molecular
systems
biology

EDITORIAL

Global systems biology, personalized medicine and molecular epidemiology

Systems biology in individuals and populations

One of the great challenges for 21st century medicine is to deliver effective therapies that are tailored to the exact biology or biological state of an individual to enable so-called 'personalized healthcare solutions'. Ideally, this would involve a system of patient evaluation that would tell clinicians the correct drug, dose or intervention for any individual before the start of therapy. A practical approach to this evaluation is the concept of patient stratification in which individuals are biologically subclassified (classically according to some genetic features) and biofeatures modelled in relation to outcome. In principle, such stratification for personalized therapy can be applied to drug safety and efficacy modelling and to more general healthcare paradigms involving optimized nutrition and lifestyle management.

Of course, truly personalized treatments, even if they can be developed and applied widely, will lamentably always be a luxury of the worlds' richest citizens and nations. So in some respects, personalized healthcare might appear to be at the opposite end of the medical spectrum to the subject of epidemiology in which disease risk factors and disease incidence are studied in populations rich and poor alike. Systems biology provides us with a common language for both describing and modelling the integrated action of regulatory networks at many levels of biological organization from the subcellular through cell, tissue and organ right up to the whole organism. The relatively new science of *molecular* epidemiology concerns the measurement of the fundamental biochemical factors that underlie population disease demography and understanding 'the health of nations' and this subject naturally lends it to systems biology approaches. Hence, systems biology is certain to have in future a major role in both the development of personalized medicine and in molecular epidemiological studies.

Populations are, of course, made up of individuals and, in principle, there are important unifying features that can be considered from a systems perspective in which biological parameter variability in individuals and their statistical description in large populations can be used to interrogate the outcomes of therapeutic interventions and global patterns of disease distribution. Personalized healthcare and molecular epidemiology are thus effectively two sides of the same 'systems biology coin'; the essential differences are with respect to the type of medical end points or outcomes that are to be modelled (Figure 1). Metabonomics (see Box 1 for definitions of terms) offers a practical approach to measuring the metabolic end points that link directly to whole system

activity and metabolic profiles are determined by both host genetic and environmental factors (Nicholson *et al*, 2002).

The majority of personalized approaches have so far been mainly based on measuring genotype variations relating to polymorphisms in drug-metabolizing enzymes such as cytochrome *P*450 isoenzymes and *N*-acetyl transferases (Meyer and Zanger, 1997; Eichelbaum and Burk, 2001; Srivastava, 2003). As there are many of examples of adverse drug reactions being linked to specific enzymatic deficiencies or mutations (Meyer and Zanger, 1997; Eichelbaum and Burk, 2001; Srivastava, 2003), it seems perfectly reasonable to pursue genetically based personalized medicine strategies. However, pharmacogenomic results have thus far proved to be surprisingly disappointing, partly because of the unexpected complexity of the human genome and the difficulties in accurately and unequivocally describing human genotypes and phenotypes (Nebert and Menon, 2001; Nebert *et al*, 2003; Nebert and Vessell, 2004). Moreover, when considering the wider aspects of human health, it is clear that most major diseases are subject to strong environmental influences, and the majority of people in the world die from what are, in the broadest sense, environmental causes. At the personal level external influences also affect drug metabolism and toxicity, and individual outcomes of a drug intervention are the result of conditional probabilistic interactions between complexes of drug-metabolizing enzyme genes, a range of metabolic regulatory genes and environmental factors such as diet (Nicholson *et al*, 2004).

Even the basic concept of a 'specified' human population is actually confusing and has often involved ill-defined notions of ethnicity, which are associated with historical culturally biased thinking rather than the genuine and usually small genetic differences between human population groups. The overall lack of genetic variation *between* populations is remarkable in itself and this is a consequence of humans having moved out of Africa only ca. 100 000 years ago. Thus, according to microsatellite studies, only about 5–10% of the total human genetic variance actually occurs between populations or ethnic groups (Cavalli-Sforza and Feldman, 2003). Of course phenotypically, population subgroups around the world vary widely, as do human disease distributions that are related to diet and environmental factors. There are also well-known differences in drug metabolism (and hence toxicity potential) associated with variations in human genotype and phenotype at both individual and population levels (Meyer and Zanger, 1997; Eichelbaum and Burk, 2001; Nebert and Menon, 2001; Nebert *et al*, 2003; Srivastava, 2003; Nebert and Vessell, 2004). Obviously, there are many connections between the health of general populations and

Box 1

Definitions of terms

Conditional metabolic phenotype: The characteristic phenotypic metabolite profile in any compartment or fluid resulting from the interaction of the host genome with environmental factors—this can be considered to be the directly measurable component of the metabonome.

Co-metabolite: A metabolite that can only be formed by the integrated biochemical actions of more than one genome, for example, the gut microbial metabolism of a mammalian metabolite or vice versa (Nicholson *et al*, 2005).

Metabolome: The quantitative description of all the low-molecular-weight components (<1 kDa) of endogenous (host genome control) metabolites in a specified biological sample. Each cell type and biological fluid will have a characteristic set of metabolites that is characteristic of a species under specific environmental conditions and fluctuates through time according to physiological demands.

Metabonome: All the theoretical sums and products of the interactions of multiple metabolomes in a complex system including extended genome, symbiotic, parasitic, environmental and co-metabolic interactions. Urine and plasma are fluids that carry metabolic signatures that result from such interactions.

Metabonomics: The quantitative measurement of the multiparametric time-related metabolic *responses* of a complex (multicellular) system to a pathophysiological intervention or genetic modification (Nicholson *et al*, 1999, 2002). Thus metabonomics seeks to assess the global system level homeostatic and pathological responses to interventions or stressors. The word origin is from the Greek (*meta* meaning change and *nomos* meaning a rule set or set of laws).

Metabolomics: The comprehensive quantitative analysis of all the metabolites of an organism or specified biological sample. There are numerous and often conflicting uses of this word in the literature, but they are all basically analytical definitions (e.g. Raamsdonk *et al*, 2001). Also it is often not exactly clear what constitutes a *metabolite* because there is a continuum of sources of molecules that can be considered to be metabolites, ranging from those entirely under host genome control to exogenous dietary compounds and drugs to those being extensively co-metabolized by the gut microbiome (Nicholson and Wilson, 2003; Nicholson *et al*, 2004).

Metabotype: A metabolic profile that defines a phenotype which relates to genetic variation of the mammal (Gavaghan *et al*, 2000).

Microbiome: The flexible consortium of microorganisms, bacteria, fungi and yeasts that live commensally or symbiotically in the gut of higher animals (Lederberg, 2000; Nicholson *et al*, 2005). The microbial species present and relative abundance vary significantly between animal species and between individuals within a species. Human individuals may have microbiome characteristics that are unique to their individual biology. There is intrinsic stability in an individuals' microbiome, but certain interventions such as drugs or antibiotics may cause perturbations.

Pharmaco-metabonomics: The prediction of the quantitative outcome of a healthcare (typically drug) intervention in a given individual based on a pre-dose mathematical model of the metabolic state constructed in a supervised learning system (Clayton *et al*, 2006).

Theranostic: A specific combined diagnostic for a particular disease with a therapy for that condition. The test, which may be genetic, results in patient stratification of the population so that the drug is more efficiently directed.

Xeno-metabolome: The characteristic profile of non-endogenous compounds observed in the biofluid as a result of individual or population exposures through their environment or diet, either passively or through self-administration e.g. drugs and their metabolites, pollutants, dietary components, herbal medicines, etc (Nicholson and Wilson, 2003; Teague *et al*, 2004).

that of the individuals that make them up, and so it is useful to consider this from a molecular systems biology viewpoint (Figure 1). However, measurement of parameters that relate system level activities to drug interventional outcomes is practically highly limited in applications involving large-scale human populations (Box 2). Population stratification (in the epidemiological rather than personal sense) according to age, gender, diet, 'ethnicity' and socioeconomic factors is complicated by the fuzziness of some of the classes, and this complicated modelling of these features in relation to systems biology (omics) metrics. Thus, bridging the subjects of personal healthcare and population epidemiology via system biology will require a pragmatic and practical approach, which leads us to the concept of 'top-down' systems biology and the derivation of metabolic parameters of 'global' system function.

'Top-down' systems biology and metabonomics

We have been advocating the use of metabolic measurement at the system level utilizing metrics obtainable from biological fluids such as urine and plasma for many years (Nicholson and Wilson, 1989). A particular advantage of biological fluid monitoring or screening is that it is minimally invasive or non-invasive and can be applied on a large scale for human population phenotyping (Nicholson *et al*, 1999, 2002). The science of metabonomics deals with understanding metabolic changes of a complete system caused by interventions (Nebert *et al*, 2003; Dumas *et al*, 2006a, b) and in particular we have noted that metabolic end points are the result of gene–environment interactions in their broadest sense, including extended genome and parasitic interactions (Wang *et al*, 2004; Dumas *et al*, 2006a, b; Martin *et al*, 2006). We have previously outlined our ideas about conditional probabilistic (Bayesian) interactions between genes and environment with respect to adverse drug reactions in individuals and have suggested a hypothetical (Pachinko) model to help study and visualize these interactions (Nicholson and Wilson, 2003). In the Pachinko model, a popular Japanese pinball machine game is used as a metaphor to underscore the idea that metabolic fate results from a sequence of conditional probabilistic interactions between metabolites and components of the cellular biochemical network. In particular, drug molecules can be thought of as a tumbling shaped charge represented as a ball in the machine. Each ball (drug molecule) hits pins

(representing drug-metabolizing enzymes—the exact position of which would analogously vary with SNP variations), which transforms the molecule sequentially and so alters its course through the machine (cell/body). Eventually, the drug is

metabolized to a state that readily leaves the body and so the exits from the machine at hypothetical 'excretion points'. The behaviour of each individual ball is difficult to predict, but the probabilistic path of the whole population can be modelled. Thus, the environmental interaction components, for example, gut microbial metabolites, chemicals or dietary compounds, can also be visualized as other balls or shaped objects tumbling through the Pachinko machine. These agents may then block or interfere with or even enhance the drug metabolism pathways. This equates to altering the probabilities of metabolic flow through the system, and the resulting changes in the pathway utilization may be modelled using Bayesian methods. These gene–environment interactions can result in many outcomes—some of which may generate metabolites that cause cellular damage or idiosyncratic (unpredictable) toxicity. Related to this is our concept of the 'conditional metabolic phenotype' or CMP (Nicholson et al, 2005) in which both genetic factors and exogenous factors, such as diet, exposure to foreign chemicals and so on, interact to determine the possible outcomes of a drug or dietary intervention (Nicholson et al, 2004, 2005; Dumas et al, 2006a, b). The most important feature of the CMP concept is that it represents a starting point of an individual in a multivariate metabolic space that is the result of the combination of many physical, chemical, genetic and environmental influences. We have hypothesized that it is the starting position irrespective of the relative contributions of the individual 'vector' components that determines the outcome of an intervention (Nicholson et al, 2004) and this is exactly the basis of the personalized healthcare paradigm.
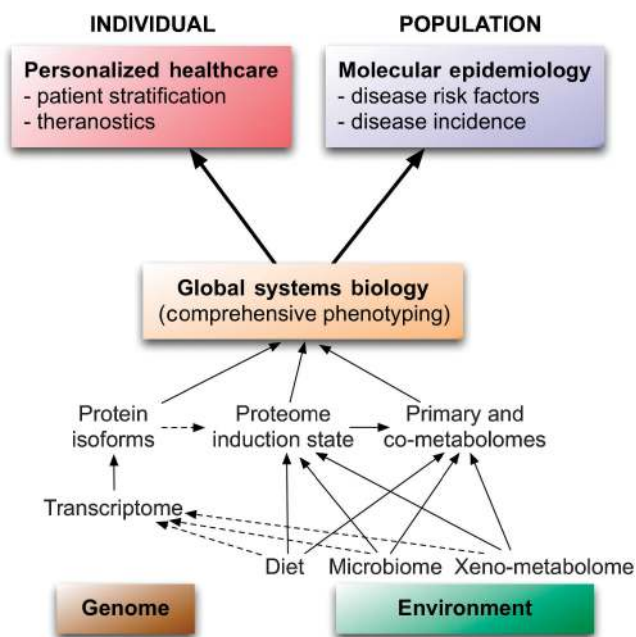

Figure 1 Relationships between systems biology, personalized healthcare and molecular epidemiology (dotted lines indicate indirect connections or influences).
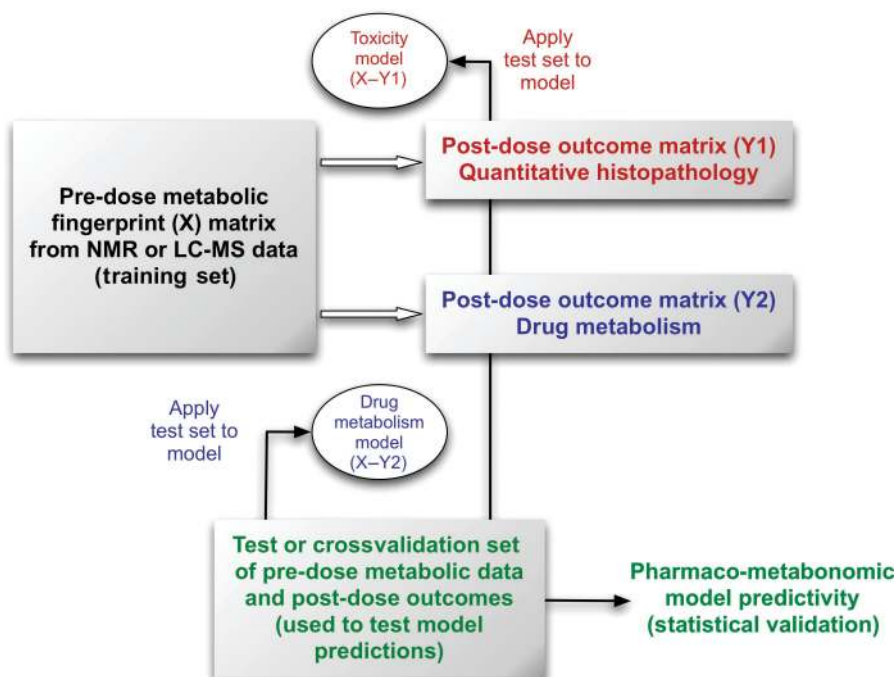

Figure 2 Pharmaco-metabonomic modelling procedure: spectroscopic data on pre-dose metabolic fingerprints (X matrix) from biofluids such as urine and plasma are statistically linked to outcome (quantitative toxicity (Y1) drug metabolism (Y2) matrixes) of a drug intervention via multivariate statistics such as partial least squares methods. Typically, 20–50% of all data is used in the training set construction. The predictive power of the models is then tested using a test set or a cross-validation set to assess model robustness. It is also possible as an additional test to avoid overfitting of data, to deliberately permute the training set matrixes to induce a false model that should have a very low predictive capability.

Box 2

Metabolic screening of individuals and populations

Current methodologies for metabolic profiling: All metabolic profiling revolves around NMR and mass spectrometric (MS) methods, as these technologies can give multiparametric information on many classes of molecules often at the same time. In the case of NMR, intact fluids can be analysed directly (Nicholson and Wilson, 1989; Nicholson *et al*, 2002). In the case of MS, chromatographic hyphenation is usually required either to types of liquid chromatography (LC–MS) (Wilson *et al*, 2005; Plumb *et al*, 2006) or to gas chromatography (GC–MS) following chemical derivatization (Raamsdonk *et al*, 2001). All technologies are low cost in comparison to other omics approaches such as proteomics and gene arrays. The typical cost per sample will vary according to the number and types of experiment performed, but might be expected to be in the range of \$10–\$150. NMR is the most rapid method and can deliver a reasonable quality spectrum with quantitative information on up to several hundred metabolites present at mid-micromolar level and above within minutes of collecting the sample. Low micromolar detection is possible with longer scanning times. NMR is also by far the most reproducible technology from an analytical point of view and has the largest linear dynamic range ($> 10^5$). Direct absolute quantification is possible by NMR, but is much more time consuming and usually is not necessary when combined with pattern recognition. GC–MS and LC–MS methods are usually more sensitive than NMR—sometimes by several orders of magnitude—but in the case of LC–MS, quantification is dependent on ionization efficiency and peak overlap. Both GC–MS and LC–MS are highly reproducible. Absolute quantification by LC–MS is difficult unless there is a standard to compare with. All the technologies can potentially handle hundreds of samples per day—in the case of MS, this is limited by the chromatographic run times.

Building metabonomic and pharmaco-metabonomic models: The basic procedure for constructing pharmaco-metabonomic models is outlined in Figure 2. The sample sizes required to build robust predictive models vary according to the application. In the case of human population metabonomic profiling, we have produced highly predictive models using a few hundred individuals (Dumas *et al*, 2006a, b). In the case of pharmaco-metabonomics for paracetamol toxicity, we obtained good models using ca. 80 animals (Clayton *et al*, 2006). Models that work on humans (e.g. for disease diagnosis) always require large training and test sets and this is likely to be true for pharmaco-metabonomics. This is so that all of the human metabolic variation and sources are well sampled in relation to the drug dosing outcome. NMR-derived metabonomics can give very strong models for population biochemical phenotyping (Brugman *et al*, 2006; Clayton *et al*, 2006) or for individuals in terms of class prediction (e.g. population of origin). In pharmaco-metabonomics, it may not be possible to give this level of precision—but it may be appropriate to express the outcome in simpler ways such as low, medium or high risk of an adverse reaction—with some statistical definition given to these classes, which will be case dependent and is as yet unknown. In the case of population studies, we are mainly concerned with the statistical description of multivariate metabolic parameters of large numbers of individuals stratified by age, gender, ethnicity, diet and environmental factors that could be measured. As this is a descriptive process, it is much easier to design molecular epidemiological studies with metabolic classifiers and outputs than it is to design pharmaco-metabonomic studies where the individuals must be subject to controlled experiments to build the initial models.

So how do we start to apply these ideas to real systems? 'Bottom-up' modelling approaches if viewed in the cold light of day can never really work in the world of gigantic human phenotypic variability. Indeed even *in vitro* to *in vivo* extrapolations of drug metabolism and toxicity within one species are notoriously unreliable, and 'bottom-up' systems biology modelling poses a vastly more complex challenge because most of the quantitative features needed to make reasonable cellular models are simply not measurable in 'intact' humans. So approaches appear to work very well in the systems biology of yeast or *Escherichia coli* cultures are not readily translatable into the modelling of either individual human or population biology.

Pharmaco-metabonomics and prediction of drug intervention outcomes

In the alternative 'top-down' approach where metrics of the systemic homeostatic activity are obtained, we have now shown a 'proof-of-concept' of a new 'pharmaco-metabonomic' approach to understanding and predicting interventional outcome of drugs (such as toxicity and xenobiotic metabolism in animal model systems) based on mathematical models of a pre-dose metabolic profiles (Clayton *et al*, 2006). In these studies, we investigated the effects of three structurally diverse hepatotoxins in rats (galactosamine, allyl alcohol and paracetamol), which act via different mechanisms, and found that pre-dose urinary profiles carried information about the degree of post-dosing toxicity, and in the case of paracetamol information about variation of drug metabolism itself. Pharmaco-metabonomics is thus the prediction of the outcome of an intervention in individuals based on pre-dose metabolic

state of that individual (Nicholson and Wilson, 2003; Clayton *et al*, 2006). In a preliminary study on galactosamine toxicity, we found that the responder/non-responder pattern of liver damage at 24 h post-dosing was reflected in the *pre-dose* metabolic profile of the urine. This was achieved using a simple principal components analysis (which is an unsupervised method that is blind to class in its construction). In a more complex study, a supervised approach, projection-on-latent structure method, was used working with animals given a threshold toxic dose of paracetamol that produced a wide range of liver toxicity between individuals (Figure 2). Here, we found once again that there was a significant association between pre-dose metabolic profile and post-dose outcome with respect to liver damage severity and indeed to drug metabolism (specifically the paracetamol to paracetamol glucuronide excretion ratio was strongly correlated with pre-dose urinary metabolite profiles). These studies imply that there may be future possibilities of applying this approach non-invasively to screening humans in populations. However, practically this is still far off, and we need to extend our knowledge on the relationships between endogenous metabolic status and drug metabolism outcomes for a much wider range of drugs. Of course, there are also significant ethical issues involved with such screening procedures in man. Furthermore, we should not forget that models obtained by the integration of various 'omics' approaches (pharmaco-genomics, pharmacoproteomics and pharmaco-metabonomics) may have improved predictive power, which might indeed be required to get personalized healthcare to work in the real world. Indeed, we have recently shown that proteomics and metabonomics can be statistically integrated to produce new trans-omic combination biomarkers to classify experimental disease states such as xenograft models of

prostatic cancer (Rantalainen *et al*, 2006). However, with current technology, the scale-up of multi-omics strategies to man would be impractical and prohibitively expensive.

The most likely near-term implementation of pharmaco-metabonomics would be in the pharmaceutical industry itself at the clinical trial or development stage when drugs are first going into man. Here pre-dose metabolic models could be built and then related to quantitative metabolic fates of compounds and any observed adverse reactions. This would then lead to knowledge about the possible contraindications of a particular drug used in certain phenotypic classes of individuals, which is effectively a type of patient stratification. In any case, both early and clinical safety studies would benefit from the improved metabolic descriptions of test subjects (animal or man) and their responses to novel therapeutic agents, good or bad. It must also be said that the pharmaco-metabonomic concept is not limited just to drug interventions. Effects of dietary modulation, pre-biotic and probiotic treatments and other lifestyle changes could also ultimately be evaluated in this way. This is important because 'personalized healthcare' means different things for different people and, in general populations, it is lifestyle management not drug therapy that is most effective for disease prevention, which of course is better than having to find a cure.

Populations and molecular epidemiology: getting systems biology into man

Getting systems biology out of the laboratory into the more general human population both for screening purposes and in order to understand our own changing health patterns is a formidable challenge. Despite relentless advances in medical technology, many major indications of population morbidity and mortality such as heart disease, diabetes, obesity and cancer (all problems in which genetic and environmental factors are closely entwined) are rising all over the world. Interestingly, many of these diseases may be related to changes in the activities or composition of the gut microbiota (microbiome), which has probably been profoundly affected by our lifestyle changes (especially antibiotic use) over the last 50 or so years. In fact, the microbiome is the exact point where host genetics meets environment and can be considered to be our most integrated and influential 'environmental' factor (Nicholson *et al*, 2005). Given that humans have slowly evolved with this 'extended genome' of the microbiota, perturbation of this close association is potentially dangerous and, controversially, may be a root cause of many of our rapidly spreading 'modern' diseases (Nicholson *et al*, 2005). Indeed recent studies by us and others have shown that gut microbiotal variations affect the development of diet-induced insulin resistance and type II diabetes mellitus (Dumas *et al*, 2006a, b) and even the development of type I diabetes in experimental animals (Brugman *et al*, 2006), which until recently was thought of as being related to purely mammalian (human) genome problems. Thus, wherever we turn we see hypercomplexity in disease development and this must be taken into account in systems biology disease modelling if we are ever going to get effective treatments that actually work in man. In examining human populations for molecular epide-miological purposes, it will probably be important to measure metagenomic features of the gut microbiome, which strongly influences exact mammalian metabolic phenotypes of mice and men (Holmes and Nicholson, 2005; Gavaghan-McKee *et al*, 2006) and so, using the pharmaco-metabonomic argument, must also influence disease development and possibly optimized therapeutic interventions in individuals and populations. So as systems biology moves forward with the strong driver of personalized medicine, we will also be able to apply these strategies for looking at the changing demography of human disease around the world. Also the creation of personalized health science for the 'rich nations' should hopefully also benefit the people of developing nations, perhaps especially those countries that are trying to Western-ize their economies and lifestyles, and in so doing are now acquiring Western disease patterns at an alarming rate.

References

Brugman S, Klatter FA, Visser JTJ, Wildeboer-Veloo ACM, Harmsen HJM, Rozing J, Bos NA (2006) Antibiotic treatment partially protects against type 1 diabetes in the Bio-Breeding diabetes-prone rat. Is the gut microflora involved in the development of type 1 diabetes. *Diabetologia* **49:** 2105–2108

Cavalli-Sforza LL, Feldman MW (2003) The application of molecular genetic approaches to the study of human evolution. *Nat Genet* **33:** 266–275

Clayton TA, Lindon JC, Antti H, Charuel C, Hanton G, Provost JP, Le Net JL, Baker D, Walley RJ, Everett JE, Nicholson JK (2006) Pharmaco-metabonomic phenotyping and personalised drug treatment. *Nature* **440:** 1073–1077

Dumas ME, Barton RH, Toye A, Cloarec O, Blancher C, Rothwell A, Fearnside J, Tatoud R, Blanc V, Lindon JC, Mitchell S, Holmes E, McCarthy MI, Scott J, Gauguier D, Nicholson JK (2006a) Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin resistant mice. *Proc Natl Acad Sci USA* **103:** 12511–12516

Dumas ME, Maibaum EC, Teague C, Ueshima H, Zhou BF, Lindon JC, Nicholson JK, Stamler J, Elliott P, Chan Q, Holmes E (2006b) Assessment of analytical reproducibility of ^{1}H NMR spectroscopy based metabonomics in large scale epidemiological studies. *Anal Chem* **78:** 2199–2208

Eichelbaum M, Burk O (2001) CYP3A (2001) genetics in drug metabolism. *Nat Med* **7:** 285–288

Gavaghan CL, Holmes E, Lenz E, Wilson ID, Nicholson JK (2000) 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

Gavaghan-McKee CL, Wilson ID, Nicholson JK (2006) Molecular phenotyping of nude and normal (Alpk:ApfCD, C57BL10J) mice via metabonomics. *J Proteome Res* **5:** 378–384

Holmes E, Nicholson JK (2005) Gut microbiota strongly influences experimental rodent phenotypes. *Toxicol Lett* **87:** 1–2

Lederberg J (2000) Infectious history. *Science* **288:** 287–293

Martin FPJ, Verdu EF, Wang Y, Dumas ME, Yap IKS, Cloarec O, Bergonzelli GE, Corthesy-Theulaz I, Kochhar S, Holmes E, Lindon JC, Collins SM, Nicholson JK (2006) Transgenomic metabolic interactions in a mouse disease model: interactions of *Trichinella spiralis* infection with dietary *Lactobacillus paracasei* supplementation. *J Proteome Res* **5:** 2185–2193

Meyer UA, Zanger UM (1997) Molecular mechanisms of genetic polymorphisms of drug metabolism. *Annu Rev Pharmacol Toxicol* **37:** 269–296

Nebert DW, Menon AG (2001) Pharmacogenomics, ethnicity and susceptibility genes. *Pharmacogenomics J* **1:** 541–549

Nebert DW, Nebert J, Vessell ES (2003) Pharmacogenomics and 'Individualized Drug Therapy'. High expectations and disappointing results. *Am J Pharmacogenomics* **3:** 361–370

Nebert DW, Vessell ES (2004) Role of cytochrome P4501A in chemical carcinogenesis: implications for human variability in expression and enzyme activity. *Pharmacogenetics* **5:** 259–274

Nicholson JK, Connelly J, Lindon JC, Holmes E (2002) Metabonomics: a platform for studying drug toxicity and gene function. *Nat Rev Drug Discov* **1:** 153–161

Nicholson JK, Holmes E, Lindon JC, Wilson ID (2004) The challenges of modelling mammalian biocomplexity. *Nat Biotechnol* **22:** 1268–1274

Nicholson JK, Holmes E, Wilson ID (2005) Gut microbes, mammalian metabolism and personalized healthcare. *Nat Rev Microbiol* **3:** 2–8

Nicholson JK, Lindon JC, Holmes E (1999) 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR data. *Xenobiotica* **29:** 1181–1189

Nicholson JK, Wilson ID (1989) High resolution proton NMR spectroscopy of biological fluids. *Prog NMR Spectrosc* **21:** 449–501

Nicholson JK, Wilson ID (2003) Understanding 'global' systems biology: metabonomics and the continuum of metabolism. *Nat Rev Drug Discov* **2:** 668–676

Plumb RS, Johnson KA, Rainville P, Smith BW, Wilson ID, Castro-Perez JM, Nicholson JK (2006) UPLC-MS[E]:a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun Mass Spectrom* **20:** 1989–1994

Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* **19:** 45–50

Rantalainen M, Cloarec O, Trygg J, Beckonert O, Wilson ID, Jackson D, Rowlinson R, Jones S, Rayner S, Nickson J, Tonge R, Wilkinson R, Mills JD, Nicholson JK, Holmes E (2006) Statistically integrated metabonomic–proteomic studies on a human prostate cancer xenograft model in mice. *J Proteome Res* (in press)

Srivastava P (2003) Drug metabolism and individualized medicine. *Curr Drug Metab* **4:** 33–44

Teague C, Holmes E, Meibaum E, Nicholson JK, Tang H, Chan Q, Elliot P, Wilson ID (2004) Ethyl glucoside in human urine following dietary exposure: detection by ^{1}H NMR spectroscopy as a result of metabonomic screening of humans. *Analyst* **129:** 259–264

Wang Y, Holmes E, Nicholson JK, Cloarec O, Chollet J, Tanner M, Singer B H, Utzinger J (2004) Metabonomic investigations in mice infected with *Schistosoma mansoni*: an approach for biomarker identification. *Proc Natl Acad Sci USA* **101:** 12676–12681

Wilson ID, Nicholson JK, Castro-Perez J, Grainger JH, Johnson KA, Smith BW, Plumb RS (2005) High resolution 'ultra performance' liquid chromatography coupled to/oa-TOF mass spectrometry as a tool for differential metabolic pathway profiling in functional genomic studies. *J Proteome Res* **4:** 591–598

Jeremy K Nicholson
Department of Biomolecular Medicine, Faculty of Medicine, Imperial College London, South Kensington, London, UK