageJ software (https://imagej.nih.gov/ij/) with the Bio-Formats plugin (download "bioformats_package.jar" from http://downloads.openmicroscopy.org/bio-formats/, and save it in the "C:\Program Files\ImageJ\plugins" folder).
2. To measure $\Delta\psi_m$, open the *z*-stack image file (e.g., a .roi file) as a Hyperstack using "File>Open" or the "Plugins>Bio-Formats>Bio-Formats Importer" function. Select "split channels" when opening the file so that the TMRM channel can be analyzed separately.
3. Select the TMRM channel, and compress the *z*-stack into one image by selecting "Image>Stacks>Z project." For

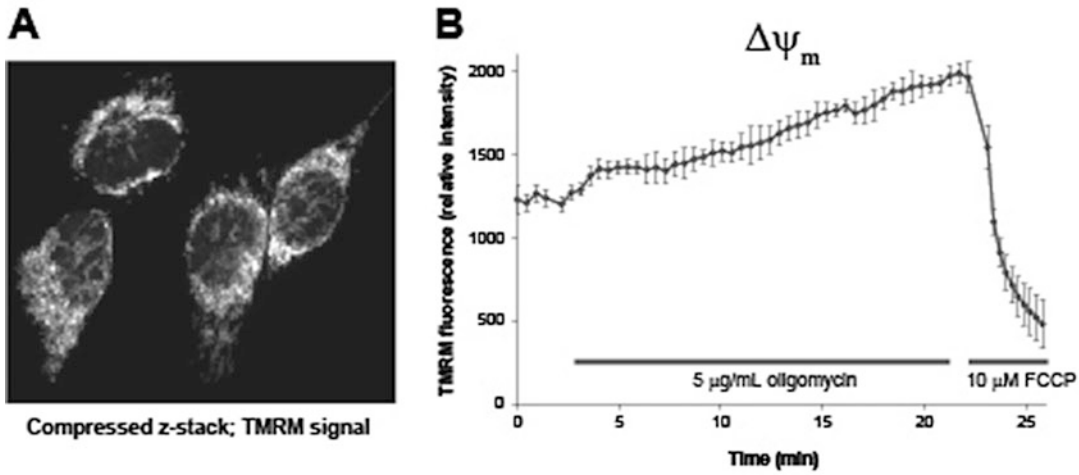


Fig. 1 Quantitative assessment of $\Delta\psi_{m.}$ (**a**) A compressed *z*-stack image showing 143B osteosarcoma cells stained with 20 nM TMRM. Each image slice was acquired using a scanning laser confocal microscope with a 60×/1.35NA oil immersion objective. (**b**) Analysis of $\Delta\psi_m$ over time. TMRM fluorescence increases slowly with the addition of oligomycin as the $\Delta\psi_m$ hyperpolarizes due to the inhibition of the F_1F_o-ATPase. Addition of the protonophore FCCP depolarizes $\Delta\psi_m$, resulting in a decrease in TMRM fluorescence and the uncoupling of oxidative phosphorylation. Data shown is mean ± SD, $n = 5$

"Projection Type," select "Sum Slices." You should now have a single, combined image of the *z*-stack (Fig. 1a).

4. To measure the intensity of the TMRM signal, select "Image>Adjust>Threshold." Check the "Dark background" box, and choose the best fit algorithm from the drop-down menu (or "Default") and "red" for threshold. Adjust the top slider until all of the TMRM signal is selected with the red threshold tool. Alternatively, the "Auto" threshold option can be used (*see* **Note 4**).
5. To obtain the mean intensity values for the TMRM signal, open the "Analyze>Set Measurements" box, and select the required variables, such as "Area," "Min & max gray value," and "Mean gray value." Also ensure that "Limit to threshold" is checked so that only the TMRM signal under the red threshold is measured. Next select "Analyze>Measure" (Ctrl+M) to obtain mean values for the whole image field. Alternatively, regions of interest (ROIs) containing specific cell/s can be selected.
6. To monitor how $\Delta\psi_m$ changes over time due to inhibitor or drug treatment, open the image file (e.g., an .oif file), click on the selection tool, and select an ROI.
7. Open "Analyze>Tools>ROI Manager," and click "Add" to include the selected ROI for intensity measurement. Multiple ROIs can be selected on a single image for measurement. Repeat previous steps until all ROIs have been added to ROI Manager.

8. Click "More>Multi Measure," make sure that "Measure all slices" is selected and that "One Row Per Slice" is deselected, then click "OK" to obtain a table of mean fluorescent intensity of the TMRM signal for each ROI at each time point.
9. Use data from all ROIs to calculate the average intensities and standard deviation for the TMRM signal. For example, oligomycin treatment causes TMRM fluorescence to increase slowly over time as $\Delta\psi_m$ hyperpolarizes due to the inhibition of the F_1F_o-ATPase (Fig. 1b).

4 Notes

1. The concentration of TMRM within the mitochondria is directly proportional to the cytosolic concentration. As such, if the plasma membrane depolarizes, the cytosolic TMRM concentration will be reduced, with a concomitant reduction in mitochondrial TMRM signal (without any actual change to $\Delta\psi_m$). To eliminate this effect, a high K^+ concentration is used in the Record Solution to depolarize the plasma membrane [15]. If plasma membrane potential is not an issue, standard isotonic Record Solution with normal K^+ concentration (156 mM NaCl, 3 mM KCl, 2 mM $MgSO_4$, 1.25 mM KH_2PO_4, 10 mM D-glucose, 2 mM $CaCl_2$, 10 mM HEPES, pH 7.35), or Hank's Balanced Salt Solution, can be used.
2. Verapamil is added to stop the export of TMRM out of the cancer cell by the plasma membrane multidrug transporter.
3. Hoechst 33342 staining of the nucleus is useful for locating the cells within the imaging field but can be left out if preferred.
4. When setting the threshold during image analysis, do not hit "Apply," as this will create a binomial mask with only two values of 0 and 255 for the signal intensity.

Acknowledgment

This work was supported by the Monash Health Translational Precinct (MHTP) Micro Imaging Platform and the Victorian Government's Operational Infrastructure Support Program.

References

1. Wallace DC (1999) Mitochondrial diseases in man and mouse. Science 283:1482–1488
2. Perry SW, Norman JP, Barbieri J, Brown EB, Gelbard HA (2011) Mitochondrial membrane potential probes and the proton gradient: a practical usage guide. BioTechniques 50:98–115
3. Chinopoulos C, Tretter L, Adam-Vizi V (1999) Depolarization of in situ mitochondria due to hydrogen peroxide-induced oxidative

stress in nerve terminals: inhibition of alpha-ketoglutarate dehydrogenase. J Neurochem 73:220–228

4. Rottenberg H, Wu S (1998) Quantitative assay by flow cytometry of the mitochondrial membrane potential in intact cells. Biochim Biophys Acta 1404:393–404
5. McKenzie M, Liolitsa D, Akinshina N, Campanella M, Sisodiya S, Hargreaves I, Nirmalananthan N, Sweeney MG, Abou-Sleiman PM, Wood NW, Hanna MG, Duchen MR (2007) Mitochondrial ND5 Gene Variation Associated with Encephalomyopathy and Mitochondrial ATP Consumption. J Biol Chem 282:36845–36852
6. McKenzie M, Duchen MR (2016) Impaired Cellular Bioenergetics Causes Mitochondrial Calcium Handling Defects in MT-ND5 Mutant Cybrids. PLoS One 11:e0154371
7. Zhang X, Lemasters JJ (2013) Translocation of iron from lysosomes to mitochondria during ischemia predisposes to injury after reperfusion in rat hepatocytes. Free Radic Biol Med 63:243–253
8. Zhang BB, Wang DG, Guo FF, Xuan C (2015) Mitochondrial membrane potential and reactive oxygen species in cancer stem cells. Fam Cancer 14:19–23
9. Ye XQ, Li Q, Wang GH, Sun FF, Huang GJ, Bian XW, Yu SC, Qian GS (2011) Mitochondrial and energy metabolism-related properties as novel indicators of lung cancer stem cells. Int J Cancer 129:820–831
10. Bonnet S, Archer SL, Allalunis-Turner J, Haromy A, Beaulieu C, Thompson R, Lee CT, Lopaschuk GD, Puttagunta L, Harry G, Hashimoto K, Porter CJ, Andrade MA, Thebaud B, Michelakis ED (2007) A mitochondria-K+ channel axis is suppressed in cancer and its normalization promotes apoptosis and inhibits cancer growth. Cancer Cell 11:37–51
11. Lamb R, Ozsvari B, Bonuccelli G, Smith DL, Pestell RG, Martinez-Outschoorn UE, Clarke RB, Sotgia F, Lisanti MP (2015) Dissecting tumor metabolic heterogeneity: Telomerase and large cell size metabolically define a sub-population of stem-like, mitochondrial-rich, cancer cells. Oncotarget 6:21892–21905
12. Britten CD, Rowinsky EK, Baker SD, Weiss GR, Smith L, Stephenson J, Rothenberg M, Smetzer L, Cramer J, Collins W, Von Hoff DD, Eckhardt SG (2000) A phase I and pharmacokinetic study of the mitochondrial-specific rhodacyanine dye analog MKT 077. Clin Cancer Res 6:42–49
13. Lo-Coco F, Avvisati G, Vignetti M, Thiede C, Orlando SM, Iacobelli S, Ferrara F, Fazi P, Cicconi L, Di Bona E, Specchia G, Sica S, Divona M, Levis A, Fiedler W, Cerqui E, Breccia M, Fioritoni G, Salih HR, Cazzola M, Melillo L, Carella AM, Brandts CH, Morra E, von Lilienfeld-Toal M, Hertenstein B, Wattad M, Lubbert M, Hanel M, Schmitz N, Link H, Kropp MG, Rambaldi A, La Nasa G, Luppi M, Ciceri F, Finizio O, Venditti A, Fabbiano F, Dohner K, Sauer M, Ganser A, Amadori S, Mandelli F, Dohner H, Ehninger G, Schlenk RF, Platzbecker U (2013) Retinoic acid and arsenic trioxide for acute promyelocytic leukemia. N Engl J Med 369:111–121
14. Lim SC, Carey KT, McKenzie M (2015) Anti-Cancer analogues ME-143 and ME-344 exert toxicity by directly inhibiting mitochondrial NADH:ubiquinone oxidoreductase (Complex I). Am J Cancer Res 5:689–701
15. Duchen MR, Surin A, Jacobson J (2003) Imaging mitochondrial function in intact cells. Methods Enzymol 361:353–389

Chapter 6

Assessment of Stabilization and Activity of the HIFs Important for Hypoxia-Induced Signalling in Cancer Cells

David Kung-Chun Chiu, Misty Shuo Zhang, Aki Pui-Wah Tse, and Carmen Chak-Lui Wong

Abstract

Blood vessels in tumors contain chaotic branching structures and leaky vessel lumens, resulting in uneven supply of oxygen in the tumor microenvironment. High metabolic and proliferation rate of tumor cells further depletes the local oxygen supply. Therefore, hypoxia is a common phenomenon in multiple solid malignancies. Hypoxia-inducible factors (HIFs) regulate the transcription of a spectrum of genes, which are vitally important for tumor cell adaption under hypoxia, and shape the tumor microenvironment to become more favorable for progression. HIFs are involved in almost every step of cancer development through inducing angiogenesis, metabolic reprogramming, metastasis, cancer stemness maintenance, chemoresistance, and immune evasion. Here, we describe methods for the assessment of HIF activity, as well as identification of novel transcriptional targets of HIFs in vitro and in vivo.

Key words Hypoxia, HIF, qPCR, Western blotting, Luciferase reporter assay, ChIP assay, Immunohistochemistry

1 Introduction

Hypoxia, decrease of oxygen (O_2) supply, is one of the most widely observed hallmarks of cancer. Inadequate oxygenation in the tumor is mainly caused by the uncontrolled duplication of cancer cells which rapidly expands the tumorous regions, which are not accessible to functional blood vessels. Moreover, cancer cells have high demand for O_2 due to their robust metabolic requirements, leading to a fast consumption of O_2 from the tumor microenvironment. The primary molecular adaptation system in hypoxia is orchestrated by hypoxia-inducible factors (HIFs). HIFs are heterodimers consisting of the constitutively stable HIF-s1β subunit and the oxygen-sensitive HIF-1/2α subunits [1]. In well-oxygenated conditions,

David Kung-Chun Chiu, Misty Shuo Zhang, and Aki Pui-Wah Tse contributed equally to this work.

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_6,

HIF-1/2α subunits are being hydroxylated at specific proline residues using O_2 and α-ketoglutarate as the substrates by prolyl hydroxylases (PHDs) [2]. Hydroxylated HIF-1/2α binds to von Hippel-Lindau (VHL) which recruits E3 ubiquitin ligase to conjugate HIF-1/2α with ubiquitin for proteasomal degradation [3]. In poorly oxygenated conditions, inhibition of hydroxylation enables the stabilization of HIF-1/2α, which then dimerizes with HIF-1β, together with other transcriptional coactivators such as p300/CBP. The complexes then bind to the genes with DNA consensus sequence 5′[A/G]CGTG-3′, also called the hypoxia-responsive elements (HREs), to turn on gene transcription [4]. The interaction of HIF-1/2α and coactivators and therefore the transactivation activity are inhibited by factor-inhibiting HIF (FIH) which hydroxylates specific asparagine residues of HIF-1/2α. Given the high occurrence of hypoxia in cancer, HIFs are frequently expressed in cancers and often correlated with poor clinical outcomes [5]. Loss of functions due to mutations in VHL, succinate dehydrogenase B (SDH-B), SDH-C, SDH-D, and fumarate hydratase (FH) prevent the ubiquitin-mediated proteasomal degradation of HIF-1/2α, thereby resulting in the stabilization of HIF even in well-oxygenated conditions in some cancer types [6–10]. Loss of function in tumor suppressors such as phosphatase and tensin homolog (PTEN) and tuberous sclerosis complex 2 (TSC2) increases the synthesis of HIF proteins in cancer [11, 12]. HIFs transcriptionally activate a wide spectrum of genes with broad functions. Some transcriptional targets allow cancer cells to adapt and survive hypoxia, while some transcriptional targets allow cancer cells to acquire more aggressive phenotypes. Through transcriptionally activating distinct subsets of genes, HIFs are involved in almost every step of cancer development through inducing angiogenesis, metabolic reprogramming, metastasis, cancer stemness maintenance, chemoresistance, and immune evasion.

1.1 Angiogenesis

Angiogenesis is a complicated but important mechanism to establish tumor vasculature, providing cancer cells nutrients. It is initiated by the removal of pericytes from the endothelium which results in the structural destabilization of blood vessels [13]. This activates endothelial cells to acquire proliferative character. Vascular endothelial growth factor (VEGF) increases the permeability of blood vessels and induces the release of proteases and matrix components. Endothelial cells migrate through the remodeled extracellular matrix (ECM) and proliferate to form tubes. Circulating mesenchymal cells move toward the new tubes or vessels and differentiate into pericytes to structurally form new blood vessels. VEGF is one of the most famous HIF transcriptional targets and represents one of the most important angiogenic genes for arterial destabilization and hypervascular permeability in cancer promotion [14]. Genes involved in increasing vascular permeability include

fms-like tyrosine kinase (FLT-1), angiopoietin-2 (ANGPT2), and Tie-2; genes involved in ECM modeling include matrix metalloproteinases (MMPs) and collagen; genes involved in migration and proliferation of endothelial cells include VEGF, placental growth factor (PLGF), ANGPT1, monocyte chemoattractant protein 1 (MCP-1), platelet-derived growth factor (PDGF), stromal cell-derived factor 1 (SDF-1), and C-X-C motif chemokine receptor 4 (CXCR4); genes involved in endothelial cell sprouting include ANGPT2 and Tie-2; genes involved in the recruitment of new pericytes include PDGF, plasminogen activator inhibitor-1 (PAI-1), ANGPT1, and Tie-2—all the aforementioned genes are regulated by HIF and summarized by Hirota et al. [13].

1.2 Metabolic Reprogramming

O_2 serves as the ultimate electron recipient of the electron transport chain (ETC) to complete oxidative phosphorylation. Due to the lack of final electron recipient in hypoxia, cells could not complete electron transfer entirely but in turn generate a large amount of reactive oxygen species (ROS) in the ETC. HIF-1 acts as a metabolic switch to divert the metabolic flux from tricarboxylic acid (TCA) cycle, which couples with oxidative phosphorylation to glycolysis, so that cells can metabolize glucose despite the lack of O_2. Glucose is converted to pyruvate which is further converted to lactate instead of acetyl-CoA, which would otherwise enter the TCA cycle. HIF-1 transcriptionally activates pyruvate dehydrogenase kinase 1 (PDK1) [15], lactate dehydrogenase A (LDHA) [16, 17], glucose transporters, and all glycolytic enzymes to achieve the metabolic switch [18–20]. Glucose transporter (GLUT) accelerates the glucose uptake to increase the turnover of glucose to compensate for the energy inefficiency in glycolysis. PDK1 inhibits pyruvate dehydrogenase (PDH), which would otherwise convert pyruvate into acetyl-CoA to initiate the TCA cycle. LDHA converts pyruvate into lactate to complete glycolysis. Through upregulating these enzymes, HIF-1 diverts pyruvate into glycolysis from TCA cycle to prevent ROS accumulation and to sustain the metabolic requirements of cancer cells. Further, HIF-1 induces the transcription of cytochrome oxidative 4-2 (COX4-2) and NDUFA4L2, less active subunits in complex IV and complex I, respectively, of the ETC, in order to reduce the electron flow rate through ETC and to alleviate ROS production [21–23]. All the aforementioned HIF-1 metabolic targets were shown to be upregulated in cancer and confer survival advantage to cancer cells.

1.3 Metastasis

Metastasis is the terminal stage of cancer development and a major cause of death in cancer patients. Metastasis involves the detachment of cancer cells from their primary organ site to another part of the same organ or to distant organs. It is a multistep process which includes the increase of cancer cell motility, invasion of cancer cells into the neighboring extracellular matrix, intravasation,

extravasation, and colonization of metastatic sites. Hypoxia and HIFs are involved in all steps of metastasis. Cancer cells undergo epithelial-mesenchymal transition (EMT) to lose adherent junction, in order to detach from the primary site. The EMT is characterized by the loss of E-cadherin, a transmembrane glycoprotein required for cell-cell junction interaction. Additional numerous properties of HIF in cancer metastasis are discussed below:

HIF activates E-cadherin repressors, Snail family transcriptional repressor 1 (SNAIL1), SNAIL2, zinc finger E-box binding homeobox (ZEB), Twist family BHLH transcription factor (TWIST), and transcription factor 3 (TCF3) [24, 25]. HIF also promotes mesenchymal properties by activating vimentin [26]. HIF additionally activates RhoA/Rho-kinase 1 (ROCK1) pathway to induce cytoskeletal remodeling to increase cell motility [27]. HIF-activated matrix metallopeptidase 2 (MMP2), MMP9, and MMP14 participate in the degradation of components of the ECM to enable cancer cell invasion into the stroma. HIF also transcriptionally activates collagen-synthesizing and collagen-modifying enzymes, such as prolyl-4-hydroxylases (P4HA1 and P4HA2), lysyl hydroxylase 2 (PLOD2), and lysyl oxidase (LOX) family members, to increase collagen synthesis and modification, in order to create anchoring substrates to facilitate cancer cell movement into the neighboring stroma [28–31]. In order to intravasate and extravasate, HIF transcriptionally activates L1 cell adhesion molecule (L1CAM), which facilitates adhesion of cancer cells to endothelial cells of blood vessels [32]. HIF also transcriptionally activates ANGPTL4 to increase the permeability of endothelial cells to facilitate the penetrance of cancer cells through the blood vessels [32]. HIF mediates the transcription of secretory enzymes, LOX family members, to cross-link collagen in the lung tissue, allowing bone marrow-derived cells (BMDCs) to adhere at the lung tissue epithelium [33, 34]. As BMDCs release angiogenic factors, BMDCs create a favorable metastatic niche for cancer cells to colonize and propagate at metastatic sites.

1.4 Cancer Stemness

Cancer stem cells, also referred as tumor-initiating cells, represent a population of cancer cells with stemness properties including abilities to self-renew and maintain pluripotency [35]. Cancer stem cells are highly refractory to chemotherapies. Cancer stem cells have been identified in various cancer types including leukemia and breast, brain, colon, liver, pancreatic, prostate, and head and neck cancers, as well as in melanoma and neuroblastoma. Multiple studies suggest that hypoxia maintains stemness properties of cancer cells. One of such studies has shown that a constitutive expression of HIF-1α induces CD133, a stem cell marker, in glioblastoma cells and augments glioblastoma stem cells [36]. HIFs also induce expression of E26 transformation-specific (ETS) transcription factors to activate prominin (PROM1), gene that encodes CD133

[37]. Hypoxia induces the expression of stemness genes, including Kruppel-like factor 4 (KLF4), Nanog homeobox (NANOG), proto-oncogene c-Myc (MYC), octamer-binding protein 4 (OCT4), and SRY-box 2 (SOX2) through HIFs [38]. Hypoxia also induces the expression of tafazzin (TAZ) and yes-associated protein (YAP) in the hippo pathway through HIF-1α, in order to maintain stemness, as shown in a breast cancer model [39, 40]. In leukemic stem cells, HIF-1α maintains stemness through inducing the Notch signalling pathway [41].

1.5 Chemoresistance

Hypoxia is a characteristic of resistance to chemotherapy. The most obvious reason is that hypoxic tumors are inaccessible to blood vessels which would otherwise deliver drugs to the vicinity. HIF-1α induces the transcription of multidrug resistance 1 (MDR1) gene, encoding membrane-resident P-glycoprotein (P-gp), a member of the ATP-binding cassette (ABC) transporters, and functions to stimulate the efflux of drugs to maintain low intracellular drug concentrations [42, 43]. P-gp was shown to efflux many chemotherapeutic drugs such as 5-fluorouracil, methotrexate, anthracyclines, and paclitaxel [42–44]. As mentioned above, cancer stem cells are more resistant to chemotherapies. Other HIF-related molecular mechanisms leading to chemotherapy failure include induction of autophagy, as well as suppression of senescence, mitochondrial activity, apoptosis, and DNA damage sensing, which are comprehensively summarized by Rohwer et al. [44]. As was demonstrated in a triple-negative breast cancer model, chemotherapy through HIF-1α was shown to induce the transcription of two glutathione synthesis genes, cystine/glutamate transporter (SLC7A11), and glutamate-cysteine ligase modifier subunit (GCLM), thereby allowing increased glutathione production, which in turn increased the chelation of copper to inactivate mitogen-activated protein kinase-extracellular signal-regulated kinases (MAPK-ERK) signalling [45]. Inactivation of MAPK-ERK promoted the nuclear translocation of forkhead box O3 (FOXO3) to transcribe stemness gene NANOG to further enrich breast cancer stem cells [45]. Chemotherapy was also shown to induce interleukin 6 (IL-6) and IL-8 through HIFs to maintain breast cancer stemness [45].

1.6 Immune Evasion

Cancer cells coexist with various immune cells in the tumor niche. Immune cells could be generally categorized into pro-tumorigenic and anti-tumorigenic. Pro-tumorigenic immune cells include tumor-associated macrophages (TAMs), myeloid-derived suppressor cells (MDSCs), and T-regulatory cells (Tregs) which counteract anti-tumorigenic immune cells such as cytotoxic T cells, helper T cells, natural killer (NK) cells, macrophages, and dendritic cells. Anti-tumorigenic immune cells provide immune surveillance and clear foreign and abnormal cells, including cancer cells. Hypoxia allows cancer cells to evade the immune surveillance. Hypoxia

favors the accumulation of pro-tumorigenic immune cells, which is evidenced by the dual and overlapping staining of macrophage or MDSC markers with hypoxia markers in different cancer models [46, 47]. Hypoxia was shown to increase C-C motif chemokine ligand 2 (CCL2), CCL5, colony-stimulating factor 1 (CSF1), VEGF, semaphorin 3A (SEMA3A), endothelial cell monocyte-activating polypeptide-II (EMAP-II), endothelin, stromal cell-derived factor 1α (SDF1α), eotaxin, and on statin M, which are factors that attract macrophages to the tumor [46, 48–53]. Hypoxia was also shown to lead to an increase of necrotic cancer cells, which release damage-associated molecular pattern (DAMP) molecules to recruit innate immune cell to the tumor [54]. Hypoxia through HIFs was shown to turn on the transcription of CCL26 and CCL28 to recruit MDSCs and Tregs to the tumor [47, 55]. Hypoxia also maintained macrophages in the vicinity by shutting down the expression of CCR2, CCR5, and neuropilin-1 (NRP1), in turn reducing their mobility [46, 56, 57]. Once kept in the tumor site, macrophages acquired angiogenic and immune escape capabilities. Furthermore, HIF-1α in macrophages increased the expression of programmed cell death 1 ligand 1 (PDL1), which, in turn, relayed inhibitory signals in T cells through ligation to receptor programmed cell death 1 (PD-l) on T cells [58]. Co-culturing hypoxic cancer cells with macrophages turned on the transcription of indoleamine 2,3-deoxygenase (IDO) expression, which depleted T cells from tryptophan, the essential nutrient for T-cell proliferation and survival [59]. Recently, we demonstrated that HIF-1α transcriptionally activated ectonucleoside triphosphate diphosphohydrolase 2 (ENTPD2) in liver cancer cells [60]. ENTPD2 catalyzed the conversion of extracellular ATP to AMP, which maintained MDSCs, impairing the proliferation of T cells [60].

New transcriptional targets of HIF-1 are continuously being discovered, which further emphasizes the diverse roles and functional importance of HIF. However, many researchers encounter difficulties in hypoxia studies as HIF-1/2α is prone to degradation. To facilitate the hypoxia/HIF-signalling research, this chapter will provide detailed protocols from setting up hypoxic conditions in cell culture to assessing the stability and activity of HIFs in cancer cell culture systems and mouse cancer models.

2 Materials

2.1 Setup of Culturing System in Hypoxia

1. Hypoxia incubator chamber (*see* Fig. 1).
2. Single flow meter.
3. Hypoxia gas tank (1% O_2).
4. Polycarbonate tubing.
5. Tissue culture incubator (37 °C).

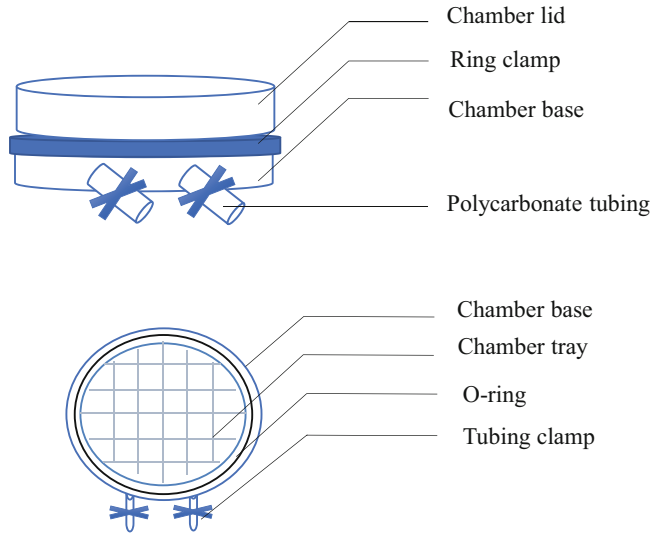


Fig. 1 Hypoxia chamber compartments. The hypoxia incubator chamber is generally composed of four parts: chamber lid, ring clamp, chamber tray, and chamber base. The chamber base contains O-ring and polycarbonate tubing with a plastic clamp. The tray is placed onto the chamber base, and the cells are put on the tray. When the chamber lid and base are locked by ring clamp, the O-ring fills the spare space between the edges of these two compartments, ensuring that they are tightly sealed

2.2 Detection of HIF-1/2 Protein by Western Blotting

1. RIPA buffer: 50 mM Tris–HCl, 150 mM NaCl, 1 mM EDTA, 0.7% (w/v) NP-40, 0.05% (w/v) SDS, pH7.4.
2. Protease inhibitor cocktail: cOmplete™, EDTA-free Protease Inhibitor Cocktail. Dissolve one tablet of cOmplete™ in 2 mL distilled water to prepare a 25× concentrated stock solution.
3. Phosphatase inhibitor: PhosSTOP™. Dissolve one tablet of PhosSTOP™ in 0.5 mL distilled water to prepare a 20× concentrated stock solution.
4. Phosphate-buffered saline (PBS; 1×): Chill the 1× PBS to 4 °C.
5. Cell scrapers.
6. Centrifuges for 15 mL conical tubes and 1.5 mL micro-centrifuge tubes: Cool centrifuges to 4 °C.
7. Bradford reagent.
8. 6× SDS sample buffer.
9. Protein ladder.
10. 10% SDS-PAGE gel.
11. SDS-PAGE running buffer (1×): 25 mM Tris base, 192 mM glycine, 0.1% (w/v) SDS.

12. PVDF membranes.
13. Western blot transfer buffer: 20 mM Tris base, 192 mM glycine, 20% (v/v) methanol.
14. Tris-buffered saline (TBS; 1×): 10 mM Tris–HCl, 150 mM NaCl, pH 7.4.
15. TBST: 1× TBS containing 0.05% Tween 20.
16. Blocking solution and diluent solution: 5% milk in TBST. Store at 4 °C.
17. Anti-HIF1α antibody (Cell Signaling, 3716S).
18. Anti-rabbit IgG, HRP-linked antibody.
19. ECL-HRP substrate.

2.3 Detection of HIF-1α Activity Using Luciferase Reporter Assay

1. Anti-rabbit IgG, HRP-linked antibody.
2. ECL-HRP substrate.
3. Transfection reagent: X-tremeGENE™ 9 DNA transfection reagent (*see* **Note 1**).
4. DMEM-HG medium without serum and antibiotics.
5. pGL2.1-5× HRE-luciferase plasmid (*see* **Note 2**).
6. pRL-CMV plasmid encoding Renilla luciferase.
7. Phosphate-buffered saline (PBS; 1×).
8. Dual-Luciferase® Reporter Assay Kit (Promega):
   (a) Passive Lysis Buffer (PLB; 5×): dilute 5× PLB to 1× with distilled water.
   (b) Luciferase Assay Reagent II (LAR II): dissolve the luciferase assay substrate in 10 mL of Luciferase Assay Buffer II (*see* **Note 3**).
   (c) Stop & Glo® Reagent: dilute 50× Stop & Glo® substrate to 1× in Stop & Glo® buffer (*see* **Note 4**).
9. White opaque 96-well microplate.
10. Luminometer with reagent auto-injectors.

2.4 In Vivo Detection of Hypoxia Markers Using Immunohistochemistry

1. Antigen retrieval buffer: 1 mM EDTA, pH 7.8.
2. Tris-buffered saline (TBS; 10×): 1.5 M NaCl, 0.1 M Tris-HCl, pH 7.4.
3. Phosphate-buffered saline (PBS; 10×).
4. Washing buffer: 1× TBS containing 0.05% Tween 20 (TBST).
5. Peroxidase quenching solution: 1× TBS containing 3% H_2O_2.
6. 10% formalin in PBS.
7. 10× casein solution.

8. Chromogen solution: 1× TBS (50 mL) containing 25 mg 3,3′-diaminobenzidine and 15 μL 100% H_2O_2.
9. Harris hematoxylin solution.
10. Scott's tap water.
11. Xylene.
12. 100% ethanol.
13. Paraffin.
14. Coverslip.
15. Mounting medium.
16. Horseradish peroxidase (HRP)-conjugated anti-mouse/rabbit secondary antibody from Dako.
17. Hypoxyprobe™-1 Kit from Hypoxyprobe.
18. Rabbit antihuman glucose transporter 1 (GLUT1) antibody (ab15309) from Abcam.
19. Rabbit antihuman carbonic anhydrase 9 (CA9) antibody (ab15086) from Abcam.

2.5 Detection of HIF Transcriptional Targets by Real-Time Quantitative PCR

1. Phosphate-buffered saline (PBS; 1×).
2. Trizol.
3. Chloroform.
4. Isopropanol (IPA).
5. 75% ethanol.
6. DEPC water.
7. Centrifuge for 1.5 mL micro-centrifuge tubes: Cool the centrifuge to 4 °C.
8. Heat block: Set the temperature to 55–60 °C.
9. Nanodrop.
10. PCR machine.
11. PCR tubes.
12. Milli-Q water.
13. 10× PCR Buffer (ABI, Ref: 4486220).
14. 25 mM $MgCl_2$.
15. Random hexamer (Invitrogen, Ref: 100026484).
16. dNTP Mix (10 mM).
17. RNase inhibitor (ABI, Ref: 100021540).
18. Reverse Transcriptase (ABI, Ref: 4308228).

3 Methods

3.1 Setup of Culturing System in Hypoxia

1. Seed appropriate number of cells in the tissue culture dish 1 day prior to hypoxia incubation (*see* **Notes 5** and **6**).
2. To assemble the hypoxia chamber, open the hypoxia chamber by removing the ring clamp and the chamber lid (*see* **Note 7**). Put the culture dishes carefully onto the tray.
3. Place the chamber lid to the base, and make sure the edge of the lid fits with O-ring exactly. Apply the ring clamp to the position where the chamber lid and base meet, and close the ring clamp slowly by pushing the buckle toward the chamber (*see* **Notes 8** and **9**).
4. To purge the chamber, connect the inlet port and outlet port of the single flow meter to the gas tank and hypoxia chamber, respectively (polycarbonate tubing can be used to extend the tubing length) (*see* Fig. 2). Ensure that both of the tubing clamps in the hypoxia chamber are opened.
5. Switch on the gas valve to let the gas flow into the chamber. Adjust the flow rate control to steady the flow rate at 30–40 L/min (LPM) (*see* **Note 10**). Flush the chamber with the gas for no less than 2 min (*see* **Note 11**).
6. To seal the chamber, close the outlet tubing clamp followed by the inlet clamp. Turn off the gas valve, and disconnect the chamber tubing and the gas supply.
7. Put the hypoxia chamber into the tissue culture incubator (*see* **Note 12**).

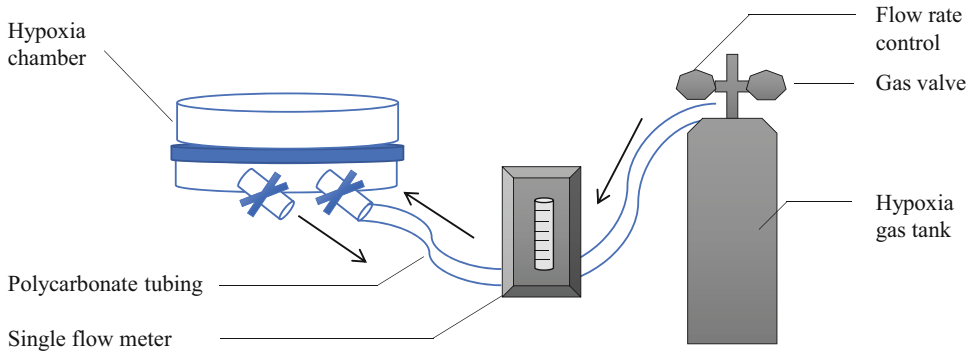


Fig. 2 Setup of hypoxia gas flow. The hypoxia gas tank is connected to the inlet port of the single flow meter by polycarbonate tubing. The outlet port of the meter is linked to one of the polycarbonate tubing at the bottom of the hypoxia chamber base. When the gas flow is turned on, the gas comes from the hypoxia tank flush into the hypoxia chamber. The air in the chamber is purged through the other polycarbonate tubing. Black arrows in the figure indicate the direction of the gas flow throughout the setting

8. To open the hypoxia chamber, take the hypoxia chamber out of the tissue culture incubator. Open the tubing clamps to release the gas in the chamber (*see* **Note 13**).
9. Release and remove the ring clamp (*see* **Note 14**). Remove the chamber lid, and collect the cells as soon as possible for further analysis (*see* **Notes 15** and **16**).

3.2 Detection of HIF-1/2 Protein by Western Blotting

1. Seed appropriate number of cells in two 60 mm dishes the day before protein extraction such that the cells should be 70–90% confluent at the time of protein extraction.
2. On the day of the protein extraction, place one set of cells in hypoxia chamber, and fill the chamber with 1% hypoxia gas as in Subheading 3.1.
3. Place the hypoxia chamber in the 37 °C incubator for 4 h.
4. Prior to the protein extraction, add protease inhibitor cocktail (25×) and phosphatase inhibitor (20×) to RIPA buffer, and dilute them to 1×.
5. For the normoxic condition, place the cells on ice. Remove the medium from cells, and wash the cells with pre-chilled 1× PBS once.
6. Add 4 mL of pre-chilled 1× PBS to cells and scrape the cells with cell scrapers.
7. Transfer the cells to 15 mL conical tubes, and keep the cells on ice.
8. For the hypoxic condition, open the hypoxia chamber, and place the cells on ice immediately. Cells are collected as in steps 5–7 (*see* **Note 17**).
9. Pellet cells for both normoxic and hypoxic conditions at 2000 × *g* for 2 min at 4 °C.
10. Remove the supernatant.
11. Resuspend the cell pellets with 20–100 μL RIPA buffer containing 1× protease inhibitor cocktail and 1× phosphatase inhibitor according to the size of the pellets.
12. Transfer the lysates into 1.5 mL micro-centrifuge tubes.
13. Vortex the lysates and incubate on ice for 15 min.
14. Centrifuge the lysates at 12,000 × *g* for 15 min at 4 °C.
15. Transfer the supernatant to new 1.5 mL micro-centrifuge tubes (*see* **Note 18**).
16. Determine protein concentration by Bradford assay.
17. Prepare 40 μg of protein per sample in 1× SDS sample buffer.
18. Heat the samples at 95 °C for 10 min before loading to the SDS-PAGE gel.

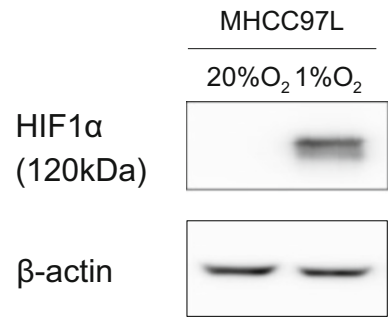


Fig. 3 HIF-1α protein expression in liver cancer cells MHCC97L exposed to 20% O_2 (normoxia) and 1% O_2 (hypoxia)

19. Load the samples and protein ladder into the 10% SDS-PAGE gel.
20. Run the SDS-PAGE according to the standard SDS-PAGE protocol.
21. Transfer the proteins to PVDF membrane according to the standard transfer protocol.
22. Block the membrane with 5% milk in TBST at room temperature for 1 h.
23. Wash the membrane with TBST for 15 min once and 5 min twice.
24. Dilute anti-HIF1α antibody 1:1000 in 5% milk in TBST.
25. Add the diluted anti-HIF1α antibody to the membrane, and incubate at 4 °C overnight.
26. Remove the anti-HIF1α antibody (*see* **Note 19**).
27. Wash the membrane with TBST for 15 min once and 5 min twice.
28. Dilute anti-rabbit IgG, HRP- linked antibody 1:2500 in 5% milk in TBST.
29. Add the diluted anti-rabbit antibody to the membrane, and incubate at room temperature for 2 h.
30. Remove the diluted anti-rabbit antibody.
31. Wash the membrane with TBST for 15 min once and 5 min twice.
32. Develop the bands with ECL-HRP substrate (*see* Fig. 3 and **Note 20**).

3.3 Detection of HIF-1α Activity Using Luciferase Reporter Assay

1. Seed the appropriate number of cells in two 24-well plates the day before transfection such that the cells should be 40–60% confluent at the time of the transfection.

2. For each well containing the cells, dilute 0.25 μg pGL2.1-5× HRE-luciferase plasmids and 1.25 ng pRL-CMV plasmids in 25 μL DMEM-HG medium without serum and antibiotics.
3. For each well of cells, dilute 0.75 μL X-tremeGENE™ 9 DNA transfection reagent in 25 μL DMEM-HG medium without serum and antibiotics.
4. Add the diluted DNA transfection reagent into the DNA-diluted solution, and incubate at room temperature for 20 min.
5. Add 50 μL of DNA-X-tremeGENE™ 9 complexes to each well of cells dropwise, and mix gently.
6. Sixteen hours after transfection, remove medium, and replenish with new growth medium.
7. Place one set of cells in the hypoxia chamber, and fill the chamber with 1% hypoxia gas as in Subheading 3.1.
8. Place the hypoxia chamber in the 37 °C incubator for 24 h.
9. For both normoxic and hypoxic conditions, remove the medium, and wash the cells with 1× PBS once.
10. Add 100 μL 1× passive lysis buffer in each well of cells.
11. Place the plates on the shaker at room temperature for 15 min to lyse the cells.
12. Mix the lysates in each well, and transfer 20 μL lysates to a white 96-well microplate (*see* **Note 21**).
13. Load LAR II and Stop & Glo® reagent into the reagent auto-injectors.
14. Insert the white 96-well plate into the luminometer that is connected to the reagent auto-injectors.
15. For each well, inject 50 μL LAR II to the well, and the firefly luminescence will be measured.
16. Inject 50 μL Stop & Glo® reagent, and the Renilla luminescence will be measured.
17. The activity of firefly luciferase is normalized to the activity of Renilla luciferase.

3.4 In Vivo Detection of Hypoxia Markers Using Immunohistochemistry

1. Dissolve solid pimonidazole HCl (Hypoxyprobe™-1) with saline. The solubility of Hypoxyprobe™-1 in saline is 116 mg/mL.
2. Thirty minutes prior to tumor harvest, administrate 60 mg/kg Hypoxyprobe™-1 solution to tumor-bearing mice via intraperitoneal injection (*see* **Notes 22** and **23**).
3. Harvest and fix the tumor tissues with 10% formalin for 24 h (*see* **Note 24**).

4. After fixation, transfer the tissues to 70% ethanol for processing or short-term storage (*see* **Note 25**).
5. Dehydrate the tissues with ethanol and xylene, and embed the tissues in paraffin according to the conventional protocols.
6. Cut the paraffin-embedded tissue blocks into sections at a thickness of 4–5 μm, and affix onto the slides. Once mounted, the slides should be dried in the 37 °C oven overnight to remove any water that may be trapped under the section.
7. Before proceeding with a staining protocol, deparaffinize the sections with xylene, and rehydrate with a gradient of ethanol and water as follows:
   (a) Xylene: 4 × 3 min
   (b) 100% ethanol: 4 × 3 min
   (c) 95% ethanol: 1 min
   (d) 85 % ethanol: 1 min
   (e) Rinse with cold tap water
8. For antigen retrieval, incubate the sections in boiling 1 mM EDTA solution (pH 7.8) for 15 min.
9. After cooling down, transfer and incubate the sections in 1× TBS containing 3% H_2O_2 at room temperature for 30 min to suppress the endogenous peroxidase activity.
10. Rinse the sections with 1× TBS four times.
11. Dilute 10× casein solution to 2× with 1× PBS. Incubate the sections in 2× casein solution at room temperature for 10 min for protein blocking.
12. Dilute the primary antibodies with antibody diluent as follows:
    (a) Rabbit anti-GLUT1 (1:1000)
    (b) Rabbit anti-CA9 (1:500)
    (c) Mouse anti-pimonidazole (1:50)
13. Remove the casein solution, and incubate the sections with diluted primary antibodies at 4 °C overnight.
14. The next day, remove the primary antibodies, and wash the sections with 1× TBST for 5 min four times with gentle agitation.
15. Incubate the sections with HRP-conjugated secondary antibody of corresponding species at room temperature for 30 min with gentle agitation.
16. Remove the secondary antibody, and wash the sections with 1× TBST for 5 min four times with gentle agitation.
17. For detection, incubate the sections with freshly prepared chromogen solution at room temperature for 1–10 min. Stop

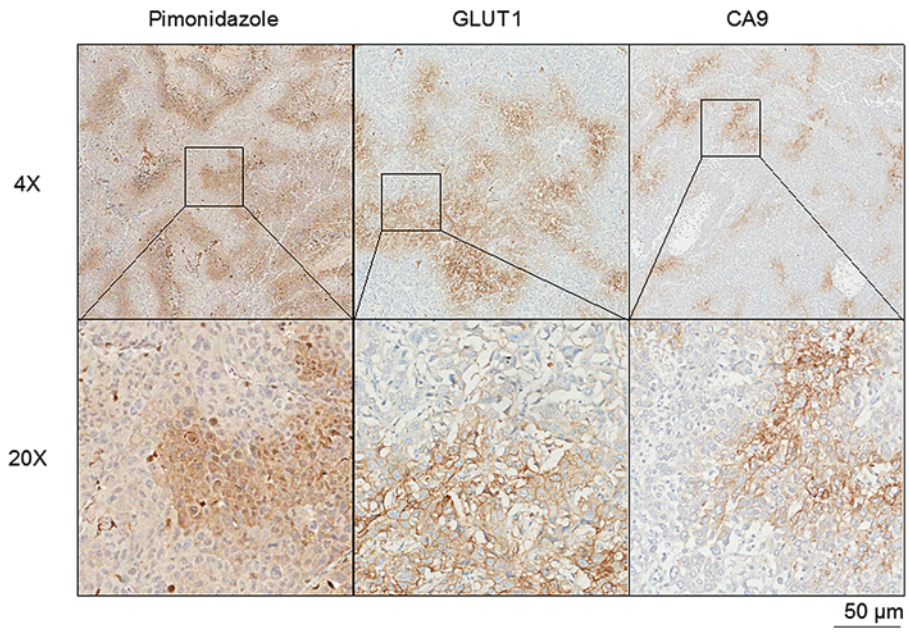


Fig. 4 Immunohistochemical staining of hypoxia markers (pimonidazole, GLUT1, and CA9) in human HCC tissues displayed a patchy pattern, which is a typical oxygen diffusion pattern

the reaction with tap water when the desired signal intensity is achieved (*see* **Note 26**).

18. Counterstain with hematoxylin solution and Scott's tap water according to the manufacturer's instruction.
19. Dehydrate the sections with a gradient of ethanol and xylene as follows:
    (a) 85 % ethanol: 1 min
    (b) 95% ethanol: 1 min
    (c) 100% ethanol: 4 × 3 min
    (d) Xylene: 4 × 3 min
20. Mount the sections (*see* Fig. 4).

3.5 Detection of HIF Transcriptional Targets by Real-Time Quantitative PCR

1. Seed the appropriate number of cells in two 6-well plates such that the cells should be 70–90% confluent at the time of RNA extraction.
2. On the next day, place one set of cells in the hypoxia chamber, and fill the chamber with 1% hypoxia gas as in Subheading 3.1.
3. Place the hypoxia chamber in the 37 °C incubator for 24 h.
4. For both normoxic and hypoxic conditions, remove the medium, and wash the cells with 1× PBS once.
5. Add 1 mL Trizol to cells.
6. Incubate at room temperature for 5 min.

7. Pipette up and down several times, and transfer to 1.5 mL micro-centrifuge tubes.
8. Add 0.2 mL chloroform.
9. Shake tubes vigorously by hand for 15 s, and incubate at room temperature for 3 min.
10. Centrifuge at 12,000 × *g* for 10 min at 4 °C.
11. Transfer the upper aqueous phase to new 1.5 mL micro-centrifuge tubes (*see* **Note 27**).
12. Add 0.5 mL 100% IPA and vortex the samples.
13. Incubate at room temperature for 10 min.
14. Centrifuge at 12,000 × *g* for 10 min at 4 °C.
15. Remove the supernatant.
16. Wash the RNA pellet with 1 mL 75% ethanol and vortex the samples.
17. Centrifuge at 7500 × *g* for 10 min at 4°C.
18. Dry the RNA pellet (*see* **Note 28**).
19. Dissolve RNA in 20–50 μL DEPC water according to the size of pellets.
20. Incubate the samples at 55–60 °C heat block for 10 min.
21. Measure the RNA concentration using nanodrop (*see* **Note 29**).
22. Aliquot 1 μg RNA into PCR tubes, and top up to 3 μL using Milli-Q water (*see* **Note 30**).
23. Close the cap tightly and heat RNA at 65 °C for 10 min.
24. Prepare Mastermix solution, and mix thoroughly (*see* **Note 31**).

Reagents	**μL**
10× PCR Buffer	2
$MgCl_2$	4
dNTP	8
Random hexamer	1
RNase inhibitor	1
Reverse transcriptase	1

25. Add 17 μL of Mastermix solution into RNA samples to make a 20 μL PCR system while pipetting to ensure a well-mixed reaction.

26. Set up PCR amplification as follows:

25 °C	10 min
42 °C	60 min
95 °C	5 min
4 °C	∞

27. 1:5 dilute cDNA sample in Milli-Q water. The diluted cDNA is ready for quantitative PCR (*see* **Notes 32** and **33**).

4 Notes

1. Other transfection reagents may be used for the transfection, depending on the transfection efficiency of used cell lines.
2. The pGL2.1-5× HRE-luciferase plasmid contains hypoxia response element from human ENO1 gene and firefly luciferase coding sequence [17].
3. Freeze-and-thaw cycles of LAR II reduce the reagent's activity. Aliquots of LAR II are recommended after the reconstitution of luciferase assay substrate.
4. Stop & Glo® reagent should be prepared freshly for each experiment.
5. For each assay which needs to be carried out after hypoxia chamber incubation, always remember to have a normoxia (20% O_2) positive control with the same experimental conditions.
6. In order to avoid the insufficient cell volume or overgrowth of the cells, a preliminary test for optimizing the cell seeding density is recommended if the growth rate of the cells is unknown under hypoxia.
7. Before using the hypoxia chamber, it is necessary to ensure that all the compartments are intact without a crack and the O-ring is evenly distributed around the flange of the chamber base.
8. When assembling and opening the ring clamp, hold the chamber firmly with hands to avoid cell spillage.
9. If there were a mismatch between the lid and the base, it would be extremely hard to close the ring clamp. Do not force to close the clamp in this case. Adjust the position of the chamber lid and base until they are properly assembled.
10. Before purging the chamber, one can shortly close the outlet tubing of the hypoxia chamber by fingers to check the leakage of the gas. If the chamber lid lifts accompanied by dropped flow rate, it indicates that the lid and base are properly sealed by ring

clamp. If leakage is found, adjust the position of the chamber lid or ring clamp until the chamber is sealed.

11. According to our experience, 2 min is the minimum time for purging the chamber. The time could be extended to 3–4 min if required and may vary depending on different experimental settings. Thus, a preliminary test is required for a successful induction of HIF activation.
12. For detecting the mRNA level of HIF targets, cells should be kept in hypoxia chamber for 24 h. For testing the change of the protein level, the incubation time should be extended to 48 h.
13. When opening the tubing clamps after hypoxia, the sound of escaping of the gas should be heard. It means that the chamber is tightly sealed after purging.
14. Normally, the air pressure is higher within the hypoxia chamber after purging. When opening the hypoxia chamber, it is better to open the tubing clamps to balance the air pressure before release of the ring clamp.
15. Since the HIF proteins are easily degraded under normoxia condition (20% O_2), the cells should be collected immediately after opening the hypoxia chamber.
16. The hypoxia chamber should be kept at room temperature when it is not in use. Both the ring clamp and the tube clamps should be released for better maintenance.
17. Cells should be handled as fast as possible until RIPA buffer is added as HIF1α degrades very quickly. The process can be done in cold room to slow the degradation of HIF1α.
18. Lysates can be stored at −80 °C for later use.
19. The diluted anti-HIF1α antibody in 5% milk in TBST can be stored at −20 °C and reused several times.
20. The band of HIF1α appears around 120 kDa.
21. The white residues in lysates do not affect the dual-luciferase reporter assay. Removing of white residues in lysates is not necessary.
22. In our model, BALB/c nude mouse is orthotopically injected with human HCC cell line, MHCC97L.
23. The plasma half-life of pimonidazole HCl is approximately 25 min.
24. The ideal fixation time depends on the size and type of the tissue. Fixation between 18 and 24 h is ideal for this application.
25. Formalin-fixed tissues are stable in 70% ethanol within weeks of this application.

Table 1
Primer sequences

Genes	Sequences (5′–3′)
Human GLUT1	Forward: CGGGCCAAGAGTGTGCTAAA Reverse: TGACGATACCGGAGCCAATG
Human HK2	Forward: CCAGTTCATTCACATCATCAG Reverse: CTTACACGAGGTCACATAGC
Human LOX	Forward: GTTCCAAGCTGGCTACTC Reverse: GGGTTGTCGTCAGAGTAC
Human LOXL2	Forward: GGAAAGCGTACAAGCCAGAG Reverse: GCACTGGATCTCGTTGAGGT
Human NDUFA4L2	Forward: AAAAGACATCCGGGGATCAT Reverse: TCCGGGTTGTTCTTTCTGTC
Human CCL26	Forward: ATACAGCCACAAGCCCCTTC Reverse: TGGTAGTGAATATCACAGCCCG
Human ENTPD2	Forward: ACCCACAGCTTCCTCTGCTA Reverse: CTGAAGAGCCCAGAAACCAG

26. 3,3′-Diaminobenzidine (DAB) is a suspected carcinogen. Wear the appropriate protective clothing when preparing the chromogen solution.
27. Three phases will be seen after centrifugation. Avoid transferring the middle and lower phases to the new tubes.
28. Air-dry the RNA pellet until the pellet becomes colorless.
29. When measuring RNA concentration, also check the A260/A280 ratio to evaluate the RNA purity. One sample with high RNA purity will have the ratio between 1.8 and 2.1.
30. Always keep RNA on ice before heating.
31. The Mastermix solution should be freshly prepared and kept on ice before using.
32. For long-term storage, the cDNA sample can be kept at −20 °C.
33. Our validated primer sequences for HIF transcriptional targets can be found in Table 1.

References

1. Semenza GL (2012) Hypoxia-inducible factors in physiology and medicine. Cell 148 (3):399–408. https://doi.org/10.1016/j.cell.2012.01.021
2. Epstein AC, Gleadle JM, McNeill LA, Hewitson KS, O'Rourke J, Mole DR, Mukherji M, Metzen E, Wilson MI, Dhanda A, Tian YM, Masson N, Hamilton DL, Jaakkola P, Barstead R, Hodgkin J, Maxwell PH, Pugh CW, Schofield CJ, Ratcliffe PJ (2001) C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate

HIF by prolyl hydroxylation. Cell 107 (1):43–54

3. Kaelin WG Jr, Ratcliffe PJ (2008) Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. Mol Cell 30 (4):393–402. https://doi.org/10.1016/j.molcel.2008.04.009
4. Wang GL, Jiang BH, Rue EA, Semenza GL (1995) Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. Proc Natl Acad Sci U S A 92(12):5510–5514
5. Semenza GL (2012) Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy. Trends Pharmacol Sci 33(4):207–214. https://doi.org/10.1016/j.tips.2012.01.005
6. Hansen WJ, Ohh M, Moslehi J, Kondo K, Kaelin WG, Welch WJ (2002) Diverse effects of mutations in exon II of the von Hippel-Lindau (VHL) tumor suppressor gene on the interaction of pVHL with the cytosolic chaperonin and pVHL-dependent ubiquitin ligase activity. Mol Cell Biol 22(6):1947–1960
7. Maxwell PH, Wiesener MS, Chang GW, Clifford SC, Vaux EC, Cockman ME, Wykoff CC, Pugh CW, Maher ER, Ratcliffe PJ (1999) The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature 399 (6733):271–275. https://doi.org/10.1038/20459
8. Morris MR, Maina E, Morgan NV, Gentle D, Astuti D, Moch H, Kishida T, Yao M, Schraml P, Richards FM, Latif F, Maher ER (2004) Molecular genetic analysis of FIH-1, FH, and SDHB candidate tumour suppressor genes in renal cell carcinoma. J Clin Pathol 57 (7):706–711. https://doi.org/10.1136/jcp.2003.011767
9. Pollard PJ, Briere JJ, Alam NA, Barwell J, Barclay E, Wortham NC, Hunt T, Mitchell M, Olpin S, Moat SJ, Hargreaves IP, Heales SJ, Chung YL, Griffiths JR, Dalgleish A, McGrath JA, Gleeson MJ, Hodgson SV, Poulsom R, Rustin P, Tomlinson IP (2005) Accumulation of Krebs cycle intermediates and over-expression of HIF1alpha in tumours which result from germline FH and SDH mutations. Hum Mol Genet 14 (15):2231–2239. https://doi.org/10.1093/hmg/ddi227
10. Selak MA, Armour SM, MacKenzie ED, Boulahbel H, Watson DG, Mansfield KD, Pan Y, Simon MC, Thompson CB, Gottlieb E (2005) Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-alpha prolyl hydroxylase. Cancer Cell 7(1):77–85. https://doi.org/10.1016/j.ccr.2004.11.022
11. Liu MY, Poellinger L, Walker CL (2003) Up-regulation of hypoxia-inducible factor 2alpha in renal cell carcinoma associated with loss of Tsc-2 tumor suppressor gene. Cancer Res 63(10):2675–2680
12. Zundel W, Schindler C, Haas-Kogan D, Koong A, Kaper F, Chen E, Gottschalk AR, Ryan HE, Johnson RS, Jefferson AB, Stokoe D, Giaccia AJ (2000) Loss of PTEN facilitates HIF-1-mediated gene expression. Genes Dev 14(4):391–396
13. Hirota K, Semenza GL (2006) Regulation of angiogenesis by hypoxia-inducible factor 1. Crit Rev Oncol Hematol 59(1):15–26. https://doi.org/10.1016/j.critrevonc.2005.12.003
14. Forsythe JA, Jiang BH, Iyer NV, Agani F, Leung SW, Koos RD, Semenza GL (1996) Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. Mol Cell Biol 16(9):4604–4613
15. Kim JW, Tchernyshyov I, Semenza GL, Dang CV (2006) HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. Cell Metab 3(3):177–185. https://doi.org/10.1016/j.cmet.2006.02.002
16. Fantin VR, St-Pierre J, Leder P (2006) Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance. Cancer Cell 9 (6):425–434. https://doi.org/10.1016/j.ccr.2006.04.023
17. Semenza GL, Jiang BH, Leung SW, Passantino R, Concordet JP, Maire P, Giallongo A (1996) Hypoxia response elements in the aldolase A, enolase 1, and lactate dehydrogenase A gene promoters contain essential binding sites for hypoxia-inducible factor 1. J Biol Chem 271(51):32529–32537
18. Iyer NV, Kotch LE, Agani F, Leung SW, Laughner E, Wenger RH, Gassmann M, Gearhart JD, Lawler AM, Yu AY, Semenza GL (1998) Cellular and developmental control of O_2 homeostasis by hypoxia-inducible factor 1 alpha. Genes Dev 12(2):149–162
19. Kim JW, Gao P, Liu YC, Semenza GL, Dang CV (2007) Hypoxia-inducible factor 1 and dysregulated c-Myc cooperatively induce vascular endothelial growth factor and metabolic switches hexokinase 2 and pyruvate dehydrogenase kinase 1. Mol Cell Biol 27 (21):7381–7393. https://doi.org/10.1128/MCB.00440-07

20. Semenza GL, Roth PH, Fang HM, Wang GL (1994) Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor 1. J Biol Chem 269 (38):23757–23763
21. Fukuda R, Zhang H, Kim JW, Shimoda L, Dang CV, Semenza GL (2007) HIF-1 regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells. Cell 129(1):111–122. https://doi.org/10.1016/j.cell.2007.01.047
22. Lai RK, Xu IM, Chiu DK, Tse AP, Wei LL, Law CT, Lee D, Wong CM, Wong MP, Ng IO, Wong CC (2016) NDUFA4L2 fine-tunes oxidative stress in hepatocellular carcinoma. Clin Cancer Res 22(12):3105–3117. https://doi.org/10.1158/1078-0432.CCR-15-1987
23. Tello D, Balsa E, Acosta-Iborra B, Fuertes-Yebra E, Elorza A, Ordonez A, Corral-Escariz-M, Soro I, Lopez-Bernardo E, Perales-Clemente E, Martinez-Ruiz A, Enriquez JA, Aragones J, Cadenas S, Landazuri MO (2011) Induction of the mitochondrial NDUFA4L2 protein by HIF-1alpha decreases oxygen consumption by inhibiting Complex I activity. Cell Metab 14(6):768–779. https://doi.org/10.1016/j.cmet.2011.10.008
24. Krishnamachary B, Zagzag D, Nagasawa H, Rainey K, Okuyama H, Baek JH, Semenza GL (2006) Hypoxia-inducible factor-1-dependent repression of E-cadherin in von Hippel-Lindau tumor suppressor-null renal cell carcinoma mediated by TCF3, ZFHX1A, and ZFHX1B. Cancer Res 66(5):2725–2731. https://doi.org/10.1158/0008-5472.CAN-05-3719
25. Yang MH, Wu MZ, Chiou SH, Chen PM, Chang SY, Liu CJ, Teng SC, Wu KJ (2008) Direct regulation of TWIST by HIF-1alpha promotes metastasis. Nat Cell Biol 10(3):295–305. https://doi.org/10.1038/ncb1691
26. Krishnamachary B, Berg-Dixon S, Kelly B, Agani F, Feldser D, Ferreira G, Iyer N, LaRusch J, Pak B, Taghavi P, Semenza GL (2003) Regulation of colon carcinoma cell invasion by hypoxia-inducible factor 1. Cancer Res 63(5):1138–1143
27. Gilkes DM, Xiang L, Lee SJ, Chaturvedi P, Hubbi ME, Wirtz D, Semenza GL (2014) Hypoxia-inducible factors mediate coordinated RhoA-ROCK1 expression and signaling in breast cancer cells. Proc Natl Acad Sci U S A 111(3):E384–E393. https://doi.org/10.1073/pnas.1321510111
28. Erler JT, Bennewith KL, Nicolau M, Dornhofer N, Kong C, Le QT, Chi JT, Jeffrey SS, Giaccia AJ (2006) Lysyl oxidase is essential for hypoxia-induced metastasis. Nature 440 (7088):1222–1226. https://doi.org/10.1038/nature04695
29. Gilkes DM, Bajpai S, Wong CC, Chaturvedi P, Hubbi ME, Wirtz D, Semenza GL (2013) Procollagen lysyl hydroxylase 2 is essential for hypoxia-induced breast cancer metastasis. Mol Cancer Res 11(5):456–466. https://doi.org/10.1158/1541-7786.MCR-12-0629
30. Gilkes DM, Chaturvedi P, Bajpai S, Wong CC, Wei H, Pitcairn S, Hubbi ME, Wirtz D, Semenza GL (2013) Collagen prolyl hydroxylases are essential for breast cancer metastasis. Cancer Res 73(11):3285–3296. https://doi.org/10.1158/0008-5472.CAN-12-3963
31. Wong CC, Tse AP, Huang YP, Zhu YT, Chiu DK, Lai RK, Au SL, Kai AK, Lee JM, Wei LL, Tsang FH, Lo RC, Shi J, Zheng YP, Wong CM, Ng IO (2014) Lysyl oxidase-like 2 is critical to tumor microenvironment and metastatic niche formation in hepatocellular carcinoma. Hepatology 60(5):1645–1658. https://doi.org/10.1002/hep.27320
32. Zhang H, Wong CC, Wei H, Gilkes DM, Korangath P, Chaturvedi P, Schito L, Chen J, Krishnamachary B, Winnard PT Jr, Raman V, Zhen L, Mitzner WA, Sukumar S, Semenza GL (2012) HIF-1-dependent expression of angiopoietin-like 4 and L1CAM mediates vascular metastasis of hypoxic breast cancer cells to the lungs. Oncogene 31(14):1757–1770. https://doi.org/10.1038/onc.2011.365
33. Erler JT, Bennewith KL, Cox TR, Lang G, Bird D, Koong A, Le QT, Giaccia AJ (2009) Hypoxia-induced lysyl oxidase is a critical mediator of bone marrow cell recruitment to form the premetastatic niche. Cancer Cell 15 (1):35–44. https://doi.org/10.1016/j.ccr.2008.11.012
34. Wong CC, Gilkes DM, Zhang H, Chen J, Wei H, Chaturvedi P, Fraley SI, Wong CM, Khoo US, Ng IO, Wirtz D, Semenza GL (2011) Hypoxia-inducible factor 1 is a master regulator of breast cancer metastatic niche formation. Proc Natl Acad Sci U S A 108 (39):16369–16374. https://doi.org/10.1073/pnas.1113483108
35. Bonnet D, Dick JE (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med 3(7):730–737
36. Bar EE, Lin A, Mahairaki V, Matsui W, Eberhart CG (2010) Hypoxia increases the expression of stem-cell markers and promotes clonogenicity in glioblastoma neurospheres. Am J Pathol 177(3):1491–1502. https://doi.org/10.2353/ajpath.2010.091021
37. Ohnishi S, Maehara O, Nakagawa K, Kameya A, Otaki K, Fujita H, Higashi R,

Takagi K, Asaka M, Sakamoto N, Kobayashi M, Takeda H (2013) hypoxia-inducible factors activate CD133 promoter through ETS family transcription factors. PLoS One 8(6):e66255. https://doi.org/10.1371/journal.pone.0066255

38. Mathieu J, Zhang Z, Zhou W, Wang AJ, Heddleston JM, Pinna CM, Hubaud A, Stadler B, Choi M, Bar M, Tewari M, Liu A, Vessella R, Rostomily R, Born D, Horwitz M, Ware C, Blau CA, Cleary MA, Rich JN, Ruohola-Baker H (2011) HIF induces human embryonic stem cell markers in cancer cells. Cancer Res 71(13):4640–4652. https://doi.org/10.1158/0008-5472.CAN-10-3320
39. Ma B, Chen Y, Chen L, Cheng H, Mu C, Li J, Gao R, Zhou C, Cao L, Liu J, Zhu Y, Chen Q, Wu S (2015) Hypoxia regulates Hippo signalling through the SIAH2 ubiquitin E3 ligase. Nat Cell Biol 17(1):95–103. https://doi.org/10.1038/ncb3073
40. Xiang L, Gilkes DM, Hu H, Takano N, Luo W, Lu H, Bullen JW, Samanta D, Liang H, Semenza GL (2014) Hypoxia-inducible factor 1 mediates TAZ expression and nuclear localization to induce the breast cancer stem cell phenotype. Oncotarget 5(24):12509–12527. https://doi.org/10.18632/oncotarget.2997
41. Wang Y, Liu Y, Malek SN, Zheng P, Liu Y (2011) Targeting HIF1alpha eliminates cancer stem cells in hematological malignancies. Cell Stem Cell 8(4):399–411. https://doi.org/10.1016/j.stem.2011.02.006
42. Comerford KM, Wallace TJ, Karhausen J, Louis NA, Montalto MC, Colgan SP (2002) Hypoxia-inducible factor-1-dependent regulation of the multidrug resistance (MDR1) gene. Cancer Res 62(12):3387–3394
43. Gottesman MM, Fojo T, Bates SE (2002) Multidrug resistance in cancer: role of ATP-dependent transporters. Nat Rev Cancer 2(1):48–58. https://doi.org/10.1038/nrc706
44. Rohwer N, Cramer T (2011) Hypoxia-mediated drug resistance: novel insights on the functional interaction of HIFs and cell death pathways. Drug Resist Updat 14 (3):191–201. https://doi.org/10.1016/j.drup.2011.03.001
45. Lu H, Samanta D, Xiang L, Zhang H, Hu H, Chen I, Bullen JW, Semenza GL (2015) Chemotherapy triggers HIF-1-dependent glutathione synthesis and copper chelation that induces the breast cancer stem cell phenotype. Proc Natl Acad Sci U S A 112(33):E4600–E4609. https://doi.org/10.1073/pnas.1513433112
46. Casazza A, Laoui D, Wenes M, Rizzolio S, Bassani N, Mambretti M, Deschoemaeker S, Van Ginderachter JA, Tamagnone L, Mazzone M (2013) Impeding macrophage entry into hypoxic tumor areas by Sema3A/Nrp1 signaling blockade inhibits angiogenesis and restores antitumor immunity. Cancer Cell 24 (6):695–709. https://doi.org/10.1016/j.ccr.2013.11.007
47. Chiu DK, Xu IM, Lai RK, Tse AP, Wei LL, Koh HY, Li LL, Lee D, Lo RC, Wong CM, Ng IO, Wong CC (2016) Hypoxia induces myeloid-derived suppressor cell recruitment to hepatocellular carcinoma through chemokine (C-C motif) ligand 26. Hepatology 64(3):797–813. https://doi.org/10.1002/hep.28655
48. Du R, Lu KV, Petritsch C, Liu P, Ganss R, Passegue E, Song H, Vandenberg S, Johnson RS, Werb Z, Bergers G (2008) HIF1alpha induces the recruitment of bone marrow-derived vascular modulatory cells to regulate tumor angiogenesis and invasion. Cancer Cell 13(3):206–220. https://doi.org/10.1016/j.ccr.2008.01.034
49. Grimshaw MJ, Wilson JL, Balkwill FR (2002) Endothelin-2 is a macrophage chemoattractant: implications for macrophage distribution in tumors. Eur J Immunol 32(9):2393–2400. https://doi.org/10.1002/1521-4141(200209)32:9<2393::AID-IMMU2393>3.0.CO;2-4
50. Leek RD, Hunt NC, Landers RJ, Lewis CE, Royds JA, Harris AL (2000) Macrophage infiltration is associated with VEGF and EGFR expression in breast cancer. J Pathol 190 (4):430–436. https://doi.org/10.1002/(SICI)1096-9896(200003)190:4<430::AID-PATH538>3.0.CO;2-6
51. Matschurat S, Knies UE, Person V, Fink L, Stoelcker B, Ebenebe C, Behrensdorf HA, Schaper J, Clauss M (2003) Regulation of EMAP II by hypoxia. Am J Pathol 162 (1):93–103. https://doi.org/10.1016/S0002-9440(10)63801-1
52. Murdoch C, Giannoudis A, Lewis CE (2004) Mechanisms regulating the recruitment of macrophages into hypoxic areas of tumors and other ischemic tissues. Blood 104 (8):2224–2234. https://doi.org/10.1182/blood-2004-03-1109
53. Tripathi C, Tewari BN, Kanchan RK, Baghel KS, Nautiyal N, Shrivastava R, Kaur H, Bhatt ML, Bhadauria S (2014) Macrophages are recruited to hypoxic tumor areas and acquire a pro-angiogenic M2-polarized phenotype via hypoxic cancer cell derived cytokines Oncostatin M and Eotaxin. Oncotarget 5 (14):5350–5368. https://doi.org/10.18632/oncotarget.2110

54. Zhang X, Mosser DM (2008) Macrophage activation by endogenous danger signals. J Pathol 214(2):161–178. https://doi.org/10.1002/path.2284
55. Facciabene A, Peng X, Hagemann IS, Balint K, Barchetti A, Wang LP, Gimotty PA, Gilks CB, Lal P, Zhang L, Coukos G (2011) Tumour hypoxia promotes tolerance and angiogenesis via CCL28 and T(reg) cells. Nature 475 (7355):226–230. https://doi.org/10.1038/nature10169
56. Bosco MC, Reffo G, Puppo M, Varesio L (2004) Hypoxia inhibits the expression of the CCR5 chemokine receptor in macrophages. Cell Immunol 228(1):1–7. https://doi.org/10.1016/j.cellimm.2004.03.006
57. Sica A, Saccani A, Bottazzi B, Bernasconi S, Allavena P, Gaetano B, Fei F, LaRosa G, Scotton C, Balkwill F, Mantovani A (2000) Defective expression of the monocyte chemotactic protein-1 receptor CCR2 in macrophages associated with human ovarian carcinoma. J Immunol 164(2):733–738
58. Kuang DM, Zhao Q, Peng C, Xu J, Zhang JP, Wu C, Zheng L (2009) Activated monocytes in peritumoral stroma of hepatocellular carcinoma foster immune privilege and disease progression through PD-L1. J Exp Med 206 (6):1327–1337. https://doi.org/10.1084/jem.20082173
59. Ye LY, Chen W, Bai XL, Xu XY, Zhang Q, Xia XF, Sun X, Li GG, Hu QD, Fu QH, Liang TB (2016) Hypoxia-induced epithelial-to-mesenchymal transition in hepatocellular carcinoma induces an immunosuppressive tumor microenvironment to promote metastasis. Cancer Res 76(4):818–830. https://doi.org/10.1158/0008-5472.CAN-15-0977
60. Chiu DK, Tse AP, Xu IM, Di Cui J, Lai RK, Li LL, Koh HY, Tsang FH, Wei LL, Wong CM, Ng IO, Wong CC (2017) Hypoxia inducible factor HIF-1 promotes myeloid-derived suppressor cells accumulation through ENTPD2/CD39L1 in hepatocellular carcinoma. Nat Commun 8(1):517. https://doi.org/10.1038/s41467-017-00530-7

Chapter 7

Determination of Polarization of Resident Macrophages and Their Effect on the Tumor Microenvironment

Ioannis S. Pateras and Tomer Cooks

Abstract

Interactions between tumor cells and their microenvironment have been long established as a cardinal hallmark of tumorigenesis and metastasis. To that end, tumor-associated macrophages (TAMs) have been studied extensively and were found to be typically correlated with poor prognosis in various cancers. TAMs are key elements of cancer-associated inflammation promoting cancer progression by increasing angiogenesis, inducing immunosuppression of the tumor tissue, and remodeling the extracellular matrix favoring invasion and metastasis. Since resident macrophages are characterized by substantial diversity and plasticity, understanding their polarization patterns in response to microenvironmental cues is a prime focus in the field. This chapter demonstrates an efficient manner to characterize polarization patterns of macrophages inside tumor tissues.

Key words Resident macrophages, Tumor-associated macrophages, Polarization, Immunohistochemistry, Flow cytometry

1 Introduction

The immense plasticity of macrophages plays a significant role in tissue homeostasis, completing essential regulatory functions as well as pathogen protection [1] recalling Metchnikoff "I rather believe that the essence of an inflammation lies in the phagocyte attack of solid pathogenic substances, be it a weakened or dead cell, a bacterium or any other foreign body." Macrophages are sensitive to an ensemble of extracellular and intercellular signals that dictate their regulatory phenotype [2]. For decades, it was believed that tissue macrophages originate from the adult hematopoietic system, derived from stem cells in the bone marrow and differentiating via monocytic precursors in the blood. It is now proven that many tissue-resident macrophage populations maintain their populations through in situ proliferation by tissue-specific mechanisms [3–6]. In addition, macrophages are involved in various diseases

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_7,

including cancer, where tumor-associated macrophages (TAMs) were reported to typically correlate with poor prognosis [7].

Often, solid tumors are compared to organ-like structures composed of cancer cells, as well as immune and mesenchymal cells, which continuously interact and affect each other. Mirroring the Th1/Th2 dichotomy of T lymphocytes, a simplistic polarization of macrophages into M1 and M2 groups is commonly used [6, 8]. The pro-inflammatory or classically activated macrophages (M1) are induced by mediators such as IFN-γ and TNF-α and release pro-inflammatory cytokines such as IL-12 and IL-23. They are important in protecting the tissue from intracellular pathogens like bacteria and viruses. Contrary, alternatively activated macrophages (M2) are involved in repair, anti-inflammatory, and nourishing mechanisms such as tissue remodeling, angiogenesis, and wound healing [6, 8–10]. In this chapter, we address two major approaches to distinguish the polarization patterns inside tissues. By flow cytometry means, a physical separation of live and fixed macrophagic populations is allowed. Using such separation to further sort and study subpopulations may be instrumental in investigating tissue-specific and tumor-specific traits using advanced methods of gene expression and single cell sequencing. Importantly, we also discuss in detail immunohistochemistry (IHC) as a powerful application implementing the fundamental principle of specificity and affinity of antibody-antigen interaction to visualize the localization of cellular components (antigens) in cells and tissues. IHC provides a rapid approach with favorable cost-effective ratio qualitative (existence and distribution) and quantitative (number of macrophages per power field) data. Within this frame, by employing serial section analysis and multiple staining on a single tissue section, it becomes feasible to detect multiple antigens in the same context.

Here, we demonstrate two IHC detection procedures that are employed for the study of TAMs on routinely processed formalin-fixed paraffin sections: (a) the traditional biotin-avidin immunoenzymatic method and (b) the more recent polymer-based assay that does not react with endogenous biotin. The rationale behind both methodologies is identical, encompassing the following steps: (1) deparaffinization and rehydration, (2) antigen unmasking, (3) blocking non-specific staining (by abrogating activity of endogenous enzymes and blocking non-specific interaction with endogenous Fc receptors), (4) incubation with the primary antibody, (5) incubation with the labeled secondary antibody, (6) incubation with HRP conjugate complex, (7) visualization of the staining pattern employing DAB chromogen, (8) counterstain in hematoxylin, and (9) dehydration and mounting. Notably, IHC can be employed in conjunction with other conventional techniques including in situ hybridization in the same tissue section.

Given that TAMs generally display an M2-like phenotype [1], they can be detected by employing anti-CD206 and anti-CD163 that recognize mannose and hemoglobin/haptoglobin receptors, respectively. Both receptors are expressed by M2 (alternative activated) macrophages [11, 12]. However, additional antibodies detecting macrophages including anti-CD64 (Fcγ receptor 1) detecting M1 macrophages as well as anti-CD14, anti-CD68 (PGM1), and anti-CD68 (KP1) can be also incorporated in the IHC analysis.

2 Materials

1. 100 mm petri dishes (Corning).
2. Surgical scalpels.
3. Forceps and surgical scissors for animal dissection.
4. 15 and 50 mL conic tubes (Corning).
5. 0.5 and 1.5 mL collection tubes (Eppendorf).
6. 10× red blood cell (RBC) lysis buffer (eBioscience).
7. Collagenase D (Sigma-Aldrich).
8. DNase I (Sigma-Aldrich).
9. 70 μm cell strainers (Corning).
10. Isoflurane.
11. Blocking antibody (Ab) CD16/CD32 (R&D).
12. Antibodies for FACS: $CD45^+$, $CD11b^+$, $F4/80^+$, $Ly6C^-$, $Ly6G^-$, $CD3^-$, $CD19^-$, and $NK1.1^-$ (R&D), and for IHC, CD206 (Abcam), CD163 (Leica), goat anti-mouse biotin conjugated, and goat anti-rabbit biotin conjugated (Thermo Fisher Scientific).
13. DAPI (Thermo Fisher Scientific).
14. gentleMACS™ Dissociator (Miltenyi Biotec).
15. C tubes (Miltenyi Biotec).
16. Vortex.
17. Lab centrifuge.
18. Lab bench centrifuge.
19. Flow cytometer.
20. Flow sorter.
21. Pipetman.
22. Pipettes.
23. Beaker.
24. Biotin and streptavidin blocking kit (Vector Labs).

25. Coplin staining jars.
26. Cover slips.
27. Glass trough.
28. Goat serum (Thermo Fisher Scientific).
29. Hematoxylin.
30. Hydrogen peroxide 30%.
31. IHC kit (providing the solutions for the blocking of non-specific interaction, the labeled secondary antibody conjugated to a polymer backbone, the immune complex that forms the bridge between the secondary antibody and the chromogen mean and the corresponding chromogen) (Thermo Fisher Scientific).
32. Immune slide staining tray.
33. Lab bench centrifuge.
34. Lab oven.
35. Microwave oven.
36. Mounting medium.
37. Refrigerator.
38. Staining jar.
39. Staining racks.
40. Steamer.
41. Streptavidin/biotin blocking kit (Vector Labs).
42. Streptavidin HRP (horseradish peroxidase) conjugate (Millipore).
43. Ultrapure water.
44. Volumetric cylinders.
45. 3,3′-diaminobenzidine (DAB).

2.1 Solutions

1. Phosphate-buffered saline (PBS).
2. Ethanol (EtOH).
3. Xylene.
4. FACS buffer: 1× PBS, 2% bovine serum, 1 mM EDTA, 0.1% sodium azide.
5. Collection medium: DMEM or RPMI 10% v/v FBS, 1% penicillin/streptomycin.
6. 1× RBC lysis buffer: dilute 10× RBC buffer in distilled water.
7. 1× TBS (Tris-buffered saline) solution: dissolve 8.1 g of dry sodium chloride (NaCl) and 1 g of Tris(hydroxymethyl) aminomethane in 1 L of ultrapure water. Titrate to pH 7.6 with

HCl. PBS (phosphate-buffered saline) 1× solution may be used alternatively instead of 1× TBS.

8. 3% hydrogen peroxide (H_2O_2) solution: dilute 10 mL 30% hydrogen peroxide to 90 mL ultrapure water.
9. Citric acid solution: dissolve 2.1 g of dry sodium nitrate in 1 L ultrapure water. Titrate to pH 6.0 with NaOH.
10. EDTA solution: dilute 2.2 mL from stock EDTA (0.5 M) in 1 L ultrapure water. Titrate to pH 8.0–8.5 with NaOH.
11. DAB solution: dissolve 50 mg of dry DAB dilute in 100 mL 1× TBS. Add 30 μL 3% H_2O_2 solution. Avoid unnecessary exposure to light. Safety precautions in the handling of DAB are necessary since it is a suspected carcinogen.

3 Methods

3.1 Isolation of Resident Macrophages from Lung Mouse Tissue (see Notes 1–5)

3.1.1 Tissue Harvesting

1. Anesthetize the mouse using CO_2 or isoflurane.
2. Spray down mouse with 70% ethanol, and make initial incision just below the rib cage. Cut through the skin and connective tissue across the mouse. Make lateral incisions on each side up to the neck of the mouse.
3. Perfuse 10 mL of cold 1× PBS through the right ventricle until the lungs are cleared of blood (to allow blood to leak out, cut a slit in the left ventricle).
4. Carefully remove the rest of the rib cage (the more bone removed, the better) and other tissues to expose the trachea; place the forceps under the trachea to keep it exposed (making sure to separate the trachea from the esophagus).
5. Dissect out the lungs by gently tugging on the trachea while snipping away the connective tissue; leave the lungs intact.
6. Put the lung in a 50 mL conical with enough 1× PBS to cover 5× of the entire tissue.

3.1.2 Tissue Dissociation into Single Cells

1. Rinse tissue in a petri dish containing 1× PBS. Separate each lobe using a scalpel.
2. Transfer a maximum of 450 mg lung tissue to a gentleMACS C tube containing 2.5 mL PBS.
3. Add a mix of collagenase D and DNase I solution to a final concentration of 2 mg/mL (collagenase) and 40 U/mL (DNase) (*see* **Note 6**).
4. Tightly close the C tube and attach it onto the gentleMACS Dissociator. Then run the program "m_lung_01."
5. Incubate for 30 min at 37 °C with automated rotation (*see* **Note 7**).

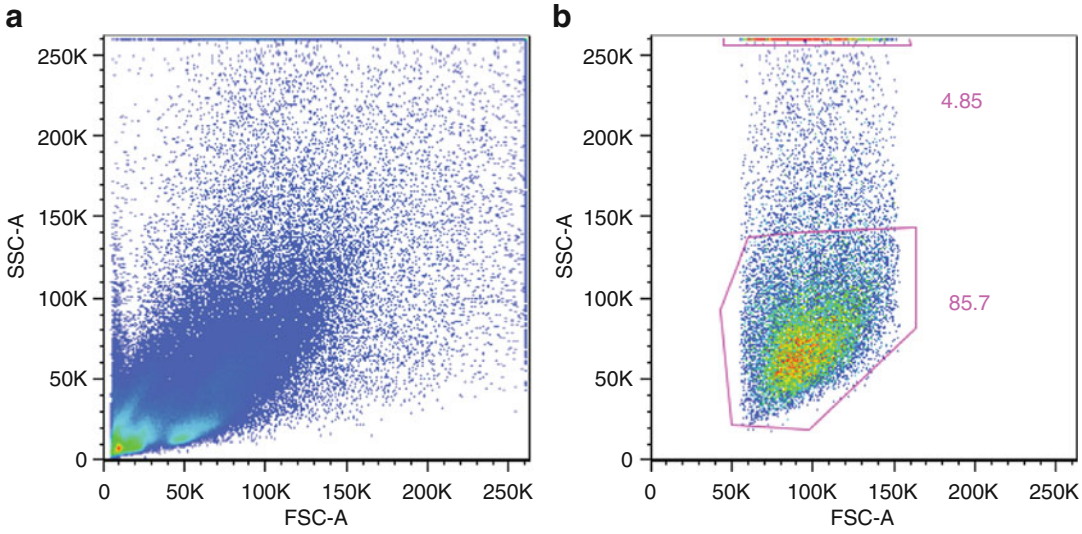


Fig. 1 Typical side and forward scatter of cells from dissociated mouse lung. (**a**) A wide range of cellular populations from the lung tissue, including epithelial, immune, and connective tissue. (**b**) Using gating strategies requires ensuring not overlooking populations beyond the scope of the forward and side scatter as in the example presented

6. Put the tube back to the gentleMACS Dissociator, and run the program "m_lung_02" (*see* **Note 8**).
7. Prepare a 50 mL tube for collecting filtered cells by replacing the cap with a 70 μm mesh cell strainer.
8. Use p1000 pipette to collect dissociated solution from the C tube, and apply the cells to the cell strainer.
9. Wash the cell strainer with 2.5 mL 1× PBS.
10. Centrifuge the cells to a pellet in a 50 mL tube at 500 × *g* for 10 min.
11. Aspirate the supernatant, and resuspend the pellet in 3 mL of 1× RBC lysis buffer.
12. Incubate for 4–5 min at room temperature.
13. Stop the lysis reaction by adding 20 mL of 1× PBS.
14. Centrifuge at 500 × *g* for 5 min, and decant the supernatant.
15. Resuspend cells in 2 mL FACS buffer, and perform a cell count and viability analysis. Figure 1 demonstrates a typical forward and side scatter of a single cell dissociation of a mouse lung tissue.

3.1.3 Blocking and Preparation for Flow Cytometry

1. Incubate cells with 5–10 μg/mL of purified anti-mouse CD16/CD32 antibody to block non-specific staining. Vortex briefly, and incubate on ice for a minimum of 30 min (vortex every 10 min).

2. Centrifuge for 5 min at 500 × *g* (4 °C). Discard the supernatant, and resuspend the pellet in 500 μL of staining buffer.
3. Add the antibody and incubate on ice for 1 h (protect from light).
4. Centrifuge for 5 min at 500 × *g* (4 °C). Discard the supernatant, and resuspend the pellet in 400 μL of staining buffer.
5. Filter the sample using filter top 5 mL polystyrene tubes (40 μm) to prevent clumps of cells clogging the cytometer (*see* **Note 9**).
6. Add 300 nM of DAPI to gate viable cells.
7. The following markers should be used for a gating strategy to isolate specific macrophages subpopulations: $CD45^+$, $CD11b^+$, $F4/80^+$, $Ly6C^-$, $Ly6G^-$, $CD3^-$, $CD19^-$, and $NK1.1^-$.
8. After optimizing these settings, more distinctive macrophagic subpopulations can be gated, such as $CD206^+$ and $CD163^+$.
9. After setting, gating, and compensation are completed, cells can be sorted and can be further used for DNA, RNA, and protein extraction as well as for biological function experiments.

3.2 Detection of Resident Macrophages and TAM Using Immunohistochemistry

3.2.1 Deparaffinization and Rehydration

1. Melt paraffin by dissolving it in xylene followed by rehydration in decreasing series of ethanol solutions. Place sections in staining rack (*see* **Notes 10** and **11**).
2. Incubate in an oven at 60 °C for 30 min.
3. Wash in xylene for 10 min, and then move the sections to a new glass trough with fresh xylene for additional 10 min.
4. Incubate in pure ethanol for 15 min.
5. Incubate in 95% ethanol for 10 min.
6. Incubate in 80% ethanol for 5 min.
7. Incubate in 70% ethanol for 5 min.
8. Incubate in 50% ethanol for 5 min.
9. Wash in 1× TBS for 5 min. From this point on, avoid drying out the sections.

3.2.2 Antigen Retrieval

1. Preheat the buffer solution for antigen retrieval [citric acid 1× (pH 6.0) or EDTA (pH, 8–8.5)] for 5 min (optional) (*see* **Note 12**).
2. Incubate sections in the buffer solution in microwave for 25 min in 700 watts or steamer for 40–50 min. To avoid drying out the slides, incubation should be performed in a humid atmosphere.
3. Allow slides to cool within the buffer solution and wash in 1× TBS for 5 min.

3.2.3 Blocking

1. A humid chamber is required to avoid evaporation of the blocking solution. Sections are incubated with the corresponding solution in an immune slide staining tray where layers of damp paper towels are placed at the bottom.
2. For blocking endogenous hydrogen peroxidase, incubate with 3% hydrogen peroxide solution for 15 min, in the dark, at room temperature. Commercially available IHC kits provide the equivalent blocking solution.
3. Wash in 1× TBS for 2 × 5 min.
4. Blocking of endogenous biotin employing a streptavidin/biotin blocking kit according to the manufacturer's instructions. This step can be omitted when employing commercially available IHC kits which are not biotin−/avidin-based immunoassays.
5. For blocking non-specific interaction with endogenous Fc receptors (FcR), an optimization step is necessary (*see* **Note 13**). For most antibodies, incubate with 5–10% normal goat serum diluted in 1× TBS for 10 min at room temperature.
6. Wash in 1× TBS for 5 min.

3.2.4 Incubation with Primary and Secondary Antibodies

1. Primary antibody incubation should be optimized accordingly (*see* **Note 14**). Sections are incubated with the corresponding antibody diluted in 1× TBS in a humid chamber as previously described. To further reduce non-specific staining, primary antibodies can be diluted in 1× TBS containing 5–10% goat serum (or a similar blocking medium).
2. Wash in 1× TBS for 3 × 5 min. It is important to remove unbound primary antibody prior to incubation with the secondary antibody.
3. Incubate with the labeled secondary antibody. Sections should be incubated in a humid atmosphere employing a biotin conjugated secondary antibody (*see* **Note 15**). IHC kits may include an equivalent to labeled secondary antibody.
4. Wash in 1× TBS for 3 × 5 min in order to remove unbound secondary antibody.

3.2.5 Incubation with HRP Conjugate Complex

1. Incubate sections in a humid chamber, and employ streptavidin HRP conjugate diluted in TBS 1× according to the manufacturer's instructions (*see* **Note 16**). Streptavidin HRP complex binds to the biotinylated secondary antibody, thus joining the peroxidase moiety to the site of target antigen. IHC kits include the appropriate immune complexes that are usually HRP conjugated but do not depend on biotin/avidin interaction. Avoid unnecessary exposure to light.

2. Wash in 1× TBS for 3 × 5 min in order to remove unbound complexes.

3.2.6 Visualization of Immunoreactivity Employing DAB Chromogen

To visualize with DAB chromogen (*see* **Note 17**), embed sections in DAB solution. Staining procedure is carried out in Coplin staining jars. Incubation time should be optimized. IHC kits include all necessary materials for DAB staining. To stop the reaction, transfer the sections to another Coplin staining jar filled with tap water, and rinse gently with running tap water.

3.2.7 Counterstaining in Hematoxylin

1. Immerse sections in a glass trough with diluted filtered hematoxylin (in distilled water). Incubation time and dilution of hematoxylin should be optimized.
2. Rinse sections with running tap water. Before proceeding to the next step, observe the sections through the microscope briefly (ensuring that they do not dry out) to examine the color contrast. In case the desired color contrast is not reached, repeat **step 1**.

3.2.8 Dehydration and Mounting

1. Incubate in 50% ethanol for 5 min.
2. Incubate in 70% ethanol for 5 min.
3. Incubate in 80% ethanol for 5 min.
4. Incubate in 95% ethanol for 10 min.
5. Incubate in pure ethanol for 15 min.
6. Wash in xylene for 10 min. Then move the sections to a new glass trough with fresh xylene for additional 10 min.
7. Mount sections by employing a mounting media, and cover them by lowering the cover slips avoiding air bubble trapping.

For the evaluation of CD206 and CD163 staining, first you should examine for the positivity of these antigens (Fig. 2). *For*

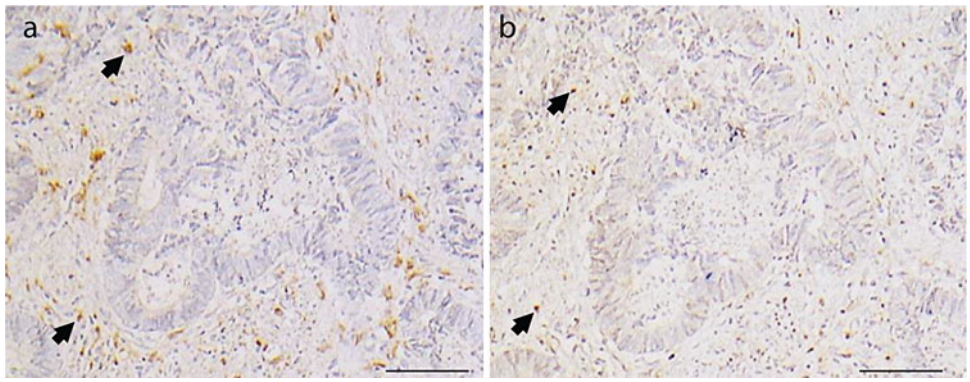


Fig. 2 Immunohistochemistry staining of serial sections for CD163 (on the left) and CD206 (on the right) in cases with sporadic colorectal carcinomas. Arrows denote positive staining. Scale bar: 100 μm

Table 1
Troubleshooting for the IHC procedure

Problem	Possible reason	Solution
Absence or low signal	Improper fixative conditions; over-fixation may be a possible cause	Modify the fixative conditions; reduce the fixation time
	Absence of target antigen	Choose a positive control
	Compromised antigen retrieval	Optimize antigen retrieval
	Low concentration of primary antibody and/or improper incubation time and temperature	Increase concentration of primary antibody; increase incubation time and/or incubation temperature
	Secondary antibody does not react with the primary antibody	Choose the appropriate secondary antibody
	Detection reagents expired	Verify the expiration date of the detection reagents
High background	Insufficient blocking	Increase the duration of blocking solution, or alter the blocking medium
	Increased concentration of secondary antibody; cross-reactivity with the secondary antibody	Reduce the concentration of the secondary antibody; include a control case omitting the primary antibody to examine the reactivity of the secondary antibody
	Decreased washes with TBS 1×	Increase the washes with TBS 1×

suboptimal immunostaining, a troubleshooting guide is provided (Table 1). Upon positive staining, it is very important to look into the distribution pattern of CD206+ and CD163+ cells. In other words examine for their presence in the invasive front and within the tumor. An insightful way to quantify them is to count the number of positive cells per high-power field (HPF) (magnification 400×) manually. Alternatively, an image analysis program may be employed.

4 Notes

1. The FACS protocol has been used extensively for the isolation of TAM from mouse lung tumors and from healthy lungs. For isolation of resident macrophages and TAM from other tissues, this protocol needs to be optimized specifically.
2. When multicolor staining is applied, fluorescence compensation is obligatory.
3. Beyond the basic markers detailed in these protocols, to delineate further subpopulations, additional macrophages and TAM markers could be used. These include also various cytokines and transcription factors.
4. In addition to the experimental samples, negative unstained controls and appropriate isotype controls must be prepared.

5. Allow 4–5 h to conclude the FACS protocol.
6. Make fresh before use. Most TAM extracellular markers are stable and sustain enzymatic digestion, but it is fundamental to optimize the digestion protocol before starting.
7. Some collagenase mixtures may contain additional proteases that can cleave extracellular markers. Liberase is another optional enzyme in such a protocol. Do not over-digest the tissue to avoid loss of cellular yield.
8. Optional: if enzymatic digestion is too harsh, dissociate the tissue in the gentleMACS using PBS only, and follow with a standard Ficoll-Histopaque separation. Collect the buffy coat and continue to the cell strainer step.
9. Primary digested tissues are particularly prone to clogging the machines. Before analysis all samples should be filtered using a 40 μm filter cap polypropylene tube.
10. Positive and negative controls for the target antigen are necessary. For optimal performance it is important to perform serial section analysis omitting individual reagents (including the primary antibody) to verify the specificity of the signal. For anti-CD206, human as well as mouse liver may serve as positive control. For anti-CD163, human liver can be employed as positive control.
11. For frozen sections there are certain differences from the procedure described above. First a fixation step should precede the IHC procedure. Sections are usually fixed for 20–30 min at room temperature employing 10% neutral buffered formalin (NBF) or acetone. Second deparaffinization and rehydration are omitted. Third antigen retrieval needs to be optimized due to the friable nature of these sections.
12. Antigen retrieval should be optimized according to the specific setting. The following parameters should be examined: (a) retrieval buffer solution (i.e., with low or high pH), (b) type of equipment (including microwave oven, steamer, pressure cooker, etc.), and (c) incubation time for antigen retrieval. Apart from heat-mediated antigen retrieval, proteolytic enzyme methods are occasionally used alone or in combination with heat-mediated methods.
13. Blocking endogenous Fc receptors is also context dependent. The blocking solution should not interfere with the primary antibody. It is usually raised from the species in which the secondary antibody is raised. Occasionally, blocking solutions may be employed, including BSA in various concentrations (1% is typically used), as well as goat serum.
14. The optimal dilution, incubation time, and temperature should be determined for each primary antibody.

15. Taking into consideration the commercially available clones of the antibodies mentioned in the Introduction, incubation with goat anti-mouse biotin conjugated or goat anti-rabbit biotin conjugated at 1:200–1:300 dilution according to the origin of the primary antibody, for 30 min at 37 °C, is a good starting point.
16. Peroxidase is a very common enzyme label. An alternative enzyme is alkaline phosphatase.
17. The most popular visualization substrate for peroxidase-based methods is DAB, although additional substrates may be employed. Be aware that DAB is a suspected carcinogen, and therefore use safety measures when handling it.

References

1. Biswas SK, Mantovani A (2010) Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm. Nat Immunol 11 (10):889–896
2. Gordon S, Pluddemann A (2017) Tissue macrophages: heterogeneity and functions. BMC Biol 15(1):53
3. Franklin RA, Li MO (2016) Ontogeny of tumor-associated macrophages and its implication in cancer regulation. Trends Cancer 2 (1):20–34
4. Hoeffel G, Ginhoux F (2018) Fetal monocytes and the origins of tissue-resident macrophages. Cell Immunol S0008-8749 (18):30001–30007
5. Okabe Y, Medzhitov R (2016) Tissue biology perspective on macrophages. Nat Immunol 17 (1):9–17
6. Lahmar Q, Keirsse J, Laoui D (2016) Tissue-resident versus monocyte-derived macrophages in the tumor microenvironment. Biochim Biophys Acta 1865(1):23–34
7. Takeya M, Komohara Y (2016) Role of tumor-associated macrophages in human malignancies: friend or foe? Pathol Int 66(9):491–505
8. Chanmee T, Ontong P, Konno K (2014) Tumor-associated macrophages as major players in the tumor microenvironment. Cancers (Basel) 6(3):1670–1690
9. Rhee I (2016) Diverse macrophages polarization in tumor microenvironment. Arch Pharm Res 39(11):1588–1596
10. Mantovani A, Marchesi F, Malesci A (2017) Tumour-associated macrophages as treatment targets in oncology. Nat Rev Clin Oncol 14 (7):399–416
11. Murdoch C, Muthana M, Coffelt SB (2008) The role of myeloid cells in the promotion of tumour angiogenesis. Nat Rev Cancer 8 (8):618–631
12. Cooks T, Pateras IS, Jenkins LM (2018) Mutant p53 cancers reprogram macrophages to tumor supporting macrophages via exosomal miR-1246. Nat Commun 9(1):771

Chapter 8

Quantitation of Macropinocytosis in Cancer Cells

Koen M. O. Galenkamp, Basheer Alas, and Cosimo Commisso

Abstract

Macropinocytosis has emerged as an important nutrient supply pathway that sustains cell growth of cancer cells within the nutrient-poor tumor microenvironment. By internalizing extracellular fluid through this bulk endocytic pathway, albumin is supplied to the cancer cells, which, after degradation, serves as an amino acid source to meet the high nutrient demands of these highly proliferating cells. Here, we describe a streamlined protocol for visualization and quantitation of macropinosomes in adherent cancer cells grown in vitro. The determination of the "macropinocytic index" provides a tool for measuring the extent to which this internalization pathway is utilized within the cancer cells and allows for comparison between different cell lines and treatments. The protocol provided herein has been optimized for reproducibility and is readily adaptable to multiple conditions and settings.

Key words Macropinocytosis, Macropinosome, Quantitation, Quantification, Cancer, Endocytosis, Nutrient uptake, Membrane ruffles, Fluorescence microscopy, Macropinocytic index

1 Introduction

Macropinocytosis, which is derived from the Greek words for large (*makros*), drink (*pino*), container (*kytos*), and process (*−osis*), involves the bulk fluid-phase uptake of extracellular fluids [1, 2]. This endocytic pathway is classically described to occur after growth factor stimulation (e.g., EGF and PDGF); however, in cancer cells, the process can be driven by oncogenes, such as *HRAS* and *KRAS* (Fig. 1) [3–7]. In cancer cells, macropinocytosis serves as a nutrient supply mechanism used to meet elevated metabolic demands and is necessary to sustain tumor growth in nutrient-depleted tumor microenvironments [3, 8–10]. The scavenging of albumin from the extracellular fluid supplies the cells with a protein source to generate protein-derived amino acids through lysosomal proteolysis. These protein-derived amino acids can fuel the TCA cycle and support the biosynthesis of macromolecules, such as nucleotides, lipids, and proteins.

Macropinocytosis is driven by small GTPase-mediated (i.e., Rac1 and Cdc42) actin remodeling, which causes protrusions of

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_8,

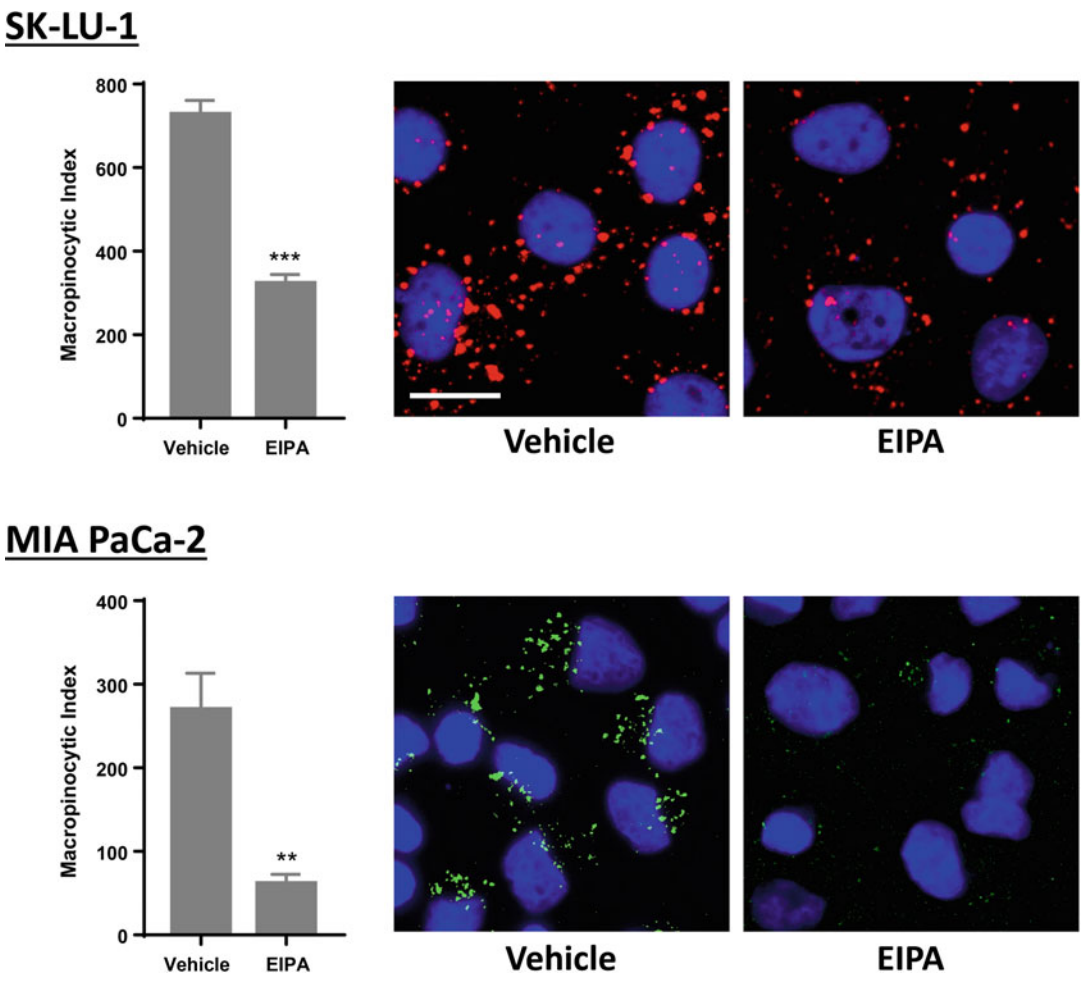


Fig. 1 Representative images of macropinocytosis in lung cancer and PDAC cells and the quantitation of the macropinocytic index as obtained by the described protocol. Mutant KRAS SK-LU-1 and MIA PaCa-2 cells were maintained in serum-free media and treated with vehicle or 75 μM EIPA for 30 min, after which macropinosomes were labeled by uptake of TMR-dextran or FITC-dextran, respectively. Scale bar = 20 μm, $^{**}p < 0.01$, $^{***}p < 0.001$ as assessed by *t*-test

the plasma membrane known as membrane ruffles [11, 12]. These membrane ruffles are motile structures that can fuse, thereby creating macropinocytic cups filled with extracellular fluid and solubilized molecules that act as cargo. Closure of the macropinocytic cups results in the formation of clathrin-independent endocytic vesicles with a diameter between 0.2 and 5 μm, which can either recycle back to the plasma membrane or undergo a maturation process and fuse with lysosomes. Due to the pH sensitivity of the small GTPases, intracellular acidification induced by inhibitors of Na^+/H^+ exchange, such as amiloride and its structural analogs (e.g., EIPA and HOE-694), has been shown to specifically suppress macropinocytosis (Fig. 1) [13, 14]. Thus, incorporation of these

inhibitors in the experimental setting provides a critical control for validating the presence of macropinocytosis in a cell line of interest.

Detection of macropinosomes can be achieved by loading the extracellular fluid with fluorophore-labeled high-molecular-weight dextran (i.e., 70 kDa), which due to size exclusion is only internalized through macropinocytosis and not via other endocytic pathways [15, 16]. When visualized by fluorescence microscopy, the macropinosomes can be recognized as clusters of intracellular fluorescent puncta (Fig. 1). In this protocol, we describe the steps necessary to specifically label macropinosomes, capture images of the macropinosomes using a standard fluorescent microscope, and finally, the quantitation of the extent of macropinocytosis—a parameter known as the "macropinocytic index" (Fig. 2) [15]. These approaches have been employed to explore the role of macropinocytosis in fueling cell metabolism and to screen for factors or agents that induce or inhibit this endocytic pathway [3, 7, 8, 15].

2 Materials

2.1 Acid-Washed Coverslips

1. 12 mm circular coverslips (0.13–0.17 mm thick).
2. 1 M HCl.
3. 95% EtOH.
4. dH_2O.
5. Large beaker.
6. Hot plate or water bath.
7. 10 cm cell culture dish.
8. Parafilm.

2.2 Culturing Cancer Cells

1. Adherent cancer cell line of interest.
2. Cell line-specific growth media with and without serum (complete and serum-free media, respectively).
3. Dulbecco's phosphate-buffered saline (DPBS) without calcium and magnesium.
4. Trypsin solution: 0.25% trypsin, 0.1% EDTA in HBSS without calcium, magnesium, and sodium bicarbonate.
5. 10 cm tissue culture-treated cell culture dish.
6. 24-well clear flat-bottom TC-treated multiwell plate.
7. Dumont #5 forceps.
8. Vacuum aspirator.
9. Cell culture hood.
10. Humidified 37 °C, 5% CO_2 cell culture incubator.

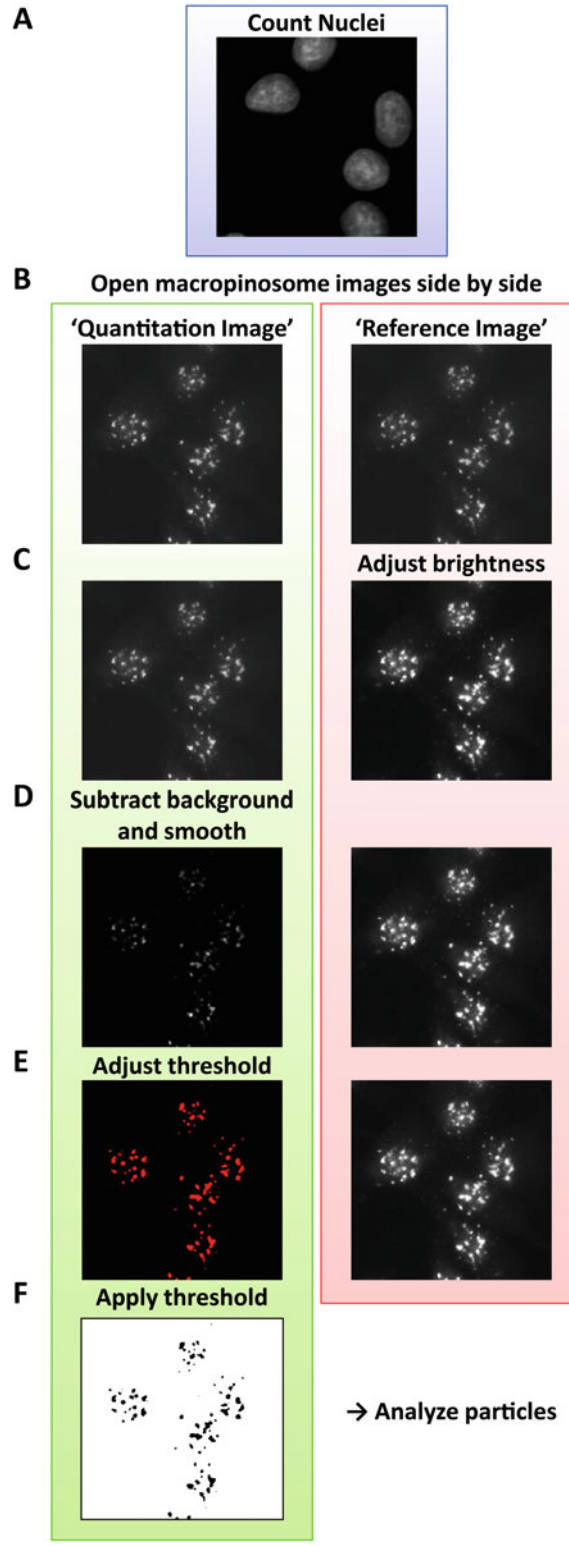


Fig. 2 Diagrammatic depiction of macropinosome quantitation using ImageJ and the steps as described in the protocol. (**a**) DAPI-stained nuclei are counted using

11. 37 °C water bath.
12. 15 mL centrifuge tube.
13. Centrifuge.

2.3 Macropinosome Labeling

1. 1× PBS (Phosphate-buffered saline).
2. 20 mg/mL 70 kDa lysine-fixable dextran labeled with FITC or tetramethylrhodamine (TMR) in 1× PBS.
3. 3.7% ACS grade formaldehyde in 1× PBS, freshly prepared on the day of the experiment.
4. 2 μg/mL DAPI in 1× PBS, freshly prepared on the day of the experiment.
5. Fluorescence mounting media.
6. Plain water-white microscope slides (75 mm × 25 mm and 0.90–1.10 mm thick).
7. Humidified 37 °C, 5% CO_2 cell culture incubator.
8. Wash bottle.
9. Dumont #5 forceps.
10. Vacuum aspirator.
11. Shaker.

2.4 Macropinosome Imaging

1. Fluorescence microscope with appropriate DAPI and FITC and/or TMR emission and excitation filters and equipped with a 40× or 63× objective.
2. 70% EtOH.
3. Q-tip.
4. Kimwipe.

2.5 Macropinosome Quantitation

1. ImageJ software (http://imagej.nih.gov/ij/).
2. Spreadsheet software (e.g., Excel).

Fig. 2 (continued) the ImageJ *Cell Counter* plug-in. (**b**) Two images of dextran uptake placed side by side. (**c**) For the "Reference Image," the brightness is adjusted such that macropinosomes are clearly visible. (**d**) From here on, the "Quantitation Image" is used for the remainder of the steps. The background is subtracted and the *Smooth* function is applied. (**e**) The threshold is adjusted such that all visible macropinosomes are labeled in red. (**f**) After applying the threshold, the image is converted to a binary image with the macropinosomes labeled in black. By using the *Analyze Particles* function, the total macropinosome area is obtained from the *Summary* window

3 Methods

3.1 Acid-Washed Coverslips

1. In a large beaker, place coverslips in 100 mL 1 M HCl, and heat at 56 °C for 24–48 h while covering the beaker with two layers of foil.
2. Then, wash the coverslips four times with distilled water followed by four washes with 95% EtOH.
3. For future use, store the acid-washed coverslips submerged in 95% EtOH at room temperature in a 10 cm cell culture dish sealed with parafilm.

3.2 Culturing Cancer Cells

All work should be performed in a sterile cell culture hood.

Media, DPBS, and trypsin solution should be preheated in a 37 °C water bath.

1. Culture the cancer cells until they are sub-confluent in a 10 cm cell culture dish.
2. Using forceps, place one circular acid-washed coverslip onto the bottom of each well of a 24-well multiwell plate (*see* **Note 1**).
3. Wash the coverslips twice with DPBS (0.5 mL per well).
4. Aspirate DPBS.
5. Gently rinse the cells with DPBS (5 mL per plate).
6. Add 1.5 mL trypsin solution to the plate, and incubate in a humidified 37 °C, 5% CO_2 cell culture incubator until cells have detached from the 10 cm dish (*see* **Note 2**).
7. Collect the detached cells in a 15 mL centrifuge tube with 10 mL of complete media.
8. Centrifuge the cell suspension for 3 min at 200 × *g*.
9. Aspirate the supernatant and resuspend the cell pellet in complete media.
10. Seed the cells on the coverslip-containing wells by adding 0.5 mL of cells suspended in complete growth media to each well (*see* **Note 3**).
11. Incubate the cells in a humidified 37 °C, 5% CO_2 cell culture incubator.

When the cells reach the desired 60–80% confluency:

12. Aspirate complete growth media, and perform one rinse with 0.5 mL DPBS per well.
13. After aspirating the DPBS, add 0.5 mL of the appropriate serum-free media to each well.
14. Incubate the cells for 16–24 h in a humidified 37 °C, 5% CO_2 cell culture incubator.

3.3 Macropinosome Labeling

At least 30 min prior to the labeling, pre-cool a PBS-filled wash bottle by placing it on ice or in a refrigerator.

1. For each well, aspirate the media, and add 190 μL of fresh serum-free media to the well (*see* **Note 4**).
2. Add 10 μL of 20 mg/mL fluorophore-labeled dextran solution to each well.
3. Mix thoroughly by moving the plate in a crosswise fashion.
4. Incubate for 30 min in a humidified 37 °C, 5% CO_2 cell culture incubator.

   Move the tissue culture plate to the bench and place on ice.
5. Carefully aspirate the dextran-containing media.
6. Quickly but gently perform five washes with ice-cold PBS using the wash bottle (*see* **Note 5**).
7. Add 0.35 mL formaldehyde (3.7% in 1× PBS) to each well.
8. Incubate for 30 min at room temperature in the dark.
9. Aspirate 3.7% formaldehyde and wash twice with 0.5 mL PBS.
10. Add 0.35 mL DAPI (2 μg/mL in 1× PBS) to each well.
11. Incubate for 15 min at room temperature in the dark.
12. Aspirate DAPI solution, and wash coverslips three times with PBS for 5 min while rocking on a shaker and protected from light.
13. Add one small drop of fluorescence mounting media per circular coverslip on a microscope slide (*see* **Note 6**).
14. Using forceps, carefully remove the circular coverslips from their wells, and place them facedown onto the drop of mounting media (*see* **Notes 6** and **7**).
15. Store the mounted coverslips on a flat surface in a dark environment for 16–24 h.
16. When the mounting media has dried, slides can be imaged or stored in a −20 °C freezer for up to 2 weeks.

3.4 Macropinosome Imaging

1. Use a Q-tip dipped in 70% ethanol to gently wipe off any excess mounting media or PBS from the surface of the circular coverslips, and then dry with a Kimwipe.
2. Using a fluorescence microscope with 40× of 63× magnification, capture an image of the DAPI stain, and save it in .tif format.
3. Next, without moving the specimen, capture an image in the channel that corresponds to the fluorophore-labeled dextran and save it in .tif format (*see* **Note 8**).
4. Repeat these **steps 10–20** times at different regions of the coverslip (*see* **Note 9**).

3.5 Macropinosome Quantitation (Fig. 2)

1. Open the first DAPI image in ImageJ, and use the counter tool to count nuclei by selecting *Plugins* → *Analyze* → *Cell Counter* (Fig. 2a) (*see* **Notes 10** and **11**).
2. Record the number of nuclei onto a spreadsheet.
3. Open the corresponding image of the dextran uptake to serve as the "Quantitation Image" (*see* **Note 11**).
4. Duplicate the "Quantitation Image" by selecting *Image* → *Duplicate*. This image will serve as a "Reference Image" and should be placed side by side to the "Quantitation Image" (Fig. 2b).
5. For the reference image, select: *Image* → *Adjust* → *Brightness/Contrast*, and adjust the maximum brightness slider such that you have a clear image of the macropinosomes (Fig. 2c).
6. For the quantitation image, select *Process* → *Subtract Background* → *Enter 10.0 pixels* for rolling ball radius, remove the checkmarks from all other options, and select *OK* (Fig. 2d).
7. Next, select: *Process* → *Smooth* (Fig. 2d).
8. Next, select: *Image* → *Adjust* → *Threshold*.
9. With the two images side by side, modify the minimum threshold value (upper slider) until you have successfully labeled the macropinosomes in red without labeling non-specific background (Fig. 2e).
10. Record the minimum threshold value for future use, only check *Dark Background* and convert the image to a binary image by selecting *Apply* (Fig. 2f).
11. To quantify the total macropinosome area, select *Analyze* → *Analyze Particles* → Only Check *Summarize* → Select *OK*.
12. From the *Summary* window, record *Total Area* value in your spreadsheet next to the nuclei counted. Total area divided by the number of nuclei for an image will give you the macropinocytic index.
13. Using the previously determined threshold value, follow the same steps for the rest of the .tif images to generate an average value for each coverslip and condition (*see* **Note 12**).

4 Notes

1. When necessary, excess EtOH can be removed by tapping the coverslip to the wall of the 10 cm storage dish, before placing the coverslip in the 24-well multiwell plate.
2. Confirming whether the cells have detached from the surface of the 10 cm dish can be achieved by visualizing the cells under a

microscope. Gently tapping the cell culture dish against the palm of your hand may aid in the detachment of loosely attached cells.

3. For each cell line, the seeding density should be optimized such that it accounts for the proliferation rate of the cell line in complete and serum-free growth media. The optimal cell density is such that the cells will reach 60–80% confluency 3 days post-seeding at which time the cells can be subjected to the serum starvation preceding the macropinosome labeling. Some cell lines may continue to proliferate under serum-free conditions, which should be accounted for to avoid overconfluence at the time of macropinosome labeling.
4. Depending on the experimental setting, the media can be replaced with fresh media, or conditioned growth media can be added back to the well. In case of using conditioned media, draw up 190 μL of media with a pipette, aspirate the leftover media in the well, and return the media from the pipette to the well.
5. Thorough washing of the cells after labeling is necessary to reduce background signal. In our hands, the best results are obtained by completely filling each well to the top with ice-cold PBS. Adding the PBS directly on top of the cells or adding it with too much pressure may cause cells to detach from the coverslip. Therefore, it is recommended to press the tip of the wash bottle to the wall of the well and add the PBS at a steady pace. To prevent cells from drying out between washes, only aspirate a reasonable number of wells at a time.
6. After placing the coverslip on top, the fluorescence mounting media should dry within 16–24 h. Excess mounting media or PBS on the coverslip will hinder optimal drying and may cause drifting of the coverslip when cleaning it before imaging. To remove excess PBS from the coverslip, gently tap the edge of the coverslip on a paper towel.
7. When having difficulties removing the coverslips from their wells, a 27G needle bent at the tip can facilitate raising the coverslips from the surface.
8. Macropinosomes are bright clusters of circular puncta. For correct quantitation of the macropinosomes, the exposure time and/or light intensity should be adjusted such that there is no overexposure, autofluorescence of the cells, or background signal. When blotches, smudges, or marks which are definitely not macropinosomes are present in the frame of the photo, retake the DAPI and macropinosome photo in a different frame as these discrepancies will skew the macropinosome quantitation further down the way.

9. To ensure proper comparison between samples, use the same exposure time and light intensity for each coverslip and sections thereof.
10. If the *Cell Counter* plug-in is missing, the *Multi-point Tool* can be used, or the plug-in can be downloaded from the following link: https://imagej.nih.gov/ij/plugins/cell-counter.html.
11. In case your images have been saved in the RGB or 16-bit format, convert them to 8-bit by selecting: *Image* → *Type* → *8-Bit*.
12. ImageJ macros allow for more rapid quantitation of large numbers of samples. Simple macros can be generated by selecting *Plugins* → *Macros* → *Record*, before applying the steps for the quantitation image. After finishing the quantitation of the first image, save the resulting macro commands by selecting *Create* and saving the macro as Macropinocytosis.ijm in the ImageJ "plugins" folder. For the resulting quantitation images, apply the macro by selecting *Plugins* → *Macropinocytosis*. Check macros for consistency, and analyze each image individually to identify spurious sources of fluorescence. In addition, the threshold should be modified accordingly for each experiment by opening the Macropinocytosis.ijm macro using word processing software.

One example of a generated macro to obtain macropinosome area is as follows with X being the decided threshold for macropinosome selection:

```
run("Subtract Background...", "rolling=10");
run("Smooth");
setAutoThreshold("Default dark");
//run("Threshold...");
setThreshold(X, 255);
setOption("BlackBackground", false);
run("Convert to Mask");
run("Analyze Particles...", " summarize");
```

References

1. Doherty GJ, McMahon HT (2009) Mechanisms of endocytosis. In: Annual review of biochemistry, vol 78. Annual Reviews, Palo Alto, CA, pp 857–902. https://doi.org/10.1146/annurev.biochem.78.081307.110540
2. Lim JP, Gleeson PA (2011) Macropinocytosis: an endocytic pathway for internalising large gulps. Immunol Cell Biol 89(8):836–843. https://doi.org/10.1038/icb.2011.20
3. Commisso C, Davidson SM, Soydaner-Azeloglu RG, Parker SJ, Kamphorst JJ, Hackett S, Grabocka E, Nofal M, Drebin JA, Thompson CB, Rabinowitz JD, Metallo CM, Vander Heiden MG, Bar-Sagi D (2013) Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. Nature 497(7451):633–637. https://doi.org/10.1038/nature12138

4. Seguin L, Camargo MF, Wettersten HI, Kato S, Desgrosellier JS, von Schalscha T, Elliott KC, Cosset E, Lesperance J, Weis SM, Cheresh DA (2017) Galectin-3, a druggable vulnerability for KRAS-addicted cancers. Cancer Discov 7(12):1464–1479. https://doi.org/10.1158/2159-8290.cd-17-0539
5. Bar-Sagi D, Feramisco JR (1986) Induction of membrane ruffling and fluid-phase pinocytosis in quiescent fibroblasts by ras proteins. Science (New York, NY) 233(4768):1061–1068. https://doi.org/10.1126/science.3090687
6. Amyere M, Payrastre B, Krause U, Van Der Smissen P, Veithen A, Courtoy PJ (2000) Constitutive macropinocytosis in oncogene-transformed fibroblasts depends on sequential permanent activation of phosphoinositide 3-kinase and phospholipase C. Mol Biol Cell 11(10):3453–3467
7. Fennell M, Commisso C, Ramirez C, Garippa R, Bar-Sagi D (2015) High-content, full genome siRNA screen for regulators of oncogenic HRAS-driven macropinocytosis. Assay Drug Dev Technol 13(7):347–355. https://doi.org/10.1089/adt.2015.660
8. Kamphorst JJ, Nofal M, Commisso C, Hackett SR, Lu WY, Grabocka E, Vander Heiden MG, Miller G, Drebin JA, Bar-Sagi D, Thompson CB, Rabinowitz JD (2015) Human pancreatic cancer tumors are nutrient poor and tumor cells actively scavenge extracellular protein. Cancer Res 75(3):544–553. https://doi.org/10.1158/0008-5472.can-14-2211
9. Recouvreux MV, Commisso C (2017) Macropinocytosis: a metabolic adaptation to nutrient stress in cancer. Front Endocrinol 8. https://doi.org/10.3389/fendo.2017.00261
10. Davidson SM, Jonas O, Keibler MA, Hou HW, Luengo A, Mayers JR, Wyckoff J, Del Rosario AM, Whitman M, Chin CR, Condon KJ, Lammers A, Kellersberger KA, Stall BK, Stephanopoulos G, Bar-Sagi D, Han J, Rabinowitz JD, Cima MJ, Langer R, Vander Heiden MG (2017) Direct evidence for cancer-cell-autonomous extracellular protein catabolism in pancreatic tumors. Nat Med 23(2):235–241. https://doi.org/10.1038/nm.4256
11. Mooren OL, Galletta BJ, Cooper JA (2012) Roles for actin assembly in endocytosis. In: Kornberg RD (ed) Annual review of biochemistry, vol 81. Annual Reviews, Palo Alto, CA, pp 661–686. https://doi.org/10.1146/annurev-biochem-060910-094416
12. Swanson JA (2008) Shaping cups into phagosomes and macropinosomes. Nat Rev Mol Cell Biol 9(8):639–649. https://doi.org/10.1038/nrm2447
13. Koivusalo M, Welch C, Hayashi H, Scott CC, Kim M, Alexander T, Touret N, Hahn KM, Grinstein S (2010) Amiloride inhibits macropinocytosis by lowering submembranous pH and preventing Rac1 and Cdc42 signaling. J Cell Biol 188(4):547–563. https://doi.org/10.1083/jcb.200908086
14. Counillon L, Scholz W, Lang HJ, Pouyssegur J (1993) Pharmacological characterization of stably transfected NA+/H+ antiporter isoforms using amiloride analogs and a new inhibitor exhibiting antiischemic properties. Mol Pharmacol 44(5):1041–1045
15. Commisso C, Flinn RJ, Bar-Sagi D (2014) Determining the macropinocytic index of cells through a quantitative image-based assay. Nat Protoc 9(1):182–192. https://doi.org/10.1038/nprot.2014.004
16. Wang JTH, Teasdale RD, Liebl D (2014) Macropinosome quantitation assay. MethodsX 1:36–41. https://doi.org/10.1016/j.mex.2014.05.002

Chapter 9

Integrated Analysis of Acetyl-CoA and Histone Modification via Mass Spectrometry to Investigate Metabolically Driven Acetylation

Simone Sidoli, Sophie Trefely, Benjamin A. Garcia, and Alessandro Carrer

Abstract

Acetylation is a highly abundant and dynamic post-translational modification (PTM) on histone proteins which, when present on chromatin-bound histones, facilitates the accessibility of DNA for gene transcription. The central metabolite, acetyl-CoA, is a substrate for acetyltransferases, which catalyze protein acetylation. Acetyl-CoA is an essential intermediate in diverse metabolic pathways, and cellular acetyl-CoA levels fluctuate according to extracellular nutrient availability and the metabolic state of the cell. The Michaelis constant (Km) of most histone acetyltransferases (HATs), which specifically target histone proteins, falls within the range of cellular acetyl-CoA concentrations. As a consequence, global levels of histone acetylation are often restricted by availability of acetyl-CoA. Such metabolic regulation of histone acetylation is important for cell proliferation, differentiation, and a variety of cellular functions. In cancer, numerous oncogenic signaling events hijack cellular metabolism, ultimately inducing an extensive rearrangement of the epigenetic state of the cell. Understanding metabolic control of the epigenome through histone acetylation is essential to illuminate the molecular mechanisms by which cells sense, adapt, and occasionally disengage nutrient fluctuations and environmental cues from gene expression. In particular, targeting metabolic regulators or even dietary interventions to impact acetyl-CoA availability and histone acetylation is a promising target in cancer therapy. Liquid chromatography coupled to mass spectrometry (LC-MS) is the most accurate methodology to quantify protein PTMs and metabolites. In this chapter, we present state-of-the-art protocols to analyze histone acetylation and acetyl-CoA. Histones are extracted and digested into short peptides (4–20 aa) prior to LC-MS. Acetyl-CoA is extracted from cells and analyzed using an analogous mass spectrometry-based procedure. Model systems can be fed with isotopically labeled substrates to investigate the metabolic preference for acetyl-CoA production and the metabolic dependence and turnover of histone acetylation. We also present an example of data integration to correlate changes in acetyl-CoA production with histone acetylation.

Key words Acetyl-CoA, Histones, Mass spectrometry, Metabolism, Proteomics

1 Introduction

Post-translational modifications (PTMs) of histone proteins are important epigenetic signals that change the physical accessibility of the genome. Histone PTMs affect DNA-protein interactions

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_9,

and, subsequently, gene transcription, without altering the underlying DNA sequence [1]. Lysine acetylation is a widespread histone PTM and is catalyzed by a class of enzymes known as histone acetyltransferases (HATs). The core histone proteins (H2A, H2B, H3, and H4) form the nucleosome structure around which DNA is tightly wrapped. Core histone proteins can be acetylated on multiple residues; notably, histones have long unstructured tails that protrude from the nucleosome that can be abundantly acetylated. Histones are enriched in basic amino acid residues (Arg and Lys) making them positively charged. Acetylation of lysine residues neutralizes this positive charge, relaxing the interaction of the nucleosomes with the negatively charged DNA. Thus, histone acetylation is usually associated with chromatin openness (*euchromatin*), characterized by accessibility to transcription factors and other DNA-associated proteins, and active gene transcription. In addition to its physical impact on nuclear architecture, acetylated histone lysines are recognized by specific protein domains (*bromodomains*) usually found in potent gene activators [2]. A varied set of enzymes, known as histone deacetylases (HDACs), catalyze a rapid and energetically inexpensive removal of acetyl moieties from histones [3]. Altogether, histone acetylation is a dynamic and reversible chromatin modification that allows cells to rapidly and potently modulate gene expression in order to adapt to extracellular stimuli. However, it is now clear that levels of histone acetylation can also be influenced by availability of the universal acetyl donor, acetyl-CoA [4–6]. Acetyl-CoA is an essential intermediate in diverse metabolic pathways, and cellular acetyl-CoA levels fluctuate according to extracellular nutrient availability and the metabolic state of the cell. Acetyl-CoA cannot cross the mitochondrial membrane, so only acetyl-CoA generated in the nucleocytoplasmic compartment is accessible for histone modification. In most cell settings, acetyl-CoA is produced in the nucleocytoplasmic compartment from citrate through the activity of ATP-citrate lyase (ACLY) (Fig. 1a). Citrate generated in the mitochondrial TCA cycle is exported to fuel this process. As glucose is the primary substrate for TCA cycle activity in a number of settings, glucose limitation has been shown to restrict acetyl-CoA availability and decrease global levels of histone acetylation [7–9] (Fig. 1b). Importantly, cancer cells, as well as highly specialized cell types, can channel TCA carbon units into acetyl-CoA to sustain high levels of histone acetylation even when glucose is limiting or poorly utilized [10–12]. For example, activation of PI3K/Akt signaling in cancer cells leads to phosphorylation of ACLY on serine 455, which increases enzymatic activity of approximately sixfold and elevates acetyl-CoA and histone acetylation levels also in glucose-limiting conditions [7]. Multiple groups have also shown that cells can utilize alternative carbon sources to produce acetyl-CoA, in particular acetate, which can be converted to acetyl-CoA by the acyl-coenzymeA synthetase short chain family

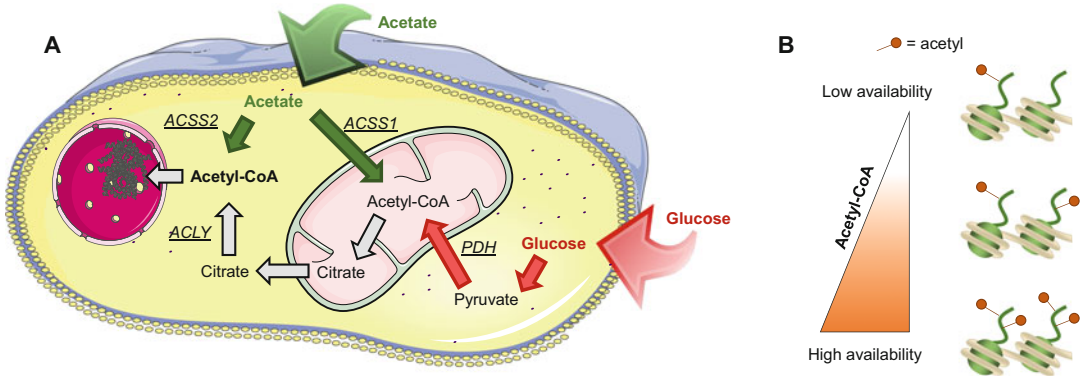


Fig. 1 Pathways leading to histone acetylation. (**a**) Acetyl-CoA is commonly produced by processing of either glucose or acetate. Citrate is converted into acetyl-CoA by the ATP-citrate lyase (ACLY), while acetate is processed into acetyl-CoA by the enzyme acetate synthetase 2 (ACSS2). (**b**) Increased availability of acetyl-CoA correlates with elevated histone acetylation

member 2 (ACSS2). Although acetate-derived acetyl-CoA is less efficiently incorporated into histones, contribution of acetate to histone acetylation is significantly enhanced under some conditions (e.g., hypoxia) [10].

Acetyl-CoA is highly unstable, so the appropriate use of isotope-labeled internal standard is recommended for accurate quantitation. Mass spectrometry (MS) is the most reliable approach for the quantification of acetyl-CoA. Enzyme-based assays for acetyl-CoA quantitation suffer from sensitivity and specificity issues and cannot incorporate appropriate internal standards to account for sample degradation [13]. Internal standards added in equal amounts are used to normalize every sample accounting for sample loss and degradation. Exact quantitation can be achieved by comparing sample ratios to a standard curve generated by a serial dilution of known amounts of acetyl-CoA standards. Internal standard can be generated cheaply in yeast through stable isotope labeling of essential nutrients in cell culture (SILEC) where cells are fed with heavy labeled [$^{13}C_3{}^{15}N_1$]-vitamin B5, which is incorporated into the CoA backbone of acetyl-CoA [14]. The characteristic fragmentation pattern of acetyl-CoA with MS/MS detection ensures highly specific quantitation [13].

Antibody-based techniques such as Western blotting have been extensively adopted to characterize histone modifications, including acetylation. However, this approach is limited in throughput and sometimes specificity, as antibodies often cross-react with similar histone marks and multiple PTMs. A recent assessment of commercial antibodies found that more than 25% fail specificity tests in dot blot and Western blot experiments and about 20% of antibodies fail in ChIP-seq experiments due to non-specific antibody binding [15]. In contrast, MS platforms can achieve high

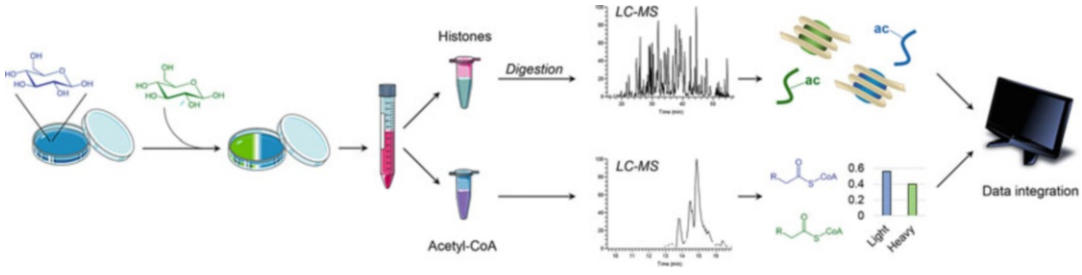


Fig. 2 Proposed workflow. Cell cultures can be grown in the presence of either normal (represented by blue medium) or labeled (green medium) nutrients. Labeled histones and metabolites can be analyzed in parallel, according to different protocols. Quantification of either is performed by LC-MS analysis. Data can then be integrated for correlation analysis

specificity and sensitivity, with automation facilitating high-throughput analyses. MS has thus become the most suitable analytical tool to study both acetyl-CoA and histone modifications in general (reviewed in [16, 17]).

In this chapter, we describe an integrated workflow to probe acetyl-CoA fluctuations and changes in histone acetylation. The protocol is designed for the analysis of multiple histone acetyl marks. Schematic representation of our approach is shown in Fig. 2. Even though the protocol is primarily prepared for cell culture studies, adaptability to in vivo analysis of histone modifications and acetyl-CoA availability will be discussed.

2 Materials

2.1 Reagents and Abbreviations

1. Acetonitrile (ACN).
2. Ammonium hydroxide (NH_4OH), 28% NH_3 in water.
3. D-glucose-$^{13}C_6$ and acetate-$^{13}C_6$ (heavy labeled).
4. Propionic anhydride and acetonitrile for propionylation mixture.
5. Sodium acetate-$^{13}C_2$ (heavy labeled).
6. Trichloroacetic acid (TCA).
7. Trifluoroacetic acid (TFA).
8. Trypsin (sequencing grade).
9. Hydrochloric acid (HCl), 32%.
10. Acetone, residue grade.
11. Coomassie® (Bradford) Protein Assay kit.
12. Coomassie® solution.
13. Ammonium bicarbonate (NH_4HCO_3).
14. Glacial acetic acid.

2.2 Buffers

1. Phosphate-buffered saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4.
2. Ammonium bicarbonate (NH_4HCO_3): 50 mM NH_4HCO_3, pH 8.0.
3. Nuclei isolation buffer (NIB-250): 15 mM Tris–HCl (pH 7.5), 15 mM NaCl, 60 mM KCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 250 mM sucrose.
4. Stage-tip loading and wash buffer: 0.1% TFA.
5. Stage-tip elution buffer: 75% acetonitrile, 0.025% TFA.
6. NanoLC buffer A (for histone peptide analysis): 0.1% formic acid in HPLC grade water.
7. NanoLC buffer B (for histone peptide analysis): 0.1% formic acid, 95% HPLC grade ACN, in HPLC grade water.
8. Extraction buffer A (for acetyl-CoA analysis): 10% (w/v) trichloroacetic acid in HPLC grade water.
9. Extraction buffer B (for acetyl-CoA analysis): HPLC grade methanol containing 25 mM ammonium acetate.
10. Extraction buffer B (for acetyl-CoA analysis): 5% (w/v) 5-sulfosalicylic acid in HPLC grade water.
11. HPLC buffer A (for acetyl-CoA analysis): 5 mM ammonium acetate in HPLC grade water.
12. HPLC buffer B (for acetyl-CoA analysis): 5 mM ammonium acetate in HPLC grade acetonitrile/water (95:5, v/v).
13. HPLC buffer C (for acetyl-CoA analysis): 0.1% formic acid, 80% HPLC grade ACN, in HPLC grade water.

2.3 Solutions

1. Protease inhibitors (add fresh to buffers prior to use): 1 M dithiothreitol (DTT) in ddH_2O (1000 ×); 200 mM AEBSF in ddH_2O (400×).
2. Phosphatase inhibitor (add fresh to buffers prior to use): 2.5 μM microcystin in 100% ethanol (500×).
3. HDAC inhibitor (add fresh to buffers prior to use): 5 M sodium butyrate, made by titration of 5 M butyric acid using NaOH to pH 7.0 (500×).
4. 10% (v/v) NP-40 alternative in ddH_2O.
5. 0.2 M H_2SO_4 in ddH_2O.
6. 100% TCA (w/v) in ddH_2O.

2.4 Equipment

1. Tissue and cell homogenizers (optional).
2. pH indicator strips (pH 0–14).
3. Liquid nitrogen.
4. 1.5 mL microcentrifuge tubes.

5. 15 and 50 mL conical tubes.
6. Pipettes from P10 to P1000 range with respective tips.
7. −80 °C refrigerator.
8. SpeedVac.
9. Heat blocks or water baths.
10. 3 M Empore™ solid phase extraction disks C_{18}.
11. 75 and 100 μm internal diameter fused silica tubings.
12. Micro-stir magnets.
13. C_{18}-AQ 3 μm bulk resin with 200–300 Å pore size for trap column and analytical column for nanoHPLC.
14. Pressure cell for capillary column packing with respective compressed gas bomb (either helium, nitrogen, or air).
15. Oasis HLB 1 cc (30 mg) solid phase extraction columns (waters).
16. Commercial 2.1 mm ID C_{18} column.

3 Methods

All procedures should be carried out at room temperature, unless specified otherwise. Samples can be frozen and stored in −80 °C at the end of each section, best if previously dried in a SpeedVac concentrator centrifuge. The scheme of the full workflow is illustrated in Fig. 2. For simplicity, we will be discussing the use of either [^{13}C]-glucose or [^{13}C]-acetate to refer to the ubiquitously labeled forms (every carbon atoms substituted with heavy [^{13}C] isotope). Both are environmental sources for the generation of nucleocytoplasmic acetyl-CoA. Glucose tracing is preferred to assess the contribution of ACLY to the existing acetyl-CoA pool. Acetate is a carbon source that does not require ACLY activity and thus gauges the contribution of ACSS2 to acetyl-CoA production. The use of one is alternative to the other, so informative data can be obtained only tracing the two carbon sources separately.

3.1 Labeling of Biological Samples Using Stable Isotopes of Glucose or Acetate

3.1.1 Cell Culture

Glucose limitation restricts acetyl-CoA availability in various cell lines, in a way that ultimately impacts global levels of histone acetylation. We recommend using this feature as positive control for the experiment. However, keep in mind that some cell types adapt to glucose limitation by using alternative substrates such as acetate for acetyl-CoA generation.

1. Plate an appropriate number of cells in a 10 cm dish (*see* **Note 1**).
2. Culture cells in standard medium for 12–36 h.

3. Replace culture medium with medium (glucose-free) supplemented with 10 mM [^{13}C]-glucose or 100 μM [^{13}C]-acetate. Use medium supplemented with dialyzed serum. Regular serum contains traceable amount of glucose or acetate, which might affect analysis. Remember to add equivalent cold counterpart (e.g., when tracing [^{13}C]-glucose, add 100 μM [^{12}C]-acetate to the medium, and vice versa). For optimization, *see* **Note 2**
4. As positive control, replace culture medium with medium supplemented with 1 mM [^{13}C]-glucose and 100 uM acetate (unlabeled).
5. Incubate for 2–24 h at 37 °C, depending on the metabolic activity of the system adopted (may need optimization).

3.1.2 Organisms

Carbon tracing into metabolites (and potentially histone proteins) has been performed to determine the metabolic fate of glucose, acetate, or glutamine in vivo. This approach has generated extremely valuable results, especially when applied in humans [18]. Nonetheless, in vivo tracing of acylated metabolites is technically demanding and presents numerous challenges. Here, we briefly outline a simplified protocol for the quantification of labeled acetyl-CoA from either [^{13}C]-glucose or [^{13}C]-acetate in mice. Note that a "stress-free" protocol for efficient and reliable tracing of metabolites in vivo has very recently been proposed [19]. Also, glucose can be more conveniently administered by IP injection or gavage [20]. In addition, authors recommend performing all the experiments described hereafter upon approval of a proper IACUC protocol (or equivalent approval from dedicated ethical committee).

1. Prepare or treat animal according to a pre-optimized study design. Remember to tag animals.
2. Restrain mouse movements (typically, use a mouse restrainer).
3. Prepare a 25% (w/v) [^{13}C]-glucose solution and/or 3 mM [^{13}C]-acetate solution in PBS. Filter the solution(s) with a 0.2 μm sterile filter.
4. Inject 80 μL of stock solution into the tail vein. Repeat the injection three times at 15 min intervals.

Subheadings 3.2–3.9 are for histone extraction and analysis. For extraction and analysis of acetyl-CoA, proceed to Subheading 3.10. The sample can be divided into two aliquots, and the two sample preparations can be performed in parallel.

3.2 Sample Harvesting

3.2.1 Harvesting of Adherent Cells

1. Remove plate(s) from the incubator and place on ice.
2. Scrape cells using a cell lifter and place them in a new tube.
3. Centrifuge cells at 300 × *g* for 5–10 min.
4. Remove supernatant.
5. Resuspend cells in ice-cold PBS and repeat **steps 3** and **4**.
6. Estimate the volume of cell pellets (approximate), and keep them frozen until use (*go to Subheading 3.3*).

3.2.2 Harvesting of Suspension-Growing Cells

1. Remove plate(s) from the incubator and put it on ice.
2. Aspirate cells and transfer them in a new tube.
3. Centrifuge cells at 300 × *g* for 5–10 min.
4. Remove supernatant.
5. Resuspend cells in ice-cold PBS and repeat **steps 3** and **4**.
6. Estimate the volume of cell pellets (approximate), and keep them frozen until use (*go to Subheading 3.3*).

3.2.3 Harvesting of Tissues

1. Sacrifice animals according to an IACUC-approved protocol.
2. Rapidly expose organ of interest using clean surgical tools.
3. Cut approximately 100 mg of tissue with sharp scissors, and quickly rinse with ice-cold PBS.
4. Homogenize the tissue explant with a Dounce homogenizer in 1 mL PBS.
5. Transfer tissue homogenate into a new tube.
6. Centrifuge tissue homogenate at 300 × *g* for 5–10 min.
7. Remove supernatant.
8. Resuspend pellet in ice-cold PBS and repeat **steps 6–8**.
9. Estimate the volume of pellets (approximate), and proceed directly to histone extraction without freezing (*go to Subheading 3.3*).

3.3 Isolation of Cell Nuclei

This section describes how to separate intact nuclei from the cell cytoplasm, membrane, and other organelles. This reduces the presence of background proteins when histones are purified. Notably, protocols bypassing nuclei isolation have been published [21].

1. Add protease inhibitors and other inhibitors to NIB-250 buffer. For 1 mL of cell pellet, approximately 50 mL of NIB-250 buffer is prepared. Add to 50 mL NIB-250 buffer 50 μL of 1 M DTT, 125 μL of 200 mM AEBSF, 100 μL of 2.5 μM microcystin, and 100 μL of 5 M sodium butyrate.
2. Lyse the cell pellet with ten times the volume of NIB-250 including inhibitors and 0.2% NP-40 alternative.

3. Properly homogenize the cell suspension. Normally, gentle pipetting is sufficient, but tissue samples might need the use of a Dounce homogenizer.
4. Incubate the suspension on ice for 5–10 min; the outer cell membranes will lyse and release nuclei.
5. Centrifuge at 1000 × *g* for 5–10 min at 4 °C. Cell nuclei are pelleted, while the supernatant contains mostly cytoplasmic components.
6. Resuspend the nuclei pellet using ten volumes of NIB-250 + inhibitors without the NP-40 alternative.
7. Centrifuge at 1000 × *g* for 5 min at 4 °C and remove supernatant.
8. Repeat **steps 6** and **7** for complete removal of residual NP-40 alternative.

3.4 Extraction and Purification of Histones

Histones are extracted exploiting their solubility in acid (H_2SO_4). Alternatively, salt extraction can be performed (*see* **Note 3**).

1. Resuspend cell nuclei with five volumes of 0.2 M H_2SO_4 by gentle pipetting.
2. Incubate the sample with gentle rotation or shaking for 2–4 h at 4 °C. Use the longer time frame in case of low abundance material, i.e., <200 μL cell pellet.
3. Centrifuge at 3400 × *g* for 5 min.
4. Transfer the supernatant to a new tube.
5. Repeat **steps 3** and **4** with the supernatant to ensure complete cleanup from pellet residuals.
6. Add 100% TCA to the sample solution with a ratio of 1:3 (v/v), to obtain a final TCA concentration of 33%. This step will precipitate histones.
7. Let the mixture precipitate on ice for at least 1 h or overnight.
8. Centrifuge at 3400 × *g* for 5 min. Remove the supernatant by aspiration without touching the precipitate. Histones are the white layer condensed around the bottom of the tube. The pellet in the very bottom of the tube normally contains other acid biomolecules, such as DNA.
9. By using a glass Pasteur pipette, rinse the tube with acetone +0.1% HCl.
10. Centrifuge at 3400 × *g* for 2 min and discard the supernatant.
11. Repeat **steps 9** and **10** using acetone without 0.1% HCl.
12. Leave the tubes open on the bench for a few minutes to dry them completely.

13. Resuspend histone proteins in 30–50 µL of ddH_2O. Rinse the borders of the tubes as best as possible, especially the white layer on the side of the tube.
14. Estimate the amount of histone proteins in the sample using either BCA, Bradford protein assay, or amino acid analysis (AAA).
15. Verify the purity of extracted histones with SDS gel and Coomassie staining (optional).
16. Histone isotypes can be separated and differentially purified using HPLC-UV equipped with a C_{18} column (*see* **Note 4**) (optional).

3.5 Propionic Anhydride Derivatization Prior Histone Digestion for Bottom-Up Analysis

In proteomics, proteins are digested into short (6–30 aa) peptides prior analysis. This approach is called "bottom-up" or "shotgun." Histones are also digested into short peptides. The canonical protein digestion protocol uses trypsin as digestion enzyme, which cleaves at the C-termini of basic amino acid residues, i.e., lysine (K) and arginine (R). Histones are highly enriched of KR residues, and thus they require either the use of alternative enzymes [22] or derivatization strategies to reduce trypsin targets on the protein sequence [23–26]. Here, we discuss the most widely adopted histone digestion protocol, adopting propionic anhydride derivatization prior and after trypsin digestion [27, 28]. Such derivatization blocks the ε-amino groups of unmodified and monomethyl lysine residues, allowing trypsin to perform proteolysis only at the C-terminal of arginine residues. Derivatization after digestion increases peptide hydrophobicity, which enhances efficient HPLC column retention for HPLC-MS.

1. Resuspend at least 20 µg of histones in 30 µL of 50 mM NH_4HCO_3, pH 8.0. Recommended amount is 50–100 µg (*see* **Note 5** for estimated histone yield from cell counts).
2. Assess the pH using a P10 pipette tip; dip it into the sample, and touch a pH indicator strip. NH_4OH and glacial acetic acid can be used to adjust the pH (*see* **Note 6** for safety instructions).
3. For three samples, prepare propionylation reagent by mixing propionic anhydride with acetonitrile (ACN) in the ratio 1:3 (v/v); i.e., mix 15 µL of propionic anhydride and 45 µL of ACN (*see* **Note 7** on alternative procedure in presence of a large number of samples).
4. Add rapidly the propionylation reaction to the histone sample with a ratio of 1:2 (v/v), i.e., 15 µL propionylation reaction for 30 µL sample.

5. Add rapidly ~7 μL of NH_4OH to re-establish pH 8.0 to the solution (*see* **Note 8** for elucidations on issues related to not optimal pH).
6. Pipette up and down for a few seconds.
7. Assess pH as described in **step 2**.
8. Incubate samples at room temperature for 15–20 min.
9. Dry samples down to 5–10 μL in a SpeedVac centrifuge.
10. Resuspend or dilute samples with 50 mM NH_4HCO_3 until achieving 30 μL of final volume.
11. Repeat **steps 3–10** to double the propionylation reaction to ensure complete derivatization.

3.6 Histone Digestion and Propionylation of Peptide N-Termini

1. Resuspend histones in 30 μL of 50 mM NH_4HCO_3.
2. Assess the pH to be around 8.0 by using pH strips.
3. Add trypsin (sequencing grade) at a 1:20 ratio (w/w), e.g., 5 μg of trypsin for 100 μg of histones.
4. Incubate at room temperature for 6 h or overnight.
5. Repeat **steps 3–12** of Subheading 3.5, double round of propionylation included. At the second round, stop at **step 10** (included).
6. Resuspend the histone peptide samples in 50 μL of ddH_2O + 0.1% trifluoroacetic acid (TFA).

3.7 Stage Tipping for Sample Desalting

To remove residuals of salt, propionylation reagent, and other debris leftover of the histone extraction, samples are passed through a tip packed with reversed-phase material for cleanup. Not doing so might result in column clogging and/or instrument contamination during HPLC-MS analysis.

1. Take a 3 M Empore™ solid phase extraction disk C_{18}, and cut a disk of ~2–3 mm diameter, e.g., by using the tip of a P1000 pipette.
2. Push this disk to the bottom of a P100/200 pipette tip, e.g., by using a fused silica capillary.
3. Repeat **steps 1** and **2** in the same P100/200 tip if you are desalting more than 10 μg of sample. This will increase the capacity of the stage tip.
4. Wash the stage tip by flushing 50 μL of 75% ACN and 0.025% TFA with air pressure, e.g., using a syringe (*see* **Note 9** for higher throughput procedure using centrifugation).
5. Equilibrate the stage tip by flushing 50 μL of 0.1% TFA by air pressure. Do not dry completely the disk(s).
6. Load the sample onto the stage tip by air pressure. Do not dry completely the disk(s).

7. Wash the sample by flushing 50 µL of 0.1% TFA by air pressure. Do not dry completely the disk(s).
8. Elute the sample by flushing 50 µL of 75% ACN and 0.025% TFA by air pressure. Collect the sample in a 1.5 mL tube.
9. Dry samples in a SpeedVac centrifuge.

3.8 Histone Peptide Analysis via Nano Liquid Chromatography Coupled to Mass Spectrometry (NanoLC-MS)

NanoLC is now the separation technique most preferred for proteomics, due to its high sensitivity and the possibility of online coupling to MS. Here, we describe how to prepare columns for nanoLC (**steps 1–8** can be omitted if using commercial columns) and how to configure the nanoLC-MS analysis. This procedure described how to prepare a picofrit column, i.e., a nano column integrated with the tip. For alternative procedure, *see* **Note 10**.

1. Cut ~30 cm of fused silica capillary with 75 or 100 µm internal diameter (ID).
2. Tape one end of the capillary on a solid surface, e.g., the bench (Fig. 3a).
3. By using a torch, fire close to the end of the capillary while gently pulling from the other end (Fig. 3a). After a few seconds, the capillary will elongate and detach from the surface creating a tip at the end (Fig. 3b). This tip might look different every time; its end might be very long, like a "hair" (Fig. 3c).
4. Gently remove the "hair" part of the tip, e.g., with the finger, until it looks like a tip for nano columns (Fig. 3d). This procedure might take a few attempts (*see* **Note 11** to overview possible consequences of failure).
5. Prepare in a clean HPLC glass vial the resin slurry for column packing. This includes (a) C_{18}-AQ reversed-phase 3 µm particles, (b) 100% methanol, and (c) a micro-stir magnet. The ratio

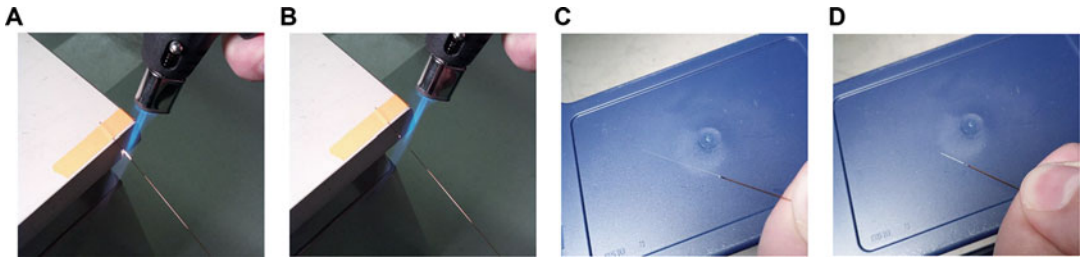


Fig. 3 Tip pulling for nano columns. (**a**) Capillaries should be secured on a rigid surface, e.g., by using tape. (**b**) Heating combined with gentle pulling elongates the capillary into a smaller and smaller internal diameter, until it detaches from the surface as a pulled needle. (**c**) This needle is frequently a nonrigid end, similar to a hair. This end has no internal diameter, preventing any liquid to pass through it. (**d**) By gently breaking the nonrigid part of the tip (e.g., with a finger), a normal size tip is created. This capillary is ready to be packed and become a nano column for LC-MS

between particles and methanol are flexible and should be optimized for a rapid and efficient column packing.

6. Place the resin slurry in a pressure bomb and turn on magnetic stirring.
7. Place the capillary in the pressure bomb, secure it, and open the gas tank (containing helium, nitrogen, or air).
8. Leave the column on the bomb until it is packed for at least ~25 cm.
9. Remove the column from the pressure bomb, and connect it to the nanoLC.
10. Program the HPLC method as follows: From 2% to 28% buffer B in 45 min, from 28% to 80% B in 5 min, and 10 min at isocratic 80% B. Buffer A and B composition are described in Subheading 2.2 (**items 3** and **4**, respectively).
11. Program the MS acquisition method to perform data-independent acquisition (DIA) (references to set up the method using SWATH™, using low-resolution instrumentation and using multiplexed DIA, respectively [29–31]). The instrument will alternate a full MS scan with MS/MS scans of the entire mass range using acquisition windows of 50 *m/z*. All other settings are in common to standard proteomic experiments.
12. Load ~1 μg of sample onto the HPLC column.
13. Run the nanoLC-MS/MS method as programmed.

3.9 Extracted Ion Chromatography (XIC) of Histone Peptides

Raw files obtained from the LC-MS runs are now ready to be processed. In our lab, we developed EpiProfile [32], a software tool that performs extracted ion chromatography (XIC) of histone peptides (Fig. 4).

1. Group the raw files into the same folder.
2. Run EpiProfile using Matlab or GNU Octave. The software will provide a table containing the XIC of the desired analytes. In the output, peptides are already normalized to obtain the relative abundance of each post-translational modification (PTM). The relative abundance is calculated by dividing the XIC of a given peptide modified form by the sum of all XIC of peptides sharing the same sequence. The software automatically discriminates isobaric species, i.e., differentially modified peptides with the same intact mass, by using the MS/MS events acquired with DIA.
3. Light and heavy labeled acetylations are considered independently by the software. The turnover of a given acetylation is calculated by dividing the area of the XIC of the peptide with

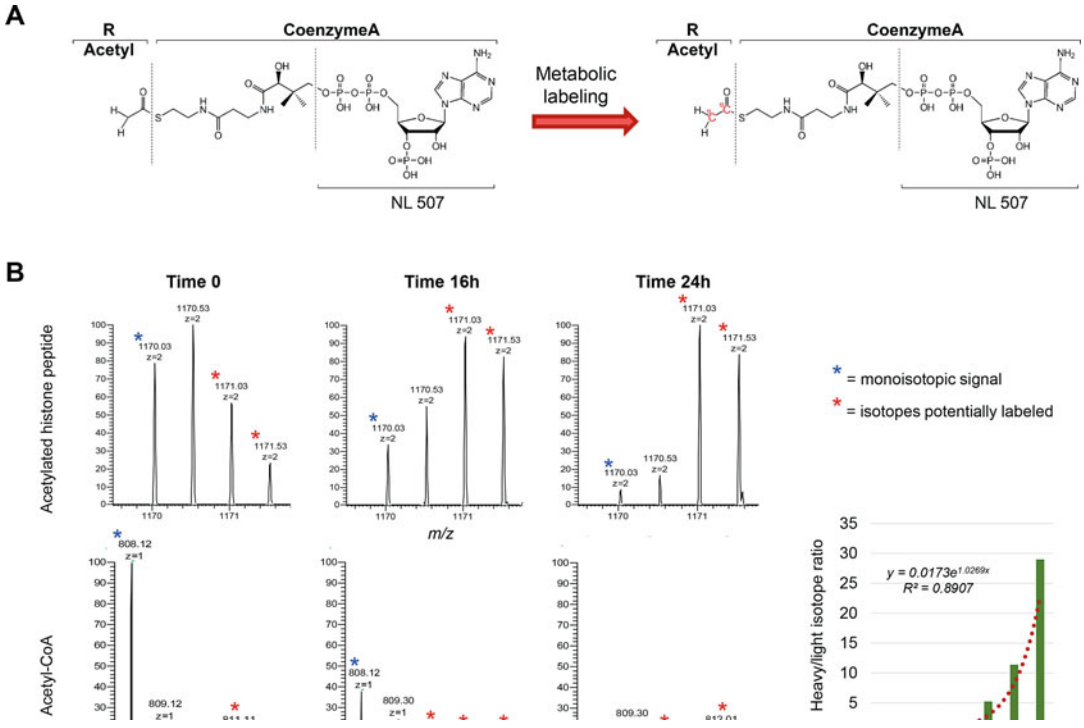


Fig. 4 Example of histone peptide and acetyl-CoA spectra before and after metabolic labeling. (**a**) Left: acetyl-CoA chemical structure. The acetyl group (R—2 carbon atoms) is attached to the coenzyme A backbone through a thioester bond. Right: glucose, acetate, and other nutrients provide the carbon atoms to the acetyl group. Using isotope labeling, ^{13}C atoms can be incorporated into acetyl-CoA (carbons highlighted in red in the structure; each +1 Da). Rate of incorporation and carbon source depends on the metabolic activity of the cell and may be studied with the approach presented here. (**b**) The spectra display how the isotopic pattern of histone peptides (top) and acetyl-CoA (bottom) changes after growing fast-replicating cells into a media containing isotopically heavy glucose (+6 Da). A histone peptide carrying a single acetylation increments its third isotope (+2 Da), as only two ^{13}C carbon atoms are incorporated into the acetyl group. With time, the relative abundance of the heavy isotope increases until the unlabeled acetylation disappears from the signal. Acetyl-CoA gets labeled in multiple carbon residues, as the CoA group is also synthesized using glucose. Therefore, the isotopes increasing their relative abundance are multiple

the heavy acetyl group by the peptide with the light acetyl group.

4. Alternatively to **steps 1–3**, the XIC can be performed with other software tools such as Skyline [33]. However, EpiProfile is currently the only available software trained for this purpose. Other software tools require extensive manual integration and post-processing to discriminate isobaric forms.

3.10 Extraction of Acetyl-CoA from Cells

Here, we describe the method for the extraction, purification, and quantification of labeled acetyl-CoA from cells. Final data will be expressed as the ratio between heavy (labeled) and light (unlabeled). The use of trichloroacetic acid as extraction buffer helps

both to stabilize the unstable thioester bond and slow acetyl-CoA degradation and precipitates protein from the samples. As acetyl-CoA is highly unstable, extraction from cells should be carried out as quickly as possible, and samples should be kept cold on ice at all times. For accurate quantitation of total levels of acetyl-CoA, we recommend the use of stable isotope-labeled internal standards added to samples as early as possible in the sample processing. With high resolution MS (>10,000), the use of [$^{13}C_3{}^{15}N_1$]-labeled internal standard for accurate quantitation can be used to generate labeling and quantitation data simultaneously [34]. Use a replicate dish for cell counting/cell volume measurement using a BD Coulter Counter or analogous instruments. *See* **Note 12** for extraction of acetyl-CoA from whole tissues.

1. Take cell dishes out of the incubator, and place them on ice (for suspension cells, spin down at 400 × *g* at 4C and place on ice).
2. Aspirate tissue culture media thoroughly (for adherent cells, tilt dishes on a slope on ice after initial aspiration to allow residual media to drain down and remove by repeating aspiration). *See* **Note 13** on why cells are not washed.
3. Add 1 mL of ice-cold 10% TCA directly to the cell plate, and scrape cells with a cell lifter (for cell pellets from suspension cells, add 1 mL of ice-cold 10% TCA and mix briefly by vortexing).
4. Transfer cell suspension to a new 1.5 mL Eppendorf tube. You may store the samples at −80 °C.
5. Sonicate samples with a probe tip sonicator (12 × 0.5 s pulses).
6. Centrifuge dismembranated cells at 13,000–17,000 × *g* for 10 min at 4 °C. the cleared supernatant contains acetyl-CoA extract; the protein pellet can be used for Western blotting and/or protein determination.
7. Take 1 mL capacity solid phase extraction columns with strongly hydrophilic, reversed-phase chemistry.
8. Wash columns with 1 mL of methanol.
9. Equilibrate columns with 1 mL of water.
10. Load columns with the supernatant from samples.
11. Desalt columns with 1 mL of water.
12. Elute columns with 1 mL of methanol containing 25 mM ammonium acetate, and recover the eluted fraction.
13. Evaporate to dryness under nitrogen flow.

3.11 HPLC-MS Analysis of Acetyl-CoA

1. Resuspend the sample purified in Subheading 3.10 in 50 μL 5% (w/v) 5-sulfosalicylic acid dissolved in HPLC grade water, and transfer to HPLC-compatible vials or 96-well plate for analysis.

Table 1
Mass transitions for relevant acetyl-CoA isotopologues in MS-positive mode

Species	Isotopologue	Parent *m/z*	Product *m/z*
Acetyl-CoA	M0	810.1	303.1
Acetyl-CoA	M1	811.1	304.1
Acetyl-CoA	M2	812.1	305.1
Acetyl-CoA	M3	813.1	306.1
Acetyl-CoA	M4	814.1	307.2
Acetyl-CoA	M5	815.1	308.2

Store at 4 °C. *See* **Note 14** for details on why the quick extraction.

2. Connect a C_{18} reversed-phase chromatographic column to an HPLC system capable of running flows at 0.2 mL/min. *See* **Note 15** for recommended commercial columns.
3. Program the HPLC gradient as follows: Isocratic 2% buffer B (98% buffer A) for 2 min, from 2% to 25% B for 3.5 min, from 25% to100% B in 0.5 min, and isocratic 100% B for 8.5 min, washed with 100% buffer C for 5 min followed by equilibration to 2% buffer B for 5 min. The composition of buffer A, B, and C is described in Subheading 2.2 (**items 11–13**). Set flow rate to 0.2 mL/min.
4. Set up the MS acquisition method as DIA or targeted. The *m/z* acquisition window should include the masses for precursors and fragments as outlined in Table 1. Use the instrument in positive mode.
5. HPLC-MS/MS method as programmed.

3.12 Extracted Ion Chromatography (XIC) of Acetyl-CoA

1. Import data into a peak detection software compatible with your MS platform. An example of universal software for peak area extraction is Skyline [33].
2. Perform extracted ion chromatography (XIC) for each acetyl-CoA species to obtain area under the curve values.
3. To confirm the specificity of the peak, check that the retention time of parent acetyl-CoA isotopologue peaks aligns with their corresponding product ion peaks.
4. Calculate percent isotopic enrichment from ^{13}C-labeled substrate by entering AUC data from labeled and unlabeled control samples into the FluxFix web tool (www.fluxfix.science) [35]. *See* **Note 16** on how to computationally correct biases due to isotopic enrichment.

3.13 Data Integration and Interpretation

EpiProfile [32] provides a result table with approximately 300 peptide isoforms. Because the list can be overwhelming for manual inspection, it is very helpful to process the data using the proper statistics, in order to detect the most significant and reliable changes. When using a limited number of replicates (<5), we recommend the use of t-test to estimate the significant differences between conditions. Nonparametric statistics is generally more appreciated, as it can be applied also if replicates do not have a Gaussian distribution; however, it is not sufficiently powerful to deal with such small number of data points. In case an overall trend of acetylation increase/decrease is observed, we recommend to correlate the observation with the levels of acetyl-CoA before jumping to conclusions like "we observe an overall higher activity of enzymes catalyzing histone acetylation."

1. Open the table containing the raw intensities (Fig. 5a) of the identified and quantified histone peptides.
2. Normalize each modified histone peptide by the sum intensity of all peptides sharing the same sequence (Fig. 5b). For

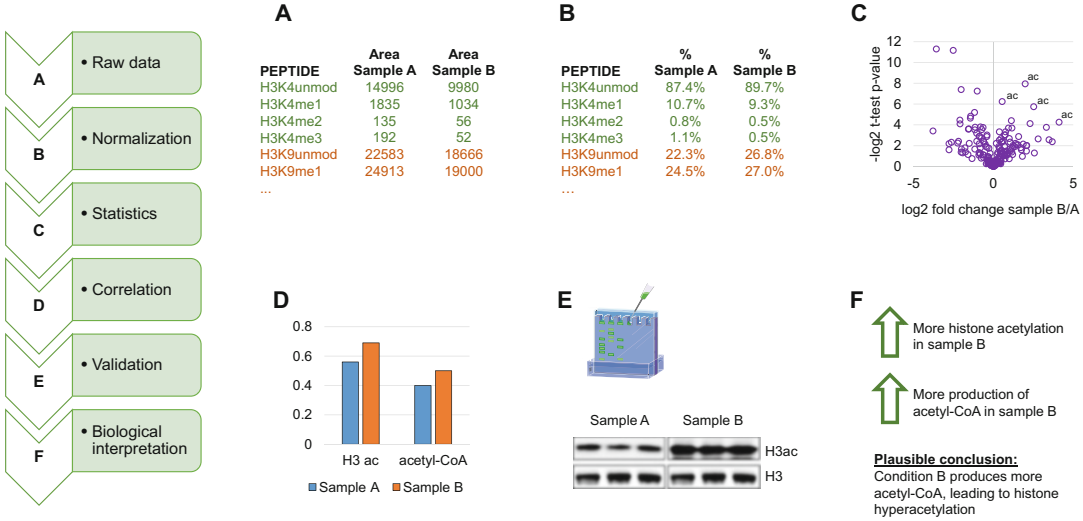


Fig. 5 Representative workflow for data analysis. (**a**) Raw data produced by EpiProfile are the area of the extracted ion chromatogram of selected histone peptides (currently ~300) and acetyl-CoA. EpiProfile can also extract ion chromatograms of histone peptides carrying isotopically labeled acetylations. (**b**) The raw intensity is automatically converted into a relative abundance by dividing the intensity of each peptide by the sum of all peptides (both unmodified and labeled with heavy carbons) sharing the same amino acid sequence. (**c**) Plotting and analysis of differentially modified histone species and/or metabolite isotopomers (recommended statistics is based on t-test or ANOVA). (**d**) The relative changes of histone acetylation can be compared with relative changes of acetyl-CoA when comparing samples A and B, so that a correlation can be performed to predict causes of global regulation of histone acetylation. (**e**) Validation can be performed by, e.g., Western blotting. (**f**) Data-driven hypotheses are easier to formulate when both histone and acetyl-CoA data are acquired

instance, the peptide containing H3K4me3 has the sequence TKQTAR in human (and most other eukaryotes). The relative abundance of the peptide modified as $TK_{me3}QTAR$ is calculated as:

intensity $TK_{me3}QTAR$*/ (intensity TKQTAR + intensity* $TK_{me1}QTAR$ *+* $TK_{me2}QTAR$ *+ intensity* $TK_{me3}QTAR$*)*

We strongly recommend to perform the extracted ion chromatography of histone peptides using EpiProfile, as this calculation is already performed, and the software automatically deals with isobaric peptides.

3. Perform the t-test when comparing two conditions or ANOVA when comparing more than two conditions. Data can be displayed as a volcano plot, using for the *x*-axis the log2 fold change between the two conditions and as *y*-axis the –log2 of the t-test *p*-value (Fig. 5c). Conventionally, the significance threshold is set at a *p*-value < 0.05, which when transformed as –log2 becomes >4.32.
4. Sum the relative abundance of all acetylated peptides, and compare them with the relative changes of acetyl-CoA (Fig. 5d). If the correlation is linear and positive, a possible biological interpretation of the data is that the acetyl-CoA levels change affecting the abundance of histone acetylation (Fig. 5f). An example of this analysis is illustrated in [36].
5. For validation of the findings, we recommend performing Western blotting (Fig. 5e) [37]. Ensure that the used antibody is specific for the acetylation site investigated.
6. If using metabolically labeled acetylations, it is also possible to estimate their turnover rate by dividing the relative abundance of the acetylated peptide with heavy labeled acetylation from the one with light acetylation. *Important*: Do not confuse the turnover rate with the relative abundance. The turnover rate indicates how frequently an acetylation is recycled with a new one, while the relative abundance indicates how much of that site is acetylated. An acetylation might change in abundance between two conditions, but maintain its turnover rate, or vice versa.
7. Once identified the PTMs that are significantly regulated between the analyzed conditions, a potential next step can be determining which histone writer is potentially responsible for this regulation. This enzyme is a potential target for complementary treatment using either inhibitors or other post-transcriptional regulation (e.g., knock-down). A comprehensive list of known histone modifications and their respective writers has been recently published in [38].

4 Notes

1. Number of cells plated for every single experiment should be optimized. It is important to avoid confluency at the moment of harvesting. Factors to take into consideration include, but not limited to:
   - Growth rate.
   - Length of the experiment.
   - Viability upon treatment.
2. Nutrient availability can be optimized for any specific cell line/condition. We found that the above-described concentrations mimic well a physiological situation in most cell lines we tested. Nutrient sensing and impact of acetyl-CoA availability on histone acetylation can be well appreciated under these culture conditions. However, some primary well-differentiated cell types display minimal nutrient consumption. We recommend lower nutrient availability when metabolic activity is low.
3. The high-salt extraction protocol [39] is alternative to acid extraction for histone purification. High-salt is a milder procedure, and it preserves acid-labile PTMs. However, it produces samples with very high concentrated salt, preventing an effective LC-MS analysis. Desalting can be performed as described in Subheading 3.7, but it is not 100% effective.
4. Fractionation of intact histone isotypes can be performed by using HPLC-UV. It requires at least 100 μg of starting material (if 2.1 mm ID column is used) or 300 μg (if 4.6 mm ID column is used). Given the optimized nanoLC-MS and the EpiProfile software for the analysis of the runs, histone fractionation is not recommended. It might be convenient for scarcely pure histone extractions and if interested in very low abundance PTMs.
5. In case estimating the yield of extracted histones is prohibitive, it is reasonable to assume that standard procedures extract about 1 μg of histone every 1 μL of cell pellet. Unpublished data from our lab demonstrate that histone analysis can be performed with as low as 50,000 cells as starting material.
6. NH_4OH, glacial acetic acid, and propionic anhydride should be used in the fume hood. The bottle of propionic anhydride must be filled with argon after its use to preserve the reagent.
7. In presence of multiple samples (>3–4), consider re-preparing the propionylation mix every 3–4 samples. This mix is very reactive, and using it for a long list of samples can prevent its efficacy. In case of preparation of a large number of samples (>20), consider performing the reaction in a 96-well plate using a multichannel pipette.

8. If the pH of the propionylation reaction is acidic, no reaction will occur. In case the pH is >10.0, other amino acid residues with higher pKa might be labeled as well, generating issues in the proper identification and quantification of histone peptides by LC-MS.
9. Buffers can be pushed through stage tips by using centrifugation instead of air pressure. Use appropriate holders on the top of the tube to place the stage tip, or drill a hole on the cap. This approach is not recommended, as it is harder to prevent complete drying of the stage tip during the procedure.
10. When preparing a nano column, the top can be capped using a "frit" instead of pulling a tip. Frits are prepared by mixing 88 μL Kasil® into a 0.5 mL tube with 16 μL formamide. One end of the capillary is dipped into this solution, and it is left for polymerization in a heater at ~110 °C for 3–4 h. fritted columns require a connection with a tip for spraying a sample into MS.
11. If the tip of the nano column was not prepared properly, two issues can be encountered. If the tip has a too small orifice, columns will be plugged and they cannot be utilized. This issue can be solved by carefully cutting the very top of the tip or by torching for a few seconds the tip while the column is packing. If the tip has been cut with a too large orifice, the column will not retain the particles, and they will geyser out from the tip while packing. In this case, immediately close the pressure bomb and discard the capillary.
12. For metabolite analysis in vivo, freeze clamping of tissue of interest in alive, deeply anesthetized is often recommended, but rarely feasible. We propose to rapidly sacrifice the animal (cervical dislocation for rodents), and quickly expose the tissue of interest. Everything must take less than 1 min. Pre-chill a tissue clamp in liquid nitrogen. Various models are commercially available and equivalent; alternatively, a toothed forceps/scissor with large, flat extremities can be used. Snap freeze a chunk of tissue then cut approximately 50 mg of tissue in a superchilled ceramic tile on dry ice. Weight the tissue with a precision scale.
13. Washing cells is not necessary as acetyl-CoA is not present in the cell medium and washing can skew metabolite quantitation. It is, however, important to minimize cell medium in the sample so as to prevent excessive salt and matrix effects from media components in extraction and MS acquisition.
14. Acyl-CoA analysis should be performed as quickly as possible after extraction to avoid sample degradation. Acetyl-CoA is relatively stable for several days at 4 °C in 5-sulfosalicylic acid, with ~50% sample loss after 15 days [40]. Low pH helps to

minimize hydrolysis of the acyl-CoA thioester bond [41]. 5-sulfosalicylic acid is also an antimicrobial agent.

15. For analysis of metabolites like acetyl-CoA, high-flow HPLC (flow rate of 100–200 μL/min) is currently preferred to nanoLC due to robustness. High-flow HPLC requires columns of larger diameter than nanoLC, thus not packed in-house. We personally tested and considered reliable the following commercial columns: Acquity HSS T3 column (2.1 × 150 mm, 1.7 μm particle size), Phenomenex HPLC Luna C18 (2.0 × 150 mm, 5 μm particle size), and waters XBridge C18 (2.1 × 150 mm, 3.5 μm particle size).
16. Calculation of % isotopic enrichment involves applying a correction matrix that compensates for the nonlinearity of isotopic enrichment [42].

Acknowledgment

S.S. and B.A.G. gratefully acknowledge the NIH grants CA196539, GM110174, and AG031862. A.C. gratefully acknowledges the AACR—Pancreatic Cancer Action Network Career Development Award—for funding his research in Kathryn E. Wellen's laboratory. S.T. acknowledges the American Diabetes Association grant #1-18-PDF-144. All authors want to acknowledge Kathryn E. Wellen for editorial support, funding of A.C.'s and S.T.'s work, and inspiring discussions.

References

1. Kouzarides T (2007) Chromatin modifications and their function. Cell 128(4):693–705. https://doi.org/10.1016/j.cell.2007.02.005
2. Fujisawa T, Filippakopoulos P (2017) Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. Nat Rev Mol Cell Biol 18(4):246–262. https://doi.org/10.1038/nrm.2016.143
3. de Ruijter AJ, van Gennip AH, Caron HN, Kemp S, van Kuilenburg AB (2003) Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem J 370 (Pt 3):737–749. https://doi.org/10.1042/BJ20021321
4. Carrer A, Wellen KE (2015) Metabolism and epigenetics: a link cancer cells exploit. Curr Opin Biotechnol 34:23–29. https://doi.org/10.1016/j.copbio.2014.11.012
5. Shi L, Tu BP (2015) Acetyl-CoA and the regulation of metabolism: mechanisms and consequences. Curr Opin Cell Biol 33:125–131. https://doi.org/10.1016/j.ceb.2015.02.003
6. Peleg S, Feller C, Ladurner AG, Imhof A (2016) The metabolic impact on histone acetylation and transcription in ageing. Trends Biochem Sci 41(8):700–711. https://doi.org/10.1016/j.tibs.2016.05.008
7. Lee S, Lee HC, Kwon YW, Lee SE, Cho Y, Kim J, Lee S, Kim JY, Lee J, Yang HM, Mook-Jung I, Nam KY, Chung J, Lazar MA, Kim HS (2014) Adenylyl cyclase-associated protein 1 is a receptor for human resistin and mediates inflammatory actions of human monocytes. Cell Metab 19(3):484–497. https://doi.org/10.1016/j.cmet.2014.01.013
8. Kinnaird A, Dromparis P, Saleme B, Gurtu V, Watson K, Paulin R, Zervopoulos S, Stenson T, Sutendra G, Pink DB, Carmine-Simmen K, Moore R, Lewis JD, Michelakis ED (2016) Metabolic modulation of clear-cell renal cell carcinoma with dichloroacetate, an inhibitor of pyruvate dehydrogenase kinase. Eur Urol 69(4):734–744. https://doi.org/10.1016/j.eururo.2015.09.014

9. Das S, Morvan F, Jourde B, Meier V, Kahle P, Brebbia P, Toussaint G, Glass DJ, Fornaro M (2015) ATP citrate lyase improves mitochondrial function in skeletal muscle. Cell Metab 21 (6):868–876. https://doi.org/10.1016/j.cmet.2015.05.006
10. Zhao S, Torres A, Henry RA, Trefely S, Wallace M, Lee JV, Carrer A, Sengupta A, Campbell SL, Kuo YM, Frey AJ, Meurs N, Viola JM, Blair IA, Weljie AM, Metallo CM, Snyder NW, Andrews AJ, Wellen KE (2016) ATP-citrate lyase controls a glucose-to-acetate metabolic switch. Cell Rep 17(4):1037–1052. https://doi.org/10.1016/j.celrep.2016.09.069
11. Mews P, Donahue G, Drake AM, Luczak V, Abel T, Berger SL (2017) Acetyl-CoA synthetase regulates histone acetylation and hippocampal memory. Nature 546(7658):381–386. https://doi.org/10.1038/nature22405
12. Comerford SA, Huang Z, Du X, Wang Y, Cai L, Witkiewicz AK, Walters H, Tantawy MN, Fu A, Manning HC, Horton JD, Hammer RE, McKnight SL, Tu BP (2014) Acetate dependence of tumors. Cell 159 (7):1591–1602. https://doi.org/10.1016/j.cell.2014.11.020
13. Basu SS, Blair IA (2011) SILEC: a protocol for generating and using isotopically labeled coenzyme A mass spectrometry standards. Nat Protoc 7(1):1–12. https://doi.org/10.1038/nprot.2011.421
14. Snyder NW, Tombline G, Worth AJ, Parry RC, Silvers JA, Gillespie KP, Basu SS, Millen J, Goldfarb DS, Blair IA (2015) Production of stable isotope-labeled acyl-coenzyme A thioesters by yeast stable isotope labeling by essential nutrients in cell culture. Anal Biochem 474:59–65. https://doi.org/10.1016/j.ab.2014.12.014
15. Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, Cheung MS, Day DS, Gadel S, Gorchakov AA, Gu TT, Kharchenko PV, Kuan S, Latorre I, Linder-Basso D, Luu Y, Ngo Q, Perry M, Rechtsteiner A, Riddle NC, Schwartz YB, Shanower GA, Vielle A, Ahringer J, Elgin SCR, Kuroda MI, Pirrotta V, Ren B, Strome S, Park PJ, Karpen GH, Hawkins RD, Lieb JD (2011) An assessment of histone-modification antibody quality. Nat Struct Mol Biol 18 (1):91–93. https://doi.org/10.1038/Nsmb.1972
16. Sidoli S, Cheng L, Jensen ON (2012) Proteomics in chromatin biology and epigenetics: elucidation of post-translational modifications of histone proteins by mass spectrometry. J Proteomics 75(12):3419–3433. https://doi.org/10.1016/j.jprot.2011.12.029
17. Simithy J, Sidoli S, Garcia BA (2018) Integrating proteomics and targeted metabolomics to understand global changes in histone modifications. Proteomics 18(18):e1700309. https://doi.org/10.1002/pmic.201700309
18. Faubert B, Li KY, Cai L, Hensley CT, Kim J, Zacharias LG, Yang C, Do QN, Doucette S, Burguete D, Li H, Huet G, Yuan Q, Wigal T, Butt Y, Ni M, Torrealba J, Oliver D, Lenkinski RE, Malloy CR, Wachsmann JW, Young JD, Kernstine K, DeBerardinis RJ (2017) Lactate metabolism in human lung tumors. Cell 171 (2):358–371 e359. https://doi.org/10.1016/j.cell.2017.09.019
19. Sun RC, Fan TW, Deng P, Higashi RM, Lane AN, Le AT, Scott TL, Sun Q, Warmoes MO, Yang Y (2017) Noninvasive liquid diet delivery of stable isotopes into mouse models for deep metabolic network tracing. Nat Commun 8 (1):1646. https://doi.org/10.1038/s41467-017-01518-z
20. Jang C, Hui S, Lu W, Cowan AJ, Morscher RJ, Lee G, Liu W, Tesz GJ, Birnbaum MJ, Rabinowitz JD (2018) The small intestine converts dietary fructose into glucose and organic acids. Cell Metab 27(2):351–361 e353. https://doi.org/10.1016/j.cmet.2017.12.016
21. Govaert E, Van Steendam K, Scheerlinck E, Vossaert L, Meert P, Stella M, Willems S, De Clerck L, Dhaenens M, Deforce D (2016) Extracting histones for the specific purpose of label-free MS. Proteomics 16(23):2937–2944. https://doi.org/10.1002/pmic.201600341
22. Zhao X, Sidoli S, Wang L, Wang W, Guo L, Jensen ON, Zheng L (2014) Comparative proteomic analysis of histone post-translational modifications upon ischemia/reperfusion-induced retinal injury. J Proteome Res 13 (4):2175–2186. https://doi.org/10.1021/pr500040a
23. Bonaldi T, Imhof A, Regula JT (2004) A combination of different mass spectroscopic techniques for the analysis of dynamic changes of histone modifications. Proteomics 4 (5):1382–1396. https://doi.org/10.1002/pmic.200300743
24. Garcia BA, Mollah S, Ueberheide BM, Busby SA, Muratore TL, Shabanowitz J, Hunt DF (2007) Chemical derivatization of histones for facilitated analysis by mass spectrometry. Nat Protoc 2(4):933–938. https://doi.org/10.1038/nprot.2007.106
25. Liao R, Wu H, Deng H, Yu Y, Hu M, Zhai H, Yang P, Zhou S, Yi W (2013) Specific and efficient N-propionylation of histones with propionic acid N-hydroxysuccinimide ester for histone marks characterization by LC-MS.

Anal Chem 85(4):2253–2259. https://doi.org/10.1021/ac303171h

26. Maile TM, Izrael-Tomasevic A, Cheung T, Guler GD, Tindell C, Masselot A, Liang J, Zhao F, Trojer P, Classon M, Arnott D (2015) Mass spectrometric quantification of histone post-translational modifications by a hybrid chemical labeling method. Mol Cell Proteomics 14(4):1148–1158. https://doi.org/10.1074/mcp.O114.046573
27. Sidoli S, Bhanu NV, Karch KR, Wang X, Garcia BA (2016) Complete workflow for analysis of histone post-translational modifications using bottom-up mass spectrometry: from histone extraction to data analysis. J Vis Exp (111). https://doi.org/10.3791/54112
28. Sidoli S, Garcia BA (2017) Characterization of individual histone posttranslational modifications and their combinatorial patterns by mass spectrometry-based proteomics strategies. Methods Mol Biol 1528:121–148. https://doi.org/10.1007/978-1-4939-6630-1_8
29. Sidoli S, Simithy J, Karch KR, Kulej K, Garcia BA (2015) Low resolution data-independent acquisition in an LTQ-Orbitrap allows for simplified and fully untargeted analysis of histone modifications. Anal Chem 87 (22):11448–11454. https://doi.org/10.1021/acs.analchem.5b03009
30. Sidoli S, Lin S, Xiong L, Bhanu NV, Karch KR, Johansen E, Hunter C, Mollah S, Garcia BA (2015) Sequential window acquisition of all theoretical mass spectra (SWATH) analysis for characterization and quantification of histone post-translational modifications. Mol Cell Proteomics 14(9):2420–2428. https://doi.org/10.1074/mcp.O114.046102
31. Sidoli S, Fujiwara R, Garcia BA (2016) Multiplexed data independent acquisition (MSX-DIA) applied by high resolution mass spectrometry improves quantification quality for the analysis of histone peptides. Proteomics 16(15–16):2095–2105. https://doi.org/10.1002/pmic.201500527
32. Yuan ZF, Lin S, Molden RC, Cao XJ, Bhanu NV, Wang X, Sidoli S, Liu S, Garcia BA (2015) EpiProfile quantifies histone peptides with modifications by extracting retention time and intensity in high-resolution mass spectra. Mol Cell Proteomics 14(6):1696–1707. https://doi.org/10.1074/mcp.M114.046011
33. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26(7):966–968. https://doi.org/10.1093/bioinformatics/btq054
34. Frey AJ, Feldman DR, Trefely S, Worth AJ, Basu SS, Snyder NW (2016) LC-quadrupole/Orbitrap high-resolution mass spectrometry enables stable isotope-resolved simultaneous quantification and (1)(3)C-isotopic labeling of acyl-coenzyme A thioesters. Anal Bioanal Chem 408(13):3651–3658. https://doi.org/10.1007/s00216-016-9448-5
35. Trefely S, Ashwell P, Snyder NW (2016) FluxFix: automatic isotopologue normalization for metabolic tracer analysis. BMC Bioinformatics 17(1):485. https://doi.org/10.1186/s12859-016-1360-7
36. Kori Y, Sidoli S, Yuan ZF, Lund PJ, Zhao X, Garcia BA (2017) Proteome-wide acetylation dynamics in human cells. Sci Rep 7(1):10296. https://doi.org/10.1038/s41598-017-09918-3
37. Kurien BT, Scofield RH (2015) Western blotting: an introduction. Methods Mol Biol 1312:17–30. https://doi.org/10.1007/978-1-4939-2694-7_5
38. Zhao Y, Garcia BA (2015) Comprehensive catalog of currently documented histone modifications. Cold Spring Harb Perspect Biol 7(9): a025064. https://doi.org/10.1101/cshperspect.a025064
39. Vonholt C, Brandt WF, Greyling HJ, Lindsey GG, Retief JD, Rodrigues JD, Schwager S, Sewell BT (1989) Isolation and characterization of histones. Methods Enzymol 170:431–523
40. Demoz A, Garras A, Asiedu DK, Netteland B, Berge RK (1995) Rapid method for the separation and detection of tissue short-chain coenzyme A esters by reversed-phase high-performance liquid chromatography. J Chromatogr B Biomed Appl 667(1):148–152
41. Bishop JE, Hajra AK (1980) A method for the chemical synthesis of 14C-labeled fatty acyl coenzyme A's of high specific activity. Anal Biochem 106(2):344–350
42. Fernandez CA, Des Rosiers C, Previs SF, David F, Brunengraber H (1996) Correction of 13C mass isotopomer distributions for natural stable isotope abundance. J Mass Spectrom 31(3):255–262. https://doi.org/10.1002/(SICI)1096-9888(199603)31:3<255::AID-JMS290>3.0.CO;2-3

Chapter 10

Methods to Measure Autophagy in Cancer Metabolism

Natalia von Muhlinen

Abstract

Autophagy, a dynamic pathway in which intracellular membrane structures sequester portions of the cytosol for degradation, plays multiple roles in physiological and pathological processes. Autophagy may have suppressive and promotive roles in the formation and progression of cancer. A growing number of methods to identify, quantify, and manipulate autophagy have been developed. Because most of these methods are semiquantitative and have significant limitations, it is important to emphasize that a combination of these assays is recommended for the analysis of autophagy. Here, I briefly discuss the autophagic process, its role in disease, and I summarize some of the best-known and most widely used methods to study autophagy in vitro in the context of cancer, including transmission electron microscopy (TEM), detection and quantification of the autophagy protein LC3 by western blot, and the use of GFP-LC3 to quantify puncta by fluorescence microscopy and tandem labeled RFP/mCherry-GFP-LC3 fluorescence microscopy to measure autophagic flux.

Key words Macroautophagy, Autophagy, Autophagy flux, Cancer, Metabolism, LC3, GFP-LC3, Atg8, Cancer therapy

1 Introduction

1.1 The Autophagic Process

Autophagy, a process evolutionary conserved from yeast to mammals, involves the formation of a small vesicular sac called the isolation membrane or phagophore, which subsequently elongates and encloses a portion of cytoplasm, resulting in the formation of the autophagosome [1–6]. Many types of autophagy have been described, including nonselective and selective processes [5]. Macroautophagy (hereafter referred to as "autophagy"), the nonselective autophagic degradation of cytosolic content, is the most extensively studied type of autophagy and will be the focus of this chapter. This evolutionary conserved mechanism is mediated by a specific organelle, the autophagosome, a double-membraned structure that delivers cytoplasmic material and autophagy substrates to the lysosome, resulting in the lysosomal degradation of the enclosed materials (Fig. 1) [3, 4].

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_10,

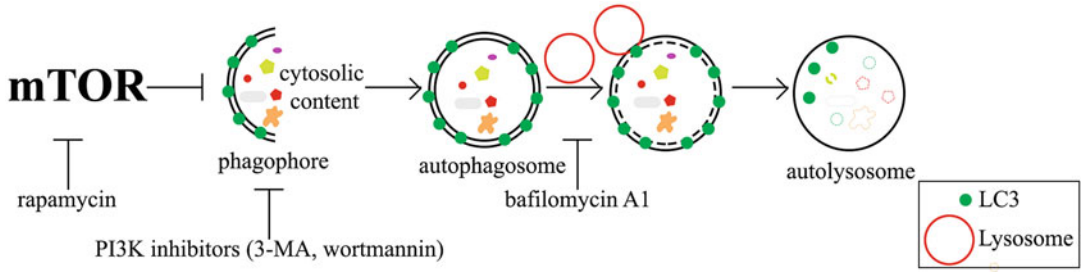


Fig. 1 The autophagy process. Summary of the main steps of the autophagy pathway. Cytosolic content, including organelles, is engulfed by an isolation membrane or phagophore to form an autophagosome. Subsequently, the outer membrane of the autophagosome fuses with the lysosome, and the enclosed cytosolic material is degraded in the autolysosome. Some autophagy inducers and inhibitors are indicated. Inhibition of mammalian target of rapamycin (mTOR) induces autophagy; 3-methyladenine (3-MA) inhibits the formation of autophagosomes; bafilomycin A1 inhibits the fusion of autophagosome with lysosomes. The LC3 protein is the only autophagy-related protein that remains attached to the autophagosomal membrane throughout the entire autophagy pathway

Upon induction of autophagy, a small isolation membrane also known as the "phagophore" sequesters portions of the cytosol and subsequently elongates, forming an enclosed double-membrane compartment (autophagosome) (Fig. 1). The outer membrane of the autophagosome then fuses with lysosomes, forming the "autolysosome," leading to the degradation of the enclosed cytosolic material [7]. This process involves a set of evolutionary conserved genes known as the Atg proteins [8, 9], collectively referred to as the "autophagic machinery," which are required for the formation and maturation of the autophagosome. Other genes encode proteins involved downstream of autophagosome formation, including proteins that mediate the autophagosome-lysosome fusion, as well as degradative lysosomal enzymes that break down the cargo, among others [10, 11]. Of all the proteins involved in autophagosome formation, only the microtubule-associated protein light chain 3 (LC3), the mammalian homologue of yeast Atg8, exists in autophagosomes and remains attached to the double membrane, and thus it is widely used as a marker of autophagosomes [12]. The methods described in this chapter detect the various stages of autophagosome formation and should be used in combination to determine the cellular autophagic status (induction or inhibition).

1.2 Autophagy in the Pathogenesis of Cancer

Autophagy has several well-characterized vital roles under physiological conditions, including but not limited to maintenance of the amino acid pool during starvation, clearance of intracellular microbes, prevention of neurodegeneration, antiaging, cancer, clearance of intracellular microbes, and regulation of immunity [13–16]. Autophagy dysfunction has been shown to contribute to several diseases, including cancer. Increasing evidence suggests the

importance of autophagy in cancer [17, 18], although whether autophagy protects or promotes cancer is still not well understood.

Autophagy affects several cellular functions through the alteration of metabolism, the proteome, organelle quantity and quality [5, 15, 16, 19]. In addition, autophagy alters the interaction between the tumor and its microenvironment by promoting stress adaptation, inhibiting innate and adaptive immunity, and supporting tumor growth in a nutrient-limited microenvironment. Genomic analysis of human cancer samples suggests core autophagy genes are rarely mutated or lost in cancer [20], while oncogenic mutations that dysregulate autophagy and lysosomal biogenesis have been identified. Functional analysis of autophagy, however, is very difficult due to the lack of methods that measure the autophagy status in tumor samples.

Several proteins and pathways that regulate autophagy are altered in cancer [17–19]. Although initially it was believed that beclin 1 (BECN1), the mammalian orthologue of yeast Atg6, was involved in cancer due to its loss in some breast, prostate, and ovarian cancers [21–23], it was later found that *BECN1* loss was associated with mutation in the well-known tumor suppressor *BRCA1* in chromosome 17 [24, 25]. Hereditary breast cancer is commonly a result of deletion of the wild-type *BRCA1* allele. Due to the proximity of *BECN1*, large deletions encompassing both *BRCA1* and *BECN1* have also been shown in breast and ovarian cancer [24, 25]. Analysis in The Cancer Genome Atlas (TCGA) demonstrated that allelic loss of *BECN1*1 is dependent deletion of *BRCA1*; therefore, deletion of *BECN1* is a consequence of *BRCA1* deletion rather than a driver of breast cancer [24, 25]. However, the reduced *BECN1* that occurs due to *BRCA1* deletion may reduce autophagy in breast and ovarian cancer, and thus, cancer response to stress and nutrient starvation may be impaired, which may be a therapeutic target for cancer therapy to consider.

Several signaling mechanisms modulate autophagy; however information about the processes that regulate autophagy in normal versus cancer cells is not well understood [26]. The signals that induce autophagy in cancer are believed to be stressful conditions including anticancer treatments or cellular starvation that results from the rapid proliferation of cancer cells [19]. Because many of these regulatory mechanisms are altered in cancer, they may be candidate therapeutic targets. For instance, the class I PI3K-AKT-mTOR pathway, activated in many types of cancer through growth factor receptors, is important in regulating autophagy [17, 18, 27, 28]. The serine-threonine kinase AKT, downstream of PI3K kinase, suppresses autophagy by activating the mTOR kinase, an autophagy inhibitor. Oncogenic forms of RAS negatively control autophagy through the activation of PI3K kinase [29]. Another regulator of the class I PI3K-AKT-mTOR pathway is the tumor suppressor gene phosphatase and tensin homologue (*PTEN*) [30]. PTEN

inhibits class I PI3K, which thereby inhibits AKT and induces autophagy initiation. Mutations in *PTEN*, common in several types of cancer, including malignant gliomas and prostate and brain cancer, result in constitutive activation of ALK and, thus, activation of mTOR and inhibition of autophagy [31]. Several in vitro and clinical studies have shown that rapamycin, an inhibitor of mTOR, induces autophagy and inhibits the proliferation of cancer cells, including malignant glioma cells, making it a putative therapeutic candidate for treatment of many tumor types [32].

In conclusion, there are several lines of evidence to suggest that autophagy is a key component of tumorigenesis and that its modulation is a potential therapeutic target. Improvement of cancer treatment may be achieved in several ways related to autophagy, including (1) the manipulation or activation of autophagy-related proteins or modulators such as BECN1 or PTEN, which may lead to cell death or inhibition of cellular proliferation; (2) using mTOR inhibitors in such way that the autophagy-inducing drug rapamycin might repress proliferation in cells with an intact autophagy pathway; and (3) in tumor cells in which autophagy mediates drug resistance, autophagy inhibitors such as bafilomycin might sensitize cells to therapeutic agents by converting the autophagic pathway to an apoptotic process. More extensive research in the role of autophagy in cancer and its potential use as a therapeutic target need to be performed. Here, I present a review of current methods to study autophagy in cancer metabolism.

2 Materials

2.1 Reagents for Cell Culture

1. Dulbecco's modified eagle's medium (DMEM) cell culture medium, supplemented with 10% v/v fetal bovine serum (FBS), 2 mM L-Glutamine and antibiotics.
2. PBS (1×): 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na2HPO4, and 1.4 mM KH_2PO_4 in deionized water (dH_2O), adjust pH to 7.4 with 2 *N* NaOH.

2.2 Reagents for Electrophoresis and Western Blot

1. RIPA lysis buffer: (150 mM sodium chloride, 1.0% NP-40 or Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS (sodium dodecyl sulfate), 50 mM Tris, pH 8.0), 1× EDTA-free protease inhibitor cocktail tablets.
2. Sample buffer: 2× SDS loading buffer (4% SDS, 4% SDS, 20% glycerol, 0.004% bromophenol blue, 0.125 M Tris–HCl, pH 6.8), 10% 2-β-mercaptoethanol.
3. 10–20% Tris-glycine electrophoresis gels.
4. Precision plus protein dual color standards.

5. 1× Tris-Glycine Running Buffer: 25 mM Tris, 192 mM glycine, 0.1% SDS.
6. 1× Transfer Buffer: 25 mM Tris, 192 mM glycine, 20% (v/v) methanol, 0.025–0.1% SDS (pH 8.3)
7. Invitrolon™ PVDF/Filter Paper Sandwich.
8. Tris-buffered saline Tween (TBST): 50 mM Tris–HCL (pH = 7.5), 150 mM NaCl, and 0.1% Tween 20.
9. Bovine serum albumin (BSA), stock solution of 10 mg/mL in dH_2O, stored at 4 °C.
10. Blocking buffer: 0.1% Tween-20 (Sigma-Aldrich) and 5% (w/v) nonfat powdered milk (commonly found in food stores) in TBST.
11. Primary antibody: anti-LC3 rabbit polyclonal antibody.
12. Primary antibody: anti-β-actin monoclonal antibody.
13. Secondary antibody: peroxidase-conjugated affiniPure goat anti-rabbit IgG (H + L).
14. Secondary antibody: peroxidase-conjugated affiniPure goat anti-mouse IgG (H + L).
15. SuperSignal West Pico Stable Peroxide Solution.

2.3 Reagents for immunofluorescence

1. BSA buffer (blocking buffer): 3 mg/mL BSA in PBS.
2. Primary antibody: anti-LC3 rabbit polyclonal antibody.
3. Secondary antibody: Anti-rabbit Alexa Fluor® 568 Goat Molecular Probes A-11036.
4. VECTASHIELD Antifade Mounting Medium with DAPI.
5. Confocal microscope: Zeiss LSM 510 (Carl Zeiss AG, Oberkochen, Germany).

3 Methods

Drugs that potentially affect autophagy are being tested in clinical trials. New potential modulators of autophagy are also being screened for therapeutic purposes. Thus, it is important to establish basic guidelines of methods to measure autophagy. Conventional methods for detecting autophagic cells include ultrastructural analysis and protein degradation assays and have been extensively reviewed previously [33–36]. Here I present a summarized guideline of how to perform and interpret these methods.

One critical point to understand is that autophagy is a highly dynamic process, and like many other pathways, it can be positively and negatively modulated at every step, impacting the generation of phagophores, the maturation of autophagosomes, and/or their fusion with lysosomes (Fig. 2). Thus, the use of multiple

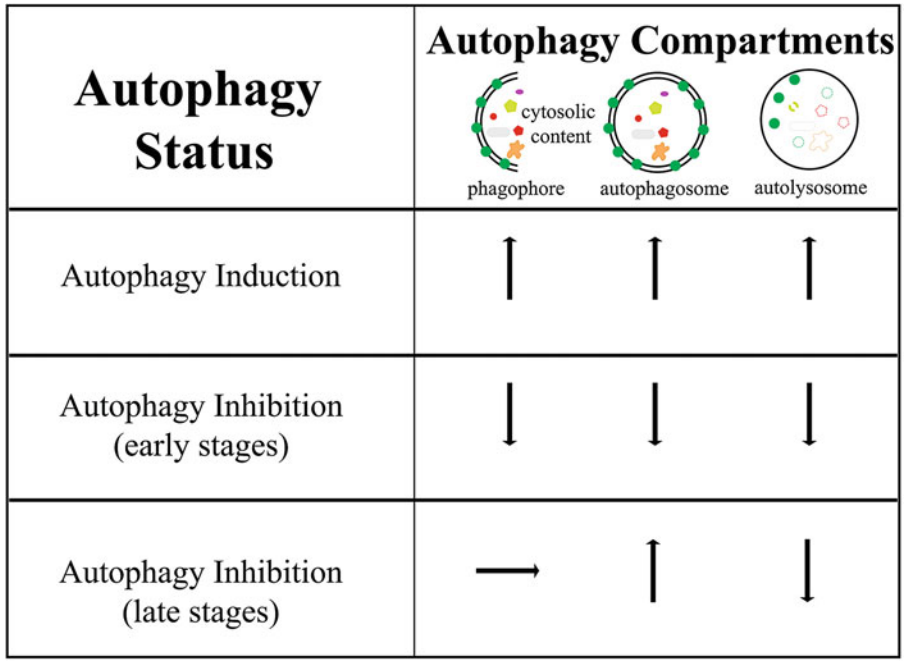


Fig. 2 Effect of autophagy induction or inhibition. The relative amounts of phagophores, autophagosome, or autolysosomes in each condition are depicted. When autophagy is induced, levels of autophagic compartments increase. In contrast, when autophagy is inhibited at the initial stages, the relative amounts of all autophagic membranes decrease. Finally, when autophagy is inhibited at the later stages (such as fusion between autophagosome and lysosomes), the levels of phagophore formation remain the same, while autophagosomes accumulate due to the inhibition of lysosomal degradation

complimentary assays is essential to measure the overall autophagy flow and to make a comprehensive interpretation of the results. The measurement of autophagic flux, that is, the entire process of autophagy, includes not only the increase in lipidation of LC3 but also the delivery of cargo to lysosomes and their subsequent breakdown and release into the cytosol. Here I summarize the most utilized current assays to study autophagy.

3.1 Transmission Electron Microscopy

1. Treat cultured cells with autophagy modulators or drugs of interest.
2. Pellet cells upon treatment for different timepoints.
3. Fix cells with a solution containing 3% glutaraldehyde plus 2% paraformaldehyde in 0.1 M cacodylate buffer.
4. Wash pellets in 0.2 M HEPES, pH 7.4.
5. Wash with PBS at least two to three times.
6. Incubate pellets in 1% osmium tetroxide in water at RT for 1 h.
7. Wash pellets in water.

8. Stain the pellets in 2% uranyl acetate in water at room temperature (RT), in the dark, for 1 h.
9. Pellets must be dehydrated using ethanol at increasing concentrations at RT: 70% ethanol for 15 min; 95% ethanol for 15 min; 100% ethanol, twice for 15 min; and propylene oxide for 20 min.
10. Infiltration with resin: Incubate the pellets in a mixture of resin and propylene oxide (1:1) at RT for 2 h first and then in 100% resin at RT overnight.
11. Transfer the pellets to fresh resin in beam capsules (Agar Scientific Cat. No. G362 or G360) (Glauert and Lewis, 1998), and incubate at RT for 4–6 h.
12. Examination is performed at 80 kV under a transmission electron microscope, on ultrathin sections (80 nm) stained with 0.1% lead citrate and 10% uranyl acetate (*see* **Notes 1** and **2**).

3.2 Monitoring LC3 by Western Blot to Determine Autophagy Status or Autophagy Flux

3.2.1 Cell Plating

1. Plate cells in 6-well tissue culture plates in technical and biological replicates to 50–70% confluency.
2. Allow cells to attach to the bottom of the plate at least overnight before performing treatments.

3.2.2 Treatment with Autophagy Inducers or Inhibitors as Controls

1. Rinse cells with PBS before adding cell culture media containing autophagy modulators.
2. For autophagy induction, use DMEM without essential amino acids (starvation media) to measure starvation-induced autophagy.
3. For inhibition of various stages of the autophagy pathway, use 3-MA or bafilomycin A1 (10 nM final concentration) (*see* **Note 3**).
4. Incubate treated cells for several timepoints (1 h, 2 h, 4 h, 24 h) to study the early and late effects in the autophagy pathway of the treatment (*see* **Note 4**).

3.2.3 Sample Preparation

1. Lyse pelleted cells using RIPA buffer (*see* Subheading 2.2) after treatment, vortex to lyse cells, and incubate on ice for 20 min.
2. Centrifuge lysate at 16,000 × *g* for 15 min at 4 °C.
3. Transfer supernatant to a new collection tube and measure protein concentration.
4. Dilute lysate 1:1 in sample buffer (*see* Subheading 2.2), vortex, boil, centrifuge at 16,000 × *g*, and run fresh sample for electrophoresis.

3.2.4 Electrophoresis

1. Use a 10–20% Tris-glycine gradient gel for better separation of cytosolic and membrane-bound LC3-I and LC3-II forms (*see* **Notes 4** and **5**).
2. Mount the gel in an electrophoresis tank according to standard protocol.
3. Fill the top and the bottom of the reservoir with 1× tris-glycine running buffer (*see* Subheading 2.2).
4. Load equal amounts of protein (10–60 μg per lane) into each well of the gel, along with molecular weight markers.
5. Run the gel electrophoresis at 100 V for approximately 2 h or until the 10 kDa molecular weight band reaches the bottom of the gel.

3.2.5 Western Blotting

1. To transfer protein samples to PVDF membrane, soak the membrane first in methanol for 10–30 s, rinse off with water, and place in transfer buffer. Disassemble gel according to manufacturer's instruction, and mount the transfer cassette with the gel and membrane according to standard protocols. Transfer for 2 h at 25 V.
2. Block the membrane using blocking buffer containing 5% dry milk in TBST (*see* Subheading 2.2) for 1 h at RT.
3. Incubate membrane with LC3 antibody diluted 1:1000 in blocking buffer (*see* Subheading 2.2) overnight in a cold room.
4. Rinse with TBST for about 20 min to remove excess of primary antibody.
5. Incubate with secondary antibody (*see* Subheading 2.2) diluted in blocking buffer for 1 h at RT.
6. Rinse with PBS for about 20–30 min at room temperature.
7. Develop using SuperSignal West Pico according to the manufacturer's instructions.
8. Capture the chemiluminescence using conventional X-ray films or the method of your choice.
9. Probe with antitubulin or anti-β-actin to use as loading control (*see* **Note 6**).
10. LC3-I, LC3-II, and housekeeping bands can be quantified using free software such as ImageJ and used to determine autophagy status (*see* **Note 5**) and autophagy flux (*see* **Notes 6** and **7**).

3.3 Use of Green Fluorescence Protein (GFP)-Labeled LC3 to Monitor Autophagy by Fluorescence Microscopy

1. Stable GFP-LC3 cell lines can be generated by infecting cells with GFP-LC3 retrovirus or lentivirus, followed by a selection with a selectable marker such as blasticidin.
2. A homogenous cellular population expressing the desired levels of GFP can be selected by flow cytometry or by single-cell cloning. These cells can be used to analyze autophagy by fluorescence microscopy (*see* below and **Notes 8–10**) or by partial proteolysis using western blot (*see* **Note 11**).
3. Plate cells on glass coverslips at least 24 h before analysis.
4. Treat cells with autophagy modulators, such as bafilomycin (autophagy inhibitor, 10 nM final concentration) or serum-starved media (autophagy inducer) for different timepoints (1 h, 2 h, 4 h).
5. Wash cells on coverslips with PBS at least three times for 5 min, and fix with 4% formaldehyde for 10 min at RT.
6. Wash cells with PBS at least for 15 min.
7. Mount coverslips on slides using ProLong Gold antifade or other equivalent mounting media.
8. Allow slides to dry for 24 h at RT.
9. Visualize using a fluorescence microscope on 40× to 100× magnification.
10. Take pictures of multiple fields representing a total of cell population of 100 cells or more, and then count the number of GFP-LC3 puncta per cells either manually or using computerized software, for instance, Image J (*see* **Notes 8–10**).

3.4 Tandem RFP/mCherry-GFP Fluorescence Microscopy

1. Using retrovirus or lentivirus, generate stable cell lines expressing RFP or mCherry-GFP-tagged LC3. Following infection, select a stable expressing population by treating cells with a selectable marker, such as blasticidin (1ug/ml) for at least 7 days.
2. A cellular population expressing the desired levels of GFP homogenously can be selected by flow cytometry or by single-cell cloning.
3. Plate cells on glass coverslips at least 24 h before treatment.
4. Incubate cells at approximately 70% confluency with the drug of interest for the desired time—several timepoints are recommended.
5. Fix with 4% formaldehyde for 10 min at RT following washes of coverslips with PBS.
6. Wash formaldehyde off the coverslips with PBS at least for 15 min.

7. Mount coverslips on slides using ProLong Gold antifade or other equivalent mounting media.
8. Allow slides to dry for 24 h at RT.
9. Visualize using a fluorescence microscope on at least a 63× magnification using the alexa 488 nm channel to detect GFP and the alexa 594 nm to detect the RFP or mCherry signals.
10. Take pictures of multiple fields representing a total of cell population of 100 cells or more.
11. Count the number of puncta per cells either manually or using computerized software. Red only puncta correspond to acid compartments where the GFP signal is quenched, like the autolysosomes. Yellow puncta correspond to phagophores or autophagosomes where the GFP and the RFP/mCherry signal co-localize (*see* **Note 12**).

4 Notes

1. TEM can be used to follow the sequential morphological changes during the autophagy process, from the appearance of the double-membrane "phagophore" to the degradation in the autolysosomes [37]. Autophagosomes can be observed as double-membraned structures under a TEM microscope. However, because the autophagic process is continuous, the further classification of compartments into their subsets may be sometimes difficult. For instance, because in the autophagosome and autolysosome stages the compartment displays a double membrane, it may be difficult at times to distinguish them, in which they may be called simply as "autophagic vacuole" [38]. Autophagosomes, referred to as initial autophagic vacuoles (AVi), typically display a double membrane, which is visualized as a compartment with two parallel membrane layers (bilayer), containing cytosolic content or organelles [38]. Fusion of the autophagosome with lysosomes results in the formation of the autolysosome, structure that can sometimes be differentiated by the degradation status of its content. Immunolabeling of samples using commercial antibodies that detect LC3, one of the few proteins that remain attached to the autophagosome throughout the autophagy pathway, is also available.
2. There are some difficulties to consider when using TEM. It requires a lot of expertise and it is time-consuming. Proper identification of the autophagy structure is also critical for qualitative and quantitative purposes, which might be the most difficult criteria to meet when using TEM since many subcellular components might have double membranes,

including mitochondria, endoplasmic reticulum (ER), and endosomes. Therefore, TEM is best used in combination with other methods to accumulate conclusive data to ensure the proper measurement of autophagy in cultured cells or tissues.

3. Longer than 4 h incubation with bafilomycin may be toxic to some cell types.
4. The mammalian orthologues of Atg8 are subdivided in two families: the LC3 subfamily and the GABARAP subfamily [12, 39]. Isoforms in the LC3 subfamily are referred to as LC3 for simplicity reasons, and since LC3 remains attached to the autophagosomal membrane throughout the entire autophagic pathway [12], it has become the most widely used marker for autophagosome and the autophagic process. During the early process of autophagosome formation, LC3 is conjugated to phosphatidylethanolamine (PE), referred to as LC3-PE, which remains attached to the autophagosomal membrane during the autophagy process, from formation of the phagophore to the degradation of the double membrane of the lysosome [12]. While membrane-bound LC3-PE (also known as LC3-II) has a higher molecular weight than free cytosolic LC3 (also referred to as LC3-I), LC3-II shows faster electrophoretic mobility in SDS-PAGE gels probably due to increased hydrophobicity. Both LC3-I (approximately 16–18 kDa) and LC3-II (approximately 14–16 kDa) are detected by western blot using the protocol described (*see* Subheading 3.2) (Fig. 3). The amount of LC3-I and LC3-II can be quantified by Western blotting, and ratios of these isoforms compared to loading controls, such as actin, can be used to estimate the overall autophagic status of cells under induction or inhibition of autophagy [33, 34, 40]. It is important to note that the levels of LC3-I or LC3-II alone cannot be used as a measurement of autophagic flux. To this end, western blotting must be combined with other methods such as those described in the Subheading 3.
5. When performing western blot detection and quantification of LC3-II, levels should be compared to housekeeping genes such as β-actin. Stain-free gels can also be used to quantify the total amount of protein loaded and can be used as an alternative to housekeeping genes [40].
6. Autophagic flux is often calculated by the turnover or difference in the amount of LC3-II measured by western blot in the presence or absence of lysosomal degradation inhibitors such as bafilomycin A1 [33, 34, 40]. Saturating levels of this inhibitor can be used to measure LC3-II transit through the autophagy pathway. If autophagic flux is occurring, LC3-II will accumulate in the presence of an autophagy inhibitor. When analyzing

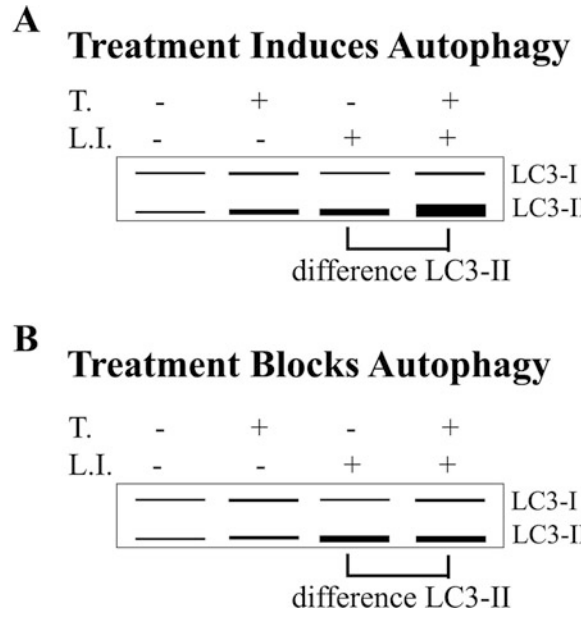


Fig. 3 Measuring autophagic flux by LC3 turnover. Detection of the conversion of cytosolic LC3-I to membrane-bound LC3-II under different treatment conditions by Western Blot. (**a**) If a treatment (T) alone increases the amount of LC3-II, and the treatment plus a lysosomal inhibitor (L.I., for instance, bafilomycin A1) has an additive effect compared to bafilomycin A1 alone may indicate that the treatment induces the formation of autophagic membranes. (**b**) If a treatment alone increases the amount of LC3-II, but the combination with bafilomycin A1 does not increase LC3-II levels compared to inhibitor alone, it may suggest that the treatment partially or completely blocks autophagy. *T* treatment; *L.I* lysosomal inhibitor

whether a treatment affects the autophagy pathway, LC3-II levels must be measured during treatment alone, inhibitor (such as bafilomycin) alone, combination of treatment and inhibitor, as well as no treatment (Fig. 3). An additive effect in increasing LC3-II levels of treatment plus inhibitor may suggest that the treatment increases autophagic flux. If the treatment plus inhibitor causes higher levels of LC3-II compared to inhibitor alone, it may suggest that the treatment increases the generation of autophagy membranes (Fig. 3a). If treatment by itself increases LC3-II but treatment plus inhibitor does not change the levels of LC3-II compared to inhibitor alone, it may suggest that the treatment causes a partial or complete block in autophagy flux (Fig. 3b) [41]. Positive control experiments with autophagy modulators that are known to induce autophagy are essential to make conclusions from these experiments.

7. When determining autophagic flux by LC3-II turnover, it is essential to use correct positive controls that induce autophagy flux, such as rapamycin or nutrient starvation. Care must be taken to use non-saturating and saturating inhibitor/treatment conditions. Furthermore, it is essential to test at least two timepoints after treatment, including early and late timepoints, for instance, 4 h and 24 h after treatment, since some compounds, such as calcium phosphate precipitates, might induce autophagy at early stages but block it at later timepoints [40].
8. When autophagy is not induced, GFP-LC3 is homogenously expressed in the cytosol. However, when autophagy is induced, GFP-LC3 puncta can be observed and quantified in a single-cell manner. It is recommended to use stable cellular transformants expressing GFP-LC3 versus using transient transfectants, since stable populations will homogenously express GFP-LC3, maximizing puncta counting efficiency [40]. As a negative control, a non-cleavable mutant of GFP-LC3 (GFP-LC3^{G120A}) can be used to measure auto fluorescence or non-specific effects of inhibitors or treatments [40]. Although this assay is time-consuming, it is highly sensitive and can be used to quantify autophagy induction and inhibition under different conditions. Importantly, it is recommended to count GFP-LC3 puncta per cell rather than the number of cells containing GFP-LC3 puncta [40] and represent the result as changes in the percentage of GFP-LC3 puncta per cell under basal or treatment conditions.
9. The GFP-tagged expression vector of LC3 is one of the most widely used methods to detect formation of autophagosome and monitor the autophagy process [12]. Measuring GFP-LC3 by fluorescence microscopy, while more time-consuming and tedious, is more sensitive and quantitative than measuring LC3-II turnover by western blot. Induction of autophagy results in formation of phagophores or pre-autophagosomes, which are labeled as GFP-LC3 structures and can be visualized as GFP-LC3 dots or puncta (*see* Subheading 3.3). GFP-LC3-expressing cell lines can also be used to monitor autophagy by quantification of cleaved GFP by western blot (*see* **Note 10**).
10. For GFP-LC3 puncta quantification, it is recommended to perform the experiment in technical and biological duplicates, using stable GFP-LC3 expressing transformants, and to quantify the number of puncta per cell rather than the number of cells containing GFP-LC3 puncta. Ideally, it is preferable that both assays are included when studying autophagy to accumulate conclusive evidence as to the status of the autophagic pathway and to compare both sets of results.

11. GFP-LC3 is a target of autophagy itself and thus has also been used to measure autophagic flux in transformed cells by Western blot. When GFP-LC3 is transported to the lysosome, the GFP-LC3 bond is sensitive to proteolysis in the lysosomal compartment, where it is cleaved and released to the cytosol as free GFP. Thus, measurement of the ratio of GFP-LC3 and free cleaved GFP can be monitored by western blot, using the western blot protocol described in Subheading 3.2. It is important to note that complementary methods such as turnover of LC3-II (*see* Subheading 3.2) and quantification of GFP-LC3 puncta (*see* Subheading 3.3) are recommended to be performed in parallel to make strong conclusions about cellular autophagy status.
12. The basis for the RFP (or mCherry)/GFP-tagged LC3 assay to measure autophagy is the sensitivity of GFP to the acidic/proteolytic environment of the lysosome, while RFP and mCherry are more stable [42]. In this assay, cells are treated with autophagy inducers or inhibitors, and localization of RFP/mCherry and GFP puncta is analyzed by microscopy to identify autophagosomal compartments following the protocol described in Subheading 3.4. Coverslips are then analyzed by fluorescence microscopy. Co-localization of RFP/mCherry with GFP (yellow puncta) indicates a phagophore or autophagosome compartment, while RFP/mCherry puncta without GFP signal corresponds to an autolysosome, since GFP signal is quenched in the lysosomal compartment. One of the advantages of this method is that it allows for the quantification of autophagic induction and flux without the need to use autophagy inhibitors. The use of early and late timepoints allows for the visualization of early phagophore formation as well as lysosomal degradation. Furthermore, this method can be used for high-throughput screening [40].

Acknowledgement

I am grateful to Dr. Curtis Harris and the National Institutes of Health, where I am currently a postdoctoral fellow.

References

1. Takeshige K, Baba M, Tsuboi S, Noda T, Ohsumi Y (1992) Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction. J Cell Biol 119 (2):301–311
2. Baba M, Takeshige K, Baba N, Ohsumi Y (1994) Ultrastructural analysis of the autophagic process in yeast: detection of autophagosomes and their characterization. J Cell Biol 124(6):903–913
3. Seglen PO, Bohley P (1992) Autophagy and other vacuolar protein degradation mechanisms. Experientia 48(2):158–172

4. Dunn WA Jr (1994) Autophagy and related mechanisms of lysosome-mediated protein degradation. Trends Cell Biol 4(4):139–143
5. Levine B, Klionsky DJ (2004) Development by self-digestion: molecular mechanisms and biological functions of autophagy. Dev Cell 6 (4):463–477
6. Klionsky DJ, Emr SD (2000) Autophagy as a regulated pathway of cellular degradation. Science 290(5497):1717–1721
7. Mizushima N, Noda T, Yoshimori T, Tanaka Y, Ishii T, George MD, Klionsky DJ, Ohsumi M, Ohsumi Y (1998) A protein conjugation system essential for autophagy. Nature 395 (6700):395–398. https://doi.org/10.1038/26506
8. Tsukada M, Ohsumi Y (1993) Isolation and characterization of autophagy-defective mutants of Saccharomyces cerevisiae. FEBS Lett 333(1-2):169–174
9. Harding TM, Morano KA, Scott SV, Klionsky DJ (1995) Isolation and characterization of yeast mutants in the cytoplasm to vacuole protein targeting pathway. J Cell Biol 131 (3):591–602
10. Rabinowitz JD, White E (2010) Autophagy and metabolism. Science 330 (6009):1344–1348. https://doi.org/10.1126/science.1193497
11. Guo JY, White E (2016) Autophagy, metabolism, and cancer. Cold Spring Harb Symp Quant Biol 81:73–78. https://doi.org/10.1101/sqb.2016.81.030981
12. Kabeya Y, Mizushima N, Ueno T, Yamamoto A, Kirisako T, Noda T, Kominami E, Ohsumi Y, Yoshimori T (2000) LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. EMBO J 19(21):5720–5728. https://doi.org/10.1093/emboj/19.21.5720
13. Deretic V, Levine B (2009) Autophagy, immunity, and microbial adaptations. Cell Host Microbe 5(6):527–549. https://doi.org/10.1016/j.chom.2009.05.016
14. Cecconi F, Levine B (2008) The role of autophagy in mammalian development: cell makeover rather than cell death. Dev Cell 15 (3):344–357. https://doi.org/10.1016/j.devcel.2008.08.012
15. Levine B, Kroemer G (2008) Autophagy in the pathogenesis of disease. Cell 132(1):27–42. https://doi.org/10.1016/j.cell.2007.12.018
16. Mizushima N, Levine B, Cuervo AM, Klionsky DJ (2008) Autophagy fights disease through cellular self-digestion. Nature 451 (7182):1069–1075. https://doi.org/10.1038/nature06639
17. Ogier-Denis E, Codogno P (2003) Autophagy: a barrier or an adaptive response to cancer. Biochim Biophys Acta 1603(2):113–128
18. Gozuacik D, Kimchi A (2004) Autophagy as a cell death and tumor suppressor mechanism. Oncogene 23(16):2891–2906. https://doi.org/10.1038/sj.onc.1207521
19. Kondo Y, Kanzawa T, Sawaya R, Kondo S (2005) The role of autophagy in cancer development and response to therapy. Nat Rev Cancer 5(9):726–734. https://doi.org/10.1038/nrc1692
20. Lebovitz CB, Robertson AG, Goya R, Jones SJ, Morin RD, Marra MA, Gorski SM (2015) Cross-cancer profiling of molecular alterations within the human autophagy interaction network. Autophagy 11(9):1668–1687. https://doi.org/10.1080/15548627.2015.1067362
21. Liang XH, Jackson S, Seaman M, Brown K, Kempkes B, Hibshoosh H, Levine B (1999) Induction of autophagy and inhibition of tumorigenesis by beclin 1. Nature 402 (6762):672–676. https://doi.org/10.1038/45257
22. Qu X, Yu J, Bhagat G, Furuya N, Hibshoosh H, Troxel A, Rosen J, Eskelinen EL, Mizushima N, Ohsumi Y, Cattoretti G, Levine B (2003) Promotion of tumorigenesis by heterozygous disruption of the beclin 1 autophagy gene. J Clin Invest 112 (12):1809–1820. https://doi.org/10.1172/JCI20039
23. Yue Z, Jin S, Yang C, Levine AJ, Heintz N (2003) Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor. Proc Natl Acad Sci U S A 100(25):15077–15082. https://doi.org/10.1073/pnas.2436255100
24. Laddha SV, Ganesan S, Chan CS, White E (2014) Mutational landscape of the essential autophagy gene BECN1 in human cancers. Mol Cancer Res 12(4):485–490. https://doi.org/10.1158/1541-7786.MCR-13-0614
25. White E, Mehnert JM, Chan CS (2015) Autophagy, metabolism, and cancer. Clin Cancer Res 21(22):5037–5046. https://doi.org/10.1158/1078-0432.CCR-15-0490
26. Meijer AJ, Codogno P (2004) Regulation and role of autophagy in mammalian cells. Int J Biochem Cell Biol 36(12):2445–2462. https://doi.org/10.1016/j.biocel.2004.02.002
27. Burman C, Ktistakis NT (2010) Regulation of autophagy by phosphatidylinositol 3-phosphate. FEBS Lett 584(7):1302–1312.

https://doi.org/10.1016/j.febslet.2010.01.011

28. Shintani T, Klionsky DJ (2004) Autophagy in health and disease: a double-edged sword. Science 306(5698):990–995. https://doi.org/10.1126/science.1099993
29. Furuta S, Hidaka E, Ogata A, Yokota S, Kamata T (2004) Ras is involved in the negative control of autophagy through the class I PI3-kinase. Oncogene 23(22):3898–3904. https://doi.org/10.1038/sj.onc.1207539
30. Arico S, Petiot A, Bauvy C, Dubbelhuis PF, Meijer AJ, Codogno P, Ogier-Denis E (2001) The tumor suppressor PTEN positively regulates macroautophagy by inhibiting the phosphatidylinositol 3-kinase/protein kinase B pathway. J Biol Chem 276(38):35243–35246. https://doi.org/10.1074/jbc.C100319200
31. Steck PA, Pershouse MA, Jasser SA, Yung WK, Lin H, Ligon AH, Langford LA, Baumgard ML, Hattier T, Davis T, Frye C, Hu R, Swedlund B, Teng DH, Tavtigian SV (1997) Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. Nat Genet 15(4):356–362. https://doi.org/10.1038/ng0497-356
32. Chan S (2004) Targeting the mammalian target of rapamycin (mTOR): a new approach to treating cancer. Br J Cancer 91(8):1420–1424. https://doi.org/10.1038/sj.bjc.6602162
33. Klionsky DJ, Abdalla FC, Abeliovich H, Abraham RT, Acevedo-Arozena A, Adeli K, Agholme L, Agnello M, Agostinis P, Aguirre-Ghiso JA, Ahn HJ, Ait-Mohamed O, Ait-Si-Ali S, Akematsu T, Akira S, Al-Younes HM, Al-Zeer MA, Albert ML, Albin RL, Alegre-Abarrategui J, Aleo MF, Alirezaei M, Almasan A, Almonte-Becerril M, Amano A, Amaravadi R, Amarnath S, Amer AO, Andrieu-Abadie N, Anantharam V, Ann DK, Anoopkumar-Dukie S, Aoki H, Apostolova N, Arancia G, Aris JP, Asanuma K, Asare NY, Ashida H, Askanas V, Askew DS, Auberger P, Baba M, Backues SK, Baehrecke EH, Bahr BA, Bai XY, Bailly Y, Baiocchi R, Baldini G, Balduini W, Ballabio A, Bamber BA, Bampton ET, Banhegyi G, Bartholomew CR, Bassham DC, Bast RC Jr, Batoko H, Bay BH, Beau I, Bechet DM, Begley TJ, Behl C, Behrends C, Bekri S, Bellaire B, Bendall LJ, Benetti L, Berliocchi L, Bernardi H, Bernassola F, Besteiro S, Bhatia-Kissova I, Bi X, Biard-Piechaczyk M, Blum JS, Boise LH, Bonaldo P, Boone DL, Bornhauser BC, Bortoluci KR, Bossis I, Bost F, Bourquin JP, Boya P, Boyer-Guittaut M, Bozhkov PV, Brady NR, Brancolini C, Brech A, Brenman JE, Brennand A, Bresnick EH, Brest P, Bridges D, Bristol ML, Brookes PS, Brown EJ, Brumell JH, Brunetti-Pierri N, Brunk UT, Bulman DE, Bultman SJ, Bultynck G, Burbulla LF, Bursch W, Butchar JP, Buzgariu W, Bydlowski SP, Cadwell K, Cahova M, Cai D, Cai J, Cai Q, Calabretta B, Calvo-Garrido J, Camougrand N, Campanella M, Campos-Salinas J, Candi E, Cao L, Caplan AB, Carding SR, Cardoso SM, Carew JS, Carlin CR, Carmignac V, Carneiro LA, Carra S, Caruso RA, Casari G, Casas C, Castino R, Cebollero E, Cecconi F, Celli J, Chaachouay H, Chae HJ, Chai CY, Chan DC, Chan EY, Chang RC, Che CM, Chen CC, Chen GC, Chen GQ, Chen M, Chen Q, Chen SS, Chen W, Chen X, Chen X, Chen X, Chen YG, Chen Y, Chen Y, Chen YJ, Chen Z, Cheng A, Cheng CH, Cheng Y, Cheong H, Cheong JH, Cherry S, Chess-Williams R, Cheung ZH, Chevet E, Chiang HL, Chiarelli R, Chiba T, Chin LS, Chiou SH, Chisari FV, Cho CH, Cho DH, Choi AM, Choi D, Choi KS, Choi ME, Chouaib S, Choubey D, Choubey V, Chu CT, Chuang TH, Chueh SH, Chun T, Chwae YJ, Chye ML, Ciarcia R, Ciriolo MR, Clague MJ, Clark RS, Clarke PG, Clarke R, Codogno P, Coller HA, Colombo MI, Comincini S, Condello M, Condorelli F, Cookson MR, Coombs GH, Coppens I, Corbalan R, Cossart P, Costelli P, Costes S, Coto-Montes A, Couve E, Coxon FP, Cregg JM, Crespo JL, Cronje MJ, Cuervo AM, Cullen JJ, Czaja MJ, D'Amelio M, Darfeuille-Michaud A, Davids LM, Davies FE, De Felici M, de Groot JF, de Haan CA, De Martino L, De Milito A, De Tata V, Debnath J, Degterev A, Dehay B, Delbridge LM, Demarchi F, Deng YZ, Dengjel J, Dent P, Denton D, Deretic V, Desai SD, Devenish RJ, Di Gioacchino M, Di Paolo G, Di Pietro C, Diaz-Araya G, Diaz-Laviada I, Diaz-Meco MT, Diaz-Nido J, Dikic I, Dinesh-Kumar SP, Ding WX, Distelhorst CW, Diwan A, Djavaheri-Mergny M, Dokudovskaya S, Dong Z, Dorsey FC, Dosenko V, Dowling JJ, Doxsey S, Dreux M, Drew ME, Duan Q, Duchosal MA, Duff K, Dugail I, Durbeej M, Duszenko M, Edelstein CL, Edinger AL, Egea G, Eichinger L, Eissa NT, Ekmekcioglu S, El-Deiry WS, Elazar Z, Elgendy M, Ellerby LM, Eng KE, Engelbrecht AM, Engelender S, Erenpreisa J, Escalante R, Esclatine A, Eskelinen EL, Espert L, Espina V, Fan H, Fan J, Fan QW, Fan Z, Fang S, Fang Y, Fanto M, Fanzani A, Farkas T, Farre JC, Faure M, Fechheimer M, Feng CG, Feng J, Feng Q, Feng Y, Fesus L, Feuer R, Figueiredo-Pereira ME, Fimia GM, Fingar DC, Finkbeiner S,

Finkel T, Finley KD, Fiorito F, Fisher EA, Fisher PB, Flajolet M, Florez-McClure ML, Florio S, Fon EA, Fornai F, Fortunato F, Fotedar R, Fowler DH, Fox HS, Franco R, Frankel LB, Fransen M, Fuentes JM, Fueyo J, Fujii J, Fujisaki K, Fujita E, Fukuda M, Furukawa RH, Gaestel M, Gailly P, Gajewska M, Galliot B, Galy V, Ganesh S, Ganetzky B, Ganley IG, Gao FB, Gao GF, Gao J, Garcia L, Garcia-Manero G, Garcia-Marcos M, Garmyn M, Gartel AL, Gatti E, Gautel M, Gawriluk TR, Gegg ME, Geng J, Germain M, Gestwicki JE, Gewirtz DA, Ghavami S, Ghosh P, Giammarioli AM, Giatromanolaki AN, Gibson SB, Gilkerson RW, Ginger ML, Ginsberg HN, Golab J, Goligorsky MS, Golstein P, Gomez-Manzano C, Goncu E, Gongora C, Gonzalez CD, Gonzalez R, Gonzalez-Estevez C, Gonzalez-Polo RA, Gonzalez-Rey E, Gorbunov NV, Gorski S, Goruppi S, Gottlieb RA, Gozuacik D, Granato GE, Grant GD, Green KN, Gregorc A, Gros F, Grose C, Grunt TW, Gual P, Guan JL, Guan KL, Guichard SM, Gukovskaya AS, Gukovsky I, Gunst J, Gustafsson AB, Halayko AJ, Hale AN, Halonen SK, Hamasaki M, Han F, Han T, Hancock MK, Hansen M, Harada H, Harada M, Hardt SE, Harper JW, Harris AL, Harris J, Harris SD, Hashimoto M, Haspel JA, Hayashi S, Hazelhurst LA, He C, He YW, Hebert MJ, Heidenreich KA, Helfrich MH, Helgason GV, Henske EP, Herman B, Herman PK, Hetz C, Hilfiker S, Hill JA, Hocking LJ, Hofman P, Hofmann TG, Hohfeld J, Holyoake TL, Hong MH, Hood DA, Hotamisligil GS, Houwerzijl EJ, Hoyer-Hansen M, Hu B, Hu CA, Hu HM, Hua Y, Huang C, Huang J, Huang S, Huang WP, Huber TB, Huh WK, Hung TH, Hupp TR, Hur GM, Hurley JB, Hussain SN, Hussey PJ, Hwang JJ, Hwang S, Ichihara A, Ilkhanizadeh S, Inoki K, Into T, Iovane V, Iovanna JL, Ip NY, Isaka Y, Ishida H, Isidoro C, Isobe K, Iwasaki A, Izquierdo M, Izumi Y, Jaakkola PM, Jaattela M, Jackson GR, Jackson WT, Janji B, Jendrach M, Jeon JH, Jeung EB, Jiang H, Jiang H, Jiang JX, Jiang M, Jiang Q, Jiang X, Jiang X, Jimenez A, Jin M, Jin S, Joe CO, Johansen T, Johnson DE, Johnson GV, Jones NL, Joseph B, Joseph SK, Joubert AM, Juhasz G, Juillerat-Jeanneret L, Jung CH, Jung YK, Kaarniranta K, Kaasik A, Kabuta T, Kadowaki M, Kagedal K, Kamada Y, Kaminskyy VO, Kampinga HH, Kanamori H, Kang C, Kang KB, Kang KI, Kang R, Kang YA, Kanki T, Kanneganti TD, Kanno H, Kanthasamy AG, Kanthasamy A, Karantza V, Kaushal GP, Kaushik S, Kawazoe Y, Ke PY, Kehrl JH, Kelekar A, Kerkhoff C, Kessel DH, Khalil H, Kiel JA, Kiger AA, Kihara A, Kim DR, Kim DH, Kim DH, Kim EK, Kim HR, Kim JS, Kim JH, Kim JC, Kim JK, Kim PK, Kim SW, Kim YS, Kim Y, Kimchi A, Kimmelman AC, King JS, Kinsella TJ, Kirkin V, Kirshenbaum LA, Kitamoto K, Kitazato K, Klein L, Klimecki WT, Klucken J, Knecht E, Ko BC, Koch JC, Koga H, Koh JY, Koh YH, Koike M, Komatsu M, Kominami E, Kong HJ, Kong WJ, Korolchuk VI, Kotake Y, Koukourakis MI, Kouri Flores JB, Kovacs AL, Kraft C, Krainc D, Kramer H, Kretz-Remy C, Krichevsky AM, Kroemer G, Kruger R, Krut O, Ktistakis NT, Kuan CY, Kucharczyk R, Kumar A, Kumar R, Kumar S, Kundu M, Kung HJ, Kurz T, Kwon HJ, La Spada AR, Lafont F, Lamark T, Landry J, Lane JD, Lapaquette P, Laporte JF, Laszlo L, Lavandero S, Lavoie JN, Layfield R, Lazo PA, Le W, Le Cam L, Ledbetter DJ, Lee AJ, Lee BW, Lee GM, Lee J, Lee JH, Lee M, Lee MS, Lee SH, Leeuwenburgh C, Legembre P, Legouis R, Lehmann M, Lei HY, Lei QY, Leib DA, Leiro J, Lemasters JJ, Lemoine A, Lesniak MS, Lev D, Levenson VV, Levine B, Levy E, Li F, Li JL, Li L, Li S, Li W, Li XJ, Li YB, Li YP, Liang C, Liang Q, Liao YF, Liberski PP, Lieberman A, Lim HJ, Lim KL, Lim K, Lin CF, Lin FC, Lin J, Lin JD, Lin K, Lin WW, Lin WC, Lin YL, Linden R, Lingor P, Lippincott-Schwartz J, Lisanti MP, Liton PB, Liu B, Liu CF, Liu K, Liu L, Liu QA, Liu W, Liu YC, Liu Y, Lockshin RA, Lok CN, Lonial S, Loos B, Lopez-Berestein G, Lopez-Otin C, Lossi L, Lotze MT, Low P, Lu B, Lu B, Lu B, Lu Z, Luciano F, Lukacs NW, Lund AH, Lynch-Day MA, Ma Y, Macian F, MacKeigan JP, Macleod KF, Madeo F, Maiuri L, Maiuri MC, Malagoli D, Malicdan MC, Malorni W, Man N, Mandelkow EM, Manon S, Manov I, Mao K, Mao X, Mao Z, Marambaud P, Marazziti D, Marcel YL, Marchbank K, Marchetti P, Marciniak SJ, Marcondes M, Mardi M, Marfe G, Marino G, Markaki M, Marten MR, Martin SJ, Martinand-Mari C, Martinet W, Martinez-Vicente M, Masini M, Matarrese P, Matsuo S, Matteoni R, Mayer A, Mazure NM, McConkey DJ, McConnell MJ, McDermott C, McDonald C, McInerney GM, McKenna SL, McLaughlin B, McLean PJ, McMaster CR, McQuibban GA, Meijer AJ, Meisler MH, Melendez A, Melia TJ, Melino G, Mena MA, Menendez JA, Menna-Barreto RF, Menon MB, Menzies FM, Mercer CA, Merighi A, Merry DE, Meschini S, Meyer CG, Meyer TF, Miao CY, Miao JY, Michels PA, Michiels C, Mijaljica D, Milojkovic A, Minucci S, Miracco C, Miranti CK, Mitroulis I, Miyazawa K, Mizushima N, Mograbi B,

Mohseni S, Molero X, Mollereau B, Mollinedo F, Momoi T, Monastyrska I, Monick MM, Monteiro MJ, Moore MN, Mora R, Moreau K, Moreira PI, Moriyasu Y, Moscat J, Mostowy S, Mottram JC, Motyl T, Moussa CE, Muller S, Muller S, Munger K, Munz C, Murphy LO, Murphy ME, Musaro A, Mysorekar I, Nagata E, Nagata K, Nahimana A, Nair U, Nakagawa T, Nakahira K, Nakano H, Nakatogawa H, Nanjundan M, Naqvi NI, Narendra DP, Narita M, Navarro M, Nawrocki ST, Nazarko TY, Nemchenko A, Netea MG, Neufeld TP, Ney PA, Nezis IP, Nguyen HP, Nie D, Nishino I, Nislow C, Nixon RA, Noda T, Noegel AA, Nogalska A, Noguchi S, Notterpek L, Novak I, Nozaki T, Nukina N, Nurnberger T, Nyfeler B, Obara K, Oberley TD, Oddo S, Ogawa M, Ohashi T, Okamoto K, Oleinick NL, Oliver FJ, Olsen LJ, Olsson S, Opota O, Osborne TF, Ostrander GK, Otsu K, Ou JH, Ouimet M, Overholtzer M, Ozpolat B, Paganetti P, Pagnini U, Pallet N, Palmer GE, Palumbo C, Pan T, Panaretakis T, Pandey UB, Papackova Z, Papassideri I, Paris I, Park J, Park OK, Parys JB, Parzych KR, Patschan S, Patterson C, Pattingre S, Pawelek JM, Peng J, Perlmutter DH, Perrotta I, Perry G, Pervaiz S, Peter M, Peters GJ, Petersen M, Petrovski G, Phang JM, Piacentini M, Pierre P, Pierrefite-Carle V, Pierron G, Pinkas-Kramarski R, Piras A, Piri N, Platanias LC, Poggeler S, Poirot M, Poletti A, Pous C, Pozuelo-Rubio M, Praetorius-Ibba M, Prasad A, Prescott M, Priault M, Produit-Zengaffinen N, Progulske-Fox A, Proikas-Cezanne T, Przedborski S, Przyklenk K, Puertollano R, Puyal J, Qian SB, Qin L, Qin ZH, Quaggin SE, Raben N, Rabinowich H, Rabkin SW, Rahman I, Rami A, Ramm G, Randall G, Randow F, Rao VA, Rathmell JC, Ravikumar B, Ray SK, Reed BH, Reed JC, Reggiori F, Regnier-Vigouroux-A, Reichert AS, Reiners JJ Jr, Reiter RJ, Ren J, Revuelta JL, Rhodes CJ, Ritis K, Rizzo E, Robbins J, Roberge M, Roca H, Roccheri MC, Rocchi S, Rodemann HP, Rodriguez de Cordoba S, Rohrer B, Roninson IB, Rosen K, Rost-Roszkowska MM, Rouis M, Rouschop KM, Rovetta F, Rubin BP, Rubinsztein DC, Ruckdeschel K, Rucker EB 3rd, Rudich A, Rudolf E, Ruiz-Opazo N, Russo R, Rusten TE, Ryan KM, Ryter SW, Sabatini DM, Sadoshima J, Saha T, Saitoh T, Sakagami H, Sakai Y, Salekdeh GH, Salomoni P, Salvaterra PM, Salvesen G, Salvioli R, Sanchez AM, Sanchez-Alcazar JA, Sanchez-Prieto R, Sandri M, Sankar U, Sansanwal P, Santambrogio L, Saran S, Sarkar S, Sarwal M, Sasakawa C, Sasnauskiene A, Sass M, Sato K, Sato M, Schapira AH, Scharl M, Schatzl HM, Scheper W, Schiaffino S, Schneider C, Schneider ME, Schneider-Stock R, Schoenlein PV, Schorderet DF, Schuller C, Schwartz GK, Scorrano L, Sealy L, Seglen PO, Segura-Aguilar J, Seiliez I, Seleverstov O, Sell C, Seo JB, Separovic D, Setaluri V, Setoguchi T, Settembre C, Shacka JJ, Shanmugam M, Shapiro IM, Shaulian E, Shaw RJ, Shelhamer JH, Shen HM, Shen WC, Sheng ZH, Shi Y, Shibuya K, Shidoji Y, Shieh JJ, Shih CM, Shimada Y, Shimizu S, Shintani T, Shirihai OS, Shore GC, Sibirny AA, Sidhu SB, Sikorska B, Silva-Zacarin EC, Simmons A, Simon AK, Simon HU, Simone C, Simonsen A, Sinclair DA, Singh R, Sinha D, Sinicrope FA, Sirko A, Siu PM, Sivridis E, Skop V, Skulachev VP, Slack RS, Smaili SS, Smith DR, Soengas MS, Soldati T, Song X, Sood AK, Soong TW, Sotgia F, Spector SA, Spies CD, Springer W, Srinivasula SM, Stefanis L, Steffan JS, Stendel R, Stenmark H, Stephanou A, Stern ST, Sternberg C, Stork B, Stralfors P, Subauste CS, Sui X, Sulzer D, Sun J, Sun SY, Sun ZJ, Sung JJ, Suzuki K, Suzuki T, Swanson MS, Swanton C, Sweeney ST, Sy LK, Szabadkai G, Tabas I, Taegtmeyer H, Tafani M, Takacs-Vellai K, Takano Y, Takegawa K, Takemura G, Takeshita F, Talbot NJ, Tan KS, Tanaka K, Tanaka K, Tang D, Tang D, Tanida I, Tannous BA, Tavernarakis N, Taylor GS, Taylor GA, Taylor JP, Terada LS, Terman A, Tettamanti G, Thevissen K, Thompson CB, Thorburn A, Thumm M, Tian F, Tian Y, Tocchini-Valentini G, Tolkovsky AM, Tomino Y, Tonges L, Tooze SA, Tournier C, Tower J, Towns R, Trajkovic V, Travassos LH, Tsai TF, Tschan MP, Tsubata T, Tsung A, Turk B, Turner LS, Tyagi SC, Uchiyama Y, Ueno T, Umekawa M, Umemiya-Shirafuji R, Unni VK, Vaccaro MI, Valente EM, Van den Berghe G, van der Klei IJ, van Doorn W, van Dyk LF, van Egmond M, van Grunsven LA, Vandenabeele P, Vandenberghe WP, Vanhorebeek I, Vaquero EC, Velasco G, Vellai T, Vicencio JM, Vierstra RD, Vila M, Vindis C, Viola G, Viscomi MT, Voitsekhovskaja OV, von Haefen C, Votruba M, Wada K, Wade-Martins R, Walker CL, Walsh CM, Walter J, Wan XB, Wang A, Wang C, Wang D, Wang F, Wang F, Wang G, Wang H, Wang HG, Wang HD, Wang J, Wang K, Wang M, Wang RC, Wang X, Wang X, Wang YJ, Wang Y, Wang Z, Wang ZC, Wang Z, Wansink DG, Ward DM, Watada H, Waters SL, Webster P, Wei L, Weihl CC, Weiss WA, Welford SM, Wen LP, Whitehouse CA, Whitton JL, Whitworth AJ, Wileman T, Wiley JW,

Wilkinson S, Willbold D, Williams RL, Williamson PR, Wouters BG, Wu C, Wu DC, Wu WK, Wyttenbach A, Xavier RJ, Xi Z, Xia P, Xiao G, Xie Z, Xie Z, Xu DZ, Xu J, Xu L, Xu X, Yamamoto A, Yamamoto A, Yamashina S, Yamashita M, Yan X, Yanagida M, Yang DS, Yang E, Yang JM, Yang SY, Yang W, Yang WY, Yang Z, Yao MC, Yao TP, Yeganeh B, Yen WL, Yin JJ, Yin XM, Yoo OJ, Yoon G, Yoon SY, Yorimitsu T, Yoshikawa Y, Yoshimori T, Yoshimoto K, You HJ, Youle RJ, Younes A, Yu L, Yu L, Yu SW, Yu WH, Yuan ZM, Yue Z, Yun CH, Yuzaki M, Zabirnyk O, Silva-Zacarin E, Zacks D, Zacksenhaus E, Zaffaroni N, Zakeri Z, Zeh HJ 3rd, Zeitlin SO, Zhang H, Zhang HL, Zhang J, Zhang JP, Zhang L, Zhang L, Zhang MY, Zhang XD, Zhao M, Zhao YF, Zhao Y, Zhao ZJ, Zheng X, Zhivotovsky B, Zhong Q, Zhou CZ, Zhu C, Zhu WG, Zhu XF, Zhu X, Zhu Y, Zoladek T, Zong WX, Zorzano A, Zschocke J, Zuckerbraun B (2012) Guidelines for the use and interpretation of assays for monitoring autophagy. Autophagy 8(4):445–544

34. Klionsky DJ, Abeliovich H, Agostinis P, Agrawal DK, Aliev G, Askew DS, Baba M, Baehrecke EH, Bahr BA, Ballabio A, Bamber BA, Bassham DC, Bergamini E, Bi X, Biard-Piechaczyk M, Blum JS, Bredesen DE, Brodsky JL, Brumell JH, Brunk UT, Bursch W, Camougrand N, Cebollero E, Cecconi F, Chen Y, Chin LS, Choi A, Chu CT, Chung J, Clarke PG, Clark RS, Clarke SG, Clave C, Cleveland JL, Codogno P, Colombo MI, Coto-Montes A, Cregg JM, Cuervo AM, Debnath J, Demarchi F, Dennis PB, Dennis PA, Deretic V, Devenish RJ, Di Sano F, Dice JF, Difiglia M, Dinesh-Kumar S, Distelhorst CW, Djavaheri-Mergny M, Dorsey FC, Droge W, Dron M, Dunn WA Jr, Duszenko M, Eissa NT, Elazar Z, Esclatine A, Eskelinen EL, Fesus L, Finley KD, Fuentes JM, Fueyo J, Fujisaki K, Galliot B, Gao FB, Gewirtz DA, Gibson SB, Gohla A, Goldberg AL, Gonzalez R, Gonzalez-Estevez C, Gorski S, Gottlieb RA, Haussinger D, He YW, Heidenreich K, Hill JA, Hoyer-Hansen M, Hu X, Huang WP, Iwasaki A, Jaattela M, Jackson WT, Jiang X, Jin S, Johansen T, Jung JU, Kadowaki M, Kang C, Kelekar A, Kessel DH, Kiel JA, Kim HP, Kimchi A, Kinsella TJ, Kiselyov K, Kitamoto K, Knecht E, Komatsu M, Kominami E, Kondo S, Kovacs AL, Kroemer G, Kuan CY, Kumar R, Kundu M, Landry J, Laporte M, Le W, Lei HY, Lenardo MJ, Levine B, Lieberman A, Lim KL, Lin FC, Liou W, Liu LF, Lopez-Berestein G, Lopez-Otin C, Lu B, Macleod KF, Malorni W, Martinet W, Matsuoka K, Mautner J, Meijer AJ, Melendez A, Michels P, Miotto G, Mistiaen WP, Mizushima N, Mograbi B, Monastyrska I, Moore MN, Moreira PI, Moriyasu Y, Motyl T, Munz C, Murphy LO, Naqvi NI, Neufeld TP, Nishino I, Nixon RA, Noda T, Nurnberg B, Ogawa M, Oleinick NL, Olsen LJ, Ozpolat B, Paglin S, Palmer GE, Papassideri I, Parkes M, Perlmutter DH, Perry G, Piacentini M, Pinkas-Kramarski R, Prescott M, Proikas-Cezanne T, Raben N, Rami A, Reggiori F, Rohrer B, Rubinsztein DC, Ryan KM, Sadoshima J, Sakagami H, Sakai Y, Sandri M, Sasakawa C, Sass M, Schneider C, Seglen PO, Seleverstov O, Settleman J, Shacka JJ, Shapiro IM, Sibirny A, Silva-Zacarin EC, Simon HU, Simone C, Simonsen A, Smith MA, Spanel-Borowski K, Srinivas V, Steeves M, Stenmark H, Stromhaug PE, Subauste CS, Sugimoto S, Sulzer D, Suzuki T, Swanson MS, Tabas I, Takeshita F, Talbot NJ, Talloczy Z, Tanaka K, Tanaka K, Tanida I, Taylor GS, Taylor JP, Terman A, Tettamanti G, Thompson CB, Thumm M, Tolkovsky AM, Tooze SA, Truant R, Tumanovska LV, Uchiyama Y, Ueno T, Uzcategui NL, van der Klei I, Vaquero EC, Vellai T, Vogel MW, Wang HG, Webster P, Wiley JW, Xi Z, Xiao G, Yahalom J, Yang JM, Yap G, Yin XM, Yoshimori T, Yu L, Yue Z, Yuzaki M, Zabirnyk O, Zheng X, Zhu X, Deter RL (2008) Guidelines for the use and interpretation of assays for monitoring autophagy in higher eukaryotes. Autophagy 4(2):151–175

35. Yoshii SR, Mizushima N (2017) Monitoring and measuring autophagy. Int J Mol Sci 18 (9). pii: E1865). https://doi.org/10.3390/ijms18091865

36. Morgan MJ, Thorburn A (2016) Measuring autophagy in the context of cancer. Adv Exp Med Biol 899:121–143. https://doi.org/10.1007/978-3-319-26666-4_8

37. Eskelinen EL (2005) Maturation of autophagic vacuoles in Mammalian cells. Autophagy 1 (1):1–10

38. Yla-Anttila P, Vihinen H, Jokitalo E, Eskelinen EL (2009) Monitoring autophagy by electron microscopy in Mammalian cells. Methods Enzymol 452:143–164. https://doi.org/10.1016/S0076-6879(08)03610-0

39. Kabeya Y, Mizushima N, Yamamoto A, Oshitani-Okamoto S, Ohsumi Y, Yoshimori T (2004) LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. J Cell Sci 117 (Pt 13):2805–2812. https://doi.org/10.1242/jcs.01131

40. Klionsky DJ, Abdelmohsen K, Abe A, Abedin MJ, Abeliovich H, Acevedo Arozena A, Adachi H, Adams CM, Adams PD, Adeli K, Adhihetty PJ, Adler SG, Agam G, Agarwal R, Aghi MK, Agnello M, Agostinis P, Aguilar PV, Aguirre-Ghiso J, Airoldi EM, Ait-Si-Ali S, Akematsu T, Akporiaye ET, Al-Rubeai M, Albaiceta GM, Albanese C, Albani D, Albert ML, Aldudo J, Algul H, Alirezaei M, Alloza I, Almasan A, Almonte-Beceril M, Alnemri ES, Alonso C, Altan-Bonnet N, Altieri DC, Alvarez S, Alvarez-Erviti L, Alves S, Amadoro G, Amano A, Amantini C, Ambrosio S, Amelio I, Amer AO, Amessou M, Amon A, An Z, Anania FA, Andersen SU, Andley UP, Andreadi CK, Andrieu-Abadie N, Anel A, Ann DK, Anoopkumar-Dukie S, Antonioli M, Aoki H, Apostolova N, Aquila S, Aquilano K, Araki K, Arama E, Aranda A, Araya J, Arcaro A, Arias E, Arimoto H, Ariosa AR, Armstrong JL, Arnould T, Arsov I, Asanuma K, Askanas V, Asselin E, Atarashi R, Atherton SS, Atkin JD, Attardi LD, Auberger P, Auburger G, Aurelian L, Autelli R, Avagliano L, Avantaggiati ML, Avrahami L, Awale S, Azad N, Bachetti T, Backer JM, Bae DH, Bae JS, Bae ON, Bae SH, Baehrecke EH, Baek SH, Baghdiguian S, Bagniewska-Zadworna A, Bai H, Bai J, Bai XY, Bailly Y, Balaji KN, Balduini W, Ballabio A, Balzan R, Banerjee R, Banhegyi G, Bao H, Barbeau B, Barrachina MD, Barreiro E, Bartel B, Bartolome A, Bassham DC, Bassi MT, Bast RC Jr, Basu A, Batista MT, Batoko H, Battino M, Bauckman K, Baumgarner BL, Bayer KU, Beale R, Beaulieu JF, Beck GR Jr, Becker C, Beckham JD, Bedard PA, Bednarski PJ, Begley TJ, Behl C, Behrends C, Behrens GM, Behrns KE, Bejarano E, Belaid A, Belleudi F, Benard G, Berchem G, Bergamaschi D, Bergami M, Berkhout B, Berliocchi L, Bernard A, Bernard M, Bernassola F, Bertolotti A, Bess AS, Besteiro S, Bettuzzi S, Bhalla S, Bhattacharyya S, Bhutia SK, Biagosch C, Bianchi MW, Biard-Piechaczyk M, Billes V, Bincoletto C, Bingol B, Bird SW, Bitoun M, Bjedov I, Blackstone C, Blanc L, Blanco GA, Blomhoff HK, Boada-Romero E, Bockler S, Boes M, Boesze-Battaglia K, Boise LH, Bolino A, Boman A, Bonaldo P, Bordi M, Bosch J, Botana LM, Botti J, Bou G, Bouche M, Bouchecareilh M, Boucher MJ, Boulton ME, Bouret SG, Boya P, Boyer-Guittaut M, Bozhkov PV, Brady N, Braga VM, Brancolini C, Braus GH, Bravo-San Pedro JM, Brennan LA, Bresnick EH, Brest P, Bridges D, Bringer MA, Brini M, Brito GC, Brodin B, Brookes PS, Brown EJ, Brown K, Broxmeyer HE, Bruhat A, Brum PC, Brumell JH, Brunetti-Pierri N, Bryson-Richardson RJ, Buch S, Buchan AM, Budak H, Bulavin DV, Bultman SJ, Bultynck G, Bumbasirevic V, Burelle Y, Burke RE, Burmeister M, Butikofer P, Caberlotto L, Cadwell K, Cahova M, Cai D, Cai J, Cai Q, Calatayud S, Camougrand N, Campanella M, Campbell GR, Campbell M, Campello S, Candau R, Caniggia I, Cantoni L, Cao L, Caplan AB, Caraglia M, Cardinali C, Cardoso SM, Carew JS, Carleton LA, Carlin CR, Carloni S, Carlsson SR, Carmona-Gutierrez D, Carneiro LA, Carnevali O, Carra S, Carrier A, Carroll B, Casas C, Casas J, Cassinelli G, Castets P, Castro-Obregon S, Cavallini G, Ceccherini I, Cecconi F, Cederbaum AI, Cena V, Cenci S, Cerella C, Cervia D, Cetrullo S, Chaachouay H, Chae HJ, Chagin AS, Chai CY, Chakrabarti G, Chamilos G, Chan EY, Chan MT, Chandra D, Chandra P, Chang CP, Chang RC, Chang TY, Chatham JC, Chatterjee S, Chauhan S, Che Y, Cheetham ME, Cheluvappa R, Chen CJ, Chen G, Chen GC, Chen G, Chen H, Chen JW, Chen JK, Chen M, Chen M, Chen P, Chen Q, Chen Q, Chen SD, Chen S, Chen SS, Chen W, Chen WJ, Chen WQ, Chen W, Chen X, Chen YH, Chen YG, Chen Y, Chen Y, Chen Y, Chen YJ, Chen YQ, Chen Y, Chen Z, Chen Z, Cheng A, Cheng CH, Cheng H, Cheong H, Cherry S, Chesney J, Cheung CH, Chevet E, Chi HC, Chi SG, Chiacchiera F, Chiang HL, Chiarelli R, Chiariello M, Chieppa M, Chin LS, Chiong M, Chiu GN, Cho DH, Cho SG, Cho WC, Cho YY, Cho YS, Choi AM, Choi EJ, Choi EK, Choi J, Choi ME, Choi SI, Chou TF, Chouaib S, Choubey D, Choubey V, Chow KC, Chowdhury K, Chu CT, Chuang TH, Chun T, Chung H, Chung T, Chung YL, Chwae YJ, Cianfanelli V, Ciarcia R, Ciechomska IA, Ciriolo MR, Cirone M, Claerhout S, Clague MJ, Claria J, Clarke PG, Clarke R, Clementi E, Cleyrat C, Cnop M, Coccia EM, Cocco T, Codogno P, Coers J, Cohen EE, Colecchia D, Coletto L, Coll NS, Colucci-Guyon E, Comincini S, Condello M, Cook KL, Coombs GH, Cooper CD, Cooper JM, Coppens I, Corasaniti MT, Corazzari M, Corbalan R, Corcelle-Termeau E, Cordero MD, Corral-Ramos C, Corti O, Cossarizza A, Costelli P, Costes S, Cotman SL, Coto-Montes A, Cottet S, Couve E, Covey LR, Cowart LA, Cox JS, Coxon FP, Coyne CB, Cragg MS, Craven RJ, Crepaldi T, Crespo JL, Criollo A, Crippa V, Cruz MT, Cuervo AM, Cuezva JM, Cui T, Cutillas PR, Czaja MJ, Czyzyk-Krzeska MF, Dagda RK, Dahmen U, Dai C, Dai W, Dai Y, Dalby KN, Dalla Valle L,

Dalmasso G, D'Amelio M, Damme M, Darfeuille-Michaud A, Dargemont C, Darley-Usmar VM, Dasarathy S, Dasgupta B, Dash S, Dass CR, Davey HM, Davids LM, Davila D, Davis RJ, Dawson TM, Dawson VL, Daza P, de Belleroche J, de Figueiredo P, de Figueiredo RC, de la Fuente J, De Martino L, De Matteis A, De Meyer GR, De Milito A, De Santi M, de Souza W, De Tata V, De Zio D, Debnath J, Dechant R, Decuypere JP, Deegan S, Dehay B, Del Bello B, Del Re DP, Delage-Mourroux R, Delbridge LM, Deldicque L, Delorme-Axford E, Deng Y, Dengjel J, Denizot M, Dent P, Der CJ, Deretic V, Derrien B, Deutsch E, Devarenne TP, Devenish RJ, Di Bartolomeo S, Di Daniele N, Di Domenico F, Di Nardo A, Di Paola S, Di Pietro A, Di Renzo L, DiAntonio A, Diaz-Araya G, Diaz-Laviada I, Diaz-Meco MT, Diaz-Nido J, Dickey CA, Dickson RC, Diederich M, Digard P, Dikic I, Dinesh-Kumar SP, Ding C, Ding WX, Ding Z, Dini L, Distler JH, Diwan A, Djavaheri-Mergny M, Dmytruk K, Dobson RC, Doetsch V, Dokladny K, Dokudovskaya S, Donadelli M, Dong XC, Dong X, Dong Z, Donohue TM Jr, Doran KS, D'Orazi G, Dorn GW 2nd, Dosenko V, Dridi S, Drucker L, Du J, Du LL, Du L, du Toit A, Dua P, Duan L, Duann P, Dubey VK, Duchen MR, Duchosal MA, Duez H, Dugail I, Dumit VI, Duncan MC, Dunlop EA, Dunn WA Jr, Dupont N, Dupuis L, Duran RV, Durcan TM, Duvezin-Caubet S, Duvvuri U, Eapen V, Ebrahimi-Fakhari D, Echard A, Eckhart L, Edelstein CL, Edinger AL, Eichinger L, Eisenberg T, Eisenberg-Lerner A, Eissa NT, El-Deiry WS, El-Khoury V, Elazar Z, Eldar-Finkelman H, Elliott CJ, Emanuele E, Emmenegger U, Engedal N, Engelbrecht AM, Engelender S, Enserink JM, Erdmann R, Erenpreisa J, Eri R, Eriksen JL, Erman A, Escalante R, Eskelinen EL, Espert L, Esteban-Martinez L, Evans TJ, Fabri M, Fabrias G, Fabrizi C, Facchiano A, Faergeman NJ, Faggioni A, Fairlie WD, Fan C, Fan D, Fan J, Fang S, Fanto M, Fanzani A, Farkas T, Faure M, Favier FB, Fearnhead H, Federici M, Fei E, Felizardo TC, Feng H, Feng Y, Feng Y, Ferguson TA, Fernandez AF, Fernandez-Barrena MG, Fernandez-Checa JC, Fernandez-Lopez A, Fernandez-Zapico ME, Feron O, Ferraro E, Ferreira-Halder CV, Fesus L, Feuer R, Fiesel FC, Filippi-Chiela EC, Filomeni G, Fimia GM, Fingert JH, Finkbeiner S, Finkel T, Fiorito F, Fisher PB, Flajolet M, Flamigni F, Florey O, Florio S, Floto RA, Folini M, Follo C, Fon EA, Fornai F, Fortunato F, Fraldi A, Franco R, Francois A, Francois A, Frankel LB, Fraser ID, Frey N, Freyssenet DG, Frezza C, Friedman SL, Frigo DE, Fu D, Fuentes JM, Fueyo J, Fujitani Y, Fujiwara Y, Fujiya M, Fukuda M, Fulda S, Fusco C, Gabryel B, Gaestel M, Gailly P, Gajewska M, Galadari S, Galili G, Galindo I, Galindo MF, Galliciotti G, Galluzzi L, Galluzzi L, Galy V, Gammoh N, Gandy S, Ganesan AK, Ganesan S, Ganley IG, Gannage M, Gao FB, Gao F, Gao JX, Garcia Nannig L, Garcia Vescovi E, Garcia-Macia M, Garcia-Ruiz C, Garg AD, Garg PK, Gargini R, Gassen NC, Gatica D, Gatti E, Gavard J, Gavathiotis E, Ge L, Ge P, Ge S, Gean PW, Gelmetti V, Genazzani AA, Geng J, Genschik P, Gerner L, Gestwicki JE, Gewirtz DA, Ghavami S, Ghigo E, Ghosh D, Giammarioli AM, Giampieri F, Giampietri C, Giatromanolaki A, Gibbings DJ, Gibellini L, Gibson SB, Ginet V, Giordano A, Giorgini F, Giovannetti E, Girardin SE, Gispert S, Giuliano S, Gladson CL, Glavic A, Gleave M, Godefroy N, Gogal RM Jr, Gokulan K, Goldman GH, Goletti D, Goligorsky MS, Gomes AV, Gomes LC, Gomez H, Gomez-Manzano C, Gomez-Sanchez R, Goncalves DA, Goncu E, Gong Q, Gongora C, Gonzalez CB, Gonzalez-Alegre P, Gonzalez-Cabo P, Gonzalez-Polo RA, Goping IS, Gorbea C, Gorbunov NV, Goring DR, Gorman AM, Gorski SM, Goruppi S, Goto-Yamada S, Gotor C, Gottlieb RA, Gozes I, Gozuacik D, Graba Y, Graef M, Granato GE, Grant GD, Grant S, Gravina GL, Green DR, Greenhough A, Greenwood MT, Grimaldi B, Gros F, Grose C, Groulx JF, Gruber F, Grumati P, Grune T, Guan JL, Guan KL, Guerra B, Guillen C, Gulshan K, Gunst J, Guo C, Guo L, Guo M, Guo W, Guo XG, Gust AA, Gustafsson AB, Gutierrez E, Gutierrez MG, Gwak HS, Haas A, Haber JE, Hadano S, Hagedorn M, Hahn DR, Halayko AJ, Hamacher-Brady A, Hamada K, Hamai A, Hamann A, Hamasaki M, Hamer I, Hamid Q, Hammond EM, Han F, Han W, Handa JT, Hanover JA, Hansen M, Harada M, Harhaji-Trajkovic L, Harper JW, Harrath AH, Harris AL, Harris J, Hasler U, Hasselblatt P, Hasui K, Hawley RG, Hawley TS, He C, He CY, He F, He G, He RR, He XH, He YW, He YY, Heath JK, Hebert MJ, Heinzen RA, Helgason GV, Hensel M, Henske EP, Her C, Herman PK, Hernandez A, Hernandez C, Hernandez-Tiedra S, Hetz C, Hiesinger PR, Higaki K, Hilfiker S, Hill BG, Hill JA, Hill WD, Hino K, Hofius D, Hofman P, Hoglinger GU, Hohfeld J, Holz MK, Hong Y, Hood DA, Hoozemans JJ, Hoppe T, Hsu C, Hsu CY, Hsu LC, Hu D, Hu G, Hu HM, Hu H,

Hu MC, Hu YC, Hu ZW, Hua F, Hua Y, Huang C, Huang HL, Huang KH, Huang KY, Huang S, Huang S, Huang WP, Huang YR, Huang Y, Huang Y, Huber TB, Huebbe P, Huh WK, Hulmi JJ, Hur GM, Hurley JH, Husak Z, Hussain SN, Hussain S, Hwang JJ, Hwang S, Hwang TI, Ichihara A, Imai Y, Imbriano C, Inomata M, Into T, Iovane V, Iovanna JL, Iozzo RV, Ip NY, Irazoqui JE, Iribarren P, Isaka Y, Isakovic AJ, Ischiropoulos H, Isenberg JS, Ishaq M, Ishida H, Ishii I, Ishmael JE, Isidoro C, Isobe K, Isono E, Issazadeh-Navikas S, Itahana K, Itakura E, Ivanov AI, Iyer AK, Izquierdo JM, Izumi Y, Izzo V, Jaattela M, Jaber N, Jackson DJ, Jackson WT, Jacob TG, Jacques TS, Jagannath C, Jain A, Jana NR, Jang BK, Jani A, Janji B, Jannig PR, Jansson PJ, Jean S, Jendrach M, Jeon JH, Jessen N, Jeung EB, Jia K, Jia L, Jiang H, Jiang H, Jiang L, Jiang T, Jiang X, Jiang X, Jiang X, Jiang Y, Jiang Y, Jimenez A, Jin C, Jin H, Jin L, Jin M, Jin S, Jinwal UK, Jo EK, Johansen T, Johnson DE, Johnson GV, Johnson JD, Jonasch E, Jones C, Joosten LA, Jordan J, Joseph AM, Joseph B, Joubert AM, Ju D, Ju J, Juan HF, Juenemann K, Juhasz G, Jung HS, Jung JU, Jung YK, Jungbluth H, Justice MJ, Jutten B, Kaakoush NO, Kaarniranta K, Kaasik A, Kabuta T, Kaeffer B, Kagedal K, Kahana A, Kajimura S, Kakhlon O, Kalia M, Kalvakolanu DV, Kamada Y, Kambas K, Kaminskyy VO, Kampinga HH, Kandouz M, Kang C, Kang R, Kang TC, Kanki T, Kanneganti TD, Kanno H, Kanthasamy AG, Kantorow M, Kaparakis-Liaskos M, Kapuy O, Karantza V, Karim MR, Karmakar P, Kaser A, Kaushik S, Kawula T, Kaynar AM, Ke PY, Ke ZJ, Kehrl JH, Keller KE, Kemper JK, Kenworthy AK, Kepp O, Kern A, Kesari S, Kessel D, Ketteler R, Kettelhut Ido C, Khambu B, Khan MM, Khandelwal VK, Khare S, Kiang JG, Kiger AA, Kihara A, Kim AL, Kim CH, Kim DR, Kim DH, Kim EK, Kim HY, Kim HR, Kim JS, Kim JH, Kim JC, Kim JH, Kim KW, Kim MD, Kim MM, Kim PK, Kim SW, Kim SY, Kim YS, Kim Y, Kimchi A, Kimmelman AC, Kimura T, King JS, Kirkegaard K, Kirkin V, Kirshenbaum LA, Kishi S, Kitajima Y, Kitamoto K, Kitaoka Y, Kitazato K, Kley RA, Klimecki WT, Klinkenberg M, Klucken J, Knaevelsrud H, Knecht E, Knuppertz L, Ko JL, Kobayashi S, Koch JC, Koechlin-Ramonatxo C, Koenig U, Koh YH, Kohler K, Kohlwein SD, Koike M, Komatsu M, Kominami E, Kong D, Kong HJ, Konstantakou EG, Kopp BT, Korcsmaros T, Korhonen L, Korolchuk VI, Koshkina NV, Kou Y, Koukourakis MI, Koumenis C, Kovacs AL, Kovacs T, Kovacs WJ, Koya D, Kraft C, Krainc D, Kramer H, Kravic-Stevovic T, Krek W, Kretz-Remy C, Krick R, Krishnamurthy M, Kriston-Vizi J, Kroemer G, Kruer MC, Kruger R, Ktistakis NT, Kuchitsu K, Kuhn C, Kumar AP, Kumar A, Kumar A, Kumar D, Kumar D, Kumar R, Kumar S, Kundu M, Kung HJ, Kuno A, Kuo SH, Kuret J, Kurz T, Kwok T, Kwon TK, Kwon YT, Kyrmizi I, La Spada AR, Lafont F, Lahm T, Lakkaraju A, Lam T, Lamark T, Lancel S, Landowski TH, Lane DJ, Lane JD, Lanzi C, Lapaquette P, Lapierre LR, Laporte J, Laukkarinen J, Laurie GW, Lavandero S, Lavie L, LaVoie MJ, Law BY, Law HK, Law KB, Layfield R, Lazo PA, Le Cam L, Le Roch KG, Le Stunff H, Leardkamolkarn V, Lecuit M, Lee BH, Lee CH, Lee EF, Lee GM, Lee HJ, Lee H, Lee JK, Lee J, Lee JH, Lee JH, Lee M, Lee MS, Lee PJ, Lee SW, Lee SJ, Lee SJ, Lee SY, Lee SH, Lee SS, Lee SJ, Lee S, Lee YR, Lee YJ, Lee YH, Leeuwenburgh C, Lefort S, Legouis R, Lei J, Lei QY, Leib DA, Leibowitz G, Lekli I, Lemaire SD, Lemasters JJ, Lemberg MK, Lemoine A, Leng S, Lenz G, Lenzi P, Lerman LO, Lettieri Barbato D, Leu JI, Leung HY, Levine B, Lewis PA, Lezoualc'h F, Li C, Li F, Li FJ, Li J, Li K, Li L, Li M, Li M, Li Q, Li R, Li S, Li W, Li W, Li X, Li Y, Lian J, Liang C, Liang Q, Liao Y, Liberal J, Liberski PP, Lie P, Lieberman AP, Lim HJ, Lim KL, Lim K, Lima RT, Lin CS, Lin CF, Lin F, Lin F, Lin FC, Lin K, Lin KH, Lin PH, Lin T, Lin WW, Lin YS, Lin Y, Linden R, Lindholm D, Lindqvist LM, Lingor P, Linkermann A, Liotta LA, Lipinski MM, Lira VA, Lisanti MP, Liton PB, Liu B, Liu C, Liu CF, Liu F, Liu HJ, Liu J, Liu JJ, Liu JL, Liu K, Liu L, Liu L, Liu Q, Liu RY, Liu S, Liu S, Liu W, Liu XD, Liu X, Liu XH, Liu X, Liu X, Liu X, Liu Y, Liu Y, Liu Z, Liu Z, Liuzzi JP, Lizard G, Ljujic M, Lodhi IJ, Logue SE, Lokeshwar BL, Long YC, Lonial S, Loos B, Lopez-Otin C, Lopez-Vicario C, Lorente M, Lorenzi PL, Lorincz P, Los M, Lotze MT, Lovat PE, Lu B, Lu B, Lu J, Lu Q, Lu SM, Lu S, Lu Y, Luciano F, Luckhart S, Lucocq JM, Ludovico P, Lugea A, Lukacs NW, Lum JJ, Lund AH, Luo H, Luo J, Luo S, Luparello C, Lyons T, Ma J, Ma Y, Ma Y, Ma Z, Machado J, Machado-Santelli GM, Macian F, MacIntosh GC, MacKeigan JP, Macleod KF, MacMicking JD, MacMillan-Crow LA, Madeo F, Madesh M, Madrigal-Matute J, Maeda A, Maeda T, Maegawa G, Maellaro E, Maes H, Magarinos M, Maiese K, Maiti TK, Maiuri L, Maiuri MC, Maki CG, Malli R, Malorni W, Maloyan A, Mami-Chouaib F, Man N, Mancias JD, Mandelkow EM, Mandell MA, Manfredi

AA, Manie SN, Manzoni C, Mao K, Mao Z, Mao ZW, Marambaud P, Marconi AM, Marelja Z, Marfe G, Margeta M, Margittai E, Mari M, Mariani FV, Marin C, Marinelli S, Marino G, Markovic I, Marquez R, Martelli AM, Martens S, Martin KR, Martin SJ, Martin S, Martin-Acebes MA, Martin-Sanz P, Martinand-Mari C, Martinet W, Martinez J, Martinez-Lopez N, Martinez-Outschoorn U, Martinez-Velazquez M, Martinez-Vicente M, Martins WK, Mashima H, Mastrianni JA, Matarese G, Matarrese P, Mateo R, Matoba S, Matsumoto N, Matsushita T, Matsuura A, Matsuzawa T, Mattson MP, Matus S, Maugeri N, Mauvezin C, Mayer A, Maysinger D, Mazzolini GD, McBrayer MK, McCall K, McCormick C, McInerney GM, McIver SC, McKenna S, McMahon JJ, McNeish IA, Mechta-Grigoriou F, Medema JP, Medina DL, Megyeri K, Mehrpour M, Mehta JL, Mei Y, Meier UC, Meijer AJ, Melendez A, Melino G, Melino S, de Melo EJ, Mena MA, Meneghini MD, Menendez JA, Menezes R, Meng L, Meng LH, Meng S, Menghini R, Menko AS, Menna-Barreto RF, Menon MB, Meraz-Rios MA, Merla G, Merlini L, Merlot AM, Meryk A, Meschini S, Meyer JN, Mi MT, Miao CY, Micale L, Michaeli S, Michiels C, Migliaccio AR, Mihailidou AS, Mijaljica D, Mikoshiba K, Milan E, Miller-Fleming L, Mills GB, Mills IG, Minakaki G, Minassian BA, Ming XF, Minibayeva F, Minina EA, Mintern JD, Minucci S, Miranda-Vizuete A, Mitchell CH, Miyamoto S, Miyazawa K, Mizushima N, Mnich K, Mograbi B, Mohseni S, Moita LF, Molinari M, Molinari M, Moller AB, Mollereau B, Mollinedo F, Mongillo M, Monick MM, Montagnaro S, Montell C, Moore DJ, Moore MN, Mora-Rodriguez R, Moreira PI, Morel E, Morelli MB, Moreno S, Morgan MJ, Moris A, Moriyasu Y, Morrison JL, Morrison LA, Morselli E, Moscat J, Moseley PL, Mostowy S, Motori E, Mottet D, Mottram JC, Moussa CE, Mpakou VE, Mukhtar H, Mulcahy Levy JM, Muller S, Munoz-Moreno-R, Munoz-Pinedo C, Munz C, Murphy ME, Murray JT, Murthy A, Mysorekar IU, Nabi IR, Nabissi M, Nader GA, Nagahara Y, Nagai Y, Nagata K, Nagelkerke A, Nagy P, Naidu SR, Nair S, Nakano H, Nakatogawa H, Nanjundan M, Napolitano G, Naqvi NI, Nardacci R, Narendra DP, Narita M, Nascimbeni AC, Natarajan R, Navegantes LC, Nawrocki ST, Nazarko TY, Nazarko VY, Neill T, Neri LM, Netea MG, Netea-Maier RT, Neves BM, Ney PA, Nezis IP, Nguyen HT, Nguyen HP, Nicot AS, Nilsen H, Nilsson P, Nishimura M, Nishino I, Nisoo-Santano M, Niu H, Nixon RA, Njar VC, Noda T, Noegel AA, Nolte EM, Norberg E, Norga KK, Noureini SK, Notomi S, Notterpek L, Nowikovsky K, Nukina N, Nurnberger T, O'Donnell VB, O'Donovan T, O'Dwyer PJ, Oehme I, Oeste CL, Ogawa M, Ogretmen B, Ogura Y, Oh YJ, Ohmuraya M, Ohshima T, Ojha R, Okamoto K, Okazaki T, Oliver FJ, Ollinger K, Olsson S, Orban DP, Ordonez P, Orhon I, Orosz L, O'Rourke EJ, Orozco H, Ortega AL, Ortona E, Osellame LD, Oshima J, Oshima S, Osiewacz HD, Otomo T, Otsu K, Ou JH, Outeiro TF, Ouyang DY, Ouyang H, Overholtzer M, Ozbun MA, Ozdinler PH, Ozpolat B, Pacelli C, Paganetti P, Page G, Pages G, Pagnini U, Pajak B, Pak SC, Pakos-Zebrucka K, Pakpour N, Palkova Z, Palladino F, Pallauf K, Pallet N, Palmieri M, Paludan SR, Palumbo C, Palumbo S, Pampliega O, Pan H, Pan W, Panaretakis T, Pandey A, Pantazopoulou A, Papackova Z, Papademetrio DL, Papassideri I, Papini A, Parajuli N, Pardo J, Parekh VV, Parenti G, Park JI, Park J, Park OK, Parker R, Parlato R, Parys JB, Parzych KR, Pasquet JM, Pasquier B, Pasumarthi KB, Patschan D, Patterson C, Pattingre S, Pattison S, Pause A, Pavenstadt H, Pavone F, Pedrozo Z, Pena FJ, Penalva MA, Pende M, Peng J, Penna F, Penninger JM, Pensalfini A, Pepe S, Pereira GJ, Pereira PC, Perez-de la Cruz V, Perez-Perez ME, Perez-Rodriguez D, Perez-Sala D, Perier C, Perl A, Perlmutter DH, Perrotta I, Pervaiz S, Pesonen M, Pessin JE, Peters GJ, Petersen M, Petrache I, Petrof BJ, Petrovski G, Phang JM, Piacentini M, Pierdominici M, Pierre P, Pierrefite-Carle V, Pietrocola F, Pimentel-Muinos FX, Pinar M, Pineda B, Pinkas-Kramarski R, Pinti M, Pinton P, Piperdi B, Piret JM, Platanias LC, Platta HW, Plowey ED, Poggeler S, Poirot M, Polcic P, Poletti A, Poon AH, Popelka H, Popova B, Poprawa I, Poulose SM, Poulton J, Powers SK, Powers T, Pozuelo-Rubio M, Prak K, Prange R, Prescott M, Priault M, Prince S, Proia RL, Proikas-Cezanne T, Prokisch H, Promponas VJ, Przyklenk K, Puertollano R, Pugazhenthi S, Puglielli L, Pujol A, Puyal J, Pyeon D, Qi X, Qian WB, Qin ZH, Qiu Y, Qu Z, Quadrilatero J, Quinn F, Raben N, Rabinowich H, Radogna F, Ragusa MJ, Rahmani M, Raina K, Ramanadham S, Ramesh R, Rami A, Randall-Demllo S, Randow F, Rao H, Rao VA, Rasmussen BB, Rasse TM, Ratovitski EA, Rautou PE, Ray SK, Razani B, Reed BH, Reggiori F, Rehm M, Reichert AS, Rein T, Reiner DJ, Reits E, Ren J, Ren X, Renna M, Reusch JE, Revuelta JL, Reyes L, Rezaie AR, Richards RI,

Richardson DR, Richetta C, Riehle MA, Rihn BH, Rikihisa Y, Riley BE, Rimbach G, Rippo MR, Ritis K, Rizzi F, Rizzo E, Roach PJ, Robbins J, Roberge M, Roca G, Roccheri MC, Rocha S, Rodrigues CM, Rodriguez CI, de Cordoba SR, Rodriguez-Muela N, Roelofs J, Rogov VV, Rohn TT, Rohrer B, Romanelli D, Romani L, Romano PS, Roncero MI, Rosa JL, Rosello A, Rosen KV, Rosenstiel P, Rost-Roszkowska M, Roth KA, Roue G, Rouis M, Rouschop KM, Ruan DT, Ruano D, Rubinsztein DC, Rucker EB 3rd, Rudich A, Rudolf E, Rudolf R, Ruegg MA, Ruiz-Roldan C, Ruparelia AA, Rusmini P, Russ DW, Russo GL, Russo G, Russo R, Rusten TE, Ryabovol V, Ryan KM, Ryter SW, Sabatini DM, Sacher M, Sachse C, Sack MN, Sadoshima J, Saftig P, Sagi-Eisenberg R, Sahni S, Saikumar P, Saito T, Saitoh T, Sakakura K, Sakoh-Nakatogawa M, Sakuraba Y, Salazar-Roa M, Salomoni P, Saluja AK, Salvaterra PM, Salvioli R, Samali A, Sanchez AM, Sanchez-Alcazar JA, Sanchez-Prieto-R, Sandri M, Sanjuan MA, Santaguida S, Santambrogio L, Santoni G, Dos Santos CN, Saran S, Sardiello M, Sargent G, Sarkar P, Sarkar S, Sarrias MR, Sarwal MM, Sasakawa C, Sasaki M, Sass M, Sato K, Sato M, Satriano J, Savaraj N, Saveljeva S, Schaefer L, Schaible UE, Scharl M, Schatzl HM, Schekman R, Scheper W, Schiavi A, Schipper HM, Schmeisser H, Schmidt J, Schmitz I, Schneider BE, Schneider EM, Schneider JL, Schon EA, Schonenberger MJ, Schonthal AH, Schorderet DF, Schroder B, Schuck S, Schulze RJ, Schwarten M, Schwarz TL, Sciarretta S, Scotto K, Scovassi AI, Screaton RA, Screen M, Seca H, Sedej S, Segatori L, Segev N, Seglen PO, Segui-Simarro JM, Segura-Aguilar J, Seki E, Sell C, Seiliez I, Semenkovich CF, Semenza GL, Sen U, Serra AL, Serrano-Puebla A, Sesaki H, Setoguchi T, Settembre C, Shacka JJ, Shajahan-Haq AN, Shapiro IM, Sharma S, She H, Shen CK, Shen CC, Shen HM, Shen S, Shen W, Sheng R, Sheng X, Sheng ZH, Shepherd TG, Shi J, Shi Q, Shi Q, Shi Y, Shibutani S, Shibuya K, Shidoji Y, Shieh JJ, Shih CM, Shimada Y, Shimizu S, Shin DW, Shinohara ML, Shintani M, Shintani T, Shioi T, Shirabe K, Shiri-Sverdlov R, Shirihai O, Shore GC, Shu CW, Shukla D, Sibirny AA, Sica V, Sigurdson CJ, Sigurdsson EM, Sijwali PS, Sikorska B, Silveira WA, Silvente-Poirot S, Silverman GA, Simak J, Simmet T, Simon AK, Simon HU, Simone C, Simons M, Simonsen A, Singh R, Singh SV, Singh SK, Sinha D, Sinha S, Sinicrope FA, Sirko A, Sirohi K, Sishi BJ, Sittler A, Siu PM, Sivridis E, Skwarska A, Slack R, Slaninova I, Slavov N, Smaili SS, Smalley KS, Smith DR, Soenen SJ, Soleimanpour SA, Solhaug A, Somasundaram K, Son JH, Sonawane A, Song C, Song F, Song HK, Song JX, Song W, Soo KY, Sood AK, Soong TW, Soontornniyomkij V, Sorice M, Sotgia F, Soto-Pantoja DR, Sotthibundhu A, Sousa MJ, Spaink HP, Span PN, Spang A, Sparks JD, Speck PG, Spector SA, Spies CD, Springer W, Clair DS, Stacchiotti A, Staels B, Stang MT, Starczynowski DT, Starokadomskyy P, Steegborn C, Steele JW, Stefanis L, Steffan J, Stellrecht CM, Stenmark H, Stepkowski TM, Stern ST, Stevens C, Stockwell BR, Stoka V, Storchova Z, Stork B, Stratoulias V, Stravopodis DJ, Strnad P, Strohecker AM, Strom AL, Stromhaug P, Stulik J, Su YX, Su Z, Subauste CS, Subramaniam S, Sue CM, Suh SW, Sui X, Sukseree S, Sulzer D, Sun FL, Sun J, Sun J, Sun SY, Sun Y, Sun Y, Sun Y, Sundaramoorthy V, Sung J, Suzuki H, Suzuki K, Suzuki N, Suzuki T, Suzuki YJ, Swanson MS, Swanton C, Sward K, Swarup G, Sweeney ST, Sylvester PW, Szatmari Z, Szegezdi E, Szlosarek PW, Taegtmeyer H, Tafani M, Taillebourg E, Tait SW, Takacs-Vellai K, Takahashi Y, Takats S, Takemura G, Takigawa N, Talbot NJ, Tamagno E, Tamburini J, Tan CP, Tan L, Tan ML, Tan M, Tan YJ, Tanaka K, Tanaka M, Tang D, Tang D, Tang G, Tanida I, Tanji K, Tannous BA, Tapia JA, Tasset-Cuevas I, Tatar M, Tavassoly I, Tavernarakis N, Taylor A, Taylor GS, Taylor GA, Taylor JP, Taylor MJ, Tchetina EV, Tee AR, Teixeira-Clerc F, Telang S, Tencomnao T, Teng BB, Teng RJ, Terro F, Tettamanti G, Theiss AL, Theron AE, Thomas KJ, Thome MP, Thomes PG, Thorburn A, Thorner J, Thum T, Thumm M, Thurston TL, Tian L, Till A, Ting JP, Titorenko VI, Toker L, Toldo S, Tooze SA, Topisirovic I, Torgersen ML, Torosantucci L, Torriglia A, Torrisi MR, Tournier C, Towns R, Trajkovic V, Travassos LH, Triola G, Tripathi DN, Trisciuoglio D, Troncoso R, Trougakos IP, Truttmann AC, Tsai KJ, Tschan MP, Tseng YH, Tsukuba T, Tsung A, Tsvetkov AS, Tu S, Tuan HY, Tucci M, Tumbarello DA, Turk B, Turk V, Turner RF, Tveita AA, Tyagi SC, Ubukata M, Uchiyama Y, Udelnow A, Ueno T, Umekawa M, Umemiya-Shirafuji R, Underwood BR, Ungermann C, Ureshino RP, Ushioda R, Uversky VN, Uzcategui NL, Vaccari T, Vaccaro MI, Vachova L, Vakifahmetoglu-Norberg H, Valdor R, Valente EM, Vallette F, Valverde AM, Van den Berghe G, Van Den Bosch L, van den Brink GR, van der Goot FG, van der Klei IJ, van der Laan LJ, van Doorn WG, van Egmond M, van

Golen KL, Van Kaer L, van Lookeren Campagne M, Vandenabeele P, Vandenberghe W, Vanhorebeek I, Varela-Nieto I, Vasconcelos MH, Vasko R, Vavvas DG, Vega-Naredo I, Velasco G, Velentzas AD, Velentzas PD, Vellai T, Vellenga E, Vendelbo MH, Venkatachalam K, Ventura N, Ventura S, Veras PS, Verdier M, Vertessy BG, Viale A, Vidal M, Vieira HL, Vierstra RD, Vigneswaran N, Vij N, Vila M, Villar M, Villar VH, Villarroya J, Vindis C, Viola G, Viscomi MT, Vitale G, Vogl DT, Voitsekhovskaja OV, von Haefen C, von Schwarzenberg K, Voth DE, Vouret-Craviari V, Vuori K, Vyas JM, Waeber C, Walker CL, Walker MJ, Walter J, Wan L, Wan X, Wang B, Wang C, Wang CY, Wang C, Wang C, Wang C, Wang D, Wang F, Wang F, Wang G, Wang HJ, Wang H, Wang HG, Wang H, Wang HD, Wang J, Wang J, Wang M, Wang MQ, Wang PY, Wang P, Wang RC, Wang S, Wang TF, Wang X, Wang XJ, Wang XW, Wang X, Wang X, Wang Y, Wang Y, Wang Y, Wang YJ, Wang Y, Wang Y, Wang YT, Wang Y, Wang ZN, Wappner P, Ward C, Ward DM, Warnes G, Watada H, Watanabe Y, Watase K, Weaver TE, Weekes CD, Wei J, Weide T, Weihl CC, Weindl G, Weis SN, Wen L, Wen X, Wen Y, Westermann B, Weyand CM, White AR, White E, Whitton JL, Whitworth AJ, Wiels J, Wild F, Wildenberg ME, Wileman T, Wilkinson DS, Wilkinson S, Willbold D, Williams C, Williams K, Williamson PR, Winklhofer KF, Witkin SS, Wohlgemuth SE, Wollert T, Wolvetang EJ, Wong E, Wong GW, Wong RW, Wong VK, Woodcock EA, Wright KL, Wu C, Wu D, Wu GS, Wu J, Wu J, Wu M, Wu M, Wu S, Wu WK, Wu Y, Wu Z, Xavier CP, Xavier RJ, Xia GX, Xia T, Xia W, Xia Y, Xiao H, Xiao J, Xiao S, Xiao W, Xie CM, Xie Z, Xie Z, Xilouri M, Xiong Y, Xu C, Xu C, Xu F, Xu H, Xu H, Xu J, Xu J, Xu J, Xu L, Xu X, Xu Y, Xu Y, Xu ZX, Xu Z, Xue Y, Yamada T, Yamamoto A, Yamanaka K, Yamashina S, Yamashiro S, Yan B, Yan B, Yan X, Yan Z, Yanagi Y, Yang DS, Yang JM, Yang L, Yang M, Yang PM, Yang P, Yang Q, Yang W, Yang WY, Yang X, Yang Y, Yang Y, Yang Z, Yang Z, Yao MC, Yao PJ, Yao X, Yao Z, Yao Z, Yasui LS, Ye M, Yedvobnick B, Yeganeh B, Yeh ES, Yeyati PL, Yi F, Yi L, Yin XM, Yip CK, Yoo YM, Yoo YH, Yoon SY, Yoshida K, Yoshimori T, Young KH, Yu H, Yu JJ, Yu JT, Yu J, Yu L, Yu WH, Yu XF, Yu Z, Yuan J, Yuan ZM, Yue BY, Yue J, Yue Z, Zacks DN, Zacksenhaus E, Zaffaroni N, Zaglia T, Zakeri Z, Zecchini V, Zeng J, Zeng M, Zeng Q, Zervos AS, Zhang DD, Zhang F, Zhang G, Zhang GC, Zhang H, Zhang H, Zhang H, Zhang H, Zhang J, Zhang J, Zhang J, Zhang J, Zhang JP, Zhang L, Zhang L, Zhang L, Zhang L, Zhang MY, Zhang X, Zhang XD, Zhang Y, Zhang Y, Zhang Y, Zhang Y, Zhang Y, Zhao M, Zhao WL, Zhao X, Zhao YG, Zhao Y, Zhao Y, Zhao YX, Zhao Z, Zhao ZJ, Zheng D, Zheng XL, Zheng X, Zhivotovsky B, Zhong Q, Zhou GZ, Zhou G, Zhou H, Zhou SF, Zhou XJ, Zhu H, Zhu H, Zhu WG, Zhu W, Zhu XF, Zhu Y, Zhuang SM, Zhuang X, Ziparo E, Zois CE, Zoladek T, Zong WX, Zorzano A, Zughaier SM (2016) Guidelines for the use and interpretation of assays for monitoring autophagy (third edition). Autophagy 12(1):1–222. https://doi.org/10.1080/15548627.2015.1100356

41. Rubinsztein DC, Cuervo AM, Ravikumar B, Sarkar S, Korolchuk V, Kaushik S, Klionsky DJ (2009) In search of an "autophagomometer". Autophagy 5(5):585–589
42. Kimura S, Noda T, Yoshimori T (2007) Dissection of the autophagosome maturation process by a novel reporter protein, tandem fluorescent-tagged LC3. Autophagy 3 (5):452–460

Chapter 11

Lipidomic Analysis of Cancer Cell and Tumor Tissues

Ramiz Islam Sk and Soumen Kanti Manna

Abstract

Due to their role in cellular structure, energetics, and signaling, characterization of changes in cellular and extracellular lipid composition is of key importance to understand cancer biology. In addition, several mass spectrometry-based profiling as well as imaging studies have indicated that lipid molecules may be useful to augment existing biochemical and histopathological methods for diagnosis, staging, and prognosis of cancer. Therefore, analysis of lipidomic changes associated with cancer cells and tumor tissues can be useful for both fundamental and translational studies. Here, we provide a high-throughput single-extraction-based method that can be used for simultaneous lipidomic and metabolomic analysis of cancer cells or healthy or tumor tissue samples. In this chapter, a modified Bligh-Dyer method is described for extraction of lipids followed by analysis of fatty acid composition by gas chromatography-mass spectrometry (GC-MS) or untargeted lipidomics using electrospray ionization mass spectrometry (ESIMS) coupled with reverse-phase (RP) ultraperformance liquid chromatography (UPLC) followed by multivariate data analysis to identify features of interest.

Key words Cancer, Lipidomics, RP-UPLC-ESIMS, Fatty acid methyl ester, GC-MS

1 Introduction

Lipids refer to a broad spectrum of molecules ranging from free fatty acids, prenols, and sterols to complex esterified and conjugated moieties such as phosphocholines (PC), phosphoethanolamines (PE), phosphoserines (PS), ceramide (Cer), sphingomyelin (SM), triglycerides (TG), and cholesteryl esters (CE) [1]. These molecules play diverse roles in cellular structure and functions such as triglycerides for long-term energy storage, phospholipids in lipid membranes, sterols in regulating membrane dynamics and hormonal functions, and arachidonic acid metabolites in inflammatory signal transduction and modulation of innate immune response to name a few. Inflammation is intricately connected to initiation and

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_11,

progression of cancer. One of the hallmarks of cancer cell is sustained cell proliferation that demands energy supply as well as synthesis of new membrane for daughter cells. Thus, it is not surprising that significant reprogramming of lipid metabolism has been identified in cancer [2, 3]. A number of studies showed that cancer cells exhibit elevated fatty acid metabolism [4, 5] as well as de novo fatty acid synthesis [6–8]. Phospholipase-mediated production of free fatty acids and subsequent β-oxidation has been shown to help cancer cells survive despite blockade of oncogenic PI3K-AKT-mTOR signaling [9]. A recent study indicated that alteration in lipid transport and metabolism is also involved in initiation of metastasis [10], which accounts for a lion share of deaths from cancer. In addition, de novo lipogenesis has been shown to be upregulated [11, 12] leading to an increase in lipid droplets in cancer stem cells [13], which have been found to play a key role in therapeutic resistance and relapse in cancer. Lipid droplets have also been shown to accumulate in circulating tumor cells [14] that are promising markers for diagnostic and prognostic purposes. Several studies have demonstrated that lipidomic profile of tumor is distinct from the adjacent normal tissue [15–23]. Monitoring of such differential lipidomic signature in patient biofluids [23–29] has been proposed for diagnostic, prognostic, and therapeutic applications in cancer. While they are lucrative targets, their robustness would hinge on correlation between changes in biofluid lipidome and metabolism of cancer cells in or outside tumor (in case of hematological malignancies and circulating tumor cells). In addition, the mechanistic connection between lipidomic changes and metabolism in cancer cells, as well as its microenvironment, needs to be investigated in detail to evaluate its value as a potential therapeutic target. Therefore, a detailed untargeted analysis of lipidomic signature of cancer cells and tissues in tandem with the analysis of metabolic signature has to be carried out.

Several methods for lipidomic analysis have been described in several studies –*see* Cajka T. and Fiehn O. (2014) for a comprehensive review [30]. In this chapter, we describe a simple workflow for high-throughput lipidomic analysis of cancer cells and tumor tissue, which also allows for simultaneous metabolomic analysis as shown in Fig. 1 and, thus, allows for an effective analysis of correlation between changes in metabolism and lipidomic signatures. This protocol is a modified version of the protocol that has been reported our earlier work [31]. The protocol starts with a biphasic extraction of lipids followed by either gas chromatography (GC) coupled with mass spectrometry (MS)-based fatty acid profiling or reverse-phase (RP) ultraperformance liquid chromatography coupled with electrospray ionization mass spectrometry (ESIMS)-based untargeted lipidomic analysis.

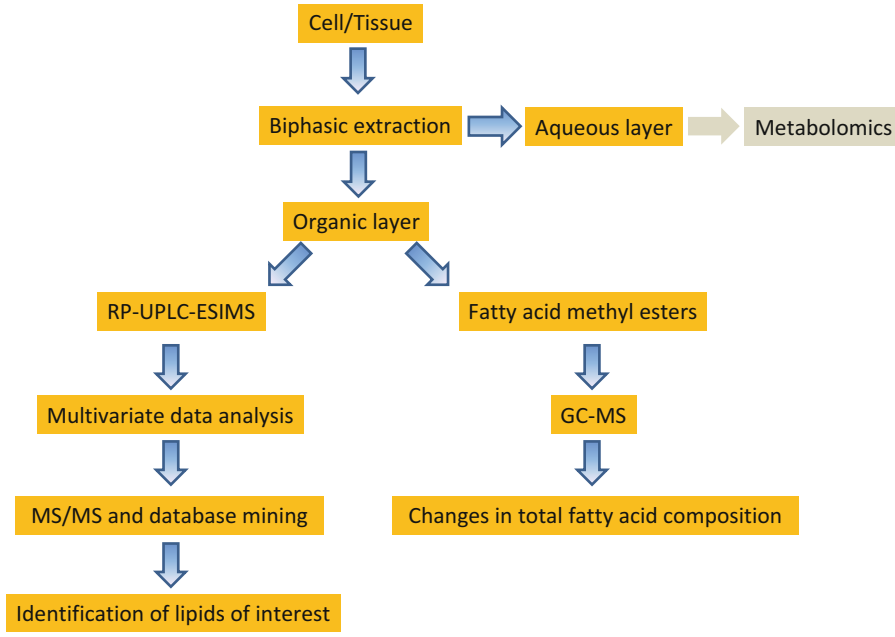


Fig. 1 Schematic representation of the analytical strategy showing a single-step biphasic extraction followed by GC-MS-based fatty acid analysis and UPLC-ESIMS-based untargeted lipidomic analysis with the option to perform metabolomic analysis on same tissue (or cell) samples in tandem. The aqueous layer can be used to proceed for RP or HILIC-based metabolomic analysis as described in Chapter 15

2 Materials

2.1 Solvents

1. Water (UHPLC-MS grade).
2. Methanol (UHPLC-MS grade).
3. Acetonitrile (UHPLC-MS grade).
4. Isopropanol (UHPLC-MS grade).
5. Chloroform (GC-MS grade).
6. n-Hexane (GC-MS grade).

2.2 Chemicals

1. Sulfadimethoxine.
2. Chlorpropamide.
3. Heptadecanoic acid.
4. Nonadecanoic acid.
5. LPC(17:0).
6. LPC(19:0).
7. LPA(17:0).
8. PC(17:0/17:0).
9. SM(17:0).
10. CE(17:0).

11. Methyl heptadecanoate.
12. Methyl nonadecanoate.
13. Sodium chloride.
14. Anhydrous sodium sulfate.
15. Ammonium formate.
16. Fatty acid methyl ester (FAME) mixture, C4–C24.
17. Formic acid.
18. 1.25 M methanolic HCl.
19. Authentic standards for lipids of interest.

2.3 Consumables

1. 1 mL glass sample vials.
2. 0.6 mL crimp top amber microvials (diameter 8 mm).
3. Crimp seals with aluminum cap and PTFE-silicone seals (diameter 8 mm).
4. PTFE non-disposable sleeves (O.D. × H 12 mm × 32 mm).
5. 2 mL screw-cap polypropylene tubes with gasket.
6. 5 mL glass tubes.
7. 15 mL screw-cap glass tube with teflon liner.
8. 1.4 mm zirconium oxide beads.
9. Graduated glass pipettes.
10. Liquid nitrogen.
11. Helium gas (purity >99.999%).
12. Nitrogen gas (purity >99.999%).
13. Argon gas (purity >99.999%).

2.4 Instruments

1. Refrigerated tissue homogenizer.
2. Pipettements (1–10 μL, 10–100 μL, 20–200 μL, 100–1000 μL).
3. Refrigerated centrifuge (capable of 18,000 × *g* at 4 °C).
4. Vortex.
5. Crimper.
6. Heating bath.
7. Sonicator bath.
8. Fume hood (with nitrogen connection).
9. Drying block fitted with nitrogen flow inlets.
10. SYNAPT G2 HDMS ESI qTOF coupled with Acquity™ UPLC platform (Waters Corp, Milford, MA) (*see* **Note 1**).
11. Agilent 7890B GC coupled with 5977B single quadrupole mass spectrometer (Agilent Technologies, Santa Clara, CA) (*see* **Note 1**).

2.5 Chromatography Columns

1. 100 mm Acquity UPLC BEH C18 (particle size 1.7 μm, i.d. 2.1 mm) for reverse-phase (RP) UPLC-ESIMS analysis (*see* **Note 1**).
2. 30 m Agilent HP5-MS column (i.d. 0.25 mm, film thickness 0.25 μm) for GC-MS analysis (*see* **Note 1**).

2.6 Buffer and Lipid Extraction Solvents

1. Phosphate-buffered saline (PBS): 10 mM Na_2HPO_4, 2.7 mM KCl, 1.8 mM KH_2PO_4, 137 mM NaCl, pH = 7.4.
2. Extraction solvent 1 (ES 1): 5 μM chlorpropamide and 4 μM nonadecanoic acid in methanol.
3. Extraction solvent 2 (ES 2): 2 μM each of LPC(17:0), PC (17:0/17:0), SM(17:0), and 5 μM LPA(17:0) in chloroform.

2.7 Solvents for Reconstitution and Dilution

1. LC reconstitution solvent: 5 μM LPC(19:0) in chloroform/methanol (1:1).
2. LC dilution solvent: Acetonitrile/isopropanol/water (1:2:1).

2.8 Solvents for Liquid Chromatography

1. Eluent A: 10 mM ammonium formate in 60% aqueous acetonitrile containing 0.1% formic acid.
2. Eluent B: 10 mM ammonium formate in a acetonitrile/isopropanol mixture (1:9) containing 0.1% formic acid.

2.9 Fatty Acid Stocks and Calibration Standards for Quantitation

1. Stocks: 10 mM solution of fatty acid of interest in hexane.
2. Diluent: 4 μM nonadecanoic acid in hexane.
3. Calibration standards: Serially diluted 100, 50, 25, 12.5, 6.25, 3.125, 1.562, 0.781, 0.390, 0.195, 0.098, 0.049, 0.024, and 0.012 μM solutions of fatty acid of interest in hexane each containing 4 μM nondecanoic acid (*see* **Note 2**).

2.10 Softwares

1. Microsoft Excel.
2. MassLynx package including MarkerLynx and TargetLynx (Waters, Milford, MA).
3. MassHunter Software package (Agilent, Santa Clara, CA).
4. SIMCA-P12+ (Umetrics, Kinnelon, NJ).
5. Prism version 6 (Graph Pad, La Jolla, CA).

2.11 Online Databases and Tools

1. LIPIDMAPS: http://www.lipidmaps.org/ [32, 33].
2. LipidBlast: http://fiehnlab.ucdavis.edu/projects/LipidBlast [34].
3. METLIN: https://metlin.scripps.edu/ [35].
4. HMDB: http://www.hmdb.ca/ [36].
5. MetaboAnalyst: http://www.metaboanalyst.ca/ [37].

3 Methods

3.1 Sample Collection and Storage

3.1.1 Tissue

Tissue samples should be placed immediately on ice after excision and flash-frozen into liquid nitrogen within 15 min (*see* **Note 3**) and eventually stored at −80 °C (*see* **Note 4**) in labelled cryovials.

3.1.2 Cell lines

Adherent Cells

1. Take supernatant media gently from the side of the adherent cell culture plate (*see* **Note 5**).
2. Gently scrape the cells, and collect them on one side of the plate with a rubber-tipped cell scraper (*see* **Note 6**).
3. Add 1 mL of the medium that was used to culture the cell to the plate and transfer it to a 2 mL screw-cap tube.
4. Take out an aliquot for counting cells (*see* **Note 7**). Process ~5 × 10^6 cells.
5. Centrifuge the rest at 300 × *g* for 5 min at 4 °C, and remove the supernatant media.
6. Wash with 1.5 mL PBS by centrifuging as above and remove the supernatant.
7. Wash with PBS for a second time and centrifuge at 500 × *g* for 5 min at 4 °C. Remove the supernatant as much as possible.
8. Proceed with lipid extraction (*see* Subheading 3.2) or flash-freeze in liquid nitrogen and store at −80 °C.

Non-adherent Cells (Suspension Culture)

1. Count cells (*see* **Note 7**). Take out volume of culture medium containing ~5 × 10^6 cells, and proceed to **step 2**.
2. Centrifuge culture medium containing cells at 300 × *g* for 5 min at 4 °C, and remove supernatant media as much as possible.
3. Gently wash the cell pellet with 1.5 mL ice-cold PBS by centrifuging as above, and remove the supernatant.
4. Wash the cell pellet with 1.5 mL PBS, centrifuge at 500 × *g* for 5 min at 4 °C, and remove supernatant.
5. Proceed for lipid extraction (*see* Subheading 3.2) or flash-freeze in liquid nitrogen, and store at −80 °C.

Patient-Derived Cells

1. Resuspend cells (*see* **Note 8**) into a suitable medium that allows cells to be viable.
2. Take an aliquot and count cells.
3. Take out medium containing minimum 2 × 10^6 cells.
4. Centrifuge medium containing cells at 300 × *g* for 5 min at 4 °C, and remove supernatant media as much as possible.

5. Gently wash the cell pellet twice with 1.5 mL ice-cold PBS by centrifuging as above, and remove the supernatant.
6. Proceed for lipid extraction (*see* Subheading 3.2) or flash-freeze in liquid nitrogen, and store at −80 °C.

3.2 Lipid Extraction

The following method is an adopted version of Bligh-Dyer method [38] (*see* **Note 9**).

3.2.1 Cell Extraction for Complex Lipid Analysis

1. Add 330 μL of ice-cold water (*see* **Note 10**) to the cell pellet to resuspend, take 30 μL out for protein estimation (using Bradford method), and transfer into 2 mL screw-capped polypropylene tubes with gasket.
2. Add 400 μL of chilled methanol containing internal standards (ES 1), and mix well by pipetting up and down.
3. Flash-freeze in liquid nitrogen, thaw, vortex vigorously for 30 s. Repeat freeze-thaw cycle 4 times.
4. Add 800 μL of chilled chloroform containing internal standards (ES 2).
5. Vortex for 30 s and shake at room temperature for 20 min.
6. Centrifuge at 18000 × *g* for 20 min at 4 °C.
7. Take out the upper aqueous layer as much as possible without disrupting the protein disk between two layers. Save the aqueous layer for metabolomic analysis (*see* **Note 10**).
8. Take out the lower organic layer carefully using a glass pipette penetrating by the side of the protein disk into 1 mL glass vial (*see* **Note 11**).
9. Transfer 400 μL from the glass vial into 5 mL glass tubes, and dry under nitrogen flow.
10. Reconstitute sample with 100 μL chloroform/methanol (1:1) containing 5 μM LPC(19:0), and transfer into 1 mL glass vials for LC-MS analysis.
11. Add three volume (300 μL) acetonitrile/isopropanol/water (1:2:1), and proceed to UPLC-ESIMS analysis (*see* Subheading 3.5.1).
12. Transfer another 200 μL from glass vial (*see* **step 8** above) into 15 mL Teflon-capped glass tubes, and proceed for the preparation of fatty acid methyl esters (*see* **step 1** of Subheading 3.3 below) for subsequent GC-MS-based fatty acid analysis (*see* Subheading 3.6).

3.2.2 Tissue Extraction for Complex Lipid Analysis

1. Take 10–20 mg tissue into 2 mL screw-capped vials with gasket.
2. Add 300 μL of ice-cold water and 400 μL chilled methanol containing internal standards.

3. Add 1.4 mm fresh zirconium oxide beads (~10–15) (*see* **Note 12**).
4. Place samples inside the homogenizer and homogenize at 6000 rpm (2 × 20 s with 30 s gap in between) at 10 °C (*see* **Note 13**).
5. Add 800 μL of chilled chloroform containing internal standards, and repeat **steps 5–12** as described above in Subheading 3.2.1.

3.3 Preparation of Fatty Acid Methyl Esters

1. Evaporate 200 μL of organic lipid extract under a stream of nitrogen in the 15 mL tube.
2. Add 1 mL methanolic HCl and screw the cap.
3. Heat at 50 °C overnight (or reflux for 2 hours).
4. Cool the reaction mixture on ice and add 4 mL of saline and 4 mL n-hexane.
5. Vortex vigorously for 2 min. Allow tubes to stand at RT for 10 min.
6. Centrifuge at 1000 × *g* for 10 min.
7. Remove the upper hexane layer with a clean glass pipette into a new glass tube.
8. Add 50 mg of anhydrous sodium sulfate, vortex, and centrifuge as mentioned above.
9. Transfer the supernatant carefully into another glass tube, and evaporate the solvent under stream of nitrogen.
10. Resuspend in 200 μL of hexane, transfer it with glass pipettes into 600 μL amber crimp top sample vials, cap and place them into PTFE sleeves, and proceed for GC-MS analysis (*see* Subheading 3.6.1).
11. Treat 200 μL each of pooled (*see* Subheading 3.4.1, **step 3**), extraction blank (*see* Subheading 3.4.2, **step 3**), and calibration standards (*see* Subheading 2.9, **item 3**) in exactly the same manner as described in **steps 1–10** above to use for GC-MS analysis (*see* Subheading 3.6.1).

3.4 Preparation of Quality Control (QC) Samples

3.4.1 Pooled Samples

1. Take 20 μL (more if the total sample number is less than 30 or less if the total sample number is more than 50) organic extract from each individual sample (after **step 8** of Subheadings 3.2.1 or 3.2.2) into a glass vial to make a pooled extract with total volume ≥ 600 μL.
2. Take 400 μL of this pooled extract and process as per **steps 9–11** of Subheading 3.2.1 to prepare a pooled sample for UPLC-ESIMS-based lipidomic analysis (*see* Subheading 3.5.1 below).

3. Take another 200 μL of the pooled extract from **step 1** above into 15 mL Teflon-capped glass tube, and prepare fatty acid methyl ester as described in **steps 1–10** of Subheading 3.3 above.

3.4.2 Extraction Blank

1. Add 300 μL water and extraction solvents in an empty 2 mL screw-capped tube, and perform all steps as described in **steps 1–8** of Subheading 3.2.1 (for cells) or Subheading 3.2.2 (for tissues). These will lead to extraction blank samples for cell and tissue extraction, respectively.
2. Take out 400 μL organic layer from the step above, and perform **steps 9**, **10**, and **12** as described in Subheading 3.2.1 to obtain extraction blank for UPLC-ESIMS analysis.
3. Take out another 200 μL organic layer, and perform **steps 1–10** as described in Subheading 3.3 to obtain extraction blank for GC-MS analysis.

3.4.3 Solvent Blank

1. Take 100 μL chloroform/methanol (1:1) into 1 mL glass vial, and add 300 acetonitrile/chloroform/water (1:2:1) containing 5 μM chlorpropamide to prepare solvent blank for UPLC-ESIMS analysis.
2. Use hexane containing 4 μM methylnonadecanoate as solvent blank for GC-MS analysis.

3.4.4 Authentic Standard Mixture

1. Standard mixture for UPLC-ESIMS: 10 μM heptadecanoic acid, 10 μM nonadecanoic acid, 2 μM LPC(17:0), 5 μM LPA (17:0), 2 μM PC(17:0/17:0), 2 μM SM(17:0), 2 μM CE (17:0) in acetonitrile/isopropanol/water (1:2:1).
2. Standard reference mixture for GC-MS: FAME mixture C4–C24.

3.5 UPLC-ESIMS-Based Untargeted Lipidomic Analysis

3.5.1 UPLC-ESIMS Experiment

1. Room temperature should be maintained at 22 ± 1 °C (*see* **Note 14**). Turn on and stabilize the source voltage and gas flow over at least 30 min.
2. Infuse 200 ng/mL sulfadimethoxine at 20 μL/min as lockspray ($m/z = 311.0814$ in ESI + and 309.0658 in ESI modes).
3. Prepare a randomized sample list using a random number generator function of Microsoft Excel with six pooled samples at the beginning (*see* **Note 15**) and pooled, extraction blank, solvent blank, and standard mixture samples injected intermittently throughout the run.
4. Optimize cone and lens voltages to achieve desired mass accuracy, peak shape, and intensity for the lock mass (sulfadimethoxine) in positive ($m/z = 311.0814$) and negative mode ($m/z = 309.0658$) as per instructions in the operator's manual of the mass spectrometer. Define these as lock masses in the MS method for respective modes.

Table 1
LC gradient condition

Time Segment (min)	Flow rate (ml/min)	Solvent A (%)	Solvent B (%)
0	0.4	60	40
0–2	0.4	57	43
2–2.1[a]	0.4	50	50
2.1–12	0.4	46	54
12–12.1[a]	0.4	30	70
12.1–18	0.4	1	99
18–18.1	0.4	60	40
18.1–20	0.4	60	40

[a]Indicates segments with ballistic gradient

5. Calibrate the mass axis (50–1000 Da) using 0.1 mg/mL sodium formate in 90% isopropanol as per instructions in operator's manual.
6. Set up typical experimental parameters (*see* **Note 16**) for the positive mode analysis of lipids in the *m/z* range 50–1000 Da in the sensitivity mode with capillary voltage at 3.1 kV, sampling cone at 55 V, source temperature at 125 °C, desolvation temperature at 400 °C, desolvation gas (N_2) at 850 L/hr., cone gas (N_2) at 50 L/h.
7. For negative mode (*see* **Note 16**), perform the analysis in the same *m/z* range in sensitivity mode with capillary voltage at 3.2 kV, sampling cone at 70 V, source temperature at 125 °C, desolvation temperature at 400 °C, desolvation gas (N_2) at 850 L/h, cone gas (N_2) at 50 L/h.
8. Perform the chromatographic separation of lipids on a 100 mm Acquity UPLC BEH C18 column using a gradient program shown in Table 1.
9. Load glass sample vials, and maintain the sample chamber at 12 °C, and inject 5 μL of each sample for analysis.

3.5.2 Data Inspection, Quality Control, and Feature Extraction

1. Inspect chromatograms and check for any sudden shift in baseline in QC samples or drop in intensity in standard mix and pooled samples.
2. Inspect chromatograms of solvent blank samples for any carry-over (*see* **Note 17**) by looking for peaks of compounds present in the immediate preceding sample.

3. Inspect extraction blanks for any contamination or artifacts (peaks other than internal standards present in ES 1 and ES 2).
4. Extract standard mix chromatograms for peaks corresponding to chlorpropamide ($m/z = 277.0415+, 275.0257-$), nonadecanoic acid ($m/z = 297.279-$), LPC(17:0) ($m/z = 510.356+$), LPC(19:0) ($m/z = 538.387+$), LPA (17:0) ($m/z = 445.233-$), SM(17:0) ($m/z = 717.591+$), and PC(17:0/17:0) ($m/z = 762.601+$) to check for consistency of retention time and mass error (in ppm).
5. Check sample chromatograms for consistency of baseline and retention time of internal standards. Figure 2a, b shows representative chromatograms of a lipid extract in positive and negative ionization modes.
6. Overlay pooled samples (*see* Fig. 3) to check for reproducibility of chromatograms, integrate extracted ion chromatograms of internal standards as well as representative endogenous peaks such as LPC(16:0) ($m/z = 496.34+$), LPC(18:2) ($m/z = 520.34+$), PC (38:4) ($m/z = 810.61+$) in ESI+ mode or LPE(18:0) ($m/z = 480.3-$), PE(38:5) ($m/z = 764.53-$), PC(36:3) ($m/z = 828.58-$, formate adduct), etc. Check mass error range.
7. Calculate coefficients of variation (CV) (*see* **Note 18**) for internal standard as well as endogenous peaks. Figure 4a shows extracted ion chromatograms of ion with $m/z = 810.61+$ (shown in Fig. 4b) that shows a CV $< 6\%$ across multiple injections of pooled sample and a retention time shift of 0.15 min (compared to its peak width of 0.4 min) across a run consisting over 50 injections. CVs should typically be less than 10% for representative peaks and retention time shift less than half of the peak width.
8. Process the centroided and integrated mass spectrometric data using MarkerLynx method with a mass window of 0.02 Da, retention time window of 0.5 min, and intensity threshold of 200 counts and with box for data deisotoping checked in the method (*see* **Note 19**).

3.5.3 Data Preprocessing

1. After total ion count (TIC)-normalization of MarkerLynx-generated data matrix comprising features (m/z-retention time pairs), perform principal components analysis of the Pareto-scaled data matrix using SIMCA-P + 12. Check if solvent blank, extraction blank, and standard mix samples are tightly clustered in the scatter plot. Otherwise, check chromatograms for quality, extraction parameters, instrument log, and repeat experiment (*see* **Note 20**).
2. Remove, solvent blanks, extraction blanks, and standard mix samples from the analysis.

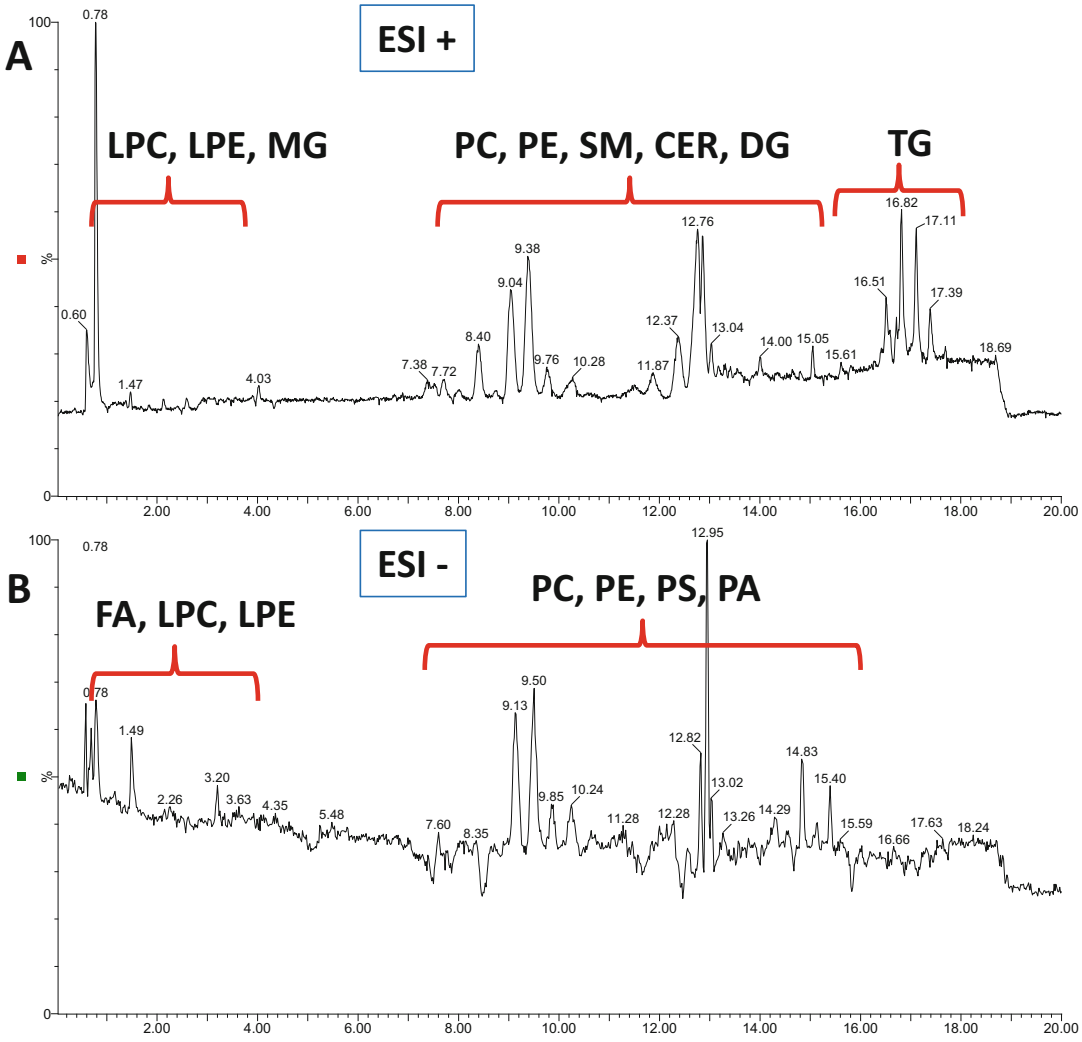


Fig. 2 Representative total ion chromatogram of an untargeted lipidomic analysis of liver tissue in ESI (**a**) positive and (**b**) negative modes. Typical retention time window for compond classes are shown. *LPC*, lyspophosphocholine; *LPE*, lysophosphoethanolamine; *MG*, monoacylglycerol; *PC*, phosphocholine; *PE*, phosphoethanolamine; *SM*, sphingomyelin; *CER*, ceramide; *DG*, diacylglycerol; *TG*, triacylglycerol; *FA*, fatty acid; *PS*, phosphatidylserine; *PA*, phosphatidic acid

3. Exclude features that show missing values in more than 10% of pooled samples. Also exclude all features with ≥75% missing values in individual samples. Replace all remaining missing values by a number ten times smaller than the lowest value in the data matrix.
4. Inspect whether pooled samples are clustered tightly compared to individual samples in the scatter plot before proceeding to the next step (*see* **Note 21**).

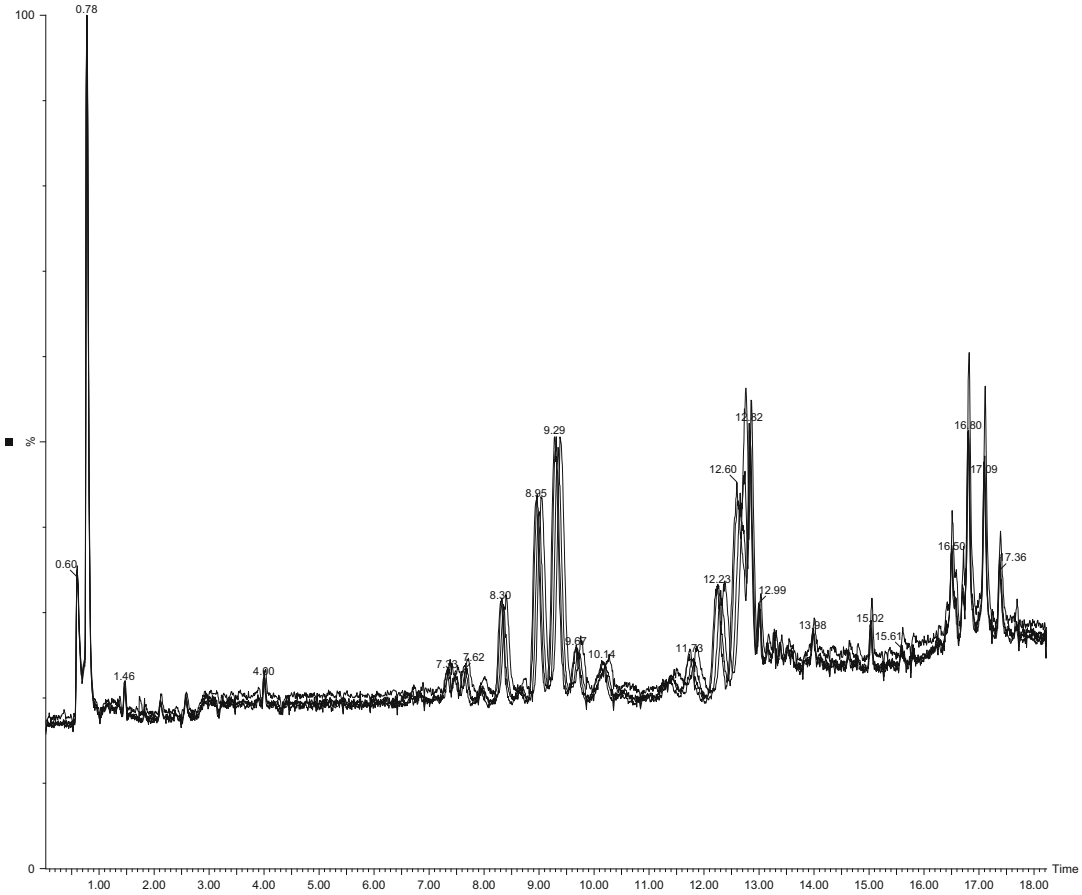


Fig. 3 Overlay of total ion chromatograms (ESI positive mode) of 5 pooled samples injected intermittently during a run of over 50 samples to check consistency of mass spectrometer performance and chromatogram reproducibility

5. Remove pooled samples from the data matrix, and upload on to the MetaboAnalyst webserver as a peak intensity table (with unpaired samples in rows and features in columns) for statistical analysis. Choose sum normalization, log transformation (*see* **Note 22**), and Pareto-scaling of the data to analyze intercorrelation between features across sample set.
6. Download the correlation matrix and screen features with correlation coefficients >0.9 (*see* **Note 23**), compare their elution profile, and remove potential adducts or fragments. For example, Fig. 5a shows the extracted chromatograms of a highly correlated co-eluting ion cluster comprising of the parent ion ($m/z = 538.391+$), K^+ adduct ($m/z = 576.343+$), Na^+ adduct ($m/z = 560.371+$) as well as in-source fragments due to loss of one water molecule ($m/z = 520.376+$) and two water molecules ($m/z = 502.377+$), as found in the mass spectrum (Fig. 5b) under the LPC(19:0) peak at 2.68 min.

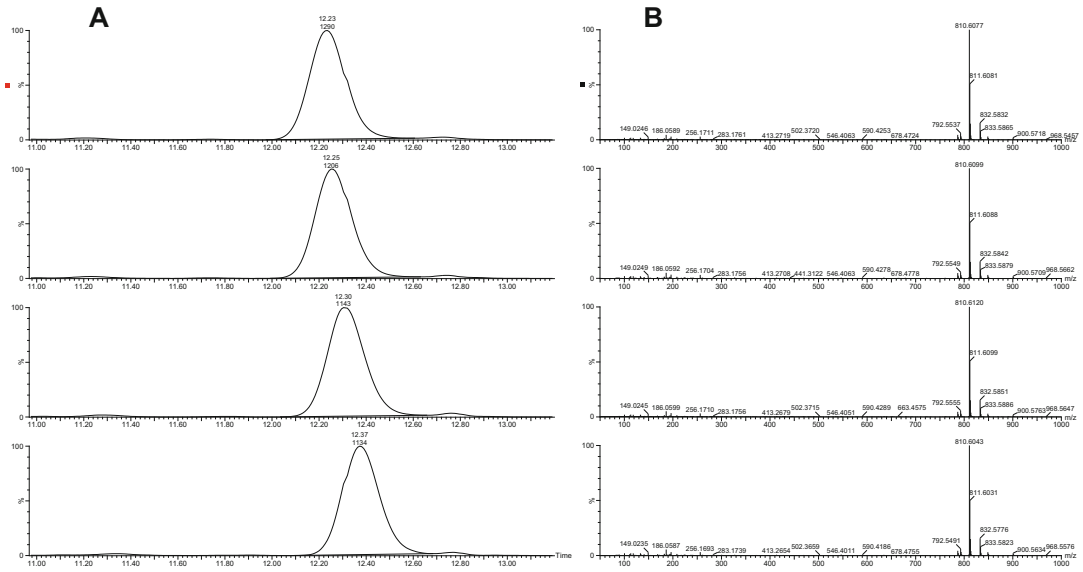


Fig. 4 (**a**) Extracted ion chromatograms for $m/z = 810.61+$ in pooled samples showing integrated area under each peak and the retention time along with (**b**) the mass spectrum under the peak on the right. The CV was calculated to be 6%, and a retention time drift of 0.15 min were observed over a run of 50 samples

7. Remove all adducts and fragments except the putative parent ion from the data matrix *(see* **Note 24**).
8. Remove all features consistently present in solvent blank and extraction blank samples (*see* Subheading 3.5.2), and use this trimmed data matrix for pattern recognition and feature identification.

3.5.4 Pattern Recognition and Identification of Features of Interest

1. For tissue samples, normalize ion counts for each feature with respect to tissue weight or total ion count. For cell lines, normalize with respect to cell count, total protein concentration or total ion count.
2. Define class variables after importing the normalized data matrix into SIMCA-P 12+. Define class variable (Y) as "0" for non-cancerous tissue or cell and "1" for cancer tissue or cell. No class variable needs to be defined for pooled samples.
3. Exclude "Y" variables, and perform a principal components analysis on Pareto-scaled data to check that clustering is tight for pooled samples compared to individual samples in the scores scatter plot.
4. Remove pooled samples from the analysis, check for any segregation of non-cancer and cancer samples. In case non-cancer and cancer samples are clustered distinctly, identify ions contributing to such segregation from the loadings plot.

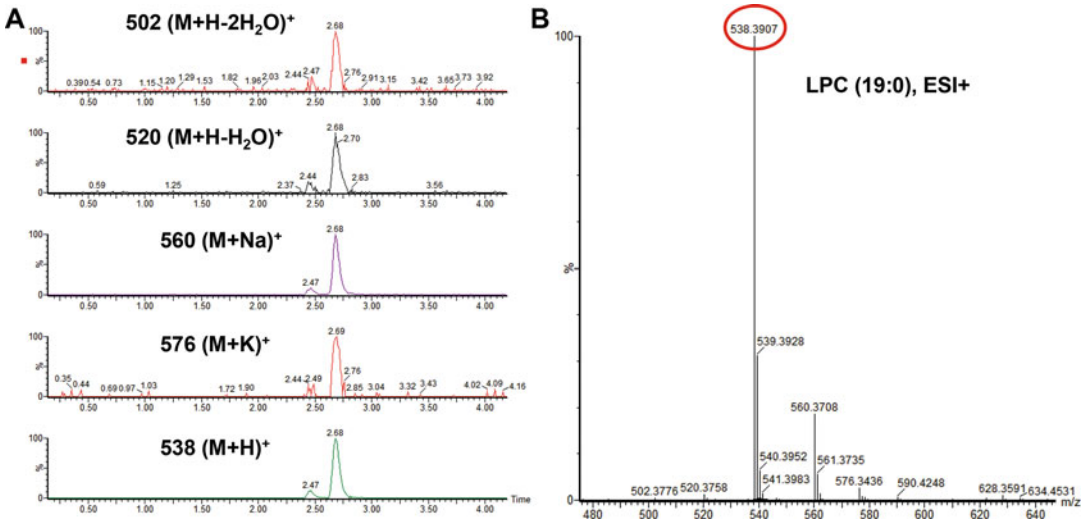


Fig. 5 (**a**) Extracted ion chromatograms of LPC(19:0) ($m/z = 538.391+$) and it's K^+ adduct ($m/z = 576.343+$), Na^+ adduct ($m/z = 560.371+$) as well as in-source fragments due to loss of one ($m/z = 520.376+$) and two water molecules ($m/z = 502.377+$) along with (**b**) the mass spectrum under the peak at 2.68 min

5. If there is no segregation of cancer and non-cancer samples, remove most abundant ions, and repeat the PCA analysis. PCA is driven by most abundant features. They tend to mask relatively low abundant features that might be of interest with respect to the correlation between metabotype and phenotype.
6. Perform supervised orthogonal projection to latent structures-discriminant analysis (OPLS-DA) (*see* **Note 25**), and identify ions contributing significantly to class prediction from the loadings S-plot with $p(corr)[1] > 0.8$ or $p(corr)[1] < -0.8$.
7. Submit normalized and class-labelled data matrix (without pooled samples) to MetaboAnalyst server and choose no normalization, log transformation, and Pareto-scaling of the data.
8. Perform heatmap analysis to get an idea about the pattern of change in lipidome. Use "volcano plot" analysis with $p < 0.05$ using nonparametric test assuming unequal variance, and apply the FDR correction to identify features that change significantly between classes.

3.5.5 Lipid Identification

1. Submit masses of features of interest to databases like LIPID-MAPS, HMDB, or METLIN. Use mass error window of according to results obtained in **steps 4** and **6** of Subheading 3.5.2 (*see* **Note 26**). In addition to protonation or deprotonation, choose common adducts (e.g., Na^+, K^+, $NH4^+$ for ESI+ and Cl^-, $HCOO^-$ for ESI-) or fragments (e.g., $-H_2O$, $-2H_2O$) (*see* **Note 27**).

Table 2
Signature fragments for lipid identification

Lipid Class[a]	Molecular ion m/z[b]	ESI Pos	ESI Neg	Signature fragments[c]	Neutral losses[d]
PC[e]	Even	Y	Y	184+, 104+, 224−, 140−	60 (−)
PE	Even	Y	Y	196−, 140−	141 (+)
PS	Even		Y	153−	185 (+), 87 (−)
PI	Odd		Y	223−, 241−, 259−, 297−, 315−	
PG	Odd		Y	153−, 227−, 171−	
PA	Odd		Y	153−, 135−	
SM[5]	Odd	Y	Y	184+, 104+	60 (−)

[a] *PC* glycerophosphocholine, *PE* glycerophosphoethanolamine, *PS* glycerophosphoserine, *PI* gycerophosphoinositol, *PG* glycerophosphoglycerol, *PA* glycerophosphate, *SM* sphingomyelin
[b] Both (M+H)+ and (M−H)− will have same odd/even characteristic.
[c] The mode of ionization indicated next to the m/z value of the ion. Numbers in Da.
[d] The mode of ionization, where the neutral loss is observed, is presented in brackets
[e] PC and SM shows up as formate adduct in negative mode due to presence of formate in the buffer and undergoes a neutral loss of 60 (HCO_2+CH_3) as seen in Figure 6c.

2. Perform MS/MS analysis on ions of interest by ramping up collision energy 5 to 50 eV (*see* **Note 28**) with argon as the collision gas. Chromatography and other mass spectrometry conditions including lockspray should be similar to that described above for the respective sample and mode of analysis (refer to **steps 6–8** in Subheading 3.5.1).
3. Compare experimental fragmentation pattern reported in the database or theoretical fragmentation pattern with the observed fragmentation pattern.
4. The integer mass of the molecular ion (M+H or M−H) of lipids (*see* Table 2) can help to reduce plausible structures. In addition, several lipid classes produce signature fragmentation patterns (*see* Table 2) that can not only help to identify the lipid class but also to identify the fatty acid side chains, specifically, number of carbons and unsaturations to help in identification of the species [39–43]. For example, Fig. 6a shows that LIPID-MAP returns three and six hits for ions with *m/z* value 480.31 and 568.36, respectively, in the negative ionization mode assuming possible Cl^-, $HCOO^-$, CH_3COO^- adducts and M-CH_3 fragments along with M−H. Presence of the 283 fragment in MS/MS spectra (Fig. 6b) unequivocally suggests the species to be LPE(18:0). Similarly, presence of the 283 fragment in the MS/MS spectra of the 568 peak suggests it to be formic acid adduct of LPC(18:0). The neutral loss of 60 mass units corresponds to a loss of (HCO_2 + CH_3) from the formic acid adduct [43].

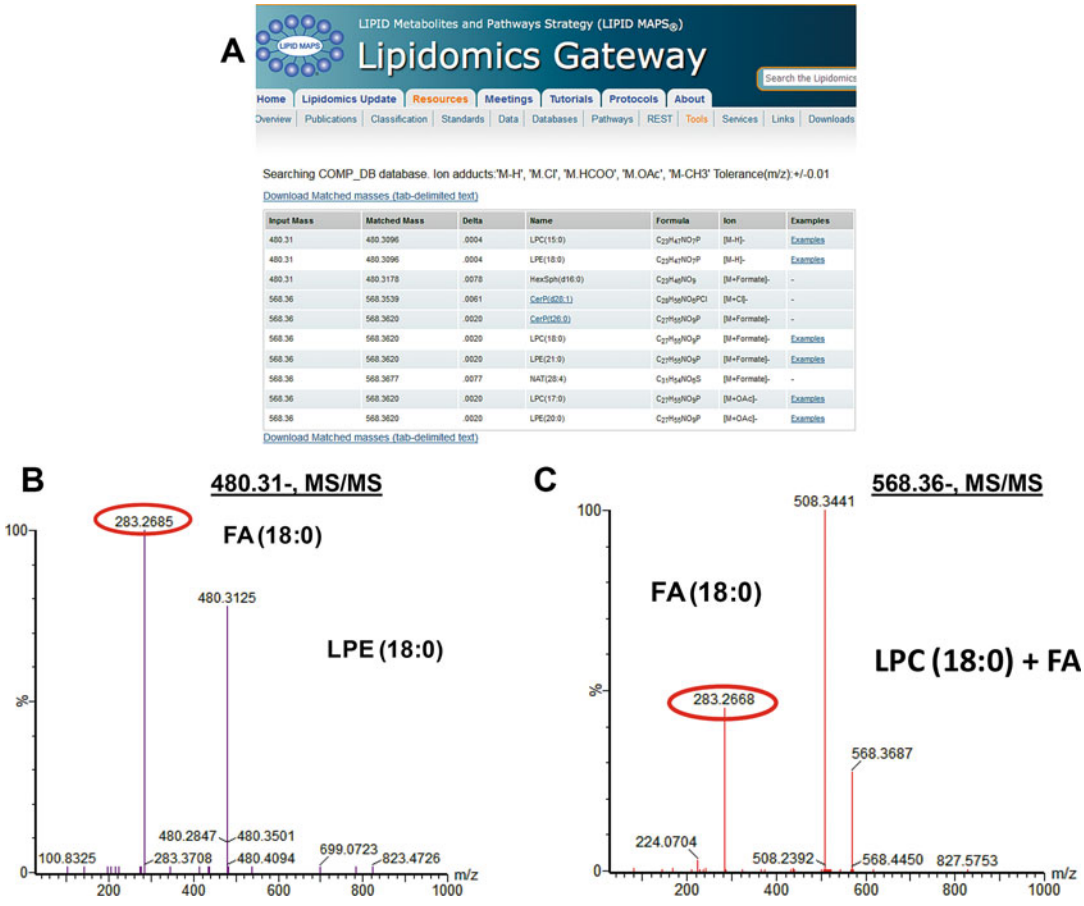


Fig. 6 (**a**) LIPIDMAPS database showing three and six possible identities for 480.31- and 568.34-, respectively, with a mass error of 0.01 units. (**b**) MS/MS analysis showing the 18:0 fatty acid ion to identify 480.31- as LPE (18:0). (**c**) MS/MS analysis showing the 18:0 fatty acid peak to identify 568.34- as formic acid adduct of LPC (18:0)

5. While selecting a particular precursor ion for MS/MS, it should be remembered that due to structural or stereoisomerism, multiple lipids in the same class may have same molecular weight eluting at different times. For example, the ion 782.57+, which returns 16 (*see* **Note 29**) hits in LIPIDMAPS search (Fig. 7a), shows two peaks at 7.27 min and 8.55 min (Fig. 7b). MS/MS analysis (Fig. 7c) shows a phospholipid signature (dotted box) along with LPC(16:0) fragment for the former and LPC(18:2) fragment for the latter, suggesting them to be PC(16:0/20:4) and PC(18:2/18:2), respectively.
6. Finally, whenever possible, run authentic standards to confirm the identity by comparing the retention time and the fragmentation pattern (*see* **Note 30**).

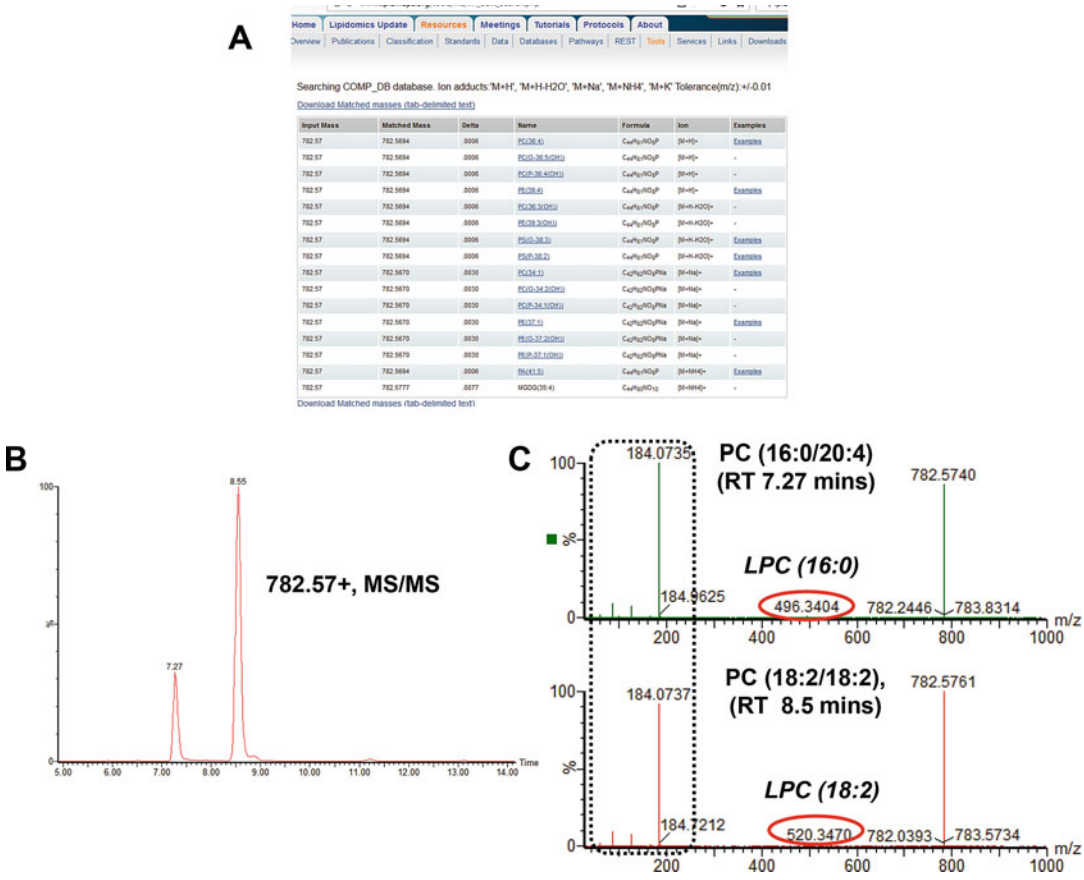


Fig. 7 (**a**) LIPIDMAPS database showing 16 possible identities for 782.57+ ion which (**b**) shows peaks at 7.27 as well as 8.55 min in the chromatogram. (**c**) MS/MS analysis revealed signature for the phosphocholine head group (showed with dotted box) under both peaks while showing LPC (16:0) and LPC (18:2) fragments under 7.27 min and 8.55 min peaks, respectively, to identify them as PC(16:0/20:4) and PC(18:2/18:2)

3.6 GC-MS-Based Analysis of Fatty Acid Methyl Esters

3.6.1 GC-MS Experiment

1. Place samples into the autosampler.
2. Use hexane and methanol for strong and weak wash of injector needle.
3. Inject at least five pooled samples before injecting individual samples (*see* **Note 15**).
4. Create a randomized sample list using random number generator function of Microsoft Excel with pooled, extraction blank, and solvent blank samples (*see* Subheading 3.3, **step 11**) and reference FAME mixture C4–C24 inserted intermittently throughout the run.
5. Add methyl esters of serially diluted calibration standard before and after the run (*see* Subheading 3.3, **step 11**).
6. Inject 1 μL of sample into the inlet, and analyze in splitless mode using helium as the carrier gas (1 mL/min) and

following instrument parameters: inlet temperature at 300 °C, transfer line temperature at 300 °C, source temperature at 230 °C, quadrupole temperature at 150 °C, mass range at 46–500 Da.

7. Equilibrate the column for 1 min at 70 °C before injection, maintain the column at 70 °C for 4 min, and ramp it at 5 °C/min to 320 °C, and hold at that temperature for 5 min.
8. In the MS section of the method, enter solvent delay of 6 min to keep the MS off during first 6 min of the run.

3.6.2 Data Analysis

1. Open chromatograms using MassHunter Qualitative analysis suite or MSD ChemStation. The chromatogram for the C2–C24 FAME reference standard is shown in Fig. 8a. Figure 8b shows the fatty acid methyl ester profile of a liver sample, while Fig. 8c shows that for HepG2 cells. Typical retention time for FAME standards are given in the Table 3.
2. Examine the solvent blank samples for any carry-over by looking for peaks present in the sample immediately preceding it and extraction blank samples for contamination.

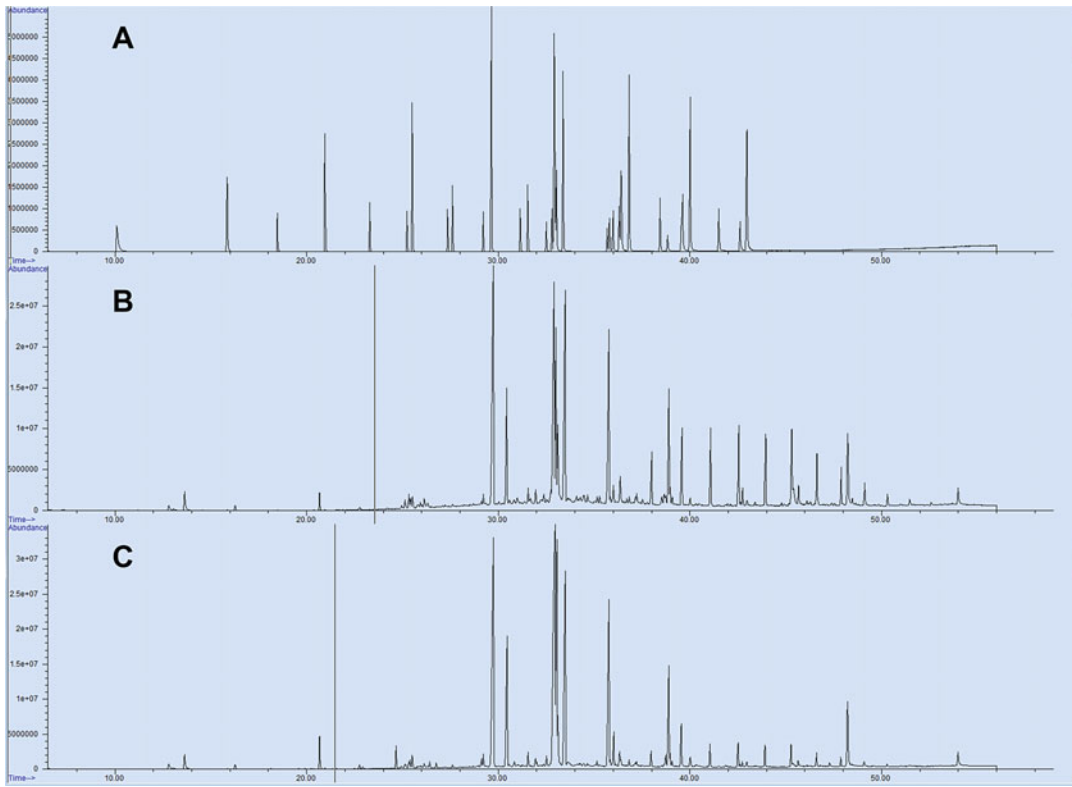


Fig. 8 Total ion chromatogram of fatty acid methyl esters from (**a**) reference fatty acid methyl ester mixture (C4–C24), (**b**) liver tissue extract, and (**c**) HepG2 cell extract. Retention time of representative fatty acid methyl esters is given in Table 3

Table 3
Retention times of fatty acid methyl ester standards

Compound	Structure[a]	Retention time (min)
Methyl octanoate	8:0	10.1
Methyl decanoate	10:0	15.84
Methyl undecanoate	11:0	18.5
Methyl dodecanoate	12:0	20.95
Methyl tridecanoate	13:0	23.29
Methyl myristoleate	14:1	25.23
Methyl myristate	14:0	25.53
Methyl *cis*-10-pentadecenoate	15:1	27.34
Methyl pentadecanoate	15:0	27.6
Methyl palmitoleate	16:1	29.2
Methyl palmitate	16:0	29.64
Methyl *cis*-10-heptadecenoate	17:1	31.1
Methyl heptadecanoate	17:0	31.5
Methyl linolenate	18:3	32.5
Methyl linoleate	18:2	32.8
Methyl oleate	18:1	33.04
Methyl stearate	18:0	33.38
Methyl arachidonate	20:4	35.68
cis-8,11,14-Eicosatrienoic acid methyl ester	20:3	36.0
cis-11,14-Eicosadienoic acid methyl ester	20:2	36.3
cis-11-Eicosenoic acid methyl ester	20:1	36.4
Methyl eicosanoiate	20:0	36.83
Methyl heneicosanoate	21:0	38.45
cis-4,7,10,13,16,19-Docosahexaenoic acid methyl ester	22:6	38.84
cis-13,16-Docosadienoic acid methyl ester	22:2	39.56
Methyl docosanoate	22:0	40.0
Methyl tricosanoate	23:0	41.51
Methyl cis-15-tetracosenoate	24:1	42.63
Methyl tetracosanoate	24:0	42.97

[a]Indicates the number of carbon atoms followed by number of double bonds

3. Check for any considerable shift (> half of the peak width) in retention time (*see* **Note 31**).
4. Identify fatty acids present in samples by comparing retention time and fragmentation pattern of the corresponding methyl esters in the C4–C24 FAME mixture reference and the sample (*see* **Note 32**).
5. Create a method for extraction of each of these peaks by defining the retention time window and assigning an abundant and characteristic ion as qualifier along with two other characteristic ions as qualifier (*see* **Note 33**).
6. Integrate the area under the extracted chromatogram of peaks for reference standards and representative peaks from the pooled samples to calculate the coefficient of variation. Typically, it should be within 5% for reference standards and, at most, 10% for pooled samples.
7. For serially diluted methyl ester calibration standards, define the highest concentration of the fatty acid of interest used, dilution factors, and number of dilutions to generate desired number of concentration levels for each standard.
8. Use nonadecanoic acid as the internal standard, and validate the method, and choose to integrate and analyze.
9. Find the limit of quantitation and linear dynamic range for $r > 0.95$. In case the concentration range of the fatty acid of interest is beyond the linear dynamic range, either concentrate or dilute samples as appropriate.
10. Finally, normalize fatty acid concentrations by tissue weight for tumor and adjacent normal tissues. For cells, normalize values using the cell count or the total protein concentration.

3.7 Statistical Tests

1. To test the significance of difference in fatty acid or lipid composition of cells or tissues, use Mann-Whitney U test with 95% confidence interval and apply correction for multiple testing (such as Bonferroni's correction) using Prism software.
2. Examine correlation between any phenotype of interest and level of fatty acid or lipids using Pearson correlation analysis.
3. Perform receiver operator characteristic (ROC) analysis to examine the ability of lipidomic signatures to distinguish between cancer and normal or any other phenotypically distinct sample sets.

4 Notes

1. The method described here was developed and optimized using aforementioned instruments (*see* Subheading 2.4) and softwares (*see* Subheading 2.10). However, the basic analytical

approach is platform-independent, and it can be adopted for similar instruments from other vendors. Similar columns for ultrahigh-performance liquid chromatography (UHPLC) or gas chromatography from other vendors could also be used.

2. Prepare 100 μM solution of fatty acid to be quantitated in hexane containing 4 μM nonadecanoic acid. This solution is serially diluted 1:1 (v/v) with hexane containing 4 μM nonadecanoic acid to produce calibration standards for respective fatty acids.

   Often, more than one fatty acid are desired to be quantitated. In such case make a mixture containing 100 μM of each of the fatty acid of interest, and dilute as described above. Proceed for preparation of fatty acid methyl ester as described in Subheading 3.3 and GC-MS analysis as described in Subheading 3.6. After GC-MS analysis, create method (*see* **step 5**, Subheading 3.6.2) for quantitation of fatty acids by defining the extraction window according to the retention time observed with the authentic standards and assigning the molecular ion of respective fatty acid as the quantifier or one of the qualifier ions.

3. Degradation starts immediately after excision of the tissue. Delay >15 min may cause changes in phospholipid profiles [44].

4. Storage is not recommended beyond 1 year. Significant degradation of proteins and lipids is associated with long-term storage [45, 46]. If multiple use of the same samples is foreseen, samples should be cut into small single-use portions on dry ice using surgical blades, and stored at −80 °C. Preferably, use separate surgical blades for each sample to avoid cross-contamination. In the case in which the same blade is used, wipe with fresh lint-free tissue paper; dip it into water, followed by methanol followed by chloroform; and air-dry before the use on the next sample.

5. Plates are preferred over T-flasks for the ease of harvesting cells, particularly, for harvesting with dry ice and scraping.

6. Avoid trypsinization, particularly, if simultaneous metabolomic analysis is foreseen or planned. It has been shown to cause significant metabolite loss [47].

7. Use trypan blue staining to count dead cells. Culture should have no more than 10% of dead cells. Both cells under comparison should have a similar number of dead cells. The lipid composition may be skewed by of the presence of dead cells or cell debris. If needed live cells should be separated from dead ones by centrifugation or using Ficoll before further processing.

8. Patient-derived cancer cells, such as those from hematological malignancies or circulating tumor cells, are collected through a variety of techniques including immune capture, density gradient centrifugation, or size selection. Many of the chemicals used in these procedures, such as Ficoll, may interfere with MS analysis. They should be removed by thorough wash before proceeding for extraction.
9. Methanol/chloroform/water-based biphasic extraction has been used by several researchers and works well for untargeted profiling. However, if specific lipid class is of interest, extraction method should be chosen accordingly. Visit following sites for more on lipid-specific extraction protocols:

   http://www.cyberlipid.org/
   http://lipidlibrary.aocs.org/History/content.cfm?ItemNumber=40362
   http://lipidlibrary.aocs.org/Analysis/content.cfm?ItemNumber=40376#G
10. Since patient-derived cells or tissues are often available in small quantities, it is a good idea to perform metabolomic analysis on the same sample. More importantly, it allows for analysis of correlation between metabolic reprogramming and lipidomic changes observed in the cancer cells or tissue and thus helps to give a better mechanistic understanding of both metabolomic and lipidomic data. Add 5 μM alpha-aminopimelic acid if simultaneous metabolomic analysis is planned. This aqueous layer can be analyzed using untargeted metabolomics approaches. Otherwise, modify the protocol to use saline instead of water. This helps to improve the recovery of lipids into the organic phase.
11. Be steady and gentle while penetrating the disk and taking out the lower organic layer to ensure that disk is not totally disrupted and no particulate matter comes into the organic layer. In case the volume of the cleanly collected organic layer is not sufficient, modify the **steps 10–12** accordingly.
12. For most normal and tumor tissues, these beads will work. However, use 2.8 mm beads for bony and fibrous tissues.
13. Typically, two cycles of 6000 rpm (20 s-30 s-20 s) suffices for most tissues. However, check samples to ensure that it has become a homogenous slurry, and use additional cycles before proceeding to the next step.
14. Mass accuracy of TOF instruments may be affected by temperature fluctuations. Check the room temperature profile if sudden changes in mass profile are noted during a long run.
15. Significant change in response and shift in retention time is observed between the first injection into a clean column and

subsequent injections of the same sample. Therefore, column should be conditioned by injecting few pooled samples after which the chromatogram becomes more reproducible.

16. For every platform these experimental parameters should always be optimized afresh.
17. Lipids are more likely to stick to the column compared to metabolites/small molecules. Strong wash with a nonpolar solvent is essential to minimize carry-over. Make sure that column is not overloaded. This would lead to peak shift and tailing. The proportion of tissue and reconstitution volume needs to be optimized.
18. While calculating coefficient of variation of higher lipid species such as PC, PE, TG, etc., remember that there may be multiple species with exactly same molecular weight due to structural isomerism. For example, Fig. 4a shows the four peaks at around 11.3, 11.8, 12.3, and 12.7 min in the extracted chromatogram of $m/z = 810.61+$, representing different structural isomers of PC(38:4) under the chromatographic conditions used here. In such case, be careful to compare only corresponding peaks to calculate CV. Achieving good chromatographic separation is essential along with retention time reproducibility across the chromatogram to successfully implement the method described here. In case significant drift in retention time is observed, retention time locking can be helpful.
19. These numbers should be based on the retention time consistency and mass error as found in earlier steps and baseline noise, i.e., data quality. Windows narrower than those observed in the experiment will lead to artifacts such as peak splitting during deconvolution; windows too broad will lead to merging of closely eluting peaks during processing. Since lipid chromatogram often has very closely eluting peaks and masses of constituents often differing by two units (such as due to increasing or decreasing unsaturation in fatty acid side chain), these values should be chosen carefully. Otherwise, it might lead to artifacts during deisotoping. This makes data preprocessing very important before proceeding to multivariate analysis for identification of features of interest that might be enriched or depleted in cancer cell line or tumor. Inspect the baseline of the blank and standard mix samples to determine an initial cutoff for ion counts. Optimize it by minimizing spurious features (with no bona fide elution profile across samples and missing values) in the dataset by repeated extraction of features at various cutoffs above and below the initial value.
20. Changes in instrument response over a long run may lead to such results. The quality control steps help to detect such

events. However, use of very narrow mass and retention time window or very low ion count cutoff in MarkerLynx can lead to peak splitting or falsely introducing noise into the dataset, counting them as real signals.

21. This may be caused by poor sample stability during a long run as well as points mentioned in **Note 19** above. Quality control steps using pooled samples are designed to capture such incidents. Inspect for any turbidity or precipitation in the sample in such case. Although very rare, this situation may also arise if the CVs of metabolic features (biological) are indeed comparable to their respective experimental CVs. In such case, check for any confounding factors, rethink if the cohort is suitable for analysis of differential signatures associated with cancer, and increase the sample number until the samples are significantly more scattered in the PCA scores scatter plot than the pooled samples.
22. In sum normalization, ion counts for each feature are divided by the total ion count for that sample. Most of the statistical analysis and model building that are performed assume a Gaussian (Bell-shaped distribution) of the data. However, owing to a huge variation in the relative abundance of individual metabolites within and across samples as well as due to a limited sample size (in most experiments), the data distribution is often skewed. Log transformation allows for the data distribution to appear more Gaussian and allows for further statistical analysis that assumes such parametric distributions.
23. This cutoff value depends on the data quality. Check correlation coefficients of known ion clusters originating from the same molecule (e.g., LPC(19:0) and ions related to it as shown in Fig. 5). The correlation coefficient for such ion clusters found in standard mixture or pooled samples can be used to decide the cutoff value for the correlation coefficient. It should also be noted that a correlation between features may also arise if they belong to biologically connected pathways.
24. Please note that there may be lipid molecules with retention times very close to one another and structurally related. Unless you are confident about an ion being a common adduct (Na^+, K^+, NH_4^+) or fragment of another ion, please refrain from excluding them. For example, if there are two ions with difference in *m/z* value of 18 and the lower *m/z* value (say, M) represents a molecule with unsaturated fatty acid side chain, the ion with higher *m/z* value (M + 18) may either be a simple non-covalent water adduct or be a different molecule with the side chain having one unsaturation less than M along with an additional hydroxyl group. Be particularly careful about excluding such related ions, if the retention time shift is

significant over the run. Identity of such moiety can eventually be confirmed using authentic standard and/or MS/MS.

25. R^2 and Q^2 values should, respectively, be at least 0.5 and 0.4. High R2 with poor Q2 value indicates over-fitting of the model. Any model with negative Q^2 value should not be used to identify features contributing to class prediction.
26. Upon analysis of mass error of authentic standards and endogenous compounds in **steps 4** and **6** of Subheading 3.5.2, find the highest error observed in these QC samples during the experiment. The mass error window for the database search should be either comparable or higher than the largest mass error observed.
27. Under ESIMS conditions described in this chapter, doubly or triply charged ions are generally not seen, so they should be excluded from query unless spacing between isotopic peaks indicates otherwise. In addition, adducts like CH_3COOH and MeOH could also be excluded since under the aforementioned experimental conditions they are unlikely to be present.
28. Some fragments may be very labile and may not produce enough intensity of characteristic fragments upon ramping the collision voltage. In such case, stepwise increase collision energy to catch labile fragments.
29. In principle, it is possible to have more candidates than those 16 that was returned by LIPIDMAPS search for $m/z = 782.57$ – due to regio- and stereoisomerism involving double bonds on fatty acid side chains. MS/MS alone is often insufficient to distinguish between such regio- and stereoisomers and to identify position of double bonds on carbon chain. For such detailed analysis, cation doping, MS^n analysis, and ion mobility mass spectrometry can be useful [48–53].
30. Authentic standards are sometime unavailable or expensive. In cases where the fragmentation pattern does not give a clear clue and leaves too many possibilities, some of them can be excluded based on an informed guess about the retention time. This depends on the lipid class as well as the structure. It helps to run representative compounds for each lipid class so that the approximate retention time window under experimental conditions can be estimated. However, this takes significant expertise on chromatographic behavior of molecules, and exclusions should be conservative.
31. If considerable shift is observed, use retention time locking.
32. Samples present several bona fide peaks other than those matching with C4–C24 standards. NIST Library can be used to identify plausible candidates (match score > 800) followed

by confirmation by matching retention time and fragmentation pattern using authentic standards.

33. Check extracted chromatograms and make sure ions are co-eluting. Analyze the spectra to choose one of the most abundant ions as the quantifier. Use the molecular ion (M^+) as the quantifier whenever possible or, at least, as one of the qualifier ions. The level of uncertainty in the ratio of the quantifier and qualifier ions for extraction of the peak of interest should be less than half of the actual abundance ratio of the ion pair. For example, if the abundance of qualifier 1 and 2 is 40% and 20%, respectively, the uncertainty level for qualifier 1 and 2 should be <20% and <10%, respectively. Examine linearity, dynamic range, and the limit of quantitation to identify best pairs for quantitation.

Acknowledgment

Authors would like to sincerely acknowledge the contribution of Mr. Kristopher W. Krausz in developing these methods and Dr. Frank J. Gonzalez (Laboratory of Metabolism, Center for Cancer Research, National Cancer Institute, Bethesda, USA) for his encouragement and support. This work was supported by Saha Institute of Nuclear Physics, Kolkata, India.

References

1. Fahy E, Cotter D, Sud M, Subramaniam S (2011) Lipid classification, structures and tools. Biochim Biophys Acta 1811 (11):637–647. https://doi.org/10.1016/j.bbalip.2011.06.009
2. Nomura DK, Cravatt BF (2013) Lipid metabolism in cancer. Biochim Biophys Acta 1831 (10):1497–1498. https://doi.org/10.1016/j.bbalip.2013.08.001
3. Beloribi-Djefaflia S, Vasseur S, Guillaumond F (2016) Lipid metabolic reprogramming in cancer cells. Oncogene 5:e189. https://doi.org/10.1038/oncsis.2015.49
4. Carracedo A, Cantley LC, Pandolfi PP (2013) Cancer metabolism: fatty acid oxidation in the limelight. Nat Rev Cancer 13(4):227–232. https://doi.org/10.1038/nrc3483
5. Liu Y (2006) Fatty acid oxidation is a dominant bioenergetic pathway in prostate cancer. Prostate Cancer Prostatic Dis 9(3):230–234. https://doi.org/10.1038/sj.pcan.4500879
6. Mashima T, Seimiya H, Tsuruo T (2009) De novo fatty-acid synthesis and related pathways as molecular targets for cancer therapy. Br J Cancer 100(9):1369–1372. https://doi.org/10.1038/sj.bjc.6605007
7. Kuhajda FP (2006) Fatty acid synthase and cancer: new application of an old pathway. Cancer Res 66(12):5977–5980. https://doi.org/10.1158/0008-5472.CAN-05-4673
8. Baenke F, Peck B, Miess H, Schulze A (2013) Hooked on fat: the role of lipid synthesis in cancer metabolism and tumour development. Dis Model Mech 6(6):1353–1363. https://doi.org/10.1242/dmm.011338
9. Lue HW, Podolak J, Kolahi K, Cheng L, Rao S, Garg D, Xue CH, Rantala JK, Tyner JW, Thornburg KL, Martinez-Acevedo A, Liu JJ, Amling CL, Truillet C, Louie SM, Anderson KE, Evans MJ, O'Donnell VB, Nomura DK, Drake JM, Ritz A, Thomas GV (2017) Metabolic reprogramming ensures cancer cell survival despite oncogenic signaling blockade. Genes Dev 31(20):2067–2084. https://doi.org/10.1101/gad.305292.117
10. Pascual G, Avgustinova A, Mejetta S, Martin M, Castellanos A, Attolini CS, Berenguer A, Prats N, Toll A, Hueto JA,

Bescos C, Di Croce L, Benitah SA (2017) Targeting metastasis-initiating cells through the fatty acid receptor CD36. Nature 541 (7635):41–45. https://doi.org/10.1038/nature20791

11. Sun P, Xia S, Lal B, Shi X, Yang KS, Watkins PA, Laterra J (2014) Lipid metabolism enzyme ACSVL3 supports glioblastoma stem cell maintenance and tumorigenicity. BMC Cancer 14:401. https://doi.org/10.1186/1471-2407-14-401
12. Yasumoto Y, Miyazaki H, Vaidyan LK, Kagawa Y, Ebrahimi M, Yamamoto Y, Ogata M, Katsuyama Y, Sadahiro H, Suzuki M, Owada Y (2016) Inhibition of fatty acid synthase decreases expression of stemness markers in glioma stem cells. PLoS One 11(1): e0147717. https://doi.org/10.1371/journal.pone.0147717
13. Darekar SD, Mushtaq M, Gurrapu S, Kovalevska L, Drummond C, Petruchek M, Tirinato L, Di Fabrizio E, Carbone E, Kashuba E (2015) Mitochondrial ribosomal protein S18-2 evokes chromosomal instability and transforms primary rat skin fibroblasts. Oncotarget 6(25):21016–21028. https://doi.org/10.18632/oncotarget.4123
14. Mitra R, Chao O, Urasaki Y, Goodman OB, Le TT (2012) Detection of lipid-rich prostate circulating tumour cells with coherent anti-Stokes Raman scattering microscopy. BMC Cancer 12:540. https://doi.org/10.1186/1471-2407-12-540
15. Dill AL, Ifa DR, Manicke NE, Costa AB, Ramos-Vara JA, Knapp DW, Cooks RG (2009) Lipid profiles of canine invasive transitional cell carcinoma of the urinary bladder and adjacent normal tissue by desorption electrospray ionization imaging mass spectrometry. Anal Chem 81(21):8758–8764. https://doi.org/10.1021/ac901028b
16. Eberlin LS, Norton I, Dill AL, Golby AJ, Ligon KL, Santagata S, Cooks RG, Agar NY (2012) Classifying human brain tumors by lipid imaging with mass spectrometry. Cancer Res 72(3):645–654. https://doi.org/10.1158/0008-5472.CAN-11-2465
17. Hilvo M, Denkert C, Lehtinen L, Muller B, Brockmoller S, Seppanen-Laakso T, Budczies J, Bucher E, Yetukuri L, Castillo S, Berg E, Nygren H, Sysi-Aho M, Griffin JL, Fiehn O, Loibl S, Richter-Ehrenstein C, Radke C, Hyotylainen T, Kallioniemi O, Iljin K, Oresic M (2011) Novel theranostic opportunities offered by characterization of altered membrane lipid metabolism in breast cancer progression. Cancer Res 71 (9):3236–3245. https://doi.org/10.1158/0008-5472.CAN-10-3894
18. Patterson AD, Maurhofer O, Beyoglu D, Lanz C, Krausz KW, Pabst T, Gonzalez FJ, Dufour JF, Idle JR (2011) Aberrant lipid metabolism in hepatocellular carcinoma revealed by plasma metabolomics and lipid profiling. Cancer Res 71(21):6590–6600. https://doi.org/10.1158/0008-5472.CAN-11-0885
19. Marien E, Meister M, Muley T, Fieuws S, Bordel S, Derua R, Spraggins J, Van de Plas R, Dehairs J, Wouters J, Bagadi M, Dienemann H, Thomas M, Schnabel PA, Caprioli RM, Waelkens E, Swinnen JV (2015) Non-small cell lung cancer is characterized by dramatic changes in phospholipid profiles. Int J Cancer 137(7):1539–1548. https://doi.org/10.1002/ijc.29517
20. Cifkova E, Holcapek M, Lisa M, Vrana D, Melichar B, Student V (2015) Lipidomic differentiation between human kidney tumors and surrounding normal tissues using HILIC-HPLC/ESI-MS and multivariate data analysis. J Chromatogr B Analyt Technol Biomed Life Sci 1000:14–21. https://doi.org/10.1016/j.jchromb.2015.07.011
21. Liu Y, Chen Y, Momin A, Shaner R, Wang E, Bowen NJ, Matyunina LV, Walker LD, McDonald JF, Sullards MC, Merrill AH Jr (2010) Elevation of sulfatides in ovarian cancer: an integrated transcriptomic and lipidomic analysis including tissue-imaging mass spectrometry. Mol Cancer 9:186. https://doi.org/10.1186/1476-4598-9-186
22. Li J, Ren S, Piao HL, Wang F, Yin P, Xu C, Lu X, Ye G, Shao Y, Yan M, Zhao X, Sun Y, Xu G (2016) Integration of lipidomics and transcriptomics unravels aberrant lipid metabolism and defines cholesteryl oleate as potential biomarker of prostate cancer. Sci Rep 6:20984. https://doi.org/10.1038/srep20984
23. Patterson NH, Alabdulkarim B, Lazaris A, Thomas A, Marcinkiewicz MM, Gao ZH, Vermeulen PB, Chaurand P, Metrakos P (2016) Assessment of pathological response to therapy using lipid mass spectrometry imaging. Sci Rep 6:36814. https://doi.org/10.1038/srep36814
24. Del Boccio P, Perrotti F, Rossi C, Cicalini I, Di Santo S, Zucchelli M, Sacchetta P, Genovesi D, Pieragostino D (2017) Serum lipidomic study reveals potential early biomarkers for predicting response to chemoradiation therapy in advanced rectal cancer: a pilot study. Adv Radiat Oncol 2(2):118–124. https://doi.org/10.1016/j.adro.2016.12.005

25. Ma S, Jiang B, Deng W, Gu ZK, Wu FZ, Li T, Xia Y, Yang H, Ye D, Xiong Y, Guan KL (2015) D-2-hydroxyglutarate is essential for maintaining oncogenic property of mutant IDH-containing cancer cells but dispensable for cell growth. Oncotarget 6 (11):8606–8620. https://doi.org/10.18632/oncotarget.3330
26. Skotland T, Ekroos K, Kauhanen D, Simolin H, Seierstad T, Berge V, Sandvig K, Llorente A (2017) Molecular lipid species in urinary exosomes as potential prostate cancer biomarkers. Eur J Cancer 70:122–132. https://doi.org/10.1016/j.ejca.2016.10.011
27. Armitage EG, Southam AD (2016) Monitoring cancer prognosis, diagnosis and treatment efficacy using metabolomics and lipidomics. Metabolomics 12:146. https://doi.org/10.1007/s11306-016-1093-7
28. Chen X, Chen H, Dai M, Ai J, Li Y, Mahon B, Dai S, Deng Y (2016) Plasma lipidomics profiling identified lipid biomarkers in distinguishing early-stage breast cancer from benign lesions. Oncotarget 7(24):36622–36631. https://doi.org/10.18632/oncotarget.9124
29. Zhou X, Mao J, Ai J, Deng Y, Roth MR, Pound C, Henegar J, Welti R, Bigler SA (2012) Identification of plasma lipid biomarkers for prostate cancer by lipidomics and bioinformatics. PLoS One 7(11):e48889. https://doi.org/10.1371/journal.pone.0048889
30. Cajka T, Fiehn O (2014) Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. Trends Anal Chem 61:192–206. https://doi.org/10.1016/j.trac.2014.04.017
31. Chen Y, Singh S, Matsumoto A, Manna SK, Abdelmegeed MA, Golla S, Murphy RC, Dong H, Song BJ, Gonzalez FJ, Thompson DC, Vasiliou V (2016) Chronic glutathione depletion confers protection against alcohol-induced steatosis: implication for redox activation of amp-activated protein kinase pathway. Sci Rep 6:29743. https://doi.org/10.1038/srep29743
32. Fahy E, Subramaniam S, Brown HA, Glass CK, Merrill AH Jr, Murphy RC, Raetz CR, Russell DW, Seyama Y, Shaw W, Shimizu T, Spener F, van Meer G, VanNieuwenhze MS, White SH, Witztum JL, Dennis EA (2005) A comprehensive classification system for lipids. J Lipid Res 46(5):839–861. https://doi.org/10.1194/jlr.E400004-JLR200
33. Fahy E, Subramaniam S, Murphy RC, Nishijima M, Raetz CR, Shimizu T, Spener F, van Meer G, Wakelam MJ, Dennis EA (2009) Update of the LIPID MAPS comprehensive classification system for lipids. J Lipid Res 50 (Suppl):S9–S14. https://doi.org/10.1194/jlr.R800095-JLR200
34. Kind T, Liu KH, Lee DY, DeFelice B, Meissen JK, Fiehn O (2013) LipidBlast in silico tandem mass spectrometry database for lipid identification. Nat Methods 10(8):755–758. https://doi.org/10.1038/nmeth.2551
35. Tautenhahn R, Cho K, Uritboonthai W, Zhu Z, Patti GJ, Siuzdak G (2012) An accelerated workflow for untargeted metabolomics using the METLIN database. Nat Biotechnol 30(9):826–828. https://doi.org/10.1038/nbt.2348
36. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vazquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, Sayeeda Z, Lo E, Assempour N, Berjanskii M, Singhal S, Arndt D, Liang Y, Badran H, Grant J, Serra-Cayuela A, Liu Y, Mandal R, Neveu V, Pon A, Knox C, Wilson M, Manach C, Scalbert A (2018) HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res 46(D1): D608–D617. https://doi.org/10.1093/nar/gkx1089
37. Xia J, Wishart DS (2016) Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. Curr Protoc Bioinformatics 55:14.10.11–14.10.91. https://doi.org/10.1002/cpbi.11
38. Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Can J Biochem Physiol 37(8):911–917. https://doi.org/10.1139/o59-099
39. Schwudke D, Schuhmann K, Herzog R, Bornstein SR, Shevchenko A (2011) Shotgun lipidomics on high resolution mass spectrometers. Cold Spring Harb Perspect Biol 3(9):a004614. https://doi.org/10.1101/cshperspect.a004614
40. Hsu FF, Turk J (2009) Electrospray ionization with low-energy collisionally activated dissociation tandem mass spectrometry of glycerophospholipids: mechanisms of fragmentation and structural characterization. J Chromatogr B Analyt Technol Biomed Life Sci 877 (26):2673–2695. https://doi.org/10.1016/j.jchromb.2009.02.033
41. Ferreira CR, Saraiva SA, Catharino RR, Garcia JS, Gozzo FC, Sanvido GB, Santos LF, Lo Turco EG, Pontes JH, Basso AC, Bertolla RP, Sartori R, Guardieiro MM, Perecin F, Meirelles FV, Sangalli JR, Eberlin MN (2010) Single embryo and oocyte lipid fingerprinting by mass spectrometry. J Lipid Res 51 (5):1218–1227. https://doi.org/10.1194/jlr.D001768
42. Bowden JA, Shao F, Albert CJ, Lally JW, Brown RJ, Procknow JD, Stephenson AH, Ford DA

(2011) Electrospray ionization tandem mass spectrometry of sodiated adducts of cholesteryl esters. Lipids 46(12):1169–1179. https://doi.org/10.1007/s11745-011-3609-2

43. Houjou T, Yamatani K, Nakanishi H, Imagawa M, Shimizu T, Taguchi R (2004) Rapid and selective identification of molecular species in phosphatidylcholine and sphingomyelin by conditional neutral loss scanning and MS3. Rapid Commun Mass Spectrom 18 (24):3123–3130. https://doi.org/10.1002/rcm.1737
44. Sugiura Y, Konishi Y, Zaima N, Kajihara S, Nakanishi H, Taguchi R, Setou M (2009) Visualization of the cell-selective distribution of PUFA-containing phosphatidylcholines in mouse brain by imaging mass spectrometry. J Lipid Res 50(9):1776–1788. https://doi.org/10.1194/jlr.M900047-JLR200
45. Lemaire R, Wisztorski M, Desmons A, Tabet JC, Day R, Salzet M, Fournier I (2006) MALDI-MS direct tissue analysis of proteins: Improving signal sensitivity using organic treatments. Anal Chem 78(20):7145–7153. https://doi.org/10.1021/ac060565z
46. Schwartz SA, Reyzer ML, Caprioli RM (2003) Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. J Mass Spectrom 38(7):699–708. https://doi.org/10.1002/jms.505
47. Dettmer K, Nurnberger N, Kaspar H, Gruber MA, Almstetter MF, Oefner PJ (2011) Metabolite extraction from adherently growing mammalian cells for metabolomics studies: optimization of harvesting and extraction protocols. Anal Bioanal Chem 399 (3):1127–1139. https://doi.org/10.1007/s00216-010-4425-x
48. Thomas MC, Mitchell TW, Harman DG, Deeley JM, Murphy RC, Blanksby SJ (2007) Elucidation of double bond position in unsaturated lipids by ozone electrospray ionization mass spectrometry. Anal Chem 79(13):5013–5022. https://doi.org/10.1021/ac0702185
49. Groessl M, Graf S, Knochenmuss R (2015) High resolution ion mobility-mass spectrometry for separation and identification of isomeric lipids. Analyst 140(20):6904–6911. https://doi.org/10.1039/c5an00838g
50. Kyle JE, Zhang X, Weitz KK, Monroe ME, Ibrahim YM, Moore RJ, Cha J, Sun X, Lovelace ES, Wagoner J, Polyak SJ, Metz TO, Dey SK, Smith RD, Burnum-Johnson KE, Baker ES (2016) Uncovering biologically significant lipid isomers with liquid chromatography, ion mobility spectrometry and mass spectrometry. Analyst 141(5):1649–1659. https://doi.org/10.1039/c5an02062j
51. Williams PE, Klein DR, Greer SM, Brodbelt JS (2017) Pinpointing double bond and sn-positions in glycerophospholipids via hybrid 193 nm ultraviolet photodissociation (UVPD) mass spectrometry. J Am Chem Soc 139 (44):15681–15690. https://doi.org/10.1021/jacs.7b06416
52. Castro-Perez J, Roddy TP, Nibbering NM, Shah V, McLaren DG, Previs S, Attygalle AB, Herath K, Chen Z, Wang SP, Mitnaul L, Hubbard BK, Vreeken RJ, Johns DG, Hankemeier T (2011) Localization of fatty acyl and double bond positions in phosphatidylcholines using a dual stage CID fragmentation coupled with ion mobility mass spectrometry. J Am Soc Mass Spectrom 22(9):1552–1567. https://doi.org/10.1007/s13361-011-0172-2
53. Hsu FF, Turk J (2010) Electrospray ionization multiple-stage linear ion-trap mass spectrometry for structural elucidation of triacylglycerols: assignment of fatty acyl groups on the glycerol backbone and location of double bonds. J Am Soc Mass Spectrom 21(4):657–669. https://doi.org/10.1016/j.jasms.2010.01.007

Chapter 12

Mass Spectrometry-Based Profiling of Metabolites in Human Biofluids

Tanushree Chakraborty and Soumen Kanti Manna

Abstract

Cancer poses a daunting challenge to researchers and clinicians alike. Early diagnosis, accurate prognosis, and prediction of therapeutic response remain elusive in most types of cancer. In addition, lacunae in our understanding of cancer biology continue to hinder advancement of therapeutic strategies. Metabolic reprogramming has been identified as integral to pathogenesis and progression of the disease. Consequently, analysis of biofluid metabolome has emerged as a promising approach to further our understanding of disease biology as well as to identify cancer biomarkers. However, unbiased identification of robust and meaningful differences in metabolic signatures remains a non-trivial task. This chapter describes a generalized strategy for global metabolic profiling of human biofluids using ultra-performance liquid chromatography (UPLC) and mass spectrometry, which together offer a sensitive, high-throughput, and versatile platform. A step-by-step protocol for performing untargeted metabolic profiling of urine and serum (or plasma), using hydrophilic interaction liquid chromatography (HILIC) or reverse-phase (RP) chromatography coupled with electrospray ionization mass spectrometry (ESI-MS) to multivariate data analysis and identification of metabolites of interest has been detailed.

Key words Urine, Serum, Plasma, Metabolomics, Untargeted profiling, HILIC, RP, UPLC-ESI-MS

1 Introduction

Initially elucidated in the phenomenon of aerobic glycolysis (Warburg effect) [1], metabolic reprogramming has been established as a hallmark of cancer. However, the role of metabolic reprogramming has expanded far beyond accelerated energy production from glucose in cancer cells. Emerging evidences suggest that metabolic reprogramming is essential to channel nutrient-derived carbon and nitrogen into de novo synthesis of nucleic acids, amino acids, amino sugars, nucleotide sugars, as well as fatty acids, which are building blocks for different macromolecules required for proliferating cells [2–5]. In addition, it has also been well-established that epigenetic

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_12,

events, such as methylation and acetylation, essential for changes in gene expression, are intricately connected to metabolic processes [6, 7]. In fact, some metabolites have been shown to directly promote the oncogenic property of cells. Consequently, metabolites like 2-hydroxyglutaric acid [8, 9], fumaric acid [10, 11], and succinic acid [12, 13] have been dubbed as 'oncometabolites'. Several studies have shown that metabolic signatures of tumor tissue [14–17] as well as biofluids, such as plasma [18, 19], serum [20–24], urine [25–30], saliva [31], bile [32], and cerebrospinal fluid [33] from cancer patients, are distinct from those from healthy individuals. These bear the promise of utility of metabolic profiling of human biofluids in providing useful minimally-invasive alternative for screening, diagnosis, and monitoring of therapeutic response in cancer patients [34, 35], subject to validation in multiple independent cohorts and establishment of mechanistic association between putative biomarkers and pathogenesis.

However, very few studies have so far established unequivocal mechanistic association between changes in a single or a set of metabolites in biofluids and cancer in humans. Some of the examples are metabolites related to specific genes that contribute to pathogenesis of the disease, such as 2-hydroxyglutaric acid, which was identified as a product of IDH1 mutants in glioblastoma and subsequently other cancers [9, 36] or fumaric acid, which accumulates as a consequence of mutation in fumarate hydratase genes [37]. This is not only because most cancers involve multiple genes, gene products, and pathways but also the fact that the human biofluid metabolome is very diverse and influenced by multiple extrinsic factors, in addition to genetic architecture; including diet, lifestyle, habitat, environmental exposure and gut flora, which are difficult to control. This essentially warrants untargeted metabolic profiling to be an integral step for identification of cancer-associated signatures.

Mass spectrometry coupled with chromatography and nuclear magnetic resonance (NMR) spectroscopy are two most widely used techniques for global metabolic profiling. However, it should be noted that biofluid may contain metabolites with concentrations varying over ten orders of magnitude and the metabolite(s) of interest may be of very low abundance. Although NMR gives quite unequivocal structural identification for many molecules and offers robust quantitation, it is relatively insensitive and, therefore, of limited use as a global profiling-based discovery tool aimed at identification of orthogonal signatures of interest. In addition, NMR cannot distinguish between molecules with similar and repetitive motifs such as lipid molecules. Thus, superior sensitivity and structural information make mass spectrometry the preferred tool for global profiling of human biofluids. Mass spectrometry is typically coupled with gas chromatography or liquid chromatography

to help in separation, ionization, identification, and quantitation of metabolites. Compared to gas chromatography, liquid chromatography not only allows analysis of biofluid metabolome with minimal sample preparation but also offers a higher throughput. Throughput is important from the viewpoint of discovery as well as translation. Recent advances in chromatography, particularly, development of ultra-performance liquid chromatography (UPLC), have significantly helped to further increase the throughput as well as sensitivity and reproducibility, making it a powerful tool for unbiased discovery of cancer-associated signatures in biofluids. It should be noted that biofluids may contain metabolites that are highly polar (such as amino acids, nucleosides) as well as non-polar (such as lipids), which behave quite differently in terms of their chromatographic properties. Therefore, there is no one--size-fits-all method for untargeted metabolic profiling of biofluids using chromatography and mass spectrometry.

As mentioned above, several studies have reported targeted as well as untargeted metabolic profiling of a variety of human biofluid samples in the context of cancer. Targeted metabolic profiling is focused and, therefore, relatively convenient to execute. Untargeted metabolic profiling, on the other hand, does not presume an association between any metabolite or metabolic pathway and pathology of interest. Thus, it offers an opportunity to identify novel features that may not only lead to orthogonal signatures for a particular type or subtype of cancer but also further our understanding of cancer biology. This was successfully demonstrated by our previous works in the context of lung cancer and colorectal cancer. While the lung cancer study identified a novel metabolite D-ribofuranosylcreatine to be a potential noninvasive biomarker for lung cancer [28], the colorectal cancer study revealed novel noninvasive metabolic signatures of aberrant methylation [38].

Therefore, the aim of this protocol is to provide a method and guidelines for global profiling of both polar and non-polar metabolites in commonly sampled biofluids such as urine, plasma, or serum using ultra-performance liquid chromatography (UPLC) coupled with electrospray ionization mass spectrometry (ESIMS). Although, the method described hereneath is for the aforementioned biofluids, it can be adopted for other biofluids, such as saliva, bile, and cerebrospinal fluid. This article describes details of experimental setup and generalized methodology for data analysis to identify signatures of interest along with recommendations and advice from our personal experience. Special emphasis has been given to quality control and data preprocessing to reduce chances of spurious associations.

2 Materials

2.1 Solvents

1. Water (UHPLC-MS grade).
2. Methanol (UHPLC-MS grade).
3. Acetonitrile (UHPLC-MS grade).
4. Isopropanol (LC-MS grade).

2.2 Chemicals

1. Chlorpropamide.
2. Difluoromethylornithine (DFMO) hydrochloride hydrate.
3. α-Aminopimelic acid.
4. 4-Nitrobenzoic acid.
5. Debrisoquine sulfate salt.
6. Ammonium acetate.
7. Ammonium hydroxide.
8. Acetic acid.
9. Sulfadimethoxine.
10. Sodium hydroxide.
11. Formic acid.
12. Authentic standards for metabolites of interest.

2.3 Stock and Working Solutions

1. Lockspray solution: 250 pg/μL sulfadimethoxine in 50% aqueous acetonitrile.
2. Chlorpropamide stock: 5 mM chlorpropamide in methanol.
3. α-Aminopimelic acid stock: 5 mM aqueous α-aminopimelic acid.
4. Difluoromethyornithine stock: 5 mM aqueous difluoromethylornithine.
5. Debrisoquine stock: 2 mM aqueous debrisoquine.
6. 4-Nitrobenzoic acid stock: 10 mM 4-nitrobenzoic acid in methanol.
7. A solution containing 1 μM debrisoquine and 5 μM difluoromethylornithine was prepared either in water for reverse-phase analysis or in acetonitrile/water/methanol solvent mixture (65:30:5) for hydrophilic interaction liquid chromatography.
8. Authentic standard cocktail: 20 μM of each authentic standards comprising of representative metabolites present in urine (*see* **Note 1**) and serum in 40% aqueous acetonitrile (for reverse-phase analysis) or in acetonitrile/water/methanol solvent mixture (65:30:5) (for hydrophilic interaction liquid chromatography).

2.4 Diluents

1. Diluent A: 5 μM chlorpropamide in 50% aqueous acetonitrile.
2. Diluent B: 10 μM chlorpropamide in 70% aqueous acetonitrile solution.
3. Diluent C: 5 μM chlorpropamide and 10 μM α-aminopimelic acid in acetonitrile/water/methanol (60:35:5).

2.5 Solvents for Chromatography

2.5.1 RP

1. Eluent A: 0.1% aqueous formic acid degassed by sonication (*see* **Note 2**).
2. Eluent B: 0.1% formic in acetonitrile degassed by sonication.

2.5.2 HILIC

1. Eluent A: 10 mM ammonium acetate in 90% acetonitrile degassed by sonication.
2. Eluent B: 10 mM ammonium acetate in 10% acetonitrile degassed by sonication.

2.6 Instruments

1. Pipettements (1–10 μL, 10–100 μL, 20–200 μL, 100–1000 μL).
2. Refrigerated centrifuge (capable of 18,000 × *g* at 4 °C).
3. Vortex.
4. Xevo G2 QTof coupled with a Acquity™ UPLC platform (Waters Corp, Milford, MA) (*see* **Note 3**).

2.7 Chromatography Columns

1. 50 mm Acquity UPLC BEH C18 (particle size 1.7 μm, i.d. 2.1 mm) for reverse phase (RP) (*see* **Note 3**).
2. 50 mm Acquity UPLC BEH amide column (particle size 1.7 μm, i.d. 2.1 mm) for hydrophilic interaction liquid chromatography (HILIC) (*see* **Note 3**).

2.8 Gasses

1. Nitrogen (purity >99.999%).
2. Argon (purity >99.999%).

2.9 Software

1. Microsoft Excel.
2. MassLynx package including MarkerLynx and TargetLynx.
3. SIMCA-P12+ (Umetrics, Kinnelon, NJ).
4. Prism version 6 (GraphPad).

2.10 Online Databases

1. METLIN: https://metlin.scripps.edu/ [39].
2. HMDB: http://www.hmdb.ca/ [40].
3. MassTRIX: http://masstrix3.helmholtz-muenchen.de/masstrix3/ [41].
4. MetaboAnalyst: http://www.metaboanalyst.ca/ [42].

3 Methods

3.1 Clinical Samples

Collect first-pass midstream urine samples from empty-stomach subjects. Fasting blood samples should be collected. It should be allowed to clot standing at room temperature for 30 min and spun at 2000 × *g* for 10 min at 4 °C to collect the serum. For plasma, blood samples are collected in tubes containing anticoagulants (preferably, lithium heparin) followed by centrifugation at 4 °C to separate cells and collect plasma. All samples are stored in −80 °C until further use (*see* **Note 4**). Since plasma and serum samples are processed in similar manner in this protocol, in the following section, only serum is mentioned for convenience.

3.2 Metadata

While samples should be identified by anonymous bar codes or sample identification number, details of age, gender, BMI, ethnicity, and clinical, pathological status, any ongoing treatment or medication and any other relevant lifestyle-related information, such as smoking and drinking status, etc. should be carefully collected for each sample (*see* **Note 5**).

3.3 Sample Preparation

3.3.1 Quality Control (QC) Samples

1. Blank: Use respective diluents as blank samples to investigate any carry-over.
2. Standard mix A: Prepare a mixture of exogenous compounds such as 1 μM debrisoquine, 10 μM 4-nitrobenzoic acid, and 5 μM DFMO in diluent A (for RP) or in diluent C for HILIC.
3. Standard mix B: Prepare a mixture of representative endogenous metabolites (*see* **Note 1**) present in urine and serum in diluent A (for RP) or in diluent C for HILIC.
4. Pooled sample: Take 50 μL aliquot from individual urine or 20 μL aliquot from individual serum samples to prepare a pooled urine or serum sample. This sample should be used to prepare pooled quality control samples for each run as described below.

3.3.2 Urine Samples for HILIC

1. Thaw urine samples on ice and vortex for 30 s.
2. Aliquot 200 μL urine sample into 1.5 mL snapcap tubes, and add 800 μL (*see* **Note 6**) pre-chilled (*see* **Note 7**) diluent C.
3. Vortex for 30 s.
4. Centrifuge samples at 18000 × *g* at 4 °C for 25 min.
5. Take tubes out (*see* **Note 8**), and carefully take out 800 μL of supernatant into glass sample tubes (*see* **Note 9**).

3.3.3 Urine Samples for RP

1. Thaw urine samples on ice and vortex for 30 s.
2. Aliquot 500 μL urine sample into 1.5 mL snapcap tubes and add 500 μL of pre-chilled diluent A (*see* **Note 6**).
3. Vortex for 30 s.

4. Centrifuge samples at 18000 × *g* at 4 °C for 25 min.
5. Take tubes out (*see* **Note 8**) and carefully take out 800 μL of supernatant into glass sample tubes (*see* **Note 9**).

3.3.4 Serum Samples for HILIC

1. Thaw serum samples on ice (*see* **Note 10**) and vortex for 30 s.
2. Add 50 μL of serum sample directly into 950 μL of chilled diluent B (*see* **Note 11**).
3. Vortex for 30 s.
4. Centrifuge samples at 18000 × *g* at 4 °C for 25 min.
5. Take tubes out (*see* **Note 8**), and carefully take 800 μL of supernatant into glass sample tubes (*see* **Note 9**).

3.3.5 Serum Samples for RP

1. Dispense 300 μL of UPLC grade water into glass vials for RP.
2. Take 300 μL of sample out of the 800 μL of supernatant mentioned above into corresponding glass vials.
3. Mix by pipetting up and down.

3.4 UPLC-ESI-MS Analysis

1. Turn the source voltages and gas flow on, and let it stabilize over at least 30 min.
2. Make sure the room temperature is well-maintained around 22 ± 1 °C (*see* **Note 12**).
3. Acquire the lockspray mass spectra with 250 pg/μL sulfadimethoxine in 50% aqueous acetonitrile; adjust cone and lens voltages as required to achieve desired mass accuracy, intensity, and peak shape in both positive ($m/z = 311.0814$) and negative mode ($m/z = 309.0658$) following guidelines as per operator's manual.
4. Calibrate the mass axis (50–1000 Da) using 0.1 mg/mL sodium formate in 90% isopropanol as per instructions in operator's manual.
5. Typical experimental parameters (*see* **Note 13**) for ESI-positive mode analysis of samples: choose m/z range of 50–850 Da and sensitivity mode with capillary voltage at 3.1 kV, sampling cone voltage at 40 V, source temperature at 150 °C, desolvation temperature at 650 °C, desolvation gas (N_2) at 850 L/h, and cone gas (N_2) at 50 L/h. Infuse lockspray solution continuously at 30 μL/min for real time mass correction. Define lock mass as 311.0814+ with lockspray capillary voltage at 2.3 kV.
6. Typical experimental parameters (*see* **Note 13**) for ESI-negative mode analysis of samples: choose m/z range of 50–850 Da and sensitivity mode with capillary voltage at 2.9 kV, sampling cone voltage at 30 V, source temperature at 150 °C, desolvation temperature at 550 °C, desolvation gas (N_2) at 850 L/h, and cone gas (N_2) at 50 L/h. Acquire data in

sensitivity mode. Infuse lockspray solution continuously at 30 μL/min, and define lock mass as 309.0658- with lockspray capillary voltage at 2.7 kV.

7. Prepare and save inlet methods separately for HILIC and RP mode for urine and serum analysis.
8. For RP mode separation of serum constituents, use the following gradient with BEH C18 column maintained at 60 °C:

Time segment (min)	Flow rate (mL/min)	Solvent A (%)	Solvent B (%)
0–0.5	0.5	95	5
0.5–4	0.5	95–40	5–60
4–8	0.5	40–0	60–100
8–9	0.5	0	100
9–9.2	0.5	0–95	100–5
9.2–11	0.5	95	5

Representative chromatograms of serum samples analyzed using this method in ESI positive and negative mode are shown in Fig. 1a and 1b, respectively.

9. For RP separation of urine constituents, use the following gradient with BEH C18 column maintained at 40 °C:

Time segment (min)	Flow rate (mL/min)	Solvent A (%)	Solvent B (%)
0–0.5	0.5	98	2
0.5–4	0.5	98–80	2–20
4–8	0.5	80–5	20–95
8–8.2	0.5	5–1	95–99
8.2–9.1	0.5	1	99
9.1–9.3	0.5	1–98	99–2
9.3–11	0.5	98	2

10. For HILIC mode separation, use the following gradient with BEH amide column maintained at 40 °C:

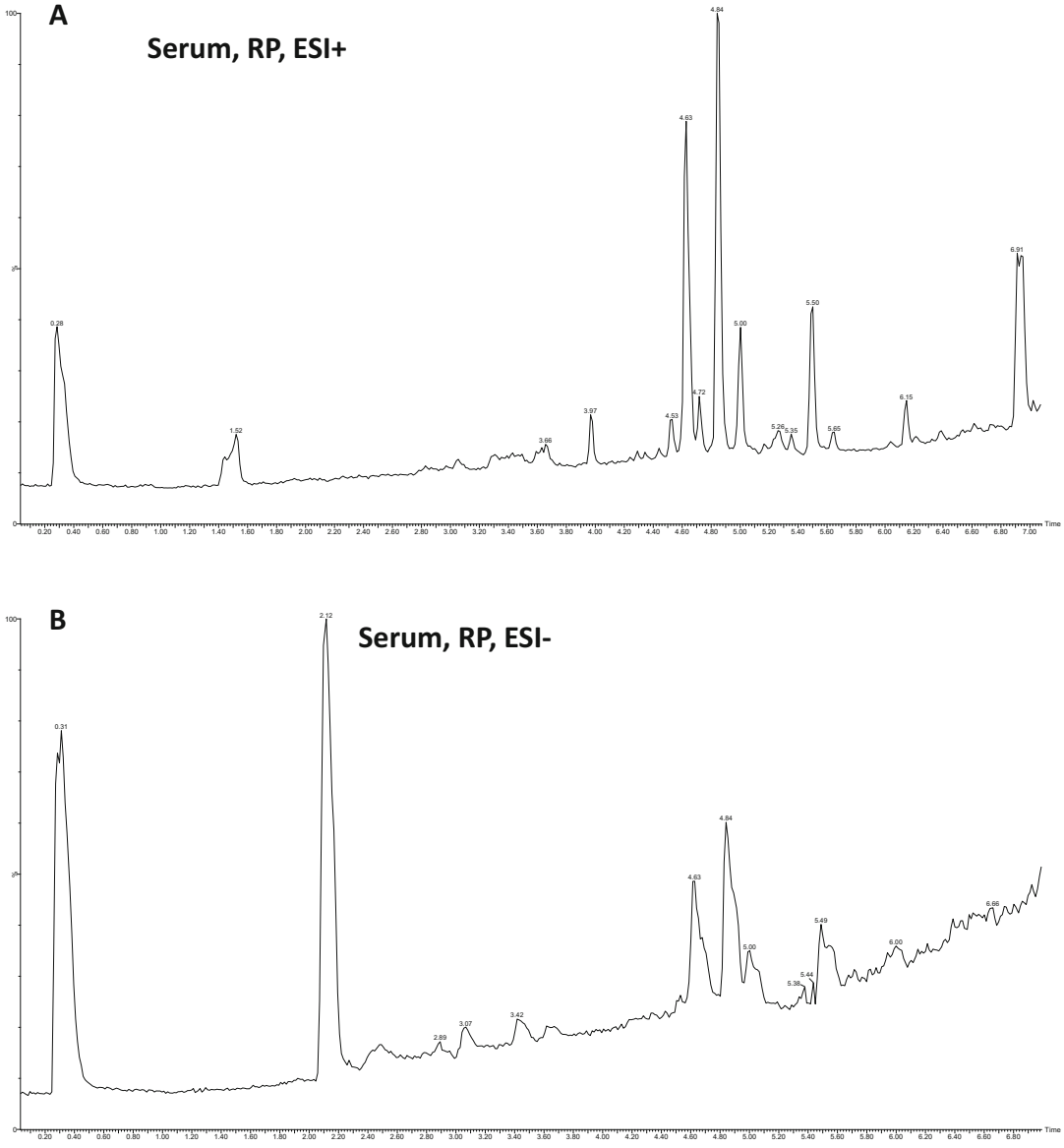


Fig. 1 Representative total ion chromatogram of serum samples analyzed by UPLC-ESIMS in (**a**) positive and (**b**) negative ionization modes with a 2.1 mm × 50 mm Acquity BEH C18 reverse-phase column (particle size 1.7 μM)

Time segment (min)	Flow rate (mL/min)	Solvent A (%)	Solvent B (%)
0–0.5	0.4	99	1
0.5–4	0.4	99–40	1–60
4–8	0.4	40–20	60–80
8–8.5	0.3	20	80
8.5–8.8	0.3	20–99	80–1
8.8–9.8	0.3	99	1
9.8–12	0.4	99	1

Representative chromatograms of urine samples analyzed using this method in ESI positive and negative mode are shown in Fig. 2a and 2b, respectively.

11. Load sample vials (or 96-well plates) into the Acquity UPLC sample chamber maintained at 12 °C and choose appropriate sample bed layout.
12. Prepare a sample list with appropriate MS and inlet methods for desired mode of analysis with 5 μL injection volume for each sample.
13. Sample list should start with couple of injections of standard mix A and B followed by at least six injections of the pooled sample before injection of individual samples (*see* **Note 14**).
14. Samples should be injected in a randomized fashion. Use random number generator function in Microsoft Excel to randomize the order of sample injections.
15. Standard mixtures and pooled samples should be injected intermittently throughout the run.

3.5 Data Inspection and Quality Control

1. Manually inspect chromatograms, check for any sudden drop or increase in counts or shift in baseline of blank, standard mix A, B, and pooled samples.
2. Check blank samples for any carry-over by looking for peaks for compounds present in the immediate preceding sample.
3. Extract these chromatograms for ions corresponding to authentic standards (standard mix A and B) and representative metabolites that are found in all samples (pooled samples) as well as internal standard (chlorpropamide in RP and α-aminopimelic acid in HILIC) to check for consistency in retention time and *m/z* values. Tabulate retention time for each peak, and determine the mean and standard deviation (*see* **Note 15**).

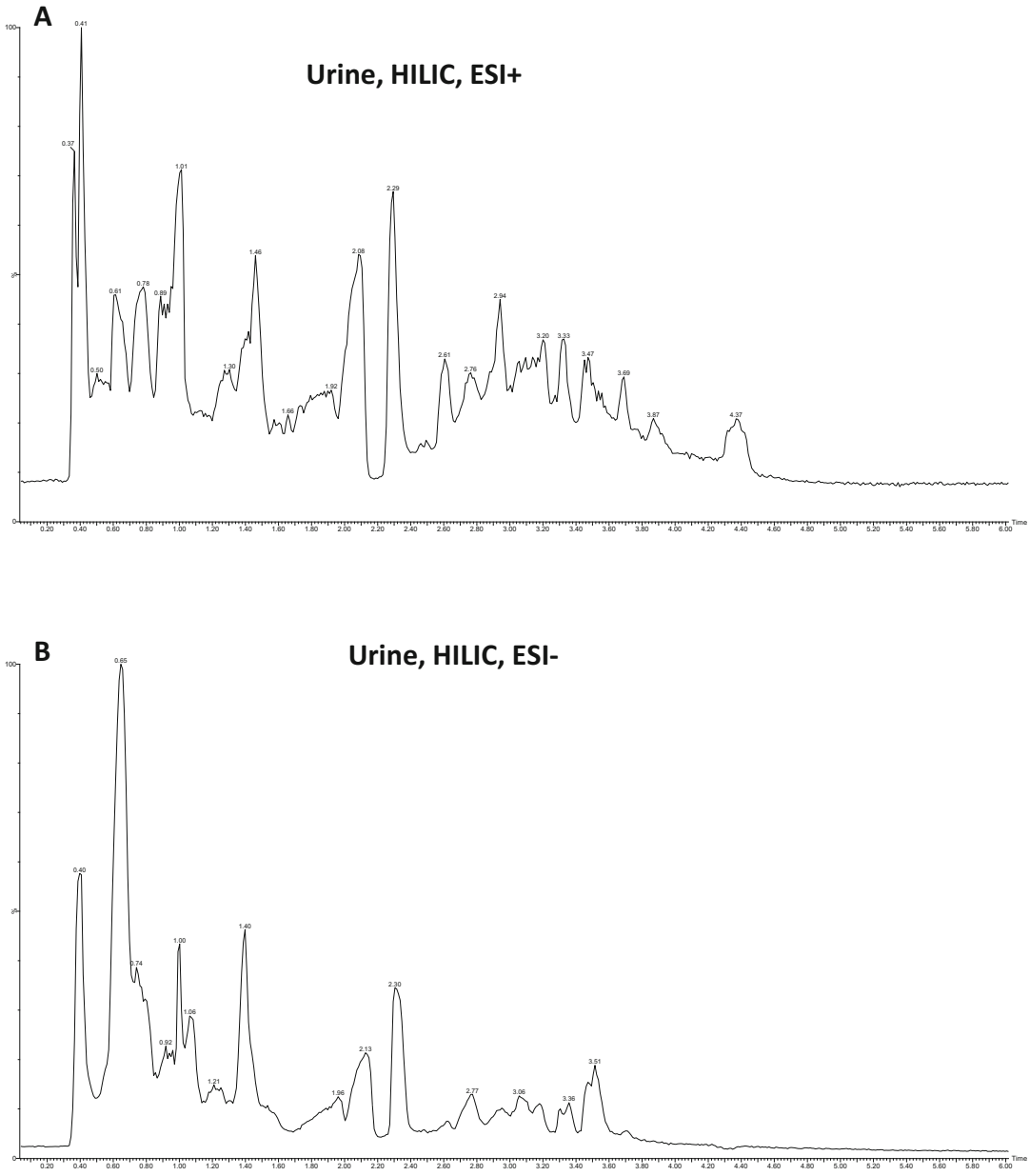


Fig. 2 Representative total ion chromatogram of urine samples analyzed by UPLC-ESIMS in (**a**) positive and (**b**) negative ionization modes with a 2.1 mm × 50 mm Acquity BEH amide hydrophilic interaction liquid chromatography (HILIC) reverse-phase column (particle size 1.7 μM)

4. Calculate the *m/z* error range (in ppm) for each of these peaks. This value should be used for feature extraction and database search in subsequent analysis (*see* Subheadings 3.10 and 3.13).
5. Integrate peaks to get areas under the curve and calculate the coefficient of variation (CV) of these representative peaks

across run (*see* **Note 16**). The CV for representative peaks should be less than 20% to proceed further with analysis.

3.6 Feature Extraction

1. Process the centroided and integrated mass spectrometric data using MarkerLynx.
2. For RP analysis of urine samples, process 0.1–6.5 min region of the chromatogram with a mass window of 0.02 Da, retention time window of 0.15 min, and intensity threshold of 200 counts, and deisotope (*see* **Note 17**).
3. For RP analysis of serum, process 0.1–7.2 min region with retention time window of 0.2 min and other parameters same as those in RP analysis of urine samples (*see* **Note 17**).
4. For HILIC mode analysis, process 0.1–5 min region of the chromatogram with retention time window of 0.3 min and other parameters same as those in RP analysis of urine samples (*see* **Note 17**).

3.7 Data Preprocessing: Exclusion and Reduction

1. Normalize intensities of individual features (*m/z*, retention time pairs) in the raw data matrix obtained after processing chromatograms using MarkerLynx with respect to total ion count (TIC) for respective samples.
2. Perform unsupervised principal components analysis (PCA) on the TIC-normalized and Pareto-scaled data matrix using SIMCA-P + 12. The blank, standard mix A, B samples should be clustered very tightly together in the PCA scores scatter plot as shown in Fig. 3a. If yes, proceed to next step. If not, inspect chromatograms, extraction parameters, and instrument log, and repeat experiment (*see* **Note 18**).
3. Remove blank, standard mix A and B sample from the analysis. Remove features that show missing values in more than 10% of pooled samples. Further remove all features that show missing values in 75% of individual samples. Replace all missing values by a number ten times smaller than the lowest value in the data matrix.
4. Inspect if the pooled samples are clustered tightly, whereas individual samples are scattered in the scores scatter plot as shown in Fig. 3b. If yes, proceed to the next step. If not, *see* **Note 19**.
5. Examine PCA scores scatter plot for outliers, and examine their association with factors other than the disease; such as diet, dietary supplements, smoking or drinking status, current medications, coexisting pathologies, etc. or any other relevant and available metadata. Specifically, in case of any current medication or dietary supplement, examine the loading plot to identify ion(s) that contribute to their segregation. Inspect if these ions represent any known drugs, dietary supplements,

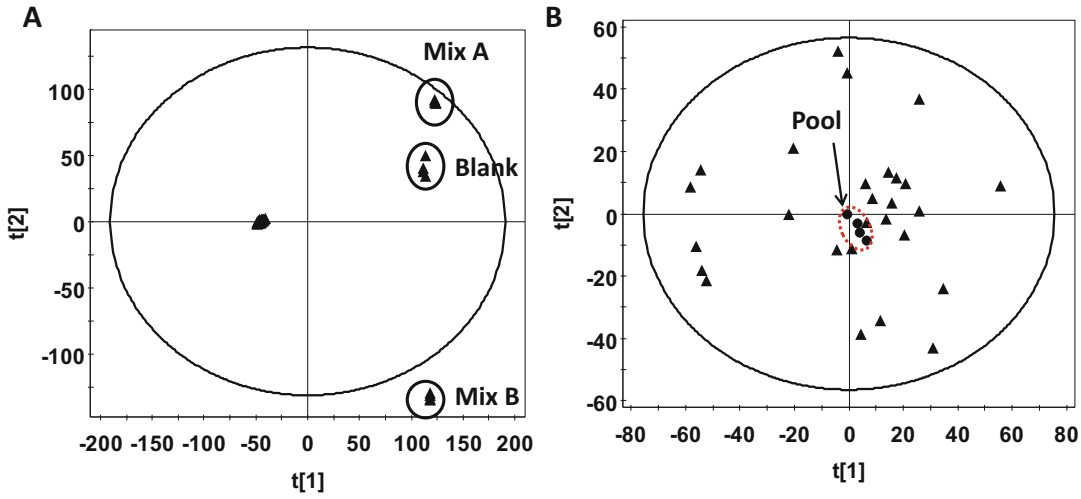


Fig. 3 Principal component analysis of the MarkerLynx processed data from reverse-phase UPLC-ESI-MS analysis of urine samples showing (**a**) tight clustering of external quality control samples, i.e., blank, standard mix A and standard mix B indicating consistent response and (**b**) tight clustering of pooled samples indicating reproducibility of ion chromatograms throughout the run

pollutants, or their metabolites by searching against databases (*see* Subheading 3.13). Remove all such ions of putative exogenous origin from the data matrix.

6. The data matrix is then uploaded on to the MetaboAnalyst web server as a peak intensity table (with unpaired samples in rows and features in columns) for statistical analysis. Choose sum normalization, log transformation (*see* **Note 20**), and Pareto scaling of the data to analyze intercorrelation between features across sample set. An example of result of analysis of intercorrelation between features in RPLC-MS analysis of urine samples is shown in Fig. 4.
7. The correlation matrix table is downloaded, and highly correlated features (Pearson correlation coefficient > 0.9; *see* **Note 21**) are screened for similarity in retention times. In case the retention times of these features are close, their elution profile is compared. In case there is a significant overlap between elution profiles, the mass differences between these ions are calculated to investigate if they correspond to common in-source fragments (e.g., $-H_2O$, $-CO_2$, $-NH_3$, –HCOOH, etc.) or adducts (Na^+, NH_4^+, etc. in the positive mode and Cl^-, $HCOO^-$, CH_3COO^-, etc. in the negative mode as well as solvent adducts) within the experimental mass error. For example, inset of Fig. 5 shows raw as well as extracted chromatograms of a number of co-eluting ions, which were eventually identified to be Na^+ -adduct (202.049+) and in-source

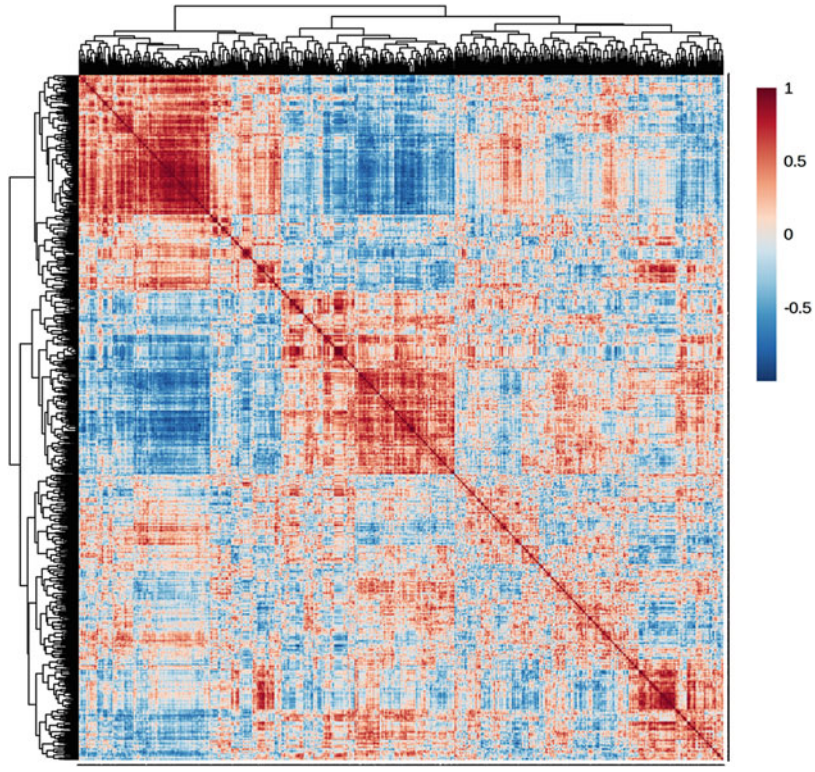


Fig. 4 Heat map representation of correlation between extracted features (*m/z*, retention time pair) across samples in reverse-phase UPLC-MS analysis of urine samples. The color scale for correlation coefficient is shown on the right side

fragments (105.035+, 77.041+) of hippuric acid (180.066+) that eluted at 2.6 min in RP analysis of urine samples.

8. Remove all adducts and fragments except the putative parent ion from the data matrix (*see* **Note 22**).

3.8 Normalization

1. After trimming of the data matrix as mentioned in Subheading 3.7, **steps 7** and **8**, the data is finally normalized before proceeding to the identification of features that distinguish the biofluid metabolome of controls and cases.
2. Normalize the urine metabolomics data either with respect to the total ion count or the creatinine peak (*see* **Note 23**).
3. For creatinine normalization, find out the creatinine peak ($m/z = 114.066+$ or $112.05-$; RT = 0.8 min in HILIC and 0.3 min in RP mode in the method described above) from the standard mix and pooled samples. Divide ion count of all peaks by creatinine ion count in the respective sample.
4. For serum, normalize the data with respect to total ion count.

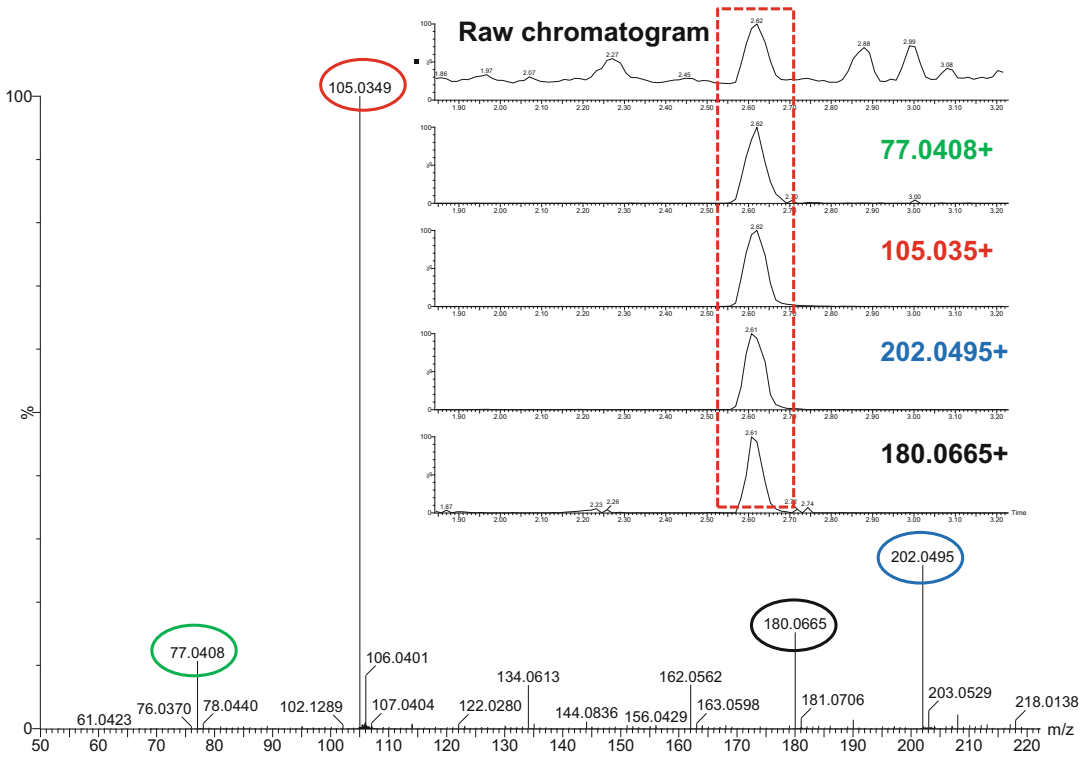


Fig. 5 Mass spectrum under the peak at 2.6 min in reverse-phase analysis of urine samples showing *in-source* fragments and adducts of hippuric acid. Raw and extracted ion chromatograms for co-eluting ions all of which originate from hippuric acid are shown in the inset. *m/z* values at 180.066+ and 202.049+ represent the protonated and sodiated species, whereas 105.035+ and 77.041+ represent the fragments

3.9 Pattern Recognition and Identification of Features of Interest

3.9.1 SIMCA-P12+

1. Import the normalized data matrix to SIMCA-P12+, and define class variable (Y) as "0" for healthy controls and "1" for cases. Pooled samples need not be given any Y value.
2. Make sure that pooled samples are forming a tight cluster, whereas individual samples are significantly scattered compared to them.
3. Exclude pooled samples and the 'Y' variable to subject the Pareto-scaled data matrix to PCA analysis. Check for any segregation of the samples in the scores scatter plot (e.g., *see* Fig. 6a). Correlate the segregation with disease status or any other information available as metadata. If the segregation is according to phenotype of interest, evaluate the loading plot, and select ions that are contributing to the segregation (*see* **Note 24**). Figure 6b shows the loading plot corresponding to Fig. 6a along with ions that contribute to segregation of controls vs. cases.
4. If the segregation is not according to any phenotype (*see* **Note 25**) of interest, remove most abundant ions and repeat the

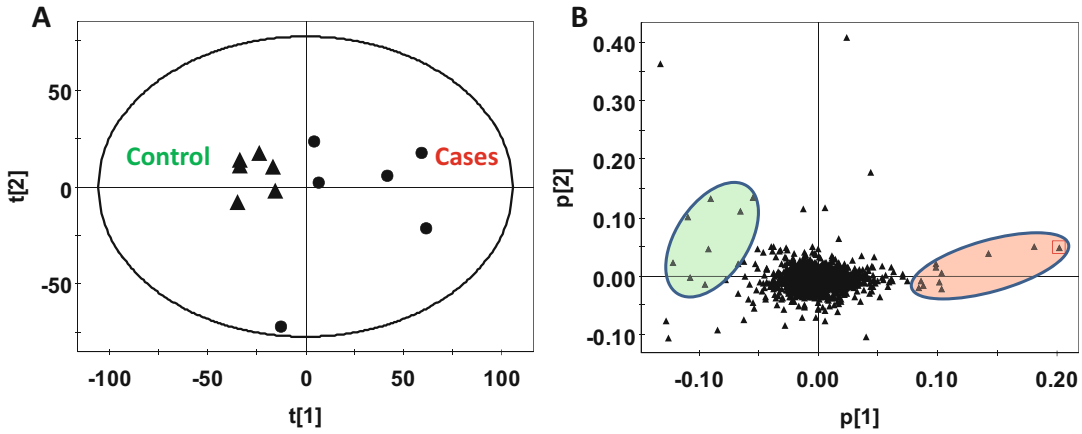


Fig. 6 (**a**) Representative scores scatter plot from principal component analysis showing segregation of controls and cases along first principal component and (**b**) corresponding loading plot. Ions that were more abundant in cases are shaded in red, whereas those more abundant in controls are shaded in green

PCA analysis. PCA is driven by most abundant features. They tend to mask relatively low abundant features that might be of interest with respect to the correlation between metabotype and phenotype.

5. If PCA indicates a good dataset (with tight clustering of pooled samples, few outliers, and most variances justified), proceed with orthogonal partial least squares discriminant analysis (OPLS-DA) by including the Y values to identify features that show significant difference between controls and cases.
6. R^2 and Q^2 values should at least be 0.5 and 0.4 for the OPLS-DA analysis. High R^2 with poor Q^2 value indicates over-fitting of the model. Discard any model with negative Q^2 value.
7. Use the loading S-plot (representative example shown in Fig. 7), and generate a list of ions with p(corr)[1] > 0.8 (elevated in cases, shaded in red in Fig. 7) or p(corr)[1] < −0.8 (depleted in cases, shaded in green in Fig. 7).

3.9.2 MetaboAnalyst

1. For feature identification using MetaboAnalyst, submit the normalized data matrix after **step 2** or **3** or **4** of Subheading 3.8 with samples labeled 0 and 1 according the phenotype category under comparison (e.g., 0 for healthy controls and 1 for cancer cases), and choose no normalization, log transformation, and Pareto scaling of the data (*see* **Note 26**).
2. Choose "volcano plot" for analysis and define fold change (typically 1.5); on the *x*-axis, choose comparison type (say 0/1), and on the *y*-axis, choose non-parametric test assuming unequal variance; define *P* value threshold (0.05 or lower), opt for FDR correction and submit.

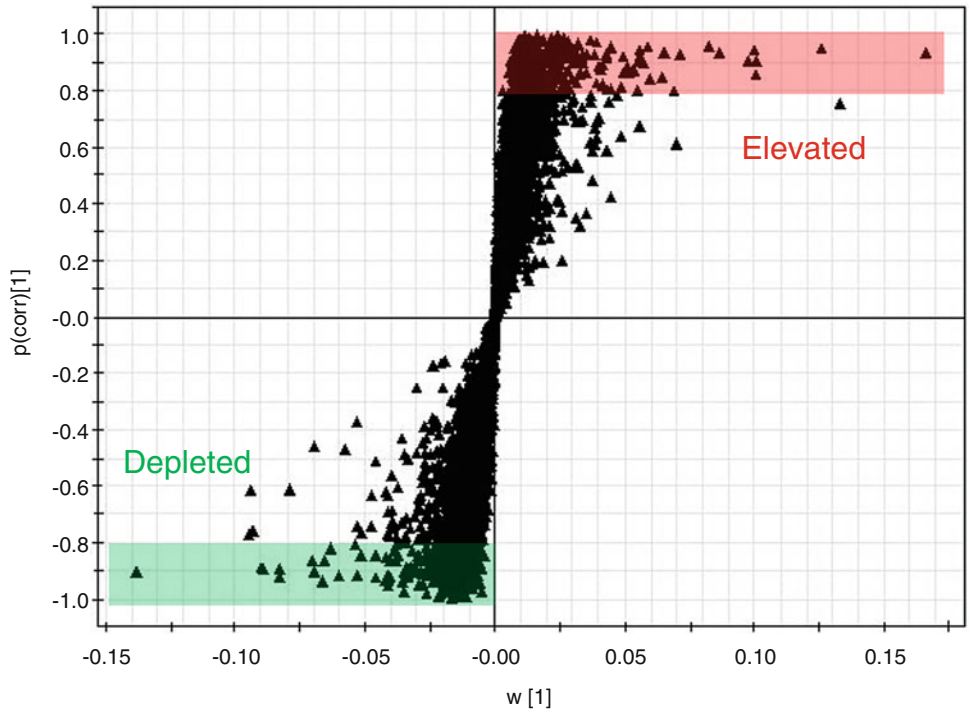


Fig. 7 Loading S-plot OPLS-DA analysis. While p(corr)[1] values represent the inter-class difference, w (1) values indicate the relative abundance of respective ions. Ions with p(corr)[1] > 0.8 (shaded in red) represent ions that were significantly more abundant in cases. Ions with p(corr)[1] < -0.8 (shaded in green) represent ions that were significantly more abundant in controls

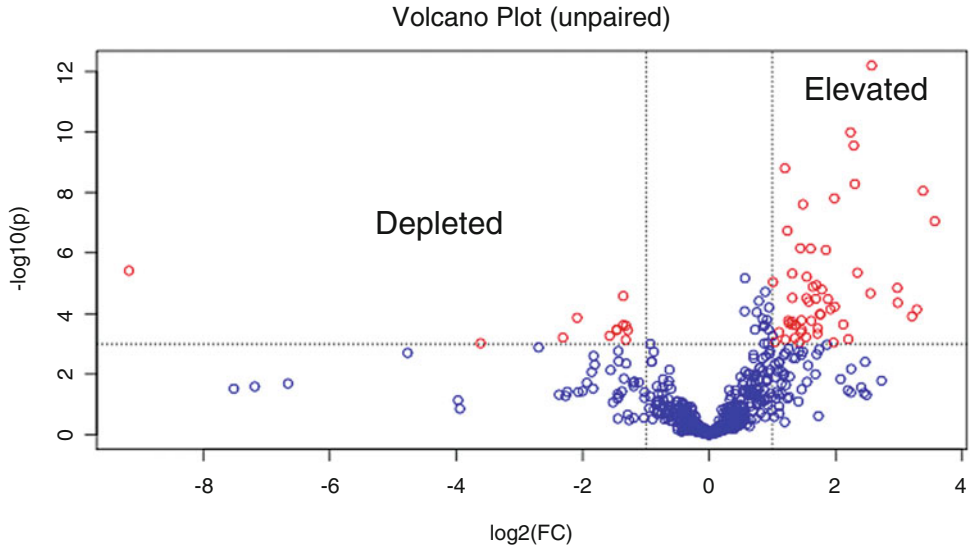


Fig. 8 Volcano plot generated through MetaboAnalyst-based analysis of normalized, log-transformed, and Pareto-scaled data matrix. P value and fold change cutoff were 0.001 and 2, respectively. Ions in the right upper third are those that were elevated in cases, and ions on the left upper third were those that were depleted in cases

3. The resulting plot should look something like that shown in Fig. 8. View detailed data table, and select features that show significant difference in abundance between controls and cases.

3.10 Pathway Analysis

1. Go to MassTRIX webpage to submit a new job.
2. Paste list of ions found to be significantly altered between control and cases.
3. Choose appropriate ionization mode along with potential adducts, e.g., Na^+ or K^+ for positive mode (*see* **Note 27**).
4. Choose mass error range according to the error range calculated from quality control samples.
5. In the database option, choose KEGG/HMDB/LIPIDMAPS without isotopes.
6. Choose Homo sapiens as the organism and submit the job.
7. Once completed, go to the result page and click on each of the pathways to find putative identity of ions belonging to the pathway.
8. Cancer is often associated with coordinate derangement of metabolism. Thus, a metabolic pathway that is altered as a result of upstream signaling events is likely to show changes in multiple metabolites. The more the association between the pathway and pathogenesis of cancer, the likelihood of differentially regulated ions belonging to the pathway increases. Thus, if an ion of interest has multiple hits, i.e., candidate compounds (*see* Subheading 3.13) from database mining, it belongs to a pathway that has an overall higher fraction of metabolites altered. Pathways may be prioritized for metabolite identification accordingly.

3.11 Statistical Tests

1. Perform final statistical tests using GraphPad Prism software.
2. Use Mann-Whitney U test with 95% confidence interval, and apply correction for multiple testing (such as Bonferroni correction) to test the significance of difference in level of a metabolite of interest between controls and cases or other groups according to metadata (demographics, lifestyle, BMI, disease status, treatment status, etc.).
3. Examine correlation between the phenotype of interest and abundance of a metabolite by Pearson correlation analysis.
4. Inspect the sensitivity and specificity of metabolites to distinguish between phenotypes of interest using receiver operating characteristics (ROC) analysis.

3.12 MS/MS Analysis

1. Perform MS/MS analysis on ions of interest that show a significant power to discriminate between controls and cases to get a preliminary idea about the type of structural and functional motifs.

2. Use the quad to select the ion of interest with unit mass resolution, and fragment it through collision-induced dissociation using argon as the collision gas and ramping the collision energy from 5 to 40 eV (*see* **Note 28**). Chromatography and other mass spec conditions, including lockspray, should be similar to that described in Subheading 3.4 for the respective sample and mode of analysis.

3.13 Database Mining

1. Submit the list of ions of interest to metabolomic databases to identify putative targets before confirmation through MS/MS analysis.
2. METLIN offers a comprehensive list of both exogenous and endogenous compounds along with fragmentation pattern for some of them (*see* **Note 29**). Choose a mass window as appropriate from the analysis of QC samples. Choose plausible adduct and fragmentation for the respective mode of ionization (*see* **Note 30**).
3. If the list of hits against the query masses contains fragmentation pattern in the databases, compare them with the fragmentation pattern of the ion obtained in your experiment, and screen according to similarity of fragmentation pattern (*see* **Note 31**).
4. Examine the structure of the hits, and screen compounds further based on an informed guess (*see* **Note 32**) considering their expected retention time and retention time of the ion of interest.

3.14 Confirmation of Identity of the Metabolite

1. In case compounds are available commercially, procure highest purity standards.
2. Make a 20–50 μM solution of standards in diluent A, B, or C as appropriate.
3. Run the following samples: standard sample, biofluid sample spiked with the standards (*see* **Note 33**), and biofluid sample alone.
4. Extract all chromatograms for the mass of the ion of interest to inspect if the retention times of the extracted peaks of these chromatograms match (*see* **Note 34**).
5. If they do, run an MS/MS analysis on the target ion, and check if fragmentation patterns match for all three spectrum to confirm the identity.
6. Even if no authentic standards are available or none of them match by retention time and/or fragmentation pattern, bona fide signatures that distinguish cases from controls or show a high correlation with a particular phenotype are of utmost interest, and they should be reported as a *m/z* retention time pair along with fragmentation pattern whenever possible (*see* **Note 35**).

Identification should be followed up by quantification of target compounds using multiple reaction monitoring (MRM) optimized using authentic standards in a tandem mass spectrometer such as a triple quad. The statistical significance of difference in abundance of individual metabolites should be tested correcting for multiple comparisons using the appropriate statistical test (e.g., with considerations such as whether the dataset is normally or non-normally distributed). For most unpaired case-control type studies, Mann-Whitney U test, which doesn't make any assumption on data distribution, with Bonferroni correction, is recommended for selecting potential biomarkers. Sensitivity and specificity of such biomarker (s) should be evaluated using receiver operating characteristic (ROC) analysis. In addition, pathway enrichment analysis is recommended to examine whether identified signatures are coordinated and connected to pathways that are known to get dysregulated in cancer [38]. Whenever feasible, correlation between gene expression signatures should also be examined to further establish such mechanistic associations [30, 38].

4 Notes

1. This may vary depending on the type of sample as well as experiment/analysis. For example, typically, we use a mixture of creatinine, creatine, deoxyuridine, L-glutamic acid, citric acid, L-arginine, hippuric acid, L-tryptophan, kynurenic acid, and indolelactic acid for urine.
2. Use of degassed solvent is a must in UPLC operations to avoid increase in back pressure and erratic chromatograms. However, due care should be exercised while sonicating organic solvents since sonication causes heating and builds up vapor inside the bottle. The cap must not be tightened too much so that pressure can be released during the process.
3. The method described here was developed and optimized using the aforementioned instruments, columns and softwares. However, the basic analytical approach is platform-independent, and it can be adopted for similar instruments from other vendors. Ultrahigh performance liquid chromatography (UHPLC) columns from other vendors could also be used.
4. As anticoagulant, EDTA should be avoided as it interferes with mass spectrometry-based metabolomic profiling, while citrate cannot be used since it is an endogenous metabolite. Multiple freeze-thaw cycle affects stability of the metabolome. Therefore, samples should be preferably divided into single-use aliquots and stored in −80 °C.

5. In most cases, analysis of human biofluids for identification of cancer-associated metabolomic signatures is carried out in a relatively small number of samples. In absence of appropriate metadata on drug, diet, and lifestyle, there is a possibility of spurious associations or masking of features of interest.
6. Urine samples may vary in overall concentration due to differences in water intake or underlying pathology. It may also vary in protein concentration due to differences in kidney function or coexisting pathologies (albuminuria), etc. If the sample is too concentrated or has a higher protein content, diluent/urine ratio should be increased. Two pooled samples containing control and cases, respectively, should be prepared and used to standardize the diluent/urine ratio. Typically, 1:5 dilution suffices for most urine samples.
7. Diluents should be kept for at least few hours in −20 °C freezer. Addition of the chilled diluent helps to precipitate proteins and other sparingly soluble compounds that may precipitate out of the solution during the course of analysis when the sample sits in a sample chamber maintained at 12 °C. Precipitation during analysis might result in artefacts and false positives.
8. Samples should not be left too long after centrifugation as the pellet starts re-suspending. Also, if samples start showing some turbidity upon standing at the sample chamber of the instrument, it might indicate incomplete deproteination. In such case, a higher amount of organic solvent or higher diluent/sample ratio should be used.
9. In case the number of samples is large, Waters 96-well Sirocco (1 mL) protein precipitation plates could be used instead of 48-well sample holder. Typically, 500 μL of sample is loaded into each well. For a large number of samples, the diluent addition, supernatant transfer, and any further dilution can also be performed using robotic liquid handlers. If available sample volume is low, a total of 300 μL of the urine and diluent mixture can be prepared keeping urine/diluent ratio the same, out of which 200 μL of supernatant can be transferred into 300 μL 12 × 32 mm polypropylene vials with PTFE septa for UPLC-MS analysis. These vials give better results compared to 350 μL 96-well Sirocco plates. While using any 96-well plate, make sure to choose the correct sample bed layout in the autosampler settings.
10. Serum samples should typically be either colorless or light yellow and slightly turbid. Note presence of significant turbidity in any sample. This is sometimes observed in advanced-stage cancer patients with significant cachexia and lipolysis [43, 44]. Also note any reddish or brownish colored samples.

This indicates hemolysis that might be due to improper handling of samples or hemolytic anemia that is observed in some cancer patients [45]. These should be specifically inspected during the analysis, particularly if they turn out to be outliers in the PCA analysis.

11. Addition of serum into the chilled diluent B should result in visible cloudiness starting from the pipette tip. In case no cloudiness is noted, organic content of the diluent may have changed. Use freshly prepared and chilled diluent B to ensure deproteination.
12. Significant temperature fluctuations may adversely affect the mass accuracy in TOF instruments. In case sudden changes in mass profile are noted during a long run, do check the room temperature profile for temperature variations.
13. These experimental parameters should always be optimized by the operator of a particular platform.
14. It is often observed that the first couple of injections of a sample give a different response than the subsequent ones of the same sample. Therefore, injecting a few pooled samples is essential to condition the column with the constituent analytes and other matrix components. Only after the repeated injection of the pooled sample starts giving reproducible signature, samples of interest should be injected. This is typically achieved after six consecutive injections.
15. In most cases, the retention time reproducibility using the method described is very good, and drift in retention time is typically less than peak width. However, dirty samples or insufficient column equilibration may result in shifts in retention time. QC samples, such as blanks or pools, should help to identify such phenomenon, and if needed, an alignment of chromatograms should be performed.
16. This can be executed by creating a method using "TargetLynx" module of the MassLynx software mentioning the m/z value of the ion of interest and retention time along with respective windows. The resulting file can be collated in excel to calculate the CV for each peak.
17. If the retention time or mass window is narrower than those experimentally observed in QC pooled samples, it might lead to peak splitting in the MarkerLynx results. Lowering the ion count cutoff also introduces noise into the data matrix. Choose retention time and mass window according to the results obtained in QC analysis. Inspect the baseline of the blank and standard mix samples to determine an initial cutoff for ion counts. Optimize it by minimizing spurious features (with no bona fide elution profile across samples and missing values) in

the dataset by repeated extraction of features at various cutoffs above and below the initial value.

18. This may be caused by changes in instrument response during the run. The quality control steps reduce such chances. However, use of very narrow mass and retention time window or very low ion count cutoff in MarkerLynx can lead to peak splitting or falsely introducing noise into the dataset, counting them as real signals.
19. In addition to the points mentioned in **Note 17**, this may also be caused by poor sample stability during a long run. Quality control steps using pooled samples are designed to capture such incidents. Inspect for any turbidity or precipitation in the sample in such case. Although very rare, this situation may also arise if the CVs of metabolic features (biological) are indeed comparable to their respective experimental CVs. In such case, check for any confounding factors, rethink if the cohort is suitable for analysis of differential signatures associated with cancer, and increase the sample number until the samples are significantly more scattered in the PCA scores scatter plot than the pooled samples.
20. Sum normalization basically refers to dividing ion count for each feature by total ion count in that sample. Most of the statistical analysis and model building that are performed assume a Gaussian (bell-shaped) distribution of the data. However, owing to a huge variation in the relative abundance of individual metabolites within and across samples as well as due to a limited sample size (in most experiments), the data distribution is often skewed (*see* Fig. 9a). Log transformation allows for the data distribution to appear more Gaussian (*see* Fig. 9b) and allows for further statistical analysis that assume such parametric distributions.
21. This cutoff depends on the data quality. Check correlation coefficients of known ion clusters originating from the same molecule (e.g., hippuric acid and its Na^+ adducts and fragments as shown in Fig. 5). The correlation coefficient between these ions in standard mixture B or pooled samples can be used to decide the cutoff value for the correlation coefficient. It should be noted that a correlation between features may also arise if they belong to biologically connected pathways.
22. This assumes that the relative abundance of adduct and fragments with respect to the putative molecular ion (protonated or deprotonated species in positive and negative ESI-MS, respectively) is not significantly different across samples, and, therefore, the putative molecular ion itself can faithfully represent the changes in the abundance of the parent compound in these samples. Although this is a reasonable assumption in

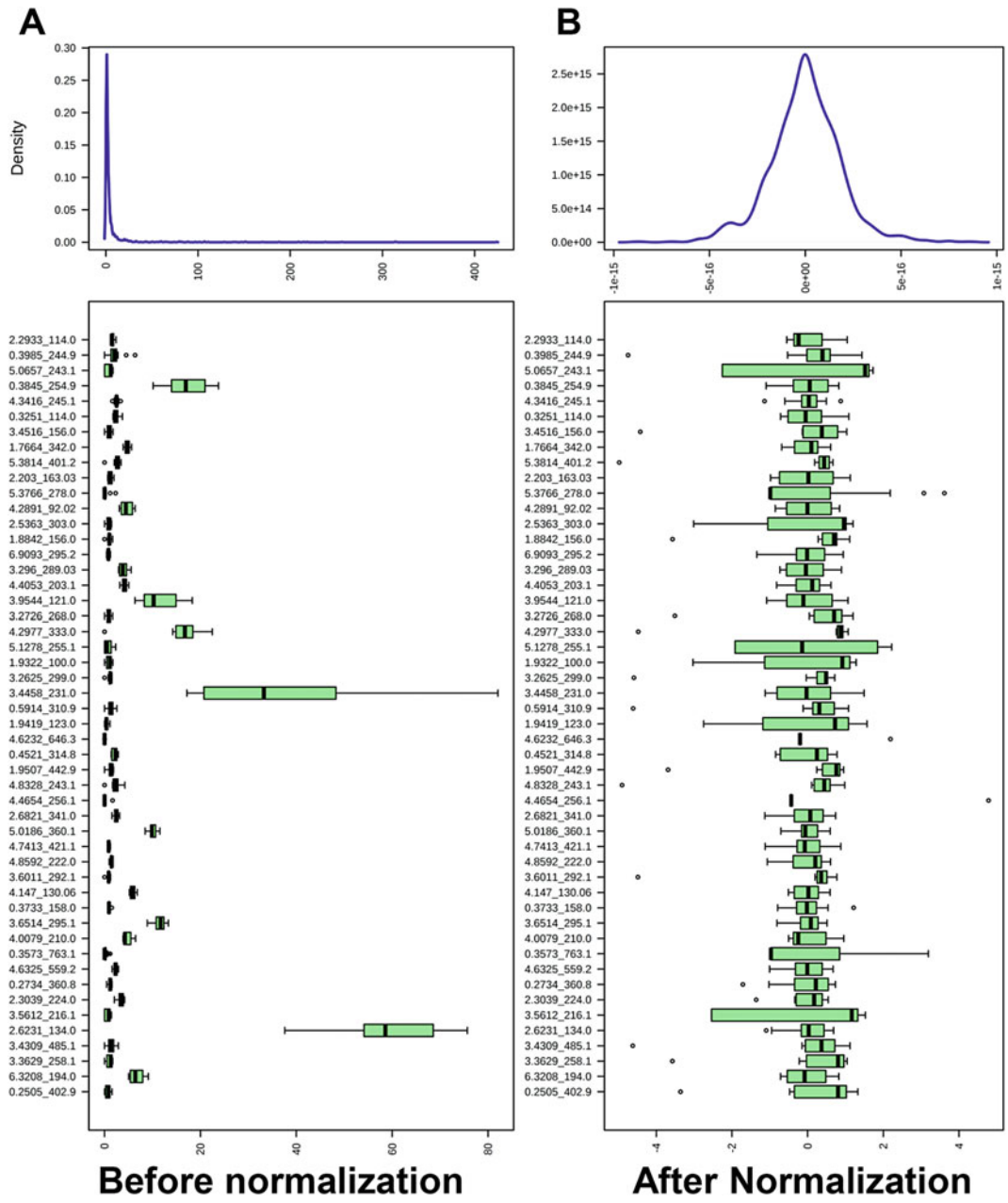


Fig. 9 Effect of data normalization on the shape of data distribution

most cases, a better alternative is to sum up the intensities of all ions arising from the putative molecular ion and use that for further statistical analysis. Progenesis QI (Nonlinear Dynamics) (or other equivalent software) examines various adduct/fragment relationships among a co-eluting cluster of ions and helps to sum up the total intensity of all such ions that are

presumably generated from the same parent compound. Such trimming of the data matrix reduces chances of model overfitting and identification of spurious associations.

23. A large number of exogenous compounds are excreted in urine. These include but are not limited to drugs, drug metabolites, metabolites related to the diet and dietary supplements, pollutants, etc. Although some ions/features related to these compounds may be identified and eliminated during data processing, many others may remain. Total ion count (TIC) normalization including these peaks would artificially alter relative abundance of endogenous peaks in samples that have different abundance of peaks related to these exogenous compounds. This would affect the supervised as well as unsupervised analysis and may lead to spurious associations between features and phenotypes. Therefore, TIC normalization may not be ideal. In addition, endogenous metabolomic signature may be affected by dilution (due to change in water intake) or changes in glomerular filtration rate (GFR). Creatinine concentration can be used as a surrogate to normalize the data for any dilution or changes in GFR. Please note that unlike TIC normalization, creatinine normalization is unlikely to artificially alter the relative abundance of an endogenous metabolite. It is recommended that the data is normalized by both methods to identify unmistakable changes associated with the disease/phenotype. Another alternative for data normalization in the biofluid metabolomic analysis is probabilistic quotient normalization that has been described elsewhere [46].
24. If any of the internal standards turns out to be a differentiating feature in PCA or OPLS-DA analysis, in spite of satisfactory results in QC analysis, this might indicate a problem with the data normalization as described in **Note 23**. Ways to deal with this problem include (1) removal of overabundant features that are potentially non-endogenous, (2) checking metadata for confounding factors, (3) recruitment of additional samples, and (4) critical analysis of the study design.
25. PCA is more appropriate for inspecting the overall data quality and not so much for the identification of features of interest, unless they are obvious. Some samples may look like an outlier in the PCA analysis. This may be caused by biological factors (significantly different metabolic profiles or presence of compounds of exogenous origin) or may be an analytical artifact. Removal of outliers to achieve better model fit and segregation of data should be avoided.
26. Data exclusion in MetaboAnalyst may be skipped since following the protocol mentioned earlier for data preprocessing (*see* Subheading 3.7), the number of features would be less than 5000 in most cases.

27. Although MetaboAnalyst also offers the option to use raw mass spectrometry data to analyze putative pathways, it doesn't offer the option to search for adducts that MassTRIX does. This gives MassTRIX an edge in identifying putative pathways and allows for biomarker identification, since adducts like Na^+ and Cl^- are quite common.
28. Using an energy ramp helps to sample different fragmentation reaction in a single go. However, some molecules may be very labile and undergo in-source fragmentation. In such cases, the fragmentation spectra should be recorded by increasing collision energy stepwise: 0, 10, 20, 30, etc.
29. In addition, the LC-MS search option in the HMDB database can also be used to search for potential hits that have been reported to be present in human samples. The criteria for mass error window and selection of potential adduct and fragments should be similar to that described for the METLIN database.
30. Although METLIN provides a host of adducts and fragments, they should be judiciously chosen during submission of a query. For example, upon selection of the positive mode ionization, solvent adducts such as CH_3CN and MeOH are shown as plausible species. However, if neither sample nor mobile phases contain the solvent (e.g., MeOH in case of RP analysis of urine samples), such solvent adducts need not be chosen during search. In addition, ions with multiple charges (e.g., M + 2H, M + 3H, M + H + Na, etc.) are generally not observed for small molecules, and they should not be selected in the METLIN search, unless the isotopic peak pattern for the ion clearly indicates that the ion is multiply charged.
31. Please note that the fragmentation pattern depends on collision energy as well as on the type of an instrument. Also, while METLIN contains fragmentation patterns at different energies, this method describes a ramp from 5 to 40 eV to sample all possible fragments. Therefore, the MS/MS spectra, particularly relative abundance of different fragments, reported in METLIN are likely to be different from those observed using the method described here. So, fragmentation at different energy levels reported in METLIN should be inspected while comparing with experimental MS/MS spectra.
32. Look at the structure of the hit and compare it with structures of compounds in the authentic standard mix B that was run as a QC sample. Assess the similarity of chromatographic behavior based on structural motifs, functional groups, and polarity. For example, in the RP mode analysis described here for urine samples, the authentic standard of hippuric acid elutes at 2.6 min. If a peak of interest has a retention time similar or less than that of hippuric acid and a hit for the ion contains an

indole or naphthol or indane moiety without any other significantly polar group (such as sulfate, glucuronide, etc.), it is very unlikely that the peak will match with the hit in retention time. On the other hand, hits with structures like nucleosides are not likely to be true if the retention time for the ion of interest is similar to or more than that of hippuric acid. Inclusion of a number of endogenous metabolites representing different structural/functional motifs and polarity in the standard mix B helps in the process of elimination and reduces the number of authentic standards to be run for compound identification. ***Caution***: This is a non-trivial step that should be attempted only by experts in chromatography. Those not very familiar with chromatography may choose to seek guidance from an expert.

33. It is recommended to run at least two biofluid samples spiked with the standard, one with the lowest abundance and the other with more than an average abundance. Both of these samples should also be run without spiking the standard so that an increase intensity of the ion of interest in the extracted chromatogram is reliably detected at the desired retention time.
34. At times, the retention profile of the pure compound may be slightly different from that of the spiked sample due to matrix effect. Using low and high abundance samples as mentioned above in **Note 33** often helps to resolve the issue.
35. A large number of features (*m/z*, retention pairs) remain unidentified. Even when potential candidates are reported in the literature, authentic standards may not be available commercially. In such case, the only option is to synthesize candidate molecules and to confirm them by comparing retention time and fragmentation pattern. This is a challenging task. However, only untargeted metabolomic profiling holds the promise of identification of such novel molecules and/or related pathways [28, 40, 47], which can significantly further our understanding of cancer biology.

Acknowledgments

Authors would like to sincerely acknowledge the contribution of Mr. Kristopher W. Krausz in developing methods for UPLC-ESIMS analysis and Dr. Frank J. Gonzalez (Laboratory of Metabolism, Center for Cancer Research, National Cancer Institute, Bethesda, USA) for his encouragement and support. This work was supported by Saha Institute of Nuclear Physics, Kolkata, India.

References

1. Warburg O (1956) On the origin of cancer cells. Science 123(3191):309–314
2. Vander Heiden MG, Cantley LC, Thompson CB (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science 324(5930):1029–1033. https://doi.org/10.1126/science.1160809
3. Pavlova NN, Thompson CB The emerging hallmarks of cancer metabolism. Cell Metab 23(1):27–47. https://doi.org/10.1016/j.cmet.2015.12.006
4. Hsu PP, Sabatini DM (2008) Cancer cell metabolism: Warburg and beyond. Cell 134 (5):703–707. https://doi.org/10.1016/j.cell.2008.08.021
5. DeBerardinis RJ, Mancuso A, Daikhin E, Nissim I, Yudkoff M, Wehrli S, Thompson CB (2007) Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. Proc Natl Acad Sci U S A 104(49):19345–19350. https://doi.org/10.1073/pnas.0709747104
6. Johnson C, Warmoes MO, Shen X, Locasale JW Epigenetics and cancer metabolism. Cancer Lett 356(2 Pt A):309–314. https://doi.org/10.1016/j.canlet.2013.09.043
7. Kinnaird A, Zhao S, Wellen KE, Michelakis ED Metabolic control of epigenetics in cancer. Nat Rev Cancer 16(11):694–707. https://doi.org/10.1038/nrc.2016.82
8. Ma S, Jiang B, Deng W, Gu ZK, Wu FZ, Li T, Xia Y, Yang H, Ye D, Xiong Y, Guan KL D-2-hydroxyglutarate is essential for maintaining oncogenic property of mutant IDH-containing cancer cells but dispensable for cell growth. Oncotarget 6 (11):8606–8620. https://doi.org/10.18632/oncotarget.3330
9. Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, Fantin VR, Jang HG, Jin S, Keenan MC, Marks KM, Prins RM, Ward PS, Yen KE, Liau LM, Rabinowitz JD, Cantley LC, Thompson CB, Vander Heiden MG, Su SM (2009) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 462(7274):739–744. https://doi.org/10.1038/nature08617
10. Kerins MJ, Vashisht AA, Liang BX, Duckworth SJ, Praslicka BJ, Wohlschlegel JA, Ooi A Fumarate mediates a chronic proliferative signal in fumarate hydratase-inactivated cancer cells by increasing transcription and translation of ferritin genes. Mol Cell Biol 37(11). pii: e00079-17). https://doi.org/10.1128/MCB.00079-17
11. Sciacovelli M, Goncalves E, Johnson TI, Zecchini VR, da Costa AS, Gaude E, Drubbel AV, Theobald SJ, Abbo SR, Tran MG, Rajeeve V, Cardaci S, Foster S, Yun H, Cutillas P, Warren A, Gnanapragasam V, Gottlieb E, Franze K, Huntly B, Maher ER, Maxwell PH, Saez-Rodriguez J, Frezza C Fumarate is an epigenetic modifier that elicits epithelial-to-mesenchymal transition. Nature 537 (7621):544–547. https://doi.org/10.1038/nature19353
12. Selak MA, Armour SM, MacKenzie ED, Boulahbel H, Watson DG, Mansfield KD, Pan Y, Simon MC, Thompson CB, Gottlieb E (2005) Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-alpha prolyl hydroxylase. Cancer Cell 7(1):77–85. https://doi.org/10.1016/j.ccr.2004.11.022
13. Yang M, Pollard PJ Succinate: a new epigenetic hacker. Cancer Cell 23(6):709–711. https://doi.org/10.1016/j.ccr.2013.05.015
14. Chan EC, Koh PK, Mal M, Cheah PY, Eu KW, Backshall A, Cavill R, Nicholson JK, Keun HC (2009) Metabolic profiling of human colorectal cancer using high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy and gas chromatography mass spectrometry (GC/MS). J Proteome Res 8(1):352–361. https://doi.org/10.1021/pr8006232
15. Denkert C, Budczies J, Kind T, Weichert W, Tablack P, Sehouli J, Niesporek S, Konsgen D, Dietel M, Fiehn O (2006) Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors. Cancer Res 66(22):10795–10804. https://doi.org/10.1158/0008-5472.CAN-06-0755
16. Hirayama A, Kami K, Sugimoto M, Sugawara M, Toki N, Onozuka H, Kinoshita T, Saito N, Ochiai A, Tomita M, Esumi H, Soga T (2009) Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. Cancer Res 69(11):4918–4925. https://doi.org/10.1158/0008-5472.CAN-08-4806
17. Rocha CM, Barros AS, Gil AM, Goodfellow BJ, Humpfer E, Spraul M, Carreira IM, Melo JB, Bernardo J, Gomes A, Sousa V, Carvalho L, Duarte IF Metabolic profiling of human lung cancer tissue by 1H high resolution magic angle spinning (HRMAS) NMR spectroscopy.

J Proteome Res 9(1):319–332. https://doi.org/10.1021/pr9006574

18. Roig B, Rodriguez-Balada M, Samino S, Lam EW, Guaita-Esteruelas S, Gomes AR, Correig X, Borras J, Yanes O, Guma J Metabolomics reveals novel blood plasma biomarkers associated to the BRCA1-mutated phenotype of human breast cancer. Sci Rep 7(1):17831. https://doi.org/10.1038/s41598-017-17897-8
19. Jove M, Collado R, Quiles JL, Ramirez-Tortosa MC, Sol J, Ruiz-Sanjuan M, Fernandez M, de la Torre Cabrera C, Ramirez-Tortosa C, Granados-Principal S, Sanchez-Rovira P, Pamplona R A plasma metabolomic signature discloses human breast cancer. Oncotarget 8(12):19522–19533. https://doi.org/10.18632/oncotarget.14521
20. Yonezawa K, Nishiumi S, Kitamoto-Matsuda J, Fujita T, Morimoto K, Yamashita D, Saito M, Otsuki N, Irino Y, Shinohara M, Yoshida M Nibu K Serum and tissue metabolomics of head and neck cancer. Cancer Genomics Proteomics 10(5):233–238
21. Nishiumi S, Kobayashi T, Ikeda A, Yoshie T, Kibi M, Izumi Y, Okuno T, Hayashi N, Kawano S, Takenawa T, Azuma T, Yoshida M A novel serum metabolomics-based diagnostic approach for colorectal cancer. PLoS One 7(7): e40459. https://doi.org/10.1371/journal.pone.0040459
22. Hadi NI, Jamal Q, Iqbal A, Shaikh F, Somroo S, Musharraf SG Serum metabolomic profiles for breast cancer diagnosis, grading and staging by gas chromatography-mass spectrometry. Sci Rep 7(1):1715. https://doi.org/10.1038/s41598-017-01924-9
23. Farshidfar F, Weljie AM, Kopciuk K, Buie WD, Maclean A, Dixon E, Sutherland FR, Molckovsky A, Vogel HJ, Bathe OF Serum metabolomic profile as a means to distinguish stage of colorectal cancer. Genome Med 4 (5):42. doi:https://doi.org/10.1186/gm341
24. Asiago VM, Alvarado LZ, Shanaiah N, Gowda GA, Owusu-Sarfo K, Ballas RA, Raftery D Early detection of recurrent breast cancer using metabolite profiling. Cancer Res 70 (21):8309–8318. https://doi.org/10.1158/0008-5472.CAN-10-1319
25. Ganti S, Weiss RH Urine metabolomics for kidney cancer detection and biomarker discovery. Urol Oncol 29(5):551–557. https://doi.org/10.1016/j.urolonc.2011.05.013
26. Haznadar M, Cai Q, Krausz KW, Bowman ED, Margono E, Noro R, Thompson MD, Mathe EA, Munro HM, Steinwandel MD, Gonzalez FJ, Blot WJ, Harris CC Urinary metabolite risk biomarkers of lung cancer: a prospective cohort study. Cancer Epidemiol Biomarkers Prev 25 (6):978–986. https://doi.org/10.1158/1055-9965.EPI-15-1191
27. Kind T, Tolstikov V, Fiehn O, Weiss RH (2007) A comprehensive urinary metabolomic approach for identifying kidney cancerr. Anal Biochem 363(2):185–195. https://doi.org/10.1016/j.ab.2007.01.028
28. Mathe EA, Patterson AD, Haznadar M, Manna SK, Krausz KW, Bowman ED, Shields PG, Idle JR, Smith PB, Anami K, Kazandjian DG, Hatzakis E, Gonzalez FJ, Harris CC Noninvasive urinary metabolomic profiling identifies diagnostic and prognostic markers in lung cancer. Cancer Res 74(12):3259–3270. https://doi.org/10.1158/0008-5472.CAN-14-0109
29. Patel D, Thompson MD, Manna SK, Krausz KW, Zhang L, Nilubol N, Gonzalez FJ, Kebebew E Unique and novel urinary metabolomic features in malignant versus benign adrenal neoplasms. Clin Cancer Res 23 (17):5302–5310. https://doi.org/10.1158/1078-0432.CCR-16-3156
30. Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J, Laxman B, Mehra R, Lonigro RJ, Li Y, Nyati MK, Ahsan A, Kalyana-Sundaram S, Han B, Cao X, Byun J, Omenn GS, Ghosh D, Pennathur S, Alexander DC, Berger A, Shuster JR, Wei JT, Varambally S, Beecher C, Chinnaiyan AM (2009) Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature 457 (7231):910–914. https://doi.org/10.1038/nature07762
31. Ishikawa S, Sugimoto M, Kitabatake K, Sugano A, Nakamura M, Kaneko M, Ota S, Hiwatari K, Enomoto A, Soga T, Tomita M, Iino M Identification of salivary metabolomic biomarkers for oral cancer screening. Sci Rep 6:31520. https://doi.org/10.1038/srep31520
32. Xu X, Cheng S, Ding C, Lv Z, Chen D, Wu J, Zheng S Identification of bile biomarkers of biliary tract cancer through a liquid chromatography/mass spectrometry-based metabolomic method. Mol Med Rep 11(3):2191–2198. https://doi.org/10.3892/mmr.2014.2973
33. Locasale JW, Melman T, Song S, Yang X, Swanson KD, Cantley LC, Wong ET, Asara JM Metabolomics of human cerebrospinal fluid identifies signatures of malignant glioma. Mol Cell Proteomics 11(6):M111.014688. https://doi.org/10.1074/mcp.M111.014688
34. Armitage EG, Barbas C Metabolomics in cancer biomarker discovery: current trends and future perspectives. J Pharm Biomed Anal

87:1–11. https://doi.org/10.1016/j.jpba.2013.08.041

35. Kim YS, Maruvada P, Milner JA (2008) Metabolomics in biomarker discovery: future uses for cancer prevention. Future Oncol 4 (1):93–102. https://doi.org/10.2217/14796694.4.1.93
36. Ward PS, Patel J, Wise DR, Abdel-Wahab O, Bennett BD, Coller HA, Cross JR, Fantin VR, Hedvat CV, Perl AE, Rabinowitz JD, Carroll M, Su SM, Sharp KA, Levine RL, Thompson CB The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. Cancer Cell 17(3):225–234. https://doi.org/10.1016/j.ccr.2010.01.020
37. Tomlinson IP, Alam NA, Rowan AJ, Barclay E, Jaeger EE, Kelsell D, Leigh I, Gorman P, Lamlum H, Rahman S, Roylance RR, Olpin S, Bevan S, Barker K, Hearle N, Houlston RS, Kiuru M, Lehtonen R, Karhu A, Vilkki S, Laiho P, Eklund C, Vierimaa O, Aittomaki K, Hietala M, Sistonen P, Paetau A, Salovaara R, Herva R, Launonen V, Aaltonen LA (2002) Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. Nat Genet 30(4):406–410. https://doi.org/10.1038/ng849
38. Manna SK, Tanaka N, Krausz KW, Haznadar M, Xue X, Matsubara T, Bowman ED, Fearon ER, Harris CC, Shah YM, Gonzalez FJ Biomarkers of coordinate metabolic reprogramming in colorectal tumors in mice and humans. Gastroenterology 146 (5):1313–1324. https://doi.org/10.1053/j.gastro.2014.01.017
39. Tautenhahn R, Cho K, Uritboonthai W, Zhu Z, Patti GJ, Siuzdak G An accelerated workflow for untargeted metabolomics using the METLIN database. Nat Biotechnol 30(9):826–828. https://doi.org/10.1038/nbt.2348
40. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vazquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, Sayeeda Z, Lo E, Assempour N, Berjanskii M, Singhal S, Arndt D, Liang Y, Badran H, Grant J, Serra-Cayuela A, Liu Y, Mandal R, Neveu V, Pon A, Knox C, Wilson M, Manach C, Scalbert A HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res 46(D1): D608–D617. https://doi.org/10.1093/nar/gkx1089
41. Suhre K, Schmitt-Kopplin P (2008) MassTRIX: mass translator into pathways. Nucleic Acids Res 36(Web Server):W481–W484. https://doi.org/10.1093/nar/gkn194
42. Xia J, Wishart DS Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. Curr Protoc Bioinformatics 55:14.10.11–14.10.91. https://doi.org/10.1002/cpbi.11
43. Agustsson T, Ryden M, Hoffstedt J, van Harmelen V, Dicker A, Laurencikiene J, Isaksson B, Permert J, Arner P (2007) Mechanism of increased lipolysis in cancer cachexia. Cancer Res 67(11):5531–5537. https://doi.org/10.1158/0008-5472.CAN-06-4585
44. Arner P, Langin D Lipolysis in lipid turnover, cancer cachexia, and obesity-induced insulin resistance. Trends Endocrinol Metab 25 (5):255–262. https://doi.org/10.1016/j.tem.2014.03.002
45. Ellis LD, Westerman MP (1965) Autoimmune hemolytic anemia and cancer. JAMA 193:962–964
46. Dieterle F, Ross A, Schlotterbeck G, Senn H (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. Anal Chem 78 (13):4281–4290. https://doi.org/10.1021/ac051632c
47. Matsubara T, Tanaka N, Krausz KW, Manna SK, Kang DW, Anderson ER, Luecke H, Patterson AD, Shah YM, Gonzalez FJ Metabolomics identifies an inflammatory cascade involved in dioxin- and diet-induced steatohepatitis. Cell Metab 16(5):634–644. https://doi.org/10.1016/j.cmet.2012.10.006

Chapter 13

Gas Chromatography-Mass Spectrometry and Analysis of the Serum Metabolomic Profile Through Extraction and Derivatization of Polar Metabolites

Jodi Rattner, Farshad Farshidfar, and Oliver F. Bathe

Abstract

Metabolite profiling in complex biological matrices such as serum requires high-throughput technologies capable of accurate and reproducible quantitative analysis and detection of slight differences in metabolite concentrations. Gas chromatography-mass spectrometry (GC-MS) is widely used for characterizing the metabolome. This chapter summarizes the necessary preparatory steps required to profile the metabolome using GC-MS. While this chapter focuses on evaluating polar metabolites in serum samples, the methods can be adapted to quantify nonpolar metabolites in other biological matrices.

Key words Gas chromatography-mass spectrometry, Serum, Metabolite profiling, Biomarker, Cancer

1 Introduction

The detection and classification of small molecule intermediaries (metabolites) in complex matrices such as serum are one way to characterize normal and perturbed biological systems [1–3] (*see* **Note 1**). Nuclear magnetic resonance (NMR) spectroscopy, liquid chromatography-mass spectrometry (LC-MS), and gas chromatography-mass spectrometry (GC-MS) are widely used analytical platforms in the field of metabolomics [4, 5]. No single platform has the capability to accurately cover the whole metabolome, and each platform has its advantages and disadvantages. All of these platforms have the sensitivity to detect even low abundance features such as metabolites [6]. 1H-NMR spectroscopy is the only fully quantitative modality, and it is highly reliable and accurate [7–10]. However, it has limited utility in a clinical laboratory because of its size, the need for a special facility to house a powerful magnet, and its limited capacity for high-throughput testing, as well as relatively low sensitivity [11]. LC-MS can detect a large number of compounds (more than GC-MS) [12, 13]. However,

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_13,

the measurement of more compounds comes with the cost of a lower signal-to-noise ratio, resulting in greater difficulty constructing a stable metabolomic model during discovery work. GC-MS allows for the identification and quantification of a diverse array of (polar and nonpolar) metabolites within a single sample, covering a broad range of metabolic pathways. GC-MS is sensitive and robust, and standard protocols have been created for instrument maintenance, as well as compound identification [14, 15]. Finally, GC-MS is capable of high-throughput testing, and it is relatively compact, typically occupying the footprint of a desktop. GC-MS is therefore easy to utilize in a clinical laboratory setting (*see* **Note 2**).

In the gas chromatograph, compounds elute from the column at different times (retention times), according to their physicochemical properties. They are transported by the mobile phase to the mass spectrometer, where they are ionized and separated further by atomic mass. The GC-MS spectrum therefore resolves metabolite fragments by retention time, atomic mass, and charge. The height of the peaks (or Gaussian curves) is proportional to each ion's concentration. The most common type of mass spectrometer paired with GC is a quadrupole MS. Time-of-flight (TOF)-MS has a number of advantages over quadrupole instruments, including greater resolution and faster scan times, which improves its high-throughput testing capabilities.

GC-MS is semiquantitative, although it is capable of detecting even small changes in metabolite concentrations. The addition of internal standards allows accurate quantification, reduces measurement variance, and smooths out inter-batch and intra-batch variability. The state-of-the-art approach to assay design for the quantification of targeted metabolites by mass spectrometry employs isotopic labeling, which involves spiking each sample with a known quantity of selected metabolites that have been labeled with a nonradioactive isotope, such as deuterium. This causes a measurable shift in the mass-to-charge spectrum, making it possible to quantify the unlabeled peak. Deuterium-labeled metabolites are relatively inexpensive and widely accessible. All deuterium-labeled metabolites should be tested to define the dynamic range of detection in a given instrument; standard curves consisting of known quantities of the corresponding non-labeled metabolites should be constructed to determine their linearity, as well as the accuracy and reliability of measurements.

The following chapter will describe the experimental approach for the detection of polar compounds in serum samples. GC-MS is also capable of detecting nonpolar metabolites. This would require a different GC column and different preparation procedures described by others [16–19]. Our description will detail the experimental design, experimentation, spectral analysis, as well as data preprocessing and cleaning. While we will touch on the statistical analysis and interpretation of results, a detailed description is out of the scope of this chapter.

2 Materials

The preparation of solutions requires the usage of analytical grade reagents and pure deionized water. The methanol-chloroform solution, chloroform H_2O solution, methoxyamine hydrochloride in pyridine solution, and alkane standard mix solution can be stored at room temperature (unless stated otherwise). *N*-Methyl-*N*-trimethylsilyltrifluor(*o*)acetamide (MSTFA) must be stored in a 4 °C fridge. All waste disposal instructions must be followed.

2.1 Extraction

1. 2:1 Methanol-Chloroform: In a 1000 mL graduated cylinder, measure 1000 mL of methanol (with >99.8% purity), and transfer to a glass beaker. Measure 500 mL of chloroform (with >99.9% purity). Transfer to same glass beaker as methanol. Store at room temperature.
2. 1:1 Chloroform H_2O: In a 1000 mL graduated cylinder, measure 500 mL of ultrapure, deionized H_2O, and transfer to a glass beaker. Measure 500 mL of chloroform (with >99.9% purity). Transfer to same glass beaker as deionized H_2O. Store at room temperature.
3. Alkane Standard Mix Solution: Weigh 3 mg each of n-Docosane, n-Hexacosane, n-Nonadecane, and n-Triacontane into glass vials. Dissolve each in 1 mL of hexane (with purity >99.9%). Measure 4 μL each of n-Decane, n-Dodecane, and n-Pentadecane into glass vials. Dilute each in 1 mL of hexane (with purity >99.9%). Transfer each solution to one glass vial. Vortex to ensure mixture. Store at room temperature.

2.2 Derivatization

1. *N*-Methyl-*N*-trimethylsilyltrifluor(*o*)acetamide (MSTFA). Store in 4 °C fridge.
2. Weigh 20 μg of methoxyamine hydrochloride. Transfer to a glass vial. Using a syringe, measure out 4 mL of pyridine, and deposit in the glass vial (*see* **Note 3**). Vortex to ensure mixture.

3 Methods

There are diverse applications for GC-MS including drug detection and identifying the metabolic effects of toxins or disease [20–22]. Typically, experiments designed to identify metabolic alterations associated with a disease state will identify a pattern of changes in metabolites [23]. This metabolomic profile could represent a biomarker or a meta-biomarker [24, 25]. If the intent of the experiment is biomarker development, then experimental design hinges on the end use of the biomarker. Metabolomic studies can

be performed on any biofluid, including serum, plasma, urine, etc. [26, 27]. Sometimes, the choice of biofluid is a pragmatic one, dependent on sample availability. Ideally, the sample type selected should be the most appropriate for addressing the clinical need. If a clinical end use is considered, then convenient, minimally invasive tests would be most widely accepted by a patient. Urine and serum (or plasma) collections are standards of practice and could easily be implemented in a clinical lab.

A suitable study cohort must be available. Ideally, the cohort should be designed prospectively, before sample procurement. There are some important features of the study cohort that should be considered. First, the samples must be a representative of the end use. For example, if a disease biomarker is being developed, then the study cohort should reflect the features of the population. Features to consider include gender distribution, ethnicity, and environmental features including diet and time of sampling. This is a particularly important experimental design consideration for metabolomic experiments, where the metabolome can be affected by genetic and environmental factors. Second, an appropriate control must be available. For a disease biomarker, that would consist of a disease-free control; for studies on drug toxicities, controls might consist of individuals on the drug who do and do not have the toxicity; for studies on the metabolic effects of drugs, appropriate controls are individuals who did not receive the drug. It is critical that the sample collection and storage conditions in comparator groups are exactly the same. Third, confounders must be considered. The experimental and control groups should be age- and gender-matched. Comorbidities that may systematically bias results should be considered, and this includes diabetes mellitus and other prevalent diseases with obvious and gross metabolic effects.

The experimental design should include separate discovery (training) experiments and validation steps. If we have sufficient samples, we randomly assign 2/3 of the samples to the discovery set and 1/3 to an independent validation cohort. A metabolomic model is constructed after testing the discovery cohort, and then the model is tested in the validation cohort to see if it accurately predicts the condition (i.e., experimental vs. control). Subsequent studies on a separate prospectively collected cohort would further validate the biomarker.

In order to accurately and reproducibly quantify metabolites on any analytical platform based on MS, internal standards are required. For GC-MS, deuterated metabolites represent the best internal controls. Internal controls should at least represent each chemical class of metabolites that will be considered and should also be representative of the range of retention indices of metabolites targeted for analysis. Internal standards are designed to reduce measurement variation that is introduced through sample

preparation, as well as variations in ionization efficiency and chromatographic retention. Therefore, internal standards should be introduced to each sample prior to the extraction phase of the preparation procedure (*see* **Note 4**).

In some instances, large numbers of samples will be tested, requiring analysis in multiple separate batches depending on instrument capacity. The distribution of samples to different analytical batches must be carefully considered, at least in the discovery phase. One approach is to randomly assign samples to batches but stratifying comparator groups. For example, each batch should contain equal numbers from the experimental and control groups; each batch could be further stratified by ensuring a similar age and gender distribution.

For the extraction procedure for metabolites, we use the method described by Bligh and Dyer [28], which separates metabolites into two phases of water (non-lipids) and chloroform (lipids). We have described a customized workflow below, derived from the protocol described by Migne et al. [29].

3.1 Extraction Experimentation: Preparation of Samples and Addition of Internal Standards

1. Prepare an ice bucket. Take the serum samples from the −80 °C freezer, and place them in the bucket to thaw. While serum samples are thawing, prepare the 2:1 methanol-chloroform by placing it into the ice bucket to cool.
2. Prepare compounds that will constitute internal standard solution. Mark and set aside four vials for each internal standard dilution mixture, and label them A, B, C, and D.
3. While the 2:1 methanol-chloroform cools and the serum samples thaw, label three sets of 2 mL microcentrifuge tubes: #1 for sample, #2 for upper fraction (UF, aqueous layer), and #3 for lower fraction (LF, organic layer). Label three sets of 2 mL tubs: #1 for mixture solutions (A, B, C, D), #2 for UF, and #3 for LF.
4. Transfer 50 μL of biofluid samples into #1 labeled tubes. Transfer 50 μL of internal standard solutions (A, B, C, and D) into #1 labeled tubes.
5. Add 300 μL of cold 2:1 methanol-chloroform and 50 μL of mixture solution B to each serum sample, and securely close the cap. Do not add to vials labeled A, B, C, and D.
6. Add 300 μL of cold 2:1 methanol-chloroform to all vials.
7. Vortex each sample until a homogenous solution is formed.
8. Remove each tube, check all caps are secure, and place into a floating rack.
9. Sonicate for at least 15 min to ensure a good mix.
10. Add 100 μL from each layer of the 1:1 chloroform-H_2O solution into each tube.

11. Centrifuge for 7 min at 16,600 × *g* (*see* **Note 5**).
12. Remove a batch of tubes and restart the centrifuge.
13. Carefully pipette the upper fraction (c. 200 μL), ensuring that none of the lower fraction, or pellet, is also present. Transfer this upper fraction into #2 labeled tubes.
14. Without greatly disturbing the pellet, remove all remaining liquid. This will contain the lower fraction and some upper fraction. Briefly allow this to settle in the pipette tip and expel into tubes #3 for lower fraction and #2 for upper fraction as relevant.
15. Repeat **steps 11–14** for all samples.
16. Load upper fraction tubes into the SpeedVac, and evaporate to dryness for 8 h (the time needed depends on the instrument). The upper fraction samples contain the aqueous metabolites suitable for gas chromatography-mass spectrometry or NMR analysis.
17. Put the tubes containing pellets and those containing lower fraction into a fume hood with caps open to allow for evaporation. These lower fraction organic metabolites are suitable for FAME or liquid chromatography-mass spectrometry lipid analysis and can be stored at −20 °C.

3.2 Derivatization Process of Samples for GC-MS Run

The derivatization step is a key step used to improve the reproducibility of experimental results in GC-MS analysis. Derivatization increases thermal stability and volatility, allowing the metabolites to pass through the gas chromatograph. Derivatization should improve the resolution between overlapping peaks and coeluting metabolites [30, 31]. The process entails the chemical transformation of metabolites within the matrix through the addition of a derivatizing agent, such as *N-methyl-N-trimethylsilyltrifluor(o) acetamide*.

1. Prepare ~20 mg/mL methoxyamine hydrochloride in pyridine.
2. Add 50 μL of 20 mg/mL methoxyamine hydrochloride/pyridine to each dried aqueous sample.
3. Shake at 37 °C for 2–3 h.
4. Add 50 μL of MSTFA to each sample. Be sure to seal remaining MSTFA with parafilm and store at 4 °C fridge.
5. Shake at 37 °C for an additional 30–60 min.
6. Dilute each sample with 500 μL of hexane.
7. Centrifuge all samples at 16,600 × *g* for 4 min to remove any solid particle. This step is important even if no particles are visible by the eye, as microparticles can interfere with the GC-MS.

8. Transfer 200 μL of supernatant to GC-MS vial with glass insert.
9. Make quality control samples (100 μL of every disease-free control sample (DFC)) into a giant 1000 μL tube. Distribute into five 200 μL glass vials to put in with every solvent-only (blank) and standard sample/vial (i.e., after ~every 10 samples) (*see* **Note 6**).
10. Create a sample list file. The sample list should contain all samples to be run through GC-MS, including samples, internal standard concentration vials, retention time standard samples, and solvent-only samples.
11. Place vials onto rack in sample list order, and begin run. Please refer to the instrument's instruction manual, as details vary based on the instrument's brand and the technology used.
12. Check the saturated peaks on the GC-MS, and continue with the adjusted concentration for the rest of your samples. Ideally, peaks from metabolites of interest should be well inside the instrument's dynamic range and should not saturate the detector if they are in the biologic range of concentration.

3.3 Analysis of Spectra

Metabolomic studies that utilize the GC-MS platform aim to identify and quantify (either the relative or absolute) levels of metabolites in complex matrices such as serum. GC-MS is capable of detecting small differences in abundance of metabolites, which makes it an excellent platform for looking at differential metabolite patterns between experimental conditions (or experimental groups). Ions from individual metabolites can be identified based on retention pattern, atomic mass, and charge. When signals from different ions (mass/charge values) are found at the same retention time, observed fragments can be assumed to originate from the same metabolite (or of the similar chemical structure). However, retention times of metabolites can vary based on factors such as the chromatography column and other features of the mass spectrometer. Retention index is used for standardization of retention times across a variety of experimental conditions and is obtained through logarithmic adjustment that relates retention time of the sample to the retention time of a standard that elutes before and after the peak of the sample. This provides a value that is system-independent and can therefore be used to compare peak intensities across many platforms (*see* **Note 7**).

To identify metabolites that are differentially abundant between experimental conditions, mass spectral data must first be processed. Software such as *MetaboliteDetector*, *SIMAT*, and *OpenChrom* are freely available tools to use for both targeted and non-targeted comprehensive analysis of spectra produced by GC-MS [32]. The following section will describe a *general* workflow for the metabolomic data processing using the software **MetaboliteDetector**.

3.4 Import and Calibrate

1. Using the software **provided by instrument's manufacturer**, convert the raw files generated by the GC-MS into CDF files.
2. *Import* sample CDF files into **MetaboliteDetector.** After importing the files, the folder containing the .CDF files will now contain .bin and .idx files.
3. Open *Tools*, and click on *RI-Calibration Wizard*. This step will standardize retention times across a variety of experimental conditions, including different instruments, columns, carrier gases, etc. This will allow for all compounds detected on chromatogram to be searched and referenced in a library for metabolite identification (*see* Fig. 1). Most of these libraries are public and specific for GC-MS metabolite analysis (e.g., Golm Metabolome Database [33]). Select one of the standards from the list containing your samples.
4. The compounds in the reference chromatogram will be detected. Open the *Settings*, and choose a high threshold value, so that the software may detect the compounds present in your standard.
5. The standard you chose should have a distinct number of peaks, depending on the standard chosen. If this is not the case, change the standard sample file as multiple standard vials should have been run (*see* Fig. 2).
6. A calibration table will appear. Ensure that the retention time values for the standard peaks are correct. This is done by clicking on one of the standard peaks, which brings up a window showing detected peaks. Check that the suggested retention time is the same as the retention time you see in the tables as you select the triangles below the peaks. Make corrections as necessary by changing the values in the calibration table—the numbers identified will automatically appear. This step is also necessary if no retention time has been suggested.
7. Choose all sample and blank chromatograms in .bin format.

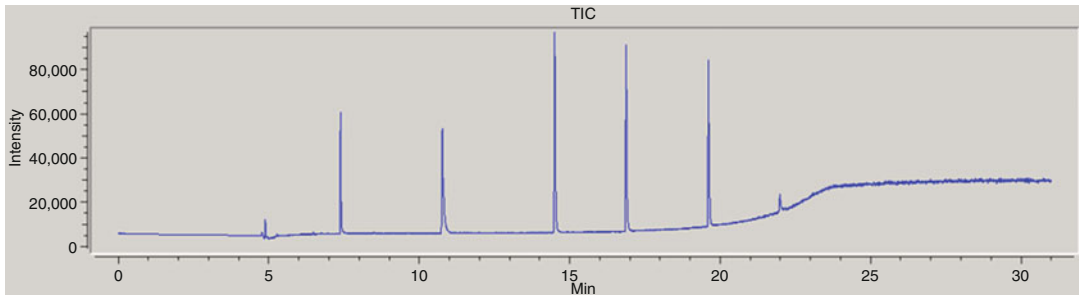
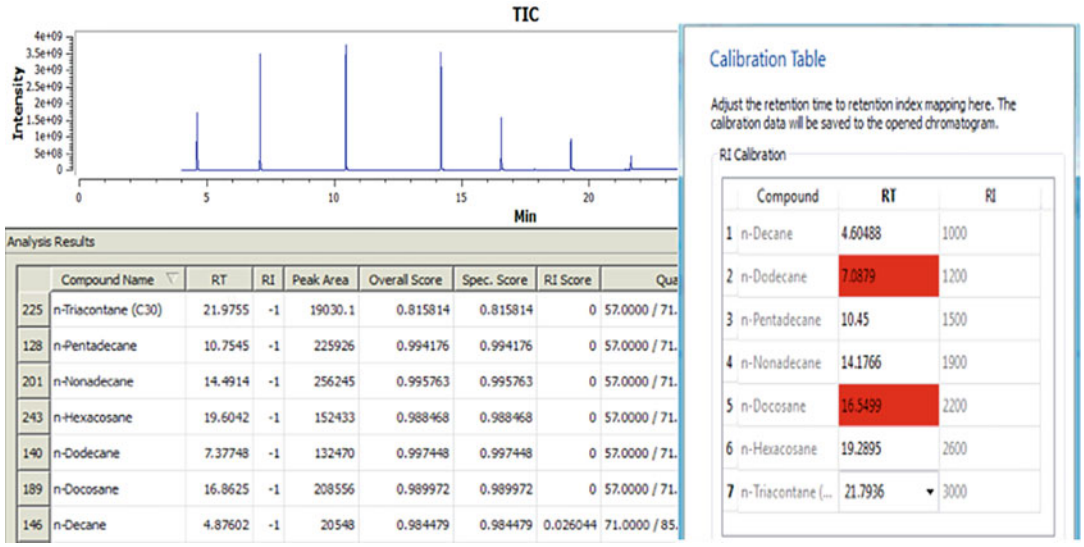


Fig. 1 A chromatogram of a standard 7-alkane solution used for retention times normalization and generation of Retention Index (RI). A wide range of time is covered by these alkanes, which provide a comprehensive basis for times standardization



	Compound Name	RT	RI	Peak Area	Overall Score	Spec. Score	RI Score	Qua
225	n-Triacontane (C30)	21.9755	-1	19030.1	0.815814	0.815814	0	57.0000 / 71.
128	n-Pentadecane	10.7545	-1	225926	0.994176	0.994176	0	57.0000 / 71.
201	n-Nonadecane	14.4914	-1	256245	0.995763	0.995763	0	57.0000 / 71.
243	n-Hexacosane	19.6042	-1	152433	0.988468	0.988468	0	57.0000 / 71.
140	n-Dodecane	7.37748	-1	132470	0.997448	0.997448	0	57.0000 / 71.
189	n-Docosane	16.8625	-1	208556	0.989972	0.989972	0	57.0000 / 71.
146	n-Decane	4.87602	-1	20548	0.984479	0.984479	0.026044	71.0000 / 85.

	Compound	RT	RI
1	n-Decane	4.60488	1000
2	n-Dodecane	7.0879	1200
3	n-Pentadecane	10.45	1500
4	n-Nonadecane	14.1766	1900
5	n-Docosane	16.5499	2200
6	n-Hexacosane	19.2895	2600
7	n-Triacontane (...	21.7936	3000

Fig. 2 A screenshot of the assignment procedure for seven alkane compounds in the RI standard solution in Metabolite Detector. Retention times and nominal retention indices (RI) can be found in the "Analysis Results" table

8. Check that all chromatograms are present. Remove any unwanted duplicates.
9. Next, open *Settings* and change the threshold from very high to a low value, so that the software may detect compounds with much lower intensity. Ensure this threshold value is appropriate in that it may identify and pick up all peaks of your interest.
10. Allow calibration to complete.

3.5 Compound Detection

1. Open *Settings*, and add mass filters for known contaminant ions that should not be identified by the software and that you want to remove. Example of such ions includes 73.00 and 146.0 in MSTFA-derivatized samples.
2. Select *Tools*, and open *Compound Detection Wizard*.
3. Choose all chromatograms of the sample set, and click "redetect all compounds." Allow software to complete detection.

3.6 Batch Quantification

1. Open *Tools* and select *Batch Quantification*. Unless samples were run in different batches, create one group containing all samples.
2. Depending on which type of analysis is preferred to be done, the software can complete both targeted and nontargeted analyses. For nontargeted analysis, select this option, and allow the software to complete this step. For targeted analysis, see **MetaboliteDetector** instruction for further instruction on how to prepare a list of metabolites for *targeted profiling*.

3. Detection and quantification results are now present in a tab on the bottom left corner of the screen. For metabolite detection, libraries such as Golm Metabolome Database, NIST, and MassBank mass spectral database may be used to create in-house libraries for identification [34, 35].
4. Export results into a .CSV file.

3.7 Data Cleaning and Preprocessing

Normalization of sample chromatographic results is necessary to compensate for sample-to-sample variation in the overall peak intensities. There are several approaches to normalization, including the incorporation of an internal standard (fully quantitative approach) or applying a calculated scaling factor such as probabilistic quotient normalization (PQN) [36] (semiquantitative approach). PQN is a robust normalization method that accounts for dilution of complex biological matrices, such as serum. PQN identifies the most common scale factor between spectra and divides the peak values for each spectrum by this scale factor. This approach reduces the influence of missing peaks in individual spectra and can be readily done in an **R-based application** or within **Microsoft Excel**.

Missing data are problematic. Replacing them with a numerical value of 0 is misleading and can significantly skew results, resulting in inaccurate quantitative and qualitative representations of metabolites when submitting them as part of a multivariate analysis. Therefore, methods for data imputation are more appropriate. Various approaches exist for data imputation, such as mean, singular variable decomposition, and half minimum [30], and should take into account whether missing values are random or systematic.

Scaling is particularly important in biomarker investigation and analysis of complex samples. Scaling allows for spectrum-distinguishing features that are often not the largest peaks or considered part of background noise. Scaling provides a mechanism for the enhancement of the contribution of the medium peaks, without inflating the background noise. There are different approaches for scaling, including unit variance and Pareto scaling, which can both be done in SIMCA (Umetrics AB, Malmö, Sweden) [37].

Intra-batch and inter-batch corrections are also essential. Within a particular batch, instrument drift and other technical factors may affect the appearance of the MS spectrum. The intermittent addition of pooled quality control samples at regular intervals within the batch enables one to monitor for unexpected perturbations. If there is a systematic drift in the peak intensities, then the quality control samples can aid in correcting for the drift. When doing larger sample sets, it may be impossible to run all of the samples in a single batch (on the same day). This compounds any potential measurement error. Batches are subject to systematic and random variations that may be quite profound. In addition to

variability in metabolite ion peak intensities that can be introduced by variations in extraction and derivatization, batches run on different days are subject to a myriad of instrument-dependent technical factors. One approach to correct for batch-dependent variation is the use of the ComBat method (part of the *sva* package in R) which can identify and estimate surrogate variables for unknown sources of variation and remove that from the high-dimensional dataset [38]. The need for any correction is diminished with the addition of internal standards, which should allow comparison between samples within a batch and between batches.

3.8 Statistical Analysis

There are multiple approaches to analyzing metabolomic data [39–41]. Typically, we perform an unsupervised principal component analysis (PCA) to identify any latent features or patterns that distinguish the experimental group and the control group [42]. We then apply a filtration step that removes the "noise" comprised of non-related compounds that are clearly no different between the experimental and control groups. Subsequently, a supervised multivariate analysis can be applied to identify metabolites that significantly contribute to the model that distinguishes the experimental group from the control group. For example, supervised orthogonal partial least squares discriminant analysis (OPLS-DA) can be used to reduce the dimensionality of the data and to identify important metabolites that distinguish comparator groups. Commercial multivariate analysis software, such as **SIMCA** (version 15.0.0; Umetrics AB, Malmö, Sweden), or open-source software such as packages in R are available (https://cran.r-project.org/web/views/ChemPhys.html).

3.9 Interpretation

Once the metabolomic model is finalized (and perhaps validated), it is important to understand the underlying biological significance of the pattern of differentially abundant metabolites. Data software packages can be used to interpret metabolomic data. By inputting the data generated from multivariate analysis into software such as **MetaboAnalyst** and **Pathview**, one can see the components of primary metabolic pathways that are perturbed [43–46].

4 Notes

1. The methods described focus on serum as the biological matrix, which requires little sample processing. Urine testing is similarly relatively straightforward. Analysis of solid tissues, which would be expected to contain higher concentrations of metabolites, requires additional sample processing and optimization. The methods in this chapter focus on the analysis of polar metabolites. Efficient elution of nonpolar components is optimized through different preparation procedures [16–19].

2. GC-MS is a sensitive method for identifying differentially abundant metabolites in two or more groups of experimental samples. GC-TOF-MS improves the resolution of individual metabolites. Good experimental design is perhaps the most important feature of a successful and valid comparison between groups of samples. Ideally, internal standards and the addition of intermittent quality control samples will be included in the experimental design. This is especially important for larger data sets, in which multiple batches are required to analyze all of the samples.
3. Use a syringe instead of a pipette to get pyridine from the stock to avoid contamination of the stock solution with plasticizers.
4. The metabolites that make up the internal standard solution are the compounds that represent different physiochemical classes to cover the portion of the metabolome that you wish to analyze. Examples would include glucose and phenylalanine. Once your compounds are chosen and the concentration gradients have been selected based on the dynamic range of quantification, choose four concentration dilutions that will cover the concentration gradient, and put them into vials A, B, C, and D. This will enable comparison of each compound found in your samples to these solutions, based on the gradients. For example, the highest concentration of the internal standard mix can be assigned to vial A, while the lowest concentration can be assigned to vial D. These internal standards should be run with every batch for comparison.
5. The outlined number is a minimum g force that should be applied for the separation of particles in the size range from serum (as the matrix). Using a greater g force will not damage the samples or separation—on the contrary, the separated particles will make a tighter layer that cannot be disturbed easily.
6. A blank is an analyte-free solvent which is carried through the analysis to examine and clean any contamination or carry-over in the experimental procedure. If not used, contamination can be carried through the whole analysis procedure. A blank vial should be placed every 5–10 samples in the run of the GC-MS.
7. The methods above describe the characteristics of a discovery phase of experiments, which involves the untargeted (or exploratory) analysis of metabolites. In such an experiment, there is a scan of the full spectrum, considering all peaks in a spectrum over a large mass-to-charge range. After discovery and training, it is possible to perform selective ion monitoring for targeted analysis.

References

1. Denkert C, Budczies J, Kind T, Weichert W, Tablack P, Sehouli J, Niesporek S, Konsgen D, Dietel M, Fiehn O (2006) Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors. Cancer Res 66(22):10795–10804
2. Dettmer K, Aronov PA, Hammock BD (2007) Mass spectrometry-based metabolomics. Mass Spectrom Rev 26(1):51–78
3. He J, Luo Z, Huang L, He J, Chen Y, Rong X, Jia S, Tang F, Wang X, Zhang R, Zhang J, Shi J, Abliz Z (2015) Ambient mass spectrometry imaging metabolomics method provides novel insights into the action mechanism of drug candidates. Anal Chem 87(10):5372–5379. https://doi.org/10.1021/acs.analchem.5b00680
4. Claudino WM, Quattrone A, Biganzoli L, Pestrin M, Bertini I, Di Leo A (2007) Metabolomics: available results, current research projects in breast cancer, and future applications. J Clin Oncol 25(19):2840–2846
5. Di Lena M, Travaglio E, Altomare DF (2013) Metabolomics: a potential powerful ally in the fight against cancer. Color Dis 16(4):235–238. https://doi.org/10.1111/codi.12523
6. Farshidfar F, Weljie AM, Kopciuk K, Buie WD, Maclean A, Dixon E, Sutherland FR, Molckovsky A, Vogel HJ, Bathe OF (2012) Serum metabolomic profile as a means to distinguish stage of colorectal cancer. Genome Med 4(5):42. https://doi.org/10.1186/gm341
7. Bayet-Robert M, Loiseau D, Rio P, Demidem A, Barthomeuf C, Stepien G, Morvan D (2010) Quantitative two-dimensional HRMAS 1H-NMR spectroscopy-based metabolite profiling of human cancer cell lines and response to chemotherapy. Magn Reson Med 63(5):1172–1183. https://doi.org/10.1002/mrm.22303
8. Beckonert O, Keun HC, Ebbels TM, Bundy J, Holmes E, Lindon JC, Nicholson JK (2007) Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat Protoc 2(11):2692–2703
9. Bertini I, Cacciatore S, Jensen BV, Schou JV, Johansen JS, Kruhoffer M, Luchinat C, Nielsen DL, Turano P (2012) Metabolomic NMR fingerprinting to identify and predict survival of patients with metastatic colorectal cancer. Cancer Res 72(1):356–364. https://doi.org/10.1158/0008-5472.CAN-11-1543
10. Brindle JT, Nicholson JK, Schofield PM, Grainger DJ, Holmes E (2003) Application of chemometrics to 1H NMR spectroscopic data to investigate a relationship between human serum metabolic profiles and hypertension. Analyst 128(1):32–36
11. Lodi A, Ronen SM (2011) Magnetic resonance spectroscopy detectable metabolomic fingerprint of response to antineoplastic treatment. PLoS One 6(10):e26155. https://doi.org/10.1371/journal.pone.0026155
12. Jonsson P, Bruce SJ, Moritz T, Trygg J, Sjostrom M, Plumb R, Granger J, Maibaum E, Nicholson JK, Holmes E, Antti H (2005) Extraction, interpretation and validation of information for comparing samples in metabolic LC/MS data sets. Analyst 130(5):701–707. https://doi.org/10.1039/b501890k
13. Nicholson JK, Lindon JC, Scarfe GB, Wilson ID, Abou-Shakra F, Sage AB, Castro-Perez J (2001) High-performance liquid chromatography linked to inductively coupled plasma mass spectrometry and orthogonal acceleration time-of-flight mass spectrometry for the simultaneous detection and identification of metabolites of 2-bromo-4-trifluoromethyl. Anal Chem 73(7):1491–1494
14. Begley P, Francis-McIntyre S, Dunn WB, Broadhurst DI, Halsall A, Tseng A, Knowles J, Goodacre R, Kell DB (2009) Development and performance of a gas chromatography-time-of-flight mass spectrometry analysis for large-scale nontargeted metabolomic studies of human serum. Anal Chem 81(16):7038–7046. https://doi.org/10.1021/ac9011599
15. Dorman FL, Overton EB, Whiting JJ, Cochran JW, Gardea-Torresdey J (2008) Gas chromatography. Anal Chem 80(12):4487–4497
16. Ichihara K, Fukubayashi Y (2010) Preparation of fatty acid methyl esters for gas-liquid chromatography. J Lipid Res 51(3):635–640. https://doi.org/10.1194/jlr.D001065
17. Roberts LD, McCombie G, Titman CM, Griffin JL (2008) A matter of fat: an introduction to lipidomic profiling methods. J Chromatogr B 871(2):174–181. https://doi.org/10.1016/j.jchromb.2008.04.002
18. Quehenberger O, Armando AM, Dennis EA (2011) High sensitivity quantitative lipidomics analysis of fatty acids in biological samples by gas chromatography–mass spectrometry. Biochimica Biophys Acta 1811(11):648–656. https://doi.org/10.1016/j.bbalip.2011.07.006

19. Lagerstedt SA, Hinrichs DR, Batt SM, Magera MJ, Rinaldo P, McConnell JP (2001) Quantitative determination of plasma C8–C26 Total fatty acids for the biochemical diagnosis of nutritional and metabolic disorders. Mol Genet Metab 73(1):38–45. https://doi.org/10.1006/mgme.2001.3170
20. Gowda GA, Zhang S, Gu H, Asiago V, Shanaiah N, Raftery D (2008) Metabolomics-based methods for early disease diagnostics. Expert Rev Mol Diagn 8(5):617–633
21. Halama A, Riesen N, Moller G, Hrabe de Angelis M, Adamski J (2013) Identification of biomarkers for apoptosis in cancer cell lines using metabolomics: tools for individualized medicine. J Intern Med 274(5):425–439. https://doi.org/10.1111/joim.12117
22. Miolo G, Muraro E, Caruso D, Crivellari D, Ash A, Scalone S, Lombardi D, Rizzolio F, Giordano A, Corona G (2016) Pharmacometabolomics study identifies circulating spermidine and tryptophan as potential biomarkers associated with the complete pathological response to trastuzumab-paclitaxel neoadjuvant therapy in HER-2 positive breast cancer. Oncotarget 7(26):39809–39822. https://doi.org/10.18632/oncotarget.9489
23. Jobard E, Blanc E, Negrier S, Escudier B, Gravis G, Chevreau C, Elena-Herrmann B, Tredan O (2015) A serum metabolomic fingerprint of bevacizumab and temsirolimus combination as first-line treatment of metastatic renal cell carcinoma. Br J Cancer 113 (8):1148–1157. https://doi.org/10.1038/bjc.2015.322
24. Kell DB (2007) Metabolomic biomarkers: search, discovery and validation. Expert Rev Mol Diagn 7(4):329–333. https://doi.org/10.1586/14737159.7.4.329
25. Kiehntopf M, Nin N, Bauer M (2013) Metabolism, metabolome, and metabolomics in intensive care: is it time to move beyond monitoring of glucose and lactate? Am J Respir Crit Care Med 187(9):906–907. https://doi.org/10.1164/rccm.201303-0414ED
26. Napoli C, Sperandio N, Lawlor RT, Scarpa A, Molinari H, Assfalg M (2012) Urine metabolic signature of pancreatic ductal adenocarcinoma by (1)h nuclear magnetic resonance: identification, mapping, and evolution. J Proteome Res 11(2):1274–1283. https://doi.org/10.1021/pr200960u
27. Nishiumi S, Kobayashi T, Ikeda A, Yoshie T, Kibi M, Izumi Y, Okuno T, Hayashi N, Kawano S, Takenawa T, Azuma T, Yoshida M (2012) A novel serum metabolomics-based diagnostic approach for colorectal cancer. PLoS One 7(7):e40459. https://doi.org/10.1371/journal.pone.0040459
28. Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Can J Biochem Physiol 37(8):911–917. https://doi.org/10.1139/o59-099
29. Migne C, Durand S, Pujos-Guillot E (2018) Exploratory GC/MS-Based Metabolomics of Body Fluids. Clinical Metabolomics: 239–246
30. Lien SK, Kvitvang HF, Bruheim P (2012) Utilization of a deuterated derivatization agent to synthesize internal standards for gas chromatography-tandem mass spectrometry quantification of silylated metabolites. J Chromatogr A 1247:118–124. https://doi.org/10.1016/j.chroma.2012.05.053
31. Liu T, Zhang J, Wang X, Yang J, Tang Z, Lu J (2014) Radiolabeled glucose derivatives for tumor imaging using SPECT and PET. Curr Med Chem 21(1):24–34
32. Hiller K, Hangebrauk J, Jager C, Spura J, Schreiber K, Schomburg D (2009) MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis. Anal Chem 81 (9):3429–3439. https://doi.org/10.1021/ac802689c
33. Hummel J, Selbig J, Walther D, Kopka J (2007) The Golm Metabolome Database: a database for GC-MS based metabolite profiling. In: Nielsen J, Jewett MC (eds) Metabolomics: a powerful tool in systems biology. Springer, Berlin, pp 75–95. https://doi.org/10.1007/4735_2007_0229
34. Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, Ojima Y, Tanaka K, Tanaka S, Aoshima K, Oda Y, Kakazu Y, Kusano M, Tohge T, Matsuda F, Sawada Y, Hirai MY, Nakanishi H, Ikeda K, Akimoto N, Maoka T, Takahashi H, Ara T, Sakurai N, Suzuki H, Shibata D, Neumann S, Iida T, Tanaka K, Funatsu K, Matsuura F, Soga T, Taguchi R, Saito K, Nishioka T (2010) MassBank: a public repository for sharing mass spectral data for life sciences. J Mass Spectrom 45(7):703–714. https://doi.org/10.1002/jms.1777
35. Stein SE (1995) Chemical substructure identification by mass spectral library searching. J Am Soc Mass Spectrom 6(8):644–655. https://doi.org/10.1016/1044-0305(95)00291-k
36. Dieterle F, Ross A, Schlotterbeck G, Senn H (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. Anal Chem 78 (13):4281–4290. https://doi.org/10.1021/ac051632c

37. Farshidfar F, Kopciuk KA, Hilsden R, McGregor SE, Mazurak VC, Buie WD, MacLean A, Vogel HJ, Bathe OF (2018) A quantitative multimodal metabolomic assay for colorectal cancer. BMC Cancer 18(1):26
38. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics (Oxford, England) 28(6):882–883. https://doi.org/10.1093/bioinformatics/bts034
39. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B 57(1):289–300
40. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2(5):401–404. https://doi.org/10.1158/2159-8290.CD-12-0095
41. Smilde AK, Jansen JJ, Hoefsloot HC, Lamers RJ, van der Greef J, Timmerman ME (2005) ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data. Bioinformatics (Oxford, England) 21(13):3043–3048. https://doi.org/10.1093/bioinformatics/bti476
42. Gu H, Pan Z, Xi B, Asiago V, Musselman B, Raftery D (2011) Principal component directed partial least squares analysis for combining nuclear magnetic resonance and mass spectrometry data in metabolomics: application to the detection of breast cancer. Anal Chim Acta 686(1–2):57–63. https://doi.org/10.1016/j.aca.2010.11.040
43. Xia J, Wishart, DS (2011) Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst Nature Protocols 6(6), 743–760
44. Luo W, Brouwer C (2013) Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 15(29):1830–1831
45. Xia J, Broadhurst DI, Wilson M, Wishart DS (2013) Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics 9(2):280–299. https://doi.org/10.1007/s11306-012-0482-9
46. Xia J, Wishart DS (2011) Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. Nat Protoc 6(6):743–760. https://doi.org/10.1038/nprot.2011.319

Chapter 14

Metabolite Profiling of Clinical Cancer Biofluid Samples by NMR Spectroscopy

Beata Mickiewicz, M. Eric Hyndman, and Hans J. Vogel

Abstract

Metabolomics is a comprehensive characterization of the small polar molecules (metabolites) in different biological systems. One of the analytical platforms commonly used to study metabolic alterations in biofluid samples is proton nuclear magnetic resonance (^{1}H NMR) spectroscopy. NMR spectroscopy is very specific, quantitative, and highly reproducible. Moreover, sample preparation for NMR experiments is very simple and straightforward, and this gives NMR spectroscopy a distinct advantage over other metabolic profiling methods. It has already been shown that ^{1}H NMR-based profiling of biological fluids can be effective in differentiating benign from malignant lesions and in investigating the efficacy of specific cancer treatments. Therefore, ^{1}H NMR spectroscopy may become a promising tool for early noninvasive diagnosis and rapid assessment of treatment effects in cancer patients. Here, we describe a detailed protocol for ^{1}H NMR metabolite profiling in serum, plasma, and urine samples, including sample collection procedures, sample preparation for ^{1}H NMR experiments, spectral acquisition and processing, and quantitative profiling of ^{1}H NMR spectra. We also discuss several aspects of appropriate study design and some multivariate statistical methods that are commonly used to analyze metabolomics datasets.

Key words Metabolomics, ^{1}H NMR, Quantitative profiling, Cancer, Serum, Plasma, Urine

1 Introduction

Metabolomics is a rapidly evolving division of the "omics" sciences. It represents the small-molecule changes induced by upstream alterations in the genome, transcriptome, and proteome. Genetic mutations in cell cycle pathways of cancer cells can influence cellular metabolism and therefore metabolic profiles [1–4]. Alternatively, mutations in metabolic enzymes such as in fumarate hydratase, as in papillary renal cell carcinoma, can directly affect metabolism and contribute to a malignant phenotype [5]. Tumor growth, invasion into adjacent tissues, and the immune response often trigger a cascade of cytokines, which, in turn, can affect metabolism of the tumor itself and the surrounding tissue [4, 6]. Unlike genomic changes, the metabolic response to malignancy can be rapid and

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_14,

dependent on many biological processes. This can lead to challenges in differentiating the malignant metabolic signal from the noise, but by being able to detect a rapid alteration in metabolic pathways, it may allow early noninvasive diagnosis and rapid assessment of treatment effects in biological fluids.

It is well known that an appropriate study design for metabolomics-based studies is crucial for obtaining optimal and reliable results. A number of studies have already demonstrated that factors such as sample size, sample collection methods and sample storage conditions, patient's gender, age, genetics, diet, gut microbiota, etc. can have an enormous impact on the quality of metabolic data [7–9]. A well-planned metabolomics study should therefore be designed in such a way that all possible confounding factors are controlled, and thus, any unwanted variance in the metabolomics data is minimized. Consequently, it is important to collect an adequate number of samples, to be consistent in sample collection, storage, and preparation, employ strong controls, and attempt to balance the groups where appropriate. Ideally, the sample size for human-based studies will be $\geq$50 samples per group [10]. Additionally, the sample preparation and sample analysis should be performed in a random order to avoid progressive bias. Furthermore, the results need to be validated in a different, independent cohort to confirm the relevance of metabolic patterns.

Currently, the most widely used analytical platforms in metabolomics studies are nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS). NMR spectroscopy is focused on measuring the magnetic properties of atomic nuclei such as ^{1}H, ^{31}P, or ^{13}C, in various compounds [11]. Each molecule has a unique NMR peak pattern, which can be used to identify and assign spectral signals to specific metabolites. NMR spectroscopy is well known to be a very specific, quantitative, and highly reproducible analytical method [11]. However, the sensitivity of NMR instruments is lower than MS-based platforms. Nonetheless, in our recent studies, the ^{1}H NMR technique was sufficiently sensitive to detect significant changes in metabolic profiles in various biofluid samples collected from cancer patients [12–15]. The NMR technique does not require a special and sophisticated sample preparation protocol. In most cases simply adding H_2O/D_2O and buffer for pH adjustment is sufficient to prepare a biological sample for the NMR experiment [16]. Because of this simplicity, NMR spectroscopy can be used to analyze a wide range of biological fluids, e.g., blood, urine, cerebrospinal fluid, synovial fluid, semen, and pancreatic juice [17–21]. The protocol presented here describes in detail the process for metabolite profiling in serum, plasma, and urine samples as these biofluids can be obtained in a relatively noninvasive manner, and they are the most often studied in our laboratory. Examples of ^{1}H NMR spectra collected for serum and

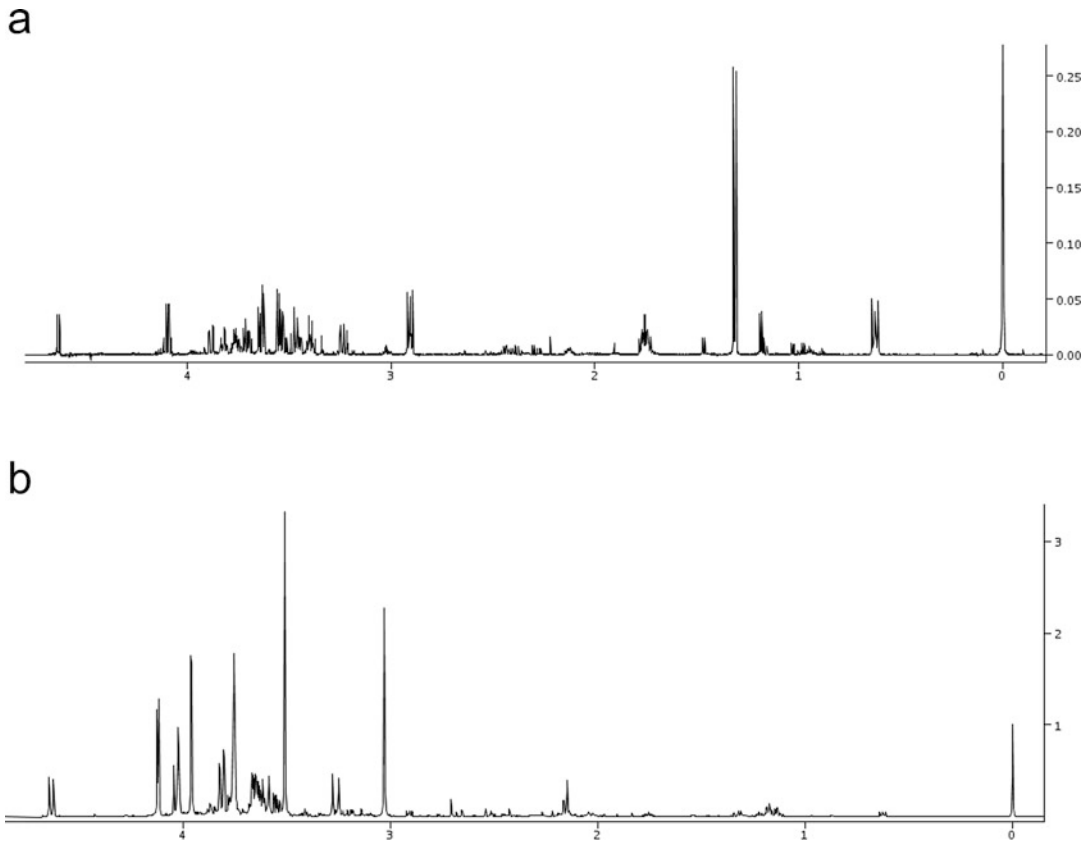


Fig. 1 Typical ^{1}H NMR spectra for (**a**) serum and (**b**) urine samples collected from a bladder cancer patient, obtained in one of our metabolomics studies [40]. Only the aliphatic region of the spectrum is shown. The spectra were obtained on a 600 MHz Bruker Ultrashield Plus NMR spectrometer (Bruker BioSpin Ltd., Canada), equipped with a 5 mm TXI probe at 298 K. A 1-h and 28-min acquisition time for each NMR spectrum was used. The singlet peaks at 0.0 ppm represent the internal reference standard (DSS) and are clearly separated from regular metabolites

urine samples from a bladder cancer patient from one of our metabolomics studies are shown in Fig. 1.

There are in principle two ways to perform metabolite profiling in ^{1}H NMR spectra: spectral binning and quantitative (sometime called "targeted") profiling [22]. Spectral binning consists of dividing the spectrum into a number of regions (bins) of a fixed width. Next, the total area of each bin is calculated and compared between different ^{1}H NMR spectra to identify which bins are changing significantly across the samples. The selected bins are analyzed in order to assign the metabolite identities found in the bin. Although spectral binning is quite fast and provides an efficient way of metabolite profiling, it does not separate metabolic peak patterns that could overlap. For these reasons, some of the metabolite identities can be missed, and significant erroneous differences in metabolite concentrations may be calculated. On the other hand, quantitative

profiling involves the identification of different compounds by their characteristic peak patterns (that are stored as individual metabolite ^{1}H NMR spectra in an external reference database) and assigning these individual clusters based on appropriate regions of the spectrum (Fig. 2) [22]. The concentration of each metabolite is represented by the area under the peak(s). Therefore, by knowing the exact concentration of an internal standard, the relative concentration of all identified compounds can be easily calculated.

Many biological samples, particularly biofluids such as plasma and serum, often contain high molecular weight metabolites (e.g., lipids vesicles or micelles and proteins). These macromolecules give rise to broad signals in the NMR spectra. The metabolite profiling approach is directed specifically toward low molecular weight compounds, and so the broad spectral resonances from lipids and proteins are unwanted as they can interfere with the quantitative analysis of the metabolites in the NMR spectra. Several approaches can be used to facilitate the selective observation of the narrow resonances from small metabolites, such as implementing the Carr-Purcell-Meiboom-Gill (CPMG) NMR pulse sequence, extraction of the sample with organic solvents or perchloric acid, or ultrafiltration (Fig. 3) [23–26]. The CPMG experiment improves the detection of the spectra of the small metabolites but does not completely suppress resonances from mobile lipids, and there is still a considerable overlap in the NMR spectrum from these lipids. Therefore, profiling and metabolite quantification of CPMG spectra remains fraught with difficulties. Extraction and ultrafiltration methods have been studied as well, and it has been shown that ultrafiltration provides high reproducibility and is superior for the removal of high molecular metabolites when compared to extraction methods [27, 28]. Additionally, in our experience ultrafiltration increases the resolution of the NMR spectrum and provides an opportunity to observe aromatic signals, which are typically not visible in a CPMG spectrum. Therefore, the sample preparation protocol that will be described here for metabolite profiling of biofluid samples includes an ultrafiltration step for optimal reduction of broad NMR resonances and improvement of the signal-to-noise ratio.

In addition to 1D ^{1}H NMR, other 2D NMR spectroscopy techniques can be used for metabolite profiling of biofluid samples, for example, homonuclear 2D *J*-resolved NMR spectroscopy [29], correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY), or heteronuclear single-quantum correlation (HSQC) spectroscopy [30, 31]. Overall, these 2D NMR techniques are normally utilized to improve metabolite signal assignment and/or to confirm a prior 1D spectral assignment [32].

The number of small-molecule metabolites that can be detected by 1D ^{1}H NMR in a biofluid sample ranges from a few dozen to hundreds. In order to analyze such big and very often heterogenic datasets, specific statistical analysis tools need to be

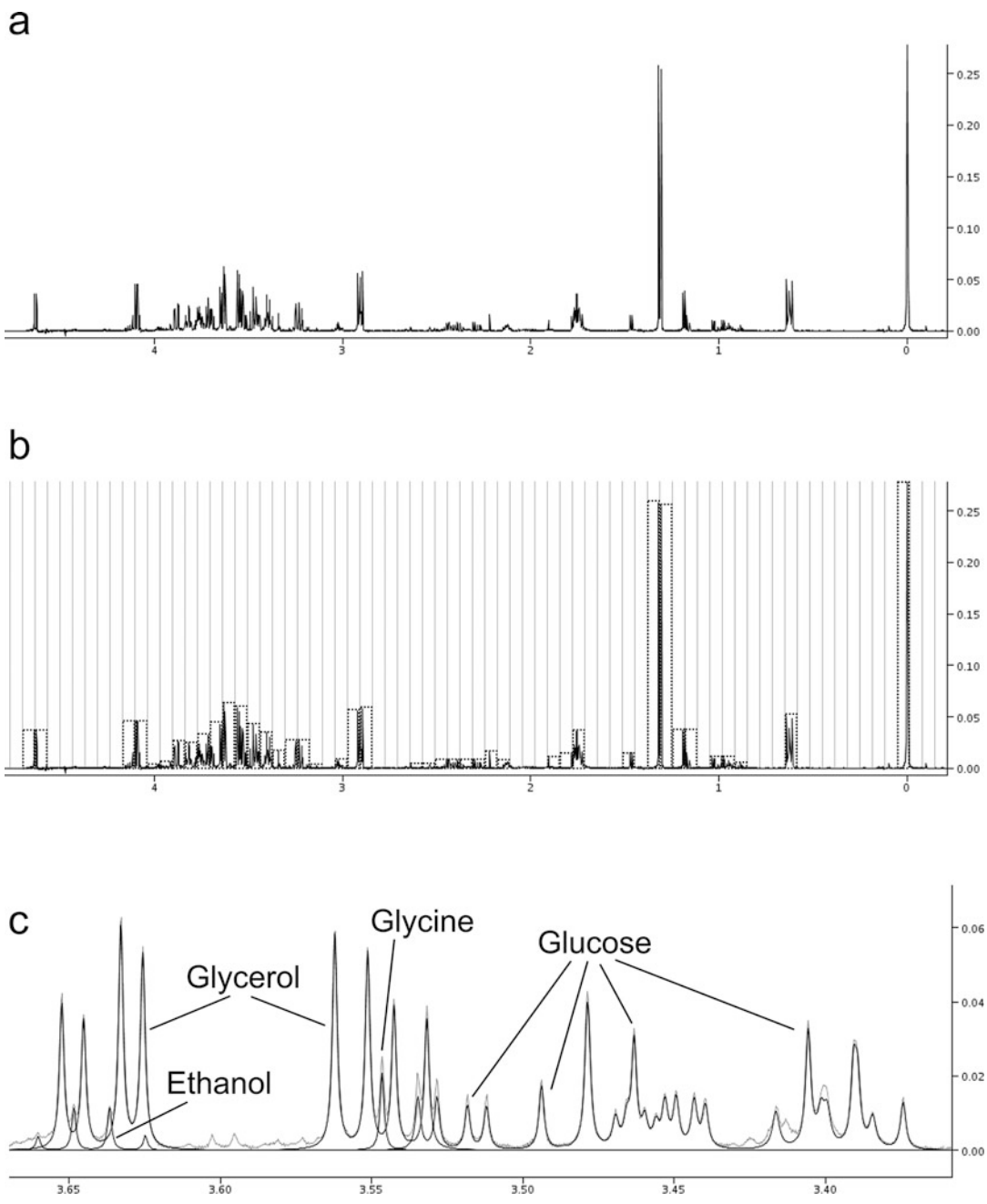


Fig. 2 Examples of metabolite profiling of ^{1}H NMR spectra: (**a**) and (**b**) spectral binning and (**c**) quantitative ("targeted") profiling. Spectral binning is based on dividing the NMR spectrum (**a**) into a number of regions (bins) of a fixed width, (**b**) while quantitative profiling (**c**) is used to identify different compounds by their characteristic peak shift assignment

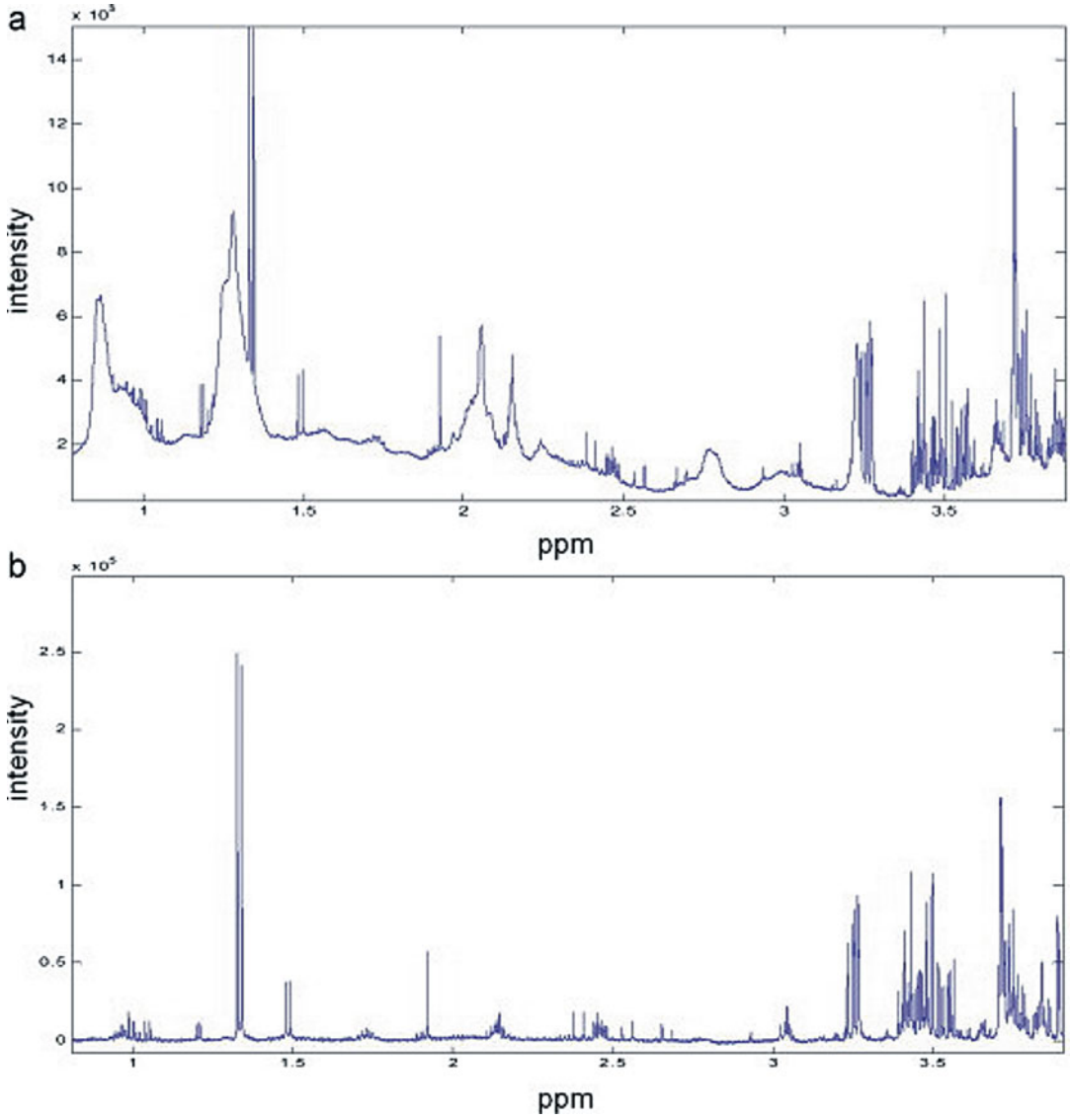


Fig. 3 The 500 MHz ^{1}H NMR spectrum of a blood plasma sample: (**a**) before and (**b**) after protein removal using 10 kDa centrifugal filters (*Reprinted from* Anal Chim Acta [51] with *permission from Elsevier*)

implemented. The most popular method applied for metabolic data analysis is multivariate data analysis, which includes principal component analysis (PCA) and partial least squares (PLS) analyses [33]. The PCA and PLS methods aim to distinguish between separate cohorts in highly complex datasets, despite a large background and the "within-class" variability. Unsupervised PCA is used to reveal initial groupings of samples in the dataset and to identify possible outliers, while the supervised PLS and orthogonal PLS (OPLS) analyses aim to determine data discrimination and classification based on the class identifiers (e.g., benign versus

malignant). It should be noted that, often, before the statistical analysis is carried out, the data needs to be normalized (e.g., total sum normalization, median fold change) and/or log transformed, scaled (e.g., unit variance scaling, Pareto scaling), and centered (e.g., mean centering) to ensure normal distribution of the data and to account for differences in sample dilution [34, 35].

While NMR has its known limitations with respect to sensitivity and initial investment compared to other analytical platforms, it does have inherent strengths that allow it to be a useful tool for assessing the cancer metabolome. Processing of fluids is simple and allows for rapid and reproducible initial setup. This straightforward preparation allows for the retrospective analysis of diverse biobanks and biofluids relatively easily. Recent advances in NMR technology allow for even more improvements in the metabolomics field [36]. High-throughput epidemiological ^{1}H NMR-based metabolomics studies have already been carried out which consisted of more than 5000 samples [37, 38]. Currently there are several large-scale clinical research projects ongoing in our laboratory, which are concentrating on bladder, prostate, and breast cancer [39, 40]. These projects involve larger number of samples, with on the order of 300–1400 samples per study. NMR-based metabolomics also has the ability to identify unknown metabolites that can be further characterized, thereby making it a discovery platform as well as an analytical tool. Direct measurements of in vivo metabolites using NMR can also be performed with magnetic resonance imaging (MRI)-based spectroscopy methods, and development of this technology may allow for noninvasive tissue metabolite analysis without sample preparation [41]. Be that as it may, most NMR-based metabolomics studies of clinical cancer samples, reported to date, have focused on metabolic disease biomarker discovery in various biofluid samples. Since collection of blood and urine samples is well established, economical, and also well accepted by patients, biofluid analysis will remain a major method for clinical assessment in the future.

2 Materials

2.1 Reagents

Prepare all solutions using ultrapure water (purified deionized water, resistivity of $\geq$18.2 MΩ·cm at 25 °C), and use only analytical grade reagents. Prepare and store all reagents at room temperature (unless otherwise indicated). Follow all safety procedures and waste disposal regulations when working in the laboratory and disposing waste materials.

1. Sample buffer: 6.8995 g monosodium phosphate monohydrate ($NaH_2PO_4 \cdot H_2O$), 0.0546 g 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (NaDSS), and 100 mL D_2O (*see*

Note 1). Final concentration of 0.5 M NaH_2PO_4 and 2.5 mM DSS (*see* **Note 2**).

2. D_2O (≥99.8%).
3. 1 M NaN_3.

2.2 Equipment

1. Sterile silicone-coated vacutainers (serum tubes or serum-separating tubes, no additive).
2. Sterile sodium heparin- or lithium heparin-coated vacutainers.
3. Sterile urine containers.
4. Centrifuge with temperature control.
5. Container with ice.
6. Cryovials/2D barcode sample storage tubes.
7. Bucket with ice.
8. 10 kDa centrifugal filters (filter and microtubes) (*see* **Note 3**).
9. Centrifuge with temperature control.
10. Vortex.
11. 1 mL serological pipettes.
12. Glass Pasteur pipettes.
13. 5 mm NMR tubes which do not require spinners (length, 103.5 mm; outside diameter, 5 mm; wall thickness, 0.43 mm) (Fig. 4) (*see* **Note 4**).

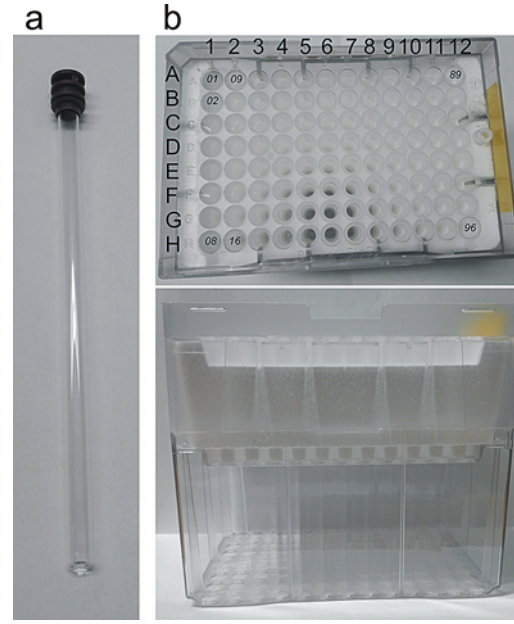


Fig. 4 NMR equipment. (**a**) 5 mm NMR tube which does not require spinners (length, 103.5 mm; outside diameter, 5 mm). (**b**) 96-well-plate-sized NMR tube rack (view from the top, top figure; side view, figure below). The tube positions are highlighted in circles and marked with numbers

14. 600 MHz Bruker Ultrashield Plus NMR spectrometer (Bruker BioSpin Ltd., Canada) equipped with the Bruker® SampleJet and a 5 mm TXI probe at 298 K (*see* **Note 5**).
15. 96-well-plate-sized NMR tube rack (Fig. 4).

2.3 Software (See Note 6)

1. TopSpin (Bruker, Canada).
2. IconNMR (Bruker, Canada).
3. Chenomx NMR Suite (Chenomx Inc., Canada).
4. SIMCA (Umetrics, Sweden).

3 Methods

3.1 Biofluid Sample Collection

For optimal clinical metabolomics-based studies, it is crucial to ensure that information about the medication and dietary intake of the patients is recorded for up to 24 h prior to biofluid sample collection.

3.1.1 Blood

1. Collect blood from an existing arterial or central venous catheter, or with intravenous insertion, or blood culture draw. If possible more than one sample from each individual should be collected.
2. Serum is obtained by collecting the blood into sterile silicone-coated vacutainers (serum tubes, red top; serum-separating tubes with no additive, gold top). The vacutainer type used in a study should be consistent throughout the whole process of sample collection. Immediately after sample collection, the blood is allowed to clot for 45 min at room temperature, and serum is subsequently isolated via centrifugation (1200 × *g* for 15 min at 4 °C) (*see* **Note 7**).
3. Plasma is obtained by collecting the blood into sterile sodium heparin (green top) or lithium heparin (light green top) containing vacutainers (*see* **Note 8**). The vacutainer type used in a particular study should be consistent throughout the whole process of sample collection. Immediately after collection, gently invert blood samples several times, and next centrifuge at 1200 × *g* for 15 min at 4 °C.
4. After centrifugation collect the supernatant (serum or plasma), transfer to cryovials or aliquot if necessary into 2D barcode sample storage tubes, and immediately freeze at −80 °C.

3.1.2 Urine

1. Urine: the samples can be obtained by Foley catheter or collected directly as midstream urine into sterile urine containers. The samples should be processed within 2 h after collection (storage at 4 °C between collection and processing) [42].

2. The samples are immediately treated with NaN_3 (final concentration of NaN_3: min 0.05% weight/volume) over ice [17] and centrifuged at 1600 × *g* for 5 min at 4 °C to remove any particulate matter [42].
3. After centrifugation the supernatant is pipetted into cryovials or aliquoted if necessary into 2D barcode sample storage tubes and stored at −80 °C.

3.2 Sample Preparation

All samples should be prepared fresh and in a random order to avoid progressive bias. Samples should be kept on ice all the time to minimize potential metabolic alterations.

1. Allow the serum/plasma/urine samples to thaw on ice.
2. Clean the 10 kDa centrifugal filters in order to remove glycerol contamination from the filter: place filters in the microtubes and fill up each filter with 600 μL ultrapure water (purified deionized water, resistivity of ≥18.2 MΩ·cm at 25 °C). Next, centrifuge the tubes at 20,000 × *g* for 8 min at 4 °C (*see* **Note 9**).
3. Dump the filtrate in the sink and repeat the filters' washing step four more times.
4. Rinse the microtubes of the 10 kDa filters and the filters with more ultrapure water to remove any remaining trace of filtrate. Insert the filters into the microtubes.
5. Properly label the washed microtubes.
6. Transfer the thawed biofluid samples (250 μL) into the 10 kDa centrifugal filters. For blank samples, transfer 250 μL of D_2O into the 10 kDa centrifugal filters. Centrifuge the microtubes at 20,000 × *g* for 60 min at 4 °C (*see* **Notes 10** and **11**).
7. After 60 min of centrifugation, add 100 μL of D_2O to the filtered protein and blank samples, and continue centrifuging at 20,000 × *g* for 45 min at 4 °C (*see* **Note 12**).
8. After 45 min of centrifugation, remove the 10 kDa centrifugal filters from the microtubes. Dispose the centrifugal filters according to the local waste disposal regulations (*see* **Note 13**).
9. To every sample (filtrate in the microtube), add 112 μL of sample buffer and 10 μL of 1 M NaN_3 aqueous solution (*see* **Note 14**).
10. Close the sample microtubes and vortex each microtube for 15 s.
11. Using a 1 mL serological pipette, measure the volume of the sample in the microtube.
12. Bring the sample volume up to 560 μL with D_2O (*see* **Note 15**). The final concentration of DSS in the sample is 0.5 mM.

13. Close the sample microtubes and vortex each microtube for 10 s.
14. Using glass Pasteur pipettes, transfer the samples into the 5 mm NMR tubes.
15. Place the sample tubes in a 96-well-plate-sized NMR tube rack (rack for tubes which do not require spinners).

3.3 Spectral Acquisition

If using the Bruker® SampleJet system and 5 mm Bruker® SampleJet NMR tubes, make sure that the 5 mm shuttle operating mode is activated in the sample changer.

1. Load a 96-well-plate-sized NMR tube rack with the sample tubes in the sample changer.
2. Open the TopSpin software and check that the temperature is set up to 298 K and the air flow is set to 400 L/min (command: edte).
3. Insert the first sample tube (representative sample) into the magnet (command: sx *tube's position*, e.g., sx 401, where 4 indicates a rack number and 01 indicates a tube's position in the rack) (Fig. 4), and wait till the magnet equilibrates.
4. Based on this representative sample, (a) "Lock" the system (command, lock; select a proper solvent from the list); (b) tune the probe (for automatic tuning use the command atma; for manual tuning, atmm); (c) shim the sample (for automatic gradient shimming, use the command gradshim).
5. Next, reduce the lock signal to ~67%, and run an example NMR spectrum (number of scans (NS) = 1; command: zg).
6. Inspect if the example NMR spectrum is correct, e.g., phase the spectrum and confirm that the upfield DSS peak is at 0.0 ppm and that it has a regular shape with a width at half-height of not more than 1.5 Hz (command, nl; peak's width at half-height range, 0.8–1.5 Hz). If the DSS peak doesn't have a good shape and/or is too wide, it can be corrected by adjusting manually the first-, second-, and third-order shims (*see* **Note 16**).
7. After confirming that the example NMR spectrum is correct, save the shims (command: wsh).
8. Next, determine the 90° pulse length (P1) and save the correct value.
9. Open the IconNMR software, and in the list, select (double click) a sampler holder number that corresponds to the sample (tube's position) which was inserted into the magnet (Fig. 4).
10. From the pulldown menu, select or type the following information: (a) "Name," name of the folder where the NMR data will be saved; (b) "No," number of the NMR spectrum; (c) "Solvent," type of sample's solvent; (d) "Experiment,"

name of the experiment (pulse program); and (e) "Title," title of the NMR dataset.

11. Next, select from the top menu bar "Parameters → Edit all acquisition parameters."
12. The TopSpin window appears. In order to set up required parameters, write a command: ased. Select a proper pulse program (PULPROG: noesypr1d), acquisition time (AQ), delay between pulses (D1), number of scans (NS), number of dummy scans (DS), and the 90° pulse length (P1, the value determined in **step 8**) (*see* **Note 17**).
13. Return to IconNMR, and copy the row with the established parameters to the remaining sampler holder numbers (each number corresponds to the sample tube in a 96-well-plate-sized NMR tube rack). Remember to update the title for each sample. "No" (number of the NMR spectrum) will be updated automatically by the software.
14. Highlight the rows and select "Submit." Next, click on the "Set Automation mode and start the run" button in the top menu bar. Mark the option "First sample in the magnet (locked and shimmed)," and then click "Start" to start the measurements.
15. When the spectral acquisition is finished, open the 1D NMR spectrum in TopSpin (command re *Number of the NMR spectrum*), and briefly check the quality of the spectrum: (a) Fourier transform the spectrum (command: efp); (b) adjust phasing (zero order and first order); (c) perform baseline correction (command: abs); (d) set DSS peak at 0.0 ppm; and (e) check the peak width at half-height (command, nl; the width should not be more than 1.5 Hz). If the spectrum has defects (e.g., uneven baseline, wide DSS peak, irregular shape of DSS peak), the 1D NMR spectrum needs to be rerun.
16. After initial inspection of the NMR spectra and confirming spectral quality, the spectra can be transferred to the Chenomx NMR Suite (Chenomx Inc.) software as Bruker 1r or fid files for further processing and metabolite profiling.

3.4 Spectral Processing

NMR spectra are normally processed using the processor module in the Chenomx NMR Suite (Chenomx Inc.) software. The NMR data can be imported as Bruker 1r or fid files to the processor module. While processing the spectra, the user can select automatic or manual spectral processing options. The protocol below describes all the steps for manual spectral processing.

1. Line broadening: in the top menu bar, select Processing and next Line Broadening (or click on the Line Broadening button

under the spectrum window). Set the Line Broadening value to 0.2 Hz. Accept the change.

2. Phasing: from the Processing menu tab, select Phase Correction (or click on the Phase Correction button under the spectrum window). Adjust the zero-order and the first-order phases by moving the pivots in a proper direction and using the Normal, Fine, and Very Fine modes (scale of changes made with changing the phases). Click the Accept button to continue. The NMR spectrum is correctly phased when its baseline looks like a smooth line (curve) along the entire spectrum (*see* **Note 18**).
3. Water deletion: from the Processing menu tab, choose Region Deletion. Before deleting the water peak, make sure that there are no peaks of interest near to the water peak (zoom in the region). To delete highlighted water peak region, click Accept. You can choose a narrower or broader region to be deleted by changing the range of ppm.
4. Baseline correction: from the Processing menu tab, select Baseline Correction (or click on the Baseline Correction button under the spectrum window). Choose Auto Spline option. You can smooth the line by manually adding and/or moving the breakpoints. A flat baseline allows for a better peak integration and subsequently for better quantification results. When finished click Accept.
5. Shim correction: from the Processing menu tab, select Shim Correction (or click on the Shim Correction button under the spectrum window). This function allows to smooth and correct the peak shapes. Zoom in the highlighted DSS peak and overlap the blue contour with the peak. You may also slightly narrow or broaden the DSS peak by moving the side pivots or changing the value of Target Linewidth (Hz). The peak width should not be more than 1.5 Hz. Accept the changes (*see* **Note 19**).
6. Calibrate CSI: click on the Calibrate CSI button under the spectrum window. This step allows one to set the upfield DSS peak to 0.0 ppm and to establish the concentration of the internal standard (0.5 mM). Zoom in the highlighted DSS peak and overlap the red contour with the peak by moving the lower pivot. Use the upper pivot to adjust the peak height and side pivots to slightly narrow or broaden the DSS peak. When finished accept the changes.
7. Save the processed spectrum as a cnx file. Such a processed spectrum can be further profiled in the Profiler module in Chenomx NMR Suite (Chenomx Inc.) software.

3.5 Metabolite Profiling

Start metabolite profiling by first analyzing the NMR signals of DSS. Next, move to other regions of the NMR spectrum to profile more compounds or search for specific compounds (that you have identified in prior or preliminary studies) using the Quick Search option.

1. Click on the Select Region button in the top menu bar. Next, highlight the area around the upfield DSS peak (0.0 ppm).
2. Right click inside of the highlighted area, and select the Search for Compounds in this Region option. The suggested compounds appear in the Compound Table under the spectrum window. The best match is marked in bold (DSS Chemical Shape Indicator) (*see* **Note 20**).
3. Inspect all signals associated with DSS by clicking on each ppm value in the left top corner of the spectrum window. The ^{1}H NMR spectrum of DSS exhibits four signals: singlet at 0.0 ppm and three multiplets at 0.6 ppm, 1.75 ppm, and 2.9 ppm (*see* **Note 21**).
4. A green ppm value indicates a good fit of the suggested compound with the actual signal. You can adjust the fitting of the suggested signal into the spectrum by moving the lower and upper pivots.
5. After profiling (fitting) all the signals of DSS, move to other regions of the NMR spectrum and search for other compounds. Repeat all the steps of the procedure described above for other peaks in the spectrum (*see* **Note 22**). Ideally, all signals observed in the NMR spectrum will be profiled and matched with specific compounds (*see* **Note 23**). To assure correct selection of the compounds, you can verify the compounds with 2D NMR spectra and available metabolite databases (e.g., Human Metabolome Database [43]).
6. When finished save the profiled spectrum as a cnx file.
7. In order to export the metabolite profiling results into a spreadsheet, select Batch Export function from the Tools (top menu bar). Follow the steps: (a) choose the type of data you wish to export (concentrations in different units, matched clusters, etc.), and select the cnx files you want to export, and provide the name of a target file for the exported data, and click next; (b) select what compounds should be exported (choose Export Data for All Profiled Compounds), and click next; (c) select (or not) additional information you would like to be included in the exported data (e.g., KEGG Compound ID, HMDB Accession Number, Formula), and click next; (d) choose (or not) additional spectrum information to export (e.g., Magnet Frequency (MHz), Pulse sequence), and click next; and (e) select data format (Default, cnx file names in rows,

compounds in columns; Transposed, compounds in rows, cnx file names in columns), and click finish.

8. Open the exported file (metabolite concentrations in profiled spectra) in Excel (Microsoft Corporation, USA), in order to perform further data processing and statistical analysis as appropriate for your study.

3.6 Statistical Analysis

The statistical analysis procedure implemented in our laboratory consists of ^{1}H NMR data preprocessing (median fold change normalization [44]) in Excel, while further data processing (logarithmic transformation, mean centering, and unit variance scaling) is completed in the commercially available SIMCA software (v. 14.1, Umetrics, Sweden). Next, the SIMCA software is used for all multivariate statistical analysis such as unsupervised PCA and supervised PLS and OPLS methods [45] (*see* **Note 24**).

1. Start the SIMCA program and create a new project from File → New → Regular project.
2. Select type of data (supported files—default setting) and find the dataset (Excel file). Next, click Open. The selected Excel spreadsheet will be imported to the SIMCA software.
3. Specify Primary ID and as many Secondary IDs as desired for both variables and observations. Specify quantitative, qualitative, and date/time variables. Click on the Finish button to finalize importing a currently open spreadsheet. Save the file as a new SIMCA project file (.usp).
4. Depending on the observation/variables ID assignment, an unsupervised PCA or a supervised PLS model will alternatively be created by the software. However, this requires further fitting and analysis (*see* Subheading 3.6, **step 5**).
5. To fit the model, select Home → Autofit. A plot with the summary of the fit of the model will appear. The fit summary plot (bar plot) shows R2X(cum) (variation of the data explained by each component; R2Y(cum) for the supervised analysis) and Q2(cum) (cross-validated R2X(cum) for the PCA model or predictive ability of the supervised model based on the sevenfold cross-validation (default setting) for the supervised analysis). The R2X (R2Y) and Q2 values range from 0.0 to 1.0 (*see* **Note 25**).
6. To see the score scatter plot for the dataset, select Home → Scores → Scatter. By default the plot displays two first components. The ellipse represents the 95% confidence interval of the Hotelling's T-squared distribution (Fig. 5) (*see* **Notes 26** and **27**).
7. To investigate potentially important metabolites, select Home → Loadings → Scatter for unsupervised analysis or

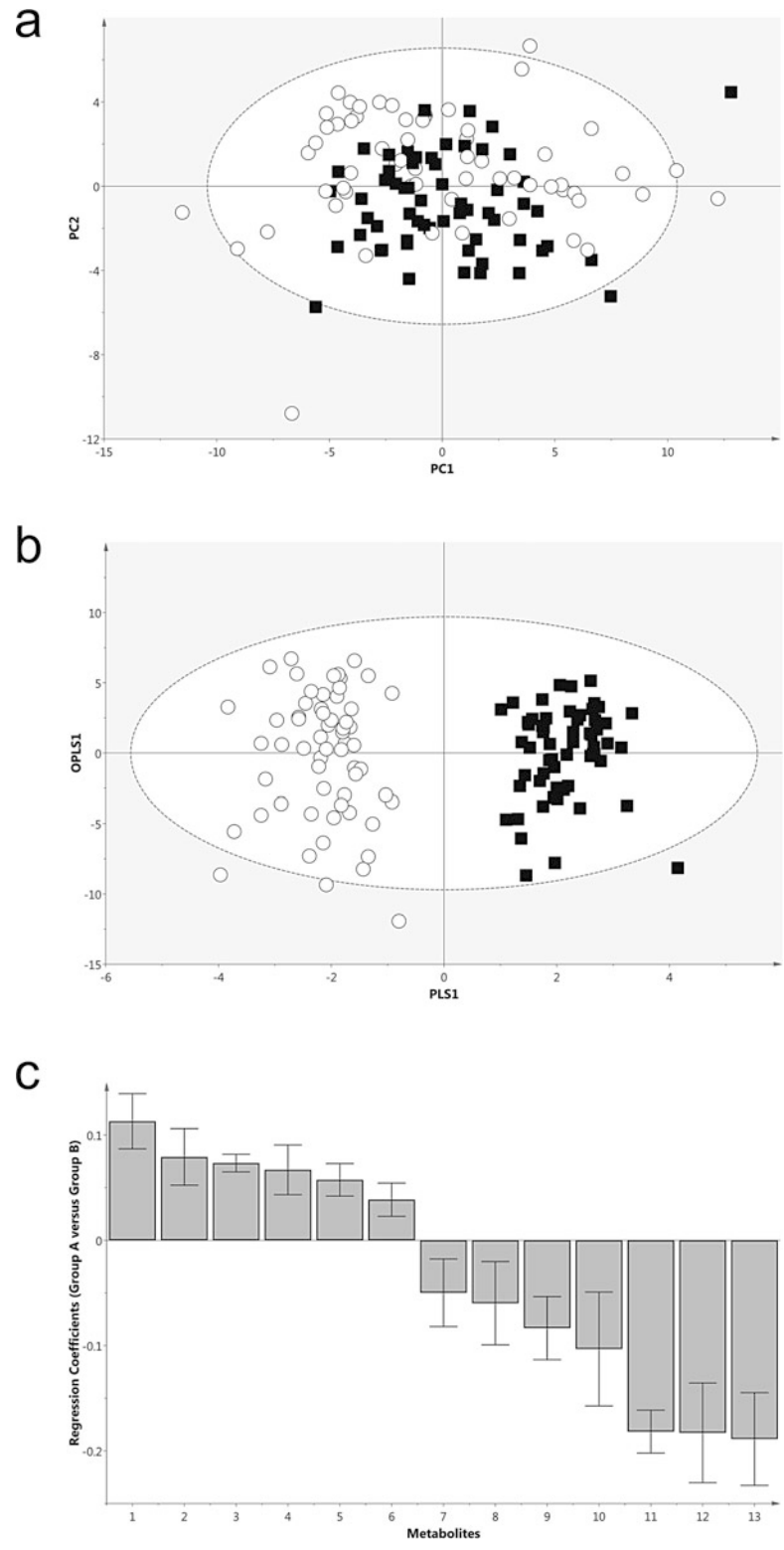


Fig. 5 Examples of (**a**) PCA score scatter plot, (**b**) OPLS score scatter plot (R2Y = 0.89, Q2 = 0.87), and (**c**) regression coefficient plot for a hypothetical

Home → Coefficients for supervised analysis. These plots demonstrate which variables (metabolites) are responsible for differences between studied patient groups (Fig. 5).

8. To save the plots, first, right click on the plot, and next select "Save as" function.
9. To save the SIMCA project, File → Save (or Save as). When finished click Close.

4 Notes

1. Dissolve 6.8995 g of $NaH_2PO_4 \cdot H_2O$ and 0.0546 g of NaDSS in approximately 90 mL of D_2O. Dissolving $NaH_2PO_4 \cdot H_2O$ and NaDSS in D_2O may require some sonication. Transfer the solution quantitatively to a 100 mL volumetric flask. Bring the solution to pH 7 using small aliquots from aqueous solutions of 1 M HCl and/or 1 M NaOH. Bring the buffer solution to a final volume (100 mL) using D_2O, and mix it thoroughly.
2. For plasma, serum, and other biofluids that contain a significant protein content, it is advisable not to use the sodium salt of 3-trimethylsilylpropionic acid (TSP) as an internal reference standard. TSP can bind to proteins, and therefore, its signal is much reduced with a very broad line width [17].
3. In our laboratory, we have also used 3 kDa centrifugal filters. Although the results that we obtained using 3 kDa and 10 kDa centrifugal filters were very similar, the 3 kDa centrifugal filters can get clogged easier via accumulated proteins at the bottom of the filter than the 10 kDa centrifugal filters.
4. In our laboratory, we typically use 5 mm Bruker® SampleJet NMR tubes without requiring spinners for 600 MHz NMR spectrometer (tube length, 103.5 mm; outside diameter, 5 mm; wall thickness, 0.43 mm). Depending on the settings of the NMR spectrometer, other NMR tubes can be used such as 3 mm Bruker® SampleJet NMR tubes (tube length, 103.5 mm; outside diameter, 3 mm; wall thickness, 0.43 mm), 1.7 mm Bruker® MicroCryoProbe NMR tubes

Fig. 5 (continued) metabolomics study: Group A (circles, $n = 61$) versus Group B (black squares, $n = 60$). The ellipse in panels (**a**) and (**b**) represents the 95% confidence interval of the Hotelling's T-squared distribution. In panel (**c**) only significant metabolites ($p < 0.05$) are shown ($n = 13$). Positive values of coefficients (the upper part of the diagrams) indicate increased metabolite concentrations in Group A (fold change >1), while negative values (the lower part of diagrams) present a decrease in metabolite concentrations, as compared to Group B (fold change <1)

(tube length, 103.5 mm; outside diameter, 1.7 mm; wall thickness, 0.4 mm), standard 5 mm medium wall tubes (tube length, 178 mm; wall thickness, 0.8 mm), or standard 5 mm thin wall tubes (tube length, 178 mm; wall thickness, 0.38 mm). The 5 mm and 3 mm NMR tubes can be easily filled and shimmed, while 1.7 mm NMR tubes are difficult to fill and are not recommended for routine metabolomics analysis.

5. In our laboratory, we use a 600 MHz Bruker Ultrashield Plus NMR spectrometer for metabolomics-based studies. Other NMR spectrometers and other frequencies can be used, too. However, the frequency of 600 MHz is a very good compromise between the signal intensity of metabolites and the cost and time of NMR experiment. Somewhat better spectral resolution is obtained at higher proton (^{1}H) frequencies, which facilitates the spectral fitting process. It should also be noted that better sensitivity of detection (and hence shorter acquisition times) can be obtained by using cryoprobes. However, in our experience the tuning and matching of cryoprobes are more sensitive to changes in salt concentrations in the samples than in regular probes.
6. TopSpin (Bruker, Canada) and IconNMR (Bruker, Canada) are integral parts of Bruker NMR spectrometer's operating system. The Chenomx (Chenomx Inc., Canada) and SIMCA (Umetrics, Sweden) software are commonly used in our laboratory for metabolite profiling and multivariate statistical data analysis. However, the protocol described here can be adapted to other available software.
7. It has been reported that to further minimize any bias associated with serum sample collection, it is recommended to allow all the blood samples to clot on ice for a defined time period between 20 and 35 min [46].
8. Blood is sometimes also collected into vacutainers with additives such as citrate or ethylenediaminetetraacetate (EDTA). However, it is important to remember that when citrate or EDTA is used as an anticoagulant, additional strong NMR signals with high intensity are observed in the NMR spectrum [17]. These interfere with the spectral analysis and reduce the number of peaks that can be reliably assigned.
9. Always clean more centrifugal filters than the number of analyzed samples (e.g., $\geq n + 2$, n: number of analyzed samples). It can happen that a few centrifugal filters may have manufacturing defects, and therefore, it is recommended to always have some extra clean and washed centrifugal filters ready. Additionally, wash and clean a few more centrifugal filters which will be used for preparation of blank samples

(at least one blank sample for every ten samples). The blank samples are used to detect potential sources of contamination during sample preparation.

10. A biofluid sample volume of 250 μL is the most optimal for clinical NMR-based metabolomics studies. However, in our laboratory we have also successfully performed metabolomics NMR measurement using smaller sample volumes such as a minimum volume of 150 μL. If there is a sufficient sample volume available, one can also prepare additional pooled QC samples.
11. The volume of D_2O used for the blank samples should be the same as the sample volume.
12. The filtrate should be clear. If it is not, immediately transfer the filtrate to a new clean and washed 10 kDa centrifugal filter and repeat.
13. The filters contain proteins which can be further analyzed for proteomics analysis, if so desired. In this case remove the filters from the microtubes, and place them in the Eppendorf tubes. Label the Eppendorf tubes accordingly. Store the Eppendorf tubes with the filter at −80 °C until subsequent analysis.
14. The volume of sample buffer varies depending on the type of the NMR tube. In our protocol we normally use 5 mm Bruker® SampleJet NMR tubes for 600 MHz NMR spectrometer (tube length, 103.5 mm; outside diameter, 5 mm). The volume of sample buffer should be 80 μL if the medium wall tubes (tube length:,178 mm; wall thickness, 0.8 mm) are used or 130 μL for the standard (thin wall) tubes (tube length, 178 mm; wall thickness, 0.38 mm).
15. The final sample volume varies depending on the type of the NMR tube. In our protocol we routinely use 5 mm Bruker® SampleJet NMR tubes for 600 MHz NMR spectrometer (tube length, 103.5 mm; outside diameter, 5 mm). The final sample volume should be 400 μL if the medium wall tubes (tube length, 178 mm; wall thickness, 0.8 mm) are used or 650 μL for the standard (thin wall) tubes (tube length, 178 mm; wall thickness, 0.38 mm).
16. In order to manually adjust the shims and correct the shape of DSS peak, press the "onaxis" button on the console and start changing the shim values. If the peak is broad and/or split on the top, the first-order shims need to be adjusted ($Z,^{1}$ X, Y). The second-order shims ($Z,^{2}$ XY, YZ, XZ, X^2-Y^2) need to be changed if the peak is skewed (asymmetric shape of the peak from its left or right side). When the peak has a side shoulder, both, the first- and second-order shims, need to be adjusted. Symmetric broad tails are often associated with incorrect third-order shims ($Z,^{3}$ $X,^{3}$ Y^3, $XZ,^{2}$ $YZ,^{2}$ ZXY, $Z(X^2$-$Y^2)$). For more

information about shim adjustment, *see* Chmurny et al. [47]. After changing each shim, run an example NMR spectrum (number of scans (NS) = 1; command, zg; when finish command, efp) to assure that the shape of DSS peak has improved. When making changes to the shims, always observe the lock intensity. Generally, continue making changes until a maximum of lock intensity is achieved (it cannot be improved anymore).

17. The noesypr1d pulse program provides an optimal water suppression pulse sequence which utilizes the first increment of a NOESY (Nuclear Overhauser Effect Spectroscopy) pulse sequence with water irradiation during the relaxation delay and the mixing time. In our laboratory for all the pilot/discovery studies, we use the noesypr1d pulse program with a mixing time of 100 ms and the following acquisition parameters: size of fid (number of time domain points) TD = 57,690, acquisition time AQ = 4 s, delay between pulses D1 = 1 s, number of scans NS = 1024, and number of dummy scans DS = 4. The 90° pulse length (P1) is determined based on a representative NMR sample. The total acquisition time is 1 h and 28 min for each NMR spectrum. Depending on the purpose of the study, the NS and AQ values may be reduced which will result in shortening the total acquisition time to even few minutes. For other pulse programs that can be applied in metabolomics studies, *see* [17].
18. Very often, the water peak has a quite distorted shape when compared to the rest of the NMR spectrum. Therefore, the water peak should not be used to adjust phasing. However, a flat baseline should be obtained on both sides of the water peak along the entire spectrum.
19. It is important to note that when artificially changing the DSS peak width in the Chenomx software, some potential spectral artifacts might appear, e.g., narrowing of the peak will reduce the signal-to-noise ratio by increasing the noise, while artificial broadening of the peak will result in spectral line broadening (effects seen on spectral baseline). Therefore, it is recommended to establish a proper DSS peak width at half-height during the spectral acquisition (*see* Subheading 3.3) instead of changing it significantly during the spectral processing in the Chenomx software.
20. Although the software gives a recommended best match, it is often advisable to check other not highlighted compounds that are also listed in the Compound Table. One should make sure that all signals of a specific compound match with the peaks in the spectrum.

21. If the DSS signals cause a problem (e.g., overlap with the signals of the metabolite of interest), a deuterated version of DSS can be used: 2,2-dimethyl-2-silapentane-5-sulfonate-D6 sodium salt (NaDSS-D6). DSS-D6 gives only one signal in ^{1}H NMR spectrum: the singlet peak at 0.0 ppm. However, the deuterated version of DSS is available at a much higher cost.
22. If one performs a targeted analysis and is only interested in certain metabolites, the metabolite profiling can solely be done on selected compounds instead of assigning all the signals detected in the ^{1}H NMR spectrum. It should also be noted that complete manual fitting of the spectra using the Chenomx software can be time-consuming. Computer programs that allow for automated peak assignments and concentration determination are becoming available [48] and may eventually replace this step. Some automation is also possible in the current version of the Chenomx software.
23. Some compounds might have signals for which the shape/position will not directly match with the Chenomx compound library, e.g., citrate and glucose. One peak from each of citrate's doublets is often shifted downfield or upfield. The glucose's signal at 3.8 ppm often has an asymmetric shape. The Chenomx software can make small adjustment for many compounds depending on pH, salt concentrations, etc. (for discussion, *see* also [48]).
24. Multivariate statistical analysis (rather than univariate statistics) is most often used to analyze the large and complex metabolomics datasets. While PCA and PLS/OPLS analysis are among the most frequently used approaches, other machine learning methods such as support vector machines, fuzzy logic, random forest, or Bayesian network are also starting to be used as alternative approaches [49, 50].
25. For the supervised analysis (PLS, OPLS) of clinical biofluid samples, the difference between R2Y and Q2 values should not be more than 0.3 units. If the differences between R2Y and Q2 are $\geq$0.3 units, it indicates that the supervised model is overfitted, and hence the model demonstrates poor predictive ability. Also, the supervised models are considered reliable if the R2Y $\geq$0.5 and Q2 $\geq$ 0.4 [33].
26. In PCA score scatter plots, sometimes differences are observed between the sexes and those patients consuming a specific drug or being exposed to drug treatment. In such cases it may be advantageous to consider these groups separately, which implies that the study design gets much bigger, as one should ideally still try to maintain $\geq$50 participants in each group. Likewise diabetes can sometimes be a cofounder that influences the outcome of a metabolomics study [12], and ideally

this should be considered as well if at all possible. Other cofounders are possible, and therefore an appropriate study design with a sufficiently powered dataset should be considered, in order to observe meaningful differences among the groups of interest.

27. In a well-designed study, it is not uncommon to see no separation in the PCA score scatter plots. However, subsequent PLS or OPLS analysis can still give good separation.

Acknowledgments

This work was funded by a CRIO-Cancer Grant from Alberta Innovates Health Solutions. We are indebted to all our colleagues that provided clinical samples, and especially to Drs. Aalim Weljie, Farshad Farshidfar, and Karen Kopciuk for discussions about methodology.

References

1. Warburg O (1956) On the origin of cancer cells. Science 123(3191):309–314
2. Otto AM (2016) Warburg effect(s)-a biographical sketch of Otto Warburg and his impacts on tumor metabolism. Cancer Metab 4:5
3. Wolpaw AJ, Dang CV (2018) MYC-induced metabolic stress and tumorigenesis. Biochim Biophys Acta Rev Cancer 1870(1):43–50
4. Dowell AC, Cobby E, Wen K, Devall AJ, During V, Anderson J, James ND, Cheng KK, Zeegers MP, Bryan RT, Taylor GS (2017) Interleukin-17-positive mast cells influence outcomes from BCG for patients with CIS: data from a comprehensive characterisation of the immune microenvironment of urothelial bladder cancer. PLoS One 12(9):e0184841
5. Srinivasan R, Ricketts CJ, Sourbier C, Linehan WM (2015) New strategies in renal cell carcinoma: targeting the genetic and metabolic basis of disease. Clin Cancer Res 21(1):10–17
6. Yang G, Shen W, Zhang Y, Liu M, Zhang L, Liu Q, Lu HH, Bo J (2017) Accumulation of myeloid-derived suppressor cells (MDSCs) induced by low levels of IL-6 correlates with poor prognosis in bladder cancer. Oncotarget 8 (24):38378–38388
7. Townsend MK, Bao Y, Poole EM, Bertrand KA, Kraft P, Wolpin BM, Clish CB, Tworoger SS (2016) Impact of pre-analytic blood sample collection factors on metabolomics. Cancer Epidemiol Biomarkers Prev 25(5):823–829
8. Xiao Q, Moore SC, Boca SM, Matthews CE, Rothman N, Stolzenberg-Solomon RZ, Sinha R, Cross AJ, Sampson JN (2014) Sources of variability in metabolite measurements from urinary samples. PLoS One 9(5):e95749
9. Sampson JN, Boca SM, Shu XO, Stolzenberg-Solomon RZ, Matthews CE, Hsing AW, Tan YT, Ji BT, Chow WH, Cai Q, Liu DK, Yang G, Xiang YB, Zheng W, Sinha R, Cross AJ, Moore SC (2013) Metabolomics in epidemiology: sources of variability in metabolite measurements and implications. Cancer Epidemiol Biomarkers Prev 22(4):631–640
10. The 5 Core Elements of a Successful Metabolomics Study. Metabolon Inc. http://metabolomics.metabolon.com/acton/media/17033/metabolomics-study-success. Accessed 30 May 2018
11. Dunn WB, Broadhurst DI, Atherton HJ, Goodacre R, Griffin JL (2011) Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. Chem Soc Rev 40 (1):387–426
12. Bathe OF, Shaykhutdinov R, Kopciuk K, Weljie AM, McKay A, Sutherland FR, Dixon E, Dunse N, Sotiropoulos D, Vogel HJ (2011) Feasibility of identifying pancreatic cancer based on serum metabolomics. Cancer Epidemiol Biomarkers Prev 20(1):140–147
13. Falegan OS, Ball MW, Shaykhutdinov RA, Pieroraio PM, Farshidfar F, Vogel HJ, Allaf ME, Hyndman ME (2017) Urine and serum

metabolomics analyses may distinguish between stages of renal cell carcinoma. Metabolites 7(1):6

14. Farshidfar F, Weljie AM, Kopciuk K, Buie WD, Maclean A, Dixon E, Sutherland FR, Molckovsky A, Vogel HJ, Bathe OF (2012) Serum metabolomic profile as a means to distinguish stage of colorectal cancer. Genome Med 4(5):42
15. McConnell YJ, Farshidfar F, Weljie AM, Kopciuk KA, Dixon E, Ball CG, Sutherland FR, Vogel HJ, Bathe OF (2017) Distinguishing benign from malignant pancreatic and periampullary lesions using combined use of (1)H-NMR spectroscopy and gas chromatography-mass spectrometry. Metabolites 7(1):3
16. Lindon JC, Nicholson JK, Holmes E, Everett JR (2000) Metabonomics: metabolic processes studied by NMR spectroscopy of biofluids. Concepts Magn Reson 12(5):289–320
17. Beckonert O, Keun HC, Ebbels TM, Bundy J, Holmes E, Lindon JC, Nicholson JK (2007) Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat Protoc 2(11):2692–2703
18. An YJ, Cho HR, Kim TM, Keam B, Kim JW, Wen H, Park CK, Lee SH, Im SA, Kim JE, Choi SH, Park S (2015) An NMR metabolomics approach for the diagnosis of leptomeningeal carcinomatosis in lung adenocarcinoma cancer patients. Int J Cancer 136(1):162–171
19. Roberts MJ, Richards RS, Chow CWK, Buck M, Yaxley J, Lavin MF, Schirra HJ, Gardiner RA (2017) Seminal plasma enables selection and monitoring of active surveillance candidates using nuclear magnetic resonance-based metabolomics: a preliminary investigation. Prostate Int 5(4):149–157
20. Wang J, Ma C, Liao Z, Tian B, Lu JP (2011) Study on chronic pancreatitis and pancreatic cancer using MRS and pancreatic juice samples. World J Gastroenterol 17(16):2126–2130
21. Mickiewicz B, Kelly JJ, Ludwig TE, Weljie AM, Wiley JP, Schmidt TA, Vogel HJ (2015) Metabolic analysis of knee synovial fluid as a potential diagnostic approach for osteoarthritis. J Orthop Res 33(11):1631–1638
22. Weljie AM, Newton J, Mercier P, Carlson E, Slupsky CM (2006) Targeted profiling: quantitative analysis of 1H NMR metabolomics data. Anal Chem 78(13):4430–4442
23. Nicholson JK, Foxall PJ, Spraul M, Farrant RD, Lindon JC (1995) 750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma. Anal Chem 67(5):793–811
24. Van QN, Chmurny GN, Veenstra TD (2003) The depletion of protein signals in metabonomics analysis with the WET-CPMG pulse sequence. Biochem Biophys Res Commun 301(4):952–959
25. Issaq HJ, Van QN, Waybright TJ, Muschik GM, Veenstra TD (2009) Analytical and statistical approaches to metabolomics research. J Sep Sci 32(13):2183–2199
26. Tiziani S, Emwas AH, Lodi A, Ludwig C, Bunce CM, Viant MR, Gunther UL (2008) Optimized metabolite extraction from blood serum for 1H nuclear magnetic resonance spectroscopy. Anal Biochem 377(1):16–23
27. Daykin CA, Foxall PJ, Connor SC, Lindon JC, Nicholson JK (2002) The comparison of plasma deproteinization methods for the detection of low-molecular-weight metabolites by (1)H nuclear magnetic resonance spectroscopy. Anal Biochem 304(2):220–230
28. Zhang B, Xie M, Bruschweiler-Li L, Bruschweiler R (2016) Nanoparticle-assisted removal of protein in human serum for metabolomics studies. Anal Chem 88(1):1003–1007
29. Ludwig C, Viant MR (2010) Two-dimensional J-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox. Phytochem Anal 21(1):22–32
30. Kruk J, Doskocz M, Jodlowska E, Zacharzewska A, Lakomiec J, Czaja K, Kujawski J (2017) NMR techniques in metabolomic studies: a quick overview on examples of utilization. Appl Magn Reson 48(1):1–21
31. Lewis IA, Schommer SC, Hodis B, Robb KA, Tonelli M, Westler WM, Sussman MR, Markley JL (2007) Method for determining molar concentrations of metabolites in complex solutions from two-dimensional 1H-13C NMR spectra. Anal Chem 79(24):9385–9390
32. Fan TW-M (1996) Metabolite profiling by one- and two-dimensional NMR analysis of complex mixtures. Prog Nucl Magn Reson Spectrosc 28(2):161–219
33. Trygg J, Holmes E, Lundstedt T (2007) Chemometrics in metabonomics. J Proteome Res 6 (2):469–479
34. Euceda LR, Giskeodegard GF, Bathen TF (2015) Preprocessing of NMR metabolomics data. Scand J Clin Lab Invest 75(3):193–203
35. van den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, van der Werf MJ (2006) Centering, scaling, and transformations: improving

the biological information content of metabolomics data. BMC Genomics 7:142

36. Nagana Gowda GA, Raftery D (2017) Recent advances in NMR-based metabolomics. Anal Chem 89(1):490–510
37. Wurtz P, Kangas AJ, Soininen P, Lawlor DA, Davey Smith G, Ala-Korpela M (2017) Quantitative serum nuclear magnetic resonance metabolomics in large-scale epidemiology: a primer on -omic technologies. Am J Epidemiol 186(9):1084–1096
38. Fischer K, Kettunen J, Wurtz P, Haller T, Havulinna AS, Kangas AJ, Soininen P, Esko T, Tammesoo ML, Magi R, Smit S, Palotie A, Ripatti S, Salomaa V, Ala-Korpela M, Perola M, Metspalu A (2014) Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons. PLoS Med 11(2):e1001606
39. Brockton NT, Gill SJ, Laborge SL, Paterson AH, Cook LS, Vogel HJ, Shemanko CS, Hanley DA, Magliocco AM, Friedenreich CM (2015) The breast cancer to bone (B2B) metastases research program: a multidisciplinary investigation of bone metastases from breast cancer. BMC Cancer 15:512
40. Mickiewicz B, Arnold Egloff S, Eskaros AH, Clark PE, Zijlstra A, Vogel HJ, Hyndman ME (2018) Metabolomics of bladder cancer: from metabolic data to clinical diagnosis and prognosis
41. Lin G, Keshari KR, Park JM (2017) Cancer metabolism and tumor heterogeneity: imaging perspectives using MR imaging and spectroscopy. Contrast Media Mol Imaging 2017:6053879
42. Bernini P, Bertini I, Luchinat C, Nincheri P, Staderini S, Turano P (2011) Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. J Biomol NMR 49(3-4):231–243
43. Human Metabolome Database (HMDB). www.hmdb.ca. Accessed 20 May 2018
44. Dieterle F, Ross A, Schlotterbeck G, Senn H (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. Anal Chem 78 (13):4281–4290
45. Eriksson L, Johansson E, Kettaneh-Wold N, Trygg J, Wikstrom C, Wold S (2006) Multi- and megavariate data analysis part I: basic principles and applications. Umetrics AB, Umeå
46. Teahan O, Gamble S, Holmes E, Waxman J, Nicholson JK, Bevan C, Keun HC (2006) Impact of analytical bias in metabonomic studies of human blood serum and plasma. Anal Chem 78(13):4307–4318
47. Chmurny GN, Hoult DI (1990) The ancient and honourable art of shimming. Concepts Magn Reson 2(3):131–149
48. Takis PG, Schafer H, Spraul M, Luchinat C (2017) Deconvoluting interrelationships between concentrations and chemical shifts in urine provides a powerful analysis tool. Nat Commun 8(1):1662
49. Gromski PS, Muhamadali H, Ellis DI, Xu Y, Correa E, Turner ML, Goodacre R (2015) A tutorial review: metabolomics and partial least squares-discriminant analysis—a marriage of convenience or a shotgun wedding. Anal Chim Acta 879:10–23
50. Cuperlovic-Culf M (2018) Machine learning methods for analysis of metabolic data and metabolic pathway modeling. Metabolites 8 (1):E4
51. Smolinska A, Blanchet L, Buydens LM, Wijmenga SS (2012) NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review. Anal Chim Acta 750:82–97

Chapter 15

Assessment of Metabolic Signature for Cancer Diagnosis Using Desorption Electrospray Ionization Mass Spectrometric Imaging

Shibdas Banerjee and Soumen Kanti Manna

Abstract

Metabolic reprogramming is a hallmark of tumor development. A technique that can map this complex biochemical shift by taking a snapshot of various metabolites in a tissue specimen (biopsy) is of high utility in the context of cancer diagnosis. Desorption electrospray ionization mass spectrometric imaging (DESI-MSI) is such a powerful and emerging analytical technique to simultaneously visualize the distributions of hundreds of metabolites, lipids, and other small molecules in the biological tissue. In DESI-MSI, a fine spray of high-velocity charged microdroplets rapidly extracts molecular species from the tissue surface and subsequently transfers them to the mass spectrometer, while the sample is continuously moved in two dimensions under the impinging spray of microdroplets. This allows a detailed multiplex molecular mapping of the tissue. DESI-MSI enables simultaneous examination of hundreds of putative metabolic biomarkers, an approach that lends much more predictive power than simply evaluating one or a few candidate biomarkers. The speed, versatility, lack of complicated sample preparation, and operation at ambient conditions make DESI-MSI extremely promising as a rapid diagnostic and prognostic tool.

Key words DESI-MSI, Tissue imaging, Cancer margin, Metabolites and lipids, Histopathology

1 Introduction

Currently, microscopic analysis in histology is the gold standard for diagnosing and grading cancer. Histopathological analysis investigates morphological abnormalities in tissues to identify a neoplastic lesion. However, because of morphological mimics, artifacts, and heterogeneity, pathological decision is often subjective, and the outcome may vary depending on the expertise of the pathologist. Moreover, an unreliable biopsy report often causes unsuccessful cancer surgery by leaving tumor cells at the resected specimen edge, which has been associated in many cancers with an increased local recurrence and decreased overall survival. This is now a global challenge to rapidly and accurately detect cancer margin during surgery (intraoperatively) [1–3] to increase the survival rate and

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_15,

decrease the operation time, chance of infection, and the dose of anesthesia.

Even when a tumor is identified and graded correctly, the response to therapy may vary widely. This is, primarily, due to intra-tumoral heterogeneity that confers cancer cells with different degrees of therapeutic resistance and potential for relapse. Intra-tumoral heterogeneity arises due to innate differences in genetic and epigenetic makeup of individual cancer cells within the same tumor that leads to differences in their biochemical phenotype [4–6]. Thus, a method that can quickly and accurately capture the distribution of tumor cells as well as enable characterization of intra-tumoral heterogeneity based on the biochemical signature is highly warranted.

Global change in metabolism is a hallmark of neoplasia [7]. Such metabolic alteration is the consequence of dysregulation of oncogenes or tumor suppressor genes [8, 9]. Oncogenes are known to regulate key genes involved in several metabolic processes, notably glycolysis (Warburg effect), pentose phosphate pathway, Krebs cycle, and lipid metabolism [10]. These phenomena are attributed to significant changes in production of several biosynthetic intermediates and increased channeling of glucose carbon into nucleic acids, amino acids, and phospholipids. Therefore, metabolic reprograming in tumorigenesis contributes to the different aspects of cellular processes, including energy production, cell growth, proliferation, signaling, differentiation, and motility [11–13]. Assessment of metabolic signature from biopsy specimen can offer a trustworthy method for detection and classification of cancer margin.

Indeed, a major headway has been made in the last decade by monitoring metabolic biomarkers in the tissue to diagnose and prognosticate cancer [14–18]. Mass spectrometry imaging (MSI) is a powerful emerging analytical technique for evaluation of the spatial distribution of tissue metabolome to diagnose different types of cancers [19]. Although a number of ionization techniques, including matrix-assisted laser desorption/ionization (MALDI) and desorption electrospray ionization (DESI), have been used in MSI, DESI has potential advantage over other ionization techniques. This is particularly because DESI requires minimum sample preparation and does not require a vacuum chamber or an enclosure for ionization, rendering this technique very fast into the operational workflow.

Desorption electrospray ionization mass spectrometric imaging (DESI-MSI), developed in the laboratory of Prof. Graham Cooks [20, 21], has been used to rapidly evaluate the tissue metabolome at ambient conditions by simultaneously characterizing hundreds of lipids and metabolites [22–25]. Since DESI-MSI offers open-air sampling, it can be applied to a fresh tissue or even a live organ/skin without any pretreatment. DESI-MSI (Fig. 1) is typically

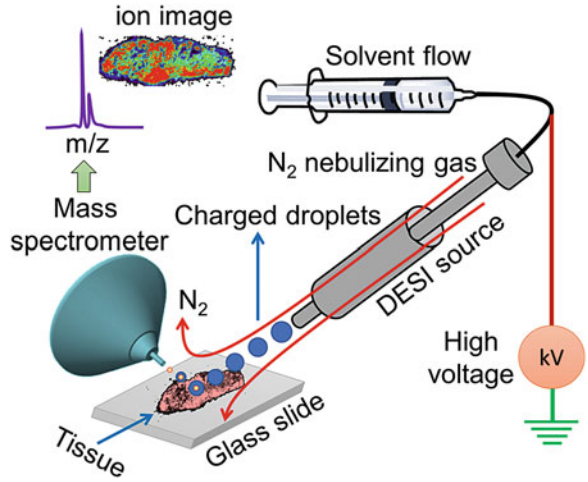


Fig. 1 Schematic diagram of the DESI-MSI setup for intercepting molecular species from the tissue

performed on a tissue section (~15 μm thickness), which is bombarded with a stream of charged microdroplets generated by electrospraying [26, 27] a solvent at high voltage and using nitrogen as nebulizing gas. These droplets dissolve hundreds of lipids and metabolites present in the tissue by wetting the tissue surface with the droplet solvent. Splashing of this liquid film, upon the arrival of subsequent primary droplets, results in the formation of secondary microdroplets containing the analyte ions (extraction of lipids/metabolites), which are afterward converted into gaseous ions for mass spectrometric analysis. The production of gas-phase analyte ions (desolvation process) from secondary microdroplets occurs through repeated solvent evaporation and Coulomb fission, akin to that of normal electrospray ionization mechanism [26, 28]. The imaging experiment is performed by scanning the tissue surface in x and y directions through an impinging spray of charged microdroplets and the corresponding analyte ion signals in pixel-to-pixel mass spectra, which can then be plotted as two-dimensional images, allowing a detailed biochemical (lipids/metabolites) mapping of the sample (tissue). The spatial resolution is approximately 150–200 μm, which compares well with the thickness of a surgical knife. Each pixel, containing the mass spectral fingerprints of hundreds of metabolite, is analyzed using unsupervised or supervised statistical methods such as principal component analysis or lasso multiscale-logistic regression model. The histopathological H&E image of the specimen section is overlaid with the metabolite ion map to teach the DESI-MSI the metabolic signature of normal tissue, indolent cancer, and aggressive cancer.

As mentioned above, DESI-MSI is a label-free and rapid method that allows multiplex analyses of hundreds to thousands

of molecules in a histological section, which can bring a new class of complex molecular information to link the expert discipline of pathology and cell metabolism. In the last decade, some reports from Zare laboratory [23, 24, 29–32] and others [22, 33–43] have demonstrated the usefulness of DESI-MSI in viewing the metabolite/lipid distribution on tissues in the context of cancer and other disease pathologies. Therefore, DESI-MSI has opened a vast possibility to be of utility in the area of cancer diagnosis and biochemistry, considering that the variability in clinical manifestation is largely driven by the underlying biochemical heterogeneity of cancer. There are more than a hundred different known cancers that affect humans. The biochemical profiles of different cancers are expected to be different. Further, unlike genome, cancer metabolome is dynamic and also dependent on diet, lifestyle, medication, race, and heredity. DESI-MSI utility in each cancer model could provide discrete results, inviting an exhaustive exploration of this nascent field.

In view of the aforementioned perspectives, this chapter provides a step-by-step protocol to use DESI-MSI as a diagnostic tool for classifying cancer and normal specimens. Details of this protocol are based on the recent work of Banerjee et al. in prostate cancer diagnosis by DESI-MSI [32]. This protocol, previously used in prostate cancer, is representative of the type of work that can be done in any tumor type.

2 Materials

2.1 Solvents and Chemicals/ Reagents

1. DESI spray solvents: HPLC grade dimethylformamide (DMF; ≥99.9% purity) and acetonitrile (ACN, ≥99.9% purity).
2. Mass spectrometer inlet cleaning solvents: Deionized ultrafiltered water (specific conductance ≤2.0 μmho/cm) and LC-MS grade methanol (≥99.9% purity).
3. H&E staining reagents/solvents: hematoxylin solution (Harris modified), eosin Y solution (intensified), bluing reagent, ammonia, absolute ethanol, methanol, xylene, distilled/Milli-Q water, histological mounting medium.
4. High purity (≥99.995%) compressed nitrogen.
5. Optimal cutting temperature (OCT compound).
6. Metabolite and lipid standards (as required).

2.2 Tissue Sectioning and Scanning

1. Cryostat for high-quality tissue sectioning.
2. Microtome blades.
3. Paint brushes.
4. High-resolution tissue slide scanner.

2.3 Mass Spectrometry

1. LTQ Orbitrap XL mass spectrometer (ThermoFisher Scientific; *see* **Note 1**).
2. Custom-built [44, 45] or commercially available DESI 2D source (Prosolia Inc.; *see* **Note 2**).
3. Extended ion transfer tube-Thermo Exactive 2D (Prosolia Inc.; *see* **Note 2**).
4. Ion transfer tube.
5. 500 μL Hamilton syringe (1750 RN series).

2.4 Other Equipment

1. −80 °C freezer.
2. Ultrasonic baths (5.5 L).
3. 0.22 micron syringe filter.
4. Vacuum desiccator with 1/8–1/6 HP diaphragm pump.
5. Microscope slides (dimension 25 × 75 × 1 mm, suitable for tissue adhesion).
6. Coverslips (dimension 24 × 50 × 0.15 mm).
7. Microscope slide box.
8. Calipers.
9. Scotch magic tape.
10. Magnifying glass or a laboratory microscope.
11. King-size black sharpie.
12. 2 mL microcentrifuge tube.
13. Single-edge lab razor blade.
14. Shandon slide holder.

2.5 Software

1. Xcalibur 2.2 software (Thermo Fisher Scientific, San Jose, CA; *see* **Note 1**).
2. Omni spray 2D software (Prosolia, Indianapolis, IN; *see* **Note 2**).
3. BioMap (freeware, https://ms-imaging.org/wp/biomap/).
4. MSConvert (ProteoWizard freeware, http://proteowizard.sourceforge.net/index.shtml).
5. imzMLConverter (freeware, http://www.cs.bham.ac.uk/~ibs/imzMLConverter/).
6. MSiReader (freeware, http://www4.ncsu.edu/~dcmuddim/msireader.html).
7. MATLAB Complier Runtime R2017b (9.2) (freeware, https://in.mathworks.com/products/compiler/matlab-runtime.html).
8. NDP.view2 software (https://www.hamamatsu.com/jp/en/U12388-01.html).

3 Methods

3.1 Tissue Preparation

1. Embed the frozen prostate tissue (or a biopsy specimen) in minimal amount of OCT compound (few drops) to fix it to the cryostat adapter (*see* **Note 3**).
2. Cut tissue sections of 5–15 μm thickness using the cryomicrotome at −20 °C (*see* **Note 4**).
3. Collect two successive tissue sections of thickness 5 and 15 μm using a paintbrush, and place them on two different microscopic glass slides by thaw mounting. Label these slides properly with the patient code (sample code), section number, and/or section thickness (*see* **Note 5**).
4. Collect the tissue slides immediately in a slide box, and store them at −80 °C to prevent degradation by ice. Keep these samples in the ultra-freezer (−80 °C) until the time of experiment (*see* **Note 6**).

3.2 Histopathology

Hematoxylin and eosin (H&E) staining is one of the most frequently used staining methods in histology and often accepted as gold standard in clinical diagnosis and cancer margin evaluation (*see* **Note 7**). Hematoxylin stains the nucleus in blue, and eosin stains the cytoplasm in pink [46]. Classification of cancer and benign specimens can be done by microscopic investigation of tissue morphology after this staining [46]. A brief method for H&E staining is as follows:

1. Take out the slide containing the tissue of thickness 5 μm from the refrigerator (*see* above), and dry it for 5–10 min in the vacuum desiccator.
2. Carefully remove the OCT film surrounding the tissue using Scotch tape, and clean the Sharpie marks (if any) from the back side of the slide by ethanol wipe.
3. Place the slides (e.g., ten slides from different patients) in the Shandon slide holder.
4. Fix the tissue slides in methanol for 2 min.
5. Rinse the slides in water by ten dips (one dip ~1 s).
6. Fix the slides in hematoxylin solution for 1.5 min.
7. Rinse the slides in water by ten dips.
8. Submerge the slides with one quick dip in the bluing solution containing 0.1% ammonia.
9. Rinse the slides in water by ten dips.
10. Fix the slides in eosin Y solution for 8 s.
11. Rinse and dehydrate the slides in absolute ethanol by ten dips.

12. Rinse the slides again and dehydrate them in absolute ethanol by ten dips.
13. Rinse the slides in xylene by six dips.
14. Rinse the slides again in xylene by six dips.
15. Next, place the slides on a plain white paper, and let them dry at room temperature for 5 h.
16. After drying, deposit a drop of mounting medium onto the stained tissue section followed by quickly covering it with a transparent glass coverslip and let dry for another 5 h.
17. Collect the H&E slides in a slide box, and take them for recording high-resolution optical images, which can be evaluated by a histopathologist. A microscope can also be used alternatively to investigate the tissue morphology for cancer diagnosis.

3.3 Optical Scanning of Tissues

The H&E-stained specimen slides can be converted into digital slides (optical image) by scanning them at high resolution (e.g., ~0.43 μm/pixel) using a digital slide scanner.

1. Scan multiple-stained slides at a specific resolution, scan mode (20× or 40×), and scan speed using the high-resolution tissue scanner (Hamamatsu NanoZoomer 2.0-RS slide scanner; *see* **Note 8**).
2. Save and store these digital slides in the computer as high-definition, high-quality digital images.
3. Use NDP.view2 software to open these image files, which can be zoomed in or out on any portion of the entire tissue image with mouse operation.
4. Demark the images with the help of expert pathologist to show the area of cancer, normal, inflammation, or any other type of pathological features (stroma, BPH, grade of cancer or aggressiveness of cancer, etc.). This pathological examination is done without the knowledge of DESI-MSI evaluation. However, results from this study (gold standard) will be used to guide the DESI-MSI analysis (*see* Subheadings 3.5 and 3.7). Figure 2 Shows an example of the histopathological analysis of a digital slide of prostate specimen that was considered in the earlier study by Banerjee et al. [32].

3.4 DESI-MSI

1. Before starting the DESI-MS imaging experiments, ensure that the instrument is properly calibrated and the MS inlet (extended ion transfer tube) is cleaned (*see* **Note 9**).
2. Attach the DESI source (either homebuilt or from Prosolia Inc.) to the MS instrument with a proper connection to the moving stage controller.

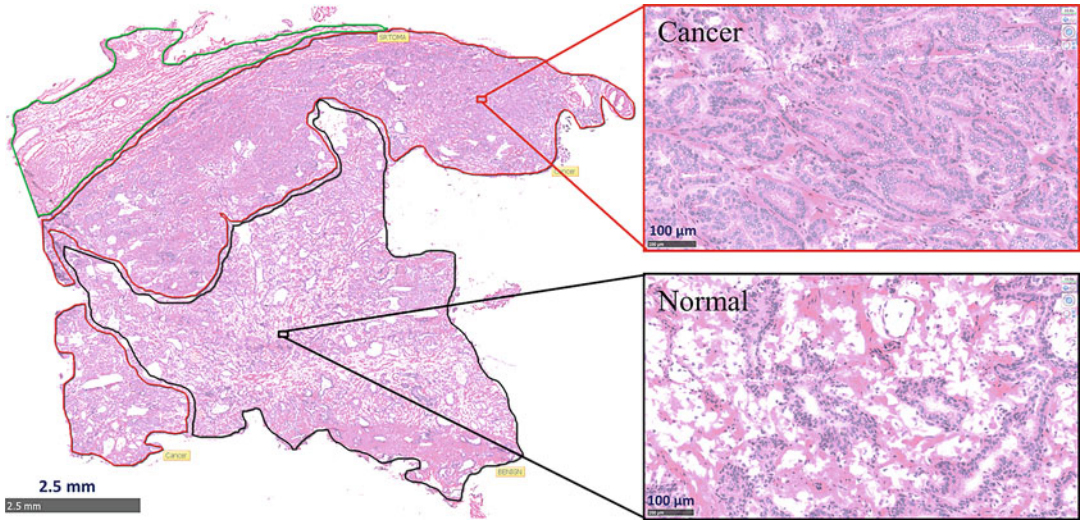


Fig. 2 High-resolution optical image of a typical prostate tissue specimen showing the area of cancer (red outline), normal (black outline), and stroma (green outline). Insert images represent 24× magnification of two typical regions of cancer and normal

3. Select the proper DESI solvent for imaging [34]. In the aforementioned reference, we had taken 1:1 mixture (*v/v*) of ACN and DMF (*see* **Note 10**).
4. In order to obtain a high-quality image, optimize the DESI-MS system by properly choosing different instrumental and imaging parameters. Table 1 Lists a number of such parameters, which were optimized in imaging a typical prostate tissue sample using a homebuilt DESI source coupled to a high-resolution Orbitrap-MS (*see* **Note 11**).
5. During an optimization of the DESI source, ensure that the diameter of the spray spot on a microscopic glass slide is around 200 µm or less, which can be seen with a naked eye (a normal human eye can see objects as small as 100 µm) (*see* **Note 12**).
6. Use a spare tissue section to optimize the signal intensity mostly by tuning different geometric parameters (*see* Table 1) and the solvent flow rate. Ion optics can also be tuned by selecting a specific peak from the *m/z* region of your interest. It has been observed that ion signal intensity ranging from 10^4 to 10^5 can be achieved in the negative ion mode imaging using the Orbitrap-MS.
7. Take out the slide containing the tissue of thickness 15 µm from the refrigerator, and dry it for approximately 10–15 min in the vacuum desiccator (*see* **Note 13**).
8. Measure the desired imaging area (rectangular) in the slide by a caliper, and note down the *x* and *y* lengths followed by a calculation of acquisition time, scan speed, and number of

Table 1
DESI-MS instrument and imaging parameters optimized for imaging a typical prostate tissue specimen using a homebuilt DESI source

Instrument parameters	Imaging parameters
Solvent: ACN/DMF (1:1, *v/v*)	Automatic gain control (AGC): Off
Flow rate: 0.7 μL/min	Spatial resolution/pixel size: 200 μm by 200 μm
Sheath gas (nitrogen) pressure: 170 psi	Scan time (ST) per pixel: 1.08 s[a]
Spray incident angle: 55°	*x* dimension of the image (slightly higher than tissue length): 14,000 μm
Droplet collection angle: 10°	*y* dimension of the image (slightly higher than the tissue width): 8000 μm
Spray tip-to-MS inlet distance: 5 mm	Number of pixels in the *x* dimension (one line scan): (14,000/200) = 70
Spray tip-to-surface distance: 2 mm	Number of pixels in the *y* dimension (number of line scans or rows): (8000/200) = 40
m/z range: 50–1000	Acquisition time (for one line scan in *x* dimension): 70 × 1.08 s = 75.6 s = 1.26 min
Ionization mode: Negative (−5 kV)	Scan speed (moving stage speed) in *x* dimension: (14,000/75.6) = 185.19 μm/s
MS resolution: 60,000	Total scan time for the whole image: ~1.26 × 40 min ~50.4 min (approximately 1 h as few seconds time lags are there between successive line scans)
Inlet capillary temperature: 275 °C	

[a]The Orbitrap-MS showed a scan time of 1.08 s per spectrum in the negative ion mode at the mass resolution 60,000. This scan time can be reduced by decreasing the mass resolution

line scans (number of pixels in the *y* dimension) as shown in Table 1 as a typical example (*see* **Note 14**). Some of these values are automatically calculated in the imaging software if the commercial DESI imaging source is used (Prosolia Inc.; *see* **Note 2**).

9. Adjust the DESI spray spot to the starting point [(0,0) position of the image] of the first-line scan of the image (*see* **Note 15**).
10. Create a sample list in the acquisition software (Xcalibur 2.2). The total number of samples in the list corresponds to the number of pixels in the y dimension (number of line scans; *see* Table 1). Set the acquisition time for each line scan in *x* dimension, e.g., 1.26 min in Table 1.
11. Set up the moving stage velocity (e.g., 185.19 μm/s in Table 1) along with sample dimensions (*x* and *y* values) and

pixel size on the moving stage controller software (*see* **Note 16**).

12. Ensure that the syringe contains enough solvent to acquire the whole image in the calculated scan time (Table 1).
13. Start the data acquisition.
14. After completion of the imaging experiment (data acquisition), store the glass slide in a closed slide box at room temperature until the H&E staining is performed on it (*see* **Note 17**).

3.5 Data Processing and Analysis

1. In order to construct the ion image, combine all Xcalibur raw data of line scans (e.g., 40 line scans in Table 1) into Analyze 7.5 format files (.img, .hdr, and .t2 m, which can be opened and read by BioMap) with a bin size of *m/z* 0.05 or less for preserving the peak resolution. Image generator software [47] was used in our study for this data conversion. Commercial DESI source (Prosolia Inc.) comes with their firefly imaging software for performing this data conversion and visualization.
2. Viewing the ion image: Load the .img file, as generated above, in the BioMap software (*see* **Note 18**), and then follow these instructions: (a) click on the window 1 (black), and then go to analysis in the menu bar > Plot > Point to open the mass spectrum window; (b) right click on the window 1, and then choose an average pixel (say 1) > click on Done; (c) expand the mass spectrum window in the desired *m/z* range, and then click the right button on a peak (e.g., *m/z* 709.4778); (d) select the desired color scale (e.g., rainbow color) from "Set the color table" tool bar in the left; and (e) change the minimum and maximum values on the slide bars to adjust the contrast of the image (e.g., Fig. 3a for *m/z* 709.4778). At this point, distribution of the metabolite/lipid can be visualized and correlated to the diagnostic feature of the tissue. For example, two species of *m/z* 709.4778 (Fig. 3a) and *m/z* 788.5409 (Fig. 3b) were differentially distributed in a typical prostate tissue specimen, which was later evaluated by histopathology (H&E). Results show that *m/z* 709.4778 is upregulated and *m/z* 788.5409 is downregulated in cancer as the left portion of the tissue is normal and the right portion of the tissue is cancer, as evaluated by histopathology (H&E not shown here). A number of ions with differential abundances in cancer and normal tissue can be identified by employing this tissue visualization technique on a large set of samples (*see* **Note 19**). One full scan (.img file), thus, provides hundreds of ion images, and the corresponding molecular species are generally found to be metabolites, lipids, sugar, etc. (*see* **Note 20**) when analyzed by tandem mass spectrometric technique (*see* Subheading 3.6).

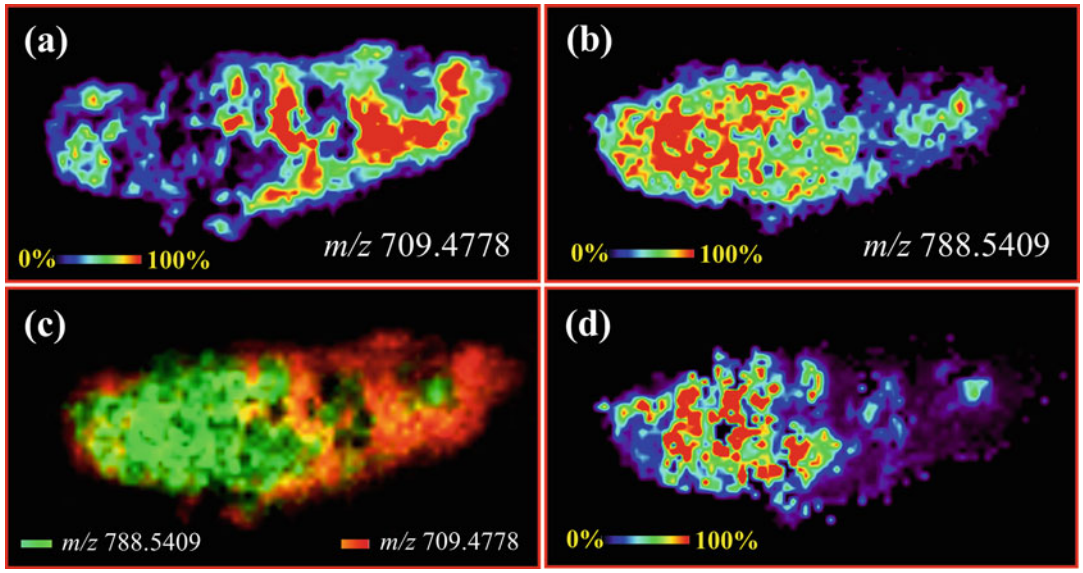


Fig. 3 Negative ion mode DESI-MS images of a typical prostate tissue specimen showing the distribution of (**a**) *m/z* 709.4778 [PA(20:4/17:0)], (**b**) *m/z* 788.5409 [PS(18:0/18:1)], (**c**) both *m/z* 709.4778 and *m/z* 788.5409 (overlaid image in bicolor), and (**d**) *m/z* 788.5409 to *m/z* 709.4778 ratio. Left portion is benign, and right portion is cancer in the tissue as evaluated by histochemistry. The abundance of the given ion in the corresponding ion image is normalized to 100%. Rainbow color order, wherever used, presents highest concentration by red and the lowest concentration by violet

3. Overlay of two ion images: Fig. 3c shows the overlay of the distribution of two species (*m/z* 709.4778 and 788.5409) on the same tissue. Instructions below will help to generate such an overlaid image in the BioMap:
   (a) Follow the above instructions to generate an ion image of *m/z* 709.4778 in the window 1.
   (b) Copy window 1 (Edit > Copy) and paste it in the window 2 (Edit > Paste > Scan).
   (c) Select *m/z* 788.5409 as described above to construct the ion image in the window 2.
   (d) Note down the "N" value corresponding to the peak of *m/z* 788.5409. This *N* value can be found underneath the mass spectrum when the cursor is kept on the desired peak (*m/z* 788.5409). In the example we give, the *N* value was found to be 26927 from the position: [788.541 (*m/z*), 973.702 (intensity)], #26927.
   (e) Copy window 2 (*m/z* 788.5409), and paste it in window 1 (*m/z* 709.4778) by clicking Edit > Paste > As Ovl;
   (f) Change the Overlay display mode to bicolor by right clicking the window 1 and then going to the Properties;

(g) Click on Windows>Ovl Control, and then change the N value accordingly ($N = 26{,}927$ in the present example) by moving the corresponding slide bar.

Contrast of the image can also be optimized by adjusting the minimum and maximum value in this window. At this point, the plot as shown in Fig. 3c is generated, showing different distributions of these two species in the cancer and normal region of the tissue.

4. Ratio of two ion images: Fig. 3d shows the distribution of the abundance ratio of two species (*m/z* 788.5409 to *m/z* 709.4778). Follow these instructions to generate such images:
   (a) Plot the ion image of *m/z* 788.5409 in window 1 as described in Subheading 3.5, **step 2**.
   (b) Expand the corresponding mass spectrum to view the peak of *m/z* 788.5409.
   (c) (c) Select this peak from left to right (a green bar will appear on the peak) by pressing the middle button of the mouse, and then change the output window to window 2 in the pop-up window box. This will create the extracted ion image of *m/z* 788.5409 in the window 2.
   (d) Similarly create another extracted ion image in the window 3.
   (e) Now click on the window 2 and then go to Tools > Calculation > Divide.
   (f) In the pop-up box, change the Output Device to Window 4, and select Windows 3 as the Reference.

At this point, the plot as shown in Fig. 3d appears, showing the distribution of this ratio (*see* **Note 21**). Contrast of this image can further be tuned as described before (Subheading 3.5, **step 2**).

3.6 Identification of Species

The marker ion signals for classifying cancer and normal tissue, as determined by BioMap analysis, need to be characterized and identified to understand their biological significance. Often, this identification is done based on high mass accuracy, isotopic distribution, and tandem mass analysis (MS/MS study). The following methods describe how to extract and identify the chemical species from an ion signal:

1. After DESI-MSI, scrape off the tissue from the microscopic glass plate using a single-edge lab razor blade, and collect the debris to a 2 mL microcentrifuge tube. Add 500 μL LC-MS grade methanol or any other desired solvent to this, and vortex for 10 min followed by sonication in an ultrasonic bath for another 5 min (*see* **Note 22**). Centrifuge this solution at 18000 × *g* in a microcentrifuge, and decant the supernatant

liquid, which should be further filtered through a 0.22 μm syringe filter. This filtered solution can be further diluted using methanol before injecting to the electrospray ionization mass spectrometer (high-resolution ESI-MS).

2. During electrospraying the above filtered solution in direct injection mode, maintain proper analyte concentration, low solution flow rate (1–5 μL/min), and optimum geometric and instrumental parameters (ion optics), and spray voltage (e.g., −5 kV in the negative ion mode) to get intense ion signals in the mass spectrum. Hundreds of peaks can be found in this MS^1 spectrum, with very high mass accuracy, ranging from very low intensity to very high intensity. Record and save this MS^1 data.
3. Mass select the desired peak with an isolation width of $\Delta m/z$ 0.8, and perform collision-induced dissociation (CID) on this species by applying the appropriate collision energy (normalized collision energy ~0–100%) so that the relative abundance of the parent ion becomes 10% in the MS^2 spectrum. Record and save this spectrum. Thus, all targeted peaks are mass selected sequentially (manually) for CID study. Likewise, if required, further CID (MS^3) can also be performed on a desired species mass selected from the MS^2 spectrum.
4. In order to identify the species, first monitor the highly resolved m/z value up to four decimal places. The list of lipids or metabolites corresponding to this mass (mostly deprotonated in the negative ion mode; *see* **Note 23**) can be searched from LIPID MAPS (http://www.lipidmaps.org/), the human Metabolome database (http://www.hmdb.ca/), or MassBank (http://www.massbank.jp/index.html?lang=en). Sort out few species from this list based on high mass accuracy (e.g., below 5 ppm) and isotopic distribution pattern (https://sites.google.com/site/isoproms/). Then look for the MS^2 fragments of those species, if available in the above databases, and compare those fragmentation patterns with that from the experimental MS^2 (from the tissue). In most cases, MS/MS fragments can be found from the aforementioned databases for the identification of the species (*see* **Note 24**). However, if the MS/MS fragments are not available in the database, a good knowledge of gas-phase ion chemistry is required to identify the structure of the ion based on its tandem mass analysis (MS^n) (*see* **Note 25**). Moreover, high-performance liquid chromatography (HPLC) and CID data of the suspected species can be compared with that of its synthetic (if possible) or commercial (if available) standard for the molecular identification (*see* **Note 26**).
5. Generally, DESI-MSI records the ion images (signals) of small metabolites (mainly observed in m/z 50–200), lipids (mainly

observed in *m/z* 200–1000), and sugars (glucose, erythrose, etc.). Lipids include fatty acid (FA), phosphatidic acid (PA), glycerophosphoglycerols (PG), glycerophosphoserines (PS), and glycerophosphoinositols (PI). For prostate cancer, a list of such molecular species can be found in the recent report by Banerjee et al. [32].

6. After finding the metabolite/lipid signatures and their up- or downregulation in cancer, try to interpret the observed metabolic profile and its biological significance to link that to the discipline of pathology (*see* **Note 27**). Understanding this biological significance can also help to investigate the potential therapeutic targets for pharmacological intervention in preclinical and clinical settings [48].

3.7 Data Extraction for Statistical Analysis

Although few species can be identified as potential markers for detecting cancer in the tissue using BioMap software, often an unbiased statistical approach (considering all detected ion signals) is adopted to identify biomarker candidates from large data set. Such approach has been shown to be very useful in determining the cancer margin [23, 30]. A training set is first constructed comprising samples with known pathology (H&E), and then a statistical classifier is developed, which is further tested with a cross validation process. To test the performance of this classifier, independent validation specimens are taken for analysis. Below are the details of data extraction method from the Xcalibur files (raw data) for statistical analysis.

1. Collect all Xcalibur line scan files (.raw) for an image in the designated folder with the patient code (say PC351 is the patient code).
2. Use MSConvert software to convert these .raw files (Xcalibur line scans) to .mzML files, and save them to a new folder (e.g., PC351_1) (*see* **Note 28**).
3. Use imzMLConverter to combine the above .mzML files into the .imzML file, and save this to a new folder (e.g., PC351_2) (*see* **Note 29**), which can be read by MSiReader software.
4. Open MSiReader (*see* **Note 30**) > click on load data > go to the specific folder (e.g., PC351_2) > open the .imzML file (e.g., PC351.imzML) > change *m/z* value to 281.2486 in the MS Navigation box (*see* Fig. 4; *see* **Note 31**) > change the *m/z* tolerance to 5 ppm > adjust the contrast of the image by changing the maximum value (e.g., 10,000 in Fig. 4) on the slide bar > enable polygon tool for region of interest (ROI) selection in the tool bar > choose a region (e.g., cancer or normal guided by histopathology; *see* Fig. 4) in the ion image by selecting multiple points > connect first and last point by double clicking to close the polygon > click on Export and/or

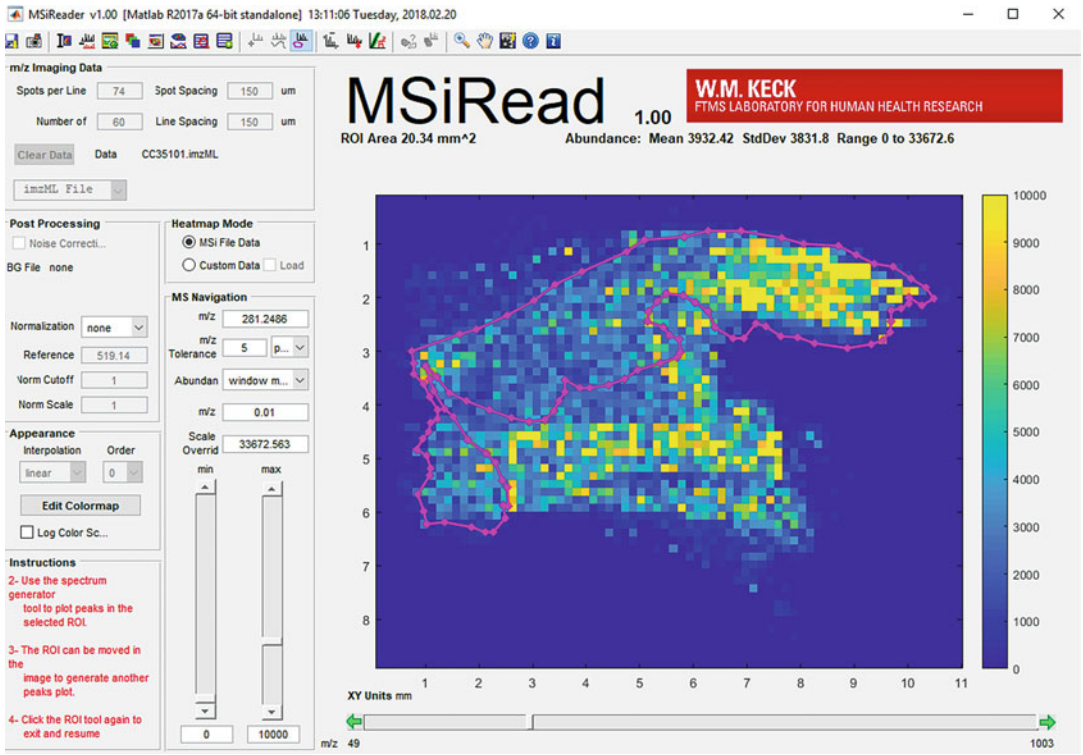


Fig. 4 Screenshot showing the selection of ROI in an ion image opened in MSiReader software. This ROI covers the pixels corresponding to the area of cancer in the prostate tissue (PC351) as analyzed by the pathologist (Fig. 2)

view mass spectrum data from selected pixels (*see* **Note 32**) > choose options like negative scans, parabolic centroid, export peaks to .txt file, export raw data for each pixels > browse the new folder (e.g., PC351_3) where the exported raw data will be saved > click OK. The raw data of each pixel from the ROI will be exported as .txt file.

5. For the convenience of statistical analysis, the above .txt file can be converted to excel or .csv file as below.

   Open Microsoft excel > click on File > go to the specific folder (PC351_3) > open the .txt file > select delimited and click on Next > click on Treat consecutive delimiters as one > click on Next > click on Finish. Data will be opened in the excel page > review the data (*m/z* vs abundance values extracted properly for all pixels in the ROI; Fig. 4) and save it as .csv file (*see* **Note 33**).

6. Although hundreds of metabolites and lipids are detected by DESI-MS, certain number of peaks, whose abundances are significant and characterized, can be selected for statistical analysis. Statistical calculation can be performed, both by

using the individual peaks and by using all possible ratios of two peaks in this list [32].

7. All mass spectral data (pixel-to-pixel) can be normalized to 100 during the statistical analysis (*see* **Note 34**).

4 Notes

1. It is not mandatary to use high-resolution mass spectrometer for DESI-MSI study. However, imaging experiment performed using a high-resolution mass spectrometer provides robust results. The present protocol uses LTQ Orbitrap XL mass spectrometer (a high-resolution mass spectrometer from Thermo Fisher Scientific), and therefore data acquisition, extraction, and processing are guided by the concerned software (Xcalibur). Nevertheless, a suitable mass spectrometer from any manufacturer (vendor) can be used in DESI-MS imaging.
2. Prosolia Inc. is the only commercial manufacturer for the DESI source (https://prosolia.com).
3. Precaution should be taken during embedding the tissues in OCT and then sectioning it. The blade should be cleaned properly with ethanol before its use for sectioning the tissue. The use of minimum amount of OCT is suggested to avoid the contamination of the tissue surface with OCT in the resected specimen. Unlike conventional H&E histopathology, DESI-MSI recognizes molecules on the tissue surface. This has been found that the DESI-MS data quality is greatly affected by a series of intense ion signals (peaks) from the polymeric components of OCT in the positive ion mode if the resected tissue surface is contaminated with traces of OCT. However, these ion signals from traces of OCT in negative ion mode are less intense, and those can be ignored. Because of the OCT interference, it is easier to pinpoint the tissue metabolites/lipids or small molecules in the negative ion mode as compared to positive ion mode of DESI-MSI analysis.
4. Temperature setting can be changed depending on the tissue type. Avoid folding and cracking of the tissue section. DESI-MSI offers good results from sections, which have smooth and flat surface.
5. The 5 μm section is sent to a histopathologist for evaluation (H&E), and the 15 μm section, taken immediately adjacent to the above 5 μm section, is used for DESI-MSI. Tissue morphologies for both of the aforementioned sections are likely to be similar, suggesting that pathological features are similar in both sections.

6. Depending on the type of samples, frozen sections can be preserved for several months in the ultra-freezer before taking them for DESI-MSI analysis. We found no significant change in metabolic/lipidomic profiles in the DESI-MSI data even if the sections were preserved for a year at −80 °C (tested on two adjacent sections from a prostate tissue specimen) [32].
7. Apart from H&E-based histopathology, immunohistochemistry is another process of staining the tissue and locating the cancer cells. This is based on the interaction of antigen (a specific protein in the tissue) and an antibody tagged to a fluorophore.
8. A large space in the hard disc is required to store these images as each image file (.ndpi) can be of several hundreds of megabytes depending on the image resolution and actual size of the tissue. Details of the operation of the instrument can be found in the instrument manual.
9. Cleaning of the inlet capillary can be done by sonicating in a mixture of (1:1 *v/v*) methanol/water for 15 min.
10. Polarity of the spray solvent plays an important role in extracting the molecular species from the tissue. Different solvent systems can extract and ionize various molecular species to different extents [34]. In the present study, we have used a mixture of ACN/DMF (1:1, *v/v*), as we have found that this solvent system can cover many metabolites/lipids in the mass range of $m/z = 50–1000$. Further, this solvent system was reported to be histologically compatible [49], which means that it does not destroy the native tissue morphology even after the spot-by-spot extraction of molecular species by DESI-MSI. Therefore, the same slide can be taken for H&E study in histopathology after the DESI-MSI experiment [32].
11. A detailed method of construction of a DESI source and setting these parameters was reported earlier in the literature [25, 34, 44, 45]. However, a commercial DESI source (Omni Spray, Prosolia, Indianapolis, IN), with a friendly user interface, can automatically calculate many of these parameters in the software (Omni spray 2D software).
12. This spray spot on the glass slide can be viewed by a naked eye if the back side of the slide is marked black by a king-size sharpie. The spray spot can be optimized to nearly a stable circular shape (minimal side splashing) of diameter ~200 μm by slightly adjusting different geometric parameters and the solvent flow rate (Table 1). High solvent flow rate can result in excessive wetting of the surface and cross contamination. The spray tip should also be blunt for providing a stable spray.

13. Dry the tissue until all visible water is gone. Typically, it takes 10–15 min when 1/8–1/6 HP vacuum pump is used. Avoid overdrying the tissue in high vacuum for a prolonged period, which might cause rupturing/cracking of the tissue.
14. We considered the resolution of the image (pixel size) 200 μm here. However, one can choose to work with 150 μm resolution if the spray spot size can be made like that. Please note that decreasing the resolution will increase the total scan time of the tissue.
15. One image is comprised of many line scans (in rows). The pixel size defines how many line scans (rows) are required for imaging a given area (Table 1). Please note that whole ion image of the tissue is not strictly required for the diagnosis. Sometimes, simple line scan or rapid scribble scanning can offer valuable diagnostic information from the tissue [3, 32], which also saves time of analysis.
16. The moving stage velocity is automatically calculated and taken by the commercial software (Omni Spray 2D) once the sample dimension and the pixel size are defined.
17. As mentioned above (*see* **Note 10**), H&E staining can be performed on this slide after DESI-MSI as the tissue morphology remains preserved. This can also be judged by comparing its high-resolution optical image (H&E) with the same of its adjacent section (5 μm) preserved for histochemistry.
18. BioMap is written in IDL. Therefore, IDL Virtual Machine (IDL VM) is necessary to be installed on the computer. Details of its installation and demo can be found here: https://ms-imaging.org/wp/biomap/.
19. Being a label-free unbiased method, DESI-MSI method can generate hundreds of ion images of different molecular species (of different m/z values as analyzed using BioMap), as mass spectrum from individual pixel provides hundreds of ion signals. As cancer cells reprogram their metabolic pathways, some potential markers (e.g., metabolites, lipids) can be identified in this way, being differentially distributed in cancer and normal specimens.
20. Imaging protein by DESI-MSI is very challenging, and we are still struggling to map proteins in tissues using this imaging technique. It appears that proteins, being large macromolecules, strongly adhere to the tissue by several interactions (hydrophilic and hydrophobic). Special type of sample treatment is required to image protein by DESI-MSI. Recently, use of nanoDESI-MSI has allowed imaging of some proteins in mouse brain sections [50].

21. Sometimes this ratio distribution offers a valuable diagnostic feature. In the study with prostate patients, it was found that distribution of glucose/citrate ratio in the tissue can identify cancer with high accuracy [32].
22. A mini tissue homogenizer can also be employed for this purpose, instead of sonication.
23. Species are mostly deprotonated in the negative ion mode and protonated in the positive ion mode [26]. However, sometimes adducts with cations or anions are also observed in positive and negative ion modes, respectively. If the species are inherently charged (e.g., cholines), they can also directly appear in the mass spectrum depending on polarity.
24. The CID spectra of the mass-selected ions from the tissue specimens are sometimes complex, although the majority of the ion fragments match with that of standards in the database. This complexity can be interpreted by the interference of isomeric/isobaric ions obtained from the biological matrix (tissue). As the position and stereochemistry of the double bond in the fatty acid (FA) complicate the structural elucidation of lipids, they are often tentatively assigned in the structures of FAs and glycerophospholipids [32]. CID also enables identification of isomeric species (same m/z values) although they produce single-ion signal in the MS^1 spectrum [32].
25. The task is like a jigsaw puzzle, where the players have been given a broken plate and asked to join them together in a sensible way to find the actual shape of the plate [26].
26. It takes a great effort to identify the molecular species by tandem mass spectrometry, particularly when the MS/MS data is not available in the database.
27. Metabolic profile of cancer cells is strikingly different from normal cells. For example, metabolism in the prostate gland presents distinctly different kinetics (metabolic flux) of the Krebs cycle compared with other organs [51]. DESI-MSI is an ideal method to exploit these differences in cancer biochemistry in vivo because it is tissue based, does not require tissue fixation, and provides reasonably good spatial resolution (~200 μm) [32]. Two important strengths of DESI-MSI are its speed and the need for little sample preparation. These characteristics could make DESI-MS a particularly promising and rapid point-of-care clinical test.
28. MSConvert can be found in the program list of your computer once you have installed the ProteoWizard freeware [52].
29. No installation is required for imzMLConverter, but extraction of all associated files in it is required for the performance of this freeware. Installation of Java is required in the computer as

imzMLConverter is written in Java [53]. Click on the imzMLConverter.jar file in the extracted folder to open the software directly.

30. No installation is required for MSiReader but extraction of all the associated files in it. Install MATLAB Complier Runtime R2017b (9.2) before launching this software, as this is a MATLAB application. Those who are unfamiliar with MATLAB language can also use MSiReader as it has a user-friendly interface [54]. User manual for this software can be found in the extracted folder. High memory resource (e.g., RAM $\geq$16 GB) is required to handle high-resolution MS data (small *m/z* increment) easily.
31. *m/z* 281.2486 corresponds to deprotonated oleic acid, which can be found in almost all biological tissues. Often, constructing the ion image of oleic acid might represent the shape of the whole tissue. However, one can select *m/z* value of any other appropriate lipids/metabolites to present the shape of the tissue.
32. To reduce the time required for data extraction, open "MSiReaderPrefs.INI" from the MSiReaderRelease folder (extracted), and change the setting "ExportToExcel" to false instead of true. This will export the file in a text format, which is up to 300 times faster than an Excel workbook.
33. These .csv files consist of the information of *m/z* values of several species and their abundances collected from each pixel in the ROI. Statistical analysis is convenient with these .csv files. Data collected from cancer and normal patients or from cancer and normal regions of the tissues (Fig. 2) are then fed into the statistical analysis for the classification of cancer and normal or determining the cancer margin.
34. Normalization helps to avoid instrumental effects on ion currents from different specimens studied at different times (days).

Acknowledgments

Authors are thankful to Prof. Richard N. Zare for his encouragement and advice. S.B. gratefully acknowledges the help from Dr. Livia S. Eberlin, Prof. Robert J. Tibshirani, Prof. James D. Brooks, Dr. Christian A. Kunder, Dr. Rosalie Nolley, Dr. Richard Fan, and Dr. Geoffrey A. Sonn. This work was supported by the Indian Institute of Science Education and Research, Tirupati, and Saha Institute of Nuclear Physics, Kolkata. S.B. is also thankful to Science and Engineering Research Board, Department of Science and Technology, Government of India, for providing Ramanujan Fellowship Research Grant (SB/S2/RJN-130/2017).

References

1. Zhang J, Rector J, Lin JQ, Young JH, Sans M, Katta N, Giese N, Yu W, Nagi C, Suliburk J, Liu J, Bensussan A, DeHoog RJ, Garza KY, Ludolph B, Sorace AG, Syed A, Zahedivash A, Milner TE, Eberlin LS (2017) Nondestructive tissue analysis for ex vivo and in vivo cancer diagnosis using a handheld mass spectrometry system. Sci Transl Med 9(406): eaan3968
2. Balog J, Sasi-Szabó L, Kinross J, Lewis MR, Muirhead LJ, Veselkov K, Mirnezami R, Dezső B, Damjanovich L, Darzi A, Nicholson JK, Takáts Z (2013) Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. Sci Transl Med 5 (194):194ra193–194ra193
3. Pirro V, Alfaro CM, Jarmusch AK, Hattab EM, Cohen-Gadol AA, Cooks RG (2017) Intraoperative assessment of tumor margins during glioma resection by desorption electrospray ionization-mass spectrometry. Proc Natl Acad Sci U S A 114(26):6700–6705
4. Hao J-J, Lin D-C, Dinh HQ, Mayakonda A, Jiang Y-Y, Chang C, Jiang Y, Lu C-C, Shi Z-Z, Xu X, Zhang Y, Cai Y, Wang J-W, Zhan Q-M, Wei W-Q, Berman BP, Wang M-R, Koeffler HP (2016) Spatial intratumoral heterogeneity and temporal clonal evolution in esophageal squamous cell carcinoma. Nat Genet 48:1500
5. Axelson H, Fredlund E, Ovenberger M, Landberg G, Påhlman S (2005) Hypoxia-induced dedifferentiation of tumor cells—a mechanism behind heterogeneity and aggressiveness of solid tumors. Semin Cell Dev Biol 16(4):554–563
6. Mazor T, Pankov A, Song JS, Costello JF (2016) Intratumoral heterogeneity of the epigenome. Cancer Cell 29(4):440–451
7. Levine AJ, Puzio-Kuter AM (2010) The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science 330(6009):1340–1344
8. Vander Heiden MG, Cantley LC, Thompson CB (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science 324(5930):1029–1033
9. Iurlaro R, León-Annicchiarico CL, Muñoz-Pinedo C (2014) Chapter three - regulation of cancer metabolism by oncogenes and tumor suppressors. In: Galluzzi L, Kroemer G (eds) Methods in enzymology, vol 542. Academic Press, New York, pp 59–80. https://doi.org/10.1016/B978-0-12-416618-9.00003-0
10. Vermeersch KA, Styczynski MP (2013) Applications of metabolomics in cancer research. J Carcinog 12:9
11. DeBerardinis RJ, Chandel NS (2016) Fundamentals of cancer metabolism. Sci Adv 2(5): e1600200
12. Baenke F, Peck B, Miess H, Schulze A (2013) Hooked on fat: the role of lipid synthesis in cancer metabolism and tumour development. Dis Model Mech 6(6):1353–1363
13. van Meer G, Voelker DR, Feigenson GW (2008) Membrane lipids: where they are and how they behave. Nat Rev Mol Cell Biol 9 (2):112–124
14. Andronesi OC, Rapalino O, Gerstner E, Chi A, Batchelor TT, Cahill DP, Sorensen AG, Rosen BR (2013) Detection of oncogenic IDH1 mutations using magnetic resonance spectroscopy of 2-hydroxyglutarate. J Clin Invest 123 (9):3659–3663
15. Leather T, Jenkinson MD, Das K, Poptani H (2017) Magnetic resonance spectroscopy for detection of 2-hydroxyglutarate as a biomarker for IDH mutation in gliomas. Metabolism 7 (2):29
16. Kurhanewicz J, Vigneron DB (2008) Advances in MR spectroscopy of the prostate. Magn Reson Imaging Clin N Am 16(4):697–69x
17. Spratlin JL, Serkova NJ, Gail Eckhardt S (2009) Clinical applications of metabolomics in oncology: a review. Clin Cancer Res 15 (2):431–440
18. Armitage EG, Southam AD (2016) Monitoring cancer prognosis, diagnosis and treatment efficacy using metabolomics and lipidomics. Metabolomics 12(9):146
19. Ifa DR, Eberlin LS (2016) Ambient ionization mass spectrometry for cancer diagnosis and surgical margin evaluation. Clin Chem 62 (1):111–123
20. Takáts Z, Wiseman JM, Gologan B, Cooks RG (2004) Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. Science 306(5695):471–473
21. Wiseman JM, Ifa DR, Song Q, Cooks RG (2006) Tissue imaging at atmospheric pressure using desorption electrospray ionization (DESI) mass spectrometry. Angew Chem Int Ed 45(43):7188–7192
22. Eberlin LS, Norton I, Orringer D, Dunn IF, Liu X, Ide JL, Jarmusch AK, Ligon KL, Jolesz FA, Golby AJ, Santagata S, Agar NYR, Cooks RG (2013) Ambient mass spectrometry for the intraoperative molecular diagnosis of human

brain tumors. Proc Natl Acad Sci U S A 110 (5):1611–1616

23. Eberlin LS, Tibshirani RJ, Zhang J, Longacre TA, Berry GJ, Bingham DB, Norton JA, Zare RN, Poultsides GA (2014) Molecular assessment of surgical-resection margins of gastric cancer by mass-spectrometric imaging. Proc Natl Acad Sci U S A 111(7):2436–2441
24. Eberlin LS, Gabay M, Fan AC, Gouw AM, Tibshirani RJ, Felsher DW, Zare RN (2014) Alteration of the lipid profile in lymphomas induced by MYC overexpression. Proc Natl Acad Sci U S A 111(29):10450–10455
25. Eberlin LS (2014) DESI-MS imaging of lipids and metabolites from biological samples. In: Raftery D (ed) Mass spectrometry in metabolomics: methods and protocols. Springer New York, New York, NY, pp 299–311. https://doi.org/10.1007/978-1-4939-1258-2_20
26. Banerjee S, Mazumdar S (2012) Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. Int J Analyt Chem 2012:282574
27. Fenn J, Mann M, Meng C, Wong S, Whitehouse C (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246(4926):64–71
28. Kebarle P, Verkerk UH (2010) On the mechanism of electrospray ionization mass spectrometry (ESIMS). In: Electrospray and MALDI mass spectrometry. John Wiley & Sons, Inc., New York, pp 1–48. https://doi.org/10.1002/9780470588901.ch1
29. Perry RH, Bellovin DI, Shroff EH, Ismail AI, Zabuawala T, Felsher DW, Zare RN (2013) Characterization of MYC-induced tumorigenesis by in situ lipid profiling. Anal Chem 85 (9):4259–4262
30. Eberlin LS, Margulis K, Planell-Mendez I, Zare RN, Tibshirani R, Longacre TA, Jalali M, Norton JA, Poultsides GA (2016) Pancreatic cancer surgical resection margins: molecular assessment by mass spectrometry imaging. PLoS Med 13(8):e1002108
31. Gouw AM, Eberlin LS, Margulis K, Sullivan DK, Toal GG, Tong L, Zare RN, Felsher DW (2017) Oncogene KRAS activates fatty acid synthase, resulting in specific ERK and lipid signatures associated with lung adenocarcinoma. Proc Natl Acad Sci U S A 114 (17):4300–4305
32. Banerjee S, Zare RN, Tibshirani RJ, Kunder CA, Nolley R, Fan R, Brooks JD, Sonn GA (2017) Diagnosis of prostate cancer by desorption electrospray ionization mass spectrometric imaging of small metabolites and lipids. Proc Natl Acad Sci U S A 114(13):3334–3339
33. Eberlin LS, Norton I, Dill AL, Golby AJ, Ligon KL, Santagata S, Cooks RG, Agar NYR (2012) Classifying human brain tumors by lipid imaging with mass spectrometry. Cancer Res 72(3):645–654
34. Eberlin LS, Ferreira CR, Dill AL, Ifa DR, Cooks RG (2011) Desorption electrospray ionization mass spectrometry for lipid characterization and biological tissue imaging. Biochim Biophys Acta 1811(11):946–960
35. Zhang J, Feider CL, Nagi C, Yu W, Carter SA, Suliburk J, Cao HST, Eberlin LS (2017) Detection of metastatic breast and thyroid cancer in lymph nodes by desorption electrospray ionization mass spectrometry imaging. J Am Soc Mass Spectrom 28(6):1166–1174
36. Dória ML, McKenzie JS, Mroz A, Phelps DL, Speller A, Rosini F, Strittmatter N, Golf O, Veselkov K, Brown R, Ghaem-Maghami S, Takats Z (2016) Epithelial ovarian carcinoma diagnosis by desorption electrospray ionization mass spectrometry imaging. Sci Rep 6:39219
37. Abbassi-Ghadi N, Golf O, Kumar S, Antonowicz S, McKenzie JS, Huang J, Strittmatter N, Kudo H, Jones EA, Veselkov K, Goldin R, Takats Z, Hanna GB (2016) Imaging of esophageal lymph node metastases by desorption electrospray ionization mass spectrometry. Cancer Res 76 (19):5647–5656
38. Alfaro CM, Jarmusch AK, Pirro V, Kerian KS, Masterson TA, Cheng L, Cooks RG (2016) Ambient ionization mass spectrometric analysis of human surgical specimens to distinguish renal cell carcinoma from healthy renal tissue. Anal Bioanal Chem 408(20):5407–5414
39. Calligaris D, Feldman DR, Norton I, Brastianos PK, Dunn IF, Santagata S, Agar NYR (2015) Molecular typing of Meningiomas by desorption electrospray ionization mass spectrometry imaging for surgical decision-making. Int J Mass Spectrom 377:690–698
40. Guenther S, Muirhead LJ, Speller AVM, Golf O, Strittmatter N, Ramakrishnan R, Goldin RD, Jones E, Veselkov K, Nicholson J, Darzi A, Takats Z (2015) Spatially resolved metabolic phenotyping of breast cancer by desorption electrospray ionization mass spectrometry. Cancer Res 75(9):1828–1837
41. Calligaris D, Caragacianu D, Liu X, Norton I, Thompson CJ, Richardson AL, Golshan M, Easterling ML, Santagata S, Dillon DA, Jolesz FA, Agar NYR (2014) Application of desorption electrospray ionization mass spectrometry imaging in breast cancer margin analysis. Proc Natl Acad Sci U S A 111(42):15184–15189

42. Abbassi-Ghadi N, Veselkov K, Kumar S, Huang J, Jones E, Strittmatter N, Kudo H, Goldin R, Takats Z, Hanna GB (2014) Discrimination of lymph node metastases using desorption electrospray ionisation-mass spectrometry imaging. Chem Commun 50 (28):3661–3664
43. Calligaris D, Norton I, Feldman DR, Ide JL, Dunn IF, Eberlin LS, Cooks RG, Jolesz FA, Golby AJ, Santagata S, Agar NY (2013) Mass spectrometry imaging as a tool for surgical decision-making. J Mass Spectrom 48 (11):1178–1187
44. Ifa DR, Wiseman JM, Song Q, Cooks RG (2007) Development of capabilities for imaging mass spectrometry under ambient conditions with desorption electrospray ionization (DESI). Int J Mass Spectrom 259(1):8–15
45. Manicke NE, Kistler T, Ifa DR, Cooks RG, Ouyang Z (2009) High-throughput quantitative analysis by desorption electrospray ionization mass spectrometry. J Am Soc Mass Spectrom 20(2):321–325
46. Kiernan J (2015) Histological and Histochemical methods: theory and practice. Scion Publishing Limited, Banbury
47. Xiong X, Xu W, Eberlin LS, Wiseman JM, Fang X, Jiang Y, Huang Z, Zhang Y, Cooks RG, Ouyang Z (2012) Data processing for 3D mass spectrometry imaging. J Am Soc Mass Spectrom 23(6):1147–1156
48. Kishton RJ, Rathmell JC (2015) Novel therapeutic targets of tumor metabolism. Cancer J 21(2):62–69
49. Eberlin LS, Ferreira CR, Dill AL, Ifa DR, Cheng L, Cooks RG (2011) Non-destructive, histologically compatible tissue imaging by desorption electrospray ionization mass spectrometry. Chembiochem 12(14):2129–2132
50. Hsu C-C, Chou P-T, Zare RN (2015) Imaging of proteins in tissue samples using nanospray desorption electrospray ionization mass spectrometry. Anal Chem 87(22):11171–11175
51. Dakubo GD (2010) The Warburg phenomenon and other metabolic alterations of cancer cells. In: Mitochondrial genetics and cancer. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 39–66. https://doi.org/10.1007/978-3-642-11416-8_2
52. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak M-Y, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL, Mallick P (2012) A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol 30:918
53. Race AM, Styles IB, Bunch J (2012) Inclusive sharing of mass spectrometry imaging data requires a converter for all. J Proteomics 75 (16):5111–5112
54. Robichaud G, Garrard KP, Barry JA, Muddiman DC (2013) MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform. J Am Soc Mass Spectrom 24(5):718–721

Chapter 16

Compositional Analysis of the Human Microbiome in Cancer Research

Elisa Morales, Jun Chen, and K. Leigh Greathouse

Abstract

Gut microbial composition has shown to be associated with obesity, diabetes mellitus, inflammatory bowel disease, colitis, autoimmune disorders, and cancer, among other diseases. Microbiome research has significantly evolved through the years and continues to advance as we develop new and better strategies to more accurately measure its composition and function. Careful selection of study design, inclusion and exclusion criteria of participants, and methodology are paramount to accurately analyze microbial structure. Here we present the most up-to-date available information on methods for gut microbial collection and analysis.

Key words Metagenomic sequencing, Gut microbiota, Taxonomic classification, Cancer research

1 Introduction

Health-related microbiome research has increased exponentially as we discover the influence of the microbiome on diseases not previously recognized to have a microbial component. Certain cancers, for example, are known to have a microbial origin, but methods to comprehensively examine the microbiome in relationship to cancer have been hindered until recently. Methodologies to analyze the microbiome are quickly evolving to provide the best techniques in order to expand our understanding of the relationship between the microbiome and diseases. The main objective of this chapter is to provide general guidelines of existing methods as well as important points to consider when conducting microbiome research (especially focused on gut microbiota). The first part of this chapter will focus on microbial ecology and its evolution to provide the reader with a general background before engaging in metagenomic applications to cancer biology. The second part of this chapter involves a comprehensive overview of experimental methods as well as several step-by-step procedures on sampling, storage, handling, nucleotide extraction, PCR amplification, and gene sequencing of the human microbiome.

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_16,

1.1 Microbial Ecology History

A sequence of key discoveries marked the evolution of microbiology, beginning with the creation of microscopes and expanding to the application of microbiology in personalized medicine. Key observations demonstrate that trillions of microbes in and on our bodies play relevant roles in health and disease at multiple body sites such as the oral cavity, the gut, the skin, the reproductive tract, and the respiratory tract [1].

The complex and adaptable microbial communities in the human body interact with each other by exchanging nutrients, genetic information, and chemical signals [1]. The study of the interactions of microorganisms with their environment, also known as microbial ecology, has recently received increased attention due to the advent of next-generation sequencing technology. Prior to this advance, species identification required the organism to grow in a pure culture (culture-dependent); however, early bacterial cultures were monogenic, preventing any conclusion being drawn at the ecosystem level [2]. Furthermore, in 1987, Brock discovered that the activity of organisms cultivated in laboratories may not necessarily reflect their activity in situ [3]. Supporting Brock's theory, Amann et al. demonstrated that a good portion of the microorganisms in the environment do not survive cultivation [4]. These observations highlighted the inability of culture-dependent approaches to describe microbial diversity and activity. Thus, the need for culture-independent approaches to determine microbial composition surpassed the capabilities of microbial cultivation. To capture the full biological and environmental influence on host-associated microbial communities, it is crucial to further our knowledge by studying microbial communities in their natural habitats, constrained to their host's biochemical and physiological context.

When Carl Woese and George Fox proposed a new domain, Archaea, they ushered forth a new strategy for microbial evolution using phylogenetic taxonomy of the 16S rRNA gene [5]. Researchers have since taken advantage of the presence of both conserved and variable regions in the 16S rRNA gene. By amplifying variable regions, researchers can easily classify bacteria within their clades (discussed in detail in a later section). The first culture-independent approach utilizing phylogenetic taxonomy of the 16S rRNA gene was introduced in the 1980s allowing the extraction, amplification, cloning, and characterization of 16S rRNA genes obtained directly from the natural environment [6]. Analysis of 16S rRNA set the stage for high-throughput sequencing, which allows identification of entire bacterial communities within hours.

1.2 The Evolution of Microbial Ecology Composition and Function Assessment with Metagenomic Sequencing

Metagenomics involves the culture-independent, genome-level characterization of microbial communities by using high-throughput sequencing analyses. Metagenomics overcomes the biggest obstacle in microbiology by being able to assess the high diversity of uncultivated viable and nonviable microbes. "Metagenomics" and "functional metagenomics" are derived from "traditional genomics" and "functional genomics," respectively. While traditional genomics focuses on sequencing individual genomes, metagenomics aims at characterizing the genes at the community level. Furthermore, functional genomics assesses the function of the genome at the organism level, while "functional metagenomics" evaluates the function of the genes present in a community [1]. In the following paragraphs, we will describe two common techniques used to profile polymicrobial communities: 16S rRNA gene sequencing and whole genome sequencing (WGS). For methodological purposes, however, we will focus mainly on methods used in analysis of the 16S rRNA gene.

A metagenomic approach that is commonly utilized as the first step in pilot and larger metagenomic projects is the analysis of the 16S ribosomal RNA gene or 16S rRNA. Sequencing the non-conserved regions of this gene is used to analyze phylogenetic taxonomy and monitor changes over time in a particular community. Although 16S rRNA gene analysis provides information about the structure and composition of the microbial community, it only delivers inferences (based on marker genes) about the functional capabilities (metabolic potential). To overcome this shortcoming, shotgun DNA or whole genome sequencing (WGS) and RNA sequencing (RNA-seq) are used to reveal more detailed information about community diversity and function as compared to 16S rRNA gene analyses [1]. Next-generation sequencing (NGS), also called massive parallel sequencing, encompasses these technologies and can determine an entire human genome by sequencing millions of small DNA fragments at a time. These fragments are joined together by comparing to a reference human genome using bioinformatic techniques to build a complete genome. Similarly, NGS can be used to sequence entire genomes of polymicrobial communities or can be constrained to individual species of interest, which include 16S rRNA gene, WGS, and RNA-seq [2, 7, 8]. Although at a higher cost, WGS and RNA-seq can provide more in-depth information regarding the composition and metabolic potential of the microbial communities in comparison to 16S rRNA gene analysis. Further, these types of sequencing, as well as full-length 16S rRNA sequencing, overcome the drawback of using individual 16S rRNA variable regions to enable species- and strain-level information.

Microbial functionality can be further analyzed by using "metatranscriptomics" (analysis of the RNA transcripts from the microorganisms present in an ecosystem (RNA-seq)), "metabolomics"

(analysis of the metabolites from the microorganisms present in an ecosystem), and/or "metaproteomics" (analysis of the proteins produced by the microorganisms present in an ecosystem). Nonetheless, the instability, short half-life of RNA, and the low correlation between RNA transcript levels and the synthesis of its corresponding proteins have lowered the use of metatranscriptomics. Whereas, proteomics has advanced quickly by using mass spectrometry, which allows rapid and precise protein identification [1, 7]. Nevertheless, the integration of "omic" disciplines has the potential to link genetic structure to functional diversity [9, 10].

1.3 Metagenomic Sequencing Applications in Cancer Biology

Research into drug sensitivity, metabolic mutations, sources of infection outbreaks, and cancer detection, among others, are some of the medical uses of NGS, and the decrease in cost of sequencing will continue to increase its use in medicine and cancer research [8, 11]. With the commercialization of NGS in 2005, cancer medicine has experienced numerous advances in early detection and treatment using molecular profiling [12]. More recently, however, we have started to apply this technology to microbial associated diseases, including cancer. While we have known for decades that specific microbial species are causing cancer (e.g., *H. pylori* and gastric cancer), less has been understood about the contribution of polymicrobial communities or biofilms not only to cancer initiation but also to cancer promotion, progression, or metastasis [13–15]. The application of metagenomics to the field of cancer biology, however, has dramatically expanded the research questions investigators can now address. Multiple lines of evidence show alterations in the microbiome are associated with both development and progression of several types of cancer including liver, esophageal, and colon cancer [16–21]. With metagenomics we can more fully understand how the structure of the bacterial community determines the function and health of the epithelium and the immune system [22, 23]. For example, putative microbial initiators of colon cancer, specifically *B. fragilis* (ETBF), *E. coli* (pks+), and *E. faecalis*, are more prevalent in human tumors and/or those with familial adenomatous polyposis as compared to healthy controls [24–28]. These findings and those from similar studies were revealed through a combination of large metagenomic studies of multiple cohorts and murine models to link associations with molecular mechanism. Further, several bacteria have been identified through metagenomic sequencing as promoters of colon cancer development, including *Fusobacterium nucleatum*, which has been isolated from patients with inflammatory bowel disease, a risk factor for colon cancer [29, 30]. Building on these findings, recent attempts to generate a classifier for distinguishing normal colon from colon adenocarcinoma using the fecal microbiome in combination with the FOBT or FIT test have shown promise in improving screening for early detection of colon cancer [31].

Response to anticancer therapy has also shown to be reliant upon both the immune system and microbiota composition in animal and human studies [23]. Using 16S rRNA sequencing, Iida et al. (2013) observed that destruction of the microbiome in mice through use of antibiotics resulted in a compromised efficacy of immunotherapy (anti-IL-10/CpG-OGN) and platinum-based chemotherapeutics (e.g., oxaliplatin) [32]. Therefore, the commensal microbiota was recognized as necessary for the cytotoxic effects of oxaliplatin. With the advent of checkpoint inhibitors for anticancer therapy against PD-1/PD-L1 and CTLA4, studies have focused on determining the influence of the microbiome in therapeutic efficacy [33]. Building on these studies, Gopalakrishnan et al. [23] used metagenomics to characterize baseline fecal samples from melanoma patients undergoing immunotherapy (e.g., anti-PD-1). They used 16S rRNA sequencing to obtain microbial community composition and predict responders and nonresponders, demonstrating the power of the microbiome and metagenomics to uncover clinically pertinent information. Thus, mounting evidence indicates that specific microbial species and community states play a role in the etiology and treatment outcome of several types of cancer and that precise methods for analyzing these microbial populations are required. Hence, the remainder of this chapter will focus on delineating those methods.

1.4 Study Design, Sampling, Storage, and Handling Protocols for Fecal Collection

Multiple factors can affect the results of human microbiome studies. The greatest sources of variability leading to spurious results come from study design, nucleotide extraction, and statistical analysis methods [34, 35]. Therefore, this section will focus on these most salient factors and provide suggestions for additional resources and readings for factors not addressed. For a comprehensive overview of microbiome study design and analysis, *see* the review by Mallick et al. [10].

One of the initial steps in studies involving the human microbiota is determining the study design, which includes proper identification of the inclusion and exclusion criteria, as well as careful calculation of sample size. Some of the key issues to consider when selecting the inclusion and exclusion criteria for microbiome studies are age, current and past medication use (especially antibiotics), diet, hormonal status, smoking status, recent GI infections, geographical location, and even physical activity levels. Since all of these factors can have an impact on the gut microbiota, they should be either considered in the inclusion/exclusion criteria or statistically controlled [36–39]. Determining the effect of age is fundamental as children under 3 years old have a relatively unstable gut microbiota that diversifies and stabilizes as the child gets older, with a slight decrease in older stages of life [40]. One of the strongest factors that contributes to confounding results is antibiotic use; having consumed antibiotics alters the gut microbiota from

6 months up to 4 years before it recovers its initial composition [41]. In addition to antibiotics, proton pump inhibitors and the diabetic drug metformin are able to significantly modify the composition of the gut bacteria [42, 43].

When evaluating the longitudinal effects of specific factors that influence microbial community, careful selection of appropriate controls should be put in place in order to differentiate between normal variation from baseline and actual change [1, 44]. For example, stronger conclusions can be made when participants act as their own control, such as when comparing participants diagnosed with a relapsing disease in longitudinal studies. Nonetheless, when this design is not possible, between group comparisons are also informative, given enough participants are sampled and appropriate controls are selected [45]. Between group comparisons require the careful selection of the interest and control groups with as many similar characteristics as possible (age, gender, physical activity level, energy intake, body composition, etc.) with exemption of the variable(s) of interest. Controlling for extraneous variables preferentially with study design or at least statically decreases their potential confounding effect on target variables. If the species of interest are already known, selection of the appropriate study design, nucleotide extraction, and statistical analysis will be simplified.

It is also essential to achieve enough power to detect real changes in the variables of interest, which can be accomplished by determining the appropriate sample size. Underpowered studies can cause spurious associations that cannot be replicated by studies with enough power. This has been observed, for example, in studies attempting to demonstrate the ability of the microbiome to distinguish between obese and nonobese individuals [46]. Factors, such as the complexity of the community, the heterogeneity of the habitat over time and space, and the specificity of the results that need to be inferred, determine the correct selection of the sample size. Software to help estimate sample size is available with the following R packages [47–49].

Although logical, it is also important to highlight that, if conclusions are to be drawn for a specific microbial community, the sample must be representative of that community's particular habitat. Community representativeness can be achieved by knowing the amplitude of microbial variation; the more variation in an environment and the lower microbial biomass (e.g., total microbial load), the higher the sample size needed. Performing pilot studies (e.g., 16S rRNA sequencing) is useful in order to assess diversity and variability, which in turn aids in determining sample size [1].

Microbial communities readily share genetic information and have rapid communication systems; consequently, dramatic shifts in community structure can occur within hours to days [50]. Thus, timing can affect the interpretation of microbial analyses. This

recommendation is critical in studies involving individuals with cancer undergoing treatment as bowel preparation for surgery, prophylactic antibiotics, and chemo- and radiotherapy significantly affect the microbiome. To overcome this issue, it is crucial to sample at multiple times.

Lastly, location can also be a confounder; in gut microbiota analysis, for example, repeated sampling from different sites of the same stool or from different spatial locations in the colon or lung can lead to variation in microbiota composition [1, 44]. Homogenization of the stool sample can overcome this bias. Thus, it is crucial to control for or keep location/site constant within and between samples.

When evaluating different sampling methods, it is paramount to consider the reproducibility, stability, and accuracy of each of the methods to compare, especially if conducting studies at multiple geographical locations. Accuracy in this context means that the sampling method must conserve the microbial signature of the sample. Ideally, one would also want the method to be stable under suboptimal field conditions and to be useful for multiple different analyses (transcriptomics, proteomics, metabolomics, etc.) [51]. Several recent studies comparing microbial sampling methods have highlighted the importance of understanding the potential limitations of each method depending on the outcome of interest. Among the different methodological steps in microbiota analysis, nucleotide extraction has the most impact on results and will be further discussed later in this section [34, 35].

Also crucial is to avoid spurious or incorrect conclusions by carefully evaluating microbial sampling methodology in order to increase the researchers' ability to obtain valid and reliable data that can be compared against other studies [34, 51, 52]. For large population studies, three collection and storage methods appear most prudent: FOBT/FTA cards, 95% ethanol, and premade homogenization collection tubes. Comparison of seven fecal sampling methods, (1) no additive, (2) RNAlater, (3) 70% ethanol, (4) EDTA, (5) dry swab, (6) predevelopment FOBT (fecal occult blood test), and (7) post-development FOBT, showed that the use of swabs, FOBT cards (both pre and post), and 70% ethanol produced the most accurate microbial diversity measures. However, 70% ethanol resulted in low microbiome stability, rendering an unstable sample across time. Likewise, although RNAlater seemed to stabilize the microbiome over time, it caused considerable modifications to the microbiota diversity, causing a change in the microbial signature. Thus, according to this study, FOBT cards and swabs appear to be the best of these seven storing options [51].

FOBT consist of cards with fecal samples smeared on them, collected by using a stick and flushable tissue paper. The smeared sample in the card dries after a couple of minutes making it easy to transport and store at lower cost. Alternatively, rectal swabs consist

of sterile moistened swabs that are inserted 1–2 cm passing the anal sphincter (rotating the swab 360°). Swabs are then stored in their original container with liquid medium and stored at −80 °C. Rectal swabs can be collected at any time, require no patient preparation, and can be easily transported within the hospital; however, amount of sample obtained is usually a potential limitation for this method [52].

Commercial storage vials, such as OMNIgene Gut kit and ZYMO DNA/RNA Shield—fecal collection tubes—are now available. OMNIgene Gut kits have been used extensively in large population studies and have shown to effectively maintain microbiota composition by stabilizing nucleotides without the need of freezing. An advantage of this device is its stabilization buffer and the presence of a metal ball for complete homogenization. This method can potentially remove the inconvenient need to store stool samples in donors' freezers which is advantageous in several ways: it (1) eliminates the hazard of placing tubes with stool samples in house freezers, (2) reduces temperature fluctuations caused by the automatic defrost cycles that home freezers undergo, and (3) decreases the cost of transportation as there is no need for dry ice [44, 53].

A recent study evaluated the efficiency of OMNIgene Gut kit (DNA Genotek), RNAlater, FTA cards (Flinders Technology Associates cards from Whatman, similar product to FOBT cards), and 70% as well as 95% ethanol [54]. The authors demonstrated that 95% ethanol, FTA cards, and the OMNIgene Gut kit are able to efficiently preserve stool samples at ambient temperature for up to 8 weeks. The similarity between the preservation in the microbiota signature caused by 95% ethanol in comparison to OMNIgene Gut kit and FTA cards and the low cost and global accessibility of ethanol makes it a suitable choice for microbiome research. However, despite the fact that ethanol preserved DNA with high reliability, its use yielded lower concentrations of DNA in comparison to OMNIgene Gut kit and FTA cards. FTA cards and OMNIgene kit resulted in microbial community changes that are similar to those observed between technical replicates, thus suggesting that any changes over time produced by these two methods are within the limits of variation produced by 16S rRNA gene analysis. Caution against the use of 70% ethanol is given as it produces changes in the microbial community comparable in proportion to the ones observed between species. Additionally, although RNAlater was found to perform similarly to the other methods, it began losing stability after 2 weeks when maintained above freezing temperatures [54]. Nonetheless, it is important to highlight that all of these methods were compared based on the efficiency and stability analyzed through 16S rRNA gene sequencing; whether this interpretation applies for whole genome sequencing, RNA sequencing, or any other analysis is the focus of ongoing studies.

Nucleic acid extraction on fresh samples is the best way to ensure conservation of the microbial signature from the sample; however this is highly impractical and sometimes impossible in large population studies. Freezing the samples at −80 °C immediately after collection is still considered the reference method to compare new alternatives for sample collection. When freezing the samples directly at −80 °C is not possible, methods such as freezing the sample at −20 °C or storing them at 4 °C for up to 48–72 h have shown to prevent from significant alterations in the microbiota. However, if the study design does not allow immediate storage at −80 °C, −20 °C, or 4 °C due to geographical distance or other issues, the use of FOBT cards, OMNIgene kits, and rectal swabs is the next, best option [51, 52, 55, 56]. Several authors, however, suggest that even in cases where immediate freezing is available, other options such as OMNIgene kits, FTA cards, and 95% ethanol should be considered as they provide good stability and increased protection from freeze-thaw cycles, which normally occur during shipment and sample processing [54]. This recommendation is also supported by changes in the *Firmicutes/Bacteroidetes* ratio observed in frozen, unfixed stool samples in comparison to fresh samples. This change in bacterial composition under freezing temperatures seems to be caused by the greater ability of gram-positive bacteria (*Firmicutes*) to preserve DNA stability in comparison to gram-negative bacteria [57]. This error can be fully or partially addressed by (1) directly analyzing the fresh samples, (2) freezing all the samples the same amount of time, or (3) using another reliable storing method that can keep the microbial signature stable for a short period of time at ambient temperature.

1.5 Sample Preparation and Nucleotide Extraction

Cultivation-dependent methods have provided a valuable but incomplete picture of the human microbiome. Cultivation-independent approaches that involve the extraction of genomic DNA represent a more accurate depiction of the microbial community [58]. One of the key steps in analyzing microbial DNA is nucleotide extraction, as it has shown to have the second largest effect on metagenomic analysis after biological differences [35]. When choosing a DNA extraction protocol, it is important to select a method that is able to extract as much genetic material as possible without causing DNA damage [44]. Extraction of high-quality, high-molecular-weight DNA is necessary for techniques such as shotgun metagenomics; however, proper lysis of heterogeneous microbial communities without damaging their genomes is a major challenge. Human samples contain large numbers of microbial cells that belong to different phyla, heterogeneous in their morphology and cellular architecture. This heterogeneity narrows the potential options for ideal extraction methods; protocols cannot be too harsh that can damage DNA quality or too weak that

cause only partial cellular lysis [58]. Maximizing DNA concentration while minimizing its fragmentation is a key factor to consider when choosing an extraction protocol. Extraction of high-quality DNA is needed to prevent misrepresentation of community composition and to perform posterior analyses with more in-depth techniques such as shotgun sequencing [35]. Studies analyzing the efficiency of different DNA extraction protocols have mainly focused on measuring DNA yield as an outcome; however, research has shown a lack of correlation between DNA yield and accurate representation of the microbial diversity [58]. Thus, we cannot guarantee that methods that ensure high DNA yield will necessarily produce good representation of the microbial community [58]. Therefore, a key challenge in sequencing metagenomics is to isolate DNA that are both high quality and representative of the sampled community, which is the focus of this section.

It has been shown that separate locations within the same fecal sample can produce different microbial readings; thus, the first step in fecal sample analysis involves the homogenization of samples. Homogenization before storage reduces within-sample and intraindividual variability, which becomes crucial when performing longitudinal studies [59, 60]. Homogenization by either a blender or a pneumatic mixer has shown to increase the proportion of gram-positive bacteria, with the former method requiring only 2 min in comparison to 10–30 min of the latter one [60].

Analyzing human fecal samples is complex due to the presence of fibers, microbes, undigested particles, nucleases, and human cells; these factors can affect the overall quality and quantity of the metagenomic DNA obtained [61]. Removal of large size insoluble impurities takes place after homogenization, producing clean microbial pellets.

Once the fecal sample has been homogenized and cleaned from impurities, a lysis buffer and proteinases are then added to the sample [61]. Lysis of microbial cells exposes their genomic DNA to intra- and extracellular nucleases making DNA susceptible to exo- and endonucleases. Therefore, it is crucial to inactivate nucleases and other enzymes by adding strong denaturing agents [62]. A couple of thermal incubations, washes, and centrifugation cycles follow the procedure in order to obtain DNA with good integrity (>1.8 kb) and purity (A260/A280 >1.85; A260/A230 ~2.0–2.2) [63].

Methods that combine physical (thermal), mechanical (bead beating), and chemical lysis to isolate community bacterial DNA from human samples appear to be more efficient in obtaining good-quality DNA than methods utilizing only one of them [62]. Specifically, protocols that include bead beating and/or mutanolysin produce the closest representation of the bacterial community. Bead beating produces a mechanical disruption of the microbial cells that increases DNA recovery, specifically from gram-positive bacteria

which are difficult to lyse by enzymatic methods due to the strength of their cell walls, while mutanolysin creates cellular lysis through enzymatic reactions [35, 58, 60]. Protocols assessing the lysis efficiency of different lytic methods demonstrate that combining enzymes is superior than using individual enzymes to achieve more effective lyses of cell membranes. This might be due to the different structures of membrane peptidoglycans among bacterial species. C-type lysozyme, for example, cleaves the glycosidic bond between the C-1 of *N*-acetylmuramic acid and the C-4 of *N*-acetylglucosamine in bacterial peptidoglycans [62]. However, not all bacteria are sensitive to C-type lysozymes. Bacteria that have O-acetylated peptidoglycan (such as *S. aureus* and *Proteus mirabilis*) are sensitive to mutanolysin instead [64]. Lysostaphin is a glycylglycine endopeptidase that is able to cleave pentaglycine bridges in the cell wall of staphylococci. Therefore, using a cocktail of lytic enzymes can likely improve the cell lysis of different bacterial species, thus allowing for a better representation of bacterial communities [58, 65, 66].

Slight variations in DNA extraction protocols can impact DNA isolation and thus DNA sequencing. The European MetaHIT and the American Human Microbiome Project (HMP) are two major consortiums on microbiota studies. A recent study comparing the DNA extraction methods used by these two consortiums demonstrated that, although both DNA extraction methods resulted in sufficient yields and purity, a more efficient eukaryotic DNA extraction was achieved with the MetaHIT protocol; however, a more efficient bacterial DNA extraction was obtained with the HMP protocol. Both protocols involve comparable procedures for bead beating (same size); however, one of the main variations between the protocols is the different heat treatments (70 °C for 1–2 h in the MetaHIT method compared to 65 °C for 10 min followed by other 10 min at 95 °C in the HMP method). Interestingly, the study shows that although there was a high correlation between the two methods, samples that were extracted with the same DNA extraction method tended to cluster together in analysis, highlighting the impact of the method used on the detected composition of the sample. Last but not the least, when analyzing a large number of samples, it is important to consider the extra amount of labor involved in the protocol. In this case the MetaHIT protocol is more time-consuming than the HMP protocol [59].

After cellular lysis takes place, metagenomic DNA is then isolated to be later visualized by agarose gel and quantified by spectrophotometry. Metagenomic DNA refers to the DNA extracted from all microbes existent in a particular environment [1]. The majority of studies use one or multiple regions of the 16S rRNA to perform sequencing; however, this technique can be affected by the choice of primer used, PCR biases, and even the selected polymerase. Comparing methods by using shotgun

metagenomics sequencing as a readout can overcome these potential sources of bias [35].

All of the steps involved in sample preparation and nucleotide extraction are potential sources of bias, creating major challenges when trying to compare microbiota studies [44]. Adoption of a DNA extraction protocol that is universally considered as the "gold standard" is fundamental to improve the comparability of human studies analyzing the microbiome. Collaboration of 21 laboratories in 11 countries on 3 continents made possible the comparison of 21 commercial and noncommercial DNA extraction methods [35]. In the first phase of the study, 5 methods were selected out of the 21 considering the quantity and integrity of extracted DNA, recovered diversity, and observed ratio of gram-positive bacteria obtained through shotgun metagenomic sequencing. Gram-positive bacteria were used as an outcome criterion for comparison between protocols, as they are more likely to be affected by the extraction method. Their results demonstrate that protocol Q is the preferred option for DNA extraction considering its ease of use, higher extraction quality, accuracy, and transferability in comparison to the rest of the selected protocols [35]. Protocol Q is further explained in "Steps for DNA Extraction."

Important considerations when performing DNA extraction involve the inclusion of positive and negative controls to evaluate the overall performance of extraction protocols. Positive controls should include known artificial communities (known as mock communities) to measure the overall accuracy (error rates of sequencing and composition) of the protocol. Negative controls should be included throughout extraction (e.g., PCR clean water and Tris-HCl buffer) in order to evaluate potential contamination [34]. Contamination sources include molecular grade water, PCR reagents, and DNA extraction kit reagents. Furthermore, pathogen contamination has been identified in sequencing reagents. All PCR amplicons should therefore be run on an agarose gel prior to sequencing. The presence of contaminating DNA represents a particular challenge for researchers working with samples containing a low microbial biomass such as those from the blood or lungs. Contamination not only affects 16S rRNA gene sequencing but also affects whole genome sequencing (shotgun metagenomics), and it seems to be mainly caused by soil- or water-dwelling bacteria. Some recommendations to decrease the impact of contaminants are (1) maximizing the starting sample biomass as microbial loads lower than 10^3 to 10^4 may not provide robust results; (2) minimizing risk of contamination at the time of sample collection; (3) recording the batches of kits used for each particular sample; (4) processing samples at random order and in replicates or triplicates; and (5) checking for taxa that are biologically unexpected for the sample of hand [67].

DNA extraction is considered as a bottleneck of an otherwise high-throughput microbiota NGS as it is manual and time-consuming. Several semiautomated fecal DNA extraction protocols have been recently created promising to expedite the process, reduce variability, and eliminate potential human errors. More research, however, is needed to evaluate their validity and reliability in comparison to already existent manual protocols [68].

1.6 PCR Amplification

Correct selection of 16S rRNA variable region gene primers is critical for downstream analysis and identification of key taxa that are biologically important for each particular study population. Further information regarding primer selection is provided in the following section.

PCR amplification is an important stopping point to assess purity of samples. The presence of 16S rRNA amplicons in the negative controls suggests reagent or user contamination of the sample extraction procedure and should be used to optimize extraction procedures (for details *see* Salter et al. 2014).

After confirming amplification of the 16S rRNA gene in samples and absence of amplification in background and/or H_2O control samples, then variable region PCR amplification can proceed. The following 16S rRNA PCR protocol from the Langille Lab is recommended prior to sequencing [69] and contains step-by-step wet lab and bioinformatic protocols for 16S rRNA gene and metagenomic sequencing.

1.7 Next-Generation Sequencing

The two main approaches for inferring microbial community composition are 16S rRNA and whole genome shotgun sequencing. Choosing between each of these methods will largely depend on the experimental setting and cost; therefore, it is important to highlight the advantages and disadvantages of each method. Despite being cheaper and less time-consuming, performing 16S rRNA gene sequencing has several other advantages over shotgun metagenomics. For example, 16S rRNA gene sequencing is able to identify a higher number of bacteria in comparison to whole genome sequencing. This is because, to this day, genomic databases contain about an order of magnitude less sequences stored than 16S rRNA databases (as amplicon analyses were the first molecular tool to be applied to the human microbiota) [70, 71]. Analysis of the 16S rRNA gene gives a qualitative and semiquantitative description of the bacteria present in a complex biological system by amplifying only a section of the DNA used to identify target genes [71]. One critical drawback to using this method for quantification is the presence of multiple 16S rRNA copies in multiple species. Thus, while it might appear as one taxon that is relatively higher in abundance, it also may be an artifact of that taxa carrying multiple 16S rRNA gene copies [72–74]. This issue can be overcome in later analysis by conducting absolute quantitative PCR

(e.g., droplet digital PCR) for the specific species identified. 16S rRNA gene analysis clusters similar sequences at a particular level of identity and counts the number of representatives of each cluster. These clusters of similar sequences are referred to as operational taxonomic units (OTU), which are generated by binning organisms together that possess either 97% or 100% of similarity for identification at the genus level. Currently, it is not recommended to generate species-level assignments using individual variable regions of the 16S rRNA gene. Newer methods now exist, however, to infer species-level identification based on sequence variance analysis (SVA), which is described by Mallick et al. (2017) [10].

Another key consideration prior to amplification of the 16S rRNA gene is the choice of variable region. Prokaryotic 16S rRNA is a 1.5 kb long gene that contains conserved and variable regions [45]. There are nine "hypervariable regions" labeled V1 to V9. Hypervariable regions are flanked by highly conserved DNA sequences that are similar among most bacteria. These conserved regions serve as anchors for universal PCR primers. On the contrary, hypervariable regions have different gene sequence among bacterial taxa allowing their identification. The nucleotide spans of the nine hypervariable regions are 69–99 (V1), 137–242 (V2), 433–497 (V3), 576–682 (V4), 822–879 (V5), 986–1043 (V6), 1117–1173 (V7), 1243–1294 (V8), and 1435–1465 (V9).

Unfortunately, no single hypervariable region is capable of differentiating among all bacteria; thus, careful selection of target sequences is highly dependent on the microbes of interest. In bacterial cultures, V1 (30 bp in length), for example, has shown to best differentiate among *Staphylococcus* sp., as well as *Streptococcus* sp. Whereas, V2 and V3 were able to distinguish all bacterial species to the genus level except for those closely related to *Enterobacteriaceae*. Additionally, V2 (106 bp in length) distinguished better among *Mycobacterium*, *Staphylococcus*, *Streptococcus*, and *Clostridium* sp., while V3 (65 bp in length) proved better resolution among *Haemophilus* sp., *Enterobacteriaceae*, *K. pneumoniae*, and *E. aerogenes*. V6 (58 bp in length) was also able to distinguish among most bacterial species except those of *Enterobacteriaceae* and proved to be the best to distinguish *B. anthracis* and *B. cereus*. On the other hand, V7, V8, and especially V5 were less useful targets for genus- or species-specific probes due to their higher degree of sequence conservation in comparison with the rest of the hypervariable regions. Combining the hypervariable regions V2, V3, and V6 increases the resolution to the genus level for all the 110 bacterial species and down to the species level for most of them (but not all). While these three hypervariable regions have the highest nucleotide heterogeneity and discriminatory power from the nine, V4 has been used extensively in the majority of fecal-based studies and thus has the strongest ability to compare across studies [75]. Once 16S rRNA gene segments have been

amplified, their sequences are aligned with existing reference genome databases (NRC 2007). This technique allows the most taxonomic information in the smallest span of nucleotides at a more affordable price [75].

The use of melting curves to evaluate for primer dimers or unwanted amplifications is also recommended [66]. After sequencing is complete, the use of software such as FastQC (Babraham Bioinformatics) can be useful to quickly perform basic quality control checks on raw sequences of data coming from high-throughput sequencing pipelines.

Several common bioinformatics processing tools are available for taking 16S rRNA reads from sequencing run output to taxonomic identification. The two most commonly applied tools are MOTHUR (https://www.mothur.org) and QIIME (https://www.qiime.org), which are publically available, regularly updated, and well cited. It is beyond the scope of this chapter to enumerate the procedures in these complementary methods. Extremely detailed tutorials are available for each tool at their websites.

Shotgun metagenomic sequencing represents a more precise and informative alternative to 16S rRNA that bypasses gene-specific amplification, thus sequencing all DNA located in the sample including that from viruses, fungi, and archaea [70]. It also provides a more accurate representation of the bacterial composition and functionality in comparison to 16S rRNA. Techniques utilizing PCR amplification, such as 16S rRNA, can be easily contaminated and require many controls at all levels of the process [71], while metagenomic shotgun sequencing provides more information with reduced biases [76]. Unlike 16S rRNA, shotgun metagenomic sequencing provides higher resolution that allows the identification of specific species and strains of bacteria as well as mutations that allow the identification of sources of outbreaks and antibiotic resistance [11, 45, 70]. Several methods for 16S rRNA sequencing and shotgun metagenomics or metatranscriptomic sequencing are available, which have been most recently reviewed [10]. One final important consideration when designing large sample studies that require multiple sequencing runs, 16S rRNA or whole genome sequencing, is to consider the impact of batch effect. Therefore, careful setup of each plate including randomly distributed sample types (e.g., cases and controls), as well as duplicates and negative and positive controls (identical), will allow for batch effect adjustment across multiple plates or sequencing runs.

This chapter, while not exhaustive in covering all methods necessary for human microbiome research, provides the necessary background and methodological details in order to initiate independent research in microbial ecology for the purpose of determining associations with cancer biology. Future research methodology will likely advance most with regard to sequencing and will hopefully converge on standardized methods for collection, extraction,

sequencing, and bioinformatic/statistical analysis. In summary, this information goes toward ensuring that high standards for microbiome research are maintained in order to produce quality reproducible data for comparison across studies in future investigations for cancer research.

2 Materials

2.1 Stool Sample Collection

1. Spatula, stick, or spoon.
2. Stool collection tube.
3. Stool collection hat.
4. Pair of gloves.
5. Biohazard bag.
6. Easy-fold mailbox.
7. Plastic clinical mailing bag.

2.2 DNA Extraction

For fecal samples:

2.2.1 Homogenization with Lysis Buffer

1. 2 mL screw top.
2. Sterile zirconia beads.
3. ASL lysis buffer.
4. Vortex.

For sputum and other liquid samples:

1. 200 mM of monobasic sodium phosphate.
2. GES buffer (guanidine thiocyanate, EDTA disodium, pH 8.0, *N*-lauroylsarcosine sodium salt).
3. 2 mL screw top tube containing 0.2 g of 0.1 mm glass beads.
4. Homogenizer.

For skin samples:

1. Lysozyme solution.
2. Heat block.
3. Swabs.
4. Spin baskets and microtubes for preprocessing samples on swabs.

For tissue samples:

1. Lysozyme solution.
2. Heat block.
3. Proteinase K.

Enzymatic/mechanical lysis:
For fecal samples:

1. Incubator.
2. Homogenizer.
3. Centrifuge.
4. 2 mL tubes.
5. ASL lysis buffer.

For sputum and other liquid samples:

1. Lysozyme (100 mg/mL).
2. RNAse A (10 mg/mL of H_2O).
3. Vortex.
4. Incubator.
5. 25% SDS.
6. Proteinase K.
7. 62.5 μL of 5 M NaCl.

For skin samples:

1. Stainless steel beads, size 5.0 mm.
2. Homogenizer.
3. Shaking heat block.
4. Ice.

For tissue samples:

1. Tubes containing 2.0–2.8 mm ceramic beads.
2. Homogenizer.

2.2.2 DNA Extraction/ Precipitation

For fecal samples:

1. 10 M ammonium acetate.
2. Incubator.
3. Centrifuge.
4. 1.5 mL tubes.
5. Isopropanol.
6. Ice.
7. 70% EtOH.

For sputum and other liquid samples:

1. Screw cap tubes.
2. 25:24:1 phenol-chloroform-isoamyl alcohol.
3. 2 mL tubes.

4. Vortex.
5. Centrifuge.

For skin samples:

1. Protein precipitation reagent.
2. Vortex.
3. Centrifuge.
4. 2 mL tubes.
5. 100% EtOH.

For tissue samples:

1. 1.5 μL tube.
2. RNAse A (5 μg/μL).
3. Heat block with shaking.
4. Ice.
5. Centrifuge.
6. Protein precipitation reagent.
7. Vortex.
8. 2 mL tubes.
9. 100% EtOH.

2.2.3 DNA Purification

For fecal samples:

1. Tris-EDTA.
2. DNase-free RNase (10 mg/mL).
3. Incubator.
4. Proteinase K.
5. AL buffer.
6. Vortex.
7. Ethanol (96–100%).
8. Spin column.
9. Wash buffer.
10. Centrifuge.

For sputum and other liquid samples:

1. DNA binding buffer.
2. 1.5 mL tubes.
3. Vortex.
4. DNA column.
5. Centrifuge.

6. Wash buffer.
7. Sterile DNase/RNase-free dH_2O.

For skin samples:

1. gDNA column.
2. Centrifuge.
3. Wash buffer.
4. Safe-Lock PCR tubes.
5. PCR water.

For tissue samples:

1. gDNA column.
2. Centrifuge.
3. Wash buffer.
4. Safe-Lock PCR tubes.
5. PCR water.

2.2.4 DNA Quantification

1. 1% agarose gel or automated electrophoresis.
2. Fluorometer.

2.3 PCR Amplification

1. 75% ETOH.
2. 10% bleach.
3. Pre-PCR-labeled pipettes.
4. Filter tips.
5. Primers: 8F (5′-AGA GTT TGA TCC TGG CTC AG-3′) and 1391R (5′-GAC GGG CGG TGW GTR CA-3′).
6. PCR water.
7. 1.5 mL tubes.
8. Takara HS LA Taq (Taq DNA polymerase and DNA proofreading polymerase optimized for long-range PCR).
9. 10× buffer.
10. dNTP mix.
11. HS LA Taq (Taq polymerase plus a monoclonal antibody specific to Taq polymerase).
12. DMSO *optional for tissue.
13. Thermocycler.
14. 1% gel.

3 Methods

3.1 Steps for Stool Collection (see Note 1)

1. Urinate first.
2. Put gloves on.
3. Place the stool collection hat into the back of the toilet.
4. Have a bowel movement into the collection device.
5. Use the spatula, the stick, or the spoon to collect sample.
6. If using collection tubes, immediately screw back the cap without spilling the liquids.

   Some tubes recommend shaking the sample for a minimum of 30 s (or until all large particles have been dissolved).

   Discard the rest of the stool into the toilet, and place the collection hat and spatula in the garbage.
7. Follow instructions for storage of the tubes/cards. Commercial stool sample collection tubes are to be kept at environmental temperature (15–25 °C), while cards can be either maintained temporarily at environmental temperature or immediately frozen at −20 or −80 °C.
8. Place the sample inside a sealed bag, and put the bag inside an easy-fold mailer (optional).
9. Place easy-fold mailer inside a plastic clinical shipping bag.
10. Ship in accordance to biological specimen regulations.

*Stool collection protocol was adapted from the Human Microbiome Project,

Stool Collection and Shipping Instructions, and the OMNIgene User instructions.

3.2 Steps for DNA Extraction from Different Sources (see Notes 2–4)

1. Select a number of samples that is manageable (~12–24).
2. Remove samples from −80 °C and thaw them on ice.
3. Homogenization with lysis buffer (this will be different for each sample type).

For fecal samples:

1. Add 100–200 mg of feces to a 2 mL screw top tube containing 0.3 g of sterile zirconia beads.
2. Add 1.0 mL lysis buffer.
3. Vortex for 2 min.

3.2.1 For Sputum and Other Liquid Samples

1. Add 300 μL of sample, 800 μL of 200 mM of monobasic sodium phosphate, and 100 μL of GES buffer into a 2 mL screw top tube containing 0.2 g of 0.1 mm glass beads.
2. Bead beat with homogenizer for 3 min at 3000 rpm.

For skin samples:

1. Add lysozyme solution.
2. Incubate in shaking heat block at 37 °C for 1 h.
3. Swabs are placed in UV-treated basket and placed back in tube for a short spin to collect excess buffer. Swabs are discarded in MPW.

For tissue samples:

1. Add lysozyme solution.
2. Incubate in shaking heat block at 37 °C for 1 h.
3. Add 40 μL of proteinase K to each tube.

Enzymatic/mechanical lysis:

For fecal samples:

1. Incubate for 15 min at 95 °C.
2. Cells are mechanically lysed by running the homogenizer instrument 1 min, resting 5 min for eight times.
3. Samples are allowed to cool down on ice for 2 min.
4. Tubes are then centrifuged at 16,000 × *g* and 4 °C for 5 min.
5. Supernatant is transferred to a new 2 mL tube.
6. The pellet is mixed with 300 μL of ASL lysis buffer, and **steps 2–5** are repeated.
7. Supernatants are pooled in new 2 mL tubes.

For sputum and other liquid samples:

1. Add 50 μL of lysozyme and 10 μL of RNAse A.
2. Mix with vortex.
3. Incubate samples in a 37 °C water bath for 1–1.5 h.
4. Add 25 μL of 25% SDS, 25 μL of proteinase K, and 62.5 μL of 5 M NaCl.
5. Mix with vortex.
6. Incubate samples in a 65 °C water bath for 0.5–1.5 h.

For skin samples:

1. Add two sterile stainless steel beads to sample with bead dispenser.
2. Bead beat using homogenizer for 2 min at 3500 rpm.
3. Incubate tubes for 30 min at 65 °C in shaking heat block.
4. Ice for 5 min.

For tissue samples:

1. Transfer contents into tube containing 2.0–2.8 mm ceramic beads.
2. Bead beat using homogenizer for 2 min at 3500 rpm.

3.2.2 DNA Extraction/ Precipitation

For fecal samples:

1. Add 260 μL of 10 M ammonium acetate to each lysate tube.
2. Mix well.
3. Incubate on ice for 5 min.
4. Centrifuge tubes at 16,000 × *g* and 4 °C for 10 min.
5. Transfer the supernatant to two 1.5 mL tubes; add one volume of isopropanol (measure the supernatant transferred into each tube to add the same volume of isopropanol).
6. Mix well.
7. Incubate on ice for 30 min.
8. Centrifuge tubes at 16,000 × *g* and 4 °C for 15 min.
9. Remove supernatant using careful aspiration without touching the solid pellet.
10. Wash nucleic acid pellet with 70% EtOH (0.5 mL), and dry the pellet under vacuum for 3 min.

For sputum and other liquid samples:

1. Centrifuge screw cap tubes at 13,500 × *g* for 5 min.
2. Add 900 μL of phenol-chloroform-isoamyl alcohol to another set of 2 mL tubes.
3. Take 900 μL of supernatant from the screw cap tubes, and add it into the 2 mL tubes, thus creating a 1:1 volume solution of sample and phenol-chloroform.
4. Mix with vortex.
5. Centrifuge at <13,000 × *g* for 10 min.

For skin samples:

1. Add 250 μL of protein precipitation solution.
2. Vortex tube 10s.
3. Spin max 10 min.
4. Transfer to new 2 mL tubes and add equal volume 100% EtOH.

For tissue samples:

1. Remove liquid from tubes and add to new 1.5 μL tube.
2. Add 1 μL of RNAse A to each tube.

3. Incubate 30 min at 65 °C in heat block with shaking.
4. Ice 5 min.
5. Max spin tubes (18,078 × *g*) for 5 min and transfer supernatant to new tube.
6. Add 250 μL of protein precipitation solution.
7. Vortex tube10s.
8. Spin max 10 min.
9. Transfer to new 2 mL tubes and add equal volume COLD 100% EtOH (~600 μL).

3.2.3 DNA Purification

For fecal samples:

1. Dissolve the nucleic acid pellet in 100 μL of Tris-EDTA.
2. Add 2 μL of DNase-free RNase.
3. Incubate at 37 °C for 15 min.
4. Add 15 μL of proteinase K and 200 μL of AL buffer to the supernatant.
5. Mix with vortex.
6. Incubate at 70 °C for 10 min.
7. Add 200 μL of ethanol (96–100%) to the lysate.
8. Mix with vortex.
9. Transfer to a spin column and centrifuge at 16,000 × *g* for 1 min at room temperature.
10. Discard flow-through.
11. Add 500 μL wash buffer and centrifuge at 16,000 × *g* for 1 min at room temperature.
12. Discard flow-through.
13. Add 500 μL wash buffer and centrifuge at 16,000 × *g* for 1 min at room temperature.
14. Dry the column by centrifugation at room temperature for 1 min.
15. Add 200 μL buffer Tris-EDTA; incubate for 1 min at room temperature.
16. Centrifuge for 1 min at 16,000 × *g* to elute DNA.

For sputum and other liquid samples:

1. Add 200 μL of DNA binding buffer to 1.75 mL tubes.
2. Carefully transfer the top layer (avoiding taking up any of the interface) of the centrifuged sample to the DNA binding buffer in the 1.5 mL tubes.
3. Mix with vortex.

4. Transfer this solution to a DNA column 600 μL at a time.
5. Spin columns at <12,000 × *g*.
6. Discard flow-through.
7. Once sample has moved through the column, add 200 μL of wash buffer to the column.
8. Spin columns at <12,000 × *g*.
9. Discard flow-through.
10. Repeat wash step once more (**steps 7–9**).
11. Place the columns into a new, sterile 1.5 mL tubes.
12. Add 50 μL of sterile DNase/RNase-free ddH_2O preheated at 65 °C to the center of each column.
13. Incubate the columns at room temperature for 5 min.
14. Elute the DNA into the 1.75 mL tubes by centrifuging the columns in the tubes at a maximum speed of 12,000 × *g* for 1 min.

For skin samples:

1. Apply sample to gDNA column.
2. Spin at maximum speed for 1 min.
3. Remove flow-through, and apply additional sample; spin until all sample is used.
4. Place column in fresh collection tube.
5. Wash column with 500 μL Buffer 1.
6. Spin at maximum speed for 1 min.
7. Discard flow-through and place column in fresh collection tube.
8. Wash column with 500 μL Buffer 2.
9. Spin and discard flow-through spin at maximum speed for 3 min to dry.
10. Discard flow-through and place column in clean Safe-Lock PCR tubes.
11. Apply 35 μL of PCR water to column.
12. Let sit ~3 min.
13. Spin at maximum speed for 30 s to elute DNA.

For tissue samples:

1. Apply sample to gDNA column.
2. Spin at maximum speed for 1 min.
3. Remove flow-through, and apply additional sample; spin until all sample is used.

4. Place column in fresh collection tube.
5. Wash column with 500 μL Buffer 1.
6. Spin at maximum speed for 1 min.
7. Discard flow-through and place column in fresh collection tube.
8. Wash column with 500 μL Buffer 2.
9. Spin and discard flow-through.
10. Spin at maximum speed for 3 min to dry.
11. Discard flow-through and place column in clean Safe-Lock PCR tubes.
12. Apply 35 μL of PCR water to column.
13. Let sit ~3 min.
14. Spin at maximum speed for 30 s to elute DNA.

3.2.4 DNA Quantification (see Note 5)

1. For quality control use 1% agarose gel or automated electrophoresis.
2. For sample concentration use a fluorometer.
3. Store DNA at −20 °C for short term <1 month (>1 month store at −80).

Fecal, liquid, and skin/tissue DNA extraction protocols were adapted from.

Costea et al. (2017), Surette et al. (2014), Segre laboratory at the National.

Institutes of Health, and our laboratory at Baylor University [35, 77, 78].

3.3 Steps for 16S rRNA gene PCR Amplification Testing

1. Clean bench space with 75% ETOH and 10% bleach.
2. Use pre-PCR-labeled pipettes and filter tips.
3. Fecal sample = 16S rRNA ×25 cycles.
4. Background and H_2O controls are generally 28 cycles.
5. Use primers: 8F and 1391R. Stock solution is at 100 μM, and working solution is at 20 μM, all diluted in PCR water (individual 1 mL tubes).
6. For each 25 μL reaction: use Takara HS LA Taq (Taq DNA polymerase and DNA proofreading polymerase optimized for long-range PCR).
   (a) μL 10× buffer,
   (b) 4 μL dNTP mix
   (c) 0.5 μL each primer (20 μM, ordered from IDT)
   (d) 0.25 μL Takara HS LA Taq
   (e) 2 μL of DNA depending on the sampling method

(f) 1 μL DMSO

(g) To volume with PCR water.

7. Thermocycling: initial denaturation at 95 °C for 5 min, followed by 25–30 cycles of 30 s at 95 °C for denaturation, 30 s annealing at 55 °C, and 1.5 min elongation at 72 °C, all followed by final extension of 10 min at 72 °C.
8. Separate PCR products on a 1% gel.
9. Do not stack lanes. Only run one row of lanes at a time.
10. Run gel long enough for ample separation.

*Adapted from the laboratories of Julie Segre and Elizabeth Grice.

3.4 Statistical Considerations for Performing Metagenomic Studies

Characteristics of microbiome data. After bioinformatics processing, microbiome sequencing data are usually summarized into an abundance table, which records the counts of all detected features (taxa/gene) in each sample. While we focus the discussion on taxa abundance data, similar principles apply to functional abundance data. There are several important characteristics for microbiome data, which need to be taken into account in the statistical analysis [79]. First, the microbiome data are count data, where the variance depends on the mean, and thus are not appropriate to be analyzed directly using traditional statistical tools based on normal assumptions. Second, the microbiome data are highly dimensional, and the number of taxa usually exceeds the number of samples. Microbiome-wide hypothesis testing, which involves hundreds or even thousands of statistical tests, should be corrected for multiple testing to control for false positives [80]. Third, the microbiome data contain excessive zeros. Those zeros are due to either undersampling of rare taxa (sampling zeros) or physical absence of the taxa (structural zeros) [81]. The sampling zeros strongly depend on the sequencing depth, while the structural zeros do not. The different nature of zeros, as well as the dependence of sampling zeros on sequencing depth, requires the application of appropriate statistical treatment in both diversity- and taxa-level analysis, especially when the sequencing depth is a confounding variable that correlates with the phenotype of interest. Finally, the taxa are phylogenetically related. Closely related taxa usually share biological functions and tend to react to the environmental perturbation similarly. Thus, the phylogenetic tree is essential prior to knowledge, which can be used to increase the efficiency and power of statistical analysis [82, 83].

Before moving to more advanced analysis, exploratory data analysis for each variable (heat maps, ordination, stacked bar plots, etc.) should be performed to assess the sequencing quality, detect outliers, and determine sources of variability. The EMPeror [84] in QIIME is an excellent software to perform exploratory data

analysis. Potential confounders can also be identified at this stage, and they should be adjusted properly in both diversity and taxa-level analysis. Critically, for large studies it is often necessary to conduct multiple sequencing runs. Thus, as mentioned earlier, keeping track of the run or batch number is paramount when conducting exploratory analysis. It has been our experience that batch effects can significantly affect total number of reads per sample, which need to be controlled or adjusted for in later statistical analyses. As part of this exploratory analysis, it is absolutely critical to conduct contamination checking, even if you did not see any 16S rRNA amplicons during the PCR testing as sequencing reagents may introduce contaminants. At this point comparing positive and negative controls should be conducted to assess potential contaminants, especially so in low-biomass samples (e.g., lungs). This can be assessed by (1) checking for consistency across runs in distribution of positive controls or taxa of interest and (2), after major analyses or association testing is complete, identifying taxa of interest that may have a signal that appears artefactual (for a detailed explanation, *see* [67, 85]).

Diversity analysis. Diversity analysis, which summarizes the complex microbiome data into diversity measures, provides a holistic view of the microbiome. Both α-diversity and β-diversity analyses have been routinely performed for microbiome data. α-Diversity (within-sample diversity) reflects species richness and evenness within the microbial community, while β-diversity (between-sample diversity) reflects the shared diversity between microbial communities. A plethora of diversity measures have been proposed, including phylogenetic vs. non-phylogenetic measures and unweighted vs. weighted measures [86]. Phylogenetic measures take into account the phylogenetic relationship, while non-phylogenetic measures ignore such relationship. Phylogenetic measures are thus more powerful to reveal "clustered" phylogenetic signals. Non-phylogenetic measures, on the other hand, are more efficient to detect "scattered" signals. Unweighted measures are defined based on the presence and absence data, while weighted measures are based on the abundance data. Since the majority of species are rare or less abundant, unweighted measures implicitly give most weights to those rare and less abundant species [87]. In contrast, weighted measures give more weights to abundant species. Many variants of weighted measures have been proposed, mainly differing in their weighting schemes. Depending on the specific condition, each measure could be the most powerful to reveal the difference. Without much prior knowledge, it may be helpful to inspect the representative measure in each category to avoid missing important signals. For α-diversity, representative non-phylogenetic measures include species richness (e.g., chao1), Shannon index, and Simpson index with increasing weights on abundant species [88], and phylogenetic measures include the

unweighted phylogenetic diversity (PD) and abundance-weighted PD [89]. For β-diversity, representative non-phylogenetic measures include the Jaccard (unweighted) and Bray-Curtis (weighted) distance, and phylogenetic measures include the unweighted, generalized, and weighted UniFrac distance with increasing weights on abundant lineages [87].

For diversity analysis, the current practice still recommends rarefaction-based normalization, which subsamples the counts of different samples to even sequencing depth [90]. Although rarefaction discards a significant proportion of reads and is not optimal from an information perspective, it effectively reduces the variability due to differential sequencing depths. It is especially crucial for unweighted measures since the diversity of a sample strongly depends on the sequencing depth and rarefaction makes the diversity comparable across samples. Without rarefaction, samples tend to cluster by sequencing depths, and the biological difference may be masked. For confounded scenarios, where the variable of interest such as case and control status is correlated with the sequencing depth, rarefaction can efficiently reduce the confounding effects. In contrast, scaling-based normalization, where the counts are divided by a normalizing factor that reflects the sequencing depth, could not address the problem since presence/absence status does not change after normalization. For weighted measures, especially those that place most weights on abundant species such as Simpson index and weighted UniFrac, they are less sensitive to variable sequencing depths [90]. Nevertheless, rarefaction does not affect statistical power substantially for weighted measures unless the rarefaction depth is too low (e.g., <1000).

To test the association between α- or β-diversity and the phenotype of interest (POI), a regression approach is preferred due to its ability to adjust for covariates. Since α-diversity measures are usually approximately normally distributed, a multiple linear regression can be used to test the association between α-diversity and POI. PERMANOVA (permutational multivariate analysis of variance based on a distance matrix) can be used to test the association between β-diversity and POI [91]. However, PERMANOVA results may have inflated type I error in the presence of differential dispersion and sample size imbalance for two-group comparison. In such case, a multivariate Welch t-test on distances could be used [92]. To combine different β-diversity measures to improve statistical power, an omnibus test that takes multiple β-diversity measures can be performed [87, 93]. A closely related method, MiRKAT (microbiome regression-based kernel association test), which treats the β-diversity as a covariate, can also be used to test the association between β-diversity and POI [94].

Differential abundance analysis. α- and β-diversity analysis focuses on the community-level change and is instrumental in establishing the overall microbiome-phenotype association

through the aforementioned association tests. The next step is to perform differential abundance analysis (DAA), which aims to identify specific taxa associated with POI while controlling for potential confounders. Although multivariate methods, which jointly model the taxa and the phenotype, are emerging, univariate methods (one taxon at a time) are still the dominant approach for DAA. Over the past a few years, many DAA methods have been proposed including MetaStats [95], DESeq2 [96], edgeR [97], LEfSe [98], metagenomeSeq [99], ALDEx2 [100], RAIDA [101], ANCOM [102], and MicrobiomeDDA [103]. These methods differ in the way to address the three statistical challenges of microbiome data, namely, variable sequencing depth, excessive zeros, and compositionality. Microbiome data are inherently compositional—the counts only convey the relative abundance information [102, 104]. An increase in the abundance of one prevalent taxon will necessarily lead to the decrease in the relative abundances of the other taxa, which makes identifying truly differential taxa challenging. Although many evaluation studies have been done [90, 105, 106], no consensus has been reached about the optimal DAA method for microbiome sequencing data. Nevertheless, there are some recommendations for performing robust and powerful DAA:

1. Test the correlation between the sequencing depth and POI first. If they are significantly correlated, sequencing depth could confound the association between the abundance and POI for those less abundant taxa especially when nonparametric methods such as Wilcoxon rank sum test and Fisher's exact test are used [90]. In such case, rarefaction may be performed for nonparametric tests, or consider using count-based regression model (*see* **step 2**).
2. Use regression methods based on over-dispersed and/or zero-inflated count data such as DESeq2 and MicrobiomeDDA. These methods directly model the counts without the need for rarefaction, so they are usually statistically more powerful. The compositional effect can be addressed through a proper normalization such as CSS [99] and GMPR [103]. Confounder adjustment and testing interaction are very natural in the regression framework. The regression coefficients are biologically interpretable, and effect size can be readily defined. They are also amenable to meta-analysis, which combines effect sizes across studies [107]. However, there are some caveats for applying these count-based methods. Firstly, since they are based on parametric assumptions, they are vulnerable to outliers. Outlier replacement strategy such as winsorization should be considered to reduce the impact of outliers [103]. Secondly, these count-based models require a normalizing factor to address variable sequencing depth. Their native normalization

procedures may not be optimal for zero-inflated count data. Consider replacing their native normalization with CSS [99] and GMPR [103] normalization; both are developed for microbiome data. Finally, the *P*-values from these methods are usually based on a large sample theory and may not be accurate for small sample sizes or rare taxa with many zeros. Filtering out the less prevalent taxa before testing is recommended: generally, those that are represented by <5% abundance and/or are present in less than five samples in total (depending on sample size).

3. Test the method before using it. It is possible that the best method depends on the data characteristics [90]. It is thus advisable that the method is tested on the label-shuffled data to ensure proper false-positive control and on data with in silico spiked taxa to evaluate the statistical power [106]. Select methods with adequate false-positive control and reasonable power.

For additional information regarding data reproducibility in microbial ecology research, a series of tutorial videos have been published by the laboratory of Patrick Schloss and can be accessed using the following weblink: riffomonas.org.

4 Notes

1. If using the spatula (as the one provided by several commercial kits), fill out about 1/4 of the spatula with sample, and transfer the fecal sample into the yellow tube top until this one is full (without pushing the sample into the tube). If using the spoon provided, fill out one complete spoon. If sticks are used (as the ones provided by commercial cards), take a small amount of sample, and smear it into the indicated applicator. If stool samples are not well formed (e.g., diarrhea), then a spoon is recommended for collection.
2. Before starting DNA extraction, make sure to clean bench, pipettes, centrifuge, and racks with 75% EtOH and 10% bleach. Sterilize all pipettes, racks, forceps, and beads with UV light for 30 min. Also, make sure to obtain a negative control sample (50–100 μL or similar to your sample volume) of clean water that will be used for each set of samples to be extracted. This will identify any background contaminants in your extraction and sequencing procedures. Lastly, make sure to keep separate tips and pipettes for pre-PCR and post-PCR. Labeling each clearly is recommended prior to beginning extraction. Best practices include further separating the areas (e.g., different rooms) for pre-PCR and post-PCR or using an enclosed

cabinet/hood for DNA extraction to reduce introduction of contaminants.

3. Before performing enzymatic lysis, if specific species of interest are identified a priori, efforts should be made to choose the most effective enzyme for lysis.
4. When performing mechanical lysis, remember that different tissues require different types and size of beads. A set of samples should be tested first to identify the most prudent methods by quantifying total DNA and integrity using precision tools such as automated electrophoresis and fluorometers for accurate quantification.
5. DNA quality (A280/A260) needs to be >1.85 (A260/A230, 2.0–2.2), with less than 25% of the sample <1.8 kb fragments. DNA quantity needs to be >500 ng/μL. For more detailed DNA quality analysis methods and procedures, *see* Olson and Morrow [63].

References

1. NRC (2007) The new science of metagenomics—revealing the secrets of our microbial planet. The National Academies Press, Washington, DC
2. Schmidt TM (2006) The maturing of microbial ecology. Int Microbiol 9(3):217–223
3. Brock TD (1987) The study of microorganisms in situ: progress and problems. Symp Soc Gen Microbiol 41:1–17
4. Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59(1):143–169
5. Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A 74 (11):5088–5090
6. Pace NR, Stahl DA, Lane DJ, Olsen GJ (1986) The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences. In: Marshall K.C. (eds) Adv Microb Ecol vol 9. Springer, Boston, MA
7. Maron PA, Ranjard L, Mougel C, Lemanceau P (2007) Metaproteomics: a new approach for studying functional microbial ecology. Microb Ecol 53(3):486–493. https://doi.org/10.1007/s00248-006-9196-8
8. Behjati S, Tarpey PS (2013) What is next generation sequencing? Arch Dis Child Educ Pract Ed 98(6):236–238. https://doi.org/10.1136/archdischild-2013-304340
9. Heintz-Buschart A, May P, Laczny CC, Lebrun LA, Bellora C, Krishna A, Wampach L, Schneider JG, Hogan A, Beaufort C, Wilmes P (2016) Erratum: Integrated multi-omics of the human gut microbiome in a case study of familial type 1 diabetes. Nat Microbiol 2:16227. https://doi.org/10.1038/nmicrobiol.2016.227
10. Mallick H, Ma S, Franzosa EA, Vatanen T, Morgan XC, Huttenhower C (2017) Experimental design and quantitative analysis of microbial community multiomics. Genome Biol 18(1):228. https://doi.org/10.1186/s13059-017-1359-z
11. Snitkin ES, Zelazny AM, Thomas PJ, Stock F, Group NCSP, Henderson DK, Palmore TN, Segre JA (2012) Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. Sci Transl Med 4(148):148ra116. https://doi.org/10.1126/scitranslmed.3004129
12. Shibata T (2015) Current and future molecular profiling of cancer by next-generation sequencing. Jpn J Clin Oncol 45 (10):895–899. https://doi.org/10.1093/jjco/hyv122
13. Vogtmann E, Goedert JJ (2016) Epidemiologic studies of the human microbiome and cancer. Br J Cancer 114(3):237–242. https://doi.org/10.1038/bjc.2015.465
14. Drewes JL, White JR, Dejea CM, Fathi P, Iyadorai T, Vadivelu J, Roslani AC, Wick EC, Mongodin EF, Loke MF, Thulasi K, Gan HM, Goh KL, Chong HY, Kumar S, Wanyiri JW, Sears CL (2017) High-resolution

bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. NPJ Biofilms Microbiomes 3:34. https://doi.org/10.1038/s41522-017-0040-3

15. Bullman S, Pedamallu CS, Sicinska E, Clancy TE, Zhang X, Cai D, Neuberg D, Huang K, Guevara F, Nelson T, Chipashvili O, Hagan T, Walker M, Ramachandran A, Diosdado B, Serna G, Mulet N, Landolfi S, Ramon YCS, Fasani R, Aguirre AJ, Ng K, Elez E, Ogino S, Tabernero J, Fuchs CS, Hahn WC, Nuciforo P, Meyerson M (2017) Analysis of Fusobacterium persistence and antibiotic response in colorectal cancer. Science 358 (6369):1443–1448. https://doi.org/10.1126/science.aal5240
16. Zackular JP, Baxter NT, Chen GY, Schloss PD (2016) Manipulation of the gut microbiota reveals role in colon tumorigenesis. mSphere 1(1):e00001-15. https://doi.org/10.1128/mSphere.00001-15
17. Zackular JP, Baxter NT, Iverson KD, Sadler WD, Petrosino JF, Chen GY, Schloss PD (2013) The gut microbiome modulates colon tumorigenesis. MBio 4(6): e00692–e00613. https://doi.org/10.1128/mBio.00692-13
18. Peters BA, Wu J, Pei Z, Yang L, Purdue MP, Freedman ND, Jacobs EJ, Gapstur SM, Hayes RB, Ahn J (2017) Oral microbiome composition reflects prospective risk for esophageal cancers. Cancer Res 77(23):6777–6787. https://doi.org/10.1158/0008-5472.CAN-17-1296
19. Yoshimoto S, Loo TM, Atarashi K, Kanda H, Sato S, Oyadomari S, Iwakura Y, Oshima K, Morita H, Hattori M, Honda K, Ishikawa Y, Hara E, Ohtani N (2013) Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. Nature 499(7456):97–101. https://doi.org/10.1038/nature12347
20. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, Clancy TE, Chung DC, Lochhead P, Hold GL, El-Omar EM, Brenner D, Fuchs CS, Meyerson M, Garrett WS (2013) Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. Cell Host Microbe 14(2):207–215. https://doi.org/10.1016/j.chom.2013.07.007
21. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, Ojesina AI, Jung J, Bass AJ, Tabernero J, Baselga J, Liu C, Shivdasani RA, Ogino S, Birren BW, Huttenhower C, Garrett WS, Meyerson M (2012) Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. Genome Res 22(2):292–298. https://doi.org/10.1101/gr.126573.111
22. McDermott AJ, Huffnagle GB (2014) The microbiome and regulation of mucosal immunity. Immunology 142(1):24–31. https://doi.org/10.1111/imm.12231
23. Gopalakrishnan V, Spencer CN, Nezi L, Reuben A, Andrews MC, Karpinets TV, Prieto PA, Vicente D, Hoffman K, Wei SC, Cogdill AP, Zhao L, Hudgens CW, Hutchinson DS, Manzo T, Petaccia de Macedo M, Cotechini T, Kumar T, Chen WS, Reddy SM, Szczepaniak Sloane R, Galloway-Pena J, Jiang H, Chen PL, Shpall EJ, Rezvani K, Alousi AM, Chemaly RF, Shelburne S, Vence LM, Okhuysen PC, Jensen VB, Swennes AG, McAllister F, Marcelo Riquelme Sanchez E, Zhang Y, Le Chatelier E, Zitvogel L, Pons N, Austin-Breneman JL, Haydu LE, Burton EM, Gardner JM, Sirmans E, Hu J, Lazar AJ, Tsujikawa T, Diab A, Tawbi H, Glitza IC, Hwu WJ, Patel SP, Woodman SE, Amaria RN, Davies MA, Gershenwald JE, Hwu P, Lee JE, Zhang J, Coussens LM, Cooper ZA, Futreal PA, Daniel CR, Ajami NJ, Petrosino JF, Tetzlaff MT, Sharma P, Allison JP, Jenq RR, Wargo JA (2018) Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. Science 359 (6371):97–103. https://doi.org/10.1126/science.aan4236
24. Dejea CM, Fathi P, Craig JM, Boleij A, Taddese R, Geis AL, Wu X, DeStefano Shields CE, Hechenbleikner EM, Huso DL, Anders RA, Giardiello FM, Wick EC, Wang H, Wu S, Pardoll DM, Housseau F, Sears CL (2018) Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. Science 359(6375):592–597. https://doi.org/10.1126/science.aah3648
25. Gagniere J, Bonnin V, Jarrousse AS, Cardamone E, Agus A, Uhrhammer N, Sauvanet P, Dechelotte P, Barnich N, Bonnet R, Pezet D, Bonnet M (2017) Interactions between microsatellite instability and human gut colonization by Escherichia coli in colorectal cancer. Clin Sci (Lond) 131 (6):471–485. https://doi.org/10.1042/CS20160876
26. Wang X, Yang Y, Huycke MM (2015) Commensal bacteria drive endogenous transformation and tumour stem cell marker expression through a bystander effect. Gut 64 (3):459–468. https://doi.org/10.1136/gutjnl-2014-307213
27. Purcell RV, Pearson J, Aitchison A, Dixon L, Frizelle FA, Keenan JI (2017) Colonization

with enterotoxigenic Bacteroides fragilis is associated with early-stage colorectal neoplasia. PLoS One 12(2):e0171602. https://doi.org/10.1371/journal.pone.0171602

28. Raisch J, Buc E, Bonnet M, Sauvanet P, Vazeille E, de Vallee A, Dechelotte P, Darcha C, Pezet D, Bonnet R, Bringer MA, Darfeuille-Michaud A (2014) Colon cancer-associated B2 Escherichia coli colonize gut mucosa and promote cell proliferation. World J Gastroenterol 20(21):6560–6572. https://doi.org/10.3748/wjg.v20.i21.6560
29. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, Barnes R, Watson P, Allen-Vercoe E, Moore RA, Holt RA (2012) Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. Genome Res 22(2):299–306. https://doi.org/10.1101/gr.126516.111
30. Strauss J, Kaplan GG, Beck PL, Rioux K, Panaccione R, Devinney R, Lynch T, Allen-Vercoe E (2011) Invasive potential of gut mucosa-derived Fusobacterium nucleatum positively correlates with IBD status of the host. Inflamm Bowel Dis 17(9):1971–1978. https://doi.org/10.1002/ibd.21606
31. Baxter NT, Ruffin MT, Rogers MA, Schloss PD (2016) Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions. Genome Med 8 (1):37. https://doi.org/10.1186/s13073-016-0290-3
32. Iida N, Dzutsev A, Stewart CA, Smith L, Bouladoux N, Weingarten RA, Molina DA, Salcedo R, Back T, Cramer S, Dai RM, Kiu H, Cardone M, Naik S, Patri AK, Wang E, Marincola FM, Frank KM, Belkaid Y, Trinchieri G, Goldszmid RS (2013) Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. Science 342 (6161):967–970. https://doi.org/10.1126/science.1240527
33. Zitvogel L, Kroemer G (2012) Targeting PD-1/PD-L1 interactions for cancer immunotherapy. Oncoimmunology 1 (8):1223–1225. https://doi.org/10.4161/onci.21335
34. Sinha R, Abu-Ali G, Vogtmann E, Fodor AA, Ren B, Amir A, Schwager E, Crabtree J, Ma S Microbiome Quality Control Project C, Abnet CC, Knight R, White O, Huttenhower C(2017) Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. Nat Biotechnol 35 (11):1077–1086. https://doi.org/10.1038/nbt.3981
35. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, Tramontano M, Driessen M, Hercog R, Jung FE, Kultima JR, Hayward MR, Coelho LP, Allen-Vercoe E, Bertrand L, Blaut M, Brown JRM, Carton T, Cools-Portier S, Daigneault M, Derrien M, Druesne A, de Vos WM, Finlay BB, Flint HJ, Guarner F, Hattori M, Heilig H, Luna RA, van Hylckama Vlieg J, Junick J, Klymiuk I, Langella P, Le Chatelier E, Mai V, Manichanh C, Martin JC, Mery C, Morita H, O'Toole PW, Orvain C, Patil KR, Penders J, Persson S, Pons N, Popova M, Salonen A, Saulnier D, Scott KP, Singh B, Slezak K, Veiga P, Versalovic J, Zhao L, Zoetendal EG, Ehrlich SD, Dore J, Bork P (2017) Towards standards for human fecal sample processing in metagenomic studies. Nat Biotechnol 35(11):1069–1076. https://doi.org/10.1038/nbt.3960
36. Clarke SF, Murphy EF, O'Sullivan O, Lucey AJ, Humphreys M, Hogan A, Hayes P, O'Reilly M, Jeffery IB, Wood-Martin R, Kerins DM, Quigley E, Ross RP, O'Toole PW, Molloy MG, Falvey E, Shanahan F, Cotter PD (2014) Exercise and associated dietary extremes impact on gut microbial diversity. Gut 63(12):1913–1920. https://doi.org/10.1136/gutjnl-2013-306541
37. Biedermann L, Zeitz J, Mwinyi J, Sutter-Minder E, Rehman A, Ott SJ, Steurer-Stey C, Frei A, Frei P, Scharl M, Loessner MJ, Vavricka SR, Fried M, Schreiber S, Schuppler M, Rogler G (2013) Smoking cessation induces profound changes in the composition of the intestinal microbiota in humans. PLoS One 8(3):e59260. https://doi.org/10.1371/journal.pone.0059260
38. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA, Biddinger SB, Dutton RJ, Turnbaugh PJ (2014) Diet rapidly and reproducibly alters the human gut microbiome. Nature 505 (7484):559–563. https://doi.org/10.1038/nature12820
39. Flores R, Shi J, Fuhrman B, Xu X, Veenstra TD, Gail MH, Gajer P, Ravel J, Goedert JJ (2012) Fecal microbial determinants of fecal and systemic estrogens and estrogen metabolites: a cross-sectional study. J Transl Med 10:253. https://doi.org/10.1186/1479-5876-10-253
40. Salazar N, Arboleya S, Valdes L, Stanton C, Ross P, Ruiz L, Gueimonde M, de Los Reyes-Gavilan CG (2014) The human intestinal microbiome at extreme ages of life. Dietary intervention as a way to counteract

alterations. Front Genet 5:406. https://doi.org/10.3389/fgene.2014.00406

41. Macfarlane S (2014) Antibiotic treatments and microbes in the gut. Environ Microbiol 16(4):919–924. https://doi.org/10.1111/1462-2920.12399
42. Seto CT, Jeraldo P, Orenstein R, Chia N, DiBaise JK (2014) Prolonged use of a proton pump inhibitor reduces microbial diversity: implications for Clostridium difficile susceptibility. Microbiome 2:42. https://doi.org/10.1186/2049-2618-2-42
43. Lee H, Ko G (2014) Effect of metformin on metabolic improvement and gut microbiota. Appl Environ Microbiol 80(19):5935–5943. https://doi.org/10.1128/AEM.01357-14
44. Hill CJ, Brown JR, Lynch DB, Jeffery IB, Ryan CA, Ross RP, Stanton C, O'Toole PW (2016) Effect of room temperature transport vials on DNA quality and phylogenetic composition of faecal microbiota of elderly adults and infants. Microbiome 4(1):19. https://doi.org/10.1186/s40168-016-0164-3
45. Thomas V, Clark J, Dore J (2015) Fecal microbiota analysis: an overview of sample collection methods and sequencing strategies. Future Microbiol 10(9):1485–1504. https://doi.org/10.2217/fmb.15.87
46. Sze MA, Schloss PD (2016) Looking for a signal in the noise: revisiting obesity and the microbiome. MBio 7(4):e01018–e01016. https://doi.org/10.1128/mBio.01018-16
47. Kelly BJ, Gross R, Bittinger K, Sherrill-Mix S, Lewis JD, Collman RG, Bushman FD, Li H (2015) Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA. Bioinformatics 31(15):2461–2468. https://doi.org/10.1093/bioinformatics/btv183
48. La Rosa PS, Brooks JP, Deych E, Boone EL, Edwards DJ, Wang Q, Sodergren E, Weinstock G, Shannon WD (2012) Hypothesis testing and power calculations for taxonomic-based human microbiome data. PLoS One 7(12):e52078. https://doi.org/10.1371/journal.pone.0052078
49. Mattiello F, Verbist B, Faust K, Raes J, Shannon WD, Bijnens L, Thas O (2016) A web application for sample size and power calculation in case-control microbiome studies. Bioinformatics 32(13):2038–2040. https://doi.org/10.1093/bioinformatics/btw099
50. Schirmer M, Franzosa EA, Lloyd-Price J, McIver LJ, Schwager R, Poon TW, Ananthakrishnan AN, Andrews E, Barron G, Lake K, Prasad M, Sauk J, Stevens B, Wilson RG, Braun J, Denson LA, Kugathasan S, McGovern DPB, Vlamakis H, Xavier RJ, Huttenhower C (2018) Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. Nat Microbiol 3(3):337–346. https://doi.org/10.1038/s41564-017-0089-z
51. Sinha R, Chen J, Amir A, Vogtmann E, Shi J, Inman KS, Flores R, Sampson J, Knight R, Chia N (2016) Collecting fecal samples for microbiome analyses in epidemiology studies. Cancer Epidemiol Biomarkers Prev 25(2):407–416. https://doi.org/10.1158/1055-9965.EPI-15-0951
52. Bassis CM, Moore NM, Lolans K, Seekatz AM, Weinstein RA, Young VB, Hayden MK, Program CDCPE (2017) Comparison of stool versus rectal swab samples and storage conditions on bacterial community profiles. BMC Microbiol 17(1):78. https://doi.org/10.1186/s12866-017-0983-9
53. Abrahamson M, Hooker E, Ajami NJ, Petrosino JF, Orwoll ES (2017) Successful collection of stool samples for microbiome analyses from a large community-based population of elderly men. Contemp Clin Trials Commun 7:158–162. https://doi.org/10.1016/j.conctc.2017.07.002
54. Song SJ, Amir A, Metcalf JL, Amato KR, Xu ZZ, Humphrey G, Knight R (2016) Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. mSystems 1(3):e00021-16. https://doi.org/10.1128/mSystems.00021-16
55. Choo JM, Leong LE, Rogers GB (2015) Sample storage conditions significantly influence faecal microbiome profiles. Sci Rep 5:16350. https://doi.org/10.1038/srep16350
56. Vogtmann E, Chen J, Kibriya MG, Chen Y, Islam T, Eunes M, Ahmed A, Naher J, Rahman A, Amir A, Shi J, Abnet CC, Nelson H, Knight R, Chia N, Ahsan H, Sinha R (2017) Comparison of fecal collection methods for microbiota studies in Bangladesh. Appl Environ Microbiol 83(10):e00361-17. https://doi.org/10.1128/AEM.00361-17
57. Bahl MI, Bergstrom A, Licht TR (2012) Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis. FEMS Microbiol Lett 329(2):193–197. https://doi.org/10.1111/j.1574-6968.2012.02523.x
58. Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ (2012) Evaluation of methods for the extraction and purification of DNA from the human microbiome. PLoS One 7(3):e33865.

https://doi.org/10.1371/journal.pone.0033865

59. Wesolowska-Andersen A, Bahl MI, Carvalho V, Kristiansen K, Sicheritz-Ponten T, Gupta R, Licht TR (2014) Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. Microbiome 2:19. https://doi.org/10.1186/2049-2618-2-19
60. Hsieh YH, Peterson CM, Raggio A, Keenan MJ, Martin RJ, Ravussin E, Marco ML (2016) Impact of different fecal processing methods on assessments of bacterial diversity in the human intestine. Front Microbiol 7:1643. https://doi.org/10.3389/fmicb.2016.01643
61. Kumar J, Kumar M, Gupta S, Ahmed V, Bhambi M, Pandey R, Chauhan NS (2016) An improved methodology to overcome key issues in human fecal metagenomic DNA extraction. Genomics Proteomics Bioinformatics 14(6):371–378. https://doi.org/10.1016/j.gpb.2016.06.002
62. Bag S, Saha B, Mehta O, Anbumani D, Kumar N, Dayal M, Pant A, Kumar P, Saxena S, Allin KH, Hansen T, Arumugam M, Vestergaard H, Pedersen O, Pereira V, Abraham P, Tripathi R, Wadhwa N, Bhatnagar S, Prakash VG, Radha V, Anjana RM, Mohan V, Takeda K, Kurakawa T, Nair GB, Das B (2016) An improved method for high quality metagenomics DNA extraction from human and environmental samples. Sci Rep 6:26775. https://doi.org/10.1038/srep26775
63. Olson ND, Morrow JB (2012) DNA extract characterization process for microbial detection methods development and validation. BMC Res Notes 5:668. https://doi.org/10.1186/1756-0500-5-668
64. Shiba T, Harada S, Sugawara H, Naitow H, Kai Y, Satow Y (2000) Crystallization and preliminary X-ray analysis of a bacterial lysozyme produced by Streptomyces globisporus. Acta Crystallogr D Biol Crystallogr 56 (Pt 11):1462–1463
65. Gill C, van de Wijgert JH, Blow F, Darby AC (2016) Evaluation of lysis methods for the extraction of bacterial DNA for analysis of the vaginal microbiota. PLoS One 11(9): e0163148. https://doi.org/10.1371/journal.pone.0163148
66. Sohrabi M, Nair RG, Samaranayake LP, Zhang L, Zulfiker AH, Ahmetagic A, Good D, Wei MQ (2016) The yield and quality of cellular and bacterial DNA extracts from human oral rinse samples are variably affected by the cell lysis methodology. J Microbiol Methods 122:64–72. https://doi.org/10.1016/j.mimet.2016.01.013
67. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW (2014) Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 12:87. https://doi.org/10.1186/s12915-014-0087-z
68. Rintala A, Pietila S, Munukka E, Eerola E, Pursiheimo JP, Laiho A, Pekkala S, Huovinen P (2017) Gut microbiota analysis results are highly dependent on the 16S rRNA gene target region, whereas the impact of DNA extraction is minor. J Biomol Tech 28 (1):19–30. https://doi.org/10.7171/jbt.17-2801-003
69. Comeau A, Douglas G, Langille M (2017) Microbiome Helper: a Custom and Streamlined Workflow for Microbiome Research. https://github.com/LangilleLab/microbiome_helper/wiki/Microbiome-Amplicon-Sequencing-Workflow. Accessed 15 Mar 2018
70. Fuks G, Elgart M, Amir A, Zeisel A, Turnbaugh PJ, Soen Y, Shental N (2018) Combining 16S rRNA gene variable regions enables high-resolution microbial community profiling. Microbiome 6(1):17. https://doi.org/10.1186/s40168-017-0396-x
71. Cox MJ, Cookson WO, Moffatt MF (2013) Sequencing the human microbiome in health and disease. Hum Mol Genet 22(R1): R88–R94. https://doi.org/10.1093/hmg/ddt398
72. Vetrovsky T, Baldrian P (2013) The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. PLoS One 8(2):e57923. https://doi.org/10.1371/journal.pone.0057923
73. Taylor DL, Walters WA, Lennon NJ, Bochicchio J, Krohn A, Caporaso JG, Pennanen T (2016) Accurate estimation of fungal diversity and abundance through improved lineage-specific primers optimized for Illumina amplicon sequencing. Appl Environ Microbiol 82(24):7217–7226. https://doi.org/10.1128/AEM.02576-16
74. Louca S, Doebeli M, Parfrey LW (2018) Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. Microbiome 6(1):41. https://doi.org/10.1186/s40168-018-0420-9
75. Chakravorty S, Helb D, Burday M, Connell N, Alland D (2007) A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J Microbiol Methods 69(2):330–339. https://doi.org/10.1016/j.mimet.2007.02.005

76. Clooney AG, Fouhy F, Sleator RD, O'Driscoll A, Stanton C, Cotter PD, Claesson MJ (2016) Comparing apples and oranges?: next generation sequencing and its impact on microbiome analysis. PLoS One 11(2): e0148028. https://doi.org/10.1371/journal.pone.0148028
77. Whelan FJ (2014) Isolation of DNA from Clinical Samples (Genomic Prep). Surette Laboratory—Microbiome & Polymicrobial Research
78. Surette MG (2014) Isolation of DNA from clinical samples (GENOMIC PREP)
79. Li H (2015) Microbiome, metagenomics, and high-dimensional compositional data analysis. Annu Rev Statist Appl 2(1):73–94. https://doi.org/10.1146/annurev-statistics-010814-020351
80. Gilbert JA, Quinn RA, Debelius J, Xu ZZ, Morton J, Garg N, Jansson JK, Dorrestein PC, Knight R (2016) Microbiome-wide association studies link dynamic microbial consortia to disease. Nature 535(7610):94–103. https://doi.org/10.1038/nature18850
81. Kaul A, Mandal S, Davidov O, Peddada SD (2017) Analysis of microbiome data in the presence of excess zeros. Front Microbiol 8:2114. https://doi.org/10.3389/fmicb.2017.02114
82. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R (2011) UniFrac: an effective distance metric for microbial community comparison. ISME J 5(2):169–172. https://doi.org/10.1038/ismej.2010.133
83. Xiao J, Cao H, Chen J (2017) False discovery rate control incorporating phylogenetic tree increases detection power in microbiome-wide multiple testing. Bioinformatics 33 (18):2873–2881. https://doi.org/10.1093/bioinformatics/btx311
84. Vázquez-Baeza Y, Pirrung M, Gonzalez A, Knight R (2013) EMPeror: a tool for visualizing high-throughput microbial community data. GigaScience 2(1):1–4. https://doi.org/10.1186/2047-217X-2-16
85. Callahan BJ, DiGiulio DB, Goltsman DSA, Sun CL, Costello EK, Jeganathan P, Biggio JR, Wong RJ, Druzin ML, Shaw GM, Stevenson DK, Holmes SP, Relman DA (2017) Replication and refinement of a vaginal microbial signature of preterm birth in two racially distinct cohorts of US women. Proc Natl Acad Sci U S A 114(37):9966–9971. https://doi.org/10.1073/pnas.1705899114
86. Kuczynski J, Liu Z, Lozupone C, McDonald D, Fierer N, Knight R (2010) Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. Nat Methods 7 (10):813–819. https://doi.org/10.1038/nmeth.1499
87. Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, Wu GD, Collman RG, Bushman FD, Li H (2012) Associating microbiome composition with environmental covariates using generalized UniFrac distances. Bioinformatics 28(16):2106–2113. https://doi.org/10.1093/bioinformatics/bts342
88. Chao A, Gotelli NJ, Hsieh TC, Sander EL, Ma KH, Colwell RK, Ellison AM (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecol Monogr 84 (1):45–67. https://doi.org/10.1890/13-0133.1
89. McCoy CO, Matsen FA IV (2013) Abundance-weighted phylogenetic diversity measures distinguish microbial community states and are robust to sampling depth. PeerJ 1(6):e157. https://doi.org/10.7717/peerj.157
90. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, Hyde ER, Knight R (2017) Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5(1):59. https://doi.org/10.1186/s40168-017-0237-y
91. McArdle BH, Anderson MJ (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology 82(1):290–297. https://doi.org/10.2307/2680104?ref=no-x-route:eace2e52c544dc8ef8f1c463e5849bd9
92. Alekseyenko AV (2016) Multivariate Welch t-test on distances. Bioinformatics 32 (23):3552–3558. https://doi.org/10.1093/bioinformatics/btw524
93. Tang Z-Z, Chen G, Alekseyenko AV (2016) PERMANOVA-S: association test for microbial community composition that accommodates confounders and multiple distances. Bioinformatics 32(17):2618–2625. https://doi.org/10.1093/bioinformatics/btw311
94. Zhao N, Chen J, Carroll IM, Ringel-Kulka T, Epstein MP, Zhou H, Zhou JJ, Ringel Y, Li H, Wu MC (2015) Testing in microbiome-profiling studies with MiRKAT, the microbiome regression-based kernel association test. Am J Hum Genet 96 (5):797–807. https://doi.org/10.1016/j.ajhg.2015.04.003

95. White JR, Nagarajan N, Pop M (2009) Statistical methods for detecting differentially abundant features in clinical metagenomic samples. PLoS Comput Biol 5(4):e1000352. https://doi.org/10.1371/journal.pcbi.1000352
96. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550. https://doi.org/10.1186/s13059-014-0550-8
97. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26 (1):139–140. https://doi.org/10.1093/bioinformatics/btp616
98. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C (2011) Metagenomic biomarker discovery and explanation. Genome Biol 12(6):R60. https://doi.org/10.1186/gb-2011-12-6-r60
99. Paulson JN, Stine OC, Bravo HC, Pop M (2013) Differential abundance analysis for microbial marker-gene surveys. Nat Methods 10(12):1200–1202. https://doi.org/10.1038/nmeth.2658
100. Fernandes AD, Reid JN, Macklaim JM, McMurrough TA, Edgell DR, Gloor GB (2014) Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. Microbiome 2(1):15. https://doi.org/10.1186/2049-2618-2-15
101. Sohn MB, Du R, An L (2015) A robust approach for identifying differentially abundant features in metagenomic samples. Bioinformatics 31(14):2269–2275. https://doi.org/10.1093/bioinformatics/btv165
102. Mandal S, Van Treuren W, White RA, Eggesbø M, Knight R, Peddada SD (2015) Analysis of composition of microbiomes: a novel method for studying microbial composition. Microb Ecol Health Dis 26(0):27663. https://doi.org/10.3402/mehd.v26.27663
103. Chen J, King E, King E, Deek R, Deek R, Wei Z, Yu Y, Grill D, Grill D, Ballman K, Stegle O (2018) An omnibus test for differential distribution analysis of microbiome sequencing data. Bioinformatics 34 (4):643–651. https://doi.org/10.1093/bioinformatics/btx650
104. Tsilimigras MCB, Fodor AA (2016) Compositional data analysis of the microbiome: fundamentals, tools, and challenges. Ann Epidemiol 26(5):330–335. https://doi.org/10.1016/j.annepidem.2016.03.002
105. Jonsson V, Österlund T, Nerman O, Kristiansson E (2016) Statistical evaluation of methods for identification of differentially abundant genes in comparative metagenomics. BMC Genomics 17(1):1. https://doi.org/10.1186/s12864-016-2386-y
106. Thorsen J, Brejnrod A, Mortensen M, Rasmussen MA, Stokholm J, Al-Soud WA, Sørensen S, Bisgaard H, Waage J (2016) Large-scale benchmarking reveals false discoveries and count transformation sensitivity in 16S rRNA gene amplicon data analysis methods used in microbiome studies. Microbiome 4(1):62. https://doi.org/10.1186/s40168-016-0208-8
107. Waldron L (2018) Data and statistical methods to analyze the human microbiome. mSystems 3(2):e00194–e00117. https://doi.org/10.1128/mSystems.00194-17

Chapter 17

Assessing Metabolic Dysregulation in Muscle During Cachexia

Myriam Y. Hsu, Paolo E. Porporato, and Elisabeth Wyart

Abstract

Cancer cachexia is a metabolic disease characterized by a negative energy balance associated with systemic weight loss and poor quality of life.

In particular, skeletal muscle, which represents almost 50% of the total body mass, is strongly affected, and metabolic alterations therein (e.g., insulin resistance and mitochondrial dysfunction) can eventually support tumor growth by facilitating nutrient scavenging by the growing mass. Interestingly, metabolic interventions on wasting muscle have been proven to be protective, advocating for the importance of metabolic regulation in the wasting muscle.

Here, we will briefly define the current knowledge of metabolic regulation in cachexia and provide a protocol to grow and differentiate in vitro myotubes for the assessment of mitochondrial metabolism during cachexia.

Key words Cachexia, Muscle wasting, Myotube differentiation, Energy metabolism, Oxygen consumption

1 Introduction

Cancer can be interpreted as metabolic disease since biochemical alterations are instrumental to tumor onset and progression. Several hallmarks of cancer, such as unbridled proliferation and cell death resistance, are processes that are highly demanding in energy, requiring increased supply of ATP, and reducing equivalents and biosynthetic intermediates. Coherently, a critical bottleneck that cancer cells need to overcome is the increased energetic requirement in a hostile environment that is poor in nutrients and characterized by leaky, dysfunctional vessels with incipient hypoxia. In order to scavenge further resources, tumor co-opt cellular microenvironment to support growth via metabolic reprogramming of the tumor stroma. Along with the altered metabolism in stromal cells in the vicinity of the tumor, it is tempting to speculate that metabolic alterations might actually occur on a systemic level.

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_17,

Although little is known concerning the potential reprogramming of the whole host metabolism by the tumor, old and recent data support this particular view connecting systemic metabolic alterations to cancer growth, such as the following: (1) In an experimental model of early-stage KRAS-driven malignant transformation, a dysregulation of circadian rhythm is already present, leading eventually to altered insulin, glucose, and lipid metabolism [1]; (2) systemic alteration of insulin sensitivity is a well-documented fact in advanced tumor [2] and is actually the first alteration reported in cancer patients in 1919 [3]; (3) in drosophila, tumor is intrinsically capable of promoting insulin resistance through factor secretion, in order to increase nutrient availability by the tumor [4]; and (4) tumor growth promotes dysbiosis of the gut microbiome leading to general wasting syndrome [5].

The aforementioned alterations, along with chronic inflammation and metastatic dissemination, are often associated to a systemic syndrome affecting different organ functions known as *cachexia*.

Cachexia ("bad condition" in Greek) is a muscle-wasting syndrome, even in absence of reduced caloric intake. It occurs in several chronic and/or systemic diseases such as sepsis, AIDS, heart failure, and, notably, cancer. Particularly in cancer, it is a negative factor of prognosis, which affects strongly the quality of life of patients, decreases the tolerance to anticancer therapies, and is estimated to be the direct cause of death in at least one quarter of the patients afflicted with cancer [2]. Understanding the underpinning molecular mechanism of the metabolic alterations in cancer cachexia can thus potentially provide new therapeutic options alleviating cancer patients.

Cancer cachexia is a slow process during which different organs are affected, mostly due to chronic systemic inflammation that is physiological (natural response from the organism to fight against the disease) but also pathological (exacerbation by cancer cells that exploit the immune cells to secrete growth factors promoting tumor growth) [6]. Notable cytokines involved in cancer cachexia and multi-organ wasting include interleukins 1 and 6, tumor necrosis factor α (TNF-α formerly known as cachectin), and interferon gamma [7, 8].

In the brain, increased levels of these pro-inflammatory cytokines and disrupted hormonal balance lead to loss of appetite (anorexia), alteration of senses, and eventually increased energy consumption [9, 10]. A noteworthy example is ghrelin (one of the main orexigenic hormones) resistance; although its levels are abnormally high in cachectic patients, food intake remains decreased [11, 12].

In the digestive system, compelling evidence from cachectic mouse models has shown altered composition of gut microbiome [5, 13]. Surprisingly, a commensal strain of *E. coli* has been recently

reported to buffer intestinal inflammation triggered by infection and to prevent muscle wasting [14].

The liver, which is the metabolic hub and the body factory of essential proteins (e.g., blood clotting factors, angiotensinogen, albumin, lipoproteins, etc.), is also involved in the wasting process. Hepatomegaly (abnormal enlargement of liver) has long been associated with tumor progression and has been shown to correlate with an increased energy expenditure in cachectic patients [15]. Indeed, not only the liver is a target of wasting but it actually contributes to the process since the increased secretion of albumin and other acute phase proteins in response to cytokines secreted by local macrophages exacerbate systemic inflammation, thereby increasing muscle atrophy [16–18].

Another salient alteration in the liver of cancer patients is the accelerated gluconeogenesis using lactate produced by aerobic glycolysis of cancer cells (a metabolic pathway known as Cori cycle, [19–21]) that is further enhanced by steatosis (lipid accumulation) due to decreased secretion of very low-density lipoproteins. The latter being also a feature observed in both cachectic patients and mice [22–25].

In fact, the first studies showing altered carbohydrate metabolism in cancer trace back to the last century and have noted that although fasting blood glucose levels are roughly identical, cancer patients have significantly decreased blood glucose clearance [3, 26]. Correlation between insulin resistance and cachexia progression in cancer patients has later been reported [20], and furthermore, glucagon secretion by the pancreas has been found to be increased, thereby boosting liver gluconeogenesis [27–29]. Moreover, since insulin intrinsically acts as an anabolic factor, decreased secretion and sensitivity therefore result in an imbalance of protein synthesis versus protein breakdown that ultimately promotes atrophy [30].

Other life-threatening defects typical in patients with advanced cancer that have often been associated with cachexia include cardiac failure and arrhythmias [2, 31]. Indeed, cardiac muscle is an important target of cachexia, and it has been documented in mouse models that both cardiac function and weight are decreased in colon cancer [32]. Similarly to liver wasting, chronic cardiac failure has also been associated with an increased energy expenditure, while the high rate of oxidative metabolism has been evidenced in cachectic rat hearts [33, 34]. More recent studies have suggested that the pathways of cardiac muscle wasting are identical to those involved in skeletal muscle atrophy, e.g., activation of ubiquitin-proteasome degradation or NF-KB pathway [35, 36].

Albeit only few studies have investigated alterations in adipose tissue during cachexia, solid evidence has indicated that white adipose tissue also enters into the wasting process [37–40]. Paradoxically, cancer patients suffering from extreme weight loss display

raised serum levels of free fatty acids, glycerol, and triacylglycerol [39, 41], which can be explained by upregulated activities of several lipases in response to increased lipid mobilization promoting factors such as pro-inflammatory cytokines [8, 38, 42]. Interestingly, this wasting in adipose tissue has been documented in cachectic mice even before skeletal muscle atrophy manifests [38] and has more recently been associated with browning, hence mitochondrial enrichment characterizing the switch from white to brown adipose tissue [43, 44]. However, unlike physiological browning upon exposure to cold environment or in obesity, this metabolic shift in cancer is more deleterious than beneficial as it contributes to cachexia development by increasing energy expenditure.

Last, but most importantly, overt muscle atrophy is usually considered as the hallmark of cachexia and seen at the late stage of the disease, mostly prevalent in pancreatic and colon cancers. Muscle wasting causes fatigue and general weakness and can drastically decrease respiratory function, making the syndrome life threatening [45]. Skeletal muscle is the protein reservoir representing over 40% of the human body weight, whose breakdown into nitrogen and carbon might nurture cancer cells through mechanisms that have not been fully elucidated [46]. Unfortunately, no treatment to date has been able to reverse cancer-associated muscle wasting, although circulating markers of muscle proteolysis have been detected in cancer patients before the diagnosis of cancer [47].

Indeed, pro-inflammatory cytokines are promoters of muscle atrophy through modulation of signaling pathways involved in protein turnover (such as PI_3K-Akt-mTOR, NF-KB, JAK/STAT pathways) [48]. For instance, cytokines IL-1 and TNF-α were first found to induce muscle atrophy [49], but later investigations have revealed that targeting one single cytokine does not suffice to counteract atrophy [50]. It is now evident that diverse cytokines, hormonal changes, and complex signaling pathways are simultaneously involved in cachexia and multiple targets must be considered to eventually achieve prevention or reversal of the wasting syndrome [48]. Moreover, the underlying mechanisms causing cachexia can actually stem from interplay of distant organs, as evidenced by the study of Das and colleagues, showing that lipolysis inhibition in murine models of cachexia offsets muscle atrophy [38].

Several pathways of intracellular protein degradation in skeletal muscle have been identified to play a role in the cachectic process. A growing body of evidence has shown that ubiquitin-mediated proteasome degradation (UPR) is induced and promotes atrophy in murine models, while fewer studies have been done in human patients, but conflicting data are nevertheless present [48, 51]. Similarly, upregulated levels of autophagy mediators have been recently reported in cachectic patients [52]. Finally, although less

investigated, it has also been suggested that calcium-dependent cysteine proteases (calpains) also contribute to cachexia, but interference of this pathway failed to halt the cachectic process in mice [50, 53].

Interestingly, a role for epigenetic modulation of cachexia has recently been established by the study of Segatto and coworkers [54] who have identified the bromodomain and extra-terminal domain protein BRD4 as an epigenetic regulator of muscle mass, and its inhibition prevented both muscle and fat wasting and further prolonged survival in murine colon cancer models.

Undoubtedly, hypermetabolism is often found in cachectic cancer patients [55], and there has been growing interest in mitochondrial metabolism. As the soil of multiple vital cellular functions including respiration, apoptosis, reactive oxygen species production, etc., alterations in mitochondria have been reported in muscle atrophy not only in the case of cancer-associated cachexia but also in other neuromuscular diseases or even in physiological muscle mass loss due to aging, known as sarcopenia [56].

In fact, it has been reported that type II (fast-twitch) glycolytic fibers are the main target of cancer-induced atrophy while type I (slow-twitch) oxidative fibers are resistant to cachexia as seen in exercised skeletal muscle that underwent fiber-type switching from glycolytic to oxidative metabolism [57, 58]. Outside the context of cancer, this switch has been shown to protect skeletal muscle from atrophy in several studies by repressing myostatin activity [59] or UPR through activation of transcriptional factor peroxisome proliferator-activated receptor gamma coactivator 1α (PGC-1α), which is widely known as the master regulator of mitochondrial biogenesis [57].

A phenotypic shift from fast- to slow-twitch fibers has also been observed during cachexia; in other words, there is an increased mitochondrial mass in cachectic skeletal muscle [56, 58]. Yet paradoxically, fluxes of tricarboxylic acid cycle and ATP synthesis have been shown to be greatly reduced in skeletal muscle of lung tumor-bearing mice, suggesting therefore a mitochondrial uncoupling [60]. Those data have been corroborated by in vitro analyses of C2C12 myotubes treated with lung cancer conditioned medium showing impaired electron transport chain (ETC) activity along with a burst in mitochondrial reactive oxygen species (mtROS) [61]. Nevertheless, a study has documented no altered ATP production efficiency nor mitochondrial uncoupling but reduced complex IV activity in cachectic muscle of peritoneal tumor-bearing mice [62].

Furthermore, transcriptomic and metabolomic investigations on human muscle cell-based models and in vivo samples have unveiled that cancer cells secrete inflammatory mediators promoting excessive fatty acid oxidation (FAO) that drives to oxidative stress in skeletal muscle, while pharmacological inhibition of FAO

was shown to efficiently counteract cachexia, both in vitro and in vivo [63]. Similarly, genetic interference of triglyceride lipase has been shown to prevent myocyte apoptosis and adipose tissue and muscle atrophy in murine models of lung cancer [38]. Furthermore, the activation of ATP citrate lyase, a cytosolic enzyme that converts mitochondrial-derived citrate to oxaloacetate and acetyl-CoA, has been reported to ameliorate mitochondrial function in skeletal muscle, while knockdown of this enzyme reduced the activity of several electron transport chain complexes [64].

Quantification of ATP levels and energy metabolism typically requires assessment of cellular glycolytic and oxidative phosphorylation rates and can provide in-depth understanding of molecular pathogenesis. Investigating skeletal muscle metabolism in vitro can however be tricky since physiological metabolic changes that occur during myotube differentiation must be considered. It is therefore important to carry out assays at a specific time point during differentiation to obtain coherent results.

To this aim, we will briefly provide protocols of myotube differentiation induction using a mouse myoblast cell line (C2C12) and of mitochondrial metabolic profiling by Seahorse MitoStress assay.

Several protocols have been set to study mitochondrial dynamics over the last decades, but most require cell membrane permeabilization and/or extraction of mitochondria, and some require a large amount of samples to meet the sensitivity range of the assay. More recently, key parameters of mitochondrial oxidative metabolism can be accurately measured without these shortcomings; the Seahorse assay allows the quantification of oxygen consumption rate (OCR) over time with pO_2-sensitive fluorophores after successive injections of compounds that target specifically a complex of the electron transport chain (ETC), namely, oligomycin (ATP synthase inhibitor), FCCP (proton ionophore of the inner mitochondrial membrane, uncoupling the ETC and causing proton leak), antimycin A, and rotenone (ETC complexes III and I inhibitor, respectively).

2 Materials

1. Proliferation medium: Dulbecco's Modified Eagle Medium (DMEM), high glucose, 10% fetal bovine serum, and 1% penicillin-streptomycin.
2. Differentiation medium*:* Dulbecco's Modified Eagle Medium (DMEM), high glucose, 2% horse serum, and 1% penicillin-streptomycin.
3. 0, 25% Trypsin +0, 53 mM EDTA solution

4. Phosphate buffer saline (PBS 1×): 10 mM PO_4^{3-}, 137 mM NaCl, and 2.7 mM KCl.
5. Seahorse XF analyzer (Agilent Technologies).
6. Seahorse XF 96-well microplate (Agilent Technologies).
7. XF Base Medium (Agilent Technologies).
8. XF calibrant.
9. 100 mM pyruvate
10. 200 mM glutamine
11. 2.5 M glucose
12. 20–200 μL multichannel pipette
13. Seahorse MitoStress kit containing:
14. Sensor cartridge, utility plate, and lid.
15. Oligomycin.
16. Phenylhydrazone (FCC).
17. Rotenone/antimycin A.

3 Methods

3.1 Cell Growth

1. Rapidly thaw a vial of cells in a 37 °C pre-warmed bath.
2. Once cells are thawed, transfer cells to a 100 mm dish with 8 ml fresh complete medium (Dulbecco's Modified Eagle Medium (DMEM), high glucose, 10% fetal bovine serum, and 1% penicillin-streptomycin). Place in incubator at 37 °C in 5% CO_2.
3. After 1 day, discard the medium, and add fresh medium.
4. When cells reach 70% confluency, split to 1:8 (*see* **Note 1**) by removing the medium, and then rinse twice with PBS, and add 1 ml of 0, 25% Trypsin +0, 53 mM EDTA solution.
5. Place the dish in incubator for few minutes. When cells detach from the plate, add complete growth medium to inhibit trypsin; resuspend and transfer 1:8 of the cell suspension to an appropriate volume (7–8 mL for a 10 cm^2 dish) on a new dish.

3.2 Differentiation Treatment

1. When the cells are fully confluent, remove culture medium, rinse the cells twice with PBS, and add freshly prepared low serum differentiation medium (Dulbecco's Modified Eagle Medium (DMEM), high glucose, 2% horse serum, and 1% penicillin-streptomycin (optional)).
2. After 4–5 days in differentiation medium, the cells are fully differentiated and ready to be used for experiment.

3.3 Induction of Atrophic Process Using Conditioned Medium from Cancer Cells

Conditioned media from cancer cells should be prepared in advance as the following:

1. Allow the cancer cell line (*see* **Note 2**) to reach full confluency in a 10 mm dish.
2. Rinse two times with PBS, and add 8 ml of serum-free growth medium.
3. After 24 h, collect the medium, and centrifuge briefly at 5000 × *g* for 5 min. The resulting conditioned medium can either be immediately used or be stored at −20 °C.
4. Treat the differentiated myotubes with 10–30% of CM diluted in low serum differentiation medium (Dulbecco's Modified Eagle Medium (DMEM), high glucose, 2% horse serum, and 1% penicillin-streptomycin) for 24 h.

3.4 Assessment of Myotube Size

1. Take pictures at the brightfield microscope (magnification ×20) (Fig. 1).
2. Assess atrophy measuring diameters of muscular fibers using ImageJ software (*see* **Note 3**).

3.5 Evaluation of Mitochondrial Metabolism: General Overview

Oxygen consumption rate (OCR) is dynamically measured with a Seahorse allowing the assessment of mitochondrial metabolism in presence of various mitochondrial poison.

OCR is first measured at basal conditions prior to any treatment, and the measurement represents the oxygen consumed by cells for energy generation, accounting mitochondrial proton leak. Then, OCR decreases following the injection of oligomycin, with the difference in OCR between basal and after oligomycin treatment reflecting the amount of oxygen used to produce ATP by complex I. FCCP is thirdly used to induce OCR peak, which shows the maximal respiratory capacity cells can reach. The difference between the maximal OCR and the one used for ATP production

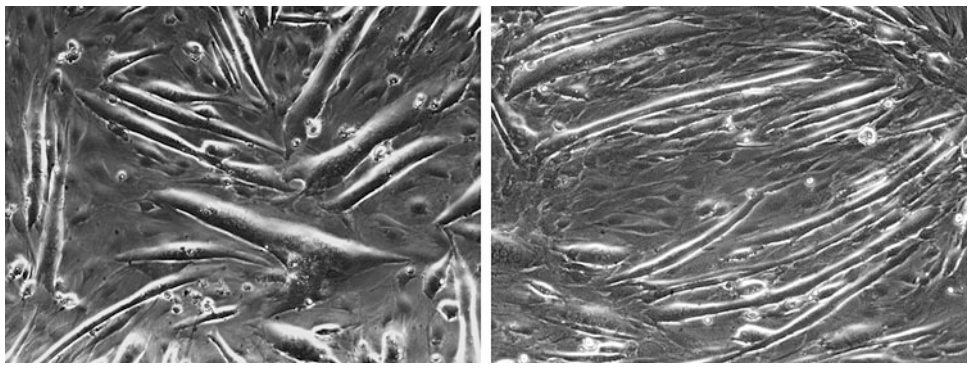


Fig. 1 C212 myotubes at 5 days of differentiation in control conditions (Left Panel) or after 24 h treatments with 10% Conditioned Medium from the pancreatic cancer cell line MiaPaca2 (Right Panel)

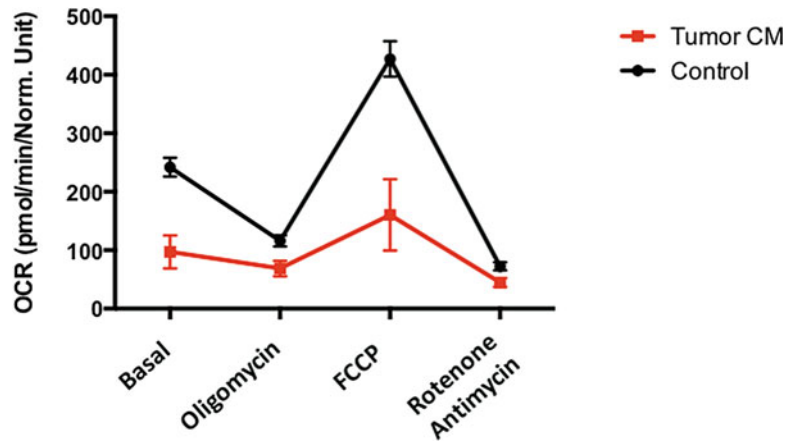


Fig. 2 Profiling of mitochondrial metabolism during cancer-induced myotube atrophy. Differentiated C2C12 myotubes have been treated with 5% conditioned media (CM) derived from colon cancer cells. Drug as indicated have been injected over time to assess mitochondrial contribution to Oxygen Consumption Rate (OCR)

(spare capacity) indicates how close to the maximal cells are using oxygen and reflecting the flexibility of energy production.

Last, antimycin A/rotenone is used to inhibit fully the electron transport chain, causing OCR to slump. The persisting OCR represents the oxygen consumed by enzymes located outside mitochondria (Fig. 2 depicting a difference in metabolism between normal myotubes and cancer cell conditioned media-treated ones). The difference between the non-mitochondrial respiration and the OCR used for ATP production gives the proton-leak OCR, i.e., the portion of basal oxygen consumption that is not coupled to ATP production. Increased proton leak can thus be a sign of mitochondrial damage. The resulted information provided by the assay will be exemplified herein.

3.6 Before the Assay (1–5 Days in Advance)

1. Seed 2000 C2C12 cells/well in 100 μL culture medium on a Seahorse XF 96-well microplate in the morning, incubate in a 5% CO_2 incubator at 37 °C, and treat with differentiation media after 6 h (*see* **Note 4**).

3.7 The Day Prior to the Assay

1. According to the experimental design, treat the myotubes with the specific factors and conditioned media.
2. Open the MitoStress kit that contains two plates: a cartridge plate containing oxygen and pH sensors (upper plate) and a utility plate for cartridge hydration (lower plate).
3. Fill the lower plate "utility plate" with 200 μL XF calibrant per well, and place the cartridge plate onto it to hydrate the sensor (*see* **Note 5**).

4. Incubate overnight (*see* **Note 6**) the assembled plates at 37 °C in an incubator without CO_2. The presence CO_2 will affect the pH of the calibrant resulting in incorrect pH measurements.
5. Switch on the XF analyzer; log into the XF 96 software to allow the instrument warm up overnight to 37 °C.

3.8 On the Day of Assay

1. Preparation of assay medium: Warm 50 ml of Seahorse XF assay medium to 37 °C, supplement the medium to have final concentrations of 1 mM pyruvate, 2 mM glutamine, and 10 mM glucose, and set the pH to 7.35 (*see* **Note 7**).
2. Preparation of 10× compounds stock solutions: Resuspend each compound of the kit with indicated volumes of prepared assay medium (Table 1) (*see* table below), pipetting up/down to homogenize (*see* **Note 8**).
3. Dilution of stock solutions: For each compound, dilute 300 μL of stock solution in 2.7 mL assay medium to obtain 3 mL of ready-to-inject solutions at 10 μM for oligomycin and FCCP and 5 μM for Rotenone/Antimycin-A. Mix well by flipping gently the tubes without vortexing.
4. Compound loading: Remove the previously hydrated cartridge from the incubator, and load each diluted compound solution into the corresponding port (*see* **Note 9**) as indicated in the below Table 2:

Table 1
Preparation of stock solutions

Compound	Assay medium to add (μL)	Final concentration 10× (μM)
Oligomycin	630	10
FCCP	720	10
Rotenone/antimycin A	540	5/5

Table 2
Cartridge loading

Compound	Volume to load (μL)	Port
Oligomycin	20	A
FCCP	22	B
Rotenone/antimycin A	25	C

3.8.1 One Hour Before Assay Starts

Use a multichannel pipette for the following filling-emptying procedure.

1. Remove the cell culture microplate from the incubator, and check briefly under the microscope for cell confluency that should be at least 80%.
2. Aspirate carefully the culture medium by placing the tip on the edge of the well. Leave a small volume (approximately 50 μL) to prevent cells from drying.
3. Gently wash the cells once with 100 μL/well of warmed assay medium by injecting the same way as previously described for cell seeding.
4. Load the wells with 100 μL/well of warmed assay medium, and place the microplate back to the incubator without CO_2 to incubate for at least 1 h.
5. Assay running: Switch on the Seahorse Analyzer, open Wave, and select template file by clicking on OPEN or design new template with the Assay Wizard. Click on RUN.
6. Place the calibration plate with the loaded cartridge on the instrument tray. Check the plate orientation, and click CONTINUE to calibrate the analyzer.
7. When the calibration is done (after 30 min), load the cell culture microplate taken out from the incubator without CO_2, and click START (*see* **Note 10**).
8. When the assay is completed, all the data will be automatically saved into a file with title as previously set. Remove the plates from the instrument, and trash the cartridge (*see* **Note 11**).
9. Proceed to quantification of cell numbers and/or protein content on the retrieved culture microplate for normalizing the data. Details explaining data analysis can be found on the instruction manual from the manufacturer.

4 Notes

1. These cells should never reach full confluency as they will start to fuse and differentiate upon cell-to-cell contact. It's a fast-growing cell line (doubling time is approximatively 12 h) that should be passed every 2–3 days.
2. Different cancer cell lines can be used to induce atrophy in vitro but not with the same power. A dose response should be effectuated to determine the optimal concentration of CM to use to induce atrophy.
3. Fiber diameters can be highly variable. To be consistent in the analysis, it's important to take pictures at different locations on

the plate and to measure a minimum of 100 myotubes per condition (take 10 pictures, and measure 10 myotubes in each picture, for example). Also, avoid the area of where nuclei are located, which is generally a bit thicker, and make sure to measure the uniform and elongated part of the fiber.

4. Leave some wells blank for background correction (at least four) by filling them with medium only. The upper part of the well is wider than the lower part, so apply the tip with an angle against the rim of the lowest part when injecting cell suspension, and leave the plate at room temperature for 10 min before placing it to the incubator to allow uniform monolayer adhesion of cells on the bottom of the wells. Eventually, start differentiation when cells are attached to the microplate (approximately 4 h), and allow differentiation into myotubes up to 4 days prior to assaying.
5. Make sure to not touch the sensors, which are extremely fragile.
6. It is possible to hydrate the sensors for up to 72 h. In this case, wrap the assembled plates with parafilm to prevent evaporation.
7. Since pH is temperature dependent, it is recommended to adjust the pH when the medium is warmed.
8. Do not refreeze the reconstituted solutions, and use within a day. Freeze-thawing can destroy the activity of the drugs.
9. When loading, place the tip vertically with no angle touching the bottom of the port; check the orientation of the plate; a notch should be found on the lower left corner. Make sure to load each compound in the right port since the instrument will inject in the A > B > C > D order. If only one or some of the parameters is/are to be measured, do not leave any port empty, but load the same volume of assay medium instead to have the correct final compound concentration and equal final volume in all the wells. Each port can contain up to 25 μL volume.
10. If needed, it is possible to change the protocol after starting the assay as long as the command has not yet been executed by clicking MODIFY, DELETE, or ADD. Results can be visualized while the assay is running by clicking RESULTS.
11. Check briefly under microscope if cells/myotubes have remained attached to the bottom of the wells on the microplate. Some treatments can loosen cell adhesion. To avoid poor cell adhesion on plates, coating wells with poly-*L*-lysine prior to cell seeding can eventually be considered.

References

1. Masri S, Papagiannakopoulos T, Kinouchi K, Liu Y, Cervantes M, Baldi P, Jacks T, Sassone-Corsi P (2016) Lung adenocarcinoma distally rewires hepatic circadian homeostasis. Cell 165 (4):896–909. https://doi.org/10.1016/j.cell.2016.04.039
2. Porporato PE (2016) Understanding cachexia as a cancer metabolism syndrome. Oncogene 5:e200. https://doi.org/10.1038/oncsis.2016.3
3. Tayek JA (1992) A review of cancer cachexia and abnormal glucose metabolism in humans with cancer. J Am Coll Nutr 11(4):445–456
4. Wagner EF, Petruzzelli M (2015) Cancer metabolism: a waste of insulin interference. Nature 521(7553):430–431. https://doi.org/10.1038/521430a
5. Bindels LB, Delzenne NM (2013) Muscle wasting: the gut microbiota as a new therapeutic target? Int J Biochem Cell Biol 45 (10):2186–2190. https://doi.org/10.1016/j.biocel.2013.06.021
6. Coussens LM, Werb Z (2002) Inflammation and cancer. Nature 420(6917):860–867. https://doi.org/10.1038/nature01322
7. Argiles JM, Lopez-Soriano FJ, Busquets S (2012) Counteracting inflammation: a promising therapy in cachexia. Crit Rev Oncog 17(3):253–262. https://doi.org/10.1615/CritRevOncog.v17.i3.30
8. Mantovani G, Maccio A, Mura L, Massa E, Mudu MC, Mulas C, Lusso MR, Madeddu C, Dessi A (2000) Serum levels of leptin and proinflammatory cytokines in patients with advanced-stage cancer at different sites. J Mol Med (Berl) 78(10):554–561
9. Argiles JM, Busquets S, Stemmler B, Lopez-Soriano FJ (2014) Cancer cachexia: understanding the molecular basis. Nat Rev Cancer 14(11):754–762. https://doi.org/10.1038/nrc3829
10. DeWys WD (1978) Changes in taste sensation and feeding behaviour in cancer patients: a review. J Hum Nutr 32(6):447–453
11. Otto B, Cuntz U, Fruehauf E, Wawarta R, Folwaczny C, Riepl RL, Heiman ML, Lehnert P, Fichter M, Tschop M (2001) Weight gain decreases elevated plasma ghrelin concentrations of patients with anorexia nervosa. Eur J Endocrinol 145(5):669–673
12. Garcia JM, Garcia-Touza M, Hijazi RA, Taffet G, Epner D, Mann D, Smith RG, Cunningham GR, Marcelli M (2005) Active ghrelin levels and active to total ghrelin ratio in cancer-induced cachexia. J Clin Endocrinol Metab 90 (5):2920–2926. https://doi.org/10.1210/jc.2004-1788
13. Bindels LB, Neyrinck AM, Claus SP, Le Roy CI, Grangette C, Pot B, Martinez I, Walter J, Cani PD, Delzenne NM (2016) Synbiotic approach restores intestinal homeostasis and prolongs survival in leukaemic mice with cachexia. ISME J 10(6):1456–1470. https://doi.org/10.1038/ismej.2015.209
14. Schieber AM, Lee YM, Chang MW, Leblanc M, Collins B, Downes M, Evans RM, Ayres JS (2015) Disease tolerance mediated by microbiome E. coli involves inflammasome and IGF-1 signaling. Science 350(6260):558–563. https://doi.org/10.1126/science.aac6468
15. Lieffers JR, Mourtzakis M, Hall KD, McCargar LJ, Prado CM, Baracos VE (2009) A viscerally driven cachexia syndrome in patients with advanced colorectal cancer: contributions of organ and tumor mass to whole-body energy demands. Am J Clin Nutr 89(4):1173–1179. https://doi.org/10.3945/ajcn.2008.27273
16. Bonetto A, Aydogdu T, Kunzevitzky N, Guttridge DC, Khuri S, Koniaris LG, Zimmers TA (2011) STAT3 activation in skeletal muscle links muscle wasting and the acute phase response in cancer cachexia. PLoS One 6(7): e22538. https://doi.org/10.1371/journal.pone.0022538
17. Patra SK, Arora S (2012) Integrative role of neuropeptides and cytokines in cancer anorexia-cachexia syndrome. Clin Chim Acta 413(13–14):1025–1034. https://doi.org/10.1016/j.cca.2011.12.008
18. Bologa RM, Levine DM, Parker TS, Cheigh JS, Serur D, Stenzel KH, Rubin AL (1998) Interleukin-6 predicts hypoalbuminemia, hypocholesterolemia, and mortality in hemodialysis patients. Am J Kidney Dis 32 (1):107–114
19. Holroyde CP, Skutches CL, Boden G, Reichard GA (1984) Glucose metabolism in cachectic patients with colorectal cancer. Cancer Res 44(12 Pt 1):5910–5913
20. Yoshikawa T, Noguchi Y, Doi C, Makino T, Okamoto T, Matsumoto A (1999) Insulin resistance was connected with the alterations of substrate utilization in patients with cancer. Cancer Lett 141(1–2):93–98
21. Tisdale MJ (2009) Mechanisms of cancer cachexia. Physiol Rev 89(2):381–410. https://doi.org/10.1152/physrev.00016.2008
22. Berriel DM, Krones-Herzig A, Metzger D, Ziegler A, Vegiopoulos A, Klingenspor M,

Muller-Decker K, Herzig S (2008) Nuclear receptor cofactor receptor interacting protein 140 controls hepatic triglyceride metabolism during wasting in mice. Hepatology 48 (3):782–791. https://doi.org/10.1002/hep.22383

23. Martignoni ME, Dimitriu C, Bachmann J, Krakowski-Rosen H, Ketterer K, Kinscherf R, Friess H (2009) Liver macrophages contribute to pancreatic cancer-related cachexia. Oncol Rep 21(2):363–369
24. Jones A, Friedrich K, Rohm M, Schafer M, Algire C, Kulozik P, Seibert O, Muller-Decker-K, Sijmonsma T, Strzoda D, Sticht C, Gretz N, Dallinga-Thie GM, Leuchs B, Kogl M, Stremmel W, Diaz MB, Herzig S (2013) TSC22D4 is a molecular output of hepatic wasting metabolism. EMBO Mol Med 5 (2):294–308. https://doi.org/10.1002/emmm.201201869
25. Sunny NE, Parks EJ, Browning JD, Burgess SC (2011) Excessive hepatic mitochondrial TCA cycle and gluconeogenesis in humans with nonalcoholic fatty liver disease. Cell Metab 14 (6):804–810. https://doi.org/10.1016/j.cmet.2011.11.004
26. Bishop JS, Marks PA (1959) Studies on carbohydrate metabolism in patients with neoplastic disease. II. Response to insulin administration. J Clin Invest 38(4):668–672. https://doi.org/10.1172/JCI103845
27. Inculet RI, Peacock JL, Gorschboth CM, Norton JA (1987) Gluconeogenesis in the tumor-influenced rat hepatocyte: importance of tumor burden, lactate, insulin, and glucagon. J Natl Cancer Inst 79(5):1039–1046
28. Bartlett DL, Charland SL, Torosian MH (1995) Reversal of tumor-associated hyperglucagonemia as treatment for cancer cachexia. Surgery 118(1):87–97
29. Hartl WH, Demmelmair H, Jauch KW, Koletzko B, Schildberg FW (1998) Effect of glucagon on protein synthesis in human rectal cancer in situ. Ann Surg 227(3):390–397
30. Gelfand RA, Barrett EJ (1987) Effect of physiologic hyperinsulinemia on skeletal muscle protein synthesis and breakdown in man. J Clin Invest 80(1):1–6. https://doi.org/10.1172/JCI113033
31. Kalantar-Zadeh K, Rhee C, Sim JJ, Stenvinkel P, Anker SD, Kovesdy CP (2013) Why cachexia kills: examining the causality of poor outcomes in wasting conditions. J Cachexia Sarcopenia Muscle 4(2):89–94. https://doi.org/10.1007/s13539-013-0111-0
32. Tian M, Nishijima Y, Asp ML, Stout MB, Reiser PJ, Belury MA (2010) Cardiac alterations in cancer-induced cachexia in mice. Int J Oncol 37(2):347–353
33. Drott C, Lundholm K (1990) Glucose uptake and amino acid metabolism in perfused hearts from tumor-bearing rats. J Surg Res 49 (1):62–68. https://doi.org/10.1016/0022-4804(90)90112-F
34. Mustafa I, Leverve X (2001) Metabolic and nutritional disorders in cardiac cachexia. Nutrition 17(9):756–760
35. Wysong A, Couch M, Shadfar S, Li L, Rodriguez JE, Asher S, Yin X, Gore M, Baldwin A, Patterson C, Willis MS (2011) NF-kappaB inhibition protects against tumor-induced cardiac atrophy in vivo. Am J Pathol 178 (3):1059–1068. https://doi.org/10.1016/j.ajpath.2010.12.009
36. Willis MS, Bevilacqua A, Pulinilkunnil T, Kienesberger P, Tannu M, Patterson C (2014) The role of ubiquitin ligases in cardiac disease. J Mol Cell Cardiol 71:43–53. https://doi.org/10.1016/j.yjmcc.2013.11.008
37. Costa G, Holland JF (1962) Effects of Krebs-2 carcinoma on the lipide metabolism of male Swiss mice. Cancer Res 22:1081–1083
38. Das SK, Eder S, Schauer S, Diwoky C, Temmel H, Guertl B, Gorkiewicz G, Tamilarasan KP, Kumari P, Trauner M, Zimmermann R, Vesely P, Haemmerle G, Zechner R, Hoefler G (2011) Adipose triglyceride lipase contributes to cancer-associated cachexia. Science 333(6039):233–238. https://doi.org/10.1126/science.1198973
39. Das SK, Hoefler G (2013) The role of triglyceride lipases in cancer associated cachexia. Trends Mol Med 19(5):292–301. https://doi.org/10.1016/j.molmed.2013.02.006
40. Wyart E, Reano S, Hsu MY, Longo DL, Li M et al (2018) Metabolic alterations in a slow-paced model of pancreatic cancer-induced wasting. Oxid Med Cell Longev 2018:6419805
41. Shaw JH, Wolfe RR (1987) Fatty acid and glycerol kinetics in septic patients and in patients with gastrointestinal cancer. The response to glucose infusion and parenteral feeding. Ann Surg 205(4):368–376
42. Argiles JM, Busquets S, Lopez-Soriano FJ (2011) Anti-inflammatory therapies in cancer cachexia. Eur J Pharmacol 668(Suppl 1): S81–S86. https://doi.org/10.1016/j.ejphar.2011.07.007
43. Kir S, White JP, Kleiner S, Kazak L, Cohen P, Baracos VE, Spiegelman BM (2014) Tumour-derived PTH-related protein triggers adipose tissue browning and cancer cachexia. Nature

513(7516):100–104. https://doi.org/10.1038/nature13528

44. Petruzzelli M, Schweiger M, Schreiber R, Campos-Olivas R, Tsoli M, Allen J, Swarbrick M, Rose-John S, Rincon M, Robertson G, Zechner R, Wagner EF (2014) A switch from white to brown fat increases energy expenditure in cancer-associated cachexia. Cell Metab 20(3):433–447. https://doi.org/10.1016/j.cmet.2014.06.011
45. Cohen S, Nathan JA, Goldberg AL (2015) Muscle wasting in disease: molecular mechanisms and promising therapies. Nat Rev Drug Discov 14(1):58–74. https://doi.org/10.1038/nrd4467
46. Baldanzi G, Pietronave S, Locarno D, Merlin S, Porporato P, Chianale F, Filigheddu N, Cantelmo AR, Albini A, Graziani A, Prat M (2011) Diacylglycerol kinases are essential for hepatocyte growth factor-dependent proliferation and motility of Kaposi's sarcoma cells. Cancer Sci 102(7):1329–1336. https://doi.org/10.1111/j.1349-7006.2011.01953.x
47. Mayers JR, Wu C, Clish CB, Kraft P, Torrence ME, Fiske BP, Yuan C, Bao Y, Townsend MK, Tworoger SS, Davidson SM, Papagiannakopoulos T, Yang A, Dayton TL, Ogino S, Stampfer MJ, Giovannucci EL, Qian ZR, Rubinson DA, Ma J, Sesso HD, Gaziano JM, Cochrane BB, Liu S, Wactawski-Wende J, Manson JE, Pollak MN, Kimmelman AC, Souza A, Pierce K, Wang TJ, Gerszten RE, Fuchs CS, Vander Heiden MG, Wolpin BM (2014) Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. Nat Med 20(10):1193–1198. https://doi.org/10.1038/nm.3686
48. Petruzzelli M, Wagner EF (2016) Mechanisms of metabolic dysfunction in cancer-associated cachexia. Genes Dev 30(5):489–501. https://doi.org/10.1101/gad.276733.115
49. Fong Y, Moldawer LL, Marano M, Wei H, Barber A, Manogue K, Tracey KJ, Kuo G, Fischman DA, Cerami A et al (1989) Cachectin/TNF or IL-1 alpha induces cachexia with redistribution of body proteins. Am J Physiol 256(3 Pt 2):R659–R665. https://doi.org/10.1152/ajpregu.1989.256.3.R659
50. Costelli P, Reffo P, Penna F, Autelli R, Bonelli G, Baccino FM (2005) Ca(2+)-dependent proteolysis in muscle wasting. Int J Biochem Cell Biol 37(10):2134–2146. https://doi.org/10.1016/j.biocel.2005.03.010
51. Baracos VE, DeVivo C, Hoyle DH, Goldberg AL (1995) Activation of the ATP-ubiquitin-proteasome pathway in skeletal muscle of cachectic rats bearing a hepatoma. Am J Physiol 268(5 Pt 1):E996–E1006
52. Aversa Z, Pin F, Lucia S, Penna F, Verzaro R, Fazi M, Colasante G, Tirone A, Rossi Fanelli F, Ramaccini C, Costelli P, Muscaritoli M (2016) Autophagy is induced in the skeletal muscle of cachectic cancer patients. Sci Rep 6:30340. https://doi.org/10.1038/srep30340
53. Pin F, Minero VG, Penna F, Muscaritoli M, De Tullio R, Baccino FM, Costelli P (2017) Interference with Ca(2+)-dependent proteolysis does not alter the course of muscle wasting in experimental cancer cachexia. Front Physiol 8:213. https://doi.org/10.3389/fphys.2017.00213
54. Segatto M, Fittipaldi R, Pin F, Sartori R, Dae Ko K, Zare H, Fenizia C, Zanchettin G, Pierobon ES, Hatakeyama S, Sperti C, Merigliano S, Sandri M, Filippakopoulos P, Costelli P, Sartorelli V, Caretti G (2017) Epigenetic targeting of bromodomain protein BRD4 counteracts cancer cachexia and prolongs survival. Nat Commun 8(1):1707. https://doi.org/10.1038/s41467-017-01645-7
55. Fearon KC, Glass DJ, Guttridge DC (2012) Cancer cachexia: mediators, signaling, and metabolic pathways. Cell Metab 16(2):153–166. https://doi.org/10.1016/j.cmet.2012.06.011
56. Dowling P, Murphy S, Ohlendieck K (2016) Proteomic profiling of muscle fibre type shifting in neuromuscular diseases. Expert Rev Proteomics 13(8):783–799. https://doi.org/10.1080/14789450.2016.1209416
57. Sandri M, Lin J, Handschin C, Yang W, Arany ZP, Lecker SH, Goldberg AL, Spiegelman BM (2006) PGC-1alpha protects skeletal muscle from atrophy by suppressing FoxO3 action and atrophy-specific gene transcription. Proc Natl Acad Sci U S A 103(44):16260–16265. https://doi.org/10.1073/pnas.0607795103
58. Ciciliot S, Rossi AC, Dyar KA, Blaauw B, Schiaffino S (2013) Muscle type and fiber type specificity in muscle wasting. Int J Biochem Cell Biol 45(10):2191–2199. https://doi.org/10.1016/j.biocel.2013.05.016
59. Ruas JL, White JP, Rao RR, Kleiner S, Brannan KT, Harrison BC, Greene NP, Wu J, Estall JL, Irving BA, Lanza IR, Rasbach KA, Okutsu M, Nair KS, Yan Z, Leinwand LA, Spiegelman BM (2012) A PGC-1alpha isoform induced by resistance training regulates skeletal muscle hypertrophy. Cell 151(6):1319–1331. https://doi.org/10.1016/j.cell.2012.10.050
60. Tzika AA, Fontes-Oliveira CC, Shestov AA, Constantinou C, Psychogios N, Righi V, Mintzopoulos D, Busquets S, Lopez-Soriano

FJ, Milot S, Lepine F, Mindrinos MN, Rahme LG, Argiles JM (2013) Skeletal muscle mitochondrial uncoupling in a murine cancer cachexia model. Int J Oncol 43(3):886–894. https://doi.org/10.3892/ijo.2013.1998

61. McLean JB, Moylan JS, Andrade FH (2014) Mitochondria dysfunction in lung cancer-induced muscle wasting in C2C12 myotubes. Front Physiol 5:503. https://doi.org/10.3389/fphys.2014.00503
62. Julienne CM, Dumas JF, Goupille C, Pinault M, Berri C, Collin A, Tesseraud S, Couet C, Servais S (2012) Cancer cachexia is associated with a decrease in skeletal muscle mitochondrial oxidative capacities without alteration of ATP production efficiency. J Cachexia Sarcopenia Muscle 3(4):265–275. https://doi.org/10.1007/s13539-012-0071-9
63. Fukawa T, Yan-Jiang BC, Min-Wen JC, Jun-Hao ET, Huang D, Qian CN, Ong P, Li Z, Chen S, Mak SY, Lim WJ, Kanayama HO, Mohan RE, Wang RR, Lai JH, Chua C, Ong HS, Tan KK, Ho YS, Tan IB, Teh BT, Shyh-Chang N (2016) Excessive fatty acid oxidation induces muscle atrophy in cancer cachexia. Nat Med 22(6):666–671. https://doi.org/10.1038/nm.4093
64. Das S, Morvan F, Jourde B, Meier V, Kahle P, Brebbia P, Toussaint G, Glass DJ, Fornaro M (2015) ATP citrate lyase improves mitochondrial function in skeletal muscle. Cell Metab 21(6):868–876. https://doi.org/10.1016/j.cmet.2015.05.006

Chapter 18

Using Seahorse Machine to Measure OCR and ECAR in Cancer Cells

Jing Zhang and Qing Zhang

Abstract

A large amount of energy used for nutrient processing and cellular functions is essential for tumorigenesis. Total intracellular adenosine triphosphate (ATP) is mainly generated by glycolysis and mitochondrial oxidative phosphorylation. Here, we provide a protocol for measurements of energy metabolism in cancer cells by using Seahorse XF24 Extracellular Flux analyzer. Specifically, this machine measures glycolysis by analyzing the extracellular acidification rate (ECAR) and measures mitochondrial oxidative phosphorylation on the basis of the oxygen consumption rate (OCR), through real-time and live cell analysis. This protocol is provided for researchers who are unfamiliar with the method and to aid them in carrying out the technique successfully.

Key words Energy metabolism, Glycolysis, Oxidative phosphorylation, Seahorse XF24 Extracellular Flux analyzer, OCR, ECAR

1 Introduction

Altered metabolism, upon which cancer cells depend heavily to support their uncontrolled cell proliferation, is a hallmark of cancer [1, 2]. Cancer metabolic programs include reprogramming of glycolysis, glutaminolysis, oxidative phosphorylation, fatty acid metabolism, one-carbon metabolism, etc., all of which provide essential energy, biosynthesis, and intermediates for tumor cell growth, division, and redox homeostasis [3]. Energy metabolic pathways mostly involve glycolysis and mitochondrial oxidative phosphorylation that are used to support all other cellular functions. Through glycolysis cycle, glucose is sequentially metabolized to pyruvate and lactic acid, with the concomitant production of small amount of adenosine triphosphate (ATP). Through mitochondrial oxidative phosphorylation, pyruvate is transported into the mitochondria, where it is oxidatively decarboxylated to acetyl-CoA. Maximal ATP is then generated through the citric acid cycle and oxidative phosphorylation, while oxygen is reduced to water

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_18,

during this process. It is worth noting that oxidative phosphorylation consumes more than 90% of oxygen in the body, the remaining 10% of which is used for non-mitochondrial respiration, including substrate oxidation and cell surface oxygen consumption [4].

Otto Warburg proposed that cancer cells rely largely on glycolysis to produce ATP even in the presence of oxygen [5]. However, increasing experimental evidence suggests that mitochondrial oxidative phosphorylation still plays an important role in cancer cells' metabolism [6–15]. Various protocols using the Seahorse instrument have been developed for analyzing glycolysis and oxidative phosphorylation in different cell types [16–18]. Here, we use renal cancer cell line 786-O as an example to describe detailed practical procedures of quantifying ECAR and OCR with the Seahorse XF24 Extracellular Flux analyzer to measure glycolysis and mitochondrial oxidative phosphorylation, respectively. The optimal seeding density was determined to be 50,000 cells/well for 786-O cells for use with this specific cell line in this protocol. The optimal concentrations of the injection compounds were 4.5 g/L glucose; 1 μM oligomycin and 50 mM 2-DG for ECAR; and 1 μM oligomycin, 1 μM FCCP, and 1 μM rotenone for OCR. For other cell types, concentrations of the injection compounds may have to be optimized.

2 Materials

2.1 Equipment

1. Seahorse XF24 Extracellular Flux analyzer.
2. Non-CO_2 isotemp incubator.
3. Seahorse XF24 FluxPak.
4. Automatic cell counter.
5. pH meter.
6. Water bath.

2.2 Cell Lines of Interest and Reagents

1. Renal cancer cell line 786-O.
2. Growth media: DMEM, 10% fetal bovine serum, 1% penicillin-streptomycin.
3. PBS.
4. Trypsin-EDTA (0.05%).
5. XF calibrant solution.
6. XF assay medium.
7. Glucose.
8. Glutamine.
9. Sodium pyruvate.
10. Oligomycin A.

11. FCCP.
12. Rotenone.
13. 2-DG.

3 Methods

3.1 The XF24 Assay: Day 1

3.1.1 Seeding Cells in XF24 Cell Culture Plate (Blue Box)

1. The XF24 cell culture plate is designated with rows A–D and columns 1–6 (*see* **Note 1**). Plate 50,000,786-O cells in 100 μL growth media per well (*see* **Note 2**), leaving background temperature correction wells (A1, B4, C3, and D6) blank with 100 μL growth media.
2. Incubate plate at 37 °C, 5% CO_2 for 2 h, check cells to ensure that they have adhered to plate reaching almost 90–100% confluence, and carefully add 150 μL growth media to each well (*see* **Notes 3** and **4**).
3. Incubate cells overnight at 37 °C, 5% CO_2 (*see* **Notes 5** and **6**).

3.1.2 Preparing Sensor Cartridge (Green Box)

1. Add 1 mL of Seahorse XF24 calibrant solution (pH 7.4) to each well of the Seahorse 24-well plate; replace green sensor cartridge on the top (*see* **Note 7**).
2. Incubate entire cartridge into a non-CO_2 incubator at 37 °C overnight or for up to 72 h (*see* **Note 8**).

3.1.3 Turning On Seahorse Instrument and the Seahorse XF24 Software (Fig. 1, See **Note 9**)

1. Click the icon of "Seahorse XF24 Analysis Software" on the desktop.
2. Continue with clicking "Standard."
3. Then choose "Seahorse Guest."

3.2 The XF24 Assay: Day 2

All reagents should be at pH 7.4 and 37 °C at all times for the XF24 assay.

3.2.1 Changing Media for XF24 Cell Plate from Subheading 3.1.1

1. Prepare XF media (*see* **Note 10**).

   For ECAR: 49 mL XF media with 500 μL 100 mM sodium pyruvate and 500 μL 200 mM glutamine.
   For OCR: 49 mL XF media with 500 μL 2.5 M glucose and 500 μL 100 mM sodium pyruvate.
2. Check cells to assure that they are even and around 90% confluent in monolayer.
3. Remove almost all growth media from all wells (*see* **Note 11**).
4. Wash once with 1 mL XF assay media. Aspirate it off, and add 500 μL XF media to each well.
5. Place plate in non-CO_2 incubator at 37 °C until ready to assay (*see* **Note 12**).

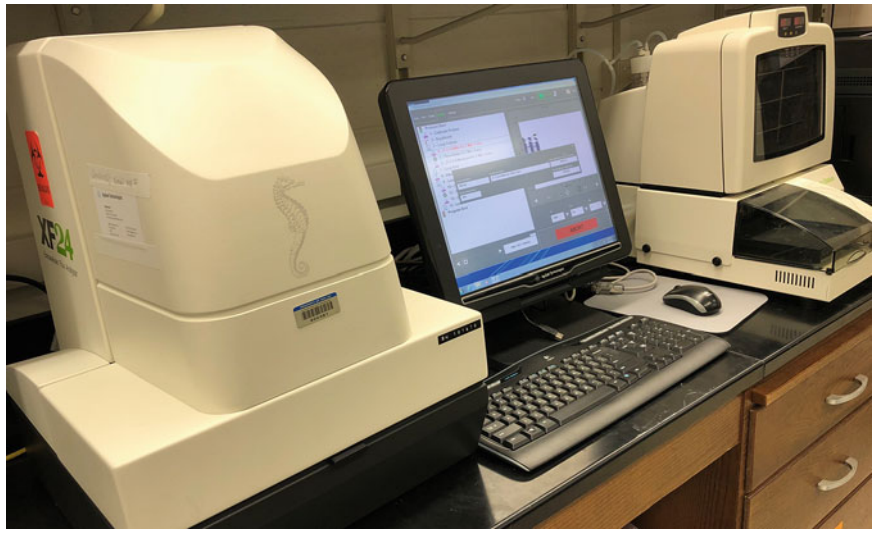


Fig. 1 Overview of a whole set of Seahorse XF24 instrument. From left to right, they are Seahorse XF24 analyzer, combination computer with touchscreen display, and non-CO_2 incubator

Table 1
Location of each port in sensor cartridge

D	C
B	A

Table 2
Compound preparation for ECAR

Compound	Injection port	XF media volume (μL)	Volume of stock compound (μL)	Injection volume (μL)
Glucose (450 μg/L)	A	1386	154	55
Oligomycin (1 mM stock)	B	1736	17.4	61.6
2-DG (1 M)	C	900	900	68.5

3.2.2 Loading Testing Compounds in the Sensor Cartridge

The sensor cartridge is equipped with four injection ports for each well (Port A, B, C, and D, the order as shown in Table 1) (*see* **Note 13**). Dilute compounds with XF media from Subheading 3.2, **step 1**, and inject them respectively into the ports according to following Tables 2 and 3 (*see* **Notes 14–16**).

3.2.3 Calibrating the Sensors and Running the Seahorse XF24 Assay

1. Click "Assay Wizard" (Fig. 2) to define all your experimental parameters, including general information (Fig. 3a), background correction (Fig. 3b), Groups&Labels (Fig. 3c), and

Table 3
Compound preparation for OCR

Compound	Injection port	XF media volume (μL)	Volume of stock compound (μL)	Injection volume (μL)
Oligomycin (1 mM stock)	A	1540	15.4	55
FCCP (10 mM stock)	B	1736	1.74	61.6
Rotenone (10 mM stock)	C	1918	1.92	68.5

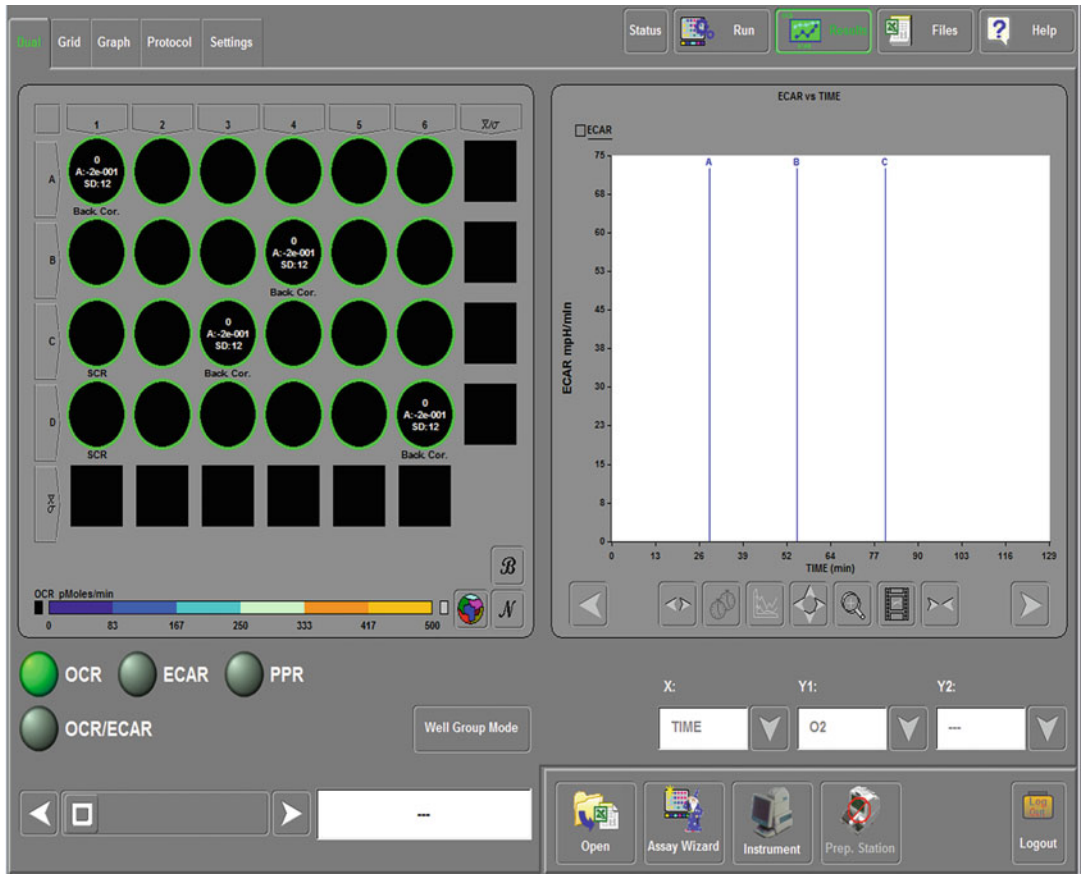


Fig. 2 The interface of the Seahorse XF24 software. "Assay Wizard" is on the bottom right line of the window; "Open" is to the left of "Assay Wizard"

running protocol (Fig. 4). Click "End," and choose "Save your template" or "End wizard." Or, click "Open" to load your saved template (Fig. 2; *see* **Note 17**).

2. Click "Start" to start the program, and then load the sensor cartridge for the calibration process (*see* **Notes 18–21**).
3. When calibration is complete (takes ~20 min), replace calibration plate with cell plate, and click "Continue."

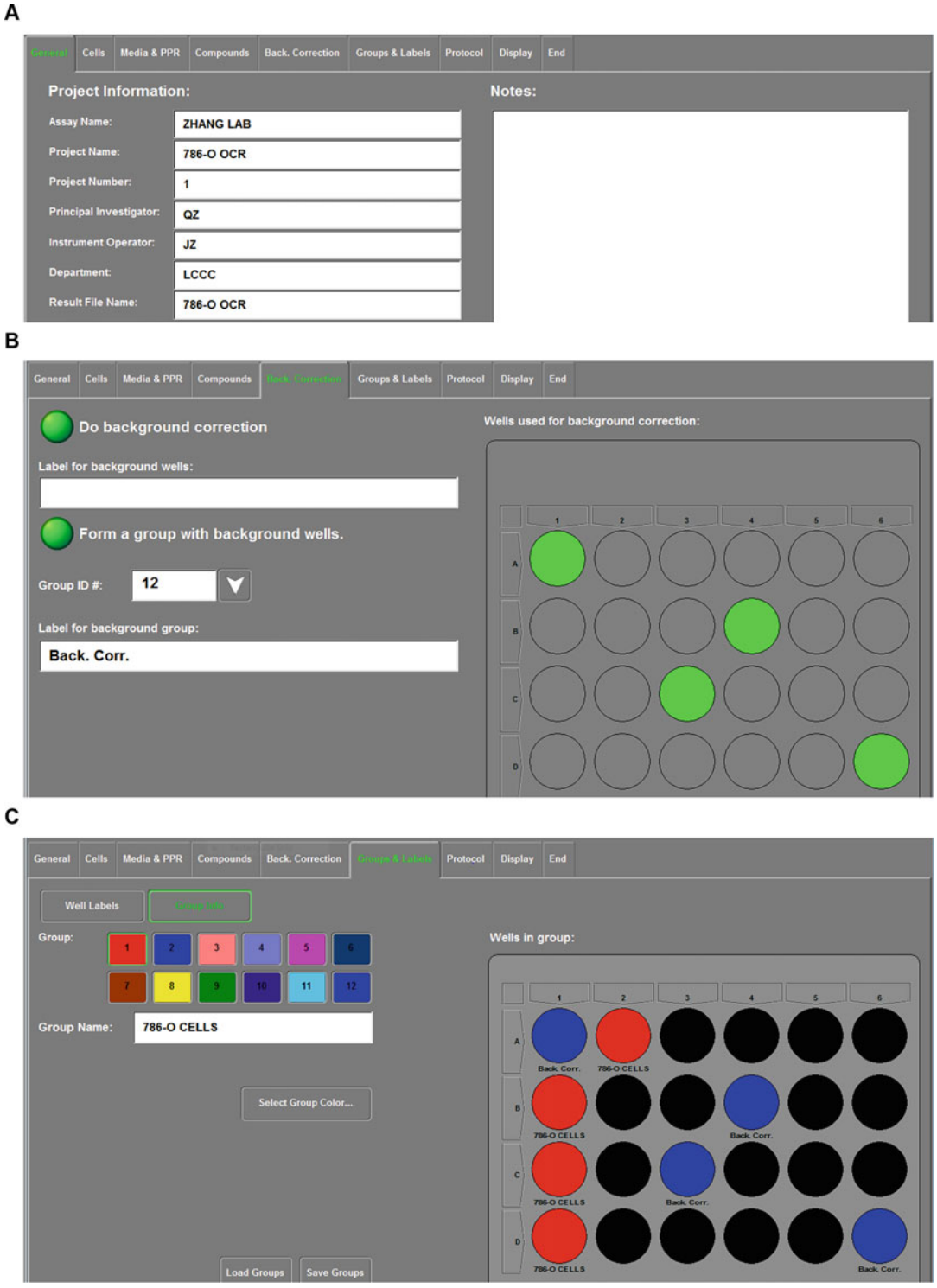


Fig. 3 The interface of the Assay Wizard related to all experimental parameters. Define all your experimental parameters from left to right. (**a**) Example of defining general project information. (**b**) Overview of choosing background correction well and "Do background correction." (**c**) Example of labeling sample groups

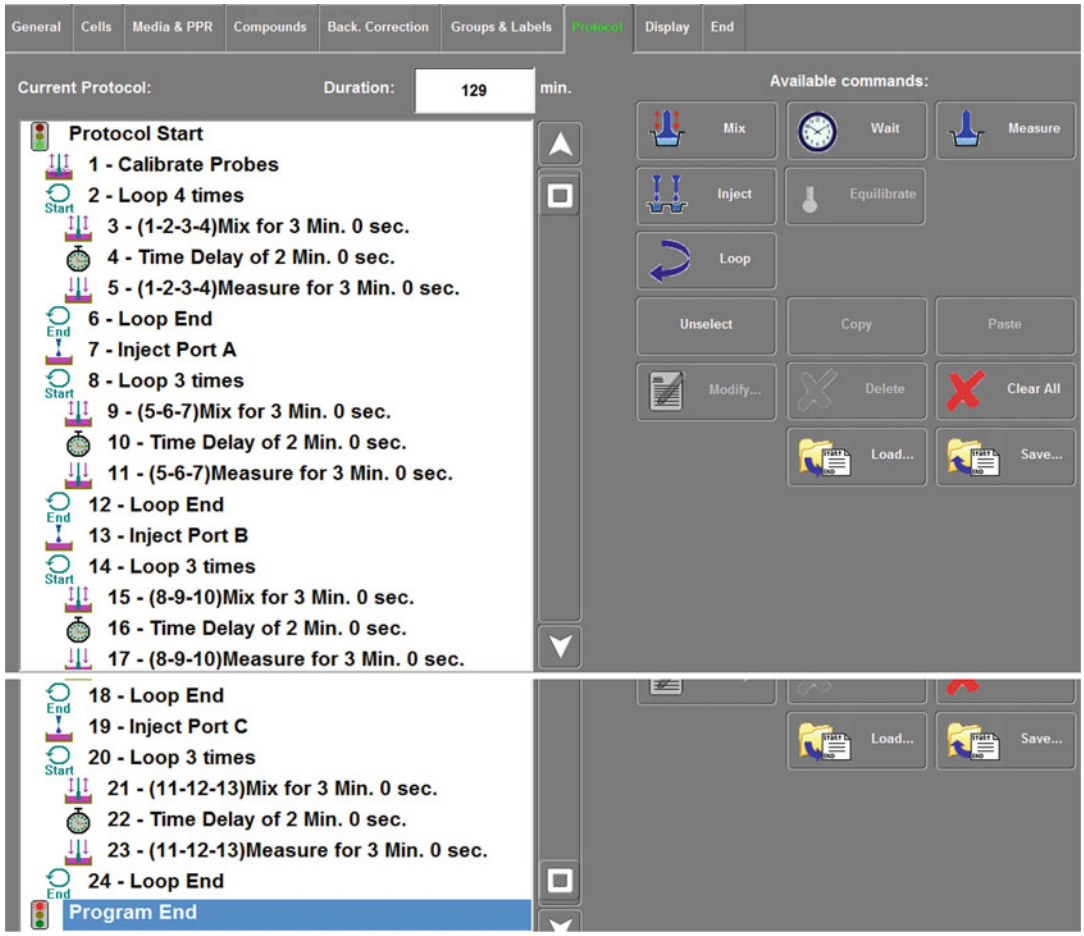


Fig. 4 The interface of running protocol. The left side of the window is ECAR and OCR running protocol; the right side "available commands" can be used to modify the protocol according to different assay purposes

3.3 Analyzing the Seahorse XF24 Data

Seahorse XF24 data can be analyzed directly by the Seahorse software, as shown in Fig. 5. Alternatively, go to the Agilent Seahorse XF Software https://www.agilent.com/en/support/cell-analysis/seahorse-xf-software, and download these useful tools. Results can be automatically generated through the software. Some researchers use other software to generate graphs manually from the raw data (*see* **Note 22**).

4 Notes

1. The seeding surface of each well of XF24 cell culture plate has the same size as a typical 96-well plate.
2. Accurate XF24 measurement requires the analyzed cells to be seeded and grown in an even and uniform monolayer configuration. Cell confluence between 90% and 100% is very important

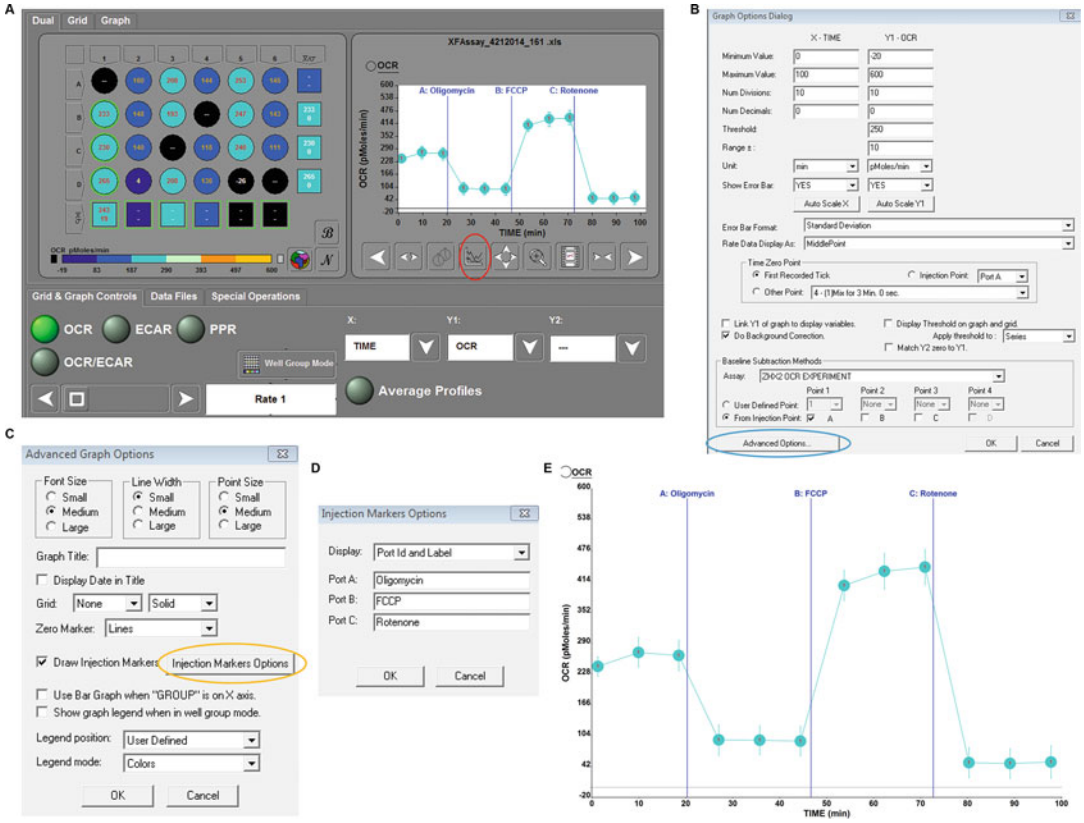


Fig. 5 The interface of analyzing Seahorse data. (**a**) Click "OCR"/"ECAR" (what you are measuring) on the bottom left line of the window, and choose OCR/ECAR (what you are measuring) for Y1 on the bottom right side; click "Well Group Mode" in between to show the comparison among groups in the graph on the right. Then click the icon in red circle to open the window as shown in (**b**); X and Y values can be adjusted according to the result. Click the icon in blue circle to further open window as shown in (**c**); font size, line width, point size, etc. for graph can be modified in detail. Click the icon in orange circle to add compound information in the graph (**d**). (**e**) Example of oxygen consumption rate (OCR) result in 786-0 cells by using the Seahorse XF24 instrument

for the XF assay. Seeding density depends on the cell size, although typically from 20,000 to 100,000 cells per well. We recommend that the cell density is first optimized before the commencement of the experiment. It is also recommended to use at least triplicate wells/group for the XF24 assay.

3. For strongly adherent cells, it generally takes around 1 h for the cells to adhere; for less adherent cells, it may take up to 6 h.
4. Slowly add 150 μL of growth medium to the side of each well in order to not disturb the newly attached cells, to bring the total volume of media to 250 μL per well.
5. The assay for comparison across different cell lines with various growth rates does not necessarily require overnight culture of

cells, and we therefore recommend to proceed to the next step within their doubling time.

6. The assay for nonadherent cells and centrifugation of cell suspension in the XF24 plate coated with Cell-Tak Cell and Tissue Adhesive will allow these cells to be attached to the bottom of wells. Thus, cells can be suspended directly in XF assay medium, after which you should proceed to the next step without the requirement of overnight incubation.
7. Each sensor cartridge has four reagent delivery ports per well, used for the injection of testing compounds into their corresponding wells during the assay.
8. If cartridge is being hydrated for more than 24 h, wrap it in parafilm to avoid evaporation. The hydration should be no less than 4 h.
9. This step is to allow the instrument and its program to stabilize at 37 °C, because the stable temperature of 37 °C of the XF instrument is required for accurate measurements. We recommend to leave the instrument to stabilize to 37 °C overnight.
10. Make fresh XF for each experiment.
11. Do not remove all media completely from wells; leave some medium (~50 μL) in each well to avoid cells being dried out. Take extra caution not to disturb the cells at the bottom of the wells.
12. A 1-h non-CO_2 incubation is optimal and necessary for de-gassing the cell plate, allowing for CO_2 diffusion.
13. Testing compounds need to be loaded into the ports before calibration of the sensor cartridge.
14. Different compounds and combinations of selected compounds can be used dependent on the purpose of the assay. Also, the concentration of each compound should be optimized based on cell type. Detailed information for these compounds and related interpretation of Seahorse results have been previously reported [16]. We provide the concentrations of the compound here, as they can be a good starting point for researchers to optimize around.
15. The injected compounds are diluted 10× for Port A, 11× for Port B, 12× for Port C, and 13× for Port D with XF24 assay media.
16. XF24 assay medium or compounds should be loaded in all the ports for all wells including background temperature correction or unused wells, serving as controls to ensure the proper injection in all wells.
17. For your experimental parameters, make sure to define all sample groups, check off “do background correction,” and

then run the protocol. Then save your template such that you can load your template easily when you start your assay in the future.

18. Before you confirm "start," you can also define or modify the file directory saved previously.
19. Plate should be at 37 °C, and it can be warmed in non-CO_2 incubator if necessary.
20. The notch on the cartridge should be located at the bottom left for a correct loading.
21. For the accurate measurement of ECAR and OCR, all the sensors in the cartridge should be individually calibrated to determine each sensor gain based on the sensor output, which is measured in the calibration reagent with known pH value and oxygen concentration.
22. Given potentially different growth rates among different cell lines, it is recommended to count cell number or measure protein concentration in each well for normalization of XF assay results.

Acknowledgment

We thank Yong Liu and Agilent agents for helpful discussions and suggestions. Jing Zhang is supported by a DoD BCRP Breakthrough Fellowship Award (W81XWH-17-1-0016). Qing Zhang is supported by grants from the National Cancer Institute (R01CA211732, R21CA223675) and American Cancer Society (RSG-18-059-01-TBE). Qing Zhang also received a career development award from the DoD (W81XWH-15-1-0599). Qing Zhang is a V Scholar, Kimmel Scholar, Komen Career Catalyst Awardee, and Mary Kay Foundation Awardee.

Reference

1. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144 (5):646–674. https://doi.org/10.1016/j.cell.2011.02.013
2. Ward PS, Thompson CB (2012) Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. Cancer Cell 21 (3):297–308. https://doi.org/10.1016/j.ccr.2012.02.014
3. Pavlova NN, Thompson CB (2016) The emerging hallmarks of cancer metabolism. Cell Metab 23(1):27–47. https://doi.org/10.1016/j.cmet.2015.12.006
4. Herst PM, Tan AS, Scarlett DJ, Berridge MV (2004) Cell surface oxygen consumption by mitochondrial gene knockout cells. Biochim Biophys Acta 1656(2-3):79–87. https://doi.org/10.1016/j.bbabio.2004.01.008
5. Warburg O, Posener K, Negelein E (1924) Über den Stoffwechsel der Carcinomzelle. Biochem Z 152:309–344
6. Wallace DC (2012) Mitochondria and cancer. Nat Rev Cancer 12(10):685–698. https://doi.org/10.1038/nrc3365
7. Guppy M, Leedman P, Zu X, Russell V (2002) Contribution by different fuels and metabolic pathways to the total ATP turnover of

proliferating MCF-7 breast cancer cells. Biochem J 364(Pt 1):309–315

8. Nakashima RA, Paggi MG, Pedersen PL (1984) Contributions of glycolysis and oxidative phosphorylation to adenosine 5′-triphosphate production in AS-30D hepatoma cells. Cancer Res 44(12 Pt 1):5702–5706
9. Gatenby RA, Gillies RJ (2004) Why do cancers have high aerobic glycolysis? Nat Rev Cancer 4 (11):891–899. https://doi.org/10.1038/nrc1478
10. Moreno-Sanchez R, Rodriguez-Enriquez S, Marin-Hernandez A, Saavedra E (2007) Energy metabolism in tumor cells. FEBS J 274(6):1393–1418. https://doi.org/10.1111/j.1742-4658.2007.05686.x
11. Zu XL, Guppy M (2004) Cancer metabolism: facts, fantasy, and fiction. Biochem Biophys Res Commun 313(3):459–465
12. Strohecker AM, Guo JY, Karsli-Uzunbas G, Price SM, Chen GJ, Mathew R, McMahon M, White E (2013) Autophagy sustains mitochondrial glutamine metabolism and growth of BrafV600E-driven lung tumors. Cancer Discov 3(11):1272–1285. https://doi.org/10.1158/2159-8290.CD-13-0397
13. Guo JY, Karsli-Uzunbas G, Mathew R, Aisner SC, Kamphorst JJ, Strohecker AM, Chen G, Price S, Lu W, Teng X, Snyder E, Santanam U, Dipaola RS, Jacks T, Rabinowitz JD, White E (2013) Autophagy suppresses progression of K-ras-induced lung tumors to oncocytomas and maintains lipid homeostasis. Genes Dev 27(13):1447–1461. https://doi.org/10.1101/gad.219642.113
14. Weinberg F, Hamanaka R, Wheaton WW, Weinberg S, Joseph J, Lopez M, Kalyanaraman B, Mutlu GM, Budinger GR, Chandel NS (2010) Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. Proc Natl Acad Sci U S A 107(19):8788–8793. https://doi.org/10.1073/pnas.1003428107
15. Guo JY, Chen HY, Mathew R, Fan J, Strohecker AM, Karsli-Uzunbas G, Kamphorst JJ, Chen G, Lemons JM, Karantza V, Coller HA, Dipaola RS, Gelinas C, Rabinowitz JD, White E (2011) Activated Ras requires autophagy to maintain oxidative metabolism and tumorigenesis. Genes Dev 25(5):460–470. https://doi.org/10.1101/gad.2016311
16. Qian W, Van Houten B (2010) Alterations in bioenergetics due to changes in mitochondrial DNA copy number. Methods 51(4):452–457. https://doi.org/10.1016/j.ymeth.2010.03.006
17. van der Windt GJ, Chang CH, Pearce EL (2016) Measuring bioenergetics in T cells using a seahorse extracellular flux analyzer. Curr Protoc Immunol 113:3 16B 11–13 16B 14. https://doi.org/10.1002/0471142735.im0316bs113
18. Zhang J, Nuebel E, Wisidagama DR, Setoguchi K, Hong JS, Van Horn CM, Imam SS, Vergnes L, Malone CS, Koehler CM, Teitell MA (2012) Measuring energy metabolism in cultured cells, including human pluripotent stem cells and differentiated cells. Nat Protoc 7(6):1068–1085. https://doi.org/10.1038/nprot.2012.048

Chapter 19

Metabolic Profiling of Live Cancer Tissues Using NAD(P)H Fluorescence Lifetime Imaging

Thomas S. Blacker, Michael D. E. Sewell, Gyorgy Szabadkai, and Michael R. Duchen

Abstract

Altered metabolism is a hallmark of cancer, both resulting from and driving oncogenesis. The NAD and NADP redox couples play a key role in a large number of the metabolic pathways involved. In their reduced forms, NADH and NADPH, these molecules are intrinsically fluorescent. As the average time for fluorescence to be emitted following excitation by a laser pulse, the fluorescence lifetime, is exquisitely sensitive to changes in the local environment of the fluorophore, imaging the fluorescence lifetime of NADH and NADPH offers the potential for label-free monitoring of metabolic changes inside living tumors. Here, we describe the biological, photophysical, and methodological considerations required to establish fluorescence lifetime imaging (FLIM) of NAD(P)H as a routine method for profiling the metabolism of living cancer cells and tissues.

Key words NADH, NADPH, Autofluorescence, Fluorescence lifetime imaging, Live-cell microscopy, Cancer metabolism

1 Introduction

Since the pioneering work of Otto Warburg almost a century ago [1], it has become clear that the metabolic reprogramming of cancer cells is crucial in facilitating their uncontrolled proliferation [2–4]. The "Warburg effect" describes an increase in glucose uptake and lactate generation upon carcinogenesis, despite the presence of abundant oxygen, arising from an apparent switch in the means of producing adenosine-3,5-triphosphate (ATP) from oxidative phosphorylation (OXPHOS) to "aerobic glycolysis." The advantage of this metabolic shift is unclear, given that the complete oxidation of a glucose molecule through OXPHOS yields almost 20 times more ATP than via glycolysis alone [5]. Warburg originally attributed both his observations and the root cause of cancer to defective mitochondria, the organelles at the heart of cell metabolism, and the location of the OXPHOS machinery, whose

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_19,

impairment has now been implicated in numerous other pathologies [6]. For many decades, the Warburg effect was presumed to be a biochemical hallmark of all cancers, forming the basis of a widely successful diagnostic tool, positron emission tomography (PET), which measures enhanced radiolabelled glucose uptake to detect tumors [7]. However, later research showed that many tumors and their precursors defy Warburg's hypothesis, predominantly maintaining high ATP/ADP ratios using mitochondrial OXPHOS [8]. Further, subsequent studies have also shown that tumors displaying the Warburg effect still possess functional mitochondria [9–11]. These findings demonstrate a significantly more complex link between cancer and metabolism than in Warburg's initial description. Perhaps unsurprisingly, therefore, a heavy focus on identifying ubiquitous characteristics of neoplastic tumors to develop a "one-size-fits-all" approach to cancer treatment has so far proven unsuccessful, the wide acceptance of the Warburg effect perhaps diverting attention away from the unique metabolic phenotypes of individual cancers [12]. As understanding the pathways active in individual tumors is crucial to developing effective treatment strategies [13], tools for the metabolic profiling of cancer tissues are vital. This chapter concerns the application of time-resolved measurements of cellular autofluorescence to provide a window into cellular metabolism in cancer, using the intrinsic photophysical properties of the reduced metabolic cofactor nicotinamide adenine dinucleotide (NADH) and its phosphorylated counterpart NADPH.

1.1 The Role of NAD(P) in Cancer

Mitochondria are cytoplasmic organelles that, in normal differentiated cells, act as the primary ATP generator through aerobic respiration [5], in addition to regulating Ca^{2+} signalling, the production of reactive oxygen species (ROS), and cell death [6]. In the presence of oxygen, most differentiated cells catabolize their main fuel, glucose, first through its cytosolic conversion in glycolysis to pyruvate and then by its subsequent transport into the mitochondria and oxidation via the tricarboxylic acid (TCA) cycle [3]. In both mitochondria and cytosol, the electrons removed during this chain of oxidation reactions are passed to NAD+, reducing it to form NADH. The NADH generated by the TCA cycle donates electrons to the electron transport chain (ETC) located on the inner mitochondrial membrane. Glycolytically derived cytosolic NADH cannot cross the mitochondrial membrane, but its reducing equivalents may contribute to this pool by being passed to NAD+ in the mitochondrial matrix via the malate-aspartate shuttle [14]. The passage of the electrons donated by NADH along the complexes of the ETC to the terminal electron acceptor, oxygen, is coupled to the pumping of protons from the mitochondrial matrix to the intermembrane space, establishing the membrane potential that powers the production of ATP by ATP synthase. By ferrying electrons from the TCA cycle to the ETC, alongside those removed

during glycolysis, NADH plays a pivotal role in linking the oxidation of carbon sources to the storage of the energy liberated in the useable form of ATP.

Disruption of the passage of electrons along the ETC allows electrons to leak directly to O_2, causing its direct reduction to superoxide O_2^-. Such conditions occur in response to hypoxia or upon exposure to respiratory chain inhibitors such as cyanide and are reflected by an increased NADH/NAD+ ratio [15]. Superoxide acts as the proximal ROS from which other toxic species such as hydrogen peroxide (H_2O_2) can be produced, causing damage inside the cell by reacting directly with proteins, lipids, and DNA [14]. The cell contains two thiol-linked antioxidant defense systems to decrease oxidative stress, mediated by glutathione and thioredoxin. Both molecules are maintained in their functional, reduced form by the electron donor NADPH [14]. The phosphate group attached to the adenine end of this NADH analogue allows enzyme-binding sites to be distinct to those of its unphosphorylated counterpart. Thus, while the NAD pool participates in ATP-yielding catabolic reactions, the NADP pool contributes specifically to anabolic reactions, such as antioxidant maintenance and nucleic acid and fatty acid synthesis [16]. In the mitochondria, the NADPH pool is primarily maintained by the nicotinamide nucleotide transhydrogenase (NNT) which, powered by the mitochondrial membrane potential, reduces NADP+ by oxidizing NADH. The NNT thereby siphons a portion of the NADH produced by the TCA cycle to act in an antioxidant capacity as NADPH, defending against oxidative damage caused by the remaining NADH, should electron leak from the ETC occur [14, 17].

The NAD and NADP pools are maintained in vastly different redox balances due to their contrasting intracellular roles. The NADPH/NADP+ ratio is maintained high due to the primary role of the NADP pool in donating electrons to anabolic reactions. In the cytosol this is largely carried out by the pentose phosphate pathway, which may compete with glycolysis for glucose, reducing NADP+ rather than NAD+. Glycolysis itself requires that the NADH/NAD+ ratio is maintained low in the cytosol, as the pathway is unable to proceed if the NAD+ pool is not restored to provide electron acceptors for the key oxidation reactions [14]. While the malate-aspartate shuttle contributes to maintaining this balance, lactate dehydrogenase plays the critical role, converting the end product of glycolysis, pyruvate, into lactate, oxidizing NADH to NAD+ in the process. High glycolytic fluxes are therefore necessarily correlated with significant lactate production [18].

The high levels of aerobic glycolysis in cancer cells observed by Warburg occur only under anaerobic conditions in normal cells [3]. Possible reasons for this include a higher ATP production rate compared to OXPHOS [18], or a reduction in ROS generation, excessive quantities of which trigger apoptotic cell death

[19, 20]. Cancer cell metabolism is also no longer coupled to oxygen availability according to the Warburg effect, providing a selective advantage within the tumor microenvironment where oxygen concentration is often low or constantly fluctuating due to a disrupted blood supply [21, 22]. Furthermore, glycolytic intermediates may enter the pentose phosphate pathways via the transketolase and transaldolase enzymes, generating NADPH to provide the reducing equivalents required for fatty acid and nucleotide synthesis, in addition to ribose-5-phosphate, forming the backbones of RNA and DNA [14]. Consequently, aerobic glycolysis may fulfil an abundance of the biochemical requirements of a proliferating cell [23]. However, this route to fulfilling the bioenergetic needs of oncogenic proliferation is now known to be but one of many [4].

The metabolic phenotype adopted by a cancer cell is largely dictated by the specific oncogenes and tumor suppressor genes expressed within it, acting on diverse metabolic targets ranging from glucose metabolism and substrate transport to redox homeostasis and protein, lipid, and nucleic acid synthesis [4, 24, 25]. p53-deficient tumors have been shown to be particularly reliant on serine metabolism, feeding into the folate and methionine cycles for the reactions involved in one-carbon metabolism [26]. This has led to the emergence of serine starvation as a particularly attractive therapeutic target [27]. The importance of glutamine to cancers, feeding directly into the TCA cycle by its conversion into α-ketoglutarate via glutamate, has long been known and appears to be regulated by the c-MYC oncogene [28]. Additionally, metabolism may not only act to provide the energy and building blocks required for proliferation but may also generate products that directly cause oncogenic transformation, so-called oncometabolites; a mutation in the enzyme isocitrate dehydrogenase creates a mutant form that can convert α-ketoglutarate into 2-hydroxyglutarate (2-HG), the accumulation of which hampers differentiation [29]. These examples show that the twenty-first-century "post-Warburg" understanding of the links between cancer and metabolism is one of a panoply of potential routes by which metabolism can both drive and support tumorigenesis [4].

Importantly for the techniques described in this chapter, the majority of the metabolic pathways commandeered by a cancer cell to ensure its proliferation are dependent on redox reactions in which the NAD and NADP pools participate, as summarized in Fig. 1. In addition to the well-known roles of these cofactors in the glycolytic and OXPHOS pathways perturbed by the Warburg effect, serine biosynthesis involves the reduction of NAD+ to NADH, the conversion of glutamine to α-ketoglutarate involves the reduction of NADP+ to NADPH, and the production of 2-HG by mutant isocitrate dehydrogenase relies on reversing its normal function from an NADP+ reducing to an NADPH oxidizing form

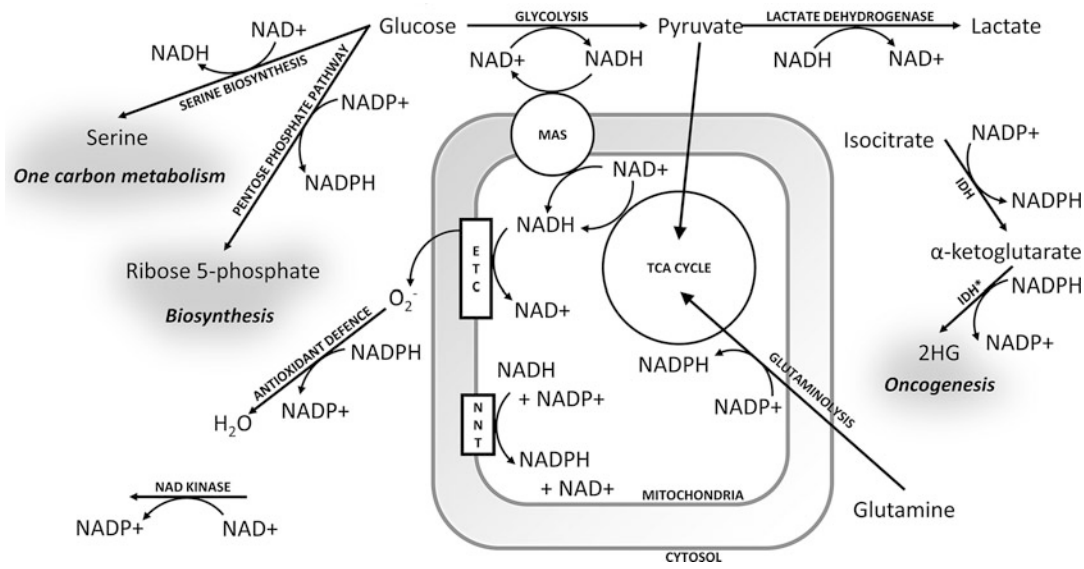


Fig. 1 NADH and NADPH play a role in a large number of the metabolic pathways known to be altered in cancer cells. NADH is involved in the pathways whose relative activities are altered by the Warburg effect, cytosolic glycolysis and mitochondrial energy metabolism, linking the tricarboxylic acid (TCA) cycle and electron transport chain (ETC). NADH is also produced during serine biosynthesis, a key feature of p53-deficient tumors. Meanwhile, NADPH is involved in the biosynthetic pathways required for cell proliferation. In the mitochondria, NADPH is primarily produced by the nicotinamide nucleotide transhydrogenase (NNT). In the cytosol, both the pentose phosphate pathway and isocitrate dehydrogenase (IDH) contribute. IDH mutation is a known oncogenic transformation, causing oxidation of NADPH and production of 2-hydroxyglutarate, an oncometabolite

[27–29]. The NAD and NADP pools therefore lie at the heart of the complex web of links between cancer and metabolism, and, as such, their biochemical status may offer insight into the specific metabolic pathways active in a given tumor. The development and application of such approaches are made all the more appealing as the reduced cofactors can be observed in living tissues using laser scanning microscopy without the need for the addition of extrinsic dyes [14].

1.2 Autofluorescence of NADH and NADPH

The reduction of NAD+ to NADH involves the transfer of a hydride ion to its nicotinamide ring. This decreases the energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) of the moiety by approximately 2 eV, shifting the wavelength of light at which it maximally absorbs from the far-ultraviolet 220 nm to the near-visible 340 nm [30]. As a complex molecule, the large number of vibrational degrees of freedom available in NADH broadens its absorption spectrum by around 30 nm on either side of this peak. Absorption of light at these wavelengths causes a transition into an

excited electronic state. In aqueous solution at room temperature, the fate of 98% of the absorption events is to cause small-scale motion of the nicotinamide ring [31]. However, the remaining 2% of the photons absorbed are reemitted as fluorescence. Excess vibrational energy dissipated to the surroundings following excitation causes the fluorescence photons to carry less energy than those absorbed. NADH therefore emits light of wavelength 460 (±25) nm [32].

The intrinsic fluorescence of NADH has been used as a reporter of metabolic state since the pioneering experiments of Britton Chance in the 1950s [33]. Chance used changes in the intensity of the fluorescence emitted from living tissues following illumination at 366 nm to monitor changes in the redox state of the NAD pool and thus interrogate the function of the ETC in intact tissues for the first time [34]. Our lab routinely uses a modern adaptation of these experiments to estimate the redox state of the mitochondria in live cell models [35]. In a confocal microscope with 351 nm excitation and 435–485 nm emission filtering, the resting NADH fluorescence level is compared with maximally oxidized and maximally reduced conditions. These are achieved, respectively, by application of the uncoupler carbonyl cyanide-p-trifluoromethoxyphenylhydrazone (FCCP), causing complex I of the ETC to oxidize NADH at its maximum rate, and cyanide, inhibiting the ETC and thus terminating NADH oxidation. However, a caveat to this method is in its assumption that NADH is the only fluorescent molecule present in this spectral region that may respond to these pharmacological perturbations. The contribution from NADPH is neglected [14].

As the phosphate group which allows enzymes to distinguish between the NADP and NAD pools lies at the adenine end of the molecule, far from the redox-active and fluorescent nicotinamide ring, NADH and NADPH are spectrally identical [32]. Conventional, intensity-based fluorescence measurements can therefore not distinguish between these two dinucleotide pools, leading to the labelling of their combined signal as NAD(P)H [14, 35]. However, such a technique would be desirable, given the drastically different elements of intracellular metabolism that the NAD and NADP pools are involved in regulating. We recently addressed this issue by applying fluorescence lifetime imaging microscopy (FLIM). The fluorescence lifetime of a molecule measures the average amount of time it spends in the excited state following absorption. Addition or removal of the pathways available for a molecule to leave the excited state will therefore alter the fluorescence lifetime [36]. Enzyme binding is known to restrict the small-scale motions that account for a large majority of the deexcitation events occurring in NADH [37], thereby decreasing the rate at which the molecules leave the excited state, increasing the

fluorescence lifetime. Concomitantly, a greater proportion of excitation events in enzyme-bound NADH will result in fluorescence compared to free NADH in solution, increasing its brightness. In solution, NADH and NADPH share a fluorescence lifetime of approximately 0.4 ns [31]. Our work suggested that, inside live cells, the binding site of NADPH increases its fluorescence lifetime to a greater extent than that of NADH, at around 4.4 ns compared to 1.5 ns [38]. Combining NAD(P)H imaging with FLIM thereby allows the relative contributions of NADH and NADPH to the total signal to be determined, allowing NAD- and NADP-associated pathways to be separately interrogated in living tissues for the first time.

1.3 Studying Cancer Metabolism Using NAD(P)H FLIM

Time-resolved measurements of NAD(P)H fluorescence in living biological samples were first performed over 25 years ago [39] when Schneckenburger and Koenig showed, in live yeast cells, that the fluorescence lifetime of NAD(P)H was increased relative to its lifetime free in solution, indicative of enzyme binding. In subsequent work [40], again in yeast, Paul and Schneckenburger demonstrated a correlation between oxygen tension and the average lifetime of the NAD(P)H signal. With the Warburg hypothesis in place for 70 years, this finding inevitably motivated the initial comparisons of the time-resolved fluorescence properties of NAD(P)H in cancers relative to healthy tissue. Pradhan and colleagues compared the fluorescence lifetime of NAD(P)H in metastatic and non-metastatic variants of a number of cancer cell lines in suspension [41]. The average lifetime was approximately threefold lower in metastatic cell lines relative to their non-metastatic counterparts. The potential for time-resolved NAD(P)H fluorescence measurements to act as a label-free "Warburg sensor" was thus established.

As discussed in Subheading 1.1, the Warburg effect is not a universal feature of all cancers; carcinogenesis is critically dependent on a wide range of metabolic shifts [4]. Given that a number of these cause alterations in the redox balance of the NAD and NADP pools and that fluorescence lifetime measurements are sensitive to the balance between the reduced forms of these two nucleotides [38], NAD(P)H FLIM measurements can be used as a direct reporter of metabolic differences between cancer tissues. The commercialization of time-resolved imaging add-ons to laser scanning microscope systems in the early 2000s has allowed the number of studies reporting differences in the fluorescence decay of NAD(P)H between healthy and cancer tissues to increase steadily from 2005 onward (Fig. 2). In this year, a now highly cited study by Bird and co-workers demonstrated the sensitivity of the NAD(P)H fluorescence lifetime to metabolic perturbation in breast epithelial cells, raising hope of an optical technique for staging breast tumors [42]. Correlations between the fluorescence lifetime and cancer

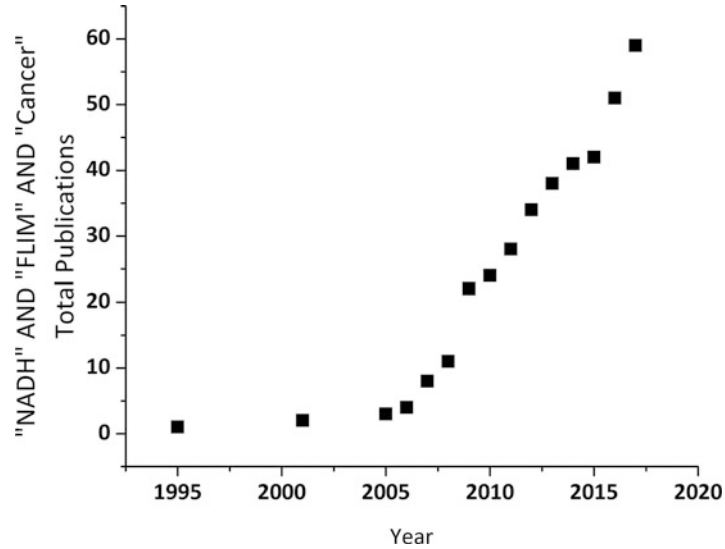


Fig. 2 Following commercialization of time-resolved fluorescence imaging technology in the early 2000s, the number of FLIM studies of NADH in cancer has increased approximately linearly since 2005 (Scopus)

stage were subsequently observed in a hamster cheek model of oral carcinogenesis [43]. In 2010, McGinty et al. demonstrated a wide-field instrument able to delineate the boundaries of colonic adenocarcinoma using time-resolved measurements of tissue autofluorescence [44]. Adur et al. have since shown that a significant increase in the average NAD(P)H lifetime inside the tumor contributes to this capability [45]. Longer average NAD(P)H fluorescence lifetimes have also been observed in malignant oral mucosa cell lines compared to their non-malignant counterparts [46] and in cancerous, relative to non-transformed, cervical tissues [47]. In contrast, Skala et al. showed in vivo that the fluorescence lifetime of NAD(P)H in healthy oral epithelial cells is longer than in the neighboring precancerous tissue [43]. Shorter average lifetimes were also observed in cancerous cell lines relative to control by Awasthi et al. [48] and in melanomas relative to healthy skin by Pastore et al. [49]. The complex changes in metabolism during carcinogenesis are therefore reflected in complex variations in NAD(P)H lifetimes. Nevertheless, FLIM of NAD(P)H has now established itself as a robust and reliable method for detecting metabolic shifts in cancers.

The majority of the studies discussed here, alongside those in the literature, are based on the time-correlated single-photon counting (TCSPC) technique. The principles of this method are outlined in Fig. 3. Its practical application for the study of metabolic shifts in cancers is the subject of the remainder of this chapter.

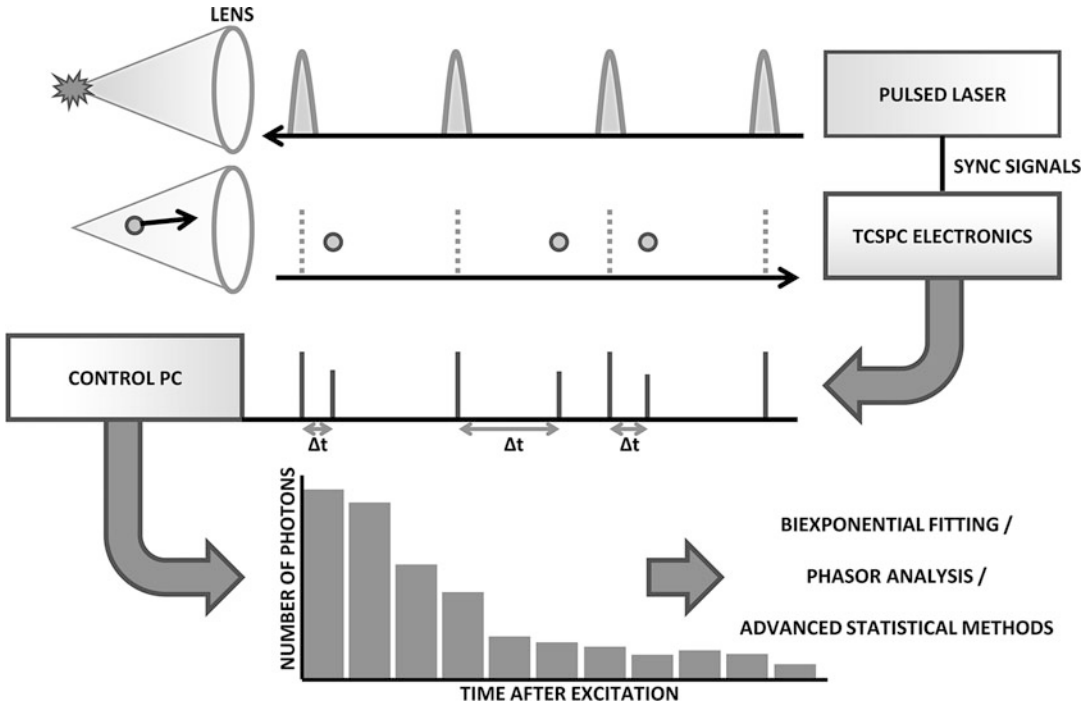


Fig. 3 A schematic overview of the time-correlated single-photon counting (TCSPC) method. Light from a pulsed laser source is focused onto a sample. Fluorescence photons emitted by the target molecules are registered by a detector. Electronics in a PC measure the time delay between the incident pulse and the fluorescence emission. The excitation-emission delay times are recorded in a histogram. This data is then analyzed to extract the fluorescence decay dynamics. FLIM combines this approach with the x-y coordinates provided by microscope scan heads to build up fluorescence decay histograms at every pixel of an image

2 Materials

2.1 Pulsed Laser Source

A pulsed excitation source is a key feature of the TCSPC technique and represents an important difference between fluorescence lifetime and routine confocal imaging [50]. Many well-equipped microscopy facilities will now offer two-photon imaging platforms, which rely on the use of a pulsed laser operating in the red to near-infrared portion of the spectrum to excite the target molecule using the near-simultaneous arrival of two photons each carrying half the energy of the single-photon transition [51]. Consequently, FLIM electronics can be added directly to these existing systems without the requirement for a new dedicated laser. The widely tuneable Ti:sapphire laser has become the workhorse instrument for such applications, providing short (~100 fs) pulses operating at ~80 MHz repetition rates [52]. This ensures that there is sufficient delay between excitation pulses (~10 ns) to measure the NAD(P)H fluorescence decay, with more than 99% of the initial excited state population having decayed by the time the next pulse arrives.

Two-photon excitation provides a means by which to achieve increased depth penetration into complex biological tissue preparations [53], making it a popular choice for clinical application of NAD(P)H FLIM. However, NAD(P)H FLIM with single-photon excitation has previously been applied [54], offering the advantage of increasing the signal at the expense of axial resolution by opening the confocal pinhole [55]. A number of excitation laser solutions could be applied for this purpose. Single wavelength diode-pumped solid state (DPSS) lasers with <100 ps pulse widths are becoming available operating at the appropriate wavelength for single-photon excitation at ~340 nm, the output of a Ti:sapphire laser can be frequency doubled using an appropriate crystal (e.g., lithium triborate), or an optical parametric oscillator system can be used to extend the tuning range of the Ti:sapphire into the visible.

2.2 Laser Scanning Microscope

The pulsed laser excitation source must be coupled to a laser scanning confocal microscope equipped with an objective lens of suitable magnification for the biological preparation to be studied. An appropriate selection of dichroic mirrors and emission filters must also be available. Short-pass dichroic mirrors are required for two-photon excitation as the fluorescence emission results at a shorter wavelength than the incident illumination. The excitation conditions used for NAD(P)H may also cause absorption in a number of other intrinsic fluorophores, particularly flavins which emit at longer wavelengths around 530(±30) nm [38]. Emission filters specific for NAD(P)H should therefore be used, such as a band pass around the emission peak at 460 nm. FLIM relies on performing the TCSPC technique at each pixel of the image, so the microscope must also provide scan signal outputs for the pixel registration of the detected photons and allow the attachment of an external detector (*see* Subheading 2.5).

2.3 Sample Mounting Considerations

Maintaining the biological integrity of the sample of interest at the microscope is a vital aspect of the imaging process. The sample must be kept sufficiently still over the time required to acquire sufficient photons to generate a FLIM image (of the order of minutes; *see* Subheading 5.3) so as to not cause spatial blurring in the resulting dataset. The choice of buffer solutions, perfusion systems, and environmentally controlled chambers to maintain the physiological integrity of the living sample will vary from preparation to preparation, and determining the correct conditions is likely to be a significant undertaking at the start of a new research project. However, live cell and tissue imaging is now a sufficiently mature field that commercial solutions to a range of sample mounting issues are available, allowing NAD(P)H FLIM studies to be performed on a diverse array of biological samples, from cell cultures on coverslips to in vivo tumors.

2.4 TCSPC Detector

The internal detectors of the microscope are unlikely to provide the single-photon sensitivity required to perform the TCSPC technique. As such, an external detector must be attached to the microscope at a suitable exit port. Microscope systems designed for multiphoton imaging should provide non-descanned ports for this purpose. As two-photon excitation does not require the use of a confocal pinhole, the path followed by fluorescence to exit a non-descanned port should contain fewer optical elements, increasing the detection efficiency and ultimately the sensitivity of the technique [56]. The most frequently used detectors for TCSPC applications are photomultiplier tubes (PMTs) [57]. Here, a single photon causes the ejection of electrons from a photocathode which, via acceleration through an electric field and subsequent collisions with further dynodes, causes a detectable current pulse at the anode. The use of a hybrid PMT should be considered for studies involving NAD(P)H fluorescence lifetimes [58]. These detectors remove a large number of the signal amplification steps, instead accelerating photoelectrons directly into a semiconductor diode. This shortens its time resolution by almost an order of magnitude from the hundreds of picoseconds for a conventional PMT, close to the fast lifetime of free NAD(P)H, increasing the accuracy with which the fluorescence decay parameters can be determined. The extreme sensitivity of these detectors makes them highly susceptible to damage from leakage of ambient light, so a shutter assembly with automatic overload sensing is crucial.

2.5 TCSPC Electronics

While the earliest implementations of the TCSPC technique were modular pieces of electronic laboratory apparatus, modern applications make use of miniaturized circuitry such that all the components can be condensed onto a hardware card inside a desktop PC [59]. Counting cards are available commercially from Becker and Hickl and PicoQuant, among others. These companies are also the preferred source of bespoke detectors and shutter assemblies optimized to both their photon-counting electronics and the microscope upon which it is to be mounted. The cards will register the detection of photons and measure the time delay between the laser pulse and emission (*see* Fig. 3), ascribing the event to the appropriate pixel using the signals provided by the microscope scanner. The hardware will interface with appropriate acquisition software to control the relevant experimental parameters, such as collection time, and allow the data to be saved for subsequent analysis.

2.6 Operation and Analysis PC

The fast rate of detected emission events induced by a MHz excitation pulse train prompts the recommendation of a dedicated PC for control of the FLIM acquisition, with background software kept to a minimum to avoid buffer overflows and data corruption. The resulting image files contain count data for hundreds of channels in each of tens of thousands of pixels, so the resulting images can be

large (10–100 MB per image). As such, access to high-volume (~TB) external storage is important for multiuser facilities. Such facilities should also consider dedicated analysis computers for extracting fluorescence decay parameters from TCSPC data as pixel-by-pixel curve-fitting approaches (*see* Subheading 4) can be time-consuming. The more powerful these machines, the faster the fitting and subsequent analysis can take place.

3 Methods

1. Tune the laser to the appropriate wavelength. For single-photon studies, this should be as close to the absorption peak at 340 nm as possible. With two-photon excitation, NAD(P)H maximally absorbs at 700 nm [60]. However, this is close to the minimum wavelength achievable using a Ti:sapphire laser, and many models may be unable to maintain a stable, mode-locked output. In such situations, a wavelength as close to 700 nm as possible should be used, in order to maximize the proportion of NAD(P)H excited relative to other intrinsic fluorophores that may absorb in this spectral range [60]. The incident power should be chosen in accordance with the considerations discussed in Subheading 5.2.
2. Select a dichroic mirror appropriate to the excitation conditions used. This optical component allows the incident illumination to be separated from the spectrally shifted fluorescence. The mirror should reflect the excitation wavelength onto the sample through the objective, allowing the fluorescence to be transmitted through it to the detectors. For single-photon excitation, the dichroic should transmit light of longer wavelength than that absorbed by the NAD(P)H. Dichroic mirrors for two-photon must transmit light of a shorter wavelength than that reflected.
3. Select emission filters appropriate for isolating the NAD(P)H signal. A band-pass filter centered on the emission peak at 460 nm is recommended. While this cannot guarantee exclusion of all autofluorescent compounds that absorb in this spectral region, we have previously estimated that use of a 460 ($\pm$25) nm emission filter with excitation between 700 nm and 740 nm is 95% specific for NAD(P)H [38].
4. Locate the sample in the microscope eyepieces using bright-field illumination and roughly bring it into focus. Start a fast scan, and use the internal detectors of the microscope to observe the NAD(P)H fluorescence from the sample. Fine-tune the focus to observe the desired focal plane and translate the image until the required region is located. Rotate and crop the image as necessary, remembering that increased zoom may

reduce the signal levels obtainable due to the increased likelihood of photodamage.

5. Adjust the microscope settings to send the NAD(P)H signal to the external detector connected to the TCSPC system. Begin a fast scan, and acquire a FLIM image for a predetermined length of time, chosen based on the considerations discussed in Subheading 5.3.
6. Repeat the imaging process across different regions of the same sample, if possible, and across biological replicates for each experimental condition. Guidance for choosing a suitable number of repeats is given in Subheading 5.5.

4 Data Analysis

The data produced by a traditional TCSPC experiment performed on a fluorophore solution in a cuvette consist of the number of photons emitted by the solution as a function of time after excitation by a laser pulse. Time-resolved fluorescence microscopy extends this into two spatial dimensions, obtaining these decay measurements at every pixel of an image. A FLIM image is therefore a significantly more complex entity than a standard intensity-based fluorescence image, the analysis of which can subsequently be more involved. Here, the principle means by which to derive metabolic information from the data contained within an NAD(P)H FLIM image are discussed.

4.1 Biexponential Fitting

NADH and NADPH may bind to a vast array of different enzymes inside the cell, all of which may induce slightly different fluorescence lifetimes [14]. The cofactors may also exist in an unbound form, exhibiting a correspondingly shorter lifetime [31]. Each pixel of an NAD(P)H FLIM image could therefore be expected to contain a highly heterogeneous mix of different fluorescent species with different fluorescent lifetimes. In such a circumstance, multi-exponential fitting can be employed. This method is generally available in the software provided by commercial vendors of FLIM systems, such as Becker & Hickl's SPCImage and PicoQuant's Fluofit. A fitting algorithm varies the lifetimes and weightings of a sum of exponential decays until a good fit is achieved between model and data [61]. The relative weightings of each decay component then represent the relative abundances of each fluorescent species. However, as fluorescence emission is a stochastic process, Poisson noise inherent in the fluorescence decay reduces the ability to separate the contributions of every species to the measured signal [38]. The signal-to-noise ratio of such a process increases as the square root of the signal [62]; however at the signal levels obtainable in live tissues, only two components can

be reliably resolved [38]. A basic rule of thumb states that ten times more signal is required to resolve every additional component [63]. Observing even a third component would therefore require a tenfold increase in the acquisition time of each image to 10 min or more, rendering the technique impractical and increasing the risk of phototoxicity [64]. Fortunately, the two lifetimes extracted have been shown to represent average values of the free and enzyme-bound species present, τ_{free} and τ_{bound} [38]. Their relative weighting α_{bound} therefore provides the percentage of total NAD(P)H that is enzyme-bound at a given pixel.

4.2 Statistical Fitting Methods

Biexponential analysis using weighted least-squares curve fitting is the most commonly employed method for extracting the NAD(P)H fluorescence decay parameters from the acquired FLIM data [14]. The method relies on minimization of the χ_R^2 fitting statistic, which compares the magnitude of deviations between the proposed model and the experimental data relative to the differences expected purely from the presence of Poisson noise [65]. If the difference between the data and the model is larger than that expected from the noise, χ_R^2 is large. If the only differences between the model and the data can be ascribed to Poisson noise, χ_R^2 is approximately 1. The software fitting algorithm acts by varying the model parameters until χ_R^2 is minimized. In conditions of low signal, such as may be encountered in cells with low NADH and NADPH levels or with high susceptibility to phototoxicity, the statistical assumptions of the least-square fitting process can break down [66], and the lifetimes become biased toward incorrect values [67]. This can be corrected by making use of a maximum likelihood estimator for the fitting statistic, which aims to find the fluorescence decay parameter values that maximize the statistical likelihood of obtaining the measured dataset [68, 69]. This approach is available in the OMERO FLIMfit software [70]. By recasting the problem as the recovery of a probability distribution of lifetimes present in the data, it is also possible to apply the maximum entropy method to further refine the correct extraction of fluorescence decay parameters [71, 72]; however this approach has, to the best of our knowledge, yet to be applied to FLIM. One potentially powerful advanced statistical method which has been successfully applied to time-resolved fluorescence imaging is the "fit free" approach of Bayesian inference [73, 74]. Here, the arrival times of each photon in turn are used to update the likelihood of a given underlying set of decay parameters. The method has been shown to give significantly more precise estimates of fluorescence lifetimes and decay amplitudes for measurements up to 10,000 photons, similar to that which may be obtained in an NAD(P)H FLIM image with acquisition times of the order of minutes [73].

4.3 Phasor Analysis

The enhanced precision analysis approaches described above will play an important role in the application of FLIM as a precise, quantitative experimental technique, providing absolute, numerical information to aid the ongoing goal of a predictive, model-based twenty-first-century biology [75]. However, to date, NAD(P)H FLIM has largely been applied as a qualitative and descriptive tool in which changes in the fluorescence decay parameters are used to infer alterations in the metabolic state of a tissue. User-friendliness and minimized data processing times are therefore a priority in the choice of analysis method. Over the last decade, phasor analysis of NAD(P)H FLIM data has become a popular means by which to fulfil these criteria, generally using the Globals software (Laboratory for Fluorescence Dynamics, Irvine, USA) [76]. The method involves using the real and imaginary components of the Fourier transform of the TCSPC data at each pixel of the FLIM image as coordinates of points in a two-dimensional phase space known as the phasor plot. Each pixel in the FLIM image will correspond to a location in the phasor plot, the coordinates of which reflect the shape of the fluorescence decay at that pixel. Pixels with similar fluorescence decay characteristics will cluster around similar locations in the phasor plot, providing a graphical means by which to deduce heterogeneity within the fluorescence lifetimes present across an image [76]. Changes in the relative abundance of two or more species can also be graphically inferred; pure solutions of freely diffusing and enzyme-bound NAD(P)H will occupy specific points in the phasor plot, and changing the proportion of free and bound NAD(P)H in a mixture of the two species will move the coordinates of its fluorescence decay along the straight line joining the two primary points. As the phasor approach does not require fitting of the fluorescence decay data, it is computationally simple and therefore fast. For this reason, its implementation has contributed significantly to the growth in applications of NAD(P)H fluorescence lifetime imaging in recent years [77–79].

5 Notes

5.1 Instrument Response Function Measurement

The TCSPC components of a FLIM system do not operate infinitely fast. In fact, a distribution of lag times may exist between arrival of a fluorescence photon at the detector and registration of an emission event in the control PC. This distribution, the instrument response function (IRF), has a width which is typically of the order of 100 ps, close to the timescale over which fluorescence from free NAD(P)H decays; its influence must therefore be taken into account. Most FLIM analysis software provides the functionality to deconvolve the effect of the IRF in the extraction of fluorescence decay parameters. The user must provide it with a measurement of the fluorescence decay of a sample with infinitely short fluorescence

lifetime, such as a scattering solution for single-photon measurements or second harmonic generation from a collagen sample for two-photon systems. Unfortunately both methods are technically inadequate as the measured signal in this calibration will not be at the wavelength of the fluorescence to be observed in the final experiment. One solution is to use a gold nanorod solution as a target, which has been shown to have a wide emission profile as well as the required negligible lifetime [80].

5.2 Laser Power Choice

Increased excitation power will increase the amount of fluorescence emitted, reducing the influence of Poisson noise on the measurements and therefore allowing fluorescence decay parameters to be determined with increased accuracy. Additionally, increased signal will allow the amount of pixel binning to be reduced (*see* Subheading 5.4), increasing the spatial resolution of the FLIM image. However, laser power cannot be increased without limit, as photodamage will begin to reduce the biological viability of the sample under study. 98% of the energy absorbed by free NAD(P)H is converted into heat [31]. Laser light can also directly induce the direct oxidation of NADH and the subsequent production of free radicals [81]. As such, NAD(P)H photobleaching is typically used as a proxy for the onset of photodamage [38, 82]. The ideal laser power can then be chosen as the highest value that does not cause photobleaching over the acquisition period. Laser repetition rate and pulse width will affect the incident photon flux and would thus likely be a factor in the level of photodamage at a given average excitation power [83]. However, these properties of the incident beam are difficult to control in a user-friendly microscope set-up. Instead, the scan speed and zoom level can be adjusted to limit laser dwell times. For extra confidence in the chosen settings, short bursts of data can be collected at the start and end of an image acquisition, and the data obtained can be binned into two single decays [38]. If the NAD(P)H fluorescence decay parameters are the same in the two decays, the user can be assured that the laser has not caused significant perturbation to the metabolism of the sample under study.

5.3 Acquisition Time Considerations

The maximum rate at which photons can be detected in a TCSPC experiment is limited to 1% of the incident laser repetition rate [84]. This rule of thumb reduces to insignificance the likelihood of more than one photon arriving in the time delay between two pulses. Based on a computational analysis of the signal-dependent uncertainty in the parameters obtained from biexponential fitting [38], we typically aim to acquire sufficient fluorescence to register at least 200 counts in the peak of each decay in the image. Acquiring at the maximum rate using an 80 MHz Ti:sapphire laser, assuming a homogeneous distribution of emission events across a 256×256 FLIM image and 256 detection channels, the peak value

is given by $I(t=0) \approx 40T$ for typical NAD(P)H fluorescence decay characteristics ($\alpha_{\text{bound}} = 0.2$, $\tau_{\text{free}} = 0.4$ ns, $\tau_{\text{bound}} = 2.5$ ns), where T is the acquisition time in minutes. Each image must therefore be acquired for around 5 min to detect sufficient signal. Unfortunately, the requirement to keep incident laser powers low enough to avoid photobleaching means fluorescence count rates can be significantly less than the maximum rate permitted by the equipment. Indeed, 5×10^4 counts per second are typical in our hands, 16 times below the pulse pileup threshold. Obtaining an image over 80 min is wholly impractical; living samples will move and perhaps alter their metabolism over this timescale. It is therefore not possible to acquire sufficient counts at each pixel for a reliable analysis based on increased acquisition times alone, and signal must be increased by other means, notably spatial binning (*see* Subheading 5.4). Meanwhile, a practical acquisition time should be chosen. This should be long enough to allow the spatial resolution of relevant features of the sample that may be analyzed separately, for example, subcellular organelles such as the nucleus or mitochondria, or different cell types in a complex tissue. The acquisition time must also be short enough such that the biological integrity of the sample is maintained, both during the collection of an individual image while the sample is exposed to the laser and over the longer time period that the sample remains at the microscope while repeats are being taken in different regions of the preparation. Acquisition times of between 1 and 5 min are typical [38, 43, 85].

5.4 Spatial Binning

The impracticality of acquiring sufficient counts at an individual pixel for reliable fitting to NAD(P)H FLIM data means that spatial binning of the collected photons is required. This procedure, typically performed by the FLIM analysis software, will combine detection events from surrounding pixels at each location of the image in order to increase signal levels and decrease the relative impact of noise on the uncertainty in the reported parameters. Typically, we will increase the binning level until 200 counts are in the peak channel of the lowest intensity pixel of interest [38]. The extent to which the data are binned is frequently reported as the "binning factor". A binning factor of 1 indicates that the data from the eight surrounding pixels are added to that in each pixel, resulting in a ninefold increase in signal, at the expense of spatial resolution. As Poisson statistics apply, this can be expected to decrease the uncertainty in the fit parameters by a factor of approximately $1/\sqrt{9}$, or to 33% of their initial value. Increasing the binning further to include the 5×5 region surrounding each pixel, a binning factor of 2, will reduce the uncertainties to $1/\sqrt{25}$ of their single-pixel value, a further improvement of 13% over the first level of binning. The gains in parameter precision by additional binning steps then diminish significantly to 6, 3, and 2% by increasing to binning factors of 3, 4 and 5, respectively. There is therefore little benefit

to applying a binning factor above 2 should the threshold of 200 peak photon counts not be reached, with increases merely reducing the spatial resolution of the FLIM image. Increased certainty in the parameters determined can instead be achieved by taking averages across regions of interest of single measurements and across repeat measurements of identical conditions.

5.5 Averaging and Repeats

A FLIM image of a homogeneous solution of a fluorescent dye with a single lifetime will still report pixel-to-pixel variations in the measured fluorescence decay rate due to the Poisson noise inherent in the TCSPC method. Interpreting these noise-induced variations as physical phenomena can lead to extreme conclusions regarding intracellular heterogeneity of biological properties, such as in the use of temperature-sensitive lifetime probes [86]. The effect of this noise can be eliminated by taking average parameter values across pixels. This can be performed by exporting lifetime images from the lifetime analysis package for analysis in ImageJ (National Institutes of Health, Bethesda, USA). Performing this procedure implies the underlying assumption that the local environment of the fluorescent probe at each pixel is the same. Inside the complex environment of the cell, this assumption does not hold. In the case of NAD(P)H, three distinct microenvironments can be resolved from the fluorescence intensity image itself; the cytosol, the darker nucleus and the brighter mitochondria. As such, we typically extract cytosolic-, nuclear-, and mitochondrial-specific NAD(P)H fluorescence decay parameters from each image we acquire. The mean NAD(P)H lifetime is typically shorter in the nucleus than the rest of the cell. We have shown this to be consistent with equal NADH levels but decreased NADPH levels in the nucleus of HEK293 cells [38]. Despite differing total NAD(P)H concentrations, the mitochondria and cytosol typically display similar fluorescence decay parameters, likely reflecting the high interconnectivity of the redox states of the NAD and NADP pools within and between these two compartments [14, 16, 87]. As with all measurements on living systems, these single region-of-interest measurements are not sufficient to support hypotheses due to the need to account for biological variability. Replicate measurements on independent samples must therefore be taken. We typically observe standard deviations in the fluorescence decay parameters of around 3%. Based on statistical power analysis, around eight biological replicates would be required for a difference in lifetime of the order of the IRF width (~100 ps) to be reported with a P value of less than 5% in 90% of experiments. This reduces to three repeats for reporting statistically significant differences of the order of two IRF widths (~200 ps). The oft-quoted "$n = 3$" should therefore be modestly exceeded in order to make inferences that exploit the full time resolution of the TCSPC method.

5.6 Experimental Design Summary

The above considerations point to the initial steps of designing an NAD(P)H FLIM experimental protocol outlined below. The procedure can then be tailored to the biological model under investigation as required.

- Choose an acquisition time of between 1 and 5 min. Longer times are preferable, but the preparation must remain still and viable over the imaging time course.
- Choose the highest laser power that doesn't cause photobleaching over the acquisition time.
- Aim to perform more than three independent biological repeats of each condition to ensure that the smallest technically feasible differences in lifetime can be reliably determined.
- In the analysis stage, increase binning to a maximum of two. Lower binning is better, but aim for the dimmest pixels of interest to contain at least 200 photon counts in the peak channel.
- Separately extract mean NAD(P)H fluorescence decay parameters from the cytosol, nucleus and mitochondria. Organelle-specific dyes (e.g., TMRM) could be used to facilitate this process if the compartments cannot be resolved by the NAD(P)H intensity alone.

6 Conclusions

The original applications of NAD(P)H FLIM in live cells and tissues were content to identify changes in the fluorescence decay characteristics of the signal as indicators of unspecified metabolic alterations. The contemporary ambition is to derive true biological understanding from intracellular NAD(P)H lifetime data by translating fluorescence decay parameters into underlying biochemical and physiological states [14]. With the Warburg effect in mind, the large number of demonstrations of altered autofluorescence lifetimes in cancer relative to healthy tissue prompted developments in this area to focus on the effect of alterations in the balance between aerobic and anaerobic energy metabolism on the NAD(P)H fluorescence decay characteristics [43].

The application of pharmacological perturbations to alter the balance in ATP production between OXPHOS and glycolysis in a given cell type has been observed to correlate with changes in the proportion of bound NAD(P)H species, α_{bound} [42]. However, these correlations may not hold for the comparison of metabolic states in two independent cells or tissues [88]. Indeed, we have previously shown that a genetically modified HEK293 cell line with increased reliance on OXPHOS exhibits the same NAD(P)H

fluorescence decay characteristics as the more glycolytic wild-type cells [38]. In addition, a differently modified HEK293 cell line with identical aerobic/anaerobic respiratory balance showed vastly different NAD(P)H fluorescence decay characteristics. The modifications in question were under- and overexpression of NAD kinase, the master regulator of NADPH levels inside the cell. These results demonstrated that the NADPH/NADH balance plays a crucial role in determining the time-resolved NAD(P)H fluorescence characteristics from a given tissue and that correlations between the bound weighting and the means of ATP production may only apply to acute, externally induced changes and may not be an overall feature of the metabolic phenotype.

Deciphering the biochemical and physiological meaning behind alterations in the subcellular photophysics of NADH and NADPH remains an active and developing area of research [89]. Nevertheless, exploiting the high sensitivity of fluorescence lifetime measurements to metabolically induced alterations in the local environment of intrinsic fluorophores continues to be applied as a robust, precise, and minimally invasive means to detect changes in the metabolism of living cells and tissues. Medical devices based on these phenomena for monitoring cancer progression and delineating the boundaries of surgically accessible tumors are being constructed [90–92]. The extent to which these developments succeed clinically is reliant on integrating fundamental scientific understanding across scales and disciplines, from excited state decay processes and mitochondrial redox dynamics to statistical analysis of single-photon counting data and metabolic coupling between cells in a complex tissue, positioning NAD(P)H FLIM as a clear demonstration of the integrative frontiers of modern biomedical research.

References

1. Koppenol WH, Bounds PL, Dang CV (2011) Otto Warburg's contributions to current concepts of cancer metabolism. Nat Rev Cancer 11:325
2. Kim J, Dang CV (2006) Cancer's molecular sweet tooth and the Warburg effect. Cancer Res 66:8927–8930
3. Vander Heiden MG, Cantley LC, Thompson CB (2009) Understanding the warburg effect: the metabolic requirements of cell proliferation. Science 324(80):1029–1033
4. Vander Heiden MG, DeBerardinis RJ (2017) Understanding the intersections between metabolism and cancer biology. Cell 168:657–669
5. Osellame LD, Blacker TS, Duchen MR (2012) Cellular and molecular mechanisms of mitochondrial function. Best Pract Res Clin Endocrinol Metab 26:711–723
6. Duchen MR, Szabadkai G (2010) Roles of mitochondria in human disease. Essays Biochem 47:115–137
7. Gambhir SS (2002) Molecular imaging of cancer with positron emission tomography. Nat Rev Cancer 2:683
8. Wallace DC (2012) Mitochondria and cancer. Nat Rev Cancer 12:685–698
9. Weinhouse S (1976) The Warburg hypothesis fifty years later. Zeitschrift für Krebsforsch und Klin Onkol 87:115–126
10. Fantin VR, St-Pierre J, Leder P (2006) Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology,

and tumor maintenance. Cancer Cell 9:425–434

11. Moreno-Sánchez R, Rodríguez-Enríquez S, Marín-Hernández A, Saavedra E (2007) Energy metabolism in tumor cells. FEBS J 274:1393–1418
12. Robertson-Tessi M, Gillies RJ, Gatenby RA, Anderson ARA (2015) Impact of metabolic heterogeneity on tumor growth, invasion, and treatment outcomes. Cancer Res 75:1567–1579
13. Granger A, Mott R, Emambokus N (2016) Hacking cancer metabolism. Cell Metab 24:643–644
14. Blacker TS, Duchen MR (2016) Investigating mitochondrial redox state using NADH and NADPH autofluorescence. Free Radic Biol Med 100
15. Murphy MP (2009) How mitochondria produce reactive oxygen species. Biochem J 417:1–13
16. Ying W (2008) NAD+/NADH and NADP+/NADPH in cellular functions and cell death: regulation and biological consequences. Antioxid Redox Signal 10:179–206
17. Nickel AG, von Hardenberg A, Hohl M et al (2015) Reversal of mitochondrial transhydrogenase causes oxidative stress in heart failure. Cell Metab 22:472–484
18. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB (2008) The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. Cell Metab 7:11–20
19. Sosa V, Moliné T, Somoza R et al (2013) Oxidative stress and cancer: an overview. Ageing Res Rev 12:376–390
20. Tomiyama A, Serizawa S, Tachibana K et al (2006) Critical role for mitochondrial oxidative phosphorylation in the activation of tumor suppressors Bax and Bak. J Natl Cancer Inst 98:1462–1473
21. Kroemer G, Pouyssegur J (2008) Tumor cell metabolism: cancer's Achilles' heel. Cancer Cell 13:472–482
22. Hsu PP, Sabatini DM (2008) Cancer cell metabolism: Warburg and beyond. Cell 134:703–707
23. Shlomi T, Benyamini T, Gottlieb E et al (2011) Genome-scale metabolic modeling elucidates the role of proliferative adaptation in causing the Warburg effect. PLoS Comput Biol 7: e1002018
24. Croce CM (2008) Oncogenes and cancer. N Engl J Med 358:502–511
25. Levine AJ, Puzio-Kuter AM (2010) The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science 330(80):1340–1344
26. Locasale JW (2013) Serine, glycine and one-carbon units: cancer metabolism in full circle. Nat Rev Cancer 13:572
27. Maddocks ODK, Berkers CR, Mason SM et al (2013) Serine starvation induces stress and p53-dependent metabolic remodelling in cancer cells. Nature 493:542–546
28. Dang CV (2010) Rethinking the Warburg effect with Myc micromanaging glutamine metabolism. Cancer Res 70:859–862
29. Dang L, White DW, Gross S et al (2009) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 462:739
30. De Ruyck JJ, Famerée M, Wouters J et al (2007) Towards the understanding of the absorption spectra of NAD(P)H/NAD(P)+ as a common indicator of dehydrogenase enzymatic activity. Chem Phys Lett 450:119–122
31. Blacker TS, Marsh RJ, Duchen MR, Bain AJ (2013) Activated barrier crossing dynamics in the non-radiative decay of NADH and NADPH. Chem Phys 422:184–194
32. Patterson GH, Knobel SM, Arkhammar P et al (2000) Separation of the glucose-stimulated cytoplasmic and mitochondrial NAD(P)H responses in pancreatic islet beta cells. Proc Natl Acad Sci U S A 97:5203–5207
33. Mayevsky A, Chance B (2007) Oxidation-reduction states of NADH in vivo: from animals to clinical use. Mitochondrion 7:330–339
34. Chance B, Cohen P, Jobsis F, Schoener B (1962) Intracellular oxidation-reduction states in vivo. Science 137(80):499–508
35. Duchen MR, Surin A, Jacobson J (2003) Imaging mitochondrial function in intact cells. Methods Enzymol 361:353–389
36. Berezin MY, Achilefu S (2010) Fluorescence lifetime measurements and biological imaging. Chem Rev 110:2641–2684
37. Meijers R, Morris RJ, Adolph HW et al (2001) On the enzymatic activation of NADH. J Biol Chem 276:9316–9321
38. Blacker TS, Mann ZF, Gale JE et al (2014) Separating NADH and NADPH fluorescence in live cells and tissues using FLIM. Nat Commun 5:3936
39. Schneckenburger H, Koenig K (1992) Fluorescence decay kinetics and imaging of NAD(P)H and flavins as metabolic indicators. Opt Eng 31:1447–1451
40. Paul RJ, Schneckenburger H (1996) Oxygen concentration and the oxidation-reduction state of yeast: determination of free/bound

NADH and flavins by time-resolved spectroscopy. Naturwissenschaften 83:32–35

41. Pradhan A, Pal P, Durocher G et al (1995) Steady state and time-resolved fluorescence properties of metastatic and non-metastatic malignant cells from different species. J Photochem Photobiol B Biol 31:101–112
42. Bird DK, Yan L, Vrotsos KM et al (2005) Metabolic mapping of MCF10A human breast cells via multiphoton fluorescence lifetime imaging of the coenzyme NADH. Cancer Res 65:8766–8773
43. Skala MC, Riching KM, Bird DK et al (2007) In vivo multiphoton fluorescence lifetime imaging of protein-bound and free nicotinamide adenine dinucleotide in normal and precancerous epithelia. J Biomed Opt 12:24014
44. McGinty J, Galletly NP, Dunsby C et al (2010) Wide-field fluorescence lifetime imaging of cancer. Biomed Opt Express 1:627–640
45. Adur J, Pelegati VB, Bianchi M, et al (2013) Multimodal nonlinear optical microscopy used to discriminate human colon cancer. In: Multiphoton microscopy in the biomedical sciences XIII. International Society for Optics and Photonics, p 85881J
46. Rueck AC, Hauser C, Mosch S, Kalinina S (2014) Spectrally resolved fluorescence lifetime imaging to investigate cell metabolism in malignant and nonmalignant oral mucosa cells. J Biomed Opt 19:96005
47. Wang Y, Song C, Wang M et al (2016) Rapid, label-free, and highly sensitive detection of cervical cancer with fluorescence lifetime imaging microscopy. IEEE J Sel Top Quantum Electron 22:228–234
48. Awasthi K, Moriya D, Nakabayashi T et al (2016) Sensitive detection of intracellular environment of normal and cancer cells by autofluorescence lifetime imaging. J Photochem Photobiol B Biol 165:256–265
49. Pastore MN, Studier H, Bonder CS, Roberts MS (2017) Non-invasive metabolic imaging of melanoma progression. Exp Dermatol 26:607–614
50. Phillips D, Drake RC, O'Connor D V, Christensen RL (1985) Time correlated single-photon counting (TCSPC) using laser excitation
51. Bain AJ (2015) Multiphoton Processes. In: Photonics. John Wiley & Sons, Inc., Hoboken, NJ, pp 279–320
52. Moulton PF (1986) Spectroscopic and laser characteristics of Ti:Al_2O_3. JOSAB 3:125–133
53. Helmchen F, Denk W (2005) Deep tissue two-photon microscopy. Nat Methods 2:932–940
54. Schneckenburger H, Wagner M, Weber P et al (2004) Autofluorescence lifetime imaging of cultivated cells using a UV picosecond laser diode. J Fluoresc 14:649–654
55. Webb RH (1996) Confocal optical microscopy. Reports Prog Phys 59:427
56. Le Grand Y, Leray A, Guilbert T, Odin C (2008) Non-descanned versus descanned epifluorescence collection in two-photon microscopy: experiments and Monte Carlo simulations. Opt Commun 281:5480–5486
57. Becker W, Bergmann A, Hink MA et al (2004) Fluorescence lifetime imaging by time-correlated single-photon counting. Microsc Res Tech 63:58–66
58. Becker W, Su B, Holub O (2011) FLIM and FCS detection in laser-scanning microscopes: increased efficiency by GaAsP hybrid detectors. Microsc Res Tech 74:804–811
59. Becker W, Bergmann A, Biskup C (2007) Multispectral fluorescence lifetime imaging by TCSPC. Microsc Res Tech 70:403–409
60. Zipfel WR, Williams RM, Christie R et al (2003) Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generation. Proc Natl Acad Sci U S A 100:7075–7080
61. Moré JJ (1978) The Levenberg-Marquardt algorithm: implementation and theory. In: Numerical analysis. Springer, New York, pp 105–116
62. Wahl P (1979) Analysis of fluorescence anisotropy decays by a least square method. Biophys Chem 10:91–104
63. Moger J, Gribbon P, Sewing A, Winlove CP (2006) The application of fluorescence lifetime readouts in high-throughput screening. J Biomol Screen 11:765–772
64. Tiede LM, Nichols MG (2006) Photobleaching of reduced nicotinamide adenine dinucleotide and the development of highly fluorescent lesions in rat basophilic leukemia cells during multiphoton microscopy. Photochem Photobiol 82:656–664
65. Enderlein J, Erdmann R (1997) Fast fitting of multi-exponential decay curves. Opt Commun 134:371–378
66. Laurence TA, Chromy BA (2010) Efficient maximum likelihood estimator fitting of histograms. Nat Methods 7:338
67. Turton DA, Reid GD, Beddard GS (2003) Accurate analysis of fluorescence decays from single molecules in photon counting experiments. Anal Chem 75:4182–4187
68. Maus M, Cotlet M, Hofkens J et al (2001) An experimental comparison of the maximum likelihood estimation and nonlinear least-squares

fluorescence lifetime analysis of single molecules. Anal Chem 73:2078–2086

69. Bajzer Ž, Therneau TM, Sharp JC, Prendergast FG (1991) Maximum likelihood method for the analysis of time-resolved fluorescence decay curves. Eur Biophys J 20:247–262
70. Warren SC, Margineanu A, Alibhai D et al (2013) Rapid global fitting of large fluorescence lifetime imaging microscopy datasets. PLoS One 8:e70687
71. Livesey AK, Brochon JC (1987) Analyzing the distribution of decay constants in pulse-fluorimetry using the maximum entropy method. Biophys J 52:693–706
72. Brochon J-C (1994) Maximum entropy method of data analysis in time-resolved spectroscopy. In: Methods in enzymology. Elsevier, Amsterdam, pp 262–311
73. Rowley MI, Coolen ACC, Vojnovic B, Barber PR (2016) Robust Bayesian fluorescence lifetime estimation, decay model selection and instrument response determination for low-intensity FLIM imaging. PLoS One 11: e0158404
74. Rowley MI, Barber PR, Coolen ACC, Vojnovic B (2011) Bayesian analysis of fluorescence lifetime imaging data. In: Multiphoton microscopy in the biomedical sciences XI. International Society for Optics and Photonics, p 790325
75. O'Shea P (2012) Future medicine shaped by an interdisciplinary new biology. Lancet 379:1544–1550
76. Digman MA, Caiolfa VR, Zamai M, Gratton E (2008) The phasor approach to fluorescence lifetime imaging analysis. Biophys J 94: L14–L16
77. Plotegher N, Stringari C, Jahid S et al (2015) NADH fluorescence lifetime is an endogenous reporter of α-synuclein aggregation in live cells. FASEB J 29:2484–2494
78. Stringari C, Cinquin A, Cinquin O et al (2011) Phasor approach to fluorescence lifetime microscopy distinguishes different metabolic states of germ cells in a live tissue. Proc Natl Acad Sci 108:13582–13587
79. Stringari C, Edwards RA, Pate KT et al (2012) Metabolic trajectory of cellular differentiation in small intestine by Phasor fluorescence lifetime microscopy of NADH. Sci Rep 2
80. Talbot CB, Patalay R, Munro I et al (2011) Application of ultrafast gold luminescence to measuring the instrument response function for multispectral multiphoton fluorescence lifetime imaging. Opt Express 19:13848–13861
81. Czochralska B, Lindqvist L (1983) Diphotonic one-electron oxidation of NADH on laser excitation at 353 nm. Chem Phys Lett 101:297–299
82. Walsh AJ, Cook RS, Sanders ME et al (2014) Quantitative optical imaging of primary tumor organoid metabolism predicts drug response in breast cancer. Cancer Res 74:5184–5194
83. Kawano H, Nabekawa Y, Suda A et al (2003) Attenuation of photobleaching in two-photon excitation fluorescence from green fluorescent protein with shaped excitation pulses. Biochem Biophys Res Commun 311:592–596
84. Salthammer T (1992) Numerical simulation of pile-up distorted time-correlated single photon counting (TCSPC) data. J Fluoresc 2:23–27
85. Tosatto A, Sommaggio R, Kummerow C et al (2016) The mitochondrial calcium uniporter regulates breast cancer progression via HIF-1α. EMBO Mol Med 8:569–585
86. Baffou G, Rigneault H, Marguet D, Jullien L (2014) A critique of methods for temperature imaging in single cells. Nat Methods 11:899
87. Cracan V, Titov DV, Shen H et al (2017) A genetically encoded tool for manipulation of NADP+/NADPH in living cells. Nat Chem Biol 13:1088
88. Guo H-W, Yu J-S, Hsu S-H et al (2015) Correlation of NADH fluorescence lifetime and oxidative phosphorylation metabolism in the osteogenic differentiation of human mesenchymal stem cell. J Biomed Opt 20:17004
89. Schaefer PM, Hilpert D, Niederschweiberer M et al (2017) Mitochondrial matrix pH as a decisive factor in neurometabolic imaging. Neurophotonics 4:45004
90. Sun Y, Phipps J, Elson DS et al (2009) Fluorescence lifetime imaging microscopy: in vivo application to diagnosis of oral carcinoma. Opt Lett 34:2081–2083
91. De Beule PA, Dunsby C, Galletly NP et al (2007) A hyperspectral fluorescence lifetime probe for skin cancer diagnosis. Rev Sci Instrum 78:123101
92. Butte PV, Pikul BK, Hever A et al (2005) Diagnosis of meningioma by time-resolved fluorescence spectroscopy. J Biomed Opt 10:64026–64029

Chapter 20

Overview of Characterizing Cancer Glycans with Lectin-Based Analytical Methods

Amanda J. Pearson and Elyssia S. Gallagher

Abstract

Glycosylation is a post-translational modification that is often altered in disease development and progression, including cancer. In cancerous patients, the abnormal expression of glycosylation enzymes leads to aberrant glycosylation, which has been linked to malignant tissues. Due to aberrant glycosylation, the presence of specific glycans can be used as biomarkers for identifying the type and stage of cancer. Glycan structures are heterogeneous, with different protein glycoforms having different functional activities. Lectins are an important tool in glycan analysis due to their specificity in binding to unique glycan linkages and monosaccharide units, which allows for the identification of unique glycan structural properties. In this review, we will discuss the use of lectins in microarrays and chromatography for characterizing glycan structures.

Key words Lectin, Cancer, Glycan, Glycosylation, Microarray, Affinity chromatography

1 Introduction

1.1 Overview

Glycomics, the study of glycans, is a relatively new area of research, especially compared with other fields characterizing biomolecules, such as proteomics and genomics. Much of this field is dedicated to developing techniques for analyzing glycan structures and identifying how the presence of glycans affects the structure and function of the lipids and proteins to which they are covalently linked. Because glycomics is still growing, research dedicated to the involvement of glycosylation in disease is a small fraction of the field [1].

Glycans are linear or branched chains of monosaccharides that are added to other biomolecules, including proteins and lipids. It is approximated that 50–70% of all eukaryotic proteins are glycosylated as they are processed in the endoplasmic reticulum (ER) and Golgi apparatus. Changes in protein glycosylation affect protein structure and function [2], which can result in biological misfunction and disease progression, as observed in autoimmunity [3],

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_20,

cancer [4], and hereditary disease [5]. Herein, we will examine the use of lectins for the structural determination of glycans [6]. Lectins, proteins that bind to glycans, recognize specific monosaccharides, linkages, and/or carbohydrate stereochemistries, making them unique tools for structural analyses. While this review will focus on the use of lectins, additional reviews describing glycan structural characterizations by chromatography [7, 8], capillary electrophoresis [9, 10], and mass spectrometry [11, 12] are present in the literature.

1.2 Lectins

Lectins are proteins that specifically bind to carbohydrates and are found throughout nature in animals, plants, and microorganisms such as fungi, bacteria, viruses, and protozoa [13]. Numerous lectins have been purified, but it is assumed that many lectins have yet to be discovered [14]. Lectins bind to glycans with specific structural motifs (monosaccharides, linkages, and stereochemistries) with binding affinities (K_d) in the μM range [15, 16]. To combat these weak binding affinities, researchers have started developing synthetic lectins, which are often boron-containing peptides and aptamers, with a similar selectivity as natural lectins to unique glycan structural characteristics but with binding affinities in the nM–pM range [17–19].

1.3 Glycan Synthesis

Glycans are classified as either *O*- or *N*-linked and are added to proteins during and after translation. The main differences between the two are in structure and linkage. *N*-linked glycans are attached to proteins via asparagine residues present in conserved NXS/T sequences, where X is any amino acid besides P, while *O*-linked glycans are attached via threonine or serine residues [20]. *N*-linked glycans are usually highly branched, whereas *O*-linked glycans tend to have minimal branching, being mostly linear. Glycan precursors, monosaccharide subunits that are the core of both *N*- and *O*-linked glycans, are first formed in the cytosol and are then transferred to the endoplasmic reticulum (ER) and Golgi apparatus for branch extension and terminal glycosylation [21–23]. Following glycosylation and folding, glycoproteins are transported to their functional locations, including cell surfaces [24]. The terminal glycosylation and branching patterns are important factors in determining glycans' biological roles and involvement in disease. Terminal sugars on plasma membrane proteins interact with other proteins and cells due to proximity and thus can have significant impacts on cellular interactions [25, 26].

In comparison to DNA replication and protein translation, glycosylation is a non-template-driven process. Thus, the structures that are formed are dependent on the glycosylation enzymes (e.g., glycosyltransferases and glycosidases) that are present in the ER and Golgi at the time of protein translation, the specificity of the enzymes for the developing glycan, and the availability of

carbohydrate substrates. The resulting glycans are heterogeneous and often contain different branching patterns or terminal sugars [22, 27]. The heterogeneity adds to the challenge of glycan analyses since each protein can be present as multiple glycoforms, with different glycoforms having different functionalities [28]. Ideal analysis methods sample the heterogeneous glycan structures to monitor changes in the glycan population during disease development and progression.

1.4 Glycans in Cancer

Though glycans have multiple roles in cancer, these roles have not been fully elucidated. Glycosyltransferases and glycosidases, enzymes that add and remove monosaccharides to developing glycans, respectively, are often suppressed or overexpressed in cancer, resulting in aberrant or altered glycosylation patterns [29–31]. Aberrant glycosylation occurs in both *N*- and *O*-linked glycans [32, 33]; however, changes in *N*-linked glycans are easier to characterize. *N*-linked glycans are more complex, containing more branching and diversity in monosaccharide subunits; thus, *N*-linked glycans have more varied structural motifs providing a greater diversity of structures for lectin analysis. Alternatively, *O*-linked glycans have relatively simple structures, so if a lectin is not able to bind to the few monosaccharides, linkages, or stereochemistries, there is a lower probability that the structural components will be identified. In addition, there are enzymes that are better equipped for removing *N*-linked glycans from tissue samples, while *O*-linked glycans do not have dependable enzymes for their removal [34]. Peptide-N-glycosidase F (PNGaseF) removes most *N*-linked glycans; however, there are cases in which complete deglycosylation can be difficult because sites are resistant to enzymatic cleavage, requiring the glycoprotein to be incubated with the enzyme for long periods of time [35]. Currently, *O*-linked glycans are removed by chemical methods (hydrazine) or enzymes (pronase), but neither is as effective as PNGaseF, making it more difficult to characterize the full *O*-linked glycan signature [36, 37].

Glycan alterations have been observed in regard to tissue type, cancer stage, and following exposure to radiotherapy [38–44]. In malignant tissues, common terminal glycosylation patterns have been observed due to overexpression of specific glycosylation enzymes [30]. Common motifs in cancer include sialyl Lewis x (sLe^x), sialyl Lewis a (sLe^a), and Lewis y (Le^y), as seen in Fig. 1 [33, 41, 45, 46]. For example, Drahos' group observed significant changes in the abundance of glycans present during and after radiotherapy, with some glycoforms increasing by fivefold and other glycoforms decreasing by fivefold. The Drahos' group showed that these changes lasted for over a month following completion of radiotherapy treatment [47]. There is interest in using the presence and abundance of certain glycoforms as biomarkers in screening and staging cancers.

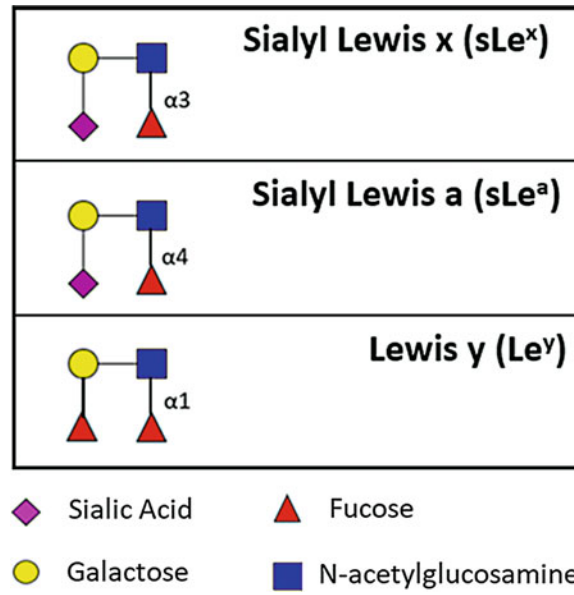


Fig. 1 Structures of terminal sugar moieties that are commonly overexpressed in malignant cancers. While the monosaccharide structures of sLe^x and sLe^a contain the same monosaccharides, the linkage between *N*-acetylglucosamine and fucose is different. In addition, Le^y has a fucose where sialic acid is present in the other two glycans

2 Structural Analytical Methods

2.1 Lectin Microarrays

2.1.1 Overview of Microarray Analysis

Microarrays are simple, high-throughput techniques that maximize the number of simultaneously analyzed samples or the number of simultaneous measurements on a single sample. While microarrays have the advantage of analyzing many samples, a disadvantage comes from the long times needed for sample incubation before detection. Microarrays are prevalent for analyses of DNA and proteins [48–50]; yet, since the technique was first published in 2005, lectin microarrays have become more widely used tools for analysis of glycan structural motifs [51–53]. An overview of a lectin microarray is shown in Fig. 2. In general, a microarray analysis involves spotting a surface with known analytes at different positions and then adding a second, unknown analyte to the surface to initiate interactions. Following incubation, the positions at which interactions occur are detected, and structural information can be inferred for the unknown analyte based on the binding affinities of the known analytes spotted on the surface. The binding interactions are generally detected by fluorescence, which gives an intensity output that correlates to binding between lectins and glycans. By comparing the relative fluorescence of different samples, the relative quantity of a particular structure can be compared to the

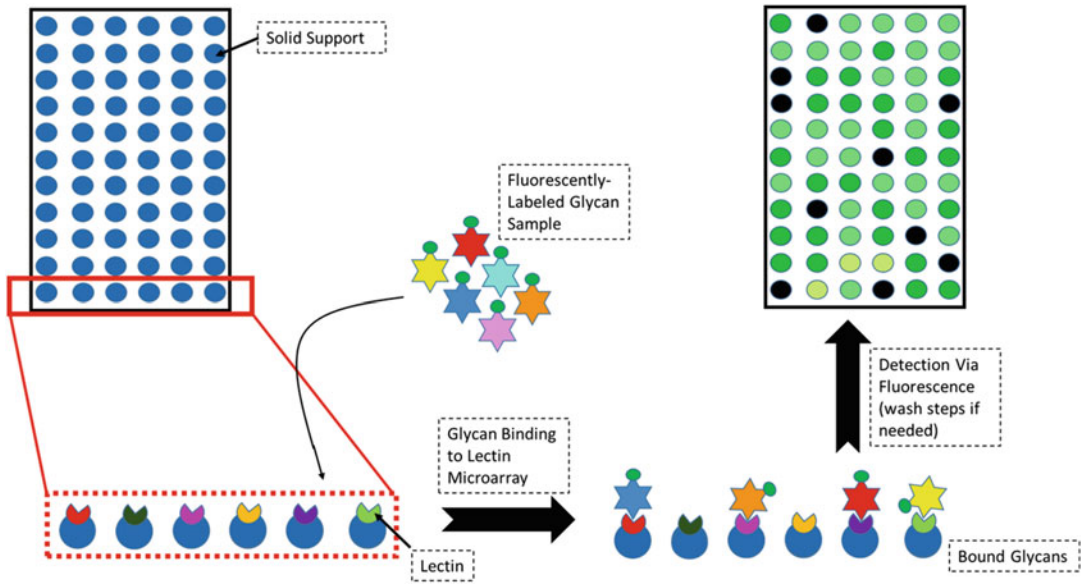


Fig. 2 Overview of a lectin microarray. Different lectins are covalently linked to designated positions on a solid support. A fluorescently labeled sample of glycans, glycoproteins, or cells is incubated with the microarray, allowing for the glycans to bind to the surface-attached lectins. The binding interaction is detected via fluorescence, giving an intensity for each position of the microarray where glycans interact with lectins. The intensity for each position is shown here in shades of green, with the brightest green representing positions with the most glycans binding and the duller greens representing positions with fewer bound glycans. Black positions represent lectins that did not interact with any component in the sample

relative quantity of other structures [54]. The data is usually more qualitative than quantitative, but there are a variety of ways to interpret the collected data. For example, if performing a comparative analysis, the intensity of the fluorescent signal illustrates differences in the glycans that are present in the two samples [55]. This is useful for the study of cancer, because one can compare samples from different tissues, stages, or donors and observe differences in the structure or quantity of glycans.

2.1.2 Microarray Preparation

Microarrays are versatile methods that vary in the type and number of lectins spotted onto the array. The first step in developing a microarray is choosing a solid support. Common supports are polystyrene well plates [56], polymer gel slides [57], and epoxy slides [58], with multiple commercially available supports available. The size of the support varies and depends on the number of lectins and the method for spotting them. If a microarrayer, or microarray printer (automated instrument used to spot lectins), is used, the size of the chip must accommodate the printer requirements. For example, an average plate is 25 × 75 mm or a 384-well plate for a NanoArrayIt printer [59]. Finally, the lectins are spotted onto the array. The size and distance between spots vary but are important factors to maximize the number of spots on the array while

ensuring that the lectins can be spatially resolved during detection [52]. Currently, it is common for the space between each lectin to be equivalent to or greater than the diameter of each lectin spot. A common method for spotting lectins is using a microarrayer; these printers draw lectin solutions from a reservoir and spot them in a predefined pattern [50]. Microarray printers are preferred, because they minimize the diameter of each spot and create more reproducible arrays. Despite these advantages, lectins can be spotted manually to reduce the cost of preparation. While there are various ways to add the lectins to the array, the goal is simply for the lectin to covalently bind an active group on the surface (e.g., epoxy, aldehyde, or amino) [60]. Once spotted, the microarray is ready for use or storage. Storage requirements depend upon the array format, but in general, arrays are stored at −20 °C and can be stored for approximately 6 months. It is best to consider the materials used to determine the activity of the array before using an array after a lengthy storage time.

2.1.3 Types of Lectins

Natural Lectins

Natural lectins are the most abundant form of lectins currently in use for lectin microarrays. They are convenient because they can be purified or purchased commercially. Lectins are a useful structural analysis tool, because they bind to various monosaccharides, linkages, and stereochemistries within both *N*- and *O*-linked glycans. *O*-linked glycans are much less studied in comparison to *N*-linked due to the lack of dependable enzymes for their removal, so the ability of lectins to simultaneously analyze both *N*- and *O*- linked glycans is an advantage [9, 61]. While this review focuses on the use of lectins in the structural analysis of glycans, other reviews focus on identification and isolation of natural lectins [15, 16, 62, 63].

Synthetic Lectins

Lectins are powerful tools for structural analyses of glycans, but their use is often limited by their weak binding affinities, K_d ~μM [15, 16]. Several groups have developed synthetic lectins with stronger binding affinities compared to natural lectins. Lectins bind to specific linkages and monosaccharides, and developing a synthetic lectin that can be selective and recognize these structural differences has been challenging [64]. The most successful synthetic lectins use boron moieties, including boron aptamers and peptide borono lectins (PBLs), which have binding affinities in the nM to pM range [18, 19]. Both the aptamers and PBLs are synthesized through similar methods, with incorporation of boronic acid into a modified nucleotide or peptide, respectively. When designing synthetic aptamers or peptides, an extensive number of lectins are made and then screened to determine which compounds are most selective for a particular structural motif. Due to the stronger binding affinity of nucleotides with glycan hydroxyls, boron aptamers have been more regularly used than peptides. The binding

interaction between boron and carbohydrates is the result of boronic acid binding to diols on carbohydrates and forming boronate esters. The stereochemistry of the carbohydrates affects the binding affinity of the boronate, with the highest affinity for *cis*-1,2-diols [65]. The development of synthetic lectins results in greater coverage and diversity in the number and types of glycans that can be identified, because synthetic lectins can be synthesized to target glycan structural motifs that are currently unable to be identified by natural lectins [66]. Other synthetic lectins have been developed outside of the boron-based lectins, such as bio-inspired lectins, aptamers, and peptides, which are reviewed elsewhere [17, 64–68].

2.1.4 Sampling Methods

One of the advantages of using microarrays is the minimal sample preparation for many analytes, including crude cells [69], glycoproteins [70], blood [71], or urine [55]. To minimize matrix effects, blood and urine are often sonicated, vortexed, and/or centrifuged [55]. Prior to analysis, samples also require labeling with a fluorescent tag, such as 2-aminobenzoic acid (2-AA) or 2-aminobenzamide (2-AB), that binds via reductive amination at the reducing end of the glycan, to aid in analyte detection [72].

Tissues can also be analyzed with microarrays but require more extensive sample preparation. Laser microdissection (LMD) allows for analysis of precise tissue sections, rather than heterogeneous tissue [73]. For example, LMD can be used to separate tissue sections, allowing for analysis of a prostate tumor, rather than prostate tissue containing a mixture of cancerous and normal cells [74]. LMD is ideal for lectin microarrays because the arrays require as little as 10 μg of sample. To sample human tissues, biopsies are collected from patients [75] followed by LMD to isolate specific cells. In preparation for LMD, the tissue sample is positioned on a glass slide. Then, a thermolabile polymer is placed above the tissue section and melted with an infrared laser, surrounding the cells of interest, which bind to the polymer and can be removed. This effectively dissects the cells of interest from excess tissue [76]. After LMD, protein extraction is used to collect the glycan or glycoprotein of interest [77]. The method of extraction is dependent on the analyte being extracted, but generally, the tissue is solubilized, sonicated, and then centrifuged to purify the glycan or glycoprotein. Once extraction is complete, the sample is labeled with a fluorescent tag prior to microarray analysis.

2.1.5 Detection Methods

Once the glycan sample has been added to the microarray, the interaction must be detected to probe the binding interactions between lectins and glycans to characterize glycan structures. For most assays, fluorescence intensity is measured, giving an estimation of the amount of a structural motif in a glycan sample. Early microarrays detected fluorescence with plate readers

[51, 78]. However, evanescent-field fluorescence (EFF), a method developed by Hirabayashi and coworkers, is the current standard for lectin microarray detection [53]. In EFF, the microarray is set on a glass slide, above a charge-coupled device (CCD) camera. An excitation light is focused into the glass slide, which causes internal reflection, creating an evanescent field. This field penetrates only ~100–200 nm above the surface, enabling detection of material at the surface with weak binding affinities. EFF allows for detection without a washing step, in which weak interactions are removed; therefore these weaker interactions can be observed with EFF. The wash steps are able to be eliminated due to the equilibrium created between the lectin and glycan and the real-time imaging as the result of internal reflection [79]. The emitted fluorescence signals are then detected from below by the CCD camera [53, 80]. This technique is advantageous due to its ability to detect materials with low binding affinity, such as binding interactions with natural lectins. When using traditional fluorescence, the microarray is washed to remove low-binding species to reduce background signal. When using EFF, a more accurate picture of the structural motifs of the analytes in the sample is collected, because even those that are not bound as tightly are detected [77, 81].

2.1.6 Commercial Lectin Microarrays

For convenience, multiple companies have developed commercial lectin microarrays. These are ready-for-use chips with spotted natural lectins. Currently, there are no commercial microarrays containing synthetic lectins. One of the most used commercial lectin arrays is the LecChip [82]. This array has 7 replicated arrays on the same chip, with each array consisting of 45 different natural lectins. This setup provides a considerable amount of data on a single chip, requiring minimal time for data collection. These commercial arrays also operate in the same manner as other microarrays, allowing for the analysis of either glycoproteins, tissue extracts, or the cell surface glycome: the glycans present on the cell surface. A disadvantage of the LecChip is that it requires a 3–20-h incubation period, depending on the glycan concentration and glycan-binding affinities. While this is a high-throughput method that requires little preparation for characterizing many samples, it is not a rapid technique. This chip also lends itself to the use of EFF, where an extra wash is not needed, yielding high sensitivity [81, 83–85]. There are other commercial lectin microarrays such as the Qproteome Glycan Array kit, which contains 20 natural lectins and, thus, has lower glycan structural coverage due to the reduced number of lectins compared to the LecChip. Because there are few commercial options and each of these uses a predefined group of natural lectins, many labs choose to create their own microarray to increase the structural details collected during their glycan analyses [58].

2.1.7 Applications of Microarrays

In developing and testing lectin microarrays, the primary analytes have been purified glycoproteins. However, as the method has become more established, it is becoming more common to analyze surface glycans of cells [54, 56, 69, 86, 87], as well as the glycans present in biological samples such as tissue, blood, and urine [43, 44, 55, 60, 69]. This discussion of applications is not meant to be exhaustive but, rather, will illustrate the variety of methods currently available for characterizing glycan structures.

Glycoproteins

Purified glycoproteins are the simplest samples for glycan analysis using lectin microarrays, and they are most often used when developing and testing new microarrays. For example, commercial glycoproteins are used to verify the function of the microarray when altering the lectin printing conditions [88] or using nanoparticles as supports for lectins [89]. For this example, we will focus on the design of synthetic lectins, specifically peptide borono lectins [18]. In a study by Lavigne's group, low-diversity peptides with phenylboronic acid substitutions were developed, creating approximately 2,000,000 unique lectins. To test the binding and selectivity of their lectins, they used four commercially purchased glycoproteins (ovalbumin, bovine submaxillary mucin, porcine stomach mucin, and carcinoembryonic antigen) to probe different interactions with the lectins. Each lectin was bound to beads, and the beads for each synthetic lectin were placed in a single well of a 96-well microplate. The glycoproteins were labeled with fluorescein isothiocyanate (FITC), a fluorescent tag, and then a single glycoprotein was added to each well containing beads with a single synthetic lectin. Bovine serum albumin (BSA) was used as a control to show the interaction of a protein with no glycans, as well as to block non-specific interactions. After binding to the glycoproteins, BSA (1%) was added to the beads, and they were washed multiple times with phosphate buffered saline (PBS). Fluorescence was then used to measure the interaction between the glycoproteins and synthetic lectins. The study illustrated that borono lectins bound selectively to the glycans of glycoproteins and that each lectin bound to different glycan structural motifs. To show the use of these synthetic lectins in cancer analyses, the authors incubated the beads with carcinoembryonic antigen (CEA) at varying concentrations. The PBLs bound to the glycans and, thus, could identify cancer targets [90]. While these lectins show potential, additional experiments are needed to show binding to other cancer targets.

Cell Surface Glycans

The cell surface glycome is a characteristic of each cell line and is also subject to changes in cell culture conditions [91]. Hirabayashi's group analyzed a cell surface glycome in 2007, using a microarray with 43 natural lectins on an epoxy slide [53, 84]. For this study, approximately 1×10^7 cells/mL were harvested and labeled with

CellTracker Orange, a fluorescent tag. The labeled cells were added to the slide and incubated for 1 h, before detection via EFF. This was one of the first papers that enabled analysis of the cell surface glycome via lectin microarray. The analysis involved comparison of glycan signatures of various cell lines, including Lec cells (a variation of the Chinese hamster ovary (CHO) cell line containing mutations in glycosylation enzymes) [92]. Wild-type CHO cells and multiple Lec cell lines (Lec1, Lec2, and Lec8 with mutations in the genes mannosyl (alpha-1,3-) glycoprotein beta-1,2-N-acetylglucosamine (*MGAT1*), solute carrier family 35 member A1 (*Slc35a1*), and solute carrier family 35 member A2 (*Slc35a2*), respectively) [93] were harvested, and the glycome was analyzed on lectin microarrays. All four cell lines showed a distinct binding pattern on the lectin microarray, indicating that they each have their own glycome signature. This was the first study published using lectin microarrays to observe live cells, and the authors suggest that this method could be used to study more complex cell lines and classify cancerous versus normal cells.

Glycans from Biological Samples

Glycans play significant biological roles, and changes in their structure are observed in disease and disease progression. Many groups have looked at changes in glycan structures in blood, urine, or tissues in patients with cancer. In 2011, Leathem's group did a study using the LecChip to analyze glycans from urine, sera, and breast tissue samples, collected from healthy patients and those with breast tumors [60]. Tissue samples were centrifuged, and xylenes and ethanol were added to remove wax and formalin, which were present for tissue storage. Serum was prepared by adding clotting additives to blood samples, centrifuging, and recovering the supernatant. Urine did not require any sample preparation prior to microarray analysis. Sera and urine were analyzed from patients who had metastatic cancer and those for which metastasis had not been observed. All samples were fluorescently labeled with cyanine 3 (Cy3). Samples were added to lectin microarrays, and after 150-min incubation, fluorescence detection was used to observe binding interactions. This study showed significant differences in the glycan patterns for healthy and cancerous patients. There was a difference seen in the presence of glycan binding, as well as the intensity of fluorescence signals for bound glycans. The data revealed patterns in the glycans but was not conclusive enough to define specific glycan signatures for either cancerous or healthy patients.

2.1.8 Microarray Conclusions

Lectin microarrays are useful for qualitative and comparative information regarding glycan structures. Furthermore, microarrays provide a way to sample the inherent heterogeneity within glycans. However, there are limitations for this technique in the lack of quantitative data. The amount of each glycan subgroup cannot be

determined, and glycans cannot be separated for further analysis using microarrays. Lectin affinity chromatography (LAC) is a complementary technique that is used to separate and identify glycans that are present in samples. We will discuss LAC in the next section of this review.

2.2 Chromatography

Chromatography is a versatile technique that allows for separation of an analyte from a mixture based on a specific property, such as size, polarity, or binding affinity. In this review, we will focus on affinity chromatography, specifically LAC and its variations; other reviews discuss additional applications of chromatography for glycan separation and analysis [43, 94–97]. In affinity chromatography, ligands are immobilized, forming the stationary phase and allowing for specific interactions with analytes in the mobile phase; for LAC, lectins are the immobilized ligands. When a sample is loaded onto the column, the glycans that interact with the immobilized lectin bind to the stationary phase and are retained, while the rest of the sample flows through the column. The glycan can then be eluted off the column with a competing glycan or carbohydrate that has a stronger affinity for the lectin than the glycan in the sample, as seen in Fig. 3.

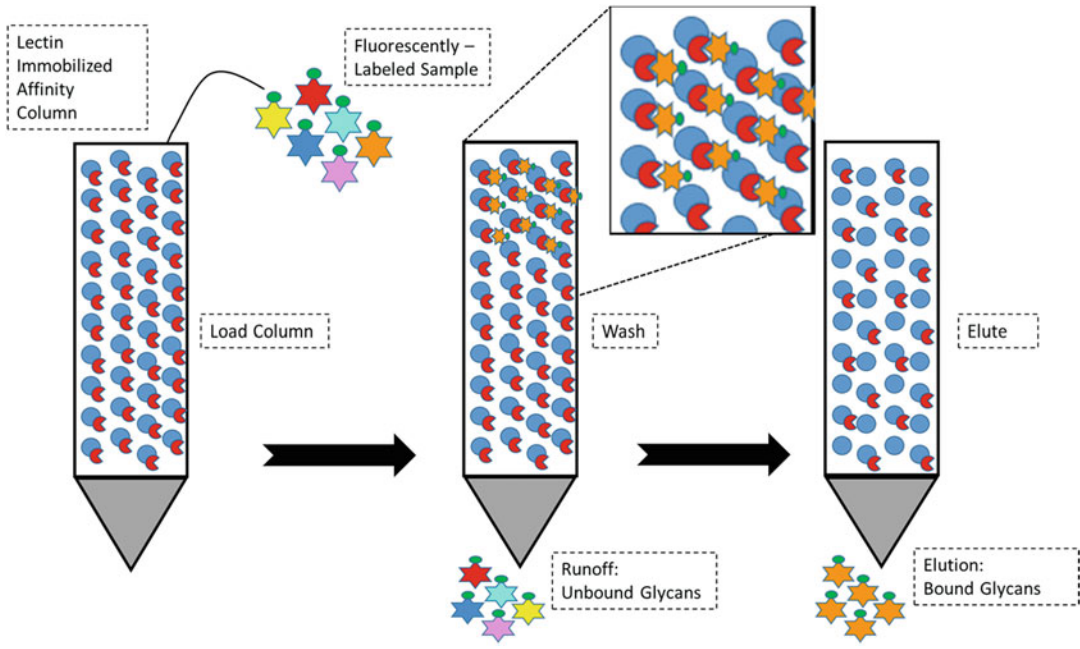


Fig. 3 Overview of a general lectin affinity column. Lectins are immobilized onto beads and packed into a column. Then, fluorescently labeled samples of glycans or glycoproteins are loaded onto the column. The glycans that bind to the immobilized lectin are retained (orange stars), while other analytes, including glycans that do not interact with the lectin, are eluted from the column (red, blue, pink, yellow, and teal stars). The bound glycans are eluted from the column following addition of a sugar with a stronger binding interaction with the immobilized lectin

There are variations and different purposes for the use of LAC. A variation of this technique is serial lectin affinity chromatography (SLAC), in which the sample is analyzed on multiple columns back to back, with each column having a different immobilized lectin to bind various glycan structural motifs. Another use of LAC is frontal affinity chromatography (FAC), which is used to quantify the binding affinities between glycans and lectins.

2.2.1 Affinity Column Preparation

While many chromatography columns can be purchased, it is often necessary to prepare and pack columns with specific lectins. Lectins must first be immobilized on stationary phase supports, which vary depending on the pressure requirements of the separation [98]. Briefly, lectins are dissolved in coupling buffer ($NaHCO_3$ and NaCl) and added to N-hydroxysuccinimide (NHS)-activated sepharose or agarose [99]. The stationary phase support is chosen to maximize lectin immobilization while minimizing non-specific interactions with glycans. Before packing, the column is filled with buffer to avoid forming bubbles in the packed particle bed. The buffer is then removed as the column is filled with lectin-immobilized gel. Once the column is packed, it can be stored at 4 °C for multiple years without loss of activity [99]. Most of the time, the stationary phase beads are packed into column housings that can withstand applied pressures, to allow for automation of separations and to decrease analysis time.

2.2.2 Sampling Requirements and Detection Methods

Unlike lectin microarrays, columns cannot be used to analyze crude samples. Many examples of LAC use serum samples to test immobilized lectins. Prior to analysis, proteins in the sample must be removed to reduce non-specific interactions [100–102]. However, it is not ideal to remove the glycans entirely, because it has been observed that lectins have stronger affinity to glycoproteins compared to free glycans [98]. The glycan samples are then tagged with a fluorescent label for detection and diluted before injecting them into the column. However, samples do not need to be labeled if using a non-fluorescent detector, such as a mass spectrometer [103]. Fluorescence detection is common with SLAC and FAC, when attempting to confirm the presence of the glycan. However, mass spectrometry is more often used following LAC, when additional structural details of the sample are desired.

2.2.3 Serial Lectin Affinity Chromatography and Its Applications

LAC is often used as a purification method prior to additional analysis techniques, such as mass spectrometry [102]. This is useful for enriching a glycan from a crude sample, because only one lectin is present. It is not an ideal technique for separations of samples containing multiple glycans and is not often used by itself to analyze the structure of glycans. However, SLAC is often used for the fractionation and analysis of a sample containing multiple glycans. An advantage of this technique is that glycan mixtures, containing

both *N*- and *O*-linked glycans, can be separated based on structural motifs using simple instrumentation [98].

One of the original studies using this technique was published in 1982, where SLAC facilitated the fractionation of *N*-linked glycans [104]. This study used three separate lectin affinity columns (Con A, pea lectin, and phytohaemagglutinin) to fractionate *N*-linked glycans removed from mouse lymphoma cells, for further structural characterization of the glycans. Since this study, there have been developments in separation automation and the structural details of these analyses have increased with the use of more specific lectins and more advanced detection methods, specifically mass spectrometry. A more recent study from 2017 combined both SLAC and mass spectrometry to separate and analyze glycoforms of a single protein [105]. For this study, human blood samples were obtained, and the glycoproteins in the plasma were separated for analysis. This study characterized the *O*-linked glycans present on glycoforms of IgA1, a human antibody. They used two lectins for SLAC, *Helix pomatia* agglutinin (HPA) and peanut agglutinin (PNA), to create four glycoprotein fractions (HPA^+/PNA^+, HPA^+/PNA^-, HPA^-/PNA^+, and HPA^-/PNA^-). The *O*-linked glycans were then released from the protein using pronase and further analyzed with mass spectrometry. Thus, this technique could be used for clinical diagnostics by analyzing serum samples of healthy and cancerous patients to observe different glycoforms in the samples.

2.2.4 Frontal Affinity Chromatography and Its Applications

FAC is a tool for determining glycan-lectin binding affinities. This technique was originally published in 1986 [106] but has since been automated using pressurized columns [99]. The principle behind FAC, seen in Fig. 4, is that an excess of diluted glycan is continuously added to a column that contains immobilized lectin. The binding of the glycan to the lectin is then measured by the elution volume or time that it takes for the analyte to exit the column. If the glycan does not interact with the lectin, it will elute rapidly. However, if the glycan binds to the lectin, it will be retarded as it flows through the column [99, 107]. This technique can be used to quantify binding affinities for specified interactions, as in Eq. 1. In which, the K_d is the binding affinity of the lectin, B_t is the total amount of immobilized lectin, V is the reference volume, V_0 is the elution volume of the glycan or glycoprotein analyte, and $[A]_0$ is the initial analyte concentration. The difference between the reference volume and the elution volume of the analyte ($V - V_0$) is the elution volume of the column.

$$[A]_0(V - V_0) = \frac{B_t[A]_0}{([A]_0 + K_d)} \tag{1}$$

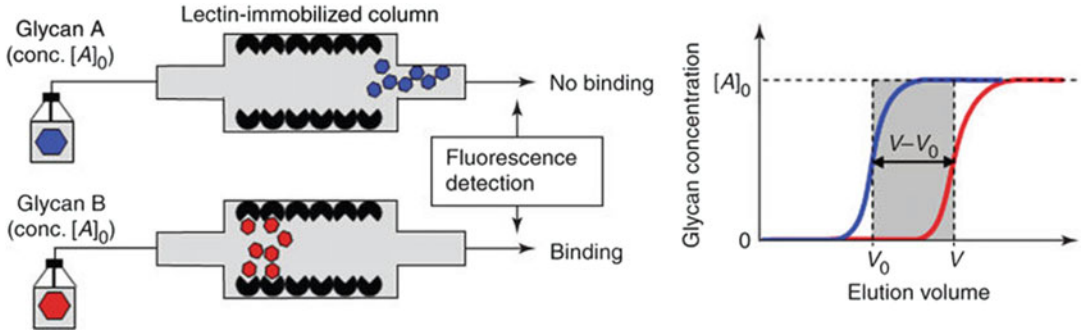


Fig. 4 Overview of frontal affinity chromatography. Glycan A is added to the column but has no interaction with the bound lectins. Thus, Glycan A passes through the column rapidly, and its elution can be used to determine the reference volume, V_0. Glycan B is then added to the column, and it interacts with the lectins; this retards the elution of Glycan B, which has an elution volume described by V. Using Eq. 2, the differences in the elution volumes can then be used to determine the K_d of the immobilized lectin and the glycan with which it interacts. This can be repeated for different glycans, calculating the binding affinity of a lectin for a variety of glycans. Reprinted by permission from Springer Nature: Springer Nature, Nature Protocols (Frontal Affinity Chromatography: sugar-protein interactions, Tateno et al.) Copyright 2007 [99]

However, the exact concentration of the analyte is often unknown, so a simplified equation is used (Eq. 2) in which the concentration of the glycan ($[A_0]$) is assumed to be significantly less than K_d.

$$K_d = \frac{B_t}{(V - V_0)}, \quad K_d \gg [A]_0 \tag{2}$$

The automated version of FAC is preferred due to its time advantage (e.g., less than 10 h are required for 100 sample analyses) [108]. This technique is highly useful when it comes to lectin microarray analyses, which were described in Subheading 2.1, because the development of FAC has quantified many binding affinities for lectins and glycans [109, 110].

Due to its high specificity, FAC is an excellent technique for identifying differences in lectin binding specificities. Galectins, a subset of lectins, are a family of proteins with very similar binding interactions to similar glycans. However, FAC was used to study six different galectins, and unique binding fingerprints were observed for each galectin [111]. This is a promising study and shows the potential for the use of FAC for identifying differences between lectins as new proteins are discovered. This is also a useful technique that could eventually be combined with synthetic lectins to close the gap of knowledge concerning the binding specificity of synthetic peptides and aptamers.

3 Conclusions

Glycans are diverse and play prominent roles in many systems, yet their structures are challenging to characterize. This review describes a few of the most recent techniques that have been developed and utilized to analyze the structures of both *N*- and *O*-linked glycans, specifically those that involve the use of lectins. Lectins are a powerful analytical tool because they bind different structural motifs of glycans, such as monosaccharides, linkages, and stereochemistries. This is particularly useful when analyzing glycans involved in cancer, because a wide variety of glycans can be identified through their unique structural characteristics. These methods allow for comparison of the type and number of glycans that are present when observing cancerous samples. These are techniques that are paving the way for monitoring glycan biomarkers in cancer.

References

1. Moremen KW, Tiemeyer M, Nairn AV (2012) Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol 13(7):448–462. https://doi.org/10.1038/nrm3383
2. Jayaprakash NG, Surolia A (2017) Role of glycosylation in nucleating protein folding and stability. Biochem J 474 (14):2333–2347. https://doi.org/10.1042/BCJ20170111
3. Maverakis E, Kim K, Shimoda M, Gershwin ME, Patel F, Wilken R, Raychaudhuri S, Ruhaak LR, Lebrilla CB (2015) Glycans in the immune system and the altered glycan theory of autoimmunity: a critical review. J Autoimmun 57:1–13. https://doi.org/10.1016/j.jaut.2014.12.002
4. Taniguchi N, Kizuka Y (2015) Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. Adv Cancer Res 126:11–51. https://doi.org/10.1016/bs.acr.2014.11.001
5. Freeze HH, Schachter H, Kinoshita T (2015) Genetic disorders of glycosylation. In: Varki A, Cummings RD, Esko JD et al (eds) Essentials of glycobiology, 3rd edn. Cold Spring Harbor, New York. https://doi.org/10.1101/glycobiology.3e.045
6. Varki A, Kornfeld S (2015) Historical Background and Overview. In: Varki A, Cummings RD, Esko JD et al (eds) Essentials of glycobiology, 3rd edn. Cold Spring Harbor, New York. https://doi.org/10.1101/glycobiology.3e.001
7. Kailemia MJ, Park D, Lebrilla CB (2017) Glycans and glycoproteins as specific biomarkers for cancer. Anal Bioanal Chem 409 (2):395–410. https://doi.org/10.1007/s00216-016-9880-6
8. Veillon L, Huang Y, Peng W, Dong X, Cho BG, Mechref Y (2017) Characterization of isomeric glycan structures by LC-MS/MS. Electrophoresis 38(17):2100–2114. https://doi.org/10.1002/elps.201700042
9. Mechref Y (2011) Analysis of glycans derived from glycoconjugates by capillary electrophoresis-mass spectrometry. Electrophoresis 32(24):3467–3481. https://doi.org/10.1002/elps.201100342
10. Suzuki S (2013) Recent developments in liquid chromatography and capillary electrophoresis for the analysis of glycoprotein glycans. Anal Sci 29(12):1117–1128
11. Furukawa J, Fujitani N, Shinohara Y (2013) Recent advances in cellular glycomic analyses. Biomol Ther 3(1):198–225. https://doi.org/10.3390/biom3010198
12. Shajahan A, Heiss C, Ishihara M, Azadi P (2017) Glycomic and glycoproteomic analysis of glycoproteins-a tutorial. Anal Bioanal Chem 409(19):4483–4505. https://doi.org/10.1007/s00216-017-0406-7
13. Nizet V, Varki A, Aebi M (2015) Microbial lectins: hemagglutinins, adhesins, and toxins. In: Varki A, Cummings RD et al (eds) Essentials of glycobiology. Cold Spring Harbor, New York, pp 481–491. https://doi.org/10.1101/glycobiology.3e.037

14. Lam SK, Ng TB (2011) Lectins: production and practical applications. Appl Microbiol Biotechnol 89(1):45–55. https://doi.org/10.1007/s00253-010-2892-9
15. Gemeiner P, Mislovicova D, Tkac J, Svitel J, Patoprsty V, Hrabarova E, Kogan G, Kozar T (2009) Lectinomics II. A highway to biomedical/clinical diagnostics. Biotechnol Adv 27 (1):1–15. https://doi.org/10.1016/j.biotechadv.2008.07.003
16. Taylor ME, Drickamer K, Schnaar RL, Etzler ME, Varki A (2015) Discovery and Classification of Glycan-Binding Proteins. In: Varki A, Cummings RD, Esko JD et al (eds) Essentials of glycobiology, 3rd edn. Cold Spring Harbor, NY. https://doi.org/10.1101/glycobiology.3e.028
17. Bicker KL, Sun J, Harrell M, Zhang Y, Pena MM, Thompson PR, Lavigne JJ (2012) Synthetic lectin arrays for the detection and discrimination of cancer associated glycans and cell lines. Chem Sci 3(4):1147–1156. https://doi.org/10.1039/c2sc00790h
18. Zou Y, Broughton DL, Bicker KL, Thompson PR, Lavigne JJ (2007) Peptide borono lectins (PBLs): a new tool for glycomics and cancer diagnostics. Chembiochem 8 (17):2048–2051. https://doi.org/10.1002/cbic.200700221
19. Li M, Lin N, Huang Z, Du L, Altier C, Fang H, Wang B (2008) Selecting aptamers for a glycoprotein through the incorporation of the boronic acid moiety. J Am Chem Soc 130(38):12636–12638. https://doi.org/10.1021/ja801510d
20. Rini J, Esko J, Varki A (2009) Glycosyltransferases and glycan-processing enzymes. In: nd VA, Cummings RD et al (eds) Essentials of Glycobiology. Cold Spring Harbor, NY
21. Ednie AR, Bennett ES (2012) Modulation of voltage-gated ion channels by sialylation. Compr Physiol 2(2):1269–1301. https://doi.org/10.1002/cphy.c110044
22. Stanley P, Taniguchi N, Aebi M (2015) N-Glycans. In: Varki A, Cummings RD, Esko JD et al (eds) Essentials of glycobiology, 3rd edn. Cold Spring Harbor, NY. https://doi.org/10.1101/glycobiology.3e.009
23. Gill DJ, Clausen H, Bard F (2011) Location, location, location: new insights into O-GalNAc protein glycosylation. Trends Cell Biol 21(3):149–158. https://doi.org/10.1016/j.tcb.2010.11.004
24. Gabius HJ (2006) Cell surface glycans: the why and how of their functionality as biochemical signals in lectin-mediated information transfer. Crit Rev Immunol 26(1):43–79
25. Hoja-Lukowicz D, Przybylo M, Duda M, Pochec E, Bubka M (2017) On the trail of the glycan codes stored in cancer-related cell adhesion proteins. Biochim Biophys Acta 1861(1 Pt A):3237–3257. https://doi.org/10.1016/j.bbagen.2016.08.007
26. Rambaruth ND, Dwek MV (2011) Cell surface glycan-lectin interactions in tumor metastasis. Acta Histochem 113(6):591–600. https://doi.org/10.1016/j.acthis.2011.03.001
27. Gagneux P, Aebi M, Varki A (2015) Evolution of glycan diversity. In: Varki A, Cummings RD, Esko JD et al (eds) Essentials of glycobiology, 3rd edn. Cold Spring Harbor, NY. https://doi.org/10.1101/glycobiology.3e.020
28. Higel F, Seidl A, Sorgel F, Friess W (2016) N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. Eur J Pharm Biopharm 100:94–100. https://doi.org/10.1016/j.ejpb.2016.01.005
29. Fuster MM, Esko JD (2005) The sweet and sour of cancer: glycans as novel therapeutic targets. Nat Rev Cancer 5(7):526–542. https://doi.org/10.1038/nrc1649
30. Pinho SS, Reis CA (2015) Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer 15(9):540–555. https://doi.org/10.1038/nrc3982
31. Meany DL, Chan DW (2011) Aberrant glycosylation associated with enzymes as cancer biomarkers. Clin Proteomics 8(1):7. https://doi.org/10.1186/1559-0275-8-7
32. Oliveira-Ferrer L, Legler K, Milde-Langosch K (2017) Role of protein glycosylation in cancer metastasis. Semin Cancer Biol 44:141–152. https://doi.org/10.1016/j.semcancer.2017.03.002
33. Veillon L, Fakih C, Abou-El-Hassan H, Kobeissy F, Mechref Y (2017) Glycosylation changes in brain cancer. ACS Chem Neurosci. https://doi.org/10.1021/acschemneuro.7b00271
34. Varki A, Kannagi R, Toole BP (2009) Glycosylation changes in cancer. In: Varki A, Cummings RD, Esko JD et al (eds) Essentials of glycobiology, 2nd edn. Cold Spring Harbor, NY
35. Magnelli P, McClung C Proteomics: fast and efficient antibody deglycosylation using rapid PNGase F. New England Biolabs Inc Glycobiology & Protein Tools (Appl. Note)
36. Maley F, Trimble RB, Tarentino AL, Plummer TH, Jr. (1989) Characterization of

glycoproteins and their associated oligosaccharides through the use of endoglycosidases. Anal Biochem 180 (2):195–204

37. Royle L, Mattu TS, Hart E, Langridge JI, Merry AH, Murphy N, Harvey DJ, Dwek RA, Rudd PM (2002) An analytical and structural database provides a strategy for sequencing O-glycans from microgram quantities of glycoproteins. Anal Biochem 304(1):70–90. https://doi.org/10.1006/abio.2002.5619
38. Kang JG, Ko JH, Kim YS (2016) Application of cancer-associated glycoforms and glycan-binding probes to an in vitro diagnostic multivariate index assay for precise diagnoses of cancer. Proteomics 16(24):3062–3072. https://doi.org/10.1002/pmic.201500553
39. Hollingsworth MA, Swanson BJ (2004) Mucins in cancer: protection and control of the cell surface. Nat Rev Cancer 4(1):45–60. https://doi.org/10.1038/nrc1251
40. Hirao Y, Matsuzaki H, Iwaki J, Kuno A, Kaji H, Ohkura T, Togayachi A, Abe M, Nomura M, Noguchi M, Ikehara Y, Narimatsu H (2014) Glycoproteomics approach for identifying Glycobiomarker candidate molecules for tissue type classification of non-small cell lung carcinoma. J Proteome Res 13(11):4705–4716. https://doi.org/10.1021/pr5006668
41. Dube DH, Bertozzi CR (2005) Glycans in cancer and inflammation--potential for therapeutics and diagnostics. Nat Rev Drug Discov 4(6):477–488. https://doi.org/10.1038/nrd1751
42. Dennis JW, Laferte S, Waghorne C, Breitman ML, Kerbel RS (1987) Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis. Science 236 (4801):582–585
43. Abbott KL, Lim JM, Wells L, Benigno BB, McDonald JF, Pierce M (2010) Identification of candidate biomarkers with cancer-specific glycosylation in the tissue and serum of endometrioid ovarian cancer patients by glycoproteomic analysis. Proteomics 10(3):470–481. https://doi.org/10.1002/pmic.200900537
44. Abbott KL, Nairn AV, Hall EM, Horton MB, McDonald JF, Moremen KW, Dinulescu DM, Pierce M (2008) Focused glycomic analysis of the N-linked glycan biosynthetic pathway in ovarian cancer. Proteomics 8 (16):3210–3220. https://doi.org/10.1002/pmic.200800157
45. Zhang S, Cordon-Cardo C, Zhang HS, Reuter VE, Adluri S, Hamilton WB, Lloyd KO, Livingston PO (1997) Selection of tumor antigens as targets for immune attack using immunohistochemistry: I. Focus on gangliosides. Int J Cancer 73(1):42–49
46. Zhang S, Zhang HS, Cordon-Cardo C, Reuter VE, Singhal AK, Lloyd KO, Livingston PO (1997) Selection of tumor antigens as targets for immune attack using immunohistochemistry: II. Blood group-related antigens. Int J Cancer 73(1):50–56
47. Toth E, Vekey K, Ozohanics O, Jeko A, Dominczyk I, Widlak P, Drahos L (2016) Changes of protein glycosylation in the course of radiotherapy. J Pharm Biomed Anal 118:380–386. https://doi.org/10.1016/j.jpba.2015.11.010
48. Heller MJ (2002) DNA microarray technology: devices, systems, and applications. Annu Rev Biomed Eng 4:129–153. https://doi.org/10.1146/annurev.bioeng.4.020702.153438
49. Hall DA, Ptacek J, Snyder M (2007) Protein microarray technology. Mech Ageing Dev 128(1):161–167. https://doi.org/10.1016/j.mad.2006.11.021
50. Cretich M, Damin F, Chiari M (2014) Protein microarray technology: how far off is routine diagnostics? Analyst 139(3):528–542. https://doi.org/10.1039/c3an01619f
51. Angeloni S, Ridet JL, Kusy N, Gao H, Crevoisier F, Guinchard S, Kochhar S, Sigrist H, Sprenger N (2005) Glycoprofiling with micro-arrays of glycoconjugates and lectins. Glycobiology 15(1):31–41. https://doi.org/10.1093/glycob/cwh143
52. Pilobello KT, Krishnamoorthy L, Slawek D, Mahal LK (2005) Development of a lectin microarray for the rapid analysis of protein glycopatterns. Chembiochem 6(6):985–989. https://doi.org/10.1002/cbic.200400403
53. Kuno A, Uchiyama N, Koseki-Kuno S, Ebe Y, Takashima S, Yamada M, Hirabayashi J (2005) Evanescent-field fluorescence-assisted lectin microarray: a new strategy for glycan profiling. Nat Methods 2(11):851–856. https://doi.org/10.1038/nmeth803
54. Chen S, Zheng T, Shortreed MR, Alexander C, Smith LM (2007) Analysis of cell surface carbohydrate expression patterns in normal and tumorigenic human breast cell lines using lectin arrays. Anal Chem 79 (15):5698–5702. https://doi.org/10.1021/ac070423k
55. Li Y, Tao SC, Bova GS, Liu AY, Chan DW, Zhu H, Zhang H (2011) Detection and verification of glycosylation patterns of glycoproteins from clinical specimens using lectin microarrays and lectin-based immunosorbent

assays. Anal Chem 83(22):8509–8516. https://doi.org/10.1021/ac201452f

56. Landemarre L, Cancellieri P, Duverger E (2013) Cell surface lectin array: parameters affecting cell glycan signature. Glycoconj J 30(3):195–203. https://doi.org/10.1007/s10719-012-9433-y
57. Chen P, Liu Y, Kang X, Sun L, Yang P, Tang Z (2008) Identification of N-glycan of alpha-fetoprotein by lectin affinity microarray. J Cancer Res Clin Oncol 134(8):851–860. https://doi.org/10.1007/s00432-008-0357-7
58. Tao SC, Li Y, Zhou J, Qian J, Schnaar RL, Zhang Y, Goldstein IJ, Zhu H, Schneck JP (2008) Lectin microarrays identify cell-specific and functionally significant cell surface glycan markers. Glycobiology 18 (10):761–769. https://doi.org/10.1093/glycob/cwn063
59. Microarrayers NanoPrint. (2017) Arrayit Corporation. http://www.arrayit.com/Products/Microarrayers/Microarray_Printer/microarray_printer.html. Accessed November 2017
60. Fry SA, Afrough B, Lomax-Browne HJ, Timms JF, Velentzis LS, Leathem AJ (2011) Lectin microarray profiling of metastatic breast cancers. Glycobiology 21 (8):1060–1070. https://doi.org/10.1093/glycob/cwr045
61. Leymarie N, Zaia J (2012) Effective use of mass spectrometry for glycan and glycopeptide structural analysis. Anal Chem 84 (7):3040–3048. https://doi.org/10.1021/ac3000573
62. Ghazarian H, Idoni B, Oppenheimer SB (2011) A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. Acta Histochem 113(3):236–247. https://doi.org/10.1016/j.acthis.2010.02.004
63. Ambrosi M, Cameron NR, Davis BG (2005) Lectins: tools for the molecular understanding of the glycocode. Org Biomol Chem 3 (9):1593–1608. https://doi.org/10.1039/b414350g
64. Davis AP (2009) Synthetic lectins. Org Biomol Chem 7(18):3629–3638. https://doi.org/10.1039/b909856a
65. Houston TA (2010) Developing high-affinity boron-based receptors for cell-surface carbohydrates. Chembiochem 11(7):954–957. https://doi.org/10.1002/cbic.201000079
66. Brighid Pappin MJK, Houston TA (2012) Boron-carbohydrate interactions. Comp Stud Glycobiol Glycotechnol. Doi:https://doi.org/10.5772/50630
67. Arnaud J, Audfray A, Imberty A (2013) Binding sugars: from natural lectins to synthetic receptors and engineered neolectins. Chem Soc Rev 42(11):4798–4813. https://doi.org/10.1039/c2cs35435g
68. Sun X, Zhai W, Fossey JS, James TD (2016) Boronic acids for fluorescence imaging of carbohydrates. Chem Commun (Camb) 52 (17):3456–3469. https://doi.org/10.1039/c5cc08633g
69. Li S, Mo C, Peng Q, Kang X, Sun C, Jiang K, Huang L, Lu Y, Sui J, Qin X, Liu Y (2013) Cell surface glycan alterations in epithelial mesenchymal transition process of Huh7 hepatocellular carcinoma cell. PLoS One 8 (8):e71273. https://doi.org/10.1371/journal.pone.0071273
70. Ito H, Kuno A, Sawaki H, Sogabe M, Ozaki H, Tanaka Y, Mizokami M, Shoda J, Angata T, Sato T, Hirabayashi J, Ikehara Y, Narimatsu H (2009) Strategy for glycoproteomics: identification of glyco-alteration using multiple glycan profiling tools. J Proteome Res 8(3):1358–1367. https://doi.org/10.1021/pr800735j
71. Sun Y, Cheng L, Gu Y, Xin A, Wu B, Zhou S, Guo S, Liu Y, Diao H, Shi H, Wang G, Tao SC (2016) A Human Lectin Microarray for Sperm Surface Glycosylation Analysis. Mol Cell Proteomics 15(9):2839–2851. https://doi.org/10.1074/mcp.M116.059311
72. Ruhaak LR, Zauner G, Huhn C, Bruggink C, Deelder AM, Wuhrer M (2010) Glycan labeling strategies and their use in identification and quantification. Anal Bioanal Chem 397 (8):3457–3481. https://doi.org/10.1007/s00216-010-3532-z
73. Korekane H, Shida K, Murata K, Ohue M, Sasaki Y, Imaoka S, Miyamoto Y (2007) Evaluation of laser microdissection as a tool in cancer glycomic studies. Biochem Biophys Res Commun 352(3):579–586. https://doi.org/10.1016/j.bbrc.2006.10.191
74. Murray GI (2007) An overview of laser microdissection technologies. Acta Histochem 109(3):171–176. https://doi.org/10.1016/j.acthis.2007.02.001
75. Sturm D, Marselli L, Ehehalt F, Richter D, Distler M, Kersting S, Grutzmann R, Bokvist K, Froguel P, Liechti R, Jorns A, Meda P, Baretton GB, Saeger HD, Schulte AM, Marchetti P, Solimena M (2013) Improved protocol for laser microdissection of human pancreatic islets from surgical specimens. J Vis Exp (71). doi:https://doi.org/10.3791/50231
76. Espina V, Wulfkuhle JD, Calvert VS, VanMeter A, Zhou W, Coukos G, Geho

DH, Petricoin EF 3rd, Liotta LA (2006) Laser-capture microdissection. Nat Protoc 1 (2):586–603. https://doi.org/10.1038/nprot.2006.85

77. Kuno A, Matsuda A, Ikehara Y, Narimatsu H, Hirabayashi J (2010) Differential glycan profiling by lectin microarray targeting tissue specimens. Methods Enzymol 478:165–179. https://doi.org/10.1016/S0076-6879
78. Koshi Y, Nakata E, Yamane H, Hamachi I (2006) A fluorescent lectin array using supramolecular hydrogel for simple detection and pattern profiling for various glycoconjugates. J Am Chem Soc 128(32):10413–10422. https://doi.org/10.1021/ja0613963
79. Lehr HP, Reimann M, Brandenburg A, Sulz G, Klapproth H (2003) Real-time detection of nucleic acid interactions by total internal reflection fluorescence. Anal Chem 75 (10):2414–2420
80. Uchiyama N, Kuno A, Koseki-Kuno S, Ebe Y, Horio K, Yamada M, Hirabayashi J (2006) Development of a lectin microarray based on an evanescent-field fluorescence principle. Methods Enzymol 415:341–351. https://doi.org/10.1016/S0076-6879(06)15021-1
81. Uchiyama N, Kuno A, Tateno H, Kubo Y, Mizuno M, Noguchi M, Hirabayashi J (2008) Optimization of evanescent-field fluorescence-assisted lectin microarray for high-sensitivity detection of monovalent oligosaccharides and glycoproteins. Proteomics 8 (15):3042–3050. https://doi.org/10.1002/pmic.200701114
82. Zhao R, Liu X, Wang Y, Jie X, Qin R, Qin W, Zhang M, Tai H, Yang C, Li L, Peng P, Shao M, Zhang X, Wu H, Ruan Y, Xu C, Ren S, Gu J (2016) Integrated glycomic analysis of ovarian cancer side population cells. Clin Proteomics 13:32. https://doi.org/10.1186/s12014-016-9131-z
83. Matsuda A, Kuno A, Ishida H, Kawamoto T, Shoda J, Hirabayashi J (2008) Development of an all-in-one technology for glycan profiling targeting formalin-embedded tissue sections. Biochem Biophys Res Commun 370 (2):259–263. https://doi.org/10.1016/j.bbrc.2008.03.090
84. Tateno H, Uchiyama N, Kuno A, Togayachi A, Sato T, Narimatsu H, Hirabayashi J (2007) A novel strategy for mammalian cell surface glycome profiling using lectin microarray. Glycobiology 17 (10):1138–1146. https://doi.org/10.1093/glycob/cwm084
85. Ltd. G (2016) GlycoTechnia Ltd. http://www.glycotechnica.com/english/products.htm. Accessed Oct 2017
86. Zheng T, Peelen D, Smith LM (2005) Lectin arrays for profiling cell surface carbohydrate expression. J Am Chem Soc 127 (28):9982–9983. https://doi.org/10.1021/ja0505550
87. Pilobello KT, Slawek DE, Mahal LK (2007) A ratiometric lectin microarray approach to analysis of the dynamic mammalian glycome. Proc Natl Acad Sci U S A 104 (28):11534–11539. https://doi.org/10.1073/pnas.0704954104
88. Pilobello KT, Agrawal P, Rouse R, Mahal LK (2013) Advances in lectin microarray technology: optimized protocols for piezoelectric print conditions. Curr Protoc Chem Biol 5 (1):1–23. https://doi.org/10.1002/9780470559277.ch120035
89. Fromell K, Andersson M, Elihn K, Caldwell KD (2005) Nanoparticle decorated surfaces with potential use in glycosylation analysis. Colloids Surf B Biointerfaces 46(2):84–91. https://doi.org/10.1016/j.colsurfb.2005.06.017
90. Burtis CA, Ashwood ER, Tietz NW (1999) Tietz textbook of clinical chemistry, 3rd edn. W.B. Saunders, New York
91. MSaM B (2015) Glycosylation in cell culture. In: Al-Rubeai M (ed) Animal cell culture. Springer, NY, pp 237–258
92. Patnaik SK, Stanley P (2006) Lectin-resistant CHO glycosylation mutants. Methods Enzymol 416:159–182. https://doi.org/10.1016/S0076-6879(06)16011-5
93. North SJ, Huang HH, Sundaram S, Jang-Lee J, Etienne AT, Trollope A, Chalabi S, Dell A, Stanley P, Haslam SM (2010) Glycomics profiling of Chinese hamster ovary cell glycosylation mutants reveals N-glycans of a novel size and complexity. J Biol Chem 285 (8):5759–5775. https://doi.org/10.1074/jbc.M109.068353
94. Cavazzini A, Pasti L, Massi A, Marchetti N, Dondi F (2011) Recent applications in chiral high performance liquid chromatography: a review. Anal Chim Acta 706(2):205–222. https://doi.org/10.1016/j.aca.2011.08.038
95. Cheung RC, Wong JH, Ng TB (2012) Immobilized metal ion affinity chromatography: a review on its applications. Appl Microbiol Biotechnol 96(6):1411–1420. https://doi.org/10.1007/s00253-012-4507-0
96. Hibbert DB (2012) Experimental design in chromatography: a tutorial review. J Chromatogr B Analyt Technol Biomed Life Sci 910:2–13. https://doi.org/10.1016/j.jchromb.2012.01.020

97. Pfaunmiller EL, Paulemond ML, Dupper CM, Hage DS (2013) Affinity monolith chromatography: a review of principles and recent analytical applications. Anal Bioanal Chem 405(7):2133–2145. https://doi.org/10.1007/s00216-012-6568-4
98. Yamashita K, Ohkura T (2014) Determination of glycan motifs using serial lectin affinity chromatography. Methods Mol Biol 1200:79–92. https://doi.org/10.1007/978-1-4939-1292-6_7
99. Tateno H, Nakamura-Tsuruta S, Hirabayashi J (2007) Frontal affinity chromatography: sugar-protein interactions. Nat Protoc 2 (10):2529–2537. https://doi.org/10.1038/nprot.2007.357
100. Zeng Z, Hincapie M, Pitteri SJ, Hanash S, Schalkwijk J, Hogan JM, Wang H, Hancock WS (2011) A proteomics platform combining depletion, multi-lectin affinity chromatography (M-LAC), and isoelectric focusing to study the breast cancer proteome. Anal Chem 83(12):4845–4854. https://doi.org/10.1021/ac2002802
101. Jung K, Cho W, Regnier FE (2009) Glycoproteomics of plasma based on narrow selectivity lectin affinity chromatography. J Proteome Res 8(2):643–650. https://doi.org/10.1021/pr8007495
102. Song E, Zhu R, Hammoud ZT, Mechref Y (2014) LC-MS/MS quantitation of esophagus disease blood serum glycoproteins by enrichment with hydrazide chemistry and lectin affinity chromatography. J Proteome Res 13(11):4808–4820. https://doi.org/10.1021/pr500570m
103. Drake PM, Schilling B, Niles RK, Prakobphol A, Li B, Jung K, Cho W, Braten M, Inerowicz HD, Williams K, Albertolle M, Held JM, Iacovides D, Sorensen DJ, Griffith OL, Johansen E, Zawadzka AM, Cusack MP, Allen S, Gormley M, Hall SC, Witkowska HE, Gray JW, Regnier F, Gibson BW, Fisher SJ (2012) Lectin chromatography/mass spectrometry discovery workflow identifies putative biomarkers of aggressive breast cancers. J Proteome Res 11 (4):2508–2520. https://doi.org/10.1021/pr201206w
104. Cummings RD, Kornfeld S (1982) Fractionation of asparagine-linked oligosaccharides by serial lectin-Agarose affinity chromatography. A rapid, sensitive, and specific technique. J Biol Chem 257(19):11235–11240
105. Lehoux S, Ju T (2017) Separation of two distinct O-glycoforms of human IgA1 by serial lectin chromatography followed by mass spectrometry O-glycan analysis. Methods Enzymol 585:61–75. https://doi.org/10.1016/bs.mie.2016.10.003
106. Kasai K, Oda Y, Nishikata M, Ishii S (1986) Frontal affinity chromatography: theory for its application to studies on specific interactions of biomolecules. J Chromatogr 376:33–47
107. Kasai K (2014) Frontal affinity chromatography: a unique research tool for biospecific interaction that promotes glycobiology. Proc Jpn Acad Ser B Phys Biol Sci 90(7):215–234
108. Nakamura-Tsuruta S, Uchiyama N, Hirabayashi J (2006) High-throughput analysis of lectin-oligosaccharide interactions by automated frontal affinity chromatography. Methods Enzymol 415:311–325. https://doi.org/10.1016/S0076-6879(06)15019-3
109. Hirabayashi J, Arata Y, Kasai K (2003) Frontal affinity chromatography as a tool for elucidation of sugar recognition properties of lectins. Methods Enzymol 362:353–368
110. Arata Y, Hirabayashi J, Kasai KI (2001) Application of reinforced frontal affinity chromatography and advanced processing procedure to the study of the binding property of a Caenorhabditis elegans galectin. J Chromatogr A 905(1–2):337–343
111. Hirabayashi J, Hashidate T, Arata Y, Nishi N, Nakamura T, Hirashima M, Urashima T, Oka T, Futai M, Muller WE, Yagi F, Kasai K (2002) Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. Biochim Biophys Acta 1572 (2–3):232–254

Chapter 21

Hyperpolarized MRI for Studying Tumor Metabolism

Mikko I. Kettunen

Abstract

Hyperpolarized magnetic resonance imaging (MRI) can be used to detect real-time in vivo tumor metabolism. Dissolution dynamic nuclear polarization method increases polarization of ^{13}C-labeled molecules, typically [1-^{13}C]pyruvate, which can be injected into an animal during MRI scanning. Increased polarization leads to a higher observed signal, which allows for the detection and imaging of the transfer of ^{13}C-label between the injected marker molecule, pyruvate, and its metabolic products, most importantly lactate. This information can be used to assess the metabolic status of the tumor, for example, during therapy. Here, the basic methodology and data analysis for a preclinical hyperpolarized pyruvate experiment are described.

Key words Hyperpolarization, Pyruvate, MR spectroscopy, Modeling

1 Introduction

Exploiting the difference between physiological and pathological metabolism is one of the main driving forces behind the development of noninvasive imaging methods. Among the most common preclinical and clinical imaging modalities, magnetic resonance (MR)-based methods are unique in that they allow for, both, a relatively high spatial resolution and a noninvasive separation of a range of metabolites (*see* **Note 1**). The higher information content provided by the latter also partially compensates for a much lower sensitivity when compared to methods such as positron emission tomography (PET). Spectroscopic analysis of metabolite distribution in the tissue using ^{1}H spectroscopy reveals a wealth of information of steady-state metabolite concentrations, but the production of metabolic maps is still time-consuming (minutes to tens of minutes). Further insight into metabolism can be acquired using dynamic metabolic studies of ^{13}C-labeled marker molecules, e.g., ^{13}C-glucose, by monitoring the transfer of ^{13}C-label to different downstream metabolites using ^{13}C MR spectroscopy. This method, however, has traditionally suffered from very low

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_21,

sensitivity, making the studies very long (tens of minutes to hours) and imaging very challenging.

Hyperpolarization amplifies the MR-visible signal of, e.g., ^{13}C-labeled molecules several thousandfold. Although there are several ways to achieve a hyperpolarized signal, dissolution dynamic nuclear polarization (dDNP) [1] (*see* **Note 2**) is currently the most used methodology because of its flexibility in terms of marker molecules and high signal amplification. dDNP increases the sensitivity of, e.g., ^{13}C-label >10,000-fold, making real-time metabolic mapping of the marker molecule and its metabolic products feasible. The signal levels remain increased for up to 3 min (*see* **Note 3**), depending on the studied molecule, and largely remove the obstacles associated with traditional ^{13}C spectroscopy while limiting studies to rapid metabolic events. Finally, as the signal is pre-amplified, it is largely independent of the actual imaging magnet used, so the dDNP technique makes rapid real-time metabolic imaging viable also at lower magnetic fields; this technique is already being implemented in clinical imaging [2].

The majority of dDNP studies have so far focused on cancer using [1-^{13}C]pyruvic acid [3]. Pyruvate is situated at the crossing point of the glycolytic pathway, which is in many ways optimal for studying tumor metabolism, with major downstream metabolism to lactate, alanine, or acetyl-CoA (releasing carbon dioxide in the process), depending on the needs of the cell. High glycolytic activity (the so-called Warburg effect) in tumors leads to an elevated lactated production, which is reflected as increased labeling of [1-^{13}C]lactate resonance following an intravenous injection of hyperpolarized [1-^{13}C]pyruvate (Fig. 1) [4]. It has been shown that the lactate labeling efficiency is related to factors such as tumor grade [5, 6] and can also be used to detect metabolic changes related to tumor therapy [3, 7].

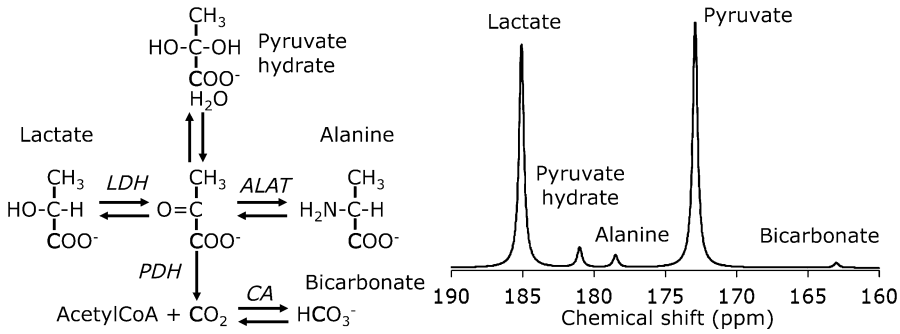


Fig. 1 Major metabolic downstream products of pyruvate (at the center) and a corresponding ^{13}C spectrum. In tumors, pyruvate and lactate signals are usually the most prominent. Enzymes catalyzing the reactions, LDH = lactate dehydrogenase, ALAT = alanine transaminase, PDH = pyruvate dehydrogenase, CA = carbonic anhydrase

2 Materials

dDNP is based on microwave-induced signal pre-amplification of the labeled marker molecule, usually ^{13}C, at a relatively high field (>3 T) and low temperature (~1 K), followed by a rapid dissolution of sample, providing a neutral room-temperature solution of the marker molecule with increased signal levels [1]. To achieve this, a marker molecule is mixed with radical-containing free electrons. Furthermore, the frozen sample needs to form an amorphous glass when frozen to allow efficient hyperpolarization. Unlike most other molecules, neat pyruvic acid forms glass when frozen, so there is no need for additional glassing agents.

2.1 Preclinical [1-^{13}C]Pyruvic Acid Sample for In Vitro or Small Animals at 3.35 T (e.g., Hypersense)

dDNP experiments require a dedicated hyperpolarizer system based around a high-field (3 T or more) magnet placed near the imaging magnet. In addition to the magnet, the main components of the hyperpolarizer include a variable temperature insert, capable of achieving ~1 K temperature to freeze the sample, a MR probe for solid-state signal buildup monitoring, a GHz-range microwave source (e.g., 94 GHz for 3.35 T), and a setup for rapid dissolution and removal of the sample (e.g., dissolution stick [1] or fluid path [8, 9]). The majority of the existing hyperpolarizer equipment are commercial systems, either preclinical Hypersense (Oxford Instruments, UK) or clinical SpinLab (Research Circle Technology, US) systems, but noncommercial designs based on the same principles are also available. The differences among systems are relatively minor, with the commercial ones usually featuring more automated sample dissolution process. In addition, a regular magnetic resonance (MR) imaging system with ^{13}C-channel will be required for signal detection.

1. [1-^{13}C]Pyruvic acid stock solution: 1 g neat [1-^{13}C]pyruvic acid (~14.1 mol/L, 99% ^{13}C) (*see* **Note 4**), 16.7 mg trityl radical (15 mmol/L, 8-carboxy-2,2,6,6-tetra-(hydroxyethyl)-benzo-[1,2–4,5]-bis- [1, 3]-dithiole-4-yl)-methyl sodium salt (OX063; Research Circle Technology, US/GE Healthcare) (*see* **Note 5**).

   The amount of radical needed for one hyperpolarized pyruvate sample is too small to be measured accurately, so a stock solution is prepared by mixing the pyruvic acid and radical (*see* **Note 6**).

2. [1-^{13}C]Pyruvic acid sample: 40 mg (~31 μL) of pyruvic acid stock solution, 1 μL of 1/10 dilution of 0.5 mmol/mL Gd chelate (~1.5 mmol/L, e.g., Dotarem, Guerbet).

   Warm up the pyruvate stock solution to room temperature, weigh pyruvic acid into the sample cup, and add an aqueous solution of a gadolinium chelate (*see* **Note 7**). Gd chelates are potentially unstable at low pH, so Gd should be added just before the sample goes into the polarizer to avoid degradation.

3. Dissolution (neutralization) buffer [8]: 1 L MilliQ water, 24.23 g (200 mM) tris-hydroxymethyl aminomethane (TRIS base), 4 g (100 mM) NaOH, 100 mg EDTA, pH ~7.6 (*see* **Note 8**).

3 Methods

3.1 Hyperpolarization

The exact steps required to achieve a polarized sample depend on the hyperpolarizer system used, so the main steps for polarization are summarized below, and the reader is advised to consult specific instructions for their system of choice.

1. During the polarization, immerse the sample into liquid helium under vacuum, so that temperatures ranged 1.0–1.4 K are reached. The variable temperature insert and probe need to be cooled down with liquid helium close to the operation temperature before the sample insertion. This avoids crystallization of the sample during the sample loading, which would lead to an inefficient polarization.
2. Prepare the sample by connecting the filled sample cup to, e.g., sample loading stick or fluid path. Load the correct amount (typically 5–6 mL for Hypersense and 14 mL for SpinLab) of prepared dissolution fluid to the system (*see* **Note 9**). When working with fluid path-based systems, all the preparation (sample loading, dissolution fluid loading, helium gas priming) needs to be done before the fluid path is inserted into the polarizer. In contrast, in dissolution stick-based systems, dissolution fluid can be usually added at a later stage (e.g., just before the dissolution), and helium priming is not usually needed as the sample is not separated from liquid helium in the system.
3. Repressurize the probe to atmospheric pressure before loading the sample. This avoids air getting sucked and frozen inside the system.
4. After loading the sample, initiate the microwave-driven polarization at the optimized microwave frequency and power. During polarization it is important to keep the sample covered in helium. Typical polarization times for pyruvate are from 30 to 120 min depending on the polarization field. The polarization process can be monitored using a small flip-angle pulse-acquire experiment and should be relatively consistent between samples (*see* **Note 10**).
5. Heat the dissolution fluid to the set temperature [e.g., 130 °C (SpinLab) or 180 °C (Hypersense)]; lift the sample out of helium, and perform the dissolution collecting the sample to a suitable container. In commercial systems, the dissolution

process is automated. The temperature is set to be high enough to allow rapid dissolution of the sample to avoid polarization losses and, in most cases, should not be changed. In systems with manual dissolution (e.g., dissolution stick), it is crucial to repressurize the system before the dissolution is performed to avoid rapid influx of air into the system. The sample is lifted a few cm above helium just before the dissolution is initiated to, on one hand, allow more efficient dissolution and, on the other, minimize signal losses when the sample is outside helium and optimal magnetic field. A significant heat load is introduced to the sample when the dissolution stick is lowered in place; therefore the dissolution process needs to be initiated as soon as the stick is in place.

A typical sample produced with the above procedure has ~80 mM [1-^{13}C]pyruvate concentration and neutral pH (~7.5) (*see* **Note 11**). The sample coming out of the polarizer can initially be hot (around 60 °C), so care should be taken when handling it.

6. Remove the sample needed for the imaging experiment to a separate syringe, and make sure it has cooled down sufficiently before using it. Because of the continuous loss of the sample, however, it is vital to perform the injection as soon as possible.

The achieved levels of polarization for [1-^{13}C]pyruvate vary greatly (usually being in the range of 20–70%) (*see* **Note 12**) depending on factors such as polarization field (higher polarization field usually yields higher polarization, albeit at a cost of longer polarization times). Polarization levels above 15% are usually required for in vivo experiments, but lower polarization can be used for in vitro experiments.

3.2 MRI Experiment

As mentioned above, injection of [1-^{13}C]pyruvate leads to ^{13}C-labeling of its major downstream metabolites, alanine, carbon dioxide, and lactate. Carbon dioxide rapidly exchanges the majority of the label to bicarbonate. Furthermore, aqueous pyruvate is in an equilibrium with pyruvate hydrate. Therefore, a ^{13}C spectrum following injection of [1-^{13}C]pyruvate contains a maximum of five peaks spanning over 20 ppm: ^{13}C-bicarbonate (~163 ppm), pyruvate (172.9 ppm), alanine (178.5 ppm), pyruvate hydrate (181 ppm), and lactate (185.1 ppm) (*see* **Note 1**). In tumor studies, the main interest are pyruvate and lactate signals (12 ppm apart), but alanine and bicarbonate can also be of interest when visible in tumor.

Hyperpolarized signal is lost relatively quickly, as the spin returns to their equilibrium state through T_1 relaxation (for pyruvate T_1 is around 50–60s ex vivo and 30s in vivo in typical magnetic fields used for research), and, furthermore, any radiofrequency (RF) irradiation destroys some hyperpolarized signal irreversibly.

Therefore, hyperpolarized MRI experiments differ from traditional MRI experiments in a few significant ways. Most importantly, low flip angles are usually employed, and their number is kept as low as possible to minimize RF-based losses. Fast imaging approaches where one metabolic map is ideally recorded after each RF pulse are, therefore, favored for imaging studies. Furthermore, because there is no need to wait for the signal to recover after RF pulsing, repetition time can be kept minimal when needed (e.g., when more than one RF pulse is required for each snapshot metabolite map). In the following, some typical MR approaches are summarized.

3.2.1 MR Spectroscopy (e.g., Cell Experiments, Bioreactors, In Vivo Experiments with Superficial Tumors)

Low flip-angle MR spectroscopy either without (ex vivo) or with (in vivo) slice selection is the simplest way to monitor tumor metabolism. The MR experiment can be usually started before the addition of the sample, so the whole metabolic curve can be recorded for analysis. All the produced peaks can be easily detected, and peak integrals can be obtained using any standard method, either in time- or Fourier-transformed frequency domain. Typical repetition time is around a second for pyruvate experiment with flip angle around 5°. Proton decoupling is not routinely used during hyperpolarized experiment, because the ^{13}C-label is in carboxylic acid making $J_{C\text{-}H}$ coupling small. When converting a proton spectroscopic sequence for ^{13}C use, it is important that fat/water suppression and outer volume suppression pulses are turned off because these could irreversibly destroy some pyruvate signal. Similarly, voxel selection based on 180° refocusing pulses is usually not used because they often lead to increased loss of polarization. Finally, it is vital that the chemical shift artifact is minimized by using high enough slice selection gradient and a suitable spectral bandwidth to cover all peaks.

3.2.2 Chemical Shift Imaging (CSI, e.g., In Vivo Metabolic Snapshots)

Addition of two phase encoding gradients to MR spectroscopic sequence allows the spatial localization of the observed signal within the slab. This is the simplest approach to collect spectroscopic imaging data with no limitations from the wide spectral bandwidth, and it is usually available in all imaging systems.

While CSI produces metabolic maps with high sensitivity, it requires a large number of low flip-angle RF pulsing (e.g., 16 × 16 matrix requires 256 pulses, although this can be reduced by not sampling the full k-space) (*see* **Note 13**) and therefore is not optimal for hyperpolarized studies as it allows typically only a single metabolic map to be acquired. A center-out-spiral phase encoding is usually applied to collect the most important part of k-space first, before significant RF-caused signal loss has occurred. Repetition times are usually kept short (around 20–50 ms) to collect data as quickly as possible. It is worth noting that, because the hyperpolarized ^{13}C spectra are very simple, with five or fewer reasonably

separated peaks typically present, full free-induction decay does not need to be sampled for accurate separation of the peaks.

Because of the limitations of CSI pulse sequence, the data acquisition is started after the injection of hyperpolarized sample. In case of tumor imaging, the main metabolite of interest is usually lactate, so the imaging time is often selected close to the time point where lactate signal is maximal (based on previous spectroscopic analysis). The peak time varies between tumor models and depends on experimental factors such as injection rate, which should, therefore, be maintained constant during the experiments. Times around 15–20s are often used to minimize signal loss due to T_1. Typical voxel size is in the order $1.5 \times 1.5 \times 5$ mm^3.

3.2.3 Fast Spectroscopic Imaging (e.g., In Vivo Metabolic Time Courses)

The limitations of CSI can be bypassed by using a number of fast spectroscopic imaging approaches, which are suitable for dynamic follow-up of metabolism or imaging of larger volumes through 3D imaging. The sparsity of ^{13}C spectrum makes them optimal for fast imaging approaches; in many tumor cases, the pyruvate and lactate are the only two prominent peaks. The most important factor to note when converting proton imaging sequences to carbon is that four times stronger gradients are needed, due to lower carbon gyromagnetic ratio. This can quickly lead to problems with the gradient performance if not accounted for. The technical challenges usually limit the resolution to around 5 mm, but the imaging parameters are highly dependent on the frequency difference between metabolites and, therefore, vary depending on the magnetic field of the imaging system and the selected imaging approach. The most common approaches are briefly summarized in the following paragraphs.

Imaging time can be significantly shortened by replacing one of the phase encoding directions with an echo-planar-type spectroscopic data collection (EPSI). Because the same frequency encoded signal is collected repeatedly following excitation, both spatial and spectroscopic information can be separated simultaneously by Fourier transformation. This reduces the number of RF pulses in case of 16×16 matrix to just 16, and a corresponding shortening in the imaging time is achieved. EPSI-type pulse sequences are widely available in imaging systems and can be set up for carbon experiments relatively easily (*see* **Note 14**). Alternatively, both phase encoding directions could be replaced with a repeated spiral-type data acquisition, in which case full metabolic map could in theory be achieved with just one excitation pulse.

As an alternative to Fourier-based spectroscopic imaging, the sparsity of ^{13}C spectra can be exploited by acquiring MR images (theoretically N + 1 images are needed to separate N peaks) at specific echo times, which are selected to give maximum phase separation of the underlying metabolic signals and then using

spectral decomposition methods to resolve the data [10, 11]. Typically, seven to eight images for time point are needed to allow the separation of pyruvate and all its downstream metabolites. When pyruvate and lactate are the only two prominent peaks, even fewer images may be sufficient. The appeal of this approach is that normal fast imaging sequences can be used to collect data, making it easier to optimize, and higher spatial resolution can be achieved. Despite its more complex post-processing, spiral-type data collection has been favored over echo-planar imaging because it allows for shorter echo times.

Finally, an approach where individual peaks are excited using spectral-spatial pulses and imaged sequentially can be adopted. In this case, the number of RF pulses per time point often matches the number of metabolites imaged, and only the peaks of interest (e.g., pyruvate and lactate) are imaged. Again, standard fast imaging sequences can be used for data collection. While spectral-spatial pulses used for this approach are not regularly available in standard MR systems (*see* **Note 15**), they offer great flexibility in designing the experiment [12]. However, great care is needed with shimming because variations in frequency offsets lead directly to signal losses and cannot be corrected in post-processing.

3.2.4 *In Vitro Experiment*

1. Confirm that the sample (e.g., a test tube with around 10^8 tumor cells or a bioreactor) is at the correct place inside the imaging magnet.
2. Shim the system at ^{1}H channel.
3. Change the system to ^{13}C and optimize the coil performance (tune/match) (*see* **Note 16**).
4. Start spectroscopic pulse sequence before the addition of a hyperpolarized compound.
5. Sample for 3–5 min until the signal has decayed away.

3.2.5 *In Vivo Experiment*

1. Anesthetize the animal (*see* **Note 17**), place an injection line, e.g., to the tail vein, and place the animal inside the magnet.
2. Shim the system at ^{1}H channel.
3. Collect reference images for metabolic maps or placing the spectroscopic slab.
4. Optimize the ^{13}C coil performance (tune/match) (*see* **Note 16**).
5. Select the ^{13}C acquisition slice/slab based on ^{1}H data.
6. Depending on the imaging approach, start the spectroscopic pulse sequence either before the intravenous injection of hyperpolarized compound or at a fixed time point after the injection.

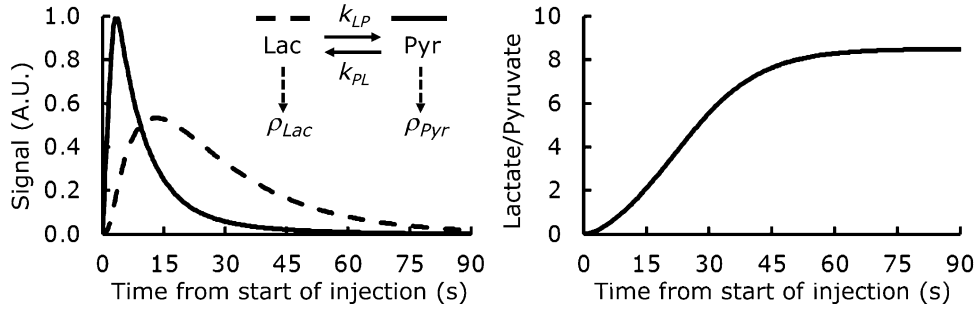


Fig. 2 Typical tumor pyruvate and lactate signal time courses and the corresponding lactate-pyruvate ratio; pyruvate injection is started at time 0 s. Time courses can be analyzed using a two-site exchange model including both exchange (k) and relaxation (ρ) parameters for each metabolite

Despite the large increase in the amount of signal available, a high-concentration bolus (3–10 mL/min) of marker molecule (e.g., 80 mM [1-^{13}C]pyruvate) still needs to be injected, typically at 5–10 mL/kg. This is in contrast, for example, to PET studies, where only trace amounts are used. Such high doses may transiently disturb the underlying metabolism and even lead to potential toxicity effects. No significant side effects have been reported for tumor experiments using hyperpolarized pyruvate, but normal physiological monitoring is advisable during the experiments.

3.3 Data Analysis

Currently there is no consensus on the best approach to analyze the data from hyperpolarized pyruvate experiments (Fig. 2). This is partially due, both, to the fact that there are several ways to collect the data, as presented above, and to the relatively young age of the field, so the methodology is still evolving.

3.3.1 Single Time Point: Lactate-Pyruvate Ratio

This is the simplest and fastest analysis for hyperpolarized data. In some cases, for example, for CSI-type snapshot imaging, it is also the only analysis method available. It is worth noting that the lactate-pyruvate ratio changes continuously during the experiment (Fig. 2), so the analysis is only meaningful if the data have been acquired at similar time points.

3.3.2 Time Courses: Kinetic Modeling

The label transfer can be analyzed in more detail when partial or the entire time course is available. The labeling of downstream products, most importantly lactate, can be described using two-site Bloch-McConnell exchange model [3, 7] (*see* **Note 18**):

$$\frac{d_{\mathrm{Pyr}}}{d_t} = -\left(k_{\mathrm{PL}} + \rho_{\mathrm{Pyr}}\right)\mathrm{Pyr} + k_{\mathrm{LP}}\mathrm{Lac}[+\mathrm{Inflow}]$$

$$\frac{d_{\mathrm{Lac}}}{d_t} = -(k_{\mathrm{LP}} + \rho_{\mathrm{Lac}})\mathrm{Lac} + k_{\mathrm{PL}}\mathrm{Pyr}$$

where k_{PL} and k_{LP} are reaction rates for pyruvate → lactate and lactate → pyruvate reactions and ρ_{Pyr} and ρ_{Lac} are the effective longitudinal relaxation rates ($1/T_1$) for pyruvate and lactate, respectively (Fig. 2) (*see* **Note 19**). Additional terms to account for inflow terms (e.g., gamma variate function or boxcar step function) or additional exchanging sites can be easily incorporated into the model. Similar equations can be written for other downstream metabolites. It is worth noting that the loss of carbon dioxide is an irreversible reaction, which, therefore, only proceeds one way. The differential equations can be solved analytically [7] or using numerical methods, and these are considered to be the gold standard (*see* **Note 20**).

Modeling approaches can be used to produce apparent reaction rates with units (s^{-1}). For in vitro experiments, these can often be converted to actual fluxes (e.g., nmol/s/cell) by taking into account the injected pyruvate concentration and the number of cells used [7]. However, the in vivo experiments probe a complex system with multiple compartments and involve both the transport (via monocarboxylate transporters) and metabolic steps. For example, pyruvate needs to be transported inside the cell (e.g., tumor cell or red blood cell), and the lactate may be exported from the cells after the label enters the blood stream. Furthermore, the uptake of marker molecule may be transport-limited, leading to a situation where the signal coming from the marker molecule will be mostly in blood stream, while the signals from downstream products are coming from tissue and/or blood. Finally, it is also worth noting that the visibility of hyperpolarized signal is dependent not only on the absolute metabolite concentration but also on the polarization level and the longitudinal relaxation rates of different metabolites. Because of this, the in vivo results are usually reported as apparent rates with unit s^{-1}. Further work on the metabolic interpretation of the experiments still remains to be conducted, and correct biological interpretation of DNP result therefore often requires careful analysis of potential signal losses and traditional biochemical analysis.

3.4 Beyond [1-^{13}C] Pyruvate Metabolism

While [1-^{13}C]pyruvate experiment is by the far the most used in cancer studies, other hyperpolarized marker molecules can also be used, although these are much less frequently studied. For example, injection of [1,4-$^{13}C_2$]fumaric acid leads to malate labeling in regions with significant necrosis [13], while [1-^{13}C]bicarbonate [14] and [^{13}C]urea [15] can be utilized to assess tumor acidosis and perfusion, respectively. [^{13}C]glucose can also be hyperpolarized to potentially reveal further insights into glycolysis and pentose phosphate pathway [16]. Finally, multiple markers could be used simultaneously to get a more detailed understanding of the metabolism, providing that reasonable separation between resonances can

be sustained [15]. This may open up previously unprecedented window to real-time metabolism.

All metabolites mentioned in the previous paragraph can be polarized using similar methods as those presented for pyruvate (*see* **Note 21**). The most important difference is that most of the other metabolites require a glassing agent (e.g., glycerol) as a part of the hyperpolarized sample for efficient polarization. Similarly, neutral dissolution buffers can be often used, as the original preparation is not highly acidic (of the metabolites mentioned above, only fumarate is polarized as acid form). Similar imaging and analysis methods can also be applied, provided the correct frequencies for each metabolite are used when setting up the experiments.

4 Notes

1. The chemical structure of the molecule leads to variations in electron shielding around the nuclei, thus causing them to resonate at different frequencies (chemical shifts). These chemical shifts are typical for molecules and can therefore be used to, e.g., identify molecules, in case of hyperpolarized experiments pyruvate and its metabolic products. Chemical shifts can be found from, e.g., online databases such as "Human Metabolome Database," http://www.hmdb.ca [17].
2. For spin = 1/2 nuclei such as ^{1}H or ^{13}C, polarization is defined as the population difference between high- and low-energy states and depends on the strength of the main magnetic field and temperature according to the Boltzmann distribution. Under normal conditions, the populations in both states are nearly equal, and polarization is therefore very low (e.g., 8 ppm for ^{13}C at 9.4 T at room temperature), making MRI not sufficiently sensitive. Polarization can be increased by increasing the magnetic field strength or lowering temperature or by using hyperpolarization methodology.

   Dissolution dynamic nuclear polarization (dDNP) methodology is based on microwave-induced signal pre-amplification of the labeled marker molecule, usually ^{13}C, at low temperature (~1 K) followed by a rapid dissolution of sample providing a room-temperature solution of the marker molecule with increased sensitivity. Microwaves induce polarization transfer between free electrons in the radical and nearby ^{13}C. The hyperpolarization is efficiently distributed throughout the amorphous solid sample through spin diffusion. The ^{13}C signal loss is slow in solid state (longitudinal (T_1) relaxation times are long) leading to a signal buildup and a significant increase in the overall signal level.

Alternative means to achieve hyperpolarization include brute-force (high magnetic field and low temperature), laser-driven optical pumping (for hyperpolarized gases), and parahydrogen-induced polarization (PHIP, SABRE), usually achieved by reacting the hyperpolarized parahydrogen with a suitable labeled precursor, yielding the final hyperpolarized marker molecule.

3. The metabolism available for hyperpolarized studies is ultimately limited by T_1 relaxation times of ^{13}C-label in the marker and product molecules because all transport, uptake, and metabolism need to occur during a time period of two to three T_1 relaxation times after which the majority (>90%) of the signal is lost. The majority of markers up to now have T_1 times in the range of 10–50 s (e.g., pyruvate has T_1 of around 50–60 s in liquid and around 30 s in vivo), giving a metabolic time window of up to 3 min. This relatively short time period limits the method to mainly catabolic reactions and possibly the initial steps of anabolic processes. Other significant factors influencing hyperpolarized experiments are the sufficient chemical shift between precursor and metabolites, especially because the precursor signal is often much larger than the products, and the relative nontoxicity of the injected metabolites, because the concentrations used are often several folds higher than those normally are present. For example, a typical preclinical pyruvate dose is injected intravenously at 80 mM and 5–10 mL/kg yielding around 8 mM blood concentration, whereas typical plasma pyruvate concentration is around ~0.2 mM.
4. [1-^{13}C]Pyruvic acid can currently be purchased in non-clinical and clinical quality; the latter is designed for clinical hyperpolarizer studies. Non-clinical quality pyruvic acid is suitable for preclinical research. Pyruvic acid forms an amorphous glass when it is frozen, so it is directly suitable for hyperpolarizer experiments.
5. The best polarization has been achieved using trityl radicals (e.g., OX063, AH15011), which to date have been only available through GE Healthcare. In most cases, there is relatively little difference between OX063 and AH15011 in terms of performance. Other radicals, such as nitroxide-based TEMPO, can also be used, but these usually achieve lower polarization levels. The optimal radical concentration depends on the polarization field and is usually slightly higher at higher magnetic fields (e.g., 20–25 mM at 7 T). Too low radical concentration leads to slow polarization, whereas too high radical concentration leads to rapid polarization but lower final polarization level.

6. Pyruvate and radical are normally stored in a cold, dry place protected from direct light to improve stability. Allow the radical and pyruvic acid to warm up to room temperature before proceeding with the sample preparation. Weigh the radical into a suitable container (e.g., conical tube) and add pyruvic acid. Vortex the sample until a dark green liquid with no visible radical residual is obtained. Sonication is usually not needed to fully dissolve the radical into small stock solutions. Store the stock solution in cold, dry place protected from direct light (e.g., in a desiccator inside −20 °C freezer), preferably as pre-made sample aliquots.
7. Addition of Gd chelate leads to a reduction in electron T_1 relaxation time, which allows higher ^{13}C polarization levels to be achieved at lower magnetic fields (e.g., 3.35 T). The addition of Gd chelate leads to a small reduction on liquid state T_1 relaxation time, but this is greatly offset by the benefits of the higher polarization levels. Due to this effect, however, the Gd concentration should be kept as low as possible. Gd^{3+} ion is also highly toxic, so Gd should always be added in a chelate form to avoid any unexpected biological side effects even at these low concentrations. The beneficial effect of Gd appears to be largely lost at higher polarization fields (4.6 T or more), due to altered polarization conditions [18].
8. Hyperpolarized pyruvic acid sample is highly acidic (pH < 1) and needs to be neutralized and made isotonic as part of the dissolution process. Both of these goals can be achieved using buffered NaOH solutions, in which any biologically compatible buffer can be used. The buffer usually contains some ethylenediaminetetraacetic acid (EDTA) to chelate impurities. TRIS base in the recipe can be replaced with Trizma crystals with preset pH 7.6. An alternative recipe based around HEPES-buffer can be found in [6]: 1 L MilliQ water, 9.54 g (40 mM) 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 3.76 g (94 mM) NaOH, 1.752 g (30 mM) NaCl, and 100 mg EDTA.

   In some cases, it is favorable to separate the dissolution and neutralization steps. For example, AH15011 radical is not highly soluble in acidic conditions and can be filtered out of the sample before the sample is neutralized using NaOH in the collection cup. It is worth noting that OX063 solubility is not highly pH dependent. For preclinical work, simultaneous dissolution and neutralization is usually favored, because of its simplicity.
9. The amounts of dissolution volumes cited here are the typically used volumes for preclinical experiments and give approximately 5 mL of the final polarized sample. This results in a significant loss of sample because a typical injected dose used in

preclinical experiments is 0.1–2 mL and the dissolved sample cannot be easily recycled. The dissolution volume can be decreased (up to 50%) to minimize the dose of pyruvic acid or to obtain higher pyruvate concentrations, as long as rapid dissolution can be achieved. The achievable volume reduction depends on the hyperpolarizer equipment used, however. For example, in Hypersense equipment, the chase gas will remove the whole sample from the system, whereas SpinLab systems do not use a chase gas, so some sample will be always lost as dead volume and significant volume reductions are more difficult to achieve. An alternative approach for the reduction of final sample volume is to replace some of the dissolution buffer with perfluorocarbon solutions, which rapidly separate from water-based buffer after dissolution [19].

10. Microwave irradiation used dDNP experiments is performed around the precession frequency of the electron (e.g., around 94.1 GHz for Hypersense operating at 3.35 T). The polarization frequency yielding the highest signal depends on the properties of a given molecule and the radical used. The polarization efficiency can be optimized by recording a series of hyperpolarized spectra while sweeping the microwave frequency (e.g., in steps of 1 MHz) around the expected maximum and using the polarization frequency giving the highest solid-state signal. While most polarizer systems allow direct monitoring of NMR, this can also be achieved by using the main imaging system for signal detection. The conversion of the obtained signal intensity to absolute polarization is not usually straightforward at this stage.
11. The exact amount of NaOH may vary slightly from system to system depending on, e.g., the dead volume of the dissolution system; therefore, the pH of the final samples needs to be confirmed and the amount of NaOH of dissolution buffer adjusted accordingly. The recipe given in materials is from SpinLab-based dissolution system. The alternative recipe given in **Note 8** has been successfully used in Hypersense-based system.
12. Measurement of polarization levels can be achieved using either a dedicated polarimeter or by comparing the hyperpolarized signal to thermal equilibrium signal measured from the same sample. While the latter is more accurate, as the same sample is used for both measurements, it is very time-consuming, due to low thermal signal; the aforementioned is not easily achieved during in vivo work.
13. Either a constant or variable flip angle can be used for the experiment. In the latter case, the flip angle steadily increases during the data collection to compensate for the loss of the

signal due to previous pulsing, which may improve the data quality. Variable flip angles have been particularly popular with fast imaging sequences.

14. The technical limitations, mainly gradient performance, limit the spatial resolution for spectroscopic imaging. The bandwidth of the collected spectra depends on the repeat rate of the read-out block, leading easily to a situation where the spectral range of the peaks exceeds the technical specifications of the instrument. The sparsity of the ^{13}C spectra allow significant folding of the peaks, which alleviates the situation, but imaging parameters need to be selected carefully to avoid overlapping the peaks of interest.
15. Matlab implementation for spectral-spatial pulse design can be found in "Hyperpolarized MRI Technology Resource Center" website (https://radiology.ucsf.edu/research/labs/hyperpolarized-mri-tech, under topic "TR&D 1: Improved DNP Methodology and HP MR Acquisition Techniques" in Dissemination and Training).
16. Calibration of flip angles is usually performed in a ^{13}C-containing phantom (e.g., ^{13}C-acetate) before the actual hyperpolarized experiment because the sample usually does not contain enough ^{13}C signal before the experiment and there is no time to perform it after the injection. As long as the used phantom is of similar composition, flip angles remain approximately the same between samples. In animals, fat signals (20–40 ppm) can also be used to check flip angles before hyperpolarized experiments provided enough fat is visible to the coil.
17. There is some evidence that, similar to PET experiments, fasting animal prior to the experiments may lead to less variation in the results [20], but further confirmations are required.
18. It has been shown that labeling of lactate signal is mainly driven by the lactate dehydrogenase (LDH)-catalyzed label exchange between the injected [1-^{13}C]pyruvate and the existing tumor lactate pool, which needs to be taken into account when interpreting the results [3, 7].
19. The effective longitudinal relaxation rates include the signal losses from RF pulses, which can be modeled more explicitly if a better estimate of actual relaxation rate is wanted. Equation

$$\rho = \frac{1}{T_{1,\mathrm{eff}}} = \frac{1}{T_1} - \frac{1}{\mathrm{TR}} \ln(\cos\alpha)$$

gives effective relaxation rate for an experiment with a constant flip angle (α) and repetition time (TR). A more accurate estimate of relaxation rates can be achieved by varying either flip angle or relaxation time or both during the scan.

20. A reasonable starting guess for the fitting is required; in vivo k_{PL} exchange rates are usually around 0.05–0.1 s^{-1} and relaxation rates $1/30$ s^{-1}. In vitro exchange rates can be several orders lower due to smaller number of cells. Simultaneous solving of all four unknowns is usually not feasible, so simplifications are made (e.g., relaxation rates are assumed to be equal or k_{LP} is assumed to be 0 s^{-1}). Even after these modifications, the fit can easily become unstable or produce unrealistic values (e.g., negative k_{LP}) especially if the beginning of the time course is not fully sampled. Nevertheless, the value for k_{PL} is usually robustly fitted regardless of the approach taken. The quality of numerical fitting can be improved by making both signals to contribute similarly to the residual, for example, by normalizing the residual data for each signal. This helps especially with in vitro data, wherein the pyruvate signal is much larger than the lactate signal.

    Several derivations of the two-site exchange model, aimed to simplify the estimation, have been proposed. These can be used to either estimate the exchange rates or derive parameters that correlate with the underlying exchange rates. While the approaches vary in detail, and the obtained absolute exchange rates may vary slightly, they all appear to report well on the most significant parameter, the forward pyruvate → lactate reaction rate, k_{PL} [21, 22]. However, unlike the two-site model, these approaches have not yet been used extensively. The most commonly used approaches are summarized below.

    (a) Precursor-product [21]

        A simple precursor-product model only fits lactate response (k_{PL} and $k_{LP} + \rho_{Lac}$) based on pyruvate input. In this approach, the number of unknowns is limited, and the factors such as pyruvate and relaxation do not need an explicit form making the fitting more stable.

    (b) Lactate-pyruvate ratio analysis [23]

        In this approach, a relation between the lactate-pyruvate ratio (R_{LP}) and exchange rates is established after a selected time point, t_0 (e.g., after pyruvate has peaked)

$$R_{LP}(t) = \frac{r[1 + R_{LP}(t_0)] + [R_{LP}(t_0) - r]e^{-s(t-t_0)}}{1 + R_{LP}(t_0) + [r - R_{LP}(t_0)]e^{-s(t-t_0)}}$$

        where $r = k_{PL}/k_{LP}$ and $s = k_{PL} + k_{LP}$. Relaxation terms are effectively removed, and only two terms need to be fitted. Alternatively, a linear fit to data describing R_{LP} and its changing rate can be used to derive exchange rates

$$q = \frac{\frac{dR_{LP}}{dt}}{1 + R_{LP}} = -k_{LP}R_{LP} + k_{PL}$$

Both of these approaches may allow an efficient estimation of exchange rates with only a few time points.

(c) Area-under-the curve analysis [24]

It has been shown that the ratio of areas under lactate and pyruvate correlates with exchange rates by a simple ratio:

$$\frac{\sum \mathrm{Lac}}{\sum \mathrm{Pyr}} = \frac{k_{PL}}{k_{LP} + \rho_{Lac}}$$

The analysis is easy to perform, signal-to-noise ratio can be maximized by summing over the time course, and no assumptions about pyruvate inflow are required. However, no exact value for k_{PL} is directly obtained, and the analysis may become inaccurate if the full-time course cannot be sampled.

(d) Lactate peak analysis

The peak signal amplitude and time of lactate signal are both closely related to k_{PL} and can, therefore, be used as a marker for the transfer. The lactate-pyruvate ratio at the time of the lactate peak has been shown to be equal to the product k_{PL} and effective T_1, while an estimate of T_1 can be obtained by estimating the lactate peak width at around 80% height [25].

$$R_{LP}\left(t_{Lac,max}\right) \approx k_{PL}T_1$$

Alternatively, the time-to-lactate peak (TPP) can be used as a surrogate for k_{PL} activity [22]:

$$\mathrm{TPP} = t_{Lac,max} - t_{Lac,0} = \frac{1}{k_{PL} + k_{LP}} \ln\left[1 + T_1(k_{PL} + k_{LP})\right]$$

21. While there is currently no centralized database for hyperpolarized preparations and new ones are published continuously, a number of most common formulations can be found from "Hyperpolarized MRI Technology Resource Center" website (https://radiology.ucsf.edu/research/labs/hyperpolarized-mri-tech, under topic "Formulation of carbon-13 labeled compounds for dissolution-DNP" in Dissemination and Training).

Acknowledgments

This work was supported by the Academy of Finland grants no. 286895 and no. 314551. The author shares dDNP-related intellectual property with GE Healthcare.

References

1. Ardenkjaer-Larsen JH, Fridlund B, Gram A et al (2003) Increase in signal-to-noise ratio of > 10,000 times in liquid-state NMR. Proc Natl Acad Sci U S A 100:10158–10163
2. Nelson SJ, Kurhanewicz J, Vigneron DB et al (2013) Metabolic imaging of patients with prostate cancer using hyperpolarized [1-^{13}C] pyruvate. Sci Transl Med 5:198ra108
3. Brindle KM, Bohndiek SE, Gallagher FA, Kettunen MI (2011) Tumor imaging using hyperpolarized ^{13}C magnetic resonance spectroscopy. Magn Reson Med 66:505–519
4. Golman K, Zandt RI, Lerche M et al (2006) Metabolic imaging by hyperpolarized ^{13}C magnetic resonance imaging for in vivo tumor diagnosis. Cancer Res 66:10855–10860
5. Albers MJ, Bok R, Chen AP et al (2008) Hyperpolarized ^{13}C lactate, pyruvate, and alanine: noninvasive biomarkers for prostate cancer detection and grading. Cancer Res 68:8607–8615
6. Serrao EM, Kettunen MI, Rodrigues TB et al (2016) MRI with hyperpolarized [1-^{13}C]pyruvate detects advanced pancreatic preneoplasia prior to invasive disease in a mouse model. Gut 65:465–475
7. Day SE, Kettunen MI, Gallagher FA et al (2007) Detecting tumor response to treatment using hyperpolarized ^{13}C magnetic resonance imaging and spectroscopy. Nat Med 13:1382–1387
8. Malinowski RM, Lipso KW, Lerche MH, Ardenkjaer-Larsen JH (2016) Dissolution Dynamic Nuclear Polarization capability study with fluid path. J Magn Reson 272:141–146
9. Ardenkjaer-Larsen JH, Leach AM, Clarke N et al (2011) Dynamic nuclear polarization polarizer for sterile use intent. NMR Biomed 24:927–932
10. Mayer D, Yen YF, Levin YS et al (2010) In vivo application of sub-second spiral chemical shift imaging (CSI) to hyperpolarized ^{13}C metabolic imaging: comparison with phase-encoded CSI. J Magn Reson 204:340–345
11. Wiesinger F, Weidl E, Menzel MI et al (2012) IDEAL spiral CSI for dynamic metabolic MR imaging of hyperpolarized [1-^{13}C]pyruvate. Magn Reson Med 68:8–16
12. Lau AZ, Chen AP, Hurd RE, Cunningham CH (2011) Spectral-spatial excitation for rapid imaging of DNP compounds. NMR Biomed 24:988–996
13. Gallagher FA, Kettunen MI, Hu DE et al (2009) Production of hyperpolarized [1,4-$^{13}C_2$]malate from [1,4-$^{13}C_2$]fumarate is a marker of cell necrosis and treatment response in tumors. Proc Natl Acad Sci U S A 106:19801–19806
14. Gallagher FA, Kettunen MI, Day SE et al (2008) Magnetic resonance imaging of pH in vivo using hyperpolarized ^{13}C-labelled bicarbonate. Nature 453:940–943
15. von Morze C, Larson PE, Hu S et al (2012) Investigating tumor perfusion and metabolism using multiple hyperpolarized ^{13}C compounds: HP001, pyruvate and urea. Magn Reson Imaging 30:305–311
16. Rodrigues TB, Serrao EM, Kennedy BW et al (2014) Magnetic resonance imaging of tumor glycolysis using hyperpolarized ^{13}C-labeled glucose. Nat Med 20:93–97
17. Wishart DS, Feunang YD, Marcu A et al (2018) HMDB 4.0 — the human metabolome database for 2018. Nucleic Acids Res 46: D608–D617
18. Johannesson H, Macholl S, Ardenkjaer-Larsen JH (2009) Dynamic nuclear polarization of [1-^{13}C]pyruvic acid at 4.6 tesla. J Magn Reson 197:167–175
19. Peterson ET, Gordon JW, Erickson MG et al (2011) Dynamic nuclear polarization system output volume reduction using inert fluids. J Magn Reson Imaging 33:1003–1008
20. Serrao EM, Rodrigues TB, Gallagher FA et al (2016) Effects of fasting on serial measurements of hyperpolarized [1-^{13}C]pyruvate metabolism in tumors. NMR Biomed 29:1048–1055
21. Kazan SM, Reynolds S, Kennerley A et al (2013) Kinetic modeling of hyperpolarized ^{13}C pyruvate metabolism in tumors using a measured arterial input function. Magn Reson Med 70:943–953
22. Daniels CJ, McLean MA, Schulte RF et al (2016) A comparison of quantitative methods for clinical imaging with hyperpolarized ^{13}C-pyruvate. NMR Biomed 29:387–399
23. Li LZ, Kadlececk S, Xu HN et al (2013) Ratiometric analysis in hyperpolarized NMR (I): test of the two-site exchange model and the quantification of reaction rate constants. NMR Biomed 26:1308–1320
24. Hill DK, Orton MR, Mariotti E et al (2013) Model free approach to kinetic analysis of real-time hyperpolarized ^{13}C magnetic resonance spectroscopy data. PLoS One 8:e71996
25. Pages G, Kuchel PW (2015) FmRα analysis: rapid and direct estimation of relaxation and kinetic parameters from dynamic nuclear polarization time courses. Magn Reson Med 73:2075–2080

Chapter 22

Overview of Glutamine Dependency and Metabolic Rescue Protocols

Shuo Qie, Dan He, and Nianli Sang

Abstract

Enhanced glutaminolysis and glycolysis are the two most remarkable biochemical features of cancer cell metabolism, reflecting increased utilization of glutamine and glucose in proliferating cells. Most solid tumors often outgrow the blood supply, resulting in a tumor microenvironment characterized by the depletion of glutamine, glucose, and oxygen. Whereas mechanisms by which cancer cells sense and metabolically adapt to hypoxia have been well characterized with a variety of cancer types, mechanisms by which different types of tumor cells respond to a dynamic change of glutamine availability and the underlying importance remains to be characterized. Here we describe the protocol, which uses cultured Hep3B cells as a model in determining glutamine-dependent proliferation, metabolite rescuing, and cellular responses to glutamine depletion. These protocols may be modified to study the metabolic roles of glutamine in other types of tumor or non-tumor cells as well.

Key words Anaplerosis, Cell proliferation, Endoplasmic reticulum stress, Glutamine depletion, Metabolism, Nitrogen anabolism

1 Introduction

Increased utilization of glutamine and glucose is a common biochemical feature of most rapid proliferating cells, indicating crucial metabolic roles in supporting cell division. Whereas glucose plays important roles in carbon anabolism, energy homeostasis, and redox balance [1], glutamine may serve as both a nitrogen source and a carbon source participating in a variety of biosynthesis, actively occurring in all types of proliferating cells [2]. In addition to its direct roles in nitrogen-dependent anabolic processes such as protein translation, nucleotide biosynthesis, and asparagine biosynthesis, glutamine serves as an important precursor of glutamate in most cells. Upon entering the mitochondria, glutamine can be readily converted into glutamate through glutaminolysis catalyzed by glutaminases [2]. A variety of transamination processes involve glutamate as amino group donor, or α-ketoglutarate as the amino

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_22,

group acceptor, to maintain the dynamic homeostasis of the intracellular pool of amino acids. Particularly, proliferating cells have increased demands for glutamate, aspartate, serine, and glycine for the biosynthesis of nucleotides, biomembranes, and glutathiones and require one-carbon units for a variety of cell functions [3–5]. The high levels of intracellular concentrations of glutamine and glutamate also serve chemical potential to power the cross-membrane transport of nutrients and metabolites, which indirectly participates in other cellular processes [6, 7]. The carbon skeleton derived from glutamine catabolism may eventually enter into various carbon metabolic pathways to support the production of NADPH and NADH directly or indirectly through anaplerotic pathways [8–10]. The multiplicity of metabolic fates and complexity of metabolic pathways have been extensively studied by using isotope tracking and mass spectrometry [11–13]; however, determining the indispensable role of glutamine in specific type of cells under specific physiological or pathological context usually demands functional assays.

Under in vivo conditions, solid tumors usually have poorly developed vasculature, leading to a microenvironment characterized by localized hypoxia and nutrient depletion. To survive in such microenvironment, tumor cells usually resort to stress response and metabolic reprogramming, both of which may involve transcriptional control of gene expression and biosignaling [14]. Hypoxia inducible factor (HIF)-mediated transcriptional and metabolic reprogramming are the best known cellular response to hypoxia [15, 16]. Under hypoxic condition, HIF activation upregulates the expression of angiogenic, glycolytic, and other relevant genes that induce the adaptive response to hypoxia [17]. Glutamine, even though classified as a nonessential amino acid at organismal levels, has been shown to be essential for most types of cells in culture [2]. Lack of glutamine has been shown to trigger both general stress responses and metabolic responses specific to glutamine depletion [14]. The general responses observed in past studies include endoplasmic reticulum stress response, cell proliferation inhibition, and activation of heat shock protein system. However, currently available studies are not sufficient to generalize glutamine sensing or signaling pathways. It remains unclear if different tumor cells or under different biological contexts may have different glutamine dependency and may respond to glutamine depletion differently.

In addition to tumor cells, normal cells may also conditionally assume a rapid proliferation status; these include endothelial cells in wound healing, activated T and B cells in responding to pathogens, hematopoietic stem cells, etc. In another scenario, normal tissues may experience ischemia-triggered hypoxia and glutamine depletion. Importance and roles of glutamine utilization in normal cells under these conditions remain to be determined.

We present the protocols we have used to study glutamine metabolism in Hep3B cells, a tumor cell line originated from hepatocytes, which depends on glutamine for proliferation. The protocols are intended for the determination of the glutamine dependency of cell functions and cellular responses to glutamine depletion. We also introduce the method to determine the critical metabolic roles in specific cell types and physiological context by functional rescuing tests. These approaches may be utilized to complement the genetic approaches and mass spectrometry-based tracking of carbon or nitrogen derived from isotope-labeled glutamine.

2 Materials

2.1 Cell Culture and Proliferation Assay Equipment and Kit

1. Hep3B cells (ATCC HB-8064).
2. Water jacketed, humidified cell culture incubator, with 5% CO_2 and 95% air atmosphere, temperature setting at 37 °C.
3. CyQUANT® GR dye (a proprietary product of molecular probes): 500× solution in dimethyl sulfoxide (DMSO) (*see* **Note 1**).
4. Fluorescence microplate reader equipped with excitation wavelength at 485 nm and emission detection at 530 nm.
5. Hank's Balanced Salt Solution (HBSS buffer 1×): NaCl 140 mM, KCl 5 mM, $CaCl_2$ 1 mM, $MgSO_4$-$7H_2O$ 0.4 mM, $MgCl_2$-$6H_2O$ 0.5 mM, Na_2HPO_4 0.3 mM, KH_2PO_4 0.4 mM, glucose 6 mM, and $NaHCO_3$ 4 mM.

2.2 Cell Culture Media and Reagents

1. Regular Dulbecco's modified Eagle's medium (DMEM): with 4.5 g/L glucose (about 25 mM) and 4 mM glutamine.
2. Glutamine-free DMEM: with 4.5 g/L glucose; without glutamine.
3. Penicillin/streptomycin solution (100× stock): penicillin 10,000 units, streptomycin 10,000 μg/mL.
4. 0.25% trypsin.
5. 1× phosphate buffered saline (PBS): 137 mM NaCl, 10 mM Na_2HPO_4, 2.7 mM KCl, 1.76 mM KH_2PO_4, pH 7.4.
6. Dimethyl sulfoxide (DMSO), cell culture grade.

2.3 Cell Culture Sera

1. Regular fetal bovine serum (FBS), heat inactivated.
2. Dialyzed (10 kD cutoff) FBS, heat inactivated (*see* **Note 2**).

2.4 Other Conditional Nutrient Supplements

1. Ammonium sulfate, 0.4 M stock.
2. Glucose, 2 M stock.

3. Glutamine, 0.2 M stock.
4. Sodium pyruvate (100× stock).
5. Nonessential amino acid mix (100× stock): L-alanine 890 mg/L, L-asparagine 1320 mg/L, L-aspartic acid 1330 mg/L, L-glutamic acid 1470 mg/L, glycine 750 mg/L, L-proline 1150 mg/L, and L-serine 1050 mg/L.
6. Dimethyl-α-ketoglutarate (DM-α-KG), 1 M stock.
7. Any other metabolites to be tested.

3 Methods

3.1 Cell Revival and Maintenance

1. Prepare regular DMEM media: warm up a bottle of regular DMEM, and add FBS to 10%, sodium pyruvate to 1 mM (final concentration), and penicillin/streptomycin solution mix to 1×.
2. Thaw cryopreserved cells at 37 °C in water bath with gentle shaking.
3. Immediately after complete thawing, transfer cells into a 15 mL centrifuge tube containing 10 mL of the prepared regular DMEM media. Gently mix.
4. Spin down at 4 °C in a centrifuge, 1000 × *g* for 5 min.
5. Dispose of the supernatant by careful aspiration.
6. Resuspend cells in the regular DMEM culture media prepared in 1.
7. Transfer the cell suspension to a 60 mm or 100 mm cell culture dish.
8. Place the culture dish in a humidified incubator with 5% CO_2/95% air atmosphere at 37 °C.
9. Trypsinize cultured Hep3B cells upon reaching 90% confluence, dilute cells in a ratio of 1:4 with regular DMEM media.
10. Continue maintaining the cells in the regular DMEM media until experiments dictate otherwise (*see* **Notes 3** and **4**).

3.2 Effects of Glutamine Depletion on Gene Expression and Signaling

The effects of acute glutamine depletion on gene expression and signaling are complicated and far-reaching [14]. In most cases, these may include cell stress response, cell cycle arrest, inhibition of the mechanistic target of rapamycin complexes (mTORC), autophagy, and metabolic reprogramming [18–20]. A common starting point to explore these aspects is to expose the cells to glutamine depletion and analyze the change of gene expression and protein markers of relevant cellular processes. The following protocol outlines the overall processes to prepare the RNA and protein samples for further studies. Methods for analyzing gene

expression, enzyme activities, protein markers, and cellular functions are outside the scope of this chapter.

1. Preparing glutamine-free DMEM culture media: warm up a bottle of glutamine-free DMEM with 4.5 g/L glucose. Add dialyzed FBS to 10%. Add sodium pyruvate to a final concentration of 1 mM and penicillin/streptomycin (1×).
2. Trypsinize freshly cultured Hep3B cells reaching 90% confluence, dilute cells in a ratio of 1:4, and culture in 10 cm culture dishes with glutamine-free DMEM prepared in **step 1**, but add glutamine to culture dishes to a final concentration of 4 mM for acclimation overnight (about 12–16 h) (*see* **Note 5**).
3. 24 h later, replace with fresh glutamine-free DMEM.
4. For control cells, add glutamine to 4 mM immediately after the replacement of fresh media (*see* **Note 6**).
5. Continue culturing all the cells for 6–48 h (*see* **Note 7**).
6. At desired time, harvest the cells, prepare protein samples, or isolate RNA for subsequent analysis.
7. Protein samples may be used to detect cellular stress responses, for example, biomarkers of cell signaling, apoptosis, autophagy, and change of metabolic enzymes.
8. RNA samples can be used to detect the expression of specific genes and noncoding RNAs or for non-biased gene expression profiling using microarrays or RNA sequencing.

3.3 Effects of Glutamine Concentration on Cell Proliferation

Overall, glutamine depletion will negatively affect the cell proliferation rate; however, glutamine-independent cell survival and proliferation have also been observed [2, 21]. We provide the following protocol to evaluate the importance of glutamine on the proliferation rates of any cell type to be tested. In addition, this protocol can be used to evaluate the rescuing effects of any metabolite that may potentially substitute glutamine in order to support cell proliferation. Data from these rescuing experiments are expected to provide insight into the metabolic roles of glutamine in a specific type of cells at given physiological conditions.

Considering glutamine may contribute carbon source for anaplerotic reaction in the mitochondria; glutamine depletion may potentially affect activities of metabolic enzymes [2, 22]. As such, the activity of metabolic enzymes may not accurately correlate to cell numbers in this specific context; instead, DNA content determination or assay based on labeled nucleotide incorporation (^{3}H-thymidine or 5-bromo-2′-deoxyuridine) should be a more reliable approach. We use CyQUANT® NF Cell Proliferation Assay Kit to determine cell proliferation rates (*see* **Note 1**). The core component of the kit, the CyQUANT® GR dye, exhibits strong fluorescence enhancement after binding with double stranded DNA (dsDNA).

As DNA content is closely proportional to cell number, the assay is designed to produce a linear analytical response in the range of 100–20,000 cells per well in a 96-well microplate. If absolute cell number determination is desired, a standard curve can be generated from plating out known cell numbers. The relative cell number stands for the ratio of cell number at indicated time to the starting cell number at the time of treatment. If absolute cell number determination is desired, a standard curve can be generated from plating out cells in a range of 100–20,000 cells/well and running the test after cells are attached (about 4–6 h after seeding). We recommend monitoring the cell proliferation daily for 5 consecutive days at precisely 24 h intervals. The protocol is modified as following:

1. Trypsinize freshly cultured Hep3B cells reaching 90% confluence.
2. Resuspend cells in 20 mL glutamine-free DMEM with 10% dialyzed FBS and 4 mM glutamine.
3. Count the cell number using a hemocytometer and an upright microscope or using an automatic cell counter.
4. Dilute the cells to make a suspension of 200 cells/100 μL in glutamine-free DMEM with 10% dialyzed FBS and 4 mM glutamine.
5. Plate 200 cells in 100 μL per well in a 96-well cell culture microplate compatible with the assay kit. Set up at least four wells for each experimental condition and one microplate for each observing day (*see* **Note 8**).
6. Allow cells to recover overnight (around 12 h).
7. After 12 h, take one microplate to measure cell numbers as the basal control.
8. For other microplates, replace regular DMEM media, which contains 4 mM glutamine, with glutamine-free DMEM media, which is supplemented with defined concentrations of glutamine, including 0 mM and 4 mM as the controls (*see* **Note 9**).
9. Follow the protocol provided by the manufacturer of the assay kit to determine the cell numbers (*see* **Note 10**).
10. Based on needs, dilute the 500× CyQUANT® GR dye solution with 1× HBSS buffer to 1× working solution.
11. Remove culture medium gently.
12. Add 50 μL of 1× CyQUANT® GR dye solution to each well of the microplates.
13. Cover the microplates with aluminum foil to avoid lights.
14. Incubate the plates at 37 °C for 30 min.

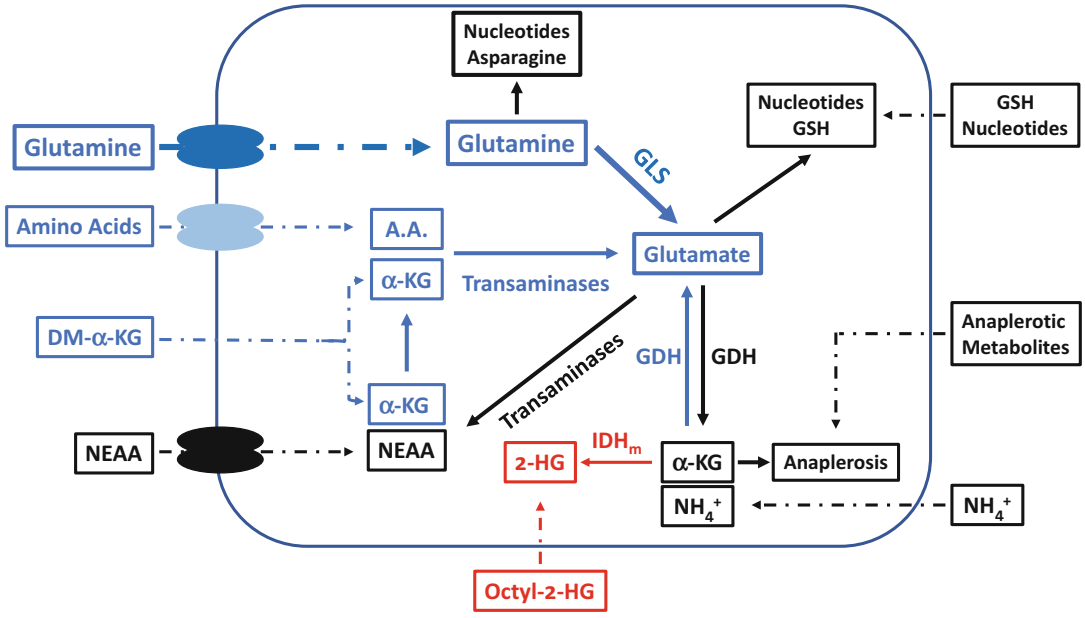


Fig. 1 Diagram showing metabolic pathways of glutamine and potential metabolites or nutrients that may functionally rescue cell proliferation under glutamine depletion Blue arrows and text indicate factors that may facilitate the production of glutamate in cells, and black arrows and text indicate the pathways or metabolites that consume glutamate. Under specific context, one or more of the metabolites may become rate limiting for cell proliferation. By examining the rescuing effects of various relevant metabolites, it will be possible to determine the key metabolic limitation caused by glutamine depletion under specific biological conditions. AA, amino acids; α-KG, α-ketoglutarate; NEAA, nonessential amino acids; GLS, glutaminase; GDH, glutamine dehydrogenase; GSH, glutathione. In addition, glutamine has been found to facilitate the production of 2-hydroxyglutarate (2-HG) in the presence of mutant forms of isocitrate dehydrogenases (IDH$_m$) or hypoxia, which is indicated by red arrows and text. The potential metabolic consequence of 2-HG production may be tested through providing cells with esterified 2-HG in culture media

15. Measure the fluorescence intensity using a fluorescence microplate reader with excitation at 485 nm and emission detection at 530 nm.
16. Analyze the data, perform statistical analysis, and create report.

3.4 Rescuing Glutamine Depletion by Glutamine-Derived Metabolites

The carbon and nitrogen from glutamine can be tracked to enter various metabolic pathways in the cells (*see* Fig. 1). Glutamine serves as a substrate directly in the production of glutamate and asparagine. Through transamination reactions, glutamate may channel the amino groups to a variety of keto acids to produce nonessential amino acids. Glutamine and some of the nonessential amino acids synthesized from nitrogen provided by glutamine are important in nucleotide synthesis and glutathione biosynthesis. In addition to serving as a nitrogen source, glutamine also provides the carbon skeleton α-KG. α-KG may be used in the Krebs cycle in the mitochondrion to produce ATP or in anaplerotic reactions to support a variety of biosynthetic processes. In some tumors

harboring mutations of cytosolic isocitrate dehydrogenases [23, 24] or under hypoxic condition [25, 26], α-KG may undergo abnormal metabolic pathways to create an oncometabolite 2-hydroxyglutarate (2-HG).

Accordingly, the metabolites that may potentially rescue glutamine depletion include glutamate, aspartate, asparagine, alanine, cysteine, serine, glycine, nucleotides (or precursors), glutathione, α-KG, and the other anaplerotic metabolites of the Krebs cycle. We use α-KG as an example in the following rescue protocol. Similarly, other metabolites can also be tested to determine the rate-limiting factors upon glutamine depletion in a specific setting. In addition to cell proliferation, other parameters such as autophagic rate, apoptotic rate, antibody production of cultured B cells, etc. can be used as readouts in rescuing experiments.

Since α-KG cannot diffuse across the cell membrane, dimethyl-α-KG (DM-α-KG) is commonly used to increase the intracellular α-KG level (*see* **Note 11**). Upon entering cells, DM-α-KG is hydrolyzed by endogenous enzymes to α-KG, which may rescue cell proliferation either by facilitating the homeostasis of glutamate (depending on the expression of transaminases and/or glutamate dehydrogenase) (*see* **Note 12**), or by providing anaplerotic carbon metabolites.

1. Trypsinize freshly cultured Hep3B cells reaching 90% confluence.
2. Resuspend cells in 20 mL glutamine-free DMEM with 10% dialyzed FBS and 4 mM glutamine.
3. Count the cell number using a hemocytometer and microscope or using an automatic cell counter.
4. Dilute the cells to make a suspension of 200 cells/100 μL in glutamine-free DMEM with 10% dialyzed FBS and 4 mM glutamine.
5. Plate 200 cells in 100 μL per well in a 96-well cell culture microplate compatible with the assay kit, and place the microplates in cell culture incubator to recover overnight. A total of six microplates are needed.
6. On the following day, take one microplate to measure cell numbers as the basal control.
7. For other five microplates, replace culture media with glutamine-free DMEM and add 0 mM glutamine as the control of glutamine depletion, 4 mM glutamine as the control of normal glutamine supply, and various concentrations of DM-α-KG. A suggested range is from 0.5 to 10 mM.
8. At 24 h interval, take one microplate and measure the cell numbers. Replace culture media for the other microplates (*see* Subheading 3.3, **step 9**).

3.5 Establishing Glutamine-Independent Cells as a Chronic Adaptive Model

Cancer cells usually have multiple ways to obtain carbon source to support the needs for energy, reducing power and biosynthesis. In most cases, glutamine-dependent proliferation is caused by the limitation of nitrogen sources to support the proliferative biosynthesis. As such, some type of tumor cells may adapt to the utilization of alternative nitrogen sources, a feature that heavily depends on the availability of specific type of metabolic enzymes to synthesize glutamate. For example, glutamate dehydrogenase may catalyze the synthesis of glutamate using ammonia and α-KG as substrates [2, 21]. On the other hand, glutamine synthetase is expressed in most cell types; accordingly, by providing sufficient glutamate, most cells may synthesize glutamine to partially compensate for glutamine removal. The following protocol takes advantage of the adequate expression of glutamate dehydrogenase in Hep3B cells to establish a chronic adaptive cell model, which utilizes ammonia to synthesize glutamate in order to support cell proliferation. The procedures may be modified to test other cells or metabolites (*see* **Note 13**).

1. Trypsinize cultured Hep3B cells reaching 90% confluence.
2. Resuspend cells in glutamine-free DMEM with 10% dialyzed FBS and 4 mM glutamine, and reseed the cells with 1:4 split.
3. 24 h later, replace media with glutamine-free DMEM including 10% dialyzed FBS and 0.8 mM ammonia.
4. Keep culturing cells at 37 °C in a humidified incubator with 5% CO_2.
5. Change the culture medium every 2 days, and pass the cells at a ratio (1:2) when reaching 90% confluence.
6. After 6–8 weeks, cell subpopulation will assume ammonia-dependent proliferation (*see* **Note 14**).
7. The cells should be maintained in glutamine-free DMEM supplemented with 10% dialyzed FBS and 0.2–0.4 mM ammonium sulfate (*see* **Note 15**).
8. To characterize these cells, split cells in a ratio of 1:2 or 1:3.
9. Feed cells with fresh media (glutamine-free DMEM with 10% dialyzed FBS and 0.4 mM ammonium sulfate) 24 h before cell harvesting for protein or RNA sample preparation.

4 Notes

1. These experiments are designed to study the metabolic roles of nutrients. Manipulation of culture nutrient may significantly affect the mitochondrial metabolism. Therefore, cell number determination based on measuring the activity of metabolic enzymes (such as succinyl CoA dehydrogenase) may create

artifacts [27, 28]. ^{3}H-thymidine and 5-bromo-2′-deoxyuridine (BrdU) incorporation assays are reliable methods, but these assays require the use of hazardous materials and complicated procedures. As DNA content is closely proportional to cell numbers, we recommend methods based on measuring DNA content as indicators of cell numbers in these experiments. There are several commonly used DNA staining dyes commercially available. In our lab we routinely use dsDNA-specific CyQUANT® GR dye, a proprietary dye supplied in 500× solution, to determine cell numbers, becasue it gives satisfactory accuracy, repeatability, and sensitivity [29]. Other staining dyes of nucleic acids may be optimized and used in combination with RNase treatment to quantify DNA contents. Some dyes may require permeating the cells or lysing the cells.

2. Regular FBS contains various levels of glucose and amino acids; therefore, to analyze the nutrient effects, it is critical to use dialyzed FBS to exclude these variable parameters. In addition, for these metabolic or nutrient dependency studies, batch to batch variations of sera may create artifacts. It is strongly recommended to order sufficient sera of the same batch from the same vendor for all set of experiments to provide a good control.
3. All cells are maintained in humidified incubator with 5% CO_2, 95% air atmosphere at 37 °C, and the cells should be split upon reaching 90% confluence. For Hep3B cells used in most cases, we usually split twice a week with a 1:4 ratio. Other fast proliferating cells may need to be split more frequently or at a lower ratio (1:5–1:8).
4. Unless freshly purchased from ATCC, tumor cell lines should be authenticated prior to the start of experiments. The use of commercially available ones is highly recommended, for it provides more reliable results. The importance of authentication is to confirm the identity of the cell lines and to prove that they are free of any contamination. This is crucial for reproducibility of findings.
5. If immunofluorescent studies are planned, 1:5 and 1:10 dilution should be used, and cells should be reseeded in chamber slides.
6. We do not recommend the use of regular culture media as a control. To use the same bottle of media for both control and experimental groups can minimize potential artifacts and can produce the most reliable results.
7. It is recommended to examine cells at various time points after exposure to glutamine depletion. To plan this experiment, it is important to carefully design the experiment in advance and determine the total dishes of cells will be needed. If more than

four dishes are needed, trypsinize two dishes (or more dishes as needed) of actively proliferating cells, and thoroughly mix them together. Use the mixed cell suspension to seed all dishes that are to be used in the whole set of experiments.

8. Initial plating cell numbers may be adjusted experientially to 100–500 cells per well, based on the proliferation rate of the cell type to be tested. Fast proliferating cells may reach the plateaus too fast if seeded at a high density.
9. Normal cell culture media uses 4 mM of glutamine, which represents the optimized condition for cell proliferation in vitro. Under physiological conditions, normal tissues may reach 50 μM glutamine, and due to poor vascularization in solid tumors, tumor cells may be exposed to glutamine concentrations below 50 μM. A careful titration of glutamine concentration between 0 and 50 μM may be more relevant to in vivo situations; but it remains important to include 4 mM glutamine as a key reference point of optimal proliferation.
10. It is important to note that keeping the precise 24 h intervals between cell number determinations will make the proliferation curves more reliable and more reproducible. At the exponential proliferating phase, several hours make a lot of differences in cell numbers.
11. The cell's ability to uptake the rescuing metabolites should be taken into consideration in experimental design. For amino acids, the efficiency of cell surface transporters may affect rescuing effect. Sometimes, it may be necessary to confirm the uptake efficiency for data interpretation. For example, most cells cannot uptake glutamate efficiently, and most organic acids need certain type of modification to increase the uptake. As a rule of thumb, when anaplerotic metabolites are tested, properly esterified precursors that can easily enter the cells should be used.
12. The utilization of certain metabolites depends on the expression of relevant metabolic enzymes in the cells. In certain cases, it may be necessary to genetically engineer the cells to express the required enzymes prior to the rescuing test.
13. By using the same procedures, we have created a glutamine-independent HeLa line, which overexpresses alanine transaminase and is able to proliferate perpetually in glutamine-free media supplemented with 4 mM of alanine.
14. Ammonium-dependent proliferation can be demonstrated by comparing the cell proliferation rates of the cells cultured in glutamine-free DMEM plus 10% dialyzed FBS with and without ammonia supplement, respectively.
15. The cells can be cryopreserved as regular cells in dialyzed FBS with 20% DMSO.

Acknowledgment

Research in Dr. Sang's laboratory at Drexel University are supported by grants from NCI (R01-CA129494) and the National Natural Science Foundation of China (81470134) and start-up fund from Drexel University.

References

1. Yin C, Qie S, Sang N (2012) Carbon source metabolism and its regulation in cancer cells. Crit Rev Eukar Gene Expr 22(1):17–35. https://doi.org/10.1615/CritRevEukarGeneExpr.v22.i1.20
2. Meng M, Chen S, Lao T, Liang D, Sang N (2010) Nitrogen anabolism underlies the importance of glutaminolysis in proliferating cells. Cell Cycle 9(19):3921–3932. https://doi.org/10.4161/cc.9.19.13139
3. Possemato R, Marks KM, Shaul YD, Pacold ME, Kim D, Birsoy K, Sethumadhavan S, Woo H-K, Jang HG, Jha AK (2011) Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. Nature 476(7360):346–350. https://doi.org/10.1038/nature10350
4. Locasale JW (2013) Serine, glycine and one-carbon units: cancer metabolism in full circle. Nat Rev Cancer 13(8):572–583. https://doi.org/10.1038/nrc3557
5. Amelio I, Cutruzzolá F, Antonov A, Agostini M, Melino G (2014) Serine and glycine metabolism in cancer. Trends Biochem Sci 39(4):191–198. https://doi.org/10.1016/j.tibs.2014.02.004
6. Bannai S, Ishii T (1988) A novel function of glutamine in cell culture: utilization of glutamine for the uptake of cystine in human fibroblasts. J Cell Physiol 137(2):360–366. https://doi.org/10.1002/jcp.1041370221
7. Timmerman LA, Holton T, Yuneva M, Louie RJ, Padró M, Daemen A, Hu M, Chan DA, Ethier SP, vant Veer LJ (2013) Glutamine sensitivity analysis identifies the xCT antiporter as a common triple-negative breast tumor therapeutic target. Cancer Cell 24(4):450–465. https://doi.org/10.1016/j.ccr.2013.08.020
8. Wise DR, Thompson CB (2010) Glutamine addiction: a new therapeutic target in cancer. Trends Biochem Sci 35(8):427–433. https://doi.org/10.1016/j.tibs.2010.05.003
9. Le A, Lane AN, Hamaker M, Bose S, Gouw A, Barbi J, Tsukamoto T, Rojas CJ, Slusher BS, Zhang H (2012) Glucose-independent glutamine metabolism via TCA cycling for proliferation and survival in B cells. Cell Metab 15(1):110–121. https://doi.org/10.1016/j.cmet.2011.12.009
10. Altman BJ, Stine ZE, Dang CV (2016) From Krebs to clinic: glutamine metabolism to cancer therapy. Nat Rev Cancer 16(10):619–180. https://doi.org/10.1016/j.trecan.2017.01.005
11. DeBerardinis RJ, Mancuso A, Daikhin E, Nissim I, Yudkoff M, Wehrli S, Thompson CB (2007) Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. Proc Natl Acad Sci U S A 104(49):19345–19350. https://doi.org/10.1073/pnas.0709747104
12. Metallo CM, Gameiro PA, Bell EL, Mattaini KR, Yang J, Hiller K, Jewell CM, Johnson ZR, Irvine DJ, Guarente L (2012) Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature 481(7381):380–384. https://doi.org/10.1038/nature10602
13. Fendt S-M, Bell EL, Keibler MA, Olenchock BA, Mayers JR, Wasylenko TM, Vokes NI, Guarente L, Vander Heiden MG, Stephanopoulos G (2013) Reductive glutamine metabolism is a function of the α-ketoglutarate to citrate ratio in cells. Nat Comm 4:2236. https://doi.org/10.1038/ncomms3236
14. Qie S, Liang D, Yin C, Gu W, Meng M, Wang C, Sang N (2012) Glutamine depletion and glucose depletion trigger growth inhibition via distinctive gene expression reprogramming. Cell Cycle 11(19):3679–3690. https://doi.org/10.4161/cc.21944
15. Chen S, Yin C, Lao T, Liang D, He D, Wang C, Sang N (2015) AMPK-HDAC5 pathway facilitates nuclear accumulation of HIF-1α and functional activation of HIF-1 by deacetylating Hsp70 in the cytosol. Cell Cycle 14(15):2520–2536. https://doi.org/10.1080/15384101.2015.1055426
16. Chen S, Sang N (2016) Hypoxia-inducible factor-1: a critical player in the survival strategy of stressed cells. J Cell Biochem 117(2):267–278. https://doi.org/10.1002/jcb.25283

17. Semenza GL (1999) Regulation of mammalian O2 homeostasis by hypoxia-inducible factor 1. Ann Rev Cell Dev Biol 15(1):551–578. https://doi.org/10.1146/annurev.cellbio.15.1.551
18. Cairns RA, Harris IS, Mak TW (2011) Regulation of cancer cell metabolism. Nat Rev Cancer 11(2):85–95. https://doi.org/10.1038/nrc2981
19. Davie E, Forte GM, Petersen J (2015) Nitrogen regulates AMPK to control TORC1 signaling. Curr Biol 25(4):445–454. https://doi.org/10.1016/j.cub.2014.12.034
20. Jewell JL, Kim YC, Russell RC, Yu F-X, Park HW, Plouffe SW, Tagliabracci VS, Guan K-L (2015) Differential regulation of mTORC1 by leucine and glutamine. Science 347:194–198. https://doi.org/10.1126/science.1259472
21. Spinelli JB, Yoon H, Ringel AE, Jeanfavre S, Clish CB, Haigis MC (2017) Metabolic recycling of ammonia via glutamate dehydrogenase supports breast cancer biomass. Science 358 (6365):941–946. https://doi.org/10.1126/science.aam9305
22. Gaglio D, Metallo CM, Gameiro PA, Hiller K, Danna LS, Balestrieri C, Alberghina L, Stephanopoulos G, Chiaradonna F (2011) Oncogenic K-Ras decouples glucose and glutamine metabolism to support cancer cell growth. Mol Sys Biol 7(1):523. https://doi.org/10.1038/msb.2011.56
23. Reitman ZJ, Jin G, Karoly ED, Spasojevic I, Yang J, Kinzler KW, He Y, Bigner DD, Vogelstein B, Yan H (2011) Profiling the effects of isocitrate dehydrogenase 1 and 2 mutations on the cellular metabolome. Proc Natl Acad Sci 108(8):3270–3275. https://doi.org/10.1073/pnas.1019393108
24. Seltzer MJ, Bennett BD, Joshi AD, Gao P, Thomas AG, Ferraris DV, Tsukamoto T, Rojas CJ, Slusher BS, Rabinowitz JD (2010) Inhibition of glutaminase preferentially slows growth of glioma cells with mutant IDH1. Cancer Res 70(22):8981–8987. https://doi.org/10.1158/0008-5472
25. Intlekofer AM, Dematteo RG, Venneti S, Finley LW, Lu C, Judkins AR, Rustenburg AS, Grinaway PB, Chodera JD, Cross JR (2015) Hypoxia induces production of L-2-hydroxyglutarate. Cell Metab 22(2):304–311. https://doi.org/10.1016/j.cmet.2015.06.023
26. Oldham WM, Clish CB, Yang Y, Loscalzo J (2015) Hypoxia-mediated increases in L-2-hydroxyglutarate coordinate the metabolic response to reductive stress. Cell Metab 22 (2):291–303. https://doi.org/10.1016/j.cmet.2015.06.021
27. Quent V, Loessner D, Friis T, Reichert JC, Hutmacher DW (2010) Discrepancies between metabolic activity and DNA content as tool to assess cell proliferation in cancer research. J Cell Mol Med 14(4):1003–1013. https://doi.org/10.1111/j.1582-4934.2010.01013.x
28. Wang P, Henning SM, Heber D (2010) Limitations of MTT and MTS-based assays for measurement of antiproliferative activity of green tea polyphenols. PLoS One 16(4): e10202. https://doi.org/10.1371/journal.pone.0010202
29. Jones LJ, Gray M, Yue ST, Haugland RP, Singer VL (2001) Sensitive determination of cell number using the CyQUANT® cell proliferation assay. J Immunol Meth 254 (1-2):85–98. https://doi.org/10.1016/S0022-1759(01)00404-5

Chapter 23

Integration of Metabolomics and Transcriptomics to Identify Gene-Metabolite Relationships Specific to Phenotype

Andrew Patt, Jalal Siddiqui, Bofei Zhang, and Ewy Mathé

Abstract

Metabolomics plays an increasingly large role in translational research, with metabolomics data being generated in large cohorts, alongside other omics data such as gene expression. With this in mind, we provide a review of current approaches that integrate metabolomic and transcriptomic data. Furthermore, we provide a detailed framework for integrating metabolomic and transcriptomic data using a two-step approach: (1) numerical integration of gene and metabolite levels to identify phenotype (e.g., cancer)-specific gene-metabolite relationships using IntLIM and (2) knowledge-based integration, using pathway overrepresentation analysis through RaMP, a comprehensive database of biological pathways. Each step makes use of publicly available R packages (https://github.com/mathelab/IntLIM and https://github.com/mathelab/RaMP-DB), and provides a user-friendly web interface for analysis. These interfaces can be run locally through the package or can be accessed through our servers (https://intlim.bmi.osumc.edu and https://ramp-db.bmi.osumc.edu). The goal of this chapter is to provide step-by-step instructions on how to install the software and use the commands within the R framework, without the user interface (which is slower than running the commands through command line). Both packages are in continuous development so please refer to the GitHub sites to check for updates.

Key words Metabolomics, Gene expression, Pathway analysis, Network, Omics integration, R packages, Gene, Metabolite

1 Introduction

Metabolomics is a powerful approach for uncovering disease biomarkers and for elucidating biological mechanisms that become disrupted by disease. As a result, metabolomics data is increasingly collected in large cohorts, as evidenced in large initiatives such as the Metabolomics Workbench [1] and the COnsortium of METabolomics Studies (https://epi.grants.cancer.gov/comets/). At the same time, the advancement of high-throughput technologies is producing a deluge of data to address clinical and translational research questions, thereby requiring bioinformatic and computational approaches that can manage, analyze, and interpret the data [2]. With this in mind, metabolomics data is increasingly integrated

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_23,

with other omics data, such as gene expression data, to fully exploit and improve interpretability of metabolomic profiles [3]. More specifically, metabolomics and transcriptomics integration helps to elucidate mechanisms that drive diseases and also to uncover putative biomarkers (metabolites) and targets (genes).

2 Materials

Integration approaches can be broadly characterized as numerical, knowledge-based, or a combination of both. In this section, we describe current approaches and tools within each category, with an emphasis on approaches that are available as R packages. While the focus is on integration of gene expression and metabolomics data, it is worth noting that most of these methods can be applied to a wider array of omics combinations, so long as any assumptions regarding the distribution of the data (e.g., normality) are met. In this section, we highlight how the combination of our linear modeling-based (numerical) approach, IntLIM (Integration through Linear Modeling), and our pathway enrichment (knowledge-based) approach, RaMP, uncovers putative gene-metabolite dependencies, which can be prominently enriched in cancer-related pathways.

2.1 Numerical Approaches

Gene expression and metabolomics data are commonly analyzed independently via univariate analysis methods, such as t-tests or analysis of variance (ANOVA). However, in these types of analysis, each feature (e.g., gene, metabolite) is considered to be independent, and the relationship between features, including potentially crucial biological information, is lost. To complement results from univariate analyses and to assess how the relationships between features change in response to a phenotype (e.g., cancer type), multivariate approaches can be applied. These methods can be categorized as supervised or unsupervised. In supervised analyses, the phenotype information is utilized, and the goal is typically to identify the set of features that best separate groups by phenotype. In unsupervised analyses, the phenotype is not utilized, thereby allowing one to assess the overall structure of the data, which may or may not relate to phenotype. Unsupervised analyses are typically applied to uncover substructures in a dataset (e.g., a subgroup of samples with similar metabolite abundances) and are useful for quality control (e.g., to evaluate possible batch effects).

The R package mixOmic*s* provides access to a wide range of multivariate analysis techniques, including principal component analysis, independent component analysis, partial least squares, and canonical correlation analysis [4]. The DiffCorr package also supports integration of multi-omics data, allowing the user to

identify pattern changes between two experimental conditions in correlation networks [5].

Other numerical integration methods focus on identifying pairwise relationships between genes and metabolites and involve calculating global correlations. Identifying these relationships is biologically relevant because functionally related genes and metabolites show coherent co-regulation patterns [6, 7]. Of note, these gene-metabolite correlations can be used to develop correlation networks, which can then be compared between different phenotypes [8, 9]. Alternatively, logistic regression models that combine gene and metabolite levels can be applied to predict phenotype.

Most numerical approaches assess global gene-metabolite relationships or use gene and metabolite levels to predict phenotype. However, these methods do not directly test whether associations between genes and metabolites depend upon phenotype. This distinction is important because global associations between genes and metabolites may not only reflect one phenotype of interest but could reflect other features (e.g., environment, histology). The discordant method [10] addresses phenotype specificity of gene-metabolite relationships although they do not capture pairs of features that are correlated in one group and not correlated in another group. These approaches are also not necessarily implemented into user-friendly frameworks.

To address these limitations, we developed IntLIM [11], which is publicly available as an R package with an associated user-friendly RShiny web application (https://github.com/mathelab/IntLIM, https://intlim.bmi.osumc.edu). IntLIM is based on a linear modeling approach that integrates gene expression and metabolite data to identify gene-metabolite associations that significantly differ by phenotype (e.g., a gene-metabolite pair that is positively correlated in one group but negatively or not correlated in another group). Specifically, the following linear model is applied to all possible gene-metabolite pairs: $metabolite \sim gene + phenotype + gene{:}phenotype$. Significant p-values of the interaction coefficient (*gene:phenotype* term) denotes phenotype-specific gene-metabolite associations by testing the hypothesis that a gene-metabolite association is significantly altered by phenotype. We have demonstrated the utility of IntLIM to identify relevant tumor-specific gene-metabolite relationships involved in known cancer-related pathways (e.g., glutamine metabolism) [11].

2.2 Knowledge-Based Approaches

Numerical integration approaches typically fail to capture complex and indirect relationships between transcripts and metabolites. For example, nonlinear reaction kinetics, metabolite-metabolite connections that regulate metabolite levels, and posttranslational modifications all contribute to the complexity of gene-metabolite relationships. Knowledge-based approaches can help capture these complex relationships. Fortunately, a multitude of databases store

annotations and biological pathways related to genes and metabolites including KEGG [12–14], Reactome [15], WikiPathways [16–18], HMDB [19–21], SMPDB [22, 23], BioCyc [24], and Pathway Commons [25], among others.

Several user-friendly tools incorporate gene and metabolite annotations for pathway enrichment or network analyses. MetaboAnalyst [26] is a comprehensive metabolomics suite incorporating annotations from KEGG [12–14], HMDB [19–21], and SMPDB [22, 23] and allowing for metabolite overrepresentation analysis, integrated topology and enrichment analysis, and integrated metabolite and gene analysis. IMPaLa [27] incorporates biological pathway information on genes and metabolites from 11 different databases and supports overrepresentation analysis given a list of genes and metabolites. Other large-scale efforts, such as Pathway Commons [25], integrate many databases to facilitate pathway enrichment and network analyses and/or method development. Notably, they do not currently integrate HMDB, which is one of the most comprehensive public repositories for metabolite annotations.

Other tools provide a wider array of analyses, from raw metabolomics data preprocessing to pathway analyses. Of these, XCMS [28, 29] uses information from METLIN [30], KEGG [12–14], HMDB [19–21], LIPID MAPS [31], NIST [32], and MassBank [33], to perform predictive pathway analysis on direct metabolite features obtained from raw mass spectrometry data. XCMS is available as an R package or through an online interface. Metabox [34] performs network analysis using an internal graph database incorporating data from KEGG [12–14], PubChem [35], UniProt [36], ENSEMBL [37], and Pathway Commons [25]. Metabox also supports overrepresentation analysis using KEGG and functional class scoring of annotation terms. Metabox is publicly available as an R package and can be run in a web browser.

While these tools offer numerous benefits to the research communities, there are still areas of development needed. First, the underlying integrated databases used for analysis are typically not accessible for inexperienced programmers, as they frequently require the use of an application programming interface (API) or database programming languages (e.g., MySQL). These technical hurdles make it difficult to perform more complicated batch queries or to leverage the information contained in integrated databases with more advanced analysis tools. Furthermore, especially in cases of pathway analysis, there is an assumption that these pathways are independent of one another. This assumption fails to recognize that pathways can be hierarchical, as is the case with KEGG [12–14] and Reactome [15] pathways. For instance, some biological pathways are completely embedded in larger pathways yet are listed separately (e.g., DNA methylation is a subpathway of gene expression). In addition, we have demonstrated that pathways from one database often overlap with pathways from other databases [38].

RaMP (Relational database of Metabolomics Pathways) has been developed to address some of these concerns [38]. RaMP integrates pathway-level information from KEGG [12–14], Reactome [15], WikiPathways [16–18], and HMDB [19–21] into a relational MySQL database. The underlying MySQL dump file is downloadable so that it can be readily integrated with other tools. Furthermore, all the code used to create the database is publicly available (https://github.com/mathelab/ramp-backend), and we provide an interface to the RaMP database through an R package, which includes a user-friendly RShiny web interface (https://github.com/mathelab/ramp-db, https://ramp-db.bmi.osumc.edu). RaMP supports simple and batch queries and supports pathway enrichment analysis of both genes and metabolites, including functional clustering of pathway-level results. This clustering capability groups enriched pathways that share overlapping genes or metabolites (based on a user-defined cutoff of overlap), thereby accounting for the hierarchical nature of many pathway annotations and facilitating interpretation of results when highly overlapping pathways are significant.

3 Methods

3.1 Identifying Phenotype-Specific Gene-Metabolite Pairs Using IntLIM

We demonstrate how to apply IntLIM [11] to *integrate* metabolomics and gene expression data collected in breast tumor and adjacent non-tumor tissue [39]. Using IntLIM, we identify gene-metabolite associations that differ in tumor tissue vs. non-tumor tissue. The formatted gene expression and metabolomics data are available through GitHub (https://github.com/Mathelab/BreastCancerAmbs_GeneMetabolite_Data). Please note that each function contains a documentation that can be accessed by preceding the function name with a question mark (e.g.,? function_name) in the R console.

3.1.1 Installing IntLIM

Running IntLIM requires the user to have R version 3.2.0 or higher. (The latest version of R can be downloaded from https://cloud.r-project.org/). RStudio offers a user-friendly interface to the R console and is available for download at https://www.rstudio.com/products/rstudio/download3/. Prior to installing IntLIM, it is necessary to install the MultiDataSet package [40]. This can be done via the following commands in R.

```
> source("https://bioconductor.org/biocLite.R")
> biocLite("MultiDataSet")
```

The next step is to install IntLIM from the GitHub repository. This requires the install_github() function from the devtools package [41] available in CRAN. The installation of both devtools and IntLIM is accomplished via the following lines.

```
> install.packages("devtools")
> library(devtools)
> install_github("mathelab/IntLIM")
```

IntLIM is then loaded via the library() function:

```
> library(IntLIM)
```

3.1.2 Inputting Data

IntLIM requires four input files:

1. geneData: a gene expression data file (rows are gene IDs, columns are sample IDs),
2. metabData: a metabolite abundance data file (rows are metabolite IDs, columns are sample IDs),
3. sampleMetaData: a sample meta-data file (rows are sample IDs, columns are phenotypic or clinical variables),
4. an "input.csv" file that provides file names of other input files.

Note that all input files should be placed in one folder, and they should be formatted as comma-separated values (CSV) files. Optionally, users can also input metadata information related to genes and metabolites, where rows are genes or metabolites.

For the metabData and geneData files, the first row contains the gene-metabolite IDs, and the first column contains the sample IDs. For the sample metadata file, the first column is assumed to be the sample ID, which should match the first row of metabolite and gene expression data. Importantly, it is required that all sample IDs in the metabolite data and gene expression are also in the sample metadata file. The sample metadata contains clinical or phenotypic information about the different samples. In our breast cancer dataset, the sample metadata contains a "DIAG" (or diagnosis) column, describing whether the sample is "TUMOR" or "NORMAL."

IntLIM additionally requires an "input.csv" file that lists the names of all other input files. The directory of all files listed in "input.csv" is assumed to be the same as the "input.csv." This input file contains two required columns: "type" and "filenames." The "type" corresponds to the file types, as listed above: metabData, geneData, and sampleMetaData are required, and metabMetaData and geneMetaData are optional. The "filename" column contains the name of the corresponding data CSV files. For convenience, all formatted files for the breast cancer data are available at https://github.com/Mathelab/BreastCancerAmbs_GeneMetabolite_Data. These formatted files allow users to readily try using IntLIM, and the files can be used to run all the code present in this protocol.

The ReadData() function reads the data into R. The location of the CSV meta-file that lists the location of the other files ("input.csv") should be passed to the filename parameter of the ReadData()

function. If meta-information is input through geneMetaData and metabMetaData files, then the "metabid" and "geneid" parameters should also be specified. These IDs correspond to the name of the column from metabolite/gene metadata to be used as IDs, and these IDs must match the metabolite/gene abundance data.

```
> inputData <- ReadData('input.csv', metabid='id', geneid='id')
[1] "CreateMultiDataSet created"
```

The ReadData() function creates a MultiDataSet [40] object containing both gene expression and metabolomics data. The ShowStats() function produces a summary of the data read in:

```
> ShowStats(inputData)
    Num_Genes          Num_Metabolites     Num_Samples_withGeneExpression
1   20254              536                 108
    Num_Samples_withMetabolomics     Num_Samples_inCommon
1   132                              108
```

Using the example dataset noted above from the breast cancer study, we have read in abundances from 20,254 genes in 108 samples and abundances from 536 metabolites in 132 samples. A total of 108 samples contain both gene expression and metabolomics data.

3.1.3 Filtering Data

Users can filter data by removing metabolites or genes with low abundances or metabolites that contain a high percentage of imputed values. For metabolomics data, we assume imputed values are imputed by a minimum value observed for a given ion or metabolite across all samples in the dataset. The FilterData() function produces a new MultiDataSet object with filtered gene and metabolite data. Here we chose to remove 10% of the lowest expressing genes and metabolites with more than 80% imputed values.

```
> inputDatafilt <- FilterData(inputData,geneperc=0.10, metabmiss = 0.80)
[1] "No metabolite filtering by percentile is applied"
> ShowStats(inputDatafilt)
    Num_Genes          Num_Metabolites     Num_Samples_withGeneExpression
1   18228              379                 108
    Num_Samples_withMetabolomics          Num_Samples_inCommon
1   132                                   108
```

The filtered dataset object contains 18,228 genes and 379 metabolites. To write the filtered dataset to a CSV file:

```
> OutputData(inputData=inputDatafilt,filename="FilteredData.zip")
```

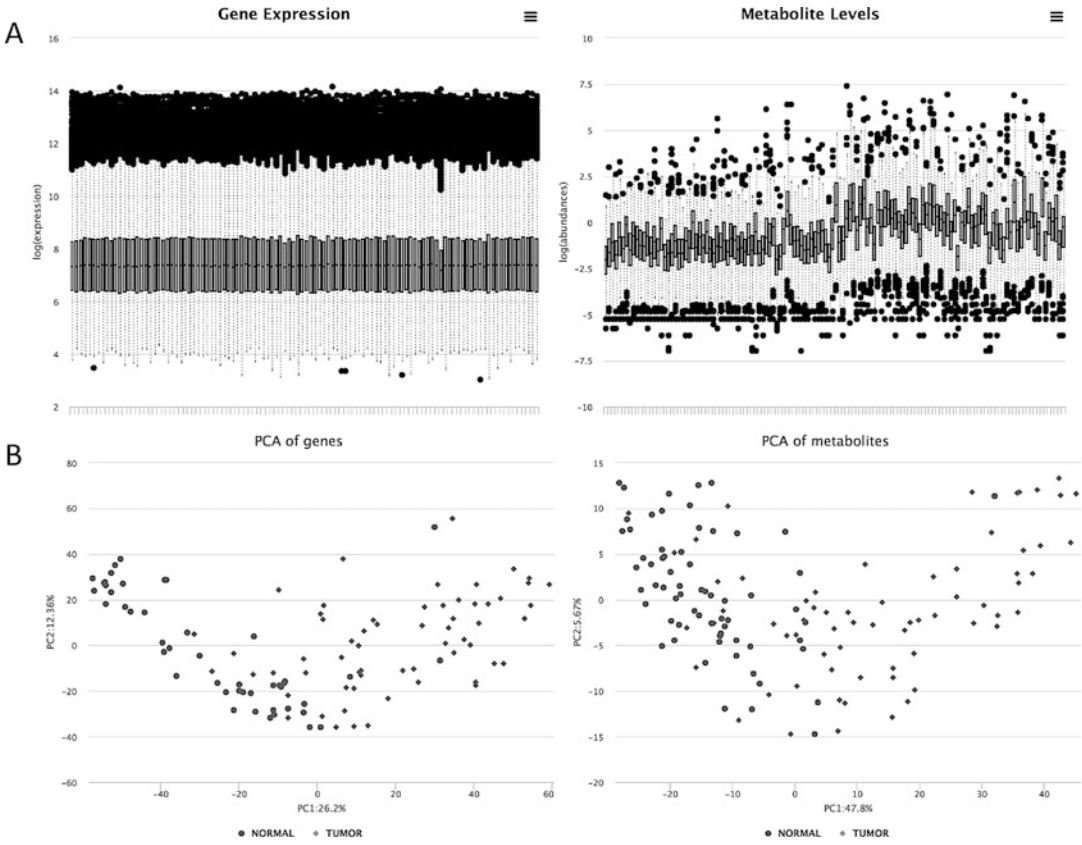


Fig. 1 Visualization of breast cancer gene expression and metabolite data on IntLIM. (**a**) Boxplot distributions (**b**) PCA plots

3.1.4 Inspecting Data

Prior to running linear models, we suggest that users visualize their data to identify potential artifacts (e.g., batch effects) or outliers. The PlotDistributions() function produces boxplots of gene and metabolite levels. Users can define the color scheme for plotting genes and metabolites via the "palette" parameter (Fig. 1a).

```
> PlotDistributions(inputDatafilt, palette = c("black", "black"))
```

The PlotPCA() function produces a principal components analysis plot (PCA) by plotting the data upon the two principal components capturing the highest amount of variation within the data. PCA reduces dimensionality of the data and allows the user to observe how different samples cluster with regard to one in high dimensional space. Users can color-code individual samples by phenotype by specifying the "stype" parameter. Our "stype" is "DIAG" or diagnosis of tissue, which is either "TUMOR" or "NORMAL." The "common" parameter specifies whether all samples, or only the samples in

common between the gene expression and metabolomics data, should be plotted.

```
> PlotPCA(inputDatafilt, stype = "DIAG", common = F)
```

From the PCA plots (Fig. 1b), we observe a clear separation of both tumor and non-tumor samples in both gene expression and metabolomics data and do not detect any clear outliers.

3.1.5 Running Linear Models

The RunIntLim() function computes linear models on all possible gene-metabolite pairs. The function takes a MultiDataSet object (containing gene and metabolite data) and the "stype" (e.g., phenotype from sample metadata) as input. For the breast cancer analysis, we use the filtered gene expression and metabolite data (inputDatafilt object) and the "stype" of "DIAG" (diagnosis of tissue as "TUMOR" or "NORMAL"). The RunIntLim() function calculates the *p*-values of the statistical interaction term gene/phenotype and adjusts the *p*-values by the false discovery rate (FDR) method.

```
> myres <- RunIntLim(inputDatafilt, stype="DIAG")
[1] "Running the analysis on"

NORMAL TUMOR
 47 61
[1] "10 % complete"
[1] "20 % complete"
[1] "30 % complete"
[1] "40 % complete"
[1] "50 % complete"
[1] "60 % complete"
[1] "70 % complete"
[1] "80 % complete"
[1] "90 % complete"
 user system elapsed
294.708 29.973 325.436
```

The results are stored in an "IntLimResults" object, named "myres" here, which includes the following slots: "interaction.pvalues," "interaction.adj.pvalues," and "filt.results." The "interaction.pvalues" and "interaction.adj.pvalues" slots contain matrices of interaction *p*-values and FDR-adjusted interaction *p*-values of gene-metabolite pairs, respectively. We can visualize the raw interaction *p*-values calculated via the DistPvalues() function.

```
> DistPvalues(myres)
```

From the distribution of *p*-values (Fig. 2a), we see a peak close to 0 indicating gene-metabolite pairs that significantly differ based on phenotype.

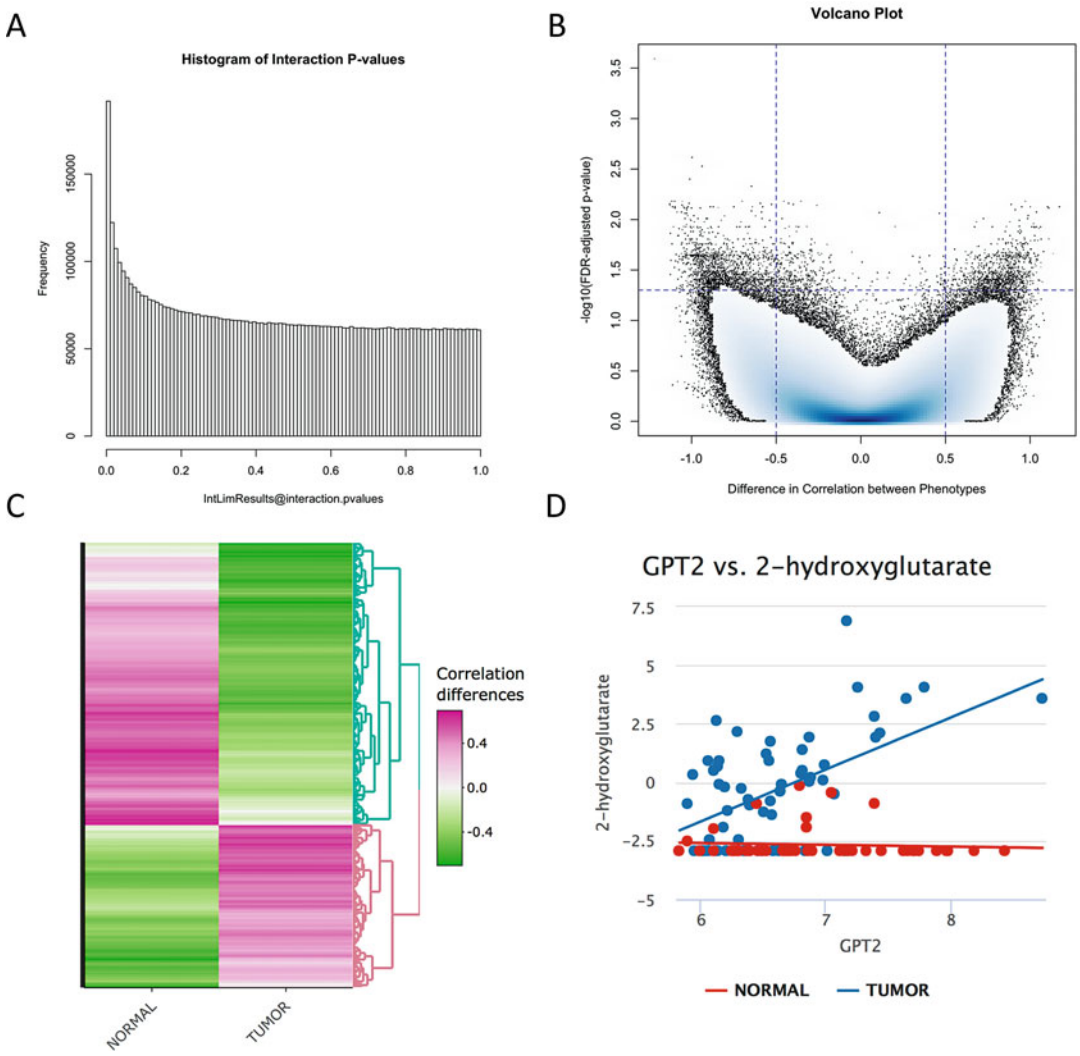


Fig. 2 Visualization of IntLIM results for breast cancer data. (**a**) Histogram of *p*-values. (**b**) Volcano plots of IntLIM results. (**c**) Heatmap of gene-metabolite pairs (FDR-adjusted *p*-value <0.05 and Spearman correlation >0.5). (**d**) Plot of GPT2 and 2-hydroxyglutarate

Linear modeling results can be filtered by FDR-adjusted interaction *p*-value and effect size, where the effect size is the interaction coefficient. For each gene-metabolite pair is the difference in Spearman correlations between the two phenotypic groups. A volcano plot can be generated to assist the user in determining FDR-adjusted *p*-value and effect size cutoffs using the pvalCorrVolcano() function. The "diffcorr" (for Spearman correlation differences) and "pvalcutoff" (for FDR-adjusted *p*-values) options allow the user to plot lines on the volcano plot to help define filtering cutoffs (Fig. 2b).

```
> pvalCorrVolcano(inputResults = myres, inputData = inputDatafilt, diffcorr = 0.5,
pvalcutoff = 0.05)
```

We select an FDR-adjusted interaction *p*-value of 0.05 and an absolute Spearman correlation difference of 0.5 as our cutoffs. The ProcessResults() function then filters gene-metabolite pairs based on those cutoffs. In addition, ProcessResults() performs hierarchical clustering of gene-metabolite pairs, based on the pattern of effect size. Specifically, Spearman correlations of each gene-metabolite pairs are calculated within each phenotype and input for hierarchical clustering. The parameter "treecuts" allows users to specify how many clusters to "cut" the results into.

```
> myres <- ProcessResults(inputResults = myres,
inputData = inputDatafilt, diffcorr = 0.50, pvalcutoff = 0.05, treecuts =2)
[1] "2842 gene-metabolite pairs found given cutoffs"
```

Based on our selected cutoffs, we identify 2842 gene-metabolite pairs that are tumor-dependent. The results are stored in the "filt.results" slot of the myres "IntLimResults" object. We can observe the first few lines of the filtered results using the head() function.

```
> head(myres@filt.results)
                    metab                          gene       NORMAL_cor   TUMOR_cor
1 1-arachidonoylglycerophosphoethanolamine*    C16orf42    -0.3374451     0.40211586
2 1-arachidonoylglycerophosphoethanolamine*    TOMM40      -0.4068953     0.52906640
3 1-arachidonoylglycerophosphoethanolamine*    C6orf186     0.3840933    -0.45903202
4 1-methylnicotinamide                         TGFB2        0.5392218    -0.13463776
5 1-methylnicotinamide                         KANK4       -0.5587962     0.06456901
6 1-methylnicotinamide                         CASC2        0.4997256    -0.31078794
       diff.corr     Pval FDRadj          Pval          cluster
1      0.7395610     1.499692e-05         0.04284062       1
2      0.9359617     1.667433e-05         0.04435878       1
3     -0.8431254     3.587729e-06         0.02777541       2
4     -0.6738596     1.703903e-05         0.04481976       2
5      0.6233652     1.128466e-05         0.03969863       1
6     -0.8105135     6.923041e-06         0.03449098       2
```

The OutputResults() function outputs result to a CSV file:

```
> OutputResults(inputResults = myres, filename = "FinalResults.csv")
```

3.1.6 Visualizing Significant Gene-Metabolite Pairs

The CorrHeatMap() function produces a heatmap of the Spearman correlations within each phenotypic group:

```
> CorrHeatmap(inputResults = myres, top_pairs = 3000)
```

The top_pairs parameter denotes the maximum number of pairs to plot (3000 in this example). We observe two major clusters (Fig. 2c). The smaller cluster (1038 gene-metabolite pairs) consists of gene-metabolite pairs positively correlated in tumor tissue but negatively correlated in normal tissue. The larger cluster (1084 gene-metabolite pairs) consists of gene-metabolite pairs negatively correlated in tumor tissue but positively correlated in normal tissue. From the heatmap, we are able to observe that most gene-metabolite associations that become altered change the directionality of their associations in tumor tissue. Integrating these results into pathway analysis tools will provide further insight into what pathways are altered in tumor tissue.

3.1.7 Extracting Relevant Genes and Metabolites for Pathway Analysis

The getMetabList() function allows users to extract metabolite names or source IDs for input into pathway analysis tools such as RaMP [38] (described below). To extract IDs, the metadata on metabolites needs to contain columns for the "sourceID" and "IDtype." The "sourceID" is the identifier for the metabolite, and the "IDtype" describes the type of identifier (kegg, hmdb, CAS, etc.). If no such columns exist, the unique metabolite names provided are extracted. By default, getMetabList() will prepend the "IDtype" to the "sourceID" since this is what is expected by RaMP. We extract the metabolite IDs as follows:

```
> metab.list <- getMetabList(inputResults = myres, inputData = inputDatafilt,
outputMetab = "id")
> length(metab.list)
[1] 145
#Output first 10 ids:
> metab.list[1:10]
[1]  "kegg:C02918"    "kegg:C03916"    "kegg:C01885"    "kegg:C04102"
 [5] "CAS:19420-57-6" "CAS:69747-55-3" "kegg:C03819"    "CAS:29743-97-3"
 [9] "CAS:73033-09-7" "kegg:C00956"
```

Analogously, the getGeneList() function retrieves a list of all gene IDs corresponding to the genes identified by the IntLIM analysis.

```
> gene.list <- getGeneList(inputResults = myres, inputData = inputDatafilt,
outputGene = "id")
> length(gene.list)
[1] 968
```

It is important to note that source IDs have the following format: "database_origin:id." This format is required for queries that use RaMP.

3.1.8 Visualizing Significant Gene-Metabolite Pairs

Lastly, IntLIM produces scatterplots of gene-metabolite pairs to visualize the alteration of their association by phenotype. In the published breast cancer study (PMCID: PMC3871244), 2-hydroxyglutarate was identified as being elevated in breast tumor tissue. With IntLIM, we were able to identify a tumor-dependent relationship between the gene GPT2 and 2-hydroxyglurate. We can visualize this using the scatterplot function (Fig. 2d).

```
> PlotGMPair(inputData = inputDatafilt, stype = "DIAG",
geneName = "GPT2", metabName = "2-hydroxyglutarate")
```

From the scatterplot, we observe that the relationship between GPT2 and 2-hydroxyglutarate is weakly negatively correlated in normal tissue ($r = -0.11$) but positively correlated in tumor tissue ($r = +0.40$). This alteration in GPT2–2-hyrdoxyglutarate may carry implications for metastasis of breast tissue.

3.2 Identifying Pathways Involved in Phenotype-Specific Gene-Metabolite Pairs Using RaMP

RaMP installation is a two-step process: (1) installation of the RaMP package and (2) installation of a local copy of the RaMP database. Detailed instructions on how to perform both of these tasks can be found at https://github.com/Mathelab/RaMP-DB. An overview of how to install RaMP is also outlined below. In this section, we outline the utility of the RaMP package with four example analyses: (1) retrieving analyte(s) from pathways of interest; (2) retrieving pathways from analyte(s) of interest and performing overrepresentation analysis (including clustering of similar enriched pathways); (3) identifying genes that catalyze reactions involving metabolites given a list of metabolites (or vice versa), and generating network visuals; and 4) retrieving ontologies (e.g., biofluid, cellular location, origin, tissue type) associated with metabolite(s).

3.2.1 Installing the RaMP R Package

Similar to IntLIM, we use the function install_github() to install RaMP:

```
> install.packages("devtools")
> library(devtools)
> install_github("mathelab/RAMP-DB")
> library(RaMP)
```

The RaMP package and its dependencies should begin to automatically download.

Next, the database connection information should be configured. The file that contains this information is within the package installation, and the directory can be found by typing the following:

```
> system.file("shinyApp", package="RaMP", mustWork=TRUE)
```

This line will return the directory location, in which you can find the "db.properties.template" file. Rename this file to "db.properties," and then edit it based on your current environment as follows:

```
host=<hostname of mysql server>
dbname=<db name on mysql>
username=<username to connect to mysql>
conpass=<password for username to connect to mysql>
```

Note that none of the RaMP associated functions will work until the RaMP database has also been installed (next step).

3.2.2 Installing the RaMP Database

To download and use a local copy of the RaMP database, a MySQL environment should be set up on your machine. If this is not already completed, a free download and instructions for setup are available at www.mysql.com/downloads. It is important to make note of the "root" user password when establishing your MySQL account. This password is used to access the RaMP database in many of the functions included in the package. If the password is lost or forgotten, it can be changed (instructions for this can be also be found on the MySQL website). Of note, one can also install the database on a remote server with a known host name, so the database could be used by multiple computers running RaMP.

Once the MySQL environment is set up, launch the MySQL server. To do so on the command line, type the following:

```
> mysql -u root -p your_password_here
```

When MySQL is running, initialize the RaMP database by typing the following:

```
mysql > create database ramp;
mysql > exit;
```

The database can now be populated with the ramp mysql dump file, which can be directly downloaded from https://github.com/Mathelab/RaMP-DB/inst/extdata/. The dump file name will have the ".sql" suffix with a number indicating the time of the latest update for the file, e.g., "ramp180302.sql.gz," where 180302 denotes the date—YearMonthDay—of the latest database update. This step can be performed with the code below (note that, similar to launching MySQL, it may be necessary to specify the location of your "myramp.sql" file or to be in the same directory as the myramp.sql file. It will most likely be in your downloads directory):

```
> mysql -u root -p ramp < myramp.sql
```

A local copy of the RaMP database should now be available on your machine. If the database is not located on a local computer, then the parameter -h can be specified with the hostname.

The line above can also be used to update the RaMP database as new versions of the RaMP dump file become available through our GitHub repository. This line will automatically overwrite the previous database with the same name, so you don't have to recreate an empty database if you have already done so. **Please note that you must manually download the updated file and update your local database**. Updating the R package will not update the local RaMP database automatically.

It is worth noting that RaMP functions that query the RaMP database take the following database-related parameters as input: (1) username (default: root), (2) conpass (password to access the database), (3) host (default: localhost), and (4) dbname (default: ramp). If the instructions above were followed to set up the RaMP database locally, only the conpass parameter will need to be passed to the functions.

3.2.3 Retrieve a List of Analytes from Pathways of Interest

Assuming that installation was successful, a copy of the RaMP SQL database is now present on your machine. One major utility of RaMP is the ability to easily search for analytes (genes or metabolites) in individual or multiple pathways of interest using the function getAnalyteFromPathway(). This function takes two mandatory parameters as input: (1) "pathway," pathway name(s), and (2) "conpass," MySQL password. In this example, analytes from the "glycolysis" pathway are retrieved.

```
> glycolysisAnalytes <- getAnalyteFromPathway(pathway="glycolysis",conpass="")
[1] "fired"
[1] "Timing .."
   user    system  elapsed
   0.092   0.002   0.709
> head(glycolysisAnalytes)
    pathwayName     pathwayCategory   pathwayType    compoundName
1   Glycolysis          smpdb2          hmdb          GALM
2   Glycolysis          smpdb2          hmdb          SLC2A2
3   Glycolysis          smpdb2          hmdb          PKM
4   Glycolysis          smpdb2          hmdb          PANK1;pantothenatekinase1
5   Glycolysis          smpdb2          hmdb          D-Glucose
6   Glycolysis          smpdb2          hmdb          PGAM2;phosphoglyceratemutase2
    sourceCompoundIDs
1  hmdb:HMDBP07277;ensembl:ENSG00000143891;kegg:130589;uniprot:Q96C23
2  uniprot:P11168;hmdb:HMDBP05474;kegg:6514; ensembl:ENSG00000163581;
    entrez:6514
3  uniprot: P14618;ensembl:ENSG00000067225;uniprot:V9HWB8;kegg:5315;
```

```
    entrez: 5315;hmdb:HMDBP00763
4   ensembl:ENSG00000152782;kegg:53354;uniprot:Q8TE04;hmdb:HMDBP01026
5   hmdb:HMDB0000122;chemspider:5589;CAS:50-997;chebi:4167;
    kegg:C00031;chebi:17634; pubchem:5793;CAS:2280-44-6
6   ensembl:ENSG00000164708;kegg:5224;uniprot:P15259;hmdb:HMDBP00267;
    entrez:5224
    geneOrCompound
1   gene
2   gene
3   gene
4   gene
5   compound
6   gene
```

Note: when inputting metabolite source IDs into getPathwayFromAnalyte(), make sure IDs are properly formatted. IDs should be prepended with their database of origin, e.g., kegg:C02712, hmdb:HMDB04824, or CAS:2566-39-4. The list of IDs that are included in RaMP are listed on our GitHub site (https://github.com/Mathelab/RaMP-DB) and currently include CAS, chebi, chemspider, enzymeNomenclature, ensembl, entrez, hmdb, kegg kegg, LIPIDMAPS, pubchem, and uniprot.

3.2.4 Retrieving Pathways Given Analytes, Performing Pathway Overrepresentation Analysis, and Clustering Similar Resulting Pathways

RaMP can assist in identifying pathways associated with analytes of interest. To start, a list of metabolites and/or genes of interest is used for input. In this example, a list of significant gene-metabolite pairs output by IntLIM is used as input for pathway analysis. The lists used in the following examples are stored in the gene.metab.res.RData dataset, which is also included in the RaMP R package (https://github.com/Mathelab/RaMP-DB/tree/dev/inst/extdata). To load the data:

```
> mydir <- system.file("extdata", package="RaMP", mustWork=TRUE)
> load(paste0(mydir,"/gene.metab.IntLIM.res.RData"))
> # To view the objects loaded:
> ls()
[1] "gene.list"      "gene.res"          "metab.list"
[4] "metab.res"       "mydir"
```

The objects metab.list and gene.list are lists of KEGG/HMDB and ensemble IDs, respectively, which correspond to analytes that were identified as members of significant gene-metabolite pairs by IntLIM in the prior section.

To retrieve all pathways from the RaMP database that include any of the analytes output by IntLIM, both "metab.list" and "gene.list" will be input in the getPathwayFromAnalyte() function through the "analytes" parameter. While metabolites and genes can be input simultaneously, either the names or source IDs of

the analytes, not a mix of names and source IDs, can be processed. To designate which type of input is used, the user can designate "names" or "IDs" for the "NameOrIds" parameter. Lastly, the user must provide the MySQL password through the conpass argument. The following code is used to run getPathwayFromAnalyte() on the example data:

```
> myPathways <- getPathwayFromAnalyte(analytes=c(gene.list,metab.list),
conpass="",NameOrIds="ids")
> head(myPathways)
     rampId            pathwayRampId              pathwayName
1   RAMP_C_000000943  RAMP_P_000051316.     Alanine, aspartate and glutamate
    metabolism
2   RAMP_C_000000943  RAMP_P_000048918      Amino Acid metabolism
3   RAMP_C_000000943  RAMP_P_000048900      Alanine and aspartate
    metabolism
4   RAMP_C_000000943  RAMP_P_000051291      Mineral absorption
5   RAMP_C_000000943  RAMP_P_000051360      Aminoacyl-tRNA biosynthesis
6   RAMP_C_000000943  RAMP_P_000051210      Metabolic pathways
    pathwaysourceId     pathwaysource       commonName
1   00250               kegg                l-asparagine
2   WP3925              wiki                l-asparagine
3   WP106               wiki                l-asparagine
4   04978               kegg                l-asparagine
5   00970               kegg                l-asparagine
6   01100               kegg                l-asparagine
```

Remember that if using IDs as input, they should be prepended with their database of origin (e.g., kegg:C02712).

Also note that while the internal RaMP IDs are returned for compounds and pathways, these are only useful for harmonizing metabolites and pathways across databases and are not used otherwise.

The function call returns all pathways from the RaMP database that includes any of the analytes in your input. Importantly, pathways and analytes in the dataframe are not necessarily unique, as a single analyte may map to several pathways, and several analytes in the user data may be associated with the same pathways.

To output the table from getPathwayFromAnalyte() as a CSV file, one could use the write_to_csv() function as follows:

```
> writePathwaysToCSV(myPathways,outputfile="myfile.csv")
```

Using the output of this table, it is now possible to perform pathway overrepresentation analysis (using the Fisher's exact test) with the runCombinedFisherTest() function. This function takes genes and metabolites as input together or separately. When both

genes and metabolites are input, overrepresentation analysis is first performed independently, and then *p*-values are combined using the Fisher method [42].

```
> myFishersDF <- runCombinedFisherTest(pathwaydf=myPathways, conpass=" ")
> print(head(myFishersDF$fishresults))
    pathwayRampId       Pval.Metab        Num_In_Path.Metab      Total_In_Path.Metab
 1  RAMP_P_000048626    NA                 NA                     NA
 5  RAMP_P_000048628    NA                 NA                     NA
 8  RAMP_P_000048630    NA                 NA                     NA
10  RAMP_P_000048636    NA                 NA                     NA
14  RAMP_P_000048648    0.2914752          2                      9
20  RAMP_P_000048650    NA                 NA                     NA
    Pval.Gene         Num_In_Path.Gene    Total_In_Path.Gene     Pval_combined
  1 0.029177925       4                   64                     0.0291779250
  5 0.122862751       3                   67                     0.1228627510
  8 0.033500948       2                   16                     0.0335009480
10 0.087152253        4                   92                     0.0871522531
14 0.001359978        4                   27                     0.0003963999
20 0.217371940        2                   48                     0.2173719405
     Pval_combined_FDR         Pval_combined_Holm
 1   0.119646102               1.0000000
 5   0.329821621               1.0000000
 8   0.131628270               1.0000000
10   0.271377549               1.0000000
14   0.005129462               0.7908178
20   0.459628927               1.0000000
       pathwayName
  1   Lung fibrosis
  5   Association Between Physico-Chemical Features and Toxicity Associated Pathways
  8   Codeine and Morphine Metabolism
10   Corticotropin-releasing hormone signaling pathway
14   Lipid Metabolism Pathway
20   Differentiation Pathway
     pathwaysourceId     pathwaysource
 1   WP3624              wiki
 5   WP3680              wiki
 8   WP1604              wiki
10   WP2355              wiki
14   WP3965              wiki
20   WP2848              wiki
```

The output of runCombinedFisherTest is a list containing two objects:

(a) fishresults: A dataframe with one unique pathway per line, displaying results of overrepresentation analysis. The number of columns will vary depending on whether genes, metabolites,

or both were input. In this example, where genes and metabolites were input, the columns are:

Column 1: The unique RaMPID for the pathway in that line.

Columns 2 to 4: The p-value for metabolites, the number of user metabolites in that pathway, the total number of metabolites in that pathway.

Columns 5 to 7: Same as 2:4, but for genes.

Columns 8 to 10: Combined *p*-values, including adjustments for multiple comparisons using the false discovery rate and Holm methods.

Columns 11 to 13: Pathway information (Common name, source ID, and source database)

(b) analyte_type: A string describing the data types that were detected (genes, metabolites, or both).

Due to the large number of statistical tests being performed (one per pathway being considered), *p*-values are adjusted for multiple comparisons using the methods by Benjamini Hochberg or Holm to correct for multiple comparisons (FDR) and Holm method. The FilterFishersResults() function filters the pathway results by FDR (via the p_fdradj_cutoff argument) or the Holm method (with the p_holmadj_cutoff argument):

```
> myPathwaysSignificant=FilterFishersResults(myFishersDF, p_holmadj_cutoff=0.1)
>
> head(myPathwaysSignificant$fishresults)
     pathwayRampId       Pval.Metab        Num_In_Path.Metab     Total_In_Path.Metab
 42  RAMP_P_000048718    NA                NA                     NA
182  RAMP_P_000048918    1.357424e-03      22                     97
211  RAMP_P_000048922    NA                NA                     NA
305  RAMP_P_000049039    4.008352e-05      5                      32
334  RAMP_P_000049067    6.351244e-03      2                      10
399  RAMP_P_000049098    NA                NA                     NA
     Pval.Gene           Num_In_Path.Gene  Total_In_Path.Gene     Pval_combined
 42  2.918568e-05        9                 87                     2.918568e-05
182  1.647419e-03        7                 94                     2.236245e-06
211  1.599853e-05        8                 62                     1.599853e-05
305  NA                  NA                NA                     4.008352e-05
334  1.177852e-03        5                 44                     7.480828e-06
399  1.700290e-06       11                 98                     1.700290e-06
     Pval_combined_FDR  Pval_combined_Holm
 42  6.502089e-04       0.060268433
182  7.106655e-05       0.004682697
211  3.928729e-04       0.033180943
305  8.095373e-04       0.082371642
334  2.155476e-04                0.015612487
399  5.741137e-05                0.003567209
     pathwayName
```

```
 42  Ectoderm Differentiation
182  Amino Acid metabolism
211  Oxidative phosphorylation
305  Metabolism of ingested SeMet, Sec, MeSec into H2Se
334  Condensation of Prophase Chromosomes
399  RUNX1 regulates transcription of genes involved in differentiation of HSCs
     pathwaysourceId     pathwaysource
 42  WP2858              wiki
182  WP3925              wiki
211  WP623               wiki
305  R-HSA-2408508       reactome
334  R-HSA-2299718       reactome
399  R-HSA-8939236       reactome
> dim(myPathwaysSignificant$fishresults)
[1] 85 13
```

In this example, we identified 85 pathways with a Holm adjusted p-value below 0.1.

To improve interpretation of our resulting list of pathways, and because many pathways have a high degree of overlap in their gene-metabolite composition, we implemented the findCluster() function. This pathway clustering function runs an unsupervised fuzzy clustering algorithm, similar to the one implemented in DAVID [43], which was designed to automatically identify and cluster overlapping pathways [38]. The algorithm identifies "seeds," which are pathways that have a user-defined degree of overlap (perc_analyte_overlap parameter) with a user-defined number of other pathways (min_pathway_tocluster parameter). These pathways are used as "seeds" for new clusters. Pathways with a user-specified degree of overlap with a seed (perc_analyte_overlap parameter) are then placed into a cluster with the corresponding seed. The overlap between two pathways is defined as the number of analytes that are overlapping between the two pathways divided by the total number of unique analytes among the pathways. Lastly, similar clusters are themselves merged together if their contained pathways overlap more than a user-defined threshold (perc_pathway_overlap). The findCluster() function takes four parameters as input:

(a) fishers_df: the raw output of runCombinedFisherTest or the filtered output from FilterFishersResults.

(b) perc_analyte_overlap: cutoff value for the percent overlap between pathways necessary to consider a pathway a neighbor to a seed. This threshold is used to identify seeds for cluster

establishment (default: 0.5). (This parameter is only used for cluster establishment).

(c) min_pathway_tocluster: cutoff value for the minimum number of similar pathways a pathway needs to be considered a seed for cluster establishment (default: 2). (This parameter is only used for cluster establishment).

(d) perc_pathway_overlap: cutoff value for the minimum percent overlap between clusters by pathway composition to merge clusters. Lower values of this threshold will result in more distinct clusters in the final output but will reduce the number of clusters identified (default: 0.5).

Using the example data, findCluster() can be run like so:

```
> myPathwaysSignificantClustered=findCluster(fishers_df=myPathwaysSignificant,
     perc_analyte_overlap=0.3, min_pathway_tocluster=2,
     perc_pathway_overlap=0.75)
> length(myPathwaysSignificantClustered$cluster_list)
[1] 21
```

The output is a list containing three objects: the pathway enrichment output with a cluster_assignment column added, the analyte type, and the cluster output in list form, with each entry of the list representing a cluster, containing a vector of pathways contained in that cluster. In the example dataset, 22 clusters were identified. To retrieve pathways for cluster 12, we use the grep function to search for "12" in the cluster assignments:

```
>head(myPathwaysSignificantClustered$fishresults[grep("12",myPathwaysSignificant-
Clustered$fishresults$cluster_assignment),])
      Pval.Metab         Num_In_Path.Metab     Total_In_Path.Metab   Pval.Gene
 867  4.678048e-05       5                     33                     0.003251840
1433  1.285221e-05       6                     43                     0.008199957
1825  2.132602e-06       6                     32                     0.160791585
      Num_In_Path.Gene   Total_In_Path.Gene    Pval_combined       Pval_combined_FDR
 867  4                  34                     1.521227e-07          7.146458e-06
1433  6                  96                     1.053876e-07          5.554696e-06
1825  3                  76                     3.429045e-07          1.482033e-05
      Pval_combined_Holm  pathwayName          pathwaysourceId       pathwaysource
 867  0.0003218916        Gluconeogenesis      R-HSA-70263           reactome
1433  0.0002235270        Glucose metabolism   R-HSA-70326           reactome
1825  0.0007242144        Glycolysis           R-HSA-70171           reactome
      cluster_assignment  rampids
 867  8, 12               RAMP_P_000049428
1433  8, 12, 17           RAMP_P_000049831
```

```
1825  12, 17              RAMP_P_000050117
```

To write pathway enrichment analysis results (before or after clustering) to a CSV file for filtering and viewing in Excel:

```
> write_FishersResults(myPathwaysSignificantClustered,
        outputfile="myfile.csv")
```

3.2.5 Identifying Analytes that Catalyze Reactions Involving Analytes in a List of User Analytes and Generating Network Visuals Based on the Result

The function rampFastCata() retrieves a list of analytes that participate in reactions involving an analyte or analytes of interest. This function takes three parameters as input:

(a) analytes: a string with a single analyte or vector of multiple analytes to be queried,

(b) conpass: the database password for the "root" user,

(c) NameOrIds: a string specifying if the input is common names ("names") or source IDs ("IDs," default).

For example:

```
> glucoseReactions <- rampFastCata(analytes = "glucose", conpass = "", NameOrIds =
"names")
[1] "Get Compound ..."
[1] 1
[1] "Getting gene Id from Compound Id ..."
[1] "Getting names from gene Id ..."
[1] "Done ..."
[1] "timing ..."
   user  system elapsed
  0.021   0.000   0.764
```

Only print the first two columns since the third column contains ID names and can be lengthy.

```
> head(glucoseReactions[,1:2])
  Input_Analyte   Input_CatalyzedBy_CommonName
1     Glucose     HK1;hexokinase1
2     Glucose     HK1
3     Glucose     LALBA
4     Glucose     LALBA;lactalbuminalpha
5     Glucose     G6PC
6     Glucose     G6PC;glucose-6-phosphatasecata
```

```
> batchReactions <- rampFastCata(analytes = c("glucose", "glycogen", "cholesterol",
"creatine", "MDM2", "TP53"),
conpass="",NameOrIds="names")
[1] "Get Compound ..."
```

```
[1] 5
[1] "Geting gene Id from Compound Id ..."
[1] "Getting names from gene Id ..."
[1] "Also find gene inside"
[1] "Get gene ..."
[1] "rampId"         "type1"          "InputAnalyte"   "rampCompoundId"
[5] "rampGeneId"     "sourceId"       "rampId2"        "IDtype"
[9] "geneOrCompound" "commonName"
[1] "Done ..."
[1] "timing ..."
   user  system elapsed
  0.079   0.013   1.613
> head(batchReactions[,1:2])
      Input_Analyte   Input_CatalyzedBy_CommonName
1     Glucose         HK1;hexokinase1
2     Glucose         HK1
3     Glucose         LALBA
4     Glucose         LALBA;lactalbuminalpha
5     Cholesterol     LCAT;lecithin-cholesterolacylt
6     Cholesterol     LCAT
```

Note that rampFastCata() can accept both genes and metabolites as input for batch queries. If the input is a metabolite, the function will output gene transcript common names and source IDs that are known to catalyze reactions in the same pathway as that metabolite. Conversely, if the input is a gene, the function will return the common name and source id of metabolites known to be catalyzed by that gene.

To visualize the reaction level as a network, use the plotCataNetwork() function:

```
> plotCataNetwork(batchReactions)
```

The output of this function is an interactive HTML plot that allows the user to pan/zoom into regions of interest (Fig. 3). User genes/metabolites are colored in blue, whereas analytes found by the function are colored in orange.

3.2.6 Retrieve a List of Ontologies Associated with User Metabolite(s)

The last RaMP functionality is to query ontologies relating to metabolites. Ontologies include biofluid location, cellular location, origin, and tissue types. The function getOntoFromMeta() accepts a list of genes or metabolites and returns a dataframe containing ontological terms associated with input analytes, as well as biofluid or cellular location. Conversely, the function getMetaFromOnto() accepts ontological terms and returns metabolites associated with those terms. Both functions require the conpass parameter, and the

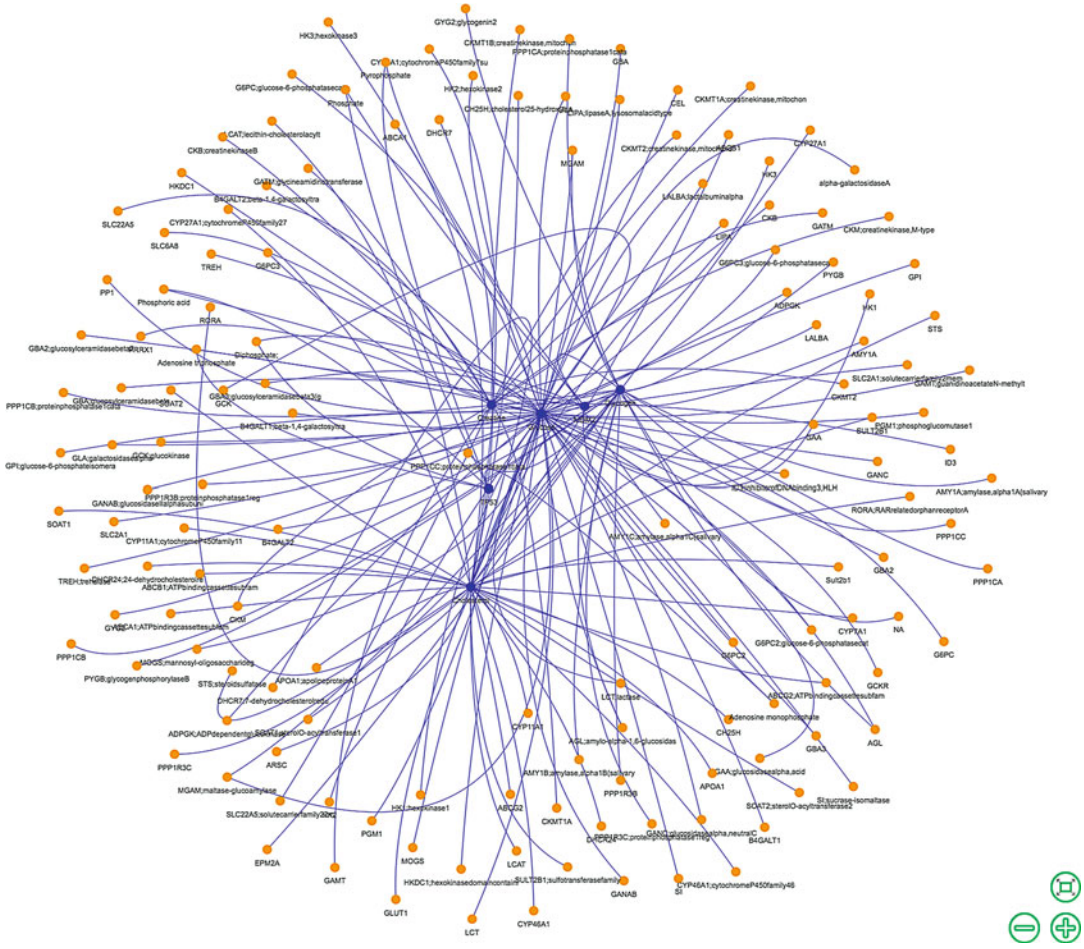


Fig. 3 Network visualization of gene-metabolite reactions

getOntoFromMeta() function requires the NameOrIds parameter, to interpret the input list. To run the function:

```
> ontoDF <- getOntoFromMeta(analytes=c("creatine","glucagon","cholesterol"),
conpass="",NameOrIds="name")
> head(ontoDF)
     Metabolites                 sourceId                  IDtype      Ontology
1   Cholesterol-ester-pool       chebi:16113               chebi       Bile
2   Cholesterol-ester-pool       CAS:57-88-5               CAS         Bile
3   Cholesterol                  chebi:1307929             chebi       Bile
4   Cholesterol                  chemspider:9200676        chemspider  Bile
5   Cholesterol                  LIPIDMAPS:LMST01010093    LIPIDMAPS   Bile
6   Cholesterol-ester-pool       hmdb:HMDB0062453          hmdb        Bile
                biofluidORcellular
1   biofluid
2   biofluid
3   biofluid
```

```
4   biofluid
5   biofluid
6   biofluid
> foodDF <- getMetaFromOnto(ontology="food",conpass="")
[1] "rampOntologyIdLocation" "commonName"             "biofluidORcellular"
[1] "rampCompoundId"         "rampOntologyIdLocation"
[1] "sourceId"       "rampId"         "IDtype"         "geneOrCompound"
[5] "commonName"
[1] "Merging 1..."
[1] "Merging 2..."
> head(foodDF)
    Metabolites                          sourceId           IDtype     Ontology
1   TG(24:1(15Z)/18:4(6Z,9Z,12Z,15       hmdb:HMDB0052301   hmdb       Food
2   TG(15:0/18:2(9Z,12Z)/22:5(4Z,7       hmdb:HMDB0043416   hmdb       Food
3   TG(18:3(6Z,9Z,12Z)/14:1(9Z)/18       hmdb:HMDB0052950   hmdb       Food
4   Annuolide G                          hmdb:HMDB0034605   hmdb       Food
5   TG(15:0/20:4(8Z,11Z,14Z,17Z)/2       hmdb:HMDB0043696   hmdb       Food
6   TG(16:1(9Z)/16:0/20:2n6)             hmdb:HMDB0048466   hmdb       Food
    biofluidORcellular
1   origins
2   origins
3   origins
4   origins
5   origins
6   origins
```

Acknowledgment

This work was supported by funding from the National Cancer Institute (1R03CA222428-01) and the Ohio State University Translational Data Analytics Institute and startup funds by the Ohio State University to Ewy Mathé, by the Ohio State University Discovery Themes Foods for Health postdoctoral fellowship to Jalal Siddiqui, and by the National Institute of General Medical Sciences of the National Institutes of Health to Andy Patt (T32GM068412). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Ritchie MD, Holzinger ER, Li R, Pendergrass SA, Kim D (2015) Methods of integrating data to uncover genotype-phenotype interactions. Nat Rev Genet 16(2):85–97. https://doi.org/10.1038/nrg3868

2. van Karnebeek CDM, Wortmann SB, Tarailo-Graovac M, Langeveld M, Ferreira CR, van de Kamp JM, Hollak CE, Wasserman WW, Waterham HR, Wevers RA, Haack TB, Wanders RJA, Boycott KM (2018) The role of the clinician in the multi-omics era: are you ready? J Inherit

Metab Dis 41(3):571–582. https://doi.org/10.1007/s10545-017-0128-1

3. Gomez-Cabrero D, Abugessaisa I, Maier D, Teschendorff A, Merkenschlager M, Gisel A, Ballestar E, Bongcam-Rudloff E, Conesa A, Tegner J (2014) Data integration in the era of omics: current and future challenges. BMC Syst Biol 8(Suppl 2):I1. https://doi.org/10.1186/1752-0509-8-S2-I1
4. Rohart F, Gautier B, Singh A, Le Cao KA (2017) mixOmics: an R package for 'omics feature selection and multiple data integration. PLoS Comput Biol 13(11):e1005752. https://doi.org/10.1371/journal.pcbi.1005752
5. Fukushima A (2013) DiffCorr: an R package to analyze and visualize differential correlations in biological networks. Gene 518(1):209–214. https://doi.org/10.1016/j.gene.2012.11.028
6. Su G, Burant CF, Beecher CW, Athey BD, Meng F (2011) Integrated metabolome and transcriptome analysis of the NCI60 dataset. BMC Bioinf 12(Suppl 1):S36. https://doi.org/10.1186/1471-2105-12-S1-S36
7. Bradley PH, Brauer MJ, Rabinowitz JD, Troyanskaya OG (2009) Coordinated concentration changes of transcripts and metabolites in Saccharomyces cerevisiae. PLoS Comput Biol 5(1):e1000270. https://doi.org/10.1371/journal.pcbi.1000270
8. Kuo TC, Tian TF, Tseng YJ (2013) 3Omics: a web-based systems biology tool for analysis, integration and visualization of human transcriptomic, proteomic and metabolomic data. BMC Syst Biol 7:64. https://doi.org/10.1186/1752-0509-7-64
9. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinf 9:559. https://doi.org/10.1186/1471-2105-9-559
10. Siska C, Bowler R, Kechris K (2016) The discordant method: a novel approach for differential correlation. Bioinformatics 32(5):690–696. https://doi.org/10.1093/bioinformatics/btv633
11. Siddiqui JK, Baskin E, Liu M, Cantemir-Stone CZ, Zhang B, Bonneville R, McElroy JP, Coombes KR, Mathe EA (2018) IntLIM: integration using linear models of metabolomics and gene expression data. BMC Bioinf 19(1):81. https://doi.org/10.1186/s12859-018-2085-6
12. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45(D1):D353–D361. https://doi.org/10.1093/nar/gkw1092
13. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28(1):27–30
14. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44(D1):D457–D462. https://doi.org/10.1093/nar/gkv1070
15. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR, Jassal B, Jupe S, Matthews L, May B, Palatnik S, Rothfels K, Shamovsky V, Song H, Williams M, Birney E, Hermjakob H, Stein L, D'Eustachio P (2014) The Reactome pathway knowledgebase. Nucleic Acids Res 42(Database issue):D472–D477. https://doi.org/10.1093/nar/gkt1102
16. Kelder T, van Iersel MP, Hanspers K, Kutmon M, Conklin BR, Evelo CT, Pico AR (2012) WikiPathways: building research communities on biological pathways. Nucleic Acids Res 40(Database issue):D1301–D1307. https://doi.org/10.1093/nar/gkr1074
17. Kutmon M, Riutta A, Nunes N, Hanspers K, Willighagen EL, Bohler A, Melius J, Waagmeester A, Sinha SR, Miller R, Coort SL, Cirillo E, Smeets B, Evelo CT, Pico AR (2016) WikiPathways: capturing the full diversity of pathway knowledge. Nucleic Acids Res 44(D1):D488–D494. https://doi.org/10.1093/nar/gkv1024
18. Slenter DN, Kutmon M, Hanspers K, Riutta A, Windsor J, Nunes N, Melius J, Cirillo E, Coort SL, Digles D, Ehrhart F, Giesbertz P, Kalafati M, Martens M, Miller R, Nishida K, Rieswijk L, Waagmeester A, Eijssen LMT, Evelo CT, Pico AR, Willighagen EL (2018) WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Res 46(D1):D661–D667. https://doi.org/10.1093/nar/gkx1064
19. Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, Djoumbou Y, Mandal R, Aziat F, Dong E, Bouatra S, Sinelnikov I, Arndt D, Xia J, Liu P, Yallou F, Bjorndahl T, Perez-Pineiro R, Eisner R, Allen F, Neveu V, Greiner R, Scalbert A (2013) HMDB 3.0--The Human Metabolome Database in 2013. Nucleic Acids Res 41(Database issue):D801–D807. https://doi.org/10.1093/nar/gks1065
20. Wishart DS, Knox C, Guo AC, Eisner R, Young N, Gautam B, Hau DD, Psychogios N, Dong E, Bouatra S, Mandal R, Sinelnikov I, Xia J, Jia L, Cruz JA, Lim E, Sobsey CA,

Shrivastava S, Huang P, Liu P, Fang L, Peng J, Fradette R, Cheng D, Tzur D, Clements M, Lewis A, De Souza A, Zuniga A, Dawe M, Xiong Y, Clive D, Greiner R, Nazyrova A, Shaykhutdinov R, Li L, Vogel HJ, Forsythe I (2009) HMDB: a knowledgebase for the human metabolome. Nucleic Acids Res 37 (Database issue):D603–D610. https://doi.org/10.1093/nar/gkn810

21. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, Fung C, Nikolai L, Lewis M, Coutouly MA, Forsythe I, Tang P, Shrivastava S, Jeroncic K, Stothard P, Amegbey G, Block D, Hau DD, Wagner J, Miniaci J, Clements M, Gebremedhin M, Guo N, Zhang Y, Duggan GE, Macinnis GD, Weljie AM, Dowlatabadi R, Bamforth F, Clive D, Greiner R, Li L, Marrie T, Sykes BD, Vogel HJ, Querengesser L (2007) HMDB: the human metabolome database. Nucleic Acids Res 35(Database issue):D521–D526. https://doi.org/10.1093/nar/gkl923
22. Frolkis A, Knox C, Lim E, Jewison T, Law V, Hau DD, Liu P, Gautam B, Ly S, Guo AC, Xia J, Liang Y, Shrivastava S, Wishart DS (2010) SMPDB: the small molecule pathway database. Nucleic Acids Res 38(Database issue):D480–D487. https://doi.org/10.1093/nar/gkp1002
23. Jewison T, Su Y, Disfany FM, Liang Y, Knox C, Maciejewski A, Poelzer J, Huynh J, Zhou Y, Arndt D, Djoumbou Y, Liu Y, Deng L, Guo AC, Han B, Pon A, Wilson M, Rafatnia S, Liu P, Wishart DS (2014) SMPDB 2.0: big improvements to the small molecule pathway database. Nucleic Acids Res 42(Database issue):D478–D484. https://doi.org/10.1093/nar/gkt1067
24. Paley S, Karp PD (2017) Update notifications for the BioCyc collection of databases. Database (Oxford). 2017. Doi:https://doi.org/10.1093/database/bax086
25. Cerami EG, Gross BE, Demir E, Rodchenkov I, Babur O, Anwar N, Schultz N, Bader GD, Sander C (2011) Pathway commons, a web resource for biological pathway data. Nucleic Acids Res 39(Database issue):D685–D690. https://doi.org/10.1093/nar/gkq1039
26. Xia J, Sinelnikov IV, Han B, Wishart DS (2015) MetaboAnalyst 3.0--making metabolomics more meaningful. Nucleic Acids Res 43(W1): W251–W257. https://doi.org/10.1093/nar/gkv380
27. Kamburov A, Cavill R, Ebbels TM, Herwig R, Keun HC (2011) Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. Bioinformatics 27 (20):2917–2918. https://doi.org/10.1093/bioinformatics/btr499
28. Gowda H, Ivanisevic J, Johnson CH, Kurczy ME, Benton HP, Rinehart D, Nguyen T, Ray J, Kuehl J, Arevalo B, Westenskow PD, Wang J, Arkin AP, Deutschbauer AM, Patti GJ, Siuzdak G (2014) Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses. Anal Chem 86 (14):6931–6939. https://doi.org/10.1021/ac500734c
29. Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G (2012) XCMS Online: a web-based platform to process untargeted metabolomic data. Anal Chem 84(11):5035–5039. https://doi.org/10.1021/ac300698c
30. Guijas C, Montenegro-Burke JR, Domingo-Almenara X, Palermo A, Warth B, Hermann G, Koellensperger G, Huan T, Uritboonthai W, Aisporna AE, Wolan DW, Spilker ME, Benton HP, Siuzdak G (2018) METLIN: a technology platform for identifying knowns and unknowns. Anal Chem 90 (5):3156–3164. https://doi.org/10.1021/acs.analchem.7b04424
31. Fahy E, Sud M, Cotter D, Subramaniam S (2007) LIPID MAPS online tools for lipid research. Nucleic Acids Res 35(Web Server issue):W606–W612. https://doi.org/10.1093/nar/gkm324
32. P.J. Linstrom and W.G. Mallard 2018) NIST Chemistry WebBook, NIST Standard Reference Database Number 69. National Institute of Standards and Technology. Doi:https://doi.org/10.18434/T4D303
33. Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, Ojima Y, Tanaka K, Tanaka S, Aoshima K, Oda Y, Kakazu Y, Kusano M, Tohge T, Matsuda F, Sawada Y, Hirai MY, Nakanishi H, Ikeda K, Akimoto N, Maoka T, Takahashi H, Ara T, Sakurai N, Suzuki H, Shibata D, Neumann S, Iida T, Tanaka K, Funatsu K, Matsuura F, Soga T, Taguchi R, Saito K, Nishioka T (2010) MassBank: a public repository for sharing mass spectral data for life sciences. J Mass Spectrom 45(7):703–714. https://doi.org/10.1002/jms.1777
34. Wanichthanarak K, Fan S, Grapov D, Barupal DK, Fiehn O (2017) Metabox: a toolbox for metabolomic data analysis, interpretation and integrative exploration. PLoS One 12(1): e0171046. https://doi.org/10.1371/journal.pone.0171046
35. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2016) PubChem substance and compound databases.

Nucleic Acids Res 44(D1):D1202–D1213. https://doi.org/10.1093/nar/gkv951

36. UniProt Consortium T (2018) UniProt: the universal protein knowledgebase. Nucleic Acids Res 46(5):2699. https://doi.org/10.1093/nar/gky092
37. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Giron CG, Gil L, Gordon L, Haggerty L, Haskell E, Hourlier T, Izuogu OG, Janacek SH, Juettemann T, To JK, Laird MR, Lavidas I, Liu Z, Loveland JE, Maurel T, McLaren W, Moore B, Mudge J, Murphy DN, Newman V, Nuhn M, Ogeh D, Ong CK, Parker A, Patricio M, Riat HS, Schuilenburg H, Sheppard D, Sparrow H, Taylor K, Thormann A, Vullo A, Walts B, Zadissa A, Frankish A, Hunt SE, Kostadima M, Langridge N, Martin FJ, Muffato M, Perry E, Ruffier M, Staines DM, Trevanion SJ, Aken BL, Cunningham F, Yates A, Flicek P (2018) Ensembl 2018. Nucleic Acids Res 46(D1):D754–D761. https://doi.org/10.1093/nar/gkx1098
38. Zhang B, Hu S, Baskin E, Patt A, Siddiqui JK, Mathe EA (2018) RaMP: a comprehensive relational database of metabolomics pathways for pathway enrichment analysis of genes and metabolites. Meta 8(1):E16. https://doi.org/10.3390/metabo8010016
39. Terunuma A, Putluri N, Mishra P, Mathe EA, Dorsey TH, Yi M, Wallace TA, Issaq HJ, Zhou M, Killian JK, Stevenson HS, Karoly ED, Chan K, Samanta S, Prieto D, Hsu TY, Kurley SJ, Putluri V, Sonavane R, Edelman DC, Wulff J, Starks AM, Yang Y, Kittles RA, Yfantis HG, Lee DH, Ioffe OB, Schiff R, Stephens RM, Meltzer PS, Veenstra TD, Westbrook TF, Sreekumar A, Ambs S (2014) MYC-driven accumulation of 2-hydroxyglutarate is associated with breast cancer prognosis. J Clin Invest 124 (1):398–412. https://doi.org/10.1172/JCI71180
40. Hernandez-Ferrer C, Ruiz-Arenas C, Beltran-Gomila A, Gonzalez JR (2017) MultiDataSet: an R package for encapsulating multiple data sets with application to omic data integration. BMC Bioinf 18(1):36. https://doi.org/10.1186/s12859-016-1455-1
41. Wickham H; Chang W (2016) devtools: tools to make developing R packages easier. R package version 1.11.1 edn
42. Fisher R (1950) Statistical methods for research workers. Oliver & Boyd, London
43. Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, Stephens R, Baseler MW, Lane HC, Lempicki RA (2007) The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 8(9):R183. https://doi.org/10.1186/gb-2007-8-9-r183

Chapter 24

Biclustering Analysis of Co-regulation Patterns in Nuclear-Encoded Mitochondrial Genes and Metabolic Pathways

Robert B. Bentham, Kevin Bryson, and Gyorgy Szabadkai

Abstract

Transcription of a large set of nuclear-encoded genes underlies biogenesis of mitochondria, regulated by a complex network of transcription factors and co-regulators. A remarkable heterogeneity can be detected in the expression of these genes in different cell types and tissues, and the recent availability of large gene expression compendiums allows the quantification of specific mitochondrial biogenesis patterns. We have developed a method to effectively perform this task. Massively correlated biclustering (MCbiclust) is a novel bioinformatics method that has been successfully applied to identify co-regulation patterns in large genesets, underlying essential cellular functions and determining cell types. The method has been recently evaluated and made available as a package in Bioconductor for R. One of the potential applications of the method is to compare expression of nuclear-encoded mitochondrial genes or larger sets of metabolism-related genes between different cell types or cellular metabolic states. Here we describe the essential steps to use MCbiclust as a tool to investigate co-regulation of mitochondrial genes and metabolic pathways.

Key words Biclustering, MCbiclust, Mitochondria, Metabolism, Gene expression

1 Introduction

Mammalian mitochondria are estimated to be composed of as many as 1500 genes [1] encoded in the nucleus along with the 13 protein-coding genes of the mitochondrial genome (mtDNA). To maintain proper mitochondrial function, the expression of the two genomes must be both coordinated and able to adapt to highly variable energetic demands. This results in a remarkable heterogeneity of mitochondrial composition, as detailed in numerous recent studies exploring the startling variety of mitochondrial function, physiology, and proteome makeup across different tissues and cell types [2–4].

Accordingly, the transcriptional regulation of mitochondrial biogenesis has been shown to be a highly complex process (*see*, e.g., [5, 6]), involving numerous transcription factors and

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_24,

co-regulators, forming a complex interaction network, which is also highly adaptable via posttranscriptional modifications. While physiological regulation of mitochondrial biogenesis and composition vary enormously across healthy tissues, it is also known to contribute to major disease states. Mitochondrial dysfunction due to defects in the mitochondrial biogenesis pathway is known to be an important factor in cancer, neuromuscular degenerative disease, and cardiomyopathies [7, 8]. Whether these changes are the primary cause of the disease or the result of adaptation or maladaptation is an important open question in many cases. For these reasons bioinformatics tools to investigate the co-regulation of nuclear-encoded mitochondrial genes not only have the potential to examine how physiological regulation works but also to reveal underlying factors that contribute to the disease.

While the direct examination of the total mitochondrial proteome affected by the transcription factor network is often technically unfeasible, the availability of good-quality, high-coverage gene expression (microarray or RNAseq) data makes it realistic to study the output of this network at the mRNA level. However, the success of this analysis relies on the ability of the applied methods to identify gene-sample "biclusters" of similar mitochondrial co-regulation, since a single dataset often contains multiple modes of control in diverse mitochondrial gene groups. Here we discuss how a recently developed novel method MCbiclust [9] can be used for this task.

2 Materials

In the following sections, we will refer to these software/manuals/datasets:

1. MCbiclust (doi:10.18129/B9.bioc.MCbiclust, current version 1.2.1), an R package available in Bioconductor [10], an open source platform for software in bioinformatics.
2. The MCbiclust package introductory manual (IM) accessed on the Bioconductor website (https://bioconductor.org/packages/release/bioc/vignettes/MCbiclust/inst/doc/MCbiclust_vignette.html).
3. The MCbiclust reference manual (RM) providing documentation to the R functions involved, accessed on the Bioconductor website (https://bioconductor.org/packages/release/bioc/manuals/MCbiclust/man/MCbiclust.pdf).
4. The MitoCarta 1.0 [2] mitochondrial geneset used in the IM.
5. The microarray dataset from the Cancer Cell Line Encyclopedia [11] also used in the IM.

3 Methods

In this section we will discuss in detail only the implications for applying the methodology to the analysis of mitochondrial biogenesis patterns. For complete understanding of the method, the theoretical considerations, and the benchmarking against other algorithms, please refer to Bentham et al. [9].

3.1 Choosing a Geneset

Once a dataset has been chosen (for details on choosing your dataset and judging whether it is suitable, *see* **Note 1**), the first step of using MCbiclust is to select a suitable geneset representing mitochondrial function with the scope of discovering co-regulation patterns in nuclear-encoded mitochondrial genes. This is not a trivial problem as there are genes with different confidence levels of evidence relating them to mitochondria, as well as genes that while not being mitochondrial are highly co-regulated with mitochondrial processes. We consider two alternative methods for geneset selection.

3.1.1 Established Databases with Mitochondrial Genesets

1. MitoCarta [12] in its latest version (2.0) contains 1158 human and mouse genes with strong support of mitochondrial localization.
2. MitoMiner 4.0 [13] is an integrated web resource of mitochondrial localization evidence and phenotype data for mammals, zebrafish, and yeast. The team behind MitoMiner developed the Integrated Mitochondrial Protein Index (IMPI), which in its current version (Q3 2017) includes 1550 genes.
3. Genes that are associated with the gene ontology [14] term "mitochondrion," which contains 1647 genes; genes in the dataset, however, have varying evidence with many inferred from in silico analysis.

The user can decide whether to use one of these datasets in order to select the mitochondrial genes to be analyzed. Alternatively, the intersection (985 genes) or union of all three datasets (1997 genes) could be used. The size of the geneset is an important factor for determining the speed at which MCbiclust completes the analysis. However, an increased geneset size does not necessarily bring any benefits (*see* **Note 2**).

3.1.2 Interaction Networks of Mitochondrial Genes

An alternative strategy to using public lists of known or predicted mitochondrial genes is to compose a list by using a single well-established mitochondrial gene and determine its interactions from the existing correlation structure in the dataset of interest. By taking a single, well-established mitochondrial gene, e.g., a component of the electron transport chain or the mitochondrial

ribosome, the remaining genes can be ordered by the strength of the Pearson's correlation coefficient to the expression of this gene across all of the samples (*see* **Note 3** for details). The geneset, used by MCbiclust to initiate the analysis, can then be selected as the top genes correlated with the mitochondrial gene of interest. The advantages of this method are that (1) it is more likely to include genes that are strongly co-regulated with mitochondrial processes, thus representing a specific function, (2) it is more likely to identify biclusters that are associated with a single mitochondrial gene of interest, and (3) the geneset can be specifically tailored for each dataset. The disadvantage of using this strategy is that the geneset will differ in each user case; thus comparison of results will become more complex or even unfeasible.

Overall, there is no "correct" way to choose a geneset, and the appropriate way should be decided on a case-by-case basis, according to the precise biclusters that are being sought. Nor should an investigator be limited to running a single geneset as the results of MCbiclust using multiple genesets can be compared (*see* **Note 4**).

3.2 Running MCbiclust to Identify Co-regulation of Nuclear-Encoded Mitochondrial Genes

Following the selection and loading the sample set and initial geneset(s) (IM 3.1), "FindSeed" is used to identify a "seed" of samples with high Pearson correlations between the genes in the geneset (IM 3.2; 3.3). Importantly, this method is stochastic and identifies the samples by a greedy search. Thus, in order to find an exhaustive and representative sample set, it is required to run "FindSeed" multiple times. The different strategies to perform this task are discussed in *see* **Note 5**.

Multiple runs of "FindSeed" result in a number of sample seeds. Once a suitable number of sample seeds have been found, the next step is to identify how many distinct modes of regulation of the geneset have been found, i.e., which samples are included and how genes are correlated in these sample seeds. Clearly, if the samples are identical in different seeds, they represent the same pattern, but it is not clear if different samples between seeds represent fundamental differences in regulation or the seed has selected different samples that are representative of the same pattern. For this reason, the different outputs of MCbiclust must be compared at the gene level using a parameter that is called the correlation vector (CV) (*see* IM 3.4). The CV is a vector that quantifies the correlation of each gene measured in the dataset to the average expression of a group of genes in the chosen geneset that are selected as "highly representative" of the bicluster. The CV for each run can then be compared to one another, after which the runs are clustered and then the Silhouette method [15] is used to identify the number of distinct biclusters found in the analysis (IM 4. and RM: SilhouetteClustGroups). The CVs can be averaged across each distinct bicluster, and consequently the samples can be ranked by how well they preserve the correlation within the

geneset. The final output of MCbiclust for each bicluster found is a correlation vector describing the strength of the correlation of each gene to the bicluster and an ordered list of samples (IM 4). Accordingly, the biclusters can be visualized with a distinctive "fork plot" with the ranked samples on the x-axis plotted against the PC1 value from a PCA analysis of the samples within the seed, with the PC1 value being fitted to the remaining samples (IM 3.10). At the beginning of the ranking, the samples separate into an upper and lower fork. By convention, the sign of the PC1 value being chosen is such that the upper fork samples will have genes with a positive correlation vector that are upregulated and genes with a negative correlation vector value that are downregulated. The lower fork samples have the opposite phenotype.

3.3 Analyzing the Resulting Biclusters

The analysis of the resulting biclusters involves the separate analysis of genes and samples. Sample analysis is dataset-specific and involves associating samples in the distribution plot with the different properties (metadata) of sample groups made available for the dataset (for previous examples, *see* Figs. 5, 6, 7, and 8 from Bentham et al. [9]). In patient-derived gene expression samples, this typically includes clinical outcome and genetic and histological subtypes of the disease. Thus, biclusters are the basis of stratification, that is, classification of disease states according to mitochondrial gene expression patterns.

On the other hand, the methods for the analysis of genes can be generalized for different biological applications and are listed below.

3.3.1 Geneset Enrichment Analysis

The simplest analysis is a geneset enrichment analysis on the values of the correlation vector (IM 3.5). The correlation vector can be viewed as a ranked list of genes with values between +1 and −1, and thus geneset enrichment analysis can be run on the entire ranked list, or on selected genesets, e.g., the top positive or negative correlation vector values. At this point, any geneset enrichment method can be used (e.g., DAVID [16], GSEA [17], g:Profiler [18]). The MCbiclust package comes with a specifically designed method that uses the entire correlation vector and applies the Mann-Whitney test to identify gene ontology terms that have significantly different distributions (either more positive or negative) as compared to the entire distribution of values. The output gives the average CV value for each significant term; thus terms that are positive in average (i.e., upregulated in the upper fork, downregulated in the lower fork) can be distinguished from those that are negative in average (i.e., downregulated in the upper fork and upregulated in the lower fork). Interpretation of the significant terms can be challenging, since standard terms often give no other detail than the list of genes that are generally related to "mitochondria." For fine-grain understanding of the differences

in pathways, the individual genes involved must be examined. Different mitochondrial pathways of interest, such as the metabolic enzymes, can each be examined individually. For these metabolic pathways, it is also possible to build diagrams of the pathways to show which parts have been regulated in different ways, e.g., with the pathview R package [19]. On the other hand, geneset enrichment analysis can be useful for identifying non-mitochondrial pathways that are also being simultaneously co-regulated with mitochondria, providing further insight into the biology behind the underlying process.

3.3.2 Comparison of Genesets Across Biclusters

In cases where two or more biclusters are found, it is appropriate to compare the differences in the co-regulation of the genes in the biclusters. In order to identify a module of genes that are regulated in the same way across different biclusters, different visualization techniques can be applied. First, co-regulation of genes in different biclusters can be compared using the CVplot function in MCbiclust (IM 4, RM/CVplot). This function plots the values of the correlation vectors against each other for all the genes, as well as genes in any chosen geneset (e.g., mitochondrial genes). In this way, modules of co-regulated genes across different biclusters can be identified. Alternatively, these groups can be identified through examining the intersection of genesets (e.g., upregulated in bicluster 1, upregulated in bicluster 2, etc.), using Venn diagrams for a small number of groups. If the number of different biclusters is large, a different technique such as UpSet plot [18] can also be used. Examples of these visualization techniques are shown in Fig. 1.

3.4 Identification of Samples in Other Datasets Matching the Bicluster

Once a bicluster has been identified and associated with a particular type of mitochondrial function, a further aim is to determine whether this type of gene expression pattern can be identified in additional datasets. Theoretically this could be achieved by running the entire MCbiclust pipeline on this new dataset and comparing the resulting correlation vectors to understand whether a similar bicluster is present. However, this approach might be time-consuming, and often datasets are not large enough for MCbiclust to reliably identify biclusters (see discussion on the required dataset size in Bentham et al. [9]). Thus, ideally a method is required that can take a small dataset or single sample and determine whether these samples fall into a particular bicluster and whether they belong to a particular branch in the fork distribution.

3.4.1 PointScore Algorithm

A method of choice included in the MCbiclust package to achieve the classification of single samples is the PointScore algorithm (RM: PointScore). This method uses the two genesets (A and B) determining the distribution of samples in the fork pattern (*see* **Note 6** for how these genesets are chosen). Geneset A includes genes

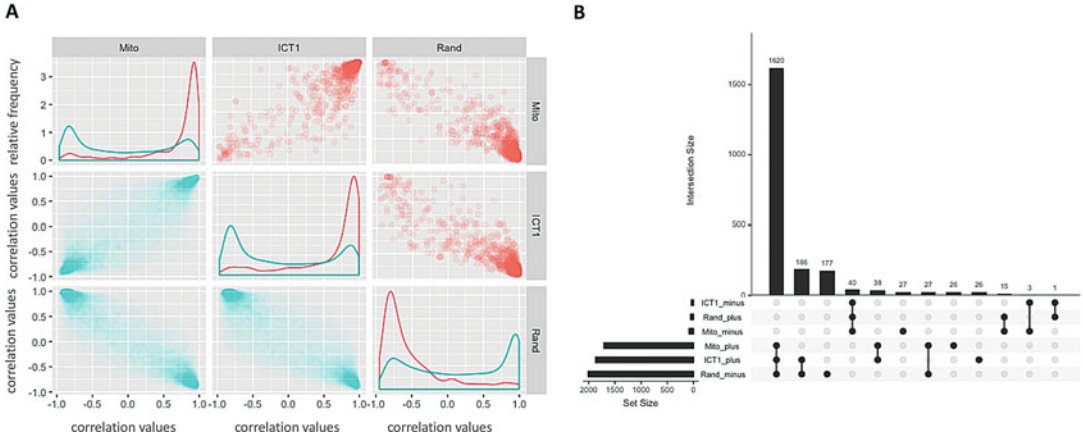


Fig. 1 Comparison of genesets across biclusters using CVplot and UpSet plots Plots produced from a RNASeq dataset from CoMMpass (Relating Clinical Outcomes in MM to Personal Assessment of Genetic Profile) IA9 study (NCT01454297) produced by the Multiple Myeloma Research Foundation (MMRF) containing transcriptomics from 734 patient samples. (**a**) Shows an output of CVplot comparing the correlation vectors from three different runs of MCbiclust with mitochondrial genes from MitoCarta (Mito), a geneset based on the most correlated genes to mitochondrial gene MRPL58 (ICT1) and random (Rand) genesets. The lower diagonal plots (cyan) represent values of the non-mitochondrial genes in the correlation vector, while the upper diagonal plots (red) represent the mitochondrial genes in the correlation vector. In this case a very similar bicluster (in terms of the genes which are most strongly correlated to it) is found from all three initial genesets used. Plots in the diagonal axis show the frequency distribution of mitochondrial (red traces) and non-mitochondrial (cyan traces) genes across the correlation values in the three biclusters. (**b**) Shows the output from the UpSet R package to determine the intersections of the significant genes identified in each of these correlation vectors from MCbiclust's custom geneset enrichment method (*see* Subheading 3.3.1). The significant genesets found in each bicluster have been split into two groups (pos and neg) depending on whether they are associated with genes with positive or negative correlation vector values. The majority of significant terms are shared between all three biclusters, again indicating that these three biclusters are close to identical

upregulated in the upper fork and downregulated in the lower fork, and geneset B contains genes downregulated in the upper fork and upregulated in the lower fork. "PointScore" scores samples based on how well they match this regulation by comparing the genes in the genesets to the median value across the entire dataset. Importantly, this method requires that the dataset contains samples that are representative of all types of regulation seen in the original dataset (where the bicluster was identified), so that the median of the genes can be used as a dividing line for resolving up- or down-regulation in samples. For this reason, the PointScore algorithm cannot be used for single or too few samples.

For single samples or datasets with very few samples, there are two further solutions detailed in Subheadings 3.4.2 and 3.4.3.

3.4.2 Single-Sample GSEA (ssGSEA)

Single-sample GSEA [20], from the Bioconductor package GSVA [21], can be applied by taking the same genesets as used in the PointScore method and calculating the ssGSEA score, based on

how the genes in each geneset are up- or downregulated, compared to other genes in the samples. Therefore, for an upper fork sample, the ssGSEA score for geneset A will be positive, and the score for geneset B will be negative.

3.4.3 First Principal Component Values

It is possible to calculate the PC1 value of the sample (using the R function lsfit from the known PC1 loadings) and compare it directly to the initial bicluster. This technique requires that this sample (or small dataset) is normalized to the original dataset. This is only reliable when the datasets are all measured on the same platform, quantile normalization is performed, and any possible batch effects are removed between experiments (e.g., by using ComBat [22]).

4 Notes

1. As a method MCbiclust is agnostic toward the data platform and can be run on both microarray and RNASeq data. However, for a successful run, the data must meet one important requirement, which the dataset contains enough samples. As a rule of thumb, at an absolute minimum, there should be at least 100 samples in the sample set. In general, the more samples are in a dataset, the more likely MCbiclust is able to find significant biclusters. If the dataset contains few samples, it can be analyzed by comparing to previously analyzed larger sets, as described in Subheading 3.4.
2. MCbiclust calculates the correlation matrices of the chosen geneset repeatedly. Thus, the larger the geneset chosen, the more computation time is needed to perform MCbiclust. In general, a geneset containing more than 1000 genes is sub-optimal and significantly slows down the computation. There is also little advantage to augment the size of the geneset past a certain point, since the biclusters we seek to find are large; as long as a significant number of genes in the geneset are contained in them, they will be found. Additionally, genes outside the geneset can easily be found to be associated with the bicluster in the correlation vector stage of the method (*see* Subheading 3.2). Thus in general, there is no need for genesets to be significantly larger than 1000.
3. This can be achieved simply using base functions such as the apply and cor function in R, e.g., vec1 ← apply(data, MARGIN = 1, FUN = function(x) cor(x, as.numeric(data[gene.loc,])), and then selecting the genes that have the highest correlation, e.g., hicor.loc ← order(abs(vec1), decreasing = TRUE)[seq_len(1000)].

4. Since the choice of the initial geneset is an important factor in determining the results of MCbiclust, running MCbiclust on different initial genesets, e.g., a general mitochondrial one from MitoCarta, as well as different genesets made up of genes that are strongly correlated with mitochondrial genes of interest is a good and recommended strategy.
5. FindSeed should be run enough times to identify all significant biclusters present in the dataset. Typically, this number should be at least 100. However, some biclusters are only identified rarely by random search, and to find these, it is necessary to run FindSeed a very large number of times. In these cases, it is of help to use high-performance computing to run the FindSeed algorithm. An alternative way to find these rare biclusters is to run FindSeed on different initial genesets or run FindSeed on the dataset after removing the most commonly selected samples in the final seed. This way the final seed is forced to include samples not yet chosen.
6. Genesets that represent the upper and lower fork can be created directly from the correlation vector selecting genes with a value greater than a certain threshold, e.g., > 0.9 for upper fork and ≤ 0.9 for lower fork.

Acknowledgment

Funding was provided by the University College London COMPLeX/British Heart Foundation Fund (SP/08/004), the Biochemical and Biophysical Research Council (BB/L020874/1, BB/P018726/1), the Wellcome Trust (097815/Z/11/Z) in the UK, and the Association for Cancer Research (AIRC, IG13447) in Italy.

References

1. Lopez MF, Kristal BS, Chernokalskaya E, Lazarev A, Shestopalov AI, Bogdanova A, Robinson M (2000) High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation. Electrophoresis 21:3427–3440
2. Pagliarini DJ, Calvo SE, Chang B, Sheth SA, Vafai SB, Ong S-E, Walford GA, Sugiana C, Boneh A, Chen WK et al (2008) A mitochondrial protein compendium elucidates complex I disease biology. Cell 134:112–123
3. Thor Johnson D, Harris RA, French S, Blair PV, You J, Bemis KG, Wang M, Balaban RS (2007) Tissue heterogeneity of the mammalian mitochondrial proteome. Am J Physiol Cell Physiol 292:689–697
4. Kuznetsov AV, Hermann M, Saks V, Hengster P, Margreiter R (2009) The cell-type specificity of mitochondrial dynamics. Int J Biochem Cell Biol 41:1928–1939
5. Scarpulla RC (2008) Transcriptional paradigms in mammalian mitochondrial biogenesis and function. Physiol Rev 88:611–638
6. Hock MB, Kralli A (2009) Transcriptional control of mitochondrial biogenesis and function. Annu Rev Physiol 71:177–203
7. Duchen MR, Szabadkai G (2010) Roles of mitochondria in human disease: figure 1. Essays Biochem 47:115–137
8. Jones AWE, Yao Z, Vicencio JM, Karkucinska-Wieckowska A, Szabadkai G (2012) PGC-1

family coactivators and cell fate: roles in cancer, neurodegeneration, cardiovascular disease and retrograde mitochondria-nucleus signalling. Mitochondrion 12:86–99

9. Bentham RB, Bryson K, Szabadkai G (2017) MCbiclust: a novel algorithm to discover large-scale functionally related genesets from massive transcriptomics data collections. Nucleic Acids Res 45:8712–8730
10. Gentleman R, Carey V, Bates D, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J et al (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5:R80
11. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehár J, Kryukov GV, Sonkin D et al (2012) The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483:603–607
12. Calvo SE, Clauser KR, Mootha VK (2015) MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res 44(D1):D1251–D1257. https://doi.org/10.1093/nar/gkv1003
13. Smith AC, Blackshaw J a, Robinson AJ (2012) MitoMiner: a data warehouse for mitochondrial proteomics data. Nucleic Acids Res 40: D1160–D1167
14. Consortium TGO (2000) Gene ontology: tool for the unification of biology. Nat Genet 25:25–29
15. Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65
16. Huang DW, Lempicki R a, Sherman BT (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57
17. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al (2005) Geneset enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102:15545–15550
18. Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H, Vilo J (2016) g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res 44:W83–W89
19. Luo W, Brouwer C (2013) Pathview: an R/bioconductor package for pathway-based data integration and visualization. Bioinformatics 29:1830–1831
20. Barbie DA, Tamayo P, Boehm JS, Kim SY, Susan E, Dunn IF, Schinzel AC, Sandy P, Meylan E, Fröhling S et al (2010) Systematic RNA interference reveals that oncogenic KRAS- driven cancers require TBK1. Nature 462:108–112
21. Hanzelmann S, Castelo R, Guinney J (2013) GSVA: geneset variation analysis for microarray and RNA-Seq data. BMC Bioinf 14:7
22. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8:118–127

Chapter 25

Using the Human Genome-Scale Metabolic Model Recon 2 for Steady-State Flux Analysis of Cancer Cell Metabolism

Lake-Ee Quek and Nigel Turner

Abstract

Flux analysis is performed to infer intracellular metabolic activity using measured rates. By applying the highly curated human metabolic reconstruction Recon 2 as the reference model, the investigation of cancer cell metabolic fluxes can encompass the full metabolic potential of a human cell. But in its full form, Recon 2 is unsuitable for conventional metabolic flux analysis due to a large number of redundant elements. Here, we describe a procedure to reduce Recon 2 to an appropriate scale for cancer cell flux analysis such that calculated flux intervals are still informative, without compromising the opportunity to explore alternative pathways encoded in Recon 2 that may reveal novel metabolic features.

Key words Flux analysis, Genome-scale model, Constraint-based, MATLAB, COBRA Toolbox, Cancer metabolism, Model reduction

1 Introduction

Metabolic processes convert material (e.g., nutrients, cofactors) into energy and biomass to achieve cell growth and maintenance. Quantifying these processes is particularly relevant to cancer cell metabolism because metabolic pathways (e.g., aerobic glycolysis) are often implicated in a tumor's adaptation for survival, proliferation, and aggressiveness [1, 2]. Metabolic experiments are generally designed with contrast in mind to induce changes in these pathways, and thus it is imperative to be able to determine the activity of these pathways, either directly or by inferring from relevant models. The latter approach is effective when many independent but indirect measurements support the same conclusion. Steady-state metabolic flux analysis (MFA), while conventionally viewed as a tool to

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-1-4939-9027-6_25) contains supplementary material, which is available to authorized users.

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6_25,

estimate absolute cellular fluxes, is a simple yet versatile technique to integrate a variety of cellular measurements to achieve this [3]. The required measurements are primarily growth and extracellular rates but can include rates estimated from stable isotopic tracing (e.g., ^{13}C-MFA) and accumulation of radiolabelled products and enzyme capacity and directionality determined from kinetics, thermodynamics, and gene/protein expression [4–9]. A persistent challenge, however, is deploying a model that fits the investigation purpose and scope.

Because of the diversity of cancer metabolism and the associated context of investigations [10], it is difficult to find a one-size-fits-all metabolic model for flux analysis. We risk prematurely eliminating pathways that may transpire to become novel cancer metabolism features [11]. But with modifications, Recon 2 can serve this purpose. Recon 2 is reconstructed from genes in the human genome [12]. It embodies years of effort in prescribing accurate gene-protein-reaction mappings for each metabolic enzyme, such as enzyme complexes, atom and charge balancing, cofactor usage (e.g., NADP vs. NAD), and compartmentalization. Recon 2 encodes the full capacity of human metabolism, and conceivably cancer metabolism is well-described within this metabolic space [13]. There is a difficult balance in tuning the dimension of the model between a narrowly scoped model with coerced flux results and a loose model with non-informative flux intervals. Rather than building one metabolic model, manually, we closed this gap by dynamically extracting a representative subset of reactions from Recon 2 based on the context of the metabolic investigation and the information/data available (Table 1).

Here, we describe a method to perform MFA of cancer cell metabolism using reactions from Recon 2 (Fig. 1). The extraction process is facilitated by metabolic data and a core reaction template that reflects investigator's preference (*see* **Note 1**). By "activating" and "deactivating" individual reactions using data (e.g., rates) and

Table 1
Recon 2.2 network size before and after the reduction

	Recon 2.2	Reduced model
Total reactions	7785	339
Exchange reactions	746	31
Infeasible reactions[a]	1872	0
Metabolites	5324	334
Degrees of freedom[b]	2840	23

[a]After performing modifications described in Subheading 3.2
[b]Number of reactions minus rank of stoichiometric matrix

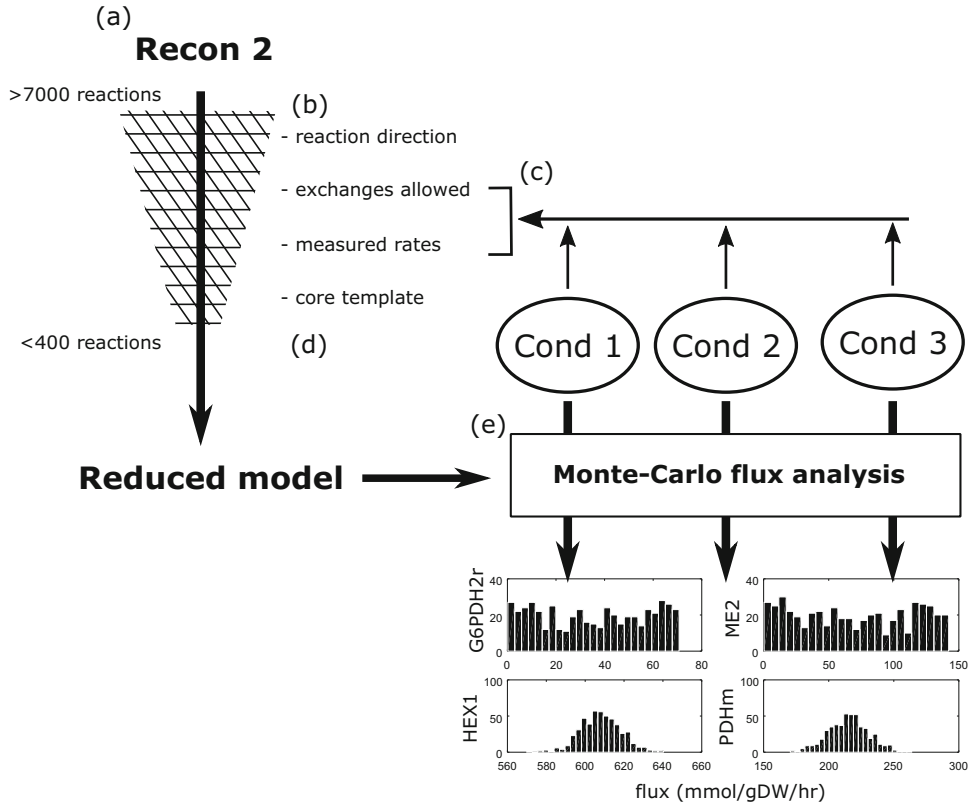


Fig. 1 Workflow for reducing Recon 2. (**a**) Starting from the full human reconstruction, (**b**) reaction directionalities are revised to restrict ATP and NADPH production. (**c**) Cell culture conditions and measured rates are used to constrain allowable exchange reactions and magnitude of network inputs/outputs. (**d**) Manual flux specifications are applied before maximizing the elimination of reactions not contained in the core template. (**e**) The reduced model is used for flux analysis. A Monte-Carlo approach is used to generate flux distributions

manual specifications, associated pathways in Recon 2 are autonomously assembled into a complete functional network, with preference for core reactions. Classical mainstream reactions tend to be selected according to the template, but the choices are superseded by the data and the manual specifications. Pathway rerouting is thus readily accomplished, and efforts can focus on picking active reactions that precipitate the desired metabolic pathways. Rates and directionality information are derived from metabolite and expression data. Hypothetical information can be incorporated as well. Computation takes less than a minute; thus the model extraction and flux analysis can be performed in rapid explorative cycles. This method requires considerable coding, but we have provided MATLAB scripts to help exemplify the method, as well as to serve as a functional application.

Flux analysis is a technique to extrapolate intracellular metabolic activity from experimental data using a model. The method described provides a simple and reproducible approach to manipulate content of Recon 2 for flux analysis of cancer cell metabolism.

2 Materials

The method described will follow closely the development of the MATLAB application (provided in the Supplementary Materials) but is not limited as a MATLAB implementation. The minimum computational requirements are software packages that can read SBML files and perform mathematical programming and optimizations (*see* **Note 2**).

2.1 Software

1. Install latest MATLAB (MathWorks) (*see* **Note 3**). The version used here is R2015a.
2. Install latest Gurobi Optimizer (Gurobi Optimization) (*see* **Note 4**). The version used here is Gurobi-7.5.2.
3. Install COBRA Toolbox (CT) for reading SBML files (*see* **Note 5)** [14].

2.2 Model and Inputs

1. Download Recon 2.2 (MODEL1603150001) as the reference model (*see* **Note 6**) [15].
2. Use Microsoft Excel or equivalent for model specification, as shown in modelSpecs.xlsx (*see* **Note 7**).
3. As guidance, use the list of Recon 2 core template reactions and allowable exchange reactions provided in the Supplementary Materials (*see* **Note 8**).
4. As guidance, use rates for cell growth and extracellular exchanges provided in the Supplementary Materials (*see* **Note 8**).

3 Methods

3.1 Prepare MATLAB Environment

1. Check that Gurobi and CT in MATLAB paths are set correctly (*see* **Note 9**).
2. Change MATLAB working directory to project folder.
3. Read Recon 2 SBML using CT function `readCbModel()`. Save Recon 2 workspace variable to file (*see* **Note 10**). Repeat this step when the SBML file is altered.

3.2 Load Recon 2 and Modify Model

1. Load Recon 2 from MATLAB file.
2. Update the directionality of the reactions by changing flux lower and upper boundaries. The focus of these changes is to

restrict reactions that can produce ATP and NADPH to a few conventional ones (*see* **Note 11**).

3. Restrict exchange reactions by changing the respective flux lower and upper boundaries to zero. The focus of these changes is to eliminate uptake of trivial substrates, i.e., not present in cell culture media (*see* **Note 11**).
4. Update coefficients of the biomass equation if required (*see* **Note 8**).

3.3 Test Feasibility with and Without Measured Rates

1. Perform linear programming (LP) to check that modifications made in Subheading 3.2, **steps 2–4** did not render the model infeasible, i.e., the biomass equation "biomass_reaction" still carries flux. If solver returns an infeasible status, revise modifications made in Subheading 3.2 until a feasible solution can be achieved before proceeding (*see* **Note 12**).
2. Perform quadratic programming (QP) to adjust measured exchange rates and growth rates such that they are balanced. Check residuals of the adjusted rates to make sure they are acceptable (*see* **Note 13**).

3.4 Reduce Recon 2 and Generate Submodel

1. Apply adjusted rates to Recon 2 by fixing the lower and upper boundaries of the exchange flux to the adjusted value. The constraints are relaxed slightly using a small slack value (1e-6) to improve model feasibility (*see* **Note 14**).
2. Perform mixed-integer linear programming (MILP) to minimize use of reactions that are not listed in the core reaction template (*see* **Note 15**). Repeat this step in case MILP produces nonunique solutions. The MATLAB script (*see* *modelReduction.m*) will generate three solutions, and choose the first solution as the reduced model by default.
3. Using LP, perform randomized elimination of reactions that have non-zero flux in the MILP flux solution (*see* **Note 16**). This step is optional.
4. Reactions that have non-zero flux are exported as the reduced model for flux analysis. Combining two or more reduced models is possible. Remove adjusted rates applied in Subheading 3.4, **step 1**, but keep modifications made in Subheading 3.2.
5. Save reduced model as MATLAB file and as a CT model (*see* **Note 5**).

3.5 Monte-Carlo Flux Analysis

1. Load reduced model from MATLAB file.
2. Apply measured rates and perform QP. This step is the same as Subheading 3.3, **step 2**, but the reduced model is used instead of Recon 2. Save adjusted rates, intracellular fluxes, and residuals as "optimum" set.

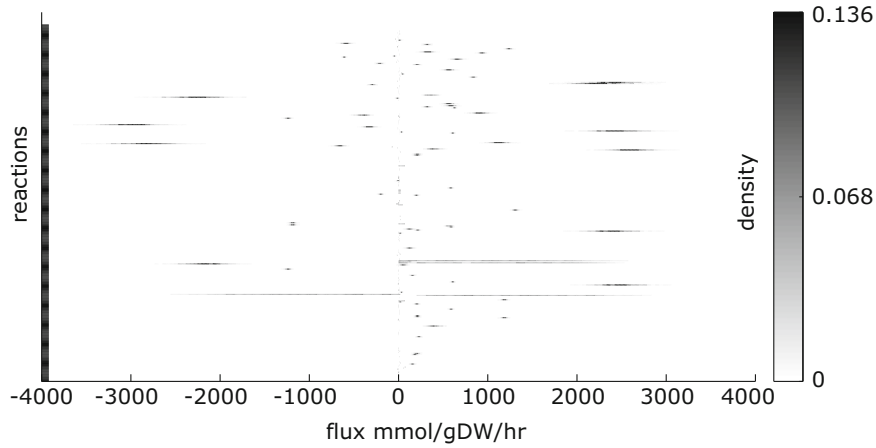


Fig. 2 Flux distributions generated by a Monte-Carlo approach. Histograms are shown from a bird's-eye view, with density represented by shading

3. Calculate the value of underdetermined fluxes by randomized uniform sampling. This is achieved by performing flux variability analysis (FVA) on a randomly chosen reaction and fixing the chosen reaction's flux to a value sampled from the FVA interval randomly with a uniform distribution. This step is performed to all reactions in a randomized order (*see* **Note 17**).
4. Generate flux distributions by a Monte-Carlo approach by repeating Subheading 3.5, **steps 1–3** (>100 iterations). With this procedure, however, measured rates are corrupted first with normally distributed noise before performing QP (*see* **Note 17**). Save adjusted rates, intracellular fluxes, and residuals as "corrupted" set.
5. Compile and plot flux results from the corrupted set (Fig. 2) (*see* **Note 18**).

4 Notes

1. The method is an abridged version of an existing approach of reducing Recon 2 for steady-state flux analysis of HEK cell culture [4]. The published method involves complicated handling of raw gene expression data and numerous subjective steps, which are now summarized into a single editable "core reaction template." This simplifies manipulation of reactions contained in Recon 2.
2. The optimization approaches required are linear programming (LP), mixed-integer linear programming (MILP), and quadratic programming (QP). The generalized formulation for these approaches is shown as Eq. 1. QP and LP are similar,

except that the matrix Q and the constant k are zero in LP. For steady-state flux analysis, the inequality constraint is reformulated as the metabolite balance equation shown as Eq. 2, where S and v are the stoichiometric matrix and flux vector, respectively [3]. In MILP, the vector x contains both continuous and binary [0,1] values.

$$\begin{aligned} &\text{Minimize } c' \tilde{n} x + x' \tilde{n} Q \tilde{n} x + k \\ &\text{such that :} \\ &lb \leq x \leq ub \\ &A \cdot x \leq b \end{aligned} \tag{1}$$

$$S \cdot v \leq 0 \tag{2}$$

3. MATLAB is preferred because both Gurobi and CT installations come with MATLAB interface. MATLAB script modelReduction.m contains the steps performed in Subheadings 3.1–3.4. MATLAB script monteCarloMFA.m contains the steps performed in Subheading 3.5.
4. Gurobi has all three LP, QP, and MILP solvers. Academic free licensing is available.
5. COBRA Toolbox (CT) is currently being updated (version 3.0 in the pipeline) (https://opencobra.github.io/). The recommended installation procedure is via GitHub, but the required MATLAB function files can also be downloaded as a ZIP file. Running `initCobraToolbox.m` will run an update procedure. CT contains a large number of functions for constraint-based modelling and model editing that may be useful. The description of these functions can be found in OpenCOBRA's website, under "Analysis" and "Reconstruction"-->"Refinement." The website also has very detailed tutorials. Typical analysis functions include FBA and FVA, which are useful for flux simulations and checking.
6. The method described is based on Recon 2.2 [15], which is a 2016 version of the human reconstruction. The SBML model (MODEL1603150001) can be downloaded from EMBL-EBI BioModels Database. This version supersedes Recon 2.04 (2015), which can be browsed and downloaded from the Virtual Metabolic Human (VMH) website (https://vmh.uni.lu). Recon 3D currently is in the pipeline.
7. A spreadsheet interface is used to control and modify the contents of Recon 2. The file (modelSpecs.xlsx) is provided in the Supplementary Materials. Recon 2.2's reactions are tabulated in this file. Excel's filter is a useful way to visualize specific groups of reactions.

8. Information used to modify and reduce Recon 2 is listed in the spreadsheet, instead of being scripted in MATLAB. This provides traceability and opportunity for version control. The information includes (1) core template reactions, (2) allowable exchange reactions, (3) measured rates and errors, (4) reaction directions, (5) flux constraints, and (6) biomass composition. The core template provided has 363 reactions. The original biomass coefficients reflect the makeup of a generic mammalian cell (in mmol/gDW), and not for any specific cell type. See "Instructions" tab for more detailed explanation of the spreadsheet. Model specification can be equally accomplished in plain text files as well.
9. Run `initCobraToolbox.m` after installing Gurobi and specifying Gurobi path in MATLAB; otherwise COBRA Toolbox may not assign the correct solver. Add Gurobi routines to MATLAB path using `gurobi_setup.m`. This file is located in Gurobi's MATLAB interface directory. Add CT scripts to MATLAB path using `initCobraToolbox.m`. This file is located in OpenCOBRA's root directory. Rerun these functions if MATLAB path was reset or when Gurobi or CT functions are no longer detected.
10. CT takes about 15 min to load the Recon 2.2 SBML file. To save time, save the model variable as a local MATLAB file, and reload this file at the start of each MATLAB session, or reset the variable.
11. Controlling the sources of ATP and NADPH is important because energy and redox balance can alter flux outcomes [4, 16]. For cancer metabolism, one can restrict the production of ATP to mainstream pathways like glycolysis, TCA cycle, and oxidative phosphorylation and NADPH to oxidative pentose phosphate pathway and malic enzymes. Restriction is accomplished by altering the directionality of reactions, i.e., ATP and NAPDH are consumed by reactions unless allowed intentionally. Controlling the network inputs and outputs is also important to properly balance the measured rates (when performing QP). Inputs and outputs should be restricted to nontrivial substrates and products and essential nutrients required by the biomass equation. Cell culture dataset tends to be amino acid-centric [4]; allowable exchanges are thus the 20 amino acids, glucose, lactate, and small molecules H^+, H_2O, Pi, O_2, CO_2, NH_3, SO_4, and urea. To incorporate other substrates or by-products (e.g., fatty acids, acetate) [6], simply allow the exchange of these metabolites of interests.
12. When the lower boundary of "biomass_reaction" flux is fixed to 1, an infeasible outcome means that biomass cannot be produced. This is caused by the deactivation of an essential

source, sink, and/or pathway route. To troubleshoot, modifications are toggled ON/OFF until a feasible solution is achieved. Conditionally essential metabolic features may require special annotation for future references.

13. While experimental measurements are acquired independently, the model inputs and outputs are occasionally coupled, for example, essential amino acid uptakes to the growth rate and glucose consumption to lactate production. QP is performed to adjust measure rates in a weighted least square fashion (Eq. 3) such that they become consistent according to the model [3]. Equation 3 can be reformulated as a QP problem, with measurement errors σ converted into the weight matrix Q. Calculated residuals show the "movement" of the rates and can be manipulated by tuning σ. For example, reduce error term of growth rate to induce greater adjustments of amino acid rates. Terms with zero residuals are free (uncoupled) fluxes.

$$\text{Minimize } \left(\frac{v_{\text{measured}} - v_{\text{simulated}}}{\sigma}\right)^2 \tag{3}$$

14. Measured rates do not necessarily need to belong to the control or treatment set per se. It is important to use rates that represent average activity and directionality of exchange reactions. If rates are opposite between conditions, then model reduction may be performed more than once, with the final model produced by overlaying two or more reduced models.
15. For the MATLAB application, the binary values 0 and 1 were formulated to represent activation and deactivation of reactions, respectively. Hence the objective of the MILP problem is set to maximize "1"s among reactions that are not listed as core reactions.
16. MILP does not produce a minimum model, i.e., all reactions are essential. Further model reduction can be accomplished by randomly eliminating active reactions, one at a time, and checking for feasibility by LP. The objective function can be used to bias the choice of reactions eliminated, e.g., keep reactions that maximize ATP production.
17. Determining flux distribution by a Monte-Carlo approach allows propagation of measurements errors to the flux of all other reactions. Measured rates are corrupted using Eq. 4 before being adjusted and used as flux constraints. Although underdetermined fluxes have uniform distributions within the flux interval, they also have complex couplings to other underdetermined fluxes. Hence, a greedy sampling approach was used. Underdetermined fluxes were constrained in a

randomized order, each to a random flux uniformly sampled from that instance of flux interval.

$$v_{\text{corrupt}} = v_{\text{measured}} + \sigma \cdot N(0,1) \quad (4)$$

18. Figure 2 is an example of a bird's-eye view plot of histograms for all reactions. MATLAB's zoom function can be used to inspect individual histograms and to produce a more legible y-axis (reaction ID).

References

1. Günther UL, Chong MG, Volpari T, Koczula KM, Atkins K, Bunce CM, Khanim FL (2015) Metabolic fluxes in cancer metabolism. In: Mazurek S, Shoshan M (eds) Tumor cell metabolism: pathways, regulation and biology. Springer, Vienna, pp 315–348. https://doi.org/10.1007/978-3-7091-1824-5_14
2. Pavlova NN, Thompson CB (2016) The emerging hallmarks of cancer metabolism. Cell Metab 23(1):27–47. https://doi.org/10.1016/j.cmet.2015.12.006
3. Quek LE, Dietmair S, Kromer JO, Nielsen LK (2010) Metabolic flux analysis in mammalian cell culture. Metab Eng 12(2):161–171. https://doi.org/10.1016/j.ymben.2009.09.002
4. Quek LE, Dietmair S, Hanscho M, Martinez VS, Borth N, Nielsen LK (2014) Reducing Recon 2 for steady-state flux analysis of HEK cell culture. J Biotechnol 184:172–178. https://doi.org/10.1016/j.jbiotec.2014.05.021
5. Liu M, Quek LE, Sultani G, Turner N (2016) Epithelial-mesenchymal transition induction is associated with augmented glucose uptake and lactate production in pancreatic ductal adenocarcinoma. Cancer Metab 4:19. https://doi.org/10.1186/s40170-016-0160-x
6. Quek LE, Liu M, Joshi S, Turner N (2016) Fast exchange fluxes around the pyruvate node: a leaky cell model to explain the gain and loss of unlabelled and labelled metabolites in a tracer experiment. Cancer Metab 4:13. https://doi.org/10.1186/s40170-016-0153-9
7. Martinez VS, Quek LE, Nielsen LK (2014) Network thermodynamic curation of human and yeast genome-scale metabolic models. Biophys J 107(2):493–503. https://doi.org/10.1016/j.bpj.2014.05.029
8. Hackett SR, Zanotelli VR, Xu W, Goya J, Park JO, Perlman DH, Gibney PA, Botstein D, Storey JD, Rabinowitz JD (2016) Systems-level analysis of mechanisms regulating yeast metabolic flux. Science 354(6311). https://doi.org/10.1126/science.aaf2786
9. Krycer JR, Fazakerley DJ, Cater RJ, CT K, Naghiloo S, Burchfield JG, Humphrey SJ, Vandenberg RJ, Ryan RM, James DE (2017) The amino acid transporter, SLC1A3, is plasma membrane-localised in adipocytes and its activity is insensitive to insulin. FEBS Lett 591(2):322–330. https://doi.org/10.1002/1873-3468.12549
10. Vander Heiden MG, DeBerardinis RJ (2017) Understanding the intersections between metabolism and cancer biology. Cell 168(4):657–669. https://doi.org/10.1016/j.cell.2016.12.039
11. Jiang L, Boufersaoui A, Yang C, Ko B, Rakheja D, Guevara G, Hu Z, DeBerardinis RJ (2017) Quantitative metabolic flux analysis reveals an unconventional pathway of fatty acid synthesis in cancer cells deficient for the mitochondrial citrate transport protein. Metab Eng 43(Pt B):198–207. https://doi.org/10.1016/j.ymben.2016.11.004
12. Thiele I, Swainston N, Fleming RM, Hoppe A, Sahoo S, Aurich MK, Haraldsdottir H, Mo ML, Rolfsson O, Stobbe MD, Thorleifsson SG, Agren R, Bolling C, Bordel S, Chavali AK, Dobson P, Dunn WB, Endler L, Hala D, Hucka M, Hull D, Jameson D, Jamshidi N, Jonsson JJ, Juty N, Keating S, Nookaew I, Le Novere N, Malys N, Mazein A, Papin JA, Price ND, Selkov E Sr, Sigurdsson MI, Simeonidis E, Sonnenschein N, Smallbone K, Sorokin A, van Beek JH, Weichart D, Goryanin I, Nielsen J, Westerhoff HV, Kell DB, Mendes P, Palsson BO (2013) A community-driven global reconstruction of human metabolism. Nat Biotechnol 31(5):419–425. https://doi.org/10.1038/nbt.2488
13. Yizhak K, Chaneton B, Gottlieb E, Ruppin E (2015) Modeling cancer metabolism on a genome scale. Mol Syst Biol 11(6):817. https://doi.org/10.15252/msb.20145307

14. Schellenberger J, Que R, Fleming RM, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmanian S, Kang J, Hyduke DR, Palsson BO (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat Protoc 6(9):1290–1307. https://doi.org/10.1038/nprot.2011.308
15. Swainston N, Smallbone K, Hefzi H, Dobson PD, Brewer J, Hanscho M, Zielinski DC, Ang KS, Gardiner NJ, Gutierrez JM, Kyriakopoulos S, Lakshmanan M, Li S, Liu JK, Martinez VS, Orellana CA, Quek LE, Thomas A, Zanghellini J, Borth N, Lee DY, Nielsen LK, Kell DB, Lewis NE, Mendes P (2016) Recon 2.2: from reconstruction to model of human metabolism. Metabolomics 12:109. https://doi.org/10.1007/s11306-016-1051-4
16. Smith AC, Eyassu F, Mazat JP, Robinson AJ (2017) MitoCore: a curated constraint-based model for simulating human central metabolism. BMC Syst Biol 11(1):114. https://doi.org/10.1186/s12918-017-0500-7

INDEX

Majda Haznadar (ed.), *Cancer Metabolism: Methods and Protocols*, Methods in Molecular Biology, vol. 1928,
https://doi.org/10.1007/978-1-4939-9027-6,

D

E

F

G

H

I

L

M

N

O

P

Q

R

S

T

U

V

W