Published in final edited form as:

Semin Nephrol. 2010 September 1; 30(5): 500–511. doi:10.1016/j.semnephrol.2010.07.007.

Analytical Approaches to Metabolomics and Applications to Systems Biology

Jeffrey H. Wang, MD [Assistant Professor]

Department of Internal Medicine Hennepin County Medical Center and University of Minnesota School of Medicine Minneapolis, Minnesota, USA

Division of Nephrology, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan 48109-5676

Jaeman Byun, PhD [Research Investigator]

Department of Internal Medicine University of Michigan Medical School Ann Arbor, Michigan, USA

Division of Nephrology, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan 48109-5676

Subramaniam Pennathur, MD [Assistant Professor]*

Department of Internal Medicine University of Michigan Medical School Ann Arbor, Michigan, USA

Division of Nephrology, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan 48109-5676

Abstract

Phenotypic expression of renal diseases encompasses a complex interaction between genetic, environmental and local tissue factors. The level of complexity requires integrated understanding of perturbations in the network of genes, proteins and metabolites. Metabolomics attempts to systematically identify and quantitate metabolites from biological samples. The small molecules represent the end result of complexity of biological processes in a given cell, tissue or organ and thus form attractive candidates to understand disease phenotypes. Metabolites represent a diverse group of low-molecular weight structures including lipids, amino acids, peptides, nucleic acids and organic acids which makes comprehensive analysis a difficult analytical challenge. The recent rapid development of a variety of analytical platforms based on mass spectrometry (MS) and nuclear magnetic resonance (NMR) have enabled separation, characterization, detection and quantification of such chemically diverse structures. Continued development of bioinformatics and analytical strategies will accelerate widespread use and integration of metabolomics into systems biology. Here, we will discuss analytical and bioinformatic techniques and highlight recent studies that utilize metabolomics in understanding pathophysiology of disease processes.

Keywords

Metabolomics; mass spectrometry; systems biology; bioinformatics; renal disease

*Address Correspondence to: Subramaniam Pennathur 1570 MSRBII, SPC 5676 1150 W. Medical Center Drive University of Michigan Ann Arbor, MI, 48109-5676 Fax: 734-763-0982 spennath@umich.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Biomedical research has been traditionally pursued using a reductionist approach. In reductionism, a complex problem is divided into smaller pieces with the assumption that it can be solved by studying the less complex subunits in isolation. Although reductionism is well-suited for simple disorders, it is suboptimal to study chronic and complex diseases such as renal disease which are driven by the dynamic interactions of many diverse factors (e.g. genetics, inflammation and environment; Ref: 1,2). In the past two decades, a new field of science called systems biology has emerged. Unlike reductionism which focuses only on one part of a disease, systems biology studies disease from a global level and examines the interrelationships and dynamics between the parts and thus may be a better tool for studying complex disorders. The goal of systems biology is to characterize an organism on a gene (genomics), transcript (transcriptomics), protein (proteomics) and metabolite (metabolomics) level. Once assembled into the “omics cascade”, these datasets comprehensively describe the response of a biological system to disease, genetic, pharmacologic and environmental perturbations (Figure 1).

Before the recent advances in molecular biology, understanding disease mechanisms was essentially a biochemical endeavor. Indeed, diagnosis of diverse metabolic diseases depends on accurate measurement of key metabolites such as glucose in diabetes, cholesterol in hypercholesterolemia and creatinine in renal failure. In principle, low-molecular weight metabolites represent the “composite output” of the cellular machinery accounting for genomic, transcriptomic and proteomic variability. Thus, metabolic changes represent the most proximal alteration of the cell, tissue or the organism to a disease process. Metabolomics which aims to catalog all small molecule metabolites in an organism – the metabolome –thus offers unique promise. Unlike genomics and proteomics which have well-described methodologies and often even core laboratories, metabolomic techniques are still evolving. Despite this, there has been recent surge in interest in metabolomics. The intent of this review is to serve as a primer for the nephrology research community who are considering utilizing metabolomics in their research.

2. An Overview of Metabolomics

The goal of metabolomics is to comprehensively identify and quantify all endogenous and exogenous small molecule metabolites in a biological system in a high-throughput manner. Although scientists have been analyzing chemicals in human specimens for centuries, modern metabolomics is a relatively new science. In 1971, Linus Pauling and Arthur Robinson proposed that quantitation of metabolites in biofluids may represent the functional status of a biological system³. The term “metabolic profiling” was first utilized by Horning when describing gas phase analytical methods for analysis of metabolites from human urine samples⁴. “Metabolome” was first used by Oliver in 1998⁵ to describe the collection of small molecules in an organism and metabolomic profiling of large numbers of small molecules was termed “metabonomics” by Nicholson⁶. Unlike the human genome, the human metabolome has not been fully mapped so the number of metabolites present in the human body is unknown. Current estimates vary, but the Human Metabolome Database (HMDB; <http://www.hmdb.ca>) has identified and experimentally confirmed 6,826 small molecules in various human tissues and biofluids⁷. Although this number pales in comparison to the estimated 20,000 genes and 100,000 transcripts in the human genome and transcriptome respectively, it likely only reflects a minute portion of the true metabolome. For example, Wishart who is the curator of the HMDB, postulates that current technology can only detect about 20% of all lipids and thus the size of the metabolome may increase exponentially with future advances in instrumentation⁷.

One of the challenges of doing metabolomic studies is the complexity of the metabolome. The metabolome contains a wide variety of chemically diverse compounds such as lipids, organic acids, carbohydrates, amino acids, nucleotides and steroids among others (reviewed in^{8–14}). In fact, in the latest iteration of the HMDB, 52 different classes of compounds were represented by the reported metabolites⁷. In comparison, genes and proteins may perhaps be more chemically homogenous as each gene is some combination of only 4 basic nucleotides while each protein is composed of a mixture of 32 amino acids. The variability in chemical structures results in a collection of analytes with very different physiochemical properties such as polarity, solubility and volatility. Another issue is that metabolites occur in a wide dynamic range of concentrations (nanomolar to millimolar) in the human body. A third hurdle is that not every metabolite is present in each tissue or biofluid. Finally, the metabolome may be “contaminated” by exogenous metabolites obtained from food or medications which also may not be uniform in each subject. Therefore, it is evident that comprehensive metabolomics is an analytical challenge. Indeed, no single metabolomics methodology is currently able to measure the entire metabolome accurately.

In the metabolomics community, research groups design metabolomic experiments with either an untargeted or targeted approach. In untargeted metabolomics also known as “unbiased or undirected metabolomics” or “metabolic fingerprinting”, the goal is to detect as many metabolites as possible in a sample in order to classify phenotypes based on a metabolite pattern. Untargeted metabolomic studies are hypothesis-generating and well-suited for biomarker discovery. Meanwhile, in targeted metabolomics which is also known as “biased or directed metabolomics” or “metabolic profiling”, the focus is limited to either a pre-determined set of metabolites or a specific chemical class of small molecules such as tricarboxylic acid cycle metabolites or lipids, respectively. Unequivocal identification and absolute quantitation are essential to a targeted experiment and require use of isotope labeled internal standards. While hundreds of either ¹³C or deuteriated standards are commercially available, this represents only a small fraction of all metabolites. In contrast to untargeted studies, targeted metabolomics are hypothesis-driven to validate critical biological pathways or confirm an untargeted study.

3. The Metabolomics Workflow

A metabolomics experiment (Figure 2) can be divided into four parts: 1) sample acquisition and preparation, 2) separation and detection of analytes, 3) data mining and extraction, 4) data analysis.

3.1 Sample Acquisition and Preparation

Metabolites can be measured in a variety of different samples including tissue, biofluids (blood, urine, feces, seminal fluid, saliva, bile, cerebrospinal fluid) and cell culture¹⁴. In nephrology, blood, urine and kidney tissue are the most commonly used biospecimens. Tissue should be quenched immediately after harvesting with liquid nitrogen in order to arrest metabolism and prevent induction of “stress” metabolites which can confound the analysis¹⁴. It is generally accepted that metabolite levels can be affected by factors such as age, gender, diet, activity level and medication use¹⁵. Sample preparation typically entails metabolite extraction and enrichment, depletion of proteins and removal of sample matrix. Processing the samples is a double-edged sword because each step will result in some degree of metabolite loss. Numerous sample preparation protocols have been developed for small molecule extraction and most involve solvents and subsequent sample clean-up with solid phase extraction. The reader is referred to excellent reviews for more details on sample preparation optimization techniques^{9,14,16}

3.2.1 Separation and Detection of Metabolites—The two main analytical platforms used in metabolomics are nuclear magnetic resonance (NMR) and mass spectrometry (MS). Both instruments are capable of reproducible and high throughput measurement of large numbers of metabolites.

3.2.2 NMR Spectroscopy—In NMR spectroscopy, a compound is placed in a magnetic field. Isotopes within the compound (e.g. ^1H , ^{13}C , ^{14}N , ^{15}N , ^{17}O) absorb the radiation and resonate at a frequency which is dependent on its location in the small molecule¹⁷. The resultant NMR spectrum is a collection of peaks at different positions and intensities and each compound has a unique pattern. NMR chemical shifts are reported relative to an internal reference. Unlike MS, NMR is non-discriminating (any compound with protons, carbon, nitrogen or oxygen can be detected), does not destroy the sample during analysis, and little to no sample preparation is required. Perhaps the greatest advantage of NMR is that it can reveal structural information which can facilitate the identification of an unknown metabolite. NMR has one major drawback, namely its low sensitivity¹⁸. As a result, low abundance metabolites are routinely missed with NMR based approaches and consequently, there has been a recent shift in the metabolomics community toward the use of more MS based techniques⁹.

3.2.3 MS—The overall strategy for identifying target analytes by MS is outlined in Figure 2. Biomolecules derived from the sample are separated by liquid chromatography (LC) or gas chromatography (GC) and ionized. The mass-to-charge (m/z) ratios of ions derived by fragmenting the ionized parent compound are determined by MS. A full scan mass spectrum obtained from target analytes can be acquired utilizing data acquisition software. Bioinformatic tools can be employed to identify differences in key metabolites. Unequivocal identity of target biomolecules can be established by utilizing spectral libraries. Alternatively, the analyte is quantified by tandem mass spectrometry (MS/MS). With chemical ionization (CI) or electron impact ionization (EI), it is possible to detect and quantify sub-femtomole levels of biomolecules. The advantages of MS based methodologies are its speed, high selectivity, ability to be linked to GC or LC and high sensitivity⁸. The primary disadvantages of MS are that it is destructive to the sample, is discriminating (not all samples ionize under a given condition and the source ion polarity may have to be modified), typically requires sample preparation which can result in metabolites losses, and cost.

3.2.4 Gas Chromatography/Mass Spectrometry (GC/MS)—GC/MS is a commonly used platform in metabolomics and is excellent for measurement of volatile compounds such as fatty acids and organic acids. As separation in GC occurs in an oven at high temperatures, analytes need to be volatile and thermally stable and it is often necessary to derivatize samples prior to analysis. While necessary, it is important to keep in mind that derivatization is additional sample processing that can result in metabolite loss and this is one of the major drawbacks of GC/MS. Samples entering the source in GC/MS are ionized by EI or CI. Untargeted and targeted metabolomics can be done with GC/MS by operating in full scan and selected ion monitoring mode, respectively. One of the most useful aspects of GC/MS is that compounds will have characteristic spectral patterns and extensive libraries are available online. Moreover, unlike LC, GC retention times are robust and reproducible even on different machines and thus they can be used in database searches to identify unknown metabolites. However, GC/MS has limited mass range and the molecular ion is often not detected due to fragmentation which hinders identification of unknown compounds. Recently, alternative strategies have been employed to improve separation and sensitivity of complex metabolite mixtures. These include GC $\times$ GC, also known as “comprehensive GC,” which separates complex samples by diverting each peak from a GC column to a second GC column. Mass accuracy (and thereby compound identification) can be improved by interfacing GC or GC $\times$ GC with a time of flight (TOF) mass analyzer¹⁹.

3.2.5 Liquid Chromatography/Mass Spectrometry (LC/MS)—Biomedical mass spectrometry has recently been revolutionized by the advent of electrospray ionization (ESI) which permits coupling of LC with MS^{20,21}. This platform is now used predominantly in metabolomics today. Its popularity stems from its tremendous versatility which enables analysis of a wide variety of small molecules. Since high temperatures and need for volatility are not involved, sample derivatization is not needed.

Achieving broad coverage of the metabolome is one of the major challenges of LC/MS. Reversed phase LC using C18 columns are widely used and provide reasonably good separation of nonpolar and weakly polar compounds. However, polar analytes will not be retained by C18 columns. Recently, Tolstikov and Fiehn reported the use of a new type of LC called hydrophilic interaction chromatography (HILIC) to analyze polar compounds in plant extracts²². In contrast to reversed phase LC, HILIC columns with modified stationary phases preferentially retain polar compounds which elute off when the mobile phase becomes more aqueous. Despite its advantages, HILIC has distinct disadvantages which include poor retention time reproducibility and analytical drift with analysis of multiple samples.

Recently, several new alternatives have emerged which improve the sensitivity and dynamic range of metabolite detection. Atmospheric pressure chemical ionization (APCI) is typically used for non-polar compounds such as lipids. APCI is a robust ionization technique which induces little fragmentation and preserves the molecular ion and is ideal for thermally stable lipid species. Capillary LC and nano-LC with low flow rates have been coupled with ESI and APCI to offer outstanding dynamic range and sensitivity. These techniques which have been used extensively in proteomics have now been adopted for metabolomic studies. The resultant microdroplets require less time for evaporation and improve both chromatographic resolution and detection by the mass analyzer. Matrix Assisted Laser Desorption Ionization (MALDI) has been used recently for analysis of amino acids and metabolites from islets and prokaryotic cells^{23,24}. However, MALDI has significant limitations which include ion suppression and matrix interference for low-molecular weight compounds.

Several mass analyzers can be used in LC/MS based metabolomic studies and the decision of which machine to use is typically predicated on the research question and more importantly, what is available at each investigator's institution. The most common analyzers that are employed in metabolomic studies are the quadrupole, ion trap and time of flight (TOF) based analyzers. Other types include Fourier Transform (FT) MS and Orbitraps. The triple Quadrupole (QQQ) combines 3 quadrupoles in series and allows MS/MS to be performed. The QQQ with multiple reaction monitoring (MRM) can be used to experimentally confirm potential biomarkers. An ion-trap mass analyzer is similar to a QQQ in that it can focus on a group of ions with a particular m/z . A unique feature of ion-traps is that they can do a multi-step MS/MS also known as MSⁿ, where a particular fragment ion is saved in the chamber, subjected to repeated disassociation. Newer linear ion-trap machines provide excellent sensitivity and very high mass resolution and accuracy within a limited mass range. Quadrupole-TOFs are high-end mass analyzers that combine the high mass resolution and high mass accuracy with the MS/MS capabilities of a QQQ. While expensive, these hybrid instruments are well-suited for all types of metabolomic projects. FTMS analyzers provide the highest mass resolution of 100,000, the best accuracy (<1 ppm mass error) and have MS/MS and MSⁿ capabilities. However, the high cost of this instrument precludes routine use for most investigators.

3.3 Data Mining and Extraction

Untargeted LC/MS and GC/MS metabolomic analysis generate large amounts of complex data sets that require analysis by specialized software to properly interpret the data. There are several commercial and free software packages available to automate the process of peak selection,

evaluation and relative quantitation. Ideally, software programs should be capable of background spectral filtering (noise elimination), appropriate peak assignment for same compound (identification of matching m/z and assigning adducts appropriately), peak alignment (matching peaks across multiple samples) and peak normalization (adjusting peak intensities and reducing analytical drift). While the current available software programs attempt to do this with varying degrees of success, there are still significant flaws with each program. Unfortunately, most mass spectrometers generate data that can only be read with proprietary software issued by the manufacturer of the instrument. For this reason, there currently is no universal metabolomics software for data extraction and analysis. While universal data formats such as $mzXML$ have been proposed, they are not widely adopted yet²⁵. For critical assessment of the alignment strategies and data optimization the readers are referred to other excellent articles in the area^{11,25,26}. For targeted analysis, data processing and analysis is relatively straight-forward as authentic standards are usually available to optimize the analytical strategy.

3.4 Data Analysis

Pattern recognition, multivariate and multidimensional statistical programs have been developed to facilitate and filter large amounts of untargeted LC/MS and GC/MS data. One approach involves chemometric techniques such as principal component analysis (PCA) and partial least squares-discriminant analysis to identify the spectral pattern and intensities of the metabolites in each sample. As this type of analysis is focused on visualizing patterns, identification and quantification of the metabolites is not done. The utility of this approach is shown in a recent paper by Kim and colleagues where they performed unbiased metabolomics on urine taken from patients with renal cell carcinoma and healthy controls²⁷. The group discovered 1,766 potential metabolites and demonstrated that PCA of the urine profiles was able to segregate patients with renal cell carcinoma from the controls. These exciting results suggest that it may be possible to develop a noninvasive screening test for renal cell carcinoma. The second approach attempts to identify and quantify each metabolite in the sample followed by multivariate statistical analyses to determine which metabolites are differentially expressed between the experimental groups. While theoretically this pathway is straightforward, in reality the inability to obtain identifications on a significant proportion of the metabolites makes it difficult. The largest and most popular human metabolite databases are HMDB⁷ and METLIN (<http://metlin.scripps.edu>); others include Kyoto Encyclopedia of Genes and Genomes (KEGG) and the Madison Metabolomics Consortium Database⁸. LIPID maps (<http://www.lipidmaps.org/>) provides an useful database to search for lipid metabolites²⁸.

An individual way to identify compounds in GC/MS is by spectral matching where the experimental mass spectrum of the unknown is compared to a spectral library such as National Institute of Standards and Technology (<http://www.webbook.nist.gov/chemistry>) and Wiley Registry (<http://www.wileyregistry.com>). EI generates reproducible spectra even among different mass spectrometers and so it is possible to identify compounds based on GC/MS profiling. However, the spectral libraries are not personalized for biological samples and may be limited in the number of analytes that can be identified and quantitated.

One major limitation of the current technology platforms is that systematic error can be introduced throughout the metabolomic analysis as minor variations in experimental conditions, sample preparation and instrumental parameters can result in significant differences. For higher specificity, two orthogonal properties should be employed. For example, retention time which depicts a physical property (hydrophobicity, volatility) and accurate mass or fragmentation pattern which depicts structure should be utilized. We advocate routine use of internal standards to normalize datasets to account for instrumental drift and ionization efficiency. Due to the above limitations, it is estimated that at most 30% of metabolites can be identified⁸. Thus, it is critical that putative identifications obtained through

accurate mass matching be experimentally validated with isotope labeled authentic standards before they can be confidently reported.

4. Applications of Metabolomic Strategies in Biomedical Research

4.1 Untargeted MS Based Metabolomics Applied to Diabetic Nephropathy (DN) Research

The pathogenesis of DN is poorly understood and likely involves genetic and environmental factors in addition to hyperglycemia. While it appears to be a heterogeneous disease with some patients progressing rapidly to end stage renal disease while others having a more insidious course, there currently are no good biomarkers that can predict patient outcomes. The complex nature of the disease makes it ideal for a metabolomic study to explore underlying unifying mechanisms of disease progression and to identify novel biomarkers of progressive DN. We analyzed the urine from streptozotocin-induced diabetic and control DBA/2J mice (a pathophysiologically relevant model for DN) treated with rosiglitazone (which reverses DN phenotype) for 10 weeks using LC/ESI/TOF²⁹. A total of 56 features showed up- or down-regulation by >2-fold in the diabetic animals. Of the 56 molecular features, 32 were identified with a mass error of <200 ppm. Of these, we were able to identify nine compounds that returned back to baseline with rosiglitazone therapy, and were therefore potential biomarkers for DN and response to treatment and reversal of DN phenotype. We obtained preliminary identification of the features with publicly available exact mass databases including METLIN Metabolite Database and the HMDB. These included compounds such as ubiquinone and indoxyl sulfate. These metabolites may be involved in both metabolic pathways that cause DN as well as those responsible for the protective effects of rosiglitazone. Interestingly, a recent study³⁰ highlighted that serum indoxyl sulfate correlates inversely with renal function and a direct relationship with aortic calcification and pulse wave velocity in 139 patients with chronic kidney disease. In survival analyses, the highest indoxyl sulfate tertile was a powerful predictor of overall and cardiovascular mortality. The predictive power of indoxyl sulfate for death was maintained after adjustment for age, gender, diabetes, albumin, hemoglobin, phosphate and aortic calcification.

4.2 Untargeted NMR Based Metabolomics in Kidney Transplantation Research

Early identification of acute rejection in kidney transplant recipients is of vital importance. While the gold standard for diagnosis, percutaneous kidney transplant biopsy, is a relatively low risk procedure, it is an invasive procedure nonetheless with potential risks. Furthermore, it is impractical to perform serial biopsies to monitor disease activity or progression. Therefore, there has been intense interest in developing noninvasive biomarkers that accurately detect rejection and monitor graft function. Several previous studies have attempted to identify urine biomarkers of graft dysfunction and virtually all of them have used untargeted NMR based metabolomic platforms^{14,31–34}. Trimethylamine-N-oxide (TMAO), an organic amine, has been consistently found to be elevated in the studies. Foxall³² et al compared urinary TMAO levels by NMR 7 days after transplantation in patients with good graft function versus patients with graft dysfunction (9 with biopsy proven acute rejection, 1 with acute tubular necrosis). The patients with graft dysfunction had significantly higher TMAO levels as compared to those with good graft function (410 ± 102 vs. 91 ± 18 μ M TMAO/mM creatinine, $p \leq 0.025$). TMAO is a marker of renal medullary damage and is believed to protect blood proteins against the effects of toxins such as guanidine and urea nitrogen that may accumulate during renal failure³¹. A number of other serum and urine metabolic alterations have also been observed in kidney graft dysfunction (e.g. elevated lactate, acetate, succinate, ethanol, urea nitrogen; decreased nitrates and nitrites)^{31,35}. Given that it is possible that no single biomarker will have enough specificity to diagnose renal transplant rejection, groups have looked at using combinations of metabolites. In one study, urine ¹H NMR spectral patterns were used to predict the outcome of a protocol biopsy in kidney transplant recipients³⁴. Urine was obtained from

68 patients prior to a protocol biopsy. Histopathology was normal in 33 patients and showed acute rejection in 35 patients. Metabolite spectral patterns appropriately classified 26 out of 33 normal patients and 27 out of 35 acute rejection patients. Most of these previous studies used NMR techniques and it will be interesting to see if ongoing projects using MS based metabolomics yield new results.

4.3 Targeted Metabolomics Applied to Cardiovascular Disease

In a recent study, targeted metabolomics was used to identify potential biomarkers of myocardial injury³⁶. Targeted metabolomic analysis was conducted on blood samples obtained in patients undergoing alcohol septal ablation treatment for hypertrophic obstructive cardiomyopathy, a model for, planned myocardial injury (PMI), spontaneous myocardial injury (SMI), and patients without myocardial injury undergoing cardiac catheterization. The authors chose metabolites involved in a wide range of metabolic pathways as well as metabolites that have been associated with cardiovascular disease. The authors discovered that metabolic alterations could be seen as early as 10 minutes after a PMI. Moreover, they identified a set of metabolites (aconitic acid, hypoxanthine, TMAO, threonine) that was differentially expressed in SMI as compared to patients undergoing cardiac catheterization without myocardial injury. These results identify a role for targeted metabolic profiling in the early detection of myocardial injury.

4.4 Combined Use of Untargeted and Targeted Metabolomic Strategies in Prostate Cancer Research

Cancer researchers are increasingly using untargeted MS-based metabolomics to discover noninvasive biomarkers for the early detection of various cancers. In a recent article by Sreekumar et al in collaboration with our group, urine, plasma and prostate tissue from prostate cancer patients and healthy controls were analyzed using untargeted GC/MS and LC/MS³⁷. 42 prostate tissue specimens (adjacent benign=16, clinically localized prostate cancer (PCa) = 12 and metastatic prostate cancer = 14) and 110 each of biopsy-proven matched urine and plasma specimens (benign = 51, PCa =59) were collected prospectively from patients, post-digital rectal exam, prior to needle biopsy. Using an untargeted MS based metabolomic platform, the authors identified and correlated over 1100 compounds in the three sample types. Evaluation of the metabolomic profiles of plasma or urine did not identify robust differences between biopsy positive and biopsy negative individuals. Thus, the initial focus was directed towards understanding the tissue metabolomic profiles as they exhibited more robust alterations. In total, high throughput profiling of the tissues quantitatively detected 626 metabolites (175 named, 19 isobars and 432 metabolites without identification), of which 515/626 were shared by the three diagnostic classes. Importantly, there were 60 metabolites found in PCa and/or metastatic tumors but not in benign prostate. These profiles are displayed as a scatter plot of the benign-based Z-transformed data (Figure 3). The dendrogram analysis demonstrated a strong separation between the metastatic and control samples and a significant separation between the control and tumor tissue types. Mapping the differential metabolomic profiles to their respective biochemical pathways utilizing KEGG revealed an increase in amino acid metabolism and nitrogen breakdown pathways during cancer progression to metastatic disease. Sarcosine, a key metabolite that was initially discovered by the untargeted approach was quantitatively confirmed by the more precise targeted approach raising the possibility that this marker could be a novel biomarker for prostate cancer progression³⁷. Moreover, sarcosine influences prostate cancer cell invasion and regulatory enzymes of sarcosine such as glycine-N-methyltransferase, sarcosine dehydrogenase, dimethyl dehydrogenase may be potential targets for prostate cancer treatment.

5. Integration of Metabolomics in Systems Biology of Complex Disorders

Although the metabolome is most predictive of the final phenotype, full understanding of an organism on a molecular level cannot be achieved without incorporating data from the whole genome-association, transcriptome, and proteome data sets. A few recent studies have shown that metabolite profiles are heritable and have identified novel gene/metabolite networks by integrating genomic, transcriptomic and metabolomic datasets^{38–41}. In one such study, targeted metabolomics, whole genome single nucleotide polymorphisms (SNP) and microarray analyses were performed on mouse liver samples from F2 cross generated from a diabetes resistant C57BL/6-*ob/ob* and diabetes-susceptible BTBR-*ob/ob* mouse models³⁸. The authors integrated the datasets with SNP analysis and showed that there were groups of liver metabolites that mapped to distinct chromosomal regions suggesting that they were under the control of genes in those regions. Shah and colleagues⁴⁰ performed MS based targeted metabolomic profiling of 66 metabolites (acylcarnitine species, amino acids, free fatty acids) in plasma from 117 individuals in 8 multiplex families of premature coronary artery disease (CAD). Heritability was calculated for each group of metabolites and several metabolites were highly heritable including β -hydroxybutyrate and C2-acylcarnitine ($h^2=0.61$), short and medium chain acylcarnitines ($h^2=0.39$), amino acids ($h^2=0.44$), long chain acylcarnitines ($h^2=0.39$), and branched-chain amino acids ($h^2=0.27$). The results suggest that some metabolic derangements in premature CAD might be determined by genetics. Another study combined SNP analysis and metabolic phenotyping in a German population (363 metabolites in serum from 284 subjects), revealing associations between frequent SNPs and alterations in specific metabolites⁴¹. The authors speculated that such an approach to personalized health care based on a combination of genotyping and metabolic characterization may provide novel application of defining disease phenotype and response to therapy. While these are early attempts to integrate metabolomics with the other “omic” disciplines, the promise of the above studies suggest this is a fertile future area of research to explore.

6. Conclusions and Future Perspectives

Metabolomics is a rapidly expanding new –omics science that can provide integrated insight into complex disorders including renal disease. In spite of significant advances there are several limitations of current technology platforms. As outlined in the review, these include lack of a single platform to comprehensively analyze the metabolome, incomplete spectral libraries and databases and flaws in current software programs with data extraction and analysis. Despite these limitations, the application of metabolomics has already led to notable advances in prostate cancer, DN, kidney transplantation and acute coronary syndrome research. Future technology development combined with more robust data analysis and bioinformatic tools will help overcome current limitations and fully integrate small molecule biochemistry to systems biology. As metabolomics is complimentary to genomics, transcriptomics and proteomics, full integration of these datasets will ultimately lead to personalized molecular diagnosis and treatment of diseases.

Acknowledgments

This work is supported in part by grants from the National Institutes of Health (HL094230, HL092237, DK082841, HL092237-02S109), the Doris Duke Foundation Clinical Scientist Development Award and by the Molecular Phenotyping Core of the Michigan Nutrition and Obesity Center(P30DK089503).

References

1. Ahn AC, Tewari M, Poon CS, Phillips RS. The limits of reductionism in medicine: could systems biology offer an alternative? *PLoS Med* 2006;3(6):e208. [PubMed: 16681415]

2. Ahn AC, Tewari M, Poon CS, Phillips RS. The clinical applications of a systems approach. *PLoS Med* 2006;3(7):e209. [PubMed: 16683861]
3. Pauling L, Robinson AB, Teranishi R, Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proc Natl Acad Sci U S A* 1971;68(10):2374–2376. [PubMed: 5289873]
4. Horning EC, Horning MG. Metabolic profiles: gas-phase methods for analysis of metabolites. *Clin Chem* 1971;17(8):802–809. [PubMed: 5105517]
5. Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. *Trends Biotechnol* 1998;16(9):373–378. [PubMed: 9744112]
6. 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):1181–1189. [PubMed: 10598751]
7. Wishart DS, Knox C, Guo AC, et al. HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res* 2009;37(Database issue):D603–610. [PubMed: 18953024]
8. Bain JR, Stevens RD, Wenner BR, et al. Metabolomics applied to diabetes research: moving from information to knowledge. *Diabetes* 2009;58(11):2429–2443. [PubMed: 19875619]
9. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. *Mass Spectrom Rev* 2007;26(1):51–78. [PubMed: 16921475]
10. Idle JR, Gonzalez FJ. Metabolomics. *Cell Metab* 2007;6(5):348–351. [PubMed: 17983580]
11. Issaq HJ, Van QN, Waybright TJ, Muschik GM, Veenstra TD. Analytical and statistical approaches to metabolomics research. *J Sep Sci* 2009;32(13):2183–2199. [PubMed: 19569098]
12. Lawton KA, Berger A, Mitchell M, et al. Analysis of the adult human plasma metabolome. *Pharmacogenomics* 2008;9(4):383–397. [PubMed: 18384253]
13. Lewis GD, Asnani A, Gerszten RE. Application of metabolomics to cardiovascular biomarker and pathway discovery. *J Am Coll Cardiol* 2008;52(2):117–123. [PubMed: 18598890]
14. Wishart DS. Metabolomics: the principles and potential applications to transplantation. *Am J Transplant* 2005;5(12):2814–2820. [PubMed: 16302993]
15. Slupsky CM, Rankin KN, Wagner J, et al. Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles. *Anal Chem* 2007;79(18):6995–7004. [PubMed: 17702530]
16. Want EJ, Nordstrom A, Morita H, Siuzdak G. From exogenous to endogenous: the inevitable imprint of mass spectrometry in metabolomics. *J Proteome Res* 2007;6(2):459–468. [PubMed: 17269703]
17. Nicholson JK, Wilson ID. Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism. *Nat Rev Drug Discov* 2003;2(8):668–676. [PubMed: 12904817]
18. Pan Z, Raftery D. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Anal Bioanal Chem* 2007;387(2):525–527. [PubMed: 16955259]
19. Humston EM, Zhang Y, Brabeck GF, McShea A, Synovec RE. Development of a GC × GC-TOFMS method using SPME to determine volatile compounds in cacao beans. *J Sep Sci* 2009;32(13):2289–2295. [PubMed: 19569109]
20. Grandori R, Santambrogio C, Brocca S, Invernizzi G, Lotti M. Electrospray-ionization mass spectrometry as a tool for fast screening of protein structural properties. *Biotechnol J* 2009;4(1):73–87. [PubMed: 19156745]
21. Ho C, Lam C, Chan M, et al. Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin Biochem Rev* 2003;24(1):3–12. [PubMed: 18568044]
22. Tolstikov VV, Fiehn O. Analysis of highly polar compounds of plant origin: combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry. *Anal Biochem* 2002;301(2):298–307. [PubMed: 11814300]
23. Edwards JL, Kennedy RT. Metabolomic analysis of eukaryotic tissue and prokaryotes using negative mode MALDI time-of-flight mass spectrometry. *Anal Chem* 2005;77(7):2201–2209. [PubMed: 15801754]
24. Dally JE, Gorniak J, Bowie R, Bentzley CM. Quantitation of underivatized free amino acids in mammalian cell culture media using matrix assisted laser desorption ionization time-of-flight mass spectrometry. *Anal Chem* 2003;75(19):5046–5053. [PubMed: 14708777]

25. Sturm M, Bertsch A, Gropl C, et al. OpenMS - an open-source software framework for mass spectrometry. *BMC Bioinformatics* 2008;9:163. [PubMed: 18366760]
26. Lange E, Tautenhahn R, Neumann S, Gropl C. Critical assessment of alignment procedures for LC-MS proteomics and metabolomics measurements. *BMC Bioinformatics* 2008;9:375. [PubMed: 18793413]
27. Kim K, Aronov P, Zakharkin SO, et al. Urine metabolomics analysis for kidney cancer detection and biomarker discovery. *Mol Cell Proteomics* 2009;8(3):558–570. [PubMed: 19008263]
28. Fahy E, Subramaniam S, Murphy RC, et al. Update of the LIPID MAPS comprehensive classification system for lipids. *J Lipid Res* 2009;50(Suppl):S9–14. [PubMed: 19098281]
29. Zhang H, Saha J, Byun J, et al. Rosiglitazone reduces renal and plasma markers of oxidative injury and reverses urinary metabolite abnormalities in the amelioration of diabetic nephropathy. *Am J Physiol Renal Physiol* 2008;295(4):F1071–1081. [PubMed: 18667486]
30. Barreto FC, Barreto DV, Liabeuf S, et al. Serum indoxyl sulfate is associated with vascular disease and mortality in chronic kidney disease patients. *Clin J Am Soc Nephrol* 2009;4(10):1551–1558. [PubMed: 19696217]
31. Bell JD, Lee JA, Lee HA, et al. Nuclear magnetic resonance studies of blood plasma and urine from subjects with chronic renal failure: identification of trimethylamine-N-oxide. *Biochim Biophys Acta* 1991;1096(2):101–107. [PubMed: 2001424]
32. Foxall PJ, Mellotte GJ, Bending MR, Lindon JC, Nicholson JK. NMR spectroscopy as a novel approach to the monitoring of renal transplant function. *Kidney Int* 1993;43(1):234–245. [PubMed: 8433564]
33. Le Moyec L, Pruna A, Eugene M, et al. Proton nuclear magnetic resonance spectroscopy of urine and plasma in renal transplantation follow-up. *Nephron* 1993;65(3):433–439. [PubMed: 8289995]
34. Rush D, Somorjai R, Deslauriers R, et al. Subclinical rejection--a potential surrogate marker for chronic rejection--may be diagnosed by protocol biopsy or urine spectroscopy. *Ann Transplant* 2000;5(2):44–49. [PubMed: 11217206]
35. Dedeoglu IO, Feld LG. Decreased urinary excretion of nitric oxide in acute rejection episodes in pediatric renal allograft recipients. *Transplantation* 1996;62(12):1936–1938. [PubMed: 8990391]
36. Lewis GD, Wei R, Liu E, et al. Metabolite profiling of blood from individuals undergoing planned myocardial infarction reveals early markers of myocardial injury. *J Clin Invest* 2008;118(10):3503–3512. [PubMed: 18769631]
37. Sreekumar A, Poisson LM, Rajendiran TM, et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 2009;457(7231):910–914. [PubMed: 19212411]
38. Ferrara CT, Wang P, Neto EC, et al. Genetic networks of liver metabolism revealed by integration of metabolic and transcriptional profiling. *PLoS Genet* 2008;4(3):e1000034. [PubMed: 18369453]
39. Keurentjes JJ, Fu J, de Vos CH, et al. The genetics of plant metabolism. *Nat Genet* 2006;38(7):842–849. [PubMed: 16751770]
40. Shah SH, Hauser ER, Bain JR, et al. High heritability of metabolomic profiles in families burdened with premature cardiovascular disease. *Mol Syst Biol* 2009;5:258. [PubMed: 19357637]
41. Gieger C, Geistlinger L, Altmaier E, et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. *PLoS Genet* 2008;4(11):e1000282. [PubMed: 19043545]

Figure 1. Systems biology of renal disease

The relationship between genome, transcriptome, proteome and metabolome is depicted. The complexity of the dataset increases from genome to transcriptome to proteome. Integration of these datasets provides comprehensive understanding of pathophysiology of renal disease.

Figure 2. Workflow for mass spectrometry based metabolomic analysis

Biological samples are processed and subjected to either gas (GC) or liquid (LC) chromatography separation. The eluent is ionized by one of several modes of ionization such as EI, CI, ESI, APCI and MALDI. Subsequently, the resultant mass spectra are derived from the mass analyzers and further processed by data and statistical analysis and the metabolite of interest is identified. EI, electron impact ionization; CI, chemical ionization; ESI, electrospray ionization; APCI, atmospheric pressure chemical ionization; MALDI, matrix assisted laser desorption ionization; TOF, time of flight; FTMS, Fourier transform mass spectrometry, MS/MS, tandem mass spectrometry; PC, principal component.

Figure 3. Metabolomic profiling of prostate cancer

Z score plots for 626 metabolites in localized prostate cancer and metastatic samples normalized to the mean of the benign prostate samples (Reproduced from³⁷).