Metabolomics as a tool in the identification of dietary biomarkers

Helena Gibbons[1, 2] and Lorraine Brennan[1, 2*]

[1] Institute of Food & Health, UCD School of Agriculture and Food Science, University College Dublin (UCD), Belfield, Dublin, Ireland [2]UCD Conway Institute of Biomolecular Research, UCD, Belfield, Dublin, Ireland

[*]Address correspondence to Prof. Lorraine Brennan, UCD School of Agriculture and Food Science, Belfield, University College Dublin, Dublin 4, Ireland. E-mail: lorraine.brennan@ucd.ie

Metabolomics in dietary biomarker discovery

Metabolomics: Dietary biomarkers: Diet and nutrition: Dietary assessment

Abstract

Current dietary assessment methods including food-frequency questionnaires (FFQs), 24-hour recalls and weighed food diaries are associated with many measurement errors. In an attempt to overcome some of these errors, dietary biomarkers have emerged as a complimentary approach to these traditional methods. Metabolomics has developed as a key technology for the identification of new dietary biomarkers and to date, metabolomics based approaches have led to the identification of a number of putative biomarkers. The three approaches generally employed when using metabolomics in dietary biomarker discovery are; i) acute interventions where participants consume specific amounts of a test food, ii) cohort studies where metabolic profiles are compared between consumers and non-consumers of a specific food and iii) the analysis of dietary patterns and metabolic profiles to identify nutritypes and biomarkers. The present review critiques the current literature in terms of the approaches used for dietary biomarker discovery and gives a detailed overview of the currently proposed biomarkers, highlighting steps needed for their full validation. Furthermore, this review also evaluates areas such as current databases and software tools which are needed to advance the interpretation of results and therefore enhance the utility of dietary biomarkers in nutrition research.

Dietary biomarkers and the concept of metabolomics

The contribution of diet to the increasing burdens of cardiovascular disease (CVD), diabetes, obesity and cancers has been recognised since the 1970s [1]. Selected foods and nutrients as well as dietary patterns are now known to interact with various metabolic processes contributing to a reduction or an increase in the risk of disease [2]. For example, it is well established that high salt consumption raises blood pressure [3] and high consumption of red meat has been associated with increased incidence of type 2 diabetes [4; 5], CVD [6] and cancers [7]. In contrast dietary patterns such as the Dietary Approaches to Stop Hypertension (DASH) diet, which emphasises consumption of fruit and vegetables, low-fat dairy foods and whole grains and reduced intake of red meats and sugars has been shown to decrease blood pressure and CVD risk [8; 9]. Similarly, the Mediterranean diet which emphasises high fruit, vegetable and olive oil consumption has been shown to reduce CVD and type 2 diabetes risk [10; 11]. As diet is a key environmental risk factor, the identification and targeting of dietary factors with the greatest prospective for reducing or increasing disease risk is of major scientific and public health importance [12]. It is therefore essential that dietary assessment methods are reliable and accurate for the advancement of our understanding of the links between diet and health.

Diet is traditionally measured via self-reporting methods such as food-frequency questionnaires (FFQs), 24-hour recalls and weighed food diaries. There is however a number of methodological issues associated with each of these assessment methods, including energy under-reporting, recall errors and difficulty in assessment of portion sizes [2; 13; 14]. Such errors can lead to reduced power, underestimated associations and false findings which may contribute to inconsistencies in the field of nutritional epidemiology [14; 15]. In an effort to address some of these measurement issues, the use of dietary biomarkers, which are found in biological samples and are related to ingestion of a specific food or food group, have emerged [16]. Currently dietary biomarkers exist for salt, protein, sucrose/fructose intake (sodium/nitrogen/sucrose and fructose measured in 24 h urine samples) and energy expenditure (the doubly labelled water technique) [2; 17]. These dietary biomarkers can be used in conjunction with traditional dietary assessment methods to improve the accuracy of dietary intake measurement and can also be used to more accurately associate dietary intake with disease risk and nutritional status [18].

In recent years, metabolomics has developed as a key technology for the identification of new dietary biomarkers. Metabolomics provides a powerful approach for the comprehensive description of all low molecular weight molecules present in biological samples [16]. In metabolomics research the analytical platforms predominantly used are nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) coupled with a chromatographic step, for

example, gas chromatography (GC) or liquid chromatography (LC). Each of these techniques are associated with a number of advantages and disadvantages, for example MS-based techniques have high sensitivity and therefore may detect metabolites below the detection limit of NMR spectroscopy, however sample treatment is necessary before MS-based analysis, while little or no pre-treatment is required for NMR [19].While in the past many articles detailed the advantages and disadvantages of different approaches there has now been a realisation that using one platform alone will not give complete coverage of the metabolite profile; therefore, a combination of technologies and approaches is usually recommended for optimal coverage. Analysis of metabolomic data is commonly performed using multivariate statistics and there are an increasing selection of databases and tools available to assist in the interpretation of these multivariate results [20].

Examination of the literature reveals that there are three approaches generally employed for dietary biomarker discovery. These can be summarised as : i) acute or medium interventions where participants consume specific amounts of a test food and biological samples are collected post consumption, ii) cohort studies where metabolic profiles are compared between consumers and non-consumers of a specific food and iii) the analysis of dietary patterns and metabolic profiles to identify nutritypes and biomarkers. Although these study designs have led to the identification of a number of biomarkers in the literature in recent years, each of these approaches have a number of limitations associated with them. Awareness of these is important in the interpretation and potential use of such biomarkers. Therefore the objective of this review is to give an overview of currently proposed biomarkers and secondly this review aims to critique the current literature in terms of approaches for dietary biomarker discovery, highlighting steps needed for their full validation.

Dietary biomarker discovery using intervention studies

Dietary intervention studies involve participants consuming specific amounts of a test food in a single meal (acute intervention) or for a short to medium term intervention the test food is consumed in repeated meals. In this approach baseline and postprandial biofluids are collected and following analysis, potential biomarkers are identified. This approach has led to the identification of a number of putative biomarkers of specific foods and beverages as summarised in Table 1. An excellent example of a biomarker successfully identified using this approach is proline betaine, a robust biomarker of citrus fruit intake. Proline betaine was originally identified by Atkinson et al. [21] and following this Heinzmann and colleagues performed an acute intervention study with a mixed-fruit meal, which consisted of apples, grapes, oranges, and grapefruit [22]. Eight participants consumed standardised meals over three days and on the second day the mixed-fruit meal was consumed [22]. Urine samples were collected and analysed using NMR spectroscopy. Following multivariate analysis proline betaine was identified as a potential biomarker. To assign the origin of

urinary proline betaine excretion after the mixed-fruit meal, concentrations of proline betaine in fruits and fruit juices were measured. Concentrations of proline betaine were higher in citrus fruit compared with other commonly available fruit and fruit juices tested. The urinary excretion profile of proline betaine was then measured in 6 individuals after consumption of orange juice. This biomarker was confirmed using data from participants in the INTERMAP U.K. cohort and demonstrated a high sensitivity and specificity for citrus fruit consumption (90.6 and 86.3% respectively) [22]. Lloyd and colleagues also identified proline betaine and a number of biotransformed products in postprandial urine samples after consumption of 200ml of orange juice as part of a standardised test breakfast [23]. Subsequent biomarker validation demonstrated sensitivities and specificities of 80.8–92.2% and 74.2–94.1% respectively, for elevated proline betaine in high reporters of citrus fruit consumption [23]. Following on from these acute studies, a medium term intervention study used MS to profile the urinary metabolomes of 12 volunteers that consumed orange juice regularly for one month as part of their habitual diet. Proline betaine was again identified as a potential marker of citrus fruit [24]. Considering the range of studies that consistently report proline betaine as a marker of citrus fruit intake the evidence base is strong to support its use.

A number of research groups have also used dietary interventions to investigate biomarkers of cruciferous vegetables [25; 26; 27]. Andersen and colleagues performed a controlled cross-over meal study with nine brassica-containing New Nordic Diet (NND) meals in 17 subjects [26]. 24 h urine samples were collected and analysed by ultra-performance liquid chromatography quadruple time-of-flight MS (UPLC-qTOF-MS). To investigate the food sources of the biomarkers found in the meal study, a range of small single food studies were performed with 3–4 participants in each. Using a sensitivity and specificity analyses to select the most promising biomarkers, a range of conjugated isothiocyanates were identified as PEMs of brassica intake [26]. Further PEMs of other foods, including fish were also identified [26]. To validate the biomarkers from this study, Andersen et al. carried out a 6-month parallel dietary intervention study where 107 participants were randomised into two distinct dietary patterns [27]. Combining LC-MS data from 24 h urine samples and data from 3-day weighed dietary data this study again identified conjugates of isothiocyanates as brassica biomarkers. However, using this approach it was only possible to verify 23% of potential biomarkers observed in the previous-meal studies [27]. As this was a less controlled intervention that included a wider selection of foods with varied amounts of intake and different preparation methods, it highlights the need for the validation of biomarkers in different subjects and study settings [27].

A number of red meat and fish biomarkers have been identified using this intervention approach [7; 28; 29]. Most recently, metabolomics has been applied to compare the different effects of

meat and fish on the plasma metabolome [30]. Ross et al. carried out an intervention study analysing the differences in the postprandial plasma metabolic response to meals containing baked beef, baked herring and pickled herring [30]. 17 males consumed three test meals in a crossover design with one week washout between the meals. Postprandial blood plasma samples were taken over seven hours and analysed by GC-MS. Concentrations of 2-aminoadipic acid, β-alanine and 4-hydroxyproline were significantly higher following the beef meal compared to the baked herring meal. Herring intake led to a greater plasma postprandial response from docosahexaenoic acid (DHA) and cetoleic acid compared with beef [30]. However, further studies are needed to confirm these dietary biomarkers and decipher their specificity.

Dietary biomarker discovery using cohort studies

Searching for new dietary biomarkers in cohort studies requires the use of self-reported dietary data to identify low and high consumers of a specific food. Following this, the metabolomic profiles are compared between low and high consumers and potential biomarkers are identified. Putative biomarkers of foods, identified using this approach, are presented in Table 2. Work in our lab combined this approach with an acute intervention to identify and confirm a panel of biomarkers indicative of sugar sweetened beverage (SSB) intake [31]. Heat map analysis was performed to identify correlations between NMR spectral regions and SSB intakes in the cohort study. A panel of 4 biomarkers; formate, citrulline, taurine and isocitrate were identified as markers of SSB intake. Following the acute consumption of the SSB all 4 metabolites were shown to increase in the urine and the panel of biomarkers were successfully identified in the SSB [31]. Another study using this cohort study approach, analysed the correlations between serum profiles and dietary data collected using FFQs in participants from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial [32]. The application of untargeted metabolomics to this epidemiologic data set detected 39 metabolites of known identity that were correlated with a total of 13 dietary groups, for example citrus intake was associated with stachydrine, chiro-inositol, scyllo-inositol and N-methyl proline, fish with 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid, DHA and EPA, peanut intake with tryptophan betaine and 4-vinylphenol sulfate and coffee intake was associated with trigonelline-N-methylnicotinate and quinate [32]. To complicate interpretation further, the intake of foods is highly correlated making identification of specific biomarkers difficult and this highlights the need for the validation of biomarkers. The majority of biomarkers identified using cohort studies have been predominantly identified in urine, this study demonstrates the potential use of serum samples in dietary biomarker discovery. However, the proposed biomarkers identified are only based on associations and some biomarkers were not food specific, for example DHA was

correlated with fish and rice intake. Further validation in intervention studies is therefore necessary to demonstrate responsiveness to intake.

Wittenbecher and colleagues also demonstrated the use of serum samples when identifying biomarkers of red meat intake in a subset of participants from the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam cohort (n=2047) [33]. Total red meat consumption was assessed using FFQs and serum samples were analysed using a targeted metabolomics approach. Ferritin, glycine, 4 diacyl phosphatidylcholines, 11 acylalkyl phosphatidylcholines, 2 lysophosphatidylcholines and 2 sphingomyelins were associated with total red meat consumption and 6 of these biomarkers were also found to be associated with type 2 diabetes risk [33]. This is the first study evaluating a large set of metabolites as potential mediators of the association between red meat intake and diabetes risk, however, dietary information relied on estimates of habitual consumption over the past year by FFQs and metabolites were measured at a single time point. Furthermore, total red meat was defined as processed and unprocessed meat and therefore did not identify biomarkers of specific types of meat. Additional study is essential to validate the biomarkers identified and to further dissect such relationships with disease risk.

Biomarkers of bread intake have also been investigated in 155 subjects from the PERIMED study [34]. A 137-item FFQ was used to stratify subjects into three groups: non-consumers of bread (n = 56), white-bread consumers (n = 48) and whole-grain bread consumers (n = 51). Fasting urine samples, analysed by untargeted high-performance liquid chromatography quadruple time-of-flight MS (HPLC–qTOF-MS), identified higher concentrations of compounds including benzoxazinoids and alkylresorcinol metabolites and compounds produced by gut microbiota (enterolactones, hydroxybenzoic and dihydroferulic acid metabolites) in bread consumers. 2, 8-dihydroxyquinoline glucuronide was also found to be more abundant in whole-grain bread consumers [34]. The biomarkers identified are based on a FFQ; therefore further validation is essential to demonstrate a direct relationship with bread consumption.

Dietary biomarker discovery using dietary patterns

The third approach; analysing dietary patterns and metabolomic profiles to identify nutritypes (ie, metabolic profiles that reflect dietary intake) and biomarkers have been demonstrated by a number of research groups (see Table 3). One of the first examples emerged from our laboratory when a k-means cluster analysis was performed on self-reporting dietary data and 3 distinct dietary patterns, which were associated with unique food intakes were identified [35]. Dietary clusters were reflected in the urinary metabolomic profiles of the 125 participants and a number of metabolites were identified and linked to the intake of specific food groups [35]. These nutritypes have the potential to aid dietary assessment by unobjectively classifying people into certain dietary patterns. Further

work within our research group, applying the concept of using biomarkers to reflect dietary patterns, has focused on lipidomics, a subfield of metabolomics that concentrates on the global study of lipids [36]. Dietary data, measured by FFQs and lipid profiles measured from serum samples, in 34 Metabolic Challenge Study (MECHE) participants were used for this analysis. PCA reduced lipid profiles into lipid patterns and these were regressed against dietary data to identify biomarkers related to the intake of certain foods and nutrients. 6 lipid patterns were identified including lipid pattern 1 which was found to be highly predictive of dietary fat intake (AUC of 0.82), lipid pattern 4 which was highly predictive of alcohol intake (AUC=0.81) and lipid pattern 6 which had a reasonably good ability to predict dietary fish intake (AUC=0.76). Lysophosphatidylcholine alkyl C18:0 (LPCeC18:0) was identified as a potential biomarker of alcohol consumption and lysophosphatidylethanolamine acyl C18:2 (LPEaC18:2) and phoshatidylethanolamine diacyl C38:4 (PEaaC38:4) were identified as potential biomarkers of fish intake [36]. This approach demonstrates the utility of serum in the identification of key dietary factors that influence the lipidomic profile. However, again validation of the biomarkers through use of intervention studies is needed.

Most recently, Andersen and colleagues used an untargeted metabolomics approach to distinguish between two dietary patterns with the purpose of developing a compliance measure [37]. In a parallel intervention study 181 participants were randomly assigned to follow a New Nordic Diet (NND) or an Average Danish Diet (ADD). 24 hour urine samples were collected, analysed by UPLC-qTOF-MS and PLS-DA was applied to develop a compliance model for ADD and NND based on the most discriminative features detected in urine. This resulted in a robust model with a misclassification rate of 19% [37]. Metabolites characterising the ADD diet and the NND diet are listed in Table 3. This study demonstrates the potential of metabolomics in discovering biomarkers indicative of dietary patterns but furthermore it highlights a promising approach that may be used to develop compliance measures that cover the most important discriminant metabolites of complex diets.

Limitations of current approaches/study designs

In general, metabolomics based approaches have produced reasonably robust models for dietary biomarker identification. However, following the discovery of any biomarker, validation in an independent study is critical to enable the generalisability of the results. This validation step is essential because factors which may not be present in traditional dietary assessment methods including genetic factors, lifestyle and physiological factors, dietary factors, the biological sample or the analytic methodology could skew biomarker measures of dietary intake [38]. For many of the

study designs discussed, validation of the biomarker is often absent, making it difficult for the translation of these biomarkers into practice.

It has been proposed that the confirmation of dietary biomarkers should occur in two stages, firstly the dose–response effect should be included in intervention studies and secondly the suitability of the candidate biomarker in a free-living population should be investigated using a (controlled) habitual diet [39]. Evaluation of the dose–response relationship is critical as it allows for the assessment of the suitability of the biomarker over a range of intakes [20]. Unfortunately, in many studies, this important step is often absent. Biomarkers identified using samples from cohort studies do not assess the direct relationships of food amounts consumed and levels of biomarkers and do not demonstrate responsiveness to intakes, therefore the relationship is only an association [16]. Such studies should ideally be combined with intervention studies to demonstrate direct relationships and dose-response relationships. Conversely, dietary biomarkers identified within acute intervention studies advantageously allow for the examination of dose-response relationships, however, to date few studies have incorporated such designs.

When using self-reporting dietary data from cohort studies in the biomarker discovery process, one should be aware of reporting errors and the potential for missing important correlations and attenuation of results. May and colleagues investigated the metabolomic profiles of participants consuming a high-phytochemical diet compared with a diet without fruits and vegetables in a randomised controlled trial and also investigated the metabolomic profiles of participants in a cross-sectional study, where high and low fruit and vegetable diets were identified based on 3-day food records and FFQs. The intervention study found forty-six putatively annotated ions, with MS/MS fragment ion support that were differentially abundant between the two intervention diets, however within the cross-sectional study only one compound annotated with MS/MS support was identified using the 3-day food records and there were no metabolites that significantly separated groups based on FFQ data [40]. This therefore demonstrates the drawbacks of using self-reported data in dietary biomarker discovery. Furthermore, when using cohort studies to identify or confirm biomarkers it is imperative that it is acknowledged that many of the foods consumed are highly correlated and therefore biomarkers identified may not be specific to the particular food of interest [20]. Following identification of putative biomarkers from cohort studies we recommend that the relationship is confirmed using an intervention study in a dose-response manner where the sensitivity and specificity of the biomarkers can also be assessed. The importance of such a step is key to the validation of the biomarkers and important to support their use.

Use of acute and medium term interventions is not without limitations in terms of dietary biomarker identification: many of the biomarkers identified using this approach are markers of acute intake. For example proline betaine is excreted rapidly in urine and excretion is almost

complete ≤24 h [22]. These acute biomarkers may therefore only be valid for people that regularly and frequently consume the particular foods. The identification of dietary biomarkers that reflect habitual intake requires longer-term studies. Furthermore, it must also be noted that the majority of the acute and medium term intervention study designs involve only a small number of participants [22; 24; 41]. The proposed dietary biomarkers identified using these approaches therefore cannot always be extrapolated to population studies in free-living individuals. However, this can be in part be dealt with by confirmation in cohort studies with a diverse range of characteristics.

While the above describes limitations in study designs, there is also the need for development of databases and software tools to advance the interpretation of metabolomics results and therefore enhance the utility of dietary biomarkers in nutrition research. Current databases such as the Human Metabolome Database (HMDB) provides access to an online database containing detailed information about small molecule metabolites (>40,000) found in the human body [42]. Since it was first described in 2007, it is constantly being expanded and updated and has become a valuable resource that contains spectroscopic, quantitative, analytic and physiological information about human metabolites [42]. The Food Metabolome Database (FooDB), is another database of >28,000 food constituents that contains information about food sources and food concentrations [43]. This resource provides an aid for the identification of new metabolites that are reflective of food intake. While this resource is valuable, the identification of metabolites originating from food remains difficult and there is a need for sharing of databases to aid identification. Most recently, a comprehensive and electronically accessible human urine metabolome database, which includes quantitative concentrations of metabolites in urine samples was established [44]. This database also represents a significant development and resource for biomarker identification and quantification. Other new software tools include BAYESIL, this system provides fully automated and fully quantitative NMR-based metabolomics of complex mixtures [45]. This will have a significant impact on NMR spectroscopy and NMR-based metabolomics.

Conclusion

The use of dietary biomarkers in nutrition research holds great promise. However, prior to having a suite of reliable dietary biomarkers that could be used in nutrition research a number of validation steps need to considered. Furthermore, the challenges identified in this review need to acknowledged and addressed. Appropriate validation steps are essential, otherwise the robustness of biomarkers will remain uncertain and the translation of these biomarkers into practice will be challenging. Longer-term studies are also needed for the identification of dietary biomarkers

reflective of habitual dietary intake. Until well validated biomarkers are identified it is unlikely we will see uptake by the research community of the emerging biomarkers. The challenge for the researchers working in this field, in the coming years, will be to develop a suite of well validated biomarkers. To this end the JPI funded programme FoodBall will address some of these issues and pave the way forward (http://foodmetabolome.org/). They may also have the potential for the assessment of compliance to dietary interventions in both a clinical and a research setting. Ultimately these dietary biomarkers will be used to further elucidate the proposed links between certain foods and disease.

Financial Support:

The authors would like to acknowledge the following funding: FP7 Project NutriTech (289511), SFI (14/JP-HDHL/B3075) and ERC (647783).

Conflict of interest:

None

Authorship

H.G. drafted the outline of the manuscript, conducted the literature search and drafted the manuscript. L. B. was responsible for critically reviewing the manuscript. Both authors read and approved the final manuscript before submission.

1. Mozaffarian D, Ludwig DS (2010) Dietary guidelines in the 21st century—a time for food. *JAMA* **304**, 681-682.
2. Bingham SA (2002) Biomarkers in nutritional epidemiology. *Public Health Nutr* **5**, 821-827.
3. Kotchen TA, Cowley AW, Frohlich ED (2013) Salt in Health and Disease — A Delicate Balance. *N Engl J Med* **368**, 1229-1237.
4. Feskens EJ, Sluik D Fau - van Woudenbergh GJ, van Woudenbergh GJ (2013) Meat consumption, diabetes, and its complications. *Curr Diab Rep* **13**, 298-306.
5. Pan A, Sun Q, Bernstein AM *et al.* (2011) Red meat consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis. *Am J Clin Nutr* **94**, 1088-1096.
6. Hu FB, Rimm EB, Stampfer MJ *et al.* (2000) Prospective study of major dietary patterns and risk of coronary heart disease in men. *Am J Clin Nutr* **72**, 912-921.
7. Cross AJ, Leitzmann MF, Gail MH *et al.* (2007) A Prospective Study of Red and Processed Meat Intake in Relation to Cancer Risk. *PLoS Med* **4**, e325.
8. Appel LJ, Moore TJ, Obarzanek E *et al.* (1997) A Clinical Trial of the Effects of Dietary Patterns on Blood Pressure. *N Engl J Med* **336**, 1117-1124.
9. Sacks FM, Svetkey LP, Vollmer WM *et al.* (2001) Effects on Blood Pressure of Reduced Dietary Sodium and the Dietary Approaches to Stop Hypertension (DASH) Diet. *N Engl J Med* **344**, 3-10.
10. Estruch R, Ros E, Salas-Salvado J *et al.* (2013) Primary prevention of cardiovascular disease with a Mediterranean diet. *N Engl J Med* **368**, 1279-1290.
11. Salas-Salvadó J, Bulló M, Babio N *et al.* (2011) Reduction in the Incidence of Type 2 Diabetes With the Mediterranean Diet: Results of the PREDIMED-Reus nutrition intervention randomized trial. *Diabetes Care* **34**, 14-19.
12. Mozaffarian D, Appel LJ, Van Horn L (2011) Components of a Cardioprotective Diet: New Insights. *Circulation* **123**, 2870-2891.
13. Kipnis V, Midthune D, Freedman L *et al.* (2002) Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public Health Nutr* **5**, 915-923.
14. Dhurandhar NV, Schoeller D, Brown AW *et al.* (2015) Energy balance measurement: when something is not better than nothing. *Int J Obes* **39**, 1109-1113.
15. Marshall JR, Chen Z (1999) Diet and health risk: risk patterns and disease-specific associations. *Am J Clin Nutr* **69**, 1351s-1356s.
16. Brennan L, Gibbons H, O'Gorman A (2015) An Overview of the Role of Metabolomics in the Identification of Dietary Biomarkers. *Curr Nutr Rep* **4**, 304-312.
17. Tasevska N, Runswick SA, McTaggart A *et al.* (2005) Urinary Sucrose and Fructose as Biomarkers for Sugar Consumption. *Cancer Epidemiol Biomarkers Prev* **14**, 1287-1294.
18. Potischman N (2003) Biologic and methodologic issues for nutritional biomarkers. *J Nutr* **133** 875s-880s.
19. Wishart DS (2008) Metabolomics: applications to food science and nutrition research. *Trends Food Sci Tech* **19**, 482-493.
20. Scalbert A, Brennan L, Manach C *et al.* (2014) The food metabolome: a window over dietary exposure. *Am J Clin Nutr* **99**, 1286-1308.
21. Atkinson W, Downer P, Lever M *et al.* (2007) Effects of orange juice and proline betaine on glycine betaine and homocysteine in healthy male subjects. *Eur J Nutr* **46**, 446-452.
22. Heinzmann SS, Brown IJ, Chan Q *et al.* (2010) Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. *Am J Clin Nutr* **92**, 436-443.
23. Lloyd AJ, Beckmann M, Favé G *et al.* (2011) Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption. *Br J Nutr* **106**, 812-824.
24. Pujos-Guillot E, Hubert J, Martin JF *et al.* (2013) Mass spectrometry-based metabolomics for the discovery of biomarkers of fruit and vegetable intake: citrus fruit as a case study. *J Proteome Res* **12**, 1645-1659.
25. Edmands WMB, Beckonert OP, Stella C *et al.* (2011) Identification of human urinary biomarkers of cruciferous vegetable consumption by metabonomic profiling. *J Proteome Res* **10**, 4513-4521.
26. Andersen M-BS, Reinbach HC, Rinnan Å *et al.* (2013) Discovery of exposure markers in urine for Brassica-containing meals served with different protein sources by UPLC-qTOF-MS untargeted metabolomics. *Metabolomics* **9**, 984-997.

27. Andersen MB, Kristensen M, Manach C *et al.* (2014) Discovery and validation of urinary exposure markers for different plant foods by untargeted metabolomics. *Anal Bioanal Chem* **406**, 1829-1844.
28. Lloyd AJ, Fave G, Beckmann M *et al.* (2011) Use of mass spectrometry fingerprinting to identify urinary metabolites after consumption of specific foods. *Am J Clin Nutr* **94**, 981-991.
29. Stella C, Beckwith-Hall B, Cloarec O *et al.* (2006) Susceptibility of human metabolic phenotypes to dietary modulation. *J Proteome Res* **5**, 2780-2788.
30. Ross AB, Svelander C, Undeland I *et al.* (2015) Herring and Beef Meals Lead to Differences in Plasma 2-Aminoadipic Acid, β-Alanine, 4-Hydroxyproline, Cetoleic Acid, and Docosahexaenoic Acid Concentrations in Overweight Men. *J Nutr* **11**, 2456-2463.
31. Gibbons H, McNulty BA, Nugent AP *et al.* (2015) A metabolomics approach to the identification of biomarkers of sugar-sweetened beverage intake. *Am J Clin Nutr* **101**, 471-477.
32. Guertin KA, Moore SC, Sampson JN *et al.* (2014) Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. *Am J Clin Nutr* **101**, 1000-1011.
33. Wittenbecher C, Mühlenbruch K, Kröger J *et al.* (2015) Amino acids, lipid metabolites, and ferritin as potential mediators linking red meat consumption to type 2 diabetes. *Am J Clin Nutr* **101**, 1241-1250.
34. Garcia-Aloy M, Llorach R, Urpi-Sarda M *et al.* (2014) Nutrimetabolomics fingerprinting to identify biomarkers of bread exposure in a free-living population from the PREDIMED study cohort. *Metabolomics* **11**, 155-165.
35. O'Sullivan A, Gibney MJ, Brennan L (2011) Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies. *Am J Clin Nutr* **93**, 314-321.
36. O'Gorman A, Morris C, Ryan M *et al.* (2014) Habitual dietary intake impacts on the lipidomic profile. *J Chromatogr B Analyt Technol Biomed Life Sci* **966**, 140-146.
37. Andersen M-BS, Rinnan Å, Manach C *et al.* (2014) Untargeted Metabolomics as a Screening Tool for Estimating Compliance to a Dietary Pattern. *J Proteome Res* **13**, 1405-1418.
38. Jenab M, Slimani N, Bictash M *et al.* (2009) Biomarkers in nutritional epidemiology: applications, needs and new horizons. *Hum Genet* **125**, 507-525.
39. Kuhnle GG (2012) Nutritional biomarkers for objective dietary assessment. *J Sci Food Agric* **92**, 1145-1149.
40. May DH, Navarro SL, Ruczinski I *et al.* (2013) Metabolomic profiling of urine: response to a randomised, controlled feeding study of select fruits and vegetables, and application to an observational study. *Br J Nutr* **110**, 1760-1770.
41. Heinzmann SS, Holmes E, Kochhar S *et al.* (2015) 2-Furoylglycine as a Candidate Biomarker of Coffee Consumption. *J Agric Food Chem* **63**, 8615-8621.
42. Wishart DS, Jewison T, Guo AC *et al.* (2013) HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Res* **41**, D801-D807.
43. Wishart D (2012) Systems Biology Resources Arising from the Human Metabolome Project. In *Genetics Meets Metabolomics: from Experiment to Systems Biology*, pp. 157-175 [K Suhre, editor]. New York, NY: Springer New York.
44. Bouatra S, Aziat F, Mandal R *et al.* (2013) The Human Urine Metabolome. *PLoS ONE* **8**, e73076.
45. Ravanbakhsh S, Liu P, Bjordahl TC *et al.* (2015) Accurate, Fully-Automated NMR Spectral Profiling for Metabolomics. *PLoS ONE* **10**, e0124219.
46. Lang R, Wahl A, Stark T *et al.* (2011) Urinary N-methylpyridinium and trigonelline as candidate dietary biomarkers of coffee consumption. *Mol Nutr Food Res* **55**, 1613-1623.
47. van Velzen EJJ, Westerhuis JA, van Duynhoven JPM *et al.* (2009) Phenotyping Tea Consumers by Nutrikinetic Analysis of Polyphenolic End-Metabolites. *J Proteome Res* **8**, 3317-3330.
48. Daykin CA, Duynhoven JPMV, Groenewegen A *et al.* (2005) Nuclear Magnetic Resonance Spectroscopic Based Studies of the Metabolism of Black Tea Polyphenols in Humans. *J Agric Food Chem* **53**, 1428-1434.
49. Van Dorsten FA, Daykin CA, Mulder TPJ *et al.* (2006) Metabonomics Approach To Determine Metabolic Differences between Green Tea and Black Tea Consumption. *J Agric Food Chem* **54**, 6929-6938.
50. Wang Y, Tang H, Nicholson JK *et al.* (2005) A Metabonomic Strategy for the Detection of the Metabolic Effects of Chamomile (Matricaria recutita L.) Ingestion. *J Agric Food Chem* **53**, 191-196.

51. Tulipani S, Llorach R, Jáuregui O *et al.* (2011) Metabolomics Unveils Urinary Changes in Subjects with Metabolic Syndrome following 12-Week Nut Consumption. *J Proteome Res* **10**, 5047-5058.
52. Cross AJ, Major JM, Sinha R (2011) Urinary Biomarkers of Meat Consumption. *Cancer Epidemiol Biomarkers Prev* **20**, 1107-1111.
53. Lehtonen H-M, Lindstedt A, Järvinen R *et al.* (2013) 1H NMR-based metabolic fingerprinting of urine metabolites after consumption of lingonberries (Vaccinium vitis-idaea) with a high-fat meal. *Food Chem* **138**, 982-990.
54. Vázquez-Fresno R, Llorach R, Alcaro F *et al.* (2012) 1H-NMR-based metabolomic analysis of the effect of moderate wine consumption on subjects with cardiovascular risk factors. *Electrophoresis* **33**, 2345-2354.
55. van Dorsten FA, Grün CH, van Velzen EJJ *et al.* (2010) The metabolic fate of red wine and grape juice polyphenols in humans assessed by metabolomics. *Mol Nutr Food Res* **54**, 897-908.
56. Jacobs DM, Fuhrmann JC, van Dorsten FA *et al.* (2012) Impact of Short-Term Intake of Red Wine and Grape Polyphenol Extract on the Human Metabolome. *J Agric Food Chem* **60**, 3078-3085.
57. Johansson-Persson A, Barri T, Ulmius M *et al.* (2013) LC-QTOF/MS metabolomic profiles in human plasma after a 5-week high dietary fiber intake. *Anal Bioanal Chem* **405**, 4799-4809.
58. Rasmussen LG, Winning H, Savorani F *et al.* (2012) Assessment of dietary exposure related to dietary GI and fibre intake in a nutritional metabolomic study of human urine. *Genes Nutr* **7**, 281-293.
59. Bondia-Pons I, Barri T, Hanhineva K *et al.* (2013) UPLC-QTOF/MS metabolic profiling unveils urinary changes in humans after a whole grain rye versus refined wheat bread intervention. *Mol Nutr Food Res* **57**, 412-422.
60. Beckmann M, Lloyd AJ, Haldar S *et al.* (2013) Hydroxylated phenylacetamides derived from bioactive benzoxazinoids are bioavailable in humans after habitual consumption of whole grain sourdough rye bread. *Mol Nutr Food Res* **57**, 1859-1873.
61. Hjerpsted JB, Ritz C, Schou SS *et al.* (2014) Effect of cheese and butter intake on metabolites in urine using an untargeted metabolomics approach. *Metabolomics* **10**, 1176-1185.
62. Zheng H, Yde CC, Clausen MR *et al.* (2015) Metabolomics Investigation To Shed Light on Cheese as a Possible Piece in the French Paradox Puzzle. *J Agric Food Chem* **63**, 2830-2839.
63. Lloyd AJ, Beckmann M, Haldar S *et al.* (2013) Data-driven strategy for the discovery of potential urinary biomarkers of habitual dietary exposure. *Am J Clin Nutr* **97**, 377-389.
64. Rothwell JA, Fillâtre Y, Martin JF *et al.* (2014) New biomarkers of coffee consumption identified by the non-targeted metabolomic profiling of cohort study subjects. *PLoS One* **9**, e93474.
65. Myint T, Fraser GE, Lindsted KD *et al.* (2000) Urinary 1-Methylhistidine Is a Marker of Meat Consumption in Black and in White California Seventh-day Adventists. *Am J Epidemiol* **152**, 752-755.
66. Edmands WMB, Ferrari P, Rothwell JA *et al.* (2015) Polyphenol metabolome in human urine and its association with intake of polyphenol-rich foods across European countries. *Am J Clin Nutr* **102**, 905-913.
67. Garcia-Aloy M, Llorach R, Urpi-Sarda M *et al.* (2014) Novel Multimetabolite Prediction of Walnut Consumption by a Urinary Biomarker Model in a Free-Living Population: the PREDIMED Study. *J Proteome Res* **13**, 3476-3483.
68. Bouchard-Mercier A, Paradis A-M, Rudkowska I *et al.* (2013) Associations between dietary patterns and gene expression profiles of healthy men and women: a cross-sectional study. *Nutr J* **12**, 1-13.
69. Altmaier E, Kastenmüller G, Römisch-Margl W *et al.* (2010) Questionnaire-based self-reported nutrition habits associate with serum metabolism as revealed by quantitative targeted metabolomics. *Eur J Epidemiol* **26**, 145-156.
70. Peré-Trepat E, Ross AB, Martin F-P *et al.* (2010) Chemometric strategies to assess metabonomic imprinting of food habits in epidemiological studies. *Chemometrics and Intelligent Laboratory Systems* **104**, 95-100.
71. Floegel A, von Ruesten A, Drogan D *et al.* (2013) Variation of serum metabolites related to habitual diet: a targeted metabolomic approach in EPIC-Potsdam. *Eur J Clin Nutr* **67**, 1100-1108.

Table 1: Summary of putative biomarkers identified using a metabolomics approach in intervention studies

Dietary Factor	**Study duration**	**No. of subjects**	**Sample**	**Metabolomic technique**	**Biomarker**	**Author**
Citrus fruit	Acute intervention	8	Fasting & postprandial urine	NMR	Proline betaine	Heinzmann et al.[22]
Citrus fruit	Acute intervention	4	24 h urine	LC-ESI-qTOF, LTQ-Orbitrap	Proline betaine, hydroxyproline betaine, hesperetin 3′-*O*-glucuronide, naringenin 7-*O*-glucuronide, limonene 8,9-diol glucuronide, nootkatone 13,14-diol glucuronide, N-Methyltyramine sulfate	Pujos-Guillot et al.[24]
Citrus fruit	4 weeks intervention	12	24 h urine	LC-ESI-qTOF, LTQ-Orbitrap	Proline betaine, hydroxyproline betaine, hesperetin 3′-*O*-glucuronide, naringenin 7-*O*-glucuronide, limonene 8,9-diol glucuronide, nootkatone 13,14-diol glucuronide, N-Methyltyramine sulfate	Pujos-Guillot et al.[24]
Citrus fruit	Acute intervention	12	Fasting & postprandial urine	FIE-FTICR-MS	Proline betaine, hydroxyproline betaine	Lloyd et al. [23]
Citrus fruit	6 month intervention	107	24 h urine	LC-qTOF	Proline betaine, hesperetin-3-glucuronide	Andersen et al.[27]
Red cabbage	6 month intervention	107	24 h urine	LC-qTOF	3-Hydroxy-3-(methylsulfinyl)propanoic acid, 3-hydroxyhippuric acid-sulfate, 3-hydroxyhippuric acid, iberin N-acetyl-cysteine, N-acetyl-S-(N-3-methylthiopropyl)cysteine, N-acetyl-S-(N-lylthiocarbamoyl)cysteine, sulforaphane N-acetylcysteine	Andersen et al.[27]
Beetroot	6 month intervention	107	24 h urine	LC-qTOF	4-Ethyl-5-aminopyrocatechol sulfate, 4-ethyl-5-methylaminopyrocatechol sulfate, 4-ethylpyridine-2-carboxylic acid glycine conjugate	Andersen et al.[27]
Walnuts	6 month intervention	107	24 h urine	LC-qTOF	5-Hydroxyindole-3-acetic acid	Andersen et al.[27]
Strawberries	6 month intervention	107	24 h urine	LC-qTOF	2,5-Dimethyl-4-methoxy-3(2H)-furanone-sulfate	Andersen et al. [27]
Chocolate	6 month intervention	107	24 h urine	LC-qTOF	6-Amino-5-(N-methylformylamino)-1-methyluracil, theobromine, 7-methyluric acid	Andersen et al.[27]
Raspberries	Acute intervention	24	Fasting & postprandial urine	FIE-FTICR-MS, GC-TOF-MS	Caffeic acid-sulfate, methylepicatechin-sulfate	Lloyd et al. [28]

Cruciferous vegetables	2 week intervention	20	Fasting & postprandial urine	NMR	S-methyl-L-cysteine sulfoxide	Edmands et al.[25]
Cruciferous vegetables	Acute intervention	17	Fasting & postprandial urine	UPLC- qTOF -MS	N-acetyl-S-(N-3- methylthiopropyl) cysteine, N-acetyl-S-(Nallylthiocarbamoyl) cysteine, Iberin N-acetyl-cysteine, N-acetyl-cysteine conjugate, 4-iminopentylisothiocyanate, Sulforaphane N-acetyl- cysteine, Erucin N-acetyl-cysteine, N-Acetyl-(N'-benzylthiocarbamoyl)- cysteine, Sulforaphane N-cysteine	Andersen et al. [26]
Broccoli	Acute intervention	24	Fasting & postprandial urine	FIE-FTICR-MS	Tetronic acid, xylonate/lyxonate, threitol/erythritol	Lloyd et al. [28]
Coffee	Acute intervention	5	Fasting & postprandial urine	NMR	2-furoylglycine	Heinzmann et al.[41]
Coffee	Acute intervention	9	Fasting, morning spot, 24 h urine	HILIC-MS/MS	N-Methylpyridinium, trigonelline	Lang et al.[46]
Black tea	Acute intervention	20	Fasting & postprandial urine	NMR	Hippuric acid, 4-hydroxyhippuric acid, 1,3-dihydrophenyl-2-*O*-sulfate, allic acid, 4-*O*-methylgallic acid	Van Velzen et al. [47]
Black tea	Acute intervention	3	24 h urine	NMR	Hippuric acid, gallic acid, 1,3-dihydroxyphenyl-2-*O*-sulfate	Daykin et al.[48]
Black and green tea	2 day intervention	17	24 h urine	NMR	Hippuric acid, 1,3-dihydrophenyl-2-*O*-sulfate	van Dorsten et al.[49]
Chamomile tea	2 week intervention	14	Spot urine	NMR	Hippuric acid	Wang et al.[50]
Mixed nuts	12 week intervention	42	24 h urine	LC-qTOF, LTQ-Orbitrap	10-Hydroxydecene-4,6-diynoic acid-sulfate, tridecadienoic/tridecynoic acidglucuronide, dodecanedioic acid, 1,3-dihydroxyphenyl-2-*O*-sulfate, *p*-coumaroyl alcohol-glucuronide and -sulfate, *N*-acetylserotonine-sulfate, 5-hydroxyindoleacetic acid, urolitin A-glucuronide, sulfate, sulfate glucuronide	Tulipani et al. [51]
Beef	Acute intervention	17	Postprandial plasma	GC-MS	2-aminoadipic acid,β-alanine, 4-hydroxyproline	Ross et al.[30]

Herring	Acute intervention	17	Postprandial plasma	GC-MS	Cetoleic acid, docosahexaenoic acid	Ross et al. [30]
Salmon	Acute intervention	24	Fasting & postprandial urine	FIE-FTICR-MS	Anserine, methylhistidine, TMAO	Lloyd et al. [28]
Red meat	15 day intervention	17	24 h urine	Ion exchange chromatography	1 and 3 methylhistidine	Cross et al. [52]
Red meat	15 day intervention	12	24 h urine	NMR	Carnitine, creatinine, TMAO, acetyl-carnitine, taurine, 1 and 3 methylhistidine	Stella et al.[29]
Cruciferous vegetables, citrus and soya	2 week intervention	10	Fasting urine	LTQ-FT LC-MS/MS	Proline betaine, sulforaphane, hippuric acid, genistein, daidzein, equol, glycitein, O-desmethylangolensin, trigonelline, (iso)valerlglycine, hydroxyphenylacetyl-glycine, nicotinuric acid	May et al.[40]
Lingonberries	Acute intervention	14	Postprandial urine	NMR	Hippuric acid, 4-hydroxyhippuric acid	Lehtonen et al.[53]
Wine	28 day intervention	61	24 h urine	NMR	Tartrate, 4-hydroxyphenylacetate, mannitol, ethanol	Vazequez-Fresno et al.[54]
Mixed red wine/grape juice extracts	4 week intervention	58	24 h urine	NMR, GC-TOF-MS	Syringic acid, 3-hydroxyhippuric acid, 4-hydroxyhippuric acid, 3-hydroxyphenylacetic acid, 4-hydroxymandelic acid, vanilmandelic acid, hippuric acid, 3-hydroxyphenylpropionic acid, 1,2,3-trihydroxybenzene, 4-hydroxybenzoic acid, homovanillic acid, dihydroferulic acid, phenylacetylglutamine	van Dorsten et al.[55]
Mixed red wine/grape juice extracts	4 day intervention	35	24 h urine	GC-MS, LC-MS	Syringic acid, 3-hydroxyhippuric acid, pyrogallol, 3-hydroxyphenylacetic acid, 3-hydroxyphenylpropionic acid, 3,4-dihydroxyphenylpropionic acid, indole-3-lactic acid, hippuric acid, catechol, 4-hydroxyhippuric acid, 3,4-dihydroxyphenylacetic acid, vanillic acid	Jacobs et al.[56]
Dietary fibres (oat bran, rye bran, & sugar beet fibres)	5 week intervention	25	Fasting plasma	LC-qTOF-MS	2-aminophenol sulphate, 2,6-dihydroxybenzoic acid, hydroxylated and glucuronidated nuatigenin	Johansson-Persson et al.[57]
Dietary fibre	6 month intervention	77	24 h urine	NMR	Hippuric acid	Rasmussen et al.[58]
Whole-grain rye bread	4 week intervention	20	24 h urine	LC-qTOF	3-(3,5-Dihydroxyphenyl)-1-propanoic acid-sulfate and -glucuronide, enterolactone- glucuronide, azelaic acid, 2-	Bondia-Pons et al.[59]

					aminophenol-sulfate, 2,4-dihydroxy-1,4-benzoxazin-3-one, 2-aminophenol-sulfate, 2-4-dihydroxy-1,4-benzoxazin-3-one-sulfate, indolylacryloylglycine, ferulic acid-sulfate, 3,5-dihydroxyphenylethanol-sulfate, 3,5-dihydroxycinnamic acid-sulfate	
Whole-grain sourdough rye bread	8 week intervention	28	24 h urine	FIE-FTICR-MS	HHPAA glucuronide, HPAA sulphate, HBOA glucuronide, N-feruloylglycine sulphate, HHPAA sulphate	Beckmann et al.[60]
Cheese	6 week intervention	23	24 h urine	UPLC-ESI-qTOF	Tyramine, sulphate, isobutyryl glycine (and other acyl glycines), xanthurenic acid, 4-hydroxyphenylacetic acid	Hjerpsted et al.[61]
Milk and cheese	14 day intervention	15	Faeces, 24 h urine	NMR	Milk; citrate, creatine, creatinine, urea, cheese; proline betaine, tyrosine, hippurate	Zheng et al.[62]

NMR, nuclear magnetic resonance. LC, liquid chromatography. ESI, electrospray ionisation. qTOF, quadrupole time-of-flight. LTQ, linear trap quadrupole. FIE, flow infusion electrospray ionisation. FTICR, Fourier transform-ion cyclotron resonance. MS, mass spectrometry. GC, gas chromatography. TOF, time-of-flight. UPLC, ultra-performance liquid chromatography. HILIC, hydrophilic liquid interaction chromatography. TMAO, trimethylamine-N-oxide. LTQ-FT. linear ion trap-Fourier transform mass spectrometer. HHPAA, 2-hydroxy-N-(2-hydroxyphenyl)acetamide. HPAA, N-(2-hydroxyphenyl)acetamide. HBOA, 2-hydroxy-1,4-benzoxazin-3-one. HHPAA, 2-hydroxy-N-(2-hydroxyphenyl)acetamide.

Table 2: Summary of putative biomarkers identified using a metabolomics approach in cohort studies

Dietary Factor	**Dietary assessment tool**	**No. of subjects**	**Sample**	**Metabolomic technique**	**Biomarkers**	**Author**
Oily fish	FFQ	68	Fasting, morning spot, 24 h urine	FIE-FTICR-MS	Methylhistidine	Lloyd et al. [63]
Citrus fruit	24-h dietary record	80	Fasting urine	LC-ESI-qTOF, LTQ-Orbitrap	Proline betaine, hydroxyproline betaine, hesperetin 3′-O-glucuronide, naringenin 7-O-glucuronide, limonene 8,9-diol glucuronide, nootkatone 13,14-diol glucuronide, N-Methyltyramine sulfate	Pujos-Guillot et al.[24]
Sugar sweetened beverages	4-day food diary	565	Fasting urine	NMR	Formate, isocitrate, citrulline, taurine	Gibbons et al.[31]
Citrus, green vegetables, red meat, shellfish, fish, peanuts, coffee etc.	FFQ	502	Fasting serum	UHPLC-MS/MS,GC-MS	Citrus; Scyllo- & chiro-inositol, Greens; CMPF, Red meat; indolepropionate, Shellfish; CMPF, Peanuts; Tryptophan betaine, 4-Vinylphenol sulfate, Coffee; trigonelline-N-methylnicotinate and quinate	Guertin et al.[32]
Coffee	24-h dietary record, FFQ	39	Morning spot urine	UPLC-qTOF-MS	Atractyligenin glucuronide, Cyclo(isoleucyl-prolyl), 1-Methylxanthine, 1,7-dimethyluric acid, kahweol oxide glucuronide, 1-methyluric acid, trigonelline, dimethylxanthine glucuronide, 5-acetylamino-6-formylamino-3-methyluracil (AMFU), hippuric acid, trimethyluric acid, paraxanthine, 3-hydroxyhippuric acid, 1,3 or 3,7-dimethyluric acid, caffeine	Rothwell et al.[64]
Coffee	FFQ	68	Fasting, morning spot, 24 h urine	FIE-FTICR-MS	Dihydrocaffeic acid	Lloyd et al.[63]
Red meat	24-h dietary record, FFQ	126	Fasting urine	Ion exchange chromatography	1-Methylhistidine	Myint et al.[65]
Red meat	FFQ	2047	Serum	FIA-MS/MS	PC aa 36:0, PC aa 36:4, PC aa 38:0, PC aa 38:4, PC ae 34:2, PC ae 34:3, PC ae 36:3, PC ae 36:4, PC ae 36:5, PC ae 38:4, PC ae 38:5, PC ae 38:6, PC ae 40:4, Lyso-PC 20:4, SM 24:1, Ferritin	Wittenbecher et al.[33]

White bread and wholegrain bread	FFQ	155	Fasting spot urine	HPLC–qTOF-MS	2-Aminophenol sulphate, HPAA glucuronide, HHPAA, HMBOA glucuronide, HBOA glycoside, HPPA, HMBOA, DHPPA glucuronide, 3,5-dihydroxyphenylethanol sulphate, DHPPTA sulphate, hydroxybenzoic acid glucuronide, dihydroferulic acid sulphate, enterolactone glucuronide, pyrraline, 3-indolecarboxylic acid glucuronide, riboflavin, 2,8-dihydroxyquinoline glucuronide	Garcia-Aloy et al. [34]
Cruciferous vegetables, citrus and soya	3-day food records, FFQ	93	Fasting urine	LTQ-FT LC-MS/MS	Proline betaine	May et al.[40]
Polyphenol rich foods	24-h dietary record, FFQ	481	24 h urine	UHPLC–qTOF-MS	Coffee; dihydroferulic acid sulfate. Red wine; gallic acid ethyl ester. Citrus fruit; naringenin glucuronide. Tea; 4-O-methylgallic acid. Apples and pears; phloretin glucuronide. Chocolate products; methyl(epi)catechin sulfate	Edmands et al. [66]
Walnuts	FFQ	381	Fasting spot urine	HPLC-qToF-MS	3-indolecarboxylic acid glucuronide, hydroxyindoleacetic acid sulfate, N-acetylserotonin sulfate, 10-hydroxy-decene-4,6-diynoic acid sulfate, tridecadienoic/tridecynoic acid glucuronide, enterolactone glucuronide, urolithins,	Garcia-Aloy et al.[67]

FFQ, food frequency questionnaire. FIE, flow infusion electrospray ionisation. FTICR, Fourier transform-ion cyclotron resonance. MS, mass spectrometry. LC, liquid chromatography. ESI, electrospray ionisation. qTOF, quadrupole time-of-flight. LTQ, linear trap quadrupole. NMR, nuclear magnetic resonance. UHPLC, ultra-high-performance liquid chromatography. GC, gas chromatography. CMPF, 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid. FIA, flow injection analysis. PC aa, diacyl phosphatidylcholines. PC ae, acylalkyl phosphatidylcholines. Lyso-PC, lysophosphatidylcholines. SM, sphingomyelin. HPLC, high-performance liquid chromatography. HPAA, N-(2-hydroxyphenyl) acetamide. HHPAA, 2-hydroxy-N-(2-hydroxyphenyl) acetamide. HMBOA, 2-hydroxy-7-methoxy-2H-1, 4-benzoxazin-3-one. HBOA, 2-hydroxy-1,4-benzoxazin-3-one. HPPA, 2-hydroxy-N-(2-hydroxyphenyl) acetamide. DHPPA, 3-(3,5-dihydroxyphenyl) propanoic acid. DHPPTA, 5-(3,5-dihydroxyphenyl) pentanoic acid.

Table 3: Summary of putative biomarkers identified using dietary patterns and metabolomic profiles

Dietary Patterns	**Dietary pattern approach**	**Sample**	**Metabolomic technique**	**Biomarkers**	**Author**
Prudent and Western dietary patterns	PCA	Fasting plasma	ESI-MS/MS	Western dietary pattern; increased amino acids and short-chain acylcarnitines	Bouchard-Mercier et al.[68]
Healthy, unhealthy, traditional Irish dietary pattern	k-means cluster analysis	Fasting urine	NMR	Healthy; glycine, phenylacetylglutamine and actetoacetate Traditional Irish; TMAO, *O*-acetylcarnitine and nndimethylglycine	O'Sullivan et al.[35]
7 dietary patterns (e.g. healthy diet, traditional Bavarian)	PCA	Fasting plasma	ESI-MS/MS	Healthy diet; decrease in the degree of saturation of the fatty acid moieties of different glycerol-phosphatidylcholines	Altmaier et al.[69]
7 dietary patterns (e.g. dietary fat lipid pattern, alcohol lipid pattern)	PCA	Fasting serum	ESI-MS/MS	Alcohol consumption; LPCeC18:0 Fish consumption; LPEaC18:2, PEaaC38:4	O'Gorman et al.[36]
5 dietary patterns (e.g. energy intake, plant versus animal based diet)	PCA	Fasting plasma	NMR	Energy intake; greater concentrations of lipids related high energy intake, higher circulating phosphatidycholine related to lower energy intake. Animal based diet; higher concentrations of lysine, arginine, glutamine/glutamate, threonine, aspartate/asparagine,citrate and polyol compounds.	Peré-Trepat et al.[70]

New Nordic Diet (NND) and an Average Danish Diet (ADD)		24 h urine	UPLC-qTOF-MS	NND diet; TMAO, hippuric acid, hydroquinone-glucuronide, (2-oxo-2,3-dihydro-1H-indol-3-yl)acetic acid and 3,4,5,6-tetrahydrohippurate. ADD diet; pyrraline, glucuronide conjugated products, theobromine, 7-methyluric acid, 3,7-dimethyluric acid, 7-methylxanthine, 6-amino-5-[N-methylformylamino]-1-methyluracil, proline betaine and glucuronides of perillic acid.	Andersen et al. [37]
Dietary patterns e.g. high intake of butter/low intake of margarine, high intake of red meat and fish/low intake of whole-grain bread, tea and coffee	RRR	Fasting serum	FIA-MS/MS	High intake of butter and low intake of margarine; acylcarnitines, acyl-alkyl-phosphatidylcholines, lyso-phosphatidylcholines and hydroxy-sphingomyelins. High intake of red meat and fish and low intake of whole-grain bread and tea; hexose and phosphatidylcholines.	Floegal et al. [71]

PCA, principal component analysis. ESI, electrospray ionisation. MS, mass spectrometry. LPCeC18:0, lysophosphatidylcholine alkyl C18:0. LPEaC18:2, lysophosphatidylethanolamine acyl C18:2. PEaaC38:4, phoshatidylethanolamine diaclyl C38:4. TMAO, trimethylamine-N-oxide. NMR, nuclear magnetic resonance. UPLC, ultra-performance liquid chromatography. qTOF, quadrupole time-of-flight. RRR, reduced rank regression. FIA, flow injection analysis.