TRANSLATIONAL CARDIOMETABOLIC GENOMIC MEDICINE

Edited by

ANNABELLE RODRIGUEZ-OQUENDO

ELSEVIER

AMSTERDAM • BOSTON • CAMBRIDGE • HEIDELBERG
LONDON • NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, UK
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN: 978-0-12-799961-6

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

Typeset by TNQ Books and Journals
www.tnq.co.in

Printed and bound in the United States of America

Contributors

Rodrigo Alonso Obesity and Lipid Units, Department of Nutrition, Clínica Las Condes, Santiago de Chile, Chile

Sahar Al Seesi Department of Computer Science & Engineering, University of Connecticut, Storrs, CT, USA; Department of Immunology, Carole and Ray Neag Comprehensive Cancer Center, University of Connecticut School of Medicine, Farmington, CT, USA

Veronica Alvarez Obesity and Lipid Units, Department of Nutrition, Clínica Las Condes, Santiago de Chile, Chile

Thomas E. Cheatham, III Department of Medicinal Chemistry, L.S. Skaggs Pharmacy Institute, University of Utah, Salt Lake City, UT, USA

Ada Cuevas Obesity and Lipid Units, Department of Nutrition, Clínica Las Condes, Santiago de Chile, Chile

Anthony M. DeAngelis University of Connecticut Health Center, Farmington, CT, USA

Fei Duan Department of Immunology, Carole and Ray Neag Comprehensive Cancer Center, University of Connecticut School of Medicine, Farmington, CT, USA

Charles R. Farber Departments of Public Health Sciences and Biochemistry and Molecular Genetics, Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

Magdalena Farías Obesity and Lipid Units, Department of Nutrition, Clínica Las Condes, Santiago de Chile, Chile

Alexis C. Frazier-Wood Department of Pediatrics, USDA/ARS Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX, USA

Lita A. Freeman Lipoprotein Metabolism Section, Cardiopulmonary Branch, National Heart, Lung and Blood Institute, Bethesda, MD, USA

Rodrigo Galindo-Murillo Department of Medicinal Chemistry, L.S. Skaggs Pharmacy Institute, University of Utah, Salt Lake City, UT, USA

David Herrington Department of Internal Medicine, Section of Cardiology, Wake Forest School of Medicine, Winston-Salem, NC, USA

Angela Kueck Carole and Ray Neag Comprehensive Cancer Center, University of Connecticut School of Medicine, Farmington, CT, USA

Ion I. Mandoiu Department of Computer Science & Engineering, University of Connecticut, Storrs, CT, USA

Gareth J. McKay Centre for Public Health, Queen's University Belfast, Belfast, Northern Ireland, UK

Larry D. Mesner Departments of Public Health Sciences and Biochemistry and Molecular Genetics, Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

Steven Myint Center for Enterprise and Development, Duke-NUS Medical School, Singapore; Nanyang Business School, Nanyang Technological University, Singapore; Inex Private Ltd, Singapore; Plexpress Oy, Finland

Adam Naj Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Waqas Qureshi Department of Internal Medicine, Section of Cardiology, Wake Forest School of Medicine, Winston-Salem, NC, USA

Theodore P. Rasmussen Department of Pharmaceutical Sciences, University of Connecticut, Storrs, CT, USA; University of Connecticut Stem Cell Institute, University of Connecticut, Storrs, CT, USA; Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, USA

Alan T. Remaley Lipoprotein Metabolism Section, Cardiopulmonary Branch, National Heart, Lung and Blood Institute, Bethesda, MD, USA

Stephen S. Rich Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

Annabelle Rodriguez-Oquendo University of Connecticut Health Center, Farmington, CT, USA

Meaghan Roy-O'Reilly University of Connecticut Health Center, Farmington, CT, USA

Pramod K. Srivastava Department of Immunology, Carole and Ray Neag Comprehensive Cancer Center, University of Connecticut School of Medicine, Farmington, CT, USA; Carole and Ray Neag Comprehensive Cancer Center, University of Connecticut School of Medicine, Farmington, CT, USA

Kasey C. Vickers Department of Medicine, Vanderbilt University, School of Medicine, Nashville, TN, USA

CHAPTER

1

Metabolomics and Cardiovascular Medicine

Waqas Qureshi, David Herrington

Department of Internal Medicine, Section of Cardiology, Wake Forest School of Medicine, Winston-Salem, NC, USA

1. INTRODUCTION

Cardiovascular disease is a complex, multifactorial disease that currently affects 1.5 billion people worldwide [1]. Many causative pathways through which risk factors act are still unclear. Metabolomics has recently emerged as a promising field that may help elucidate intricate details of the relationships between changes in human biology and complex cardiovascular disease phenotypes.

Metabolites are the end-products of metabolic processes occurring in cellular organelles. These are molecules smaller than 1 kDa ($1\ \text{Da} = 1.66 \times 10^{-27}$ kg). Historically, studies of metabolism have typically focused on a narrow range of metabolites or specific metabolic pathways. In contrast, metabolomics generally involves a more comprehensive assessment of many metabolites and pathways. The Metabolomics Society defines metabolomics as "comprehensive characterization of the small molecule metabolites in biological systems. It can provide an overview of the metabolic status and global biochemical events associated with a cellular or biological system" [2]. Similarly, Nicholson [3] defined metabolomics as a "quantitative measurement of the dynamic multi-parametric metabolic response of living systems to pathophysiologic stimuli or genetic modification…" [3]. The term "metabolome" was first coined by Oliver et al. [4] and Tweeddale et al. [5] in 1998. Fiehn et al. used the term "metabolomics" for the first time in 2001 [6]. Although the term was initially used in the literature pertaining to plants and agricultural research, with advancement in analytical tools and sample extraction methods, human and animal medical fields have also embraced these terms and approaches.

A major advantage of studying the metabolome is that metabolites are generally considered to be molecular phenotypes much closer to clinical traits of interest—reflecting the integrated effects of both upstream molecular signals (e.g., genome, epigenome, transcriptome, and proteome) and environmental factors (e.g., diet, psychological stressors, microbiome, and environmental exposures—collectively referred to as the exposome). This concept was built on central dogma from the 1950s that there is a linear and unidirectional flow of information from genome to phenome. More recently, this concept has evolved into a complex network-based framework to account for multiple interconnected factors at all levels that determine a clinical phenotype. Nevertheless, the metabolome remains at the intersection of many factors that ultimately reflect or determine the health of an individual (Figure 1). Thus, it is essential to understand both the metabolome and its interactions with other -omics to apply metabolomics to practice.

The application of this concept in clinical and population research is still relatively new. However, already many important insights about the relationship between the metabolome and clinical cardiovascular disease have emerged, and several studies have identified key genome−metabolome associations that better characterize the molecular signals contributing to pathogenesis of atherosclerosis, hypertension, and other cardiovascular conditions. In this chapter, we provide an overview

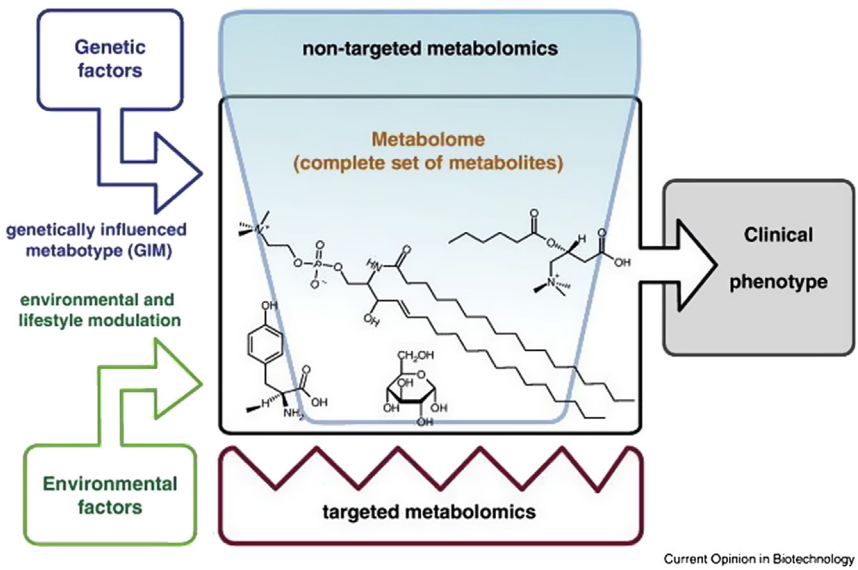

FIGURE 1 System biology network. *Adapted from Adamski et al. Curr Opin Biotechnol 2013; 24:39−47.*

of metabolomics trends and techniques, and review current understanding of the relationships among the metabolome, the genome, and the cardiovascular disease.

2. METABOLOMIC METHODS

A metabolomic study is composed of several discrete steps including bio-sampling, separation, metabolite detection, and data analysis (Figure 2). Each step is critically important to ensure valid and interpretable results.

2.1 Bio-specimen Selection and Handling

Particular attention should be given to the type of bio-sample and sampling techniques, as these determine the types of metabolites that need to be studied (Table 1). While collecting bio-samples, care must be taken to avoid degradation and contamination of collected sample. Some necessary considerations for sampling include timing (e.g., 24-h sampling versus collection after a 12- to14-h fasting period), sudden dietary changes (potentially interfere with the metabolome for over a week), bacterial or fungal contamination (especially for urine samples), medication intake (metabolites of medications may change the metabolome), diurnal variation in the hormones (anabolic and catabolic hormones vary during the day and may alter the metabolome), randomization, and transport and storage conditions of samples.

Serum is generally the preferred blood bio-specimen for metabolomic studies. This is because plasma retains fibrinogen and other coagulation cascade proteolytic enzymes that can continue to be active ex vivo and alter the metabolome or interfere with metabolite detection. Furthermore, ethylenediaminetetraacetic acid is typically used as the anticoagulant in plasma samples and can denature chromatograms and proteins in a way that may influence results. Serum has fewer potentially interfering coagulation cascade enzymes and proteins. However, even in serum there remain active enzymes that can continue to modify the metabolome ex vivo. To counteract these enzymatic processes, serum can be collected on ice. It is recommended to use samples stored at $-20\,^{\circ}C$ within 7 days and samples stored at $-80\,^{\circ}C$ within 1 month. However, in one report, changes in the metabolome were negligible after storage for 2.5 years [7]. In the case of repeated usage, fewer than three freeze−thaw cycles is advisable [7].

Until now, urine has been the most common sample for human metabolomics studies. It has many advantages, such as a noninvasive method of collection, lack of special preparation for collection, almost no

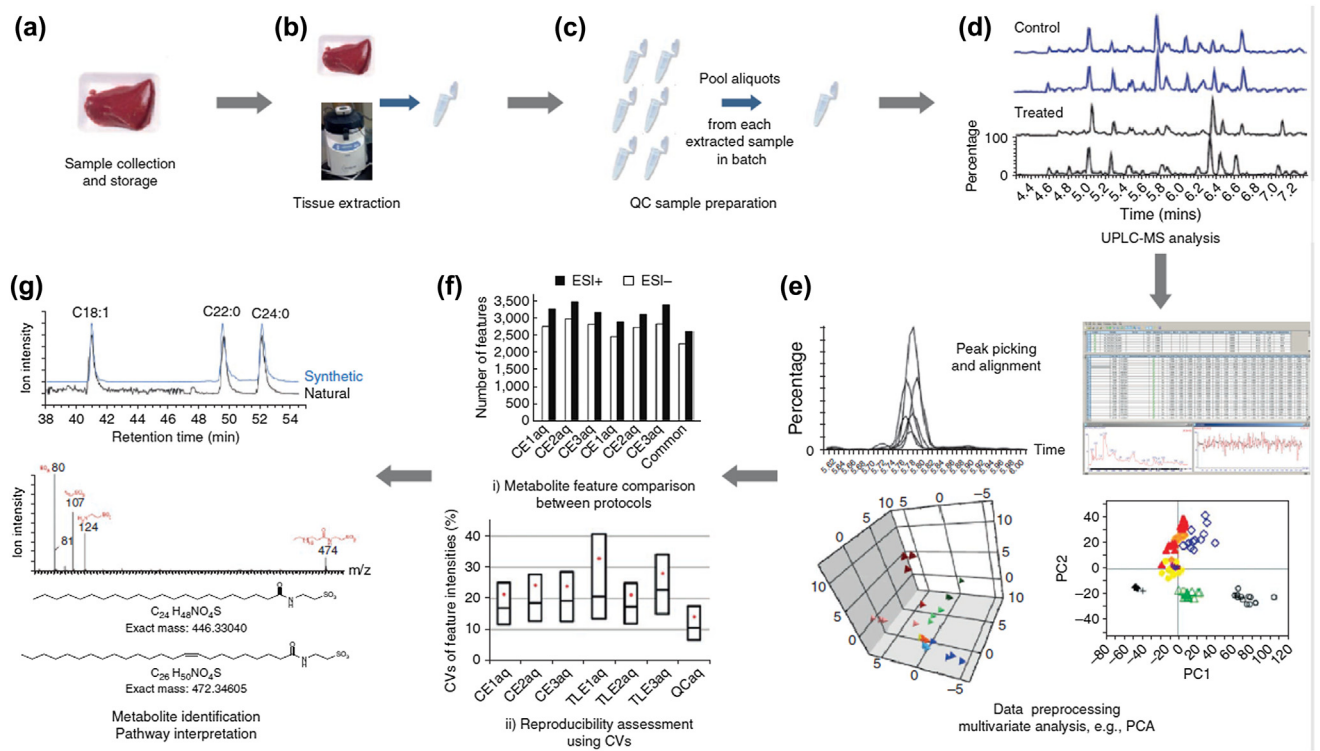

FIGURE 2 Schematic view of steps of a metabolomic study. *Adapted from Want et al. Nature Protoc 2013; 8:17–32.*

TABLE 1 Sample Types Used in Metabolomics Studies

Sample	Advantage	Disadvantage	Approximate volume
Plasma/serum	Minimally invasive, readily available, provides information of metabolic footprint of many metabolic reactions	May not reflect specific tissue level changes (e.g., cardiac changes) Tissue level metabolic fluxes cannot be measured Requires deproteination of mass spectrometer Anticoagulant used for plasma collection may interfere with the endogenous metabolites	150–550 μL
Urine	Noninvasive, readily available, provides information of metabolic footprint of many metabolic reactions	May not reflect specific tissue level changes (e.g., cardiac changes) Tissue level metabolic fluxes cannot be measured Contains high amount of urea that can damage mass spectrometer Differences in pH of urine leads to difficulty in evaluation of various metabolites especially when nuclear magnetic resonance spectroscopy was used	500–1000 μL
Primary cell culture	Tissue level information can be obtained, metabolic fluxes can be studied	Takes time, expensive, technically demanding, and requires complex procedures, phenotype may change over time	100–600 μL
Tissue	Tissue level information can be obtained, metabolic fluxes can be studied, accurate description of phenotype	Takes time, expensive, technically demanding, and requires complex procedures, invasive for some of the specific tissues	>20 mg

sample pretreatment, and lower protein content, which helps to increase the sensitivity of identifying other metabolites [8]. Urine, like blood or plasma, provides a metabolomic "footprint" of the whole body, not necessarily a single organ. Usually a 24-h urine sample is collected for metabolomics studies; this method requires that participants receive

certain instructions and follow specific procedures. Improper procedures may lead to bacterial overgrowth secondary to either contamination or infection that can affect the urine metabolome. Many other types of samples have also been used in metabolomics research (e.g., cell lysates and spinal fluid), although these types of bio-samples are less amenable to large-scale clinical and population research.

Prior to sample separation, typical sample preparation steps are required. Standardization of these sample preparation steps is pivotal to avoid random or, even worse, systematic bias in the results. For example, alcohol-based extraction is a common procedure employed in sample extraction. However, the pyruvate pathway is up-regulated when boiling ethanol or freeze–thaw extraction is used, but down-regulated when cold methanol is used [9]. The ideal method should be highly reproducible, simple, rapid, unselective, and include a metabolism-quenching step. The latter is not frequently used in human studies, but is important because many active enzymes may influence metabolites in the collected sample (especially in plasma or serum samples). If the enzymes continue to remain active, they may change the concentrations of metabolites of interest. Special attention should be paid to this matter when evaluating metabolites that are affected by rate-limiting steps, for example, glucose-6-phosphate or ATP.

2.2 Metabolite Separation Methods

Chromatography is the principal separation technique used to obtain fine resolution of metabolites for metabolomics studies. It plays a key role in obtaining analytical data needed in metabolic profiling. The separation of metabolites is achieved by interactions of analytes with solvent and a variety of stationary phases (e.g., solid or liquid chromatography columns and so on—see below).

The basic theory of chromatography was first described in 1903 when a chromatogram made up of calcium carbonate (stationary phase) was inoculated with a leaf extract (sample) mixed with ethanol (mobile phase). For metabolomics research, a sample (or sample extract) is dissolved in a mobile phase (liquid or gas). This mobile phase is then forced through an immiscible stationary phase. The component of sample that has an affinity for the stationary phase will travel slowly, whereas the component that repels or does not dissolve in the stationary phase is the first one to reach the other end of chromatography column. By making changes in these phases, metabolites can be separated with a high degree of resolution, and a high-quality chromatogram can be obtained.

There are a number of definitions and key concepts that are useful to keep in mind when considering metabolomic studies.

2.2.1 Retention Time

At equilibrium, an analyte in mobile phase and in stationary phase can be given as

$$M_{mobile} \rightleftharpoons M_{stationary}$$

In this equilibrium reaction, the *equilibrium constant K* is defined as the ratio of concentration (in moles) of an analyte in the stationary versus the mobile phase. This is also called the *partition coefficient*. The time needed from injection of sample to formation of peak for a given analyte is known as *retention time* (T_R). The time taken for the mobile phase to pass through column is T_M. A constant called the retention factor (k') is defined as the migration rate of the analyte through the column. This is given as

$$k' = \frac{T_R - T_M}{T_M}$$

The ideal retention factor is from 1 to 5. Too quick a retention factor may obscure the sensitivity of detection, but if the retention factor is slow, chromatography will be too time-consuming [10].

2.2.2 Solvent Strength

Another important factor to consider is the change in solvent (mobile phase) strength during a run. The term *isocratic* is used for solvents that do not change their strength during the entire run (e.g., 60:40 alcohol:water). If the solvent composition changes through the run, it is called a gradient mobile phase. The change may be gradual (linear gradient) or stepwise (step gradient) [11].

2.2.3 Resolution

Resolution (R) is the degree of separation between two analytes. It is given as

$$R = \frac{2(T_{R1} - T_{R2})}{W_1 + W_2}$$

Subscripts 1 and 2 are for analytes 1 and 2. The *selectivity factor* (α) describes the separation between the centers of two peaks of analytes. Resolution is a function of efficiency of the column, selectivity of analytes, and retention time of analytes. These relationships are analyzed using the Purnell equation, which describes the resolution between two peaks on a chromatogram as follows:

$$R = \frac{\sqrt{N_2}}{4}\left(\frac{\alpha - 1}{\alpha}\right)\left(\frac{k'_2}{1 + k'_2}\right)$$

N_2 is the number of theoretical plates in a column of the second peak (a measure of efficiency of the chromatogram), and k'_2 is the retention factor for the second peak.

Baseline resolution occurs when $R > 1.5$ is achieved. Resolution is important for peak identification and helps in accurate integration and quantification of analytes. Improving the resolution of peaks is a primary goal of chromatography. The retention factor can be changed by altering the temperature in gas chromatography (GC) or by changing the composition of the mobile phase in liquid chromatography (LC).

2.2.4 Peak Capacity

Peak capacity (P_c) is the maximum number of peaks that can be separated in a chromatogram with a resolution of 1. Gradient solvents can give higher peak capacities than isocratic solvents. Peak capacity is described in the following equation:

$$P_c = 1 + \left(\frac{T_g}{W_1}\right)$$

where T_g is gradient time and W_1 is width of the peak.

A *Van Deemter plot* depicts relationships between column height and mobile phase velocity (Figure 3). It is usually a J-shaped curve that shows the optimal flow rate of a particular solvent for a given column height.

Below, we discuss the three major forms of separation techniques: GC, LC, and capillary electrophoresis (not exactly chromatography).

The GC is mainly used for volatile organic metabolites. This technique offers high resolution and high sensitivity. It is inexpensive and requires small amount of samples. The technique is limited by the length of time it takes, which increases when derivatization procedures are needed for nonvolatile substances. It is also not suitable for thermally labile samples. This type of chromatography is usually coupled with mass spectrometry

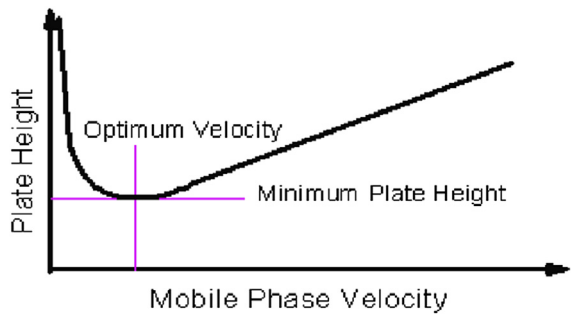

FIGURE 3 Van Deemter plot.

(GC-MS) to provide sensitive and accurate quantification of low-molecular-weight metabolites. Examples of usual known metabolites separated by this technique include cholesterol, fatty acids, hydrophilic vitamins, terpenoids, sugars, and some alkaloids [12].

The LC uses the liquid mobile phase to separate analytes. A smaller particle size (<1 kDa) leads to higher back pressure, and a higher pressure is required to move the particle through the mobile phase. This form of specialized LC is called high-performance liquid chromatography (HPLC). There are many types of HPLC, and the main types used in metabolomics are normal-phase chromatography, reversed-phase chromatography (RPLC), hydrophilic interaction chromatography (HILIC), ion exchange chromatography, size exclusion chromatography, chiral chromatography, and supercritical fluid chromatography.

There are several solvents used in LC that can be distinguished based on their polarity:

$$\text{Polar solvents: Water} > \text{methanol} > \text{acetonitrile} > \text{ethanol} > \text{oxydipropionitrile}$$

$$\text{Non-polar solvents: N} - \text{decane} > \text{N} - \text{hexane} > \text{N} - \text{pentane} > \text{Cyclohexane}$$

HILIC and RPLC are frequently used forms of LC; their relative advantages and disadvantages are listed in Table 2. These techniques are mainly used for separation of secondary neutral metabolites, lipids, flavonoids, and hydrophobic vitamins. Another way to increase resolution of metabolites is utilized in ultra-performance liquid chromatography (UPLC), where the particle size is reduced at the cost of an increase in back pressure.

Capillary electrophoresis is a different type of separation technique that uses an electric field applied to sub-millimeter channels (capillaries) to separate analytes. The separation usually takes about 20 min [13]. This method is useful for the separation of nucleic acids, vitamins (hydrophobic), amino acids, coenzymes, and amino acids.

There are several advantages of linking these separation techniques with analytical tools such as nuclear magnetic resonance (NMR) and MS. Processing samples with these analytical tools will show peaks; however, if the analyte of interest exists in very small concentrations (as is usually true in biological fluids) other compounds with similar-sized peaks may be difficult to differentiate from the analyte of interest. Chromatography uses the physical and chemical properties of the mobile and stationary phases to further increase resolution of this detection process.

TABLE 2 Advantages and Disadvantages of LC Methods

	RPLC	HILIC
Advantages	Well-understood retention mechanisms Widely applicable Faster equilibrium Wide variety of mobile phases available	Retains highly polar analytes not retained by RPC Less interference with matrix components Complementary selectivity to RPC Polar compounds retained more than parent compound High organic mobile phases promotes enhanced ESI-MS response Direct injection of precipitate supernatant without dilution is possible Facilitates use of lower volume sample Increased throughput
Disadvantages	Poor peak shape with basic analytes	Sensitive to sample diluent Mechanism not well understood Longer equilibration time

RPLC; reversed phase liquid chromatography, HILIC; hydrophilic interaction liquid chromatography.

The interface between the separation apparatus and the detectors is also critical. Snyder and Kirkland note desirable characteristics of such an interface [14]:

1. There should be no reduction in chromatographic performance over time.
2. There should be no uncontrolled chemical modification of the analyte during chromatography.
3. The sample transfer to mass spectrometer should be high and efficient.
4. The interface should be reliable, easy to use, and inexpensive.
5. The interface should give a low chemical background.
6. The interface should be capable of operating across a wide variety of chromatography conditions.
7. The interface should not affect the vacuum requirements of the mass spectrometer and should be compatible with all capabilities of the mass spectrometer.
8. The interface should provide quantitative information with reproducibility better than 10%, with low limits of detection. This response should be linear over a wide range of sample sizes.

Flow-injection analysis refers to a technique when the sample is ionized and injected directly into the mass spectrometer without a

preceding separation step. This method is mainly used for targeted metabolomic analysis where the unique mass spectra are known beforehand and are easily identified in the mass spectrum. This method takes a few minutes.

2.3 Metabolite Detection

NMR spectroscopy and MS are the primary analytical technologies of metabolomics. Both should be considered as complementary technologies rather than separate methods for a metabolomic study. Almost all of the high abundance molecules can be detected by NMR spectroscopy. For low-abundance molecules, MS is more suitable.

NMR spectroscopy is a technique by which a spectrum is created from resonance of nuclei that, when in a magnetic field, absorb and emit electromagnetic radiation. There are several types of NMR spectroscopy techniques and several different isotopes used for NMR analysis including ^{1}H, ^{13}C, and ^{31}P. Various NMR experiments useful for metabolomics are listed in Table 3.

NMR utilizes the principle that atomic nuclei with nonzero spin when subjected to a strong magnetic field and irradiated with a characteristic radiofrequency wave, emit radiofrequency signals. These signals, which can be detected by an antenna or coil, reflect the number and proximity of the specific nuclear isotopes in question that can be used to infer the quantity and structural features of the compounds in the sample [15]. Because NMR is nondestructive, a sample can be reused many times. Another advantage is that the sample does not require extensive preparation before use.

Limitations of NMR include its lower sensitivity; it can only detect high-abundance metabolites with concentrations >100 nmol/L or 5–10 μM in size. Also, if one-dimensional NMR is used, many metabolites may have overlapping peaks, especially those in low abundance. This limitation has been resolved by developing sophisticated two-dimensional NMR methods (not discussed in this chapter). Use of LC as a method of separating metabolites improves the sensitivity of analyses (typically <100 metabolites can be resolved at one time).

In contrast, the MS takes advantage of the fact that the velocity of charged particles in an electronic field vary in a precise manner as a function of mass and charge. This permits positive identification of a wide range of molecules including those that are typically part of the metabolome. Since its invention in 1946, time of flight mass spectroscopy (TOF-MS) has grown significantly in its capabilities. The mechanism of ionization has continued to evolve in analyzing biological samples, to address the fact that organic analytes tend to be fragile and often dissociate and break down when hard ionizing techniques are used. Matrix-