Title	þÿ M e t a b o l o m i c s B a s e d D i e t a r y B i o m a r k e r s i n N u t r i t i o n a l E p i d Future Opportunities
Authors(s)	Brennan, Lorraine; Hu, Frank B.
Publication date	2018-04-24
Publication information	Molecular Nutrition and Food Research, 63 (1):
Publisher	Wiley
Item record/more information	http://hdl.handle.net/10197/9977
Publisher's statement	This is the author's version of the following article: Brennan, L., Hu, F.B. (2018) þÿ " M e t a b o l o m i c s B a s e d D i e t a r y B i o m a r k e r s i n N u t r i t i o n a l E p i Future Opportunities" Molecular Nutrition & Food Research VOL 63, ISSUE 1), which has been published in final form at http://dx.doi.org/10.1002/mnfr.201701064.
Publisher's version (DOI)	10.1002/mnfr.201701064

Metabolomics based dietary biomarkers in nutritional epidemiology- current status and future opportunities

Lorraine Brennan[1] and Frank B. Hu[2,3,4]

[1] UCD School of Agriculture and Food Science, UCD, Belfield, Dublin 4, Ireland

[2] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA. [3] Department of Nutrition, Harvard TH Chan School of Public Health, Boston, MA. [4]Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA.

Professor Lorraine Brennan
UCD School of Agriculture and Food Science,
UCD Institute of Food and Health,
UCD, Belfield, Dublin 4, Ireland.

Email: lorraine.brennan@ucd.ie
Phone: 00 353 1 7162811

Abstract

The application of metabolomics in nutrition epidemiology holds great promise and there is a high expectation that it will play a leading role in deciphering the interactions between diet and health. However, while significant progress has been made in identification of putative biomarkers more work is needed to address the use of the biomarkers in dietary assessment. The aim of this review to critically evaluate progress in these areas and to identify challenges that need to be addressed going forward. The notable applications of dietary biomarkers in nutritional epidemiology include (1) Determination of food intake based on biomarkers levels and calibration equations from feeding studies (2) Classification of individuals into dietary patterns based on the urinary metabolic profile and (3) Application of metabolome-wide-association studies. Further work is needed to address some specific challenges to enable biomarkers to reach their full potential.

Keywords: metabolomics, dietary biomarkers, food intake, dietary assessment

1. Introduction

The application of metabolomics in nutrition epidemiology holds great promise and there is a high expectation that it will play a big role in deciphering out the interactions between diet and health. The cornerstone method for measuring food intake in epidemiology has been the food frequency questionnaires (FFQs) and it is well documented that this self-reported instrument is associated with a number of issues such as random and systematic errors. Examples of such errors include under-reporting, recall errors and difficulty in assessment of portion sizes [1-3].

Such errors can lead to reduced power, underestimated associations and false findings which may contribute to inconsistencies in the field of nutritional epidemiology [3-5]. To overcome such issues there has been intense interest in the potential of dietary biomarkers as objective measures of intake. Currently dietary biomarkers exist for salt, protein, sucrose/fructose intake (sodium/nitrogen/sucrose and fructose measured in 24h urine samples) and energy expenditure (the doubly labelled water technique) [1, 6]. The use of the biomarkers to develop calibrated equations to correct for estimates for protein and energy intake have been shown to significantly impact on one's ability to examine diet-disease associations: for example inclusion of biomarker calibrated data in the Women's Health Initiative cohorts allowed for disease associations to be revealed that otherwise would have not been identified [7, 8]. Using urinary nitrogen and doubly labelled water as biomarkers Tinker and colleagues calibrated the intakes of protein and energy in a group of postmenopausal women [7]. Using biomarker calibrated equations they identified associations between protein and energy intake with diabetes risk; these associations were not identified in the uncalibrated data. Similarly, Prentice and colleagues reported that following biomarker correction energy intake was positively associated with coronary heart disease incidence and protein density was inversely associated. These relationships were only apparent when the biomarker data was employed [8]. These examples clearly highlight the potential of validated dietary biomarkers in nutritional epidemiology.

With the clear potential of such biomarkers there has been increased interest in using metabolomics for identifying biomarkers that are associated with dietary/food intake. In recent years there has been a plethora of publications demonstrating correlations between metabolites and food intake data highlighting the potential for the field [9-11]. Putative

biomarkers now exist for a range of foods including but not limited to red meat, coffee, nuts, wine, vegetables, legumes, citrus fruit, tea, sugar sweetened beverages [12]. While this range may seem extensive the challenge faced in moving these putative biomarkers into use is that many of these have not been validated in dose response intervention studies or examined for their sensitivity of specificity. A complete review of the these putative biomarkers and examination of their validity is underway by the FoodBall consortium (http://foodmetabolome.org/). Furthermore, a new flexible classification scheme for biomarkers related to food intake has been recently presented [13]. With this in mind, the present review will not discuss an exhaustive list of potential biomarkers but rather focus on the potential use of well validated dietary biomarkers in nutritional epidemiology.

While there is no doubt that there is great potential for the role of dietary biomarkers in nutritional epidemiology there are also a number of issues and limitations that need to addressed if it is to reach its full potential. The present review will examine some of the potential applications of dietary biomarkers in nutritional epidemiology while discussing pertinent issues and highlighting key challenges that exist (Figure 1).

2. Use biomarkers to estimate dietary intake: is it a reality?

Traditionally in epidemiology dietary biomarkers have been measured in a sub population of the cohort and regression calibration equations used to correct the self-reported data [5, 14]. This has been extremely successful for biomarkers of energy and protein intake. The same principles could apply for dietary biomarkers that reflect food intake (often referred to as food intake biomarkers) - where the biomarkers could be used to estimate intake from calibration equations and then applied using regression calibration to the larger cohort. Although there has been a sustained interest in food intake biomarkers the majority of publications demonstrate correlations between the metabolites and food intake data and there is a paucity of papers demonstrating that biomarkers can actually determine intake and thus could be used in larger epidemiological studies to either estimate intake or to correct self-reported data as is often done with energy and protein biomarkers.

A recent paper by Lampe and colleagues illustrated the potential of candidate biomarkers for dietary assessment in terms of nutrient intake [15]. The work demonstrated that a series of dietary biomarkers including carotenoids, tocopherols, folate, and vitamin B12 performed as

well as established energy (doubly labelled water) and protein biomarkers (urinary nitrogen) in representing nutrient intake. Importantly, the person characteristics did not contribute substantially to the overall model of biomarkers against intake. While the work didn't specifically examine food intake it illustrates very clearly that the biomarkers were as good at estimating nutrient intake as the well-established biomarkers. Furthermore, it demonstrated the potential for blood based biomarkers. The calibration equations developed within this study have the potential to be applied in future studies to correct for the measurement error in self-reported nutrient intake: this will be particularly important when examining diet and disease associations.

A similar approach for correction of food intake is also feasible if suitable feeding studies have been performed to establish calibration equations. Two recent publications are of note in progressing this concept and demonstrating feasibility. Using NMR spectroscopy Garcia-Perez and colleagues analysed urine samples from a controlled intervention with the objective to identify a biomarker of grape intake [16]. Urinary tartaric acid was identified as a biomarker displaying a dose response relationship allowing the authors to construct calibration curves of urinary tartaric acid against red grape intake (g/day). The agreement between estimated intake and actual intake was good and a correlation coefficient of R^2=0.9 was reported. Gibbons et al reported a similar approach, however, the calibration curves were also used in an independent population study to estimate intake [17]. Using a controlled dietary intervention approach participants consumed standardized breakfasts for three consecutive days over three weeks [17] where the quantity of the food of interest was varied. In this instance, intake of orange juice decreased over a 3 week period from average of 520 g/day to 30 g/day. Calibration curves were constructed with the urinary proline betaine concentration against the known orange juice intake (g/day). A correlation of 0.92 was reported between actual intake and predicted intake again highlighting the high level of agreement. To the best of our knowledge this biomarker has stronger correlations with intake compared to any other proposed biomarker of citrus intake (for example citrus flavonoids)[18, 19]. Furthermore, the ability of the biomarker to estimate intake was tested in an independent cross sectional study of 560 individuals. Using the calibration curves determined in the controlled intervention study the citrus intake (g/day) was estimated from the urinary concentration of proline betaine. There was excellent agreement between the self-report intake (estimated from a 4 day semi-weighed food diary) and the estimated intake from the biomarker. The agreement was assessed by Bland and Altman analysis and mean

difference between the methods was 4.3g. The significance of this study lies with the fact that it clearly demonstrates how biomarkers may be used in a larger cohort/population setting to estimate food intake. The challenge going forward will be to ascertain if such biomarkers can give information on long term intake.

In conclusion, these studies indicate the huge potential of biomarkers in the determination of food intake. It is also noteworthy that the examples above have employed either urine or blood samples and there is great potential in using the combination of the two biofluids. However, for markers to function as estimates of intake it is essential that calibration curves are established in controlled interventions. These curves can then be used to estimate intake in the larger population based study. To make a significant impact it is now imperative that such work is performed on other foods and nutrients of interest in order to achieve a good capture of a person's daily intake. With this in mind, it is also important to acknowledge that identification and validation of specific biomarkers for all foods is unlikely and that combining biomarker and traditional approaches will be essential. For example, in very large cohort studies it may not be feasible to measure biomarkers in all samples- in this case a sub-sample may have estimates of intake from biomarker data and a regression-calibration equation could be derived to correct the remaining self-reported data. Such an approach would be very important for studying diet-disease relationships in large cohort studies. It is also possible to use these food intake biomarkers directly and examine the relationships between intake and disease or disease markers; this concept is discussed further in the metabolome wide association studies below.

3. Use of biomarkers to estimate dietary patterns: what is the current status?

An interest in dietary patterns in relation to health outcomes has emerged in recent years as the importance of the whole diet as opposed to specific nutrients is emerging [20-22]. The ultimate goal of using dietary biomarkers within the context of dietary patterns will be to classify individuals into certain patterns or to confirm adherence/non-adherence to certain predefined scores. However, while there has been many publications demonstrating the link between dietary patterns and metabolomic profiles there are only a limited number of publications that demonstrated the ability to classify or assign people into certain dietary patterns based on biomarkers.

Andersen and colleagues used an untargeted UPLC-qTOF-MS metabolic phenotyping approach to distinguish between two dietary patterns with the purpose of developing a compliance measure for adherence to certain dietary patterns [23]. As part of this study the participants (n=181) were randomly assigned to follow a New Nordic Diet (NND) or an Average Danish Diet (ADD) for 6 months. Using the urinary metabolic profile a multivariate model was established that could distinguish the two dietary patterns. Examination of the model revealed that it performed well with a misclassification error rate of 19% (estimated using a validation set of samples that were not used to develop the model); this work demonstrates how one could use a biomarker approach to identify non-compliance to a certain diet. Using a controlled feeding study Esko and colleagues defined 3 diets that differed in macronutrient composition: low fat (60% carbohydrate, 20% fat, 20% protein), low glycemic index (40% carbohydrate, 40% fat, 20% protein) and very-low carbohydrate (10% carbohydrate, 60% fat, 30% protein) [24]. Using a targeted metabolomics approach plasma metabolites that were different between the diets were identified and then used to build a classification model to distinguish between the diets. The importance of such a model could be in the checking of adherence to specific diets but could also be use in large epidemiological studies to identify dietary patterns and link to health outcomes.

Garcia-Perez and colleagues used controlled interventions to define dietary patterns and built multivariate models to classify people into a dietary pattern based in urinary metabolomics data [25] . The four diets used in this study were designed to have a step variance in the WHO healthy eating guidelines that are planned to prevent non-communicable diseases (NCDs). The multivariate models were based on the urinary metabolomic profiles and were confirmed in independent studies. Interestingly, in these studies the individuals were classified into dietary patterns that were associated with higher or lower non-communicable disease risk. Further development of this work to obtain even more rigorous classification will be important for advancement of this field and the potential use of the approach for examining the links between diet and health. Furthermore, work should be directed at establishing if such approaches can pick up long term dietary patterns [26].

In work from Brennan and colleagues a cross sectional study design was used to derive a model based on urinary metabolomic data which could classify subjects into either a healthy dietary pattern or an unhealthy dietary pattern [27]. The healthy dietary patterns was characterised by higher intakes of breakfast cereals and porridge , low fat and skimmed

milks, fruit and fish while the unhealthy dietary patterns was classified by higher intakes of chips and processed potatoes, savoury snacks and meat products. Furthermore, the dietary patterns were supported by significant differences in blood parameters such as higher folate and 25(OH)-vitamin D in the healthy dietary pattern. The developed model was then tested in an independent study and it was found that it had a high correct classification rate. Further development of the model to include a more diverse range of biomarkers and more dietary patterns would be an important next step. The work presented by these three example demonstrate the huge potential there is in the use of dietary biomarkers to assign dietary patterns to individuals. Further refinement and development of the models and concepts should allow for rapid and objective classification of individuals which in turn could be used in large epidemiological studies to examine the associations between certain dietary patterns and health outcomes.

While the above four examples are particularly encouraging in terms of classification into dietary patterns there are also other noteworthy potential applications of metabolomics in terms of dietary patterns [28] [29, 30]. Recently, the relationship between 4 dietary patterns estimated by 4 diet quality indices and serum metabolites was reported by Playdon and colleagues [29]. For example the World Health Organisation Healthy Diet Indicator (WHO-HDI) correlated with metabolites related to polyunsaturated fat and fibre. The alternative Mediterranean Diet Score (aMED) was correlated with metabolites related to fish, nuts, fruit and vegetables. Further analysis also revealed that the dietary patterns were associated with certain metabolic pathways thus identifying potential mechanisms influenced by diet quality. The examination of relationship between dietary patterns and endogenous metabolites has the potential to further our understanding of the influence of diet on certain metabolic pathways. In fact in this context, the relationship between endogenous metabolic pathways and dietary patterns and indices will be more informative than examining correlations with biomarkers perceived to be related to food intake. Collectively, these studies provide a strong evidence base for the potential of metabolomics based biomarkers as a tool for assessment of dietary intake, evaluating compliance to a dietary pattern and identifying and evaluating relationships between diet patterns and disease.

4. Potential of metabolome wide association studies.

Application of metabolomics to obtain an understanding of mechanisms underlying diet related diseases in large epidemiological studies has the potential to identify early prevention strategies. In this scenario, attention needs to given to a number of factors including the highly correlated nature of metabolomics data, the potential confounders and appropriate strategies for multiple comparison corrections. One approach that has been particularly successful is the metabolome-wide-association strategy (MWA): using this approach one can identify new biomarkers that can link with a particular condition or a quantitative trait [31, 32]. The modelling involved can account for the highly correlated metabolomic data, multiple comparisons and control for potential confounding variables [33, 34]. Recent advancements have improved the visualization of the results which in turn can aid interpretation [35]. This approach has been effective in a number of studies but to date has been under used in metabolomics studies. Application of the MWA approach to the INTERMAP study has indicated potential causal factors for high blood pressure across geographically diverse populations, however, further work is needed to confirm these findings as the results to date are only associations [36]. They identified a number of metabolites that were associated with blood pressure: both formate and hippurate were inversely associated with systolic and diastolic blood pressure and alanine was positively associated with blood pressure. These metabolites are influenced by a number of factors including diet and the gut microbiome: however, by placing the focus on the metabolite irrespective of whether it was correlated with self-reported dietary data allowed the authors to examine multiple exposures and factors simultaneously. Further application of the MWA approach should reveal patterns of metabolites associated with certain diet related diseases. The potential of also combining other omic data opens up new avenues.

5. Challenges in the field

One of main challenges in the assessment of dietary intake is obtaining good estimates of long term intake. From a biomarker viewpoint this also remains a challenge and one of the potential routes of addressing this is examination of repeated samples (2 to 3 repeats) over a time frame of 6- 12 months depending on the feasibility. A recent study demonstrated that urinary excretion of various biomarkers was reasonably reproducible and concluded that three urine samples are sufficient for the long-term exposure status in epidemiologic studies [37]. A further challenge that exists is the large inter-individual variation in response to foods making it difficult to identify biomarkers that respond reproducibly across populations.

Furthermore, for biomarkers of long term intake one also needs to consider the potential contribution of the gut microbiota.

In the examples given in this manuscript for estimation of food intake- a single biomarker estimated intake of a single food. As we move forward in this field we will arrive at more complex situations such as combinations of biomarkers predicting intake of a single food- in this scenario we will need new mathematical models to develop appropriate calibration curves. In essence, for the field to grow and mature close collaboration between data scientists, metabolomics experts and nutritional epidemiologists is essential.

As our ability to measure more and more biomarkers increases we need to realise that many endogenous biomarkers may be correlated to food intake but will not be a good estimate for food intake as their metabolism including degradation and production will be tightly regulated and influenced by other physiological factors. Such biomarkers may have a potential role in examining the mechanisms underpinning health effects of certain diets. However, care needs to be taken that correct adjustments are performed to account for influence of other factors. Additionally, we need to employ recently developed assessment systems to examine the validity of the biomarkers in question [38]. This proposed system includes examination of biological plausibility, dose-response relationships, robustness (including specificity) and reliability.

To allow for cross cohort comparisons and pooling of data it is imperative that we move to measuring quantitative data- this will bring challenges to the analytical and data pipelines. Nevertheless, to ensure that we get meaningful data that we can link to dietary intake this will be essential. The measurement of quantitative data is inherent in the NMR approach, however, the lack of sensitivity means that we are limited to abundant metabolites. Furthermore, the identification of unknown compounds in samples is a major bottleneck for the identification of new biomarkers of food intake. A combined effort from the community will be essential to make significant progress in this field. Combining 2D-NMR experiments with LC-MS based approaches has the potential to aid identification of unknown compounds.

Finally, while challenges remain in the application of dietary biomarkers in nutritional epidemiology significant progress has also been achieved in the last 5 years. With global

efforts, it will be possible to address these and realise the full potential of metabolomics in nutritional epidemiology.

Acknowledgements

LB acknowledges the following funding: The European Research Council ERC (647783) and SFI (JPI-HDHL/B3075, Foodball).

Conflict of Interest

The authors have no conflict of interests to declare.

Figure Legends

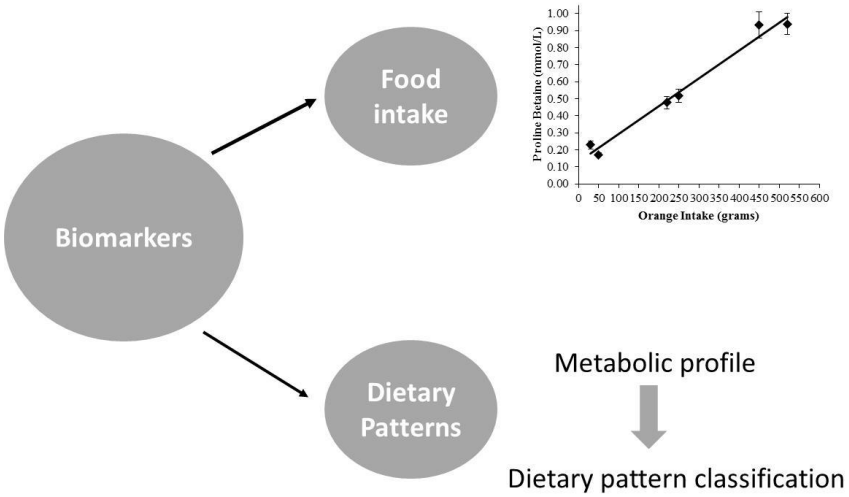

Figure 1. An overview of potential applications of biomarkers. Biomarkers can be used to determine food intake through the use of well-defined calibration curves. Using a biomarker level the amount of food intake in g/day can be estimated. In terms of dietary patterns one can use the metabolic profile to classify individuals into dietary patterns.

References

[1] Bingham, S. A., Biomarkers in nutritional epidemiology. *Public Health Nutr* 2002, *5*, 821-827.
[2] Kipnis, V., Midthune, D., Freedman, L., Bingham, S., Day, N. E., Riboli, E., Ferrari, P., Carroll, R. J., Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public health nutrition* 2002, *5*, 915-923.
[3] Dhurandhar, N. V., Schoeller, D., Brown, A. W., Heymsfield, S. B., Thomas, D., Sorensen, T. I. A., Speakman, J. R., Jeansonne, M., Allison, D. B., the Energy Balance Measurement Working, G., Energy balance measurement: when something is not better than nothing. *Int J Obes* 2015, *39*, 1109-1113.
[4] Marshall, J. R., Chen, Z., Diet and health risk: risk patterns and disease-specific associations. *Am J Clin Nutr* 1999, *69*, 1351s-1356s.
[5] Spiegelman, D., McDermott, A., Rosner, B., Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr* 1997, *65*, 1179S-1186S.
[6] Tasevska, N., Runswick, S. A., McTaggart, A., Bingham, S. A., Urinary Sucrose and Fructose as Biomarkers for Sugar Consumption. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2005, *14*, 1287-1294.
[7] Tinker, L. F., Sarto, G. E., Howard, B. V., Huang, Y., Neuhouser, M. L., Mossavar-Rahmani, Y., Beasley, J. M., Margolis, K. L., Eaton, C. B., Phillips, L. S., Prentice, R. L., Biomarker-calibrated dietary energy and protein intake associations with diabetes risk among postmenopausal women from the Women's Health Initiative. *Am J Clin Nutr* 2011, *94*, 1600-1606.
[8] Prentice, R. L., Huang, Y., Kuller, L. H., Tinker, L. F., Horn, L. V., Stefanick, M. L., Sarto, G., Ockene, J., Johnson, K. C., Biomarker-calibrated energy and protein consumption and cardiovascular disease risk among postmenopausal women. *Epidemiology* 2011, *22*, 170-179.
[9] Guertin, K. A., Loftfield, E., Boca, S. M., Sampson, J. N., Moore, S. C., Xiao, Q., Huang, W. Y., Xiong, X., Freedman, N. D., Cross, A. J., Sinha, R., Serum biomarkers of habitual coffee consumption may provide insight into the mechanism underlying the association between coffee consumption and colorectal cancer. *Am J Clin Nutr* 2015, *101*, 1000-1011.
[10] Playdon, M. C., Sampson, J. N., Cross, A. J., Sinha, R., Guertin, K. A., Moy, K. A., Rothman, N., Irwin, M. L., Mayne, S. T., Stolzenberg-Solomon, R., Moore, S. C., Comparing metabolite profiles of habitual diet in serum and urine. *Am J Clin Nutr* 2016, *104*, 776-789.
[11] Edmands, W. M., Ferrari, P., Rothwell, J. A., Rinaldi, S., Slimani, N., Barupal, D. K., Biessy, C., Jenab, M., Clavel-Chapelon, F., Fagherazzi, G., Boutron-Ruault, M. C., Katzke, V. A., Kuhn, T., Boeing, H., Trichopoulou, A., Lagiou, P., Trichopoulos, D., Palli, D., Grioni, S., Tumino, R., Vineis, P., Mattiello, A., Romieu, I., Scalbert, A., Polyphenol metabolome in human urine and its association with intake of polyphenol-rich foods across European countries. *Am J Clin Nutr* 2015, *102*, 905-913.
[12] Scalbert, A., Brennan, L., Manach, C., Andres-Lacueva, C., Dragsted, L. O., Draper, J., Rappaport, S. M., van der Hooft, J. J., Wishart, D. S., The food metabolome: a window over dietary exposure. *Am J Clin Nutr* 2014, *99*, 1286-1308.
[13] Gao, Q., Pratico, G., Scalbert, A., Vergeres, G., Kolehmainen, M., Manach, C., Brennan, L., Afman, L. A., Wishart, D. S., Andres-Lacueva, C., Garcia-Aloy, M., Verhagen, H., Feskens, E. J. M., Dragsted, L. O., A scheme for a flexible classification of dietary and health biomarkers. *Genes Nutr* 2017, *12*, 34.
[14] Freedman, L. S., Midthune, D., Carroll, R. J., Tasevska, N., Schatzkin, A., Mares, J., Tinker, L., Potischman, N., Kipnis, V., Using regression calibration equations that combine self-reported intake and biomarker measures to obtain unbiased estimates and more powerful tests of dietary associations. *Am J Epidemiol* 2011, *174*, 1238-1245.
[15] lampe, J., Huang, Y., Neuhouser, M. L., Tinker, L., Song, X., Schoeller, D., Kim, S., Raftery, D., Di, C., Zheng, C., Schwarz, Y., Van Horn, L., Thomson, C., Mossavar-Rahmani, Y., Beresford, S. A.,

Prentice, R., Dietary Biomarker Evaluation in a controlled feeding study in women from the womens health initative cohort. *American Journal of Clinical Nutrition* 2016.
[16] Garcia-Perez, I., Posma, J. M., Chambers, E. S., Nicholson, J. K., J, C. M., Beckmann, M., Draper, J., Holmes, E., Frost, G., An Analytical Pipeline for Quantitative Characterization of Dietary Intake: Application To Assess Grape Intake. *J Agric Food Chem* 2016, *64*, 2423-2431.
[17] Gibbons, H., Michielsen, C. J. R., Rundle, M., Frost, G., McNulty, B. A., Nugent, A. P., Walton, J., Flynn, A., Gibney, M. J., Brennan, L., Demonstration of the utility of biomarkers for dietary intake assessment; proline betaine as an example. *Mol Nutr Food Res* 2017.
[18] Krogholm, K. S., Bysted, A., Brantsaeter, A. L., Jakobsen, J., Rasmussen, S. E., Kristoffersen, L., Toft, U., Evaluation of flavonoids and enterolactone in overnight urine as intake biomarkers of fruits, vegetables and beverages in the Inter99 cohort study using the method of triads. *Br J Nutr* 2012, *108*, 1904-1912.
[19] Nielsen, S. E., Freese, R., Kleemola, P., Mutanen, M., Flavonoids in human urine as biomarkers for intake of fruits and vegetables. *Cancer Epidemiol Biomarkers Prev* 2002, *11*, 459-466.
[20] de Koning, L., Chiuve, S. E., Fung, T. T., Willett, W. C., Rimm, E. B., Hu, F. B., Diet-quality scores and the risk of type 2 diabetes in men. *Diabetes Care* 2011, *34*, 1150-1156.
[21] InterAct, C., Adherence to predefined dietary patterns and incident type 2 diabetes in European populations: EPIC-InterAct Study. *Diabetologia* 2014, *57*, 321-333.
[22] Nanri, A., Mizoue, T., Shimazu, T., Ishihara, J., Takachi, R., Noda, M., Iso, H., Sasazuki, S., Sawada, N., Tsugane, S., Japan Public Health Center-Based Prospective Study, G., Dietary patterns and all-cause, cancer, and cardiovascular disease mortality in Japanese men and women: The Japan public health center-based prospective study. *PLoS One* 2017, *12*, e0174848.
[23] Andersen, M.-B. S., Rinnan, Å., Manach, C., Poulsen, S. K., Pujos-Guillot, E., Larsen, T. M., Astrup, A., Dragsted, L. O., Untargeted Metabolomics as a Screening Tool for Estimating Compliance to a Dietary Pattern. *J Proteome Res* 2014, *13*, 1405-1418.
[24] Esko, T., Hirschhorn, J. N., Feldman, H. A., Hsu, Y. H., Deik, A. A., Clish, C. B., Ebbeling, C. B., Ludwig, D. S., Metabolomic profiles as reliable biomarkers of dietary composition. *Am J Clin Nutr* 2017, *105*, 547-554.
[25] Garcia-Perez, I., Posma, J. M., Gibson, R., Chambers, E. S., Hansen, T. H., Vestergaard, H., Hansen, T., Beckmann, M., Pedersen, O., Elliott, P., Stamler, J., Nicholson, J. K., Draper, J., Mathers, J. C., Holmes, E., Frost, G., Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. *The lancet. Diabetes & endocrinology* 2017, *5*, 184-195.
[26] Bhupathiraju, S. N., Hu, F. B., One (small) step towards precision nutrition by use of metabolomics. *The lancet. Diabetes & endocrinology* 2017, *5*, 154-155.
[27] Gibbons, H., Carr, E., McNulty, B. A., Nugent, A. P., Walton, J., Flynn, A., Gibney, M. J., Brennan, L., Metabolomic-based identification of clusters that reflect dietary patterns. *Mol Nutr Food Res* 2017, *61*.
[28] Floegel, A., von Ruesten, A., Drogan, D., Schulze, M. B., Prehn, C., Adamski, J., Pischon, T., Boeing, H., Variation of serum metabolites related to habitual diet: a targeted metabolomic approach in EPIC-Potsdam. *Eur J Clin Nutr* 2013, *67*, 1100-1108.
[29] Playdon, M. C., Moore, S. C., Derkach, A., Reedy, J., Subar, A. F., Sampson, J. N., Albanes, D., Gu, F., Kontto, J., Lassale, C., Liao, L. M., Mannisto, S., Mondul, A. M., Weinstein, S. J., Irwin, M. L., Mayne, S. T., Stolzenberg-Solomon, R., Identifying biomarkers of dietary patterns by using metabolomics. *Am J Clin Nutr* 2017, *105*, 450-465.
[30] Bouchard-Mercier, A., Rudkowska, I., Lemieux, S., Couture, P., Vohl, M. C., The metabolic signature associated with the Western dietary pattern: a cross-sectional study. *Nutrition journal* 2013, *12*, 158.
[31] Nicholson, J. K., Holmes, E., Elliott, P., The metabolome-wide association study: a new look at human disease risk factors. *J Proteome Res* 2008, *7*, 3637-3638.
[32] Yap, I. K., Brown, I. J., Chan, Q., Wijeyesekera, A., Garcia-Perez, I., Bictash, M., Loo, R. L., Chadeau-Hyam, M., Ebbels, T., De Iorio, M., Maibaum, E., Zhao, L., Kesteloot, H., Daviglus, M. L.,

Stamler, J., Nicholson, J. K., Elliott, P., Holmes, E., Metabolome-wide association study identifies multiple biomarkers that discriminate north and south Chinese populations at differing risks of cardiovascular disease: INTERMAP study. *J Proteome Res* 2010, *9*, 6647-6654.
[33] Rodriguez-Martinez, A., Posma, J. M., Ayala, R., Neves, A. L., Anwar, M., Petretto, E., Emanueli, C., Gauguier, D., Nicholson, J. K., Dumas, M. E., MWASTools: an R/Bioconductor package for metabolome-wide association studies. *Bioinformatics* 2017.
[34] Chadeau-Hyam, M., Ebbels, T. M., Brown, I. J., Chan, Q., Stamler, J., Huang, C. C., Daviglus, M. L., Ueshima, H., Zhao, L., Holmes, E., Nicholson, J. K., Elliott, P., De Iorio, M., Metabolic profiling and the metabolome-wide association study: significance level for biomarker identification. *J Proteome Res* 2010, *9*, 4620-4627.
[35] Castagne, R., Boulange, C. L., Karaman, I., Campanella, G., Santos Ferreira, D. L., Kaluarachchi, M. R., Lehne, B., Moayyeri, A., Lewis, M. R., Spagou, K., Dona, A. C., Evangelos, V., Tracy, R., Greenland, P., Lindon, J. C., Herrington, D., Ebbels, T. M. D., Elliott, P., Tzoulaki, I., Chadeau-Hyam, M., Improving Visualization and Interpretation of Metabolome-Wide Association Studies: An Application in a Population-Based Cohort Using Untargeted 1H NMR Metabolic Profiling. *J Proteome Res* 2017, *16*, 3623-3633.
[36] Holmes, E., Loo, R. L., Stamler, J., Bictash, M., Yap, I. K., Chan, Q., Ebbels, T., De Iorio, M., Brown, I. J., Veselkov, K. A., Daviglus, M. L., Kesteloot, H., Ueshima, H., Zhao, L., Nicholson, J. K., Elliott, P., Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* 2008, *453*, 396-400.
[37] Sun, Q., Bertrand, K. A., Franke, A. A., Rosner, B., Curhan, G. C., Willett, W. C., Reproducibility of urinary biomarkers in multiple 24-h urine samples. *Am J Clin Nutr* 2017, *105*, 159-168.
[38] Pratico, G., Gao, Q., Scalbert, A., Vergeres, G., Kolehmainen, M., Manach, C., Brennan, L., Pedapati, S. H., Afman, L. A., Wishart, D. S., Vazquez-Fresno, R., Lacueva, C. A., Garcia-Aloy, M., Verhagen, H., Feskens, E. J. M., Dragsted, L. O., Guidelines for Biomarker of Food Intake Reviews (BFIRev): how to conduct an extensive literature search for biomarker of food intake discovery. *Genes Nutr* 2018, *13*, 3.