INNOVATION

Metabonomics: a platform for studying drug toxicity and gene function

Jeremy K. Nicholson, John Connelly, John C. Lindon and Elaine Holmes

The later that a molecule or molecular class is lost from the drug development pipeline, the higher the financial cost. Minimizing attrition is therefore one of the most important aims of a pharmaceutical discovery programme. Novel technologies that increase the probability of making the right choice early save resources, and promote safety, efficacy and profitability. Metabonomics is a systems approach for studying *in vivo* metabolic profiles, which promises to provide information on drug toxicity, disease processes and gene function at several stages in the discovery-and-development process.

The 'old' testing paradigms often fail to prevent many 'doomed' molecules from entering development. There is a need for methodologies that can describe altered gene expression and cellular protein profiles in terms of their early metabolic consequences, and relate these to developing, established or regressing pathology. Studying the effects of drugs on whole organisms by metabonomics relies on multiparametric measurement of alterations in metabolism over time in response to a stressor or intervention[1–10]. This approach can also be readily adapted to investigate the functional consequences of genetic variation and transgenesis[3,11–13], which is potentially of great importance in the creation and validation of new models of human disease and efficacy. There is considerable scope for the application of metabonomic approaches in the pharmaceutical industry, from discovery through to clinical development and beyond. In the discovery phase, these include early *in vivo* toxicological testing, lead compound selection and pre-lead prioritization, and *in vivo* efficacy screening in animal models. In the development phase, applications include finding new preclinical safety biomarkers and mechanisms, metabotyping and the validation of animal models against human disease profiles, and the discovery of new clinical safety and efficacy biomarkers.

In this article, we will explore the relationships between the findings of the new 'omics' sciences, describe analytical technologies for measuring multiparametric metabolic responses, with particular emphasis on NMR-based approaches, discuss multivariate statistical methods that optimize information recovery from multivariate NMR data sets, and consider the application of metabonomics to drug discovery and development, particularly in minimizing attrition.

Metabonomics in pathophysiology

The realization that obtaining the genome sequence of humans or other species does not in itself explain the fundamental nature of many disease processes has triggered a marked increase in interest in approaches that relate gene expression to phenotypic outcome[1,2,14–16]. Several technologies are being developed to achieve this end, namely: genomics and transcriptomics, which examine genetic complement and gene expression, respectively; proteomics, which involves the analysis of protein synthesis and cell signalling; metabolomics, which investigates metabolic regulation and fluxes in individual cells or cell types; and metabonomics — the determination of systemic biochemical profiles and regulation of function in whole organisms by analysing biofluids and tissues.

In complex organisms, these levels of biomolecular organization and control are interdependent, and are affected by environmental events and stresses throughout life. Their characterization, by appropriate analytical methods, describes changes in biological activity by using complex multivariate data sets that can be analysed using various chemometric and bioinformatic tools[15,17,18]. The aim of such procedures is to extract latent biochemical information that is of diagnostic or prognostic value, and which reflects 'actual' biological events rather than the 'potential' for disease or toxicity that is offered by the collection of gene expression and proteomic data after exposure to a drug or stressor. It is therefore necessary to relate real-world or end-point observations to the measurements that are provided by the 'omics' technologies (FIG. 1). This allows an understanding of the relationships between the inputs that change 'omics' responses and the outputs of those responses. Metabolic changes are real-world end points, whereas gene expression changes are not; they merely indicate the potential for an end-point change. As such, metabonomics provides a useful connection between the 'omics' platforms and actual tissue histology.

Post-genomic analytical technologies

To investigate the complex metabolic consequences of disease processes, toxic reactions and genetic manipulation, non-selective, but specific, 'information-rich' analytical approaches are required (BOX 1). Several spectroscopic methods in addition to NMR can produce metabolic signatures of biomaterials, including mass spectrometry (MS)[19,20], gas chromatography/mass spectrometry (GC/MS)[21], high-performance liquid chromatography (HPLC)[22,23] and optical spectroscopic techniques[24]. Bioanalytically, NMR and MS are powerful means of generating multivariate metabolic data. NMR has the advantages of being non-destructive, applicable to intact biomaterials, and intrinsically more information rich with respect to the determination of molecular structures, especially in complex-mixture analyses. Furthermore, a technique known as magic angle spinning (MAS)-NMR can be used to carry out biochemical studies on intact tissues and cells, which, if carefully conducted, can preserve the samples for other studies and allow abnormal molecular compartmentation and interactions to be studied in intact tissues. MS is analytically more sensitive than NMR, but differential ionization suppression can make pattern quantification difficult, and extraction and derivatization might be necessary. The choice between NMR and MS approaches is ultimately matrix or problem dependent. Both technologies require further development, especially of high-throughput and data-processing methods, to optimize their use in complex metabolism studies. However, ^{1}H-NMR spectroscopy can be efficiently applied to

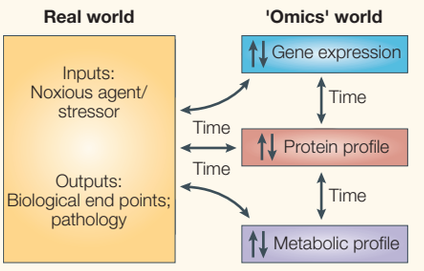
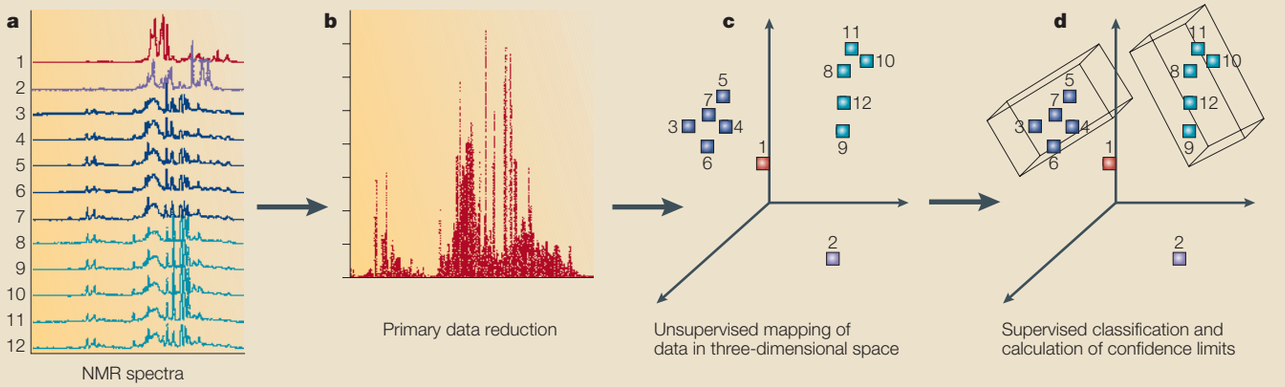


Figure 1 | **Relating the real world to the 'omics' world.** Relationships between real-world inputs (stressors) and outputs, and their time-displaced connections with the measured 'omics' responses. Note time differentials at all interaction stages.

Box 1 | **Important techniques and procedures in metabonomics**



NMR spectroscopy. Some atomic nuclei possess a non-zero magnetic moment. This property is quantized, and leads to discrete energy states in a magnetic field. Nuclei such as ^{1}H, ^{13}C, ^{15}N, ^{19}F and ^{31}P can undergo transitions between these states when radio-frequency pulses of appropriate energy are applied. The exact frequency of a transition depends on the type of nucleus and its electronic environment in a molecule. For example, ^{1}H nuclei in a molecule give NMR peaks at frequencies (chemical shifts) that are characteristic of their chemical environment. NMR spectroscopy is used extensively as a structural tool, and information on isomers and molecular conformations can be obtained by interpretation of the chemical shifts as well as the splitting patterns due to indirect nuclear interactions (J-couplings). In metabonomics, the patterns that occur with time when many biochemical entities are detected simultaneously in the mixture by NMR are interpreted[63,82].

Pattern-recognition methods. Pattern recognition and related multivariate statistical approaches can be used to discern significant patterns in complex data sets, and are particularly appropriate in situations in which there are more variables than samples in the data set. The general aim of pattern recognition is to classify objects — in this case, ^{1}H-NMR spectra — or to predict the origin of objects, by identifying inherent patterns in a set of indirect measurements. Pattern-recognition methods can reduce the dimensionality of complex data sets by means of two- or three-dimensional mapping procedures, thereby facilitating the visualization of inherent patterns in the data.

Supervised and unsupervised techniques. Methods such as principal components analysis (PCA) are termed 'unsupervised' techniques, in that no *a priori* knowledge of the class of the samples is required, and they are based on the calculation of latent variables. Principal components are linear combinations of the original descriptors, such that they are uncorrelated, and describe decreasing amounts of data variance (that is, PC1>PC2>PC3 and so on). Use of PCA enables the 'best' representation, in terms of biochemical variation in the data set, to be shown in two or three dimensions. In addition, multiparametric data can be modelled, so that the class of a sample from an independent data set can be predicted on the basis of a series of mathematical models that are derived from the original data or 'training' set. These methods are known as 'supervised' methods, and use class information to maximize the separation between classes. Supervised methods, such as soft independent modelling of classification analogy (SIMCA), partial least squares (PLS) analysis and PLS discriminant analysis (PLS-DA), can be used to predict objects that are unknown to the system on the basis of their NMR spectral properties, and are therefore valuable for generating models for predicting drug toxicity[83–86].

Strategy for metabonomic analysis. An NMR data-analysis procedure is shown in the figure. After spectra are accumulated and processed (panel **a**), a primary data reduction is carried out that digitizes the one-dimensional spectrum into a series of typically 250–1,000 integrated regions (panel **b**)[36,59]. After removal of redundant signals and appropriate scaling, primary data analysis is used to map the samples according to their biochemical composition, using methods such as PCA. Samples that are generated from animals that are in a similar pathophysiological state are generally intrinsically similar in composition, and therefore occupy neighbouring positions in the PC space (panel **c**). Each class of samples is then modelled separately, and class boundaries and confidence limits are calculated to construct a model for the prediction of independent data (panel **d**)[11].

measure the metabolite profiles of biofluids and tissues. NMR measurements, coupled with multivariate statistical, chemometric methods for the purpose of latent-information extraction and sample classification, offer a powerful new approach to whole-system diagnostics and metabolic function. The efficient application of NMR-based metabonomics in toxicological and clinical investigations has been shown[1,2,25–40], and we have recently shown its fundamental value in characterizing the metabolic consequences of genetic variation in mammalian systems, and in identifying the 'metabotypes' or metabolic phenotypes that result from a combination of genetic and environmental factors[3,11,12,41,42]. Metabonomics is also a powerful tool for investigating phenotypic abnormalities in mutant animals[13] and human diseases[32,43–48], and in modelling physiological variation in experimental animals and man[42,49].

Biochemical changes over time

In an integrated biosystem, it is axiomatic that the initiation of functionally connected gene expression events, cell signalling, protein-synthesis changes and metabolic responses to a stressor must be essentially sequential. Maturation and persistence of changes in gene expression, protein synthesis and post-translational modification, and subsequent effects on metabolic processes, also differ significantly. An event must therefore be evaluated in relation to time at each level of biomolecular organization if molecular responses are to be accurately associated with their macroscopic consequences in an organism. So, in metabolic studies, it is of paramount importance to measure time-dependent patterns of change in response to stimuli, because metabolic fluxes occur very rapidly, even in normal homeostasis, and consideration of the 'metabolite content' at only a fixed point in time can be misleading. The same considerations also apply to proteomic and transcriptomic data. Levels of messenger RNA and associated protein products might correlate poorly, even in carefully conducted studies on

unicellular organisms, such as yeast[50] and bacteria[51], or in higher organisms[52]. In fact, it might not be appropriate to try to identify simple correlations between transcriptomic and proteomic data that are collected at single time points after exposure to a stressor, as these might not exist at all because of the non-linearity of many gene expression and protein synthesis relationships.

In multicellular organisms, these time differentials will probably vary unpredictably with gene and tissue, and modulation by other external stressors in addition to, for example, a drug, is possible. It is important, therefore, to choose appropriate windows of time to study transcriptomic and proteomic responses to exposure to drug candidates. An important potential role for metabonomics is to direct the use and timing of proteomic and genomic analyses in order to maximize the probability of observing biological transitions that predict functional outcomes; this principle applies to human, animal and microbial systems. In the case of a single exposure to a toxic drug, there will be a response that takes time to complete, and the patterns that are observed in gene expression, proteins and metabolites will therefore vary according to when the measurements are made (FIG. 2a). If the measurements are made long after dosing, it is possible that the only profile changes will be due to biomarkers of recovery or cellular repair. In the case of a multi-dose study (FIG. 2b), the second and subsequent doses of a compound might arrive before the effects of the first dose are cleared, complicating the profiles further. As the doses continue, there might be a rising curve of toxicity (measured by whatever means), and there could then be superimposed profile changes due to cell death and regeneration.

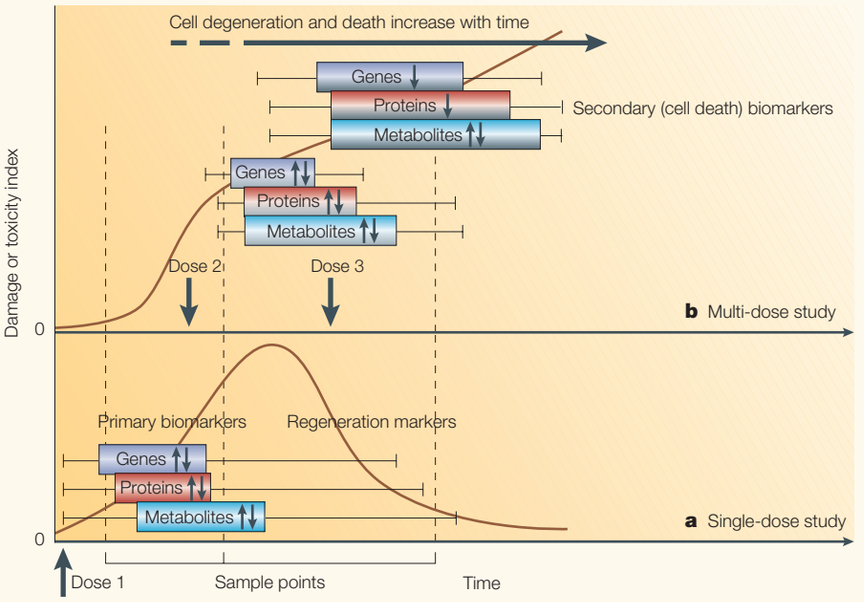


Figure 2 | **Effect of time in toxicological studies.** Theoretical relationships between dosing and 'omics' responses in **a** | single-dose and **b** | multi-dose situations. The observed changes in gene expression, proteomic or metabolic profiles are dependent on when samples are taken, and will vary in a complex way according to the severity and time taken to develop maximal pathology (for example, cell damage as determined by histology) — shown for simplicity as a bell-curve in part **a** and as a rising curve in part **b**. The horizontal bars reflect the ranges of times over which up- and downregulation of responses can occur. In a single-dose study, measurement points before maximal damage reflect primary biomarkers of toxicity. Those measured after maximal damage must also contain biomarkers that relate to cellular regeneration. In a multi-dose study, second and subsequent doses change the biomarker responses, as the system is not at equilibrium before the next dose is given. Effects might also change due to enzyme-induction processes that change the drug-metabolite patterns. As cell death increases, the 'omics' responses increasingly reflect end-stage changes that carry little mechanistic information.

^{1}H-NMR spectroscopy of biomatrices

^{1}H-NMR spectroscopy of biofluids such as urine and plasma has been successfully applied to investigate numerous diseases and toxic processes[53–62]. Because biofluids fulfil diverse biological purposes, their metabolic composition varies with their role and the functional integrity of the organ systems that are communicating with them, and ultimately with the physiological status of the whole organism. So, each biofluid yields a characteristic ^{1}H-NMR spectroscopic fingerprint in which the spectral intensity distribution is determined by the relative concentrations of solutes, and in some cases by their intermolecular interactions[61,63]. High-frequency ^{1}H-NMR spectroscopy is particularly useful in biochemical investigations, in that it is sensitive (low-nanogram detection limits are possible with appropriate instrumentation), and nearly all metabolic intermediates have unique ^{1}H-NMR signatures[61,63]. Simple one-dimensional spectra typically take only a few minutes to acquire, with no sample preparation other than buffering and addition of D_2O to provide a reference frequency. The large interfering signal that arises from water in all biofluids is easily eliminated using appropriate solvent-suppression methods[61].

Metabolic profiles of biofluids such as plasma, cerebrospinal fluid and urine reflect both normal variation and the pathophysiological impact of toxicity or disease on single or multiple organ systems[53,55,64,65]. Urine and plasma are obtained in a non- or minimally invasive manner, and are therefore appropriate for clinical-trial monitoring and disease diagnosis. Even a one-dimensional high-frequency ^{1}H-NMR spectra (600 MHz or greater) of urine typically contains many thousands of sharp lines from hundreds or potentially thousands of metabolites[63]. ^{1}H-NMR spectra of urine are dominated by low-molecular-weight compounds, whereas plasma contains both low- and high-molecular-weight components, which give a wide range of signal line widths: protein and lipoprotein signals dominate simple one-dimensional ^{1}H-NMR spectra, with small-molecule fingerprints superimposed on them[65]. Spin-echo experiments minimize the broad signal contributions from proteins, other macromolecules and micelles[53,65].

An important challenge posed by NMR spectroscopy of biofluids is how to efficiently recover metabolic information that allows diagnosis or classification of disease or toxicity. The information that is needed for useful classification of spectra from biological samples can often be obtained without detailed structural chemical (spectroscopic) analyses. More sophisticated interrogation involves the identification of the molecular species that differentiate 'pathological' from 'normal' states, and a statistical description of the biomarkers that aid understanding of disease or toxic mechanisms. So, NMR spectra of biofluids serve in two distinct, but closely related, modes; that is, as quantitative metabolic-fingerprinting tools, and as a means of determining metabolite (biomarker) structure[63].

The detailed structural assignment of a biofluid NMR spectrum can be a complex procedure, involving the application of several

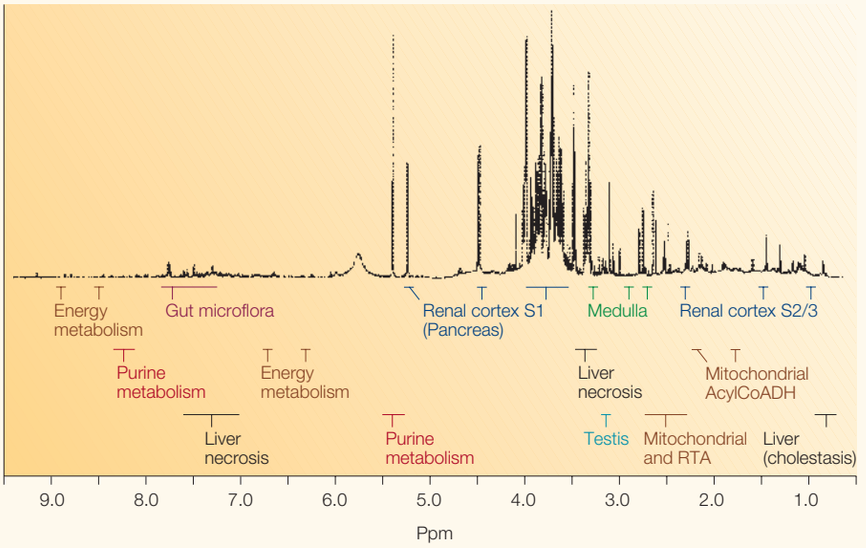


Figure 3 | **A functional NMR spectrum of rat urine.** A 600-MHz ^{1}H-NMR spectrum of rat urine is shown, which is colour-coded to indicate spectral biomarker windows that are diagnostic for a subset of diverse pathophysiological conditions. One NMR spectrum can carry information on a wide range of pathological or toxic processes (potentially hundreds of disease classes in a single spectral measurement). AcylCoADH, acyl coenzyme A dehydrogenase dysfunction; RTA, renal tubular acidosis.

NMR techniques, including ^{1}H–^{1}H and ^{1}H–^{13}C two-dimensional experiments[65]. Diffusion-ordered and diffusion-edited NMR spectra can also be of value, especially when solutes cover a wide range of molecular weights, as with blood plasma[66]. Directly coupled chromatography–NMR spectroscopic methods can be used, especially for determining the structure of drug metabolites — the most powerful of these 'hyphenated' approaches is HPLC–NMR–MS[67]. Complete assignment of NMR spectra of biofluids is extremely difficult, because NMR-detectable metabolite concentrations (excluding water) vary over six orders of magnitude. This causes difficulties in the interpretation of two-dimensional spectra of minor components in the presence of peaks from compounds at much higher concentrations. Furthermore, unlike single proteins, which also give complex spectra, it is not possible to uniformly enrich biofluid metabolite mixtures with ^{13}C or ^{15}N labels, limiting the range of heteronuclear-correlation NMR methods that can be used. In practice, however, neither issue markedly limits the usefulness of ^{1}H-NMR spectroscopy of biofluids as a diagnostic and biomarker-identification tool. Biofluid NMR is useful for rapid screening, especially when carried out using flow-injection methods[68]. However, target-organ toxicity or disease fingerprints can also be investigated by using MAS-NMR of intact tissues[69–76], and this can give deep insight into toxic mechanisms, such as cadmium-induced nephrotoxicity[26,77].

Metabonomics in drug toxicology

Although transcriptomic/genomic and proteomic measurements respond to the administration of toxic agents, it is difficult to relate findings to classical toxicological end points, and hence influence the drug attrition rate. Metabonomics offers a complementary approach that gives information on whole-organism functional integrity over time after drug exposure. Target tissues or processes, and biomarkers, can be identified by characteristic changes in the pattern of concentrations of endogenous metabolites in biofluids that relate to the site and mechanism of toxicity.

^{1}H-NMR analysis of biofluids has uncovered novel metabolic markers of organ-specific toxicity in the laboratory rodent, and this 'exploratory' role is one in which biofluid NMR spectroscopy excels. For example, a combination of changes in the urinary levels of trimethylamine-*N*-oxide, *N*,*N*-dimethylglycine, dimethylamine and succinate, together indicate renal papillary damage, for which no biochemical markers existed previously[36,38]. The biomarker information that is present in the NMR spectra of biofluids is potentially very rich, as hundreds of compounds that represent a variety of metabolic pathways are measured simultaneously[61,63]. The NMR spectrum of a biofluid can be conveniently thought of as a series of 'biomarker windows', which are spectral regions that contain signals from metabolites that are associated with specific targets for toxicity. An NMR spectrum of rat urine, with regions labelled by the biological or functional changes in the animal, is shown in FIG. 3, and contrasts with the conventional way of labelling an NMR spectrum according to molecular structure. This diagram shows a series of 'biomarker windows', in which there are combinations of pattern and intensity changes according to the site and mechanism of organ dysfunction. Many such windows exist in NMR spectra of biofluids, and diagnostic information on hundreds of different types of disease or toxic process can potentially be derived from a simple NMR measurement, as shown in FIG. 4 using examples of liver and kidney toxicity.

Although toxins can affect gene regulation or expression directly, significant responses might be completely unrelated to gene switching (for example, enzyme inhibitors). In such cases, transcriptomic and proteomic methods are likely to be unrevealing. However, many drug-induced effects involve disturbed endogenous metabolite concentration fluxes or ratios that result from direct chemical reactions, altered binding to macromolecules, modified control mechanisms, and induction or inhibition of enzymes, to name but a few. If these disturbances overwhelm compensatory or adaptive mechanisms, consequences that are recognized as toxicity occur. As metabolite concentrations in several key body fluids relate to cell and tissue processes, so toxin- or disease-induced disequilibria are reflected in those fluids.

Using pattern-recognition methods, NMR spectra can be used to: classify the sample as being normal or abnormal (this is useful in the control of spectrometer automation using flow injection); classify target-organ toxicity and the site and mechanism of action within the organ; identify biomarkers of toxic effect; and evaluate the time course of the effect; for example, the onset, evolution and regression of toxicity. The information that is derived from databases of NMR spectra can be maximized using appropriate chemometric and multivariate analytical strategies. Preliminary analysis involves the application of 'unsupervised' pattern-recognition methods, such as principal components analysis (PCA) or cluster algorithms that assume no previous knowledge of sample class. Specific toxic challenges can be characterized by PCA trajectories, in which clustering of sample coordinates reflects an intrinsic similarity in biochemical composition. Such trajectories are shown in BOX 2 for a model liver toxin and its metabolites. Having established the presence of site- or mechanism-related metabolic responses, more sophisticated supervised algorithms can be applied to the

data, in which the sample class can be used to optimize the differentiation between classes. Information that relates to biomarkers of toxicity or recovery can be extracted from the analysis with a view to furthering our understanding of the mechanisms of toxicity[78].

Time-resolved metabonomic experiments can also be used to deconvolve the overlapping biochemical effects of drugs and their metabolites, which might have different toxicities.

As with all 'omics' platforms, metabonomics has certain limitations in terms of the recovery of biological information. In the case of toxicity assessment, it is possible to generate false-positive data in situations in which the compound of interest causes significant metabolic changes without associated toxicity, because of marked physiological

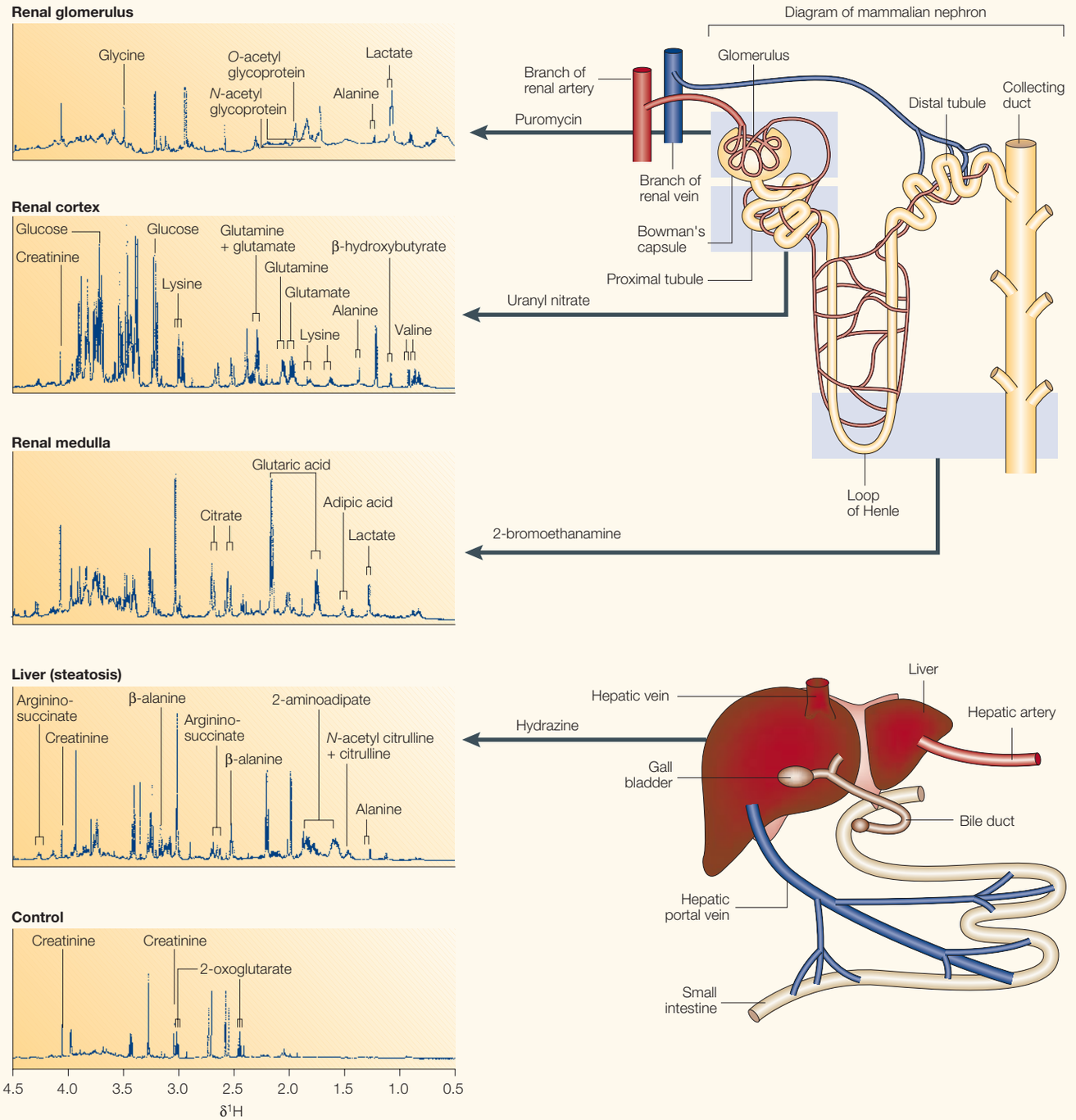
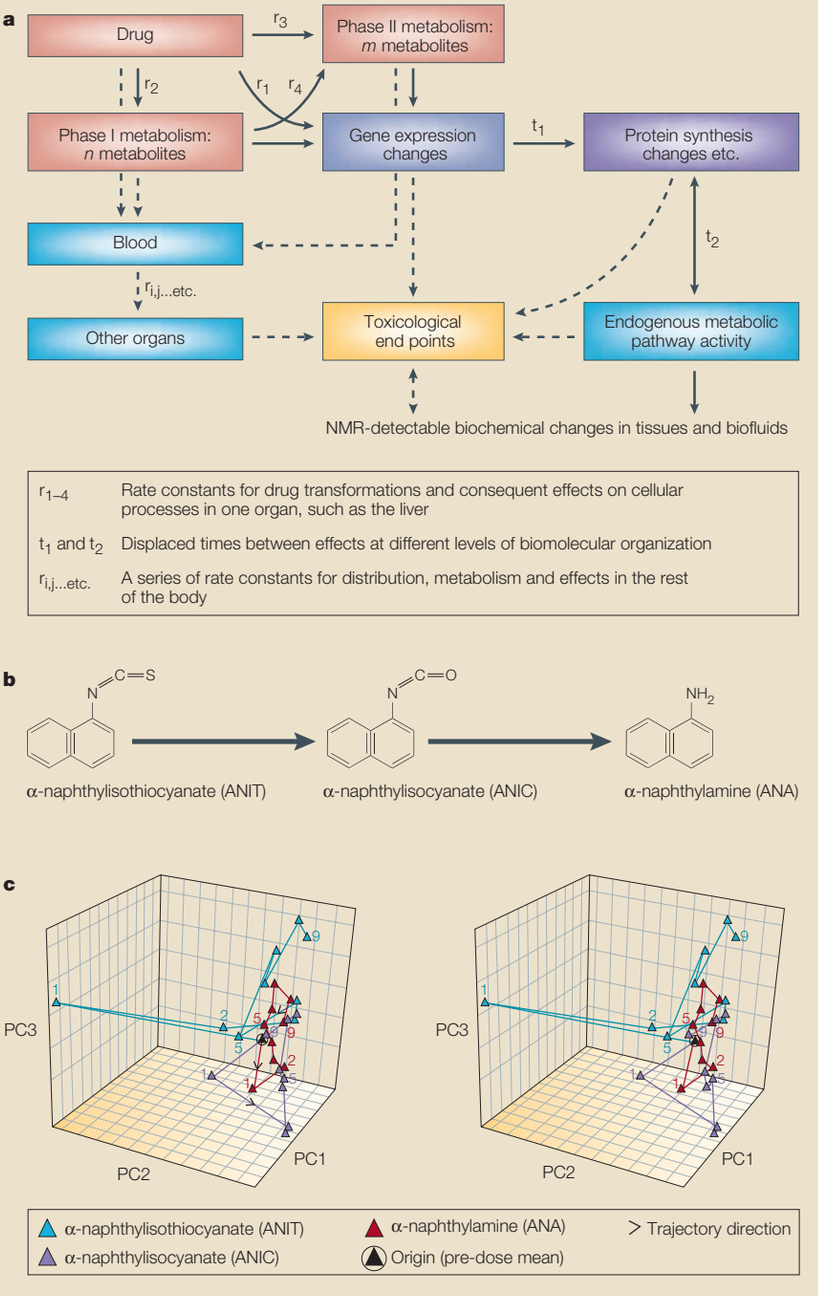


Figure 4 | **Metabonomic detection of liver and kidney toxicity.** Stack plot of NMR spectra showing characteristic metabolic fingerprints of tissue-specific toxicity produced by different site-selective xenobiotics given in single doses to rats in relation to an untreated control. Each 600-MHz spectrum represents one time point after dosing for each toxic compound. The xenobiotics affect specific regions within the organs (depicted in the figure by shaded boxes): puromycin affects the renal glomeruli; uranyl nitrate affects the lower regions of the proximal tubules; 2-bromoethanamine affects the renal medulla, including the loop of Henle and the collecting ducts; and hydrazine affects the hepatic parenchymal cells.

Box 2 | **Convolution of biochemical and molecular events caused by drug metabolism**

Most drugs are extensively biotransformed to metabolites with markedly different toxicological properties. Such biotransformations take variable times in different tissues and in different species, convolving the cellular toxicological effects. Drug/metabolite proportions also change with time as the parent drug becomes absorbed, distributed and cleared by metabolism and excretion (panel a). This is an important problem, as snapshots of net toxic effect using any 'omics' approach cannot be readily interpreted, and this is illustrated with respect to the hepatotoxin α-naphthylisothiocyanate, which undergoes the changes shown in panel b.

When administered separately to rats, these compounds produce markedly different metabolic trajectories, as they act by different mechanisms[36]. Trajectory plots map the metabolic response, show the extent and type of the lesion, and indicate whether there is functional recovery[36,59]. A three-dimensional principal components analysis (PCA) stereo pair plot, which consists of two images that are designed to be viewed with stereo glasses to give a three-dimensional effect, shows trajectories for each compound (panel c). Each connected point represents the mean metabolic position of five animals at given time points (numbered 1–9 in the diagram) after dosing. The α-naphthylisothiocyanate (ANIT) trajectory represents the aggregate biochemical effects of the metabolic flux (ANIT to α-naphthylisocyanate (ANIC) to α-naphthylamine (ANA)). There is no functional recovery of the ANIT-induced lesion over the seven-day course of the experiment (the trajectory does not return to origin). The ANIC trajectory is simpler, and is convolved only by its primary metabolite, ANA — there is functional recovery. The first metabolic direction of the ANA-dosed animals is in the third principal component (PC3). For ANIC, the second direction is in PC3, as it takes time for ANA to be produced. For ANIT, the PC3 direction change takes longer, as ANA is formed after ANIC. This approach, termed 'metabolic trajectory deconvolution', allows effects of drugs and their metabolites to be separated for mechanistic purposes. Endogenous metabolic changes that are caused by the effects of successive drug metabolites must also mirror complex time-related changes in the transcriptomic and proteomic patterns, thus limiting the value of single-time-point measurements of genes and proteins.



or pharmacological effects. For example, acetazolamide is a renal carbonic anhydrase inhibitor that massively reduces the excretion of intermediates in the citric-acid cycle. Misinterpretation can, however, be minimized by using supervised methods that include models of such effects — such models are now under construction in our laboratory. Conversely, certain pathologies, such as liver fibrosis, are associated with negligible effects on biofluids, as metabolic derangement does not occur until there is significant tissue damage. In the case of low-potency compounds, there might be particular difficulties in separating toxicological from physiological effects. However, in previous dose–response studies, NMR-based metabonomic methods were at least as sensitive as conventional methods for detecting lesions at the 'threshold-dose' level[26,27,56,61], and even minor physiological changes were detected in normal animals[41,42,49].

There are obvious limitations in terms of choice of biofluid; for instance, urine might not be as appropriate as cerebrospinal fluid for studying neuropathology. There is also the potential for confusion with mixed-toxicity drugs that, for example, affect both liver and kidney, as the biomarkers of toxicity will be a complex combination that relates to both sites and possibly to multiple mechanisms. However, this is offset by the fact that mixed toxicities often have different timescales, and

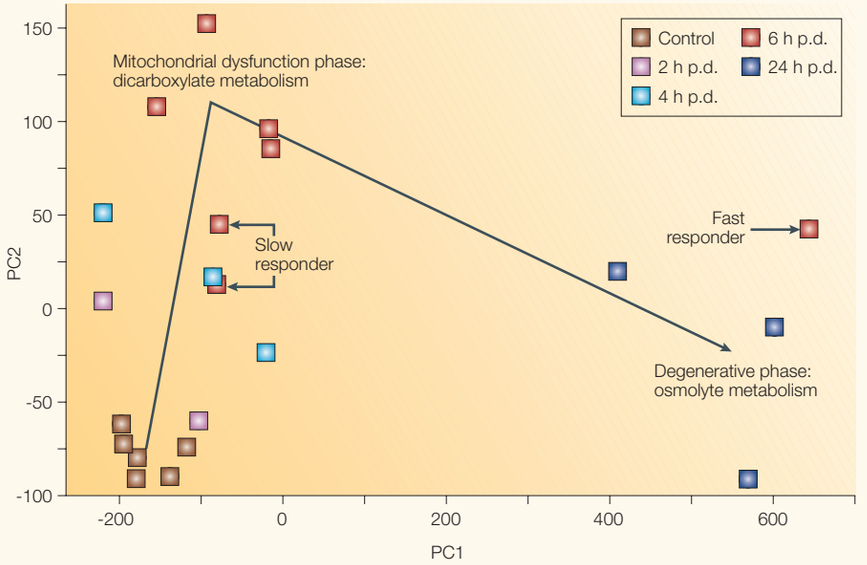


Figure 5 | **Principal components analysis (PCA) of rat renal papillary damage.** The figure depicts a PCA map that shows the trajectory of biochemical changes in the intact rat renal papilla (detected using magic angle spinning (MAS)-NMR) after the administration of 2-bromoethanamine. Some animals respond to the intoxication faster than others, even though they are of uniform age and sex and were raised under the same conditions. This is a typical type of response, with 'slow' and 'fast' responders being characteristic of many drugs and toxins. The coloured boxes refer to the sampling time point after treatment. PC, principal component; p.d., post-dosing.

such effects can therefore be deconvolved by making repeated sequential measurements in individual animals. The distinction between adaptive and toxic effects remains a challenge with all the 'omics' platforms. This applies to the variables that change due to the initial adverse (mechanism-related) interactions, the homeostatic response to the cellular derangement (which could reflect entirely positive reactions of a healthy cell or tissue), and the changes due to cell death. Although it is possible that these responses might all occur at the same time in a complex organism, there will be a progressive change in their contributions over time, which could hold the key to their future deconvolution.

MAS-NMR spectroscopic measurements on intact tissues allow the direct correlation of tissue biomarkers with histological change. The development of a renal papillary lesion, as observed by MAS-NMR analysis of intact renal papilla samples after the administration of 2-bromoethanamine to rats[79], revealed a consistent response over time, with the exception of a few rats that were deemed to be either 'fast' or 'slow' responders to the treatment[79] (FIG. 5). Indeed, the metabonomic determination of individual differences in response to drug therapies offers great potential. We have recently shown the value of obtaining several NMR data sets from biofluid samples and tissues of the same animals collected at different time points. This procedure is termed 'integrated metabonomics'[80], and can be used to describe the changes in metabolic chemistry in different body compartments that are caused by exposure to toxic drugs. Such timed profiles in multiple compartments are themselves characteristic of particular types and mechanisms of toxicity, and can be used to give a more complete description of the consequences of toxicity than can be obtained from one fluid or tissue alone. The use of integrated metabonomics in relation to the conventional screening procedures that are now used in drug discovery is illustrated in FIG. 6.

Metabonomics in functional genomics

Metabonomics can be used to separate classes of experimental animals, such as mice and rats, according to their strain on the basis of the endogenous metabolite patterns in their biofluids[3,11–13]. This is possible because differences in 'silent-gene' function between strains can influence the fluxes of metabolites through many key intermediary pathways, resulting in distinct animal 'metabotypes'[11,12]. Such differences might help to explain the differential toxicity of drugs between strains in which the metabolic fate and receptor populations, and hence potential toxicity of the drug itself, is linked to the activity of endogenous pathways. There is also a strong indication for the use of metabonomics in the phenotyping of mutant or transgenic animals and the investigation of the consequences of transgenesis. In a recent study, we have shown that the transfection process itself can cause marked metabolic differences in hepatoma cell lines due to the disruption of host cell membranes[81]. It is important to differentiate such unintended consequences of the genetic engineering process from the intended result, as these can potentially confuse the interpretation of the function of particular genes or gene classes when the cell system or organism is examined for the physical effects of the intervention. This is important to pharmaceutical companies that are trying to genetically engineer new animal models of disease using biochemically invasive transfection procedures. Furthermore, metabonomic approaches can give deep insight into the metabolic similarities or differences between mutant or transgenic animals, and the human disease processes that they are actually intended to simulate. If the veracity of an animal model can be established using metabolic criteria — that is, biomarkers of the disease process — then it might also be possible to monitor the efficacy of novel therapeutic agents (normalization of the biochemical profile) using metabonomic criteria. Such approaches might be of great future value to the pharmaceutical industry in the quest for discovering safe and efficacious new drugs.

"…integration of data types will also pave the way to understanding the relationships between gene function and metabolic control in health and disease."

Conclusions

Metabonomics is now recognized as an independent and widely used technique for evaluating the toxicity of drug-candidate compounds, and has been adopted by several pharmaceutical companies into their drug development protocols. It is possible to identify the target organ of the toxicity, derive the biochemical mechanism of the toxicity, and determine the combination of biochemical biomarkers for the onset, progression and regression of the lesion. Furthermore, the

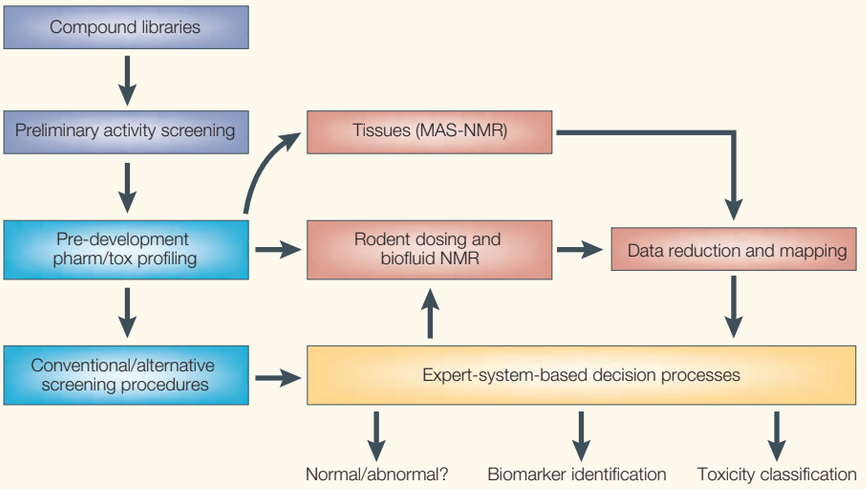


Figure 6 | **Integrated metabonomics.** Incorporation of metabonomics-based expert systems for toxicity classification (presence of abnormality, type of abnormality and biomarkers of abnormality) in relation to other drug discovery procedures.

technique can provide a metabolic fingerprint of an organism (metabotyping) — a key to functional genomics — and hence has applications in the design of drug clinical trials and the evaluation of genetically modified animals as disease models. Finally, using metabonomics, it has proved possible to derive new biochemically based assays for disease diagnosis, and to identify combination biomarkers for disease, which can then be used to monitor the efficacy of drugs in clinical trials. Clearly, metabonomics is not a panacea for all future drug-safety studies and clinical investigations, and will complement, rather than entirely supplant, conventional methods, particularly bearing in mind the needs of regulatory bodies. There are certainly particular areas within drug discovery for which metabonomics will probably not prove useful, because of sensitivity or specificity issues, and these limitations require further exploration. One potential area of importance in the future is an understanding of the reasons for the rare human 'idiosyncratic' toxicity of certain drugs that are already on the market or are in late-stage development. It is possible to foresee future situations in which metabonomic, gene expression and proteomic data are interrogated using multivariate approaches to provide a holistic picture of complex organisms that are undergoing physiological stress, using optimized biomarker combinations from all three platforms. Such integration of data types will also pave the way to understanding the relationships between gene function and metabolic control in health and disease.

Jeremy K. Nicholson, John Connelly, John C. Lindon and Elaine Holmes are at the Biological Chemistry Section, Biomedical Sciences Division, Faculty of Medicine, Imperial College of Science, Technology and Medicine, South Kensington, London SW7 2AZ, UK. Correspondence to J.K.N.
e-mail: j.nicholson@ic.ac.uk

DOI: 10.1038/nrd728

1. Nicholson, J. K., Lindon, J. C. & Holmes, E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR data. *Xenobiotica* **29**, 1181–1189 (1999).
2. Lindon, J. C., Nicholson, J. K., Holmes, E. & Everett, J. R. Metabonomics: metabolic processes studied by NMR spectroscopy of biofluids. *Concepts Magn. Reson.* **12**, 289–320 (2000).
3. Holmes, E., Nicholson, J. K. & Tranter, G. Metabonomic classification of genetic variations in toxicological and metabolic responses using probabalistic neural networks. *Chem. Res. Toxicol.* **48**, 182–191 (2001).
4. Robertson, D. G. *et al.* Metabonomics: evaluation of nuclear magnetic resonance (NMR) and pattern recognition technology for rapid *in vivo* screening of liver and kidney toxicants. *Toxicol. Sci.* **57**, 326–337 (2000).
5. Gartland, K. P. R. *et al.* Pattern recognition analysis of high resolution ^{1}H NMR spectra of urine: a nonlinear mapping approach to the classification of toxicological data. *NMR Biomed.* **3**, 166–172 (1990).
6. Anthony, M. L., Beddell, C. R., Lindon, J. C. & Nicholson, J. K. Studies on the comparative toxicity of *S*-(1,2-dichlorovinyl)-L-cysteine, *S*-(1,2-dichlorovinyl)-homocysteine and 1,1,2-trichloro-3,3,3-trifluoro-1-propene in the Fischer 344 rat. *Arch. Toxicol.* **69**, 99–110 (1994).
7. Nicholson, J. K., Higham, D., Timbrell, J. A. & Sadler, P. J. Quantitative ^{1}H NMR urinalysis studies on the biochemical effects of acute cadmium exposure in the rat. *Mol. Pharmacol.* **36**, 398–404 (1989).
8. Griffin, J. L., Williams, H. J., Sang, E. & Nicholson, J. K. Abnormal lipid profile of dystrophic cardiac tissue as demonstrated by one- and two-dimensional magic angle spinning NMR spectroscopy. *Magn. Reson. Med.* **46**, 249–255 (2001).
9. Phipps, A. N., Stewart, J., Wright, B. & Wilson, I. D. Effect of diet on the urinary excretion of hippuric acid and other dietary derived aromatics in the rat. *Xenobiotica* **28**, 527–537 (1998).
10. Holmes, E. *et al.* Identification of novel biomarkers of renal toxicity using automatic data reduction techniques and PCA of proton NMR spectra of urine. *Chemometr. Intell. Lab. Syst.* **44**, 251–261 (1998).
11. Holmes, E. *et al.* Chemometric models for toxicity based on NMR spectra of biofluids. *Chem. Res. Toxicol.* **13**, 471–478 (2000).
12. Gavaghan, C. L., Holmes, E., Lenz, E., Wilson, I. D. & Nicholson, J. K. An NMR-based metabonomic approach to investigate the biochemical consequences of genetic strain differences; application to the C57BL10J and Alpk:ApfCD mouse. *FEBS Lett.* **484**, 169–174 (2000).
13. Griffin, J. L. *et al.* Metabolic profiling of genetic disorders: a multi-tissue ^{1}H NMR spectroscopic and pattern recognition study into dystrophic tissue. *Anal. Biochem.* **293**, 16–21 (2001).
14. Smith, L. L. Key challenges for toxicologists in the 21st century. *Trends Pharmacol. Sci.* **22**, 281–285 (2001).
15. Delneri, D., Brancia, F. L. & Oliver, S. G. Towards a truly integrative biology through the functional genomics of yeast. *Curr. Opin. Biotechnol.* **12**, 87–91 (2001).
16. Raamsdonk, L. M. *et al.* A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnol.* **19**, 45–50 (2001).
17. Kettaneh, N., Berglund, A. & Wold, S. in *Proceedings of the PLS '01 International Symposium* (eds Vinzi, V. E., Lauro, C., Morineau, A. & Tenenhaus, M.) 1 (CICIA Ceresta, Paris, 2001).
18. Johnston, M. The yeast genome: on the road to the Golden Age. *Curr. Opin. Genet. Dev.* **10**, 617–623 (2000).
19. Goodacre, R., Rooney, P. J. & Kell, D. B. Discrimination between methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* using pyrolysis mass spectrometry and artificial neural networks. *J. Antimicrob. Chemother.* **41**, 27–34 (1998).
20. Freeman, R. *et al.* Rapid identification of species within the *Mycobacterium tuberculosis* complex by artificial neural network analysis of pyrolysis mass spectra. *J. Med. Microbiol.* **40**, 170–173 (1994).
21. Halket, J. M. *et al.* Deconvolution of gas chromatography/mass spectrometry of urinary organic acids — potential for pattern recognition and automated identification of metabolic disorders. *Rapid Commun. Mass Spectrom.* **13**, 279–284 (1999).
22. Ramos, L. S. Characterisation of *Mycobacteria* species by HPLC and pattern recognition. *J. Chromatogr. Sci.* **32**, 219–227 (1994).
23. Vigneau-Calahan, K. E., Shestopalov, A. I., Milbury, P. E., Matson, W. R. & Kristal, B. S. Characterisation of diet-dependent metabolic serotypes: analytical and biological variability issues in rats. *J. Nutr.* **131**, S924–S932 (2001).
24. Winson, M. K. *et al.* Diffuse reflectance absorbance spectroscopy taking in chemometrics (DRASTIC). A hyperspectral FT-IR-based approach to rapid screening for metabolite overproduction. *Anal. Chim. Acta* **348**, 273–282 (1997).
25. Holmes, E. *et al.* Development of a model for classification of toxin-induced lesions using ^{1}H NMR spectroscopy of urine combined with pattern recognition. *NMR Biomed.* **11**, 1–10 (1998).
26. Griffin, J. L. *et al.* The initial pathogenesis of cadmium induced renal toxicity. *FEBS Lett.* **478**, 147–150 (2000).
27. Nicholls, A., Nicholson, J. K., Haselden, J. N. & Waterfield, C. J. A metabonomic approach to the investigation of drug-induced phospholipidosis. *Biomarkers* **5**, 410–423 (2000).
28. Espina, J. R. *et al.* Detection of *in vivo* biomarkers of phospholipidosis using NMR-based metabonomic approaches. *Magn. Reson. Chem.* **39**, 559–565 (2001).
29. Howells, S. L., Maxwell, R. J., Peet, A. C. & Griffiths, J. R. An investigation of tumour ^{1}H nuclear magnetic resonance spectra by the application of chemometric techniques. *Magn. Reson. Med.* **28**, 214–236 (1992).
30. Maxwell, R. J. *et al.* Pattern recognition analysis of ^{1}H NMR spectra from perchloric avid extracts of human brain biopsies. *Magn. Reson. Med.* **39**, 869–877 (1998).
31. El-Deredy, W. Pattern recognition approaches in biomedical and clinical magnetic resonance spectroscopy: a review. *NMR Biomed.* **10**, 99–124 (1997).
32. Tomlins, A. M. *et al.* High resolution ^{1}H NMR spectroscopic studies on dynamic biochemical processes in incubated human seminal fluid samples. *Biochim. Biophys. Acta* **1379**, 367–380 (1998).
33. Griffin, J. L., Walker, L. A., Shore, R. F. & Nicholson, J. K. Metabolic profiling of chronic cadmium exposure in the rat. *Chem. Res. Toxicol.* **14**, 1428–1434 (2001).
34. Foxall, P. J. D. & Nicholson, J. K. Nuclear magnetic resonance spectroscopy: a non-invasive probe of kidney metabolism and function. *Exp. Nephrol.* **6**, 409–414 (1998).
35. Bairaktari, E., Katopodis, K., Siamopoulos, K. C. & Tsolas, O. Paraquat-induced renal injury studied by ^{1}H NMR spectroscopy of urine. *Clin. Chem.* **44**, 1256–1261 (1998).

36. Beckwith-Hall, B. M. *et al.* Nuclear magnetic resonance spectroscopic and principal components analysis investigations into biochemical effects of three model hepatotoxins. *Chem. Res. Toxicol.* **11**, 260–272 (1998).
37. Ghauri, F., McLean, A., Beales, D., Wilson, I. D. & Nicholson, J. K. Induction of 5-oxoprolinuria in the rat following chronic feeding with *N*-acetyl 4-aminophenol (paracetamol). *Biochem. Pharmacol.* **46**, 953–957 (1993).
38. Holmes, E., Bonner, F. & Nicholson, J. K. Comparative studies on the nephrotoxicity of 2-bromoethanamine hydrobromide in the Fisher 344 rat and the multimammate desert mouse (*Mastomys natalensis*). *Arch. Toxicol.* **70**, 89–95 (1995).
39. Holmes, E., Bonner, F. & Nicholson, J. K. Comparative biochemical effects of low doses of mercury II chloride in the F344 rat and the multimammate mouse (*Mastomys natalensis*). *Comp. Biochem. Physiol. C* **114**, 7–15 (1996).
40. Holmes, E., Bonner, F. & Nicholson, J. K. ^{1}H NMR spectroscopic and histopathological studies on propylene imine induced renal papillary necrosis in the rat and the multimammate desert mouse (*Mastomys natalensis*). *Comp. Biochem. Physiol. C* **116**, 125–134 (1997).
41. Bollard, M. E. *et al.* Investigations into biochemical changes due to diurnal variation and estrus cycle in female rats using high resolution ^{1}H NMR spectroscopy of urine and pattern recognition. *Anal. Biochem.* **295**, 194–202 (2001).
42. Gavaghan, C. L. *et al.* HPLC–NMR spectroscopic and chemometric studies on metabolic variation in Sprague Dawley rats. *Anal. Biochem.* **291**, 245–252 (2001).
43. Moolenaar, S. H. *et al.* β-Ureidopropionase deficiency: a novel inborn error of metabolism discovered using NMR spectroscopy on urine. *Magn. Reson. Med.* **46**, 1014–1017 (2001).
44. Moolenaar, S. H. *et al. In vivo* and *in vitro* NMR spectroscopy reveal a putative novel inborn error involving polyol metabolism. *NMR Biomed.* **14**, 167–176 (2001).
45. Burns, S. P., Holmes, H. C., Chalmers, R. A., Johnson, A. & Iles, R. A. Proton NMR spectroscopic analysis of multiple acyl-CoA dehydrogenase deficiency — capacity of the choline oxidation pathway for methylation *in vivo*. *Biochim. Biophys. Acta* **1406**, 274–282 (1998).
46. Neild, G. H., Foxall, P. J. D., Lindon, J. C., Holmes, E. & Nicholson, J. K. Uroscopy in the 21st century: high field NMR spectroscopy. *Nephrol. Dial. Transplant.* **12**, 404–417 (1997).
47. Foxall, P. J. D., Bewley. S., Neild, G., Rodeck, C. H. & Nicholson, J. K. Analysis of fetal and neonatal urine using proton NMR spectroscopy. *Arch. Dis. Child.* **73**, F153–F157 (1995).
48. Foxall, P. J. D. *et al.* Proton nuclear magnetic resonance studies of early ifosfamide-induced nephrotoxicity in man: ifosfamide-induced encephalopathy and glutaric aciduria. *Clin. Cancer Res.* **3**, 1507–1518 (1997).
49. Holmes, E. *et al.* Automatic data reduction and pattern recognition methods for analysis of ^{1}H nuclear magnetic resonance spectra of human urine from normal and pathological states. *Anal. Biochem.* **220**, 284–296 (1994).
50. Gygi, S. P., Rochon, Y., Franza, B. R. & Aebersold, R. Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720–1730 (1999).
51. Tweeddale, H., Notley-McRobb, L. & Ferenci, T. Effect of slow growth on metabolism of *Escherichia coli*, as revealed by global metabolite pool ('metabolome') analysis. *J. Bacteriol.* **180**, 5109–5116 (1988).
52. Cutler, P. *et al.* An integrated proteomic approach to studying glomerular nephrotoxicity. *Electrophoresis* **20**, 3647–3658 (1999).
53. Nicholson, J. K., Buckingham, M. J. & Sadler, P. J. High resolution proton NMR studies of vertebrate blood and plasma. *Biochem. J.* **211**, 605–615 (1983).
54. Bales, J. R., Higham, D. P., Howe, I., Nicholson, J. K. & Sadler, P. J. Use of high resolution proton nuclear magnetic resonance spectroscopy for rapid multi-component analysis of urine. *Clin. Chem.* **30**, 426–432 (1984).
55. Nicholson, J. K. *et al.* Proton NMR studies of serum, plasma and urine from fasting normal, and diabetic subjects. *Biochem. J.* **217**, 365–375 (1984).
56. Nicholson, J. K., Timbrell, J. A., & Sadler, P. J. Proton NMR spectra of urine as indicators of renal damage: mercury nephrotoxicity in rats. *Mol. Pharmacol.* **27**, 644–651 (1985).
57. Halligan, S. *et al.* A study of the nephrotoxicity of three cephalosporins in rabbits using ^{1}H NMR spectroscopy. *Toxicol. Lett.* **81**, 15–21 (1995).
58. Gartland, K. P. R., Beddell, C., Lindon, J. C. & Nicholson, J. K. The application of pattern recognition methods to the analysis and classification of toxicological data derived from NMR spectroscopy of urine. *Mol. Pharmacol.* **39**, 629–642 (1991).
59. Holmes, E. *et al.* NMR spectroscopy and pattern recognition analysis of the biochemical processes associated with the progression and recovery from nephrotoxic lesions in the rat induced by mercury (II) chloride and 2-bromoethanamine. *Mol. Pharmacol.* **42**, 922–930 (1992).
60. Ghauri, F., Wilson, I. D., Parkes, H. G. & Nicholson, J. K. Asymptomatic 5-oxoprolinuria in man detected by proton NMR spectroscopy of urine. *Clin. Chem.* **39**, 1341 (1993).
61. Nicholson, J. K. & Wilson, I. D. High resolution proton NMR spectroscopy of biological fluids. *Prog. Nucl. Magn. Reson. Spectrosc.* **21**, 449–501 (1989).
62. Lynch, M. J. & Nicholson, J. K. Proton MRS of human prostatic fluid: correlations between citrate, spermine and *myo*-inositol and changes with disease. *Prostate* **30**, 248–255 (1997).
63. Lindon, J. C., Holmes, E. & Nicholson, J. K. Pattern recognition methods and applications in biomedical magnetic resonance. *Prog. Nucl. Magn. Reson. Spectrosc.* **39**, 1–40 (2001).
64. Ghauri, F. Y. K. *et al.* NMR spectrosocopy of human post mortem cerebrospinal fluid: distinction of Alzheimer's disease from control using pattern recognition and statistics. *NMR Biomed.* **6**, 163–167 (1993).
65. Nicholson, J. K., Foxall, P., Spraul, M., Farrant, R. D. & Lindon, J. C. 750 MHz ^{1}H and ^{1}H–^{13}C NMR spectroscopy of human blood plasma. *Anal. Chem.* **67**, 793–811 (1995).
66. Liu, M., Nicholson, J. K., Parkinson, J. A. & Lindon, J. C. Measurement of biomolecular diffusion coefficients in blood plasma using two-dimensional ^{1}H–^{1}H diffusion-edited total correlation spectroscopy. *Anal. Chem.* **69**, 1504–1509 (1997).
67. Shockcor, J. *et al.* Combined hyphenation of HPLC, NMR spectroscopy and ion-trap mass spectrometry (HPLC–NMR–MS) with application to the detection and characterization of xenobiotic and endogenous metabolites in human urine. *Anal. Chem.* **68**, 4431–4435 (1996).
68. Spraul, M. *et al.* Flow injection NMR spectroscopy of biofluids for rapid structural studies and high throughput biochemical screening. *Anal. Commun.* **34**, 339–341 (1997).
69. Moka, D. *et al.* Magic angle spinning proton nuclear magnetic resonance spectroscopic analysis of intact kidney tissue samples. *Anal. Commun.* **34**, 107–109 (1997).
70. Cheng, L. L. *et al.* Enhanced resolution of proton NMR spectra of malignant lymph nodes using magic angle spinning. *Magn. Reson. Med.* **36**, 653–658 (1996).
71. Cheng, L. L. *et al.* Quantitative neuropathology by high resolution magic angle spinning proton magnetic resonance spectroscopy. *Proc. Natl Acad. Sci. USA* **94**, 6408–6413 (1997).
72. Tomlins, A. *et al.* High resolution magic angle spinning ^{1}H nuclear magnetic resonance analysis of intact prostatic hyperplastic and tumour tissues. *Anal. Commun.* **35**, 113–115 (1998).
73. Garrod, S. L. *et al.* High resolution MAS-^{1}H NMR spectroscopic studies on intact rat renal cortex and medulla. *Magn. Reson. Med.* **41**, 1108–1118 (1999).
74. Waters, N. J. *et al.* High resolution magic angle spinning NMR spectroscopy of intact liver and kidney: optimisation of sample preparation procedures and biochemical stability of tissue during spectral acquisition. *Anal. Biochem.* **282**, 16–23 (2000).
75. Bollard. M., Holmes, E., Lindon, J. C. & Nicholson J. K. MAS-NMR studies on rat liver: assignment of glycogen resonances. *Magn. Reson. Med.* **44**, 201–207 (2000).
76. Millis, K., Maas, E., Cory, D. G. & Singer, S. Gradient enhanced high resolution magic angle spinning nuclear magnetic resonance spectroscopy of human adipocyte tissue. *Magn. Reson. Med.* **38**, 399–403 (1997).
77. Griffin, J. L. *et al.* The biochemical profile of rat testicular tissue as measured by magic angle spinning NMR spectroscopy. *FEBS Lett.* **486**, 225–229 (2000).
78. Nicholls, A. W. *et al.* Metabonomic investigations into hydrazine toxicity in the rat. *Chem. Res. Toxicol.* **14**, 975–987 (2001).
79. Garrod, S. *et al.* High resolution ^{1}H NMR and magic angle spinning NMR spectroscopic investigation of the biochemical effects of 2-bromoethanamine in intact renal and hepatic tissue. *Magn. Reson. Med.* **45**, 781–790 (2001).
80. Waters, N. J. *et al.* NMR and pattern recognition studies on the time-related metabolic effects of α-naphthylisothiocyanate on liver, urine, and plasma in the rat: an integrative metabonomic approach. *Chem. Res. Toxicol.* **14**, 1401–1412 (2001).
81. Griffin, J. L., Mann, C. J., Scott, J., Shoulders, C. C. & Nicholson, J. K. Choline containing metabolites during cell transfection: an insight into magnetic resonance spectroscopy detectable changes. *FEBS Lett.* **509**, 263–266 (2001).
82. Evans, J. N. S. *Biomolecular NMR Spectroscopy* (Oxford Univ. Press, Oxford, 1996).
83. Manley, B. F. J. *Multivariate Statistical Methods: a Primer* (Chapman & Hall, London, 1986).
84. Beebe, K. R., Pell, R. J. & Seahsholt, M. B. *Chemometrics: a Practical Guide* (John Wiley & Sons, New York, 1998).
85. Jurs, P. C. Pattern recognition used to investigate multivariate data in analytical chemistry. *Science* **232**, 1219–1224 (1986).
86. Eriksson, L., Johansson, E., Kettanah-Wold, N. & Wold, S. *Introduction to Multi- and Megavariate Data Analysis Using Projection Methods (PCA and PLS)* (UMETRICS, Sweden, 1999).

Acknowledgements

We thank I. Wilson, (AstraZeneca Pharmaceuticals, UK), J. Everett, (Pfizer Global Research, UK) and D. Robertson (Pfizer Global Research, USA) for helpful discussions on the philosophy of metabonomics. We are grateful to M. Spraul and Bruker Analytik GmbH, Germany, for help with many metabonomics projects. Particular thanks are due to H. Antti, A. Nicholls, C. Gavaghan and J. Bundy for data collection and help with data processing. We thank J. Azmi for her help with the α-naphthylisothiocyanate data and analysis. We also thank The Engineering and Physical Sciences Research Council, the Medical Research Council, the Natural Environment Research Council and the Biotechnology and Biological Sciences Research Council for funding this and related work over many years.

Online links

FURTHER INFORMATION
Biological Chemistry Section:
http://www.med.ic.ac.uk/divisions/1/home.asp
Metabometrix: http://www.metabometrix.com/
Access to this interactive links box is free online.

Jeremy Nicholson is Professor and Head of Biological Chemistry in the Faculty of Medicine at Imperial College, London, where he leads one of the world's largest research groups in metabolic science. He is the author of more than 350 scientific papers and articles, and has won the Royal Society of Chemistry Gold and Silver Medals for Analytical Science, and the Chromatographic Society Jubilee Silver Medal for his work on drug metabolism, analytical toxicology and clinical biochemistry. He is a Fellow of the Royal Society of Chemistry and the Royal College of Pathologists.

Elaine Holmes is a lecturer in biological chemistry at Imperial College, London. She obtained her B.Sc. in Biology from London University, and a Ph.D. in Chemistry from Birkbeck College, London Univeristy, on NMR spectroscopy of biofluids and pattern-recognition analysis. She and her group are working on new approaches to enhanced information recovery from complex biological spectroscopic data, and novel NMR methods in clinical diagnosis.

John Connelly obtained his B.Sc. and Ph.D. from Surrey University, United Kingdom, in toxicology. He was subsequently elected as a Fellow of the Institute of Biomedical Sciences. For several years, he was Contracts Manager at the Robens Institute, Surrey University, and subsequently at Beecham, SmithKline Beecham and GlaxoSmithKline, where he held positions as Head of Clinical Pathology, Director of Drug Development Sciences, Director of Investigative Toxicology and Group Director of the Safety Assessment Executive Committee, with international responsibilities. He formed the Biofluid NMR Group at SmithKline Beecham in 1995. He is now Director of Biology at Metabometrix Ltd (an Imperial College spin-off company), and a visiting researcher in biological chemistry at Imperial College.

John Lindon holds B.Sc., Ph.D. and D.Sc. degrees in chemistry from Birmingham University, United Kingdom, and is a Fellow of the Royal Society of Chemistry. He has an extensive research background in NMR spectroscopy of complex systems, including liquid crystals, biofluids and tissues, and has published more than 300 research papers and articles. He has many years of experience in pharmaceutical research and development with the Wellcome Foundation, Beckenham, United Kingdom. His main research interests are NMR spectroscopy, applications of spectroscopy in general to biomedical research, computational chemistry and chemometrics. He is a senior research investigator and professor at Imperial College in the Biological Chemistry section.