Horizons in Nutritional Science

Metabolomics: an emerging post-genomic tool for nutrition

Phillip D. Whitfield1*, Alexander J. German2 and Peter-John M. Noble2

1Department of Veterinary Preclinical Sciences and
2Small Animal Teaching Hospital, Faculty of Veterinary Science, University of Liverpool, Crown Street, Liverpool L69 7ZJ, UK

(Received 9 June 2004 – Revised 21 June 2004 – Accepted 21 June 2004)

The post-genomic era has been driven by the development of technologies that allow the function of cells and whole organisms to be explored at the molecular level. Metabolomics is concerned with the measurement of global sets of low-molecular-weight metabolites. Metabolite profiles of body fluids or tissues can be regarded as important indicators of physiological or pathological states. Such profiles may provide a more comprehensive view of cellular control mechanisms in man and animals, and raise the possibility of identifying surrogate markers of disease. Metabolomic approaches use analytical techniques such as NMR spectroscopy and MS to measure populations of low-molecular-weight metabolites in biological samples. Advanced statistical and bioinformatic tools are then employed to maximise the recovery of information and interpret the large datasets that are generated. Metabolomics has already been used to study toxicological mechanisms and disease processes and offers enormous potential as a means of investigating the complex relationship between nutrition and metabolism. Examples include the metabolism of dietary substrates, drug-induced disturbances of lipid metabolites in type 2 diabetes mellitus and the therapeutic effects of vitamin supplementation in the treatment of chronic metabolic disorders.

Metabolomics: Metabonomics: Metabolite profiling

With the advent of the post-genomic era, biological, medical and veterinary research has witnessed an explosion in strategies that provide an integrative view of the molecular regulation of cells and whole organisms (Watkins & German, 2002). These advances have been driven by the development of novel technologies that can analyse global sets of gene products. Transcriptomics defines the population of mRNA species in a cell at a specific time and set of conditions; proteomics addresses the challenging problem of defining changes in protein expression, protein dynamics and post-translational modifications; whilst the emerging field of metabolomics measures changes in populations of low-molecular-weight metabolites under a given set of conditions (Fiehn et al. 2001).

Low-molecular-weight metabolites represent the endproducts of cell regulatory processes and as such advertise the response of biological systems to a variety of genetic and environmental influences (Fiehn, 2002). Metabolomic strategies aim to detect changes in the distribution and concentration of a broad range of metabolites and can be applied to multiple levels of biological organisation from single cells to whole organisms. These analyses involve the use of modern analytical techniques to measure global populations of metabolites in biological samples. Advanced statistical and bioinformatic tools are then employed to maximise the recovery of information and to aid interpretation of the very large datasets that are generated.

In man and animals, metabolite profiles can be regarded as important indicators of normal phenotype and pathology, and offer the possibility of identifying surrogate biomarkers of disease states. Abnormal cellular processes may lead to disturbances in the profile of endogenous low-molecular-weight metabolites. Metabolomic-based strategies can, therefore, provide an insight into metabolite perturbations caused by differential gene expression, toxicological insult, pathophysiological processes and altered nutritional status (Nicholson et al. 2002). The present article will review the emerging post-genomic science of metabolomics. Metabolomic-based technologies will be outlined and a number of applications of metabolomic analyses described. The implications of metabolomics for the nutritional sciences will also be discussed.

Terminology

The field of metabolomics has already produced a vast array of specialist terms; however, their precise definitions are still evolving. The greatest debate relates to the terms

Abbreviation: HIF-1, hypoxia-inducible factor-1.

* Corresponding author: Dr P. D. Whitfield, fax +44 151 794 4243, email pdw01@liv.ac.uk
used to describe the global measurement of low-molecular-weight molecules in biological systems; metabolomics and metabonomics. Metabolomics has been defined as the comprehensive analysis of the whole metabolome under a given set of conditions (Goodacre et al. 2004). Metabonomics has been described as the quantitative measurement of time-related multiparametric metabolic responses of multicellular systems to pathophysiological stimuli or genetic modification (Nicholson et al. 1999). It has recently been suggested that metabolomics should be used as a parallel term to genomics, transcriptomics and proteomics, and that a more appropriate term for metabonomics would be metabolic fingerprinting (Fiehn, 2002; Sumner et al. 2003). Clearly, some research will overlap these definitions and a consensus on nomenclature needs to be established within the field. A list of terms and definitions is given in Table 1.

### Metabolomic technologies

A major goal of all metabolite-profiling strategies is to examine large populations of low-molecular-weight metabolites within a biological system. Since body fluids such as plasma and urine contain hundreds of small molecules and there may be as many as 200 000 metabolites in the plant kingdom (Fiehn, 2002), this represents a considerable analytical challenge. Unlike genomics, transcriptomics and proteomics, where essentially just one class of compound is analysed, metabolomic strategies need to detect a broad spectrum of molecules with diverse properties (Weckwerth, 2003).

Conventional methods for the analysis of low-molecular-weight metabolites have focused on specific compound classes, with isolation and detection protocols being tailored to the chemical and physical characteristics of the metabolites of interest. This approach has been widely employed to profile the metabolites that accumulate in the body fluids and tissues of patients with inborn errors of metabolism such as aminoacidopathies, fatty acid oxidation defects and organic acidemias (Clayton, 2001; Rashed, 2001). In contrast, metabolomic analyses must avoid bias for specific molecules and ideally should be able to detect every individual metabolite. Achieving the broadest overview of metabolite composition is difficult and requires an integrated strategy for metabolite analysis and data processing.

Suitable analytical techniques for the global analysis of metabolites must be sensitive, robust and have the capacity for high-throughput analysis required to screen large numbers of samples. In recent years, considerable advances have been made in the development of analytical technologies to measure and interpret complex metabolite profiles. However, given the wide dynamic and chemical range of low-molecular-weight metabolites in biological mixtures, it is not yet possible to analyse all low-molecular-weight metabolites with a single analytical platform (Glassbrook & Ryals, 2001). The techniques most frequently employed for metabolomic studies are NMR spectroscopy and MS.

The data acquired by either NMR spectroscopy or MS from biological mixtures produce a complex spectral profile which reflects the metabolic status of the organism (Shockcor & Holmes, 2002). A single spectrum can contain thousands of signals. In order to interpret the biochemical perturbations, metabolomic-based strategies employ powerful bioinformatic and statistical methods (Fig. 1).

### Nuclear magnetic resonance spectroscopy

To date, NMR spectroscopy has been the major analytical technique used to profile global populations of low-molecular-weight metabolites. This use of NMR spectroscopy in metabolomic studies has been pioneered by Nicholson and colleagues (Nicholson et al. 1999, 2002; Nicholson & Wilson, 2003).

NMR-based analysis of biological samples has significant advantages for metabolomic applications. NMR spectroscopy requires little or no sample preparation and, as it is non-destructive, is capable of generating a comprehensive profile of low-molecular-weight metabolites from intact biofluids and tissues (Reo, 2002). NMR spectroscopy is also inherently quantitative. Despite these considerable advantages, NMR does suffer from certain drawbacks. NMR is not as sensitive as MS and requires relatively large sample volumes. The issue of sensitivity has been addressed by the development of magnets with increased field strength and improvements in the design of NMR detectors (Griffin, 2004); however, the analysis of low-abundance metabolites by NMR spectroscopy can still prove problematic.

### Mass spectrometry

The high sensitivity of MS combined with its wide dynamic range makes the technique an extremely powerful tool for the analysis of large populations of metabolites.

---

**Table 1. Terminology of metabolomics**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metabolomics</td>
<td>The comprehensive analysis of the whole metabolome under a given set of conditions (Goodacre et al. 2004)</td>
</tr>
<tr>
<td>Metabonomics</td>
<td>The quantitative measurement of time-related multiparametric metabolic responses of multicellular systems to pathophysiological stimuli or genetic modification (Nicholson et al. 1999)</td>
</tr>
<tr>
<td>Metabolome</td>
<td>The full set of low-molecular-weight metabolites within, or that can be secreted by, a given cell type or tissue (Nicholson &amp; Wilson, 2003)</td>
</tr>
<tr>
<td>Metabolic fingerprinting</td>
<td>The application of any technological approach whose output is processed with pattern-recognition software and without differentiation of individual metabolites (Sumner et al. 2003)</td>
</tr>
<tr>
<td>Lipidomics</td>
<td>The characterisation of chemically distinct lipid species in cells and the molecular mechanisms through which they facilitate cellular function (Han &amp; Gross, 2003)</td>
</tr>
</tbody>
</table>
However, as with NMR spectroscopy, MS has a number of weaknesses. Mass spectrometric analyses require extensive sample preparation. Low-molecular-weight metabolites from biological samples typically have to be extracted in organic solvents, which can result in losses of certain compounds. Furthermore, variable ionisation efficiencies of metabolites mean that MS is unable to provide absolute quantification (Weckwerth & Fiehn, 2002), although the concentrations of low-molecular-weight metabolites can be accurately determined by MS using stable-isotope internal standards.

Ion suppression and matrix effects can also lead to difficulties in detecting certain compounds. These problems can be alleviated by coupling a chromatography system to the mass spectrometer. The chromatographic step is able to reduce the number of competing analytes entering the mass spectrometer in addition to separating complex mixtures of metabolites (Pham-Tuan et al. 2003). As a result, liquid chromatography–MS is increasingly being used to profile low-molecular-weight metabolites (Plumb et al. 2002, 2003; Idborg-Bjorkman et al. 2003). Similarly, GC–MS has been widely employed for metabolomic applications, despite the need to chemically derivatise many biological compounds before analysis (Fiehn et al. 2000a,b; Jonsson et al. 2004; Weckwerth et al. 2004). The recent developments in Fourier transform ion cyclotron resonance MS have also aroused interest within the field of metabolomics. The capability of Fourier transform ion cyclotron resonance MS to generate spectral data relating to the elemental composition of metabolites suggests that this analytical technique may play a major role in future metabolomic strategies (Aharoni et al. 2002).

**Statistical analysis**

The profiling of global sets of low-molecular-weight metabolites in biological samples generates a wealth of information relating to the physiological or pathophysiological status of an organism (Shockcor & Holmes, 2002). To provide evidence of metabolite perturbations, pattern-recognition tools are employed. These statistical methods provide an efficient, non-biased procedure for interpreting the complex, multivariate datasets and allow the correlation of metabolic responses in biological systems to be fully investigated (Holmes & Antti, 2002).

Pattern-recognition methods fall into two general classes; supervised and unsupervised. Unsupervised methods, such as principal component analysis, determine pattern sets according to their properties and without prior knowledge of the sample class. These approaches are typically employed to identify clustering within sample populations. Supervised methods are more powerful statistical tools that use additional information on the datasets, such as clinical data, to determine similarities and differences between predefined groups. An example of a supervised method is partial least squares data analysis (Holmes & Antti, 2002; Lindon et al. 2003a).

**Identification of low-molecular-weight metabolites**

The identification of metabolites in biological systems can lead to the targeting of specific metabolic pathways which may provide greater insight into the mechanisms of certain metabolic disturbances. However, the bioinformatic tools required to deconvolute spectral data and identify low-molecular-weight metabolites are still being developed (Mendes, 2002; Kell, 2004). Investigators in the fields of genomics and proteomics can access databases such as Genbank or Swiss-Prot to rapidly determine the identity of a gene or protein, but there is currently no equivalent database for metabolomic analyses. Whilst there are some spectral libraries available for the comparison of NMR or MS data, the structure of low-molecular-weight metabolites must largely be determined manually, which is a time-consuming process.

**Applications of metabolomics**

Whilst still in its infancy, metabolomics is already beginning to make a significant impact in the fields of biology...
and medicine. This section provides an outline of the major applications arising from strategies which aim to screen global populations of low-molecular-weight metabolites.

**Gene function**

Metabolite profiling techniques have been used to explore the function of genes in plants, yeast and experimental animals. In one such study, the metabolite profiles of leaf extracts from *Arabidopsis thaliana* were examined. Analysis of wild-type and transgenic plants demonstrated that the individual plant types each had a distinct metabolite profile (Fiehn et al. 2000a). These findings suggest a potential use for metabolomics in the monitoring of genetically modified foods (Kuiper et al. 2001; Roessner et al. 2001; Sumner et al. 2003; Trehewe, 2004).

Metabolomic strategies have also been used to define gene function in the yeast, *Saccharomyces cerevisiae*. Raamsdonk et al. (2001) employed NMR spectroscopy to characterise the phenotype of silent mutations of specific genes. In addition, investigators have employed MS-based metabolomic strategies to obtain a ‘metabolic footprint’ of extracellular metabolites derived from *S. cerevisiae* (Allen et al. 2003). This study was able to distinguish between the wild-type and mutant strains of yeast during different growth phases.

Another investigation characterised the urinary metabolite profiles from two phenotypically normal strains of mice (Gavaghan et al. 2000). Using NMR spectroscopy and pattern-recognition methods, differences were observed in tricarboxylic acid cycle intermediates and methylene pathway between the strains of mice. This technology has obvious potential in evaluating animal models of human disease states and further demonstrates the capability of metabolomics in determining the metabolic signature of complex living organisms.

**Toxicology and drug discovery**

A major application of metabolite profiling is in the field of toxicology where it provides a means of rapidly assessing the health risks of specific toxins. This strategy has been used to determine the disturbances in metabolite populations of body fluids and tissues following the administration of various chemicals and drugs to experimental animals (Beckwith-Hall et al. 1998; Nicholls et al. 2001; Waters et al. 2002; Lenz et al. 2004). The Consortium for Metabonomic Toxicology (COMET) is using metabolite-profiling methods to evaluate the safety of candidate drugs (Lindon et al. 2003b). This group, which comprises Imperial College of Science Technology and Medicine, London (UK) and five major pharmaceutical companies (Bristol-Myers-Squibb, Eli Lily and Company, Hoffman-LaRoche, NovoNordisk and Pfizer Incorporated), is employing NMR spectroscopy in conjunction with multivariate statistical analysis, to screen the urine and serum of laboratory rodents for markers of hepatic and renal toxicity. This experimental approach promises to play a central role in improving the drug discovery process through faster and better assessments of drug efficacy and toxicity.

**Disease mechanisms and diagnosis**

A number of complex diseases, such as atherosclerosis, type 2 diabetes mellitus and cancer, may result from a chronic imbalance of normal metabolism. Therefore, metabolomic-based approaches may be of use in identifying surrogate biomarkers of pathological states. The ability to clinically evaluate diseases using biochemical parameters has major implications for diagnostic procedures, prognostic evaluation and the monitoring of treatments.

A groundbreaking study used NMR-based methods to identify individuals with CHD (Brindle et al. 2002). By characterising metabolite disturbances in serum, individuals with CHD could be differentiated from patients with angiographically normal coronary arteries. This approach was also employed to examine the disease-related metabolic fingerprints of osteoarthritic guinea-pigs (Lamers et al. 2003). Analyses showed significant disturbances in the urinary concentrations of lactic acid, malic acid, hypoxanthine and alanine in affected animals, indicating possible metabolic pathways which may be involved in the osteoarthritic process.

A similar experimental strategy was used to explore the molecular pathogenesis of a specific form of cancer. Grifiths & Stubbs (2002, 2003) employed NMR spectroscopy to measure metabolite changes in mouse-derived tumour cells that were deficient in the transcription factor, hypoxia-inducible factor-1 (HIF-1). HIF-1 up regulates the genes of certain glycolytic enzymes and the glucose transporters GLUT-1 and GLUT-3. The absence of this messenger is known to slow tumour growth. Data analysis showed that ATP production was significantly impaired in HIF-1-deficient tumour cells. Interestingly, the cells also had lower concentrations of glycine, betaine and various choline metabolites. These metabolic intermediates are used in the *de novo* synthesis of purines. The findings indicate that the primary function of HIF-induced up regulation of glycolysis and glucose transport in tumours is to provide an increased supply of precursors for the production of high-energy phosphates. These studies demonstrate the potential of metabolomics to characterise disease mechanisms.

Finally, a study has used metabolomic strategies to profile the metabolite disturbances in the *mdm* mouse model of Batten’s disease (Griffin et al. 2002). Batten’s disease belongs to a group of lysosomal storage disorders known as the neuronal ceroid lipofuscinoses which are characterised by progressive neurodegeneration. Metabolomic-based analyses demonstrated altered metabolite profiles in the neural tissues and sera of *mdm* mice compared with control animals. Free radical damage to brain tissues is thought to play a role in the pathogenesis of neuronal ceroid lipofuscinoses and the *mdm* mouse model has a profound deficiency of vitamin E, a key physiological antioxidant. As a result, the investigators examined the therapeutic effects of dietary supplementation with vitamin E. Metabolomic approaches were able to show that the treatment reversed some of the metabolic disturbances associated with the disease state.
Metabolomics and nutrition

The arrival of the post-genomic era has enabled researchers to investigate the effects of nutrients on physiological functions in man and animals at the molecular level (Elliot & Ong, 2002; Orduovas & Moozer, 2004; van Ommen, 2004). The influence of nutrition in the development of disease is widely understood and it is well known that chronic diseases can have both dietary and genetic components (Go et al. 2003; Milner, 2003). Metabolomics is uniquely suited to explore the complex relationship between nutrition and metabolism and to investigate the role that dietary components play in health and disease (Watkins et al. 2001).

In addition, metabolomics has the potential to explore homeostatic control and how this metabolic balance may be disturbed by deficiencies or excesses of dietary components (German et al. 2003b). The potential value of metabolomics to nutrition has already been demonstrated in studies which have examined the metabolism of ethyl glucoside (Teague et al. 2004) and isoflavones (Solanky et al. 2003) found in the diet.

Investigators have also begun utilising metabolomic-based strategies to analyse broad classes of endogenous metabolites and in particular, lipids. Lipids are amongst the most heavily studied molecules within the nutritional sciences. They play a central role in many physiological processes and their concentrations are often disturbed in metabolic diseases (German et al. 2003a). In recognition of the importance of lipids in cellular metabolism the National Institute of General Medical Sciences in the USA recently funded a US $35 million project entitled ‘Lipid Metabolites and Pathway Strategy’ (Lipid MAPS; www.lipidmaps.org). This research programme aims to fully characterise the lipid populations and signalling pathways in the mouse macrophage.

The global analysis of lipid classes has been termed ‘lipidomics’ (Han & Gross, 2003). Lipidomics-based analyses have already been used to examine the effects of the peroxisome proliferator-activated receptor-γ agonist, rosiglitazone, in a murine model of type 2 diabetes mellitus (Watkins et al. 2002). This study revealed that the administration of this drug had a wide range of effects on lipids in the body fluids and tissues of the treated mice. Rosiglitazone lowered the lipid concentrations of plasma, significantly altered cardioliopin metabolism in the heart, and led to an unusual profile of PUFA in adipose tissue.

A similar study recently employed lipidomics to explore the changes in lipid metabolism of cultured 3T3-L1 adipocytes (Su et al. 2004). The investigations demonstrated that following differentiation, there was a substantial accumulation of odd-chain fatty acids in all of the major lipid classes of the fat cells, including triacylglycerols and phospholipids. The findings suggest an important role for peroxisomes in the processing of adipocyte fatty acids.

Conclusions and future perspectives

Metabolomics is an emerging post-genomic science that is already being employed to investigate gene function, toxicological mechanisms and the pathogenesis of disease states. Metabolomics now promises to play a major role in the nutritional sciences. However, despite the considerable advances in this field, huge challenges still remain. Metabolite-profiling strategies, in many respects, only provide a ‘snapshot’ of the metabolic status of a cell, tissue or organism. Metabolism is a dynamic process and strategies for determining the turnover of low-molecular-weight metabolites will provide further insight into the complexities of metabolic mechanisms (Hellerstein, 2003, 2004). Improved technologies for metabolite analysis and better computational tools for data handling will, therefore, need to be established if metabolomics is to realise its full potential. It is highly probable that the field of metabolomics will continue to be driven by technological developments for the foreseeable future.

Acknowledgements

The authors gratefully acknowledge The Royal Society, The Waltham Foundation and The PetPlan Charitable Trust for funding work in their laboratory. The authors also thank Professor Rob Beynon for his helpful comments and suggestions.

References


