Q&A

SYSTEMS BIOLOGY

Metabonomics

Jeremy K. Nicholson and John C. Lindon

Organisms often respond in complex and unpredictable ways to stimuli that cause disease or injury. By measuring and mathematically modelling changes in the levels of products of metabolism found in biological fluids and tissues, metabonomics offers fresh insight into the effects of diet, drugs and disease.

What are the origins of the field?

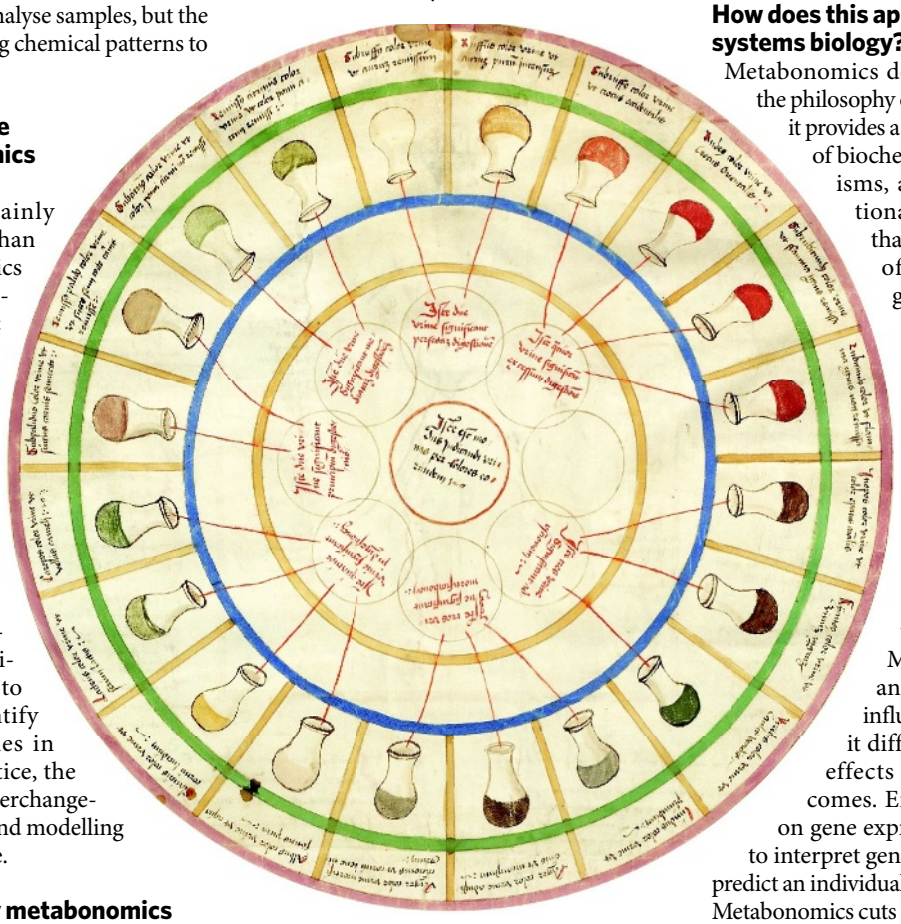
The idea that changes in tissues and biological fluids are indicative of disease goes back at least as far as ancient Greece. Diagnostic 'urine charts' were widely used from the Middle Ages onwards (Fig. 1). These charts linked the colours, smells and tastes of urine to various medical conditions. Such features are, of course, metabolic in origin. Metabonomics, and the related field of metabolomics, uses modern techniques to analyse samples, but the basic principle of relating chemical patterns to biology is the same.

What's the difference between metabonomics and metabolomics?

The distinction is mainly philosophical, rather than technical. Metabonomics broadly aims to measure the global, dynamic metabolic response of living systems to biological stimuli or genetic manipulation. The focus is on understanding systemic change through time in complex multicellular systems. Metabolomics seeks an analytical description of complex biological samples, and aims to characterize and quantify all the small molecules in such a sample. In practice, the terms are often used interchangeably, and the analytical and modelling procedures are the same.

How did modern-day metabonomics begin?

There were two, largely independent, starting points. The first was metabolic-control analysis, a mathematical method developed in the 1960s for modelling metabolism in cells. This required metabolite concentrations to be quantified, and so methods were developed to do this — often using gas chromatography (GC) or GC coupled to mass spectrometry (MS). The second contributing factor was the development of nuclear magnetic resonance (NMR) spectroscopy. By the mid-1980s, NMR was sensitive enough to identify metabolites in unmodified biological fluids. This led to the discovery that altered metabolite profiles are caused by certain diseases or by adverse side effects to drugs. MS techniques were also developed for profiling biological fluids. But metabonomics really took off with the realization that pattern-recognition methods (also known as chemometrics or multivariate statistical analysis) could help to interpret the complex data sets that result from these studies.



THE ROYAL LIBRARY, COPENHAGEN

Figure 1 | Metabonomics of yore. This urine wheel was published in 1506 by Ullrich Pinder, in his book *Epiphanie Medicorum*. It describes the possible colours, smells and tastes of urine, and uses them to diagnose disease.

How does this approach fit in with systems biology?

Metabonomics dovetails beautifully with the philosophy of systems biology, because it provides a 'top-down', integrated view of biochemistry in complex organisms, as opposed to the traditional 'bottom-up' approach that investigates the network of interactions between genes, proteins and metabolites in individual cell types. A problem with systems biology is that each level of biological organization and control — genomics, gene expression, protein expression and metabolism — operates on a markedly different timescale from the others, making it difficult to find causal linkages. Moreover, environmental and lifestyle factors greatly influence metabolism, making it difficult to disentangle their effects from gene-related outcomes. Environmental influences on gene expression also make it hard to interpret genomic data, for example to predict an individual's susceptibility to diseases. Metabonomics cuts through these problems by monitoring the global outcome of all the influencing factors, without making assumptions about the effect of any single contribution to that outcome. Yet in so doing, the individual contributions can be teased out.

What analytical techniques are used for metabonomics?

Usually NMR spectroscopy and MS. NMR is generally used to detect hydrogen atoms in metabolites. In a typical biological-fluid sample, all hydrogen-containing molecules in the sample — including nearly all metabolites — will give an ^{1}H NMR spectrum, as long as they are present in concentrations above the detection limit. The NMR spectrum of a biological fluid is therefore the superposition of the spectra of all of the metabolites in the fluid (Fig. 2). An advantage of NMR is that the biological fluid doesn't require any physical or chemical treatment prior to the analysis. MS studies, on the other hand, usually require the metabolites to be separated from the biological fluid before detection, typically by using high-performance liquid chromatography (or modern variants). Alternatively, the metabolites can be chemically modified to make them more volatile, so that GC–MS can be used.

What are the other pros and cons of these techniques?

Both methods yield information on a wide range of metabolites in a single measurement, without having to preselect which analytes to detect. Furthermore, both can be used to identify the metabolites' structures, and to measure the relative and absolute concentrations of the molecules (although NMR is more reliable for determining concentrations). MS is usually more sensitive than NMR, but many compounds in complex mixtures give variable responses in MS experiments, which can be highly misleading. Because NMR doesn't damage analytes, it is particularly useful for studying metabolite levels in intact tissues, such as biopsy samples, which can then be used in further experiments. The non-destructive nature of NMR also enables the dynamics and sequestration of metabolites in tissue samples to be observed; such information is lost in MS experiments, because this technique disrupts the structures and interactions of molecular complexes.

How are the resulting data interpreted?

The spectra of samples from organisms of interest (such as those with a specific disease) are compared with those from controls, so that the spectral features caused by the disease state can be determined. Precise metabolite concentrations are not always necessary to formulate hypotheses about the mechanism of the disease. But if only a few metabolites turn out to be important, then knowledge of their concentrations might be instructive, and these can subsequently be measured.

Doesn't the quantity and complexity of the data make interpretation impossible?

Not if you use multivariate statistical analysis — a collection of techniques in which the intensities of peaks in a spectrum are used as coordinates in multidimensional plots of metabolic activity. This allows distinctive patterns in the data to be identified more easily than by looking at the original spectra. The multidimensional plots can even be reduced to two- or three-dimensional graphs, to help visualize any clustering of points that might be used to characterize the data (Fig. 2d).

What exactly do these analyses involve?

One frequently used technique, known as principal-components analysis, attempts to find a small set of variables that explain the original data set. Alternatively, in what are known as supervised methods, mathematical models are derived from a 'training' set of multiparametric data, so that individual biological samples can be classified (for example as being characteristic of a disease state). A widely used supervised method is called partial least squares (PLS) analysis. PLS is often combined with discriminant analysis — a method that finds a linear combination of features that can be used to classify data into sets. The combination of PLS with discriminant analysis defines a surface that can be placed in a multidimensional plot to separate data into classes.

How are metabolites associated with disease states identified?

There are various methods based on NMR and MS, but one exciting new approach is an NMR technique termed statistical total correlation spectroscopy (STOCSY). In this method, correlations are found between the intensities of spectral peaks across a cohort of samples. This enables all the peaks from a given metabolite to be identified, so that the structure of that molecule can be determined. The beauty of STOCSY is that additional information can be gleaned by examining lower correlations between peaks, or even negative correlations, as these indicate connections between two or more molecules that are involved in the same biochemical process. Statistical heterospectroscopy is a powerful extension of STOCSY that allows the co-analysis of data sets acquired using combinations of different spectroscopic techniques.

Once metabolites have been identified, what then?

They can be used as diagnostic biomarkers for biological events. Metabonomics thus allows real-world, medical observations to be related to data from all the other '-omics' technologies, which are less directly related to actual biological outcomes than metabolism is (Fig. 3, overleaf). And because samples of biological fluids (usually urine or blood) can be collected fairly easily, the time-dependent fluctuations of metabolites that occur in response to disease, drug effects or other stimuli can be easily studied in detail.

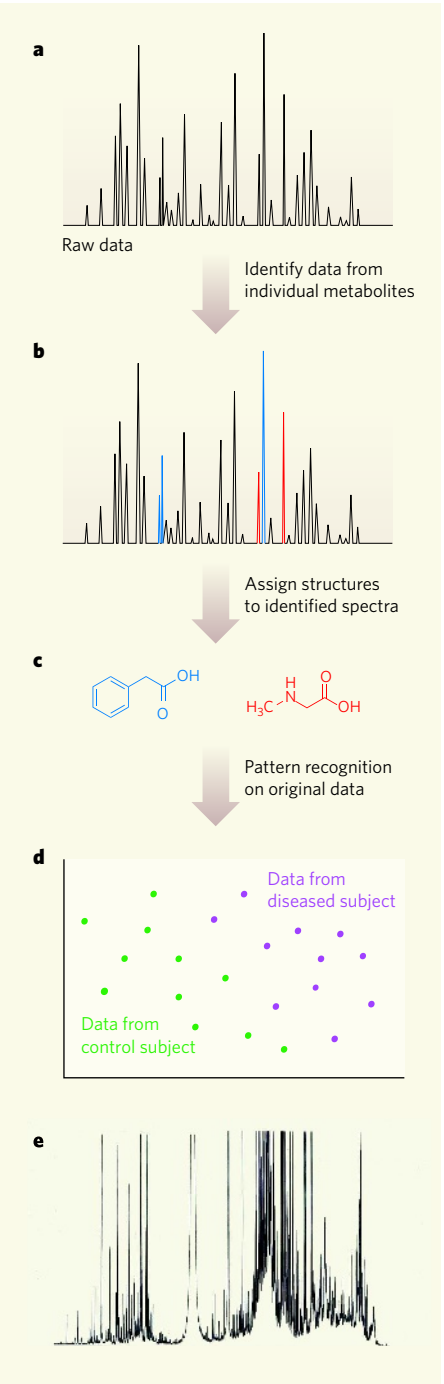


Figure 2 | The discovery process in metabonomics. Metabonomics analyses the metabolites in biological fluids to determine the metabolic response of an organism to a physiological stimulus. **a**, A typical procedure might start with the NMR spectrum of a biological fluid, which contains signals from hundreds of metabolites. **b**, The individual spectra for each metabolite are identified. **c**, This enables the structure of the metabolites to be determined. **d**, Pattern-recognition techniques can be used to work out how the spectra of biological fluids from individuals who have a disease differ from those of healthy subjects. Here, 'principal-component analysis' has reduced multivariate data to a two-dimensional plot. **e**, Although the procedure above is conceptually simple, this NMR spectrum of human urine reveals how complicated the raw data can be.

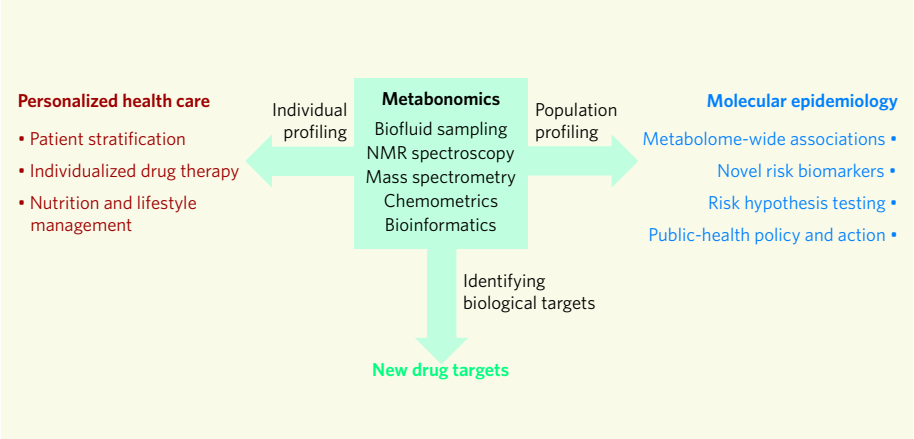


Figure 3 | Applications of metabonomics. There are three broad areas that might benefit from metabonomics. Metabolic profiling of individuals could be used in personalized health care to work out patients' susceptibilities to disease or their responses to medicines, and to tailor their lifestyles and drug therapies accordingly. Metabolic profiling of populations could allow the development of 'molecular epidemiology' — the ability to work out the susceptibilities of specific groups to disease. This might allow metabolites to be identified as risk identifiers (biomarkers) for diseases, with implications for health screening programmes. Finally, by identifying biochemical pathways for disease, metabonomics could uncover new targets for drug discovery.

Are there any drawbacks to metabonomics?

The need to use complex data-interpretation techniques and combinations of analytical methods isn't ideal. Another problem is that the number of metabolites produced by any given system cannot be predicted — compare this with genome sequencing, where the number of genes is known. But this problem isn't insurmountable, and is similar to that of other fields, such as epigenetics.

Is there an equivalent of the Human Genome Project?

Yes indeed. The Human Metabolome Project in Canada strives to provide a repository of all human metabolites. So far, it has collected a few thousand of these compounds and created an accompanying database of spectroscopic data. Both the repository and the database are publicly available as resources for metabolic research. Similarly, the LIPID Metabolites And Pathways Strategy (LIPID MAPS) consortium in the United States is attempting to characterize all the lipids in the human macrophage, and currently has a database of more than 10,000 characterized substances. Such resources are invaluable, but what they can't reveal are the all-important dynamic interactions of metabolites in space and time, or how metabolite profiles vary across human populations.

How has metabonomics been used for drug discovery?

Its use in evaluating drug toxicity has been comprehensively assessed by the Consortium for Metabonomic Toxicology (COMET), a collaboration between five pharmaceutical companies and Imperial College London. COMET produced a database of NMR spectra of rodent urine and blood serum, taken from animals that had been dosed with a range of toxins. This database now forms the basis of a successful system for predicting the liver and kidney toxicity of drug candidates. A follow-up project, COMET-2, is currently investigating the detailed biochemical mechanisms of toxicity, and seeks a better understanding of inter-subject variation in metabonomic analyses. It has also been demonstrated in animals that the metabolic profile of an individual's urine can be used to predict both how that individual will metabolize a given drug and their susceptibility to the side effects of that drug. If this principle can be applied widely in humans, it will have enormous implications for personalized health care and in optimizing clinical trials.

What insights into diseases have emerged?

Many metabolites have been identified as flags for a variety of diseases. They include markers for schizophrenia found in cerebrospinal fluid; markers of coronary-artery occlusion found in plasma; and even indicators of spinal-trauma-induced infertility in men, found in seminal fluid. The concentrations of these metabolites often vary in response to therapy for the disease in question. Furthermore, the biomarkers carry information about the sites and mechanisms of disease. Metabolites have also been found that act as indicators for disease risk, individual susceptibility, or as markers of recovery from an illness.

What else has metabonomics taught us?

It has revealed much about humans' symbiotic relationship with their gut flora. Disruption of gut microbial activity seems to be central to certain diseases in the gut, liver, pancreas and even the brain. But there are thousands of species of microorganism (the microbiome) in the human gut, and it is impossible to study each one in isolation to work out what they do. Large research programmes have therefore been devised to study the combined genetic structure of humans and their microbes — the metagenome. This genetic information is invaluable, but it says nothing about the actual activity of the microbial community, or its interactions with the host at a physiological level. Metabonomic modelling, statistically coupled to metagenomic analysis, has allowed possible communication networks between species to be identified, and has also shown which metabolic pathways are strongly influenced by which members of the microbiome.

What will be the next big thing in metabonomics?

We believe it will be metabolome-wide association (MWA) studies, which identify relationships between metabolic profiles and disease risks for both individuals and populations. In this approach, the metabolic profiles of thousands of people are captured spectroscopically, and are then statistically linked to disease risk factors such as obesity and diabetes. The beauty of MWA is that vast, well-curated collections of biological-fluid samples are available from epidemiological studies of many disease processes. These can be explored retrospectively for markers of disease risk. We need such indicators as part of a strategy for disease prevention, which is essential to drive down health costs as the world's population increases. Effective prevention requires knowledge of risk, which for most modern diseases involves unfavourable gene–environment interactions. Understanding how these interactions affect metabolic regulation and phenotype allows new testable hypotheses to be developed about future disease risk. It is in this field that metabonomics might prove to be strongest and of most value to humanity. ■

Jeremy K. Nicholson and John C. Lindon are in the Department of Biomolecular Medicine, Imperial College London, Sir Alexander Fleming Building, South Kensington, London SW7 2AZ, UK.
e-mails: j.nicholson@imperial.ac.uk; j.lindon@imperial.ac.uk

FURTHER READING
Nicholson, J. K. & Wilson, I. D. Understanding 'global' systems biology: Metabonomics and the continuum of metabolism. *Nature Rev. Drug Discov.* **2,** 668–676 (2003).
Dethlefsen, L., McFall-Ngai, M. & Relman, D. A. An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature* **449,** 811–818 (2007).
Lenz, E. M. & Wilson, I. D. Analytical strategies in metabonomics. *J. Proteome Res.* **6,** 443–458 (2007).
Wishart, D. S. *et al.* HMDB: The Human Metabolome Database. *Nucleic Acids Res.* **35,** D521–D526 (2007).
Holmes, E., Wilson, I. D. & Nicholson, J. K. Metabolic phenotyping in health and disease. *Cell* **134,** 714–717 (2008).
Holmes, E. *et al.* Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* **453,** 396–400 (2008).