Metabolomics in Toxicology: Preclinical and Clinical Applications

Donald G. Robertson,*1 Paul B. Watkins,†‡ and Michael D. Reily*

*Applied and Investigative Metabolomics, Bristol-Myers Squibb Co. Princeton, New Jersey 08543; †Hamner-University of North Carolina Institute for Drug Safety Sciences, Research Triangle Park, North Carolina, 27709; and ‡Department of Medicine, University of North Carolina, Chapel Hill, Chapel Hill, North Carolina 27599

1To whom correspondence should be addressed at Applied and Investigative Metabolomics, Bristol-Myers Squibb Co., Route 206 & Province Line Road, Princeton, NJ 08543. Fax: (609) 252-7156. E-mail: don.robertson@bms.com.

Received October 21, 2010; accepted November 19, 2010

Key Words: metabolomics; metabonomics; metabolic profiling; biomarkers; mechanisms of toxicity.

HISTORICAL PERSPECTIVE

Metabolomics or Metabonomics: What is in a Name?

Metabolomics first burst into the toxicology arena approximately 10–12 years ago, experiencing rapid growth as a science in the intervening period (Fig. 1). Though metabolomics still trails the publication rates of its sister technologies, genomics and proteomics, it is closing the gap in terms of both the numbers of publications and the quality of those publications. One continuing source of confusion is differentiating the terms metabolomics, metabonomics, and metabolic profiling. Although metabonomics was the first term formally defined (Nicholson et al., 1999), the term metabolomics came into usage shortly thereafter (Fiehn, 2002). Metabolic profiling is used by some as a generic term to avoid confusion of the two aforementioned “omic” terms, but that phrase itself can be confused with the comprehensive metabolite analysis of xenobiotics, so it is debatable whether its use adds any clarity to the situation. Whatever the case, it is now clear that metabolomics is the term preferred by most practitioners (Fiehn, 2002). Metabolic profiling is used by some as a generic term to avoid confusion of the two aforementioned “omic” terms, but that phrase itself can be confused with the comprehensive metabolite analysis of xenobiotics, so it is debatable whether its use adds any clarity to the situation. Whatever the case, it is now clear that metabolomics is the term preferred by most practitioners (Fig. 1). Therefore, metabolomics will be used in this review with the understanding that it represents all three terms. For most people, the names can be used interchangeably, but the reader is still advised to utilize all three terms when performing literature searches.

Now that we have determined what to call it, we need to define what it is. The literature is replete with fine-tuned definitions, but the most succinct definition still appears to be most appropriate, and that is that metabolomics is “the comprehensive and quantitative analysis of all metabolites” (Fiehn, 2001). Every word in that phrase could be subject to debate as no analytical technique can measure “all” metabolites and what exactly qualifies as a metabolite? Do dietary-derived metabolites qualify, what about gut flora–derived metabolites? Rather than wade into this academic argument, we will leave it to the reader to define metabolomics as he or she chooses. For the toxicologist, the important point is that a metabolomics approach has the potential to reveal novel biochemical sequelae of toxicant administration that can lead to mechanistic insights and identification of biomarkers of cause and/or effect. Furthermore, the technology has the potential to characterize models or disease states to provide biochemical bases for observed interactions with xenobiotics.

Given the explosion of recent literature, this review will not revisit the origins of metabolomics to any extent but will focus on recent advances in the field. Although botanical, environmental, and nutritional applications of the technology have expanded with the field, there is now such a wealth of literature this review will not be able to cover those burgeoning areas of metabolomic research. After providing some historical perspective, this review focuses on the biomedical applications of metabolomic technology with emphasis on new analytical approaches, preclinical, and clinical applications of the technology, particularly those areas relevant to practicing toxicologists.

Repositioning the Field

Toxicology played a significant role in the early development of the technology (Robertson, 2005); yet, the relative contribution of toxicology applications (in broadest definition of that term) to the field has been decreasing steadily since 2000 (Fig. 2). Rather than signaling an apparent decline in interest by toxicologists, this trend simply reflects the fact that...
the expansion of the technology in nontoxicology-related endeavors has far exceeded the rate of expansion for toxicology applications. Additionally, although studies profiling metabolic responses to various toxins elicited a lot of interest by toxicologists early on, the field has moved beyond these descriptive studies to more focused mechanistic and biomarker identification/validation studies. These studies tend to be more complex and difficult to conduct and interpret. The result has been fewer, though in our opinion, more important publications. This said, the continuing importance of the technology in the field is emphasized by the fact that in 2009, the relative rate of toxicology-associated metabolomic publications exceeded the relative rate of toxicology-associated genomics or proteomic publications by two- to threefold (Fig. 2).

A Change of Players

A tidal change in metabolomic technology in toxicology over the past several years has been the apparent shift in the user base from industry (including contract research organizations, instrument vendors, and software developers) to academic, governmental, and medical organizations. An examination of first author affiliations of toxicology-associated metabolomics publications over time reveals that approximately 46% (12/26) of first author affiliations were industrial in 2002, 36% (24/67) in 2005, and 13% (17/136) in 2009, representing decreases in both relative and absolute terms. What does this mean? First, it must be considered that publication by industry scientists is frequently hampered by intellectual property concerns so that the publication rate is not absolutely reflective of industry research, as much industrial research remains unpublished. Still, assuming that barriers to publication by industrial scientists in 2009 are similar to what they were in 2002, the drop in publication rate is disconcerting. There are probably several reasons for this, but economic considerations are probably the main cause. The last several years have been extremely trying for industry with downsizing and outsourcing much more prevalent than in the first half of the decade. Although metabolomics groups were seen as “nice to have” within industrial toxicology departments, they were not considered by many to be “need to have”; therefore, when budget or headcount cuts were mandated, many dedicated metabolomics groups were reconfigured, down sized, or eliminated entirely. While trying not to sound too self-serving, we believe this to be a strategic mistake as will hopefully become clear in this review. In addition, early applications of metabolomics to toxicological problems focused on enhanced descriptions of biochemical sequelae to administration of xenobiotics. Although that was a useful exercise, it did little to aid decision making processes within toxicology departments—the real “bottom line” for any new technology in an industrial setting. With new metabolomic research identifying specific mechanisms and novel biomarkers in both the preclinical and clinical settings, it is evident that metabolomics is moving beyond descriptive science into the arena of a value-added activity.

One thing that has not changed in recent years is that the majority of toxicology-associated metabolomic work is
conducted outside the United States. This was evidenced by 69% (94/136) of first author affiliations being based outside the United States (primarily Europe, China, and Japan) in 2009 as compared with 61% (41/67) in 2005.

ANALYTICAL TRENDS

Measuring the Metabolome

The concept of the metabolome as the “total complement of metabolites in a cell” (Tweeddale et al., 1998) has since been broadened to include not only endogenous small molecules but also those introduced and modified by diet, medication, environmental exposure, and coexisting organisms (Dunn, 2008). From an analytical (and pharmaceutical) perspective, this broader definition of the metabolome is the most useful and most closely fits the concept under which the term “metabolomics” was defined (Nicholson et al., 1999). Regardless of the exact definition, the analytical goal of metabolomics is to achieve a comprehensive measurement of the metabolome and how it changes in response to stressors, with the biological payoff being illumination of the relationship between the perturbations and effected biochemical pathways. The systematic practice of correlating analytical observables in biofluids to human health dates back to Hippocrates (Kouba et al., 2007), and the first truly modern vestiges of metabolomics date back nearly 40 years to Linus Pauling and his idea of “orthomolecular medicine” (Pauling et al., 1971). Hippocrates had the instruments of his own senses (sight, smell, and taste), which were superseded in Pauling’s era by high-resolution gas chromatography (GC) systems coupled with various types of detectors, including mass spectrometers (Jellum et al., 1971). Modern metabolomics has benefited and contributed to the development of ever more sophisticated instruments that can measure hundreds of components simultaneously along with the methods for analyzing and quantifying them.

Metabolomics Approaches

Within the field of metabolomics, there is a continuum of approaches that range from more spectroscopy-centric fingerprinting to nontargeted profiling to targeted analysis. A common denominator among them is a high-information content data set that directly reflects the biochemical state of an organism’s metabolome at a given point in time.

Metabolomic fingerprinting is typically conducted on minimally prepared peripheral biofluids, and sample differentiation is based on spectroscopic or chromatographic data. These “analytical fingerprints” can contain thousands of individual data points that relate to the composition of the sample, which may or may not be annotated as to their molecular origin. Such a data set can be used as input into multivariate statistical tools such as principal component analysis (PCA), described more below. The advantage of this approach is that it avoids the frequently rate-limiting step of data annotation and therefore may be rapidly applied for establishing differentiation between individuals or groups, with applications in in vivo screening of similar drug candidates (Dieterle et al., 2006; Robertson et al., 2007; Robosky et al., 2002), prescreening animals prior to extensive toxicological evaluations, and building of toxicological classification models (Lindon et al., 2005).

Nontargeted metabolomics goes a step beyond fingerprinting and seeks to assign as many individual chromatographic or spectroscopic peaks as possible (Vinayavekhin and Saghatelian, 2010). This approach produces a comprehensive list of metabolites with absolute or relative quantification and has been referred to as metabolomics (Fiehn et al., 2001) or metabolic profiling (Brown et al., 2005). The changes in components can be mapped to specific pathways and provide biomarkers and/or mechanistic information. It should be clear that nontargeted metabolomics always provides an analytical fingerprint, but the inverse is true only if the analytical data are readily annotated. Vibrational spectroscopies such as infrared (IR) can provide a highly reproducible and detailed spectrum that reflects the molecular composition of a complex mixture, but assignment of IR bands to individual molecular species is not normally possible. This makes IR a good metabolomic fingerprinting method but not useful in a nontargeted approach. The requirement of unambiguous annotation essentially limits the nontargeted metabolomic analytical techniques to nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS)–based approaches.

Targeted metabolomics differs from those discussed above in that measured analytes have been selected a priori, usually to address certain specific biological questions within a study. Because the exact analytes are known, it is possible and desirable to measure their absolute concentrations with appropriate use of internal standards. The lines between targeted metabolomics and traditional multiplexed assay development are arguably ill defined but usually center on the number of metabolites measured simultaneously. For example, an assay developed to measure serum acetate is not targeted metabolomics, but one to measure 13 or 14 organic acids in a urine sample probably qualifies. Another distinction between targeted metabolomics and traditional multianalyte assays is that the former should preserve the ability to also identify things that are not necessarily on the predefined list. Often, targeted approaches can be developed for specific compound classes such as lipids (Wolf and Quinn, 2008), bile acids (Want et al., 2010a), or amino acids (Kimura et al., 2009; Mayboroda et al., 2007) that are pertinent to a particular study. Targeted metabolomic assays often result from the desire to more accurately quantify specific sets of molecules observed in an nontargeted study.

In summary, the concepts discussed above actually represent a continuum of approaches that go from largely qualitative fingerprinting to more quantitative targeted approaches. Most metabolomic studies seek to provide a little from all three concepts—a rapid qualitative read on what the data are
indicating combined with as much quantitative information as possible while ideally preserving the ability to discover the unexpected.

Analytical Instrumentation

The ideal characteristics of analytical data that will be most useful in the approaches described above include information content, reproducibility, and amenability of the data to facile annotation. Although there are many analytical techniques that meet one or two of these requirements, only two approaches stand out as meeting all three: MS and NMR. The vast majority of metabolomic publications in the past decade use one or both of these techniques, which are highly complementary (Robertson et al., 2007). It is interesting to note that the majority of publications utilize one method, probably reflecting the analytical strengths of the reporting laboratory. In recent years, the rapid growth of MS-based publications has been noteworthy (Fig. 3). On the other hand, the growth of reports that utilize both platforms has been keeping pace with the general trend, and it is difficult to argue against the improved confidence in annotation and breath of coverage that a multi-platform approach affords (Gowda et al., 2009; Lenz and Wilson, 2007; Pan and Raferty, 2007).

The earliest literature surrounding the use of metabolomics for toxicity assessment was based on NMR spectroscopy of biofluids that began in the early 1980’s and has been extensively summarized (Ala-Korpela, 2008; Daykin and Wulfert, 2006; Lindon et al., 1999; Maher et al., 2009; Reily and Lindon, 2005). Although there continues to be significant developments in the field of NMR and its application to metabolomics, one could argue that it has reached a certain level of maturity. The reason that NMR will continue to be a useful metabolomics tool lies in its unique ability to report quantitatively on the chemical and dynamic environment of individual atoms within an isolated molecule by itself or in the presence of other molecules in solution with a linear dynamic range of \(10^5\). Thus, NMR can be used as a selective and quantitative tool, which is suited for a key analytical requirement of toxicological metabolomics: the quantitative measurement of many molecules in a complex mixture. The interlaboratory and longitudinal reproducibility of the method has also been well established (Keun et al., 2002b). NMR’s principal liability is its low sensitivity, and as a result, the majority of developments in NMR hardware have focused on improving this parameter. Developments such as ultralow-temperature detection circuits (so-called cryoprobes) (Keun et al., 2002a; Robosky et al., 2007) and low-volume probes (Martin, 2005; Mukhopadhyay, 2007) have pushed the limits of detection for modern NMR spectrometers into the \(10^{-5}\)M range, requiring only low microgram quantities of typical endogenous metabolites. Also because of the inherent sensitivity problem, the bulk of NMR work in the field relies on observing the most ubiquitous and sensitive nucleus, the proton. Fortunately, the proton lies at the heart of every hydrogen atom and thus almost every molecule of interest is festooned with numerous NMR spectroscopic probes for exploitation. Although NMR-based metabolomic research has and will focus on proton spectroscopy, recent trends in the NMR metabolomics literature have exploited other nuclides, such as \(^{13}\)C at natural abundance (Shaykhutdinov et al., 2009) or isotope-enriched (Fan et al., 2009), \(^{31}\)P (DeSilva et al., 2009), and multidimensional methods (Ludwig and Viant, 2010; Van et al., 2008), and such nonhydrogen NMR work provides exciting prospects for the future.

Unlike proton NMR, MS is a collection of very diverse platforms and methodologies, which have significant potential as metabolic tools, many of which have been reviewed (Koal and Deigner; Lu et al., 2008b; Scalbert et al., 2009; Villas-Boas et al., 2005; Want et al., 2007). Broadly, MS-based metabolomics fall into two categories: those that rely on physical separation of complex mixtures prior to mass analysis (hyphenated methods, such as liquid chromatography [LC]-MS or GC-MS) and those that measure mixtures directly (direct injection [DI] methods). Perhaps, the most significant development in the past 5+ years that affects almost all areas of MS research is the increasing availability of high-resolution accurate mass (HRAM) detectors. These enable the measurement of mass-to-charge ratio of ions with low ppm accuracy; enough to narrow down the number of possible atomic compositions of an ion and in some cases assign it a unique molecular formula, greatly enhancing the ability to annotate peaks (Hu et al., 2005). This is particularly useful in untargeted approaches where the need to identify unknowns that change can greatly benefit from this additional information. Many nominal mass (non-HRAM) MS-based techniques lend themselves to a targeted approach. Multiple reaction monitoring (MRM) using tandem quadrupole MS allows highly selective monitoring of precursor ions and product ions of each.
metabolite and enables high sensitivity over a broad dynamic range without the need for isotope-labeled standards. The MRM approach is preferred when a limited number of analytes (< 50) are to be measured, whereas a full scan, high-resolving power HRAM system has advantages with larger target analyte lists (Lu et al., 2008a).

From a data generation standpoint, the biggest challenge in MS of complex mixtures is the simultaneous ionization and detection of multiple analytes. Placing a charge on a molecule is prerequisite to MS measurement, and for most isolated analytes, this is usually readily achievable. This must be done with enough energy to induce ionization, but not enough to destroy the molecule, thereby preserving the molecular ion (M) or its +/− hydrogen form (M + H/M − H). This is most commonly accomplished using electrospray ionization (ESI), wherein a charge is applied to a droplet and transferred to the analytes contained therein as the droplet evaporates under vacuum, or by atmospheric chemical ionization, where charged reagent molecules combine with analytes through collision to confer charge. Thus, when trying to apply a charge to many components at once, competition for the applied energy can favor one or a few particular components, and difficult-to-ionize analytes may go completely undetected. The most common way to minimize this ion suppression phenomenon is to physically separate the mixture prior to ionization. The most recent developments in chromatography-hyphenated MS methods have been the improvements of LC separation technologies, the rediscovery of GC-MS, and the development of capillary electrophoresis (CE) as a front-end separation platform (Garcia et al., 2008) and chemical ionization approaches combined with HRAM detectors that, unlike EI, can preserve the intact molecular information (Carrasco-Pancorbo et al., 2009). One drawback to GC-MS is the need to render all analytes more volatile, generally though derivatization chemistry, introducing one additional source of analytical variability that one must account for.

It is true that the utility of the MS is greatly enhanced by predetector separation of individual analytes, but the chromatography itself introduces additional analytical variability. Realizing this, approaches have been developed that attempt to avoid chromatography by introducing the sample into the mass spectrometer via DI (Beckmann et al., 2008; Hansen and Smedsgaard, 2007). An example of a metabolomic fingerprinting approach is described in the study of yeast cell extracts for the investigation of cadmium toxicity utilizing DI-MS that revealed that about 400 out of several thousand measured ions correlated with increased glutathione (GSH) synthesis (Madalinski et al., 2008). Although this approach is simple and rapid, the detrimental effects of salts and charge competition and the inability to distinguish isobaric compounds need to be considered. Miniaturization partially mitigates these problems, and increased salt tolerance has been demonstrated with chip-based nanoelectrospray compared with regular flow electrospray (Boernsen et al., 2005). A relatively new method that avoids chromatography and requires virtually no sample preparation is desorption electrospray ionization (DESI), which ionizes samples under open atmospheric conditions using a stream of ionizing gas to desorb, ionize, and transfer the analytes into the MS instrument. DESI data from urine samples blotted directly onto a paper strip were able to differentiate several different inborn errors of metabolism (Pan et al., 2007). Indeed, most MS instrument manufacturers have begun offering higher pressure LC systems to complement their MS detectors in the past few years.

Long a favorite method for analyzing mixtures in environmental and plant science, GC is making a resurgence into biomedical metabolomics applications (Fiehn, 2008; Pasikanti et al., 2008a,b). Unlike most LC-MS approaches that seek to preserve the molecular ion, the majority of GC-MS studies rely on electron impact (EI) ionization, which breaks apart each analyte into many ionized fragments and produces a diagnostic “fragmentation pattern” which can be compared with the patterns from known metabolites for identification purposes. Although this is ideal for some applications, it also presents challenges when analyzing complex mixtures containing molecules that have very similar fragmentation, such as may be encountered in toxicological metabolomics. New approaches that differentiate GC-MS-based metabolomics from its more traditional roles include the development of more comprehensive mammalian-centric EI fragmentation libraries (Kind et al., 2009; Kopka et al., 2005; Tautenhahn et al., 2008) and chemical ionization approaches combined with HRAM detectors that, unlike EI, can preserve the intact molecular information (Carrasco-Pancorbo et al., 2009). One drawback to GC-MS is the need to render all analytes more volatile, generally though derivatization chemistry, introducing one additional source of analytical variability that one must account for.

Data Evaluation and Interpretation

Once the metabolomic data are collected, the data have to be processed. Each analytical approach and platform obviously has specific requirements that can only be briefly touched on in this review. Fingerprinting methods generally rely on the application of multivariate statistical analysis to raw or minimally reduced data. In the case of NMR, this has traditionally been done by parsing the spectra into frequency
bins, or buckets, that are small enough to contain a limited number of metabolites and large enough to smooth out the sample-to-sample variation caused by small differences in sample pH (Holmes et al., 1994). In practice, bins of 0.01–0.04 ppm meet these criteria. These are then normalized and integrated, thereby reducing each 16K–32K data point NMR spectrum by a factor of ~10^2. These so-called “bucketed” spectra each contain 200–500 data points, with each point representing the sum of contributions of all the measured nuclei within each frequency bucket. Many such spectra can then be subjected to various types of supervised and unsupervised multivariate analyses (Holmes and Antti, 2002; Holmes et al., 1994; Robertson et al., 2005, 2007). New spectral alignment approaches are rapidly making binning passe, and utilizing the entire NMR spectrum for multivariate analysis is becoming established (Staab et al.; Stoyanova et al., 2004). For hyphenated MS-based fingerprinting, data features represented as retention time (RT) and mass/charge (m/z) pairs, along with measured intensity, have similarly been used as multivariate input. Normalization strategies are required to account for biological variability and depend on the method and matrix. For blood fluids and tissues, volume or weight normalization generally suffices because these matrices tend to be homeostatic. On the other hand, urine concentration depends on factors such as food and water intake and diurnal variation, and normalizing to the total signal of all endogenous components is the norm (Warrack et al., 2009). In the case of proton NMR, this measure can be thought of as a surrogate for the concentration of total dissolved organic content because most polar organic compounds contain at least one hydrogen. In our experience, urinary total NMR signal correlates strongly with urinary creatinine concentration, the standard normalizing factor for most urine clinical chemistry tests, unless there are perturbations to the urine creatinine itself. If complete 24-h collections are made, one can, in principle, use urine volume to normalize the data. Although these chemometric procedures are invaluable for identification of outliers, classification, and statistical evaluation of buckets (which can later be annotated), care must be taken to avoid misinterpretation and overfitting of the data (Robertson et al., 2007).

As useful as pattern recognition and multivariate analysis are, when the dust settles, the ultimate users of metabolomics results are biologists. Experience has shown that when data analysis ends with colorful PCA or partial least squares plots, the real impact of a metabolomics study is not realized. Ultimately, it is not sufficient to tell the customer that you can elegantly differentiate dosed animals from controls or females from males—the biologist or toxicologist already knows that! The real value of a metabolomics study is a detailed knowledge of what molecules are changing and by how much. Indeed, multivariate analysis can help you get to this information, but it is not an end unto itself. On the surface, the task of getting from raw LC-MS, GC-MS, or NMR data to a quantitative metabolite table seems straightforward, but there are numerous obstacles that are encountered depending on the platform. These barriers include annotation of the data (assigning the peaks to specific molecules) and quantitation of the analytes.

Because NMR integrals are directly proportional to the concentration of the corresponding nucleus (atom), the process of preparing a quantitation table is relatively straightforward. One simply needs to find an isolated peak that represents a nucleus in an analyte and integrate it to obtain a relative concentration. Comparison with the integral of an external or internal standard of known concentration yields absolute concentration. Unfortunately, it is rare to find a peak that is not overlapping with another component in a typical biofluid NMR spectrum (Fig. 4). Thus, one of the biggest problems facing NMR-based metabolomics is one of deconvolution of overlapping peaks. Recent work that promises to address these issues includes peak fitting to library spectra (Wishart, 2008), advanced mathematical deconvolution analysis (Rubtsov et al., 2010), two dimensional (2D) NMR (Van et al., 2008; Xia et al., 2008), pseudo 2D NMR (Ludwig and Viant, 2010; Ludwig et al., 2009; Parsons et al., 2009), statistical correlations (Sands et al., 2009), or a combination of these methods. Annotation of the spectral peaks is normally done by comparing with authentic spectra and ultimately by co-addition of authentic standards. Several commercial (SBASE, Bruker Biospin, Inc.; NMR Suite, Chenomx Inc.) and publicly available databases (Cui et al., 2008; Wishart et al., 2009) exist that contain many 1D and 2D spectra of endogenous metabolites.

For MS, the issues are a bit different. Because modern spectrometers allow selective filtering of the data through very high-resolution mass measurement, when combined with high-resolution chromatography, overlap can be minimized (although not entirely eliminated). Nonetheless, quantification of MS data suffers from two major problems. First, through isotopic distribution, molecular fragmentation, and adduct formation, each molecule produces many ions, even under the most gentle ionization conditions. As a result, a typical ESI LC-MS data set from a biofluid sample may contain tens of thousands of peaks (millions if one considers noise) that represent only hundreds of compounds. Thus, an important theme in the application of MS to metabolomics is to relate these individual ions to molecular species (Fig. 5). Various software approaches have been proposed for this massive data reduction problem, but no one solution has emerged as a best practice (Katajamaa and Oresic, 2007; Reilly et al., 2011). Second, unlike the quantum mechanical transitions measured in NMR, the ionization of an analyte in a mass spectrometer is highly dependent on the nature of the molecule itself and the other analytes that might simultaneously be competing for ionization energy, as discussed above. Hence, absolute quantitation is only possible with calibration against multiple concentrations of authentic standards in the relevant matrix—a tedious and time-consuming process. Thus, such analyses are rarely conducted in a nontargeted fashion. More commonly, quantitation is relegated to subsequent targeted analysis. For
nontargeted analysis, the assumption can be made that a given analyte has a similar ionization response across many samples, and therefore, relative changes can be determined. However, even in that instance, it must first be determined that detector saturation is not occurring \cite{Reily2011}. Highly multiplexed DI-MS methods that utilize SIL standards of representative compounds represent a compromise between rigorous quantitation and high throughput \cite{Altmaier2008}. One potential approach that has not been extensively tested would be to utilize an auxiliary detector that has a more predictable concentration response such as UV or flame ionization for LC-MS and GC-MS, respectively. As with NMR, annotation relies on comparing RTs and/or m/z values with those of known standards run under identical conditions. For GC-MS, the US National Institute of Science and Technology database \cite{Anonymous2010a} has been the gold standard for decades and has recently been augmented by more metabolomics-relevant incarnations \cite{Hummel2007,Kind2009}. The METLIN database \cite{Sana2008,Smith2005} is the most extensive public repository for LC-MS data and includes fragmentation libraries for many molecules measured under standard conditions. Although the applicability of this library as conditions and instrumentation stray from those used to build it has not been thoroughly tested, its establishment is an important scientific achievement, and the proliferation of useful commonly available libraries is an important goal for this field.

**PRECLINICAL APPLICATIONS**

**Model Characterization**

Routinely, metabolomics evaluation is nontargeted enabling users to gain a comprehensive evaluation of the systemic response of the subject (preclinical or clinical) to pathophysiological stimuli or genetic modification \cite{Nicholson1999}. In other words, it represents a powerful tool for evaluation of systems biology \cite{Goodacre2004}. Metabolomics has been used to demonstrate biochemical sequelae of strain differences \cite{Gavaghan2000,Plumb2003}, gender differences \cite{Hodson2007,Kochhar2006}, diurnal variation \cite{Bollard2001,Minami2009,Slupsky2007}, and age \cite{Gu2009,Robertson2000,Williams2005}. This ability to discern the biochemical underpinnings of many common animal model variables \cite{Bollard2005} demonstrates the power of the technology to characterize animal models \cite{Griffin2004,Griffin2006}. Phenotyping of genetically modified mouse models is a powerful tool for discovering and understanding the downstream ramifications of target gene (mal)function \cite{Rull2009}. Over the past 5 years, metabolomic evaluations of these mouse models have frequently been reported (Table 1). Beyond genetically modified models, other in vivo models investigated using metabolomic technology include various dietary models of...
various aspects of metabolic syndrome in rats (An et al., 2004; Kim et al., 2009b; Koves et al., 2008; Li et al., 2008, 2010a), mice (Connor et al., 2010; Dumas et al., 2006; Kim et al., 2005; Toye et al., 2007), and primates (Cox et al., 2009). Metabolomics evaluations in lactating (Rudolph et al., 2007) and developing mouse models (Shimizu et al., 2009) have also been reported.

The ability to discern global biochemical responses to single-gene alterations or focused physiological or pharmacological interventions is one of the great strengths of metabolomic technology, and it forces practitioners to see all the systemic ramifications of experimental design and model manipulation, some of which they might later wish they never started to investigate (Robertson, 2008). For example, several years ago while in the process of running a number of metabolomic studies in rats, we noted that control rats from several studies had remarkably altered urine metabolic spectra compared with the majority of other studies we had conducted (Fig. 6). Further investigation revealed that these unusual samples came from a single room (colony) of rats at the Charles River Raleigh facility, whereas animals in the room right next to it had a phenotype consistent with our historical controls (Robosky et al., 2005, 2006). We were able to attribute the cause of this phenotypic difference to altered gut flora most probably because of a failure to acquire a normal gut flora because of stringent animal room hygiene and the fact that new rat colonies are typically originated from a foundation colony with limited gut flora (Rohde et al., 2007). Since that time, we have had at least two other instances (as yet unpublished) where similar findings were noted in control C57BL/6 mice suggesting this phenomenon may be more common than might be realized.

The above examples highlight that a key area of metabolomic research over the past 10 years has been with regard to the role gut microflora plays with systemic biochemistry (Calvani et al., 2010; Dumas et al., 2006; Fava et al., 2006; Martin et al., 2007; Nicholson, 2008; Nicholson and Wilson, 2003; Nicholson et al., 2005; Wikoff et al., 2009). One interesting study combining both gut flora research and model evaluation using metabolomics was reported by Martin et al. (2008) in which a metabolomic evaluation of the effects of probiotic treatment of germ-free mice colonized with human gut flora was monitored. The probiotics altered microflora content with subsequent effects on lipid metabolism, lowered plasma lipoproteins, and stimulated glycolysis as well as disrupting a number of other biochemical pathways. These results demonstrate the key role gut flora can play in various central metabolic processes and also highlight the fact that the interaction between gut flora and toxic responses is seldom evaluated and poorly understood by practicing toxicologists. A key advantage of nontargeted metabolomics is that it forces the scientific practitioner to look at all aspects of metabolic response, not just those of preselected target organs. It is this facet of the technology that has revealed that we have been
missing a key “target organ” evaluation—that of the intestinal flora which, in humans, weigh approximately 1.5 kg and contain a genome with 100-fold more genes than our own (Kinross et al., 2008). In the past, short of frank gastrointestinal effects (e.g., stool changes or histopathology), evaluating effects on gut flora was a difficult task. Metabolomic analyses simplify that task to the point where gut flora as a potential target should no longer be routinely ignored in toxicology studies.

### Toxicity Screening

On of the earliest promises of metabolomic technology was with regard to the potential for rapid throughput in vivo toxicity screening particularly within the pharmaceutical industry (Harrigan and Yates, 2006; Robertson et al., 2000, 2002). It was speculated that relatively rapid analytical assessment of biofluids followed by extensive chemometric analysis of the spectra would reveal both mechanisms and biomarkers of toxicity. Furthermore, the nature of the chemometric analyses would obviate the necessity for the time-consuming process of peak (metabolite) annotation. One of the goals for the COMET I consortium was to generate predictive toxicology models based on over 100 in vivo studies in both rats and mice (Lindon et al., 2003, 2005). Although the consortium did deliver predictive models (Ebbels et al., 2007), these models have gained little traction in the industry as it was unclear what exactly was being modeled. Was it cause (mechanism) or effect (both direct and indirect)? For example, metabolomic analysis of urine or serum from animals given a compound inducing a hepatic toxicity because of oxidative stress may reveal metabolites directly indicative of that stress (e.g., changes in the reduced and oxidized glutathione ratio), the resultant hepatic pathology (indicating hepatic dysfunction), metabolites associated indirectly with the toxicity (e.g., weight loss), or a metabolic profile reflective of some combination of these effects. Furthermore, the metabolite profile may reflect effects on minor or unrecognized target organs that nevertheless produce disproportionate metabolic derangements. In recent years, it has become clear that not only is metabolite annotation critical for proper toxicological evaluation but also a background in pathway biochemistry is necessary to gain a thorough understanding of the role various endogenous metabolites play.

Recent studies conducted in the Bristol-Myers Squibb laboratories demonstrate the power of metabolomics in early in vivo safety screening (prior to good laboratory practice toxicology evaluation). Some shrewd metabolomic detective work on samples from early in vivo studies led to the identification of ascorbate and gulonic acids as preclinical markers of microsomal induction (Aranibar et al., 2009). Since that time, several screening studies have shown increases in both acids indicative of microsomal induction. Although it could be argued that the fact that the compounds in question were inducers would have eventually been identified by traditional methods, we first noted it in vivo using noninvasive metabolomic technology. The work further provides a nice example of an integrated metabolomic/transcriptomic approach to biomarker discovery. A transcriptomic and metabolomic analysis was conducted on samples collected from rats treated with various known inducers (Fig. 7). Increased urine ascorbic and gulonic acids were noted after treatment with phenobarbital, diallyl sulfide, and DMP904 (another microsomal inducer; Wong et al., 2005) but not β-naphthoflavone. Importantly, transcript analysis of treatments increasing urinary gulonic and ascorbic acids revealed upregulated hepatic aldo-keto reductase 1A1 (glucuronate reductase)—an enzyme that catalyzes the formation of gulonic acid from glucurionate, with concurrent decreased expression of both regucalcin (gulonolactonase), the enzyme responsible for conversion of gulonic acid to gulono-1, 4-lactone, and gulonolactone oxidase, the rate-limiting enzyme in ascorbate biosynthesis. Furthermore, increased hepatic expression of enzymes regulating ascorbic acid reutilization including glutaredoxin reductase and thioredoxin reductase was noted. These data demonstrate the use of metabolomic data to aid in transcriptomic interpretation of subtle changes that might have been missed or dismissed if analyzed in isolation.

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Model</th>
<th>Disease being modeled</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJD1</td>
<td>Spinocerebellar ataxia</td>
<td>Griffin et al. (2004)</td>
</tr>
<tr>
<td>PKb−−</td>
<td>Ischemia-reperfusion injury</td>
<td>May et al. (2004)</td>
</tr>
<tr>
<td>Cln3</td>
<td>Batten disease</td>
<td>Pears et al. (2005)</td>
</tr>
<tr>
<td>R6/2</td>
<td>Huntington’s disease</td>
<td>Tsang et al. (2006)</td>
</tr>
<tr>
<td>peroxisome</td>
<td>Metabolic syndrome</td>
<td>Atherton et al. (2006), Makowski et al. (2009), and Zhen et al. (2007)</td>
</tr>
<tr>
<td>PPARα−−</td>
<td>Huntington’s disease</td>
<td>Underwood et al. (2006)</td>
</tr>
<tr>
<td>PSCr−/−</td>
<td>Obesity-induced inflammation</td>
<td>Mutch et al. (2007)</td>
</tr>
<tr>
<td>Interleukin 10−/−</td>
<td>Inflammatory bowel disease</td>
<td>Lin et al. (2009), Martin et al. (2009), and Murdoch et al. (2008)</td>
</tr>
<tr>
<td>apoE−/−</td>
<td>Atherosclerosis</td>
<td>Leo and Darrow (2009) and May et al. (2008)</td>
</tr>
<tr>
<td>TRAMP</td>
<td>Prostate cancer</td>
<td>Teichert et al. (2008)</td>
</tr>
<tr>
<td>Apc(Min/+</td>
<td>gastrointestinal tumorigenesis</td>
<td>Backshall et al. (2009)</td>
</tr>
<tr>
<td>PXR−/−</td>
<td>Transactivation</td>
<td>Cho et al. (2009, 2010)</td>
</tr>
<tr>
<td>and Fxr−/−</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRND8</td>
<td>Alzheimer’s disease</td>
<td>Salek et al. (2010)</td>
</tr>
<tr>
<td>LDLr−/−</td>
<td>Fatty liver disease</td>
<td>Vinaixa et al. (2010)</td>
</tr>
<tr>
<td>Sec61q</td>
<td>Diabetes</td>
<td>Lloyd et al. (2010)</td>
</tr>
</tbody>
</table>

*Note.* PPAR, peroxisome proliferator activated receptor; TRAMP, transgenic adenocarcinoma of the mouse prostate.
In another example, a compound under development was demonstrated to produce significant and dose-related increases in rat serum cholesterol by routine clinical pathology assessment. Metabolomics analysis revealed that not only was serum cholesterol increased but several measured phytosterols also increased in nearly an identical fashion. This strongly suggested that the increased cholesterol was because of an effect on sterol absorption at the level of the gut that was not specific to cholesterol. Serum bile acid profiles (which were part of the metabolomic screening analysis) were unaffected, suggesting that the changes in cholesterol were not because of altered bile acid kinetics. Combined, these data brought a level of mechanistic understanding to the finding of increased cholesterol that was not apparent using traditional toxicology endpoints.

It is difficult to assess how many companies may now be utilizing metabolomics in such a fashion in early discovery, though some certainly are (Dieterle et al., 2006). However, in many cases, intellectual property concerns may prevent publication of such examples at least until compounds are discontinued and more likely until the entire chemical series or therapeutic program is discontinued. Although anonymized examples (like those above) demonstrate the point, they are intellectually far from satisfying for the typical toxicologist or typical journal editor.

**In Vitro Applications**

The growth of MS-oriented metabolomic approaches has greatly expanded the potential for utilization of the technology for *in vitro* applications, though NMR approaches to *in vitro* applications are still being pursued (Duarte et al., 2009; Gottschalk et al., 2008). The field was recently reviewed by Cuperlovic-Culf et al. (2010) and is rapidly growing. The
single-cell focus of most in vitro approaches greatly simplifies evaluation of metabolomic data as nonspecific metabolic changes are greatly reduced compared with in vivo assessment of biofluids. Not only can the target cell (or cell extract) be analyzed but the flux of metabolic changes in the media is also a great source of biochemical insight (Nagy et al., 2008). The disadvantage of in vitro approaches is no different for metabolomics than for any other application—relevance to in vivo response.

One technique particularly suited for in vitro applications is the evaluation of single metabolite dynamics. An interesting review by Wheatley (2005) reports on the role of arginine in both normal and tumor cell lines as assessed by metabolomic technology. Assessment of metabolic flux is also much more readily investigated on the single-cell level (Wang and Bodovitz, 2010), and metabolomics is a valuable approach for concurrently monitoring numerous metabolites (Lee et al., 2006; Rabinowitz, 2007).

Beyond pure biochemical applications, in vitro evaluation of various aspects of reproductive biology is one area where metabolomics has made significant inroads over the past 5 years (Assou et al., 2010; Botros et al., 2008; Bromer and Seli, 2008; Revelli et al., 2009; Seli et al., 2008). Within the toxicology community, in vitro metabolomic approaches have been used to assess the toxicity of cell-penetrating peptides in CHO cells (Kilk et al., 2009), the toxicity of cigarette smoke on lung epithelial cells (Vulimiri et al., 2009), the neurotoxicity of mercury and caffeine in primary reaggregating brain cell cultures (van Vliet et al., 2008), and evaluation of developmental toxicity potential in human stem cells (West et al., 2010).

**Mechanisms and Biomarkers**

It can be argued that if you define the mechanism of efficacy or toxicity of a xenobiotic, you can identify an appropriate biomarker. Of course, whether that biomarker is practical or analytically feasible is another question. Theoretically, one can identify a biomarker without understanding the mechanism, but that is far from ideal. Early on in the development of the technology, mechanisms and biomarkers were recognized as key deliverables of metabolomic technology (Bailey and Ulrich, 2004; Colatsky and Sumner, 2003; Dunckley et al., 2005; Holmes and Shockcor, 2000; Kleno et al., 2004; Mortishire-Smith et al., 2004; Robertson, 2005; Roy et al., 2004). Early attempts at biomarker discovery associated observed metabolic disruptions with disease or toxicity and voilà a biomarker was born. The early literature is replete with examples of “the usual suspects” (Robertson, 2005) (e.g., citrate, 2-oxoglutarate, creatine, trimethylamine, etc.) as being biomarkers of just about everything. To be fair, as these molecules do change in many toxicities and disease states, they are indeed biomarkers, at least by some definition. The problem of course is that they are nonspecific enough to be essentially useless as diagnostic biomarkers. Although there may be noble attempts to make these nonspecific biomarkers mechanistically relevant, it is a bit like “six-degrees of molecular separation”—no molecule is separated from any mechanism by any more than six steps on a metabolic pathway chart . . . Despite these growing pains, metabolomics technology has begun to deliver on the promise of biomarkers of disease and toxicity.

Metabolomics has been promulgated as an approach for mechanistic understanding and biomarker discovery for neuroscience applications (Dunckley et al., 2005; Pendyala et al., 2007; Quinones and Kaddurah-Daouk, 2009), oncology (Claudino et al., 2007), cardiovascular disease (Vasan, 2006), tuberculosis (Parida and Kaufmann, 2010), cystic fibrosis (Wetmore et al., 2010), and metabolic syndrome (Oresic, 2010) to name a few applications. Within the toxicology community, metabolomics has not been ignored (Clarke and Haselden, 2008), with the technology proposed as a tool for discovering biomarkers of both renal (Sieber et al., 2009) and hepatic toxicity (Antoine et al., 2009; McBurney et al., 2009).

Beyond these reviews, specific biomarkers have been identified. Select biomarkers attributed to various toxicities, disease states, or physiological changes are summarized in Table 2, with several discussed in more detail below. In an important paper by Soga et al. (2006), a metabolomics investigation utilizing a CE-time of flight-MS approach was employed to evaluate an acetaminophen model of hepatic oxidative stress in the mouse. The authors identified ophthalmic acid as a potential biomarker of oxidative stress associated with decreased GSH levels in both the liver and the serum. Their proposed mechanism is summarized in Figure 8. The synthesis of ophthalmic acid from 2-aminobutyrate uses the same enzymatic machinery used to synthetize GSH from cysteine. The key point is that GSH produces feedback regulation of γ-glutamylcysteine synthase (GCS). As GSH is depleted because of the rigors of oxidative stress, the feedback inhibition of GCS is reduced leading to increased biosynthesis of GSH, and because ophthalmic acid is synthesized using the same machinery, it is also increased. However, unlike GSH, ophthalmic acid is not depleted by oxidative stress leading to its accumulation. Importantly, this work went beyond a simple survey of changed metabolites. Extensive mechanistic work to validate the finding was reported enabling other researchers to follow-up on their conclusions (Kombu et al., 2009).

Though often criticized as a “fishing expedition,” recent work in our laboratory highlights how use of metabolomic technology can lead to the discovery of novel biomarkers (Aranibar et al., 2010). As part of a mechanistic study of myotoxins, a metabolomic analysis was conducted on urine samples from rats treated with up to 1 mg/kg/day cerivastatin for 14 days. Detailed statistical analysis of the resultant NMR spectra revealed two resonance peaks contributing to the PCA separation of samples from control and treated animals. Further analytical work eventually identified these peaks as...
N-acetylated 1- and 3-methylhistidine. Although urinary 3-methylhistidine has been used as a biomarker of muscle protein turnover for many years (Young et al., 1973), neither isomer has been used as a biomarker of myotoxicity in rats. Continuing work in our laboratory suggests that these biomarkers are also useful in serum and are as sensitive and specific as any competing biomarker for statin myopathy with the added advantage that these markers increase with muscle atrophy and decrease with muscle hypertrophy (Waterfield et al., 1995), a diagnostic capacity not currently available with other biomarkers of myotoxicity.

Exciting developments in clinical biomarker discovery utilizing metabolomic technology include the identification of sarcosine as a mechanistically linked biomarker of prostate cancer progression (Sreekumar et al., 2009). Sarcosine was increased in the tissue and urine of patients with prostate cancer and in prostate cancer cell lines. Importantly, knockdown of glycine n-methyl transferase (responsible for sarcosine synthesis from glycine) decreased cancer invasiveness, whereas supplementation with sarcosine induced an invasive phenotype in benign prostate epithelial cells. Elegant work by Newgard et al. (2009) proposed a branched chain amino acid (BCAA) “signature” of obesity and insulin resistance in human subjects. The changes in BCAAs were accompanied by specific increases in C3 and C5 acylcarnitines and several other intermediate metabolites. They attributed a rise in circulating BCAAs in obese subjects to low insulin-like growth factor 1 levels. Their hypothesis is that increased catabolism of BCAAs leads to activation of the mammalian target of rapamycin/S6K1 kinase pathway and phosphorylation of insulin receptor substrate 1 on multiple serines, contributing to insulin resistance. Independently, Gall et al. (2010) proposed α-hydroxybutyrate as an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. The authors concluded that the increase in α-hydroxybutyrate may be because of two mechanisms including elevation of hepatic oxidative stress resulting in an increased demand for GSH production and by elevation of the reduced and oxidized nicotinamide adenine dinucleotide ratio because of increased lipid oxidation and observed that they saw some evidence of both mechanisms in their study. This group noted similar metabolic changes as those described by Newgard et al. (2009).
but concluded that this single biomarker may prove more useful. It will be interesting to see which approach (if any) gains more traction in clinical practice. Clearly, metabolomics technology is moving beyond the stage of "showing promise" to actual delivery of specific biomarkers.

In addition to biomarker discovery, metabolomics has been utilized as a tool to reexamine classic toxicity models for novel mechanistic insights. Galactosamine has been utilized as a hepatotoxicity model for many years (Coen et al., 1968). Early mechanistic studies identified uridine nucleotide depletion as a probable mechanism responsible for the hepatic pathology as UTP supplementation ameliorated the toxicity (Keppler et al., 1969). A confounding finding was the observation that glycine supplementation also ameliorated the toxicity, which was attributed to an alternative mechanistic explanation involving Kupffer cells (Stachlewitz et al., 1999). Metabolomic evaluation of a galactosamine model revealed the novel observation that glycine increased hepatic uridine, which was most likely responsible for the ameliorating effects in the model (Coen et al., 2007) nicely demonstrating how metabolomics can bring new insights to "settled science."

CLINICAL APPLICATIONS
Overview

Measurements of specific endogenous metabolites have long been a part of the practice of medicine from the measurement of urea, creatinine, glucose, and cholesterol that are part of most clinical blood test panels to measurements routinely made in the blood of newborns to exclude treatable diseases such as phenylketonuria. There is a growing enthusiasm for metabolomics among clinicians, which in part reflects disappointment in the clinical application of genetics to personalized medicine (Limdi and Veenstra, 2010). Factors such as epigenetics, disease, aging, environmental exposures, and diet can all significantly impact susceptibility to a disease, the natural history of that disease, as well as the beneficial and adverse responses to any treatment option. Metabolomics holds the promise of characterizing the phenotype of the patient at any given moment, potentially providing the physician with the ultimate information to personalize care. Moreover, metabolomics can be performed on bodily fluids readily accessible to the physician, including urine, blood, saliva, breath condensate, sweat, ascites, amniotic, and cerebrospinal fluid. Although it is likely that the metabolome in each of these fluids can fluctuate as a function of time of day, environmental factors, and diet (Kochhar et al., 2006; Lenz et al., 2004; Slupsky et al., 2007; Stella et al., 2006), where studied, it appears that large portions of the metabolome remain relatively invariant within an individual (Assafal et al., 2008; Bernini et al., 2009; Holmes et al., 2008) and can therefore provide low "noise" levels for detection of potential biomarkers. This has allowed application of metabolomics approaches to broad areas of medicine, including, cardiovascular disease (Goonewardena et al., 2010), hypertension (Holmes et al., 2008), neuropsychiatric diseases (Quinones and Kaddurah-Daouk, 2009), inflammatory bowel disease (Bezabeh et al., 2009), infectious diseases (Slupsky, 2010), respiratory diseases (Basanta et al., 2010), autoimmune diseases (Seeger, 2009), pre- and neonatal diseases (Atzori et al., 2009; Graca et al., 2009), male infertility (Deepinder et al., 2007), and even eye diseases (Young and Wallace, 2009). Additional general reviews of the potential role of metabolomics in the clinic have been recently published (Goldsmith et al., 2010; Gowda et al., 2008; Madsen et al., 2010; Oresic, 2009; Vinayavekhin et al., 2010). This review will focus on metabolomic approaches to cancer and adverse drug reactions (ADRs).

Cancer

Metabolomics has had its biggest clinical impact in the field of oncology, and this has already generated several reviews (Chung and Griffiths, 2007; Kim et al., 2008; Lane et al., 2008, 2009; Serkova et al., 2007; Spratlin et al., 2009). Oncology is fertile ground for metabolomics for several reasons. First, the concept of personalized therapy is a comfortable one to the oncologist. Patient-specific parameters, such as body surface area, clinical laboratory assessments of kidney and liver function, and more recently host genetic factors, are routinely used in the selection of therapy and individualization of dose (Walko and McLeod, 2009). The oncologist often selects the treatment options based on cancer tissue histology and, increasingly, on molecular phenotyping such as selecting herceptin treatment only from those patients whose breast cancer expresses Her 2 neu (Gonzalez-Angulo et al., 2010). In addition, the oncologist is comfortable with following serum-based biomarkers of response to treatment, such as serial measurements of prostate-specific antigen to gauge response to treatment for prostate cancer (Payne and Cornford, 2010). Moreover, oncologists are especially eager to try new approaches to optimize therapy for their patients because in spite of their best efforts, treatments often fail or cause toxicities with tragic consequences.

An advantage of the field of oncology is the routine surgical removal of tumor tissue that allows extensive metabolic characterization in the laboratory. This can be performed in whole tissue with "magic angle" NMR techniques (Sitter et al., 2010) or in tissue extracts with either MS or NMR techniques. Tissue studies, together with transcriptomic and proteomic characterizations, have revealed that cancer cells tend to have metabolic phenotypes that are distinct from that of healthy cells, and this property may be useful to characterize various cancers in terms of prognosis and response to treatment (Bathen et al., 2007; Oakman et al., 2010; Sitter et al., 2006). A variety of clinical studies have attempted to identify cancer-specific changes in the metabolome that would be reflected in readily accessible bodily fluids (Table 3). Development of such
biomarkers could facilitate wider cancer screening in the population and might provide information on tumor response to treatment.

Both MS and magnetic resonance imaging (MRI) approaches have been used in cancer biomarker discovery (Table 3). One advantage of using MRI-based metabolomic approaches is that once specific discriminating metabolites are identified, it may be possible to use MR technology to visualize and accurately locate tumors in the body. For example, cancer cells typically contain elevated phospholipid levels characterized by an elevation in choline-containing lipids (Glunde and Serkova, 2006). This property has led to a clinical role for magnetic resonance spectroscopy imaging (MRSI) in the detection and management of breast (Bartella and Huang, 2007), prostate (Sciarra et al., 2010), and brain (Dowling et al., 2001) cancers. One recent study suggested that MRSI imaging of choline had a 100% sensitivity for detected malignant versus benign breast tumors and suggested that this approach could eliminate the need for over 50% of breast biopsies (Bartella et al., 2007). MRSI imaging to detect breast and prostate cancer is no longer considered experimental, and the costs for these tests are reimbursed by major health insurance plans. This has provided an attractive path for metabolomic biomarker discovery and clinical translation in oncology.

Based on the successes, there has been substantial interest within the oncology community to expand research on the use of metabolomics in the detection and management of cancers (Evelhoch et al., 2005).

### Adverse Drug Reactions

ADRs remain a major problem in the clinic despite the best efforts of the pharmaceutical industry to design and screen for these liabilities. The most problematic toxicities are those that are “idiosyncratic,” meaning they are not clearly dose related, often occur late in therapy, and are relatively rare. The liability for idiosyncratic toxicities might not be evident during preclinical and clinical development of a drug and only become apparent postmarketing when a large population of patients is exposed. There is an urgent need for improved biomarkers that could identify drugs in development with the potential for serious ADRs. There is also need for biomarkers that could identify the onset of an ADR before the patient becomes ill and ideally identify the susceptible patient before they actually receive the drug. Such biomarkers would improve the safety of existing drugs and provide a path to approval for drugs found to have ADR liabilities during development.

The organs most frequently affected by serious ADRs are the liver and kidney. Currently, available biomarkers do not reliably detect or quantitate risk for ADRs involving these organs, nor do they predict which patients will develop these injuries before they occur. Blood urea nitrogen and creatinine are the blood tests most commonly used in the clinic to detect kidney dysfunction, but these become abnormal only after severe kidney damage has occurred. Available urine tests are neither sensitive nor specific for many types of kidney ADRs. To identify liver injury at the early, presymptomatic stage, it is often recommended that the patients receiving potentially hepatotoxic drugs undergo frequent monitoring of serum alanine aminotransferase (ALT) levels. However, there are no data to support the effectiveness of this practice (Senior, 2009). Moreover, it is not necessary to stop treatment in all patients experiencing asymptomatic elevations in serum ALT because these elevations will frequently resolve despite continuing drug treatment (Senior, 2009; Watkins, 2009; Watkins et al., 2008). This phenomenon of “adaptation” does

---

**TABLE 3**

Examples of the Use of Metabolomics to Improve Detection of Cancers

<table>
<thead>
<tr>
<th>Fluid</th>
<th>Cancer</th>
<th>Technique</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum</td>
<td>Breast</td>
<td>GC-MS</td>
<td>Wang et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>Kidney</td>
<td>GC/LC-MS</td>
<td>Gao et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>Prostate</td>
<td>LC-MS</td>
<td>Osl et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>Ovarian</td>
<td>NMR, LC/MS</td>
<td>Guan et al. (2009) and Odunsi et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>Liver</td>
<td>NMR</td>
<td>Gao et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>Oral</td>
<td>NMR</td>
<td>Tiziani et al. (2009)</td>
</tr>
<tr>
<td>Plasma</td>
<td>Pancreas</td>
<td>LC-MS</td>
<td>Urayama et al. (2010)</td>
</tr>
<tr>
<td>Saliva</td>
<td>Breast, oral, pancreatic</td>
<td>CE-TOF-MS</td>
<td>Sugimoto et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>Oral</td>
<td>LC-MS</td>
<td>Yan et al. (2008)</td>
</tr>
<tr>
<td>Urine</td>
<td>Kidney</td>
<td>LC/GC-MS</td>
<td>Kim et al. (2009a)</td>
</tr>
<tr>
<td></td>
<td>Bladder</td>
<td>GC-MS</td>
<td>Issaq et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>Breast</td>
<td>GC-MS</td>
<td>Nam et al. (2009)</td>
</tr>
<tr>
<td>Breath</td>
<td>Lung</td>
<td>GC-MS</td>
<td>Phillips et al. (2007) and Poli et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>Breast</td>
<td>GC-MS</td>
<td>Phillips et al. (2006)</td>
</tr>
<tr>
<td>Feces</td>
<td>Colon</td>
<td>NMR</td>
<td>Monleon et al. (2009)</td>
</tr>
<tr>
<td>Prostatic fluid</td>
<td>Prostate</td>
<td>NMR</td>
<td>Cheng et al. (2001)</td>
</tr>
<tr>
<td>Bile</td>
<td></td>
<td>NMR</td>
<td>Wen et al. (2010)</td>
</tr>
</tbody>
</table>

---

**Downloaded from toxsci.oxfordjournals.org on August 16, 2011 at Acquisition Services**
not occur in all patients, and liver injury may become irreversible in these patients even if drug therapy is discontinued (Senior, 2009; Watkins, 2009; Watkins et al., 2008). It is not currently possible to distinguish patients with benign and reversible ALT elevations from those in whom the liver injury will progress. The combination of elevations in serum ALT and bilirubin has been proposed as a reliable sign of serious liver toxicity, but this denotes serious and potentially irreversible liver injury (Watkins et al., 2008).

Liver and kidney ADRs are attractive targets for metabolomic biomarker discovery (Boudonck et al., 2010)—the liver because it is the major organ in the body that generates the metabolome in blood and the kidney because it largely influences the urinary metabolome. With kidney toxicity, most metabolomics studies to date have been performed in rodents (Boudonck et al., 2009) and few human studies have been reported. The paucity of human research may reflect current enthusiasm for newly discovered protein markers of renal injury that can be measured in urine (Dieterle et al., 2010). However, in a recent study in 13 healthy volunteers (Klawitter et al., 2010a), changes in urine metabolome were observed after a single dose of the potentially nephrotoxic drug cyclosporin A and these changes were similar to those observed in rats experiencing nephrotoxicity because of the same drug (Klawitter et al., 2010b).

The potential of metabolomics to predict liver toxicity was first suggested by Clayton et al. (2006). In this study, Sprague-Dawley rats were given a single toxic threshold dose of acetaminophen (APAP). NMR-based metabolomic profiles of urine samples collected prior to dosing were found to significantly correlate with the extent of liver injury observed subsequently in each rat. The use of predose metabolomic measurements to predict postdose outcomes has been termed “pharmaco-metabonomics,” defined as “the prediction of the outcome of a drug or xenobiotic intervention in an individual based on a mathematical model of pre-intervention metabolite signatures” (Clayton et al., 2006).

Acetaminophen causes liver injury by conversion to a reactive metabolite N-acetyl-p-benzo-quinone (NAPQI) in the liver, and reactive metabolites are believed to underlie most drug-induced liver injury (Srivastava et al., 2010). The ability of the pretreatment metabolome to predict individual susceptibility to acetaminophen hepatotoxicity may be explained if portions of the metabolome correlate with the liver’s ability to produce and accumulate the NAPQI. This is plausible because many of the enzymes and transporters involved in the disposition of drugs are also known to be involved in the disposition and metabolism of endogenous substances (Amacher, 2010). In this case, the high level of an enzyme or transporter in a person could directly result in parallel increases in the production of endogenous metabolites and the reactive acetaminophen metabolite. Alternatively, or in addition, endogenous metabolites might compete with acetaminophen for certain nontoxic metabolic pathways, influencing the fraction of the consumed dose that is converted to the reactive metabolite. In this case, a low concentration of an endogenous metabolite would be correlated with a high concentration of the reactive metabolite.

There is experimental support for these hypotheses. In the rat study referenced above (Clayton et al., 2006), the urine metabolome not only predicted individual susceptibility to acetaminophen toxicity but also the extent of production of the major metabolite (glucuronide) of acetaminophen excreted in urine. In a study of healthy human volunteers given a single oral dose of acetaminophen, those with high predose levels of p-cresol-sulfate in their urine tended to have reduced urinary elimination of the sulfate metabolite (Clayton et al., 2009). p-Cresol is known to be produced from by the gut bacteria. The authors concluded that the p-cresol coming from the gut via enterohepatic circulation competed with the acetaminophen in the liver for sulfation. Further support for the endogenous metabolome being able to predict drug disposition in man has come from a recent study in healthy volunteers who received a single oral dose of the immunosuppressant drug tacrolimus (Phapale et al., 2010). Using an unbiased LC/MS profile and partial least squares modeling, it was found that the level of just four endogenous metabolites yielded a significant prediction of the systemic exposure to tacrolimus (assessed as the serum area under the concentration curve). Because tacrolimus has a narrow therapeutic index and a high degree of individual variation in oral clearance (Venkataramanan et al., 1995), it was proposed that quantitation of these four metabolites in urine might be useful in the clinic to guide personalized dosing of tacrolimus.

Following up on the observations of Clayton et al. (2006) in rats, serial characterization of the urine metabolome was performed in 58 healthy volunteers (Winnike et al., 2010) who received recurrent treatment with acetaminophen that produces mild liver injury (elevations in ALT, aspartate aminotransferase, and alpha glutathione-S-transferase) in about one-third of subjects (Winnike et al., 2006). Unlike what was observed in the rat studies, NMR profiles of 24-h urine collections obtained before dosing could not significantly separate those subjects who would develop ALT elevations after commencing acetaminophen dosing from those who would not. However, this separation was achieved in urine obtained shortly after dosing but before the onset of ALT elevations. The application of metabolomics to predict adverse events early in the dosing, but prior to the occurrence of phenotypic change, was termed “early intervention pharmaco-metabonomics” (Winnike et al., 2010). An interesting finding of this study was that the acetaminophen mercapturate and cysteine conjugates, which are the urinary excretion products of NAPQI, tended to be higher in responders. However, the urinary levels of these metabolites were not predictive by themselves but only became predictive when changes in the endogenous metabolome were also considered. This demonstrates the potential power of combined analysis of endogenous and exogenous metabolites in predicting outcomes of drug therapy.
Another recent study has supported the ability of pretreatment metabolome to predict patients who are most likely to develop liver injury caused by the oral anticoagulant, ximelagatran (Andersson et al., 2009; Keisu and Andersson, 2010). This drug was withdrawn from worldwide markets because of rare incidence of severe hepatic injury (Lee et al., 2005). To identify potential biomarkers that could be used to identify susceptible patients, unbiased NMR- and MS-based metabolomics analysis was conducted on plasma samples obtained from patients before they were treated with ximelagatran in a large clinical trial (Olsson, 2003). It was observed that patients who would develop liver injury tended to have lower pretreatment levels of serum pyruvate than did those patients who would not develop elevations. A limitation of this study is that the blood samples analyzed in subjects with and without ALT elevations were not obtained at the same intervals prior to treatment.

The above studies suggest that pharmaco-metabonomics, perhaps combined with genetic testing, may have a future role in personalized medicine.

**CURRENT NEEDS/FUTURE DIRECTIONS**

Clearly, metabolomics has grown and changed over the past several years evolving into a technology with wide application across the spectrum of biological sciences. Although toxicology is no longer at the center of the metabolomics universe, the technology has demonstrated tremendous potential for toxicological applications. Current needs for the technology include some sort of standardization in MS approaches so that different laboratories running different instrumentation can still generate comparable data on the same set of samples. Currently, this is not the case as different MS configurations and separation strategies can give very different biochemical pictures of a biofluid or tissue extract. The current atmosphere has a bit of a “Wild West” feel to it with most investigators using their own approaches that tend to vary from laboratory to laboratory. Although this is probably a necessary stage of development of the technology, as no single approach or strategy can yet claim superiority, it is inhibiting to those wishing to pursue regulatory applications of the technology. Although some attempts have been made to standardize MS approaches and their reporting (Jones et al., 2007; Kanani et al., 2008; Sumner et al., 2007; Want et al., 2010b), we still have a long way to go. Reviews by Gomase (Gomase, 2008) and Go (2010) provide nice summaries of current database resources for metabolomic analysis. Although these databases are increasingly valuable, they are dependent on the approaches used to generate them so that unless identical methods are employed, they can become difficult to utilize. A fundamental tenet of scientific research is that experiments be repeatable. Unless and until metabolomics can routinely (and easily) meet that requirement questions as to its applicability as a decision making tool will remain.

In general, NMR and MS data sets, even when they are acquired on the same samples, are interpreted separately, and the molecular information is then collated and explained using “gray matter.” Clearly, both of these approaches are just tools for extracting the information of the metabolome, and future developments will be aimed at making these tools more interactive and seamless (Lindon and Nicholson, 2008). In the end, we want a thorough and accurately annotated assessment of the metabolome, and the tools we use to arrive at that are inconsequential.

Apart from analytical issues is the ever-present problem of information overload. Even the simplest of experiments can generate hundreds or even thousands of metabolic differences. Assuming that all these metabolic changes could be correctly annotated, making sense of all that data is a herculean task. Software solutions are evolving, particularly those that integrate metabolomics data with transcriptomic data, but they are not yet so user friendly that one needs to simply enter data and out pops a definitive diagnosis or mechanism. A significant gray matter investment is still required. This is probably the area where those that adopted the technology early, only to drop it shortly thereafter, veered farthest of course. No matter how sophisticated our analytical techniques and software applications become, they will not eliminate the need for talented biochemists and toxicologists to interpret the data. Metabolomics should not be viewed as a technology to do our thinking for us but rather as a tool to unveil hidden aspects of biochemistry and pathophysiology that have previously proven elusive.

One clear advantage metabolomics has over transcriptomics is that the former directly conveys phenotype and can be obtained from peripheral samples. This makes the technology very attractive from a translation standpoint. As the routine collection of biopsy samples is unattractive from a clinical standpoint, identification of preclinical biomarkers in urine, serum, or other accessible biofluid is easily transferred to the clinical setting. The question remains about how predictive preclinical biochemistry is of clinical biochemistry, but there is no reason to suspect that metabolomic endpoints are any less predictive than currently employed markers of efficacy or toxicity across species. We currently routinely use clinical chemistry assessment of metabolic biomarkers (e.g., glucose, cholesterol, bilirubin, creatinine, etc.) with great confidence in the biological significance of these markers in both the preclinical and clinical settings. It is not difficult to imagine that 1 day we will have the same confidence in dozens if not hundreds of additional metabolic markers made accessible by metabolomic technology. As preclinical and metabolomic research gains traction, characterization of these “novel” biomarkers will naturally follow. Proteomics could theoretically fulfill a similar role as metabolomics; however, sample throughput and data turnaround time for proteomics approaches need to be markedly improved. This translational aspect of the metabolomic technology is one of the most exciting prospects of metabolomics research.
What does the future hold? It is clear that metabolomics has now established itself as a distinct field with applications across the spectrum of biological sciences from single-cell experiments to complex clinical studies. As our tools get more sophisticated, integrated omic approaches will begin to merge these scientific areas, such that the commonalities of disease mechanisms and toxic mechanisms will become apparent and the line where physiological response drifts into toxic response will be more sharply drawn. Within the toxicology community, mechanisms and biomarkers are where the technology is currently beginning to payoff and where the greatest gains will come in the near future.

REFERENCES


