Ernst Schering Foundation Symposium Proceedings, Vol. 4, pp. 227–249
DOI 10.1007/2789_2008_096

Published Online: 03 July 2008

Human Metabolic Phenotyping and Metabolome Wide Association Studies

E. Holmes(✉), J.K. Nicholson

Divsion of Surgery, Oncology, Reproductive Biology and Anaesthetics (SORA), Faculty of Medicine, Sir Alexander Fleming Building, South Kensington, SW7 2AZ London, UK
email: *Elaine.holmes@imperial.ac.uk*

Abstract. Metabolic phenotyping in large-scale population studies can yield crucial information regarding the impact and interaction of genetic and environmental factors with regard to the prevalence and risk of chronic diseases. Spectroscopic technologies such as nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) can be used to generate multi-parameter profiles of biological samples and together with automated sample delivery and mathematical modelling systems, can be used as a high throughput screening tool. The adaptation of these metabolic profiling tools from pre-clinical studies in animal models to population studies in man is explored and an overview of the current and future roles of metabolic phenotyping is described, including the idea of "Metabolome Wide Association Screening" focussing on key dis-

ease areas such as cardiovascular disease and metabolic syndrome, cancers and neurodegeneration.

1 Introduction

Recognition of the inadequacy of the genome sequence to explain the fundamental nature of many disease processes has precipitated a marked increase in the evaluation of approaches that relate gene expression to phenotypic outcomes. There is also increasing recognition of biological complexity and the conceptual paradigm has been shifted from simple univariate measurements of response to the need to integrate technologies and their outputs in order to operate at a systems biology level. Interactions of genes, proteins and metabolites at different levels of biomolecular organization can be probed by various technologies and integrated using bioinformatic and chemometric strategies to extract latent information that carries a diagnostic or even prognostic signature. One of the major goals of twenty-first century medicine will be the introduction of personalized health care through a holistic understanding of an individual's overall biochemical status. In order to achieve this aim, the effects of both genetic predisposition and a wide range of environmental factors such as diet, drug intake, smoking habits, stress and amount of physical activity, etc., need to be taken into account. Metabonomics (variously referred to as metabolomics or metabolic profiling) (Nicholson et al. 1999, 2002; Fiehn et al. 2000) is a rapidly emerging field of research combining sophisticated analytical tools such as nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry with multivariate statistical analysis to generate complex metabolic profiles of biofluids and tissues. Pathological stimuli or genetic modification influence metabolite profiles in a characteristic and consistent manner, involving adjustment of the intra- and extracellular fluids as the organism strives to maintain homeostatic equilibrium. By harnessing appropriate mathematical and pattern recognition procedures to interrogate the data produced by high-resolution spectral analysis, characteristic profiles of physiological or pathological responses can be established.

The global objective of this chapter is to review the potential of metabolic profiling methods for characterizing the complex metabolic phenotype of humans in health and disease. Metabolic profiling has been successfully applied across a wide range of fields in plant and animal biology such as characterization of natural products (Bailey et al. 2002), monitoring response to therapeutic or nutritional interventions (Neild et al. 1997; Lamers et al. 2003; Wang et al. 2004), toxicology (Ebbels et al. 2007), drug metabolism (Foxall et al. 1996; Plumb et al. 2003), functional genomics (Gavaghan et al. 2000) and disease diagnosis and prognosis (Brindle et al. 2002; Yi et al. 2006; Clayton et al. 2006). The vast majority of metabolic profiling studies have been conducted in laboratory models of disease or toxicity where control over genetic and environmental conditions can be exercised. However, given the substantial array of animal studies that identify the metabolic response to controlled interventions, it is now appropriate to expand the available knowledge to address more complex phenotypes and, in particular, to extend the methodology to investigate human metabolism. The potential of metabolic profiling to address complex human clinical and even epidemiological questions has vastly increased due to recent advances in both analytical and mathematical technology; including capacity for higher throughput of samples, increased analytical sensitivity and the evolution of mathematical methods for accommodating analytical and biological variation. In this chapter, illustrations of research where metabolic profiling has already been employed in investigating human health and disease is summarized, and potential areas which would benefit from application of such technology are outlined.

2 Defining the "Normal" Phenotype

Prior to utilizing metabolic profiling technology for diagnostic purposes in human studies, it is first necessary to define the metabolic range covered by normal physiological variation. Only then can robust and specific biomarkers of disease be extracted. Metabolic variation is dependent on both genetic and environmental parameters, and each biological tissue or fluid has its own unique metabolic signature (Fig. 1). Ethnicity, gender, age, activity, nutritional status, medication, stress, polymor-

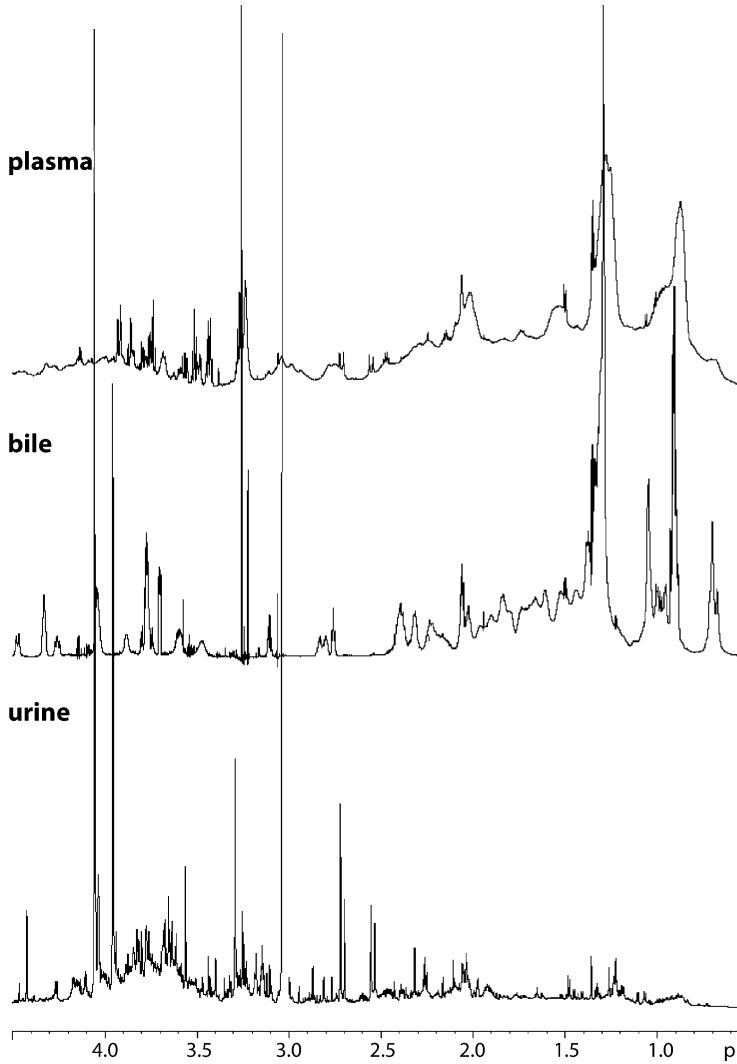

Fig. 1. Standard 600 MHz ^{1}H NMR spectra showing characteristic profiles for urine, plasma and bile

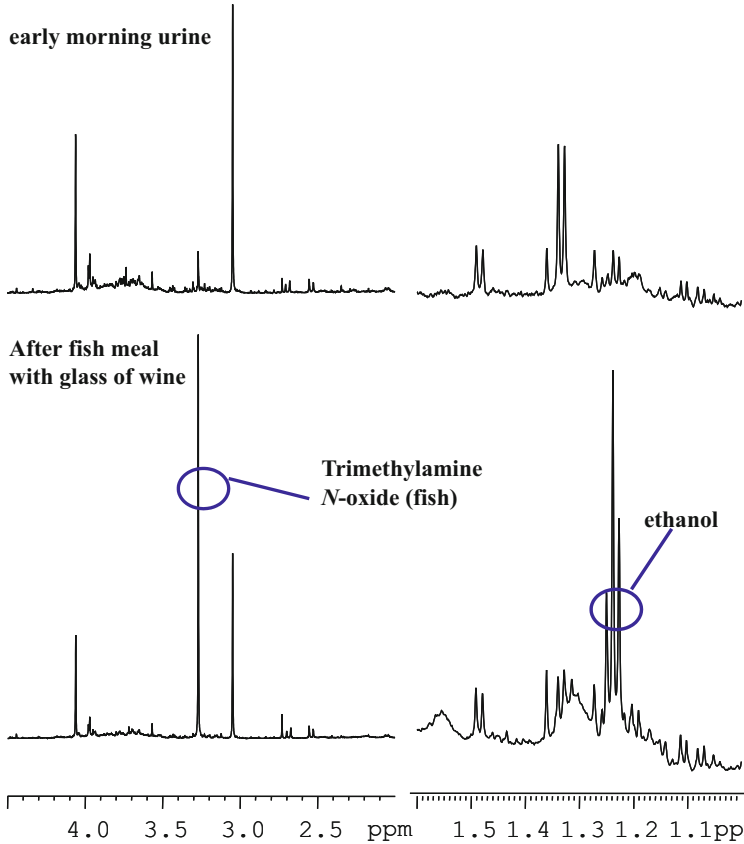

Fig. 2. Selected regions of standard 600-MHz ^{1}H NMR spectrum of urine samples from a healthy male before and after consuming an evening meal showing characteristic metabolic changes associated with the consumption of fish and a glass of wine (vertical scale for region on right hand side of plot x5)

phisms, hormone levels and circadian cycles are all known to impact upon mammalian metabolite profiles (Holmes et al. 1994; Slupsky et al. 2007; Williams et al. 2006; Bollard et al. 2001; Teague et al. 2004, 2006) (Fig. 2). Evaluation of normal ranges of mammalian metabolite

composition under various physiological and analytical conditions can be found in the literature for several biological matrices, including urine (Holmes et al. 1994; Maher et al. 2007), plasma (Teahan et al. 2006; Lenz et al. 2003), cerebrospinal fluid (CSF) (Koschorek et al. 1993), feces (Saric et al. 2008) and various tissues (Tsang et al. 2005; Wang et al. 2008; Garrod et al. 1999). The extent of variation and the dynamic ranges of metabolite concentrations in metabolite profiles are dependent upon the influence of homeostatic mechanisms on that biological matrix (Fig. 1). For example, plasma composition is maintained under homeostatic control and metabolite concentrations are found to be relatively stable in terms of both qualitative and quantitative differences in comparison with excretory biofluids such as urine, where metabolite concentrations vary greatly in terms of both the presence and quantity.

3 Detecting Pathophenotypes: Diagnostics

In many instances, diagnosis of the presence of a disease is achievable by routine and inexpensive clinical assays or genetic tests, for example type 2 diabetes, inborn errors of metabolism such as phenylketonuria, and many neurodegenerative disorders (Guthrie and Susi 1963; International Huntington Association and the World Federation of Neurology Research Group on Huntington's Chorea 1994). However, for some diseases, early diagnosis remains the key issue, and even for those diseases that are easily diagnosed by simple assays, in some cases the stage of disease is harder to determine accurately. Therefore, improved diagnostics are required in order to establish the optimal therapeutic management. Here the application of metabolic profiling can be an efficient tool for differential diagnosis of various disease conditions, as has been shown for a wide range of diseases, including cardiovascular, intestinal disorders, cancers, renal disease (Fig. 3), osteopathies and neuropathologies. Several examples are discussed in the following sections.

3.1 Metabolic Profiling of Insulin Resistance

Insulin resistance (IR) is one of the fastest growing human pathological conditions, and is now an increasing health burden in the develop-

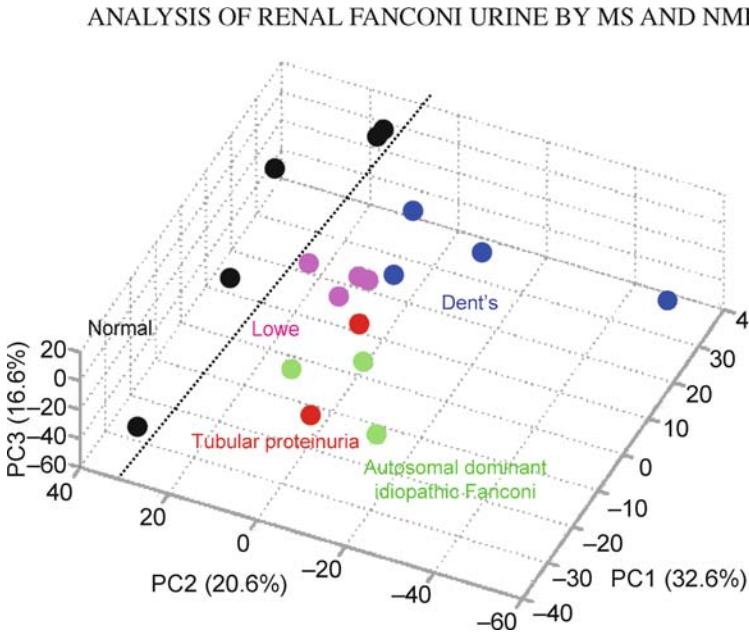

Fig. 3. Principal Components scores plot derived from the NMR and MS profiles of urine obtained from humans with different types of Fanconi syndrome. (Adapted from Vilasi 2007)

ing world as well as westernized societies. IR has been studied across a wide range of animal models using metabolic profiling and large-scale epidemiological studies are now being undertaken in human populations. Several studies on models of insulin resistance and type 2 diabetes have been undertaken in animal models. For example, the effects of streptozotocin-induced diabetes have been profiled using NMR with principal components analysis (PCA) (Nemoto et al. 2007). The effects of a high-fat diet were explored in inbred mouse strains selected for their resistance (BALB/c) and susceptibility (129S6) to IR and nonalcoholic fatty liver disease (NAFLD). High plasma concentrations of phosphocholine and increased urinary excretion of methylamines, associated with changes in gut microflora were found (Dumas et al. 2006a). Several

studies have also been conducted on the Zucker rat, which is a common animal model for IR and obesity (Dumas et al. 2006b; Yi et al. 2006). Several recent EU-funded initiatives such as MolPAGE, FGENTCARD, PROCARDIS involve or even focus on metabolic profiling of the human IR phenotype. Most of the early publications arising from these studies encompass an exploration of variation in human biofluids (Teague et al. 2004; Maher et al. 2007; Plumb et al. 2005), but several smaller studies targeting IR have identified specific metabolic phenotypes or metabotypes associated with IR and type 2 diabetes (Williams et al. 2005; Atherton et al. 2006). Indeed, type 2 diabetes was first profiled using NMR spectroscopy in 1984 (Bales et al. 1984). Although as yet there are few substantive papers exploring the more subtle and substantive metabolic consequences of IR, an explosion in the literature reporting on some of the major epidemiological studies is imminent.

3.2 Cardiovascular Disease

Like IR, cardiovascular disease (CVD) is also a part of the metabolic syndrome spectrum and is also growing at an alarming rate. CVD has been studied across several small populations using metabolic profiling approaches. Predominantly NMR spectroscopy-based studies on plasma or serum, in particular, have yielded metabolic profiles that are differentiated from control or healthy profiles in both the lipoprotein profiles, choline metabolites and in some of the lower molecular weight metabolic components (Brindle et al. 2002; Kirschenlohr et al. 2006). One LC-MS study, conducted on patients with myocardial ischaemia, some of whom demonstrated inducible ischemia and some of whom did not, was able to separate the two groups clearly on the basis of citric and lactic acid amongst other metabolites. However, of the 23 metabolites identified as candidate biomarkers, few were identified and the study lacked a matched control group. Nevertheless, the potential of LC-MS methodology to characterize myocardial ischemia was clearly demonstrated (Sabatine et al. 2005). Although early studies have produced promising results in terms of obtaining a diagnostic signature, due to confounders such as medication, the higher prevalence of the disease in men and the high dependency on diet and lifestyle, there is

still a requirement for larger-scale definitive studies in this area. Identification of an early diagnostic for CVD, or even a prognostic signature, would undoubtedly be one of the "Holy Grails" of metabolic profiling. Several population-based epidemiology studies have been designed to test hypotheses regarding the relationship between the development of hypertension, a condition that predisposes to CVD, and various lifestyle factors. In particular, the INTERMAP study (International Study of Macronutrients and Blood Pressure) was launched in 1996 to investigate the relationship of dietary intake of macronutrients and other factors to blood pressure across four countries: China, Japan, the United Kingdom and the United States (Stamler et al. 2003). This study involved collection of two 24-h urine samples from 4,680 participants, in addition to blood pressure measurements and NMR spectroscopy; the first results are beginning to emerge (Dumas et al. 2006b; Homes et al. 2007). Furthermore we have recently introduced the concept of the "Metabolome-wide association study" (Holmes et al. 2008) demonstrating broad metabolite profile screening is linked statistically to disease risk factor data to identify new molecular targets in metabolism that can be physiologically tested. Thus, whilst large epidemiological studies present a practical and logistical challenge, they are at least feasible and metabolic profiling is well suited to characterizing the metabolic phenotypes of populations, which have high risk and prevalence of pathological or prepathological conditions such as CVD and hypertension.

3.3 Metabolic Investigations of Neuropathological Disease

Disease progression in many neurodegenerative and psychological disorders is difficult to assess with batteries of cognitive or psychological tests forming part of the diagnostic for disease stage. For these pathologies, it would be ideal to have a metabolic indicator of disease stage in order to achieve the optimal therapeutic intervention strategy. In the neurodegeneration field, there are many more studies using magnetic resonance spectroscopy (MRS) of tissues than high-resolution NMR spectroscopy; however, MRS profiles lack sensitivity in comparison. Again, in experimental models such as the transgenic R6/2 mouse model of juvenile Huntington's disease (Bates et al. 1997), NMR-based metabonomic studies in the R6/2 mouse characterized the metabolic

signature of HD in several tissues and body fluids (urine, plasma, skeletal muscle, striatum, cerebral cortex, cerebellum and brain stem) at 4, 8 and 12 weeks of age (Tsang et al. 2006b). This study supported previous results obtained by Jenkins et al. (1993) using MRS, but additionally was able to resolve choline and glycerophosphocholine resonances. Choline levels were observed to decrease in most of the neuroanatomical regions analysed in the R6/2 mouse, whereas glycerophosphocholine increased suggestive of a pro-catabolic phenotype in the R6/2 mouse model. This has also been shown in a small number of HD patients where glycerophosphocholine levels correlated with disease progression (Underwood et al. 2006). Currently, the European Huntington's Disease Network is focused on collaborating across European cohorts to establish biomarkers of HD and is actively employing a systems biology approach combining transcriptomic, proteomic and metabonomic data from HD patients and age-matched controls. For schizophrenia, the metabolic profiling strategy has been taken one step further, and not only has the metabolic phenotype of the pathology been defined (Tsang et al. 2006a), but a preliminary study evaluating response to therapeutic intervention with antipsychotics has been profiled, showing that those patients treated on the first episode of the disease were able to achieve normalization of their spectral profiles (Holmes et al. 2006).

3.4 Intestinal Disorders

Although it is relatively easy to diagnose irritable bowel disorders, discriminating between them, for example Crohn's disease (CD) and ulcerative colitis, can provide more of a challenge. Moreover, monitoring the condition generally involves an invasive series of surgical procedures such as colonoscopy. Recently spectroscopic methods have been applied to stool samples from CD, ulcerative colitis, polyposis and colon cancer to achieve discriminatory profiles for each of these conditions (Marchesi et al. 2007; Scanlan et al. 2008). Faecal water profiles from each of these conditions were found to have higher levels of amino acids, lower levels of short chain fatty acids and characteristic bile acid signatures, although the specific amino acids and short chain fatty acids that changed were different for each condition.

3.5 Cancers

Cancer is one area where a specific and distinctive metabolic profile has remained elusive. The first studies performed on profiling cancer using NMR spectroscopy as a diagnostic tool were unfortunately badly confounded (Fossel et al. 1986), which resulted in avoidance of this area for many years. Two of the main problems with cancer diagnostics are the lack of specificity and the fact that many of the metabolic changes are associated with inflammation. Now, however, with the recent advances in technology, several studies on small cohorts of patients have produced promising results. In one such study Odunsi et al. were able to differentiate between patients with ovarian cancer and matched controls using NMR analysis of blood plasma (Odunsi et al. 2005), whilst in another study, excised tumour tissue was analysed using GC-MS and ovarian cancers were differentiated from borderline tumours with high sensitivity (Denkert et al. 2006). Because of the difficulty of finding cancer-specific biomarkers, several studies have employed more than one "omics" platform. For example, renal cell carcinoma has been characterized using a combined proteomic and MS-based metabolic profiling approach (Perroud et al. 2007). Due to the obvious effect on glycolysis in tumours, studies on cancer cell lines often employ ^{13}C-labelled glucose. Using this labelling approach, characterization was achieved for a breast cancer mammary epithelial line from a normal mammary epithelial line (Yang et al. 2007) using a combination of NMR and GC-MS. Whilst such studies can potentially throw light on mechanisms and aid drug target discovery, the metabolic situation is very different and inherently more complex inside the human, and one must bear in mind the biomarkers discovered via metabolic profiling of cell lines may not always be translatable.

3.6 Infectious Diseases

There have been a number of metabolic profiling initiatives in the infectious disease area, including parasitic infection, tuberculosis and meningitis (Singer et al. 2006; Glickman et al. 1994; Coen et al. 2005). In reality, infectious diseases are predominant in developing countries and therefore relatively little metabolic profiling work has been done in this

area due to financial constraints and practicality. Further complications of applying the technology in this area is the fact that multiple infection is the norm for many of these populations (Buck et al. 1978), thereby rendering extraction of a panel of biomarkers for single infection difficult, although arguably it would be preferable to profile the multiple diseases simultaneously. Since metabolic profiling is an inexpensive technology, particularly when used in an exploratory capacity with subsequent development of biomarker assays, it has great potential in the diagnosis and surveillance of infectious diseases.

4 Defining Biomarkers

In order to be truly useful, a biomarker must be quantifiable, reproducible and analytically simple to measure (Atkinson et al.). Other desirable qualities of biomarkers are that the biomarker is inexpensive to measure, its concentration or level does not vary across a large range, it is specific to the condition of interest and that it is not affected by co-morbid factors.

The capacity for metabolic profiling approaches to generate diagnostic molecular signatures has been demonstrated for a range of conditions in human studies, but many studies have been preliminary in nature and now require extensive validation across larger cohorts on individuals.

Biomarker detection plays a key role in the discovery and development of new treatments for human disease and therefore there has been a great deal of method development in the area of improving biomarker detection and extraction from large multivariate data sets.

Increased sensitivity of analytical detection is useless without the means to interpret the greater number of candidate molecules or signals generated by an analytical platform. The three major analytical platforms—GC-MS, LC-MS and NMR spectroscopy—have strengths and weaknesses that are partially determined by the nature of the disease or intervention under investigation. Although GC-MS typically requires time-consuming derivatization steps, there are several good databases for molecular identification once the data are acquired. To improve molecular identification, GC-MS data can be deconvolved using hierarchical multivariate curve resolution to resolve the spectra into

pure profiles of compounds (Jonsson et al. 2006). GC-MS is well suited to measuring diseases where targeted analysis can be applied to a set of molecules which are known to carry a signature for a particular disease, for example the measurement of organic acids for characterization of many inborn errors of metabolism. LC-MS, and more particularly UPLC-MS, with its enhanced resolution and sensitivity, provides a comprehensive signature of metabolic perturbation and is becoming increasingly useful as the databases associated with molecular identification from retention-time–*m/z* pairs improve. NMR spectroscopy is the most reproducible of the three techniques and the least prone to artefact. Sensitivity remains lower than MS methods even with the use of cryoprobes, but the technique is inherently more amenable to structural elucidation. For a few well-funded laboratories, the obvious choice is to employ all three techniques. For the rest then, a sensible choice has to be made based on cost, laboratory infrastructure and the clinical, therapeutic or nutritional areas of interest.

Whatever the platform of choice, there is a continuous stream of new chemometric and bioinformatics processing and preprocessing techniques for optimization of the analysis of spectral data including algorithms for curve resolution (Jonsson et al. 2006), peak alignment (Jonsson et al. 2005; Csenki et al. 2007; Stoyanova et al. 2004), normalization (Dieterle et al. 2006) and quantification (Vehtari et al. 2007) in order to provide the best chance of capturing potential biomarkers. Other methods focus more on the identification of correlation within the data structure in order to provide as much information as possible regarding the identity of biomarkers, for example statistical correlation spectroscopy (Cloarec et al. 2005; Crockford et al. 2006).

5 The Way Forward

Arguably, the most valuable type of biomarker is either an early diagnostic or even prognostic, i.e. one which allows detection of a disease prior to the manifestation of clinical symptoms. Identification of prognostic biomarkers can result in prevention of the development of that pathology, or even the reversal of the pathology. Several recent studies have indicated that for certain conditions metabolic profiling can un-

cover a prognostic signature. For example, from the predose ^{1}H NMR urine profile, it is possible to predict animals who will develop toxicity after an oral dose of galactosamine and to predict toxicity associated with paracetamol ingestion (Clayton et al. 2006).

Psychological and physiological stress have also been shown to predispose individuals to a number of illnesses and conditions and the concept of allostatic load, the indication of wear-and-tear on multiple biological systems as they adapt and respond, within the individual, to life's demands (McEwen 2002). It was found that for men at highest risk of mortality, a cluster of five biomarkers are usually present at elevated levels, namely CRP, IL-6, fibrinogen, norepinephrine, and epinephrine (Gruenewald et al. 2006). Application of metabolic profiling strategies to epidemiological cohorts should finally give enough power to make associations between gene-gene and gene-environment factors and the associated consequence on the metabolic phenotype.

As the world turns towards systems biology, there is a pressing need to begin to integrate multiple "-omics" data sensibly. Simply making lists of genes, proteins and metabolites that are altered by a particular disease, mathematical modelling solutions can help to extract latent information and can use aspects of one "-omics" data set to strengthen another. This integration has been attempted at a preliminary level in animal models with relatively small group sizes, but has yet to be applied on a large scale to human clinical or population studies. Examples of co-analysis of "-omics" data include integrating metabolic profiles with quantitative trait locus data in a diabetic rat model (Dumas et al. 2007), combining metabolic and proteomic data for a mouse model of prostate cancer (Rantalainen et al. 2006). Bayesian methods for establishing correlations between two disparate sets of data show promise and have the added advantage of being nonlinear. Preliminary studies linking clinically measured lipoprotein measurements to ^{1}H NMR spectra enable some resolution of the highly overlapped lipoproteins in the NMR plasma spectra (Vehtari et al. 2007).

Metabolic profiling technology has come a long way since its origins in small-scale animal studies looking at gross metabolic changes. It is now an exquisitely sensitive tool for profiling multiple dynamic biological processes and cannot only accommodate the high degree of

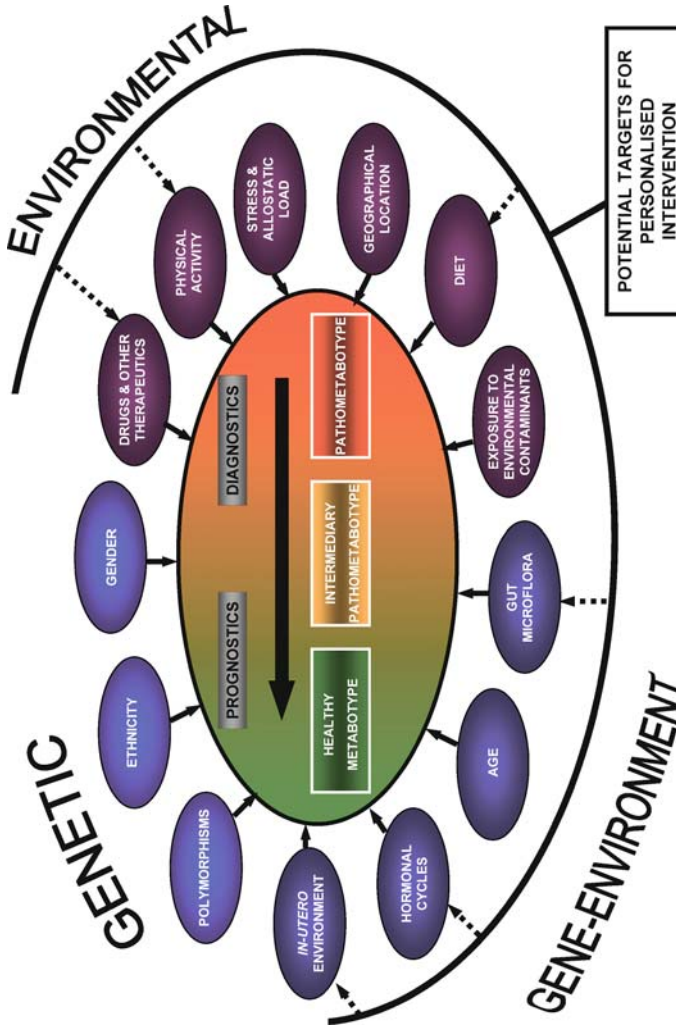

Fig. 4. Schematic showing genetic and environmental factors that contribute to the global metabolic profile of humans

metabolic variation typical of human data sets, but can help to unravel the various contributions from a range of genetic, epigenetic and environmental influences (Fig. 4). We are now standing on the threshold of a new era as this technology comes of age and is increasingly applied in the systems biology arena. With judicious application, this technology promises to deliver advances both in personalized health care and in population screening.

References

Atherton HJ, Bailey NJ, Zhang W, Taylor J, Major H, Shockcor J, Clarke K, Griffin JL (2006) A combined ^{1}H NMR spectroscopy- and mass spectrometry-based metabolomic study of the PPAR-alpha null mutant mouse defines profound systemic changes in metabolism linked to the metabolic syndrome. J Physiol Genomics 27:178–186

Atkinson AJ, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, Oates JA, Peck CC, Schooley RT, Spilker BA et al. Biomarkers and surrogate end points: preferred definitions and conceptual framework

Bailey NJ, Sampson J, Hylands PJ, Nicholson JK, Holmes E (2002) Multicomponent metabolic classification of commercial feverfew preparations via high-field ^{1}H-NMR spectroscopy and chemometrics. Planta Med 68:734–738

Bales JR, Higham DP, Howe I, Nicholson JK, Sadler PJ (1984) Use of high-resolution proton nuclear magnetic resonance spectroscopy for rapid multicomponent analysis of urine. Clin Chem 30:426–432

Bates GP, Mangiarini L, Mahal A, Davies SW (1997) Transgenic models of Huntington's disease. Hum Mol Genet 6:1633–1637

Bollard ME, Holmes E, Lindon JC, Mitchell SC, Branstetter D, Zhang W, Nicholson JK (2001) 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

Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HW, Clarke LS, Schofield PM, McKilligin E, Mosedale DE, Grainger DJ (2002) Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using ^{1}H-NMR-based metabonomics. Nat Med 8:1439–1444

Buck AA, Anderson RI, MacRae AA (1978) Epidemiology of poly-parasitism. I. Occurrence, frequency and distribution of multiple infections in rural communities in Chad, Peru, Afghanistan, and Zaire. Tropenmed Parasitol 29:61–70

Clayton TA, Lindon JC, Cloarec O, Antti H, Charuel C, Hanton G, Provost JP, Le-Net JL, Baker D, Walley RJ, Everett JR, Nicholson JK (2006) Pharmacometabonomic phenotyping and personalized drug treatment. Nature 440: 1073–1075

Cloarec O, Dumas ME, Craig A, Barton RH, Trygg J, Hudson J, Blancher C, Gauguier D, Lindon JC, Holmes E, Nicholson J (2005) Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H-NMR data sets. Anal Chem 77:1282–1289

Coen M, O'Sullivan M, Bubb WA, Kuchel PW, Sorrell T (2005) Proton nuclear magnetic resonance-based metabonomics for rapid diagnosis of meningitis and ventriculitis. Clin Infect Dis 41:1582–1590

Crockford DJ, Holmes E, Lindon JC, Plumb RS, Zirah S, Bruce S, Rainville P, Stumpf CL, Nicholson JK (2006) Statistical heterospectroscopY (SHY), a new approach to the integrated analysis of NMR, UPLC-MS data sets: application in metabonomic toxicology studies. Anal Chem 78:363–371

Csenki L, Alm E, Torgrip RJ, Aberg RJ, Nord LI, Schuppe-Kostinen I, Lindberg J (2007) Proof of principle of a generalized fuzzy Hough transform approach to peak alignment of one dimensional ^{1}H NMR data. Anal Bioanal Chem 389:8775–885

Denkert C, Budczies J, Kind T, Weichert W, Tablack P, Sehouli J, Niesporek S, Könsgen D, Dietel M, Fiehn O (2006) Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors. J Mass Spectrom 41:1546–1553

Dieterle F, Ross A, Schlotterbeck G, Senn H (2006) Probabalistic quotient normalization as a robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics. Anal Chem 78:4281–4290

Dumas M-E, Barton RH, Toye A, Cloarec O, Craig A, Blancher C, Rothwell A, Fearnside J, Tatoud R, Blanc V, Lindon JC, Mitchell S, Holmes E, McCarthy MI, Scott J, Gaugier D, Nicholson JK (2006a) Metabolic profiling reveals a contribution of gut microbiota to insulin resistance phenotype in mice. Proc Natl Acad Sci U S A 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 for large-scale epidemiological research: the INTERMAP study. Anal Chem 78:2199–2208

Dumas ME, Wilder SP, Bihoreau MT, Barton RH, Fearnside JF, Argoud K, D'Amato L, Wallis RH, Blancher C, Keun HC, Baunsgaard D, Scott J, Sidelmann UG, Nicholson JK, Gauguier D (2007) Direct quantitative trait locus mapping of mammalian metabolic phenotypes in diabetic and normoglycemic rat models. Nat Genet 39:666–672

Ebbels TM, Keun HC, Beckonert OP, Bollard ME, Lindon JC, Holmes E, Nicholson JK (2007) Prediction and classification of drug toxicity using probabilistic modeling of temporal metabolic data: the consortium on metabonomic toxicology screening approach. J Proteome Res 6:4407–4422

Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmjtzer L (2000) Metabolite profiling for plant functional genomics. Nat Biotechnol 18: 1157–1161

Fossel ET, Carr JM, McDonagh J (1986) Detection of malignant tumors. Water-suppressed proton nuclear magnetic resonance spectroscopy of plasma. N Engl J Med 315:1369–1376

Foxall PJ, Lenz EM, Lindon JC, Neild GH, Wilson ID, Nicholson JK (1996) Nuclear magnetic resonance and high-performance liquid chromatography-nuclear magnetic resonance studies on the toxicity and metabolism of ifosfamide. Ther Drug Monit 18:498–505

Garrod SL, Humpfer E, Spraul M, Connor SC, Polley S, Connelly J, Lindon JC, Nicholson JK, Holmes E (1999) High resolution magic-angle-spinning ^{1}H NMR spectroscopic studies on intact rat renal cortex and medulla. Magn Res Med 41:1108–1118

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, Alpk:ApfCD mouse. FEBS Lett 484:169–174

Glickman SE, Kilburn JO, Butler WR, Ramos LS (1994) Rapid identification of mycolic acid patterns of mycobacteria by high-performance liquid chromatography using pattern recognition software and a *Mycobacterium* library. J Clin Microbiol 32:740–745

Gruenewald TL, Seeman TE, Ryff CD, Singer BH (2006) Early warning biomarkers: what combinations predict later life mortality? Proc Natl Acad Sci 103:14158–14163

Guthrie R, Susi A (1963) A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. Pediatrica 132:328–343

Holmes E, Tsang TM, Huang JTJ, Leweke M, Koethe D, Gerth CW, Nolden BM Gross S, Schreiber D, Nicholson JK, Bahn S (2006) Metabolic profiling of CSF: a new tool for monitoring schizophrenia and response to therapeutic intervention. J PLOS Med 3:1420–1428

Holmes E, Foxall PJD, Nicholson JK, Neild GH, Brown SM, Beddell CR, Sweatman BC, Rahr E, Lindon JC, Spraul M, Neidig P (1994) 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

Holmes E, Loo RL, Cloarec O, Coen M, Tang H, Maibaum E, Bruce S, Chan Q, Elliott P, Stamler J, Wilson ID, Lindon JC, Nicholson JK (2007) Detectionof urinary drug metabolite (xenometabolome) signatures in molecular epidemiology studies via statistical total correlation spectroscopy. Anal Chem 79:2629–2640

Holmes E, Loo RL, Stamler J, Bictash M, Yap IK, Chan Q, Ebbels T, De Iorio M, Brown IJ, Veselkov KA, Daviglus ML, Ueshima H, Nicholson JK, Elliott P (2008) Human metabolic phenotype diversity and its association with diet and blood pressure. Nature 453:396–400

International Huntington Association and the World Federation of Neurology Research Group on Huntington's chorea (1994) Guidelines for the molecular genetics predictive test in Huntington's disease. J Med Genet. 31:555–559

Jenkins BG, Koroshetz WJ, Beal MF, Rosen BR (1993) Evidence for impairment of energy metabolism in vivo in Huntington's disease using localized ^{1}H NMR spectroscopy. Neurology 43:2689–2695

Jonsson P, Bruce SJ, Moritz T, Trygg J, Sjostrom M, Plumb R, Granger J, Maibaum E, Nicholson JK, Holmes E, Antti H (2005) Extraction, interpretation and validation of information for comparing samples in metabolic LC/MS data sets. Analyst 130:701–707

Jonsson P, Johansson ES, Wuolikainen A, Lindberg J, Schuppe-Koistinen I, Kusano M, Sjöström M, Trygg J, Moritz T, Antti H (2006) Predictive metabolite profiling applying hierarchical multivariate curve resolution to GC-MS data – a potential tool for multi-parametric diagnosis. J Proteome Res 5:1407–1414

Kirschenlohr HL, Griffin JL, Clarke SC, Rhydwen R, Grace AA, Schofield PM, Brindle KM, Metcalfe JC Proton (2006) NMR analysis of plasma is a weak predictor of coronary artery disease. Nat Med 12:705–710

Koschorek F, Offermann W, Stelten J, Braunsdorf WE, Stellar U, Gremmel H, Liebfritz D (1993) High resolution ^{1}H NMR spectroscopy of cerebrospinal fluid in spinal disease. Neurosurg Rev 16:302–315

Lamers RJ, DeGroot J, Spies-Faber EJ, Jellema RH, Kraus VB, Verzijl N, TeKoppele JM, Spijksma GK, Vogels JT, van der Greef J, van Nesselrooij JH (2003) Identification of disease- and nutrient-related metabolic fingerprints in osteoarthritic Guinea pigs. J Nutr 133:1776–1780

Lenz EM, Bright J, Wilson ID, Morgan SR, Nash AF (2003) A ^{1}H NMR-based metabonomic study of urine and plasma samples obtained from healthy human subjects. J Pharm Biomed Anal 33:1103–1115

Maher AD, Zirah SF, Holmes E, Nicholson JK (2007) Experimental and analytical variation in human urine in ^{1}H NMR spectroscopy based metabolic phenotyping studies. Anal Chem 79:5204–5211

Marchesi J, Holmes E, Khan F, Kochhar S, Scanalan P, Shanahan F, Wilson ID, Wang Y (2007) Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. J Proteome Res 6:546–551

McEwen B (2002) Sex, stress and the hippocampus: allostasis, allostatic load and the aging process. Neurobiol Aging 23:921–939

Neild GH, Foxall PJ, Lindon JC, Holmes EH, Nicholson JK (1997) Uroscopy in the 21st century: high field NMR spectroscopy. Nephrol Dial Transplant 12:404–417

Nemoto T, Ando I, Kataoka T, Arifuku K, Kanazawa K, Natori Y, Fujiwara M (2007) NMR metabolic profiling combined with two-step principal component analysis for toxin-induced diabetes model rat using urine. J Toxicol Sci 32:429–435

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 spectroscopic data. Xenobiotica 11:181–1189

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

Odunsi K, Wollman RM, Ambrosone CB, Hutson A, McCann SE, Tammela J, Geisler JP, Miller G, Sellers T, Cliby W, Qian F, Keitz B, Intengan M, Lele S, Alderfer JL (2005) Detection of epithelial ovarian cancer using H-1-NMR-based metabonomics. Int J Cancer 113:782–788

Perroud B, Lee J, Valkova N, Dhirapong A, Lin PY, Fiehn O, Kültz D, Weiss RH (2007) Pathway analysis of kidney cancer using proteomics and metabolic profiling. Anal Chem 79:6995–7004

Plumb RS, Stumpf CL, Granger JH, Castro-Perez J, Haselden JN, Dear GJ (2003) Use of liquid chromatography/time-of-flight mass spectrometry and multivariate statistical analysis shows promise for the detection of drug metabolites in biological fluids. Rapid Commun Mass Spectrom 17:2632–2638

Plumb RS, Granger JH, Stumpf CL, Johnson KA, Castro-Perez J, Wilson ID, Nicholson JK (2005) A rapid screening approach to metabonomics using UPLCand oa-TOF mass spectrometry: application to age, gender and diurnal variation in normal/Zucker obese rats and black and white nude mice. Analyst 130:844–849

Rantalainen M, Cloarec O, Beckonert O, Wilson ID, Jackson D, Tonge R, Rowlinson R, Rayner S, Nickson J, Wilknson RW, Mills JD, Trygg J, Nicholson JK, Holmes E (2006) Statistically integrated metabonomic-proteomic studies on human prostate cancer xenograft model in mice. J Proteome Res 5:2642–2655

Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, Wiegand R, Berriz GF, Roth FP, Gerszten RE (2005) Metabolomic identification of novel biomarkers of myocardial ischemia. Circulation 112(25):3868–3875

Saric J, Wang Y, Li JV, Coen M, Utzinger J, Marchesi JR, Keiser J, Veselkov K, Lindon JC, Nicholson JK, Holmes E (2008) Species variation in the fecal metabolome gives insight into differential gastrointestinal function. J Proteome Res 7:352–360

Scanalan PD, Shanahan F, Clune Y, Collins JK, O'Sullivan GC, O'Riordan M, Holmes E, Wang Y, Marchesi J (2008) Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis. Env Microbiol 10:789–798

Singer BH, Utzinger J, RYff CD, Wang Y, Holmes E (2006) Exploiting the potential of metabonomics in large population studies: three venues. In: Lindon JC, Nicholson JK, Holmes E (eds) The handbook of metabonomics and metabolomics, 1st edn. Elsevier, pp. 289–327

Slupsky CM, Rankin KN, Wagner J, Fu H, Chang D, Weljie AM, Saude EJ, Lix B, Adamko DJ, Shah S, Greiner R, Sykes BD, Marrie TJ (2007) Investigations of the effects of gender, diurnal variation and age in human urinary metabolomic profiles. Anal Chem 79:6995–7004

Stamler J, Elliott P, Dennis B, Dyer A, Kesteloot H, Liu K, Ueshima H, Zhou BF, for the INTERMAP Research Group. (2003) INTERMAP: background, aim, design, methods and descriptive statistics (non-dietary). J Hum Hypertension 17:591–608

Stoyanova R, Nicholson JK, Lindon JC, Brown TR (2004) Sample classification based on Bayesian spectral decomposition of metabonomic NMR data sets. Anal Chem 76:3666–3674

Teague C, Holmes E, Maibaum E, Nicholson J, Tang H, Chan Q, Elliot P, Wilson I (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

Teague CR, Dhabhar FS, Beckwith-Hall B, Holmes E, Powell J, Cobain M (2006) Metabonomic studies on the physiological effects of acute and chronic psychological stress in Sprague-Dawley rats. J Proteome Res 6:2080–2093

Teahan O, Gamble S, Holmes E, Waxman J, Nicholson JK, Bevan C, Keun HC (2006) Impact on analytical bias in metabonomic studies of human blood serum and plasma. Anal Chem 78:4307–4318

Tsang TM, Griffin JL, Haselden J, Holmes E (2005) Metabolic characterization of distinct neuroanatomical regions in rats by magic angle spinning H-1 nuclear magnetic resonance spectroscopy. Magn Res Med 53:1018–1024

Tsang TM, Huang JT, Holmes E, Bahn S (2006a) Metabolic profiling of plasma from discordant schizophrenia twins: correlation between lipid signals and global functioning in female schizophrenia patients. J Proteome Res 5:756–760

Tsang TM, Woodman B, McLoughlin GA, Griffin JL, Tabrizi SJ, Bates GP Holmes E (2006b) Metabolic characterization of the R6/2 transgenic mouse model of Huntington's disease by high-resolution MAS ^{1}H NMR spectroscopy. J Proteome Res 5:483–492

Underwood BR, Broadhurst D, Dunn WB, Ellis DI, Michell AW, Vacher C et al. (2006) Huntington disease patients and transgenic mice have similar procatabolic serum metabolite profiles. Brain 129:877–886

Vehtari A, Mäkinen VP, Soininen P, Ingman P, Mäkelä SM, Savolainen MJ, Hannuksela ML, Kaski K, Ala-Korpela M (2007) A novel Bayesian approach to quantify clinical variables and to determine their spectroscopic counterparts in ^{1}H NMR metabonomic data. BMC Bioinformatics 8 [Suppl 2]:S8

Vilasi A, Cutillas PR, Maher AD, Zirah SF, Capasso G, Norden AW, Holmes E, Nicholson JK, Unwin RJ (2007) Combined proteomic and metabonomic studies in three genetic forms of the renal Fanconi syndrome. Am J Physiol Renal Physiol 293:456–467

Wang Y, Holmes E, Comelli E, Fotopoulos G, Dorta G, Tang H, Rantalainen M, Lindon JC, Corthesy-Theulaz Fay LB, Kochhar S, Nicholson JK (2008) Topographical variation in metabolic signatures of human gut epithelial biopsies revealed by high-resolution magic-angle-spinning ^{1}H NMR spectroscopy. J Proteome Res 6:3944–3951

Wang YL, Tang HR, Nicholson JK et al. (2004) Metabolomic strategy for the classification and quality control of phytomedicine: a case study of chamomile flower (*Matricaria recutita* L) Planta Medica 70:250–255

Williams R, Lenz EM, Evans JA, Granger JH, Plumb RS, Stumpf CL (2005) A combined ^{1}H NMR, HPLC-MS-based metabonomics study of urine from obese (fa/fa) Zucker and normal Wistar-derived rats. J Pharm Biomed Anal 38:465–471

Williams RE, Lenz EM, Rantalainen M, Wilson ID (2006) The comparative metabonomics of age-related changes in the urinary composition of male Wistar-derived and Zucker (fa/fa) obese rats. Mol BioSystems 2:193–202

Yang C, Richardson AD, Smith JW, Osterman A (2007) Comparative metabolomics of breast cancer. Pac Symp Biocomput 181–192

Yi LZ, He J, Liang YZ, Yuan DL, Chau FT (2006) Plasma fatty acid metabolic profiling and biomarkers of type 2 diabetes mellitus based on GC/MS, PLS-LDA. FEBS Lett 580:6837–6845