Metabolomics Ready for the Prime Time?

Manuel Mayr, MD, PhD

Metabolomics is one of the most rapidly growing areas of contemporary science. Although classic genetics aims to link variations in the DNA sequence directly to distinct phenotypes, “-omic” technologies allow us to shift the focus from the specific gene to the actual effects of the gene itself. Because neither the transcriptional or protein profile can be directly correlated with metabolite concentrations, the importance of measuring small-molecule metabolites has become increasingly clear. In view of the rapid progress in metabolomic techniques, metabolomics is expected to become more widely applied to cardiovascular research. Metabolomics brings the promise of the identification of potential biomarkers and alterations in biochemical pathways, which will facilitate the transition from a reductionistic approach to a more integrated science. Because the relative lack of attention given to the system behavior hampers our progress in translating basic science research into clinical applications, the holistic nature of these emerging techniques may yield valuable new strategies for the prevention and treatment of cardiovascular diseases.

By analogy to the genome, the metabolome is defined as the total complement of small-molecule metabolites found in or produced by an organism. The most recent estimates place the number of endogenous metabolites (metabolites synthesized by enzymes encoded in the human genome) at approximately a few thousand, far less than had been previously predicted. Importantly, the size of the exogenous metabolome (metabolites not synthesized in the body but consumed as food or generated by host-specific microbes) is far greater, and there is often a spatial separation between metabolite synthesis and use. Hence, although genes, proteins, and metabolites are intimately connected in biological systems and their interactions with environmental changes are reflected in the metabolome, gene or protein expression may not directly correlate to metabolite concentrations from the same region (Figure 1). Thus, there is a clear need for an additional readout at the metabolite level, and the promise of “metabolomic profiling” is to achieve a quantitative and qualitative assessment of a subset of metabolites in complex samples such as bodily fluids and tissues.

Metabolomic Technologies

The huge diversity of chemical compounds with different physicochemical properties means that coverage of the human metabolome is practically impossible to achieve with a single analytic method. Two of the most commonly used analytic tools, nuclear magnetic resonance spectroscopy (NMR) and mass spectrometry (MS), provide detailed information on metabolite structure, which is necessary for successful metabolite identification (the Table).

NMR Studies

Magnetic resonance spectroscopy can detect a wide range of biochemical metabolites, provided that they contain an NMR nucleus and are present in sufficiently high abundance. NMR nuclei (\(^{1}\)H, \(^{13}\)C, \(^{19}\)F, \(^{31}\)P, etc) possess a property known as nuclear spin, the spinning motion of the nucleus about its own axis. In contrast, molecules with an even number of both protons and neutrons have zero spin because of the tendency for both neutrons and protons to form pairs to the effect that the individual spins cancel each other out. Consequently, the abundant isotopes of carbon and oxygen, \(^{12}\)C and \(^{16}\)O, do not produce NMR signals. Protons (\(^{1}\)H) have been used mostly for magnetic resonance spectroscopy because of their high natural abundance in organic compounds. Phosphorus-NMR (\(^{31}\)P) is of particular interest for studies on energy metabolism and phospholipid analysis.

NMR-based methods have proved to be very robust and reproducible, and metabolites can be identified by chemical shift measurement. Chemical shift, the separation of resonance frequencies from an arbitrarily chosen reference frequency, usually is expressed in terms of the dimensionless units of parts per million. The resonance frequency of a given nucleus is modified slightly (typically by a few parts per million) by its molecular environment because of the screening effect of the electron cloud. This allows the distinction and identification of different molecules containing the given nucleus. Spectra are plotted with decreasing frequency left to right. The parameters that characterize each peak include its resonance frequency, its height, and its width at half-height. The height (maximum peak intensity) or the area under the peak yields a relative measure of the concentration of nuclei. An internal standard is added to the samples for chemical shift calibration and quantification.

Different methods sharing the same fundamental technology are available for performing analyses on small volumes of bodily fluids (a few microliters), tissue extracts, and intact samples such as bodily fluids and tissues.
tissues. For example, $^{31}$P magnetic resonance spectroscopy can measure high-energy metabolites in vivo, but the cardiac and breathing motion has to be tagged to synchronized acquisition of magnetic resonance spectra. Magic angle spinning is used to obtain metabolic profiles from small pieces of intact tissues ex vivo. Both techniques circumvent the need for metabolite extraction and offer the advantage that the tissue is preserved; however, the major disadvantage is that the sensitivity and resolution are further compromised compared with NMR analysis on tissue extracts. Despite its relative insensitivity, NMR-based metabolomics has been used successfully to study cardiovascular diseases because many of the NMR-detectable metabolites are found at central hubs of metabolism (Figure 2). However, this also implies that metabolites detected by NMR may be poor markers for specific diseases because they can be perturbed by numerous conditions.

### MS Studies

Because of its superior sensitivity compared with NMR, MS is being used increasingly as an alternative approach in the assessment of metabolites. The quantity measured by MS is

### Table. Comparison of Metabolomic Technologies

<table>
<thead>
<tr>
<th>Criteria</th>
<th>NMR</th>
<th>GC-MS</th>
<th>LC-MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>Poor</td>
<td>Good</td>
<td>Excellent</td>
</tr>
<tr>
<td>Metabolites detected</td>
<td>High-concentration organic compounds in a solution</td>
<td>Ideal for volatile, nonpolar compounds</td>
<td>Ideal for nonvolatile compounds in a solution</td>
</tr>
<tr>
<td>Robustness</td>
<td>Good</td>
<td>Reasonable</td>
<td>Reasonable</td>
</tr>
<tr>
<td>Speed</td>
<td>Rapid</td>
<td>Depending on chromatography</td>
<td>Depending on chromatography</td>
</tr>
<tr>
<td>Quantification</td>
<td>Good; no standards needed</td>
<td>Poor; standards are needed</td>
<td>Poor; standards are needed</td>
</tr>
<tr>
<td>Identification</td>
<td>By chemical shift calibration</td>
<td>By mass and fragmentation</td>
<td>By mass and fragmentation</td>
</tr>
<tr>
<td>Problems</td>
<td>Peak overlap</td>
<td>Volatility of metabolites</td>
<td>Ionization of metabolites</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>Exact adjustment of pH required after metabolite extraction</td>
<td>Nonvolatile compounds have to be derivatized to be amenable to analysis</td>
<td>Problem of ion suppression and adduct formation</td>
</tr>
<tr>
<td></td>
<td>In vivo magnetic resonance spectroscopy and magic angle spinning circumvent the need for extraction but result in a further loss of sensitivity</td>
<td>Some metabolites cannot be made volatile even with derivatization</td>
<td>Different metabolites detected in positive and negative scanning mode</td>
</tr>
</tbody>
</table>

Metabolite identification is difficult because of the lack of comprehensive spectral libraries.

---

LC indicates liquid chromatography.
the mass-to-charge ratio of ions formed from molecules, usually separated by chromatography because the power of this technology depends on separation along with detection. The current application to metabolomics has 2 major platforms: gas chromatography MS (GC-MS) and liquid chromatography MS (the Table).

GC-MS is used widely in the analysis of volatile nonpolar metabolites. In GC-MS, the sample is vaporized and carried through the chromatography column in the gas phase. Two-dimensional chromatography before MS has been shown to substantially increase metabolite identifications. Several involatile compounds (including polar metabolites) also may be made amenable to GC-MS analysis by chemical derivatization techniques, but this is a time-consuming procedure, and involatile metabolites are more readily analyzed by mass spectrometers equipped with an electrospray ionization source. Electrospray ionization, the most commonly used technique in organic MS, permits the measurement of virtually any compound that can be dissolved. In principle, biological fluids such as plasma and urine can be introduced directly after some degree of pretreatment such as protein precipitation. Electrospray ionization, however, is affected by the ionizability of the metabolites, and ionization suppression (known as matrix suppression) is a major concern. Matrix suppression arises when particular analytes preferentially ionize over less polar metabolites in the complex mixture. Thus, quantitative changes may be misrepresented as a result of matrix effects, causing either suppression (underestimation) or enhancement (overestimation) of the target analyte response. Hence, appropriate steps have to be taken to minimize matrix effects throughout the application method.

The most common ways to address the challenge of ionization suppression are sample cleanup technologies such as liquid-liquid extraction, solid-phase extraction, protein precipitation, and separation of complex metabolite mixtures by using liquid chromatography. Nanospray, with its lower flow rates (100 to 200 nL/min), is a promising alternative to the more routinely used higher flow rates in conventional electrospray analysis (200 to 2000 μL/min) because of the inherent sensitivity enhancement of low flow rate operation, but it is technically and operationally more challenging. The reduction in the flow rate and the internal diameter of the tip lead to a decrease in the initial size of liquid droplets, smaller sample consumption, and an increase of the ionization yield with a noticeable reduction in the adverse effects of ion suppression. Nonetheless, as with all MS approaches, reliable quantification can be made only for metabolites for which stable isotope-labeled internal standards or closely matched analogs are available.

Importantly, liquid chromatography MS is lagging behind GC-MS with respect to data analysis and the assignment of detected ions using library matching. Although top-end instruments such as Fourier transform ion cyclotron resonance mass spectrometers offer unsurpassed mass accuracy (<1 ppm), allowing the empirical formula of metabolites to be calculated from their accurate mass alone, unequivocal identifications can be obtained only by elucidation of the chemical structure. Unfortunately, the current metabolite databases lack comprehensive spectral libraries, which would allow the masses of the observed fragmentation products to be compared with fragmentation patterns of known metabolites, in part as a result of the substantial variation in spectrum appearances across different techniques. Thus, the identification of unknown metabolites remains one of the biggest analytic challenges in MS, but efforts such as the Human Metabolome Project1 (http://metabolomics.ca/) and the Metlin metabolite database (http://metlin.scripps.edu/) aim to create these much-needed data repositories.

Apart from GC-MS and liquid chromatography MS, researchers are examining the advantages of other types of MS, including matrix-assisted laser desorption/ionization. Matrix-assisted laser desorption/ionization MS has often been restricted to the analysis of higher-molecular-weight metabolites (>500 Da) because inherent matrix cluster ions create a multiplicity of signals in the low-mass range of the spectrum, which can interfere with the detection of low-molecular-weight metabolites. However, novel matrixes have been developed that produce minimal spectral noise in the low-molecular-weight region of interest, ie, 9-aminoacridine. Advantages of matrix-assisted laser desorption/ionization MS are that sample preparation is extremely fast and easy, no derivatization is required, and higher levels of buffer or salt contamination can be tolerated compared with electrospray ionization. Although electrospray ionization is regarded as the more versatile ionization method, matrix-assisted laser desorption/ionization MS has been applied for sugar and lipid analyses. Its potential in cardiovascular metabolomics has recently been demonstrated by Sun et al,12 who profiled 285 metabolites from murine myocardium and identified 90 metabolites.

The Quest for Biomarkers

The past several years have seen an increase in the number of metabolomic studies, with the overwhelming majority of work being carried out in plasma and other bodily fluids. Metabolic profiling has recognized potential for revealing novel biomarkers of disease; however, akin to the beginnings of genomics and proteomics, there are often incorrect views of the immediate impact of new technologies on the biomarker pipeline. A key limitation is the fact that the human metabolome is not well characterized, and although metabolomic technologies are rapidly evolving, the software tools for data analysis are still in their infancy. Besides data analysis, the challenges in metabolomics are the integration of different analytic approaches and the control of preanalytical, analytical, and postanalytical variables.

Patient and Sample Preparation

Careful consideration of experimental and biological variation and potential bias in the selection of study groups is essential for metabolomic analysis because dietary and environmental factors affect metabolite measurements. Ideally, patients must fast and abstain from smoking before samples are taken at a specified time during the day to minimize the effects of circadian variation. Notably, if samples are collected from symptomatic patients, they might be more anxious when sampled, unlike control subjects whose samples may be taken as part of a routine checkup. In this case, the
observed differences in the metabolic profiles may simply be a consequence of different levels of stress hormones and have nothing to do with whether the patient has the disease. Sample storage is another cause of artifacts in metabolomic analysis. Some metabolites such as phosphocreatine are so labile that they can be quantified accurately only in vivo. Although most metabolites are preserved if samples are immediately snap-frozen in liquid nitrogen and kept at temperatures below -80°C, differences in storage time may still account for the classification obtained between patient and control samples. Finally, the procedure used for metabolite extraction has to be robust and highly reproducible. Bodily fluids may be analyzed directly after precipitating proteins. Tissues are first pulverized under liquid nitrogen. While perchloric acid extraction is widely used in NMR spectroscopy to isolate water-soluble metabolites from tissues, dual-phase extraction with chloroform/methanol allows the simultaneous assessment of water-soluble and lipid metabolites.

**Data Analysis**

Given the potential “noise” in metabolomic measurements, models with hundreds of metabolites are not acceptable. Robust models tend to have <25 significant variables derived from 3 to 10 metabolites. Otherwise, the effects of overfitting in the analysis of multivariate data produced by multiplexed “-omic” technologies are of major concern. Central steps in defining metabolic differences are pattern recognition techniques such as principal-component analysis. Principal-component analysis replaces a group of variables with a smaller number of new variables, called principal components, which are linear combinations of the original variables. The first principal components capture the rough shape of the signals contained in the data sets, whereas finer details are contained in subsequent principal components. Thus, a principal-component analysis decomposes the signals into a sum of other signals. Projecting the observation on one of these axes generates a new variable designed to maximize the description of the variance in the data set. After the principal-component analysis, each sample can be represented by its set of scores, which can then be used as variables for other classification methods, ie, linear discriminant analysis. Linear discriminant analysis constructs a separating hyperplane from an optimal projection that maximizes the distances between groups while minimizing the distances within the groups. Besides primary statistics, metabolite expression can be analyzed in pathway analysis programs such as Ingenuity (Ingenuity systems) and MetaCore (GeneGo). The latter also provides a portal for chemical structures. In addition, researchers are working on standardizing metabolomic measurements and reporting to allow better comparison and exchange of metabolomic data.

**Plasma Metabolite Markers of Coronary Artery Disease**

Multivariate analysis of 1H-NMR spectra of blood sera was shown previously to predict angiographically defined advanced coronary artery disease with >90% accuracy and specificity. The reported predictive power in this pioneering study was based on a relatively small number of samples and depended mainly on the lipid regions of the spectra. Common variables, including gender, statin and hormonal status, diet, and exercise, are known to affect the amount and composition of lipids in the circulation. Later studies demonstrated that the value of 1H-NMR diagnostic assays was reduced substantially by the extent to which these confounding factors affect the accuracy and specificity of prediction, and the predictive power for coronary artery disease was particularly poor in patients treated with statins. Thus, metabolomic technologies are unlikely to replace angiography in the foreseeable future.

**Plasma Metabolite Markers of Myocardial Injury**

Using a combination of liquid chromatography and selective reaction monitoring on a triple-quadrupole mass spectrometer, Sabatine et al quantified 500 metabolites released after “exercise-induced” myocardial injury in a cohort of 36 patients. Plasma levels of γ-aminobutyric acid decreased strikingly in the cases but remained unchanged in controls. In addition, members of the citric acid pathway were significantly overrepresented in the metabolites that changed specifically in the setting of myocardial ischemia. This finding is consistent with previous reports that there is a constant efflux of citric acid cycle intermediates from cardiac muscle (cataplerosis) that falls in the acute settings of ischemia to defend ATP production. Notably, citric acid cycle intermediates such as succinate and α-ketoglutarate are present in micromolar concentrations in blood and have unexpected signaling functions by acting as ligands for orphan G-protein–coupled receptor (GPR). Succinate infusion, for example, increased the blood pressure in animals via the renin-angiotensin system in wild-type but not in GPR91-null mice. However, it is important to acknowledge that many of the low-molecular-weight peaks in the plasma samples were not identified and that most of the plasma metabolites implicated as potential biomarkers for the cardiovascular system may originate from noncardiac sources. In addition, troponin I and T, the current gold standard for diagnosis of myocardial infarction, provide reliable rule-in and rule-out markers within the first hours of the event. Thus, although an additional panel of defined markers to fine-tune the diagnosis would be desirable, finding a better diagnostic test for myocardial injury will be a significant challenge. Nonetheless, in a recent NMR-based study, exercise-induced myocardial ischemia was even predicted by metabolic analysis of blood samples obtained before exercise.

**Advanced Lipoprotein Profiling**

NMR spectroscopy also is used for lipoprotein analysis; the particle concentrations of the different-sized lipoprotein sub-classes are given by the measured amplitudes of the characteristic lipid methyl group NMR signals they emit. Altered lipoprotein particle composition has been detected by NMR in prediabetic individuals and in patients with hypertension. The latter finding was confirmed and refined in another cohort by the use of an adaptive, intelligent binning algorithm for data processing of NMR spectra. Studies on lipoprotein abnormalities in prediabetic individuals re-
revealed that very low-density lipoprotein particle size and small high-density lipoprotein particles, rather than low-density lipoprotein particle size, may act as significant contributors to incident diabetes. The preponderance of small low-density lipoprotein particles appears to reflect underlying insulin resistance rather than a causal relation between this phenomenon and the development of diabetes. Notably, lipidomics is a separate branch of metabolomics that focuses on a systems-based study of all lipids within the cell. Its application to cardiovascular disorders has been reviewed extensively elsewhere.

Metabolomics in Human Cardiac Tissue

An alternative approach to finding biomarker candidates for the empirical diagnosis of cardiovascular disease is to study samples close to the source of the disease. This approach offers 2 advantages: The potential biomarkers should be markedly enriched and easier to detect compared with plasma, and direct measurements of metabolites in cardiac tissue allow an assessment of how metabolic substrates are used in the tissue of interest. Using microarray analysis, Barth et al found that in patients with permanent atrial fibrillation (AF), there were a ventricularization of gene expression and a prominent upregulation of transcripts involved in metabolic activities, including several glycolytic enzymes, suggesting a switch to glucose metabolism in permanent AF. Using high-resolution 1H-NMR spectroscopy, we showed ketone metabolism to be affected during persistent AF, which was further substantiated by findings from proteomic studies that complemented the metabolomic data set.

Studies relating to changes in persistent AF, however, do not enable extrapolation of metabolic changes associated with postoperative AF, which develops in 30% of patients undergoing cardiac surgery. We therefore analyzed cardiac tissues from patients undergoing 2 types of common cardiac surgeries, coronary artery bypass grafting and nonrheumatic valve surgery, and grouped them into patients who maintained sinus rhythm and those who developed AF postoperatively after cardiac surgery. These further studies aimed at assessing the role of metabolic derangements in instigating AF have shown a discordant regulation in glycolytic and lipid metabolism in cardiac patients who developed AF, regardless of the underlying cardiac pathophysiology. Moreover, the ratio of glycolytic end products to end products of lipid metabolism correlated positively with the time of onset of postoperative AF. These observations are consistent with a previous finding in rats that glycolytic inhibition predisposes to AF. Thus, before the onset of AF, metabolite changes were observed in small but highly matched cohorts, illustrating that a pervasive physiological abnormality might underlie the susceptibility to AF after cardiac surgery. Larger studies should establish whether interrogation of cardiac tissue samples such as atrial tissues samples routinely removed and discarded during the cardiac surgery can be used as a diagnostic tool to stratify patients according to risk and to effectively target preventive therapy.

Integrating Proteomics and Metabolomics to Understand Models of Human Diseases

Besides its application in the discovery of disease biomarkers, metabolomics can contribute to the elucidation of pathophysiological mechanisms. Current research is based heavily on reductionism, whereby complex biological systems are investigated by detailed examination of the properties of its constituent parts in hypothesis-driven experiments. Our profound knowledge of individual signaling pathways and molecules is an obvious success of this approach. Biological systems, however, are not an assembly of linear signaling pathways but function as scale-free networks, and “-omic” technologies will be essential to bridge the gap between molecular and systems biology. A single gene mutation may cause alterations of metabolites of seemingly unrelated biochemical pathways, which is likely to happen when genes are constitutively overexpressed or knocked out by targeted disruption of the endogenous gene. Although transgenic mice have provided valuable insights into cardiovascular disease mechanisms, there is a general lack of control for compensatory changes and spurious off-target effects after
genetic manipulations. Furthermore, many transgenic and knockout mice are generated on a mixed background, and the background of disease models can be dynamic in nature. Strain background has a major impact on the global metabolic phenotype, contributing to the significant strain-dependent variation in cardiac disease processes, e.g., in coronary and aortic vascular responsiveness to nitric oxide and the development of atherosclerosis. Thus, it is important to acknowledge that the observed phenotype may not necessarily be a direct result of the genetic alteration. By using the potential of “-omic” techniques, one can characterize models of cardiovascular diseases in a systems biology context. Reducionism, the basis of “traditional” scientific methods, is influenced by covariation and needs constantly novel hypotheses to test. In contrast, “-omic” techniques defy common concepts of “proof,” lack simplicity of interpretation, and are technically demanding. However, in combination with metabolomics, transcriptomic and proteomic studies gain a strong functional correlate at the level of a metabolic readout.

Conclusions
Whereas metabolomics has already shown promising results in the area of toxicology, physicians and scientists are just starting to realize the potential of these techniques in cardiovascular medicine. Because metabolite profiles reflect environmental and genetic influences in patients, metabolomics may reveal new links between metabolites and cardiovascular diseases that could be exploited in the early diagnosis, prognostication, and monitoring of patients’ response to drug therapy. Similarly, metabolic profiling may provide a molecular basis for assessing cardiac tissue metabolism at the time of surgery, which could help to judge the vulnerability of the myocardium and to identify patients at risk of intraoperative and postoperative complications. Finally, better insights into cardiac metabolism could point to novel applications of metabolic modulators such as trimetazidine or ranolazine, which are currently undergoing trials for ischemic heart disease and heart failure. Because many companies have identified metabolomics as future areas of growth, we can expect significant improvements in the current workflows that will facilitate the wider use of these technologies by researchers and clinicians.

Acknowledgments
Work in the Dr Mayr’s laboratory is supported by grants from the British Heart Foundation. Dr Mayr would like to acknowledge the contribution of Drs Yuen-Li Chung (Institute of Cancer Research, Sutton, UK) and Christophe Ladroue (University of Warwick, Warwick, UK) for providing figures and Dr Anthony Sullivan for proofreading the manuscript.
Disclosures

None.

References


**KEY WORDS:** metabolism, mass spectrometry, spectroscopy, proteomics
Metabolomics: Ready for the Prime Time?
Manuel Mayr

doi: 10.1161/CIRCGENETICS.108.808329
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2008 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/1/1/58

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Genetics can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Genetics is online at:
http://circgenetics.ahajournals.org//subscriptions/