UNIVERSITÀ DI PARMA

# UNIVERSITÀ DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
"FOOD SCIENCE"

CICLO XXXIV

# *Metabolomics applications to nutritional intervention studies*

Coordinatore:
Chiar.ma Prof.ssa **Chiara Dall'Asta**

Tutore:
Chiar.mo Prof. **Daniele Del Rio**

Dottoranda: **Claudia Favari**

Anni Accademici 2018/2019 – 2020/2021

*“Albert leggeva Kant e seguiva a tempo perso lezioni all’Università di Pavia: per divertimento, senza essere iscritto né fare esami. È così che si diventa scienziati sul serio.”*

Carlo Rovelli, “Sette brevi lezioni di fisica”

# Table of contents

**Figures** ..... **iv**

**Tables** ..... **vii**

**Abbreviations** ..... **viii**

**Abstract** ..... **xi**

**Chapter 1 – Introduction** ..... **1**

*1.1. The role of plant bioactives in human health* ..... *2*

*1.2. The phenolic compounds: bioavailability and health effects* ..... *4*

1.2.1. General information ..... 4

Chemistry and main dietary sources of phenolic compounds ..... 4

Absorption, distribution, metabolism, and excretion of phenolic compounds ..... 9

The effects on health of phenolic compounds and their derived circulating metabolites ..... 11

1.2.2. Focus on flavan-3-ols and their derived colonic metabolites ..... 13

Chemistry and main dietary sources of flavan-3-ols ..... 13

Bioavailability of flavan-3-ols ..... 15

Health benefits of flavan-3-ols and their derived circulating metabolites ..... 24

*1.3. Metabolomics as a tool in nutrition research* ..... *31*

1.3.1. Evaluation of dietary compounds bioavailability ..... 32

1.3.2. Metabotyping ..... 33

1.3.3. Study of the metabolic changes following dietary interventions ..... 34

*References* ..... *37*

**Chapter 2 – Aims of the thesis** ..... **57**

**Chapter 3 – Selected studies** ..... **59**

*3.1. Study 1* ..... *60*

Abstract ..... 61

3.1.1. Introduction ..... 62

3.1.2. Material and Methods ..... 63

Materials ..... 63

Study Design ..... 63

UHPLC-ESI-QqQ-MS/MS analysis ..... 63

Data analysis ..... 64

3.1.3. Results ..... 64
Plasma kinetics ..... 64
Urinary excretion ..... 72
Plasma and urinary dose-response ..... 74
Inter-individual variability ..... 75
3.1.4. Discussion ..... 77
3.1.5. Conclusion ..... 79
Electronic Supplementary Information (ESI) ..... 80
References ..... 84

*3.2. Study 2* ..... *88*
Abstract ..... 89
3.2.1. Introduction ..... 90
3.2.2. Material and Methods ..... 91
Intervention studies ..... 91
Sample analysis ..... 92
Unsupervised analyses ..... 93
Supervised analysis: Partial least square-discriminant analysis ..... 94
Univariate statistics ..... 95
3.2.3. Results ..... 95
Effect of data pre-treatment on resulting PCA models ..... 95
Cluster definition ..... 99
PLS-DA models to explore differences between clustering methods and the biological relevance of each group ..... 101
Univariate statistics confirm the differences in metabolite excretion between groups ..... 104
Sulfate/glucuronide ratio as proxy of individual variability in phase II metabolism ..... 105
3.2.4 Discussion ..... 110
3.2.5. Conclusion ..... 113
Electronic Supplementary Information (ESI) ..... 116
References ..... 128

*3.3. Study 3* ..... *136*
Abstract ..... 138
3.3.1. Introduction ..... 139
3.3.2. Materials and methods ..... 140
Chemicals ..... 140
Subjects ..... 140
Dosage information and study design ..... 141
Urine sample preparation ..... 141
UHPLC-TWIMS-QTOF analysis ..... 142

Data processing and multivariate modelling ..... 142
Metabolic pathway analysis ..... 143
3.3.3. Results and Discussion ..... 144
Multivariate modelling and metabolite identification ..... 144
Metabolism of coffee and cocoa phytochemicals (biomarkers of intake) ..... 145
Biological interpretation of significant markers ..... 153
3.3.4. Concluding remarks ..... 158
Electronic Supplementary Information (ESI) ..... 160
References ..... 167

**Chapter 4 – Overall discussion and conclusions ..... 175**

*4.1. General remarks ..... 176*

*4.2. Analytical considerations ..... 179*

*References ..... 181*

**Acknowledgements ..... 185**

**About the author ..... 187**

*Curriculum Vitae ..... 187*
Scientific activities ..... 187
List of publications ..... 187
Oral communications at National and International congresses ..... 188
Poster communications at National and International congresses ..... 189

# Figures

Figure 1. Graphical abstract of the doctoral thesis..... xiii
Figure 2. Main dietary phytochemicals found in plant-based foods (Modified from Fraga et al. [10])..3
Figure 3. The basic skeleton of flavonoid. .....4
Figure 4. Stereoisomers of monomeric flavan-3-ols with different patterns of hydroxylation. .....14
Figure 5. Dimeric B-type and A-type proanthocyanidins.....15
Figure 6. Primary green and black tea flavan-3-ol derivatives. .....16
Figure 7. Structures of leading native and colonic flavan-3-ol metabolites. .....19
Figure 8. Exemplified catabolism of procyanidin B2 and (−)-epicatechin by gut microbiota. Pathways proposed, according to Appeldoorn et al. [78] and Stoupi et al. [79] .....20
Figure 9. Applications of metabolomics in the field of nutrition.....32
Figure 10. Overview of the concept of metabotyping. Individuals are grouped into metabotypes based on their metabolic phenotypes and dietary advice is delivered tailored to the metabotype. .....33
Figure 11. Kinetic profile of 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide (a), 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-glucuronide (b), 5-(hydroxyphenyl)-γ-valerolactone-sulfate (3', 4') (c) and of the sum of quantified PVLs (d) for the different treatments. Data are expressed as nM, mean ± SEM (*n* = 10). Treatment 0 TF corresponds to the control drink, other treatments to cranberry juice drinks with growing amounts of total flavan-3-ols. .....71
Figure 12. (a) Cumulative urinary excretion of PVLs at 24 h after treatment with higher flavan-3-ols content (1741 mg TF). Data are expressed as µmol. (b) Relative contribution of each PVL to the total urinary excretion at 24 h after treatment with higher flavan-3-ols content (1741 mg TF). (c) Relative contribution of derivatives of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone compared to derivatives of 5-(3'-hydroxyphenyl)-γ-valerolactone and 5-(4'-hydroxyphenyl)-γ-valerolactone to the total urinary excretion at 24 h after treatment with higher flavan-3-ols content (1741 mg TF). (d) Relative contribution of glucuronide and sulfate conjugates to the total urinary excretion at 24 h after treatment with higher flavan-3-ols content (1741 mg TF). .....74
Figure 13. Relation between total amount of ingested flavan-3-ols and Area under the curve ($AUC_{0-24}$) of the concentration in plasma of PVLs over 24h. Data are expressed as nM*h, mean ± SEM, *n* = 10.....75

Figure 14. Relation between total amount of ingested flavan-3-ols and cumulative quantity of excreted metabolites. Data are expressed as µmol, mean ± SEM, *n* = 10. .....76
Figure 15. (a) Inter-individual variability in circulating concentrations of the sum of quantified PVLs after consumption of cranberry juice containing increasing amount of total flavan-3-ols. Data are expressed as Area under the curve (AUC) of the concentration in plasma over 24h in nM*h. (b) Inter-individual variability in urinary excretion of the sum of quantified PVLs after consumption of cranberry juice containing increasing amount of total flavan-3-ols. Data are expressed as cumulative quantity of excreted metabolites in µmol. .....77
Figure 16. Score (A, C, E) and loading (B, D, F) plots resulting after PCA analysis on non-transformed, centered data for individual metabolites (A, B), non-transformed, centered and unit variance scaled data for individual metabolites (C, D), non-transformed, centered data for sums of metabolites belonging to the same aglycone compound (E, F).....97
Figure 17. Two-classes cluster plot resulted from the application of different clustering methods on the datasets with individual metabolites (A) and with sums of metabolites belonging to the same aglycone compound (B). .....100
Figure 18. PLS-DA models (score and loading plots) considering individual metabolites and the clusters obtained from different clustering methods: (A) final consensus – FC –, (B) k-means – Kmeans –, (C) expectation-maximization – EM – and PC score-based models for 2 (D) or 3 (E) groups. .....104
Figure 19. Mean urinary excretion (µmol) over 24 h of sums of metabolites belonging to the same aglycone compound (3'OH-PVLs, sum of conjugates from the aglycone 5-(3'-hydroxyphenyl)-γ-valerolactone; 4'OH-PVLs, 5-(4'-hydroxyphenyl)-γ-valerolactone; 3',4'diOH-PVLs, 5-(3',4'-dihydroxyphenyl)-γ-valerolactone; HPPs, 3-(hydroxyphenyl)propanoic acid), calculated both before and after cluster analysis ("Individual" and "Sums", respectively). Clustering has been performed on the basis of: Final Consensus (first column), k-means (second column), expectation-maximization algorithm (third column), PC score forming 2 groups (fourth column), PC score forming 3 groups (fifth column). Different letters indicate statistically significant differences ($p < 0.05$) among groups 1 and 2 or 1,2 and 3. .....109
Figure 20. Inter-individual variability in phase II metabolism illustrated by the sulfate (SULF)/glucuronide (GLUC) ratio of the sums of respective conjugated metabolites (A) and of 5-(3',4'-dihydroxyphenyl)-γ-valerolactones (dOH-PVL) (B) in urine samples. (C) Relationship between the

sulfate/glucuronide ratio of all the metabolites and the sulfate/glucuronide ratio of 5-(3',4'-dihydroxyphenyl)-γ-valerolactones. .....110
Figure 21. (A) PLS-DA model built with positive ionization data ($R^2Y$ = 0.95, $Q^2$ = 0.651) and (B) negative ionization data ($R^2Y$ = 0.906, $Q^2$ = 0.769). The intervention treatment groups (1C, 3C, PC) are color coded accordingly. .....144
Figure 22. Pathway analysis performed with all the significant metabolites identified. The output displays metabolic pathways arranged by scores from pathway enrichment (y-axis) and topology analysis (x-axis). The color and size of each circle are based on p-values and pathway impact values, respectively (from yellow to red, the –log(*p*) increases, the bigger the circle size, the higher the pathway impact value). .....153
Figure 23. Detailed metabolic pathways involved and the significant metabolites (endogenous and exogenous) identified in urine after coffee or coffee and cocoa intake and their inter-connections. Information on pathways has been drawn from KEGG database. * indicates level I identified metabolites. Phe: phenylalanine; Tyr: tyrosine; Trp: tryptophan; Arg: arginine; Pro: proline; 1-Mx: 1-methylxanthine, AFMU: 5-Acetylamino-6-formylamino-3-methyluracil; AAMU: 5-Acetylamino-6-amino-3-methyluracil; TCA cycle: tricarboxylic acid cycle/citrate cycle. .....155

# Tables

Table 1. Basic structures, main representative compounds and dietary sources of the most important subclasses of phenolic compounds. ..... 6

Table 2. Examples of dietary sources of flavan-3-ols and content of their main flavan-3-ols (PhenolExplorer [76]). ..... 16

Table 3. Mass spectral characteristic of PVLs and PVAs identified in plasma and urine ..... 64

Table 4. Kinetic parameters of PVLs detected in plasma after cranberry juice consumption (mean ± SEM, *n* = 10). The coefficient of variation (CV%) of $C_{max}$ is reported in brackets. ..... 66

Table 5. Cumulative excretion of PVLs in urine after 24h following cranberry juice consumption. Data are expressed in µmol (mean ± SEM, *n* = 10). Coefficient of variation (CV%) is reported in brackets. ..... 72

Table 6. Statistics of computed PCA models illustrating the effect of transformation and scaling methods on the datasets considering individual metabolites and sums of metabolites belonging to the same aglycone compound. ..... 96

Table 7. Statistics of computed PLS-DA models, considering the class resulted from different clustering methods and individual metabolites or sums of metabolites belonging to the same aglycone compound. ..... 102

Table 8. Urinary excretion of individual metabolites and sums of metabolites sulfate or glucuronide per each cluster defined after applying different clustering methods (final consensus -FC-, expectation-maximization -EM-, k-means -Kmeans-, and PC score-based models for 2 or 3 groups) on the dataset with individual metabolites. Data are expressed in µmol (mean ± SD). Mean differences between groups were considered significant for $p < 0.05$ and significant *p*-values are highlighted in bold. N indicates the number of observations per cluster. ..... 106

Table 9. Statistically significant urine metabolites annotated using UHPLC-TWIMS-QTOF. Metabolite levels that increase in response to the first considered treatment are shaded red and metabolite levels that decrease are colored green. The fold change value is calculated as the ratio between the highest mean concentration reported in a specific treatment (1C, 3C or PC) for a metabolite and the lowest mean concentration reported for the same metabolite in the respective treatment. A red-shaded value of 1.0 for the mentioned ratio means that the real value was, for example, 1.023 rounded to 1.0, indicating an increase, although small. ..... 146

# Abbreviations

| | |
|---|---|
| **[2-$^{14}$C]EC** | [2-$^{14}$C](–)-Epicatechin |
| **ADD** | Average danish diet |
| **ADME** | Absorption, distribution, metabolism and excretion |
| **AGAT** | Arginine:glycine amidinotransferase |
| **$AT_{1/2}$** | Apparent elimination half-life |
| **ATP** | Adenosine triphosphate |
| **AUC** | Area under the curve |
| **CBG** | Cytosolic β-glucosidase |
| **CBPCC** | Cocoa-based products containing coffee |
| **CCS** | Collision cross section |
| **$C_{max}$** | Maximum concentration |
| **COMT** | Catechol-*O*-methyltransferases |
| **COST-POSITIVe** | European Cooperation in Science and Technology-Plant food bioactives interindividual variation |
| **CVDs** | Cardiovascular diseases |
| **CV-ANOVA** | Cross validation-one way analysis of variance |
| **CV%** | Coefficient of variation |
| **DIA** | Data-independent approach |
| **EGCG** | (–)-Epigallocatechin-gallate |
| **EM** | Expectation-maximization |
| **ESI** | Electronic supplementary information |
| **FC** | Final consensus |
| **FDR** | False discovery rate |
| **FMD** | Flow-mediated dilatation |
| **FoodBAll** | Food Biomarkers Alliance |
| **GC-MS** | Gas chromatography-mass spectrometry |
| **GI tract** | Gastro-intestinal tract |
| **HCPC** | Hierarchical clustering on principal components |
| **HDL-C** | High-density lipoprotein cholesterol |
| **HMDB** | Human Metabolome Database |
| **HPLC-RC** | High performance liquid chromatography-radiocounting |
| **HPPs** | 3-(Hydroxyphenyl)propanoic acids |
| **H-fuzzy** | Hierarchical fuzzy |
| **H-Kmeans** | Hierarchical k-means |
| **H-PAM** | Hierarchical partition around medoids |
| **Kmeans** | k-means |
| **LC-MS** | Liquid chromatography-mass spectrometry |
| **LDL-C** | Low-density lipoprotein cholesterol |
| **LPH** | Lactase phlorizin hydrolase |
| **NADPH** | Nicotinamide adenine dinucleotide phosphate |
| **NIH** | National Institutes of Health |
| **NCDs** | Non-communicable diseases |
| **NMR** | Nuclear magnetic resonance spectroscopy |
| **NND** | New nordic diet |
| **PACs** | Proanthocyanidins |

| | |
|---|---|
| **PAM** | Partition around medoids |
| **PC** | Principal components |
| **PCA** | Principal component analysis |
| **PLS-DA** | Partial least-squares discriminant analysis |
| **PREDIMED** | Prevención con Dieta Mediterránea |
| **PVAs** | Phenylvaleric acids |
| **PVLs** | Phenyl-γ-valerolactones |
| **QC** | Quality control |
| **RCTs** | Randomized controlled trials |
| **SD** | Standard deviation |
| **SEM** | Standard error of the mean |
| **SGLT1** | Sodium-glucose transporter 1 |
| **SREMs** | Structurally-related (–)-epicatechin metabolites |
| **SRM** | Selective reaction monitoring |
| **SULT** | Sulfotransferases |
| **TF** | Total flavan-3-ols |
| **$T_{max}$** | Time to reach $C_{max}$ |
| **UGTs** | Uridine-5′-diphosphate glucuronosyltransferases |
| **UHPLC-ESI-QqQ-MS/MS** | Ultra high-performance liquid chromatography coupled with triple quadrupole mass spectrometry, equipped with an electrospray ionization source |
| **UHPLC-IMS-HRMS** | Ultra high performance liquid chromatography- ion mobility spectrometry-high resolution mass spectrometry |
| **UHPLC-TWIMS-QTOF** | Ultra high performance liquid chromatography-travelling wave ion mobility spectrometry-quadrupole time-of-flight |
| **UPEC** | Uropathogenic *Escherichia coli* |
| **UPLC-MS** | Ultra performance liquid chromatography-mass spectrometry |
| **UTIs** | Urinary tract infections |
| **UV** | Unit variance scaling |
| **VIP** | Variable Importance in Projection |
| **WHO** | World Health Organization |

# Abstract

Plant phytochemicals might not be essential throughout life, but they may promote health and well-being by preventing, for example, certain non-communicable diseases. Among phytochemicals, flavan-3-ols are particularly relevant, as they are extensively consumed within diet, being characteristic compounds of tea, cocoa, wine, pome fruits, berries, and nuts. The existing literature shows a clear evidence that flavan-3-ols exert beneficial effects on health, however, discrepancies of results have been highlighted when encouraging results obtained using observational or animal studies are compared with more wary results from *in vivo* human intervention studies. The reason for this inconsistency may lay in the profound structural modifications that these compounds undergo once in contact with the human body, and in the inter-individual variability described in the capacity of our body to transform and absorb them. So, assessment of flavan-3-ol bioavailability and associated inter-individual variability are crucial factors to understand to what extent and in which forms they are available for the individual internal compartments, a fundamental information to unravel the biological effects of these compounds in human health. This information is then completely lacking for cranberry flavan-3-ols.

Among dietary sources rich in phytochemicals, coffee is particularly relevant as it is one of the most consumed drinks worldwide. In many epidemiological studies, regular coffee consumption has been associated with a reduced risk of several non-communicable diseases, but no association has yet been found between circulating and excreted coffee bioactives and derived metabolites and physiological responses, making the mechanisms through which coffee exerts its potential preventive effects still widely undisclosed. In this frame, metabolomics may aid in achieving insights into such open questions, thanks to its several applications in the field of nutrition.

The global objective of this doctoral thesis was to apply metabolomics to nutritional intervention studies to pursue several aims. Firstly, to assess the bioavailability of flavan-3-ols from cranberry through the analysis of their main colonic metabolites (phenyl-γ-valerolactones and phenyl-valeric acids), also considering their potential as biomarkers of flavan-3-ol intake and inter-individual variability in their appearance in plasma and urine. Secondly, to elucidate the presence of different metabotypes (metabolic phenotypes) in the urinary excretion of flavan-3-ol colonic metabolites as a strategy to manage the associated inter-individual variability, also assessing the impact of the statistical technique used in the process. Lastly, to study the changes occurring at metabolomic level

following consumption of different patterns of coffee and the impact on known metabolic pathways.

Plasma and/or urine samples collected in previously conducted human intervention trials were analysed trough LC-MS based metabolomics approaches, either targeted or untargeted. Specifically, a targeted analysis of all the potential conjugated forms of phenyl-γ-valerolactones and phenylvaleric acid revealed that sulfate and glucuronide conjugates of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone are the main circulating and excreted metabolites after cranberry flavan-3-ols intake, reaching maximum plasma concentration at about 4–6 h post consumption. These metabolites show a clear dose-response relationship with the amount of cranberry flavan-3-ols consumed, supporting their suitability as biomarkers of cranberry flavan-3-ol intake. Glucuronidation seems to be the most favored phase II conjugation after cranberry juice intake, despite wide differences among individuals. A high inter-individual variability is also reported in circulating and urinary metabolite levels, with some subjects displaying a greater efficiency in metabolizing flavan-3-ols and producing phenyl-γ-valerolactones.

On the basis of this evidence and based on the results of a preliminary set of experiments, where a different urinary production of phenyl-γ-valerolactones and 3-(hydroxyphenyl)propanoic acids upon green tea flavan-3-ol consumption led to the elucidation of three putative metabotypes, a targeted analysis of these metabolites revealed different profiles in their urinary excretion upon cranberry consumption in two diverse experimental settings. A methodological workflow for metabotype definition and validation, based on broadly accessible multivariate and univariate tools, has been proposed, highlighting the importance of data pre-treatment and clustering methods on the final outcomes when analysing datasets of flavan-3-ol metabolites. Non-transformed, centered, and UV-scaled data has been key to unravel metabolic patterns based on colonic metabolism. Cluster analysis based on k-means and a final consensus algorithm led to quantitative-based models, whereby the distribution of the clusters was due to the amount of metabolites excreted (high vs. low). Differently, the expectation-maximization algorithm and clustering according to principal component scores yield metabotypes characterized by quali-quantitative differences in the excretion of colonic metabolites. PLS-DA, together with univariate analyses, served to validate the urinary metabotypes in the production of flavan-3-ol metabolites and to confirm the robustness of the methodological approach. This metabotyping strategy may be key to manage the inter-individual variability reported in the colonic metabolism of flavan-3-ols and to further investigate its consequences in the impact on the observed health effects attributed to this class of compounds.

Finally, the use of an untargeted approach has allowed the detection of changes in the metabolome due to the intake of different patterns of coffee. Indeed, besides metabolites specifically derived from coffee intake, endogenous metabolites have been detected whose levels were modulated by the different patterns of coffee consumed. New metabolic pathways have therefore been identified to be modulated by the intake of different patterns of coffee, such as the metabolism and biosynthesis of specific amino acids, that may in turn potentially influence human health.

In conclusion, the application of metabolomics approaches to nutrition intervention studies previously conducted has offered the possibility to gain new and further insights in the research field of dietary bioactives, namely the bioavailability of flavan-3-ols and the associated inter-individual variability, the elucidation of a metabotyping strategy to manage it, and the metabolic routes by which coffee and coffee-related metabolites may exert effects on human health, opening the door to new hypotheses about the health effects of coffee consumption and the underlying mechanisms.

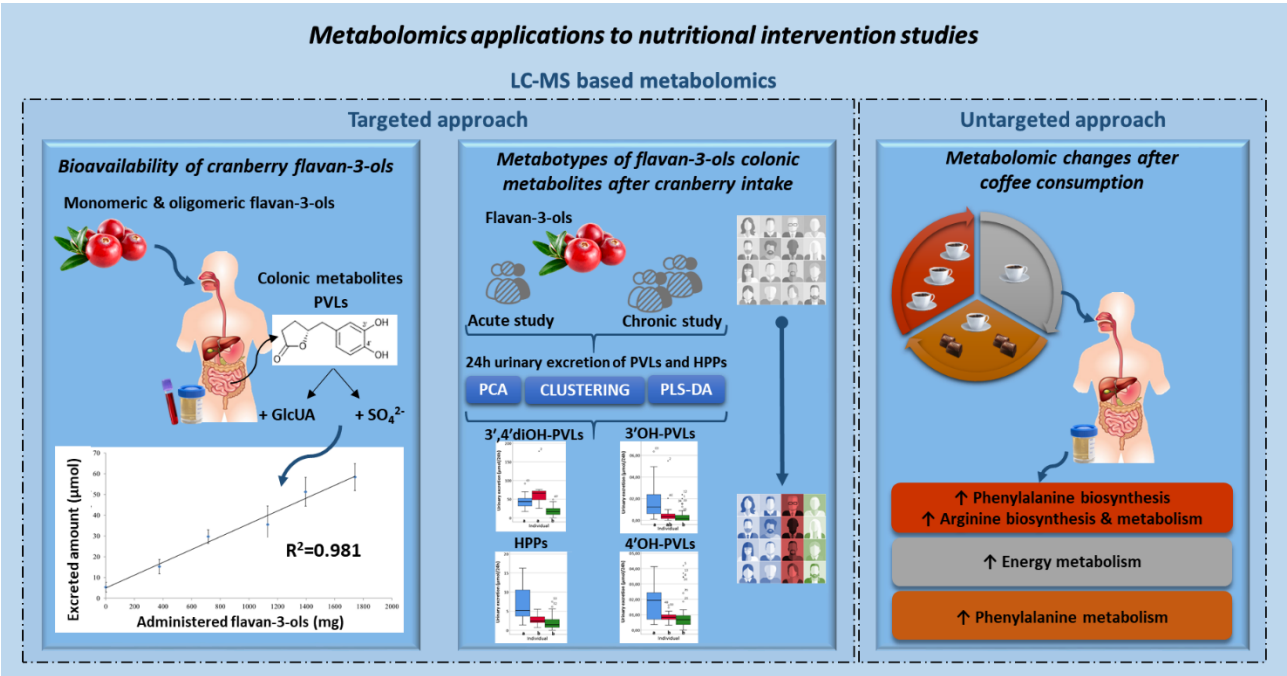



**Figure 1.** Graphical abstract of the doctoral thesis.

# Chapter 1 – Introduction

# 1.1. The role of plant bioactives in human health

The World Health Organization (WHO) has estimated that in 2016 non-communicable diseases (NCDs) collectively caused 41 million deaths worldwide, equivalent to 71% of all global deaths [1]. Non-communicable diseases are non-infectious health conditions, which usually last for a long period of time and therefore are also known as chronic diseases [2]. These diseases are caused by a combination of genetic, physiological, environmental and behavioural factors, and are considered a major public health concern in developing as well as developed regions of the world. The major NCDs are cardiovascular diseases (CVDs), cancer, diabetes and chronic respiratory diseases. At a global level, CVDs cause more premature deaths than cancer [1]. Unhealthy diet is considered one of the most important preventable risk factor for many NCDs [2,3], such as CVDs, type 2 diabetes and certain types of cancer. The WHO and the EAT-Lancet commission on Food, Planet, Health recommend a radical transformation of the global food system, with drastic shifting towards mostly plant-based dietary patterns, demonstrated to be both healthier and more sustainable [2,4,5]. Doubling the global consumption of fruits, vegetables, nuts, whole grains and legumes by 2050 may be key to prevent the onset of certain NCDs (with an estimated reduction of around 19.0-23.6% deaths per year) as well as to fight environmental degradation, benefiting both human and planet health [2,3,6].

From ten years now, an explosion of research is shedding light on the health benefits of plant-based diets and specific plant foods. Plant-based foods have generally a low caloric density and a low glycemic index. In addition, they provide dietary fibre, micronutrients and a wide variety of biologically active phytochemicals (i.e. bioactives) [7]. Phytochemicals are secondary metabolites synthetized by plants primarily as defence against pathogens and stress. They can be specific to a given plant species or widely distributed in plants and their final concentration in food products is affected by industrial and home processing. Dietary intake of phytochemicals often exceeds 1g/day [8]. Phenolic compounds are the most abundant and heterogeneous family of phytochemicals and are widespread in fruits, vegetables and beverages like tea, coffee and wine [9] (Figure 2). Other relevant categories of phytochemicals found in plant-based foods are: alkylresorcinols, carotenoids, phytosterols, alkaloids, glucosinolates and thiosulfinates [10–16] (Figure 2). Plant phytochemicals might not be essential throughout life, but they may promote health and well-being by preventing for example obesity, type 2 diabetes, cardiovascular and other diseases [7,17]. Large intervention studies, as the PREDIMED (Prevención con Dieta Mediterránea) trial, have been instrumental in

demonstrating the health benefits of switching to plant food-rich dietary patterns, including the Mediterranean diet [18] and the New Nordic diet [19]. In parallel, a number of randomized controlled trials (RCTs) and meta-analyses have provided good evidence of improvement by particular plant foods or phytochemicals of risk factors and pathways relevant to weight control and NCDs, such as glucose and lipid metabolism, insulin resistance, inflammation, blood pressure and endothelial dysfunction [20–25]. Some plant foods have shown substantially larger health benefits than others for specific outcomes, likely due to their phytochemical composition that is unique to each species [26].

However, one of the main hindrances to unravel and exploit the potential of plant bioactives for the prevention of NCDs is the heterogeneity in the metabolism and individual response to their consumption, which may lead to inconclusive results. Evidence indicates that some individuals may benefit from the health effects of these phytochemicals more than others, being age, sex, ethnicity, and health status some of the factors behind the heterogeneity in the responsiveness to plant bioactives [27–29]. The gut microbiome, the occurrence of other dietary components able to shape gut

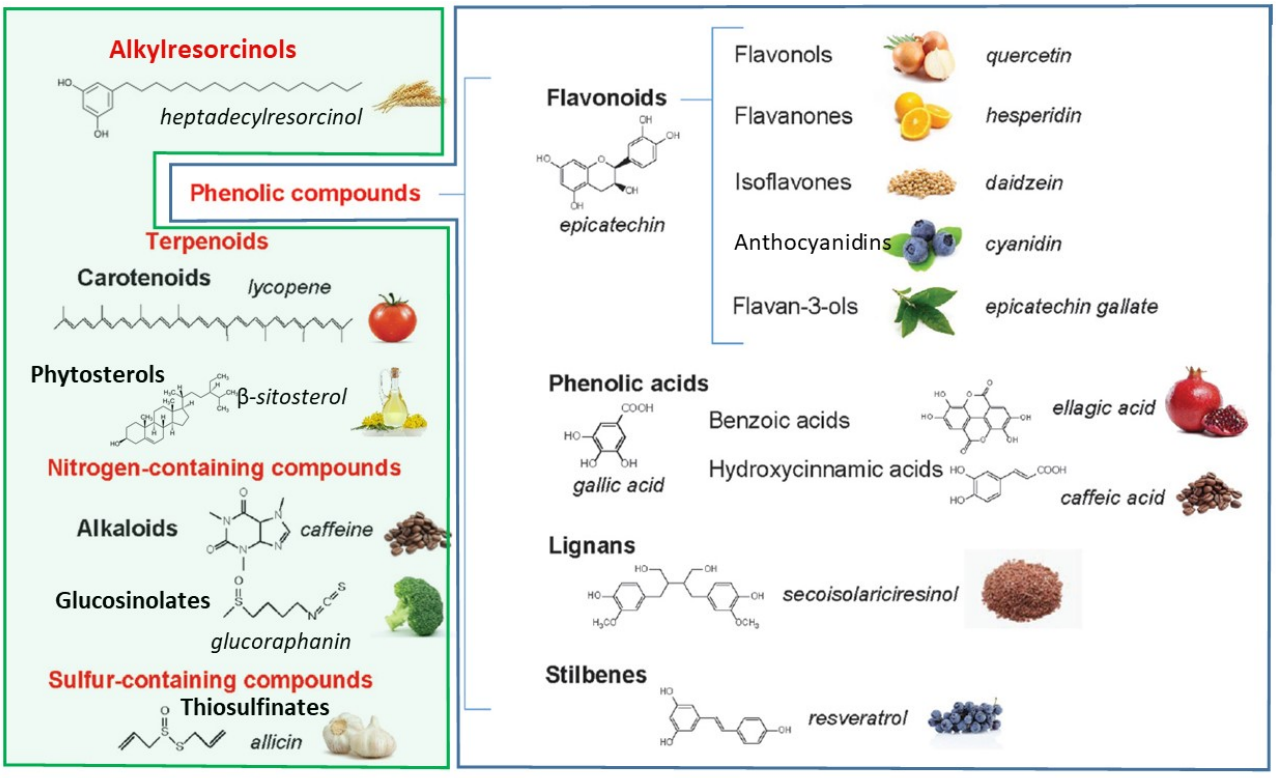



**Figure 2.** Main dietary phytochemicals found in plant-based foods (Modified from Fraga et al. [10]).

microbiota composition and functionality and to be transformed by the gut microbiota into metabolites able to interact with the host physiology, and the adaptive and innate immunity related to lymphocytes and monocytes may also condition the metabolism and individual response to plant

bioactives [30–35]. However, the contribution of all these factors is poorly understood across major plant phytochemicals classes and health outcomes, even more in the context of plant-based diets where dozens of different families of plant bioactives coexist.

# 1.2. The phenolic compounds: bioavailability and health effects

## 1.2.1. General information

*Chemistry and main dietary sources of phenolic compounds*

Phenolic compounds (or (poly)phenols) are characterized by having at least one aromatic ring with one or more hydroxyl groups attached [36]. Their structures range from simple, low molecular weight, single aromatic-ring compounds, such as phenolic acids, to complex, polymerized molecules of relatively high molecular weight, such as hydrolysable and condensed tannins [37]. Phenolics occurring in plant can be classified into two groups based on the number and arrangement of the carbon atoms in their structure: the flavonoids and the non-flavonoids [36] (Table 1). Flavonoids are composed of two phenol rings (A and B) linked through a three-carbon chain that forms a heterocyclic pyran ring (C) containing one oxygen atom (Figure 3). On the basis of the degree of oxidation, saturation, and hydroxylation of the central pyran ring, flavonoids can be divided into different subclasses. The main subclasses occurring in plant-based foods are: flavonols, flavanones, isoflavones, anthocyanidins, flavan-3-ols and flavones (Table 1). Other minor dietary flavonoids are dihydroflavonols, flavan-3,4-diols, coumarins, chalcones, dihydrochalcones and aurones. Non-flavonoids class includes both compounds with a simpler chemical structure than that of the flavonoids as well as large and complex molecules. The main dietary non-flavonoid phenolics are: phenolic acids (hydroxycinnamic and benzoic acids), the hydrolyzable tannins (ellagitannins and gallotannins, derivatives of ellagic acid and gallic acid respecitvely), stilbenes, coumarins, and lignans (Table 1). In general, both flavonoids and non-flavonoids occur *in planta* in the conjugated form rather than as free compounds, with sugars and organic acids. Sugar residues are linked by β-glycosidic bonds to a hydroxyl group (*O*-glycosides) or a carbon atom of the aromatic ring (*C*-glycosides). The associated sugars can be monosaccharides,

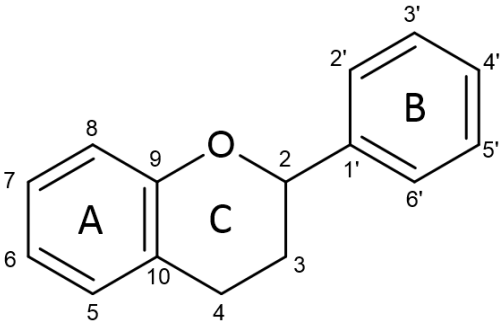



**Figure 3.** The basic skeleton of flavonoid.

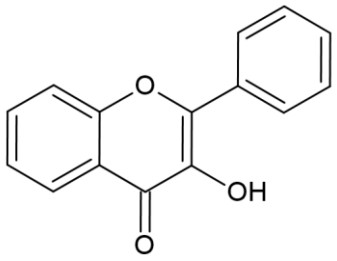
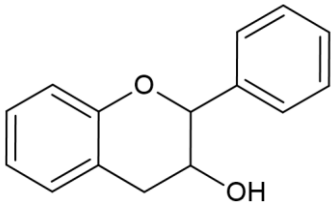
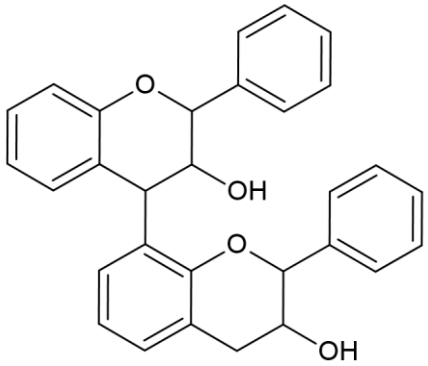
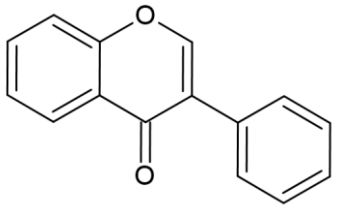
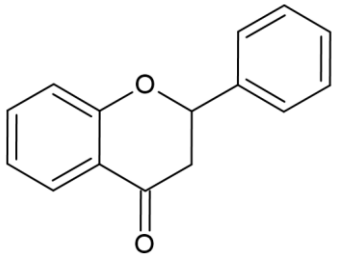
disaccharides or even oligosaccharides, glucose being the most common followed by others such as galactose, rhamnose, xylose, arabinose [36,37].

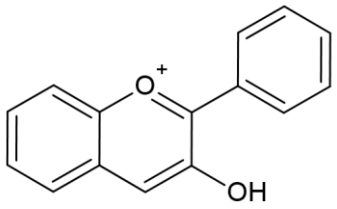
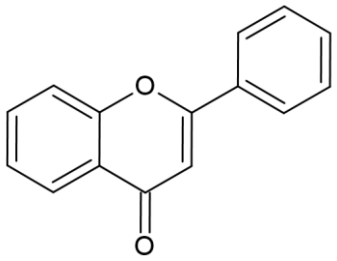
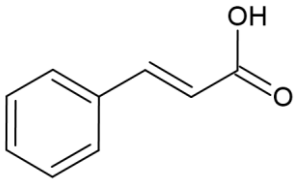
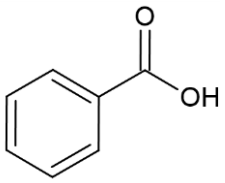
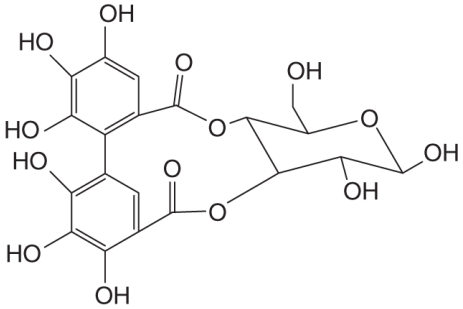
Phenolic compounds are widely distributed throughout the plant kingdom as constituents of almost all vegetables, fruits, cereals, plant-derived beverages such as coffee, tea and red wine and other plant-derived foods [37]. Most plant foods contain complex mixtures of (poly)phenols, however some phenolics are mainly present in specific foods such as flavanones in citrus fruits and isoflavones in legumes. The phenolic content of foods of plant origin can present sizeable differences in seemingly similar products due to varietal differences, local growing conditions and seasonal changes[9]. Flavonoids and phenolic acids are extensively found in most plant foods, while hydrolyzable tannins, stilbenes, coumarins, and lignans are less common and occur in particular foodstuffs [36–38]. Specifically, flavonols are the most ubiquitous flavonoids: they mainly occur in yellow and red onions, capers, lettuce, parsley, carrot, pumpkin, Brassicaceae and Aliaceae, some berries, mango, apples and figs [37]. The most common flavonols are conjugated quercetin, kaempferol, isorhamnetin and myricetin [9]. Flavan-3-ols are the characteristic (poly)phenols of tea, cocoa, red wine, pome fruits (as apple and pear), berries and nuts, but they are also found in stone fruits (as apricot and peach), cereals and legumes [39,40]. This subclass of compounds is the main dietary source of flavonoids in Western diets [8,41,42]. Differently from other flavonoids, flavan-3-ols exist *in planta* predominantly as aglycones rather than glycosides or esterified with gallic acid [36]. Flavan-3-ols can occur as monomers, the most common being (+)-catechin and (–)-epicatechin, or as oligomers and polymers (the so-called proanthocyanidins, PACs, or condensed tannins). Isoflavones are almost exclusively found in legumes, with the highest amount found in the cultivated soybean (*Glycine max* (L.)). The main isoflavones are genistein, daidzein and glycitein [37]. Flavanones occur in high concentrations mainly in citrus fruits and their juices (orange, grapefruit, lemon, lime, etc.) and in smaller amounts in artichokes, tomatoes and certain aromatic plants such as oregano [36,37]. The main flavanone glycosides are hesperidin (hesperetin-7-*O*-rutinoside), naringin (naringenin-7-*O*-neohesperidoside), neohesperidin (hesperetin-7-*O*-neohesperidoside) and eriocitrin (eriodictyol-7-*O*-rutinoside) [37]. Anthocyanins (the glycosidic form of anthocyanidins) are those responsible for the blue, purple and red pigments found in flowers, fruits, leaves, and roots. The most common types of anthocyanidins are cyanidin (responsible for reddish-purple pigment), delphinidin (responsible for blue-reddish or purple pigment), pelargonidin (responsible for red and orange pigment), peonidin (responsible for reddish-purple pigment), malvidin (responsible for purple-blue and red pigment), and petunidin (responsible for dark red or purple pigment). Among coloured fruits, the main dietary sources are

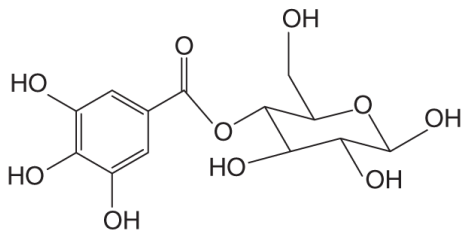
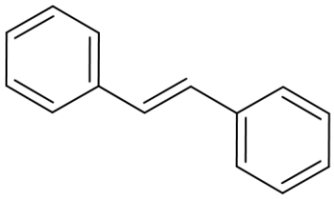
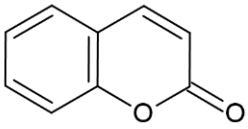
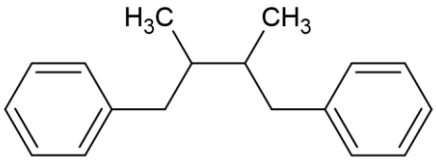
berries, cherries, plums, pomegranates, some tropical fruits and fruit-derived products (red wine, fruit juices and jams). Among dark-coloured vegetables and cereals, anthocyanins are found in red cabbage, eggplant, red onions, radishes, hazelnuts, pistachio nut, black and red beans, as well as certain varieties of herbal medicinal plants including red clover, red hibiscus, and purple passion flower [37]. Flavones are the least common flavonoids in food. They occur in relatively high amounts in parsley and celery (apigenin and luteolin). They are also present in lettuce, artichoke, herbs (such as rosemary, thyme, oregano, mint, sage, etc.), citrus fruits, cereal grains and sweet peppers [37]. Regarding phenolic acids, hydroxycinnamic acids (*p*-coumaric, caffeic, ferulic and sinapic acids) are more abundant in plant foods and are found in coffee and some vegetables, fruits, and cereals (particularly in potatoes, broccoli, spinach, lettuce, cabbage, apples, pears, cherries, apricots, peaches, blackcurrants, blueberries, asparagus, wine and rye bread) [9]. Differently, benzoic acids, such as gallic, protocatechuic, syringic and vanillic acids, are found in very few edible plant foods, mainly in some berries such as cloudberry and raspberry, red cabbage, chestnut and tea [38]. Then, concerning hydrolyzable tannins, ellagitannins such as sanguiin H6, punicalagin or pedunculagin occur in significant amounts in many berries, in particular strawberries and raspberries, as well as in pomegranate, muscadine grapes, walnuts, chestnuts, hazelnuts, mango, green and black tea, and are also found in oak-aged beverages (wine, whiskey, etc.) [36,37]. Gallotannins are rarely found in plant-based foods and occur almost exclusively in mango, chestnuts, and red sword bean [37]. Stilbenes are present in low quantities in the human diet, resveratrol being the most characteristic, mainly found in grapes and red wine and in smaller amounts in peanuts, pistachios, some berries, red cabbage and spinach [36,37]. Coumarins are components of plant foods of the Rutaceae (citrus) and Apiaceae (parsley, celery, etc.) families. Lignans are mainly found in flaxseed (secoisolariciresinol diglucoside and matairesinol) and sesame seeds (pinoresinol and lariciresinol). Other minor sources include cereals (triticale and wheat), legumes (soybeans and lentils), vegetables (garlic, asparagus, broccoli, carrots), and fruits (pears, prunes, strawberries, lingonberries, blackcurrants) [37]. In Table 1 are reported the basic structures, main representative compounds and dietary sources of the most important subclasses of phenolic compounds.

**Table 1.** Basic structures, main representative compounds and dietary sources of the most important subclasses of phenolic compounds.

| | **Basic Structure** | **Main Compounds** | **Food Sources** |
|---|---|---|---|

| | | | |
|---|---|---|---|
| **Flavonoids** | **Flavonols**<br>  | Quercetin<br>Kaempferol<br>Myricetin<br>Isorhamnetin | Yellow and red onion, caper, lettuce, parsley, carrot, pumpkin, kale, cabbage, broccoli, garlic, berries, mango, apples, figs |
| | **Flavan-3-ols**<br>Monomers<br>  <br>Proanthocyanidins<br>  | (+)-Catechin<br>(–)-Epicatechin<br>(–)-Epicatechin gallate<br>(+)-Gallocatechin<br><br>Procyanidin B1<br>Procyanidin B2 | Black and green tea, chocolate, red wine, nuts, apple, pear, apricot, peach, grape, berries, cereals |
| | **Isoflavones**<br>  | Daidzein<br>Genistein<br>Glycitein | Soybean and soy-derived foods, green bean, lentil, chickpea, pea, mung bean, broad bean |
| | **Flavanones**<br>  | Naringenin<br>Hesperetin<br>Neohesperidin<br>Eriocitrin | Orange, grapefruit, lemon, lime |

| | | | |
|---|---|---|---|
| | **Anthocyanidins**<br>  | Cyanidin<br>Delphinidin<br>Pelargonidin<br>Peonidin<br>Malvidin<br>Petunidin | Berries, cherries, plum, pomegranate, red wine, red cabbage, eggplant, red onion, radish, hazelnut, pistachio nut, black and red bean, medicinal herbs |
| | **Flavones**<br>  | Apigenin<br>Luteolin | Parsley, celery, lettuce, artichoke, herbs (rosemary, thyme, oregano, etc.), citrus fruits, cereal grains, sweet peppers |
| **Non-Flavonoids** | **Phenolic acids**<br>Hydroxycinnamic acid<br>  | *p*-Coumaric acid<br>Caffeic acid<br>Ferulic acid<br>Sinapic acid | Coffee, potato, broccoli, spinach, lettuce, cabbage, apple, pear, cherries, apricot, peach, blackcurrant, blueberry, asparagus, wine, rye bread |
| | Benzoic acid<br>  | Gallic acid<br>Protocatechuic acid<br>Syringic acid<br>Vanillic acid | Cloudberry, raspberry, red cabbage, chestnut, tea |
| | **Hydrolyzable Tannins**<br>Ellagitannins<br>  | Sanguiin H6<br>Punicalagin<br>Pedunculagin | Strawberry, raspberry, blackberry, pomegranate, walnut, chestnut, hazelnut, mango, green and black tea, oak-aged beverages |

| | | | |
|---|---|---|---|
| | Gallotannins  | Galloyl-hexoside<br>Digalloyl-hexoside | Mango, chestnut, red sword bean |
| | **Stilbenes**  | Resveratrol | Red wine, grape, peanuts, pistachios, some berries, red cabbage, spinach |
| | **Coumarins**  | Umbelliferone<br>Esculetine<br>Scoparone | Citrus, parsley, celery |
| | **Lignans**  | Secoisolariciresinol<br>Matairesinol<br>Pinoresinol<br>Lariciresinol | Flaxseed, sesame seeds, triticale and wheat, soybeans and lentils, garlic, asparagus, broccoli, carrots, pears, prunes, strawberries, lingonberries, blackcurrants |

*Absorption, distribution, metabolism, and excretion of phenolic compounds*

When dietary phenolics are consumed, only a small proportion is absorbed in the small intestine, while a large amount reaches the colon and becomes a suitable substrate for the local microbiota. As mentioned above, phenolic compounds, with the only exception of flavan-3-ols, exist *in planta* predominantly as glycosides and their absorption is firstly associated with a hydrolytic reaction releasing aglycones. In the small intestine there are two possible routes by which the glucoside conjugates are hydrolyzed with the resultant aglycones appearing in the epithelial cells, namely

lactase phlorizin hydrolase (LPH)-diffusion and transport-cytosolic β-glucosidase (CBG) [43]. In the first case, hydrolysis is the result of the action of LPH in the brush border of the small intestine epithelial cells [36]. The released aglycone can enter the epithelial cells by passive diffusion, thanks to its increased lipophilicity and its proximity to the enterocyte membrane. LPH exhibits broad substrate specificity for flavonoid-*O*-β-D-glucosides [44]. Alternatively, the hydrolytic step is mediated by CBG within the epithelial cells and so, in order for CBG-catalyzed hydrolysis to occur, the polar glucosides must be transported into the epithelial cells, possibly with the involvement of the active sodium-dependent glucose transporter SGLT1 [45]. In the enterocytes, (poly)phenol aglycones undergo phase II metabolism before entering systemic circulation [10]. Sulfate, glucuronide and/or methylated metabolites are formed through the respective action of sulfotransferases (SULT), uridine-5′-diphosphate glucuronosyltransferases (UGTs) and catechol-*O*-methyltransferases (COMT). Efflux back into the lumen of the small intestine of some of the metabolites occurs, and this is thought to involve members of the adenosine triphosphate (ATP)-binding cassette family of transporters, including multidrug resistance protein and P-glycoprotein [36]. Once in the portal bloodstream, metabolites pass on to the liver, where they may be subjected to further phase II metabolism before being excreted in urine [36]. The significant amounts of phenolic compounds not absorbed in the first gastro-intestinal (GI) tract pass to the colon, where the local microbiota can cleave conjugated moieties and the resultant aglycones undergo ring fission, leading to the production of phenolic catabolites and aromatic compounds. These low molecular weight catabolites can be absorbed in colonocytes and subjected to phase II metabolism before entering circulatory system and reaching the liver, where they may undergo further phase II metabolism in order to be excreted in urine [36]. In most instances, the quantity of excreted metabolites of colonic origin is well in excess of that of metabolites that enter systemic circulation *via* the small intestine [36]. Although there is speculation [46], up to now there is no compelling evidence that enterohepatic recirculation in humans results in substantial recycling of (poly)phenol metabolites back into the small intestine via biliary excretion [47].

For decades, (poly)phenols were thought to be poorly bioavailable, due to the limited detection in circulation or in urine of their phase II metabolites occurring after absorption in the first GI tract. However, when colonic catabolites started to be included in the overall estimation of absorption, distribution, metabolism, and excretion (ADME) of phenolic compounds (which is synonymous with bioavailability in nutrition research according to the European Food Information Council 2010 [48]), it was evident that many (poly)phenols are highly bioavailable. Nonetheless, it has to be considered

that several factors, first of all the complexity of the colonic microbiota composition, have been associated with inter-individual variability in the ADME of phenolic compounds and the classification of subjects as non-, low- and high-producers of phenolic catabolites. Well established examples of this inter-individual variability are the formation of urolithins from ellagitannins and equol from the isoflavone daidzein, for which metabolic phenotypes (metabotypes) have been described [49,50]. In bioavailability studies it is of note that while plasma profiles can supply useful information, they are not an accurate quantitative guide of absorption because the presence of metabolites and catabolites in the circulatory system is transient as they are rapidly removed from the bloodstream via renal excretion. More appropriate estimates of absorption are obtained by determining cumulative urinary excretion, typically over a 24-48 h period post-intake [9,51,52].

*The effects on health of phenolic compounds and their derived circulating metabolites*

Originally, the primary health benefit attributed to (poly)phenols was thought to lie in their direct antioxidant effects. However, these effects were demonstrated *in vitro* using compounds as found *in planta* or aglycones [9]. Nowadays, the antioxidant properties are no longer considered to be as relevant *in vivo*, as these compounds do not reach concentrations high enough to have a significant effect in terms of scavenging free radicals in most tissues [10]. Nonetheless, a number of other possible biochemical and molecular mechanisms of action of (poly)phenols have been identified, testing phenolic metabolites and catabolites at physiological concentrations. These other mechanisms include several effects within intra- and inter-cellular signalling pathways, such as regulation of nuclear transcription factors and fat metabolism and modulation of the synthesis of inflammatory mediators, notably cytokines tumor necrosis factor α, interleukin (IL)-1β and IL-6 [10]. For example, certain flavonoids have been shown to have a role in glucoregulation through downstream signalling that increases insulin secretion, reduces apoptosis, promotes β-cell proliferation and reduces insulin resistance, inflammation and oxidative stress in muscle and other cells [10].

In addition to *in vitro* studies, fundamental to unravel molecular mechanisms, in the last decades a number of observational and intervention studies has suggested that (poly)phenol-rich diets beneficially affect cardiometabolic health and cognitive functioning. Observational studies can determine if there are associations between a factor (e.g., intake of total phenolics) and disease incidence (e.g., occurrence of cardiovascular events) and assess the strength of these associations. Such studies also help in generating hypotheses to be tested in subsequent RCTs, the gold standards for producing scientific evidence. RCTs are intervention studies, where investigators do intervene and look at the effects of the intervention (or treatment) on a certain outcome. If well conducted,

RCTs can establish causal relationships between treatment (e.g., intake of flavan-3-ols from cocoa) and outcome (e.g., CVDs and/or clinically significant risk factors associated with such disorders). For example, regarding cardiometabolic health, a cohort study within the PREDIMED trial found an inverse association between total polyphenol intake and risk of cardiovascular-related events, independent of other dietary and non-dietary risk factors [53]. Recent systematic reviews with meta-analysis of prospective cohort studies found significant risk reductions of type 2 diabetes and CVDs when comparing subjects reporting the highest and lowest levels of intakes of different classes of phenolic compounds [54–58]. More than 25 RCTs demonstrated that flavonoid-rich foods can prevent endothelial dysfunction, an early event in the development of atherosclerosis, with an average of 20-30% improved flow-mediated dilatation after acute or short-term intake [59]. Interestingly, measurement of flavonoid metabolites in plasma showed a perfect time concordance with the observed effects [60–63]. Pioneering studies also demonstrated the effects of phenolic compounds by using randomized controlled cross-over studies to compare the impact of foods to the equivalent foods enriched or deprived of these compounds [60,62,63]. Recent meta-analyses examining the effects of (poly)phenol-rich foods on biomarkers of cardiometabolic risk revealed that overweight and obese subjects would receive more benefits from flavanols, anthocyanins and ellagitannins than normal weight people [22,23]. Moreover, a certain influence of individuals' metabolic phenotype on the effects on cardiometabolic health associated with ellagitannin and isoflavone intake has been observed. Actually, clustering healthy overweight-obese individuals according to their urolithin metabotypes revealed that only subjects characterized by urolithin metabotype B could benefit from the consumption of pomegranate by lowering of some serum CVD risk markers [64], and acute benefits of equol on arterial stiffness, an established marker of vascular function, were observed after soy intake only in men characterized by an equol producer phenotype [65]. The influence of inter-individual variability in health outcomes needs to be further assessed in new generation intervention studies [66].

Besides cardiometabolic health, phenolic compounds are also thought to have beneficial effects on cognition. A prospective study of adults in midlife found a positive association between total (poly)phenol intake and cognitive factors, such as language and verbal memory, assessed over 13 years [10,67]. In longitudinal studies, regular consumption of chocolate within the diet has been shown to reduce the risk of experiencing cognitive decline [68,69]. In a meta-analysis of 17 independent observational studies, tea consumption was found to have an inverse linear relationship with the incidence of cognitive disorders [70]. Chocolate and tea are important dietary sources of flavan-3-ols.

In long-term intervention studies, the strongest evidence of flavan-3-ols exerting benefits on cognition comes from two studies with similar methodologies conducted in healthy elderly participants [71] and sufferers of age-related cognitive impairment (*n* = 90 per study) [72]. Participants consumed drinks containing increasing amounts of flavan-3-ols - low (control), medium (520 mg) or high (990 mg) - over 8 weeks. Both studies reported that the high-flavan-3-ol drink improved attention and executive function, besides reducing insulin resistance, blood pressure and lipid peroxidation [10]. A deeper insight into the benefits of flavan-3-ols and their derived metabolites on brain functioning will be further presented in section 1.2.2.

There are some other putative health benefits attributed to phenolic compounds, such as anticarcinogenic activity and estrogenic modulation, that at present still lack of strong *in vivo* evidence in their support [37].

## 1.2.2. Focus on flavan-3-ols and their derived colonic metabolites

(The content of this section has also been published as chapter 8 "*Flavan-3-ols: Catechins and Proanthocyanidins*" [73], by **Claudia Favari**, Pedro Mena, Claudio Curti, Daniele Del Rio, Donato Angelino, in "*Dietary Polyphenols*" [37], edited by Francisco A. Tomás-Barberán, Antonio González-Sarrías, Rocío García-Villalba, 2020, Wiley-Blackwell. Permission to reuse the content in the present thesis has been granted by the Publisher. The content of the chapter has been slightly modified to be conformed with the current thesis, in accordance with the license obtained n. 1139321-1)

### *Chemistry and main dietary sources of flavan-3-ols*

Flavan-3-ols are the most complex subclass of flavonoids, with structures of various molecular weight ranging from simple monomers to oligomers and polymers of up to 190 units. The monomeric flavan-3-ol presents two stereogenic centers at C2 and C3 that produce four possible stereoisomers (two diastereomeric couples of enantiomers) for each level of B-ring hydroxylation and shows two hydroxyl group in C5 and C7. The hydroxyl groups in the B ring can range from one to three, yielding different structures: (epi)afzelechin, with one hydroxyl at C4'; (epi)catechin, two hydroxyls at C3' and C4'; and (epi)gallocatechin, three hydroxyls at C3', C4', and C5' (Figure 4). (+)-Catechin and (–)-epicatechin are the most common flavan-3-ol monomers, being widespread in nature. Monomers can also undergo esterification with gallic acid, forming (epi)(gallo)catechin derivatives [9,36]. Differently from other flavonoids, flavan-3-ols exist *in planta* predominantly as aglycones rather than glycosides. Proanthocyanidins are the oligomeric and polymeric structures of flavan-3-ols, also known as condensed tannins. Two types of PACs are distinguished: type B and type A. Type B PACs are formed from flavan-3-ol monomers by oxidative coupling between the C4 of the upper unit and the C6 or C8 of the adjacent lower or extension unit (Figure 5). Type A PACs have an

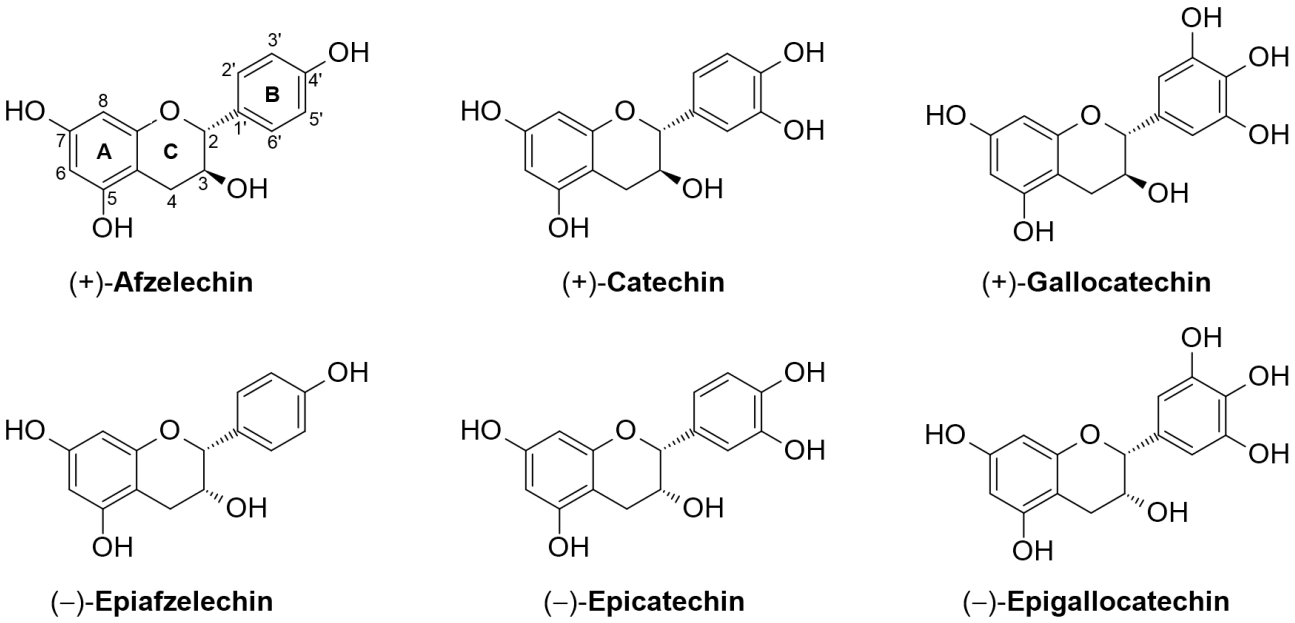



**Figure 4.** Stereoisomers of monomeric flavan-3-ols with different patterns of hydroxylation.

additional ether bond between C2 and C7 (Figure 5) and are less widespread in the plant kingdom than type B PACs. Oligomeric and polymeric structures have an additional chiral center at C4 of every additional flavan-3-ol unit. Condensed tannins made up exclusively of (epi)catechin units are called procyanidins, whereas propelargonidins and prodelphinidins are the oligomeric and polymeric structures consisting of (epi)afzelechin or (epi)gallocatechin subunits, respectively. Procyanidins are the most abundant type of PACs in plants, while propelargonidins and prodelphinidins are comparatively rare. Many PACs contain more than one type of monomers, and also linkages with anthocyanins and flavonols can occur [9,36,51].

Flavan-3-ols are characteristic polyphenols of tea, cocoa, wine, pome fruits (as apple and pear), berries, and nuts, but they are also found in stone fruits and legumes [39,40] (Table 2). This subclass of compounds is the main source of flavonoids in Western diets [8,41]. Different dietary sources contain distinct types of flavan-3-ols (Table 2). Green tea, the beverage prepared from the leaves of *Camellia* spp., contains very high levels of monomers, with the main components being (–)-epigallocatechin-3-*O*-gallate, (–)-epicatechin-3-*O*-gallate, and (–)-epigallocatechin (Figure 6). Black tea, the product of the fermentation of green leaves, presents lower amounts of monomers, as a result of the action of polyphenol oxidase, and a concomitant accumulation of theaflavins and thearubigins. Theaflavins and thearubigins are dimer- and polymer-like structures, that can also be esterified with gallic acid [9,36] (Figure 6). These compounds, that are the result of the monomer transformation during tea leaf fermentation, also belong to the class of flavan-3-ols. Cocoa (*Theobroma cacao*) beans and their derived products (cocoa powder and chocolate) are rich sources of (+)-catechin, (–)-epicatechin and

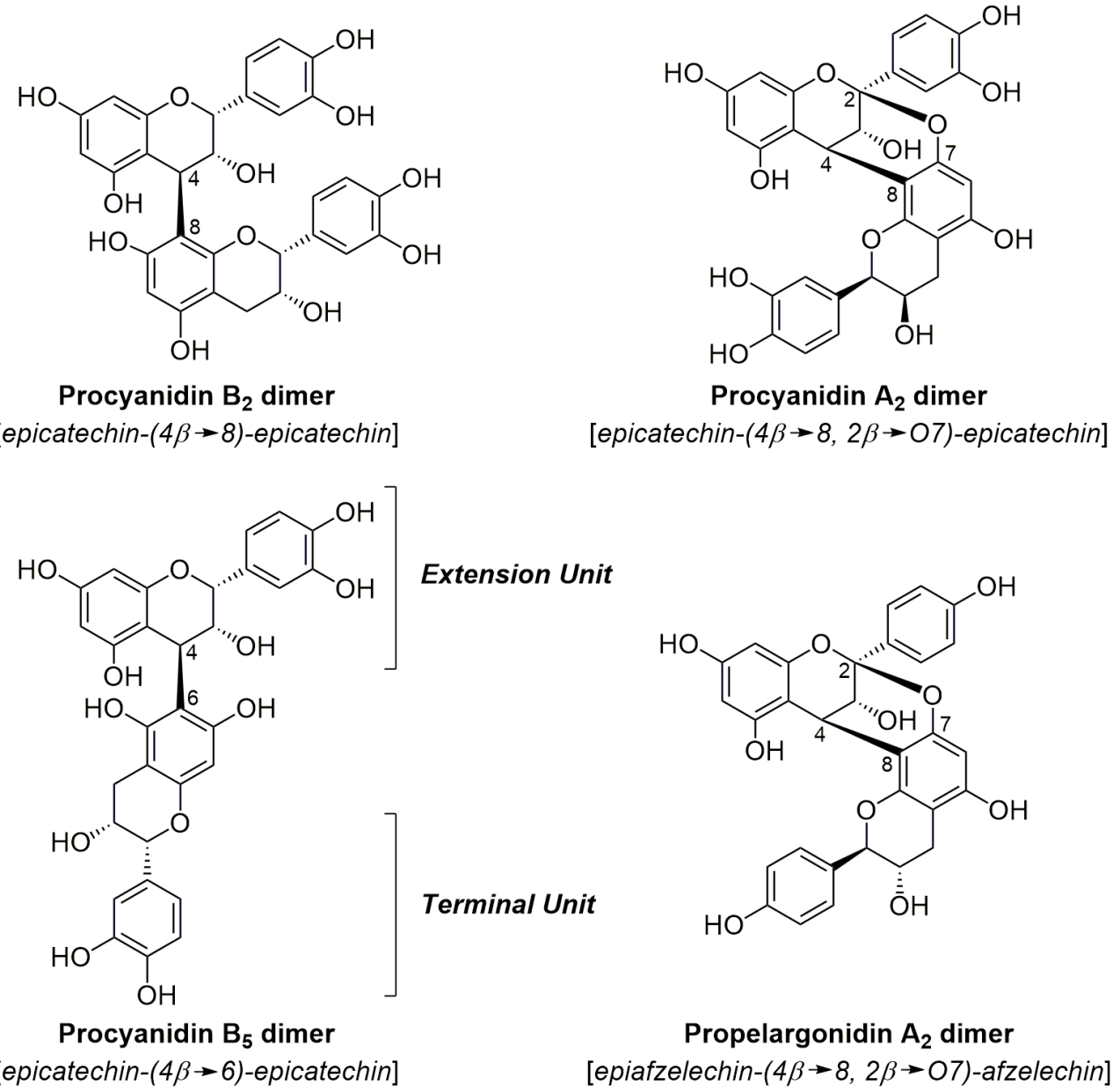



**Figure 5.** Dimeric B-type and A-type proanthocyanidins.

oligomeric B-type procyanidins (from dimers to decamers) [36]. Wine, the alcoholic fermented juice of *Vitis vinifera* grapes, as well as apples (*Malus domestica*), pears (*Pyrus communis*), and berries are also characterized by high contents of oligomeric and polymeric PACs [36,51].

*Bioavailability of flavan-3-ols*

The bioavailability of a dietary compound is a crucial factor to understand to what extent and in which forms it is available for the individual internal compartments. Over the last decades, significant results have been accomplished in the elucidation of the mechanisms of ADME of (poly)phenolic compounds, and among them of flavan-3-ols, after dietary intake [74]. This information is fundamental to unravel the biological effects of these compounds in human health. At first

instance, it has to be considered that many factors affect the bioavailability of flavan-3-ols: the stereochemical configuration, the degree of polymerization, the flavanol-containing food matrix ingested, as well as inter-individual variations (including sex, genetic background, gut microbiota composition, ethnicity, age, dietary habits, etc.) [75].

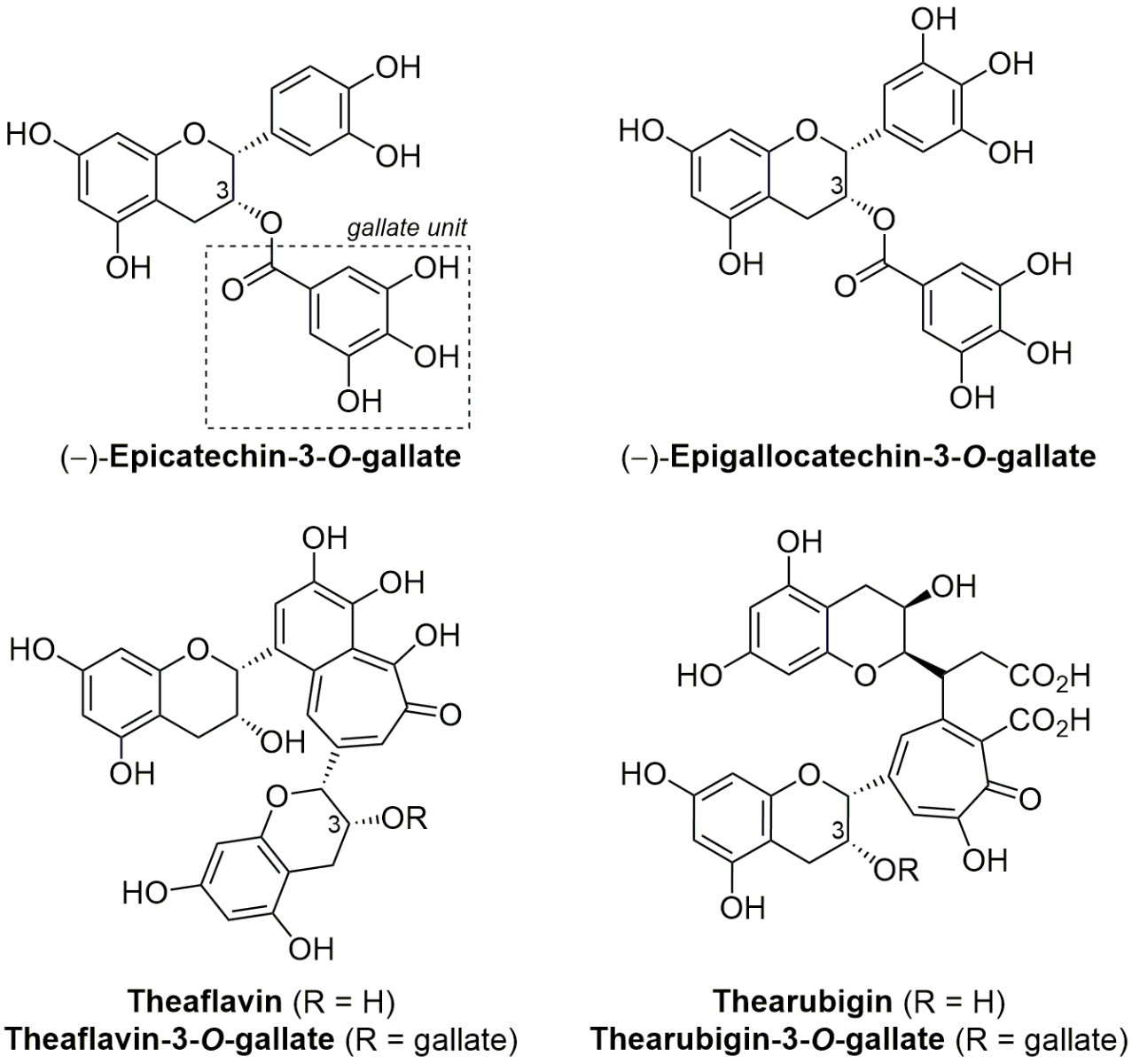



**Figure 6.** Primary green and black tea flavan-3-ol derivatives.

**Table 2.** Examples of dietary sources of flavan-3-ols and content of their main flavan-3-ols (PhenolExplorer [76]).

| Dietary Source | Main Flavan-3-ols | Content* (mean ± SD) |
|---|---|---|
| Green tea, *infusion* (*Camellia sinensis* L. Kuntze) | (–)-Epigallocatechin-3-*O*-gallate | 27.16 ± 39.91 mg/100 mL |
| | (–)-Epigallocatechin | 19.68 ± 25.11 mg/100 mL |
| | (–)-Epicatechin-3-*O*-gallate | 7.50 ± 10.19 mg/100 mL |

| | | |
|---|---|---|
| Black tea, *infusion* (*Camellia sinensis* L. Kuntze) | (–)-Epigallocatechin-3-*O*-gallate<br>(–)-Epicatechin-3-*O*-gallate<br>(–)-Epigallocatechin | 9.12 ± 12.67 mg/100 mL<br>7.34 ± 7.10 mg/100 mL<br>7.19 ± 10.87 mg/100 mL |
| Cocoa, *powder* (*Theobroma cacao*) | (–)-Epicatechin<br>(+)-Catechin<br>Procyanidin dimer B1 | 158.30 ± 86.33 mg/100 g FW<br>107.75 ± 63.95 mg/100 g FW<br>112.00 ± 0.00 mg/100 g FW |
| Chocolate (dark) | (–)-Epicatechin<br>Procyanidin tetramer D<br>Procyanidin dimer B2 | 70.36 ± 29.54 mg/100 g FW<br>53.83 ± 20.09 mg/100 g FW<br>36.50 ± 11.69 mg/100 g FW |
| Red wine (*Vitis vinifera* L.) | Procyanidin dimer B3<br>Procyanidin dimer B4<br>(+)-Catechin | 9.47 ± 4.29 mg/100 mL<br>7.29 ± 3.78 mg/100 mL<br>6.81 ± 6.24 mg/100 mL |
| Grape, black (*Vitis vinifera* L.) | (+)-Catechin<br>(–)-Epicatechin<br>(–)-Epicatechin-3-*O*-gallate | 5.46 ± 5.74 mg/100 g FW<br>5.24 ± 5.61 mg/100 g FW<br>1.68 ± 1.87 mg/100 g FW |
| Apple (*Malus domestica*) | Procyanidin dimer B2<br>(–)-Epicatechin<br>(+)-Catechin | 14.56 ± 9.19 mg/100 g FW<br>8.37 ± 3.67 mg/100 g FW<br>1.22 ± 0.83 mg/100 g FW |
| Pear (*Pyrus communis*) | (–)-Epicatechin<br>(+)-Catechin | 3.77 ± 2.65 mg/100 g FW<br>0.28 ± 0.35 mg/100 g FW |
| Apricot (*Prunus armeniaca* L.) | (–)-Epicatechin<br>(+)-Catechin<br>Procyanidin dimer B1 | 3.47 ± 4.27 mg/100 g FW<br>2.96 ± 3.28 mg/100 g FW<br>0.09 ± 0.00 mg/100 g FW |
| Peach (*Prunus persica* L.) | (+)-Catechin | 2.33 ± 0.00 mg/100 g FW |
| Plum (*Prunus domestica* L.) | Procyanidin trimer C1<br>Procyanidin dimer B1<br>Procyanidin trimer isomer | 10.01 ± 0.00 mg/100 g FW<br>8.84 ± 0.00 mg/100 g FW<br>7.73 ± 0.00 mg/100 g FW |
| Hazelnut (*Corylus* L.) | (–)-Epigallocatechin<br>(+)-Catechin<br>(–)-Epigallocatechin-3-*O*-gallate | 2.80 ± 0.00 mg/100 g FW<br>1.20 ± 0.00 mg/100 g FW<br>1.10 ± 0.00 mg/100 g FW |
| Almond (*Prunus dulcis*) | (–)-Epigallocatechin<br>(+)-Catechin<br>(–)-Epicatechin | 2.60 ± 0.00 mg/100 g FW<br>1.28 ± 1.04 mg/100 g FW<br>0.59 ± 0.35 mg/100 g FW |

| | | |
|---|---|---|
| Lentil (*Lens culinaris*) | (+)-Catechin-3-*O*-glucose | 3.15 ± 0.00 mg/100 g FW |
| | Procyanidin dimer B3 | 0.71 ± 0.33 mg/100 g FW |
| | Prodelphinidin dimer B3 | 0.45 ± 0.00 mg/100 g FW |

* Results expressed as mg/100 g FW are relative to the Fresh Weight of the product.

Flavan-3-ols undergo an extensive metabolism once introduced into the GI tract. After ingestion, they pass through the oral cavity and the stomach remaining almost unchanged, thus reaching the small intestine. Here, some monomers are absorbed in the enterocytes by passive diffusion and subjected to some degree of phase-II enzymatic metabolism. Sulfotransferases (SULT), uridine-5'-diphosphate-glucuronosyl-transferases (UGTs) and catechol-*O*-methyltransferases (COMT) originate sulfated, glucuronidated and *O*-methylated metabolites respectively (Figure 7), which pass through the portal vein to the liver or efflux back into the lumen mediated via members of the adenosine-binding cassette (ABC) family of transporters [9,51]. Unlike simple monomers, flavan-3-ols with a 3-gallate moiety do not necessarily undergo conjugation by phase-II enzymes since they have been detected in the circulation unmetabolized. Besides, a very small percentage of oligomers (< 1%, mainly dimers) seems to be absorbed in the small intestine, getting into the portal bloodstream in unconjugated forms [74]. Absorbed flavan-3-ols rapidly reach the liver, where they may be further subjected to phase II metabolism before entering the systemic circulation and eventually undergoing renal excretion. A small amount of phase II metabolites may also return from the liver to the small intestine via enterohepatic recirculation in the bile [9,51]. However, most of ingested flavan-3-ols (> 70%) are not absorbed in the upper part of the GI tract and reach the colon, where they are extensively metabolized by the host microbiota [74].

Regarding the microbial metabolism of flavan-3-ol monomers, free (epi)(gallo)catechins present in the large intestine undergo: (1) C-ring fission by the action of specific bacterial species yielding a di- or trihydroxyphenylpropan-2-ol derivative, (2) subsequent conversion into phenyl-γ-valerolactones and 4-hydroxy-phenylvaleric acids by the action of *Flavonifractor plautii*, (3) possible dehydroxylations, and (4) further transformation into low molecular weight phenolics [77–79] (Figure 8). Free (epi)(gallo)catechins subject to these microbial metabolic pathways can derive from: (a) undigested simple monomers, (b) galloyl-moiety removal from 3-gallate monomers by microbial esterases, (c) PAC interflavan link cleavage operated by specific microorganisms. The number and the position of hydroxyls in the phenyl group of these microbial-derived catabolites depend on the characteristics of the (epi)(gallo)catechin precursor. In particular, free (epi)catechin is catabolized into a 1-(3',4'-dihydroxyphenyl)-3-(2'',4'',6''-trihydroxyphenyl)propan-2-ol, before being converted

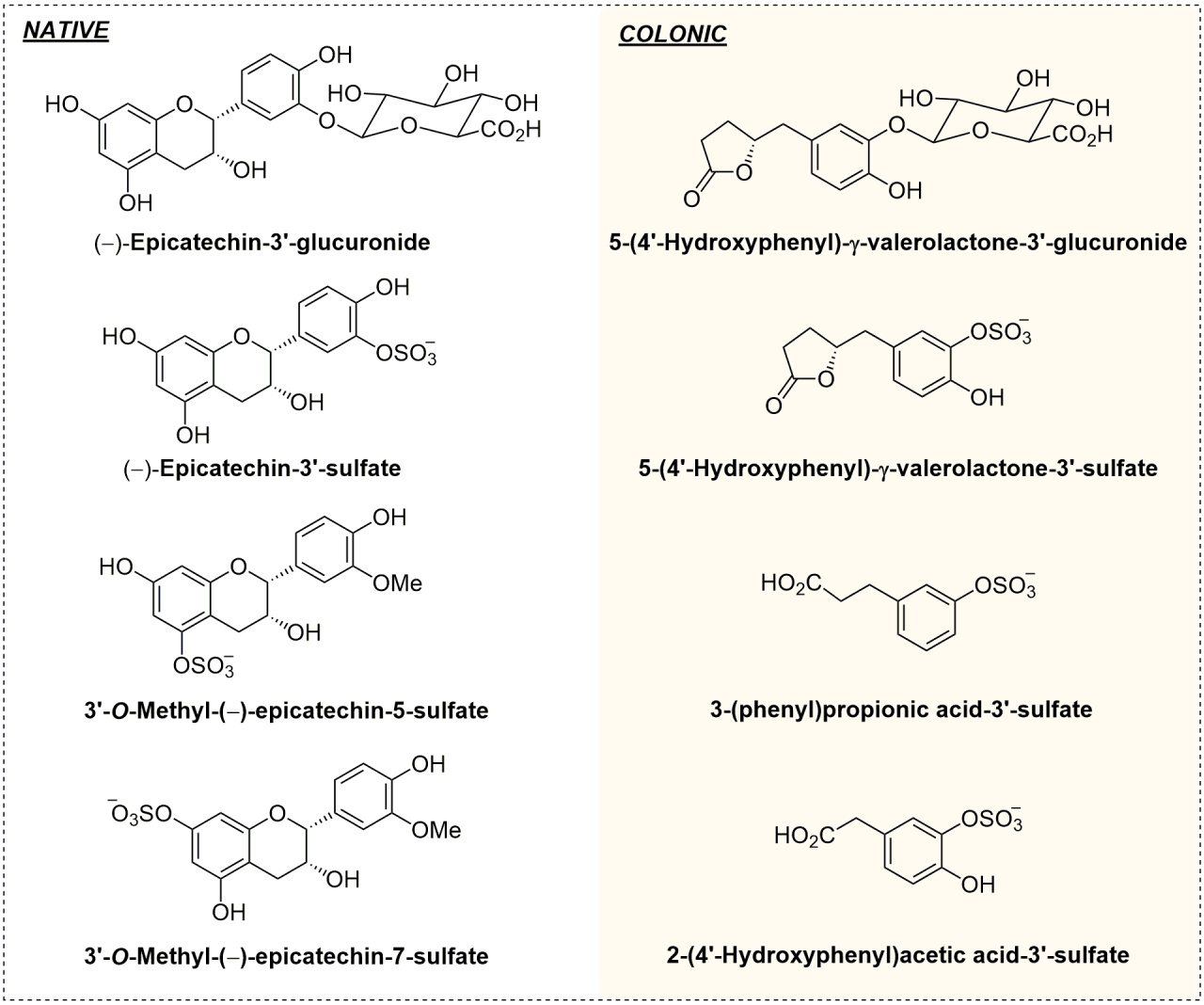



**Figure 7.** Structures of leading native and colonic flavan-3-ol metabolites.

into 5-(3′,4′-dihydroxyphenyl)-γ-valerolactone and/or 4-hydroxy-5-(3′,4′-dihydroxyphenyl)valeric acid [77–79] (Figure 8). The γ-valerolactone ring can also be opened to 4-hydroxy-5-(3′,4′-dihydroxyphenyl)valeric acid and/or be later dehydroxylated to 5-(3′,4′-dihydroxyphenyl)valeric acid[78] (Figure 8). When (epi)gallocatechin is the colonic precursor, 1-(3′,4′,5′-trihydroxyphenyl)-3-(2′′,4′′,6′′-trihydroxyphenyl)propan-2-ol is produced and converted into 5-(3′,4′,5′-trihydroxyphenyl)-γ-valerolactone and 4-hydroxy-5-(3′,4′,5′-trihydroxyphenyl)valeric acid. These trihydroxyphenyl catabolites can be dehydroxylated to the analogous 3′,5′-dihydroxyphenyl and 3′,4′-dihydroxyphenyl derivatives, both of which can subsequently yield the 3′-hydroxyphenyl derivative, and the 3′,4′-dihydroxyphenyl derivative potentially also the 4′-hydroxyphenyl derivative [74].

Phenyl-γ-valerolactones (PVLs) and their related phenylvaleric acids (PVAs) are the main catabolites produced, but these compounds can be further converted by the intestinal microbiota into other

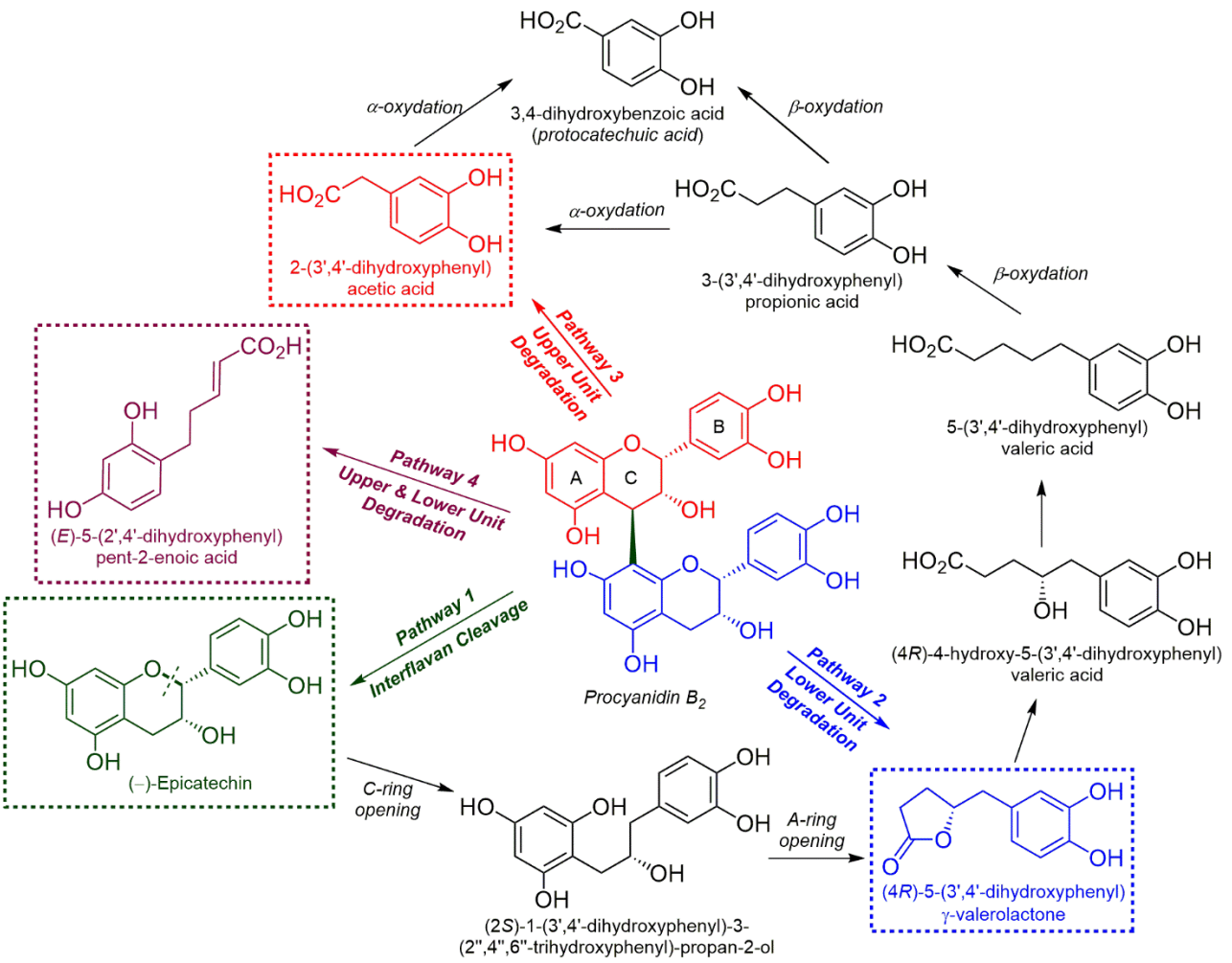



**Figure 8.** Exemplified catabolism of procyanidin B2 and (−)-epicatechin by gut microbiota. Pathways proposed, according to Appeldoorn et al. [78] and Stoupi et al. [79]

low molecular weight phenolics, such as phenylpropionic, benzoic and cinnamic acid derivatives, by successive loss of carbon atoms from the side chain through β-oxidation [79]. In this sense, a single monomeric precursor can lead to a broad array of phenolic metabolites. Concerning B-type PAC dimers and oligomers, besides the possible depolymerisation into monomeric units previously described (which represents a minor catabolic route, Figure 8 pathway 1), the main degradation pathways involve the direct production of PVLs [78,79] (Figure 8 pathway 2) and other low molecular weight phenolics [78,79] (Figure 8 pathway 3 and 4). For example, for procyanidin B2, 5-(3',4'-dihydroxyphenyl)-γ-valerolactone could result from the direct degradation of its lower unit (Figure 8 pathway 2), while 2-(3',4'-dihydroxyphenyl)acetic acid could derive from the cleavage of its upper unit (Figure 8 pathway 3). Moreover, other microbial metabolites, such as 5-(2',4'-dihydroxyphenyl)-2-ene-valeric acid, could arise from the simultaneous degradation of the upper and lower units [79] (Figure 8 pathway 4). Differently, less is known about the microbial catabolism of A-type PACs, which are more resistant to microbial breakdown than B-type. This resistance could be probably due to

their more rigid interflavan ether bonds [80,81]. Like B-type dimer catabolism, degradation of A-type procyanidins starts with the cleavage of monomeric unit C-rings, followed by the production of various phenolic acids [81]. The occurrence of PVA derivatives after incubation of A-type dimers with colonic microbiota has been observed, while the formation of PVLs from A-type PACs has not been reported to date [80]. The products of the colonic microbial catabolism of flavan-3-ols (mainly PVLs, PVAs, and phenolic acids) can be absorbed in colonocytes by passive transport and be subjected to enzymatic phase II metabolism, before getting into the portal bloodstream to reach the liver. In hepatocytes, these catabolites can be further conjugated by phase II enzymes before passing to the systemic circulation or getting expelled back into the intestinal lumen via enterohepatic recirculation. The circulating fraction is then excreted in urine [74]. Lastly, undigested flavan-3-ols (mostly high molecular weight PACs with a degree of polymerization, DP, > 4) and unabsorbed catabolites are voided in faeces. Faeces may also contain conjugated metabolites released from enterocytes or excreted through the bile [82,83].

Over the years, numerous human intervention studies have been carried out for evaluating the profiles of absorption and excretion of flavan-3-ols to understand to what extent and in which forms these compounds are bioavailable for the human organism. Ottaviani and colleagues investigated the ADME of (–)-epicatechin, one of the most widely consumed flavan-3-ols, using radiolabeled and stereochemically pure [2-$^{14}$C](–)-epicatechin ([2-$^{14}$C]EC) [84]. This fundamental study, that will be followed as an example along this section, revealed that levels of recovered radioactivity were substantially higher in urine than in plasma, as determined by liquid scintillation counting, with mean total radioactivity in plasma never exceeding 2% of intake, while reaching 82% ± 5% of ingested [2-$^{14}$C]EC in urine, over the 0-48 h period after consumption. Inter-individual differences were observed, especially in the recovery of radioactivity in urine, with individual values ranging from 49 to 90% [82]. Analysis of blood samples showed that radioactivity was associated almost exclusively with plasma rather than blood cellular components. The pharmacokinetic profile was biphasic with maximum peaks at ~1 h and 6 h after [2-$^{14}$C]EC intake. A $T_{max}$ of 1.0 ± 0.1 h is indicative of absorption in the small intestine and was related to the presence of 12 native phase II metabolites, mainly in the forms of (–)-epicatechin-3'-glucuronide, 3'-*O*-methyl-(–)-epicatechin-5-sulfate, and (–)-epicatechin-3'-sulfate. Native phase II (–)-epicatechin metabolites reached an overall $C_{max}$ of 1223 ± 104 nmoL/L, before declining rapidly with an apparent elimination half-life ($AT_{1/2}$) of 1.9 ± 0.1 h and disappearing from the circulatory system within eight h. On the other hand, a $T_{max}$ of 5.8 ± 0.4 h is indicative of absorption in the distal GI tract and it was characterized by the

occurrence of colonic phase II catabolites, namely the PVL and PVA derivatives. The main ones were 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-sulfate and 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide. Microbial-derived metabolites attained a combined $C_{max}$ of 588 ± 102 nmoL/L, but they were retained in the circulation longer than the native phase-II metabolites, showing an $AT_{1/2}$ of 5.7 ± 0.7 h and area under the curve (AUC) concentration ~3-fold higher. In fact, ~0.2% of [2-$^{14}$C]EC intake was still present in plasma 24 h after consumption. Radioactivity associated to phenolic acid derivatives in plasma was low and this precluded their detection by HPLC-RC, making it impossible to quantify in plasma how much of the ingested (–)-epicatechin was converted into these low molecular weight phenolics. Inter-individual variations in plasma $C_{max}$ of metabolites were reported, with higher differences observed for colonic catabolites concentrations [82,84]. Urinary excretion reflected the plasma pharmacokinetic profiles. Native phase-II (–)-epicatechin metabolites were mostly excreted in the first 0-4 h collection period and represented 20% ± 2% of the ingested [2-$^{14}$C]EC, absorbed in the proximal GI tract. As in plasma, the major urinary structurally-related (–)-epicatechin metabolites (SREMs) were (–)-epicatechin-3'-glucuronide, 3'-*O*-methyl-(–)-epicatechin-5-sulfate, and (–)-epicatechin-3'-sulfate. PVL and PVA derivatives were excreted later, mainly over the 4-8 h, 8-12 h, and 12-24 h collection periods. Their urinary excretion corresponded to 42 ± 5 % of the consumed [2-$^{14}$C]EC and, as in plasma, the main representatives were 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-sulfate and 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide. Concerning low molecular weight phenolics, their urinary excretion continued over 48 h post [2-$^{14}$C]EC intake and represented 28 ± 3 % of the ingested radioactivity. Thus, recovery of microbial-derived catabolites absorbed in the colon was about 70% of the consumed [2-$^{14}$C]EC. Variability among individuals in the levels of urinary excretion was 2.9-fold for native phase II (–)-epicatechin metabolites and 3.2-fold for PVL and PVA derivatives, while there was a much more substantial variation with low molecular weight phenolics, but these were relatively minor metabolites [82,84]. When considering dimeric flavan-3-ols, native procyanidin B1 can be rapidly absorbed in enterocytes, though in small quantities, and appear in circulation mainly unconjugated. It is also quickly excreted, after undergoing enzymatic methylation and then glucuronidation or sulfation. Colonic-derived 5-(3',4'-dihydroxyphenyl)-γ-valerolactone conjugates were the most important metabolites detected later after procyanidin B1 intake, exhibiting high plasma levels and significant urine excretion. High inter-individual variability was also reported [85].

It must be taken into account that, unless consumed as food supplements, (–)-epicatechin and other flavan-3-ols are not consumed as pure compounds, but as part of more complex food matrixes,

which may affect their bioavailability and pharmacokinetic profiles. Plenty of human intervention studies tried to assess ADME of flavan-3-ols using their primary dietary sources, in particular tea and cocoa products, but also grape and grape-by products, cranberry, apples and almonds [86,87,96–101,88–95]. In general, native phase-II metabolites show shorter plasma $T_{max}$ than colonic phase-II metabolites, reflecting a proximal GI tract absorption versus a distal GI tract absorption, as well as higher plasma $C_{max}$ but lower $AT_{1/2}$. Cumulative urinary excretion of colonic phase-II metabolites is instead more elevated than that of native phase II metabolites. Monoglucuronide and monosulfate forms of flavan-3-ol monomers are the dominant native phase II metabolites, as well as monoglucuronide and monosulfate derivatives of PVLs are the main colonic phase-II catabolites. In particular, glucuronide- and sulfate-conjugates of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone are the most important flavan-3-ol metabolites, regardless of the food source, and have therefore been proposed as biomarker of flavan-3-ol intake [102]. High inter-individual variability is always recorded, especially in the production of PVLs and PVAs, likely due to individual differences in the gut microbiota composition. Thus, the possible existence of different metabotypes (metabolic phenotypes) in the production of flavan-3-ol colonic metabolites has been proposed. They were observed in a free-living study in which volunteers were daily supplemented with very high amounts of green tea flavan-3-ols [103]. Three putative metabotypes related to green tea flavan-3-ols were distinguished and characterized by: (1) an elevated excretion of tri- and dihydroxyphenyl-γ-valerolactones and a reduced excretion of 3-(hydroxyphenyl)propionic acids (metabotype 1); (2) a medium excretion of dihydroxyphenyl-γ-valerolactones and reduced excretion of trihydroxyphenyl-γ-valerolactones and 3-(hydroxyphenyl)propionic acids (metabotype 2); and (3) a high excretion of 3-(hydroxyphenyl)propionic acids and limited production of phenyl-γ-valerolactones (metabotype 3). Recently, a study on this topic has somehow denied the existence of metabotypes coming from nut proanthocyanidins. Differences in the flavan-3-ols consumed in each study (green tea monomers vs. nut oligomers) might account for these differences. However, it is important to emphasise that the results between both studies are not contradictory. The work by Cortés-Martín and colleagues [104] observed clusters of individuals using the same metabolites that were considered in the tea study by Mena and collaborators [103]. In particular, Cortés-Martín and colleagues observed four clusters representing groups of volunteers with different concentrations of hydroxyphenyl-γ-valerolactones and hydroxyphenylpropionic acids [104], fully in line with the preliminary study [103]. A relevant aspect supporting the robustness of these clear inter-individual differences in the production of flavan-3-ol colonic catabolites is related to a limitation of the work carried out with

nuts proanthocyanidins, since the amount of proanthocyanidins provided under free living conditions was very low (54 mg per day) and cannot be considered a good source of flavan-3-ol catabolites (80% of the nut flavan-3-ols presented a mean degree of polymerization ranging from 2.4 to 8.5). In this sense, it could be considered that Cortés-Martín and collaborators confirmed the presence of different metabolic profiles in a work closer to a free-diet setting than to an intervention with sources of flavan-3-ols able to yield colonic metabolites [104]. The only contradictory point between these studies is related to the concept of metabotypes that both groups have considered: while the study by Mena and colleagues [103] considered the definition of metabotyping as the classification of individuals in subgroups according to their metabolic profile (as usually done in the nutrition field [105]), Cortés-Martín and collaborators adhered to a more restrictive definition considering "gut microbiota polyphenol metabotypes", which are characterised by the presence–absence of specific metabolites for specific individuals [104]. Therefore, beyond definitions, it is important to highlight that Cortés-Martín and colleagues confirmed the existence of a high inter-individual variability in the production of colonic metabolites of flavan-3-ols and that volunteers can be grouped according to different metabolic profiles [104]. Nevertheless, further research is needed to confirm the existence and stability of these metabotypes and to understand how they may vary based on the flavan-3-ol profile of the food source consumed. The elucidation of metabotypes is of particular interest to fully understand the potential health benefits associated with flavan-3-ol consumption and their microbiota-derived metabolites on an individual basis.

*Health benefits of flavan-3-ols and their derived circulating metabolites*

Being categorized as "natural antioxidants", together with all the phenolic compounds, the first studies on the health effects of favan-3-ols focused on their role on oxidative stress and redox state of the cellular environment [106,107]. Today, the most relevant health benefits attributed to flavan-3-ols are those towards cognitive decline and inflammatory status and cardiometabolic health. Moreover, flavan-3-ols and their derived circulating and excreted metabolites might exert a beneficial action against urinary tract infections (UTIs).

Regarding the beneficial action on nervous system functions, Schaffer and Halliwell [108] suggested that flavan-3-ols might alter brain function at three locations: i) outside the central nervous system (i.e. by improving cerebral blood flow), ii) at the blood-brain barrier (i.e. by altering multi-drug resistant protein-dependent influx and efflux mechanisms of various biomolecules) and iii) inside the central nervous system (i.e. by directly modifying the activity of neurons and glial cells). The first hypothesized mechanism has been proven in a recent RCT involving sedentary older adults

consuming, for 3 months, a high (900 mg/d) or low (45 mg/d) dose of cocoa flavan-3-ol supplements [109]. Results highlighted that high flavan-3-ol supplementation significantly enhanced the hippocampal dentate gyrus function by increasing cerebral blood volume (measured using functional magnetic resonance imaging) compared to those who were supplemented with low-flavan-3-ol cocoa. This study confirmed a previous one from the same group, where young adults, after the consumption of pure (–)-epicatechin (up to 3 mg/kg) or a flavan-3-ol-rich cocoa drink (917 mg/d), showed a significant transient increase in the flow-mediated vasodilation response at 1–4 h post-consumption compared to baseline and a low-flavan-3-ol cocoa beverage (37 mg/d) [110]. Authors demonstrated that the increase of (–)-epicatechin and its metabolites, such as 4′-O-methyl-epicatechin-glucuronide, 4′-O-methyl-epicatechin, and epicatechin-7-glucuronide, significantly correlated with increasing concentrations of plasma and urinary nitric oxide, as well as with the induction of nitric oxide synthesis. Authors concluded that this might be the mechanism behind amelioration of flow-mediated vasodilation after flavan-3-ol consumption [110].

The close link between the cerebral blood flow (including oxygen and nutrient delivery) and brain cell activities is defined as neurovascular coupling. A direct association has been underlined between impaired neurovascular coupling and cognitive diseases [111]. Sorond and colleagues investigated whether flavan-3-ol-rich cocoa could affect neurovascular coupling through the improvement in endothelial and cognitive functions in a parallel-arm, double-blind clinical trial, where sixty older volunteers, with or without impaired neurovascular coupling, consumed, for 30 days, up to 1200 mg/d of cocoa flavan-3-ols [112]. Cocoa consumption was associated with significant increased neurovascular coupling for impaired individuals and amelioration of their cognitive performance after the intervention, compared to healthy elderlies. Other than the above-mentioned biological effects, authors hypothesized also cholinergic, anti-inflammatory, antioxidant, and β-amyloid–reducing properties for cocoa flavan-3-ols [112].

Flavan-3-ols have been consistently studied for their preventive role towards as mild as severe cognitive decline models, such as dementia and Alzheimer's disease. Concerning the former, one of the leading studies in the field is the Cocoa, Cognition, and Aging (CoCoA) Study, a RCT where 90 elderlies with mild cognitive impairment were supplemented daily, for 8 weeks, a drink containing high, intermediate or low cocoa flavan-3-ols (990, 520, or 45 mg/day, respectively) [72,113]. Not only cognitive functions, assessed by Mini-Mental State Examination, Trail Making Test A and B, and verbal fluency test, but also insulin resistance, blood pressure, and lipid peroxidation markers improved at eight weeks following the consumption of the high flavan-3-ol beverage in comparison

to the intermediate and lower ones. Consequently, it was hypothesized that circulating flavan-3-ol metabolites might have improved the vascular reactivity, such as nitric oxide-dependent endothelial function, or markers of oxidative stress and, in turn, they promoted neuronal functions [113].

Oxidative damage has been indicated as one of the most dangerous enhancers of neuronal degeneration, caused by the accumulation of iron or radical species – mainly derived by oxygen or nitrogen – which trigger neuroinflammation and depletion of endogenous antioxidant enzymes and compounds [114]. Green tea catechins have been considered among the most effective polyphenols counteracting neuroinflammation and oxidative processes. Epigallocatechin-gallate (EGCG) is regarded as a "multimodal acting molecule", as it seems to be involved in various cellular neuroprotection mechanisms including iron-chelation, scavenging of oxygen and nitrogen radical species, and activation of protein kinase-C signaling pathway [115]. The biological activity of EGCG also shows a characteristic biphasic pattern, where higher doses exhibit pro-oxidant and pro-apoptotic effects, supposedly involved in its anti-tumorigenic role, whereas low doses have neuroprotective effects [116]. Other potential mechanisms involve EGCG binding to peroxisome proliferator-activated receptors, which pleiotropically act as transcription factors for a large number of downstream target genes [117]. The protective effects of green tea catechins on Alzheimer's disease have also been considered in several human intervention studies. Three Japanese studies focusing on green tea beverages confirmed the improvement in cognitive functions (by means of Mini-Mental State Examination or Dementia Scales) of volunteers with different shades of cognitive impairment and dysfunction, after the consumption of more than 200 mg/day green tea catechins, up to 12 months, compared to placebo [118].

Very few studies focused on the putative biological role of flavan-3-ol gut microbial metabolites, PVLs and PVAs, on neuronal improvement, so far. Unno and collaborators treated human SH-SY5Y neuroblastoma cells with 5-(3',5'-dihydroxyphenyl)-γ-valerolactone, 5-(5'-hydroxyphenyl)-γ-valerolactone-3'-sulfate, and 5-(5'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide, in the range of 0.01–1.0 mM for 48 h [119]. These colonic metabolites induced an increase in neurite numbers [119]. A recent work carried out by Angelino and colleagues, in the frame of an international collaboration, have used five experimental models, with increased complexisty and experimental realism (including *in silico*, *in vitro* and *in vivo* animal studies, to understand whether PVLs and PVAs are able to cross the blood-brain barrier [120]. Results showed the blood-brain barrier permeability of one of the main flavan-3-ol microbial metabolites, 5-(hydroxyphenyl)-γ-valerolactone-sulfate, supporting the hypothesis that those microbial derivatives might be present in brain tissues. Therefore, further

works with this metabolite may shed light on the prospects of flavan-3-ol colonic metabolites for the prevention of neurological disorders.

Several different flavan-3-ol-rich foods have been considered for their beneficial effects towards inflammation, as observational studies reported an association between flavan-3-ol intake and positive impact on cardiometabolic outcomes, i.e. cholesterol levels, blood pressure, and myocardial infarction [10]. Among these, a prospective study on the Cancer Prevention Study II Nutrition cohort showed that flavan-3-ol intake was positively associated with a lower risk of fatal cardiovascular disease [121], while a meta-analysis of cohort, case-control, and cross-sectional studies reported more moderate risk of any cardiovascular disease and diabetes and a reduced risk of stroke in individuals who consumed higher levels of cocoa and chocolate [122]. Similarly, a recent meta-analysis focused on inter-individual variability, also indicated that flavan-3-ols derived from green tea, apple, and cocoa products might improve blood lipid levels, in particular in overweight adults [123].

A recent meta-analysis of Berends and colleagues summarized findings from RCTs focused on both short- and long-term consumption of chocolate and cocoa-based products, and it evidenced significant decrease in blood pressure (both systolic and diastolic), improvements in insulin resistance, in flow-mediated dilatation (FMD), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C) [124]. Again, another review which considered RCTs on the protective role of berry (poly)phenols, among which flavan-3-ols, towards markers of cardiovascular disease, pointed out that the most active positive effects were shown on blood pressure, lipid-related markers, and FMD [125]. Exciting and contrasting results have been highlighted when not only flaval-3-ols, but also their colonic metabolites, have been considered as biological effectors. A recent RCT found that the consumption of cocoa extract, containing 130 mg (–)-epicatechin and 560 mg procyanidins, by 45 healthy men for 4 weeks significantly improved FMD, decreased blood pressure, arterial stiffness markers (pulse wave velocity and augmentation index) and total cholesterol, compared to those who consumed 20 mg (–)-epicatechin and 560 mg procyanidins, and control [126]. Authors hypothesized that flavan-3-ol monomers and their derived SREMs could be more effective in the improvement of vascular function than procyanidins and circulating PVLs, while reduction on total cholesterol was related to the procyanidin content rather than to the monomers.

Nevertheless, plasma 5-(3’-hydroxyphenyl)-γ-valerolactone-4’-sulfate was found positively correlated with an increase in FMD 4h and 8h after cranberry juice consumption (containing up to 1910 mg of total (poly)phenols), compared to control. Some improvement in blood pressure after

cranberry juice intake was observed compared to the baseline in the intervention group, but no significant results were observed when comparing changes in blood pressure, pulse wave velocity or augmentation index to the control [127]. Contrasting results on endothelial function markers have been found when considering the health status of the participants and the chronic/acute consumption of the supplement. While a crossover RCT involving volunteers with coronary artery disease showed a decrease in carotid−femoral pulse wave velocity after consumption of cranberry juice for 4 weeks [128], these results were not confirmed when healthy men consumed a single supplementation of a blueberry drink [129].

Both cocoa- and berry-containing products have been found to significantly improve platelet function by means of the use of a platelet function analyser. A significant increase in adenosine-5'-diphosphate/collagen closure time has been found i) after 2h and 6h consumption of dark chocolate, by healthy volunteers [130], ii) after 8-week supplementation of 100 g berries, by subjects with metabolic syndrome [131] and iii) after 4h dark chocolate consumption, by healthy male volunteers [132]. Particularly in this last study, a focus on the putative role of SREMs and PVLs in increasing adenosine-5'-diphosphate/collagen closure time was paid, revealing a significantly positive correlation with single and total SREMs but not with PVLs.

Several molecular mechanisms underlying the biological effects of flavan-3-ols and their derivatives have been hypothesized. Flavan-3-ols have been found to increase nitric oxide availability through the activation of several different phosphorylation cascades, leading to the activation of the endothelial nitric oxide synthase, the major enzyme responsible for the production of nitric oxide [133]. An intriguing hypothesis by Rodriguez-Mateos and colleagues claims that some phenolic metabolites have a structure similar to the Nicotinamide Adenine Dinucleotide Phosphate (NADPH) oxidase inhibitor apocynin (acetovanillone) [125], which has been showed to act as a potent NAPDH oxidase inhibitor in *in vitro* models of endothelial cells. A decrease in NADPH oxidase activity has been proposed as the primary mechanism mediating the short-term improvements in FMD observed after consumption of flavan-3-ol-rich foods. Among these, Bernatoniene and collaborators focused on sources of catechins, i.e. tea, cocoa and wine, and their potential beneficial effects on cardiovascular diseases [134]. Catechins from green tea decreased blood pressure and reduced the risk of stroke and coronary heart disease, alleviated conditions associated with vascular dysfunction, including vascular inflammation and smooth muscle cell proliferation, blood platelet aggregation, lipoprotein oxidation, altered lipid profile, and vascular reactivity. Epigallocatechin gallate enhanced expression of p53, p21, and NF-kB, induced the apoptosis of vascular smooth muscle cells and

prevented the development of atherosclerosis. Besides, catechins reduced the accumulation of cholesterol and its oxidation products in artery walls *in vivo*, thus improving blood circulation [134].

A third potential beneficial action of flavan-3-ols and derived metabolites on human health worth mentioning is the one against UTIs. Urinary tract infections are caused by bacteria and certain fungi, with uropathogenic *Escherichia coli* (UPEC) as the most common cause. UTIs can be uncomplicated (e.g., cystitis and pyelonephritis affecting the lower and upper urinary tract, respectively, but without structural abnormalities in the tract) or complicated, when able to compromise the affected parts (e.g., obstruction of the urinary tract) [135]. UTIs are among the most pervasive bacterial infections and represent a significant economic and medical burden worldwide. They account for several millions of outpatient hospital visits and millions of emergency room visits, with a vast annual direct cost for national healthcare systems [136]. Cranberry consumption has been reported to be effective in decreasing the occurrence and severity of UTI in women [137]. The preventive effect of cranberry on UTIs has been attributed to its bioactives and/or their metabolites during the phase of bacterial adherence to the uroepithelial cells, disabling or inhibiting the adhesion of UPEC and preventing bacterial colonization and progression of UTI. Importantly, cranberries are supposed to reduce UTI-related symptoms by suppressing inflammatory cascades as an immunologic response to bacterial invasion [138]. Another interesting theory, which considers the large intestine a reservoir for uropathogenic bacteria, claims that A-type PACs might specifically decrease the transient intestinal colonisation by UPEC, consequently reducing the risk of UTI incidence [138]. Despite several studies pointed out that the beneficial effects towards UTI-related diseases and recurrence might be driven from type-A PACs [139], Peron and colleagues pointed out that A-type PACs might not be the sole responsible for the anti-adhesive activity, as these compounds present in the original food have been detected in very low concentrations (< nM) in urine samples of women who consumed cranberry products [140]. There is still a debate in literature concerning which is or are the compound(s) present in urine after cranberry consumption that prevent(s)/inhibit(s) the effective adherence of uropathogens to uroepithelial cells. Three recent *in vitro* investigations focused on the major gut microbial derivatives from cranberry flavan-3-ols and their anti-adhesive activity against UPEC on human bladder cells. Gonzalez de Llano and collaborators found that catechol, benzoic acid, vanillic acid, phenylacetic acid and 3,4-dihydroxyphenylacetic acid showed anti-adhesive activity against UPEC on human bladder T24 cells in a concentration-dependent manner from 100–500 µM, whereas procyanidin A2, widely reported as an inhibitor of UPEC adherence on uroepithelium, was only statistically significant at 500 µM [138]. Mena and colleagues investigated the

role of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone derivatives on the inhibition of UPEC adherence to bladder epithelial cells, showing that all the compounds significantly decreased the UPEC adherence at 100 µM, up to 50 µM for 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-sulfate [141]. Gonzalez de Llano and collaborators found that 3',4'-dihydroxyphenylacetic acid (but not its methyl derivative) inhibited UPEC adherence to both Gram-positive and Gram-negative bacteria up to 250 µM [142].
Despite the *in vitro* findings support the biological effects of flavan-3-ol gut microbial derivatives, *in vivo* studies have reported conflicting results. One of the primary limits worth to be considered is the wide inter-individual variability observed in the human urinary/plasmatic phenolic metabolite profile. As previously described, inter-individual differences in the production of flavan-3-ol metabolites might lead to the existence of urinary metabotypes, which group subjects with similar capacity to metabolize flavan-3-ols [143]. It remains to be proven whether the intake of cranberry polyphenols would also cluster subjects according to their phenolic metabolite profile in physiological fluids (urine, plasma, and feces) and whether these hypothetical metabotypes could be one of the factors putatively associated with the considerable variability observed in the efficacy of cranberry products.

The existing literature is consistent with the evidence that flavan-3-ols are bioactive compounds with beneficial effects on health. Despite this, discrepancies of results have been highlighted when encouraging results obtained using observational and *in vivo* animal studies are compared with more wary results from *in vivo* human interventions studies [144]. Most of the reasons fall in the profound structural modifications that these compounds undergo once in contact with the human body, and in the inter-individual variability that defines every single organism as "unique". Recently, essential viewpoints and position papers have been published to fill this evidence gap. Ottaviani and colleagues produced an interesting handbook of rules worth to be considered when studying the health effects of flavan-3-ols, among which i) an in-depth characterization of the food/supplement flavan-3-ol profile, ii) the study of ADME of the compounds, with the identification of both flavan-3-ol conjugates as well as gut microbial derivatives, iii) the choice of outcomes/endpoints and of adequate statistical methods [145]. A recent viewpoint of the European COST-POSITIVe network (https://www6.inra.fr/cost-positive) has pointed out that the failing results in *in vivo* studies might be linked to the high variability in ADME among volunteers, particularly because of the different gut microbiota metabolism or the presence of polymorphisms in phase I and II metabolism enzymes [146]. To fully understand the effects of flavan-3-ol intake on human health at individual level, further

research is needed. This research should be undeniably conducted taking into account the evidence collected to date and considering real life situations of flavan-3-ol intake [147].

## 1.3. Metabolomics as a tool in nutrition research

Metabolomics is the field of "omics" sciences that has grown more rapidly in the recent years and it focuses on the analysis of the metabolome, namely the set of small molecules (<1500 Da) derived from cellular metabolism in complex specimens, as biofluids, tissues and cells [148,149]. It can also be described as the application of high throughput analytical technologies, such as liquid chromatography or gas chromatography coupled to mass spectrometry (LC-MS or GC-MS) and nuclear magnetic resonance spectroscopy (NMR), aimed at characterizing the metabolome [150]. Two separate approaches have been distinguished: targeted and untargeted metabolomics. Targeted metabolomics is hypothesis-driven and looks for specific and previously defined metabolites in a sample, often related to a specific metabolic pathway (e.g., fatty acids, acylcarnitines, amino acids or particular classes/subclasses of phytochemicals) [151,152]. Untargeted metabolomics, instead, is the analysis of all the measurable metabolites in a sample, including unknown metabolites, and it can be useful to generate new hypotheses [153]. Whether targeted or untargeted, metabolomics generates large and complex datasets that necessitate advanced statistical and bioinformatics tools for their interpretation. In addition to univariate statistical methods, multivariate analyses are required, including principal component analysis (PCA), clustering methods and partial least-squares discriminant analysis (PLS-DA) [150,153].

The potential of metabolomics applied to nutrition research is wide (Figure 9), allowing the study of the ADME of dietary compounds, the discovery and validation of food intake biomarkers (and, through them, better measuring food consumption), the study of biological responses to dietary interventions and of diet-related disease, the identification of novel biomarkers of disease, the determination of metabotypes (metabolic phenotypes) [149,153,154]. In particular, three metabolomics applications - study of the ADME of dietary compounds, determination of metabotypes, study of the metabolic changes following dietary interventions - will be dealt with more in detail in the following subparagraphs 1.3.1, 1.3.2, 1.3.3. These applications are those adopted in the PhD project reported in this thesis.


**Figure 9.** Applications of metabolomics in the field of nutrition.

### 1.3.1. Evaluation of dietary compounds bioavailability

In bioavailability studies, targeted metabolomics analysis of plasma and/or urine samples can provide information on absorption, metabolism and excretion of dietary compounds [88,155,156]. This information is key to figure out in which forms, what concentrations and how long compounds from the diet circulate within the human body and are available to target tissues and organs, where they can exert biological effects. To this end, specific analytical methods are developed that allows absolute or semi quantification of selected metabolites, on the basis of reference standards availability. First of all, metabolites of interest are extracted and purified from the biological samples. Then, they are separated by liquid or gas chromatography prior to mass spectrometry detection [157]. Also NMR can be employed, but this technique has a lower sensitivity than MS, which narrows its coverage of the food metabolome to predominant nutrients, sugars and metabolites present in the range of millimolar to micromolar concentrations [150], whereas derivative metabolites from bioactive compounds are usually present in the range of micromolar to nanomolar concentrations. Today also, technological advancements contribute to increasing the number of compounds that can be quantified simultaneously in a single analysis. For example, using selected retention times with ultra-performance liquid chromatography and multiple-reaction monitoring transitions with triple-quadrupole instruments, hundreds of metabolites can be monitored in a single run [157].

In addition, this information can help in the validation process for candidate food intake biomarkers, as defined by the FoodBAll (Food Biomarkers Alliance) consortium (https://foodmetabolome.org/) [158]. In fact, it provides data on dose-response and time-response, two of the eight criteria comprising the developed validation procedure.

## 1.3.2. Metabotyping

The concept of metabotyping, or metabolic phenotyping, describes the classification of individuals into homogeneous subgroups, the so-called metabotypes or metabolic phenotypes, on the basis of their metabolic or phenotypic characteristics [159] (Figure 10). This concept implies that individuals within a metabotype show a high metabolic similarity than those grouped in different metabotypes. Metabolomics has emerged as a tool for determining metabotypes, as metabolomic profiles or combinations of specific metabolites can be used for stratifying individuals into subgroups [105]. In nutrition research, metabotyping was initially adopted to distinguish individuals with or without diet-related diseases, but it has rapidly developed to identify those subjects at metabolic risk and to evaluate different responses to dietary interventions [105,159]. With an increased interest in the observed inter-individual variation in response to dietary interventions, metabotyping represents an unbiased method of identifying these differential responses. The ultimate goal would be to harness the metabotyping approach for delivering personalised nutrition, i.e., targeted dietary advice to individuals based on which metabotype they belong to [105] (Figure 10).

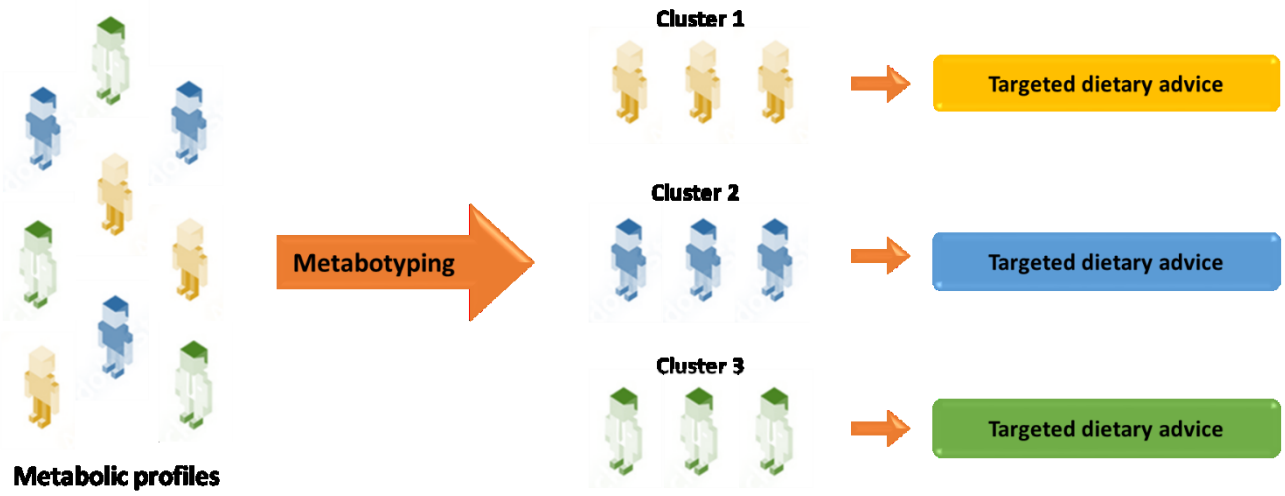



**Figure 10.** Overview of the concept of metabotyping. Individuals are grouped into metabotypes based on their metabolic phenotypes and dietary advice is delivered tailored to the metabotype.

In this sense, considering the inter-individual variability characterizing the ADME of plant food bioactives and the biological responses to their consumption, subjects presenting similar metabolic capacity towards plant food bioactives can be grouped into metabotypes. As emphasised by the COST-POSITIVe network (https://www6.inra.fr/cost-positive), the stratification of individuals into different metabotypes may represent a potential strategy to identify population subgroups that could gain particular benefit from the consumption of specific plant food bioactives [66]. To this end,

both targeted and untargeted metabolomics appear to be valid approaches. Targeted metabolomics can help in grouping subjects based on their ability to produce specific metabolites. The metabolite production can be either dichotomic (production vs. non production of specific metabolites), as in the case of gut-microbiota mediated metabolites from ellagitannins and isoflavones [49,160], or characterized by different proportions of certain metabolites belonging to a catabolic pathway. An untargeted approach can instead provide a comprehensive phenotype after exposure to a diversity of food phytochemicals [66]. Metabotyping can be used both *ex post* to stratify individuals according to their internal exposure to phytochemical metabolites and also to prospectively establish correlations among metabolites signatures and health-related outcomes [66]. Well-described metabotypes in literature are those of gut-derived metabolites from ellagitannins and isoflavones, based on individuals' capacity to produce or not, respectively, certain urolithins and equol, as mentioned before [49,160]. In human intervention studies, clustering individuals according to their urolithin or equol metabotype has proven that only subjects characterized by a certain metabotype show improvement of cardiometabolic markers following ellagitannin or isoflavone intake [64,65]. However, for most families of plant food bioactives, a lot of work still needs to be carried out in defining reliable metabotypes. First of all, shared statistical procedures for metabotype definition are lacking, and there is an on-going discussion on which statistical methods should be used to obtain the best separation among subgroups [159]. Then, the key determinants responsible for the different metabotypes, and, in particular, the role of genetics and the microbiome still have to be deciphered. This knowledge will be crucial to further develop a more robust metabotyping approach. As a final aim, metabotyping may be useful in refining the recommendations of plant foods rich in specific bioactives, and also in providing science-based tailored dietary advice for their consumption to individuals [66].

### 1.3.3. Study of the metabolic changes following dietary interventions

The food we eat impacts on our metabolic pathways. Metabolomics is actually well-suited to examine the influence of certain diets, food items or specific dietary components on endogenous metabolic routes, with the ultimate goal of understanding the effects of such diets, foods or dietary components on human health, and unravelling the underlying mechanisms [161]. Both targeted and untargeted approaches can serve the purpose. On the one hand, targeted metabolomics can monitor specific health biomarkers, or a group of pre-defined metabolites related to a specific metabolic pathway, and quantify their levels in biological fluids after a dietary intervention, perhaps together with selected diet-related chemicals or metabolites. For instance, Chiuve and colleagues

studied the influence of two specific diet-derived molecules, betaine and choline, on reducing plasma levels of homocysteine, a well-established risk factor for cardiovascular disease and many other chronic diseases [162]. Likewise, Zheng and collaborators used a targeted approach to measure plasma amino acid concentrations in response to weight-loss diet interventions in two independent studies, and revealed significant decreases in branched chain and aromatic amino acids, both previously associated with insulin resistance and type 2 diabetes [163]. This study, applying metabolomics to a pre- and post-dietary intervention scenario, provided insights into potential mechanisms linked to the health benefits of certain dietary regimens. On the other hand, untargeted metabolomics comprehensively monitors all the measurable metabolites in a sample, both of exogenous and endogenous origin, including unknown metabolites. This allows the detection of both metabolites specifically associated with the consumption of a given food or diet and metabolites from endogenous metabolic pathways, whose levels may be modulated by the intake of that particular food or diet. As an example, the untargeted GC-MS metabolomics analysis of plasma samples from 145 individuals who followed either the New Nordic Diet (NND) or the Average Danish Diet (ADD) for a six-month period revealed significant differences in the blood metabolome between the the two diets [164]. The metabolites found to be at higher levels in the NND individuals included those directly reflecting the higher intake of fish, whole grain and vegetables, including crucifers, and other metabolites mainly reflecting the impact of NND on energy metabolism to increase gluconeogenesis and ketosis, leading to hypothesise mechanisms through which the NND elicits its beneficial effects. Actually, one of the main challenges of this approach lies in the mapping of observed changes in metabolite levels to metabolic routes and processes [153]. This task is presently supported by useful tools (such as MetaboAnalyst [165] or ChemRICH [166]), that, however, still need further improvement. In addition, since food is also made of components that are transformed into products partly identical to some endogenous metabolites, sometimes it is hard to distinguish the diet induced changes in the metabolome impacting health from changes deriving from the food itself [167].

Collaboration between scientists from different disciplines is required to obtain meaningful interpretation of the data [161], and the identification of unknown metabolites still remains a major bottleneck in metabolomics, with the resulting consequence that biological interpretation has to be performed on a small number of identified metabolites [161,167]. Indeed, endogenous metabolites are currently better covered in reference databases, like the Human Metabolome Database (HMDB) [168],

than food-derived metabolites, which are characterized by a large chemical complexity and diversity, and are still widely undocumented.

The great efforts the nutrimetabolomics community is carrying out to update databases, improve software tools, and promote sharing of data and resources should significantly contribute to overcome the current limitations associated to metabolomics experiments and to help deciphering the complex interactions between dietary components and our physiology, potentially responsible of effects, either beneficial or detrimental, on human health [150,167].

# References

1 World Health Organization, World health statistics 2019: monitoring health for the SDGs, sustainable development goals, https://apps.who.int/iris/handle/10665/324835, (accessed 20 July 2021).

2 W. Willett, J. Rockström, B. Loken, M. Springmann, T. Lang, S. Vermeulen, T. Garnett, D. Tilman, F. DeClerck, A. Wood, M. Jonell, M. Clark, L. J. Gordon, J. Fanzo, C. Hawkes, R. Zurayk, J. A. Rivera, W. De Vries, L. Majele Sibanda, A. Afshin, A. Chaudhary, M. Herrero, R. Agustina, F. Branca, A. Lartey, S. Fan, B. Crona, E. Fox, V. Bignet, M. Troell, T. Lindahl, S. Singh, S. E. Cornell, K. Srinath Reddy, S. Narain, S. Nishtar and C. J. L. Murray, Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems, *Lancet*, 2019, **393**, 447–492.

3 A. Afshin, P. J. Sur, K. A. Fay, L. Cornaby, G. Ferrara, J. S. Salama, E. C. Mullany, K. H. Abate, C. Abbafati, Z. Abebe, M. Afarideh, A. Aggarwal, S. Agrawal, T. Akinyemiju, F. Alahdab, U. Bacha, V. F. Bachman, H. Badali, A. Badawi, I. M. Bensenor, E. Bernabe, S. K. K. Biadgilign, S. H. Biryukov, L. E. Cahill, J. J. Carrero, K. M. Cercy, L. Dandona, R. Dandona, A. K. Dang, M. G. Degefa, M. El Sayed Zaki, A. Esteghamati, S. Esteghamati, J. Fanzo, C. S. e. S. Farinha, M. S. Farvid, F. Farzadfar, V. L. Feigin, J. C. Fernandes, L. S. Flor, N. A. Foigt, M. H. Forouzanfar, M. Ganji, J. M. Geleijnse, R. F. Gillum, A. C. Goulart, G. Grosso, I. Guessous, S. Hamidi, G. J. Hankey, S. Harikrishnan, H. Y. Hassen, S. I. Hay, C. L. Hoang, M. Horino, F. Islami, M. D. Jackson, S. L. James, L. Johansson, J. B. Jonas, A. Kasaeian, Y. S. Khader, I. A. Khalil, Y. H. Khang, R. W. Kimokoti, Y. Kokubo, G. A. Kumar, T. Lallukka, A. D. Lopez, S. Lorkowski, P. A. Lotufo, R. Lozano, R. Malekzadeh, W. März, T. Meier, Y. A. Melaku, W. Mendoza, G. B. M. Mensink, R. Micha, T. R. Miller, M. Mirarefin, V. Mohan, A. H. Mokdad, D. Mozaffarian, G. Nagel, M. Naghavi, C. T. Nguyen, M. R. Nixon, K. L. Ong, D. M. Pereira, H. Poustchi, M. Qorbani, R. K. Rai, C. Razo-García, C. D. Rehm, J. A. Rivera, S. Rodríguez-Ramírez, G. Roshandel, G. A. Roth, J. Sanabria, T. G. Sánchez-Pimienta, B. Sartorius, J. Schmidhuber, A. E. Schutte, S. G. Sepanlou, M. J. Shin, R. J. D. Sorensen, M. Springmann, L. Szponar, A. L. Thorne-Lyman, A. G. Thrift, M. Touvier, B. X. Tran, S. Tyrovolas, K. N. Ukwaja, I. Ullah, O. A. Uthman, M. Vaezghasemi, T. J. Vasankari, S. E. Vollset, T. Vos, G. T. Vu, L. G. Vu, E. Weiderpass, A. Werdecker, T. Wijeratne, W. C. Willett, J. H. Wu, G. Xu, N. Yonemoto, C. Yu and C. J. L. Murray, Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017,

*Lancet*, 2019, **393**, 1958–1972.

4 World Health Organization, *GLOBAL ACTION PLAN FOR THE PREVENTION AND CONTROL OF NONCOMMUNICABLE DISEASES 2013-2020*, 2013.

5 B. A. Swinburn, V. I. Kraak, S. Allender, V. J. Atkins, P. I. Baker, J. R. Bogard, H. Brinsden, A. Calvillo, O. De Schutter, R. Devarajan, M. Ezzati, S. Friel, S. Goenka, R. A. Hammond, G. Hastings, C. Hawkes, M. Herrero, P. S. Hovmand, M. Howden, L. M. Jaacks, A. B. Kapetanaki, M. Kasman, H. V. Kuhnlein, S. K. Kumanyika, B. Larijani, T. Lobstein, M. W. Long, V. K. R. Matsudo, S. D. H. Mills, G. Morgan, A. Morshed, P. M. Nece, A. Pan, D. W. Patterson, G. Sacks, M. Shekar, G. L. Simmons, W. Smit, A. Tootee, S. Vandevijvere, W. E. Waterlander, L. Wolfenden and W. H. Dietz, The Global Syndemic of Obesity, Undernutrition, and Climate Change: The Lancet Commission report, *Lancet*, 2019, **393**, 791–846.

6 D. Aune, E. Giovannucci, P. Boffetta, L. T. Fadnes, N. Keum, T. Norat, D. C. Greenwood, E. Riboli, L. J. Vatten and S. Tonstad, Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality—a systematic review and dose-response meta-analysis of prospective studies, *Int. J. Epidemiol.*, 2017, **46**, 1029–1056.

7 R. H. Liu, Health-promoting components of fruits and vegetables in the diet, *Adv. Nutr.*, 2013, **4**, 384S-392S.

8 R. Zamora-Ros, V. Knaze, J. A. Rothwell, B. Hémon, A. Moskal, K. Overvad, A. Tjønneland, C. Kyrø, G. Fagherazzi, M.-C. Boutron-Ruault, M. Touillaud, V. Katzke, T. Kühn, H. Boeing, J. Förster, A. Trichopoulou, E. Valanou, E. Peppa, D. Palli, C. Agnoli, F. Ricceri, R. Tumino, M. S. de Magistris, P. H. M. Peeters, H. B. Bueno-de-Mesquita, D. Engeset, G. Skeie, A. Hjartåker, V. Menéndez, A. Agudo, E. Molina-Montes, J. M. Huerta, A. Barricarte, P. Amiano, E. Sonestedt, L. M. Nilsson, R. Landberg, T. J. Key, K.-T. Khaw, N. J. Wareham, Y. Lu, N. Slimani, I. Romieu, E. Riboli and A. Scalbert, Dietary polyphenol intake in Europe: the European Prospective Investigation into Cancer and Nutrition (EPIC) study, *Eur. J. Nutr.*, 2016, **55**, 1359–1375.

9 D. Del Rio, A. Rodriguez-Mateos, J. P. E. Spencer, M. Tognolini, G. Borges and A. Crozier, Dietary (Poly)phenolics in Human Health: Structures, Bioavailability, and Evidence of Protective Effects Against Chronic Diseases, *Antioxid. Redox Signal.*, 2013, **18**, 1818–1892.

10 C. G. Fraga, K. D. Croft, D. O. Kennedy and F. A. Tomás-Barberán, The effects of polyphenols and other bioactives on human health, *Food Funct.*, 2019, **10**, 514–528.

11 M. Rodriguez-Concepcion, J. Avalos, M. L. Bonet, A. Boronat, L. Gomez-Gomez, D. Hornero-Mendez, M. C. Limon, A. J. Meléndez-Martínez, B. Olmedilla-Alonso, A. Palou, J. Ribot, M. J. Rodrigo, L. Zacarias and C. Zhu, A global perspective on carotenoids: Metabolism, biotechnology, and benefits for nutrition and health, *Prog. Lipid Res.*, 2018, **70**, 62–93.

12 R. A. Moreau, L. Nyström, B. D. Whitaker, J. K. Winkler-Moser, D. J. Baer, S. K. Gebauer and K. B. Hicks, Phytosterols and their derivatives: Structural diversity, distribution, metabolism, analysis, and health-promoting uses, *Prog. Lipid Res.*, 2018, **70**, 35–61.

13 A. Abellán, R. Domínguez-Perles, D. A. Moreno and C. García-Viguera, Sorting out the value of cruciferous sprouts as sources of bioactive compounds for nutrition and health, *Nutrients*, 2019, **11**, 1–22.

14 S. Petropoulos, F. Di Gioia and G. Ntatsi, Vegetable Organosulfur Compounds and their Health Promoting Effects, *Curr. Pharm. Des.*, 2017, **23**, 2850–2875.

15 K. Kempf, H. Kolb, B. Gärtner, G. Bytof, H. Stiebitz, I. Lantz, R. Lang, T. Hofmann and S. Martin, Cardiometabolic effects of two coffee blends differing in content for major constituents in overweight adults: a randomized controlled trial, *Eur. J. Nutr.*, 2015, **54**, 845–854.

16 R. Landberg, M. Marklund, A. Kamal-Eldin and P. Åman, An update on alkylresorcinols - Occurrence, bioavailability, bioactivity and utility as biomarkers, *J. Funct. Foods*, 2014, **7**, 77–89.

17 A. F. G. Cicero, A. Colletti, G. Bajraktari, O. Descamps, D. M. Djuric, M. Ezhov, Z. Fras, N. Katsiki, M. Langlois, G. Latkovskis, D. B. Panagiotakos, G. Paragh, D. P. Mikhailidis, O. Mitchenko, B. Paulweber, D. Pella, C. Pitsavos, Ž. Reiner, K. K. Ray, M. Rizzo, A. Sahebkar, M. C. Serban, L. S. Sperling, P. P. Toth, D. Vinereanu, M. Vrablík, N. D. Wong and M. Banach, Lipid lowering nutraceuticals in clinical practice: Position paper from an International Lipid Expert Panel, *Arch. Med. Sci.*, 2017, **13**, 965–1005.

18 R. Estruch, E. Ros, J. Salas-Salvadó, M.-I. Covas, D. Corella, F. Arós, E. Gómez-Gracia, V. Ruiz-Gutiérrez, M. Fiol, J. Lapetra, R. M. Lamuela-Raventos, L. Serra-Majem, X. Pintó, J. Basora, M. A. Muñoz, J. V. Sorlí, J. A. Martínez, M. Fitó, A. Gea, M. A. Hernán and M. A. Martínez-González, Primary Prevention of Cardiovascular Disease with a Mediterranean Diet Supplemented with Extra-Virgin Olive Oil or Nuts, *N. Engl. J. Med.*, 2018, **378**, e34.

19 M. Uusitupa, K. Hermansen, M. J. Savolainen, U. Schwab, M. Kolehmainen, L. Brader, L. S. Mortensen, L. Cloetens, A. Johansson-Persson, G. Önning, M. Landin-Olsson, K. -H. Herzig, J. Hukkanen, F. Rosqvist, D. Iggman, J. Paananen, K. J. Pulkki, M. Siloaho, L. Dragsted, T. Barri, K. Overvad, K. E. Bach Knudsen, M. S. Hedemann, P. Arner, I. Dahlman, G. I. A. Borge, P. Baardseth, S. M. Ulven, I. Gunnarsdottir, S. Jónsdóttir, I. Thorsdottir, M. Orešič, K. S. Poutanen, U. Risérus and B. Åkesson, Effects of an isocaloric healthy Nordic diet on insulin sensitivity, lipid profile and inflammation markers in metabolic syndrome – a randomized study (SYSDIET), *J. Intern. Med.*, 2013, **274**, 52–66.

20 M. A. Beydoun, X. Chen, K. Jha, H. A. Beydoun, A. B. Zonderman and J. A. Canas, Carotenoids, vitamin A, and their association with the metabolic syndrome: a systematic review and meta-analysis, *Nutr. Rev.*, 2019, **77**, 32–45.

21 I. Demonty, R. T. Ras, H. C. M. van der Knaap, G. S. M. J. E. Duchateau, L. Meijer, P. L. Zock, J. M. Geleijnse and E. A. Trautwein, Continuous Dose-Response Relationship of the LDL-Cholesterol–Lowering Effect of Phytosterol Intake, *J. Nutr.*, 2009, **139**, 271–284.

22 M.-T. García-Conesa, K. Chambers, E. Combet, P. Pinto, M. Garcia-Aloy, C. Andrés-Lacueva, S. de Pascual-Teresa, P. Mena, A. Konic Ristic, W. Hollands, P. Kroon, A. Rodríguez-Mateos, G. Istas, C. Kontogiorgis, D. Rai, E. Gibney, C. Morand, J. Espín and A. González-Sarrías, Meta-Analysis of the Effects of Foods and Derived Products Containing Ellagitannins and Anthocyanins on Cardiometabolic Biomarkers: Analysis of Factors Influencing Variability of the Individual Responses, *Int. J. Mol. Sci.*, 2018, **19**, 694.

23 A. González-Sarrías, E. Combet, P. Pinto, P. Mena, M. Dall'Asta, M. Garcia-Aloy, A. Rodríguez-Mateos, E. R. Gibney, J. Dumont, M. Massaro, J. Sánchez-Meca, C. Morand and M.-T. García-Conesa, A Systematic Review and Meta-Analysis of the Effects of Flavanol-Containing Tea, Cocoa and Apple Products on Body Composition and Blood Lipids: Exploring the Factors Responsible for Variability in Their Efficacy, *Nutrients*, 2017, **9**, 746.

24 L. Hooper, C. Kay, A. Abdelhamid, P. A. Kroon, J. S. Cohn, E. B. Rimm and A. Cassidy, Effects of chocolate, cocoa, and flavan-3-ols on cardiovascular health: a systematic review and meta-analysis of randomized trials, *Am. J. Clin. Nutr.*, 2012, **95**, 740–751.

25 E. Trautwein, M. Vermeer, H. Hiemstra and R. Ras, LDL-Cholesterol Lowering of Plant Sterols and Stanols—Which Factors Influence Their Efficacy?, *Nutrients*, 2018, **10**, 1262.

26 L. Blekkenhorst, M. Sim, C. Bondonno, N. Bondonno, N. Ward, R. Prince, A. Devine, J. Lewis and J. Hodgson, Cardiovascular Health Benefits of Specific Vegetable Types: A Narrative Review, *Nutrients*, 2018, **10**, 595.

27 D. Milenkovic, C. Morand, A. Cassidy, A. Konic-Ristic, F. Tomás-Barberán, J. M. Ordovas, P. Kroon, R. De Caterina and A. Rodriguez-Mateos, *Adv. Nutr.*, 2017, **8**, 558–570.

28 C. Manach, D. Milenkovic, T. Van de Wiele, A. Rodriguez-Mateos, B. de Roos, M. T. Garcia-Conesa, R. Landberg, E. R. Gibney, M. Heinonen, F. Tomás-Barberán and C. Morand, Addressing the inter-individual variation in response to consumption of plant food bioactives: Towards a better understanding of their role in healthy aging and cardiometabolic risk reduction, *Mol. Nutr. Food Res.*, 2017, **61**, 1–16.

29 E. R. Gibney, D. Milenkovic, E. Combet, T. Ruskovska, A. Greyling, A. González-Sarrías, B. de Roos, F. Tomás-Barberán, C. Morand and A. Rodriguez-Mateos, *Eur. J. Nutr.*, 2019, **58**, 37–47.

30 A. A. Kolodziejczyk, D. Zheng and E. Elinav, Diet–microbiota interactions and personalized nutrition, *Nat. Rev. Microbiol.*, 2019, **17**, 742–753.

31 A. L. Barabási, G. Menichetti and J. Loscalzo, The unmapped chemical complexity of our diet, *Nat. Food*, 2020, **1**, 33–37.

32 H. C. Descamps, B. Herrmann, D. Wiredu and C. A. Thaiss, The path toward using microbial metabolites as therapies, *EBioMedicine*, 2019, **44**, 747–754.

33 A. W. C. Man, H. Li and N. Xia, Resveratrol and the interaction between gut microbiota and arterial remodelling, *Nutrients*, 2020, **12**, 119.

34 P. Vernocchi, F. Del Chierico and L. Putignani, Gut microbiota metabolism and interaction with food components, *Int. J. Mol. Sci.*, 2020, **21**, 3688.

35 M. do C. G. Peluzio, J. A. Martinez and F. I. Milagro, Postbiotics: Metabolites and mechanisms involved in microbiota-host interactions, *Trends Food Sci. Technol.*, 2021, **108**, 11–26.

36 A. Crozier, I. B. Jaganath and M. N. Clifford, Dietary phenolics: Chemistry, bioavailability and effects on health, *Nat. Prod. Rep.*, 2009, **26**, 1001–1043.

37 F. A. Tomas-Barberan, A. González-Sarrías and R. García-Villalba, *Dietary Polyphenols*, Wiley, 2020.

38 C. Manach, A. Scalbert, C. Morand, C. Rémésy and L. Jiménez, Polyphenols: food sources and bioavailability, *Am. J. Clin. Nutr.*, 2004, **79**, 727–747.

39 R. Zamora-Ros, N. G. Forouhi, S. J. Sharp, C. A. González, B. Buijsse, M. Guevara, Y. T. Van Der Schouw, P. Amiano, H. Boeing, L. Bredsdorff, F. Clavel-Chapelon, G. Fagherazzi, E. J. Feskens, P. W. Franks, S. Grioni, V. Katzke, T. J. Key, K. T. Khaw, T. Kühn, G. Masala, A. Mattiello, E. Molina-Montes, P. M. Nilsson, K. Overvad, F. Perquier, J. Ramón Quirós, I. Romieu, C. Sacerdote, A. Scalbert, M. Schulze, N. Slimani, A. M. W. Spijkerman, A. Tjonneland, M. J. Tormo, R. Tumino, D. L. Van Der A., C. Langenberg, E. Riboli and N. J. Wareham, The association between dietary flavonoid and lignan intakes and incident type 2 diabetes in european populations: The EPIC-InterAct study, *Diabetes Care*, 2013, **36**, 3961–3970.

40 I. Zanotti, M. Dall'Asta, P. Mena, L. Mele, R. Bruni, S. Ray and D. Del Rio, Atheroprotective effects of (poly) phenols: A focus on cell cholesterol metabolism, *Food Funct.*, 2015, **6**, 13–31.

41 F. Ferrucci, G. Gaudiosi, A. Hirn and R. Nicolich, Ionian Basin and Calabria Arc: Some new elements from DSS data, *Tectonophysics*, 1991, **195**, 411–419.

42 N. Ziauddeen, A. Rosi, D. Del Rio, B. Amoutzopoulos, S. Nicholson, P. Page, F. Scazzina, F. Brighenti, S. Ray and P. Mena, Dietary intake of (poly)phenols in children and adults: cross-sectional analysis of UK National Diet and Nutrition Survey Rolling Programme (2008–2014), *Eur. J. Nutr.*, 2019, **58**, 3183–3198.

43 G. Williamson, C. D. Kay and A. Crozier, The Bioavailability, Transport, and Bioactivity of Dietary Flavonoids: A Review from a Historical Perspective, *Compr. Rev. Food Sci. Food Saf.*, 2018, **17**, 1054–1112.

44 A. J. Day, F. J. Cañada, J. C. Díaz, P. A. Kroon, R. Mclauchlan, C. B. Faulds, G. W. Plumb, M. R. A. Morgan and G. Williamson, Dietary flavonoid and isoflavone glycosides are hydrolysed by the lactase site of lactase phlorizin hydrolase, *FEBS Lett.*, 2000, **468**, 166–170.

45 J. M. Gee, M. S. DuPont, A. J. Day, G. W. Plumb, G. Williamson and I. T. Johnson, Intestinal Transport of Quercetin Glycosides in Rats Involves Both Deglycosylation and Interaction with the Hexose Transport Pathway, *J. Nutr.*, 2000, **130**, 2765–2771.

46 C. Garcia-Muñoz and F. Vaillant, Metabolic Fate of Ellagitannins: Implications for Health, and Research Perspectives for Innovative Functional Foods, *Crit. Rev. Food Sci. Nutr.*, 2014, **54**,

1584–1598.

47 A. Crozier, Absorption, metabolism, and excretion of (–)-epicatechin in humans: an evaluation of recent findings, *Am. J. Clin. Nutr.*, 2013, **98**, 861–862.

48 Nutrient bioavailability: Getting the most out of food (EUFIC), https://www.eufic.org/en/food-today/article/nutrient-bioavailability-getting-the-most-out-of-food, (accessed 23 July 2021).

49 F. A. Tomás-Barberán, R. García-Villalba, A. González-Sarrías, M. V. Selma and J. C. Espín, Ellagic Acid Metabolism by Human Gut Microbiota: Consistent Observation of Three Urolithin Phenotypes in Intervention Trials, Independent of Food Source, Age, and Health Status, *J. Agric. Food Chem.*, 2014, **62**, 6535–6538.

50 K. D. R. Setchell, N. M. Brown, S. Summer, E. C. King, J. E. Heubi, S. Cole, T. Guy and B. Hokin, Dietary Factors Influence Production of the Soy Isoflavone Metabolite S-(-)Equol in Healthy Adults, *J. Nutr.*, 2013, **143**, 1950–1958.

51 A. Rodriguez-Mateos, D. Vauzour, C. G. Krueger, D. Shanmuganayagam, J. Reed, L. Calani, P. Mena, D. Del Rio and A. Crozier, Bioavailability, bioactivity and impact on health of dietary flavonoids and related compounds: an update, *Arch. Toxicol.*, 2014, **88**, 1803–1853.

52 G. Pereira-Caro, G. Borges, J. van der Hooft, M. N. Clifford, D. Del Rio, M. E. J. Lean, S. A. Roberts, M. B. Kellerhals and A. Crozier, Orange juice (poly)phenols are highly bioavailable in humans, *Am. J. Clin. Nutr.*, 2014, **100**, 1378–1384.

53 A. Tresserra-Rimbau, E. B. Rimm, A. Medina-Remón, M. A. Martínez-González, R. de la Torre, D. Corella, J. Salas-Salvadó, E. Gómez-Gracia, J. Lapetra, F. Arós, M. Fiol, E. Ros, L. Serra-Majem, X. Pintó, G. T. Saez, J. Basora, J. V. Sorlí, J. A. Martínez, E. Vinyoles, V. Ruiz-Gutiérrez, R. Estruch and R. M. Lamuela-Raventós, Inverse association between habitual polyphenol intake and incidence of cardiovascular events in the PREDIMED study, *Nutr. Metab. Cardiovasc. Dis.*, 2014, **24**, 639–647.

54 X. Liu, Y. Liu, Y. Huang, H. Yu, S. Yuan, B. Tang, P. Wang and Q. He, Dietary total flavonoids intake and risk of mortality from all causes and cardiovascular disease in the general population: A systematic review and meta-analysis of cohort studies, *Mol. Nutr. Food Res.*, 2017, **61**, 1601003.

55 J. Rienks, J. Barbaresko, K. Oluwagbemigun, M. Schmid and U. Nöthlings, Polyphenol exposure and risk of type 2 diabetes: dose-response meta-analyses and systematic review of prospective cohort studies, *Am. J. Clin. Nutr.*, 2018, **108**, 49–61.

56 Z-M. Wang, D. Zhao, Z-L. Nie, H. Zhao, B. Zhou, W. Gao, L-S. Wang and Z-J. Yang, Flavonol intake and stroke risk: a meta-analysis of cohort studies, *Nutrition*, 2014, **30**, 518–523.

57 H. Xu, J. Luo, J. Huang and Q. Wen, Flavonoids intake and risk of type 2 diabetes mellitus: A meta-analysis of prospective cohort studies, *Medicine*, 2018, **97**, e0686.

58 A. Micek, J. Godos, D. Del Rio, F. Galvano and G. Grosso, Dietary Flavonoids and Cardiovascular Disease: A Comprehensive Dose–Response Meta-Analysis, *Mol. Nutr. Food Res.*, 2021, **65**, 1–11.

59 K. D. Croft, Dietary polyphenols: Antioxidants or not?, *Arch. Biochem. Biophys.*, 2016, **595**, 120–124.

60 A. Rodriguez-Mateos, G. Istas, L. Boschek, R. P. Feliciano, C. E. Mills, C. Boby, S. Gomez-Alonso, D. Milenkovic and C. Heiss, Circulating Anthocyanin Metabolites Mediate Vascular Benefits of Blueberries: Insights From Randomized Controlled Trials, Metabolomics, and Nutrigenomics, *Journals Gerontol. Ser. A*, 2019, **74**, 967–976.

61 A. Rodriguez-Mateos, C. Rendeiro, T. Bergillos-Meca, S. Tabatabaee, T. W. George, C. Heiss and J. P. Spencer, Intake and time dependence of blueberry flavonoid–induced improvements in vascular function: a randomized, controlled, double-blind, crossover intervention study with mechanistic insights into biological activity, *Am. J. Clin. Nutr.*, 2013, **98**, 1179–1191.

62 H. Schroeter, C. Heiss, J. Balzer, P. Kleinbongard, C. L. Keen, N. K. Hollenberg, H. Sies, C. Kwik-Uribe, H. H. Schmitz and M. Kelm, (-)-Epicatechin mediates beneficial effects of flavanol-rich cocoa on vascular function in humans, *Proc. Natl. Acad. Sci.*, 2006, **103**, 1024–1029.

63 V. Habauzit, M.-A. Verny, D. Milenkovic, N. Barber-Chamoux, A. Mazur, C. Dubray and C. Morand, Flavanones protect from arterial stiffness in postmenopausal women consuming grapefruit juice for 6 mo: a randomized, controlled, crossover trial, *Am. J. Clin. Nutr.*, 2015, **102**, 66–74.

64 A. González-Sarrías, R. García-Villalba, M. Romo-Vaquero, C. Alasalvar, A. Örem, P. Zafrilla, F.

A. Tomás-Barberán, M. V. Selma and J. C. Espín, Clustering according to urolithin metabotype explains the interindividual variability in the improvement of cardiovascular risk biomarkers in overweight-obese individuals consuming pomegranate: A randomized clinical trial, *Mol. Nutr. Food Res.*, 2017, **61**, 1–14.

65 S. Hazim, P. J. Curtis, M. Y. Schär, L. M. Ostertag, C. D. Kay, A.-M. Minihane and A. Cassidy, Acute benefits of the microbial-derived isoflavone metabolite equol on arterial stiffness in men prospectively recruited according to equol producer phenotype: a double-blind randomized controlled trial, *Am. J. Clin. Nutr.*, 2016, **103**, 694–702.

66 C. Morand, B. De Roos, M. T. Garcia-Conesa, E. R. Gibney, R. Landberg, C. Manach, D. Milenkovic, A. Rodriguez-Mateos, T. Van de Wiele and F. Tomas-Barberan, Why interindividual variation in response to consumption of plant food bioactives matters for future personalised nutrition, *Proc. Nutr. Soc.*, 2020, **79**, 225–235.

67 E. Kesse-Guyot, L. Fezeu, V. A. Andreeva, M. Touvier, A. Scalbert, S. Hercberg and P. Galan, Total and Specific Polyphenol Intakes in Midlife Are Associated with Cognitive Function Measured 13 Years Later, *J. Nutr.*, 2012, **142**, 76–83.

68 A. Moreira, M. J. Diógenes, A. de Mendonça, N. Lunet and H. Barros, Chocolate Consumption is Associated with a Lower Risk of Cognitive Decline, *J. Alzheimer's Dis.*, 2016, **53**, 85–93.

69 G. E. Crichton, M. F. Elias and A. Alkerwi, Chocolate intake is associated with better cognitive function: The Maine-Syracuse Longitudinal Study, *Appetite*, 2016, **100**, 126–132.

70 X. Liu, X. Du, G. Han and W. Gao, Association between tea consumption and risk of cognitive disorders: A dose-response meta-analysis of observational studies, *Oncotarget*, 2017, **8**, 43306–43321.

71 D. Mastroiacovo, C. Kwik-uribe, D. Grassi, S. Necozione, A. Raffaele, L. Pistacchio, R. Righetti, R. Bocale, M. C. Lechiara, C. Marini, C. Ferri and G. Desideri, Cocoa flavanol consumption improves cognitive function, blood pressure control, and metabolic profile in elderly subjects: the Cocoa, Cognition, and Aging (CoCoA) Study - a randomized controlled trial, *Am. J. Clin. Nutr.*, 2015, **101**, 538–548.

72 G. Desideri, C. Kwik-Uribe, D. Grassi, S. Necozione, L. Ghiadoni, D. Mastroiacovo, A. Raffaele, L. Ferri, R. Bocale, M. C. Lechiara, C. Marini and C. Ferri, Benefits in cognitive function, blood

pressure, and insulin resistance through cocoa flavanol consumption in elderly subjects with mild cognitive impairment: The cocoa, cognition, and aging (CoCoA) study, *Hypertension*, 2012, **60**, 794–801.

73 C. Favari, P. Mena, C. Curti, D. Rio and D. Angelino, in *Dietary Polyphenols*, Wiley, 2020, 283–317.

74 P. Mena, L. Bresciani, N. Brindani, I. A. Ludwig, G. Pereira-Caro, D. Angelino, R. Llorach, L. Calani, F. Brighenti, M. N. Clifford, C. I. R. Gill, A. Crozier, C. Curti and D. Del Rio, Phenyl-γ-valerolactones and phenylvaleric acids, the main colonic metabolites of flavan-3-ols: synthesis, analysis, bioavailability, and bioactivity, *Nat. Prod. Rep.*, 2019, **36**, 714–752.

75 J. I. Ottaviani, T. Y. Momma, G. K. Kuhnle, C. L. Keen and H. Schroeter, Structurally related (-)-epicatechin metabolites in humans: Assessment using de novo chemically synthesized authentic standards, *Free Radic. Biol. Med.*, 2012, **52**, 1403–1412.

76 V. Neveu, J. Perez-Jimenez, F. Vos, V. Crespy, L. du Chaffaut, L. Mennen, C. Knox, R. Eisner, J. Cruz, D. Wishart and A. Scalbert, Phenol-Explorer: an online comprehensive database on polyphenol contents in foods, *Database*, 2010, **2010**, bap024–bap024.

77 M. Kutschera, W. Engst, M. Blaut and A. Braune, Isolation of catechin-converting human intestinal bacteria, *J. Appl. Microbiol.*, 2011, **111**, 165–175.

78 M. M. Appeldoorn, J.-P. Vincken, A.-M. Aura, P. C. H. Hollman and H. Gruppen, Procyanidin Dimers Are Metabolized by Human Microbiota with 2-(3,4-Dihydroxyphenyl)acetic Acid and 5-(3,4-Dihydroxyphenyl)-γ-valerolactone as the Major Metabolites, *J. Agric. Food Chem.*, 2009, **57**, 1084–1092.

79 S. Stoupi, G. Williamson, J. W. Drynan, D. Barron and M. N. Clifford, A comparison of the *in vitro* biotransformation of (-)-epicatechin and procyanidin B2 by human faecal microbiota, *Mol. Nutr. Food Res.*, 2010, **54**, 747–759.

80 K. Ou, P. Sarnoski, K. R. Schneider, K. Song, C. Khoo and L. Gu, Microbial catabolism of procyanidins by human gut microbiota, *Mol. Nutr. Food Res.*, 2014, **58**, 2196–2205.

81 A. Engemann, F. Hübner, S. Rzeppa and H. U. Humpf, Intestinal metabolism of two A-type procyanidins using the pig cecum model: Detailed structure elucidation of unknown catabolites with Fourier transform mass spectrometry (FTMS), *J. Agric. Food Chem.*, 2012, **60**,

749–757.

82 G. Borges, J. I. Ottaviani, J. J. J. van der Hooft, H. Schroeter and A. Crozier, Absorption, metabolism, distribution and excretion of (−)-epicatechin: A review of recent findings, *Mol. Aspects Med.*, 2018, **61**, 18–30.

83 A. Stalmach, W. Mullen, H. Steiling, G. Williamson, M. E. J. Lean and A. Crozier, Absorption, metabolism, and excretion of green tea flavan-3-ols in humans with an ileostomy, *Mol. Nutr. Food Res.*, 2010, **54**, 323–334.

84 J. I. Ottaviani, G. Borges, T. Y. Momma, J. P. E. Spencer, C. L. Keen, A. Crozier and H. Schroeter, The metabolome of [2-14C](-)-epicatechin in humans: Implications for the assessment of efficacy, safety, and mechanisms of action of polyphenolic bioactives, *Sci. Rep.*, 2016, **6**, 1–10.

85 S. Wiese, T. Esatbeyoglu, P. Winterhalter, H. P. Kruse, S. Winkler, A. Bub and S. E. Kulling, Comparative biokinetics and metabolism of pure monomeric, dimeric, and polymeric flavan-3-ols: A randomized cross-over study in humans, *Mol. Nutr. Food Res.*, 2015, **59**, 610–621.

86 F. Castello, G. Costabile, L. Bresciani, M. Tassotti, D. Naviglio, D. Luongo, P. Ciciola, M. Vitale, C. Vetrani, G. Galaverna, F. Brighenti, R. Giacco, D. Del Rio and P. Mena, Bioavailability and pharmacokinetic profile of grape pomace phenolic compounds in humans, *Arch. Biochem. Biophys.*, 2018, **646**, 1–9.

87 S. Sang, M.-J. Lee, I. Yang, B. Buckley and C. S. Yang, Human urinary metabolite profile of tea polyphenols analyzed by liquid chromatography/electrospray ionization tandem mass spectrometry with data-dependent acquisition, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 1567–1578.

88 L. Calani, D. Del Rio, M. Luisa Callegari, L. Morelli and F. Brighenti, Updated bioavailability and 48 h excretion profile of flavan-3-ols from green tea in humans, *Int. J. Food Sci. Nutr.*, 2012, **63**, 513–521.

89 D. Del Rio, L. Calani, C. Cordero, S. Salvatore, N. Pellegrini and F. Brighenti, Bioavailability and catabolism of green tea flavan-3-ols in humans, *Nutrition*, 2010, **26**, 1110–1116.

90 J. J. J. van der Hooft, R. C. H. de Vos, V. Mihaleva, R. J. Bino, L. Ridder, N. de Roo, D. M. Jacobs, J. P. M. van Duynhoven and J. Vervoort, Structural Elucidation and Quantification of Phenolic

Conjugates Present in Human Urine after Tea Intake, *Anal. Chem.*, 2012, **84**, 7263–7271.

91 J. van Duynhoven, J. J. J. van der Hooft, F. A. van Dorsten, S. Peters, M. Foltz, V. Gomez-Roldan, J. Vervoort, R. C. H. de Vos and D. M. Jacobs, Rapid and Sustained Systemic Circulation of Conjugated Gut Microbial Catabolites after Single-Dose Black Tea Extract Consumption, *J. Proteome Res.*, 2014, **13**, 2668–2678.

92 S. Roowi, A. Stalmach, W. Mullen, M. E. J. Lean, C. A. Edwards and A. Crozier, Green Tea Flavan-3-ols: Colonic Degradation and Urinary Excretion of Catabolites by Humans, *J. Agric. Food Chem.*, 2010, **58**, 1296–1304.

93 M. Urpi-Sarda, M. Monagas, N. Khan, R. M. Lamuela-Raventos, C. Santos-Buelga, E. Sacanella, M. Castell, J. Permanyer and C. Andres-Lacueva, Epicatechin, procyanidins, and phenolic microbial metabolites after cocoa intake in humans and rats, *Anal. Bioanal. Chem.*, 2009, **394**, 1545–1556.

94 P. Vitaglione, R. Barone Lumaga, R. Ferracane, S. Sellitto, J. R. Morelló, J. Reguant Miranda, E. Shimoni and V. Fogliano, Human bioavailability of flavanols and phenolic acids from cocoa-nut creams enriched with free or microencapsulated cocoa polyphenols, *Br. J. Nutr.*, 2013, **109**, 1832–1843.

95 L. Y. Rios, M.-P. Gonthier, C. Rémésy, I. Mila, C. Lapierre, S. A. Lazarus, G. Williamson and A. Scalbert, Chocolate intake increases urinary excretion of polyphenol-derived phenolic acids in healthy human subjects, *Am. J. Clin. Nutr.*, 2003, **77**, 912–918.

96 M. Boto-Ordóñez, M. Urpi-Sarda, M. I. Queipo-Ortuño, D. Corella, F. J. Tinahones, R. Estruch and C. Andres-Lacueva, Microbial Metabolomic Fingerprinting in Urine after Regular Dealcoholized Red Wine Consumption in Humans, *J. Agric. Food Chem.*, 2013, **61**, 9166–9175.

97 O. Khymenets, C. Andres-Lacueva, M. Urpi-Sarda, R. Vazquez-Fresno, M. M. Mart, G. Reglero, M. Torres and R. Llorach, Metabolic fingerprint after acute and under sustained consumption of a functional beverage based on grape skin extract in healthy human subjects, *Food Funct.*, 2015, **6**, 1288–1298.

98 R. P. Feliciano, A. Boeres, L. Massacessi, G. Istas, M. R. Ventura, C. Nunes dos Santos, C. Heiss and A. Rodriguez-Mateos, Identification and quantification of novel cranberry-derived plasma and urinary (poly)phenols, *Arch. Biochem. Biophys.*, 2016, **599**, 31–41.

99 K. Trošt, M. M. Ulaszewska, J. Stanstrup, D. Albanese, C. De Filippo, K. M. Tuohy, F. Natella, C. Scaccini and F. Mattivi, Host: Microbiome co-metabolic processing of dietary polyphenols – An acute, single blinded, cross-over study with different doses of apple polyphenols in healthy subjects, *Food Res. Int.*, 2018, **112**, 108–128.

100 I. Garrido, M. Urpi-Sarda, M. Monagas, C. Gómez-Cordovés, P. J. Martín-Álvarez, R. Llorach, B. Bartolomé and C. Andrés-Lacueva, Targeted Analysis of Conjugated and Microbial-Derived Phenolic Metabolites in Human Urine After Consumption of an Almond Skin Phenolic Extract, *J. Nutr.*, 2010, **140**, 1799–1807.

101 M. Urpi-Sarda, I. Garrido, M. Monagas, C. Gómez-Cordovés, A. Medina-Remón, C. Andres-Lacueva and B. Bartolomé, Profile of Plasma and Urine Metabolites after the Intake of Almond [Prunus dulcis (Mill.) D.A. Webb] Polyphenols in Humans, *J. Agric. Food Chem.*, 2009, **57**, 10134–10142.

102 J. I. Ottaviani, R. Fong, J. Kimball, J. L. Ensunsa, A. Britten, D. Lucarelli, R. Luben, P. B. Grace, D. H. Mawson, A. Tym, A. Wierzbicki, K. T. Khaw, H. Schroeter and G. G. C. Kuhnle, Evaluation at scale of microbiome-derived metabolites as biomarker of flavan-3-ol intake in epidemiological studies, *Sci. Rep.*, 2018, **8**, 1–11.

103 P. Mena, I. A. Ludwig, V. B. Tomatis, A. Acharjee, L. Calani, A. Rosi, F. Brighenti, S. Ray, J. L. Griffin, L. J. Bluck and D. Del Rio, Inter-individual variability in the production of flavan-3-ol colonic metabolites: preliminary elucidation of urinary metabotypes, *Eur. J. Nutr.*, 2019, **58**, 1529–1543.

104 A. Cortés-Martín, M. V. Selma, J. C. Espín and R. García-Villalba, The Human Metabolism of Nuts Proanthocyanidins does not Reveal Urinary Metabolites Consistent with Distinctive Gut Microbiota Metabotypes, *Mol. Nutr. Food Res.*, 2019, **63**, 1800819.

105 E. Hillesheim and L. Brennan, Metabotyping and its role in nutrition research, *Nutr. Res. Rev.*, 2020, **33**, 33–42.

106 M. Serafini, R. Bugianesi, G. Maiani, S. Valtuena, S. De Santis and A. Crozier, Plasma antioxidants from chocolate, *Nature*, 2003, **424**, 1013–1013.

107 C. A. Rice-evans, N. J. Miller, P. G. Bolwell, P. M. Bramley and J. B. Pridham, The Relative Antioxidant Activities of Plant-Derived Polyphenolic Flavonoids, *Free Radic. Res.*, 1995, **22**,

375–383.

108 S. Schaffer and B. Halliwell, Do polyphenols enter the brain and does it matter? Some theoretical and practical considerations, *Genes Nutr.*, 2012, **7**, 99–109.

109 A. M. Brickman, U. A. Khan, F. A. Provenzano, L.-K. Yeung, W. Suzuki, H. Schroeter, M. Wall, R. P. Sloan and S. A. Small, Enhancing dentate gyrus function with dietary flavanols improves cognition in older adults., *Nat. Neurosci.*, 2014, **17**, 1798–803.

110 H. Schroeter, C. Heiss, J. Balzer, P. Kleinbongard, C. L. Keen, N. K. Hollenberg, H. Sies, C. Kwik-Uribe, H. H. Schmitz and M. Kelm, (-)-Epicatechin mediates beneficial effects of flavanol-rich cocoa on vascular function in humans., *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 1024–9.

111 H. Girouard and C. Iadecola, Neurovascular coupling in the normal brain and in hypertension, stroke, and Alzheimer disease, *J. Appl. Physiol.*, 2006, **100**, 328–335.

112 F. A. Sorond, S. Hurwitz, D. H. Salat, D. N. Greve and N. D. L. Fisher, Neurovascular coupling, cerebral white matter integrity, and response to cocoa in older people, *Neurology*, 2013, **81**, 904–909.

113 D. Mastroiacovo, C. Kwik-Uribe, D. Grassi, S. Necozione, A. Raffaele, L. Pistacchio, R. Righetti, R. Bocale, M. C. Lechiara, C. Marini, C. Ferri and G. Desideri, Cocoa flavanol consumption improves cognitive function, blood pressure control, and metabolic profile in elderly subjects: the Cocoa, Cognition, and Aging (CoCoA) Study--a randomized controlled trial., *Am. J. Clin. Nutr.*, 2015, **101**, 538–548.

114 O. Weinreb, S. Mandel, T. Amit and M. B. H. Youdim, Neurological mechanisms of green tea polyphenols in Alzheimer's and Parkinson's diseases, *J. Nutr. Biochem.*, 2004, **15**, 506–516.

115 O. Weinreb, T. Amit, S. Mandel and M. B. H. Youdim, Neuroprotective molecular mechanisms of (−)-epigallocatechin-3-gallate: a reflective outcome of its antioxidant, iron chelating and neuritogenic properties, *Genes Nutr.*, 2009, **4**, 283–296.

116 L. Ortiz-López, B. Márquez-Valadez, A. Gómez-Sánchez, M. D. C. Silva-Lucero, M. Torres-Pérez, R. I. Téllez-Ballesteros, M. Ichwan, M. A. Meraz-Ríos, G. Kempermann and G. B. Ramírez-Rodríguez, Green tea compound epigallo-catechin-3-gallate (EGCG) increases neuronal survival in adult hippocampal neurogenesis *in vivo* and *in vitro*, *Neuroscience*, 2016, **322**, 208–220.

117 S.-J. Lee and Y. Jia, The effect of bioactive compounds in tea on lipid metabolism and obesity through regulation of peroxisome proliferator-activated receptors, *Curr. Opin. Lipidol.*, 2015, **26**, 3–9.

118 K. Ide, N. Matsuoka, H. Yamada, D. Furushima, K. Kawakami, K. Ide, N. Matsuoka, H. Yamada, D. Furushima and K. Kawakami, Effects of Tea Catechins on Alzheimer's Disease: Recent Updates and Perspectives, *Molecules*, 2018, **23**, 2357.

119 K. Unno, M. Pervin, A. Nakagawa, K. Iguchi, A. Hara, A. Takagaki, F. Nanjo, A. Minami and Y. Nakamura, Blood-Brain Barrier Permeability of Green Tea Catechin Metabolites and their Neuritogenic Activity in Human Neuroblastoma SH-SY5Y Cells, *Mol. Nutr. Food Res.*, 2017, **5**, 1700294.

120 D. Angelino, D. Carregosa, C. Domenech-Coca, M. Savi, I. Figueira, N. Brindani, S. Jang, S. Lakshman, A. Molokin, J. F. Urban, C. D. Davis, M. A. Brito, K. S. Kim, F. Brighenti, C. Curti, C. Bladé, J. M. del Bas, D. Stilli, G. I. Solano-Aguilar, C. N. dos Santos, D. del Rio and P. Mena, 5-(Hydroxyphenyl)-γ-Valerolactone-Sulfate, a Key Microbial Metabolite of Flavan-3-ols, Is Able to Reach the Brain: Evidence from Different in Silico, *In Vitro* and *In Vivo* Experimental Models, *Nutrients*, 2019, **11**, 2678.

121 M. L. McCullough, J. J. Peterson, R. Patel, P. F. Jacques, R. Shah and J. T. Dwyer, Flavonoid intake and cardiovascular disease mortality in a prospective cohort of US adults, *Am. J. Clin. Nutr.*, 2012, **95**, 454–464.

122 A. Buitrago-Lopez, J. Sanderson, L. Johnson, S. Warnakula, A. Wood, E. Di Angelantonio and O. H. Franco, Chocolate consumption and cardiometabolic disorders: Systematic review and meta-analysis, *BMJ*, 2011, **343**, 1–8.

123 A. González-Sarrías, E. Combet, P. Pinto, P. Mena, M. Dall'Asta, M. Garcia-Aloy, A. Rodríguez-Mateos, E. R. Gibney, J. Dumont, M. Massaro, J. Sánchez-Meca, C. Morand and M.-T. García-Conesa, A Systematic Review and Meta-Analysis of the Effects of Flavanol-Containing Tea, Cocoa and Apple Products on Body Composition and Blood Lipids: Exploring the Factors Responsible for Variability in Their Efficacy, *Nutrients*, 2017, **9**, 746.

124 L. M. Berends, V. van der Velpen and A. Cassidy, Flavan-3-ols, theobromine, and the effects of cocoa and chocolate on cardiometabolic risk factors, *Curr. Opin. Lipidol.*, 2015, **26**, 10–19.

125 A. Rodriguez-Mateos, C. Heiss, G. Borges and A. Crozier, Berry (Poly)phenols and Cardiovascular Health, *J. Agric. Food Chem.*, 2014, **62**, 3842–3851.

126 A. Rodriguez-Mateos, T. Weber, S. S. Skene, J. I. Ottaviani, A. Crozier, M. Kelm, H. Schroeter and C. Heiss, Assessing the respective contributions of dietary flavanol monomers and procyanidins in mediating cardiovascular effects in humans: randomized, controlled, double-masked intervention trial, *Am. J. Clin. Nutr.*, 2018, **108**, 1229–1237.

127 A. Rodriguez-Mateos, R. P. Feliciano, A. Boeres, T. Weber, C. N. dos Santos, M. R. Ventura and C. Heiss, Cranberry (poly)phenol metabolites correlate with improvements in vascular function: A double-blind, randomized, controlled, dose-response, crossover study, *Mol. Nutr. Food Res.*, 2016, **60**, 2130–2140.

128 M. M. Dohadwala, M. Holbrook, N. M. Hamburg, S. M. Shenouda, W. B. Chung, M. Titas, M. a Kluge, N. Wang, J. Palmisano, P. E. Milbury, J. B. Blumberg and J. A. Vita, Effects of cranberry juice consumption on vascular function in patients with coronary artery disease., *Am. J. Clin. Nutr.*, 2011, **5**, 934–940.

129 A. Rodriguez-Mateos, C. Rendeiro, T. Bergillos-Meca, S. Tabatabaee, T. W. George, C. Heiss and J. P. Spencer, Intake and time dependence of blueberry flavonoid–induced improvements in vascular function: a randomized, controlled, double-blind, crossover intervention study with mechanistic insights into biological activity, *Am. J. Clin. Nutr.*, 2013, **98**, 1179–1191.

130 R. R. Holt, D. D. Schramm, C. L. Keen, S. A. Lazarus and H. H. Schmitz, Chocolate consumption and platelet function, *JAMA*, 2002, **287**, 2212–2213.

131 I. Erlund, R. Koli, G. Alfthan, J. Marniemi, P. Puukka, P. Mustonen, P. Mattila and A. Jula, Favorable effects of berry consumption on platelet function, blood pressure, and HDL cholesterol, *Am. J. Clin. Nutr.*, 2008, **87**, 323–331.

132 M. Montagnana, E. Danese, D. Angelino, P. Mena, A. Rosi, M. Benati, M. Gelati, G. L. Salvagno, E. J. Favaloro, D. Del Rio and G. Lippi, Dark chocolate modulates platelet function with a mechanism mediated by flavan-3-ol metabolites, *Medicine*, 2018, **97**, e13432.

133 R. Jiménez, J. Duarte and F. Perez-Vizcaino, Epicatechin: Endothelial Function and Blood Pressure, *J. Agric. Food Chem.*, 2012, **60**, 8823–8830.

134 J. Bernatoniene and D. Kopustinskiene, The Role of Catechins in Cellular Responses to Oxidative Stress, *Molecules*, 2018, **23**, 965.

135 M. E. Terlizzi, G. Gribaudo and M. E. Maffei, UroPathogenic Escherichia coli (UPEC) Infections: Virulence Factors, Bladder Responses, Antibiotic, and Non-antibiotic Antimicrobial Strategies, *Front. Microbiol.*, 2017, **8**, 1566.

136 A. L. Flores-Mireles, J. N. Walker, M. Caparon and S. J. Hultgren, Urinary tract infections: epidemiology, mechanisms of infection and treatment options, *Nat. Rev. Microbiol.*, 2015, **13**, 269–284.

137 L. E. Nicolle, Cranberry for Prevention of Urinary Tract Infection?, *JAMA,* 2016, **316**, 1873–1874.

138 D. González de Llano, A. Esteban-Fernández, F. Sánchez-Patán, P. Martínlvarez, M. Moreno-Arribas and B. Bartolomé, Anti-Adhesive Activity of Cranberry Phenolic Compounds and Their Microbial-Derived Metabolites against Uropathogenic Escherichia coli in Bladder Epithelial Cell Cultures, *Int. J. Mol. Sci.*, 2015, **16**, 12119–12130.

139 K. L. Kaspar, A. B. Howell and C. Khoo, A randomized, double-blind, placebo-controlled trial to assess the bacterial anti-adhesion effects of cranberry extract beverages, *Food Funct.*, 2015, **6**, 1212–1217.

140 G. Peron, S. Sut, A. Pellizzaro, P. Brun, D. Voinovich, I. Castagliuolo and S. Dall'Acqua, The antiadhesive activity of cranberry phytocomplex studied by metabolomics: Intestinal PAC-A metabolites but not intact PAC-A are identified as markers in active urines against uropathogenic Escherichia coli, *Fitoterapia*, 2017, **122**, 67–75.

141 P. Mena, D. González de Llano, N. Brindani, A. Esteban-Fernández, C. Curti, M. V. Moreno-Arribas, D. Del Rio and B. Bartolomé, 5-(3',4'-Dihydroxyphenyl)-γ-valerolactone and its sulphate conjugates, representative circulating metabolites of flavan-3-ols, exhibit anti-adhesive activity against uropathogenic Escherichia coli in bladder epithelial cells, *J. Funct. Foods*, 2017, **29**, 275–280.

142 D. González de Llano, H. Liu, C. Khoo, M. V. Moreno-Arribas and B. Bartolomé, Some New Findings Regarding the Antiadhesive Activity of Cranberry Phenolic Compounds and Their Microbial-Derived Metabolites against Uropathogenic Bacteria, *J. Agric. Food Chem.*, 2019,

**67**, 2166–2174.

143 P. Mena, I. A. Ludwig, V. B. Tomatis, A. Acharjee, L. Calani, A. Rosi, F. Brighenti, S. Ray, J. L. Griffin, L. J. Bluck and D. Del Rio, Inter-individual variability in the production of flavan-3-ol colonic metabolites: preliminary elucidation of urinary metabotypes, *Eur. J. Nutr.*, 2019, **58**, 1529–1543

144 F. Potì, D. Santi, G. Spaggiari, F. Zimetti and I. Zanotti, Polyphenol Health Effects on Cardiovascular and Neurodegenerative Disorders: A Review and Meta-Analysis, *Int. J. Mol. Sci.*, 2019, **20**, 351.

145 J. I. Ottaviani, C. Heiss, J. P. E. Spencer, M. Kelm and H. Schroeter, *Mol. Aspects Med.*, 2018, **61**, 63–75.

146 C. Morand and F. A. Tomás-Barberán, Interindividual Variability in Absorption, Distribution, Metabolism, and Excretion of Food Phytochemicals Should Be Reported, *J. Agric. Food Chem.*, 2019, **67**, 3843–3844.

147 P. Mena and D. Del Rio, Gold Standards for Realistic (Poly)phenol Research, *J. Agric. Food Chem.*, 2018, **66**, 8221–8223.

148 M. M. Ulaszewska, C. H. Weinert, A. Trimigno, R. Portmann, C. Andres Lacueva, R. Badertscher, L. Brennan, C. Brunius, A. Bub, F. Capozzi, M. Cialiè Rosso, C. E. Cordero, H. Daniel, S. Durand, B. Egert, P. G. Ferrario, E. J. M. Feskens, P. Franceschi, M. Garcia-Aloy, F. Giacomoni, P. Giesbertz, R. González-Domínguez, K. Hanhineva, L. Y. Hemeryck, J. Kopka, S. E. Kulling, R. Llorach, C. Manach, F. Mattivi, C. Migné, L. H. Münger, B. Ott, G. Picone, G. Pimentel, E. Pujos-Guillot, S. Riccadonna, M. J. Rist, C. Rombouts, J. Rubert, T. Skurk, P. S. C. Sri Harsha, L. Van Meulebroek, L. Vanhaecke, R. Vázquez-Fresno, D. Wishart and G. Vergères, Nutrimetabolomics: An Integrative Action for Metabolomic Analyses in Human Nutritional Studies, *Mol. Nutr. Food Res.*, 2019, **63**, 1800384.

149 D. S. Wishart, Metabolomics: applications to food science and nutrition research, *Trends Food Sci. Technol.*, 2008, **19**, 482–493.

150 A. Scalbert, L. Brennan, C. Manach, C. Andres-Lacueva, L. O. Dragsted, J. Draper, S. M. Rappaport, J. J. van der Hooft and D. S. Wishart, The food metabolome: a window over dietary exposure, *Am. J. Clin. Nutr.*, 2014, **99**, 1286–1308.

151 U. Vrhovsek, D. Masuero, M. Gasperotti, P. Franceschi, L. Caputi, R. Viola and F. Mattivi, A Versatile Targeted Metabolomics Method for the Rapid Quantification of Multiple Classes of Phenolics in Fruits and Beverages, *J. Agric. Food Chem.*, 2012, **60**, 8831–8840.

152 A. Gucciardi, P. Pirillo, I. M. Di Gangi, M. Naturale and G. Giordano, A rapid UPLC–MS/MS method for simultaneous separation of 48 acylcarnitines in dried blood spots and plasma useful as a second-tier test for expanded newborn screening, *Anal. Bioanal. Chem.*, 2012, **404**, 741–751.

153 H. Gibbons, A. O'Gorman and L. Brennan, Metabolomics as a tool in nutritional research, *Curr. Opin. Lipidol.*, 2015, **26**, 30–34.

154 L. Brennan, Use of metabotyping for optimal nutrition, *Curr. Opin. Biotechnol.*, 2017, **44**, 35–38.

155 I. A. Ludwig, P. Mena, L. Calani, G. Borges, G. Pereira-Caro, L. Bresciani, D. Del Rio, M. E. J. Lean and A. Crozier, New insights into the bioavailability of red raspberry anthocyanins and ellagitannins, *Free Radic. Biol. Med.*, 2015, **89**, 758–769.

156 G. Pereira-Caro, T. Polyviou, I. A. Ludwig, A. M. Nastase, J. M. Moreno-Rojas, A. L. Garcia, D. Malkova and A. Crozier, Bioavailability of orange juice (poly)phenols: The impact of short-term cessation of training by male endurance athletes, *Am. J. Clin. Nutr.*, 2017, **106**, 791–800.

157 G. Astarita and J. Langridge, An Emerging Role for Metabolomics in Nutrition Science, *J. Nutrigenet. Nutrigenomics*, 2013, **6**, 181–200.

158 L. O. Dragsted, Q. Gao, A. Scalbert, G. Vergères, M. Kolehmainen, C. Manach, L. Brennan, L. A. Afman, D. S. Wishart, C. Andres Lacueva, M. Garcia-Aloy, H. Verhagen, E. J. M. Feskens and G. Praticò, Validation of biomarkers of food intake—critical assessment of candidate biomarkers, *Genes Nutr.*, 2018, **13**, 14.

159 A. Riedl, C. Gieger, H. Hauner, H. Daniel and J. Linseisen, Metabotyping and its application in targeted nutrition: an overview, *Br. J. Nutr.*, 2017, **117**, 1631–1644.

160 T. Larkin, W. E. Price and L. Astheimer, The Key Importance of Soy Isoflavone Bioavailability to Understanding Health Benefits, *Crit. Rev. Food Sci. Nutr.*, 2008, **48**, 538–552.

161 L. Brennan, Metabolomics in nutrition research–a powerful window into nutritional metabolism, *Essays Biochem.*, 2016, **60**, 451–458.

162 S. E. Chiuve, E. L. Giovannucci, S. E. Hankinson, S. H. Zeisel, L. W. Dougherty, W. C. Willett and E. B. Rimm, The association between betaine and choline intakes and the plasma concentrations of homocysteine in women, *Am. J. Clin. Nutr.*, 2007, **86**, 1073.

163 Y. Zheng, U. Ceglarek, T. Huang, L. Li, J. Rood, D. H. Ryan, G. A. Bray, F. M. Sacks, D. Schwarzfuchs, J. Thiery, I. Shai and L. Qi, Weight-loss diets and 2-y changes in circulating amino acids in 2 randomized intervention trials–, *Am. J. Clin. Nutr.*, 2016, **103**, 505–511.

164 B. Khakimov, S. K. Poulsen, F. Savorani, E. Acar, G. Gürdeniz, T. M. Larsen, A. Astrup, L. O. Dragsted and S. B. Engelsen, 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, *J. Proteome Res.*, 2016, **15**, 1939–1954.

165 J. Chong, O. Soufan, C. Li, I. Caraus, S. Li, G. Bourque, D. S. Wishart and J. Xia, MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis, *Nucleic Acids Res.*, 2018, **46**, W486–W494.

166 D. K. Barupal and O. Fiehn, Chemical Similarity Enrichment Analysis (ChemRICH) as alternative to biochemical pathway mapping for metabolomic datasets, *Sci. Rep.*, 2017, **7**, 14567.

167 A. Scalbert, L. Brennan, O. Fiehn, T. Hankemeier, B. S. Kristal, B. van Ommen, E. Pujos-Guillot, E. Verheij, D. Wishart and S. Wopereis, Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research, *Metabolomics*, 2009, **5**, 435–458.

168 D. S. Wishart, T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat, E. Dong, S. Bouatra, I. Sinelnikov, D. Arndt, J. Xia, P. Liu, F. Yallou, T. Bjorndahl, R. Perez-Pineiro, R. Eisner, F. Allen, V. Neveu, R. Greiner and A. Scalbert, HMDB 3.0-The Human Metabolome Database in 2013, *Nucleic Acids Res.*, 2013, **41**, 801–807.

# Chapter 2 – Aims of the thesis

The objective of this doctoral thesis was to apply metabolomics for the analysis of biological samples collected in nutrition intervention studies in humans, to gain new and further insights in the field of nutrition research focused on dietary bioactive compounds. In particular, within my PhD project, I applied:

(I) targeted metabolomics to evaluate the absorption, metabolism and excretion of flavan-3-ols from cranberry through the analysis of their main colonic metabolites (phenyl-γ-valerolactones and phenyl-valeric acids), considering also their potential use as biomarkers of flavan-3-ol intake and the inter-individual variability in their appearance in plasma and urine (**3.1** Study 1 – *Kinetic profile and urinary excretion of phenyl-γ-valerolactones upon consumption of cranberry: a dose–response relationship*),

(II) targeted metabolomics to elucidate the presence of different metabotypes in the urinary excretion of principal flavan-3-ol colonic metabolites after consumption of cranberry products, also assessing the impact of the statistical technique used for metabotyping (**3.2** Study 2 – *Metabotypes of flavan-3-ol colonic metabolites after cranberry intake: elucidation and statistical approaches*),

(III) untargeted metabolomics to investigate the changes occurring at metabolomic level following consumption of different patterns of coffee and the impact on metabolic pathways (**3.3** Study 3 – *Metabolomic changes after coffee consumption: new paths on the block*).

# Chapter 3 – Selected studies

# 3.1. Study 1

## Kinetic profile and urinary excretion of phenyl-γ-valerolactones upon consumption of cranberry: a dose-response relationship

**Claudia Favari**[1], Pedro Mena[1], Claudio Curti[2], Geoffrey Istas[3], Christian Heiss[4], Daniele Del Rio[5, 6,*], Ana Rodriguez-Mateos[3,*]

[1] Human Nutrition Unit, Department of Food & Drug, University of Parma, Parma, Italy

[2] Department of Food & Drug, University of Parma, Parma, Italy

[3] Department of Nutritional Sciences, Faculty of Life Sciences and Medicine, King's College London, London, UK

[4] Department of Clinical and Experimental Medicine, Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK

[5] Human Nutrition Unit, Department of Veterinary Science, University of Parma, Parma, Italy

[6] School of Advanced Studies on Food and Nutrition, University of Parma, Parma, Italy

[*] Corresponding authors: Via Volturno 39, 43125 Parma, Italy; Tel.: +39 0521-903841; E-mail: daniele.delrio@unipr.it; Franklin Wilkins Building, 150 Stamford Street, SE1 9NH London, UK; Tlf: +44 20 7848 4349; E-mail: ana.rodriguez-mateos@kcl.ac.uk

The content of this section has also been published in Food & Function (Favari C, Mena P, Curti C, et al (2020) Kinetic profile and urinary excretion of phenyl-γ-valerolactones upon consumption of cranberry: a dose–response relationship. Food Funct 11:3975–3985. https://doi.org/10.1039/D0FO00806K).

## Abstract

Cranberries are a rich source of poly(phenols), mainly monomeric and oligomeric flavan-3-ols. However, information on the appearance of their main circulating microbial metabolites, namely phenyl-γ-valerolactones and phenylvaleric acid, is lacking despite its relevance to understanding the health effects attributed to cranberries. The aim of this study was to evaluate the absorption, metabolism and urinary excretion of cranberry flavan-3-ols through the targeted analysis of phenyl-γ-valerolactones and their related phenylvaleric acids, considering also their potential as biomarkers of flavan-3-ol intake and inter-individual variability in their appearance in plasma and urine. A six-arm acute crossover, randomized, double-blinded, controlled intervention trial was performed in ten healthy males who consumed a cranberry juice drink (375, 716, 1131, 1396, 1741 mg of total flavan-3-ols) or an isocaloric control drink with one-week washout. Plasma and urine were analyzed by UHPLC-ESI-QqQ-MS/MS and 22 compounds were identified. Glucuronide and sulfate conjugates of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone were the main circulating and excreted metabolites after cranberry juice intake, with glucuronidation appearing to be the most favorable conjugation route. These compounds reached maximum plasma concentration at about 4-6 h. Plasma and urinary concentrations of the sum of the metabolites increased in relation to the amounts of cranberry flavan-3-ols provided by the drink, showing a clear and linear dose-dependent relationship and underscoring their potential as biomarkers of flavan-3-ol intake. A high inter-individual variability in circulating and urinary metabolite levels was observed and, interestingly, some subjects seemed to display a greater efficiency in metabolizing flavan-3-ols and producing phenyl-γ-valerolactones.

**Keywords**

Cranberry; phenolic metabolites; flavan-3-ols; proanthocyanins; phenyl-γ-valerolactones

## 3.1.1. Introduction

Growing evidence from human intervention studies reveals beneficial actions of cranberry (*Vaccinium macrocarpon Ait.*) consumption on human health [1–4], mainly related to prevention of urinary tract infections [5,6] and improvements in vascular function [7]. These health benefits have been attributed to cranberry phenolic content, which is mainly constituted by flavan-3-ols, flavonols, phenolic acids and anthocyanins [8]. In particular, flavan-3-ols, in their monomeric, but mainly oligomeric form (the so called proanthocyanidins) are the most abundant compounds in cranberries [9]. It has to be pointed out that the ability of these compounds to impart positive health effects is closely related to their bioavailability [5,10,11], as flavan-3-ols are poorly absorbed in the upper gastrointestinal tract and mostly reach the large intestine, where they are extensively metabolized by the host microbiota into smaller metabolites, namely phenyl-γ-valerolactones (PVLs), phenylvaleric acids (PVAs) and other low molecular weight phenolics, such as phenylpropionic, benzoic and cinnamic acid derivatives [12]. These smaller phenolics can be absorbed and in turn subjected to phase II metabolism at colonocyte and/or hepatocyte level, in order to produce conjugated metabolites (as sulfate, glucuronide, methoxy or combination thereof) that can circulate through the system and then be excreted in urine. These compounds are therefore the ones potentially available to target tissues and organs [13] and may have an effect on (patho)physiological scenarios where cranberry may show protective effects [7,14,15].

In particular, conjugated PVLs and their related acids (PVAs) are the main circulating metabolites of flavan-3-ols of colonic origin, and are specific for this class of compounds [12]. However, information on their production after cranberry intake is limited [16,17]. Also, little is known about whether they follow a dose-dependent response *in vivo* at different levels of intake of cranberry flavan-3-ols. Therefore, the principal aim of this work was to evaluate the absorption, metabolism and urinary excretion of cranberry flavan-3-ols through the targeted analysis of PVLs and their related PVAs after consumption of a cranberry juice containing increasing amounts of total flavan-3-ols (TF). This may serve to further support the suitability of PVLs as biomarkers of flavan-3-ol intake. Moreover, since the gut microbiota composition varies among individuals, it may result in differences in the production of PVLs and PVAs [18,19]. Consequently, the second aim of the present study was to assess the extent of this inter-individual variability.

## 3.1.2. Material and Methods

*Materials*

All chemicals and solvents used in this study were of analytical grade. 5-Phenyl-γ-valerolactone-3'-sulfate, 5-phenyl-γ-valerolactone-4'-sulfate, 5-phenyl-γ-valerolactone-3'-glucuronide, 5-(3',4'-dihydroxyphenyl)-γ-valerolactone, 5-(3',5' dihydroxyphenyl)-γ-valerolactone, 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-sulfate, 5-(5'-hydroxyphenyl)-γ-valerolactone-3'-sulfate and 5-(5'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide were synthesized in house [20,21]. All solvents and reagents were purchased from Sigma (St. Louis, MO, USA). Ultrapure water from MilliQ system (Millipore, Bedford, MA, USA) was used throughout the experiment.

*Study Design*

The study protocol has already been described [7,17]. Briefly, ten healthy men aged between 18 and 35 years took part in a six-arm acute crossover, randomized, double-blinded, controlled intervention trial registered under the NIH ClinicalTrials.gov website (NCT02517775). The study was conducted in accordance to the guidelines stated in the current revision of the Declaration of Helsinki, and informed consent was obtained for all subjects. All procedures involving human subjects were approved by the University of Dusseldorf Research Ethics Committee (ref: 14-012). Volunteers were asked to consume a cranberry drink containing 375, 716, 1131, 1396, 1741 mg of total flavan-3-ols (TF) or an isocaloric control (0 mg TF) drink with one-week washout. Participants were instructed to follow a low-(poly)phenol diet for 3 days before and during the study day and had to fast for 12 h before the study day. Blood samples were taken before and at 1, 2, 4, 6, 8 and 24 h after drink intake; urine samples were collected at baseline, between 0-8 h and 8-24 h after drink intake.

*UHPLC-ESI-QqQ-MS/MS analysis*

Plasma and urine samples were prepared according to validated protocols [17,20] and then both biological samples were analyzed through UHPLC DIONEX Ultimate 3000 fitted with a TSQ Vantage triple quadrupole mass spectrometer (Thermo Fisher Scientific Inc., San Jose, CA, USA) equipped with a heated-electrospray ionization source (H-ESI-II; Thermo Fisher Scientific Inc.). Chromatographic and ionization parameters were set following a validated method optimized for the analysis of PVLs [20]. Metabolite identification was carried out by comparison of the retention time with in-house synthesized standards and/or MS/MS fragmentation patterns. Up to 72 compounds were monitored in selective reaction monitoring (SRM) mode. Quantification was performed with calibration curves of standards, when available. When not available, metabolites

were quantified with the most structurally similar compound, as in the case of 5-phenyl-γ-valerolactone-4'-glucuronide, quantified with its isomer 5-phenyl-γ-valerolactone-3'-glucuronide; 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide, 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-glucuronide, 5-phenyl-γ-valerolactone-methoxy-glucuronide isomer (3',4') and 5-phenyl-γ-valerolactone-sulfate-glucuronide isomer (3',4'), quantified with 5-(5'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide; and the 5-phenyl-γ-valerolactone-methoxy-sulfate (3',4') isomers 1 and 2, quantified with 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-sulfate.

*Data analysis*

All data were expressed as mean values ± SEM. Kinetic parameters were calculated using the PKSolver add-in software for Microsoft Excel [22]. Associations between total plasma concentrations or urinary excretions of quantified compounds and flavan-3-ol content in the intervention drinks, both using average and individual values, were assessed with linear regression models using Microsoft Excel. Inter-individual variability was assessed according to the recommendations suggested by the COST Action POSITIVe "Interindividual variation in response to consumption of plant food bioactives and determinants involved" [23].

## 3.1.3. Results

*Plasma kinetics*

A comprehensive identification and quantification of PVLs and PVAs was performed in both plasma and urine, thanks to specifically synthesized standards. Twenty-two compounds were identified (Table 3). Phase-II conjugated forms of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone were the most representative compounds, followed by derivatives of 5-(monohydroxyphenyl)-γ-valerolactone (3' or 4' isomers). Free PVLs and derivatives of 5-(3',5'-dihydroxyphenyl)-γ-valerolactone were present in minor amounts, while phase-II conjugated forms of 5-(3',4',5'-trihydroxyphenyl)-γ-valerolactone and 5-(3',4',5'-trihydroxyphenyl)valeric acid were found at trace levels in just a few samples and consequently they were identified but not quantified.

**Table 3.** Mass spectral characteristic of PVLs and PVAs identified in plasma and urine

| Compound | Retention time (min) | Parent ion [M-H]- (*m/z*) | Product ions | |
|---|---|---|---|---|
| | | | Quantifier (*m/z*) | Qualifier(s) (*m/z*) |
| 5-(dihydroxyphenyl)-γ-valerolactone-glucuronide (3',4',5') | 1.57 | 399 | 223 | 175 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 5-(5′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide | 2.38 | 383 | 207 | 113 | | | |
| 5-(3′,5′-dihydroxyphenyl)-γ-valerolactone | 3.78 | 207 | 123 | 163 | 123 | 122 | 121 |
| 5-(dihydroxyphenyl)-γ-valerolactone-sulfate (3′,4′,5′) | 3.80 | 303 | 223 | 179 | 138 | | |
| 5-phenyl-γ-valerolactone-4′-glucuronide | 3.87 | 367 | 191 | 113 | 147 | | |
| 5-(3′-hydroxyphenyl)-γ-valerolactone-4′-glucuronide | 3.92 | 383 | 207 | 113 | | | |
| 5-phenyl-γ-valerolactone-sulfate-glucuronide isomer (3′,4′) | 4.03 | 463 | 287 | 207 | 383 | 163 | 175 |
| 5-(4′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide | 4.08 | 383 | 207 | 113 | | | |
| 4-hydroxy-5-(hydroxyphenyl)valeric acid-sulfate (3′/4′) isomer 1 | 4.15 | 305 | 225 | 207 | | | |
| 4-hydroxy-5-(hydroxyphenyl)valeric acid-glucuronide (3′/4′) | 4.15 | 401 | 225 | 207 | 175 | 113 | |
| 5-(3′,4′-dihydroxyphenyl)-γ-valerolactone | 4.17 | 207 | 122 | 163 | 122 | 121 | 123 |
| 5-(5′-hydroxyphenyl)-γ-valerolactone-3′-sulfate | 4.26 | 287 | 207 | 163 | 122 | | |
| 5-phenyl-γ-valerolactone-methoxy-glucuronide isomer (3′,4′) | 4.30 | 397 | 221 | 206 | 113 | | |
| 5-hydroxyphenyl-γ-valerolactone-methoxy-glucuronide (3′,4′,5′) | 4.31 | 413 | 237 | 222 | 175 | | |
| 5-phenyl-γ-valerolactone-3′-glucuronide | 4.33 | 367 | 191 | 113 | 147 | | |
| 4-hydroxy-5-(hydroxyphenyl)valeric acid-sulfate (3′/4′) isomer 2 | 4.34 | 305 | 225 | 207 | | | |
| 5-(hydroxyphenyl)-γ-valerolactone-methoxy-sulfate (3′,4′,5′) | 4.47 | 317 | 237 | 222 | | | |
| 5-(hydroxyphenyl)-γ-valerolactone-sulfate (3′,4′ isomers) | 4.88 | 287 | 207 | 163 | 122 | | |
| 5-phenyl-γ-valerolactone-4′-sulfate | 4.94 | 271 | 191 | 147 | 106 | | |
| 5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 1 | 5.02 | 301 | 221 | 206 | 125 | | |
| 5-phenyl-γ-valerolactone-3′-sulfate | 5.10 | 271 | 191 | 147 | 106 | | |
| 5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 2 | 5.20 | 301 | 221 | 206 | 125 | | |

Thirteen compounds were quantified in plasma. Quantitatively, 5-(4′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide, 5-(3′-hydroxyphenyl)-γ-valerolactone-4′-glucuronide and 5-

(hydroxyphenyl)-γ-valerolactone-sulfate (which is the sum of the two 3',4' isomers) were the most representative compounds (Table 4).

**Table 4**. Kinetic parameters of PVLs detected in plasma after cranberry juice consumption (mean ± SEM, *n* = 10). The coefficient of variation (CV%) of $C_{max}$ is reported in brackets.

| Compound | $C_{max}$ (nM) | $T_{max}$ (h) | $AUC_{0-24}$ (nM h) |
|---|---|---|---|
| **5-phenyl-γ-valerolactone-3'-sulfate** | | | |
| Treatment 375 mg TF | 1.01 ± 0.88 (36%) | 0.20 ± 0.20 | 11.60 ± 11.39 (32%) |
| Treatment 716 mg TF | 0.58 ± 0.33 (56%) | 1.00 ± 0.54 | 1.10 ± 0.64 (54%) |
| Treatment 1131 mg TF | 1.69 ± 1.69 (32%) | 0.20 ± 0.20 | 10.87 ± 10.87 (32%) |
| Treatment 1396 mg TF | 0.87 ± 0.45 (61%) | 1.00 ± 0.54 | 4.08 ± 2.76 (47%) |
| Treatment 1741 mg TF | 0.88 ± 0.53 (53%) | 0.80 ± 0.44 | 3.89 ± 2.16 (57%) |
| **5-phenyl-γ-valerolactone-4'-sulfate** | | | |
| Treatment 375 mg TF | 0.13 ± 0.09 (46%) | 2.60 ± 2.39 | 0.52 ± 0.41 (41%) |
| Treatment 716 mg TF | 0.42 ± 0.24 (55%) | 2.60 ± 2.38 | 2.70 ± 1.86 (46%) |
| Treatment 1131 mg TF | 0.60 ± 0.43 (45%) | 2.40 ± 2.40 | 3.39 ± 3.17 (34%) |
| Treatment 1396 mg TF | 0.23 ± 0.12 (58%) | 2.70 ± 2.38 | 0.94 ± 0.61 (48%) |
| Treatment 1741 mg TF | 0.61 ± 0.24 (80%) | 3.20 ± 2.33 | 2.48 ± 1.26 (62%) |
| **5-phenyl-γ-valerolactone-3'-glucuronide** | | | |
| Treatment 375 mg TF | 0.62 ± 0.62 (32%) | 0.60 ± 0.60 | 2.60 ± 2.60 (32%) |
| Treatment 716 mg TF | 0.87 ± 0.45 (62%) | 1.20 ± 0.80 | 1.73 ± 1.08 (51%) |

| | | | |
|---|---|---|---|
| Treatment 1131 mg TF | 2.16 ± 1.24 (55%) | 0.50 ± 0.27 | 4.93 ± 2.98 (52%) |
| Treatment 1396 mg TF | 2.61 ± 1.32 (62%) | 2.00 ± 0.94 | 11.04 ± 6.66 (52%) |
| Treatment 1741 mg TF | 2.84 ± 2.14 (42%) | 0.60 ± 0.43 | 12.44 ± 9.91 (40%) |
| **5-phenyl-γ-valerolactone-4'-glucuronide** | | | |
| Treatment 375 mg TF | 2.98 ± 0.96 (98%) | 2.70 ± 2.37 | 7.52 ± 2.63 (91%) |
| Treatment 716 mg TF | 4.78 ± 0.95 (160%) | 4.40 ± 2.30 | 20.66 ± 7.93 (82%) |
| Treatment 1131 mg TF | 12.01 ± 1.75 (216%) | 6.20 ± 3.01 | 50.60 ± 12.10 (132%) |
| Treatment 1396 mg TF | 10.83 ± 1.63 (210%) | 3.90 ± 2.26 | 28.45 ± 3.84 (235%) |
| Treatment 1741 mg TF | 9.62 ± 3.12 (97%) | 1.80 ± 0.59 | 26.74 ± 8.07 (105%) |
| **5-(hydroxyphenyl)-γ-valerolactone-sulfate (3',4' isomers)** | | | |
| Treatment 375 mg TF | 49.27 ± 16.19 (96%) | 3.90 ± 0.91 | 377.90 ± 93.91 (127%) |
| Treatment 716 mg TF | 87.38 ± 23.75 (116%) | 2.70 ± 0.62 | 737.89 ± 158.11 (148%) |
| Treatment 1131 mg TF | 70.46 ± 9.29 (240%) | 6.90 ± 2.04 | 812.22 ± 164.47 (156%) |
| Treatment 1396 mg TF | 127.82 ± 26.43 (153%) | 4.40 ± 0.65 | 1265.28 ± 312.37 (128%) |
| Treatment 1741 mg TF | 166.44 ± 53.95 (98%) | 4.20 ± 0.76 | 1712.65 ± 624.63 (87%) |
| **5-(5'-hydroxyphenyl)-γ-valerolactone-3'-sulfate** | | | |
| Treatment 375 mg TF | 0.64 ± 0.38 (53%) | 0.80 ± 0.59 | 0.99 ± 0.70 (45%) |

| | | | |
|---|---|---|---|
| Treatment 716 mg TF | 0.08 ± 0.08 (32%) | 0.10 ± 0.10 | 0.08 ± 0.08 (32%) |
| Treatment 1131 mg TF | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| Treatment 1396 mg TF | 0.12 ± 0.12 (32%) | 0.00 ± 0.00 | 0.06 ± 0.06 (32%) |
| Treatment 1741 mg TF | 0.00 ± 0.00 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| **5-(4′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide** | | | |
| Treatment 375 mg TF | 36.47 ± 6.59 (175%) | 4.70 ± 0.67 | 309.80 ± 43.95 (223%) |
| Treatment 716 mg TF | 85.66 ± 10.58 (256%) | 4.70 ± 0.67 | 801.06 ± 91.50 (277%) |
| Treatment 1131 mg TF | 116.83 ± 17.73 (208%) | 6.20 ± 0.47 | 1098.50 ± 130.97 (265%) |
| Treatment 1396 mg TF | 153.12 ± 20.07 (241%) | 5.80 ± 0.36 | 1572.50 ± 196.48 (253%) |
| Treatment 1741 mg TF | 175.69 ± 18.74 (297%) | 6.40 ± 0.50 | 1871.47 ± 222.26 (266%) |
| **5-(3′-hydroxyphenyl)-γ-valerolactone-4′-glucuronide** | | | |
| Treatment 375 mg TF | 10.77 ± 3.13 (109%) | 4.10 ± 0.90 | 70.57 ± 20.76 (108%) |
| Treatment 716 mg TF | 20.46 ± 2.43 (266%) | 6.50 ± 2.03 | 152.08 ± 23.34 (206%) |
| Treatment 1131 mg TF | 25.64 ± 3.74 (217%) | 5.60 ± 0.50 | 217.85 ± 36.93 (187%) |
| Treatment 1396 mg TF | 35.48 ± 5.71 (197%) | 6.20 ± 0.55 | 365.04 ± 59.58 (194%) |
| Treatment 1741 mg TF | 38.60 ± 7.93 (154%) | 8.00 ± 1.86 | 388.55 ± 60.65 (203%) |

| | | | |
|---|---|---|---|
| **5-(5′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide** | | | |
| Treatment 375 mg TF | 3.24 ± 1.45 (71%) | 1.90 ± 0.90 | 24.24 ± 11.99 (64%) |
| Treatment 716 mg TF | 4.01 ± 1.39 (91%) | 3.40 ± 1.27 | 32.81 ± 13.31 (78%) |
| Treatment 1131 mg TF | 1.42 ± 0.99 (45%) | 1.60 ± 1.07 | 12.79 ± 8.89 (45%) |
| Treatment 1396 mg TF | 4.80 ± 1.51 (100%) | 4.00 ± 1.15 | 26.07 ± 9.33 (88%) |
| Treatment 1741 mg TF | 2.10 ± 1.11 (60%) | 2.20 ± 1.13 | 12.48 ± 7.12 (55%) |
| **5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 1** | | | |
| Treatment 375 mg TF | 0.36 ± 0.16 (69%) | 1.50 ± 0.83 | 1.90 ± 1.25 (48%) |
| Treatment 716 mg TF | 0.72 ± 0.36 (63%) | 3.80 ± 2.33 | 3.56 ± 1.31 (86%) |
| Treatment 1131 mg TF | 0.65 ± 0.16 (125%) | 4.70 ± 2.33 | 3.61 ± 1.73 (66%) |
| Treatment 1396 mg TF | 0.89 ± 0.43 (65%) | 4.00 ± 0.94 | 5.14 ± 2.45 (66%) |
| Treatment 1741 mg TF | 1.08 ± 0.27 (128%) | 4.70 ± 0.90 | 7.31 ± 2.85 (81%) |
| **5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 2** | | | |
| Treatment 375 mg TF | 0.50 ± 0.13 (123%) | 2.00 ± 0.87 | 2.80 ± 1.28 (69%) |
| Treatment 716 mg TF | 0.61 ± 0.13 (143%) | 2.50 ± 0.75 | 3.20 ± 1.22 (83%) |
| Treatment 1131 mg TF | 0.72 ± 0.16 (140%) | 6.40 ± 2.23 | 7.57 ± 1.87 (128%) |
| Treatment 1396 mg TF | 0.96 ± 0.27 (112%) | 5.80 ± 0.81 | 8.74 ± 2.62 (105%) |

| | | | |
|---|---|---|---|
| Treatment 1741 mg TF | 1.05 ± 0.33 (101%) | 6.20 ± 2.16 | 10.10 ± 3.68 (87%) |
| **5-phenyl-γ-valerolactone-methoxy-glucuronide isomer (3',4')** | | | |
| Treatment 375 mg TF | 3.66 ± 1.59 (73%) | 0.10 ± 0.10 | 14.69 ± 8.62 (54%) |
| Treatment 716 mg TF | 5.45 ± 1.86 (93%) | 3.40 ± 1.15 | 41.34 ± 11.39 (115%) |
| Treatment 1131 mg TF | 7.80 ± 1.28 (192%) | 7.00 ± 2.16 | 61.49 ± 11.10 (175%) |
| Treatment 1396 mg TF | 9.76 ± 1.74 (178%) | 6.60 ± 0.52 | 99.90 ± 20.33 (155%) |
| Treatment 1741 mg TF | 10.54 ± 2.79 (120%) | 3.30 ± 0.92 | 78.39 ± 15.47 (160%) |
| **5-phenyl-γ-valerolactone-sulfate-glucuronide isomer (3',4')** | | | |
| Treatment 375 mg TF | 6.28 ± 4.74 (42%) | 1.20 ± 0.68 | 12.16 ± 9.47 (41%) |
| Treatment 716 mg TF | 2.23 ± 1.00 (71%) | 2.20 ± 0.92 | 5.72 ± 2.86 (63%) |
| Treatment 1131 mg TF | 2.67 ± 1.13 (75%) | 4.20 ± 2.39 | 22.36 ± 11.05 (64%) |
| Treatment 1396 mg TF | 3.93 ± 1.32 (94%) | 3.30 ± 1.14 | 30.09 ± 11.48 (83%) |
| Treatment 1741 mg TF | 3.53 ± 1.54 (72%) | 2.80 ± 0.95 | 24.69 ± 12.76 (61%) |

In particular, 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide was present at the highest concentration, and its kinetic profiles after each of the six treatments are shown in Figure 11a. No increase in the plasma concentration was observed after the consumption of the control drink without flavan-3-ols, as expected, while after the intake of the cranberry juice drinks, there was a rapid increase in the plasma concentration of the metabolite, reaching its maximum concentration

($C_{max}$) at about 5 h post intake for treatments with lower amounts of flavan-3-ols and at 6 h post intake for treatments with higher amounts. After reaching its maximum, plasma concentration

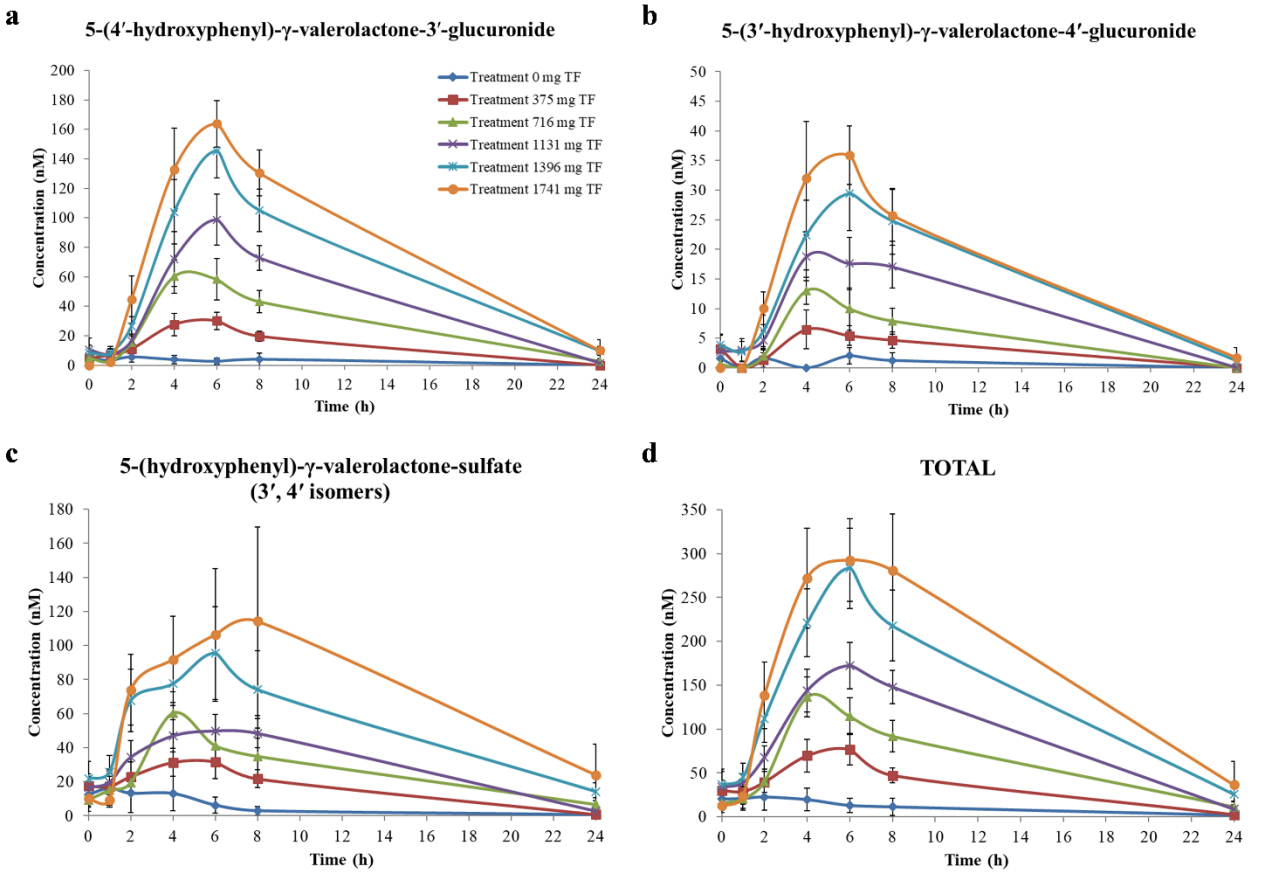



**Figure 11.** Kinetic profile of 5-(4′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide (**a**), 5-(3′-hydroxyphenyl)-γ-valerolactone-4′-glucuronide (**b**), 5-(hydroxyphenyl)-γ-valerolactone-sulfate (3′, 4′) (**c**) and of the sum of quantified PVLs (**d**) for the different treatments. Data are expressed as nM, mean ± SEM (*n* = 10). Treatment 0 TF corresponds to the control drink, other treatments to cranberry juice drinks with growing amounts of total flavan-3-ols.

decreased, going back to baseline levels at 24 h. The isomer of this metabolite, 5-(3′-hydroxyphenyl)-γ-valerolactone-4′-glucuronide, was present at lower concentration but, similarly, its plasma concentration rapidly increased after cranberry juice intake, reaching its maximum after 4 h for treatments with lower amounts of flavan-3-ols and after 6 h for treatments with higher amounts (Figure 11b). After reaching its $C_{max}$, the concentration decreased, going back to basal levels at 24 h post intake.

A similar trend was also observed for 5-(hydroxyphenyl)-γ-valerolactone-sulfate (the sum of the two 3′,4′ isomers), which was one of the most abundant compounds despite the high inter-individual variability and greater differences between treatments (Figure 11c). This metabolite reached its $C_{max}$ at about 6 h for treatments 2, 4, 5, while peaked sooner, at 4 h, for treatment 3 and later, at 8 h, for treatment 6. The kinetic profiles of the other derivatives of 5-(3′,4′-dihydroxyphenyl)-γ-

valerolactone did not differ from the previous ones, despite the much lower plasmatic concentration reached by these metabolites. Differently, derivatives of 5-(3'-hydroxyphenyl)-γ-valerolactone and 5-(4'-hydroxyphenyl)-γ-valerolactone showed an earlier $T_{max}$ (1-4 h), with not so clear kinetic profiles (Supplementary Figure 1 e-n). This could be likely due to the low circulating amounts and the extensive variability recorded for these minor metabolites.

When considering the kinetic profiles of the sum of quantified PVLs, the trend observed for the main metabolites is consistent, with $C_{max}$ reached at about 6 h for almost all the treatments (Figure 11d). As a sum, quantified PVLs in plasma attained maximum levels ranging from 76.9 ± 57.1 nM to 292.6 ± 149.7 nM after consumption of increasing amounts of total flavan-3-ols.

*Urinary excretion*

The type and amount of PVLs excreted in urine over the 24 h after cranberry juice consumption (treatment 375-1741 mg TF) are shown in Table 5. Consistently with the observations made for plasma, 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide, 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-glucuronide and 5-(hydroxyphenyl)-γ-valerolactone-sulfate were excreted in high amounts. The cumulative excretion of PVLs for the different urine samples collected is presented in Supplementary Figure 2a-n.

**Table 5.** Cumulative excretion of PVLs in urine after 24h following cranberry juice consumption. Data are expressed in µmol (mean ± SEM, *n* = 10). Coefficient of variation (CV%) is reported in brackets.

| Compound | Treatment 375 mg TF | Treatment 716 mg TF | Treatment 1131 mg TF | Treatment 1396 mg TF | Treatment 1741 mg TF |
|---|---|---|---|---|---|
| **5-phenyl-γ-valerolactone-3'-sulfate** | 0.04 ± 0.02 (63%) | 0.03 ± 0.01 (77%) | 0.19 ± 0.17 (35%) | 0.11 ± 0.09 (40%) | 0.05 ± 0.02 (88%) |
| **5-phenyl-γ-valerolactone-3'-glucuronide** | 0.29 ± 0.16 (59%) | 0.26 ± 0.09 (92%) | 0.58 ± 0.46 (40%) | 0.43 ± 0.20 (68%) | 0.43 ± 0.18 (75%) |
| **5-phenyl-γ-valerolactone-4'-glucuronide** | 0.59 ± 0.27 (68%) | 0.61 ± 0.21 (94%) | 0.85 ± 0.37 (72%) | 0.82 ± 0.21 (127%) | 1.13 ± 0.18 (197%) |
| **5-(3',4'-dihydroxyphenyl)-γ-valerolactone** | 0.05 ± 0.03 (53%) | 0.11 ± 0.04 (79%) | 0.10 ± 0.04 (89%) | 0.21 ± 0.07 (98%) | 0.31 ± 0.09 (107%) |

| | | | | | |
|---|---|---|---|---|---|
| **5-(hydroxyphenyl)-γ-valerolactone-sulfate (3′,4′)** | 3.16 ± 0.93 (107%) | 5.68 ± 1.69 (106%) | 6.83 ± 2.07 (104%) | 10.71 ± 2.39 (142%) | 10.35 ± 2.57 (127%) |
| **5-(5′-hydroxyphenyl)-γ-valerolactone-3′-sulfate** | 0.02 ± 0.01 (60%) | 0.02 ± 0.01 (89%) | 0.05 ± 0.02 (88%) | 0.08 ± 0.04 (64%) | 0.03 ± 0.01 (70%) |
| **5-(4′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide** | 6.21 ± 1.47 (134%) | 13.98 ± 1.37 (322%) | 17.37 ± 3.97 (138%) | 23.76 ± 3.07 (245%) | 30.41 ± 3.24 (297%) |
| **5-(3′-hydroxyphenyl)-γ-valerolactone-4′-glucuronide** | 1.79 ± 0.51 (111%) | 3.66 ± 0.51 (225%) | 3.88 ± 0.95 (129%) | 7.96 ± 2.30 (110%) | 7.00 ± 0.92 (239%) |
| **5-(5′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide** | 0.55 ± 0.14 (121%) | 0.75 ± 0.16 (146%) | 0.46 ± 0.16 (93%) | 0.71 ± 0.20 (111%) | 0.82 ± 0.20 (129%) |
| **5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 1** | 0.01 ± 0.00 (70%) | 0.02 ± 0.01 (121%) | 0.02 ± 0.01 (96%) | 0.03 ± 0.01 (111%) | 0.02 ± 0.01 (119%) |
| **5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 2** | 0.03 ± 0.01 (148%) | 0.07 ± 0.02 (110%) | 0.07 ± 0.03 (74%) | 0.08 ± 0.02 (145%) | 0.10 ± 0.02 (129%) |
| **5-phenyl-γ-valerolactone-methoxy-glucuronide isomer (3′,4′)** | 0.66 ± 0.17 (121%) | 1.15 ± 0.19 (190%) | 1.07 ± 0.17 (194%) | 1.53 ± 0.18 (273%) | 1.55 ± 0.18 (268%) |
| **5-phenyl-γ-valerolactone-sulfate-glucuronide isomer (3′,4′)** | 1.84 ± 0.49 (120%) | 3.37 ± 0.53 (199%) | 4.06 ± 1.42 (90%) | 4.89 ± 0.99 (156%) | 6.16 ± 1.11 (175%) |
| **Total** | 15.25 ± 3.50 (138%) | 29.71 ± 3.25 (289%) | 35.52 ± 8.92 (126%) | 51.32 ± 6.96 (233%) | 58.42 ± 6.48 (285%) |

When considering the consumption of the cranberry juice with highest flavan-3-ol content (1741 mg TF), 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide was the main excreted metabolite for all the subjects, followed by 5-(hydroxyphenyl)-γ-valerolactone-sulfate or 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-glucuronide (Figure 12a and 12b). Overall, derivatives of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone were the main excreted metabolites when compared to derivatives of 5-(3'-hydroxyphenyl)-γ-valerolactone and 5-(4'-hydroxyphenyl)-γ-valerolactone (Figure 12c). Glucuronide conjugates were, for all subjects, more abundant than sulfate (Figure 12d).

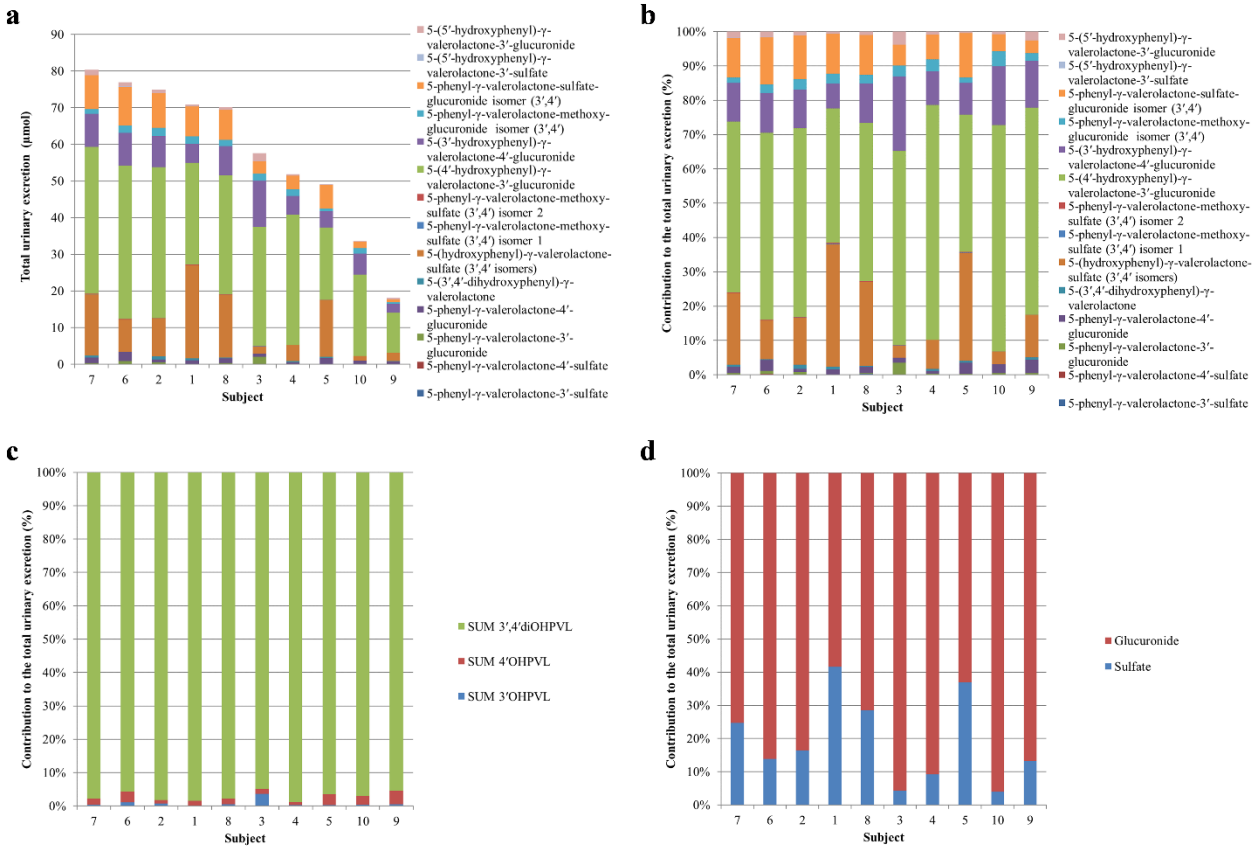



**Figure 12.** (**a**) Cumulative urinary excretion of PVLs at 24 h after treatment with higher flavan-3-ols content (1741 mg TF). Data are expressed as µmol. (**b**) Relative contribution of each PVL to the total urinary excretion at 24 h after treatment with higher flavan-3-ols content (1741 mg TF). (**c**) Relative contribution of derivatives of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone compared to derivatives of 5-(3'-hydroxyphenyl)-γ-valerolactone and 5-(4'-hydroxyphenyl)-γ-valerolactone to the total urinary excretion at 24 h after treatment with higher flavan-3-ols content (1741 mg TF). (**d**) Relative contribution of glucuronide and sulfate conjugates to the total urinary excretion at 24 h after treatment with higher flavan-3-ols content (1741 mg TF).

### *Plasma and urinary dose-response*

From the kinetic profiles of both the main and total metabolites, a dose-dependent effect was observed, as the metabolite concentration in plasma increased with increasing amounts of cranberry flavan-3-ols provided by the intervention drinks (Figure 11 a-b-c-d). When plotting the

area under the curve ($AUC_{0-24}$) values of the sum of PVLs versus the flavan-3-ol content in the intervention drinks, a positive linear dose-response relation was clearly observed (regression of $R^2$ = 0.98) (Figure 13).

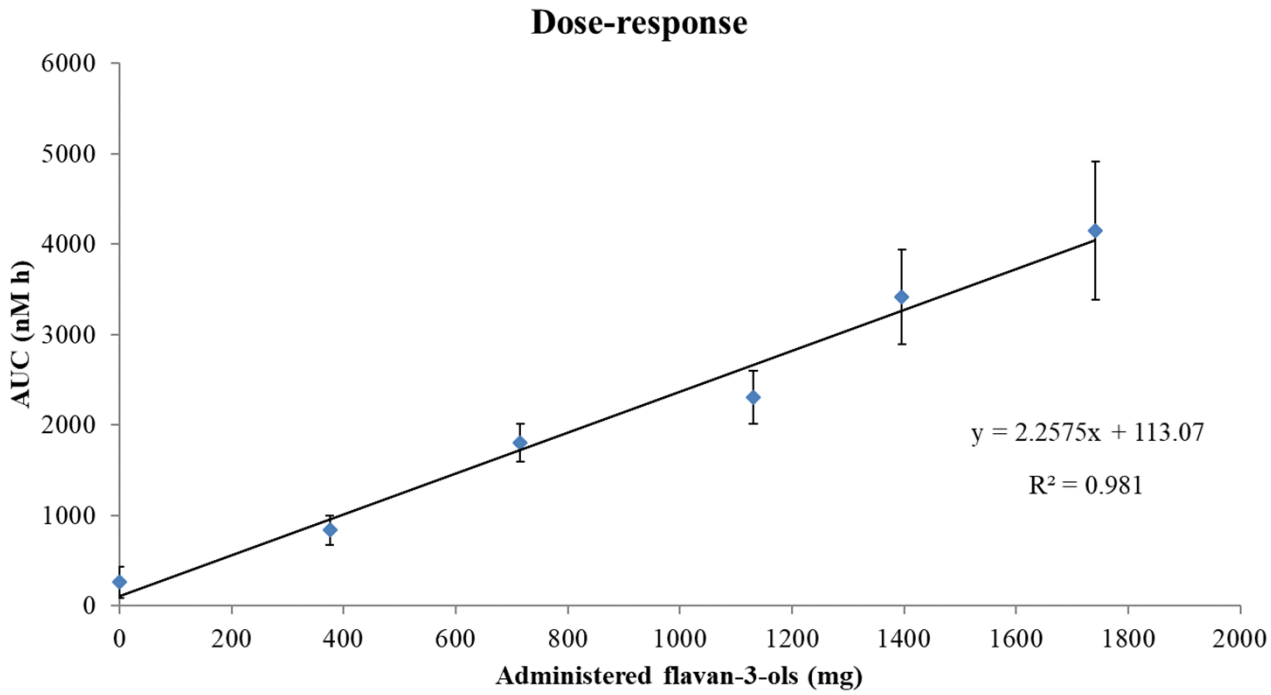



**Figure 13.** Relation between total amount of ingested flavan-3-ols and Area under the curve ($AUC_{0-24}$) of the concentration in plasma of PVLs over 24h. Data are expressed as nM*h, mean ± SEM, *n* = 10.

Similarly to what was observed in plasma, the urinary excretion of the main metabolites increased with increasing amounts of cranberry flavan-3-ols provided by the intervention drinks, showing a dose-dependent effect. When plotting the total urinary excretion of PVLs over 24h versus the flavan-3-ol content in the intervention drinks, a positive linear dose-response relation is clearly observed (regression of $R^2$ = 0.98) (Figure 14).

*Inter-individual variability*

A high variability among subjects was observed in the production of PVLs, as indicated by the high coefficients of variation (CV%) reported in brackets in Table 4 for plasma and in Table 5 for the urinary excretion.

Following the highest flavan-3-ol intake (treatment 1741 mg TF), $AUC_{0-24}$ values of the sum of PVLs varied between 1034 nM*h and 8839 nM*h and total urinary excretion of PVLs ranged from 18.07 µmol to 80.34 µmol. Every subject produced different amounts and proportions of each metabolite (Figure 12a and 12b), with subjects 9 and 10 excreting lower amounts of PVLs compared to others.

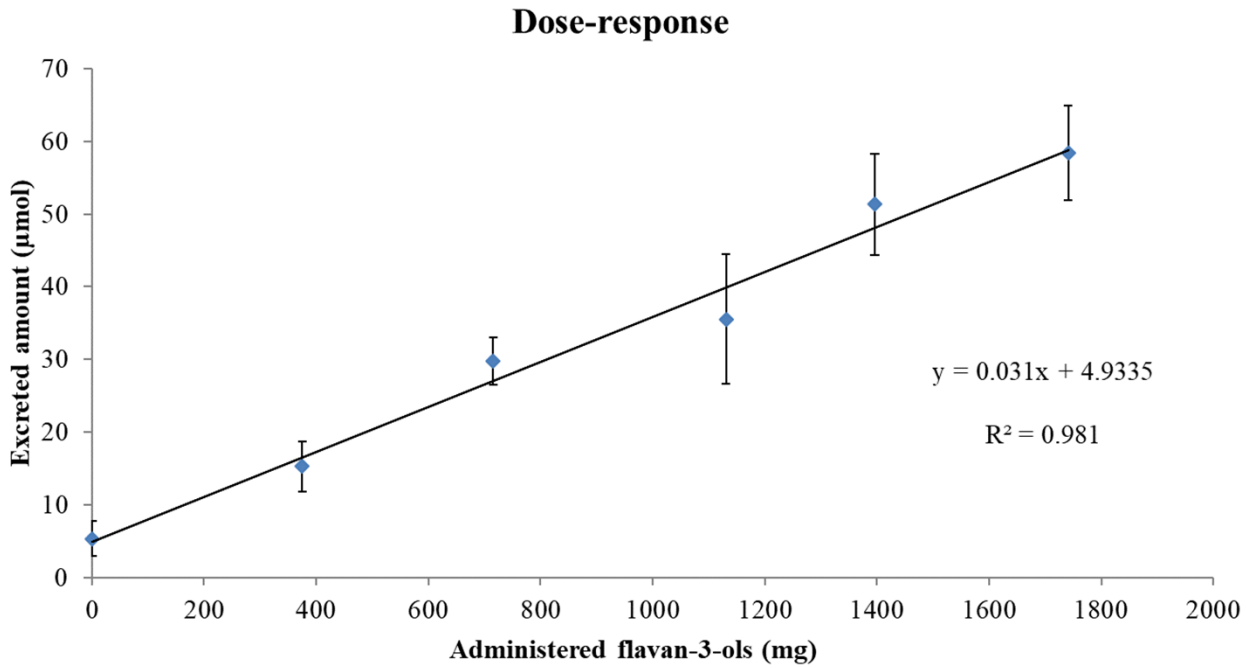



**Figure 14.** Relation between total amount of ingested flavan-3-ols and cumulative quantity of excreted metabolites. Data are expressed as µmol, mean ± SEM, *n* = 10.

For all the subjects, conjugates of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone were prevalent respect to conjugates of 5-(3'-hydroxyphenyl)-γ-valerolactone and 5-(4'-hydroxyphenyl)-γ-valerolactone (Figure 12b), with percentages varying from 91% to 98% of total excreted PVLs. In general, derivatives of 5-(3'-hydroxyphenyl)-γ-valerolactone were less present than 5-(4'-hydroxyphenyl)-γ-valerolactone derivatives, with the sole exception of subject 3 who excreted 3.53% of 5-(3'-hydroxyphenyl)-γ-valerolactone derivatives compared to excretions varying from 0.15% to 1.12% for the other subjects. This trend was also confirmed for the rest of the treatments (375-1396 mg TF). Regarding phase II conjugation, glucuronide conjugates were the main metabolites compared to sulfate and methoxy ones (Figure 12b and 12d), with percentages ranging from 51% to 91% of total excreted PVLs.

Interestingly, it was noted that the variation in PVL production among subjects remained consistent, subjects 1, 2 and 7 were almost always the highest PVL producers while subject 9 the lowest PVL producer, irrespective of the dose. This trend was observed for both plasma concentrations and urinary excretions (Figure 15a and 15b) and was confirmed on the basis of the good coefficients of determination observed for individual dose-response relationships ($R^2$>0.75 for most of the individuals, Tables S1a and b in Supplementary Information). The individual relations were obtained plotting the area under the curve ($AUC_{0-24}$) values or the total urinary excretion values of the sum

of quantified PVLs of each subject versus the flavan-3-ol content in the intervention drinks for plasma concentrations (Figure 15a) and urinary excretions (Figure 15b), respectively.

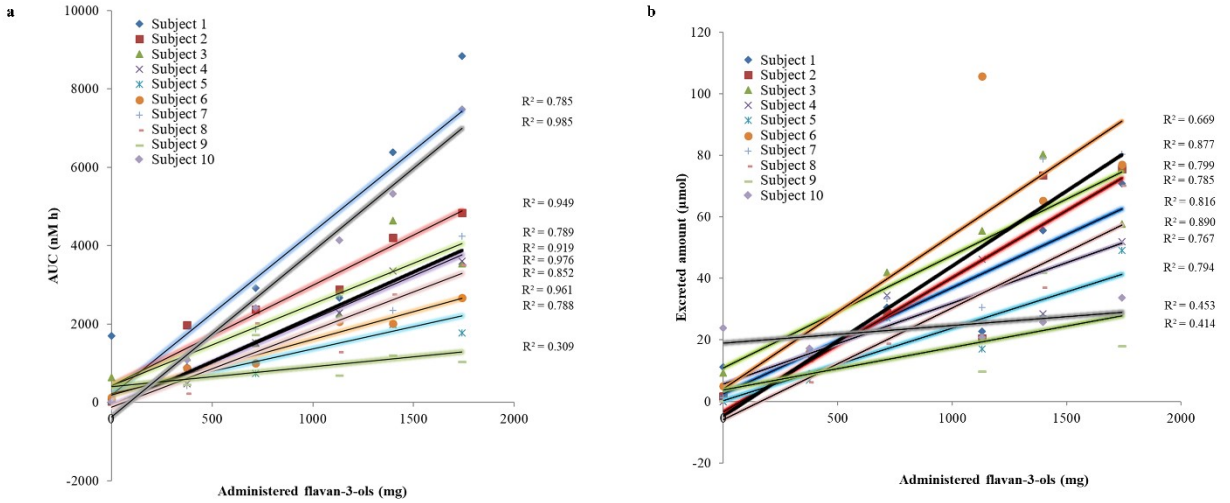



**Figure 15.** (**a**) Inter-individual variability in circulating concentrations of the sum of quantified PVLs after consumption of cranberry juice containing increasing amount of total flavan-3-ols. Data are expressed as Area under the curve (AUC) of the concentration in plasma over 24h in nM*h. (**b**) Inter-individual variability in urinary excretion of the sum of quantified PVLs after consumption of cranberry juice containing increasing amount of total flavan-3-ols. Data are expressed as cumulative quantity of excreted metabolites in µmol.

## 3.1.4. Discussion

In this work, up to 22 compounds among PVLs and PVAs were identified in human biofluids upon consumption of cranberry drinks. Several studies have shown that PVLs concentration exceeds PVAs concentration [12], and this observation was also reflected in our results. Thirteen PVLs were quantified in plasma and urine, while PVAs were only found at trace levels in a few samples. The availability of reference molecules that were specifically synthesized [20,21] allowed the absolute quantification of most of the cranberry-derived PVLs detected (8 compounds), as well as an accurate semi-quantification of the rest of the metabolites (5 compounds). To the best of our knowledge, this is the first study investigating the kinetic profile of cranberry-derived PVLs and PVAs in humans. Glucuronide and sulfate conjugates of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone were the main circulating and excreted metabolites. In particular, 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide was the main metabolite present in plasma and urine for all the subjects, differently from what was observed by Anesi and colleagues [24], who reported higher amounts of 5-(hydroxyphenyl)-γ-valerolactone-sulfate upon consumption of apple flavan-3-ols. Also, Anesi and colleagues found a ratio "5-hydroxyphenyl-γ-valerolactone-sulfate / -glucuronide" either nearly equivalent to 1 or in favour of sulfate-conjugates rather than glucuronide-conjugates. Similar results

to those of the present study, were observed by Cortés-Martín and colleagues in a trial where volunteers were fed mixed nuts [25], although in this case authors did not use a glucuronidated PVL to quantify 5-hydroxyphenyl-γ-valerolactone-glucuronide isomers, and thus data cannot be compared to the present work.

The relationship between the total amount of flavan-3-ols provided by cranberry juice and the levels of PVLs in plasma and urine was also examined. The sum of plasma and urinary PVLs positively correlated with the amount of ingested flavan-3-ols, showing a linear dose-response, with glucuronide and sulfate derivatives of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone being the key metabolites driving the overall correlation. This result suggests that 5-(3',4'-dihydroxyphenyl)-γ-valerolactone, in its sulfate and glucuronide conjugated forms, could serve as biomarker of intake for cranberry flavan-3-ols. These metabolites have already been proposed as biomarker of flavan-3-ol intake in a large epidemiological study [26], but information on cranberry was missing. Cranberry is a rich source of flavan-3-ols, not only monomeric but mainly oligomeric, contributing to the dietary intake of this class of compounds. This study has shown that consumption of cranberry flavan-3-ols leads to the production of the same metabolites originated after intake of cocoa-derived products, wine, and green tea contributing to the pool of circulating PVLs. In addition, our results indicate that sulfate and glucuronide conjugated forms of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone could be used in clinical trials to check compliance with interventions made of standardized cranberry products. A wide variability among subjects was registered for all the assessed metabolites, as CV% of $C_{max}$ in plasma and of cumulative amounts excreted in urine was up to 297% and 322%, respectively. Not all the compounds exhibited the same level of variability, as, for example, for 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide a greater CV% was recorded. This could be due to differences at individual level in metabolizing flavan-3-ols, which is a result of the individual genotype and gut microbiota composition. When considering the sum of all the metabolites, a broad variability was found too, suggesting that there were some subjects able to metabolize flavan-3-ols more efficiently than others, producing higher amounts of PVLs.

Finally, we acknowledge that, although volunteers were instructed to follow dietary restrictions before the study days, we had limited control over this aspect. This could have led to an increase in the inter-individual variability and also to differences in the $T_{max}$ reported for the cranberry juices, as mentioned by Feliciano and colleagues [17], as well as for the small amounts of PVLs detected at baseline for some treatments.

## 3.1.5. Conclusion

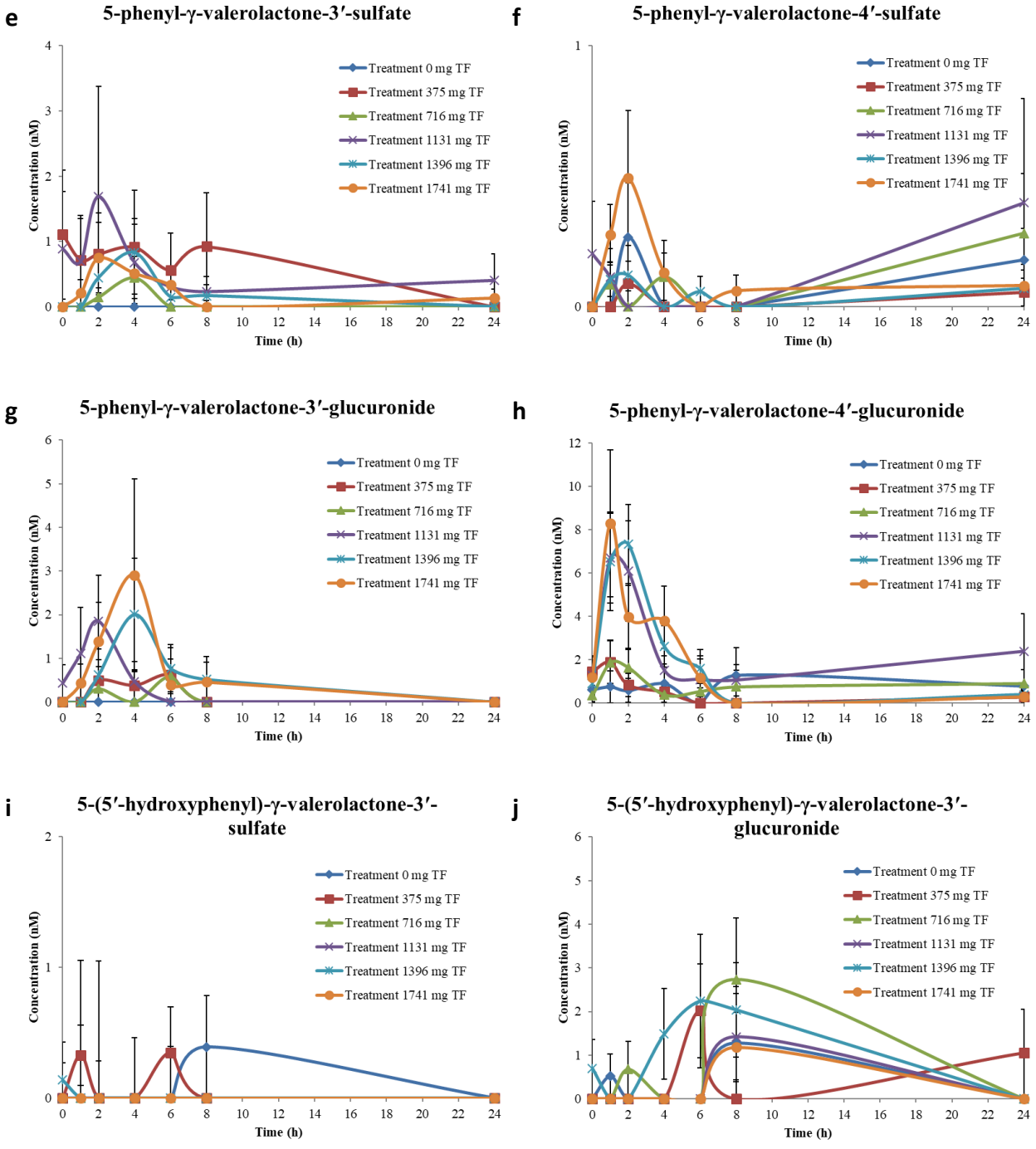
This study shows that glucuronide and sulfate conjugated forms of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone were the most relevant PVLs produced, circulating and excreted after cranberry juice intake. Glucuronidation seems to be the most favorable phase II metabolic pathway after cranberry juice intake, despite the fact that a wide inter-individual variability exists. We noted that most of the metabolites showed a dose-dependent relation, as reflected in a positive and highly significant correlation between the total PVLs, both circulating and excreted, and the amount of flavan-3-ols ingested with cranberry juice. Glucuronide and sulfate conjugates of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone drove the overall correlation and seems to be promising biomarkers of cranberry flavan-3-ols intake. Moreover, certain subjects displayed greater efficiency in metabolizing flavan-3-ols and producing PVLs with respect to others, and this might explain the differences in the putative health effects associated to these compounds. Further studies with a higher sample size are needed to better explore the inter-individual variability existing in the production of these colonic metabolites.

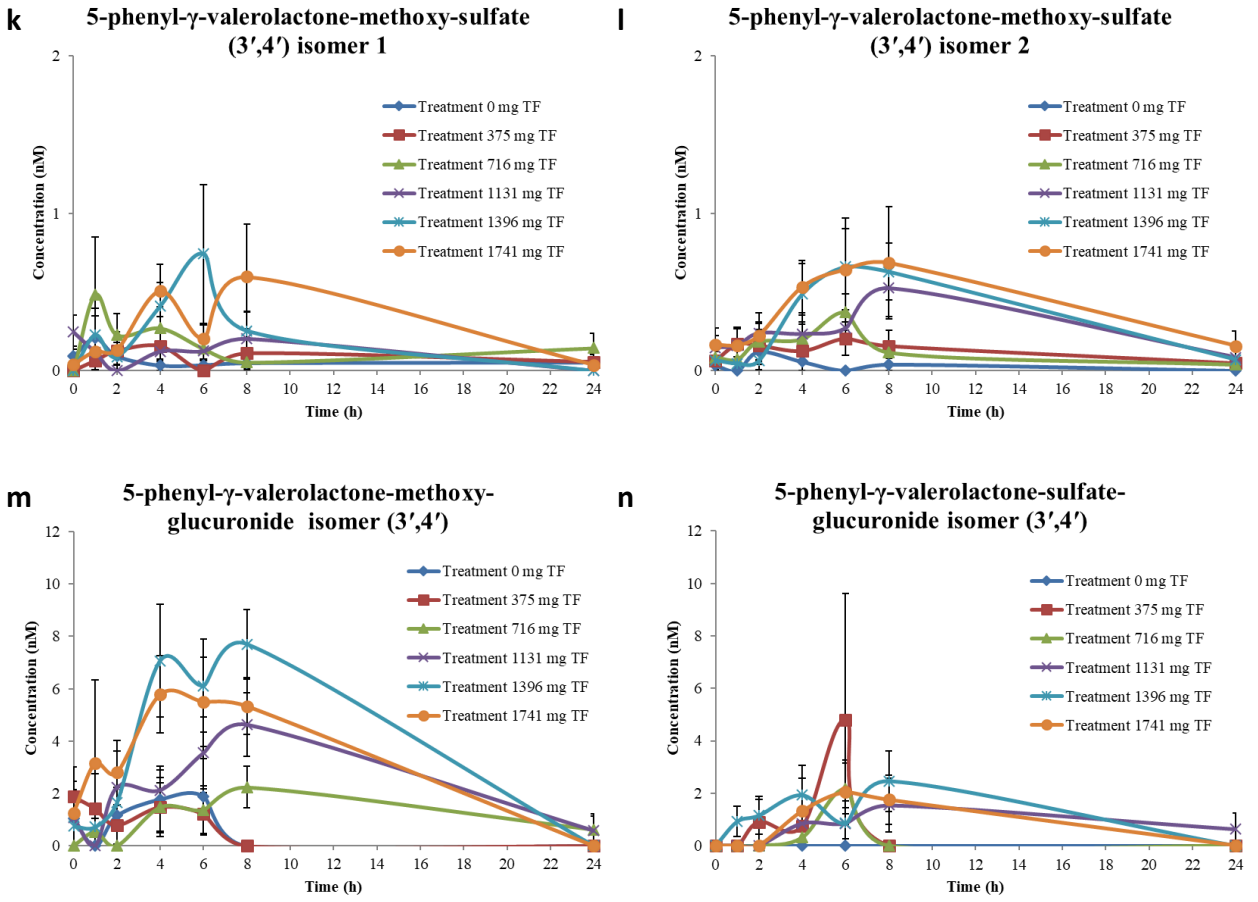
**Conflicts of Interest**

This study was funded by the Cranberry Institute. Ocean Spray donated the cranberry juices used in the human study and provided composition analyses. The funders of the study had no input on the design, implementation, analysis or interpretation of the data. The authors have no other conflicts of interest to declare.

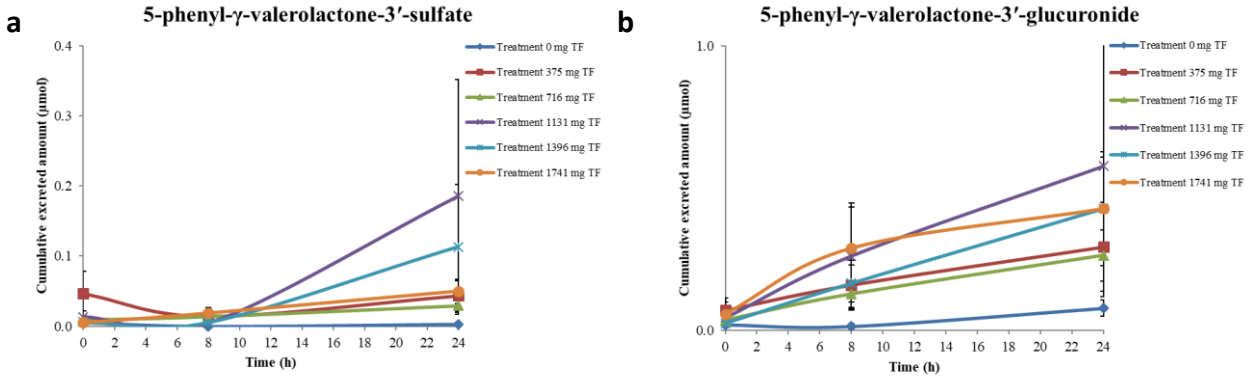
## Electronic Supplementary Information (ESI)

**Supplementary Figure 1e-n.** Pharmacokinetic profile of 5-phenyl-γ-valerolactone-3′-sulfate (**e**), 5-phenyl-γ-valerolactone-4′-sulfate (**f**), 5-phenyl-γ-valerolactone-3′-glucuronide (**g**), 5-phenyl-γ-valerolactone-4′-glucuronide (**h**), 5-(5′-hydroxyphenyl)-γ-valerolactone-3′-sulfate (**i**), 5-(5′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide (**j**), 5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 1 (**k**), 5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 2 (**l**), 5-phenyl-γ-valerolactone-methoxy-glucuronide isomer (3′,4′) (**m**) and 5-phenyl-γ-valerolactone-sulfate-glucuronide isomer (3′,4′) (**n**) for the different treatments. Data are expressed as nM, mean ± SEM (*n* = 10). Treatment 0 TF corresponds to the control drink, other treatments to cranberry juice drinks with growing amounts of total flavan-3-ols.

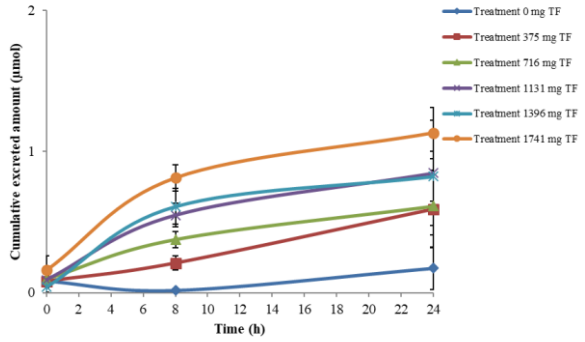




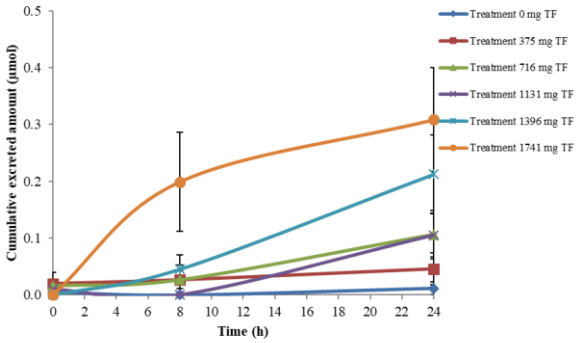

**Supplementary Figure 2a-n.** Cumulative urinary excretion profile of 5-phenyl-γ-valerolactone-3′-sulfate (**a**), 5-phenyl-γ-valerolactone-3′-glucuronide (**b**), 5-phenyl-γ-valerolactone-4′-glucuronide (**c**), 5-(3′,4′-dihydroxyphenyl)-γ-valerolactone (**d**), 5-(hydroxyphenyl)-γ-valerolactone-sulfate (3′,4′ isomers) (**e**), 5-(5′-hydroxyphenyl)-γ-valerolactone-3′-sulfate (**f**), 5-(4′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide (**g**), 5-(3′-hydroxyphenyl)-γ-valerolactone-4′-glucuronide (**h**), 5-(5′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide (**i**), 5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 1 (**j**), 5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 2 (**k**), 5-phenyl-γ-valerolactone-methoxy-glucuronide isomer (3′,4′) (**l**), 5-phenyl-γ-valerolactone-sulfate-glucuronide isomer (3′,4′) (**m**) and of the sum of quantified PVLs (**n**) for the different treatments. Data are expressed as µmol, mean ± SEM (*n* = 10). Treatment 0 TF corresponds to the control drink, other treatments to cranberry juice drinks with growing amounts of total flavan-3-ols.



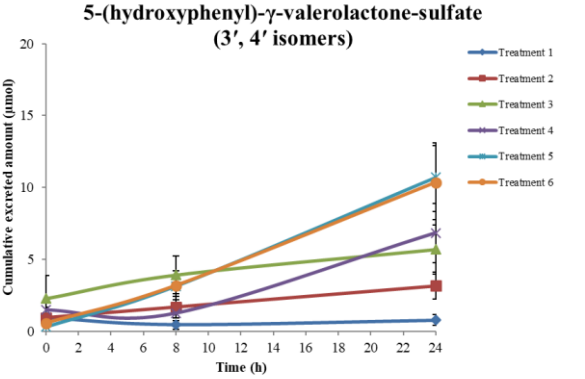
**c**


5-phenyl-γ-valerolactone-4′-glucuronide
Cumulative excreted amount (µmol)
Time (h)
Treatment 0 mg TF
Treatment 375 mg TF
Treatment 716 mg TF
Treatment 1131 mg TF
Treatment 1396 mg TF
Treatment 1741 mg TF

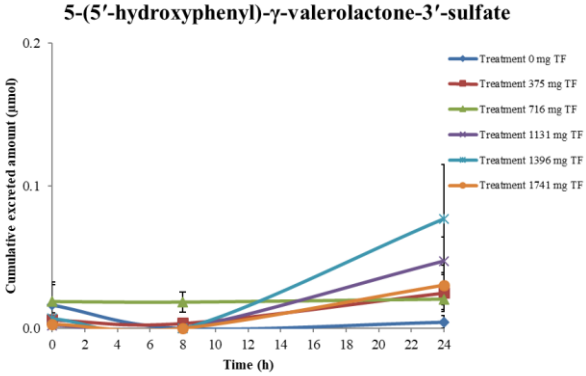
**d**


5-(3′,4′-dihydroxyphenyl)-γ-valerolactone
Cumulative excreted amount (µmol)
Time (h)
Treatment 0 mg TF
Treatment 375 mg TF
Treatment 716 mg TF
Treatment 1131 mg TF
Treatment 1396 mg TF
Treatment 1741 mg TF

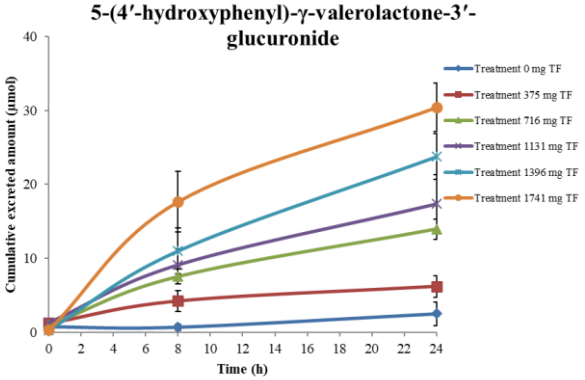
**e**


5-(hydroxyphenyl)-γ-valerolactone-sulfate (3′, 4′ isomers)
Cumulative excreted amount (µmol)
Time (h)
Treatment 1
Treatment 2
Treatment 3
Treatment 4
Treatment 5
Treatment 6

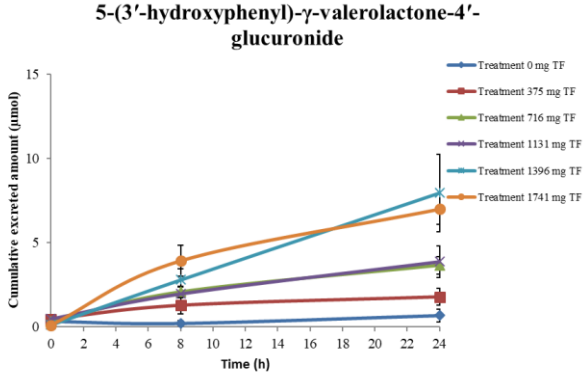
**f**


5-(5′-hydroxyphenyl)-γ-valerolactone-3′-sulfate
Cumulative excreted amount (µmol)
Time (h)
Treatment 0 mg TF
Treatment 375 mg TF
Treatment 716 mg TF
Treatment 1131 mg TF
Treatment 1396 mg TF
Treatment 1741 mg TF

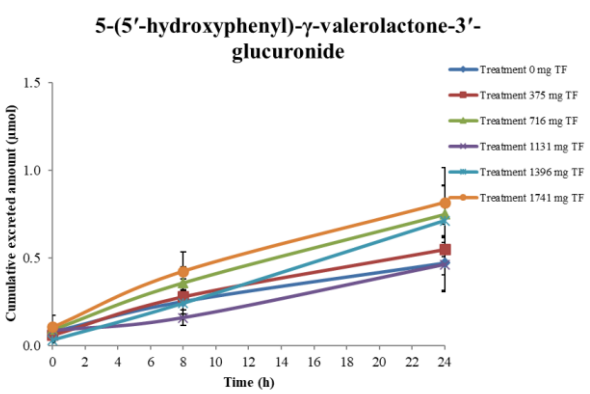
**g**


5-(4′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide
Cumulative excreted amount (µmol)
Time (h)
Treatment 0 mg TF
Treatment 375 mg TF
Treatment 716 mg TF
Treatment 1131 mg TF
Treatment 1396 mg TF
Treatment 1741 mg TF

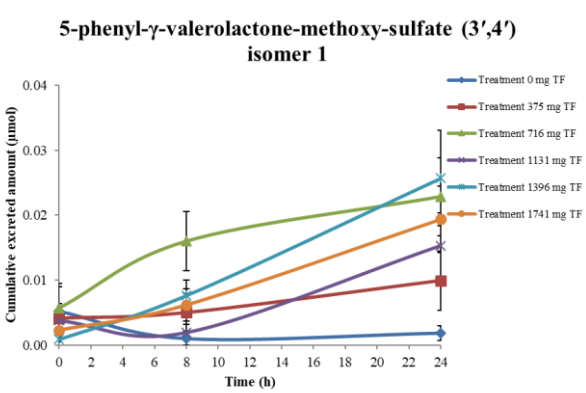
**h**


5-(3′-hydroxyphenyl)-γ-valerolactone-4′-glucuronide
Cumulative excreted amount (µmol)
Time (h)
Treatment 0 mg TF
Treatment 375 mg TF
Treatment 716 mg TF
Treatment 1131 mg TF
Treatment 1396 mg TF
Treatment 1741 mg TF

**i**


5-(5′-hydroxyphenyl)-γ-valerolactone-3′-glucuronide
Cumulative excreted amount (µmol)
Time (h)
Treatment 0 mg TF
Treatment 375 mg TF
Treatment 716 mg TF
Treatment 1131 mg TF
Treatment 1396 mg TF
Treatment 1741 mg TF

**j**


5-phenyl-γ-valerolactone-methoxy-sulfate (3′,4′) isomer 1
Cumulative excreted amount (µmol)
Time (h)
Treatment 0 mg TF
Treatment 375 mg TF
Treatment 716 mg TF
Treatment 1131 mg TF
Treatment 1396 mg TF
Treatment 1741 mg TF

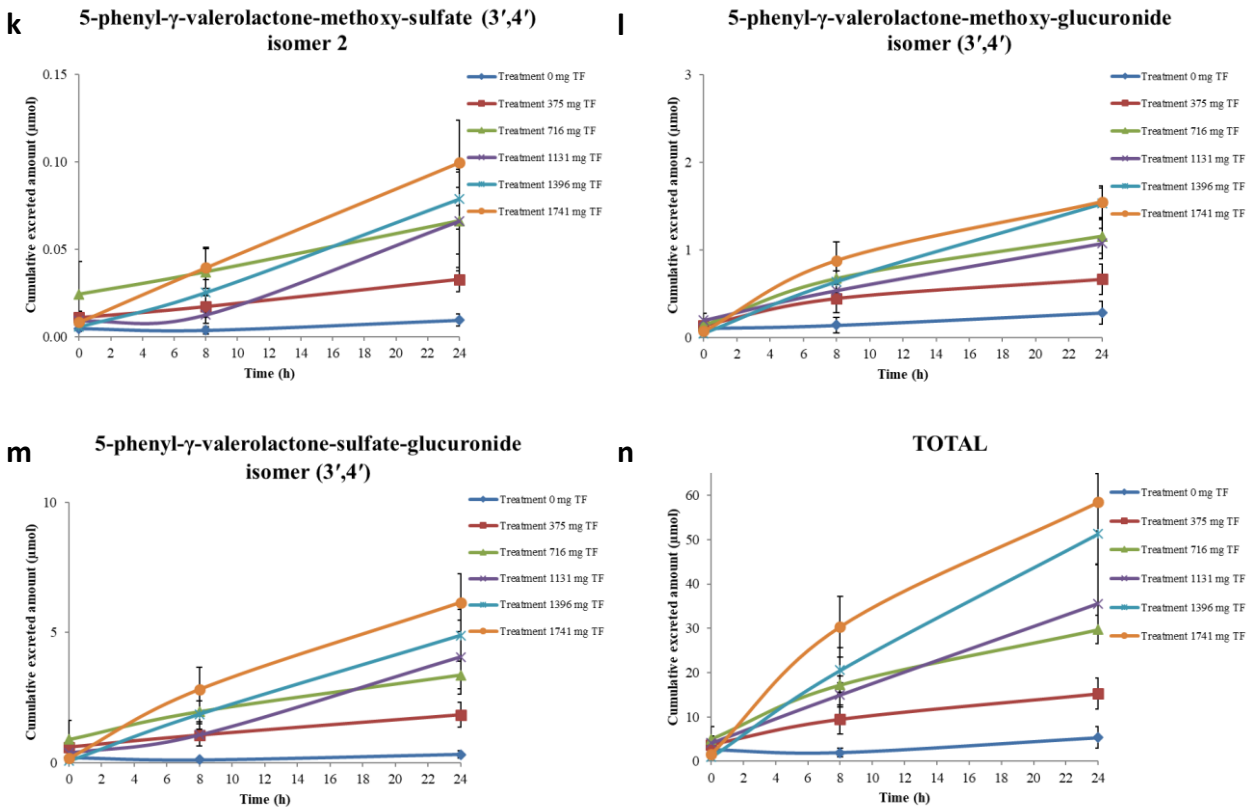



**Table S1.** Individual dose-response relation between **(a)** total amount of ingested flavan-3-ols and Area under the curve (AUC) of the concentration in plasma of PVLs over 24h and **(b)** total amount of ingested flavan-3-ols and cumulative quantity of excreted PVLs in urine.

**a**

| Subject | Equation | $R^2$ |
|---|---|---|
| **All (*n* = 10)** | y = 2.2575x + 113.07 | 0.981 |
| **1** | y = 4.1603x + 183.67 | 0.785 |
| **2** | y = 2.5589x + 425.06 | 0.949 |
| **3** | y = 2.0872x + 421.29 | 0.789 |
| **4** | y = 2.2223x - 102.44 | 0.976 |
| **5** | y = 1.1354x + 228.06 | 0.788 |
| **6** | y = 1.4211x + 184.62 | 0.961 |
| **7** | y = 2.2859x - 107.54 | 0.920 |
| **8** | y = 1.962x - 123.01 | 0.852 |
| **9** | y = 0.5098x + 395.69 | 0.310 |
| **10** | y = 4.2317x - 374.7 | 0.985 |

**b**

| Subject | Equation | $R^2$ |
|---|---|---|
| **All (*n* = 10)** | y = 0.031x + 4.9335 | 0.982 |
| **1** | y = 0.0345x + 2.4507 | 0.816 |
| **2** | y = 0.0436x - 3.4747 | 0.785 |
| **3** | y = 0.0366x + 10.938 | 0.799 |
| **4** | y = 0.0263x + 5.7057 | 0.767 |
| **5** | y = 0.0235x + 0.3171 | 0.794 |
| **6** | y = 0.05x + 4.1262 | 0.669 |
| **7** | y = 0.0488x - 4.7871 | 0.878 |
| **8** | y = 0.0363x - 5.895 | 0.890 |
| **9** | y = 0.0139x + 3.6589 | 0.414 |
| **10** | y = 0.0057x + 18.91 | 0.453 |

## References


1 K. C. Maki, K. L. Kaspar, C. Khoo, L. H. Derrig, A. L. Schild and K. Gupta, Consumption of a cranberry juice beverage lowered the number of clinical urinary tract infection episodes in women with a recent history of urinary tract infection, *Am. J. Clin. Nutr.*, 2016, **103**, 1434–1442.

2 J. A. Novotny, D. J. Baer, C. Khoo, S. K. Gebauer and C. S. Charron, Cranberry juice consumption lowers markers of cardiometabolic risk, including blood pressure and circulating C-reactive protein, triglyceride, and glucose concentrations in adults, *J. Nutr.*, 2015, **145**, 1185–1193.

3 M. P. Nantz, C. A. Rowe, C. Muller, R. Creasy, J. Colee, C. Khoo and S. S. Percival, Consumption of cranberry polyphenols enhances human γδ-T cell proliferation and reduces the number of symptoms associated with colds and influenza: a randomized, placebo-controlled intervention study, *Nutr. J.*, 2013, **12**, 1–9.

4 A. Basu, N. M. Betts, J. Ortiz, B. Simmons, M. Wu and T. J. Lyons, Low-energy cranberry juice decreases lipid oxidation and increases plasma antioxidant capacity in women with metabolic syndrome, *Nutr. Res.*, 2011, **31**, 190–196.

5 A. Rodriguez-Mateos, D. Vauzour, C. G. Krueger, D. Shanmuganayagam, J. Reed, L. Calani, P. Mena, D. Del Rio and A. Crozier, Bioavailability, bioactivity and impact on health of dietary flavonoids and related compounds: an update, *Arch. Toxicol.*, 2014, **88**, 1803–1853.

6 C.-H. Wang, C.-C. Fang, N.-C. Chen, S. S.-H. Liu, P.-H. Yu, T.-Y. Wu, W.-T. Chen, C.-C. Lee and S.-C. Chen, Cranberry-containing products for prevention of urinary tract infections in susceptible populations, *Arch. Intern. Med.*, 2012, **172**, 988–996.

7 A. Rodriguez-Mateos, R. P. Feliciano, A. Boeres, T. Weber, C. N. dos Santos, M. R. Ventura and C. Heiss, Cranberry (poly)phenol metabolites correlate with improvements in vascular function: a double-blind, randomized, controlled, dose-response, crossover study, *Mol. Nutr. Food Res.*, 2016, **60**, 2130–2140.

8 E. Pappas and K. M. Schaich, Phytochemicals of cranberries and cranberry products: characterization, potential health effects, and processing stability, *Crit. Rev. Food Sci. Nutr.*, 2009, **49**, 741–781.

9 L. Gu, M. A. Kelm, J. F. Hammerstone, G. Beecher, J. Holden, D. Haytowitz, S. Gebhardt and R. L. Prior, Concentrations of proanthocyanidins in common foods and estimations of normal consumption, *J. Nutr.*, 2004, **134**, 613–617.

10 D. Del Rio, A. Rodriguez-Mateos, J. P. E. E. Spencer, M. Tognolini, G. Borges and A. Crozier, Dietary (poly)phenolics in human health: structures, bioavailability, and evidence of protective effects against chronic diseases, *Antioxid. Redox Signal.*, 2013, **18**, 1818–92.

11 A. Rodriguez-Mateos, R. Del Pino-García, T. W. George, A. Vidal-Diez, C. Heiss and J. P. E. Spencer, Impact of processing on the bioavailability and vascular effects of blueberry (poly)phenols, *Mol. Nutr. Food Res.*, 2014, **58**, 1952–1961.

12 P. Mena, L. Bresciani, N. Brindani, I. A. Ludwig, G. Pereira-Caro, D. Angelino, R. Llorach, L. Calani, F. Brighenti, M. N. Clifford, C. I. R. Gill, A. Crozier, C. Curti and D. Del Rio, Phenyl-γ-valerolactones and phenylvaleric acids, the main colonic metabolites of flavan-3-ols: synthesis, analysis, bioavailability, and bioactivity, *Nat. Prod. Rep.*, 2019, **36**, 714–752.

13 L. Bresciani, D. Martini, P. Mena, M. Tassotti, L. Calani, G. Brigati, F. Brighenti, S. Holasek, D. E. Malliga, M. Lamprecht and D. Del Rio, Absorption profile of (poly)phenolic compounds after consumption of three food supplements containing 36 different fruits, vegetables, and berries, *Nutrients*, 2017, **9**, 1–17.

14 P. Mena, D. González de Llano, N. Brindani, A. Esteban-Fernández, C. Curti, M. V. Moreno-Arribas, D. Del Rio and B. Bartolomé, 5-(3’,4’-Dihydroxyphenyl)-γ-valerolactone and its sulphate conjugates, representative circulating metabolites of flavan-3-ols, exhibit anti-adhesive activity against uropathogenic Escherichia coli in bladder epithelial cells, *J. Funct. Foods*, 2017, **29**, 275–280.

15 D. González de Llano, A. Esteban-Fernández, F. Sánchez-Patán, P. Martínlvarez, M. Moreno-Arribas and B. Bartolomé, Anti-adhesive activity of cranberry phenolic compounds and their microbial-derived metabolites against uropathogenic Escherichia coli in bladder epithelial cell cultures, *Int. J. Mol. Sci.*, 2015, **16**, 12119–12130.

16 R. P. Feliciano, A. Boeres, L. Massacessi, G. Istas, M. R. Ventura, C. Nunes dos Santos, C. Heiss and A. Rodriguez-Mateos, Identification and quantification of novel cranberry-derived plasma and urinary (poly)phenols, *Arch. Biochem. Biophys.*, 2016, **599**, 31–41.

17 R. Feliciano, C. Mills, G. Istas, C. Heiss and A. Rodriguez-Mateos, Absorption, metabolism and excretion of cranberry (poly)phenols in humans: a dose response study and assessment of inter-individual variability, *Nutrients*, 2017, **9**, 268.

18 C. Manach, D. Milenkovic, T. Van de Wiele, A. Rodriguez-Mateos, B. de Roos, M. T. Garcia-Conesa, R. Landberg, E. R. Gibney, M. Heinonen, F. Tomás-Barberán and C. Morand, Addressing the inter-individual variation in response to consumption of plant food bioactives: towards a better understanding of their role in healthy aging and cardiometabolic risk reduction, *Mol. Nutr. Food Res.*, 2017, **61**, 1–16.

19 P. Mena, I. A. Ludwig, V. B. Tomatis, A. Acharjee, L. Calani, A. Rosi, F. Brighenti, S. Ray, J. L. Griffin, L. J. Bluck and D. Del Rio, Inter-individual variability in the production of flavan-3-ol colonic metabolites: preliminary elucidation of urinary metabotypes, *Eur. J. Nutr.*, 2019, **58**, 1529–1543.

20 N. Brindani, P. Mena, L. Calani, I. Benzie, S. W. Choi, F. Brighenti, F. Zanardi, C. Curti and D. Del Rio, Synthetic and analytical strategies for the quantification of phenyl-γ-valerolactone conjugated metabolites in human urine, *Mol. Nutr. Food Res.*, 2017, **61**, 6–10.

21 C. Curti, N. Brindani, L. Battistini, A. Sartori, G. Pelosi, P. Mena, F. Brighenti, F. Zanardi, D. Del Rio, Catalytic, enantioselective vinylogous mukaiyama aldol reaction of furan-based dienoxy silanes: a chemodivergent approach to γ-valerolactone flavan-3-ol metabolites and δ-lactone analogues, *Adv. Synth. Catal.*, 2015, **357**, 4082–4092.

22 Y. Zhang, M. Huo, J. Zhou and S. Xie, PKSolver: an add-in program for pharmacokinetic and pharmacodynamic data analysis in Microsoft Excel, *Comput. Methods Programs Biomed.*, 2010, **99**, 306–314.

23 M. Nikolic, A. Konic Ristic, A. González-Sarrías, G. Istas, M. Urpi-Sarda, M. Dall'Asta, L.-E. Monfoulet, L. Cloetens, B. Bayram, M. R. Tumolo, M. Chervenkov, E. Scoditti, M. Massaro, N. Tejera, D. Abadjieva, K. Chambers, I. Krga, F. A. Tomás-Barberán, C. Morand, R. Feliciano, R. García-Villalba, M. Garcia-Aloy and P. Mena, Improving the reporting quality of intervention trials addressing the inter-individual variability in response to the consumption of plant bioactives: quality index and recommendations, *Eur. J. Nutr.*, 2019, **58**, 49–64.

24 A. Anesi, P. Mena, A. Bub, M. Ulaszewska, D. Del Rio, S. E. Kulling and F. Mattivi, Quantification of urinary phenyl-γ-valerolactones and related valeric acids in human urine on consumption

of apples, *Metabolites*, 2019, **9**, 254.

25 A. Cortés-Martín, M. V. Selma, J. C. Espín and R. García-Villalba, The human metabolism of nuts proanthocyanidins does not reveal urinary metabolites consistent with distinctive gut microbiota metabotypes, *Mol. Nutr. Food Res.*, 2019, **63**, e1800819.

26 J. I. Ottaviani, R. Fong, J. Kimball, J. L. Ensunsa, A. Britten, D. Lucarelli, R. Luben, P. B. Grace, D. H. Mawson, A. Tym, A. Wierzbicki, K. T. Khaw, H. Schroeter and G. G. C. Kuhnle, Evaluation at scale of microbiome-derived metabolites as biomarker of flavan-3-ol intake in epidemiological studies, *Sci. Rep.*, 2018, **8**, 1–11.

## 3.2. Study 2

# Metabotypes of flavan-3-ol colonic metabolites after cranberry intake: elucidation and statistical approaches

Pedro Mena[1,*,+], **Claudia Favari**[1,+], Animesh Acharjee[2,3,4], Saisakul Chernbumroong[2,3], Letizia Bresciani[1], Claudio Curti[5], Furio Brighenti[1], Christian Heiss[6], Ana Rodriguez-Mateos[7], Daniele Del Rio[1, 8]

[1] Human Nutrition Unit, Department of Food & Drug, University of Parma, Parma, Italy

[2] College of Medical and Dental Sciences, Institute of Cancer and Genomic Sciences, Centre for Computational Biology, University of Birmingham, B15 2TT, UK

[3] Institute of Translational Medicine, University Hospitals Birmingham NHS, Foundation Trust, B15 2TT, UK

[4] NIHR Surgical Reconstruction and Microbiology Research Centre, University Hospital Birmingham, Birmingham B15 2WB, UK

[5] Department of Food & Drug, University of Parma, Parma, Italy

[6] Department of Clinical and Experimental Medicine, Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK

[7] Department of Nutritional Sciences, Faculty of Life Sciences and Medicine, King's College London, London, UK

[8] School of Advanced Studies on Food and Nutrition, University of Parma, Parma, Italy

[+] Equal contributors

[*] Corresponding author: Via Volturno 39, 43125 Parma, Italy; Tel.: +39 0521-903841; E-mail: pedro.mena@unipr.it

The content of this section has also been published in the European Journal of Nutrition (Mena P, Favari C, Acharjee A, et al (2021) Metabotypes of flavan-3-ol colonic metabolites after cranberry intake: elucidation and statistical approaches. Eur J Nutr. https://doi.org/10.1007/s00394-021-02692-z).

## Abstract

**Purpose:** Extensive inter-individual variability exists in the production of flavan-3-ol metabolites. Preliminary metabolic phenotypes (metabotypes) have been defined, but there is no consensus on the existence of metabotypes associated with the catabolism of catechins and proanthocyanidins. This study aims at elucidating the presence of different metabotypes in the urinary excretion of main flavan-3-ol colonic metabolites after consumption of cranberry products and at assessing the impact of the statistical technique used for metabotyping.

**Methods:** Data on urinary concentrations of phenyl-γ-valerolactones and 3-(hydroxyphenyl)propanoic acid derivatives from two human interventions has been used. Different multivariate statistics, principal component analysis (PCA), cluster analysis, and partial least square-discriminant analysis (PLS-DA), have been considered.

**Results:** Data pre-treatment plays a major role on resulting PCA models. Cluster analysis based on k-means and a final consensus algorithm lead to quantitative-based models, while the expectation-maximization algorithm and clustering according to principal component scores yield metabotypes characterized by quali-quantitative differences in the excretion of colonic metabolites. PLS-DA, together with univariate analyses, has served to validate the urinary metabotypes in the production of flavan-3-ol metabolites and to confirm the robustness of the methodological approach.

**Conclusions**: This work proposes a methodological workflow for metabotype definition and highlights the importance of data pre-treatment and clustering methods on the final outcomes for a given dataset. It represents an additional step toward the understanding of the inter-individual variability in flavan-3-ol metabolism.



**Trial registration number**

The acute study was registered at clinicaltrials.gov as NCT02517775, August 7, 2015; the chronic study was registered at clinicaltrials.gov as NCT02764749, May 6, 2016.

**Keywords**

Metabotypes; flavan-3-ols; inter-individual variation; phenolic metabolites; phenyl-γ-valerolactones

## 3.2.1. Introduction

Flavan-3-ols are characteristic polyphenols of tea, cocoa, wine, pome fruits (as apple and pear), berries, and nuts, but they are also found in stone fruits and legumes [1,2]. This subclass of compounds is the main dietary source of flavonoids in Western diets [3–5] and has been associated with beneficial effects on the prevention of cardiometabolic diseases [6–9]. In addition, other putative benefits have been observed against cognitive decline [10,11] and urinary tract infections [12,13]. In plant-based foods, they occur as simple monomers or as oligomers and polymers of up to 190 units (also known as proanthocyanidins or condensed tannins) [14]. When ingested, both monomeric and high molecular weight flavan-3-ols are poorly absorbed and metabolized in the first gastrointestinal tract, reaching the colon and becoming a suitable substrate for the local microbiota[15]. These compounds undergo an extensive microbial metabolism leading to the formation of specific metabolites, namely phenyl-γ-valerolactones (PVLs) and phenylvaleric acids (PVAs), as well as of common end-products of (poly)phenol colonic catabolism, such as phenylpropanoic, phenylacetic, and benzoic acid derivatives [16–19]. The microbial metabolites are then absorbed by colonocytes before reaching the liver and are converted into phase II conjugated derivatives. Conjugated PVLs (sulfate, glucuronide, methoxy, and combinations thereof) are the main colonic circulating metabolites after ingestion of monomeric and polymeric flavan-3-ols by humans [19–23] and, in particular, sulfate and glucuronide derivatives of 5-(3′,4′-dihydroxyphenyl)-γ-valerolactone have been proposed as biomarkers of flavan-3-ol intake [23,24]. These metabolites may be responsible for the health effects attributed to flavan-3-ols, as they are circulating molecules potentially available to target tissues and organs prior to be excreted in urine[19].

An extensive inter-individual variability is reported in the production of flavan-3-ol metabolites [20–23,25–29], possibly affecting, at individual level, the health benefits associated with this class of compounds [19]. This variability might be due to personal differences in gut microbiota composition, resulting in different metabolic phenotypes or metabotypes (i.e. different profiles of circulating and consequently excreted metabolites), likely impacting their effects on health, as it happens for other phenolic metabolites of colonic origin. Well-described examples of these differences are equol production from isoflavones and urolithin production from ellagitannins [30–32]. Stratification of individuals according to their equol/urolithin metabotype has proven to be necessary to understand the health effects associated to isoflavone and ellagitannin intake [33–35]. However, the information on flavan-3-ol colonic metabolites is much less defined. *In vitro* anaerobic incubations of (−)-epicatechin revealed inter-individual differences in its colonic metabolism and the formation of

certain metabolites was correlated with specific microbial phyla [36]. In a recent preliminary study, three putative metabotypes after green tea flavan-3-ol consumption were defined *in vivo* on the basis of a different urinary production of PVLs and 3-(hydroxyphenyl)propanoic acids (HPPs), through explorative partial least squares-discriminant analysis (PLS-DA) models [26]. Similar results were obtained after consumption of nut proanthocyanidins in nearly free living conditions, using the k-means clustering algorithm [25], but the authors did not associate the different profiles of PVLs and HPPs to metabotypes as they adhered to a more restrictive definition of phenolic metabotypes, characterised by the presence/absence of specific metabolites. However, the urinary profiles there described could be defined as flavan-3-ol colonic metabotypes when considering a broader definition of the term, commonly accepted in the nutrition field as "subgroups of individuals sharing the same metabolic profile" [37]. Beyond terminology, it is clear that there is a lack of information on how to handle the inter-individual variability in the production of phenolic metabolites to define metabotypes in those cases where all the subjects produce all the phenolic metabolites of a catabolic pathway, but in different proportions, as it happens for flavan-3-ols and for the main dietary classes of (poly)phenols.

The primary aim of the present study was to evaluate the existence of metabotypes, based on the urinary excretion of flavan-3-ol metabolites after consumption of flavan-3-ols from cranberry products, to shed light on this key aspect associated with the metabolism of these major phenolics. Secondly, this work aimed at investigating the impact of the statistical techniques used for the definition of phenolic metabotypes, defining an approach to specifically seek for metabotypes when they are not characterized by the dichotomic production/non-production of specific phenolic metabolites.

## 3.2.2. Material and Methods

*Intervention studies*

The dataset for this study consisted of urinary concentrations of several gut microbiota-derived metabolites of flavan-3-ols, namely monohydroxyPVLs (isomers 3' and 4'), dihydroxyPVLs (3',4'), and HPPs, quantified in urine samples collected in two different cranberry feeding studies, one with an acute design and one chronic. These metabolites were chosen according to previous evidence [26].

The acute study was a crossover, randomized, controlled intervention trial registered under the NIH ClinicalTrials.gov website (NCT02517775). The study was conducted in accordance with the guidelines stated in the current revision of the Declaration of Helsinki, and informed consent was obtained for all subjects. All procedures involving human subjects were approved by the University

of Dusseldorf Research Ethics Committee (ref: 14-012). Briefly, ten healthy men had to consume a cranberry drink containing increasing amounts of total flavan-3-ols (TF) or an isocaloric control (0 mg TF) drink with one-week washout [23,38]. Participants were instructed to follow a low-(poly)phenol diet for 3 days before and during the study day and had to fast for 12 h before the study day. Urine samples were collected at baseline, between 0-8 h and 8-24 h after drink intake. For the study purpose, data on cumulative urinary excretion (0-24 h) of the metabolites after higher flavan-3-ol intake (716, 1131, 1396, and 1741 mg TF) were considered, for a total of 40 observations. Quantitative data on the urinary excretion of PVLs have been previously reported [23], while data on HPPs is novel.

The chronic study was a parallel, randomized, controlled trial in which 22 healthy participants were asked to consume a cranberry powder containing 0.5 mg of flavan-3-ol monomers and 374 mg of proanthocyanidins every day for one month, without any other dietary restriction or recommendation. The study was registered under the NIH ClinicalTrials.gov website (NCT02764749) and was conducted according to the guidelines laid down in the current revision of the Declaration of Helsinki. Informed consent was obtained for all participants and all procedures involving human subjects were approved by the University of Dusseldorf Research Ethics Committee (Ref: 5360R). Cumulative 24-h urine samples from the first (v1) and the last (v2) intervention day were collected and analyzed to obtain data on metabolite concentration, for a total of 43 observations (1 sample from v1 was missing). In this case, both data on PVLs and HPPs are new. To sum up, a reasonable number of observations ($n$ = 83) was used for subsequent statistical analyses.

*Sample analysis*

Urine samples were prepared according to a previous report [39] and then analyzed through UHPLC DIONEX Ultimate 3000 fitted with a TSQ Vantage triple quadrupole mass spectrometer, equipped with a heated-electrospray ionization (H-ESI-II) source (Thermo Fisher Scientific Inc., San Jose, CA, USA). Chromatographic and ionization parameters were set following a validated method optimized for the analysis of PVLs [39]. Metabolite identification was carried out by comparison of the retention time with in-house synthesized standards and/or MS/MS fragmentation patterns. Up to 76 compounds among PVLs, PVAs and HPPs were simultaneously monitored in selective reaction monitoring (SRM) mode. Eleven metabolites were quantified in urine samples from the two interventions, namely 5-phenyl-γ-valerolactone-3’-sulfate, 5-phenyl-γ-valerolactone-3’-glucuronide, 5-phenyl-γ-valerolactone-4’-glucuronide, 5-(3’,4’-dihydroxyphenyl)-γ-valerolactone, 5-(hydroxyphenyl)-γ-valerolactone-sulfate (3’,4’ isomers), 5-(4’-hydroxyphenyl)-γ-valerolactone-3’-

glucuronide, 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-glucuronide, 5-phenyl-γ-valerolactone-methoxy-sulfate isomer (3',4'), 5-phenyl-γ-valerolactone-sulfate-glucuronide isomer (3',4'), 3-(phenyl)propanoic acid-sulfate and 3-(phenyl)propanoic acid-glucuronide. This nomenclature follows the current recommendations for (poly)phenol catabolites [40]. Quantification was performed with calibration curves of standards, when available. When not available, metabolites were quantified with the most structurally similar compound, as in the case of 5-phenyl-γ-valerolactone-4'-glucuronide, quantified as its isomer 5-phenyl-γ-valerolactone-3'-glucuronide; 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide, 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-glucuronide and 5-phenyl-γ-valerolactone-sulfate-glucuronide isomer (3',4'), quantified as 5-(5'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide; 5-phenyl-γ-valerolactone-methoxy-sulfate isomer (3',4'), quantified with 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-sulfate (prepared in house using reported procedures, [39]); and the 3-(phenyl)propanoic acid-sulfate and –glucuronide, quantified as 3-(4'-hydroxyphenyl)propanoic acid-3'-sulfate and 3-(4'-hydroxyphenyl)propanoic acid-3'-glucuronide (Toronto Research Chemicals, Toronto, Canada), respectively. Sums of metabolites belonging to the same aglycone compound were calculated: namely 5-phenyl-γ-valerolactone-3'-sulfate and 5-phenyl-γ-valerolactone-3'-glucuronide for 5-(3'-hydroxyphenyl)-γ-valerolactone aglycone; 5-phenyl-γ-valerolactone-4'-glucuronide for 5-(4'-hydroxyphenyl)-γ-valerolactone; 5-(3',4'-dihydroxyphenyl)-γ-valerolactone, 5-(hydroxyphenyl)-γ-valerolactone-sulfate (3',4' isomers), 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide, 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-glucuronide, 5-phenyl-γ-valerolactone-methoxy-sulfate isomer and 5-phenyl-γ-valerolactone-sulfate-glucuronide isomer for 5-(3',4'-dihydroxyphenyl)-γ-valerolactone; and 3-(phenyl)propanoic acid-sulfate and 3-(phenyl)propanoic acid-glucuronide for 3-(hydroxyphenyl)propanoic acid aglycone. This way, two different datasets, one consisting of 83 observations (samples) and 11 variables corresponding to individual metabolites, and one consisting of 83 observations (samples) and 4 variables corresponding to the sums of metabolites belonging to the same aglycone, were considered. All the metabolite data are expressed as μmol excreted in 24 h.

*Unsupervised analyses*

**Principal component analysis**

Principal component analysis (PCA) was performed using SIMCA 16.0.1 software (Sartorius Stedim Data Analytics, Umea, Sweden). Both datasets were subjected to several transformations (no transformation, logarithmic transformation, and power transformation) and four mean centering plus scaling methods (1, neither centering nor scaling, 2) only centering, 3) centering plus unit

variance scaling or autoscaling, and 4) centering with Pareto scaling) of the variables, resulting in a total of 24 PCA models. In particular, logarithmic transformation applied was a 10-based logarithm Log (C1*X + C2) where C1 = 1 and C2 = 0 and the power transformation was $(C1*X + C2)^{C3}$ with C1 = 1, C2 = 0 and C3 = 2 [41]. No scaling was taken into consideration as many variables already presented values close to zero. Centering converted all the data to variations around zero instead of around the mean of the data; unit variance scaling used standard deviation as scaling factor, while Pareto scaling used the square root of standard deviation [42]. All models were presented by default with two principal components (PC). The parameters used to assess the quality of each model and subsequent data interpretability were $R^2(X)$ and $Q^2$, namely the model fit (or explained variation) and the predictive ability, respectively.

**Cluster analysis: cluster identification and consensus**

The two datasets, individual metabolites or the sums of metabolites belonging to the same aglycone compound, were separately submitted to cluster analysis after being centered and unit variance scaled. The cluster analysis was carried out using R version 3.6.1. [43]. The cluster analysis was performed using nine different algorithms, namely hierarchical clustering on principal components (HCPC) [44], hierarchical k-means (H-Kmeans), hierarchical partition around medoids (H-PAM), hierarchical fuzzy (H-fuzzy), partition around medoids (PAM) [45], k-means (Kmeans) [46], fuzzy c-means [47], hierarchical [48] and expectation-maximization (EM) [49]. The number of clusters between two and ten clusters were experimented. To maintain results stability, the process was repeated for five times. Twenty-five internal cluster indexes such as Ball Hall, Banfield Raftery, C index, etc., were applied to measure how compact the clusters were. The optimal number of clusters were selected based on the majority voting scheme. Using the identified optimal number of clusters, we developed nine clustering models using the aforementioned algorithms. Majority voting was used to identify the final cluster assignments (final consensus, FC). In addition, clustering was carried out taking into account the scores of each observation for each principal component (PC) after conducing PCAs with autoscaled data.

*Supervised analysis: Partial least square-discriminant analysis*

PLS-DA on both datasets was performed using SIMCA 16.0.1 software (Sartorius Stedim Data Analytics, Umea, Sweden). Observations were assigned to classes based on the results of cluster analysis and their PC scores. Variables of both datasets were centered and unit variance scaled (autoscaled). Model validity was assessed by $R^2(X)$, $Q^2$, the random permutation test, and CV-ANOVA within the SIMCA package. The identification of the most relevant metabolites from the whole set

of metabolites (variable selection) was performed using the Variable Importance in Projection (VIP) scores, estimating the importance of each variable in the projection used in a PLS model [50]: variables with VIP scores greater than 1 were considered important in the given model.

*Univariate statistics*

The urinary excretion of individual metabolites, sums of metabolites belonging to the same aglycone compound, and sums of sulfate or glucuronide metabolites per each cluster defined after applying different clustering methods (FC, EM, Kmeans, and PC score-based) were expressed as mean ± standard deviation. The normality of data distribution was checked through the Kolmogorov–Smirnov test. Data homoscedasticity was tested with Levene's test. Comparisons between two clusters were performed using independent sample *t*-test for normally distributed variables or non-parametric Mann–Whitney *U* test for non-normally distributed variables. Comparisons among three clusters were investigated by one-way ANOVA with *post hoc* Dunnett's test (all variables were heteroscedastic) for normally distributed variables or non-parametric Kruskal-Wallis test with *post hoc* pairwise multiple comparison for non-normally distributed variables. Differences were considered significant at *p*-value < 0.05. Boxplots were built using the urinary excretion of sums of metabolites belonging to the same aglycone compound. All these univariate statistical analyses were performed using IBM SPSS Statistics version 26 (IBM, Chicago, IL, USA).

## 3.2.3. Results

*Effect of data pre-treatment on resulting PCA models*

To evaluate the influence of data pre-treatment on the resulting PCA models and derived biological outcomes, several transformations and scaling methods were applied to the two distinct datasets.

Considering individual metabolites, applying no data transformation resulted in better models compared to logarithmic and power transformations, which yielded worse $R^2(X)$ and $Q^2$ values than non-transformed models (Table 6). In fact, logarithmic transformation of variables returned many missing values, due to its inability to deal with zero value [42], while power transformation increased data skewness, which is not advisable. This was likely due to the presence of many excretion values close to zero, as it happened for minor excreted metabolites. Regarding scaling methods, for every applied transformation, higher quality models were obtained when using, in order, no scaling > centering > centering with Pareto scaling > centering with unit variance scaling (UV) (or autoscaling) (Table 6). Looking at every model, it was possible to identify patterns of metabolite (variable) distribution in the loading plot (Fig. 16 and Figure S1A-L). Three types of patterns were observed

(Table 1): (1) one reflecting differences in phase II metabolism (P), as sulfate and glucuronide metabolites grouped separately in the loading plot (Figure 16B); (2) one reflecting differences in colonic metabolism (C), on the basis of the derivatives originating from a certain aglycone (Figure 16D); and (3) one resembling a random distribution (R), as a biological interpretation was not found.

**Table 6.** Statistics of computed PCA models illustrating the effect of transformation and scaling methods on the datasets considering individual metabolites and sums of metabolites belonging to the same aglycone compound.

| **Individual metabolites** | | | | |
|---|---|---|---|---|
| **Data pre-treatment** | | **Model quality parameters** | | **Pattern*** |
| **Transformation** | **Mean centering + scaling** | **$R^2X$ (cum)** | **$Q^2$ (cum)** | |
| None | None | 0.965 | 0.781 | C |
| | Centering | 0.935 | 0.681 | P |
| | Centering + UV | 0.621 | 0.247 | C |
| | Centering + Pareto | 0.818 | 0.551 | P |
| Log | None | 0.862 | 0.622 | P |
| | Centering | 0.704 | 0.303 | R |
| | Centering + UV | 0.622 | 0.230 | R |
| | Centering + Pareto | 0.648 | 0.257 | R |
| Power | None | 0.991 | 0.527 | P |
| | Centering | 0.990 | 0.511 | P |
| | Centering + UV | 0.621 | -0.070 | C |
| | Centering + Pareto | 0.926 | 0.560 | P |
| **Sums of metabolites** | | | | |
| **Data pre-treatment** | | **Model quality parameters** | | **Pattern*** |
| **Transformation** | **Mean centering + scaling** | **$R^2X$ (cum)** | **$Q^2$ (cum)** | |
| None | None | 0.999 | 0.282 | C |
| | Centering | 0.998 | 0.094 | C |
| | Centering + UV | 0.715 | -0.139 | C |
| | Centering + Pareto | 0.949 | 0.076 | C |
| Log | None | 0.933 | 0.396 | C |
| | Centering | 0.728 | -0.210 | C |
| | Centering + UV | 0.705 | -0.210 | C |
| | Centering + Pareto | 0.704 | -0.210 | C |
| Power | None | 1.00 | 0.063 | C |
| | Centering | 1.00 | 0.045 | C |
| | Centering + UV | 0.720 | -0.210 | C |
| | Centering + Pareto | 0.998 | -0.020 | C |

Notes: The two parameters $R^2X$ (cum) and $Q^2$ (cum) represent respectively the model fit (or explained variation) and the predictive ability. The higher these values, the better the model. Abbreviations: UV: Unit Variance. Centering + UV is so-called autoscaling.

* "Pattern" stands for "pattern of metabolite distribution": P, data distribution on the basis of the phase II metabolism; C, distribution on the basis of the colonic metabolism; R, random distribution (no biological explanation).

A phase II metabolism-based distribution pattern was mainly observed after applying centering or centering with Pareto, while a colonic metabolism-based pattern was shown after autoscaling, regardless of the transformation used (Table 6). This colonic metabolism pattern accounted for the

existence of potential flavan-3-ol colonic metabotypes. Interestingly, random distribution was only seen after applying logarithmic transformation. No pattern was associated to the intervention study (acute or chronic) or the treatment/visit type, indicating that these aspects did not influence the variability registered in our data (Figure 16A and 16C and Figure S1A-L).

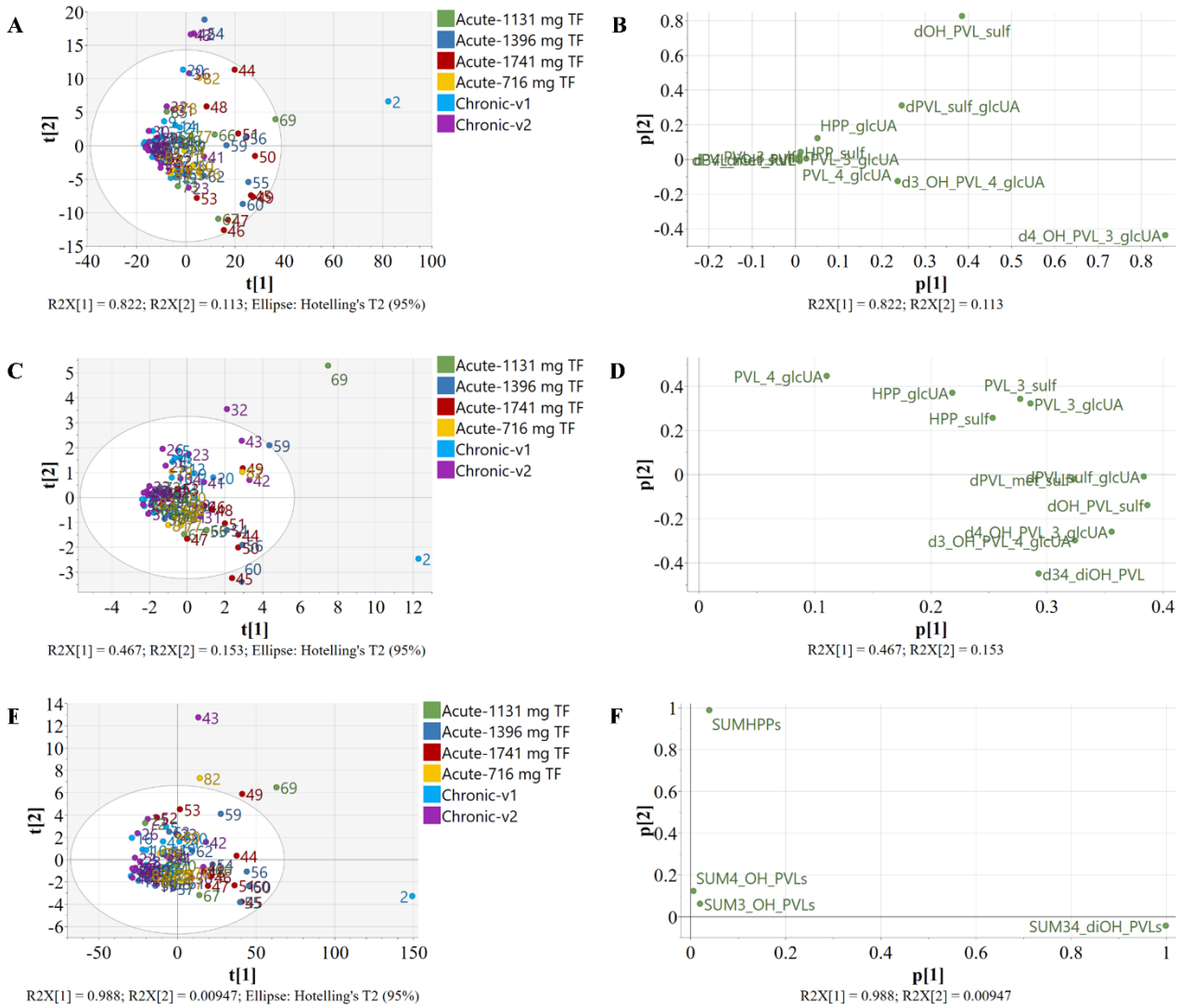



**Figure 16.** Score (**A, C, E**) and loading (**B, D, F**) plots resulting after PCA analysis on non-transformed, centered data for individual metabolites (**A, B**), non-transformed, centered and unit variance scaled data for individual metabolites (**C, D**), non-transformed, centered data for sums of metabolites belonging to the same aglycone compound (**E, F**).

The best PCA model showing a phase II metabolism-based distribution was obtained after non transforming and centering data (Figure 16A and B). Samples in the top right quadrant (Figure 16A) were characterized by a more abundant excretion of 5-(hydroxyphenyl)-γ-valerolactone-sulfate (3',4' isomers) (Figure 16B), while samples in the bottom right quadrant by a higher excretion of glucuronide derivatives of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone, namely 5-(4'-hydroxyphenyl)-γ-valerolactone-3'-glucuronide and 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-glucuronide. These

three compounds are the most excreted after consumption of flavan-3-ols [19,23,24] and probably phase II metabolism of some individuals favours sulfation, while in some others glucuronidation is predominant [25,51]. As a matter of fact, when looking at every subject during the different treatments (acute study) or intervention days (chronic study), it was possible to observe that most of the related samples appeared close together, suggesting that the pattern of conjugation of flavan-3-ol catabolites is preserved in different subjects (as for example samples 44 and 54, 50 and 60, 45 and 55, 47 and 67 in Figure 16A). On the other hand, the best PCA model representing a colonic metabolism pattern was obtained after non transforming and autoscaling the data (Figure 16C and D). In this case, samples in the top right quadrant were characterized by a higher urinary concentration of 3-(hydroxyphenyl)propanoic acid and 5-(hydroxyphenyl)-γ-valerolactone (both 3' and 4') derivatives, while samples in the bottom right quadrant were described by a more abundant excretion of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone derivatives, suggesting differences in the microbial production and urinary excretion of flavan-3-ol catabolites. As in the previous case, samples belonging to the same person fell close and within the same quadrant (as for example samples 49, 59 and 69, 32 and 42, 50 and 60, 44 and 54, 45 and 55, 56 and 66, 47 and 67 in Figure 16C). This was relevant as it accounted for the conservation of the metabolic pattern in the short time.

When considering sums of metabolites belonging to the same aglycone, the information about phase II metabolism was obviously lost, but it served to better highlight differences in the colonic metabolism of flavan-3-ols. Also in this case, not transformed data returned higher quality models compared to logarithmic- and power-transformed data (Table 6). Power transformation resulted in overfitted models, especially when coupled to any scaling or centering. Better models were obtained when applying, in order, no scaling > centering > centering with pareto scaling > centering with unit variance scaling (or autoscaling) to data matrix (Table 6), as for the individual metabolite dataset. All the models showed a colonic metabolism-based distribution pattern, even though not all the models displayed similar distributions for samples and metabolites in the score and loading plots, respectively (Figure 16E, F and Figure S2A-M). The model resulting after non transforming and centering the variables in the dataset is shown in Figure 16E, F, as an example of a high quality PCA model considering sums of metabolites belonging to the same aglycone. The information gathered from these plots was that the samples placed in the top right quadrant were characterized by a higher excretion of 3-(hydroxyphenyl)propanoic acids and those placed in the bottom right quadrant by a higher excretion of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone derivatives, while 5-

(hydroxyphenyl)-γ-valerolactones (both 3' and 4' isomers) did not account for sample variability (Figure 16E, F). In general, the conservative pattern of metabolite production among subjects was also observed using the sum of metabolites (i.e., samples 49, 59 and 69, 46 and 56, 50 and 60, 45 and 55, 44 and 54 in Figure 16E).

In all these unsupervised analyses, samples out of the Hotelling's circle (as visible in Figure 16A, C, and E) were not defined as outliers to be removed, since there are no profiles of metabolite excretion that can be judged as correct and incorrect when it comes to the individual production of flavan-3-ol metabolites.

*Cluster definition*

Once PCA highlighted a notable variability in the production of PVLs and HPPs, attention was paid into sample grouping. Testing several clustering criteria on nine clustering algorithms identified two clusters as the optimal number of groups best describing the data. This was done when considering both datasets (individual metabolites, Figure 17A, and sums of metabolites belonging to the same aglycone compound, Figure 17B) and all the clustering methods tested performed similarly, except for the EM algorithm when individual data were taken into account (Figure 17A). Then, a FC on clustering was voted, based on each observation frequency to fall within a group (Table S1). The results of three clustering methods were then selected to be used for the PLS-DA, namely EM, since it performed differently to the rest of the clustering algorithms; Kmeans, as it is widely applied [37,52] and has already been used when studying the potential metabotypization of flavan-3-ol colonic metabolites [25] and FC, because it merged and summarized all the results obtained after testing all the different clustering algorithms. Regarding the distribution of the observations between clusters, one cluster was larger than the other (about 60 observations vs. about 20) for all the three algorithms chosen. Of note, a subject allocated in a group after applying a clustering method on the dataset with individual metabolites was not necessarily then allocated in the same group when the same clustering method was applied to the dataset with sums of metabolites belonging to the same aglycone.

Clustering was also carried out according to the scores of each observation for the PCA models obtained after autoscaling. In particular, two (PC score-based, 2 groups) and three clusters (PC score-based, 3 groups) were defined for both datasets. Two clusters were obtained by allocating the observations with a positive PC2 in a group and the observations with a negative PC2 in another

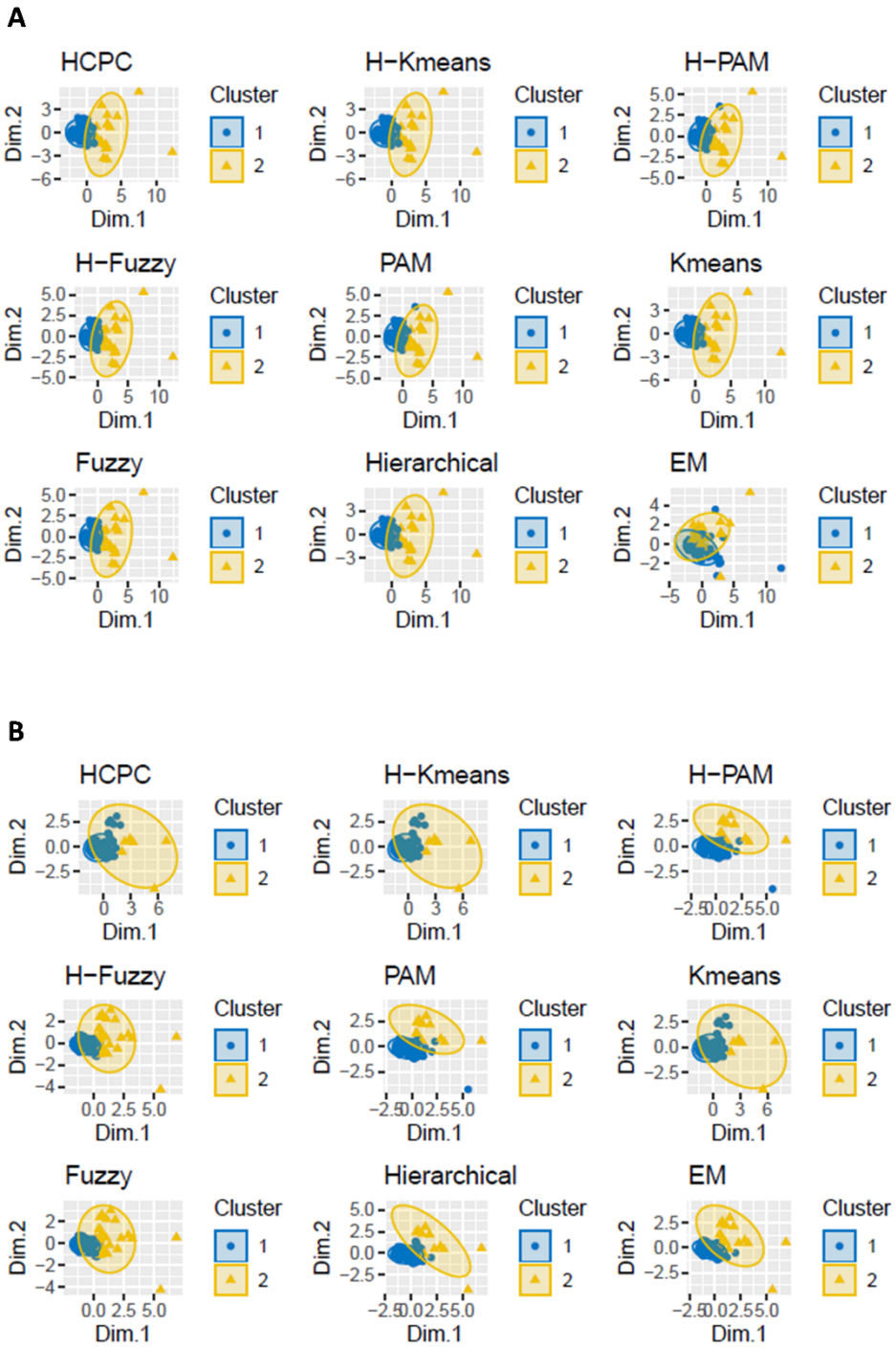



**Figure 17.** Two-classes cluster plot resulted from the application of different clustering methods on the datasets with individual metabolites (**A**) and with sums of metabolites belonging to the same aglycone compound (**B**).

group, the two clusters including a similar number of samples. The three clusters were set by allocating the observations with positive PC1 and PC2 scores in one group, the observations with a positive PC1 and a negative PC2 score in a second group, and the observations with a negative PC1 score in a third group. Of these three groups, one was notably more numerous than the other two.

*PLS-DA models to explore differences between clustering methods and the biological relevance of each group*

All the PLS-DA models performed with individual metabolites taking into account the groups from the selected clustering methods showed a good explained variance ($R^2X$) (Table 7). The models performed upon grouping by clustering algorithms (FC, EM, Kmeans) showed a better predictive ability ($Q^2 > 0.5$) than the models clustering observations by using PC scores ($Q^2 < 0.5$). All the models passed cross-validation by CV-ANOVA (*p*-value in Table 7) and by random permutation (Figure S3A-E). These results asserted model validation and excluded data overfitting. The two PLS-DA models performed using the classes defined by FC and Kmeans were very similar (Figure 18A, C). In both models, the distribution of the clusters was due to the amount of metabolites excreted (high vs. low) (VIP values for all the models are reported in Table S2). Differently, the PLS-DA model performed using the classes defined by EM clustering (Figure 18B) identified two groups characterized by the excretion of different metabolites: group 1 presented a higher excretion of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone metabolites, while group 2 had a higher excretion of 3-(hydroxyphenyl)propanoic acid derivatives and 5-phenyl-γ-valerolactone-4'-glucuronide, these differences being attributed to differences in the colonic metabolism of flavan-3-ols. A similar trend was observed in the PLS-DA models performed using groups defined according to the PC scores (Figure 18D and E), as they were described by the excretion of different colonic metabolites. When 2 groups were considered (Figure 18D) (PC score-based, 2groups), group 1 showed a higher excretion of (monohydroxyphenyl)-γ-valerolactone (both 3' and 4') and 3-(hydroxyphenyl)propanoic acid derivatives, while group 2 presented a higher excretion of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone metabolites. When 3 groups were taken into account (Figure 18E) (PC score-based, 3groups), group 1 was characterized by a greater excretion of 5-(hydroxyphenyl)-γ-valerolactone (both 3' and 4') and 3-(hydroxyphenyl)propanoic acid conjugates, group 2 by a higher excretion of mono-conjugated 5-(3',4'-dihydroxyphenyl)-γ-valerolactones and group 3 by a limited excretion of metabolites. None of the 5 PLS-DA models showed a distribution of the observations within each group due to phase II metabolism (Figure 18A-E).

When considering sums of metabolites belonging to the same aglycone, all the PLS-DA models exhibited good explained variance ($R^2X$) (Table 7), even better than what observed for the PLS-DA models with individual metabolites (Table 7). The models performed using classes defined by FC and EM also showed quite good predictive ability ($Q^2 > 0.5$), differently from Kmeans and clustering by PC scores ($Q^2 < 0.5$). All the models passed cross-validation by CV-ANOVA (*p*-value in Table 7) and

by random permutation (Figure S4A-E). These results validated the model and excluded overfitting of the data.

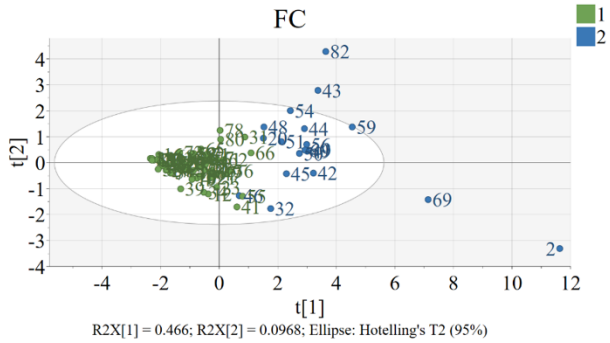
**Table 7.** Statistics of computed PLS-DA models, considering the class resulted from different clustering methods and individual metabolites or sums of metabolites belonging to the same aglycone compound.

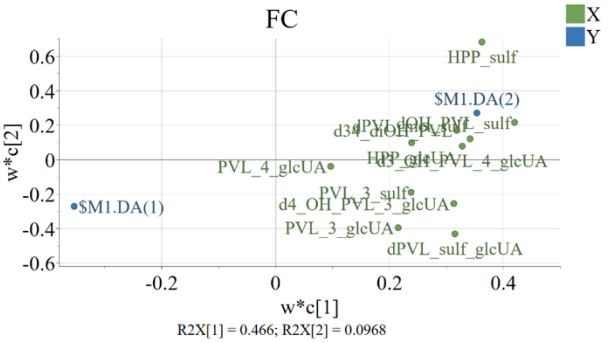
| **Individual Metabolites** | | | | |
|---|---|---|---|---|
| **Clustering Method** | **Model Quality** | | **Model Reliability** | **Pattern*** |
| | **$R^2X$ (cum)** | **$Q^2$ (cum)** | **p value CV-ANOVA** | |
| **FC** | 0.562 | 0.518 | $4.2e^{-12}$ | A |
| **Kmeans** | 0.565 | 0.590 | $4.6e^{-15}$ | A |
| **EM** | 0.583 | 0.538 | $9.6e^{-14}$ | C |
| **PC score-based, 2 groups** | 0.594 | 0.481 | $2.3e^{-10}$ | C |
| **PC score-based, 3 groups** | 0.618 | 0.451 | $3.8e^{-20}$ | C |
| **Sums of Metabolites** | | | | |
| **Clustering Method** | **Model Quality** | | **Model Reliability** | **Pattern*** |
| | **$R^2X$ (cum)** | **$Q^2$ (cum)** | **p value CV-ANOVA** | |
| **FC** | 0.712 | 0.601 | $6.1e^{-14}$ | A |
| **Kmeans** | 0.635 | 0.492 | $3.4e^{-10}$ | A |
| **EM** | 0.711 | 0.58 | $2.5e^{-13}$ | A |
| **PC score-based, 2 groups** | 0.713 | 0.402 | $1.0e^{-08}$ | C |
| **PC score-based, 3 groups** | 0.713 | 0.452 | $4.3e^{-11}$ | C |

Notes: The two parameters $R^2X$ (cum) and $Q^2$ (cum) represent the model fit (or explained variation) and the predictive ability, respectively. The higher these values, the better the model.
Abbreviations: *p*-value CV-ANOVA is the *p*-value resulting from cross-validation analysis assessing the reliability of the model. The model is valid for *p*-value < 0.05.
* "Pattern" stands for "pattern of metabolite distribution": A, data distribution on the basis of the amount of metabolites excreted (high vs. low); C, data distribution on the basis of the colonic metabolism.

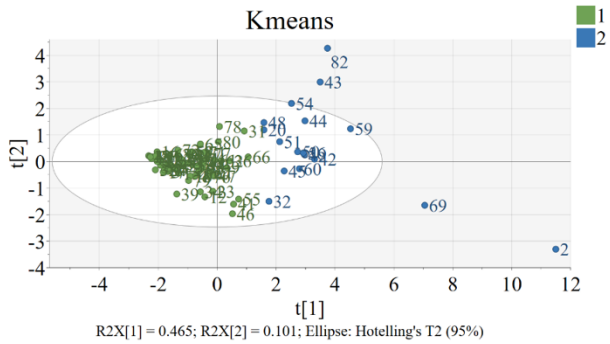
The idea behind sums of metabolites is to highlight colonic metabolism by hiding individual differences attributed to phase II metabolism. Nevertheless, PLS-DA models performed using classes defined by clustering algorithms (FC, EM, Kmeans) yielded similar outputs and presented data distributions based mainly on the amount of metabolites excreted (Table 7, Figure S5A-C, VIP values in Table S2). Briefly, most of the samples were characterized by a low excretion of metabolites whereas a smaller group showed a higher excretion of all of them. A distribution into groups reflecting a different colonic metabolism was however observed in the PLS-DA models obtained using clusters defined by the PC scores (Figure S5D, E). When 2 groups were defined, one group of observations was characterized by the prevalent excretion of 5-(3',4'-dihydroxyphenyl)-γ-valerolactones and 5-(3'-hydroxyphenyl)-γ-valerolactones, while the other group was characterized by the excretion of 5-(4'-hydroxyphenyl)-γ-valerolactone and 3-(hydroxyphenyl)propanoic acid derivatives (Figure S5D). When observations were clustered into 3 groups (Figure S5E), one group showed a high excretion of 5-(4'-hydroxyphenyl)-γ-valerolactone and 3-(hydroxyphenyl)propanoic

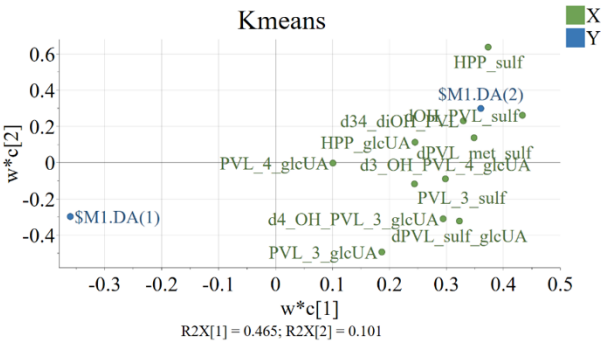
**A**


FC
t[1]
t[2]
R2X[1] = 0.466; R2X[2] = 0.0968; Ellipse: Hotelling's T2 (95%)

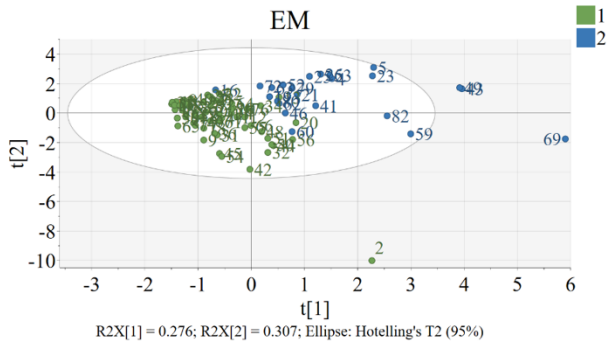

FC
HPP_sulf
$M1.DA(2)
PVL_4_glcUA
$M1.DA(1)
PVL_3_sulf
d4_OH_PVL_3_glcUA
PVL_3_glcUA
dPVL_sulf_glcUA
w*c[1]
w*c[2]
R2X[1] = 0.466; R2X[2] = 0.0968

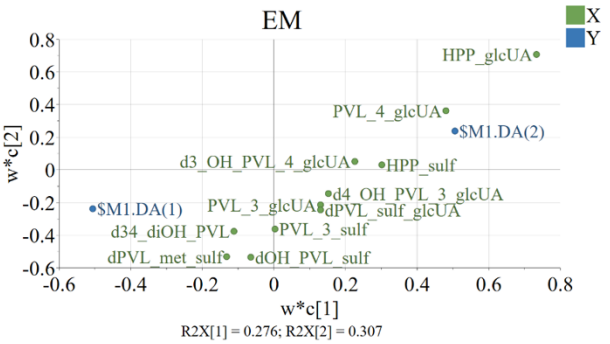
**B**


Kmeans
t[1]
t[2]
R2X[1] = 0.465; R2X[2] = 0.101; Ellipse: Hotelling's T2 (95%)

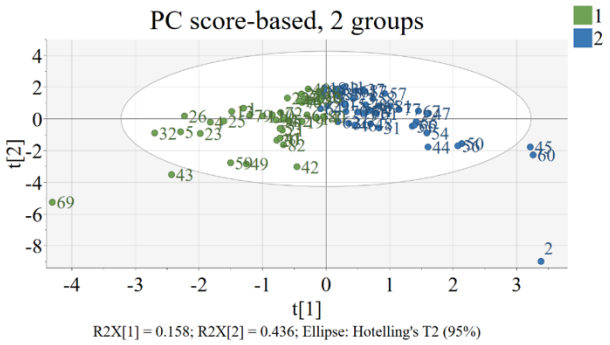

Kmeans
HPP_sulf
$M1.DA(2)
PVL_4_glcUA
$M1.DA(1)
PVL_3_sulf
d4_OH_PVL_3_glcUA
PVL_3_glcUA
dPVL_sulf_glcUA
w*c[1]
w*c[2]
R2X[1] = 0.465; R2X[2] = 0.101

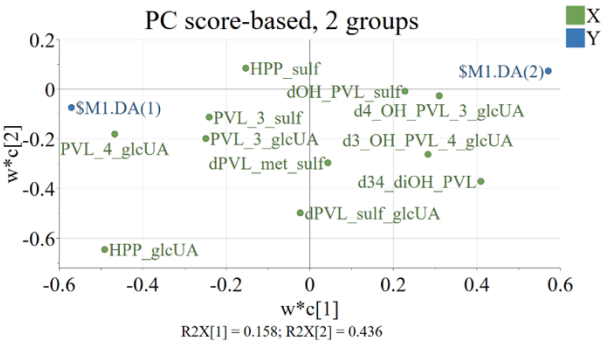
**C**


EM
t[1]
t[2]
R2X[1] = 0.276; R2X[2] = 0.307; Ellipse: Hotelling's T2 (95%)

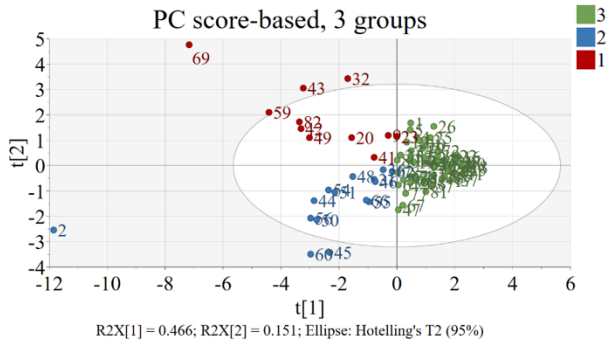

EM
HPP_glcUA
PVL_4_glcUA
$M1.DA(2)
d3_OH_PVL_4_glcUA
HPP_sulf
d4_OH_PVL_3_glcUA
$M1.DA(1)
PVL_3_glcUA
dPVL_sulf_glcUA
d34_diOH_PVL
PVL_3_sulf
dPVL_met_sulf
dOH_PVL_sulf
w*c[1]
w*c[2]
R2X[1] = 0.276; R2X[2] = 0.307

**D**


PC score-based, 2 groups
t[1]
t[2]
R2X[1] = 0.158; R2X[2] = 0.436; Ellipse: Hotelling's T2 (95%)


PC score-based, 2 groups
HPP_sulf
$M1.DA(2)
dOH_PVL_sulf
$M1.DA(1)
d4_OH_PVL_3_glcUA
PVL_3_sulf
PVL_4_glcUA
PVL_3_glcUA
d3_OH_PVL_4_glcUA
dPVL_met_sulf
d34_diOH_PVL
dPVL_sulf_glcUA
HPP_glcUA
w*c[1]
w*c[2]
R2X[1] = 0.158; R2X[2] = 0.436

**E**


PC score-based, 3 groups
t[1]
t[2]
R2X[1] = 0.466; R2X[2] = 0.151; Ellipse: Hotelling's T2 (95%)

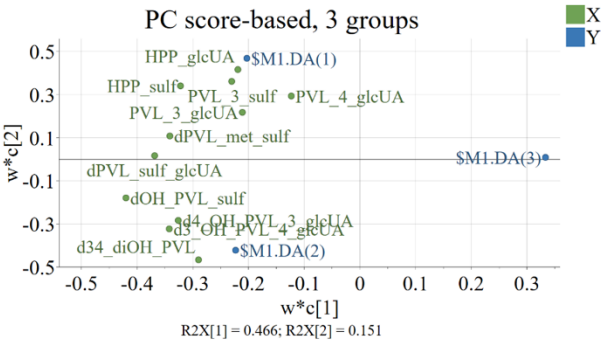

PC score-based, 3 groups
HPP_glcUA
$M1.DA(1)
HPP_sulf
PVL_3_sulf
PVL_4_glcUA
PVL_3_glcUA
dPVL_met_sulf
$M1.DA(3)
dPVL_sulf_glcUA
dOH_PVL_sulf
d4_OH_PVL_3_glcUA
d3_OH_PVL_4_glcUA
d34_diOH_PVL
$M1.DA(2)
w*c[1]
w*c[2]
R2X[1] = 0.466; R2X[2] = 0.151

**Figure 18.** PLS-DA models (score and loading plots) considering individual metabolites and the clusters obtained from different clustering methods: (**A**) final consensus – FC –, (**B**) k-means – Kmeans –, (**C**) expectation-maximization – EM – and PC score-based models for 2 (**D**) or 3 (**E**) groups.

acid derivatives, a second group of 5-(3',4'-dihydroxyphenyl)-γ-valerolactones and 5-(3'-hydroxyphenyl)-γ-valerolactones, and a third larger group was associated to a scarce excretion of metabolites.

*Univariate statistics confirm the differences in metabolite excretion between groups*

Results from PLS-DA models and information on discriminating metabolites were confirmed by univariate statistics. Considering individual metabolites (Table 8), after clustering on the basis of FC and Kmeans, the small group of high excretors of flavan-3-ol catabolites showed statistically significant differences in the urinary excretion of most of the metabolites in comparison with the larger group of low excretors (Table 8), except for 5-phenyl-γ-valerolactone-4'-glucuronide (and 5-phenyl-γ-valerolactone-3'-glucuronide in the case of the Kmeans-based groups). On the contrary, EM clustering identified significant differences only in the quantities of 5-phenyl-γ-valerolactone-4'-glucuronide, 5-(3'-hydroxyphenyl)-γ-valerolactone-4'-glucuronide and 3-(phenyl)propanoic acid glucuronide, these compounds being excreted at higher levels in the group with a lower number of observations. PC score-based clustering into two groups distinguished a cluster of subjects excreting higher amounts of 5-(hydroxyphenyl)-γ-valerolactones (both 3' and 4' derivatives) and 3-(phenyl)propanoic acid glucuronide and lower quantities of some 5-(3',4'-dihydroxyphenyl)-γ-valerolactone derivatives, while the other cluster showed an inverse excretion pattern, clearly marked by the differential colonic metabolism of flavan-3-ols (Table 8). When three PC score-based groups were considered, one larger group was characterized by a low excretion of all metabolites (low excretors), another smaller group by a high excretion of 5-(hydroxyphenyl)-γ-valerolactone (both 3' and 4' isomers) and 3-(hydroxyphenyl)propanoic acid conjugates, as well as a moderate excretion of main 5-(3',4'-dihydroxyphenyl)-γ-valerolactone derivatives, and a third small group by a low/moderate excretion of 5-(hydroxyphenyl)-γ-valerolactone (3'/4') and 3-(hydroxyphenyl)propanoic acid conjugates, and a high excretion of main 5-(3',4'-dihydroxyphenyl)-γ-valerolactone derivatives (Table 8). In this model, all the metabolites reported statistically significant differences among groups.

In order to favour comparisons between data processing strategies, data for individual metabolites were also pooled, once clusters were defined. Results for sums of metabolites belonging to the same aglycone reflected the same trend previously described for individual metabolites on the basis of

each clustering approach (Figure 19). The general trend described for the dataset with individual metabolites was also confirmed when clustering was performed on the datasets with sums of metabolites belonging to the same aglycone. Nevertheless, some differences in comparison to the previous results were observed. In particular, 5-(3',4'-dihydroxyphenyl)-γ-valerolactone was not significant between the two clusters for FC model, while differences in 5-(4'-hydroxyphenyl)-γ-valerolactone were statistically significant (Figure 19). For the Kmeans model, 5-(3'-hydroxyphenyl)-γ-valerolactone and 5-(4'-hydroxyphenyl)-γ-valerolactone were significantly different between the two clusters. 5-(3'-hydroxyphenyl)-γ-valerolactone was significantly different as well in EM model. PC score-based model with 2 groups did not yield statistically significant differences for the sum of 5-(3'-hydroxyphenyl)-γ-valerolactones. Differences in 5-(3',4'-dihydroxyphenyl)-γ-valerolactone excretion between groups for the PC score-based model with 3 groups were the same as reported when individual data was considered, while for the other aglycones some differences in their excretion were found (Figure 19).

The sum of metabolites was also conducted by considering phase II metabolism (sulfation and glucuronidation) using the dataset for individual metabolites. After clustering by FC and Kmeans, the small observation groups presented higher excretion of sulfate and glucuronide derivatives (Table 8). Something similar was seen for the PC score-based model on 3 groups, with 2 groups having a higher excretion of phase II conjugates than the third group, but without differences between the 2 groups with a high excretion of PVLs and HPPs. The PC score-based model on 2 groups did not yield differences in the rate of conjugations between groups, while EM clustering returned a group characterized by the higher excretion of glucuronide (Table 8).

*Sulfate/glucuronide ratio as proxy of individual variability in phase II metabolism*

To investigate deeper the inter-individual variability in phase II metabolism, an aspect that was previously described [25,51], the sulfate/glucuronide ratios of the sums of all metabolites (Figure 20A), 5-(3',4'-dihydroxyphenyl)-γ-valerolactones (Figure 20B), 5-(hydroxyphenyl)-γ-valerolactones (Figure S6A) and 3-(hydroxyphenyl)propanoic acids (Figure S6B) were calculated. 5-phenyl-γ-valerolactone-sulfate-glucuronide isomer (3',4') was excluded from the calculation. In general, the sulfate/glucuronide ratio of the sums of all metabolites (Figure 20A) and of the sums of aglycones (Figure 20B, Figure S6A and B) was in favour of glucuronide conjugates. Just very few subjects ($n$ = 2 for the sum of all metabolites, $n$ = 7 for 5-(3',4'-dihydroxyphenyl)-γ-valerolactones) presented a ratio higher than 1.1, meaning higher excretion of sulfate conjugates. This observation was in line

**Table 8.** Urinary excretion of individual metabolites and sums of metabolites sulfate or glucuronide per each cluster defined after applying different clustering methods (final consensus -FC-, expectation-maximization -EM-, k-means -Kmeans-, and PC score-based models for 2 or 3 groups) on the dataset with individual metabolites. Data are expressed in µmol (mean ± SD). Mean differences between groups were considered significant for $p < 0.05$ and significant *p*-values are highlighted in bold. N indicates the number of observations per cluster.

| Metabolite | Cluster N. | FC | | | Kmeans | | | EM | | | PC score-based, 2 groups | | | PC score-based, 3 groups | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N. | Mean ± SD | *p* value | N. | Mean ± SD | *p* value | N. | Mean ± SD | *p* value | N. | Mean ± SD | *p* value | N. | Mean ± SD | *p* value |
| **5-Phenyl-γ-valerolactone-3'-glucuronide** | 1 | 65 | 0.33 ± 0.42 | **0.041** | 66 | 0.35 ± 0.47 | 0.077 | 60 | 0.42 ± 0.67 | 0.261 | *40* | *0.68 ± 0.91* | ***0.014*** | *11* | *1.26 ± 1.39 a* | ***0.006*** |
| | 2 | 18 | 1.06 ± 1.40 | | 17 | 1.01 ± 1.42 | | 23 | 0.64 ± 1.06 | | *43* | *0.31 ± 0.64* | | *15* | *0.64 ± 1.00 ab* | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | *57* | *0.29 ± 0.42 b* | |
| **5-Phenyl-γ-valerolactone-3'-sulfate** | 1 | *65* | *0.07 ± 0.16* | ***0.000*** | *66* | *0.07 ± 0.16* | ***0.000*** | 60 | 0.15 ± 0.40 | 0.975 | *40* | *0.25 ± 0.49* | ***0.039*** | *11* | *0.67 ± 0.76 a* | ***0.000*** |
| | 2 | *18* | *0.47 ± 0.72* | | *17* | *0.49 ± 0.74* | | 23 | 0.16 ± 0.38 | | *43* | *0.07 ± 0.26* | | *15* | *0.16 ± 0.42 b* | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | *57* | *0.05 ± 0.13 b* | |
| **5-Phenyl-γ-valerolactone-4'-glucuronide** | 1 | 65 | 0.90 ± 0.94 | 0.118 | 66 | 0.90 ± 0.94 | 0.102 | *60* | *0.72 ± 0.58* | ***0.002*** | *40* | *1.42 ± 1.18* | ***0.000*** | 11 | 1.84 ± 1.28 a | **0.005** |
| | 2 | 18 | 1.30 ± 0.95 | | 17 | 1.32 ± 0.98 | | *23* | *1.69 ± 1.33* | | *43* | *0.58 ± 0.38* | | 15 | 0.86 ± 0.37 b | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | 57 | 0.85 ± 0.92 b | |
| **5-(3',4'-Dihydroxyphenyl)-γ-valerolactone** | 1 | *65* | *0.05 ± 0.10* | ***0.000*** | *66* | *0.04 ± 0.10* | ***0.000*** | 60 | 0.12 ± 0.22 | 0.339 | *40* | *0.02 ± 0.08* | ***0.000*** | *11* | *0.08 ± 0.14 b* | ***0.000*** |
| | 2 | *18* | *0.32 ± 0.32* | | *17* | *0.34 ± 0.32* | | 23 | 0.07 ± 0.15 | | *43* | *0.18 ± 0.25* | | *15* | *0.39 ± 0.32 a* | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | *57* | *0.04 ± 0.08 b* | |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **5-(4′-Hydroxyphenyl)-γ-valerolactone-3′-glucuronide** | 1 | *65* | *10.8 ± 8.28* | ***0.000*** | *66* | *11.13 ± 8.64* | ***0.000*** | 60 | 13.52 ± 13.89 | 0.191 | *40* | *10.66 ± 10.02* | ***0.003*** | 11 | 18.78 ± 14.09 ab | **0.000** |
| | 2 | *18* | *28.85 ± 18.56* | | *17* | *28.63 ± 19.11* | | 23 | 17.84 ± 11.93 | | *43* | *18.49 ± 15.15* | | 15 | 30.77 ± 17.65 a | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | 57 | 9.71 ± 7.33 b | |
| **5-(3′-Hydroxyphenyl)-γ-valerolactone-4′-glucuronide** | 1 | *65* | *2.65 ± 2.04* | ***0.000*** | *66* | *2.80 ± 2.37* | ***0.000*** | *60* | *3.40 ± 3.72* | ***0.015*** | 40 | 2.79 ± 2.35 | **0.013** | 11 | 5.14 ± 2.74 a | **0.000** |
| | 2 | *18* | *8.73 ± 6.60* | | *17* | *8.50 ± 6.73* | | *23* | *5.47 ± 5.39* | | 43 | 5.07 ± 5.35 | | 15 | 9.52 ± 6.92 a | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | 57 | 2.29 ± 1.61 b | |
| **5-(Hydroxyphenyl)-γ-valerolactone-sulfate (3′,4′ isomers)** | 1 | *65* | *3.11 ± 3.54* | ***0.000*** | *66* | *3.09 ± 3.51* | ***0.000*** | 60 | 6.40 ± 8.12 | 0.574 | *40* | *4.41 ± 6.04* | *0.074* | 11 | 11.89 ± 7.03 a | **0.000** |
| | 2 | *18* | *16.93 ± 8.87* | | *17* | *17.80 ± 8.31* | | 23 | 5.34 ± 6.46 | | *43* | *7.68 ± 8.71* | | 15 | 15.80 ± 9.90 a | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | 57 | 2.44 ± 2.74 b | |
| **5-Phenyl-γ-valerolactone-sulfate-glucuronide isomer (3′,4′)** | 1 | *65* | *3.19 ± 3.04* | ***0.000*** | *66* | *3.19 ± 3.02* | ***0.000*** | 60 | 4.23 ± 4.90 | 0.260 | 40 | 4.71 ± 4.44 | 0.847 | *11* | *9.44 ± 4.51 a* | ***0.000*** |
| | 2 | *18* | *9.71 ± 6.49* | | *17* | *10.08 ± 6.49* | | 23 | 5.57 ± 4.57 | | 43 | 4.50 ± 5.20 | | *15* | *8.72 ± 6.89 a* | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | *57* | *2.59 ± 2.31 b* | |
| **5-Phenyl-γ-valerolactone-methoxy-sulfate isomer (3′,4′)** | 1 | *65* | *0.01 ± 0.02* | ***0.000*** | *66* | *0.01 ± 0.02* | ***0.000*** | 60 | 0.02 ± 0.03 | 0.252 | 40 | 0.02 ± 0.03 | 0.715 | 11 | 0.05 ± 0.04 a | **0.000** |
| | 2 | *18* | *0.05 ± 0.04* | | *17* | *0.05 ± 0.04* | | 23 | 0.01 ± 0.02 | | 43 | 0.02 ± 0.03 | | 15 | 0.04 ± 0.03 a | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | 57 | 0.01 ± 0.01 b | |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **3-(Phenyl)propanoic acid-sulfate** | 1 | *65* | *0.10 ± 0.17* | ***0.000*** | *66* | *0.10 ± 0.17* | ***0.000*** | 60 | 0.18 ± 0.27 | 0.081 | 40 | 0.35 ± 0.70 | 0.208 | 11 | 1.08 ± 1.02 a | |
| | 2 | *18* | *0.92 ± 0.83* | | *17* | *0.97 ± 0.82* | | 23 | 0.52 ± 0.88 | | 43 | 0.20 ± 0.28 | | 15 | 0.42 ± 0.34 a | **0.000** |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | 57 | 0.08 ± 0.16 b | |
| **3-(Phenyl)propanoic acid-glucuronide** | 1 | *65* | *1.98 ± 1.75* | ***0.001*** | *66* | *1.98 ± 1.73* | ***0.001*** | *60* | *1.46 ± 1.20* | ***0.000*** | *40* | *3.73 ± 2.96* | ***0.000*** | 11 | 6.00 ± 3.88 a | **0.000** |
| | 2 | *18* | *4.54 ± 3.61* | | *17* | *4.68 ± 3.67* | | *23* | *5.34 ± 2.81* | | *43* | *1.43 ± 1.17* | | 15 | 2.33 ± 1.28 b | |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | 57 | 1.92 ± 1.80 b | |
| **SUM of SULFATE conjugates** | 1 | *65* | *3.28 ± 3.65* | ***0.000*** | *66* | *3.27 ± 3.63* | ***0.000*** | 60 | 6.76 ± 8.47 | 0.718 | *40* | *5.03 ± 6.90* | *0.113* | 11 | 13.69 ± 7.89 a | |
| | 2 | *18* | *18.38 ± 9.17* | | *17* | *19.32 ± 8.50* | | 23 | 6.03 ± 7.55 | | *43* | *7.97 ± 9.07* | | 15 | 16.42 ± 10.34 a | **0.000** |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | 57 | 2.58 ± 2.86 b | |
| **SUM of GLUCURONIDE conjugates** | 1 | *65* | *16.66 ± 10.7* | ***0.000*** | *66* | *17.17 ± 11.38* | ***0.000*** | *60* | *19.51 ± 18.24* | ***0.001*** | 40 | 19.27 ± 15.2 | 0.107 | 11 | 33.02 ± 20.05 a | |
| | 2 | *18* | *44.47 ± 24.65* | | *17* | *44.15 ± 25.37* | | *23* | *30.99 ± 17.51* | | 43 | 25.88 ± 21.07 | | 15 | 44.11 ± 24.35 a | **0.000** |
| | 3 | - | - | - | - | - | - | - | - | - | - | - | - | 57 | 15.06 ± 9.31 b | |

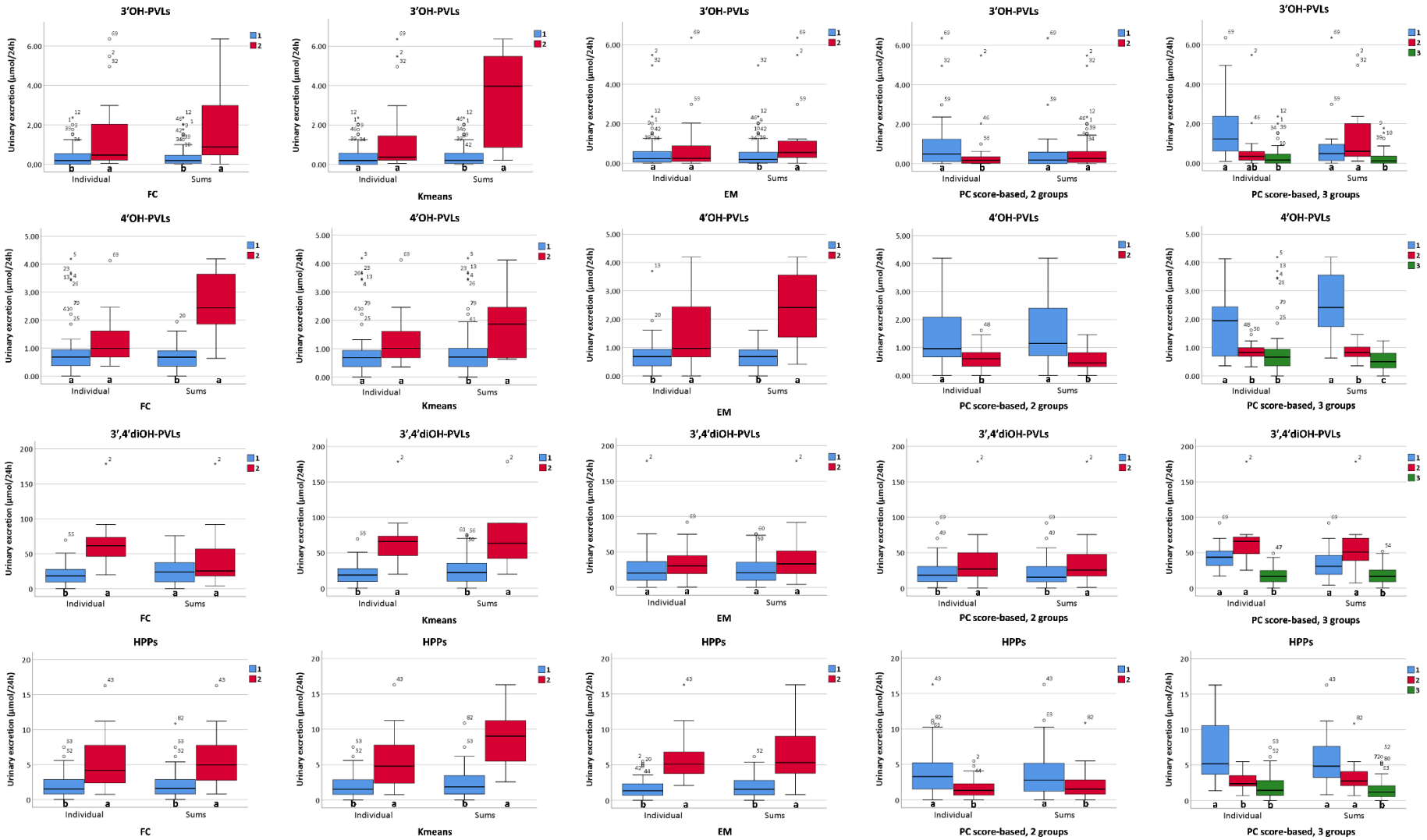



**Figure 19.** Mean urinary excretion (µmol) over 24 h of sums of metabolites belonging to the same aglycone compound (3'OH-PVLs, sum of conjugates from the aglycone 5-(3'-hydroxyphenyl)-γ-valerolactone; 4'OH-PVLs, 5-(4'-hydroxyphenyl)-γ-valerolactone; 3',4'diOH-PVLs, 5-(3',4'-dihydroxyphenyl)-γ-valerolactone; HPPs, 3-(hydroxyphenyl)propanoic acid), calculated both before and after cluster analysis ("Individual" and "Sums", respectively). Clustering has been performed on the basis of: Final Consensus (first column), k-means (second column), expectation-maximization algorithm (third column), PC score forming 2 groups (fourth column), PC score forming 3 groups (fifth column). Different letters indicate statistically significant differences ($p < 0.05$) among groups 1 and 2 or 1,2 and 3.

with one of the PCA models characterized by a more abundant excretion of 5-(hydroxyphenyl)-γ-valerolactone-sulfate (3',4' isomers, Figure 16A,B).

A positive linear correlation was clearly observed (r = 0.977, $p$<0.001) between the sulfate/glucuronide ratio of the sums of all the metabolites and the sulfate/glucuronide ratio of 5-(3',4'-dihydroxyphenyl)-γ-valerolactones (Figure 20C), suggesting that the ratio of the sums of all the sulfate or glucuronide metabolites is mainly influenced by the sulfate/glucuronide ratio of 5-(3',4'-dihydroxyphenyl)-γ-valerolactones, which were actually the main excreted metabolites. Weaker correlations were observed between the sulfate/glucuronide ratio of the sums of metabolites and the sulfate/glucuronide ratio of both 5 (hydroxyphenyl)-γ-valerolactones (r = 0.547, $p$<0.001, Figure S6C) or 3 (hydroxyphenyl)propanoic acids (r = 0.724, $p$<0.001, Figure S6D).

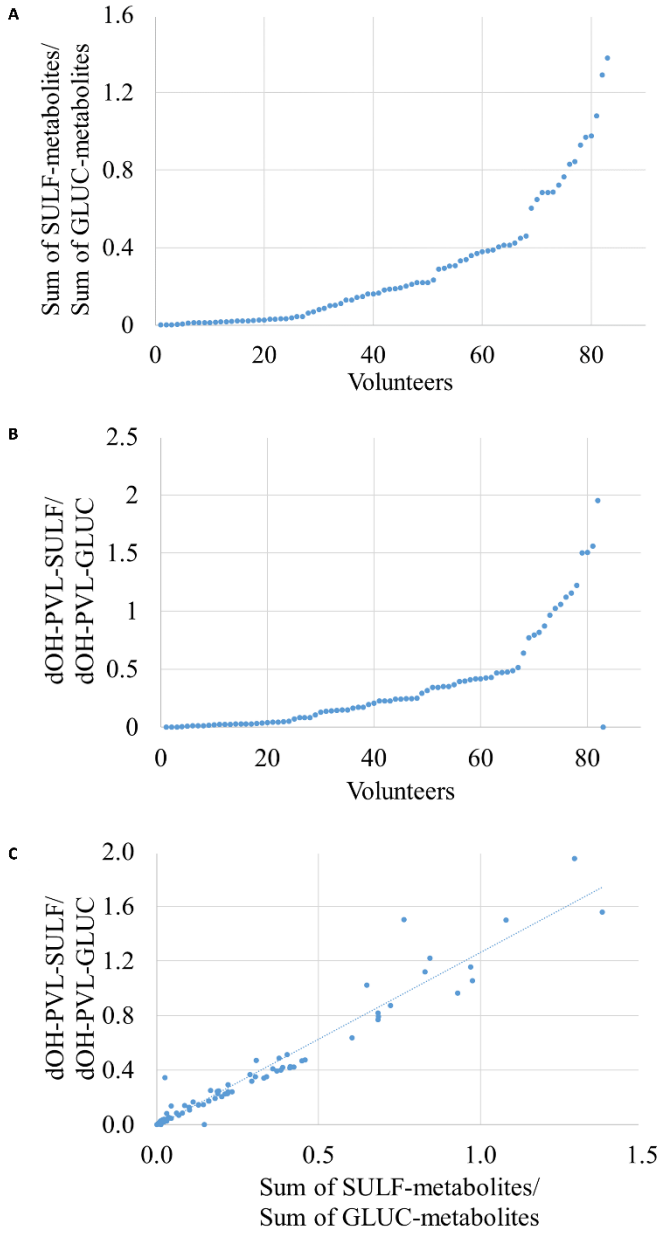



**Figure 20.** Inter-individual variability in phase II metabolism illustrated by the sulfate (SULF)/glucuronide (GLUC) ratio of the sums of respective conjugated metabolites (**A**) and of 5-(3',4'-dihydroxyphenyl)-γ-valerolactones (dOH-PVL) (**B**) in urine samples. (**C**) Relationship between the sulfate/glucuronide ratio of all the metabolites and the sulfate/glucuronide ratio of 5-(3',4'-dihydroxyphenyl)-γ-valerolactones.

## 3.2.4 Discussion

When flavan-3-ols are consumed, different patterns of production and excretion of their main colonic metabolites, namely PVLs and HPPs, can be observed due to the unique genetic asset and microbiota composition of each individual [20–23,25–29]. For instance, some subjects may be more efficient in metabolizing flavan-3-ols and producing higher quantities of metabolites compared to others [25,53], or glucuronidation may be favoured in some individuals rather than sulfation [25,51]. Metabotyping may be a strategy to manage this individual variability and to further investigate its consequences in the impact on the observed health effects

attributed to flavan-3-ols. This work demonstrates that, by applying different strategies of multivariate data analysis, it is possible to cluster all these different metabolic patterns, characterizing groups of individuals. This outcome confirmed thus the existence of metabotypes in the urinary excretion of flavan-3-ol colonic metabolites, as previously hypothesised [26], even though results were different from those preliminary ones.

Data pre-treatment deeply influenced the observations gathered from PCA. No transformation of the data returned higher quality models. Mean centering and mean centering + Pareto scaling highlighted different patterns of phase II metabolism (sulfation vs. glucuronidation), while centering + UV scaling showed different patterns of colonic metabolism. These facts emphasised the importance of data pre-treatment when analysing datasets of flavan-3-ol metabolites, as it is well acknowledged that pre-treatment procedures in metabolomics studies may greatly influence the biological relevance of the results[42].

The inter-individual variability in phase II metabolism observed using certain PCA models revealed that most of the subjects excreted higher quantities of glucuronide derivatives, and this was supported by the analysis of the sulfate/glucuronide ratio and, especially, of the sulfate/glucuronide ratio of 5-(3',4'-dihydroxyphenyl)-γ-valerolactones, quantitatively the main excreted metabolites, as previously discussed [23]. This variability could be related to genetic polymorphisms in phase II enzymes [54–56], but also to the influence of the dose of flavan-3-ols, since the sulfonation pathway has higher affinity but lower capacity than the glucuronidation one, so that when the consumed amount of flavan-3-ols increases, a shift from sulfation toward glucuronidation might occur [57]. This may be an explanation with respect to other works reporting a higher excretion of sulfate derivatives [51,58], together with the lack of the respective reference compound for metabolite quantification [59], but further research is needed to better understand the reasons behind these differences in phase II metabolism. Therefore, to overcome experimental limitations associated with the production and quantification of phase II metabolites, the sums of metabolites belonging to the same aglycone were taken into account, also as a strategy that should lead to a better assessment of colonic metabolism. However, this reductive approach did not yield any benefits in comparison with processing individual metabolite data and calculating sums at the end of the procedure, in line with a previous report [26]. It is worth mentioning that different PCA data pre-treatments should be assessed to fully understand and summarize what happens at colonic level, as well as in phase II conjugation.

Regarding clustering algorithms, EM, rather than the broadly used Kmeans algorithm, was useful in clustering individuals on the basis of their pattern of excretion of colonic metabolites (i.e. flavan-3-ol colonic metabotypes). Kmeans served its purpose to identify groups of individuals with different metabolic profiles in the production of flavan-3-ol colonic metabolites [25], but, in the present work, it was quite influenced by the overall amount of metabolites excreted. The PLS-DA model built using EM clustering was also affected by the excreted amount of metabolites, but it showed a trend towards a different metabolic profiles, as a group of individuals was characterized by a relatively high excretion of 5-(4'-hydroxyphenyl)-γ-valerolactone and HPP derivatives, while the other by a reduced production of these metabolites. The models better highlighting urinary metabotypes of flavan-3-ol colonic metabolites were the two models built using PC scores for clustering. The model with two groups of observations suggested that a small group of subjects is more able to metabolize flavan-3-ols into smaller metabolites (5-(hydroxyphenyl)-γ-valerolactones -both 3' and 4'- and HPPs), while 5-(3',4'-dihydroxyphenyl)-γ-valerolactones was predominant in the main group, fully in line with previous data [25,26]. The model with three groups was able to discriminate observations on the basis of both the total amount excreted and the pattern of colonic metabolites, leading to metabotypes deserving to be further investigated in future bioactivity and functional studies. It should be noted that the PC score-based strategy had the power of well describing the data, but a limited predictive ability. Nevertheless, generalizable predictive models in flavan-3-ol colonic catabolism are not expected due to the chemical complexity of this family of polyphenols and to their variability in dietary sources administered. For example, flavan-3-ols from green tea are rich in trihydroxylated precursors and may lead to more complex metabolic pathways [26], flavan-3-ols from cranberries, the case of the present work, are poor of trihydroxylated precursors and rich in dihydroxylated ones, while the intervention by Cortés-Martín and colleagues [25] consisted of a supplementation of 54.5 mg/d of nut procyanidins, but in free-diet conditions, so that the presence of both dihydroxylated and trihydroxylated precursors is foreseeable according to epidemiological data on the consumption of flavan-3-ols in similar populations [4,5,60,61].

Faecal fermentation of (−)-epicatechin has also highlighted the possible existence of metabotypes among 24 individuals using PCA and hierarchical clustering, where different patterns of (−)-epicatechin catabolism were observed [36]. Common features on metabolic phenotypes have been observed, regardless of the experimental setting, like, for example, the ability of some individuals to metabolise 5-(3',4'-dihydroxyphenyl)-γ-valerolactone into 5-(hydroxyphenyl)-γ-valerolactones and HPPs at a faster pace and the presence of low producers of all metabolites. However, the

allocation of most of the individuals into specific metabotypes will depend on the dataset used, unless further insights on key discriminant metabolites arise, or massive epidemiological evidence is collected to establish specific thresholds.

The biological causes behind the observed metabotypes may rely on the differences in gut microbiota composition of individuals, as this has been reported to be the most important factor modulating the inter-individual variability reported in the colonic metabolism of phenolic compounds [62] and, in particular, of flavan-3-ols [19,36]. Information on specific bacterial strains and enzymes involved in the bioconversion of flavan-3-ols to PVLs and low-molecular weight phenolic acids, as well as on factors that may modulate their activities, is very limited. Up to now, *Adlercreutzia equolifaciens*, *Eggerthella lenta*, *Flavonifractor plautii*, and *Lactobacillus plantarum* IFPL935 are the only bacteria identified as responsible for the catabolism of flavan-3-ols into 5-(3',4'-dihydroxyphenyl)-γ-valerolactone and 5-(3'-hydroxyphenyl)-γ-valerolactone [18,63,64], but, for example, 5-(4'-hydroxyphenyl)-γ-valerolactone has not been described as one of their catabolic products. In addition, no microorganisms responsible for further β-oxidation into 3-(hydroxyphenyl)propanoic acids have been reported, yet.

Besides a better understanding of the metabolism and bioavailability of flavan-3-ols, the importance of identifying different metabotypes relies on the possibility of unravelling the health effects associated with flavan-3-ol consumption and associated to their microbiota-derived metabolites, as it has been described for isoflavones (with equol production) and ellagitannins (with urolithin production) [33–35]. For instance, *in vitro* findings support the biological effects of flavan-3-ol colonic metabolites against uropathogenic *Escherichia coli* adherence to uroepithelial cells [12,13], while it is well-known that human studies administering cranberry flavan-3-ols to prevent urinary tract infections (UTIs) have reported conflicting results [65,66]. This might be due to different profiles of excreted metabolites, exerting different biological effects. In this sense, clustering subjects according to their urinary metabotype of flavan-3-ol colonic metabolites may provide new insights in the actual effect of flavan-3-ols on UTI prevention, not only through cranberries but potentially also from other flavan-3-ol food sources like cocoa, wine, pome fruits, other berries, and nuts.

## 3.2.5. Conclusion

The current work shed light on the existence of metabotypes in the urinary excretion of flavan-3-ol metabolites, which are not characterized by the production/non-production of specific metabolites, but by different quali-quantitative metabolic profiles. A series of univariate and multivariate tools, all broadly accessible to the research community, highlighted the importance of data pre-treatment

and clustering methods on the final outcomes for a given dataset. Different profiles in the urinary excretion of PLVs and HPPs were observed upon cranberry consumption in two diverse experimental settings, these metabolic profiles being related to not only specific pathways of phase II metabolism but also the type of metabolites produced at colonic level. Insights depended on PCA data pre-treatment: non-transformed, centered, and UV-scaled data were key to unravel metabolic patterns based on colonic metabolism, while other approaches favoured differences in phase II metabolism. Regarding clustering, while Kmeans and a FC algorithm highlighted differences in the overall production of PVLs and HPPs, the EM algorithm and PC score-based clustering yielded well-defined metabotypes in the urinary excretion of these metabolites. The true physiological relevance of each metabotyping model, whether based on phase II or colonic metabolism, will be related to the application of these inter-individual differences to explore their potential impact on the biological activity of this major (poly)phenol subclass. When applied to physiological outcomes, different ways of metabotyping may lead to different biological observations, fostering the understanding of the impact of flavan-3-ols on human health. The unambiguous elucidation of metabotypes and the allocation of subjects into a metabotype or another, when dealing with (poly)phenols not characterized by the selective production of specific metabolites, will likely depend on the datasets considered for the development of further predictive models. Therefore, these results, both the proposed metabotypes and the defined procedures, should be validated in larger datasets involving a higher number of participants, more phenolic metabolites from the flavan-3-ol metabolic pathway, and different sources of flavan-3-ols. In any case, this work represents an additional step toward the understanding of the bioavailability of flavan-3-ols and the inter-individual variability associated to these compounds and it will be useful for future studies aiming to investigate metabolic phenotypes in the production and urinary excretion of other classes of phenolic metabolites.

**Funding**

This study was partially funded by the Cranberry Institute. Ocean Spray donated the cranberry juices and powder used in the human studies and provided composition analyses. This funder had no input on the design, implementation, analysis or interpretation of the data. In addition, PM has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (PREDICT-CARE project, grant agreement No 950050).

**Conflicts of Interest**

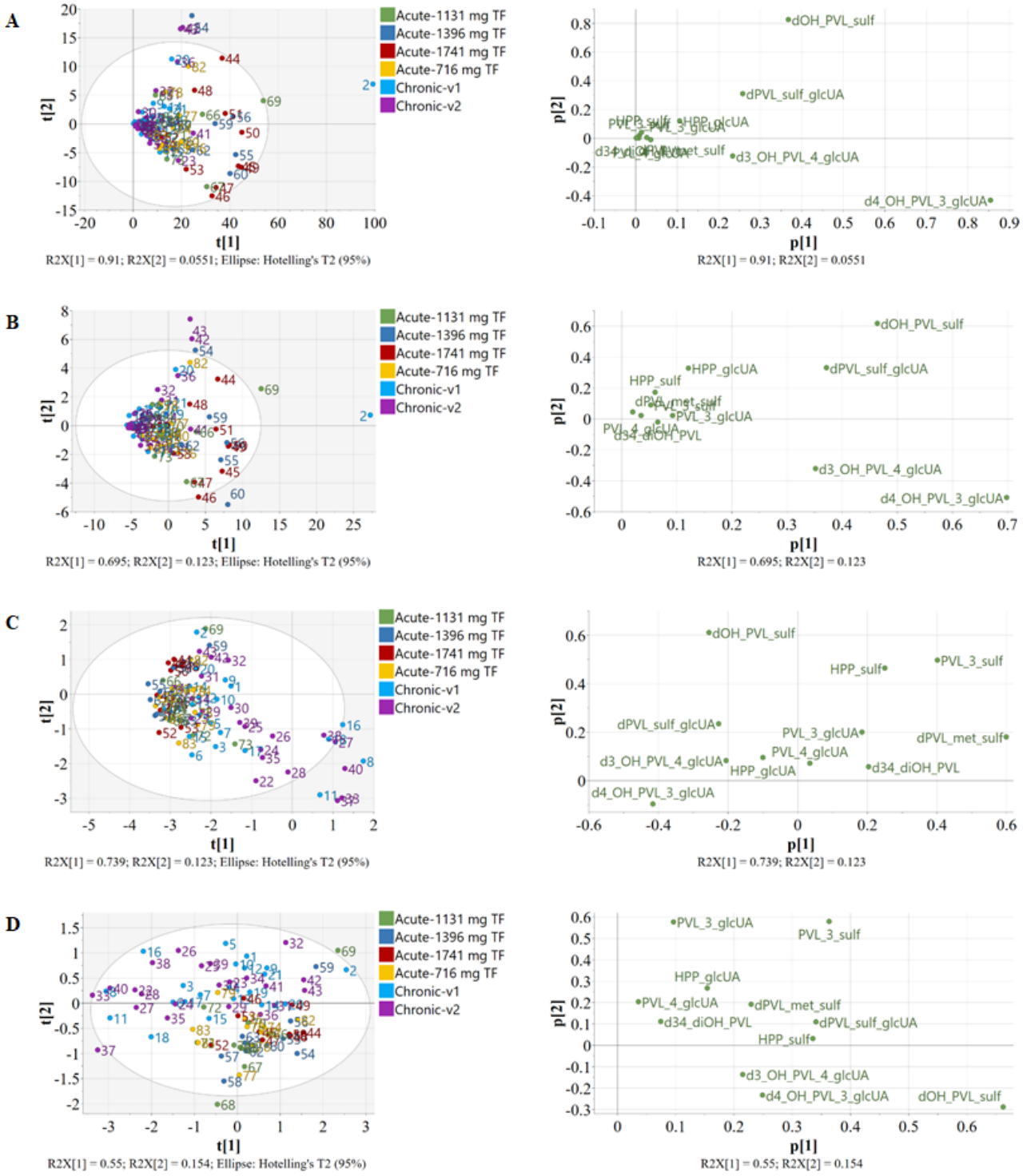
PM, ARM, and DDR received a research grant from the Cranberry Institute. The rest of the authors declare no conflict of interest.

**Authors' contributions**

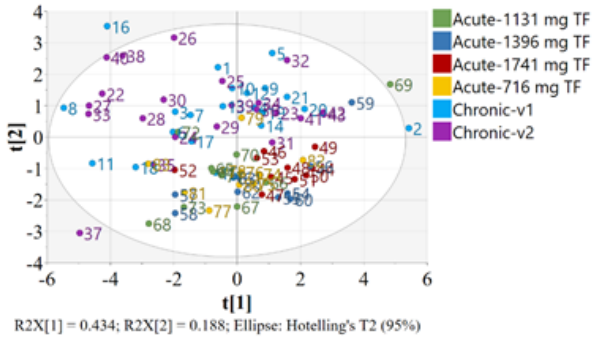
PM: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing - original draft; CF: Formal analysis, Investigation, Methodology, Visualization, Writing - original draft; AA: Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing - review & editing; SC: Formal analysis, Methodology, Software, Visualization, Writing - review & editing; LB: Formal analysis, Writing - review & editing; CC: Resources, Methodology, Writing - review & editing; FB: Resources, Methodology, Writing - review & editing; CH: Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing - review & editing; ARM: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing - review & editing; DDR: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing - review & editing. All authors approved the final version of the manuscript to be submitted.

## Electronic Supplementary Information (ESI)

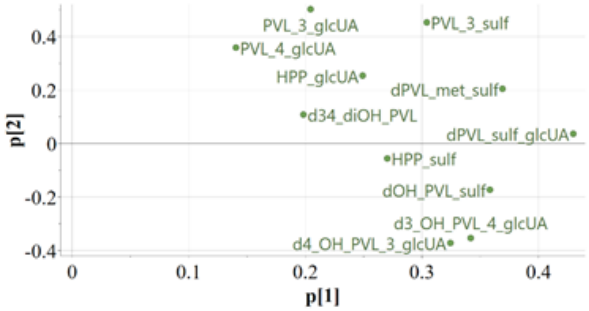
**Figure S1A-L.** PCA models (score and loading plots) for individual metabolites on: (**A**) non-transformed non-scaled data, (**B**) non-transformed centered and Pareto scaled data, (**C**) log-transformed non-scaled data, (**D**) log-transformed centered data, (**E**) log-transformed autoscaled data, (**F**) log-transformed centered and Pareto scaled data, (**G**) power-transformed non-scaled data, (**H**) power-transformed centered data, (**I**) power-transformed autoscaled data, (**L**) power-transformed centered and Pareto scaled data.



**E**

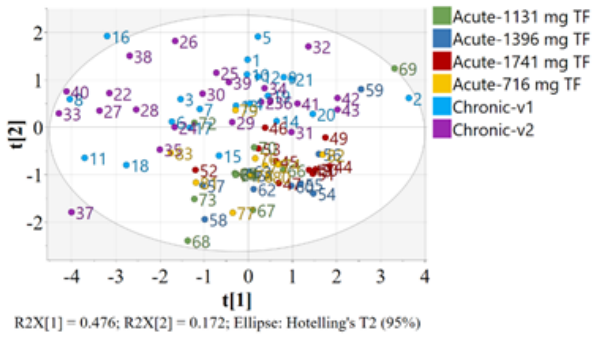



R2X[1] = 0.434; R2X[2] = 0.188; Ellipse: Hotelling's T2 (95%)

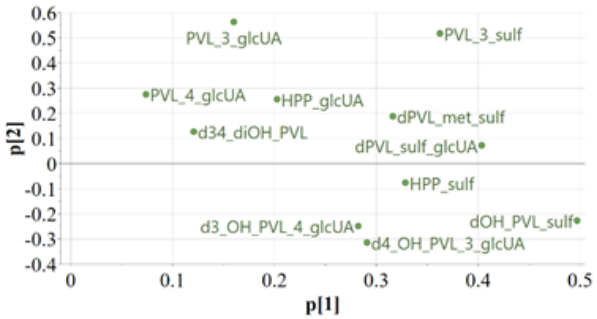



R2X[1] = 0.434; R2X[2] = 0.188

**F**

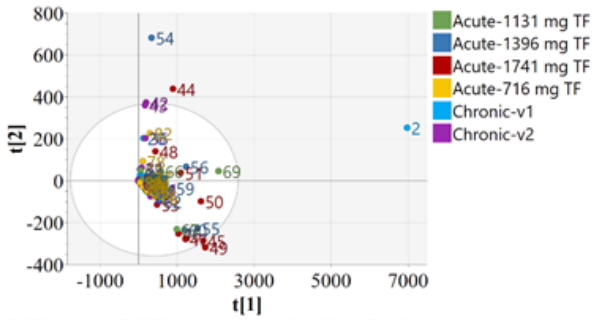



R2X[1] = 0.476; R2X[2] = 0.172; Ellipse: Hotelling's T2 (95%)

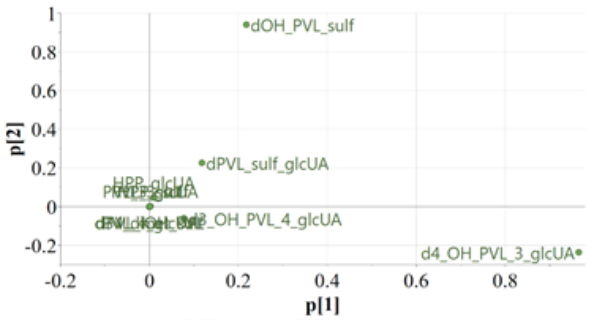



R2X[1] = 0.476; R2X[2] = 0.172

**G**

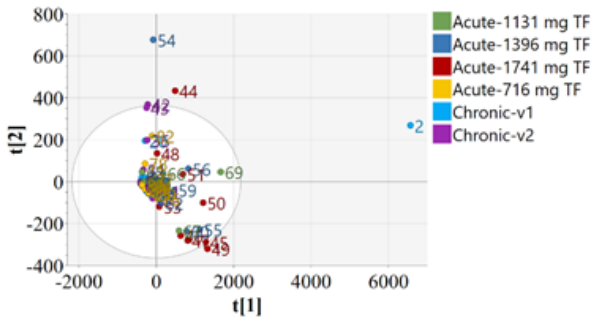



R2X[1] = 0.969; R2X[2] = 0.0221; Ellipse: Hotelling's T2 (95%)

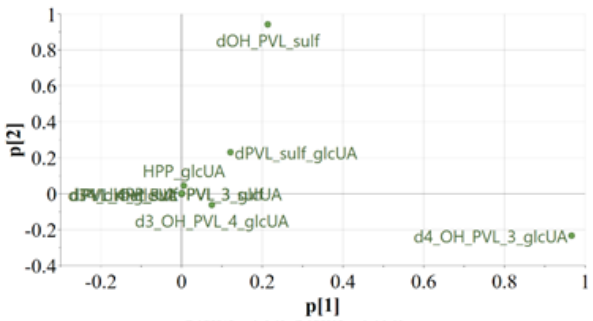



R2X[1] = 0.969; R2X[2] = 0.0221

**H**

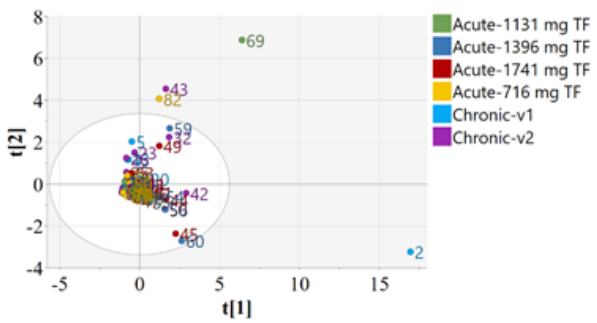



R2X[1] = 0.963; R2X[2] = 0.0268; Ellipse: Hotelling's T2 (95%)

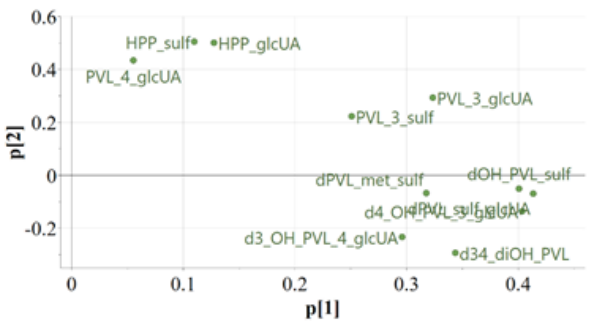



R2X[1] = 0.963; R2X[2] = 0.0268

**I**




R2X[1] = 0.458; R2X[2] = 0.164; Ellipse: Hotelling's T2 (95%)




R2X[1] = 0.458; R2X[2] = 0.164

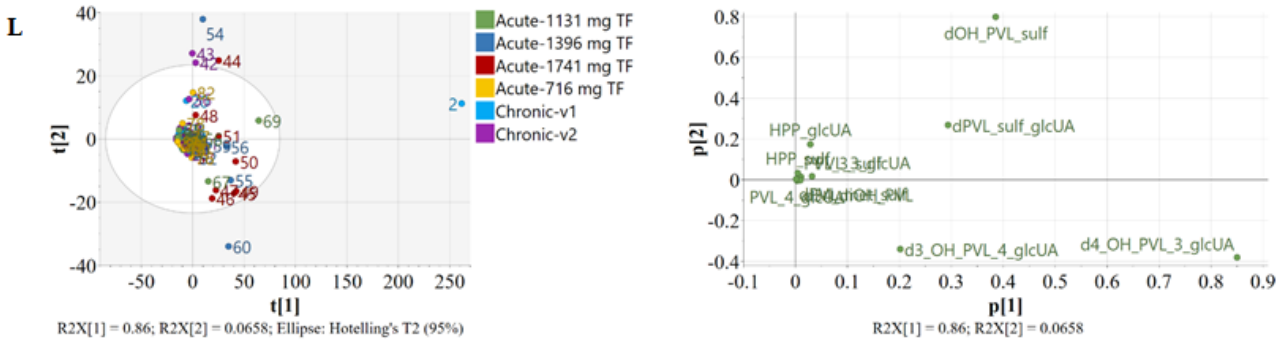
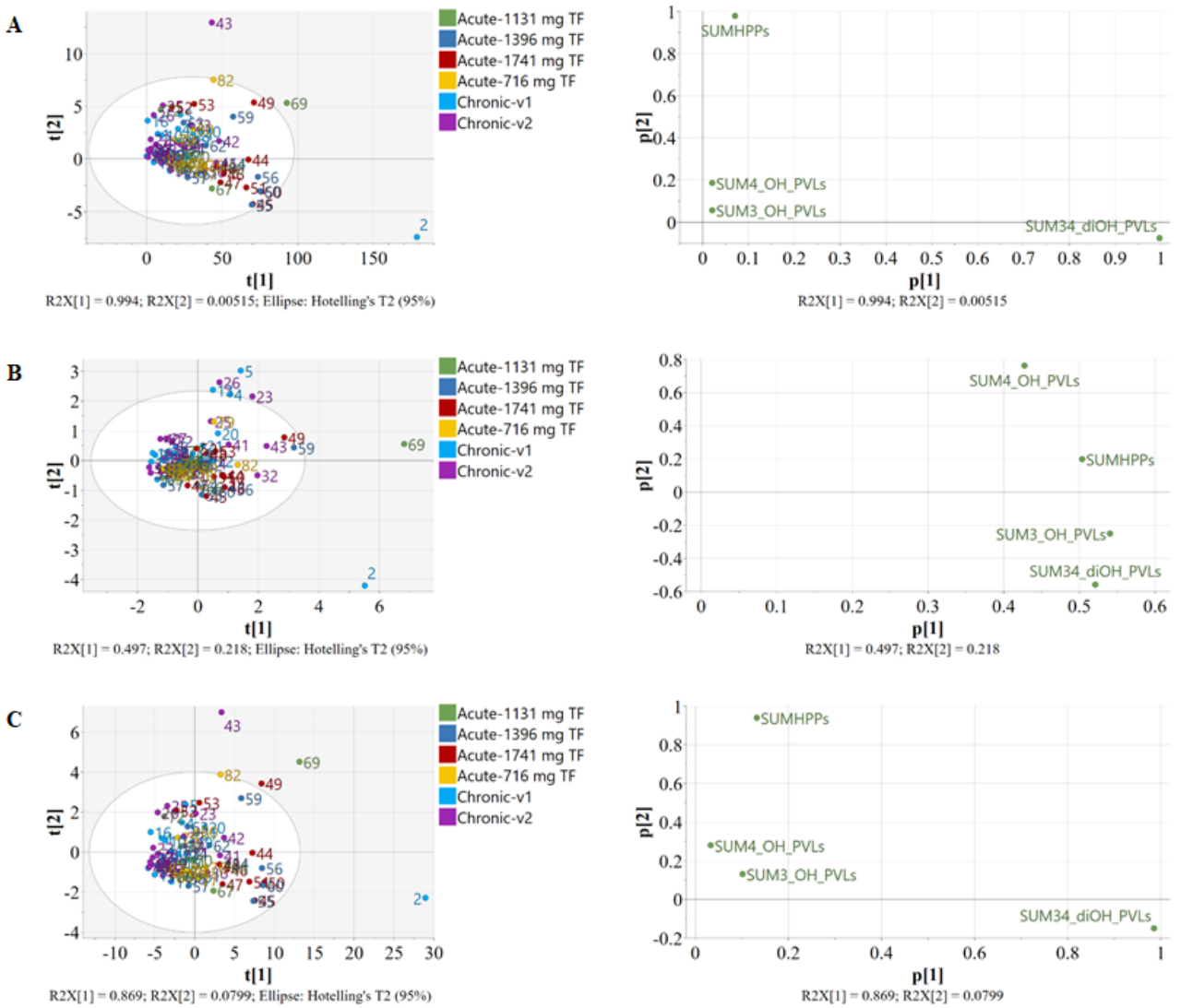
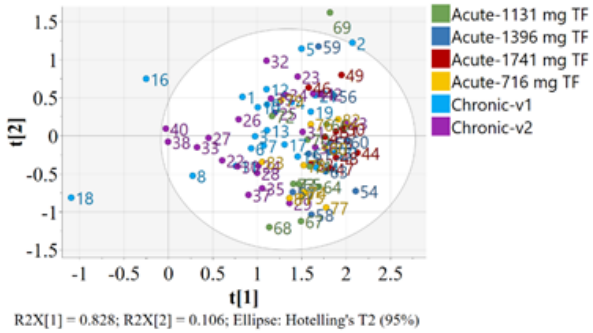


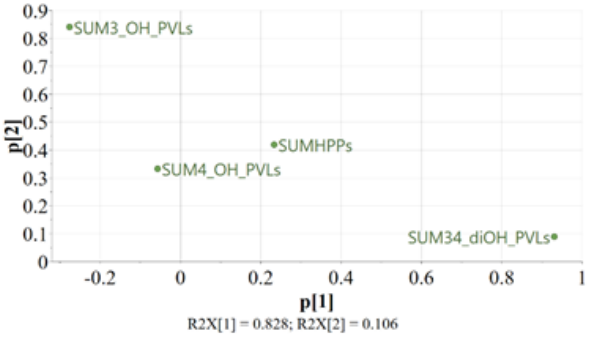

**Figure S2A-M.** PCA models (score and loading plots) for sums of metabolites belonging to the same aglycone compound on: (**A**) non-transformed non-scaled data, (**B**) non-transformed autoscaled data, (**C**) non-transformed centered and Pareto scaled data, (**D**) log-transformed non-scaled data, (**E**) log-transformed centered data, (**F**) log-transformed autoscaled data, (**G**) log-transformed centered and Pareto scaled data, (**H**) power-transformed non-scaled data, (**I**) power-transformed centered data, (**L**) power-transformed autoscaled data, (**M**) power-transformed centered and Pareto scaled data.



**D**

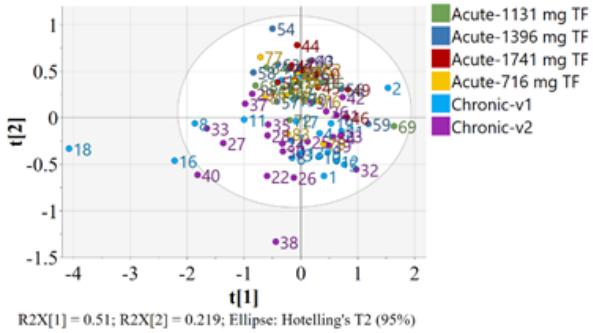



R2X[1] = 0.828; R2X[2] = 0.106; Ellipse: Hotelling's T2 (95%)

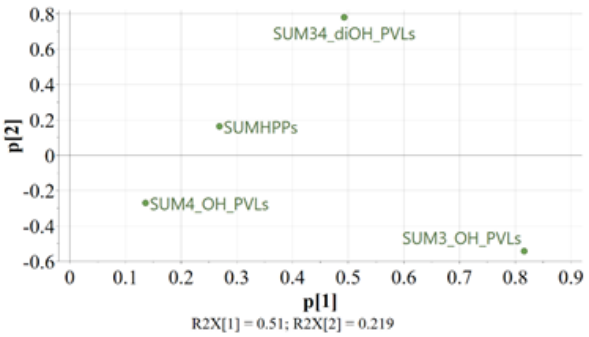



R2X[1] = 0.828; R2X[2] = 0.106

**E**

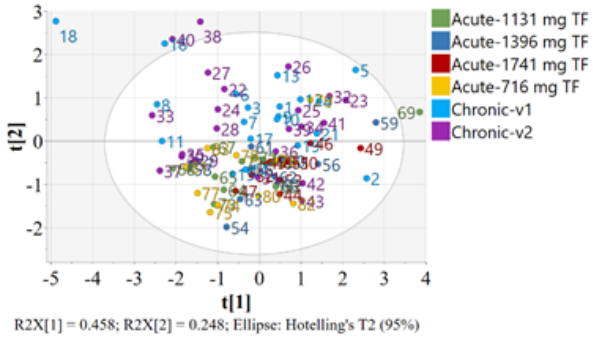



R2X[1] = 0.51; R2X[2] = 0.219; Ellipse: Hotelling's T2 (95%)

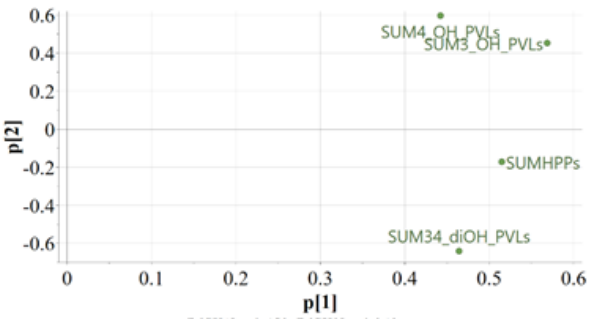



R2X[1] = 0.51; R2X[2] = 0.219

**F**

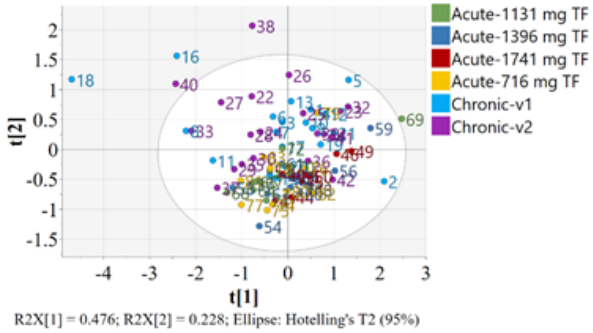



R2X[1] = 0.458; R2X[2] = 0.248; Ellipse: Hotelling's T2 (95%)

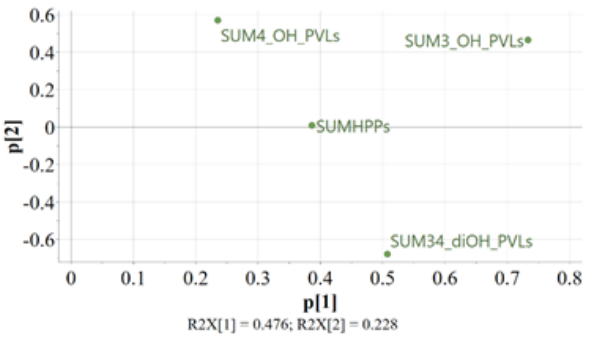



R2X[1] = 0.458; R2X[2] = 0.248

**G**

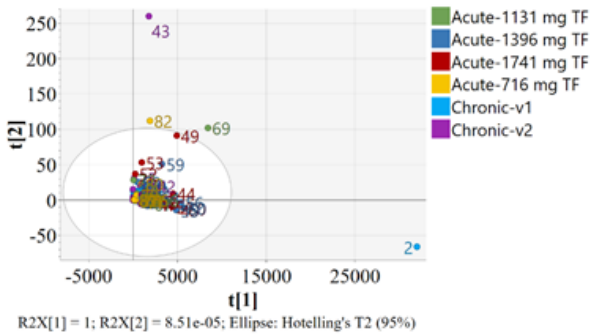



R2X[1] = 0.476; R2X[2] = 0.228; Ellipse: Hotelling's T2 (95%)

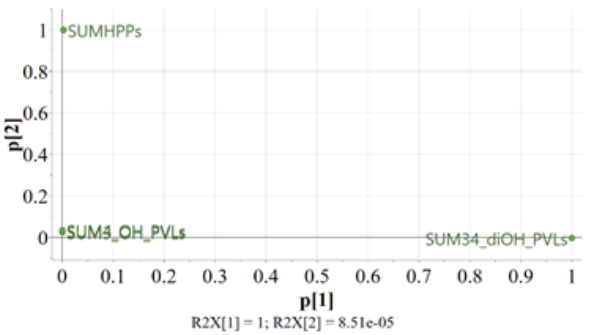



R2X[1] = 0.476; R2X[2] = 0.228

**H**




R2X[1] = 1; R2X[2] = 8.51e-05; Ellipse: Hotelling's T2 (95%)




R2X[1] = 1; R2X[2] = 8.51e-05

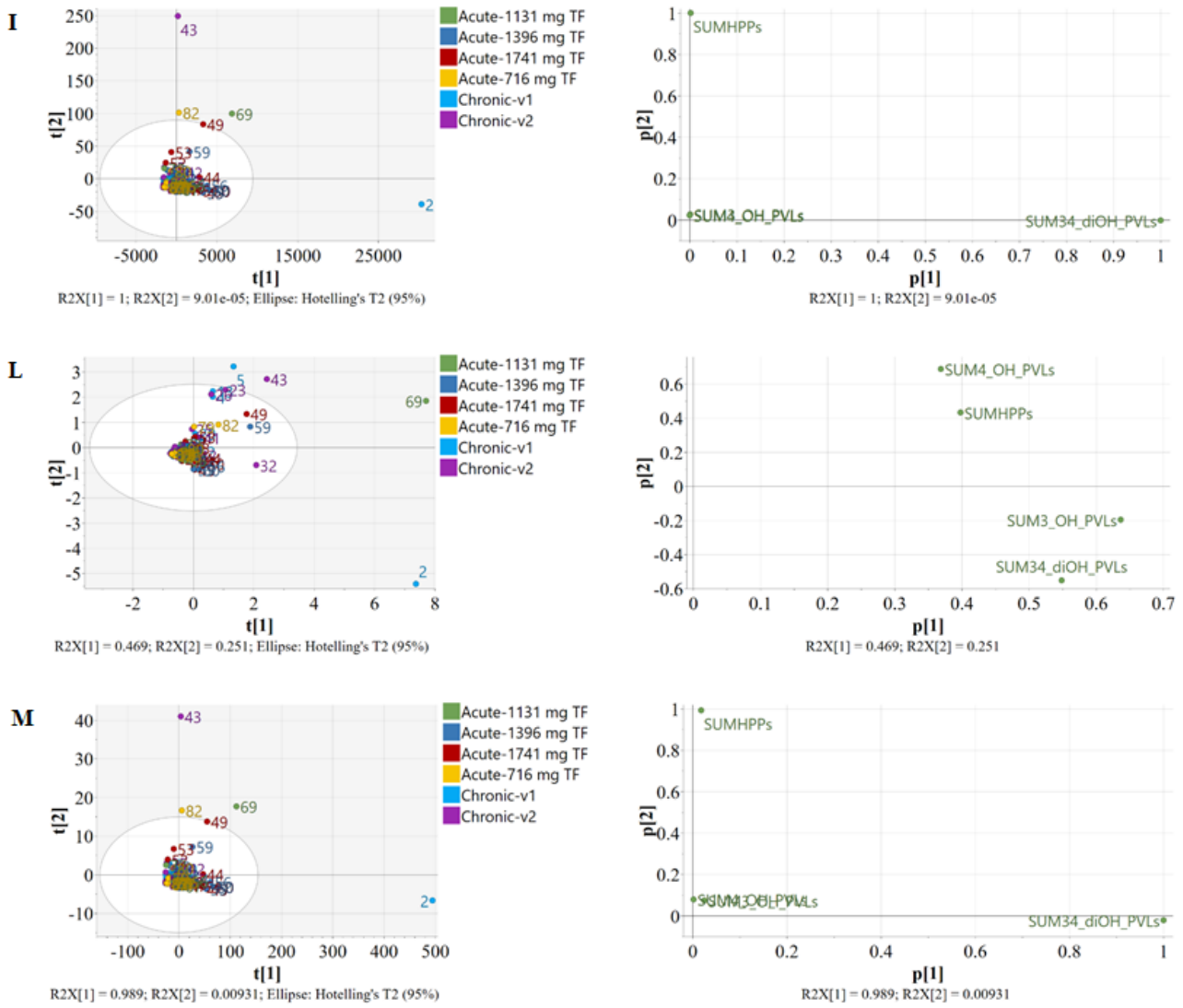
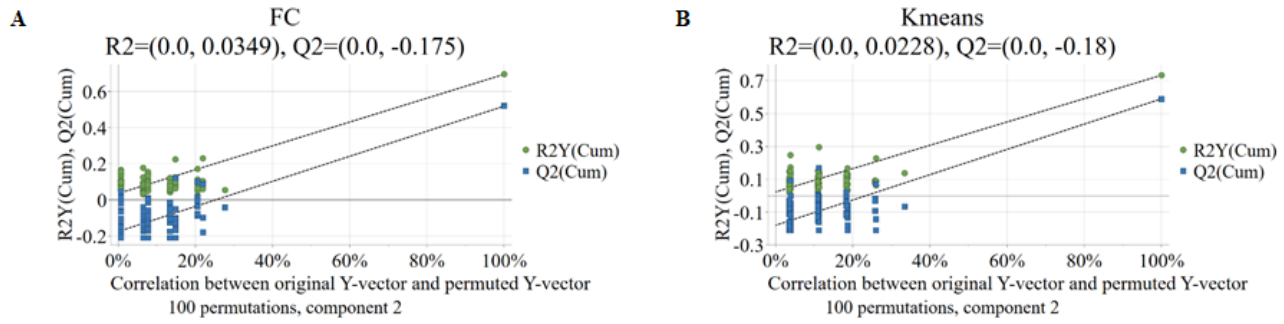



**Figure S3A-E.** Permutation plots of the PLS-DA models considering individual metabolites and the clusters obtained from different clustering methods: (**A**) final consensus – FC –, (**B**) k-means – Kmeans –, (**C**) expectation-maximization – EM – and PC score-based models for 2 (**D**) or 3 (**E**) groups.



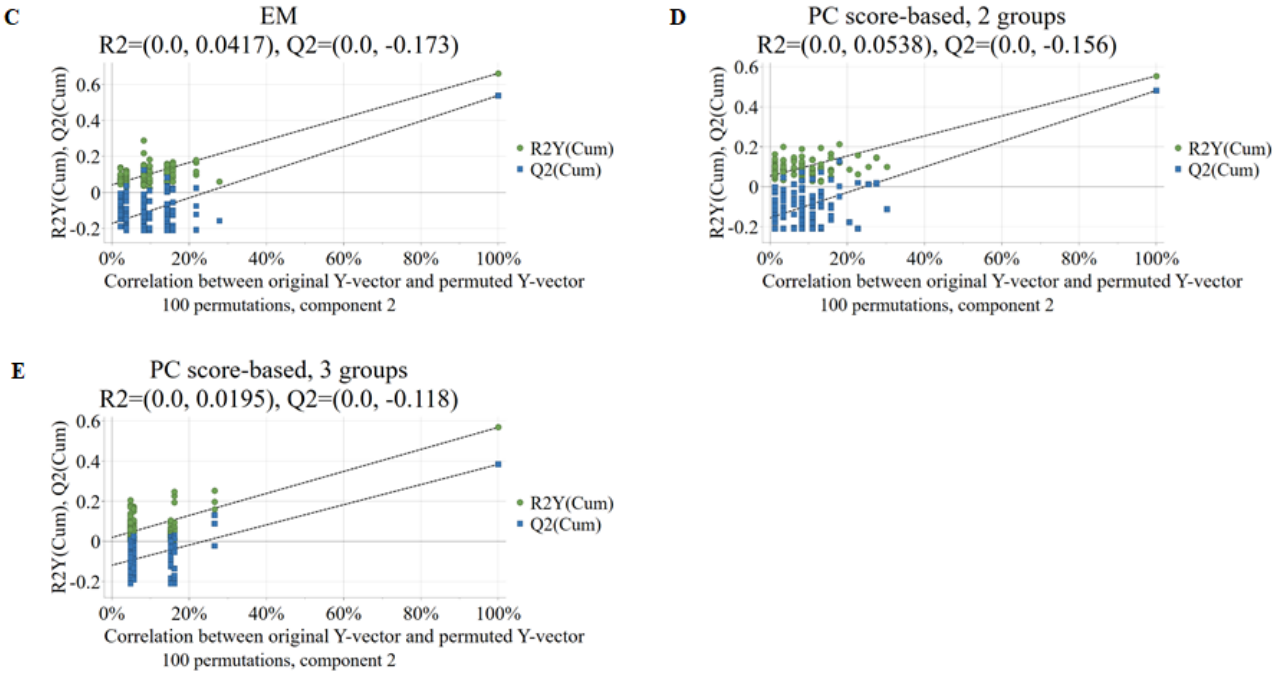
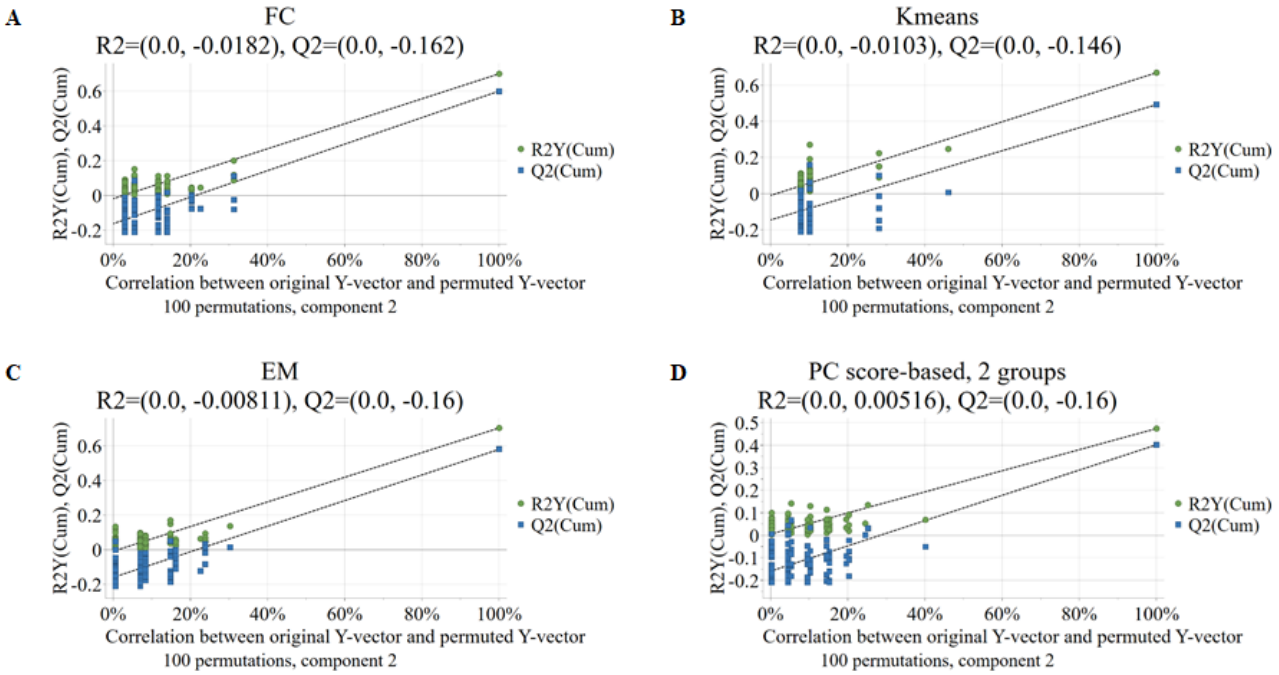



**Figure S4A-E.** Permutation plots of the PLS-DA models considering sums of metabolites belonging to the same aglycone family and the clusters obtained from different clustering methods: (**A**) final consensus – FC –, (**B**) k-means – Kmeans –, (**C**) expectation-maximization – EM – and PC score-based models for 2 (**D**) or 3 (**E**) groups.



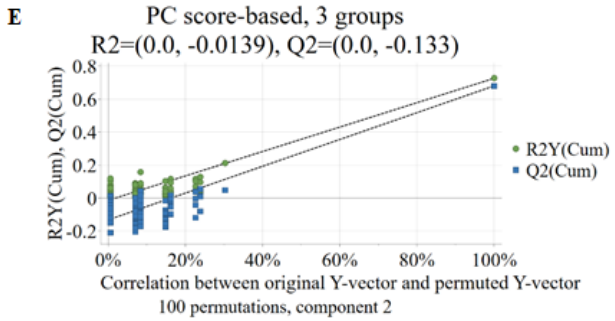
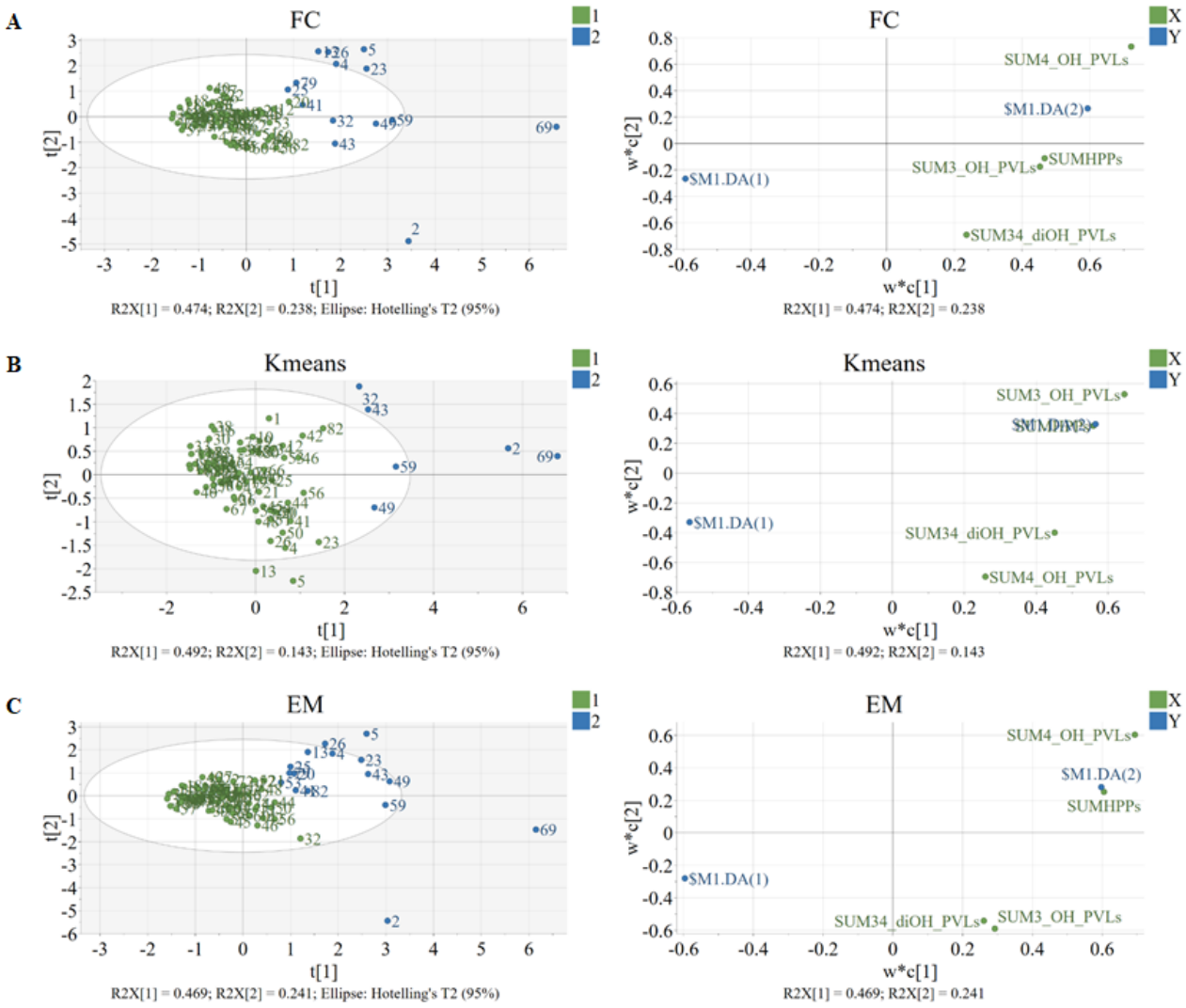



**Figure S5A-E.** PLS-DA models (score and loading plots) considering sums of metabolites belonging to the same aglycone family and the clusters obtained from different clustering methods: (**A**) final consensus – FC –, (**B**) k-means – Kmeans –, (**C**) expectation-maximization – EM – and PC score-based models for 2 (**D**) or 3 (**E**) groups.



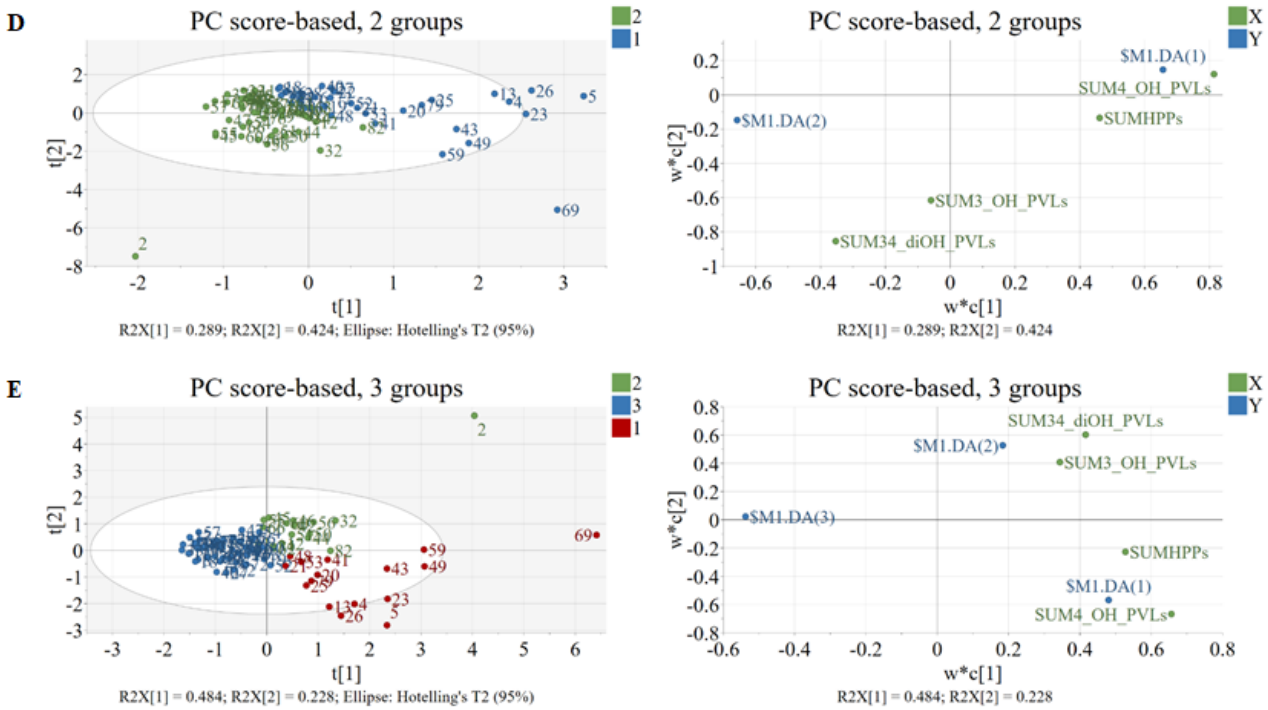



**Figure S6**. Inter-individual variability in phase II metabolism illustrated by the sulfate (SULF)/glucuronide (GLUC) ratio of 5-(monohydroxyphenyl)-γ-valerolactones (monoOH-PVLs) (**A**) and 3-(hydroxyphenyl)propanoic acids (HPPs) (**B**) in urine samples. Relation between the ratio of the sums of all the sulfate or glucuronide metabolites values and the sulfate/glucuronide ratio of 5-(monohydroxyphenyl)-γ-valerolactones (**C**) or 3-(hydroxyphenyl)propanoic acids (**D**).

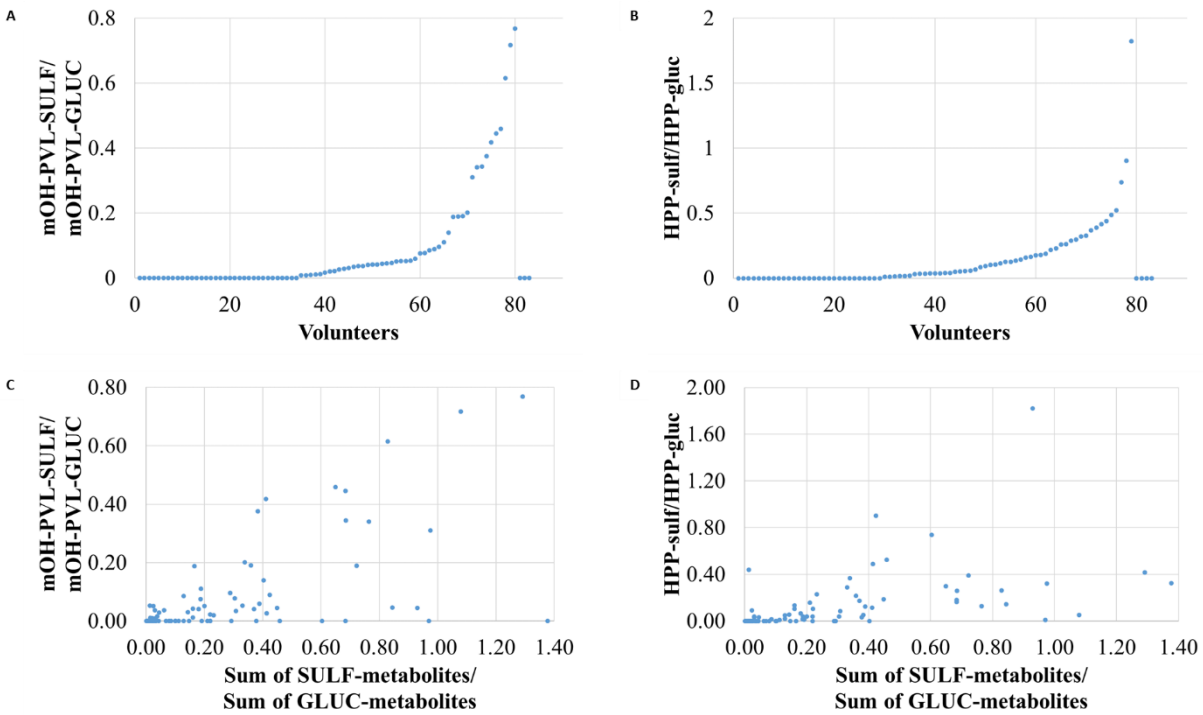


**Table S1.** Classification in groups according to every clustering method.

| Study | Subject | Identifier | INDIVIDUAL METABOLITES | | | | | SUMS OF METABOLITES BELONGING TO THE SAME AGLYCONE COMPOUND | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | FC | Kmeans | EM | PC score-based, 2 groups | PC score-based, 3 groups | FC | Kmeans | EM | PC score-based, 2 groups | PC score-based, 3 groups |
| Chronic-v1 | 1 | 1 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 2 | 2 |
| Chronic-v1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| Chronic-v1 | 3 | 3 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v1 | 5 | 4 | 1 | 1 | 2 | 1 | 3 | 2 | 1 | 2 | 1 | 1 |
| Chronic-v1 | 9 | 5 | 1 | 1 | 2 | 1 | 3 | 2 | 1 | 2 | 1 | 1 |
| Chronic-v1 | 10 | 6 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v1 | 11 | 7 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v1 | 12 | 8 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v1 | 14 | 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v1 | 17 | 10 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v1 | 19 | 11 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v1 | 21 | 12 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 2 | 2 |
| Chronic-v1 | 24 | 13 | 1 | 1 | 1 | 1 | 3 | 2 | 1 | 2 | 1 | 1 |
| Chronic-v1 | 25 | 14 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v1 | 27 | 15 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v1 | 32 | 16 | 1 | 1 | 2 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v1 | 34 | 17 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v1 | 37 | 18 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v1 | 39 | 19 | 1 | 1 | 2 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v1 | 41 | 20 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| Chronic-v1 | 42 | 21 | 1 | 1 | 2 | 1 | 3 | 1 | 1 | 1 | 1 | 1 |
| Chronic-v2 | 1 | 22 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v2 | 2 | 23 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 1 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Chronic-v2 | 3 | 24 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v2 | 5 | 25 | 1 | 1 | 2 | 1 | 3 | 2 | 1 | 2 | 1 | 1 |
| Chronic-v2 | 9 | 26 | 1 | 1 | 2 | 1 | 3 | 2 | 1 | 2 | 1 | 1 |
| Chronic-v2 | 10 | 27 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v2 | 11 | 28 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v2 | 12 | 29 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v2 | 14 | 30 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v2 | 16 | 31 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v2 | 17 | 32 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 |
| Chronic-v2 | 19 | 33 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v2 | 21 | 34 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 2 | 2 |
| Chronic-v2 | 24 | 35 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v2 | 25 | 36 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v2 | 27 | 37 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v2 | 32 | 38 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v2 | 34 | 39 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 2 | 3 |
| Chronic-v2 | 37 | 40 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Chronic-v2 | 39 | 41 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 1 |
| Chronic-v2 | 41 | 42 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| Chronic-v2 | 42 | 43 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 |
| Acute-1741 mg TF | 1 | 44 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Acute-1741 mg TF | 2 | 45 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Acute-1741 mg TF | 3 | 46 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Acute-1741 mg TF | 4 | 47 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1741 mg TF | 5 | 48 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| Acute-1741 mg TF | 6 | 49 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 |
| Acute-1741 mg TF | 7 | 50 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Acute-1741 mg TF | 8 | 51 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Acute-1741 mg TF | 9 | 52 | 1 | 1 | 2 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Acute-1741 mg TF | 10 | 53 | 1 | 1 | 2 | 1 | 3 | 1 | 1 | 2 | 1 | 1 |
| Acute-1396 mg TF | 1 | 54 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 3 |
| Acute-1396 mg TF | 2 | 55 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Acute-1396 mg TF | 3 | 56 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Acute-1396 mg TF | 4 | 57 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1396 mg TF | 5 | 58 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1396 mg TF | 6 | 59 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 |
| Acute-1396 mg TF | 7 | 60 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Acute-1396 mg TF | 8 | 61 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1396 mg TF | 9 | 62 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Acute-1396 mg TF | 10 | 63 | 1 | 1 | 2 | 1 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1131 mg TF | 1 | 64 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1131 mg TF | 2 | 65 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1131 mg TF | 3 | 66 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Acute-1131 mg TF | 4 | 67 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1131 mg TF | 5 | 68 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1131 mg TF | 6 | 69 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 1 |
| Acute-1131 mg TF | 7 | 70 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-1131 mg TF | 8 | 71 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 1 | 3 |
| Acute-1131 mg TF | 9 | 72 | 1 | 1 | 2 | 1 | 3 | 1 | 1 | 1 | 1 | 3 |
| Acute-1131 mg TF | 10 | 73 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-716 mg TF | 1 | 74 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-716 mg TF | 2 | 75 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-716 mg TF | 3 | 76 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-716 mg TF | 4 | 77 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-716 mg TF | 5 | 78 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-716 mg TF | 6 | 79 | 1 | 1 | 2 | 1 | 3 | 2 | 1 | 2 | 1 | 1 |
| Acute-716 mg TF | 7 | 80 | 1 | 1 | 2 | 1 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-716 mg TF | 8 | 81 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |
| Acute-716 mg TF | 9 | 82 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| Acute-716 mg TF | 10 | 83 | 1 | 1 | 1 | 2 | 3 | 1 | 1 | 1 | 2 | 3 |

**Table S2.** VIP values for individual metabolites and aglycone compounds of each PLS-DA model considering classes defined by each clustering method selected (final consensus – FC –, k-means – Kmeans –, expectation-maximization – EM – and PC score-based models for 2 or 3 groups).

| Metabolites | VIP value | | | | |
|---|---|---|---|---|---|
| | FC | Kmeans | EM | PC score-based, 2groups | PC score-based, 3groups |
| **5-Phenyl-γ-valerolactone-3'-glucuronide** | 0.80 | 0.83 | 0.66 | 0.82 | 0.71 |
| **5-Phenyl-γ-valerolactone-3'-sulfate** | 0.78 | 0.78 | 0.63 | 0.79 | 0.96 |
| **5-Phenyl-γ-valerolactone-4'-glucuronide** | 0.31 | 0.31 | **1.36** | **1.52** | 0.69 |
| **5-(3',4'-Dihydroxyphenyl)-γ-valerolactone** | **1.01** | **1.05** | 0.59 | **1.36** | **1.22** |
| **5-(4'-Hydroxyphenyl)-γ-valerolactone-3'-glucuronide** | **1.03** | **1.00** | 0.62 | **1.02** | **1.03** |
| **5-(3'-Hydroxyphenyl)-γ-valerolactone-4'-glucuronide** | **1.04** | 0.94 | 0.68 | 0.94 | **1.11** |
| **5-(Hydroxyphenyl)-γ-valerolactone-sulfate (3',4' isomers)** | **1.34** | **1.37** | 0.85 | 0.74 | **1.15** |
| **5-Phenyl-γ-valerolactone-sulfate-glucuronide isomer (3',4')** | **1.11** | **1.09** | 0.70 | 0.29 | 0.96 |
| **5-Phenyl-γ-valerolactone-methoxy-sulfate isomer (3',4')** | **1.08** | **1.09** | 0.83 | 0.22 | 0.91 |
| **3-Phenylpropanoic acid sulfate** | **1.32** | **1.33** | 0.92 | 0.50 | **1.09** |
| **3-Phenylpropanoic acid glucuronide** | 0.76 | 0.77 | **2.09** | **1.63** | **1.03** |
| **3'OH-PVLs** | 0.88 | **1.26** | 0.70 | 0.35 | 0.7 |
| **4'OH-PVLs** | **1.42** | 0.65 | **1.34** | **1.56** | **1.25** |
| **3',4'diOH-PVLs** | 0.64 | 0.91 | 0.63 | 0.79 | **1.04** |
| **HPPs** | 0.90 | **1.08** | **1.14** | 0.90 | 0.87 |

3'OH-PVLs, sum of conjugates from the aglycone 5-(3'-hydroxyphenyl)-γ-valerolactone; 4'OH-PVLs, 5-(4'-hydroxyphenyl)-γ-valerolactone; 3',4'diOH-PVLs, 5-(3',4'-dihydroxyphenyl)-γ-valerolactone; HPPs, 3-(hydroxyphenyl)propanoic acid.

## References

1 R. Zamora-Ros, N. G. Forouhi, S. J. Sharp, C. A. González, B. Buijsse, M. Guevara, Y. T. Van Der Schouw, P. Amiano, H. Boeing, L. Bredsdorff, F. Clavel-Chapelon, G. Fagherazzi, E. J. Feskens, P. W. Franks, S. Grioni, V. Katzke, T. J. Key, K. T. Khaw, T. Kühn, G. Masala, A. Mattiello, E. Molina-Montes, P. M. Nilsson, K. Overvad, F. Perquier, J. Ramón Quirós, I. Romieu, C. Sacerdote, A. Scalbert, M. Schulze, N. Slimani, A. M. W. Spijkerman, A. Tjonneland, M. J. Tormo, R. Tumino, D. L. Van Der A., C. Langenberg, E. Riboli and N. J. Wareham, The association between dietary flavonoid and lignan intakes and incident type 2 diabetes in european populations: The EPIC-InterAct study, *Diabetes Care*, 2013, **36**, 3961–3970.

2 I. Zanotti, M. Dall'Asta, P. Mena, L. Mele, R. Bruni, S. Ray and D. Del Rio, Atheroprotective effects of (poly) phenols: A focus on cell cholesterol metabolism, *Food Funct.*, 2015, **6**, 13–31.

3 F. Ferrucci, G. Gaudiosi, A. Hirn and R. Nicolich, Ionian Basin and Calabria Arc: Some new elements from DSS data, *Tectonophysics*, 1991, **195**, 411–419.

4 R. Zamora-Ros, V. Knaze, J. A. Rothwell, B. Hémon, A. Moskal, K. Overvad, A. Tjønneland, C. Kyrø, G. Fagherazzi, M.-C. Boutron-Ruault, M. Touillaud, V. Katzke, T. Kühn, H. Boeing, J. Förster, A. Trichopoulou, E. Valanou, E. Peppa, D. Palli, C. Agnoli, F. Ricceri, R. Tumino, M. S. de Magistris, P. H. M. Peeters, H. B. Bueno-de-Mesquita, D. Engeset, G. Skeie, A. Hjartåker, V. Menéndez, A. Agudo, E. Molina-Montes, J. M. Huerta, A. Barricarte, P. Amiano, E. Sonestedt, L. M. Nilsson, R. Landberg, T. J. Key, K.-T. Khaw, N. J. Wareham, Y. Lu, N. Slimani, I. Romieu, E. Riboli and A. Scalbert, Dietary polyphenol intake in Europe: the European Prospective Investigation into Cancer and Nutrition (EPIC) study, *Eur. J. Nutr.*, 2016, **55**, 1359–1375.

5 N. Ziauddeen, A. Rosi, D. Del Rio, B. Amoutzopoulos, S. Nicholson, P. Page, F. Scazzina, F. Brighenti, S. Ray and P. Mena, Dietary intake of (poly)phenols in children and adults: cross-sectional analysis of UK National Diet and Nutrition Survey Rolling Programme (2008–2014), *Eur. J. Nutr.*, 2019, **58**, 3183–3198.

6 A. Rodriguez-Mateos, R. P. Feliciano, A. Boeres, T. Weber, C. N. dos Santos, M. R. Ventura and C. Heiss, Cranberry (poly)phenol metabolites correlate with improvements in vascular

function: A double-blind, randomized, controlled, dose-response, crossover study, *Mol. Nutr. Food Res.*, 2016, **60**, 2130–2140.

7 R. Sansone, A. Rodriguez-Mateos, J. Heuel, D. Falk, D. Schuler, R. Wagstaff, G. G. C. Kuhnle, J. P. E. Spencer, H. Schroeter, M. W. Merx, M. Kelm and C. Heiss, Cocoa flavanol intake improves endothelial function and Framingham Risk Score in healthy men and women: a randomised, controlled, double-masked trial: the Flaviola Health Study, *Br. J. Nutr.*, 2015, **114**, 1246–1255.

8 A. Rodriguez-Mateos, D. Vauzour, C. G. Krueger, D. Shanmuganayagam, J. Reed, L. Calani, P. Mena, D. Del Rio and A. Crozier, Bioavailability, bioactivity and impact on health of dietary flavonoids and related compounds: an update, *Arch. Toxicol.*, 2014, **88**, 1803–1853.

9 A. González-Sarrías, E. Combet, P. Pinto, P. Mena, M. Dall'Asta, M. Garcia-Aloy, A. Rodríguez-Mateos, E. R. Gibney, J. Dumont, M. Massaro, J. Sánchez-Meca, C. Morand and M. T. García-Conesa, A systematic review and meta-analysis of the effects of flavanol-containing tea, cocoa and apple products on body composition and blood lipids: exploring the factors responsible for variability in their efficacy, *Nutrients*, 2017, **9**, 746.

10 M. C. Morris, C. C. Tangney, Y. Wang, F. M. Sacks, L. L. Barnes, D. A. Bennett and N. T. Aggarwal, MIND diet slows cognitive decline with aging, *Alzheimer's Dement.*, 2015, **11**, 1015–1022.

11 K. Ide, N. Matsuoka, H. Yamada, D. Furushima, K. Kawakami, K. Ide, N. Matsuoka, H. Yamada, D. Furushima and K. Kawakami, Effects of Tea Catechins on Alzheimer's Disease: Recent Updates and Perspectives, *Molecules*, 2018, **23**, 2357.

12 D. González de Llano, A. Esteban-Fernández, F. Sánchez-Patán, P. Martínlvarez, M. Moreno-Arribas and B. Bartolomé, Anti-Adhesive Activity of Cranberry Phenolic Compounds and Their Microbial-Derived Metabolites against Uropathogenic Escherichia coli in Bladder Epithelial Cell Cultures, *Int. J. Mol. Sci.*, 2015, **16**, 12119–12130.

13 P. Mena, D. González de Llano, N. Brindani, A. Esteban-Fernández, C. Curti, M. V. Moreno-Arribas, D. Del Rio and B. Bartolomé, 5-(3',4'-Dihydroxyphenyl)-γ-valerolactone and its sulphate conjugates, representative circulating metabolites of flavan-3-ols, exhibit anti-adhesive activity against uropathogenic Escherichia coli in bladder epithelial cells, *J. Funct. Foods*, 2017, **29**, 275–280.

14 D. Del Rio, A. Rodriguez-Mateos, J. P. E. Spencer, M. Tognolini, G. Borges and A. Crozier, Dietary (Poly)phenolics in Human Health: Structures, Bioavailability, and Evidence of Protective Effects Against Chronic Diseases, *Antioxid. Redox Signal.*, 2013, **18**, 1818–1892.

15 L. Bresciani, D. Martini, P. Mena, M. Tassotti, L. Calani, G. Brigati, F. Brighenti, S. Holasek, D.-E. Malliga, M. Lamprecht and D. Del Rio, Absorption Profile of (Poly)Phenolic Compounds after Consumption of Three Food Supplements Containing 36 Different Fruits, Vegetables, and Berries, *Nutrients*, 2017, **9**, 194.

16 M. M. Appeldoorn, J.-P. Vincken, A.-M. Aura, P. C. H. Hollman and H. Gruppen, Procyanidin Dimers Are Metabolized by Human Microbiota with 2-(3,4-Dihydroxyphenyl)acetic Acid and 5-(3,4-Dihydroxyphenyl)-γ-valerolactone as the Major Metabolites, *J. Agric. Food Chem.*, 2009, **57**, 1084–1092.

17 S. Stoupi, G. Williamson, J. W. Drynan, D. Barron and M. N. Clifford, A comparison of the *in vitro* biotransformation of (-)-epicatechin and procyanidin B2 by human faecal microbiota, *Mol. Nutr. Food Res.*, 2010, **54**, 747–759.

18 M. Kutschera, W. Engst, M. Blaut and A. Braune, Isolation of catechin-converting human intestinal bacteria, *J. Appl. Microbiol.*, 2011, **111**, 165–175.

19 P. Mena, L. Bresciani, N. Brindani, I. A. Ludwig, G. Pereira-Caro, D. Angelino, R. Llorach, L. Calani, F. Brighenti, M. N. Clifford, C. I. R. Gill, A. Crozier, C. Curti and D. Del Rio, Phenyl-γ-valerolactones and phenylvaleric acids, the main colonic metabolites of flavan-3-ols: synthesis, analysis, bioavailability, and bioactivity, *Nat. Prod. Rep.*, 2019, **36**, 714–752.

20 J. I. Ottaviani, G. Borges, T. Y. Momma, J. P. E. Spencer, C. L. Keen, A. Crozier and H. Schroeter, The metabolome of [2-14C](-)-epicatechin in humans: Implications for the assessment of efficacy, safety, and mechanisms of action of polyphenolic bioactives, *Sci. Rep.*, 2016, **6**, 1–10.

21 G. Borges, J. I. Ottaviani, J. J. J. van der Hooft, H. Schroeter and A. Crozier, Absorption, metabolism, distribution and excretion of (−)-epicatechin: A review of recent findings, *Mol. Aspects Med.*, 2018, **61**, 18–30.

22 F. Castello, G. Costabile, L. Bresciani, M. Tassotti, D. Naviglio, D. Luongo, P. Ciciola, M. Vitale, C. Vetrani, G. Galaverna, F. Brighenti, R. Giacco, D. Del Rio and P. Mena, Bioavailability and

pharmacokinetic profile of grape pomace phenolic compounds in humans, *Arch. Biochem. Biophys.*, 2018, **646**, 1–9.

23 C. Favari, P. Mena, C. Curti, G. Istas, C. Heiss, D. Del Rio and A. Rodriguez-Mateos, Kinetic profile and urinary excretion of phenyl-γ-valerolactones upon consumption of cranberry: a dose–response relationship, *Food Funct.*, 2020, **11**, 3975–3985.

24 J. I. Ottaviani, R. Fong, J. Kimball, J. L. Ensunsa, A. Britten, D. Lucarelli, R. Luben, P. B. Grace, D. H. Mawson, A. Tym, A. Wierzbicki, K. T. Khaw, H. Schroeter and G. G. C. Kuhnle, Evaluation at scale of microbiome-derived metabolites as biomarker of flavan-3-ol intake in epidemiological studies, *Sci. Rep.*, 2018, **8**, 1–11.

25 A. Cortés-Martín, M. V. Selma, J. C. Espín and R. García-Villalba, The Human Metabolism of Nuts Proanthocyanidins does not Reveal Urinary Metabolites Consistent with Distinctive Gut Microbiota Metabotypes, *Mol. Nutr. Food Res.*, 2019, **63**, 1800819.

26 P. Mena, I. A. Ludwig, V. B. Tomatis, A. Acharjee, L. Calani, A. Rosi, F. Brighenti, S. Ray, J. L. Griffin, L. J. Bluck and D. Del Rio, Inter-individual variability in the production of flavan-3-ol colonic metabolites: preliminary elucidation of urinary metabotypes, *Eur. J. Nutr.*, 2019, **58**, 1529–1543.

27 E. J. J. van Velzen, J. A. Westerhuis, C. H. Grün, D. M. Jacobs, P. H. C. Eilers, T. P. Mulder, M. Foltz, U. Garczarek, R. Kemperman, E. E. Vaughan, J. P. M. van Duynhoven and A. K. Smilde, Population-based nutrikinetic modeling of polyphenol exposure, *Metabolomics*, 2014, **10**, 1059–1073.

28 J. van Duynhoven, J. J. J. van der Hooft, F. A. van Dorsten, S. Peters, M. Foltz, V. Gomez-Roldan, J. Vervoort, R. C. H. de Vos and D. M. Jacobs, Rapid and Sustained Systemic Circulation of Conjugated Gut Microbial Catabolites after Single-Dose Black Tea Extract Consumption, *J. Proteome Res.*, 2014, **13**, 2668–2678.

29 F. Sánchez-Patán, C. Cueva, M. Monagas, G. E. Walton, G. R. Gibson, P. J. Martín-Álvarez, M. Victoria Moreno-Arribas and B. Bartolomé, Gut microbial catabolism of grape seed flavan-3-ols by human faecal microbiota. Targetted analysis of precursor compounds, intermediate metabolites and end-products, *Food Chem.*, 2012, **131**, 337–347.

30 F. A. Tomás-Barberán, R. García-Villalba, A. González-Sarrías, M. V. Selma and J. C. Espín, Ellagic Acid Metabolism by Human Gut Microbiota: Consistent Observation of Three Urolithin Phenotypes in Intervention Trials, Independent of Food Source, Age, and Health Status, *J. Agric. Food Chem.*, 2014, **62**, 6535–6538.

31 T. Larkin, W. E. Price and L. Astheimer, The Key Importance of Soy Isoflavone Bioavailability to Understanding Health Benefits, *Crit. Rev. Food Sci. Nutr.*, 2008, **48**, 538–552.

32 I. R. Rowland, H. Wiseman, T. A. B. Sanders, H. Adlercreutz and E. A. Bowey, Interindividual Variation in Metabolism of Soy Isoflavones and Lignans: Influence of Habitual Diet on Equol Production by the Gut Microflora, *Nutr. Cancer*, 2000, **36**, 27–32.

33 M. V. Selma, A. González-Sarrías, J. Salas-Salvadó, C. Andrés-Lacueva, C. Alasalvar, A. Örem, F. A. Tomás-Barberán and J. C. Espín, The gut microbiota metabolism of pomegranate or walnut ellagitannins yields two urolithin-metabotypes that correlate with cardiometabolic risk biomarkers: Comparison between normoweight, overweight-obesity and metabolic syndrome, *Clin. Nutr.*, 2018, **37**, 897–905.

34 A. González-Sarrías, R. García-Villalba, M. Romo-Vaquero, C. Alasalvar, A. Örem, P. Zafrilla, F. A. Tomás-Barberán, M. V. Selma and J. C. Espín, Clustering according to urolithin metabotype explains the interindividual variability in the improvement of cardiovascular risk biomarkers in overweight-obese individuals consuming pomegranate: A randomized clinical trial, *Mol. Nutr. Food Res.*, 2017, **61**, 1–14.

35 S. Hazim, P. J. Curtis, M. Y. Schär, L. M. Ostertag, C. D. Kay, A.-M. Minihane and A. Cassidy, Acute benefits of the microbial-derived isoflavone metabolite equol on arterial stiffness in men prospectively recruited according to equol producer phenotype: a double-blind randomized controlled trial, *Am. J. Clin. Nutr.*, 2016, **103**, 694–702.

36 C. Liu, J. Vervoort, K. Beekmann, M. Baccaro, L. Kamelia, S. Wesseling and I. M. C. M. Rietjens, Interindividual Differences in Human Intestinal Microbial Conversion of (−)-Epicatechin to Bioactive Phenolic Compounds, *J. Agric. Food Chem.*, 2020, **68**, 14168–14181.

37 E. Hillesheim and L. Brennan, Metabotyping and its role in nutrition research, *Nutr. Res. Rev.*, 2020, **33**, 33–42.

38 R. Feliciano, C. Mills, G. Istas, C. Heiss and A. Rodriguez-Mateos, Absorption, Metabolism and Excretion of Cranberry (Poly)phenols in Humans: A Dose Response Study and Assessment of Inter-Individual Variability, *Nutrients*, 2017, **9**, 268.

39 N. Brindani, P. Mena, L. Calani, I. Benzie, S. W. Choi, F. Brighenti, F. Zanardi, C. Curti and D. Del Rio, Synthetic and analytical strategies for the quantification of phenyl-γ-valerolactone conjugated metabolites in human urine, *Mol. Nutr. Food Res.*, 2017, **61**, 6–10.

40 C. D. Kay, M. N. Clifford, P. Mena, G. J. McDougall, C. Andres-Lacueva, A. Cassidy, D. Del Rio, N. Kuhnert, C. Manach, G. Pereira-Caro, A. Rodriguez-Mateos, A. Scalbert, F. Tomás-Barberán, G. Williamson, D. S. Wishart and A. Crozier, Recommendations for standardizing nomenclature for dietary (poly)phenol catabolites, *Am. J. Clin. Nutr.*, 2020, **112**, 1051–1068.

41 F. Tugizimana, P. Steenkamp, L. Piater and I. Dubery, A Conversation on Data Mining Strategies in LC-MS Untargeted Metabolomics: Pre-Processing and Pre-Treatment Steps, *Metabolites*, 2016, **6**, 40.

42 R. A. van den Berg, H. C. J. Hoefsloot, J. A. Westerhuis, A. K. Smilde and M. J. van der Werf, Centering, scaling, and transformations: improving the biological information content of metabolomics data, *BMC Genomics*, 2006, **7**, 142.

43 R. C. Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria., https://www.r-project.org/.

44 S. Lê, J. Josse and F. Husson, FactoMineR: An R package for multivariate analysis, *J. Stat. Softw.*, 2008, **25**, 1–18.

45 E. Schubert and P. J. Rousseeuw, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer, 2019, vol. 11807 LNCS, pp. 171–187.

46 J. A. Hartigan and M. A. Wong, Algorithm AS 136: A K-Means Clustering Algorithm, *J. R. Stat. Soc. Ser. C (Applied Stat.*, 1979, **28**, 100–108.

47 N. R. Pal, J. C. Bezdek and R. J. Hathaway, Sequential competitive learning and the fuzzy c-means clustering algorithms, *Neural Networks*, 1996, **9**, 787–796.

48 G. N. Lance, A general theory of classificatory sorting strategies: II. Clustering systems, *Comput. J.*, 1967, **10**, 271–277.

49 C. Fraley, A. E. Raftery and T. B. Murphy, *mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation*, 2012.

50 A. Acharjee, R. Finkers, R. G. Visser and C. Maliepaard, Comparison of Regularized Regression Methods for ~Omics Data, *J. Postgenomics Drug Biomark. Dev.*, 2012, **03**, 126.

51 A. Anesi, P. Mena, A. Bub, M. Ulaszewska, D. Del Rio, S. E. Kulling and F. Mattivi, Quantification of Urinary Phenyl-γ-Valerolactones and Related Valeric Acids in Human Urine on Consumption of Apples, *Metabolites*, 2019, **9**, 254.

52 A. Riedl, C. Gieger, H. Hauner, H. Daniel and J. Linseisen, Metabotyping and its application in targeted nutrition: an overview, *Br. J. Nutr.*, 2017, **117**, 1631–1644.

53 A. Esteban-Fernández, C. Ibañez, C. Simó, B. Bartolomé and V. Moreno-Arribas, Metabolome-based clustering after moderate wine consumption, *Oeno One*, 2020, **54**, 455–467.

54 M. Inoue-Choi, J.-M. Yuan, C. S. Yang, D. J. Van Den Berg, M.-J. Lee, Y.-T. Gao and M. C. Yu, Genetic Association Between the COMT Genotype and Urinary Levels of Tea Polyphenols and Their Metabolites among Daily Green Tea Drinkers., *Int. J. Mol. Epidemiol. Genet.*, 2010, **1**, 114–123.

55 J. O. Miners, R. A. McKinnon and P. I. Mackenzie, Genetic polymorphisms of UDP-glucuronosyltransferases and their functional significance, *Toxicology*, 2002, **181**–**182**, 453–456.

56 S. Haldar, N. T. Hernandez, L. Ostertag, P. Curtis, A. Cassidy and A. M. Minihane, Genetic and phenotypic determinants of flavonoid absorption and metabolism: the COB study, *Arch. Public Heal.*, 2014, **72**, O3.

57 H. Koster, I. Halsema, E. Scholtens, M. Knippers and G. J. Mulder, Dose-dependent shifts in the sulfation and glucuronidation of phenolic compounds in the rat *in vivo* and in isolated hepatocytes, *Biochem. Pharmacol.*, 1981, **30**, 2569–2575.

58 F. Castello, G. Costabile, L. Bresciani, M. Tassotti, D. Naviglio, D. Luongo, P. Ciciola, M. Vitale, C. Vetrani, G. Galaverna, F. Brighenti, R. Giacco, D. Del Rio and P. Mena, Bioavailability and pharmacokinetic profile of grape pomace phenolic compounds in humans, *Arch. Biochem. Biophys.*, 2018, **646**, 1–9.

59 J. I. Ottaviani, R. Y. Fong, G. Borges, H. Schroeter and A. Crozier, Use of LC-MS for the quantitative analysis of (poly)phenol metabolites does not necessarily yield accurate results: Implications for assessing existing data and conducting future research, *Free Radic. Biol. Med.*, 2018, **124**, 97–103.

60 A. Vogiatzoglou, A. A. Mulligan, R. N. Luben, M. A. H. Lentjes, C. Heiss, M. Kelm, M. W. Merx, J. P. E. Spencer, H. Schroeter and G. G. C. Kuhnle, Assessment of the dietary intake of total flavan-3-ols, monomeric flavan-3-ols, proanthocyanidins and theaflavins in the European Union, *Br. J. Nutr.*, 2014, **111**, 1463–1473.

61 A. Rosi, P. Mena, F. Castello, D. Del Rio and F. Scazzina, Comprehensive dietary evaluation of Italian primary school children: food consumption and intake of energy, nutrients and phenolic compounds, *Int. J. Food Sci. Nutr.*, 2021, **72**, 70–81.

62 C. Manach, D. Milenkovic, T. Van de Wiele, A. Rodriguez-Mateos, B. de Roos, M. T. Garcia-Conesa, R. Landberg, E. R. Gibney, M. Heinonen, F. Tomás-Barberán and C. Morand, Addressing the inter-individual variation in response to consumption of plant food bioactives: Towards a better understanding of their role in healthy aging and cardiometabolic risk reduction, *Mol. Nutr. Food Res.*, 2017, **61**, 1600557.

63 A. Takagaki and F. Nanjo, Bioconversion of (−)-Epicatechin, (+)-Epicatechin, (−)-Catechin, and (+)-Catechin by (−)-Epigallocatechin-Metabolizing Bacteria, *Biol. Pharm. Bull.*, 2015, **38**, 789–794.

64 F. Sánchez-Patán, R. Tabasco, M. Monagas, T. Requena, C. Peláez, M. V. Moreno-Arribas and B. Bartolomé, Capability of lactobacillus plantarum IFPL935 to catabolize flavan-3-ol compounds and complex phenolic extracts, *J. Agric. Food Chem.*, 2012, **60**, 7142–7151.

65 D. González de Llano, M. V. Moreno-Arribas and B. Bartolomé, Cranberry Polyphenols and Prevention against Urinary Tract Infections: Relevant Considerations, *Molecules*, 2020, **25**, 3523.

66 O. A. Gbinigie, E. A. Spencer, C. J. Heneghan, J. J. Lee and C. C. Butler, Cranberry Extract for Symptoms of Acute, Uncomplicated Urinary Tract Infection: A Systematic Review, *Antibiotics*, 2020, **10**, 12.

## 3.3. Study 3

# Metabolomic changes after coffee consumption: new paths on the block

**Claudia Favari**[1‡], Laura Righetti[2‡*], Michele Tassotti[1], Lee A. Gethings[3], Daniela Martini[1,4], Alice Rosi[1], Monica Antonini[5], Josep Rubert[6], Claudine Manach[7], Alessandra Dei Cas[5], Riccardo Bonadonna[5], Furio Brighenti[1], Chiara Dall'Asta[2], Pedro Mena[1*], Daniele Del Rio[8,9]

[1] Human Nutrition Unit, Department of Food and Drugs, University of Parma, Medical School Building C, Via Volturno, 39, 43125 Parma, Italy

[2] Department of Food and Drugs, University of Parma, Viale delle Scienze 17/A, 43124 Parma, Italy

[3] Waters Corporation, Altrincham Road, Wilmslow, SK9 4AX, United Kingdom

[4] Department of Food, Environmental and Nutritional Sciences (DeFENS), Università degli Studi di Milano, 20122 Milan, Italy

[5] Division of Endocrinology and Metabolic Diseases, Department of Medicine and Surgery, University of Parma, Via Gramsci 14, 43126, Parma, Italy

[6] Interdisciplinary research structure of biotechnology and biomedicine, Department of Biochemistry and Molecular Biology, Universitat de Valencia, 46100 Burjassot (València), Spain.

[7] Université Clermont Auvergne, INRAE, UNH, F-63000 Clermont–Ferrand, France

[8] Human Nutrition Unit, Department of Veterinary Science, University of Parma, Parma, Italy

[9] School of Advanced Studies on Food and Nutrition, University of Parma, Parma, Italy

[‡]authors equally contributed to the study

*Corresponding authors details:

Pedro Mena, Via Volturno 39, 43125 Parma, Italy; Tel.: +39 0521-903841; pedro.mena@unipr.it.

Laura Righetti, Viale delle Scienze 17/A, 43124 Parma, Italy; Tel.: +39 0521-906196; laura.righetti@unipr.it.

The content of this section has also been published in Molecular Nutrition & Food Research (Favari C, Righetti L, Tassotti M, et al (2021) Metabolomic Changes after Coffee Consumption: New Paths on the Block. Mol Nutr Food Res 65:2000875. https://doi.org/10.1002/mnfr.202000875).

## Abstract

**Scope:** Several studies suggest that regular coffee consumption may help preventing chronic diseases, but the impact of daily intake and the contribution of coffee metabolites in disease prevention are still unclear. The present study aimed at evaluating whether and how different patterns of coffee intake (one cup of espresso coffee/day, three cups of espresso coffee/day, one cup of espresso coffee/day and two cocoa-based products containing coffee two times per day) might impact endogenous molecular pathways.

**Methods and results:** A three-arm, randomized, cross-over trial was performed in 21 healthy volunteers who consumed each treatment for one month. Urine samples were collected to perform untargeted metabolomics based on UHPLC-IMS-HRMS. A total of 153 discriminant metabolites were identified. Several molecular features were associated with coffee consumption, while others were linked with different metabolic pathways, such as phenylalanine, tyrosine, energy metabolism, steroid hormone biosynthesis and arginine biosynthesis and metabolism.

**Conclusion:** This information has provided new insights into the metabolic routes by which coffee and coffee-related metabolites may exert effects on human health.

**Keywords**

Biomarker, cocoa, coffee, metabolomics, xenobiotics.

## 3.3.1. Introduction

Coffee is one of the most appreciated and consumed beverages worldwide. Besides the pleasant aroma and taste, it is considered an important source of bioactive compounds, mainly caffeine, trigonelline, chlorogenic acids, cafestol, and kahweol [1]. In many epidemiological studies, regular coffee consumption has been associated with a reduced risk of several chronic diseases, such as type 2 diabetes, atherosclerotic heart disease, and stroke, as well as of neurodegenerative conditions, like Parkinson's and Alzheimer's diseases [2–8]. Most meta-analyses have shown an apparent dose-response effect, with the lowest disease risk achieved with the consumption of about 3–5 cups/day [9,10]. However, the dose – a cup of coffee – is not a standardized measurement, and compound content of a dose also varies with the brewing method [11]. Of note, no association has already been found between circulating coffee-related metabolites and physiological responses, making the mechanisms through which coffee exerts its potential preventive effects still widely undisclosed.

Among the plant matrixes with high content in bioactive phytochemicals, cocoa is also gaining increasing attention in nutrition research [12–14]. Cocoa and its derived products mainly contain flavan-3-ols and theobromine, a closely related analogue of caffeine [13]. Cocoa products may enhance the preventive effects of regular coffee consumption, and, in turn, cocoa-based products containing coffee, combining the phytochemical content of both coffee and cocoa, may be regarded as a potential candidate to increment the levels of putatively protective metabolites in the context of a balanced diet [15]. Since for some individuals the consumption of more than 1–2 coffees daily is unlikely, the substitution of coffee with cocoa-based products containing coffee may represent an alternative strategy to supply both cocoa and coffee bioactives and to ensure their putative bioactivity.

It is worth noting that, except for trigonelline, coffee and cocoa-related phytochemicals are extensively transformed by human metabolism, and the gut microbial catabolism. These derived compounds, rather than the parent molecules, are circulating molecules that might exert a beneficial action in human health [1]. To date, the complete pool of circulating metabolites resulting from coffee and cocoa consumption still needs to be disclosed. In this frame, metabolomics allows a comprehensive description of the metabolites in a biological sample, providing information on exposure to exogenous metabolites and on levels of endogenous metabolites from metabolic pathways, thus allowing the study of biochemical processes modulation [16]. In most recent years, the number of metabolomic studies applied to answer nutritional questions has increased. In particular,

metabolomic profiling has been widely used to map biomarkers of intake, which are metabolites generated from compounds present in a specific food [17–21]. On the other hand, untargeted metabolomics approaches aim at identifying not only biomarkers specifically associated with a given food like coffee, but also metabolites that may reflect the biological effects of specific dietary components. This strategy may help in elucidating the contribution of coffee metabolites in disease prevention and in shedding light on the underlying mechanisms [22,23]. For instance, a comprehensive metabolomic analysis of serum samples following coffee intake (up to 8 cups of coffee/day) revealed that metabolites from the endocannabinoid and fatty acid acylcholine pathway decreased in response to coffee consumption whilst those of the steroid pathway generally increased [24]. Moreover, induction of fatty acid metabolism, mainly related to carnitine derivatives, was observed in urine after 30 days following green coffee bean extract consumption [25]. Shi and colleagues have also identified several plasma metabolites specifically associated with filtered and boiled coffee consumption and used them to estimate filtered or boiled coffee intake and to find associations with type 2 diabetes risk [26].

Based on those previous studies, there is still room for discovering novel pathways of coffee metabolic effects using metabolomics. This untargeted metabolomics study has revealed new molecular pathways affected by coffee and cocoa intake, linking metabolic pathways to different levels of coffee and cocoa intake.

## 3.3.2. Materials and methods

### *Chemicals*

HPLC-grade methanol, acetonitrile, and acetic acid were purchased from Sigma-Aldrich (Taufkirchen, Germany); bidistilled water was obtained using Milli-Q System (Millipore, Bedford, MA, USA). MS-grade formic acid from Fisher Chemical (Thermo Fisher Scientific Inc., San Jose, CA, USA) and ammonium acetate (Fluka, Chemika-Biochemika, Basil, Switzerland) were also used. Leucine-enkephalin, used as lock mass standard and Major Mix for collisional cross-sectional (CCS) calibration were purchased from Waters (Milford, USA.).

### *Subjects*

Twenty-one volunteers were recruited in Parma (Northern Italy) to participate in the study. Inclusion and exclusion criteria and main subject clinical characteristics have already been published [27]. Briefly, 21 subjects, 10 males (2 smokers) and 11 females (6 smokers), aged 25.9 ± 0.5, BMI 22.3 ± 0.6 kg/m$^2$, were enrolled. The study was conducted according to the guidelines of Good Clinical

Practice and the Declaration of Helsinki. All subjects provided written informed consent before study entry, and they all completed the intervention study.

*Dosage information and study design*

A three-arm, randomized cross-over trial was performed in 21 healthy volunteers, as previously reported [15]. The study was approved by the Ethics Committee for Parma Hospital and University (AZOSPR/0015693/6.2.2.) and registered on ClinicalTrials.gov on May 21, 2017 (NCT03166540). Briefly, participants had to consume three different treatments in a random order for one month: (1) one cup of espresso coffee/day (at 9.00 AM, namely 1C group), (2) three cups of espresso coffee/day (at 9.00 AM, 12.00 noon, and 3.00 PM, namely 3C group) and (3) one cup of espresso coffee/day and two cocoa-based products containing coffee (CBPCC) twice per day (coffee at 9.00 AM and two CBPCC at 12.00 PM and 3.00 PM, namely PC group). The randomization list was generated using Random Number Generator Pro (Segobit Software). Volunteers were supplied with a single-serve coffee machine (Essenza EN 97.W, De' Longhi Appliances S.r.l., Treviso, Italy) and coffee capsules (Capriccio, Nespresso Italia S.p.a., Assago, Italy) to standardize raw material, brewing method and cup volume, and also with the CBPCC (Pocket Coffee, Ferrero Commerciale Italia S.r.l., Alba, Italy). The number of CBPCC was chosen considering that the amount of caffeine provided by a cup of espresso coffee was ≈60 mg per serving, while two CBPCC provided, together, ≈30 mg of caffeine and thus were approximately the same as half coffee cup in terms of caffeine content. Minimal dietary restrictions were given to volunteers two days before and on each sampling day to exclude other sources of coffee/cocoa-related phytochemicals apart from those provided by the assigned treatment. The sampling day corresponded to the last day of each intervention period. On the sampling day, urine from each volunteer was collected at baseline (t0) and different collection periods within 0–3 h, 3–6 h, 6–9 h, and 9–24 h. Samples used for this study corresponded to the period 9-24 h. The volume of urine collected during each period was measured and two 2 mL samples were stored at −80 °C until analysis.

*Urine sample preparation*

Urine samples were prepared as described elsewhere [28]. Briefly, urine samples were thawed on ice before analysis and centrifuged for 10 min at 10,000g to remove particulates. 50 µL of supernatant were diluted with 100 µL of Milli-Q water. Quality control samples consisting of all urine samples to form a pool were analyzed for the study and injected every 9 samples to allow for the performance of the analytical system in terms of retention times, mass accuracy and signal intensities to be

evaluated. Three technical replicates of each sample were injected. All samples were acquired in a randomized order.

*UHPLC-TWIMS-QTOF analysis*

ACQUITY I-Class UPLC separation system coupled to a VION IMS QTOF mass spectrometer (Waters, Wilmslow, UK) equipped with electrospray ionization (ESI) interface was employed.

Samples were injected (5 µL) and chromatographically separated using a reversed-phase C18 HSS T3 ACQUITY column 2.1 × 100 mm, 1.7 µm particle size (Waters, Milford, MA, USA). A gradient profile, as previously described was applied [28]. In short, water (eluent A) and acetonitrile (eluent B), both acidified with 0.1% formic acid, were used as mobile phases. Initial conditions were set at 1% B followed by a linear change to 15% B in 3 min and 50% B in 3 min. Finally, 95% B was achieved at 9 min prior to holding at 95% for 1 min to allow for column washing before returning to initial conditions. Column recondition was completed over 3 min, providing a total run time of 14 min. The column was maintained at 40 °C and a flow rate of 0.5 mL/min used.

Mass spectrometry data were collected in both positive and negative electrospray mode over the mass range of *m/z* 70−1000. Source settings were maintained using a capillary voltage, 2.5 kV; cone voltage, 40 V; source temperature, 120 °C; desolvation temperature, 500 °C and desolvation gas flow, 800 L/h. The TOF analyzer was operated in "sensitivity mode" and data acquired using HDMSE [29], which is a data-independent approach (DIA) coupled with ion mobility. The ion mobility device within the Vion was calibrated using the Major Mix IMS calibration kit (Waters, Wilmslow, UK) to allow for CCS values to be determined in nitrogen. The calibration covered the CCS range from 130-306 $Å^2$. The TOF was also calibrated prior to data acquisition using sodium formate (Waters, Wilmslow, UK) and covered the mass range from 151 Da to 1013 Da. TOF and CCS calibrations were performed for both positive and negative ion mode. Data acquisition was conducted using UNIFI 1.8 (Waters, Wilmslow, UK).

*Data processing and multivariate modelling*

Data processing and compound identification were conducted using Progenesis QI Informatics (Nonlinear Dynamics, Newcastle, UK). Each UPLC-MS run was imported as an ion-intensity map, including *m/z* (*m/z* range 70-1000) and retention time, that were then aligned in the retention-time direction (0-8.5 min). From the aligned runs, an aggregate run representing the compounds in all samples was used for peak picking. This aggregate was then compared with all runs, so that the same ions are detected in every run. Isotope and adduct deconvolution were applied, to reduce the

number of features detected. Data were normalized according to creatinine intensity in each sample.

Unsupervised principal components analysis (PCA) with pareto scaling was performed to check the quality of the raw data. Afterward, the variables were filtered, retaining entities with coefficients of variation (CV) lower than 30% across the QCs. From the analysis of the variance (ANOVA) significant features were selected, retaining those presenting, simultaneously, fold change > 2, and Benjamini-Hochberg FDR adjusted *p*-value (q value) < 0.01. In parallel, multivariate supervised models, including least-squares discriminant analysis (PLS-DA) were built and validated using SIMCA software (v. 16.0.2, Sartorius Stedim Data Analytics, Sweden). Cross-validation of the PLS-DA model using one-third leaving out approach and permutation testing were applied to validate and to exclude overfitting by inspecting model parameters (goodness-of-fit $R^2Y$ and goodness-of-prediction $Q^2Y$). The variable influence in projection analysis (VIP) was further used to identify the compounds that have the highest discrimination potential (VIP value threshold > 1.2). The resulting significant features to both ANOVA *p*-values < 0.01 and VIP > 1.2 were subjected to the identification.

Metabolites were identified by publicly available database searches including Lipid Metabolites and Pathways Strategy (LIPID MAPS) [30], Human Metabolome database (HMDB) [31], and METLIN [32], as well as by fragmentation patterns, retention times and collision cross-sections. CCS values were searched against "MetCCS Predictor" database containing *m/z* and CCS values by selecting a ΔCCS of 5% for metabolite matching [33]. Based on the Metabolomics Standards Initiative [34], metabolites were annotated as level III (putatively characterized), level II (putatively identified compounds) and level I (identified compound), as reported in Table 9. Level I identification was performed by comparison of rt and fragmentation pattern with the standard collect in our UNIFI library, created by running a mix of standards with the same analytical method.

*Metabolic pathway analysis*

Identified metabolites were submitted to the Pathway and Network Analysis modules in MetaboAnalyst 4.0 using HMDB identifiers [35]. For the former analysis, Fishers' exact test and relative-betweenness centrality were the algorithms respectively selected to perform pathway enrichment analysis and pathway topology analysis, using the current KEGG version of "homo sapiens" library. For the network analysis, the Metabolite-Metabolite Interaction Network mode was chosen.

## 3.3.3. Results and Discussion

*Multivariate modelling and metabolite identification*

An untargeted metabolomics approach was used to explore metabolome changes in urine in response to different patterns of coffee consumption. UHPLC-TWIMS-QTOF data sets, obtained in positive and negative ionization modes, were separately submitted for data analysis. A total of 15714 and 19591 features were initially peak picked for positive and negative modes, respectively. Most likely, the high number of detected features was due to the use of ion mobility between LC and MS detector. Indeed, Rainville and co-authors have quantified that the features detected in urine increased up to 41% when adding a further dimension of separation as provided by ion mobility between the LC system and the Q-TOF, most likely due to a combination of separation of co-eluting compounds and noise reduction [36].

At first, the PCA of non-averaged samples was employed to explore the data obtained. Score scatter plots for both positive and negative ionization data are depicted in Figure S1. Both PCA plots demonstrated a grouping of samples associated with coffee consumption. Afterward, technical replicates were merged, and both unsupervised and supervised models were constructed. PLS-DA (Figure 21) applied on positive and negative datasets showed a clear separation between 1C and the other two treatments (PC and 3C), displaying excellent goodness-of-fit ($R^2Y$) and good prediction ability ($Q^2$). Cross-validation of both PLS-DA models indicates that 100% of urine samples analyzed in positive ionization mode were correctly classified, while in ESI(-) the percentage of total correct classification was 98.4% since one sample was not correctly predicted (1C sample predicted as PC). Permutation plots are depicted in Figure S2.

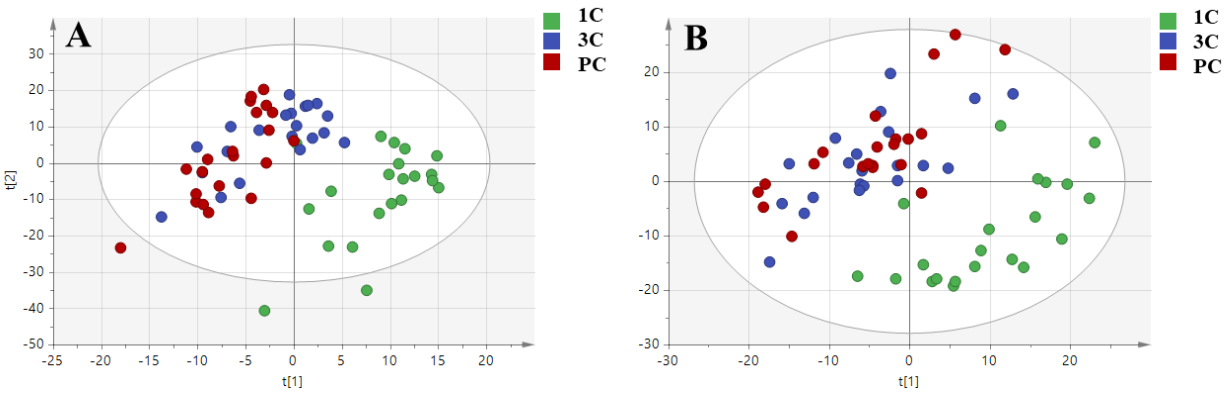



**Figure 21. (A)** PLS-DA model built with positive ionization data ($R^2Y$ = 0.95, $Q^2$ = 0.651) and (**B**) negative ionization data ($R^2Y$ = 0.906, $Q^2$ = 0.769). The intervention treatment groups (1C, 3C, PC) are color coded accordingly.

Subsequently, significant features were selected, retaining those presenting, simultaneously, fold change >2, and FDR adjusted *p*-value (q value) < 0.01 and merged with those showing VIP > 1.2. This filtering step returned a dataset with 3590 significant features for both polarities, which were subjected to the identification. This last step is considered the bottleneck of the whole metabolomics workflow, which remains a major analytical challenge. With mass fragmentation and CCS matching, 153 identifications were assigned out of 3590, meaning that less than 5% of the significant features were translated into knowledge. All significant features are reported, for completeness, in Table S1. We assume that the vast majority of the unidentified features may correspond to coffee or cocoa chemicals and their metabolites. These beverages have a tremendous chemical complexity, and their derived metabolites are largely undocumented and thus absent from reference databases. Endogenous metabolites involved in human metabolic pathways are currently better covered in databases. The annotation of discriminant metabolites, information regarding their biochemical class, and statistical parameters (ANOVA *p*-value and fold change) of each metabolite are reported in Table 9. Further analytical details on accurate mass, detected adduct, formula, error ppm, CCS value and retention time are summarized in Table S2. In parallel to 3-groups modelling, 2-groups comparison was performed between 3C vs. 1C, PC vs. 1C and 3C vs. PC. PLS-DA models were built, and their plots showing excellent clustering are summarized in Figure S3. Significant metabolites from the binary comparisons with their fold changes are reported in Table 9.

*Metabolism of coffee and cocoa phytochemicals (biomarkers of intake)*

More than one hundred out of the 153 annotated metabolites were generated from phytochemicals present in coffee and the CBPCC. These metabolites belonged to various chemical classes, including cinnamic acids, imidazopyrimidines, naphthofurans, pyridine derivatives, phenols, and benzene derivatives, among others (Table 9).

As expected, we observed an increase in urinary levels of caffeine and its related metabolites (dimethylxanthines, monomethylxanthines, and methyluric acids, with the sole exceptions of theobromine and its metabolites) upon increased coffee consumption, reaching the highest intensities for the 3C treatment (3 cups of coffee per day). After being rapidly absorbed and metabolized in the liver, caffeine and its metabolites are slowly removed from the circulation [37].

Trigonelline followed the same trend in urine samples, with the highest mean value for 3C and lowest in 1C group. Conversely to caffeine, trigonelline is excreted unmetabolized [17], and has already

**Table 9.** Statistically significant urine metabolites annotated using UHPLC-TWIMS-QTOF. Metabolite levels that increase in response to the first considered treatment are shaded red and metabolite levels that decrease are colored green. The fold change value is calculated as the ratio between the highest mean concentration reported in a specific treatment (1C, 3C or PC) for a metabolite and the lowest mean concentration reported for the same metabolite in the respective treatment. A red-shaded value of 1.0 for the mentioned ratio means that the real value was, for example, 1.023 rounded to 1.0, indicating an increase, although small.

| Class | Sub Class | Putative identification | q value | Fold Change | | |
|---|---|---|---|---|---|---|
| | | | | 3C/1C | 3C/PC | PC/1C |
| Azoles | Imidazoles | 3-(Imidazol-5-yl)lactate[c] | 6.3E-03 | 0.5 | 0.4 | 1.3 |
| | | 4-Amino-1H-imidazole-5-carboxylic acid[*, c] | 1.0E-16 | 2.4 | 1.2 | 2.0 |
| Benzene and substituted derivatives | Aniline and substituted anilines | 4-Aminophenol[*, c] | 1.0E-16 | 2.8 | 0.8 | 3.5 |
| | | Aniline[*, b] | 3.8E-09 | 1.9 | 0.6 | 3.2 |
| | Anilines | 4-Methoxyaniline[c] | 1.2E-11 | 3.8 | 0.9 | 4.4 |
| | Benzaldehydes | 3-Hydroxybenzaldehyde[*, c] | 4.5E-12 | 1.9 | n.s. | n.s. |
| | Benzenesulfonic acids and derivatives | 3,4-Dihydroxybenzenesulfonic acid[b] | 2.8E-08 | 1.4 | 0.9 | 1.6 |
| | Benzoic acids and derivatives | Hydroxybenzoic acid[*, c] | 1.0E-16 | 2.5 | 1.1 | 2.2 |
| | | Benzoic acid-4-sulfate[*, c] | 3.9E-07 | 1.1 | 0.5 | 2.3 |
| | Benzyl alcohols | Benzyl Alcohol[*, c] | 1.0E-07 | n.s. | 1.6 | n.s. |
| | Phenethylamines | Tyramine[c] | 1.3E-13 | 8.8 | 0.7 | 13.4 |
| | | Phenethylamine[*,b] | 1.0E-16 | 7.0 | 0.8 | 8.5 |
| | Phenols and derivatives | Benzene-1,3,5-triol[c] | 7.7E-03 | 1.3 | 1.1 | 1.2 |
| | | Benzene-1,3-diol[c] | 2.7E-08 | 1.4 | 0.9 | 1.6 |
| | Phenylacetaldehydes | Phenylacetaldehyde[b] | 4.0E-02 | 0.8 | 0.3 | 3.1 |
| | Phenylmethylamines | 4-Hydroxybenzylamine[c] | 1.0E-16 | 5.9 | 1.0 | 6.2 |
| | | Phenylmethanamine[*, c] | 1.0E-16 | 3.8 | 0.8 | 4.5 |
| | Phenylpyruvic acid derivatives | 3-(4-Hydroxyphenyl)Pyruvate[*, c] | 1.3E-07 | n.s. | 3.5 | n.s. |
| Benzenoids | Phenols | 6-Hydroxydopamine[*, c] | 2.4E-11 | 1.8 | n.s. | n.s. |
| Carboxylic acids and derivatives | Amino acids, peptides, and analogues | Glutamine[*, c] | 1.4E-10 | 2.9 | n.s. | n.s. |
| | | 3,4-Dihydroxy-L-Phenylalanine[c] | 8.6E-10 | 1.4 | 1.0 | 1.3 |
| | | 6-acetamido-3-aminohexanoic acid[c] | 2.1E-09 | 1.2 | 0.5 | 2.5 |
| | | Arginine[c] | 1.0E-16 | 7.7 | n.s. | n.s. |
| | | Phenylalanine[c] | 7.2E-14 | 3.2 | 1.2 | 2.7 |
| | | 3-Methoxy-L-Tyrosine[c] | 1.3E-04 | 1.7 | n.s. | n.s. |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | 4-Hydroxy-L-Phenylglycine[*, c] | 4.4E-10 | 1.7 | n.s. | n.s. |
| | | Dihydrodipicolinic acid[b] | 2.2E-02 | 3.0 | 1.9 | 1.6 |
| | | N,N-Dimethylarginine (ADMA)[c] | 3.5E-02 | 0.7 | 0.6 | 3.6 |
| | | N5-(1-Carboxyethyl)-ornithine[c] | 3.2E-12 | 1.7 | 0.5 | 3.6 |
| | | N-Acetylleucine[c] | 9.6E-04 | 2.0 | 0.8 | 1.7 |
| | | N-Acetyl-L-lysine[c] | 4.3E-05 | 2.2 | 1.1 | 2.1 |
| | | O-Acetyl-L-homoserine[b] | 9.1E-13 | 4.8 | 4.5 | 1.1 |
| | | Phosphocreatine[*, c] | 6.4E-13 | 2.0 | 1.4 | n.s. |
| | | Tilarginine acetate[*, c] | 2.9E-04 | 0.9 | 0.7 | n.s. |
| | | 4-Methylene-L-glutamate[c] | 2.9E-12 | 4.4 | 1.3 | 3.4 |
| | Dicarboxylic acids and derivatives | Glutarate[c] | 2.4E-03 | 1.5 | n.s. | n.s. |
| | Tricarboxylic acids and derivatives | Citric Acid[c] | 6.2E-04 | n.s. | n.s. | 0.6 |
| Cinnamic acids and derivatives | Cinnamic acids | Cinnamic acid[*, c] | 2.0E-03 | 1.7 | 1.3 | 1.3 |
| | | 4′-methoxycinnamic acid[b] | 1.3E-04 | 2.0 | 1.8 | 1.1 |
| | Hydroxycinnamic acids and derivatives | Cinnamic acid-sulfate[*,b] | 1.0E-16 | 3.6 | 1.3 | 2.8 |
| | | Dimethoxycinnamic acid [b] | 1.7E-04 | 1.1 | 2.1 | 0.5 |
| | | 3′-Hydroxycinnamic acid-4′-sulfate[b] | 1.6E-04 | 1.9 | 1.7 | 1.1 |
| | | 4′-Hydroxy-3′-methoxycinnamic acid[b] | 2.7E-03 | 1.6 | 1.8 | 0.9 |
| | | 3′-Hydroxycinnamic acid[c] | 3.7E-02 | 1.6 | 1.2 | 1.3 |
| | | 4′-Hydroxycinnamic acid[c] | 1.2E-02 | 1.7 | 2.0 | 0.9 |
| | | 3-Propanoic acid-3′-sulfate[b] | 2.8E-02 | 0.7 | 1.4 | 0.5 |
| Coumarins and derivatives | Hydroxycoumarins | 7-Hydroxy-3-(4-methoxyphenyl)-4-propylcoumarin[c] | 1.2E-04 | 1.4 | 1.0 | 1.3 |
| Diazines | Pyrazines | 3,5-diethyl-2-methylpyrazine[b] | 1.0E-16 | 1.2 | 0.2 | 5.9 |
| | Pyrimidines and pyrimidine derivatives | 5-Acetylamino-6-formylamino-3-methyluracil[b] | 1.6E-02 | 2.4 | 1.2 | 2.0 |
| Dihydrofurans | Furanones | L-Ascorbic acid 2-glucoside[b] | 2.3E-07 | 3.1 | 1.3 | 2.4 |
| | | 3-Hydroxy-4,5-dimethyl-2(5H)-furanone[c] | 7.1E-07 | 2.6 | 2.2 | 1.2 |
| Fatty Acyls | Fatty acid esters | Sorbate[*, c] | 9.1E-10 | 1.8 | n.s. | 0.9 |
| | Fatty acids and conjugates | Ethylmalonic Acid[c] | 1.4E-02 | 1.7 | n.s. | 1.3 |
| | | Hydnocarpic acid[b] | 6.3E-04 | 2.6 | 1.3 | 2.1 |
| | | Hydroxy-methylsuccinic acid[*, c] | 1.0E-16 | 1.7 | 0.3 | 5.3 |
| | Lineolic acids and derivatives | Methyl Jasmonate[c] | 3.3E-11 | n.s. | n.s. | 2.3 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Glycerophospholipids | Glycerophosphates | Sn-Glycero-3-Phosphocholine[c] | 1.0E-16 | n.s. | n.s. | 7.1 |
| Hydroxy acids and derivatives | Beta hydroxy acids and derivatives | O-Methyl-(epi)catechin-sulfate[*, b] | 1.0E-16 | 2.8 | 0.4 | 7.3 |
| Imidazopyrimidines | Purines and purine derivatives | 1,3,7-Trimethyluric acid*[b] | 1.9E-06 | 4.5 | 1.3 | 3.4 |
| | | 1,3-Dimethyluric acid*[,a] | 1.2E-06 | 2.8 | 1.3 | 2.1 |
| | | 1,7-Dimethyluric acid*[,a] | 1.1E-06 | 2.5 | 1.1 | 2.2 |
| | | 1-Methyluric acid[*,a] | 1.0E-16 | 2.1 | 1.1 | 2.0 |
| | | 3,7-Dimethyluric acid[*,a] | 1.0E-16 | 1.9 | 0.2 | 8.1 |
| | | 3-Methyluric acid[a] | 1.0E-16 | 2.1 | 0.4 | 5.8 |
| | | 3-Methylxanthine[*, b] | 1.0E-16 | 1.8 | 0.3 | 5.4 |
| | | 7-Methylxanthine[*, b] | 1.0E-16 | 1.7 | 0.3 | 4.9 |
| | | 1-Methylxanthine[b] | 1.5E-06 | 2.5 | 1.3 | 1.9 |
| | | Caffeine[*, a] | 5.3E-09 | 4.0 | 1.2 | 3.3 |
| | | Paraxanthine[*, a] | 1.0E-16 | 1.8 | 0.3 | 6.7 |
| | | Theobromine[*, a] | 4.0E-07 | 2.2 | 1.1 | 2.0 |
| | | Theophylline[*,a] | 1.5E-07 | 2.9 | 1.3 | 2.3 |
| | | Uric acid*[,c] | 1.2E-02 | 1.1 | 1.2 | 1.1 |
| Indoles and derivatives | Indoles and derivatives | Indole[c] | 4.1E-04 | n.s. | n.s. | 1.6 |
| | Indolyl carboxylic acids and derivatives | 3-Methylindolepyruvate*[,b] | 4.6E-10 | 4.7 | 1.9 | 2.4 |
| | | Indole-3-methylethanoate[c] | 2.1E-10 | 3.3 | 1.8 | 1.9 |
| | Tryptamines and derivatives | 5-Hydroxy-L-Tryptophan[c] | 1.1E-04 | 0.9 | 1.4 | 0.7 |
| Keto acids and derivatives | Gamma-keto acids and derivatives | Hydroxy-oxovaleric acid[c] | 7.0E-09 | 1.9 | 1.3 | 1.5 |
| | Medium-chain keto acids and derivatives | 2-Oxoadipate[c] | 1.3E-02 | n.s. | n.s. | 1.5 |
| | Short-chain keto acids and derivatives | Oxo-pentenoic acid[c] | 1.4E-02 | 1.4 | 1.1 | 1.3 |
| Organonitrogen compounds | N-arylamides | 5-Acetylamino-6-amino-3-methyluracil[*,b] | 1.0E-16 | 2.7 | 1.1 | 2.3 |
| Organooxygen compounds | Alcohols and polyols | 3-Dehydroshikimate[*] | 6.2E-12 | 2.1 | n.s. | n.s. |
| | | Pantothenic acid[b] | 4.2E-11 | 4.5 | 2.4 | 1.9 |
| | Carbohydrates and carbohydrate conjugates | D-Sorbitol[c] | 1.0E-16 | 7.8 | n.s. | n.s. |
| | | Galactitol*[,c] | 3.8E-02 | 1.2 | 0.6 | 2.0 |
| | | 1-Ribosylnicotinamide[b] | 2.4E-08 | 2.7 | 1.4 | 2.0 |
| | | D-Glucuronic Acid[c] | 3.0E-04 | 2.0 | n.s. | n.s. |
| | | Inosine monophosphate[c] | 1.0E-16 | 9.0 | n.s. | n.s. |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | Mannitol[c] | 1.0E-16 | 10.0 | n.s. | n.s. |
| | | Raffinose[c] | 1.0E-16 | n.s. | n.s. | 8.8 |
| | Carbonyl compounds | 1-Methyl-2-pyrrolecarboxaldehyde*[,c] | 1.1E-03 | 1.8 | n.s. | 1.8 |
| Peptidomimetics | Hybrid peptides | Carnosine[c] | 9.0E-09 | 5.5 | n.s. | n.s. |
| Phenyl-γ-valerolactones and phenylvaleric acids | Phenyl-γ-valerolactones | 5-(Hydroxyphenyl)-γ-valerolactone*[,b] | 1.0E-16 | n.s. | n.s. | 17.4 |
| | | 5-(Phenyl)-γ-valerolactone-sulfate*[,b] | 1.0E-16 | n.s. | 0.05 | n.s. |
| | | 5-(3′,4′-Dihydroxyphenyl)-γ-valerolactone[b] | 1.0E-16 | n.s. | 0.02 | n.s. |
| | | 5-(Hydroxyphenyl)-γ-valerolactone-sulfate*[,b] | 1.0E-16 | n.s. | 0.04 | 1.2 |
| | | 5-(Hydroxyphenyl)-γ-valerolactone-sulfate[b] | 1.7E-02 | 1.1 | 0.85 | 1.2 |
| | | 5-(Hydroxyphenyl)-γ-valerolactone-glucuronide*[,b] | 1.0E-16 | n.s. | 0.04 | n.s. |
| | | 5-(Methoxyphenyl)-γ-valerolactone-sulfate[c] | 1.0E-16 | n.s. | n.s. | 8.3 |
| | | 5-(Dihydroxyphenyl)-γ-valerolactone-sulfate (3′,4′,5′)*[,c] | 1.4E-08 | n.s. | n.s. | 2.4 |
| | Phenylvaleric acids | 5-(Hydroxyphenyl)valeric acid*[,c] | 1.7E-02 | 0.5 | 1.0 | 2.1 |
| | | 5-(Phenyl)valeric acid-sulfate*[,c] | 1.9E-03 | 0.3 | 0.8 | 3.1 |
| Phenols | Benzenediols | 3,4-Dihydroxyphenylacetic acid[c] | 4.5E-03 | 1.8 | n.s. | n.s. |
| | | 3-Hydroxymethyl-phenol*[,c] | 6.0E-03 | 1.7 | 1.5 | 1.1 |
| | | Dopamine*[,c] | 1.8E-08 | 1.4 | 1.0 | 1.4 |
| | Benzenetriols and derivatives | Benzene-1,2,4-triol[c] | 5.7E-09 | 3.0 | 2.8 | 1.1 |
| | | 4-Methylphenol-2-sulfate*[,b] | 1.1E-08 | 1.4 | 0.8 | 1.7 |
| | | 4-Methylbenzenediol-sulfate*[, c] | 1.0E-10 | 1.9 | 1.3 | 1.5 |
| | Methoxyphenols | 2-Hydroxy-2-(4′-hydroxy-3′-methoxyphenyl)acetic acid [c] | 2.1E-07 | n.s. | n.s. | 2.4 |
| | | 3-Methoxytyramine*[,c] | 1.0E-04 | 1.9 | n.s. | n.s. |
| | | 4-Hydroxy-3-Methoxyphenylglycol [c] | 2.1E-04 | n.s. | n.s. | 2.3 |
| | | DL-Normetanephrine*[,c] | 3.6E-03 | n.s. | n.s. | 1.5 |
| | | 3-methoxybenzaldehyde-4-sulfate*[,b] | 1.0E-16 | 2.5 | 0.7 | 3.4 |
| Prenol lipids | Sesquiterpenoids | Farnesyl Diphosphate[c] | 1.3E-02 | n.s. | n.s. | 2.7 |
| Purine nucleosides | Purine 2'-deoxyribonucleosides | Deoxyadenosine[c] | 1.0E-16 | 17.6 | n.s. | 3.9 |
| | | Deoxyadenosine monophosphate[c] | 1.1E-11 | 3.2 | n.s. | n.s. |
| Pyridine nucleotides | Nicotinamide nucleotides | Nicotinamide Mononucleotide[c] | 1.0E-16 | n.s. | 0.03 | n.s. |
| Pyridines and derivatives | Pyridinecarboxylic acids and derivatives | 5-Hydroxy-6-methylnicotinic acid *[,b] | 1.7E-03 | 3.3 | 2.2 | 1.5 |
| | | 6-amino nicotinamide*[, b] | 1.0E-16 | 2.0 | 0.3 | 6.6 |
| | | N,N-Diethylnicotinamide[b] | 5.7E-03 | 1.4 | 0.8 | 1.6 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Hydropyridines | 2-Hydroxypyridine[*, c] | 6.7E-12 | 3.0 | 1.1 | 2.7 |
| | | 2,6-Dihydroxypyridine[c] | 2.7E-05 | 1.6 | 1.2 | 3.0 |
| Steroids and steroid derivatives | Steroidal glycosides | estriol 3-glucuronide[*, b] | 3.5E-13 | 3.6 | 0.9 | 4.1 |
| | Bile acids, alcohols and derivatives | Taurocholic Acid[c] | 2.9E-13 | n.s. | 14.2 | n.s. |
| | Estrane steroids | Estriol *[,b] | 9.6E-05 | 2.7 | 1.2 | 2.3 |
| | | 2-methoxy-17beta-estradiol 3-glucosiduronic acid[b] | 1.5E-02 | 2.4 | 1.3 | 1.9 |
| | | 3-O-(Carboxymethyl)estrone *[,b] | 2.7E-03 | 2.5 | 1.5 | n.s. |
| | Pregnane steroids | Progesterone[c] | 1.0E-16 | 3.8 | n.s. | n.s. |
| Benzene and substituted derivatives | Miscellaneous | 2-Hydroxy-2-phenylacetic acid *[,c] | 1.4E-03 | 1.8 | 2.0 | 0.9 |
| | | Phenylacetic Acid *[,b] | 1.1E-03 | 1.7 | 0.9 | 2.0 |
| | | Phenylethanol[c] | 1.8E-07 | 2.1 | n.s. | n.s. |
| Benzimidazoles | | Dimethylbenzimidazole[c] | 3.1E-06 | 5.8 | n.s. | n.s. |
| Miscellaneous | | ((17-Oxoestra-1,3,5(10)-trien-3-yl)oxy)acetic acid *[,b] | 7.5E-09 | 3.2 | 0.8 | 3.8 |
| | | 2-Deoxyinosose[*, b] | 1.0E-16 | 1.6 | 0.3 | 5.4 |
| | | 4-Oxocyclohexanecarboxylic acid *[,b] | 4.4E-09 | 1.5 | 0.8 | 1.8 |
| | | Guanidino-butanol[b] | 5.1E-02 | 1.3 | 2.7 | 0.5 |
| | | Hydroxy-(indol-yl)ethanamine[c] | 1.1E-02 | 1.0 | 1.2 | 0.8 |
| | | Trigonelline[*, b] | 2.9E-10 | 2.2 | 1.3 | 1.7 |
| | | Trihydroxy-5alpha-cholan-24-yl sulfate[*, c] | 1.0E-16 | 3.7 | 1.7 | n.s. |
| | | Kahweol oxide glucuronide[*, b] | 2.6E-13 | 3.7 | 0.8 | 4.4 |
| Phenol esters | | Phenylacetic acid isomer[*,c] | 1.0E-16 | 4.0 | n.s. | n.s. |
| Phenylpropanoic acids | | 3-(2-Hydroxyphenyl)propanoic acid *[,c] | 4.8E-05 | 1.0 | 1.9 | 0.5 |
| | | 3-(3'-Hydroxy-4'-methoxyphenyl)propanoic acid[c] (Dihydroisoferulic acid) | 1.1E-03 | 1.8 | 2.2 | n.s. |
| | | 3-(3'-Methoxy-4'-hydroxyphenyl)propanoic acid[b] (Dihydroferulic acid) | 3.9E-03 | 1.8 | 2.1 | 0.5 |
| | | 2-Hydroxy-3-(4'-hydroxyphenyl)propanoic acid [b] | 9.1E-07 | 2.2 | n.s. | n.s. |
| Purine nucleosides | | 2-Aminoadenosine[*, c] | 1.4E-03 | 0.9 | 0.7 | 1.3 |
| | | Inosine[*, c] | 4.4E-03 | n.s. | n.s. | 1.8 |
| Pyrrolizines | | 2,3-Dihydro-5-(3-hydroxypropanoyl)-1H-pyrrolizine[b] | 8.4E-03 | 3.0 | 1.3 | 2.3 |

Class and Sub class have been taken from Human Metabolome Database (HMDB); nomenclature for dietary (poly)phenolic catabolites is in accordance with the recent recommendations proposed by Kay and co-authors [73], nomenclature for the other metabolites is in accordance with the common names present in the DBs used (mainly Metlin, PubChem and HMDB).

q value is the FDR adjusted *p*-value. n.s.: not significant according to *p*-value < 0.01 in the 2-groups comparison.

[*] indicates metabolites common to both corrected *p*-Value < 0.01 and VIP > 1.2.

[a] identified metabolites (level I).

[b] putatively identified metabolites (level II).

[c] putatively characterized metabolites (level III)

1C, 3C and PC are the three treatments under investigation, respectively characterized by the consume for a month of one cup of espresso coffee/day, three cups of espresso coffee/day and one cup of espresso coffee/day and two cocoa-based products containing coffee (CBPCC) twice per day.

been found to significantly correlate with coffee consumption. Indeed, it has been proposed as a biomarker of coffee intake, alone [20] or in combination with 1-methylxanthine and cyclo(isoleucylprolyl) in previous studies [38]. However, neither cyclo(isoleucylprolyl) nor the diterpene atractyligenin glucuronide, recognized as specific biomarkers of coffee consumption, were detected in the present study [39]. This observation reinforces the potential of trigonelline to serve as a candidate biomarker of coffee intake.

Another important class of coffee-derived metabolites is that of cinnamic acids, which originated mainly from the metabolism of chlorogenic acids, the main phenolic compounds found in coffee [1]. This metabolism takes place both at the upper and lower level of the gastrointestinal tract, with the latter involving the gut microbiota. Once absorbed, coffee hydroxycinnamates are then subjected to phase II metabolism at the hepatocyte level and enter into circulation [1]. In line with this prediction, coumaric acid-sulfate was the most significant marker of this class of compounds, with the highest intensity in the 3C treatment and lowest in the 1C group. However, these compounds cannot serve as selective biomarkers of coffee intake, because of their very poor specificity and their colonic origin [17,38], which is inevitably affected by a high inter-individual variation due to intrinsic variability of the human gut microbiota [40].

Investigating the PC intervention, in which coffee and cocoa intake were combined, an increased amount of theobromine derivatives, kahweol oxide glucuronide and nicotinamides was observed in urine (Table 9). Theobromine is known to be found in cocoa and to be rapidly absorbed and converted in, among other metabolites, 3,7-methyluric acid, 3-methylxanthine, 7-methylxanthine and 3-methyluric acid [41]. These four metabolites were found to be greatly excreted after the PC treatment and, in particular, 3,7-methyluric acid and 3-methyluric acid were good discriminant markers for both positive and negative ionization modes (Table 9 and Table S2). On the other hand, although kahweol oxide glucuronide and nicotinamides mostly derive from coffee consumption, they changed notably as a consequence of the PC treatment. Both metabolites might be considered markers of different roasting processes or coffee brewing styles, different from the espresso [42,43]. Thus, this result might indicate that the cocoa-based products most likely contained a coffee with a phytochemical composition different to that of the espresso coffees consumed by the volunteers.

Moreover, the significant presence of compounds coming exclusively from cocoa constituents rather than coffee, like flavan-3-ol metabolites and phenethylamines, was observed. Eight phenyl-γ-valerolactone glucuronides and sulfates were detected as markers for the PC treatment, these having been reported to originate by colonic microbial catabolism of flavan-3-ols [44]. In particular, 5-

(3',4'-dihydroxyphenyl)-γ-valerolactone and its glucuronide and sulfate derivatives were among the metabolites showing the most significant fold changes compared to the espresso coffee treatments (1C and 3C).

*Biological interpretation of significant markers*

Untargeted metabolomics allowed the identification of unrelated metabolites to coffee and cacao. The role of these metabolites was evaluated by pathway analysis.

**Pathway analysis**

The pathway analysis shows the main metabolic routes modulated by the modes of coffee consumption under study (Figure 22). All the identified pathways resulted in being upregulated

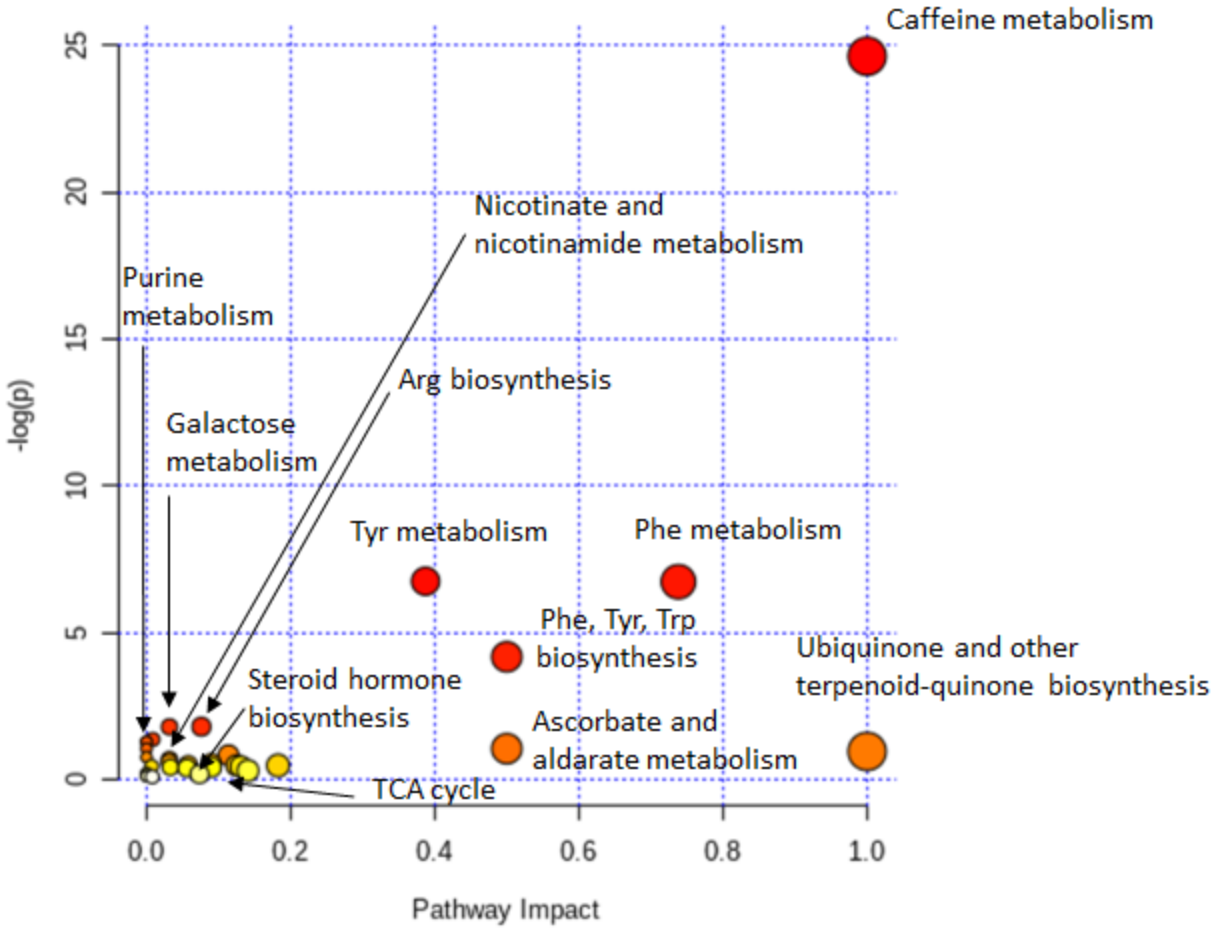



**Figure 22.** Pathway analysis performed with all the significant metabolites identified. The output displays metabolic pathways arranged by scores from pathway enrichment (y-axis) and topology analysis (x-axis). The color and size of each circle are based on p-values and pathway impact values, respectively (from yellow to red, the –log(*p*) increases, the bigger the circle size, the higher the pathway impact value).

following coffee intake (Table S3). As previously stated, caffeine is the main bioactive compound in coffee beans and abundantly present in coffee. Once ingested, it is rapidly absorbed and metabolized into more hydrophilic metabolites that can be excreted in the urine. Up to nine known

metabolites corresponding to this pathway were identified in urine samples following coffee intake, highly impacting on the urinary metabolome. Other metabolic routes influenced by the intake of coffee were the metabolism and biosynthesis of specific amino acids (in particular phenylalanine, tyrosine and arginine), ascorbate and aldarate metabolism, ubiquinone and other terpenoid-quinone biosynthesis, galactose metabolism, purine metabolism, nicotinate and nicotinamide metabolism, steroid hormone biosynthesis and citrate cycle (TCA cycle), as depicted in Figure 22. Phenylalanine biosynthesis and metabolism, arginine, ubiquinone and other terpenoid-quinone biosyntheses are paths that had never been linked to coffee consumption before.

Phenylalanine metabolism resulted as the second most perturbed pathway (Figure 22). In particular, the route arising from the catabolism of phenylalanine into phenylacetic acid was significantly altered. In this sense, phenylalanine metabolism, phenylalanine, tyrosine and tryptophan biosynthesis and tyrosine metabolism were influenced by coffee consumption. Other pathways involving amino acids affected by the mode of coffee consumption were arginine biosynthesis and metabolism. Arginine is a semi-essential amino acid that has many functions, including being involved in the urea cycle, as a precursor of nitric oxide, creatine, glutamate, and proline, and it can be converted into glucose and glycogen if needed [31]. On the other hand, the activation of nicotinate and nicotinamide metabolism is probably a consequence of the presence of trigonelline and other pyridines (*N*-methylpyridinium and niacin) in coffee [43], as well as of caffeine for purine metabolism. The impact on the TCA cycle (via citric acid) suggests an influence on energy metabolism. The effects of coffee and its constituents on energy metabolism has been observed and extensively studied [45–47], with several mechanisms of action having been proposed and reviewed [9,48], but none involving the TCA cycle. A similar effect towards the TCA cycle was observed by Takahashi and colleagues in mice when, through an integrated multi-omics study, researchers found that TCA cycle-related proteins in mice were upregulated upon coffee consumption [49]. Among these upregulated proteins, NADH dehydrogenase (ubiquinone), which may explain the presence of ubiquinone and other terpenoid-quinone biosynthesis among the main perturbed pathways. The underlying mechanism by which coffee, through one or more of its components, determines this upregulation of TCA cycle enzymes is still unknown.

**Functional response metabolites (biomarkers of effect)**

A dose-dependent increase following coffee intake (1 and 3 cups of espresso coffee per day) was also observed for unusual pathways that did not involve bioactive coffee compounds directly (Figure 23). This was the case of some amino acids, purine nucleosides and steroids.

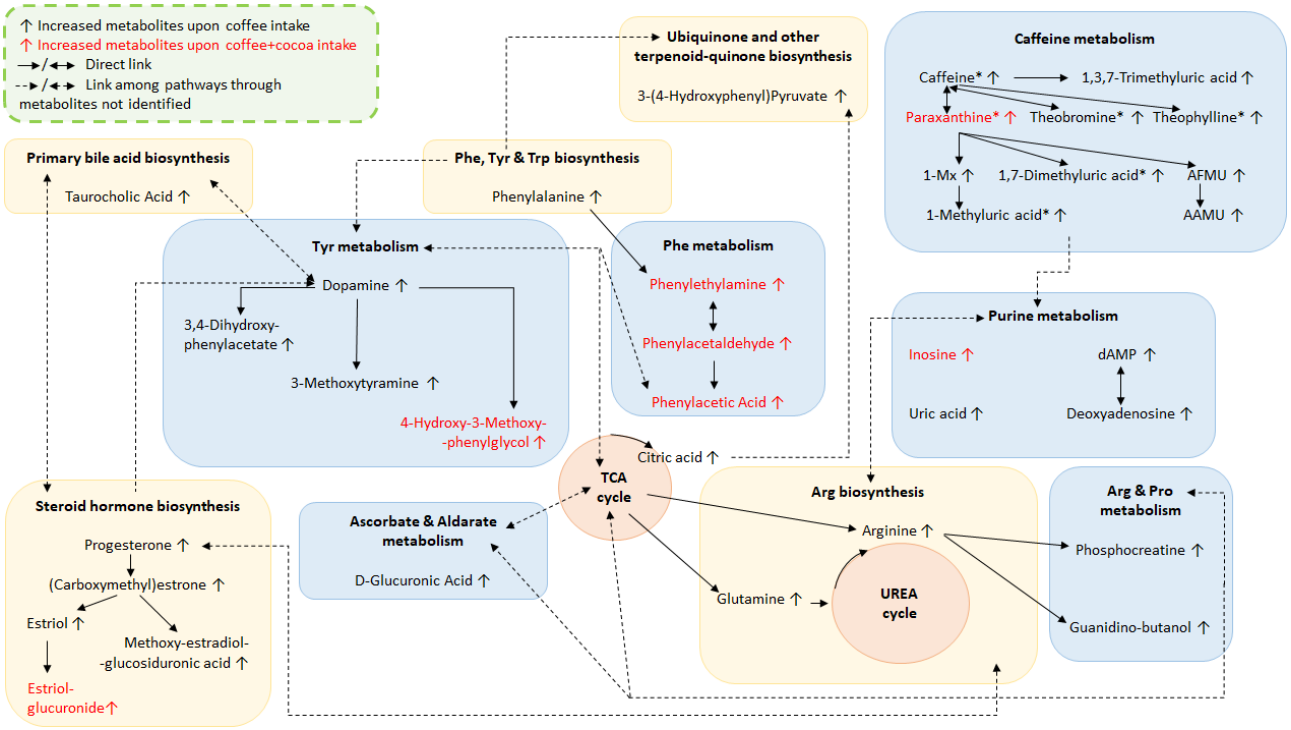



**Figure 23.** Detailed metabolic pathways involved and the significant metabolites (endogenous and exogenous) identified in urine after coffee or coffee and cocoa intake and their inter-connections. Information on pathways has been drawn from KEGG database. * indicates level I identified metabolites. Phe: phenylalanine; Tyr: tyrosine; Trp: tryptophan; Arg: arginine; Pro: proline; 1-Mx: 1-methylxanthine, AFMU: 5-Acetylamino-6-formylamino-3-methyluracil; AAMU: 5-Acetylamino-6-amino-3-methyluracil; TCA cycle: tricarboxylic acid cycle/citrate cycle.

*Amino acids*

Several amino acids are naturally present in coffee. However, the significant presence of arginine and phenylalanine in the urinary metabolome upon coffee consumption (3C) (Table 9) is more likely due to the activation of metabolic pathways leading to their syntheses, as described above [1,50]. To the best of our knowledge, the link of arginine with coffee seems to be new, since no other study identified changes in arginine levels upon coffee consumption, nor suggested a potential effect of coffee on human health through the modulation of arginine biosynthesis. Cornelis and collaborators did not report arginine but found a slight decrease in homoarginine levels following coffee intake [24]. Homoarginine is an endogenous non-proteogenic amino acid produced from arginine and lysine by the catalytic action of arginine:glycine amidinotransferase (AGAT) [51,52]. Recent epidemiological studies have demonstrated an association between low circulating concentrations of L-homoarginine and an increased risk of cardiovascular and all-cause mortality [53,54]. The high levels of free arginine in urine, after the 3C treatment, could have different implications—for example, a decreased conversion into homoarginine or reduced production of dimethylarginines, among others. If homoarginine favors human health by promoting nitric oxide synthesis, endogenous dimethylated derivatives of arginine (asymmetric dimethylarginine and symmetric

dimethylarginine) are generally accepted cardiovascular risk factors [52,55]. Notably, asymmetric dimethylarginine slightly increased (3.6 fold-change) after the PC treatment compared to a lower coffee dose (1C). These observations should be investigated more in-depth to explore any possible association with disease prevention or development.

Phenylalanine is the essential amino acid precursor of tyrosine. Via tyrosine metabolism, it is also a precursor for catecholamines, like dopamine. Regarding phenylalanine, increased levels of this amino acid were registered for the first time as a consequence of high coffee consumption (3C vs. 1C and PC vs. 1C) (Table 9), likely having an impact on the synthesis and metabolism of tyrosine and dopamine. This result is in contrast with the findings of the EPIC-Potsdam Study on the evaluation of various biomarkers as potential mediators of the association between coffee consumption and incident type 2 diabetes [56]. Actually, in that study authors reported an inverse association between coffee consumption and plasma levels of phenylalanine in men. However, they couldn't suggest any plausible biological explanation for this association and a consequent linkage with type 2 diabetes. Dopamine levels also raised after higher coffee intake, highlighting an increased amount of catecholamines as a consequence of coffee consumption. Increased dopamine levels may represent a potential neuroprotective mechanism exerted by coffee [57].

*Purine nucleosides*

A higher excretion of deoxyadenosine was observed for the 3C treatment, mostly as a nucleoside (with a fold change of 17.6 when comparing 3C vs. 1C) but also esterified with a phosphate group forming the corresponding nucleotide (with a fold change of 3.2 when comparing 1C vs. 3C). This compound is a critical component of DNA, and it is linked with coffee consumption for the first time. Its increased levels at the highest coffee intake might be a result of enhanced purine metabolism, as shown by the pathway analysis. Disorders in purine metabolism have been associated with various diseases, such as gout [58], multiple sclerosis [59], and certain cancers [60]. However, the purine compounds showing altered levels in those studies were different from deoxyadenosine and an important role was played by uric acid, the final product of purine degradation. In our study, uric acid increased only slightly following coffee intake (Table 9), but this is likely due to caffeine metabolism [61]. This enhanced deoxyadenosine production, not accompanied by increased uric acid excretion, may necessitate further investigation.

*Steroids and steroid derivatives*

The link between coffee and steroid metabolism is not new and has been largely studied mainly because of the possible association with female cancer risk [62–66] and adverse effects during

pregnancy [67,68]. In particular, in the present study, an impact on steroid hormone biosynthesis and augmented levels of progesterone, estriol, (carboxymethyl)estrone, methoxy-estradiol glucosiduronic acid and estriol glucuronide (Table 9) were reported. While most are inactive metabolites usually excreted in the urine, progesterone and estriol are hormones that can affect human health, particularly in female subjects in delicate conditions, such as pregnancy, lactation, or menopause [69,70]. Cornelis and colleagues also found a significant plasmatic enrichment of steroid metabolites after coffee intake but linked to the androgen pathway [24]. This could be due to differences in the study population (at elevated risk of type 2 diabetes vs. healthy subjects), sampling times, analysed samples matrix (plasma vs. urine) or in the types of intervention since subjects in that study had to consume higher doses of coffee (4 to 8 cups per day), and the brewing method was not standardized. Augmented levels of steroid hormones (estriol and progesterone) after 3C treatment compared to 1C are very relevant to consider, but their association with female cancers is controversial, both in pre- and post-menopausal women [62,63,66,71,72]. Further research aimed at elucidating increased hormone levels will be of great interest. A recent meta-analysis associated coffee intake with probable decreased risk of breast and endometrial cancers, among other pathologies, with the lowest risk reached with the consumption of about 4-5 cups/day [9], thus reducing any concern about this specific association.

*Others*

The boost to ascorbate and aldarate metabolism to produce high amounts of D-glucuronate is possibly due to the increased need, following coffee consumption, to remove the many xenobiotic substances introduced, considering that glucuronidation of exogenous compounds is the first phase II mechanism involved in the detoxification of reactive electrophiles and the production of polar metabolites that diffuse less across membranes. Regarding the functional effect of the treatment including cocoa (PC treatment) on pathways associated to specific metabolic responses, it should be noted that the only major route, beyond those related to the metabolism of cocoa xenobiotics, regulated by this intervention in comparison to the other treatments was phenylalanine catabolism into phenylacetic acid.

This work is part of a bigger study having as primary outcome the bioavailability assessment of the three main groups of phytochemicals in roasted coffee (methylxanthines, phenolic compounds, and trigonelline), their modulation by the level of consumption, and the definition of the daily average concentration of coffee derived plasma circulating metabolites following the consumption of different doses of coffee, even after its partial replacement with cocoa-based confectionary

containing coffee in a real-life setting. As a limitation of our observation, however intrinsic in the original study design, the absence of a treatment providing only cocoa products (without coffee) may prevent us from seeing further functional effects of the PC treatment.

### 3.3.4. Concluding remarks

In conclusion, this study showed that – in controlled but realistic conditions – it is possible to detect changes in the metabolome that are associated with different modes of coffee consumption. First, we demonstrated that the use of a holistic approach, such as untargeted metabolomics, may disclose not only the fate of coffee components after ingestion, but also how coffee can modulate endogenous metabolome changes. Indeed, besides the already known coffee and cocoa biomarkers of intake, we detected endogenous metabolites from the phenylalanine, tyrosine and arginine biosynthesis and metabolism, energy metabolism and steroid hormone biosynthesis that were affected by the three modes of coffee consumption and may, in turn, potentially influence human health. Although only 5% of the detected features were identified, our results unveil the complex metabolic pathways that may be modulated by coffee and cocoa consumption, some of which being reported for the first time.

Inter-individual variability, day-to-day variation, and differences in coffee composition, among other aspects that may lead to a more comprehensive evaluation of the effect of coffee on the metabolome, were not assessed due to the complex nature of this study. These shortcomings might be considered in future studies. Actually, the observed changes should be further validated by quantitative measurement in the kinetics of all key metabolites of the modulated pathways, and their biological meaning and potential implications in disease prevention should also be investigated in specifically designed intervention studies.

**Acknowledgments**

Authors thank the volunteers for participating in the study, and Yevgeniy Topolskyy, Francesca Zicca, Debora Fede, Luciano Notarpietro, Matteo Chini, and Chiara Bandini for their assistance in the conduction of the study.

**Author contributions**

C.F. and L.R. contributed equally to this work. P.M., D.D.R., and F.B. designed the study. M.T., A.R., D.M., M.A., A.D.C., and R.B. conducted the study. L.R. and L.A.G. performed the analysis. L.R. and C.F. analyzed and interpreted data and drafted the manuscript. P.M., J.R., C.M., C.D.A. and D.D.R. edited the manuscript. All authors critically read and approved the final version of the manuscript.

**Sources of funding**

This work was partially funded by Soremartec Italia S.r.l. (Alba, Italy). The funder approved the final trial protocol prior to its implementation, but it was not involved in the design of the study, data analysis and interpretation, nor the drafting of this manuscript.

**Conflict of Interest**

Daniele Del Rio and Furio Brighenti have received research grants from Soremartec Italia S.r.l. Lee A. Gethings is a Waters' employee. The rest of the authors declare no conflict of interest.

**Data availability**

Data matrices have been deposited in the following dataset: Righetti, Laura; Favari, Claudia; Mena, Pedro (2020), "Metabolomic changes after coffee consumption: new paths on the block", Mendeley Data, V2, doi: 10.17632/5w936rthzg.2, https://data.mendeley.com/datasets/5w936rthzg/2

## Electronic Supplementary Information (ESI)

**Figure S1.** Unsupervised principal components analysis (PCA) models built from non-averaged samples run in positive ($R^2X$ = 0.808, $Q^2$ = 0.716) and negative ($R^2X$ = 0.807, $Q^2$ = 0.564) ionization modes (A, B). The intervention treatment groups (1C, 3C, PC) are color coded accordingly.
Both PCA plots demonstrated natural grouping of samples by coffee consumption., with some differences between positive and negative modes. Particularly, positive PCA (**Figure S1A**) showed a clear separation between PC and the other two treatments (1C and 3C) (PC1 + PC2 = 59% of total variance). This might be due to metabolites specifically associated with cocoa intake, which were nicely ionized in ESI positive mode. Indeed, the intensities of theobromine metabolites and aromatic amines (as tyramine and phenethylamine) were found to be significantly higher in the PC group. On the other hand, PCA resulting from negative mode proved a nice separation of the three treatments (**Figure S1B**). The first two PCs accounted for the 33% of the total variance.

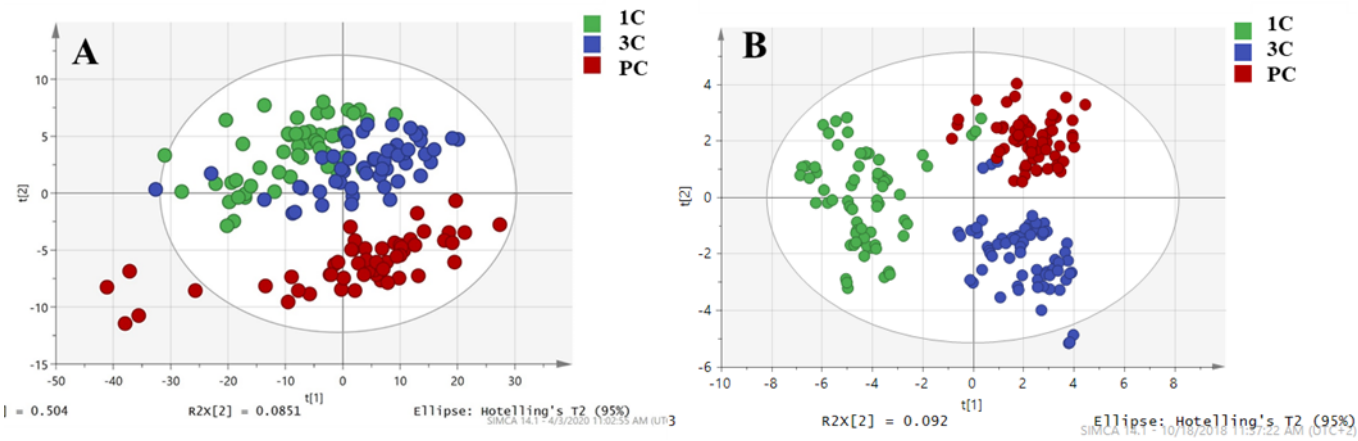



**Figure S2.** Permutation plots obtained for (A) positive and (B) negative ionization modes of 3-groups comparison data.

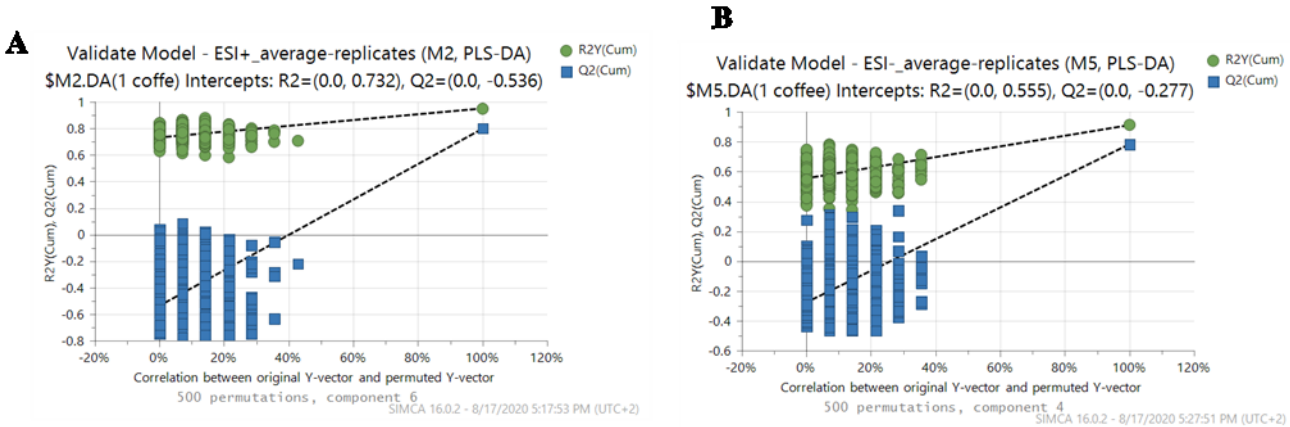


**Figure S3.** PLS-DA plots for two-groups comparison.

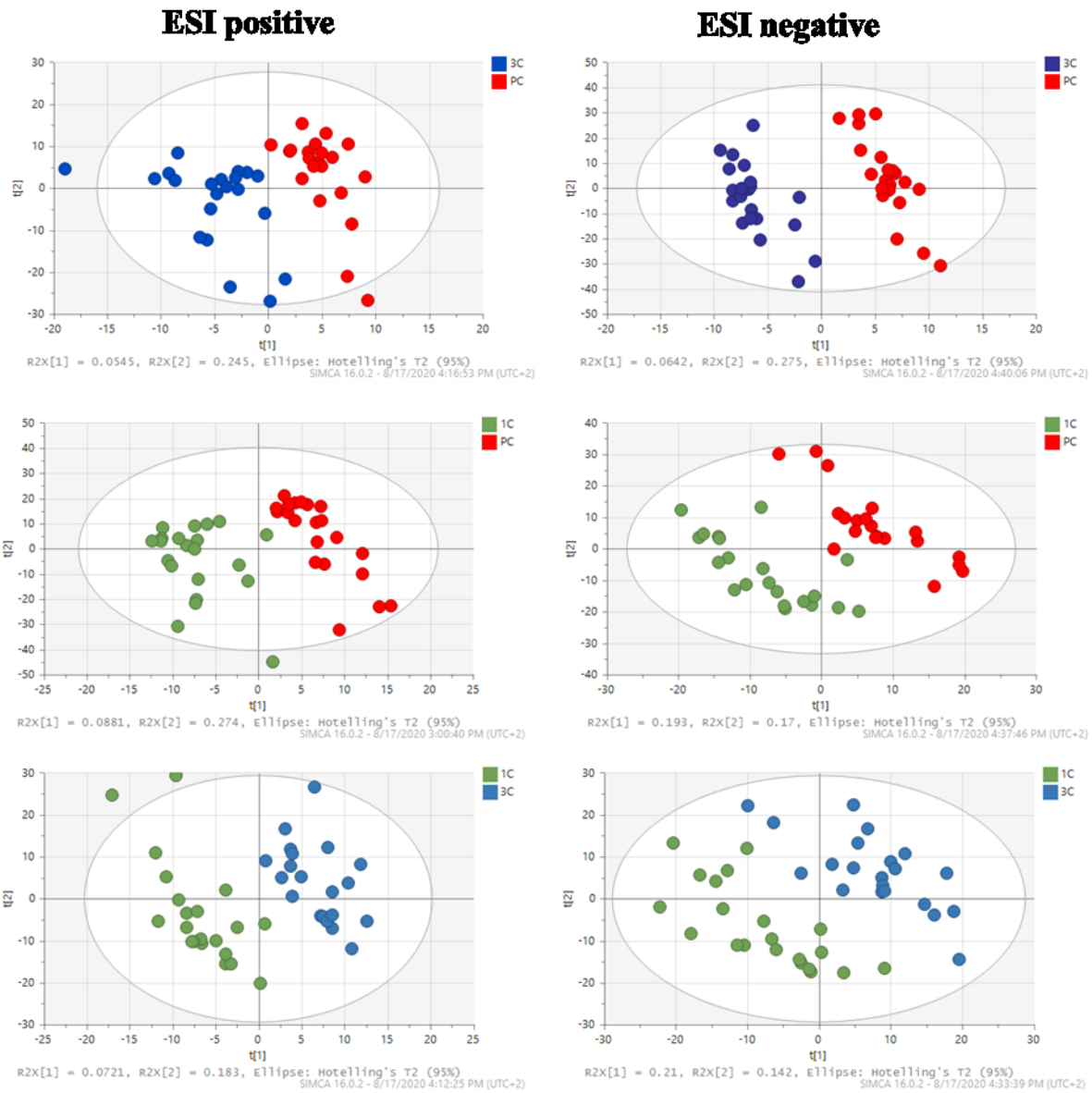



**Table S1.** ESI positive and ESI negative significant features ranked by ANOVA P-Value, including annotated and non-identified metabolites. The annotated metabolites are depicted in bold. *Available at* doi.org/10.1002/mnfr.202000875.

**Table S2.** Chromatographic and spectrometric characteristic of significant annotated metabolites.

| Putative identification | Formula | Adducts | Detected *m/z* | Mass Error [ppm] | RT [min] | CCS [Å²] | dCCS (Å²) |
|---|---|---|---|---|---|---|---|
| 3-(Imidazol-5-yl)lactate [c] | C6H8N2O3 | M+H | 157.0601 | -4.5 | 0.5 | 130 | |
| 4-Amino-1H-imidazole-5-carboxylic acid [*, c] | C4H5N3O2 | M-H2O-H | 108.0198 | -4.0 | 2.1 | 127.5 | |
| 4-Aminophenol [*, c] | C6H7NO | M+H | 110.0606 | 5.6 | 0.5 | 117.9 | 1.6 |
| Aniline [*, b] | C6H7N | M+H | 94.0647 | -4.7 | 0.5 | 117.4 | 2.4 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 4-Methoxyaniline [c] | C7H9NO | M+H | 124.0753 | -2.8 | 0.9 | 121.8 | 0.3 |
| 3-Hydroxybenzaldehyde [*, c] | C7H6O2 | M-H | 121.0291 | -3.0 | 2.9 | 118.1 | -1.2 |
| 3,4-Dihydroxybenzenesulfonic acid [b] | C6H6O5S | M-H | 188.9858 | -2.9 | 2.7 | 128.2 | |
| Hydroxybenzoic acid [*, c] | C7H6O3 | M-H | 137.0223 | 7.2 | 2.7 | 130.4 | |
| Benzyl Alcohol [*, c] | C7H8O | M-H | 107.0498 | -4.3 | 3.8 | 119.8 | -0.7 |
| 5-(Phenyl)valeric acid-sulfate [*, c] | C11H14O6S | M+H | 275.0565 | 7.0 | 5.1 | 156.1 | |
| 2-Hydroxy-2-phenylacetic acid [*, c] | C8H8O3 | M+H | 153.0543 | -2.4 | 3.1 | 140.6 | |
| Phenylacetic Acid [*, b] | C8H8O2 | M+H | 137.0593 | -3.3 | 2.4 | 120.5 | -0.2 |
| Phenylethanol [c] | C8H10O | M-H | 121.0652 | -5.9 | 3.6 | 149.9 | |
| Tyramine [c] | C8H11NO | M+H | 138.0903 | -7.8 | 1.7 | 127.4 | |
| Phenethylamine [*, b] | C8H11N | M+H | 122.0960 | -3.8 | 1 | 123.4 | |
| Benzene-1,3,5-triol [c] | C6H6O3 | M-H | 125.0237 | -3.0 | 1.2 | 136.4 | |
| Benzene-1,3-diol [c] | C6H6O2 | M-H | 109.0291 | -3.5 | 2.7 | 132.1 | |
| Phenylacetaldehyde [b] | C8H8O | M+H | 121.0642 | -4.5 | 3.8 | 122.1 | 1.8 |
| 4-Hydroxybenzylamine [c] | C7H9NO | M+H | 124.0753 | -3 | 0.7 | 123.2 | 0.4 |
| Phenylmethanamine [*, c] | C7H9N | M+H | 108.0803 | -4 | 0.6 | 118.1 | |
| 3-(4-Hydroxyphenyl)Pyruvate [*, c] | C9H8O4 | M-H2O-H | 161.0235 | -5.3 | 3.8 | 144.2 | |
| 6-Hydroxydopamine [*, c] | C8H11NO3 | M-H2O-H | 150.0556 | -2.4 | 2.7 | 145.1 | |
| Dimethylbenzimidazole [c] | C9H10N2 | M+Cl | 181.0521 | -11.8 | 3.3 | 377.3 | |
| Glutamine [*] | C5H10N2O3 | M-H2O-H | 127.0505 | -5.8 | 1.2 | 145.9 | |
| 3,4-Dihydroxy-L-Phenylalanine [c] | C9H11NO4 | M-H2O-H | 178.0499 | -5.6 | 3.4 | 142.9 | |
| 6-acetamido-3-aminohexanoic acid [c] | C8H16N2O3 | M+Na | 211.1051 | -1.3 | 2.7 | 141 | |
| Arginine [c] | C6H14N4O2 | M+Cl | 209.0802 | -4.9 | 3.9 | 158.7 | |
| Phenylalanine [c] | C9H11NO2 | M+Na | 188.0696 | 8.2 | 2.5 | 136.5 | |
| 3-Methoxy-L-Tyrosine [c] | C10H13NO4 | M-H2O-H | 192.0654 | -5.8 | 4.0 | 145.2 | |
| 4-Hydroxy-L-Phenylglycine [*, c] | C8H9NO3 | M-H2O-H | 148.0397 | -4.5 | 2.7 | 127.9 | |
| Dihydrodipicolinic acid [b] | C7H7NO4 | M+ACN+H | 211.0711 | -1.4 | 2.3 | 147 | |
| N,N-Dimethylarginine [c] | C8H18N4O2 | M+H | 203.1497 | -2.7 | 0.5 | 144.4 | |
| N5-(1-Carboxyethyl)-ornithine [c] | C8H16N2O4 | M+Na | 227.1000 | -1 | 2.7 | 147.6 | |
| N-Acetylleucine [c] | C8H15NO3 | M+Na | 196.0951 | 3.7 | 0.9 | 143.3 | -3.4 |
| N-Acetyl-L-lysine [c] | C8H16N2O3 | M+Na | 211.1052 | -0.5 | 3.6 | 145.4 | 2.3 |
| O-Acetyl-L-homoserine [b] | C6H11NO4 | M+Na | 184.0587 | 4.1 | 3.3 | 136.7 | |
| Phosphocreatine [*] | C4H10N3O5P | M+Cl | 246.0056 | 1.9 | 2.6 | 148.9 | |
| Tilarginine acetate [*] | C9H20N4O4 | M+Cl | 283.1170 | -3.6 | 5.0 | 165.1 | |
| 4-Methylene-L-glutamate [c] | C6H9NO4 | M+Na | 182.0419 | -3.2 | 2.7 | 131.1 | |
| Glutarate [c] | C5H8O4 | M-H2O-H | 113.0238 | -4.9 | 3.6 | 187.0 | |
| Citric Acid [c] | C6H8O7 | M+Cl | 226.9978 | 7.3 | 3.6 | 145.0 | |
| Cinnamic acid [*, c] | C9H8O2 | M+H | 149.0592 | -3.4 | 4.1 | 126.4 | 1.2 |
| 4′-methoxycinnamic acid [b] | C10H10O3 | M+H | 179.0685 | 10.0 | 4.1 | 134.1 | |
| Cinnamic acid-sulfate [*, b] | C9H8O6S | M-H | 242.9955 | 1.2 | 3.2 | 142.7 | |
| Dimethoxycinnamic acid [b] | C11H12O4 | M+H | 209.0797 | 5.4 | 4.2 | 145.6 | |
| 3′-Hydroxycinnamic acid-4′-sulfate [b] | C9H8O7S | M+H | 261.0069 | -2.3 | 3.7 | 156.7 | |
| 4′-Hydroxy-3′-methoxycinnamic acid [b] | C10H10O4 | M+H | 195.0641 | -5.8 | 3.1 | 140.4 | -0.8 |
| 3′-Hydroxycinnamic acid [c] | C9H8O3 | M+H | 165.0540 | -4 | 3.7 | 129.4 | 2.2 |
| 4′-Hydroxycinnamic acid [c] | C9H8O3 | M+H | 165.0536 | -6.5 | 3.4 | 128 | 0.8 |
| 7-Hydroxy-3-(4-methoxyphenyl)-4-propylcoumarin [c] | C19H18O4 | M-H2O-H | 291.1025 | -0.6 | 2.9 | 166.4 | |
| 3,5-diethyl-2-methylpyrazine [b] | C9H14N2 | M+H | 151.1218 | 11.1 | 2.6 | 131.9 | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 5-Acetylamino-6-formylamino-3-methyluracil [b] | C8H10N4O4 | M+H | 227.0772 | 1.3 | 1.2 | 140.2 | |
| L-Ascorbic acid 2-glucoside [b] | C12H18O11 | M+H | 339.0915 | -2 | 2.5 | 169.6 | |
| 3-Hydroxy-4,5-dimethyl-2(5H)-furanone [c] | C6H8O3 | M+H | 129.0540 | -4.7 | 2 | 122.6 | |
| Sorbate [*, a] | C6H8O2 | M-H2O-H | 93.0342 | -3.5 | 2.7 | 137.9 | |
| Ethylmalonic Acid [c] | C5H8O4 | M-H2O-H | 113.0240 | -3.4 | 3.0 | 190.5 | |
| 5-(Hydroxyphenyl)valeric acid [*, c] | C11H14O3 | M+H | 195.1007 | 4.7 | 4.9 | 140.4 | |
| Hydnocarpic acid [b] | C16H28O2 | M+Na | 275.1992 | 4 | 5 | 167.1 | |
| Hydroxy-methylsuccinic acid [*] | C5H8O5 | M+H | 149.0447 | 1.8 | 2.1 | 127.8 | |
| Oxo-pentenoic acid [c] | C5H6O3 | M-H | 113.0239 | -4.9 | 4.4 | 190.5 | |
| Methyl Jasmonate [c] | C13H20O3 | M+Cl | 259.1110 | 1.4 | 3.1 | 156.3 | |
| Sn-Glycero-3-Phosphocholine [c] | C8H21NO6P+ | M-H2O-H | 239.0902 | -10.2 | 3.8 | 157.2 | |
| O-Methyl-(epi)catechin-sulfate [*, b] | C16H16O9S | M-H | 383.0409 | 5.7 | 3.7 | 191.9 | |
| 1,3,7-Trimethyluric acid [*, b] | C8H10N4O3 | M+H | 211.0817 | -4.2 | 3 | 141 | |
| 1,3-Dimethyluric acid [*, a] | C7H8N4O3 | M+H | 197.0667 | -1.1 | 2.4 | 135.9 | |
| 1,7-Dimethyluric acid [*, a] | C7H8N4O3 | M+H | 197.0668 | -0.6 | 2.7 | 137.4 | |
| 1,7-Dimethyluric acid [*, a] | C7H8N4O3 | M-H | 195.0518 | -3.0 | 2.7 | 134.1 | |
| 1-Methyluric acid [*, a] | C6H6N4O3 | M-H | 181.0363 | -2.5 | 1.8 | 128.7 | |
| 3,7-Dimethyluric acid [*, a] | C7H8N4O3 | M-H | 195.0516 | -4.1 | 2.1 | 134.1 | |
| 3,7-Dimethyluric acid [*, a] | C7H8N4O3 | M+H | 197.0664 | -2.9 | 2.1 | 133 | |
| 3-Methyluric acid [a] | C6H6N4O3 | M-H | 181.0362 | -2.9 | 1.5 | 128.7 | |
| 3-Methyluric acid [a] | C6H6N4O3 | M+H | 183.0510 | -1.2 | 1.5 | 131 | |
| 3-Methylxanthine [*, b] | C6H6N4O2 | M+H | 167.0563 | -0.4 | 2.1 | 129.2 | |
| 7-Methylxanthine [*, b] | C6H6N4O2 | M+H | 167.0563 | -0.6 | 1.9 | 129.2 | |
| 1-Methylxanthine [b] | C6H6N4O2 | M+H | 167.0553 | 1.9 | 2.5 | 130.7 | |
| Caffeine [*, a] | C8H10N4O2 | M+H | 195.0875 | -0.8 | 3.4 | 137.5 | 0.5 |
| Paraxanthine [*, a] | C7H8N4O2 | M+H | 181.0719 | -0.6 | 2.5 | 131.1 | 0.2 |
| Theobromine [*, a] | C7H8N4O2 | M+H | 181.0716 | -2.2 | 0.7 | 132.6 | |
| Theophylline [*, a] | C7H8N4O2 | M-H | 179.0568 | -3.4 | 2.8 | 130.3 | 0.3 |
| Theophylline [*, a] | C7H8N4O2 | M+H | 181.0719 | -0.5 | 2.8 | 132.6 | -1.1 |
| Uric acid [*, c] | C5H4N4O3 | M-H | 167.0206 | -2.8 | 0.6 | 123.5 | 0.7 |
| Indole [c] | C8H7N | M+Cl | 152.0261 | -10.2 | 1.4 | 153.0 | |
| 3-Methylindolepyruvate [*, b] | C12H11NO3 | M+H | 218.0803 | -4 | 1.4 | 146.6 | |
| Indole-3-methylethanoate [c] | C11H11NO2 | M+H | 190.0852 | -5.4 | 2.1 | 139.3 | 1.4 |
| 5-Hydroxy-L-Tryptophan [c] | C11H12N2O3 | M-H2O-H | 219.0758 | -3.7 | 1.5 | 148.5 | -1.4 |
| Hydroxy-oxovaleric acid [c] | C5H8O4 | M-H2O-H | 113.0240 | -3.2 | 3.4 | 179.9 | |
| 2-Oxoadipate [c] | C6H8O5 | M-H2O-H | 141.0186 | -4.5 | 0.7 | 139.5 | |
| Kahweol oxide glucuronide [c] | C26H32O10 | M-H | 503.1917 | 1.0 | 5.0 | 215.0 | |
| Kahweol oxide glucuronide [*, b] | C26H32O10 | M+H | 505.2053 | 3.0 | 5.0 | 210.9 | |
| ((17-Oxoestra-1,3,5(10)-trien-3-yl)oxy)acetic acid [*, b] | C20H24O4 | M+H | 329.1746 | -0.4 | 5 | 171.5 | |
| 2-Deoxyinosose [*, b] | C6H10O5 | M+H | 163.0598 | -1.7 | 2.5 | 129.5 | |
| 4-Oxocyclohexanecarboxylic acid [*, b] | C7H10O3 | M-H2O-H | 123.0445 | -4.9 | 3.4 | 138.3 | |
| Guanidino-butanol [b] | C5H13N3O | M+Na | 154.0952 | 1.0 | 0.6 | 131.7 | |
| Hydroxy-(indol-yl)ethanamine [c] | C10H12N2O | M-H2O-H | 157.0762 | -4.9 | 1.5 | 146.1 | |
| Trigonelline [*, b] | C7H7NO2 | M+H | 138.0548 | -0.9 | 0.5 | 124.5 | |
| Trihydroxy-5alpha-cholan-24-yl sulfate [*, c] | C24H42O7S | M+Cl | 509.2356 | 2.3 | 4.8 | 218.3 | |
| 5-(Phenyl)-γ-valerolactone-sulfate [*,b] | C11H12O6S | M-H | 271.0271 | -0.1 | 4.0 | 157.5 | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 4-Methylbenzenediol-sulfate [*] | C7H8O6S | M-H | 218.9942 | 7.4 | 2.4 | 139.1 | |
| 5-(Dihydroxyphenyl)-γ-valerolactone-sulfate (3',4',5') [*, c] | C11H12O8S | M-H | 303.0149 | -6.6 | 3.1 | 164.4 | |
| Benzoic acid-4-sulfate [*,c] | C7H6O6S | M+H | 218.9962 | -1.8 | 3.3 | 143.5 | |
| 5-(Hydroxyphenyl)-γ-valerolactone-sulfate [*, b] | C11H12O7S | M-H | 287.0222 | -0.7 | 3.8 | 160.1 | |
| 5-(Hydroxyphenyl)-γ-valerolactone-sulfate [b] | C11H12O7S | M-H | 287.0207 | 3.1 | 4.6 | 155.3 | |
| 5-(Methoxyphenyl)-γ-valerolactone-sulfate | C12H14O7S | M-H | 301.0366 | -3.5 | 3.9 | 169.4 | |
| 3-Propanoic acid-3'-sulfate [b] | C9H10O6S | M+H | 247.0238 | 13.3 | 2.4 | 155.8 | |
| 4-Methylphenol-2-sulfate [*, b] | C7H8O5S | M-H | 203.0012 | -1.5 | 3.4 | 133.6 | |
| 3-methoxybenzaldehyde-4-sulfate [*, b] | C8H8O6S | M-H | 230.9958 | 0.0 | 3.3 | 141.6 | |
| 5-Acetylamino-6-amino-3-methyluracil [*, b] | C7H10N4O3 | M-H | 197.0662 | 3.8 | 0.8 | 140.2 | |
| 3-Dehydroshikimate [*] | C7H8O5 | M-H2O-H | 153.0183 | -6.2 | 2.4 | 141.6 | |
| Pantothenic acid [b] | C9H17NO5 | M+H | 220.1179 | -0.3 | 1.2 | 146.5 | |
| D-Sorbitol [c] | C6H14O6 | M+Cl | 217.0478 | -3.4 | 5.1 | 159.9 | |
| Galactitol [*, c] | C6H14O6 | M+Na | 205.0669 | -7.5 | 0.5 | 139.8 | 0.0 |
| 1-Ribosylnicotinamide [b] | C11H16N2O5 | M+H | 257.1110 | 8.7 | 2.4 | 155.3 | |
| 5-(Hydroxyphenyl)-γ-valerolactone-glucuronide [*, b] | C17H20O10 | M-H | 383.0974 | 0.3 | 3.6 | 181.8 | |
| D-Glucuronic Acid [c] | C6H10O7 | M+Cl | 229.0112 | -4.4 | 2.3 | 162.5 | |
| Inosine monophosphate [c] | C10H13N4O8P | M-H2O-H | 329.0310 | 4.9 | 4.2 | 162.0 | |
| Mannitol [c] | C6H14O6 | M+Cl | 217.0493 | 5.0 | 4.2 | 158.3 | |
| Raffinose [c] | C18H32O16 | M-H2O-H | 485.1526 | 2.7 | 3.6 | 208.4 | |
| 1-Methyl-2-pyrrolecarboxaldehyde [*, c] | C6H7NO | M+H | 110.0604 | 3.4 | 0.8 | 119.3 | 1.6 |
| 2,6-Dihydroxypyridine [c] | C5H5NO2 | M+H | 112.0388 | -4.8 | 0.8 | 119 | |
| Carnosine [c] | C9H14N4O3 | M+Cl | 261.0754 | -2.6 | 5.2 | 159.5 | |
| Phenylacetic acid isomer [*, c] | C8H8O2 | M-H | 135.0448 | -2.9 | 3.4 | 127.5 | 4.4 |
| 3,4-Dihydroxyphenylacetic acid [c] | C8H8O4 | M-H | 167.0340 | -5.7 | 4.4 | 128.0 | 0.8 |
| 3-Hydroxymethyl-phenol [*, c] | C7H8O2 | M+Na | 147.0431 | 11.6 | 3.7 | 125.1 | 1.9 |
| 5-(3',4'-Dihydroxyphenyl)-γ-valerolactone [*,b] | C11H12O4 | M-H | 207.0646 | 2.8 | 3.6 | 182.1 | |
| Dopamine [*, c] | C8H11NO2 | M-H2O-H | 134.0601 | -6.5 | 3.4 | 141.8 | |
| Benzene-1,2,4-triol [c] | C6H6O3 | M+H | 127.0390 | 0.5 | 3.3 | 122.8 | |
| 2-Hydroxy-2-(4'-hydroxy-3'-methoxyphenyl)acetic acid [c] | C9H10O5 | M-H2O-H | 179.0341 | -4.4 | 4.2 | 146.0 | |
| 3-Methoxytyramine [*, c] | C9H13NO2 | M-H2O-H | 148.0758 | -6.2 | 4.0 | 146.9 | |
| 4-Hydroxy-3-Methoxyphenylglycol [c] | C9H12O4 | M-H | 183.0658 | -2.5 | 2.4 | 149.0 | |
| DL-Normetanephrine [*, c] | C9H13NO3 | M-H2O-H | 164.0711 | -3.1 | 3.8 | 152.0 | |
| 5-(Hydroxyphenyl)-γ-valerolactone [*, b] | C11H12O3 | M-H | 191.0699 | 1.9 | 4.0 | 158.2 | |
| 3-(2-Hydroxyphenyl)propanoic acid [*, c] | C9H10O3 | M+H | 167.0694 | -5.4 | 4.2 | 129.2 | -2.6 |
| 3-(3'-Hydroxy-4'-methoxyphenyl)propanoic acid (Dihydroisoferulic acid) [c] | C10H12O4 | M+H | 197.0795 | 7.0 | 3.7 | 188.2 | |
| 3-(3'-Methoxy-4'-hydroxyphenyl)propanoic acid (Dihydroferulic acid) [b] | C10H12O4 | M+H | 197.0798 | 5.2 | 3.4 | 188.2 | |
| 2-Hydroxy-3-(4'-hydroxyphenyl)propanoic acid [b] | C9H10O4 | M-H2O-H | 163.0392 | -4.5 | 3.7 | 152.1 | |
| Farnesyl Diphosphate [c] | C15H28O7P2 | M-H | 381.1234 | -0.8 | 4.1 | 181.9 | |
| 2-Aminoadenosine [*] | C10H14N6O4 | M-H | 281.1009 | 1.8 | 4.6 | 195.3 | |
| Inosine [*] | C10H12N4O5 | M+Cl | 303.0521 | 7.0 | 4.5 | 162.8 | |
| Deoxyadenosine [c] | C10H13N5O3 | M-H | 250.0917 | -11.5 | 1.7 | 155.1 | 1.2 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Deoxyadenosine [c] | C10H13N5O3 | M+H | 252.1070 | -8.6 | 1.7 | 155.6 | 0.8 |
| Deoxyadenosine monophosphate [c] | C10H14N5O6P | M-H | 330.0627 | 5.6 | 3.6 | 170.1 | 1.9 |
| Nicotinamide Mononucleotide [c] | C11H15N2O8P | M+Cl | 369.0256 | -1.1 | 3.4 | 172.3 | |
| 5-Hydroxy-6-methylnicotinic acid [*, b] | C7H7NO3 | M+H | 154.0495 | -2.2 | 2.3 | 127.4 | |
| 6-amino nicotinamide [*, b] | C6H7N3O | M+H | 138.0645 | 12.1 | 2.5 | 128.8 | |
| N,N-Diethylnicotinamide [b] | C10H14N2O | M+H | 179.1165 | 7.8 | 1.6 | 135.6 | |
| 2-Hydroxypyridine [*, c] | C5H5NO | M+H | 96.0441 | -3.3 | 0.7 | 117.1 | 3.2 |
| 2,3-Dihydro-5-(3-hydroxypropanoyl)-1H-pyrrolizine [b] | C10H13NO2 | M+H | 180.1000 | -10.9 | 1.1 | 137 | |
| estriol 3-glucuronide [*, b] | C24H32O9 | M+Na | 487.1933 | -1.2 | 5 | 211.3 | |
| Taurocholic Acid [c] | C26H45NO7S | M-H2O-H | 496.2701 | -7.3 | 5.7 | 213.4 | |
| Estriol [*, b] | C18H24O3 | M+Na | 311.1625 | 2.6 | 4.9 | 167.3 | |
| 2-methoxy-17beta-estradiol 3-glucosiduronic acid [b] | C25H34O9 | M+H | 479.2274 | -0.2 | 4.3 | 206.2 | |
| 3-O-(Carboxymethyl)estrone [*, b] | C20H24O4 | M+H | 329.1743 | -1.4 | 4.1 | 171.5 | |
| Progesterone [c] | C21H30O2 | M+Cl | 349.1935 | -1.5 | 4.4 | 181.1 | |

Nomenclature for dietary (poly)phenolic catabolites is in accordance with the recent recommendations proposed by Kay and co-authors (Kay et al., 2020), nomenclature for the other metabolites is in accordance with the common names present in the DBs used (mainly Metlin, PubChem and HMDB).

[*] indicates metabolites common to both ANOVA *p*-value<0.01 and VIP>1.2.

[a] identified metabolites (level I).

[b] putatively identified metabolites (level II).

[c] putatively characterized metabolites (level III)

RT is the Retention Time measured in minutes.

CCS is the observed Collision Cross Section value measured in $Å^2$.

dCCS is the deviation of the observed CCS value from the expected one, reported in $Å^2$.

**Table S3.** Results from the pathway analysis, performed with all the significant metabolites. In particular, the main metabolite routes perturbed by coffee consumption, the number of compounds matched on the total number of compounds in the pathway and the pathway impact value calculated from pathway topology analysis are reported.

| Pathway Name | Match Status | Impact |
|---|---|---|
| Caffeine metabolism | 09/10 | 1 |
| Tyrosine metabolism | 08/42 | 0.39 |
| Phenylalanine metabolism | 04/10 | 0.74 |
| Phenylalanine, tyrosine and tryptophan biosynthesis | 02/04 | 0.5 |
| Arginine biosynthesis | 02/14 | 0.08 |
| Galactose metabolism | 03/27 | 0.03 |
| Purine metabolism | 05/65 | 0.01 |
| D-Glutamine and D-glutamate metabolism | 01/06 | 0 |
| Nitrogen metabolism | 01/06 | 0 |

| | | |
|---|---|---|
| Taurine and hypotaurine metabolism | 01/08 | 0 |
| Ascorbate and aldarate metabolism | 01/08 | 0.5 |
| Ubiquinone and other terpenoid-quinone biosynthesis | 01/09 | 1 |
| Alanine, aspartate and glutamate metabolism | 02/28 | 0.11 |
| Aminoacyl-tRNA biosynthesis | 03/48 | 0 |
| Glyoxylate and dicarboxylate metabolism | 02/32 | 0.03 |
| Nicotinate and nicotinamide metabolism | 01/15 | 0.03 |
| Histidine metabolism | 01/16 | 0.09 |
| Arginine and proline metabolism | 02/38 | 0.06 |
| Pentose and glucuronate interconversions | 01/18 | 0.13 |
| Terpenoid backbone biosynthesis | 01/18 | 0.18 |
| Pantothenate and CoA biosynthesis | 01/19 | 0.01 |
| Tryptophan metabolism | 02/41 | 0.13 |
| Fructose and mannose metabolism | 01/20 | 0.03 |
| Citrate cycle (TCA cycle) | 01/20 | 0.09 |
| beta-Alanine metabolism | 01/21 | 0.06 |
| Lysine degradation | 01/25 | 0.14 |
| Inositol phosphate metabolism | 01/30 | 0 |
| Steroid hormone biosynthesis | 03/85 | 0.07 |
| Pyrimidine metabolism | 01/39 | 0 |
| Steroid biosynthesis | 01/42 | 0 |
| Primary bile acid biosynthesis | 01/46 | 0.01 |

## References


1 I. A. Ludwig, M. N. Clifford, M. E. J. Lean, H. Ashihara and A. Crozier, Coffee: biochemistry and potential impact on health, *Food Funct.*, 2014, **5**, 1695–1717.

2 S. Bidel and J. Tuomilehto, The emerging health benefits of coffee with an emphasis on type 2 diabetes and cardiovascular disease, *Eur. Endocrinol.*, 2013, **9**, 99–106.

3 M. Ding, S. N. Bhupathiraju, A. Satija, R. M. Van Dam and F. B. Hu, Long-term coffee consumption and risk of cardiovascular disease: A systematic review and a dose-response meta-analysis of prospective cohort studies, *Circulation*, 2014, **129**, 643–659.

4 A. Elbaz, L. Carcaillon, S. Kab and F. Moisan, Epidemiology of Parkinson's disease, *Rev. Neurol. (Paris).*, 2016, **172**, 14–26.

5 R. Huxley, C. M. Y. Lee, F. Barzi, L. Timmermeister, S. Czernichow, V. Perkovic, D. E. Grobbee, D. Batty and M. Woodward, Coffee, decaffeinated coffee, and tea consumption in relation to incident type 2 diabetes mellitus: A systematic review with meta-analysis, *Arch. Intern. Med.*, 2009, **169**, 2053–2063.

6 S. C. Larsson, Coffee, Tea, and Cocoa and Risk of Stroke, *Stroke*, 2014, **45**, 309–314.

7 S. Malerba, F. Turati, C. Galeone, C. Pelucchi, F. Verga, C. La Vecchia and A. Tavani, A meta-analysis of prospective studies of coffee consumption and mortality for all causes, cancers and cardiovascular diseases, *Eur. J. Epidemiol.*, 2013, **28**, 527–539.

8 L. Wu, D. Sun and Y. He, Coffee intake and the incident risk of cognitive disorders: A dose–response meta-analysis of nine prospective cohort studies, *Clin. Nutr.*, 2017, **36**, 730–736.

9 G. Grosso, J. Godos, F. Galvano and E. L. Giovannucci, Coffee, Caffeine, and Health Outcomes: An Umbrella Review, *Annu. Rev. Nutr.*, 2017, **37**, 131–156.

10 R. Poole, O. J. Kennedy, P. Roderick, J. A. Fallowfield, P. C. Hayes and J. Parkes, Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes, *BMJ*, 2017, **359**, j5024.

11 I. A. Ludwig, P. Mena, L. Calani, C. Cid, D. Del Rio, M. E. J. Lean and A. Crozier, Variations in caffeine and chlorogenic acid contents of coffees: what are we drinking?, *Food Funct.*, 2014, **5**, 1718–1726.

12 EFSA, Scientific Opinion on the substantiation of a health claim related to cocoa flavanols and maintenance of normal endothelium-dependent vasodilation pursuant to Article 13(5) of Regulation (EC) No 1924/2006, *EFSA J.*, 2012, **10**, 1–21.

13 J. Kim, J. Kim, J. Shim, C. Y. Lee, K. W. Lee and H. J. Lee, Cocoa Phytochemicals: Recent Advances in Molecular Mechanisms on Health, *Crit. Rev. Food Sci. Nutr.*, 2014, **54**, 1458–1472.

14 R. Sansone, A. Rodriguez-Mateos, J. Heuel, D. Falk, D. Schuler, R. Wagstaff, G. G. C. Kuhnle, J. P. E. Spencer, H. Schroeter, M. W. Merx, M. Kelm and C. Heiss, Cocoa flavanol intake improves endothelial function and Framingham Risk Score in healthy men and women: a randomised, controlled, double-masked trial: the Flaviola Health Study, *Br. J. Nutr.*, 2015, **114**, 1246–1255.

15 P. Mena, M. Tassotti, D. Martini, A. Rosi, F. Brighenti and D. Del Rio, The Pocket-4-Life project, bioavailability and beneficial properties of the bioactive compounds of espresso coffee and cocoa-based confectionery containing coffee: Study protocol for a randomized cross-over trial, *Trials*, 2017, **18**, 1–11.

16 A. Scalbert, L. Brennan, C. Manach, C. Andres-Lacueva, L. O. Dragsted, J. Draper, S. M. Rappaport, J. J. van der Hooft and D. S. Wishart, The food metabolome: a window over dietary exposure, *Am. J. Clin. Nutr.*, 2014, **99**, 1286–1308.

17 F. Madrid-Gambin, M. Garcia-Aloy, R. Vázquez-Fresno, E. Vegas-Lozano, M. C. R. de Villa Jubany, K. Misawa, T. Hase, A. Shimotoyodome and C. Andres-Lacueva, Impact of chlorogenic acids from coffee on urine metabolome in healthy human subjects, *Food Res. Int.*, 2016, **89**, 1064–1070.

18 C. C. J. R. Michielsen, E. Almanza-Aguilera, E. M. Brouwer-Brolsma, M. Urpi-Sarda and L. A. Afman, Biomarkers of food intake for cocoa and liquorice (products): a systematic review, *Genes Nutr.*, 2018, **13**, 22.

19 L. H. Münger, A. Trimigno, G. Picone, C. Freiburghaus, G. Pimentel, K. J. Burton, F. P. Pralong, N. Vionnet, F. Capozzi, R. Badertscher and G. Vergères, Identification of Urinary Food Intake Biomarkers for Milk, Cheese, and Soy-Based Drink by Untargeted GC-MS and NMR in Healthy Humans., *J. Proteome Res.*, 2017, **16**, 3321–3335.

20 J. A. Rothwell, P. Keski-Rahkonen, N. Robinot, N. Assi, C. Casagrande, M. Jenab, P. Ferrari, M. C. Boutron-Ruault, Y. Mahamat-Saleh, F. R. Mancini, H. Boeing, V. Katzke, T. Kühn, K. Niforou, A. Trichopoulou, E. Valanou, V. Krogh, A. Mattiello, D. Palli, C. Sacerdote, R. Tumino and A. Scalbert, A Metabolomic Study of Biomarkers of Habitual Coffee Intake in Four European Countries, *Mol. Nutr. Food Res.*, 2019, **63**, 1–21.

21 N. Vázquez-Manjarrez, C. H. Weinert, M. M. Ulaszewska, C. I. Mack, P. Micheau, M. Pétéra, S. Durand, E. Pujos-Guillot, B. Egert, F. Mattivi, A. Bub, L. O. Dragsted, S. E. Kulling and C. Manach, Discovery and Validation of Banana Intake Biomarkers Using Untargeted Metabolomics in Human Intervention and Cross-sectional Studies, *J. Nutr.*, 2019, **149**, 1701–1713.

22 H. Gibbons, A. O'Gorman and L. Brennan, Metabolomics as a tool in nutritional research, *Curr. Opin. Lipidol.*, 2015, **26**, 30–34.

23 D. S. Wishart, Metabolomics: applications to food science and nutrition research, *Trends Food Sci. Technol.*, 2008, **19**, 482–493.

24 M. C. Cornelis, I. Erlund, G. A. Michelotti, C. Herder, J. A. Westerhuis and J. Tuomilehto, Metabolomic response to coffee consumption: application to a three-stage clinical trial, *J. Intern. Med.*, 2018, **283**, 544–557.

25 G. Peron, D. Santarossa, D. Voinovich, S. Dall'Acqua and S. Sut, Urine metabolomics shows an induction of fatty acids metabolism in healthy adult volunteers after supplementation with green coffee (*Coffea robusta* L.) bean extract, *Phytomedicine*, 2018, **38**, 74–83.

26 L. Shi, C. Brunius, I. Johansson, I. A. Bergdahl, O. Rolandsson, B. van Guelpen, A. Winkvist, K. Hanhineva and R. Landberg, Plasma metabolite biomarkers of boiled and filtered coffee intake and their association with type 2 diabetes risk, *J. Intern. Med.*, 2020, **287**, 405–421.

27 D. Martini, A. Rosi, M. Tassotti, M. Antonini, M. Dall'Asta, L. Bresciani, F. Fantuzzi, V. Spigoni, R. Domínguez-Perles, D. Angelino, C. Ricci, S. Del Pozo-Luengo, P. L. Tornel, F. Scazzina, A. Gil-Izquierdo, A. Dei Cas, F. Brighenti, R. Bonadonna, D. Del Rio and P. Mena, Effect of coffee and cocoa-based confectionery containing coffee on markers of cardiometabolic health: results from the pocket-4-life project, *Eur. J. Nutr.*, 2020, **60**, 1453-1463.

28 E. J. Want, I. D. Wilson, H. Gika, G. Theodoridis, R. S. Plumb, J. Shockcor, E. Holmes and J. K. Nicholson, Global metabolic profiling procedures for urine using UPLC-MS, *Nat. Protoc.*, 2010, **5**, 1005–1018.

29 E. Rodriguez-Suarez, C. Hughes, L. Gethings, K. Giles, J. Wildgoose, M. Stapels, K. E. Fadgen, S. J. Geromanos, J. P.C. Vissers, F. Elortza and J. I. Langridge, An Ion Mobility Assisted Data Independent LC-MS Strategy for the Analysis of Complex Biological Samples, *Curr. Anal. Chem.*, 2013, **9**, 199–211.

30 E. Fahy, S. Subramaniam, R. C. Murphy, M. Nishijima, C. R. H. Raetz, T. Shimizu, F. Spener, G. Van Meer, M. J. O. Wakelam and E. A. Dennis, Update of the LIPID MAPS comprehensive classification system for lipids, *J. Lipid Res.*, 2009, **50**, S9–S14.

31 D. S. Wishart, T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat, E. Dong, S. Bouatra, I. Sinelnikov, D. Arndt, J. Xia, P. Liu, F. Yallou, T. Bjorndahl, R. Perez-Pineiro, R. Eisner, F. Allen, V. Neveu, R. Greiner and A. Scalbert, HMDB 3.0-The Human Metabolome Database in 2013, *Nucleic Acids Res.*, 2013, **41**, 801–807.

32 C. A. Smith, G. O'Maille, E. J. Want, C. Qin, S. A. Trauger, T. R. Brandon, D. E. Custodio, R. Abagyan and G. Siuzdak, METLIN: A metabolite mass spectral database, *Ther. Drug Monit.*, 2005, **27**, 747–751.

33 Z. Zhou, X. Shen, J. Tu and Z. J. Zhu, Large-scale prediction of collision cross-section values for metabolites in ion mobility-mass spectrometry, *Anal. Chem.*, 2016, **88**, 11084–11091.

34 L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reily, J. J. Thaden and M. R. Viant, Proposed minimum reporting standards for chemical analysis, *Metabolomics*, 2007, **3**, 211–221.

35 J. Chong, O. Soufan, C. Li, I. Caraus, S. Li, G. Bourque, D. S. Wishart and J. Xia, MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis, *Nucleic Acids Res.*, 2018, **46**, W486–W494.

36 P. D. Rainville, I. D. Wilson, J. K. Nicholson, G. Issacs, L. Mullin, J. I. Langridge and R. S. Plumb, Ion mobility spectrometry combined with ultra performance liquid chromatography/mass spectrometry for metabolic phenotyping of urine: Effects of column length, gradient

duration and ion mobility spectrometry on metabolite detection, *Anal. Chim. Acta*, 2017, **982**, 1–8.

37 S. Martínez-López, B. Sarriá, G. Baeza, R. Mateos and L. Bravo-Clemente, Pharmacokinetics of caffeine and its metabolites in plasma and urine after consuming a soluble green/roasted coffee blend by healthy subjects, *Food Res. Int.*, 2014, **64**, 125–133.

38 J. A. Rothwell, F. Madrid-Gambin, M. Garcia-Aloy, C. Andres-Lacueva, C. Logue, A. M. Gallagher, C. Mack, S. E. Kulling, Q. Gao, G. Praticò, L. O. Dragsted and A. Scalbert, Biomarkers of intake for coffee, tea, and sweetened beverages, *Genes Nutr.*, 2018, **13**, 15.

39 J. A. Rothwell, Y. Fillâtre, J.-F. Martin, B. Lyan, E. Pujos-Guillot, L. Fezeu, S. Hercberg, B. Comte, P. Galan, M. Touvier and C. Manach, New Biomarkers of Coffee Consumption Identified by the Non-Targeted Metabolomic Profiling of Cohort Study Subjects, *PLoS One*, 2014, **9**, e93474.

40 A. Bento-Silva, V. M. Koistinen, P. Mena, M. R. Bronze, K. Hanhineva, S. Sahlstrøm, V. Kitrytė, S. Moco and A.-M. Aura, Factors affecting intake, metabolism and health benefits of phenolic acids: do we understand individual variability?, *Eur. J. Nutr.*, 2020, **59**, 1275–1293.

41 S. Martínez-López, B. Sarriá, M. Gómez-Juaristi, L. Goya, R. Mateos and L. Bravo-Clemente, 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**, 446–455.

42 G. Gross, E. Jaccaud and A. C. Huggett, Analysis of the content of the diterpenes cafestol and kahweol in coffee brews, *Food Chem. Toxicol.*, 1997, **35**, 547–554.

43 R. Lang, N. Dieminger, A. Beusch, Y. M. Lee, A. Dunkel, B. Suess, T. Skurk, A. Wahl, H. Hauner and T. Hofmann, Bioappearance and pharmacokinetics of bioactives upon coffee consumption, *Anal. Bioanal. Chem.*, 2013, **405**, 8487–8503.

44 P. Mena, L. Bresciani, N. Brindani, I. A. Ludwig, G. Pereira-Caro, D. Angelino, R. Llorach, L. Calani, F. Brighenti, M. N. Clifford, C. I. R. Gill, A. Crozier, C. Curti and D. Del Rio, Phenyl-γ-valerolactones and phenylvaleric acids, the main colonic metabolites of flavan-3-ols: synthesis, analysis, bioavailability, and bioactivity, *Nat. Prod. Rep.*, 2019, **36**, 714–752.

45 D. Bracco, J. M. Ferrarra, M. J. Arnaud, E. Jequier and Y. Schutz, Effects of caffeine on energy metabolism, heart rate, and methylxanthine metabolism in lean and obese women, *Am. J. Physiol. - Endocrinol. Metab.*, 1995, **269**, 671–678.

46 A. G. Dulloo, C. A. Geissler, A. Collins and D. S. Miller, Normal caffeine consumption: Influence on thermogenesis and daily energy expenditure in lean and postobese human volunteers, *Am. J. Clin. Nutr.*, 1989, **49**, 44–50.

47 A. Astrup, S. Toubro, S. Cannon, P. Hein, L. Breum and J. Madsen, Caffeine: a double-blind, placebo-controlled study of its thermogenic, metabolic, and cardiovascular effects in healthy volunteers, *Am. J. Clin. Nutr.*, 1990, **51**, 759–767.

48 S. J. Stohs and V. Badmaev, A Review of Natural Stimulant and Non-stimulant Thermogenic Agents, *Phyther. Res.*, 2016, **30**, 732–740.

49 S. Takahashi, K. Saito, H. Jia and H. Kato, An integrated multi-omics study revealed metabolic alterations underlying the effects of coffee consumption, *PLoS One*, 2014, **9**, 1–8.

50 F. Kulapichitr, C. Borompichaichartkul, I. Suppavorasatit and K. R. Cadwallader, Impact of drying process on chemical composition and key aroma components of Arabica coffee, *Food Chem.*, 2019, **291**, 49–58.

51 R. N. Rodionov, E. Oppici, J. Martens-Lobenhoffer, N. Jarzebska, S. Brilloff, D. Burdin, A. Demyanov, A. Kolouschek, J. Leiper, R. Maas, B. Cellini, N. Weiss and S. M. Bode-Böger, A novel pathway for metabolism of the cardiovascular risk factor Homoarginine by alanine:glyoxylate aminotransferase 2, *Sci. Rep.*, 2016, **6**, 1–12.

52 D. Tsikas and G. Wu, Homoarginine, arginine, and relatives: analysis, metabolism, transport, physiology, and pathology, *Amino Acids*, 2015, **47**, 1697–1702.

53 S. Pilz, A. Meinitzer, M. Gaksch, M. Grübler, N. Verheyen, C. Drechsler, B. Hartaigh, F. Lang, I. Alesutan, J. Voelkl, W. März and A. Tomaschitz, Homoarginine in the renal and cardiovascular systems, *Amino Acids*, 2015, **47**, 1703–1713.

54 D. Atzler, E. Schwedhelm and C. Choe, L-Homoarginine and cardiovascular disease, *Curr. Opin. Clin. Nutr. Metab. Care*, 2015, **18**, 83–88.

55 N. Jarzebska, A. A. Mangoni, J. Martens-Lobenhoffer, S. M. Bode-Böger and R. N. Rodionov, The second life of methylarginines as cardiovascular targets, *Int. J. Mol. Sci.*, 2019, **20**, 4592.

56 S. Jacobs, J. Kröger, A. Floegel, H. Boeing, D. Drogan, T. Pischon, A. Fritsche, C. Prehn, J. Adamski, B. Isermann, C. Weikert and M. B. Schulze, Evaluation of various biomarkers as potential mediators of the association between coffee consumption and incident type 2 diabetes in the EPIC-Potsdam Study, *Am. J. Clin. Nutr.*, 2014, **100**, 891–900.

57 L. M. de Lau and M. M. Breteler, Epidemiology of Parkinson's disease, *Lancet Neurol.*, 2006, **5**, 525–535.

58 P. Richette and T. Bardin, Gout, *Lancet*, 2010, **375**, 318–328.

59 A. M. Amorini, A. Petzold, B. Tavazzi, J. Eikelenboom, G. Keir, A. Belli, G. Giovannoni, V. Di Pietro, C. Polman, S. D'Urso, R. Vagnozzi, B. Uitdehaag and G. Lazzarino, Increase of uric acid and purine compounds in biological fluids of multiple sclerosis patients, *Clin. Biochem.*, 2009, **42**, 1001–1006.

60 W. Struck-Lewicka, R. Kaliszan and M. J. Markuszewski, Analysis of urinary nucleosides as potential cancer markers determined using LC-MS technique, *J. Pharm. Biomed. Anal.*, 2014, **101**, 50–57.

61 V. C. Myers and E. L. Wardell, THE INFLUENCE OF THE INGESTION OF METHYL XANTHINES ON THE EXCRETION OF URIC ACID., *J. Biol. Chem.*, 1928, 697–722.

62 R. L. Ferrini and E. Barrett-Connor, Caffeine intake and endogenous sex steroid levels in postmenopausal women: The Rancho Bernardo study, *Am. J. Epidemiol.*, 1996, **144**, 642–644.

63 J. Kotsopoulos, A. H. Eliassen, S. A. Missmer, S. E. Hankinson and S. S. Tworoger, Relationship between caffeine intake and plasma sex hormone concentrations in premenopausal and postmenopausal women, *Cancer*, 2009, **115**, 2765–2774.

64 S. C. Larsson, L. Bergkvist and A. Wolk, Coffee and black tea consumption and risk of breast cancer by estrogen and progesterone receptor status in a Swedish cohort, *Cancer Causes Control*, 2009, **20**, 2039–2044.

65 J. Lucero, B. L. Harlow, R. L. Barbieri, P. Sluss and D. W. Cramer, Early follicular phase hormone levels in relation to patterns of alcohol, tobacco, and coffee use, *Fertil. Steril.*, 2001, **76**, 723–729.

66 J. S. Sisti, S. E. Hankinson, N. E. Caporaso, F. Gu, R. M. Tamimi, B. Rosner, X. Xu, R. Ziegler and A. H. Eliassen, Caffeine, coffee, and tea intake and urinary estrogens and estrogen metabolites in premenopausal women, *Cancer Epidemiol. Biomarkers Prev.*, 2015, **24**, 1174–1183.

67 C. Dinçer, T. Apaydin and D. Gogas Yavuz, Endocrine effects of coffee consumption, *Turkish J. Endocrinol. Metab.*, 2020, **24**, 72–86.

68 C. Doepker, K. Franke, E. Myers, J. J. Goldberger, H. R. Lieberman, C. O'brien, J. Peck, M. Tenenbein, C. Weaver and D. Wikoff, Key findings and implications of a recent systematic review of the potential adverse effects of caffeine consumption in healthy adults, pregnant women, adolescents, and children, *Nutrients*, 2018, **10**, 1536.

69 K. Ito, Hormone Replacement Therapy and Cancers: The Biological Roles of Estrogen and Progestin in Tumorigenesis are Different between the Endometrium and Breast, *Tohoku J. Exp. Med.*, 2007, **212**, 1–12.

70 J. R. Pasqualini, Enzymes involved in the formation and transformation of steroid hormones in the fetal and placental compartments, *J. Steroid Biochem. Mol. Biol.*, 2005, **97**, 401–415.

71 D. Ganmaa, W. C. Willett, T. Y. Li, D. Feskanich, R. M. Van Dam, E. Lopez-Garcia, D. J. Hunter and M. D. Holmes, Coffee, tea, caffeine and risk of breast cancer: A 22-year follow-up, *Int. J. Cancer*, 2008, **122**, 2071–2076.

72 J. Li, P. Seibold, J. Chang-Claude, D. Flesch-Janys, J. Liu, K. Czene, K. Humphreys and P. Hall, Coffee consumption modifies risk of estrogen-receptor negative breast cancer, *Breast Cancer Res.*, 2011, **13**, R49.

73 C. D. Kay, M. N. Clifford, P. Mena, G. J. McDougall, C. Andres-Lacueva, A. Cassidy, D. Del Rio, N. Kuhnert, C. Manach, G. Pereira-Caro, A. Rodriguez-Mateos, A. Scalbert, F. Tomás-Barberán, G. Williamson, D. S. Wishart and A. Crozier, Recommendations for standardizing nomenclature for dietary (poly)phenol catabolites, *Am. J. Clin. Nutr.*, 2020, **112**, 1051–1068.

# Chapter 4 – Overall discussion and conclusions

# 4.1. General remarks

In the present doctoral thesis, three specific metabolomics applications have been addressed.

The first application was aimed at evaluating the bioavailability of flavan-3-ols from cranberry in realtion to the absorption, metabolism, and excretion of PVLs and PVAs, the main and specific colonic metabolites of this subclass of flavonoids. Their potential as biomarkers of flavan-3-ol intake and inter-individual variability in their appearance in plasma and urine was also considered (**3.1.** Study 1 – *Kinetic profile and urinary excretion of phenyl-γ-valerolactones upon consumption of cranberry: a dose–response relationship*). Information on PVLs and PVAs production following cranberry intake was limited, despite its relevance for understanding the health effects attributed to these berries [1,2]. Cranberry flavan-3-ols are not only monomeric, but mainly oligomeric and characterized by A-type bonds, known to be less accessible than B-type bonds, which are typical, for example, of oligomeric flavan-3-ols found in cocoa [3]. Within this study, it was shown that consumption of cranberry flavan-3-ols leads to the production of the same colonic metabolites originated after intake of other dietary sources of flavan-3-ols, as cocoa-derived products, wine, and green tea, contributing to the pool of circulating PVLs. Actually, PVLs were mainly found in human biofluids upon consumption of cranberry flavan-3-ols, while PVAs were only found at trace levels in a few samples, reflecting the observations obtained in previous studies [4]. In particular, sulfate and glucuronide conjugated forms of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone were the main circulating and excreted metabolites. Conjugated forms of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone reached maximum plasma concentration at about 4–6 h post consumption, in accordance with their colonic origin, and then decreased, going back to baseline levels at 24 h. Differently, derivatives of 5-(3'-hydroxyphenyl)-γ-valerolactone and 5-(4'-hydroxyphenyl)-γ-valerolactone showed an earlier $T_{max}$ with not so clear kinetic profiles, but this could have been likely due to the low circulating amounts and the extensive variability recorded for these minor metabolites.

Plasma and urinary concentrations of the sum of the metabolites increased in relation to the amounts of cranberry flavan-3-ols provided by the intervention drink, showing a clear and linear dose-dependent relationship with the intake, with glucuronide and sulfate derivatives of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone being the key metabolites driving the overall correlation. This result suggested that 5-(3',4'-dihydroxyphenyl)-γ-valerolactone, in its sulfate and glucuronide conjugated forms, may serve as biomarker of intake for cranberry flavan-3-ols. Actually, these

metabolites have already been proposed as biomarkers of flavan-3-ol intake in a large epidemiological study [5], but information on cranberry was missing. In addition, our results indicated that sulfate and glucuronide derivatives of 5-(3',4'-dihydroxyphenyl)-γ-valerolactone could be used in clinical trials to check compliance with interventions made of standardized cranberry products. Glucuronidation seemed to be the most favorable phase II metabolic pathway after cranberry juice intake, even though a wide inter-individual variability existed. A high inter-individual variability was also reported in circulating and urinary metabolite levels and, interestingly, some subjects seemed to display a greater efficiency in metabolizing flavan-3-ols and producing PVLs. This might indicate that metabotypes (i.e., subjects presenting different metabolic profiles) of flavan-3-ols colonic metabolites could be present and this might explain the differences in the putative health effects associated to flavan-3-ols.

On the basis of the previous evidence and the results of a preliminary work [6], where a different urinary production of PVLs and HPPs upon green tea flavan-3-ols consumption led to the definition of three putative metabotypes, the second metabolomics application was aimed at elucidating the presence of different metabotypes in the urinary excretion of the main flavan-3-ol colonic metabolites after consumption of cranberry products in two diverse experimental settings, and at assessing the impact of the statistical technique used for metabotyping (**3.2.** Study 2 – *Metabotypes of flavan-3-ol colonic metabolites after cranberry intake: elucidation and statistical approaches*). A methodological workflow for metabotype definition has been proposed, trying to fill in the present gap on how to handle the inter-individual variability in the production of phenolic metabolites to define metabotypes in those cases where the production does not result dichotomic (production vs. non-production of specific phenolic metabolites) [7-9], and all the subjects produce instead all the phenolic metabolites of a catabolic pathway, but in different proportions. The methodological workflow proposed consisted in the subsequent application of different multivariate statistics – PCA, cluster analysis and PLS-DA -, followed by univariate analyses to confirm the results from PLS-DA models and validate the robustness of the methodological approach. Data pre-treatment played a major role on resulting PCA models, as non-transformed, centered, and UV-scaled data were key to unravel metabolic patterns based on colonic metabolism, while other pre-treatments highlighted differences in phase II metabolism. Regarding clustering, k-means and a final consensus algorithm highlighted differences in the overall production of PVLs and HPPs, leading to quantitative-based models whereby the distribution of the clusters was due to the amount of metabolites excreted (high vs. low). Differently, the expectation-maximization algorithm and clustering according to

principal component scores yielded well-defined metabotypes characterized by quali-quantitative differences in the excretion of the colonic metabolites. The true physiological relevance of each metabotyping model will be related to the application of these inter-individual differences to explore their potential impact on the biological activity attributed to flavan-3-ols. For instance, *in vitro* findings support the biological effects of flavan-3-ol colonic metabolites against uropathogenic *Escherichia coli* adherence to uroepithelial cells [10,11], while it is well-known that human studies administering cranberry flavan-3-ols to prevent urinary tract infections have reported conflicting results [12,13]. This might be due to different profiles of excreted metabolites, exerting different biological effects. In this sense, clustering subjects according to their urinary metabotype of flavan-3-ol colonic metabolites may provide new insights in the actual effect of flavan-3-ols on UTI prevention, not only through cranberries but potentially also from other flavan-3-ol food sources like cocoa, wine, pome fruits, other berries, and nuts.

The results of this work are not conclusive and should be validated in larger datasets involving a higher number of participants, more phenolic metabolites from the flavan-3-ol metabolic pathway, and different sources of flavan-3-ols. However, this work represents an additional step toward the understanding of the inter-individual variability in flavan-3-ol metabolism, which is key to further investigate its impact on the observed health effects attributed to this major subclass of (poly)phenols.

Finally, the last metabolomics application was aimed at investigating the changes occurring at metabolomic level and the impact on metabolic pathways upon daily consumption, for a whole month, of different patterns of coffee (**3.3.** Study 3 – *Metabolomic changes after coffee consumption: new paths on the block*). The adoption of an untargeted approach has allowed the simultaneous identification not only of metabolites specifically associated with coffee consumption, that are principally of exogenous origin, but also of endogenous metabolites, the levels of which were influenced by the different treatments, reflecting the physiological effects of consuming different patterns of coffee. All the pathways that were significantly impacted by coffee intake were upregulated. Caffeine metabolism resulted the principal pathway influenced by coffee consumption, highly affecting urinary metabolome and this was expected, being caffeine the principal bioactive compound found in coffee and, once ingested, rapidly absorbed and metabolized [22]. Besides caffeine metabolism, endogenous metabolic pathways, such as the metabolism and biosynthesis of specific amino acids, were modulated and may, in turn, potentially influence human health. Specifically, an upregulation of phenylalanine biosynthesis and metabolism, arginine

biosynthesis and ubiquinone biosynthesis have been associated to coffee consumption for the first time. Of note, these results should be confirmed by further studies, but the most important aspect is that this study opens the door to new hypotheses about the health effects of coffee consumption, and the possible underlying mechanisms.

## 4.2. Analytical considerations

All the metabolomics applications reported in this doctoral thesis have been performed by using MS, in view of its sensitivity and selectivity, coupled to LC, due to its breadth of coverage of the metabolome [23].

Specifically, ultra-high performance liquid chromatography coupled to a triple quadrupole mass spectrometer has been adopted for the targeted analyses performed in study 1 (**3.1**.) and 2 (**3.2.**), in view of the increased sensitivity and specificity of this technology, yielding lower detection and quantitation limits. The availability of specifically sinthesized analytical standards [24,25] together with a validated method optimized for the analysis of PVLs [24] allowed the absolute quantification of most of the PVLs detected, as well as an accurate semi-quantification of the remaining metabolites (PVLs, PVAs and HPPs). Indeed, the lack of reference compounds for *in vivo* metabolites (as phase II glucuronide, sulfate and methylated metabolites of phenolic compounds) is one of the major issues in bioavailability studies, precluding both level I identification [26] of the metabolites present in biofluids and accuracy in their quantitative estimate [27].

Ultra-high performance liquid chromatography coupled to high resolution MS equipped with ion mobility separation has been employed for the untargeted experiment developed in study 3 (**3.3.**). In this case, the use of ion mobility has added a further dimension of separation, increasing the number of detected features, and a further parameter for molecules description (the collision cross section value, CCS) that has enhanced putative identification of significant features (level II identification[26]) in those cases where analytical standards were not available. Metabolite identification remains one of the most challenging step of the whole metabolomics workflow and, unfortunately, only a small percentage of detected features has been translated into knowledge, limiting biological interpretation. However, as a discovery strategy, the untargeted metabolomics approach adopted in study 3 has generated robust and comprehensive datasets, containing large amounts of information, and, undoubtedly, this strategy has created new hypotheses to be investigated in the future.

Overall, the results contained in this thesis constitute an important piece in the advancement of the research on dietary bioactives and associated health benefits.

# References


1 R. P. Feliciano, A. Boeres, L. Massacessi, G. Istas, M. R. Ventura, C. Nunes dos Santos, C. Heiss and A. Rodriguez-Mateos, Identification and quantification of novel cranberry-derived plasma and urinary (poly)phenols, *Arch. Biochem. Biophys.*, 2016, **599**, 31–41.

2 R. Feliciano, C. Mills, G. Istas, C. Heiss and A. Rodriguez-Mateos, Absorption, metabolism and excretion of cranberry (poly)phenols in humans: a dose response study and assessment of inter-individual variability, *Nutrients*, 2017, **9**, 268.

3 L. Gu, M. A. Kelm, J. F. Hammerstone, G. Beecher, J. Holden, D. Haytowitz, S. Gebhardt and R. L. Prior, Concentrations of proanthocyanidins in common foods and estimations of normal consumption, *J. Nutr.*, 2004, **134**, 613–617.

4 P. Mena, L. Bresciani, N. Brindani, I. A. Ludwig, G. Pereira-Caro, D. Angelino, R. Llorach, L. Calani, F. Brighenti, M. N. Clifford, C. I. R. Gill, A. Crozier, C. Curti and D. Del Rio, Phenyl-γ-valerolactones and phenylvaleric acids, the main colonic metabolites of flavan-3-ols: synthesis, analysis, bioavailability, and bioactivity, *Nat. Prod. Rep.*, 2019, **36**, 714–752.

5 J. I. Ottaviani, R. Fong, J. Kimball, J. L. Ensunsa, A. Britten, D. Lucarelli, R. Luben, P. B. Grace, D. H. Mawson, A. Tym, A. Wierzbicki, K. T. Khaw, H. Schroeter and G. G. C. Kuhnle, Evaluation at scale of microbiome-derived metabolites as biomarker of flavan-3-ol intake in epidemiological studies, *Sci. Rep.*, 2018, **8**, 1–11.

6 P. Mena, I. A. Ludwig, V. B. Tomatis, A. Acharjee, L. Calani, A. Rosi, F. Brighenti, S. Ray, J. L. Griffin, L. J. Bluck and D. Del Rio, Inter-individual variability in the production of flavan-3-ol colonic metabolites: preliminary elucidation of urinary metabotypes, *Eur. J. Nutr.*, 2019, **58**, 1529–1543.

7 M. V. Selma, A. González-Sarrías, J. Salas-Salvadó, C. Andrés-Lacueva, C. Alasalvar, A. Örem, F. A. Tomás-Barberán and J. C. Espín, The gut microbiota metabolism of pomegranate or walnut ellagitannins yields two urolithin-metabotypes that correlate with cardiometabolic risk biomarkers: Comparison between normoweight, overweight-obesity and metabolic syndrome, *Clin. Nutr.*, 2018, **37**, 897–905.

8 A. González-Sarrías, R. García-Villalba, M. Romo-Vaquero, C. Alasalvar, A. Örem, P. Zafrilla, F. A. Tomás-Barberán, M. V. Selma and J. C. Espín, Clustering according to urolithin

metabotype explains the interindividual variability in the improvement of cardiovascular risk biomarkers in overweight-obese individuals consuming pomegranate: A randomized clinical trial, *Mol. Nutr. Food Res.*, 2017, **61**, 1–14.

9 S. Hazim, P. J. Curtis, M. Y. Schär, L. M. Ostertag, C. D. Kay, A.-M. Minihane and A. Cassidy, Acute benefits of the microbial-derived isoflavone metabolite equol on arterial stiffness in men prospectively recruited according to equol producer phenotype: a double-blind randomized controlled trial, *Am. J. Clin. Nutr.*, 2016, **103**, 694–702.

10 D. González de Llano, A. Esteban-Fernández, F. Sánchez-Patán, P. Martínlvarez, M. Moreno-Arribas and B. Bartolomé, Anti-Adhesive Activity of Cranberry Phenolic Compounds and Their Microbial-Derived Metabolites against Uropathogenic Escherichia coli in Bladder Epithelial Cell Cultures, *Int. J. Mol. Sci.*, 2015, **16**, 12119–12130.

11 P. Mena, D. González de Llano, N. Brindani, A. Esteban-Fernández, C. Curti, M. V. Moreno-Arribas, D. Del Rio and B. Bartolomé, 5-(3',4'-Dihydroxyphenyl)-γ-valerolactone and its sulphate conjugates, representative circulating metabolites of flavan-3-ols, exhibit anti-adhesive activity against uropathogenic Escherichia coli in bladder epithelial cells, *J. Funct. Foods*, 2017, **29**, 275–280.

12 D. González de Llano, M. V. Moreno-Arribas and B. Bartolomé, Cranberry Polyphenols and Prevention against Urinary Tract Infections: Relevant Considerations, *Molecules*, 2020, **25**, 3523.

13 O. A. Gbinigie, E. A. Spencer, C. J. Heneghan, J. J. Lee and C. C. Butler, Cranberry Extract for Symptoms of Acute, Uncomplicated Urinary Tract Infection: A Systematic Review, *Antibiotics*, 2020, **10**, 12.

14 I. A. Ludwig, M. N. Clifford, M. E. J. Lean, H. Ashihara and A. Crozier, Coffee: biochemistry and potential impact on health, Food Funct., 2014, 5, 1695–1717.

15 S. Bidel and J. Tuomilehto, The emerging health benefits of coffee with an emphasis on type 2 diabetes and cardiovascular disease, *Eur. Endocrinol.*, 2013, **9**, 99–106.

16 M. Ding, S. N. Bhupathiraju, A. Satija, R. M. Van Dam and F. B. Hu, Long-term coffee consumption and risk of cardiovascular disease: A systematic review and a dose-response meta-analysis of prospective cohort studies, *Circulation*, 2014, **129**, 643–659.

17 A. Elbaz, L. Carcaillon, S. Kab and F. Moisan, Epidemiology of Parkinson's disease, *Rev. Neurol. (Paris).*, 2016, **172**, 14–26.

18 R. Huxley, C. M. Y. Lee, F. Barzi, L. Timmermeister, S. Czernichow, V. Perkovic, D. E. Grobbee, D. Batty and M. Woodward, Coffee, decaffeinated coffee, and tea consumption in relation to incident type 2 diabetes mellitus: A systematic review with meta-analysis, *Arch. Intern. Med.*, 2009, **169**, 2053–2063.

19 S. C. Larsson, Coffee, Tea, and Cocoa and Risk of Stroke, *Stroke*, 2014, **45**, 309–314.

20 S. Malerba, F. Turati, C. Galeone, C. Pelucchi, F. Verga, C. La Vecchia and A. Tavani, A meta-analysis of prospective studies of coffee consumption and mortality for all causes, cancers and cardiovascular diseases, *Eur. J. Epidemiol.*, 2013, **28**, 527–539.

21 L. Wu, D. Sun and Y. He, Coffee intake and the incident risk of cognitive disorders: A dose–response meta-analysis of nine prospective cohort studies, *Clin. Nutr.*, 2017, **36**, 730–736.

22 S. Martínez-López, B. Sarriá, G. Baeza, R. Mateos and L. Bravo-Clemente, Pharmacokinetics of caffeine and its metabolites in plasma and urine after consuming a soluble green/roasted coffee blend by healthy subjects, *Food Res. Int.*, 2014, **64**, 125–133.

23 A. Scalbert, L. Brennan, C. Manach, C. Andres-Lacueva, L. O. Dragsted, J. Draper, S. M. Rappaport, J. J. van der Hooft and D. S. Wishart, The food metabolome: a window over dietary exposure, *Am. J. Clin. Nutr.*, 2014, **99**, 1286–1308.

24 N. Brindani, P. Mena, L. Calani, I. Benzie, S. W. Choi, F. Brighenti, F. Zanardi, C. Curti and D. Del Rio, Synthetic and analytical strategies for the quantification of phenyl-γ-valerolactone conjugated metabolites in human urine, *Mol. Nutr. Food Res.*, 2017, **61**, 6–10.

25 C. Curti, N. Brindani, L. Battistini, A. Sartori, G. Pelosi, P. Mena, F. Brighenti, F. Zanardi, D. Del Rio and D. Delrio, Catalytic, enantioselective vinylogous mukaiyama aldol reaction of furan-based dienoxy silanes: a chemodivergent approach to γ-valerolactone flavan-3-ol metabolites and δ-lactone analogues, *Adv. Synth. Catal.*, 2015, **357**, 4082–4092.

26 L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reily, J. J. Thaden and M. R. Viant, Proposed minimum reporting standards for chemical analysis, *Metabolomics*, 2007, **3**, 211–221.

27 J. I. Ottaviani, R. Y. Fong, G. Borges, H. Schroeter and A. Crozier, Use of LC-MS for the quantitative analysis of (poly)phenol metabolites does not necessarily yield accurate results: Implications for assessing existing data and conducting future research, *Free Radic. Biol. Med.*, 2018, **124**, 97.

# Acknowledgements

“One must always acknowledge his/her mentors”, this is one of the many lessons I learned during the past three years from my supervisor, Professor Daniele Del Rio. At the end of this PhD journey it is now my turn to thank him for all the incredible opportunities and priceless advice he has given me. Also thank you for always believing in me.

Besides my mentor, there are several persons that I must thank for always supporting (and tolerating) me during the past three years.

I have to express a huge thank you to Dr. Pedro Mena, for guiding me on my PhD path. I am greatful for his availability, scientific expertise and all the teachings.

Then, I would like to thank all the persons – present and past – of the Human Nutrition Unit of the University of Parma: each and everyone of you has teached me something that I will never forget. In particular, I would like to thank Giuseppe, Rossella, Cinzia, Michele, Veronica, Irene, Nicole, Beatrice, Francesca, Cecilia and Elena to be more than just colleagues: working with you has been a lot of fun. Also, thanks to all the students who helped me with my work.

Thanks to all the researchers I had the possibility to work with, especially Professor Ana Rodriguez-Mateos, Dr. Laura Righetti, Professor Chiara Dall’Asta, Dr. Animesh Acharjee and Professor Claudio Curti. Also, a special thanks to Dr. Luca Calani, for all his precious scientific advice.

Infine, ringrazio i miei genitori per il loro smisurato amore e per aver sempre creduto in me e nelle mie capacità. Inoltre dico grazie a tutta la mia famiglia: zie, zii, cugine e cugini, per essere i miei primi e più grandi sostenitori, sebbene non gli sia ancora ben chiaro che lavoro svolgo.

Ringrazio le mie amiche di sempre, il *fregname*, per i loro incoraggiamenti e per esserci sempre nei momenti importanti.

Ringrazio Francesca, Vittoria, Vincenzo, Giuseppe e Rossella per il loro sostegno, le lunghe chiacchierate e gli spunti di riflessione che sanno offrirmi. Sono amici molto preziosi per me.

Ringrazio Rebecca ed Ilaria, perché anche se sono tanti i km che ci separano nella vita di tutti i giorni, so che posso e potrò sempre contare su di loro.

Ultimo, ma non di certo per importanza, ringrazio Andrea, per supportarmi e sopportarmi ogni giorno e per essere il primo a gioire dei miei successi. La mia vita è più bella da quando ci sei entrato tu.

# About the author

## Curriculum Vitae

Claudia Favari was born on June 11th, 1991 in Castel San Giovanni (Piacenza, Italy).

She performed her academic studies at the University of Parma, earning both Bachelor and Master's Degree in Food Science and Technology at the Department of Food and Drug. After graduating, she spent 8 months working in the laboratory of Professor Del Rio, before as an intern and then as a research fellow. Three years ago, in 2018, Claudia started the PhD course in Food Science at the Department of Food and Drug of the University of Parma, under the supervision of Professor Daniele Del Rio. Her doctoral research aimed at applying metabolomics approaches to nutrition intervention studies in humans to gain further insights into the bioavailability of cranberry flavan-3-ols and associated inter-individual variability, and the metabolic routes modulated by coffee intake, as this is fundamental information to understand the health effects attributed to these dietary sources of bioactive compounds.
During the past three years, Claudia has served as tutor for students of the M. Sc. in Food Science and Technology.

Claudia is author of scientific articles published in ranked journals.

### Scientific activities

*List of publications*

- L. Bresciani, M. Dall'Asta, **C. Favari**, L. Calani, D. Del Rio and F. Brighenti, An *in vitro* exploratory study of dietary strategies based on polyphenol-rich beverages, fruit juices and oils to control trimethylamine production in the colon, *Food Funct.*, 2018, **9**, 6470-6483.

- R. Llorach, **C. Favari**, D. Alonso, M. Garcia-Aloy, C. Andres-Lacueva and M. Urpi-Sarda, Comparative metabolite fingerprinting of legumes using LC-MS based untargeted metabolomics, *Food Res. Int.*, 2019, **126**, 108666.

- L. Bresciani, **C. Favari**, L. Calani, V. Francinelli, A. Riva, G. Petrangolini, P. Allegrini, P. Mena and D. Del Rio, The effect of formulation of curcuminoids on their metabolism by human colonic microbiota, *Molecules*, 2020, **25**, 940.

- F. Martelli, **C. Favari**, P. Mena, S. Guazzetti, A. Ricci, D. Del Rio, C. Lazzi, E. Neviani and V. Bernini, Antimicrobial and fermentation potential of *Himanthalia elongata* in food applications, *Microorganisms*, 2020, **8**, 248.

- **C. Favari**, P. Mena, C. Curti, G. Istas, C. Heiss, D. Del Rio and Ana Rodriguez-Mateos, Kinetic profile and urinary excretion of phenyl-γ-valerolactones upon consumption of cranberry: a dose–response relationship, *Food Funct.*, 2020, **11**, 3975-3985.

- A. Micek, J. Godos, T. Brzostek, A. Gniadek, **C. Favari**, P. Mena, M. Libra, D. Del Rio, F. Galvano and G. Grosso, Dietary phytoestrogens and biomarkers of their intake in relation to cancer survival and recurrence: a comprehensive systematic review with meta-analysis, *Nutr. Rev.*, 2020, **79**, 42–65.

- **C. Favari**, P. Mena, C. Curti, D. Del Rio and D. Angelino, Flavan-3-ols: Catechins and Proanthocyanidins (book chapter in Dietary Polyphenols edited by Francisco A. Tomás-Barberán, Antonio González-Sarrías, Rocío García-Villalba), John Wiley & Sons, Inc., 2020, 283-317.

- L. Bresciani, G. Di Pede, **C. Favari**, L. Calani, V. Francinelli, A. Riva, G. Petrangolini, P. Allegrini, P. Mena and Daniele Del Rio, *In vitro* (poly)phenol catabolism of unformulated- and phytosome-formulated cranberry (*Vaccinium macrocarpon*) extracts, *Food Res. Int.*, 2021, **141**, 110137.

- **C. Favari**+, L. Righetti+, M. Tassotti, L. A. Gethings, D. Martini, A. Rosi, M. Antonini, J. Rubert, C. Manach, A. Dei Cas, R. Bonadonna, F. Brighenti, C. Dall'Asta, P. Mena and Daniele Del Rio, Metabolomic Changes after Coffee Consumption: New Paths on the Block, *Mol. Nutr. Food Res.*, 2021, **65**, 2000875.

- P. Mena+, **C. Favari**+, A. Acharjee, S. Chernbumroong, L. Bresciani, C. Curti, F. Brighenti, C. Heiss, A. Rodriguez-Mateos and D. Del Rio, Metabotypes of flavan-3-ol colonic metabolites after cranberry intake: elucidation and statistical approaches, *Eur. J. Nutr.*, 2021.

- S. Quarta, E. Scoditti, M. A. Carluccio, N. Calabriso, G. Santarpino, F. Damiano, L. Siculella, M. Wabitsch, T. Verri, **C. Favari**, D. Del Rio, P. Mena, R. De Caterina and M. Massaro, Coffee bioactive N-methylpyridinium attenuates tumor necrosis factor (TNF)-α-mediated insulin resistance and inflammation in human adipocytes, *Biomolecules*, 2021, **11**, 1545.

- C. Heiss, G. Istas, R. P. Feliciano, T. Weber, B. Wang, **C. Favari**, P. Mena, D. Del Rio and A. Rodriguez-Mateos, Daily consumption of cranberry improves endothelial function in healthy adults: a double blind randomized controlled trial. (Submitted to *Food Funct.*).

*Oral communications at National and International congresses*

- **C. Favari** - Plasma and urinary levels of phenyl-γ-valerolactones derived from cranberry flavan-3-ols: a dose-response study. *BERRY HEALTH BENEFITS SYMPOSIUM*, Portland, Oregon (OR), 7$^{th}$-9$^{th}$ May 2019.

- **C. Favari** - Livelli plasmatici e urinari di fenil-γ-valerolattoni da flavan-3-oli del cranberry: uno studio dose-risposta. *XL Congresso Nazionale 2019 della Società Italiana di Nutrizione Umana – SINU*, Genova, 27th-29th November 2019.

- **C. Favari** - Metabolomic changes after coffee consumption: new paths on the block. *XLI Congresso Nazionale 2021 della Società Italiana di Nutrizione Umana – SINU*, virtual edition, 9th/10th and 16th/17th April 2021. Awarded with a "PREMIO PER MIGLIOR COMUNICAZIONE ORALE".

- **C. Favari** - Metabotypes of flavan-3-ol colonic metabolites after cranberry intake. *XXX International Conference on Polyphenols – ICP2020*, held as a virtual event hosted by the University of Turku, 13th-15th July 2021.

- **C. Favari** - Metabolomics application to nutritional intervention studies. *I TELEMATIC WORKSHOP ON THE DEVELOPMENTS IN THE ITALIAN PHD RESEARCH ON FOOD SCIENCE TECHNOLOGY AND BIOTECHNOLOGY*, held as a virtual event hosted by the University of Palermo, 14th-15th September 2021.

*Poster communications at National and International congresses*

- **C. Favari** - Plasma and urinary levels of phenyl-γ-valerolactones derived from cranberry flavan-3-ols: a dose-response study. *BERRY HEALTH BENEFITS SYMPOSIUM*, Portland, Oregon (OR), 7th-9th May 2019.

- **C. Favari** - Biomarkers of food intake: a metabolomics approach. *XXIV WORKSHOP ON THE DEVELOPMENTS IN THE ITALIAN PHD RESEARCH ON FOOD SCIENCE TECHNOLOGY AND BIOTECHNOLOGY*, Firenze, 11th-13th September 2019.