Chapter 7
Role of Metabolomics in Personalized Medicine

Metabolomics and Metabonomics

The human metabolome is best understood by analogy to the human genome, i.e., where the human genome is the set of all genes in a human being, the human metabolome is the set of all metabolites in a human being. In a systems biology approach, metabolomics provides a functional readout of changes determined by the genetic blueprint, regulation, protein abundance and modification, and environmental influence. Metabolomics is the study of the small molecules, or metabolites, contained in a human cell, tissue, or organ (including fluids) and involved in primary and intermediary metabolism. By definition, the metabolome should exclude enzymes, genetic material and structural molecules such as glycosaminoglycans, and other polymeric units that are degraded to small molecules but do not otherwise participate in metabolic reactions.

A related term, metabonomics is the use of nuclear magnetic resonance (NMR) technology to study metabolomics. According to the Metabolomics Society, "Metabolomics is the study of metabolic changes. It encompasses metabolomics, metabolite target analysis, metabolite profiling, metabolic fingerprinting, metabolic profiling, and metabonomics". Examination of a sample using multiple mass spectrometry-based technologies, nuclear magnetic resonance, integration of the data, and analysis by proprietary software and algorithms enables faster and more accurate understanding of a disease than previously possible. In spite of the broader scope of metabolomics to include metabonomics, the two terms still continue to be used interchangeably.

Researchers at the University of Alberta (Edmonton, Canada), funded by Genome Canada, have completed the first draft of the human metabolome. They categorized 2,500 metabolites, 1,200 drugs, and 3,500 food components, which can be found in the human body. The metabolome has been gathered into the human metabolome database (HMDB), which will enable researchers to find out what metabolites are associated with which diseases, what the normal and abnormal concentrations are, where the metabolites are found or what genes are associated with which metabolites (Wishart et al. 2007). Application of metabolomics to diagnostics, drug research, and nutrition might be integral to improved health and personalized medicine (Hunter 2009).

K.K. Jain, *Textbook of Personalized Medicine*,
DOI 10.1007/978-1-4419-0769-1_7, © Springer Science+Business Media, LLC 2009

Metabolomics Bridges the Gap Between Genotype and Phenotype

In general, the phenotype is not necessarily predicted by the genotype. The gap between the genotype and the phenotype is spanned by many biochemical reactions, each with individual dependencies on various influences, including drugs, nutrition, and environmental factors. In this chain of biomolecules from the genes to the phenotype, metabolites are the quantifiable molecules with the closest link to the phenotype. Many phenotypic and genotypic states, such as a toxic response to a drug or disease prevalence are predicted by differences in the concentrations of functionally relevant metabolites within biological fluids and tissues.

Metabolomics provides the capability to analyze large arrays of metabolites for extracting biochemical information that reflects true functional end-points of overt biological events whereas other functional genomics technologies such as transcriptomics and proteomics merely indicate the potential cause for phenotypic response. Therefore they cannot necessarily predict drug effects, toxicological response, or disease states at the phenotype level unless functional validation is added.

Metabolomics bridges this information gap by depicting, in particular, such functional information because metabolite differences in biological fluids and tissues provide the closest link to the various phenotypic responses. Such changes in the biochemical phenotype are of direct interest to pharmaceutical, biotech, and health industries once appropriate technology allows the cost-efficient mining and integration of this information.

A genome-wide association (GWA) study has been carried out with metabolic traits as phenotypic traits (Gieger et al. 2008). Genetically determined variants in metabolic phenotype (metabotype) have been identified by simultaneous measurements of single nucleotide polymorphism (SNPs) and serum concentrations of endogenous organic compounds in human population. Four of these polymorphisms are located in genes. Individuals with polymorphisms in genes coding for well-characterized enzymes of the lipid metabolism have significantly different metabolic capacities with respect to the synthesis of some polyunsaturated fatty acids, the beta-oxidation of short- and medium-chain fatty acids, and the breakdown of triglycerides. Thus, the concept of "genetically determined metabotype" as an intermediate phenotype provides a measurable quantity in the framework of GWA studies with metabolomics and might help to better understand the pathogenesis of common diseases and gene-environment interactions.

The use of this approach to screen previous GWA studies to look for associations between the SNPs of interest and clinical measurements influencing cardiovascular disease, revealed overlap between several SNPs that seem to affect both metabolite biochemistry and clinical outcomes. These metabotypes, in interactions with environmental factors such as nutrition and lifestyle, may influence the susceptibility of an individual for certain phenotypes. For example, there are potential links between long-chain fatty acid metabolism and attention deficit hyperactivity syndrome. Understanding these connections, in turn, may eventually lead to more targeted

nutrition or therapies and more refined disease risk stratification. These could result in a step towards personalized health care and nutrition based on a combination of genotyping and metabolic characterization.

Metabolomics, Biomarkers and Personalized Medicine

Metabolomics has been used to identify biomarkers for disease and to identify off-target side effects in marketed drugs and new chemical entities in development. Compared to 25,000 genes and approximately a million proteins, there are only 2,500 metabolites (small molecules). Their limited number enables an easier, more quantitative method of analysis. Examination of a sample using multiple mass spectrometry-based technologies, integration of the data and analysis by proprietary software and algorithms enables faster and more accurate understanding of a disease than previously possible. Plasma samples obtained from patients can be analyzed for signatures of neurodegenerative disorders by measuring the spectrum of biochemical changes and mapping these changes to metabolic pathways. This technology can be applied to discover biomarkers for diabetic nephropathy in type 1 diabetes. It is hoped that metabolomic profiling would be included in personalized medicine.

Metabolomic Technologies

Within the last few years, metabolomics has developed into a technology that complements proteomics and transcriptomics. In combination with techniques for functional analysis of genes, it is hoped that a holistic picture of metabolism can be formed. In addition to the genome analysis and proteome analyses, the exhaustive analysis of metabolites is important for a comprehensive understanding of cellular functions because the dynamic behavior of metabolites cannot be predicted without information regarding the metabolome.

In view of the chemical and physical diversity of small biological molecules, the challenge remains of developing protocols to gather the whole ‘metabolome’. No single technique is suitable for the analysis of different types of molecules, which is why a mixture of techniques has to be used. In the field of metabolomics, the general estimations of the size and the dynamic range of a species-specific metabolome are at a preliminary stage. Metabolic fingerprinting and metabonomics with high sample throughput but decreased dynamic range and the deconvolution of individual components achieve a global view of the in vivo dynamics of metabolic networks. The technologies used include NMR, direct infusion mass spectrometry, and/or infrared spectroscopy. Gas chromatography (GC)–MS and liquid chromatography-mass spectrometry LC-MS technology achieve a lower sample throughput but provide unassailable identification and quantification of individual compounds in a

complex samples. Major steps forward in these technologies have made it possible to match specific demands with specific instruments and novel developments in the performance of mass analyzers.

However, it is important to note that each type of technology exhibits a bias towards certain compound classes, mostly due to ionization techniques, chromatography and detector capabilities. GC-MS has evolved as an imperative technology for metabolomics because of its comprehensiveness and sensitivity. The coupling of GC to time-of-flight (TOF) mass analyzers is an emerging technology. High scan rates provide accurate peak deconvolution of complex samples. GC-TOF-MS capabilities provide an improvement over conventional GC-MS analysis in the analysis of ultracomplex samples, which is particularly important for the metabolomics approach. Ultracomplex samples contain hundreds of co-eluting compounds that vary in abundance by several orders of magnitude. Thus, accurate mass spectral deconvolution and a broad linear dynamic range represent indispensable prerequisites for high quality spectra and peak shapes. Modern GC-TOF-MS applications and incorporated mass spectral deconvolution algorithms fulfill these requirements.

The advantages of metabolomics technologies are:

- Ability to analyze all bodily fluids such as blood, CSF, and urine as well as cultured or isolated cells and biopsy material
- High throughput capability enabling simultaneous monitoring of biological samples
- Ability to analyze multiple pathways and arrays of metabolites simultaneously from microliter sample quantities

Urinary Profiling by Capillary Electrophoresis

Metabolomic approaches have become particularly important for the discovery of biomarkers in urine. The analytical technology for urine profiling must be efficient, sensitive, and offer high resolution. Until recently these demands were commonly met by HPLC-MS, GC-MS and NMR. The analytical armory for urine profiling has now been extended to include cyclodextrin-modified micellar electrokinetic capillary chromatography (CD-MECC), which enables highly cost-effective, rapid, and efficient profiling with minimal sample volume and preparation requirements. The CD-MECC profiles typically show separation for over 80 urinary metabolites. These profiles have been visualized using novel advanced pattern recognition tools. Visualization of pattern changes has been achieved through development of the novel Automated Comparison of Electropherograms (ACE) software which not only removes errors due to baseline shifts but also allows for rapid reporting of semiquantitative profile differences. The method has been applied in the investigation of biomarkers characteristic of alcoholics or Down's syndrome persons.

Lipid Profiling

Modern medicine has come to rely on a small suite of single biomarkers, such as plasma cholesterol or triglycerides, to assess the risk of certain diseases. However, such single-biomarker assessments overlook the inherent complexity of metabolic disorders involving hundreds of biochemical processes. Assessing the full breadth of lipid metabolism is what drives the field of lipomic profiling. However, unlike the other "-omic" technologies, in which only a small portion of the genes or proteins is known, lipid metabolic pathways are well characterized. Another limitation of "-omics" technologies is that they produce so many false positive results that it is difficult to be sure that the findings are valid. Metabolomics is not immune to this problem but, when practiced effectively, the technology can reliably produce information to aid in decision making. Focused metabolomics platforms, which restrict their target analytes to those measured well by the technology, can produce data with properties that maximize sensitivity and minimize the false discovery problem. The most developed focused metabolomics area is lipid profiling. TrueMass® (Lipomic Technologies) analysis produces lipomic profiles – comprehensive and quantitative lipid metabolite profiles of biological samples. With TrueMass, Lipomics measures hundreds of lipid metabolites from each small quantity of tissue, plasma, or serum sample. Because the resulting data are quantitative, TrueMass data can be seamlessly integrated with pre-existing or future databases.

Data-dependent acquisition of MS/MS spectra from lipid precursors enables us to emulate the simultaneous acquisition of an unlimited number of precursors and neutral loss scans in a single analysis (Schwudke et al. 2006). This approach takes full advantage of rich fragment patterns in tandem mass spectra of lipids and enables their profiling by complex scans, in which masses of several fragment ions are considered within a single logical framework. No separation of lipids is required, and the accuracy of identification and quantification is not compromised, compared to conventional precursor and neutral loss scanning.

Role of Metabolomics in Biomarker Identification and Pattern Recognition

Metabolomics research has increased significantly over recent years owing to advances in analytical measurement technology and the advances in pattern recognition software enabling one to visualize changes in levels of hundreds or even thousands of chemicals simultaneously. Multivariate metabolomic and proteomic data and time-series measurements can be combined to reveal protein-metabolite correlations (Weckwerth and Morgenthal 2005). Different methods of multivariate statistical analysis can be explored for the interpretation of these data. The discrimination of the samples enables the identification of novel components. These components are interpretable as inherent biological characteristics.

Biomarkers that are responsible for these different biological characteristics can easily be classified because of the optimized separation using independent components analysis and an integrated metabolite-protein dataset. Evidently, this kind of analysis depends strongly on the comprehensiveness and accuracy of the profiling method, in this case metabolite and protein detection. Assuming that the techniques will improve, more proteins and metabolites can be identified and accurately quantified; the integrated analysis will have great promise.

Validation of Biomarkers in Large-Scale Human Metabolomics Studies

A strategy for data processing and biomarker validation has been described in a large metabolomics study that was performed on 600 plasma samples taken at four time points before and after a single intake of a high fat test meal by obese and lean subjects (Bijlsma et al. 2006). All samples were analyzed by a LC-MS lipidomic method for metabolic profiling. Such metabolomics studies require a careful analytical and statistical protocol. A method combining several well-established statistical methods was developed for processing this large data set in order to detect small differences in metabolic profiles in combination with a large biological variation. The strategy included data preprocessing, data analysis, and validation of statistical models. After several data preprocessing steps, partial least-squares discriminate analysis (PLS-DA) was used for finding biomarkers. To validate the found biomarkers statistically, the PLS-DA models were validated by means of a permutation test, biomarker models, and noninformative models. Univariate plots of potential biomarkers were used to obtain insight in up- or down-regulation.

Pharmacometabonomics

A major factor underlying interindividual variation in drug effects is variation in metabolic phenotype, which is influenced not only by genotype but also by environmental factors such as nutritional status, the gut microbiota, age, disease and the co- or pre-administration of other drugs. Thus, although genetic variation is clearly important, it seems unlikely that personalized drug therapy will be enabled for a wide range of major diseases using genomic knowledge alone. Metabolite patterns that are characteristic of the individual can be used to diagnose diseases, predict an individual's future illnesses, and their responses to treatments.

A 'pharmacometabonomic' approach to personalizing drug treatment, developed by scientists at the Imperial College London in collaboration with Pfizer, uses a combination of pre-dose metabolite profiling and chemometrics to model and predict the responses of individual subjects (Clayton et al. 2006). A proof-of-principle

for this new approach, which is sensitive to both genetic and environmental influences, is provided with a study of paracetamol (acetaminophen) administered to rats. Predose prediction of an aspect of the urinary drug metabolite profile and an association between predose urinary composition and the extent of liver damage sustained after paracetamol administration was shown. The new approach, if successful, requires the analysis of the metabolite profiles of an individual from a urine, or other biofluid, sample. This new technique is potentially of great importance for the future of healthcare and the pharmaceutical industry and for the development of personalized medicine. The new method is expected to be synergistic with existing pharmacogenomic approaches. Pharmacometabonomics is in the early stage of development and will be studied in humans to evaluate its possible clinical application. Pharmacometabonomics could be used to preselect volunteers at key stages of the clinical drug development process. This would enable stratification of subjects into cohorts, which could minimize the risk of adverse events, or focus on those individuals with a characteristic disease phenotype for assessment of efficacy (Haselden and Nicholls 2006).

Metabonomic Technologies for Toxicology Studies

Metabonomics studies demonstrate its potential impact in the drug discovery process by enabling the incorporation of safety endpoints much earlier in the drug discovery process, reducing the likelihood (and cost) of later stage attrition.

Global metabolic profiling (metabonomics/metabolomics) has shown particular promise in the area of toxicology and drug development. A metabolic profile need not be a comprehensive survey of composition, nor need it be completely resolved and assigned, although these are all desirable attributes. For the profile to be useful across a range of problems, however, it must be amenable to quantitative interpretation and it should be relatively unbiased in its scope. In addition to explicit quantification of individual metabolites, analytical profiles such as NMR spectra are effectively functions of the concentrations of the constituents of the sample and hence can be handled directly as metabolic profiles. A further requirement for the platform used to generate profiles is that the analytical variation introduced after collection be less than the typical variation in the normal population of interest, so as not to reduce significantly the opportunity to detect treatment/group-related differences. Fulfilling this condition is very dependent on the actual system and question in hand and is probably best tested in each new application.

In both preclinical screening and mechanistic exploration, metabolic profiling can offer rapid, noninvasive toxicological information that is robust and reproducible, with little or no added technical resources to existing studies in drug metabolism and toxicity. Extended into the assessment of efficacy and toxicity in the clinic, metabonomics may prove crucial in making personalized therapy and pharmacogenomics a reality.

Metabonomics/Metabolomics and Personalized Nutrition

It is possible to profile metabolic diseases before symptoms appear. Metabonomic testing is important in obesity/metabolic syndromes, in which several metabolic pathways interact to produce symptoms and could be an important guide to select diets and exercise programs tailored to metabolic states.

It is considered desirable to establish a human "metabonome" parallel to the human genome and proteome but it will be a formidable undertaking requiring analysis of at least half a million people. Some projects are examining metabonomic patterns in a series of patients with metabolic syndromes and comparing them with normal people. Other studies are examining how a person's unique metabonomic profile can be used as a guide to personalize diet and exercise regimens for obesity.

It is now possible to measure hundreds or thousands of metabolites in small samples of biological fluids or tissues. This makes it possible to assess the metabolic component of nutritional phenotypes and will enable individualized dietary recommendations. The relation between diet and metabolomic profiles as well as between those profiles and health and disease needs to be established. The American Society for Nutritional Sciences (ASNS) should take action to ensure that appropriate technologies are developed and that metabolic databases are constructed with the right inputs and organization. ASNS also should consider the social implications of these advances and plan for their appropriate utilization.

Summary

Whereas the human genome is the set of all genes in a human being, the human metabolome is the set of all metabolites in a human being. Metabolomics bridges the gap between the genotype and the phenotype and is an important basis of personalized medicine. Metabolomics has been used to identify biomarkers for disease and to identify the effects of drugs. Various metabolomic technologies include nuclear magnetic resonance, GC, and mass spectrometry. Pharmacometabonomic approach to personalizing drug treatment uses a combination of pre-dose metabolite profiling and chemometrics to model and predict the responses of individual subjects. Metabolomics/metabonomics also have a role to play in assessing drug toxicity and in guiding nutrition.