Metabolomics in nutrition research: current status and perspectives

Lorraine Brennan
UCD Institute of Food and Health, Belfield, University College Dublin, Dublin 4, Ireland

Abstract
Metabolomics is the study of metabolites present in biological samples such as biofluids, tissue/cellular extracts and culture media. Combining metabolomic data with multivariate data analysis tools allows us to study alterations in metabolic pathways following different perturbations. Examples of perturbations can be disease state, drug or nutritional interventions with successful applications in the fields of drug toxicology, biomarker development and nutrition research. Application of metabolomics to nutrition research is increasing and applications range from assessing novel biomarkers of dietary intake to application of metabolomics in intervention studies. The present review highlights the use of metabolomics in nutrition research.

Introduction
Metabolomics is an ‘-omics’ technology that has been applied to an ever-increasing number of fields over the last 5–10 years. Metabolomics involves the study of small molecules or metabolites present in biological samples with the aim to study alterations in metabolism under different conditions. By applying metabolomics, one aims to profile all of the metabolites present in the samples: this complement of metabolites is referred to as the metaboome. The study of metabolites reveals useful biological information as the metabolites represent biological end points and are now implicated in the development of a number of human diseases [1–3].

Comprehensive reviews on the experimental strategies in metabolomics can be found elsewhere [4–7]. In general terms, the following steps are involved in metabolomics study: (i) sample preparation, (ii) data acquisition, (iii) statistical analysis, and (iv) pathway mapping (see Figure 1). The present review provides an overview of current technologies, data analysis techniques and focus on applications in nutrition research. The commonly used technologies in metabolomics studies include NMR- and MS-based techniques. MS-based techniques are coupled with chromatography such as GC or LC (liquid chromatography). Each of these techniques has its own advantages and disadvantages and a detailed comparison of the techniques is beyond the scope of the present review. NMR-based techniques have high reproducibility and little inter-laboratory variation, but suffer from lower sensitivity. Even though it requires higher sample volume, the technique is not destructive and the sample can be re-used. GC–tandem MS is best suited to volatile metabolites and to increase the volatility the samples are often derivitised. Examples of molecules routinely analysed using this approach include fatty acids, organic acids and sugars. LC-based techniques are sensitive and suited to soluble and lipophilic molecules.

At present, no one technology will measure the whole metaboome. As a consequence, different technologies have been developed and a combination of these usually yields optimal coverage of the metaboome. To add a layer of complication, most of the analytical platforms can be run and analysed in a targeted or non-targeted fashion. Depending on the research question/design, a strategy should be developed and tailored to the question. In hypothesis-free studies such as disease biomarker identification, non-targeted metabolomics is usually performed and data can be acquired using NMR-, GC-MS- and LC-MS-based techniques. In this mode, the analysis provides a holistic overview of the metabolic changes occurring under different conditions such as disease compared with control. In general, the data obtained are semi-quantitative as absolute concentration calculation requires the use of multiple internal standards. When hypotheses relating to specific metabolite classes are to be tested, targeted metabolomics is the method of choice. In general, this specific analysis delivers quantitative data that have been determined by use of reference standards. Techniques commonly used in this type of data analysis are NMR, GC–MS, LC–tandem MS and flow injection–tandem MS. Independent of whether acquisition is for targeted or untargeted analysis, optimal coverage of the metaboome is achieved through the use of multiple platforms.

Metabolomics is amenable to a number of different samples types and the research question drives the optimal selection of samples. In human studies, metabolomics analysis can be performed on biofluids such as saliva, urine, blood and biopsy samples. In the case of biopsy samples, the metabolites have to be extracted first and the extraction method depends on the metabolite class to be analysed. For cellular studies, both

Key words: metabolic pathway, metabolism, metabolomics, nutrition
Abbreviations used: LC, liquid chromatography; PCA, principal component analysis; PLS-DA, partial least-square discriminant analysis; O-PLS-DA, orthogonal PLS-DA; PPCA, PCA with covariates.
1email lorraine.brennan@ucd.ie
intracellular and extracellular metabolites can be profiled to yield information for certain biological processes.

**Data analysis**
A characteristic of metabolomics is the large amount of data generated and an important part of any metabolomics study is the analysis of these data using multivariate statistics. At present, there are many statistical methods available with some of the most commonly used methods including PCA (principal component analysis), PLS-DA (partial least-square discriminant analysis) and O-PLS-DA (orthogonal PLS-DA). PCA is an unsupervised technique that allows visualization of the data with the aim to identify inherent grouping of samples as a result of the similarity of the metabolic composition. Supervised techniques such as PLS-DA and O-PLS-DA require *a priori* knowledge of the class membership and are used to identify metabolites that are differing between groups. A key point in using these supervised techniques is the necessity to validate any models developed. Methods used to validate include cross-validation where samples are left out and their class membership is predicted. Another approach is permutation testing where randomly assigned class labels are used and the predictability of the working model is compared with these data. In recent years, other tools are emerging as being useful in metabolomics analysis [8,9]. An example of one of these tools is a method for performing PPCCA (PCA with covariates) [10]. This is extremely important when studying human metabolism as a number of phenotypic factors are known to impact on the metabolic profile. Controlling for these in analysis is extremely important to reduce the impact of false positives.

**Applications of metabolomics in nutrition research**
The applications of metabolomics have expanded across biological disciplines in recent years; the present review focuses on applications in nutrition research. In general, the applications in the nutrition field can be divided into two categories: (i) dietary intervention studies, and (ii) biomarker discovery studies. Applications of metabolomics to dietary intervention studies enhance the understanding of the effects of certain diets or food items on metabolic pathways. Some examples include the effects of tea [11,12], chocolate and cocoa [13,14], vitamins [15] and fish oils [16] on the metabolic profiles of urine and plasma. To illustrate the
use of metabolomics in such an application, one example is discussed. Application of a metabolomics approach to biofluids and tissues from animals following a 12-week high-fat diet revealed a number of metabolic perturbations [17]. Of note was the significant alteration in the one-carbon metabolic pathway in the liver. The hepatic oxidative capacity and the lipid milieu were significantly altered, which resulted in the proposal that this may play a key role in the development of insulin resistance. Additionally, high choline levels were observed following the high-fat diet and, overall, the results suggested that the coupling of high levels of choline and low levels of methionine play an important role in the development of insulin resistance and liver steatosis.

To enhance our understanding of the relationship between dietary habits and chronic disease, it is imperative that reliable and accurate dietary assessments are used. Current methods for assessing dietary intake include 24-h recalls, weighed food diaries and food frequency questionnaires. The accuracy of food intake data obtained from these methods can be influenced by random and systematic errors including energy under-reporting [18]. To improve dietary assessment methods, novel biomarkers of dietary intake are needed. In recent years, a number of studies have emerged demonstrating the utility of metabolomics for the identification of biomarkers of dietary intake. The design of these studies is usually consumption of foods followed by collection of biofluids over a period of time. Analysis of these postprandial samples in an untargeted fashion can lead to the identification of specific biomarkers. Validation of the biomarkers in an independent study is desirable following the discovery stage. A large body of data exists in the literature with respect to the metabolite levels in various foods and careful examination of this literature can drive a targeted approach for biomarker discovery. To this end, there are multiple international efforts to develop online databases with information on food metabolites and their kinetics. Examples of such initiatives include FooDB (http://www.foodb.ca/) and the PolyPhenol Explorer (http://www.phenol-explorer.eu/) [19]. In the present literature, successful identification of biomarkers of dietary intake from metabolomics studies exist for numerous food including red meat, citrus fruits, fish, whole grains and vegetables [20–24]. To demonstrate application in this field, the study leading to the identification of biomarkers for citrus fruit is highlighted. Heinzmann et al. [24] performed an acute discovery study where proline betaine was identified as a biomarker of citrus fruit intake. Urine was collected multiple times following consumption of the citrus fruit. 1H-NMR and PLS-DA identified the urinary excretion of proline betaine as a biomarker of citrus fruit intake. Kinetic studies revealed the excretion profile following orange juice consumption. Validation of this marker was performed using data from participants of the INTERMAP U.K. cohort [25]. A ROC (receiver operating characteristic) curve resulted with an AUC (area under the curve) of 0.92 with a sensitivity and specificity of 90.6 and 86.3% respectively. Independently of this study, another group identified proline betaine as a marker for citrus fruit intake following analysis of samples taken using an acute breakfast challenge [26]. Further analysis of the data demonstrated sensitivities of 80.8–92.2% and specificities of 74.2–94.1% for elevated levels of proline betaine in those volunteers who reported a high consumption.

An alternative approach has emerged in recent years where patterns of intake are related to metabolic profiles. Work from our research group demonstrated that application of dietary pattern analysis to food intake data recorded from 125 subjects using 3-day food diaries revealed three dietary patterns [21]. Interrogation of these patterns revealed that they were reflected in the urinary metabolic profiles leading to the concept that subjects could be grouped into different classes that were reflective of their dietary intake. A similar approach was used by Altmaier et al. [27], where seven dietary patterns were identified. Assessment of these in relation to the metabolomics profiles revealed that the dietary patterns were reflected in the plasma samples. Following on from these findings and other work [21], the concept of ‘nutritype’ has emerged where subjects can be classed into a certain dietary pattern on the basis of their metabolomics profile. The true value of such an approach will lie in the integration of metabolomics data with classical dietary data. Development of strategies to achieve this will be a focus in the future and will require significant interaction between nutritional epidemiologists and metabolomics.

Concluding remarks

The use of metabolomics in nutrition-related research has made a significant impact. However, some challenges remain in its application and, for the field to reach its full potential, these need to be addressed. Challenges include the translation of altered levels of metabolites into altered metabolic pathways that aid biological interpretation. Translation of metabolic changes in urine following a dietary intervention into organ-specific biologically meaningful information is not trivial. Similarly for plasma, translation of findings into specific alterations in certain pathways at an organ level is difficult. To advance this area, the application of flux studies offers immense possibilities. Application of such studies should help to decipher the meaning of metabolite changes and their relationship with flux between the organs. Furthermore, coverage of the metabolome needs to improve. Notwithstanding these challenges, metabolomics has enabled significant progress in the field of nutrition and will continue to do so in the future.

Funding

L.B. was supported by funding for Food for Health Research initiative [NDP 2007–2013, 07FHRUCD01] by the Department of Agriculture, Fisheries and Food, the Health Research Board and the Department of Health and Children.
References


