Published in final edited form as:
Crit Rev Biomed Eng. 2013 ; 41(3): 205–221.

Metabolomic Fingerprinting: Challenges and Opportunities

Alyssa K. Kosmides[1], **Kubra Kamisoglu**[2], **Steve E. Calvano**[3], **Siobhan A. Corbett**[3], and **Ioannis P. Androulakis**[1,2,3,*]

[1]Department of Biomedical Engineering, Rutgers, The State University of New Jersey, Piscataway, NJ

[2]Department of Chemical and Biochemical Engineering, Rutgers, The State University of New Jersey, Piscataway, NJ

[3]Department of Surgery, Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, New Brunswick, NJ

Abstract

Systems biology has primarily focused on studying genomics, transcriptomics, and proteomics and their dynamic interactions. These, however, represent only the potential for a biological outcome since the ultimate phenotype at the level of the eventually produced metabolites is not taken into consideration. The emerging field of metabolomics provides complementary guidance toward an integrated approach to this problem: It allows global profiling of the metabolites of a cell, tissue, or host and presents information on the actual end points of a response. A wide range of data collection methods are currently used and allow the extraction of global or tissue-specific metabolic profiles. The great amount and complexity of data that are collected require multivariate analysis techniques, but the increasing amount of work in this field has made easy-to-use analysis programs readily available. Metabolomics has already shown great potential in drug toxicity studies, disease modeling, and diagnostics and may be integrated with genomic and proteomic data in the future to provide in-depth understanding of systems, pathways, and their functionally dynamic interactions. In this review we discuss the current state of the art of metabolomics, its applications, and future potential.

Keywords

metabolomics; bioinformatics; data analysis; disease modeling

I. METABOLOMICS–METABONOMICS: BASIC PRINCIPLES AND CONCEPTS

Understanding the mechanisms by which biological systems respond to physiological and pathophysiological stimuli is of great scientific and clinical interest. Systems biology studies interactions between interconnected networks that involve changes at the genomic,

*Address all correspondence to: Ioannis P. Androulakis, Biomedical Engineering Department, Rutgers University, 599 Taylor Road, Piscataway, NJ 08854; Tel.: (732) 445-4500 ext. 6212; Fax: (732) 445-3753; yannis@rci.rutgers.edu.

proteomic, and metabolomic levels under homeostatic conditions and in response to stimuli. Biomedical scientists are ultimately interested in determining links between experimental conditions and clinically relevant phenotypic changes, motivating experimental techniques that can quantify high-level responses to complement transcriptome-level information that can be difficult to link to biological function. Evaluating changes at a higher organizational level allows for a primarily data-driven approach because of the ability to quickly collect and analyze data on the state of an organism. It provides novel information on phenotypic characteristics and therefore the potential to investigate the output of complex, interconnected networks. This review focuses on monitoring changes at the metabolite level by discussing techniques, potential applications, opportunities, and challenges in studying the dynamics of a host response.

The terms *metabolomics* and *metabonomics* both refer to studying metabolites of a biosample. Metabolomics, or metabolic phenotyping, is the study of the quantitative description of all low-molecular-weight (<1 kDa) components in a biological sample. [1,2] These may consist of metabolites solely under endogenous control and may also involve those originating from exogenous sources (microbiome,[3] diet,[4] drugs,[5,6] etc.). Metabonomics is the study of the interactions of these metabolites over time in a complex system.[1,2] Although metabonomics is in fact a subcategory of metabolomics, many use these terms interchangeably.

Metabolomics is a unique top-down approach for studying complex systems.[7] Rather than attempting to decipher the interactions among transcriptional and translational level data, metabolomics studies the end result. Each type of cell and tissue has a characteristic metabolic composition that is uniquely altered in response to physiological and pathophysiological stimuli. These phenotypes reflect the collective effects of epigenetic factors, heterogeneous distributions of molecules, and differential reaction rates. The metabolome of a sample, that is, the concentrations of these metabolites at a given time, can be thought of as a metabolic “fingerprint” representative of the state of the organism at that time.[8] Metabolomics involves the quantification of these metabolites, often in a temporal manner, to track the developing response to a stimulus.

Genomics and proteomics are the -omics techniques most widely used to study the effects of stimuli on integrated systems and individual pathways. Genomics and transcriptomics are used to study the genome and gene expression levels of an organism, respectively. The data obtained by these methods is at the lowest level of organization, shedding light on the origin of specific phenotypes. Proteomics quantifies the abundance of proteins, elucidating the next level up from gene expression data. The integration of these fields provides a unified picture of cellular-level responses from transcription through translation.

While genomic studies have linked genetic factors to disease predisposition[9,10] and proteomic studies have identified proteins that enable monitoring disease progression,[11] both proteomics and genomics are limited in their potential applications.[2] Genetics alone cannot fully explain differences in disease predisposition.[1] Only about 5–10% of the total human genetic variance occurs across populations and ethnic groups, although disease distributions and drug toxicity may vary greatly. Broadly speaking, genomics does not

account for differences in phenotype.[2] Although a gene may be expressed and a protein may be synthesized, this protein may not be in the proper form to induce a metabolic change and therefore induce a phenotypic effect. Xenobiotics and other environmental factors may also cause gene expression to provoke differential phenotypic responses.[5] In addition to the lack of information on phenotypic changes using genomics and proteomics, true end point results are difficult to unravel from this low-level data alone. Gene expression and protein concentrations vary within each organ and tissue, and organs and tissues interact through complex signaling pathways.[3] Variables must be measured frequently to achieve accurate results based on gene-protein coupling, which is time-consuming and expensive.[12] Infrequent genomic and proteomic measurements, even taken together, may lead to incorrect conclusions about the relationship between a specific gene and its complementary protein because of the time lag between protein synthesis and gene transcription.[3]

The microbiome, an assemblage of more than 100 trillion microorganisms belonging to 300–500 different species that live inside of every human being[13] also represents an obstacle in the context of using genomics or proteomics to study responses.[1,3,7,14] Each individual's microbiome is as unique as a fingerprint, [14] with each person sharing as little as 1% of the same species.[15] The species contained in a single human changes with diet, drugs, age, disease, and medical or surgical intervention.[14] The gut microbiome interacts with the endogenous systems in the human body and has metabolic, trophic, and protective functions.[13,16] It influences levels of cytochrome P450, an important group of enzymes that metabolize both endogenous and exogenous materials,[3] and plays a large role in obesity in humans[14,17–19] and rodents, [19,20] sepsis,[21,22] inflammatory bowel disease, irritable bowel syndrome, and colon cancer.[13] The gut microbiome contributes to heterogeneous responses to drugs and interindividual variability in drug toxicities and may contribute to the carcinogenicity of certain compounds by metabolizing substances that otherwise would not be broken down.[23] Metabolites that originate from the gut microbiome merge with endogenous metabolites, directly altering the metabolome without necessarily influencing gene and protein expression.

Metabolomics complements more traditional -omics techniques by allowing the investigation of properties that cannot be directly assessed through proteomics and genomics. Epigenetics, or the study of heritable changes in DNA expression that are not explained by the underlying DNA sequence,[24] can be studied at the end point level. Heritable epigenetic factors predispose to some types of cancer, autoimmune diseases, mental disorders, and diabetes, all of which can be passed on through generations.[25,26] Although not used in epigenomics to date, metabolomics has the potential to play an important role in the emerging field of epigenetic therapy,[25] which has already shown promising results in curing some types of leukemia and anemia.[27]

Integration of metabolomics with genomics and proteomics is also possible and can help make the relationship between the levels of information produced by each technique more clear. Using data from one level to predict function at another is not always accurate. By obtaining data at multiple levels through the application of multiple -omics techniques, the interactions between the genome, proteome, and metabolome can be further studied and

understood. Changes in gene expression levels and protein concentrations can then be linked to physiological changes and interpreted in the biological context.

In addition to offering a different level of information about biological processes that cannot be obtained through other -omics approaches, metabolomics simplifies certain aspects of the analysis. It presents several benefits over lower-level approaches that uncover only the potential for a phenotypic result.[28] Because it provides data on the phenotypic response, it eliminates the need for any assumptions to be made about the origin of a phenotype, although the individual contributions can still be drawn out.[7] Xenobiotic and protein kinetics do not have to be predicted, although kinetic data can be reverse engineered from metabolomic data. Low experimental variability results in the generated data being generally reproducible and transferable[29]—2 factors that are of extreme importance when studying biological systems. Experimental[30] and analytical[30,31] variations of commonly used methods are several orders of magnitudes smaller than biological variation and therefore do not undermine the applications of metabolomics in modeling and classification.

It is often difficult or even impossible to pursue certain studies on humans because of ethical reasons, even though applications in humans are typically the ultimate goals of biomedical researchers. Although metabolic rates may differ among mammalian species, metabolic pathways are highly conserved.[29] This enables the information collected about common laboratory animals, such as rats and mice, to provide useful information that can be relatively easily related to humans. Metabolomics also will allow for a greater understanding of the role and mode of action of the gut microbiome[18,19] since the samples of the microbiome are easily attainable. There is great potential for determining cause-and-effect relationships between microbiome and metabolite profiles considering the critical role that the microbiome plays in the development of many pathologies mentioned earlier. Urine and blood—the 2 biofluids most commonly used in metabolomics studies—are relatively cheap and easy to collect, enabling time series measurements to facilitate the study of temporal changes.[7] Metabolomics may therefore aid in directing the use and timing of more complex procedures to maximize efficiency and minimize the collection of insignificant data.[29] Because certain combinations and ratios of metabolites are specific to individual conditions, metabolic profiles of urine can also indicate the region of an organ and mode of toxicity of a response through a biomarker-like approach. By observing changes in metabolic phenotype after exposure to a toxin that causes a specific response, biomarkers indicative of the toxicity can be uncovered and used as an indicator in future studies.[29]

II. DATA SOURCES, COLLECTION, AND ANALYTICAL METHODS

A. Sample Sources

Because each type of cell and tissue has its own unique metabolic fingerprint, one must choose what cell or tissue type to study or measure biofluids that represent the combined output of interactions between multiple organs.[32] Urine and plasma are the most commonly used biofluids in metabolomics[33] because they are reasonably easy to obtain[7,34] and are collected relatively noninvasively,[2,35] which enables high-frequency sampling even in critically ill patients. They are always at dynamic equilibrium with the body, rapidly reflecting metabolic changes within the host.[12] It must be kept in mind, however, that while

plasma represents a snapshot of the state of an organism, urine provides a time-averaged representation.[30] Less common samples for study include cerebrospinal fluid, saliva, and semen,[33] as well as various tissue samples.[36] Tissue samples, unlike biofluids, can be used to quantify organ-specific metabolic fingerprints,[32] making it possible to study the origins of metabolites.

B. Data Generation

1. Nuclear Magnetic Resonance Spectroscopy—Nuclear magnetic resonance (NMR) spectroscopy is the most common technique used to generate metabolomic data from biofluids.[37] While it is less sensitive than some other techniques described later, it offers several advantages. NMR spectroscopy is high-throughput, taking only a few minutes per sample,[38] has relatively low per-sample cost, and requires no a priori knowledge[34,39] of what metabolites to study[7] since it outputs a superposition of the spectra of all detectable metabolites. NMR spectroscopy also includes a large range of metabolites per measurement and provides information on the chemical structure, chemical environment, dynamic molecular motions, and molecular interactions between metabolites.[38] NMR spectroscopy is nondestructive, so it can be applied to samples before they are subjected to further destructive analysis.[7] Downsides to NMR spectroscopy include low spectral resolution and sensitivity, although both of these can be mitigated by applying stronger magnetic fields.[39]

NMR spectroscopy can be performed on any spin-active nuclei,[38] although 1H and 13C are most commonly used. 1H allows high sensitivity since it is ubiquitous in organic materials and has a natural abundance of 99.98%, but it has a smaller chemical shift range than 13C, resulting in greater peak overlaps that make the data more difficult to analyze.[38] 13C benefits from a chemical shift range that is about 20 times greater than that of 1H40 and therefore provides much greater spectral resolution. Although 13C has a much lower sensitivity because of its low natural abundance, it has been demonstrated that using a cryogenic probe can drastically improve such results.[40] By using a combination of these different NMR spectroscopy methods, sensitivity and resolution can be optimized to increase the accuracy and integrity of generated metabolomic data.

2. Mass Spectrometry—Mass spectrometry (MS)–coupled techniques,[38] including liquid chromatography MS (LC-MS),[41] gas chromatography MS (GC-MS),[42] and high-performance liquid chromatography (HPLC-MS),[20,41] have proven to be very useful in generating metabolomic data. Metabolites are typically separated from the biological fluid[7] before MS, which causes these methods to be slower and more complex than NMR spectroscopy. However, they provide a much higher sensitivity[38] and therefore enable the quantitative measurement of a broader spectrum of metabolites.

Compared to other MS-coupled techniques, GC-MS has the highest resolving power, and 2-dimensional GC-MS can further increase resolution.[42] This method is selective, however, and can be used to analyze only certain substances. Several classes of compounds (sugars, nucleosides, amino acids, etc.) cannot be analyzed directly because of their polarity and lack of volatility.[38] Although LC-MS does not have the high resolving power as GC-MS,[38] it is popular because of the minimal sample preparation[41] and small sample size.[43] High-

performance liquid chromatography also has shown great advancements in recent years, especially in combination with 1H NMR spectroscopy, since the 2 techniques identify complimentary sets of metabolites.[20] Other MS methods also have been used to measure metabolomic data from biofluids, including microbore LC-MS, ultra performance LC, and capillary electrophoresis–MS (CE-MS) for metabonomics[38], although to date these techniques have not been widely applied.

3. Magic-Angle Spinning NMR Spectroscopy—While the above techniques are used predominantly on biofluids, magic-angle spinning (MAS)–NMR spectroscopy has been extensively used for metabolic phenotyping of tissue samples. MAS-NMR enables the acquisition of tissue-specific metabolic phenotypes[38] in contrast to the integrated metabotypes obtained by biofluid samples. It can be applied after the biofluid analysis to confirm the origin of certain biomarkers,[29] which can play an important role in drug toxicity studies. MAS-NMR spectroscopy can also provide valuable and unique information on the compartmentalization of metabolites *in vivo*,[38] as some pathologies are characterized by a redistribution of metabolites rather than a change in their concentration. A modified MAS 1H NMR method that makes use of changes in the apparent diffusion coefficient of metabolites is capable of detecting changes in the cellular environment.[44,45] Like other NMR spectroscopy techniques, it requires little sample preparation, is not destructive, and requires only small amount of samples (~20 mg).[46]

III. DATA ANALYSIS METHODS AND SOFTWARE

One of the main challenges in metabolomics is that the large volume of data produced requires the use of complex multivariate analysis techniques. Although none of the currently available data collection methods can capture the quantitative and qualitative information on all metabolites in a given sample,[41,43] all methods still generate and process immense amounts of information. Data collected by 1H NMR spectroscopy contains information on up to 100 metabolites in urine and up to 30 in plasma and tissue extracts,[47] whereas data collected by MS-coupled techniques can contain information on more than 1000 metabolites per sample.[42] Extracting meaningful data for biological interpretation from this vast amount of data requires the use of robust computational techniques, which are being developed for and widely used in other application areas of systems biology.

A. Computational Data Analysis

1. Metabolite Identification—NMR spectroscopy creates data in the form of spectra consisting of multiple peaks that are the superposition of the spectra of all detected metabolites. Each of these peaks, or combinations of peaks, corresponds to a unique compound. Compound-specific peak combinations can be determined by conducting NMR spectroscopy on the metabolite of interest in water alone or simply referencing the vast amounts of published data in the literature.[48–51] Statistical total correlation spectroscopy can be used to aid in identifying molecules in NMR spectra by recognizing highly correlated peak intensities, leading to detection of all of the peaks of a certain molecule. This information on highly positively and negatively correlated peaks can also support the identification of molecules in the same pathway.[52]

Although peaks that correspond to a specific molecule are always in nearly the same position on the spectra, pH differences cause these peaks to shift slightly,[34,38] despite the use of buffers.[34] This can be addressed by integrating the spectra over small chemical shift windows (~0.04 ppm)[53,54] or by using peak alignment algorithms[55] before analysis to reduce the probability of incorrectly characterizing metabolites. However, sometimes the use of these methods can be disadvantageous. Using raw data, and therefore the information on magnitude and direction of chemical shifts, could improve the accuracy of models and add to the understanding of physicochemical variations in metabonomic data sets.[52,56] Preprocessing of MS datasets consists of similar steps. Peak alignment and deconvolution techniques often are used to adjust for variations due to temperature, column variability,[41] instrument parameters, and other sources.[42]

2. Unsupervised Methods: Principal Component Analysis—In most metabonomic studies, including the majority of disease models and drug toxicity studies, the metabolites of interest are not known a priori. Thus, unsupervised methods—those in which no prior knowledge of class membership is assumed—are a common first step in data analysis. The most common unsupervised method used in metabonomics is principal component analysis (PCA), an algorithm that reduces a high-dimensional data set to a small number of dimensions that explain as much of the variation in the data as possible. Each principal component is a linear combination of the original variables. The principal components (PCs) are orthogonal (uncorrelated), and the first few PCs contain the largest portion of the variation, with each subsequent PC containing correspondingly smaller amounts.[12]

Once the PCs have been found, each sample can be plotted on a PCA map to give a visual representation of the results. Because the PCs encompass most of the variability in the data, the data points typically appear clustered in the PCA map, with each cluster being representative of a different metabolic fingerprint. These metabolic fingerprints then are used to find potential biomarkers—metabolites that vary most between classes. To better visualize the interacting effects of macroscopic factors, influence vectors can be calculated and placed in the multidimensional PCA map. Influence vectors display the general magnitude and direction of the effects that a certain influence factor (e.g., age, sex) has on the principal components so that each point can be understood as a result of the complex interaction between these factors. For example, an influence vector representing the effect of age may point in the direction of general metabolic trends during increasing age, so that a point that lies further along the vector exhibits a profile indicative of older age. A challenge exists, however, in calculating the simplest vectors without compromising accuracy, as the relationships are often highly nonlinear.[5] This would provide an understanding of what factors cause the differences between fingerprints and how certain metabolites are affected. To determine what metabolites differ most between classes, a loadings plot can be used to help interpret the PCA map. A loadings plot is unique for each PC, where the loadings (eigenvector components) are plotted against each of the original variables. These plots give a graphical representation of what spectral regions (metabolites) contribute most to each PC, thus showing the metabolites that differ most between each class.[46]

While PCA is very good at detecting clusters and outliers, its results often are used only in directing future analysis since its accuracy can be improved by using supervised methods.[8]

Although PCs describe the largest portion of the variation, this variation may not closely correspond to the separation between classes[57]; therefore, PCA usually is used only as a first step to aid in the development of a model using better-classifying supervised methods.

3. Supervised Methods—Once potential biomarkers have been found, supervised methods can be used to maximize the separation between classes and identify the most robust biomarkers.[58] Because supervised methods use information on class membership, they are much better at developing classifiers and predicting where a sample falls with respect to those already classified. [28] Commonly used supervised methods include partial least squares with discriminate analysis (PLS-DA),[58] orthogonal PLS-DA,[30] and orthogonal partial least squares.[14,52]

PLS-DA uses both the descriptor matrix (e.g., class) and the results (e.g., spectra) to define a surface in *n*-dimensional space that separates data into classes.[7] However, in the presence of noise, PLS-DA models can be less accurate and difficult to interpret. Because noise is so common in biological data sets, especially those involving humans, methods are needed to filter out noisy and unrelated variation that has no correlation with class identification.[59] Orthogonal signal correction (OSC) is commonly used to improve the integrity of data sets and has been shown to increase model accuracy in many cases,[47,58–60] OSC reduces noise by filtering out the variation in the descriptor matrix X that is orthogonal to the variation in the results matrix Y.[57] In a similar way, orthogonal PLS-DA is an extension of PLS-DA that includes an integrated OSC filter[57] and models the variation common to X and Y separate from the uncorrelated variation. This not only leads to models that are easier to interpret (since the noise is modeled separately) but also provides the opportunity to examine the resultant noise.[52]

4. Geometric Trajectory Analysis—While PCA and partial least squares are the initial methods typically used to analyze metabolomic data, numerous other techniques add to the extracted information and aid in the interpretation of results. The “trajectory” of a response, where each point represents a sequential time point plotted on a PCA map, can be used to study responses where temporal variability is of interest, such as recovery from a toxic insult or progression of a disease.[35,53,61] It is important to study the time-related changes of a response because metabolite concentrations often fluctuate, and taking measurements at only one time point may give misleading results.[28] Response trajectories aid in the visual interpretation of the magnitude of a response because a trajectory that deviates further from homeostasis will plot farther away from the control time point and a slower recovery will be visible as a higher number of points deviating from the control measurement.[35] This allows for the observation of which toxins generate more severe reactions and which toxins produce similar responses, although the latter is often more difficult to interpret. Different drugs may cause the same response in 2 different tissues but at different rates, and changes in sample size can cause 2 similar trajectories to be unrecognized because of the differences in magnitude. Keun et al.[53] developed a method to make the recognition of similar responses more accurate, despite magnitude and time scale changes. Their technique, scaled-to-maximum, aligned, and reduced trajectories (SMART) analysis, determines if 2 responses are homothetic, that is, they have shapes that are related by expansion or geometric

contraction and/or translation. Two homothetic responses share many characteristics, such as correlations in the relative size and direction of metabolic changes. This method corrects for starting position, scaling differences, and differing number of samples to determine whether 2 responses are the same, and it has proven successful in analyzing interlaboratory reproducibility of results and the differences in the responses of 2 different rodents to the same toxin.[53]

5. Entropy-Based Modeling—Entropy-based models have been proposed to study the uncertainty of a group of responses as well as the uncertainty of a particular response given specific starting concentrations of metabolites, a value termed *configurational entropy* by Veselkov et al.[62] Initial concentrations are important to consider because very slight changes can lead to very different responses that seem "disordered." Veselkov et al. also coined the terms *R-potential*, the deviation of a response from homeostasis as measured by the metabolic cost to re-achieve homeostasis, and *relative entropy*, the collective divergence of metabolic phenotypes of a group of subjects from homeostasis. This type of modeling has been successfully used to study the effects of toxins and stressors in terms of the extent and uncertainties of the responses.[62]

6. Genetic Algorithms—Genetic algorithms improve the accuracy of models and aid in finding the most robust biomarkers.[63,64] The large data sets that are collected in metabolomics experiments makes them effective tools in the interpretation of metabolomic data. Genetic algorithms mimic biological evolution (through concepts such as mutation, inheritance, breeding, natural selection, etc.) as the working principle by evolving solutions to a problem over many runs of the algorithm, ranking variables' importance by their frequency of selection in "good" runs. Furthermore, genetic algorithms that simultaneously select variables and samples for optimal use in a classifier have been shown to improve the accuracy of models more than those that select variables sequentially.[63] By using one or many of these methods, significant inferences can be made about the metabolic changes in an organism in response to a wide variety of stimuli.

B. Analysis Software

Several software programs have been developed to aid in data analysis, from data preprocessing to biomarker identification. Many of these software packages are freely available, which is important to gain popularity in the young, growing field of metabolomics. rNMR is a graphically based, open-source software designed to make the identification and quantification of metabolites from 1- and 2-dimensional NMR spectroscopy easier. Rather than the commonly used peak lists, which give only limited information, it uses regions of interest that contain all underlying NMR data within the range of regions of interest. MeltDB[65] and MetaboAnalyst[50] are 2 web-based metabolomic data processing tools that accept a variety of inputs (NMR spectra, MS peak lists, etc.) and support data preprocessing, metabolite identification, and analysis methods such as PCA and PLS-DA. These integrative programs encourage the growth of the field by making the data analysis process simpler and more efficient without requiring the installation of complex software packages. Metabolite Set Enrichment Analysis (MSEA),[66] also available as a web-based tool, aids in the detection of biologically important patterns in groups of metabolite

concentrations that may be overlooked by other methods. It contains a library of about 1000 metabolite sets that vary according to certain metabolic pathways, disease states, biofluids, and tissue locations that are compared to the uploaded metabolomic data. MassTrix[67] is a web-based tool that accepts high-precision mass spectra and presents the identified metabolites on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway maps. It is unique in that it focuses on the pathway rather than the individual metabolites so that the course of a response can be analyzed and understood in a more mechanistic context. Tools that contain a peak alignment function[68] and peak alignment algorithms also have been proposed to align NMR and MS signals that vary from sample to sample because of pH, temperature, and various other differences. Such algorithms attempt to preserve shape while moving only the *x*-axis location of peaks to reduce misinterpretations during analysis.[55] Many more programs have been developed to aid in the interpretation and analysis of metabolomic data,[69,70] all of which are freely accessible.

IV. CURRENT STATUS

A. Reference Information

1. The Consortium for Metabonomic Toxicology—Crucial for development of metabolomics as a field is the widespread availability of reference information, which has been developed in both large- and small-scale experiments. The Consortium for Metabonomic Toxicology, a collaboration between 5 major pharmaceutical companies and the Imperial College, London, assessed the use of metabonomics to study xenobiotic toxicity for approximately 150 model toxins in rats and mice. A predictive expert system that determined the organ of toxicity based on spectra then was developed, providing excellent results.[71]

2. Normal Human Metabotype—The compilation of normal human urine metabotypes through the use of 1H NMR spectroscopy has provided thorough reference information for those conducting studies of the effects of toxins. Statistical parameters, including mean and standard deviations for inter- and intraindividual variability for the major urinary metabolites, were calculated to provide baseline information and confidence intervals for future studies. This has led to the quantification of changes in metabotype based on differences in sex and diet, giving information on intra- and inter-individual variability under normal (no toxin) conditions. [72] Cross-study comparisons also have been made more efficient through the development of reference tables of metabolites that vary in response to certain toxins by integrating the information across myriad studies.[73]

3. Effects of Storage on Metabolites—Metabolomic variation that occurs because of storage techniques is important to take into account to accurately categorize metabolic fingerprints into classes. The effects of long-term storage,[34] storage temperature,[30] and borate[74] (a commonly used antibacterial preservative) have been investigated so that studies can be planned accordingly. Although borate affects the 1H-NMR peaks of citrate, mannitol, α-hydroxyisobutyric acid, and methylmalonic acid, it is a highly effective antimicrobial agent, and its effects on the 1H NMR spectra are negligible when compared to

interindividual biochemical variation.[74] Urine can be stored for up to 9 months (at −40 °C) with no significant changes to the 1H-NMR spectra.[34]

4. Sources of Biological Variation—Metabolites that vary with time of day,[20,59,75,76] age,[77,78] sex,[79] strain,[20,79] and diet[4] are important to quantify so that these known sources of variability are not misinterpreted in metabolomic experiments. Metabolite concentrations vary most greatly in urine,[33] and therefore more studies have measured this variation than any other biofluid. The effect of age on metabolic fingerprints has been investigated in both adults[77] and children,[78] which is important to consider in studies where subjects are sampled from a wide age range. One study[4] researched the urinary metabolites that vary with blood pressure across a wide range of geographic populations. Similar studies lead to more robust biomarkers by identifying metabolites that consistently change in response to specific conditions, regardless of other parameters.[4] While this variation can be a nuisance, it is often systematic and therefore can be accounted for in metabolomic data analysis.[30] It also has been shown that OSC can be used to filter data and remove diurnal metabolic variation,[59] and this method can most likely be used to filter out other unwanted sources of variation as well, although this should be studied further. Finally, easy-to-follow protocols on how to collect and analyze biofluid[73] and tissue samples[80] for metabolomic analysis have been published.

B. Pharmacometabonomics

1. Predisposition for Toxicity—One of the most promising uses of metabolomics is to noninvasively detect the site and mode of action of toxicity of xenobiotics as well as to use predose metabolomes as predictors of individual toxicity. To date, this subfield, known as pharmacometabonomics, has shown significant potential in its ability to contribute to personalized healthcare. If individuals can be quickly and easily tested for adverse drug reactions before a drug is administered, more drugs will be able to reach the market and help those who currently cannot benefit from them because of their toxicity in some patients. Winnike et al.[6] has recently distinguished those who would endure drug-induced liver injury from prolonged use of acetaminophen from those who would not before liver injury occurred. Although 1H-NMR spectroscopy of predose urine samples were not able to characterize these 2 groups, noninvasive early detection of drug-induced liver injury shows the great potential of metabolomics in personalizing healthcare. A similar study was conducted using rats and found a weak but statistically significant difference in the predose urinary metabolome of responders and nonresponders.[81]

2. Drug Toxicity Studies—Numerous studies have been performed to discover biomarkers indicative of the organ and mode of action of drug toxicity, especially for the liver and kidney. These studies will aid in drug discovery and testing by quickly and noninvasively diagnosing toxicities that are not currently easy to detect. The similarities and differences of the NMR spectra of rodents treated with various hepatotoxins have been characterized,[35] proving that NMR spectroscopy is capable of distinguishing the mechanism of xenobiotic action. Distinction between the changes in urinary metabotypes in response to tubule-directed and renal medullary nephrotoxins has been accomplished by studying the similarities and differences between different nephrotoxins.[61,82,83] Models that predict the

organ of toxicity also have been developed and show robust results.[63,84] From the perspective of clinical applications, pharmacometabonomics is still in the early stages of accurately and noninvasively testing patients for adverse drug reactions before treatment, but the above studies illustrate the significant potential of the field.

C. Disease Models

1. Critical Illness—Metabolomics has been widely used in creating models of disease either to detect the presence or severity of an illness or to follow its progression. Many studies of metabolic staging during critical illness[85] have used the metabolomes of patients with systemic inflammatory response syndrome (SIRS), multiple organ dysfunction syndrome (MODS), and sepsis. The use of metabolomics has allowed critically ill patients to be investigated from a wide variety of angles, from serum changes to microbiome alterations to variations in cerebrospinal fluid. Patients with SIRS can be distinguished from those with MODS through a combined pattern recognition and NMR spectroscopy approach, leading to the earlier detection of MODS.[58] Evidence for the theory of the "gut origin of sepsis" has been uncovered by studies that focus on changes in the microbiome during SIRS and sepsis.[21,22] Although causation has not been proved, that critical illness is associated with gut bacterial overgrowth and that this leads to a predisposition to sepsis have been confirmed.[21] The gut microbiomes of patients with SIRS have fewer obligate anaerobic bacteria and a higher pH, again revealing that a disruption in the delicate balance of commensal gut flora is associated with SIRS but providing no clear information on the underlying mechanism that causes this alteration. Sepsis is also complicated by septic encephalopathy, or septic brain dysfunction, in up to 70% of all patients; this has been studied through the use of metabolomics on cerebrospinal fluid.[11] Such information was not previously easily accessible through other approaches, but it can now be studied thoroughly.

2. Cancer—Metabolomics also has been used to investigate several types of cancer. Metabolomic data can be combined with the abundant genomic, proteomic, and other "-omics" data already available to refine our understanding. To date, metabolomics has been used primarily as a potential diagnostic tool in colorectal cancer,[86] prostate cancer,[36] and breast cancer.[87–89] Although urine and serum have not yet been successful in correctly classifying benign and cancerous prostates, sarcosine levels in tissue samples have shown excellent results, characterizing benign prostates, clinically localized prostate cancer, and metastatic disease.[36] Metabolic profiling of colon mucosae also has shown potential in the diagnosis of colorectal cancer, which currently is not usually diagnosed until late in disease progression.[86] Noninvasive detection of cancer is a realistic goal for metabolomic studies; numerous urinary nucleosides that vary with certain types of cancer (e.g., leukemia, lymphoma, nasopharyngeal cancer, breast cancer, colorectal cancer, bronchogenic carcinoma) have been measured through electromigration.[90] However, more thorough databases must be created to include the variations of these nucleosides with age, sex, diet, and other potential confounders before a conclusive test can be performed. Metabolomics also has shown potential in directing the treatment of cancer patients. Weight gain after breast cancer chemotherapy results in a decreased overall survival rate. However, various metabolite levels in urine before chemotherapy correlate with weight gain (or absence of weight gain); hence such information could target patients for intervention.[89]

3. Obesity, High Blood Pressure, and Coronary Heart Disease—Several other diseases have been studied through the use of metabolomics. Obese Zucker rats have been used as a model of type 2 diabetes, an illness that continues to escalate in Western countries. The 2 strains of rat, normal and Zucker, were determined based on metabotype alone, shedding light on the metabolites that vary with type 2 diabetes. Correlations between blood pressure and metabotypes have been investigated,[4] particularly within the context of using metabolomics to diagnose the presence and severity of coronary heart disease.[60] The relationship between metabolic phenotypes and obesity has revealed that obesity may often be caused by variations in the gut microflora,[17] in most cases by the relative abundance of *Bacteroidetes* and *Firmicutes*.[19]

D. Integration of "–Omics" Fields

Genes code for proteins, which are synthesized when the gene is expressed, and these proteins are processed into metabolites while also being used as the machinery that process and control cellular reactions. The integration of knowledge about all 3 steps through the use of genomics, proteomics, and metabolomics can greatly expand our knowledge of biological systems and the time scales of events between these 3 levels. In the absence of metabolomics, it would be difficult to validate the estimated metabolic outcomes of the lower levels. Now, metabolomic results can be combined with those of other fields to draw complete pictures of biological pathways and their interactions. It is extremely important that changes in gene expression be understood with relation to metabolic activity at the level of the whole organism to fully understand their biological functions.

1. Cytokines and Metabolites—Metabolic phenotypes have been studied in relation to cytokine levels during parasitic infections of rats to decipher connections between the immune system and metabolic pathways. While correlations have been observed between various cytokines and metabolites, further studies must be conducted to determine whether such correlations are a result of mechanistic links or simply covariation.[91] Pro- and anti-inflammatory cytokines have been found to increase in the blood of human patients with severe SIRS, although no significant correlations were found between any cytokine and anaerobic bacteria count or organic acid concentration in feces.[22] While other studies have found some possible correlations between cytokines and the gut microflora,[92,93] further studies are needed to clarify these relationships.

2. Integrative Omics-Metabolic Analysis—Integrative omics-metabolic analysis—a constraint-based method—recently has been introduced as a way to integrate proteomic and metabonomic data with genome-scale metabolic models. Integrative omics-metabolic analysis is constructed as a quadratic programming problem that aims to find steady-state flux equations that follow mass-balance and enzyme directionality constraints and is consistent with fluxes estimated by Michaelis-Menten kinetics. It has been able to successfully predict changes in fluxes both in the central metabolism of *Escherichia coli* under various genetic perturbations and in a simulated kinetic model of red blood cells. This method presents a unique and informative approach to the integration of the -omics fields, as metabolic fluxes shed a great deal of light on the state of cells and tissues.[94]

3. Trace on KEGG Pathways—Gene expression data has been linked with metabolomic data and traced using the KEGG to give a more visual representation of the pathways affected by orotic acid–induced fatty liver. Using PLS, correlations were found between gene expression and metabolomic data.[95] Studying phenotypic and metabolomic changes during genomic analysis is essential for genomic changes to be understood in a biological context rather than an artificial experimental time scale. Although to date few studies have combined the -omics fields because of the lack of metabolic information relative to other fields, the breadth of metabolomic data is quickly growing as fast and economical methods such as NMR spectroscopy become more widespread.

V. OUTLOOK AND OPEN QUESTIONS

A. Biological Variation

While metabolomics has great potential for more thoroughly understanding pathways in a systems biology sense, it has certain downsides, just as any other field of study. Biological variability in metabolomic data[31] can make finding widely applicable metabolites indicative of disease difficult to find. Metabolites also vary with circadian rhythms, diet, age, sex, and weight,[75–77, 96–98] which can be difficult to control, especially in human studies, where there are many ethical and economic limitations. Most of these factors, however, have been shown to vary systematically, so their effects can potentially be avoided through the use of automated filters to reduce unwanted variations. [59] If samples are to be assessed over time, the unwanted variability, such as metabolite degradation, can be reduced with the use of a quality control sample consisting of aliquots from each sample. This quality control set can then be randomly analyzed throughout the analytical run to see general trends in the metabolic composition over time.[99]

B. Insensitivity of Most Common Data Generation Methods

1H-NMR spectroscopy, the most commonly used method because of its economical and high-throughput properties, cannot measure metabolites at low concentrations. Many of the measured high-concentration metabolites are found in multiple pathways and therefore are not unique to a specific response. Although these high-concentration metabolites can vary greatly during a perturbation, their ubiquity decreases their ability to be used as robust biomarkers[47] because tracing exactly which pathways are disturbed can be a difficult task. However, the high-throughput nature of metabolomics has the ability to compensate for this; many different metabolites can be measured and the subtle changes can potentially be clarified.

C. Unknown Metabolites

Unknown metabolites are sometimes observed,[99] but this can potentially be avoided when sufficient metabolomic studies have been performed and large databases are created. Until that point, LC-MS, either alone[100] or in conjunction[101] with other techniques, can be used to identify unknown metabolites.

D. Complex Data

Finally, metabolomics creates a vast amount of complex data that requires multivariate analysis techniques. However, this is becoming less of a problem as the field grows and high-quality, practical preprocessing and analysis software becomes available.

E. Personalized Healthcare

Metabolomics has only recently begun to significantly affect biological and pharmacological research, but it is quickly becoming a commonly used technique offering a critical advantage in disease diagnostics.[102] It will have a large role in pharmacology in the future and will help make personalized healthcare possible. It can be used in drug development[1]: in silico models can be created to predict the effects of drugs on metabolic phenotypes before *in vivo* testing. Metabolic prescreening could also be used to predict the outcome of an intervention, thus improving the success of treatment plans. Prescreening could also allow more drugs that currently are not sold because of the adverse effects they have in some patients to reach the market and help those who could greatly benefit from their availability. Finally, metabolomics has shed light on the constituents of the gut microbiota, including the bacterial species associated with certain diseases, and therefore will influence the development of drugs that account for these factors.[14]

F. Integration of '-omics'

The integration of genomics, proteomics, and metabolomics will be the greatest contribution of metabolomics because it will improve our fundamental biological knowledge and impact many areas of biomedical research.[39] Rather than assuming or attempting to calculate unknown information, the outcomes at each level can be explicitly measured and integrated to more accurately model disease progression and drug intervention. Changes in gene expression and protein concentrations can be understood at a cellular and whole-organism level, which will help to decipher their biological functions. Interactions between the environment and the organism can also be studied at the phenotypic level and may be linked to genomic changes or inform epigenetic therapy.[27]

G. Diagnostics

Because metabolomics produces large amounts of data through relatively cheap and noninvasive techniques, it has great potential to improve disease diagnostics. [60] For this ideal to be fully reached, more sensitive data collection techniques must be used to allow for the identification of more robust biomarkers. To date, NMR spectroscopy is most commonly used, and most metabolites measured by this technique are not specific to certain conditions. Changes in the metabolome with age, diet, exercise, and sex must be more extensively studied and transcribed so that they may be accounted for before biomarker identification. Metabolomics applied to clinically accessible urine or blood samples ultimately can be used for diagnosis in a wide range of scenarios, including even critically ill patients. In that direction recent advances in microfluidics offer a promising technological development toward a broader application of metabolomics profiling.[103]

VI. CONCLUDING REMARKS

The field of metabolomics has come a long way in the past decade, but it still has a long way to go. Its potential applications in the evolution of healthcare and biomedical sciences are immense. Metabolic phenotyping has and will aid in drug discovery, disease diagnosis, personalization of healthcare, and noninvasive diagnosis of the mode of drug toxicity. Because of its current lack of sensitivity, metabolomics may not have an immediate role in all areas, although this will change as researchers continue to develop methods to improve sensitivity. The field will most likely gain popularity as more published information facilitates the comparative analysis of metabolomic data. Finally, when combined with genomics and proteomics, metabolomics represents a critical puzzle piece in the understanding of biological systems—and with that understanding, much more will possible.

Acknowledgments

The authors acknowledge the support of and motivation for this work by the late Dr. S.F. Lowry. IPA and KK acknowledge support from National Institutes of Heath grant no. GM082974. SEC was supported, in part, from National Institutes of Heath grant no. GM34695.

References

1. Nicholson JK. Global systems biology, personalized medicine and molecular epidemiology. Mol Syst Biol. 2006; 2:52. [PubMed: 17016518]
2. Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. Cell. 2008; 134(5):714–7. [PubMed: 18775301]
3. Nicholson JK, Holmes E, Lindon JC, Wilson ID. The challenges of modeling mammalian biocomplexity. Nat Biotechnol. 2004; 22(10):1268–74. [PubMed: 15470467]
4. Holmes E, Loo RL, Stamler J, Bictash M, Yap IK, Chan Q, Ebbels T, De Iorio M, Brown IJ, Veselkov KA, Daviglus ML, Kesteloot H, Ueshima H, Zhao L, Nicholson JK, Elliott P. Human metabolic phenotype diversity and its association with diet and blood pressure. Nature. 2008; 453(7193):396–400. [PubMed: 18425110]
5. Nicholson JK, Wilson ID. Opinion: understanding “global” systems biology: metabonomics and the continuum of metabolism. Nat Rev Drug Discov. 2003; 2(8):668–76. [PubMed: 12904817]
6. Winnike JH, Li Z, Wright FA, Macdonald JM, O’Connell TM, Watkins PB. Use of pharmaco-metabonomics for early prediction of acetaminophen-induced hepatotoxicity in humans. Clin Pharmacol Ther. 2010; 88(1):45–51. [PubMed: 20182423]
7. Nicholson JK, Lindon JC. Systems biology: metabonomics. Nature. 2008; 455(7216):1054–6. [PubMed: 18948945]
8. Chemometrics. Multivariate solutions to metabonomic profiling and functional genomics, part 2. Sep 2. 2006 cited 23 Nov 2013 .Available at: http://chemometrics.se/2006/09/02/multivariate-solutions-to-metabonomic-profiling-and-functional-genomics-part-2/
9. Kölsch H, Jessen F, Wiltfang J, Lewczuk P, Dichgans M, Teipel SJ, Kornhuber J, Frölich L, Heuser I, Peters O, Wiese B, Kaduszkiewicz H, van den Bussche H, Hüll M, Kurz A, Rüther E, Henn FA, Maier W. Association of SORL1 gene variants with Alzheimer’s disease. Brain Res. 2009; 1264:1–6. [PubMed: 19368828]
10. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cézard JP, Belaiche J, Almer S, Tysk C, O’Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn’s disease. Nature. 2001; 411(6837):599–603. [PubMed: 11385576]
11. Hinkelbein J, Feldmann RE Jr, Peterka A, Schubert C, Schelshorn D, Maurer MH, Kalenka A. Alterations in cerebral metabolomics and proteomic expression during sepsis. Curr Neurovasc Res. 2007; 4(4):280–8. [PubMed: 18045154]

12. 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(11):1181–9. [PubMed: 10598751]
13. Guarner F, Malagelada JR. Gut flora in health and disease. Lancet. 2003; 361(9356):512–9. [PubMed: 12583961]
14. Kinross JM, von Roon AC, Holmes E, Darzi A, Nicholson JK. The human gut microbiome: implications for future health care. Curr Gastroenterol Rep. 2008; 10(4):396–403. [PubMed: 18627653]
15. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA. Diversity of the human intestinal microbial flora. Science. 2005; 308(5728):1635–8. [PubMed: 15831718]
16. Kinross J, von Roon AC, Penney N, Holmes E, Silk D, Nicholson JK, Darzi A. The gut microbiota as a target for improved surgical outcome and improved patient care. Curr Pharm Des. 2009; 15(13):1537–45. [PubMed: 19442171]
17. Calvani R, Miccheli A, Capuani G, Tomassini Miccheli A, Puccetti C, Delfini M, Iaconelli A, Nanni G, Mingrone G. Gut microbiome-derived metabolites characterize a peculiar obese urinary metabotype. Int J Obes (Lond). 2010; 34(6):1095–8. [PubMed: 20212498]
18. Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H, Zhang Y, Shen J, Pang X, Zhang M, Wei H, Chen Y, Lu H, Zuo J, Su M, Qiu Y, Jia W, Xiao C, Smith LM, Yang S, Holmes E, Tang H, Zhao G, Nicholson JK, Li L, Zhao L. Symbiotic gut microbes modulate human metabolic phenotypes. Proc Natl Acad Sci U S A. 2008; 105(6):2117–22. [PubMed: 18252821]
19. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006; 444(7122):1027–31. [PubMed: 17183312]
20. Williams RE, Lenz EM, Evans JA, Wilson ID, Granger JH, Plumb RS, Stumpf CL. A combined (1)H NMR and HPLC-MS-based metabonomic study of urine from obese (fa/fa) Zucker and normal Wistar-derived rats. J Pharm Biomed Anal. 2005; 38(3):465–71. [PubMed: 15925248]
21. MacFie J, O'Boyle C, Mitchell CJ, Buckley PM, Johnstone D, Sudworth P. Gut origin of sepsis: a prospective study investigating associations between bacterial translocation, gastric microflora, and septic morbidity. Gut. 1999; 45(2):223–8. [PubMed: 10403734]
22. Shimizu K, Ogura H, Goto M, Asahara T, Nomoto K, Morotomi M, Yoshiya K, Matsushima A, Sumi Y, Kuwagata Y, Tanaka H, Shimazu T, Sugimoto H. Altered gut flora and environment in patients with severe SIRS. J Trauma. 2006; 60(1):126–33. [PubMed: 16456446]
23. Nicholson JK, Holmes E, Wilson ID. Gut microorganisms, mammalian metabolism and personalized health care. Nat Rev Microbiol. 2005; 3(5):431–8. [PubMed: 15821725]
24. Callinan PA, Feinberg AP. The emerging science of epigenomics. Hum Mol Genet. 2006; 15(Spec No 1):R95–101. [PubMed: 16651376]
25. Bonetta L. Epigenomics: the new tool in studying complex diseases. Nat Educ. 2008; 1(1):178.
26. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. Nat Rev Genet. 2007; 8(4):253–62. [PubMed: 17363974]
27. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. Nature. 2004; 429(6990):457–63. [PubMed: 15164071]
28. Nicholson JK, Connelly J, Lindon JC, Holmes E. Metabonomics: a platform for studying drug toxicity and gene function. Nat Rev Drug Discov. 2002; 1(2):153–61. [PubMed: 12120097]
29. Griffin JL, Bollard ME. Metabonomics: its potential as a tool in toxicology for safety assessment and data integration. Curr Drug Metab. 2004; 5(5):389–98. [PubMed: 15544433]
30. Maher AD, Zirah SF, Holmes E, Nicholson JK. Experimental and analytical variation in human urine in 1H NMR spectroscopy-based metabolic phenotyping studies. Anal Chem. 2007; 79(14): 5204–11. [PubMed: 17555297]
31. Keun HC, Ebbels TM, Antti H, Bollard ME, Beckonert O, Schlotterbeck G, Senn H, Niederhauser U, Holmes E, Lindon JC, Nicholson JK. Analytical reproducibility in (1)H NMR-based metabonomic urinalysis. Chem Res Toxicol. 2002; 15(11):1380–6. [PubMed: 12437328]
32. Lin C, Wu H, Tjeerdema R, Viant M. Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics. Metabolomics. 2007; 3(1):55–67.

33. Lindon JC, Nicholson JK, Holmes E, Everett JR. Metabonomics: metabolic processes studied by NMR spectroscopy of biofluids. Concepts Magn Reson. 2000; 12(5):289–320.
34. Beckonert O, Keun HC, Ebbels TM, Bundy J, Holmes E, Lindon JC, Nicholson JK. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat Protoc. 2007; 2(11):2692–703. [PubMed: 18007604]
35. Beckwith-Hall BM, Nicholson JK, Nicholls AW, Foxall PJ, Lindon JC, Connor SC, Abdi M, Connelly J, Holmes E. Nuclear magnetic resonance spectroscopic and principal components analysis investigations into biochemical effects of three model hepatotoxins. Chem Res Toxicol. 1998; 11(4):260–72. [PubMed: 9548796]
36. Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J, Laxman B, Mehra R, Lonigro RJ, Li Y, Nyati MK, Ahsan A, Kalyana-Sundaram S, Han B, Cao X, Byun J, Omenn GS, Ghosh D, Pennathur S, Alexander DC, Berger A, Shuster JR, Wei JT, Varambally S, Beecher C, Chinnaiyan AM. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature. 2009; 457(7231):910–4. [PubMed: 19212411]
37. Kell DB. Metabolomics and systems biology: making sense of the soup. Curr Opin Microbiol. 2004; 7(3):296–307. [PubMed: 15196499]
38. Lenz EM, Wilson ID. Analytical strategies in metabonomics. J Proteome Res. 2007; 6(2):443–58. [PubMed: 17269702]
39. Reo NV. NMR-based metabolomics. Drug Chem Toxicol. 2002; 25(4):375–82. [PubMed: 12378948]
40. Keun HC, Beckonert O, Griffin JL, Richter C, Moskau D, Lindon JC, Nicholson JK. Cryogenic probe 13C NMR spectroscopy of urine for metabonomic studies. Anal Chem. 2002; 74(17):4588–93. [PubMed: 12236374]
41. Lu X, Zhao X, Bai C, Zhao C, Lu G, Xu G. LC-MS-based metabonomics analysis. J Chromatogr B Analyt Technol Biomed Life Sci. 2008; 866(1–2):64–76.
42. O'Hagan S, Dunn WB, Knowles JD, Broadhurst D, Williams R, Ashworth JJ, Cameron M, Kell DB. Closed-loop, multiobjective optimization of two-dimensional gas chromatography/mass spectrometry for serum metabolomics. Anal Chem. 2007; 79(2):464–76. [PubMed: 17222009]
43. An Z, Chen Y, Zhang R, Song Y, Sun J, He J, Bai J, Dong L, Zhan Q, Abliz Z. Integrated ionization approach for RRLC-MS/MS-based metabonomics: finding potential biomarkers for lung cancer. J Proteome Res. 2010; 9(8):4071–81. [PubMed: 20560663]
44. Rooney OM, Troke J, Nicholson JK, Griffin JL. High-resolution diffusion and relaxation-edited magic angle spinning 1H NMR spectroscopy of intact liver tissue. Magn Reson Med. 2003; 50(5): 925–30. [PubMed: 14587002]
45. Griffin JL, Troke J, Walker LA, Shore RF, Lindon JC, Nicholson JK. The biochemical profile of rat testicular tissue as measured by magic angle spinning 1H NMR spectroscopy. FEBS Lett. 2000; 486(3):225–9. [PubMed: 11119708]
46. Coen M, Kuchel PW. Metabonomics based on NMR spectroscopy. Chem Aust. 2004; 71(6):13–17.
47. Griffin JL. The Cinderella story of metabolic profiling: does metabolomics get to go to the functional genomics ball? Philos Trans R Soc Lond B Biol Sci. 2006; 361(1465):147–61. [PubMed: 16553314]
48. Cui Q, Lewis IA, Hegeman AD, Anderson ME, Li J, Schulte CF, Westler WM, Eghbalnia HR, Sussman MR, Markley JL. Metabolite identification via the Madison Metabolomics Consortium Database. Nat Biotechnol. 2008; 26(2):162–4. [PubMed: 18259166]
49. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, Fung C, Nikolai L, Lewis M, Coutouly MA, Forsythe I, Tang P, Shrivastava S, Jeroncic K, Stothard P, Amegbey G, Block D, Hau DD, Wagner J, Miniaci J, Clements M, Gebremedhin M, Guo N, Zhang Y, Duggan GE, Macinnis GD, Weljie AM, Dowlatabadi R, Bamforth F, Clive D, Greiner R, Li L, Marrie T, Sykes BD, Vogel HJ, Querengesser L. HMDB: the Human Metabolome Database. Nucleic Acids Res. 2007; 35(Database issue):D521–6. [PubMed: 17202168]

50. Xia J, Psychogios N, Young N, Wishart DS. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. Nucleic Acids Res. 2009; 37(Web Server issue):W652–60. [PubMed: 19429898]
51. Xia J, Bjorndahl TC, Tang P, Wishart DS. MetaboMiner– semi-automated identification of metabolites from 2D NMR spectra of complex biofluids. BMC Bioinformatics. 2008; 9:507. [PubMed: 19040747]
52. Cloarec O, Dumas ME, Craig A, Barton RH, Trygg J, Hudson J, Blancher C, Gauguier D, Lindon JC, Holmes E, Nicholson J. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. Anal Chem. 2005; 77(5):1282–9. [PubMed: 15732908]
53. Keun HC, Ebbels TM, Bollard ME, Beckonert O, Antti H, Holmes E, Lindon JC, Nicholson JK. Geometric trajectory analysis of metabolic responses to toxicity can define treatment specific profiles. Chem Res Toxicol. 2004; 17(5):579–87. [PubMed: 15144214]
54. Azmi J, Griffin JL, Antti H, Shore RF, Johansson E, Nicholson JK, Holmes E. Metabolic trajectory characterisation of xenobiotic-induced hepatotoxic lesions using statistical batch processing of NMR data. Analyst. 2002; 127(2):271–6. [PubMed: 11913873]
55. Torgrip RJO, Åberg M, Karlberg B, Jacobsson S. Peak alignment using reduced set mapping. J Chemom. 2003; 17(11):573–82.
56. Cloarec O, Dumas ME, Trygg J, Craig A, Barton RH, Lindon JC, Nicholson JK, Holmes E. Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in 1H NMR spectroscopic metabonomic studies. Anal Chem. 2005; 77(2):517–26. [PubMed: 15649048]
57. Bylesjo M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. J Chemom. 2006; 20(8–10):341–51.
58. Mao H, Wang H, Wang B, Liu X, Gao H, Xu M, Zhao H, Deng X, Lin D. Systemic metabolic changes of traumatic critically ill patients revealed by an NMR-based metabonomic approach. J Proteome Res. 2009; 8(12):5423–30. [PubMed: 19835422]
59. Gavaghan CL, Wilson ID, Nicholson JK. Physiological variation in metabolic phenotyping and functional genomic studies: use of orthogonal signal correction and PLS-DA. FEBS Lett. 2002; 530(1–3):191–6. [PubMed: 12387891]
60. Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HW, Clarke S, Schofield PM, McKilligin E, Mosedale DE, Grainger DJ. Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR- based metabonomics. Nat Med. 2002; 8(12): 1439–44. [PubMed: 12447357]
61. Holmes E, Bonner FW, Sweatman BC, Lindon JC, Beddell CR, Rahr E, Nicholson JK. Nuclear magnetic resonance spectroscopy and pattern recognition analysis of the biochemical processes associated with the progression of and recovery from nephrotoxic lesions in the rat induced by mercury(II) chloride and 2-bromoethanamine. Mol Pharmacol. 1992; 42(5):922–30. [PubMed: 1435756]
62. Veselkov KA, Pahomov VI, Lindon JC, Volynkin VS, Crockford D, Osipenko GS, Davies DB, Barton RH, Bang JW, Holmes E, Nicholson JK. A metabolic entropy approach for measurements of systemic metabolic disruptions in patho-physiological states. J Proteome Res. 2010; 9(7):3537–44. [PubMed: 20423051]
63. Cavill R, Keun HC, Holmes E, Lindon JC, Nicholson JK, Ebbels TM. Genetic algorithms for simultaneous variable and sample selection in metabonomics. Bioinformatics. 2009; 25(1):112–8. [PubMed: 19010803]
64. Li Z, Srivastava S, Mittal S, Yang X, Sheng L, Chan C. A Three Stage Integrative Pathway Search (TIPS) framework to identify toxicity relevant genes and pathways. BMC Bioinformatics. 2007; 8:202. [PubMed: 17570844]
65. Neuweger H, Albaum SP, Dondrup M, Persicke M, Watt T, Niehaus K, Stoye J, Goesmann A. MeltDB: a software platform for the analysis and integration of metabolomics experiment data. Bioinformatics. 2008; 24(23):2726–32. [PubMed: 18765459]

66. Xia J, Wishart DS. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. Nucleic Acids Res. 2010; 38(Suppl):W71–7. [PubMed: 20457745]

67. Suhre K, Schmitt-Kopplin P. MassTRIX: mass translator into pathways. Nucleic Acids Res. 2008; 36(Web Server issue):W481–4. [PubMed: 18442993]

68. Izquierdo-Garcia JL, Rodriguez I, Kyriazis A, Villa P, Barreiro P, Desco M, Ruiz-Cabello J. A novel R-package graphic user interface for the analysis of metabonomic profiles. BMC Bioinformatics. 2009; 10:363. [PubMed: 19874601]

69. Zhao Q, Stoyanova R, Du S, Sajda P, Brown TR. HiRes– a tool for comprehensive assessment and interpretation of metabolomic data. Bioinformatics. 2006; 22(20):2562–4. [PubMed: 16895927]

70. Antonov AV, Dietmann S, Wong P, Mewes HW. TICL–a web tool for network-based interpretation of compound lists inferred by high-throughput metabolomics. FEBS J. 2009; 276(7): 2084–94. [PubMed: 19292876]

71. Lindon J, Keun H, Ebbels T, Pearce J, Holmes E, Nicholson J. The Consortium for Metabonomic Toxicology (COMET): aims, activities and achievements. Pharmacogenomics. 2005; 6(7):691–9. [PubMed: 16207146]

72. Zuppi C, Messana I, Forni F, Rossi C, Pennacchietti L, Ferrari F, Giardina B. 1H NMR spectra of normal urines: reference ranges of the major metabolites. Clin Chim Acta. 1997; 265(1):85–97. [PubMed: 9352132]

73. Shockcor JP, Holmes E. Metabonomic applications in toxicity screening and disease diagnosis. Curr Top Med Chem. 2002; 2(1):35–51. [PubMed: 11899064]

74. Smith LM, Maher AD, Want EJ, Elliott P, Stamler J, Hawkes GE, Holmes E, Lindon JC, Nicholson JK. Large-scale human metabolic phenotyping and molecular epidemiological studies via 1H NMR spectroscopy of urine: investigation of borate preservation. Anal Chem. 2009; 81(12):4847–56. [PubMed: 19453167]

75. Bollard ME, Holmes E, Lindon JC, Mitchell SC, Branstetter D, Zhang W, Nicholson JK. Investigations into biochemical changes due to diurnal variation and estrus cycle in female rats using high-resolution (1)H NMR spectroscopy of urine and pattern recognition. Anal Biochem. 2001; 295(2):194–202. [PubMed: 11488622]

76. Slupsky CM, Rankin KN, Wagner J, Fu H, Chang D, Weljie AM, Saude EJ, Lix B, Adamko DJ, Shah S, Greiner R, Sykes BD, Marrie TJ. Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles. Anal Chem. 2007; 79(18):6995– 7004. [PubMed: 17702530]

77. Psihogios NG, Gazi IF, Elisaf MS, Seferiadis KI, Bairaktari ET. Gender-related and age-related urinalysis of healthy subjects by NMR-based metabonomics. NMR Biomed. 2008; 21(3):195–207. [PubMed: 17474139]

78. Gu H, Pan Z, Xi B, Hainline BE, Shanaiah N, Asiago V, Gowda GA, Raftery D. 1H NMR metabolomics study of age profiling in children. NMR Biomed. 2009; 22(8):826–33. [PubMed: 19441074]

79. Gavaghan McKee CL, Wilson ID, Nicholson JK. Metabolic phenotyping of nude and normal (Alpk:ApfCD, C57BL10J) mice. J Proteome Res. 2006; 5(2):378–84. [PubMed: 16457604]

80. Beckonert O, Coen M, Keun HC, Wang Y, Ebbels TM, Holmes E, Lindon JC, Nicholson JK. High-resolution magic-angle-spinning NMR spectroscopy for metabolic profiling of intact tissues. Nat Protoc. 2010; 5(6):1019–32. [PubMed: 20539278]

81. Clayton TA, Lindon JC, Cloarec O, Antti H, Charuel C, Hanton G, Provost JP, Le Net JL, Baker D, Walley RJ, Everett JR, Nicholson JK. Pharmaco-metabonomic phenotyping and personalized drug treatment. Nature. 2006; 440(7087):1073–7. [PubMed: 16625200]

82. Anthony ML, Sweatman BC, Beddell CR, Lindon JC, Nicholson JK. Pattern recognition classification of the site of nephrotoxicity based on metabolic data derived from proton nuclear magnetic resonance spectra of urine. Mol Pharmacol. 1994; 46(1):199–211. [PubMed: 8058053]

83. Gartland KP, Bonner FW, Nicholson JK. Investigations into the biochemical effects of region-specific nephrotoxins. Mol Pharmacol. 1989; 35(2):242–50. [PubMed: 2918855]

84. Holmes E, Nicholls AW, Lindon JC, Ramos S, Spraul M, Neidig P, Connor SC, Connelly J, Damment SJ, Haselden J, Nicholson JK. Development of a model for classification of toxin-

induced lesions using 1H NMR spectroscopy of urine combined with pattern recognition. NMR Biomed. 1998; 11(4–5):235–44. [PubMed: 9719578]

85. Aller MA, Arias JI, Alonso-Poza A, Arias J. A review of metabolic staging in severely injured patients. Scand J Trauma Resusc Emerg Med. 18(1):27. [PubMed: 20478066]
86. Chan EC, Koh PK, Mal M, Cheah PY, Eu KW, Backshall A, Cavill R, Nicholson JK, Keun HC. Metabolic profiling of human colorectal cancer using high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy and gas chromatography mass spectrometry (GC/MS). J Proteome Res. 2009; 8(1):352–61. [PubMed: 19063642]
87. Fan X, Bai J, Shen P. Diagnosis of breast cancer using HPLC metabonomics fingerprints coupled with computational methods. Conf Proc IEEE Eng Med Biol Soc. 2005; 6:6081–4. [PubMed: 17281650]
88. Frickenschmidt A, Frohlich H, Bullinger D, Zell A, Laufer S, Gleiter CH, Liebich H, Kammerer B. Metabonomics in cancer diagnosis: mass spectrometry-based profiling of urinary nucleosides from breast cancer patients. Biomarkers. 2008; 13(4):435–49. [PubMed: 18484357]
89. Keun HC, Sidhu J, Pchejetski D, Lewis JS, Marconell H, Patterson M, Bloom SR, Amber V, Coombes RC, Stebbing J. Serum molecular signatures of weight change during early breast cancer chemotherapy. Clin Cancer Res. 2009; 15(21):6716–23. [PubMed: 19861449]
90. Markuszewski MJ, Struck W, Waszczuk-Jankowska M, Kaliszan R. Metabolomic approach for determination of urinary nucleosides as potential tumor markers using electromigration techniques. Electrophoresis. 2010; 31(14):2300–10. [PubMed: 20564268]
91. Saric J, Li JV, Swann JR, Utzinger J, Calvert G, Nicholson JK, Dirnhofer S, Dallman MJ, Bictash M, Holmes E. Integrated cytokine and metabolic analysis of pathological responses to parasite exposure in rodents. J Proteome Res. 2010; 9(5):2255–64. [PubMed: 20092362]
92. Yang R, Han X, Uchiyama T, Watkins SK, Yaguchi A, Delude RL, Fink MP. IL-6 is essential for development of gut barrier dysfunction after hemorrhagic shock and resuscitation in mice. Am J Physiol Gastrointest Liver Physiol. 2003; 285(3):G621–9. [PubMed: 12773301]
93. Guo W, Magnotti LJ, Ding J, Huang Q, Xu D, Deitch EA. Influence of gut microflora on mesenteric lymph cytokine production in rats with hemorrhagic shock. J Trauma. 2002; 52(6): 1178–85. discussion 85. [PubMed: 12045650]
94. Yizhak K, Benyamini T, Liebermeister W, Ruppin E, Shlomi T. Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. Bioinformatics. 2010; 26(12):i255–60. [PubMed: 20529914]
95. Griffin JL, Bonney SA, Mann C, Hebbachi AM, Gibbons GF, Nicholson JK, Shoulders CC, Scott J. An integrated reverse functional genomic and metabolic approach to understanding orotic acid-induced fatty liver. Physiol Genomics. 2004; 17(2):140–9. [PubMed: 14747661]
96. Kochhar S, Jacobs DM, Ramadan Z, Berruex F, Fuerholz A, Fay LB. Probing gender-specific metabolism differences in humans by nuclear magnetic resonance-based metabonomics. Anal Biochem. 2006; 352(2):274–81. [PubMed: 16600169]
97. Gu H, Chen H, Pan Z, Jackson AU, Talaty N, Xi B, Kissinger C, Duda C, Mann D, Raftery D, Cooks RG. Monitoring diet effects via biofluids and their implications for metabolomics studies. Anal Chem. 2007; 79(1):89–97. [PubMed: 17194125]
98. Bollard ME, Stanley EG, Lindon JC, Nicholson JK, Holmes E. NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. NMR Biomed. 2005; 18(3):143–62. [PubMed: 15627238]
99. Sangster TP, Wingate JE, Burton L, Teichert F, Wilson ID. Investigation of analytical variation in metabonomic analysis using liquid chromatography/mass spectrometry. Rapid Commun Mass Spectrom. 2007; 21(18):2965–70. [PubMed: 17680628]
100. Ma S, Chowdhury SK, Alton KB. Application of mass spectrometry for metabolite identification. Curr Drug Metab. 2006; 7(5):503–23. [PubMed: 16787159]
101. Dear GJ, Ayrton J, Plumb R, Sweatman BC, Ismail IM, Fraser IJ, Mutch PJ. A rapid and efficient approach to metabolite identification using nuclear magnetic resonance spectroscopy, liquid chromatography/mass spectrometry and liquid chromatography/nuclear magnetic resonance spectroscopy/sequential mass spectrometry. Rapid Commun Mass Spectrom. 1998; 12(24):2023–30.

102. Ellis DI, Dunn WB, Griffin JL, Allwood JW, Goodacre R. Metabolic fingerprinting as a diagnostic tool. Pharmacogenomics. 2007; 8(9):1243–66. [PubMed: 17924839]

103. Kraly JR, Holcomb RE, Guan Q, Henry CS. Review: microfluidic applications in metabolomics and metabolic profiling. Anal Chim Acta. 2009; 653(1):23–35. [PubMed: 19800473]