STATE-OF-THE-ART PAPER

Application of Metabolomics to Cardiovascular Biomarker and Pathway Discovery

Gregory D. Lewis, MD,*†‡ Aarti Asnani, BS,*† Robert E. Gerszten, MD*†‡

Charlestown, Boston, and Cambridge, Massachusetts

Emerging technologies based on mass spectrometry and nuclear magnetic resonance enable the monitoring of hundreds of metabolites from tissues or body fluids, that is, "metabolomics." Because metabolites change rapidly in response to physiologic perturbations, they represent proximal reporters of disease phenotypes. The profiling of low molecular weight biochemicals, including lipids, sugars, nucleotides, organic acids, and amino acids, that serve as substrates and products in metabolic pathways is particularly relevant to cardiovascular diseases. In addition to serving as disease biomarkers, circulating metabolites may participate in previously unanticipated roles as regulatory signals with hormone-like functions. Cellular metabolic pathways are highly conserved among species, facilitating complementary functional studies in model organisms to provide insight into metabolic changes identified in humans. Although metabolic profiling technologies and methods of pattern recognition and data reduction remain under development, the coupling of metabolomics with other functional genomic approaches promises to extend our ability to elucidate biological pathways and discover biomarkers of human disease. (J Am Coll Cardiol 2008;52:117–23) © 2008 by the American College of Cardiology Foundation

Metabolism refers to the body's conversion of food stores into energy currencies that can be used to perform work. Although decades of research in biochemistry, nutrition, and physiology have revealed specific metabolic pathways, systematic surveys of pathways altered in human disease states, such as diabetes, obesity, and cardiovascular disease, have yet to be performed. An emerging set of tools, based on mass spectrometry, nuclear magnetic resonance, and other technologies, enables the monitoring of dozens to hundreds of metabolites from biological samples. Although these technologies are still under development, they complement other functional genomic approaches, such as high-throughput genome sequencing, ribonucleic acid expression analysis, and proteomics, and promise to transform our ability to profile samples with the goal of illuminating biology and discovering valuable clinical biomarkers.

The Birth of Metabolomics

Small biochemicals are the end result of all regulatory complexity present in the cell, tissue, or organism, including

transcriptional regulation, translational regulation, and post-translational modification (Fig. 1). Metabolic changes are thus the most proximal reporters of alterations in the body in response to a disease process or drug therapy. In 1971, Arthur Robinson and Linus Pauling conceived the core idea that information-rich data reflecting the functional status of a complex biological system resides in the quantitative and qualitative pattern of metabolites in body fluids (1). In the same year, Horning and Horning (2) first used the term metabolic profiling to describe the gas chromatography (GC) output from a patient sample. This emerging approach to the quantitative metabolic profiling of large numbers of small molecules in biofluids was ultimately termed "metabonomics" by Nicholson et al. (3), and "metabolomics" by others. Recently, more focused analyses of specific metabolite families or subsets have even given rise to new terms such as "lipidomics." Although the majority of biomarkers have emerged as extensions of "targeted" physiological studies, it has become evident that a metabolite profile derived in an unbiased manner may be informative even if the constituents or their relationships to the disease are initially unknown (Fig. 2).

To date, the majority of metabolomics studies have been performed in model organisms. Studies have elucidated the genetic control of metabolites in plants, such as arabidopsis, and have determined "metabolic footprints" of genetically altered yeast (*S. cerevisiae*) (4,5). In the latter report, metabolic profiling of conditioned media was used to "diagnose" otherwise silent mutant phenotypes. Tandem mass spectrometry (MS) has also been used to profile 36 acylcarnitine species in mice overexpressing hepatic malonyl-coenzyme A

decarboxylase, yielding novel information regarding muscle beta-hydroxybutyrate levels and insulin sensitivity (6).

The vision for metabolic profiling to diagnose human disease, however, extends from seminal studies of inborn errors of metabolism in infants. Millington and colleagues pioneered the use of tandem MS-based methods for monitoring fatty acid oxidation, as well as organic and selected amino acids (7). Their work has culminated in universal neonatal screening for metabolism disorders in the state of North Carolina (8). It is anticipated that a global metabolomic analysis of more common diseases might identify new biomarkers or spotlight pathways for dietary or drug modulation. The application of metabolomics to complex cardiovascular diseases, however, is likely to be more difficult than its application to inherited inborn errors of metabolism.

Technologies to Define the Human Metabolome

The global collection of metabolites in a cell or organism is often called the metabolome; this refers to all small molecules that exclude nucleic acids and proteins (Fig. 1, Table 1). Present estimates suggest that the human metabolome consists of approximately 3,000 endogenous metabolites (Human Metabolome Project [9,10], Kyoto Encyclopedia of Genes and Genomes [11]). As with the human genome, the exact size of the human metabolome remains under debate. Estimates of the metabolome will likely be revised as technologies to detect metabolites become more sensitive and comprehensive. More-

over, some argue that nutritional compounds, xenobiotics modified by human enzymes, as well as microbial metabolites present in the gut must also be taken into consideration when defining the human metabolome.

The metabolome spans a variety of chemical compound classes, including those that are anionic versus cationic and lipophilic versus hydrophilic (Table 1). Metabolites in tissue or body fluids are present across a broad range of concentrations. Therefore, no single analytical method is capable of analyzing all metabolites. However, capturing a subset of "sentinel" metabolites in critical pathways may prove to be a more tractable problem than proteomics. Estimates suggest that post-translational modifications may bring the total number of protein species to $>10^6$, and perhaps 10^8 to 10^9 if immunoglobulins are included. Thus, the metabolome may be less complex than the human proteome. Cellular metabolic pathways are highly conserved across species. Therefore, once metabolic changes are identified in humans, complementary functional studies in model organisms may rapidly provide insight into homeostatic and disease pathways.

Metabolites can be measured by several available analytical methods (for reviews of metabolomics technologies, see articles by Dunn et al. [12,13] and Lindon et al. [14]). Chromatographic procedures such as GC, high-performance liquid chromatography (LC), and capillary electrophoresis have been used to identify and quantify specific metabolite subsets (e.g., amino acids [15,16] or purine metabolites [16,17]) but are best used for initial compound separation in combination with other detection techniques. Recently, 2 high-throughput technologies have garnered the most use for profiling a large number of metabolites simultaneously: nuclear magnetic resonance (NMR) spectroscopy and MS. Mass spectrometry distinguishes metabolites on the basis of mass/charge ratio (m/z)
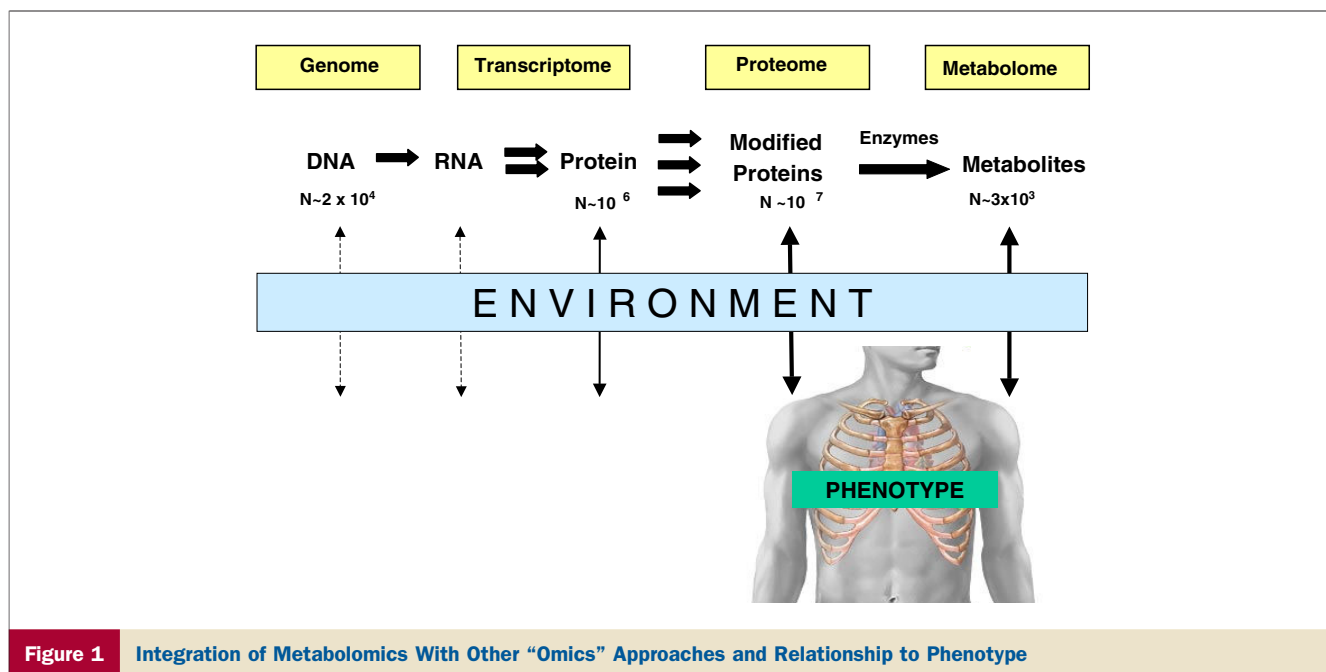

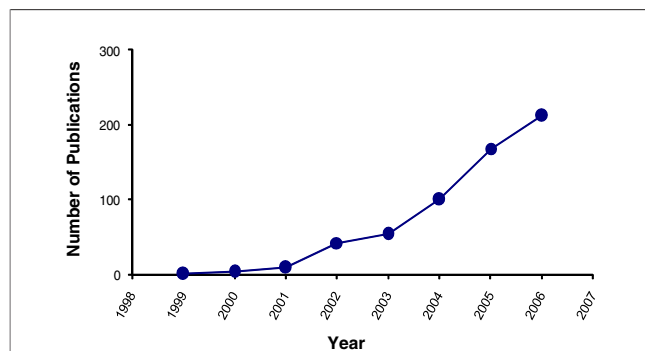**Figure 1**   **Integration of Metabolomics With Other "Omics" Approaches and Relationship to Phenotype**

Figure 2 **Rapid Incorporation of Metabolomics in Biomedical Research**

Criteria for inclusion of publications included use of "metabolome," "metabolomics," or "metabonomics" as a keyword or in the title.

and requires a separation of the metabolite components using either GC after chemical derivitization or LC, with a new method of ultraperformance LC being increasingly used. Mass spectrometry also permits absolute quantification of metabolite levels via the standard addition method, which entails using spiked-in internal standards across a range of concentrations. When available, isotope-labeled standards can be easily differentiated from the endogenous metabolite by the appropriate number of mass units.

Nuclear magnetic resonance spectroscopy uses magnetic properties of nuclei to determine the number and type of chemical entities in a molecule. Proton (^{1}H) NMR spectroscopy can detect soluble proton-containing molecules with a molecular weight of ~20 kD or less. The NMR spectra serve as the raw material for pattern recognition analyses, which simplifies the complex multivariate data into 2 or 3 dimensions that can be readily understood and evaluated. Both NMR and LC-MS systems can be applied to in vivo tissues or to biological fluids, such as serum, plasma, urine, and so on, obtained from humans (18). The advantages of NMR are that it requires relatively little sample preparation, is nondestructive, and can provide information about the precise structure of metabolites (4). However, NMR sensitivity is related to magnet strength, and presently available instrumentation can unambiguously detect only the most abundant metabolites in plasma. However, more sensitive systems are rapidly evolving.

In contrast, the most important advantage of MS coupled with up-front chromatography is far greater sensitivity than NMR. Mass spectrometry-based systems have been used to resolve compounds in the nanomole/liter to picomole/liter and even femtomole/liter range, whereas identification of compounds by ^{1}H-NMR requires concentrations of 1 nmol/l or higher (19). In human plasma, limits of detection between 0.1 and 1 μmol/l for a series of compounds analyzed by GC-MS have been described. Normal plasma concentrations for these metabolites are in the micromole/

liter range, well above the limits of detection established for most MS technologies.

Targeted Versus Pattern Recognition Analyses

Perturbations of the metabolome that arise either as a cause or consequence of disease manifest as particular patterns of

Table 1	Endogenous Compounds Profiled With Current Metabolomics Technologies
Metabolite Class	
Acetylcarnitines	
Acylglycines	
Alcohols	
Aldehydes	
Alkanes/alkenes	
Amides/amines	
Amino acids	
Aromatics	
Bile acids	
Branched-chain hydrocarbons	
Carbohydrates	
Carnitines	
Catecholamines	
Cholesterols	
Coenzyme A derivatives	
Esters	
Ethers	
Fatty acids/alcohols/aldehydes	
Glycerols/glycerophospholipids	
Glycols/glycolipids	
Heterocyclics	
Histidine metabolites	
Hydroxyacids	
Imidazoles	
Isoprenoids	
Ketones	
Lactams	
Lactones	
Lipids	
Minerals	
Nitrogenous compounds	
Nucleosides/nucleotides	
Organic acids	
Peptides	
Phenols	
Phosphates	
Porphyrins	
Prostaglandins	
Pterins	
Purines	
Pyrimidines	
Quinones	
Sphingolipids	
Steroids/sterols	
Sugars	
Sulfates	
Thiols	
Vitamins	

Data from Wishart et al. (**9**).

metabolites in a tissue or body fluid. This patterning concept has been the basis for recent efforts to discover proteomic or metabolomic "signatures" in tissue or serum. Mass spectrometry and NMR techniques can rapidly generate well-defined sets of peaks from a sample across a broad range of mass/charge. A growing controversy is whether such "metabolite signatures" can be used to accurately distinguish disease states from normal. A significant time advantage of direct profile comparisons derives from skipping the far more laborious task of unambiguously identifying the entities that underlie the peaks. Thus, rapid screening of patient samples is possible.

Using a pattern of peaks to diagnose disease without knowing the represented metabolites, however, raises some concerns. One issue is that of reproducibility. Because most MS or NMR instruments were not designed as clinical tools, it is hard to generate consistent results from machine to machine or from operator to operator. Some contend that the patterns are mostly "noise" and do not discriminate biologically meaningful information. Without unequivocal identifications, one cannot independently confirm findings with complementary technologies. Others contend that the peaks profiled by the methodologies used to date only represent the most abundant plasma or tissue constituents. The most important consequence of not unequivocally identifying spectral peaks, however, is that little insight is gained into the biology, either to understand disease pathways through basic cellular mechanisms or as a check on the biological consistency and reasonableness of the data. Overfitting of data is also a common problem when algorithms are generated from hundreds or thousands of peaks. Blinded prospective studies must ultimately be organized to better address the controversy.

Human metabolomics studies are also complicated by potentially confounding clinical variables such as diet or drug effects, particularly if NMR- or MS-based profiling techniques are used in which metabolite peaks are not unambiguously identified. Because of the various limitations inherent to pattern discovery, many have championed metabolomics applications in targeted approaches. The user targets a predefined set of metabolites to be quantified by monitoring specific chromatographic retention times, as well as parent and daughter mass-to-charge ratios of analytes. With the targeted approach being more focused, relying on a predefined set of entities, researchers have more confidence in the end results, because they know what is giving rise to the signals. Although this approach has many advantages, it is blind to changes in metabolites whose retention times and MS characteristics have not been incorporated into the analysis method. As efforts to define the human metabolome grow (20), we anticipate increasingly comprehensive targeted platforms for biomarkers and pathway discovery. Improvements in MS and databases to enable identification of unknown peaks will also be critical.

Statistical Approaches to Metabolomic Data Reduction and Pathway Analysis

Although a high-throughput metabolomics approach to biomarker discovery brings many advantages, it also brings a danger of generating false-positive associations due to multiple testing and overfitting of data, as noted in the preceding text. Application of traditional statistical approaches (e.g., Bonferroni correction) in this setting tends to levy an insurmountable statistical penalty that can obscure biologically relevant associations. Even newer statistical techniques, such as advanced resampling methods or control of the false discovery rate (21), do not adequately address the fundamental problem of how to detect subtle but important changes in multiple variables identified in an "omics" approach.

For metabolites participating in known biological pathways, a bioinformatics approach using pathways analysis can harness the vast information gathered in genomics or metabolomics experiments and turn it into a strength. Specifically, although measurement error in the marker discovery phase often prevents high confidence in any one particular metabolite's correlation, the observation that multiple metabolites in a particular biological pathway are moving in tandem brings confidence that a particular pathway, and therefore any biomarkers in that pathway, truly is correlated with the perturbation. By using a more principled selection process for candidate marker triage, this approach increases the likelihood that candidate biomarkers will be validated in subsequent prospective validation studies. This approach also enhances one's ability to use the metabolomic data collected in the biomarker discovery phase to gain insight into disease biology.

Systematic analysis of functional trends has become widespread and important in the analysis of deoxyribonucleic acid microarray data from model organisms (22). The value of this approach in human studies was illustrated in a recent analysis of high-throughput differential messenger ribonucleic acid (mRNA) expression (23). Expression of mRNA was assessed on over 22,000 genes comparing patients with type 2 diabetes mellitus and unaffected control subjects (patients with normal glucose tolerance). A group of genes with depressed expression in diabetes versus control subjects was identified and tested for association with a collection of other gene characteristics. It was found that this gene set was enriched for genes involved in oxidative phosphorylation. Although individual oxidative phosphorylation genes were not dramatically reduced in expression, as a group the trend was highly significant. Furthermore, the effect was attributable to a subset of oxidative phosphorylation genes regulated by peroxisome proliferator-activated receptor coactivator 1, a regulator of mitochondrial biogenesis. Thus, the analysis of trends among differentially expressed genes led directly to insight into altered metabolism in diabetes patients and hinted at therapeutic hypoth-

eses involving the modulation of oxidative phosphorylation pathways.

There are several statistical issues complicating functional trends analysis of high-throughput data that have been rigorously addressed in software under development, including "FuncAssociate," recently described by Berriz et al. (24). Although the analysis software was developed for use with high-throughput mRNA expression data, the general approach may be used in conjunction with essentially any high-throughput experimental approach for identifying or ranking "interesting genes," FuncAssociate has generally been used in conjunction with controlled-vocabulary functional annotation (e.g., Gene Ontology annotation), but can be used in conjunction with many different sources of gene/protein/metabolite annotation (e.g., expression pattern in other studies, phenotype, protein complex membership, disease association, or phylogenetic profile).

Several data reduction strategies using supervised learning multivariate analysis can be used to construct multivariate metabolite biomarker profiles (25). In supervised learning, an algorithm is used to transform the multivariate data from metabolite profiles into a lower dimensionality with biological interest (e.g., health vs. disease). Metabolite data (inputs) and disease status (outputs or targets) form pairs that are used in the calibration of the model, with the goal of the model being to correctly associate the inputs with the targets. Discriminant analysis is a cluster analysis-based algorithm for categorical variables (26), partial least squares is a popular linear regression-based method, and artificial neural networks offer the advantage of a machine-based method that can learn nonlinear mappings (27).

Principal components analysis applies raw data to a normalized matrix, which is then projected onto a specific scoring schema (4,28). This schema is based upon a series of orthogonal principal components, the first of which describes the maximum variance in the original dataset. The Principal components analysis can be used for dimensionality reduction in a data set by retaining those characteristics of the dataset that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. Similarly, Fisher discriminant analysis creates normalized data matrices that decrease sample variability within a specific condition (e.g., cells grown aerobically) while maximizing sample variability between different conditions (e.g., aerobic vs. anaerobic growth) (29). These relatively unbiased methods of reducing large datasets have the potential to define previously unknown relationships between metabolites in a given physiologic state.

A limitation of current analytical approaches in metabolomics is that they rely on the relative or absolute concentrations of a metabolite in a given tissue or plasma sample and do not take into account varying enzymatic activities controlling metabolic flux through biological pathways. Studies examining metabolic flux incorporate isotope-labeled metabolites to map the fate of enzymatic substrates. Studies of tricarboxylic acid (TCA)-cycle anaplerosis in energy metabolism (30) and fractional synthetic rate of fatty acids and cholesterol (31) have generally focused on small, well characterized pathways in specific tissues such as muscle or liver. The integration of current analytical methods with assessment of enzymatic activities contributing to metabolic flux will help to elucidate biologically relevant metabolic changes.

Novel Roles for Metabolites in Human Physiology

Secreted small ligands such as catecholamines play central roles in cardiovascular physiology. A growing body of literature suggests previously unanticipated roles for metabolites that have been traditionally thought to function exclusively as intracellular signals. He et al. (32) recently discovered that the "orphan" G-protein–coupled receptors GPR91 and GPR99, which are highly expressed in the kidney, bind the TCA-cycle intermediates succinate and α-ketoglutarate, respectively. The working hypothesis is that a local mismatch of energy supply and demand, altered metabolism of TCA-cycle intermediates, or injury leads to mitochondrial dysfunction and the release of succinate and α-ketoglutarate from tissues. Once released into the circulation, the metabolites function in a hormone-like manner, binding their receptors in the renal cortex and triggering the release of renin and activation of the renin-angiotensin system. In the case of tissue ischemia from volume loss, this process might be adaptive to match metabolic demands. In other conditions associated with high succinate production, such as congestive heart failure, resultant increases in blood pressure might prove maladaptive. This recent work highlights roles for new types of circulating metabolites functioning as hormones in the body.

Application of Metabolomics to Unique Human Cardiovascular Disease Models

Novel metabolomics techniques still suffer from signal-to-noise issues, however, and applications to humans may be limited by interindividual variability. Although recent studies have evaluated the diurnal and even seasonal variation of hemostatic and inflammatory proteins (e.g., fibrinogen, D-dimer, and C-reactive protein), systematic studies have yet to be performed for metabolites in humans. Studies to identify novel disease-related pathways are also restricted by the inherent unpredictability of the onset of pathological states. As noted previously, human metabolomics studies are also at high risk for potential clinical confounders, such as diet or drug effects, as well as age, gender, and comorbidities. It has been advocated that the analysis of samples from large patient cohorts, stratified by known risk factors or exposures, may minimize the impact of clinical confounding variables (33). However, the throughput of most current metabolomics technologies, particularly those that are MS based, precludes the analysis of large patient cohorts.

To help circumvent these problems, investigators have begun to apply these emerging technologies to unique clinical scenarios where serial sampling can be performed in patients both before and after a controlled perturbation, thereby allowing each patient to serve as his or her own biological control. Clinical cardiology is uniquely suited for such investigation. As proof of principle, a targeted MS-based metabolomics platform was applied to patients undergoing exercise stress testing with myocardial perfusion imaging (34). Eighteen patients had no evidence of ischemia (control group), whereas 18 patients had inducible ischemia (case group). Plasma was fractionated by high-performance LC and metabolites analyzed using a triple quadrupole mass spectrometer to monitor hundreds of ion pairs by targeted MS. The majority of metabolites displayed concordant changes in cases and controls (i.e., increased in both or decreased in both). For example, lactic acid, as well as hypoxanthine and inosine, end products of adenosine monophosphate catabolism, increased in both case and control groups. However, 6 metabolites yielded highly significant discordant changes in case and control groups. Using a metabolic risk score derived from metabolites with discordant changes in the 2 groups, case subjects could be distinguished from control subjects with a high degree of accuracy (p < 0.0001; c-statistic = 0.95).

Strategies emphasizing the in-depth analysis of small, extremely well phenotyped patient cohorts are ideal in light of current technological limitations. However, such an approach has potential limitations that should be considered. First, although serial sampling in patients serving as their own biological controls helps diminish interindividual variability and signal-to-noise issues, populations studied to date are nevertheless small. Further testing in larger cohorts may be powered to detect more subtle metabolic changes and will provide sufficient precision in the estimates of the utility of each marker to allow for appropriate relative weighting of each component.

Second, although metabolite profiling of serum or plasma offers the advantage of simple sample collection, and may reflect the sum of metabolic changes occurring throughout the body, sampling specific tissues that serve as proximal sources of metabolites (35,36) enables localization of metabolic changes and may help to gauge the sensitivity and specificity of signature metabolic profiles in plasma. Alternatively, for metabolites that are rapidly cleared from the circulation, it may be more appropriate to perform metabolic profiling on urine samples.

Integration of Metabolomics With Other "Omics" Technologies in a Systems Biology Approach to Cardiovascular Disease

The identification of new pathways and biomarkers in cardiovascular disease will depend on the complementary power of genetics, transcriptional profiling, proteomics, and metabolomics. For example, Mayr et al. (36) recently used metabolomics and proteomics to characterize metabolic profiles of atrial tissue that predispose patients to the development of atrial fibrillation. Genome-wide association studies that provide an unbiased scan of genomic sequence variants will also catalyze integrative "omics" approaches. For example, 3 groups recently identified several loci, including chromosome 9p21, associated with early-onset myocardial infarction (37–39). The chromosomal regions identified to date do not contain genes recognizably associated with established coronary heart disease risk factors such as plasma lipoproteins. However, the integration of metabolic and proteomic data from these same patients may provide clues as to how the variants modulate the atherosclerotic process.

Conclusions

An emerging set of analytical and bioinformatics tools have made it possible to profile hundreds of metabolites in complex mixtures such as plasma. Although these technologies are still under development, when coupled with other functional genomic approaches, metabolomics promises to transform our ability to profile samples with the goal of elucidating biological pathways and discovering valuable clinical biomarkers.

Reprint requests and correspondence: Dr. Robert E. Gerszten, Cardiology Division and Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital East-8307, 149 13th Street, Charlestown, Massachusetts 02129. Email: rgerszten@partners.org.

REFERENCES

1. Pauling L, Robinson AB, Teranishi R, Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. Proc Natl Acad Sci U S A 1971;68:2374–6.
2. Horning EC, Horning MG. Metabolic profiles: gas-phase methods for analysis of metabolites. Clin Chem 1971;17:802–9.
3. Nicholson JK, Lindon JC, Holmes E. "Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica 1999;29:1181–9.
4. Raamsdonk LM, Teusink B, Broadhurst D, et al. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. Nat Biotechnol 2001;19:45–50.
5. Allen J, Davey HM, Broadhurst D, et al. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. Nat Biotechnol 2003;21:692–6.
6. An J, Muoio DM, Shiota M, et al. Hepatic expression of malonyl-CoA decarboxylase reverses muscle, liver and whole-animal insulin resistance. Nat Med 2004;10:268–74.
7. Roe CR, Millington DS, Maltby DA. Identification of 3-methylglutarylcarnitine. A new diagnostic metabolite of 3-hydroxy-3-methylglutaryl-coenzyme A lyase deficiency. J Clin Invest 1986;77:1391–4.
8. Frazier DM, Millington DS, McCandless SE, et al. The tandem mass spectrometry newborn screening experience in North Carolina: 1997–2005. J Inherit Metab Dis 2006;29:76–85.
9. Wishart DS, Tzur D, Knox C, et al. HMDB: the Human Metabolome Database. Nucleic Acids Res 2007;35:D521–6.
10. Human Metabolome Project. The Human Metabolome Database. Available at: http://www.hmdb.ca. Accessed March 2008.

11. KEGG: Kyoto Encyclopedia of Genes and Genomes. Available at: http://www.genome.jp/kegg/. Accessed March 2008.

12. Dunn WB, Bailey NJ, Johnson HE. Measuring the metabolome: current analytical technologies. Analyst 2005;130:606–25.

13. Dunn WB, Ellis D. Metabolomics: current analytical platforms and methodologies. Trends Anal Chem 2005;4:285–94.

14. Lindon JC, Holmes E, Nicholson JK. Metabonomics techniques and applications to pharmaceutical research and development. Pharm Res 2006;23:1075–88.

15. Backstrom T, Goiny M, Lockowandt U, Liska J, Franco-Cereceda A. Cardiac outflow of amino acids and purines during myocardial ischemia and reperfusion. J Appl Physiol 2003;94:1122–8.

16. Mei DA, Gross GJ, Nithipatikom K. Simultaneous determination of adenosine, inosine, hypoxanthine, xanthine, and uric acid in microdialysis samples using microbore column high-performance liquid chromatography with a diode array detector. Anal Biochem 1996;238: 34–9.

17. Zemgulis V, Ronquist G, Bjerner T, et al. Energy-related metabolites during and after induced myocardial infarction with special emphasis on the reperfusion injury after extracorporeal circulation. Acta Physiol Scand 2001;171:129–43.

18. Cheng LL, Chang IW, Louis DN, Gonzalez RG. Correlation of high-resolution magic angle spinning proton magnetic resonance spectroscopy with histopathology of intact human brain tumor specimens. Cancer Res 1998;58:1825–32.

19. Zhang X, Wei D, Yap Y, Li L, Guo S, Chen F. Mass spectrometry-based "omics" technologies in cancer diagnostics. Mass Spectrom Rev 2007;26:403–31.

20. Human Metabolome Project. What is metabolomics? Available at: http://www.metabolomics.ca/About/overview.htm. Accessed March 2008.

21. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003;100:9440–5.

22. Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM. Systematic determination of genetic network architecture. Nat Genet 1999;22:281–5.

23. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003;34:267–73.

24. Berriz GF, King OD, Bryant B, Sander C, Roth FP. Characterizing gene sets with FuncAssociate. Bioinformatics 2003;19:2502–4.

25. van der Greef J, Stroobant P, van der Heijden R. The role of analytical sciences in medical systems biology. Curr Opin Chem Biol 2004;8: 559–65.

26. Manly B. Multivariable Statistical Methods: A Primer. London: Chapman & Hall, 1994.

27. Bishop C. Neural Networks for Pattern Recognition. Oxford: Clarendon Press, 1995.

28. Baxter IR, Borevitz JO. Mapping a plant's chemical vocabulary. Nat Genet 2006;38:737–8.

29. Villas-Boas SG, Moxley JF, Akesson M, Stephanopoulos G, Nielsen J. High-throughput metabolic state analysis: the missing link in integrated functional genomics of yeasts. Biochem J 2005;388:669–77.

30. Gibala MJ, Young ME, Taegtmeyer H. Anaplerosis of the citric acid cycle: role in energy metabolism of heart and skeletal muscle. Acta Physiol Scand 2000;168:657–65.

31. Bederman IR, Reszko AE, Kasumov T, et al. Zonation of labeling of lipogenic acetyl-CoA across the liver: implications for studies of lipogenesis by mass isotopomer analysis. J Biol Chem 2004;279: 43207–16.

32. He W, Miao FJ, Lin DC, et al. Citric acid cycle intermediates as ligands for orphan G-protein-coupled receptors. Nature 2004;429: 188–93.

33. Kirschenlohr HL, Griffin JL, Clarke SC, et al. Proton NMR analysis of plasma is a weak predictor of coronary artery disease. Nat Med 2006;12:705–10.

34. Sabatine MS, Liu E, Morrow DA, et al. Metabolomic identification of novel biomarkers of myocardial ischemia. Circulation 2005;112:3868–75.

35. Howarth KR, LeBlanc PJ, Heigenhauser GJ, Gibala MJ. Effect of endurance training on muscle TCA cycle metabolism during exercise in humans. J Appl Physiol 2004;97:579–84.

36. Mayr M, Yusuf S, Weir G, et al. Combined metabolomic and proteomic analysis of human atrial fibrillation. J Am Coll Cardiol 2008;51:585–94.

37. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. N Engl J Med 2007;357:443–53.

38. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. Science 2007; 316:1488–91.

39. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007;316:1491–3.