APPLICATION OF NUTRITIONAL METABOLOMICS IN EPIDEMIOLOGICAL STUDIES

The application of metabolomics in assessment of nutrition, sources of variation in food-related metabolites, and identification of -omics features of childhood obesity

By Talha Rafiq, BPH., M.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2022)

Hamilton, Ontario (Medical Sciences)

TITLE: The Application of Metabolomics in Assessment of Nutrition, Sources of Variation in Food-related Metabolites, and Identification of -Omics Features of Childhood Obesity

AUTHOR: Talha Rafiq, BPH. (Brock University), M.Sc. (Brock University)

SUPERVISOR: Dr. Koon Teo

NUMBER OF PAGES: 313

Abstract

Ideally, a nutritional biomarker serves as an objective measure of the intake of a particular food or nutrient, may provide a reflection of health and disease processes, and can aid in the development of personalized nutritional recommendations. However, few food biomarkers have been validated and most have yet to be critically appraised in the literature. With the increased use of metabolomics in population-based studies, it is important to identify the sources of variability in nutritional biomarkers that may be attributed to intrinsic physiologic characteristics and extrinsic factors so that exposure-outcome associations can be examined more accurately. Additionally, circulating metabolites are associated with obesity-related changes in gut microbiome but there has been limited integration of metabolomics with microbiome in childhood obesity, and even less is known in non-white populations. This dissertation presents a series of studies that provide direct support for utility of nutritional biomarkers in population-based studies. The first study, presented in Chapter 2, contributes to the growing literature on food-based biomarkers by generating a comprehensive list of metabolites associated with a comprehensive list of all individual foods and food groups, and rated the evidence based on interstudy repeatability and study design. Chapter 3 identifies sources of variability in serum metabolite concentrations in White Europeans and South Asian pregnant women, thereby guiding appropriate statistical modeling when utilizing metabolomics in nutritional epidemiological studies. Chapter 4 provides results from a multi-omics integration analysis of serum metabolites and amplicon sequence variants of 16S ribosomal RNA genes to identify biomarkers that discriminate children with and

without obesity. Collectively, the results showed that a specific food/food group may give rise to many metabolites, however in several cases, a single metabolite can be a good indicator of food intake. Dietary factors explained the highest proportion of variability in exogenous food-based biomarkers relative to non-dietary factors, whereas the contribution of non-dietary factors was either similar or lower for metabolites that can either be produced endogenously, biotransformed by gut microbiota, and/or derived from more than one food source. Most of the circulating metabolites differed by ethnicity (South Asian and White Europeans). Biomarkers with good evidence can be considered direct surrogates for food intake, however, they can be influenced by several non-dietary factors, which require appropriate consideration during the statistical analyses of the data. Finally, the results showed notable differences in serum metabolome and specific gut bacterial species, and between specific metabolites and bacterial species related to childhood obesity. Obesity related metabolic pathways such as glutamate and carnitine metabolism may provide insight into the metabolic processes related to early onset of obesity in childhood.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Koon Teo for mentoring and supporting me in research and for his guidance in my personal growth. I am grateful to him for allowing me the freedom and platform to complete this research.

I would also like to thank Dr. Russel de Souza for his time, guidance, wisdom, patience, and encouragement over the course of this process. It has been a privilege to work with him since he is so dedicated to providing the environment required for his students to achieve their full potential which made him an excellent mentor and leader. I would also like to thank my committee members, Dr. Philip Britz-McKibbin, Dr. Lehana Thabane, Dr. Sonia Anand, and Dr. Katherine Morrison, for their invaluable contributions to all aspects of this research. Their efforts to give their time for reviewing my thesis and input for improvement will always be appreciated.

My thanks to Dr. Sandi Azab for her assistance with my critical review, sharing her knowledge of metabolomics, and being a wonderful role model. I am also thankful for Dr. Amel Lamri for sharing her knowledge of genetics and teaching me to code in R. To all students and colleagues whom I worked with, thank you for your encouragement and support.

Last, but not least, to my family and friends, I am forever grateful for the love and support – I cannot thank you enough.

Table of Contents

List of Figures

Supporting Figures

List of Tables

Supporting Tables

List of Abbreviations and Symbols

14:0	Myristic acid
15:0	Pentadecanoic acid
17:0	Heptadecanoic acid
ω-3 PUFA	Omega-3 polyunsaturated fatty acids
AAMU	5-acetamido-6-amino-3-methyluracil
AAAs	Aromatic amino acids
ADD	Average Danish Diet
ADMA	Asymmetric dimethylarginine
AIC	Akaike Information Criterion
ASV	Amplicon sequence variant
AUC	Area under the curve
BCAAs	Branched-chain amino acids
BFIs	Biomarkers of food intake
BiB	Born in Bradford
BIC	Bayesian Information Criterion
BMI	Body mass index
CE	Capillary electrophoresis
CI	Confidence intervals
CLR	Centered Log Ratio
CMPF	3-carboxy-4-methyl-5-propyl-2-furanproponante
DADA2	Divisive Amplicon Denoising Algorithm 2
DASH	Dietary Approaches to Stop Hypertension
DHA	Docosahexaenoic acid
DHBA	3,5-dihydroxybenzoic acid
DHPPA	3-(3,5-dihyroxyphenyl)-1-propanoic acid
DIABLO	Data Integration Analysis for Biomarker discovery using Latent cOmponent
DQS	Diet quality score
EPA	Eicosapentaenoic acid
FAMILY	Family Atherosclerosis Monitoring In earLy life
FFQ	Food frequency questionnaire
FooDB	The Food Database
FOODBALL	Food biomarker alliance
GABA	γ-aminobutyric acid
GC	Gas chromatography
GDM	Gestational diabetes mellitus
HEI	Healthy Eating Index
HLM	Hierarchical linear models
HMDB	Human Metabolome Database

HPLC	High performance liquid chromatography
IADPSG	International Association of the Diabetes and Pregnancy Study Groups
ICC	Intraclass correlation
IDO	Indoleamine-2,3-dioxygenase
IQR	Interquartile range
LC	Liquid chromatography
MeSH	Medical Subject Headings
MS	Mass spectrometry
MSI-CE-MS	Multisegment injection-capillary-electrophoresis-mass spectrometry
NAD+	Nicotinamide adenine dinucleotide
NEFA	Non-esterified fatty acids
NMR	Nuclear magnetic resonance
NND	New Nordic Diet
OLS	Ordinary Least Squares
OR	Odds ratios
PC-PR2	Principal component partial R-square
PCA	Principal component analysis
PLS	Partial least squares
PUFAs	Polyunsaturated fatty acids
QC	Quality control
RP	Reversed-phase
SCFAs	Short-chain fatty acids
SD	Standard deviation
SDI	Social disadvantage index
SDMA	Symmetric dimethylarginine
SFA	Saturated fatty acids
SFN-NAC	Sulforaphane N-acetylcysteine
sGCCA	Sparse Generalized Canonical Correlation Analysis
sPLS-DA	Sparse-PLS-discriminant analysis
SHARE	Study of Health and Risk in Ethnic Groups
START	SouTh Asian biRth cohorT
TMAO	Trimethylamine-N-oxide
TOF-MS	Time-of-flight mass spectrometer
UPLC	Ultra-high performance liquid chromatography
WHO	World Health Organization

Declaration of Academic Achievement

The following material has been previously published and is reprinted with written permission:

Chapter 2 – Rafiq, T., Azab, S. M., Teo, K. K., Thabane, L., Anand, S. S., Morrison, K. M., de Souza, R. J., & Britz-McKibbin, P. (2021). Nutritional Metabolomics and the Classification of Dietary Biomarker Candidates: A Critical Review. Advances in nutrition (Bethesda, Md.), 12(6), 2333–2357. https://doi.org/10.1093/advances/nmab054

Chapter 3 – Rafiq, T., Azab, S. M., Anand, S. S., Thabane, L., Shanmuganathan, M., Morrison, K. M., Atkinson, S. A., Stearns, J. C., Teo, K. K., Britz-McKibbin, P., & de Souza, R. J. (2022). Sources of Variation in Food-Related Metabolites during Pregnancy. Nutrients, 14(12), 2503. https://doi.org/10.3390/nu14122503

Chapter 1 – Introduction

1.1 Overview of metabolomics

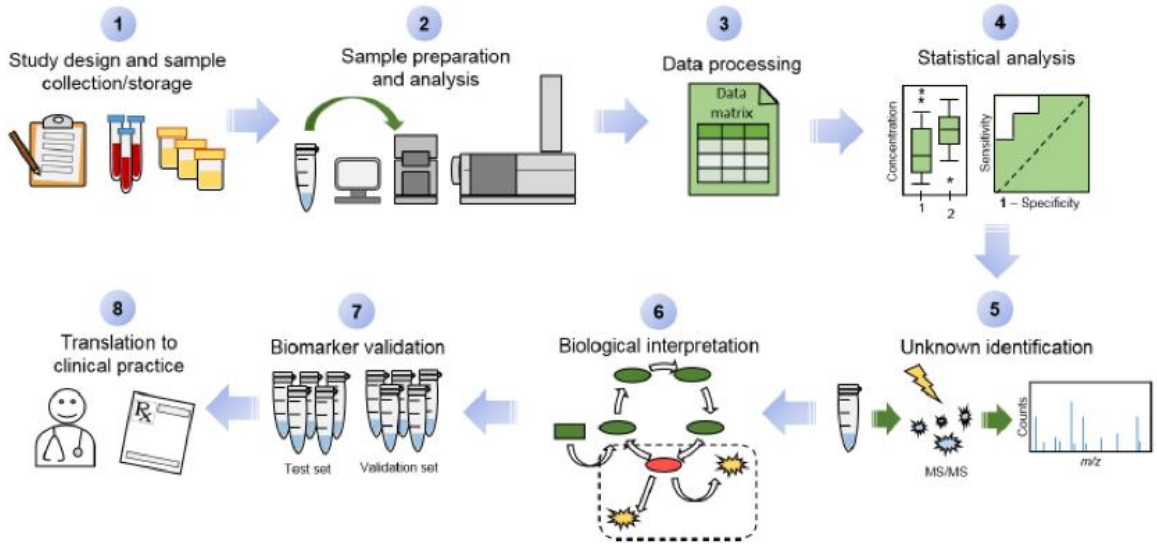
Metabolomics is one of the '-omics' approaches that is used to characterize, identify, and quantify small molecules (metabolites) in a biological system (1). Metabolites are low-molecular weight (up to 1000 to 1500 Dalton) chemical substrates, intermediates, or end products of biochemical reactions. A set of metabolites in an organism constitutes the metabolome (2). The metabolome provides a snapshot of the active metabolic processes which are influenced by internal factors including the genome, epigenome, transcriptome, and proteome, and external factors such as lifestyle and health behaviours including diet, gut microbiome, and environmental factors (1). The human metabolome is estimated to be comprised of thousands of metabolites, with the latest report in the Human Metabolome Database (HMDB) from December 2020 listing approximately 8,000 endogenous metabolites (3), and 35,000 exogenous metabolites from sources of diet, medications, microbes, and environmental exposures (3). The application of metabolomics has increased exponentially in the past decade and is now routinely applied to identify correlative or predictive markers for various human diseases. Metabolome analysis can be a useful approach for not only the identification of biomarkers that can be used for diagnosis, treatment, and prevention of diseases, but also for providing novel insights into the molecular mechanisms and to generate hypotheses for future research.

Metabolites have diverse chemical properties, thus, no single analytical approach is capable of measuring all metabolites (2). There are two fundamentally different approaches used for metabolite profiling: targeted and untargeted. The untargeted

approach is used to measure as many known and unknown metabolites as possible, often in a non-quantitative manner (2). The main limitation of this approach is the difficulty with identification of unknown metabolites (4). In contrast, the targeted approach provides absolute concentrations of a fixed number of known metabolites using authentic chemical standards and calibration curves for each metabolite (2). The targeted metabolomics approach is more accurate and precise than the non-targeted approach, but is more labor intensive and biased towards metabolites for which the appropriate internal standards are available for quantification (2).

There are several analytical platforms that can applied for both targeted and untargeted approaches, with mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR) being the two most commonly used (5). MS measures metabolites based on their mass-to-charge ratios (m/z) (6). MS has higher sensitivity and lower limits of detection, higher resolution and accuracy, is relatively faster and requires smaller sample volumes compared to NMR (2). However, the samples can come in direct contact with the detector so there is a potential risk of contamination, thus quality control samples are required. MS is often combined with pre-fractionation methods such as gas chromatography (GC), liquid chromatography (LC), and capillary electrophoresis (CE) to reduce sample complexities (2, 7). GC provide a high resolution and sensitivity and is the most suitable technique for quantifying volatile and semi-volatile compounds (8). However, this technique requires intensive pre-processing and longer run times. While the traditional LC does not have a very high resolution, the development of high and ultra-high performance liquid chromatography (HPLC and UPLC, respectively) has

improved separation sensitivity and lowered run times, thereby making these methods highly efficient for large-scale studies (9). NMR applies a magnetic field to an atomic nucleus and uses radiofrequency waves to characterize the resonant frequency of the atomic nucleus in relation to its chemical or environmental surroundings (10, 11). Unlike MS, NMR offers advantages for metabolites that are difficult to ionize or require derivatization (i.e., requires no separation and derivatization), and this method is faster than GC-MS and LC-MS (12). NMR is inherently robust and reliable as it provides



quantitative and structural information for the identification of compounds (12). Moreover, it is nondestructive so the samples can be reused for future analyses (13). The limitation of NMR is its relatively low sensitivity compared to MS, making it unsuitable to detect metabolites that are present in low abundance (2). Generally, the metabolomics workflow consists of a series of independent steps for identification of metabolites as shown in Figure 1.1.

Figure 1.1: Overview of the data workflow in metabolomics for identification of biomarkers relevant to human diseases in clinical medicine.
Retrieved from: https://britz.mcmaster.ca/research/directed-metabolite-analyses

1.2 Metabolomics in nutrition

Objective, valid, and reliable assessment tools are needed to measure dietary exposure in order to gain a better understanding of the causal links between nutrition and health outcomes (14). However, measuring food intake using methods that are both accurate and applicable to free-living individuals remains a challenge (15). Dietary intake in epidemiological studies is traditionally assessed using self-reported, and often memory-based approaches including dietary records, weighted food diaries, 24-hour dietary recalls, and food frequency questionnaires (FFQs) (15). The reliability and validity of these tools for measuring dietary intake have been questioned as they are prone to serious systematic errors (16, 17). With the exception of prospectively-collected weighted diet records, these methods rely on participants' memory and are therefore prone to over- and under-reporting of certain foods (18). Additionally, day-to-day variability in dietary intake and estimation of portion size makes it difficult to accurately measure intake of food and nutrients (19). To overcome these limitations, metabolomics has been increasingly applied in large-scale epidemiological studies to identify and validate biomarkers of food intake (20, 21). Biomarkers offer a complement (or, in some cases, an alternative) to self-report tools for a more objective assessment of food exposures because they account for nutrient bioavailability and metabolism. Although studies using FFQs have shown to provide adequate precision to distinguish high from low consumers

of foods and foods varying across populations, studies using metabolomics better characterize dose–response relations (1).

The ‘food metabolome’ has traditionally been defined as “the sum of all metabolites directly derived from the digestion of foods, their absorption in the gut, and biotransformation by the host tissues and the microbiota” (22). More recently, others have defined it as the whole set of food constituents in any food (1, 23). Although the application of metabolomics in nutrition has a relatively long history in identification of food biomarkers related to known constituents of food chemistry (using a targeted approach), the application of untargeted metabolomics only started to gain prominence in the mid-2000s and has greatly expanded in the last decade. Only a few biomarkers of foods have been validated to this date. For example, proline betaine has been validated in a large-scale observational study, where it was highly sensitive (86.3%) and specific (90.6%) for citrus fruit consumption and it has been shown to be minimally metabolized in the body (Figure 1.2) (24).

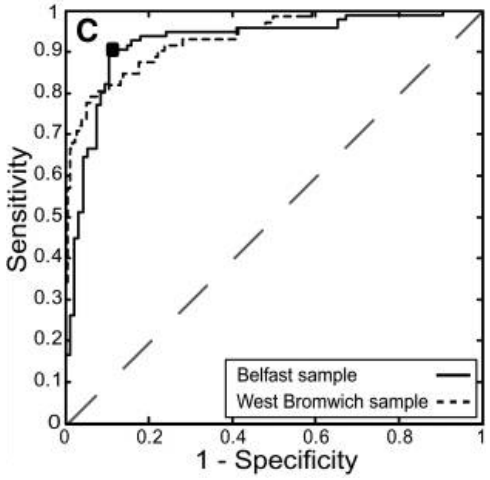


Figure 1.2: Receiver operating characteristic curves to assess the predictive ability of excretion of proline betaine for discrimination of citrus fruit intake and no citrus fruit intake

Adapted from Heinzmann et al (24).

In contrast, other metabolites that may be derived from a variety of precursors (e.g., trimethylamine *N*-oxide (TMAO)) or the occurrence of the same precursor in various foods (e.g., carnitine) are less robust in populations (Figure 1.3) (25).

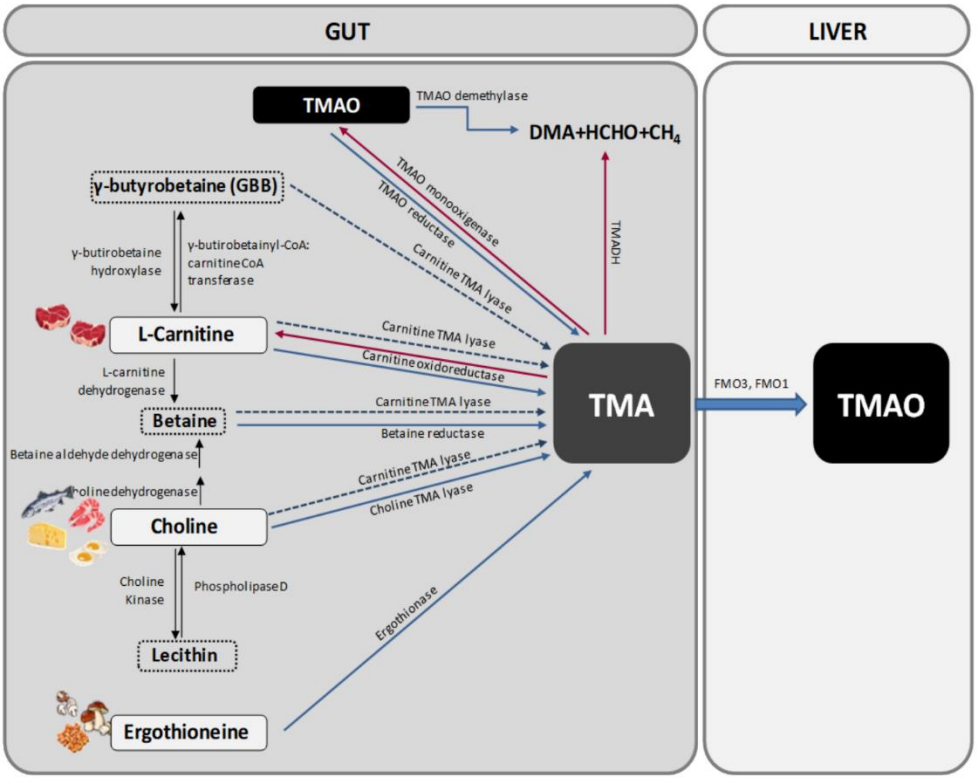


Figure 1.3: Pathways for trimethylamine N-oxide (TMAO) formation from dietary sources

Adapted from Janeiro et al (26).

For this reason, it is important to consider the characteristic of the food biomarker. Identifying robust biomarkers will depend on the sensitivity of the analytical technique used and the quality of dietary data against which the metabolites are correlated (21). Additionally, the chance of identifying biomarkers is limited if there is lower accuracy and lower number of foods documented in food questionnaires (21). The food metabolome is highly complex, with a varying composition based on the diet, and because of its responsiveness to change in dietary intake, has the potential to measure dietary exposure with a high level of detail and precision.

1.3 Sources of variability in metabolite concentrations

The ideal food biomarker is one that can be readily measured in the biological sample of interest at the population level, is highly specific for one food item or food group, shows a distinct dose- and time-dependent response, and is ideally neither generated in vivo nor extensively transformed by the microbiota and host tissue upon consumption. However, food matrices (i.e., the nutrient and non-nutrient components of foods) are complex since nutrients can be derived from various food sources rather than exclusively originate from a single food or nutrient and can show intercorrelation between metabolic processes (27). Further, the human metabolome can exhibit variation due to intrinsic physiologic characteristics, such as age, sex, ethnicity, hormonal levels, and the gut microbiome, as well as extrinsic factors such as dietary intake, drug use, lifestyle, and psychological stress. For example, TMAO is a naturally occurring small organic dietary compound that is produced by gut bacteria from choline, betaine, and carnitine and is abundant in diets rich in TMA precursors such as fish, beef, and eggs

(25). Further, many food-specific metabolites such as dietary quercetin-3-rutinoside or catechin are absorbed from the colon and are further metabolized in the liver at different rates (28), which can contribute to greater variability in metabolite concentration depending on the enzyme activity. Gut microbes influence the metabolism of food derived nutrients such as polyphenols. Hippuric acid, which is largely derived from vegetables, is derived from polyphenol contents of food via gut microbiota producing phenylpropionic acids (29, 30).

There is evidence of sex differences in the human metabolome with one study identifying one-third of the metabolites measured in serum to be significantly different between males and females. Even among females, differences in metabolome have been observed between pregnant and non-pregnant women (31, 32). Several studies have used metabolomic approaches to better understand physiological changes associated with pregnancy and to examine metabolomic differences between pregnant and non-pregnant women as well as between women with normal pregnancy and those with pregnancy complications including pre-eclampsia, fetal growth restriction, and preterm delivery among other outcomes (33-37). Although studies have suggested the application of metabolomics to understand the underlying mechanisms and prognostic and predictive value of biomarkers in relation to pregnancy complications, they have not considered the influence of sociodemographic, environmental, and lifestyle factors when examining the association between metabolites and outcomes (38, 39). For example, a recent systematic review examining biomarkers associated with preterm birth noted that only three of the 14 included studies statistically adjusted for confounding factors such as maternal age

and maternal weight or body mass index (BMI) even though all studies collected these data (39). To prevent bias and to accurately assess metabolite-health associations, it is essential to identify factors beyond dietary intake that may contribute to variability in metabolite concentrations in order to appropriately adjust for them in the analysis.

Ethnicity is another important factor influencing, the metabolite concentrations and the susceptibility, incidence, and response to risk factors and health conditions such as obesity and type 2 diabetes (40, 41). In addition to genetics, differences in metabolic rate, dietary intake, socioeconomic and lifestyle factors, and access to health care may partly explain the ethnic disparities in health outcomes (42-44). It is therefore important to account for ethnicity when examining sources of variation in metabolite concentrations and when utilizing metabolomics to understand disease pathologies (45, 46). It is also not well known whether non-dietary factors contributing to variability in metabolite concentrations are similar or unique among different ethnic groups. Therefore, the purpose of the second study in this dissertation is to determine the extent to which non-dietary factors explain the variability in the concentrations of the putative biomarkers of food intake among White and South Asian populations.

It is important to identify the non-dietary sources of food biomarkers and determine to what extent these factors explain differences in metabolite concentration. Understanding these sources of variation will impact the utility of biomarkers of food intake and help advance the field by identifying robust and reproducible biomarkers that can be used routinely in public health and clinical research. The nutrients and macromolecules present in the habitual diet are metabolized and excreted, so they will

have some influence on metabolic profiles. Thus, utilizing a holistic approach, metabolomics, could prove invaluable for identification, annotation, and characterization of metabolic signatures associated with health outcomes.

1.4 Obesity

Obesity is a multifactorial process with complex etiology and has become a significant public health concern reaching ‘epidemic’ status worldwide (47). The prevalence of childhood obesity has progressively increased globally in the past few decades (48). According to the World Health Organization (WHO), the prevalence of childhood obesity has increased by 33% between 1990 and 2016, and if this trend continues, an estimated 70 million children worldwide will be overweight or obese by 2025 (49). The most rapid increase in weight gain among children occurs between the ages of 2 and 6 years, and 90% of children obese at 3 years of age present overweight or obesity during adolescence (50). Although childhood obesity is a growing problem worldwide, its prevalence is higher in nonwhite populations (51). Childhood obesity is a risk factor for future cardiovascular and metabolic complications such as dyslipidemia and type 2 diabetes, which become more prevalent with increasing age (52). In fact, obesity is the most common comorbidity of type 2 diabetes in children (53, 54). In recent decade, the increasing prevalence of type 2 diabetes in children is mainly due to the increase in obesity rates observed in children (55, 56). Over 85% of children with type 2 diabetes are overweight or obese at the time of diagnosis (54). On a societal level, childhood obesity has been shown to have significant social and economic consequences. The annual cost of childhood obesity in Canada is estimated to be $22 billion in lost productivity and

health-care costs (57). Additionally, the total lifetime cost attributable to overweight/obesity are estimated to be five times higher in females and three times higher in males with a history of childhood obesity compared to children with normal weight (58).

Factors that likely contribute to the disparities in childhood obesity are many, involving genetics, physiology, socioeconomic status, culture, environment, and the interactions between them; but also other factors that have not yet been sufficiently researched such as the gut microbiome (43). The origin of obesity involves an imbalance between energy intake and energy expenditure resulting from the complex interplay between genetic, behavioural, social and environmental factors (59, 60). Adding to that is the variability in the composition of the gut microbiota, which may contribute to nutrient acquisition, energy regulation, and adipose storage (61). More recently, metabolomics has attracted great interest in obesity research because metabolic profiling can be used to detect subtle changes in metabolic networks and provide predictive biomarkers or biomarker patterns relevant to the biological mechanism of childhood obesity. Lifestyle interventions are effective but not sufficient in addressing childhood obesity (62). A meta-analysis of randomized controlled trials examining the efficiency of behavioral family lifestyle interventions for childhood obesity found a small-to-moderate effect (standardized effect size of 0.47) of interventions on improving weight outcomes (63). Therefore, in addition to lifestyle interventions, identifying novel risk factors is important to provide insight into the etiopathogenesis of childhood obesity.

1.5 Obesity and Metabolomics

Metabolic dysregulation is closely related to an individual's predisposition to develop low-grade inflammation and oxidative stress, both of which contribute to pathogenesis and progression of obesity and associated downstream consequences such as the metabolic syndrome, type 2 diabetes and cardiovascular disease (64). Several metabolic traits (for example fasting blood glucose, triglycerides, and cholesterol levels) have long been used as biomarkers of obesity-related phenotypes (65). Using metabolites as biomarkers to understand pathogenesis and progression of obesity is useful because their levels are likely regulated by genetic and environmental factors, therefore they may be more so related to obesity in comparison to these factors (genetic and environmental) alone.

Metabolomic profiling of children with obesity has revealed several metabolic signatures. A recent systematic review of 41 studies that reported metabolic profiles of blood and urine samples of children with obesity found that the most commonly and consistently reported metabolites are branched-chain amino acids (BCAAs), aromatic amino acids (AAAs), short chain acylcarnitines (more commonly free carnitine and acetylcarnitine), lipids, and steroids (66). However, most of the studies included in this review were cross-sectional in design, which severely limits causal inference. Many studies used a targeted approach (i.e., assessing a list of known metabolites), which further limits the scope of their findings by possibly missing novel metabolites of obesity. Nevertheless, several biomarkers of childhood obesity reported in cross-sectional studies

may likely reflect obesity because they have been identified in interventional studies for weight loss and a causal study in young adults (67, 68).

There is evidence to suggest that some metabolites such as BCAAs and metabolites of tryptophan metabolism may be predictors of childhood obesity (69-72). However, whether they are markers or direct contributors to obesity, remains inconclusive. For example, BCAAs reflect dietary proteins and have frequently been linked with obesity (73), but paradoxically, a BCAA-rich diet improves metabolic health including satiety and regulation of body weight (74, 75). Further, metabolomic analyses can simultaneously assess metabolic changes related to diet, products of microbial metabolism, and physiological changes (76). For instance, the association between p-cresol sulfate (L-tyrosine-derived microbial metabolite) and childhood obesity is inconsistent, likely because p-cresol sulfate can be affected by both tyrosine levels and altered microbial composition (77, 78). However, only a few studies have discovered metabolites (e.g., metabolites related to tryptophan metabolism and BCAAs) as determinants by metabolome at birth or during infancy and associated them BMI later in childhood (79, 80).

There is ample evidence to support the role of endogenously produced or derived metabolites in response to host diet and biotransformed by gut microbiota, which may influence host metabolic processes (energy gain or inflammation) related to obesity (66). For this reason, there is a growing interest in the integration of metabolomics with other omics-based approaches to better characterize molecular changes of obesity necessary to gain a better understanding of this pathology. Therefore, integrative multi-omics analysis

of microbial and metabolomics data in a prospective cohort design has potential for identifying microbial influence on host physiology through production or modification of bioactive metabolites to gain a more complete understanding of the molecular changes that contribute to childhood obesity.

Multi-omics aims to integrate two or more -omics datasets and offers a more system-based approach to explore molecular processes from many different perspectives. More recently, mixOmics provides a set of supervised and unsupervised multivariate methods for the integration of -omics datasets with a particular focus on variable selection. The package allows integration of multi-omics datasets on the same individuals (vertical integration) or across studies on the same variables (horizontal integration) to classify or cluster samples. Data Integration Analysis for Biomarker discovery using Latent cOmponent (DIABLO), a framework in mixOmics package, extends generalized canonical correlation analysis to identify correlated multi-omics features to discriminate subtypes of an outcome variable (81, 82). DIABLO is a multivariate dimensionality reduction method that selects correlated variables by maximizing the covariance between the linear combination of variables (latent component scores) from each omics dataset (Figure 1.4).

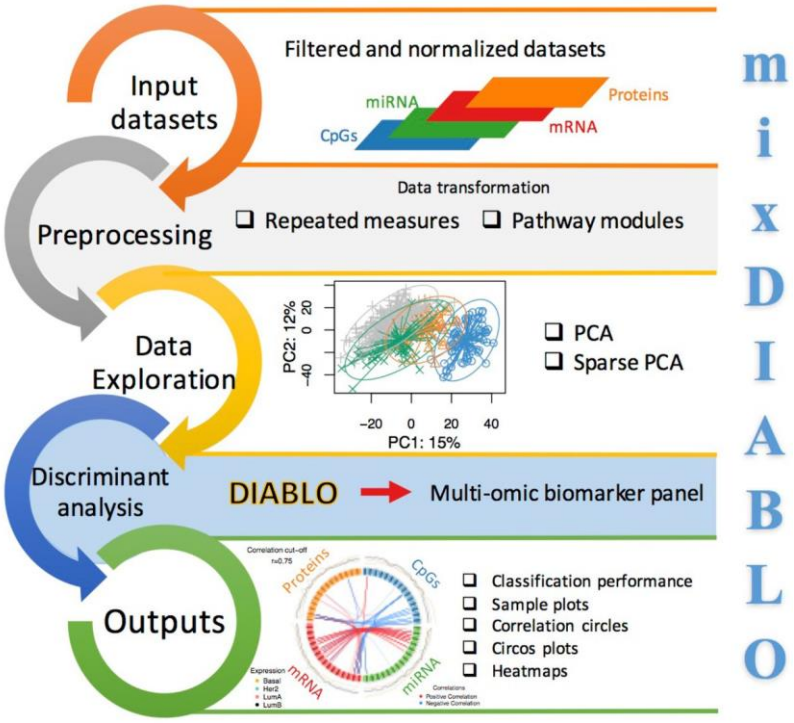


Figure 1.4: A framework for multi-omics data integration and identification of molecular features using DIABLO

Adapted from Singh et al (83).

1.6 Rationale

Given the emergent application of metabolomics in food science and its potential for global metabolic assessment, discovery of food derived biomarkers, and monitoring the progression of diet-related metabolic diseases such as obesity, it is inevitable that more researchers will adopt metabolomics in future research. While there are a few metabolomics-derived food composition databases that serves as a resource for the scientific community, it remains a challenge to critically appraise and classify robust dietary biomarkers in a rapidly evolving field (84). Therefore, it is important to generate a comprehensive list of food associated metabolites and rate the evidence of these

biomarkers. It is also important to recognize that although the use of food metabolome is more objective, it exhibits variability due to the human intrinsic physiologic characteristics and extrinsic factors. Understanding the sources of variation in biomarkers of food intake that are not attributed to changes in food intake are critical to advancing the application/field of food intake biomarkers, because they will further reduce exposure misclassification. Finally, the focus of nutritional biomarkers has largely been on the discovery of specific metabolites associated with food consumption and its impact on chronic disease risk. It is highly relevant to identify biomarkers to characterize changes in metabolic profile of children with obesity in very early stages of life to prevent future chronic diseases.

1.7 Aims and Objectives

The overall aim of this dissertation was to contribute to the development of metabolomics in the field of nutrition and identify alterations in the metabolic environment using metabolome and microbiome during the early developmental stage of life to those who are at risk of developing obesity in childhood. The specific objectives of each study are listed below:

Study 1 (Chapter 2):

1) Generate a comprehensive list of metabolites associated with individual food and food groups in apparently healthy individuals
2) Report on the study designs, metabolomic approaches, and biospecimen used.
3) Using a scoring system, rate the empirical evidence of metabolites as candidate biomarker of food intake based on interstudy repeatability and study design.

Study 2 (Chapter 3):

1) Examine the associations of non-dietary factors, including demographics, lifestyle, and pregnancy-related factors with serum metabolite concentrations using a panel of commonly identified biomarkers derived from food intake and/or gut microbiota in pregnant women of two ethnically diverse groups.
2) Determine the extent to which non-dietary factors explain the variability in the concentrations of the putative biomarkers of food intake.

Study 3 (Chapter 4):

1) Employ multi-omics approach of 16S rRNA gene amplicon sequence variant (ASV) and serum metabolome data to identify integrated molecular features that characterize risk of obesity in children.

1.8 References

1. Wishart DS. Metabolomics: applications to food science and nutrition research. Trends in Food Science & Technology 2008;19(9):482-93. doi: https://doi.org/10.1016/j.tifs.2008.03.003.
2. Dunn WB, Broadhurst DI, Atherton HJ, Goodacre R, Griffin JL. Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. Chem Soc Rev 2011;40(1):387-426. doi: 10.1039/b906712b.
3. Wishart DS, Guo A, Oler E, Wang F, Anjum A, Peters H, Dizon R, Sayeeda Z, Tian S, Lee Brian L, et al. HMDB 5.0: the Human Metabolome Database for 2022. Nucleic Acids Research 2021;50(D1):D622-D31. doi: 10.1093/nar/gkab1062.
4. Scalbert A, Brennan L, Fiehn O, Hankemeier T, Kristal BS, van Ommen B, Pujos-Guillot E, Verheij E, Wishart D, Wopereis S. Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. Metabolomics 2009;5(4):435. doi: 10.1007/s11306-009-0168-0.
5. Letertre MPM, Giraudeau P, de Tullio P. Nuclear Magnetic Resonance Spectroscopy in Clinical Metabolomics and Personalized Medicine: Current Challenges and Perspectives. Front Mol Biosci 2021;8:698337. doi: 10.3389/fmolb.2021.698337.
6. Xiao JF, Zhou B, Ressom HW. Metabolite identification and quantitation in LC-MS/MS-based metabolomics. Trends Analyt Chem 2012;32:1-14. doi: 10.1016/j.trac.2011.08.009.
7. Lindon JC, Nicholson JK, Holmes E. Chromatographic and electrophoretic separations combined with mass spectrometry for metabonomics. The Handbook of Metabonomics and Metabolomics 2011;4:149.
8. Fiehn O. Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling. Curr Protoc Mol Biol 2016;114:30.4.1-.4.2. doi: 10.1002/0471142727.mb3004s114.
9. Wales TE, Fadgen KE, Gerhardt GC, Engen JR. High-speed and high-resolution UPLC separation at zero degrees Celsius. Anal Chem 2008;80(17):6815-20. doi: 10.1021/ac8008862.
10. Trock BJ. Application of metabolomics to prostate cancer. Urol Oncol 2011;29(5):572-81. doi: 10.1016/j.urolonc.2011.08.002.
11. Ross A, Schlotterbeck G, Dieterle F, Senn H. NMR spectroscopy techniques for application to metabonomics. The handbook of metabonomics and metabolomics 2007:55-112.
12. Markley JL, Brüschweiler R, Edison AS, Eghbalnia HR, Powers R, Raftery D, Wishart DS. The future of NMR-based metabolomics. Current Opinion in Biotechnology 2017;43:34-40. doi: https://doi.org/10.1016/j.copbio.2016.08.001.

13. Emwas AH, Roy R, McKay RT, Tenori L, Saccenti E, Gowda GAN, Raftery D, Alahmari F, Jaremko L, Jaremko M, et al. NMR Spectroscopy for Metabolomics Research. Metabolites 2019;9(7). doi: 10.3390/metabo9070123.
14. Bingham SA. Biomarkers in nutritional epidemiology. Public Health Nutr 2002;5(6a):821-7. doi: 10.1079/phn2002368.
15. Shim JS, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health 2014;36:e2014009. doi: 10.4178/epih/e2014009.
16. Freedman LS, Potischman N, Kipnis V, Midthune D, Schatzkin A, Thompson FE, Troiano RP, Prentice R, Patterson R, Carroll R, et al. A comparison of two dietary instruments for evaluating the fat-breast cancer relationship. Int J Epidemiol 2006;35(4):1011-21. doi: 10.1093/ije/dyl085.
17. Kipnis V, Midthune D, Freedman L, Bingham S, Day NE, Riboli E, Ferrari P, Carroll RJ. Bias in dietary-report instruments and its implications for nutritional epidemiology. Public Health Nutr 2002;5(6A):915-23. doi: 10.1079/phn2002383.
18. Macdiarmid J, Blundell J. Assessing dietary intake: Who, what and why of under-reporting. Nutr Res Rev 1998;11(2):231-53. doi: 10.1079/nrr19980017.
19. Young LR, Nestle MS. Portion sizes in dietary assessment: issues and policy implications. Nutr Rev 1995;53(6):149-58. doi: 10.1111/j.1753-4887.1995.tb01542.x.
20. Llorach R, Garcia-Aloy M, Tulipani S, Vazquez-Fresno R, Andres-Lacueva C. Nutrimetabolomic strategies to develop new biomarkers of intake and health effects. J Agric Food Chem 2012;60(36):8797-808. doi: 10.1021/jf301142b.
21. Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, Rappaport SM, van der Hooft JJ, Wishart DS. The food metabolome: a window over dietary exposure. Am J Clin Nutr 2014;99(6):1286-308. doi: 10.3945/ajcn.113.076133.
22. Fardet A, Llorach R, Orsoni A, Martin JF, Pujos-Guillot E, Lapierre C, Scalbert A. Metabolomics provide new insight on the metabolism of dietary phytochemicals in rats. J Nutr 2008;138(7):1282-7. doi: 10.1093/jn/138.7.1282.
23. Cevallos-Cevallos JM, Reyes-De-Corcuera JI, Etxeberria E, Danyluk MD, Rodrick GE. Metabolomic analysis in food science: a review. Trends in Food Science & Technology 2009;20(11-12):557-66.
24. Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas ME, Kochhar S, Stamler J, Holmes E, Elliott P, Nicholson JK. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. Am J Clin Nutr 2010;92(2):436-43. doi: 10.3945/ajcn.2010.29672.
25. Cuparencu C, Praticó G, Hemeryck LY, Sri Harsha PSC, Noerman S, Rombouts C, Xi M, Vanhaecke L, Hanhineva K, Brennan L, et al. Biomarkers of meat and seafood intake: an extensive literature review. Genes & nutrition 2019;14:35. doi: 10.1186/s12263-019-0656-4.
26. Janeiro MH, Ramírez MJ, Milagro FI, Martínez JA, Solas M. Implication of Trimethylamine N-Oxide (TMAO) in Disease: Potential Biomarker or New Therapeutic Target. Nutrients 2018;10(10):1398. doi: 10.3390/nu10101398.

27. Noh H, Freisling H, Assi N, Zamora-Ros R, Achaintre D, Affret A, Mancini F, Boutron-Ruault M-C, Flogel A, Boeing H, et al. Identification of Urinary Polyphenol Metabolite Patterns Associated with Polyphenol-Rich Food Intake in Adults from Four European Countries. Nutrients 2017;9(8):796. doi: https://dx.doi.org/10.3390/nu9080796.
28. Hollman PCH. Absorption, Bioavailability, and Metabolism of Flavonoids. Pharmaceutical Biology 2004;42(sup1):74-83. doi: 10.3109/13880200490893492.
29. Rechner AR, Kuhnle G, Bremner P, Hubbard GP, Moore KP, Rice-Evans CA. The metabolic fate of dietary polyphenols in humans. Free Radic Biol Med 2002;33(2):220-35. doi: 10.1016/s0891-5849(02)00877-8.
30. Walsh MC, Brennan L, Pujos-Guillot E, Sébédio JL, Scalbert A, Fagan A, Higgins DG, Gibney MJ. Influence of acute phytochemical intake on human urinary metabolomic profiles. Am J Clin Nutr 2007;86(6):1687-93. doi: 10.1093/ajcn/86.5.1687.
31. Krumsiek J, Mittelstrass K, Do KT, Stückler F, Ried J, Adamski J, Peters A, Illig T, Kronenberg F, Friedrich N, et al. Gender-specific pathway differences in the human serum metabolome. Metabolomics 2015;11(6):1815-33. doi: 10.1007/s11306-015-0829-0.
32. Costanzo M, Caterino M, Sotgiu G, Ruoppolo M, Franconi F, Campesi I. Sex differences in the human metabolome. Biol Sex Differ 2022;13(1):30. doi: 10.1186/s13293-022-00440-4.
33. Monni G, Atzori L, Corda V, Dessolis F, Iuculano A, Hurt KJ, Murgia F. Metabolomics in Prenatal Medicine: A Review. Front Med (Lausanne) 2021;8:645118. doi: 10.3389/fmed.2021.645118.
34. Dunn WB, Brown M, Worton SA, Davies K, Jones RL, Kell DB, Heazell AEP. The metabolome of human placental tissue: investigation of first trimester tissue and changes related to preeclampsia in late pregnancy. Metabolomics 2012;8(4):579-97. doi: 10.1007/s11306-011-0348-6.
35. Auray-Blais C, Raiche E, Gagnon R, Berthiaume M, Pasquier J-C. Metabolomics and preterm birth: What biomarkers in cervicovaginal secretions are predictive of high-risk pregnant women? International Journal of Mass Spectrometry 2011;307(1):33-8. doi: https://doi.org/10.1016/j.ijms.2011.02.009.
36. Dessì A, Atzori L, Noto A, Visser GH, Gazzolo D, Zanardo V, Barberini L, Puddu M, Ottonello G, Atzei A, et al. Metabolomics in newborns with intrauterine growth retardation (IUGR): urine reveals markers of metabolic syndrome. J Matern Fetal Neonatal Med 2011;24 Suppl 2:35-9. doi: 10.3109/14767058.2011.605868.
37. Handelman SK, Romero R, Tarca AL, Pacora P, Ingram B, Maymon E, Chaiworapongsa T, Hassan SS, Erez O. The plasma metabolome of women in early pregnancy differs from that of non-pregnant women. PLoS One 2019;14(11):e0224682. doi: 10.1371/journal.pone.0224682.
38. Clinton CM, Bain JR, Muehlbauer MJ, Li Y, Li L, O'Neal SK, Hughes BL, Cantonwine DE, McElrath TF, Ferguson KK. Non-targeted urinary metabolomics

in pregnancy and associations with fetal growth restriction. Sci Rep 2020;10(1):5307. doi: 10.1038/s41598-020-62131-7.
39. Carter RA, Pan K, Harville EW, McRitchie S, Sumner S. Metabolomics to reveal biomarkers and pathways of preterm birth: a systematic review and epidemiologic perspective. Metabolomics 2019;15(9):124. doi: 10.1007/s11306-019-1587-1.
40. Hill-Briggs F, Adler NE, Berkowitz SA, Chin MH, Gary-Webb TL, Navas-Acien A, Thornton PL, Haire-Joshu D. Social Determinants of Health and Diabetes: A Scientific Review. Diabetes Care 2020;44(1):258-79. doi: 10.2337/dci20-0053.
41. Park SH, Nicolaou M, Dickens BSL, Yang Q, Tan KW, van Dam RM. Ethnicity, Neighborhood and Individual Socioeconomic Status, and Obesity: The Singapore Multiethnic Cohort. Obesity (Silver Spring) 2020;28(12):2405-13. doi: 10.1002/oby.22995.
42. Manuel JI. Racial/Ethnic and Gender Disparities in Health Care Use and Access. Health Serv Res 2018;53(3):1407-29. doi: 10.1111/1475-6773.12705.
43. Caprio S, Daniels SR, Drewnowski A, Kaufman FR, Palinkas LA, Rosenbloom AL, Schwimmer JB. Influence of race, ethnicity, and culture on childhood obesity: implications for prevention and treatment: a consensus statement of Shaping America's Health and the Obesity Society. Diabetes Care 2008;31(11):2211-21. doi: 10.2337/dc08-9024.
44. Deboer MD. Ethnicity, obesity and the metabolic syndrome: implications on assessing risk and targeting intervention. Expert Rev Endocrinol Metab 2011;6(2):279-89. doi: 10.1586/eem.11.17.
45. Golden SH, Brown A, Cauley JA, Chin MH, Gary-Webb TL, Kim C, Sosa JA, Sumner AE, Anton B. Health disparities in endocrine disorders: biological, clinical, and nonclinical factors--an Endocrine Society scientific statement. J Clin Endocrinol Metab 2012;97(9):E1579-639. doi: 10.1210/jc.2012-2043.
46. Vasishta S, Ganesh K, Umakanth S, Joshi MB. Ethnic disparities attributed to the manifestation in and response to type 2 diabetes: insights from metabolomics. Metabolomics 2022;18(7):45. doi: 10.1007/s11306-022-01905-8.
47. OECD. The Heavy Burden of Obesity, 2019.
48. Sahoo K, Sahoo B, Choudhury AK, Sofi NY, Kumar R, Bhadoria AS. Childhood obesity: causes and consequences. J Family Med Prim Care 2015;4(2):187-92. doi: 10.4103/2249-4863.154628.
49. World Health Organization. Report of the commission on ending childhood obesity: World Health Organization, 2016.
50. Geserick M, Vogel M, Gausche R, Lipek T, Spielau U, Keller E, Pfäffle R, Kiess W, Körner A. Acceleration of BMI in Early Childhood and Risk of Sustained Obesity. New England Journal of Medicine 2018;379(14):1303-12. doi: 10.1056/NEJMoa1803527.
51. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, Mullany EC, Biryukov S, Abbafati C, Abera SF, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014;384(9945):766-81. doi: 10.1016/s0140-6736(14)60460-8.

52. Turta O, Rautava S. Antibiotics, obesity and the link to microbes - what are we doing to our children? BMC Med 2016;14:57. doi: 10.1186/s12916-016-0605-7.
53. Panagiotopoulos C, Hadjiyannakis S, Henderson M. Type 2 Diabetes in Children and Adolescents. Can J Diabetes 2018;42 Suppl 1:S247-s54. doi: 10.1016/j.jcjd.2017.10.037.
54. Pulgaron ER, Delamater AM. Obesity and type 2 diabetes in children: epidemiology and treatment. Curr Diab Rep 2014;14(8):508. doi: 10.1007/s11892-014-0508-y.
55. Valaiyapathi B, Gower B, Ashraf AP. Pathophysiology of Type 2 Diabetes in Children and Adolescents. Curr Diabetes Rev 2020;16(3):220-9. doi: 10.2174/1573399814666180608074510.
56. Kao KT, Sabin MA. Type 2 diabetes mellitus in children and adolescents. Aust Fam Physician 2016;45(6):401-6.
57. O'Brien Institute Staff. Internet: https://ucalgary.ca/news/ignoring-needs-canadian-kids-costing-economy-billions-analysis.
58. Sonntag D. Why Early Prevention of Childhood Obesity Is More Than a Medical Concern: A Health Economic Approach. Ann Nutr Metab 2017;70(3):175-8. doi: 10.1159/000456554.
59. Canoy D, Bundred P. Obesity in children. BMJ Clin Evid 2011;2011.
60. Hill JO, Wyatt HR, Peters JC. Energy balance and obesity. Circulation 2012;126(1):126-32. doi: 10.1161/circulationaha.111.087213.
61. Heiss CN, Olofsson LE. Gut Microbiota-Dependent Modulation of Energy Metabolism. J Innate Immun 2018;10(3):163-71. doi: 10.1159/000481519.
62. Smith JD, Fu E, Kobayashi MA. Prevention and Management of Childhood Obesity and Its Psychological and Health Comorbidities. Annual review of clinical psychology 2020;16:351-78. doi: 10.1146/annurev-clinpsy-100219-060201.
63. Janicke DM, Steele RG, Gayes LA, Lim CS, Clifford LM, Schneider EM, Carmody JK, Westen S. Systematic review and meta-analysis of comprehensive behavioral family lifestyle interventions addressing pediatric obesity. J Pediatr Psychol 2014;39(8):809-25. doi: 10.1093/jpepsy/jsu023.
64. Newgard CB. Metabolomics and Metabolic Diseases: Where Do We Stand? Cell Metab 2017;25(1):43-56. doi: 10.1016/j.cmet.2016.09.018.
65. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 2002;106(25):3143-421.
66. Handakas E, Lau CH, Alfano R, Chatzi VL, Plusquin M, Vineis P, Robinson O. A systematic review of metabolomic studies of childhood obesity: State of the evidence for metabolic determinants and consequences. Obes Rev 2022;23 Suppl 1:e13384. doi: 10.1111/obr.13384.
67. Reinehr T, Wolters B, Knop C, Lass N, Hellmuth C, Harder U, Peissner W, Wahl S, Grallert H, Adamski J, et al. Changes in the serum metabolite profile in obese

children with weight loss. European Journal of Nutrition 2015;54(2):173-81. doi: 10.1007/s00394-014-0698-8.
68. Würtz P, Wang Q, Kangas AJ, Richmond RC, Skarp J, Tiainen M, Tynkkynen T, Soininen P, Havulinna AS, Kaakinen M, et al. Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. PLoS Med 2014;11(12):e1001765. doi: 10.1371/journal.pmed.1001765.
69. Moran-Ramos S, Ocampo-Medina E, Gutierrez-Aguilar R, Macias-Kauffer L, Villamil-Ramirez H, Lopez-Contreras BE, Leon-Mimila P, Vega-Badillo J, Gutierrez-Vidal R, Villarruel-Vazquez R, et al. An Amino Acid Signature Associated with Obesity Predicts 2-Year Risk of Hypertriglyceridemia in School-Age Children. Scientific Reports 2017;7:5607. doi: 10.1038/s41598-017-05765-4.
70. McCormack SE, Shaham O, McCarthy MA, Deik AA, Wang TJ, Gerszten RE, Clish CB, Mootha VK, Grinspoon SK, Fleischman A. Circulating branched-chain amino acid concentrations are associated with obesity and future insulin resistance in children and adolescents. Pediatric Obesity 2013;8(1):52-61. doi: 10.1111/j.2047-6310.2012.00087.x.
71. Ho JE, Larson MG, Ghorbani A, Cheng S, Chen MH, Keyes M, Rhee EP, Clish CB, Vasan RS, Gerszten RE, et al. Metabolomic Profiles of Body Mass Index in the Framingham Heart Study Reveal Distinct Cardiometabolic Phenotypes. PLoS One 2016;11(2):e0148361. doi: 10.1371/journal.pone.0148361.
72. Butte NF, Liu Y, Zakeri IF, Mohney RP, Mehta N, Voruganti VS, Goring H, Cole SA, Comuzzie AG. Global metabolomic profiling targeting childhood obesity in the Hispanic population. Am J Clin Nutr 2015;102(2):256-67. doi: 10.3945/ajcn.115.111872.
73. Hellmuth C, Uhl O, Kirchberg FF, Grote V, Weber M, Rzehak P, Carlier C, Ferre N, Verduci E, Gruszfeld D, et al. Effects of Early Nutrition on the Infant Metabolome. Nestle Nutr Inst Workshop Ser 2016;85:89-100. doi: 10.1159/000439491.
74. Layman DK, Walker DA. Potential importance of leucine in treatment of obesity and the metabolic syndrome. J Nutr 2006;136(1 Suppl):319s-23s. doi: 10.1093/jn/136.1.319S.
75. Blouet C, Jo YH, Li X, Schwartz GJ. Mediobasal hypothalamic leucine sensing regulates food intake through activation of a hypothalamus-brainstem circuit. J Neurosci 2009;29(26):8302-11. doi: 10.1523/jneurosci.1668-09.2009.
76. Bardanzellu F, Puddu M, Peroni DG, Fanos V. The Human Breast Milk Metabolome in Overweight and Obese Mothers. Front Immunol 2020;11:1533. doi: 10.3389/fimmu.2020.01533.
77. Troisi J, Pierri L, Landolfi A, Marciano F, Bisogno A, Belmonte F, Palladino C, Guercio Nuzio S, Campiglia P, Vajro P. Urinary Metabolomics in Pediatric Obesity and NAFLD Identifies Metabolic Pathways/Metabolites Related to Dietary Habits and Gut-Liver Axis Perturbations. Nutrients 2017;9(5). doi: 10.3390/nu9050485.
78. Lau CHE, Siskos AP, Maitre L, Robinson O, Athersuch TJ, Want EJ, Urquiza J, Casas M, Vafeiadi M, Roumeliotaki T, et al. Determinants of the urinary and

serum metabolome in children from six European populations. Bmc Medicine 2018;16. doi: 10.1186/s12916-018-1190-8.
79. Isganaitis E, Rifas-Shiman SL, Oken E, Dreyfuss JM, Gall W, Gillman MW, Patti ME. Associations of cord blood metabolites with early childhood obesity risk. Int J Obes (Lond) 2015;39(7):1041-8. doi: 10.1038/ijo.2015.39.
80. Sorrow P, Maguire R, Murphy SK, Belcher SM, Hoyo C. Elevated metabolites of acetaminophen in cord blood of children with obesity. Pediatr Obes 2019;14(1). doi: 10.1111/ijpo.12465.
81. Singh A, Shannon CP, Gautier B, Rohart F, Vacher M, Tebbutt SJ, KA LC. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. Bioinformatics 2019;35(17):3055-62. doi: 10.1093/bioinformatics/bty1054.
82. Rohart F, Gautier B, Singh A, KA LC. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS Comput Biol 2017;13(11):e1005752. doi: 10.1371/journal.pcbi.1005752.
83. Singh A, Gautier B, Shannon CP, Vacher M, Rohart F, Tebbutt SJ, Cao K-AL. DIABLO – an integrative, multi-omics, multivariate method for multi-group classification. bioRxiv 2016:067611. doi: 10.1101/067611.
84. Naveja JJ, Rico-Hidalgo MP, Medina-Franco JL. Analysis of a large food chemical database: chemical space, diversity, and complexity. F1000Res 2018;7. doi: 10.12688/f1000research.15440.2.

CHAPTER 2 – Nutritional Metabolomics and the Classification of Dietary Biomarker Candidates: A Critical Review

Thesis chapter is derived from a published peer-reviewed article:

Rafiq, T., Azab, S. M., Teo, K. K., Thabane, L., Anand, S. S., Morrison, K. M., de Souza, R. J., & Britz-McKibbin, P. (2021). Nutritional Metabolomics and the Classification of Dietary Biomarker Candidates: A Critical Review. Advances in nutrition (Bethesda, Md.), 12(6), 2333–2357. https://doi.org/10.1093/advances/nmab054

The authors' responsibilities were as follows— TR: drafting of the manuscript; TR, SMA, RJdS, and PB-M: designed and conceptualized the research; TR, SMA, LT, RJdS, and PB-M: development and implementation of literature search strategy; TR and SMA: acquisition of data, including review of literature search results and data abstraction; TR, SMA, RJdS, and PB-M: analysis and interpretation of data; TR SMA, KKT, LT, SSA, KMM, RJdS, and PB-M: critical revision of the manuscript for important intellectual content; KKT, RJdS, and PB-M: administrative, technical, and material support; RJdS and PB-M: study supervision. We acknowledge Laura E Banfield, MLIS (McMaster Health Sciences Library) for assisting with the development of the search strategy.

2.1 Abstract

Recent advances in metabolomics allows for more objective assessment of contemporary food exposures, which have been proposed as an alternative or complement to self-reporting of food intake. However, the quality of evidence supporting the utility of dietary biomarkers as valid measures of habitual intake of foods or complex dietary patterns in diverse populations has not been systematically evaluated. We reviewed nutritional metabolomics studies reporting metabolites associated with specific foods or food groups, evaluate the inter-study repeatability of dietary biomarker candidates, and report study design, metabolomic approach, analytical technique(s), and type of biofluid analyzed. A comprehensive literature search of five databases (PubMed, EMBASE, Web of Science, BIOSIS, and CINAHL) was conducted from inception through December 2020. This review included 244 studies: 169 (69%) of which were interventional studies (9 of these were replicated in free-living participants), and 151 (62%) of which measured the metabolomic profile of serum and/or plasma. Food-based metabolites identified in more than one study and/or biofluid were associated with 11 food-specific categories or dietary patterns: 1) fruits; 2) vegetables; 3) high fiber foods (grain-rich); 4) meats; 5) seafood; 6) pulses, legumes, and nuts; 7) alcohol; 8) caffeinated beverages, teas, and cocoas; 9) dairy and soya; 10) sweet and sugary foods; and 11) complex dietary patterns and other foods. We conclude that 69 metabolites represent good candidate biomarkers of food intake. Quantitative measurement of these metabolites will advance our understanding of the relationship between diet and chronic disease risk and support evidence-based dietary guidelines for global health.

Keywords: metabolomics, dietary biomarkers, nutrition, omics, food exposures

2.2 Introduction

Diet plays an important role in modulating the risk of chronic diseases including obesity, diabetes, cardiovascular disease, and certain cancers (1). Food intake in epidemiological studies has traditionally been assessed using self-reported and often memory-based approaches, including 24-hour dietary recalls, weighted food diaries, or food frequency questionnaires (FFQs). The reliability and validity of these tools have been questioned due to presence of potentially serious systematic and random measurement errors (2, 3). Errors such as misreporting of total energy intake and food portion sizes by 30-88% (4, 5) have hindered efforts to disentangle diet-disease relationships. Over the last decade, metabolomics has emerged as a valuable tool for revealing changes in metabolic profiles induced by recent or long-term/habitual diets (6, 7). High-throughput platforms for metabolomics enable comprehensive characterization of low molecular weight metabolites in biological samples, and offer a complement (or in some cases, an alternative) to self-report tools for objective assessment of "true" food exposures. Metabolomic studies may also better characterize dose-response relationships, which would be an advance over FFQs, which generally offer sufficient precision only to distinguish high from low consumers of food and food groups varying considerably across populations (8).

The primary focus of nutritional metabolomics has been the discovery of specific metabolites associated with food consumption and its impact on chronic disease risk. Such studies have led to the discovery of atherogenic trimethylamine-*N*-oxide (TMAO), a

metabolite produced by the gut microbiome from dietary nutrients such as choline, betaine, and *L*-carnitine that are prevalent in eggs, red meat, and fish (9, 10). The ability to discriminate metabolites of foods in a robust and generalizable manner depends on intrinsic factors such as characteristics of the study population (e.g., genetics, ethnicity, food habits) and extrinsic factors such as quantity and duration of food exposure. This problem is further exacerbated because there is no clear consensus on the choice of optimal study designs, sample size, metabolomic approach, biospecimen type, and methods used for metabolite identification and quantification (11).

The two main analytical techniques used in metabolomics are mass spectrometry (MS) and nuclear magnetic resonance (NMR); the latter method is highly robust, requires minimal sample handling, but is less sensitive. In contrast, MS-based approaches are usually preceded by more extensive sample preparation and chromatographic separations based on liquid chromatography (LC), gas chromatography (GC), or capillary electrophoresis (CE) for broader metabolome coverage with improved selectivity, including isomer resolution (12, 13). Recent advances in high resolution MS, in particular, the implementation of standardized LC-MS methods, have made it possible to detect thousands of molecular features when performing nontargeted metabolomics for hypothesis generation; however, rigorous data filtering approaches are needed to identify and authenticate metabolites while reducing data set redundancy and artifact signals to prevent false discoveries (8, 14, 15). On the other hand, targeted metabolomics is also widely used to quantify known list/group of known metabolites for hypothesis testing using validated analytical methods. Alternatively, both targeted and nontargeted

strategies using more than one analytical platform are increasingly used in large-scale metabolomic studies depending on sample volume requirements, sample throughput, and operational costs.

There are several thousands of low molecular weight compounds derived from foods. The food biomarker alliance (FOODBALL) is a joint initiative across 11 countries aimed at discovery and validation of dietary biomarkers (http://foodmetabolome.org/foodball/). *The Food Database (FooDB)* (http://www.foodb.ca/) is the most comprehensive database with over 70,000 metabolites derived from foods and food constituents (16). Also, *Exposome-Explorer* (http://exposome-explorer.iarc.fr) is a manually curated database of exposome chemicals including dietary and pollutant biomarkers (17). While these databases are comprehensive and useful, it is challenging for the scientific community to critically appraise and classify robust dietary biomarkers in a rapidly evolving field. Furthermore, recent nutritional metabolomic reviews do not distinguish between health/disease states of participants, and thus disease status may confound the association between dietary intake and their biomarkers (18, 19).

The purpose of this review is to 1) to generate a comprehensive list of metabolites associated with individual food and food groups in apparently healthy individuals, 2) report on the study designs, metabolomic approaches and biospecimen used, and 3) rate the evidence based on the inter-study repeatability and study design.

2.3 Methods

A comprehensive literature search was developed in collaboration with an information scientist. We searched Medline through OVID, EMBASE, Web of Science, BIOSIS, and CINAHL and included published articles from inception until December 2020. We used a comprehensive search strategy including a combination of Medical Subject Headings (MeSH) terms and key words related to study design, population, individual foods and food groups, and metabolomics. For the details of our search strategy, please see Supplemental Methods. References of the included studies were manually searched to identify any further relevant studies. Search results from all databases were merged and duplicates were removed with the use of EndNote citation manager (version X9, Thomson Reuters). Articles were initially screened based on title and/or abstract and full text of potential articles was retrieved and evaluated independently by two reviewers (TR and SMA). Any disagreement was resolved through discussion and if necessary, a third investigator made the final decision (RJdS).

2.3.1 Eligibility Criteria

Studies were eligible to be included in our review if 1) they were conducted in healthy adults or children of any sex or ethnicity, 2) used nontargeted or targeted approaches to identify metabolites of individual foods (e.g., oranges or red meat), complex dietary patterns (e.g., Mediterranean diet or meat-based diet), and/or specific nutrients or non-nutrients (e.g., *trans*-fats or carotenoids), 3) examined the relationship (observational studies) or the effect (intervention studies) of food on metabolites primarily in serum, plasma, or urine samples. We restricted the results to individual foods

and food groups but excluded dietary supplements, given that we were interested in reporting metabolites derived from food intake. We excluded studies: 1) that had examined food intake in conjunction with other interventions or lifestyle changes such as weight loss to ensure that a biomarker is specific to food and not some other intervention, 2) without a control group, 3) that enrolled participants with existing disease to ensure that identified biomarkers are not a result of a pathologic process or pharmacological intervention.

2.3.2 Study Selection Criteria

We identified 14,179 records across the five databases, and 12,177 remained after removal of duplicates (PRISMA flow diagram, Figure 2.1). The number of potentially relevant studies narrowed to 539 after title and abstract screening. After full-text review, a total of 244 studies remained eligible and were included in this systematic review.

2.3.3 Data extraction and Analysis

We extracted information regarding publication details including name of first author and year of publication, and study characteristics including age, country, type of study (e.g., feeding study or cross-sectional study), sample size, length of follow-up, specification of analytical technique, biological sample (urine or blood), exposure and/or comparator details, method of dietary assessment (only for observational studies), and all resulting metabolites following diet exposure (Appendix Tables 1 and 2). Given the large number and chemical diversity of food metabolites, a data reduction approach was applied where only those metabolites that were identified in at least two different studies and/or biofluids (blood and urine) are presented and discussed in the text of this review.

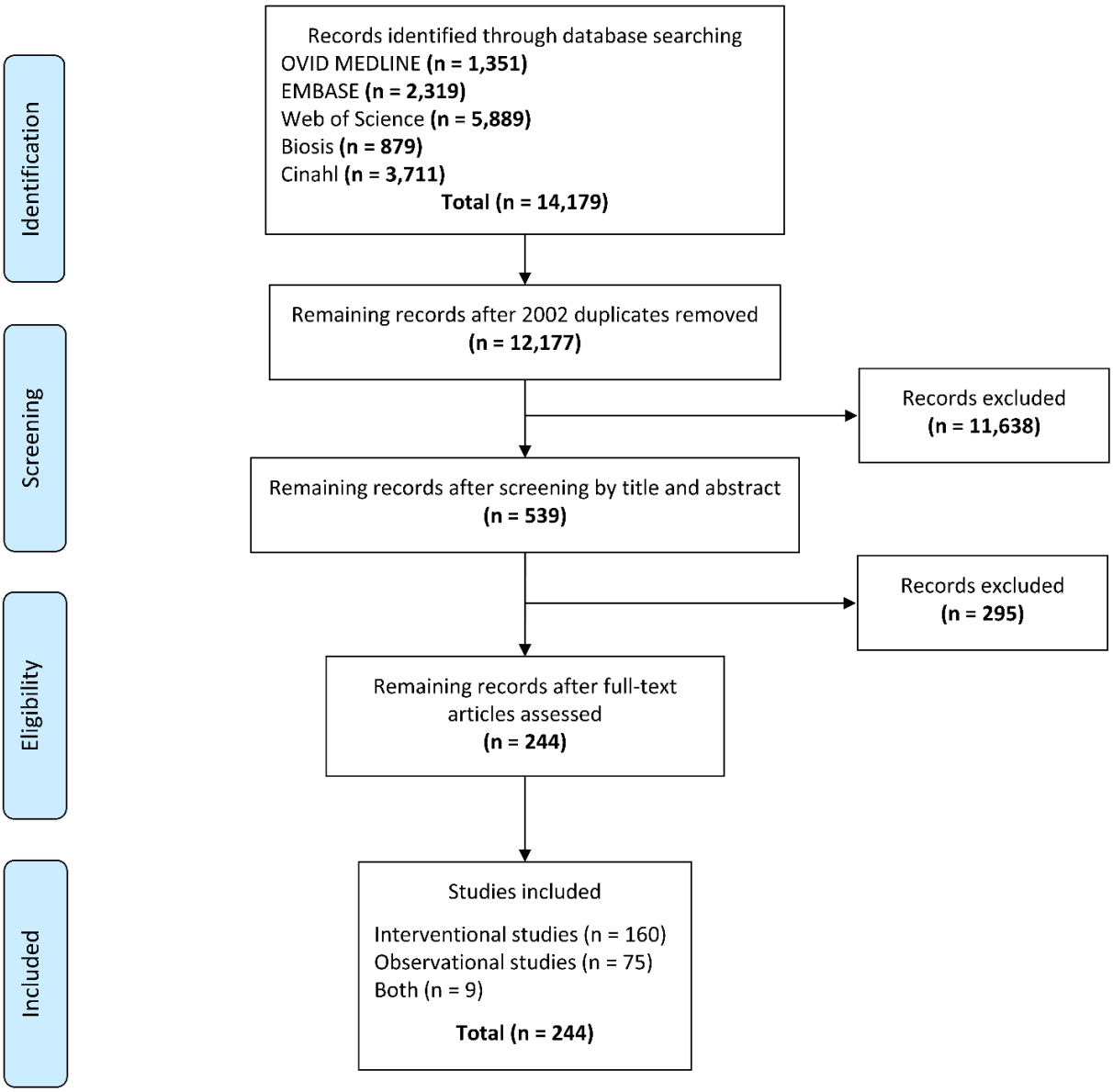


Figure 2.1: Flow diagram of the literature search process

2.3.4 Assessing Level of Evidence

We developed a scoring system to rate the evidence for each metabolite as a candidate biomarker of food intake into one of three mutually exclusive categories: 'good', 'fair', or 'poor'. The rating is based on empirical evidence of inter-study repeatability and study design.

Repeatability: Metabolites identified in more than one study were assigned a score of 2 points for each of these studies that was an interventional study plus 1 additional point for each observational study. Only metabolites that were replicated were assigned a score. The following algorithms were used to assess replication: (1) Two independent publications. A metabolite identified by one observational study and one interventional study was assigned a total score of 3 points (1x1 point for observational study and 1x2 points for interventional study); (2) A single publication reporting results from two independent cohorts/studies of a metabolite of a food, and both were congruent, was assigned a score of 3 points (1x1 points for observational study and 1x2 points for interventional study); (3) Two different biological fluids for the same cohort (urine and blood). For example, a biomarker identified in both urine and blood sample was assigned a score of 2 points if identified in an observational study (1x1[urine] + 1x1[blood]) and a score of 4 points if identified in an interventional study (1x2[urine] + 1x2[blood]).

Thus, the lowest score for replicated metabolite was 2 points, classified as poor evidence to be a score of 2 points, a score of 3-4 was considered as fair evidence, and a score of ≥5 points was considered good evidence (Table 2.1). While this scoring system has not been published previously in the literature, we have carefully designed it to be a

tool for assessing the extent of evidence of metabolites as related to recent or habitual food consumption. Certain metabolites recently recognized in the scientific community as "strong" biomarkers of food intake (BFIs) such as proline betaine for citrus fruits, were also correctly classified as "good" using our scoring system.

Table 2.1: List and Scoring of Food Metabolites Replicated in Literature[1]

Food Name	Good (≥ 5)	Fair (4-3)	Poor (2)
Fruits	Proline Betaine (5)[2]	Hippuric acid (4)[2]	
Strawberry	Pelargonidin Glucuronide (6)		
Apple		Epicatechin sulfate (4) Hydroxyphenylvaleric acid sulfate (4) Xylose (3)[2]	
Banana		3-methoxytyramine sulfate (3)[2] Dopamine sulfate (3)[2] Methoxyeugenol glucuronide (3)[2] Salsolinol sulfate 1 (3)[2]	
Fruit Juice		Proline Betaine (4) *N*-methylproline (3) *Scyllo*-inositol (3)	
Cranberry Juice		Ferulic acid sulfate (4) Sinapic acid (4) Quinic acid (4) Hippuric acid (4)	
Orange Juice	Proline Betaine (7)[2] Hippuric acid (6) 4'-Hydroxyhippuric acid (6) 3'-hydroxyhippuric acid (6) 4-hydroxyphenylacetic acid (6)	3-(3'-Hydroxy-4'-methoxyphenyl)hydracrylic acid (4) 3-(3'-Hydroxy-4'-methoxyphenyl)propionic acid (4) 3-(4'-Methoxyphenyl)propionic acid-3'-sulfate (4)	
Orange	Proline Betaine (5)[2]		
Citrus Fruit	Proline Betaine (9)[2]	*N*-methylproline (4) Naringenin (3)[2] Hesperetin (3)[2] *Chiro*-inositol (3) *Scyllo*-inositol (3)	
Broccoli	Sulforaphane (8) Sulforaphane *N*-acetylcysteine (8) Sulforaphane cysteine (8) Isothiocyanates (6)	Sulforaphane cysteinylglycine (4) Erucin-cysteine (4) Erucin-*N*-acetylcysteine (4)	
Broccoli Sprouts	Sulforaphane (8)	Erucin (4)	
Cruciferous Vegetables		*S*-Methyl-*L*-cysteine-sulfoxide (3)[2]	
Green Leafy Vegetables			3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
Mushroom			Ergothioneine

Table 1: (Continued)

Food Name	Good (≥ 5)	Fair (4-3)	Poor (2)
High-Fiber (grain rich)	Alkylresorcinols (8) 3-(3,5-dihydroxyphenyl)-1-propanoic acid (DHPPA) (5)[2]	2-aminophenol sulfate (4)[2] 3,5-dihydroxybenzoic acid (DHBA) (3)[2]	Daidzein Genistein
Whole grain Rye Bread		Alkylresorcinols (4) DHPPA sulphate (3)[2]	
Meat	Creatinine (6)[2]	Creatine (5)[2] *O*-acetyl-*L*-carnitine (3)[2] 4-hydroxyproline (3)[2] Glutamine (3)[2]	
Chicken/ Poultry	3-methylhistidine (11)[2]	Anserine (4)[2] Carnosine (4)[2] *O*-acetyl-*L*-carnitine (3)[2]	Pyroglutamine[3]
Processed Meat	*O*-acetyl-*L*-carnitine (6)[2]		
Red Meat	*O*-acetyl-*L*-carnitine (6)[2]	Trimethylamine-*N*-oxide (TMAO) (4) Carnosine (4)[2] Carnitine (3)[2] Anserine (3)[2]	
Seafood	Docosahexaenoic acid (DHA, 22:6n–3) (5)	CMPF (3) Eicosapentaenoic acid (20:5n–3) (3)	Docosapentaenoic acid (22:5n–3)
Fatty Fish	Docosahexaenoic acid (DHA, 22:6n–3) (5)[2]	Eicosapentaenoic acid (20:5n–3) (4)[2]	
Fish	Trimethylamine-*N*-oxide (TMAO) (19)[2] Docosahexaenoic acid (DHA, 22:6n–3) (12)[2] CMPF (7)[2] Creatine (7)[2] Eicosapentaenoic acid (20:5n–3) (7)[2] Dimethylamine (5)[2]	1-methylhistidine (4) 1,2,3,4-Tetrahydro-β-carboline-3-carboxylic acid (4) Arsenobetaine (4) 1-Docosahexaenoylglycero-phosphocholine (3) Docosapentaenoic acid (22:5n–3) (3)[2] Acetylcarnitine (3)[2]	Lysine Methionine Tryptophan Tyrosine
Seafood (Lean)		Trimethylamine-*N*-oxide (TMAO) (4)	
Seafood & Plant Protein			Docosahexaenoic acid (DHA, 22:6n–3)
Shellfish		3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF) (4)	2-hydroxybutyrate
Pulses/ Legumes/ Nuts		Trigonelline (4) 3-methylhistidine (4) Dimethylglycine (4) Trimethylamine (4) Lysine (4)	

Table 1: (Continued)

Food Name	Good (≥ 5)	Fair (4-3)	Poor (2)
Dry Bean Enriched Diet		Trigonelline (4) Pipecolic acid (4) S-methylcysteine (4)	
Nuts (Mixed)		Tryptophan betaine (4)	4-vinylphenol sulfate
Peanut		Tryptophan betaine (3) 4-vinylphenol sulfate (3)	
Alcohol		4-Androsten-3β-diol disulfate (3) 2-aminobutyrate (3)	α-Hydroxyisovalerate β-hydroxyisovalerate 5α-androstan-3β-diol disulfate 2-hydroxybutyrate 4-methyl-2-oxopentanoate Pipecolate Docosapentaenoic acid (22:5n–3) Stearidonate (18:4n3) Piperine Ethyl glucuronide Palmitoleate (16:1n–7) Dihomo-linoleate (20:2n–6) Malate 17β-diol disulfate 17β-diol disulfate 1
Liquor		Ethyl glucuronide (4)	
De-alcoholized Red Wine	Methylgallic sulfate (6) Σ(Epi)catechin glucuronides (6) 3-Hydroxyphenylacetic acid (6) *p*-Coumaric acid (6)	Ethylgallate sulfate (4) Ethylgallate (glucuronide 1) (4) Ethylgallate (glucuronide 2) (4) Σ Methyl(epi)catechin glucuronides (4) Σ Dihydroxyphenyl-γ-valerolactone glucuronide (4) ΣDihydroxyphenyl-γ-valerolactone sulfates (4) ΣMethoxy-hydroxyphenyl-γ-valerolactone glucuronide (4) 2,4-Dihydroxybenzoic acid (4) 2,6-Dihydroxybenzoic acid (4) 2,5-Dihydroxybenzoic acid (4) 3,5-Dihydroxybenzoic acid (4) 4-Hydroxybenzoic acid (4) 3-Hydroxybenzoic acid (4) Gallic acid (4) Methylgallic acid (4) 2-Hydroxyphenylacetic acid (4) Caffeic acid (4) Ferulic acid (4) 3-(3-hydroxyphenyl) propionic acid (4) Enterolactone (4) Pyrogallol (4) Syringic acid (4)	

Table 1: (Continued)

Food Name	Good (≥ 5)	Fair (4-3)	Poor (2)
		Ethylgallate (4) 3,4-Dihydroxyphenylacetic acid (4) Dihydrocaffeic acid (4) (Epi)catechin sulfates (4) Enterolactone (4)	
Wine		Ethyl glucuronide (3)	2,3-dihydroxyisovalerate 2,3-butanediol *Scyllo*-inositol
Red Wine	ΣMethyl(epi)catechin glucuronides (6) Methylgallic acid sulfate (5)[2]	Gallic acid (4) Methylgallic acid (4) 3-Hydroxyphenylacetic acid (4) p-Coumaric acid (4) (Epi)catechin glucuronide (4) dihydroxyphenyl-γ-valerolactone (DHPV) (4) DHPV 2 (4) ΣDHPV glucuronides (4) Ethylgallate (3)[2]	
Cocoa	3-methylxanthine (7)[2] 3-Methyluric acid (5)[2] 7-Methylxanthine (5)[2] Theobromine (5)[2]	Epicatechin-glucuronide (4) 5-(3′,4′-dihydroxyphenyl)-γ-valerolactone glucuronide (3)[2]	
Coffee	Paraxanthine (13)[2] Caffeine (13)[2] 1-methylxanthine (10)[2] Quinate (9)[2] Theophylline (10)[2] Hippuric acid (9)[2] Trigonelline (9)[2] 5-acetylamino-6-amino-3-methyluracil (8)[2] Dihydroferulic acid (8) 1,7-dimethylurate (6)[2] 1,3,7-Trimethylurate (6)[2] 3-hydroxyhippurate (6)[2] 1,3-dimethylurate (7)[2] Catechol sulfate (5)[2] Dihydrocaffeic acid (6) Caffeic acid (7)[2] Ferulic acid (5)[2] Feruloylquinic acid (5)[2] Isoferulic acid (6)[2] N-(2-furoyl)glycine (5)[2] Theobromine (5)[2]	3-caffeoylquinic acid (4) 3-methyl catechol sulfate (4)[2] 3-methylxanthine (4)[2] 4-caffeoylquinic acid (4) Dihydrocaffeic acid-3-O-sulfate (4) 1-methylurate (4)[2] 3-hydroxypyridine sulfate (3)[2] 7-methylguanine (3)[2] Caffeic acid sulfate (3)[2] Citraconate (3)[2] Cyclo(leu-pro) (3) Gallic acid (3)[2] Kynurenic acid (3)[2]	3,7-dimethyluric acid
Green Tea	Hippuric acid (10)	*O*-methyl-epicatechin-*O*-sulfates (4) *O*-me-epigallocatechin-*O*-glucuronide (4) (-)-epigallocatechin-3-gallate (4)	

Table 1: (Continued)

Food Name	Good (≥ 5)	Fair (4-3)	Poor (2)
Black Tea	4-*O*-methylgallic acid (5)[2]	Hippuric acid (4)	
Chocolate	Theobromine (7)[2] 7-methyluric (5)[2]	7-methylxanthine (4)[2] 6-Amino-5-[N-methylformylamino]-1-methyluracil (6-AMMU) (4) 3,7-dimethyluric acid (3)[2]	
Dark Chocolate		4-hydroxyphenyl acetate (4)	
Sweet and Sugary Beverages		Citrulline (3)[2] Taurine (3)[2] Isocitrate (3)[2]	Carbon isotopic signatures (δ13C)
Dairy			Pantothenic acid (vitamin B5)
Butter		10-undecenoic acid (11:1n1) (3)	Pentadecanoate (15:0) Methyl palmitic isomers
Cheese		3-phenyllactic (4)[2] Proline (4)[2] Methionine (4)[2]	
Milk	Galactonic acid (5)[2]	Galactose (4) Lactose (4) Galactono-1,5-lactone (4) Urea (4)	Uridine
High Soy Diet	Daidzein (9)[2] Genistein (8)[2] *O*-desmethylangolensin (*O*-DMA) (5)[2]	Equol (4) Glycitein (3)	Total isoflavonoids
Soy-Based Drink		Pinitol (4)	4-ethylphenylsulfate
Soy-Based Cheese		Daidzein (4) Genistein (4)	
Whey		Leucine/Isoleucine (4)	
Average Danish Diet		Theobromine (4) Proline betaine (4)	
DASH Diet		β-Cryptoxanthin (3)[2]	
Fruits & Vegetables	Hippuric acid (5)[2]	β-Carotene (3)[2] Genistein (3)[2] Total carotenoid (3)[2]	
Healthy Eating Index		3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF) (3)[2] Eicosapentaenoic acid (20:5n–3) (3)[2] Hippuric acid (3)[2]	Docosahexaenoylcholine Docosahexaenoic acid (DHA, 22:6n–3) Carotene diol Ergothioneine

Table 1: (Continued)

Food Name	Good (≥ 5)	Fair (4-3)	Poor (2)
High Carotenoid Diet		α-Carotene (3)[2] β-Carotene (3)[2] Total carotenoids (3)[2]	
Mediterranean Diet		Docosahexaenoic acid (DHA, 22:6n–3) (4)[2]	3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
New Nordic Diet	Trimethylamine-*N*-oxide (TMAO) (6)	Hippuric acid (4)	
Vegetarian			Lysine Methionine Tryptophan Tyrosine
Vegan		Alanine (3)[2] Glycine (3)	

[1]Metabolites identified in at least 2 studies
[2]Robust biomarker (i.e., reported using both interventional and observational study design)
[3]Inverse association

Interstudy repeatability score: interventional studies (2x); observational studies (1x) – Example: metabolite found in 2 interventional studies and 1 observational study will have a score of 5

Good = score of 5 or more
Fair = score of 4-3
Poor = score of 2

Mass-to-charge ratio (*m/z)* for good metabolites only reported using untargeted analysis: Proline Betaine for orange (*m/z* = 144.0988), Trigonelline (*m/z* = 138.0550), 1,7-dimethylurate (*m/z* = 195.0524), 1,3,7-Trimethylurate (*m/z* = 209.068), 3-hydroxyhippurate (*m/z* = 194.0459), 1,3-dimethylurate (*m/z* = 197.0669), and Catechol sulfate (*m/z* = 188.9863) for coffee, Theobromine for chocolate (*m/z* = 181.0720), and Trimethylamine-*N*-oxide (TMAO) for New Nordic Diet (NND) (*m/z* = 76.0757)

2.4 Results

This review included 244 studies: 169 (69%) of which were interventional studies (9 of these were replicated in free-living participants), and 101 (4 1%) of which measured metabolomic profile of urine, plasma (n = 64), serum (n= 46), or in both plasma/serum and urine samples (n = 41). A total of 7,273 individuals contributed data to 169 interventional studies (average of 42 participants per study), and 79,256 individuals participated in 84 observational studies (average of 922 participants per study). Most

studies focused on adult population with only two intervention and seven observational studies including children and/or adolescents. All but two intervention and three observational studies did not provide information on age of participants, and nearly all studies reported sex-related information. The dietary biomarkers were measured in blood (plasma or serum) and/or urine sample and were detected using LC–MS (mainly with either reversed-phase (RP) or hydrophilic interaction (HILIC) modes), GC–MS, ^{1}H NMR, or other analytical methods (e.g., flow-injection electrospray ionization-MS, capillary electrophoresis-mass spectrometry, or inductively coupled plasma MS) (Supplemental Figure S2.1). Each metabolite was scored based on the interstudy repeatability and study design score system described above. As expected, proline betaine was classified to have good evidence (score >5) for intake of citrus fruits as it appeared in two interventional studies (score=4) and five observational studies (score=5), for a combined score of 9. Meanwhile, ergothioneine for intake of mushrooms appeared in two observational studies and was classified to have poor evidence (score=2). Overall, our review concluded that 69 metabolites are good, 161 are fair, and 48 are poor biomarkers of foods.

Most food-derived exogenous compounds are biotransformed into one or more metabolites following primary and secondary metabolism, and have an optimal detection window within a 24-hour period depending on dose and frequency of food intake (mostly with urine sample), though some extend to 48-hours or longer. In this section, we discuss robust dietary biomarkers associated with intake of specific foods or complex dietary patterns. Metabolites identified in more than one study or biofluid were grouped into the

following eleven categories: 1) Fruits; 2) Vegetables; 3) High Fiber (grain-rich); 4) Meats; 5) Seafood; 6) Pulses, Legumes, and Nuts; 7) Alcohol; 8) Caffeinated Beverages, Teas, and Cocoas; 9) Dairy and Soya; 10) Sweet and Sugary Foods; and 11) Complex Dietary Patterns and Other Foods.

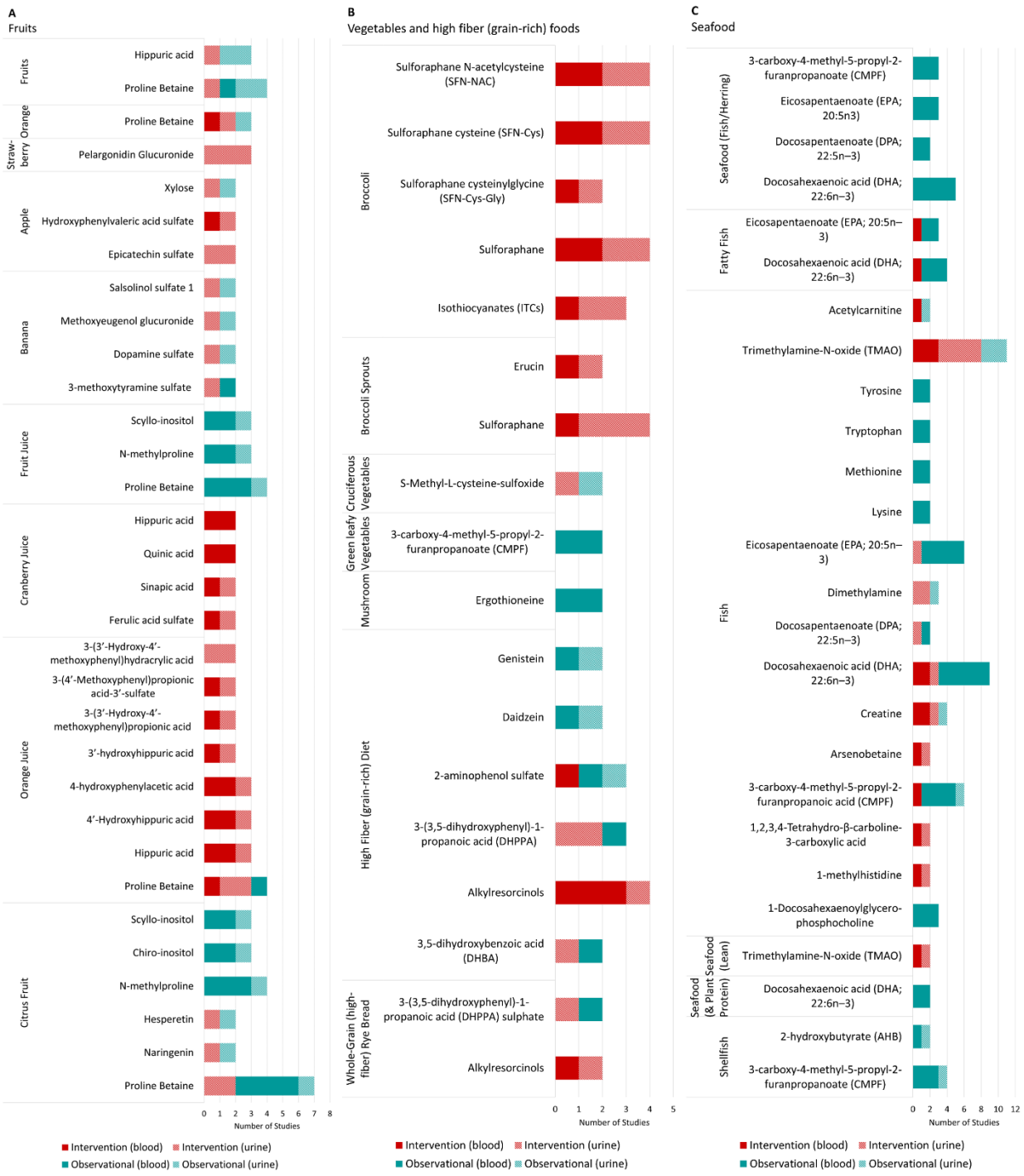

A
Fruits
Fruits
Hippuric acid
Proline Betaine
Straw-berry
Orange
Proline Betaine
Pelargonidin Glucuronide
Apple
Xylose
Hydroxyphenylvaleric acid sulfate
Epicatechin sulfate
Banana
Salsolinol sulfate 1
Methoxyeugenol glucuronide
Dopamine sulfate
3-methoxytyramine sulfate
Fruit Juice
Scyllo-inositol
N-methylproline
Proline Betaine
Cranberry Juice
Hippuric acid
Quinic acid
Sinapic acid
Ferulic acid sulfate
Orange Juice
3-(3′-Hydroxy-4′-methoxyphenyl)hydracrylic acid
3-(4′-Methoxyphenyl)propionic acid-3′-sulfate
3-(3′-Hydroxy-4′-methoxyphenyl)propionic acid
3′-hydroxyhippuric acid
4-hydroxyphenylacetic acid
4′-Hydroxyhippuric acid
Hippuric acid
Proline Betaine
Citrus Fruit
Scyllo-inositol
Chiro-inositol
N-methylproline
Hesperetin
Naringenin
Proline Betaine
0 1 2 3 4 5 6 7 8
Number of Studies
Intervention (blood)
Intervention (urine)
Observational (blood)
Observational (urine)
B
Vegetables and high fiber (grain-rich) foods
Broccoli
Sulforaphane N-acetylcysteine (SFN-NAC)
Sulforaphane cysteine (SFN-Cys)
Sulforaphane cysteinylglycine (SFN-Cys-Gly)
Sulforaphane
Isothiocyanates (ITCs)
Broccoli Sprouts
Erucin
Sulforaphane
Green leafy Cruciferous Vegetables
S-Methyl-L-cysteine-sulfoxide
Mushroom Vegetables
3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
Ergothioneine
High Fiber (grain-rich) Diet
Genistein
Daidzein
2-aminophenol sulfate
3-(3,5-dihydroxyphenyl)-1-propanoic acid (DHPPA)
Alkylresorcinols
3,5-dihydroxybenzoic acid (DHBA)
Whole-Grain (high-fiber) Rye Bread
3-(3,5-dihydroxyphenyl)-1-propanoic acid (DHPPA) sulphate
Alkylresorcinols
0 1 2 3 4 5
Number of Studies
Intervention (blood)
Intervention (urine)
Observational (blood)
Observational (urine)
C
Seafood
Seafood (Fish/Herring)
3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
Eicosapentaenoate (EPA; 20:5n3)
Docosapentaenoate (DPA; 22:5n–3)
Docosahexaenoic acid (DHA; 22:6n–3)
Fatty Fish
Eicosapentaenoate (EPA; 20:5n–3)
Docosahexaenoic acid (DHA; 22:6n–3)
Fish
Acetylcarnitine
Trimethylamine-N-oxide (TMAO)
Tyrosine
Tryptophan
Methionine
Lysine
Eicosapentaenoate (EPA; 20:5n–3)
Dimethylamine
Docosapentaenoate (DPA; 22:5n–3)
Docosahexaenoic acid (DHA; 22:6n–3)
Creatine
Arsenobetaine
3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF)
1,2,3,4-Tetrahydro-β-carboline-3-carboxylic acid
1-methylhistidine
1-Docosahexaenoylglycero-phosphocholine
Seafood (& Plant Seafood Protein) (Lean)
Trimethylamine-N-oxide (TMAO)
Docosahexaenoic acid (DHA; 22:6n–3)
Shellfish
2-hydroxybutyrate (AHB)
3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
0 2 4 6 8 10 12
Number of Studies
Intervention (blood)
Intervention (urine)
Observational (blood)
Observational (urine)

Figure 2.2: Metabolites identified from fruits, vegetables and high fiber (grain-rich) foods, and seafood by number of studies, type of study design, and type of biofluid

2.4.1 Fruits

A total of 29 subcategories of fruits were identified in this systematic review, of which, nine categories had reported at least one metabolite that was replicated (Table 2.1, Figure 2.2). Metabolites for intake of fruits were analyzed in two interventional (20, 21) and eight observational studies (22-29), fruit juices in one interventional (20) and three observational studies (22, 30, 31), citrus fruits in three interventional (20, 32, 33) and six observational studies (22, 27, 31, 34-36), orange in three interventional (20, 33, 37) and one observational study (22), orange juice in seven interventional (21, 38-43) and one observational study (34), apple in five interventional (20, 44-47) and two observational study (22, 45), banana in one interventional (48) and four observational studies (31, 34, 48, 49), strawberry in four interventional studies (50-53), and cranberry juice in four interventional studies (54-57). Several studies reported higher concentration of proline betaine with intake of fruits in general (21, 24-26). Proline betaine was also identified as the most frequent biomarker of citrus fruit (20, 22, 27, 31, 32, 34) and orange juice (21, 34, 38, 40), fruit juice (22, 30, 31), and the only metabolomic signature of orange fruit (22, 33, 37). Proline betaine was specific to the habitual consumption of citrus fruit or fruit juice due to its high natural abundance with appreciable amounts found in less commonly eaten foods, such as *Stachys affinis* or Chinese artichoke (58). In addition, studies reported higher levels of hippuric acid with intake of fruits (21, 24, 25), orange juice (39, 42, 43), and cranberry juice (55, 56). Other metabolites (including

biotransformed hippuric acid metabolites excreted in urine) identified were 3'-hydroxyhippuric acid (39, 42, 43), 4'-hydroxyhippuric acid (41-43), 4-hydroxyphenylacetic acid (39, 41, 43), 3-(3'-hydroxy-4'-methoxyphenyl)hydracrylic acid (41, 42), 3-(3'-hydroxy-4'-methoxyphenyl)propionic acid (42, 43), and 3-(4'-methoxyphenyl)propionic acid-3'-sulfate (41, 43) for orange juice, naringenin (20, 35), hesperetin (20, 35), *N*-methylproline (22, 27, 34), *chiro*-inositol (22, 27), and *scyllo*-inositol (22, 27) for citrus fruits, epicatechin sulfate (45, 47), hydroxyphenyl valeric acid sulfate (47), and xylose (45) for apple, and *N*-methylproline, *chiro*-inositol, and *scyllo*-inositol for intake of fruit juice (22, 30, 31). Meanwhile, pelargonidin, the main anthocyanin highly specific to strawberries was the only dietary biomarker reported at high concentration after intake of strawberry (50, 51, 53), 3-methoxytyramine sulfate after intake of banana (34, 48), and one study reported higher urine and plasma concentration of ferulic acid sulfate and sinapic acid (54), and quinic acid (55, 56) following intake of cranberry juice.

2.4.2 Vegetables

Five of the total 20 vegetable subcategories had identified at least one replicated metabolite as a dietary biomarker (Table 2.1, Figure 2.2). Metabolites associated with intake of broccoli were analyzed in five interventional studies (59-63), broccoli sprouts in four interventional studies (64-67), cruciferous vegetables in two interventional (68, 69) and three observational studies (22, 27, 34), green leafy vegetables in three observational studies (22, 27, 31), and mushroom in two observational studies (31, 34). Studies reported increased concentration of sulforaphane as the frequently identified metabolite,

which is derived from hydrolysis of glucosinolates by myrosinase, to be associated with intake of broccoli (62, 63) and broccoli sprouts (64, 65, 67). Additionally, sulfur-containing isothiocyanate exogenous compound prevalent in cruciferous vegetables was another more frequently identified metabolite for intake of broccoli (60, 61), as well as related sulforaphane metabolites/thiol conjugates excreted in urine such as sulforaphane cysteinylglycine (62), sulforaphane cysteine (62, 63), and sulforaphane *N*-acetylcysteine (SFN-NAC) (62, 63). In addition, a higher concentration of erucin was found in urine or blood with intake of broccoli sprouts (65), erucin-cysteine and erucin *N*-acetylcysteine with broccoli (63), and *S*-methyl-*L*-cysteine-sulfoxide, 3-carboxy-4-methyl-5-propyl-2-furanproponante (CMPF), and ergothioneine were the only metabolites associated with cruciferous vegetables (34, 69), green leafy vegetables (27, 31), and mushrooms (31, 34), respectively.

2.4.3 High Fiber (Grain-rich) Foods

The subcategories of high fiber (grain-rich) foods and whole-grain rye bread had identified at least one metabolite that was replicated (Table 2.1, Figure 2.2). Metabolites for a high fiber diet were examined in six interventional (70-75) and nine observational studies (22, 23, 26, 29, 31, 76-79) and whole-grain rye bread in seven interventional (74, 80-85) and one observational study (86). Higher concentration of urinary and blood alkylresorcinols**,** well-known phenolic lipids that are prevalent in whole grain wheat and rye, and 3-(3,5-dihyroxyphenyl)-1-propanoic acid (DHPPA), which can be measured as free molecules or as glucuronide or sulfonate conjugates resulting from phase I and II metabolism, was reported with intake of a high fiber diet (72-75) and whole-grain rye

bread (80, 81, 86). Meanwhile, studies reported higher intake of higher dietary fiber to be associated with greater urinary excretion and blood concentration of 3,5-dihydroxybenzoic acid (DHBA) (75, 77), 2-aminophenol sulfate (26, 70, 79), as well as daidzein (23, 76), and genistein (23, 76). The latter two phytochemicals are not specific to fiber intake since they are also prevalent in soya products, which have long been associated with habitual dietary patterns, and cancer and chronic disease risk (87).

2.4.4 Seafood

Six of the total eight subcategories had at least one metabolite that was replicated (Table 2.1, Figure 2.2). Metabolites for intake of seafood in general were identified in five observational studies (13, 26, 30, 31, 88), fatty fish in one interventional study (89) and three observational studies (90-92), fish in nine interventional studies (69, 93-99) and 15 observational studies (22, 24, 27, 31, 34, 52, 88, 90, 94, 99-105), lean seafood in two interventional studies (106, 107), seafood in combination with plant protein in two observational studies (26, 108), and shellfish in four observational studies (22, 27, 34, 105). Docosahexaenoic acid (DHA, 22:6n–3), an essential omega-3 fatty acid, was the most frequently reported dietary biomarker of seafood in general (13, 26, 30, 31, 88), fatty fish (89-92), and seafood in combination with plant protein (26, 108). Meanwhile, TMAO (a gut microbiota-generated metabolite) was the most frequently reported metabolite associated with intake of fish (52, 69, 94, 97-99, 105). DHA was the second most frequently reported metabolite associated with fish intake (22, 27, 31, 34, 90, 94, 95, 97, 102), and CMPF for seafood in general (26, 30, 31), and shellfish (22, 27, 34). Further, two other omega-3 fatty acids, docosapentaenoic acid (22:5n–3) and

eicosapentaenoic acid (20:5n–3) were both reported higher after intake of seafood (13, 26, 30, 31), fish (22, 27, 31, 34, 90, 95), and higher concentration of eicosapentaenoic acid (20:5n–3) was reported with intake of fatty fish (89-91). Also, elevated levels of CMPF (22, 27, 31, 34, 96), creatine (97, 98, 105), and dimethylamine (69, 99, 105) were associated with intake of fish, and TMAO for lean seafood (106, 107). Few other metabolites were replicated for intake of fish, and another metabolite for shellfish was 2-hydroxybutyrate (22) an endogenous metabolite also associated with threonine metabolism and oxidative stress (109). We therefore do not consider it a specific biomarker for shellfish.

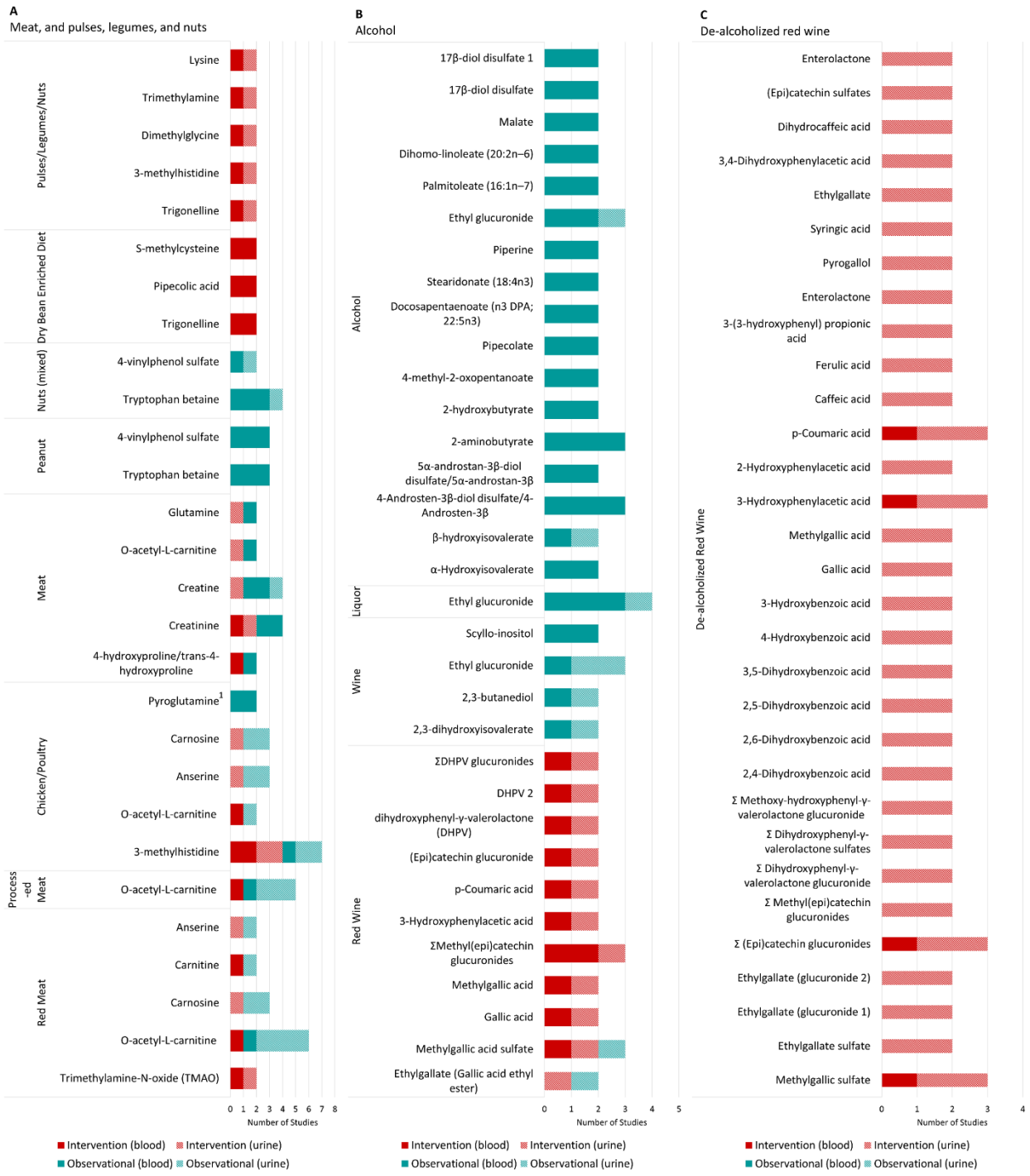

A
Meat, and pulses, legumes, and nuts
Pulses/Legumes/Nuts
Lysine
Trimethylamine
Dimethylglycine
3-methylhistidine
Trigonelline
Dry Bean Enriched Diet
S-methylcysteine
Pipecolic acid
Trigonelline
Nuts (mixed)
4-vinylphenol sulfate
Tryptophan betaine
Peanut
4-vinylphenol sulfate
Tryptophan betaine
Meat
Glutamine
O-acetyl-L-carnitine
Creatine
Creatinine
4-hydroxyproline/trans-4-hydroxyproline
Chicken/Poultry
Pyroglutamine[1]
Carnosine
Anserine
O-acetyl-L-carnitine
3-methylhistidine
Process-ed Meat
O-acetyl-L-carnitine
Red Meat
Anserine
Carnitine
Carnosine
O-acetyl-L-carnitine
Trimethylamine-N-oxide (TMAO)
0 1 2 3 4 5 6 7 8
Number of Studies
Intervention (blood)
Intervention (urine)
Observational (blood)
Observational (urine)
B
Alcohol
Alcohol
17β-diol disulfate 1
17β-diol disulfate
Malate
Dihomo-linoleate (20:2n–6)
Palmitoleate (16:1n–7)
Ethyl glucuronide
Piperine
Stearidonate (18:4n3)
Docosapentaenoate (n3 DPA; 22:5n3)
Pipecolate
4-methyl-2-oxopentanoate
2-hydroxybutyrate
2-aminobutyrate
5α-androstan-3β-diol disulfate/5α-androstan-3β
4-Androsten-3β-diol disulfate/4-Androsten-3β
β-hydroxyisovalerate
α-Hydroxyisovalerate
Liquor
Ethyl glucuronide
Wine
Scyllo-inositol
Ethyl glucuronide
2,3-butanediol
2,3-dihydroxyisovalerate
Red Wine
ΣDHPV glucuronides
DHPV 2
dihydroxyphenyl-γ-valerolactone (DHPV)
(Epi)catechin glucuronide
p-Coumaric acid
3-Hydroxyphenylacetic acid
ΣMethyl(epi)catechin glucuronides
Methylgallic acid
Gallic acid
Methylgallic acid sulfate
Ethylgallate (Gallic acid ethyl ester)
0 1 2 3 4 5
Number of Studies
Intervention (blood)
Intervention (urine)
Observational (blood)
Observational (urine)
C
De-alcoholized red wine
De-alcoholized Red Wine
Enterolactone
(Epi)catechin sulfates
Dihydrocaffeic acid
3,4-Dihydroxyphenylacetic acid
Ethylgallate
Syringic acid
Pyrogallol
Enterolactone
3-(3-hydroxyphenyl) propionic acid
Ferulic acid
Caffeic acid
p-Coumaric acid
2-Hydroxyphenylacetic acid
3-Hydroxyphenylacetic acid
Methylgallic acid
Gallic acid
3-Hydroxybenzoic acid
4-Hydroxybenzoic acid
3,5-Dihydroxybenzoic acid
2,5-Dihydroxybenzoic acid
2,6-Dihydroxybenzoic acid
2,4-Dihydroxybenzoic acid
Σ Methoxy-hydroxyphenyl-γ-valerolactone glucuronide
Σ Dihydroxyphenyl-γ-valerolactone sulfates
Σ Dihydroxyphenyl-γ-valerolactone glucuronide
Σ Methyl(epi)catechin glucuronides
Σ (Epi)catechin glucuronides
Ethylgallate (glucuronide 2)
Ethylgallate (glucuronide 1)
Ethylgallate sulfate
Methylgallic sulfate
0 1 2 3 4
Number of Studies
Intervention (blood)
Intervention (urine)
Observational (blood)
Observational (urine)

Figure 2.3: Metabolites identified from meats, pulses, legumes, and nuts, alcohol, and dealcoholized red wine by number of studies, type of study design, and type of biofluid

[1]Metabolites in lower concentration compared to control

2.4.5 Meats

Six meat subcategories were identified in this systematic review, of which four categories of overall meat intake, chicken/poultry, processed meat, and red meat had reported at least one metabolite that was replicated (Table 2.1, Figure 2.3). Examination of potential metabolites for meats was analyzed in two interventional (93, 110) and seven observational studies (24, 31, 90, 101, 104, 111, 112), poultry/chicken in four interventional (69, 94, 113, 114) and six observational studies (22, 27, 31, 34, 94, 114), processed meat in one interventional (94) and five observational (22, 27, 31, 94, 115), and red meat in five interventional (94, 114, 116-118) and seven observational studies (22, 27, 34, 94, 114, 115, 119). The most frequently identified metabolites include creatinine (93, 104, 110, 112) for meat, which was first identified a few decades ago and is degraded from creatine during cooking, *O*-acetyl-*L*-carnitine for red meat (22, 94, 115, 118), and a modified amino acid, 3-methylhistidine for chicken/poultry (34, 69, 94, 113, 114) which has long been used as a biomarker for muscle protein turnover. Other replicated markers include 4-hydroxyproline (31, 93), glutamine (110, 112), creatine (24, 110, 112), and *O*-acetyl-*L*-carnitine (104, 110) for meat, TMAO (116), anserine (114), carnosine (94, 114), and *L*-carnitine (22, 116) specifically for red meat, *O*-acetyl-*L*-carnitine for processed meat (22, 94, 115), as well as higher anserine and carnosine (94, 114) and lower pyroglutamine level (27, 31) for chicken/poultry.

Pulses, Legumes, and Nuts

Four of the seven subcategories including mixes of pulses, legumes, nuts or dry-bean enriched diets, mixed nuts, and peanut had reported at least one replicated metabolite (Table 2.1, Figure 2.3). Metabolites for intake of pulses/legumes/nuts in general were analyzed in three interventional studies (120-122) and four cross-sectional studies (23, 27, 123, 124), dry-bean enriched diet in two interventional studies (125, 126), mixed nuts in one interventional (127) and three observational studies (22, 26, 34), and peanuts in four observational studies (27, 34, 123, 128). Studies reported higher levels of tryptophan betaine (an indole alkaloid) and 4-vinylphenol sulfate (a xenobiotic associated with benzoate metabolism) with intake of mixed nuts (22, 26, 34) and peanuts (27, 34, 128). In addition, increased concentration in a vitamin B3 metabolite, trigonelline was reported with intake of pulses/legumes/nuts (120) and dry-bean enriched diet (125, 126). Other dietary biomarkers include 3-methylhistidine, dimethylglycine, trimethylamine, and lysine for pulses/legumes/nuts (120) and pipecolic acid and *S*-methylcysteine for dry-bean enriched diet (125, 126).

2.4.6 Alcohol

The subcategories of alcohol, liquor, wine, red wine, and de-alcoholized red wine had reported at least one metabolite that was replicated (Table 2.1, Figure 2.3B-C). Metabolites for intake of alcohol were analyzed in 12 observational studies (22, 27, 31, 34, 90, 101, 128-133), liquor in four observational studies (22, 27, 31, 34), wine in three interventional (20, 134, 135) and six observational studies (22, 27, 31, 34, 136, 137), red wine in three interventional (138-140) and three observational studies (34, 36, 136), and

de-alcoholized red wine in two interventional studies (138, 141). While several metabolomic signatures were identified to be associated with intake of alcohol, red wine, and de-alcoholized red wine, the more frequently reported metabolites include 4-androsten-3β-diol disulfate (27, 31, 128) and 2-aminobutyrate (31, 128, 130) for alcohol, the sum of methyl(epi)catechin glucuronides (138, 140) for red wine, and the sum of (epi)catechin glucuronides, 3-hydroxyphenylacetic acid, and *p*-coumaric acid for dealcoholized red wine (138, 141). Additional metabolites associated with intake of de-alcoholized red wine were methylgallic sulfate, 3-hydroxyphenylacetic acid, and *p*-coumaric acid (138). Additionally, higher concentration of ethyl glucuronide, a common secondary metabolite of ethanol excreted in urine, was most frequently reported with intake of wine (22, 34, 137) and liquor (22, 27, 34).

2.4.7 Caffeinated Beverages, Teas, and Cocoas

The subcategories of black tea, green tea, cocoa, and coffee intake had reported at least one metabolite that was replicated (Table 2.1, Figure 2.4A-B). Metabolites for intake of black tea was analyzed in four interventional (20, 142-144) and three observational studies (31, 145, 146), green tea in eight interventional (143, 144, 147-152) and one observational study (146), cocoa in six interventional (153-158) and one observational study (159), and coffee in 10 interventional (20, 46, 160-167) and 16 observational studies (22, 27, 30, 31, 34, 36, 136, 146, 168-175). Paraxanthine (22, 27, 30, 34, 160, 162, 170-172, 174) and 1,3,7-trimethylxanthine (coffee) (22, 30, 34, 160, 162, 169, 171-174) were the most frequently identified markers for coffee intake. Among many, some of the other metabolites more frequently identified for coffee intake include

hippuric acid (formed by the conjugation of benzoic acid with glycine) (22, 136, 162, 163, 171, 174), and well-known coffee constituent theobromine and its metabolites 1-methylxanthine (22, 27, 30, 34, 171, 172, 176) and 3-methylxanthine (136, 160, 162, 172). Meanwhile, 4-*O*-methylgallic acid, a methyl ether derivative of gallic acid, and hippuric acid were the most frequently identified metabolites for intake of black tea (20, 142, 145) and green tea (143, 144, 147, 148, 152), respectively. Furthermore, higher levels of the well-known coffee constituent theobromine and its metabolite 3-methylxanthine (most frequently) were associated with intake of cocoa (153, 156, 157, 159). There were no biomarkers for decaffeinated coffee that were reported in at least 2 studies, suggesting that the metabolites associated with coffee may likely be metabolites of caffeine and not specific to coffee. 1-methyluric acid and 5-acetamido-6-amino-3-methyluracil (AAMU) are also widely measured end-products of caffeine metabolism prevalent in urine that are associated with caffeinated beverage intake in large populations (172).

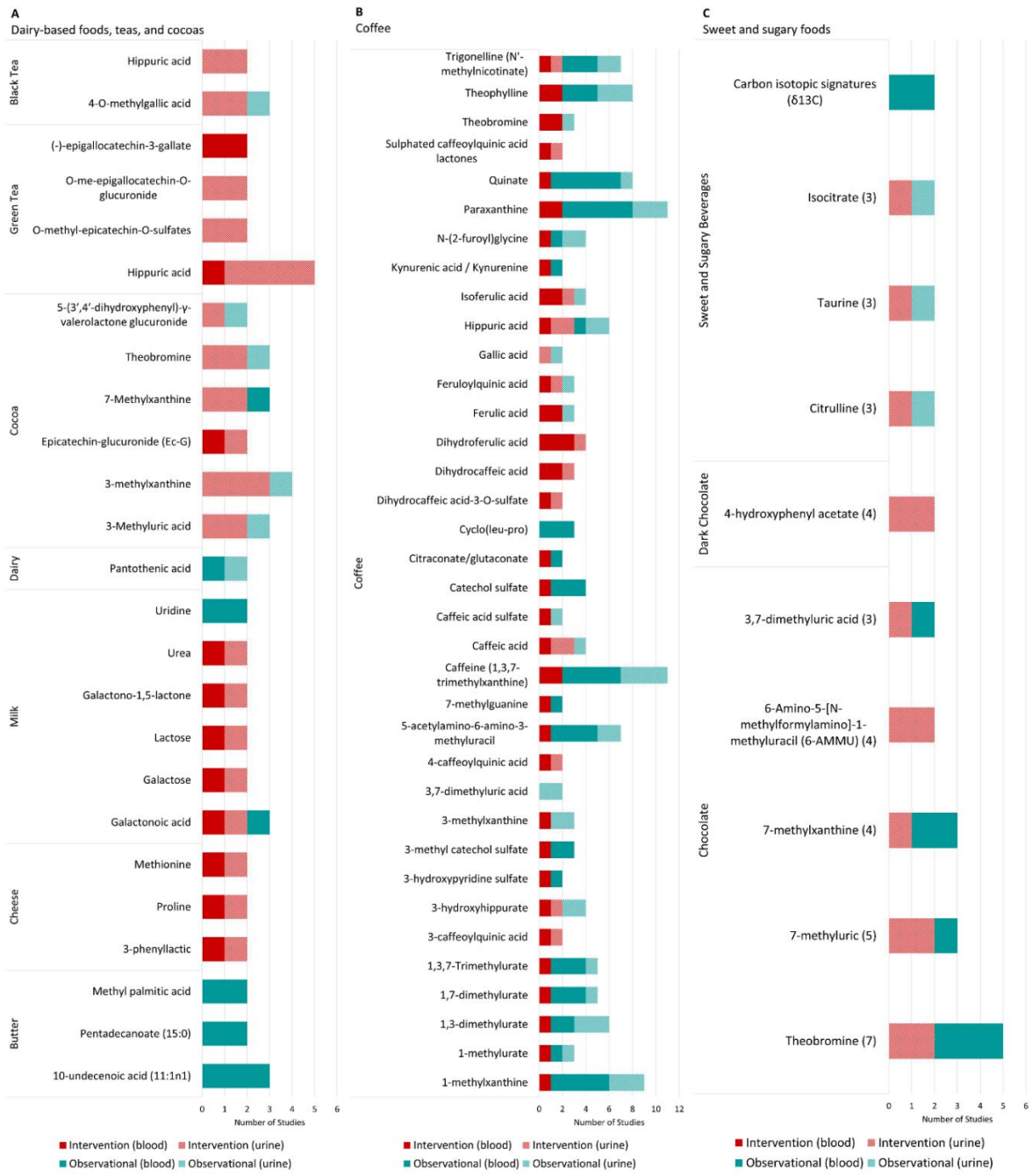


Figure 2.4: Metabolites identified from dairy-based foods, teas, cocoas, coffee, and sweet and sugary foods by number of studies, type of study design, and type of biofluid

2.4.8 Dairy

Three of the five subcategories including intake of dairy products, butter, cheese, and milk had reported at least one metabolite that was replicated (Table 2.1, Figure 2.4A). Metabolites for intake of dairy products was analyzed in two interventional studies (177, 178) and five observational studies (26, 90, 179-181), butter in four observational studies (22, 27, 31, 34), cheese in three interventional studies (182-184), milk in five interventional (178, 182-185) and seven observational studies (23, 27, 29, 34, 180, 181, 186). Galactonic acid (derived from galactose being oxidized via galactono-1,5-lactone) for milk (34, 182, 183) and 10-undecenoic acid (11:1n1) for butter intake (27, 31, 34) were identified to be the most frequently reported dietary biomarkers. A small number of other metabolites were also reported in higher concentration for these subcategories including galactose (182, 183), lactose (182, 183), galactono-1,5-lactone (182, 183), uridine (181, 186), and urea for milk (184, 185), and pentadecanoate (15:0) and methyl palmitic acid isomers for butter (27, 31). Additionally, studies reported higher concentration of 3-phenullactic (182, 183), proline (182, 183), and methionine (183, 184) for cheese intake, and pantothenic acid (vitamin B5) (179, 181) for dairy products.

2.4.9 Sweet and Sugary Foods

The subcategories of chocolate, dark chocolate, and sweet and sugary beverages had reported at least replicated metabolite (Table 2.1, Figure 2.4C). Metabolites for intake of chocolate were analyzed in two interventional (33, 52) and four observational studies (27, 31, 34, 36), dark chocolate in two interventional studies (187, 188), and sweet and sugary beverages in one interventional (189) and seven observational studies

(22, 24, 27, 189-192). While several metabolomic signatures were associated with intake of chocolate, theobromine (an alkaloid from cocoa plant) (27, 31, 33, 34, 52) followed by its endogenous metabolite 7-methyluric acid (33, 34, 52) were the most frequently reported metabolites. Meanwhile, 4-hydroxyphenyl was the only biomarker reported in higher concentration for intake of dark chocolate (187, 188). Further, citrulline, taurine, isocitrate, carbon isotopic signatures (δ13C) were reported in higher concentration after intake of sweet and sugary beverages (189-191). Various artificial sweeteners can also serve as specific/exogenous biomarkers reflecting intake of low caloric beverages and processed foods prevalent in a Western diet, including acesulfame K, aspartame, saccharin, sucralose, and steviol glycoside (193).

2.4.10 Complex Dietary Patterns and Other Foods

A number of dietary patterns and other food subcategories had identified metabolites that were replicated (Table 2.1, Figure 2.5A-B). Metabolites for intake of the Average Danish Diet (ADD) was analyzed in four interventional studies (33, 194-196), Dietary Approaches to Stop Hypertension (DASH) diet in one interventional (197) and two observational study (108, 198), Healthy Eating Index (HEI) in one interventional (96) and three observational studies (26, 108, 198), Mediterranean Diet in three interventional (199-201) and five observational studies (26, 108, 198, 202, 203), New Nordic Diet (NND) in four interventional studies (33, 194-196), vegetarian in three observational studies (100, 104, 204), vegan in one interventional (205) and three observational studies (100, 104, 112), high carotenoid in one interventional (206) and one observational study (207), fruits and vegetables in four interventional (69, 208-210) and

eight observational studies (28, 30, 76, 108, 133, 168, 176, 211), whey in three interventional studies (97, 185, 212), soy-based drink in three interventional (178, 182, 183) and two observational studies (31, 34), high-soy diet in two interventional (213, 214) and five observational studies (23, 76, 215-217), and soy-based cheese in one interventional (urine and plasma) study (218). Elevated levels of theobromine and proline betaine was reported for ADD (33, 194), β-cryptoxanthin for DASH Diet (108, 197),TMAO (33, 194, 196) and hippuric acid (33, 196) for NND, and CMPF (96, 108), eicosapentaenoic acid (20:5n–3) (96, 108), and hippuric acid (96, 198) for HEI. Meanwhile, higher levels of DHA (22:6n–3) was found to be associated with consumption of a Mediterranean Diet (26, 108, 199). Additionally, higher concentration of hippuric acid for fruits and vegetables (69, 176, 208), glycine for vegan (100, 104, 112), common dietary isoflavones daidzein and genistein for high-soy diet (23, 213-217) and soy-based cheese (218), and pinitol for soy-based drink (182, 183) were identified to be the most frequently reported markers.

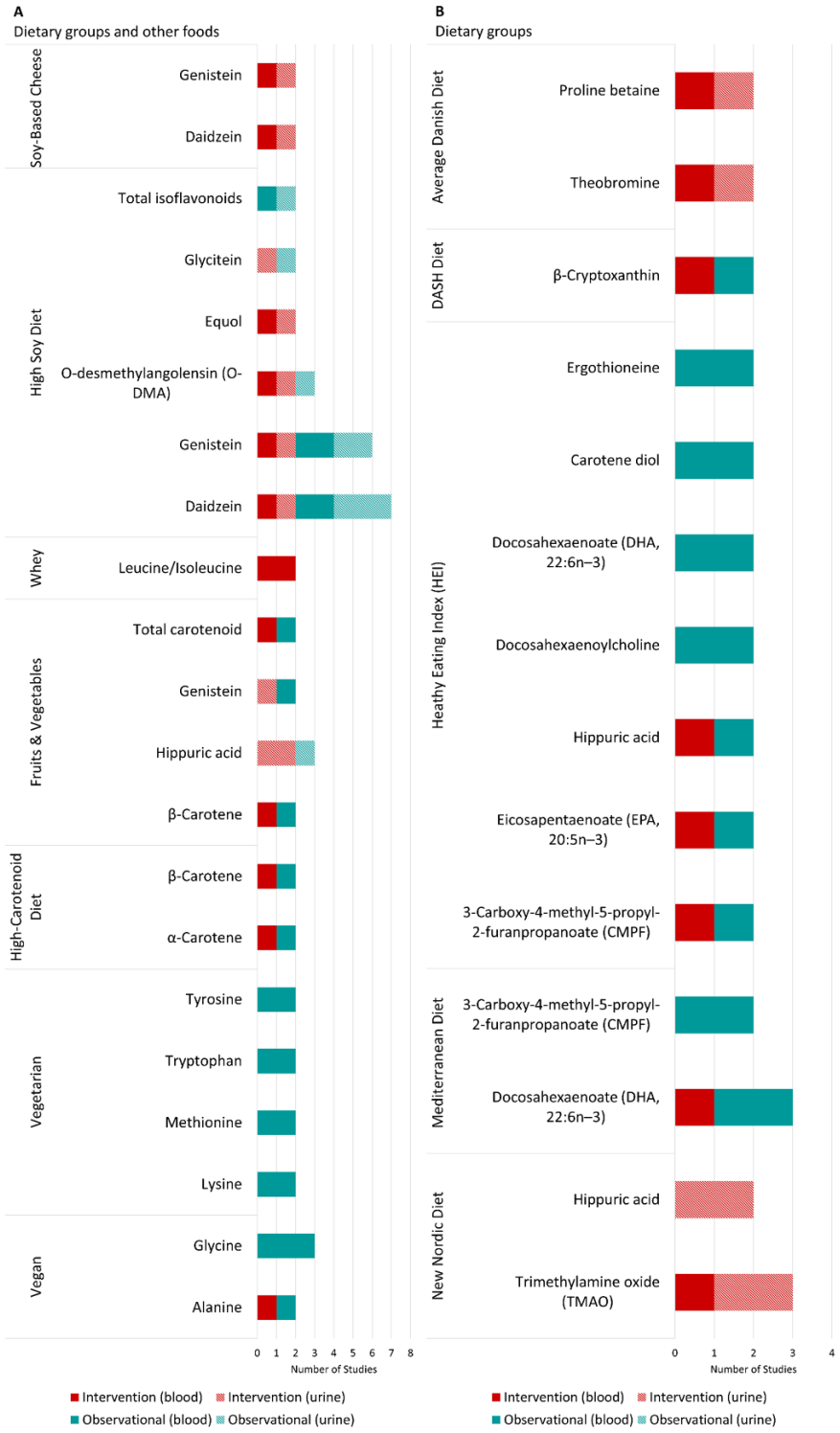

A
Dietary groups and other foods
Soy-Based Cheese
Genistein
Daidzein
High Soy Diet
Total isoflavonoids
Glycitein
Equol
O-desmethylangolensin (O-DMA)
Genistein
Daidzein
Whey
Leucine/Isoleucine
Fruits & Vegetables
Total carotenoid
Genistein
Hippuric acid
β-Carotene
High-Carotenoid Diet
β-Carotene
α-Carotene
Vegetarian
Tyrosine
Tryptophan
Methionine
Lysine
Vegan
Glycine
Alanine
0 1 2 3 4 5 6 7 8
Number of Studies
Intervention (blood)
Intervention (urine)
Observational (blood)
Observational (urine)
B
Dietary groups
Average Danish Diet
Proline betaine
Theobromine
DASH Diet
β-Cryptoxanthin
Healthy Eating Index (HEI)
Ergothioneine
Carotene diol
Docosahexaenoate (DHA, 22:6n–3)
Docosahexaenoylcholine
Hippuric acid
Eicosapentaenoate (EPA, 20:5n–3)
3-Carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
Mediterranean Diet
3-Carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
Docosahexaenoate (DHA, 22:6n–3)
New Nordic Diet
Hippuric acid
Trimethylamine oxide (TMAO)
0 1 2 3 4
Number of Studies
Intervention (blood)
Intervention (urine)
Observational (blood)
Observational (urine)

Figure 2.5: Metabolites identified from dietary patterns and other foods by number of studies, type of study design, and type of biofluid

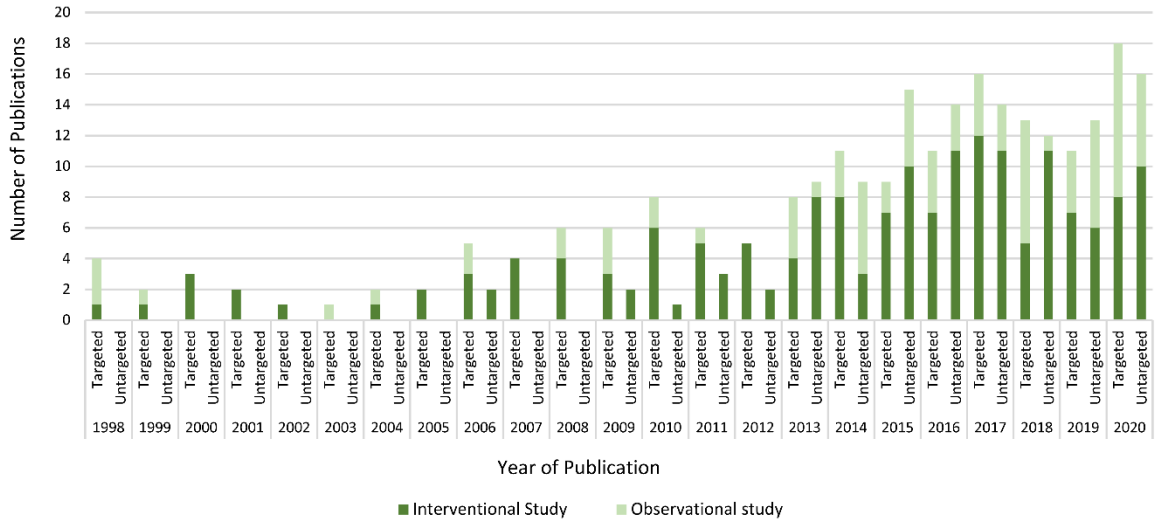


Figure 2.6: Number of publications in nutritional metabolomics

Note: Data presented is based on the inclusion criteria of this review

2.5 Discussion

This review included 244 articles (169 interventional studies, of which, 9 studies were replicated in free-living participants) that assessed the association between metabolites measured in common biofluids (i.e., urine, serum or plasma) and intake of individual food or food groups published between 1998–2020. Although, there has been a relatively long history of studies using a targeted approach to identify dietary biomarkers related to known constituents of food chemistry, the application of untargeted metabolomics only started to gain prominence in the mid-2000s and has greatly expanded in the past 5–10 years (Figure 2.6). Additionally, earlier studies have mainly been

interventional in design, but the number of observational studies has increased since the early 2000s. Given this trend and combined with recent advances in metabolomics, the application of metabolomics in nutritional epidemiology holds substantial promise.

2.5.1 Metabolites Associated with Foods or Food Groups

Based on our review, we rated the repeatability of 69 metabolites as good, 161 as fair, and 48 as poor markers of specific foods. Specifically, results from this review indicate that proline betaine for fruits in general, but also for orange, orange juice, and citrus fruit was the most repeatable (based on interstudy repeatability and study design). Additionally, pelargonidin glucuronide for strawberry; sulforaphane, sulforaphane cysteinylglycine, sulforaphane *N*-acetylcysteine, and sulforaphane cysteine for broccoli; sulforaphane for broccoli sprouts; alkylresorcinols for high fiber (grain-rich); creatinine for meat; 3-methylhistidine for chicken/poultry; acetylcarnitine for red meat and processed meat; DHA for seafood in general and fatty fish; TMAO for fish and NND; Σmethyl(epi)catechin glucuronides for red wine; methylgallic sulfate for dealcoholized red wine; 4-*O*-methylgallic acid for black tea; hippuric acid for green tea; 3-methylxanthine for cocoa; paraxanthine and caffeine for coffee; theobromine for chocolate; galactonate for milk; daidzein for high-soy diet; and hippuric acid for fruits and vegetables have good evidence and were also highly repeatable. This subset of metabolites is consistent with several previous reviews (18, 219-221).

It is important to mention that several other metabolites also had good evidence (i.e., 5 or more points) but were not found to be among the most repeated markers (i.e., not with the most points within the good category) (Table 1). For example, DHPPA for

high fiber (grain rich) foods appeared in two interventional studies (score: 2 x 2) and one observation study (score: 1 x 1), and we can classify this to be of good evidence (i.e., score ≥ 5) (Table 1, Figure 2B). However, since alkylresorcinols appeared in 4 interventional studies (score = 8) for the same food group, it was considered to be the "best" candidate metabolite (i.e., highest score).

2.5.2 Study Designs

All articles included in this review identified metabolites in human samples, and nearly 70 percent of the studies included were interventional in design (>50% of studies used a crossover design). Of the 69 metabolites with good evidence, 48 were reported in both interventional and observational studies, 20 were found only in interventional studies and one found only in observational studies. A cross-over design is ideal for assessing metabolites as participants act as their own control, which lowers variability due to physiological variation between individuals, lifestyle factors, and reporting bias (222). Most of the included interventional studies reflect short-to medium-term effects of diet or focus on a single food (e.g., orange juice) or food group (e.g., meats), whereas observational studies can be more informative as most have a large sample size (on average, 42 and 922 participants per interventional and observational study, respectively) and focus on multiple foods and/or food groups simultaneously. However, understanding the potential for biomarkers in observational designs is important because they are most likely to suffer from biases due to misreporting. Although we identified several markers in both study designs, there were a few markers that have not yet been identified in observational designs, likely because of lack of observational studies examining these

biomarkers. For example, while our review proposed isothiocyanates as a candidate dietary biomarker for broccoli consumption, it is yet to be studied in observational studies to assess their robustness (223). Having said this, we are confident that isothiocyanates may serve as a quantitative measure of short-term broccoli consumption since this biomarker was not associated with any other commonly consumed food. In contrast, we are less confident in a biomarker representing a particular food if that marker is yet to be identified in a free-living population and is not specific to a food (e.g., hippuric acid, a candidate marker for green tea, is also shown to be a candidate marker for fruits in general, fruits and vegetable, and coffee). In comparison, 3-methylhistidine as a biomarker for chicken/poultry consumption might serve as a valid marker for both short- and long-term intake as it was shown in both interventional and observational studies, with the conclusion that both study designs will provide important and unique information necessary to advance dietary biomarkers research (224).

2.5.3 Metabolomic Approaches

Of the 69 metabolites with good evidence, 38 were identified using both untargeted and targeted approaches. Some metabolites with good evidence were reported using only an untargeted approach (n=9), while others were reported using only a targeted approach (n=22) that benefits from the use of validated assays for their quantitative analysis. Though informative, a drawback of targeted analysis is that it aims to quantify *a priori* known subset of metabolites that are usually of related chemical structure and/or biological activity, and therefore discovery of novel markers cannot be achieved (8, 225). Meanwhile, an untargeted approach provides the broadest metabolite

coverage despite lengthy and complex post-analytical procedures for data filtering and unknown identification that are prone to bias or incomplete structural elucidation if not confirmed by mass spectral comparison and co-elution using an authentic standard. Nonetheless, there are potentially yet to be discovered metabolites that may be better indicators for some food groups. For these reasons, whenever possible, both approaches should be applied.

2.5.4 Analytical Techniques

All but one (DHPPA for high-fiber foods – using HPLC without MS) metabolite with good evidence were identified using LC-MS, 37 were identified using GC-MS, and 17 were identified using ^{1}H NMR. Less than one-fifth of studies in this review employed cross-platform metabolomic analysis likely due to costs, volume requirements, and throughput constraints. Moreover, due to the complexity of the metabolome, it is not possible to analyze 'each all' metabolites present in a biological sample using one or more analytical techniques due to their wide dynamic range in concentration and diverse physiochemical properties. Additionally, many metabolites are derived from specific foods infrequently consumed in a population or present at low concentration levels below detection limits resulting in missing value inputs. For this reason, it is often necessary to perform sample workup procedures prior to analysis, such as solvent extractions for sample enrichment or background matrix cleanup, noting that a non-selective solvent for sample preparation is preferred for untargeted approach, while targeted approaches sometimes rely on sample preparation procedures optimized for specific chemical groups (226). The results of this review showed that more than half of the food-specific

metabolites with good evidence were reported using at least two independent analytical platforms with acceptable mutual agreement (bias <10%) in measured concentrations, such as urinary iodide (227).

2.5.5 Concordance Between Biological Samples

A greater number of studies in this review were based on the analysis of blood (plasma or serum) than urine sample. Notwithstanding that 59 of the 69 food metabolites with good evidence were replicated in both blood and urine sample, DHA for seafood and fatty fish and catechol sulfate for coffee were detected only in blood, and dimethylamine for fish, pelargonidin glucuronide for strawberry intake, hippuric acid for fruits and vegetables intake and all four metabolites for cocoa intake were detected only in urine. The answer to the critical question of which biological sample (urine or blood) best characterizes intake of these foods thus remains unclear, with some evidence suggesting urine to be the superior biological sample to study nutrient intake or to identify BFIs (8).

2.5.6 Understanding Discordance Between Biological Samples

While few studies in this review that had used both blood and urine samples identified the same metabolites in both biospecimens (42% of the metabolites with good evidence) such as acetylcarnitine for red meat, other studies using both samples did not always find similar results. Urine has higher levels of exogenous metabolites compared to blood, which may be either phytochemicals, xenobiotics, or chemical by-products of cooking (8). The non-nutrient compounds derived from food intake are converted into more polar metabolites to decrease their renal threshold and are thus readily excreted in

urine (228). This may explain why fewer metabolites are more likely to be found in blood, because blood carries many more non-polar lipids than urine. Urine is a noninvasive biofluid, and cheaper and easier to collect in repeat measures and large-scale studies (especially children) than blood for better adherence, and it reflects a wider range of dietary biomarkers and time window to assess recent food exposures, so it is often considered the preferred sample for identification of food metabolites (22).

The biological variance of metabolites in urine is generally much greater than blood and requires adjustment for hydration status (e.g., creatinine, osmolality, and specific gravity) when relying on single-point/random collections. In contrast, 24-hour urine sampling is ideal for better assessment of average food exposures in observational or nutritional intervention trials, such as DASH style diets (46), but it is more difficult to collect consistently in large populations. Further, excretion site can influence detection of metabolites. An example of this is the detection of catechol sulfate after coffee intake in blood but not in urine. Catechol, a derivative of coffee processing, is conjugated to catechol sulfate in plasma to facilitate absorption and is generally eliminated in feces (229).

Finally, it is important to consider the time period during which the biological sample is collected; and the storage condition of the sample. Most food-specific metabolites are present in human blood and urine for approximately 5-10 hours, with some extending to 48-hours (230). Again, whenever possible, it is recommended to use a 24-48-hour model where multiple biological samples are collected and integrated over this longer time period to examine change in metabolite concentration over time or to

obtain an average value to represent "true" concentration. Typically, metabolite concentrations change rapidly in blood relative to urine biofluid, so the use of non-fasting sample adds heterogeneity to the results, but some biomarkers are best measured postprandially. Additionally, another potential influencing factor in metabolomics may be introduced with improper storage conditions (i.e., temperature, light, or duration), which may possibly lead to metabolite degradation or oxidation such as polyunsaturated fatty acids (PUFAs). There are also concerns of chemical stability if urine samples are not frozen promptly, and thus require the use of preserving agents such as sodium azide or boric acid to prevent bacterial growth (231).

2.6 Strengths and Limitations

A major strength of this review is that it provides a detailed and concise summary of all nutritional metabolomic studies reporting metabolites associated with individual foods and food groups that were conducted in healthy participants. We also provided a set of objective, transparent criteria for evaluating repeatability.

However, this review has a few limitations. First, we focused only on blood or urine metabolites and excluded studies using other less common biological samples, such as adipose tissue, feces, breath condensates, and saliva. Second, we were unable to conduct a quantitative analysis due to the variability in metabolite targets and approaches among studies, which makes it challenging to directly compare metabolite concentrations across studies. In addition, the variability in the portion size of foods and/or frequency of food intake (e.g., once versus repeated) can impose an important limitation when aiming to synthesize and integrate results from individual studies. Third, because the purpose of

the review was to rate the evidence of biomarkers based on repeatability, other validation criteria (e.g., specificity) were not assessed in this review (223). Fourth, urine and plasma measured within the same study were each counted as separate investigations and given equal weight because the samples were collected independently with the added advantage for researchers to evaluate whether either specimen could be used due to sample availability. For instance, 3-methylhistidine and proline betaine were consistently demonstrated as robust dietary biomarkers of a Prudent diet in both single-spot urine and fasting plasma samples collected from the same participants, which were also associated with self-reported intake of protein and citrus fruit, respectively (12). However, this may have inflated the score for some of the biomarkers. While we strived towards a reasonable, accurate yet simple score, the score may be biased by the biomarker's physicochemical properties e.g., detection and concentration, where lower nanomolar-picomolar metabolites or less-readily ionizable compounds are less likely to be detected, and thus identified. Finally, only a limited number of labs have investigated biomarkers associated with food intake and therefore we were unable to examine interlaboratory variability as required for nutritional epidemiology.

2.7 Conclusion

This review has reviewed and summarized metabolites associated with all possible food and food groups. The results show that while many metabolites can be identified from a specific food, there are many cases where a single metabolite is a good indicator of food intake. Findings obtained from this review have important public health implications. Dietary advice is an important component of chronic disease prevention and

management. Identifying good metabolites associated with food intake in generally healthy populations is an integral step towards examining diet as a risk factor for chronic disease more objectively (232). We recommend that future studies validate these metabolites by using criteria developed by Dragsted and colleagues (223) that includes biological plausibility, dose-response, time-response, robustness, reliability, stability, analytical performance, and inter-laboratory reproducibility to further advance the use of BFIs in nutritional research.

2.8 References

1. Bingham SA. Biomarkers in nutritional epidemiology. Public health nutrition 2002;5(6a):821-7.
2. Freedman LS, Potischman N, Kipnis V, Midthune D, Schatzkin A, Thompson FE, Troiano RP, Prentice R, Patterson R, Carroll R. A comparison of two dietary instruments for evaluating the fat–breast cancer relationship. International journal of Epidemiology 2006;35(4):1011-21.
3. Kipnis V, Midthune D, Freedman L, Bingham S, Day NE, Riboli E, Ferrari P, Carroll RJ. Bias in dietary-report instruments and its implications for nutritional epidemiology. Public health nutrition 2002;5(6a):915-23.
4. Rennie KL, Coward A, Jebb SA. Estimating under-reporting of energy intake in dietary surveys using an individualised method. British Journal of Nutrition 2007;97(6):1169-76.
5. Poslusna K, Ruprich J, de Vries JH, Jakubikova M, van't Veer P. Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice. British Journal of Nutrition 2009;101(S2):S73-S85.
6. Llorach R, Garcia-Aloy M, Tulipani S, Vazquez-Fresno R, Andres-Lacueva C. Nutrimetabolomic Strategies To Develop New Biomarkers of Intake and Health Effects. Journal of Agricultural and Food Chemistry 2012;60(36):8797-808. doi: 10.1021/jf301142b.
7. Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, Rappaport SM, van der Hooft JJ, Wishart DS. The food metabolome: a window over dietary exposure. The American Journal of Clinical Nutrition 2014;99(6):1286-308. doi: 10.3945/ajcn.113.076133.
8. Wishart DS. Metabolomics: applications to food science and nutrition research. Trends in food science & technology 2008;19(9):482-93.
9. Papandreou C, Moré M, Bellamine A. Trimethylamine N-Oxide in Relation to Cardiometabolic Health—Cause or Effect? Nutrients 2020;12(5):1330.
10. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, DuGar B, Feldstein AE, Britt EB, Fu X, Chung Y-M, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature 2011;472(7341):57-63. doi: 10.1038/nature09922.
11. Maruvada P, Lampe JW, Wishart DS, Barupal D, Chester DN, Dodd D, Djoumbou-Feunang Y, Dorrestein PC, Dragsted LO, Draper J. Perspective: dietary biomarkers of intake and exposure—exploration with omics approaches. Advances in Nutrition 2020;11(2):200-15.
12. Wellington N, Shanmuganathan M, de Souza RJ, Zulyniak MA, Azab S, Bloomfield J, Mell A, Ly R, Desai D, Anand SS, et al. Metabolic Trajectories Following Contrasting Prudent and Western Diets from Food Provisions: Identifying Robust Biomarkers of Short-Term Changes in Habitual Diet. Nutrients 2019;11(10). doi: 10.3390/nu11102407.

13. Azab SM, de Souza RJ, Teo KK, Anand SS, Williams NC, Holzschuher J, McGlory C, Philips SM, Britz-McKibbin P. Serum nonesterified fatty acids have utility as dietary biomarkers of fat intake from fish, fish oil, and dairy in women. J Lipid Res 2020;61(6):933-44. doi: 10.1194/jlr.D120000630.
14. Mahieu NG, Patti GJ. Systems-Level Annotation of a Metabolomics Data Set Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. Analytical Chemistry 2017;89(19):10397-406. doi: 10.1021/acs.analchem.7b02380.
15. Vinayavekhin N, Saghatelian A. Untargeted metabolomics. Current protocols in molecular biology 2010;90(1):30.1. 1-.1. 24.
16. Naveja JJ, Rico-Hidalgo MP, Medina-Franco JL. Analysis of a large food chemical database: chemical space, diversity, and complexity. F1000Research 2018;7.
17. Neveu V, Nicolas G, Salek RM, Wishart DS, Scalbert A. Exposome-Explorer 2.0: an update incorporating candidate dietary biomarkers and dietary associations with cancer risk. Nucleic acids research 2020;48(D1):D908-D12.
18. Rothwell JA, Madrid-Gambin F, Garcia-Aloy M, Andres-Lacueva C, Logue C, Gallagher AM, Mack C, Kulling SE, Gao Q, Praticò G, et al. Biomarkers of intake for coffee, tea, and sweetened beverages. Genes Nutr 2018;13:15. doi: 10.1186/s12263-018-0607-5.
19. Vázquez-Manjarrez N, Ulaszewska M, Garcia-Aloy M, Mattivi F, Praticò G, Dragsted LO, Manach C. Biomarkers of intake for tropical fruits. Genes Nutr 2020;15(1):11. doi: 10.1186/s12263-020-00670-4.
20. Mennen LI, Sapinho D, Ito H, Bertrais S, Galan P, Hercberg S, Scalbert A. Urinary flavonoids and phenolic acids as biomarkers of intake for polyphenol-rich foods. British Journal of Nutrition 2006;96(1):191-8. doi: 10.1079/bjn20061808.
21. Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas ME, Kochhar S, Stamler J, Holmes E, Elliott P, Nicholson JK. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. Am J Clin Nutr 2010;92(2):436-43. doi: 10.3945/ajcn.2010.29672.
22. Playdon MC, Sampson JN, Cross AJ, Sinha R, Guertin KA, Moy KA, Rothman N, Irwin ML, Mayne ST, Stolzenberg-Solomon R, et al. Comparing metabolite profiles of habitual diet in serum and urine. Am J Clin Nutr 2016;104(3):776-89. doi: 10.3945/ajcn.116.135301.
23. Frankenfeld CL. Dairy consumption is a significant correlate of urinary equol concentration in a representative sample of US adults. Am J Clin Nutr 2011;93(5):1109-16. doi: 10.3945/ajcn.111.011825.
24. Lau CE, Siskos AP, Maitre L, Robinson O, Athersuch TJ, Want EJ, Urquiza J, Casas M, Vafeiadi M, Roumeliotaki T, et al. Determinants of the urinary and serum metabolome in children from six European populations. BMC Medicine 2018;16(1):202.
25. Oude Griep LM, Chekmeneva E, Stamler J, Van Horn L, Chan Q, Ebbels TMD, Holmes E, Frost GS, Elliott P. Urinary hippurate and proline betaine relative to fruit intake, blood pressure, and body mass index. Proceedings of the Nutrition Society 2016;75 (OCE3):E178.

26. Playdon MC, Moore SC, Derkach A, Reedy J, Subar AF, Sampson JN, Albanes D, Fangyi G, Kontto J, Lassale C, et al. Identifying biomarkers of dietary patterns by using metabolomics. Am J Clin Nutr 2017;105(2):450-65. doi: 10.3945/ajcn.116.144501.
27. Guertin KA, Moore SC, Sampson JN, Huang W-Y, Xiao Q, Stolzenberg-Solomon RZ, Sinha R, Cross AJ. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr 2014;100(1):208-17. doi: https://dx.doi.org/10.3945/ajcn.113.078758.
28. 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-. doi: 10.1186/1475-2891-12-158.
29. Toffano RBD, Hillesheim E, Mathias MG, Coelho-Landell CA, Salomao RG, Almada MORV, Camarneiro JM, Barros TT, Camelo-Junior JS, Rezzi S, et al. Validation of the brazilian healthy eating index-revised using biomarkers in children and adolescents. Nutrients 2018;10 (2) (no pagination)(154).
30. Zheng Y, Yu B, Alexander D, Steffen LM, Boerwinkle E. Human Metabolome Associates With Dietary Intake Habits Among African Americans in the Atherosclerosis Risk in Communities Study. Am J Epidemiol 2014;179(12):1424-33. doi: aje/kwu073.
31. Pallister T, Jennings A, Mohney RP, Yarand D, Mangino M, Cassidy A, MacGregor A, Spector TD, Menni C. Characterizing Blood Metabolomics Profiles Associated with Self-Reported Food Intakes in Female Twins. Plos One 2016;11(6). doi: 10.1371/journal.pone.0158568.
32. Pujos-Guillot E, Hubert J, Martin JF, Lyan B, Quintana M, Claude S, Chabanas B, Rothwell JA, Bennetau-Pelissero C, Scalbert A, et al. Mass Spectrometry-based Metabolomics for the Discovery of Biomarkers of Fruit and Vegetable Intake: Citrus Fruit as a Case Study. Journal of Proteome Research 2013;12(4):1645-59. doi: 10.1021/pr300997c.
33. Andersen MBS, Rinnan A, Manach C, Poulsen SK, Pujos-Guillot E, Larsen TM, Astrup A, Dragsted LO. Untargeted Metabolomics as a Screening Tool for Estimating Compliance to a Dietary Pattern. Journal of Proteome Research 2014;13(3):1405-18. doi: 10.1021/pr400964s.
34. Wang Y, Gapstur SM, Carter BD, Hartman TJ, Stevens VL, Gaudet MM, McCullough ML. Untargeted Metabolomics Identifies Novel Potential Biomarkers of Habitual Food Intake in a Cross-Sectional Study of Postmenopausal Women. Journal of Nutrition 2018;148(6):932-43. doi: 10.1093/jn/nxy027.
35. Noh H, Freisling H, Assi N, Zamora-Ros R, Achaintre D, Affret A, Mancini F, Boutron-Ruault MC, Flogel A, Boeing H, et al. Identification of Urinary Polyphenol Metabolite Patterns Associated with Polyphenol-Rich Food Intake in Adults from Four European Countries. Nutrients 2017;9(8):25.
36. Edmands WM, Ferrari P, Rothwell JA, Rinaldi S, Slimani N, Barupal DK, Biessy C, Jenab M, Clavel-Chapelon F, Fagherazzi G, et al. Polyphenol metabolome in

human urine and its association with intake of polyphenol-rich foods across European countries. Am J Clin Nutr 2015;102(4):905-13.

37. Lacalle-Bergeron L, Portolés T, López FJ, Sancho JV, Ortega-Azorín C, Asensio EM, Coltell O, Corella D. Ultra-Performance Liquid Chromatography-Ion Mobility Separation-Quadruple Time-of-Flight MS (UHPLC-IMS-QTOF MS) Metabolomics for Short-Term Biomarker Discovery of Orange Intake: A Randomized, Controlled Crossover Study. Nutrients 2020;12(7):1916. doi: 10.3390/nu12071916.
38. Rangel-Huerta OD, Aguilera CM, Perez-de-la-Cruz A, Vallejo F, Tomas-Barberan F, Gil A, Mesa MD. A serum metabolomics-driven approach predicts orange juice consumption and its impact on oxidative stress and inflammation in subjects from the BIONAOS study. Molecular Nutrition & Food Research 2017;61(2). doi: 10.1002/mnfr.201600120.
39. Schär MY, Curtis PJ, Hazim S, Ostertag LM, Kay CD, Potter JF, Cassidy A. Orange juice–derived flavanone and phenolic metabolites do not acutely affect cardiovascular risk biomarkers: a randomized, placebo-controlled, crossover trial in men at moderate risk of cardiovascular disease. Am J Clin Nutr 2015;101(5):931-8. doi: 10.3945/ajcn.114.104364.
40. Gibbons H, Michielsen CJR, Rundle M, Frost G, McNulty BA, Nugent AP, Walton J, Flynn A, Gibney MJ, Brennan L. Demonstration of the utility of biomarkers for dietary intake assessment; proline betaine as an example. Molecular Nutrition & Food Research 2017;61(10). doi: 10.1002/mnfr.201700037.
41. Pereira-Caro G, Polyviou T, Ludwig IA, Nastase AM, Moreno-Rojas JM, Garcia AL, Malkova D, Crozier A. Bioavailability of orange juice (poly)phenols: the impact of short-term cessation of training by male endurance athletes. Am J Clin Nutr 2017;106(3):791-800.
42. Pereira-Caro G, Borges G, van der Hooft J, Clifford MN, Del Rio D, Lean ME, Roberts SA, Kellerhals MB, Crozier A. Orange juice (poly)phenols are highly bioavailable in humans. Am J Clin Nutr 2014;100(5):1378-84.
43. Pereira-Caro G, Clifford MN, Polyviou T, Ludwig IA, Alfheeaid H, Moreno-Rojas JM, Garcia AL, Malkova D, Crozier A. Plasma pharmacokinetics of (poly)phenol metabolites and catabolites after ingestion of orange juice by endurance trained men. Free Radical Biology and Medicine 2020;160:784-95. doi: http://dx.doi.org/10.1016/j.freeradbiomed.2020.09.007.
44. Lee J, Ebeler SE, Zweigenbaum JA, Mitchell AE. UHPLC-(ESI)QTOF MS/MS profiling of quercetin metabolites in human plasma postconsumption of applesauce enriched with apple peel and onion. Journal of Agricultural & Food Chemistry 2012;60(34):8510-20.
45. McNamara AE, Collins C, Harsha PSCS, Gonzalez-Pena D, Gibbons H, McNulty BA, Nugent AP, Walton J, Flynn A, Brennan L. Metabolomic-Based Approach to Identify Biomarkers of Apple Intake. Molecular nutrition & food research 2020;64(11):e1901158. doi: http://dx.doi.org/10.1002/mnfr.201901158.

46. Reisdorph NA, Hendricks AE, Tang M, Doenges KA, Reisdorph RM, Tooker BC, Quinn K, Borengasser SJ, Nkrumah-Elie Y, Frank DN, et al. Nutrimetabolomics reveals food-specific compounds in urine of adults consuming a DASH-style diet. Scientific reports 2020;10(1):1157. doi: https://dx.doi.org/10.1038/s41598-020-57979-8.
47. Ulaszewska MM, Koutsos A, Trošt K, Stanstrup J, Garcia-Aloy M, Scholz M, Fava F, Natella F, Scaccini C, Vrhovsek U, et al. Two apples a day modulate human:microbiome co-metabolic processing of polyphenols, tyrosine and tryptophan. European Journal of Nutrition 2020;59(8):3691-714. doi: 10.1007/s00394-020-02201-8.
48. Vázquez-Manjarrez N, Weinert CH, Ulaszewska MM, Mack CI, Micheau P, Pétéra M, Durand S, Pujos-Guillot E, Egert B, Mattivi F, et al. Discovery and Validation of Banana Intake Biomarkers Using Untargeted Metabolomics in Human Intervention and Cross-sectional Studies. Journal of Nutrition 2019;149(10):1701-13. doi: 10.1093/jn/nxz125.
49. Guertin KA, Moore SC, Sampson JN, Huang WY, Xiao Q, Stolzenberg-Solomon RZ, Sinha R, Cross AJ. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr 2014;100(1):208-17. doi: 10.3945/ajcn.113.078758.
50. Cuparencu CS, Andersen MBS, Gurdeniz G, Schou SS, Mortensen MW, Raben A, Astrup A, Dragsted LO. Identification of urinary biomarkers after consumption of sea buckthorn and strawberry, by untargeted LC-MS metabolomics: a meal study in adult men. Metabolomics 2016;12(2). doi: 10.1007/s11306-015-0934-0.
51. Carkeet C, Clevidence BA, Novotny JA. Anthocyanin excretion by humans increases linearly with increasing strawberry dose. Journal of Nutrition 2008;138(5):897-902.
52. Andersen MB, Kristensen M, Manach C, Pujos-Guillot E, Poulsen SK, Larsen TM, Astrup A, Dragsted L. Discovery and validation of urinary exposure markers for different plant foods by untargeted metabolomics. Analytical & Bioanalytical Chemistry 2014;406(7):1829-44.
53. Henning SM, Seeram NP, Zhang Y, Li L, Gao K, Lee R, Wang DC, Zerlin A, Karp H, Thames G, et al. Strawberry consumption is associated with increased antioxidant capacity in serum. Journal of Medicinal Food 2010;13(1):116-22. doi: 10.1089/jmf.2009.0048.
54. Feliciano RP, Mills CE, Istas G, Heiss C, Rodriguez-Mateos A. Absorption, Metabolism and Excretion of Cranberry (Poly)phenols in Humans: A Dose Response Study and Assessment of Inter-Individual Variability. Nutrients 2017;9(3):268. doi: 10.3390/nu9030268.
55. Liu H, Garrett TJ, Su Z, Khoo C, Gu L. UHPLC-Q-Orbitrap-HRMS-based global metabolomics reveal metabolome modifications in plasma of young women after cranberry juice consumption. Journal of Nutritional Biochemistry 2017;45:67-76.
56. Zhao S, Liu H, Su Z, Khoo C, Gu L. Identifying Cranberry Juice Consumers with Predictive OPLS-DA Models of Plasma Metabolome and Validation of Cranberry

Juice Intake Biomarkers in a Double-Blinded, Randomized, Placebo-Controlled, Cross-Over Study. Molecular nutrition & food research 2020;64(11):e1901242. doi: http://dx.doi.org/10.1002/mnfr.201901242.

57. Favari C, Mena P, Curti C, Istas G, Heiss C, Del Rio D, Rodriguez-Mateos A. Kinetic profile and urinary excretion of phenyl-gamma-valerolactones upon consumption of cranberry: a dose-response relationship. Food & function 2020;11(5):3975-85. doi: http://dx.doi.org/10.1039/d0fo00806k.
58. Lang R, Lang T, Bader M, Beusch A, Schlagbauer V, Hofmann T. High-Throughput Quantitation of Proline Betaine in Foods and Suitability as a Valid Biomarker for Citrus Consumption. Journal of Agricultural and Food Chemistry 2017;65(8):1613-9. doi: 10.1021/acs.jafc.6b05824.
59. Lloyd AJ, Fave G, Beckmann M, Lin WC, Tailliart K, Xie L, Mathers JC, Draper J. Use of mass spectrometry fingerprinting to identify urinary metabolites after consumption of specific foods. Am J Clin Nutr 2011;94(4):981-91. doi: 10.3945/ajcn.111.017921.
60. Baenas N, Suarez-Martinez C, Garcia-Viguera C, Moreno DA. Bioavailability and new biomarkers of cruciferous sprouts consumption. Food Res Int 2017;100(Part 1):497-503. doi: 10.1016/j.foodres.2017.07.049.
61. Conaway CC, Getahun SM, Liebes LL, Pusateri DJ, Topham DK, Botero-Omary M, Chung F. Disposition of Glucosinolates and Sulforaphane in Humans After Ingestion of Steamed and Fresh Broccoli. Nutrition & Cancer 2000;38(2):168-78.
62. Hauder J, Winker S, Bub A, Ruifer CE, Pignitter M, Somoza V. LC-MS/MS Quantification of Sulforaphane and Indole-3-carbinol Metabolites in Human Plasma and Urine after Dietary Intake of Selenium-Fortified Broccoli. Journal of Agricultural and Food Chemistry 2011;59(15):8047-57. doi: 10.1021/jf201501x.
63. Charron CS, Vinyard BT, Jeffery EH, Ross SA, Seifried HE, Novotny JA. BMI Is Associated With Increased Plasma and Urine Appearance of Glucosinolate Metabolites After Consumption of Cooked Broccoli. Frontiers in Nutrition 2020;7. doi: 10.3389/fnut.2020.575092.
64. Egner PA, Chen JG, Wang JB, Wu Y, Sun Y, Lu JH, Zhu J, Zhang YH, Chen YS, Friesen MD, et al. Bioavailability of Sulforaphane from two broccoli sprout beverages: results of a short-term, cross-over clinical trial in Qidong, China. Cancer Prevention Research 2011;4(3):384-95.
65. Clarke JD, Hsu A, Riedl K, Bella D, Schwartz SJ, Stevens JF, Ho E. Bioavailability and inter-conversion of sulforaphane and erucin in human subjects consuming broccoli sprouts or broccoli supplement in a cross-over study design. Pharmacological Research 2011;64(5):456-63.
66. Housley L, Magana AA, Hsu A, Beaver LM, Wong CP, Stevens JF, Choi J, Jiang Y, Bella D, Williams DE, et al. Untargeted Metabolomic Screen Reveals Changes in Human Plasma Metabolite Profiles Following Consumption of Fresh Broccoli Sprouts. Molecular Nutrition & Food Research 2018;62(19). doi: 10.1002/mnfr.201700665.

67. Villano D, Lopez-Chillon MT, Zafrilla P, Moreno DA. Bioavailability of broccoli sprouts in different human overweight populations. Journal of Functional Foods 2019;59:337-44. doi: 10.1016/j.jff.2019.05.052.
68. Kristensen M, Krogholm KS, Frederiksen H, Bugel SH, Rasmussen SE. Urinary excretion of total isothiocyanates from cruciferous vegetables shows high dose-response relationship and may be a useful biomarker for isothiocyanate exposure. European Journal of Nutrition 2007;46(7):377-82.
69. Garcia-Perez I, Posma JM, Gibson R, Chambers ES, Hansen TH, Vestergaard H, Hansen T, Beckmann M, Pedersen O, Elliott P, et al. Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. Lancet Diabetes & Endocrinology 2017;5(3):184-95. doi: 10.1016/s2213-8587(16)30419-3.
70. Johansson-Persson A, Barri T, Ulmius M, Onning G, Dragsted LO. LC-QTOF/MS metabolomic profiles in human plasma after a 5-week high dietary fiber intake. Analytical and Bioanalytical Chemistry 2013;405(14):4799-809. doi: 10.1007/s00216-013-6874-5.
71. Ross AB, Pere-Trépat E, Montoliu I, Martin F-PJ, Collino S, Moco S, Godin J-P, Cléroux M, Guy PA, Breton I, et al. A whole-grain-rich diet reduces urinary excretion of markers of protein catabolism and gut microbiota metabolism in healthy men after one week. Journal of Nutrition 2013;143(6):766-73. doi: 10.3945/jn.112.172197.
72. McKeown N, Marklund M, Ma J, Ross A, Lichtenstein A, Livingston K, Jacques P, Rasmussen H, Blumberg J, Chen CY. Comparison of plasma alkylresorcinols (AR) and urinary AR metabolites as biomarkers of compliance in a short-term, whole-grain intervention study. European Journal of Nutrition 2016;55(3):1235-44. doi: 10.1007/s00394-015-0936-8.
73. Ross AB, Bourgeois A, Macharia HN, Kochhar S, Jebb SA, Brownlee IA, Seal CJ. Plasma alkylresorcinols as a biomarker of whole-grain food consumption in a large population: results from the WHOLEheart Intervention Study. Am J Clin Nutr 2012;95(1):204-11.
74. Hanhineva K, Keski-Rahkonen P, Lappi J, Katina K, Pekkinen J, Savolainen O, Timonen O, Paananen J, Mykkanen H, Poutanen K. The Postprandial Plasma Rye Fingerprint Includes Benzoxazinoid-Derived Phenylacetamide Sulfates. Journal of Nutrition 2014;144(7):1016-22. doi: 10.3945/jn.113.187237.
75. Cuff J, Sanders TAB, Reidlinger DP, Hall WL, Gray R, Darzi J. Urinary alkylresorcinol metabolites as a biomarker of dietary wholegrain intake and of compliance in a randomised dietary intervention trial: results from the CRESSIDA Study. Proceedings of the Nutrition Society 2015;74(OCE1):E42. doi: 10.1017/s0029665115000579.
76. Frankenfeld CL, Patterson RE, Horner NK, Neuhouser ML, Skor HE, Kalhorn TF, Howald WN, Lampe JW. Validation of a soy food-frequency questionnaire and evaluation of correlates of plasma isoflavone concentrations in postmenopausal women. Am J Clin Nutr 2003;77(3):674-80.

77. Aubertin-Leheudre M, Koskela A, Samaletdin A, Adlercreutz H. Plasma alkylresorcinol metabolites as potential biomarkers of whole-grain wheat and rye cereal fibre intakes in women. British Journal of Nutrition 2010;103(3):339-43. doi: 10.1017/s0007114509992315.
78. Landberg R, Wierzbicka R, Shi L, Nybacka S, Kamal-Eldin A, Hedblad B, Lindroos AK, Winkvist A, Forslund HB. New alkylresorcinol metabolites in spot urine as biomarkers of whole grain wheat and rye intake in a Swedish middle-aged population. Eur J Clin Nutr 2018;72(10):1439-46. doi: 10.1038/s41430-017-0079-5.
79. Garcia-Aloy M, Llorach R, Urpi-Sarda M, Tulipani S, Salas-Salvado J, Angel Martinez-Gonzalez M, Corella D, Fito M, Estruch R, Serra-Majem L, et al. Nutrimetabolomics fingerprinting to identify biomarkers of bread exposure in a free-living population from the PREDIMED study cohort. Metabolomics 2015;11(1):155-65. doi: 10.1007/s11306-014-0682-6.
80. Bondia-Pons I, Barri T, Hanhineva K, Juntunen K, Dragsted LO, Mykkanen H, Poutanen K. UPLC-QTOF/MS metabolic profiling unveils urinary changes in humans after a whole grain rye versus refined wheat bread intervention. Molecular Nutrition & Food Research 2013;57(3):412-22. doi: 10.1002/mnfr.201200571.
81. Landberg R, Aman P, Friberg LE, Vessby B, Adlercreutz H, Kamal-Eldin A. Dose response of whole-grain biomarkers: alkylresorcinols in human plasma and their metabolites in urine in relation to intake. Am J Clin Nutr 2009;89(1):290-6. doi: 10.3945/ajcn.2008.26709.
82. Lankinen M, Schwab U, Seppanen-Laakso T, Mattila I, Juntunen K, Mykkanen H, Poutanen K, Gylling H, Oresic M. Metabolomic Analysis of Plasma Metabolites That May Mediate Effects of Rye Bread on Satiety and Weight Maintenance in Postmenopausal Women. Journal of Nutrition 2011;141(1):31-6. doi: 10.3945/jn.110.131656.
83. Lappi J, Aura A-M, Katina K, Nordlund E, Kolehmainen M, Mykkanen H, Poutanen K. Comparison of postprandial phenolic acid excretions and glucose responses after ingestion of breads with bioprocessed or native rye bran. Food & function 2013;4(6):972-81. doi: https://dx.doi.org/10.1039/c3fo60078e.
84. Moazzami AA, Bondia-Pons I, Hanhineva K, Juntunen K, Antl N, Poutanen K, Mykkanen H. Metabolomics reveals the metabolic shifts following an intervention with rye bread in postmenopausal women- a randomized control trial. Nutrition Journal 2012;11. doi: 10.1186/1475-2891-11-88.
85. Shi L, Brunius C, Lindelof M, Abou Shameh S, Wu HX, Lee I, Landberg R, Moazzami AA. Targeted metabolomics reveals differences in the extended postprandial plasma metabolome of healthy subjects after intake of whole-grain rye porridges versus refined wheat bread. Molecular Nutrition & Food Research 2017;61(7). doi: 10.1002/mnfr.201600924.
86. Hanhineva K, Brunius C, Andersson A, Marklund M, Juvonen R, Keski-Rahkonen P, Auriola S, Landberg R. Discovery of urinary biomarkers of whole grain rye intake in free-living subjects using nontargeted LC-MS metabolite

profiling. Molecular Nutrition & Food Research 2015;59(11):2315-25. doi: 10.1002/mnfr.201500423.
87. Verkasalo PK, Appleby PN, Allen NE, Davey G, Adlercreutz H, Key TJ. Soya intake and plasma concentrations of daidzein and genistein: validity of dietary assessment among eighty British women (Oxford arm of the European Prospective Investigation into Cancer and Nutrition). Br J Nutr 2001;86(3):415-21. doi: 10.1079/bjn2001424.
88. Markhus MW, Graff IE, Dahl L, Seldal CF, Skotheim S, Braarud HC, Stormark KM, Malde MK. Establishment of a seafood index to assess the seafood consumption in pregnant women. Food & Nutrition Research 2013;57:1-11. doi: 10.3402/fnr.v57i0.19272.
89. Meuronen T, Lankinen MA, Fauland A, Shimizu BI, de Mello VD, Laaksonen DE, Wheelock CE, Erkkila AT, Schwab US. Intake of Camelina Sativa Oil and Fatty Fish Alter the Plasma Lipid Mediator Profile in Subjects with Impaired Glucose Metabolism - A Randomized Controlled Trial. Prostaglandins Leukotrienes and Essential Fatty Acids 2020;159:102143. doi: http://dx.doi.org/10.1016/j.plefa.2020.102143.
90. Thiébaut ACM, Rotival M, Gauthier E, Lenoir GM, Boutron-Ruault M, Joulin V, Clavel-Chapelon F, Chajès V. Correlation between serum phospholipid fatty acids and dietary intakes assessed a few years earlier. Nutrition & Cancer 2009;61(4):500-9. doi: 10.1080/01635580802710717.
91. Philibert A, Vanier C, Abdelouahab N, Chan HM, Mergler D. Fish intake and serum fatty acid profiles from freshwater fish. Am J Clin Nutr 2006;84(6):1299-307.
92. Hustad KS, Rundblad A, Ottestad I, Christensen JJ, Holven KB, Ulven SM. Comprehensive lipid and metabolite profiling in healthy adults with low and high consumption of fatty fish: A cross-sectional study. British Journal of Nutrition 2020. doi: http://dx.doi.org/10.1017/S0007114520002305.
93. Ross AB, Svelander C, Undeland I, Pinto R, Sandberg A-S. 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. Journal of Nutrition 2015;145(11):2456-63. doi: 10.3945/jn.115.214262.
94. Cheung W, Keski-Rahkonen P, Assi N, Ferrari P, Freisling H, Rinaldi S, Slimani N, Zamora-Ros R, Rundle M, Frost G, et al. A metabolomic study of biomarkers of meat and fish intake. Am J Clin Nutr 2017;105(3):600-8. doi: 10.3945/ajcn.116.146639.
95. Chiang YL, Haddad E, Rajaram S, Shavlik D, Sabate J. The effect of dietary walnuts compared to fatty fish on eicosanoids, cytokines, soluble endothelial adhesion molecules and lymphocyte subsets: a randomized, controlled crossover trial. Prostaglandins Leukotrienes and Essential Fatty Acids 2012;87(4-5):111-7. doi: 10.1016/j.plefa.2012.07.007.
96. Hanhineva K, Lankinen MA, Pedret A, Schwab U, Kolehmainen M, Paananen J, de Mello V, Sola R, Lehtonen M, Poutanen K, et al. Nontargeted Metabolite

Profiling Discriminates Diet-Specific Biomarkers for Consumption of Whole Grains, Fatty Fish, and Bilberries in a Randomized Controlled Trial. Journal of Nutrition 2015;145(1):7-17. doi: 10.3945/jn.114.196840.

97. Stanstrup J, Schou SS, Holmer-Jensen J, Hermansen K, Dragsted LO. Whey Protein Delays Gastric Emptying and Suppresses Plasma Fatty Acids and Their Metabolites Compared to Casein, Gluten, and Fish Protein. Journal of Proteome Research 2014;13(5):2396-408. doi: 10.1021/pr401214w.
98. Hagen IV, Helland A, Bratlie M, Midttun Ø, McCann A, Sveier H, Rosenlund G, Mellgren G, Ueland PM, Gudbrandsen OA. TMAO, creatine and 1-methylhistidine in serum and urine are potential biomarkers of cod and salmon intake: a randomised clinical trial in adults with overweight or obesity. European Journal of Nutrition 2020;59(5):2249-59. doi: 10.1007/s00394-019-02076-4.
99. Yin X, Gibbons H, Rundle M, Frost G, McNulty BA, Nugent AP, Walton J, Flynn A, Brennan L. The Relationship between Fish Intake and Urinary Trimethylamine-N-Oxide. Molecular nutrition & food research 2020;64(3):e1900799. doi: http://dx.doi.org/10.1002/mnfr.201900799.
100. Schmidt JA, Rinaldi S, Scalbert A, Ferrari P, Achaintre D, Gunter MJ, Appleby PN, Key TJ, Travis RC. Plasma concentrations and intakes of amino acids in male meat-eaters, fish-eaters, vegetarians and vegans: a cross-sectional analysis in the EPIC-Oxford cohort. Eur J Clin Nutr 2016;70(3):306-12. doi: 10.1038/ejcn.2015.144.
101. O'Gorman A, Morris C, Ryan M, O'Grada CM, Roche HM, Gibney ER, Gibney MJ, Brennan L. Habitual dietary intake impacts on the lipidomic profile. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences 2014;966:140-6. doi: 10.1016/j.jchromb.2014.01.032.
102. Reeves JL, Otahal P, Magnussen CG, Dwyer T, Kangas AJ, Soininen P, Ala-Korpela M, Venn AJ, Smith KJ. DHA mediates the protective effect of fish consumption on new episodes of depression among women. British Journal of Nutrition 2017;118(9):743-9. doi: 10.1017/s0007114517002768.
103. Baylin A, Kabagame EK, Siles X, Campos H. Adipose tissue biomarkers of fatty acid intake. Am J Clin Nutr 2002;76(4):750-7.
104. Schmidt JA, Rinaldi S, Ferrari P, Carayol M, Achaintre D, Scalbert A, Cross AJ, Gunter MJ, Fensom GK, Appleby PN, et al. Metabolic profiles of male meat eaters, fish eaters, vegetarians, and vegans from the EPIC-Oxford cohort. Am J Clin Nutr 2015;102(6):1518-26. doi: 10.3945/ajcn.115.111989.
105. Gibson R, Lau C-HE, Loo RL, Ebbels TMD, Chekmeneva E, Dyer AR, Miura K, Ueshima H, Zhao L, Daviglus ML, et al. The association of fish consumption and its urinary metabolites with cardiovascular risk factors: the International Study of Macro-/Micronutrients and Blood Pressure (INTERMAP). American Journal of Clinical Nutrition 2020;111(2):280-90. doi: 10.1093/ajcn/nqz293.
106. Schmedes M, Aadland EK, Sundekilde UK, Jacques H, Lavigne C, Graff IE, Eng O, Holthe A, Mellgren G, Young JF, et al. Lean-seafood intake decreases urinary markers of mitochondrial lipid and energy metabolism in healthy subjects:

Metabolomics results from a randomized crossover intervention study. Molecular Nutrition & Food Research 2016;60(7):1661-72. doi: 10.1002/mnfr.201500785.
107. Schmedes M, Balderas C, Aadland EK, Jacques H, Lavigne C, Graff IE, Eng O, Holthe A, Mellgren G, Young JF, et al. The Effect of Lean-Seafood and Non-Seafood Diets on Fasting and Postprandial Serum Metabolites and Lipid Species: Results from a Randomized Crossover Intervention Study in Healthy Adults. Nutrients 2018;10(5). doi: 10.3390/nu10050598.
108. McCullough ML, Maliniak ML, Stevens VL, Carter BD, Hodge RA, Wang Y. Metabolomic markers of healthy dietary patterns in US postmenopausal women. Am J Clin Nutr 2019;109(5):1439-51. doi: 10.1093/ajcn/nqy385.
109. Lord RS, Bralley JA. Clinical applications of urinary organic acids. Part I: Detoxification markers. Altern Med Rev 2008;13(3):205-15.
110. Stella C, Beckwith-Hall B, Cloarec O, Holmes E, Lindon JC, Powell J, van der Ouderaa F, Bingham S, Cross AJ, Nicholson JK. Susceptibility of human metabolic phenotypes to dietary modulation. Journal of Proteome Research 2006;5(10):2780-8. doi: 10.1021/pr060265y.
111. Allen NE, Grace PB, Ginn A, Travis RC, Roddam AW, Appleby PN, Key T. Phytanic acid: measurement of plasma concentrations by gas-liquid chromatography-mass spectrometry analysis and associations with diet and other plasma fatty acids. British Journal of Nutrition 2008;99(3):653-9.
112. Lindqvist HM, Rådjursöga M, Malmodin D, Winkvist A, Ellegård L. Serum metabolite profiles of habitual diet: evaluation by 1 H-nuclear magnetic resonance analysis. American Journal of Clinical Nutrition 2019;110(1):53-62. doi: 10.1093/ajcn/nqz032.
113. Yin X, Gibbons H, Rundle M, Frost G, McNulty BA, Nugent AP, Walton J, Flynn A, Gibney MJ, Brennan L. Estimation of chicken intake by adults using metabolomics-derived markers. Journal of Nutrition 2017;147(10):1850-7.
114. Cuparencu C, Rinnan A, Silvestre MP, Poppitt SD, Raben A, Dragsted LO. The anserine to carnosine ratio: an excellent discriminator between white and red meats consumed by free-living overweight participants of the PREVIEW study. European Journal of Nutrition 2020. doi: http://dx.doi.org/10.1007/s00394-020-02230-3.
115. Wedekind R, Kiss A, Keski-Rahkonen P, Viallon V, Rothwell JA, Cross AJ, Rostgaard-Hansen AL, Sandanger TM, Jakszyn P, Schmidt JA, et al. A metabolomic study of red and processed meat intake and acylcarnitine concentrations in human urine and blood. American Journal of Clinical Nutrition 2020;112(2):381-8. doi: 10.1093/ajcn/nqaa140.
116. Wang ZN, Bergeron N, Levison BS, Li XMS, Chiu S, Jia X, Koeth RA, Li L, Wu YP, Tang WHW, et al. Impact of chronic dietary red meat, white meat, or non-meat protein on trimethylamine N-oxide metabolism and renal excretion in healthy men and women. Eur Heart J 2019;40(7):583-+. doi: 10.1093/eurheartj/ehy799.
117. Carrizo D, Chevallier OP, Woodside JV, Brennan SF, Cantwell MM, Cuskelly G, Elliott CT. Untargeted metabolomic analysis of human serum samples associated

with different levels of red meat consumption: A possible indicator of type 2 diabetes? Food Chemistry 2017;221:214-21.
118. O'Sullivan A, Gibney MJ, Brennan L. Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies. Am J Clin Nutr 2011;93(2):314-21. doi: 10.3945/ajcn.110.000950.
119. Pallister I, Morris RM, Lloyd T, Marsden NJ, Wright T, Gilbert M, Phillips J. A novel method to correctly place the fasciotomy incision for decompression of the anterior and peroneal compartments of the leg. Injury 2016;47(4):962-8. doi: 10.1016/j.injury.2016.02.007.
120. Madrid-Gambin F, Brunius C, Garcia-Aloy M, Estruel-Amades S, Landberg R, Andres-Lacueva C. Untargeted 1H NMR-Based Metabolomics Analysis of Urine and Serum Profiles after Consumption of Lentils, Chickpeas, and Beans: An Extended Meal Study To Discover Dietary Biomarkers of Pulses. Journal of agricultural and food chemistry 2018;66(27):6997-7005.
121. Hutchins AM, Martini MC, Olson BA, Thomas W, Slavin JL. Flaxseed influences urinary lignan excretion in a dose-dependent manner in postmenopausal women. Cancer Epidemiology Biomarkers & Prevention 2000;9(10):1113-8.
122. Garcia-Aloy M, Ulaszewska M, Franceschi P, Estruel-Amades S, Weinert CH, Tor-Roca A, Urpi-Sarda M, Mattivi F, Andres-Lacueva C. Discovery of Intake Biomarkers of Lentils, Chickpeas, and White Beans by Untargeted LC-MS Metabolomics in Serum and Urine. Molecular nutrition & food research 2020;64(13):e1901137. doi: http://dx.doi.org/10.1002/mnfr.201901137.
123. Malik VS, Guasch-Ferre M, Hu FB, Townsend MK, Zeleznik OA, Eliassen AH, Tworoger SS, Karlson EW, Costenbader KH, Ascherio A, et al. Identification of Plasma Lipid Metabolites Associated with Nut Consumption in US Men and Women. Journal of Nutrition 2019;16:16.
124. Rabassa M, Zamora-Ros R, Palau-Rodriguez M, Tulipani S, Minarro A, Bandinelli S, Ferrucci L, Cherubini A, Andres-Lacueva C. Habitual Nut Exposure, Assessed by Dietary and Multiple Urinary Metabolomic Markers, and Cognitive Decline in Older Adults: The InCHIANTI Study. Molecular nutrition & food research 2020;64(2):e1900532. doi: http://dx.doi.org/10.1002/mnfr.201900532.
125. Perera T, Young MR, Zhang Z, Murphy G, Colburn NH, Lanza E, Hartman TJ, Cross AJ, Bobe G. Identification and monitoring of metabolite markers of dry bean consumption in parallel human and mouse studies. Molecular Nutrition & Food Research 2015;59(4):795-806.
126. Li KJ, Borresen EC, Jenkins-Puccetti N, Luckasen G, Ryan EP. Navy Bean and Rice Bran Intake Alters the Plasma Metabolome of Children at Risk for Cardiovascular Disease. Frontiers in Nutrition 2018;4. doi: 10.3389/fnut.2017.00071.
127. Mora-Cubillos X, Tulipani S, Garcia-Aloy M, Bullo M, Tinahones FJ, Andres-Lacueva C. Plasma metabolomic biomarkers of mixed nuts exposure inversely correlate with severity of metabolic syndrome. Molecular Nutrition & Food Research 2015;59(12):2480-90. doi: 10.1002/mnfr.201500549.

128. Zheng Y, Yu B, Alexander D, Steffen LM, Nettleton JA, Boerwinkle E. Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study. Am J Clin Nutr 2014;99(6):1470-8.
129. van Roekel EH, Trijsburg L, Assi N, Carayol M, Achaintre D, Murphy N, Rinaldi S, Schmidt JA, Stepien M, Kaaks R, et al. Circulating Metabolites Associated with Alcohol Intake in the European Prospective Investigation into Cancer and Nutrition Cohort. Nutrients 2018;10(5):654. doi: 10.3390/nu10050654.
130. Harada S, Takebayashi T, Kurihara A, Akiyama M, Suzuki A, Hatakeyama Y, Sugiyama D, Kuwabara K, Takeuchi A, Okamura T, et al. Metabolomic profiling reveals novel biomarkers of alcohol intake and alcohol-induced liver injury in community-dwelling men. Environmental Health and Preventive Medicine 2016;21(1):18-26.
131. Dorgan JF, Jung S, Dallal CM, Zhan M, Stennett CA, Zhang Y, Eckert RL, Snetselaar LG, Van Horn L. Alcohol consumption and serum metabolite concentrations in young women. Cancer Causes & Control 2020;31(2):113-26. doi: 10.1007/s10552-019-01256-1.
132. Du D, Bruno R, Blizzard L, Venn A, Dwyer T, Smith KJ, Magnussen CG, Gall S. The metabolomic signatures of alcohol consumption in young adults. European Journal of Preventive Cardiology 2020;27(8):840-9. doi: 10.1177/2047487319834767.
133. Lécuyer L, Dalle C, Micheau P, Pétéra M, Centeno D, Lyan B, Lagree M, Galan P, Hercberg S, Rossary A, et al. Untargeted plasma metabolomic profiles associated with overall diet in women from the SU.VI.MAX cohort. European Journal of Nutrition 2020;59(8):3425-39. doi: 10.1007/s00394-020-02177-5.
134. Zamora-Ros R, Urpi-Sarda M, Lamuela-Raventos RM, Estruch R, Vazquez-Agell M, Serrano-Martinez M, Jaeger W, Andres-Lacueva C. Diagnostic performance of urinary resveratrol metabolites as a biomarker of moderate wine consumption. Clinical Chemistry 2006;52(7):1373-80. doi: 10.1373/clinchem.2005.065870.
135. Regueiro J, Vallverdu-Queralt A, Simal-Gandara J, Estruch R, Lamuela-Raventos RM. Urinary tartaric acid as a potential biomarker for the dietary assessment of moderate wine consumption: a randomised controlled trial. British Journal of Nutrition 2014;111(9):1680-5. doi: 10.1017/s0007114513004108.
136. Noh H, Freisling H, Assi N, Zamora-Ros R, Achaintre D, Affret A, Mancini F, Boutron-Ruault M-C, Flogel A, Boeing H, et al. Identification of Urinary Polyphenol Metabolite Patterns Associated with Polyphenol-Rich Food Intake in Adults from Four European Countries. Nutrients 2017;9(8):796. doi: https://dx.doi.org/10.3390/nu9080796.
137. Vazquez-Fresno R, Llorach R, Marinic J, Tulipani S, Garcia-Aloy M, Espinosa-Martos I, Jimenez E, Rodriguez JM, Andres-Lacueva C. Urinary metabolomic fingerprinting after consumption of a probiotic strain in women with mastitis. Pharmacological Research 2014;87:160-5.
138. Urpi-Sarda M, Boto-Ordonez M, Queipo-Ortuno MI, Tulipani S, Corella D, Estruch R, Tinahones FJ, Andres-Lacueva C. Phenolic and microbial-targeted metabolomics to discovering and evaluating wine intake biomarkers in human

urine and plasma. Electrophoresis 2015;36(18):2259-68. doi: 10.1002/elps.201400506.
139. Donovan JL, Bell JR, Kasim-Karakas S, German JB, Walzem RL, Hansen RJ, Waterhouse AL. Catechin is present as metabolites in human plasma after consumption of red wine. Journal of Nutrition 1999;129(9):1662-8.
140. Tsang C, Higgins S, Duthie GG, Duthie SJ, Howie M, Mullen W, Lean ME, Crozier A. The influence of moderate red wine consumption on antioxidant status and indices of oxidative stress associated with CHD in healthy volunteers. British Journal of Nutrition 2005;93(2):233-40.
141. Boto-Ordonez M, Urpi-Sarda M, Queipo-Ortuno MI, Corella D, Tinahones FJ, Estruch R, Andres-Lacueva C. Microbial Metabolomic Fingerprinting in Urine after Regular Dealcoholized Red Wine Consumption in Humans. Journal of Agricultural and Food Chemistry 2013;61(38):9166-75. doi: 10.1021/jf402394c.
142. Hodgson JM, Morton LW, Puddey IB, Beilin LJ, Croft KD. Gallic acid metabolites are markers of black tea intake in humans. Journal of Agricultural & Food Chemistry 2000;48(6):2276-80.
143. Van Dorsten FA, Daykin CA, Mulder TP, Van Duynhoven JP. Metabonomics approach to determine metabolic differences between green tea and black tea consumption. Journal of Agricultural & Food Chemistry 2006;54(18):6929-38.
144. Mulder TP, Rietveld AG, van Amelsvoort JM. Consumption of both black tea and green tea results in an increase in the excretion of hippuric acid into urine. Am J Clin Nutr 2005;81(1):256S-60S.
145. Hodgson JM, Devine A, Puddey IB, Beilby J, Prince RL. Drinking tea is associated with lower plasma total homocysteine in older women. Asia Pacific Journal of Clinical Nutrition 2006;15(2):253-8.
146. Seow WJ, Yanwen DL, Pan WC, Gunther SH, Sim X, Torta F, Herr DR, Kovalik JP, Jianhong C, Khoo CM, et al. Coffee, Black Tea, and Green Tea Consumption in Relation to Plasma Metabolites in an Asian Population. Molecular nutrition & food research 2020:e2000527. doi: http://dx.doi.org/10.1002/mnfr.202000527.
147. Clarke KA, Dew TP, Watson RE, Farrar MD, Bennett S, Nicolaou A, Rhodes LE, Williamson G. High performance liquid chromatography tandem mass spectrometry dual extraction method for identification of green tea catechin metabolites excreted in human urine. Journal of Chromatography B: Analytical Technologies in the Biomedical & Life Sciences 2014;972:29-37.
148. Vetrani C, Rivellese AA, Annuzzi G, Adiels M, Borén J, Mattila I, Orešič M, Aura A-M. Metabolic transformations of dietary polyphenols: comparison between in vitro colonic and hepatic models and in vivo urinary metabolites. Journal of Nutritional Biochemistry 2016;33:111-8. doi: 10.1016/j.jnutbio.2016.03.007.
149. Del Rio D, Calani L, Cordero C, Salvatore S, Pellegrini N, Brighenti F. Bioavailability and catabolism of green tea flavan-3-ols in humans. Nutrition 2010;26(11-12):1110-6.
150. Li C, Meng X, Winnik B, Lee MJ, Lu H, Sheng S, Buckley B, Yang CS. Analysis of urinary metabolites of tea catechins by liquid chromatography/electrospray

ionization mass spectrometry. Chemical Research in Toxicology 2001;14(6):702-7.
151. Renouf M, Marmet C, Guy PA, Beaumont M, Lepage M, Williamson G, Dionisi F. Dose-response plasma appearance of green tea catechins in adults. Molecular Nutrition & Food Research 2013;57(5):833-9.
152. Hodgson AB, Randell RK, Boon N, Garczarek U, Mela DJ, Jeukendrup AE, Jacobs DM. Metabolic response to green tea extract during rest and moderate-intensity exercise. Journal of Nutritional Biochemistry 2013;24(1):325-34. doi: 10.1016/j.jnutbio.2012.06.017.
153. Ibero-Baraibar I, Romo-Hualde A, Gonzalez-Navarro CJ, Zulet MA, Martinez JA. The urinary metabolomic profile following the intake of meals supplemented with a cocoa extract in middle-aged obese subjects. Food & Function 2016;7(4):1924-31. doi: 10.1039/c5fo01191d.
154. Roura E, Andrés-Lacueva C, Estruch R, Lourdes Mata Bilbao M, Izquierdo-Pulido M, Lamuela-Raventós RM. The effects of milk as a food matrix for polyphenols on the excretion profile of cocoa (-)-epicatechin metabolites in healthy human subjects. British Journal of Nutrition 2008;100(4):846-51.
155. Martinez-Lopez S, Sarria B, Gomez-Juaristi M, Goya L, Mateos R, Bravo-Clemente L. Theobromine, caffeine, and theophylline metabolites in human plasma and urine after consumption of soluble cocoa products with different methylxanthine contents. Food Res Int 2014;63:446-55. doi: 10.1016/j.foodres.2014.03.009.
156. Llorach R, Urpi-Sarda M, Jauregui O, Monagas M, Andres-Lacueva C. An LC-MS-Based Metabolomics Approach for Exploring Urinary Metabolome Modifications after Cocoa Consumption. Journal of Proteome Research 2009;8(11):5060-8. doi: 10.1021/pr900470a.
157. Roura E, Almajano MP, Bilbao MLM, Andres-Lacueva C, Estruch R, Lamuela-Raventos RM. Human urine: Epicatechin metabolites and antioxidant activity after cocoa beverage intake. Free Radical Research 2007;41(8):943-9.
158. Roura E, Andres-Lacueva C, Estruch R, Mata-Bilbao ML, Izquierdo-Pulido M, Waterhouse AL, Lamuela-Raventos RM. Milk does not affect the bioavailability of cocoa powder flavonoid in healthy human. Annals of Nutrition and Metabolism 2007;51(6):493-8. doi: 10.1159/000111473.
159. Garcia-Aloy M, Llorach R, Urpi-Sarda M, Jauregui O, Corella D, Ruiz-Canela M, Salas-Salvado J, Fito M, Ros E, Estruch R, et al. A metabolomics-driven approach to predict cocoa product consumption by designing a multimetabolite biomarker model in free-living subjects from the PREDIMED study. Molecular Nutrition & Food Research 2015;59(2):212-20.
160. Kempf K, Herder C, Erlund I, Kolb H, Martin S, Carstensen M, Koenig W, Sundvall J, Bidel S, Kuha S, et al. Effects of coffee consumption on subclinical inflammation and other risk factors for type 2 diabetes: a clinical trial. Am J Clin Nutr 2010;91(4):950-7. doi: 10.3945/ajcn.2009.28548.

161. Stalmach A, Williamson G, Crozier A. Impact of dose on the bioavailability of coffee chlorogenic acids in humans. Food & Function 2014;5(8):1727-37. doi: 10.1039/c4fo00316k.
162. Cornelis MC, Erlund I, Michelotti GA, Herder C, Westerhuis JA, Tuomilehto J. Metabolomic response to coffee consumption: application to a three-stage clinical trial. Journal of Internal Medicine 2018;283(6):544-57. doi: 10.1111/joim.12737.
163. Madrid-Gambin F, Garcia-Aloy M, Vazquez-Fresno R, Vegas-Lozano E, Jubany M, Misawa K, Hase T, Shimotoyodome A, Andres-Lacueva C. Impact of chlorogenic acids from coffee on urine metabolome in healthy human subjects. Food Res Int 2016;89:1064-70. doi: 10.1016/j.foodres.2016.03.038.
164. Stalmach A, Mullen W, Barron D, Uchida K, Yokota T, Cavin C, Steiling H, Williamson G, Crozier A. Metabolite profiling of hydroxycinnamate derivatives in plasma and urine after the ingestion of coffee by humans: identification of biomarkers of coffee consumption. Drug Metabolism & Disposition 2009;37(8):1749-58.
165. Kremer JI, Gompel K, Bakuradze T, Eisenbrand G, Richling E. Urinary Excretion of Niacin Metabolites in Humans After Coffee Consumption. Molecular Nutrition & Food Research 2018;62(7):e1700735.
166. Mills CE, Flury A, Marmet C, Poquet L, Rimoldi SF, Sartori C, Rexhaj E, Brenner R, Allemann Y, Zimmermann D, et al. Mediation of coffee-induced improvements in human vascular function by chlorogenic acids and its metabolites: Two randomized, controlled, crossover intervention trials. Clinical Nutrition 2017;36(6):1520-9. doi: 10.1016/j.clnu.2016.11.013.
167. Felberg I, Farah A, Monteiro MC, Godoy RLD, Pacheco S, Calado V, Donangelo CM. Effect of simultaneous consumption of soymilk and coffee on the urinary excretion of isoflavones, chlorogenic acids and metabolites in healthy adults. Journal of Functional Foods 2015;19:688-99. doi: 10.1016/j.jff.2015.09.059.
168. Menni C, Zhai GJ, MacGregor A, Prehn C, Romisch-Margl W, Suhre K, Adamski J, Cassidy A, Illig T, Spector TD, et al. Targeted metabolomics profiles are strongly correlated with nutritional patterns in women. Metabolomics 2013;9(2):506-14. doi: 10.1007/s11306-012-0469-6.
169. Papandreou C, Hernández-Alonso P, Bulló M, Ruiz-Canela M, Yu E, Guasch-Ferré M, Toledo E, Dennis C, Deik A, Clish C, et al. Plasma Metabolites Associated with Coffee Consumption: A Metabolomic Approach within the PREDIMED Study. Nutrients 2019;11(5):1032-. doi: 10.3390/nu11051032.
170. Klebanoff MA, Levine RJ, Dersimonian R, Clemens JD, Wilkins DG. Serum caffeine and paraxanthine as markers for reported caffeine intake in pregnancy. Ann Epidemiol 1998;8(2):107-11.
171. Rothwell JA, Fillatre Y, Martin J-F, Lyan B, Pujos-Guillot E, Fezeu L, Hercberg S, Comte B, Galan P, Touvier M, et al. New biomarkers of coffee consumption identified by the non-targeted metabolomic profiling of cohort study subjects. PloS one 2014;9(4):e93474-e. doi: https://dx.doi.org/10.1371/journal.pone.0093474.

172. Rybak ME, Sternberg MR, Pao C-I, Ahluwalia N, Pfeiffer CM. Urine excretion of caffeine and select caffeine metabolites is common in the U.S. population and associated with caffeine intake. Journal of Nutrition 2015;145(4):766-74. doi: 10.3945/jn.114.205476.
173. Rothwell JA, Keski-Rahkonen P, Robinot N, Assi N, Casagrande C, Jenab M, Ferrari P, Boutron-Ruault M-C, Mahamat-Saleh Y, Mancini FR, et al. A Metabolomic Study of Biomarkers of Habitual Coffee Intake in Four European Countries. Molecular nutrition & food research 2019;63(22):e1900659. doi: https://dx.doi.org/10.1002/mnfr.201900659.
174. Rios-Leyvraz M, Bochud M, Tabin R, Genin B, Russo M, Rossier MF, Eap CB, Bovet P, Chiolero A. Monitoring caffeine intake in children with a questionnaire and urine collection: a cross-sectional study in a convenience sample in Switzerland. European Journal of Nutrition 2020;59(8):3537-43. doi: 10.1007/s00394-020-02187-3.
175. Mack CI, Egert B, Liberto E, Weinert CH, Bub A, Hoffmann I, Bicchi C, Kulling SE, Cordero C. Robust Markers of Coffee Consumption Identified Among the Volatile Organic Compounds in Human Urine. Molecular nutrition & food research 2019;63(10):e1801060. doi: https://dx.doi.org/10.1002/mnfr.201801060.
176. Shiokawa Y, Date Y, Kikuchi J. Application of kernel principal component analysis and computational machine learning to exploration of metabolites strongly associated with diet. Scientific Reports 2018;8:3426. doi: 10.1038/s41598-018-20121-w.
177. Zheng H, Lorenzen JK, Astrup A, Larsen LH, Yde CC, Clausen MR, Bertram HC. Metabolic Effects of a 24-Week Energy-Restricted Intervention Combined with Low or High Dairy Intake in Overweight Women: An NMR-Based Metabolomics Investigation. Nutrients 2016;8(3). doi: 10.3390/nu8030108.
178. Fuchsmann P, Stern MT, Munger LH, Pimentel G, Burton KJ, Vionnet N, Vergeres G. Nutrivolatilomics of urinary and plasma samples to identify candidate biomarkers after cheese, milk, and soy-based drink intake in healthy humans. Journal of Proteome Research 2020;19(10):4019-33. doi: http://dx.doi.org/10.1021/acs.jproteome.0c00324.
179. Lau CHE, Siskos AP, Maitre L, Robinson O, Athersuch TJ, Want EJ, Urquiza J, Casas M, Vafeiadi M, Roumeliotaki T, et al. Determinants of the urinary and serum metabolome in children from six European populations. Bmc Medicine 2018;16. doi: 10.1186/s12916-018-1190-8.
180. Zong G, Sun Q, Yu D, Zhu J, Sun L, Ye X, Li H, Jin Q, Zheng H, Hu FB, et al. Dairy consumption, type 2 diabetes, and changes in cardiometabolic traits: a prospective cohort study of middle-aged and older chinese in beijing and shanghai. Diabetes Care 2014;37(1):56-63. doi: 10.2337/dc13-0975.
181. Hruby A, Dennis C, Jacques PF. Dairy Intake in 2 American Adult Cohorts Associates with Novel and Known Targeted and Nontargeted Circulating Metabolites. Journal of Nutrition 2020;150(5):1272-83. doi: 10.1093/jn/nxaa021.
182. Munger LH, Trimigno A, Picone G, Freiburghaus C, Pimentel G, Burton KJ, Pralong FP, Vionnet N, Capozzi F, Badertscher R, et al. Identification of Urinary

Food Intake Biomarkers for Milk, Cheese, and Soy-Based Drink by Untargeted GC-MS and NMR in Healthy Humans. Journal of Proteome Research 2017;16(9):3321-35. doi: 10.1021/acs.jproteome.7b00319.

183. Trimigno A, Munger L, Picone G, Freiburghaus C, Pimentel G, Vionnet N, Pralong F, Capozzi F, Badertscher R, Vergeres G. GC-MS Based Metabolomics and NMR Spectroscopy Investigation of Food Intake Biomarkers for Milk and Cheese in Serum of Healthy Humans. Metabolites 2018;8(2). doi: 10.3390/metabo8020026.
184. Zheng H, Yde CC, Clausen MR, Kristensen M, Lorenzen J, Astrup A, Bertram HC. Metabolomics investigation to shed light on cheese as a possible piece in the French paradox puzzle. Journal of Agricultural & Food Chemistry 2015;63(10):2830-9.
185. Zheng H, Yde CC, Dalsgaard TK, Arnberg K, Molgaard C, Michaelsen KF, Larnkjaer A, Bertram HC. Nuclear magnetic resonance-based metabolomics reveals that dairy protein fractions affect urinary urea excretion differently in overweight adolescents. European Food Research and Technology 2015;240(3):489-97. doi: 10.1007/s00217-014-2347-0.
186. Pallister T, Haller T, Thorand B, Altmaier E, Cassidy A, Martin T, Jennings A, Mohney R, Gieger C, MacGregor A, et al. Metabolites of milk intake: a metabolomic approach in UK twins with findings replicated in two European cohorts. European Journal of Nutrition 2017;56(7):2379-91. doi: 10.1007/s00394-016-1278-x.
187. Ostertag LM, Philo M, Colquhoun IJ, Tapp HS, Saha S, Duthie GG, Kemsley EK, de Roos B, Kroon PA, Le Gall G. Acute Consumption of Flavan-3-ol-Enriched Dark Chocolate Affects Human Endogenous Metabolism. Journal of Proteome Research 2017;16(7):2516-26. doi: 10.1021/acs.jproteome.7b00089.
188. Martin FP, Rezzi S, Pere-Trepat E, Kamlage B, Collino S, Leibold E, Kastler J, Rein D, Fay LB, Kochhar S. Metabolic effects of dark chocolate consumption on energy, gut microbiota, and stress-related metabolism in free-living subjects. Journal of Proteome Research 2009;8(12):5568-79.
189. Gibbons H, McNulty BA, Nugent AP, Walton J, Flynn A, Gibney MJ, Brennan L. A metabolomics approach to the identification of biomarkers of sugar-sweetened beverage intake. Am J Clin Nutr 2015;101(3):471-7.
190. Yeung EH, Saudek CD, Jahren AH, Kao WHL, Islas M, Kraft R, Coresh J, Anderson CAM. Evaluation of a Novel Isotope Biomarker for Dietary Consumption of Sweets. Am J Epidemiol 2010;172(9):1045-52. doi: aje/kwq247.
191. MacDougall CR, Hill CE, Jahren AH, Savla J, Riebl SK, Hedrick VE, Raynor HA, Dunsmore JC, Frisard MI, Davy BM. The δ13C Value of Fingerstick Blood Is a Valid, Reliable, and Sensitive Biomarker of Sugar-Sweetened Beverage Intake in Children and Adolescents. Journal of Nutrition 2018;148(1):147-52. doi: 10.1093/jn/nxx017.
192. Perng W, Tang L, Song PXK, Goran M, Rojo MMT, Cantoral A, Peterson KE. Urate and Nonanoate Mark the Relationship between Sugar-Sweetened Beverage

Intake and Blood Pressure in Adolescent Girls: A Metabolomics Analysis in the ELEMENT Cohort. Metabolites 2019;9(5). doi: 10.3390/metabo9050100.
193. Magnuson BA, Carakostas MC, Moore NH, Poulos SP, Renwick AG. Biological fate of low-calorie sweeteners. Nutrition Reviews 2016;74(11):670-89. doi: 10.1093/nutrit/nuw032.
194. Acar E, Gurdeniz G, Khakimov B, Savorani F, Korndal SK, Larsen TM, Engelsen SB, Astrup A, Dragsted LO. Biomarkers of Individual Foods, and Separation of Diets Using Untargeted LC-MS-based Plasma Metabolomics in a Randomized Controlled Trial. Molecular Nutrition & Food Research 2019;63(1). doi: 10.1002/mnfr.201800215.
195. Khakimov B, Poulsen SK, Savorani F, Acar E, Gurdeniz G, Larsen TM, Astrup A, Dragsted LO, Engelsen SB. New Nordic Diet versus Average Danish Diet: A Randomized Controlled Trial Revealed Healthy Long-Term Effects of the New Nordic Diet by GC-MS Blood Plasma Metabolomics. Journal of Proteome Research 2016;15(6):1939-54. doi: 10.1021/acs.jproteome.6b00109.
196. Trimigno A, Khakimov B, Savorani F, Poulsen SK, Astrup A, Dragsted LO, Engelsen SB. Human urine 1H NMR metabolomics reveals alterations of the protein and carbohydrate metabolism when comparing habitual Average Danish diet vs. healthy New Nordic diet. Nutrition 2020;79-80:110867. doi: http://dx.doi.org/10.1016/j.nut.2020.110867.
197. Rebholz CM, Lichtenstein AH, Zheng Z, Appel LJ, Coresh J. Serum untargeted metabolomic profile of the Dietary Approaches to Stop Hypertension (DASH) dietary pattern. Am J Clin Nutr 2018;108(2):243-55. doi: 10.1093/ajcn/nqy099.
198. Walker ME, Song RJ, Xu X, Gerszten RE, Ngo D, Clish CB, Corlin L, Ma J, Xanthakis V, Jacques PF, et al. Proteomic and metabolomic correlates of healthy dietary patterns: The framingham heart study. Nutrients 2020;12(5):1476. doi: http://dx.doi.org/10.3390/nu12051476.
199. Michielsen C, Hangelbroek RWJ, Feskens EJM, Afman LA. Disentangling the Effects of Monounsaturated Fatty Acids from Other Components of a Mediterranean Diet on Serum Metabolite Profiles: A Randomized Fully Controlled Dietary Intervention in Healthy Subjects at Risk of the Metabolic Syndrome. Molecular Nutrition & Food Research 2019;63(9):e1801095.
200. Davis C, Hodgson J, Bryan J, Garg M, Woodman R, Murphy K. Older Australians Can Achieve High Adherence to the Mediterranean Diet during a 6 Month Randomised Intervention; Results from theMedley Study. Nutrients 2017;9(6). doi: 10.3390/nu9060534.
201. Vazquez-Fresno R, Llorach R, Urpi-Sarda M, Lupianez-Barbero A, Estruch R, Corella D, Fito M, Aros F, Ruiz-Canela M, Salas-Salvado J, et al. Metabolomic Pattern Analysis after Mediterranean Diet Intervention in a Nondiabetic Population: A 1- and 3-Year Follow-up in the PREDIMED Study. Journal of Proteome Research 2015;14(1):531-40. doi: 10.1021/pr5007894.
202. De Filippis F, Pellegrini N, Vannini L, Jeffery IB, La Storia A, Laghi L, Serrazanetti DI, Di Cagno R, Ferrocino I, Lazzi C, et al. High-level adherence to a

Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut 2016;65(11):1812-21. doi: 10.1136/gutjnl-2015-309957.
203. Tong TYN, Koulman A, Griffin JL, Wareham NJ, Forouhi NG, Imamura F. A Combination of Metabolites Predicts Adherence to the Mediterranean Diet Pattern and Its Associations with Insulin Sensitivity and Lipid Homeostasis in the General Population: The Fenland Study, United Kingdom. Journal of Nutrition 2020;150(3):568-78. doi: 10.1093/jn/nxz263.
204. Szeto YT, Kwok TCY, Benzie IFF. Effects of a long-term vegetarian diet on biomarkers of antioxidant status and cardiovascular disease risk. Nutrition 2004;20(10):863-6.
205. Radjursoga M, Lindqvist HM, Pedersen A, Karlsson BG, Malmodin D, Ellegard L, Winkvist A. Nutritional Metabolomics: Postprandial Response of Meals Relating to Vegan, Lacto-Ovo Vegetarian, and Omnivore Diets. Nutrients 2018;10(8). doi: 10.3390/nu10081063.
206. Pezdirc K, Hutchesson MJ, Williams RL, Rollo ME, Burrows TL, Wood LG, Oldmeadow C, Collins CE. Consuming High-Carotenoid Fruit and Vegetables Influences Skin Yellowness and Plasma Carotenoids in Young Women: A Single-Blind Randomized Crossover Trial. Journal of the Academy of Nutrition & Dietetics 2016;116(8):1257-65. doi: 10.1016/j.jand.2016.03.012.
207. Chung HY, Ferreira AL, Epstein S, Paiva SA, Castaneda-Sceppa C, Johnson EJ. Site-specific concentrations of carotenoids in adipose tissue: relations with dietary and serum carotenoid concentrations in healthy adults. Am J Clin Nutr 2009;90(3):533-9. doi: 10.3945/ajcn.2009.27712.
208. May DH, Navarro SL, Ruczinski I, Hogan J, Ogata Y, Schwarz Y, Levy L, Holzman T, McIntosh MW, Lampe JW. Metabolomic profiling of urine: response to a randomised, controlled feeding study of select fruits and vegetables, and application to an observational study. British Journal of Nutrition 2013;110(10):1760-70. doi: 10.1017/s000711451300127x.
209. Macdonald HA, Hardcastle AC, Duthie GG, Duthie SJ, Aucott L, Sandison R, Shearer MJ, Reid DM. Changes in vitamin biomarkers during a 2-year intervention trial involving increased fruit and vegetable consumption by free-living volunteers. British Journal of Nutrition 2009;102(10):1477-86. doi: 10.1017/s0007114509990377.
210. Jahns L, Johnson LK, Mayne ST, Cartmel B, Picklo Sr MJ, Ermakov IV, Gellermann W, Whigham LD. Skin and plasma carotenoid response to a provided intervention diet high in vegetables and fruit: uptake and depletion kinetics. Am J Clin Nutr 2014;100(3):930-7. doi: 10.3945/ajcn.114.086900.
211. Aguilar SS, Wengreen HJ, Lefevre M, Madden GJ, Gast J. Skin Carotenoids: A Biomarker of Fruit and Vegetable Intake in Children. Journal of the Academy of Nutrition & Dietetics 2014;114(8):1174-80. doi: 10.1016/j.jand.2014.04.026.
212. Amer B, Clausen MR, Bertram HC, Bohl M, Nebel C, Zheng H, Skov T, Larsen MK, Gregersen S, Hermansen K, et al. Consumption of Whey in Combination with Dairy Medium-Chain Fatty Acids (MCFAs) may Reduce Lipid Storage due to Urinary Loss of Tricarboxylic Acid Cycle Intermediates and Increased Rates of

MCFAs Oxidation. Molecular Nutrition & Food Research 2017;61(12). doi: 10.1002/mnfr.201601048.
213. Wiseman H, Casey K, Bowey EA, Duffy R, Davies M, Rowland IR, Lloyd AS, Murray A, Thompson R, Clarke DB. Influence of 10 wk of soy consumption on plasma concentrations and excretion of isoflavonoids and on gut microflora metabolism in healthy adults. Am J Clin Nutr 2004;80(3):692-9.
214. Silva FD, Lemos TCC, Sand?ra D, Monteiro M, Perrone D. Fermentation of soybean meal improves isoflavone metabolism after soy biscuit consumption by adults. Journal of the Science of Food and Agriculture 2020;100(7):2991-8. doi: 10.1002/jsfa.10328.
215. Maskarinec G, Singh S, Meng LX, Franke AA. Dietary soy intake and urinary isoflavone excretion among women from a multiethnic population. Cancer Epidemiology Biomarkers & Prevention 1998;7(7):613-9.
216. Seow A, Shi CY, Franke AA, Hankin JH, Lee HP, Yu MC. Isoflavonoid levels in spot urine are associated with frequency of dietary soy intake in a population-based sample of middle-aged and older Chinese in Singapore. Cancer Epidemiology Biomarkers & Prevention 1998;7(2):135-40.
217. Jaceldo-Siegl K, Fraser GE, Chan J, Franke A, Sabaté J. Validation of soy protein estimates from a food-frequency questionnaire with repeated 24-h recalls and isoflavonoid excretion in overnight urine in a Western population with a wide range of soy intakes. Am J Clin Nutr 2008;87(5):1422-7.
218. Vergne S, Bennetau-Pelissero C, Lamothe V, Chantre P, Potier M, Asselineau J, Perez P, Durand M, Moore N, Sauvant P. Higher bioavailability of isoflavones after a single ingestion of a soya-based supplement than a soya-based food in young healthy males. British Journal of Nutrition 2008;99(2):333-44.
219. Cuparencu C, Praticó G, Hemeryck LY, Sri Harsha PSC, Noerman S, Rombouts C, Xi M, Vanhaecke L, Hanhineva K, Brennan L, et al. Biomarkers of meat and seafood intake: an extensive literature review. Genes & Nutrition 2019;14(1):35. doi: 10.1186/s12263-019-0656-4.
220. Brouwer-Brolsma EM, Brandl B, Buso MEC, Skurk T, Manach C. Food intake biomarkers for green leafy vegetables, bulb vegetables, and stem vegetables: a review. Genes and Nutrition 2020;15(1). doi: 10.1186/s12263-020-00667-z.
221. Garcia-Aloy M, Hulshof PJM, Estruel-Amades S, Osté MCJ, Lankinen M, Geleijnse JM, de Goede J, Ulaszewska M, Mattivi F, Bakker SJL, et al. Biomarkers of food intake for nuts and vegetable oils: an extensive literature search. Genes & Nutrition 2019;14(1):7. doi: 10.1186/s12263-019-0628-8.
222. Ulaszewska MM, Weinert CH, Trimigno A, Portmann R, Andres Lacueva C, Badertscher R, Brennan L, Brunius C, Bub A, Capozzi F, et al. Nutrimetabolomics: An Integrative Action for Metabolomic Analyses in Human Nutritional Studies. Mol Nutr Food Res 2019;63(1):e1800384. doi: 10.1002/mnfr.201800384.
223. Dragsted LO, Gao Q, Scalbert A, Vergères G, Kolehmainen M, Manach C, Brennan L, Afman LA, Wishart DS, Andres Lacueva C, et al. Validation of

biomarkers of food intake—critical assessment of candidate biomarkers. Genes & Nutrition 2018;13(1):14. doi: 10.1186/s12263-018-0603-9.
224. Brennan L. Moving toward Objective Biomarkers of Dietary Intake. Journal of Nutrition 2018;148(6):821-2. doi: 10.1093/jn/nxy067.
225. Kim SJ, Kim SH, Kim JH, Hwang S, Yoo HJ. Understanding Metabolomics in Biomedical Research. Endocrinol Metab (Seoul) 2016;31(1):7-16. doi: 10.3803/EnM.2016.31.1.7.
226. Kirkwood JS, Maier C, Stevens JF. Simultaneous, untargeted metabolic profiling of polar and nonpolar metabolites by LC-Q-TOF mass spectrometry. Curr Protoc Toxicol 2013;Chapter 4:Unit4.39. doi: 10.1002/0471140856.tx0439s56.
227. de Macedo AN, Macri J, Hudecki PL, Saoi M, McQueen MJ, Britz-McKibbin P. Validation of a Capillary Electrophoresis Assay for Monitoring Iodine Nutrition in Populations for Prevention of Iodine Deficiency: An Interlaboratory Method Comparison. The Journal of Applied Laboratory Medicine 2019;1(6):649-60. doi: 10.1373/jalm.2016.022848.
228. Gibney MJ, Walsh M, Brennan L, Roche HM, German B, van Ommen B. Metabolomics in human nutrition: opportunities and challenges. The American Journal of Clinical Nutrition 2005;82(3):497-503. doi: 10.1093/ajcn/82.3.497.
229. Manach C, Scalbert A, Morand C, Rémésy C, Jiménez L. Polyphenols: food sources and bioavailability. The American Journal of Clinical Nutrition 2004;79(5):727-47. doi: 10.1093/ajcn/79.5.727.
230. Seeram NP, Henning SM, Zhang Y, Suchard M, Li Z, Heber D. Pomegranate juice ellagitannin metabolites are present in human plasma and some persist in urine for up to 48 hours. J Nutr 2006;136(10):2481-5. doi: 10.1093/jn/136.10.2481.
231. Wang X, Gu H, Palma-Duran SA, Fierro A, Jasbi P, Shi X, Bresette W, Tasevska N. Influence of Storage Conditions and Preservatives on Metabolite Fingerprints in Urine. Metabolites 2019;9(10). doi: 10.3390/metabo9100203.
232. Wishart DS. Metabolomics for Investigating Physiological and Pathophysiological Processes. Physiol Rev 2019;99(4):1819-75. doi: 10.1152/physrev.00035.2018.

Supplemental Methods

We used the following search terms, with Medical Subject Headings (MeSH) words where available: ("Nutrition Therapy" OR "diet" OR "vegetarian*" OR "Fruit" OR "diet, food, and nutrition" OR "nutrition*" OR "Dietary Carbohydrates" OR "VITAMINS" OR "whole grain*" OR "fruit*" OR "vitamin*" OR "vegetable*" OR "nut*" OR "legume*" OR "bean*" OR "egg*" OR "dairy*" OR "dairies*" OR "milk*" OR "yogurt*" OR "cheese*" " fish*" OR "seafood*" OR "meat*" OR "processed meat*" OR "citrus fruit*" OR "citrus*" OR "grain*" OR "refined grain*" OR "cereal*" OR "rice*" OR "potato*" OR "oil*" OR "spice*" OR "Sodium, Dietary" OR "vegan*" OR "Diet, Vegan" OR "Dietary Sugars" OR "beverage*" OR "caffeine*" OR "starch*" OR "Fats" OR "Cholesterol, Dietary" OR "red meat*" OR "wholegrain*" OR "wholewheat*" "lentil*" OR "soy*" OR "coffee*" OR "Dietary Proteins" OR "Calcium, Dietary" OR "Potassium, Dietary" OR "Folic Acid" OR "Dietary Fiber") AND ("Metabolomics" OR "Metabolome" OR "metabolom*" OR "metabonom" OR "metabolite*" OR "Magnetic Resonance Spectroscopy" OR "nmr*" OR "spectrometry*" OR "Chromatography, High Pressure Liquid" OR "HPLC*" OR "Gas chromatography" "Chromatography, Gas") AND ("Cross-Sectional Studies" OR "Cohort Studies" OR "Case-Control Studies" OR "Clinical Trial" OR "cohort study" OR "historical cohort" OR "Retrospective Studies" OR "retrospective stud*" OR "retrospective cohort" OR "cohort analysis" OR "case-control" OR "cross-sectional" OR "Randomized Controlled Trial" OR "Randomized Controlled Trial*" OR "clinical trial*" OR "case-cohort*" OR "nested case-control").

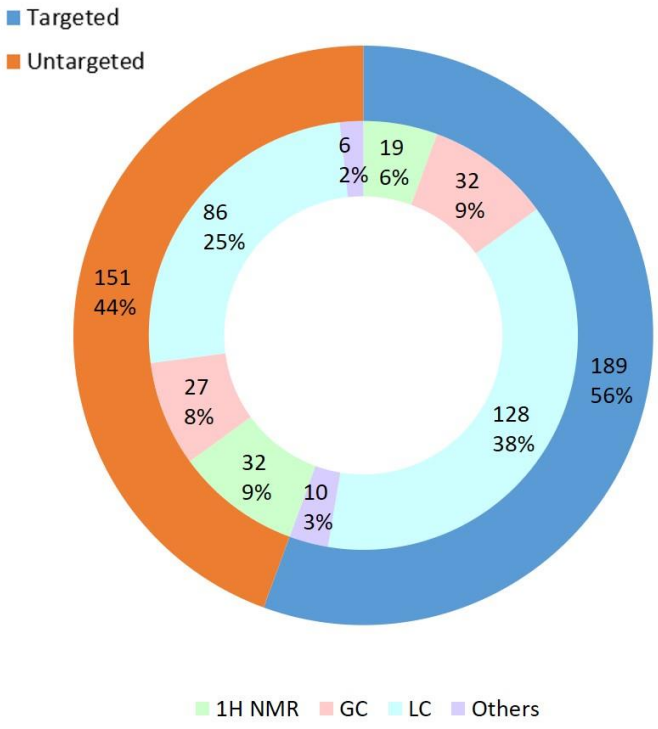


Figure S2.1: Analytical techniques by metabolomic approach (targeted versus untargeted)

CHAPTER 3 – Sources of variation in food-related metabolites during pregnancy

Thesis chapter is derived from a published peer-reviewed article:

Rafiq, T., Azab, S. M., Anand, S. S., Thabane, L., Shanmuganathan, M., Morrison, K. M., Atkinson, S. A., Stearns, J. C., Teo, K. K., Britz-McKibbin, P., & de Souza, R. J. (2022). Sources of Variation in Food-Related Metabolites during Pregnancy. *Nutrients*, *14*(12), 2503. https://doi.org/10.3390/nu14122503

The authors' responsibilities were as follows— TR led the drafting of the manuscript; RJdS and TR: designed the research; TR, RJdS and LT led the statistical analysis; PB-M, SMA and MS led the analytical chemistry interpretations; TR analyzed the data and performed statistical analyses; PB-M, SMA, MS and study participants provided essential reagents or provided essential materials; RJdS, SSA, KKT, JCS, SAA, KMM and PB-M were principal investigators responsible for the oversight of each cohort's data; RJdS had primary responsibility for final content; and all authors read, edited, and approved the final manuscript.

3.1 Abstract: The extent to which variation in food-related metabolites are attributable to non-dietary factors remains unclear, which may explain inconsistent food-metabolite associations observed in population studies. This study examined the association between non-dietary factors and serum concentrations of food-related biomarkers and quantified the amount of variability in metabolite concentrations explained by non-dietary factors. Pregnant women (n=600) from two Canadian birth cohorts completed a validated semi-quantitative food frequency questionnaire, and serum metabolites were measured by multisegment injection-capillary electrophoresis-mass spectrometry. Hierarchical linear modelling and principal component partial R-square (PC-PR2) were used for data analysis. For proline betaine and DHA (mainly exogenous), citrus foods and fish/fish oil intake, respectively, explained the highest proportion of variability relative to non-dietary factors. The unique contribution of dietary factors was similar (15:0, 17:0, hippuric acid, TMAO) or lower (14:0, tryptophan betaine, 3-methylhistidine, carnitine) compared to non-dietary factors (i.e., ethnicity, maternal age, gestational age, prepregnancy BMI, physical activity, and smoking) for metabolites that can either be produced endogenously, biotransformed by gut microbiota, and/or derived from multiple food sources. The results emphasize the importance of adjusting for non-dietary factors in future analyses to improve the accuracy and precision of measures of food intake, and their associations with health and disease.

3.2 Introduction

Accurate assessment of dietary intake remains a major challenge in human nutrition research due to the complex nature of food exposure, and reliance on self-reporting, which often leads to biased or unreliable measures of food intake. While most studies use self-reported dietary intake methods such as food frequency questionnaires (FFQ), 24-h dietary recalls, and food records, they may be subject to recall, misclassification, and measurement biases (1). To circumvent this problem, metabolomics—the global analysis of low molecular weight metabolites in biological samples—have been increasingly applied in large-scale epidemiological studies for the discovery and validation of food intake biomarkers (2).

Biomarkers can provide a more objective assessment of food exposures than self-reported dietary intake because they account for nutrient bioavailability and metabolism. An ideal biomarker of food intake is one that can be readily measured in human biofluid (blood or urine) at the population level, highly specific for one food item or food group, shows a dose- and time-dependent response, and is not extensively transformed by the microbiota and host tissue upon consumption. However, complex interpretative challenges exist since nutrients are derived from various food sources and can display intercorrelation between other metabolic processes (3). Furthermore, the human metabolome exhibits variability due to intrinsic physiologic characteristics such as age, sex, hormonal levels, and the gut microbiome, as well as due to extrinsic factors such as habitual diet and lifestyle. Further, many putative biomarkers of food intake do not exclusively originate from a single food or nutrient. For example, trimethylamine *N*-

oxide (TMAO) is formed from TMA-containing nutrient such as choline, which is abundant in fish, beef and eggs, but can also be produced from carnitine in red meat (2, 4). Moreover, many of the gut-microbiome dependent metabolites and other food-specific metabolites are metabolized in the liver at different rates depending on hepatic enzyme activity (5), which may contribute to the greater variability observed in the range of metabolite measured in the biological samples (6). Consequently, it is important to identify potential non-dietary sources of food-related biomarkers and examine the extent to which these factors explain differences in metabolite concentration.

In most cases, food intake explains a relatively small proportion (R^2 <10%) of the total variation in a given metabolite concentration, and other determinants are typically unknown, unmeasured, or if measured, the extent of measurement error is not clear (7). Biomarkers derived from food intake and gut microbiota are influenced by non-dietary factors (8, 9), however the extent to which these factors compromise the validity of the metabolite as a food intake biomarker may depend on the specificity of the biomarker (well-established, uncertain, or weak biomarker of the particular food), whether the biomarker is endogenously produced, biotransformed by gut microbiota, and/or derived from more than one food source. Understanding the sources of variation in biomarkers of food intake that are not attributed to changes in food intake are critical to advance the application/field of food intake biomarkers. If sources of variation are not clearly understood, then using these biomarkers as markers of food/nutrient intake may simply exchange one source of measurement error (self-misreport) for others (changes in the biomarker intake unrelated to changes in food intake).

Carefully designed studies examining the association between non-dietary factors and biomarker concentrations are sparse and especially lacking in women during pregnancy. Observational studies, specifically birth cohort studies, are useful designs to learn about pregnancy exposures and birth outcomes (10). Women experience a series of metabolic modifications during pregnancy, likely affected by pre-pregnancy and intrapartum factors, which in turn may affect maternal health and disease at critical stages of fetal development (11, 12). Moreover, metabolite concentrations during gestation and pre-pregnancy, and pregnancy related factors such as GDM also differ between ethnic groups (e.g., White Europeans and South Asians) (9). The purpose of this study was to examine the associations of non-dietary factors including demographics, lifestyle, and pregnancy-related factors with serum metabolite concentrations using a panel of commonly identified biomarkers derived from food intake and/or gut microbiota including proline betaine, five fatty acids (even-chain saturated fatty acids (SFA) myristic acid (14:0), odd-chain SFA pentadecanoic acid (15:0) and heptadecanoic acid (17:0), and omega-3 polyunsaturated fatty acids (ω-3 PUFA) docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA)), hippuric acid, TMAO, 3-methylhistidine, carnitine, and tryptophan betaine, in pregnant women of two ethnically diverse groups; and to determine the extent to which non-dietary factors explain the variability in concentrations of putative biomarkers of food intake.

3.3 Materials and Methods

3.3.1 Data Source and Participants

This study used data from two longitudinal Canadian birth cohorts of pregnant women: Family Atherosclerosis Monitoring In earLy life (FAMILY) study and SouTh Asian biRth cohorT (START). The FAMILY study included White European women and the START cohort included women of South Asian ethnic background. Design and methodology of these two studies have been described in detail elsewhere (13, 14). Briefly, the FAMILY study was designed to understand the environmental, genetic, and biochemical factors important in the development of obesity and cardiovascular disease risk factors in childhood. A total of 857 families (901 newborns) were recruited between 2002 and 2009 in the Hamilton area, Ontario, Canada. Women were recruited between 24 and 36 weeks of gestation. The START study enrolled 1,012 South Asian (people who originate from the Indian sub-continent: India, Pakistan, Sri Lanka, or Bangladesh) mother–child pairs between 2011 and 2015 from the Peel Region of Ontario to investigate the influence of diverse environmental exposures and genetics on early life adiposity, growth trajectory, and cardiometabolic risk. Ancestral origin of both the woman, her partner, and both offspring's grandparents were required to be classified as South Asian.

All enrolled participants provided full informed consent, and both studies obtained ethics approval from the McMaster Hamilton Integrated Research Ethics Board [START (HiREB #10–640) and FAMILY (HiREB #02–060)].

Clinical and demographic data was harmonized across the two cohorts. When questions were not identical between studies (e.g., physical activity level during pregnancy), comparable categories were constructed with the available data to satisfy the same definition. Within each cohort, 300 pregnant women were randomly selected for serum metabolomics analysis as previously described (15). This selection was based on the contrasting diet quality score (DQS), where 100 mothers were randomly selected from the 3 DQS groups (>90th percentile ["high" diet quality], <10th percentile ["low" diet quality], and between 10th and 90th percentile ["intermediate" diet quality]). A total of 600 pregnant mothers were included in the current analysis (Supplementary Figure S1).

3.3.2 Maternal Serum Metabolome Analyses

A validated multiplexed separation platform based on multisegment injection-capillary-electrophoresis-mass spectrometry (MSI-CE-MS) was used for targeted and nontargeted profiling of polar/ionic metabolites measured consistently in serum filtrate samples with stringent quality control (QC). A standardized method protocol was used for identification and quantification of the maternal serum metabolome as described in more detail elsewhere (15). Briefly, a total of 66 and 67 polar ionic metabolites from serum filtrate samples satisfied selection criteria for their analysis in the FAMILY and START cohorts, respectively, and 53 of these were measured consistently across both cohorts. Serum metabolites were reported only if they satisfied two additional criteria: 1) metabolites that were detected in majority of the individual samples (≥75%) in a cohort (i.e., frequency filter), and 2) with acceptable technical precision based on repeated

analysis of QC samples (i.e., QC filter), to reduce false discoveries and data overfitting. Metabolites with nondetectable or missing values were replaced with half of the lowest detected value for the compound in each cohort. Moreover, a QC-based batch correction algorithm was applied for robust correction of long-term monitoring of signal drift in MSI-CE-MS (16). Among metabolites measured consistently in the two cohorts, six metabolites including proline betaine, 3-methylhistidine, hippuric acid, TMAO, carnitine, and tryptophan betaine were selected for our current analysis as they were previously determined to be associated with self-report of dietary intake (2, 17). Further, they offer a combination of evidence (good, fair, or poor) for candidate biomarkers of food intake that are produced exogenously, endogenously, biotransformed by gut microbiota, and/or derived from more than one food source (2, 17, 18). The reference interval for these serum metabolites in different birth cohorts from across Canada, and their technical/biological variance, and interclass correlation coefficients have been reported previously (15).

3.3.3 Assessment of Dietary Intake

Maternal dietary intake during pregnancy was collected at 24-28 weeks gestation. Semi-quantitative validated food-frequency questionnaires (157 items in the FAMILY and 163-items in the START) developed and validated as part of the Study of Health and Risk in Ethnic Groups (SHARE) Study were used (19-21). Participants were asked to report on the frequency (daily, weekly, monthly, yearly, or never) and amount in serving size of each food or food group on average in the past 12 months. For our study, food items were either used as separate items (chicken, canned fish, fried fish) or classified

into main food groups: citrus food (citrus fruit and citrus juice), red meat, eggs (boiled and fried eggs), seafood, nuts and legumes, and fruits and vegetables. Nutrient intakes were calculated using the ESHA Food Processor Nutrient Analysis Software (ESHA Research, version 6.11, 1996, Salem, OR), derived from the 1991 Canadian Nutrient File and the US Department of Agriculture nutrient food composition databases. Fiber intake and total energy intake were also estimated from the FFQ (17, 19). Data were logarithm-transformed to correct for skewness prior to including them in the regression analysis, and nutrient intakes were adjusted for energy intake using the residual approach (22).

3.3.4 Non-Dietary Factors

Non-dietary factors included ethnicity (White European or South Asian), maternal age (years), gestational age (i.e., weeks of pregnancy), parity, pre-pregnancy body mass index (BMI, kg/m^2), smoking history (current or former smoker and never smoker), physical activity (mainly sedentary, mild activity, moderate activity, and strenuous activity), social disadvantage index (SDI), and gestational diabetes (GDM). For SDI, derived using a previously validated index based on employment status, income, and marital status, higher values indicate greater socioeconomic disadvantage (23). A case of GDM was defined based on the Born in Bradford (BiB) oral glucose tolerance test criteria, self-reported GDM, and insulin use in pregnancy in START cohort; whereas the International Association of the Diabetes and Pregnancy Study Groups [IADPSG] criteria [75-g OGTT with fasting glucose ≥5.1 mmol/L, 1 hour ≥10.0 mmol/L, 2 hours ≥8.5 mmol/L] was used in FAMILY cohort. We selected these factors based on the

known and plausible associations with the selected metabolites and/or they are commonly adjusted in population-based nutritional metabolomics studies.

3.3.5 Statistical Analysis

Descriptive statistics for categorical variables were summarized using frequency and percentages, and continuous data were summarized using mean and standard deviation (SD) or median and interquartile range (IQR). Random-effects hierarchical linear models (HLM) were fit whereby each of the natural logarithm-transformed food-metabolite concentration was regressed on dietary and non-dietary factors after adjusting for other covariates including total energy intake (kcal), total fiber intake (g/day), and period of time between the day FFQ information was collected and blood was drawn (FFQ before blood, FFQ after blood, and both taken on the same day) (24).

The data had a nested (clustered) structure where individuals within the same cohort represented a cluster because they were more similar to one another with regards to dietary and non-dietary factors. Therefore, we used hierarchical linear modeling (HLM) to accommodate the dependent nature of observations in clustered data. HLM allows nesting effects to be incorporated into the model, producing more accurate estimates, and corrects for the error structure violations (non-independent errors) to provide robust conclusions (25, 26). First unconditional (intercept-only) HLM models were tested to determine whether serum metabolite concentrations were nested within cohort using an intraclass correlation (ICC) calculated based on the covariance parameter estimates. An ICC refers to amount of variation attributed to level-two (study-level) factor. An ICC can be determined from an intercept-only model and any relationship with

an ICC of 2% or greater suggests the presence of level-two effects (24). The results showed an ICC of 3.9% for proline betaine, 25.6% for 3-methylhistidine, 1.5% for carnitine, 0% for hippuric acid, 46.0% for tryptophan betaine, and 7.0% for TMAO. A sensitivity analysis using an Ordinary Least Squares (OLS) multivariable linear regression was conducted for carnitine and hippuric acid (Supplementary Table S1). Next, the association between dietary factors as level 1 predictors (fixed) previously shown to be associated with a specific metabolite (e.g., citrus fruit and proline betaine)

Intercept-only Model (Unconditional Model)

$$\text{Metabolite}_{ij} = \beta 0_j + e_{ij}$$

$$\beta 0_j = \gamma 00 + u 0_j$$

Random Intercept ($u0_j$) with Fixed Level-1 Factors (Dietary factors, $\gamma 10$)

$$\text{Metabolite}_{ij} = \beta 0_j + \beta 1_j \textbf{Dietaryfactor} + e_{ij}$$

$$\beta 0_j = \gamma 00 + u 0_j$$

$$\beta 1_j = \gamma 10$$

Random Intercept ($u0_j$) with Fixed Level-1 Factors (Dietary ($\gamma 10$) and Non-dietary factors ($\gamma 20$…))

$$\text{Metabolite}_{ij} = \beta 0_j + \beta 1_j \textbf{Dietaryfactor} + \beta 2_j \textbf{Age …} + e_{ij}$$

$$\beta 0_j = \gamma 00 + u 0_j$$

$$\beta 1_j = \gamma 10$$

$$\beta 2_j = \gamma 20$$

was examined. Finally, in addition to the dietary factors, all non-dietary factors were also added as level 1 predictors. These HLM procedures produced the following three models: The goodness-of-fit statistics including the Akaike Information Criterion [AIC], Bayesian Information Criterion [BIC], and the change in deviance statistic were used to evaluate model fit in terms of the clustering variable. Smaller values of these statistics indicate a

better model fit (27). The AIC and BIC consider error and model parsimony simultaneously. An OLS multivariable linear regression was conducted for NEFAs as these data were only available in FAMILY cohort. Regression estimates (b), 95% confidence intervals (95% CI), and p-values were reported, and statistical analysis was conducted using SAS software version 9.4.

Finally, principal component partial R-square (PC-PR2) analysis was used to quantify the sources of systematic variability in serum metabolite concentrations (28). The PC-PR2 method combines features of principal component analysis (PCA) and the partial R-square statistic in multivariable linear regression and allows for some degree of inter-correlation between explanatory variables. The mathematical details of the PC-PR2 method are described elsewhere (28). Data reduction component was not necessary because the analytic strategy was applied to a single metabolite. The partial R^2 statistic was calculated for each explanatory variable, which quantifies the amount of variability in metabolite explained by that variable, conditional on all other covariates included in the model. The PC-PR2 method was conducted using the R software, version 1.2.5.

3.4 Results

3.4.1 Association of dietary and non-dietary factors with food-related metabolites

Descriptive characteristics of the participants overall and by ethnicity are shown in Table 1. Model fit statistics from the HLM examining the dietary and non-dietary factors associated with food-intake biomarkers are presented in Supplementary Table S3.2 and the regression estimates and 95% CI are presented in Table 3.2. Three regression models including an unconditional model (Model 1), the random intercept

model with level-one dietary factors (Model 2), and random intercept model with level-one dietary and non-dietary factors (Model 3) were examined (Supplementary Table S3.2). For each metabolite outcome, the log likelihood, AIC, and BIC statistics decreased considerably after adding the non-dietary covariates, indicating better model fit. Thus, the regression estimates presented in Table 3.2 are based on Model 3. As expected, most of the dietary food sources were significantly associated with their respective metabolite concentrations, except for carnitine ($p>0.05$) (Table 3.2). For exogenous metabolites specific to a single food source, higher citrus food intake was positively associated with proline betaine concentration (b: 0.27; 95% CI: 0.20, 0.34), and higher intake of nuts and legumes was positively associated with tryptophan betaine concentration (b: 0.02; 95% CI: 0.00, 0.03). For metabolites with both endogenous metabolic and exogenous sources and obtained from multiple food sources, such as hippuric acid, higher intake of fruits and vegetables were associated with higher hippuric acid concentration (b: 0.22; 95% CI: 0.08, 0.36), but no such association was found with tea and coffee intake. Higher intake of chicken (b: 0.02; 95% CI: 0.00, 0.04) and red meat (b: 0.03; 95% CI: 0.01, 0.06) were positively associated with 3-methyl-histidine concentration, while seafood intake was positively associated with TMAO concentration (b: 0.08; 95% CI: 0.04, 0.12) (Table 2).

For non-dietary factors, maternal age, gestational age, and smoking history were associated with serum concentration of some metabolites after adjusting for the diet-related factors (Table 2). Higher maternal age was associated with a higher concentration of proline betaine (b: 0.04; 95% CI: 0.01, 0.07) and TMAO (b: 0.02; 95% CI: 0.00, 0.04), and higher gestational age of pregnancy was associated with a higher concentration of 3-

methyl-histidine (b: 0.01; 95% CI: 0.00, 0.02) and lower concentration of carnitine (b: -0.01; 95% CI: -0.02, -0.01). Participants who indicated ever smoking cigarettes had lower concentration of proline betaine (b: -0.60; 95% CI: -0.95, -0.25) and higher concentration of carnitine (b: 0.06; 95% CI: 0.02, 0.10) compared to those who never smoked cigarettes (Table 3.2). Parity, GDM, pre-pregnancy BMI, physical activity, SDI, and timing of the administration of the FFQ (before or after blood draw relative to at the same time as the blood draw) were found to not be associated with any of the six metabolite concentration outcomes. The results for the HLM models examining the association of dietary and non-dietary factors with food-related metabolites stratified by ethnicity (White European and South Asians) are presented in Supplementary Tables S3.3 and S3.4, respectively. The results between the two cohorts were generally similar to those reported for the overall sample.

The results from the OLS regression models examining the association of dietary and non-dietary factors with NEFAs are presented in Table 3.3. Higher intake of full-fat dairy was positively associated with odd-chain SFAs 15:0 (b: 0.06; 95% CI: 0.03, 0.10) and 17:0 (b: 0.04; 95% CI: 0.01, 0.07), and higher fish/fish oil daily servings was positively associated with DHA (b: 0.11; 95% CI: 0.07, 0.14) and EPA+DHA (b: 0.08; 95% CI: 0.04, 0.12). For non-dietary factors, higher gestational age of pregnancy was associated with lower odd-chain SFAs 15:0 (b: -0.02; 95% CI: -0.03, -0.01) and 17:0 (b: -0.01; 95% CI: -0.02, -0.01), higher pre-pregnancy BMI was associated with both lower percentage concentrations (mol%) of even-chain and odd-chain SFAs 14:0, 15:0, and 17:0 (b: -0.01; 95% CI: -0.02, -0.00) and lower DHA (b: -0.01; 95% CI: -0.02, -0.00),

and higher physical activity was associated with lower 17:0 (b: -0.10; 95% CI: -0.17, -0.02). The results examining the association of dietary fish intake and ω-3 PUFA are presented in Supplementary Table S3.5.

3.4.2 Results from PC-PR2 analysis

The PC-PR2 analysis was utilized to quantify the sources of systematic variability in serum metabolite concentrations, and the results for the overall sample are displayed in Figures 3.1-2 and stratified by cohort are displayed in Supplementary Figures S3.2-3.7. For largely exogenous metabolites such as proline betaine, hippuric acid, and tryptophan betaine, dietary food intake explained the greater proportion of variability in the metabolite than non-dietary factors. Citrus fruit intake explained the largest proportion of variation in proline betaine concentration with a $R^2_{partial}$ value of 10.8%, followed by smoking history (2.5%), maternal age (1.2%), and ethnicity/cohort (1.2%) (Figure 3.1A). Similarly, for hippuric acid, fruits and vegetables intake displayed the largest $R^2_{partial}$ value of 2.0%, followed closely by energy intake (1.4%) (Figure 3.1B). For tryptophan betaine, intake of nuts and legumes, fiber intake, and overall energy intake explained between 1.2% and 1.9% of the variability. Meanwhile, ethnicity has quite a substantial impact on tryptophan betaine levels as the $R^2_{partial}$ value of cohort was 10.2% (Figure 3.1C). When the model was stratified by cohort, nuts and legumes explained the most variability (3.6%) in the FAMILY cohort (primarily White European women), while fiber intake (4.2%), energy intake (2.2%), and GDM (1.5%) explained most of the variability in tryptophan betaine in the START cohort (exclusively South Asian women) (Supplementary Figure S3.4).

For endogenous (less food-specific) metabolites, the dietary factors explained the most variability for two of the metabolites (3-methyl-histidine and TMAO) while non-dietary factors such as gestational age ($R^2_{partial}$ value: 5.7%) and smoking history ($R^2_{partial}$ value: 1.9%) appeared to play a more prominent role in explaining the variability in carnitine (Figure 3.1D). This latter finding is also consistent with the results obtained from HLM showing no dietary factor was associated with carnitine concentration. Seafood intake explained the greatest proportion of variability in TMAO with a $R^2_{partial}$ value of around 3.0%, followed by maternal age ($R^2_{partial}$ value: 1.2%) (Figure 3.1E). For 3-methyl-histidine, red meat intake had the highest $R^2_{partial}$ value of 1.2% (Figure 3.1F). There was evidence of differences by ethnicity/cohort where red meat explained 5.8% of the variability in 3-methyl-histidine in the START cohort but a negligible amount in the FAMILY cohort. Each of the remaining explanatory variables explained negligible amount of total variation in the metabolite concentrations. Although there were some differences in findings between the two cohorts, overall, the results obtained from PC-PR2 are congruent with those obtained from the HLM analysis.

For NEFAs, pre-pregnancy BMI ($R^2_{partial}$ value: 1.8%) explained the most variability in even-chain SFA 14:0 (Figure 3.2A). Gestational age explained most variability in odd-chain SFAs 15:0 ($R^2_{partial}$ value: 6.9%) and 17:0 ($R^2_{partial}$ value: 3.6%), followed by full-fat dairy intake ($R^2_{partial}$ value: 5.9%) and pre-pregnancy BMI ($R^2_{partial}$ value: 2.4%) for 15:0, and physical activity ($R^2_{partial}$ value: 3.0%) and full-fat dairy intake ($R^2_{partial}$ value: 2.6%) for 17:0 (Figures 3.2B-C). Fish/fish oil intake explained the greatest

proportion of variability in DHA ($R^2_{partial}$ value: 11.2%), followed by followed by pre-pregnancy BMI ($R^2_{partial}$ value: 2.5%) (Figure 3.2E).

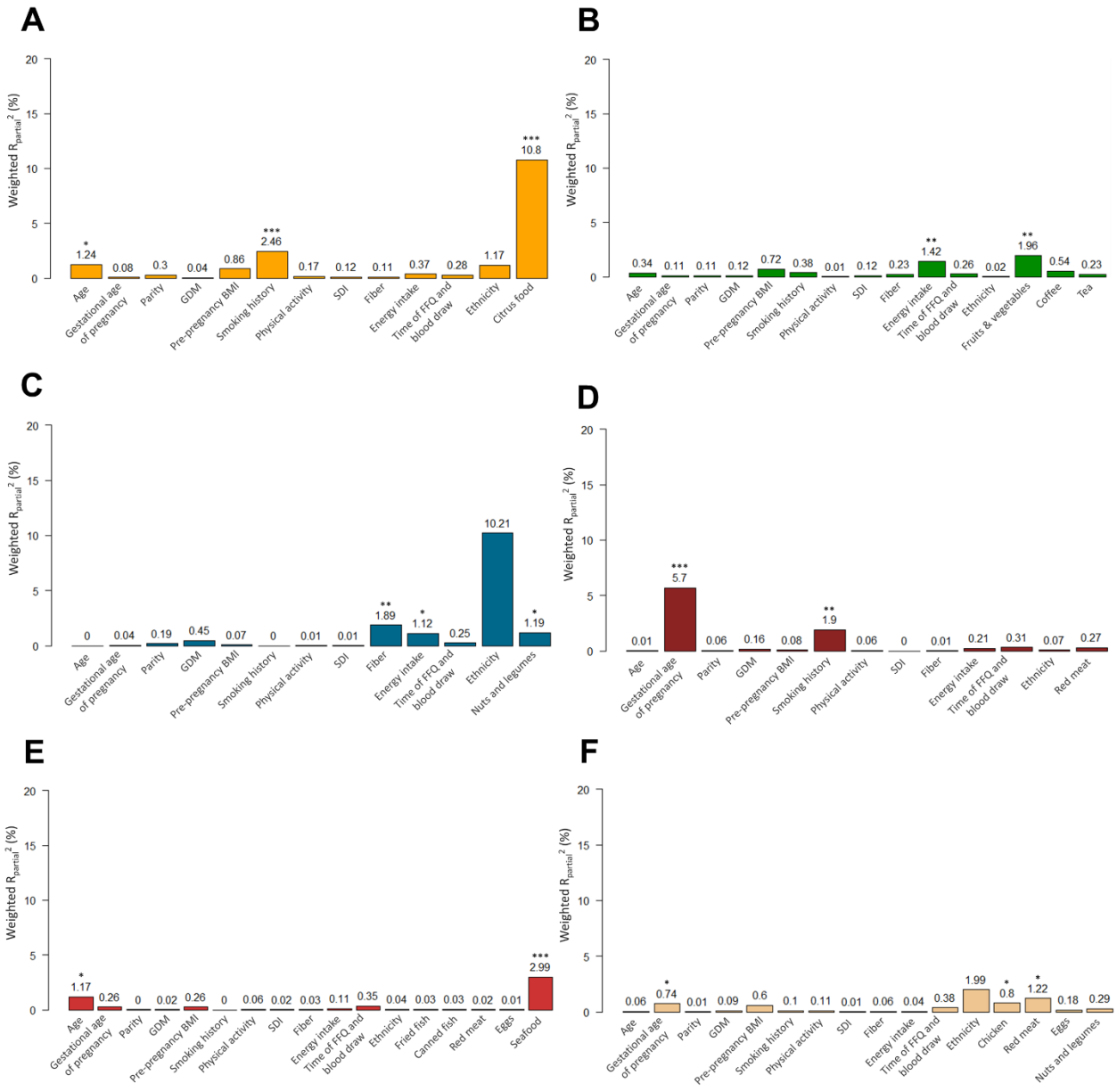


Figure 3.1: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in: (A) Proline betaine, (B) Hippuric acid, (C) Tryptophan betaine, (D) Carnitine, (E) trimethylamine N-oxide (TMAO), and (F) 3-methylhistidine**.** Statistical

significance was based on hierarchical linear models. * $p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. Intraclass correlation suggested a cluster effect by ethnicity (level-two factor) for proline betaine (ICC = 3.9%), tryptophan betaine (ICC = 46.0%), TMAO (ICC = 7.0%), and 3-methylhistidine (ICC = 25.6%), and did not suggest a cluster effect by ethnicity for hippuric acid (ICC = 0.0%) and carnitine (ICC = 1.5%).

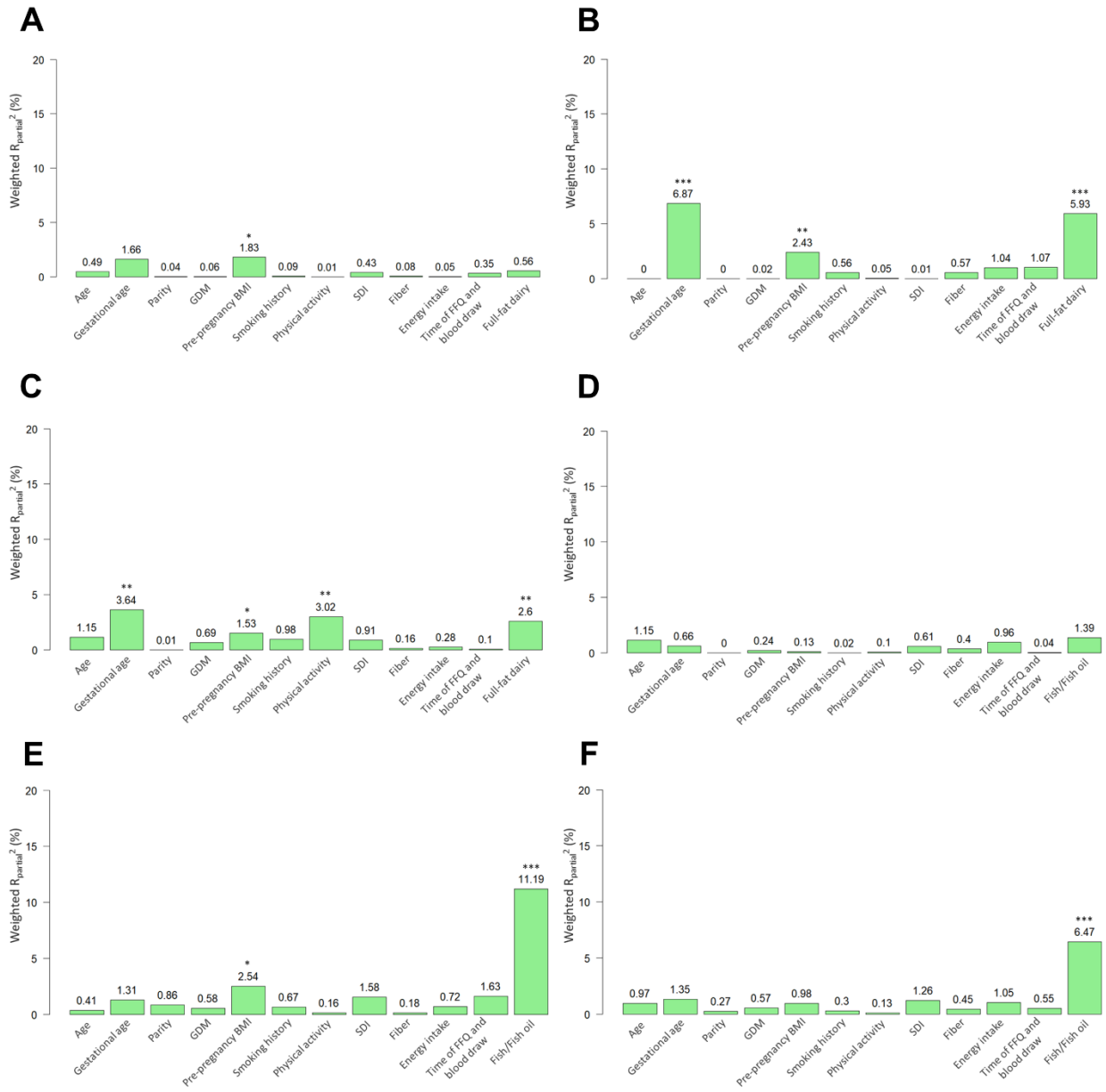


Figure 3.2: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in: (A) Myristic acid (14:0), (B) Pentadecanoic acid (15:0), (C) Heptadecanoic acid (17:0), (D) Eicosapentaenoic acid (EPA, 20:5n-3), (E) Docosahexaenoic acid (DHA; 22:6n-3), and (F) EPA + DHA in FAMILY cohort. Statistical significance was based on ordinary least squares regression. * $p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

3.5 Discussion

Using data from two birth cohorts representing two ethnically diverse groups, the results showed that for exogenous biomarkers such as proline betaine and (largely) DHA, dietary factors explained higher proportion of variability whereas the contribution of nondietary factors was relatively little. On the contrary, for metabolites that can either be produced endogenously, biotransformed by gut microbiota, and/or derived from more than one food source, the unique contribution of dietary factors was similar (15:0, 17:0, hippuric acid, and TMAO) or lower (14:0, tryptophan betaine, 3-methylhistidine, and carnitine) compared to non-dietary factors (ethnicity, maternal age, gestational age, pre-pregnancy BMI, physical activity, and smoking history). Further, there was an ethnicity effect for all metabolites, except carnitine and hippuric acid (Supplemental Figure S9).

For the non-dietary factors, higher maternal age was positively associated and ever smoking was inversely associated with proline betaine concentrations after adjusting for citrus foods. Evidence indicates that older women are more likely to make healthier choices including increasing their consumption of fruits and vegetables from pre-pregnancy to pregnancy compared to younger women (29, 30). Many studies have also shown that smokers have lower concentrations of antioxidants and elevated concentration

of 8-isoprostane (31, 32), which may be due to low consumption of antioxidants (33), reduced vitamin C absorption, or decreased turnover of vitamin C by free radicals produced from smoking (34). Proline betaine (stachydrine), a marker of citrus foods, which are rich in vitamin C (potent water-soluble antioxidant), has been shown to inhibit cell proliferation and production of reactive oxygen species in in-vitro and in-vivo studies (35, 36). As expected, higher citrus food intake was associated with proline betaine concentration and explained the largest proportion of variation in proline betaine concentration relative to non-dietary factors. In kinetics studies, proline betaine is excreted rapidly and nearly completely in urine within 24 hours (37), and therefore it is considered to be minimally metabolized in humans. Further, proline betaine was previously validated in a large-scale observational study, where it was highly sensitive (86.3%) and specific (90.6%) for citrus fruit consumption (37), and thus considered a robust biomarker for citrus food intake.

Even-chain SFA (14:0) can be derived from both exogenous sources (via dietary intake) and endogenous synthesis (via de novo lipogenesis) (38, 39), whereas odd-chain SFAs (15:0 and 17:0) mainly reflect dietary intake of full-fat dairy (40), though the possible contribution of endogenous sources cannot be ruled out (41, 42). As expected, both 15:0 and 17:0 were associated with full-fat dairy intake, and 14:0 was not. Full-fat dairy intake did not, however, explain the largest variance in 15:0 or 17:0 levels. Rather, non-dietary factors including higher gestational age and pre-pregnancy BMI were associated with lower odd chain SFA (15:0 and 17:0) and low physical activity level was associated with lower 17:0. In a previous longitudinal analysis, odd-chain SFA (sum of

15:0 and 17:0) progressively declined during pregnancy (43). Although the exact mechanism for the gestational alterations in these SFAs remain unclear, it is possible that pregnancy associated physiologic changes and increase in adipose deposition throughout pregnancy may be important factors contributing to the observed differences (44). In several population-based studies, higher circulating odd-chain SFAs (15:0 and 17:0) were inversely associated with obesity and cardiometabolic diseases (45, 46). ω-3 PUFAs (DHA more than EPA) have been considered robust biomarkers of habitual fish/fish oil intake (2). This association was demonstrated for DHA in the current study where fish/fish oil intake explained the largest proportion of variation in DHA relative to non-dietary factors. Fish/fish oil daily servings explains about twice the amount of variation in ω-3 PUFAs compared with dietary fish intake, indicating that it is important to account for EPA and DHA sources from both diet and supplements.

For other metabolites, non-dietary factors were associated with metabolite concentrations, however, their overall contribution was minimal, except for carnitine which was mostly explained by gestational age. Carnitine mainly reflects consumption of amino acids and fatty acid-containing foods and as a result is considered a generic marker for foods of animal origin but also may be synthesized from essential amino acids lysine and methionine (4, 47). A decline in carnitine across trimesters during pregnancy was previously reported (9, 48). A significant rise in acylcarnitine in pregnant women as pregnancy progresses may reflect enhanced fatty acid oxidation in later periods of gestation (48). This distribution may suggest a greater uptake of carnitine in the fatty acid

β-oxidation process leading to lower free carnitine substrate and resulting lower total body carnitine pool in pregnant women (49, 50).

For all metabolites except for proline betaine and two NEFAs (15:0 and DHA), the unique contribution of food sources was similar to or lower than non-dietary factors. This may reflect endogenous production, microbial synthesis, or multiple food sources of some of these metabolites. Interindividual variability in hippuric acid (51, 52), TMAO (53, 54), and tryptophan betaine (55, 56) may partly be due to differences in intestinal microbiota. However, the potential variation in these metabolites attributable to the gut microbiome could not be accounted for in our study. Further, variation in an endogenous metabolite concentration such as carnitine may reflect general intake of foods of animal origin and/or physiological changes that take place during pregnancy, and is influenced by factors such as age and health status, and thus may not be a suitable biomarker of red meat at the population level (4, 47).

Metabolite concentration may also vary widely across cultures and ethnic groups as the type of food, method of consumption, and food preparation techniques may vary (57). In our multi-level analysis, there was an ethnicity effect for all metabolites, except carnitine and hippuric acid. Proline betaine concentration was shown to vary to some extent by cohort, likely attributable to differences in citrus food intake in the two cohorts (Table 3.1). Also, some of this variability may be attributed to differences in lifestyle factors between members of the two cohorts, such as smoking status. Regardless, citrus fruit consumption still explained the largest amount of variance in proline betaine in both cohorts, suggesting that non-dietary factors do not contribute substantially to proline

betaine variation (Supplementary Figure S3.2). However, mixed results were shown by cohort for metabolites that are synthesized or modified by gut bacteria. Tryptophan betaine concentration was shown to be vary considerably between the two cohorts, with higher tryptophan betaine associated with higher nuts and legumes intake in the FAMILY cohort, and with higher fiber intake and lower kilocalories in the START cohort. A possible explanation for this discrepancy may be that nuts and legumes is a heterogeneous food group so the type of nuts and preparation/cooking methods for legumes may play an important part (58). Further, it is also likely that the association of nuts or legumes intake with tryptophan betaine may be confounded by fiber intake in the START cohort as fiber intake is higher in this cohort, and tryptophan betaine has been identified in fiber-rich plant-based foods and linked to gut microbiota in fiber-enriched diets (55).

Hippuric acid was one of the metabolites that did not vary by ethnic cohort but was only associated with greater fruit and vegetable intake in the FAMILY cohort despite greater intake in the START cohort. An explanation for this may be related to metabolism of different dietary polyphenols (59). Evidence suggests that differences in excretion of hippuric acid may reflect altered gut microbial metabolism (60). Generally, amount of variability in the food consumption may also affect the robustness of the association. For example, the IQR for certain foods such as chicken and red meat were higher in FAMILY compared to START, whereas variability for other foods such as fruits and vegetables, tea, eggs, and nuts and legumes were higher in START compared

to FAMILY. This may explain inconsistencies in the results for at least some serum metabolites such as association between red meat and TMAO in the START cohort.

In other comparisons, TMAO varied slightly by cohort, but may be explained by a relatively lower consumption of meats including red meat, canned and fried fish, and seafood in START cohort compared to FAMILY cohort. Despite this, higher seafood intake was positively associated with TMAO concentration in both cohorts. Differences in TMAO production and excretion may partly be related to metabolic precursors such as choline, betaine, and carnitine. TMAO concentration increases postprandially (within 15 mins) after consumption of fish (61), but it takes more time after consumption of meat (62), suggesting that free TMAO in seafood may be readily absorbed after fish consumption without much involvement of gut microbiota. Finally, although the association of 3-methylhistidine with chicken and red meat was significant in the overall sample, these associations were attenuated when analysis was stratified by cohort. This is likely because the intra-cohort variability was small, or intakes of these foods were highly correlated (as was the case in the START cohort).

Finally, biomarkers with ‘good’ evidence are considered as direct surrogates for food intake (63). However, there are several factors, in addition to food exposure, that can influence variation in food-related metabolites concentration and thus require appropriate consideration during the statistical analyses of the data (64). In line with previous research (7), in most cases, our study found that dietary factors explained less than 10% of the total variation in metabolite concentration. While some of the source of error is explained by measurement error (self-report), other can be related to non-dietary

factors. Therefore, future studies should account for non-dietary factors and differences by ethnicity to control for some of the inter-individual variation in food-related metabolites.

3.6 Limitations

Our study has several strengths including a large sample size that allowed for stratification by ethnicity, use of fasting serum samples, and comparing a diverse set of metabolites reflecting commonly consumed foods which have been previously reported in free-living population studies (2). We adopted a novel methodological approach to address an unanswered question regarding non-dietary sources of metabolites variation in the field of nutritional metabolomics and biomarkers of food intake. Our study also has some limitations. We included only pregnant women from white European and South Asian backgrounds, and thus generalizability of our findings is limited to these populations. Dietary assessment was based on a self-reported FFQ and maybe prone to some measurement error, however, FFQs are commonly used in nutritional epidemiology. The period of dietary assessment of 12 months may not be indicative of recent intake of foods or intake of foods only during pregnancy, but since our aim was to identify sources of variability in metabolites of foods that reflect habitual dietary intake, a 12-month intake was more appropriate. Samples were collected at one point in pregnancy and data on changes in dietary intake during pregnancy were not collected and therefore, not available for the analysis.

3.7 Conclusions

Overall, the results emphasize that serum metabolites that reflect specific foods are also influenced by non-dietary factors (ethnicity, maternal age, gestational age, prepregnancy BMI, physical activity, and smoking history) but to differing degrees. The results of this study provide insight into the external factors that impact serum metabolite concentrations and provide guidance on appropriate modeling when metabolomics is used in nutritional epidemiological studies to identify diet-disease associations. Identifying robust and generalized food related biomarkers in diverse populations remains a challenge, but appropriate adjustment for non-dietary factors is necessary for an unbiased assessment of metabolite concentration. Future work will explore the role of maternal nutrition and food exposures on health outcomes later in life, such as childhood obesity and metabolic syndrome.

3.8 References

1. Kipnis V, Midthune D, Freedman L, Bingham S, Day NE, Riboli E, Ferrari P, Carroll RJ. Bias in dietary-report instruments and its implications for nutritional epidemiology. Public Health Nutr 2002;5(6A):915-23. doi: 10.1079/phn2002383.
2. Rafiq T, Azab SM, Teo KK, Thabane L, Anand SS, Morrison KM, de Souza RJ, Britz-McKibbin P. Nutritional Metabolomics and the Classification of Dietary Biomarker Candidates: A Critical Review. Adv Nutr 2021;12(6):2333-57. doi: 10.1093/advances/nmab054.
3. Noh H, Freisling H, Assi N, Zamora-Ros R, Achaintre D, Affret A, Mancini F, Boutron-Ruault MC, Flögel A, Boeing H, et al. Identification of Urinary Polyphenol Metabolite Patterns Associated with Polyphenol-Rich Food Intake in Adults from Four European Countries. Nutrients 2017;9(8):796. doi: 10.3390/nu9080796.
4. Cuparencu C, Praticó G, Hemeryck LY, Sri Harsha PSC, Noerman S, Rombouts C, Xi M, Vanhaecke L, Hanhineva K, Brennan L, et al. Biomarkers of meat and seafood intake: an extensive literature review. Genes & nutrition 2019;14:35. doi: 10.1186/s12263-019-0656-4.
5. Crews HM, Olivier L, Wilson LA. Urinary biomarkers for assessing dietary exposure to caffeine. Food Addit Contam 2001;18(12):1075-87. doi: 10.1080/02652030110056630.
6. Vernocchi P, Del Chierico F, Putignani L. Gut Microbiota Metabolism and Interaction with Food Components. Int J Mol Sci 2020;21(10). doi: 10.3390/ijms21103688.
7. Floegel A, von Ruesten A, Drogan D, Schulze MB, 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(10):1100-8. doi: 10.1038/ejcn.2013.147.
8. Dogan B, Karaer A, Tuncay G, Tecellioglu N, Mumcu A. High-resolution (1)H-NMR spectroscopy indicates variations in metabolomics profile of follicular fluid from women with advanced maternal age. Journal of assisted reproduction and genetics 2020;37(2):321-30. doi: 10.1007/s10815-020-01693-x.
9. Ryckman KK, Donovan BM, Fleener DK, Bedell B, Borowski KS. Pregnancy-Related Changes of Amino Acid and Acylcarnitine Concentrations: The Impact of Obesity. AJP reports 2016;6(3):e329-36. doi: 10.1055/s-0036-1592414.
10. Canova C, Cantarutti A. Population-Based Birth Cohort Studies in Epidemiology. Int J Environ Res Public Health 2020;17(15). doi: 10.3390/ijerph17155276.
11. Li PY, Zhong CR, Li SZ, Sun TP, Huang H, Chen X, Zhu YL, Hu XL, Peng XB, Zhang X, et al. Plasma concentration of trimethylamine-N-oxide and risk of gestational diabetes mellitus. Am J Clin Nutr 2018;108(3):603-10. doi: 10.1093/ajcn/nqy116.
12. Janeiro MH, Ramírez MJ, Milagro FI, Martínez JA, Solas M. Implication of Trimethylamine N-Oxide (TMAO) in Disease: Potential Biomarker or New Therapeutic Target. Nutrients 2018;10(10):1398. doi: 10.3390/nu10101398.

13. Morrison KM, Atkinson SA, Yusuf S, Bourgeois J, McDonald S, McQueen MJ, Persadie R, Hunter B, Pogue J, Teo K. The Family Atherosclerosis Monitoring In earLY life (FAMILY) study: Rationale, design, and baseline data of a study examining the early determinants of atherosclerosis. Am Heart J 2009;158(4):533-9. doi: 10.1016/j.ahj.2009.07.005.
14. Anand SS, Vasudevan A, Gupta M, Morrison K, Kurpad A, Teo KK, Srinivasan K. Rationale and design of South Asian Birth Cohort (START): a Canada-India collaborative study. BMC Public Health 2013;13:79. doi: 10.1186/1471-2458-13-79.
15. Shanmuganathan M, Kroezen Z, Gill B, Azab S, de Souza RJ, Teo KK, Atkinson S, Subbarao P, Desai D, Anand SS, et al. The maternal serum metabolome by multisegment injection-capillary electrophoresis-mass spectrometry: a high-throughput platform and standardized data workflow for large-scale epidemiological studies. Nature protocols 2021;16(4):1966-94. doi: 10.1038/s41596-020-00475-0.
16. DiBattista A, McIntosh N, Lamoureux M, Al-Dirbashi OY, Chakraborty P, Britz-McKibbin P. Metabolic Signatures of Cystic Fibrosis Identified in Dried Blood Spots For Newborn Screening Without Carrier Identification. J Proteome Res 2019;18(3):841-54. doi: 10.1021/acs.jproteome.8b00351.
17. de Souza RJ, Shanmuganathan M, Lamri A, Atkinson SA, Becker A, Desai D, Gupta M, Mandhane PJ, Moraes TJ, Morrison KM, et al. Maternal Diet and the Serum Metabolome in Pregnancy: Robust Dietary Biomarkers Generalizable to a Multiethnic Birth Cohort. Curr Dev Nutr 2020;4(10):nzaa144. doi: 10.1093/cdn/nzaa144.
18. Wellington N, Shanmuganathan M, de Souza RJ, Zulyniak MA, Azab S, Bloomfield J, Mell A, Ly R, Desai D, Anand SS, et al. Metabolic Trajectories Following Contrasting Prudent and Western Diets from Food Provisions: Identifying Robust Biomarkers of Short-Term Changes in Habitual Diet. Nutrients 2019;11(10):2407. doi: 10.3390/nu11102407.
19. Kelemen LE, Anand SS, Vuksan V, Yi Q, Teo KK, Devanesen S, Yusuf S. Development and evaluation of cultural food frequency questionnaires for South Asians, Chinese, and Europeans in North America. J Am Diet Assoc 2003;103(9):1178-84. doi: 10.1016/s0002-8223(03)00985-4.
20. Chevli PA, Hari KJ, Kanaya AM, Talegawkar SA, Needham BL, Herrington D. Association of Alcohol Consumption and Ideal Cardiovascular Health Among South Asians: The Mediators of Atherosclerosis in South Asians Living in America (MASALA) Study. Alcohol Clin Exp Res 2020;44(9):1825-33. doi: 10.1111/acer.14422.
21. Talegawkar SA, Kandula NR, Gadgil MD, Desai D, Kanaya AM. Dietary intakes among South Asian adults differ by length of residence in the USA. Public Health Nutr 2016;19(2):348-55. doi: 10.1017/s1368980015001512.
22. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr 1997;65(4 Suppl):1220S-8S; discussion 9S-31S. doi: 10.1093/ajcn/65.4.1220S.

23. Anand SS, Razak F, Davis AD, Jacobs R, Vuksan V, Teo K, Yusuf S. Social disadvantage and cardiovascular disease: development of an index and analysis of age, sex, and ethnicity effects. Int J Epidemiol 2006;35(5):1239-45. doi: 10.1093/ije/dyl163.
24. Raudenbush SW, Bryk AS. Hierarchical linear models: Applications and data analysis methods. Second ed. Thousand Oaks: Sage Publications, 2002.
25. Heo M, Faith MS, Mott JW, Gorman BS, Redden DT, Allison DB. Hierarchical linear models for the development of growth curves: an example with body mass index in overweight/obese adults. Stat Med 2003;22(11):1911-42. doi: 10.1002/sim.1218.
26. Feldman HA. Families of lines: random effects in linear regression analysis. J Appl Physiol 1988;64(4):1721-32. doi: 10.1152/jappl.1988.64.4.1721.
27. Raftery AE. Bayesian Model Selection in Social Research. Sociological Methodology 1995;25:111-63. doi: 10.2307/271063.
28. Fages A, Ferrari P, Monni S, Dossus L, Floegel A, Mode N, Johansson M, Travis RC, Bamia C, Sanchez-Perez MJ, et al. Investigating sources of variability in metabolomic data in the EPIC study: the Principal Component Partial R-square (PC-PR2) method. Metabolomics 2014;10(6):1074-83. doi: 10.1007/s11306-014-0647-9.
29. Crozier SR, Robinson SM, Borland SE, Godfrey KM, Cooper C, Inskip HM. Do women change their health behaviours in pregnancy? Findings from the Southampton Women's Survey. Paediatr Perinat Epidemiol 2009;23(5):446-53. doi: 10.1111/j.1365-3016.2009.01036.x.
30. Skreden M, Bere E, Sagedal LR, Vistad I, Øverby NC. Changes in fruit and vegetable consumption habits from pre-pregnancy to early pregnancy among Norwegian women. BMC Pregnancy Childbirth 2017;17(1):107. doi: 10.1186/s12884-017-1291-y.
31. Morrow JD, Frei B, Longmire AW, Gaziano JM, Lynch SM, Shyr Y, Strauss WE, Oates JA, Roberts LJ, 2nd. Increase in circulating products of lipid peroxidation (F2-isoprostanes) in smokers. Smoking as a cause of oxidative damage. N Engl J Med 1995;332(18):1198-203. doi: 10.1056/nejm199505043321804.
32. Eiserich JP, van der Vliet A, Handelman GJ, Halliwell B, Cross CE. Dietary antioxidants and cigarette smoke-induced biomolecular damage: a complex interaction. Am J Clin Nutr 1995;62(6 Suppl):1490S-500S. doi: 10.1093/ajcn/62.6.1490S.
33. Zondervan KT, Ocké MC, Smit HA, Seidell JC. Do dietary and supplementary intakes of antioxidants differ with smoking status? Int J Epidemiol 1996;25(1):70-9. doi: 10.1093/ije/25.1.70.
34. Marangon K, Herbeth B, Lecomte E, Paul-Dauphin A, Grolier P, Chancerelle Y, Artur Y, Siest G. Diet, antioxidant status, and smoking habits in French men. Am J Clin Nutr 1998;67(2):231-9. doi: 10.1093/ajcn/67.2.231.
35. Wang M, Shu ZJ, Wang Y, Peng W. Stachydrine hydrochloride inhibits proliferation and induces apoptosis of breast cancer cells via inhibition of Akt and ERK pathways. Am J Transl Res 2017;9(4):1834-44.

36. Cao TT, Chen HH, Dong Z, Xu YW, Zhao P, Guo W, Wei HC, Zhang C, Lu R. Stachydrine Protects Against Pressure Overload-Induced Cardiac Hypertrophy by Suppressing Autophagy. Cell Physiol Biochem 2017;42(1):103-14. doi: 10.1159/000477119.
37. Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas ME, Kochhar S, Stamler J, Holmes E, Elliott P, Nicholson JK. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. American Journal of Clinical Nutrition 2010;92(2):436-43. doi: 10.3945/ajcn.2010.29672.
38. Hudgins LC, Hellerstein M, Seidman C, Neese R, Diakun J, Hirsch J. Human fatty acid synthesis is stimulated by a eucaloric low fat, high carbohydrate diet. J Clin Invest 1996;97(9):2081-91. doi: 10.1172/jci118645.
39. Hellerstein MK, Schwarz JM, Neese RA. Regulation of hepatic de novo lipogenesis in humans. Annu Rev Nutr 1996;16:523-57. doi: 10.1146/annurev.nu.16.070196.002515.
40. Hodson L, Skeaff CM, Fielding BA. Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake. Prog Lipid Res 2008;47(5):348-80. doi: 10.1016/j.plipres.2008.03.003.
41. Ratnayake WM. Concerns about the use of 15:0, 17:0, and trans-16:1n-7 as biomarkers of dairy fat intake in recent observational studies that suggest beneficial effects of dairy food on incidence of diabetes and stroke. Am J Clin Nutr 2015;101(5):1102-3. doi: 10.3945/ajcn.114.105379.
42. Pfeuffer M, Jaudszus A. Pentadecanoic and Heptadecanoic Acids: Multifaceted Odd-Chain Fatty Acids. Adv Nutr 2016;7(4):730-4. doi: 10.3945/an.115.011387.
43. Lepsch J, Farias DR, Vaz Jdos S, de Jesus Pereira Pinto T, da Silva Lima N, Freitas Vilela AA, Cunha M, Factor-Litvak P, Kac G. Serum saturated fatty acid decreases plasma adiponectin and increases leptin throughout pregnancy independently of BMI. Nutrition 2016;32(7-8):740-7. doi: 10.1016/j.nut.2016.01.016.
44. Ukkola O, Santaniemi M. Adiponectin: a link between excess adiposity and associated comorbidities? J Mol Med (Berl) 2002;80(11):696-702. doi: 10.1007/s00109-002-0378-7.
45. Aglago EK, Biessy C, Torres-Mejía G, Angeles-Llerenas A, Gunter MJ, Romieu I, Chajès V. Association between serum phospholipid fatty acid levels and adiposity in Mexican women. J Lipid Res 2017;58(7):1462-70. doi: 10.1194/jlr.P073643.
46. Khaw KT, Friesen MD, Riboli E, Luben R, Wareham N, Khaw K-T, Friesen MD, Riboli E, Luben R, Wareham N. Plasma phospholipid fatty acid concentration and incident coronary heart disease in men and women: the EPIC-Norfolk prospective study. PLoS Medicine 2012;9(7):e1001255-e. doi: 10.1371/journal.pmed.1001255.
47. Mitchell ME. Carnitine metabolism in human subjects. I. Normal metabolism. Am J Clin Nutr 1978;31(2):293-306. doi: 10.1093/ajcn/31.2.293.

48. Talián GC, Komlósi K, Decsi T, Koletzko B, Melegh B. Determination of carnitine ester patterns during the second half of pregnancy, at delivery, and in neonatal cord blood by tandem mass spectrometry: complex and dynamic involvement of carnitine in the intermediary metabolism. Pediatr Res 2007;62(1):88-92. doi: 10.1203/PDR.0b013e3180676cca.
49. Oey NA, van Vlies N, Wijburg FA, Wanders RJ, Attie-Bitach T, Vaz FM. L-carnitine is synthesized in the human fetal-placental unit: potential roles in placental and fetal metabolism. Placenta 2006;27(8):841-6. doi: 10.1016/j.placenta.2005.10.002.
50. Grube M, Meyer Zu Schwabedissen H, Draber K, Präger D, Möritz KU, Linnemann K, Fusch C, Jedlitschky G, Kroemer HK. Expression, localization, and function of the carnitine transporter octn2 (slc22a5) in human placenta. Drug Metab Dispos 2005;33(1):31-7. doi: 10.1124/dmd.104.001560.
51. van Duynhoven J, van der Hooft JJ, van Dorsten FA, Peters S, Foltz M, Gomez-Roldan V, Vervoort J, de Vos RC, Jacobs DM. Rapid and sustained systemic circulation of conjugated gut microbial catabolites after single-dose black tea extract consumption. Journal of Proteome Research 2014;13(5):2668-78.
52. Williams HRT, Cox IJ, Walker DG, Cobbold JFL, Taylor-Robinson SD, Marshall SE, Orchard TR. Differences in gut microbial metabolism are responsible for reduced hippurate synthesis in Crohn's disease. Bmc Gastroenterology 2010;10. doi: 10.1186/1471-230x-10-108.
53. Wang Z, Bergeron N, Levison BS, Li XS, Chiu S, Jia X, Koeth RA, Li L, Wu Y, Tang WHW, et al. Impact of chronic dietary red meat, white meat, or non-meat protein on trimethylamine N-oxide metabolism and renal excretion in healthy men and women. Eur Heart J 2019;40(7):583-94. doi: 10.1093/eurheartj/ehy799.
54. Li J, Li Y, Ivey KL, Wang DD, Wilkinson JE, Franke A, Lee KH, Chan A, Huttenhower C, Hu FB, et al. Interplay between diet and gut microbiome, and circulating concentrations of trimethylamine N-oxide: findings from a longitudinal cohort of US men. Gut 2021. doi: 10.1136/gutjnl-2020-322473.
55. Koistinen VM, Kärkkäinen O, Borewicz K, Zarei I, Jokkala J, Micard V, Rosa-Sibakov N, Auriola S, Aura AM, Smidt H, et al. Contribution of gut microbiota to metabolism of dietary glycine betaine in mice and in vitro colonic fermentation. Microbiome 2019;7(1):103. doi: 10.1186/s40168-019-0718-2.
56. Wu GD, Compher C, Chen EZ, Smith SA, Shah RD, Bittinger K, Chehoud C, Albenberg LG, Nessel L, Gilroy E, et al. Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. Gut 2016;65(1):63-72. doi: 10.1136/gutjnl-2014-308209.
57. Sharma S, Sheehy T, Kolonel LN. Ethnic differences in grains consumption and their contribution to intake of B-vitamins: results of the Multiethnic Cohort Study. Nutr J 2013;12:65. doi: 10.1186/1475-2891-12-65.
58. Wang N, Hatcher DW, Tyler RT, Toews R, Gawalko EJ. Effect of cooking on the composition of beans (Phaseolus vulgaris L.) and chickpeas (Cicer arietinum L.). Food Res Int 2010;43(2):589-94. doi: 10.1016/j.foodres.2009.07.012.

59. Temellini A, Mogavero S, Giulianotti PC, Pietrabissa A, Mosca F, Pacifici GM. Conjugation of benzoic acid with glycine in human liver and kidney: a study on the interindividual variability. Xenobiotica 1993;23(12):1427-33. doi: 10.3109/00498259309059451.
60. Pallister T, Jackson MA, Martin TC, Zierer J, Jennings A, Mohney RP, MacGregor A, Steves CJ, Cassidy A, Spector TD, et al. Hippurate as a metabolomic marker of gut microbiome diversity: Modulation by diet and relationship to metabolic syndrome. Sci Rep 2017;7(1):13670. doi: 10.1038/s41598-017-13722-4.
61. Cheung W, Keski-Rahkonen P, Assi N, Ferrari P, Freisling H, Rinaldi S, Slimani N, Zamora-Ros R, Rundle M, Frost G, et al. A metabolomic study of biomarkers of meat and fish intake. Am J Clin Nutr 2017;105(3):600-8. doi: 10.3945/ajcn.116.146639.
62. Cho CE, Taesuwan S, Malysheva OV, Bender E, Tulchinsky NF, Yan J, Sutter JL, Caudill MA. Trimethylamine-N-oxide (TMAO) response to animal source foods varies among healthy young men and is influenced by their gut microbiota composition: A randomized controlled trial. Mol Nutr Food Res 2017;61(1). doi: 10.1002/mnfr.201600324.
63. Picó C, Serra F, Rodríguez AM, Keijer J, Palou A. Biomarkers of Nutrition and Health: New Tools for New Approaches. Nutrients 2019;11(5). doi: 10.3390/nu11051092.
64. Maruvada P, Lampe JW, Wishart DS, Barupal D, Chester DN, Dodd D, Djoumbou-Feunang Y, Dorrestein PC, Dragsted LO, Draper J, et al. Perspective: Dietary Biomarkers of Intake and Exposure—Exploration with Omics Approaches. Adv Nutr 2020;11(2):200-15. doi: 10.1093/advances/nmz075.

Table 3.1: Descriptive statistics of participants overall and by ethnicity.

Factor	Overall n = 600	White European n = 300	South Asian n = 300	p-value
Age (years), mean (SD)	31.20 (4.50)	32.35 (4.89)	30.01 (3.73)	<0.0001
Gestational age (weeks), mean (SD)	28.06 (3.27)	29.50 (3.76)	26.61 (1.75)	<0.0001
Pre-pregnancy BMI (kg/m^2), mean (SD)	25.35 (5.63)	26.77 (6.39)	23.94 (4.33)	<0.0001
Parity, n (%)				
0	240 (42.33)	145 (48.33)	95 (35.58)	0.0528
1	229 (40.39)	110 (36.67)	119 (44.57)	
2	76 (13.40)	34 (11.33)	42 (15.73)	
≥3	22 (3.88)	11 (3.67)	11 (4.12)	
Gestational diabetes (GDM), n (%)[a]	169 (28.94)	50 (17.54)	119 (39.80)	<0.0001
Smoking history (ever smoked), n (%)	104 (17.48)	104 (35.25)	0 (0.00)	<0.0001
Physical activity (moderate/vigorous), n (%)	144 (24.04)	84 (28.00)	60 (20.07)	0.0231
Social disadvantage index, mean (SD)[b]	1.31 (1.37)	0.85 (1.22)	1.84 (1.35)	<0.0001

Fiber intake (g/day), mean (SD)	22.52 (10.24)	20.66 (9.23)	24.38 (10.85)	<0.0001
Energy Intake (kcal), mean (SD)	2165.39 (772.06)	2327.86 (766.33)	2002.92 (744.26)	<0.0001
Time of FFQ and blood draw, n (%)				
FFQ and blood draw on same day	354 (60.31)	88 (29.33)	266 (92.68)	<0.0001
FFQ before blood draw[c]	221 (37.65)	206 (68.67)	15 (5.23)	
FFQ after blood draw[c]	12 (2.04)	6 (2.00)	6 (2.09)	
Food items (servings/day), median (IQR)				
Citrus food	0.57 (0.95)	0.64 (0.99)	0.43 (0.89)	<0.0001
Fruits and vegetables	6.28 (5.74)	5.12 (4.26)	7.85 (6.06)	<0.0001
Tea	0.43 (0.98)	0.14 (0.57)	1.0 (1.36)	<0.0001
Coffee	0 (0.14)	0.02 (0.64)	0 (0.00)	<0.0001
Canned fish	0 (0.03)	0.03 (0.07)	0 (0.00)	<0.0001
Fried fish	0 (0.03)	0.01 (0.03)	0 (0.02)	<0.0001
Seafood	0 (0.01)	0.01 (0.02)	0 (0.00)	<0.0001
Chicken	0.10 (0.29)	0.14 (0.21)	0 (0.14)	<0.0001

Eggs	0.21 (0.40)	0.20 (0.32)	0.29 (0.57)	0.9927
Red meat	0.20 (0.44)	0.41 (0.35)	0.01 (0.15)	<0.0001
Nuts and legumes	0.71 (0.92)	0.62 (0.83)	0.85 (0.97)	<0.0001
Full-fat dairy	–	1.05 (1.11)	–	–
Fish/fish oil	–	0.08 (0.15)	–	–
Metabolite concentration, median (IQR)				
Proline betaine	1.81 (3.82)	2.33 (5.52)	1.40 (2.47)	<0.0001
Hippuric acid	10.01 (9.87)	9.68 (9.03)	10.07 (10.36)	0.8848
TMAO	2.53 (1.95)	2.68 (1.96)	2.24 (1.99)	<0.0001
3-methylhistidine	7.17 (4.12)	8.64 (4.90)	6.14 (2.24)	<0.0001
Carnitine	15.61 (3.82)	15.35 (3.69)	15.89 (3.98)	0.0117
Tryptophan betaine	1.27 (0.37)	1.19 (0.14)	1.47 (0.37)	<0.0001
Fatty acids, median (IQR)[d]				
Myristic acid (14:0)	–	2.19 (0.74)	–	–
Pentadecanoic acid (15:0)	–	0.24 (0.08)	–	–

Heptadecanoic acid (17:0)	–	0.69 (0.23)	–	–
Eicosapentaenoic acid (EPA or 20:5n-3)	–	0.51 (0.26)	–	–
Docosahexaenoic acid (DHA or 22:6n-3)	–	0.67 (0.29)	–	–

FFQ = Food frequency questionnaire; TMAO = trimethylamine N-oxide Wilcoxon's rank sum test was used to compare continuous variables, and chi-square to compare categorical variables by cohort. [a]GDM was defined based on the Born in Bradford oral glucose tolerance test criteria, self-reported GDM, and insulin use in pregnancy in START cohort, whereas the International Association of the Diabetes and Pregnancy Study Groups criteria [75-g OGTT with fasting glucose ≥5.1 mmol/L, 1 hour ≥10.0 mmol/L, 2 hours ≥8.5 mmol/L] was used in FAMILY cohort. [b]The maximum social disadvantage index was five, and the lowest possible score was zero, reflecting the least social disadvantage. [c]FFQ was implemented within a one-year time period of the blood draw. [d]Fatty acids data were only available in FAMILY cohort.

Table 3.2: Results from random effects hierarchical modelling examining the association of dietary and non-dietary factors with food-based metabolites.

	Proline betaine	**Hippuric acid**	**3-Methyl histidine**	**Carnitine**	**Tryptophan betaine**	**TMAO**
Factor	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**
Age (years)	0.04* (0.01, 0.07)	0.01 (0.00, 0.03)	0.00 (-0.01, 0.01)	0.00 (0.00, 0.00)	0.00 (0.00, 0.03)	0.02* (0.00, 0.04)
Gestational age (weeks)	0.02 (-0.03, 0.06)	0.01 (-0.01, 0.03)	0.01* (0.00, 0.02)	-0.01*** (-0.02, -0.01)	0.00 (0.00, 0.00)	0.01 (-0.01, 0.03)
Parity	-0.10 (-0.25, 0.06)	0.03 (-0.05, 0.11)	-0.01 (-0.05, 0.03)	0.01 (-0.01, 0.02)	-0.01 (-0.02, 0.01)	0.01 (-0.07, 0.09)
Gestational diabetes (GDM)	0.05 (-0.24, 0.35)	0.06 (-0.10, 0.21)	0.02 (-0.05, 0.10)	0.02 (-0.02, 0.05)	0.02 (-0.01, 0.05)	0.03 (-0.13, 0.19)
Pre-pregnancy BMI (kg/m^2)	-0.02 (-0.05, 0.00)	-0.01 (-0.02, 0.00)	-0.01 (-0.01, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	-0.01 (-0.02, 0.01)
Smoking history	-0.60***	-0.12	0.04	0.06**	0.00	-0.01

(ever vs. never smoked)	(-0.95, -0.25)	(-0.30, 0.06)	(-0.06, 0.13)	(0.02, 0.10)	(-0.03, 0.03)	(-0.20, 0.17)
Physical activity (low vs. high)	-0.13	0.02	-0.03	-0.01	0.00	-0.04
	(-0.42, 0.17)	(-0.14, 0.18)	(-0.10, 0.05)	(-0.04, 0.02)	(-0.03, 0.03)	(-0.21, 0.12)
Social disadvantage index	-0.05	-0.02	0.00	0.00	0.00	-0.01
	(-0.15, 0.06)	(-0.08, 0.03)	(-0.03, 0.02)	(-0.01, 0.01)	(-0.01, 0.01)	(-0.06, 0.05)
Fiber intake (g/day)	0.01	0.01	0.00	0.00	2.68×10^{-3}**	0.00
	(-0.01, 0.02)	(-0.01, 0.02)	(-0.01, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(-0.01, 0.01)
Energy intake (kcal)	0.00	-1.6×10^{-4}**	0.00	0.00	-3×10^{-5}*	0.00
	(0.00, 0.00)	(-0.00, -0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
FFQ before blood draw vs. FFQ at	0.02	0.11	-0.05	0.00	-0.01	0.09
the same time as blood draw	(-0.30, 0.35)	(-0.05, 0.27)	(-0.13, 0.04)	(-0.03, 0.04)	(-0.04, 0.02)	(-0.08, 0.26)
FFQ after blood draw vs. FFQ at	0.50	0.08	0.04	0.06	-0.04	-0.11
the same time as blood draw	(-0.34, 1.34)	(-0.37, 0.54)	(-0.18, 0.26)	(-0.04, 0.16)	(-0.12, 0.04)	(-0.56, 0.35)
Citrus food	0.27***					
(servings/day)	(0.20, 0.34)					

Fruits and vegetables	0.22**			
(servings/day)	(0.08, 0.36)			
Tea	0.01			
(servings/day)	(-0.01, 0.04)			
Coffee	0.02			
(servings/day)	(0.00, 0.04)			
Chicken		0.02*		
(servings/day)		(0.00, 0.04)		
Red meat		0.03*	0.00	0.00
(servings/day)		(0.01, 0.06)	(0.00, 0.01)	(-0.04, 0.04)
Eggs		0.01		0.00
(servings/day)		(-0.01, 0.02)		(-0.03, 0.04)
Nuts and legumes		0.02	0.02*	
(servings/day)		(-0.02, 0.06)	(0.00, 0.03)	
Canned fish				0.01

(servings/day)	(-0.03, 0.04)
Fried fish	0.01
(servings/day)	(-0.03, 0.05)
Seafood	0.08***
(servings/day)	(0.04, 0.12)

*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001 FFQ = Food frequency questionnaire; TMAO = trimethylamine *N*-oxide.

Table 3.3: Results from ordinary least squares regression examining the association of dietary and non-dietary factors with serum non-esterified fatty acid (NEFA) in FAMILY cohort.

	even-chain SFA	**odd-chain SFA**		**ω-3 PUFA**		
	14:0	**15:0**	**17:0**	**EPA**	**DHA**	**EPA + DHA**
Variable	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**
Age (years)	4.24×10^{-3}	-3.54×10^{-4}	-0.01	-0.01	-4.77×10^{-3}	-0.01
	(-0.00, 0.01)	(-0.01, 0.01)	(-0.01, 0.00)	(-0.03, 0.00)	(-0.01, 0.00)	(-0.02, 0.00)
Gestational age (weeks)	-0.01	-0.02***	-0.01**	-0.01	-0.01	-0.01
	(-0.02, 0.00)	(-0.03, -0.01)	(-0.02, -0.00)	(-0.03, 0.01)	(-0.02, 0.00)	(-0.02, 0.00)
Parity	-0.01	2.105 × 10-4	2.21×10^{-3}	-2.07×10^{-5}	-0.03	-0.02
	(-0.04, 0.03)	(-0.03, 0.03)	(-0.03, 0.03)	(-0.06, 0.06)	(-0.08, 0.01)	(-0.07, 0.03)
Gestational diabetes (GDM)	0.02	-0.01	-0.06	-0.06	-0.07	-0.07
	(-0.07, 0.10)	(-0.09, 0.07)	(-0.14, 0.03)	(-0.22, 0.10)	(-0.18, 0.04)	(-0.19, 0.05)
Pre-pregnancy BMI (kg/m^2)	-0.01*	-0.01**	-0.01*	-2.86×10^{-3}	-0.01*	-0.01
	(-0.01, -0.00)	(-0.01, -0.00)	(-0.01, -0.00)	(-0.01, 0.01)	(-0.02, -0.00)	(-0.01, 0.00)
Smoking history	-0.02	-0.04	-0.05	-0.01	-0.05	-0.04

(ever vs. never smoked)	(-0.08, 0.05)	(-0.10, 0.03)	(-0.12, 0.01)	(-0.13, 0.10)	(-0.14, 0.03)	(-0.12, 0.05)
Physical activity	-0.01	-0.01	-0.10**	-0.03	-0.03	-0.03
(low vs. high)	(-0.09, 0.08)	(-0.09, 0.06)	(-0.17, -0.02)	(-0.18, 0.11)	(-0.12, 0.06)	(-0.13, 0.08)
Social disadvantage index	-0.02	-1.79×10^{-3}	0.02	0.04	0.04	0.04
	(-0.04, 0.01)	(-0.03, 0.03)	(-0.01, 0.05)	(-0.03, 0.10)	(-0.00, 0.08)	(-0.01, 0.08)
Fiber intake	-1.12×10^{-3}	2.84×10^{-3}	1.45×10^{-3}	4.51×10^{-3}	2.01×10^{-3}	3.48×10^{-3}
(g/day)	(-0.01, 0.01)	(-0.00, 0.01)	(-0.00, 0.01)	(-0.00, 0.01)	(-0.00, 0.01)	(-0.00, 0.01)
Energy intake (kcal)	-1.05×10^{-5}	-4.76×10^{-5}	-2.42×10^{-5}	-8.23×10^{-5}	-4.77×10^{-5}	-6.21×10^{-5}
	(-0.00, -0.00)	(-0.00, 0.00)	(-0.00, 0.00)	(-0.00, 0.00)	(-0.00, 0.00)	(-0.00, 0.00)
FFQ before blood draw vs. FFQ at the	-0.03	0.06	0.01	-3.01×10^{-3}	0.05	0.02
same time as blood draw	(-0.10, 0.04)	(-0.01, 0.13)	(-0.06, 0.08)	(-0.14, 0.14)	(-0.04, 0.15)	(-0.07, 0.12)
FFQ after blood draw vs. FFQ at the	-0.05	0.02	0.04	0.06	0.24*	0.16
same time as blood draw	(-0.26, 0.16)	(-0.09, 0.13)	(-0.10, 0.19)	(-0.26, 0.38)	(0.02, 0.46)	(-0.07, 0.40)
Full-fat dairy	0.02	0.06***	0.04**			
(servings/day)	(-0.02, 0.06)	(0.03, 0.10)	(0.01, 0.07)			

Fish/Fish oil	0.05	0.11***	0.08***
(servings/day)	(-0.00, 0.11)	(0.07, 0.14)	(0.04, 0.12)

*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001

Table S3.1: Results from ordinary least squares regression examining the association of dietary and non-dietary factors with food-based metabolites

	Hippuric acid	Carnitine
Variable	**b (95% CI)**	**b (95% CI)**
Age (years)	0.01 (-0.01, 0.03)	0.00 (-0.00, 0.00)
Gestational age (weeks)	0.01 (-0.01, 0.03)	-0.01** (-0.02, -0.01)
Parity	0.03 (-0.05, 0.11)	0.00 (-0.01, 0.02)
Gestational diabetes (GDM)	0.06 (-0.10, 0.23)	0.02 (-0.02, 0.05)
Pre-pregnancy BMI (kg/m^2)	-0.01 (-0.03, 0.00)	0.00 (-0.00, 0.00)
Smoking history (ever vs. never smoked)	-0.13 (-0.29, 0.03)	0.06* (0.02, 0.10)
Physical activity (low vs. high)	0.02 (-0.11, 0.15)	-0.01 (-0.04, 0.02)
Social disadvantage index	-0.02 (-0.08, 0.04)	0.00 (-0.01, 0.01)
Fiber intake (g/day)	0.01 (-0.00, 0.02)	0.00 (-0.00, 0.00)
Energy intake (kcal)	0.00* (-0.00, -0.00)	0.00 (-0.00, 0.00)
FFQ before blood draw vs. FFQ at the same time as blood draw	0.10 (-0.07, 0.26)	0.01 (-0.03, 0.05)
FFQ after blood draw vs. FFQ at the same time as blood draw	0.08 (-0.18, 0.34)	0.06 (-0.07, 0.19)
Ethnicity (White European vs. South Asian)	0.04 (-0.18, 0.26)	-0.02 (-0.07, 0.04)
Fruits and vegetables (servings/day)	0.22* (0.08, 0.36)	
Tea (servings/day)	0.01 (-0.01, 0.04)	
Coffee (servings/day)	0.02 (-0.00, 0.04)	
Red meat (servings/day)		0.00 (-0.00, 0.01)

*$p \leq 0.01$, **$p \leq 0.001$

Table S3.2: Results of model fitting analyses examining the association of dietary and non-dietary factors with food metabolites

	−2 Log L	BIC	AIC	S_b^2	S_w^2
Proline Betaine					
Model 1	–	–	–	0.14	0.13
Model 2	2100.0	2101.4	2104.0	0.02	0.11
Model 3	1779.0	1780.4	1783.0	0.20	0.12
Hippuric acid					
Model 1	–	–	–	0.00	0.03
Model 2	1394.7	1395.3	1396.7	0.00	0.03
Model 3	1182.1	1182.8	1184.1	0.00	0.03
3-methylhistidine					
Model 1	–	–	–	0.07	0.01
Model 2	515.3	518.8	525.3	0.02	0.01
Model 3	496.0	497.4	500.0	0.02	0.01
Carnitine					
Model 1	–	–	–	0.00	0.00
Model 2	-387.0	-385.6	-383.0	0.00	0.00
Model 3	-298.6	-297.9	-296.6	0.00	0.00
Tryptophan betaine					
Model 1	–	–	–	0.03	0.00
Model 2	-588.1	-586.7	-584.1	0.02	0.00
Model 3	-489.4	-488.0	-485.4	0.02	0.00
TMAO					
Model 1	–	–	–	0.07	0.04
Model 2	1421.5	1422.9	1425.5	0.02	0.03
Model 3	1211.1	1211.8	1213.1	0.00	0.04

AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion, -2 Log L = -2 log likelihood,
S_b^2 = Sum of square between, S_w^2 = Sum of square within
Model 1: Unconditional (intercept only) model
Model 2: Random Intercept with Fixed Level-1 Factors (dietary factors)
Model 3: Random Intercept with Fixed Level-1 Factors (dietary and non-dietary factors)

Table S3.3: Results from ordinary least squares regression examining the association of dietary and non-dietary factors with food-based metabolites in FAMILY cohort

	Proline betaine	Hippuric acid	3-Methyl histidine	Carnitine	Tryptophan betaine	TMAO
Variable	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**
Age (years)	0.06** (0.02, 0.10)	0.02 (-0.01, 0.04)	0.00 (-0.01, 0.02)	0.00 (-0.00, 0.01)	0.00 (-0.00, 0.00)	0.02 (-0.01, 0.04)
Gestational age (weeks)	0.02 (-0.02, 0.10)	0.00 (-0.02, 0.02)	0.01 (-0.01, 0.02)	-0.01*** (-0.02, 0.01)	0.00 (-0.00, 0.00)	0.01 (-0.01, 0.04)
Parity	-0.06 (-0.26, 0.15)	-0.02 (-0.11, 0.08)	0.00 (-0.07, 0.07)	-0.01 (-0.03, 0.01)	-0.01 (-0.02, 0.00)	-0.04 (-0.12, 0.05)
Gestational diabetes (GDM)	-0.15 (-0.65, 0.35)	0.01 (-0.23, 0.24)	0.06 (-0.07, 0.19)	0.03 (-0.02, 0.08)	-0.01 (-0.04, 0.02)	-0.07 (-0.29, 0.15)
Pre-pregnancy BMI (kg/m^2)	-0.03* (-0.06, -0.00)	-0.01 (-0.03, 0.01)	-0.00 (-0.01, 0.00)	-0.00 (-0.01, 0.00)	0.00 (-0.00, 0.00)	-0.01 (-0.02, 0.01)
Smoking history (ever vs. never smoked)	-0.53** (-0.91, 0.14)	-0.10 (-0.26, 0.07)	0.03 (-0.07, 0.13)	0.06** (0.02, 0.10)	0.00 (-0.02, 0.00)	0.01 (-0.17, 0.19)
Physical activity (low vs. high)	0.06 (-0.31, 0.43)	-0.02 (-0.18, 0.14)	-0.05 (-0.15, 0.06)	-0.02 (-0.06, 0.03)	0.00 (-0.03, 0.02)	-0.02 (-0.21, 0.18)
Social disadvantage index	-0.03 (-0.22, 0.17)	-0.01 (-0.10, 0.07)	0.00 (-0.04, 0.05)	-0.00 (-0.02, 0.01)	0.00 (-0.01, 0.01)	0.04 (-0.03, 0.11)
Fiber intake (g/day)	-0.00 (-0.03, 0.02)	0.01 (-0.00, 0.02)	-0.00 (-0.01, 0.01)	0.00 (-0.00, 0.00)	0.00 (-0.00, 0.00)	0.00 (-0.01, 0.02)
Energy intake (kcal)	-0.00 (-0.00, 0.00)	-0.00* (-0.00, -0.00)	-0.00 (0.00, 0.00)	-0.00 (-0.00, 0.00)	0.00 (-0.00, 0.00)	-0.00 (-0.00, 0.00)
FFQ before blood draw vs. FFQ at the same time as blood draw	0.06 (-0.29, 0.40)	0.09 (-0.09, 0.26)	-0.06 (-0.16, 0.05)	0.02 (-0.02, 0.06)	-0.01 (-0.03, 0.02)	0.10 (-0.08, 0.28)

FFQ after blood draw vs. FFQ at the same time as blood draw	0.08 (-0.70, 0.86)	-0.02 (-0.46, 0.41)	0.10 (-0.30, 0.50)	0.05 (-0.10, 0.21)	-0.02 (-0.11, 0.06)	-0.14 (-0.72, 0.45)
Citrus food (servings/day)	0.40*** (0.24, 0.57)					
Fruits and vegetables (servings/day)		0.30*** (0.16, 0.45)				
Tea (servings/day)		0.00 (-0.02, 0.03)				
Coffee (servings/day)		0.01 (-0.01, 0.04)				
Chicken (servings/day)			0.03 (-0.00, 0.06)			
Red meat (servings/day)			0.01 (-0.03, 0.05)	0.02* (0.00, 0.04)		0.04 (-0.02, 0.10)
Eggs (servings/day)			0.03 (-0.00, 0.06)			0.03 (-0.04, 0.10)
Nuts and legumes (servings/day)			0.06* (0.00, 0.12)		0.02* (0.00, 0.04)	
Canned fish (servings/day)						0.02 (-0.01, 0.06)
Fried fish (servings/day)						0.02 (-0.02, 0.07)
Seafood (servings/day)						0.07** (0.02, 0.11)

*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001

Table S3.4: Results from ordinary least squares regression examining the association of dietary and non-dietary factors with food-based metabolites in START cohort

	Proline betaine	Hippuric acid	3-Methyl histidine	Carnitine	Tryptophan betaine	TMAO
Variable	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**
Age (years)	-0.01 (-0.07, 0.04)	-0.01 (-0.04, 0.02)	0.00 (-0.01, 0.01)	-0.00 (-0.01, 0.00)	0.00 (-0.01, 0.01)	0.03 (-0.01, 0.06)
Gestational age (weeks)	0.02 (-0.08, 0.13)	0.05 (-0.01, 0.11)	-0.00 (-0.03, 0.02)	-0.02*** (-0.03, -0.01)	-0.01 (-0.02, 0.00)	0.02 (-0.03, 0.07)
Parity	-0.18 (-0.45, 0.10)	0.11 (-0.02, 0.24)	-0.03 (-0.09, 0.04)	0.03 (-0.00, 0.06)	-0.01 (-0.04, 0.03)	0.07 (-0.06, 0.21)
Gestational diabetes (GDM)	0.18 (-0.18, 0.55)	0.09 (-0.14, 0.32)	-0.02 (-0.11, 0.07)	0.01 (-0.03, 0.06)	0.04 (-0.00, 0.09)	0.05 (-0.18, 0.27)
Pre-pregnancy BMI (kg/m^2)	0.01 (-0.04, 0.05)	-0.02 (-0.05, 0.00)	-0.01 (-0.02, 0.00)	0.00 (-0.00, 0.01)	0.00 (-0.00, 0.01)	-0.02 (-0.04, 0.01)
Physical activity (low vs. high)	-0.57* (-0.12, 0.03)	-0.05 (-0.27, 0.17)	-0.01 (-0.11, 0.09)	-0.00 (-0.05, 0.05)	0.03 (-0.02, 0.08)	-0.04 (-0.36, 0.28)
Social disadvantage index	-0.05 (-0.18, 0.09)	-0.03 (-0.10, 0.05)	-0.00 (-0.04, 0.03)	0.01 (-0.01, 0.02)	-0.00 (-0.02, 0.02)	-0.07 (-0.15, 0.01)
Fiber intake (g/day)	0.01 (-0.02, 0.04)	0.01 (-0.01, 0.03)	-0.00 (-0.01, 0.00)	-0.00 (-0.01, 0.00)	0.00** (0.00, 0.01)	-0.00 (-0.02, 0.01)
Energy intake (kcal)	-0.00 (-0.00, 0.00)	0.00* (0.00, 0.00)	0.00 (-0.00, 0.00)	0.00** (0.00, 0.00)	-0.00* (-0.00, -0.00)	0.00 (-0.00, 0.00)
FFQ before blood draw vs. FFQ at the same time as blood draw	-0.76 (-1.58, 0.07)	0.24 (-0.25, 0.72)	0.00 (-0.18, 0.19)	0.05 (-0.14, 0.05)	-0.01 (-0.08, 0.05)	0.02 (-0.37, 0.41)
FFQ after blood draw vs. FFQ at the same time as blood draw	0.94** (0.29, 1.59)	0.14 (-0.23, 0.50)	-0.07 (-0.28, 0.13)	0.06 (-0.14, 0.27)	-0.06 (-0.24, 0.12)	-0.19 (-0.64, 0.26)

Citrus food (servings/day)	0.18*** (0.17, 0.28)					
Fruits and vegetables (servings/day)		0.11 (-0.09, 0.30)				
Tea (servings/day)		0.03 (-0.02, 0.08)				
Coffee (servings/day)		0.02 (-0.04, 0.07)				
Chicken (servings/day)			-0.00 (-0.03, 0.03)			
Red meat (servings/day)			0.06** (0.02, 0.10)	-0.00 (-0.01, 0.01)		0.00 (-0.06, 0.06)
Eggs (servings/day)			-0.01 (-0.02, 0.01)			0.00 (-0.03, 0.07)
Nuts and legumes (servings/day)			-0.05 (-0.11, 0.01)		0.02 (-0.01, 0.05)	
Canned fish (servings/day)						-0.08 (-0.24, 0.07)
Fried fish (servings/day)						-0.00 (-0.08, 0.07)
Seafood (servings/day)						0.11* (0.01, 0.21)

*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001

Table S3.5: Results from ordinary least squares regression examining the association of dietary and non-dietary factors with serum non-esterified fatty acid (NEFA) in FAMILY cohort

	ω-3 PUFA		
	EPA	**DHA**	**EPA + DHA**
Variable	**b (95% CI)**	**b (95% CI)**	**b (95% CI)**
Age (years)	-0.01 (-0.02, 0.00)	-2.92×10^{-3} (-0.01, 0.01)	-0.01 (-0.02, 0.00)
Gestational age (weeks)	-0.01 (-0.03, 0.00)	-0.01* (-0.02, -0.00)	-0.01* (-0.02, -0.00)
Parity	-0.01 (-0.07, 0.06)	-0.05* (-0.09, -0.00)	-0.03 (-0.08, 0.02)
Gestational diabetes (GDM)	-0.06 (-0.22, 0.10)	-0.06 (-0.17, 0.05)	-0.07 (-0.18, 0.05)
Pre-pregnancy BMI (kg/m^2)	-3.34×10^{-3} (-0.01, 0.01)	-0.01** (-0.02, -0.00)	-0.01 (-0.01, 0.00)
Smoking history (ever vs. never smoked)	-0.01 (-0.13, 0.11)	-0.04 (-0.13, 0.04)	-0.03 (-0.12, 0.05)
Physical activity (low vs. high)	-0.03 (-0.18, 0.12)	-0.02 (-0.12, 0.08)	-0.02 (-0.13, 0.09)
Social disadvantage index	0.03 (-0.04, 0.10)	0.03 (-0.01, 0.07)	0.03 (-0.02, 0.08)
Fiber intake (g/day)	0.01 (-0.00, 0.01)	3.71×10^{-3} (-0.00, 0.01)	4.88×10^{-3} (-0.00, 0.01)
Energy intake (kcal)	-7.71×10^{-5} (-0.00, 0.00)	-3.76×10^{-5} (-0.00, 0.00)	-5.39×10^{-5} (-0.00, 0.00)
FFQ before blood draw vs. FFQ at the same time as blood draw	-0.01 (-0.15, 0.13)	0.04 (-0.05, 0.13)	0.02 (-0.08, 0.11)
FFQ after blood draw vs. FFQ at the same time as blood draw	0.06 (-0.26, 0.39)	0.25 (-0.04, 0.54)	0.17 (-0.11, 0.44)
Fish (servings/day)	0.02 (-0.01, 0.05)	0.04** (0.01, 0.06)	0.03* (0.01, 0.05)

*$p \leq 0.05$, **$p \leq 0.01$

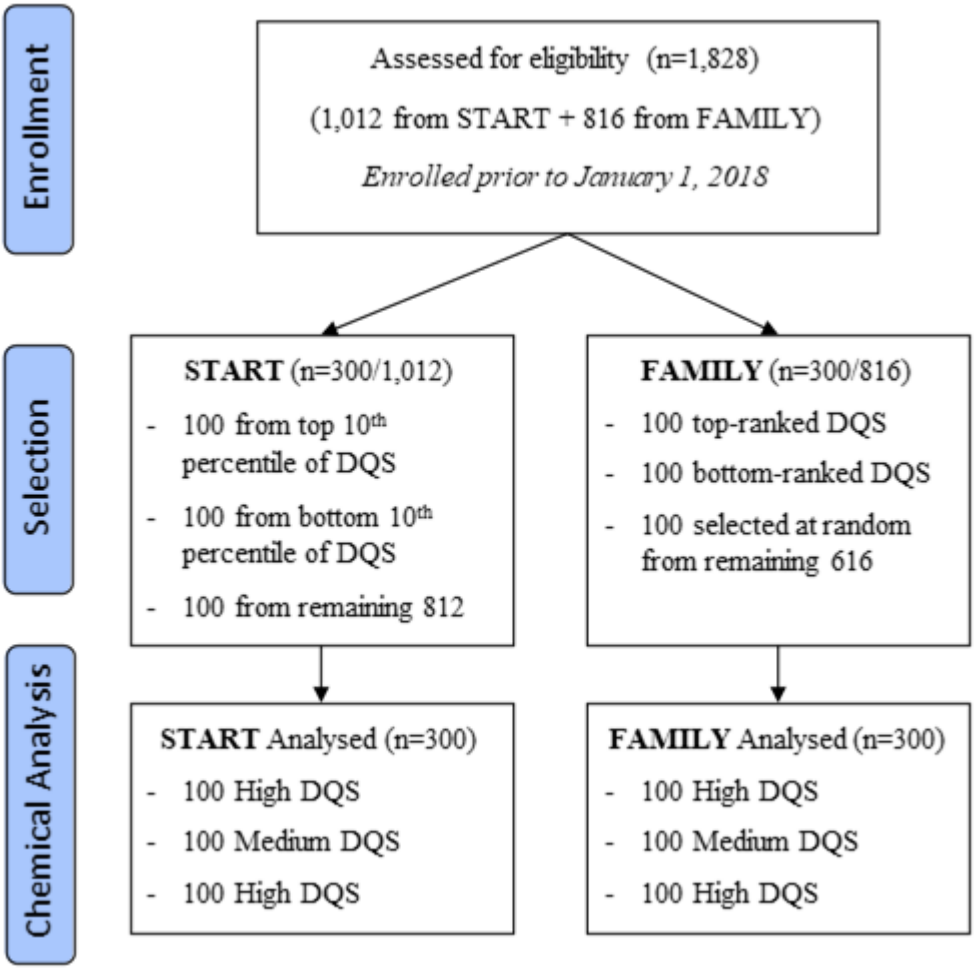


Figure S3.1: Consort flow diagram outlining selection criteria used in a cross-sectional study involving participants from the FAMILY and START birth cohorts

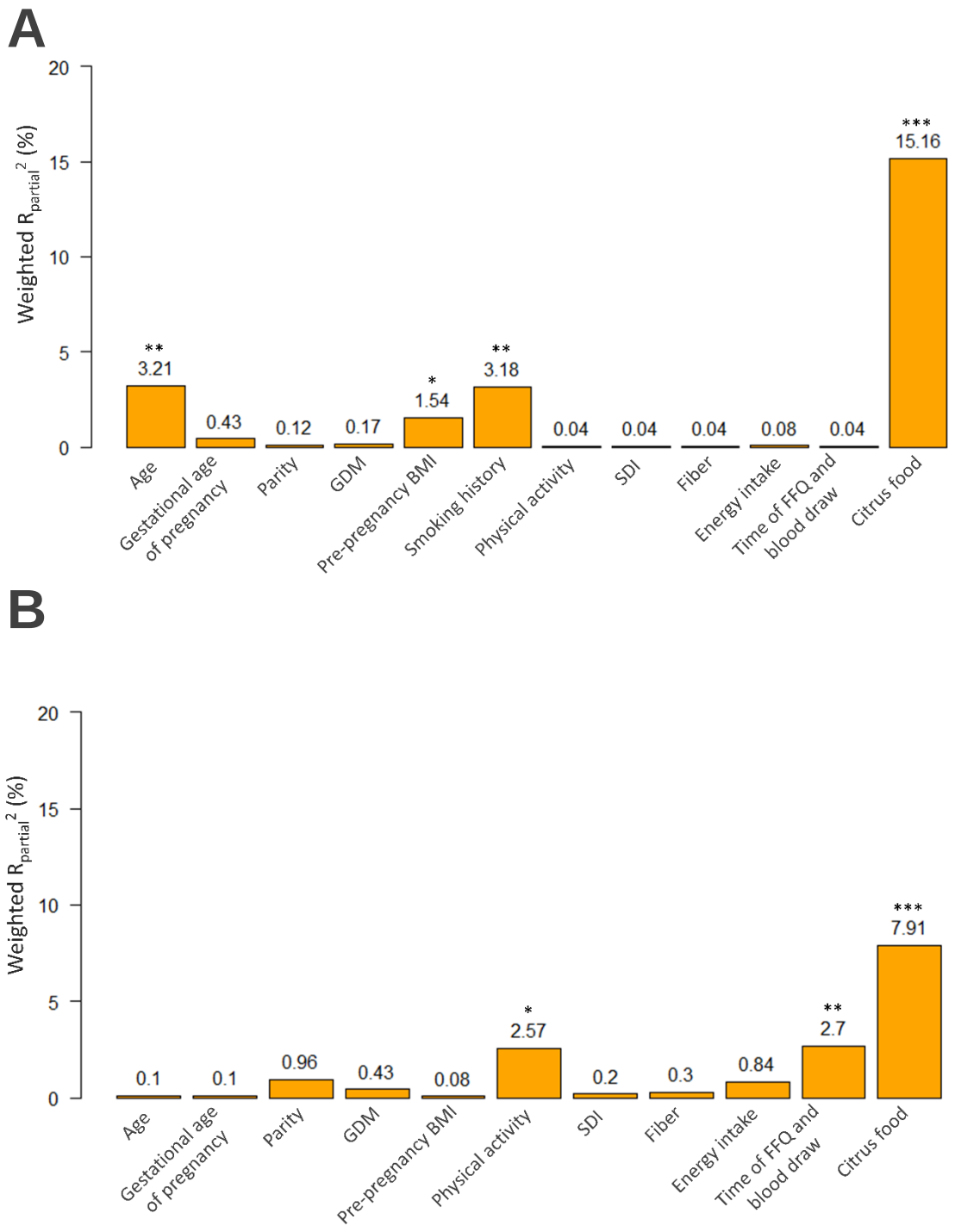


Figure S3.2: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in Proline betaine in (A) FAMILY and (B) START cohort

Statistical significance was based on hierarchical linear models. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

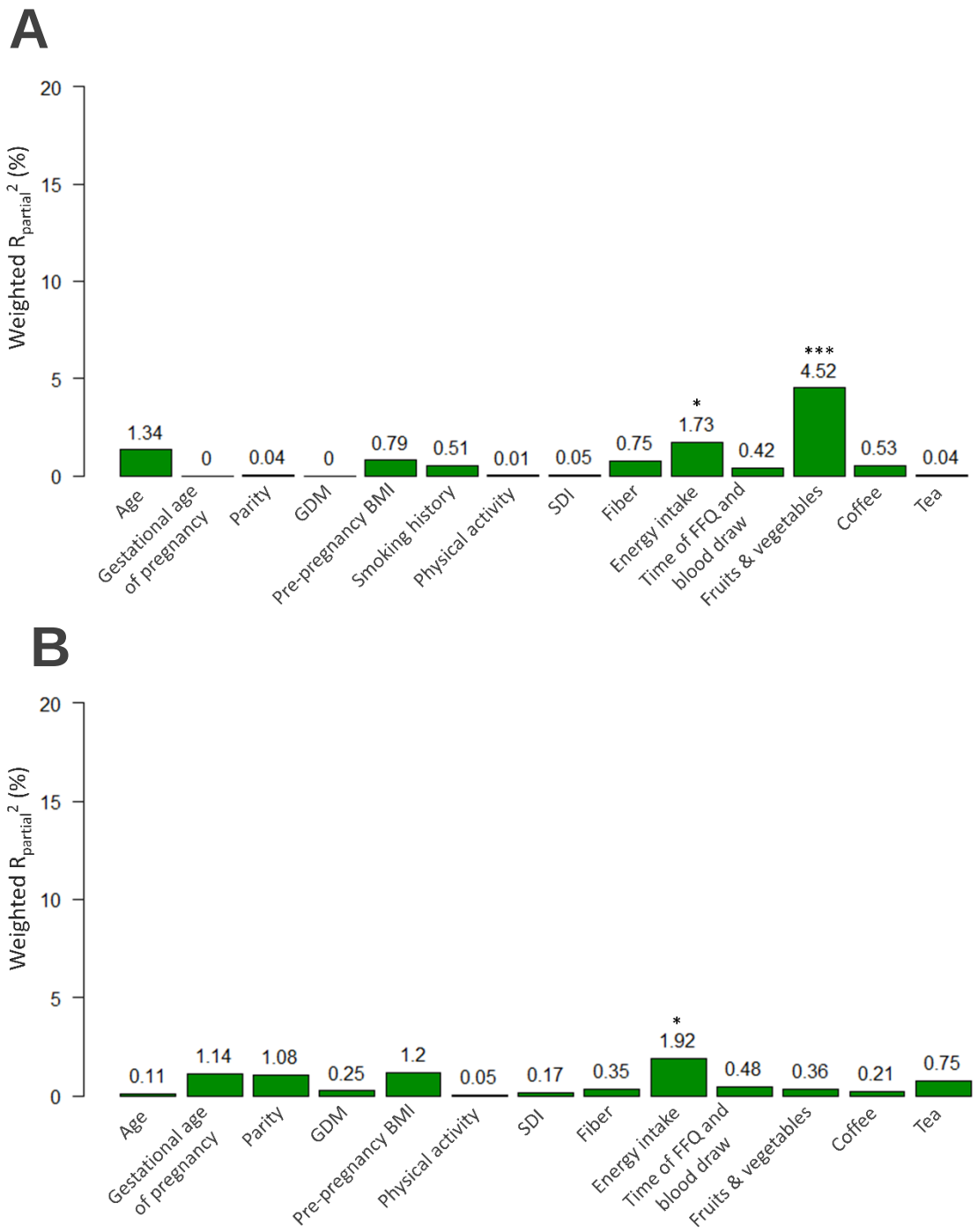


Figure S3.3: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in Hippuric acid in (A) FAMILY and (B) START cohort

Statistical significance was based on hierarchical linear models. * $p \leq 0.05$, ***$p \leq 0.001$

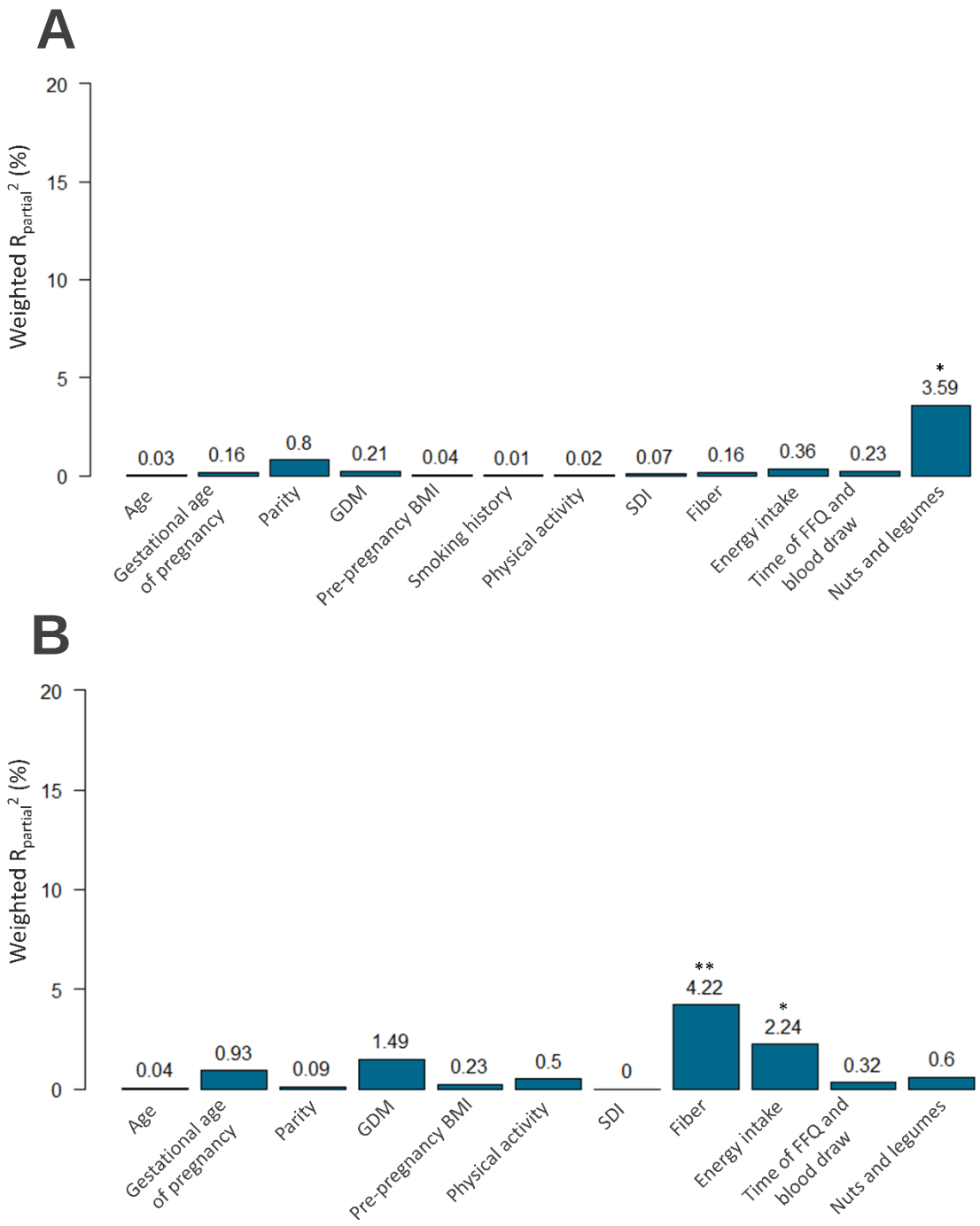


Figure S3.4: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in Tryptophan betaine in (A) FAMILY and (B) START cohort

Statistical significance was based on hierarchical linear models. * $p \leq 0.05$, **$p \leq 0.01$

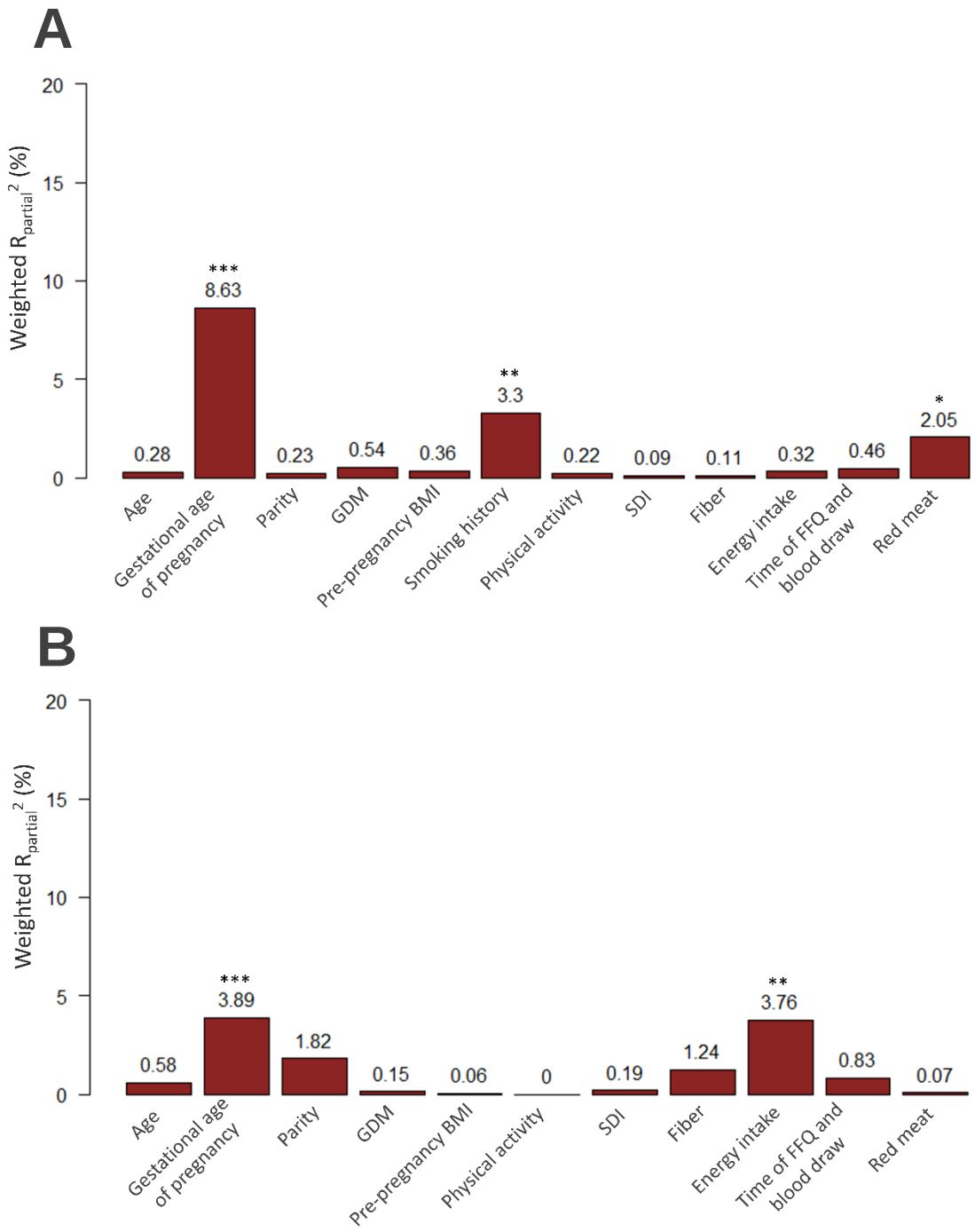


Figure S3.5: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in Carnitine in (A) FAMILY and (B) START cohort

Statistical significance was based on hierarchical linear models. * $p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$

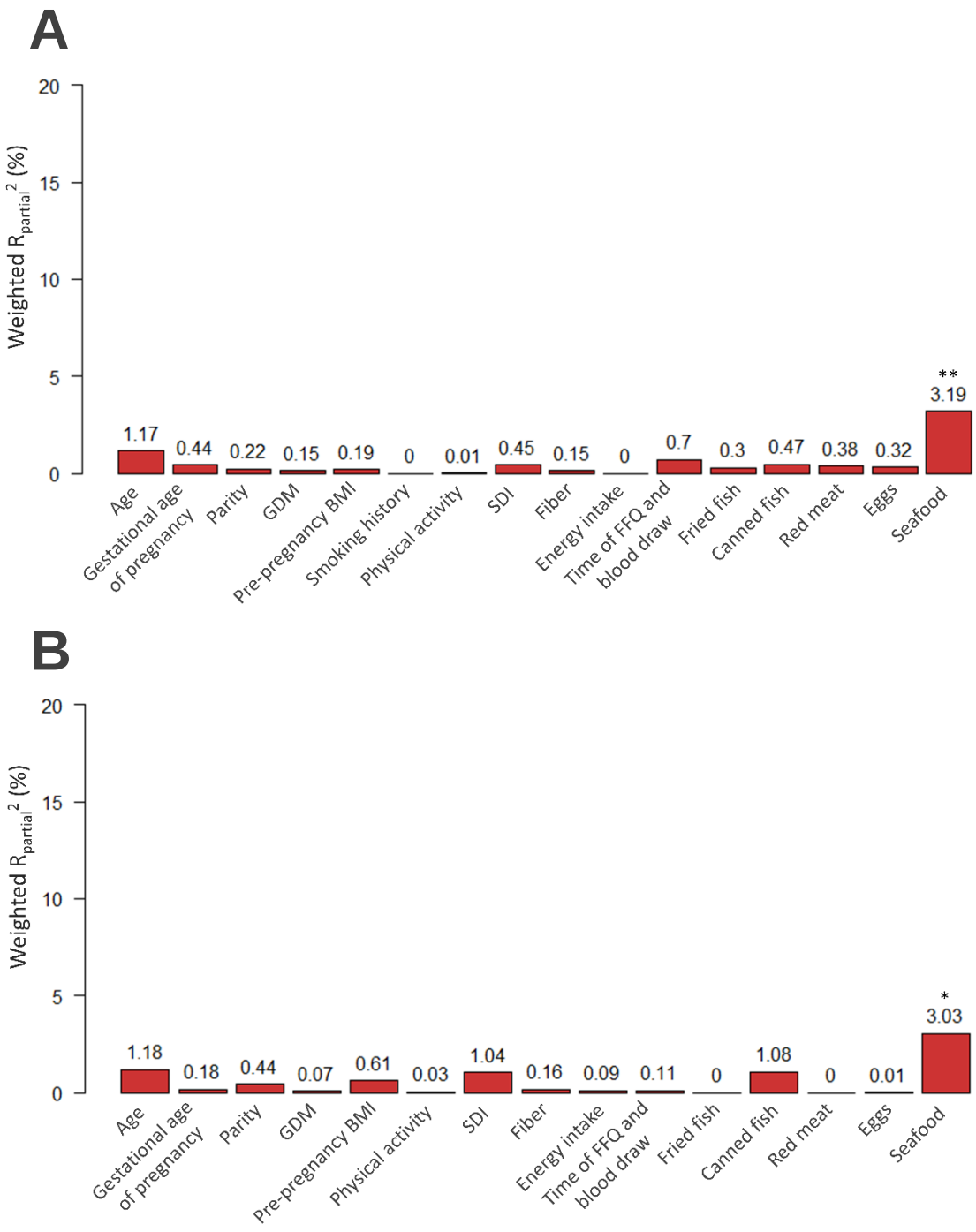


Figure S3.6: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in trimethylamine *N*-oxide (TMAO) in (A) FAMILY and (B) START cohort

Statistical significance was based on hierarchical linear models. * $p \leq 0.05$, ** $p \leq 0.01$

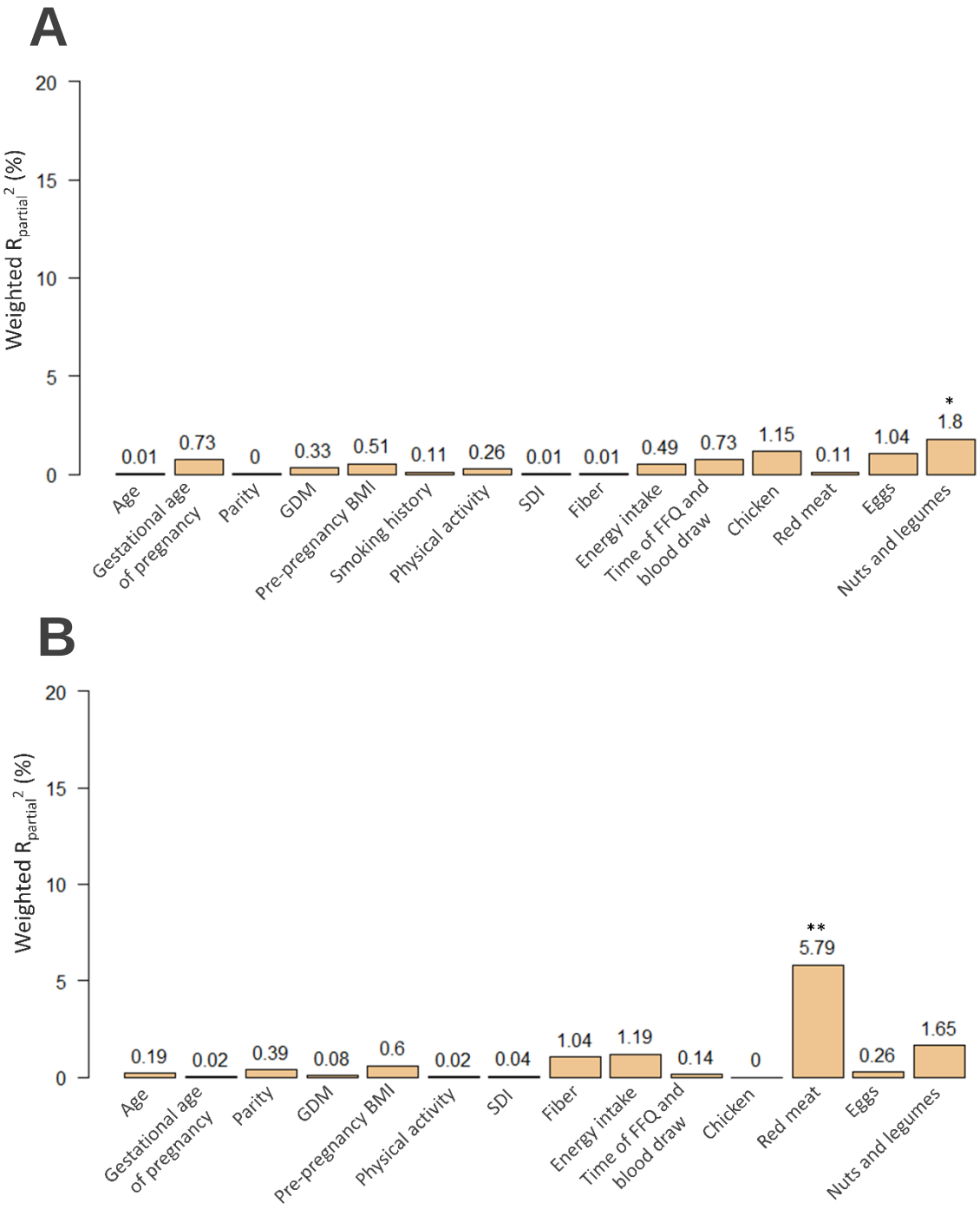


Figure S3.7: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in 3-methylhistidine in (A) FAMILY and (B) START cohort

Statistical significance was based on hierarchical linear models. * $p \leq 0.05$, ** $p \leq 0.01$

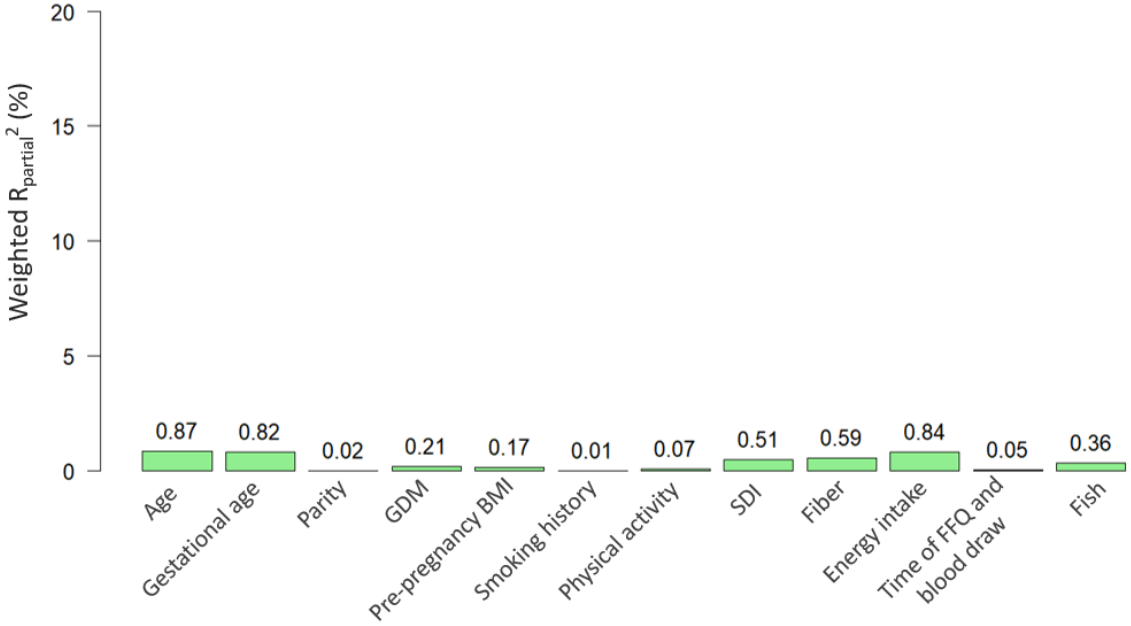


Figure S3.8: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in Eicosapentaenoic acid (EPA, 20:5n-3) in FAMILY cohort

Statistical significance was based on ordinary least squares regression.

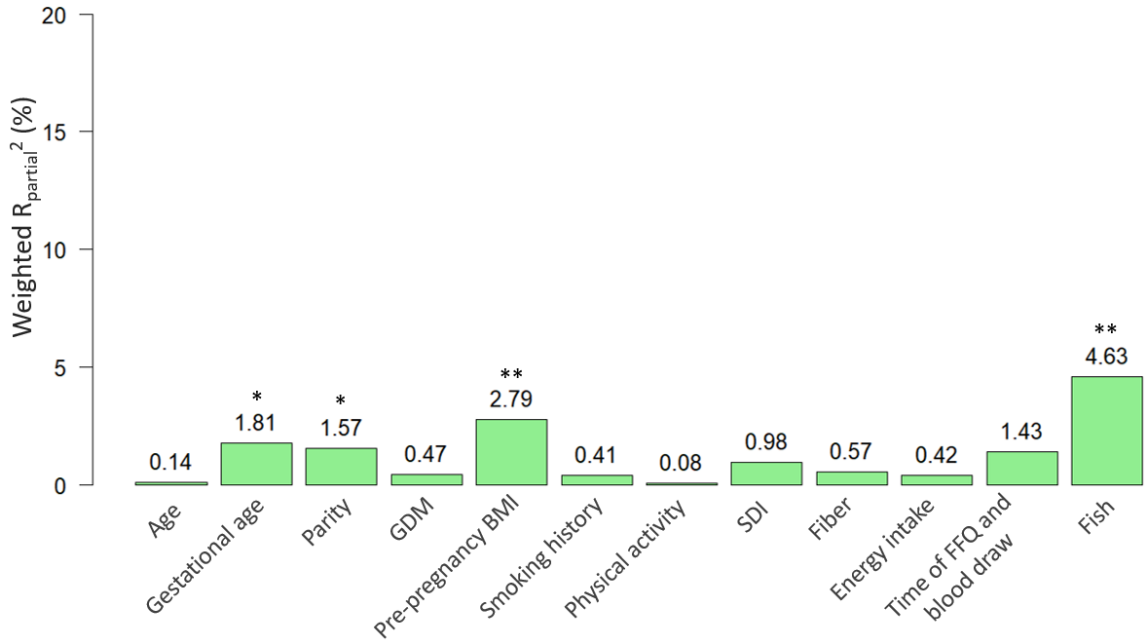


Figure S3.9: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in Docosahexaenoic acid (DHA; 22:6n-3) in FAMILY cohort

Statistical significance was based on ordinary least squares regression. * $p \leq 0.05$, ** $p \leq 0.01$

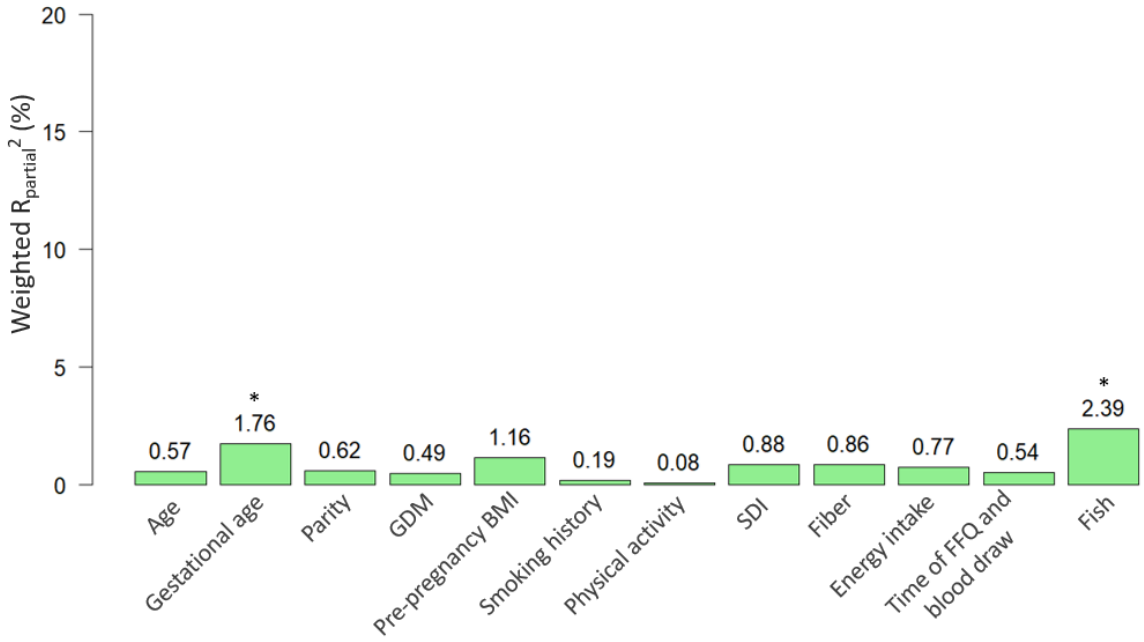


Figure S3.10: Weighted $R_{partial}^2$ for each factor showing the percentage of explained variability in EPA + DHA in FAMILY cohort

Statistical significance was based on ordinary least squares regression. * $p \leq 0.05$

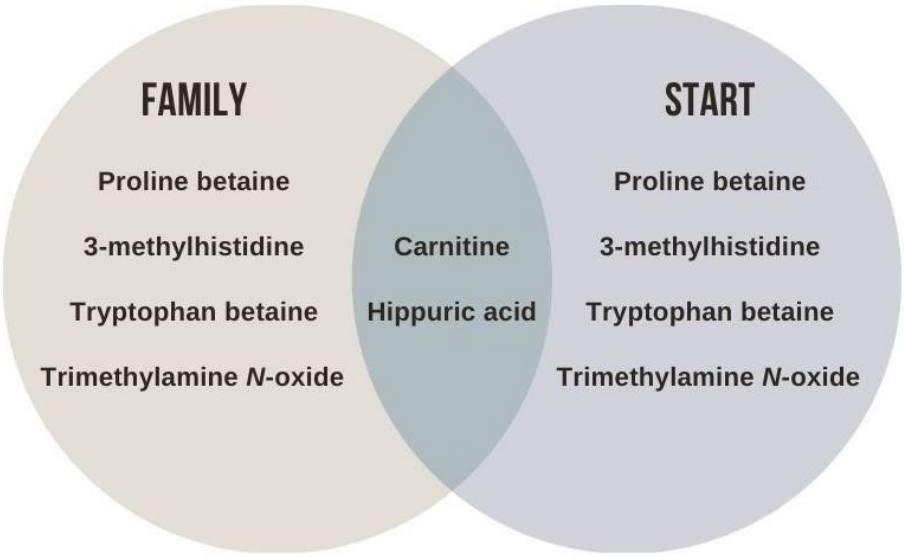


Figure S3.11: Venn diagram showing overlap of serum metabolites based on the cluster effect by ethnicity/cohort

CHAPTER 4 – Integrative multiomics analysis of infant gut microbiome and serum metabolome reveals key molecular biomarkers of childhood obesity

Talha Rafiq, Jennifer C. Stearns, Meera Shanmuganathan, Sandi M. Azab, Sonia S. Anand, Lehana Thabane, Natalie C. Williams, Katherine M. Morrison, Koon K. Teo, Philip Britz-McKibbin and Russell J. de Souza

The authors' responsibilities were as follows— TR led the drafting of the manuscript; TR, RJdS and PB-M: designed the research; TR, RJdS and LT led the statistical analysis; PB-M, SMA and MS led the analytical chemistry interpretations; TR analyzed the data and performed statistical analyses; PB-M, SMA, MS and study participants provided essential reagents or provided essential materials; JCS, MS, SMA, SSA, LT, NCW, KMM, KKT, PB-M, and RJdS provided critical revision of the manuscript for important intellectual content.

4.1 Abstract

Background: Emerging evidence supports the complex interplay of gut microbiome and host metabolism as important regulators of obesity. The metabolome profile and microbiome activity from dietary exposures may also contribute to greater obesity risk in children.

Objective: We aimed to identify features that discriminated children with overweight/obesity and normal weight by integrating 16S rRNA gene amplicon sequence variant (ASV) and serum metabolome data.

Methods: This prospective analysis included 50 South Asian children from the SouTh Asian biRth cohorT (START). Serum metabolites were measured by multisegment injection-capillary electrophoresis-mass spectrometry and the relative abundance of bacterial species was evaluated with 16S rRNA sequencing (V3 region) from stool samples at 1 year of age. Cumulative body mass index (BMI_{AUC}) and skinfold thickness (SSF_{AUC}) scores were calculated using 4 measurements obtained from birth to 3-years as the total area under the growth curve (AUC). BMI_{AUC} and/or SSF_{AUC} >85^{th} percentile was used to classify overweight/obesity children. Data Integration Analysis for Biomarker discovery using Latent cOmponent (DIABLO) was used to identify discriminant features associated with childhood adiposity. Logistic regression was used to examine the association between the identified features and anthropometric measures in young children.

Results: Several serum metabolites including glutamic acid, acetylcarnitine, carnitine, and threonine were positively, whereas γ-aminobutyric acid (GABA), symmetric

dimethylarginine (SDMA), and asymmetric dimethylarginine (ADMA) were negatively associated with childhood adiposity. The bacterium *Akkermansia* was positively correlated with GABA, whereas *Pseudobutyrivibrio* and *Lactobacillus* were inversely correlated with GABA and SDMA.

Conclusion: Metabolites related to protein and fat metabolism were associated with specific gut bacterial species. By integrating infant metabolome and microbiome, this study provides insights into metabolomic features and microbial composition associated with obesity trajectories in childhood.

4.2 Introduction

Obesity is a multifactorial process characterized by excess adipose tissue accompanied by chronic low-grade inflammation resulting from complex interactions between genetic, behavioural, social, and environmental factors. However, the pathophysiological mechanisms leading to excess adiposity early in life remains poorly understood. Recent advances in –omics technologies and approaches have allowed for a comprehensive characterization of metabolic networks to decipher underlying biological responses that lead to phenotypic obesity (1) due to dynamic changes at the genetic, epigenetic, protein, and metabolites levels. For instance, emerging evidence has shown that the distribution of bacterial communities found in the intestine (i.e., the microbiome) of children who are overweight or obese differs from those who are of normal weight (2, 3). Environmental factors and changes in the gut microbiota may therefore jointly underlie the phenotypic expression of obesity and may also signal changes in other –omics markers (4). Adiposity is also characterized by lower gut microbial diversity as compared to normal weight controls (5). Although host-derived factors are genetically hardwired, the microbiome can be regulated by environmental factors such as habitual diet. However, it is unclear whether imbalances in the microbiome composition (i.e., dysbiosis) that cause or increase risk of disease and accompanying changes in the metabolome are a cause or consequence of childhood adiposity. General strategy to provide evidence for causality is to observe change in adiposity over time in the correct temporal sequence.

The colonization of the gut microbiota starts from birth and alterations in maturation during infancy is a potential contributor to obesity and metabolic traits (3, 6).

Alterations in intestinal microbiota composition, specifically a higher *Firmicutes*-to-*Bacteroidetes* ratio and lower microbiota diversity are shown to be associated with obesity in children as young as 7 years of age (7), and can lead to disruption in energy acquisition and regulation (8). The gut microbiome associated with diet-induced obesity may be more efficient at extracting energy from indigestible polysaccharides, contributing to greater caloric uptake (9). Moreover, elevated *Firmicutes*-to-*Bacteroidetes* ratio can lead to more efficient hydrolysis and fermentation of the indigestible dietary polysaccharides to generate short-chain fatty acids (SCFAs), which can increase the host's ability to extract energy from the food components entering the GI tract and activate the lipogenic pathways (10). However, some studies report higher *Bacteroides* abundance and *Bacteroides*/*Prevotella* to be associated with obesity in children (11-13). These inconsistent results have been attributed to bias caused by a lack of standardized methods for stool collection and storage during gut microbiome sampling (14), and non-technical factors that are rarely or insufficiently controlled, such as inter-individual genetic, environmental, and dietary variance. In addition to SCFAs, fecal metabolome studies have showed that obesity incidence is associated with higher levels of branched-chain amino acids (BCAAs), including leucine, isoleucine, and valine, and aromatic amino acids (AAs), including phenylalanine, tryptophan, and tyrosine (15, 16). A recent systematic review of 41 studies using blood and urine samples reported a consistent metabolic profile of children with obesity comprising higher levels of amino acids including BCAAs and AAAs, carnitines, lipids, and steroids (17). These findings support the role of that certain metabolites can be endogenously produced or derived in

response to diet exposure and gut microbiota activity, which may influence host metabolic processes that increase obesity due to altered energy balance and/or inflammation. It is therefore of interest to use integrative multi-omics analyses to characterize molecular changes of childhood obesity necessary to gain new insights into its etiology.

Obesity and its complications are disproportionately more prevalent in non-white populations (18). Children of South Asian ancestry have an increased cardiometabolic risk at lower body mass index (BMI) than other ethnic groups which has been attributed to lower lean mass and higher abdominal fat mass at the same BMI (18-20). However, studies of metabolome or microbiome and childhood obesity have been performed primarily in white Europeans, and thus need to be examined in other ethnic populations, such as South Asians (17). Another research gap is the integration of metabolomics with microbiome, which is of great interest in child obesity research, but little research has been done and it remains largely understudied in non-white populations. To address these gaps, we employed a multi-omics approach of serum metabolomics and amplicon sequence variants (ASVs) of 16S rRNA genes to identify integrated molecular features that characterize risk of obesity in young children.

4.3 Methods

4.3.1 Data source and participants

The South Asian Birth Cohort (START) is a prospective birth cohort designed to study the influence of diverse environmental exposures and genetics on early life adiposity, growth trajectory and cardiometabolic health of South Asians living in Canada. In brief,

START enrolled 1,012 South Asian mother–child pairs from the Peel Region of Ontario, Canada. Participants were recruited through physician referrals between 2011 and 2015, and followed up at 1-, 2-, and 3-years. Ancestral origin of both the woman, her partner, and both offspring's grandparents were required to be South Asian (from the Indian sub-continent: India, Pakistan, Sri Lanka, or Bangladesh). Further details on the START design and methodology have been described in detail elsewhere (21). Of the 182 infants who provided fecal samples for microbiome analyses, our analytic dataset includes 50 infants who provided complete data on microbiome and serum metabolome at 1-year, and anthropometric measures at birth, 1, 2, and 3 years. A primary caregiver of all enrolled participants provided full informed consent. The study was approved by the McMaster Hamilton Integrated Research Ethics Board [START (HiREB #10–640).

4.3.2 Anthropometrics

Child anthropometric measurements were obtained by a trained research assistant following a standard procedure as previously described (21). Specifically, length was measured using the O'Leary Pediatric Length Board at birth, 1-, and 2-years, and a stadiometer was used to measure height after 24 months. Infant birth weight was obtained from the birth delivery reports, and weight at each follow-up visit was measured using an electronic scale. BMI was calculated as weight in kilograms divided by squared height or length in meters. The skinfold thickness of triceps and subscapular sites were measured in triplicate using calipers (Holtain Tanner/Whitehouse, UK) to the nearest 0.2 mm, and summed to create "sum of skinfolds" (SSF) (21).

4.3.3 Area under the curve (AUC) of BMI and SSF

For each child, we calculated area under the curve (AUC) of BMI_{AUC} and SSF_{AUC} from birth to 3-years as a cumulative exposure to summarize the duration and degree of body mass. The BMI_{AUC} and SSF_{AUC} were calculated separately using the following formula:

AUC = Average (BMI or SSF value at age 1, BMI or SSF value at birth) x (1 – 0) +

Average (BMI or SSF value at age 2, BMI or SSF value at age 1) x (2 – 1) +

Average (BMI or SSF value at age 3, BMI or SSF value at age 2) x (3 – 2)

In the analysis, children with BMI_{AUC} and/or SSF_{AUC} at or above internally derived 85th percentile were classified as being overweight/obese and those with BMI_{AUC} and SSF_{AUC} below 85th percentile were classified as normal weight (22). A contingency table comparing the classification of overweight/obesity using the BMI_{AUC} and SSF_{AUC} is presented in Supplementary Table S4.1.

4.3.4 Serum metabolome analyses

A validated multiplexed separation platform based on multisegment injection-capillary-electrophoresis-mass spectrometry (MSI-CE-MS) was used for targeted and nontargeted analyses of 73 polar ionic metabolites measured in serum filtrate samples with stringent quality control (QC) (23). This multiplexed separation platform takes advantage of a serial sample injection format comprising seven (or more) serum filtrates analyzed within a single run using an Agilent 6230 time-of-flight mass spectrometer (TOF-MS) with a coaxial sheath liquid Jetstream electrospray ion source coupled to an Agilent G7100A capillary electrophoresis (CE) unit (23). A standardized protocol for identification and quantification of circulating serum metabolites under positive and negative ion mode

detection is described in more detail elsewhere (23). Briefly, an iterative data workflow was used to effectively filter out spurious signals, redundant peaks, and background ions, when performing targeted and nontargeted metabolite profiling based on analysis of a pooled serum sample that served as a QC sample to monitor technical precision (24). Serum metabolites were reported if they were detected in majority of the individual samples (≥75%) with an acceptable technical precision (CV <30%) based on repeated analysis of QC samples to minimize false discoveries and data overfitting. Missing values (below method detection limit) were set as half of the lowest detected value for each metabolite. Unambiguous identification of most serum metabolites (level 1) in this work was achieved after spiking a pooled serum sample with authentic standards based on their co-migration and accurate mass with low mass error (<5 ppm). These authenticated metabolites were quantified in terms of their absolute concentration (μM) using external calibration curves, where the ion response for each compound was normalized to an internal standard (i.e., relative peak area). Unknown serum metabolites were otherwise annotated based on their relative migration time, accurate mass and most likely molecular formula.

4.3.5 Microbiome data acquisition

A fecal sample was collected from infants at the 1-year visit. Mothers were instructed to collect stool sample from a regular diaper and record the time and date of the sample and place it in a sterile bag in the freezer until their scheduled appointment. Upon arrival, the stool samples were divided into four pre-labeled cryovials and transferred to the lab in a cooler, weighed, and stored at −80 °C. Sample storage, DNA extraction, 16S rRNA gene

sequencing, and analysis has been described in detail previously (25). In brief, the V3 region of 16S rRNA gene (150 base pairs) was sequenced in the McMaster Genomics Facility with 250-base pair sequencing on the MiSeq sequencer (Illumina, Inc.). Adapter, primer and barcode sequences were trimmed from sequencing reads using cutadapt (v1.2.1) (26), and ASVs were inferred using the Divisive Amplicon Denoising Algorithm 2 (DADA2) package in R (27). The naive Bayesian classifier method in DADA2 was used to assign taxonomy using the SILVA 16S rRNA gene reference file.

4.3.6 Statistical Analysis

Data Integration Analysis for Biomarker discovery using Latent cOmponent (DIABLO) was used to integrate 73 serum metabolites and 55 16S rRNA ASVs data to identify discriminant features between children with overweight/obesity and those with normal weight (28, 29). DIABLO is a supervised learning approach based on partial least squares (PLS) that builds on sparse Generalized Canonical Correlation Analysis (sGCCA) and aims to maximize covariance between linear combination of variables (latent component scores) and a response variable. Before proceeding with data integration, individual sparse-PLS-discriminant analysis (sPLS-DA) was used to understand major sources of variation in each dataset and guide the integration process (Supplementary Figures S4.1-2). A 10% prevalence filter was used to remove low-prevalence ASVs and then transformed using the Centered Log Ratio (CLR) with the ALDEx2 package in R (30). Serum metabolome data were transformed using natural logarithm. First, the block.splsda() function was used to determine the optimum number of components based on the performance of the model considering the centroid distance technique and lowest

balanced error rate with a 5-fold cross-validation (repeated 500 times). One component was selected for use in the final model based on the lowest balanced error rate of 39% with a centroid distance metric. Next, a tune.block.splsda() function was applied to choose the optimal number of variables from each data on each component. Furthermore, a plotDiablo() function was used to generate a plot to show the overall correlation between the most discriminant ASVs and metabolites, and circosPlot() to visualize correlations greater than 0.5 between them. Finally, the plot.loadings() function is used to visualize the set of loading vectors assigned to each selected variables in each component. For discriminant analysis, the magnitude of the median value corresponds to the importance of each variable and the colour corresponds to the outcome group (overweight/obese and normal weight) in which the variable is most abundant.

Logistic regression models were also used to examine the association of the identified discriminant ASVs and metabolites with overweight/obesity. We estimated overweight/obesity per standard deviation increase in log-transformed serum metabolite level. Odds ratios (OR), 95% confidence intervals (95% CI), and p-values were reported. All analyses were carried out using R software, version 1.2.5.

4.4 Results

4.4.1 Descriptive statistics

The distribution of demographic characteristics in the overall cohort and by child adiposity status are presented in Table 4.1. A total of 11 (22.0%) children were classified as overweight/obese and about 36% of these children were males. Children with overweight/obesity compared to those with normal weight had a lower mean gestational

age (38.49 vs. 39.26 months) and social disadvantage index (0.91 vs. 2.06), and were more likely to have been born preterm (18.2% vs. 7.7% corresponding to mean gestational age of 35.52 vs 39.52 months). As sensitivity analysis, we compared the AUC derived variable with BMI z-score using World Health Organization (WHO) Child Growth Standards and found that on average those classified as overweight/obese using AUC had higher BMI z-scores at ages 1-3 years compared those classified as normal weight using AUC (Supplementary Figure S4.3). A contingency table comparing overweight/obese status using the AUC method with the BMI z-score calculated using WHO Child Growth Standards are presented in Supplementary Table S4.1.

4.4.2 Integrative analysis of ASVs and serum metabolome

The DIABLO analysis revealed a weak correlation (r=0.28) between discriminant ASVs and circulating serum metabolites (Figure 1A). The optimal feature panel consisted of 9 ASVs and 10 serum metabolites, which produced the highest correlations across the datasets and discriminated children with overweight/obesity and those with normal weight. The contribution of each selected feature based on its loading weights is shown in Figure 1B. The most important serum metabolites associated with overweight/obesity in this cohort were glutamic acid, acetylcarnitine, threonine, carnitine, tryptophan, and asparagine, and the most important ASVs at the genus level were *Pseudobutyrivibrio*, *Lactobacillus*, *Rothia*, and *Lachnospira*. On the other hand, the most important serum metabolites associated with normal weight relative to children with overweight/obesity were γ-aminobutyric acid (GABA), symmetric dimethylarginine (SDMA), asymmetric dimethylarginine (ADMA), and uric acid, whereas ASVs were members of the genera

Clostridium sensu stricto 1, *Akkermansia*, *Hungatella*, *Roseburia*, and *Erysipelatoclostridium*. A circos plot displays correlated features between selected ASVs and serum metabolites using a minimum cut-off value of r = 0.4 (Figure 4.1C). In this case, *Pseudobutyrivibrio* and *Lactobacillus* were inversely correlated with GABA (r= -0.43 and r= -0.41, respectively) and SDMA (r= -0.42 and r= -0.40, respectively), and *Akkermansia* was positively correlated with GABA (r= 0.43) and SDMA (r= 0.41). Pearson correlation between the discriminatory metabolites and ASVs are depicted in Figures 4.2-3.

4.4.3 Associations of selected metabolites and ASVs with overweight/obesity

The associations between serum metabolites and ASVs with overweight/obesity status are presented in Table 2. For metabolites involved in glutamate metabolic pathway, higher serum glutamic acid was positively associated with odds of childhood adiposity (OR per SD=2.9; 95% CI=1.3, 7.4), whereas higher serum GABA was negatively associated (OR per SD=0.5; 95% CI=0.2, 0.8) with the odds of childhood adiposity compared to children with normal weight. Two main metabolites involved in carnitine metabolism, namely carnitine (OR per SD=5.0; 95% CI=1.4, 32.1) and acetylcarnitine (OR per SD=3.6; 95% CI=1.3, 14.6), were both also positively associated with overweight/obesity. Serum concentration for SDMA (OR per SD=0.4; 95% CI=0.2, 0.8) and ADMA (OR per SD=0.46; 95% CI=0.2, 0.9), both isomers generated via methylation of arginine, were negatively associated with childhood adiposity. Threonine, related to glycine, serine, and threonine metabolic pathway, was positively associated with

childhood adiposity. However, tryptophan, uric acid, and asparagine were not statistically associated with childhood anthropometric measures in this study.

The abundance of a *Pseudobutyrivibrio* (OR=1.3; 95% CI=1.0, 1.7) and *Lactobacillus* (OR=1.2; 95% CI=1.0, 1.5) ASVs were positively associated and the abundance of *Clostridium* sensu stricto 1 (OR=0.7; 95% CI=0.5, 0.9) and *Akkermansia* ASVs (OR=0.2; 95% CI=0.1, 0.7) was inversely associated with children with overweight/obesity compared to normal weight children. ASVs assigned as *Rothia*, *Lachnospira*, *Hungatella*, *Roseburia*, and *Erysipelatoclostridium* were not statistically associated with overweight/obesity. Figure 4.4 shows the distribution of discriminatory serum metabolites and ASVs between children with overweight/obesity and those with normal weight.

4.5 Discussion

This study aimed to identify multi-omic molecular features that discriminated children with overweight/obesity from normal weight children. The results from DIABLO showed coherent patterns between ASVs and circulating metabolites with respect to overweight/obesity and significant associations between the identified features. In the regression models, children with overweight/obese had higher levels of several metabolites at 1 year of age, including glutamic acid, acetylcarnitine, carnitine, and threonine, and lower levels of GABA, SDMA, and ADMA compared to normal weight children. Our results also showed higher abundance of members of the genera *Pseudobutyrivibrio* and *Lactobacillus*, and lower abundance of *Clostridium* sensu stricto 1 and *Akkermansia* in the feces of children with overweight/obesity. *Akkermansia* was

positively correlated with GABA, and *Pseudobutyrivibrio* and *Lactobacillus* were inversely correlated with GABA and SDMA.

Glutamic acid, an α-amino acid necessary for the biosynthesis of glutamate, acquired from foods common to omnivore diets (such as meats, poultry, fish, eggs, and dairy products) was identified to have the greatest discriminatory power and present in higher level in children with overweight/obesity. Elevated glutamate levels is a proposed indicator of future risk of cardiometabolic disorders since it is found in higher concentration prior to development of type 2 diabetes and coronary artery disease in adults (31). Few studies have shown higher levels of glutamic acid in children with obesity (32, 33). In vitro studies have indicated that higher glutamate contributes to glucose-toxicity in pancreatic β-cells (34). Moreover, it is proposed that elevated glutamate levels increase the transamination of pyruvate to alanine, which can lead to the development of obesity-related insulin resistance (35). We observed a moderate correlation between glutamic acid and alanine ($r = 0.57$, $p = <0.001$), and on average higher levels of alanine in children with overweight/obesity compared to those with normal weight (mean = 4.43 vs. 3.64 RPA, $p = 0.0331$). Several factors likely contribute to the variation in circulating levels of glutamic acid, including dietary factors, genetic variation, metabolic complications, and gut microbiota. We used food-frequency data to compare serum levels of glutamine and its associated foods (measured per 1000 kilocalories). Higher glutamic acid was positively correlated with red or processed meat intake ($r = 0.33$, $p = 0.0350$), but it did not correlate with any other protein-based dietary factor (Supplementary Figure S4.4).

In contrast to glutamic acid, GABA, synthesized from glutamate by glutamic acid decarboxylase, was present in higher levels in the normal weight sample, and this finding is in agreement with a previous study (36). Hepatic GABA synthesis may modulate insulin and glucagon secretion, homeostatic model assessment for insulin resistance (HOMA-IR), type 2 diabetes, and BMI (37). Recent evidence showed that alterations in gut microbiome can influence changes in plasma concentration of glutamate and GABA levels (38, 39). In our study, GABA was positively correlated with member of the genus *Akkermansia*, and negatively correlated with *Pseudobutyrivibrio* and *Lactobacillus*. Higher abundance of *Akkermansia* has been shown to impact the net production capacity of GABA (40). Although there is no evidence in children, increased abundance of *Akkermansia* have been inversely associated with higher fasting glucose, waist-to-hip ratio, and subcutaneous adipocyte diameter in adults (41, 42). Meanwhile, *Pseudobutyrivibrio* (11) and *Lactobacillus* (12) have been found in higher abundance in feces of obese children, and higher abundance of *Lactobacillus* was correlated with plasma inflammatory marker C-reactive protein (12). Previously, *Lactobacillus* species have been identified in microbiota of breast milk and can be transmitted to infants through breastfeeding. Our data confirms this finding where children who were breastfed until 1-year had higher abundance of *Lactobacillus* compared to those who were breastfed less than 1 year or never breastfed. Although these results are intriguing and somewhat consistent, the mechanism involved are not clear, thus further studies are required to understand the role of these ASVs and metabolites in childhood obesity.

We also observed higher levels of carnitine and acetylcarnitine in children who were overweight/obese. Carnitine is largely acquired from breast milk and formula milk in newborns but can also be synthesized endogenously from two essential amino acids, lysine and methionine (43, 44). However, although our data showed a strong correlation between lysine and methionine ($r = 0.80$, $p = <0.001$), we did not find any correlation of carnitine levels with lysine and methionine. Carnitine is essential for the transport of long-chain fatty acids from cytoplasm into mitochondria for β-oxidation and energy production, and therefore it has a vital regulatory role in lipid metabolism and body composition (45). Supplementation with carnitine can increase fat oxidation in individuals with overweight/obesity, and therefore has been widely studied for weight loss (46). Thus, the positive effect of carnitine supplementation on body composition conversely suggests that higher body fat might be related to dysfunction of carnitine or lipid metabolism, which would result in higher carnitine levels. Additionally, greater body fat may overload β-oxidation of fatty acids and lead to higher amounts of short- or medium-chain-acylcarnitines (47). Short chain acylcarnitines such as acetylcarnitine and carnitine are associated with higher BMI in children (32, 48, 49). Further, both carnitine and acetylcarnitine have been linked to protein-rich diets (44, 50). However, we could not confirm these associations due to the low consumption of meats in our cohort (Supplemental Figure S4.4).

Another metabolite, threonine, was shown to be present in higher level in children who were overweight/obese, which is consistent with a previous study (51). Dietary threonine restriction may protect against metabolic alterations associated with obesity and

improve metabolism via regulation of liver-derived hormone fibroblast growth factor 21 (51).

In addition to GABA (discussed above), SDMA (52, 53) and its structural isomer ADMA was present in higher levels in normal weight children. Although the underlying cause of the inverse association between SDMA and obesity is unclear, studies have attributed this to increased cellular uptake and hepatic extraction of SDMA, where both mechanisms have been related to increased insulin levels associated with obesity-induced insulin resistance (54, 55).

Tryptophan, an essential aromatic amino acid, acquired from whole foods (such as oats, poultry, fish, eggs, and milk) was present in higher levels in the overweight/obese samples in the DIABLO analysis, but was not statistically significant in the regression analysis, although the effect size supports a potential association (OR = 1.9; p = 0.0667). Our data shows that higher tryptophan levels were correlated with consumption of red meat (r = 0.29, p = 0.0382) and eggs (r = 0.34, p = 0.0151). Overnutrition may lead to excess tryptophan uptake and availability (56, 57). Approximately 90-95% of tryptophan is metabolized by the kynurenine pathway in the liver via indoleamine-2,3-dioxygenase (IDO) into co-enzyme nicotinamide adenine dinucleotide (NAD+) and other bioactive metabolites; and residual tryptophan is largely used for serotonin synthesis (58). Tryptophan catabolism is shifted towards the kynurenine pathway in human obesity induced by inflammatory biomarkers (TNFα and IL-6) and oxidative stress (59-61). Elevated tryptophan (62, 63), IDO activity (62, 64), and kynurenine levels (62, 65), and reduced serotonin production (66) has been shown to be associated with obesity and

related metabolic diseases. Further, alterations in tryptophan metabolism may also be driven by gut microbiome in obesity, as previously shown to be disrupted at the compositional and functional level in individuals with obesity (67). Taken together, the evidence suggests that obesity may induce concomitant alterations of host (kynurenine pathway) and microbial (indole) tryptophan metabolic pathways, both of which are associated with obesity-related inflammation (66).

The optimal feature panel did not include BCAAs (leucine, isoleucine, and valine) and aromatic AAs (phenylalanine and tyrosine), which are considered "good" biomarkers of childhood obesity (17). Nevertheless, our data shows a significant positive correlation between BCAAs and glutamic acid, which is produced during the transamination reaction (first step) in BCCA catabolism (Supplementary Figure S4.5). Higher BCAAs have been shown to disrupt the balance of essential amino acids including tryptophan and threonine (both were correlated with BCAAs in our data), which we observed in higher levels in children with overweight/obesity.

The gut microbiome provides essential capacities for fermentation of non-digestible substrates such as complex plant carbohydrates (dietary fibre) (68). Differences in gut microbiota composition can influence an individual's capacity to extract more energy from diet which in turn can activate lipogenic pathways (69). Several studies have shown children with obesity to have higher levels of bacteria in the Firmicutes phylum and lower in the Bacteroidetes phylum (70), and it is proposed that Firmicutes are more efficient at extracting energy from dietary fiber than Bacteroidetes (71). Given that South Asians have a carbohydrate rich diet, it is possible that they have an elevated risk of

obesity as their microbiome is enriched with bacteria that is more efficient at extracting energy and absorbing more calories.

This study has some limitations that should be considered when interpreting the results. This study had a relatively small sample size; thus, future studies with larger sample sizes are needed to confirm these findings. Nevertheless, the sample size is consistent with a number of previous studies in this context (72-75). Our study was limited to polar ionic metabolites and should be expanded to include fatty acids and lipids in future research. Some of the biological variation is likely related to gene expression, which should be integrated and analyzed simultaneously to provide a better functional connection of the gut microbiome on the host metabolome. Also, the choice of -omic platforms and biological sample (urine, serum, tissue, and faeces) can influence the performance of data integration and comparison. For instance, intestinal bacterial products may represent a larger proportion of fecal metabolome that enter the systemic circulation and have a greater impact on the host compared to serum metabolites. Further, the number of features per dataset may determine the integration process and classification performance. This may explain why we did not observe a significant overlap between the two -omics datasets. Another likely explanation for the identification of small number of ASVs is due to the high-interindividual variation in gut microbiome, which is well-known to confound studies with smaller sample size. Finally, we were unable to adjust for covariates in the regression analysis and therefore causal association cannot be established (76).

4.6 Conclusions

Our study suggests the potential role of integrated molecular analysis for identifying biomarkers to discriminate between children who were overweight/obese and normal weight to unravel biological pathways involved in the pathophysiology of childhood obesity. Notable differences were found in serum metabolome, and between specific metabolites and ASVs associated with overweight/obesity. In particular, several metabolites including glutamic acid, acetylcarnitine, carnitine, and threonine were higher, and GABA, SDMA, and ADMA were lower in children with overweight/obesity. Additionally, *Akkermansia* was positively correlated and *Pseudobutyrivibrio* was inversely with both GABA and SDMA, and *Lactobacillus* was inversely correlated with GABA. Further studies on a larger scale and using a larger panel of omics features are required to validate the biomarkers that were found to be associated with overweight/obesity in children and identifying potential therapeutic targets. Understanding the functional capacity of these biomarkers and potential modifiable risk factor (e.g., diet) early in life may lead to targeted early-life screening and interventions, thereby offer a novel approach for prevention of obesity in children.

4.7 References

1. Townshend T, Lake A. Obesogenic environments: current evidence of the built and food environments. Perspect Public Health 2017;137(1):38-44. doi: 10.1177/1757913916679860.
2. Isolauri E, Salminen S, Rautava S. Early Microbe Contact and Obesity Risk: Evidence Of Causality? J Pediatr Gastroenterol Nutr 2016;63 Suppl 1:S3-5. doi: 10.1097/mpg.0000000000001220.
3. Geng J, Ni Q, Sun W, Li L, Feng X. The links between gut microbiota and obesity and obesity related diseases. Biomedicine & Pharmacotherapy 2022;147:112678. doi: https://doi.org/10.1016/j.biopha.2022.112678.
4. Hivert MF, Perng W, Watkins SM, Newgard CS, Kenny LC, Kristal BS, Patti ME, Isganaitis E, DeMeo DL, Oken E, et al. Metabolomics in the developmental origins of obesity and its cardiometabolic consequences. J Dev Orig Health Dis 2015;6(2):65-78. doi: 10.1017/s204017441500001x.
5. Chen X, Sun H, Jiang F, Shen Y, Li X, Hu X, Shen X, Wei P. Alteration of the gut microbiota associated with childhood obesity by 16S rRNA gene sequencing. PeerJ 2020;8:e8317. doi: 10.7717/peerj.8317.
6. Wopereis H, Oozeer R, Knipping K, Belzer C, Knol J. The first thousand days - intestinal microbiology of early life: establishing a symbiosis. Pediatr Allergy Immunol 2014;25(5):428-38. doi: 10.1111/pai.12232.
7. Xu P, Li M, Zhang J, Zhang T. Correlation of intestinal microbiota with overweight and obesity in Kazakh school children. BMC Microbiol 2012;12:283. doi: 10.1186/1471-2180-12-283.
8. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. Nature 2006;444(7122):1022-3. doi: 10.1038/4441022a.
9. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 2006;444(7122):1027-31. doi: 10.1038/nature05414.
10. Bäckhed F, Ding H, Wang T, Hooper LV, Koh GY, Nagy A, Semenkovich CF, Gordon JI. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci U S A 2004;101(44):15718-23. doi: 10.1073/pnas.0407076101.
11. Riva A, Borgo F, Lassandro C, Verduci E, Morace G, Borghi E, Berry D. Pediatric obesity is associated with an altered gut microbiota and discordant shifts in Firmicutes populations. Environ Microbiol 2017;19(1):95-105. doi: 10.1111/1462-2920.13463.
12. Bervoets L, Van Hoorenbeeck K, Kortleven I, Van Noten C, Hens N, Vael C, Goossens H, Desager KN, Vankerckhoven V. Differences in gut microbiota composition between obese and lean children: a cross-sectional study. Gut Pathog 2013;5(1):10. doi: 10.1186/1757-4749-5-10.

13. Borgo F, Verduci E, Riva A, Lassandro C, Riva E, Morace G, Borghi E. Relative Abundance in Bacterial and Fungal Gut Microbes in Obese Children: A Case Control Study. Child Obes 2017;13(1):78-84. doi: 10.1089/chi.2015.0194.
14. Short MI, Hudson R, Besasie BD, Reveles KR, Shah DP, Nicholson S, Johnson-Pais TL, Weldon K, Lai Z, Leach RJ, et al. Comparison of rectal swab, glove tip, and participant-collected stool techniques for gut microbiome sampling. BMC Microbiol 2021;21(1):26. doi: 10.1186/s12866-020-02080-3.
15. Vallianou N, Stratigou T, Christodoulatos GS, Dalamaga M. Understanding the Role of the Gut Microbiome and Microbial Metabolites in Obesity and Obesity-Associated Metabolic Disorders: Current Evidence and Perspectives. Curr Obes Rep 2019;8(3):317-32. doi: 10.1007/s13679-019-00352-2.
16. Vernocchi P, Del Chierico F, Putignani L. Gut Microbiota Profiling: Metabolomics Based Approach to Unravel Compounds Affecting Human Health. Front Microbiol 2016;7:1144. doi: 10.3389/fmicb.2016.01144.
17. Handakas E, Lau CH, Alfano R, Chatzi VL, Plusquin M, Vineis P, Robinson O. A systematic review of metabolomic studies of childhood obesity: State of the evidence for metabolic determinants and consequences. Obes Rev 2022;23 Suppl 1:e13384. doi: 10.1111/obr.13384.
18. Sivasubramanian R, Malhotra S, Fitch AK, Singhal V. Obesity and Metabolic Care of Children of South Asian Ethnicity in Western Society. Children (Basel) 2021;8(6). doi: 10.3390/children8060447.
19. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004;363(9403):157-63. doi: 10.1016/s0140-6736(03)15268-3.
20. de Wilde JA, Dekker M, Middelkoop BJC. BMI-for-age in South Asian children of 0-20 years in the Netherlands: secular changes and misclassification by WHO growth references. Ann Hum Biol 2018;45(2):116-22. doi: 10.1080/03014460.2018.1445288.
21. Anand SS, Vasudevan A, Gupta M, Morrison K, Kurpad A, Teo KK, Srinivasan K. Rationale and design of South Asian Birth Cohort (START): a Canada-India collaborative study. BMC Public Health 2013;13:79. doi: 10.1186/1471-2458-13-79.
22. Bedogni G, Iughetti L, Ferrari M, Malavolti M, Poli M, Bernasconi S, Battistini N. Sensitivity and specificity of body mass index and skinfold thicknesses in detecting excess adiposity in children aged 8-12 years. Ann Hum Biol 2003;30(2):132-9. doi: 10.1080/0301446021000033409.
23. Shanmuganathan M, Kroezen Z, Gill B, Azab S, de Souza RJ, Teo KK, Atkinson S, Subbarao P, Desai D, Anand SS, et al. The maternal serum metabolome by multisegment injection-capillary electrophoresis-mass spectrometry: a high-throughput platform and standardized data workflow for large-scale epidemiological studies. Nature protocols 2021;16(4):1966-94. doi: 10.1038/s41596-020-00475-0.
24. DiBattista A, McIntosh N, Lamoureux M, Al-Dirbashi OY, Chakraborty P, Britz-McKibbin P. Metabolic Signatures of Cystic Fibrosis Identified in Dried Blood

Spots For Newborn Screening Without Carrier Identification. J Proteome Res 2019;18(3):841-54. doi: 10.1021/acs.jproteome.8b00351.
25. Stearns JC, Zulyniak MA, de Souza RJ, Campbell NC, Fontes M, Shaikh M, Sears MR, Becker AB, Mandhane PJ, Subbarao P, et al. Ethnic and diet-related differences in the healthy infant microbiome. Genome Medicine 2017;9(1):32. doi: 10.1186/s13073-017-0421-5.
26. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 2011;17(1):3. doi: 10.14806/ej.17.1.200.
27. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 2016;13(7):581-3. doi: 10.1038/nmeth.3869.
28. Singh A, Shannon CP, Gautier B, Rohart F, Vacher M, Tebbutt SJ, KA LC. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. Bioinformatics 2019;35(17):3055-62. doi: 10.1093/bioinformatics/bty1054.
29. Rohart F, Gautier B, Singh A, KA LC. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS Comput Biol 2017;13(11):e1005752. doi: 10.1371/journal.pcbi.1005752.
30. Fernandes AD, Reid JNS, Macklaim JM, McMurrough TA, Edgell DR, Gloor GB. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. Microbiome 2014;2(1):15. doi: 10.1186/2049-2618-2-15.
31. Ottosson F, Smith E, Melander O, Fernandez C. Altered Asparagine and Glutamate Homeostasis Precede Coronary Artery Disease and Type 2 Diabetes. J Clin Endocrinol Metab 2018;103(8):3060-9. doi: 10.1210/jc.2018-00546.
32. Butte NF, Liu Y, Zakeri IF, Mohney RP, Mehta N, Voruganti VS, Goring H, Cole SA, Comuzzie AG. Global metabolomic profiling targeting childhood obesity in the Hispanic population. Am J Clin Nutr 2015;102(2):256-67. doi: 10.3945/ajcn.115.111872.
33. McCormack SE, Shaham O, McCarthy MA, Deik AA, Wang TJ, Gerszten RE, Clish CB, Mootha VK, Grinspoon SK, Fleischman A. Circulating branched-chain amino acid concentrations are associated with obesity and future insulin resistance in children and adolescents. Pediatric Obesity 2013;8(1):52-61. doi: 10.1111/j.2047-6310.2012.00087.x.
34. Davalli AM, Perego C, Folli FB. The potential role of glutamate in the current diabetes epidemic. Acta Diabetol 2012;49(3):167-83. doi: 10.1007/s00592-011-0364-z.
35. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, Haqq AM, Shah SH, Arlotto M, Slentz CA, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab 2009;9(4):311-26. doi: 10.1016/j.cmet.2009.02.002.
36. Short KR, Chadwick JQ, Teague AM, Tullier MA, Wolbert L, Coleman C, Copeland KC. Effect of Obesity and Exercise Training on Plasma Amino Acids

and Amino Metabolites in American Indian Adolescents. J Clin Endocrinol Metab 2019;104(8):3249-61. doi: 10.1210/jc.2018-02698.
37. Geisler CE, Ghimire S, Bruggink SM, Miller KE, Weninger SN, Kronenfeld JM, Yoshino J, Klein S, Duca FA, Renquist BJ. A critical role of hepatic GABA in the metabolic dysfunction and hyperphagia of obesity. Cell Rep 2021;35(13):109301. doi: 10.1016/j.celrep.2021.109301.
38. Ottosson F, Brunkwall L, Ericson U, Nilsson PM, Almgren P, Fernandez C, Melander O, Orho-Melander M. Connection Between BMI-Related Plasma Metabolite Profile and Gut Microbiota. J Clin Endocrinol Metab 2018;103(4):1491-501. doi: 10.1210/jc.2017-02114.
39. Strandwitz P. Neurotransmitter modulation by the gut microbiota. Brain Res 2018;1693(Pt B):128-33. doi: 10.1016/j.brainres.2018.03.015.
40. Baldini F, Hertel J, Sandt E, Thinnes CC, Neuberger-Castillo L, Pavelka L, Betsou F, Krüger R, Thiele I. Parkinson's disease-associated alterations of the gut microbiome predict disease-relevant changes in metabolic functions. BMC Biol 2020;18(1):62. doi: 10.1186/s12915-020-00775-7.
41. Dao MC, Everard A, Aron-Wisnewsky J, Sokolovska N, Prifti E, Verger EO, Kayser BD, Levenez F, Chilloux J, Hoyles L, et al. Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. Gut 2016;65(3):426-36. doi: 10.1136/gutjnl-2014-308778.
42. Depommier C, Everard A, Druart C, Plovier H, Van Hul M, Vieira-Silva S, Falony G, Raes J, Maiter D, Delzenne NM, et al. Supplementation with Akkermansia muciniphila in overweight and obese human volunteers: a proof-of-concept exploratory study. Nat Med 2019;25(7):1096-103. doi: 10.1038/s41591-019-0495-2.
43. Reuter SE, Evans AM. Carnitine and acylcarnitines: pharmacokinetic, pharmacological and clinical aspects. Clin Pharmacokinet 2012;51(9):553-72. doi: 10.1007/bf03261931.
44. Rafiq T, Azab SM, Teo KK, Thabane L, Anand SS, Morrison KM, de Souza RJ, Britz-McKibbin P. Nutritional Metabolomics and the Classification of Dietary Biomarker Candidates: A Critical Review. Adv Nutr 2021;12(6):2333-57. doi: 10.1093/advances/nmab054.
45. Bremer J. Carnitine--metabolism and functions. Physiol Rev 1983;63(4):1420-80. doi: 10.1152/physrev.1983.63.4.1420.
46. Wutzke KD, Lorenz H. The effect of l-carnitine on fat oxidation, protein turnover, and body composition in slightly overweight subjects. Metabolism 2004;53(8):1002-6. doi: 10.1016/j.metabol.2004.03.007.
47. Yoo H-H, Yoon H-J, Shin H-J, Lee S-H, Yoon H-R. Characterization of plasma carnitine level in obese adolescent Korean women. Biomolecules & Therapeutics 2009;17(2):181-7.
48. Hirschel J, Vogel M, Baber R, Garten A, Beuchel C, Dietz Y, Dittrich J, Körner A, Kiess W, Ceglarek U. Relation of Whole Blood Amino Acid and Acylcarnitine

Metabolome to Age, Sex, BMI, Puberty, and Metabolic Markers in Children and Adolescents. Metabolites 2020;10(4). doi: 10.3390/metabo10040149.
49. Moran-Ramos S, Ocampo-Medina E, Gutierrez-Aguilar R, Macias-Kauffer L, Villamil-Ramirez H, Lopez-Contreras BE, Leon-Mimila P, Vega-Badillo J, Gutierrez-Vidal R, Villarruel-Vazquez R, et al. An Amino Acid Signature Associated with Obesity Predicts 2-Year Risk of Hypertriglyceridemia in School-Age Children. Scientific Reports 2017;7:5607. doi: 10.1038/s41598-017-05765-4.
50. Bouter KE, van Raalte DH, Groen AK, Nieuwdorp M. Role of the Gut Microbiome in the Pathogenesis of Obesity and Obesity-Related Metabolic Dysfunction. Gastroenterology 2017;152(7):1671-8. doi: https://doi.org/10.1053/j.gastro.2016.12.048.
51. Yap YW, Rusu PM, Chan AY, Fam BC, Jungmann A, Solon-Biet SM, Barlow CK, Creek DJ, Huang C, Schittenhelm RB, et al. Restriction of essential amino acids dictates the systemic metabolic response to dietary protein dilution. Nature Communications 2020;11(1):2894. doi: 10.1038/s41467-020-16568-z.
52. Cetinalp-Demircan P, Can A, Bekpinar S, Unlucerci Y, Orhan Y. Unchanged asymmetric dimethylarginine levels in non-diabetic, premenopausal obese women who have common risk factors for cardiovascular disease. Endocrine 2007;31(2):100-4. doi: 10.1007/s12020-007-0005-z.
53. Gruber HJ, Mayer C, Meinitzer A, Almer G, Horejsi R, Möller R, Pilz S, März W, Gasser R, Truschnig-Wilders M, et al. Asymmetric dimethylarginine (ADMA) is tightly correlated with growth in juveniles without correlations to obesity related disorders. Exp Clin Endocrinol Diabetes 2008;116(9):520-4. doi: 10.1055/s-2008-1062712.
54. Zsuga J, Török J, Magyar MT, Valikovics A, Gesztelyi R, Lenkei A, Csiba L, Kéki S, Zsuga M, Bereczki D. Dimethylarginines at the crossroad of insulin resistance and atherosclerosis. Metabolism 2007;56(3):394-9. doi: 10.1016/j.metabol.2006.10.023.
55. Schutte AE, Schutte R, Huisman HW, van Rooyen JM, Fourie CM, Malan L, Malan NT, Schwedhelm E, Strimbeanu S, Anderssohn M, et al. Dimethylarginines: their vascular and metabolic roles in Africans and Caucasians. Eur J Endocrinol 2010;162(3):525-33. doi: 10.1530/eje-09-0865.
56. Strasser B, Berger K, Fuchs D. Effects of a caloric restriction weight loss diet on tryptophan metabolism and inflammatory biomarkers in overweight adults. European Journal of Nutrition 2015;54(1):101-7. doi: 10.1007/s00394-014-0690-3.
57. Brandacher G, Hoeller E, Fuchs D, Weiss HG. Chronic immune activation underlies morbid obesity: is IDO a key player? Curr Drug Metab 2007;8(3):289-95. doi: 10.2174/138920007780362590.
58. Badawy AA. Kynurenine Pathway of Tryptophan Metabolism: Regulatory and Functional Aspects. Int J Tryptophan Res 2017;10:1178646917691938. doi: 10.1177/1178646917691938.

59. Wang Q, Liu D, Song P, Zou MH. Tryptophan-kynurenine pathway is dysregulated in inflammation, and immune activation. Front Biosci (Landmark Ed) 2015;20(7):1116-43. doi: 10.2741/4363.
60. Chaves Filho AJM, Lima CNC, Vasconcelos SMM, de Lucena DF, Maes M, Macedo D. IDO chronic immune activation and tryptophan metabolic pathway: A potential pathophysiological link between depression and obesity. Prog Neuropsychopharmacol Biol Psychiatry 2018;80(Pt C):234-49. doi: 10.1016/j.pnpbp.2017.04.035.
61. Sforzini L, Nettis MA, Mondelli V, Pariante CM. Inflammation in cancer and depression: a starring role for the kynurenine pathway. Psychopharmacology (Berl) 2019;236(10):2997-3011. doi: 10.1007/s00213-019-05200-8.
62. Groer M, Fuchs D, Duffy A, Louis-Jacques A, D'Agata A, Postolache TT. Associations Among Obesity, Inflammation, and Tryptophan Catabolism in Pregnancy. Biol Res Nurs 2018;20(3):284-91. doi: 10.1177/1099800417738363.
63. Theofylaktopoulou D, Midttun Ø, Ulvik A, Ueland PM, Tell GS, Vollset SE, Nygård O, Eussen SJ. A community-based study on determinants of circulating markers of cellular immune activation and kynurenines: the Hordaland Health Study. Clin Exp Immunol 2013;173(1):121-30. doi: 10.1111/cei.12092.
64. Boulet MM, Chevrier G, Grenier-Larouche T, Pelletier M, Nadeau M, Scarpa J, Prehn C, Marette A, Adamski J, Tchernof A. Alterations of plasma metabolite profiles related to adipose tissue distribution and cardiometabolic risk. Am J Physiol Endocrinol Metab 2015;309(8):E736-46. doi: 10.1152/ajpendo.00231.2015.
65. Carayol M, Leitzmann MF, Ferrari P, Zamora-Ros R, Achaintre D, Stepien M, Schmidt JA, Travis RC, Overvad K, Tjonneland A, et al. Blood Metabolic Signatures of Body Mass Index: A Targeted Metabolomics Study in the EPIC Cohort. Journal of Proteome Research 2017;16(9):3137-46. doi: 10.1021/acs.jproteome.6b01062.
66. Cussotto S, Delgado I, Anesi A, Dexpert S, Aubert A, Beau C, Forestier D, Ledaguenel P, Magne E, Mattivi F, et al. Tryptophan Metabolic Pathways Are Altered in Obesity and Are Associated With Systemic Inflammation. Frontiers in Immunology 2020;11. doi: 10.3389/fimmu.2020.00557.
67. Kaur H, Bose C, Mande SS. Tryptophan Metabolism by Gut Microbiome and Gut-Brain-Axis: An in silico Analysis. Front Neurosci 2019;13:1365. doi: 10.3389/fnins.2019.01365.
68. Francescangeli F, De Angelis ML, Zeuner A. Dietary Factors in the Control of Gut Homeostasis, Intestinal Stem Cells, and Colorectal Cancer. Nutrients 2019;11(12). doi: 10.3390/nu11122936.
69. Boroni Moreira AP, Fiche Salles Teixeira T, do CGPM, de Cássia Gonçalves Alfenas R. Gut microbiota and the development of obesity. Nutr Hosp 2012;27(5):1408-14. doi: 10.3305/nh.2012.27.5.5887.
70. Indiani C, Rizzardi KF, Castelo PM, Ferraz LFC, Darrieux M, Parisotto TM. Childhood Obesity and Firmicutes/Bacteroidetes Ratio in the Gut Microbiota: A Systematic Review. Child Obes 2018;14(8):501-9. doi: 10.1089/chi.2018.0040.

71. Magne F, Gotteland M, Gauthier L, Zazueta A, Pesoa S, Navarrete P, Balamurugan R. The Firmicutes/Bacteroidetes Ratio: A Relevant Marker of Gut Dysbiosis in Obese Patients? Nutrients 2020;12(5). doi: 10.3390/nu12051474.
72. Pekmez CT, Larsson MW, Lind MV, Vazquez Manjarrez N, Yonemitsu C, Larnkjaer A, Bode L, Mølgaard C, Michaelsen KF, Dragsted LO. Breastmilk Lipids and Oligosaccharides Influence Branched Short-Chain Fatty Acid Concentrations in Infants with Excessive Weight Gain. Mol Nutr Food Res 2020;64(3):e1900977. doi: 10.1002/mnfr.201900977.
73. Clos-Garcia M, Andrés-Marin N, Fernández-Eulate G, Abecia L, Lavín JL, van Liempd S, Cabrera D, Royo F, Valero A, Errazquin N, et al. Gut microbiome and serum metabolome analyses identify molecular biomarkers and altered glutamate metabolism in fibromyalgia. EBioMedicine 2019;46:499-511. doi: 10.1016/j.ebiom.2019.07.031.
74. Konigsberg IR, Borie R, Walts AD, Cardwell J, Rojas M, Metzger F, Hauck SM, Fingerlin TE, Yang IV, Schwartz DA. Molecular Signatures of Idiopathic Pulmonary Fibrosis. Am J Respir Cell Mol Biol 2021;65(4):430-41. doi: 10.1165/rcmb.2020-0546OC.
75. Giallourou N, Urbaniak C, Puebla-Barragan S, Vorkas PA, Swann JR, Reid G. Characterizing the breast cancer lipidome and its interaction with the tissue microbiota. Commun Biol 2021;4(1):1229. doi: 10.1038/s42003-021-02710-0.
76. Rafiq T, Azab SM, Anand SS, Thabane L, Shanmuganathan M, Morrison KM, Atkinson SA, Stearns JC, Teo KK, Britz-McKibbin P, et al. Sources of Variation in Food-Related Metabolites during Pregnancy. Nutrients 2022;14(12). doi: 10.3390/nu14122503.

Table 4.1: Descriptive statistics of maternal and infant characteristics overall and by overweight/obesity status of children in START cohort

Variable	**Overall n = 50**	**Owt/Ob n = 11**	**Normal Weight n = 39**
Sex (Male), n (%)	25 (50.0)	4 (36.36)	21 (53.85)
Maternal age (years), mean (SD)	30.64 (3.88)	31.18 (2.82)	30.49 (4.15)
Gestational age (weeks), mean (SD)	39.09 (1.54)	38.49 (2.23)	39.26 (1.28)
Gestational weight gain (kg), mean (SD)	13.92 (6.52)	14.72 (8.33)	13.68 (8.33)
Pre-pregnancy BMI (kg/m^2), mean (SD)	24.77 (4.78)	24.82 (4.46)	24.76 (4.92)
Gestational diabetes (GDM), n (%)*[1]*	21 (42.0)	7 (63.64)	22 (56.41)
Preterm birth (Yes), n (%)	5 (10.0)	2 (18.18)	3 (7.69)
Mode of delivery, n (%)			
Vaginal	34 (68.0)	9 (81.82)	25 (35.90)
Caesarean section (planned or emergency)	16 (32.0)	2 (18.18)	14 (35.90)
Antibiotic use in labour (Yes), n (%)	26 (53.06)	7 (70.0)	19 (48.72)
Breastfeeding status at 1-year, n (%)			
Yes, and child is still being breast fed	22 (44.0)	4 (36.36)	18 (46.15)
Yes, child was breast fed but now stopped	26 (52.0)	6 (54.55)	20 (51.28)
Child was never breast fed	2 (4.0)	1 (9.09)	1 (2.56)
Time of solid food introduction (months), mean (SD)	6.02 (1.36)	6.0 (0.89)	6.03 (1.48)
Maternal physical activity (min per day), mean (SD)	11.94 (18.05)	13.64 (17.04)	11.45 (18.52)
Social disadvantage index, mean (SD)*[2]*	1.77 (1.38)	0.91 (0.94)	2.06 (1.39)
Total fibre intake at 1 year (grams), mean (SD)	18.35 (8.36)	18.06 (10.44)	18.43 (7.83)
Energy intake at 1 year (kcal), mean (SD)	1872.35 (1016.98)	1981.66 (1486.02)	1841.52 (864.92)
Birthweight (kg), mean (SD)	3.23 (0.46)	3.23 (0.71)	3.23 (0.38)
BMI$_{AUC}$, mean (SD)	48.47 (5.06)	54.57 (5.37)	46.74 (3.41)

SSF_{AUC}, mean (SD)	52.17 (8.90)	63.33 (9.10)	49.02 (5.84)

[1]GDM was defined based on the Born in Bradford oral glucose tolerance test criteria, self-reported GDM, and insulin use in pregnancy.
[2]The maximum social disadvantage index was five, and the lowest possible score was zero, reflecting the least social disadvantage.
BMI = Body mass index, SSF = Sum of skinfold, AUC = Area under the curve

Table 4.2: Results from logistic regression models examining the association of discriminatory metabolites and ASVs with overweight/obesity among children in the START cohort

	Odds ratio[1] (95% CI)	**p-value**		
Metabolites			**Sub-Pathway**	**Super-Pathway**
Glutamic acid	2.86 (1.32, 7.37)	0.0144	Glutamate metabolism	Amino Acid
GABA	0.45 (0.21, 0.86)	0.0204	Glutamate metabolism	Amino acid
Symmetric dimethylarginine	0.43 (0.20, 0.85)	0.0217	Urea cycle; arginine and proline metabolism	Amino acid
Acetylcarnitine	3.60 (1.34, 14.63)	0.0326	Carnitine metabolism	Lipid
Threonine	3.13 (1.29, 10.12)	0.0270	Glycine, serine and threonine metabolism	Amino acid
Carnitine	5.01 (1.44, 32.07)	0.0453	Carnitine metabolism	Lipid
Asymmetric dimethylarginine	0.46 (0.20, 0.92)	0.0378	Urea cycle; arginine and proline metabolism	Amino acid
Uric acid	0.46 (0.19, 0.96)	0.0564	Purine metabolism	Purine derivative
Tryptophan	1.89 (0.98, 4.01)	0.0667	Tryptophan metabolism	Amino acid
Asparagine	2.24 (1.06, 5.60)	0.0527	Alanine and aspartate metabolism	Amino Acid
ASVs			**Phylum**	**Family**
Pseudobutyrivibrio	1.30 (1.04, 1.67)	0.0258	Firmicutes	Lachnospiraceae
***Clostridium* sensu stricto 1**	0.69 (0.47, 0.96)	0.0344	Firmicutes	Clostridiaceae 1

Akkermansia	0.23 (0.05, 0.67)	0.0301	Verrucomicrobia	Akkermansiaceae
Lactobacillus	1.23 (1.01, 1.54)	0.0497	Firmicutes	Lactobacillaceae
Rothia	1.79 (1.06, 3.53)	0.0534	Actinobacteria	Micrococcaceae
Lachnospira	1.27 (0.98, 1.67)	0.0684	Firmicutes	Lachnospiraceae
Hungatella	0.70 (0.43, 1.0)	0.0831	Firmicutes	Lachnospiraceae
Roseburia	0.75 (0.45, 1.00)	0.1237	Firmicutes	Lachnospiraceae
Erysipelatoclostridium	0.73 (0.49, 1.03)	0.0920	Firmicutes	Erysipelotrichaceae

ORs are estimated per standard deviation (SD) increase in log-transformed metabolite levels.

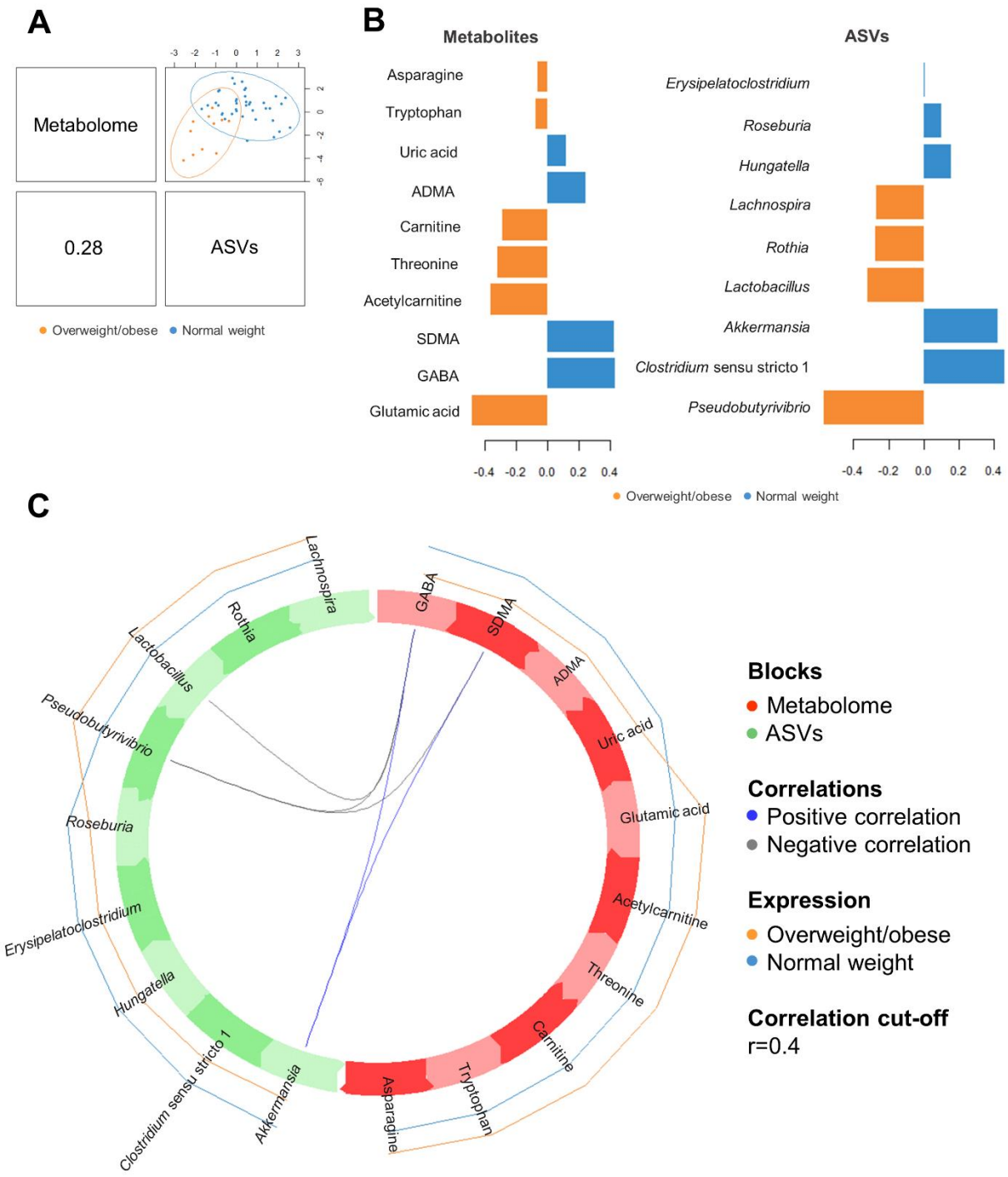


Figure 4.1: DIABLO integrative analysis of metabolome and ASVs discriminatory between overweight/obese and normal weight groups. (A) Matrix scatter plot shows the clustering of samples based on the first component in each dataset and the correlation between the datasets. (B) Loading weights of the selected discriminant metabolites and ASVs. Colours indicate the group in which the median relative abundance is maximum, and values indicate the contribution to the first component. (C) Circos plot showing correlations between the most discriminatory metabolites and ASVs. Positive correlations are displayed using blue line-connectors.

γ-aminobutyric acid (GABA); Symmetric dimethylarginine (SDMA); Asymmetric dimethylarginine (ADMA).

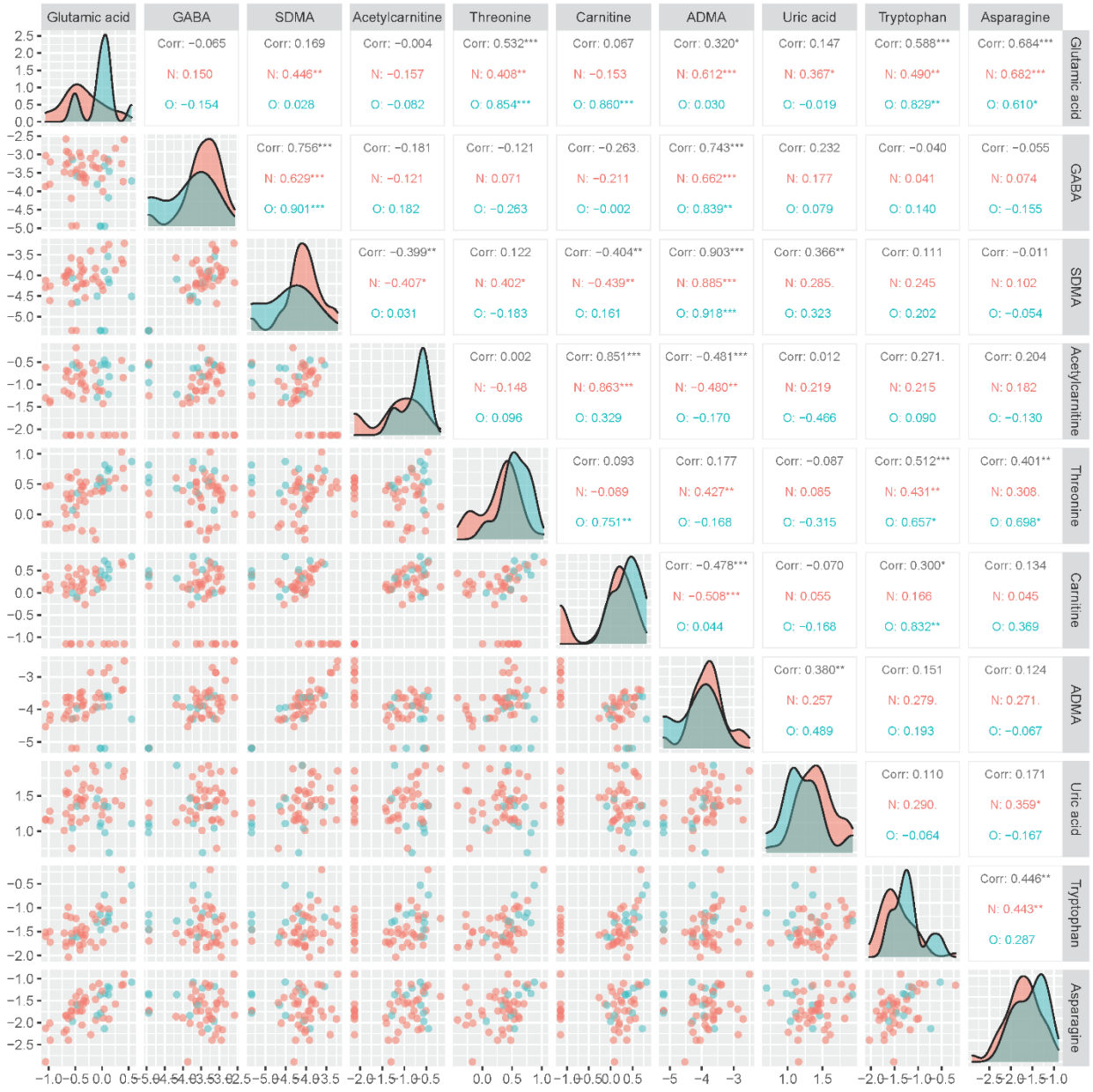


Figure 4.2: Correlation between most discriminatory metabolites in participants overall and by overweight/obese (O) and normal weight (N) groups.

γ-aminobutyric acid (GABA); Symmetric dimethylarginine (SDMA); Asymmetric dimethylarginine (ADMA).

Figure 4.3: Correlation between most discriminatory ASVs in participants overall and by overweight/obese (O) and normal weight (N) groups.

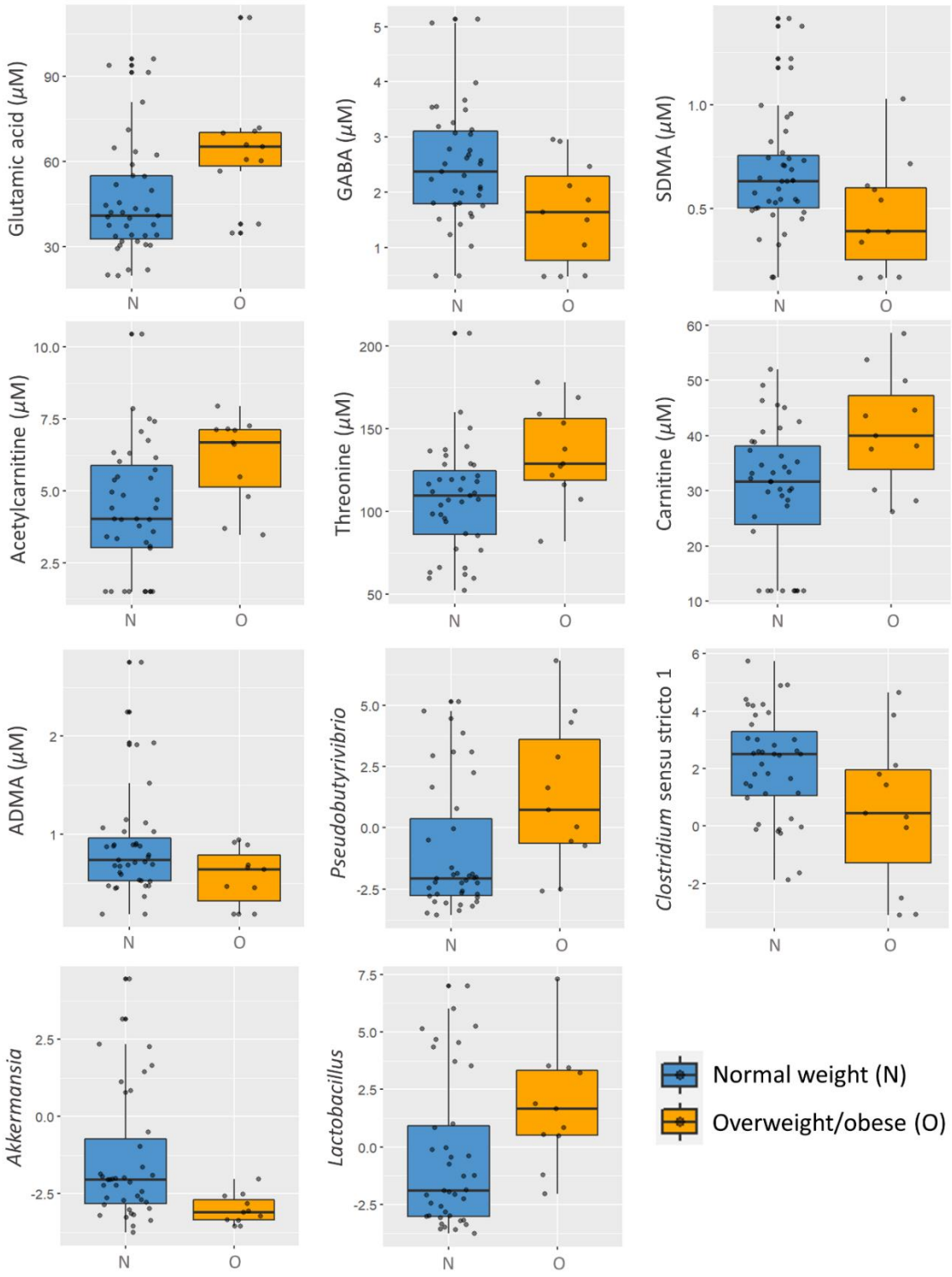


Figure 4.4: Distribution of significantly different metabolites (concentration) and ASV between children who were overweight/obese (O) and normal weight (N).

ASV counts were transformed using CLR-transformation.

Table S4.2: A contingency table illustrating the outcomes of a comparison between BMI_{AUC} and SSF_{AUC}

		Area under the growth curve using Body Mass Index at or above 85th percentile		
		Overweight/obese	Normal weight	Total
Area under the growth curve using sum of skinfold at or above 85th percentile	Overweight/obese	5	3	8
	Normal weight	3	39	42
	Total	8	42	50

Highlighted cells include children with BMI_{AUC} and/or SSF_{AUC} at or above internally derived 85th percentile and were classified as being overweight/obese.

Table S4.2: A contingency table illustrating the outcomes of a comparison between AUC derived and World Health Organization (WHO) Child Growth Standards at 1-3 years

		Area under the growth curve	
		Overweight/obese	Normal weight
BMI z-score at 3 years using the World Health Organization (WHO) Child Growth Standards	Overweight/obese	5	1
	Normal weight	6	38
	Total	11	39

		Area under the growth curve	
		Overweight/obese	Normal weight
BMI z-score at 2 year using World Health Organization (WHO) Child Growth Standards	Overweight/obese	7	6
	Normal weight	4	33
	Total	11	39

		Area under the growth curve	
		Overweight/obese	Normal weight
BMI z-score at 1 year using World Health Organization (WHO) Child Growth Standards	Overweight/obese	9	24
	Normal weight	2	15
	Total	11	39

Cumulative body mass index (BMI_{AUC}) and skinfold thickness (SSF_{AUC}) scores were calculated using from birth to 3-years as the total area under the growth curve (AUC). BMI_{AUC} and/or SSF_{AUC} >85th percentile was used to classify overweight/obesity children.

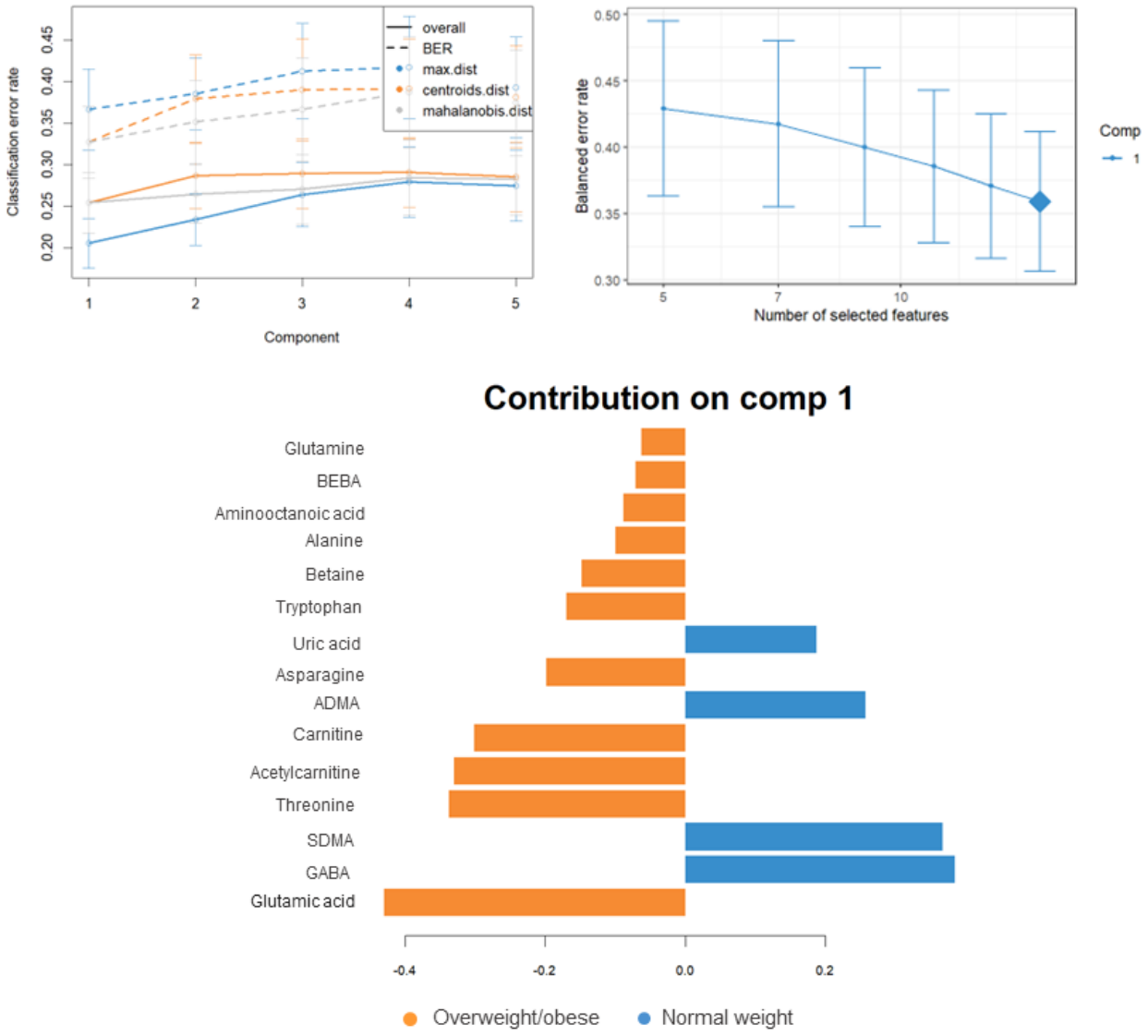


Figure S4.1: PLS-DA analysis of the metabolomics data

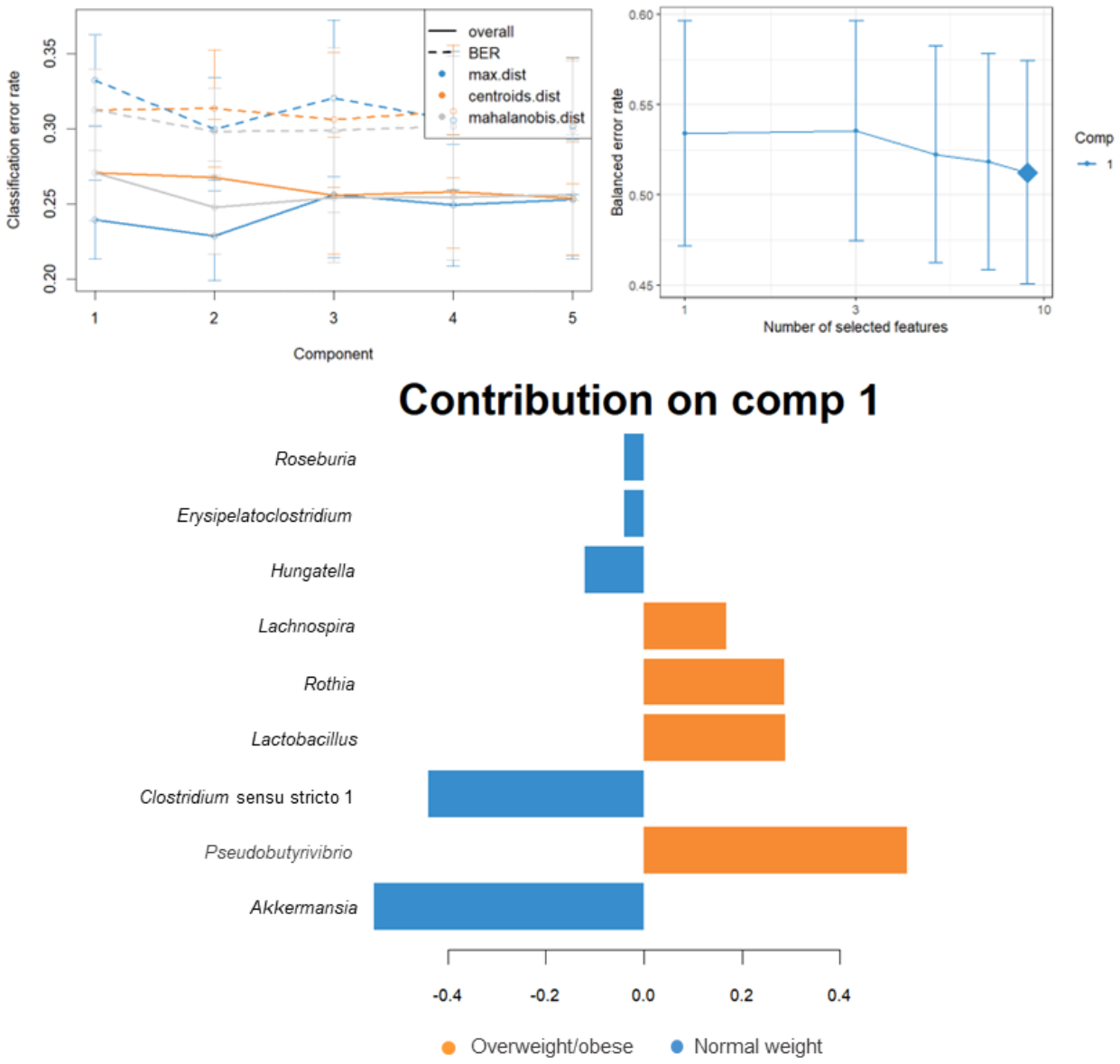


Figure S4.2: PLS-DA analysis of the microbiome data

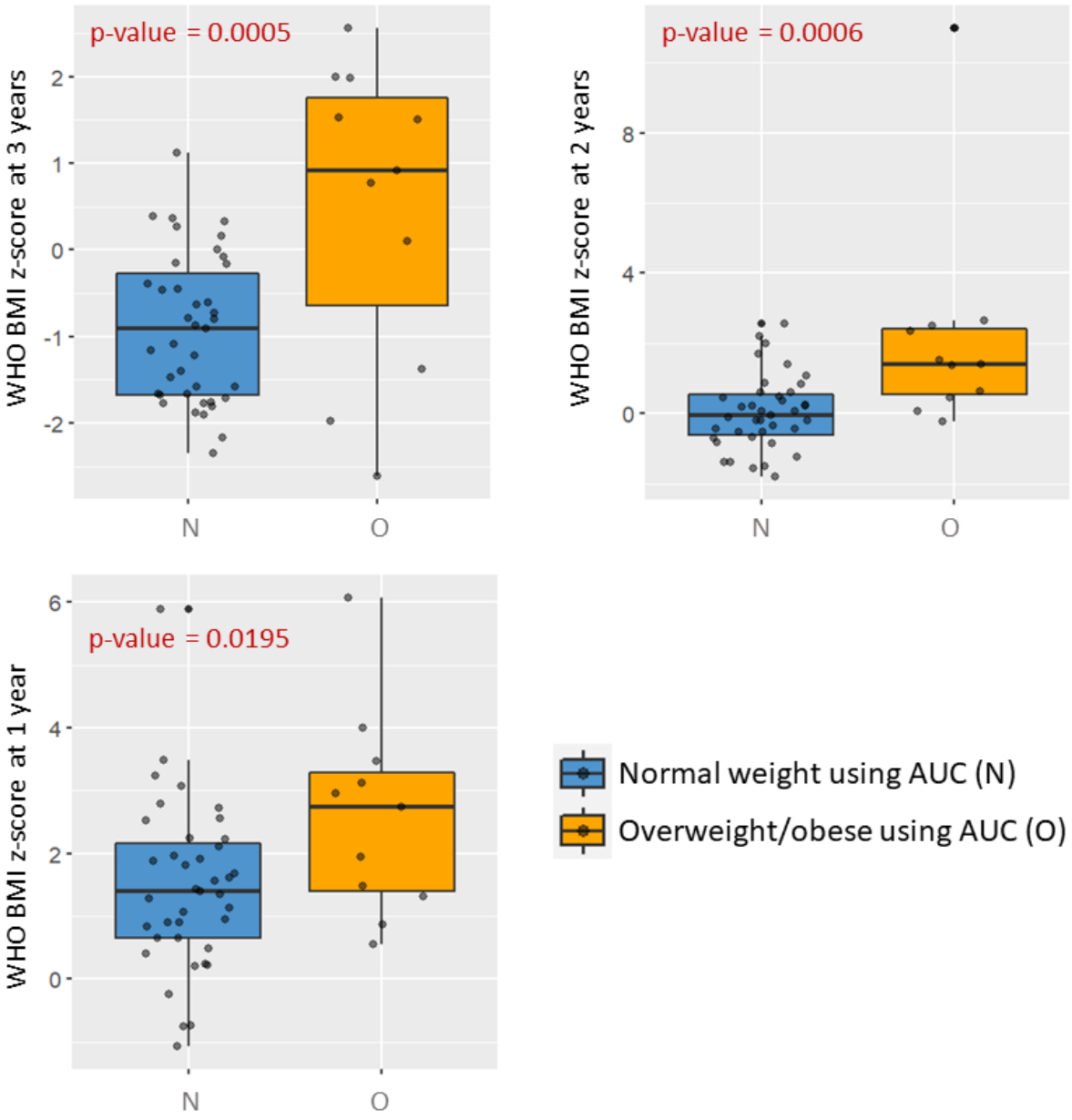


Figure S4.3: Boxplots of Body Mass Index (BMI) z-scores at 1-3 years calculated from World Health Organization (WHO) Child Growth Standards by overweight/obese status using area under the growth curve (AUC) of children

Cumulative body mass index (BMI_{AUC}) and skinfold thickness (SSF_{AUC}) scores were calculated using from birth to 3-years as the total area under the growth curve (AUC). BMI_{AUC} and/or SSF_{AUC} >85th percentile was used to classify overweight/obesity children.

Figure S4.4: Pearson correlation of discriminant metabolites with their associated foods in participants overall and by overweight/obese status of children

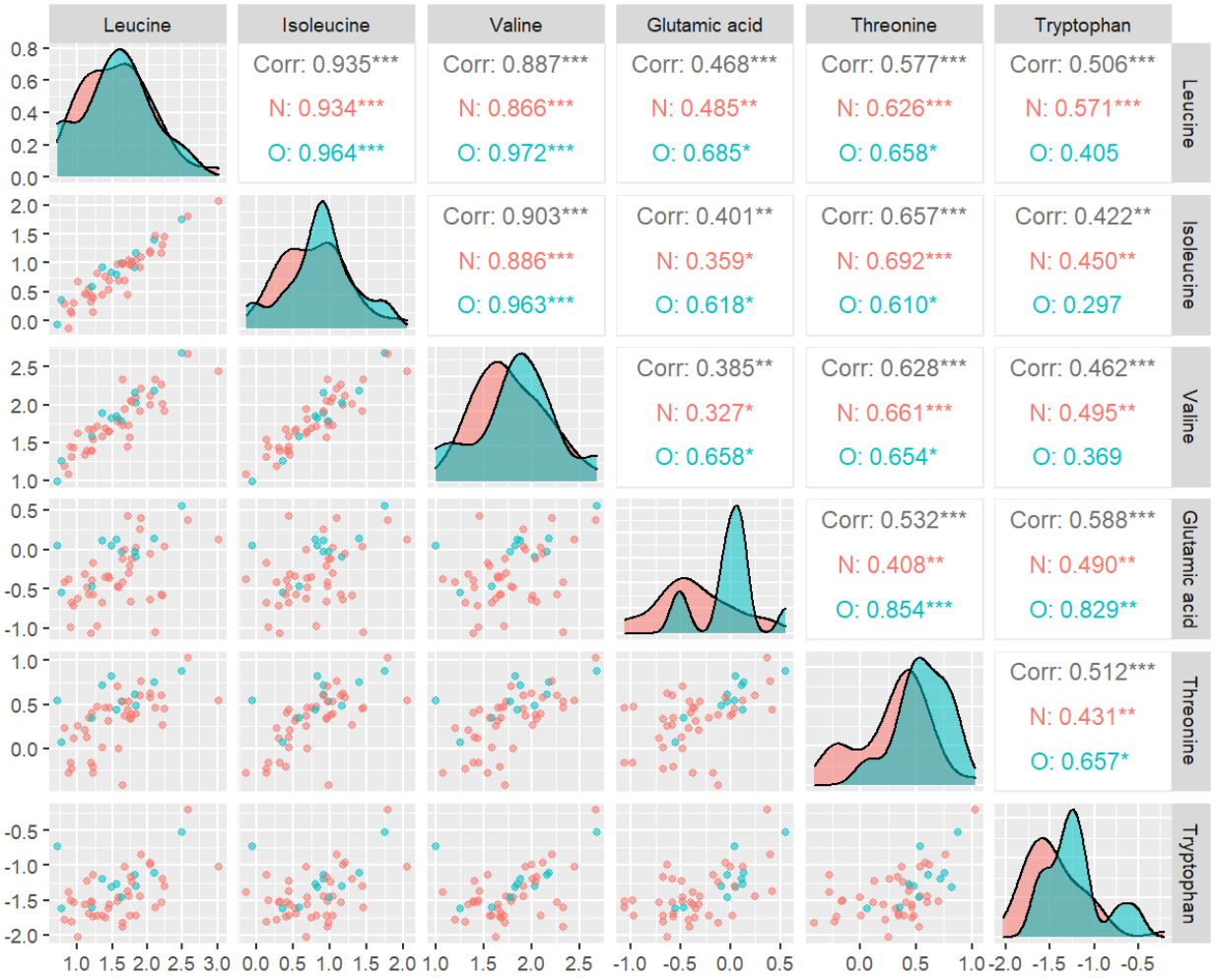


Figure S4.5: Pearson correlation between BCAAs (leucine, isoleucine, and valine), glutamic acid, threonine, and tryptophan in participants overall and by overweight/obese status of children

Chapter 5 – Epilogue and Conclusions

5.1 Overview of major thesis contributions

The aims of this dissertation were (a) to contribute to the growing literature on food-based biomarkers by generating a comprehensive list of food-associated metabolites, (b) provide information on sources of variability to guide appropriate modeling when utilizing metabolomics in nutritional epidemiological studies, and (c) employ multi-omics integration analysis of serum metabolites and ASVs of 16S rRNA genes to identify molecular features that characterize risk of obesity in young children.

The first study (*Chapter II*) of this dissertation was a review of the literature and provided a comprehensive list of metabolites associated with a comprehensive inventory of foods and food groups in apparently healthy individuals, reported on study designs, metabolomic approaches, and biospecimen used, and rated the evidence based on interstudy repeatability and study design. We developed a scoring system based on empirical evidence to rate metabolites as candidate biomarker of food intake into one of three mutually exclusive categories: ‘good’, ‘fair’, or ‘poor’. The results showed that many metabolites can be identified from a specific food, however in several cases, a single metabolite can be a “good” indicator of food intake. While the scoring system has not been published previously, there has been another publication that used a similar quality assessment score to systematically collating metabolites (1). This is an active area of method development, and with more common use of metabolomics in nutrition studies, efforts to improve standardization and methodological quality of such studies is important.

Having identified "good" metabolites of foods in the first study, the purpose of the second study (*Chapter III*) was to provide insight into the external factors that impact serum metabolite concentrations in White Europeans and South Asian pregnant women. The results of this study demonstrated that dietary factors explained the highest proportion of variability in exogenous food-based biomarkers relative to non-dietary factors, whereas the contribution of non-dietary factors was either similar or lower for metabolites that can either be produced endogenously, biotransformed by gut microbiota, and/or derived from more than one food source. Metabolite concentration also differed by ethnicity (South Asian and White Europeans). We confirmed that biomarkers with "good" evidence can be considered as direct surrogates for food intake. However, several factors other than food exposure can influence variation in food-related metabolites concentration and therefore measurement and modeling of these factors requires consideration during the planning and final statistical analyses of the data.

The third study of this dissertation (*Chapter IV*) employed a multi-omic molecular analysis for identifying biomarkers that discriminate children with from those without adiposity in order to provide insight into abnormalities associated with early onset of obesity in childhood. The results showed notable differences in serum metabolome and specific gut bacterial species, and between specific metabolites and bacterial species related to adiposity.

5.2 Clinical significance and contribution of the research

This dissertation has provided useful information on the application of metabolomics in characterizing habitual dietary exposure, insight into the external factors that impact

metabolite concentration during pregnancy, and its role in deciphering interactions between diet, gut microbiome, and childhood obesity. There are several thousand metabolites derived from food and given the rapid pace of development of the field, it has been challenging for the scientific community to critically appraise and classify robust dietary biomarkers. This dissertation comprehensively reviews state-of-the art models, and examines methodological issues related to confounders and data integration to assist nutrition researchers interested in gaining a broader insight into the potential of its application, including the historical transition, recent focuses, and issues that remain to be overcome in future research. The first study of this dissertation generated a comprehensive list of metabolites associated with individual food and food groups which can potentially be employed to (a) correct self-reported data, (b) be used in conjunction with existing dietary assessment methods to measure dietary intake, and/or (c) assess compliance to dietary interventions. In terms of public health implications, the integration of metabolomics into nutritional epidemiology can, to some extent, identify differences in dietary components so that interventions (e.g., dietary recommendations) can be tailored to specific groups of people.

This dissertation also provides insight into the interindividual variation in metabolite levels due to intrinsic physiologic characteristics and extrinsic factors in two ethnically diverse groups during pregnancy. The results showed that for exogenous biomarkers such as proline betaine, dietary factors explained higher proportion of variability compared to nondietary factors. For metabolites that can either be produced endogenously, biotransformed by gut microbiota, and/or derived from more than one

food source, the unique contribution of dietary factors was similar (15:0, 17:0, hippuric acid, and TMAO) or lower (14:0, tryptophan betaine, 3-methylhistidine, and carnitine) compared to non-dietary factors including ethnicity, maternal age, gestational age, pre-pregnancy BMI, physical activity, and smoking history. Further, there was an ethnicity effect for all metabolites, except carnitine and hippuric acid. Reducing measurement error in epidemiological studies is critical because of the potential risk of biasing the effect estimates of the exposure-outcome association (2). When examining associations between a biomarker of food intake and a health outcome, investigators should select a valid biomarker, and analyze and interpret data with the knowledge of the potential factors that may influence metabolite concentrations (3). Similarly, in nutritional epidemiological studies, using a common cutpoint to classify biomarker values may not be applicable to all subgroups and populations. Researchers must consider cutpoints by subgroups formed by sociodemographic and lifestyle variables as well as by ethnicity (3). Further, these findings imply that if appropriate consideration for non-dietary factors during statistical analysis of data is not made then using these biomarkers as markers of food/nutrient intake may simply exchange one source of measurement error (self-misreport) for others (changes in the biomarker intake unrelated to changes in food intake). Careful assessment of nutritional biomarkers in metabolomic studies will assist in understanding the causal association between a dietary exposure and outcome.

There is evidence that biomarkers of food intake are associated with childhood obesity and this association is influenced by microbiome composition. The third study in this dissertation used a multi-omic analysis to identify biomarkers that discriminate

children with and without adiposity in order to provide insight into abnormalities associated with early onset of obesity in childhood. The results demonstrated alterations in a panel of -omics features that discriminated children with and without obesity. The findings suggest that obesity related metabolic pathways such as glutamate and carnitine metabolism may provide insight into the metabolic processes related to early onset of obesity in childhood. Childhood obesity is associated with obesity in adolescence and adulthood, and is associated with an increased risk of a wide range of adverse health outcomes including cardiovascular and metabolic diseases, orthopedic problems, and psychological disorders (4-7). There is strong evidence suggesting that childhood obesity is associated with alternations in the biochemical and physiological processes, which in turn can lead to chronic diseases in later years of life (4). If validated, the findings have the potential to help in the development of early-life screening, and personalized prognostic and intervention approaches.

5.3 Future research directions

Nutritional metabolomics is still in its initial stages and there are several issues that need to be resolved before dietary biomarkers can be utilized in population-based studies. For instance, many of the food metabolites identified in our review were of endogenous origin and generated in the human body during metabolic processes, thus it can be challenging to determine their origin and measure the amount of variance explained by food. Further, the heterogenous composition of many foods (i.e., lack of specificity) hinders the ability to critically appraise and classify robust biomarkers. Therefore, future studies need to validate the BFIs using the criteria developed by Dragsted and colleagues

(8) that include biological plausibility, dose– response, time–response, robustness, reliability, stability, analytical performance, and interlaboratory reproducibility to evaluate the utility of these biomarkers in nutritional research. Currently, the consensus is to combine validated BFIs with self-reported dietary data to account for measurement error in assessment of dietary intake (9). Additionally, it is important that the heterogeneity of intake of certain foods in different populations or ethnicities in different geographical areas are taken into account in statistical analyses examining metabolite and health associations. Thus, not only is the specificity of many foods affecting the assessment and validation of BFIs, the interpretation (i.e., generalizability) of these biomarkers to other diverse populations may be limited.

Candidate BFIs identified in feeding (experimental) studies should subsequently be confirmed in observational studies. However, researchers should acknowledge the potential confounding factors in observational studies, and results must be interpreted with caution if confounders have not been considered in the statistical analysis. For instance, the findings from the second study in the dissertation showed that most of the commonly used biomarkers of habitual food intake were influenced by non-dietary factors, especially ethnicity. As a result, candidate food biomarkers should not be directly used in different ethnic groups for assessment of diet or disease classification without appropriate adjustment for non-dietary factors for an unbiased assessment of metabolite concentration. Moreover, different non-dietary factors may influence metabolite concentration and to varying degrees in different populations. Future research is therefore

required to account for differences in non-dietary factors by ethnicity to control for some of the inter-individual variation in food-related metabolites.

In order to strengthen the evidence for the application of biomarkers derived from food sources, future research should examine how these biomarkers are linked to health outcomes. Study three of this dissertation showed independent associations of metabolites and bacterial species with childhood obesity, as well as the integration of the two -omics dataset to improve the classification of children with obesity from those with normal weight. However, this study had a relatively smaller sample size which did not allow adjustment for covariates. Studies using a longitudinal design with a larger sample size and panel of -omics features are required to validate the results and further provide insight into the underlying mechanisms that could initiate the onset of obesity in children. Future research should examine the metabolic pathways such as glutamate, carnitine, and tryptophan metabolism, that were proposed in study three and their association with childhood obesity. Few additional issues were raised from the integrated multi-omics analysis which needs to be addressed in future studies. It is unclear what constitutes a physiologically relevant difference in -omics pool size, and whether it will differ across different -omics datasets. Also, it can be challenging to sort through extensive data to determine relevant information, especially when some of the findings have not been well-replicated in literature. Given that obesity is a multifactorial process characterized by changes at different molecular levels, future analysis should utilize the integrative approach based on multiomics datasets to understand the interplay between molecular features and their role in the mechanisms underlying obesity. Further, studies need to

determine the time at which dysbiosis of gut microbiota may lead to obesity and whether it is possible to reverse these features to treat obesity. It is also important to determine whether the -omics features differ by the type of obesity (central versus peripheral obesity).

5.4 Conclusion

Overall, the application of metabolomics in nutrition research holds great promise for the assessment of dietary intake using biomarkers. Our results demonstrate that many metabolites can be identified from a specific food, and there are many cases in which a single metabolite is a good indicator of food intake. However, metabolites that reflect specific foods are influenced by non-dietary factors (ethnicity, maternal age, gestational age, pre-pregnancy BMI, physical activity, and smoking history) and to differing degrees. Further, using an integrative multi-omic analyses to characterize molecular features of obesity, notable differences were observed in serum metabolome and specific gut bacterial species, and between specific metabolites and bacterial species related to childhood obesity.

5.5 References

1. U MRA, Shen EY, Cartlidge C, Alkhatib A, Thursz MR, Waked I, Gomaa AI, Holmes E, Sharma R, Taylor-Robinson SD. Optimized Systematic Review Tool: Application to Candidate Biomarkers for the Diagnosis of Hepatocellular Carcinoma. Cancer Epidemiol Biomarkers Prev 2022;31(7):1261-74. doi: 10.1158/1055-9965.Epi-21-0687.
2. Holford TR, Stack C. Study design for epidemiologic studies with measurement error. Statistical methods in medical research 1995;4(4):339-58. doi: 10.1177/096228029500400405.
3. Blanck HM, Bowman BA, Cooper GR, Myers GL, Miller DT. Laboratory issues: use of nutritional biomarkers. J Nutr 2003;133 Suppl 3:888s-94s. doi: 10.1093/jn/133.3.888S.
4. Deckelbaum RJ, Williams CL. Childhood obesity: the health issue. Obes Res 2001;9 Suppl 4:239s-43s. doi: 10.1038/oby.2001.125.
5. Park MH, Falconer C, Viner RM, Kinra S. The impact of childhood obesity on morbidity and mortality in adulthood: a systematic review. Obes Rev 2012;13(11):985-1000. doi: 10.1111/j.1467-789X.2012.01015.x.
6. Dimitri P. The Impact of Childhood Obesity on Skeletal Health and Development. J Obes Metab Syndr 2019;28(1):4-17. doi: 10.7570/jomes.2019.28.1.4.
7. Reilly JJ, Methven E, McDowell ZC, Hacking B, Alexander D, Stewart L, Kelnar CJ. Health consequences of obesity. Arch Dis Child 2003;88(9):748-52. doi: 10.1136/adc.88.9.748.
8. Dragsted LO, Gao Q, Scalbert A, Vergères G, Kolehmainen M, Manach C, Brennan L, Afman LA, Wishart DS, Andres Lacueva C, et al. Validation of biomarkers of food intake-critical assessment of candidate biomarkers. Genes & nutrition 2018;13:14. doi: 10.1186/s12263-018-0603-9.
9. Brennan L, Hu FB, Sun Q. Metabolomics Meets Nutritional Epidemiology: Harnessing the Potential in Metabolomics Data. Metabolites 2021;11(10). doi: 10.3390/metabo11100709.

APPENDIX I – Study Characteristics of Interventional Studies and Metabolites of Foods and Food group

Supplemental Reference	Study population, location	Study design	N (follow-up)	Technique (metabolite targets)	Biofluid	Food category and/or food item
Acar et al., 2019 (1)	18-65 y, Denmark	Feeding Study	146 (26 weeks)	UPLC-Q-ToF-MS (Untargeted)	Plasma	New Nordic diet (NND) (organic diet high in fruit, vegetables, whole grains, and fish). Average Danish Diet (ADD) (high in imported and processed foods)
New Nordic Diet (NND)	Diet related metabolites: pipecolic acid betaine; trimethylamine oxide (TMAO); prolyl hydroxyproline Fat metabolism: polyunsaturated phosphatidylcholines					
Average Danish Diet (ADD)	Diet related metabolites: theobromine; proline betaine Amino acid: indolelactic acid; hydroxy-3-methylbutyrate Fat metabolism: butyryl carnitine					
Allen et al., 2002 (2)	22-39 y,	Crossover Feeding Study	24 (3 days)	HPLC (Targeted – lycopene concentration and changes in lycopene isomer patterns)	Plasma	Control (low lycopene); fresh tomatoes; or processed tomato juice
Tomato	Fresh tomato: Total lycopene; 5-cis lycopene; other cis-lycopene; all-trans-lycopene Processed tomato sauce: Total lycopene; all-trans-lycopene					
Amer et al., 2017 (3)	≥ 18 y, Denmark	Feeding Study	52 (12 weeks)	GC–MS and LC–MS (Untargeted)	Urine & Plasma	Low amount of MCFAs + Whey (WL); High amount of MCFAs + Whey (WH); Low amount of MCFAs + Casein (CL); High amount of MCFAs + Casein (CH)
MCFAs + Whey (WL)	Plasma (GC-MS): Myristic acid; Threonic acid Plasma (LC-MS): Leucine/isoleucine; 3-Hydroxyundecanoic acid; Pivaloylcarnitine; Phenylalanine; Cinnamic acid; 4-Hydroxyphenylacetic acid					
High amount of MCFAs + Whey (WH)	Plasma (GC-MS): Sebacic acid; Succinic acid; Glycerol; Valine; Pyroglutamic acid; Sarcosine; Threonic acid Plasma (LC-MS): Valine; Leucine/isoleucine; Tyrosine; Phenylpyruvic acid; 2-Phenylacetamide; 3-Hydroxyundecanoic acid; Pivaloylcarnitine Urine (GC-MS): Fumaric acid; Citric acid; Succinic acid; Adipic acid; Threonine					
Low amount of MCFAs + Casein (CL)	Plasma (GC): α-Hydroxybutyric acid; Sarcosine; Threonic acid; Alanine Urine (GC-MS): Pyroglutamic acid; Myo-Inositol					

High amount of MCFAs + Casein (CH)		Plasma (GC): α-Hydroxybutyric acid; β-Hydroxybutyric acid; 3,4-Dihydroxybutyric acid; Pyroglutamic acid; Threonic acid; Threonine; Myo-Inositol Urine (GC-MS): Adipic acid				
Andersen et al., 2014 (4)	NND: 44 ± 13 y ADD: 41 ± 13 y, Denmark	Feeding Study	107 (6 months)	UPLC-Q-ToF-MS (Untargeted)	Urine	New Nordic Diet (NND) or an Average Danish Diet (ADD)
Strawberry		2,5-Dimethyl-4-methoxy-3(2H)-furanone sulphate				
Orange/citrus		Proline betaine; Hesperetin glucuronide				
Beetroot		4-Ethyl-5-amino-pyrocatechol sulphate; 4-Ethyl-5-methylamino-pyrocatechol sulphate; 4-Methylpyridine-2-carboxylic acidglycine conjugate				
Green beans		Unsaturated aliphatic hydroxy-dicarboxylic acid				
Red cabbage		3-Hydroxy-3-(methyl-sulphinyl)propanoic acid; 3-Hydroxy-hippuric acid sulphate; 3-Hydroxy-hippuric acid				
Red cabbage (brussels sprouts, pointed cabbage)		Iberin N-acetyl-cysteine (IB-NAC); N-acetyl-S-(N-3-methylthiopropyl)cysteine				
Red cabbage (brussels sprouts, horseradish)		N-acetyl-S-(N-allylthiocarbamoyl)cysteine (AITC-NAC)				
Red cabbage (Brussels sprouts)		Sulphoraphane N-acetyl-cysteine (SFN-NAC)				
Walnut		5-Hydroxyindole-3-acetic acid				
Chocolate		6-Amino-5-[N-methylformylamino]-1-methyluracil (6-AMMU); 7-Methyluric acid; Theobromine				
Andersen et al., 2014 (5)	18–65 y, Denmark	Feeding Study	107 (6 months)	UPLC-Q-ToF-MS (Untargeted)	Urine	40 food groups
Average Danish Diet (ADD)		Octanoyl-glucuronide; 3-indoleacetic acid glucuronide Heat treatment cluster: pyrraline Chocolate cluster: theobromine; 7-methyluric acid; 6-amino-5-[N-methylformylamino]-1- methyluracil; 3,7-dimethyluric acid; 7-methylxanthine Citrus cluster: proline betaine; pyroglutamyl proline Limonene cluster: *p*-menth-1-ene-6,8,9-triol; perillic acid-8,9-diol-glucuronide; limonene-8,9-diol-glucuronide; dihydroperillic acid glucuronide; limonene-1,2-diol glucuronide				

New Nordic Diet (NND)		Hydroquinone glucuronide; (2-oxo-2,3-dihydro-1H-indol-3-yl)acetic acid; 3,4,5,6-tetrahydrohippurate; hippuric acid Fish: trimethylamine N-oxide (TMAO)				
Baenasa et al., 2017 (6)	27-36 y, Spain	Crossover Feeding Study	14 women (7 days)	UHPLC-QqQ-MS/MS (Targeted – glucosinolates, Isothiocyanates, and indoles)	Urine	Broccoli and radish sprouts
Broccoli		Isothiocyanates; indole-3-carbinol (I3C)				
Barnes et al., 2019 (7)	Lean control 23.7 (5.0) Lean mango 25.6 (4.2) Obese mango 27.8 (8.3), US	Feeding Study	32 (6 weeks)	LC-MS (Targeted)	Urine & Plasma	Mango consumption
Mango		Plasma: 4-*O*-methylgallic acid Urine: 4-*O*-methylgallic acid-3-*O*-sulfate; sum of gallotannin metabolites				
Beckmann et al., 2015 (8)	29.9 ± 4.7 y, UK	Feeding Study	90 females (Acute)	FIE-MS and GC-ToF-MS (Untargeted)	Urine & Plasma	0, 50, or 100 g sucrose in 500 mL water
Sucrose		Urine: fructose; erythronic acid; 3-hydroxybutanoic acid Plasma: 3-Hydroxybutanoic acid; dihydroxybutanoic acid				
Bondia-Pons et al., 2013 (9)	28-56 y, Finland	Crossover Feeding Study	20 (4 weeks)	UPLC-Q-ToF-MS (Untargeted)	Urine	Whole gain rye bread (RB) versus refined wheat bread (WB, control)
Whole grain rye bread		3-(3,5-dihydroxyphenyl)-1-propanoic acid (DHPPA) sulphate; Ascorbic acid; 2-Aminophenol sulphate; Nonanedioic acid; DHPPA glucuronide; Indolylacryloylglycine; Enterolactone glucuronide; Ferulic acid-4-O-sulphate; 2,4-Dihydroxy-1,4-benzoxazin-3-one sulphate; 3,5-Dihydroxyphenylethanol sulphate; 1,3,4,5-Tetrahydroxycyclohexane-1-carboxylic acid				
Boto-Ordóñez et al., 2013 (10)	61 ± 9 y, Spain	Feeding Study	36 males (1 month)	UPLC-MS/MS (Targeted – phenolics and microbial derived phenolic acids)	Urine	Dealcoholized red wine (DRW)
Dealcoholized red wine		Phase II Metabolites of (Epi)catechin, Hydroxyphenylvalerolactones, and Hydroxybenzoic Acids Hydroxybenzoic acids Gallic acid metabolites: methylgallic sulfate Ethylgallate metabolites: ethylgallate sulfate; ethylgallate (glucuronide 1); ethylgallate (glucuronide 2)				

		Flavan-3-ols: (Epi)catechin (glucuronide 1); (epi)catechin (glucuronide 2); (epi)catechin (glucuronide 3); (epi)catechin (glucuronide 4); (epi)catechin sulfate 1; (epi)catechin sulfate 2; (epi)catechin sulfate 3; methyl(epi)catechin (glucuronide 1); methyl(epi)catechin (glucuronide 2); methyl(epi)catechin (glucuronide 3) Hydroxyphenylvalerolactones: Dihydroxyphenyl-γ-valerolactone (DHPV) glucuronide 1; DHPV glucuronide 2; DHPV sulfate 1; DHPV sulfate 2; methoxy-hydroxyphenyl-γ-valerolactone (MHPV) glucuronide 1; MHPV sulfate 1; MHPV sulfate 2 <u>Microbial Phenolic Acids Metabolites</u> Hydroxybenzoic acids: 2,4-dihydroxybenzoic acid; 2,6-dihydroxybenzoic acid; 2,5-dihydroxybenzoic acid; 3,5-dihydroxybenzoic acid; protocatechuic acid; syringic acid; 4-hydroxybenzoic acid; 3-hydroxybenzoic acid; 4-hydroxyhippuric acid; 3-hydroxyhippuric acid; <u>Gallic acid metabolites</u>: gallic acid; methylgallic acid <u>Ethylgallate metabolites</u>: ethylgallate Hydroxyphenylacetic acids: Phenylacetic acid; 3-hydroxyphenylacetic acid; 2-hydroxyphenylacetic acid; 3,4-dihydroxyphenylacetic acid; homovanillic acid Hydroxycinnamic acids: m-coumaric acid; o-coumaric acid; p-coumaric acid; caffeic acid; ferulic acid; sinapic acid Hydroxyphenylpropionic acids: 3-(4-hydroxyphenyl) propionic acid; 3-(3-hydroxyphenyl) propionic acid; dihydrocaffeic acid Glycinates: Vanilloylglycine; feruloylglycine Hydroxyphenylvalerolactones: DHPV 1; DHPV 2 Other polyphenols: Enterolactone; pyrogallol				
Bub et al., 2001 (11)	31 ± 4 y, Germany	Crossover Feeding Study	6 (Acute)	RP-HPLC (Targeted – malvidin-3-glucoside)	Urine & Plasma	M-3-G quantities: red wine 68mg, dealcoholized red wine 58mg, and red grape juice 117mg
Malvidin-3-glucoside drinks		No significant differences in M-3-G excretion was found between the 3 beverages				

Carkeet et al., 2008 (12)	45 ± 8.4 y, US	Crossover Feeding Study	12 (24 hours)	HPLC-DAD-ion-trap MS (Targeted – anthocyanin (pelargonidin))	Urine	100, 200, & 400 g of pureed strawberries, delivering 15, 30, and 60 mmol anthocyanin, respectively
Strawberry	Pelargonidin 3-glucoside and 3 metabolites of pelargonidin 3-glucoside (monoglucuronides)					
Carrizo et al., 2017 (13)	30.5 y, UK	Feeding Study	31	UPLC-Q-ToF-MS (Untargeted)	Serum	Remove all meat from diet (high red meat intake group), or continue their usual diet (control group)
Red meat	(1) 7-Cyclopentyl-5-(4-phenoxyphenyl)-7H-pyrrolo [2,3‑d] pyrimidin-4-ylamine (2) Glycerophosphocholines [GP01], Monoacylglycero-phosphocholines [GP0105], LysoPC(18:2(9Z,12Z)) (3) Glycochenodeoxycholic acid (4) 2-Aminoethyl 2-[(1E)-1-hexadecen-1-yloxy]-3-hydroxypropyl hydrogen phosphate (5) Docosahexaenoic acid (DHA) (6) Glycerophosphoethanolamines [GP02], 1Z-alkenylglycerophosphoethanolamines [GP0207], PE(P-18:0/0:0) (7) Glycerophosphocholines [GP01], Monoacylglycero-phosphocholines [GP0105], GPCho(14:0/22:4) (8) Sphingomyelin d18:1-C18:0 (9) Phosphatidylserine 18:0-18:1 (10) Glycerophosphoethanolamines [GP02], PE(P-16:0/20:3) 1Z-alkenylglycerophosphoethanolamines [GP0207] (11) Phosphatidylethanolamine alkenyl 18:0-18:2, PE(18:1(9Z)/18:1(9Z) (12) 2-Acetylthiophene (13) n-(tert-butoxycarbonyl)-s-trityl-l-cysteine (14) Glycerophosphocholines [GP01], Monoacylglycero-phosphocholines [GP0105], PC(18:3/0:0) (15) Glycerophosphocholines [GP01], Monoacylglycero-phosphocholines [GP0105], PC(18:2/0:0) (16) Glycerophosphocholines [GP01], Monoacylglycero-phosphocholines [GP0105], PC(20:3/0:0) (17) Glycerophosphocholines [GP01], Monoacylglycero-phosphocholines [GP0105], GPCho(6:0/26:2) (18) Glycerophosphocholines [GP01], Monoacylglycero-phosphocholines [GP0105], GPCho(16:0/17:2) (19) Glycerophosphoethanolamines [GP02], 1Z-alkenylglycerophosphoethanolamines [GP0207]					
Charron et al., 2020 (14)	54.3 ± 9.2, US	Crossover Feeding Study	17 (16 days)	LC-MS (Targeted – glucosinolate metabolites)	Urine & Plasma	Cooked broccoli
Cooked broccoli	Plasma: erucin-cysteineglycine (ER-CG (50% of total)); sulforaphane-cysteineglycine (SF-CG (14%)), sulforaphane (SF (13%)), erucin-cysteine (ER-C (7%)), sulforaphane-glutathione (SF-GSH (6%)), sulforaphane-*N*-acetylcysteine (SF-NAC (5%)), sulforaphane-cysteine (SF-C) and erucin-N-acetylcysteine (ER-NAC (each 2%)), and erucin-glutathione (ER-GSH (<1%))					

		Urine: ER-NAC (39% of total) and SF-NAC (38%), SF-C (11%), ER-C (7%), and SF (4%).				
Chen et al., 2017 (15)	US	Crossover Feeding Study	15	HPLC-ESI-MS/MS (Targeted – caffeine; and it's three major metabolites (paraxanthine, theobromine, and theophylline))	Plasma	100 mg of caffeine by administration of an energy drink and by oral inspiration of a fine powder
Energy drink		Caffeine; paraxanthine				
Cheung et al., 2017 (16)	59.4 ± 4.1 y, UK	Feeding Study	40 (3 weeks)	UHPLC-Q-ToF-MS (Targeted)	Plasma	Chicken, red meat, processed meat, and fish
Fish		Interventional (blood): acetylcarnitine; trimethylamine-N-oxide (TMAO); propionylcarnitine				
Poultry/chicken		Interventional (blood): methylhistidine; acetylcarnitine				
Processed meat (cooked ham)		Interventional (blood): acetylcarnitine				
Red meat		Interventional (blood): acetylcarnitine				
Chiang et al., 2012 (17)	23-65 y, US	Crossover Feeding Study	25 (4 weeks)	GC-FID (Targeted – phospholipid fatty acid profile)	Plasma	Walnut diet incorporating 42.5 g of walnuts per 10.1 mJ 6 times per week (1.8% of energy n-3 fat); fish diet providing 113 g of fatty fish per 10.1 mJ 2 times per week (0.8% of energy n-3 fat), or a control diet (no nuts or fish, 0.4% of energy n-3 fat)
Walnut diet		Oleic acid (18:1n–9); Polyunsaturated fatty acids; Omega-6 fatty acids; Linoleic acid (18:2n–6); y-Linolenic acid (18:3n–6); Dihomo-y-linolenic acid (20:3n–6); Arachidonic acid (20:4n–6); Omega-3 fatty acids; a-Linolenic acid (18:3n–3)				
Fish diet		Linoleic acid (18:2n–6); y-Linolenic acid (18:3n–6); Dihomo-y-linolenic acid (20:3n-6); Eicosatetraenoic acid (20:4n–3); Eicosapentaenoic acid (20:5n–3); Docosapentaenoic acid (22:5n–3); Docosahexaenoic acid (22:6n–3)				
Clarke et al., 2011 (18)	19-50 y, US	Crossover Feeding Study	12 (Acute)	HPLC-MS/MS (Targeted – isothiocyanates (sulforaphane and erucin))	Urine & Plasma	40 grams of fresh broccoli sprouts vs. 6 pills of a broccoli supplement
Broccoli sprouts		Sulforaphane; erucin				
Clarke et al., 2014 (19)	18-65 y, UK	Feeding Study	50 (3 months)	LC–MS/MS (Targeted – free and conjugated forms	Urine	Green tea and vitamin C supplements or a placebo

				of catechins and metabolites)		
Green tea		Epicatechin-O-glucuronide; Epicatechin-O-sulphate; O-me-epicatechin-O-sulphate; Epigallocatechin-O-glucuronide; O-me-epigallocatechin-O-sulphate; O-me-epigallocatechin-O-glucuronide; Quercetin-O-glucuronide; M4-O-sulphate; O-me-M4-O-sulphate; M6/M6'-O-glucuronide; M6/M6'-O-sulphate; Hippuric acid				
Conaway et al., 2000 (20)	34 ± 8.1 y, US	Crossover Feeding Study	12 males	HPLC-MS (Targeted – isothiocyanates (ITCs), hydrolysis products of glucosinolates)	Urine & Plasma	200 g of fresh or steamed broccoli
Broccoli		Isothiocyanates (ITCs); N-*acetyl-L-cysteine conjugate of sulforaphane (SFN-NAC)*				
Cornelis et al., 2018 (21)	<65 y of age, Finland	Crossover Feeding Study	47 (3 months)	UPLC-ESI-MS/MS (Untargeted)	Serum	Refrained from drinking coffee for 1 month, consumed four cups of coffee/day in the 2nd month and eight cups/day in the 3rd month
Coffee		Amino Acid Creatine metabolism: Creatinine; Guanidinoacetate Histidine metabolism: Hydantoin-5-propionic acid; Imidazole lactate Leucine, isoleucine and valine metabolism: Isovalerylcarnitine Methionine, cysteine, SAM and taurine metabolism: Cysteine; Methionine sulfone Polyamine metabolism: 4-Acetamidobutanoate; N-acetylputrescine Tryptophan metabolism: 5-Bromotryptophan; Indolelactate; Kynurenine Tyrosine metabolism: 2-Hydroxyphenylacetate Urea cycle, arginine and proline metabolism: Homoarginine Carbohydrate Aminosugar metabolism: Glucuronate Glycolysis, gluconeogenesis, and pyruvate: 1,5-Anhydroglucitol (1,5-AG) Cofactors and vitamins Nicotinate and nicotinamide metabolism: Trigonelline (N'-methylnicotinate) Energy Oxidative phosphorylation: Phosphate TCA cycle: Citraconate/glutaconate				

	Lipid Diacylglycerol: Linoleoyl-linoleoyl-glycerol (18:2/18:2) Endocannabinoid: Linoleoyl ethanolamide; N-oleoyltaurine; Palmitoyl ethanolamide; Stearoyl ethanolamide Fatty acid metabolism (acyl choline): Arachidonoylcholine; Dihomo-linolenoyl-choline; Docosahexaenoylcholine; Oleoylcholine; Palmitoloelycholine; Palmitoylcholine Glycerolipid metabolism: Glycerol 3-phosphate Phospholipid metabolism: Choline Polyunsaturated fatty acid (n3 and n6): Arachidonate (20:4n6); Docosapentaenoate (n6 DPA) Secondary bile acid metabolism: Glycocholenate sulfate Sphingolipid metabolism: Palmitoyl dihydrosphingomyelin Steroid: 4-Androsten-3alpha,17alpha-diol monosulfate (3); 4-Androsten-3beta,17beta-diol monosulfate (2); Epiandrosterone sulfate; Etiocholanolone glucuronide; Pregn steroid monosulfate Sterol: 3Beta,7alpha-dihydroxy-5-; Campesterol Nucleotide Purine metabolism, (hypo)xanthine/inosine: Urate Purine metabolism, adenine containing: N6-carbamoylthreonyladenosine Purine metabolism, guanine containing: 7-Methylguanine Pyrimidine metabolism, uracil containing: 2'-Deoxyuridine Peptide Dipeptide derivative: N-acetylcarnosine Fibrinogen cleavage peptide: DSGEGDFXAEGGGVR Xenobiotics: Benzoate metabolism: 3-(3-hydroxyphenyl)propionate; 3-(3-hydroxyphenyl)propionate sulfate; 3-Hydroxyhippurate; 3-Methyl catechol sulfate (1); 3-Phenylpropionate; 4-Vinylphenol sulfate; Catechol sulfate; Hippurate; O-methylcatechol sulfate Chemical: 3-Hydroxypyridine sulfate; N-methylpipecolate; Succinimide Food component/plant: Cinnamoylglycine; Dihydroferulic acid; Homostachydrine; N-(2-furoyl)glycine; Pyrraline; Quinate Xanthine metabolism: 1,3,7-Trimethylurate; 1,3-Dimethylurate; 1,7-Dimethylurate; 1-Methylurate; 1-Methylxanthine; 3,7-Dimethylurate; 3-Methylxanthine; 5-Acetylamino-6-amino-3-methyluracil; 7-Methylxanthine; Caffeic acid sulfate; Caffeine; Paraxanthine; Theobromine; Theophylline

Cuparencu et al., 2016 (22)	28.2 ± 7.3 y,	Crossover Feeding Study	16 males (Acute)	UPLC-ESI-Q-ToF-MS (Untargeted)	Urine	Sea buckthorn puree, strawberry puree or an iso-caloric control drink.
Berries (sea buckthorn & strawberry)	Both (Berries): Catechin Sulphate					
Strawberry	4-Hydroxyhippuric acid; Furaneol glucuronide; Pelargonidin glucuronide; P-coumaric acid sulphate; Dihydrokaempferol glucuronide; Furaneol sulphate; Mesifurane sulphate; Leucopelargonidin sulphate; Dihydrokaempferol glucuronide isomer					
Sea buckthorn (berry)	5-Hydroxyindole-3-acetic acid; xi-2,3-Dihydro-2-oxo-1H-indole-3-acetic acid; Hippuric acid; Cyclohexane carboxylic acid glycine; 1-Cyclohexene carboxylic acid glycine; Cyclohexadiene carboxylic acid glycine; N-methyl hippuric acid; Isorhamnetin glucuronide; Pyrocatechol sulphate; Dihydroxycyclohexane carboxylic acid; Protocatechuic acid glucoside					
Cuparencu et al., 2020 (23)	18-70 y, Denmark	Crossover Feeding Study	10 (acute)	UPLC-ESI-q-TOF-MS (Untargeted)	Urine	Beef, pork, and chicken
Poultry	Carnosine; anserine; 3-methylhistidine					
Red meat	Carnosine; anserine					
Cuff et al., 2015 (24)	40-70 y, UK	Feeding Study	162 (12 weeks)	GC-MS (Targeted – alkylresorcinols metabolites (3,5-dihydroxybenzoic acid (DHBA) and 3-(3,5-dihydroxyphenyl)-propanoic acid (DHPPA))	Urine	Diet high in refined cereals (CON) or a diet close to dietary guidelines, with an increased content of whole grain (DG)
Whole grain diet	3,5-dihydroxybenzoic acid (DHBA); 3-(3,5-dihydroxyphenyl)-propanoic acid (DHPPA)					
Davis et al., 2017 (25)	71 ± 4.9 y, Australia	Feeding Study	137 (6 months)	HPLC (Targeted – carotenoids)	Serum	Mediterranean diet (MedDiet) and the habitual diet (HabDiet)
Mediterranean diet	Lycopene; β-carotene; total erythrocyte saturated fat (SFA, %); total erythrocyte MUFA (%); magnesium					
de Oliveira Silva el al., 2020 (26)	20-45 y, Brazil	Crossover Feeding Study	18	HPLC-DAD-FLD (Targeted)	Urine	Soy bean meal and fermented soy bean meal biscuits
Fermented soy bean meal (FSBM) biscuits	Glycitein; genistein; daidzein; dihydrodadzein; ODMA; dihydrogenistein; 6-hydroxy-ODMA; equol					

Del Rio et al., 2010 (27)	26 ± 5 y, Italy	Feeding Study	20 (Acute)	HPLC-MS/MS (Targeted – flavan-3-ol catabolism)	Urine & Plasma	Green tea (containing approximately 400 mmol of flavan-3-ols)
Green tea		Plasma: (-)-epigallocatechin-3-gallate Urine: polyhydroxyphenyl-γ-valerolactones				
Derkach et al., 2017 (28)	31-55 y (majority), US	Crossover Feeding Study	119 (30 days)	UHPLC-MS/MS and GC–MS (Untargeted)	Plasma	High (150 nmol or 3450 mg), medium (100 nmol or 2300 mg), and low (50 nmol or 1150 mg) amounts of sodium
Sodium Diet		High- to low-sodium: - Fatty acid (Isovalerate; Butyrylcarnitine; Valerylcarnitine) - Food component or plant group (4-allylphenol sulfate; methyl glucopyranoside (α plus β); Nacetylalliin; methyl indole-3-acetate; gluconate; homostachydrine; erothioneine) - Benzoate metabolism pathway (4-ethylphenysulfate; 4-Methycatechol sulfate) - γ-glutamyl amino acid group (γ-glutamylvaline; γ-glutamylisoleucine; γ-glutamylleucine; γ-glutamylmethionine; γ-glutamylglutamate; γ-glutamylphenylalanine; γ-glutamyltyrosine) - Methionine metabolism pathway (methionine sulfone; a-ketobutyrate; S-adenosylhomocysteine; N-formylmethionine; N-acetylmethionine; methionine; methionine sulfoxide) - Tryptophan group (indoleacetate; indolebutyrate; methyl indole-3-acetate; tryptophan betaine; indoleacetylglutamine; C-glycosyl tryptophan)				
Díaz-Rubio et al., 2015 (29)	29.5 ± 4.1 y, Spain	Feeding Study	28 (8 weeks)	LC-Q-ToF-MS (Untargeted)	Urine & Plasma	200 mL of ARJ (pomegranate and grape) daily
Pomegranate & Grape juice		Urolithin A glucuronide; Ascorbic acid sulfate; Pyrogallol sulfate				
Dickson et al., 2018 (30)	18 – 38 y, Brazil	Crossover Feeding Study	16 (Acute)	UHPLC-MS (Untargeted)	Urine	Genipap (a native fruit from Amazonia)
Genipap (native fruit from Amazonia)		Dihydroxyhydrocinnamic acid; (1R,6R)-6-hydroxy-2-succinylcyclohexa-2,4-diene-1-carboxylate; hydroxyhydrocinnamic acid; genipic acid; 12-demethylated-8-hydroxygenipinic acid; 3(7)-dehydrogenipinic acid; genipic acid glucuronide; Nonate; 3,4-dihydroxyphenylacetate				
Donovan et al., 1999 (31)	29 ± 3 y, US	Crossover Feeding Study	9 (Acute)	GC-MS (Targeted - catechin and its metabolite 3'-*O*-methylcatechin)	Plasma	120 mL of red wine one day and de-alcoholized red wine
Red wine		3'-*O*-methylcatechin				

Egner et al., 2011 (32)	29-62 y, China	Crossover Feeding Study	48 (7 days)	UPLC-QqQ-ESI-MS (Targeted – glucoraphanin, sulforaphane and sulforaphane thiol conjugates)	Urine	Broccoli sprout-derived beverages: one glucoraphanin-rich (GRR) and the other sulforaphane-rich (SFR)
Sulforaphane-rich (SFR) broccoli sprout		Sulforaphane				
Ellinger et al., 2020 (33)	26.9 ± 4.1 y, Germany	Crossover Feeding Study	12 (acute)	HPLC with coulometric electrode array detection (Targeted)	Plasma	Milk-based cocoa beverage
Milk-based cocoa beverage		Epicatechin				
Erlund et al., 2006 (34)	Int 1: 19-48 y, Int 2: 60 y, Int 3: 19-52 y, Finland	Int-1: Crossover Feeding Study; Int-2: Feeding Study; Int-3: Feeding Study	Int-1: 18 (2 days); Int-2: 40 (8 weeks); Int 3: 80 (6 weeks)	HPLC-ECD (Targeted – flavonol (quercetin))	Plasma	Berries
Berries		Quercetin				
Favari et al., 2020 (35)	18-35 y, Germany	Crossover Feeding Study	10 (acute)	UHPLC-ESI-QqQ-MS/MS (Targeted)	Urine & Plasma	Cranberry juice
Cranberry Juice		5-(Dihydroxyphenyl)-γ-valerolactone-glucuronide (3’,4’,5’); 5-(5’-Hydroxyphenyl)-γ-valerolactone-3’-glucuronide; 5-(3’,5’-Dihydroxyphenyl)-γ-valerolactone; 5-(Dihydroxyphenyl)-γ-valerolactone-sulfate (3’,4’,5’); 5-Phenyl-γ-valerolactone-4’-glucuronide; 5-(3’-Hydroxyphenyl)-γ-valerolactone-4’-glucuronide; 5-Phenyl-γ-valerolactone-sulfate-glucuronide isomer (3’,4’); 5-(4’-Hydroxyphenyl)-γ-valerolactone-3’-glucuronide; 4-Hydroxy-5-(hydroxyphenyl)valeric acid-sulfate (3’/4’) isomer 1; 4-Hydroxy-5-(hydroxyphenyl)valeric acid-glucuronide (3’/4’); 5-(3’,4’-Dihydroxyphenyl)-γ-valerolactone; 5-(5’-Hydroxyphenyl)-γ-valerolactone-3’-sulfate; 5-Phenyl-γ-valerolactone-methoxy-glucuronide isomer (3’,4’); 5-Hydroxyphenyl-γ-valerolactone-methoxy-glucuronide (3’,4’,5’); 5-Phenyl-γ-valerolactone-3’-glucuronide; 4-Hydroxy-5-(hydroxyphenyl)valeric acid-sulfate (3’/4’) isomer 2; 5-(Hydroxyphenyl)-γ-valerolactone-methoxy-sulfate (3’,4’,5’); 5-(Hydroxyphenyl)-γ-valerolactone-sulfate (3’,4’ isomers); 5-Phenyl-γ-valerolactone-4’-sulfate; 5-Phenyl-γ-valerolactone-methoxy-sulfate (3’,4’) isomer 1; 5-Phenyl-γ-valerolactone-3’-sulfate; 5-Phenyl-γ-valerolactone-methoxy-sulfate (3’,4’) isomer 2				

Felberg et al., 2015 (36)	24-42 y, Brazil	Crossover Feeding Study	6 (Acute)	LC–MS (Targeted – isoflavones and chlorogenic acids (CGA))	Urine	Soy (79.7 µmol ISO), coffee (561.2 µmol CGA) and soy–coffee beverage (79.7 µmol ISO and 561.2 µmol CGA)
Soy, coffee and soy–coffee beverages	3-caffeoylquinic acid, 4-caffeoylquinic acid, 5-caffeoylquinic acid, 3,4-dicaffeoylquinic acid, 3,5-dicaffeoylquinic acid and 4,5-dicaffeoylquinic acid, and 15 metabolites: caffeic, ferulic, isoferulic, vanillic, gallic, p-hydroxybenzoic dihydrocaffeic, syringic, sinapic, hippuric, trans-3-hydroxycinnamic, 3,4-dihydroxyphenylacetic, benzoic, 2,4-dihydroxybenzoic and 3-(4-hydroxyphenyl) propionic acids					
Feliciano et al., 2017 (37)	18-35 y, Germany	Crossover Feeding Study	10 males (Acute)	UPLC-Q-ToF-MS (Targeted – (poly)phenols)	Urine & Plasma	Cranberry juices containing 409, 787, 1238, 1534 and 1910 mg total (poly)phenols
Cranberry juice - ((poly)phenols) metabolites	Plasma: 2,4-dihydroxybenzoic acid; 2,5-dihydroxybenzoic acid; 3-hydroxybenzoic acid; (4R)-5-(30-hydroxyphenyl)-γ-valerolactone-40-O-sulfate; 4-methylgallic acid-3-O-sulfate; caffeic acid 3-O-ß-D-glucuronide; caffeic acid 4-O-ß-D-glucuronide; ferulic acid; ferulic acid 4-O-sulfate; ferulic acid 4-O-ß-D-glucuronide; isoferulic acid 3-O-sulfate; quercetin-3-O-ß-D-glucuronide; sinapic acid; syringic acid; vanillic acid-4-O-sulfate Urine: 2,3-dihydrobenzoic acid; 2,4-dihydrobenzoic acid; dihydrocaffeic-3-O-sulfate; ferulic-O-4-sulfate; o-courmaric acid; quercetin-3-O-ß-d-glucuronide; 2,5-dihydroxybenzoic acid; chlorogenic acid; p-coumaric acid; sinapic acid; benzoic acid; isoferulic acid					
Fuchsmann et al., 2020 (38)	– Switzerland	Crossover Feeding Study	11 (acute)	GC-MS	Urine & Plasma	Cheese, Milk, and Soy-Based Drink
Cheese	Plasma: heptan-2-one; undecan-2-one Urine: heptan-4-one; medium-chain fatty acid ester; medium-chain fatty acid ester					
Dairy	Plasma: 3,5-dimethyloctan-2-one Urine: 1-methoxy-2-propyl acetate; medium-chain fatty acid ester; 9-decenoic acid, methyl ester; medium-chain fatty acid ester					
Soy-based drink	Urine: 1,3-octadiene; 2,4-octadiene; 1-octen-3-ol, methyl ether; monoterpene; methoxycyclooctanet; naphthalene derivative; 1-octen-3-ol; acetophenone; 1,1,6-trimethyl-1,2-dihydronaphthalene; coumarin derivative; methyl tetradecanoate					
Milk	3-ethylphenol					
Fujioka et al., 2014 (39)	35.8 ± 12 y, US	Crossover Feeding Study	25 (3 days)	LC-ESI-MS/MS-SRM (Targeted – 3,3'-diindolylmethane	Urine	50 g of either raw 'Jade Cross' Brussels sprouts (high glucobrassicin concentration) or 'Blue Dynasty' cabbage (low glucobrassicin concentration)

				(DIM, a metabolite of indole-3-carbinol))		
Brussels sprouts (high glucobrassicin concentration)	3,3'-diindolylmethane (DIM)					
Fujioka et al., 2016 (40)	31.5 ± 1.5 y, US	Crossover Feeding Study	45 (2 days)	LC-ESI-MS/MS-SRM (Targeted – 3,3'-diindolylmethane (DIM))	Urine	Mixture of Brussels sprouts and/or cabbage, at 1 of 7 discrete dose levels of glucobrassicin ranging from 25 to 500 µmol
Brassica vegetables	3,3'-diindolylmethane (DIM)					
Garcia-Aloy et al., 2020 (41)	28 ± 6 y, Spain	Crossover Feeding Study	11 (acute)	LC-HR-MS (Untargeted)	Urine & Serum	Lentils, chickpeas, and white beans
Lentils	Hydroxyarginine; Oxoarginine; (Epi)catechin sulfate; 4-Hydroxy-5-(dihydroxyphenyl)-valeric acid-O-sulfate; 5-(3',4'-Dihydroxyphenyl)-y-valerolactone-3-O-sulfate; 5-(3',4'-Dihydroxyphenyl)-y-valerolactone-glucuronide; 5-(3',5'-Dihydroxyphenyl)-y-valerolactone-methylglucuronide; Vanillic acid sulfate; Dopamine sulfate					
Chickpeas	Asp-Met; Asp-(i)Leu-Pro / (i)Leu-Asp-Pro; Asp-Ala-(i)Leu / Ala-Asp-(i)Leu; Asp-Gly-Tyr / Gly-Asp-Tyr; Cyclo((i)Leu-Phe); Asp-Thr-Pro / Thr-Asp-Pro; Protocatechuic acid glucoside; Ascorbic acid					
White beans	2-Hydroxyhippuric acid; Hydroxyjasmonic acid; Hydroxydihydrojasmonic acid (I); Hydroxydihydrojasmonic acid (II); Methylcysteine; Pipecolic acid; Trigonelline					
Garcia-Perez et al., 2016 (42)	22-32 y, UK	Crossover Feeding Study	6 (4 days)	^{1}H NMR (Untargeted)	Urine	Grapes
Grapes	Tartaric acid					
Garcia-Perez et al., 2017 (43)	21-65 y, UK	Crossover Feeding Study	19 (Acute)	^{1}H NMR (Untargeted)	Urine	WHO healthy eating guidelines (increase fruits, vegetables, whole grains, and dietary fibre; decrease fats, sugars, and salt)
WHO healthy eating guidelines (increase fruits, vegetables, whole grains, and dietary fibre; decrease fats, sugars, and salt)	Hippurate (a marker of fruit and vegetable consumption); (N-acetyl-)S-methyl-Lcysteine-sulfoxide (cruciferous vegetables); dimethylamine and TMAO (fish); and 1-methylhistidine and 3-methylhistidine (oily fish and chicken)					
Garg et al., 2016 (44)	27.8 ± 6.5 y, Ireland	Crossover Feeding Study	14 (Acute)	^{1}H NMR (Untargeted)	Urine	Minimally processed bran or aleurone

Wheat bran/aleurone		Lactate; alanine; N-acetylaspartate acid; N-acetylaspartylglutamate; betaine				
Gibbons et al., 2015 (45)	27 ± 3 y, Ireland	Feeding Study	10 (Acute)	^{1}H NMR (Untargeted - Panel of biomarkers indicative of sugar-sweetened beverages)	Urine	Sugar-sweetened beverages (330-mL can of cola)
Sugar-sweetened beverage		Citrulline; formate; isocitrate; taurine				
Gibbons et al., 2017 (46)	59 ± 5, England	Feeding Study	50 (3 days for 3 weeks)	^{1}H NMR (Targeted - proline betaine)	Urine	Orange juice
Orange juice		Proline betaine				
Gómez-Juaristi et al., 2019 (47)	26.67 ± 3.21 y, Spain	Crossover Feeding Study	13 (acute)	HPLC-ESI-QToF-MS (Targeted)	Urine & Plasma	Two soluble cocoa products: a conventional and a flavanol-rich product
Two soluble cocoa products: a conventional and a flavanol-rich product		Plasma Flavanols: Epicatechin-3'-glucuronide; Epicatechin-3'-sulfate; Epicatechin-methoxy-sulfate (isomer 1); Epicatechin-methoxy-sulfate (isomer 2) Phenyl-γ-Valerolactone (PVL) derivatives: 5-(30,40-Dihydroxyphenyl)-γ-valerolactone (DHPVL); 5-(40-Hydroxyphenyl)-γ-valerolactone-30-glucuronide (HPVL-3'-glucuronide); 5-(Hydroxyphenyl)-γ-valerolactone-sulfate (HPVL-sulfate); 5-Phenyl-γ-valerolactone-methoxy-glucuronide (PVL-methoxy-glucuronide); 5-Phenyl-γ-valerolactone-3'-sulfate (PVL-30-sulfate) Phenylvaleric acid derivatives: 4-Hydroxy-5-(hydroxyphenyl)valeric acid-sulfate (HHPVA-sulfate) Other microbial metabolites: 3,4-Dihydroxyphenylpropionic acid; 3-Methoxy-4-hydroxyphenylpropionic acid Urine Flavanols: Epicatechin-3'-glucuronide; Epicatechin-3'-methoxy-glucuronide; Epicatechin-3'-sulfate; Epicatechin-methoxy-sulfate (isomer 1); Epicatechin-methoxy-sulfate (isomer 2); Epicatechin-methoxy-sulfate (isomer 3) Phenyl-γ-Valerolactone (PVL) derivatives: 5-(30,40-Dihydroxyphenyl)-γ-valerolactone (DHPVL); 5-(30-Hydroxyphenyl)-γ-valerolactone-40-glucuronide (HPVL-4'-glucuronide); 5-(40-Hydroxyphenyl)-γ-valerolactone-30-glucuronide (HPVL-3'-glucuronide); 5-(Hydroxyphenyl)-γ-valerolactone-sulfate (HPVL-sulfate); 5-Phenyl-γ-valerolactone-methoxy-glucuronide (PVL-methoxy-glucuronide); 5-Phenyl-γ-valerolactone-methoxy-sulfate (PVL-methoxy-sulfate); 5-(30-Hydroxyphenyl)-γ-valerolactone (HPVL); 5-Phenyl-γ-valerolactone-30-glucuronide (PVL-30-glucuronide); 5-Phenyl-γ-valerolactone-3'-sulfate (PVL-30-sulfate)				

		Phenylvaleric acid derivatives: 4-Hydroxy-5-(3',4'-dihydroxyphenyl)valeric acid (HDHPVA); 4-Hydroxy-5-(hydroxyphenyl)valeric acid-glucuronide (HHPVA-glucuronide); 4-Hydroxy-5-(hydroxyphenyl)valeric acid-sulfate (HHPVA-sulfate) Other microbial metabolites: 3,4-Dihydroxyphenylpropionic acid; 3-Methoxy-4-hydroxyphenylpropionic acid; 3-Hydroxyphenylpropionic acid; 3,4-Dihydroxyphenylacetic acid; 3-Methoxy-4-hydroxyphenylacetic acid; 3-Hydroxyphenylacetic acid; Ferulic acid; Isoferulic acid; 3,4-Dihydroxybenzoic acid; 4-Hydroxyhippuric acid; 3-Hydroxyhippuric acid; Hydroxybenzoic acid				
Gouado et al., 2007 (48)	22-27 y, Cameroon	Feeding Study	14 (Acute)	HPLC (Targeted – carotenoids provitamin A, lycopene and lutein)	Serum	Mango and papaya in three types of meal treatments (juice, fresh and dried fruit)
Mango		a-Carotene; b-Carotene; Cryptoxanthin; Zeaxanthin				
Papaya		a-Carotene; Lycopene; Cryptoxanthin; Zeaxanthin				
Gu et al., 2013 (49)	17.8-52.0 y, China	Feeding Study	75	UPLC-Q-ToF-MS and GC-ToF-MS (Untargeted)	Serum	Very low carbohydrate diet
Very low carbohydrate diet		4 weeks of VLCD: 2-aminobutyrate, proline, ornithine, tryptophan, methionine, threonine, theanine, cis-11,14-eicosadienoate, cis-11,14,17-eicosatrienoate, arachidonate, acetyl-carnitine, 3-hydroxybutyrate, p-cresol, 3-aminophenol, threitol, urea, glutamate, alanine, cysteine, ribose, mannose, succinate, cis-5,8,11,14,17-eicosapentaenoate, palmitoleate, stearate, pentadecanoate, fumarate, 2-hydroxybutyrate, 2,3-dihydroxybutanoate, adenine, and nicotinamide 8 weeks of VLCD: proline, ornithine, tryptophan, threonine, p-cresol, threitol, urea, glutamate, alanine, ribose, mannose, succinate, palmitoleate, elaidate, stearate, oleamide, xanthine, 4-hydroxy-3-methoxymandelate, nicotinamide, chenodeoxycholate, glyceraldehyde, and glycerol				
Gürdeniz et al., 2016 (50)	18−60 y, Denmark	Crossover Feeding Study	18 (Acute)	UPLC-Q-ToF-MS (Untargeted)	Urine & Plasma	Four different test beverages: strong, regular, and nonalcoholic beers and a soft drink
Beer		Blood & Urine: N-methyl tyramine sulfate; sum of iso-α-acids; Iso-cohumulone Urine only: Tricyclohumols				
Hagen et al., 2020 (51)	18-69 y, Norway	Feeding Study	62 (8 weeks)	LC or GC combined with MS/MS (Targeted)	Urine & Serum	Cod and salmon
Cod		Serum: trimethylamine *N*-oxide (TMAO); creatine; 1-methylhistidine Urine: trimethylamine *N*-oxide (TMAO); creatine; 1-methylhistidine				

Salmon		Serum: 1-methylhistidine; creatine Urine: 1-methylhistidine; creatine				
Halder et al., 2018 (52)	23.7 ± 2 y, Singapore	Crossover Feeding Study	17 Chinese males (Acute)	UHPLC-MS/MS (Targeted – phenolic acids)	Plasma	Curry meal containing 0 g, 6 g, and 12 g of mixed spices
Spices		Cinnamic acid; phenylacetic acid				
Hanhineva et al., 2013 (53)	57 ± 9 y, Finland	Crossover Feeding Study	12 (Acute)	LC-Q-ToF-MS (Untargeted)	Plasma	100% whole-grain sourdough rye bread or white wheat bread enriched with native unprocessed rye bran or bioprocessed rye bran
Whole-wheat rye or rye bran		Hydroxy-N-(2-hydroxyphenyl) acetamide (HHPAA); N-(2-hydroxyphenyl) acetamide (HPAA)				
Hanhineva et al., 2015 (54)	40–70 y, Finland	Feeding Study	106 (12 weeks)	UHPLC-Q-ToF-MS (Untargeted)	Plasma	1) whole-grain products, fatty fish, and bilberries (HD); 2) a whole-grain–enriched diet with the same grain products as in the HD intervention but with no change in fish or berry consumption; and 3) refined-wheat breads and restrictions on fish and berries (control diet)
Fish		3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF)				
Berries		3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF); Hippuric acid				
Whole grain		Alkenylresorcinol (AenR) 21:1-Gln; alkylresorcinol (AR) 19:0-Gln; Pipecolic acid betaine; y-Butyrobetaine				
Healthy diet (HD)		Pyrocatechol sulfate; Hippuric acid; CMPF; 3-Carboxy-4-methyl-5-pentyl-2-furanpropionic acid; EPA; Nonadecyl-benzenediol glucuronide (AR 19:0-Gln); Heneicosenyl-benzenediol-glucuronide (AenR 21:1-Gln); lysophosphatidylcholine (LPC) (20:5) minor isomer; lysophosphatidylcholine LPC (20:5)				
Whole-grain–enriched diet		Nonadecyl-benzenediol glucuronide (AR 19:0-Gln); Heneicosenyl-benzenediol-glucuronide (AenR 21:1-Gln)				
Hauder et al., 2011 (55)	50-82 y, Germany	Feeding Study	76 (4 weeks)	HPLC-MS/MS (Targeted – sulforaphane (SFN) and indole-3-carbinol metabolites)	Urine & Plasma	200 g of a daily dose of regular blanched broccoli, selenium-fertilized blanched broccoli, or placebo
Regular broccoli		Plasma: Sulforaphane; sulforaphane cysteinylglycine (SFN-Cys-Gly); sulforaphane cysteine (SFN-Cys); sulforaphane N-acetylcysteine (SFN-NAC)				

				Urine: Sulforaphane; sulforaphane cysteinylglycine (SFN-Cys-Gly); sulforaphane cysteine (SFN-Cys); sulforaphane N-acetylcysteine (SFN-NAC)		
Selenium-fertilized broccoli				Plasma: Sulforaphane; sulforaphane cysteinylglycine (SFN-Cys-Gly); sulforaphane N-acetylcysteine (SFN-NAC) Urine: Sulforaphane; sulforaphane cysteinylglycine (SFN-Cys-Gly); sulforaphane cysteine (SFN-Cys); sulforaphane N-acetylcysteine (SFN-NAC)		
Heinzmann et al., 2010 (56)	Study 1: 28-45 y, Study 2: 24-46 y, UK	Crossover Feeding Study	Study 1: 8 Study 2: 6 (Acute)	^{1}H NMR (Untargeted and targeted – proline betaine)	Urine	Study 1: mixed-fruit meal (apple, orange, grapes, and grapefruit) Study 2: orange juice
Mixed fruits				Proline betaine; tartaric acid; hippuric acid		
Orange juice				Proline betaine		
Henning et al., 2010 (57)	29 ± 6.3 y, US	Feeding Study	21 females (3 weeks)	HPLC–MS/MS (Targeted – pelargonidin-glucuronide, urolithin A-glucuronide, and 2,5-dimethyl-4-hydroxy-3-[2H]furanoneglucuronide)	Urine	Frozen strawberries (250 g) administered daily
Strawberry				Pelargonidin (Pg)-glucuronide; urolithin A (UA)-glucuronide; 2,5-Dimethyl-4-hydroxy-3-[^{2}H]furanone (DMHF)-glucuronide		
Hernández-Alonso et al., 2017 (58)	55.3 ± 2 y, Spain	Crossover Feeding Study	39 prediabetic subjects (4 months)	^{1}H NMR (Targeted – metabolites related with gut microbiota metabolism)	Urine	Pistachio-supplemented diet (PD, 50% carbohydrates, 33% fat, including 57 g/d of pistachios daily) and a control diet (CD, 55% carbohydrates, 30% fat)
Pistachio-supplemented diet				Hippurate; p-cresol sulfate; dimethylamine; cis-aconitate (intermediate of the tricarboxylic acid (TCA)); creatinine; trimethylamine N-oxide (TMAO); N-methyl-trans-4-hydroxy-L-proline		
Hernández-Alonso et al., 2019 (59)	30-60 y, Spain	Feeding Study	102 (6 months)	GC-Q-ToF-MS and HPLC-Q-ToF-MS, and ^{1}H NMR (Targeted)	Plasma	Low-glycemic index (LGI) diet, high-glycemic index (HGI) diet, and a low-fat (LF) diet
Low-glycemic index (LGI)				Serine; tyrosine; glycine; leucine; valine; C32:1 SM; C42:3 SM; C20:3 lysophosphatidylcholine (LPC); C18:2 LPC; C32:1; C34:2e; C36:2e; C36:5e; C38:5; C40:6 PC		
Hodgson et al., 2000 (60)	61.8 ± 2.7 y, Australia	Crossover Feeding Study	10 (4 weeks)	GC-MS (Targeted – gallic acid metabolites)	Urine	5 cups per day of black tea

Black tea		4-O-methylgallic acid, 3-O-methylgallic acid, and 3,4-O-dimethylgallic acid				
Hodgson et al., 2013 (61)	22 ± 8 y, UK	Feeding Study	27 males (1 week)	LC–MS/MS and GC–MS (Untargeted and targeted – catecholamines)	Plasma	Green tea extracts (1200 mg catechins, 240 mg caffeine/day) or placebo drinks
Green tea		Urea; Cholesterylester C18:1; Cholesterylester C18:2; Glycerol (lipid fraction); Docosahexaenoic acid (C22:cis[4,7,10,13,16,19]6); Isopalmitic acid (C16:0); TAG (C16:0,C16:1); TAG (C16:0,C18:2); TAG (C18:2,C18:2); TAG (C18:2,C18:3); Phosphatidylcholine(C18:0,C22:6); Sphingomyelin (d18:1,C16:0); 3-hydroxybutyrate; Citrate; 5-hydroxy-3-indoleacetic acid (5-HIAA); Homovanillic acid (HVA); Caffeine; Glycerol (polar fraction); Hippuric acid				
Housley et al., 2018 (62)	36 ± 5.3 y, US	Crossover Feeding Study	10 (Acute)	HPLC–MS/MS (Untargeted)	Plasma	Fresh broccoli sprouts (a rich dietary source of bioactive sulforaphane)
Broccoli sprouts		Glutathione; Cysteine; Glutamine; Dehydroepiandrosterone (DHEA); Deoxyuridine monophosphate (dUMP); FA 14:0; FA 14:1; FA 16:0; FA 16:1; FA 18:0; FA 18:1				
Hövelmann et al., 2019 (63)	23-32 y, Germany	Feeding Study	7 (Acute)	LC-ESI-MS (Untargeted)	Urine	Tomato juice (1 L) + standardized breakfast
Tomato juice		Hydroxyesculeogenin B isomer; Esculeogenin B sulfonate; Esculeogenin B isomer; Tomatidine				
Hutchins et al., 2000 (64)	66.9 ± 8.2 y, US	Crossover Feeding Study	31 females (7 week)	GC-MS (Targeted – dietary estrogens, such as lignans)	Urine	Habitual diets plus 0, 5, or 10 grams of ground flaxseed per day
Flaxseed		Enterodiol; enterolactone; total lignan				
Ibero-Baraibar et al., 2016 (65)	57.5 ± 5.3 y, Spain	Feeding Study	47 (4 weeks)	HPLC-ToF-MS (Untargeted)	Urine	Ready-to-eat meals containing a cocoa extract (with 1.4 g of cocoa extract (645 mg polyphenols)) vs. meals without cocoa
Cocoa diet		Theobromine metabolism (3-Methylxanthine; 3-Methyluric acid) Food processing (L-Beta-aspartyl-L-phenylalanine) Flavonoid metabolism (2,5,7,3',4'-Pentahydroxyflavanone 5-O-glucoside; 7,4'-Dimethoxy-6-C-methylflavanone) Catecholamine metabolism (3-Methoxy-4-hydroxyphenylglycol (MHPG) sulphate) Endogenous metabolism (Uridine monophosphate)				
Jahns et al., 2014 (66)	32.1 ± 2.5 y, US	Feeding Study	29 (8 weeks)	HPLC (Targeted – carotenoids)	Plasma	Fruits and vegetables intake
Fruits and vegetables		Total carotenoid				
Jin et al., 2011 (67)	44.6 ± 13.3 y, UK	Crossover Feeding Study	20 (Acute)	LC–MS and GC–MS (Targeted – anthocyanins)	Urine	20% blackcurrant juice drink (250 ml)

Blackcurrant juice		Anthocyanins: Delphinidin-3-rutinoside; Cyanidin-3-rutinoside; Delphinidin-3-glucoside[1]				
Johansson-Persson et al., 2013 (68)	(58.6 ± 1 y), Sweden	Crossover Feeding Study	25 (5 weeks)	LC-Q-ToF/MS (Untargeted)	Plasma	High dietary fiber intake (consisted of oat bran, rye bran, and sugar beet fiber)
High fiber diet		2,6-dihydroxybenzoic acid (DHBA); 2-aminophenol sulfate				
Kempf et al., 2010 (69)	54 ± 9 y, Germany	Feeding Study	47 (2 months)	HPLC and GC–MS (Targeted – polyphenols and methylxanthines)	Plasma	Coffee (refrained for 1 month, then 4 and 8 cups per day in the 2nd and 3rd months, respectively)
Coffee		Caffeine; paraxanthine; theobromine; theophylline; caffeic acid; dihydrocaffeic acid; Coumaric acid; dihydro-3-coumaric acid; ferulic acid; isoferulic acid; dihydroferulic acid; dihydroisoferulic acid; dimethoxycinnamic acid; 3-(3,4-Dimethoxyphenyl)-propionic acid				
Khakimov et al., 2016 (70)	20-66 y, Denmark	Feeding Study	145 (6 months)	GC-MS (Untargeted)	Plasma	New Nordic Diet (NND) or Average Danish Diet (ADD)
New Nordic Diet		glycine; 3-hydroxybutanoic acid; 2,3-dihydroxybutanoic acid; erythritol, 2-hydroxybenzoic acid; aspartic acid; threonic acid; pyrophosphate; xylitol; 2,5 diisopropylnaphthalene; N-acetylaspartic acid; 3-(2,5-dimethoxyphenyl)propionic acid; palmitoleic acid				
Average Danish Diet		lactic acid; oxalic acid; alanine; threonine; phenylalanine; diethyl phthalate; 2,6-diisopropylnaphthalene (2,6-DIPN); citric acid, cholesterol				
Khymenets et al., 2015 (71)	25-55 y, Spain	Crossover Feeding Study	31 (Acute and 15 days)	HPLC-Q-ToF-MS (Untargeted)	Urine	Grape skin extract/polyphenol (drink)
Grape-derived extract/polyphenols (functional beverage)		Unique for Acute: tyrosine Unique for Sustained beverage: two 4-hydroxy-5-(dihydroxyphenyl)-valeric acid glucuronides; 4-hydroxy-5-(dihydroxyphenyl)-valeric acid sulphate; two 5-(dihydroxyphenyl)-y-valerolactone glucuronide; 5-(hydroxyphenyl)-y-valerolactone glucuronide; 5-(hydroxy-methoxy-phenyl)-y-valerolactone glucuronide Common: 4'-hydroxyhippuric acid; two hydroxy-dimethoxybenzoic acid glucuronide; vanillic acid glucuronide; vanilloylglycine; (epi)catechin glucuronide; two (epi)catechin sulphates; dihydrosinapic acid glucuronide				
Kremer et al., 2018 (72)	22-27 y, Germany	Feeding Study	10 (2 days)	HPLC-ESI-MS/MS (Targeted – niacin metabolites)	Urine	500 mL coffee beverage
Coffee		nicotinic acid (NA); nicotinamide (NAM); N^1-methyl-nicotinamide (NMNAM), and N^1-methyl-2-pyridon-5-carboxamide (2-Py)				

Kristensen et al., 2007 (73)	26-30 y, Denmark	Crossover Feeding Study	6 (2 days)	HPLC-MS (Targeted – isothiocyanates)	Urine	A basic diet supplemented with 80 or 350 g of mixed cruciferous vegetables
Cruciferous vegetables		Isothiocyanates (ITCs)				
Kuntz et al., 2015 (74)	23-27 y, Germany	Crossover Feeding Study	10	UPLC-MS (Targeted – anthocyanins)	Urine & Plasma	0.33 litres of juice or smoothie (made from an eighty/twenty mixture of red grapes and blueberries)
Berries & grapes juice or smoothie		Delphinidin-3-glucoside; cyanidin-3-glucoside; petunidin-3-glucoside; malvidin-3-glucoside; peonidin-3-glucoside; malvidin-3-glucuronides; peonidin-3-glucuronides; 3,4-dihydroxybenzoic acid				
Lacalle-Bergeron et al., 2020 (75)	25.0 ± 2.8 y, Spain	Crossover Feeding Study	30 (acute)	UHPLC-IMS-HRMS (Untargeted)	Plasma	Orange
Orange		Synephrime hydrogen sulfate; N-methyltyramine hydrogen sulfate; Hesperitin hydrogen sulfate; N-methyl-proline; Betonicine; Stachydrine				
Landberg et al., 2009 (76)	30.6 ± 10.3 y, Sweden	Crossover Feeding Study	16 (1 week)	Urine: HPLC-MS Plasma: GC-MS (Targeted – alkylresorcinols and their metabolites)	Urine & Plasma	Rye bran flakes containing 11, 22, or 44 mg total alkylresorcinols (AR)
Rye bran flakes		Alkylresorcinols				
Langer et al., 2018 (77)	23 ± 3 y, UK	Crossover Feeding Study	9 (Acute)	LC–ESI-Q-Orbitrap-MS (Untargeted)	Urine & Plasma	250 g of fresh blueberries either as the whole fruit or after juicing
Blueberries		<u>Urine</u> - <u>Whole fruit</u>: Ferulic acid 4-sulphate; Caffeic acid 4-sulphate; Ferulic acid 4-sulphate; Abscisic acid; 2'-Methoxy-3-(2,4-dihydroxyphenyl)-1,2-propanediol 4'-glucoside <u>Plasma</u> - <u>Juice</u>: N-(7-Sulfanylheptanoyl)-L-threonine - <u>Whole fruit</u>: Deoxynivalenol 3-glucoside				
Lankinen et al., 2011 (78)	58.7 ± 5.8 y,	Crossover Feeding Study	39 females (8 weeks)	UPLC-ESI-MS, GC and UPLC (Untargeted)	Plasma	High-fiber rye bread (RB) or white wheat bread (WB)
High-fiber rye bread		Ribitol; Ribonic acid; 1H-Indole-3-acetic acid				

Lappi et al., 2013 (79)	35-65 y, Finland	Crossover Feeding Study	15 (Acute)	GC-MS (Targeted – phenolic acids and their metabolites)	Urine	White wheat breads fortified with bioprocessed or native rye bran, and wholegrain rye bread and white wheat bread as controls
White wheat bread fortified with rye bran	Ferulic acid; Sinapic acid					
Lee et al., 2012 (80)	23.3 ± 2.8 y, US	Crossover Feeding Study	16 (Acute)	UHPLC-ESI-Q-ToF-MS/MS (Targeted – quercetin metabolites)	Plasma	(1) Apple peel powder-enriched applesauce (2) Onion powder-enriched applesauce
Apple sauce (apple peal or onion powder)	Quercetin sulfate; quercetin glucuronide; quercetin diglucuronide					
Lennerz et al., 2015 (81)	(22-34 y), US	Crossover Feeding Study	14 overweight subjects (Acute)	LC–MS/MS (Targeted – benzoate and hippurate)	Plasma	Sodium benzoate (a widely used food preservative)
Sodium benzoate (a widely used food preservative)	Benzoate; hippurate; anthranilic acid (tryptophan metabolite); acetylglycine					
Li et al., 2001 (82)	31-35 y, US	Crossover Feeding Study	5 (Acute)	HPLC-ESI-MS (Targeted – polyphenols glucuronides and sulphates)	Urine	Green tea
Green tea	Monoglucuronides and monosulfates of (-)-epigallocatechin (EGC) and (-)-epicatechin; *O*-methyl-EGC-*O*-glucuronides and -*O*-sulfates and *O*-methyl-epicatechin-*O*-sulfates; (-)-5-(3′,4′,5′-Trihydroxyphenyl)-γ-valero- lactone (M4); (-)-5-(3′,4′-dihydroxyphenyl)-γ-valerolactone (M6); the ring-fission metabolites of EGC and (-)-epicatechin					
Li et al., 2018 (83)	10 ± 0.8 y, US	Feeding Study	38 (4 weeks)	UPLC–MS/MS (Untargeted)	Plasma	(1) no navy beans or rice bran (control), (2) 17.5 g/day cooked navy beans, (3) 15 g/day heat-stabilized rice bran, or (4) a combination of 9 g/day navy bean and 8 g/day rice bran
Navy bean	Trigonelline; ferulic acid 4-sulfate; Pipecolate; S-methylcysteine, and *S*-methylcysteine sulfoxide					
Rice bran	Methionine sulfone; alpha-hydroxycaproate; linoleoyllinolenoyl-glycerol; palmitoyl-linolenoylglycerol; pyridoxal; 2-hydroxyhippurate; salicylate; gamma-glutamylglutamate; gamma-glutamylthreonine; hypoxanthine; dihydroorotate					
Li et al., 2020 (84)	28.7 ± 6.5 y, UK	Crossover Feeding Study	15 (2 weeks)	UHPLC-QqQ-MS (Targeted)	Urine	Blood orange juice

Blood orange juice		Hesperetin-3'-glucuronide; hesperetin-7-glucuronide				
Liu et al., 2017 (85)	21-29 y, US	Crossover Feeding Study	17 (3 days)	UHPLC-Q-Orbitrap-HRMS (Untargeted)	Plasma	Cranberry juice or apple juice
Cranberry juice		**Exogenous metabolites:** quinic acid; vanilloloside; catechol sulfate; 3,4-dihydroxyphenyl ethanol sulfate; coumaric acid sulfate; ferulic acid sulfate; 5-(trihydroxphenyl)-gamma-valerolactone; 3-(hydroxyphenyl)proponic acid; hydroxyphenylacetic acid; trihydroxybenzoic acid **Endogenous metabolites:** citramalic acid; aconitic acid, hydroxyoctadecanoic acid; hippuric acid; 2-hydroxyhippuric acid; Vanilloylglycine; 4-acetamido-2-aminobutanoic acid; dihydroxyquinoline; glycerol 3-phosphate				
Llorach et al., 2009 (86)	18–50 y, Spain	Crossover Feeding Study	10 (3 days)	HPLC-Q-ToF-MS (Untargeted)	Urine	(a) 40 g of cocoa powder with 250 mL of water (b) 40 g of cocoa powder with 250 mL of milk and (c) 250 mL of milk as a control
Cocoa		3-methyluric acid; 3-methylxanthine; 3'-methoxy-4'-hydroxyphenylvalerolactone glucuronide; 3,5-Diethyl-2-Methylpyrazine; 3,7-dimethyluric acid; 4-hydroxy-5-(3,4-dihydroxyphenyl)-valeric acid; 5-(3',4'-dihydroxyphenyl)-γ-valerolactone glucuronide; 6-amino-5-[N-methylformylamino]-1-methyluracil (AMMU); 7-methyluric acid; 7-methylxanthine; Caffeine; Cyclo(Ser-Tyr); Cyclo(Pro-Pro); Epicatechin-O-sulfate; hydroxyacetophenone; Hydroxynicotinic acid; O-Methylepicatechin; Theobromine; Tyrosine; Trigonelline; Vanillic acid; Vanilloylglycine				
Lloyd et al., 2011 (87)	44.2 ± 18.2 UK	Crossover Feeding Study	24 (Acute)	FI-ESI-MS and GC-ToF-MS (Untargeted and targeted)	Urine	(a) Oily fish (60 g smoked salmon trimmings) (b) A cruciferous vegetable (200 g steamed broccoli florets) (c) A berry fruit (200 g raspberries) (d) A whole-grain wheat cereal (37.5 g; 2 biscuits) with 125 mL ultra-high temperature–treated semi-skimmed milk
Salmon		Anserine; 1- and 3-Methylhistidine; Trimethylamine-N-oxide (TMAO)				
Raspberry		Caffeoyl sulfate; Ascorbate; Methyl-epicatechin sulfate; 3-Hydroxyhippuric acid; Naringenin glucuronide				
Broccoli		Ascorbate; Tetronic acids; Tetronic acid derivative; L-Xylonate/L-lyxonate; Threitol/erythritol; Naringenin glucuronide; Hesperitin glucuronide plus other glucuronides				
Macdonald et al., 2009 (88)	Intervention: 59.3 ± 2.1 Control:	Feeding Study	226 (2 years)	HPLC (Targeted – vitamins (E, C, K), carotenoids, folate and homocysteine)	Plasma & Serum	300 g additional fruit & vegetables per day, placebo, or potassium citrate

	59.8 ± 2.2, UK					
Fruit & vegetables		α-Tocopherol; α-Tocopherol (cholesterol); γ-Tocopherol; β-Carotene; β-Cryptoxanthin				
Madrid-Gambin et al., 2016 (89)	25-44 y, Spain	Crossover Feeding Study	10 males (28 days)	^{1}H NMR (Untargeted)	Urine	Coffee extract beverage (CEB: 223 mg/100 ml of CGAs) or a control beverage with equal caffeine dose
Coffee extract beverage		Acute intervention: Trigonelline; 2-Furoylglycine; Citric acid; Succinic acid; 3-Methyl-2-oxovaleric acid; Isobutyric acid Sustained intervention: Trigonelline; Hippuric acid; 3-(3-Hydroxyphenyl)-3-hydroxypropionic acid (HPHPA); 3-Hydroxyhippuric acid				
Madrid-Gambin et al., 2018 (90)	19-37 y, Spain	Crossover Feeding Study	11 (2 days)	^{1}H NMR (Untargeted)	Urine & Serum	Pulses (i.e., white beans, chickpeas, and lentils)
Pulses/Legumes		Trigonelline; 3-methylhistidine; dimethylglycine; trimethylamine; lysine				
Madrid-Gambin et al., 2019 (91)	25-44 y, Spain	Crossover Feeding Study	10 (28 days)	^{1}H NMR (Untargeted)	Urine	Functional high-catechin tea (HCT, 350 mL containing 187 mg/100 mL of catechins) or control (containing caffeine similar to HCT group)
Catechin tea		Acute intake: 3-methyl-2-oxovalerate; theanine; gallate; epicatechin (EC); epigallocatechin (EGC) Sustained intake: 2-hydroxyisobutyrate; succinate; pyrogallol sulfate				
Mahale et al., 2018 (92)	26-35 y, UK	Feeding Study	4 (Acute)	LC-ESI-MS/MS (Targeted – curcuminoids)	Plasma	Turmeric-containing food consisting of soup, a sandwich, and an oat bar
Turmeric-containing food		Curcumin; Curcumin glucuronide; Demethoxycurcumin glucuronide; Curcumin sulfate				
Martin et al., 2009 (93)	18-35 y, Netherlands	Feeding Study	30 subjects classified as low and high anxiety (14 days)	^{1}H NMR (Untargeted), GC-MS, and LC-MS/MS (Targeted)	Urine	Dark chocolate
Dark chocolate		4-Hydroxyphenylacetate; Adrenaline; Asparagine; Corticosterone; Cortisol; Cystine; Glucose-6-phosphate; Normetanephrine; Phenylacetylglutamine; *p*-Cresol sulfate; Threonic acid				
Martínez-López et al., 2014 (94)	18-45 y, Spain	Crossover Feeding Study	13 (Acute)	HPLC-DAD (plasma) LC-QTOF and quantified by LC-DAD (urine)	Urine & Plasma	Methylxanthines in two soluble cocoa products, one containing methylxanthines naturally occurring in cocoa (CC) and a

				(Targeted – methylxanthines)		product enriched in methylxanthines (CC-MX)
Cocoa product enriched in methylxanthines		Caffeine; Paraxanthine; Theobromine; Theophylline; 1-Methylxanthine; 3-Methylxanthine; 7-Methylxanthine; 1-Methyluric acid; 1,3-Methyluric acid; 1,7-Methyluric acid; 3,7-Methyluric acid; 1,3,7-Methyluric acid				
May et al., 2013 (95)	29 ± 4.9 y, US	Feeding Study	10 (2 weeks)	LC-ESI-LTQ-Orbitrap-MS (Untargeted)	Urine	A diet rich in cruciferous vegetables, citrus and soy (F&V), and a fruit- and vegetable-free (basal) diet
Fruits & vegetables diet		Sulforaphane; Proline betaine; Hippuric acid; Genistein; Daidzein; Equol; Glycitein; O-Desmethylangolensin; 7C-aglycone; Enterolactone; Trigonelline; Isovalerylglycine; Valerylglycine; Hydroxyphenylacetylglycine; Nicotinuric acid; Adenosine; 5-methylcytidine				
Basal diet		Iopterin; D-Biopterin; Dyspropterin; Orinapterin; Primapterin; Sepiapterin; N1-Methyl-4-pyridone-3-Carboxamide; Porphobilinogen; Riboflavin; 1-Pyrroline-4-hydroxy-2-carboxylate; N-Acryloylglycine; Pyroglutamic acid; Pyrrolidonecarboxylic acid; 2-Methylbutyroylcarnitine; Isovalerylcarnitine; 2,6 Dimethylheptanoyl carnitine; L-Carnitine; L-Acetylcarnitine; L-Kynurenine; Kynurenic acid; Xanthurenic acid; 3-Hydroxyhippuric acid; Salicyluric acid; Argininosuccinic acid; L-Histidine				
McKeown et al., 2016 (96)	18-40 y, US	Crossover Feeding Study	19 (1 week)	GS-MS (plasma) & HPLC (urine) (Targeted - Alkylresorcinols metabolites [3,5-dihydroxybenzoic acid and 3-(3,5-dihydroxyphenyl)-propanoic acid])	Urine & Plasma	Whole grain wheat and rye consumption
Whole grain		Urine: Alkylresorcinols (including its metabolites: 3,5-dihydroxybenzoic acid (DHBA) ; 3,5-dihydroxyphenylpropanoic acid (DHPPA) Serum: Alkylresorcinols				
McNamara et al., 2020 (97)	34 ± 12 (acute) 29 ± 10 (short-term), Ireland	Feeding Study	17 (acute) 32 (short-term)	LC-MS (Untargeted)	Urine	Apple
Apple		NMR: 3-hydroxyisovalerate; acetylsalicylate; glycine; xylose LC-MS: Ethyl 2-aminobenzoate; Pro Leu; 1-(Malonylamino)cyclopropanecarboxylic acid; Epicatechin sulfate; Dopachrome o-semiquinone; 4-Pyridoxic acid; L-Suberyl carnitine; D-Xylono-1,5 lactone; Glucodistylin;				

Mennen et al., 2006 (98)	35-60 y, France	Feeding Study	53 (2 days)	HPLC–ESI–MS/MS (Targeted – 13 polyphenols and metabolites)	Urine	Polyphenol-rich foods

Apple	Phloretin
Red fruits	*m*-coumaric acid; kaempferol
Grapefruit	naringenin
Orange	caffeic acid; hesperetin
Citrus fruit	caffeic acid; naringenin; hesperetin
Fruits	Kaempferol; naringenin; phloretin
Fruit juices	gallic acid; 4-*O*-methylgallic acid; isorhamnetin; naringenin; hesperetin
Fruits and/or fruit juices	*m*-coumaric acid; kaempferol; hesperetin; naringenin; phloretin
Vegetables	gallic acid
Wine	gallic acid; 4-*O*-methylgallic acid
Coffee	caffeic acid; chlorogenic acid
Black tea	chlorogenic acid; m-coumaric acid; gallic acid; 4-O-methylgallic acid; kaempferol

Meuronena et al., 2020 (99)	58.9 ± 6.5 y Finland	Feeding Study	79 (12 weeks)	LC-MS/MS (Targeted)	Plasma	Fatty fish, lean fish, and camelina sativa oil

Fatty fish	4- and 17-HDoHE; 19,20- dihydroxy-docosapentaenoic acid (19,20-DiHDPA); hydroxydocosahexaenoic acids (5- and 18-HEPE); docosahexaenoic acid (DHA); eicosapentaenoic acid (EPA)
Camelina sativa oil	15-hydroxyeicosadienoic acid (15-HEDE); ALA-derived hydroxyoctadecatrienoic acids (9- and 13-HOTrE); 12(13)-epoxy-octadecadienoic acid (12(13)-EpODE)

Michielsen et al., 2019 (100)	SFA: 51 ± 7y, MUFA:58 ± 5y, MED: 57 ± 5y, Netherlands	Feeding Study	47 (8 weeks)	^{1}H NMR (Targeted – circulating lipids, lipoprotein particles, lipoprotein composition, and low-molecular-weight metabolites, including amino acids)	Serum	Mediterranean (MED), monounsaturated fatty acid (MUFA), or saturated fatty acid (SFA)

Mediterranean (MED)		LDL related fractions; subset of the cholesterol fractions (serum cholesterol, very-low-density lipoprotein (VLDL)-cholesterol, free cholesterol, and remnant cholesterol); multiple VLDL related fractions; mainly in the XL-, L-, and M-VLDL subclasses; total VLDL-TG concentration; total triglyceride (TG); ApoB[1]; ApoB to ApoA1 ratio; albumin; DHA[1]; DHA to FA[1]; FAω3 to FA ratio[1]; CLA[1]; CLA to FA ratio[1]; MUFA to FA ratio[1]				
Monounsaturated fatty acid (MUFA)		LDL related fractions; subset of the cholesterol fractions (serum cholesterol, very-low-density lipoprotein (VLDL)-cholesterol, free cholesterol, and remnant cholesterol); ApoB[1]; ApoB to ApoA1 ratio; albumin; FAω3[1]; FAω3 to FA ratio[1]; CLA[1]; CLA to FA ratio[1]; MUFA to FA ratio[1]				
Mills et al., 2017 (101)	26.3 ± 1.6 y, Switzerland	Crossover Feeding Study	15 males (Acute)	UPLC-ESI-MS (Targeted – phenolic acids)	Plasma	Low polyphenol coffee (89 mg CGA), high polyphenol coffee (310 mg CGA), or control (0 mg CGA)
Low polyphenol coffee (89 mg CGA) and high polyphenol coffee (310 mg CGA)		3-caffeoylquinic acid (3CQA); 4-caffeoylquinic acid (4CQA); caffeic-40-O-sulfate (CA4S); 3-feruloylquinic acid (3FQA); 4-feruloylquinic acid (4FQA); 5-feruloylquinic acid (5FQA); ferulic acid (FA); isoferulic acid (iFA); methylferulic acid (MeFA); ferulic-40-O-glucuronide (F4G); isoferulic-30-O-glucuronide (iF3G); ferulic-40-O-sulfate (F4S); and isoferulic-30-O-sulfate (iF4S)				
Moazzami et al., 2012 (102)	58.8 ± 5.8 y, Finland	Crossover Feeding Study	33 females (8 weeks)	^{1}H NMR (Untargeted)	Serum	"A minimum of 20% of their daily energy intake as high fiber WG rye bread (RB) or refined wheat bread (WB)"
Rye bread		Isoleucine; Leucine; Betaine; N,N-dimethylglycine				
Mora-Cubillos et al., 2015 (103)	18-65 y, Spain	Feeding Study	50 (12 weeks)	LC-ESI-Q-ToF-MS (Untargeted)	Plasma	Daily supplement of 30 g of raw mixed nuts with skin (15 g of walnuts, 7.5 g of almonds, and 7.5 g of hazelnuts), while the control was recommended to avoid consumption of nuts
Mixed nuts		Urolithin A glucuronide; Sebacic acid; Dodecanedioic acid				
Mulder et al., 2005 (104)	18-70 y, Netherlands	Crossover Feeding Study	17 males (2 days)	HPLC-MS/MS (Targeted – Hippuric acid)	Urine	Daily dose of 6 g green tea solids, 6 g black tea solids, or 360 mg caffeine
Green Tea & Black Tea		Hippuric acid				
Münger et al., 2017 (105)	18-40 y, Switzerland	Crossover Feeding Study	11 (Acute)	GC-MS and ^{1}H NMR (Untargeted)	Urine	Single intake of milk and cheese as test products, and soy-based drink as control
Milk		Lactose; galactose; galactonate; allantoin; hippurate; galactitol; galactono-1,5-lactone				
Cheese		3-phenyllactic; alanine, proline, and pyroglutamic acid				

Soy-based drink		Pinitol; trigonelline				
Nieman et al., 2012 (106)	49–75 y, US	Feeding Study	62 Overweight/ obese, post-menopausal women (each day for 10 weeks)	GC–MS (Targeted – fatty acids)	Plasma	Chia seed (whole or milled) and placebo (poppy seed) groups
Milled chia seed		α-Linolenic acid (ALA); eicosapentaenoic acid				
Nilsson et al., 2010 (107)	25.9 ± 3.2 y, Sweden	Crossover Feeding Study	15 (Acute)	GC (Targeted – short-chain fatty acids (SCFA))	Plasma	Eight cereal-based evening test meals with different GI and contents of indigestible carbohydrates
High-amylose barley kernels or high-b-glucan barley kernels		Butyric acid				
O'Sullivan et al., 2011 (108)	35.5 ± 12 y, Ireland	Feeding Study	125 (4 weeks)	^{1}H NMR (plasma and urine), and GC (plasma fatty acid profiling) (Untargeted)	Urine & Plasma	3 clusters were identified and characterized on the basis of the food groups that were distinct contributors to the total energy intake in each cluster
Higher energy contribution from whole-meal bread, whole milk, fish, confectionary, and ice cream and desserts; Lower contribution from low-energy beverages		Glycine; phenylacetylglutamine; actetoacetate				
Higher energy contributions of white bread, sugars and preserves, butter and spreads, red meat, red-meat dishes, meat products, and alcohol;		Oleic acid (18:1); y-linoleic acid (18:3n-6); docosapentaenoic acid (DPA; 22:5n-3); total monounsaturated fatty acid (MUFA); trimethylamine N-oxide (TMAO); O-acetylcarnitine; nn-dimethylglycine				

Lower contribution from vegetables						
Red meat	O-acetylcarnitine					
Vegetable intake	Phenylacetylglutamine					
Ostertag et al., 2017 (109)	23–65 y, Scotland	Crossover Feeding Study	42 (Acute)	^{1}H NMR and HPLC-ToF-MS (Untargeted)	Urine	Flavan-3-ol-enriched dark chocolate and standard dark chocolate compared with white chocolate
Flavan-3-ol-enriched dark and Dark Chocolate	Dietary Markers: 3- and 7-methylxanthines; caffeine; hydroxynicotinate; theobromine (3,7-dimethylxanthine); vanilloylglycine; xanthine purine rings; epicatechin monosulfate; 4-hydroxy-5-(3,4-dihydroxyphenyl)-valerate; 5-(3',4'-dihydroxyphenyl)-γ-valerolactone; 3'-methoxy-4'-hydroxyphenylvalerolactone Endogenous Markers: creatinine; arginine; valine; alanine; glycine; N-methylnicotinamide; N-acetylated compounds; dimethylamine; 3-hydroxyisovalerate; 2-hydroxyisobutyrate; 3-hydroxyisobutyrate; lactate; pyruvate; 4-hydroxyphenyl acetate; tyrosine					
Paetau et al., 1998 (110)	45.7 ± 7.6 y, US	Crossover Feeding Study	15 (4 weeks)	HPLC–MS (Targeted)	Plasma	Lycopene-rich tomato juice, tomato oleoresin, lycopene beadlets, and a placebo
Lycopene rich tomato juice	Lycopene; Cyclolycopene (2,6-cyclolycopene-1,5-diol); Lutein; β-Carotene; ζ-Carotene; Phytofluen; Phytoene					
Pereira-Caro et al., 2014 (111)	23-60 y, UK	Crossover Feeding Study	12 (Acute)	HPLC–MS and GC–MS (Targeted – flavonoids & aromatic acids)	Urine	Pulp-enriched orange juice (250 mL)
Orange juice	Metabolites: Hesperetin-*O*-glucuronides; naringenin-*O*-glucuronides; hesperetin-3'-*O*-sulfate Catabolites: 3-(3'-methoxy-4'-hydroxyphenyl)propionic acid; 3-(3'-hydroxy-4'-methoxyphenyl)propionic acid, 3-(3'-hydroxy-4'-methoxyphenyl)hydracrylic acid; 3-(3'-hydroxyphenyl)hydracrylic acid; 3'-methoxy-4'-hydroxyphenylacetic acid, hippuric acid, 3'-hydroxyhippuric acid, and 4'-hydroxyhippuric acid					
Pereira-Caro et al., 2017 (112)	31.8 ± 5.7 y, Scotland	Crossover Feeding Study	10 males (Acute)	HPLC-MS (Targeted – flavanone metabolites & (poly)phenol catabolites)	Urine	500 mL of orange juice containing 398 mmol of (poly)phenols, of which 330 mmol was flavanones
Orange juice	Cinnamic acids: Ferulic acid-4'-sulfate; Isoferulic acid-3'-O-glucuronide Phenylhydracrylic acid: 3-(3'-Hydroxy-4'-methoxyphenyl)hydracrylic acid Phenylpropionic acids: 3-(4'-Methoxyphenyl)propionic acid-3'-O-glucuronide; 3-(4'-Methoxyphenyl)propionic acid-3'-sulfate Phenylacetic acid: 4'-Hydroxyphenylacetic acid					

	Benzoic acid: 3-Hydroxybenzoic acid-4-sulfate Mandelic acid: 4'-Hydroxymandelic acid Hippuric acid: 4'-Hydroxyhippuric acid					
Pereira-Caro et al., 2020 (113)	31.8 ± 5.7 y, Scotland	Feeding Study	10 (acute)	HPLC-HR-MS (Targeted)	Plasma	Orange juice
Orange juice	<u>Cinnamic acid derivatives</u>: 4'-Hydroxy-3'-methoxycinnamic; 3'-Methoxycinnamic acid-4'-glucuronide; 3'-Methoxycinnamic acid-4'-sulfate; 4'-Methoxycinnamic acid-3'-glucuronide; 3'-Hydroxy-4'-methoxycinnamic; Cinnamic acid-4'-glucuronide; 3'-Hydroxycinnamic acid-4'-sulfate; 4'-Hydroxycinnamic acid-3'-sulfate <u>Phenylpropanoic acid derivatives</u>: 3-Hydroxy-3-(3'-hydroxy-4'-methoxyphenyl)propanoic acid; 3-Hydroxy-3-(3'-hydroxyphenyl)propanoic acid; 3-(3'-Methoxyphenyl)propanoic acid-4'-glucuronide; 3-(3'-Methoxyphenyl)propanoic acid-4'-sulfate; 3-(3'-Hydroxy-4'-methoxyphenyl)propanoic acid; 3-(4'-Methoxyphenyl)propanoic acid-3'-glucuronide; 3-(4'-Methoxyphenyl)propanoic acid-3'-sulfate; 3-(3'-Hydroxyphenyl)propanoic acid; 3-(4'-Hydroxyphenyl)propanoic acid; 3-(Phenyl)propanoic acid; 3-(3',4'-Dihydroxyphenyl)propanoic acid; 3-(3'-Hydroxyphenyl)propanoic acid-4'-sulfate; 3-(4'-Hydroxyphenyl)propanoic acid-3'-sulfate; 3-(4'-Hydroxy-3'-methoxyphenyl)propanoic acid <u>Phenylacetic acid derivatives</u>: 2-Hydroxy-2-(4'-hydroxyphenyl)acetic; 3'-Hydroxyphenylacetic acid; 4'-Hydroxyphenylacetic acid; 2-Hydroxy-2-(4'-hydroxy-3'-methoxyphenyl); 3'-Hydroxyphenylacetic acid-4'-sulfate; Methoxyphenylacetic acid-glucuronide; 4'-Methoxyphenylacetic acid-3'-sulfate; 3'-Methoxyphenylacetic acid-4'-sulfate <u>Benzoic acid derivatives</u>: 3,4-Dihydroxybenzoic acid; Benzoic acid-4-sulfate; Benzoic acid-3-sulfate <u>Benzoylglycine derivatives</u>: Hippuric acid; 3'-Hydroxyhippuric acid; 4'-Hydroxyhippuric acid					
Perera et al., 2016 (114)	35-75 y, US	Feeding Study	46 males (4 weeks)	LC-ESI-MS, LC-MS, and GC-MS (Targeted)	Serum	Dry beans
Dry bean-enriched diet (250 g/d)	Pipecolic acid; S-methyl-cysteine; N-Acetylornithine; Trigonelline; Indole propionate					
Pezdirc et al., 2016 (115)	22.0 ± 4.2	Crossover Feeding Study	30 (4 weeks)	HPLC (Targeted – lycopene oxidation products)	Plasma	High-carotenoid-containing fruits and vegetables F/V (HCFV) (176,425 mg beta carotene/wk) vs. low-carotenoid F/V (LCFV) (2,073 mg beta carotene/wk)
Carotenoids from fruits and vegetables (F/V)	Alpha carotene, beta carotene, lutein, total carotenoids					

Pujos-Guillot et al., 2013 (116)	Short-term intervention (STI): 33 ± 7 y Medium-term intervention (MTI): 56 ± 1 y, France	Feeding Study	STI: 4 (Acute) MTI: 24 (4-weeks)	RP-LC-ESI-ToF-MS (Untargeted)	Urine	(a) consumed an acute dose of orange or grapefruit juice, (b) consumed orange juice regularly for one month, and (c) reported high or low consumption of citrus products for a large cohort study
Citrus fruit		Proline betaine; flavanone glucuronides; two terpene metabolites (limonene 8,9-diol glucuronide and nootkatone 13,14-diol glucuronide)				
Quifer-Rada et al., 2014 (117)	28 ± 3 y, Spain	Crossover Feeding Study	41 (Acute)	LC-MS/MS (Targeted – isoxanthohumol)	Urine	30 g of ethanol/d as gin or beer, or an equivalent amount of polyphenols as nonalcoholic beer
Beer		Isoxanthohumol (IX)				
Rådjursöga et al., 2018 (118)	Males: 27.0 ± 6.6 33.2 ± 13.2 Females 25.9 ± 10.1 31.9 ± 8.2, Sweden	Crossover Feeding Study	32 (3 days)	^{1}H NMR (Untargeted)	Serum	Breakfast meals corresponding to vegan (VE), lacto ovo-vegetarian (LOV), and omnivore (OM) diets
Breakfast meals		Vegan (VE): Acetate; Acetone; Alanine; Arginine & Lysine; Glucose (alpha, beta); Asparagine; Betaine; Creatinine; Creatinine & Creatine & Creatine phosphate; Leucine; Lipids/FFA; Mannose; Methionine; myo-Inositol; O-Phosphocholine & 3-Hydroxybutyrate; Ornithine; Succinic acid; Threonine Lacto ovo-vegetarian (LOV): 3-hydroxyisobutyrate; Acetate; Glucose (alpha, beta); Arginine & Lysine; Carnitine & Acetoacetate; Leucine & Arginine; Mannose; N-Acetylcysteine & Proline & Glutamate; Ornithine; Proline; Propylene glycol; Tyrosine; Valine Omnivore (OM): 3-Hydroxyisobutyrate; Acetate; Arginine & Lysine; Betaine; Carnitine & Acetoacetate; Choline; Isoleucine; Leucine & Arginine; Lipids/FFA; Lysine; Mannose; Methionine; myo-Inositol; Ornithine; Proline; Tyrosine; Valine				
Rangel-Huerta et al., 2017 (119)	22–63 y, Spain	Crossover Feeding Study	30 (12 weeks)	UHPLC-MS (Untargeted)	Serum	Normal-polyphenol orange juice or a high-polyphenol orange juice (500 ml of orange juice)
Orange juice		Stachydrine; methyl glucopyranoside (alpha+beta); betonicine; galactonate				

Rasmussen et al., 2012 (120)	37–45 y, Denmark	Feeding Study	77 overweight (6 months)	^{1}H NMR (Untargeted)	Urine	Low protein, low GI (LP/LGI); low protein, high GI (LP/HGI); high-protein, low GI (HP/LGI); high-protein, high GI (HP/HGI); or control (CTR) diet
High protein diets	Creatine; nitrogen					
Low protein diets	Citric acid					
Rebholz et al., 2018 (121)	31–55 y, US	Feeding Study	329 (8 weeks)	GC–MS and LC–MS (Untargeted)	Serum	DASH diet, the fruit and vegetables diet, or a control diet
Dash diet	N-methylproline; stachydrine; tryptophan betaine; theobromine; 7-methylurate; chiro-inositol; 3-methylxanthine; methyl glucopyranoside; ß-cryptoxanthin; 7-methylxanthine					
Regueiro et al., 2014 (122)	30.7 ± 5.9 y, Spain	Crossover Feeding Study	21 males	LC-ESI-MS/MS (Targeted – tartaric acid)	Urine	100-, 200-, or 300-ml wine
Wine	Tartaric acid					
Reisdorph et al., 2020 (123)	61 ± 2 y, US	Crossover Feeding Study	19 (6 weeks)	RP-LC/MS (Untargeted)	Urine	DASH diet
Apple	4-Hydroxydiphenylamine; Diphenylamine; Presqualene diphosphate; 28-Hydroxymangiferonic acid; 3,4-Dihydro-6-hydroxy-2,5,7,8-tetramethyl-2H-1-benzopyran-2-carboxylic acid; 3-Methylellagic acid 2-(4-galactosylglucoside); 3-O-p-cis-Coumaroylalphitolic acid; 4-HYDROXYBENZOATE; 4-Hydroxyxanthone; 5,7,8-Trihydroxy-3,6,4'-trimethoxy-flavone 8-isovalerate; Afzelechin; Carthamin; Cinnamtannin A2; D-Chicoric acid; Dianhydroaurasperone C; Evoxine; Gnididilatin; Hibiscitrin; Mangiferoleanone; Manglupenone; Morellin; Procyanidin B8; Quercetin 3-(6''-malonyl-glucoside); Rubroskyrin; Scutellarein 6-xyloside						
Apple juice	Deoxyuridine; (S)-N-Acetylmethionine; 2,5-Furandicarboxylic acid; Ascorbigen; Cobalt-precorrin 6; Cyanidin 3-(6''-acetylglucoside); Fraxetin; Gerberinol; Glucoputranjivin; Leucocyanidin; Leucodelphinidin 3-O-alpha-L-rhamnopyranoside; Morellic acid; PG(18:4/0:0); Rubraflavone B ; Rubroskyrin; Scutellarein 7-glucuronide-6-ferulate; xi-8-Acetonyldihydrosanguinarine						
Apple and apple juice	4-Hydroxydiphenylamine; Diphenylamine; Mammeigin; Mangiferoleanone; Manglupenone; Presqualene diphosphate; Vomifoliol 9-[xylosyl-(1-6)-glucoside]						
Beef	1-(beta-D-Ribofuranosyl)-1,4-dihydronicotinamide; 1alpha-hydroxy-22-[3-(1-hydroxy-1-methylethyl)phenyl]-23,24,25,26,27-pentanorvitamin D3; Alpha-linolenyl carnitine; Undecanoylcarnitine; Cycloheximide; Amifloxacin; Artenolide; b-D-Glucopyranosiduronic acid, (3a,5b)-24-[(carboxymethyl)amino]-24-oxocholan-3-yl; Desmosterol; Ferrioxamine; Hydroxyhexanoycarnitine; L-Urobilin; LysoPC(18:3); LysoPC(20:5); LysoPE(0:0/18:3); LysoPE(0:0/20:4);						

	LysoPE(0:0/20:5); N'-Formylkynurenine; PE(P-16:0/18:3); Pristanoylglycine; Rifaximin; Sodium taurocholate; Tryptophyl-Arginine; 1-(9Z,12Z-octadecadienoyl)-2-(9Z-nonadecenoyl)-glycero-3-phosphoserine; 1-(9Z-heptadecenoyl)-glycero-3-phosphoserine; 1-(9Z-tetradecenoyl)-2-dodecanoyl-glycero-3-phosphoethanolamine; 1-dodecanoyl-2-tetradecanoyl-glycero-3-phosphoserine; 1-hexadecanoyl-2-(2E-propionyl)-sn-glycero-3-phosphocholine; 1-hexadecanoyl-2-(9-carboxy-nonanoyl)-sn-glycero-3-phosphocholine; 1-hexadecanoyl-2-(9-oxo-nonanoyl)-sn-glycero-3-phosphocholine; 1-hexadecanoyl-2-glutaroyl-sn-glycero-3-phosphocholine; 1-tridecanoyl-2-eicosanoyl-glycero-3-phosphoserine; N-hexadecanoyl-phenylalanine E; N-hexadecanoyl-taurine; N-octadecanoyl-tyrosine; PI(18:0/0:0); PI(18:1/0:0)
Blueberries	(14alpha,17beta,20S,22R)-14,20-Epoxy-17-hydroxy-1-oxowitha-3,5,24-trienolide; Dalpanol O-glucoside; 1,2-Dimethoxy-13-methyl-[1,3]benzodioxolo[5,6-c]phenanthridine; 6-C-Fucosylluteolin; 8-Hydroxy-2-methoxy-6-methyl-1,4-naphthoquinone; Artobiloxanthone; Artomunoxanthentrione; Avicularin; Chlorogenoquinone delta6-Dehydroferruginol; Dihydroroseoside; Fisetin; Kuwanon A; L-Citronellol glucoside; Leonuriside A; Myricetin 3-[galloyl-(-2)-4-acetyl-a-L-rhamnoside]; N-(2,5-Dihydroxyphenyl)pyridinium(1+); Naphthoherniarin; Physalin E acetate; Rubraflavone C; Sandoricin; 4a-peroxy-tetrahydrobiopterin; Withaperuvin E; 5,6-Dimethoxy-3',4'methylenedioxy-6'',6''-dimethylpyrano[2'',3'':7,8]isoflavone; 5,7,3',4',5'-Pentahydroxy-3,6,8-trimethoxyflavone; 5,7-Dihydroxy-3,6,8,4'-tetramethoxyflavone 7-glucosyl-(1-3)-galactoside; 6,8-Di-C-beta-D-arabinopyranosylapigenin; Bruceantinol; Eupachloroxin; Europinidin 3-glucoside; Neolinderatone; Ornithine; Rotundifoline
Broccoli	Isorhamnetin 3-sophoroside 7-glucoside; Muzanzagenin; Chlorophyll c; Cyclobrassinin; Dihydroneopterin phosphate; Glucoraphanin; Hirsutin; Pyrophaeophorbide a; Raphanusamic acid; Bn-NCC-2; Cabbage identification factor 1; 3' 5'-cyclic AMP; 5-Methyltetrahydropteroyltri-L-glutamate; Eremosulphoxinolide B; Linamarin; Malvin; Asclepin; CL(22:6/18:2/18:2/16:1); Cnidimol 7-glucoside; Isotriglochinin
Coffee	2,3-Dihydro-5-(5-methyl-2-furanyl)-1H-pyrrolizine; Furfuryl acetate; Na-p-Hydroxycoumaroyltryptophan; N-Caffeoyltryptophan; 11-Methylgerberinol; 2-Hexylbenzothiazole; 5,6,7-Trihydroxy-4'-methoxyflavanone 7-(2-p-coumaroylglucoside); 5,7-Dihydroxy-3,6-dimethoxyflavone; 7-Ethoxy-4-methyl-2H-1-benzopyran-2-one; 8-Caffeoyl-3,4-dihydro-5,7-dihydroxy-4-phenylcoumarin; 9-Hydroxy-4-methoxypsoralen; Benzylamine Cnidimol 7-glucoside; Cyclohexylamine; Formononetin 7-(2-p-hydroxybenzoylglucoside); Gamma-glutamyl-Glycine; Gentianine; Glycerol 1-propanoate; Gynocardin; Harmalol; Hypoglycin B; Meteloidine; N6-Galacturonyl-L-lysine; N-eicosanoyl-ethanolamine; N'-Formylkynurenine; p-Anisidine; Pantothenic acid; Physagulin B; Riboflavin reduced; Sciadopitysin; Secoclausenamide; Trietazine
Chicken	Avermectin A2b; Pivaloylcarnitine; (all-Z)-7,10,13-Docosatrienoic acid; 1,2-di-(6Z,9Z,12Z,15Z-octadecatetraenoyl)-sn-glycero-3-phospho-(1'-myo-inositol); 3-(7'-Methylthio)heptylmalic acid; Gamma-glutamyl-Asparagine; Gamma-glutamyl-Cysteine; Gamma-glutamyl-Glycine; Levofloxacin; Methyl bisnorbiotinyl ketone; Prolyl-Arginine; PS(17:2/0:0); TG(18:1/14:0/20:4); TG(18:3/16:0/18:3); TG(20:4/18:0/18:3)
Cucumber	(R)-Byakangelicin; Gibberellin A4; Lipoyllysine; N-ACETYL-DL-METHIONINE; Procyanidin B3 7-glucoside; Tokinolide A 5alpha-Stigmasta-7,22,25-trien-3beta-ol; 4-Methylene-2-pyrrolidinecarboxylic acid; 2-O-alpha-D-Galactopyranosyl-1-deoxynojirimycin; 2-Phenyl-2-butenal; 3'-O-Methylcyanidin 3-O-beta-D-glucoside; alpha-Carissanol

	alpha-Ionol O-[arabinosyl-(1-6)-glucoside]; Corchoionoside B; D-Glutamine; LysoPE(0:0/18:4); 7(14)-Bisabolene-2,3,10,11-tetrol; 7-Hydroxycostal; Asparenyol; Fluazifop; Formylfusarochromanone; Fumitremorgin B; PE(14:1/14:0); PE(18:0/20:5)
Grapefruit	(4S,8S,12S,16S,20S-Pentamethylhexacosanyl)-beta-D-mannosyl phosphate; 4-[(6,7-Dihydroxy-3,7-dimethyl-2-octenyl)oxy]-7H-furo[3,2-g][1]benzopyran-7-one; 6''-(3-Hydroxy-3-methylbutanoyl)astragalin; Citbismine B; Citrusin F Dolichyl phosphate D-mannose; Hesperetin 7-(2,6-dirhamnosylglucoside); Hesperetin 7-neohesperidoside; Isolimonic acid; Nobiletin; Nomilin; Nomilinic acid; Obacunone; Obacunone 17-O-beta-D-glucoside; (E)-Suberenol; 4-Hydroxyphenylacetaldehyde E; alpha-CEHC; Cyclocalamin; Deacetylnomilin; Deacetylnomilinic acid; Ichangin 4-glucoside; Margrapine B; Naringin 6''-rhamnoside; Natsudaidain 3-(4-O-3-hydroxy-3-methylglutaroylglucoside); Neoacrimarine B; Paradisin C; Subaphylline; Bis(5-hydroxynoracronycine); 1'',2''-Dihydro-8-hydroxyisopentanyl-2'-methoxy-4'-O-methylalpinumisoflavone; 1-Hydroxy-3,5-dimethoxy-2-prenylxanthone; 1-O-Caffeoylglucose; 3-Vinylbacteriochlorophyllide a; 4-(4-Hydroxyphenyl)-2-butanone glucoside; 4,4'-Diapophytoene; 7-(4-Carboxy-3-hydroxy-3-methylbutanoyl)sudachitin 4'-glucoside; 7alpha-1(10-19)-Abeo-7-acetoxyobacun-9(11)-ene; 8-Hydroxycarapin, 3,8-Hemiacetal; 8-Epiiridodial glucoside tetraacetate; Apigenin 7-(6''-O-alpha-rhamnosyl-beta-glucoside); Apo-10'-violaxanthal; Chrysoeriol 7-(3'',6''-di-(E)-p-coumaroylglucoside); Cinnamic acid; Dukunolide A; Dulxanthone H; Heteroartonin A; Isomangiferin; Kaempferide 3-[rhamnopyranosyl-(1-6)-glucoside] 7-rhamnoside;Mammea A/AC cyclo F; Methylsyringin Esi+8.138; Moracin N; Morusignin B; Mumefural; Myricetin 3-glucoside; Normammein; Paucine; Physalin I; Phytoene; Pinostrobin 5-glucoside; Prupaside; Ramontoside; Sapidolide A; Ssioriside; 1-Phenyl-6,7-dihydroxy-isochroman; 3-alpha(S)-Strictosidine; 3-hydroxyundecanoyl carnitine; 7-[(6-Hydroxy-3,7-dimethyl-2,7-octadienyl)oxy]-2H-1-benzopyran-2-one
Peanut butter	Biotin; gamma-Tocopherol; Linoelaidic acid; 1,4,5-Naphthalenetriol; DG(16:0/18:3/0:0); DG(18:0/18:3/0:0); Glyceryl lactooleate; alpha-Linolenic acid; Arachin; 19-hydroxy-nonadecanoic acid; Homoarecoline; Vinyl caffeate; 1-Pyrrolidinecarboxaldehyde; (1S,2R,4R,8S)-p-Menthane-2,8,9-triol 2-glucoside; (1xi,3xi)-1,2,3,4-Tetrahydro-1-methyl-beta-carboline-3-carboxylic acid; (R)-Byakangelicin 2'-glucoside; 1-Cyclopropyl-4-methyl-1,3-cyclohexanediol; 4'-Apo-3,4-didehydrolycopene; 4'-Apo-beta,psi-caroten-4'-al; 4-Demethylsimmondsin 2'-(E)-ferulate; 7,9-Hexacosanedione Docosanamide; L-3-Amino-2-(oxalylamino)propanoic acid; MG(0:0/20:0/0:0); Perilloside A; Phloroacetophenone 6'-[xylosyl-(1-6)-glucoside]; Sorbitan stearate; Villinol; (R)-2,4-Dihydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one 2-glucoside; 1-(14-methyl-pentadecanoyl)-2-(8-[3]-ladderane-octanyl)-sn-glycerol; 1-(4-O-beta-D-glucopyranosyl-3-methoxyphenyl)-3,5-dihydroxydecane; 1-(9Z-hexadecenoyl)-2-(4Z,7Z,10Z,13Z,16Z,19Z-docosahexaenoyl)-glycero-3-phospho-(1'-myo-inositol); 1-(9Z-nonadecenoyl)-glycero-3-phospho-(1'-myo-inositol); 1-(9Z-pentadecenoyl)-2-(9Z-octadecenoyl)-glycero-3-phospho-(1'-sn-glycerol); 1-dodecanoyl-2-heneicosanoyl-glycero-3-phospho-(1'-sn-glycerol); 1-eicosyl-glycero-3-phospho-(1'-myo-inositol); 1-O-(1Z-Tetradecenyl)-2-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine; 1-octadecyl-2-dodecanoyl-glycero-3-phospho-(1'-myo-inositol); 3-oxo-tricosanoic acid; 4-Amino-2-methyl-1-naphthol; 6,8-Icosanedione; 6-bromo-tricosa-5E,9Z-dienoic acid; CPA(18:2/0:0); D-Cysteine; Glucosamine-1P; LysoPC(24:0); ysoPE(0:0/18:1); LysoPE(0:0/18:2); Methyl (E)-2-decene-4,6,8-triynoate; N-(11Z-eicosaenoyl)-ethanolamine; N-(9Z-octadecenoyl)-histidine; N-(9Z-octadecenoyl)-phenylalanine; NBD-Stearoyl-2-Arachidonoyl-sn-glycerol; NeuAcalpha2-

		3Galbeta-Cer(d18:1/24:1); N-Oleoylethanolamine; Octadecanedioic acid; PC(22:6/22:5); PE(P-16:0/14:1); PE-Cer(d14:1(4E)/20:0(2OH)); PE-Cer(d14:1(4E)/20:1(2OH)); PE-Cer(d14:2(4E,6E)/22:1); Phellogenic acid; Propylene glycol stearate; TG(20:5/20:5/20:5); TG(22:4/22:6/22:5)				
Pork		4,8 dimethylnonanoyl carnitine; 1-(1Z-eicosenyl)-glycero-3-phosphoethanolamine; Galabiosylceramide (d18:1/18:0); Gamma Hydroxybutyric Acid; Tetracosahexaenoic acid				
Tilapia		TRANS-4-HYDROXYPROLINE; CITCO; 9-methyl-sphinga-4E,8E,10E-trienine; Cer(d18:1/14:0); Ceramide; Dodecanoylcarnitine; Glutarylglycine				
Renouf et al., 2013 (124)	18-65 y, Switzerland	Crossover Feeding Study	12 (Acute)	LC-ESI-MS/MS (Targeted – catechins)	Plasma	Green tea (3, 5, and 7 g tea leaves in 400 mL water)
Green Tea		Epicatechin (EC), 4'-O-Me-epigallocatechin (4-O-Me-EGC); epigallocatechin (EGC); epigallocatechin-3-gallate (EGCg)				
Rodriguez-Mateos et al., 2012 (125)	27 ± 3 (SEM) y, UK	Crossover Feeding Study	15	HPLC-fluorescence detection/UV (Targeted – flavonoids)	Plasma	High-flavanol (266 mg) chocolate containing maltitol, a high-flavanol (251 mg) chocolate with sucrose or a low-flavanol (48 mg) chocolate with sucrose
Chocolate (high-flavanol chocolate with sucrose)		Non-methylated flavanols; 3'-O-methylated flavanols; 4'-O-methylated flavanols				
Ross et al., 2012 (126)	45 ± 10 y, UK	Crossover Feeding Study	266 (16 weeks)	HPLC with Coularray electrochemical detection (Targeted – alkylresorcinols)	Plasma	Control (no dietary change), intervention 1 (60 g whole grain/d for 16 wk), or intervention 2 (60 g whole grain/d for 8 wk followed by 120 g whole grain/d for 8 wk)
Whole grain		Alkylresorcinols				
Ross et al., 2013 (127)	20 – 50 y, Switzerland	Crossover Feeding Study	17 (2 weeks)	^{1}H NMR and GC-MS (plasma) (Untargeted)	Urine & Plasma	Whole grain (WG)-rich or Refined grain-rich foods
Whole grain		Urine: N-acetylcarnitine; taurine; 4-hydroxyphenylacetate^{1}; dimethylamine^{1}; trimethylamine; urea^{1}; methylguanadine^{1}; creatinine; ^{1}pyruvate; citrate^{1}; succinate^{1}; fumarate^{2}; N-acetyl-glycoproteins^{1}; 3-hydroxyisovalerate^{1} Plasma: urea				
Ross et al., 2015 (128)	41-67 y, Sweden	Crossover Feeding Study	17 overweight men (Acute)	GC–MS (Untargeted and targeted)	Plasma	Baked herring, pickled herring, and baked beef
Meat intake		Beta-alanine; 4-hydroxyproline; 2-Aminoadipic acid; Creatinine				
Herring intake		Docosahexaenoic acid (DHA); cetoleic acid				

Roura et al., 2007 (129)	18-50 y, Spain	Crossover Feeding Study	21 (Acute)	LC-MS/MS (Targeted - epicatechin ((–)-Ec) metabolites)	Urine	Polyphenol-rich food (PRF) (cocoa beverage containing 40 g of cocoa powder and 250 ml water) and a polyphenol-free food (PFF) (250 ml of whole milk)
Polyphenol-rich food (cocoa beverage containing 40 g of cocoa powder and 250 ml water)	(–)-epicatechin-sulphate 1; (–)-epicatechin-sulphate 2; epicatechin-sulphate 3; (–)-epicatechin-glucuronide					
Roura et al., 2007 (130)	25.7 ± 6.9 y, Spain	Crossover Feeding Study	21 (Acute)	LC-MS/MS (Targeted – (–)-epicatechin ((–)-Ec))	Plasma	250 ml of whole milk (M-c) (control), 40 g of cocoa powder dissolved in 250 ml of whole milk (CC-M), and 40 g of cocoa powder dissolved with 250 ml of water (CC-W)
Cocoa (40 g of cocoa powder dissolved in 250 ml of whole milk and with 250 ml of water	(–)- epicatechin glucuronide ((–)-Ec-glucuronide)					
Roura et al., 2008 (131)	25.7 ± 6.9 y, Spain	Crossover Feeding Study	21 (Acute)	LC–MS/MS (Targeted – (–)-epicatechin metabolite)	Urine	(1) 250 ml whole milk as a control; (2) 40 g cocoa powder dissolved in 250 ml whole milk (CC–M); (3) 40 g cocoa powder dissolved in 250 ml water (CC–W)
Cocoa beverages	Epicatechin-sulfate 2 (Ec-S^2); epicatechin-glucuronide (Ec-G)					
Schär et al., 2015 (132)	60.6 ± 5.6 y, UK	Crossover Feeding Study	14 males (Acute)	HPLC-ESI-MS (Targeted – flavanone metabolites)	Plasma	Orange juice or a hesperidin supplement (both providing 320 mg hesperidin) or control (all matched for sugar and vitamin C content)
Orange juice	Flavanones: 1, **hesperetin-glucuronide**; 2, naringenin-7-glucuronide; 3, hesperetin-glucuronide; 4, hesperetin-diglucuronide; 5, hesperetindiglucuronide; 6, naringenin-glucuronide; 7, hesperetin; 8, naringenin Phenolic metabolites: 1, **hippuric acid**; 2, dihydroferulic acid; 3, dihydroferulic acid–3-glucuronide; 4, 4-hydroxyphenylacetic acid; 5, vanillic acid; 6, hydroxyhippuric acid; 7, iso/ferulic acid–glucuronide; 8, 3-hydroxyhippuric acid; 9, isovanillic acid; 10, 3-hydroxyphenylacetic acid; 11, vanillic acid–glucuronide; 12, isovanillic acid–glucuronide; 13, iso/vanillic acid–glucuronide; 14, 4-hydroxy-benzoic acid; and 15, benzoic acid–4-glucuronide					

Schär et al., 2018 (133)	25-62 y, UK	Crossover Feeding Study	7 (Acute)	UPLC–ESI–MS/MS (Targeted - phenolic acids and avenanthramides)	Urine	Oat bran (60 g) or a phenolic low (control) diet
Oat Bran	Avenanthramide A; Ferulic acid; p-coumaric acid; Dihydroferulic acid; Isovanillic acid; Syringic acid; 2,4-dihydroxybenzoic acid; 2,5-dihydroxybenzoic acid; **Vanillic acid**; 2-hydroxyhippuric acid; **3-hydroxyhippuric acid**; **4-hydroxyhippuric acid**; Feruloylglycine; Caffeic acid-sulfate; **(Iso)ferulic acid-O-sulfate**; Isoferulic acid-3-O-sulfate; Syringic acid-O-sulfate; Sinapic acid-O-sulfate; Dihydroxybenzoic acid-O-sulfate; **Benzoic acid-O-sulfate**; Vanillin or hydroxyphenylacetic acid-O-sulfate; Syringaldehyde OR homovanillic acid OR dihydroxyhydrocinamic acid sulfate; Ferulic Acid-4-O-glucuronide; Isoferulic acid-O-glucuronide; Dihydroferulic acid-4-O-glucuronide; Dihydro(iso)ferulic acid-O-glucuronide; Benzoic acid-O-glucuornide; Syringaldehyde OR Homovanillic acid OR 3,4-dihydroxyhydrocinamic acid-O-glucuronide; Vanillin OR 4-hydroxyphenylacetic acid-O-glucuronide				
Scheffler et al., 2016 (134)	25–33 y, Germany	Feeding Study	12 (Acute)	GC-MS/O (Untargeted and targeted)	Urine	3 g raw garlic (equals 1–2 garlic cloves)
Garlic	Allyl methyl sulfide (AMS); allyl methyl sulfoxide (AMSO); allyl methyl sulfone (AMSO2); Dimethyl trisulfide (DMTS)					
Schmedes et al., 2016 (135)	18-65 y, Norway	Crossover Feeding Study	20 (4 weeks)	^{1}H NMR and LC-MS analyses (Untargeted)	Urine	Lean seafood & non-seafood diets
Lean-seafood diet	L-carnitine; 2,6-dimethylheptanoylcarnitine; N-methyl-2-pyridone-5-carboxamide; trimethylamine-N-oxide (TMAO); Hypoxanthine; dimethylamine					
Non-seafood diet	Guanidinoacetate; 3-methylhistidine					
Schmedes et al., 2018 (136)	50.6 ± 3.4 y, Norway	Crossover Feeding Study	19 (4 weeks)	^{1}H NMR and LC-MS (Untargeted)	Serum	Two diets that varied mainly in protein source: lean seafood versus non-seafood proteins
Sea-food diets (vary by protein)	Lean sea-food diet**:** isoleucine; valine; TMAO[1]; lactate[1]; citrate; BCAAs (isoleucine)[1]; Non-sea food diet: certain ceramides[2]; lysophosphatidylcholines; free fatty acids; lysophosphatidylethanolamines; phosphatidic acids; phosphatidylethanolamines; phosphatidylinositols; phosphatidylglycerols; phosphatidylserines; phosphatidylcholines; triacylglycerol (TAG)[1]					
Shi et al., 2017 (137)	23-60 y, Sweden	Crossover Feeding Study	21 (Acute)	^{1}H NMR (36 plasma metabolites) and GC-MS (short chain fatty acids) (Targeted)	Plasma	Plain whole-grain rye porridges (40 and 55 g), rye porridge enriched with different inulin: gluten proportions (9:3 g; 6:6 g; 3:9 g), and a 55 g refined wheat bread (control)
Whole-grain rye porridges	RPHI: porridge. 40 g whole-grain rye + 9 g inulin + 3 g gluten: Acetate; Glucose					

		RPIG: porridge. 40 g whole-grain rye + 6 g inulin + 6 g gluten: Acetate; Leucine; Isoleucine; Valine; 2-oxoisocaproate; Phenylalanine RPHG: porridge. 40 g whole-grain rye + 3 g inulin + 9 g gluten: Isoleucine; Phenylalanine; Leucine; Valine; 2-oxoisocaproate; Lysine				
Sri Harsha et al., 2018 (138)	25 ± 4.2 y, Denmark	Crossover Feeding Study	11 (Acute)	UPLC-Q-ToF-MS (Untargeted)	Urine	Peas
Peas		2-Isopropylmalic acid (2-IPMA); Asparaginyl valine; N-Carbamoyl-2-amino-2-(4-hydroxyphenyl) acetic acid (NC)				
Stalmach et al., 2009 (139)	19-35 y, UK	Feeding Study	11 (Acute)	HPLC-MS^n: MS^2, daughter ions produced from [M-H]– fragmentation; MS^3, daughter ions produced from fragmentation of MS^2 base ion (Targeted – chlorogenic acids)	Urine & Plasma	Coffee containing 412 µmol of chlorogenic acids
Coffee		Urine: 3-*O*-Caffeoylquinic acid-*O*-sulfate; Dihydrocaffeic acid-3-*O*-sulfate; Dihydrocaffeic acid-3-*O*-glucuronide; 4-*O*-Caffeoylquinic acid-*O*-sulfate; Caffeic acid-4-*O*-sulfate; Dihydroferulic acid-4-*O*-sulfate; Caffeic acid-3-*O*-sulfate; Dihydrocaffeic acid; Dihydroferulic acid-4-*O*-glucuron; Ferulic acid-4-*O*-sulfate; 3-*O*-Feruloylquinic acid; Isoferulic acid-3-*O*-sulfate; Dihydro(iso)ferulic acid-3-*O*-glucuronide; Isoferulic acid-3-*O*-glucuronide; Feruloylglycine; 3-*O*-Caffeoylquinic acid lactone-*O*-sulfate; 4-*O*-Caffeoylquinic acid lactone-*O*-sulfate; 4-*O*-Feruloylquinic acid; Dihydroferulic acid; 5-*O*-Feruloylquinic acid Plasma: Dihydrocaffeic acid-3-*O*-sulfate; Caffeic acid-4-*O*-sulfate; Dihydroferulic acid-4-*O*-sulfate; Caffeic acid-3-*O*-sulfate; Dihydrocaffeic acid; Dihydroferulic acid-4-*O*-glucuron; Ferulic acid-4-*O*-sulfate; 5-*O*-Caffeoylquinic acid; 3-*O*-Feruloylquinic acid; Isoferulic acid-3-*O*-sulfate; Isoferulic acid-3-*O*-glucuronide; 3-*O*-Caffeoylquinic acid lactone-*O*-sulfate; 4-*O*-Caffeoylquinic acid lactone-*O*-sulfate; 4-*O*-Feruloylquinic acid; Dihydroferulic acid; 5-*O*-Feruloylquinic acid				
Stalmach et al., 2014 (140)	19-35 y, Scotland	Crossover Feeding Study	11 (Acute)	HPLC-MS^n: MS^2, daughter ions produced from [M-H]– fragmentation; MS^3, daughter ions produced from fragmentation of MS^2 base ion (Targeted – chlorogenic acids)	Urine & Plasma	Single servings of coffee beverage containing low (412 µmol), medium (635 µmol) and high (795 µmol) amounts of chlorogenic acids
Coffee		Feruloylquinic acids; sulphated caffeoylquinic acid lactones				

Stanstrup et al., 2014 (141)	40-68 y, Denmark	Crossover Feeding Study	11 (Acute)	LC-Q-ToF-MS (Untargeted)	Urine & Plasma	Whey (WI), casein (CAS), cod (COD) or gluten (GLU) protein
Whey	Plasma: Leucine/Isoleucine; γ-glutamyl-leucine; Tryptophan Urine: *N*-acetyl-tyrosine					
Casein	Plasma: Methionine sulfoxide; N-phenylacetyl-methionine; Isoleucine; Paracetamol; Threonine; γ-glutamyl-methionine; Lysine; β-hydroxyisobutyric acid; Methionine; γ-glutamyl-valine; Paracetamol sulfate; Kynurenine; Paracetamol glucuronide; α-keto-3-methylvaleric acid; Valine; Citrulline; 3-Hydroxy-2-methylbutyric acid; Glutamic acid; Propionylcarnitine; α-hydroxydecanoic acid; Lauric acid; Myristic acid; Hydroxybutyric acid isomers Urine: *N*-phenylacetyl-Methionine sulfoxide; *N*-phenylacetyl-methionine; *β*-asp-Leu					
Cod	Plasma: TMAO; Creatine; Proline; Arsenobetaine; 1-Methyl-Histidinec and 3-Methyl-Histidinec mixture; 1,2,3,4-Tetrahydro-β-carboline-3-carboxylic acid; Phenylalanine; Taurine; Docosahexaenoic acid Urine: TMAO; N^6,N^6,N^6-trimethyl-lysine; 1,2,3,4-Tetrahydro-*β*-carboline-3-carboxylic acid; Arsenobetaine; 1-Methyl-Histidinec and 3-Methyl-Histidinec mixture					
Gluten	Plasma: Dopamine-3-*O*-sulfate					
Stea et al., 2008 (142)	18–26 y, Norway	Feeding Study	750 (5 months)	HPLC (Targeted –tHcy, cysteine, riboflavin, flavin adenine dinucleotide, flavin mononucleotide, folate and vitamin B12)	Plasma	Vegetables, fruits and bread
Fruits, vegetables, and bread	Total homocysteine (tHcy); cysteine; folate; riboflavin; flavin adenine dinucleotide (FAD)					
Stella et al., 2006 (143)	25-74 y, UK	Crossover Feeding Study	12 males (15 days)	^{1}H NMR (Untargeted)	Urine	Vegetarian, low meat, and high meat diet
High meat diet	Creatinine; creatine; acetylcarnitine; TMAO; taurine; 1- and 3- methylhistidine; glutamine					
Trimigno et al., 2018 (144)	19-31 y, Switzerland	Crossover Feeding Study	11 (Acute)	^{1}H NMR and GC-MS (Untargeted and Targeted (candidate metabolites previously reported such as galactitol, galactonate, and galactono-1,5-lactone	Serum	Isocaloric amounts of milk, cheese, and a soy drink as non-dairy alternative

				(milk), 3-phenyllactic acid (cheese), and pinitol (soy).		
Milk, cheese, and soy drink		Milk: lactose; galactose; galactonate; galactono-1,5-lactone Cheese: methionine; proline; leucine; glutamic acid; 3-phenyllatic acid; methionine; proline; leucine Soy: dodecanoic acid, linoleic acid, y-tocopherol, pinitol; maltol; sucrose; guaiacol; catechol				
Trimigno et al., 2020 (145)	42 y, Denmark	Feeding Study	142 (6 months)	^{1}H NMR (Untargeted)	Urine	New Nordic Diet (NND) & Average Danish Diet (ADD)
New Nordic Diet (NND)		Glycine betaine; Glucose; Glucose, lactose, maltose; Hippurate; Phenylalanine; Taurine; TMAO				
Average Danish Diet (ADD)		Dimethyl sulfone; Propylene glycol; Tartrate				
Tsang et al., 2005 (146)	23-50 y, UK	Feeding Study	20 (2 weeks)	HPLC- MS/MS (Targeted – phenolics)	Plasma	375 ml red wine or control
Red wine		(+)-Catechin glucuronide; (-)-Epicatechin glucuronide; Methyl catechin glucuronide; Methyl epicatechin glucuronide				
Ulaszewska et al., 2016 (147)	26-70 y, UK	Feeding Study	126 (18 weeks)	LS-MS/MS (Untargeted)	Urine	High flavonoid diet, low flavonoid diet, or habitual diet as a control
Flavonoid diet		High flavonoid diet: Coumaroyl malic acid; Dihydroxyphenyl-y-valerolactone-O- sulfate; Fatty Acid Hydroxy-C13:1-GLC (Hydroxy-tridecenoic acid glucuronide (Isomer I)); Fatty Acid Hydroxy-C13:1-GLC (Hydroxy-tridecenoic acid glucuronide (Isomer II)); Fatty Acid Hydroxy-C13:1-GLC (Hydroxy-tridecenoic acid glucuronide (Isomer III)); Fatty Acid Dicarboxylic C14:1-GLC (Dicarboxylic Tetradecenoic Acid Glucuronide); Fatty Acid C15:2-GLC (Pentadienoic acid glucuronide); Fatty Acid Hydroxy-C15:3-GLC (Pentatrienoic acid glucuronide); Fatty Acid Dicarboxylic C17:0-GLC (Dicarboxylic heptadecanoic Acid Glucuronide); Ferulic acid; Iberin N-acetyl-cysteine; Hydroxyphenylacetic acid; Methyl-gallic acid sulfate; Phenylacetic acid; Proline Betaine; Trihydroxyphenyl-y-valerolactone; Vanillic Acid Sulfate Low flavonoid diet: Absicic Acid Glucuronide; Cyclohexadiene carboxylic acid glycine; Cyclohexane carboxylic acid glycine; Trihydroxybenzoic acid (Gallic Acid)				
Ulaszewska et al., 2020 (148)	29–69 y, UK	Crossover Feeding Study	40 (8 weeks)	LC-MS (Untargeted)	Urine & Plasma	Apple
Apple		Urine: Phloretin glucuronide; Phloretin glucuronide sulfate; (Epi)catechin sulfate (I); (Epi)catechin sulfate (II); (Epi)catechin glucuronide; (Epi)catechin methyl sulfate; Methyl (epi)catechin; Hydroxyphenyl-y-valerolactone sulfate; Hydroxyphenyl-y-valerolactone glucuronide (I); Hydroxyphenyl-y-valerolactone glucuronide (II); Dihydroxyphenyl-y-valerolactone sulfate; Dihydroxyphenyl-y-valerolactone glucuronide; Dihydroxyphenyl-y-valerolactone methyl				

						glucuronide; Dihydroxyphenyl-γ-valerolactone glucuronide sulfate; Hydroxyphenyl valeric acid sulfate; Hydroxyphenyl valeric acid-glucuronide; Dihydroxyphenyl valeric acid; Dihydroxyphenyl valeric acid glucuronide; Hydroxy(dihydroxyphenyl) valeric acid glucuronide; Hydroxy(dihydroxy)phenyl valeric acid sulfate; Hydroxy(dihydroxyphenyl) valeric acid methyl; glucuronide Hydroxy(dihydroxy)phenyl valeric acid methyl sulfate; Feruloylquinic acid; Hydroxycinnamic acid; Cinnamoyl glycine; Cinnamic acid glycine; Hydroxyphenylpropionic acid sulfate; Dihydroxyphenyl propionic acid (I); Dihydroxyphenyl propionic acid (II); Hydroxyphenylacetic acid; Hydroxyphenylacetic acid sulfate; Glucosyl-phenyl propionic acid; Hydroxyhippuric acid (I); Hydroxyhippuric acid (II); Hydroxyhippuric acid sulfate; Methoxy-benzylalcohol glucuronide; Hydroxybenzoic acid sulfate; Fatty Acid C10:2; Hydroxy fatty acid glucuronide (OH)C13:0-GLC (I); Hydroxy fatty acid glucuronide (OH)C13:0-GLC (II); Hydroxy fatty acid glucuronide (OH)C13:0-GLC (III); Hydroxy fatty acid glucuronide (OH)C13:1-GLC; Fatty acid glucuronide C13:2-GLC (I); Fatty acid glucuronide C13:2-GLC (II); Fatty acid glucuronide C13:2-GLC (III); Hydroxy fatty acid (OH)C13:2-GLC; Hydroxy fatty acid glucuronide (OH)C13:3-GLC (I); Fatty acid GLC OH-C13:3-GLC (II); Hydroxy fatty acid glucuronide (OH)C15:1-GLC; Fatty Acid Glucuronide C15:2-GLC (I); Fatty acid glucuronide C15:2-GLC(II); Hydroxy fatty acid glucuronide (OH)C15:3-GLC; Hydroxy fatty acid (OH)C15:4; Dicarboxylic fatty acid glucuronide C15:1-GLC; Indole-acetylaspartic acid; Indolylacryloylglycine; Hydroxyphenylacetylglutamine sulfate; Hydroxyphenylacetylglutamine; N-Acetylglutamic acid; Hydroxy-glutaric acid (position 3-OH); Methylglutaconic acid glucuronide <u>Plasma</u>: Hydroxyphenylvaleric acid sulfate; Hydroxyphenylvaleric acid; (Hydroxyphenyl) propionic acid sulfate; (3-Hydroxyphenyl) acetic acid; 3-Hydroxyhippuric acid; Dicarboxylic Fatty Acid C12:3; Dicarboxylic Fatty Acid C14:1; Dicarboxylic Fatty Acid C14:0
Urpi-Sarda et al., 2015 (149)	61 ± 9 y, Spain	Crossover Feeding Study	36 males (4 weeks)	UPLC-MS/MS (Targeted – phenolics)	Urine & Plasma	Red wine or gin, or dealcoholized red wine
Urpi-Sarda et al., 2015	Red wine or gin, or dealcoholized red wine					Urine <u>Red wine</u>: 2,4-Dihydroxybenzoic acid; 2,6-Dihydroxybenzoic acid; 2,5-Dihydroxybenzoic acid; Syringic acid; Gallic acid; Methylgallic acid; Methylgallic sulfate; Ethylgallate; Ethylgallate sulfate; Ethylgallate glucuronide 1; Ethylgallate glucuronide 2; 3-Hydroxyphenylacetic acid; *p*-Coumaric acid; Σ(Epi)catechin glucuronides; Σ(Epi)catechin sulfates; ΣMethyl(epi)catechin glucuronides; ΣMethyl(epi)catechin sulfates; Vanilloylglycine; dihydroxyphenyl-γ-valerolactone (DHPV) 1; DHPV 2; ΣDHPV glucuronides; ΣDHPV sulfates; Resveratrol Biomarker; ΣResveratrol Microbial Metabolites; ΣTotal resveratrol metabolites; Pyrogallol <u>Dealcoholized red wine</u>: 4-Hydroxybenzoic acid; 3-Hydroxybenzoic acid; 2,4-Dihydroxybenzoic acid; 2,6-Dihydroxybenzoic acid; 2,5-Dihydroxybenzoic acid; 3,5-Dihydroxybenzoic acid; Syringic acid; Gallic acid; Methylgallic acid; Methylgallic sulfate; Ethylgallate; Ethylgallate sulfate; Ethylgallate glucuronide 1;

		Ethylgallate glucuronide 2; 3-Hydroxyphenylacetic acid; 2-Hydroxyphenylacetic acid; 3,4-Dihydroxyphenylacetic acid; *p*-Coumaric acid; Caffeic acid; Ferulic acid; 3-(3-Hydroxyphenyl) propionic acid; Dihydrocaffeic acid; Σ(Epi)catechin glucuronides; Σ (Epi)catechin sulfates; ΣMethyl(epi)catechin glucuronides; ΣMethyl(epi)catechin sulfates; DHPV 1; DHPV 2; ΣDHPV glucuronides; ΣDHPV sulfates; MHPV glucuronide; ΣMHPV sulfates; Resveratrol Biomarker; ΣResveratrol Microbial Metabolites; ΣTotal resveratrol metabolites; Enterolactone; Pyrogallol Gin: *m*-Coumaric acid Plasma Red wine: Gallic acid; Methylgallic acid; Methylgallic sulfate; 3-Hydroxyphenylacetic acid; *p*-Coumaric acid; (Epi)catechin glucuronide; Methyl(epi)catechin glucuronide; DHPV 1; DHPV 2; ΣDHPV glucuronides Dealcoholized red wine: Methylgallic sulfate; 3-Hydroxyphenylacetic acid; *p*-Coumaric acid; (Epi)catechin glucuronide				
Van Dorsten et al., 2006 (150)	20-70 y, Netherlands	Crossover Feeding Study	17 males (Acute)	^{1}H NMR (Untargeted)	Urine	Black tea, green tea, or caffeine
Tea (both black & green)		Hippuric acid; 1,3-dihydroxyphenyl-2-O-sulfate (1,3-DHPS)				
Vázquez-Fresno et al., 2015 (151)	53-79 y, Spain	Feeding Study	98 (3 years)	^{1}H-NMR (Untargeted)	Urine	Mediterranean diet supplemented with either extra-virgin olive oil (MD + EVOO) or nuts (MD + Nuts), to those on advice to follow a control low-fat diet (LFD)
Mediterranean diet		Mediterranean diet**:** 3-hydroxybutyrate; amino acids (proline, glycine); the branched-chain amino acid (BCAA) leucine (its derived metabolites (isobutyric acid and 2-oxoisovaleric acid), the threonine metabolite (4-deoxythreonic acid (4-DTEA)); the N-acetyl groups of glycoproteins (N-Ac); gut microbiota cometabolite p-cresol, fatty acid (oleic acid), and its breakdown product (suberic acid) Mediterranean diet supplemented with nuts**:** amino acid glutamine (phenylacetylglutamine and Nacetylglutamine); creatine Mediterranean diet supplemented with either extra-virgin olive oil: creatinine; two intermediates of the tricarboxylic acid cycle (TCA); citrate; cis-aconitate Otherwise, hippurate, trimethylamine-N-oxide, histidine and derivates (methylhistidines, carnosine, and anserine), and xanthosine were predominant after LFD.				

		Low-fat diet: hippurate; trimethyl-N-oxide (TMAO); histidine and its derived metabolites (3-methylhistidine [3-MH], 1-methylhistidine [1-MH]), carnosine, and anserine); proline-betaine; xanthosine				
Vázquez-Manjarrez et al., 2019 (152)	30.0 ± 4.9 y, France	Crossover Feeding Study	12 (3 days)	UPLC-QTOF-MS (24-hour urine samples) & GC × GC-MS (Kinetic urine)	Urine	Banana pulp
Banana		Salsolinol sulfate 1; Methoxyeugenol glucuronide; 2-Isopropylmalic acid; Dopamine sulfate; Salsolinol sulfate 2; N-acetyldopamine sulfate; Mevalonic acid; 6-Hydroxy-1-methyl-1,2,3,4-tetrahydro β-carboline sulfate; 5-Hydroxyindole-3-acetic acid; Xanthurenic acid; Kynurenic acid; 3-Methoxytyramine sulfate 5-Hydroxyindole acetic acid 3TMS; Dopamine 4TMS; Fructose MEOX-5TMS derivative 2; Fructose MEOX-5TMS derivative 1; 3,4-Dihydroxyphenylacetic acid 3TMS; N-methyl-2-pyridone-5-carboxylic acid TMS; Norepinephrine 5TMS; 1,5-Anhydrosorbitol 4TMS; 2-Ethyl-3-hydroxypropionic acid 2TMS; 4-Hydroxy-3-methoxyphenylacetic acid 2TMS				
Vergne et al., 2008 (153)	20-29 y, France	Crossover Feeding Study	12 males (Acute)	HPLC (Targeted – isoflavones)	Urine & Plasma	soya-derivative product containing 35 mg isoflavones, as either supplements or cheese
Soy-based cheese		Daidzein; Genistein				
Vetrani et al., 2016 (154)	54.5 ± 8.6 y, Italy	Feeding Study	72 overweight/ obese individuals (8 weeks)	GC×GC–ToF MS (Targeted – polyphenols)	Urine	High-polyphenol diet (2868 mg/day) or a low-polyphenol (363 mg/day) control diet
Green tea		3-OHBA; 3,4-diOHPPr; ferulic acid; hippuric acid				
Villañoa et al., 2019 (155)	45 ± 6 (males) 41 ± 6 (pre-menopausal women) 55 ± 3 (post-menopausal women)	Feeding Study	69 (5 weeks)	UHPLC–MS/MS (Targeted – glucosinolates)	Urine	Broccoli sprouts
Broccoli sprouts		Sulphoraphane-N-acetylcysteine (SFN-NAC); sulphoraphane-cysteine (SFN-CYS); sulphoraphane (SFN); 3–3-diindolylmethane (3,3'-DIM)				

Wang et al., 2019 (156)	21-65 y, US	Crossover Feeding Study	113 (4 weeks)	HPLC-MS/MS (Targeted – TMAO)	Urine & Plasma	Red meat, white meat, or non-meat
Red meat	Urine & Plasma: Trimethylamine N-oxide (TMAO) Plasma: choline; betaine; carnitine; γ-butyrobetaine; crotonobetaine					
Wedekind et al., 2019 (157)	31 ± 5.2 y, –	Crossover Feeding Study	12 (3 days)	LC-MS (Untargeted)	Urine	3 processed meat products (bacon, salami, and hot dog)
Smoked meat product (hot dog)	Syringol sulfate; 4-Methylsyringol sulfate; 4-Ethylsyringol sulfate; 4-Allylsyringol sulfate isomer II					
Wedekind et al., 2020 (158)	31 ± 5.2 y, Finland	Crossover Feeding Study	12 (3 days)	LC-MS (Untargeted)	Urine & Serum	Fried fresh pork versus tofu
Fried fresh pork	Acylcarnitines (C0, 2:0, 3:0, 4:0 (OH), 5:0, 6:0, 6:0 (OH), 6:0 (DC); 7:0, 8:0 (OH), 8:0 (OH.2), 8:2 (OH), 10:4, 10:0 (OH), 11:1, 12:2 (OH), 12:1 (OH)) Note: OH = hydroxyl group on fatty acid moiety; DC = dicarboxylic acid					
Wellington et al., 2019 (159)	47 y, Canada	Feeding Study	42 (2 weeks)	MSI-CE-MS (Untargeted) LC-MS/MS (Untargeted) GC-MS (Targeted – fatty acid)	Urine & Plasma	Prudent or Western diets
Prudent diet	Plasma: 3-methylhistidine; proline betaine; ketoleucine; ketovaline Urine: 3-methylhistidine; proline betaine; imidazole propionate; hydroxypipecolic acid; dihydroxybenzoic acid; enterolactone glucuronide					
Western diet	Plasma: myristic acid; linoelaidic acid; linoleic acid; alpha-linoleic acid; pentadecanoic acid; alanine; proline; carnitine; deoxycarnitine Urine: acesulfame K					
Wiczkowski et al., 2016 (160)	29 ± 5 y, Poland	Crossover Feeding Study	13 (2 days)	HPLC–MS/MS (Targeted – anthocyanins)	Urine & Plasma	Fresh and fermented red cabbage
Red cabbage	Native anthocyanins: Cyanidin-3-diglucoside-5-glucoside; 2 Cyanidin-3-glucoside-5-glucoside; Cyanidin-3-(sinapoyl)-diglucoside-5-glucoside; Cyanidin-3-(sinapoyl)-triglucoside-5-glucoside; Cyanidin-3-(caffeoyl)(p-coumaroyl)-diglucosides-5-glucoside; Cyanidin-3-(feruloyl)-triglucosides-5-glucoside; Cyanidin-3-(sinapoyl)-triglucoside-5-glucoside; Cyanidin-3-(feruloyl)(feruloyl)-triglucoside-5-glucoside; Cyanidin-3-(feruloyl)(sinapoyl)-triglucoside-5-glucoside; Cyanidin-3-(caffeoyl)-diglucoside-5-glucoside; Cyanidin-3-(p-coumaroyl)-diglucoside-5-glucoside; Cyanidin-3-(feruloyl)-diglucoside-5-glucoside; Cyanidin-3-					

		(sinapoyl)-diglucoside-5-glucoside; Cyanidin-3-(feruloyl)-glucoside-5-glucoside; Cyanidin-3-(sinapoyl)-glucoside-5-glucoside; Cyanidin-3-(feruloyl)(feruloyl)-diglucoside-5-glucoside; Cyanidin-3-(feruloyl)(sinapoyl)-diglucoside-5-glucoside; Cyanidin-3-(sinapoyl)(sinapoyl)-diglucoside-5-glucoside <u>Anthocyanin metabolites:</u> Cyanidin 3-glucoside; Peonidin; Peonidin 3-glucoside; Cyanidin monoglucuronide; Methylated cyanidin diglucoside; Methylated cyanidin triglucoside; Methylated cyanidin monoglucuronide; Methylated cyanidin glucoside monoglucuronide sulfate; Methylated cyanidin (p-coumaroyl)-triglucoside; Methylated cyanidin (caffeoyl)-triglucoside; Methylated cyanidin (feruloyl)-triglucoside; Methylated cyanidin (sinapoyl)-triglucoside				
Wiseman et al., 2004 (161)	24 ± 6.7 y, UK	Feeding Study	76 (10 weeks)	LC-MS (urine) and GC–MS (plasma) (Targeted – isoflavones)	Urine & Plasma	High-soy diet (104 ± 24 mg total isoflavones/d) with a low-soy diet (0.54 ± 0.58 mg total isoflavones/d)
High-soy diet		Genistein; Daidzein; Equol; *O*-desmethylangolensin (*O*-DMA)				
Yin et al., 2017 (162)	62 ±1 y, England	Feeding Study	10 (3 weeks)	^{1}H NMR (urine) and HPLC-MS (plasma) (Untargeted and targeted)	Urine & Plasma	Chicken, in increasing amounts from 88 to 290 g/d
Chicken		<u>Urine</u>: Guanidoacetate <u>Plasma</u>: 3-methylhistidine				
Yin et al., 2020 (163)	62 ± 1 y, London	Feeding Study	10 (3 weeks)	^{1}H NMR (Untargeted)	Urine	Fish
Fish		TMAO; dimethylamine; dimethyl sulfone				
Zamora-Ros et al., 2006 (164)	Males: 28.2 ± 7.3 y Females: 38.1 ± 9.2 y, Spain	Feeding Study	20 (28 days)	LC–MS/MS (Targeted – resveratrols)	Urine	30 g of ethanol/day as sparkling wine or gin for 28 days
Wine		Total resveratrol metabolites (trans and cis-resveratrol-3-O-glucuronide) – red wine > white wine > sparkling				
Zhao et al., 2020 (165)	21–29 y, US	Crossover Feeding Study	16 (21 days)	UHPLC-Q-Orbitrap-HRMS (Targeted)	Plasma	Cranberry Juice
Cranberry juice		Quinic acid; 3-(Hydroxyphenyl) propionic acid; (S)-Homostachydrine; Glycerol 3-phosphatel; Dihydroxyquinolin; Ethyl (methylthio)methyl disulfide; Hippuric acid; Hydroxypyruvic acid; 3,4-Dihydroxyphenylglycol; Guanidoacetic acid;				

	Catechol sulfate; 2-Phenylacetamide; Tyrosine; 3-Isopropylmalate; 2-Chloromaleylacetate; Lanthionine ketimine; S-Acetyl dihydroasparagusic acid; Pyrocatechol; Prolyl-Hydroxyproline; Guaiacol; Tyramine-O-sulfate; (3,4,5,6-tetrahydroxyoxan-2-yl)methyl 4-hydroxybenzoate; (4-[2,3-dioxo-3-(2,4,6-trihydroxy-3-methoxyphenyl)propyl]-2-hydroxy-6-methoxyphenyl)+oxidanesulfonic acid					
Zheng et al., 2015 (166)	12-15 y, Denmark	Feeding Study	192 (12 weeks)	^{1}H NMR (Untargeted)	Urine	1 L/day of casein (citrate content: 3.27 mol/L), whey (citrate content: 0.04 mol/L), skim milk, or water
Cheese and milk	Cheese: proline betaine; tyrosine levels; creatine; creatinine; choline; TMAO Milk: citrate; hippurate; urea; TMAO					
Zheng et al., 2015 (167)	18-50 y, Denmark	Crossover Feeding Study	15 males (14 days)	^{1}H NMR (Untargeted)	Urine	Three isocaloric diets with similar fat contents (i) high in milk, (ii) high in cheese with equal amounts of dairy calcium, or (iii) a control diet
Casein	Urea					
Skim milk	Urea					
Whey	Citrate					
Zheng et al., 2016 (168)	18-60 y, Denmark	Feeding Study	38 overweight/ obese females (24 weeks)	^{1}H NMR (Untargeted)	Urine & Plasma	Dairy intake (high vs. low intake)
High dairy intake	Citrate; creatinine; urea; trimethylamine-N-oxide (TMAO); hippurate					
Zhong et al., 2017 (169)	20-45 y, US	Crossover Feeding Study	12 (Acute)	UHPLC-Q-ToF-MS/MS (Untargeted and targeted (anthocyanins, cholinergic acid & their metabolites))	Plasma	Wild blueberries (WBB) beverage (25 g freeze dried WBB powder)
Wild blueberries (WBB)	3-CGA (3-chlorogenic acid); peonidin glycoside; delphinidin glycoside; cyanidin glycoside; petunidin glycoside					

APPENDIX II – Study Characteristics of Observational Studies and Metabolites of Foods and Food group

Supplemental Reference	Study population, location	Study design	N (follow-up)	Technique (Method)	Biofluid	Food category and/or food item	Dietary assessment tool
Aguilar et al., 2014 (170)	5-17 y, US	Cross-sectional	45	HPLC (Targeted – carotenoids)	Serum	Fruits and vegetables intake	27-itemFFQ and an automated multiple-pass 24-hour daily recall
Fruits and vegetables	Total carotenoid; beta carotene						
Azab et al., 2019 (171)	32 y, Canada	Cross-sectional	50	MSI-NACE-MS (Targeted)	Serum	Dietary fat intake	Semiquantitative FFQ
Total ω-3 PUFAs	EPA (20:5n–3); DHA (22:6n–3); [EPA + DHA]						
Fish/Seafood	EPA (20:5n–3); DHA (22:6n–3); [EPA + DHA]						
Healthy nutrient-rich foods	EPA (20:5n–3); DHA (22:6n–3); [EPA + DHA]						
Full-fat dairy intake	Myristic acid (14:0); Pentadecanoic acid (15:0)						
Allen et al., 2008 (172)	20–69 y, UK	Cross-sectional	96 females	GC–MS (Targeted – fatty acids)	Plasma	Meat	Semi-quantitative FFQ
Meat eaters	phytanic acid; pentadecanoic acid; heptadecanoic acid						

Aubertin-Leheudre et al., 2010 (173)	46 ± 13 y, Finland	Cross-sectional	56 females	HPLC (Targeted – alkylresorcinols)	Serum	Cereal fibre intake	3-d dietary record
Cereal fibre intake	3,5-dihydroxybenzoic acid (DHBA); 3-(3,5-dihydroxyphenyl)-1-propanoic acid (DHPPA)						
Bouchard-Mercier et al., 2013 (174)	34.6 ± 9.2 y, Canada	Cross-sectional	37	LC-MS (Targeted – 14 amino acids and 41 acylcarnitines)	Plasma	The Prudent dietary pattern (higher intakes of vegetables, fruits, whole grain products, non-hydrogenated fat and lower intakes of refined grain products; Western dietary pattern (higher intakes of refined grain products, desserts, sweets and processed meats)	91-items FFQ
PC1 and/or PC2							
Western Diet	PC1; PC2						
Fruits	PC2						
Desserts	PC2						
Saturated fat	PC2						
Total fat	PC2						
Individual metabolites							
Vegetables and fruits	octadecadienyl-L-carnitine; xleucine						
Fruits only	methionine						
Non-hydrogenated fats	tetradecadienyl-L-carnitine; octadecadienyl-L-carnitine; histidine						

Desserts		3 amino acids (methionine, phenylalanine, xleucine); methionine; hydroxyoctadecenoyl-L-carnitine; glutaconyl-L-carnitine					
Saturated fat		valeryl-L-carnitine; octadecadienyl-L-carnitine					
Monounsaturated fat		octenoyl-L-carnitine; methylglutaryl-L-carnitine					
Polyunsaturated fat		methylglutaryl-L-carnitine; proline; decadienyl-L-carnitine					
Protein		ornithine; histidine					
Carbohydrates		ornithine					
Chandler et al., 2020 (175)	67–68 y, US	Cross-sectional	904	LC–MS/MS (Untargeted)	Serum	Western and Prudent dietary patterns	122-item FFQ
Prudent dietary pattern		C58:8 TAG; C58:9 TAG; C60:12 TAG; eicosapentaenoate; C40:10 PC; C38:6 PC; C20:5 CE; C38:4 PC plasmalogen-B; C58:1 TAG; C40:6 PC-B; tetradecanedioate; docosahexaenoic acid; C56:9 TAG; C58:10 TAG; C56:8 TAG; C56:10 TAG; C36:5 PC C36:1 PC plasmalogen; uracil; C22:6 CE; C22:6 LPC; C18:2 SM; C36:2 PC plasmalogen-A; C34:0 PS; C38:7 PC plasmaloge uridine; C58:7 TAG; C22:6 LPE-B; C56:7 TAG; C34:0 PC plasmalogen; C54:9 TAG; C54:8 TAG; C38:6 PE; C7 carnitine; C38 PC plasmalogen-A; indole-3-propionate; C36:5 PC plasmalogen-A; C38:7 PE plasmalogen; C18:0 LPC plasmalogen; pipecolic acid; docosatrienoic acid; C22:6 LPE-A; C12 carnitine; C34:5 PC; NMMA; thiamine; 4-pyridoxate; C38:6 PC plasmalogen-A; C16:1 CE; C10 carnitine; C36:3 PC plasmalogen-A; C18:2 LPE; C8 carnitine; C36:3 PE plasmalogen; C12:1 carnitine; N-acetylornithine; C18 carnitine; trimethylbenzene; hippurate; C4-OH carnitine; C40:6 PE; C54:7 TAG; C36:4 plasmalogen; beta-hydroxybutyrate; pantothenate; C52:7 TAG; C20 carnitine; C23:0 Ceramide (d18:1); C6 carnitine; N-alpha-acetylarginine; myristoleic acid; C22:1 MAG; C36:2 PE plasmalogen; sebacate; N-carbamoyl-beta-alanine; C36:2 C22:0 Ceramide (d18:1); C38:1 PC; C14 carnitine; 7-methylxanthine; C24:1 SM; C36:3 PC plasmalogen-B; C34:2 PC plasmalogen-A; C34:3 PE plasmalogen; C16:0 CE; C14:2 carnitine; hypoxanthine; 2-aminooctanoate; C20:4 LPE-A; C36:1 PS plasmalogen; C16:0 LPE; C3 carnitine; C14:1 carnitine; homoarginine; C36:2 PC plasmalogen-B; UDP-glucuronate; docosapentaenoic acid; eicosanedioate; C32:1 PC plasmalogen-A; C18:2 LPC; ectoine					
Western dietary pattern		Tetradecanedioate; C16:1 CE; C38:4 PC plasmalogen-B; C34:2 PC plasmalogen-A; C18:2 SM; C36:2 PC plasmalogen-A; C carnitine; C36:3 PC plasmalogen-B; C34:3 PE plasmalogen; C36:3 PE plasmalogen; C36:3 PC plasmalogen-A; C36:4 PE plasmalogen; docosatrienoic acid; N-methylproline; proline betaine; C38:2 PE; C58:8 TAG; C36:1 PC plasmalogen; docosahexaenoic acid; eicosapentaenoate; C4-OH carnitine; C14:0 CE; C38:6 PC plasmalogen-B; C10:2 carnitine; C20:5 CE; C36:2 PE plasmalogen; C20:3 CE; Trimethylbenzene; C34:5 PC plasmalogen; C60:12 TAG; C36:2 PS plasmalogen; indole-3-propionate; C16:0 LPE; C32:1 PC; C58:9 TAG; C36:2 PC; C51:3 TAG; C23:0 Ceramide (d18:1); eicosanedioate; C36:2 PC plasmalogen-B; beta-hydroxybutyrate; glucose; C34:0 PC plasmalogen; C18:0 LPC plasmalogen; C38:6 PC; C9 carnitine; C36:5 PC; C36:4 PC plasmalogen-B; C34:3 PC plasmalogen; C36:4 DAG-A; C52:4 TAG; C58:7 TAG; C38:6 PE; C40:6 PC-B; C40:10 PC; C22:6 LPE-B; C8 carnitine					

Cheung et al., 2017 (16)	54.2 ± 9 y, 10 European countries	Cross-sectional	481	LC-MS/MS (Targeted)	Urine	Chicken, red meat, processed meat, and fish	24-h dietary recall
Fish		Observational (urine): acetylcarnitine; trimethylamine-N-oxide (TMAO)					
Poultry/chicken		Observational (urine): 3-methylhistidine, anserine, carnosine, acetylcarnitine					
Processed meat (cooked ham)		Observational (urine): carnosine; acetylcarnitine					
Red meat		Observational (urine): carnosine; acetylcarnitine					
Chun et al., 2009 (176)	19 y or older, US	Cross-sectional	2,908	(HN-APCI)-MS (Targeted – isoflavones)	Urine	Isoflavones	24-hour dietary recall
Isoflavones		Genistein; Daidzein; *O*-desmethylangolensin (*O*-DMA)					
Chung et al., 2009 (177)	31.9 ± 2.0 y, US	Cross-sectional	25	HPLC (Targeted – carotenoids)	Serum	Dietary carotenoid	100-item Health Habits and History FFQ
Dietary carotenoid		α-carotene; β-carotene; β-cryptoxanthin; total carotenoids					
Cuparencu et al., 2020 (23)	– New Zealand	Cross-sectional	158	UPLC-ESI-q-TOF-MS (Untargeted)	Urine	Meat-related food groups, in servings: red meat (e.g., beef, pork, lamb), poultry (e.g., chicken, turkey), seafood (e.g., fish and shellfish high or low in n-3 LC-PUFA), processed meats and organ meats.	Food diaries
Poultry		Carnosine; anserine; 3-methylhistidine					
Red meat		Carnosine; anserine					
De Filippis et al., 2016 (178)	Vegetarians 39 ± 9 y Vegans	Cross-sectional	153	GC-MS/SPME (Targeted – short chain fatty acids (SCFA))	Urine	Mediterranean diet	7-day weighed food diary

	37 ± 10 y Omnivores 37 ± 9 y						
Mediterranean diet		Trimethylamine oxide (TMAO)					
de Pee et al., 1999 (179)	27 ± 6 y, Indonesia	Cross-sectional	600	HPLC (Targeted – retinol)	Serum	Vitamin A	Semi-quantitative 24 h recall questionnaire
Vitamin A		Retinol					
Dorgan et al., 2019 (180)	25-29 y, US	Cross-sectional	211 females	UPLC-MS/MS (Untargeted)	Serum	Alcohol	Questionnaire or three non-consecutive 24-h dietary recalls
Alcohol		Amino acid (sarcosine) Lipid (eicosapentaenoate (EPA; 20:5n3)) Steroid (4-Androsten-3beta,17beta-diol monosulfate (2)) Cofactor/Vitamin (gamma-carboxyethyl hydroxychroman (CEHC))					
Du et al., 2020 (181)	26-36 y, Australia	Cross-sectional	1785	NMR (Targeted)	Serum	Alcohol	12-month FFQ
Alcohol		Fatty acids: total fatty acids (FAs); saturated FAs (SFAs); MUFAs; omega-3 FAs; docosahexaenoic acid (DHA); omega-3 FAs; DHA levels to total FAs; omega-6 FA ratio; polyunsaturated FA (PUFA) ratio; linoleic acid ratio to total FAs Glycine; isoleucine; valine; phenylalanine; citrate					
Edmands et al., 2015 (182)	55.3 ± 8.4 y 10 European Countries	Cross-sectional	481	HPLC-Q-ToF-MS (Untargeted)	Urine	6 polyphenol-rich foods (coffee, tea, red wine, citrus fruit, apples and pears, and chocolate products)	24-h dietary recalls (acute) and FFQ (habitual)
Citrus fruit		Naringenin glucuronide; Hesperetin glucuronide sulfate; Hesperetin glucuronide (I);					
Apples & pears		Methyl(epi)catechin sulfate (I); Phloretin glucuronide; Dihydroxyphenyl-y-valerolactone sulfate					
Chocolate		Methyl(epi)catechin sulfate (I); 4-Hydroxy-(3',4'-dihydroxyphenyl)valeric acid sulfate; Dihydroxyphenyl-y-valerolactone glucuronide; Vanillic acid sulfate					

Coffee		Dihydroferulic acid sulfate; Guaiacol glucuronide; Feruloylquinic acid; Ferulic acid sulfate; Feruloylquinic acid glucuroni 3-*O*-Caffeoylquinic acid; *p*-Coumaric acid sulfate; Caffeic acid sulfate; Ferulic acid glucuronide; Hydroxyhippuric acid; Dihydrocaffeic acid sulfate; m-Coumaric acid sulfate; Dihydroferulic acid glucuronide; *p*-Hydroxyphenyllactic acid; Guaiacol sulfate; Ethylcatechol glucuronide					
Tea		Methylgallic acid sulfate (I); 4-*O*-Methylgallic acid; Dihydroxyphenyl-g-valerolactone sulfate; Pyrogallol sulfate (I); Hydroxyphenylvaleric acid glucuronide; Methyl(epi)catechin sulfate (I); Dihydroxyphenyl-y-valerolactone glucuronide					
Red wine		*m*-Coumaric acid sulfate; Gallic acid ethyl ester sulfate; Hydroxytyrosol sulfate; Dihydroresveratrol glucuronide; Syringi acid sulfate; Methylgallic acid sulfate (I); 4-O-Methylgallic acid					
Floegel et al., 2013 (183)	49.8 ± 8.9 y, Germany	Cross-sectional	2,380	FIA–MS/MS (Targeted – acylcarnitines, amino acids, hexose and choline-containing phospholipids)	Serum	45 food groups	148-item FFQ
Cornflakes, crisps (Positive loading)		Acylcarnitines (C14:1 (2.02), C2 (1.97), C18 (1.84), C14:2 (1.83), C18:1 (1.70), C7-DC (1.30), C16 (0.99),C16:2 (0.95), C6-OH/C5-DC (0.81), C10 (0.78), C18:2 (0.54), C8:1 (0.46), C0 (0.40), C10:2 (0.40), C5-OH/ C3-DCM (0.17), C9 (–0.06), C3 (0.05)					
Fish, other vegetable fat, whole grain bread, cooked vegetables, garlic, nuts, tea, cabbage, sweet bread spreads, cake, cookies, high-fat cheese (Negative loading)							
Canned fruit, fried potatoes, legumes, cake, cookies (Positive loading)		Amino acids (Ser (1.34), Tyr (–1.02), Gly (0.67), Thr (0.65), Val (–0.61), His (0.25), Ile (–0.21), Pro (–0.13), Gln (0.02), Me (–0.01), Phe (0), Orn (0), Trp (0), Arg (0)					
Water, low-fat cheese, fish, whole grain bread, grain flakes, muesli (Negative loading)							
Sauce, butter (Positive loading)		Diacylphosphatidylcholines (C36:1 (3.67), C28:1 (3.51), C34:1 (3.08), C30:0 (3.00), C32:1 (2.73), C40:4 (2.18), C42:2 (–1.57), C32:2 (1.44), C40:5 (1.42), C34:4 (1.37), C36:0 (–1.37), C42:0 (–1.32), C38:0 (–1.28), C42:1 (–1.23), C32:0 (1.23), C36:3 (1.04), C34:3 (1.03), C40:3 (–0.99), C38:3 (0.79), C40:2 (0.67), C38:6 (–0.62), C38:5 (0.58), C36:4 (0.56), C38:1 (–0.50), C38:4 (0.49), C40:6 (–0.41), C32:3 (0.35), C36:2 (0.06), C36:5 (–0.04), C42:4 (0.03), C42:6 (0.03), C34:2 (0.01), C42:5 (0), C36:6 (0)					
Fish, whole grain bread, tea, grain flakes, muesli (Negative loading)							

Red meat, processed meat, poultry, margarine, non-wholegrain bread (Positive loading)				Acyl-alkylphosphatidylcholines (C36:4 (7.43), C38:5 (7.18), C36:5 (6.57), C38:6 (4.04), C30:0 (–3.89), C36:3 (3.22), C36:1 2.96), C34:0 (–2.86), C34:3 (2.78), C34:1 –1.56), C34:2 (1.49), C38:4 (1.24), C42:1 (1.08), C44:6 (0.81), C40:1 (0.66), C38:3 (–0.64), C44:5 (0.63), C42:5 (0.62), C40:4 (0.61), C36:2 (–0.(C30:2 (–0.57), C40:2 (–0.56), C30:1 (–0.34), C40:5 (0.31), C44:3 (0.30), C40:3 (–0.23), C36:0 (0.22), C42:4 (0.18), C42:3 (0.17), C38:1 (–0.16), C40:6 (0.07), C38:2 (0.05), C32:1 (0.05), C38:0 (0.02), C42:2 (–0.01), C32:2 (–0.01), C44:4 (0)				
Sweet bread spreads, butter, desserts, high-fat dairy products, tea, vegetarian dishes (Negative loading)							
Margarine, non-whole grain bread, processed meat, red meat, coffee (Positive loading)				Lysophosphatidylcholines C20:4 (2.91), C18:2 (2.19), C18:0 (1.36), C20:3 (1.26), C17:0 (–0.97), C28:1 (–0.79), C16:0 (0.(C18:1 (0.63), C14:0 (–0.56), C16:1 (–0.11)			
Butter, pasta, rice, tea, desserts, soup (Negative loading)							
Butter, sweet bread spreads, high-fat cheese, fresh fruit, whole grain bread, desserts, cake, cookies, high-fat dairy products (Negative loading)				Sphingomyelins C24:1 (4.41), OH-C14:1 (–1.95), C24:0 (1.95), C26:1 (1.70), C16:0 (1.45), C18:0 (1.44), C18:1 (1.21), C16 (1.11), OH-C16:1 (–0.75), OH-C22:2 (–0.09), OHC24:1 (–0.03), C26:0 (0.01), C20:2 (0), OH-C22:1 (0)				
Frankenfeld et al., 2003 (184)	61 ± 8 y, US	Cross-sectional	96 post menopausal females	LC-MS (Targeted – isoflavone)	Serum	Soy-based foods	122-item FFQ
Soy intake				Daidzein; Genistein			
Fiber				Daidzein; Genistein			
Caffeine				Daidzein			
Fruit and vegetables				Genistein			
Frankenfeld 2011 (185)	45.0 + 0.5 y, US	Cross-sectional	3,115	HPLC-MS/MS (Targeted – isoflavone and daidzein metabolite concentrations)	Urine	Dairy consumption	24-h dietary recall

Legumes, nuts, and seeds		Daidzein; Genistein; *O*-desmethylangolensin (ODMA)					
Grain products		Daidzein; Genistein; Equol; *O*-desmethylangolensin (ODMA)					
Fruits		Daidzein; Equol; *O*-desmethylangolensin (ODMA)					
Milk and milk products		Equol					
High soy		Daidzein; Genistein; *O*-desmethylangolensin (ODMA)					
Garcia-Aloy et al., 2014 (186)	55-80 y, Spain	Cross-sectional	195	HPLC-Q-ToF-MS (Untargeted)	Urine	Walnut consumption	9 category FFQ
Garcia-Aloy et al., 2014	Walnut	3-indolecarboxylic acid glucuronide; hydroxyindoleacetic acid sulfate; N-acetylserotonin sulfate; 10-hydrc decene-4,6-diynoic acid sulfate; urolithin C glucuronide; urolithin A glucuronide; urolithin A sulfoglucuronide; tridecadienoic/ tridecynoic acid glucuronide; urolithin B glucuronide; enterolactone glucuronide; urolithin C sulfate; urolithin A sulfate					
Garcia-Aloy et al., 2015 (187)	67.0 ± 6.3 y, Spain	Cross-sectional	64	HPLC-Q-ToF-MS (Untargeted)	Urine	Cocoa or derived products	137-item FFQ
Cocoa		5-acetylamino-6-amino-3-methyluracil (AMMU); 3-methyluric acid; 7- and 3-methylxanthine; 3,7-dimethyluric acid; theobromine; methoxyhydroxyphenylvalerolactone; 5-(3',4'-dihydroxyphenyl)-valerolactone (DHPV) glucuronide; 5-(3',4'-dihydroxyphenyl)-valerolactone (DHPV) glucuronide sulphate					
Garcia-Aloy et al., 2015 (188)	68 ± 6 y, Spain	Cross-sectional	155	HPLC–Q-ToF-MS (Untargeted)	Urine	Bread	Validated semi-quantitative 137-item FFQ
Whole grain bread		2-Aminophenol sulphate; N-(2-hydroxyphenyl) acetamide (HPAA) glucuronide; 2-hydroxy-N-(2-hydroxyphenyl) acetam (HHPAA)[1]; 2-hydroxy-1,4-benzoxazin-3-one (HBOA) glycoside[1]; 2-hydroxy-N-(2-hydroxyphenyl) acetamide (HPPA); 2-hydroxy-7-methoxy-2H-1,4-benzoxazin-3-one (HMBOA); 3-(3,5-dihydroxyphenyl) propanoic acid (DHPPA) glucuronide[1] 3,5-Dihydroxyphenylethanol sulphate; 5-(3,5-dihydroxyphenyl) pentanoic acid (DHPPTA) sulphate[1]; Hydroxybenzoic ac glucuronide; Dihydroferulic acid sulphate[1]; Enterolactone glucuronide[1]; Pyrraline; 3-Indolecarboxylic acid glucuronide[1] Riboflavin[1]; N-a-Acetylcitrulline; 2,8-Dihydroxyquinoline glucuronide[1]					
White-bread consumers		2-Aminophenol sulphate; N-(2-hydroxyphenyl) acetamide (HPAA) glucuronide; 2-hydroxy-7-methoxy-2H-1,4-benzoxaz 3-one (HMBOA) glucuronide; 2-hydroxy-7-methoxy-2H-1,4-benzoxazin-3-one (HMBOA); 3-(3,5-dihydroxyphenyl) propanoic acid (DHPPA) glucuronide; Hydroxybenzoic acid glucuronide; Riboflavin					
Gibbons et al., 2015 (45)	47 ± 16 y, Ireland	Cross-sectional/ RCT	565	^{1}H NMR (Untargeted)	Urine	Sugar-sweetened beverages	4-day semi-weighed food record

Sugar-sweetened beverage		Amino Acids: citrulline; taurine Other: formate isocitrate					
Gibbons et al., 2017 (189)	Cluster 1 52 ± 16 y Cluster 2 42 ± 15 y, Ireland	Cross-sectional	567	^{1}H NMR (Untargeted)	Urine	Healthy versus unhealthy dietary clusters	4-day semi-weighed food diaries
Healthy cluster		Higher intakes of breakfast cereals, low fat and skimmed milks, potatoes, fruit, fish and fish dishes: Hippurate; N-phenylacetylglutamine; Anserine; 3-hydroxybutyrate; 2-aminoadipate; Citrate					
Unhealthy cluster		Higher intakes of chips/processed potatoes, meat products, savory snacks and high-energy beverages: Creatinine; Glycylproline; Theophylline; N-acetylglutamate; Tryptophan					
Gibson et al., 2020 (190)	40-59 y, Japan, China, the United Kingdom, and United States	Cross-sectional	4680	NMR	Urine (Targeted)	Fish	24-h dietary recalls
Fish		TMAO; taurine; creatine; homarine; ethyl glucuronide; trimethyllysine; dimethylamine					
Shellfish		Homarine					
Griep et al., 2016 (191)	40-59 y, US & UK	Cross-sectional	2,032 (US) 449 (UK)	^{1}H NMR (Targeted – hippurate and proline betaine)	Urine	Fruits intake (including 100% fruit juices)	4 in-depth 24-hr dietary recalls
Fruits		Hippurate; proline betaine					
Guertin et al., 2014 (192)	64 ± 5 y, US	Cross-sectional	502	UHPLC–MS/MS and GC–MS (Untargeted)	Serum	36 dietary groups	137-item FFQ
Citrus		Stachydrine; Chiro-inositol; Scyllo-inositol; N-methyl proline					
Berries		1-Palmitoylglycero-phospho-inositol					
Apples, pears		13-HODE + 9-HODE					
Melon		Pregnenolone sulfate					

Bananas	y-Tocopherol
Other	Pyridoxate
Cruciferous	a-CEHC glucuronide
Greens	3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF)
Yellow/orange vegetables	Creatinine
Starchy vegetables	Cyclo (-Leu-Pro)
Alliums (garlic, onions)	3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF)
Other	Docosahexaenoic acid (DHA)
Red meat	Indolepropionate
Poultry: chicken	Pyroglutamine
Fish (excluding shellfish)	3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF); docosahexaenoic acid (DHA); eicosapentaenoic acid (EPA) 1-Docosahexaenoylglycero-phosphocholine
Shellfish	3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF)
Processed meat	Lathosterol
Baked sweets	Glutamine
Chocolate	Theobromine
Candy (non-chocolate)	Leucylleucine
Chips	Docosahexaenoic acid (DHA)
Tofu	4-Ethylphenylsulfate
Beans	*S*-Methylcysteine
Eggs	Indolepropionate
Added fats	δ-Tocopherol
Butter	Methyl palmitate (15 or 2); Pentadecanoate (15:0); 10-Undecenoate (11:1n–1);
Peanuts	Tryptophan betaine; 4-Vinylphenol sulfate
Rice (white)	Docosahexaenoic acid (DHA)
Dairy: milk	Homostachydrine

Coffee		Quinate; 1-Methylxanthine; Paraxanthine; N-2-furoyl-glycine; Catechol sulfate					
Sugar-sweetened beverages		Quinate					
Beer		16-Hydroxypalmitate					
Wine		Scyllo-inositol					
Liquor		Ethyl glucuronide					
Total alcohol[1]		Ethyl glucuronide; 4-Androsten-3b,17b-diol disulfate 1; 5-a-Androstan-3b,17b-diol disulfate; Cyclo (-Leu-Pro); Bilirubin (E,Z or Z,E); 16-Hydroxypalmitate; Dihomo-linoleate (20:2n–6); Palmitoleate (16:1n–7)					
Vitamins/supplements Multivitamins		Pantothenate; Pyridoxate; α-Tocopherol; γ-Tocopherol; Threonate; β-Tocopherol					
Hanhineva et al., 2015 (193)	44 ± 17 y, Sweden	Cross-sectional	91	LC-Q-ToF-MS (Untargeted)	Urine	Whole grain rye intake	3-day weighed food records
Whole grain rye intake		Hydroxyhydroxyphenyl acetamide (HHPAA) sulfate; 3,5-dihydroxyphenylpropionic acid sulfate (3,5- DHPPA); caffeic ac sulfate; hydroxyphenyl acetamide (HPAA) sulfate					
Harada et al., 2016 (194)	35–74 y, Japan	Cross-sectional	896 males	CE-ToF-MS (Untargeted)	Serum	Alcohol	Self-administered questionnaire
Alcohol		2-Oxoglutarate; Arg; Carnitine; Gln; Guanidinosuccinate; Hippurate; Hypoxanthine; Kynurenine; Lys; Malate; Malonate Mucate; Octanoate; Phthalate; Quinate; Trigonelline; 2-Aminobutyrate; 2-Hydroxybutyrate; 4-Methyl-2-oxopentanoat Alpha-Aminoadipate; Choline; Citrate; Creatine; CSSG; Glycerophosphorylcholine; Hydroxyproline; Ile; Leu; N,N-Dimethylglycine; Ornithine; Pipecolate; Taurine; Thr; Threonate; Tyr; Val					
Hodgson et al., 2006 (195)	70-85 y, Australia	Cross-sectional	232 females	GC-MS (Targeted – 4-O-methylgallic acid)	Plasma	Tea (black or green)	Interviewer-administered 24-h dietary recall
Tea (black or green)		4-*O*-methylgallic acid (4OMGA); homocysteine (tHcy)					
Hruby et al., 2020 (196)	55.1 ± 9.8 y, US	Cross-sectional	2,205	HILIC LC-MS (Untargeted & Targeted)	Serum	Total dairy, milk, cream/butter, cheese, and yogurt	Harvard semiquantitative 126-item FFQ
Total dairy		cis/trans-Hydroxyproline; pantothenate; uridine					
Milk		cis/trans-Hydroxyproline; uridine					
Cream/butter		C54:4 TAG;					
Cheese		C46:0 TAG; C54:4 TAG; C54:5 TAG; C54:6 TAG					

Yogurt		C20:5 CE					
Hustad et al., 2020 (197)	55 y, Norway	Cross-sectional	517	NMR (Targeted)	Plasma	Fatty fish	Semi-quantitative FFQ
Fatty fish		Fatty acids (Unsaturation; DHA; *n-3* FA; DHA:FA; *n-3* FA:FA) Amino Acids (Valine; Tyrosine)					
Jaceldo-Siegl et al., 2008 (198)	31-93 y, US & Canada	Cross-sectional	100	HPLC–PDA–MS (Targeted – isoflavonoids (daidzein, genistein, total isoflavonoids, and equol)	Urine	Soy protein	Three 24-h recalls
Soy-protein		Daidzein; genistein; total isoflavonoids					
Klebanoff et al., 1998 (199)	– US	Cross-sectional	60 females	HPLC (Targeted)	Serum	Caffeine	24-hour dietary recall
Caffeine		Paraxanthine					
Landberg et al., 2018 (200)	50–64 y, Sweden	Cross-sectional	40	GC-MS (Targeted – alkylresorcinol (AR) metabolites)	Urine	Whole grain intake	4-day food record
Whole grain intake		(3,5-dihydroxycinnamic acid (DHCA); 2-(3,5-dihydroxybenzamido)acetic acid (DHBA-glycine)					
Lau et al., 2018 (201)	6–11 y, Six European countries	Cross-sectional	1,192 children	^{1}H NMR – urine LC-MS/MS – serum (Targeted – amino acids, biogenic amines, acylcarnitines, glycerophospholipids, sphingolipids, and sum of hexoses)	Urine & Serum	Dietary intake habits	Short FFQ (consumption per-week)
Bakery product		Serum: lysoPC a C17:0; PC aa C36:6; PC aa C38:0; PC aa C38:6; PC aa C40:6; PC ae C38:0; PC ae C38:6; PC ae C40:2; PC a C40:6; SM (OH) C14:1; SM (OH) C16:1					
Beverages (soft & fizzy drinks)		Urine: 3-hydroxybutyrate/ 3-aminoisobutyrate Serum: C5:1; C6:1; SM (OH) C16:1					

Dairy				Urine: pantothenate Serum: PC aa C28:1; PC aa C30:0; PC aa C38:0; PC ae C30:0; PC ae C38:6; SM (OH) C14:1			
Fish				Serum: PC aa C36:0; PC aa C36:5; PC aa C36:6; PC aa C38:0; PC aa C38:6; PC aa C40:1; PC aa C40:4; PC aa C40:6; PC aa C42:2; PC ae C38:0; PC ae C38:6; PC ae C40:6			
Fruits				Urine: leucine; threonine/ lactate; alanine; succinate; glutamine; *scyllo*-inositol; hippurate; proline betaine; *N*-methylnicotinic acid Serum: Acetylornithine			
Meat				Urine: creatine Serum: PC ae C36:3; PC ae C36:4; PC ae C36:5; PC ae C38:5			
Potatoes				Urine: acetate Serum: PC ae C30:0; PC ae C34:0; SM (OH) C14:1			
Sweets				Serum: PC aa C38:0; PC aa C38:6; PC ae C38:5; PC ae C38:6; PC ae C40:6			
Vegetables				Urine: hippurate			
Lécuyer et al., 2020 (202)	48.3 ± 6.7 France	Cross-sectional	160 females	UPLC-QToF-MS (Untargeted)	Plasma	Overall diet	24-h dietary record
Alcohol				Ethyl-β-D-glucopyranoside			
Fruits and vegetables				Pipecolic acid; 2,6 dimethylheptanoyl carnitine; nonanoyl-l-carnitine; acylcarnitine C9:1; Phenylalanine			
Vegetables added fat				LysoPC (17:1)			
Milk and dairy products				Phenylalanyl-phenylalanine			
Sweetened food				Cholic acid			
Western diet				Pipecolic acid; Piperine			
Healthy diet				Pipecolic acid; 2,6 dimethylheptanoyl carnitine; nonanoyl-l-carnitine; acylcarnitine C9:1; Dihydro4mercapto-3(2H) furanone			
Lindqvist et al., 2019 (203)	120 y, Sweden	Cross-sectional	120	^{1}H-NMR (Untargeted)	Serum	meat-eaters from non–meat eaters and vegans from nonvegans	FFQ and a 4-d weighed food diary
Meat eaters				2-aminobutyrate; 3-hydroxyisobutyrate; Creatine; Creatine + lysine; Creatinine; Glutamine; Glycine; Isoleucine; Leucine isoleucine; Lysine + arginine; Trimethylamine; Valine			

Vegan		2-aminobutyrate; 3-hydroxyisobutyrate; Creatine; Creatine + lysine; Glutamine; Glycine; Isoleucine; Leucine; Leucine + isoleucine; Lysine + arginine; Trimethylamine; Valine					
MacDougall et al., 2018 (204)	Children 9 ± 2 y Adolescents 15 ± 2 y, US	Cross-sectional	326	Stable isotope-MS (Targeted – δ13C biomarker)	Plasma	Added sugar & sugar-sweetened beverages	24-h dietary recall
Added sugar & sugar-sweetened beverages		^{13}C:^{12}C (reported as δ13C)					
Mack et al., 2019 (205)	49.2 y, Germany	Cross-sectional	97	GC×GC-MS (Untargeted)	Urine	Coffee	24 h dietary recall
Coffee		3,4-dimethyl-2,5-furandione; 2-Methyl-furan; Guaiacol; 2-/3-Methyl-butanoic acid; 2-Vinylfuran					
Malik et al., 2019 (206)	3 ongoing studies: (1) Nurses' Health Study 30–55 y (2) NHS II 25–42 y (3) Health Professionals Follow-up Study 40–75 y, US	Cross-sectional	1,099	LC-MS (Targeted – lipid metabolites)	Plasma	Nut Consumption	Validated FFQs
Total Nut consumption		C24:0 SM; C36:3 phosphatidylcholine (PC) plasmalogen-A; C36:2 PC plasmalogen; C24:0 ceramide; C36:1 PC plasmalog C22:0 SM; C34:1 PC plasmalogen; C36:2 phosphatidylethanolamine plasmalogen; C34:3 diacylglycerol (DAG); C16:1 lysophosphatidylcholine (LPC); C16:1 cholesterol ester (CE); C32:1 DAG, C22:6 lysophosphatidylethanolamine (LPE); C2 LPC; C18:0 sphingomyelin (SM); C50:2 TG; C34:2 DAG					
Peanuts & peanut butter		C24:0 SM; C24:0 ceramide; C22:0 SM					
Other nuts		C16:1 LPC; C16:1 CE; C34:3 DAG; C22:6 LPC; C22:6 LPE; C32:1 DAG					

Markhus et al., 2013 (207)	29.6 ± 4.8 y, Norway	Cross-sectional		LC–MS/MS (Targeted – omega-3 index, the omega-3 HUFA score, and serum 25OH vitamin D)	Serum	Seafood consumption	Short FFQ
Oily fish dinner index		$25OHD_3$					
Seafood spread index		DHA; Omega-3 index[1]; RBC omega-3 HUFA score[2]					
Oily fish spread index		DHA; Omega-3 index[1]; RBC omega-3 HUFA score[2]					
Maskarinec et al., 1998 (208)	36-80 y, US	Cross-sectional	102 females	RP-HPLC (Targeted – isoflavones)	Urine	Soy intake	12-item soy-based foods questionnaire
Soy protein intake		Urinary isoflavones (Daidzein at the highest rate, followed by Genistein and Glycitein)					
McNamara et al., 2020 (97)	18-90 y, Ireland	Cross-sectional	565	^{1}H-NMR & LC-MS (Untargeted)	Urine	Apple	Four-day semi-weighed food diary
Apple		xylose					
Menni et al., 2013 (209)	58.5 ± 10.5 y, UK	Cross-sectional	1,003 females	FIA-MS/MS (Targeted – acylcarnitines, Hydroxylacylcarnitines and dicarboxyl-acylcarnitines, amino acids, sugar, sphingomyelins and sphingomyelin-derivatives, and glycerophospholipids)	Serum	Nutritional patterns	131-item FFQ
Garlic		Acylcarnitines - C8:1; C5–DC(C6–OH)					
Coffee		Acylcarnitines - C10:1					
Hypo-caloric dieting		Acylcarnitines - C9 Glycerophospholipids - PC ae C38:3					

Fruit and vegetables		Glycerophospholipids - PC aa C36:6; PC aa C38:6; PC aa C40:6; PC ae C38:6; PC ae C40:6 Sphingolipids - SM C26:1					
McCullough et al., 2019 (210)	68.3 ± 5.7 y, US	Cross-sectional	1,367 post-menopausal females	UPLC-MS/MS (Untargeted)	Serum	4 diet pattern scores—Mediterranean diet score (aMED), alternate Healthy Eating Index (AHEI)-2010, the Dietary Approaches to Stop Hypertension (DASH) diet, and the Healthy Eating Index (HEI)	152-item modified semi-quantitative Harvard FFQ
aMED		General: Sphingomyelin (d17:1/16:0, d18:1/15:0, d16:1/17:0); γ -Tocopherol/β-tocopherol; Orotidine; N-Acetylalanine Fish: Sphingomyelin (d18:2/18:1); Hydroxy-CMPF; Docosahexaenoate (DHA, 22:6n–3); Eicosapentaenoate (EPA, 20:5n- CMPF; Docosahexaenoylcholine; Eicosapentaenoylcholine; 1-Docosahexaenoylglycerol (22:6)					
AHEI		General: Hydroxy-CMPF; CMPF; Docosahexaenoate (DHA, 22:6n–3); Sphingomyelin (d18:2/18:1); Eicosapentaenoate (EPA, 20:5n–3); γ -Tocopherol/β-tocopherol; Eicosapentaenoylcholine; Ergothioneine Red/processed meat: Docosahexaenoylcholine Fruits & vegetables: Carotene diol					
DASH		General: Sphingomyelin (d18:2/18:1); γ -Tocopherol/β-tocopherol; Hydroxy-CMPF; Carotene diol; Threonate Fruits: β-Cryptoxanthin Low-fat dairy: Galactonate					
HEI		General: Docosahexaenoylcholine; 1-Docosahexaenoylglycerol (22:6); Eicosapentaenoylcholine; Eicosapentaenoate (EI 20:5n–3); 3-Methylxanthine; 7-Methylurate Seafood & plant protein: Docosahexaenoate (DHA, 22:6n–3); Hydroxy-CMPF Greens & beans: Hydroxy-CMPF; Carotene diol Fruits: β-Cryptoxanthin Added Sugar: Ergothioneine					

Noh et al., 2017 (211)	53.9 ± 8.5 y, Four European countries (Germany, France, Italy, and Greece)	Cross-sectional	475	UPLC-ESI-MS/MS (Targeted – polyphenol metabolites)	Urine	Polyphenol-rich foods	24-h dietary recalls and 158-266 items dietary questionnaire (12 months)
Apple & pears	Phloretin; Epicatechin						
Citrus	Naringenin; Hesperetin; 3,4-Dihydroxyphenylacetic acid; Catechin						
Coffee	Caffeic acid; Ferulic acid; 3,4-Dihydroxyphenylacetic acid; Gallic acid; Apigenin; Quercetin; Homovanillic acid; Protocatechuic acid; Daidzein						
Olives	Hydroxytyrosol; Tyrosol; Quercetin; 3,4-Dihydroxyphenylacetic acid; Gallic acid ethyl ester						
Tea	Gallic acid; Hydroxytyrosol; Protocatechuic acid; 3,4-Dihydroxyphenylacetic acid; 3-Hydroxybenzoic acid; m-Coumaric acid; 3,5-Dihydroxyphenylpropionic acid; Resveratrol; Gallic acid ethyl ester						
All wine	Hydroxytyrosol; Gallic acid ethyl ester; Homovanillic acid; 3-Hydroxybenzoic acid						
Red wine	Gallic acid ethyl ester						
O'Gorman et al., 2014 (212)	32 ± 12 (Male) 37±14 (Female), Ireland	Cross-sectional	34	ESI–MS/MS (Targeted – lipidomic patterns)	Serum	Lipidomic patterns with dietary data	The European Prospective Investigation into Cancer (EPIC) FFQ
Saturated fatty acid	PCaeC38:3; PCaeC36:2; PEaaC22:2; PEaaC34:0; PEaeC40:4 SMC21:0; SMC20:2; LPCaC18:2						
Polyunsaturated fatty acid	PEaaC22:2; PEaaC34:0; PEaeC40:4; SMC15:0; SMC20:2; SMC21:1; LPEeC18:0						
Monounsaturated fatty acid	PCaeC36:2; PCaeC38:3; PEaaC22:2; PEaaC34:0; PEaeC40:4; SMC19:0						
Meat	PEaaC38:5; PEaaC36:1						
Alcohol	LPCeC18:0						
Fish	LPEaC18:2; PEaaC38:4						
Vegetable	PSaaC36:2						

Pallister et al., 2016 (213)	52.5 ± 13 y, UK	Cross-sectional	3,559 Females	LC-MS and GC-MS (Untargeted and targeted)	Blood	Food groups and individual food items	131-item FFQ
Alcohol	Amino Acids: 2-aminobutyrate; 2-hydroxybutyrate (AHB); 3-(4-hydroxyphenyl)lactate; 3-methyl-2-oxobutyrate; 4-meth 2-oxopentanoate; alpha-hydroxyisovalerate; beta-hydroxyisovalerate; pipecolate Lipids: 4-androsten-3beta,17beta-diol disulfate 1[1]; 5alpha-androstan-3beta,17beta-diol disulfate; arachidonate (20:4n caprate (10:0); caprylate (8:0); docosahexaenoate (DHA; 22:6n3); docosapentaenoate (n3 DPA; 22:5n3); eicosapentaenoate (EPA; 20:5n3); epiandrosterone sulfate; myo-inositol; scyllo-inositol; stearidonate (18:4n3); X-1264 1-docosahexaenoylglycerophosphoethanolamine Xenobiotics: benzoate; piperine; theophylline						
Seafood	Amino Acid: pyroglutamine[1]; X-12696--3,4-dihydroxyphenylacetate sulfate Carbohydrates: 1,5-anhydroglucitol (1,5-AG) Lipids: 1-docosahexaenoylglycerophosphocholine[1]; 1-eicosatrienoylglycerophosphocholine[1]; 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF); docosapentaenoate (n3 DPA; 22:5n3); docosahexaenoate (DHA; 22:6n3); eicosapentaenoate (EPA; 20:5n3)						
Meat	Amino Acids: creatine; pyroglutamine[1]; trans-4-hydroxyproline						
Meat dishes	Amino Acids: 3-phenylpropionate (hydrocinnamate); Lipids: scyllo-inositol Xenobiotics: ergothioneine						
Fats and oils	Lipids: 10-undecenoate (11:1n1); 10-nonadecenoate (19:1n9); 15-methylpalmitate (isobar with 2-methylpalmitate); X-13431--nonanoylcarnitine[1]; eicosapentaenoate (EPA; 20:5n3); myristate (14:0); pentadecanoate (15:0)						
Tea and coffee	Lipids: Phosphatidylcholine acyl-alkyl C38:3; Phosphatidylcholine acyl-alkyl C40:4; Phosphatidylcholine acyl-alkyl C40:3 Phosphatidylcholine acyl-alkyl C38:2; Phosphatidylcholine diacyl C42:4; Phosphatidylcholine acyl-alkyl C38:1 Peptide: cyclo(leu-pro) Xenobiotics: 1-methylxanthine; X-12039--3-hydroxypyridine sulfate; X-12217--O-methyl catechol sulfate; X-13741--3-methyl catechol sulfate 1; catechol sulfate; quinate						

Allium vegetables	Amino Acids: tryptophan betaine
Apples/pears	Amino Acids: 3-phenylpropionate (hydrocinnamate); indolepropionate Carbohydrates: threitol
Avocado	Lipids: 1-docosahexaenoylglycerophosphocholine[1]; 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF); eicosapentaenoate (EPA; 20:5n3)
Baked sweets	Lipids: docosahexaenoate (DHA; 22:6n3); eicosapentaenoate (EPA; 20:5n3); scyllo-inositol
Banana	Amino Acids: indolepropionate
Beef Burger	Amino Acids: trans-4-hydroxyproline
Black tea	Peptide: cyclo(leu-pro) Xenobiotics: X-12039--3-hydroxypyridine sulfate; X-13741--3-methyl catechol sulfate 1; quinate
Butter	Lipids: 10-undecenoate (11:1n1); 10-nonadecenoate (19:1n9); 15-methylpalmitate (isobar with 2-methylpalmitate); X-13431--nonanoylcarnitine[1]; myristate (14:0); pentadecanoate (15:0)
Chocolate	Xenobiotics: 7-methylxanthine; theobromine
Citrus fruit	Carbohydrates: glycerate Xenobiotics: stachydrine
Coffee	Peptide: cyclo(leu-pro) Xenobiotics: 1-methylxanthine; X-12039--3-hydroxypyridine sulfate; X-12217--O-methyl catechol sulfate; X-13741--3-methyl catechol sulfate 1; catechol sulfate; quinate
Confectionary/jam	Amino Acids: pipecolate Carbohydrates: glycerate
Cream	Lipids: lysoPhosphatidylcholine acyl; C17:0 Hydroxysphingomyeline C14:1; lysoPhosphatidylcholine acyl C28:1
Fried fish	Amino Acids: 3-phenylpropionate (hydrocinnamate) Lipids: scyllo-inositol
Fruit juice	Xenobiotics: stachydrine
Green leafy vegetables	Lipids: 1-docosahexaenoylglycerophosphocholine[1]; 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)

Hamburgers	Amino Acids: trans-4-hydroxyproline
Herbal tea	Lipids: Phosphatidylcholine acyl-alkyl C38:3; Phosphatidylcholine acyl-alkyl C40:4; Phosphatidylcholine acyl-alkyl C40:3 Phosphatidylcholine acyl-alkyl C38:2; Phosphatidylcholine diacyl C42:4; Phosphatidylcholine acyl-alkyl C38:1
Salad dressing (high fat)	Lipids: eicosapentaenoate (EPA; 20:5n3)
High fiber cereal	Cofactors and vitamins: pyridoxate
Meat	Amino Acids: creatine; pyroglutamine[1]; trans-4-hydroxyproline
Mushrooms	Xenobiotics: ergothioneine
Oily fish	Lipids: 1-arachidonoylglycerophosphoethanolamine[1]; 1-docosahexaenoylglycerophosphocholine[1]; 1-eicosatrienoylglycerophosphocholine[1]; 1-linoleoylglycerophosphoethanolamine[1]; 1-oleoylglycerophosphoethanolamin 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF); docosapentaenoate (n3 DPA; 22:5n3); docosahexaenoate (DHA; 22:6n3); eicosapentaenoate (EPA; 20:5n3);
Other fish/ seafood	Amino Acid: pyroglutamine[1]; X-12696--3,4-dihydroxyphenylacetate sulfate Carbohydrates: 1,5-anhydroglucitol (1,5-AG) Lipids: 1-docosahexaenoylglycerophosphocholine[1]; 1-eicosatrienoylglycerophosphocholine[1]; 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF); docosapentaenoate (n3 DPA; 22:5n3); docosahexaenoate (DHA; 22:6n3); eicosapentaenoate (EPA; 20:5n3);
Porridge	Xenobiotics: X-12253--2-aminophenol sulfate
Poultry/chicken	Amino Acids: creatine; pyroglutamine[1]
Processed meat	Amino Acids: 3-phenylpropionate (hydrocinnamate)
Red meat	Amino Acids: creatine; pyroglutamine[1]; trans-4-hydroxyproline
Refined grain (white/brown bread)	Lipids: Octenoylcarnitine
Savoury pies	Amino Acids: 3-phenylpropionate (hydrocinnamate) Xenobiotics: ergothioneine
Soy/other milk	Xenobiotics: 4-ethylphenylsulfate
Spirits/liquor	Lipids: 4-androsten-3beta,17beta-diol disulfate 1[1]
Sweet baked products	Lipids: docosahexaenoate (DHA; 22:6n3); eicosapentaenoate (EPA; 20:5n3); scyllo-inositol;

Tomatoes		Carbohydrates: glycerate					
Wine		Amino Acids: 2-aminobutyrate; 2-hydroxybutyrate (AHB); 3-(4-hydroxyphenyl)lactate; 3-methyl-2-oxobutyrate; 4-meth 2-oxopentanoate; alpha-hydroxyisovalerate; beta-hydroxyisovalerate; pipecolate Lipids: 4-androsten-3beta,17beta-diol disulfate 1[1]; 5alpha-androstan-3beta,17beta-diol disulfate; arachidonate (20:4n caprate (10:0); caprylate (8:0); docosahexaenoate (DHA; 22:6n3); docosapentaenoate (n3 DPA; 22:5n3); eicosapentaenoate (EPA; 20:5n3); epiandrosterone sulfate; myo-inositol; scyllo-inositol; stearidonate (18:4n3); X-1264 1-docosahexaenoylglycerophosphoethanolamine Xenobiotics: benzoate; piperine; theophylline					
Pallister et al., 2017 (214)	TwinUK 55.3 ± 13.4 y EGCUT 37.9 ± 15.7 y KORA 64.1 ± 5.5 y	Cross-sectional	TwinUK (3,559) EGCUT (1,109) KORA (1,593)	FIA-MS/MS (Untargeted and targeted)	Serum	Milk intake	131-item FFQ
Milk		Trimethyl-N-aminovalerate (5-trimethylaminovalerate); Uridine; Phenylalanine; Tyrosine; Valine; 1,5-Anhydroglucitol; Erythronate; Diacylphosphatidylcholine C28:1; Hydroxy-sphingomyelin C14:1					
Papandreou et al., 2019 (215)	67.14 ± 6 y, Spain	Cross-sectional	1,664	LC-MS (Untargeted)	Plasma	Coffee	Validated semi-quantitative 137-item FFQ
Total coffee[1]		5-Acetylamino-6-amino-3-methyluracil (AAMU); Caffeine; Cotinine; C24:0 sphingomyelin (SM); Proline betaine; Kynure acid; Glycocholate; Lactate; Glyco-deoxy-chenodeox; Sucrose; 7-methylguanine					
Caffeinated coffee		Caffeine; 5-Acetylamino-6-amino-3-methyluracil (AAMU); C24:0 sphingomyelin (SM); Cotinine; Sucrose; Proline betain Acetaminophen; C16:0 lyso-phosphatidylethanolamine (LPE); Piperine; Hypoxanthine					
Decaffeinated coffee		Hydroxyhippurate; Alpha-glycerophosphate; C24:0 sphingomyelin (SM); Hippurate; C40:6 phosphatidylcholine (PC); C16:0 lyso-phosphatidylethanolamine (LPE); Phosphocreatine; Allantoin					
Parmenter et al., 2018 (216)	77.2 y, Ireland	Cross-sectional	346	UHPLC-ESI-MS (Targeted – (poly)phenol (phenyl-γ-valerolactones))	Plasma	Dietary (poly)phenols	Interviewer led FFQ
Dietary (poly)phenol intake		phenyl-γ-valerolactones metabolites: ▪ 5-(3′,4′-dihydroxyphenyl)-γ-valerolactone-3′-O-sulfate (3′4′-DiOH-VL-3′-O-Sulph)					

		▪ 5-(3',4'-dihydroxyphenyl)-γ-valerolactone-3'-O-glucuronide (3'4'-DiOHVL-3'-O-Gluc) ▪ 5-(3',5'-dihydroxyphenyl)-γ-valerolactone-3'-O-glucuronide (3'5'-DiOH-VL-3'-O-Gluc)					
Perng et al., 2019 (217)	8-14 y, Mexico	Cross-sectional	242	LC-MS (Untargeted)	Plasma	Sugar-sweetened beverage (SSB)	Age-specific semi-quantitative FFQ
Sugar-sweetened beverage (SSB)		Females: 5-methyl-tetrohydrofolate; phenylephrine; urate; nonanoate; deoxyuridine; sn-glycero-3-phosphocholine Males: 2-piperidinone; octanoylcarnitine; catechol					
Philibert et al., 2006 (218)	18-74 y, Canada	Cross-sectional	243	Inductively coupled plasma (ICP)-MS (Targeted – fatty acids)	Serum	Total fish consumption	Interviewer-administered FFQ
Total fatty fish		eicosapentaenoic acid (EPA) + docosahexaenoic acid (DHA)					
Total salmonid		eicosapentaenoic acid (EPA) + docosahexaenoic acid (DHA)					
Playdon et al., 2016 (219)	57 ± 9, US	Case-control	253	UHPLC-MS/MS and GC-MS (Untargeted)	Urine & Serum	Habitual diet	100-item FFQ
Apple		<u>Urine</u> Lipids: Trimethylamine N-oxide (TMAO)					
Beer		<u>Urine</u> Amino acids: Beta-hydroxyisovalerate; Homovanillate sulfate; hydroxyisovaleroyl carnitine; Proline Lipids: Glycerol 3-phosphate (G3P); Phosphoethanolamine					
Butter		<u>Urine</u> Xenobiotics: 2-hydroxyacetaminophen sulfate; 3-(cystein-S-yl)acetaminophen					
Caffeinated coffee		<u>Urine</u> Amino acids: C-glycosyltryptophan; Homovanillate sulfate; Methionine; Proline; Pyroglutamylvaline; Sarcosine (N-Methylglycine) Xenobiotics: 1-methylurate; 1-methylxanthine; 1,3-dimethylurate; 1,3,7-trimethylurate; 1,7-dimethylurate; 5-acetylamino-6-amino-3-methyluracil; 3-methylxanthine; 5-acetylamino-6-formylamino-3-methyluracil; Caffeine; Hippurate; N-(2-furoyl)glycine; Paraxanthine; Quinate; Theophylline; Cofactors and vitamins: Nicotinate; Trigonelline (N'-methylnicotinate) Lipids: Phosphoethanolamine Nucleotide: Pseudouridine Energy: 2-methylcitrate; Succinylcarnitine <u>Serum</u> Amino acids: Cyclo(leu-pro)					

	Xenobiotics: 1-methylxanthine; 1,3-dimethylurate; 1,3,7-trimethylurate; 1,7-dimethylurate; Caffeine; Catechol sulfate; (2-furoyl)glycine; Paraxanthine; Quinate; Theophylline Cofactors and vitamins: Trigonelline (N'-methylnicotinate)
Chicken	Urine Nucleotide: 7,8-dihydroneopterin
Citrus	Urine Amino acids: N-methylproline or N-methyl proline; N-methylglutamate Xenobiotics: Betonicine or 4-hydroxyproline betaine; Ectoine; Quinate; Stachydrine Lipids: Chiro-inositol; Myo-inositol; Scyllo-inositol Serum Amino acids: N-methylproline or N-methyl proline Xenobiotics: Stachydrine Lipids: Chiro-inositol; Scyllo-inositol
Coffee	Urine Amino acids: 2-aminobutyrate; 3-methoxytyrosine; 3-methyl-2-oxovalerate; C-glycosyltryptophan; Alanine; Homovanillate sulfate; Isoleucine; Methionine; Pyroglutamylvaline; Sarcosine (N-Methylglycine); Threonine; Vanillylmandelate (VMA) Xenobiotics: 1-methylxanthine; 1-methylurate; 1,3-dimethylurate; 1,7-dimethylurate; 3-hydroxyhippurate; 5-acetylamino-6-formylamino-3-methyluracil; 5-acetylamino-6-amino-3-methyluracil; Caffeine; Hippurate; Quinate; N-(2 furoyl)glycine; Paraxanthine; Theophylline Lipids: Phosphoethanolamine Carbohydrates: 3-sialyllactose; Lactose Nucleotide: Pseudouridine Energy: 2-methylcitrate; Succinylcarnitine Serum (caffeinated and decaffeinated) Amino acids: Cyclo(leu-pro); Glutarylcarnitine (C5) or glutaroyl carnitine Xenobiotics: 1-methylxanthine; 1,3,7-trimethylurate; Caffeine; Catechol sulfate; Quinate; N-(2-furoyl)glycine; Paraxanthine; Theophylline Cofactor and vitamins: Trigonelline (N'-methylnicotinate) Serum (decaffeinated) Amino acids: 3-methoxytyrosine Xenobiotics: 1,7-dimethylurate
Corn	Urine Amino acids: 3,4-dihydroxyphenylacetatel; Cysteine; Homovanillate (HVA); Homovanillate sulfate Xenobiotics: 2-oxo-1-pyrrolidinepropionate; 4-hydroxymandelate; Erythritol

	Lipids: 3-hydroxy-3-methylglutarate; 21-hydroxypregnenolone disulfate; Azelate (nonanedioate); Pimelate (heptanedioate); Suberate (octanedioate) Carbohydrates: 3-sialyllactose; 6-sialyl-N-acetyllactosamine; Arabitol; Cytosine; Erythronate; Mannitol; Xylonate Nucleotide: N1-methylguanosine; N2,N2-dimethylguanosine; N2-methylguanosine; N4-acetylcytidine; Pseudouridine Cofactor and vitamins: Pantothenate Energy: 2-methylcitrate
Cruciferous vegetables	Urine Lipids: Acetylcarnitine; Cholate Serum Amino acids: Kynurenine
Desserts	Urine Xenobiotics: 3-methylxanthine; 7-methylxanthine
Fish	Urine Lipids: 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF) Serum Lipids: 1-docosahexaenoylglycerophosphocholine (22:6n3); 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF); Docosahexaenoate (DHA; 22:6n3); Eicosapentaenoate (EPA; 20:5n3);
Green leafy vegetables	Urine Carbohydrates: 6-sialyl-N-acetyllactosamine Serum Amino acids: Glutamate Cofactor and vitamins: Threonate
Grapefruit	Serum Lipids: Deoxycarnitine
High fiber grains	Urine Lipids: Octanoylcarnitine Carbohydrates: Sucrose Cofactor and vitamins: Quinolinate
Liquor	Urine Xenobiotics: Ethyl glucuronide Serum Xenobiotics: Ethyl glucuronide
Margarine	Urine Carbohydrates: Pyruvate

Meat fat	Urine Amino acids: 4-hydroxyphenylacetate; 5-aminovalerate; Creatine; Lysine; Leucine; N-acetyltyrosine; N-acetylglutamine N-acetylleucine; Lipids: Acetylcarnitine Carbohydrates: Xylitol
Multivitamin use	Urine Cofactor and vitamins: Alpha-CEHC sulfate (X - 12435); Alpha-CEHC glucuronide; Pantothenate; Pyridoxate; Riboflavin (Vitamin B2) Serum Cofactor and vitamins: Alpha-tocopherol; Pantothenate; Pyridoxate; Riboflavin (Vitamin B2)
Nuts	Urine Amino acids: Tryptophan betaine Xenobiotics: 4-vinylphenol sulfate Serum Amino acids: Tryptophan betaine
Orange	Urine Amino acids: Anserine Xenobiotics: Stachydrine Nucleotide: Allantoin
Orange/Yellow vegetables	Urine Amino acids: 5-aminovalerate
Other fruit	Urine Xenobiotics: Hydroxycotinine; Nicotine Carbohydrates: Xylitol Energy: Isocitrate Serum Carbohydrates: Mannitol
Other vegetables	Urine Xenobiotics: Stachydrine
Processed meat	Urine Lipids: Acetylcarnitine; Carnitine Serum Carbohydrates: Glycerate
Red meat	Urine

	Xenobiotics: Cinnamoylglycine; Ethyl glucuronide; Methyl-alpha-glucopyranoside Lipids: 3-dehydrocarnitine; Acetylcarnitine; Carnitine Carbohydrates: Sorbitol; Xylitol
Shellfish	Urine Amino acids: 2-hydroxybutyrate (AHB); Alpha-hydroxyisovalerate; Creatine; Ciliatine (2-aminoethylphosphonate); Lysir N-acetylglycine; Taurine Xenobiotics: Sulforaphane-cysteine Lipids: 3-hydroxybutyrate (BHBA); 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF) Serum Amino acids: 2-hydroxybutyrate (AHB) Lipids: 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
Sugar sweetened beverage	Urine Xenobiotics: Stachydrine Serum Amino acids: Methylglutaroylcarnitine (3-methylglutarylcarnitine)
Tea	Urine Amino acids: 3-(4-hydroxyphenyl)lactate; 5-aminovalerate
Total alcohol	Urine Amino acids: 3-hydroxykynurenine; 5-hydroxyindoleacetate; Beta-hydroxyisovalerate; Gamma-glutamylisoleucine; Gamma-glutamylvaline; Homovanillate sulfate; Isobutyrylcarnitine; N-acetylthreonine; Proline Xenobiotics: Ethyl glucuronide; Nicotine Lipids: Glycerol 3-phosphate (G3P) Carbohydrates: 1,5-anhydroglucitol (1,5-AG); 3-sialyllactose Energy: Succinylcarnitine Serum Amino acids: Beta-hydroxypyruvate Cofactor and vitamins: Oxalate (ethanedioate)
Vitamin C	Urine Cofactor and vitamins: Ascorbate (Vitamin C) Serum Carbohydrates: Glycerate
Vitamin E	Urine Cofactor and vitamins: Alpha-CEHC sulfate (X - 12435); Pantothenate

		Serum Cofactor and vitamins: Alpha-CEHC glucuronide; Alpha-tocopherol; Gamma-tocopherol					
Wine		Urine Xenobiotics: 2,3-dihydroxyisovalerate; 2,4,6-trihydroxybenzoate; 2-isopropylmalate; Ethyl glucuronide; Methyl-alpha-glucopyranoside; Nicotine Carbohydrates: 2,3-butanediol; Xylitol					
Playdon et al., 2017 (220)	54–62 y, US	Cross-sectional	1,336 Male smokers	LC-MS, LC-MS/MS, and GC-MS (Untargeted)	Serum	4 diet quality indexes [the Healthy Eating Index (HEI) 2010, the Alternate Mediterranean Diet Score (aMED), the WHO Healthy Diet Indicator (HDI), and the Baltic Sea Diet (BSD)]	FFQ (203 foods and 73 mixed dishes)
HEI-2010							
Whole grain		Xenobiotics: Homostachydrine					
Total protein		Xenobiotics: Homostachydrine					
Seafood and plant protein		Amino acids: Pyroglutamine Lipids: Docosahexaenoate (22:6n–3); Stearidonate (18:4n–3) Xenobiotics: Ergothioneine					
Limit solid fats and added sugars		Amino acids: Pyroglutamine Lipids: Linoleate (18:2n–6); Ergothioneine; Dihomo-linoleate (20:2n–6); Stearidonate (18:4n–3); 1-Linoleoylglycerophosphoinositol					
Dairy		Lipids: Linoleate (18:2n–6) Xenobiotics: Ergothioneine					
Ratio of mono- & poly-unsaturated fat to saturated fat		Lipids: Linoleate (18:2n–6); Dihomo-linoleate (20:2n–6); 1-Linoleoylglycerophosphoinositol					
Seafood		Lipids: Docosapentaenoate (n–3 DPA; 22:5n–3)					

Limit refined grains	Lipids: Docosapentaenoate (n–3 DPA; 22:5n–3)
Limit sodium	Xenobiotics: Ergothioneine
Vegetables	Amino acids: N-δ-acetylornithine Cofactor and vitamins: Threonate Xenobiotics: Ergothioneine
Whole grain	Xenobiotics: 2-Aminophenol sulfate (X-12253)
Total or whole fruit	Cofactor and vitamins: Threonate
aMED	
Ratio of monounsaturated fat to saturated fat	Amino acids: Indolebutyrate Cofactor and vitamins: γ-CEHC Lipids: 1-Myristoleoylglycerophosphocholine (14:1); Mead acid (20:3n–9); cis-4-Decenoyl carnitine; Linoleate (18:2n–6 Linolenate (α or γ; 18:3n–3 or 18:3n–6); 1-Linoleoylglycerol (1-monolinolein) Xenobiotics: Phytanate
Vegetables	Cofactor and vitamins: γ-CEHC; Threonate Xenobiotics: Ergothioneine
Fish & seafood	Lipids: 3-Carboxy-4-methyl-5-propyl-2-furanpropanoate; DHA (22:6n–3) Xenobiotics: Ergothioneine
Fruits	Amino acids: N-methylproline or N-methyl proline Carbohydrate: Threitol Cofactor and vitamins: Threonate Xenobiotics: Stachydrine
Nuts	Amino acids: Tryptophan betaine
HDI	
Polyunsaturated fat 6–10%	Lipids: cis-4-Decenoyl carnitine; Linoleate (18:2n–6); 1-Linoleoylglycerophosphoinositol; Linolenate (a or g; 18:3n–3 or 18:3n–6) Cofactor and vitamins: γ-CEHC; γ-CEHC glucuronide
Fiber	Xenobiotics: 2-Aminophenol sulfate (X-12253); Homostachydrine
BSD	
Ratio of polyunsaturated fat to saturated and trans fat	Amino acids: 3-Hydroxy-2-ethylpropionate Cofactor and vitamins: a-Tocopherol

		Lipids: cis-4-Decenoyl carnitine; y-CEHC; 1-Palmitoleoylglycerophosphoinositol; Linoleate (18:2n–6); 10-Undecenoate (11:1n–1)					
Reduce total fat percentage		Carbohydrate: Threitol Lipids: Linoleate (18:2n–6)					
Vegetables		Cofactor and vitamins: Threonate					
Fruits		Carbohydrate: Threitol Cofactor and vitamins: Threonate					
Pujos-Guillot et al., 2013 (116)	Middle aged adults, France	Cohort	80	RP-LC-ESI-ToF-MS (Untargeted)	Urine	Citrus fruits	Dietary questionnaire (cohort study)
Citrus fruit		Proline betaine; flavanone glucuronides; two terpene metabolites (limonene 8,9-diol glucuronide and nootkatone 13,1 diol glucuronide)					
Rabassa et al., 2020 (221)	73 y, Italy	Cross-sectional	119	HPLC-Q-ToF-MS (Untargeted)	Urine	Nuts	Italian version of the FFQ
Nuts		Urolithin A; Urolithin A glucuronide; Urolithin A sulfate; Urolithin A sulphoglucuronide; Urolithin B; Urolithin B glucuronide; Hydroxyhippuric acid; 2-Hydroxyphenylacetic acid; Resveratrol-sulfate; Dodecanedioic acid; Dimethylglutaric acid; Indole-3-acetic acid glucuronide; Indoxyl sulfate/Indoxylsulphuric acid; Dihydroxy-benzoxazinon					
Reeves et al., 2017 (222)	26-36 y, Australia	Cohort	546 females	^{1}H NMR (Targeted – n-3 fatty acids or tyrosine)	Serum	Fish	127-item FFQ
Fish		DHA; n-3 PUFA; n-3:n-6; Total PUFA; Tyrosine; Tyrosine:LNAA					
Rios-Leyvraz et al., 2020 (223)	10.6 ± 2.9 y, Switzerland	Cross-sectional	91	UHPLC-ESI-MS/MS (Targeted)	Urine	Caffeine	Semiquantitative questionnaire
Caffeine		Caffeine					
Rothwell et al., 2014 (224)	– France	Cross-sectional	39	HPLC-Q-ToF-MS (Untargeted)	Urine	Coffee	Six detailed 24 h records and an FFQ
Coffee		Atractyligenin glucuronide; Cyclo(isoleucyl-prolyl); 1-Methylxanthine; 1,7 Dimethyluric acid; Kahweol oxide glucuronid 1-Methyluric acid; Trigonelline; Dimethylxanthine (Paraxanthine or Theophylline) glucuronide; 3-methyluracil (AFMU); Kahweol oxide glucuronide analogue; Hippuric acid; Trimethyluric acid; Paraxanthine; 3-hydroxyhippuric acid; 1,3- or 3 dimethyluric acid; Caffeine					

Rothwell et al., 2019 (225)	52.5 ± 6.9 France, 53.2 ± 8.8 Germany, 58.2 ± 10.8 Greece, 54.6 ± 7.5 Italy	Cross-sectional	451	UHPLC–MS (Untargeted)	Serum	Coffee intake	Center specific FFQ
Coffee		Trigonelline; Paraxanthine; AAMU; Caffeine; Cyclo(prolyl-valyl); Quinic acid; Cyclo(isoleucyl-prolyl); Pyrocatechol sulfat Hippuric acid; Cyclo(leucyl-prolyl); Theophylline					
Rybak et al., 2015 (226)	≥ 6 y, US	Cross-sectional	2,261	LC-ESI-MS/MS (Targeted – caffeine and caffeine metabolites)	Urine	Caffeine intake (from foods, beverages and dietary supplements)	24-h dietary recall interview
Coffee		Caffeine (1,3,7-trimethylxanthine); theophylline (1,3-dimethylxanthine); paraxanthine (1,7-dimethylxanthine); 1-methylxanthine; 1-methyluric acid; 1,3-dimethyluric acid; 1,7- dimethyluric acid; 1,3,7-trimethyluric acid; 5-acetylamin 6-amino-3-methyluracil; 3,7-dimethyluric acid[1]; theobromine (3,7-dimethylxanthine)[1]; 3-methyluric acid[1]; 3-methylxanthine[1]; 7-methyluric acid[1]; 7-methylxanthine[1]					
Schmidt et al., 2015 (227)	30–49 y, UK	Cross-sectional	379 Males	LC-MS (Targeted – acylcarnitines, amino acids, biogenic amines, glycerophospholipids, hexose, and sphingolipids)	Plasma	Meat eaters, fish eaters, vegetarians, and vegans	Semi-quantitative (FFQ)
Meat eaters		Acylcarnitines: C0; C3; C4; C5; C16 Amino Acids: alanine; creatinine Glycerophospholipids: highest concentration (majority) – (largest difference: PC aa 36:6) Sphingolipids: highest concentration (majority)					
Fish eaters		Amino Acids: glutamate; glutamine; leucine; lysine; methionine; tryptophan; tyrosine; valine					

Vegetarians		Amino Acids: citrulline; kynurenine; leucine; lysine; methionine; tryptophan; tyrosine; valine					
Vegans		Acylcarnitines: C18:2 Amino Acids: citrulline; glycine; ornithine Glycerophospholipids: lowest concentration (majority) Sphingolipids: lowest concentration (majority) – especially: SM(OH)24:1					
Schmidt et al., 2016 (228)	30-49 y, UK	Cross-sectional	392 males	LC-MS/MS (Targeted – amino acids)	Plasma	Meat-eaters, fish-eaters, vegetarians, and vegans	Semi-quantitative FFQ
Habitual diet group		Fish-eaters: lysine; methionine; tryptophan; tyrosine; alanine; Glycine Vegetarians: lysine; methionine; tryptophan; tyrosine; alanine; Glycine Vegans: lysine; methionine; tryptophan; tyrosine; alanine; Glycine					
Seow et al., 1998 (229)	45–74 y, Singapore	Cross-sectional	147	HPLC (Targeted – isoflavonoids)	Urine	Soy intake	Structured food frequency/portion size questionnaire
Soy intake		Daidzein; Sum of isoflavonoids (Daidzein + Genistein + Glycitein)					
Seow et al., 2020 (230)	49.7 y, Singapore	Cross-sectional	1,104 (coffee) 2,302 (black tea) 2,075 (green tea)	LC-MS (Targeted – amino acids, acylcarnitines, and sphingolipids)	Plasma	Coffee, Black Tea, and Green Tea Consumption	Semi-quantitative 169-item validated FFQ
Coffee		SM C30:1; SM C30:1; SM d18:1/14:0; HexCer d16:1/24:0; Hex2Cer d18:1/14:0; SM d18:1/16:0; HexCer d16:1/16:0; SM C32:1; HexCer d18:1/24:0; HexCer d16:1/22:0; SM d16:1/24:0; Hex2Cer d16:1/16:0; SM d16:1/20:0; SM C40:1; SM d16:1/16:0; SM d16:1/22:0; SM d18:2/14:0; SM C38:1; SM C42:1; Cer d16:1/24:0; SM C32:2; S1P d16:1; Cer d16:1/16:(HexCer d18:1/26:0; SM d18:1/22:0; SM d18:1/24:0; Cer d18:1/24:0; HexCer d18:1/18:0; Cer d18:1/16:0; HexCer d18:1/23:0; SM d16:1/24:1; SM C42:2; HexCer d18:2/24:0; C18:2 (Acylcarnitine); Cer d16:1/18:0; SM C36:1; SM C34:1; SM d18:2/24:0; HexCer d18:1/22:0; C16:2 Acylcarnitine; C14:2 Acylcarnitine; SM d16:1/18:0; C18-DC/C20-OH Acylcarnitine; Cer d16:1/20:0; C2 Acylcarnitine; SM d18:2/22:0; SM d18:1/23:0; HexCer d18:2/22:0; SM C40:2; Cer d16:1/22:0; Alanine; SPH d16:1; C14:1-OH; SM d18:1/18:0; SM d18:1/20:0; Cer d18:2/24:0; HexCer d18:1/16:0; Hex2C					

		d18:2/16:0; Cer d18:1/26:0; Cer d18:1/22:0; C10:1; HexCer d18:1/25:0; C10:2; Hex2Cer d18:1/16:0; Cer d18:1/20:0; Ce d18:1/24:1; C7-DC; Hex2Cer d18:1/24:0; HexCer d18:1/24:1; C8:1-OH/C6-DC; C12:1; SM d18:2/23:0; Cer d16:1/24:1; C16:1-OH/C14:1-DC; Cer d16:1/23:0; SM d18:1/24:1; SM d18:2/16:0; C10:3					
Black Tea		SM d16:1/18:0; S1P d17:1; SM d16:1/20:0; Hex2Cer d16:1/16:0; S1P d16:1; SM d18:2/18:0; SM C36:2; SM C43:2; Glyci					
Green Tea		GM3 d18:1/16:0; C8:1-OH/C6-DC; SM d18:2/18:0; SM d18:1/14:0					
Shiokawa et al., 2018 (231)	23–43 y, Japan	Cross-sectional	8	^{1}H NMR (Targeted)	Urine	Fruits and vegetables	309 nutritional datasets of daily dietary intake records
Fruits and vegetables		Hippurate					
Szeto et al., 2004 (232)	Vegetarian 44.2 ± 9.0 y, Non-vegetarian 44.0 ± 9.2 y, Hong Kong	Cross-sectional	30 vegetarians	HPLC (Targeted – amino acids)	Plasma	Vegetarian diet	Ate no meat or fish owing to their religious (Taoist) beliefs
Vegetarians		Ascorbic acid; hsCRP; triacylglycerol; uric acid					
Thiébaut et al., 2009 (233)	56.8 ± 6.4 y, France	Cross-sectional	1,114 women	Capillary-GC (Targeted – phospholipid fatty acid)	Serum	Fatty acid composition as a biomarker of dietary fat	208-item diet history questionnaire
Sunflower oil		Monounsaturates (*cis*16:1n-7; *cis*18:1n-9) n-6 PUFAs (*cis*18:2n-6; 20:4n-6; Total n-6) n-3 PUFAs (18:3n-3; 20:5n-3) Ratios (18:0/18:1)					
Olive oil		Monounsaturates (*cis*18:1n-9) n-6 PUFAs (*cis*18:2n-6; 20:4n-6; Total n-6) n-3 PUFAs (18:3n-3; 20:5n-3; 22:6n-3; Total n-3) Ratios (18:0/18:1; n-6/n-3)					
Dairy products		Saturates (15:0; 17:0)					
Margarine		Monounsaturates (*cis*18:1n-9; *trans*16:1n-7; *trans*18:1n-9; *trans*-MUFA[1]) n-6 PUFAs (*cis*18:2n-6; 20:4n-6; Total n-6)					

				Ratios (18:0/18:1)			
Total fish				Saturates (16:0) Monounsaturates (*cis*18:1n-9) n-6 PUFAs (*cis*18:2n-6; 20:4n-6; Total n-6) n-3 PUFAs (20:5n-3; 22:6n-3; Total n-3) Ratios (n-6/n-3)			
Fatty fish				n-6 PUFAs (Total n-6) n-3 PUFAs (20:5n-3; 22:6n-3; Total n-3) Ratios (n-6/n-3)			
Meat				n-6 PUFAs (20:4n-6)			
Manufactured food				Monounsaturates (*trans*18:1n-9; *trans*-MUFA[1])			
Alcoholic beverages				Saturates (15:0; 17:0; 16:0; 18:0) Monounsaturates (cis16:1n-7; *cis*18:1n-9; *trans*16:1n-7; *trans*18:1n-9; *trans*-MUFA[1]) n-6 PUFAs (*cis*18:2n-6) n-3 PUFAs (20:5n-3) Ratios (18:0/18:1)			
Toffano et al., 2018 (234)	11.7 ± 1.1 y, Brazil	Cross-sectional	167	UPLC (α-tocopherol and β-carotene); RP-LC-ESI-MS (riboflavin and pyridoxine); LC-MS/MS (S-adenosyl-L-homocysteine) (Targeted)	Plasma	Dietary exposure	Dietary intakes (mean scores) - averaging 24-h recalls at baseline (visit 1), and six (visit 2) and 12 (visit 3) weeks
Total vegetables				linoleic acid (LA); α-linolenic acid (ALA); eicosapentaenoic fatty acid (EPA); docosahexaenoic fatty acid (DHA); arachido fatty acid (ARA); β-carotene; Creatine			
Total fruits				linoleic acid (LA); α-linolenic acid (ALA); eicosapentaenoic fatty acid (EPA); docosahexaenoic fatty acid (DHA); arachido fatty acid (ARA); β-carotene			
Dark green and orange vegetables & legumes				α-linolenic acid (ALA); β-carotene; Retinol; Creatine			
Meat, eggs and legumes				α-linolenic acid (ALA); docosahexaenoic fatty acid (DHA); β-carotene; Creatine			
Dark green and orange vegetables WITHOUT legumes				docosahexaenoic fatty acid (DHA); Retinol; S-adenosyl-homocysteine (SAH)			

Whole fruits		β-carotene; Riboflavin					
Milk & dairy		Retinol; Pyridoxal					
Whole grains		5 methyl tetrahydrofolate (5-MTHF)					
Tong et al., 2020 (235)	48,4 y, UK	Cross-sectional	10,806	LC-ESI-MS/MS (Targeted – acylcarnitines, amines, sphingolipids, and phospholipids)	Plasma	Mediterranean diet	130-item semiquantitative FFQ
Mediterranean diet		Acylcarnitines (18:0, 18:2, 16:0, 14:2, 10:1, and carnitine) Amino Acid/Biogenic Amines (trans-hydroxyproline, cis-hydroxyproline, acetylornithine, creatinine, citrulline, threonin isoleucine, proline, tryptophan, glutamate) Lysophosphatidylcholines (20:3, 20:4, 24:0, 16:1, 18:1, 28:0) Phosphatidylcholines, Acyl-Alkyl (42:4, 38:0, 40:6, 36:4, 38:4, 40:2, 38:6, 38:1, 36:3, 40:1, 34:1, 42:1, 42:3, 38:3, 44:6, 40:4, 40:5, 34:2, 42:0, 36:5, 36:1) Phosphatidylcholines, Diacyl (38:6, 40:6, 38:0, 36:5, 36:6, 36:0, 36:1, 34:1, 38:3, 32:1, 38:4, 40:5, 40:1, 42:0, 38:1, 36:3, 34:3, 42:2, 36:4, 40:2, 42:4) Sphingomyelins (18:0, 24:1)				
Van Roekel et al., 2018 (236)	35–70 y, 10 European countries	Cross-sectional	2,974	HPLC-MS/MS (Targeted – acylcarnitines, amino acids, biogenic amines, a sum of hexoses, phosphatidylcholines (PCs) including lysoPCs, diacyl PCs, and acyl–alkyl PCs, and sphingomyelins)	Plasma & Serum	Alcohol	Validated country-specific or center-specific questionnaires
Alcohol		Amino Acids: citrulline Lipids: diacylphosphatidylcholines; lysophosphatidylcholines; sphingomyelins Other: acylcarnitines; acyl-alkyl-phosphatidylcholines					

Vázquez-Fresno et al., 2014 (237)	53-79 y, Spain	Cohort study	91 (5-year)	[1]NMR (Untargeted)	Urine	Wine	137-item FFQ
Wine		Food metabolome metabolites: tartrate; ethyl glucuronide [EtG]; 2,3-butanediol; mannitol; ethanol Endogenous response to wine exposure: 3-methyl-2-oxovalerate					
Vázquez-Manjarrez et al., 2019 (152)	47 y, Germany	Cross-sectional	78	UPLC-QTOF-MS (Targeted)	Urine	Banana	24-h dietary recall
Banana		Methoxyeugenol glucuronide; Dopamine sulfate; Salsolinol sulfate 1; 6-Hydroxy-1-methyl-1,2,3,4-tetrahydro β-carboline sulfate					
Walker et al., 2020 (238)	55 y, US	Cross-sectional	2208	LC-MS/MS	Serum	Alternative Healthy Eating Index (AHEI), the Dietary Approaches to Stop Hypertension (DASH) diet; and a Mediterranean-style (MDS) diet	Harvard semi-quantitative FFQ
Alternative Healthy Eating Index (AHEI) diet		Aconitate; Cholesterol ester (C22:6); Hippurate; Isocitrate; Lysophosphatidylcholine (C20:5); Lysophosphatidylcholine (C22:6); Phosphatidylcholine (C38:6); Phosphatidylcholine (40:6); Sphingomyelin (C18:0); Sphingomyelin (C18:1); Triacylglycerol (C56:7); Triacylglycerol (C56:8); Triacylglycerol (C58:10); Triacylglycerol (C58:8); Triacylglycerol (C58:9); Uridine					
Dietary Approaches to Stop Hypertension (DASH)		Aconitate; Cis/trans-hydroxyproline; Cotinine; Hippurate; Isocitrate; Lysophosphatidylcholine (C20:5); Lysophosphatidylcholine (C22:6); Ornithine; Oxalate; Pantothenate; Phosphatidylcholine (C38:6); Phosphatidylcholine (40:6); Serine; Sphingomyelin (C18:0); Sphingomyelin (C18:1); Thiamine; Triacylglycerol (C56:7); Triacylglycerol (C56:8) Triacylglycerol (C58:10); Triacylglycerol (C58:8); Triacylglycerol (C58:9); Uridine					
Mediterranean-style (MDS) diet		Cholesterol ester (C22:6); Cis/trans-hydroxyproline; Cotinine; Hippurate; Isocitrate; Lysophosphatidylcholine (C22:6); Lysophosphatidylethanolamine (C20:4); Oxalate; Phosphatidylcholine (C38:6); Phosphatidylcholine (40:6); Sphingomye (C14:0); Sphingomyelin (C18:0); Sphingomyelin (C18:1); Triacylglycerol (C54:7); Triacylglycerol (C56:7); Triacylglycerol (C56:8); Triacylglycerol (C58:10); Triacylglycerol (C58:8); Triacylglycerol (C58:9)					
Wanders et al., 2018 (239)	67.9 ± 7.1, Netherlands	Cross-sectional	1,171	GC-FID (Targeted – fatty acids (linoleic acid (LA),	Plasma (CODAM)	Linoleic acid (LA), alpha-linolenic acid	79-item semi-quantitative FFQ

			(pooled data from the CODAM and Hoorn studies)	alpha-linolenic acid (ALA), eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA))	& Serum (Hoorn)	(ALA), eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA)	
Self-reported LA, ALA, EPA, and DHA		Strong association with circulating LA, EPA and DHA Weaker association with circulating ALA					
Wang et al., 2018 (240)	68.3 ± 5.7 y, US	Cross-sectional	369 Nonsmoking post-menopausal women	UHPLC–MS/MS) (Untargeted)	Serum	91 food groups or items	152-item FFQ

Total citrus fruits and juices	Amino acids: N-methylproline Xenobiotics: stachydrine (proline betaine); methyl glucopyranoside (α + β); β-cryptoxanthin
Orange juice	Amino acids: N-methylproline Lipids: chiro-inositol Xenobiotics: stachydrine (proline betaine)
Banana	Amino acid: dopamine 3-O-sulfate; dopamine 4-sulfate; S-methylmethionine; 3-methoxytyramine sulfate; 5-hydroxyindoleacetate
Prunes	Amino acid: 5-hydroxymethyl-2-furoic acid Xenobiotics: hippurate; benzoylcarnitine; catechol sulfate
Cruciferous vegetables	Amino acid: S-methylcysteine sulfoxide
Mushrooms	Xenobiotics: ergothioneine
Allium vegetables	Amino acid: N-methyltaurine Cofactor/vitamin: γ –CEHC; γ -CEHC glucuronide Xenobiotics: N-acetylalliin; piperine; ergothioneine
Onion	Amino acid: N-methyltaurine
Garlic	Amino acid: N-methyltaurine Cofactor/vitamin: γ -CEHC glucuronide; γ -CEHC Xenobiotics: N-acetylalliin; S-allylcysteine; ergothioneine; alliin
Eggs	Lipids: 1-(1-enyl-stearoyl)-2-arachidonoyl-GPE (P-18:0/20:4)
Red meat	Lipids: 1-(1-enyl-stearoyl)-2-arachidonoyl-GPE (P-18:0/20:4); 1-(1-enyl-stearoyl)-2-oleoyl-GPE (P-18:0/18:1)

Poultry	Amino acid: 3-methylhistidine
Total fish	Lipids: 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF); docosahexaenoic acid (DHA); docosahexaenoylcholine 1-docosahexaenoylglycerol (22:6); eicosapentaenoic acid (EPA); eicosapentaenoylcholine
Dark fish	Lipids: 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF); docosahexaenoic acid (DHA); eicosapentaenoic acid (EPA); docosahexaenoylcholine; sphingomyelin (d18:2/18:1); eicosapentaenoylcholine; 1-docosahexaenoylglycerol (22:6); docosapentaenoate (n-6 DPA; 22:5n-6)
Shellfish	Lipids: 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
Total nuts	Amino acids: tryptophan betaine Peptide: γ -glutamylvaline Lipids: lignoceroylcarnitine (C24); behenoylcarnitine (C22); sphingomyelin (d18:2/23:1) Xenobiotics: 4-vinylphenol sulfate
Peanuts	Amino acids: tryptophan betaine Peptide: γ -glutamylvaline Lipids: lignoceroylcarnitine (C24); behenoylcarnitine (C22); sphingomyelin (d18:2/23:1) Xenobiotics: 4-vinylphenol sulfate
Milk	Carbohydrates: galactonate Peptide: phenylacetylglycine Xenobiotics: 2,8-quinolinediol sulfate
Soy milk	Xenobiotics: 4-ethylphenylsulfate
Butter	Lipids: caprate (10:0); 10-undecenoate (11:1n-1); sphingomyelin (d18:1/25:0, d19:0/24:1, d20:1/23:0, d19:1/24:0); caprylate (8:0); sphingomyelin (d17:1/16:0, d18:1/15:0, d16:1/17:0)
French fries	Lipids: eicosanodioate
Chocolate candies	Xenobiotics: 3-methylxanthine; 7-methylurate; 3,7-dimethylurate; theobromine; 7-methylxanthine
Desserts	Lipids: sphingomyelin (d18:2/18:1) Xenobiotics: ergothioneine
Total alcohol	Lipids: 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF); sphingomyelin (d18:2/18:1) Xenobiotics: ethyl glucuronide; caffeine
Total wine	Lipids: 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF); sphingomyelin (d18:2/18:1); oleoyl-linoleoyl-glycerol (18:1/18:2) (2); androstenediol (3β,17β) monosulfate (2) Xenobiotics: ethyl glucuronide; 2,3-dihydroxyisovalerate
Red wine	Xenobiotics: ethyl glucuronide; 2,3-dihydroxyisovalerate
White wine	Xenobiotics: ethyl glucuronide; 2,3-dihydroxyisovalerate

Liquor		Lipids: androstenediol (3β,17β) disulfate (1); androstenediol (3β,17β) monosulfate (2); 5α-androstan-3β,17β-diol disulfate; 5α-androstan-3α,17β-diol disulfate Xenobiotics: ethyl glucuronide					
Total coffee		Cofactor/vitamin: trigonelline (N'-methylnicotinate) Energy: citraconate/glutaconate Xenobiotics: quinate; 3-hydroxypyridine sulfate; 3-methyl catechol sulfate (1)					
Caffeinated coffee		Xenobiotics: 1-methylxanthine; Paraxanthine; 1-methylurate; 5-acetylamino-6-amino-3-methyluracil; 1,3-dimethylurat 1,7-dimethylurate; theophylline; caffeine; 1,3,7-trimethylurate					
Decaffeinated coffee		Cofactor/vitamin: trigonelline (N'-methylnicotinate) Xenobiotics: 3-hydroxypyridine sulfate; quinate; 2,3-dihydroxypyridine					
Total tea		Xenobiotics: theanine					
Non-herbal tea		Xenobiotics: theanine					
Diet soft drinks		Xenobiotics: saccharin					
Wedekind et al., 2019 (157)	53.9 ± 8.5 y, Germany Italy France Greece	Cross-sectional	474	LC-MS (Targeted)	Urine	Smoked meat intake	24HDR and FFQ
Smoked meat product		Syringol sulfate; 4-Methylsyringol sulfate; 4-Ethylsyringol sulfate; 4-Allylsyringol sulfate isomer II					
Wedekind et al., 2020 (158)	53. 9± 8.52 (urine) 54.2 ± 8.5 (serum) Germany, Italy, France, and Greece	Cross-sectional	474 (urine) 451 (serum)	LC-MS (Targeted – acylcarnitines)	Urine & Serum	Red and processed meat	24-h dietary recall
Red & processed meat		Acylcarnitines (C0, 2:0, 3:0, 4:0 (OH), 5:0, 7:0, 8:0 (OH), 8:0 (OH.2), 10:0 (OH), 11:1) Note: OH = hydroxyl group on fatty acid moiety					
Yeung et al., 2010 (241)	60–80 y, US	Cross-sectional	186	Stable isotope-MS (Targeted – δ13C biomarker)	Serum	Food products containing corn- and cane-based sweeteners	Validated Willett semi-quantitative FFQ

						(measured as sweetened beverage intake)	
Sweetened beverage		Carbon isotopic signatures ($\delta^{13}C$)					
Yin et al., 2020 (163)	45 ± 2 y, Ireland	Cross-sectional	100	^{1}H NMR (Targeted)	Urine	Fish	FFQ
Fish		TMAO					
Zheng et al., 2014 (242)	52.9 ± 5.8 y, (Discovery) 52.7 ± 5.7 y (Replication)	Cross-sectional	1,977 African Americans	GC-MS and LC-MS (Untargeted)	Serum	Dietary intake habits	66-item interviewer-administered semi-quantitative FFQ
Sugar-rich foods and beverages		Unsaturated long chain: Docosapentaenoate (n3 DPA; 22:5n3) Fatty acids: 10-Nonadecenoate (19:1n9); Adrenate (22:4n6); Dihomo-linoleate (20:2n6); Eicosenoate (20:1n9 or 11); Oleate (18:1n9); Palmitoleate (16:1n7) 2-Hydroxybutyrate–related metabolites: 2-Aminobutyrate; 2-Hydroxybutyrate; 2-Hydroxyisobutyrate; 3-Hydroxyisobutyrate; α-Hydroxyisovalerate Sex steroids: 4-Androsten-3β,17β-diol disulfate 1; 5α-Androstan-3β,17β-diol disulfate γ-Glutamyl dipeptides: γ-Glutamylglutamate; γ-Glutamylisoleucine; γ-Glutamylleucine; γ-Glutamylthreonine; γ-Glutamyltyrosine In amino acid pathway: Creatine In Krebs cycle pathway: Malate In xanthine metabolism: Theobromine					
Fruits and vegetables		In carbohydrate pathway: Glycerate					
Coffee		Quinate; Paraxanthine 5-Acetylamino-6-amino-3-methyluracil; 1,7-Dimethylurate; 1-Methylurate; 1-Methylxanthine; Caffeine; 1,3,7-Trimethylurate; 7-Methylxanthine					
Eggs		Docosapentaenoate (n6 DPA; 22:5n6)					
Fish and seafood		Eicosapentaenoate (EPA; 20:5n3); Docosahexaenoate (DHA; 22:6n3); 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)					
Fruit juice		Glycerate; Stachydrine; N-methyl proline; Threonate; Scyllo-inositol; Homostachydrine					
Nuts and peanut butter		Tryptophan betaine; 2-Methylbutyroylcarnitine; 4-Vinylphenol sulfate; 5α-Androstan-3β,17β-diol disulfate; 4-Androste 3β,17β-diol disulfate 1					

Dietary Sucrose		10-Nonadecenoate; 2-Aminobutyrate; 2-Hydroxybutyrate; 3-Hydroxyisobutyrate; 4-Androsten-3β,17β-diol disulfate 1; 5α-Androstan-3β,17β-diol disulfate; α-Hydroxyisovalerate; Dihomo-linoleate (20:2n6); Eicosenoate (20:1n9 or 11); γ-Glutamylglutamate; γ-Glutamylisoleucine; γ-Glutamylleucine; γ-Glutamylthreonine; γ-Glutamyltyrosine; Malate; Oleat (18:1n9); Theobromine					
Carbohydrate		10-Nonadecenoate; 2-Aminobutyrate; 2-Hydroxybutyrate; 3-Hydroxyisobutyrate; 4-Androsten-3β,17β-diol disulfate 1; 5α-Androstan-3β,17β-diol disulfate; Adrenate (22:4n6); α-Hydroxyisovalerate; γ-Glutamylglutamate; γ-Glutamylisoleucine; γ-Glutamylleucine; γ-Glutamylthreonine; γ-Glutamyltyrosine					
Zheng et al., 2014 (243)	52.9 ± 5.8 y, (Discovery) 52.8 ± 5.6 y (Replication)	Cross-sectional	1,977 African Americans	GC-MS and LC-MS (Untargeted)	Serum	Alcohol	Interviewer-administered dietary FFQ
Alcohol		Amino acid: 2-Aminobutyrate; α-Hydroxyisovalerate; 2-Hydroxyisobutyrate; α-Hydroxyisocaproate; 2-Hydroxy-3-methylvalerate; 5-oxoproline; Indolelactate Lipid: Docosapentaenoate (n-3 docosapentaenoic acid; 22:5n23); Palmitoleate (16:1n27); Adrenate (22:4n26); Dihomolinoleate (20:2n26); 10-Heptadecenoate [17:1n2Cyclo(leu-pro)7]; Eicosenoate (20:1n2Cyclo(leu-pro)9 or 11); Oleate (18:1n2Cyclo(leu-pro)9); Myristoleate (14:1n2Cyclo(leu-pro)5); Palmitate (16:0); Myristate (14:0); Stearidonate (18:4n23); 5-Hete; 5-Hepe; 1-Palmitoleoylglycerophosphocholine; 1-Stearoylglycerophosphoethanolamine; 1-Pentadecanoylglycerophosphocholine; 2-Arachidonoylglycerophosphoethanolamine; 4-Androsten-3b,17b-diol disulfat 1; 5a-Androstan-3b,17b-diol disulfate; Isovalerate Peptide: Leucylleucine; Cyclo(leu-pro)2; γ-Glutamyl valine; γ-Glutamyl phenylalanine; γ-Glutamyl leucine; γ-Glutamyl isoleucine; γ-Glutamyl tyrosine; γ-Glutamyl glutamate; γ-Glutamyl alanine Energy: Malate Xenobiotics: Piperine					
Zong et al., 2014 (244)	50–70 y, China	Cross-sectional	2,091	GC-FID (Targeted – Erythrocyte fatty acids)	Plasma	Dairy consumption	74-item FFQ
Dairy products		*Trans*-18:1 isomers					
Milk		*Trans*-18:1 isomers					

References

1. Acar E, Gurdeniz G, Khakimov B, Savorani F, Korndal SK, Larsen TM, Engelsen SB, Astrup A, Dragsted LO. Biomarkers of Individual Foods, and Separation of Diets Using Untargeted LC-MS-based Plasma Metabolomics in a Randomized Controlled Trial. Molecular Nutrition & Food Research 2019;63(1). doi: 10.1002/mnfr.201800215.
2. Allen CM, Smith AM, Clinton SK, Schwartz SJ. Tomato consumption increases lycopene isomer concentrations in breast milk and plasma of lactating women. Journal of the American Dietetic Association 2002;102(9):1257-62.
3. Amer B, Clausen MR, Bertram HC, Bohl M, Nebel C, Zheng H, Skov T, Larsen MK, Gregersen S, Hermansen K, et al. Consumption of Whey in Combination with Dairy Medium-Chain Fatty Acids (MCFAs) may Reduce Lipid Storage due to Urinary Loss of Tricarboxylic Acid Cycle Intermediates and Increased Rates of MCFAs Oxidation. Molecular Nutrition & Food Research 2017;61(12). doi: 10.1002/mnfr.201601048.
4. Andersen MBS, Rinnan A, Manach C, Poulsen SK, Pujos-Guillot E, Larsen TM, Astrup A, Dragsted LO. Untargeted Metabolomics as a Screening Tool for Estimating Compliance to a Dietary Pattern. Journal of Proteome Research 2014;13(3):1405-18. doi: 10.1021/pr400964s.
5. Andersen MBS, Kristensen M, Manach C, Pujos-Guillot E, Poulsen SK, Larsen TM, Astrup A, Dragsted L. Discovery and validation of urinary exposure markers for different plant foods by untargeted metabolomics. Analytical and Bioanalytical Chemistry 2014;406(7):1829-44. doi: 10.1007/s00216-013-7498-5.
6. Baenas N, Suarez-Martinez C, Garcia-Viguera C, Moreno DA. Bioavailability and new biomarkers of cruciferous sprouts consumption. Food Res Int 2017;100(Part 1):497-503. doi: 10.1016/j.foodres.2017.07.049.
7. Barnes RC, Kim H, Fang C, Bennett W, Nemec M, Sirven MA, Suchodolski JS, Deutz N, Britton RA, Mertens-Talcott SU, et al. Body Mass Index as a Determinant of Systemic Exposure to Gallotannin Metabolites during 6-Week Consumption of Mango (Mangifera indica L.) and Modulation of Intestinal Microbiota in Lean and Obese Individuals. Molecular Nutrition & Food Research 2019;63(2). doi: 10.1002/mnfr.201800512.
8. Beckmann M, Joosen AM, Clarke MM, Mugridge O, Frost G, Engel B, Taillart K, Lloyd AJ, Draper J, Lodge JK. Changes in the human plasma and urinary metabolome associated with acute dietary exposure to sucrose and the identification of potential biomarkers of sucrose intake. Molecular Nutrition & Food Research 2016;60(2):444-57. doi: 10.1002/mnfr.201500495.
9. Bondia-Pons I, Barri T, Hanhineva K, Juntunen K, Dragsted LO, Mykkanen H, Poutanen K. UPLC-QTOF/MS metabolic profiling unveils urinary changes in humans after a whole grain rye versus refined wheat bread intervention. Molecular Nutrition & Food Research 2013;57(3):412-22. doi: 10.1002/mnfr.201200571.

10. Boto-Ordonez M, Urpi-Sarda M, Queipo-Ortuno MI, Corella D, Tinahones FJ, Estruch R, Andres-Lacueva C. Microbial Metabolomic Fingerprinting in Urine after Regular Dealcoholized Red Wine Consumption in Humans. Journal of Agricultural and Food Chemistry 2013;61(38):9166-75. doi: 10.1021/jf402394c.
11. Bub A, Watzl B, Heeb D, Rechkemmer G, Briviba K. Malvidin-3-glucoside bioavailability in humans after ingestion of red wine, dealcoholized red wine and red grape juice. European Journal of Nutrition 2001;40(3):113-20. doi: 10.1007/s003940170011.
12. Carkeet C, Clevidence BA, Novotny JA, Carkeet C, Clevidence BA, Novotny JA. Anthocyanin excretion by humans increases linearly with increasing strawberry dose. Journal of Nutrition 2008;138(5):897-902.
13. Carrizo D, Chevallier OP, Woodside JV, Brennan SF, Cantwell MM, Cuskelly G, Elliott CT. Untargeted metabolomic analysis of human serum samples associated with different levels of red meat consumption: A possible indicator of type 2 diabetes? Food Chemistry 2017;221:214-21.
14. Charron CS, Vinyard BT, Jeffery EH, Ross SA, Seifried HE, Novotny JA. BMI Is Associated With Increased Plasma and Urine Appearance of Glucosinolate Metabolites After Consumption of Cooked Broccoli. Frontiers in Nutrition 2020;7. doi: 10.3389/fnut.2020.575092.
15. Chen F, Hu ZY, Parker RB, Laizure SC. Measurement of caffeine and its three primary metabolites in human plasma by HPLC-ESI-MS/MS and clinical application. Biomedical Chromatography 2017;31 (6) (no pagination)(e3900).
16. Cheung W, Keski-Rahkonen P, Assi N, Ferrari P, Freisling H, Rinaldi S, Slimani N, Zamora-Ros R, Rundle M, Frost G, et al. A metabolomic study of biomarkers of meat and fish intake. Am J Clin Nutr 2017;105(3):600-8. doi: 10.3945/ajcn.116.146639.
17. Chiang YL, Haddad E, Rajaram S, Shavlik D, Sabate J. The effect of dietary walnuts compared to fatty fish on eicosanoids, cytokines, soluble endothelial adhesion molecules and lymphocyte subsets: a randomized, controlled crossover trial. Prostaglandins Leukotrienes and Essential Fatty Acids 2012;87(4-5):111-7. doi: 10.1016/j.plefa.2012.07.007.
18. Clarke JD, Hsu A, Riedl K, Bella D, Schwartz SJ, Stevens JF, Ho E. Bioavailability and inter-conversion of sulforaphane and erucin in human subjects consuming broccoli sprouts or broccoli supplement in a cross-over study design. Pharmacological Research 2011;64(5):456-63.
19. Clarke KA, Dew TP, Watson REB, Farrar MD, Bennett S, Nicolaou A, Rhodes LE, Williamson G. High performance liquid chromatography tandem mass spectrometry dual extraction method for identification of green tea catechin metabolites excreted in human urine. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences 2014;972:29-37. doi: 10.1016/j.jchromb.2014.09.035.
20. Conaway CC, Getahun SM, Liebes LL, Pusateri DJ, Topham DK, Botero-Omary M, Chung F. Disposition of Glucosinolates and Sulforaphane in Humans After Ingestion of Steamed and Fresh Broccoli. Nutrition & Cancer 2000;38(2):168-78.

21. Cornelis MC, Erlund I, Michelotti GA, Herder C, Westerhuis JA, Tuomilehto J. Metabolomic response to coffee consumption: application to a three-stage clinical trial. Journal of Internal Medicine 2018;283(6):544-57. doi: 10.1111/joim.12737.
22. Cuparencu CS, Andersen MBS, Gurdeniz G, Schou SS, Mortensen MW, Raben A, Astrup A, Dragsted LO. Identification of urinary biomarkers after consumption of sea buckthorn and strawberry, by untargeted LC-MS metabolomics: a meal study in adult men. Metabolomics 2016;12(2). doi: 10.1007/s11306-015-0934-0.
23. Cuparencu C, Rinnan A, Silvestre MP, Poppitt SD, Raben A, Dragsted LO. The anserine to carnosine ratio: an excellent discriminator between white and red meats consumed by free-living overweight participants of the PREVIEW study. European Journal of Nutrition 2020. doi: http://dx.doi.org/10.1007/s00394-020-02230-3.
24. Cuff J, Sanders TAB, Reidlinger DP, Hall WL, Gray R, Darzi J. Urinary alkylresorcinol metabolites as a biomarker of dietary wholegrain intake and of compliance in a randomised dietary intervention trial: results from the CRESSIDA Study. Proceedings of the Nutrition Society 2015;74(OCE1):E42. doi: 10.1017/s0029665115000579.
25. Davis C, Hodgson J, Bryan J, Garg M, Woodman R, Murphy K. Older Australians Can Achieve High Adherence to the Mediterranean Diet during a 6 Month Randomised Intervention; Results from theMedley Study. Nutrients 2017;9(6). doi: 10.3390/nu9060534.
26. Silva FD, Lemos TCC, Sand?ra D, Monteiro M, Perrone D. Fermentation of soybean meal improves isoflavone metabolism after soy biscuit consumption by adults. Journal of the Science of Food and Agriculture 2020;100(7):2991-8. doi: 10.1002/jsfa.10328.
27. Del Rio D, Calani L, Cordero C, Salvatore S, Pellegrini N, Brighenti F. Bioavailability and catabolism of green tea flavan-3-ols in humans. Nutrition 2010;26(11-12):1110-6.
28. Derkach A, Sampson J, Joseph J, Playdon MC, Stolzenberg-Solomon RZ. Effects of dietary sodium on metabolites: the Dietary Approaches to Stop Hypertension (DASH)-Sodium Feeding Study. American Journal of Clinical Nutrition 2017;106(4):1131-41. doi: 10.3945/ajcn.116.150136.
29. Diaz-Rubio ME, Perez-Jimenez J, Martinez-Bartolome MA, Alvarez I, Saura-Calixto F. Regular Consumption of an Antioxidant-rich Juice Improves Oxidative Status and Causes Metabolome Changes in Healthy Adults. Plant Foods for Human Nutrition 2015;70(1):9-14. doi: 10.1007/s11130-014-0455-4.
30. Dickson L, Tenon M, Svilar L, Fanca-Berthon P, Lugan R, Martin JC, Vaillant F, Rogez H. Main Human Urinary Metabolites after Genipap (Genipa americana L.) Juice Intake. Nutrients 2018;10(9):24.
31. Donovan JL, Bell JR, Kasim-Karakas S, German JB, Walzem RL, Hansen RJ, Waterhouse AL. Catechin is present as metabolites in human plasma after consumption of red wine. Journal of Nutrition 1999;129(9):1662-8.
32. Egner PA, Chen JG, Wang JB, Wu Y, Sun Y, Lu JH, Zhu JA, Zhang YH, Chen YS, Friesen MD, et al. Bioavailability of Sulforaphane from Two Broccoli Sprout

Beverages: Results of a Short-term, Cross-over Clinical Trial in Qidong, China. Cancer Prevention Research 2011;4(3):384-95. doi: 10.1158/1940-6207.capr-10-0296.

33. Ellinger S, Reusch A, Henckes L, Ritter C, Zimmermann BF, Ellinger J, Galensa R, Stehle P, Helfrich H-P. Low Plasma Appearance of (+)-Catechin and (-)-Catechin Compared with Epicatechin after Consumption of Beverages Prepared from Nonalkalized or Alkalized Cocoa-A Randomized, Double-Blind Trial. Nutrients 2020;12(1). doi: https://dx.doi.org/10.3390/nu12010231.
34. Erlund I, Freese R, Marniemi J, Hakala P, Alfthan G. Bioavailability of quercetin from berries and the diet. Nutrition & Cancer 2006;54(1):13-7.
35. Favari C, Mena P, Curti C, Istas G, Heiss C, Del Rio D, Rodriguez-Mateos A. Kinetic profile and urinary excretion of phenyl-gamma-valerolactones upon consumption of cranberry: a dose-response relationship. Food & function 2020;11(5):3975-85. doi: http://dx.doi.org/10.1039/d0fo00806k.
36. Felberg I, Farah A, Monteiro MC, Godoy RLD, Pacheco S, Calado V, Donangelo CM. Effect of simultaneous consumption of soymilk and coffee on the urinary excretion of isoflavones, chlorogenic acids and metabolites in healthy adults. Journal of Functional Foods 2015;19:688-99. doi: 10.1016/j.jff.2015.09.059.
37. Feliciano RP, Mills CE, Istas G, Heiss C, Rodriguez-Mateos A. Absorption, Metabolism and Excretion of Cranberry (Poly)phenols in Humans: A Dose Response Study and Assessment of Inter-Individual Variability. Nutrients 2017;9(3):268. doi: 10.3390/nu9030268.
38. Fuchsmann P, Stern MT, Munger LH, Pimentel G, Burton KJ, Vionnet N, Vergeres G. Nutrivolatilomics of urinary and plasma samples to identify candidate biomarkers after cheese, milk, and soy-based drink intake in healthy humans. Journal of Proteome Research 2020;19(10):4019-33. doi: http://dx.doi.org/10.1021/acs.jproteome.0c00324.
39. Fujioka N, Ainslie-Waldman CE, Upadhyaya P, Carmella SG, Fritz VA, Rohwer C, Fan YH, Rauch D, Le C, Hatsukami DK, et al. Urinary 3,3 '-Diindolylmethane: A Biomarker of Glucobrassicin Exposure and Indole-3-Carbinol Uptake in Humans. Cancer Epidemiology Biomarkers & Prevention 2014;23(2):282-7. doi: 10.1158/1055-9965.epi-13-0645.
40. Fujioka N, Ransom BW, Carmella SG, Upadhyaya P, Lindgren BR, Roper-Batker A, Hatsukami DK, Fritz VA, Rohwer C, Hecht SS. Harnessing the Power of Cruciferous Vegetables: Developing a Biomarker for Brassica Vegetable Consumption Using Urinary 3,3'-Diindolylmethane. Cancer Prevention Research 2016;9(10):788-93.
41. Garcia-Aloy M, Ulaszewska M, Franceschi P, Estruel-Amades S, Weinert CH, Tor-Roca A, Urpi-Sarda M, Mattivi F, Andres-Lacueva C. Discovery of Intake Biomarkers of Lentils, Chickpeas, and White Beans by Untargeted LC-MS Metabolomics in Serum and Urine. Molecular nutrition & food research 2020;64(13):e1901137. doi: http://dx.doi.org/10.1002/mnfr.201901137.
42. Garcia-Perez I, Posma JM, Chambers ES, Nicholson JK, Mathers JC, Beckmann M, Draper J, Holmes E, Frost G. An Analytical Pipeline for Quantitative

Characterization of Dietary Intake: Application To Assess Grape Intake. Journal of Agricultural and Food Chemistry 2016;64(11):2423-31. doi: 10.1021/acs.jafc.5b05878.
43. Garcia-Perez I, Posma JM, Gibson R, Chambers ES, Hansen TH, Vestergaard H, Hansen T, Beckmann M, Pedersen O, Elliott P, et al. Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. Lancet Diabetes & Endocrinology 2017;5(3):184-95. doi: 10.1016/s2213-8587(16)30419-3.
44. Garg R, Brennan L, Price RK, Wallace JM, Strain JJ, Gibney MJ, Shewry PR, Ward JL, Garg L, Welch RW. Using NMR-Based Metabolomics to Evaluate Postprandial Urinary Responses Following Consumption of Minimally Processed Wheat Bran or Wheat Aleurone by Men and Women. Nutrients 2016;8(2):96.
45. Gibbons H, McNulty BA, Nugent AP, Walton J, Flynn A, Gibney MJ, Brennan L. A metabolomics approach to the identification of biomarkers of sugar-sweetened beverage intake. American Journal of Clinical Nutrition 2015;101(3):471-7.
46. Gibbons H, Michielsen CJR, Rundle M, Frost G, McNulty BA, Nugent AP, Walton J, Flynn A, Gibney MJ, Brennan L. Demonstration of the utility of biomarkers for dietary intake assessment; proline betaine as an example. Molecular Nutrition & Food Research 2017;61(10). doi: 10.1002/mnfr.201700037.
47. Gómez-Juaristi M, Sarria B, Martínez-López S, Bravo Clemente L, Mateos R. Flavanol Bioavailability in Two Cocoa Products with Different Phenolic Content. A Comparative Study in Humans. Nutrients 2019;11(7):1441-. doi: 10.3390/nu11071441.
48. Gouado I, Schweigert FJ, Ejoh RA, Tchouanguep MF, Camp JV. Systemic levels of carotenoids from mangoes and papaya consumed in three forms (juice, fresh and dry slice). European Journal of Clinical Nutrition 2007;61(10):1180-8. doi: 10.1038/sj.ejcn.1602841.
49. Gu Y, Zhao A, Huang F, Zhang Y, Liu J, Wang C, Jia W, Xie G, Jia W. Very low carbohydrate diet significantly alters the serum metabolic profiles in obese subjects. Journal of Proteome Research 2013;12(12):5801-11.
50. Gurdeniz G, Jensen MG, Meier S, Bech L, Lund E, Dragsted LO. Detecting Beer Intake by Unique Metabolite Patterns. Journal of Proteome Research 2016;15(12):4544-56. doi: 10.1021/acs.jproteome.6b00635.
51. Hagen IV, Helland A, Bratlie M, Midttun Ø, McCann A, Sveier H, Rosenlund G, Mellgren G, Ueland PM, Gudbrandsen OA. TMAO, creatine and 1-methylhistidine in serum and urine are potential biomarkers of cod and salmon intake: a randomised clinical trial in adults with overweight or obesity. European Journal of Nutrition 2020;59(5):2249-59. doi: 10.1007/s00394-019-02076-4.
52. Haldar S, Lee SH, Tan JJ, Chia SC, Henry CJ, Chan ECY. Dose-Dependent Increase in Unconjugated Cinnamic Acid Concentration in Plasma Following Acute Consumption of Polyphenol Rich Curry in the Polyspice Study. Nutrients 2018;10(7). doi: 10.3390/nu10070934.

53. Hanhineva K, Keski-Rahkonen P, Lappi J, Katina K, Pekkinen J, Savolainen O, Timonen O, Paananen J, Mykkanen H, Poutanen K. The Postprandial Plasma Rye Fingerprint Includes Benzoxazinoid-Derived Phenylacetamide Sulfates. Journal of Nutrition 2014;144(7):1016-22. doi: 10.3945/jn.113.187237.
54. Hanhineva K, Lankinen MA, Pedret A, Schwab U, Kolehmainen M, Paananen J, de Mello V, Sola R, Lehtonen M, Poutanen K, et al. Nontargeted Metabolite Profiling Discriminates Diet-Specific Biomarkers for Consumption of Whole Grains, Fatty Fish, and Bilberries in a Randomized Controlled Trial. Journal of Nutrition 2015;145(1):7-17. doi: 10.3945/jn.114.196840.
55. Hauder J, Winker S, Bub A, Ruifer CE, Pignitter M, Somoza V. LC-MS/MS Quantification of Sulforaphane and Indole-3-carbinol Metabolites in Human Plasma and Urine after Dietary Intake of Selenium-Fortified Broccoli. Journal of Agricultural and Food Chemistry 2011;59(15):8047-57. doi: 10.1021/jf201501x.
56. Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas ME, Kochhar S, Stamler J, Holmes E, Elliott P, Nicholson JK. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. American Journal of Clinical Nutrition 2010;92(2):436-43. doi: 10.3945/ajcn.2010.29672.
57. Henning SM, Seeram NP, Zhang Y, Li L, Gao K, Lee R, Wang DC, Zerlin A, Karp H, Thames G, et al. Strawberry consumption is associated with increased antioxidant capacity in serum. Journal of Medicinal Food 2010;13(1):116-22. doi: 10.1089/jmf.2009.0048.
58. Hernandez-Alonso P, Canueto D, Giardin S, Salas-Salvado J, Canellas N, Correig X, Bullo M. Effect of pistachio consumption on the modulation of urinary gut microbiota-related metabolites in prediabetic subjects. Journal of Nutritional Biochemistry 2017;45:48-53. doi: 10.1016/j.jnutbio.2017.04.002.
59. Hernandez-Alonso P, Giardina S, Canueto D, Salas-Salvado J, Canellas N, Bullo M. Changes in Plasma Metabolite Concentrations after a Low-Glycemic Index Diet Intervention. Molecular Nutrition & Food Research 2019;63(1). doi: 10.1002/mnfr.201700975.
60. Hodgson JM, Morton LW, Puddey IB, Beilin LJ, Croft KD. Gallic acid metabolites are markers of black tea intake in humans. Journal of Agricultural & Food Chemistry 2000;48(6):2276-80.
61. Hodgson AB, Randell RK, Boon N, Garczarek U, Mela DJ, Jeukendrup AE, Jacobs DM. Metabolic response to green tea extract during rest and moderate-intensity exercise. Journal of Nutritional Biochemistry 2013;24(1):325-34. doi: 10.1016/j.jnutbio.2012.06.017.
62. Housley L, Magana AA, Hsu A, Beaver LM, Wong CP, Stevens JF, Choi J, Jiang Y, Bella D, Williams DE, et al. Untargeted Metabolomic Screen Reveals Changes in Human Plasma Metabolite Profiles Following Consumption of Fresh Broccoli Sprouts. Molecular Nutrition & Food Research 2018;62(19). doi: 10.1002/mnfr.201700665.

63. Hovelmann Y, Jagels A, Schmid R, Hubner F, Humpf HU. Identification of potential human urinary biomarkers for tomato juice intake by mass spectrometry-based metabolomics. European Journal of Nutrition 2019.
64. Hutchins AM, Martini MC, Olson BA, Thomas W, Slavin JL. Flaxseed influences urinary lignan excretion in a dose-dependent manner in postmenopausal women. Cancer Epidemiology Biomarkers & Prevention 2000;9(10):1113-8.
65. Ibero-Baraibar I, Romo-Hualde A, Gonzalez-Navarro CJ, Zulet MA, Martinez JA. The urinary metabolomic profile following the intake of meals supplemented with a cocoa extract in middle-aged obese subjects. Food & Function 2016;7(4):1924-31. doi: 10.1039/c5fo01191d.
66. Jahns L, Johnson LK, Mayne ST, Cartmel B, Picklo Sr MJ, Ermakov IV, Gellermann W, Whigham LD. Skin and plasma carotenoid response to a provided intervention diet high in vegetables and fruit: uptake and depletion kinetics. American Journal of Clinical Nutrition 2014;100(3):930-7. doi: 10.3945/ajcn.114.086900.
67. Jin Y, Alimbetov D, George T, Gordon MH, Lovegrove JA. A randomised trial to investigate the effects of acute consumption of a blackcurrant juice drink on markers of vascular reactivity and bioavailability of anthocyanins in human subjects. European Journal of Clinical Nutrition 2011;65(7):849-56. doi: 10.1038/ejcn.2011.55.
68. Johansson-Persson A, Barri T, Ulmius M, Onning G, Dragsted LO. LC-QTOF/MS metabolomic profiles in human plasma after a 5-week high dietary fiber intake. Analytical and Bioanalytical Chemistry 2013;405(14):4799-809. doi: 10.1007/s00216-013-6874-5.
69. Kempf K, Herder C, Erlund I, Kolb H, Martin S, Carstensen M, Koenig W, Sundvall J, Bidel S, Kuha S, et al. Effects of coffee consumption on subclinical inflammation and other risk factors for type 2 diabetes: a clinical trial. American Journal of Clinical Nutrition 2010;91(4):950-7. doi: 10.3945/ajcn.2009.28548.
70. Khakimov B, Poulsen SK, Savorani F, Acar E, Gurdeniz G, Larsen TM, Astrup A, Dragsted LO, Engelsen SB. New Nordic Diet versus Average Danish Diet: A Randomized Controlled Trial Revealed Healthy Long-Term Effects of the New Nordic Diet by GC-MS Blood Plasma Metabolomics. Journal of Proteome Research 2016;15(6):1939-54. doi: 10.1021/acs.jproteome.6b00109.
71. Khymenets O, Andres-Lacueva C, Urpi-Sarda M, Vazquez-Fresno R, Mart MM, Reglero G, Torres M, Llorach R. Metabolic fingerprint after acute and under sustained consumption of a functional beverage based on grape skin extract in healthy human subjects. Food & Function 2015;6(4):1288-98.
72. Kremer JI, Gompel K, Bakuradze T, Eisenbrand G, Richling E. Urinary Excretion of Niacin Metabolites in Humans After Coffee Consumption. Molecular Nutrition & Food Research 2018;62(7):e1700735.
73. Kristensen M, Krogholm KS, Frederiksen H, Bugel SH, Rasmussen SE. Urinary excretion of total isothiocyanates from cruciferous vegetables shows high dose-response relationship and may be a useful biomarker for isothiocyanate exposure. European Journal of Nutrition 2007;46(7):377-82.

74. Kuntz S, Rudloff S, Asseburg H, Borsch C, Frohling B, Unger F, Dold S, Spengler B, Rompp A, Kunz C. Uptake and bioavailability of anthocyanins and phenolic acids from grape/blueberry juice and smoothie in vitro and in vivo. British Journal of Nutrition 2015;113(7):1044-55. doi: 10.1017/s0007114515000161.
75. Lacalle-Bergeron L, Portolés T, López FJ, Sancho JV, Ortega-Azorín C, Asensio EM, Coltell O, Corella D. Ultra-Performance Liquid Chromatography-Ion Mobility Separation-Quadruple Time-of-Flight MS (UHPLC-IMS-QTOF MS) Metabolomics for Short-Term Biomarker Discovery of Orange Intake: A Randomized, Controlled Crossover Study. Nutrients 2020;12(7):1916. doi: 10.3390/nu12071916.
76. Landberg R, Aman P, Friberg LE, Vessby B, Adlercreutz H, Kamal-Eldin A. Dose response of whole-grain biomarkers: alkylresorcinols in human plasma and their metabolites in urine in relation to intake. American Journal of Clinical Nutrition 2009;89(1):290-6. doi: 10.3945/ajcn.2008.26709.
77. Langer S, Kennel A, Lodge JK. The influence of juicing on the appearance of blueberry metabolites 2 h after consumption: a metabolite profiling approach. British Journal of Nutrition 2018;119(11):1233-44. doi: 10.1017/S0007114518000855.
78. Lankinen M, Schwab U, Seppanen-Laakso T, Mattila I, Juntunen K, Mykkanen H, Poutanen K, Gylling H, Oresic M. Metabolomic Analysis of Plasma Metabolites That May Mediate Effects of Rye Bread on Satiety and Weight Maintenance in Postmenopausal Women. Journal of Nutrition 2011;141(1):31-6. doi: 10.3945/jn.110.131656.
79. Lappi J, Aura AM, Katina K, Nordlund E, Kolehmainen M, Mykkanen H, Poutanen K. Comparison of postprandial phenolic acid excretions and glucose responses after ingestion of breads with bioprocessed or native rye bran. Food & Function 2013;4(6):972-81. doi: 10.1039/c3fo60078e.
80. Lee J, Ebeler SE, Zweigenbaum JA, Mitchell AE. UHPLC-(ESI)QTOF MS/MS profiling of quercetin metabolites in human plasma postconsumption of applesauce enriched with apple peel and onion. Journal of Agricultural & Food Chemistry 2012;60(34):8510-20.
81. Lennerz BS, Vafai SB, Delaney NF, Clish CB, Deik AA, Pierce KA, Ludwig DS, Mootha VK. Effects of sodium benzoate, a widely used food preservative, on glucose homeostasis and metabolic profiles in humans. Molecular Genetics & Metabolism 2015;114(1):73-9.
82. Li C, Meng X, Winnik B, Lee MJ, Lu H, Sheng S, Buckley B, Yang CS. Analysis of urinary metabolites of tea catechins by liquid chromatography/electrospray ionization mass spectrometry. Chemical Research in Toxicology 2001;14(6):702-7.
83. Li KJ, Borresen EC, Jenkins-Puccetti N, Luckasen G, Ryan EP. Navy Bean and Rice Bran Intake Alters the Plasma Metabolome of Children at Risk for Cardiovascular Disease. Frontiers in Nutrition 2018;4. doi: 10.3389/fnut.2017.00071.

84. Li L, Lyall GK, Martinez-Blazquez JA, Vallejo F, Tomas-Barberan FA, Birch KM, Boesch C. Blood Orange Juice Consumption Increases Flow-Mediated Dilation in Adults with Overweight and Obesity: A Randomized Controlled Trial. Journal of Nutrition 2020;150(9):2287-94. doi: 10.1093/jn/nxaa158.
85. Liu H, Garrett TJ, Su Z, Khoo C, Gu L. UHPLC-Q-Orbitrap-HRMS-based global metabolomics reveal metabolome modifications in plasma of young women after cranberry juice consumption. Journal of Nutritional Biochemistry 2017;45:67-76.
86. Llorach R, Urpi-Sarda M, Jauregui O, Monagas M, Andres-Lacueva C. An LC-MS-Based Metabolomics Approach for Exploring Urinary Metabolome Modifications after Cocoa Consumption. Journal of Proteome Research 2009;8(11):5060-8. doi: 10.1021/pr900470a.
87. Lloyd AJ, Fave G, Beckmann M, Lin WC, Tailliart K, Xie L, Mathers JC, Draper J. Use of mass spectrometry fingerprinting to identify urinary metabolites after consumption of specific foods. American Journal of Clinical Nutrition 2011;94(4):981-91. doi: 10.3945/ajcn.111.017921.
88. Macdonald HA, Hardcastle AC, Duthie GG, Duthie SJ, Aucott L, Sandison R, Shearer MJ, Reid DM. Changes in vitamin biomarkers during a 2-year intervention trial involving increased fruit and vegetable consumption by free-living volunteers. British Journal of Nutrition 2009;102(10):1477-86. doi: 10.1017/s0007114509990377.
89. Madrid-Gambin F, Garcia-Aloy M, Vazquez-Fresno R, Vegas-Lozano E, Jubany M, Misawa K, Hase T, Shimotoyodome A, Andres-Lacueva C. Impact of chlorogenic acids from coffee on urine metabolome in healthy human subjects. Food Research International 2016;89:1064-70. doi: 10.1016/j.foodres.2016.03.038.
90. Madrid-Gambin F, Brunius C, Garcia-Aloy M, Estruel-Amades S, Landberg R, Andres-Lacueva C. Untargeted 1H NMR-Based Metabolomics Analysis of Urine and Serum Profiles after Consumption of Lentils, Chickpeas, and Beans: An Extended Meal Study To Discover Dietary Biomarkers of Pulses. Journal of agricultural and food chemistry 2018;66(27):6997-7005.
91. Madrid-Gambin F, Garcia-Aloy M, Vazquez-Fresno R, Vegas-Lozano E, Sanchez-Pla A, Misawa K, Hase T, Shimotoyodome A, Andres-Lacueva C. Metabolic Signature of a Functional High-Catechin Tea after Acute and Sustained Consumption in Healthy Volunteers through ¹H NMR Based Metabolomics Analysis of Urine. Journal of Agricultural & Food Chemistry 2019;67(11):3118-24.
92. Mahale J, Singh R, Howells LM, Britton RG, Khan SM, Brown K. Detection of Plasma Curcuminoids from Dietary Intake of Turmeric-Containing Food in Human Volunteers. Molecular Nutrition & Food Research 2018;62(16). doi: 10.1002/mnfr.201800267.
93. Martin FP, Rezzi S, Pere-Trepat E, Kamlage B, Collino S, Leibold E, Kastler J, Rein D, Fay LB, Kochhar S. Metabolic effects of dark chocolate consumption on energy, gut microbiota, and stress-related metabolism in free-living subjects. Journal of Proteome Research 2009;8(12):5568-79.

94. Martinez-Lopez S, Sarria B, Gomez-Juaristi M, Goya L, Mateos R, Bravo-Clemente L. Theobromine, caffeine, and theophylline metabolites in human plasma and urine after consumption of soluble cocoa products with different methylxanthine contents. Food Research International 2014;63:446-55. doi: 10.1016/j.foodres.2014.03.009.
95. May DH, Navarro SL, Ruczinski I, Hogan J, Ogata Y, Schwarz Y, Levy L, Holzman T, McIntosh MW, Lampe JW. Metabolomic profiling of urine: response to a randomised, controlled feeding study of select fruits and vegetables, and application to an observational study. British Journal of Nutrition 2013;110(10):1760-70. doi: 10.1017/s000711451300127x.
96. McKeown NM, Marklund M, Ma JT, Ross AB, Lichtenstein AH, Livingston KA, Jacques PF, Rasmussen HM, Blumberg JB, Chen CYO. Comparison of plasma alkylresorcinols (AR) and urinary AR metabolites as biomarkers of compliance in a short-term, whole-grain intervention study. European Journal of Nutrition 2016;55(3):1235-44. doi: 10.1007/s00394-015-0936-8.
97. McNamara AE, Collins C, Harsha PSCS, Gonzalez-Pena D, Gibbons H, McNulty BA, Nugent AP, Walton J, Flynn A, Brennan L. Metabolomic-Based Approach to Identify Biomarkers of Apple Intake. Molecular nutrition & food research 2020;64(11):e1901158. doi: http://dx.doi.org/10.1002/mnfr.201901158.
98. Mennen LI, Sapinho D, Ito H, Bertrais S, Galan P, Hercberg S, Scalbert A. Urinary flavonoids and phenolic acids as biomarkers of intake for polyphenol-rich foods. British Journal of Nutrition 2006;96(1):191-8. doi: 10.1079/bjn20061808.
99. Meuronen T, Lankinen MA, Fauland A, Shimizu BI, de Mello VD, Laaksonen DE, Wheelock CE, Erkkila AT, Schwab US. Intake of Camelina Sativa Oil and Fatty Fish Alter the Plasma Lipid Mediator Profile in Subjects with Impaired Glucose Metabolism - A Randomized Controlled Trial. Prostaglandins Leukotrienes and Essential Fatty Acids 2020;159:102143. doi: http://dx.doi.org/10.1016/j.plefa.2020.102143.
100. Michielsen C, Hangelbroek RWJ, Feskens EJM, Afman LA. Disentangling the Effects of Monounsaturated Fatty Acids from Other Components of a Mediterranean Diet on Serum Metabolite Profiles: A Randomized Fully Controlled Dietary Intervention in Healthy Subjects at Risk of the Metabolic Syndrome. Molecular Nutrition & Food Research 2019;63(9):e1801095.
101. Mills CE, Flury A, Marmet C, Poquet L, Rimoldi SF, Sartori C, Rexhaj E, Brenner R, Allemann Y, Zimmermann D, et al. Mediation of coffee-induced improvements in human vascular function by chlorogenic acids and its metabolites: Two randomized, controlled, crossover intervention trials. Clinical Nutrition 2017;36(6):1520-9. doi: 10.1016/j.clnu.2016.11.013.
102. Moazzami AA, Bondia-Pons I, Hanhineva K, Juntunen K, Antl N, Poutanen K, Mykkanen H. Metabolomics reveals the metabolic shifts following an intervention with rye bread in postmenopausal women- a randomized control trial. Nutrition Journal 2012;11. doi: 10.1186/1475-2891-11-88.
103. Mora-Cubillos X, Tulipani S, Garcia-Aloy M, Bullo M, Tinahones FJ, Andres-Lacueva C. Plasma metabolomic biomarkers of mixed nuts exposure inversely

correlate with severity of metabolic syndrome. Molecular Nutrition & Food Research 2015;59(12):2480-90. doi: 10.1002/mnfr.201500549.
104. Mulder TP, Rietveld AG, van Amelsvoort JM. Consumption of both black tea and green tea results in an increase in the excretion of hippuric acid into urine. American Journal of Clinical Nutrition 2005;81(1):256S-60S.
105. Munger LH, Trimigno A, Picone G, Freiburghaus C, Pimentel G, Burton KJ, Pralong FP, Vionnet N, Capozzi F, Badertscher R, et al. Identification of Urinary Food Intake Biomarkers for Milk, Cheese, and Soy-Based Drink by Untargeted GC-MS and NMR in Healthy Humans. Journal of Proteome Research 2017;16(9):3321-35. doi: 10.1021/acs.jproteome.7b00319.
106. Nieman DC, Gillitt N, Jin F, Henson DA, Kennerly K, Shanely RA, Ore B, Su M, Schwartz S. Chia Seed Supplementation and Disease Risk Factors in Overweight Women: A Metabolomics Investigation. Journal of Alternative and Complementary Medicine 2012;18(7):700-8. doi: 10.1089/acm.2011.0443.
107. Nilsson AC, Ostman EM, Knudsen KEB, Holst JJ, Bjorck IME. A Cereal-Based Evening Meal Rich in Indigestible Carbohydrates Increases Plasma Butyrate the Next Morning. Journal of Nutrition 2010;140(11):1932-6. doi: 10.3945/jn.110.123604.
108. O'Sullivan A, Gibney MJ, Brennan L. Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies. American Journal of Clinical Nutrition 2011;93(2):314-21. doi: 10.3945/ajcn.110.000950.
109. Ostertag LM, Philo M, Colquhoun IJ, Tapp HS, Saha S, Duthie GG, Kemsley EK, de Roos B, Kroon PA, Le Gall G. Acute Consumption of Flavan-3-ol-Enriched Dark Chocolate Affects Human Endogenous Metabolism. Journal of Proteome Research 2017;16(7):2516-26. doi: 10.1021/acs.jproteome.7b00089.
110. Paetau I, Khachik F, Brown ED, Beecher GR, Kramer TR, Chittams J, Clevidence BA. Chronic ingestion of lycopene-rich tomato juice or lycopene supplements significantly increases plasma concentrations of lycopene and related tomato carotenoids in humans. American Journal of Clinical Nutrition 1998;68(6):1187-95.
111. Pereira-Caro G, Borges G, van der Hooft J, Clifford MN, Del Rio D, Lean MEJ, Roberts SA, Kellerhals MB, Crozier A. Orange juice (poly)phenols are highly bioavailable in humans. Am J Clin Nutr 2014;100(5):1378-84. doi: 10.3945/ajcn.114.090282.
112. Pereira-Caro G, Polyviou T, Ludwig IA, Nastase A-M, Moreno-Rojas JM, Garcia AL, Malkova D, Crozier A. Bioavailability of orange juice (poly)phenols: the impact of short-term cessation of training by male endurance athletes. Am J Clin Nutr 2017;106(3):791-800.
113. Pereira-Caro G, Clifford MN, Polyviou T, Ludwig IA, Alfheeaid H, Moreno-Rojas JM, Garcia AL, Malkova D, Crozier A. Plasma pharmacokinetics of (poly)phenol metabolites and catabolites after ingestion of orange juice by endurance trained men. Free Radical Biology and Medicine 2020;160:784-95. doi: http://dx.doi.org/10.1016/j.freeradbiomed.2020.09.007.

114. Perera T, Young MR, Zhang Z, Murphy G, Colburn NH, Lanza E, Hartman TJ, Cross AJ, Bobe G. Identification and monitoring of metabolite markers of dry bean consumption in parallel human and mouse studies. Molecular Nutrition & Food Research 2015;59(4):795-806.
115. Pezdirc K, Hutchesson MJ, Williams RL, Rollo ME, Burrows TL, Wood LG, Oldmeadow C, Collins CE. Consuming High-Carotenoid Fruit and Vegetables Influences Skin Yellowness and Plasma Carotenoids in Young Women: A Single-Blind Randomized Crossover Trial. Journal of the Academy of Nutrition & Dietetics 2016;116(8):1257-65. doi: 10.1016/j.jand.2016.03.012.
116. Pujos-Guillot E, Hubert J, Martin JF, Lyan B, Quintana M, Claude S, Chabanas B, Rothwell JA, Bennetau-Pelissero C, Scalbert A, et al. Mass Spectrometry-based Metabolomics for the Discovery of Biomarkers of Fruit and Vegetable Intake: Citrus Fruit as a Case Study. Journal of Proteome Research 2013;12(4):1645-59. doi: 10.1021/pr300997c.
117. Quifer-Rada P, Martinez-Huelamo M, Chiva-Blanch G, Jauregui O, Estruch R, Lamuela-Raventos RM. Urinary Isoxanthohumol Is a Specific and Accurate Biomarker of Beer Consumption. Journal of Nutrition 2014;144(4):484-8. doi: 10.3945/jn.113.185199.
118. Radjursoga M, Lindqvist HM, Pedersen A, Karlsson BG, Malmodin D, Ellegard L, Winkvist A. Nutritional Metabolomics: Postprandial Response of Meals Relating to Vegan, Lacto-Ovo Vegetarian, and Omnivore Diets. Nutrients 2018;10(8). doi: 10.3390/nu10081063.
119. Rangel-Huerta OD, Aguilera CM, Perez-de-la-Cruz A, Vallejo F, Tomas-Barberan F, Gil A, Mesa MD. A serum metabolomics-driven approach predicts orange juice consumption and its impact on oxidative stress and inflammation in subjects from the BIONAOS study. Molecular Nutrition & Food Research 2017;61(2). doi: 10.1002/mnfr.201600120.
120. Rasmussen LG, Winning H, Savorani F, Toft H, Larsen TM, Dragsted LO, Astrup A, Engelsen SB. Assessment of the Effect of High or Low Protein Diet on the Human Urine Metabolome as Measured by NMR. Nutrients 2012;4(2):112-31. doi: 10.3390/nu4020112.
121. Rebholz CM, Lichtenstein AH, Zheng ZH, Appel LJ, Coresh J. Serum untargeted metabolomic profile of the Dietary Approaches to Stop Hypertension (DASH) dietary pattern. American Journal of Clinical Nutrition 2018;108(2):243-55. doi: 10.1093/ajcn/nqy099.
122. Regueiro J, Vallverdu-Queralt A, Simal-Gandara J, Estruch R, Lamuela-Raventos RM. Urinary tartaric acid as a potential biomarker for the dietary assessment of moderate wine consumption: a randomised controlled trial. British Journal of Nutrition 2014;111(9):1680-5. doi: 10.1017/s0007114513004108.
123. Reisdorph NA, Hendricks AE, Tang M, Doenges KA, Reisdorph RM, Tooker BC, Quinn K, Borengasser SJ, Nkrumah-Elie Y, Frank DN, et al. Nutrimetabolomics reveals food-specific compounds in urine of adults consuming a DASH-style diet. Scientific reports 2020;10(1):1157. doi: https://dx.doi.org/10.1038/s41598-020-57979-8.

124. Renouf M, Marmet C, Guy PA, Beaumont M, Lepage M, Williamson G, Dionisi F. Dose-response plasma appearance of green tea catechins in adults. Molecular Nutrition & Food Research 2013;57(5):833-9.
125. Rodriguez-Mateos A, Oruna-Concha MJ, Kwik-Uribe C, Vidal A, Spencer JPE. Influence of sugar type on the bioavailability of cocoa flavanols. British Journal of Nutrition 2012;108(12):2243-50. doi: 10.1017/s0007114512000475.
126. Ross AB, Bourgeois A, Macharia HN, Kochhar S, Jebb SA, Brownlee IA, Seal CJ. Plasma alkylresorcinols as a biomarker of whole-grain food consumption in a large population: results from the WHOLEheart Intervention Study. American Journal of Clinical Nutrition 2012;95(1):204-11.
127. Ross AB, Pere-Trepat E, Montoliu I, Martin FPJ, Collino S, Moco S, Godin JP, Cleroux M, Guy PA, Breton I, et al. A Whole-Grain-Rich Diet Reduces Urinary Excretion of Markers of Protein Catabolism and Gut Microbiota Metabolism in Healthy Men after One Week. Journal of Nutrition 2013;143(6):766-73. doi: 10.3945/jn.112.172197.
128. Ross AB, Svelander C, Undeland I, Pinto R, Sandberg A-S. 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. Journal of Nutrition 2015;145(11):2456-63. doi: 10.3945/jn.115.214262.
129. Roura E, Almajano MP, Bilbao MLM, Andres-Lacueva C, Estruch R, Lamuela-Raventos RM. Human urine: Epicatechin metabolites and antioxidant activity after cocoa beverage intake. Free Radical Research 2007;41(8):943-9.
130. Roura E, Andres-Lacueva C, Estruch R, Mata-Bilbao ML, Izquierdo-Pulido M, Waterhouse AL, Lamuela-Raventos RM. Milk does not affect the bioavailability of cocoa powder flavonoid in healthy human. Annals of Nutrition and Metabolism 2007;51(6):493-8. doi: 10.1159/000111473.
131. Roura E, Andrés-Lacueva C, Estruch R, Lourdes Mata Bilbao M, Izquierdo-Pulido M, Lamuela-Raventós RM. The effects of milk as a food matrix for polyphenols on the excretion profile of cocoa (-)-epicatechin metabolites in healthy human subjects. British Journal of Nutrition 2008;100(4):846-51.
132. Schaer MY, Curtis PJ, Hazim S, Ostertag LM, Kay CD, Potter JF, Cassidy A. Orange juice-derived flavanone and phenolic metabolites do not acutely affect cardiovascular risk biomarkers: a randomized, placebo-controlled, crossover trial in men at moderate risk of cardiovascular disease. American Journal of Clinical Nutrition 2015;101(5):931-8. doi: 10.3945/ajcn.114.104364.
133. Schar MY, Corona G, Soycan G, Dine C, Kristek A, Alsharif SNS, Behrends V, Lovegrove A, Shewry PR, Spencer JPE. Excretion of Avenanthramides, Phenolic Acids and their Major Metabolites Following Intake of Oat Bran. Molecular Nutrition & Food Research 2018;62(2):01.
134. Scheffler L, Sauermann Y, Heinlein A, Sharapa C, Buettner A. Detection of volatile metabolites derived from garlic (Allium sativum) in human urine. Metabolites 2016;6 (4) (no pagination)(43).

135. Schmedes M, Aadland EK, Sundekilde UK, Jacques H, Lavigne C, Graff IE, Eng O, Holthe A, Mellgren G, Young JF, et al. Lean-seafood intake decreases urinary markers of mitochondrial lipid and energy metabolism in healthy subjects: Metabolomics results from a randomized crossover intervention study. Molecular Nutrition & Food Research 2016;60(7):1661-72. doi: 10.1002/mnfr.201500785.
136. Schmedes M, Balderas C, Aadland EK, Jacques H, Lavigne C, Graff IE, Eng O, Holthe A, Mellgren G, Young JF, et al. The Effect of Lean-Seafood and Non-Seafood Diets on Fasting and Postprandial Serum Metabolites and Lipid Species: Results from a Randomized Crossover Intervention Study in Healthy Adults. Nutrients 2018;10(5). doi: 10.3390/nu10050598.
137. Shi L, Brunius C, Lindelof M, Abou Shameh S, Wu HX, Lee I, Landberg R, Moazzami AA. Targeted metabolomics reveals differences in the extended postprandial plasma metabolome of healthy subjects after intake of whole-grain rye porridges versus refined wheat bread. Molecular Nutrition & Food Research 2017;61(7). doi: 10.1002/mnfr.201600924.
138. Harsha P, Wahab RA, Cuparencu C, Dragsted LO, Brennan L. A Metabolomics Approach to the Identification of Urinary Biomarkers of Pea Intake. Nutrients 2018;10(12). doi: 10.3390/nu10121911.
139. Stalmach A, Mullen W, Barron D, Uchida K, Yokota T, Cavin C, Steiling H, Williamson G, Crozier A. Metabolite profiling of hydroxycinnamate derivatives in plasma and urine after the ingestion of coffee by humans: identification of biomarkers of coffee consumption. Drug Metabolism & Disposition 2009;37(8):1749-58.
140. Stalmach A, Williamson G, Crozier A. Impact of dose on the bioavailability of coffee chlorogenic acids in humans. Food & Function 2014;5(8):1727-37. doi: 10.1039/c4fo00316k.
141. Stanstrup J, Schou SS, Holmer-Jensen J, Hermansen K, Dragsted LO. Whey Protein Delays Gastric Emptying and Suppresses Plasma Fatty Acids and Their Metabolites Compared to Casein, Gluten, and Fish Protein. Journal of Proteome Research 2014;13(5):2396-408. doi: 10.1021/pr401214w.
142. Stea TH, Mansoor MA, Wandel M, Uglem S, Frolich W. Changes in predictors and status of homocysteine in young male adults after a dietary intervention with vegetables, fruits and bread. European Journal of Nutrition 2008;47(4):201-9. doi: 10.1007/s00394-008-0714-y.
143. Stella C, Beckwith-Hall B, Cloarec O, Holmes E, Lindon JC, Powell J, van der Ouderaa F, Bingham S, Cross AJ, Nicholson JK. Susceptibility of human metabolic phenotypes to dietary modulation. Journal of Proteome Research 2006;5(10):2780-8. doi: 10.1021/pr060265y.
144. Trimigno A, Munger L, Picone G, Freiburghaus C, Pimentel G, Vionnet N, Pralong F, Capozzi F, Badertscher R, Vergeres G. GC-MS Based Metabolomics and NMR Spectroscopy Investigation of Food Intake Biomarkers for Milk and Cheese in Serum of Healthy Humans. Metabolites 2018;8(2). doi: 10.3390/metabo8020026.

145. Trimigno A, Khakimov B, Savorani F, Poulsen SK, Astrup A, Dragsted LO, Engelsen SB. Human urine 1H NMR metabolomics reveals alterations of the protein and carbohydrate metabolism when comparing habitual Average Danish diet vs. healthy New Nordic diet. Nutrition 2020;79-80:110867. doi: http://dx.doi.org/10.1016/j.nut.2020.110867.
146. Tsang C, Higgins S, Duthie GG, Duthie SJ, Howie M, Mullen W, Lean MEJ, Crozier A. The influence of moderate red wine consumption on antioxidant status and indices of oxidative stress associated with CHD in healthy volunteers. British Journal of Nutrition 2005;93(2):233-40.
147. Ulaszewska MM, Trost K, Stanstrup J, Tuohy KM, Franceschi P, Chong MFF, George T, Minihane AM, Lovegrove JA, Mattivi F. Urinary metabolomic profiling to identify biomarkers of a flavonoid-rich and flavonoid-poor fruits and vegetables diet in adults: the FLAVURS trial. Metabolomics 2016;12(2). doi: 10.1007/s11306-015-0935-z.
148. Ulaszewska MM, Koutsos A, Trošt K, Stanstrup J, Garcia-Aloy M, Scholz M, Fava F, Natella F, Scaccini C, Vrhovsek U, et al. Two apples a day modulate human:microbiome co-metabolic processing of polyphenols, tyrosine and tryptophan. European Journal of Nutrition 2020;59(8):3691-714. doi: 10.1007/s00394-020-02201-8.
149. Urpi-Sarda M, Boto-Ordonez M, Queipo-Ortuno MI, Tulipani S, Corella D, Estruch R, Tinahones FJ, Andres-Lacueva C. Phenolic and microbial-targeted metabolomics to discovering and evaluating wine intake biomarkers in human urine and plasma. Electrophoresis 2015;36(18):2259-68. doi: 10.1002/elps.201400506.
150. Van Dorsten FA, Daykin CA, Mulder TPJ, Van Duynhoven JPM. Metabonomics approach to determine metabolic differences between green tea and black tea consumption. Journal of Agricultural and Food Chemistry 2006;54(18):6929-38. doi: 10.1021/jf061016x.
151. Vazquez-Fresno R, Llorach R, Urpi-Sarda M, Lupianez-Barbero A, Estruch R, Corella D, Fito M, Aros F, Ruiz-Canela M, Salas-Salvado J, et al. Metabolomic Pattern Analysis after Mediterranean Diet Intervention in a Nondiabetic Population: A 1- and 3-Year Follow-up in the PREDIMED Study. Journal of Proteome Research 2015;14(1):531-40. doi: 10.1021/pr5007894.
152. Vázquez-Manjarrez N, Weinert CH, Ulaszewska MM, Mack CI, Micheau P, Pétéra M, Durand S, Pujos-Guillot E, Egert B, Mattivi F, et al. Discovery and Validation of Banana Intake Biomarkers Using Untargeted Metabolomics in Human Intervention and Cross-sectional Studies. Journal of Nutrition 2019;149(10):1701-13. doi: 10.1093/jn/nxz125.
153. Vergne S, Bennetau-Pelissero C, Lamothe V, Chantre P, Potier M, Asselineau J, Perez P, Durand M, Moore N, Sauvant P. Higher bioavailability of isoflavones after a single ingestion of a soya-based supplement than a soya-based food in young healthy males. British Journal of Nutrition 2008;99(2):333-44.
154. Vetrani C, Rivellese AA, Annuzzi G, Adiels M, Boren J, Mattila I, Oresic M, Aura AM. Metabolic transformations of dietary polyphenols: comparison between

in vitro colonic and hepatic models and in vivo urinary metabolites. Journal of Nutritional Biochemistry 2016;33:111-8. doi: 10.1016/j.jnutbio.2016.03.007.
155. Villano D, Lopez-Chillon MT, Zafrilla P, Moreno DA. Bioavailability of broccoli sprouts in different human overweight populations. Journal of Functional Foods 2019;59:337-44. doi: 10.1016/j.jff.2019.05.052.
156. Wang ZN, Bergeron N, Levison BS, Li XMS, Chiu S, Jia X, Koeth RA, Li L, Wu YP, Tang WHW, et al. Impact of chronic dietary red meat, white meat, or non-meat protein on trimethylamine N-oxide metabolism and renal excretion in healthy men and women. European Heart Journal 2019;40(7):583-+. doi: 10.1093/eurheartj/ehy799.
157. Wedekind R, Keski-Rahkonen P, Robinot N, Viallon V, Ferrari P, Engel E, Boutron-Ruault MC, Mahamat-Saleh Y, Mancini FR, Kuhn T, et al. Syringol metabolites as new biomarkers for smoked meat intake. American Journal of Clinical Nutrition 2019;110(6):1424-33. doi: http://dx.doi.org/10.1093/ajcn/nqz222.
158. Wedekind R, Kiss A, Keski-Rahkonen P, Viallon V, Rothwell JA, Cross AJ, Rostgaard-Hansen AL, Sandanger TM, Jakszyn P, Schmidt JA, et al. A metabolomic study of red and processed meat intake and acylcarnitine concentrations in human urine and blood. American Journal of Clinical Nutrition 2020;112(2):381-8. doi: 10.1093/ajcn/nqaa140.
159. Wellington N, Shanmuganathan M, de Souza RJ, Zulyniak MA, Azab S, Bloomfield J, Mell A, Ly R, Desai D, Anand SS, et al. Metabolic Trajectories Following Contrasting Prudent and Western Diets from Food Provisions: Identifying Robust Biomarkers of Short-Term Changes in Habitual Diet. Nutrients 2019;11(10). doi: 10.3390/nu11102407.
160. Wiczkowski W, Szawara-Nowak D, Romaszko J. The impact of red cabbage fermentation on bioavailability of anthocyanins and antioxidant capacity of human plasma. Food Chemistry 2016;190:730-40. doi: 10.1016/j.foodchem.2015.06.021.
161. Wiseman H, Casey K, Bowey EA, Duffy R, Davies M, Rowland IR, Lloyd AS, Murray A, Thompson R, Clarke DB. Influence of 10 wk of soy consumption on plasma concentrations and excretion of isoflavonoids and on gut microflora metabolism in healthy adults. American Journal of Clinical Nutrition 2004;80(3):692-9.
162. Yin X, Gibbons H, Rundle M, Frost G, McNulty BA, Nugent AP, Walton J, Flynn A, Gibney MJ, Brennan L. Estimation of chicken intake by adults using metabolomics-derived markers. Journal of Nutrition 2017;147(10):1850-7.
163. Yin X, Gibbons H, Rundle M, Frost G, McNulty BA, Nugent AP, Walton J, Flynn A, Brennan L. The Relationship between Fish Intake and Urinary Trimethylamine-N-Oxide. Molecular nutrition & food research 2020;64(3):e1900799. doi: http://dx.doi.org/10.1002/mnfr.201900799.
164. Zamora-Ros R, Urpi-Sarda M, Lamuela-Raventos RM, Estruch R, Vazquez-Agell M, Serrano-Martinez M, Jaeger W, Andres-Lacueva C. Diagnostic performance

of urinary resveratrol metabolites as a biomarker of moderate wine consumption. Clinical Chemistry 2006;52(7):1373-80. doi: 10.1373/clinchem.2005.065870.
165. Zhao S, Liu H, Su Z, Khoo C, Gu L. Identifying Cranberry Juice Consumers with Predictive OPLS-DA Models of Plasma Metabolome and Validation of Cranberry Juice Intake Biomarkers in a Double-Blinded, Randomized, Placebo-Controlled, Cross-Over Study. Molecular nutrition & food research 2020;64(11):e1901242. doi: http://dx.doi.org/10.1002/mnfr.201901242.
166. Zheng H, Yde CC, Dalsgaard TK, Arnberg K, Molgaard C, Michaelsen KF, Larnkjaer A, Bertram HC. Nuclear magnetic resonance-based metabolomics reveals that dairy protein fractions affect urinary urea excretion differently in overweight adolescents. European Food Research and Technology 2015;240(3):489-97. doi: 10.1007/s00217-014-2347-0.
167. Zheng H, Yde CC, Clausen MR, Kristensen M, Lorenzen J, Astrup A, Bertram HC. Metabolomics investigation to shed light on cheese as a possible piece in the French paradox puzzle. Journal of Agricultural & Food Chemistry 2015;63(10):2830-9.
168. Zheng H, Lorenzen JK, Astrup A, Larsen LH, Yde CC, Clausen MR, Bertram HC. Metabolic Effects of a 24-Week Energy-Restricted Intervention Combined with Low or High Dairy Intake in Overweight Women: An NMR-Based Metabolomics Investigation. Nutrients 2016;8(3). doi: 10.3390/nu8030108.
169. Zhong S, Sandhu A, Edirisinghe I, Burton-Freeman B. Characterization of Wild Blueberry Polyphenols Bioavailability and Kinetic Profile in Plasma over 24-h Period in Human Subjects. Molecular Nutrition & Food Research 2017;61(12):1700405. doi: 10.1002/mnfr.201700405.
170. Aguilar SS, Wengreen HJ, Lefevre M, Madden GJ, Gast J. Skin Carotenoids: A Biomarker of Fruit and Vegetable Intake in Children. Journal of the Academy of Nutrition & Dietetics 2014;114(8):1174-80. doi: 10.1016/j.jand.2014.04.026.
171. Azab SM, de Souza RJ, Teo KK, Anand SS, Williams NC, Holzschuher J, McGlory C, Philips SM, Britz-McKibbin P. Serum nonesterified fatty acids have utility as dietary biomarkers of fat intake from fish, fish oil, and dairy in women. J Lipid Res 2020;61(6):933-44. doi: 10.1194/jlr.D120000630.
172. Allen NE, Grace PB, Ginn A, Travis RC, Roddam AW, Appleby PN, Key T. Phytanic acid: measurement of plasma concentrations by gas-liquid chromatography-mass spectrometry analysis and associations with diet and other plasma fatty acids. British Journal of Nutrition 2008;99(3):653-9.
173. Aubertin-Leheudre M, Koskela A, Samaletdin A, Adlercreutz H. Plasma alkylresorcinol metabolites as potential biomarkers of whole-grain wheat and rye cereal fibre intakes in women. British Journal of Nutrition 2010;103(3):339-43. doi: 10.1017/s0007114509992315.
174. Bouchard-Mercier A, Rudkowska I, Lemieux S, Couture P, Vohl MC. The metabolic signature associated with the Western dietary pattern: a cross-sectional study. Nutrition Journal 2013;12. doi: 10.1186/1475-2891-12-158.
175. Chandler PD, Balasubramanian R, Paynter N, Giulianini F, Fung T, Tinker LF, Snetselaar L, Liu S, Eaton C, Tobias DK, et al. Metabolic signatures associated

with Western and Prudent dietary patterns in women. American Journal of Clinical Nutrition 2020;112(2):268-83. doi: 10.1093/ajcn/nqaa131.
176. Chun OK, Chung SJ, Song WO. Urinary isoflavones and their metabolites validate the dietary isoflavone intakes in US adults. Journal of the American Dietetic Association 2009;109(2):245-54. doi: 10.1016/j.jada.2008.10.055.
177. Chung HY, Ferreira AL, Epstein S, Paiva SA, Castaneda-Sceppa C, Johnson EJ. Site-specific concentrations of carotenoids in adipose tissue: relations with dietary and serum carotenoid concentrations in healthy adults. American Journal of Clinical Nutrition 2009;90(3):533-9. doi: 10.3945/ajcn.2009.27712.
178. De Filippis F, Pellegrini N, Vannini L, Jeffery IB, La Storia A, Laghi L, Serrazanetti DI, Di Cagno R, Ferrocino I, Lazzi C, et al. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut 2016;65(11):1812-21. doi: 10.1136/gutjnl-2015-309957.
179. de Pee S, Bloem MW, Tjiong R, Martini E, Satoto, Gorstein J, Shrimpton R, Muhilal. Who has a high vitamin A intake from plant foods, but a low serum retinol concentration? Data from women in Indonesia. European Journal of Clinical Nutrition 1999;53(4):288-97. doi: 10.1038/sj.ejcn.1600724.
180. Dorgan JF, Jung S, Dallal CM, Zhan M, Stennett CA, Zhang Y, Eckert RL, Snetselaar LG, Van Horn L. Alcohol consumption and serum metabolite concentrations in young women. Cancer Causes & Control 2020;31(2):113-26. doi: 10.1007/s10552-019-01256-1.
181. Du D, Bruno R, Blizzard L, Venn A, Dwyer T, Smith KJ, Magnussen CG, Gall S. The metabolomic signatures of alcohol consumption in young adults. European Journal of Preventive Cardiology 2020;27(8):840-9. doi: 10.1177/2047487319834767.
182. Edmands WMB, Ferrari P, Rothwell JA, Rinaldi S, Slimani N, Barupal DK, Biessy C, Jenab M, Clavel-Chapelon F, Fagherazzi G, et al. Polyphenol metabolome in human urine and its association with intake of polyphenol-rich foods across European countries. American Journal of Clinical Nutrition 2015;102(4):905-13. doi: 10.3945/ajcn.114.101881.
183. Floegel A, von Ruesten A, Drogan D, Schulze MB, Prehn C, Adamski J, Pischon T, Boeing H. Variation of serum metabolites related to habitual diet: a targeted metabolomic approach in EPIC-Potsdam. European Journal of Clinical Nutrition 2013;67(10):1100-8. doi: 10.1038/ejcn.2013.147.
184. Frankenfeld CL, Patterson RE, Horner NK, Neuhouser ML, Skor HE, Kalhorn TF, Howald WN, Lampe JW. Validation of a soy food-frequency questionnaire and evaluation of correlates of plasma isoflavone concentrations in postmenopausal women. American Journal of Clinical Nutrition 2003;77(3):674-80.
185. Frankenfeld CL. Dairy consumption is a significant correlate of urinary equol concentration in a representative sample of US adults. Am J Clin Nutr 2011;93(5):1109-16. doi: 10.3945/ajcn.111.011825.
186. Garcia-Aloy M, Llorach R, Urpi-Sarda M, Tulipani S, Estruch R, Martinez-Gonzalez MA, Corella D, Fito M, Ros E, Salas-Salvado J, et al. Novel

multimetabolite prediction of walnut consumption by a urinary biomarker model in a free-living population: the PREDIMED study. Journal of Proteome Research 2014;13(7):3476-83.

187. Garcia-Aloy M, Llorach R, Urpi-Sarda M, Jauregui O, Corella D, Ruiz-Canela M, Salas-Salvado J, Fito M, Ros E, Estruch R, et al. A metabolomics-driven approach to predict cocoa product consumption by designing a multimetabolite biomarker model in free-living subjects from the PREDIMED study. Molecular Nutrition & Food Research 2015;59(2):212-20.
188. Garcia-Aloy M, Llorach R, Urpi-Sarda M, Tulipani S, Salas-Salvado J, Angel Martinez-Gonzalez M, Corella D, Fito M, Estruch R, Serra-Majem L, et al. Nutrimetabolomics fingerprinting to identify biomarkers of bread exposure in a free-living population from the PREDIMED study cohort. Metabolomics 2015;11(1):155-65. doi: 10.1007/s11306-014-0682-6.
189. Gibbons H, Carr E, McNulty BA, Nugent AP, Walton J, Flynn A, Gibney MJ, Brennan L. Metabolomic-based identification of clusters that reflect dietary patterns. Molecular Nutrition & Food Research 2017;61(10):1601050. doi: 10.1002/mnfr.201601050.
190. Gibson R, Lau C-HE, Loo RL, Ebbels TMD, Chekmeneva E, Dyer AR, Miura K, Ueshima H, Zhao L, Daviglus ML, et al. The association of fish consumption and its urinary metabolites with cardiovascular risk factors: the International Study of Macro-/Micronutrients and Blood Pressure (INTERMAP). American Journal of Clinical Nutrition 2020;111(2):280-90. doi: 10.1093/ajcn/nqz293.
191. Oude Griep LM, Chekmeneva E, Stamler J, Van Horn L, Chan Q, Ebbels TMD, Holmes E, Frost GS, Elliott P. Urinary hippurate and proline betaine relative to fruit intake, blood pressure, and body mass index. Proceedings of the Nutrition Society 2016;75 (OCE3):E178.
192. Guertin KA, Moore SC, Sampson JN, Huang W-Y, Xiao Q, Stolzenberg-Solomon RZ, Sinha R, Cross AJ. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr 2014;100(1):208-17. doi: https://dx.doi.org/10.3945/ajcn.113.078758.
193. Hanhineva K, Brunius C, Andersson A, Marklund M, Juvonen R, Keski-Rahkonen P, Auriola S, Landberg R. Discovery of urinary biomarkers of whole grain rye intake in free-living subjects using nontargeted LC-MS metabolite profiling. Molecular Nutrition & Food Research 2015;59(11):2315-25. doi: 10.1002/mnfr.201500423.
194. Harada S, Takebayashi T, Kurihara A, Akiyama M, Suzuki A, Hatakeyama Y, Sugiyama D, Kuwabara K, Takeuchi A, Okamura T, et al. Metabolomic profiling reveals novel biomarkers of alcohol intake and alcohol-induced liver injury in community-dwelling men. Environmental Health and Preventive Medicine 2016;21(1):18-26.
195. Hodgson JM, Devine A, Puddey IB, Beilby J, Prince RL. Drinking tea is associated with lower plasma total homocysteine in older women. Asia Pacific Journal of Clinical Nutrition 2006;15(2):253-8.

196. Hruby A, Dennis C, Jacques PF. Dairy Intake in 2 American Adult Cohorts Associates with Novel and Known Targeted and Nontargeted Circulating Metabolites. Journal of Nutrition 2020;150(5):1272-83. doi: 10.1093/jn/nxaa021.
197. Hustad KS, Rundblad A, Ottestad I, Christensen JJ, Holven KB, Ulven SM. Comprehensive lipid and metabolite profiling in healthy adults with low and high consumption of fatty fish: A cross-sectional study. British Journal of Nutrition 2020. doi: http://dx.doi.org/10.1017/S0007114520002305.
198. Jaceldo-Siegl K, Fraser GE, Chan J, Franke A, Sabaté J. Validation of soy protein estimates from a food-frequency questionnaire with repeated 24-h recalls and isoflavonoid excretion in overnight urine in a Western population with a wide range of soy intakes. Am J Clin Nutr 2008;87(5):1422-7.
199. Klebanoff MA, Levine RJ, Dersimonian R, Clemens JD, Wilkins DG. Serum caffeine and paraxanthine as markers for reported caffeine intake in pregnancy. Ann Epidemiol 1998;8(2):107-11.
200. Landberg R, Wierzbicka R, Shi L, Nybacka S, Kamal-Eldin A, Hedblad B, Lindroos AK, Winkvist A, Forslund HB. New alkylresorcinol metabolites in spot urine as biomarkers of whole grain wheat and rye intake in a Swedish middle-aged population. European Journal of Clinical Nutrition 2018;72(10):1439-46. doi: 10.1038/s41430-017-0079-5.
201. Lau CE, Siskos AP, Maitre L, Robinson O, Athersuch TJ, Want EJ, Urquiza J, Casas M, Vafeiadi M, Roumeliotaki T, et al. Determinants of the urinary and serum metabolome in children from six European populations. BMC Medicine 2018;16(1):202.
202. Lécuyer L, Dalle C, Micheau P, Pétéra M, Centeno D, Lyan B, Lagree M, Galan P, Hercberg S, Rossary A, et al. Untargeted plasma metabolomic profiles associated with overall diet in women from the SU.VI.MAX cohort. European Journal of Nutrition 2020;59(8):3425-39. doi: 10.1007/s00394-020-02177-5.
203. Lindqvist HM, Rådjursöga M, Malmodin D, Winkvist A, Ellegård L. Serum metabolite profiles of habitual diet: evaluation by 1 H-nuclear magnetic resonance analysis. American Journal of Clinical Nutrition 2019;110(1):53-62. doi: 10.1093/ajcn/nqz032.
204. MacDougall CR, Hill CE, Jahren AH, Savla J, Riebl SK, Hedrick VE, Raynor HA, Dunsmore JC, Frisard MI, Davy BM. The δ13C Value of Fingerstick Blood Is a Valid, Reliable, and Sensitive Biomarker of Sugar-Sweetened Beverage Intake in Children and Adolescents. Journal of Nutrition 2018;148(1):147-52. doi: 10.1093/jn/nxx017.
205. Mack CI, Egert B, Liberto E, Weinert CH, Bub A, Hoffmann I, Bicchi C, Kulling SE, Cordero C. Robust Markers of Coffee Consumption Identified Among the Volatile Organic Compounds in Human Urine. Molecular nutrition & food research 2019;63(10):e1801060. doi: https://dx.doi.org/10.1002/mnfr.201801060.
206. Malik VS, Guasch-Ferre M, Hu FB, Townsend MK, Zeleznik OA, Eliassen AH, Tworoger SS, Karlson EW, Costenbader KH, Ascherio A, et al. Identification of Plasma Lipid Metabolites Associated with Nut Consumption in US Men and Women. Journal of Nutrition 2019;16:16.

207. Markhus MW, Graff IE, Dahl L, Seldal CF, Skotheim S, Braarud HC, Stormark KM, Malde MK. Establishment of a seafood index to assess the seafood consumption in pregnant women. Food & Nutrition Research 2013;57:1-11. doi: 10.3402/fnr.v57i0.19272.
208. Maskarinec G, Singh S, Meng LX, Franke AA. Dietary soy intake and urinary isoflavone excretion among women from a multiethnic population. Cancer Epidemiology Biomarkers & Prevention 1998;7(7):613-9.
209. Menni C, Zhai GJ, MacGregor A, Prehn C, Romisch-Margl W, Suhre K, Adamski J, Cassidy A, Illig T, Spector TD, et al. Targeted metabolomics profiles are strongly correlated with nutritional patterns in women. Metabolomics 2013;9(2):506-14. doi: 10.1007/s11306-012-0469-6.
210. McCullough ML, Maliniak ML, Stevens VL, Carter BD, Hodge RA, Wang Y. Metabolomic markers of healthy dietary patterns in US postmenopausal women. Am J Clin Nutr 2019;109(5):1439-51. doi: 10.1093/ajcn/nqy385.
211. Noh H, Freisling H, Assi N, Zamora-Ros R, Achaintre D, Affret A, Mancini F, Boutron-Ruault M-C, Flogel A, Boeing H, et al. Identification of Urinary Polyphenol Metabolite Patterns Associated with Polyphenol-Rich Food Intake in Adults from Four European Countries. Nutrients 2017;9(8):796. doi: https://dx.doi.org/10.3390/nu9080796.
212. O'Gorman A, Morris C, Ryan M, O'Grada CM, Roche HM, Gibney ER, Gibney MJ, Brennan L. Habitual dietary intake impacts on the lipidomic profile. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences 2014;966:140-6. doi: 10.1016/j.jchromb.2014.01.032.
213. Pallister T, Jennings A, Mohney RP, Yarand D, Mangino M, Cassidy A, MacGregor A, Spector TD, Menni C. Characterizing Blood Metabolomics Profiles Associated with Self-Reported Food Intakes in Female Twins. Plos One 2016;11(6). doi: 10.1371/journal.pone.0158568.
214. Pallister T, Haller T, Thorand B, Altmaier E, Cassidy A, Martin T, Jennings A, Mohney R, Gieger C, MacGregor A, et al. Metabolites of milk intake: a metabolomic approach in UK twins with findings replicated in two European cohorts. European Journal of Nutrition 2017;56(7):2379-91. doi: 10.1007/s00394-016-1278-x.
215. Papandreou C, Hernández-Alonso P, Bulló M, Ruiz-Canela M, Yu E, Guasch-Ferré M, Toledo E, Dennis C, Deik A, Clish C, et al. Plasma Metabolites Associated with Coffee Consumption: A Metabolomic Approach within the PREDIMED Study. Nutrients 2019;11(5):1032-. doi: 10.3390/nu11051032.
216. Parmenter B, Moore K, Angelino D, Del Rio D, McNulty H, Molloy AM, Cunningham C, Ward M, Pucci B, Jarrett H, et al. Validity of plasma phenyl-gamma-valerolactones as novel biomarkers of dietary (poly)phenols: Preliminary analysis from the VALID project. Proceedings of the Nutrition Society Conference: Irish Section Conference 2018;77(OCE3).
217. Perng W, Tang L, Song PXK, Goran M, Rojo MMT, Cantoral A, Peterson KE. Urate and Nonanoate Mark the Relationship between Sugar-Sweetened Beverage

Intake and Blood Pressure in Adolescent Girls: A Metabolomics Analysis in the ELEMENT Cohort. Metabolites 2019;9(5). doi: 10.3390/metabo9050100.
218. Philibert A, Vanier C, Abdelouahab N, Chan HM, Mergler D. Fish intake and serum fatty acid profiles from freshwater fish. Am J Clin Nutr 2006;84(6):1299-307.
219. Playdon MC, Sampson JN, Cross AJ, Sinha R, Guertin KA, Moy KA, Rothman N, Irwin ML, Mayne ST, Stolzenberg-Solomon R, et al. Comparing metabolite profiles of habitual diet in serum and urine. Am J Clin Nutr 2016;104(3):776-89. doi: 10.3945/ajcn.116.135301.
220. Playdon MC, Moore SC, Derkach A, Reedy J, Subar AF, Sampson JN, Albanes D, Fangyi G, Kontto J, Lassale C, et al. Identifying biomarkers of dietary patterns by using metabolomics. Am J Clin Nutr 2017;105(2):450-65. doi: 10.3945/ajcn.116.144501.
221. Rabassa M, Zamora-Ros R, Palau-Rodriguez M, Tulipani S, Minarro A, Bandinelli S, Ferrucci L, Cherubini A, Andres-Lacueva C. Habitual Nut Exposure, Assessed by Dietary and Multiple Urinary Metabolomic Markers, and Cognitive Decline in Older Adults: The InCHIANTI Study. Molecular nutrition & food research 2020;64(2):e1900532. doi: http://dx.doi.org/10.1002/mnfr.201900532.
222. Reeves JL, Otahal P, Magnussen CG, Dwyer T, Kangas AJ, Soininen P, Ala-Korpela M, Venn AJ, Smith KJ. DHA mediates the protective effect of fish consumption on new episodes of depression among women. British Journal of Nutrition 2017;118(9):743-9. doi: 10.1017/s0007114517002768.
223. Rios-Leyvraz M, Bochud M, Tabin R, Genin B, Russo M, Rossier MF, Eap CB, Bovet P, Chiolero A. Monitoring caffeine intake in children with a questionnaire and urine collection: a cross-sectional study in a convenience sample in Switzerland. European Journal of Nutrition 2020;59(8):3537-43. doi: 10.1007/s00394-020-02187-3.
224. Rothwell JA, Fillatre Y, Martin J-F, Lyan B, Pujos-Guillot E, Fezeu L, Hercberg S, Comte B, Galan P, Touvier M, et al. New biomarkers of coffee consumption identified by the non-targeted metabolomic profiling of cohort study subjects. PloS one 2014;9(4):e93474-e. doi: https://dx.doi.org/10.1371/journal.pone.0093474.
225. Rothwell JA, Keski-Rahkonen P, Robinot N, Assi N, Casagrande C, Jenab M, Ferrari P, Boutron-Ruault M-C, Mahamat-Saleh Y, Mancini FR, et al. A Metabolomic Study of Biomarkers of Habitual Coffee Intake in Four European Countries. Molecular nutrition & food research 2019;63(22):e1900659. doi: https://dx.doi.org/10.1002/mnfr.201900659.
226. Rybak ME, Sternberg MR, Pao C-I, Ahluwalia N, Pfeiffer CM. Urine excretion of caffeine and select caffeine metabolites is common in the U.S. population and associated with caffeine intake. Journal of Nutrition 2015;145(4):766-74. doi: 10.3945/jn.114.205476.
227. Schmidt JA, Rinaldi S, Ferrari P, Carayol M, Achaintre D, Scalbert A, Cross AJ, Gunter MJ, Fensom GK, Appleby PN, et al. Metabolic profiles of male meat

eaters, fish eaters, vegetarians, and vegans from the EPIC-Oxford cohort. Am J Clin Nutr 2015;102(6):1518-26. doi: 10.3945/ajcn.115.111989.
228. Schmidt JA, Rinaldi S, Scalbert A, Ferrari P, Achaintre D, Gunter MJ, Appleby PN, Key TJ, Travis RC. Plasma concentrations and intakes of amino acids in male meat-eaters, fish-eaters, vegetarians and vegans: a cross-sectional analysis in the EPIC-Oxford cohort. Eur J Clin Nutr 2016;70(3):306-12. doi: 10.1038/ejcn.2015.144.
229. Seow A, Shi CY, Franke AA, Hankin JH, Lee HP, Yu MC. Isoflavonoid levels in spot urine are associated with frequency of dietary soy intake in a population-based sample of middle-aged and older Chinese in Singapore. Cancer Epidemiology Biomarkers & Prevention 1998;7(2):135-40.
230. Seow WJ, Yanwen DL, Pan WC, Gunther SH, Sim X, Torta F, Herr DR, Kovalik JP, Jianhong C, Khoo CM, et al. Coffee, Black Tea, and Green Tea Consumption in Relation to Plasma Metabolites in an Asian Population. Molecular nutrition & food research 2020:e2000527. doi: http://dx.doi.org/10.1002/mnfr.202000527.
231. Shiokawa Y, Date Y, Kikuchi J. Application of kernel principal component analysis and computational machine learning to exploration of metabolites strongly associated with diet. Scientific Reports 2018;8:3426. doi: 10.1038/s41598-018-20121-w.
232. Szeto YT, Kwok TCY, Benzie IFF. Effects of a long-term vegetarian diet on biomarkers of antioxidant status and cardiovascular disease risk. Nutrition 2004;20(10):863-6.
233. Thiébaut ACM, Rotival M, Gauthier E, Lenoir GM, Boutron-Ruault M, Joulin V, Clavel-Chapelon F, Chajès V. Correlation between serum phospholipid fatty acids and dietary intakes assessed a few years earlier. Nutrition & Cancer 2009;61(4):500-9. doi: 10.1080/01635580802710717.
234. Toffano RBD, Hillesheim E, Mathias MG, Coelho-Landell CA, Salomao RG, Almada MORV, Camarneiro JM, Barros TT, Camelo-Junior JS, Rezzi S, et al. Validation of the brazilian healthy eating index-revised using biomarkers in children and adolescents. Nutrients 2018;10 (2) (no pagination)(154).
235. Tong TYN, Koulman A, Griffin JL, Wareham NJ, Forouhi NG, Imamura F. A Combination of Metabolites Predicts Adherence to the Mediterranean Diet Pattern and Its Associations with Insulin Sensitivity and Lipid Homeostasis in the General Population: The Fenland Study, United Kingdom. Journal of Nutrition 2020;150(3):568-78. doi: 10.1093/jn/nxz263.
236. van Roekel EH, Trijsburg L, Assi N, Carayol M, Achaintre D, Murphy N, Rinaldi S, Schmidt JA, Stepien M, Kaaks R, et al. Circulating Metabolites Associated with Alcohol Intake in the European Prospective Investigation into Cancer and Nutrition Cohort. Nutrients 2018;10(5):654. doi: 10.3390/nu10050654.
237. Vazquez-Fresno R, Llorach R, Urpi-Sarda M, Khymenets O, Bullo M, Corella D, Fito M, Martinez-Gonzalez M, Estruch R, Andres-Lacueva C. An NMR metabolomics approach reveals a combined-biomarkers model in a wine interventional trial with validation in free-living individuals of the PREDIMED study. Metabolomics 2015;11(4):797-806. doi: 10.1007/s11306-014-0735-x.

238. Walker ME, Song RJ, Xu X, Gerszten RE, Ngo D, Clish CB, Corlin L, Ma J, Xanthakis V, Jacques PF, et al. Proteomic and metabolomic correlates of healthy dietary patterns: The framingham heart study. Nutrients 2020;12(5):1476. doi: http://dx.doi.org/10.3390/nu12051476.
239. Wanders AJ, Alssema M, De Hoon SEM, Feskens EJM, van Woudenbergh GJ, van der Kallen CJ, Zock PL, Refsum H, Drevon CA, Elshorbagy A, et al. Circulating Polyunsaturated Fatty Acids as Biomarkers for Dietary Intake across Subgroups: The CODAM and Hoorn Studies. Annals of Nutrition & Metabolism 2018;72(2):117-25. doi: 10.1159/000486244.
240. Wang Y, Gapstur SM, Carter BD, Hartman TJ, Stevens VL, Gaudet MM, McCullough ML. Untargeted Metabolomics Identifies Novel Potential Biomarkers of Habitual Food Intake in a Cross-Sectional Study of Postmenopausal Women. Journal of Nutrition 2018;148(6):932-43. doi: 10.1093/jn/nxy027.
241. Yeung EH, Saudek CD, Jahren AH, Kao WHL, Islas M, Kraft R, Coresh J, Anderson CAM. Evaluation of a Novel Isotope Biomarker for Dietary Consumption of Sweets. Am J Epidemiol 2010;172(9):1045-52. doi: aje/kwq247.
242. Zheng Y, Yu B, Alexander D, Steffen LM, Boerwinkle E. Human Metabolome Associates With Dietary Intake Habits Among African Americans in the Atherosclerosis Risk in Communities Study. Am J Epidemiol 2014;179(12):1424-33. doi: aje/kwu073.
243. Zheng Y, Yu B, Alexander D, Steffen LM, Nettleton JA, Boerwinkle E. Metabolomic patterns and alcohol consumption in African Americans in the Atherosclerosis Risk in Communities Study. Am J Clin Nutr 2014;99(6):1470-8.
244. Zong G, Sun Q, Yu D, Zhu J, Sun L, Ye X, Li H, Jin Q, Zheng H, Hu FB, et al. Dairy consumption, type 2 diabetes, and changes in cardiometabolic traits: a prospective cohort study of middle-aged and older chinese in beijing and shanghai. Diabetes Care 2014;37(1):56-63. doi: 10.2337/dc13-0975.