

Workshop Report

Metabolomics and human nutrition

Sandy Primrose¹, John Draper², Rachel Elsom^{3*}, Verity Kirkpatrick³, John C. Mathers⁴, Chris Seal⁴, Manfred Beckmann⁴, Sumanto Haldar⁴, John H. Beattie⁵, John K. Lodge⁶, Mazda Jenab⁷, Hector Keun⁸ and Augustin Scalbert⁹

¹Prosciagen Consulting, High Wycombe, UK

²IBERS, University of Aberystwyth, Aberystwyth, UK

³Nutrition Division, Food Standards Agency, 125 Kingsway, London WC2 6NH, UK

⁴Human Nutrition Research Centre, Newcastle University, Newcastle upon Tyne, UK

⁵Rowett Institute of Nutrition and Health, University of Aberdeen, Aberdeen, UK

⁶Cranfield Health, Cranfield University, Cranfield, UK

⁷International Agency for Research on Cancer (WHO), Lyon, France

⁸Imperial College, London, UK

⁹French National Institute for Agricultural Research (INRA), Clermont-Ferrand, France

(Received 28 July 2010 – Revised 29 July 2010 – Accepted 29 July 2010 – First published online 24 January 2011)

Abstract

The present report summarises a workshop convened by the UK Food Standards Agency (Agency) on 25 March 2010 to discuss the current Agency's funded research on the use of metabolomics technologies in human nutrition research. The objectives of this workshop were to review progress to date, to identify technical challenges and ways of overcoming them, and to discuss future research priorities and the application of metabolomics in public health nutrition research and surveys. Results from studies nearing completion showed that by using carefully designed dietary and sampling regimens, it is possible to identify novel biomarkers of food intake that could not have been predicted from current knowledge of food composition. These findings provide proof-of-principle that the metabolomics approach can be used to develop new putative biomarkers of dietary intake. The next steps will be to validate these putative biomarkers, to develop rapid and inexpensive assays for biomarkers of food intake of high public health relevance, and to test their utility in population cohort studies and dietary surveys.

Key words: Metabolomics: Diet: Biomarkers: Food Standards Agency

On 25 March 2010, the UK Food Standards Agency (Agency) convened a workshop, chaired by Dr Augustin Scalbert, to review the Agency's funded nutrition research projects using metabolomics technologies. This workshop aimed to establish priorities and directions for future research in the field.

Diet is a key modifiable factor affecting the long-term health and well-being of individuals. To determine the effects of specific dietary components and/or patterns on health, detailed knowledge of habitual dietary exposure is required. Currently, nutritional studies investigating associations between diet and health rely on study participants either

reporting or recalling their dietary intake over a period of time using FFQ or diet diaries. However, these approaches have limitations because individuals rarely keep accurate and complete records of what they eat, and the task of keeping records can influence the foods that individuals eat during the recording period^(1,2). One potential solution is to examine body fluids such as plasma and urine for biomarkers specific to particular foods or diets using metabolomics approaches. However, the lack of robust, validated biomarkers of different foods is currently a barrier.

Following consumption, many of the chemical components in foods are absorbed either directly or after

Abbreviations: Agency, UK Food Standards Agency; EPIC, European Prospective Investigation into Cancer and Nutrition; FIE-MS, flow infusion electrospray mass spectrometry; GC-TOF-MS, gas chromatography-time-of-flight-mass spectrometry; MEDE, METabolomics to characterise Dietary Exposure; SOP, Standard operating procedures.

* **Corresponding author:** R. Elsom, fax +44 20 7972 4346, email rachel.elsom@dh.gsi.gov.uk

digestion, undergo transformations in the gastrointestinal tract and/or liver and appear in the plasma. Later, these compounds may appear, with or without further transformation and metabolism, in the urine. These metabolites derived from food form what has been called the food metabolome^(3,4). Investigation of the food metabolome in biofluids has the potential to give objective information about the short-term or long-term dietary intake of individuals, to identify potential biomarkers and to reduce the current reliance on subjective dietary records.

Dr Scalbert presented two ways of identifying potential biomarkers: a targeted 'top-down' approach involving analysis of the composition of foods and selection of candidate molecules that are likely to appear in the plasma or urine^(5,6) and an untargeted 'bottom-up' approach. In the untargeted approach, volunteers consume different diets or foods, and metabolomics technologies are used to identify potential biomarkers from the food metabolome. The advantage of this approach⁽⁷⁻⁹⁾ is that it makes no assumptions about candidate molecules. Experience has shown that where biomarkers have been identified using the untargeted approach, the candidate molecules identified often are unexpected based on current incomplete knowledge of the composition of the foods consumed, their absorption and subsequent metabolism.

However, where an untargeted approach is used in biomarker identification studies, particularly those using urine, care is required in selecting the experimental procedures. Other issues raised included the specificity of biomarkers for particular foodstuffs, the variability of composition of individual foods and inter-individual variation in the absorption and metabolism of different food components.

In closing the introduction to the workshop, Dr Scalbert made two key points. First, that proper biocuration of food composition and metabolomics data will be essential, and second, it is highly desirable that research moves to testing metabolomics technologies in population cohort studies and dietary surveys.

Identification of biomarkers of dietary intake using metabolomics

Professor John Mathers presented the protocols used in the MEtabolomics to characterise Dietary Exposure (MEDE) study. This study focused on providing 'proof-of-principle' that an untargeted ('bottom-up') metabolomics approach could identify specific chemical signals typical of individual meal components. To minimise the effects of potential confounding factors, carefully prepared standard operating procedures (SOP) were developed in parallel with the GrainMark study (see later). These SOP covered key aspects of volunteer handling, volunteer behaviours (including water consumption), test-meal preparation and timing, biofluid (blood, urine and saliva) collection, processing and storage to reduce potential confounding

effects. The SOP developed in the MEDE study were applied to the other Agency-funded projects.

Three studies were undertaken. In MEDE 1, twelve volunteers were studied, each on one occasion to assess the robustness of the protocols. In MEDE 2, a further twelve volunteers underwent two identical test procedures, receiving a standard evening meal (a ready-prepared roast chicken dinner with a chocolate éclair and still mineral water) and a standard breakfast (orange juice, croissant, tea and cornflakes with milk) several weeks apart to investigate the intra-individual variation between the two cohorts. In MEDE 3, twenty-four volunteers were allocated to receive the standard evening meal and breakfast twice or variations of the standard breakfast (known as 'test breakfast') in which one component (cornflakes with milk) was replaced with one of four food components of high public health interest (a cruciferous vegetable (broccoli), an oily fish (smoked salmon), berry fruit (raspberries) or a whole-grain cereal food (wheat biscuits)) in a four 6 × 6 Latin square design. This part of the study tested whether metabolomics analysis could distinguish between the metabolome from the standard breakfast and the test breakfasts. In addition, data from MEDE 3 were used to identify novel putative biomarkers of intake of each of the test foods.

Professor John Draper presented the detailed results obtained from analysis of the MEDE study samples. It was observed that metabolome models of plasma/serum are dominated by individual differences in blood composition, whereas for postprandial urine samples (which are relatively easy to collect non-invasively), the metabolome content is dominated by recent dietary exposure. Two techniques were used for metabolite fingerprinting (flow infusion electrospray mass spectrometry (FIE-MS) and gas chromatography-time-of-flight-mass spectrometry (GC-TOF-MS)), and both were informative. GC-TOF-MS resulted in more robust models for discrimination of samples collected after the different test breakfasts. However, FIE-MS allowed the analysis of many more samples in a much shorter period of time. Key findings from the initial analysis of the MEDE study samples were as follows:

- (1) Consistent fasting and postprandial changes were detected in the urine in all subjects following consumption of the standard breakfast and when test meals were consumed several weeks apart.
- (2) Urine and plasma samples collected 3 h after test meals were the most informative about recent dietary exposure.
- (3) The same metabolite signals were always highly ranked when the same standard breakfast was consumed. This meant that they could be 'filtered' from the data and made it much easier to identify metabolite signals associated with the test foods.

- (4) Inter-individual differences were also observed, possibly because of differences in genotype for genes encoding enzymes and other proteins involved in the transport, metabolism and excretion of food constituents.
- (5) Several behavioural, anthropomorphic and physiological factors appeared to have a minimal effect on biomarker discovery.

The conclusion from this part of the study was that the analytical procedures and data analysis tools used are robust and suitable for biomarker discovery. The data generated have been mined for biomarkers of particular dietary ingredients. Potential putative biomarkers were identified for consumption of orange juice, raspberries, smoked salmon and broccoli. This demonstrated the considerable potential of metabolomics as a route to identify and develop novel and robust biomarkers of dietary exposure.

Three findings from the data analysis were particularly striking. First, biomarkers that might have the potential to be general indicators of total fruit and vegetable consumption were identified, and it is worth noting that many of these represented metabolites biotransformed in the body which would not have been discovered if analysis had been limited to any pre-determined chemical found in the foods concerned. Second, for the twenty-four volunteers as a whole, it was not possible to distinguish the presence of cornflakes from wheat biscuits (a whole-grain cereal) in the standard breakfast. However, when a subset of volunteers with similar metabolite profiles was examined, good separation of the two breakfast cereals was achieved. Although it is possible that the analytical approaches used may not have been sufficiently sensitive to distinguish between the two breakfast cereals, an alternative explanation could be that inter-individual differences in the metabolism of food (metabotypes – which may be genetically controlled) obscured differences in metabolite fingerprints between the breakfast cereals. Finally, it was observed that provision of adequate water intake following the evening meal minimised variation in the concentration of solutes in the following morning's fasting urine samples.

Professor Chris Seal and Dr Manfred Beckmann presented data on the identification and utility of biomarkers for monitoring whole grain intake (the GrainMark study). This study involved both a traditional targeted approach 'top-down' and an untargeted 'bottom-up' approach utilising metabolomics. In the targeted approach, an assessment was made of the utility of alkylresorcinols and mammalian lignans (enterodiols and enterolactone) as biomarkers of whole grain intake. Alkylresorcinols are phenolic lipids, which are found exclusively in the bran fractions of cereals (except oats and rice of the commonly consumed cereals). Enterodiol and enterolactone are metabolites of the plant lignans formed in the large intestine by the gut microbiota. Plant lignans are found in a variety of plant foods including fruits, vegetables, seeds and whole-grain cereals. In the

untargeted approach, the objective was to identify plasma and urinary metabolome profiles indicative of diets rich in whole-grain wheat and whole-grain rye.

The GrainMark study involved sixty-eight healthy, non-smoking, normal-weight adult volunteers. After a 4-week washout period in which volunteers avoided all whole-grain foods, they consumed three servings of either whole-grain wheat or whole-grain rye per day for 4 weeks. Volunteers then doubled their dose of whole-grain foods for the next 4 weeks. Blood and urine samples were collected at the beginning and end of each intervention period, and the protocols developed in parallel with the MEDE study were used for the standard evening meal and urine sample collection. Samples were analysed for alkylresorcinols (GC-MS) and lignan metabolites (HPLC), and metabolite fingerprints were generated by FIE-MS and GC-TOF-MS.

The targeted approach was successful as the plasma concentrations of alkylresorcinols were strongly and significantly correlated with whole grain intake for both wheat and rye groups. In addition, the ratio of C17:C21 lipid chain-length alkylresorcinols was a good indicator of the type of whole grain consumed. The plasma concentrations of lignan metabolites were less informative, but urinary excretion of these metabolites could be used as a biomarker of whole grain intake.

When the untargeted approach was used to analyse GrainMark samples, a correlation between the metabolome profile and whole grain intake was found using both the GC-TOF-MS and FIE-MS methods. However, quality assessment of the data showed that only GC-TOF-MS had real utility. This result was not too surprising given that FIE-MS involves sample classification by rapid, global analysis of metabolites without an initial chromatography step. In addition, the metabolites that are potential biomarkers of bran consumption are easier to detect by GC-TOF-MS, which also is a global profiling tool. The aim of the GrainMark study was to look for potentially subtle differences in metabolite profile associated with long-term dietary changes. As such, samples from this study are akin to those which might be available from population cohort studies and dietary surveys.

Analysis of spot urine samples proved to be more informative than plasma analysis, as this yielded a greater number of metabolites that discriminated strongly the different intervention phases. This is an important finding because it reduces the likelihood that 24 h urine collection will be required. Two of the promising metabolites identified were 3,5-dihydroxybenzoic acid and 3-(3,5-dihydroxyphenyl)-1-propanoic acid, both of which are metabolites of alkylresorcinols. A number of other metabolites whose identity is currently unknown were found to be indicative of whole-grain rye intake. This provides evidence that use of the untargeted approach in nutrition biomarker studies can generate additional, and potentially more informative, biomarkers to those identified by the targeted approach.

Professor John Beattie described a project that is underway in collaboration with the University of Aberystwyth and the University of Colorado: the Human Biomarkers of Zinc Status study. Some groups of the UK population have a low Zn intake, and this may have important public health implications. Unfortunately, there are no sensitive, specific and reliable biomarkers of Zn dietary exposure and status. The study is being conducted in 100 healthy male volunteers who will be supplied with all their food during the 10-week study period. Throughout the trial, twenty of the volunteers will be on a high-Zn diet. The remaining eighty volunteers will be maintained on a low-Zn diet throughout the trial but will be randomly allocated to four groups of twenty individuals who will receive supplements of 0, 3, 6 or 9 mg of Zn/d, for the final 4 weeks. Blood and urine samples will be analysed using metabolite fingerprinting and profiling to search for potential biomarkers of Zn status.

Metabolomics analysis will identify metabolic pathways that are influenced by Zn status, which will be quantified by measuring the exchangeable Zn pool size using a stable isotope tracer technique. Blood plasma proteomics will also be used to search for potential protein biomarkers of Zn status and to investigate associations between protein and metabolite profiles. Thus, the Human Biomarkers of Zinc Status study not only focuses on products of endogenous metabolism, but also considers Zn-related influences on the metabolism of food-derived compounds. This study is due to report at the end of 2012.

Dr John Lodge described a project that is underway in collaboration with the University of Aberystwyth to identify dose-responsive biomarkers of sucrose intake in healthy female volunteers. Carbohydrate represents the major energy source in the UK diet, but there are no accurate biomarkers of carbohydrate intake. Highly refined or extrinsic sugar intake can lead to consumption of excess energy, which could increase the risk of weight gain and so accurate measures of intake are of major nutritional importance. As in the MEDE study, volunteers are given a standard evening meal and urine is collected overnight. The next day, volunteers are randomised (thirty volunteers per group) to consume a drink containing either 0, 50 or 100 g of sucrose in 500 ml of water while in the fasted state. This is followed by a standard lunch and evening meal. Blood and urine samples are taken immediately before sucrose ingestion and at intervals for the next 24 hours, and these biofluids will be analysed by metabolite fingerprinting. This project is due to report in the second half of 2011.

In each of the studies described earlier, the participants were required to conform to a controlled dietary regimen. However, a key objective in the application of metabolomics technology to nutrition studies is to develop new biomarkers that can be used in free-living populations to estimate intakes of specific foods or dietary patterns. Such biomarkers may be used to validate, or as an alternative to, conventional measures of dietary intake in

large-scale surveys or to explore diet–health associations in observational studies, particularly in large multi-centre or international settings. There is information in the literature⁽¹⁰⁾ showing that metabolic profiles can be used to distinguish between subjects from different population groups and between those with high vegetable protein intake *v.* high animal protein intake.

Dr Mazda Jenab and Dr Hector Keun described the potential application of biological samples collected in observational studies for metabolomics and biomarker discovery. Specifically, they highlighted the suitability of a cross-sectional component of the European Prospective Investigation into Cancer and Nutrition (EPIC) study to assist in translating metabolomics studies to whole populations. The EPIC cross-sectional study is a subset of the larger EPIC cohort study^(11,12), and is composed of subjects (stratified by sex, age range and smoking status) recruited from sixteen European countries from Scandinavia to southern Greece and including the UK with different dietary and lifestyle patterns. Baseline dietary intake information for these subjects is available from both detailed country-specific FFQ and standardised and computerised 24 h dietary recalls collected using the EPIC-Soft program. Information on lifestyle (including smoking, alcohol consumption and physical activity) was collected at recruitment using standardised questionnaires. Anthropometric measures as well as biological samples (stored under liquid nitrogen) were also collected at baseline and are available for all subjects. In addition, a large number of traditional dietary biomarkers (carotenoids, tocopherols, retinol and phospholipid fatty acids) have already been measured in the dataset^(13–15). To explore the suitability of EPIC biological samples for metabolomics analyses, a pilot study was conducted in a small subset of samples. Metabolic profiles of blinded serum samples from a total of forty subjects from the British, French and Italian cohorts were analysed by NMR and MS. In addition, ten blinded samples were analysed in duplicate along with several quality-control samples. The results of this pilot study showed that:

- (1) most samples could be differentiated by the country of origin;
- (2) similar discrimination could be obtained with either NMR or MS;
- (3) blinded duplicate samples could be identified, indicating methodological robustness.

These results are encouraging and suggest that a metabolomic fingerprinting approach may be applicable to population studies using biological samples not previously collected specifically with metabolomics analyses in mind.

Conclusions and key issues

From the workshop, it is clear that suitable methodology for volunteer handling and for sample collection,

preparation and analysis is available to permit use of an untargeted metabolomics approach to identify new biomarkers of dietary exposure. However, three caveats must be considered. First, adequate quality-control methods need to be developed to ensure that the analytical procedures used are suitable for the analysis of large batches or to compare different batches of samples. Second, careful thought needs to be given to the quality of all consumables and SOP used in sample collection, preparation, storage and analysis to facilitate comparison of the results obtained in various independent studies. Third, existing mass spectral databases lack information on many of the food-derived metabolites identified in metabolic fingerprints. Thus, such databases need further development to facilitate rapid identification of potential biomarkers of food intake.

Metabolomics methods generate large amounts of data and the Agency's previously funded research developed methods for automated analysis of such data^(16–18). These methods permit an assessment of the quality of data models and those mass spectral signatures that explain differences between sample classes. These methods were used in the MEDE and GrainMark studies, and their utility has been confirmed. Nevertheless, much more attention should be paid to the biocuration of metabolomics data, with the aims of archiving both existing knowledge on the food metabolome (nature of compounds, concentration in foods and spectral data) in a form that can be easily queried, and the new data generated in new metabolomics studies.

Although both plasma and urine can be analysed for the presence of food-derived biomarkers, with the present state of development of nutritional metabolomics, urine appears to be the preferred biofluid because:

- (1) a greater diversity of metabolites derived from the food metabolome is observed in the urine;
- (2) collecting urine samples is relatively easy and non-invasive;
- (3) spot urine samples collected at particular times in relation to meals and sleep period ('behavioural phase' urines) can be informative, and collection of 24 h urine samples may not be required.

However, the possibility of using plasma to identify biomarkers of food intake should not be ignored. For example, the lipophilic fraction of the plasma (food) metabolome may contain several biomarkers of long-term food consumption such as alkylresorcinols or carotenoids. Looking to the future, increases in the sensitivity of metabolite measurements in the plasma may enable characterisation of long-term (habitual) dietary exposure by detecting and quantifying metabolites which have been sequestered in body tissues and which appear in the blood through normal tissue turnover.

In the past, the Agency has funded a number of nutritional intervention trials that have not included

metabolomics analysis. Samples from these trials have been stored, and they could now be re-examined to determine if relevant biomarkers can be identified. A potential confounding issue is the long-term stability of each metabolite originally present in the samples. However, such possible instability should not be a barrier to analysis provided it is acknowledged as an issue and the storage history of samples is known.

As a number of participants in the workshop highlighted, there is now a need to move from the discovery of potential biomarkers of food intake in carefully controlled study groups to their validation and application in dietary surveys and prospective cohort studies where there are also data on health outcomes. The EPIC study has provided proof-of-principle that the metabolomics methodology can be translated to these types of studies. However, untargeted metabolomics approaches, particularly those using MS, are not yet sufficiently robust to analyse very large numbers of samples. Such untargeted metabolomics approaches could be applied to a limited number of selected samples in cross-sectional, case-control or nested studies within large cohort studies⁽¹⁹⁾, which have robust data on food intake obtained using conventional dietary assessment tools to help in the validation of novel putative biomarkers of food intake. Other validation approaches may include additional testing in controlled dietary intervention studies. Following validation, the next steps would be the development of targeted analytical methods for low-cost, high-throughput quantification of the new biomarkers and their application in larger cohort studies and dietary surveys.

Lastly, individuals differ in the way in which they handle food metabolites because of differences in the activity of transmembrane transporters and/or enzymes that transform metabolites in the liver and other tissues. These different 'metatypes' may reflect different genotypes, although differences in gut microbiota also could contribute to inter-individual differences for some metabolites. Such differences in metatype, particularly if prevalent in the population, may prevent validation of some newly discovered markers for food intake and confound attempts to link dietary intake (based on metabolite biomarkers) to health outcomes. Conversely, once identified, stratification of subjects on the basis of their metatypes may explain differences in the strength of associations between intake of a particular nutrient or food and health/disease outcomes. Much useful information on the genetic and biochemical basis of metatypes is available from pharmacological and toxicological research because similar biochemical pathways are involved in the transport, metabolism and excretion of xenobiotics, and of non-nutrient food components. Research is needed to develop rapid methods for determining individual metatypes that can be applied in a nutritional context.

Recommendations for future research

A number of research recommendations were suggested that would advance progress towards the overall goal of using metabolomics to assess dietary intake. These recommendations included the following:

- (1) Determine whether particular biomarkers (individual compounds or groups of compounds) can be used as quantitative or qualitative indicators of consumption of particular classes of nutrients (micro and macro) or foods (e.g. fruits and vegetables) or of adoption of particular dietary patterns (e.g. Mediterranean diet).
- (2) Develop a database of food-derived metabolites (food metabolome) to facilitate the rapid identification of biomarkers from unannotated metabolite signals in urine or plasma fingerprints.
- (3) Build a toolbox of experimental and informatics methods to facilitate biomarker discovery and validation.
- (4) Once suitable biomarkers have been identified, develop targeted methods of analysis using more conventional analytical procedures (e.g. enzyme-linked immunoassays) to allow their high-throughput quantification.
- (5) Develop and validate robust metabolomics methodologies that can be applied to larger numbers of urine and plasma samples to explore the direct application of metabolomics to population cohort studies and dietary surveys.
- (6) Determine the influence of metabolic phenotype (differences in metabolotype) on the relationships between dietary exposure and metabolite concentrations/patterns in the blood and urine. This recommendation will advance understanding of the mechanisms responsible for inter-individual variability in food-related biomarkers. Outcomes from such research could also provide key information, which could be used to strengthen conclusions of causality in studies of associations between dietary intake and health/disease outcomes.
- (7) Develop metabolomics approaches to identify undeclared confounding behaviour (e.g. intake of alcohol and prescription drugs) in studies of links between diet and health.

Acknowledgements

S. P. was contracted by the Food Standards Agency to be rapporteur of the workshop. He therefore produced the first draft. All other authors contributed substantially in terms of drafting, critical reviewing and editing subsequent versions of the manuscript. The authors declare no conflict of interest. The participants in the workshop were as follows: Dr Augustin Scalbert, Professor Sandy Primrose, Professor John Draper, Dr Amanda Lloyd, Dr Manfred

Beckmann, Professor John Mathers, Professor Chris Seal, Dr Sumanto Haldar, Professor John Beattie, Dr Shelley Rhodes, Dr John Lodge, Professor Gary Frost, Dr Mazda Jenab, Dr Hector Keun, Dr Clair Baynton, Dr Alison Tedstone, Dr Elaine Stone, Alisdair Wotherspoon, Rachel Elsom, Verity Kirkpatrick, Vicki Pyne, Dr Lorraine Brennan, Dr Ian Colquhoun, Dr Janet Flatley, Dr Simon Kerley, Dr Albert Koulman, Dr Julie Lovegrove, Dr Gordon McDougall, Professor David Thurnham and Professor Ian Young.

References

1. Goris AHC, Meijer EP & Westerterp KR (2001) Repeated measurement of habitual food intake increases under-reporting and induces selective under-reporting. *Br J Nutr* **85**, 629–634.
2. Livingstone MBE, Prentice AM, Strain JJ, *et al.* (1990) Accuracy of weighed dietary records in studies of diet and health. *Br Med J* **300**, 708–712.
3. Fardet A, Llorach R, Orsoni A, *et al.* (2008) Metabolomics provide new insights on the metabolism of dietary phytochemicals in rats. *J Nutr* **138**, 1282–1287.
4. Wishart DS (2008) Metabolomics: applications to food science and nutrition research. *Trends Food Sci Technol* **19**, 482–493.
5. Mennen L, Saphino D, Ito H, *et al.* (2006) Urinary flavonoids and phenolic acids as biomarkers of intake for polyphenol-rich foods. *Br J Nutr* **96**, 191–198.
6. Jenab M, Slimani N, Bictash M, *et al.* (2009) Biomarkers in nutritional epidemiology: applications, needs and new horizons. *Hum Genet* **125**, 507–525.
7. Scalbert A, Brennan L, Fiehn O, *et al.* (2009) Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics* **5**, 435–458.
8. Manach C, Hubert J, Llorach R, *et al.* (2009) The complex links between dietary phytochemicals and human health deciphered by metabolomics. *Mol Nutr Food Res* **53**, 1303–1315.
9. Favé G, Beckmann M, Draper JH, *et al.* (2009) Measurement of dietary exposure: a challenging problem which may be overcome thanks to metabolomics? *Genes Nutr* **4**, 135–141.
10. Holmes E, Loo RL, Stampler J, *et al.* (2008) Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* **453**, 396–400.
11. Riboli E, Hunt KJ, Slimani N, *et al.* (2002) European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* **5**, 1113–1124.
12. Riboli E (1992) Nutrition and cancer: background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann Oncol* **3**, 783–791.
13. Saadatian-Elahi M, Slimani N, Chajès V, *et al.* (2009) Plasma phospholipid fatty acid profiles and their association with food intakes: results from a cross-sectional study within the European Prospective Investigation into Cancer and Nutrition. *Am J Clin Nutr* **89**, 331–346.
14. Al-Delaimy WK, Ferrari P, Slimani N, *et al.* (2005) Plasma carotenoids as biomarkers of intake of fruits and vegetables: individual-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Eur J Clin Nutr* **59**, 1387–1396.

15. Al-Delaimy WK, Slimani N, Ferrari P, *et al.* (2005) Plasma carotenoids as biomarkers of intake of fruits and vegetables: ecological-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Eur J Clin Nutr* **59**, 1397–1408.
16. Beckmann M, Parker D, Enot DP, *et al.* (2008) High throughput, non-targeted metabolite fingerprinting using nominal mass flow-injection electrospray mass spectrometry. *Nature Protocols* **3**, 486–504.
17. Overy DP, Enot DP, Tailliat K, *et al.* (2008) Explanatory signal interpretation and metabolite identification strategies for nominal mass FIE-MS metabolite fingerprints. *Nature Protocols* **3**, 471–485.
18. Enot DP, Lin W, Beckmann M, *et al.* (2008) Preprocessing, classification modelling and feature selection using flow injection electrospray mass spectrometry metabolite fingerprint data. *Nature Protocols* **3**, 446–470.
19. Ritchie SA, Ahiaonu PWK, Jayasinghe D, *et al.* (2010) Reduced levels of hydroxylated, polyunsaturated ultra long-chain fatty acids in the serum of colorectal cancer patients: implications for early screening and detection. *BMC Med* **8**, 13.