universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation / Title of the Doctoral Thesis

„Investigating Polyphenol Exposure at the Omic-scale in Biological Matrices by Liquid Chromatography Coupled to Mass Spectrometry“

verfasst von / submitted by

Ian Oesterle, BSc MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, 2023 / Vienna, 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 796 605 419
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Chemie / Chemistry
Betreut von / Supervisor:	Univ.-Prof. Dr. Benedikt Warth Univ.-Prof. Dr. Annette Rompel

Acknowledgements

I would like to express my gratitude to my two supervisors, Prof. Benedikt Warth and Prof. Annette Rompel, for allowing me the opportunity to do this thesis. I would like to acknowledge the support and guidance they have provided me during this time. Moreover, their insights, feedback, and contributions have helped me grow both professionally and academically. I would then like to extend my sincere thanks to all the members of both working groups. Their help and feedback has been invaluable during my thesis, and they have foster a great working atmosphere. I would like to also extend my gratitude to the members of the Mass Spectrometry Center for their help and technical support. Finally, I would like to express my deepest appreciation to all my friends and family, notably my mother and grandmother, that supported in all aspects during this journey, especially during the difficult times.

Abstract

Polyphenols, bioactive molecules present in plants and fungi, have been known for centuries to positively influence human wellbeing. Besides impacting human health directly, polyphenols can have an indirect effect either through modulating the microbiome by promoting or inhibiting the growth of certain microbes, or through combinatory effects, both synergistic and antagonistic, with other exogenous compounds that humans are exposed to. Studying polyphenols poses a challenge as they are an immensely diverse chemical class, containing a variety of isomers and biotransformation products originating from both human and microbial metabolism. Therefore, innovative workflows based on liquid chromatography coupled with mass spectrometry (LC-MS) were developed and benchmarked to better investigate polyphenols in biological matrices and gain insight into their impact on human health. Firstly, a sensitive targeted LC-MS method and a high-throughput sample preparation for 90 distinct polyphenols, that represent all the major chemical sub-classes, was developed. The method was then validated in-house for three different human matrices (urine, serum, and plasma). Secondly, using the targeted method as reference, an untargeted LC-MS workflow was developed and benchmarked. This workflow showed the potential and suitability of untargeted approaches to investigate polyphenols present in humans beyond those with readily available reference standards. Thirdly, the applicability of the developed workflows was demonstrated in two separate studies. The interchangeability of the workflows between different biological matrices was demonstrated by applying them to nine different plant and mushroom species. This allowed the comprehensive profiling of polyphenols present in these samples in order to better understand the selectivity of polyphenol oxidases found naturally in the samples. The workflows were then applied in a pilot study involving mother-infant pairs to investigate changes in infant exposure to dietary xenobiotics when complementary foods are introduced to their diet. Additionally, correlations between xenobiotics and the infant gut microbiome were explored. The results of the developed workflows demonstrated their potential, especially the untargeted platform, to gain a better understanding of the high variety of polyphenols present in different biological matrices, and their potential link to human health.

Kurzfassung

Seit Jahrhunderten ist bekannt, dass gewisse Pflanzen heilende Wirkung aufweisen. Der Grund dafür sind unter anderem Polyphenole, bioaktive Moleküle, die in den meisten pflanzlichen Lebensmitteln vorhanden sind. Deshalb ist die Erforschung der wirksamen Eigenschaften dieser Moleküle und deren Potenzial, die Gesundheit der Menschen zu beeinflussen, von großem Interesse. Neben einer direkten Wirkung können Polyphenole auch indirekte Effekte erzielen, indem sie das Mikrobiom beeinflussen und dadurch das Wachstum bestimmter Mikroben fördern oder hemmen; oder es entstehen kombinatorische Effekte, sowohl synergistische als auch antagonistische, mit anderen exogenen Verbindungen, denen Menschen ausgesetzt sind. Polyphenole sind eine variantenreiche chemische Klasse mit einer Vielfalt von Isomeren und verschiedenen Metaboliten, die beim menschlichen und mikrobiellen Stoffwechsel entstehen. Dies stellt für die Erforschung von Polyphenolen eine Herausforderung dar und erfordert modernste analytische Verfahren, wie die Flüssigchromatographie in Verbindung mit Massenspektrometrie (LC-MS). Ziel dieser Arbeit war es, neue LC-MS-basierte Methoden für die Erforschung von Polyphenolen zu entwickeln und zu bewerten. Das erste Ziel war, eine hochempfindliche und gerichtete LC-MS Methode mit einer Hochdurchsatz-Probenvorbereitung für 90 Referenzstandards, die alle wichtigen Klassen von Polyphenolen repräsentieren, zu optimieren. Die Methode wurde intern für drei verschiedene menschliche Matrices (Urin, Serum und Plasma) validiert. Zweitens wurde mit Hilfe der gerichteten Methode als Referenz ein ungerichteter LC-MS-Arbeitsablauf entwickelt und einem Benchmarking unterzogen. Dieser Arbeitsablauf zeigte das Potenzial und die Eignung von ungerichteten Workflows, um weitere Polyphenole zu entdecken, als in den verfügbaren Referenzstandards zu finden sind. Drittens wurde die Anwendbarkeit der entwickelten Workflows in zwei separaten Studien nachgewiesen. Die Austauschbarkeit der Workflows zwischen verschiedenen biologischen Matrices wurde durch ihre Anwendung auf neun verschiedene Pflanzen und Pilze Spezies nachgewiesen. Dies ermöglichte eine umfassende Darstellung der vorhandenen Polyphenole und führte zu einem besseren Verständnis der Selektivität von natürlich vorkommenden Polyphenoloxidasen in diesen proben. Die Arbeitsabläufe wurden dann in einer Pilotstudie mit Mutter-Kind-Paaren angewandt, um Veränderungen in der Exposition von Säuglingen gegenüber ernährungsbedingten Xenobiotika zu untersuchen, wenn Beikost in ihre Ernährung aufgenommen wird. Außerdem wurden die Zusammenhänge zwischen Xenobiotika und dem Darmmikrobiom des Säuglings untersucht. Die Ergebnisse der entwickelten Workflows zeigten ihr Potenzial, insbesondere die nicht gezielte Plattform, um ein besseres Verständnis der großen Vielfalt von Polyphenolen in verschiedenen biologischen Matrices und ihre potenziellen Auswirkungen auf die menschliche Gesundheit zu erlangen.

Contents

1. Structure of the Thesis

This thesis is written as a cumulative dissertation made up of the four main parts described below.

The first part constitutes of the introduction, which outlines and describes the basis for this thesis. The background of polyphenols, such as their structure, classifications, and occurrence in nature is explained. Moreover, the human biotransformation and metabolism of polyphenols and their impact on human health is reviewed. This consequently leads to the presentation and explanation of the field of exposomics.

The second part discusses and describes the methodologies used in this thesis. The theory for liquid chromatography and mass spectrometry are described, such as discussing different relevant types of mass analyzers. Furthermore, the various techniques for data acquisition and evaluation used during this thesis are also explained and compared.

The subsequent third section presents the review article and four original research works that were either published in peer-reviewed journals or submitted for publication as a result of this dissertation.

The final fourth section summarizes the main conclusion and major breakthroughs from the research presented in this dissertation. Remaining questions and future endeavours are also presented.

2. Scope and Aims

Polyphenols are a class of molecules which gain a growing interest due to their potential in both clinical and nutritional settings such as personalized medicine. They are abundant secondary plant metabolites [1], and the interest in these chemicals stems from protective effects they exhibit in humans for various chronic diseases, such as antioxidant or antibacterial properties [2], [3]. The microbiome, a key modulator in human health, can also be influenced by polyphenols through, e.g., promoting the growth of beneficial bacterial species [4]. However, polyphenols can also potentially have a negative impact on health, such as the phytoestrogenic effects of isoflavones [5]. In addition, polyphenols are intertwined with the wide-range of xenobiotics that humans are exposed to, influencing their health implications, as seen with the ability of polyphenols to potentially reduce the negative health effects of mycotoxins [6]. Therefore, being able to investigate and quantify polyphenol exposure holistically, both short and long term, in humans would be valuable. Moreover, studying polyphenol exposure during early-life and their influence on chronic disease development would be relevant, especially from the Developmental Origin of Health and Disease perspective [7].

Besides studying human exposure to polyphenols, exploring the polyphenols present in various plants and fungi is also important. This would allow to adjust the different plant-based foods in one's diet to modulate for specific polyphenol exposure, and it would help the investigation of polyphenol oxidases (PPOs), a key enzyme responsible for enzymatic browning of plants and fungi [8]. PPOs are type-III dicopper metalloproteins that catalyze the *o*-hydroxylation of monophenols into *o*-diphenols, and the oxidation of *o*-diphenols to *o*-quinones [9]. Though their function *in vivo* is still not yet fully understood, these enzymes are involved in converting polyphenols into *o*-quinones that can then polymerize to melanin [10]. PPO activity is currently believed to be a plant defense mechanism by creating reactive oxygen species [11]. Therefore, comprehensive polyphenolic profiling of plants and fungi that contain PPOs would aid in understanding their specificity and reactivity [12].

Investigating polyphenols in both plants and humans is a challenging task. It is currently hypothesized that over 8'000 dietary polyphenols exist [13] with a plethora of isomers, especially as polyphenols can easily be conjugated due to the available hydroxyl groups [14]. Therefore, innovative methods based on liquid chromatography coupled to mass spectrometry (LC-MS) are required. Various methods analyzing polyphenols in a targeted manner exist, yet they mainly focus on a limited number of polyphenols and typically on only a couple of related classes [15]. Very few methods that involve a comprehensive look at all the different polyphenol classes exist. Even fewer include conjugated polyphenols, such as biotransformation products, even though the majority of polyphenols previously found in humans were conjugated [16]. Similarly, untargeted analytical workflows for

investigating polyphenols, especially in humans, are scarce, though it is an advantageous technique as it allows to study analytes beyond those that reference standards are readily available [15].

Therefore, the work performed within the framework of this thesis was to develop and benchmark new pioneering workflows, using both targeted and untargeted approaches, to investigate polyphenols and their metabolites in humans. The main aims of the thesis are listed as follows:

- Development and in-house validation of a targeted LC-MS/MS method in human bio-fluids for 90 different polyphenols that represent all of the main chemical sub-classes [17].
- Establishing an untargeted workflow (LC-HRMS method and data analysis) using the targeted method as a reference, in order to more comprehensively study polyphenols and their metabolites [18].
- Apply the workflows to study polyphenol exposure in humans and investigate their potential effects on the gut microbiome and with other xenobiotics humans are exposed to [19].
- Apply the workflows on plants and fungi to show the transferability of the methods and to aid in understanding PPO specificity by investigating the natural polyphenols present [20].

3. Key Outcomes

The key outcomes achieved from the research of the presented thesis are as follows:

- Wrote and published a review paper discussing the current state of human biomonitoring of polyphenols and the challenges required to overcome in this field of research.
- Developed a targeted LC-MS/MS method for 90 polyphenol analytes representing the thirteen main polyphenol sub-classes and validated the method in-house for human urine, serum, and plasma.
- Even though the targeted LC-MS/MS sample preparation used low sample volumes (40 μL) and was high-throughput, it showed high sensitivity, and its suitability was proved in a pilot study.
- Successfully established an untargeted LC-HRMS platform based on the targeted LC-MS/MS approach, and quantified the gap in sensitivity between the two approaches.
- The untargeted platform was developed to also include suspect screening of polyphenols and biomarkers of environmental exposure, and it tested and benchmarked various non-targeted tools to find additional polyphenol related features.
- The untargeted workflow was applied in a longitudinal pilot study involving mother-infant pairs to investigate the change in dietary exposure as complementary foods are introduced in the infant's diet.
- Potential correlations between xenobiotic exposure and the infant gut microbiome were investigated in the mother-infant pairs.
- The developed workflows were applied to nine plant and mushroom species to investigate the polyphenols naturally present and better understand polyphenol oxidases, and to show the interchangeability of the workflows between different biological matrices.

4. Introduction

Different plants and fungi, thought to be beneficial, have been used for centuries in traditional medicine [21]. Modern research has shown that the phytochemicals present and their bioactive properties may be responsible for the effects on human health [22]. One major group of phytochemicals are polyphenols. Their presence in various plants and fungi makes them common in human nutrition and may have an influence on the wellbeing of people. Therefore, the interest in this chemical group has been growing since the use of certain plants in traditional medicine has shown to have numerous beneficial properties for humans [23]. Being able to investigate human exposure to these chemicals will allow a more comprehensive understanding of their physiological impact and aid in various aspects such as personalized medicine.

4.1. Polyphenols

Polyphenols are an immensely large family of chemicals. It is currently presumed that over 8'000 dietary polyphenols exist [13]. This chemical family can be separated into two main classes: flavonoids and non-flavonoids, which are subsequently broken down into several sub-classes. The major polyphenol classes and sub-classes, along with their main chemical structures are depicted in Figure 4.1.

4.1.1. Flavonoids

The first main class of polyphenols are flavonoids. Molecules part of this class all contain a similar backbone with two phenyl rings and one heterocyclic ring (C6-C3-C6). These molecules can be further divided into seven different sub-classes depending on the saturation of the heterocyclic ring and the position of the secondary phenyl ring (Figure 4.1) [23], [24]. From the backbone, polyphenols can have a variety of substitutions on the different available carbon atoms, as shown by the "R" positions in Figure 4.1. These substitutions are mainly hydroxyl groups, which easily allow for conjugations to occur, as can be seen in plants and fungi where the majority of polyphenols are naturally conjugated with one or more sugar moieties [14].

- **Anthocyanidins**

  Anthocyanidins are the only sub-class of polyphenols that contain a natural positive charge, found on the oxygen present in the heterocyclic ring. Anthocyanidins are prevalent in many fruits and vegetables, especially colored fruits such as berries or grapes. They are pigments giving colors namely purple, blue, or red, depending on

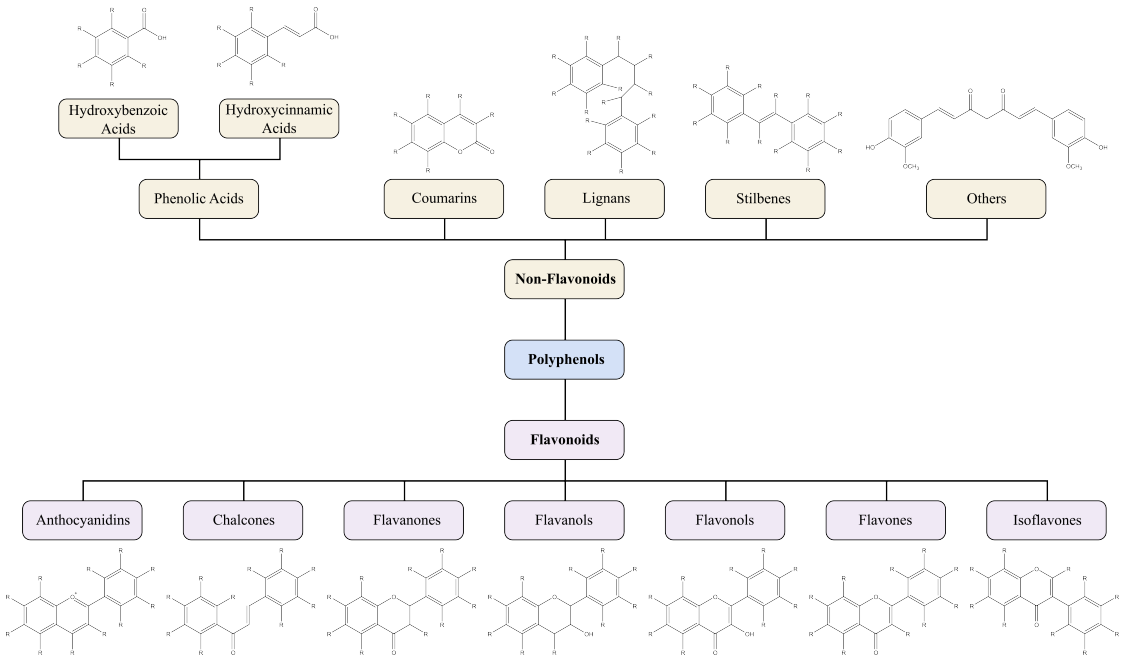


Figure 4.1.: A diagram of the main polyphenol classes and sub-classes. For each sub-class, the basic backbone of the structure is shown, where "R" represents a possible location for a functional group, typically a hydroxyl group.

the pH. They are often found naturally in conjugated forms with sugar moieties, and the most common anthocyanidins are cyanidins, e.g. cyanidin-3-glucoside [25].

- **Flavanols**

  Flavanols can either exist as monomers (catechins) or polymers (proanthocyanidins) [26]. As the heterocyclic ring of flavanols is saturated and substituted, many different stereoisomers are possible. Moreover, catechins are commonly found esterified to gallic acid. Catechins are present in high abundance in green tea [27] and chocolate [28], with the most prevalent being (-)-epigallocatechin-3-gallate [29]. While proanthocyanidins are mainly found in berries and fruits, and the most prevalent proanthocyanidins are proanthocyanidin Bs [30].

- **Flavonols**

  The most ubiquitous sub-class of flavonoids are flavonols, which are found in the majority of plants. Besides glycosylation, flavonols readily undergo hydroxylation and methylation [23]. Foods high in flavonols include onions, leek, and blueberries [26]. The main and most common flavonols are quercetin, kaempferol, and rutin, which is quercetin conjugated with a rutinoside.

- **Flavones**

Flavones are not as prevalent as flavonols, and are typically found conjugated as 7-O-glucosides [31]. They are found in a variety of plants, especially herbs. For example, chamomile and parsley contain high concentrations of flavones [31]. The most common flavones include apigenin and luteolin [26].

- **Isoflavones**

  Isoflavones are similar to flavones, but the position of the secondary phenyl ring is changed. Isoflavones are distinct as they are not as commonly found as other flavonoids, rather they are found in leguminous plants, especially in soybeans [24]. The most popular and well-known isoflavones are daidzein, genistein, and glycitein.

- **Flavanones**

  Flavanones are similar to flavanols, but the only difference is that the heterocyclic ring is fully saturated. They are present in several herbs and vegetables, with the highest concentrations found in citrus fruits [32], where they are responsible for the bitter taste [26]. The most common examples of flavanones are hesperetin, eriodictyol and naringenin [24].

- **Chalcones**

  The last sub-class of flavonoids are chalcones, a unique class, as they are the only sub-class with no heterocyclic ring. Therefore, they are typically referred to as "open-chain" flavonoids [23]. Chalcones are commonly found in various plants, notably tomatoes and licorice [33]. Examples of chalcones include phloridzin, phloretin, and naringenin chalcone [24].

4.1.2. Non-Flavonoids

The second main class of polyphenols are non-flavonoids. These include a wide range of polyphenols that do not have the general backbone of flavonoids. Non-flavonoids can be broken down into five main sub-classes. Again, these polyphenols can be found in conjugated forms.

- **Phenolic acids**

  The largest sub-class of non-flavonoids are phenolic acids. Molecules in this sub-class contain one phenol ring and can be split into two main groups: hydroxybenzoic acids and hydroxycinnamic acids. In addition, the molecules can exist in their free or esterified forms [26]. Hydroxybenzoic acids are present in various plants but in low concentrations, except for gallic acid that is found abundantly in tea and fruits such as pomegranate and raspberry [34]. Hydrolyzable tannins, such as ellagitannin, are made from hydroxybenzoic acids [26]. Unlike hydroxybenzoic acids, hydroxycinnamic acids are more prevalent and found in a variety of plants and fungi. Common examples of these acids include ferulic acid, coumaric acid, and chlorogenic acid [26]. Chlorogenic acid is most likely the most well known

hydroxycinnamic acid, as it is the esterification product of cinnamic acid and quinic acid, and found in high concentrations in coffee [35].

- **Lignans**

  Lignans are formed by two phenylpropane units, and depending on the structure can again be split into eight groups, e.g. furofuran, thus many different types exist [36]. Lignans are found in a variety of plants, and common sources of dietary lignans include flaxseed and sesame seeds [36]. Examples of lignans include matairesinol and secoisolariciresinol [26].

- **Stilbenes**

  The next sub-class that is present in limited concentrations in dietary foods are stilbenes. Molecules in this sub-class contain two phenyl rings connected by an ethylene moiety (C6-C2-C6). Stilbenes can be found in a variety of foods, notably grapes and wine [37]. The most popular and commonly found stilbene is resveratrol [26]. Stilbenes can also be found naturally as oligomers, such as viniferin [38].

- **Coumarins**

  Coumarins are based on a benzopyrone structure and thus have a similar backbone to flavonoids, except there is no second phenyl ring [23]. Currently it is thought that at least 800 different types of coumarins exist, which can be further divided into four different groups, such as furanocoumarins or pyranocoumarins [39]. Coumarins are present in a variety of plants, whereupon one prevalent source of dietary coumarins is cinnamon [40]. Common coumarins include umbelliferone and aescultin [41].

- **Others**

  Finally, the last sub-class of non-flavonoids includes all the other polyphenols that do not pertain to any of the other sub-classes. One example is curcumin, a diarylheptanoid prevalent in tumeric [42]. Another example is capsaicin, a polyphenolic amide found in chili peppers [43].

4.1.3. Polyphenols in Plants and Fungi

Not only do many different polyphenols exist, they are also highly prevalent in plants and fungi, as exhibited on Phenol-Explorer, a database containing the phenolic profile for over 400 foodstuffs [44]. Polyphenols are secondary metabolites of plants and fungi, mainly derived from the shikimate and phenylpropanoid pathways. The main steps of the shikimate pathway consists of converting phosphoenolpyruvic acid and D-erythrose-4-phosphate into shikimic acid, which is then converted into chorismic acid. Chorismic acid can then be converted into hydroxybenzoic acids or into L-phenylalanine and L-tyrosine that are the building blocks for polyphenols [45]. Whereas the phenylpropanoid pathway involves converting phosphoenolpyruvic acid and D-erythose-4-phosphate into dihydroshikimate. Dihydroshikimate can then either be converted into gallic acid to

produce tannins, or into shikimate that is then converted into phenylalanine to produce different polyphenols [46].

As polyphenols are secondary metabolites, they play various functions in plants and fungi. Their production is related to both biotic and abiotic stresses that the plants and fungi experience [45]. For example, polyphenols are produced to protect organisms from ultraviolet radiation or herbivores [47]. Moreover, it is thought that certain defense mechanisms can involve the conversion of polyphenols into reactive oxygen species that trigger additional defense pathways [11]. Additionally, in the case of polyphenols, these reactive species can polymerize into melanins, which is the cause of enzymatic browning in plants and fungi [8]. The conversion of polyphenols into reactive oxygen species is catalyzed by polyphenol oxidases, a family of type-III dicopper enzymes [9]. This family can be divided into two main groups: tyrosinases (TYRs) (EC 1.14.18.1) and catechol oxidases (COs) (EC 1.10.3.1). TYRs can convert monophenols into *o*-diphenols through *o*-hydroxylation while reducing O_2. Then, both TYRs and COs can convert *o*-diphenols into *o*-quinones, again by reducing O_2 [12].

4.2. Biotransformation of Polyphenols

Once ingested, polyphenols, like many other exogenous molecules, undergo a variety of biotransformation processes, further expanding and complexifying this chemical class. In humans, an array of biotransformation processes occur, mainly at the liver, in order to make the chemicals more polar to ease their excretion [48]. These processes can be broken down into three phases: phase I, II, and III, which can occur either sequentially or simultaneously [49]. Phase I involves increasing the polarity by functionalization reactions, such as introducing or revealing a polar group. Several different reactions are part of this phase, and the main reaction is oxidation with the aid of cytochrome P450 [50]. Other reactions that take place include hydrolysis, reduction, or dehydrogenation. Phase II reactions involve conjugation reactions with a polar conjugate. These conjugation reactions occur either on the functional groups already present on the chemical or to function groups introduced from phase I metabolism. The main types of reactions include conjugations with sulfates (sulfation), methyls (methylation), acetyls (acetylation), glutathiones, amino acids, and, most importantly, glucuronic acids (glucuronidation) [51], [52]. In these reactions, a variety of enzymes are involved, for example uridine-5'-diphospho-glucuronosyltransferases for glucuronidation [53]. Moreover, multiple conjugation reactions can occur depending on the available functional groups, such as introducing two glucuronide conjugates or one glucuronide and one sulfate conjugate. The last phase (phase III) occurs after phase II and includes further metabolism of the phase II products and transporter-mediated excretion from the body typically via bile, sweat, or urine [49]. The two major transporters used for elimination are adenosine triphosphate-binding cassette proteins [54] and solute carrier transporters [55].

Besides human biotransformation processes, biotransformations from the microbiome also occur. The microbiome is a complex ecosystem composed of all bacteria, archaea, and viruses found in the human body [56]. It is estimated that there are 10 times more cells

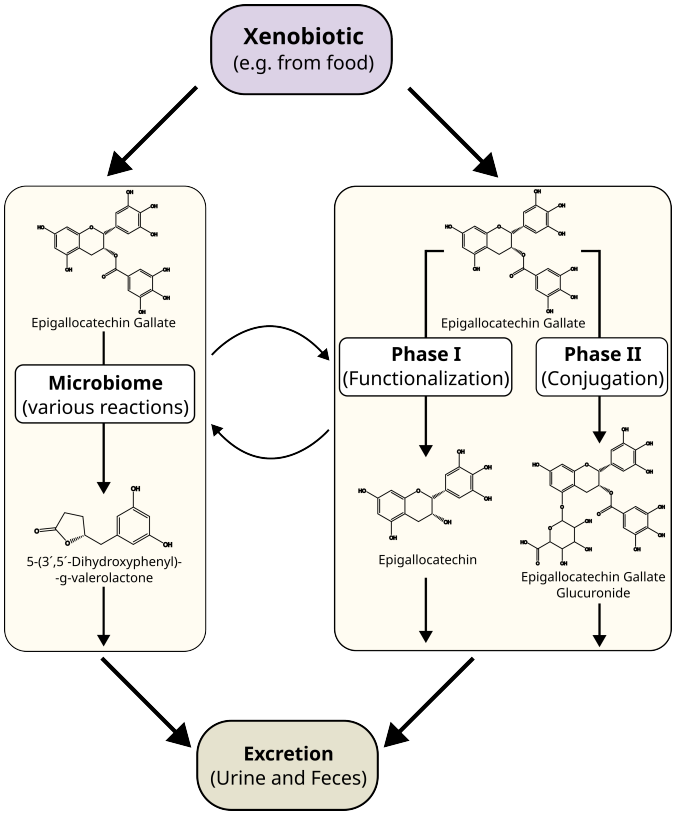


Figure 4.2.: An overview of the biotransformation processes from both human and microbial origin. The biotransformation of epigallocatechin gallate is shown as an example.

from the microbiome than human cells, originating from more than 1'000 different species, with the major species being *Bacteroidetes* and *Firmicutes* [57], [58]. The microbiome changes during an individual's lifetime, especially in the first few years of life [59], and it is influenced by various factors such as age, diet, environment and medication [60]. Depending on the bacterial species present, a variety of biotransformation processes can occur, typically producing smaller and more bio-active molecules [61]. These processes include reactions such as hydrolysis, cyclization, methylation, or redox [62].

Dietary polyphenols undergo both human and microbial biotransformations, yielding a plethora of metabolites. As natural polyphenols are typically found either conjugated with sugars, esterified, or as polymers (e.g. proanthocyanidins), they are not readily absorbed, with only about 5-10% absorbed in the intestine [63]. Therefore, polyphenols are first broken down either during phase I metabolism or through the microbiome. In both processes, the sugar conjugates are readily removed, but with microbial biotransformation, the aglycones are easily further digested [26]. For example, certain *Eubacterium* strains can produce chalcones from flavonoids and certain *Bifidobacterium* strains produce phenolic acids, such as hydrolyzing chlorogenic acid into caffeic acid [62]. The aglycones and

products from microbial biotransformation then readily undergo phase II metabolism, mainly glucuronidation, sulfation, and methylation reactions [64], [65]. For instance, one study reported that over 60% of the isoflavone metabolites detected were conjugated with a glucuronide moiety [16]. Moreover, due to the availability of various functional groups on polyphenols, multiple phase II reactions can occur, producing metabolites such as di-glucuronides [66]. Finally, the numerous polyphenol metabolites are then excreted from the body, mainly through urine and feces. An overview of both human and microbial biotransformation processes with epigallocatechin gallate as an example is depicted in Figure 4.2 [67].

4.3. The Impact of Polyphenols on the Microbiome and Human Health

Besides the high prevalence of polyphenols in an individual's diet, the interest in these molecules arises from their bio-active properties. Polyphenols show a wide range of beneficial health effects, such as anticancer, anti-inflammatory, and antimicrobial properties [66], [68]. The main mode of actions for these beneficial properties stems from the fact that polyphenols reduce oxidative stress by both, acting as free radical scavengers and producing hydrogen peroxide [1]. However, polyphenols can also exhibit adverse effects on health such as pro-oxidative effects in high concentrations [69]. In addition, certain polyphenols, e.g. isoflavones, are also phytoestrogens, meaning that they have an estrogen-like structure, and therefore can exhibit estrogenic and antiestrogenic-like activity [5].

Human health is also closely intertwined with the microbiome, for instance, a dysbiosis in the gut microbiome is linked to irritable bowel syndrome [70]. Since polyphenols are metabolized extensively by the microbiome, exposure to polyphenols can thus have an impact on human health through modulating the microbiome. For instance, polyphenols can induce the growth of the *Akkermansia* species, which can potentially reduce inflammation [71]. Additionally, polyphenols have shown antimicrobial properties [72], namely they can inhibit the growth of pathogens, as illustrated by the consumption of catechins from green tea which inhibit the *Candida* genus, or the intake of flavanones from citrus fruits, which inhibit *Helicobacter pylori* [73].

During their lifetime, humans are exposed to a variety of xenobiotics that can have adverse effects on health. One prominent example of such a xenobiotic is bisphenol A, which was found to be an endocrine disrupting chemical [74]. These xenobiotics can not only have a direct influence on human health but also indirect effects through impacting the microbiome [75]. Moreover, mixture toxicological effects can occur, such as drug-drug interactions or interactions between xenobiotics and drugs [76]. Polyphenols therefore play a key role in mixture toxicological effects among polyphenols and other xenobiotics. For instance, flavonoids from hops were found to have a synergistic estrogenic effect with endocrine disrupting pesticides [77]. In contrast, certain polyphenols can help prevent the negative health impact of mycotoxins [6].

Exposure to xenobiotics, as previously discussed, can have profound effects on human

health. In addition, it is believed that chronic diseases may have their roots in early-life [78], [79], which has led to the concept of Developmental Origins of Health and Disease [7], [80]. Given that early exposure to certain xenobiotics can influence the development of chronic diseases [81], polyphenol exposure of infants, as well as of the mother and father, during early life is assumed to have a profound impact on future health development [82]. Moreover, during early life, the microbiome is still in the process of being colonized, and the composition of the gut microbiome, which is fully developed after about three years of age, can then have important health implications [59]. As mentioned before polyphenols can modulate the microbiome, making exposure to polyphenols during early life a key factor for the healthy development of humans.

4.4. The Role of Polyphenols in the Context of Human Biomonitoring and Exposome Research

As previously discussed, exposure to various xenobiotics may have profound health implications, especially during early-life [83], leading to the term "exposome", which was first introduced by Christopher Wild in 2005 [84]. The exposome can be defined as the total exposure to exogenous chemicals, e.g. through diet or environment, and their corresponding biological impact during an individual's lifetime [85]. This represents an expansion of the metabolome, that is defined as all low molecular weight molecules and metabolites that are present in a biological specimen [86]. Hence, the goal of exposomics is to investigate xenobiotics and their corresponding biotransformation products that an individual is exposed to, as well as finding biomarkers that link exposure to health outcomes [87]. The interaction and role of the microbiome with the exposome is also included [88].

Conventionally, dietary exposure to xenobiotics was investigated and related to health outcomes through food frequency questionnaires. However, this method has multiple limitations such as not providing accurate quantitative data or being subject to biases [89]. Therefore, more precise and advanced human biomonitoring approaches are required to investigate the exposome, especially as measuring exogenous and endogenous chemicals is challenging since they are chemically diverse and are found in a wide range of concentrations. For example, median blood concentrations of endogenous molecules were at 0.94 μM, while pharmaceutical drugs were at 0.34 μM and pollutants at 2.4 x 10^{-4} μM [90]. Besides varying concentrations depending on the type of xenobiotic investigated, the chemicals that will be detected are subject to the type of biological matrix chosen. Typically, blood or urine are chosen as biological matrices as their collection is minimally [91]. Urine is a method of excretion, thus contains mainly waste and by-products, such as phase II conjugation products [92]. In contrast, blood is a dynamic biological matrix that is used by a variety of metabolomic pathways, thus diverse metabolites can be detected [91]. Besides selecting the biological matrix, the sampling time is also key, e.g. there are large variances in urine depending on the time of the day it was collected [93].

The majority of polyphenols have similar structures, which makes them advantageous over other xenobiotics for investigation. Previous research has also shown that polyphenols

are detected in measurable concentrations between 0.3 and 1800 nM in plasma [94] and between 0.02 and 1000 μM in urine [95]. Depending on the biological matrix chosen, different polyphenols are observed. For instance, in urine mainly phenolic microbial metabolites and conjugated polyphenols are present [96], while in breast milk, mainly flavonols, catechins, and flavanones are observed [97]. Furthermore, polyphenols can also be detected in a variety of other biological matrices, such as in healthy and malignant tissue [98] or bones [99].

As mentioned previously, food frequency questionnaires were a traditional method to measure dietary exposure, especially in the context of human exposure of polyphenols. However, not only do the types of polyphenols present and their quantities vary depending on the plant or fungi, but there are also variances of the same plant species, based on factors such as the location they are grown [100]. Therefore, food frequency questionnaires can not give an adequate representation of the quantities of polyphenols a person is exposed to. Another conventional technique that can be employed is to approximate polyphenol intake by measuring the total phenol content of certain foods, giving a more quantitative value on the amount of polyphenols ingested, but the drawback is that information on the type of polyphenols present is missing [101], [102]. As a plethora of polyphenols exist, knowing which types of polyphenols have a beneficial impact on human well being is of great significance. Therefore, novel metabolomic and exposomic approaches with cutting-edge instruments, such as mass spectrometry, are required to measure human polyphenol exposure and its impact on health [15].

5. Methodology

Determining the different chemicals present in a sample matrix is a challenging endeavor because of the complexity and the multitude of compounds present, especially in the ever changing human bio-fluids. The determination of chemicals of interest, or analytes, in a sample can be broken down into three main steps. The first step is to prepare and extract the analytes from the sample. The second step is to separate the different analytes in the extracted mixture, and the third and final step is to determine and potentially quantify the separated analytes.

Sample preparation is a key step as careful preparation simplifies the separation step and ameliorates the detection step by reducing interferences with other compounds present. Moreover, depending on the sample type, thorough sample preparation is required to allow the extraction of analytes, e.g. extracting them out of cells. An assortment of sample preparation techniques exist [103], for example solid-phase extraction, which is typically a time consuming procedure, but it selectively isolates analytes from complex samples [104]. Overall, the sample preparation procedure used depends on the analytical goals of the workflow, such as being high-throughput if a large number of samples will be analyzed, or on the techniques used in the rest of the workflow, such as the type of chromatography.

A fundamental technique in analytical chemistry is chromatography, a process for separating different constituents in a mixture. A variety of chromatographic methods exist, but they all work with the same principle. The basic working principle is exemplified in Figure 5.1, where a mixture is carried by a mobile phase through a stationary phase. The components in this mixture will interact with both the stationary and mobile phases. From these interactions, they pass through the stationary phase at different speeds, thus allowing for a separation of the initial mixture. The stationary phase can be in either solid or liquid phase, while the mobile phase is either a gas or liquid phase [105], yielding a variety of chromatographic techniques.

Following the separation of components in a mixture by chromatography, various detectors can be used for the qualification and quantification of the components. The type of detector to use depends on the analytes that are investigated, and on the requirements of the analytical procedure. For example, fluorescent chemicals can easily be analyzed with a fluorescence detector, but this detector will obviously not function with non-fluorescent chemicals [106]. Non-specific and more commonly used detectors include flame ionization or thermal conductivity [107]. Nowadays though, nuclear magnetic resonance (NMR) and mass spectrometry (MS) are becoming increasingly popular, especially the latter due to its high versatility and sensitivity [108].

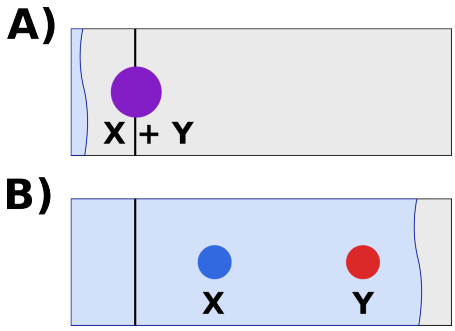


Figure 5.1.: The basic working principle of chromatography. **A)** A mixture of X+Y is placed on the stationary phase (gray). **B)** As the mobile phase (blue) moves through the stationary phase (grey), the mixture is then separated into its components X and Y as these components travel with different speeds through the stationary phase due to their interactions with both the stationary and mobile phases.

5.1. Liquid Chromatography

Liquid chromatography (LC) is one of the most widely used chromatographic techniques to separate analytes in a liquid mixture. Contrary to gas chromatography (GC), in LC, the mobile phase is in a liquid state rather than a gas state, offering the advantage to separate a wider range of molecules [109]. Additionally, in LC, the stationary phase is typically a solid phase placed on a solid support, mainly beads, that are packed in a column.

The mobile phase in LC can be made up of a single solvent, such as water, acetonitrile, or methanol, but, typically, a mixture of at least two different solvents is used. Mixing two or more solvents at different ratios allows for a better control over the polarity of the mobile phase, and thus control over the analytes' interactions with both the stationary and mobile phases. Depending on the types of phases used, the analytes will have varying affinities to the phases and pass through the stationary phase at different times. The time required for an analyte to pass through the entire stationary phase is known as elution time, or retention time [109]. Moreover, a gradient is usually applied to the mobile phase, meaning that the percentages among two or more solvents are varied over time, which allows even more control over the elution of the analytes. For example, the change in polarity of the mobile phase over time can avoid a peak broadening effect that occurs for later eluting analytes when the mobile phase is isocratic [109]. The solvents of the mobile phase, also called eluents, can be further modified by adding various additives such as ammonium fluoride or formic acid. These additives are used for two main reasons: to aid in the ionization of the analytes in the detector, and to control the pH of the mobile phase [110]. The pH of the mobile phase is critical, as depending on the pH, the analytes can be either fully or partially protonated or deprotonated, thus producing different interactions with the stationary phase and impacting retention times [111].

As previously mentioned, a variety of different LC techniques exists, for instance normal-phase, size-exclusion, or hydrophilic interaction; and the choice depends on the type of stationary and mobile phases used [109]. One common technique is reverse-phase (RP) [112], meaning the stationary phase is non-polar, typically made from alkyl chain groups, namely C18. The mobile phase is then composed of polar solvents, usually water. Thus, non-polar analytes will elute later than polar analytes due to their higher affinity to the non-polar stationary phase rather than the polar mobile phase. As mentioned earlier, separation and peak shapes can be improved by employing a gradient on the mobile phase. In RP chromatography, this involves making the mobile phase more non-polar over time by reducing the percentage of the polar solvent (ex. water), and increasing the quantity of a more non-polar solvent (ex. acetonitrile). Thus, the non-polar analytes that are interacting with the stationary phase start having a higher affinity to the mobile phase and elute more quickly. Additionally, the mobile phase can also be set to pass through the stationary phase at different flow rates, thereby affecting the separation of the mixture as the analytes will have more or less time to interact with the stationary phase.

Besides controlling the type of LC system, type of stationary phase, and the constitution and flow rate of the mobile phase, a variety of other parameters, especially for the stationary phase, can be chosen in order to fine-tune the chromatographic separation [113]. For example, the size of the particles on which the stationary phase is attached to in the column can be changed, with the result: the smaller the size, the higher the efficiency and so the resolution, i.e. the separation between two analytes. Similarly, resolution and efficiency can be increased by lengthening the column, with the draw-back that a longer column increases elution times. Besides these two parameters, column radius, flow rate, and viscosity of the mobile phase have an effect on the column back pressure. Column back pressure is an important parameter to consider as it will impact the pump system used and columns are only stable within certain pressure ranges. Heating the column can decrease back pressure, but also has an impact on selectivity and retention [114]. Overall, based on the sample used and the desired goals, all of these parameters are optimized, and compromises among them are typically needed.

An overview of a general LC system with the different parts is depicted in Figure 5.2. To start, the mobile phase is created by mixing different eluents together. The eluents needed are filled in bottles and attached to a pump system, which delivers a precise flow of eluents to the whole system. The different eluents first flow into a mixer, or gradient valve, where they are mixed at different compositions, depending on the desired constitution or gradient used to create the mobile phase. Then, the mobile phase continues through an injection chamber and into the column, which contains the stationary phase. In the injection chamber, the sample or mixture to separate is introduced. The column is typically placed in a column oven so that a specific temperature can be set. While the mobile phase, including the sample, flows through the column, the analytes in the sample move at different speeds through the stationary phase causing a separation among them. The separated analytes then elute out of the column and enter the detector.

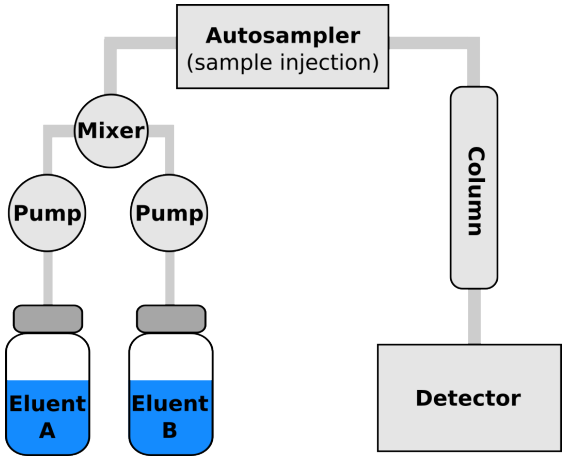


Figure 5.2.: The overview of a LC system with its different components, such as the mixing of different eluents to create the mobile phase, the column that contains the stationary phase, and the sample introduction (autosampler).

5.2. Mass Spectrometry

After separation by LC, the objective is to determine the analytes present by using a detector. A wide-range of detectors are available, and the type of detector used in a set-up depends on the analytical goals, and on the chromatographic technique. Moreover, if desired, multiple detectors can be set in parallel to increase the range of analytes to be detected. For example, one of the most common detectors used is an ultraviolet-visible light (UV-Vis) spectrometer, because of its convenience and reliability. Its basic working principle involves measuring the amount of ultraviolet or visible light absorbed or transmitted through the exit flow of an LC system, given that the amount absorbed at various wavelengths changes when different analytes elute [115]. Another common detector is a mass spectrometer. This type of detector is extremely useful due to its high sensitivity and selectivity compared to other detectors [108]. Moreover, mass spectrometers allow for better confirmation of the analytes that elute, especially if several analytes elute at the same time.

Mass spectrometers are detectors that function through the application of electric or magnetic fields. These detectors can be broken down into three different sections. The first section is an ionization source, as typically analytes entering the mass spectrometer are neutral, and thus need to become charged in order for the electric or magnetic fields to apply a force on the analytes. The second section is a mass analyzer which separates the analytes by their mass-to-charge ratios (m/z). The third and final section is a detector which measures the number of ions that reach the detector.

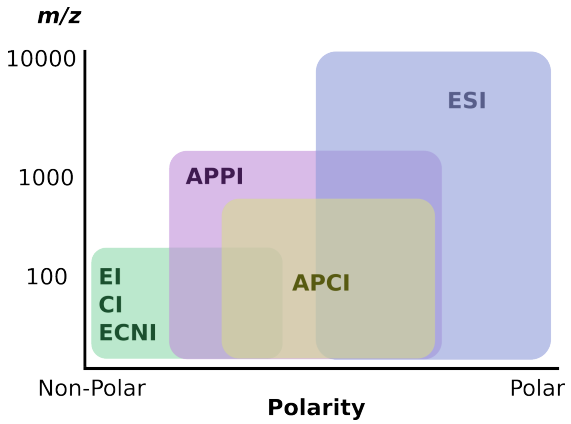


Figure 5.3.: The polarity and m/z range of analytes that can be ionized by the different ionization techniques. The techniques displayed are: electron ionization (EI), chemical ionization (CI), electron capture negative ionization (ECNI), atmospheric pressure photoionization (APPI), atmospheric pressure chemical ionization (APCI), electrospray ionization (ESI).

5.2.1. Ionization Source

An assortment of ionization techniques currently exists, with many more being developed all the time. The type of ionization will depend on both the chromatographic technique used and on the analytes of interest [116]. Figure 5.3 displays some of the most commonly used ionization techniques and the range of m/z and polarity that each technique works for. Moreover, ionization techniques differ in the number of fragments they create when ionizing an analyte, thus its selection again depends on the objective of the analytical method [117].

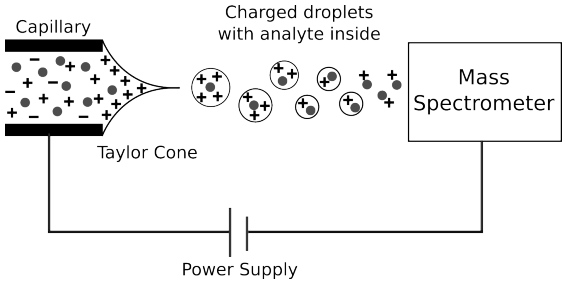


Figure 5.4.: The basic working principle of an electrospray ionization source.

One of the main ionization techniques used with LC for xenobiotics such as polyphenols is electrospray ionization (ESI) [118]. ESI is a soft ionization technique, meaning that it creates little to no fragments, and its basic working principle is displayed in Figure 5.4. ESI works by having the exit flow of the LC pass through a narrow capillary where a voltage is applied. This creates a Taylor cone that leads to a spray of charged droplets.

The solvent of these droplets is evaporated as they flow towards the mass analyzer. During evaporation, the droplets reach their Rayleigh limit, and the surface becomes unstable, causing Coulomb fission and creating smaller charged droplets. This process continues and is repeated until single ions remain. The voltage applied, temperature and flow of the heating gas, and the distance between the capillary and the orifice of the mass analyzer are all optimized to improve the ionization efficiency of the analytes investigated. Moreover, at the orifice of the mass analyzer, a curtain gas can be utilized to reduce contamination by repulsing non-evaporated solvent droplets [119]. Similarly, a declustering potential can be applied at the orifice to prevent clustering of the ions as they enter the mass analyzer. The declustering potential again is optimized depending on the analytes investigated [120].

5.2.2. Mass Analyzers

A variety of mass analyzers exists, including quadrupoles, cyclotrons or ion traps, that can either be used by themselves or in different combinations with each other depending on the need. Each type of mass analyzer contributes to the mass spectrometers' resolution and sensitivity in its own way [121], [122]. Sensitivity can be defined as the signal-to-noise (S/N) ratio for a specific concentration of analyte, while resolution is the ability to separate two m/zs. Typically, there is a trade off between resolution and sensitivity, thus each type of mass analyzer has its own advantages and disadvantages. In the context of this thesis, three relevant types of mass analyzers are described: quadrupoles, time-of-flight (TOF), and orbitrap. These analyzers are the main types used in metabolomics and exposomics due to the state-of-the art sensitivity achieved by triple quadrupoles required for targeted workflows, and the high resolution attained by TOFs and orbitraps, which is essential in untargeted workflows [123].

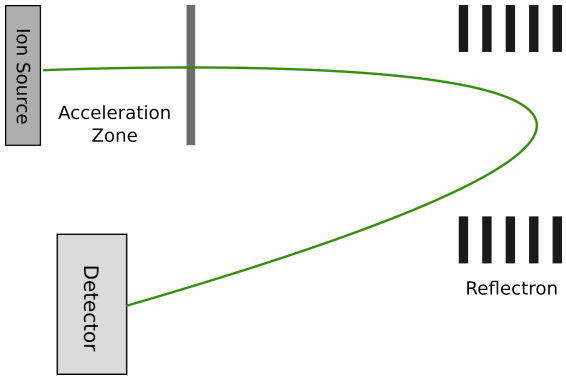


Figure 5.5.: The working principle of a time-of-flight mass analyzer with the trajectory of an ion depicted in green.

A TOF is the most basic type of mass spectrometer, and its basic working principle is shown in Figure 5.5 [124]. Ions leaving the ion source enter the first part of the mass analyzer, an acceleration zone. In this zone, a potential is applied so that the ions are

accelerated with a certain kinetic energy. Given that the kinetic energy (E_k) is related to the mass (m) and velocity (v) of an object, as shown in Equation 5.1, the ions entering the TOF that are given the same kinetic energy, will have different velocities depending on the mass of the ions. Therefore, the ions will take different times to travel through the TOF chamber and reach the detector. To increase the resolution of a TOF, mainly for ions with larger masses, a reflectron can be added, which is a series of electrodes that reflect the ions in order to increase the trajectory length, as depicted in Figure 5.5 [124]. The detector used in a TOF is typically a microchannel plate or an electron multiplier, such as a channeltron. Both of these types of detectors work the same way: when an ion hits the detector, it creates a cascade of electrons that is multiplied until a signal is detected [125].

$$E_k = \frac{1}{2} * m * v^2 \tag{5.1}$$

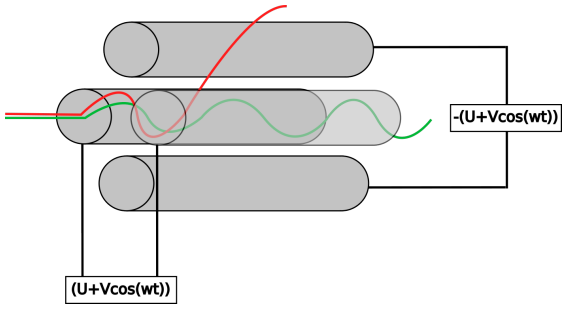


Figure 5.6.: The working principle of a quadrupole mass analyzer showing the stable trajectory of an ion (green) and the unstable trajectory of an ion (red). The two pairs of opposite electrodes are applied the same but opposite voltage with the equation shown where U is the amplitude of the direct current applied, V is the amplitude of the alternating current, w is the frequency of the alternating current, and t is the time.

The most commonly used mass analyzer is a quadrupole, and its basic working principle is depicted in Figure 5.6 [126]. A quadrupole consists of four parallel metal rods that theoretically are of hyperbolic shape. In reality though, cylinders are typically used instead because they are easier and cheaper to produce and approximate well the hyperbolic electric field required. The two pairs of opposing rods are applied with a direct current (U) and an alternating current (V) of equal and opposite amplitude. The alternating current is used in a sinusoidal manner with frequency w. The full equation of the voltages applied to the rods is shown in Equation 5.2. Because equal but oppositely charged voltages are utilized to each pair of electrodes, a saddle-type electric field is formed. Within this electric field, an ion is attracted to one pair of electrodes, while being reflected by the other pair of electrodes. Moreover, by applying a sinusoidal voltage, each pair of electrodes is constantly changing between a positive and negative charge at the specified frequency, thus constantly changing the electric field. Therefore, the rods that were once attracting

the ions are then reflecting them, and so forth, moving the ions in a corkscrew manner. Depending on the amplitude of the U and V voltages applied, only ions with a certain m/z value have a stable trajectory inside the quadrupole without either hitting one of the rods or spinning out of the mass analyzer. If the m/z is too large for the voltages applied, the electric forces acting on the ions from the rods are too weak to change the momentum of the ion; while if the m/z is too small, the electric forces are too strong causing the ion to accelerate too quickly. In both cases, this leads to an unstable trajectory of the ion through the quadrupole. Therefore, only ions with a stable trajectory will pass through the quadrupole and reach the detector. The detectors used are usually the same type as those previously mentioned with TOF mass analyzers. Also, a quadrupole can have only an alternating current applied, allowing it to act as an ion guide instead of a mass filter, thus all ions can pass through it [126].

$$U + V * cos(w * t) \tag{5.2}$$

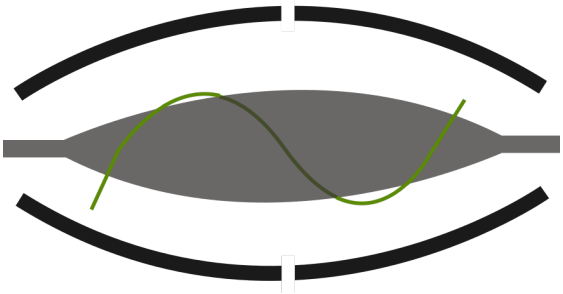


Figure 5.7.: The working principle of an orbitrap mass analyzer with the trajectory of an ion shown in green.

Orbitraps have been growing in popularity since their invention in the 1990s [127]. As shown in Figure 5.7, this mass analyzer has a more complex working mechanism than a quadrupole [128]. An orbitrap is made up of three electrodes: two outer electrodes and one central inner electrode. Ions that enter the orbitrap are trapped between the two outer electrodes and oscillate around the central electrode. Ions of different m/z values oscillate at different frequencies. Therefore, the voltages can be applied in a way to either select only ions of a specific m/z to separate or to remain trapped inside of the orbitrap; or to detect all ions that are inside the orbitrap. When used as a detector, no voltage is applied on the outer electrodes, and the ions oscillating inside the orbitrap then induce frequencies onto the outer electrodes due to their motion. These induced frequencies can be measured and then related back to the m/z values of the ions by Fourier Transformation.

To improve the capabilities of mass spectrometry, two or more mass analyzers can be combined, also known as tandem mass spectrometry (MS/MS or MS^2) [129], [130]. Different combinations are possible, but the three most common types are: quadrupole-orbitrap, quadrupole-TOF, and triple quadrupole. A triple quadrupole is a mass spectrometer

that has three quadrupoles in series. The first one filters a specific m/z (precursor or parent ion), while the second one can either allow all m/z to pass through or act as a collision-induced dissociation cell. If the second quadrupole is used as a collision-induced dissociation cell, a neutral gas is introduced at a certain energy, known as collision energy, causing the ions within to fragment into smaller ions [129]. As fragmentation is an energy based process, the bonds that create more stable fragments are more easily broken, thus a fragmentation pattern based on the relative intensities among the fragments can be created for each molecular structure [131]. The third quadrupole then acts again as a mass filter for specific fragments m/z (product ion). The quadrupole-orbitrap and quadrupole-TOF mass spectrometers both work similarly as a triple quadrupole, but rather than having a third quadrupole at the end, an orbitrap or TOF is used instead [130].

Tandem mass spectrometers are useful as they can be employed in a variety of modes, with the five main modes displayed in Figure 5.8 [129], [132], [133]. The first mode, called full scan, is where the first quadrupole of the MS/MS is in scanning mode, meaning that it acts as a mass filter cycling through a variety of m/z over time, while the other mass analyzers simply allow the ions to pass through. The next mode is called neutral loss ion scan and where the first and third mass analyzers are in scanning mode, and the second mass analyzer acts as a collision cell. The two scanning mass analyzers scan with an offset to see when a certain fragment is lost for all ions entering the instrument, such as a water or sulfate moiety loss. A precursor ion scan again involves the first mass analyzer being in scan mode and the second mass analyzer acting as a collision cell. The third mass analyzer is then set as a mass filter for only specific fragment ions, therefore allowing it to detect only ions with specific fragments. Selected reaction monitoring (SRM), or multiple reaction monitoring (MRM) if more than one SRMs is used at the same time, involves the first and third mass analyzers acting as mass filters for specific ions, while the second mass analyzer acts as a collision chamber. The choice of a parent ion for the first mass analyzer and product ion for the third mass analyzer is also known as a transition. The next mode called product ion scan or data-dependent acquisition (DDA) is where the first mass analyzer filters a specific m/z, the second acts as a collision cell, and the third is in scan mode. Product ion scan allows to see all fragment ions produced from a specific parent ion and their relative intensities to one another enabling to build a fragmentation pattern or MS^2 spectra for the parent ion. The last mode is data-independent acquisition (DIA) where the first and third mass analyzers are in scan mode, and the second mass analyzer acts as a collision cell, allowing to see all fragment ions produced from all ions.

5.3. Types of Data Acquisition and its Analysis

Two different approaches can be used to acquire data from detectors: targeted and untargeted. Each kind has its own advantages and disadvantages, as shown in 5.9, and each requires its own data evaluation procedure.

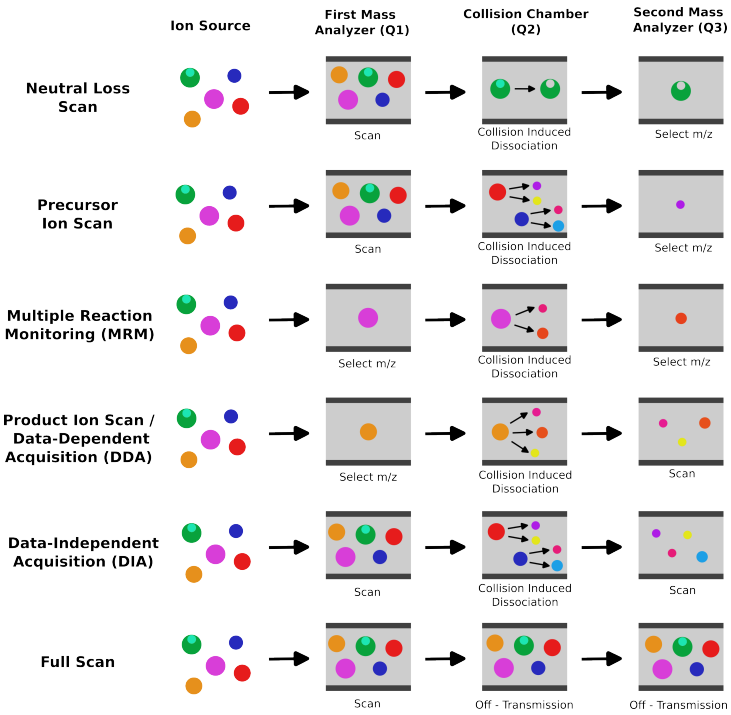


Figure 5.8.: Different acquisition modes that can be used with tandem mass spectrometers.

5.3.1. Targeted Approach

A targeted approach is the more traditional and classical method as it involves the use of authentic reference standards. These standards are applied to develop an analytical workflow, allowing to know the retention times of the analytes in the chromatography step and to compare the signals acquired in the detection step. As the retention times are known, a variety of detectors can be utilized depending on the type and complexity of the sample analyzed. Mass spectrometry can easily be used as a detector for this approach, and the type used typically favors selectivity and sensitivity over resolution as only a certain number of known analytes are investigated, which will ideally be separated in the chromatography step [134]. Currently, triple quadrupole is one of the most sensitive mass spectrometers [122] making it the most commonly used in targeted workflows. The triple quadrupole is typically run in MRM mode with two transitions for each analyte. Both transitions of each analyte have the same parent ion mass, but different product ion masses, whereby one transition is labeled as the quantifier while the other as the qualifier. Multiple transitions for the same analyte are selected to improve selectivity since for a certain molecular mass, different chemical formulas are possible, especially as the mass becomes larger. Moreover, different molecules of the same mass can have one or more similar fragments. Therefore, using multiple transitions allows to increase the certainty that the desired analyte is observed, because it is unlikely that different molecules have similar fragments with the same ratios among the fragment intensities, unless they are

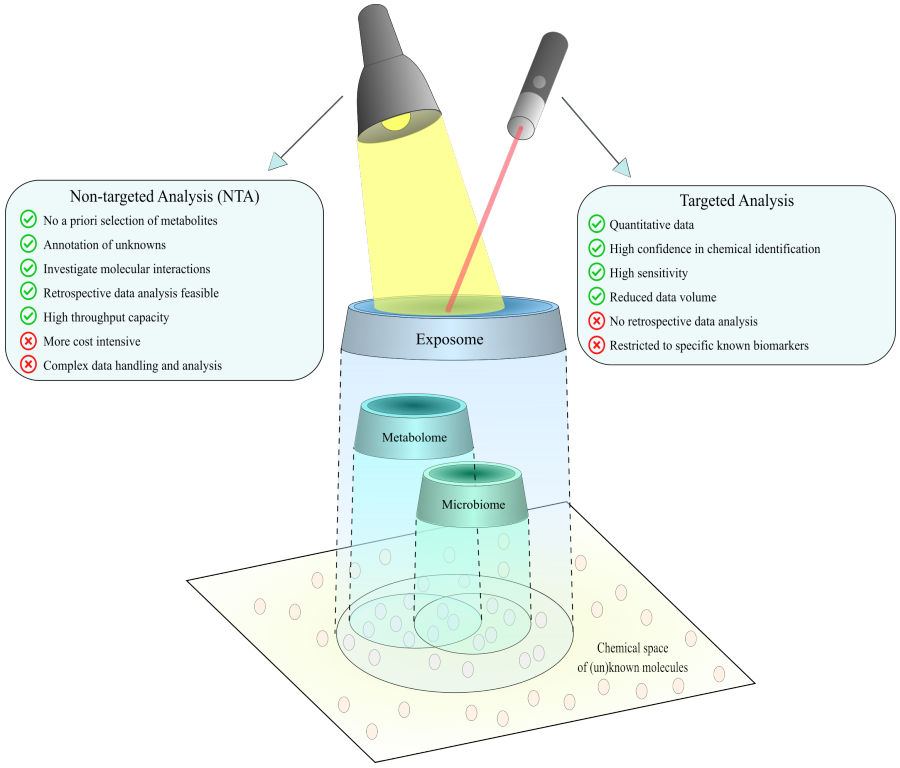


Figure 5.9.: The advantages and disadvantages summarized and compared between targeted and untargeted approaches. Reproduced with permission from the Annual Review of Food Science and Technology, Volume 12 © 2021 by Annual Reviews, http://www.annualreviews.org [15].

isomers. Thus, the ion ratio, meaning the intensity ratio between the two transitions for an analyte, is an important factor to consider. This ratio is a means of confirming the presence of an analyte. Additionally, for each transition, a variety of parameters of the triple quadrupole, e.g. collision energy, can be optimized to improve sensitivity.

In a targeted workflow, the different parameters, e.g. retention times, ion ratios, or MRM transitions, are determined and optimized with reference standards. Once the targeted method is optimized, the method is validated in order to evaluate other key criteria. These criteria prove the stability and suitability of the method, and are used to correct quantification results. They can be calculated in different ways depending on the established guidelines followed during method validation [135]. The first criteria was already previously mentioned and involves determining the selectivity and sensitivity of the method. The selectivity is determined by comparing blank samples with samples enriched with the analyte of interest to observe the difference in signal. Then, the sensitivity, typically defined by the limit of detection, is calculated to find the minimum concentration needed of an analyte to make it detectable. A further criterion is the limit of quantification, which is the minimum concentration needed to be able to quantify

the analyte, as even though a signal and the presence of an analyte can be observed, it is not always feasible to quantify the analyte properly. The next criteria is the signal suppression and enhancement effect. This is the impact interferents present in the sample have on the analyte ionization efficiency, and it is usually determined by comparing the signal of the analyte in a neat solvent to its signal in the sample matrix. Another criteria is extraction efficiency, which involves determining the percentage of the analytes lost during the sample preparation. The last key criteria is the precision of the measurements, which is usually determined by multiple replicate measurements of the same sample and of samples prepared the same way.

5.3.2. Untargeted Approach

The other data acquisition approach is called untargeted, and it is a broader and more complex approach that allows to investigate a larger variety, including potentially unknown, molecules. In an untargeted approach, rather than looking at only a few specific standards, everything ionizable in a sample is measured during the entire chromatography run. Therefore, the complexity of analyzing the data is greatly increased as three dimensions are investigated: m/z, retention time, and intensity of the m/z. In addition, it is unknown from which molecules the signals originate from, as analytes beyond those with reference standards readily available are investigated. Therefore, increasing the accuracy of the determined m/zs is vital to aid in identification, thus mass spectrometers with high-resolution are employed such as quadrupole-TOF or quadrupole-orbitrap [136].

In untargeted workflows, the first step is to acquire a full scan, or MS^1 scan, in order to detect which parent masses are present, and at what intensity and retention time. Then, the mass spectrometer is run in one of two modes, DIA or DDA, to acquire fragmentation spectras which is needed for molecular identification [137], [138]. In DDA, a specific parent mass is selected by the first mass analyzer to be fragmented; while in DIA everything is fragmented. DIA has the advantage that once the data is acquired, no matter which feature is of interest, the fragmentation spectra can always be investigated retrospectively without having to re-acquire the samples. The drawback of DIA is that data deconvolution is complex, and the fragmentation spectra typically have more noise. In contrast, fragmentation spectra in DDA can be acquired either where a certain number of the most intense ions of each MS^1 scan is fragmented, or in case the parent masses are already known, with an inclusion list. An inclusion list is a list of parent masses and when the relevant masses are detected in a MS^1 scan, they are selected and acquired for fragmentation. The retention times of the relevant masses can also be given if they are known. The disadvantage of an inclusion list is that the parent masses need to be known beforehand. Thus, if features of interest are determined after sample acquisition, the samples would need to be re-acquired to get their fragmentation spectra. During the first acquisition of the samples, certain molecules of interest can be screened for their presence by including their parent masses in the inclusion list. For most mass spectrometers, the top most intense ions and inclusion list can be acquired at the same time if desired.

Untargeted approaches can be further broken down into two main types of workflows: suspect screening and non-targeted analysis [139]. In suspect screening, molecules that

are suspected to be present in a sample are searched for. Fragmentation spectra for suspects can either be done by creating an inclusion list with the suspects, as mentioned previously, or by other untargeted data acquisition techniques. For annotation of these molecules, their experimentally acquired fragmentation spectra is matched with reference spectra available in a spectral library. In non-targeted analysis, everything is studied, in order to, e.g., discover new biomarkers of exposure. Different techniques are available such as deploying statistical analysis to find features of interest or molecular networking to find similar features.

5.3.3. Evaluation of Targeted MS/MS Data

Data analysis with results from targeted acquisition is comparatively straight forward. The MRM transitions are imported along with the raw data files in a program, such as Skyline [140], and the chromatographic peaks are integrated. To confirm the presence of an analyte in unknown samples, the retention time, intensities of the transitions and ion ratio between the transitions are examined and compared to samples enriched with the standards. Moreover, quantification of unknown samples can be applied using both internal and external calibration curves [135]. Typically, an internal standard is utilized to normalize for any variations that can occur between the individual samples. Then, quantification of the analytes in the unknown samples is done with an external calibration curve that involves enriching a blank sample at different concentrations. Overall, targeted workflows are useful when specific molecules to investigate have been determined, such as biomarkers for polyphenol exposure, and the goal is to characterize and quantify these analytes in samples of interest with high confidence and sensitivity.

5.3.4. Bioinformatic Data Processing of HRMS Data

The acquired data from untargeted workflows is complex and usually quite large, containing lots of information. Therefore, in order to solve the desired analytical goals, several steps are required. The first step involves extracting features from the raw data [141]. The raw data needs to be pre-processed, by converting to the right file format depending on the program used, and the noise level needs to be set. Choosing the right noise level is complicated, because when setting a high level, it is more certain that the signal comes from actual analytes and not noise, though only abundant analytes will be retained. However, if a low noise level is selected, more analytes are included, especially lower abundant analytes, but also more noise [142]. Next, all the features are extracted from the raw data by building extracted ion chromatograms for each mass to detect chromatographic peaks [143], [144]. A variety of algorithms can be used, e.g. continuous wavelet transformation [145]. In addition, as co-elution occurs or detected features can overlap, the features need to be resolved in order to separate these co-eluting features. Redundant features, such as those created from isotopes, also need to be removed. These steps will result in a feature list for each acquired sample. However, as there may be slight variations in both detected masses and retention times between each sample, in order to combine together the individual feature lists of the different samples, they need to be

aligned [146]. Lastly, the aligned feature list should be gap-filled to correct for certain features that are absent in various samples from the data processing steps [143]. A variety of parameters can be fine-tuned at each step when extracting the features, whereby the number and types of parameters that can be modified will depend on the software used. Moreover, stringent parameters will create features from only highly abundant analytes with ideal chromatographic peak shapes. In case less strict parameters are selected, lower abundant analytes with less ideal peak shapes are also extracted, but with the drawback that a higher number of the features created are in reality noise. Hence a compromise needs to be found among all the parameters depending on the overall objective.

In a typical untargeted workflow, thousands of features can be extracted as all ionizable molecules entering the detector are acquired. Thus, the next step involves applying quality assurance and control techniques. This will allow to, for example, remove features that originate from the background, such as those arising from the solvents used or chemical residues from the sample preparation [147]. Moreover, during quality control, features are corrected for different variations, e.g. inter-batch effects [148]. Typically, this still leaves an abundant number of features, thus depending on the experimental set-up and objectives, different approaches can be applied to extract features of interest. The most common approach is statistical analysis as commonly an experiment is set-up so that the samples acquired have some different conditions, e.g. one group of participants receiving medication and the other a placebo. Various statistical tests can be applied depending on the experimental set-up and the number of variables, such as analysis of variance to find features which show statistical significance [149].

While features of interest have now been extracted, their identity is yet unknown, and feature identification is still a challenge. As mentioned previously, knowing only the parent mass is not sufficient to determine the analyte´s formula or structure because multiple possibilities exist, especially as the mass becomes larger. Therefore, at least the fragmentation spectra (MS^2) of the features need to be acquired. Ideally, a reference standard is available to compare the retention time and the MS^2 spectra with the acquired feature. Unfortunately, this is usually not the case as obtaining a lot of reference standards is costly and the objective of untargeted analysis is to screen for all possibilities. Therefore, two other techniques are typically used to determine the analyte from the MS^2 spectra [150]. The first is spectral library matching, which involves matching the acquired MS^2 spectra with a database containing MS^2 spectra. These databases contain experimentally acquired spectra of standards from a variety of sources and instruments, such as from other research groups. The second technique is *in silico* fragmentation, which involves using computational tools based on different principles, e.g. machine learning or quantum chemistry [150], to determine the analyte from the acquired spectra. Usually, spectral library matching gives more confidence in the analyte annotation than *in silico* fragmentation, but one thing to note is that each mass spectrometer and/or brand comes with slight differences in their design so the MS^2 spectra may vary between instruments. Therefore, the most reliable results are obtained when matching the acquired spectra with spectral library spectras from the same instrument. The complexity of feature annotation has led to an accepted system that defines five different levels of confidence

for the structural identification of a feature depending on criteria met [151]. Level 1 is the most confident identification, meaning that a reference standard was used. Level 2 is for a probable structure through, e.g. annotation using spectral libraries. Level 3 is then a tentative structural candidate based on experimental data. Lastly, if no structure can be proposed, the feature is identified either as Level 4 if a molecular formula is known and Level 5 if only the exact mass is known.

Unlike in targeted workflows, quantification of the features extracted is a key impediment of untargeted workflows because molecules beyond those that reference standards are available on hand are investigated. Therefore, for the features that do not have a reference standard available, building a calibration curve is nearly impossible. Relative quantification can be used to overcome this problem [152]. In relative quantification, a feature's chromatographic peak area is compared among samples (as used during statistical analysis). Similarly, a calibration curve can be created from a reference standard and used to quantify analytes with similar structures as the reference standard.

5.4. Current Challenges in Investigating Polyphenol Exposure

Though LC-MS is currently the best technique to investigate polyphenols, there are still many obstacles that need to be overcome, especially as polyphenols are a vast chemical class that produces numerous metabolites. Therefore, having reference standards for a large selection of polyphenols is difficult because for many polyphenols, notably for conjugated polyphenols, reference standards are either not available or expensive [15]. Hence, a move from targeted methods to untargeted platforms is a must in order to comprehensively investigate polyphenols.

Previously, polyphenols present in plants or fungi were analyzed with targeted methods utilizing available reference standards that mainly consisted of phenolic acids or several flavonoid aglycones. Targeted methods with the standards for the polyphenols known to be present in the foods of one's diet were utilized to study human polyphenol exposure. Comprehensively investigating polyphenols present in both food and humans was not as common. With recent advances in technology and growing interest in polyphenols for human wellbeing, new targeted and untargeted workflows have been and are being developed. However, investigating polyphenols with untargeted workflows is complicated as a vast number of isomers exists and their MS^2 spectra are relatively similar, making annotation challenging [18]. This is especially true for conjugated polyphenols where a variety of positional isomers are possible, all having extremely similar MS^2 spectra. Moreover, as mentioned before, there is still an insufficiency of available polyphenol reference standards, making their quantification a challenge, meaning there would also be a lack in available MS^2 spectra in spectral libraries. Therefore, there is a need to further develop both targeted and untargeted workflows that investigate all different types of polyphenols such that it can easily be applied and transferred to various biological matrices, coming from both human and plant origin, to better study their presence and potential impact.

6. Original Works

6.1. Overview of the Original Research and Review Articles Included in this Thesis

Publication #1 (Review): Polyphenol Exposure, Metabolism, and Analysis: A Global Exposomics Perspective
Ian Oesterle, Dominik Braun, David Berry, Lukas Wisgrill, Annette Rompel, Benedikt Warth. "Polyphenol Exposure, Metabolism, and Analysis: A Global Exposomics Perspective", *Annual Reviews in Food Science and Technology*, **2021**, Vol. 12:461-484, DOI: 10.1146/annurev-food-062220-090807

Publication #2 (Original Research): Quantifying up to 90 polyphenols simultaneously in human bio-fluids by LC-MS/MS
Ian Oesterle, Dominik Braun, Annette Rompel, Benedikt Warth. "Quantifying up to 90 polyphenols simultaneously in human bio-fluids by LC-MS/MS", *Analytica Chimica Acta*, **2022**, Vol. 1216:339977, DOI: 10.1016/j.aca.2022.339977

Publication #3 (Original Research): Exposomic Biomonitoring of Polyphenols by Non-Targeted Analysis and Suspect Screening
Ian Oesterle, Manuel Pristner, Sabrina Berger, Mingxun Wang, Vinicius Verri Hernandes, Annette Rompel, Benedikt Warth. "Exposomic Biomonitoring of Polyphenols by Non-Targeted Analysis and Suspect Screening", *Analytical Chemistry*, **2023**, Vol. 95(28):10686-10694, DOI: 10.1021/acs.analchem.3c01393

Publication #4 (Original Research): Insights into the early-life chemical exposome of Nigerian infants and potential correlations with the developing gut microbiome
Ian Oesterle, Kolawole I. Ayeni, Chibundu N. Ezekiel, David Berry, Annette Rompel, Benedikt Warth. "Insights into the early-life chemical exposome of Nigerian infants and potential correlations with the developing gut microbiome", *bioRxiv*, **2023**, DOI: 10.1101/2023.11.08.566030

Publication #5 (Original Research): Polyphenolic Profiling of Plants and Edible Mushrooms to Aid Characterizing Polyphenol Oxidase Selectivity
Ian Oesterle, Mathias Pretzler, Annette Rompel, Benedikt Warth. "Polyphenolic Profiling of Plants and Edible Mushrooms to Aid Characterizing Polyphenol Oxidase Selectivity", *ChemRxiv*, **2023**, DOI: 10.26434/chemrxiv-2023-phq9f

6.2. Publication #1 (Review): Oesterle et al. 2021

Status	Published
Title	**Polyphenol Exposure, Metabolism, and Analysis: A Global Exposomics Perspective**
Authors	Ian Oesterle[1,2], Dominik Braun[1], David Berry[3,4], Lukas Wisgrill[5], Annette Rompel[2], Benedikt Warth[1]
Affiliations	[1]Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, 1090, Vienna, Austria [2]Department of Biophysical Chemistry, Faculty of Chemistry, University of Vienna, 1090 Vienna, Austria [3]Centre for Microbiology and Environmental Systems Science, Department of Microbiology and Ecosystem Science, Division of Microbial Ecology, University of Vienna, 1090 Vienna, Austria [4]The Joint Microbiome Facility of the Medical University of Vienna and the University of Vienna, 1090 Vienna, Austria [5]Division of Neonatology, Pediatric Intensive Care and Neuropediatrics, Department of Pediatrics and Adolescent Medicine, Medical University of Vienna, 1090 Vienna, Austria
Year	2021
Journal	*Annual Review of Food Science and Toxicology*
Journal Information	Vol. 12:461-484
Accepted	December 22, 2020
DOI	https://doi.org/10.1146/annurev-food-062220-090807
Contribution	Ian Oesterle performed literature research, and took the lead in discussion and summarizing the ideas presented in the work, and in writing the manuscript.

ANNUAL REVIEWS

Annual Review of Food Science and Technology

Polyphenol Exposure, Metabolism, and Analysis: A Global Exposomics Perspective

Ian Oesterle,[1,2] Dominik Braun,[1] David Berry,[3,4] Lukas Wisgrill,[5] Annette Rompel,[2] and Benedikt Warth[1]

[1]Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, 1090 Vienna, Austria; email: ian.oesterle@univie.ac.at, dominik.braun@univie.ac.at, benedikt.warth@univie.ac.at

[2]Department of Biophysical Chemistry, Faculty of Chemistry, University of Vienna, 1090 Vienna, Austria; email: annette.rompel@univie.ac.at

[3]Centre for Microbiology and Environmental Systems Science, Department of Microbiology and Ecosystem Science, Division of Microbial Ecology, University of Vienna, 1090 Vienna, Austria; email: david.berry@univie.ac.at

[4]The Joint Microbiome Facility of the Medical University of Vienna and the University of Vienna, 1090 Vienna, Austria

[5]Division of Neonatology, Pediatric Intensive Care and Neuropediatrics, Department of Pediatrics and Adolescent Medicine, Medical University of Vienna, 1090 Vienna, Austria; email: lukas.wisgrill@meduniwien.ac.at

Annu. Rev. Food Sci. Technol. 2021. 12:461–84

First published as a Review in Advance on December 22, 2020

The *Annual Review of Food Science and Technology* is online at food.annualreviews.org

https://doi.org/10.1146/annurev-food-062220-090807

Keywords

polyphenols, human biomonitoring, food metabolomics, microbiome, food bioactives, biotransformation, exposome

Abstract

Polyphenols are generally known for their health benefits and estimating actual exposure levels in health-related studies can be improved by human biomonitoring. Here, the application of newly available exposomic and metabolomic technology, notably high-resolution mass spectrometry, in the context of polyphenols and their biotransformation products, is reviewed. Comprehensive workflows for investigating these important bioactives in biological fluids or microbiome-related experiments are scarce. Consequently, this new era of nontargeted analysis and omic-scale exposure assessment offers a unique chance for better assessing exposure to, as well as metabolism of, polyphenols. In clinical and nutritional trials, polyphenols can be investigated simultaneously with the plethora of other chemicals to which we are exposed, i.e., the exposome, which may interact abundantly and modulate bioactivity. This research direction aims at ultimately eluting into a

true systems biology/toxicology evaluation of health effects associated with polyphenol exposure, especially during early life, to unravel their potential for preventing chronic diseases.

INTRODUCTION

Polyphenols are a large class of bioactive compounds present in plant-based foods that each contains at least one phenyl ring and one hydroxyl group. These molecules can be divided into two main classes, flavonoids and nonflavonoids, with either of these two classes containing a variety of subclasses, as shown in **Figure 1**, such as flavonols, anthocyanins, and isoflavones (Crozier et al. 2009, del Rio et al. 2012).

The presence of dietary polyphenols has been described in numerous plants, vegetables, and other food sources, such as coffee or chocolate, each having a different polyphenolic profile. Notably, 452 food items containing a total of 502 different polyphenols can be found in the Phenol-Explorer database, an important resource for polyphenol content in foods, metabolism, and fate during food processing (Neveu et al. 2010). For instance, in leaves of *Moringa oleifera* up to 291 different polyphenols have been annotated (Rocchetti et al. 2020). The phenolic profile may vary slightly for the same food item depending on factors such as region, climate, and soil. This was, for example, demonstrated in the *M. oleifera* leaves from trees that were grown in different regions of China (Zhu et al. 2020).

Polyphenols have been described as protective agents for neurodegenerative and other chronic diseases, mainly via their antibacterial, anticancer, and anti-inflammatory properties (Shahidi & Yeo 2018). The health and protective effects of polyphenols are related to not only their antioxidant activity but also a variety of other modes of action, such as modulatory interaction with enzymes or receptors (Figueira et al. 2017). For instance, ferulic acid may induce anticancer activity by affecting the behavior of cells in a human pancreatic cancer cell line by changing the expression of certain apoptosis and cell cycle genes (Fahrioğlu et al. 2016). Another example is kaempferol, which displays antioxidant activity by reducing reactive oxygen species and anti-inflammatory activity by inhibiting proinflammatory enzymes such as cyclooxygenase-1 and -2 enzymes (Devi et al. 2015).

However, polyphenols can also have prooxidant activity, especially when high amounts of certain polyphenols are present, e.g., when complementing a diet with concentrated polyphenols (Granato et al. 2020, Martin 2009). The prooxidant activity of polyphenols is also shown to occur when redox-active metals are present (Sulpizio et al. 2018) and can increase the formation of reactive oxygen species, leading to potential DNA damage (Anantharaju et al. 2016, Kyselova 2011). However, the prooxidant effect of polyphenols can in turn help activate and increase the production of enzymes and proteins that protect against the damaging effects of reactive oxygen species (Roos & Duthie 2015).

The traditional approach of human biomonitoring (HBM), i.e., measuring a dietary or environmental exposure in a biological fluid, is currently being extended by so-called exposomic approaches. Here, the intention is to go beyond single biomarkers and assess chemical exposures at the omic scale. Thus, exposomic research involves studying a vast number of biomarkers from human matrices to assess and better understand complex exposures and their impact on health and disease (Dennis et al. 2017). Importantly, newer definitions of the exposome, i.e., the totality of lifetime exposures, include the measurable biological response to these exposures. Hence, the metabolome (here defined as the totality of endogenous human metabolites) is a relevant part of the exposome for studying the various chemical pathways and processes as a response that

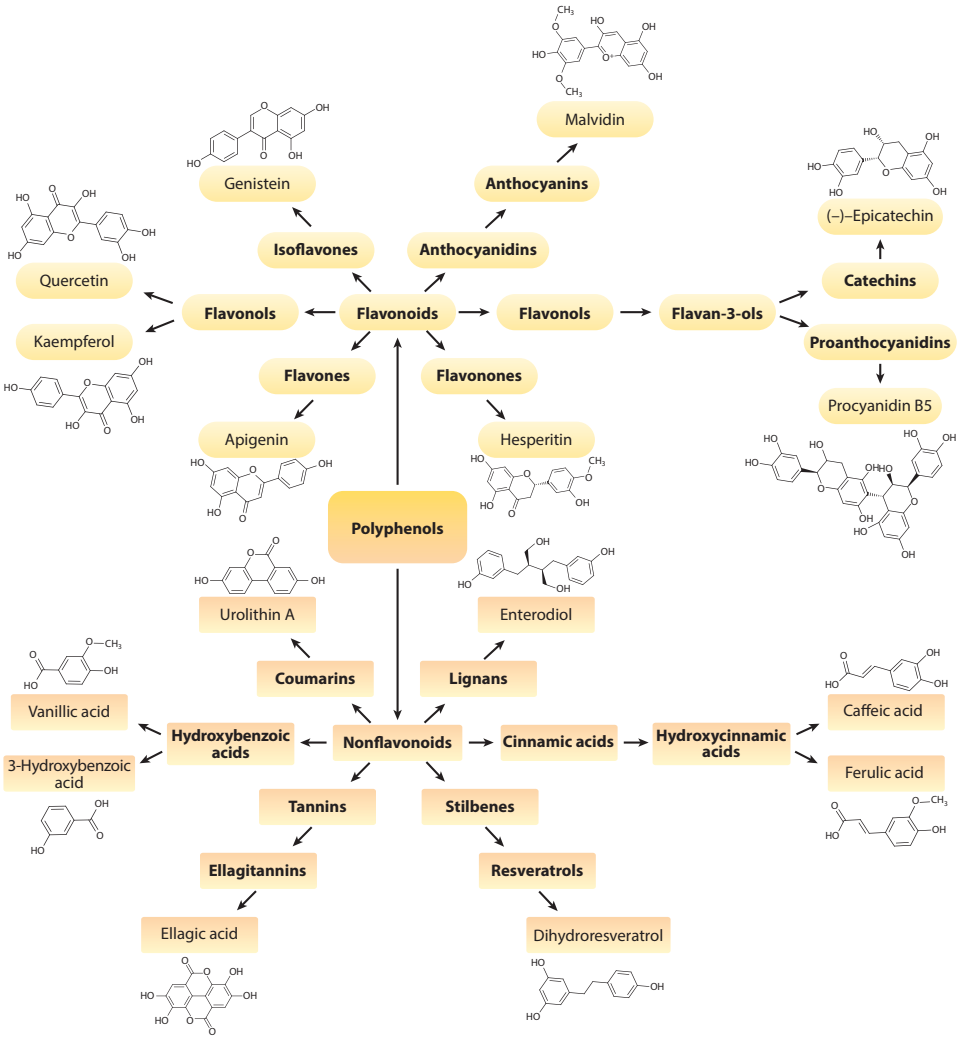


Figure 1

Flow chart depicting polyphenol classes and selected examples of each class.

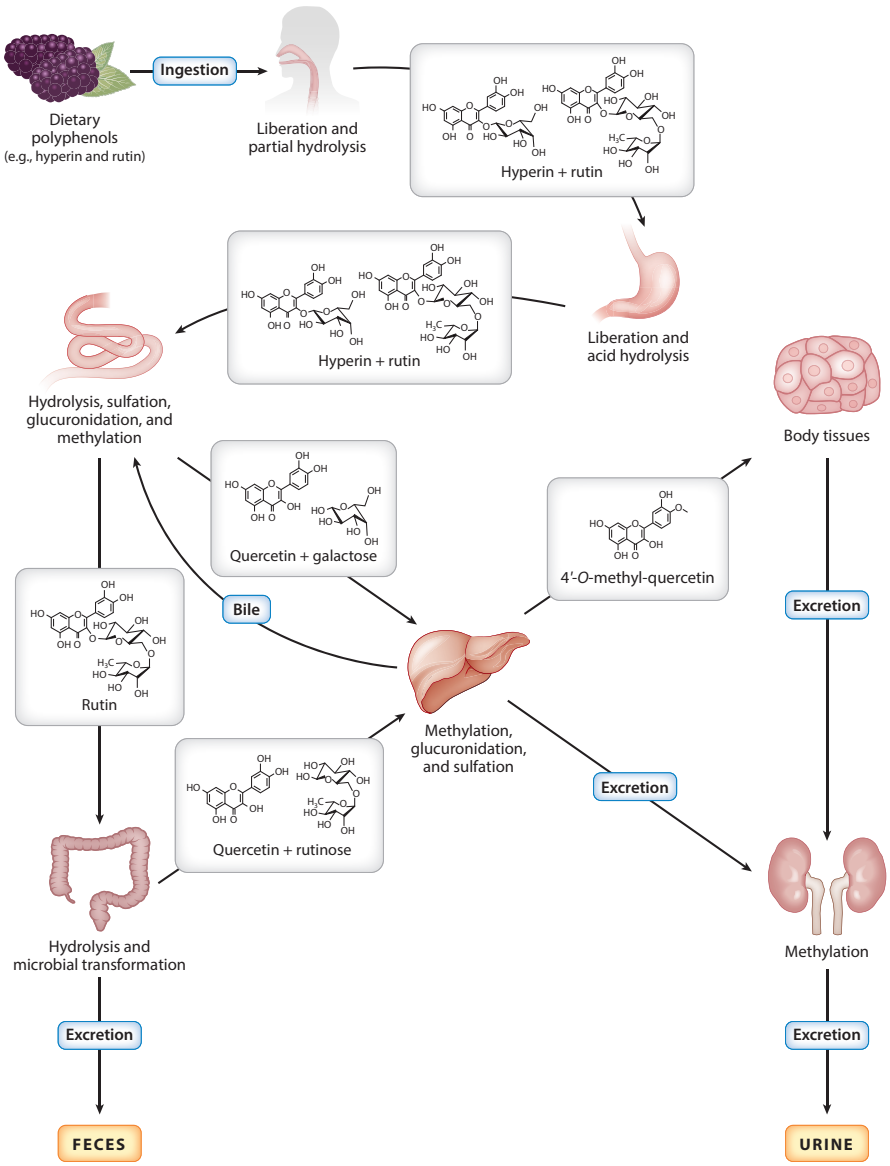
can help to understand bioactivity and toxicity. According to the Human Metabolome Database (HMDB), there are more than 110,000 different metabolites found in the human body, including those formed by the microbiota (Wishart et al. 2018). Therefore, due to the numerous potential effects of polyphenols, both beneficial and adverse, and the copious amount of at least 50,000 polyphenols present in plants, of which 8,000 were identified as dietary polyphenols (Ziaullah & Rupasinghe 2015), research initiatives for better characterization of exposure and health impact are warranted.

HBM typically applies targeted analytical approaches utilizing authentic reference standards for accurate quantification, whereas in exposomics, mostly untargeted workflows, often referred to as nontargeted analysis/screening (NTA/NTS), are developed and applied. A targeted approach involves analyzing specific biomarkers, which are often metabolites that represent exposure and/or effect. Contrarily, the untargeted approach involves studying all measurable metabolites present in a sample (Dennis et al. 2017). Both analytical approaches are currently applied in polyphenol research, and clearly both come with prospects and challenges that are discussed in this article. Moreover, a variety of analytical techniques can be utilized for both these approaches. The techniques typically involve using either liquid or gas chromatography for separation, followed by measuring the analytes by a connected detector. Diode array or ultraviolet/visible-light detectors were previously used extensively, despite challenges in metabolite identification (Scalbert et al. 2009). However, with the advancement of mass spectrometry (MS) and nuclear magnetic resonance (NMR), these detectors have become more common. Even though NMR detectors have the advantages of being nondestructive, highly reproducible, and require almost no sample preparation, NMR is less selective and sensitive and can detect a lower number of metabolites simultaneously compared to MS detectors (Emwas 2015). Thus, the focus in this article is on MS-based detectors, mainly those with liquid chromatography (LC) separation because of its high sensitivity and ability to analyze a variety of compounds (Shao et al. 2019). In addition, the application of new analytical methods and the impact of polyphenols on the gut microbiome are reviewed. Finally, future perspectives for polyphenol research, especially in the context of infant health and early-life exposure, are highlighted.

BIOMONITORING OF POLYPHENOLS IN HUMAN SPECIMENS

The metabolism of polyphenols is an intricate process, as an array of dietary polyphenols are found in various food sources, and, once ingested, they undergo a range of metabolic processes. This is, for example, reflected in the Phenol-Explorer database, where 375 different polyphenol metabolites found in urine and plasma from humans and animals are currently listed (Rothwell et al. 2013). Therefore, choosing the right biomarkers, if possible representing absorption and the effect of different polyphenols in the human body, is a critical endeavor.

As depicted in **Figure 2**, ingested dietary polyphenols can undergo two main types of biotransformations: phase I, which includes functionalization reactions such as hydrolysis or oxidation, and phase II, which includes detoxifying conjugation reactions such as glucuronidation. These transformations occur mainly in organs such as the liver or small intestine either before or after their tissue absorption or are caused by the gut microbiota found in the large intestine, yielding an assortment of microbial metabolites (Scalbert et al. 2014). For instance, ellagitannins are first partially metabolized by acid hydrolysis in the stomach or proximal small intestine, forming ellagic acid. Then, in the small intestine or colon, the gut microbiota further metabolizes the ellagitannins and ellagic acid into urolithins, which are absorbed and can undergo phase II transformations by glucuronosyltransferases and/or methyltransferases (Crozier et al. 2009). Other biotransformation reactions that may modify phenolic acids, such as cinnamic acids, are methylation in the small



(Caption appears on following page)

Figure 2 (*Figure appears on preceding page*)

Scheme of human polyphenol metabolism and biotransformation reactions, such as sulfation and methylation, that can occur, exemplified by the deconjugation and transformation of the flavonoid quercetin (Almeida et al. 2018, Heleno et al. 2015, Marín et al. 2015).

intestine and sulfation in the liver (Heleno et al. 2015). Generally, deciphering the full degree of polyphenol metabolism in the body is a complex task.

Targeted Analytical Approaches

There are two main ways targeted approaches have been used with polyphenols in HBM. First, because each foodstuff has its own polyphenolic profile, the exposure to a certain food source can be studied by looking at the biotransformation products of the polyphenols that are present in higher concentrations in that specific food source. Second, certain biotransformation products or unmetabolized polyphenols can be targeted to analyze how different types of exposures may change their concentrations in the body. An example of a targeted polyphenol metabolomics approach would be to measure levels of certain phenolic acid metabolites, such as isoferulic acid, to quantify the intake of tea or coffee (Hodgson et al. 2004). This approach can be used to better evaluate the validity of self-reported food intake data from study participants via food frequency questionnaires or food diaries, as these are often erroneous and do not provide quantitative data. Typically, for targeted approaches, reference standards are used for the unambiguous identification and quantification of either the parent molecule (typically after enzymatic deconjugation) or the respective metabolic products (frequently glucuronide and/or sulfate conjugates). However, because of a lack of commercially available reference standards for the conjugated forms, often only the aglycones are measured or used as reference standards for (semi)quantification in urine, although it is known that many compounds are fully conjugated. Also, if enzymatic deconjugation is applied, certain limitations apply. Frequently, the enzymatic conversion is incomplete, and the efficiency of deconjugation may differ between different polyphenols or even for the same polyphenol if it is conjugated at different positions. This can result in inaccurate quantification (Ottaviani et al. 2018). Furthermore, because of the choice of biomarkers, certain metabolites might not be detected by certain targeted methods, making the calculation of factors such as the bioavailability of polyphenols more difficult (Manach et al. 2005). Estimation errors can occur and certain metabolites can be missed because, even though polyphenols exist in a range of food sources, they are sometimes present at low concentrations, and as a consequence, their respective concentrations in human plasma or urine are even lower, typically below 1 nmol/L (Gleichenhagen & Schieber 2016).

Many methods for investigating polyphenol metabolites in human matrices have been published and are reported in **Table 1** (targeted) and **Table 2** (untargeted). Most of these apply a targeted approach with 2–50 different analytes included, and only a limited number of studies investigated more than 30 polyphenol metabolites, with the current highest number investigated being 54 metabolites (Zhong et al. 2017). However, this method also highlights some typical limitations that exist in quantification and can occur in targeted approaches. It employs neither deconjugation nor any reference standards for conjugated metabolites. Therefore, as mentioned above, it is very likely to severely underestimate exposure because most polyphenol metabolites are present in their conjugated forms following ingestion and biotransformation. Currently, the method with the highest number of quantifiable metabolites determines 46 analytes, including several glucuronide and sulfate conjugates; however, these metabolites were selected, as they are potential marker molecules for orange juice consumption (Ordóñez et al. 2018). Thus, there is still

Table 1 Selected MS-based methods for the targeted analysis of polyphenols and their metabolites in human-derived matrices

Matrix	Number of polyphenol analytes	Number of polyphenols quantified	Sample preparation	Enzyme treatment	Chromatography	Detection	Time (min)	LOQ (ng/mL)	LOQ (nM)	Reference
Plasma	10	5	Protein precipitation	None	UHPLC	ESI (negative)–Q-Orbitrap–MS/MS (PRM)	19	1.38–5.48	6–24	Svilar et al. 2019
Plasma	21	18	SPE	None	UHPLC	ESI (positive and negative)–QqQ-MS/MS (MRM)	22 (anthocyanin metabolites and urolithins) and 16 (phenolic acid metabolites)	NA	NA	Sandhu et al. 2018
Plasma	38	38	Enzymatic hydrolysis, followed by solvent extraction and $^{13}C_2$ and $^{12}C_2$ dansylation	β-Glucuronidase type H-1 (EC 3.2.1.31) from *Helix pomatia*	UHPLC	ESI (positive)–QTrap–MS/MS (MRM)	20.5	0.03–13.9	0.11–44	Achaintre et al. 2018
Plasma	54	20	SPE then centrifugation	None	UHPLC	ESI (positive and negative)–QTOF–MS/MS	52 (anthocyanin metabolites) and 16 (phenolic acid metabolites)	NA	NA	Zhong et al. 2017
Urine	11	0	Enzymatic hydrolysis, SPE, and silylation	β-Glucuronidase type H-3 (EC 3.2.1.31) and sulfatase type V (EC 3.1.6.1)	GC	EI–MS	70	NA	NA	Rios et al. 2003
Urine	11	11	Enzymatic hydrolysis, solvent extraction, and centrifugation	β-Glucuronidase and sulfatase from *H. pomatia*	HPLC	ESI (negative)–QqQ–MS/MS (MRM) and DAD (for hippuric acid)	15	NA	NA	Rios et al. 2003
Urine	15	15	Enzymatic hydrolysis, solvent extraction, and centrifugation	β-Glucuronidase type H-2 from *H. pomatia*	HPLC	ESI (negative)–QqQ–MS/MS (MRM)	6	5.97–1,450	20–5 × 10^3	Ito et al. 2005

(*Continued*)

Table 1 (*Continued*)

Matrix	Number of polyphenol analytes	Number of polyphenols quantified	Sample preparation	Enzyme treatment	Chromatography	Detection	Time (min)	LOQ (ng/mL)	LOQ (nM)	Reference
Urine	21	21	Enzymatic hydrolysis, SPE	β-Glucuronidase and sulfatase from *H. pomatia*	HPLC	ESI (negative)–QqQ–MS/MS (MRM)	10	0.5–100	1.68–558	Urpi-Sarda et al. 2009
Urine	28	28	SPE (with SDB-L or HLB cartridges) and silylation	None	Split GC	EI–MS	44	400–15,600	2,210–79,900	Ordóñez et al. 2018
Urine	41	18	SPE	None	HPLC	ESI (positive)–Q'Trap–MS/MS (MRM)	52	NA	NA	Kalt et al. 2017
Urine	43	24	Protein precipitation, enzymatic hydrolysis, and membrane filtration	β-Glucuronidase and sulfatase type H-2 from *H. pomatia*	UHPLC	ESI (positive and negative)–Q'Trap–MS/MS (SRM)	10	0.05–40.0	0.13–263	Magiera et al. 2012
Urine	46	46	SPE with SDB-L or HLB cartridges, or direct injection	None	UHPLC	ESI (negative)–Q-Orbitrap	45	9.87–807	27–5,920	Ordóñez et al. 2018
In vitro fecal fermentation (using pig samples)	6	6	Centrifugation and cellulose syringe filtration	None	UHPLC	ESI (positive)–Q-Orbitrap	13	NA	NA	Rocchetti et al. 2019
In vitro fermentation by human gut microbiota	18	18	Solvent extraction by centrifugation	None	UHPLC	ESI (negative)–QqQ–MS/MS (dMRM)	18	0.69–8.11	4.2–60	Zhao et al. 2018
In vitro fecal fermentation	18	18	Solvent extraction and silylation	None	Split GC	QqQ–MS/MS (MRM)	38	0.48–12.4	3.48–69.0	Carry et al. 2018
In vitro fecal fermentation	27	3	Solvent extraction by centrifugation	None	HPLC	ESI (negative)–Q'Trap–MS/MS (MRM)	38	NA	NA	de Santiago et al. 2019
Batch-culture fecal fermentation	30	6	Centrifugation and PVDF filtration	None	UHPLC	DAD–ESI (negative)–QqQ–MS (MRM)	18	9–200	15.6–231	Sánchez-Patán et al. 2012

(*Continued*)

Table 1 (*Continued*)

Matrix	Number of polyphenol analytes	Number of polyphenols quantified	Sample preparation	Enzyme treatment	Chromatography	Detection	Time (min)	LOQ (ng/mL)	LOQ (nM)	Reference
Plasma and urine	6	3	SPE and lipid removal for plasma samples	None	HPLC	ESI (positive)–QTrap–MS	68	NA	NA	Wu et al. 2002
Plasma and urine	8	1	SPE and enzymatic hydrolysis for quantification	β-Glucuronidase type VII-A and sulfatases type VIII and H-5	HPLC	ESI (negative)–QqQ–MS/MS (MRM)	9	13.3	45.8	Roura et al. 2005
Plasma and urine	14	5	SPE	None	HPLC	ESI (positive and negative)–QTrap–MS/MS (full scan) and ESI (negative)–QTOF–MS (full scan)	45	NA	NA	Teixeira et al. 2017
Plasma and urine	15	9 (8 standards were synthesized in-house)	Protein and phospholipid removal and filtration; half of the samples then went through enzymatic hydrolysis	β-Glucuronidase type H-1 and sulfatases from *H. pomatia*	UHPLC	ESI (negative)–QqQ–MS/MS (MRM)	8	5.6–14 (plasma) and 222–420 pg (urine)	12–30 (plasma) and 0.6–0.9 pmol (urine)	Actis-Goretta et al. 2012
Plasma, urine, and in vitro fecal fermentation	11 (urine), 6 (plasma), and 15 (feces)	0	Enzymatic hydrolysis (for all samples but feces) and silylation	β-D-glucuronidase type H-5 from *H. pomatia*	Split GC	EI–TOF–MS	35	NA	NA	Grün et al. 2008
Plasma, urine, and feces	45	45 (10 standards were made in-house)	SPE	None	HPLC	ESI (positive and negative)–QTrap–MS/MS (MRM)	28	NA	NA	de Ferrars et al. 2014
Normal and malignant breast tissue	15	15	Homogenization, centrifugation, filtration through a PVDF filter, enzymatic hydrolysis, and solvent extraction	β-Glucuronidase and sulfatase from *H. pomatia* for breast tissue samples	UHPLC	ESI (positive and negative)–QTOF – MS/MS (full scan)	22	0.18–29.7	0.6–130	Ávila-Gálvez et al. 2019

Abbreviations: EC, Enzyme Commission; DAD, diode array detector; dMRM, dynamic multiple reaction monitoring; EI, electron impact; ESI, electrospray ionization; GC, gas chromatography; HLB, hydrophilic-lipophilic-balanced; HPLC, high-performance liquid chromatography; LOQ, limit of quantification; MRM, multiple reaction monitoring; MS, mass spectrometer; NA, not applicable; PRM, parallel reaction monitoring; PVDF, polyvinylidene fluoride; Q-Orbitrap, hybrid quadrupole Orbitrap mass spectrometer; QqQ, triple quadrupole mass spectrometer; QTOF, quadrupole time-of-flight mass spectrometer; QTrap, triple quadrupole ion trap mass spectrometer; SDB-L, styrene divinylbenzene; SPE, solid phase extraction; SRM, single reaction monitoring; TOF, time-of-flight mass spectrometer; UHPLC, ultra-high-performance liquid chromatography.

Table 2 Selected MS-based methods for the untargeted analysis of polyphenols and their metabolites in human-derived matrices

Matrix	Number of annotated polyphenol metabolites	Number of identified polyphenol metabolites	Identification strategy and confidence level	Sample preparation	Enzyme treatment	Chromatography	Detection	Time (min)	Reference
Urine	NA	35 (23 not including isomers)	MS-based molecular networking using a cosine-similarity score between 0.7 and 1	SPE and PTFE HPLC filtration	None	UHPLC	ESI (positive and negative)–QTOF–MS/MS (auto mode)	55	Hakeem Said et al. 2020
Urine	NA	31	By comparing the exact mass (±5 mDa) to the values in free databases such as the Human Metabolome Database	Centrifugation and dilution in Milli-Q water	None	HPLC	ESI (positive and negative)–QTOF–MS/MS (full scan)	12	Garcia-Aloy et al. 2015
Urine[a]	65	46	With reference standards; MSI level 1	SPE with SDB-L or HLB cartridges, or direct injection	None	UHPLC	ESI (negative)–Q-Orbitrap	45	Ordóñez et al. 2018
Urine	629	83	Monoisotopic and standard MS–MS matching with databases or literature	Dilution until all urine samples had the same specific gravity (measured by a digital refractometer) and centrifugation	None	UHPLC	ESI (negative)–QTOF–MS/MS	12	Edmands et al. 2015

(*Continued*)

Table 2 (*Continued*)

Matrix	Number of annotated polyphenol metabolites	Number of identified polyphenol metabolites	Identification strategy and confidence level	Sample preparation	Enzyme treatment	Chromatography	Detection	Time (min)	Reference
In vitro fecal fermentation (using pig samples)[b]	75	NA	Annotation using data from Phenol-Explorer; COSMOS MSI level 2	Centrifugation followed by cellulose syringe filtration	None	UHPLC	ESI (positive)–QTOF–MS (scan mode)	35	Rocchetti et al. 2019
Plasma and urine	61 (urine) and 9 (plasma)	6 (urine) and 3 (plasma)	Matching masses and retention times with authentic standards using the same analytical method	Millipore PVDF SPE extraction	None	HPLC	ESI (positive and negative)–Q-FT-Orbitrap–MS/MS (DDA)	13	Ulaszewska et al. 2020
Plasma, urine, and normal and malignant breast tissue	90 (urine), 60 (plasma), 31 (normal breast tissue), and 27 (malignant breast tissue)	33	With reference standards; MSI level 1	Homogenization for breast tissue samples; centrifugation and filtration through a PVDF filter for all samples	None	UHPLC	ESI (positive and negative)–QTOF–MS/MS	22	Ávila-Gálvez et al. 2019

[a]A targeted screening method was also developed and applied (see **Table 1**).
[b]A targeted method was also developed and used (see **Table 1**).
Abbreviations: ESI, electrospray ionization; HLB, hydrophilic-lipophilic-balanced; HPLC, high-performance liquid chromatography; MS, mass spectrometry; MSI, Metabolomics Standards Initiative; NA, not applicable; PTFE, polytetrafluorethylene; PVDF, polyvinylidene fluoride; Q-FT-Orbitrap, hybrid linear ion trap Fourier transform Orbitrap mass spectrometer; Q-Orbitrap, hybrid quadrupole Orbitrap mass spectrometer; QTOF, quadrupole time-of-flight mass spectrometer; SDB-L, styrene divinylbenzene; SPE, solid phase extraction; UHPLC, ultra-high-performance liquid chromatography.

a vast potential and need to develop broader and more comprehensive methods for the accurate quantification of polyphenols and their human metabolites.

Most studies tend to develop methods for a limited number of related subclasses of polyphenols to address a specific biological question. Thus, the methods are optimized to have favorable separation for one chemical class or similar subclasses. This narrow focus can be a necessity to detect polyphenol metabolites where concentrations are very low. However, a method that was optimized for a subclass of polyphenols may not be optimal for another, thus making it difficult to study chemically diverse polyphenols simultaneously. Therefore, such a method is appropriate only for looking at a few desired subclasses of polyphenols but is not ideal to gain a broader and more complete picture, which would be required to assess the real impact of food bioactives on human health in a systems-wide manner. Creating such a quantitative multiclass assay would allow for studying exposure via a variety of food sources in-depth for more holistic cohort studies. Moreover, having such an extensive and broad multi-analyte LC-MS/MS method could support benchmarking and standardization initiatives for NTA. For such a method, both targeted and untargeted LC-MS-based methods can be applied, with the untargeted approach having the advantage of involving potential unknown molecules (Bocato et al. 2019).

Untargeted Liquid Chromatography–High-Resolution Mass Spectrometry–Based Metabolomics

With the advent of more affordable and easier to use high-resolution mass spectrometric instrumentation during the past two decades, untargeted approaches have emerged as an alternative to strictly targeted assays in many fields. This includes the screening (and, partly, quantification) of known and unknown chemical exposures in the food and exposure sciences. An untargeted approach requires high-resolution MS (HRMS) for full scan capacities, which thus creates massive data files that consequently require advanced bioinformatic solutions to extract relevant information from the obtained metabolic features (Bocato et al. 2019). For polyphenol metabolites in human matrices, relatively little research utilizing the power of untargeted methods has been performed to date because of the complexity of this methodology. However, there is considerable usefulness in this approach, illustrated in **Figure 3**, as it screens numerous polyphenol metabolites, allowing the identification of as yet unknown metabolites and finding metabolites that can act as better biomarkers in subsequent targeted approaches. For example, a study demonstrated that following consumption of six different polyphenol-rich foods, 83 different metabolites were annotated and used as potential dietary biomarkers (Edmands et al. 2015). Similarly, in another experiment, out of 187 potential polyphenol metabolites, 90 were annotated, of which 33 were identified with reference standards (Ávila-Gálvez et al. 2019). These are two of the methods with the highest number of annotated metabolites. However, confident identification of polyphenols and metabolites is still a major challenge. According to the Metabolomics Standards Initiative (MSI), authentic reference standards are required for confidently identifying a certain molecule (Sumner et al. 2007). This again underlines the importance of reference standard availability, notably of the conjugated metabolites. Currently, the method with the highest number of identified polyphenol metabolites that have a reference standard for each metabolite quantitatively assesses 46 polyphenols and metabolites (Ordóñez et al. 2018). Clearly, there is ample potential in further expanding the untargeted toolbox, whether it involves annotating or identifying more metabolites, looking at exposure and its correlation with environmental and food factors, developing bioinformatic pipelines, or observing other human sample matrices. **Table 2** presents an overview of published untargeted MS-based methods for the analysis of polyphenols and their metabolites in human samples.

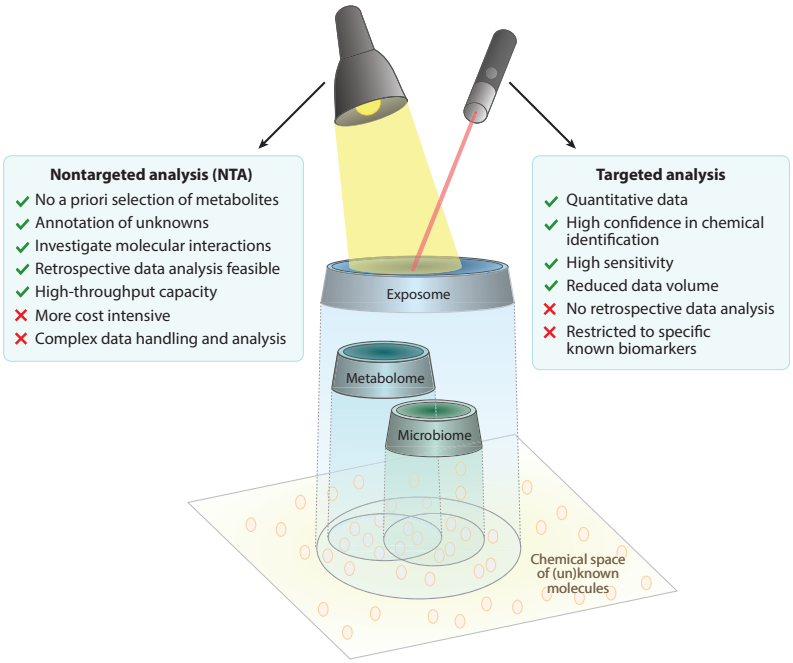


Figure 3

The exposome comprises xenobiotic exposures, including food bioactives such as polyphenols, and associated biological responses that may be reflected in the endogenous metabolome. Contrary to the selective targeted analysis of food/environmental exposures for biomonitoring purposes (symbolized by the laser), nontargeted analysis (symbolized by the floodlight) allows the more holistic investigation of chemical exposure and endogenous as well as microbial metabolites.

Polyphenol Metabolites in Human Matrices

As stated above, polyphenols tend to be metabolized by biotransformation, and various metabolites can be determined in different biological samples. As reported in **Tables 1** and **2**, the main matrices investigated for HBM purposes were plasma and urine. Plasma is often available in HBM and clinical studies and has an advantage in that it interacts with all living cells in the body; thus, a broad range of metabolites may be present (Wallace et al. 2016). Urine is frequently used in HBM studies, as the majority of absorbed metabolites are excreted via the kidneys, it is noninvasive, and large volumes can be collected. Moreover, it can easily be used in routine sample collection, and longitudinal experimental designs are easier to perform than with other matrices (Smolders et al. 2009). However, polyphenol metabolites can also be present in other sample types, e.g., 39 different metabolites were found in breast tissue (Ávila-Gálvez et al. 2019). Studying rather uncommon sample matrices holds special potential for novel insights. For instance, few studies use saliva as a matrix, even though various metabolites have been detected. One study, for example, examined the intake of red wine by studying associated metabolites found in saliva, although it only measured the levels of melatonin and 11 phenylpropanoids (Varoni et al. 2013).

Even less research has been performed on polyphenol metabolism in matrices other than biofluids or stool. An interesting option might be the investigation of biomarkers in bones, as it has recently been shown that various polyphenols, such as urolithins, from pigs fed with plant-based diets were detectable in their bones. This may be applied to humans' postmortem for assessing life-long exposure (Alldritt et al. 2019). Another possible matrix could involve teeth, as one study has found that it is possible to measure a variety of compounds, from metals to organic molecules such as pollutants like bisphenol A, in primary teeth, indicating prenatal exposure (Andra et al. 2015). The advantage of such an approach for exposure assessment is derived from the fact that organic compounds are likely more stable in the inorganic tooth matrix than in organic matrices, thus allowing better storage and cumulative exposure over a period of time. Hair is another potentially interesting matrix for the analysis of polyphenol metabolomics, as it is readily available, can be sampled noninvasively, and has the potential to deliver longitudinal exposure monitoring.

MICROBIAL POLYPHENOL METABOLITES AND MICROBIOME STUDIES

The microbiome is a vital, yet still incompletely understood, component of the human body. The composition and the metabolic activities of the microbiome can be modified by diverse factors including diet, drug treatment, and lifestyle (Moco et al. 2012). Each person harbors a unique microbiota, and the variation in microbiota composition can have dramatic effects on the extent of polyphenol metabolism. For example, soy isoflavones can be transformed into equol, an estrogen receptor agonist, by a restricted number of low-abundance intestinal bacteria, most notably *Adlercreutzia* species (Clavel et al. 2014). As not everyone has *Adlercreutzia* in their gut microbiota, there is dramatic interindividual variability in the production of equol (Setchell & Clerici 2010). It is likely that variation in the microbiota is a key underlying determinant in the observed variability between individuals in the metabolism of many polyphenols. Because of the high complexity of the microbiome, most studies exploring microbial polyphenol metabolism and its impact are performed either in vivo in mice, by in vitro models in which stool samples or derived fecal water is fermented, or in single-species bacterial culture.

Consumption of certain polyphenols has been shown to act as a prebiotic and promote the growth of specific microorganisms and thereby modulate the gut microbiome (Han & Xiao 2020), resulting in several reported effects. For example, in mice consuming a diet supplemented with bioactive dietary polyphenols, the gut microbiota was changed in a way that led to an improvement in sleep-deprived induced cognitive impairment (Frolinger et al. 2019). Conversely, polyphenols may be inhibitory to some microorganisms and reduce their activity or abundance and thereby produce a positive health effect. For example, an increase in consumption of procyanidins has been shown to reduce the amount of *Lachnospiraceae* species and thereby elicit an antiobesity effect (Gowd et al. 2019). Moreover, the effects of polyphenols, whether inhibiting or promoting, on different species in the gut microbiota depends on the polyphenols' backbone and its conjugation. For instance, the flavanol (+)-catechin promoted the growth of *Eubacterium rectale* (Etxeberria et al. 2013). For flavonols and flavanone aglycones to inhibit the activity of gut microbiota, they require a 4-carbonyl group attached to the C-ring in their backbone; this activity transpires only with aglycones and is not present in the glycosides (Duda-Chodak et al. 2015).

Examples of potential metabolic pathways for the biotransformation of polyphenols are outlined in **Table 3**. Some main mechanisms for polyphenol metabolism by the microbiota involve cleavage of various bonds in their heterocyclic carbon ring mediated by dioxygenases, hydrogenations of alkene moieties, or dehydroxylation (Stevens & Maier 2016). As an example, proanthocyanidins are cleaved into smaller phenolic acids such as 2-(*p*-hydroxyphenyl)-propionic acid

Table 3 Examples of microbial metabolites formed from polyphenols in the human body and details on the respective studies

Microbial metabolite	Matrix where metabolite was found[a]	CAS number	Microbes responsible for metabolite[a]	Parent polyphenol (and example of a food source)	Chromatography	Detection	Min, max, and mean concentrations (μg/mL)	Example of biological activity of the metabolite	Reference
Dihydroresveratrol	NA	58436-28-5	NA	Resveratrol (peanuts)	UHPLC	ESI–QTOF–MS	Min: 0 Max: 0.7 ± 0.4 Mean: NA (from in vitro fecal fermentation)	Reduces lung injury from acute pancreatitis in mice (Lin et al. 2016)	Jarosova et al. 2019
5-(3′,4′-Dihydroxyphenyl)-γ-valerolactone	Urine, blood, and feces	191666-22-5	NA	Flavan-3-ol (grape seed)	UHPLC	DAD–ESI–QqQ–MS	Min: 0 Max: 62 Mean: 8.5 (from in vitro fecal fermentation)	Inhibits TNF-α-stimulated proinflammatory responses (Lee et al. 2017)	Sánchez-Patán et al. 2012
3-(3-Hydroxyphenyl)-propionic	Urine, blood, and feces	621-54-5	*Escherichia coli* and *Eubacterium oxidoreducens*	Proanthocyanidin (legume seeds)	UHPLC	ESI–QqQ–MS/MS	Min: 0.012 Max: 18.3 Mean: 1.50 ± 3.50 (from fecal sample)	Antioxidant properties (Rios et al. 2003)	Gutiérrez-Díaz et al. 2017
3-Hydroxybenzoic acid	Urine, blood, and feces	99-06-9	*Pseudomonas*	Hydroxybenzoic acid (pineapple)	UHPLC	ESI–QqQ–MS/MS	Min: 0.009 Max: 0.174 Mean: 0.047 ± 0.042 (from fecal sample)	Analgesic and stress response desensitizing activities; overall has similar activities as aspirin (Khan et al. 2016)	Gutiérrez-Díaz et al. 2017
Gallic acid	Urine, blood, and feces	149-91-7	NA	Anthocyanin (red wine)	UHPLC	ESI–QqQ–MS/MS	Min: NA Max: NA Mean: 2.19 ± 0.24 μg/g feces (from fecal sample)	Numerous effects against, e.g., cardiovascular, gastrointestinal, and neuropsychological diseases (such as improving passive avoidance memory) (Kahkeshani et al. 2019)	Muñoz-González et al. 2013
3,4-Dihydroxy-hydrocinnamic acid	Urine, blood, feces, and cytoplasm	1078-61-1	*Lactobacillus gasseri*, *Bifidiobacterium lactis*, and *E. coli*	Flavone (chamomile)	UHPLC	DAD–ESI–QqQ–MS	Min: 0 Max: 0.33 Mean: 0.084 (from in vitro fecal fermentation)	Reduces ferric ions in plasma (DeGraft-Johnson et al. 2007)	Sánchez-Patán et al. 2012

(*Continued*)

Table 3 (*Continued*)

Microbial metabolite	Matrix where metabolite was found[a]	CAS number	Microbes responsible for metabolite[a]	Parent polyphenol (and example of a food source)	Chromatography	Detection	Min, max, and mean concentrations (μg/mL)	Example of biological activity of the metabolite	Reference
Vanillic acid	Urine, blood, feces, cytoplasm, and saliva	121-34-6	*Pseudomonas putida*, *Pseudomonas* sp. AZ10 UPM, and *Delftia acidovorans*	Anthocyanin (grape)	UHPLC	ESI–QqQ–MS/MS	Min: 0.385 Max: 0.808 Mean: 0.484 ± 0.105 (from fecal sample)	Inhibits anticoagulation effect of *Naja naja* venom by inhibiting various enzymes such as 5′ nucleotidase (Dhananjaya et al. 2006)	Gutiérrez-Díaz et al. 2017
Protocatechuic acid	Urine, blood, feces, fibroblast, and testicle	99-50-3	*Azomonas macrocytogenes*	Flavanol (apple)	UHPLC	ESI–QqQ–MS/MS	Min: 0.002 Max: 0.927 Mean: 0.150 ± 0.177 (from fecal sample)	Antioxidant neuroprotective properties by, e.g., stimulating cell proliferation (Guan et al. 2009)	Gutiérrez-Díaz et al. 2017
Urolithin A	Urine, feces, membrane, and cell membrane	1143-70-0	NA	Ellagic acid and ellagitannin (pomegranate)	UHPLC	DAD–ESI–QTOF–MS	NA	Inhibits inflammation in the bowel by improving gut barrier health (Singh et al. 2019)	García-Villalba et al. 2019
Ferulic acid	Urine, blood, feces, cytoplasm, fibroblast, and epidermis	537-98-4	*Pseudomonas acidovorans* and *Pseudomonas fluorescens*	Cinnamic acid (oats)	UHPLC	ESI–QqQ–MS/MS	Min: 0.001 Max: 0.975 Mean: 0.066 ± 0.158 (from fecal sample)	Antioxidant effect by protecting neuronal cells and synaptosomal membrane from oxidation (Mancuso et al. 2007)	Gutiérrez-Díaz et al. 2017
Hippuric acid	Urine, blood, feces, breast milk, saliva, cerebrospinal fluid, cytoplasm, kidney, liver, prostate, and placenta	495-69-2	*Actinomycetaceae*, *Clostridium*, *Eubacterium*, *Ruminococcus*, and *Faecalibacterium prausnitzii*	Phenolic acid (blueberry)	UHPLC	ESI–QqQ–MS/MS	Min: 0.008 Max: 0.123 Mean: 0.042 ± 0.040 (from fecal sample)	Potential biomarker for obesity and hypertension (Vernocchi et al. 2016)	Gutiérrez-Díaz et al. 2017

[a]These columns have data extracted from the Human Metabolome Database (Wishart et al. 2018).

Abbreviations: CAS, Chemical Abstracts Service; DAD, diode array detector; ESI, electrospray ionization; MS, mass spectrometry; NA, not applicable; QqQ, triple quadrupole mass spectrometer; QTOF, quadrupole time-of-flight mass spectrometer; TNF, tumor necrosis factor; UHPLC, ultra-high-performance liquid chromatography.

(Déprez et al. 2000). Quercetin, as another example, was shown to be transformed into homoprocatechuic acid or 4-hydroxybenzoic acid by species such as *Bifidiobacterium* B-9 and *Streptococcus* S-2 (Santangelo et al. 2019).

Because of the extreme chemical diversity of polyphenol metabolites formed by the gut microbiota, untargeted HRMS techniques offer a unique chance for the discovery of microbial metabolites in a similar manner as described above for human metabolic products. However, distinguishing metabolites derived from human or microbial transformation activity is not straightforward. Dissecting the origin of a specific microbial metabolite to a bacterial species is even more challenging. However, it is expected that novel technologies based on stable isotopes will, in the near future, provide new avenues for approaching this topic (Flasch et al. 2020).

As mentioned above, the common technique to better understand the processes of polyphenol metabolism by the gut microbiota is to mimic the large intestine in vitro. A specific method was developed to analyze 19 different metabolites from the biotransformation of grape polyphenols by the gut microbiota (Zhao et al. 2018). Another study used an untargeted approach to annotate 75 polyphenol metabolites during fecal fermentation (Rocchetti et al. 2019). However, certain biotransformation reactions that occur in other organs or during digestion can be missed when using this in vitro model. A polyphenol could be absorbed and transformed in the liver and then released back to the small intestine with the bile and finally reach the gut microbes for further transformation. To date, a limited number of studies have used untargeted LC-HRMS for polyphenol metabolomics in gut microbiome research. This results in knowledge gaps that could be highly relevant when systematically investigating the impact of polyphenol exposure on health parameters in clinical trials.

EARLY-LIFE EXPOSURE AND IMPACT ON HUMAN HEALTH

Following the developmental origin of health and disease hypothesis (Wadhwa et al. 2009), studying early-life exposure is vital, and studying polyphenols more holistically by untargeted LC-HRMS workflows would certainly be of high interest because of their reported bioactive properties. Exposure to environmental factors, including bioactives, contaminants, and toxins, especially during gestation and the first 1,000 days of life, is believed to often result in chronic health issues later in life. However, because of limitations in measuring the thousands of (low-dose) exposures and the complexity of toxicological interactions, hard evidence is currently lacking for many pathologies. In the case of undernourished human fetuses, changes in metabolism and decreased growth rates were demonstrated to ultimately lead to health issues such as hypertension or heart disease (Osmond & Barker 2000). In addition, during embryonic and fetal development there are critical windows of susceptibility during which exposure is particularly detrimental (Selevan et al. 2000); for example, pregnant mothers that were exposed to air pollution during the second trimester gave birth to children who later in life had a higher risk of developing asthma (Hsu et al. 2015). Because polyphenols are able to reduce certain toxic effects, there is a clear rationale to coinvestigate them in future exposomic-scale environmental health studies.

Current research in early-life exposomics is usually done in animal models, as it is challenging to assess fetal exposure to xenobiotics because various biomarkers in plasma and urine have a short half-life. Numerous studies in animal models have shown the benefits of polyphenols in protecting against chronic diseases. Silva et al. (2019) showed that consuming resveratrol from gestation until postnatal day 21 helped reduce hypertension development in rats; however, as a side effect, postnatal growth was restricted (Care et al. 2016, Silva et al. 2019). Studies investigating the impact of polyphenol exposure in human early-life models are rare, and those that have been done tend to look at short-term effects (Silva et al. 2019). Furthermore, some studies have also shown negative

effects of consuming polyphenols during pregnancy, e.g., consumption of polyphenol-rich foods during the third trimester could have a negative impact on the fetus by causing ductal constriction, which can increase the risk of neonatal pulmonary hypertension (Zielinsky et al. 2010, 2013). However, in this work, polyphenol exposure was estimated based on food frequency questionnaires rather than biomonitoring. Polyphenols can also be used as chemosensors for various ions that can have negative or positive impacts on health, such as using curcumin as the complexing agent against ions like Hg^{2+} or F^{-} (Khorasani et al. 2019). Similarly, polyphenol oxidases can be used as chemosensors to detect polyphenols in, for example, urine; however, issues have arisen because of their lack of specificity (Gul et al. 2017). Clearly, investigating polyphenol exposure and metabolism during early life using novel high-end technology offers unique opportunities for attaining a better picture of the combinatory impact of diverse environmental exposures.

Importantly, the composition of the early-life gut microbiome is thought to be of relevance for chronic disease development, especially with regard to inflammatory and autoimmune disorders (Kelly et al. 2007, Stiemsma & Michels 2018). Because polyphenols have been shown to influence the gut microbiome, it would be beneficial to study the associated effect in larger and more systematic multi-omic studies during early life. Research in Asian women consuming a high soy diet indicated a reduced risk of breast cancer (Wu et al. 2008). These protective properties are believed to follow exposure to soy-related isoflavones and their effect on the gut microbiome occurs during early life (Warri et al. 2008). The protective effects of genistein could be due to a change in the composition of the gut microbiota (Paul et al. 2017) or the microbial metabolites formed (Hullar et al. 2014).

CONCLUSION AND OUTLOOK

The field of polyphenol research and analysis is clearly of high relevance in the context of novel omic-scale exposure assessment and systems toxicology approaches. Recent advances in analytical instrumentation, namely HRMS, bioinformatic capacities, and tailored structural and biological databases allow for a more holistic and in-depth investigation of polyphenols and their abundant human and microbial metabolites.

Future research efforts should include the development of a very broad (targeted) reference method to analyze >100 polyphenols and key metabolites simultaneously to get a more complete understanding of the various food bioactives present. Investigating only a limited number of subclasses is not sufficient, as mixture effects are likely to occur. Developing methods to analyze polyphenols in matrices other than blood, urine, or feces, and continuing the development of untargeted methods along with tailored bioinformatic solutions for the resulting big data will pave the way for identifying potential new biomarkers of exposure and effect. In addition, possibly new correlations between metabolites from the consumption of specific polyphenol-rich foods or different influential factors may be derived.

Other research areas of interest may include the investigation of polyphenol-related enzymes, e.g., the polyphenol oxidases from various sources such as apples (Kampatsikas et al. 2019), walnuts (Panis et al. 2020), and mushrooms (Pretzler et al. 2017). These enzymes can exhibit bioactivity, as quinones can be derived via polyphenol transformation by these enzymes (Queiroz et al. 2008) or lignans such as nordihydroguaiaretic acid, which is a powerful antioxidant, can show antiviral properties, and is derived from larreatricin with the help of polyphenol oxidases (Martin et al. 2018). The bioactivity of these enzymes may also aid the development of polyphenol chemosensors with higher specificity. Finally, we encourage better investigation of polyphenol exposure during different stages of development from early to late life and the potential correlations with

the etiology and/or prevention of chronic disease in the perspective of the expanding exposome paradigm.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors would like to acknowledge Augustine Scalbert (IARC, Lyon, France), Doris Marko and Gudrun Pahlke (University of Vienna, Austria), and Gary Siuzdak (Scripps Research, La Jolla, CA) for critical feedback and fruitful discussions. All members of the Warth lab are acknowledged for providing feedback on content and figure style. This work was partly supported by the University of Vienna, and the European Research Council (D.B., Starting Grant FunKeyGut 741623).

LITERATURE CITED

Achaintre D, Gicquiau A, Li L, Rinaldi S, Scalbert A. 2018. Quantification of 38 dietary polyphenols in plasma by differential isotope labelling and liquid chromatography electrospray ionization tandem mass spectrometry. *J. Chromatogr. A.* 1558:50–58

Actis-Goretta L, Lévèques A, Giuffrida F, Romanov-Michailidis F, Viton F, et al. 2012. Elucidation of (−)-epicatechin metabolites after ingestion of chocolate by healthy humans. *Free Radic. Biol. Med.* 53(4):787–95

Alldritt I, Whitham-Agut B, Sipin M, Studholme J, Trentacoste A, et al. 2019. Metabolomics reveals diet-derived plant polyphenols accumulate in physiological bone. *Sci. Rep.* 9:8047

Almeida AF, Borge GIA, Piskula M, Tudose A, Tudoreanu L, et al. 2018. Bioavailability of quercetin in humans with a focus on interindividual variation. *Compr. Rev. Food Sci. Food Saf.* 17(3):714–31

Anantharaju PG, Gowda PC, Vimalambike MG, Madhunapantula SV. 2016. An overview on the role of dietary phenolics for the treatment of cancers. *Nutr. J.* 15(1):99

Andra SS, Austin C, Wright RO, Arora M. 2015. Reconstructing pre-natal and early childhood exposure to multi-class organic chemicals using teeth: towards a retrospective temporal exposome. *Environ. Int.* 83:137–45

Ávila-Gálvez MÁ, García-Villalba R, Martínez-Díaz F, Ocaña-Castillo B, Monedero-Saiz T, et al. 2019. Metabolic profiling of dietary polyphenols and methylxanthines in normal and malignant mammary tissues from breast cancer patients. *Mol. Nutr. Food Res.* 63(9):1801239

Bocato MZ, Ximenez JPB, Hoffmann C, Barbosa F. 2019. An overview of the current progress, challenges, and prospects of human biomonitoring and exposome studies. *J. Toxicol. Environ. Health B* 22(5–6):131–56

Care AS, Sung MM, Panahi S, Gragasin FS, Dyck JRB, et al. 2016. Perinatal resveratrol supplementation to spontaneously hypertensive rat dams mitigates the development of hypertension in adult offspring. *Hypertension* 67(5):1038–44

Carry E, Zhao D, Mogno I, Faith J, Ho L, et al. 2018. Targeted analysis of microbial-generated phenolic acid metabolites derived from grape flavanols by gas chromatography–triple quadrupole mass spectrometry. *J. Pharm. Biomed. Anal.* 159:374–83

Clavel T, Lepage P, Charrier C. 2014. The Family Coriobacteriaceae. In *The Prokaryotes*: *Actinobacteria*, ed. E Rosenberg, EF DeLong, S Lory, E Stackebrandt, F Thompson, pp. 201–38. Berlin: Springer

Crozier A, Jaganath IB, Clifford MN. 2009. Dietary phenolics: chemistry, bioavailability and effects on health. *Nat. Prod. Rep.* 26(8):1001–43

de Ferrars RM, Czank C, Saha S, Needs PW, Zhang Q, et al. 2014. Methods for isolating, identifying, and quantifying anthocyanin metabolites in clinical samples. *Anal. Chem.* 86(20):10052–58

de Roos B, Duthie GG. 2015. Role of dietary pro-oxidants in the maintenance of health and resilience to oxidative stress. *Mol. Nutr. Food Res.* 59(7):1229–48

de Santiago E, Gill CIR, Carafa I, Tuohy KM, De Peña M-P, Cid C. 2019. Digestion and colonic fermentation of raw and cooked *Opuntia ficus-indica cladodes* impacts bioaccessibility and bioactivity. *J. Agric. Food Chem.* 67(9):2490–99

DeGraft-Johnson J, Kolodziejczyk K, Krol M, Nowak P, Krol B, Nowak D. 2007. Ferric-reducing ability power of selected plant polyphenols and their metabolites: implications for clinical studies on the antioxidant effects of fruits and vegetable consumption. *Basic Clin. Pharmacol. Toxicol.* 100(5):345–52

del Rio D, Rodriguez-Mateos A, Spencer JPE, Tognolini M, Borges G, Crozier A. 2012. Dietary (poly)phenolics in human health: structures, bioavailability, and evidence of protective effects against chronic diseases. *Antioxid. Redox Signal.* 18(14):1818–92

Dennis KK, Marder E, Balshaw DM, Cui Y, Lynes MA, et al. 2017. Biomonitoring in the era of the exposome. *Environ. Health Perspect.* 125(4):502–10

Déprez S, Brezillon C, Rabot S, Philippe C, Mila I, et al. 2000. Polymeric proanthocyanidins are catabolized by human colonic microflora into low-molecular-weight phenolic acids. *J. Nutr.* 130(11):2733–38

Devi KP, Malar DS, Nabavi SF, Sureda A, Xiao J, et al. 2015. Kaempferol and inflammation: from chemistry to medicine. *Pharmacol. Res.* 99:1–10

Dhananjaya BL, Nataraju A, Rajesh R, Gowda RCD, Sharath BK, et al. 2006. Anticoagulant effect of *Naja naja* venom 5′nucleotidase: demonstration through the use of novel specific inhibitor, vanillic acid. *Toxicon* 48(4):411–21

Duda-Chodak A, Tarko T, Satora P, Sroka P. 2015. Interaction of dietary compounds, especially polyphenols, with the intestinal microbiota: a review. *Eur. J. Nutr.* 54(3):325–41

Edmands WM, Ferrari P, Rothwell JA, Rinaldi S, Slimani N, et al. 2015. Polyphenol metabolome in human urine and its association with intake of polyphenol-rich foods across European countries. *Am. J. Clin. Nutr.* 102(4):905–13

Emwas A-HM. 2015. The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. In *Metabonomics: Methods and Protocols*, ed. JT Bjerrum, pp. 161–93. New York: Springer

Etxeberria U, Fernández-Quintela A, Milagro FI, Aguirre L, Martínez JA, Portillo MP. 2013. Impact of polyphenols and polyphenol-rich dietary sources on gut microbiota composition. *J. Agric. Food Chem.* 61(40):9517–33

Fahrioğlu U, Dodurga Y, Elmas L, Seçme M. 2016. Ferulic acid decreases cell viability and colony formation while inhibiting migration of MIA PaCa-2 human pancreatic cancer cells in vitro. *Gene* 576(1):476–82

Figueira I, Menezes R, Macedo D, Costa I, Nunes dos Santos C. 2017. Polyphenols beyond barriers: a glimpse into the brain. *Curr. Neuropharmacol.* 15(4):562–94

Flasch M, Bueschl C, Woelflingseder L, Schwartz-Zimmermann HE, Adam G, et al. 2020. Stable isotope-assisted metabolomics for deciphering xenobiotic metabolism in mammalian cell culture. *ACS Chem. Biol.* 15(4):970–81

Frolinger T, Sims S, Smith C, Wang J, Cheng H, et al. 2019. The gut microbiota composition affects dietary polyphenols-mediated cognitive resilience in mice by modulating the bioavailability of phenolic acids. *Sci. Rep.* 9:3546

Garcia-Aloy M, Llorach R, Urpi-Sarda M, Jáuregui O, Corella D, et al. 2015. A metabolomics-driven approach to predict cocoa product consumption by designing a multimetabolite biomarker model in free-living subjects from the PREDIMED study. *Mol. Nutr. Food Res.* 59(2):212–20

García-Villalba R, Selma MV, Espín JC, Tomás-Barberán FA. 2019. Identification of novel urolithin metabolites in human feces and urine after the intake of a pomegranate extract. *J. Agric. Food Chem.* 67(40):11099–107

Gleichenhagen M, Schieber A. 2016. Current challenges in polyphenol analytical chemistry. *Curr. Opin. Food Sci.* 7:43–49

Gowd V, Karim N, Shishir MRI, Xie L, Chen W. 2019. Dietary polyphenols to combat the metabolic diseases via altering gut microbiota. *Trends Food Sci. Technol.* 93:81–93

Granato D, Mocan A, Câmara JS. 2020. Is a higher ingestion of phenolic compounds the best dietary strategy? A scientific opinion on the deleterious effects of polyphenols in vivo. *Trends Food Sci. Technol.* 98:162–66

Grün CH, van Dorsten FA, Jacobs DM, Le Belleguic M, van Velzen EJJ, et al. 2008. GC-MS methods for metabolic profiling of microbial fermentation products of dietary polyphenols in human and in vitro intervention studies. *J. Chromatogr. B* 871(2):212–19

Guan S, Ge D, Liu T-Q, Ma X-H, Cui Z-F. 2009. Protocatechuic acid promotes cell proliferation and reduces basal apoptosis in cultured neural stem cells. *Toxicol. In Vitro* 23(2):201–8

Gul I, Ahmad MS, Naqvi SMS, Hussain A, Wali R, et al. 2017. Polyphenol oxidase (PPO) based biosensors for detection of phenolic compounds: a review. *J. Appl. Biol. Biotech.* 5(3):72–85

Gutiérrez-Díaz I, Fernández-Navarro T, Salazar N, Bartolomé B, Moreno-Arribas MV, et al. 2017. Adherence to a Mediterranean diet influences the fecal metabolic profile of microbial-derived phenolics in a Spanish Cohort of middle-age and older people. *J. Agric. Food Chem.* 65(3):586–95

Hakeem Said I, Truex JD, Heidorn C, Retta MB, Petrov DD, et al. 2020. LC-MS/MS based molecular networking approach for the identification of cocoa phenolic metabolites in human urine. *Food Res. Int.* 132:109119

Han Y, Xiao H. 2020. Whole food-based approaches to modulating gut microbiota and associated diseases. *Annu. Rev. Food. Sci. Technol.* 11:119–43

Heleno SA, Martins A, Queiroz MJRP, Ferreira ICFR. 2015. Bioactivity of phenolic acids: metabolites versus parent compounds: a review. *Food Chem*. 173:501–13

Hodgson JM, Chan SY, Puddey IB, Devine A, Wattanapenpaiboon N, et al. 2004. Phenolic acid metabolites as biomarkers for tea- and coffee-derived polyphenol exposure in human subjects. *Br. J. Nutr.* 91(2):301–5

Hsu H-HL, Chiu Y-HM, Coull BA, Kloog I, Schwartz J, et al. 2015. Prenatal particulate air pollution and asthma onset in urban children. Identifying sensitive windows and sex differences. *Am. J. Respir. Crit. Care Med.* 192(9):1052–59

Hullar MAJ, Burnett-Hartman AN, Lampe JW. 2014. Gut microbes, diet, and cancer. *Cancer Treat. Res.* 159:377–99

Ito H, Gonthier M-P, Manach C, Morand C, Mennen L, et al. 2005. Polyphenol levels in human urine after intake of six different polyphenol-rich beverages. *Br. J. Nutr.* 94(4):500–9

Jarosova V, Vesely O, Marsik P, Jaimes JD, Smejkal K, et al. 2019. Metabolism of stilbenoids by human faecal microbiota. *Molecules* 24(6):1155

Kahkeshani N, Farzaei F, Fotouhi M, Alavi SS, Bahramsoltani R, et al. 2019. Pharmacological effects of gallic acid in health and diseases: a mechanistic review. *Iran. J. Basic Med. Sci.* 22(3):225–37

Kalt W, McDonald JE, Liu Y, Fillmore SAE. 2017. Flavonoid metabolites in human urine during blueberry anthocyanin intake. *J. Agric. Food Chem.* 65(8):1582–91

Kampatsikas I, Bijelic A, Pretzler M, Rompel A. 2019. A peptide-induced self-cleavage reaction initiates the activation of tyrosinase. *Angew. Chem. Int. Ed.* 58(22):7475–79

Kelly D, King T, Aminov R. 2007. Importance of microbial colonization of the gut in early life to the development of immunity. *Mutat. Res.* 622(1):58–69

Khan SA, Chatterjee SS, Kumar V. 2016. Low dose aspirin like analgesic and anti-inflammatory activities of mono-hydroxybenzoic acids in stressed rodents. *Life Sci.* 148:53–62

Khorasani MY, Langari H, Sany SBT, Rezayi M, Sahebkar A. 2019. The role of curcumin and its derivatives in sensory applications. *Mater. Sci. Eng. C* 103:109792

Kyselova Z. 2011. Toxicological aspects of the use of phenolic compounds in disease prevention. *Interdiscip. Toxicol.* 4(4):173–83

Lee CC, Kim JH, Kim JS, Oh YS, Han SM, et al. 2017. 5-(3′,4′-Dihydroxyphenyl-γ-valerolactone), a major microbial metabolite of proanthocyanidin, attenuates THP-1 monocyte-endothelial adhesion. *Int. J. Mol. Sci.* 18(7):1363

Lin Z-S, Ku CF, Guan Y-F, Xiao H-T, Shi X-K, et al. 2016. Dihydro-resveratrol ameliorates lung injury in rats with cerulein-induced acute pancreatitis. *Phytother. Res.* 30(4):663–70

Magiera S, Baranowska I, Kusa J. 2012. Development and validation of UHPLC-ESI-MS/MS method for the determination of selected cardiovascular drugs, polyphenols and their metabolites in human urine. *Talanta* 89:47–56

Manach C, Williamson G, Morand C, Scalbert A, Rémésy C. 2005. Bioavailability and bioefficacy of polyphenols in humans. I. Review of 97 bioavailability studies. *Am. J. Clin. Nutr.* 81(1):230S–42

Mancuso C, Scapagnini G, Currò D, Giuffrida Stella AM, de Marco C, et al. 2007. Mitochondrial dysfunction, free radical generation and cellular stress response in neurodegenerative disorders. *Front. Biosci.* 12(1):1107–23

Marín L, Miguélez EM, Villar CJ, Lombó F. 2015. Bioavailability of dietary polyphenols and gut microbiota metabolism: antimicrobial properties. *BioMed Res. Int.* 2015:905215

Martin HJ, Kampatsikas I, Oost R, Pretzler M, Al-Sayed E, et al. 2018. Total synthesis stereochemical assignment, and divergent enantioselective enzymatic recognition of larreatricin. *Chem. Eur. J.* 24:15756–60

Martin KR. 2009. Polyphenols as dietary supplements: a double-edged sword. *Nutr. Diet. Suppl.* 2:1–12

Moco S, Martin F-PJ, Rezzi S. 2012. Metabolomics view on gut microbiome modulation by polyphenol-rich foods. *J. Proteome Res.* 11(10):4781–90

Muñoz-González I, Jiménez-Girón A, Martín-Álvarez PJ, Bartolomé B, Moreno-Arribas MV. 2013. Profiling of microbial-derived phenolic metabolites in human feces after moderate red wine intake. *J. Agric. Food Chem.* 61(39):9470–79

Neveu V, Perez-Jiménez J, Vos F, Crespy V, du Chaffaut L, et al. 2010. Phenol-Explorer: an online comprehensive database on polyphenol contents in foods. *Database* 2010:bap024

Ordóñez JL, Pereira-Caro G, Ludwig I, Muñoz-Redondo JM, Ruiz-Moreno MJ, et al. 2018. A critical evaluation of the use of gas chromatography- and high performance liquid chromatography-mass spectrometry techniques for the analysis of microbial metabolites in human urine after consumption of orange juice. *J. Chromatogr. A.* 1575:100–12

Osmond C, Barker DJ. 2000. Fetal, infant, and childhood growth are predictors of coronary heart disease, diabetes, and hypertension in adult men and women. *Environ. Health Perspect.* 108(Suppl. 3):545–53

Ottaviani JI, Fong RY, Borges G, Schroeter H, Crozier A. 2018. 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.* 124:97–103

Panis F, Kampatsikas I, Bijelic A, Rompel A. 2020. Conversion of walnut tyrosinase into a catechol oxidase by site directed mutagenesis. *Sci. Rep.* 10:1659

Paul B, Royston KJ, Li Y, Stoll ML, Skibola CF, et al. 2017. Impact of genistein on the gut microbiome of humanized mice and its role in breast tumor inhibition. *PLOS ONE* 12(12):e0189756

Pretzler M, Bijelic A, Rompel A. 2017. Heterologous expression and characterization of functional mushroom tyrosinase (*Ab*PPO4). *Sci. Rep.* 7:1810

Queiroz C, Lopes MLM, Fialho E, Valente-Mesquita VL. 2008. Polyphenol oxidase: characteristics and mechanisms of browning control. *Food Rev. Int.* 24(4):361–75

Rios LY, Gonthier M-P, Rémésy C, Mila I, Lapierre C, et al. 2003. Chocolate intake increases urinary excretion of polyphenol-derived phenolic acids in healthy human subjects. *Am. J. Clin. Nutr.* 77(4):912–18

Rocchetti G, Bhumireddy SR, Giuberti G, Mandal R, Lucini L, Wishart DS. 2019. Edible nuts deliver polyphenols and their transformation products to the large intestine: an in vitro fermentation model combining targeted/untargeted metabolomics. *Food Res. Int.* 116:786–94

Rocchetti G, Pagnossa JP, Blasi F, Cossignani L, Hilsdorf Piccoli R, et al. 2020. Phenolic profiling and in vitro bioactivity of *Moringa oleifera* leaves as affected by different extraction solvents. *Food Res. Int.* 127:108712

Rothwell JA, Perez-Jimenez J, Neveu V, Medina-Remón A, M'Hiri N, et al. 2013. Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol content. *Database* 2013:bat070

Roura E, Andrés-Lacueva C, Jáuregui O, Badia E, Estruch R, et al. 2005. Rapid liquid chromatography tandem mass spectrometry assay to quantify plasma (−)-epicatechin metabolites after ingestion of a standard portion of cocoa beverage in humans. *J. Agric. Food Chem.* 53(16):6190–94

Sánchez-Patán F, Cueva C, Monagas M, Walton GE, Gibson GR, et al. 2012. 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.* 131(1):337–47

Sandhu AK, Miller MG, Thangthaeng N, Scott TM, Shukitt-Hale B, et al. 2018. Metabolic fate of strawberry polyphenols after chronic intake in healthy older adults. *Food Funct.* 9(1):96–106

Santangelo R, Silvestrini A, Mancuso C. 2019. Ginsenosides, catechins, quercetin and gut microbiota: current evidence of challenging interactions. *Food Chem. Toxicol.* 123:42–49

Scalbert A, Brennan L, Fiehn O, Hankemeier T, Kristal BS, et al. 2009. Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics* 5(4):435

Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, et al. 2014. The food metabolome: a window over dietary exposure. *Am. J. Clin. Nutr.* 99(6):1286–308

Selevan SG, Kimmel CA, Mendola P. 2000. Identifying critical windows of exposure for children's health. *Environ. Health Perspect.* 108(suppl 3):451–55

Setchell KDR, Clerici C. 2010. Equol: history, chemistry, and formation. *J. Nutr.* 140(7):1355S–62S

Shahidi F, Yeo J. 2018. Bioactivities of phenolics by focusing on suppression of chronic diseases: a review. *Int. J. Mol. Sci.* 19(6):1573

Shao B, Li H, Shen J, Wu Y. 2019. Nontargeted detection methods for food safety and integrity. *Annu. Rev. Food Sci. Technol.* 10:429–55

Silva LBAR, Pinheiro-Castro N, Novaes GM, de Freitas Laiber Pascoal G, Ong TP. 2019. Bioactive food compounds, epigenetics and chronic disease prevention: focus on early-life interventions with polyphenols. *Food Res. Int.* 125:108646

Singh R, Chandrashekharappa S, Bodduluri SR, Baby BV, Hegde B, et al. 2019. Enhancement of the gut barrier integrity by a microbial metabolite through the Nrf2 pathway. *Nat. Commun.* 10(1):89

Smolders R, Schramm K-W, Nickmilder M, Schoeters G. 2009. Applicability of non-invasively collected matrices for human biomonitoring. *Environ. Health.* 8(1):8

Stevens JF, Maier CS. 2016. The chemistry of gut microbial metabolism of polyphenols. *Phytochem. Rev.* 15(3):425–44

Stiemsma LT, Michels KB. 2018. The role of the microbiome in the developmental origins of health and disease. *Pediatrics* 141(4):e20172437

Sulpizio C, Breibeck J, Rompel A. 2018. Recent progress in synthesis and characterization of metal chalcone complexes and their potential as bioactive agents. *Coord. Chem. Rev.* 374:497–524

Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, et al. 2007. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 3(3):211–21

Svilar L, Martin J-C, Defoort C, Paut C, Tourniaire F, Brochot A. 2019. Quantification of trans-resveratrol and its metabolites in human plasma using ultra-high performance liquid chromatography tandem quadrupole-orbitrap mass spectrometry. *J. Chromatogr. B.* 1104:119–29

Teixeira LL, Costa GR, Dörr FA, Ong TP, Pinto E, et al. 2017. Potential antiproliferative activity of polyphenol metabolites against human breast cancer cells and their urine excretion pattern in healthy subjects following acute intake of a polyphenol-rich juice of grumixama (*Eugenia brasiliensis* Lam.). *Food Funct.* 8(6):2266–74

Ulaszewska MM, Koutsos A, Trošt K, Stanstrup J, Garcia-Aloy M, et al. 2020. Two apples a day modulate human:microbiome co-metabolic processing of polyphenols, tyrosine and tryptophan. *Eur. J. Nutr.* 59:3691–714

Urpi-Sarda M, Monagas M, Khan N, Lamuela-Raventos RM, Santos-Buelga C, et al. 2009. Epicatechin, procyanidins, and phenolic microbial metabolites after cocoa intake in humans and rats. *Anal. Bioanal. Chem.* 394(6):1545–56

Varoni EM, Vitalini S, Contino D, Lodi G, Simonetti P, et al. 2013. Effects of red wine intake on human salivary antiradical capacity and total polyphenol content. *Food Chem. Toxicol.* 58:289–94

Vernocchi P, del Chierico F, Putignani L. 2016. Gut microbiota profiling: metabolomics based approach to unravel compounds affecting human health. *Front. Microbiol.* 7:1144

Wadhwa PD, Buss C, Entringer S, Swanson JM. 2009. Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. *Semin. Reprod. Med.* 27(5):358–68

Wallace MAG, Kormos TM, Pleil JD. 2016. Blood-borne biomarkers and bioindicators for linking exposure to health effects in environmental health science. *J. Toxicol. Environ. Health B* 19(8):380–409

Warri A, Saarinen NM, Makela S, Hilakivi-Clarke L. 2008. The role of early life genistein exposures in modifying breast cancer risk. *Br. J. Cancer* 98(9):1485–93

Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, et al. 2018. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46(D1):D608–17

Wu AH, Yu MC, Tseng C-C, Pike MC. 2008. Epidemiology of soy exposures and breast cancer risk. *Br. J. Cancer* 98(1):9–14

Wu X, Cao G, Prior RL. 2002. Absorption and metabolism of anthocyanins in elderly women after consumption of elderberry or blueberry. *J. Nutr.* 132(7):1865–71

Zhao D, Yuan B, Carry E, Pasinetti GM, Ho L, et al. 2018. Development and validation of an ultra-high performance liquid chromatography/triple quadrupole mass spectrometry method for analyzing microbial-derived grape polyphenol metabolites. *J. Chromatogr. B.* 1099:34–45

Zhong S, Sandhu A, Edirisinghe I, Burton-Freeman B. 2017. Characterization of wild blueberry polyphenols bioavailability and kinetic profile in plasma over 24-h period in human subjects. *Mol. Nutr. Food Res.* 61(12):1700405

Zhu Y, Yin Q, Yang Y. 2020. Comprehensive investigation of *Moringa oleifera* from different regions by simultaneous determination of 11 polyphenols using UPLC-ESI-MS/MS. *Molecules* 25(3):676

Ziaullah, Vasantha Rupasinghe HP. 2015. Application of NMR spectroscopy in plant polyphenols associated with human health. In *Applications of NMR Spectroscopy*, Vol. 2, ed. A ur-Rahman, MI Choudhary, pp. 3–92. Sharah, UAE: Bentham Sci. Publ.

Zielinsky P, Piccoli AL Jr., Manica JL, Nicoloso LH, Menezes H, et al. 2010. Maternal consumption of polyphenol-rich foods in late pregnancy and fetal ductus arteriosus flow dynamics. *J. Perinatol.* 30(1):17–21

Zielinsky P, Piccoli AL Jr., Vian I, Zílio AM, Naujorks AA, et al. 2013. Maternal restriction of polyphenols and fetal ductal dynamics in normal pregnancy: an open clinical trial. *Arq. Bras. Cardiol.* 101(3):217–25

6.3. Publication #2 (Original Research): Oesterle et al. 2022

Status	Published
Title	**Quantifying up to 90 polyphenols simultaneously in human bio-fluids by LC-MS/MS**
Authors	Ian Oesterle[a,b,c], Dominik Braun[a], Annette Rompel[c], Benedikt Warth[a]
Affiliations	[a]Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, 1090 Vienna, Austria [b]Doctoral School of Chemistry, University of Vienna, 1090 Vienna, Austria [c]Universität Wien, Fakultät für Chemie,Institut für Biophysikalische Chemie, 1090 Wien, Österreich
Year	2022
Journal	*Analytica Chimica Acta*
Journal Information	Volume 1216, 11 July 2022, 339977
Accepted	May 21, 2022
DOI	https://doi.org/10.1016/j.aca.2022.339977
Contribution	Ian Oesterle performed the sample preparation and LC-MS/MS optimization, performed the method validation and result evaluation, and wrote the initial draft of the manuscript.

Analytica Chimica Acta 1216 (2022) 339977

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Quantifying up to 90 polyphenols simultaneously in human bio-fluids by LC-MS/MS

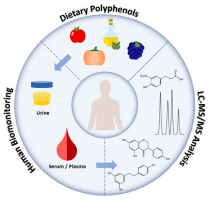
Ian Oesterle [a,b,c], Dominik Braun [a], Annette Rompel [c], Benedikt Warth [a,*]

[a] Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, 1090, Vienna, Austria
[b] Doctoral School of Chemistry, University of Vienna, 1090, Vienna, Austria
[c] Universität Wien, Fakultät für Chemie, Institut für Biophysikalische Chemie, 1090, Wien, Österreich

HIGHLIGHTS

- Exposome method for 90 polyphenol analytes developed.
- Low sample volumes of 40 μL feasible.
- Validated for human urine, serum, and plasma.
- High sensitivity despite high-throughput sample preparation.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Human biomonitoring (HBM)
Exposome
Urine
Serum
Plasma

ABSTRACT

Establishing a method for human biomonitoring (HBM) of polyphenols enables the assessment of internal concentrations of these food bio-actives and the correlation with potential health effects such as antioxidant or anti-inflammatory properties. Thus, a targeted LC-MS/MS method for quantifying up to 90 analytes, representing the main polyphenol classes including flavanones, isoflavones, stilbenes, and phenolic acids, was developed for human urine, serum, and plasma. The method was established for low sample volumes and with a cost and time efficient sample preparation protocol for high-throughput, which is critical for its application in large cohort and exposome-wide association studies. On average, the sample preparation yielded extraction efficiencies of 98% for urine, 98% for serum, and 87% for plasma. Limits of detection were between 0.11 ng mL^{-1} and 300 ng mL^{-1} for urine, 0.12 ng mL^{-1} and 190 ng mL^{-1} for serum, and 0.12 ng mL^{-1} and 340 ng mL^{-1} for plasma, excluding one analyte. In-house validation revealed that 66, 49, and 64 analytes for urine, serum, and plasma, respectively, fulfilled all stringent requirements, that are usually utilized for tailored single analyte methods, at all evaluated concentration levels. After validation, this method was applied in a proof-of-principle study that detected 39 polyphenols in urine. Changes in the concentrations of the analytes after the ingestion of a high polyphenol smoothie was examined over 24 h. The study further confirmed that the majority of polyphenols detected were phenolic acids, and phase II conjugated metabolites were more abundant than their respective non-conjugated forms.

* Corresponding author.
E-mail address: benedikt.warth@univie.ac.at (B. Warth).

https://doi.org/10.1016/j.aca.2022.339977
Received 23 December 2021; Received in revised form 27 April 2022; Accepted 21 May 2022
Available online 23 May 2022

1. Introduction

In recent years, interest in polyphenols has increased due to their various health benefits, such as anti-inflammatory and antibacterial properties [1–5], and their interaction with the gut microbiome [6]. Even if not supplemented, humans are frequently exposed to significant concentrations of polyphenols as they are present in the majority of plant-based foodstuff [3,7,8]. This is displayed in the "Phenol-Explorer" database, where currently 502 different polyphenols from 452 different food items are listed [9]. However, significantly more polyphenols exist, with over 8′000 polyphenols having been identified [10]. The vast number of polyphenols is just one of the complexities of this class of molecules.

Polyphenols are typically separated into two main families: flavonoids and non-flavonoids. These two families can be separated into several subfamilies, or classes, such as isoflavones, anthocyanidins, and lignans, depending on the molecular structure, as depicted in Fig. 1 [11–13]. Due to the available hydroxyl groups, polyphenols are often conjugated, yielding numerous structural possibilities. In foodstuffs, glycoside structures are frequently present, with the sugar moiety quickly cleaved off once ingested. In the human body, glucuronidated and sulfated conjugates are mainly formed through phase II biotransformation [14,15]. Other processes in humans involving polyphenols occur, many involving the gut microbiome, yielding a plethora of metabolites [16,17].

Successfully quantifying polyphenols in humans would be valuable because of how common they are in one's diet and their effect on health, both positive and negative, like the estrogenic effects of isoflavones [18–20]. To be able to analyze such a complex family of molecules, sensitive state-of-the-art techniques are required. LC-MS/MS offers the possibility to separate complex mixtures using liquid chromatography (LC), and the opportunity to analyze multiple compounds together through mass spectrometry (MS) due to its high sensitivity and resolution [21,22].

Multiple methods targeting specific polyphenol analytes for human biomonitoring (HBM) purposes have been established. However, these techniques typically either focus on less than 50 polyphenol analytes or on only a few specific polyphenol classes as reviewed by Oesterle et al. [23]. Therefore, developing a method that is broader in scope and includes a larger number of analytes from all the main polyphenol classes would be instrumental to aid in the development of more comprehensive targeted methods. Such a method could also be used as a reference method for developing non-targeted analysis (NTA) workflow, an approach investigating all the analytes present rather than targeting selected analytes [24].

Consequently, we developed an in-house validated LC-MS/MS method that focuses on 90 different polyphenols from the following polyphenol classes: anthocyanins, lignans, dihydrochalcones, flavones, flavanones, proanthocyanidins, stilbenes, flavonols, hydroxybenzoic acids, hydroxyphenylacetic acids, catechins, hydroxycinnamic acid, isoflavones, and some other polyphenols. To date, the majority of published HBM methods focused on one human matrix exclusively, typically urine as it is the most easily acquired human matrix, while the method here was developed and validated for urine, serum and plasma simultaneously. Finally, the newly developed method was applied in a pilot study where urine was sampled after the consumption of polyphenols from a fruit smoothie rich in polyphenols to investigate the concentration change of the polyphenols and prove the method is fit-for-purpose.

2. Experimental section

2.1. Reagents, solvents, and chemicals

Information regarding the chemicals, reagents and solvents used, where they were acquired, and the preparation of stock and working solutions is listed in detail in the Supplementary Information (page 1

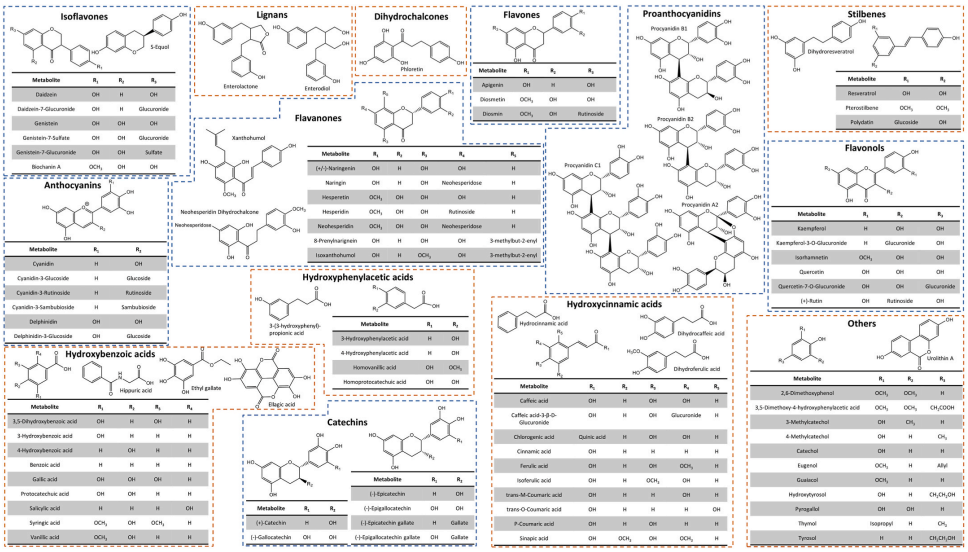


Fig. 1. Molecular structures of the analytes included in this study derived from 14 polyphenol classes, highlighting the high chemical diversity covered. Flavonoid classes are framed in blue and non-flavonoids classes in orange. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

and Table S1).

2.2. Sample preparation

A sample aliquot (either plasma, urine, or serum) was gently thawed and 40 μL was transferred into a 0.5 mL clear micro-reaction tube. 160 μL extraction solvent (acetonitrile (ACN): methanol (MeOH): formic acid (FA); 49.5:49.5:1; v:v:v) was added. The tube was shortly vortexed and sonicated in an ice bath for 10 min. After a freeze-out step (−20 °C, 2 h), the tube was centrifuged at 4 °C and 18′000×*g* for 10 min. The supernatant (75 μL) was transferred to a 1.5 mL amber LC glass vial with a 300 μL glass insert. 225 μL of the reconstitution solvent (ACN: water (H_2O): FA; 33:66:1; v:v:v) was added.

2.3. LC-MS/MS instrumentation and parameters

The neat solvent and matrix samples were analyzed on a UHPLC-ESI-QTrap-MS/MS system consisting of a 1290 Infinity II LC (Agilent) connected to a QTrap 6500+ MS (Sciex, Vienna, Austria) equipped with a heated electrospray ionization source (ESI). Data was acquired using fast polarity switching and scheduled multiple reaction monitoring (sMRM) mode. The optimized LC method utilized an Acquity UPLC HSS T3 column (1.8 μm, 2.1 × 100 mm, Waters, Vienna, Austria) attached to a VanGuard precolumn (1.8 μm, Waters, Vienna, Austria). The mobile phase was composed of 0.1% v/v FA in H_2O (eluent A) and 0.1% v/v FA in ACN (eluent B). The column oven and autosampler temperatures were set to 30 °C and 7 °C, respectively. The flow rate was set to 0.6 mL min^{-1}, and the injection volume was 5 μL. The gradient started with 5% eluent B, which after a 2-min hold was raised linearly to 64% within 10 min. Eluent B was then directly increased to 95% for a 2-min hold, and then decreased to 5% for a final 2-min re-equilibration.

MS parameters were optimized for each analyte individually by direct infusion experiments at concentrations ranging from 0.2 to 1 mg L^{-1} and a flow rate of 10 μL min^{-1} to optimize eight transitions, of which the two (in some cases three) with the highest signal-to-noise (S/N) ratios were selected for method development. The optimized MS and MS/MS parameters are reported in Table S2. Following flow-injection analysis and optimization, the final ESI operation parameters were set as follows: source temperature 550 °C, curtain gas 35 arb, sheath gas (GS1) 90 arb, drying gas (GS2) 90 arb, collision gas set to medium, and entrance potential of −10 V in negative mode and 10 V in positive mode. The ion capillary voltage was set to −4500 V in negative mode and 5500 V in positive mode.

2.4. Validation experiments

The LC-MS/MS method was validated in-house following the guidelines set by the EU commission decision 2002/657/EC [25] and Eurachem [26]. A detailed description on the analytical parameters and their calculation is discussed in the Supplementary Information (page 3). The criteria for successfully validating the parameters were as follows: a regression coefficient of at least 0.95, an extraction efficiency between 50 and 120%, and an intermediate precision and repeatability less than 25%, 30%, and 45% for the high, middle, and low spiking levels, respectively. The intermediate precision and repeatability limits were determined using the Horwitz equation [25].

2.5. Biological samples and case study

Pooled human urine, serum, and Li-heparin plasma were used for method optimization and validation. Information regarding the suppliers for serum and plasma is listed in Table S1, while pooled human urine was acquired from a female volunteer after two days of a low polyphenol diet. The purchased and acquired matrices were aliquoted into 1 mL samples and stored at −80 °C. For method validation, a new aliquot of each matrix was gently thawed for each individual experiment.

In a proof-of-principle study, four volunteers (2 male and 2 female), all healthy, non-smokers, and between 26 and 31 years old, provided urine samples. Following a 2-day polyphenol washout phase, the participants consumed a high polyphenol smoothie in the morning within 15 min. The smoothie composed of 250 mL of gunpowder green tea, 125 g of soy yoghurt, 125 g of frozen berry mix, 125 g of fresh strawberries, and 1/4 of a fresh pomelo. More information regarding the ingredients of the smoothie is listed in Table S1. The participants continued to follow a low polyphenol diet for the rest of the day. The following and last day, the participants were told to consume a high polyphenol diet of their choice.

Individual spot urine samples were collected on the 2nd day of the washout phase with the time and total volume urinated noted. Urine at each time point was aliquoted into individual 15 mL tube. Based on the total volume urinated at each sampling, a 24-h urine sample was prepared retrospectively. On the third day, morning urine was sampled (around 7 a.m.), and following the ingestion of the above-mentioned smoothie (8 a.m.), urine was acquired at the following time points: 9 a.m., 10 a.m., 11 a.m., 12 p.m., 2 p.m., 4 p.m., 6 p.m., 8 p.m., 10 p.m., 7 a.m. of the next day, and at any time during the night if needed, again with the total volume urinated noted at each time point. For each participant, another mixed 24-h urine was created for this day based on the recorded volume urinated at each sampling point. Finally, the urine from the high polyphenol day was collected in the same manner as during the wash-out such that another 24-h urine was prepared for each participant. All urine samples were stored at 4 °C after their acquisition. Aliquots and the 24-h urines were made at the end of the study and all samples were stored at −80 °C until needed. The acquisition of the urine samples was taken following signed informed consent of the participants and approval by the University of Vienna ethics committee under the authorization number #00650.

3. Results and discussion

3.1. Method optimization

As polyphenols are a vast and complex group of bio-active molecules with different physico-chemical behavior for each class, choosing the best compromise on LC conditions is a critical endeavor. Preferable conditions that are ideal for one class may not necessarily be suitable for another. Therefore, many factors were tested, including three frequently utilized columns: an Acquity UPLC HSS T3 column (1.8 μm, 2.1 × 100 mm, Waters, Vienna, Austria), a Kinetex Biphenyl column (2.6 μm, 3.0 × 150 mm Phenomenex, Aschaffenburg, Germany), and an Acquity BEH C18 column (1.7 μm, 2.1 × 100 mm, Waters, Vienna, Austria). Three different eluent systems were tested: 0.1% v/v FA in both H_2O and ACN, 0.5% v/v FA in both H_2O and ACN, and 0.1% v/v acetic acid with 5 mM ammonium acetate in H_2O and 0.1% v/v acetic acid with 5 mM ammonium acetate and 13% MeOH in ACN, the MeOH was added to prevent precipitation of the ammonium acetate in ACN. Different flow rates between 0.2 and 0.6 mL min^{-1}, column temperatures between 25 and 40 °C, and gradients were explored. The final method parameters and components were chosen based on comparing the S/N values and peak shapes for 28 representative analytes listed in the Supplementary Information (page 2). These analytes were deemed the most relevant and represented the different polyphenol subclasses [11].

Initially, the MRM transitions were acquired during the entire run due to retention time shifts with different LC parameters, but after the final LC method selection, a fixed scheduled MRM window was set. Up to three MRM transitions were used for certain analytes during the optimization process, to both further verify the selectivity of the transitions while changing matrices (e.g. urine to plasma) and to determine the best performing transitions. For the final method the two transitions with the highest S/N values in all three matrices were chosen. The optimization of the ion source parameters is described in the

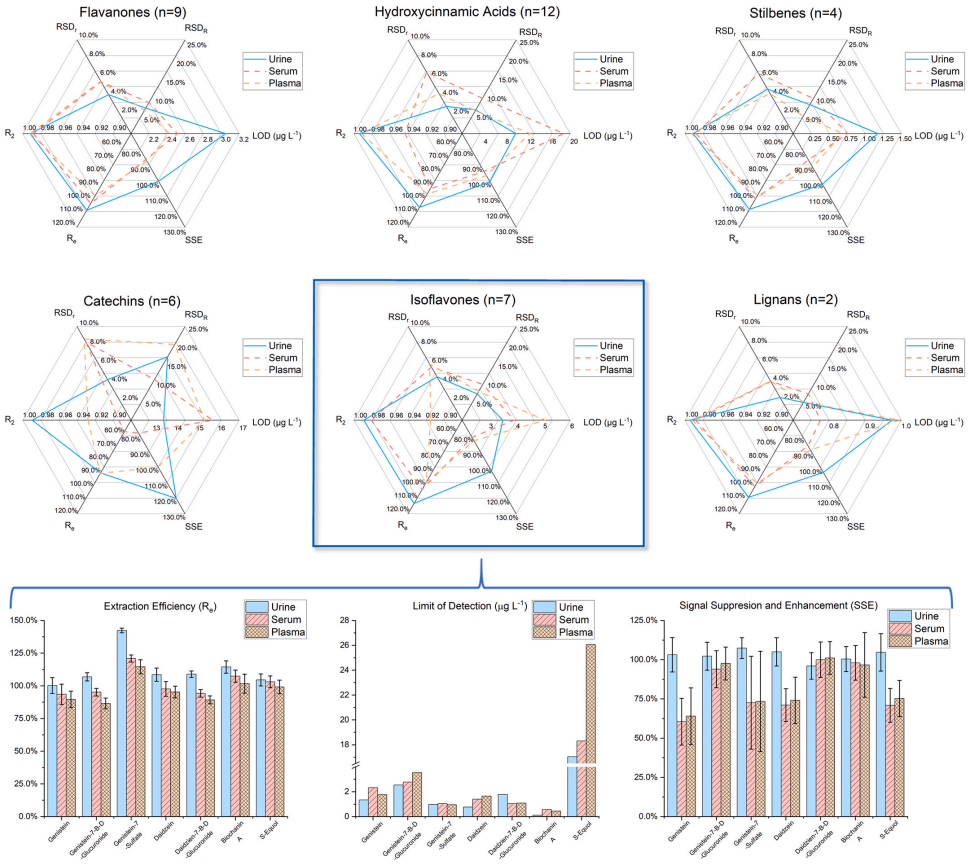


Fig. 2. Analytical figures of merit of six selected polyphenol classes for the three biological matrices as determined during method validation reported as average values. The individual values for all analytes of the class of isoflavones are additionally singled out as a representative example to showcase the underlying data.

Supplementary Information (page 2). The inclusion of several internal standards would have been beneficial but was omitted due to their high cost and a partial lack of availably. This limitation might influence the extent of matrix effects in individual samples.

In pre-experiments, different sample preparation procedures were tested, e.g. an evaporation step using a vacuum concentrator at 4 °C after the centrifugation step. The samples were then reconstituted in the initial LC starting conditions to achieve an overall dilution of 1:50. This time-consuming step did not significantly enhance sensitivity for most analytes, and consequently was omitted. Similarly, two other dilution factors were tested (1:10 and 1:20) to evaluate possible differences in sensitivity and in signal suppression and enhancement (SSE) effects. As expected, higher dilution resulted in lower SSE effects and better S/N ratios but also in less favorable limits of detection (LODs), and vice versa. Therefore, the final sample preparation involved an overall 1:20 dilution as this was deemed the best compromise between SSE and sensitive detection and quantitation.

During method development, carryover was observed for a couple analytes, notably the anthocyanins. This might arise from the anthocyanin having a positively charged oxygen atom requiring a more acidic LC method. However, when more acidic eluents were tested, other analytes yielded less favorable results. As a compromise, the carryover was reduced to <2% for most analytes by injecting twice a "cleaning" blank, consisting of isopropanol:H_2O:FA (25:70:5), after samples with high concentrations in the sequence, such as after the highest calibration standards.

3.2. Validation experiments

Overall, the in-house method validation was majoritarily successful, especially when bearing in mind that the guidelines and criteria used were established for tailored assays and not for highly-sensitive, multi-analyte assays in complex biological matrices for HBM purpose. For urine, 66 out of the 90 polyphenols analytes met all validation criteria at all spiking levels, while for serum and plasma this number was 49 and 64, respectively. For plasma, and most notably serum, fewer analytes

were successfully validated because the concentrations chosen for spiking and calibration were based on preliminary results for urine, meaning that certain analytes had spiking levels that were too low.

The criteria which were not met for the different analytes are shown in Table S5. The analytes that were not successfully validated were mainly anthocyanins, and some flavonols, catechins, and hydroxybenzoic acids. This could be attributed to the aforementioned carry-over observed for anthocyanins, and a couple of flavonols and catechins. While, the hydroxybenzoic acids that failed either had a non-ideal MRM transitions, e.g. ellagic acid which only showed a chromatographic peak in one transition at high concentrations; or had a large contamination in the matrix, such as for hippuric acid. Standard addition was used for evaluation of the analytes that had matrix contaminations. In contrast, several classes performed very well such as stilbenes, flavones, and flavanones, which had all their analytes fulfill the validation criteria.

The selectivity, evaluated by comparing the matrix-matched calibration standards to blanks and solvent calibrations, showed no interference for all analytes. However, two separate chromatographic peaks were observed for six analytes, and both of these peaks were analyzed and validated separately, hence the A and B nomination in Tables S2–5. The two peaks likely arise from potential impurities, e.g. an isomer, of the reference standards. In addition, isoferulic and ferulic acid both had the same MRM transitions and retention time, but were acquired individually in the method.

All calculated analytical parameters for each individual analyte are listed in Tables S3 and S4 with the averages for each polyphenol class are reported in Figs. 2 and S1. Moreover, the chromatograms for each analyte in the three different matrices are shown in Fig. S3. As apparent from Figs. 2 and S1, the polyphenol classes, except for those previously mentioned, yield similar average results for all analytical parameters, apart from their LODs. Moreover, the three different matrices yielded resembling results, with urine performing best as expected. As exemplified in Fig. 2 for the isoflavones, most analytes in one class behave comparably to each other.

The extraction efficiency of the different analytes had averages of 98%, 98%, and 87%, and were in the range of 27%–150%, 42%–190%, and 22%–160%, for urine, serum, and plasma, respectively. Moreover, 88% of the analytes had an extraction efficiency between 80% and 120% for urine at the lowest spiking level and excluding the analytes that are not feasible to calculate at this spiking level. For serum and plasma, it was 78% and 73% of the analytes, respectively. The average extraction efficiency is near 100% due to the sample preparation procedure involving few other steps besides dilution. The SSE for urine is in the range of 23%–460% with an average of 110%, for serum in a range of 31%–250% with an average of 94%, and for plasma between 30% and 230% with an average of 98%. Again, the average SSE effect is near 100% meaning that for the majority of analytes, the matrix does not have a pronounced effect on the chromatographic peak. The linearity of the analytes in the calibration range was represented by regression coefficients that were on average 0.97 for urine, 0.95 for serum, and 0.96 for plasma. The developed method showed adequate precision, measured from sequences run on different days (intermediate precision) and on the same day (repeatability). The repeatability averages for urine, serum, and plasma were: 20%, 22%, and 24%, respectively, while the intermediate precision for the three matrices were: 23%, 23%, and 22%.

The LOD values varied greatly between the different analytes and polyphenols classes. In urine, the LODs were between 0.11 ng mL^{-1} and 300 ng mL^{-1} with a median of 4.8 ng mL^{-1}. While for serum, in a range of 0.12 ng mL^{-1} and 190 ng mL^{-1} with a median of 5.9 ng mL^{-1}; and for plasma between 0.12 ng mL^{-1} and 340 ng mL^{-1} with a median of 5.8 ng mL^{-1}. It is seen that the lower limits and the medians of the LODs were similar for the three matrices.

In general, it is difficult to compare this method to other published assays as the majority of these other assays either focus on fewer analytes or on only a few specific polyphenol families, so only a limited number of LC-MS/MS methods actually cover such a broad range of polyphenols. The sensitivity for urine in the method presented here resembles to those of other published methods, with around half of the analytes being more sensitive [27–29]. Moreover, a recently developed method targeted 119 different analytes, including numerous conjugates, but it mainly focused on phenolic acids [30]. Though, when comparing the sensitivity, roughly half of the analytes that were in common had higher LODs than the method developed here.

In contrast to urine, fewer LC-MS methods for polyphenols in plasma and serum have been published. Sensitivity for plasma in this method is typically lower compared to other published assays, while comparable for serum [29–31]. Again, it should be considered that our method optimization was rather focused on urine while aiming to keep a solid balance between matrices and polyphenol classes.

3.3. Applicability in urine samples

The newly developed method was successfully applied in a proof-of-principle study investigating 56 urine samples obtained at different time points from four subjects. The mean, minimum, and maximum of the results from this study are shown in Table 1. The table also shows the concentrations of the analytes found in the three mixed 24-h urine samples: from the "washout" day, from the day where only a polyphenol-rich smoothie was ingested, and from the day where a high polyphenol diet was consumed. An example of chromatographic peaks from a solvent blank, matrix-matched blank, one matrix-matched calibrant, and one of the proof-of-principle sample for urolithin A and hippuric acid are shown in Fig. 3a. Only the samples which had both quantifier and qualifier ions observed and a deviation of less than 20% from the matrix-matched calibration for the ion ratio were included. For certain analytes, the concentrations were close to the LOD, thus an ion ratio of up to 50% was allowed as at these concentration levels, noise had a strong impact on the ion ratios. Moreover, only samples with minimal retention time deviations (±0.05 min) to the matrix-matched calibration standards were included.

A total of 39 analytes (listed in Table 1) out of the 90 included were observed, and several analytes had more than one chromatographic peak in the MRM window observed, even though only one peak was present in the matrix-matched calibration standards (e.g. Fig. 3d). Four peaks were observed for genistein-7-*O*-glucuronide and kaempferol-3-*O*-glucuronide, three peaks for quercetin-7-*O*-β-D-glucuronide and vanillic acid, and two peaks were found for (−)-epicatechin, (+)-catechin, 3-(3-hydroxyphenyl)propionic acid, dihydrocaffeic acid, ethyl gallate, resveratrol, and sinapic acid. These other peaks might be isomers that are present in the urine samples, such as the glucuronide conjugated on position C-5 instead of C-7 for genistein or *cis*-rather than *trans*-isomer for resveratrol. However, as no reference standards were available, confirmation is difficult.

As reported in Table 1 and exemplified in Fig. 3e, glucuronidated metabolites were quantitated in higher concentrations than their unconjugated or sulfated forms. Ingested polyphenols readily undergo phase II biotransformation, and it was previously observed that roughly 70% of isoflavone metabolites detected in urine were glucuronidated [15,32,33]. The higher concentration of glucuronidated metabolites compared to non-glucuronidated metabolites is also observed for caffeic acid. This is illustrated in Fig. S2 and in Table 1, where it is seen that caffeic-acid-3-β-D-glucuronide has a higher concentration.

In addition, Fig. 3e depicts a clear change in isoflavone concentration over time with an increase from 4-8 h after the consumption of the smoothie, followed by a 10 h decrease to its initial level. In Figs. 3c and

Table 1
Average concentrations (n = 4) of the 39 polyphenol analytes that were identified and quantitated in the proof-of-principle study in urine using the newly established LC-MS/MS multi-class polyphenol method.

Analyte	LOQ (ng mL^{-1})	Min (ng mL^{-1})	Max (ng mL^{-1})	Mean (ng mL^{-1})	24 h urine "washout" (ng mL^{-1})	24 h urine after smoothie (ng mL^{-1})	24 h urine high diet (ng mL^{-1})
Dihydrochalcones							
Phloretin	0.70	<LOQ	2.7	0.90 ± 0.93	–	<LOQ	<LOQ
Flavanones							
(+/−)-Naringenin	1.5	<LOQ	65	12 ± 16	–	6.0	2.2
Flavonols							
Kaempferol-3-*O*-glucuronide	3.5	8.9	180	66 ± 54	–	44	54
Quercetin	15	<LOQ	56	<LOQ	<LOQ	<LOQ	<LOQ
Quercetin-7-*O*-β-D-glucuronide	48	<LOQ	220	69 ± 69	<LOQ	<LOQ	<LOQ
Hydroxybenzoic Acids							
4-Hydroxybenzoic acid	27	180	1300	490 ± 300	500	410	890
Ethyl gallate	1.4	<LOQ	<LOQ	<LOQ	–	–	7.0
Gallic acid	5.2	12	27	19 ± 6.3	–	–	18
Hippuric acid	410	1400	41000	14000 ± 9500	13000	11000	29000
Protocatechuic acid	6.9	14	130	52 ± 32	28	18	97
Salicylic acid	6.5	<LOQ	180	25 ± 41	13	17	28
Syrngic acid	5.8	17	650	100 ± 120	56	67	380
Vanillic acid	11	31	710	210 ± 160	120	210	760
Hydroxycinnamic Acids							
Caffeic acid	30	<LOQ	<LOQ	<LOQ	–	–	–
Caffeic acid-3-β-D-glucuronide	1.1	<LOQ	160	22 ± 32	<LOQ	14	32
Chlorogenic acid	4.4	10	84	35 ± 34	–	–	89
Dihydroferulic acid	4.6	14	730	130 ± 140	86	130	390
Ferulic acid	7.6	<LOQ	60	18 ± 11	10	18	57
Sinapic acid	9.5	<LOQ	600	120 ± 140	14	110	130
trans-*O*-Coumaric acid	5.9	8.3	8.3	8.3 ± 0	–	–	21
Hydroxyphenylacetic Acids							
3-(3-Hydroxyphenyl) propionic acid	4.0	<LOQ	50	21 ± 14	6.9	8.3	63
3-Hydroxyphenylacetic acid	91	510	9800	2600 ± 1900	3100	1900	3100
4-Hydroxyphenylacetic acid	290	2400	20000	7800 ± 4100	13000	6300	8100
Homovanillic acid	140	460	8900	2700 ± 1900	3000	1900	3500
Homoprotocatechuic acid	29	67	2500	600 ± 520	580	420	880
Isoflavones							
Daidzein	1.5	<LOQ	390	43 ± 80	1.8	32	50
Daidzein-7-β-D-glucuronide	3.6	29	29000	3200 ± 5000	61	2500	4800
Genistein	2.7	<LOQ	230	26 ± 48	–	19	17
Genistein-7-β-D-glucuronide	5.1	17	18000	1500 ± 2900	24	1100	1200
Genistein-7-sulfate	2.0	<LOQ	670	76 ± 120	–	45	33
Lignans							
Enterolactone	2.6	<LOQ	4.2	2.0 ± 0.83	<LOQ	<LOQ	<LOQ
Others							
4-Methylcatechol	21	<LOQ	240	80 ± 53	77	55	74
Catechol	83	<LOQ	750	330 ± 230	–	310	540
Guaiacol	940[a]	10000	28000	21000 ± 8100	–	–	25000
Pyrogallol	11	12	1500	290 ± 400	41	100	140
Tyrosol	3500	<LOQ	10000	5200 ± 3100	–	–	–
Urolithin A	3.9	<LOQ	750	92 ± 190	140	41	65
Stilbenes							
Polydatin	3.2	–	–	–	–	–	<LOQ
Resveratrol	4.0	–	–	–	–	–	<LOQ

[a] The value presented here is calculated using the medium spiking level results as no signal was detected in the low spiking level.

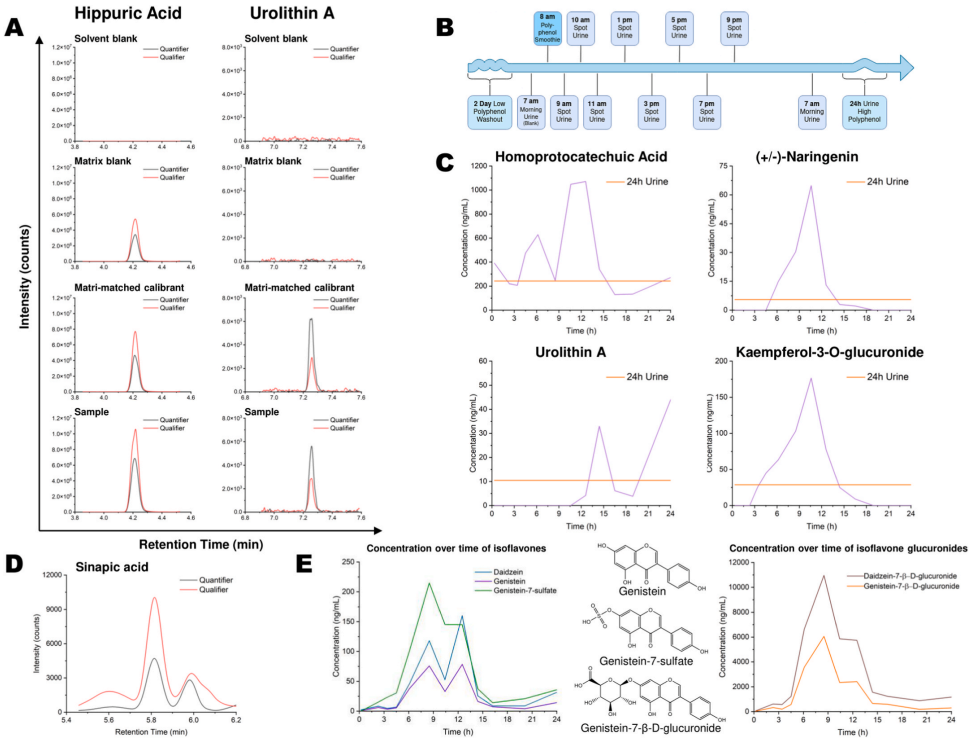


Fig. 3. **A)** MRM chromatograms in a solvent blank, a non-spiked matrix sample ('matrix blank'), a matrix-matched calibrant (1600 ng mL^{-1} for hippuric acid and 0.88 ng mL^{-1} for urolithin A), and a sample analyzed during the proof-of-principle study. **B)** Visual representation of the urine collection time points during the proof-of-principle study. **C)** Concentration for one participant over time plots for four selected polyphenols with their respective concentration found in the 24-h urine in the pilot study. **D)** MRM chromatogram for sinapic acid showing a second peak in the pilot study samples. **E)** Two graphs showing the change of the isoflavone analytes' concentration average of the four participants over time after the ingestion of a high polyphenol smoothie during the pilot study.

S2, similar curves for the concentration over time following the smoothie ingestion were observed for other detected analytes. One exception is urolithin A where the increase starts after 16 h rather than 2–6 h for the other detected analytes. This might be due to urolithin A being a metabolite produced by the gut microbiota. Another exception is sinapic acid (Fig. S2), where the increase in concentration happens very rapidly within the first 3 h and then returns to background concentration after 15 h.

Besides the phase II conjugated metabolites, the majority of polyphenols observed were hydroxybenzoic, hydroxycinnamic, and hydroxyphenylacetic acids, or other small molecules such as urolithin A or catechol. They might arise from phenolic acids and other small molecules being degradation products of other larger polyphenols, such as proanthocyanidins [34–36]. Table 1 also depicts that the concentration for the majority of the detected polyphenols increased from the 24-h urine taken at the end of the "washout" to the 24-h urine after consumption of the high polyphenol smoothie. This concentration then increased even more in the 24-h urine from the high polyphenol diet day, even though not all polyphenols had their concentration return to baseline 24-h after having ingested the smoothie. The increase in concentration following each subsequent day is due to an increase in polyphenols included in the diet for each day.

4. Conclusion and outlook

We report the development of an LC-MS/MS method to identify and quantify up to 90 polyphenols from a variety of different chemical classes in three complex biological matrices. This comprehensive method was optimized and successfully validated for the majority of the analytes in urine, serum, and plasma. Compared to other currently validated methods, the assay developed here had similar sensitivity, although it involved a broader scope and simplified sample preparation. Moreover, several phase II conjugated metabolites were included in the method presented here.

The method developed and validated here presented several advantages, notably in the sample preparation. The sample preparation is simple as it is composed of only a few dilution steps, cheap as low volumes of both reagents and human sample are used, and fast because no evaporation, purification, or other enrichment steps, e.g. SPE, were included. The method is by far more reliable and quantitative than food

frequency questionnaires/diaries or food recalls, and will therefore be beneficial for use in conjunction with these questionnaires in nutritional studies. Thus, this method offers high throughput and is scalable, easily allowing its use in various applications such as large cohort studies.

Finally, this method will serve as a reference method to establish and benchmark non-targeted LC-HRMS approaches [24] which might serve in offering more holistic insights into human exposure of polyphenols in the context of exposome research. These exposomic applications can involve drug-exposome interactions [37] or the potentially protective or harmful effects of polyphenols during food contaminant exposure [38].

CRediT authorship contribution statement

Ian Oesterle: Conceptualization, Methodology, Validation, Investigation, Writing – original draft. **Dominik Braun:** Methodology, Writing – review & editing. **Annette Rompel:** Conceptualization, Resources, Writing – review & editing. **Benedikt Warth:** Conceptualization, Resources, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge all members of the Warth and Rompel labs for providing feedback and valuable support during method development. They further express their gratitude to the Mass Spectrometry Center of the Faculty of Chemistry for technical support, and to Prof. Doris Marko for fruitful discussion. The authors also thank the volunteers participating in the proof-of-principle experiment.

Funding

This work was supported by the University of Vienna and the Austrian Science Fund FWF: P33188 (B.W.) and P32326 (A.R.).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2022.339977.

References

[1] F. Shahidi, J. Yeo, Bioactivities of phenolics by focusing on suppression of chronic diseases: a review, *Int. J. Mol.* Sci. 19 (6) (2018) 1573, https://doi.org/10.3390/ijms19061573.

[2] N. Pap, M. Fidelis, L. Azevedo, M.A.V. do Carmo, D. Wang, A. Mocan, E.P. R. Pereira, D. Xavier-Santos, A.S. Sant'Ana, B. Yang, D. Granato, Berry polyphenols and human health: evidence of antioxidant, anti-inflammatory, microbiota modulation, and cell-protecting effects, Curr. Opin. Food Sci. 42 (2021) 167–186, https://doi.org/10.1016/j.cofs.2021.06.003.

[3] H. Cory, S. Passarelli, J. Szeto, M. Tamez, J. Mattei, The role of polyphenols in human health and food systems: a mini-review, Front. Nutr. 5 (2018) 87, https://doi.org/10.3389/fnut.2018.00087.

[4] C. Sulpizio, J. Breibeck, A. Rompel, Recent progress in synthesis and characterization of metal chalcone complexes and their potential as bioactive agents, Coord. Chem. Rev. 374 (2018) 497–524, https://doi.org/10.1016/j.ccr.2018.05.023.

[5] G. Williamson, C.D. Kay, A. Crozier, The bioavailability, transport, and bioactivity of dietary flavonoids: a review from a historical perspective, Compr. Rev. Food Sci. Food Saf. 17 (5) (2018) 1054–1112, https://doi.org/10.1111/1541-4337.12351.

[6] U. Etxeberria, A. Fernández-Quintela, F.I. Milagro, L. Aguirre, J.A. Martínez, M. P. Portillo, Impact of polyphenols and polyphenol-rich dietary sources on gut microbiota composition, J. Agric. Food Chem. 61 (40) (2013) 9517–9533, https://doi.org/10.1021/jf402506c.

[7] I. Kampatsikas, A. Bijelic, A. Rompel, Biochemical and structural characterization of tomato polyphenol oxidases provide novel insights into their substrate specificity, Sci. Rep. 9 (1) (2019) 4022, https://doi.org/10.1038/s41598-019-39687-0.

[8] A. Derardja, M. Pretzler, I. Kampatsikas, M. Barkat, A. Rompel, Inhibition of apricot polyphenol oxidase by combinations of plant proteases and ascorbic acid, Food Chem. X 4 (2019), 100053, https://doi.org/10.1016/j.fochx.2019.100053.

[9] V. Neveu, J. Perez-Jiménez, F. Vos, V. Crespy, L. du Chaffaut, L.I. Mennen, C. Knox, R. Eisner, J.G. Cruz, D.S. Wishart, A. Scalbert, Phenol-explorer: an online comprehensive database on polyphenol contents in foods, Database 2010 (bap024) (2010), https://doi.org/10.1093/database/bap024.

[10] K.B. Pandey, S.I. Rizvi, Plant polyphenols as dietary antioxidants in human health and disease, Oxid. Med. Cell. Longev. 2 (5) (2009) 270–278, https://doi.org/10.4161/oxim.2.5.9498.

[11] D. del Rio, A. Rodriguez-Mateos, J.P.E. Spencer, M. Tognolini, G. Borges, A. Crozier, Dietary (Poly)Phenolics in human health: structures, bioavailability, and evidence of protective effects against chronic diseases, Antioxidants Redox Signal. 18 (14) (2012) 1818–1892, https://doi.org/10.1089/ars.2012.4581.

[12] S. Quideau, D. Deffieux, C. Douat-Casassus, L. Pouységu, Plant polyphenols: chemical properties, biological activities, and synthesis, Angew. Chem. Int. Ed. 50 (3) (2011) 586–621, https://doi.org/10.1002/anie.201000044.

[13] 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, A. Crozier, Recommendations for standardizing nomenclature for dietary (Poly)Phenol catabolites, Am. J. Clin. Nutr. 112 (4) (2020) 1051–1068, https://doi.org/10.1093/ajcn/nqaa204.

[14] M. Sobhani, M.H. Farzaei, S. Kiani, R. Khodarahmi, Immunomodulatory; anti-inflammatory/antioxidant effects of polyphenols: a comparative review on the parental compounds and their metabolites, Food Rev. Int. 37 (8) (2021) 759–811, https://doi.org/10.1080/87559129.2020.1717523.

[15] Y. Zhang, S. Hendrich, P.A. Murphy, Glucuronides are the main isoflavone metabolites in women, J. Nutr. 133 (2) (2003) 399–404, https://doi.org/10.1093/jn/133.2.399.

[16] A. Scalbert, L. Brennan, C. Manach, C. Andres-Lacueva, L.O. Dragsted, J. Draper, S. M. Rappaport, J.J. van der Hooft, D.S. Wishart, The food metabolome: a window over dietary exposure, Am. J. Clin. Nutr. 99 (6) (2014) 1286–1308, https://doi.org/10.3945/ajcn.113.076133.

[17] 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, D. D. Rio, Phenyl-γ-Valerolactones and phenylvaleric acids, the main colonic metabolites of flavan-3-ols: synthesis, analysis, bioavailability, and bioactivity, Nat. Prod. Rep. 36 (5) (2019) 714–752, https://doi.org/10.1039/C8NP00062J.

[18] D. Granato, A. Mocan, J.S. Câmara, Is a higher ingestion of phenolic compounds the best dietary strategy? A scientific opinion on the deleterious effects of polyphenols in vivo, Trends Food Sci. Technol. 98 (2020) 162–166, https://doi.org/10.1016/j.tifs.2020.01.010.

[19] D.C. Vitale, C. Piazza, B. Melilli, F. Drago, S. Salomone, Isoflavones: estrogenic activity, biological effect and bioavailability, Eur. J. Drug Metab. Pharmacokinet. 38 (1) (2013) 15–25, https://doi.org/10.1007/s13318-012-0112-y.

[20] X. Zhong, J. Ge, S. Chen, Y. Xiong, S. Ma, Q. Chen, Association between dietary isoflavones in soy and legumes and endometrial cancer: a systematic review and meta-analysis, J. Acad. Nutr. Diet. 118 (4) (2018) 637–651, https://doi.org/10.1016/j.jand.2016.09.036.

[21] A.-H.M. Emwas, The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research, in: J.T. Bjerrum (Ed.), Metabonomics: Methods and Protocols, Methods in Molecular Biology, Springer, New York, NY, 2015, pp. 161–193, https://doi.org/10.1007/978-1-4939-2377-9_13.

[22] B. Shao, H. Li, J. Shen, Y. Wu, Nontargeted detection methods for food safety and integrity, Annu. Rev. Food Sci. Technol. 10 (1) (2019) 429–455, https://doi.org/10.1146/annurev-food-032818-121233.

[23] I. Oesterle, D. Braun, D. Berry, L. Wisgrill, A. Rompel, B. Warth, Polyphenol exposure, metabolism, and analysis: a global exposomics perspective, Annu. Rev. Food Sci. Technol. 12 (1) (2021) 461–484, https://doi.org/10.1146/annurev-food-062220-090807.

[24] B.J. Place, E.M. Ulrich, J.K. Challis, A. Chao, B. Du, K. Favela, Y.-L. Feng, C. M. Fisher, P. Gardinali, A. Hood, A.M. Knolhoff, A.D. McEachran, S.L. Nason, S. R. Newton, B. Ng, J. Nuñez, K.T. Peter, A.L. Phillips, N. Quinete, R. Renslow, J. R. Sobus, E.M. Sussman, B. Warth, S. Wickramasekara, A.J. Williams, An introduction to the benchmarking and publications for non-targeted analysis working group, Anal. Chem. 93 (2021) 16289–16296, https://doi.org/10.1021/acs.analchem.1c02660.

[25] European Commission 2002/657/EC, Commission decision of 12 august 2002 implementing council directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Off. J. Eur. Commun. 221 (2002) 8–36.

[26] B. Magnusson, U. Örnemark, In Eurachem Guide: the Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, second ed., 2014, ISBN 978-91-87461-59-0.

[27] D. Achaintre, A. Buleté, C. Cren-Olivé, L. Li, S. Rinaldi, A. Scalbert, Differential isotope labeling of 38 dietary polyphenols and their quantification in urine by liquid chromatography electrospray ionization tandem mass spectrometry, Anal. Chem. 88 (5) (2016) 2637–2644, https://doi.org/10.1021/acs.analchem.5b03609.

[28] J.L. Ordóñez, G. Pereira-Caro, I. Ludwig, J.M. Muñoz-Redondo, M.J. Ruiz-Moreno, A. Crozier, J.M. Moreno-Rojas, A critical evaluation of the use of gas chromatography- and high performance liquid chromatography-mass spectrometry techniques for the analysis of microbial metabolites in human urine after consumption of orange juice, J. Chromatogr. A 1575 (2018) 100–112, https://doi.org/10.1016/j.chroma.2018.09.016.

[29] R.M. de Ferrars, C. Czank, S. Saha, P.W. Needs, Q. Zhang, K.S. Raheem, N. P. Botting, P.A. Kroon, C.D. Kay, Methods for isolating, identifying, and quantifying anthocyanin metabolites in clinical samples, Anal. Chem. 86 (20) (2014) 10052–10058, https://doi.org/10.1021/ac500565a.
[30] M. Domínguez-Fernández, Y. Xu, P. Young Tie Yang, W. Alotaibi, R. Gibson, W. L. Hall, L. Barron, I.A. Ludwig, C. Cid, A. Rodríguez-Mateos, Quantitative assessment of dietary (Poly)Phenol intake: a high-throughput targeted metabolomics method for blood and urine samples, J. Agric. Food Chem. 69 (1) (2021) 537–554, https://doi.org/10.1021/acs.jafc.0c07055.
[31] D. Achaintre, A. Gicquiau, L. Li, S. Rinaldi, A. Scalbert, Quantification of 38 dietary polyphenols in plasma by differential isotope labelling and liquid chromatography electrospray ionization tandem mass spectrometry, J. Chromatogr. A 1558 (2018) 50–58, https://doi.org/10.1016/j.chroma.2018.05.017.
[32] S. Bolca, M. Urpi-Sarda, P. Blondeel, N. Roche, L. Vanhaecke, S. Possemiers, N. Al-Maharik, N. Botting, D. De Keukeleire, M. Bracke, A. Heyerick, C. Manach, H. Depypere, Disposition of soy isoflavones in normal human breast tissue, Am. J. Clin. Nutr. 91 (4) (2010) 976–984, https://doi.org/10.3945/ajcn.2009.28854.
[33] M.Á. Ávila-Gálvez, R. García-Villalba, F. Martínez-Díaz, B. Ocaña-Castillo, T. Monedero-Saiz, A. Torrecillas-Sánchez, B. Abellán, A. González-Sarrías, J. C. Espín, Metabolic profiling of dietary polyphenols and methylxanthines in normal and malignant mammary tissues from breast cancer patients, Mol. Nutr. Food Res. 63 (9) (2019), 1801239, https://doi.org/10.1002/mnfr.201801239.
[34] K. Gao, A. Xu, C. Krul, K. Venema, Y. Liu, Y. Niu, J. Lu, L. Bensoussan, N.P. Seeram, D. Heber, S.M. Henning, Of the major phenolic acids formed during human microbial fermentation of tea, citrus, and soy flavonoid supplements, only 3,4-dihydroxyphenylacetic acid has antiproliferative activity, J. Nutr. 136 (1) (2006) 52–57, https://doi.org/10.1093/jn/136.1.52.
[35] A.R. Rechner, G. Kuhnle, P. Bremner, G.P. Hubbard, K.P. Moore, C.A. Rice-Evans, The metabolic fate of dietary polyphenols in humans, Free Radical Biol. Med. 33 (2) (2002) 220–235, https://doi.org/10.1016/S0891-5849(02)00877-8.
[36] S. Déprez, C. Brezillon, S. Rabot, C. Philippe, I. Mila, C. Lapierre, A. Scalbert, Polymeric proanthocyanidins are catabolized by human colonic microflora into low-molecular-weight phenolic acids, J. Nutr. 130 (11) (2000) 2733–2738, https://doi.org/10.1093/jn/130.11.2733.
[37] M. Pristner, B. Warth, Drug–exposome interactions: the next frontier in precision medicine, Trends Pharmacol. Sci. 41 (12) (2020) 994–1005, https://doi.org/10.1016/j.tips.2020.09.012.
[38] G. Aichinger, G. Bliem, D. Marko, Systemically achievable doses of beer flavonoids induce estrogenicity in human endometrial cells and cause synergistic effects with selected pesticides, Front. Nutr. 8 (2021) 293, https://doi.org/10.3389/fnut.2021.691872.

Supplementary Information

A supplementary data file (pdf) is available online for this publication that contains various tables and figures, such as a table with all the MRM parameters.

DOI: https://doi.org/10.1016/j.aca.2022.339977

6.4. Publication #3 (Original Research): Oesterle et al. 2023a

Status	Published
Title	**Exposomic Biomonitoring of Polyphenols by Non-Targeted Analysis and Suspect Screening**
Authors	Ian Oesterle[a,b,c], Manuel Pristner[a,b], Sabrina Berger[a], Mingxun Wang[e], Vinicius Verri Hernandes[a,d], Annette Rompel[c], Benedikt Warth[a,d]
Affiliations	[a]Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Vienna 1090, Austria [b]Doctoral School of Chemistry, University of Vienna, Vienna 1090, Austria [c]Fakultät für Chemie, Institut für Biophysikalische Chemie, Universität Wien, Wien 1090, Österreich [d]Exposome Austria, Research Infrastructure and National EIRENE Hub, Vienna 1090, Austria [e]Department of Computer Science, University of California Riverside, Riverside, California 92521, USA
Year	2023
Journal	*Analytical Chemistry*
Journal Information	Volume 95, Issue 28, 18 July 2023, Pages 10686-10694
Accepted	June 16, 2023
DOI	https://doi.org/10.1021/acs.analchem.3c01393
Contribution	Ian Oesterle performed the sample preparation and the LC-HRMS measurements, optimized the feature processing and data evaluation workflow for polyphenols, interpreted the results, and wrote the original manuscript.

analytical chemistry

pubs.acs.org/ac Article

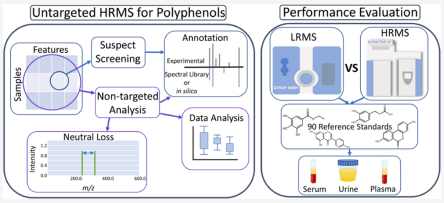
Exposomic Biomonitoring of Polyphenols by Non-Targeted Analysis and Suspect Screening

Ian Oesterle, Manuel Pristner, Sabrina Berger, Mingxun Wang, Vinicius Verri Hernandes, Annette Rompel, and Benedikt Warth*

Cite This: *Anal. Chem.* 2023, 95, 10686−10694 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Polyphenols, prevalent in plants and fungi, are investigated intensively in nutritional and clinical settings because of their beneficial bioactive properties. Due to their complexity, analysis with untargeted approaches is favorable, which typically use high-resolution mass spectrometry (HRMS) rather than low-resolution mass spectrometry (LRMS). Here, the advantages of HRMS were evaluated by thoroughly testing untargeted techniques and available online resources. By applying data-dependent acquisition on real-life urine samples, 27 features were annotated with spectral libraries, 88 with *in silico* fragmentation, and 113 by MS^1 matching with PhytoHub, an online database containing >2000 polyphenols. Moreover, other exogenous and endogenous molecules were screened to measure chemical exposure and potential metabolic effects using the Exposome-Explorer database, further annotating 144 features. Additional polyphenol-related features were explored using various non-targeted analysis techniques including MassQL for glucuronide and sulfate neutral losses, and MetaboAnalyst for statistical analysis. As HRMS typically suffers a sensitivity loss compared to state-of-the-art LRMS used in targeted workflows, the gap between the two instrumental approaches was quantified in three spiked human matrices (urine, serum, plasma) as well as real-life urine samples. Both instruments showed feasible sensitivity, with median limits of detection in the spiked samples being 10−18 ng/mL for HRMS and 4.8−5.8 ng/mL for LRMS. The results demonstrate that, despite its intrinsic limitations, HRMS can readily be used for comprehensively investigating human polyphenol exposure. In the future, this work is expected to allow for linking human health effects with exposure patterns and toxicological mixture effects with other xenobiotics.



A large class of phytochemicals known as polyphenols has been of great interest because they are highly prevalent in plants and fungi and are associated with a variety of beneficial and protective properties in humans, for example, antioxidant[1] and antibacterial[2] effects. Moreover, the gut microbiome, known for its impact on numerous health conditions, is affected and modulated by polyphenols.[3] Polyphenols also hold the potential for negative health impacts, such as exhibiting estrogen-like activity[4] that might be of relevance in mixture toxicology.[5] Thus, investigating polyphenols and their metabolites in humans would aid in understanding the correlation between diet and diseases.[2]

As exogenous compounds to the human body, polyphenols are included in the exposome, that is, the total burden of exposure of an individual in a lifetime and its health related effects.[6] Currently, different exposomic approaches are used[7] to better understand and quantify the complex exposure of xenobiotics in humans. These approaches often utilize liquid chromatography coupled with mass spectrometry (LC-MS) in two distinct acquisition strategies.[8] The first, called traditional human biomonitoring or targeted, focuses on acquiring LC-MS data for only specific known analytes of interest using available reference standards.[9] The second strategy, known as untargeted, involves acquiring spectrometric data on all compounds that can be ionized and are sufficiently abundant.[10] The untargeted approach can be split into suspect screening, which involves identifying the unknown chemical features by matching different parameters (e.g., monoisotopic mass) with compounds of interest present in databases; and into non-targeted analysis, which includes various techniques such as annotation by *in silico* fragmentation and/or finding relevant features by statistical analysis.[8,9]

Identifying the chemical features in untargeted datasets is a complex endeavor. Since several chemical formulas can have similar monoisotopic masses, especially for larger masses, at least the fragmentation spectra of the features (MS^2) are needed for tentative compound identification. Data-dependent acquisition (DDA) is one of the main strategies to acquire MS^2

Received: March 31, 2023
Accepted: June 16, 2023
Published: July 6, 2023

https://doi.org/10.1021/acs.analchem.3c01393
Anal. Chem. 2023, 95, 10686−10694

spectra. In this strategy, either a certain number of the most intense ions from the MS^1 scan or ions with defined mass-to-charge ratio (m/z) are fragmented.[11]

Different types of mass spectrometers are used depending on the analytical strategy and the objectives of a certain study, as there is a tradeoff between sensitivity and resolution of the instruments. Targeted workflows typically use sensitive instruments equipped with low-resolution mass analyzers, for example, triple quadrupole. In contrast, untargeted workflows use instruments with lower sensitivity but high resolution, achieved with mass analyzers such as quadrupole-orbitrap or quadrupole-time-of-flight, leading to increased accuracy of the measured m/z values, which is critical in aiding identification. However, with recent developments in mass spectrometry, the sensitivity of HRMS is nearing that of LRMS, notably for low molecular weight compounds such as xenobiotics.[12,13]

Both of these approaches have been applied to investigate polyphenols, however, they were mainly applied on foodstuffs[14] rather than human matrices,[15] especially for untargeted workflows.[16] Thus, the aim of this study was to transfer a previously validated method for a vast set of polyphenol standards[17] from a targeted LRMS to an untargeted HRMS platform. The HRMS workflow was developed using suspect screening and four different non-targeted analysis approaches on urine samples from a pilot study. This allowed for a thorough test of untargeted data analysis techniques, evaluation of available online resources, and exploration of potential limitations. As HRMS is known to be less sensitive than LRMS, this gap was quantified to better characterize the applicability of HRMS on polyphenol analysis in human biofluids. This was achieved by comparing the limits of detection (LODs) of both workflows for 90 polyphenol reference standards from a variety of classes. Finally, both workflows were applied to case study samples to compare their performance at naturally occurring concentrations.

EXPERIMENTAL SECTION

Reagents, Solvents, and Chemicals. Authentic reference standards, reagents, and solvents were purchased from various sources and are provided in Table S1. The reference compounds were diluted in pure methanol to a concentration of 1 mg/mL, considering the purity and density at 20 °C. The individual reference analyte solutions were diluted and mixed together in methanol at various concentrations to prepare working mixes for spiking.

Sample Preparation. The urine samples from the case study and the spiked urine, plasma, and serum samples for the LOD comparison were prepared using the optimized protocol by Oesterle *et al.*[17] The samples were stored at −80 °C until analysis. The concentrations of the spiked samples are reported in Table S2. Enzymatic hydrolysis of glucuronides and sulfates was intentionally not performed in this work to be able to perform neutral loss queries of the features.

LC-HRMS Instrumentation and Parameters. A Vanquish UHPLC (Thermo Fisher) coupled to a QExactive HF quadrupole-Orbitrap (Thermo Fisher) with a heated electrospray ionization source was used for the LC-HRMS measurements. The same LC conditions were used as described in Oesterle *et al.*[17] to best compare the high- and low-resolution measurements, except that eluent B consisted of 0.1% v/v formic acid and 3% v/v water in acetonitrile. Thus, the gradient was adapted and started at 5.2% B, after a 2 min hold, it increased linearly to 66% B within 10 min. B was then set to 97.9% and held for 2 min, before being decreased to 5.2% for a final 2 min hold.

MS^1 and MS^2 data of the case study samples were acquired with DDA of the top 10 most intense peaks in negative polarity. The parameters were set with a scan window of 100 to 1100 Da, a full scan resolution of 60,000, MS^2 resolution of 30,000, a stepped normalized collision energy of 10, 30, and 50 eV, and the additional parameters listed in Table S3. For each biological sample, an iterative exclusion list was prepared with IE-Omics,[18] until either five injections or the maximum length of the exclusion list (5000 features) was reached.

For the LOD comparison, full scan mode from 62 to 900 Da at a resolution of 120,000 was used. Data was acquired for both positive and negative polarities, separately. Additional MS parameters are listed in Table S3. A total of 10 technical replicates were acquired in each polarity at three concentration levels over the linear range: low, medium, and high, corresponding to levels 3, 5, and 7, respectively in Table S2.

Data Analysis. The acquired DDA data was first converted from raw files into MGF files using MSConvert (v3.0.22067).[19] MZMine (v3.1.0)[20] was used for feature preprocessing and extraction of the MS^2 data from the MGF files, for MS^1 screening of the features with the PhytoHub[21] and Exposome-Explorer[22] databases, and for spectral library matching with a library from MS-DIAL composed of databases such as MassBank, ReSpect, and GNPS.[23] To determine molecular formulas and predict structures, *in silico* fragmentation was performed in SIRIUS (v5.5.7)[24] with CSI/FingerID,[25] COSMIC,[26] and ZODIAC.[27] Features of interest from MS^1 matching or *in silico* fragmentation were further matched for annotation with METLIN Gen2 spectral library.[28] All parameters selected for MZMine, METLIN, and SIRIUS are listed in Table S3. The features with MS^2 spectra were additionally screened for sulfate or glucuronide neutral losses with MassQL[29] on GNPS, as described in Table S3. The MassQL queries are also listed in Table S3. MetaboAnalyst (v5.0) was used for statistical analysis.[30] The programming language R (v4.2.0) was used for making plots and performing Van Krevelen analysis.[31] The NTA Study Reporting Tool (SRT) was used as a guideline for reporting the various components of the developed workflow.[32,33]

The MS^1 data acquired for the LOD value comparison was analyzed and evaluated with Skyline (v21.2).[34] The actual concentrations of the analytes were calculated using the area of their chromatographic peaks and a $1/x$ weighted calibration curve. Standard addition was applied and evaluated in Excel if a chromatographic peak was present in the blank samples of the calibration curve. The concentrations of the detected analytes in the case study samples were corrected for extraction efficiency using values previously determined[17] and dilution factor.

The LOD was calculated by first dividing the standard deviation of the concentration of the 10 spiked technical replicates with the square root of the number of replicates and multiplying this value by 3, as advised in the Eurachem guideline[35] (Table S3). The LODs were then multiplied by 20, the dilution factor of the sample preparation.

Biological Samples, Case Study, and Approval to Use Human Matrices. The case study samples were collected as individual 24 h pool from four volunteers at three different time points. The 3 days were: after a polyphenol washout, after consuming a high polyphenol smoothie in the morning, and after a day following a high polyphenol diet, as described in

https://doi.org/10.1021/acs.analchem.3c01393
Anal. Chem. 2023, 95, 10686−10694

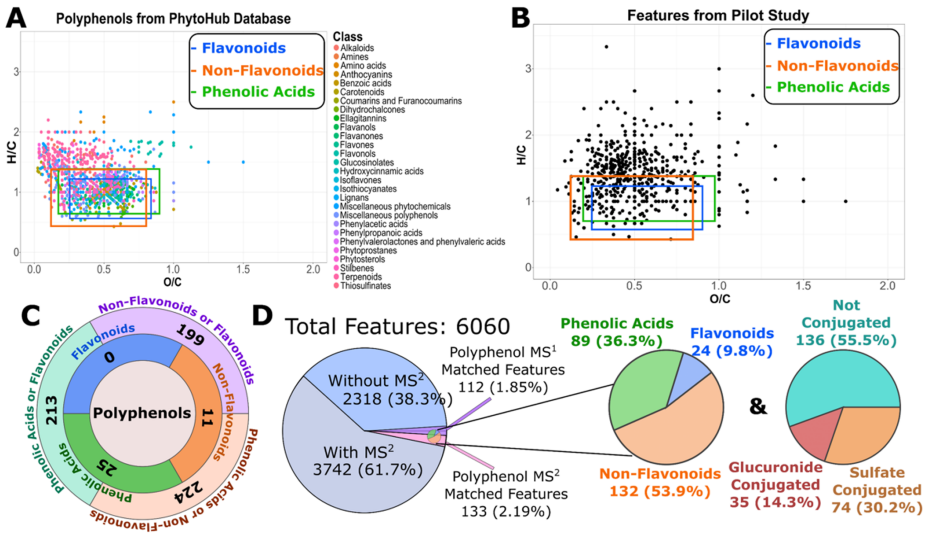


Figure 1. (A) Van Krevelen plot of all >2000 entries in the PhytoHub database. (B) Van Krevelen plot of the features not annotated from the case study samples acquired by DDA which had a molecular formula generated by SIRIUS with a Zodiac score >0.8. (C) A pie chart with the number of unknown features that are potentially polyphenols and classification based on the region they were located on the Van Krevelen plot as shown in (A,B). (D) Pie charts representing the features extracted and the types of polyphenols found from suspect screening.

Oesterle *et al.*[17] The sources of pooled human serum and Li-heparin plasma used for the LOD value comparison are listed in Table S1. The urine used for the matrix matched calibration curve and LOD comparison originated from 24 h pooled urine acquired from a female volunteer after following a low polyphenol diet for two days. The University of Vienna ethics committee, under authorization number #00650, approved the collection and measurement of samples from the participants following receiving their signed consent.

RESULTS AND DISCUSSION

Suspect Screening of Polyphenols and the Exposome in Real-Life Urine Samples. To detect polyphenols and other xenobiotics along with their biotransformation products present in human urine, a suspect screening workflow was developed. However, as urine is an extremely complex and dynamically changing biofluid, these analytes are likely not the most abundant features present. Thus, iterative exclusion DDA was applied to get better coverage of all the features present.[18] While the focus of this work was on polyphenols, the additional screening for other xenobiotics and endogenous metabolites, should highlight the potential of the HRMS approach for comprehensive exposure and effect analysis.

A total of 6060 features were detected, of which 3742 had MS^2 data. To assess which features were potentially polyphenols, the 6060 features were matched by their monoisotopic mass with the PhytoHub database, one of the largest online databases on polyphenols containing currently 2268 analytes.[21] To screen for additional exposome-related features, the features were matched by MS^1 with the Exposome-Explorer database, which contains currently 1212 biomarkers for environmental exposure.[22] Features that were MS^1 matched and had MS^2 data were then annotated by matching to several open and online spectral libraries, or by *in silico* fragmentation.

Feature annotation was based on the identification levels from Schymanski *et al.*[36] but with level 3 split into four sublevels. Features with a library match but for which the correct isomer structure was unknown were labeled as level 3a. Features annotated using *in silico* fragmentation were assigned as level 3b if a tentative structure was known, and level 3c if the elucidated isomer was unknown. Finally, features that were putatively annotated by monoisotopic matching only were assigned a level 3d. For polyphenolic features, a total of 8, 18, 70, and 112 features were assigned to levels 3a−d, respectively. Additionally, 26 features were annotated as level 2a and 11 features could be fully identify at level 1. All of the identified and annotated polyphenolic features are listed in Table S4. From these 245 features, 89 were assigned to phenolic acids, 24 to flavonoids, and 132 to non-flavonoids. These results are displayed in Figure 1d.

When investigating the full range of exposure and effect markers, rather than only polyphenol related features, four additional features were identified as level 1 and 19 annotated as level 2a. Moreover, 6, 23, 13, and 79 additional features were assigned as levels 3a−d, respectively. These features annotated using the Exposome-Explorer database are listed in Table S5.

As a large number of the annotated features were matched with conjugated analytes, the new MassQL algorithm[29] was used to aid in confirming if the features had a glucuronide (176.0321 Da) or sulfate (79.9573 Da) moiety. From the 245 polyphenolic features, 74 features indicated a sulfate

https://doi.org/10.1021/acs.analchem.3c01393
Anal. Chem. 2023, 95, 10686−10694

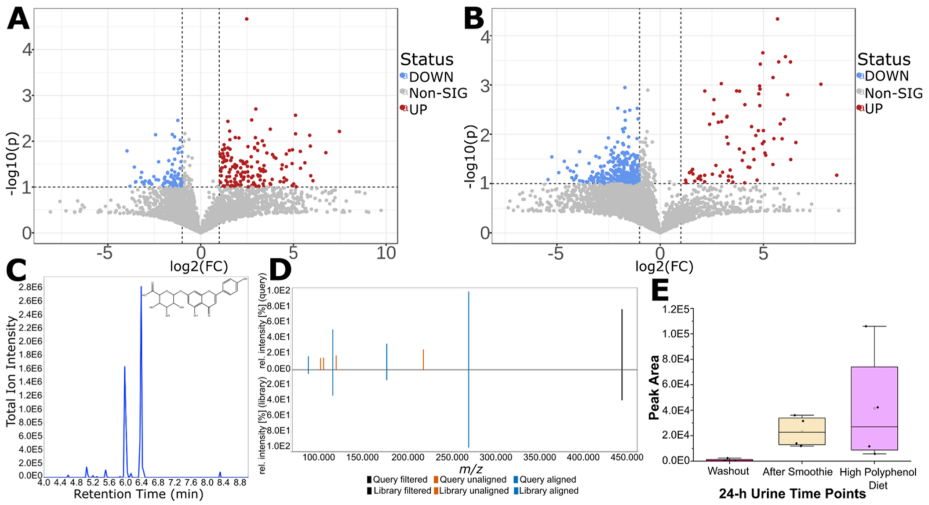


Figure 2. (A) Volcano plot showing the features acquired by iterative DDA of the case study samples which show fold change (FC) versus T-test (p), with significance if $p < 0.1$ and FC > 2, between the 24 h urine samples after a washout and after a day of high polyphenol diet. (B) Similar volcano plot as part (A), but between the 24 h urine samples after a washout and after a high polyphenol smoothie. (C) Extracted ion chromatogram from apigenin-7-*O*-glucuronide (m/z 445.077, RT 6.4), a feature showing significance. (D) Mirror plot of the experimental and spectral library match MS^2 for the selected feature from (C). (E) Box plot showing the change in peak area between the three different time points for the selected feature from (C).

conjugation and 35 a glucuronide conjugate. For the non-polyphenolic features, six indicated glucuronide conjugation and 12 a sulfate conjugate. These results complement a previous study by Jarmusch *et al.*[29] in which seven metabolites were identified with MassQL to have a sulfate moiety.

Suspect screening of polyphenols in human biofluids is typically employed to determine association between polyphenols and the intake of a certain food.[37] The results here are similar to previous research, with the majority of the polyphenol metabolites found in human urine being phenolic acids, including many that are either glucuronidated or sulfated. These results further emphasize the importance of biotransformation products, notably conjugated metabolites, when investigating human urine and potentially other biofluids,[38] which many targeted methods still omit.[6]

Exploring Additional Potential Polyphenol-Related Features through Non-Targeted Analysis. Untargeted data acquisition leads to a considerable number of features, which not only complicates data analysis but many features are a result of noise or of non-pertinent analytes. Therefore, different approaches are used to simplify the data sets and extract unknown features of relevance. Here, four separate approaches were tested and evaluated to extract features that may be polyphenolic but were not able to be annotated previously by *in silico* fragmentation, MS^1 or MS^2 matching. These features are thus labeled as level 4 if a molecular formula was predicted and level 5 if nothing else is known about the feature.

The first approach involved plotting the ratio of hydrogens to carbons (H/C) versus the ratio of oxygens to carbons (O/C) in a compound's molecular formula, also known as a Van Krevelen plot.[31] As seen in previous research,[39] polyphenols tend to aggregate in specific regions of the plot. This aggregation was further exemplified when plotting the entries of PhytoHub (Figure 1a). This region can be split into three parts: phenolic acids, flavonoids, or non-flavonoids. The unknown features which had a molecular formula with a Zodiac score of at least 0.8 were then plotted (Figure 1b). Features were labeled as phenolic acids if they had a H/C between 0.7 and 1.4 and an O/C between 0.13 and 0.9. They were labeled as non-flavonoids if H/C was between 0.5 and 1.4 and O/C between 0.06 and 0.7. Finally, they were labelled as flavonoids if H/C was between 0.6 and 1.3, and O/C between 0.25 and 0.8. This resulted in 224 features listed in Table S6 and depicted in Figure 1c. The issue with this approach is that due to the similarity of the formulas for each of these three groups, the three sections overlap, for example, 199 of the features may be a flavonoid or a non-flavonoid due to this overlap. Moreover, as the entire region of flavonoids overlaps with the other classes, no feature could be indicated as purely flavonoid. Other issues with this approach are that other molecules than polyphenols can fall in the same region of the plot, and it is highly dependent on the accuracy of the molecular formula annotation of the features. Therefore, Van Krevelen plots can be used to easily exclude non-relevant features.

The second approach involved statistical analysis with MetaboAnalyst,[30] which allows to easily filter for relevant features. To apply this approach, the pilot study samples were divided into three groups: 24 h urine following washout, 24 h urine following the ingestion of a high polyphenol smoothie, and 24 h urine following a day of a high polyphenol diet.

https://doi.org/10.1021/acs.analchem.3c01393
Anal. Chem. 2023, 95, 10686−10694

Missing values in the data set were imputed with 1/5 of the minimum positive values. Volcano plots were prepared showing fold change (FC) versus T-test (p) between these groups, with significance if FC > 2 and p < 0.1 (Figure 2a,b). This allowed the detection of 446 significant features (Table S7). Moreover, as for some features no chromatographic peaks were detected in the washout urine samples, unique features from the other two sample groups were screened for, which yielded 226 features (Table S8). It is important to highlight though that from the 672 features, not all of them are expected to be polyphenolic features. These statistical analysis techniques were also applied to the annotated features (Table S4), which would improve confidence of level 3d matched features as significant features should be polyphenol related. For example, the annotated feature apigenin-7-*O*-glucuronide (level 2a) showed a high fold change of 25 from the washout to the 24 h urine following the consumption of a polyphenol smoothie, and of 44 from the washout to the urine after following a high-polyphenol diet (Table S4 and Figure 2e). Moreover, the extracted ion chromatogram for this feature is shown in Figure 2c, and the MS^2 spectra with the library match in Figure 2d.

The third approach to filter for polyphenol-related features involved CANOPUS, a useful tool that allows for compound class annotation from the fragmentation spectra.[40] Features that were labeled with phenol class were filtered, yielding 222 features (Table S9). The advantage of CANOPUS is that structural reference data is not required, but it is prone to errors due to being a computational approximation.

MassQL can be a powerful tool for polyphenols, as specific fragments can be queried, such as backbones of certain polyphenol classes. Additionally, MassQL can be used for neutral loss queries, such as sulfate and glucuronide loss as done here on the 24 h urine samples. These two biotransformation products were chosen because the suspect screening revealed that a large percentage of the features annotated were conjugated xenobiotics. The two queries yielded 502 features with a sulfate loss (Table S10) and 160 features with a glucuronide loss (Table S11). The relationship between the features detected from MassQL, suspect screening, and the other non-targeted analysis techniques is shown in Figure 3. The same queries were then applied to the raw mzML data to find potentially analytes lost during feature processing, for example, with "non-ideal" peak shapes. These queries yielded an additional 1025 hits (Table S12) for glucuronide loss and 10208 hits (Table S13) for sulfate loss that were not previously determined. It has to be noted that the high number of hits is because these are not features yet but rather refer to an MS^2 scan. The number of hits were reduced by clustering and grouping the hits by their m/z and retention times. Another problem with MassQL are identification errors, for example, false positive results, which depend on the selected parameters, such as the ppm deviation.

Limitations of Untargeted HRMS. A general obstacle with polyphenols that complicates feature annotation is that they are a chemical class that contains a vast number of positional, constitutional and stereoisomers, which is complexified if human biotransformation products are included. For example, for many features in Table S4 it is difficult to tell which position the conjugation occurs as MS^2 spectra are extremely similar. Additionally, different unique features were matched with the same analyte, for example, two features with retention times 4.5 and 5.1 min both had a spectral library

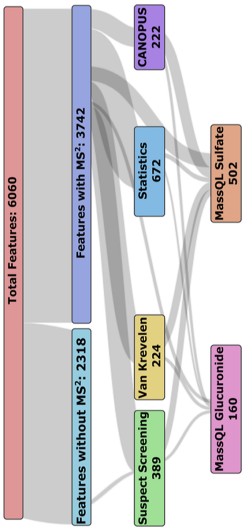


Figure 3. Sankey diagram displaying the number of features extracted from the 24 h urine samples with the various non-targeted analysis techniques and suspect screening.

match with either m-, p-, or o-cresol (Table S4). Chromatographic separation of isomers is a challenging endeavor, as seen here where apigenin and genistein both have the same monoisotopic mass and their retention times are both 8.33 (Table S4). Therefore, they will have the same chromatographic peak and their MS^2 spectra will be acquired together. This is not an issue in targeted workflows as these two compounds will have distinct quantifier/qualifier ions. This may be overcome in untargeted workflows by adding ion mobility as a complementary dimension to HRMS to separate co-eluting compounds and increase the confidence of structural elucidation.[41]

The main challenge in untargeted approaches is compound annotation and identification. One limitation with this is the size and quality of spectral libraries, especially as different types of mass spectrometers or collision energies yield different MS^2 spectra. Other techniques can be used for structural annotation of more features, such as *in silico* fragmentation, as seen here with more features annotated at levels 3b and 3c than levels 2a and 3a. However, *in silico* fragmentation has its own pitfalls as it is a computational prediction of MS^2 spectra. Moreover, *in silico* fragmentation techniques require reference structures as an input and thus would not be suitable for analytes which no structure has been previously described. Being able to elucidate structures from features of completely unknown analytes is extremely difficult, but combining various approaches, such as those described here, can yield valuable insight into the nature of these unknown features. Moreover, the use of complementary approaches can increase the confidence in their results.

Finally, one intrinsic limitation of untargeted workflows is that the HRMS instruments used typically trade-off sensitivity

https://doi.org/10.1021/acs.analchem.3c01393
Anal. Chem. 2023, 95, 10686−10694

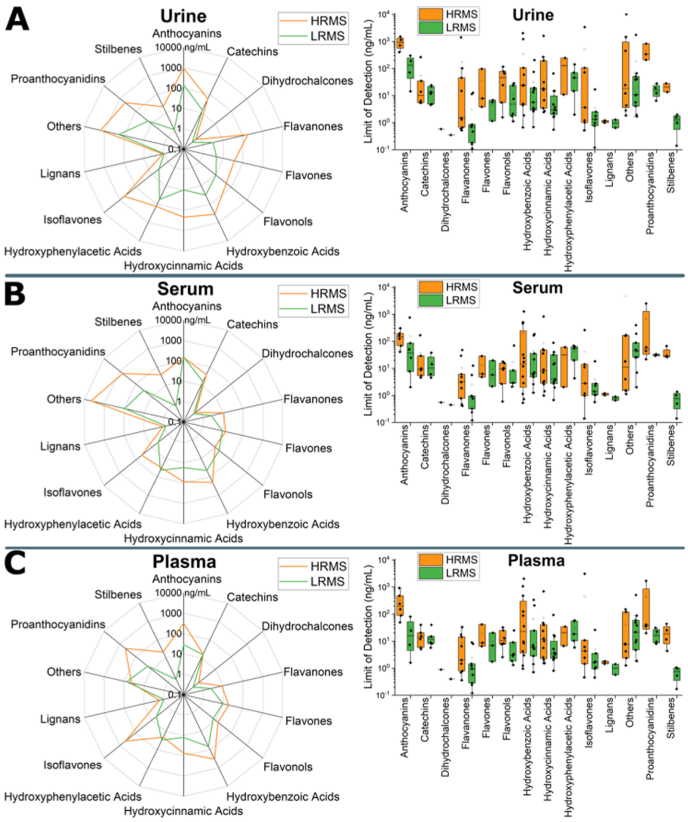


Figure 4. Spider and box plots of the LODs of 90 polyphenols depicting the difference in sensitivity between the HRMS (orange) and LRMS (green) workflows in three different human biofluids: (a) urine, (b) serum, and (c) plasma. LRMS data were reported before by Oesterle *et al.*[17]

for having higher resolution. Though as not all analytes and instruments perform the same, it is valuable to know the sensitivity of an HRMS workflow to better consider itś applicability in a specific study.

Comparing the Sensitivity of HRMS and LRMS Platforms Using Spiked Samples. To quantitatively evaluate the difference between the LRMS and HRMS workflows for polyphenols, a sensitivity comparison was performed by calculating the LODs of reference standards (Table S14) in HRMS spiked at a low level (level 2, Table S2) in the three matrices. To keep conditions as similar as possible, the same concentrations were used as in the LRMS method. However, due to the different sensitivities of the instruments, a chromatographic peak could not be integrated for certain analytes. Thus, the LOD calculations were adapted to the number of replicates, or if less than three replicates were available, medium or high spiked samples were used instead.

There are various techniques to calculate LODs. One common technique is to use the signal-to-noise ratio and calculate at which concentration this ratio would equal to 3. However, in both LRMS and HRMS, there is not always a defined background noise level. Therefore, another approach was used which involves calculating the LOD from standard deviation of multiple technical replicates, as recommended in the Eurachem guideline[35] and used in the LRMS workflow. Though different HRMS acquisition modes are available, such as product ion scan, data were acquired in full scan as it is the most commonly used data acquisition mode and typically one of the first steps in untargeted workflows.

The majority of the 90 model analytes listed in Table S14 were detected in the HRMS measurements with only 16, eight, and nine analytes not found at any spiked concentration in urine, serum, and plasma, respectively (reported as "n.d." in Table S14). Several analytes had background contamination in the matrix matched "blank", thus standard addition was applied. The analytes were 3-methylcatechol, 4-hydroxybenzoic acid, enterolactone, and hippuric acid for urine; 3,3-hydroxyphenylpropanoic acid, hippuric acid, and salicylic acid

https://doi.org/10.1021/acs.analchem.3c01393
Anal. Chem. 2023, 95, 10686−10694

for serum; and 3,3-hydroxyphenylpropanoic acid, benzoic acid, hippuric acid, and salicylic acid for plasma.

Figure 4 and Table S14 illustrate that for urine, 12 out of 90 analytes (13%) displayed a higher sensitivity in HRMS than in LRMS. This value is 29 (32%) for serum and 22 (24%) for plasma. Moreover, the plots in Figure 4 show that the majority of the polyphenol classes have similar average LODs between the two instruments. The median of the LODs of urine, serum, and plasma in the HRMS instrument were 18, 10, and 11 ng/mL, respectively, compared to 4.8, 5.8, and 5.2 ng/mL, respectively, in LRMS. However, the box plots of Figure 4 show that HRMS has a larger variance between the different analytes in each polyphenol class than LRMS.

It is reported in literature that depending on the analytes and instruments chosen, HRMS may perform better or worse than LRMS.[42] In one study comparing HRMS and LRMS for a multi-class xenobiotic method,[12] the authors concluded that LRMS would be more suited and the preferred platform for quantification of such compounds. Different from those small molecules, polyphenols are a more homogenous class, present in higher concentrations and with overall better ionization efficiency due to stabilization of proton loss by their aromatic rings and lower background noise in negative ESI mode.[43] Therefore, as shown here, the sensitivity of HRMS is similar to that of LRMS for many polyphenols, with the additional potential of a broader screening that includes their secondary metabolites (*e.g.*, sulfates and glucuronides).

Qualitative and Quantitative Comparison in Real-Life Urine Samples. Samples spiked with authentic reference standards are useful when developing workflows. However, comparing the results of HRMS and LRMS platforms in real-life samples yields additional information of the sensitivity gap between LRMS and HRMS.

With the HRMS instrument, 23 out of the 90 selected model polyphenols were detected, while 37 analytes were detected with LRMS.[17] From these analytes, 18 were detected in both platforms, with their concentrations around a factor 3 lower in HRMS than in LRMS (Table S15). Thus, the 5 analytes that were detected with HRMS but not LRMS could be from the analyte having a better LOD in the HRMS, such as is the case for caffeic acid. Alternatively, it could be a false positive or interference with another analyte of the same MS^1 mass since in HRMS only a MS^1 scan was acquired and no fragmentation data.

A difference in determined concentrations between HRMS and LRMS workflows is normal, for example, up to a 20% difference was found for pesticides in various foodstuffs by del Mar Gómez-Ramos *et al.*[44] Though here the difference between the two workflows was a factor of 3, most likely due to human biofluids being more complex matrices than foodstuffs and having more interferences. Although the difference here being larger than in previous research, the majority of the analytes are still detected, showing that HRMS is suitable for the qualitative and semi-quantitative analysis of polyphenols and their metabolites in human samples.

CONCLUSION AND OUTLOOK

Despite several limitations of untargeted approaches, HRMS can successfully be utilized for investigating the exposome, notably polyphenols and their metabolites in human samples besides many other exogenous and endogenous small molecules. For polyphenols, although HRMS has shown to have lower sensitivity for most of them, the differences are still in the range where HRMS is able to detect—and better characterize—the vast majority of polyphenols. Due to their higher naturally occurring concentration, for example, concentrations ranging from 0.01 μM to over 1000 μM in human urine,[15] and the more in-depth knowledge about these molecules compared to many other xenobiotics,[12] HRMS can readily be used instead of LRMS for qualitative and semi-quantitative analysis of polyphenols with a great gain of information. Although targeted LRMS methods are still desired for benchmarking, HRMS has the advantage that it allows to study the exposome agnostically. Currently, polyphenols are frequently under-investigated in human matrices, but workflows such as the one presented here would allow to better investigate both positive and negative effects of these xenobiotics. For example, this workflow can be applied to investigate toxicological and pharmacological mixture effects of polyphenols with other xenobiotics, such as decreasing the potency of potentially harmful xenobiotics or interfering with drug treatment.[45] Therefore, the authors strongly recommend the use of HRMS in complement to LRMS in the future investigation of exposure and effects of polyphenols and other xenobiotics in human nutritional and health studies.

ASSOCIATED CONTENT

Data Availability Statement

The raw data files have been submitted to the Metabolights data repository (MTBLS7564).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.3c01393.

Supporting Information (Excel) is included which contains tables with additional information on the material and methods, and the detailed results from the workflow, such as the LODs for each reference standard or the features filtered by MassQL (XLSX)

Moreover, the completed BP4NTA Study Reporting Tool is included for enhanced reporting confidence (XLSX)

AUTHOR INFORMATION

Corresponding Author

Benedikt Warth − *Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Vienna 1090, Austria; Exposome Austria, Research Infrastructure and National EIRENE Hub, Vienna 1090, Austria;* orcid.org/0000-0002-6104-0706; Phone: +43-1-4277-70806; Email: benedikt.warth@univie.ac.at

Authors

Ian Oesterle − *Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Vienna 1090, Austria; Fakultät für Chemie, Institut für Biophysikalische Chemie, Universität Wien, Wien 1090, Austria; Doctoral School of Chemistry, University of Vienna, Vienna 1090, Austria;* orcid.org/0000-0003-4106-5452

Manuel Pristner − *Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Vienna 1090, Austria; Doctoral School of Chemistry, University of Vienna, Vienna 1090, Austria*

Sabrina Berger − *Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Vienna 1090, Austria*

https://doi.org/10.1021/acs.analchem.3c01393
Anal. Chem. 2023, 95, 10686−10694

Mingxun Wang – *Department of Computer Science, University of California Riverside, Riverside, California 92521, United States*

Vinicius Verri Hernandes – *Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Vienna 1090, Austria; Exposome Austria, Research Infrastructure and National EIRENE Hub, Vienna 1090, Austria*

Annette Rompel – *Fakultät für Chemie, Institut für Biophysikalische Chemie, Universität Wien, Wien 1090, Austria;* orcid.org/0000-0002-5919-0553

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.3c01393

Author Contributions

The manuscript was written through contributions of all authors.

Funding

Open Access is funded by the Austrian Science Fund (FWF).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to acknowledge all members of the Warth and Rompel labs for providing feedback and valuable support during method development and optimization. They also express their gratitude to the Mass Spectrometry Center of the Faculty of Chemistry, University of Vienna, for technical support. This work was supported by the University of Vienna via the Exposome Austria Research Infrastructure and the Austrian Science Fund FWF: P32326 (A.R.).

■ REFERENCES

(1) Pahlke, G.; Ahlberg, K.; Oertel, A.; Janson-Schaffer, T.; Grabher, S.; Mock, H.-P.; Matros, A.; Marko, D. *Mol. Nutr. Food Res.* **2021**, *65*, 2100229.
(2) Shahidi, F.; Yeo, J. *Int. J. Mol. Sci.* **2018**, *19*, 1573.
(3) Baky, M. H.; Elshahed, M.; Wessjohann, L.; Farag, M. A. *Br. J. Nutr.* **2022**, *128*, 577−591.
(4) Aichinger, G.; Bliem, G.; Marko, D. *Front. Nutr.* **2021**, *8*, 293.
(5) Preindl, K.; Braun, D.; Aichinger, G.; Sieri, S.; Fang, M.; Marko, D.; Warth, B. *Anal. Chem.* **2019**, *91*, 11334−11342.
(6) Miller, G. W.; Jones, D. P. *Toxicol. Sci.* **2014**, *137*, 1−2.
(7) Oesterle, I.; Braun, D.; Berry, D.; Wisgrill, L.; Rompel, A.; Warth, B. *Annu. Rev. Food Sci. Technol.* **2021**, *12*, 461−484.
(8) Bocato, M. Z.; Bianchi Ximenez, J. P.; Hoffmann, C.; Barbosa, F. *J. Toxicol. Environ. Health - Part B Crit. Rev.* **2019**, *22*, 131−156.
(9) Gorrochategui, E.; Jaumot, J.; Lacorte, S.; Tauler, R. *TrAC Trends Anal. Chem.* **2016**, *82*, 425−442.
(10) Dmitrenko, A.; Reid, M.; Zamboni, N. *Bioinformatics* **2023**, *39*, btad096.
(11) Guo, J.; Huan, T. *Anal. Chem.* **2020**, *92*, 8072−8080.
(12) Flasch, M.; Köllensperger, G.; Warth, B. *ChemRxiv* **2022**, DOI: 10.26434/chemrxiv-2022-x4kk7.
(13) Rochat, B. *TrAC Trends Anal. Chem.* **2016**, *84*, 151−164.
(14) Flamini, R.; De Rosso, M.; Bavaresco, L. *J. Anal. Methods Chem.* **2015**, *2015*, 1−10.
(15) Achaintre, D.; Buleté, A.; Cren-Olivé, C.; Li, L.; Rinaldi, S.; Scalbert, A. *Anal. Chem.* **2016**, *88*, 2637−2644.
(16) Ávila-Gálvez, M. Á.; García-Villalba, R.; Martínez-Díaz, F.; Ocaña-Castillo, B.; Monedero-Saiz, T.; Torrecillas-Sánchez, A.; Abellán, B.; González-Sarrías, A.; Espín, J. C. *Mol. Nutr. Food Res.* **2019**, *63*, 1801239.
(17) Oesterle, I.; Braun, D.; Rompel, A.; Warth, B. *Anal. Chim. Acta* **2022**, *1216*, 339977.
(18) Koelmel, J. P.; Kroeger, N. M.; Gill, E. L.; Ulmer, C. Z.; Bowden, J. A.; Patterson, R. E.; Yost, R. A.; Garrett, T. J. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 908−917.
(19) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. *Nat. Biotechnol.* **2012**, *30*, 918−920.
(20) Schmid, R.; Heuckeroth, S.; Korf, A.; Smirnov, A.; Myers, O.; Dyrlund, T. S.; Bushuiev, R.; Murray, K. J.; Hoffmann, N.; Lu, M.; Sarvepalli, A.; Zhang, Z.; Fleischauer, M.; Dührkop, K.; Wesner, M.; Hoogstra, S. J.; Rudt, E.; Mokshyna, O.; Brungs, C.; Ponomarov, K.; Mutabdžija, L.; Damiani, T.; Pudney, C. J.; Earll, M.; Helmer, P. O.; Fallon, T. R.; Schulze, T.; Rivas-Ubach, A.; Bilbao, A.; Richter, H.; Nothias, L.-F.; Wang, M.; Orešič, M.; Weng, J.-K.; Böcker, S.; Jeibmann, A.; Hayen, H.; Karst, U.; Dorrestein, P. C.; Petras, D.; Du, X.; Pluskal, T. *Nat. Biotechnol.* **2023**, *41*, 447−449.
(21) Giacomoni, F.; Fillatre, Y.; Rothwell, J. A.; Knox, C.; Eisner, R.; Cesaire, D.; Quintana, M.; Comte, B.; Pujos-Guillot, E.; Manach, C. PhytoHUB: a new database dedicated to dietary phytochemicals and their human metabolites for nutritional metabolomics. *9th Annual Conference of the Metabolomics Society*: Glasgow, GBR, 2013.
(22) Neveu, V.; Nicolas, G.; Salek, R. M.; Wishart, D. S.; Scalbert, A. *Nucleic Acids Res.* **2020**, *48*, D908−D912.
(23) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nat. Methods* **2015**, *12*, 523−526.
(24) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. *Nat. Methods* **2019**, *16*, 299−302.
(25) Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 12580−12585.
(26) Hoffmann, M. A.; Nothias, L.-F.; Ludwig, M.; Fleischauer, M.; Gentry, E. C.; Witting, M.; Dorrestein, P. C.; Dührkop, K.; Böcker, S. *bioRxiv* **2021**, DOI: 10.1101/2021.03.18.435634.
(27) Ludwig, M.; Nothias, L.-F.; Dührkop, K.; Koester, I.; Fleischauer, M.; Hoffmann, M. A.; Petras, D.; Vargas, F.; Morsy, M.; Aluwihare, L.; Dorrestein, P. C.; Böcker, S. *Nat. Mach. Intell.* **2020**, *2*, 629−641.
(28) Guijas, C.; Montenegro-Burke, J. R.; Domingo-Almenara, X.; Palermo, A.; Warth, B.; Hermann, G.; Koellensperger, G.; Huan, T.; Uritboonthai, W.; Aisporna, A. E.; Wolan, D. W.; Spilker, M. E.; Benton, H. P.; Siuzdak, G. *Anal. Chem.* **2018**, *90*, 3156−3164.
(29) Jarmusch, A. K.; Aron, A. T.; Petras, D.; Phelan, V. V.; Bittremieux, W.; Acharya, D. D.; Ahmed, M. M. A.; Bauermeister, A.; Bertin, M. J.; Boudreau, P. D.; Borges, R. M.; Bowen, B. P.; Brown, C. J.; Chagas, F. O.; Clevenger, K. D.; Correia, M. S. P.; Crandall, W. J.; Crüsemann, M.; Damiani, T.; Fiehn, O.; Garg, N.; Gerwick, W. H.; Gilbert, J. R.; Globisch, D.; Gomes, P. W. P.; Heuckeroth, S.; James, C. A.; Jarmusch, S. A.; Kakhkhorov, S. A.; Kang, K. B.; Kersten, R. D.; Kim, H.; Kirk, R. D.; Kohlbacher, O.; Kontou, E. E.; Liu, K.; Lizama-Chamu, I.; Luu, G. T.; Knaan, T. L.; Marty, M. T.; McAvoy, A. C.; McCall, L.-I.; Mohamed, O. G.; Nahor, O.; Niedermeyer, T. H. J.; Northen, T. R.; Overdahl, K. E.; Pluskal, T.; Rainer, J.; Reher, R.; Rodriguez, E.; Sachsenberg, T. T.; Sanchez, L. M.; Schmid, R.; Stevens, C.; Tian, Z.; Tripathi, A.; Tsugawa, H.; Nishida, K.; Matsuzawa, Y.; van der Hooft, J. J. J.; Vicini, A.; Walter, A.; Weber, T.; Xiong, Q.; Xu, T.; Zhao, H. N.; Dorrestein, P. C.; Wang, M. *bioRxiv* **2022**, DOI: 10.1101/2022.08.06.503000.
(30) Pang, Z.; Zhou, G.; Ewald, J.; Chang, L.; Hacariz, O.; Basu, N.; Xia, J. *Nat. Protoc.* **2022**, *17*, 1735−1761.
(31) Rivas-Ubach, A.; Liu, Y.; Bianchi, T. S.; Tolić, N.; Jansson, C.; Paša-Tolić, L. *Anal. Chem.* **2018**, *90*, 6152−6160.

(32) Peter, K. T.; Phillips, A. L.; Knolhoff, A. M.; Gardinali, P. R.; Manzano, C. A.; Miller, K. E.; Pristner, M.; Sabourin, L.; Sumarah, M. W.; Warth, B.; Sobus, J. R. *Anal. Chem.* **2021**, *93*, 13870−13879.
(33) BP4NTA. *NTA Study Reporting Tool (Excel)*; Figshare, 2022. https://doi.org/10.6084/m9.figshare.19763503.
(34) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. *Bioinformatics* **2010**, *26*, 966−968.
(35) Magnusson, B.; Örnemark, U. *Eurachem Guide: The Fitness for Purpose of Analytical Methods—A Laboratory Guide to Method Validation and Related Topics*, 2nd ed., 2014.
(36) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. *Environ. Sci. Technol.* **2014**, *48*, 2097−2098.
(37) Zhu, Y.; Wang, P.; Sha, W.; Sang, S. *Sci. Rep.* **2016**, *6*, 36278.
(38) Fareed, Y.; Braun, D.; Flasch, M.; Globisch, D.; Warth, B. *Exposome* **2022**, *2*, osac008.
(39) Roullier-Gall, C.; Lucio, M.; Noret, L.; Schmitt-Kopplin, P.; Gougeon, R. D. *PLoS One* **2014**, *9*, No. e97615.
(40) Dührkop, K.; Nothias, L.-F.; Fleischauer, M.; Reher, R.; Ludwig, M.; Hoffmann, M. A.; Petras, D.; Gerwick, W. H.; Rousu, J.; Dorrestein, P. C.; Böcker, S. *Nat. Biotechnol.* **2021**, *39*, 462−471.
(41) de Bruin, C. R.; Hennebelle, M.; Vincken, J.-P.; de Bruijn, W. J. C. *Anal. Chim. Acta* **2023**, *1244*, 340774.
(42) Cavaliere, C.; Antonelli, M.; Capriotti, A. L.; La Barbera, G.; Montone, C. M.; Piovesana, S.; Laganà, A. *J. Agric. Food Chem.* **2019**, *67*, 4885−4896.
(43) Fulcrand, H.; Mané, C.; Preys, S.; Mazerolles, G.; Bouchut, C.; Mazauric, J.-P.; Souquet, J.-M.; Meudec, E.; Li, Y.; Cole, R. B.; Cheynier, V. *Phytochem* **2008**, *69*, 3131−3138.
(44) del Mar Gómez-Ramos, M.; Rajski, Ł.; Heinzen, H.; Fernández-Alba, A. R. *Anal. Bioanal. Chem.* **2015**, *407*, 6317−6326.
(45) Warth, B.; Raffeiner, P.; Granados, A.; Huan, T.; Fang, M.; Forsberg, E. M.; Benton, H. P.; Goetz, L.; Johnson, C. H.; Siuzdak, G. *Cell Chem. Bio.* **2018**, *25*, 291−300.e3.

https://doi.org/10.1021/acs.analchem.3c01393
Anal. Chem. 2023, 95, 10686−10694

Supplementary Information

Two supplementary data files are available online for this publication. The first is an xlsx that contains various additional tables, such as the detailed results and the MS parameters. The second file is the completed BP4NTA Study Reporting Tool (xlsx).

DOI: https://doi.org/10.1021/acs.analchem.3c01393

6.5. Publication #4 (Original Research): Oesterle et al. 2023b

Status	Submitted for publication
Title	**Insights into the early-life chemical exposome of Nigerian infants and potential correlations with the developing gut microbiome**
Authors	Ian Oesterle[a,b,c], Kolawole I. Ayeni[a,d], Chibundu N. Ezekiel[d,e], David Berry[f], Annette Rompel[b], Benedikt Warth[a,g]
Affiliations	[a]University of Vienna, Faculty of Chemistry, Department of Food Chemistry and Toxicology, 1090 Vienna, Austria [b]Universität Wien, Fakultät für Chemie, Institut für Biophysikalische Chemie, 1090 Wien, Austria; www.bpc.univie.ac.at [c]University of Vienna, Vienna Doctoral School of Chemistry (DoSChem), 1090 Vienna, Austria [d]Department of Microbiology, Babcock University, Ilishan-Remo, Ogun State, Nigeria [e]University of Natural Resource and Life Science Vienna (BOKU), Department of Agrobiotechnology (IFA-Tulln), Institute for Bioanalytics and Agro-Metabolomics, Konrad-LorenzStr. 20, 3430 Tulln, Austria [f]University of Vienna, Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, Centre for Microbiology and Environmental Systems Science, 1090 Vienna, Austria [g]Exposome Austria, Research Infrastructure and National EIRENE Node, Austria
Year	2023
Journal	*bioRxiv*
Journal Information	Pre-print
Accepted	n.a.
DOI	https://doi.org/10.1101/2023.11.08.566030
Contribution	Ian Oesterle took part in designing the experiment, took the lead in sample preparation, performed the LC-HRMS measurements, performed feature processing and result evaluation, took part in the statistical analysis and result interpretation, and wrote the original draft of the manuscript.

Insights into the early-life chemical exposome of Nigerian infants and potential correlations with the developing gut microbiome

Ian Oesterle[a,b,c], Kolawole I. Ayeni[a,d], Chibundu N. Ezekiel[d,e], David Berry[f], Annette Rompel[b], Benedikt Warth[a,g,*]

[a]University of Vienna, Faculty of Chemistry, Department of Food Chemistry and Toxicology, 1090 Vienna, Austria
[b]Universität Wien, Fakultät für Chemie, Institut für Biophysikalische Chemie, 1090 Wien, Austria; www.bpc.univie.ac.at
[c]University of Vienna, Vienna Doctoral School of Chemistry (DoSChem), 1090 Vienna, Austria
[d]Department of Microbiology, Babcock University, Ilishan-Remo, Ogun State, Nigeria
[e]University of Natural Resource and Life Science Vienna (BOKU), Department of Agrobiotechnology (IFA-Tulln), Institute for Bioanalytics and Agro-Metabolomics, Konrad-LorenzStr. 20, 3430 Tulln, Austria.
[f]University of Vienna, Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, Centre for Microbiology and Environmental Systems Science, 1090 Vienna, Austria
[g]Exposome Austria, Research Infrastructure and National EIRENE Node, Austria
*Corresponding author: Benedikt Warth, benedikt.warth@univie.ac.at, +43-1-4277-70806.

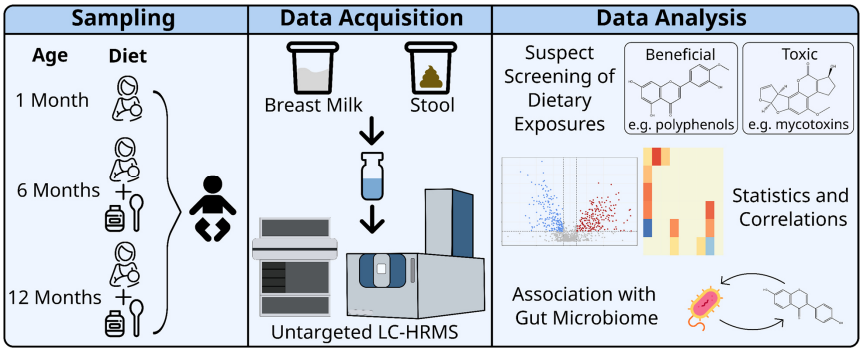
Abstract

Early-life exposure to natural and/or synthetic chemicals can impact acute and chronic health conditions. Here, a suspect screening workflow anchored on high-resolution mass spectrometry was applied to elucidate xenobiotics in breast milk and matching stool samples collected from Nigerian mother-infant pairs (n = 11) at three-time points. Potential correlations between xenobiotics and the gut microbiome, as determined by 16S rRNA gene amplicon sequencing, were subsequently explored. Overall, 12,192 and 16,461 features were acquired in the breast milk and stool samples, respectively. Following quality control and suspect screening, 562 and 864 features remained, respectively, with 149 of these features present in both matrices. Taking advantage of 242 authentic reference standards measured for confirmatory purposes of both, potentially beneficial and adverse xenobiotics, 34 features in breast milk and 68 features in stool were unambiguously identified and subsequently semi-quantified. Moreover, 51 and 78 features were annotated with spectral library matching, as well as 416 and 652 with *in silico* fragmentation in breast milk and stool, respectively. Despite that the workflow was originally optimized for polyphenols, a diverse range of other chemical classes were simultaneously identified including mycotoxins, endocrine-disrupting chemicals (EDCs), antibiotics, plasticizers, perfluorinated alkylated substances, and pesticides. Spearman rank correlation of the identified features revealed significant correlations between chemicals of the same classification such as polyphenols. One-way ANOVA and differential abundance analysis of the data obtained from stool samples revealed that features of plant-based origin were elevated when complementary foods were introduced to the infants' diets. Features in the stool deemed significant by ANOVA, such as tricetin, positively correlated with the genus *Blautia*. Moreover, vulgaxanthin negatively correlated with *Escherichia-Shigella*. Despite the limited sample size, this exploratory study provides high-quality exposure data of matched biospecimens obtained from mother-infant pairs in sub-Saharan Africa, and showed potential correlations between the chemical exposome and the gut microbiome.

Keywords

Non-targeted screening/analysis, Exposomics, Infant stool, Human breast milk, Food bioactives, Toxicants

Graphical abstract



1. Introduction

In 2005, the "exposome", an idea that expands and brings together the already established "-ome" research fields, was introduced by Wild (2005). The concept can be summarized as investigating the totality of chemicals that an individual is exposed to during their lifetime, and the consequent health-related effects (Miller and Jones 2014). The main focus of current exposome research lies on xenobiotics and their biotransformation products, and can be further complemented with other -omic techniques, e.g. proteomics or metagenomics, to characterize the biological impact of these exogenous chemicals (Vitale et al. 2021, Kalia et al. 2022).

Exposure to various xenobiotics during one´s lifetime, especially during the first thousand days, significantly influences human health and the development of chronic diseases (Heindel et al. 2015). For example, exposure to endocrine disrupting chemicals, such as bisphenol A, have been associated with an increased risk of diseases like obesity and neurodevelopmental disorders (Braun 2017). Early-life exposure to mycotoxins, such as aflatoxins and fumonisins, can adversely influence child growth (IARC 2015, Rasheed et al. 2021). Conversely, exposure to other xenobiotic classes, for instance polyphenols, can have potentially beneficial health effects such as reduced risk for breast cancer due to soy isoflavones consumption in early life (Wu et al. 2008). The diet of infants is a major source of exposure to both potentially toxic or beneficial xenobiotics. For example, consumption of complementary foods rather than breast milk can expose infants to higher levels of mycotoxins (Krausová et al. 2023).

Exposome research has mainly focused on adverse exposures and often does not include molecules that might be able to mitigate these actions. One vital class of natural xenobiotics is polyphenols, which are prevalent in plant- and fungi-based foods. They can impact the microbiome by, for example, promoting the growth of beneficial bacteria (Gowd et al. 2019). Additionally, polyphenols are bioactive compounds in humans, exhibiting properties that have shown to be either positive, like antimicrobial or anti-inflammatory properties (Pandey and Rizvi 2009), or potentially negative, such as estrogenic effects of isoflavones (Křížová 2019). Polyphenols can also reduce the toxicity of other xenobiotics, as seen with mycotoxins (Rasouli et al. 2022). Therefore, assessing human exposure to polyphenols is essential. However, this is a challenging task as polyphenols are an immense class of molecules that, once ingested, yield a variety of biotransformation products, from both human and microbial metabolism (Oesterle et al. 2021).

Since xenobiotics, including polyphenols, contain chemically diverse molecules, assessing human exposure to these chemicals requires holistic and cutting-edge exposomic workflows anchored on mass spectrometry (MS) (Bocato et al. 2019, Flasch et al. 2022b, Gu et al. 2023). Untargeted approaches, compared to targeted approaches, are advantageous as they allow the exploration of a broader spectrum of analytes rather than simply those with an available reference standard (Oesterle et al. 2021). However, identification of the analytes in untargeted approaches is challenging as only features, defined as a detected signal intensity and an associated *m/z* and retention time value, are acquired. Therefore, at least the fragmentation spectra of the features are required. One untargeted approach to acquire both features and their fragmentation spectra simultaneously is sequential window acquisition theoretical hold (SWATH) data-independent acquisition (DIA), which involves fragmenting all ions in sequential windows of monoisotopic masses (Bonner and Hopfgartner 2019). Though deconvolution of the acquired data is more complex than data-

dependent acquisition (DDA), SWATH DIA yields fragmentation spectra for all features present rather than specified features (Guo and Huan 2020), allowing retrospective feature annotation. As untargeted approaches generate a very high number of features, suspect screening can be employed to filter the features for analytes of interest thought to be present in the samples (Pourchet et al. 2020).

Besides xenobiotics, the gut microbiome has a major and highly complex impact on human health (Hou et al. 2022). The composition of the gut microbiome can be modulated through diet (Claesson et al. 2012, David et al. 2014, Wilson et al. 2020) and environmental exposures (Claus et al. 2016). Conversely, the microbiome can influence xenobiotics (Collins and Patterson 2020), leading to a bi-directional relationship between the gut microbiome and xenobiotics (Clarke et al. 2019). Xenobiotics can contribute to, or drive, dysbiosis during early-life (Chi et al. 2021, Ayeni et al. 2022), especially as the microbiome of infants is not yet fully established (Wopereis et al. 2014). Thus, research efforts have been geared towards understanding xenobiotic and early-life microbiome interactions.

Several studies have reported on metabolomic profiles of healthy children (Chen et al. 2019, Holzhausen et al. 2023), children with severe acute malnutrition (McMillan et al. 2016), or comparing breastfed infants with infants fed formula (Brink et al. 2020, Silner et al. 2021). However, to our knowledge, there is limited data on the longitudinal exposomic/metabolomic profiles of neonates and infants via breast milk and matching stool samples, which is necessary for making correlations with the gut microbiome. Thus, the objectives of this study were to apply a recently developed untargeted exposomic biomonitoring workflow (Oesterle et al. 2023) with SWATH DIA on breast milk and matching stool collected longitudinally from Nigerian mother-infant pairs. This workflow allowed us to 1) elucidate exposure profiles of polyphenols and other potential beneficial or toxic xenobiotics present in the samples, 2) investigate changes and correlations of xenobiotics as complementary foods are introduced in the diet, and 3) correlate the xenobiotics detected in the infants' stool to the gut microbiome.

2. Materials and methods

2.1. Study design, sampling, and ethical approval

The longitudinal pilot study involved human breast milk and matched infant stool samples. Details of study location and sample collection were previously described by Ayeni et al. (under review). The same set of samples was used before to investigate specific exposure classes: polyphenols in the breast milk samples (Berger et al. submitted), and mycotoxins in the stool samples (Krausová et al. 2022) and breast milk samples (Ayeni et al. under review). Briefly, samples were provided by eleven Nigerian infant-mother pairs from Ilishan-Remo, Ogun state. Breast milk and infant stool samples were collected at the same time point by the mothers, and temporarily stored at 4°C prior to further cold storage. Samples were then stored at -20°C until analysis. Food frequency questionnaires were administered to the mothers to ascertain dietary patterns and health status of the infants (unpublished data). Ethical approval was obtained from the Ethical Committee of Babcock University (BUHREC 421/21R, BUHREC 466/23). All mothers were properly informed before providing their written consent to be included in the study.

2.2. Reagents and chemicals

A total of 242 authentic reference standards were used in this study, allowing for Level 1

identifications of relevant xenobiotics, benchmarking the approach, and enabling absolute quantitation for a high number of relevant environmental and food-related exposures. The selection represented different xenobiotics, both synthetic and natural, containing, for example, mycotoxins, plasticizers, antibiotics, and all the main polyphenol classes. These standards were acquired at the highest possible purity and their acquisition information and molecular class are listed in Table A.1. Acetonitrile (ACN) and methanol (MeOH), LC-MS grade, were acquired from Honeywell. Water (H_2O), LC-MS grade, was acquired from VWR, and formic acid (FA), UPLC-MS Optigrade, was acquired from Bartelt. Anhydrous magnesium sulfate ($MgSO_4$) and sodium chloride (NaCl) were acquired from Sigma-Aldrich. The stock solutions were prepared by diluting the solid standards (Table A.1) in either MeOH (polyphenols) or ACN (other standards) and used to make various working mixes to enrich the samples.

2.3. Sample preparation

The breast milk samples were prepared following a protocol optimized by Berger et al. (submitted) for diverse polyphenols. In brief, 200 µL aliquots of breast milk were diluted 1:2 with acidified ACN (1% v/v FA) in a micro-reaction tube. The samples were vortexed for 3 min. Then, 4 mg of $MgSO_4$ and 1 mg of NaCl per 10 µL of matrix were added and the samples were again vortexed for 3 min. Next, the samples were centrifuged for 10 min at 4°C and 2,000 x g, and the supernatant was transferred to a new micro-reaction tube and placed at -20°C for 2 h. Following protein precipitation, the samples were centrifuged for 2 min at 4°C and 18,000 x g. The supernatant was transferred again to a new micro-reaction tube and diluted 1:1 with acidified H_2O (1% v/v FA). Lastly, the samples were centrifuged again for 5 min at 4°C and 18,000 x g, and the supernatant was transferred to an amber LC vial.

The infant stool samples were prepared following the method of Krausová et al. (2022) with minor modifications. In brief, approximately 80 mg of the wet infant stool was weighed in a micro-reaction tube and dried in a vacuum concentrator (Labconco). H_2O was then added to the samples at 40 µL per 20 mg of dried stool, followed by 160 µL per 20 mg of dried stool of ACN:MeOH (1:1) + 1% v/v FA. The samples were then vortexed and ultrasonicated on ice for 15 min, and subsequently placed at -20°C overnight to allow for protein precipitation. After, they were centrifuged for 10 min at 4°C and 18,000 x g. The supernatant was transferred to a new micro-reaction tube and diluted 10 times with ACN:H_2O (1:1) + 1% v/v FA, and finally the samples were passed through a PTFE filter into an amber LC vial. The quantities of wet stool and dry stool for each sample is listed in Table A.2.

2.4. LC-HRMS instrumentation and parameters

A UHPLC-ESI-QTOF-HRMS system consisting of an Agilent 1290 Infinity II UHPLC and a *Sciex* ZenoTOF 7600 MS was used to analyze the samples. The LC parameters used were previously optimized (Oesterle et al. 2022). In brief, a *Waters* Acquity HSS T3 (2.1 x 100 mm, 1.8 µm) column with a *Waters* Vanguard precolumn at a temperature of 30°C, an autosampler temperature of 4°C, and an injection volume of 5 µL was used. The eluents were composed of H_2O with 0.1% v/v FA as eluent A and ACN with 0.1% v/v FA as eluent B, at a flow rate of 0.6 mL/min, and the gradient given in Table A.3.

The MS was operated in negative polarity, and the source parameters consisted of a CAD gas of 9 arb. unit, a curtain gas of 35 psi, an ion source gas 1 and 2 of 50 psi, a source temperature of

550°C, and a spray voltage of -4500 V. The TOF MS was operated with a scan window of *m/z* 100 to 1000, an accumulation time of 0.25 s, a declustering potential of -70 V, and a collision energy of -10 V. The SWATH parameters used involved scanning for fragments from *m/z* 100 to 1000 with 10 windows, a declustering potential of -70 V, an accumulation time of 0.05 s, and different collision energies for each window. The SWATH windows were optimized for each matrix separately using the total ion chromatogram from the injection of a pooled quality control and are listed in Table A.4 along with the collision energies applied to each window.

2.5. Data processing of the acquired LC-HRMS data

The acquired raw data files were first converted to ABF file format with Reifycs Analysis Base File Converter (v2011-2020), before they were further processed in MS-Dial (v4.9.221218) (Tsugawa et al. 2015). MS-Dial was used for feature pre-processing, e.g., building extracted ion chromatograms, and for feature annotation with a spectral library created by MS-Dial that combines various databases such as GNPS and MassBank (Tsugawa et al. 2015). The parameters used in MS-Dial for the two biological matrices are listed in Table A.4. Features were further annotated with *in silico* fragmentation using MS-Finder (v3.52) (Tsugawa et al. 2016). Prior to feature annotation, the feature lists were filtered by MS^1 matching with the Exposome-Explorer (Neveu et al. 2020) and Phytohub (Giacomoni et al. 2017) databases. R (v4.3.1) (R Core Team 2023) was used for MS^1 matching and feature clean-up with the process blank and pooled quality control (QC) samples.

The identification levels of the feature annotation were given based on the levels previously defined by Schymanski et al. (2014). In brief, features identified with authentic reference standards were labeled as Level 1 and features annotated with spectral libraries as Level 2a. Level 3 was then split in a similar manner as described in Oesterle et al. (2023), and features annotated by in silico fragmentation were labeled as Level 3a, and features putatively annotated by their MS^1 with the two databases as Level 3b. The chemical classes of the features were also determined using the ChemRICH MeSH prediction tool (Barupal and Fiehn 2017), the classes listed in MS-Finder and MS-Dial, and the classes listed in the entries of the two online databases (PhytoHub and Exposome-Explorer).

2.6. Quality control of the LC-HRMS measurements

QC samples were prepared for each biological matrix. The stool QC was prepared by pooling aliquots of each infants' wet stool in respective quantities (Table A.2). The breast milk QC was prepared by combining 10 µL of each sample. The QC samples were processed following the same procedure as the experimental samples. For each biological matrix, the respective QC was used to condition the LC column prior to the acquisition of the experimental samples. Moreover, after column conditioning, three technical replicates of the QCs were measured. The QCs were routinely analyzed after every five experimental samples to continuously check the reliability of the instrument as well as to correct for any signal drifts that may occur within the acquisition batch. In addition, for each QC, a five-point serial dilution series with a constant dilution factor of four was prepared with the respective dilution solvent of the biological matrix's sample preparation procedure. Moreover, matrix-matched calibration curves for the 242 reference standards were created with the QCs (Table A.5). For each biological matrix, along with the samples, a process blank was prepared by leaving the micro-reaction tube empty rather than taking

an aliquot of the biological matrix. For the stool protocol, the process blank was assumed to contain 20 mg of dry weight, therefore 40 µL of H_2O and 160µL of ACN:MeOH (1:1) + 1% v/v FA were added. Both process blanks were measured in triplicates.

Features with an average chromatographic peak area in the process blank measurements greater than one-third of the average in the samples were removed (Kirwan et al. 2014). In addition, features were removed unless they were detected in at least two of the triplicate measurements of the QCs and their peak area in the QC triplicate measurements had a relative standard deviation <30% (Dudzik et al. 2018). Features with an average signal-to-noise value of less than three, as calculated by MS-Dial, were also removed. Finally, the QC dilution series was developed to assess the reliability of extracted features, meaning a feature's signal in the QCs should decrease as the QC is diluted. Therefore, Spearman rank correlation was applied to the QC dilution series for assessing the association between a feature´s signal and the overall dilution factor, whereby features with a correlation value less than 0 were removed.

2.7. 16S rRNA gene amplicon sequencing of the infant stool

Gut microbiome data of the infants' stool samples obtained at month 1, 6 and 12 post-delivery were retrieved from Ayeni et al. (under review). Detailed procedure of DNA extraction, polymerase chain reaction and 16S rRNA gene amplicon sequencing applied to the infants' stool was previously published (Pjevac at al. 2021). Sequencing was done at the Joint Microbiome Facility of the University of Vienna and the Medical University of Vienna. Raw sequence reads of the infants' microbiome are available under the BioProject accession number PRJNA1013123.

2.8. Statistical analysis

Following data pre-processing and annotation, MetaboAnalyst (v5.0) (Pang et al. 2022) was used for statistical analysis. Prior to analysis in MetaboAnalyst, the chromatographic peak areas of the features extracted from the stool samples were normalized by their sample dry weight (Table S2). After importing the data into MetaboAnalyst, the features in each matrix were both normalized by median and applied $\log_{10}$ transformation. Principal Component Analysis (PCA) of the features in the breast milk and stool was applied to investigate clustering of the samples. Then, for the stool samples, one-way Analysis of Variance (ANOVA) was applied to determine significant features over time, while volcano plots (fold change versus T-tests) were used to determine fold changes between the time points. Heatmaps with Ward hierarchical clustering of the annotated features were created to compare the distribution in each matrix.

ChemRICH (v4.0) (Barupal and Fiehn 2017) was used to generate chemical enrichment plots of the stool samples with the statistical analysis results from MetaboAnalyst. Boxplots of the quantification results were made with the *ggplot2* (v3.4.3) (Wickham 2016) and *cowplot* (v1.1.1) (Wilke 2020) packages. Spearman rank correlation between the features in both matrices was calculated using *stats* (v3.6.2) package in R (v4.3.1) (R Core Team 2023). The *pheatmap* package (v1.0.12) (Kolde 2018) was applied to generate heatmaps of the correlation results. Benjamini-Hochberg was applied for multiple testing (Benjamini and Hochberg 1995).

Microbiome data were loaded into R (v4.3.1) and filtered using the *ampvis2* package (v2.8.3) (Andersen et al. 2018). Spearman rank correlations with Benjamini-Hochberg correction between features and the microbiome was done using the *stats* (v3.6.2) package and visualized as heatmaps with the *pheatmap* (v1.0.12) package.

To further visualize the correlations between the features and the microbiome, a network was created in Cytoscape (v3.9.1) (Shannon et al. 2003).

3. Results and discussion

3.1. Suspect screening of breast milk and stool samples

Overall, the utilization of the LC-HRMS methods in negative mode yielded a total of 12192 features acquired in the breast milk samples and 16461 features in the stool samples. After quality control, 4347 and 6905 features remained in breast milk and stool, respectively. A suspect screening workflow optimized for polyphenols but also capable of covering many other xenobiotics and endogenous metabolites (Oesterle et al. 2023) was then used to extract features of interest in the biological matrices. This workflow involved a suspect list from two databases: PhytoHub, an online database containing 2268 phytochemicals, mainly polyphenols (Giacomoni et al. 2017), and Exposome-Explorer, a database containing 1262 chemicals known to be biomarkers for exposure to environmental and lifestyle factors such as diet, pollutants, or contaminants (Neveu et al. 2020). The application of this workflow resulted in a total of 542 matched features in the breast milk and 864 in the stool samples. From the 542 breast milk features, 34 were identified as Level 1, 51 annotated as Level 2a, 416 as Level 3a, and 41 as Level 3b. While from the 864 features in the stool, 68 were identified as Level 1, 78 annotated as Level 2a, 652 as Level 3a, and 66 putatively annotated as Level 3b. The annotated features are listed in Table A.6 for breast milk and Table A.7 for stool.

In the breast milk, many of the features detected were fatty acids, peptides, saccharides, or amino acids (Figure 1a). This was expected as breast milk is rich in nutrients, especially lipids, proteins, and carbohydrates, required for infant growth and development (Boudry et al. 2021). The breast milk also contained polyphenols such as flavonols, flavones, and isoflavones. Several studies have shown that polyphenols are abundant in human breast milk and they could be beneficial to the infants (Lu et al. 2021, Song et al. 2013, Carregosa et al. 2023). In contrast, toxic xenobiotics, such as mycotoxins, were detected at a low prevalence in breast milk. This class of xenobiotics are typically found at low concentrations in breast milk (Adejumo et al. 2013, Braun et al. 2020, Braun et al. 2022). Moreover, the heatmap of the features detected in the breast milk (Figure 1b) depicts that the features cluster in two main groups, though there is no clear separation of the types of features in each of the two groups.

More diverse chemical classes, predominantly of plant-based origin, were present in the stool (Figure 2a). One of the main chemical classes detected were polyphenols, which was expected as the LC-MS method was originally optimized for these analytes (Oesterle et al. 2022, Oesterle et al. 2023). A wide range of different polyphenol classes were detected, including 55 flavonols, 32 flavones, 29 isoflavones, four chalcones, and 27 hydroxycinnamic acids. Besides polyphenols, several potentially toxic xenobiotics were also detected. For instance, one feature (m/z 254.952, retention time: 5.8 min, Level 3b) was putatively annotated as a polychlorinated biphenyl-16. It was previously reported that prenatal exposure to PCBs were associated with adverse effects on birth weight (Govarts et al. 2020). In addition, naphthalene epoxide (Level 3a), a metabolite of naphthalene, was detected. Naphthalene exposure has been associated with various negative health implications, especially for respiratory health (Cakmak et al. 2014). Several fatty acids were also detected in the stool samples, including arachidonic acid (Level 2a)

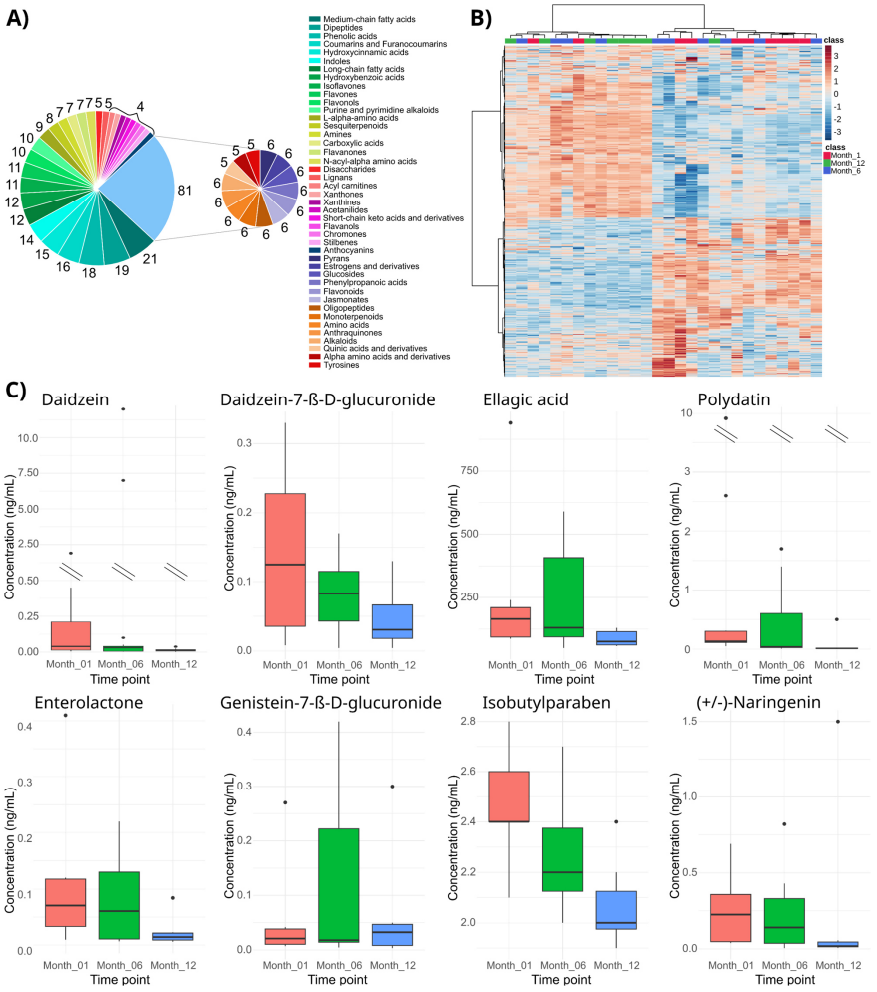


Figure 1. A) Pie chart representing the major chemical classes of the features identified and annotated in the breast milk samples obtained from Nigerian mothers at three different time points. **B)** Heatmap with Ward clustering displaying the features detected in all the breast milk samples. **C)** Box plots showing quantitative differences (Table A.8) over time for selected analytes identified at Level 1 that represent both potentially beneficial and toxic xenobiotics in the breast milk samples. The boxplots of all the identified features are shown in Figure B.2.

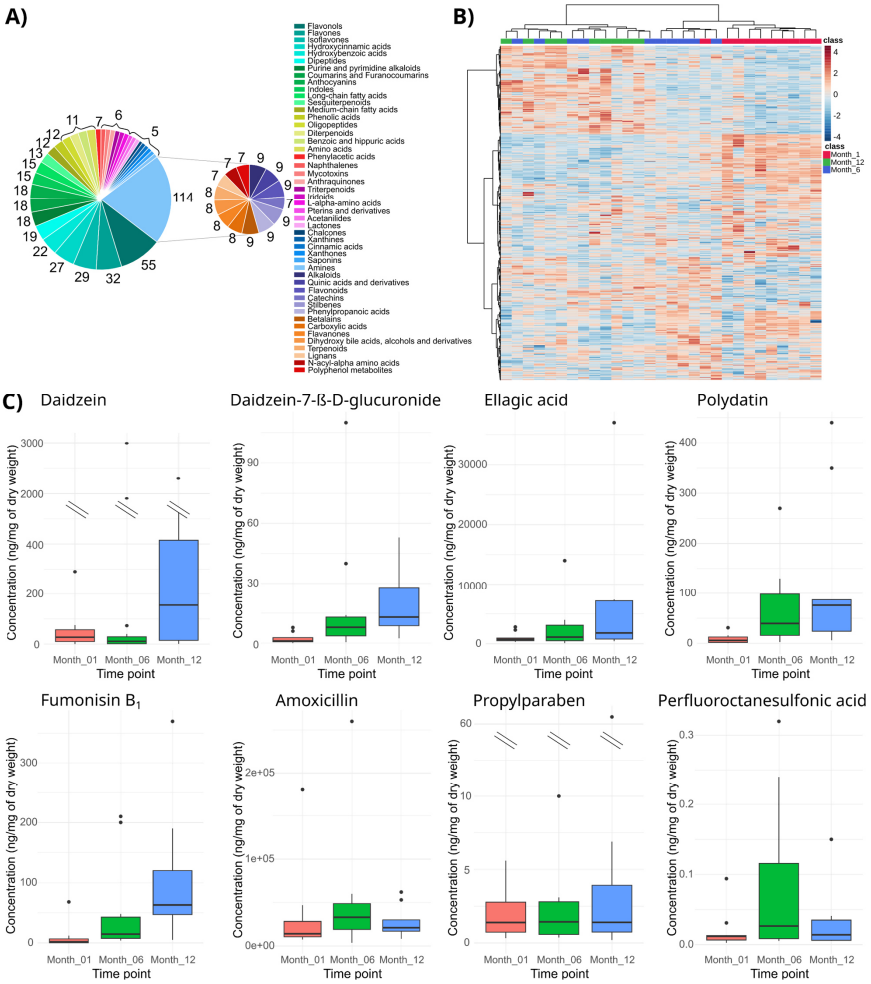


Figure 2. A) Pie chart representing the major chemical classes of the features identified and annotated in the stool samples from Nigerian infants at three different time points. **B)** Heatmap with Ward clustering displaying the features detected with the suspect screening workflow in the stool samples. **C)** Box plots of the semi-quantification (Table A.8) at the three distinct infant ages of several identified features (Level 1) that represent both potentially beneficial and toxic xenobiotics in the stool samples. These plots show that complementary foods increase exposure to various xenobiotics, and breast feeding keeps exposure levels low, despite potential lactational transfer. The box plots of all the identified features are shown in Figure B.3.

and eicosapentanaenoic acid (Level 3a). The levels of these two fatty acids were previously reported to be higher in stool of infants that were breastfed compared to infants fed that were formula-fed (Sillner et al. 2021).

Similar to the breast milk heatmap, the features in the stool (Figure 2b) also form two clusters, with one cluster from the samples during the time of breastfeeding (month 1) and the other for those acquired after the introduction of complementary foods (months 6 and 12). In addition, many of the features detected in the stool samples were conjugated with sugar moieties. This could be attributed to phase II metabolism, as it was previously reported that metabolites of genistein included hexose and pentose conjugates in cells (Flasch et al. 2022a); or to their low intestinal absorption, as, for example, seen with the bioavailability of large polyphenols like proanthocyanidins (Scalbert et al. 2002). Moreover, the bioavailability in infants may be further reduced as metabolic pathways, e.g. phase I and II metabolism (Lu and Rosenbaum 2014), or their microbiome (Wopereis et al. 2014) are not yet fully developed.

The features from each biological matrix were then compared with each other with a retention time deviation of 0.1 min and a mass error of 20 ppm. This resulted in a total of 149 features that were detected in both matrices, of which 32, 24, 91, and four were annotated as Levels 1, 2a, 3a, and 3b, respectively. One of the most prevalent chemical classes in both biological matrices was phenolic acids. Phenolic acids are a polyphenol class comprising of many human and microbial metabolites, including metabolites from other larger polyphenols such as anthocyanins (de Ferrars et al. 2014). Besides phenolic acids, several other biotransformation products were detected in both matrices including caffeic acid-3-ß-D-glucuronide, daidzein-7-ß-D-glucuronide (Figure 3b), and genistein-7-ß-D-glucuronide. This observation suggests that the chemicals were either transferred directly from the breast milk to the infant and the corresponding stool samples or that the infant's metabolism conjugated the parent compounds in the colon or liver by UDP-glucuronosyltransferases (Rowland et al. 2013). The overlap in features between the two matrices showed that xenobiotics, for example polyphenols, are transferred from the mother to the infant. Though scant data exist, the lactational transfer of various xenobiotics have been previously described in humans. This includes persistent organic pollutants (Haddad et al. 2015), ellagitannins and their metabolites (Henning et al. 2022), and pharmaceutical drugs and environmental pollutants (Dubbelboer et al. 2023). In addition, the heatmap (Figure B.1) of all the features detected in both matrices showed that the two matrices are highly different, as only 149 from the 1259 total features are present in both matrices. This further underscores the differences previously seen in the types of chemical classes found in each matrix (Figure 1a and 2a).

3.2. Semi-quantification of Level 1 identified features in breast milk and stool

A total of 242 authentic reference standards representing different xenobiotics were utilized for identification purposes. These standards included air pollutants, disinfection by-products, endogenous estrogens, food processing by-products, industrial side-products, pesticides, mycotoxins, perfluorinated alkylated substances, personal care product / pharmaceuticals, phytotoxins, plasticizer / plastic components, antibiotics, and polyphenols (Table A.1). In the breast milk, 34 features were identified that consisted of one antibiotic, one estrogen, one mycotoxin, one personal care product / pharmaceutical, and 30 polyphenols. While in the stool samples, 68 features were identified that consisted of five antibiotics, three mycotoxins,

one perfluorinated alkylated substance, one pesticide, one plasticizer / plastic component, three personal care product / pharmaceuticals, and 54 polyphenols. Polyphenols were more abundantly detected in the samples as they are typically found in higher concentrations than other xenobiotics in human biofluids (Rappaport et al. 2014, Achaintre et al. 2018) and also the applied workflow was originally optimized towards the balanced and decent performance in detection of polyphenols.

The identified features (Level 1) in the breast milk (34) and stool (68) were semi-quantified (Table A.8) with calibration curves created from the reference standards (Table A.5). To approximate signal suppression and enhancement effects, the calibration curves were matrix-matched with the pooled QC of the respective biological matrix. However, standard addition was required for quantification since all the Level 1 identified features already had a chromatographic peak present in the pooled QCs. Each identified feature in breast milk and stool samples at the three different time points was represented in (Figure 1 and B.2) and (Figure 2 and B.3), respectively.

Overall, there was no clear pattern in the concentration of features detected in the breast milk at the three time points, but a decrease over time was observed for daidzein, daidzein 7-ß-D-glucoronide, estradiol 17-glucuronide, isobutylparaben and naringenin (Figure 1c). For instance, daidzein was detected at concentration ranges between 0.0035 to 12 ng/mL (Table A.8, Figure 1c). The concentration of daidzein was similar to the reported concentrations from a longitudinal study of an Austrian mother (Jamnik et al. 2022). Moreover, compared to the Austrian study, higher concentrations were recorded for of other polyphenols including enterodiol (range: 0.024 - 23 ng/mL) and enterolactone (range: 0.0079 to 0.41 ng/mL) (Table A.8). In addition, the breast milk samples from this cohort were previously analyzed with a targeted LC-MS/MS workflow focusing on polyphenols (Berger et al. submitted). Despite the targeted assay detecting higher quantities, similar analytes and patterns were observed with both assays. Besides polyphenols, only two potentially toxic xenobiotics, alternariol monomethyl ether and isobutylparaben, were identified at Level 1 in the breast milk samples. Alternariol monomethyl ether was previously reported as a prevalent mycotoxin in breast milk from a Nigerian cohort (Braun et al. 2022, Ezekiel et al. 2022). While there is scarce reported data on parabens in breast milk from Nigeria, these xenobiotics have been reported in breast milk from other regions (Fisher et al. 2017, Kim et al. 2023). Lastly, in the breast milk samples, an antibiotic, azithromycin, was detected in only one sample at a concentration of 1100 ng/mL (Table A.8). Interestingly, the participant who donated this sample used azithromycin for medical treatment purposes (unpublished data). Moreover, azithromycin was detected at a high concentration (140000 ng/mg of dry weight) in the corresponding infant stool sample (Table A.8). It should be noted that the concentrations are semi-quantitative and would therefore require absolute quantification with targeted assays. However, this observation suggests that azithromycin can be transferred from mother's breast milk to infant (Roca et al. 2015).

In contrast, patterns could be observed in the features identified and subsequently quantified in the stool samples. Many of the features belonging to catechins, chalcones, flavanones, flavones, flavonols, isoflavones, phenolic acids, stilbenes, and mycotoxins increased in concentration from month 1 to 12. Moreover, the concentrations of the identified features in stool changed when complementary foods were introduced to the infants at 6 months showing a substantially higher quantity at month 6 compared to month 1 (Figure 3). This observation suggests that complementary foods have a considerable

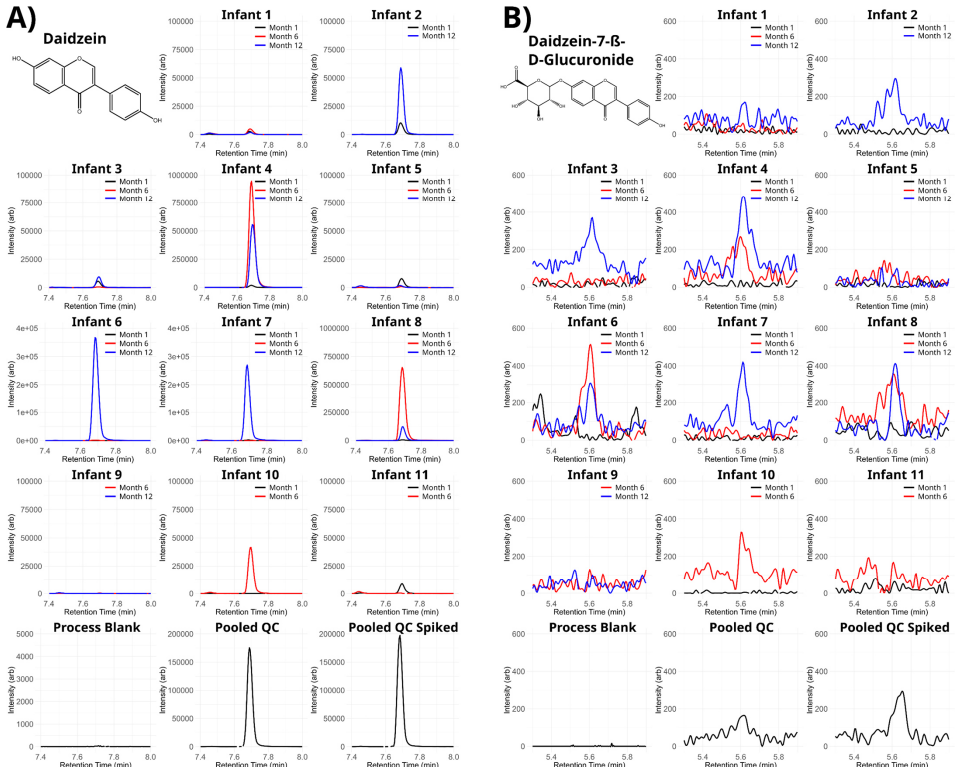


Figure 3. Extracted ion chromatograms of daidzein (*m/z* 253.050, retention time: 7.7 min, Level 1) (**A**) and daidzein-7-ß-d-glucuronide (*m/z* 429.079, retention time: 5.6 min, Level 1) (**B**) in all the infants at the three distinct time points. In addition, the process blank, pooled QC, and pooled QC spiked with the authentic reference standards (14 ng/mL for daidzein and 1.1 ng/mL for daidzein-7-ß-D-glucuronide) are depicted. These chromatograms illustrate that xenobiotics of plant origin significantly increase in concentrations as complementary foods are introduced in the infants' diets.

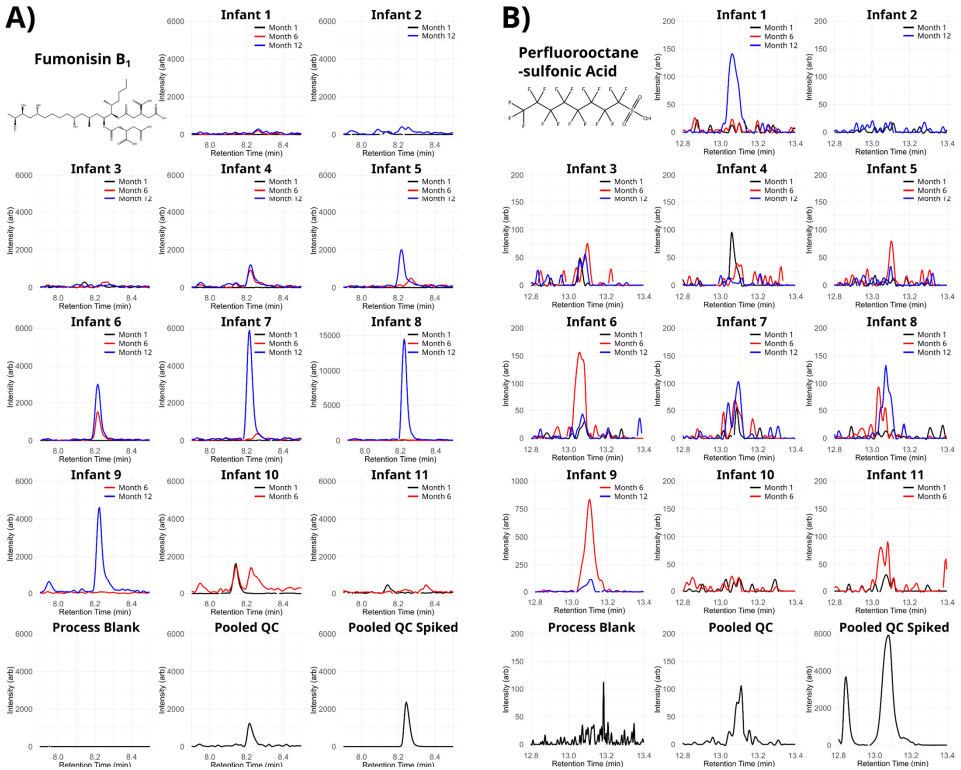


Figure 4. Extracted ion chromatograms of fumonisin B_1 (*m/z* 720.379, retention time: 8.2 min, Level 1) (**A**) and perfluorooctanesulfonic acid (*m/z* 498.929, retention time: 13.1 min, Level 1) (**B**) in all the infants at the three distinct time points. In addition, the process blank, pooled QC, and pooled QC spiked with the authentic reference standards (6.3 ng/mL for fumonisin B1 and 0.38 ng/mL for perfluorooctanesulfonic acid) are depicted. These chromatograms illustrate that exposure levels of adverse xenobiotics during exclusive breastfeeding were typically low and an increase was observed as complementary foods were introduced.

influence on the features in stool. Moreover, the presence of various features when the infants were one-month old showed that bioactive compounds, e.g. polyphenols, are transferred from the breast milk to the infants. For instance, one analyte that is most likely transferred via breast milk due to its stable concentrations over time (Figure B.3) is urolithin A, a microbial metabolite that showed properties beneficial to health (D'Amico et al. 2021).

Some of the features detected in the infant stool were toxic xenobiotics, such as the mycotoxins fumonisin B_1 (Figure 4a) and nivalenol. Fumonisin B_1 has been previously associated with neural tube defects and the etiology of esophageal cancer in humans (Marasas et al. 2004, Missmer et al. 2006). Moreover, this mycotoxin was previously found in stool of the same cohort using a targeted LC-MS/MS-based assay (Krausová et al. 2022, Ayeni et al. under review), and the longitudinal pattern of occurrence was comparable (Ayeni et al., under review). These observations add an extra layer of confidence to the results herein and highlight that targeted and non-targeted LC-MS/MS approaches can be complementary (Flasch et al. 2023). The detection of antibiotics such as amoxicillin, azithromycin, oxytetracycline and trimethoprim in the stool (Figure B.3) can be attributed to the infants taking these antibiotics for medical treatment, or transferred via breastmilk. Florfenicol, a veterinary drug, decreased in concentration in the stool from month 1 to 12, suggesting exposure to the drug was from breast milk. Other xenobiotics exclusively detected in the stool were personal care products (e.g., methylparaben, propylparaben and triclosan), pesticides (e.g., p-nitrophenol), and plasticizer (e.g., mono-2-ethylhexyl phthalate). In addition, perfluorooctanesulfonic acid, a toxic perfluorinated alkylated substance, was detected (Figure 2c and 4b). For all infants, except one, the concentrations were lower during exclusive breastfeeding (month 1 to months 6) compared to when complementary foods were introduced (months 6 to 12). Moreover, perfluorooctanesulfonic acid was detected in Ogun river (>10 ng/L) (Ololade et al. 2018), a river in the region were the samples were collected. While many of the identified toxicants came from dietary sources, others may be attributed to other routes of exposure such as air, water, or cosmetics.

3.3. Correlating xenobiotics within and between each biological matrix

As xenobiotics can exhibit synergistic or antagonistic effects, correlations of the features in the breast milk and stool samples were explored. Hence, Spearman rank correlation (p. adj < 0.05) was applied among the 34 identified features in the breast milk (Table A.9, Figure 5a), and the 68 in the stool (Table A.10, Figure 5b). In the breast milk samples, many of the polyphenols showed moderate to strong correlations (ρ = 0.37 - 0.88) (Table A.9). For example, the two microbial metabolites of matairesinol, enterolactone and enterodiol, strongly correlated with each other as expected (ρ = 0.67). Additionally, polyphenols also correlated with other xenobiotics, e.g. ellagic acid and dihydrocaffeic acid showed moderate correlation with alternariol monomethyl ether (ρ = 0.38 and 0.38, respectively). Estradiol-17-glucuronide also showed moderate to strong correlations with numerous polyphenols, including enterolactone (ρ = 0.41), 4-hydroxybenzoic acid (ρ = 0.58), and salicylic acid (ρ = 0.70). It was previously observed in mice that enterolactone activates estrogen-sensitive reporter gene expression (Penttinen et al. 2007). Hippuric acid showed negative correlation with other phenolic acids, such as benzoic acid (ρ = -0.73), which is most likely due to hippuric acid being a metabolite of benzoic acid (Lees et al. 2013).

Figure 5. A) Spearman rank correlation (p. adj < 0.05) heatmap between the 34 identified features (Level 1) in the breast milk samples. **B)** Spearman rank correlation (p. adj < 0.05) heatmap between the 67 identified features (Level 1) in the stool samples.

All the identified features in the stool, except for one combination, positively correlated with one another, especially features within the same chemical class. For example, two lignan metabolites, enterodiol and enterolactone, moderately correlated with each other (ρ = 0.52); and two isoflavone metabolites, daidzein 7-ß-D-glucuronide and genistein 7-ß-D-glucuronide, showed strong correlation (ρ = 0.82). Additionally, phloretin strongly correlated with diosmetin (ρ = 0.87), kaempferol (ρ = 0.81), and quercetin (ρ = 0.81). Besides polyphenols, the potentially toxic xenobiotics: nivalenol, amoxicillin, mono-2-ethylhexyl-phthalate, alternariol monomethyl ether, trimethoprim, p-nitrophenol, and florfenicol, showed moderate to strong correlations with one another (ρ = 0.40 - 0.90) (Table A.10). Of the xenobiotics, the highest correlation was recorded between mono-2-ethylhexyl phthalate and p-nitrophenol (ρ = 0.90). Apart from individual effects, mixture toxicity of these xenobiotics can lead to more severe health consequences especially during early-life (Hamid et al. 2021, Krausová et al. 2023). The only negative correlation was between alternariol monomethyl ether, an *Alternaria* mycotoxin, and genistein-7-sulfate, a metabolite of genistein (ρ = -0.41). Previously, it was observed *in vitro* that genistein had antagonistic effects on the genotoxicity of alternariol, another *Alternaria* mycotoxin (Aichinger et al. 2017).

Spearman rank correlation was further applied to explore possible correlations between the identified features in breast milk and infant stool. Only samples obtained at month 1 were selected because the infants predominantly consumed breast milk at this time point. No significant correlations were found (p. adj > 0.05) (Table A.11 and Figure B.4) which may be attributed to the low sample size (n = 10). However, with the raw p-value ($p < 0.05$), many of the polyphenols in the breast milk positively correlated with polyphenols in the stool, especially between phenolic acids. For example, benzoic acid in breast milk strongly correlated with 4-hydroxybenzoic acid (ρ = 0.79) in the infant stool. Alternariol monomethyl ether in the breast milk did not correlate with alternariol monomethyl ether in the stool, but showed strong negative correlation with nivalenol (ρ = -0.72). Interestingly, alternariol monomethyl ether in the breast milk had strong negative correlations with several polyphenols in the stool. For example, alternariol monomethyl ether strongly negatively correlated with (-)-epicatechin (ρ = -0.98) and urolithin A (ρ = -0.69). Urolithin C, also a metabolite of ellagitannins and structurally similar to urolithin A, was shown to influence the metabolism of alternariol *in vitro* (Crudo et al. 2021). Overall, further investigations are necessary to access the toxicological potentials of many of the observed correlations in *in vitro* models.

3.4. The impact of age and diet on the infants' chemical exposome

The influence of age on the features detected in breast milk and stool was investigated by PCA plots (Figures 6a and b). The breast milk samples showed no clear grouping at the three different time points (Figure 6a). On the contrary, the infant stool samples showed clear grouping, though there was some overlap between the clusters at months six and twelve (Figure 6b). The results of both PCAs are in agreement with the hierarchical clustering in the feature heatmaps (Figure 1b and 2b).

One-way ANOVA was applied on the features detected in the infant stool, yielding 325 features that showed significance across all time points (Table A.7). To further investigate the changes between each time point, volcano plots (fold change versus t-test p values) were generated (Figure 6c-e), with the significant features (fold change > 2, $p < 0.05$) of each

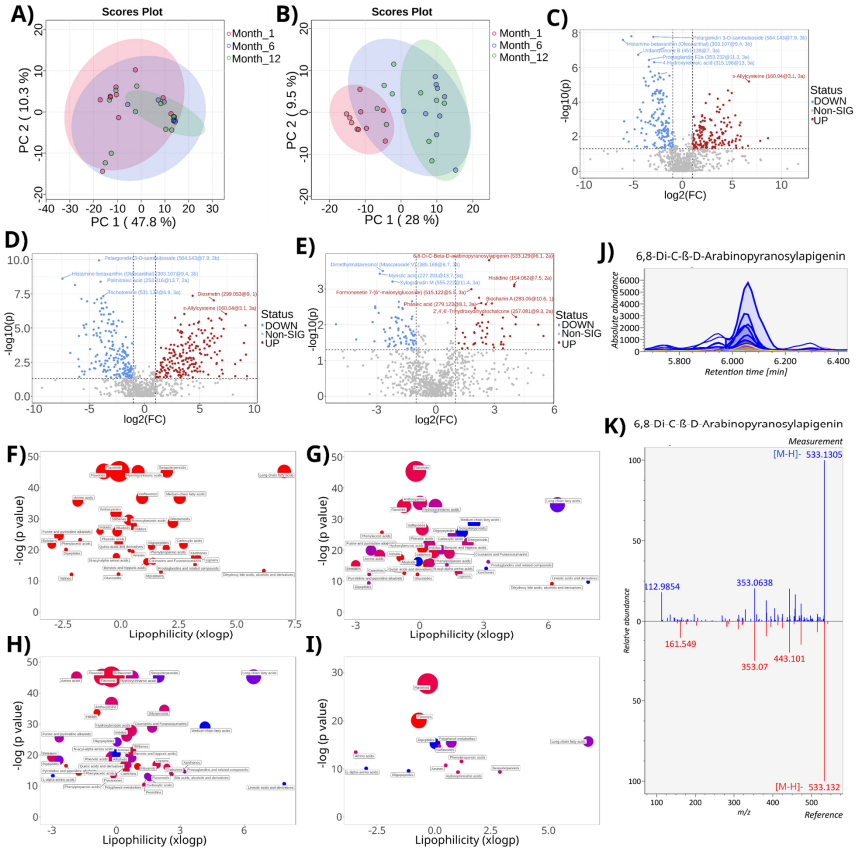


Figure 6. A) Principal component analysis (PCA) plot of the breast milk samples demonstrates that the milk composition did not cluster, indicating no major overall changes of the breast milk composition over time. This makes the study design suitable to explore the impact of solid food intake introduced during early life. **B)** PCA plot of the infant stool samples shows grouping by age with a clear separation of the first month after delivery. **C)** Volcano plot ($p < 0.05$, fold change > 2) of the features between the samples from infants at one month old and six months old. **D)** Volcano plot ($p < 0.05$, fold change > 2) of the features between the samples from infants at one month old and twelve months old. **E)** Volcano plot ($p < 0.05$, fold change > 2) of the features between the samples from infants at six months old and twelve months old. **F)** Chemical enrichment (ChemRICH) plot of the significant features from ANOVA **G)** ChemRICH plot of the significant features from C) with the chemical classes in blue are down-regulated and those in red are up-regulated. **H)** ChemRICH plot of the significant features from D). **I)** ChemRICH plot of the significant features from E). **J)** The overlapping extracted ion chromatograms from 6,8-Di-C-ß-D-arabinopyranosylapigenin which showed high fold change and significance in E). **K)** Mirror plot of the experimental MS^2 and the MS^2 from the spectral library match of the feature from J).

volcano plots listed in Table A.7. A total of 303 features were significant between months one and six, with 151 showing an increase and 153 a decrease. Then from six to twelve months, 118 were significant of which 48 increased and 70 decreased. Finally, from one to twelve months, there were 388 significant features, with 197 increasing and 191 decreasing. A greater number of features were significant between the ages of one to six months and one to twelve months than from six to twelve months, showing the impact of complementary foods. Overall, significant features from ANOVA were similar to the majority of the significant features from the volcano plots, especially between months one and twelve.

ChemRICH plots generated from the volcano plot and ANOVA results (Figure 6f-i) depicted clusters of chemical classes that were either up- or down-regulated. Months one to six and months one to twelve showed up-regulation in chemical classes from plant-origin such as polyphenols, e.g. flavonols or flavones (Figure 6f-h). Additionally, a down-regulation of chemical classes, such as amino acids and fatty acids, were observed. Fatty acids that were down-regulated over time included arachidonic and eicosapentanaenoic acid, which is most likely related to the infants' reduced consumption of breast milk (Sillner et al. 2021). Conversely, no clear pattern with the chemical classes were observed for months six to twelve (Figure 6i), which may be due to the large variance of phytochemicals in foods. These variances may be based on several factors, such as duration of food storage and/or cooking practices (Arfaoui 2021).

3.5. Correlations between significant features and the developing infant gut microbiome

Spearman rank correlation was applied to explore correlations between the features that showed significance through ANOVA and the stool microbiome. It should be noted that samples from all the time points were included, thus other confounding factors, such as age and dietary changes, may influence the correlations. Besides creating a heatmap from the correlation matrix (Figure 7a), a network was generated to better visualize the correlations (Figure 7b). Several significant correlations between some taxa and features in the stool (p. adj < 0.05) were found (Table A.12, Figure 7). Specifically, 58 microbe-feature pairs had positive correlations, while 12 microbe-feature pairs had negative correlations. The correlated features consisted mainly of phytochemicals, such as polyphenols and terpenoids. For example, the flavone tricetin showed a strong positive correlation with *Blautia* ($\rho = 0.69$). *Blautia* has been shown to biotransform flavonoids e.g., polymethoxyflavones into several demethylated flavones (Kim et al. 2014, Liu et al. 2021). Flavonols, e.g. kaempferol-4'-methylether and quercetin-3-arabinoside, showed strong correlations with *Clostridium sensu stricto* 1 ($\rho = 0.69$ and 0.67, respectively) (Figure 7). Members of *Clostridium* are known to degrade flavonols (Zhang et al. 2014, Steed et al. 2017). Fumonisin B_1 showed strong correlation with *Streptococcus* ($\rho = 0.75$). Previously, it was observed that *Streptococcus* can bind to fumonisin B_1 (Niderkorn et al. 2006). Several flavonoids, including kaempferol and biochanin A, strongly correlated with *Romboutsia* ($\rho = 0.70$ and 0.68, respectively). An extract containing flavonoids was previously shown to increase relative abundance of *Romboutsia* in mice (Wang et al. 2022).

Vulgaxanthin I showed a strong negative correlation with *Escherichia-Shigella* ($\rho = -0.70$). Extracts containing vulgaxanthin I exhibited antibacterial activities against Gram-negative bacteria including *Escherichia* (Vulić et al. 2013). The alkaloid, 1,3-dimethylurate, had strong negative correlations with *Clostridioides*

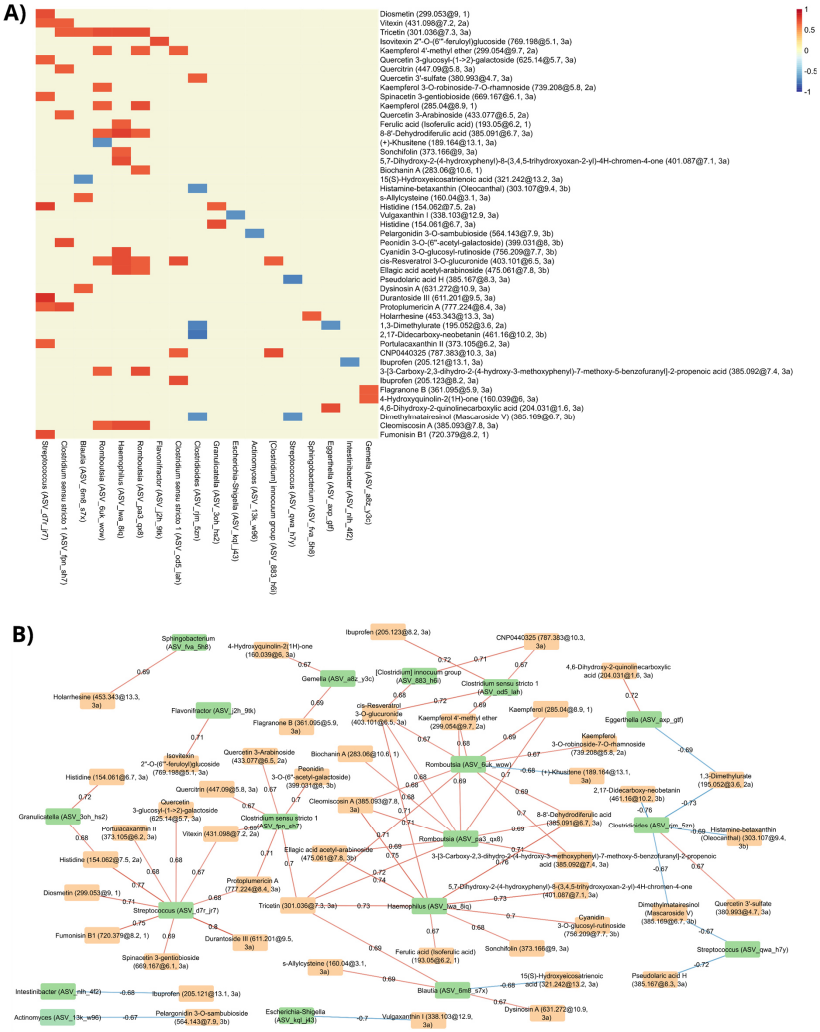


Figure 7. A) Spearman rank correlation (p.adj < 0.05) heatmap between the microbes detected by 16S rRNA gene amplicon sequencing (with their species and ASV identifier given) and the features that showed significance by ANOVA (with their *m/z*, retention time, and identification level given). **B)** A network of the Spearman rank correlation results with microbes (with their genus-level classification and ASV identifier given) colored in green and features (with their *m/z*, retention time, and identification level given) in orange.

(ρ = -0.73) and *Eggerthella* (ρ = -0.69). Similarly, the terpenoids (+)-khusitene negatively correlated with *Romboutsia* (ρ = -0.68) and pseudolaric acid H with *Streptococcus* (ρ = -0.72). The negative correlations observed for alkaloids and terpenoids may be attributed to their antibacterial properties (Yang et al. 2020, Cushnie et al. 2014). While the causal link between the features derived from xenobiotics and members of the infant gut microbiome remains to be elucidated, the results highlight the need to explore xenobiotic-microbiome interactions in less complex *in-vitro* models and its consequent health implications.

3.6. Limitations

Several limitations that are essential to consider have become apparent in the applied analytical workflow and the overall study design. Many of the analytes investigated, especially polyphenols, have a wide range of isomers, making annotation of the features a complex task. This is reflected in the suspect screening results where many features (Table A.6 and A.7) have the same annotation though they were distinct analytes. Additionally, DIA MS^2 spectra typically have more noise than DDA MS^2 spectra (Guo and Huan 2020), further complicating annotation. The MS^2 spectra with increased noise have a strong influence on *in silico* fragmentation, hence a high number of features annotated at level 3a is reported. The uncertainty in feature annotation also impacts the biological relevance of the statistical and correlation analyses. In addition, the number of participants in the study was small (n = 11) and as such, the correlation results must be interpreted with caution. Moreover, due to the low sample size, all time points were included in the correlations between the infant stool and gut microbiome. Therefore, the correlations may reflect indirect drivers such as the dynamic development of the microbiome in early life or changes in diet with age. A larger sample size and comparison among cohorts from other geographical settings would be needed for a more comprehensive analysis. Data is sparse on longitudinal metabolomic/exposomic profiles of healthy Nigerian mother-infant pairs anchored on high-resolution mass spectrometric analysis of breast milk and matching stool samples. Thus, the data herein gives an important snapshot of the chemical exposome in biofluids of a selected population that has been considered to be at a relatively high exposure levels to many beneficial and adverse xenobiotics.

4. Conclusion

The influence of natural and/or synthetic chemicals in early life is known to have a considerable impact on the development of humans. This study provides a comprehensive overview of potentially beneficial as well as potentially toxic xenobiotics in breast milk and stool from Nigerian mother-infant pairs. Several xenobiotics detected in the breast milk were also present in the corresponding stool samples, although the stool samples contained, as expected, a higher number of different xenobiotics. Correlations were observed between xenobiotics and certain members of the gut microbiome of the infants. Exposure to xenobiotics and their impact on the health of the infants significantly increased with the introduction of complementary foods. However, the toxicological relevance of these correlations needs to be further explored in larger cohorts and validated in *in vitro* models. Despite the limited sample size, the longitudinally study design and the advanced exposomic/metabolomic workflow applied allowed for the detailed assessment of the chemical exposome in breast milk and stool. The next steps should be the application of such workflows in larger cohorts and in different populations, especially in long-term studies, to better characterize the influence that exposure to

various chemicals and their impact on health and microbiome development have.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgment

The authors would like to thank the mothers and their infants for providing the samples. They would also like to thank and acknowledge all the members of their working groups, the Warth and Rompel labs, for their help, support, and feedback. The authors would like to express their gratitude to the Mass Spectrometry Center of the Faculty of Chemistry at the University of Vienna for technical support during the measurements, and to the Joint Microbiome Facility of the University of Vienna and the Medical University of Vienna. This work was supported by the University of Vienna through the Exposome Austria Research Infrastructure (B.W.), the Austrian Federal Ministry of Education, Science and Research (project DigiOmics4AT, B.W.), the Austrian Federal Ministry for Climate Protection, Environment, Energy, Mobility, Innovation and technology (BMK, B.W.), and the Austrian Science Fund (FWF) grant P32326 (A.R.).

Data availability

LC-MS raw data files have been submitted to the MetaboLights data repository (MTBLS8792). The 16S rRNA gene amplicon data is available on the BioProject accession number PRJNA1013123.

Appendix

Supplementary file A (Excel) is given that contains all of the tables mentioned in the text, e.g. the suspect screening results for breast milk and stool.

Supplementary file B (PDF) is given that contains various figures, including boxplots of the semi-quantification results and heatmaps from the correlation analysis.

CRediT author contributions

Ian Oesterle: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Kolawole I. Ayeni:** Conceptualization, Software, Formal analysis, Writing - Review & Editing. **Chibundu N. Ezekiel:** Writing - Review & Editing, Resources. **David Berry:** Writing - Review & Editing. **Annette Rompel:** Writing - Review & Editing, Supervision, Funding acquisition, Resources. **Benedikt Warth:** Conceptualization, Writing - Review & Editing, Supervision, Funding acquisition, Resources.

References

Achaintre, D., Gicquiau, A., Li, L., Rinaldi, S., Scalbert, A., 2018. Quantification of 38 dietary polyphenols in plasma by differential isotope labelling and liquid chromatography electrospray ionization tandem mass spectrometry. J. Chromatogr. A 1558, 50–58. https://doi.org/10.1016/j.chroma.2018.05.017

Adejumo, O., Atanda, O., Raiola, A., Somorin, Y., Bandyopadhyay, R., Ritieni, A., 2013. Correlation between aflatoxin M1 content of breast milk, dietary exposure to aflatoxin B1 and socioeconomic status of lactating mothers in Ogun State, Nigeria. Food Chem. Toxicol. 56, 171–177. https://doi.org/10.1016/j.fct.2013.02.027

Aichinger, G., Beisl, J., Marko, D., 2017. Genistein and delphinidin antagonize the genotoxic effects of the mycotoxin alternariol in human colon carcinoma cells. Mol. Nutr. Food. Res. 61, 1600462. https://doi.org/10.1002/mnfr.201600462

Andersen, K.S., Kirkegaard, R.H., Karst, S.M., Albertsen, M., 2018. ampvis2: an R package to analyse and visualise 16S rRNA amplicon data (preprint). bioRxiv https://doi.org/10.1101/299537

Arfaoui, L., 2021. Dietary Plant Polyphenols: Effects of Food Processing on Their Content and Bioavailability. Molecules 26, 2959. https://doi.org/10.3390/molecules26102959

Ayeni, K.I., Berry, D., Wisgrill, L., Warth, B., Ezekiel, C.N., 2022. Early-life chemical exposome and gut microbiome development: African research perspectives within a global environmental health context. Trends Microbiol. 30, 1084–1100. https://doi.org/10.1016/j.tim.2022.05.008

Ayeni, K.I., Seki, D., Pjevac, P., Hausmann, B., Krausová, M., Braun, D., Wisgrill, L., Berry, D., Warth, B., & Ezekiel, C.N. Biomonitoring associates dietary mycotoxin exposure with infant gut microbiome composition in a high-risk mycotoxin region. Submitted

Barupal, D.K., Fiehn, O., 2017. Chemical Similarity Enrichment Analysis (ChemRICH) as alternative to biochemical pathway mapping for metabolomic datasets. Sci. Rep. 7, 14567. https://doi.org/10.1038/s41598-017-15231-w

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Series B Stat. Methodol. 57, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berger, S., Oesterle, I., Ayeni, K.I., Ezekiel, C. N., Rompel, A., Warth, B., 2023. Polyphenol exposure of mothers and infants assessed by LC-MS/MS-based biomonitoring in breast milk. Submitted

Bocato, M.Z., Ximenez, J.P.B., Hoffmann, C., Barbosa, F., 2019. An overview of the current progress, challenges, and prospects of human biomonitoring and exposome studies. J. Toxicol. Environ. Health B Crit. Rev. 22, 131–156. https://doi.org/10.1080/10937404.2019.1661588

Bonner, R., Hopfgartner, G., 2019. SWATH data independent acquisition mass spectrometry for metabolomics. TrAC Trends Anal. Chem. 120, 115278. https://doi.org/10.1016/j.trac.2018.10.014

Boudry, G., Charton, E., Le Huerou-Luron, I., Ferret-Bernard, S., Le Gall, S., Even, S., Blat, S., 2021. The Relationship Between Breast Milk Components and the Infant Gut Microbiota. Front. Nutr. 8, 629740. https://doi.org/10.3389/fnut.2021.629740

Braun, D., Abia, W.A., Šarkanj, B., Sulyok, M., Waldhoer, T., Erber, A.C., Krska, R., Turner, P.C., Marko, D., Ezekiel, C.N., Warth, B., 2022. Mycotoxin-mixture assessment in mother-infant pairs in Nigeria: From mothers' meal to infants' urine. Chemosphere 287, 132226. https://doi.org/10.1016/j.chemosphere.2021.132226

Braun, D., Schernhammer, E., Marko, D., Warth, B., 2020. Longitudinal assessment of mycotoxin co-exposures in exclusively breastfed infants. Environ. Int. 142, 105845. https://doi.org/10.1016/j.envint.2020.105845

Braun, J.M., 2017. Early Life Exposure to Endocrine Disrupting Chemicals and Childhood Obesity and Neurodevelopment. Nat. Rev. Endocrinol. 13, 161–173. https://doi.org/10.1038/nrendo.2016.186

Brink, L.R., Mercer, K.E., Piccolo, B.D., Chintapalli, S.V., Elolimy, A., Bowlin, A.K., Matazel, K.S., Pack, L., Adams, S.H., Shankar, K., Badger, T.M., Andres, A., Yeruva, L., 2020. Neonatal diet alters fecal microbiota and metabolome profiles at different ages in infants fed breast milk or formula. Am. J. Clin. Nutr. 111, 1190–1202. https://doi.org/10.1093/ajcn/nqaa076

Cakmak, S., Dales, R.E., Liu, L., Kauri, L.M., Lemieux, C.L., Hebbern, C., Zhu, J., 2014. Residential exposure to volatile organic compounds and lung function: Results from a population-based cross-sectional survey. Environ. Pollut. 194, 145–151. https://doi.org/10.1016/j.envpol.2014.07.020

Carregosa, D., Silva, I.P., Teixeira, C., Baltazar, M., García-Villalba, R., Vieira, F.S., Marçal, M., Tuna, M., Santos, C.N., 2023. (Poly)phenols in Human Breast Milk and

their health benefits for the newborn (preprint). medRxiv. https://doi.org/10.1101/2023.03.27.23287781

Chen, C.-H.S., Kuo, T.-C., Kuo, H.-C., Tseng, Y.J., Kuo, C.-H., Yuan, T.-H., Chan, C.-C., 2019. Metabolomics of Children and Adolescents Exposed to Industrial Carcinogenic Pollutants. Environ. Sci. Technol. 53, 5454–5465. https://doi.org/10.1021/acs.est.9b00392

Chi, L., Tu, P., Ru, H., Lu, K., 2021. Studies of xenobiotic-induced gut microbiota dysbiosis: from correlation to mechanisms. Gut Microbes 13, 1921912. https://doi.org/10.1080/19490976.2021.1921912

Claesson, M.J., Jeffery, I.B., Conde, S., Power, S.E., O'Connor, E.M., Cusack, S., Harris, H.M.B., Coakley, M., Lakshminarayanan, B., O'Sullivan, O., Fitzgerald, G.F., Deane, J., O'Connor, M., Harnedy, N., O'Connor, K., O'Mahony, D., van Sinderen, D., Wallace, M., Brennan, L., Stanton, C., Marchesi, J.R., Fitzgerald, A.P., Shanahan, F., Hill, C., Ross, R.P., O'Toole, P.W., 2012. Gut microbiota composition correlates with diet and health in the elderly. Nature 488, 178–184. https://doi.org/10.1038/nature11319

Clarke, G., Sandhu, K.V., Griffin, B.T., Dinan, T.G., Cryan, J.F., Hyland, N.P., 2019. Gut Reactions: Breaking Down Xenobiotic–Microbiome Interactions. Pharmacol. Rev. 71, 198–224. https://doi.org/10.1124/pr.118.015768

Claus, S.P., Guillou, H., Ellero-Simatos, S., 2016. The gut microbiota: a major player in the toxicity of environmental pollutants? NPJ Biofilms Microbiomes 2, 16003. https://doi.org/10.1038/npjbiofilms.2016.3

Collins, S.L., Patterson, A.D., 2020. The gut microbiome: an orchestrator of xenobiotic metabolism. Acta Pharm. Sin. B 10, 19–32. https://doi.org/10.1016/j.apsb.2019.12.001

Crudo, F., Barilli, A., Mena, P., Rotoli, B.M., Rio, D.D., Dall'Asta, C., Dellafiora, L., 2021. An in vitro study on the transport and phase II metabolism of the mycotoxin alternariol in combination with the structurally related gut microbial metabolite urolithin C. Toxicol. Lett. 340, 15–22. https://doi.org/10.1016/j.toxlet.2021.01.007

Cushnie, T.P.T., Cushnie, B., Lamb, A.J., 2014. Alkaloids: An overview of their antibacterial, antibiotic-enhancing and antivirulence activities. Int. J. Antimicrob. Agents 44, 377–386. https://doi.org/10.1016/j.ijantimicag.2014.06.001

D'Amico, D., Andreux, P.A., Valdés, P., Singh, A., Rinsch, C., Auwerx, J., 2021. Impact of the Natural Compound Urolithin A on Health, Disease, and Aging. Trends Mol. Med. 27, 687–699. https://doi.org/10.1016/j.molmed.2021.04.009

David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A., Biddinger, S.B., Dutton, R.J., Turnbaugh, P.J., 2014. Diet rapidly and reproducibly alters the human gut microbiome. Nature 505, 559–563. https://doi.org/10.1038/nature12820

de Ferrars, R.M., Czank, C., Zhang, Q., Botting, N.P., Kroon, P.A., Cassidy, A., Kay, C.D., 2014. The pharmacokinetics of anthocyanins and their metabolites in humans. Br. J. Pharmacol. 171, 3268–3282. https://doi.org/10.1111/bph.12676

Dubbelboer, I.R., Le Roux-Pullen, L., Gehring, R., 2023. Systematic review of physiologically based kinetic lactation models for transfer of xenobiotic compounds to milk. Toxicol. Appl. Pharmacol. 467, 116495. https://doi.org/10.1016/j.taap.2023.116495

Dudzik, D., Barbas-Bernardos, C., García, A., Barbas, C., 2018. Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. J. Pharm. Biomed. Anal., Review issue 2017 147, 149–173. https://doi.org/10.1016/j.jpba.2017.07.044

Ezekiel, C.N., Abia, W.A., Braun, D., Šarkanj, B., Ayeni, K.I., Oyedele, O.A., Michael-Chikezie, E.C., Ezekiel, V.C., Mark, B.N., Ahuchaogu, C.P., Krska, R., Sulyok, M., Turner, P.C., Warth, B., 2022. Mycotoxin exposure biomonitoring in breastfed and non-exclusively breastfed Nigerian

children. Environ. Int. 158, 106996. https://doi.org/10.1016/j.envint.2021.106996

Fisher, M., MacPherson, S., Braun, J.M., Hauser, R., Walker, M., Feeley, M., Mallick, R., Bérubé, R., Arbuckle, T.E., 2017. Paraben Concentrations in Maternal Urine and Breast Milk and Its Association with Personal Care Product Use. Environ. Sci. Technol. 51, 4009–4017. https://doi.org/10.1021/acs.est.6b04302

Flasch, M., Bueschl, C., Del Favero, G., Adam, G., Schuhmacher, R., Marko, D., Warth, B., 2022a. Elucidation of xenoestrogen metabolism by non-targeted, stable isotope-assisted mass spectrometry in breast cancer cells. Environ. Int. 158, 106940. https://doi.org/10.1016/j.envint.2021.106940

Flasch, M., Fitz, V., Rampler, E., Ezekiel, C.N., Koellensperger, G., Warth, B., 2022b. Integrated Exposomics/Metabolomics for Rapid Exposure and Effect Analyses. JACS Au 2, 2548–2560. https://doi.org/10.1021/jacsau.2c00433

Flasch, M., Koellensperger, G., Warth, B., 2023. Comparing the sensitivity of a low- and a high-resolution mass spectrometry approach for xenobiotic trace analysis: An exposome-type case study. Anal. Chim. Acta 1279, 341740. https://doi.org/10.1016/j.aca.2023.341740

Giacomoni, F., Silva, A.B., Bronze, M. do R., Gladine, C., Hollman, P.C., Kopec, R.E., Yanwen, D.L., Micheau, P., Santos, M.C. dos, Pavot, B., Schmidt, G., Morand, C., Sardà, M.U., Manjarrez, N.V., Verny, M.-A., Wiczkowski, W., Knox, C., Manach, C., 2017. PhytoHub, an online platform to gather expert knowledge on polyphenols and other dietary phytochemicals. Presented at the Environmental Science.

Govarts, E., Portengen, L., Lambrechts, N., Bruckers, L., Den Hond, E., Covaci, A., Nelen, V., Nawrot, T.S., Loots, I., Sioen, I., Baeyens, W., Morrens, B., Schoeters, G., Vermeulen, R., 2020. Early-life exposure to multiple persistent organic pollutants and metals and birth weight: Pooled analysis in four Flemish birth cohorts. Environ. Int. 145, 106149. https://doi.org/10.1016/j.envint.2020.106149

Gowd, V., Karim, N., Shishir, M.R.I., Xie, L., Chen, W., 2019. Dietary polyphenols to combat the metabolic diseases via altering gut microbiota. Trends Food Sci. Technol. 93, 81–93. https://doi.org/10.1016/j.tifs.2019.09.005

Gu, Y., Peach, J.T., Warth, B., 2023. Sample preparation strategies for mass spectrometry analysis in human exposome research: Current status and future perspectives. TrAC Trends in Anal. Chem. 166, 117151. https://doi.org/10.1016/j.trac.2023.117151

Guo, J., Huan, T., 2020. Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography–Mass Spectrometry Based Untargeted Metabolomics. Anal. Chem. 92, 8072–8080. https://doi.org/10.1021/acs.analchem.9b05135

Haddad, S., Ayotte, P., Verner, M.-A., 2015. Derivation of exposure factors for infant lactational exposure to persistent organic pollutants (POPs). Regul. Toxicol. Pharmacol. 71, 135–140. https://doi.org/10.1016/j.yrtph.2014.12.003

Hamid, N., Junaid, M., Pei, D.-S., 2021. Combined toxicity of endocrine-disrupting chemicals: A review. Ecotoxicol. Environ. Saf. 215, 112136. https://doi.org/10.1016/j.ecoenv.2021.112136

Heindel, J.J., Balbus, J., Birnbaum, L., Brune-Drisse, M.N., Grandjean, P., Gray, K., Landrigan, P.J., Sly, P.D., Suk, W., Slechta, D.C., Thompson, C., Hanson, M., 2015. Developmental Origins of Health and Disease: Integrating Environmental Influences. Endocrinology 156, 3416–3421. https://doi.org/10.1210/en.2015-1394

Henning, S.M., Yang, J., Lee, R.-P., Huang, J., Thames, G., Korn, M., Ben-Nissan, D., Heber, D., Li, Z., 2022. Pomegranate juice alters the microbiota in breast milk and infant stool: a pilot study. Food Funct. 13, 5680–5689. https://doi.org/10.1039/D2FO00280A

Holzhausen, E.A., Shen, N., Chalifour, B., Tran, V., Li, Z., Sarnat, J.A., Chang, H.H., Jones, D.P., Goran, M.I., Liang, D., Alderete, T.L., 2023. Longitudinal profiles of the fecal metabolome during the first 2 years of life. Sci. Rep. 13, 1886. https://doi.org/10.1038/s41598-023-28862-z

Hou, K., Wu, Z.-X., Chen, X.-Y., Wang, J.-Q., Zhang, D., Xiao, C., Zhu, D., Koya, J.B., Wei, L., Li, J., Chen, Z.-S., 2022. Microbiota in health and diseases. Signal Transduct. Target Ther. 7, 1–28. https://doi.org/10.1038/s41392-022-00974-4

International Agency for Research on Cancer, Lyon (FR), 2015. Mycotoxin control in low- and middle-income countries. PMID: 27030861.

Jamnik, T., Flasch, M., Braun, D., Fareed, Y., Wasinger, D., Seki, D., Berry, D., Berger, A., Wisgrill, L., Warth, B., 2022. Next-generation biomonitoring of the early-life chemical exposome in neonatal and infant development. Nat. Commun. 13, 2653. https://doi.org/10.1038/s41467-022-30204-y

Kalia, V., Belsky, D.W., Baccarelli, A.A., Miller, G.W., 2022. An exposomic framework to uncover environmental drivers of aging. Exposome 2, osac002. https://doi.org/10.1093/exposome/osac002

Kim, J.H., Moon, N., Ji, E., Moon, H.-B., 2023. Effects of postnatal exposure to phthalate, bisphenol a, triclosan, parabens, and per- and poly-fluoroalkyl substances on maternal postpartum depression and infant neurodevelopment: a korean mother-infant pair cohort study. Environ. Sci. Pollut. Res. 30, 96384–96399. https://doi.org/10.1007/s11356-023-29292-0

Kim, M., Kim, N., Han, J., 2014. Metabolism of Kaempferia parviflora Polymethoxyflavones by Human Intestinal Bacterium Bautia sp. MRG-PMF1. J. Agric. Food Chem. 62, 12377–12383. https://doi.org/10.1021/jf504074n

Kirwan, J.A., Weber, R.J.M., Broadhurst, D.I., Viant, M.R., 2014. Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. Sci. Data 1, 140012. https://doi.org/10.1038/sdata.2014.12

Kolde, R., 2018. pheatmap: Pretty Heatmaps.

Krausová, M., Ayeni, K.I., Wisgrill, L., Ezekiel, C.N., Braun, D., Warth, B., 2022. Trace analysis of emerging and regulated mycotoxins in infant stool by LC-MS/MS. Anal. Bioanal. Chem. 414, 7503–7516. https://doi.org/10.1007/s00216-021-03803-9

Krausová, M., Braun, D., Buerki-Thurnherr, T., Gundacker, C., Schernhammer, E., Wisgrill, L., Warth, B., 2023. Understanding the Chemical Exposome During Fetal Development and Early Childhood: A Review. Annu. Rev. Pharmacol. Toxicol. 63, 517–540. https://doi.org/10.1146/annurev-pharmtox-051922-113350

Křížová, L., Dadáková, K., Kašparovská, J., Kašparovský, T., 2019. Isoflavones. Molecules 24, 1076. https://doi.org/10.3390/molecules24061076

Lees, H.J., Swann, J.R., Wilson, I.D., Nicholson, J.K., Holmes, E., 2013. Hippurate: The Natural History of a Mammalian–Microbial Cometabolite. J. Proteome Res. 12, 1527–1546. https://doi.org/10.1021/pr300900b

Liu, X., Mao, B., Gu, J., Wu, J., Cui, S., Wang, G., Zhao, J., Zhang, H., Chen, W., 2021. Blautia—a new functional genus with potential probiotic properties? Gut Microbes 13, 1875796. https://doi.org/10.1080/19490976.2021.1875796

Lu, H., Rosenbaum, S., 2014. Developmental Pharmacokinetics in Pediatric Populations. J. Pediatr. Pharmacol. Ther. 19, 262–276. https://doi.org/10.5863/1551-6776-19.4.262Lu, Z., Chan, Y.-T., Lo, K.K.-H., Wong, V.W.-S., Ng, Y.-F., Li, S.-Y., Ho, W.-W., Wong, M.-S., Zhao, D., 2021. Levels of polyphenols and phenolic metabolites in breast milk and their association with plant-based food intake in Hong Kong lactating women. Food Funct. 12, 12683–12695. https://doi.org/10.1039/D1FO02529E

Marasas, W.F.O., Riley, R.T., Hendricks, K.A., Stevens, V.L., Sadler, T.W., Gelineau-van

Waes, J., Missmer, S.A., Cabrera, J., Torres, O., Gelderblom, W.C.A., Allegood, J., Martínez, C., Maddox, J., Miller, J.D., Starr, L., Sullards, M.C., Roman, A.V., Voss, K.A., Wang, E., Merrill, A.H., 2004. Fumonisins Disrupt Sphingolipid Metabolism, Folate Transport, and Neural Tube Development in Embryo Culture and In Vivo: A Potential Risk Factor for Human Neural Tube Defects among Populations Consuming Fumonisin-Contaminated Maize. J. Nutr. 134, 711–716. https://doi.org/10.1093/jn/134.4.711

McMillan, A., Orimadegun, A.E., Sumarah, M.W., Renaud, J., da Encarnacao, M.M., Gloor, G.B., Akinyinka, O.O., Reid, G., Allen, S.J., 2016. Metabolic derangements identified through untargeted metabolomics in a cross-sectional study of Nigerian children with severe acute malnutrition. Metabolomics 13, 13. https://doi.org/10.1007/s11306-016-1150-2

Miller, G.W., Jones, D.P., 2014. The Nature of Nurture: Refining the Definition of the Exposome. Toxicol. Sci. 137, 1–2. https://doi.org/10.1093/toxsci/kft251

Missmer, S.A., Suarez, L., Felkner, M., Wang, E., Merrill, A.H., Rothman, K.J., Hendricks, K.A., 2006. Exposure to Fumonisins and the Occurrence of Neural Tube Defects along the Texas–Mexico Border. Environ. Health Perspect. 114, 237–241. https://doi.org/10.1289/ehp.8221

Neveu, V., Nicolas, G., Salek, R.M., Wishart, D.S., Scalbert, A., 2020. Exposome-Explorer 2.0: an update incorporating candidate dietary biomarkers and dietary associations with cancer risk. Nucleic Acids Res. 48, D908–D912. https://doi.org/10.1093/nar/gkz1009

Niderkorn, V., Boudra, H., Morgavi, D.P., 2006. Binding of Fusarium mycotoxins by fermentative bacteria in vitro. J. Appl. Microbiol. 101, 849–856. https://doi.org/10.1111/j.1365-2672.2006.02958.x

Oesterle, I., Braun, D., Berry, D., Wisgrill, L., Rompel, A., Warth, B., 2021. Polyphenol Exposure, Metabolism, and Analysis: A Global Exposomics Perspective. Annu. Rev. Food Sci. Technol. 12, 461–484. https://doi.org/10.1146/annurev-food-062220-090807

Oesterle, I., Braun, D., Rompel, A., Warth, B., 2022. Quantifying up to 90 polyphenols simultaneously in human bio-fluids by LC-MS/MS. Anal. Chim. Acta 1216, 339977. https://doi.org/10.1016/j.aca.2022.339977

Oesterle, I., Pristner, M., Berger, S., Wang, M., Verri Hernandes, V., Rompel, A., Warth, B., 2023. Exposomic Biomonitoring of Polyphenols by Non-Targeted Analysis and Suspect Screening. Anal. Chem. https://doi.org/10.1021/acs.analchem.3c01393

Ololade, I.A., Oladoja, N.A., Ololade, O.O., Oloye, F.F., Adeola, A.O., Alabi, A.B., Ademila, O., Adanigbo, P., Owolabi, M.B., 2018. Geographical distribution of perfluorooctanesulfonate and perfluorooctanoate in selected rivers from Nigeria. J. Environ. Chem. Eng. 6, 4061–4069. https://doi.org/10.1016/j.jece.2018.06.020

Pandey, K.B., Rizvi, S.I., 2009. Plant polyphenols as dietary antioxidants in human health and disease. Oxid. Med. Cell. Longev. 2, 270–278. https://doi.org/10.4161/oxim.2.5.9498

Pang, Z., Zhou, G., Ewald, J., Chang, L., Hacariz, O., Basu, N., Xia, J., 2022. Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. Nat. Protoc. 17, 1735–1761. https://doi.org/10.1038/s41596-022-00710-w

Penttinen, P., Jaehrling, J., Damdimopoulos, A.E., Inzunza, J., Lemmen, J.G., van der Saag, P., Pettersson, K., Gauglitz, G., Mäkelä, S., Pongratz, I., 2007. Diet-Derived Polyphenol Metabolite Enterolactone Is a Tissue-Specific Estrogen Receptor Activator. Endocrinology 148, 4875–4886. https://doi.org/10.1210/en.2007-0289

Pjevac, P., Hausmann, B., Schwarz, J., Kohl, G., Herbold, C.W., Loy, A., Berry, D., 2021. An Economical and Flexible Dual Barcoding, Two-Step PCR Approach for Highly Multiplexed Amplicon Sequencing. Front. Microbiol. 12. https://doi.org/10.3389/fmicb.2021.669776

Pourchet, M., Debrauwer, L., Klanova, J., Price, E.J., Covaci, A., Caballero-Casero, N., Oberacher, H., Lamoree, M., Damont, A., Fenaille, F., Vlaanderen, J., Meijer, J., Krauss, M., Sarigiannis, D., Barouki, R., Le Bizec, B., Antignac, J.-P., 2020. Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues. Environ. Int. 139, 105545. https://doi.org/10.1016/j.envint.2020.105545

R Core Team, 2023. R: A Language and Environment for Statistical Computing.

Rappaport, S.M., Barupal, D.K., Wishart, D., Vineis, P., Scalbert, A., 2014. The Blood Exposome and Its Role in Discovering Causes of Disease. Environ. Health Perspect. 122, 769–774. https://doi.org/10.1289/ehp.1308015

Rasheed, H., Xu, Y., Kimanya, M.E., Pan, X., Li, Z., Zou, X., Shirima, C.P., Holmes, M., Routledge, M.N., Gong, Y.Y., 2021. Estimating the health burden of aflatoxin attributable stunting among children in low income countries of Africa. Sci. Rep. 11, 1619. https://doi.org/10.1038/s41598-020-80356-4

Rasouli, H., Nayeri, F.D., Khodarahmi, R., 2022. May phytophenolics alleviate aflatoxins-induced health challenges? A holistic insight on current landscape and future prospects. Front. Nutr. 9. https://doi.org/10.3389/fnut.2022.981984

Roca, A., Oluwalana, C., Camara, B., Bojang, A., Burr, S., Davis, T.M.E., Bailey, R., Kampmann, B., Mueller, J., Bottomley, C., D'Alessandro, U., 2015. Prevention of bacterial infections in the newborn by pre-delivery administration of azithromycin: Study protocol of a randomized efficacy trial. BMC Pregnancy Childbirth 15, 302. https://doi.org/10.1186/s12884-015-0737-3

Rowland, A., Miners, J.O., Mackenzie, P.I., 2013. The UDP-glucuronosyltransferases: Their role in drug metabolism and detoxification. Int. J. Biochem. Cell Biol. 45, 1121–1132. https://doi.org/10.1016/j.biocel.2013.02.019

Scalbert, A., Morand, C., Manach, C., Rémésy, C., 2002. Absorption and metabolism of polyphenols in the gut and impact on health. Biomed. Pharmacother. 56, 276–282. https://doi.org/10.1016/S0753-3322(02)00205-6

Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. Environ. Sci. Technol. 48, 2097–2098. https://doi.org/10.1021/es5002105

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. 13, 2498–2504. https://doi.org/10.1101/gr.1239303

Sillner, N., Walker, A., Lucio, M., Maier, T.V., Bazanella, M., Rychlik, M., Haller, D., Schmitt-Kopplin, P., 2021. Longitudinal Profiles of Dietary and Microbial Metabolites in Formula- and Breastfed Infants. Front. Mol. Biosci. 8. https://doi.org/10.3389/fmolb.2021.660456

Song, B.J., Jouni, Z.E., Ferruzzi, M.G., 2013. Assessment of phytochemical content in human milk during different stages of lactation. Nutrition 29, 195–202. https://doi.org/10.1016/j.nut.2012.07.015

Steed, A.L., Christophi, G.P., Kaiko, G.E., Sun, L., Goodwin, V.M., Jain, U., Esaulova, E., Artyomov, M.N., Morales, D.J., Holtzman, M.J., Boon, A.C.M., Lenschow, D.J., Stappenbeck, T.S., 2017. The microbial metabolite desaminotyrosine protects from influenza through type I interferon. Science 357, 498–502. https://doi.org/10.1126/science.aam5336

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., Arita, M., 2015. MS-DIAL: Data Independent MS/MS Deconvolution for Comprehensive Metabolome Analysis. Nat. Methods 12, 523–526. https://doi.org/10.1038/nmeth.3393

Tsugawa, H., Kind, T., Nakabayashi, R., Yukihira, D., Tanaka, W., Cajka, T., Saito, K., Fiehn, O., Arita, M., 2016. Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. Anal. Chem. 88, 7946–7958. https://doi.org/10.1021/acs.analchem.6b00770

Vitale, C.M., Price, E.J., Miller, G.W., David, A., Antignac, J.-P., Barouki, R., Klánová, J., 2021. Analytical strategies for chemical exposomics: exploring limits and feasibility. Exposome 1, osab003. https://doi.org/10.1093/exposome/osab003

Vulić, J.J., Ćebović, T.N., Čanadanović, V.M., Ćetković, G.S., Djilas, S.M., Čanadanović-Brunet, J.M., Velićanski, A.S., Cvetković, D.D., Tumbas, V.T., 2013. Antiradical, antimicrobial and cytotoxic activities of commercial beetroot pomace. Food Funct. 4, 713–721. https://doi.org/10.1039/C3FO30315B

Wang, K., Guo, J., Chang, X., Gui, S., 2022. Painong-San extract alleviates dextran sulfate sodium-induced colitis in mice by modulating gut microbiota, restoring intestinal barrier function and attenuating TLR4/NF-κB signaling cascades. J. Pharm. Biomed. Anal. 209, 114529. https://doi.org/10.1016/j.jpba.2021.114529

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis.

Wild, C.P., 2005. Complementing the Genome with an "Exposome": The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. Cancer Epidemiol. Biomark. Prev. 14, 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456

Wilke, C., 2020. cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2."

Wilson, A.S., Koller, K.R., Ramaboli, M.C., Nesengani, L.T., Ocvirk, S., Chen, C., Flanagan, C.A., Sapp, F.R., Merritt, Z.T., Bhatti, F., Thomas, T.K., O'Keefe, S.J.D., 2020. Diet and the Human Gut Microbiome: An International Review. Dig. Dis. Sci. 65, 723–740. https://doi.org/10.1007/s10620-020-06112-w

Wopereis, H., Oozeer, R., Knipping, K., Belzer, C., Knol, J., 2014. The first thousand days - intestinal microbiology of early life: establishing a symbiosis. Pediatr. Allergy Immunol. 25, 428–438. https://doi.org/10.1111/pai.12232

Wu, A.H., Yu, M.C., Tseng, C.-C., Pike, M.C., 2008. Epidemiology of soy exposures and breast cancer risk. Br. J. Cancer 98, 9–14. https://doi.org/10.1038/sj.bjc.6604145

Yang, W., Chen, X., Li, Y., Guo, S., Wang, Z., Yu, X., 2020. Advances in Pharmacological Activities of Terpenoids. Nat. Prod. Commun. 15, 1934578X20903555. https://doi.org/10.1177/1934578X20903555

Zhang, Z., Peng, X., Li, S., Zhang, N., Wang, Y., Wei, H., 2014. Isolation and Identification of Quercetin Degrading Bacteria from Human Fecal Microbes. PLoS One 9, e90531. https://doi.org/10.1371/journal.pone.0090531

Supplementary Information

Two supplementary data files are available online for this publication. The first is an xlsx that contains the mentioned tables, such as the results and the MS parameters. The second file is a pdf that contains various additional figures, such as the boxplots from the semi-quantification results of each identified feature.

DOI: https://doi.org/10.1101/2023.11.08.566030

6.6. Publication #5 (Original Research): Oesterle et al. 2023c

Status	Submitted for publication
Title	**Polyphenolic Profiling of Plants and Edible Mushrooms to Aid Characterizing Polyphenol Oxidase Selectivity**
Authors	Ian Oesterle[a,b,c], Mathias Pretzler[c], Annette Rompel[c], Benedikt Warth[a,d]
Affiliations	[a]Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Vienna 1090, Austria [b]Doctoral School of Chemistry, University of Vienna, Vienna 1090, Austria [c]Fakultät für Chemie, Institut für Biophysikalische Chemie, Universität Wien, Wien 1090, Österreich [d]Exposome Austria, Research Infrastructure and National EIRENE Hub, Vienna 1090, Austria
Year	2023
Journal	*ChemRxiv*
Journal Information	Pre-print
Accepted	n.a.
DOI	https://doi.org/10.26434/chemrxiv-2023-phq9f
Contribution	Ian Oesterle aided in the sample preparation optimization, performed the final sample preparation and LC-HRMS measurements, feature extraction and result evaluation, took part in result interpretation, and took the lead in writing the manuscript.

Polyphenolic Profiling of Plants and Edible Mushrooms to Aid Characterizing Polyphenol Oxidase Selectivity

Ian OESTERLE[a,b,c,1], Matthias PRETZLER[b,2], Annette ROMPEL[b,*], Benedikt WARTH[a,d,*]

[a]University of Vienna, Faculty of Chemistry, Department of Food Chemistry and Toxicology, 1090 Vienna, Austria
[b]Universität Wien, Fakultät für Chemie, Institut für Biophysikalische Chemie, 1090 Wien, Austria; www.bpc.univie.ac.at/
[c]University of Vienna, Vienna Doctoral School of Chemistry (DoSChem), 1090 Vienna, Austria
[d]Exposome Austria, Research Infrastructure and National EIRENE Hub, Austria
[1]ian.oesterle@univie.ac.at
[2]matthias.pretzler@univie.ac.at

*Corresponding authors: Benedikt Warth, benedikt.warth@univie.ac.at, +43-1-4277-70806; Annette Rompel, annette.rompel@univie.ac.at, +43-1-4277-52502

Highlights

- Suspect screening identified polyphenol profiles in nine plant and fungi species.
- Overall, 401 features were annotated with 221 flavonoids and 180 non-flavonoids.
- 64% of flavonoids and 17% of non-flavonoids were conjugated with a sugar moiety.
- Quantitative analysis with 90 polyphenol standards was applied to the samples.
- The most abundant class was flavanones, mainly conjugated with sugar moieties.

Abstract

Polyphenol oxidases are important metalloenzymes that catalyze the oxidation of polyphenols. Organisms can contain multiple isoenzymes whose possible different functionality is not understood. Moreover, their substrate preference is still unexplored, and a working hypothesis assumes their specificity is tailored to the substrates naturally present. Polyphenols are a broad molecular class whose presence in an organism is frequently unknown. Therefore, nine different mushrooms and plants were applied a LC-HRMS-based suspect screening workflow to investigate the polyphenols present. Overall, 401 features were annotated, composed of 221 flavonoids and 180 non-flavonoids. As 64% of the flavonoids were conjugated with a

glycone, (semi-)quantification was applied when a reference standard of the aglycone was available. Isomers of hesperetin-O-glucoside were ubiquitous in all investigated samples. Additionally, a targeted LC-MS/MS assay covering 90 polyphenols was applied, showing phenolic acids as prevalent. This study provides new insights into the complex polyphenol profiles of nine mushroom and plant species.

Keywords

tyrosinase, catechol oxidase, nontargeted analysis, suspect screening, mass spectrometry, food bio-actives

1. Introduction

Polyphenols are a complex molecular family that can be separated into multiple chemical classes based on their structure (Figure 1a) (Pietta *et al.* 2003). They are derived from the secondary metabolite pathway, e.g. the shikimate pathway (Santos-Sánchez *et al.* 2019), which is not completely understood yet. Polyphenols are involved in various processes such as protecting against UV light (Donkin and Martin 1981) or producing melanin, the end result of enzymatic browning (Derardja *et al.* 2022). Additionally, secondary metabolites, including polyphenols, can beneficially influence human health (Weigmann and Hoensch 2022).

Enzymatic browning is a process catalyzed by polyphenol oxidases (PPOs), a family of metalloenzymes. PPOs are type-III dicopper enzymes, mainly consisting of tyrosinases (TYRs) and catechol oxidases (COs). TYRs (Scheme S1) reduce O_2 to catalyze the *o*-hydroxylation of monophenols and the subsequent oxidation of the intermediate (or exogenous) *o*-diphenols to *o*-quinones (EC 1.14.18.1) (Mason and Wright 1949, Molitor *et al.* 2016), referred to as monophenolase activity. COs (Scheme S1) show only the latter reaction, referred to as diphenolase activity (EC 1.10.3.1). One theory proposes that PPOs are involved in a defense mechanism by creating reactive oxygen species that both trigger defense pathways and reduce the nutritional value of the affected tissue (Boeckx *et al.* 2015). Understanding a PPO's function requires understanding its substrate scope. It is assumed that the amino acids surrounding the active site of PPOs differ and therefore the enzymes accept different substrates (Derardja *et al.* 2022). Thus, knowing which polyphenols are present in specific species would assist in better understanding PPOs and their involvement in biological pathways (Pretzler and Rompel 2018, Kampatsikas and Rompel 2021).

As numerous polyphenols exist (Figure 1a), methods based on liquid chromatography (LC) coupled to mass spectrometry (MS) are the analytical platform of choice for measuring a multitude of polyphenols simultaneously. Furthermore, untargeted LC-MS workflows allow to investigate analytes for which reference standards are not (or hardly) commercially available (Oesterle *et al.* 2021). However, analyte annotation/identification in untargeted LC-MS is a challenging endeavor, and fragmentation spectra (MS^2)

are important for gaining structural information (Aydoğan 2020). One MS^2 acquisition technique is data-dependent acquisition (DDA) and it involves selecting specific parent mass-to-charge (*m/z*) ratios for fragmentation (Guo and Huan 2020). Moreover, suspect screening, an intermediate approach between targeted and untargeted methods, allows to screen for analytes with known structures that may be present in the samples (Pourchet *et al.* 2020). Therefore, suspect screening can be combined with DDA to selectively investigate polyphenols in biological matrices.

In this work, nine different plants and fungi containing PPOs were subjected to a suspect screening DDA LC-HRMS workflow and a targeted LC-MS/MS workflow to assess polyphenols present. These diverse samples and the types of PPOs present included oyster mushroom (*Pleurotus ostreatus,* TYRs), white button mushroom (*Agaricus bisporus*, TYRs (Pretzler *et al.* 2017)), brown button mushroom (*Agaricus bisporus*, TYRs), moringa leaves (*Moringa oleifera*, TYRs and COs), enoki mushroom (*Flammulina filiformis*, TYRs), tickseed petals (*Coreopsis grandiflora*, TYRs, COs and aurone synthase (Kaintz *et al.* 2014, Molitor *et al.* 2016)), shiitake mushroom (*Lentinula edodes*, TYRs), king trumpet mushroom (*Pleurotus eryngii*, TYRs), and chanterelle mushroom (*Cantharellus cibarius*, TYRs).

2. Materials and method

2.1. Reagents, solvents, and chemicals

90 authentic polyphenol reference standards were deployed for identification and (semi-)quantitation purposes. The commercial sources and purity of the reference standards and the used solvents are listed in Table S1. The standards were prepared with methanol to create stock solutions and working mixtures at various concentrations to spike the samples and create calibration curves (Table S2).

2.2. Sample acquisition and preparation

The mushroom and plant samples were purchased from local stores (Table S3). After acquisition, each sample was freeze-dried, and approximately 1 g was ground to a fine powder in a mortar and pestle. Then, roughly 50 mg was weighed into a 2 mL micro reaction tube (Table S3), and 500 µL extraction solvent

(water:acetonitrile:methanol, 1:1:1, with 1% v/v formic acid) was added. The sample was briefly vortexed, placed in a thermoshaker (1 h, 75°C, 500 rpm), and centrifuged (10 min, 22°C, 18000 x g). The supernatant was subsequently transferred to an LC vial. The extraction procedure was repeated a second time, and the supernatants combined. For the untargeted workflow, approximately 11 mg of Moringa leaves were extracted with 300 µL of extraction solvent since a higher dilution was needed as this sample contained elevated levels of polyphenols compared to the other organic material. The samples were prepared independently a second time for the targeted measurements (Table S3), with a 7-point calibration curve in neat solvent (Table S2). Two additional samples (50.5 and 50.1 mg, respectively) of white button mushroom were prepared and spiked with the reference standards (spike matrix-matched, Table S2) to estimate extraction efficiencies and signal suppression and enhancement (SSE) effects.

2.3. LC-(HR)MS(/MS) instrumentation and parameters

Both, the LC-HRMS and the LC-MS/MS system were connected to a 1290 Infinity II UHPLC (Agilent), with chromatographic parameters previously optimized by Oesterle *et al.* 2022 (Table S4). The LC-HRMS system consisted of an ESI source and a ZenoTOF 7600 (SCIEX) MS, with the source and MS parameters listed in Table S5. Data was acquired in DDA mode with the 10 most intense ions selected for fragmentation and an MS^2 inclusion list comprised of the combined 3223 entries from PhytoHub (Giacomoni *et al.* 2017) and Exposome-Explorer (Neveu *et al.* 2020). The LC-MS/MS system consisted of an ESI source and a QTrap 6500+ (SCIEX) MS, with the source and MS parameters listed in Table S6. Data was acquired in scheduled multiple reaction monitoring mode (MRM) as previously described by Oesterle *et al.* 2022.

2.4. Data analysis

The LC-HRMS data analysis followed a workflow adapted from Oesterle *et al.* 2023. Briefly, the raw files were converted to MGF files with MSConvert (v3.0.22067) (Chambers *et al.* 2012), and then imported into MZMine (v3.3.0) (Schmid *et al.* 2023) for feature processing. Suspect screening was done in MZMine

by monoisotopic mass matching of the features with the entries in PhytoHub (Giacomoni *et al.* 2017) and Exposome-Explorer (Neveu *et al.* 2020). MZMine was used for annotation with a spectral library from MS-DIAL Metabolomics (Tsugawa *et al.* 2015) using a minimum cosine score (normalized dot-product) of 0.7. SIRIUS (v5.5.7) (Dührkop *et al.* 2019) was used for *in silico* fragmentation (minimum score of -50). All the steps and parameters applied are listed in Table S7. Conversely, the LC-MS/MS results were analyzed with Skyline (v21.2) (MacLean *et al.* 2010).

2.5. Quality control

A process blank and a pooled quality control (QC) were prepared following the sample preparation procedure. The process blank sample was left empty while the QC sample consisted of combining approximately 5 mg of each sample (Table S3). During each extraction step of the QC, the supernatant was split by transferring 150 µL into two separate vials. One of these samples was spiked with the reference standard mixture (Table S2) and the other was used for a three-point dilution series with a constant dilution factor of two and diluting with the extraction solvent. The QCs and process blank were measured in triplicates. Features were removed during data analysis if they were not present in all three replicates of the QC, had a relative standard deviation greater than 30% in the QC measurements, or had an average chromatographic peak area in the process blank sample greater than 1/3 of the average area in the QC (Kirwan *et al.* 2014). Finally, features that did not have a Spearman correlation coefficient greater than 0 in the QC dilution series were removed.

3. Results and discussion

From all nine samples, a total of 401 features were annotated with the described workflow (Table S8). As polyphenols are a molecular class rich in isomers, establishing structural annotation confidence is not straightforward. Therefore, the annotation of the relevant features was split into different confidence levels adapted from Schymanski *et al.* 2014 and described in detail by Oesterle *et al.* 2023. From the 401 features, 39 were identified with reference standards (level 1) and 81 were annotated by spectral libraries (level 2a).

An additional 163 features had a spectral library match, but it was unknown which isomer was the correct structure (level 3a). Utilizing *in silico* fragmentation, an additional six features with a known structure were annotated (level 3b) and 112 features with the isomeric structure unknown (level 3c). A heatmap was generated to show all the features following log10 normalization (Figure S3). Moreover, 170 (42%) of the features were conjugated with one or multiple sugar moieties contributing to the difficulty in identifying the correct isomeric structure (Table S8). Following feature annotation, the polyphenol class of each feature (Figure 1b) was determined resulting in 221 (55%) flavonoids and 180 (45%) non-flavonoids overall.

Quantitation in untargeted workflows is challenging as reference standards are typically lacking. As 42% of the features were conjugated with a glycone (Figure 2a-c and S4), the concentration of these conjugated features (Table S9) was approximated based on the relative chromatographic peak area to their corresponding aglycone if a reference standard was available (Table S2). Standard addition was applied as the reference standards were spiked in the QC sample (Table S2). For certain analytes, the calibration curve in neat solvent was used as no change in peak area was observed after spiking due to high natural concentrations in the QC. The samples were then quantified with a targeted LC-MS/MS method (Table S10). In this method, extraction efficiency and SSE were approximated by spiking extra samples of white button mushrooms. Table S11 lists the five features of each sample from the untargeted workflow that had the largest chromatographic peak areas, the largest relative concentration, and the five analytes with the highest concentration from the targeted analysis. It is not unexpected, that the most intense/concentrated polyphenols determined by the two alternative approaches often do not match while some of the concentrations are in line with each other. This is caused by the far broader chemical space of the untargeted method but also its limited capacity for delivering full quantitative results.

The seven mushroom samples had similar polyphenolic profiles (Table S8 and Figures 1b, 2a-c, S3, and S4) with 191, 171, 211, 175, 211, 188, and 272 features determined in the white button, brown button, oyster, king trumpet, shiitake, enoki, and chanterelle mushroom samples, respectively. In all mushroom samples, features were on average as follows: 28% flavanones, 15% hydroxycinnamic acids, 13% hydroxybenzoic acids, 9% flavones, and 6% flavonols. Moreover, on average 50% of the features were

156 conjugated with a sugar moiety (Figure S4), which included 74% of the flavanones, 87% of the flavones,
157 18% of the hydroxybenzoic acids, and 15% of the hydroxycinnamic acids. Various flavanones, notably
158 isomers of hesperetin-O-glucoside (Figures 2d-e, and S5), were among the most abundant polyphenols
159 determined in the samples (Table S8 and S11). For example, an isomer of hesperetin-O-glucoside (7.0 min,
160 level 3c), had concentrations between 26 and 110 µg/g of dry weight. Besides flavanones, mainly phenolic
161 acids, e.g. homoprotocatechuic acid and tyrosol (Figure S5), were highly prevalent, especially by targeted
162 LC-MS/MS.
163 Some minor differences between the mushroom species was observed. For example, shiitake was the
164 only mushroom sample with urolithin M5 (level 3c) (Figure S5) detected. Moreover, only two flavonols
165 (1%) were detected in the king trumpet mushroom, while more was detected in the other mushrooms, e.g.
166 enoki had 16 (9%) flavonols. A prevalent analyte of white and brown button mushrooms was γ-L-
167 gluatminyl-4-hydroxybenzene (Figure S5) with concentrations of 16 and 25 µg/g of dry weight,
168 respectively. This analyte was previously observed as prevalent in white button mushroom (Weijn *et al.*
169 2013).
170 Previous studies utilizing targeted methods focusing on phenolic acids have been applied to some of
171 the tested species (Gąsecka *et al.* 2018, Reis *et al.* 2012). In the work at hand, phenolic acids were also
172 prevalent, and in similar concentrations as described previously. Besides phenolic acids, the other
173 polyphenol classes have not been investigated as thoroughly by LC-MS. One study investigated catechins
174 in king trumpet mushrooms (Calabretti *et al.* 2021), but these polyphenols were not detected here, which
175 could be from the non-optimal nature of the workflow this polyphenol class. Untargeted methods were
176 previously applied to mushrooms to study the general composition (Uffelman *et al.* 2023) or changes, e.g.
177 during browning (Fu *et al.* 2022). However, studies focusing on polyphenols are scarce. For example,
178 untargeted analysis of shiitake mushrooms revealed 17 metabolites that included phenolic acids, e.g. gallic
179 acid or syringic acid, flavones, e.g. apigenin, and catechins (Nam *et al.* 2021). Some of the annotated
180 polyphenols were also detected in our study, e.g. gallic acid, but others, e.g. catechin, were not detected

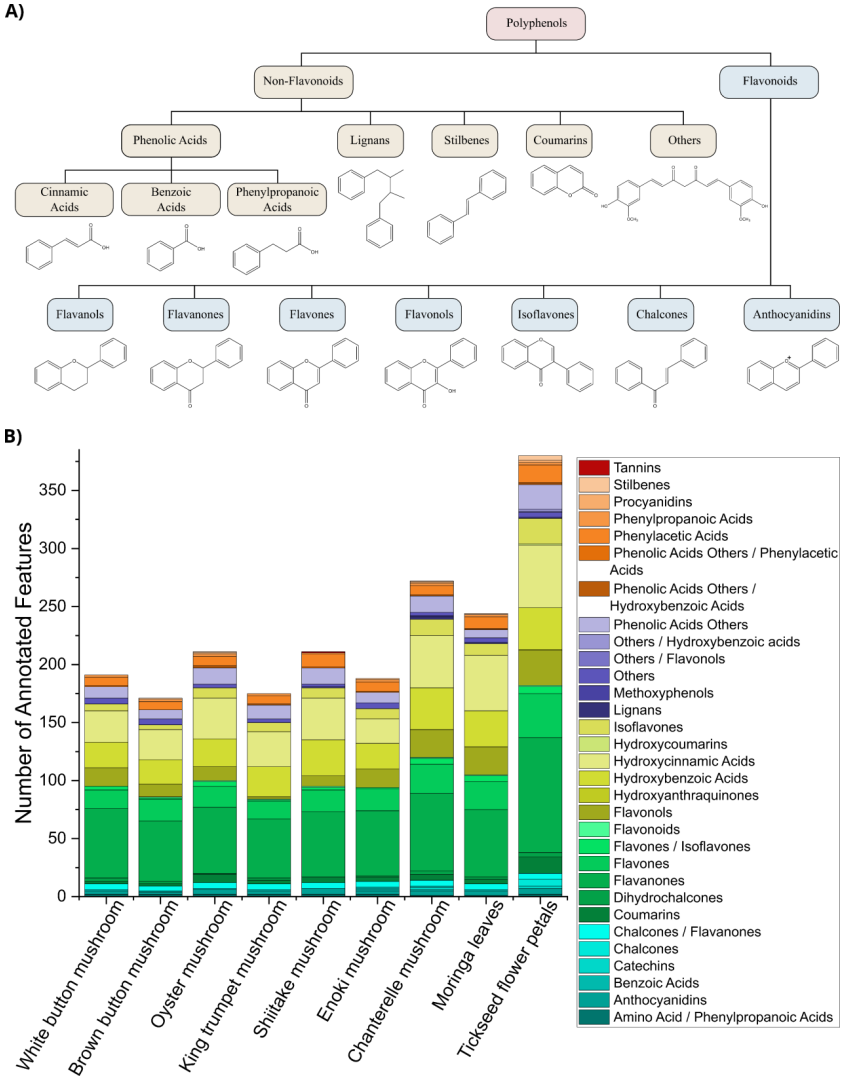


Figure 1. A) Diagram of polyphenol classes based on their structure, along with their backbone (Pietta *et al.* 2003).

B) Stacked bar chart indicating the number of features detected in each polyphenol class for the nine tested species.

181 even with the targeted method which raises concerns on the identification confidence of previous work not

182 relying on a high number of authentic reference standards.

183 Contrary to the mushroom samples, more polyphenols were detected and at higher concentrations in
184 the two plant samples. 244 features were detected in moringa leaves, and the main classes were again
185 flavanones (24%), hydroxycinnamic acids (20%), hydroxybenzoic acids (13%), flavones (10%), and
186 flavonols (10%). Tickseed flower petals showed the greatest number of polyphenol-related signals with 380
187 features detected, mainly flavanones (26%), hydroxycinnamic acids (14%), flavones (10%),
188 hydroxybenzoic aids (9%), and flavonols (8%). Moreover, 55% of the features in moringa leaves were
189 conjugated with a glycone while 42% of the features in tickseed flower petals were conjugated. Similar to
190 the mushroom samples, phenolic acids were prevalent in moringa leave, but it also showed high
191 concentrations of flavonols including rutin (Table S11). The most prevalent features in tickseed flowers
192 were flavanones, especially various isomers of hesperetin-O-glucoside. Moreover, targeted LC-MS
193 revealed flavanones and flavonols in high concentrations, while in the other samples phenolic acids
194 predominated.

195 To the best of our knowledge, no previous investigation of polyphenols in tickseed flower petals exists,
196 either targeted or untargeted. Contrarily, multiple studies of polyphenols in moringa leaves exist. For
197 example, 291 polyphenols were previously annotated, of which 39 were anthocyanins (Rocchetti *et al.*
198 2020). A higher number of anthocyanins was assigned than here (5 analytes), most likely due to the LC
199 conditions used here not being optimized for anthocyanins (Oesterle *et al.* 2022). Similar to our work (Table
200 S11), previous research found that moringa leaves contain high concentrations of caffeoylquinic isomers,
201 gallic acid glucoside, and quercetin and kaempferol with sugar moieties (Rocchetti *et al.* 2020, Oldoni *et*
202 *al.* 2019).

203 How differences in the polyphenol profiles of the species relate to their respective PPOs is challenging
204 to establish as comprehensive information on PPOs is still limited (Table S3). Genomic data for moringa
205 leaves, and enoki, king trumpet, and chanterelle mushrooms are only available at the assembly level.
206 Moreover, no genomic data for tickseed flower petals exists, however, sequences for two isoforms of aurone
207 synthases are available (Table S3) (Kaintz *et al.* 2014). For the remaining mushroom species, partial

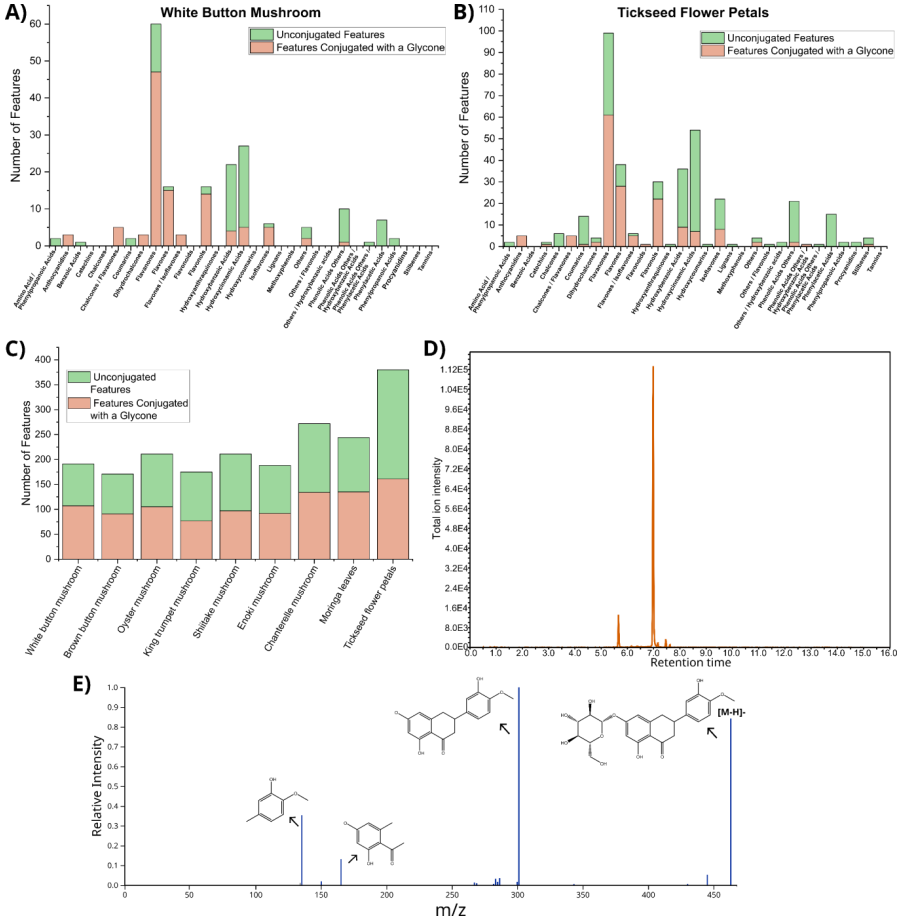


Figure 2. Stacked bar chart illustrating the number of features conjugated with a glycone versus unconjugated in white button mushroom (**A**), tickseed flower (**B**), and over the nine different samples (**C**). **D)** The extracted ion chromatogram of *m/z* 463.124 in king trumpet mushroom. **E)** MS^2 spectra and fragment structures of hesperetin-7-glucoside (*m/z* 463.124) at 7.0 min.

208 information on the TYRs present is available (Table S3), with two isoforms found in oyster mushrooms,

three in shiitake mushrooms, and six in button mushrooms.

The active center of PPOs contains two copper ions that activate molecular oxygen and catalyze the reaction (Scheme S1). The metal ions are chelated by three histidine residues each (Figures S1 and S2). Substrate access to the active site is regulated by the amino acids lining the entrance to the cavity holding the copper ions (Figure S2) (Bijelic *et al.* 2015), and the most prominent are the two residues located immediately after the second and third histidine of copper B, the activity controllers (Kampatsikas *et al.* 2017). All the fungal enzymes feature either aspartic acid or asparagine at the first activity controller position, while for the two plant enzymes it is a threonine. The second activity controller involves a more diverse set of amino acids, and the most common in the fungal enzymes are the small and uncharged valine and leucine. Moreover, the even smaller glycine, the slightly bigger methionine, and the bulky phenylalanine exist in one enzyme each. While for the two plant enzymes, *Cg*AUS1 and *Cg*AUS2, the bulky and positively charged arginine is featured, most likely to restrict their activity on monophenolic substrates (Kampatsikas *et al.* 2020). A similar picture is seen at the gatekeeper position (Bijelic *et al.* 2015), which is occupied by small amino acids (proline, valine, alanine, leucine and glycine) for most of the fungal enzymes. Additionally, in a set of isoenzymes that contain small amino acids at this position, one fungal enzyme exists with the polar asparagine and arginine in each. As typical for plant PPOs, the two enzymes from tickseed exhibit a bulky phenylalanine at this position (Tran *et al.* 2012). While this seemed to interfere with the spatial needs of the hydroxylation reaction on monophenols, structural investigations on walnut tyrosinase (Bijelic *et al.* 2015) showed that this residue is, in fact, flexible enough to not only allow the passage of substrate but may even enhance the binding of incoming phenols via interactions with their aromatic ring. From the limited available data on PPOs in the sampled fungi and plants, the isoenzymes present in any one organism seem to complement each other in the accessibility of their active center for substrates. Therefore, further investigations are required, especially among isolated isoenzymes and the different polyphenols found naturally.

5. Conclusion

A suspect screening LC-HRMS workflow for polyphenols was successfully applied to seven mushroom and two plant species known to contain polyphenol oxidases. This allowed for the comprehensive identification and annotation of polyphenols present in these samples. Moreover, (semi-)quantification was performed on conjugated and unconjugated polyphenols from the untargeted data, and the samples were additionally quantified with a targeted LC-MS/MS assay covering a wide range of polyphenol standards. Overall, a similar number of flavonoids and non-flavonoids were detected, with roughly two-thirds of the flavonoids conjugated to a sugar moiety. Tickseed flower petals showed a higher total number of polyphenols and their respective concentrations compared to moringa leaves. The mushroom species exhibited similar polyphenol profiles, especially the two button mushrooms. Besides yielding insights on the polyphenols naturally present in these species, the comprehensive profiling is intended to aid in defining and interpreting the selectivity of PPOs in the future.

Supporting information

Two supplementary information files are included. A Word file that contains schemes and additional figures, and an Excel file that contains tables with additional information on the material and methods, and the suspect screening annotation results along with the relative concentrations found in each sample.

Declaration of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the University of Vienna through the Exposome Austria Research Infrastructure and the Austrian Science Fund FWF: P32326 (A.R.).

CRediT author contribution statement

Ian Oesterle: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Matthias Pretzler:** Conceptualization, Formal Analysis, Writing - Review & Editing. **Annette Rompel:** Conceptualization, Writing - Review & Editing, Funding, Resources, Supervision. **Benedikt Warth:** Conceptualization, Writing - Review & Editing, Funding, Resources, Supervision.

Data availability

The raw data files have been submitted to the MetaboLights data repository (MTBLS8782).

Acknowledgment

The authors would like to thank all members of the Warth and Rompel labs for the feedback and valuable support they gave, especially Iris Neuburger for her aid in optimizing the sample preparation. They would also like to thank the Mass Spectrometry Center of the Faculty of Chemistry for technical support.

References

Aydoğan, C. (2020). Recent advances and applications in LC-HRMS for food and plant natural products: A critical review. *Analytical and Bioanalytical Chemistry*, *412*(9), 1973–1991. https://doi.org/10.1007/s00216-019-02328-6

Bijelic, A., Pretzler, M., Molitor, C., Zekiri, F., & Rompel, A. (2015). The Structure of a Plant Tyrosinase from Walnut Leaves Reveals the Importance of "Substrate-Guiding Residues" for Enzymatic Specificity. *Angewandte Chemie International Edition*, *54*(49), 14677–14680. https://doi.org/10.1002/anie.201506994

Boeckx, T., Winters, A. L., Webb, K. J., & Kingston-Smith, A. H. (2015). Polyphenol oxidase in leaves: Is there any significance to the chloroplastic localization? *Journal of Experimental Botany*, *66*(12), 3571–3579. https://doi.org/10.1093/jxb/erv141

278 Calabretti, A., Mang, S. M., Becce, A., Castronuovo, D., Cardone, L., Candido, V., & Camele, I. (2021). Comparison
279 of Bioactive Substances Content between Commercial and Wild-Type Isolates of Pleurotus eryngii.
280 *Sustainability*, *13*(7), Article 7. https://doi.org/10.3390/su13073777
281 Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt,
282 B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y.,
283 Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S. L., Nuwaysir, L. M., Lefebvre, B.,
284 Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B.,
285 Chadick, T., Holly, K., Eckels, J., Deutsch, E. W., Moritz, R. L., Katz, J. E., Agus, D. B., MacCoss, M., Tabb,
286 D. L., Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*,
287 *30*(10), Article 10. https://doi.org/10.1038/nbt.2377
288 Derardja, A., Pretzler, M., Kampatsikas, I., Radovic, M., Fabisikova, A., Zehl, M., Barkat, M., & Rompel, A. (2022).
289 Polyphenol oxidase and enzymatic browning in apricot (*Prunus armeniaca* L.): Effect on phenolic composition
290 and deduction of main substrates. *Current Research in Food Science*, *5*, 196–206.
291 https://doi.org/10.1016/j.crfs.2021.12.015
292 Donkin, M. E., & Martin, E. S. (1981). Blue Light Absorption by Guard Cells of *Commelina communis* and *Allium*
293 *cepa*. *Zeitschrift Für Pflanzenphysiologie*, *102*(4), 345–352. https://doi.org/10.1016/S0044-328X(81)80206-1
294 Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., Dorrestein, P. C., Rousu, J.,
295 & Böcker, S. (2019). SIRIUS 4: A rapid tool for turning tandem mass spectra into metabolite structure
296 information. *Nature Methods*, *16*(4), Article 4. https://doi.org/10.1038/s41592-019-0344-8
297 Fu, Y., Yu, Y., Tan, H., Wang, B., Peng, W., & Sun, Q. (2022). Metabolomics reveals dopa melanin involved in the
298 enzymatic browning of the yellow cultivars of East Asian golden needle mushroom (*Flammulina filiformis*).
299 *Food Chemistry*, *370*, 131295. https://doi.org/10.1016/j.foodchem.2021.131295
300 Gąsecka, M., Magdziak, Z., Siwulski, M., & Mleczek, M. (2018). Profile of phenolic and organic acids, antioxidant
301 properties and ergosterol content in cultivated and wild growing species of Agaricus. *European Food Research*
302 *and Technology*, *244*(2), 259–268. https://doi.org/10.1007/s00217-017-2952-9
303 Giacomoni, F., Pavot, B., Schmidt, G., Morand, C., Urpi-Sarda, M., Vazquez-Manjarrez, N., Verny, M.-A.,
304 Wiczkowski, W., Knox, C., & Manach, C. (2017). *PhytoHub, an online platform to gather expert knowledge on*
305 *polyphenols and other dietary phytochemicals*.

306 Guo, J., & Huan, T. (2020). Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in
307 Liquid Chromatography–Mass Spectrometry Based Untargeted Metabolomics. *Analytical Chemistry*, *92*(12),
308 8072–8080. https://doi.org/10.1021/acs.analchem.9b05135
309 Kaintz, C., Molitor, C., Thill, J., Kampatsikas, I., Michael, C., Halbwirth, H., & Rompel, A. (2014). Cloning and
310 functional expression in E. coli of a polyphenol oxidase transcript from *Coreopsis grandiflora* involved in aurone
311 formation. *FEBS Letters*, *588*(18), 3417–3426. https://doi.org/10.1016/j.febslet.2014.07.034
312 Kampatsikas, I., Bijelic, A., Pretzler, M., & Rompel, A. (2017). Three recombinantly expressed apple tyrosinases
313 suggest the amino acids responsible for mono- versus diphenolase activity in plant polyphenol oxidases.
314 *Scientific Reports*, *7*(1), Article 8860. https://doi.org/10.1038/s41598-017-08097-5
315 Kampatsikas, I., Pretzler, M., & Rompel, A. (2020). Identification of Amino Acid Residues Responsible for C−H
316 Activation in Type-III Copper Enzymes by Generating Tyrosinase Activity in a Catechol Oxidase. *Angewandte*
317 *Chemie International Edition*, *59*(47), 20940–20945. https://doi.org/10.1002/anie.202008859
318 Kampatsikas, I., & Rompel, A. (2021). Similar but Still Different: Which Amino Acid Residues Are Responsible for
319 Varying Activities in Type-III Copper Enzymes? *ChemBioChem*, *22*(7), 1161–1175.
320 https://doi.org/10.1002/cbic.202000647
321 Kirwan, J. A., Weber, R. J. M., Broadhurst, D. I., & Viant, M. R. (2014). Direct infusion mass spectrometry
322 metabolomics dataset: A benchmark for data processing and quality control. *Scientific Data*, *1*(1), Article
323 140012. https://doi.org/10.1038/sdata.2014.12
324 MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler,
325 D. C., & MacCoss, M. J. (2010). Skyline: An open source document editor for creating and analyzing targeted
326 proteomics experiments. *Bioinformatics*, *26*(7), 966–968. https://doi.org/10.1093/bioinformatics/btq054
327 Mason, H. S., & Wright, C. I. (1949). The Chemistry of Melanin: V. Oxidation of Dihydroxyphenylalanine by
328 Tyrosinase. *Journal of Biological Chemistry*, *180*(1), 235–247. https://doi.org/10.1016/S0021-9258(18)56739-
329 9
330 Molitor, C., Mauracher, S. G., & Rompel, A. (2016). Aurone synthase is a catechol oxidase with hydroxylase activity
331 and provides insights into the mechanism of plant polyphenol oxidases. *Proceedings of the National Academy*
332 *of Sciences*, *113*(13), E1806–E1815. https://doi.org/10.1073/pnas.1523575113

333 Nam, M., Choi, J. Y., & Kim, M.-S. (2021). Metabolic Profiles, Bioactive Compounds, and Antioxidant Capacity in
334 *Lentinula edodes* Cultivated on Log versus Sawdust Substrates. *Biomolecules*, *11*(11), Article 11.
335 https://doi.org/10.3390/biom11111654
336 Neveu, V., Nicolas, G., Salek, R. M., Wishart, D. S., & Scalbert, A. (2020). Exposome-Explorer 2.0: An update
337 incorporating candidate dietary biomarkers and dietary associations with cancer risk. *Nucleic Acids Research*,
338 *48*(D1), D908–D912. https://doi.org/10.1093/nar/gkz1009
339 Oesterle, I., Braun, D., Berry, D., Wisgrill, L., Rompel, A., & Warth, B. (2021). Polyphenol Exposure, Metabolism,
340 and Analysis: A Global Exposomics Perspective. *Annual Review of Food Science and Technology*, *12*(1), 461–
341 484. https://doi.org/10.1146/annurev-food-062220-090807
342 Oesterle, I., Braun, D., Rompel, A., & Warth, B. (2022). Quantifying up to 90 polyphenols simultaneously in human
343 bio-fluids by LC-MS/MS. *Analytica Chimica Acta*, *1216*, 339977. https://doi.org/10.1016/j.aca.2022.339977
344 Oesterle, I., Pristner, M., Berger, S., Wang, M., Verri Hernandes, V., Rompel, A., & Warth, B. (2023). Exposomic
345 Biomonitoring of Polyphenols by Non-Targeted Analysis and Suspect Screening. *Analytical Chemistry, 95*(28),
346 10686-10694. https://doi.org/10.1021/acs.analchem.3c01393
347 Oldoni, T. L. C., Merlin, N., Karling, M., Carpes, S. T., Alencar, S. M. de, Morales, R. G. F., Silva, E. A. da, & Pilau,
348 E. J. (2019). Bioguided extraction of phenolic compounds and UHPLC-ESI-Q-TOF-MS/MS characterization of
349 extracts of *Moringa oleifera* leaves collected in Brazil. *Food Research International*, *125*, 108647.
350 https://doi.org/10.1016/j.foodres.2019.108647
351 Pietta, P., Minoggio, M., & Bramati, L. (2003). Plant Polyphenols: Structure, Occurrence and Bioactivity. In A.
352 Rahman (Ed.), *Studies in Natural Products Chemistry* (Vol. 28, pp. 257–312). Elsevier.
353 https://doi.org/10.1016/S1572-5995(03)80143-6
354 Pourchet, M., Debrauwer, L., Klanova, J., Price, E. J., Covaci, A., Caballero-Casero, N., Oberacher, H., Lamoree, M.,
355 Damont, A., Fenaille, F., Vlaanderen, J., Meijer, J., Krauss, M., Sarigiannis, D., Barouki, R., Le Bizec, B., &
356 Antignac, J.-P. (2020). Suspect and non-targeted screening of chemicals of emerging concern for human
357 biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and
358 harmonisation issues. *Environment International*, *139*, 105545. https://doi.org/10.1016/j.envint.2020.105545
359 Pretzler, M., Bijelic, A., & Rompel, A. (2017). Heterologous expression and characterization of functional mushroom
360 tyrosinase (AbPPO4). *Scientific Reports*, *7*(1), Article 1810. https://doi.org/10.1038/s41598-017-01813-1

361 Pretzler, M., & Rompel, A. (2018). What causes the different functionality in type-III-copper enzymes? A state of the
362 art perspective. *Inorganica Chimica Acta*, *481*, 25–31. https://doi.org/10.1016/j.ica.2017.04.041
363 Reis, F. S., Martins, A., Barros, L., & Ferreira, I. C. F. R. (2012). Antioxidant properties and phenolic profile of the
364 most widely appreciated cultivated mushrooms: A comparative study between in vivo and in vitro samples. *Food*
365 *and Chemical Toxicology*, *50*(5), 1201–1207. https://doi.org/10.1016/j.fct.2012.02.013
366 Rocchetti, G., Pagnossa, J. P., Blasi, F., Cossignani, L., Hilsdorf Piccoli, R., Zengin, G., Montesano, D., Cocconcelli,
367 P. S., & Lucini, L. (2020). Phenolic profiling and in vitro bioactivity of *Moringa oleifera* leaves as affected by
368 different extraction solvents. *Food Research International*, *127*, 108712.
369 https://doi.org/10.1016/j.foodres.2019.108712
370 Santos-Sánchez, N. F., Salas-Coronado, R., Hernández-Carlos, B., Villanueva-Cañongo, C., Santos-Sánchez, N. F.,
371 Salas-Coronado, R., Hernández-Carlos, B., & Villanueva-Cañongo, C. (2019). Shikimic Acid Pathway in
372 Biosynthesis of Phenolic Compounds. In *Plant Physiological Aspects of Phenolic Compounds*. IntechOpen.
373 https://doi.org/10.5772/intechopen.83815
374 Schmid, R., Heuckeroth, S., Korf, A., Smirnov, A., Myers, O., Dyrlund, T. S., Bushuiev, R., Murray, K. J., Hoffmann,
375 N., Lu, M., Sarvepalli, A., Zhang, Z., Fleischauer, M., Dührkop, K., Wesner, M., Hoogstra, S. J., Rudt, E.,
376 Mokshyna, O., Brungs, C., Ponomarov, K., Mutabdžija, L., Damiani, T., Pudney, C. J., Earll, M., Helmer, P. O.,
377 Fallon, T. R., Schulze, T., Rivas-Ubach, A., Bilbao, A., Richter, H., Nothias, L.-F., Wang, M., Orešič, M., Weng,
378 J.-K., Böcker, S., Jeibmann, A., Hayen, H., Karst, U., Dorrestein, P. C., Petras, D., Du, X., Pluskal, T. (2023).
379 Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nature Biotechnology*, *41*(4), Article
380 4. https://doi.org/10.1038/s41587-023-01690-2
381 Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., & Hollender, J. (2014). Identifying Small
382 Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science &*
383 *Technology*, *48*(4), 2097–2098. https://doi.org/10.1021/es5002105
384 Tran, L. T., Taylor, J. S., & Constabel, C. P. (2012). The polyphenol oxidase gene family in land plants: Lineage-
385 specific duplication and expansion. *BMC Genomics*, *13*(1), 395. https://doi.org/10.1186/1471-2164-13-395
386 Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., &
387 Arita, M. (2015). MS-DIAL: Data Independent MS/MS Deconvolution for Comprehensive Metabolome
388 Analysis. *Nature Methods*, *12*(6), 523–526. https://doi.org/10.1038/nmeth.3393

389 Uffelman, C. N., Doenges, K. A., Armstrong, M. L., Quinn, K., Reisdorph, R. M., Tang, M., Krebs, N. F., Reisdorph,
390 N. A., & Campbell, W. W. (2023). Metabolomics Profiling of White Button, Crimini, Portabella, Lion's Mane,
391 Maitake, Oyster, and Shiitake Mushrooms Using Untargeted Metabolomics and Targeted Amino Acid Analysis.
392 *Foods*, *12*(16), Article 16. https://doi.org/10.3390/foods12162985
393 Weigmann, B., & Hoensch, H. P. (2022). Nutritional Health and Bioflavonoids. *Recent Progress in Nutrition*, *2*(3),
394 Article 3. https://doi.org/10.21926/rpn.2203017
395 Weijn, A., van den Berg-Somhorst, D. B. P. M., Slootweg, J. C., Vincken, J.-P., Gruppen, H., Wichers, H. J., & Mes,
396 J. J. (2013). Main Phenolic Compounds of the Melanin Biosynthesis Pathway in Bruising-Tolerant and Bruising-
397 Sensitive Button Mushroom (*Agaricus bisporus*) Strains. *Journal of Agricultural and Food Chemistry*, *61*(34),
398 8224–8231. https://doi.org/10.1021/jf4020558

Supplementary Information

Two supplementary data files are available online for this publication. The first file is a word document that contains various additional figures and schemes. The second file is a xlsx that contains all of the tables mentioned, such as the detailed results.

DOI: https://doi.org/10.26434/chemrxiv-2023-phq9f

7. Conclusion and Outlook

Polyphenols and their metabolites have a significant impact on human health and well-being. Therefore, it is vital to better understand human exposure to these xenobiotics. Typically, methods for analyzing polyphenols only focused on a limited number of analytes from a few specific sub-classes rather than attempting to study their totality. Thus, a targeted LC-MS method was developed for 90 polyphenol analytes representing all the main sub-classes, including human polyphenol metabolites. This analytical procedure allowed to investigate numerous chemically diverse polyphenols simultaneously, and in a manner that the assay can easily be adapted to different applications, such as for different human and plant matrices. As the focus of this thesis was on the comprehensive study of human polyphenol exposure, this targeted LC-MS method was optimized for three different biological matrices: urine, serum, and plasma. Through in-house validation for these three biological matrices, the method showed its suitability and stability. In addition, the LC-MS method showed high sensitivity with optimal extraction efficiency and minimal signal suppression and enhancement effect, despite using low sample volumes and a high-throughput sample preparation. Thus, overall, the method showed its readiness to be applied in future large cohorts.

Compared to targeted workflows, untargeted workflows allow to go beyond analytes for which reference standards are commercially available. In the context of polyphenols, this is crucial as diverse analytes exist, and reference standards, especially for conjugated polyphenols, are expensive or non-existent. Therefore, an untargeted workflow was developed using the targeted LC-MS method as a benchmark. As untargeted workflows use high-resolution mass spectrometers, they typically suffer from lower sensitivity than targeted workflows. Therefore, the sensitivity gap between the two workflows was quantified for polyphenols, with the untargeted workflow showing feasible sensitivity in comparison to the targeted method. Moreover, the untargeted workflow was applied to real-life samples, and showed comparable performance with the targeted workflow for polyphenols at naturally occurring concentrations. As numerous non-targeted analysis tools exist, various tools, such as *in silico* fragmentation or neutral loss querying, were benchmarked, allowing to better understand their advantages and limitations to find features that are potentially polyphenol-related. Lastly, the workflow was also developed to include suspect screening of polyphenols and other biomarkers of environmental exposure. This allowed to investigate the presence of a variety of known polyphenols whose reference standards were not readily available. In addition, it demonstrated the workflow's potential to simultaneously study other xenobiotics and their metabolites. Therefore, in the future, the workflow can be applied to investigate co-exposure effects between polyphenols and other xenobiotics.

Following their development, the suitability and applicability of the targeted and

untargeted workflows were tested in two individual studies. Firstly, both workflows were applied for the polyphenolic profiling of nine distinct plant and fungi species. Consequently, the sample preparation procedure was adapted for these types of samples. This demonstrated that the workflows can be easily transferred among different human, microbial, and plant matrices. Moreover, the obtained results provides in-depth insights into the polyphenol profiles of important plant and fungi species. In addition, this allows to further investigate the selectivity of polyphenol oxidases and their involvement in biological pathways. Secondly, the untargeted workflow was adapted and applied in a longitudinal study involving samples collected at three distinct time points from mother-infant pairs. This allowed to investigate polyphenols, as well as other xenobiotics, in matching breast milk and stool samples from a highly exposed population. Different xenobiotic classes, including polyphenols, antibiotics, plasticizers, and mycotoxins were detected in the samples. Moreover, an increase in exposure to both potentially beneficial and adverse xenobiotics was observed as complementary foods were introduced in the infant's diet.

In summary, thoroughly investigating human exposure to polyphenols and their metabolites is a complex task. As numerous diverse polyphenols exist, cutting-edge analytical techniques are required. The development of the LC-MS workflows and their applications in different pilot studies showed their significant potential. Though, intrinsic limitations still exist such as the need to improve the sensitivity of high-resolution instruments used in untargeted approaches or the feature annotation tools used, e.g. expanding spectral libraries. The mother-infant pilot study showed correlations between xenobiotics with each other and with the gut microbiome, displaying the capacity of these workflows to gain potential insights on the impact of xenobiotics with future health development. This is key as exposures during early life and a healthy evolution of the microbiome has an impact on the development of chronic diseases. Therefore future research with applying these workflows to larger cohorts, especially long-term longitudinal studies, is vital to better understand the influence of polyphenols on human health.

Bibliography

[1] H. Cory, S. Passarelli, J. Szeto, M. Tamez and J. Mattei, "The role of polyphenols in human health and food systems: A mini-review", *Frontiers in Nutrition*, vol. 5, p. 87, 21st Sep. 2018, ISSN: 2296-861X. DOI: 10.3389/fnut.2018.00087.

[2] F. Shahidi and J. Yeo, "Bioactivities of phenolics by focusing on suppression of chronic diseases: A review", *International Journal of Molecular Sciences*, vol. 19, no. 6, p. 1573, Jun. 2018. DOI: 10.3390/ijms19061573.

[3] C. Costa, A. Tsatsakis, C. Mamoulakis, M. Teodoro, G. Briguglio, E. Caruso, D. Tsoukalas, D. Margina, E. Dardiotis, D. Kouretas and C. Fenga, "Current evidence on the effect of dietary polyphenols intake on chronic diseases", *Food and Chemical Toxicology*, vol. 110, pp. 286–299, 1st Dec. 2017, ISSN: 0278-6915. DOI: 10.1016/j.fct.2017.10.023.

[4] F. Cardona, C. Andrés-Lacueva, S. Tulipani, F. J. Tinahones and M. I. Queipo-Ortuño, "Benefits of polyphenols on gut microbiota and implications in human health", *The Journal of Nutritional Biochemistry*, vol. 24, no. 8, pp. 1415–1422, 1st Aug. 2013, ISSN: 0955-2863. DOI: 10.1016/j.jnutbio.2013.05.001.

[5] L. Křížová, K. Dadáková, J. Kašparovská and T. Kašparovský, "Isoflavones", *Molecules*, vol. 24, no. 6, p. 1076, 19th Mar. 2019, ISSN: 1420-3049. DOI: 10.3390/molecules24061076.

[6] H. Rasouli, F. D. Nayeri and R. Khodarahmi, "May phytophenolics alleviate aflatoxins-induced health challenges? a holistic insight on current landscape and future prospects", *Frontiers in Nutrition*, vol. 9, 2022, ISSN: 2296-861X. DOI: 10.3389/fnut.2022.981984.

[7] J. J. Heindel, J. Balbus, L. Birnbaum, M. N. Brune-Drisse, P. Grandjean, K. Gray, P. J. Landrigan, P. D. Sly, W. Suk, D. C. Slechta, C. Thompson and M. Hanson, "Developmental origins of health and disease: Integrating environmental influences", *Endocrinology*, vol. 156, no. 10, pp. 3416–3421, 1st Oct. 2015, ISSN: 0013-7227. DOI: 10.1210/en.2015-1394.

[8] D. F. Holderbaum, T. Kon, T. Kudo and M. P. Guerra, "Enzymatic browning, polyphenol oxidase activity, and polyphenols in four apple cultivars: Dynamics during fruit development", *HortScience*, vol. 45, no. 8, pp. 1150–1154, 1st Aug. 2010, ISSN: 0018-5345, 2327-9834. DOI: 10.21273/HORTSCI.45.8.1150.

[9] H. S. Mason and C. I. Wright, "The chemistry of melanin: V. oxidation of dihydroxyphenylalanine by tyrosinase", *Journal of Biological Chemistry*, vol. 180, no. 1, pp. 235–247, 1st Aug. 1949, ISSN: 0021-9258. DOI: 10.1016/S0021-9258(18)56739-9.

[10] G. Carletti, G. Nervo and L. Cattivelli, "Flavonoids and melanins: A common strategy across two kingdoms", *International Journal of Biological Sciences*, vol. 10, no. 10, pp. 1159–1170, 29th Oct. 2014, ISSN: 1449-2288. DOI: 10.7150/ijbs.9672.

[11] T. Boeckx, A. L. Winters, K. J. Webb and A. H. Kingston-Smith, "Polyphenol oxidase in leaves: Is there any significance to the chloroplastic localization?", *Journal of Experimental Botany*, vol. 66, no. 12, pp. 3571–3579, 1st Jun. 2015, ISSN: 0022-0957. DOI: 10.1093/jxb/erv141.

[12] M. Pretzler and A. Rompel, "What causes the different functionality in type-III-copper enzymes? a state of the art perspective", *Inorganica Chimica Acta*, Bioinorganic and bio-inspired chemistry of copper, vol. 481, pp. 25–31, 1st Sep. 2018, ISSN: 0020-1693. DOI: 10.1016/j.ica.2017.04.041.

[13] Ziaullah and H. P. V. Rupasinghe, "Chapter 1 - application of NMR spectroscopy in plant polyphenols associated with human health", in *Applications of NMR Spectroscopy*, A. ur-Rahman and M. I. Choudhary, Eds., Bentham Science Publishers, pp. 3–92, ISBN: 978-1-60805-999-7. DOI: 10.1016/B978-1-60805-999-7.50001-X.

[14] M. d. S. S. Chagas, M. D. Behrens, C. J. Moragas-Tellis, G. X. M. Penedo, A. R. Silva and C. F. Gonçalves-de-Albuquerque, "Flavonols and flavones as potential anti-inflammatory, antioxidant, and antibacterial compounds", *Oxidative Medicine and Cellular Longevity*, vol. 2022, p. 9 966 750, 6th Sep. 2022, ISSN: 1942-0900. DOI: 10.1155/2022/9966750.

[15] I. Oesterle, D. Braun, D. Berry, L. Wisgrill, A. Rompel and B. Warth, "Polyphenol exposure, metabolism, and analysis: A global exposomics perspective", *Annual Review of Food Science and Technology*, vol. 12, no. 1, pp. 461–484, 25th Mar. 2021, ISSN: 1941-1413. DOI: 10.1146/annurev-food-062220-090807.

[16] Y. Zhang, S. Hendrich and P. A. Murphy, "Glucuronides are the main isoflavone metabolites in women", *The Journal of Nutrition*, vol. 133, no. 2, pp. 399–404, 1st Feb. 2003, Publisher: Oxford Academic, ISSN: 0022-3166. DOI: 10.1093/jn/133.2.399.

[17] I. Oesterle, D. Braun, A. Rompel and B. Warth, "Quantifying up to 90 polyphenols simultaneously in human bio-fluids by LC-MS/MS", *Analytica Chimica Acta*, vol. 1216, p. 339 977, 11th Jul. 2022, ISSN: 0003-2670. DOI: 10.1016/j.aca.2022.339977.

[18] I. Oesterle, M. Pristner, S. Berger, M. Wang, V. Verri Hernandes, A. Rompel and B. Warth, "Exposomic biomonitoring of polyphenols by non-targeted analysis and suspect screening", *Analytical Chemistry*, vol. 95, no. 28, pp. 10 686–10 694, 6th Jul. 2023, ISSN: 0003-2700. DOI: 10.1021/acs.analchem.3c01393.

[19] I. Oesterle, M. Pretzler, A. Rompel and B. Warth, "Polyphenolic profiling of plants and edible mushrooms to aid characterizing polyphenol oxidase selectivity", *ChemRxiv*, 20th Oct. 2023, pre-print. DOI: 10.26434/chemrxiv-2023-phq9f.

[20] I. Oesterle, K. I. Ayeni, C. N. Ezekiel, D. Berry, A. Rompel and B. Warth, "Insights into the early-life chemical exposome of nigerian infants and potential correlations with the developing gut microbiome", *bioRxiv*, 11th Nov. 2023, pre-print. DOI: 10.1101/2023.11.08.566030.

[21] M. L. Leporatti and K. Ghedira, "Comparative analysis of medicinal plants used in traditional medicine in italy and tunisia", *Journal of Ethnobiology and Ethnomedicine*, vol. 5, no. 1, p. 31, 26th Oct. 2009, ISSN: 1746-4269. DOI: 10.1186/1746-4269-5-31.

[22] S. Pathania, S. M. Ramakrishnan and G. Bagler, "Phytochemica: A platform to explore phytochemicals of medicinal plants", *Database*, vol. 2015, bav075, 1st Jan. 2015, ISSN: 1758-0463. DOI: 10.1093/database/bav075.

[23] C. Hano and D. Tungmunnithum, "Plant polyphenols, more than just simple natural antioxidants: Oxidative stress, aging and age-related diseases", *Medicines*, vol. 7, no. 5, p. 26, 9th May 2020, ISSN: 2305-6320. DOI: 10.3390/medicines7050026.

[24] A. N. Panche, A. D. Diwan and S. R. Chandra, "Flavonoids: An overview", *Journal of Nutritional Science*, vol. 5, e47, 29th Dec. 2016, ISSN: 2048-6790. DOI: 10.1017/jns.2016.41.

[25] H. E. Khoo, A. Azlan, S. T. Tang and S. M. Lim, "Anthocyanidins and anthocyanins: Colored pigments as food, pharmaceutical ingredients, and the potential health benefits", *Food & Nutrition Research*, vol. 61, no. 1, p. 1 361 779, 13th Aug. 2017, ISSN: 1654-661X. DOI: 10.1080/16546628.2017.1361779.

[26] C. Manach, A. Scalbert, C. Morand, C. Rémésy and L. Jiménez, "Polyphenols: Food sources and bioavailability", *The American Journal of Clinical Nutrition*, vol. 79, no. 5, pp. 727–747, 1st May 2004, ISSN: 0002-9165. DOI: 10.1093/ajcn/79.5.727.

[27] C. Musial, A. Kuban-Jankowska and M. Gorska-Ponikowska, "Beneficial properties of green tea catechins", *International Journal of Molecular Sciences*, vol. 21, no. 5, p. 1744, 4th Mar. 2020, ISSN: 1422-0067. DOI: 10.3390/ijms21051744.

[28] I. C. Arts, P. C. Hollman and D. Kromhout, "Chocolate as a source of tea flavonoids", *The Lancet*, vol. 354, no. 9177, p. 488, 7th Aug. 1999, ISSN: 0140-6736. DOI: 10.1016/S0140-6736(99)02267-9.

[29] B. N. Singh, S. Shankar and R. K. Srivastava, "Green tea catechin, epigallocatechin-3-gallate (EGCG): Mechanisms, perspectives and clinical applications", *Biochemical pharmacology*, vol. 82, no. 12, pp. 1807–1821, 15th Dec. 2011, ISSN: 0006-2952. DOI: 10.1016/j.bcp.2011.07.093.

[30] A. Rauf, M. Imran, T. Abu-Izneid, Iahtisham-Ul-Haq, S. Patel, X. Pan, S. Naz, A. Sanches Silva, F. Saeed and H. A. Rasul Suleria, "Proanthocyanidins: A comprehensive review", *Biomedicine & Pharmacotherapy*, vol. 116, p. 108 999, 1st Aug. 2019, ISSN: 0753-3322. DOI: 10.1016/j.biopha.2019.108999.

[31] G. L. Hostetler, R. A. Ralston and S. J. Schwartz, "Flavones: Food sources, bioavailability, metabolism, and bioactivity12", *Advances in Nutrition*, vol. 8, no. 3, pp. 423–435, 5th May 2017, ISSN: 2161-8313. DOI: 10.3945/an.116.012948.

[32] M. K. Khan, Zill-E-Huma and O. Dangles, "A comprehensive review on flavanones, the major citrus polyphenols", *Journal of Food Composition and Analysis*, vol. 33, no. 1, pp. 85–104, 1st Feb. 2014, ISSN: 0889-1575. DOI: 10.1016/j.jfca.2013.11.004.

[33] S. H. Mah, "Chalcones in diets", in *Handbook of Dietary Phytochemicals*, J. Xiao, S. D. Sarker and Y. Asakawa, Eds., Singapore: Springer, 2019, pp. 1–52, ISBN: 9789811317453. DOI: 10.1007/978-981-13-1745-3_10-1.

[34] M. Kalinowska, E. Gołębiewska, G. Świderski, S. Męczyńska-Wielgosz, H. Lewandowska, A. Pietryczuk, A. Cudowski, A. Astel, R. Świsłocka, M. Samsonowicz, A. B. Złowodzka, W. Priebe and W. Lewandowski, "Plant-derived and dietary hydroxybenzoic acids—a comprehensive study of structural, anti-/pro-oxidant, lipophilic, antimicrobial, and cytotoxic activity in MDA-MB-231 and MCF-7 cell lines", *Nutrients*, vol. 13, no. 9, p. 3107, 4th Sep. 2021, ISSN: 2072-6643. DOI: 10.3390/nu13093107.

[35] F. Wei and M. Tanokura, "Chapter 17 - organic compounds in green coffee beans", in *Coffee in Health and Disease Prevention*, V. R. Preedy, Ed., San Diego: Academic Press, 1st Jan. 2015, pp. 149–162, ISBN: 978-0-12-409517-5. DOI: 10.1016/B978-0-12-409517-5.00017-6.

[36] C. Rodríguez-García, C. Sánchez-Quesada, E. Toledo, M. Delgado-Rodríguez and J. J. Gaforio, "Naturally lignan-rich foods: A dietary tool for health promotion?", *Molecules*, vol. 24, no. 5, p. 917, 6th Mar. 2019, ISSN: 1420-3049. DOI: 10.3390/molecules24050917.

[37] M. Reinisalo, A. Kårlund, A. Koskela, K. Kaarniranta and R. O. Karjalainen, "Polyphenol stilbenes: Molecular mechanisms of defence against oxidative stress and aging-related diseases", *Oxidative Medicine and Cellular Longevity*, vol. 2015, p. 340 520, 2015, ISSN: 1942-0900. DOI: 10.1155/2015/340520.

[38] S. Fuloria, M. Sekar, F. S. Khattulanuar, S. H. Gan, N. N. I. M. Rani, S. Ravi, V. Subramaniyan, S. Jeyabalan, M. Y. Begum, K. Chidambaram, K. V. Sathasivam, S. Z. Safi, Y. S. Wu, R. Nordin, M. N. H. Maziz, V. Kumarasamy, P. T. Lum and N. K. Fuloria, "Chemistry, biosynthesis and pharmacology of viniferin: Potential resveratrol-derived molecules for new drug discovery, development and therapy", *Molecules*, vol. 27, no. 16, p. 5072, 9th Aug. 2022, ISSN: 1420-3049. DOI: 10.3390/molecules27165072.

[39] E. Küpeli Akkol, Y. Genç, B. Karpuz, E. Sobarzo-Sánchez and R. Capasso, "Coumarins and coumarin-related compounds in pharmacotherapy of cancer", *Cancers*, vol. 12, no. 7, p. 1959, 19th Jul. 2020, ISSN: 2072-6694. DOI: 10.3390/cancers12071959.

[40] M. Lončar, M. Jakovljević, D. Šubarić, M. Pavlić, V. Buzjak Služek, I. Cindrić and M. Molnar, "Coumarins in food and methods of their determination", *Foods*, vol. 9, no. 5, p. 645, 18th May 2020, ISSN: 2304-8158. DOI: 10.3390/foods9050645.

[41] M. A. Musa, J. S. Cooperwood and M. O. F. Khan, "A review of coumarin derivatives in pharmacotherapy of breast cancer", *Current medicinal chemistry*, vol. 15, no. 26, pp. 2664–2679, 2008, ISSN: 0929-8673. DOI: 10.2174/092986708786242877.

[42] T. Jin, Z. Song, J. Weng and I. G. Fantus, "Curcumin and other dietary polyphenols: Potential mechanisms of metabolic actions and therapy for diabetes and obesity", *American Journal of Physiology-Endocrinology and Metabolism*, vol. 314, no. 3, E201–E205, Mar. 2018, ISSN: 0193-1849. DOI: 10.1152/ajpendo.00285.2017.

[43] Y. Zhang, L. Cheng, Y. Liu, S. Zhan, Z. Wu, S. Luo and X. Zhang, "Dietary flavonoids: A novel strategy for the amelioration of cognitive impairment through intestinal microbiota", *Journal of the Science of Food and Agriculture*, vol. 103, no. 2, pp. 488–495, 2022, ISSN: 1097-0010. DOI: 10.1002/jsfa.12151.

[44] V. Neveu, J. Perez-Jiménez, F. Vos, V. Crespy, L. du Chaffaut, L. I. Mennen, C. Knox, R. Eisner, J. G. Cruz, D. S. Wishart and A. Scalbert, "Phenol-explorer: An online comprehensive database on polyphenol contents in foods", *Database: The Journal of Biological Databases and Curation*, vol. 2010, bap024 8th Jan. 2010, ISSN: 1758-0463. DOI: 10.1093/database/bap024.

[45] N. F. Santos-Sánchez, R. Salas-Coronado, B. Hernández-Carlos, C. Villanueva-Cañongo, N. F. Santos-Sánchez, R. Salas-Coronado, B. Hernández-Carlos and C. Villanueva-Cañongo, "Shikimic acid pathway in biosynthesis of phenolic compounds", in *Plant Physiological Aspects of Phenolic Compounds*, IntechOpen, 31st Jan. 2019, ISBN: 978-1-78984-034-6. DOI: 10.5772/intechopen.83815.

[46] T. Vogt, "Phenylpropanoid biosynthesis", *Molecular Plant*, vol. 3, no. 1, pp. 2–20, Jan. 2010, ISSN: 16742052. DOI: 10.1093/mp/ssp106.

[47] S. Singh, I. Kaur and R. Kariyat, "The multifunctional roles of polyphenols in plant-herbivore interactions", *International Journal of Molecular Sciences*, vol. 22, no. 3, p. 1442, 1st Feb. 2021, ISSN: 1422-0067. DOI: 10.3390/ijms22031442.

[48] U. A. Meyer, "Overview of enzymes of drug metabolism", *Journal of Pharmacokinetics and Biopharmaceutics*, vol. 24, no. 5, pp. 449–459, Oct. 1996, ISSN: 0090-466X. DOI: 10.1007/BF02353473.

[49] O. A. Almazroo, M. K. Miah and R. Venkataramanan, "Drug metabolism in the liver", *Clinics in Liver Disease*, Drug Hepatotoxicity, vol. 21, no. 1, pp. 1–20, 1st Feb. 2017, ISSN: 1089-3261. DOI: 10.1016/j.cld.2016.08.001.

[50] F. P. Guengerich, "Common and uncommon cytochrome p450 reactions related to metabolism and chemical toxicity", *Chemical Research in Toxicology*, vol. 14, no. 6, pp. 611–650, 1st Jun. 2001, ISSN: 0893-228X. DOI: 10.1021/tx0002583.

[51] S. Ge, Y. Tu and M. Hu, "Challenges and opportunities with predicting in vivo phase II metabolism via glucuronidation from in vitro data", *Current pharmacology reports*, vol. 2, no. 6, pp. 326–338, Dec. 2016, ISSN: 2198-641X. DOI: 10.1007/s40495-016-0076-8.

[52] J. A. Timbrell and T. C. Marrs, "Biotransformation of xenobiotics", in *General, Applied and Systems Toxicology*, John Wiley & Sons, Ltd, 2011, ISBN: 978-0-470-74430-7.

[53] E. P. Allain, M. Rouleau, E. Lévesque and C. Guillemette, "Emerging roles for UDP-glucuronosyltransferases in drug resistance and cancer progression", *British Journal of Cancer*, vol. 122, no. 9, pp. 1277–1287, Apr. 2020, ISSN: 1532-1827. DOI: 10.1038/s41416-019-0722-0.

[54] M. Dean, Y. Hamon and G. Chimini, "The human ATP-binding cassette (ABC) transporter superfamily", *Journal of Lipid Research*, vol. 42, no. 7, pp. 1007–1017, 1st Jul. 2001, ISSN: 0022-2275. DOI: 10.1016/S0022-2275(20)31588-1.

[55] L. Schaller and V. M. Lauschke, "The genetic landscape of the human solute carrier (SLC) transporter superfamily", *Human Genetics*, vol. 138, no. 11, pp. 1359–1377, 2019, ISSN: 0340-6717. DOI: 10.1007/s00439-019-02081-x.

[56] A. B. Shreiner, J. Y. Kao and V. B. Young, "The gut microbiome in health and in disease", *Current opinion in gastroenterology*, vol. 31, no. 1, pp. 69–75, Jan. 2015, ISSN: 0267-1379. DOI: 10.1097/MOG.0000000000000139.

[57] B. Zhu, X. Wang and L. Li, "Human gut microbiome: The second genome of human body", *Protein & Cell*, vol. 1, no. 8, pp. 718–725, 1st Aug. 2010, ISSN: 1674-800X. DOI: 10.1007/s13238-010-0093-z.

[58] C. A. Lozupone, J. I. Stombaugh, J. I. Gordon, J. K. Jansson and R. Knight, "Diversity, stability and resilience of the human gut microbiota", *Nature*, vol. 489, no. 7415, pp. 220–230, Sep. 2012, ISSN: 1476-4687. DOI: 10.1038/nature11550.

[59] H. Wopereis, R. Oozeer, K. Knipping, C. Belzer and J. Knol, "The first thousand days - intestinal microbiology of early life: Establishing a symbiosis", *Pediatric Allergy and Immunology: Official Publication of the European Society of Pediatric Allergy and Immunology*, vol. 25, no. 5, pp. 428–438, Aug. 2014, ISSN: 1399-3038. DOI: 10.1111/pai.12232.

[60] M. Kim and B. A. Benayoun, "The microbiome: An emerging key player in aging and longevity", *Translational Medicine of Aging*, vol. 4, pp. 103–116, 1st Jan. 2020, ISSN: 2468-5011. DOI: 10.1016/j.tma.2020.07.004.

[61] M. H. R. Ansari, S. Saher, R. Parveen, W. Khan, I. A. Khan and S. Ahmad, "Role of gut microbiota metabolism and biotransformation on dietary natural products to human health implications with special reference to biochemoinformatics approach", *Journal of Traditional and Complementary Medicine*, vol. 13, no. 2, pp. 150–160, 1st Mar. 2023, ISSN: 2225-4110. DOI: 10.1016/j.jtcme.2022.03.005.

[62] Y. Zhao, X. Zhong, J. Yan, C. Sun, X. Zhao and X. Wang, "Potential roles of gut microbes in biotransformation of natural products: An overview", *Frontiers in Microbiology*, vol. 13, p. 956 378, 29th Sep. 2022, ISSN: 1664-302X. DOI: 10.3389/fmicb.2022.956378.

[63] T. A. F. Corrêa, M. M. Rogero, N. M. A. Hassimotto and F. M. Lajolo, "The two-way polyphenols-microbiota interactions and their effects on obesity and related metabolic diseases", *Frontiers in Nutrition*, vol. 6, 2019, ISSN: 2296-861X. DOI: 10.3389/fnut.2019.00188.

[64] S. A. Heleno, A. Martins, M. J. R. P. Queiroz and I. C. F. R. Ferreira, "Bioactivity of phenolic acids: Metabolites versus parent compounds: A review", *Food Chemistry*, vol. 173, pp. 501–513, 15th Apr. 2015, ISSN: 0308-8146. DOI: 10.1016/j.foodchem.2014.10.057.

[65] A. Scalbert and G. Williamson, "Dietary intake and bioavailability of polyphenols", *The Journal of Nutrition*, vol. 130, no. 8, 2073S–2085S, 1st Aug. 2000, Publisher: Oxford Academic, ISSN: 0022-3166. DOI: 10.1093/jn/130.8.2073S.

[66] D. del Rio, A. Rodriguez-Mateos, J. P. Spencer, M. Tognolini, G. Borges and A. Crozier, "Dietary (poly)phenolics in human health: Structures, bioavailability, and evidence of protective effects against chronic diseases", *Antioxidants & Redox Signaling*, vol. 18, no. 14, pp. 1818–1892, 15th Jul. 2012, ISSN: 1523-0864. DOI: 10.1089/ars.2012.4581.

[67] J. D. Lambert, S. Sang and C. S. Yang, "Biotransformation of green tea polyphenols and the biological activities of those metabolites", *Molecular Pharmaceutics*, vol. 4, no. 6, pp. 819–825, 1st Dec. 2007, ISSN: 1543-8384. DOI: 10.1021/mp700075m.

[68] K. B. Pandey and S. I. Rizvi, "Plant polyphenols as dietary antioxidants in human health and disease", *Oxidative Medicine and Cellular Longevity*, vol. 2, no. 5, pp. 270–278, 2009, ISSN: 1942-0900. DOI: 10.4161/oxim.2.5.9498.

[69] K. R. Martin, "Polyphenols as dietary supplements: A double-edged sword", *Nutrition and Dietary Supplements*, vol. 2, pp. 1–12, 16th Dec. 2009, ISSN: 1179-1489. DOI: 10.2147/NDS.S6422.

[70] M. J. Bull and N. T. Plummer, "Part 1: The human gut microbiome in health and disease", *Integrative Medicine: A Clinician's Journal*, vol. 13, no. 6, pp. 17–22, Dec. 2014, ISSN: 1546-993X.

[71] D. E. Roopchand, R. N. Carmody, P. Kuhn, K. Moskal, P. Rojas-Silva, P. J. Turnbaugh and I. Raskin, "Dietary polyphenols promote growth of the gut bacterium akkermansia muciniphila and attenuate high-fat diet–induced metabolic syndrome", *Diabetes*, vol. 64, no. 8, pp. 2847–2858, 6th Apr. 2015, ISSN: 0012-1797. DOI: 10.2337/db14-1916.

[72] T. Manso, M. Lores and T. de Miguel, "Antimicrobial activity of polyphenols and natural polyphenolic extracts on clinical isolates", *Antibiotics*, vol. 11, no. 1, p. 46, 30th Dec. 2021, ISSN: 2079-6382. DOI: 10.3390/antibiotics11010046.

[73] A. Duda-Chodak, T. Tarko, P. Satora and P. Sroka, "Interaction of dietary compounds, especially polyphenols, with the intestinal microbiota: A review", *European Journal of Nutrition*, vol. 54, no. 3, pp. 325–341, 1st Apr. 2015, ISSN: 1436-6215. DOI: 10.1007/s00394-015-0852-y.

[74] F. S. vom Saal and L. N. Vandenberg, "Update on the health effects of bisphenol a: Overwhelming evidence of harm", *Endocrinology*, vol. 162, no. 3, bqaa171, 23rd Sep. 2020, ISSN: 0013-7227. DOI: 10.1210/endocr/bqaa171.

[75] S. P. Claus, H. Guillou and S. Ellero-Simatos, "The gut microbiota: A major player in the toxicity of environmental pollutants?", *NPJ Biofilms and Microbiomes*, vol. 2, p. 16 003, 4th May 2016, ISSN: 2055-5008. DOI: 10.1038/npjbiofilms.2016.3.

[76] M. Pristner and B. Warth, "Drug–exposome interactions: The next frontier in precision medicine", *Trends in Pharmacological Sciences*, vol. 41, no. 12, pp. 994–1005, 1st Dec. 2020, ISSN: 0165-6147. DOI: 10.1016/j.tips.2020.09.012.

[77] G. Aichinger, G. Bliem and D. Marko, "Systemically achievable doses of beer flavonoids induce estrogenicity in human endometrial cells and cause synergistic effects with selected pesticides", *Frontiers in Nutrition*, vol. 8, p. 293, 2021, ISSN: 2296-861X. DOI: 10.3389/fnut.2021.691872.

[78] P. D. Wadhwa, C. Buss, S. Entringer and J. M. Swanson, "Developmental origins of health and disease: Brief history of the approach and current focus on epigenetic mechanisms", *Seminars in Reproductive Medicine*, vol. 27, no. 5, pp. 358–368, Sep. 2009, ISSN: 1526-8004, 1526-4564. DOI: 10.1055/s-0029-1237424.

[79] A. D. Stein, A. M. Thompson and A. Waters, "Childhood growth and chronic disease: Evidence from countries undergoing the nutrition transition", *Maternal & Child Nutrition*, vol. 1, no. 3, pp. 177–184, 28th Jun. 2005, ISSN: 1740-8695. DOI: 10.1111/j.1740-8709.2005.00021.x.

[80] T. P. Fleming, A. J. Watkins, M. A. Velazquez, J. C. Mathers, A. M. Prentice, J. Stephenson, M. Barker, R. Saffery, C. S. Yajnik, J. J. Eckert, M. A. Hanson, T. Forrester, P. D. Gluckman and K. M. Godfrey, "Origins of lifetime health around the time of conception: Causes and consequences", *The Lancet*, vol. 391, no. 10132, pp. 1842–1852, 5th May 2018, ISSN: 0140-6736, 1474-547X. DOI: 10.1016/S0140-6736(18)30312-X.

[81] I. Tamayo-Uria, L. Maitre, C. Thomsen, M. J. Nieuwenhuijsen, L. Chatzi, V. Siroux, G. M. Aasvang, L. Agier, S. Andrusaityte, M. Casas, M. de Castro, A. Dedele, L. S. Haug, B. Heude, R. Grazuleviciene, K. B. Gutzkow, N. H. Krog, D. Mason, R. R. C. McEachan, H. M. Meltzer, I. Petraviciene, O. Robinson, T. Roumeliotaki, A. K. Sakhi, J. Urquiza, M. Vafeiadi, D. Waiblinger, C. Warembourg, J. Wright, R. Slama, M. Vrijheid and X. Basagaña, "The early-life exposome: Description and patterns in six european countries", *Environment International*, vol. 123, pp. 189–200, 1st Feb. 2019, ISSN: 0160-4120. DOI: 10.1016/j.envint.2018.11.067.

[82] L. B. A. R. Silva, N. Pinheiro-Castro, G. M. Novaes, G. d. F. L. Pascoal and T. P. Ong, "Bioactive food compounds, epigenetics and chronic disease prevention: Focus on early-life interventions with polyphenols", *Food Research International*, vol. 125, p. 108 646, 1st Nov. 2019, ISSN: 0963-9969. DOI: 10.1016/j.foodres.2019.108646.

[83] M. Vrijheid, R. Slama, O. Robinson, L. Chatzi, M. Coen, d. H. P. van den, C. Thomsen, J. Wright, T. J. Athersuch, N. Avellana, ñ. X. Basaga, C. Brochot, L. Bucchini, M. Bustamante, A. Carracedo, M. Casas, X. Estivill, L. Fairley, G. D. van, J. R. Gonzalez, B. Granum, ž. R. Gra, K. B. Gutzkow, J. Julvez, H. C. Keun, M. Kogevinas, R. R. C. McEachan, H. M. Meltzer, ó. E. Sabid, P. E. Schwarze, V. Siroux, J. Sunyer, E. J. Want, F. Zeman and M. J. Nieuwenhuijsen, "The human early-life exposome (HELIX): Project rationale and design", *Environmental Health Perspectives*, vol. 122, no. 6, pp. 535–544, Jun. 2014, Publisher: Environmental Health Perspectives. DOI: 10.1289/ehp.1307204.

[84] C. P. Wild, "Complementing the genome with an "exposome": The outstanding challenge of environmental exposure measurement in molecular epidemiology", *Cancer Epidemiology, Biomarkers & Prevention*, vol. 14, no. 8, pp. 1847–1850, 15th Aug. 2005, ISSN: 1055-9965. DOI: 10.1158/1055-9965.EPI-05-0456.

[85] G. W. Miller and D. P. Jones, "The nature of nurture: Refining the definition of the exposome", *Toxicological Sciences*, vol. 137, no. 1, pp. 1–2, 1st Jan. 2014, ISSN: 1096-6080. DOI: 10.1093/toxsci/kft251.

[86] C. B. Clish, "Metabolomics: An emerging but powerful tool for precision medicine", *Cold Spring Harbor Molecular Case Studies*, vol. 1, no. 1, a000588, Oct. 2015, ISSN: 2373-2865. DOI: 10.1101/mcs.a000588.

[87] J. Sun, R. Fang, H. Wang, D.-X. Xu, J. Yang, X. Huang, D. Cozzolino, M. Fang and Y. Huang, "A review of environmental metabolism disrupting chemicals and effect biomarkers associating disease risks: Where exposomics meets metabolomics", *Environment International*, vol. 158, p. 106 941, 1st Jan. 2022, ISSN: 0160-4120. DOI: 10.1016/j.envint.2021.106941.

[88] K. I. Ayeni, D. Berry, L. Wisgrill, B. Warth and C. N. Ezekiel, "Early-life chemical exposome and gut microbiome development: African research perspectives within a global environmental health context", *Trends in Microbiology*, vol. 30, no. 11, pp. 1084–1100, 1st Nov. 2022, ISSN: 0966-842X, 1878-4380. DOI: 10.1016/j.tim.2022.05.008.

[89] A. R. Kristal, U. Peters and J. D. Potter, "Is it time to abandon the food frequency questionnaire?", *Cancer Epidemiology, Biomarkers & Prevention*, vol. 14, no. 12, pp. 2826–2828, 19th Dec. 2005, ISSN: 1055-9965. DOI: 10.1158/1055-9965.EPI-12-ED1.

[90] S. M. Rappaport, D. K. Barupal, D. Wishart, P. Vineis and A. Scalbert, "The blood exposome and its role in discovering causes of disease", *Environmental Health Perspectives*, vol. 122, no. 8, pp. 769–774, Aug. 2014. DOI: 10.1289/ehp.1308015.

[91] K. K. Dennis, E. Marder, D. M. Balshaw, Y. Cui, M. A. Lynes, G. J. Patti, S. M. Rappaport, D. T. Shaughnessy, M. Vrijheid and D. B. Barr, "Biomonitoring in the era of the exposome", *Environmental Health Perspectives*, vol. 125, no. 4, pp. 502–510, Apr. 2017, ISSN: 0091-6765. DOI: 10.1289/EHP474.

[92] Y. Fareed, D. Braun, M. Flasch, D. Globisch and B. Warth, "A broad, exposome-type evaluation of xenobiotic phase II biotransformation in human biofluids by LC-MS/MS", *Exposome*, vol. 2, no. 1, osac008, 1st Jan. 2022, ISSN: 2635-2265. DOI: 10.1093/exposome/osac008.

[93] L. Smith, J. Villaret-Cazadamont, S. P. Claus, C. Canlet, H. Guillou, N. J. Cabaton and S. Ellero-Simatos, "Important considerations for sample collection in metabolomics studies with a special focus on applications to liver functions", *Metabolites*, vol. 10, no. 3, p. 104, 12th Mar. 2020, ISSN: 2218-1989. DOI: 10.3390/metabo10030104.

[94] D. Achaintre, A. Gicquiau, L. Li, S. Rinaldi and A. Scalbert, "Quantification of 38 dietary polyphenols in plasma by differential isotope labelling and liquid chromatography electrospray ionization tandem mass spectrometry", *Journal of Chromatography A*, vol. 1558, pp. 50–58, 13th Jul. 2018, ISSN: 0021-9673. DOI: 10.1016/j.chroma.2018.05.017.

[95] D. Achaintre, A. Buleté, C. Cren-Olivé, L. Li, S. Rinaldi and A. Scalbert, "Differential isotope labeling of 38 dietary polyphenols and their quantification in urine by liquid chromatography electrospray ionization tandem mass spectrometry", *Analytical Chemistry*, vol. 88, no. 5, pp. 2637–2644, 1st Mar. 2016, ISSN: 0003-2700. DOI: 10.1021/acs.analchem.5b03609.

[96] R. Zamora-Ros, D. Achaintre, J. A. Rothwell, S. Rinaldi, N. Assi, P. Ferrari, M. Leitzmann, M.-C. Boutron-Ruault, G. Fagherazzi, A. Auffret, T. Kühn, V. Katzke, H. Boeing, A. Trichopoulou, A. Naska, E. Vasilopoulou, D. Palli, S. Grioni, A. Mattiello, R. Tumino, F. Ricceri, N. Slimani, I. Romieu and A. Scalbert, "Urinary excretions of 34 dietary polyphenols and their associations with lifestyle factors in the EPIC cohort study", *Scientific Reports*, vol. 6, no. 1, p. 26 905, 7th Jun. 2016, ISSN: 2045-2322. DOI: 10.1038/srep26905.

[97] J. Ríos, V. Valero-Jara and S. Thomas-Valdés, "Phytochemicals in breast milk and their benefits for infants", *Critical Reviews in Food Science and Nutrition*, vol. 62, no. 25, pp. 6821–6836, 1st Sep. 2022, ISSN: 1040-8398. DOI: 10.1080/10408398.2021.1906627.

[98] M. Á. Ávila-Gálvez, R. García-Villalba, F. Martínez-Díaz, B. Ocaña-Castillo, T. Monedero-Saiz, A. Torrecillas-Sánchez, B. Abellán, A. González-Sarrías and J. C. Espín, "Metabolic profiling of dietary polyphenols and methylxanthines in normal and malignant mammary tissues from breast cancer patients", *Molecular Nutrition & Food Research*, vol. 63, no. 9, p. 1 801 239, 2019, ISSN: 1613-4133. DOI: 10.1002/mnfr.201801239.

[99] I. Alldritt, B. Whitham-Agut, M. Sipin, J. Studholme, A. Trentacoste, J. A. Tripp, M. G. Cappai, P. Ditchfield, T. Devièse, R. E. M. Hedges and J. S. O. McCullagh, "Metabolomics reveals diet-derived plant polyphenols accumulate in physiological bone", *Scientific Reports*, vol. 9, no. 1, pp. 1–12, 29th May 2019, ISSN: 2045-2322. DOI: 10.1038/s41598-019-44390-1.

[100] Y. Zhu, Q. Yin and Y. Yang, "Comprehensive investigation of moringa oleifera from different regions by simultaneous determination of 11 polyphenols using UPLC-ESI-MS/MS", *Molecules*, vol. 25, no. 3, p. 676, Jan. 2020. DOI: 10.3390/molecules25030676.

[101] J. C. Sánchez-Rangel, J. Benavides, J. B. Heredia, L. Cisneros-Zevallos and D. A. Jacobo-Velázquez, "The folin–ciocalteu assay revisited: Improvement of its specificity for total phenolic content determination", *Analytical Methods*, vol. 5, no. 21, pp. 5990–5999, 10th Oct. 2013, ISSN: 1759-9679. DOI: 10.1039/C3AY41125G.

[102] M. Platzer, S. Kiese, T. Herfellner, U. Schweiggert-Weisz and P. Eisner, "How does the phenol structure influence the results of the folin-ciocalteu assay?", *Antioxidants*, vol. 10, no. 5, p. 811, May 2021, ISSN: 2076-3921. DOI: 10.3390/antiox10050811.

[103] Y. Gu, J. T. Peach and B. Warth, "Sample preparation strategies for mass spectrometry analysis in human exposome research: Current status and future perspectives", *TrAC Trends in Analytical Chemistry*, vol. 166, p. 117 151, 1st Sep. 2023, ISSN: 0165-9936. DOI: 10.1016/j.trac.2023.117151.

[104] B. Buszewski and M. Szultka, "Past, present, and future of solid phase extraction: A review", *Critical Reviews in Analytical Chemistry*, vol. 42, no. 3, pp. 198–213, 1st Jul. 2012, ISSN: 1040-8347. DOI: 10.1080/07373937.2011.645413.

[105] O. Coskun, "Separation techniques: Chromatography", *Northern Clinics of Istanbul*, vol. 3, no. 2, pp. 156–160, 11th Nov. 2016, ISSN: 2536-4553. DOI: 10.14744/nci.2016.32757.

[106] P. Froehlich and E. L. Wehry, "Fluorescence detection in liquid and gas chromatography", in *Modern Fluorescence Spectroscopy*, ser. Modern Analytical Chemistry, E. L. Wehry, Ed., Boston, MA: Springer US, 1981, pp. 35–94, ISBN: 978-1-4684-1092-1. DOI: 10.1007/978-1-4684-1092-1_2.

[107] R. P. W. Scott, *Chromatographic Detectors: Design: Function, and Operation* (Chromatographic Science Series). CRC Press, 17th Jul. 1996, vol. 73, 544 pp., ISBN: 978-0-8493-8367-0.

[108] J.-L. Wolfender, G. Marti, A. Thomas and S. Bertrand, "Current approaches and challenges for the metabolite profiling of complex natural extracts", *Journal of Chromatography A*, Editors' Choice IX, vol. 1382, pp. 136–164, 20th Feb. 2015, ISSN: 0021-9673. DOI: 10.1016/j.chroma.2014.10.091.

[109] Y. I. Yashin and A. Y. Yashin, "Chapter 10 - liquid chromatography", in *Chemical Analysis of Food: Techniques and Applications*, Y. Picó, Ed., Boston: Academic Press, 1st Jan. 2012, pp. 285–310, ISBN: 978-0-12-384862-8. DOI: 10.1016/B978-0-12-384862-8.00010-8.

[110] R. Kostiainen and T. J. Kauppila, "Effect of eluent on the ionization process in liquid chromatography–mass spectrometry", *Journal of Chromatography A*, Editors' Choice III, vol. 1216, no. 4, pp. 685–699, 23rd Jan. 2009, ISSN: 0021-9673. DOI: 10.1016/j.chroma.2008.08.095.

[111] R. Veigure, K. Lossmann, M. Hecht, E. Parman, R. Born, I. Leito, K. Herodes and K. Kipper, "Retention of acidic and basic analytes in reversed phase column using fluorinated and novel eluent additives for liquid chromatography-tandem mass spectrometry", *Journal of Chromatography A*, vol. 1613, p. 460 667, 22nd Feb. 2020, ISSN: 0021-9673. DOI: 10.1016/j.chroma.2019.460667.

[112] M. C. García-Alvarez-Coque, J. J. Baeza-Baeza and G. Ramis-Ramos, "Cahpter 8 - reversed phase liquid chromatography", in *Analytical Separation Science*, vol. 1, John Wiley & Sons, Ltd, 2015, pp. 159–198, ISBN: 978-3-527-67812-9.

[113] P. Žuvela, M. Skoczylas, J. Jay Liu, T. Baczek, R. Kaliszan, M. W. Wong and B. Buszewski, "Column characterization and selection systems in reversed-phase high-performance liquid chromatography", *Chemical Reviews*, vol. 119, no. 6, pp. 3674–3729, 27th Mar. 2019, ISSN: 0009-2665. DOI: 10.1021/acs.chemrev.8b00246.

[114] J. W. Dolan, "Temperature selectivity in reversed-phase high performance liquid chromatography", *Journal of Chromatography A*, Application of Theory to the Practice and Understanding of Chromatography, vol. 965, no. 1, pp. 195–205, 2nd Aug. 2002, ISSN: 0021-9673. DOI: 10.1016/S0021-9673(01)01321-8.

[115] "UV-VIS absorption detectors for HPLC", in *Liquid Chromatography Detectors*, ser. Chromatographic Science Series, T. M. Vickrey and R. L. Stevenson, Eds., vol. 23, CRC Press, 28th Oct. 2020, pp. 23–86, ISBN: 978-1-00-011036-4.

[116] Y. Wang, S. Liu, Y. Hu, P. Li and J.-B. Wan, "Current state of the art of mass spectrometry-based metabolomics studies – a review focusing on wide coverage, high throughput and easy identification", *RSC Advances*, vol. 5, no. 96, pp. 78 728–78 737, 15th Sep. 2015, ISSN: 2046-2069. DOI: 10.1039/C5RA14058G.

[117] T. Portolés, E. Pitarch, F. J. López, F. Hernández and W. M. A. Niessen, "Use of soft and hard ionization techniques for elucidation of unknown compounds by gas chromatography/time-of-flight mass spectrometry: Soft and hard ionization for elucidation of unknowns", *Rapid Communications in Mass Spectrometry*, vol. 25, no. 11, pp. 1589–1599, 15th Jun. 2011, ISSN: 09514198. DOI: 10.1002/rcm.5028.

[118] A. P. Bruins, "Mechanistic aspects of electrospray ionization", *Journal of Chromatography A*, vol. 794, no. 1, pp. 345–357, 23rd Jan. 1998, ISSN: 0021-9673. DOI: 10.1016/S0021-9673(97)01110-2.

[119] B. B. Schneider, D. J. Douglas and D. D. Y. Chen, "Ion fragmentation in an electrospray ionization mass spectrometer interface with different gases", *Rapid Communications in Mass Spectrometry*, vol. 15, no. 4, pp. 249–257, 2001, ISSN: 1097-0231. DOI: 10.1002/rcm.221.

[120] B. A. Thomson, "Declustering and fragmentation of protein ions from an electrospray ion source", *Journal of the American Society for Mass Spectrometry*, vol. 8, no. 10, pp. 1053–1058, 1st Oct. 1997, ISSN: 1044-0305. DOI: 10.1016/S1044-0305(97)00129-3.

[121] B. Rochat, "From targeted quantification to untargeted metabolomics: Why LC-high-resolution-MS will become a key instrument in clinical labs", *TrAC Trends in Analytical Chemistry*, vol. 84, pp. 151–164, Nov. 2016, ISSN: 01659936. DOI: 10.1016/j.trac.2016.02.009.

[122] C. Li, S. Chu, S. Tan, X. Yin, Y. Jiang, X. Dai, X. Gong, X. Fang and D. Tian, "Towards higher sensitivity of mass spectrometry: A perspective from the mass analyzers", *Frontiers in Chemistry*, vol. 9, 2021, ISSN: 2296-2646. DOI: 10.3389/fchem.2021.813359.

[123] T. Cajka and O. Fiehn, "Toward merging untargeted and targeted methods in mass spectrometry-based metabolomics and lipidomics", *Analytical Chemistry*, vol. 88, no. 1, pp. 524–545, 5th Jan. 2016, ISSN: 0003-2700. DOI: 10.1021/acs.analchem.5b04491.

[124] B. A. Mamyrin, "Time-of-flight mass spectrometry (concepts, achievements, and prospects)", *International Journal of Mass Spectrometry*, vol. 206, no. 3, pp. 251–266, 22nd Mar. 2001, ISSN: 1387-3806. DOI: 10.1016/S1387-3806(00)00392-4.

[125] J. Ladislas Wiza, "Microchannel plate detectors", *Nuclear Instruments and Methods*, vol. 162, no. 1, pp. 587–601, 1st Jun. 1979, ISSN: 0029-554X. DOI: 10.1016/0029-554X(79)90734-1.

[126] P. E. Miller and M. B. Denton, "The quadrupole mass filter: Basic operating concepts", *Journal of Chemical Education*, vol. 63, no. 7, p. 617, Jul. 1986, ISSN: 0021-9584, 1938-1328. DOI: 10.1021/ed063p617.

[127] Q. Hu, R. J. Noll, H. Li, A. Makarov, M. Hardman and R. Graham Cooks, "The orbitrap: A new mass spectrometer", *Journal of Mass Spectrometry*, vol. 40, no. 4, pp. 430–443, 2005, ISSN: 1096-9888. DOI: 10.1002/jms.856.

[128] A. Makarov, "Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis", *Analytical Chemistry*, vol. 72, no. 6, pp. 1156–1162, 1st Mar. 2000, Publisher: American Chemical Society, ISSN: 0003-2700. DOI: 10.1021/ac991131p.

[129] C. Ho, C. Lam, M. Chan, R. Cheung, L. Law, L. Lit, K. Ng, M. Suen and H. Tai, "Electrospray ionisation mass spectrometry: Principles and clinical applications", *The Clinical Biochemist Reviews*, vol. 24, no. 1, pp. 3–12, Feb. 2003, ISSN: 0159-8090.

[130] L. Lin, H. Lin, M. Zhang, X. Dong, X. Yin, C. Qu and J. Ni, "Types, principle, and characteristics of tandem high-resolution mass spectrometry and its applications", *RSC Advances*, vol. 5, no. 130, pp. 107 623–107 636, 17th Dec. 2015, ISSN: 2046-2069. DOI: 10.1039/C5RA22856E.

[131] T. O. Nicolescu, "Interpretation of mass spectra", in *Mass Spectrometry*, M. Aliofkhazraei, Ed., IntechOpen, 7th Jun. 2017, pp. 23–78, ISBN: 978-953-51-3223-3.

[132] E. de Hoffmann, "Tandem mass spectrometry: A primer", *Journal of Mass Spectrometry*, vol. 31, no. 2, pp. 129–137, 1996, ISSN: 1096-9888. DOI: 10.1002/(SICI)1096-9888(199602)31:2<129::AID-JMS305>3.0.CO;2-T.

[133] C. Ludwig, L. Gillet, G. Rosenberger, S. Amon, B. C. Collins and R. Aebersold, "Data-independent acquisition-based SWATH-MS for quantitative proteomics: A tutorial", *Molecular Systems Biology*, vol. 14, no. 8, e8126, 13th Aug. 2018, ISSN: 1744-4292. DOI: 10.15252/msb.20178126.

[134] J. Zhou and Y. Yin, "Strategies for large-scale targeted metabolomics quantification by liquid chromatography-mass spectrometry", *Analyst*, vol. 141, no. 23, pp. 6362–6373, 14th Nov. 2016, ISSN: 1364-5528. DOI: 10.1039/C6AN01753C.

[135] A. Kruve, R. Rebane, K. Kipper, M.-L. Oldekop, H. Evard, K. Herodes, P. Ravio and I. Leito, "Tutorial review on validation of liquid chromatography–mass spectrometry methods: Part i", *Analytica Chimica Acta*, vol. 870, pp. 29–44, 22nd Apr. 2015, ISSN: 0003-2670. DOI: 10.1016/j.aca.2015.02.017.

[136] B. Rochat, R. Mohamed and P.-E. Sottas, "LC-HRMS metabolomics for untargeted diagnostic screening in clinical laboratories: A feasibility study", *Metabolites*, vol. 8, no. 2, p. 39, Jun. 2018, ISSN: 2218-1989. DOI: 10.3390/metabo8020039.

[137] J. Guo and T. Huan, "Comparison of full-scan, data-dependent, and data-independent acquisition modes in liquid chromatography–mass spectrometry based untargeted metabolomics", *Analytical Chemistry*, vol. 92, no. 12, pp. 8072–8080, 16th Jun. 2020, Publisher: American Chemical Society, ISSN: 0003-2700. DOI: 10.1021/acs.analchem.9b05135.

[138] E. Defossez, J. Bourquin, S. von Reuss, S. Rasmann and G. Glauser, "Eight key rules for successful data-dependent acquisition in mass spectrometry-based metabolomics", *Mass Spectrometry Reviews*, vol. 42, no. 1, pp. 131–143, 18th Jun. 2021, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mas.21715, ISSN: 1098-2787. DOI: 10.1002/mas.21715.

[139] M. Pourchet, L. Debrauwer, J. Klanova, E. J. Price, A. Covaci, N. Caballero-Casero, H. Oberacher, M. Lamoree, A. Damont, F. Fenaille, J. Vlaanderen, J. Meijer, M. Krauss, D. Sarigiannis, R. Barouki, B. Le Bizec and J.-P. Antignac, "Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues", *Environment International*, vol. 139, p. 105 545, 1st Jun. 2020, ISSN: 0160-4120. DOI: 10.1016/j.envint.2020.105545.

[140] B. MacLean, D. M. Tomazela, N. Shulman, M. Chambers, G. L. Finney, B. Frewen, R. Kern, D. L. Tabb, D. C. Liebler and M. J. MacCoss, "Skyline: An open source document editor for creating and analyzing targeted proteomics experiments", *Bioinformatics*, vol. 26, no. 7, pp. 966–968, 1st Apr. 2010, ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btq054.

[141] E. Gorrochategui, J. Jaumot, S. Lacorte and R. Tauler, "Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: Overview and workflow", *TrAC Trends in Analytical Chemistry*, vol. 82, pp. 425–442, 1st Sep. 2016, ISSN: 0165-9936. DOI: 10.1016/j.trac.2016.07.004.

[142] J. Houriet, W. S. Vidar, P. K. Manwill, D. A. Todd and N. B. Cech, "How low can you go? selecting intensity thresholds for untargeted metabolomics data preprocessing", *Analytical Chemistry*, vol. 94, no. 51, pp. 17 964–17 971, 27th Dec. 2022, Publisher: American Chemical Society, ISSN: 0003-2700. DOI: 10.1021/acs.analchem.2c04088.

[143] T. Pluskal, A. Korf, A. Smirnov, R. Schmid, T. Fallon, X. Du and J.-K. Weng, "CHAPTER 7. metabolomics data analysis using MZmine", in *Processing Metabolomics and Proteomics Data with Open Software: A Practical Guide*, ser. New Developments in Mass Spectrometry 8, R. Winkler, Ed., The Royal Society of Chemistry, 16th Mar. 2020, pp. 232–254, ISBN: 978-1-78801-721-3. DOI: 10.1039/9781788019880-00232.

[144] X. Du, A. Smirnov, T. Pluskal, W. Jia and S. Sumner, "Metabolomics data preprocessing using ADAP and MZmine 2", *Methods in Molecular Biology*, vol. 2104, pp. 25–48, 2020, ISSN: 1064-3745. DOI: 10.1007/978-1-0716-0239-3_3.

[145] P. Du, W. A. Kibbe and S. M. Lin, "Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching", *Bioinformatics*, vol. 22, no. 17, pp. 2059–2065, 1st Sep. 2006, ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btl355.

[146] R. Smith, D. Ventura and J. T. Prince, "LC-MS alignment in theory and practice: A comprehensive algorithmic review", *Briefings in Bioinformatics*, vol. 16, no. 1, pp. 104–117, 1st Jan. 2015, ISSN: 1467-5463. DOI: 10.1093/bib/bbt080.

[147] D. Dudzik, C. Barbas-Bernardos, A. García and C. Barbas, "Quality assurance procedures for mass spectrometry untargeted metabolomics. a review", *Journal of Pharmaceutical and Biomedical Analysis*, Review issue 2017, vol. 147, pp. 149–173, 5th Jan. 2018, ISSN: 0731-7085. DOI: 10.1016/j.jpba.2017.07.044.

[148] J. Kuligowski, Á. Sánchez-Illana, D. Sanjuán-Herráez, M. Vento and G. Quintás, "Intra-batch effect correction in liquid chromatography-mass spectrometry using quality control samples and support vector regression (QC-SVRC)", *Analyst*, vol. 140, no. 22, pp. 7810–7817, 26th Oct. 2015, ISSN: 1364-5528. DOI: 10.1039/C5AN01638J.

[149] B. Xi, H. Gu, H. Baniasadi and D. Raftery, "Statistical analysis and modeling of mass spectrometry-based metabolomics data", *Methods in molecular biology (Clifton, N.J.)*, vol. 1198, pp. 333–353, 2014, ISSN: 1064-3745. DOI: 10.1007/978-1-4939-1258-2_22.

[150] I. Blaženović, T. Kind, J. Ji and O. Fiehn, "Software tools and approaches for compound identification of LC-MS/MS data in metabolomics", *Metabolites*, vol. 8, no. 2, p. 31, Jun. 2018, ISSN: 2218-1989. DOI: 10.3390/metabo8020031.

[151] E. L. Schymanski, J. Jeon, R. Gulde, K. Fenner, M. Ruff, H. P. Singer and J. Hollender, "Identifying small molecules via high resolution mass spectrometry: Communicating confidence", *Environmental Science & Technology*, vol. 48, no. 4, pp. 2097–2098, 18th Feb. 2014, ISSN: 0013-936X. DOI: 10.1021/es5002105.

[152] J. P. McCord, L. C. Groff and J. R. Sobus, "Quantitative non-targeted analysis: Bridging the gap between contaminant discovery and risk characterization", *Environment International*, vol. 158, p. 107 011, 1st Jan. 2022, ISSN: 0160-4120. DOI: 10.1016/j.envint.2021.107011.

A. Appendix

A.1. List of Additional Contributions to Peer-Reviewed Publications

Submitted for Publication	Berger, S.; Oesterle, I.; Ayeni, K. I.; Ezekiel, C. N.; Rompel, A.; Warth, B. "Polyphenol exposure of mothers and infants assessed by LC-MS/MS quantification in breast milk", *Submitted for Publication*
Pre-print	Mildau, K.; Ehlers, H.; Oesterle, I.; Pristner, M.; Warth, B. Doppler, M.; Bueschl, C.; Zanghellini, J.; van der Hooft, J. "Tailored mass spectral data exploration using the specXplore interactive dashboard", *bioRxiv*, **2023**, DOI: 10.1101/2023.10.03.560677

A.2. List of Oral Presentations

06/2023	**Human exposome, immune function, and human health**, Stockholm, Sweden Oesterle I., Berger, S., Ayeni, K., Rompel A., Warth B. *Polyphenol Profiling & Non-Targeted Assessment*
05/2023	**JunganalytikerInnen Forum 2023**, Leoben, Austria Oesterle I., Berger, S., Ayeni, K., Rompel A., Warth B. *Polyphenol Biomonitoring in Human Biofluids by LC-(HR)MS(/MS)*
04/2022	**Österreichische Lebensmittelchemikertage**, Online Oesterle I., Rompel A., Warth B. *Novel LC-MS/MS Approaches to Explore Polyphenol Exposure in Humans*
10/2021	**DoSChem Panel C Retreat 2021**, Admont, Austria Oesterle I., Braun D., Rompel A., Warth B. *Targeted LC-MS/MS for the Analysis of Polyphenol Exposure in Humans**
06/2021	**JunganalytikerInnen Forum 2021**, Online Oesterle I., Braun D., Rompel A., Warth B. *Targeted LC-MS/MS for the Analysis of Polyphenol Exposure in Humans**

*The same abstract and oral presentation was presented due to the near time-frame of these events.

A.3. List of Poster Presentations

09/2022	**DoSChem Symposium**, Vienna, Austria Oesterle I., Rompel A., Warth B. *Metabolomic LC-MS Approaches to Study Polyphenol Exposure in Humans**
07/2022	**MassSpec-Forum Vienna**, Vienna, Austria Oesterle I., Pristner, M., Rompel A., Warth B. *Metabolomic LC-MS Approaches to Study Polyphenol Exposure in Humans**
06/2022	**Annual Conference of the Metabolomics Society - Metabolomics 2022**, Valencia, Spain Oesterle I., Pristner, M., Rompel A., Warth B. *Metabolomic LC-MS Approaches to Study Polyphenol Exposure in Humans**
05/2022	**DoSChem Panel C Retreat 2022**, Frankenfels, Austria Oesterle I., Rompel A., Warth B. *Non-targeted LC-HRMS Analysis for Polyphenol in Humans*

*The same abstract and poster was presented due to the near time-frame of these events.

A.4. List of Abbreviations

APCI, Atmospheric Pressure Chemical Ionization; APPI, Atmospheric Pressure Photoionization; CI, Chemical Ionization; CO, Catechol Oxidase; DDA, Data-Dependent Acquisition; DIA, Data-Independent Acquisition; E_k, Kinetic Energy; ECNI, Electron Capture Negative Ionization; EI, Electron Ionization; ESI, Electrospray Ionization; GC, Gas Chromatography; LC, Liquid Chromatography; LC-HRMS, Liquid Chromatography coupled to High-Resolution Mass Spectrometry; LC-MS/MS, Liquid Chromatography coupled to Tandem Mass Spectrometry; m, Mass; MRM, Multiple Reaction Monitoring; MS, Mass Spectrometry; m/z, Mass-to-Charge; NMR, Nuclear Magnetic Resonance; PPO, Polyphenol Oxidase; RP, Reverse-Phase; S/N, Signal-to-Noise; SRM, Selected Reaction Monitoring; TOF, Time-of-Flight; TYR, Tyrosinase; U, Direct Current; UV-Vis, Ultraviolet-Visible Light; v, Velocity; V, Alternating Current; w, Frequency.

A.5. List of Figures

A.6. Overview of Measured LC-(HR)MS(/MS) Sequences and Their Corresponding Raw Data Storage Location

Batch Name	Purpose of Experiment	Methods	Raw Data Storage	Instrument
Pophe analyte tuning	Targeted method development - MS conditions	Automatic tuning with direct injection	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\Tuning	QTrap 6500+
FIA compound optimization	Targeted method development - MS conditions	200703_CompoundOpt_FIA_Mix1_v1	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\FIAOpt_20200703123605	QTrap 6500+
200706_StdsMix_Eluent_v1	Targeted method development - LC conditions	200703_sol_groupA_v1 200703_sol_groupB_v1 200706_QC_PoPhe_v1	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200706_StdsMixs_Eluent	QTrap 6500+
200708_SdsMix_Eluent_v1_1	Targeted method development - LC conditions	200703_sol_groupA_v1_1 (_2, _3, _4, _5, _6, _7, _8 and _9) 200708_sol_groupB_v1_1 (_2, _3, _4, _5, _6, _7, _8 and _9) 200706_QC_PoPhe_v1	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200708_StdsMixs_Eluent_2	QTrap 6500+
200720_SdsMix_Eluent_v2	Targeted method development - LC conditions	200708_sol_groupA_v1_1 (_2, _3, _4, _5, _7, _8 and _9) 200708_sol_groupB_v1_1 (_2, _3, _4, _5, _7, _8 and _9) 200706_QC_PoPhe_v1	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200720_StdsMixs_Eluent_3	QTrap 6500+
200721_SdsMix_Eluent_v3	Targeted method development - LC conditions	20072_sol_groupA_v2_1 (2_2,2_3,2_4, 3_1, 3_2, 3_4 and 3_4) 200721_sol_groupB_v2_1 (2_2,2_3,2_4, 3_1, 3_2, 3_4 and 3_4) 200721_QC_PoPhe_v2_1 (and _2)	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200721_StdsMixs_Eluent_4	QTrap 6500+
200811_SdsMix_Eluent_v1_2	Targeted method development - LC conditions	200811_sol_groupC_v1_1 (_2, _3, _4, _5, _6, _7, _8 and _9) 200811_sol_groupD_v1_1 (_2, _3, _4, _5, _6, _7, _8 and _9) 200706_QC_PoPhe_v1	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200811_StdsMixs_Eluent_1_2	QTrap 6500+
200812_StdsMix_Eluent_v1_3	Targeted method development - LC conditions	200811_sol_groupC_v1_3 (_2, _4, _5, _7, _8 and _9) 200811_sol_groupD_v1_3 (_2, _4, _5, _7, _8 and _9) 200708_sol_groupA_v1_3 (_2, _4, _5, _7, _8 and _9) 200708_sol_groupB_v1_3 (_2, _4, _5, _7, _8 and _9) 200706_QC_PoPhe_v1	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200812_StdsMixs_Eluent_1_3	QTrap 6500+
200820_SdsMix_Eluent_v4	Targeted method development - LC conditions	200820_sol_groupE_v1_1 (_2, and _3) 200820_sol_groupF_v1_1 (_2, and _3) 200820_sol_groupG_v1_1 (_2, and _3) 200706_QC_PoPhe_v1	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200821_StdsMixs_Eluent_4	QTrap 6500+
200820_SdsMix_Eluent_v5	Targeted method development - LC conditions	200828_sol_groupE_v1_4_1 (_4_2, _4_3, _5_1, _5_2, _5_3 and _6_1) 200828_sol_groupF_v1_4_1 (_4_2, _4_3, _5_1, _5_2, _5_3 and _6_1) 200828_sol_groupG_v1_4_1 (_4_2, _4_3, _5_1, _5_2, _5_3 and _6_1) 200706_QC_PoPhe_v1	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200821_StdsMixs_Eluent_4	QTrap 6500+
200901_SdsMix_Eluent_v6	Targeted method development - LC conditions	200828_sol_groupE_v1_4_1 (_4_2, _4_3, _5_1, _5_2, _5_3 and _6_1) 200828_sol_groupF_v1_4_1 (_4_2, _4_3, _5_1, _5_2, _5_3 and _6_1) 200828_sol_groupG_v1_4_1 (_4_2, _4_3, _5_1, _5_2, _5_3 and _6_1) 200706_QC_PoPhe_v1	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200901_StdsMixs_Eluent_6	QTrap 6500+
200911_U_&_S_w_Stds_E1&E2_v1	Targeted method development - LC conditions	200828_sol_groupE_v1_6_1 (and _5_1) 200911_sol_groupE_v1_6_2 (_6_2_E2, _6_3, _6_3_E2, _7_1, and _7_3_E2) 200911_sol_groupF_v1_6_2 (_6_2_E2, _6_3, _6_3_E2, _7_1, and _7_3_E2) 200911_sol_groupG_v1_6_2 (_6_2_E2, _6_3, _6_3_E2, _7_1, and _7_3_E2) 200911_QC_PoPhe_v1_2	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200911_U_&_S_w_Stds_E1&E2	QTrap 6500+
FIA compound optimization at 2 flow rates	Targeted method development - MS conditions	200928_CompoundOpt_FIA_Mix2_v1 200928_CompoundOpt_Mix2_v2	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\FIAOpt_20200928155349	QTrap 6500+
FIA compound optimization at 2 flow rates	Targeted method development - MS conditions	200928_CompoundOpt_FIA_Mix2_v2 200928_CompoundOpt_Mix2_v3	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\FIAOpt_20200928164250	QTrap 6500+
FIA compound optimization at 2 flow rates	Targeted method development - MS conditions	200928_CompoundOpt_FIA_Mix2_v1 200928_CompoundOpt_Mix2_v2	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\FIAOpt_20200929093135	QTrap 6500+
FIA compound optimization at 2 flow rates	Targeted method development - MS conditions	200928_CompoundOpt_FIA_Mix2_v2 200928_CompoundOpt_Mix2_v3	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\FIAOpt_20200929101422	QTrap 6500+
200928_U_&_S_w_Stds_E1&E4_v2	Targeted method development - LC conditions	200928_groupE_v2_1_1 (_1_2, _1_3, _1_4, _2_1, and _2_2) 200928_groupF_v2_1_1 (_1_2, _1_3, _1_4, _2_1, and _2_2) 200928_groupG_v2_1_1 (_1_2, _1_3, _1_4, _2_1, and _2_2) 200922_PoPhe_QC_2	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\200929_U_&_S_w_Stds_E1&E4_v2	QTrap 6500+
201016_U_&_S_w_Stds_E1_v3	Targeted method development - LC conditions	201016_mixVIC_m2_06ml_40C (_30C) 201016_mixE_m2_06ml_40C 201016_mixF_m2_06ml_40C 201016_mixG_m2_06ml_40C _201016_PoPhe_QC_v3	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\201016_U_&_S_w_Stds_E1_v3	QTrap 6500+
201019_U_&_S_w_Stds_E1_v3_1	Targeted method development - LC conditions	201016_mixVIC_m2_06ml_40C 201016_PoPhe_QC_v3	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\201019_U_&_S_w_Stds_E1_v3_1	QTrap 6500+
201022_U_&_S_w_Stds_E1_v3_2	Targeted method development - LC conditions	201022_PoPhe_QC_v4 201022_mixE_m2_06mL_30C_grad1 (and grad2 and grad3) 201022_mixF_m2_06mL_30C_grad1 (and grad2 and grad3) 201022_mixG_m2_06mL_30C_grad1 (and grad2 and grad3)	lmch40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\201022_U_&_S_w_Stds_E1_v3_2	QTrap 6500+

Batch Name	Purpose of Experiment	Methods	Raw Data Storage	Instrument
201022_U_&_S_w_Stds_E1_v3_2	Targeted method development - LLOQ conditions	201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\201022_U_&_S_w_Stds_E1_v3_2	QTrap 6500+
201022_U_&_S_w_Stds_E1_v3_3	Targeted method development - LLOQ conditions	201022_mixVIC_6dmL_30C_grad1 (and grad2 and grad3) 201016_mixVIC_m2_06mL_40C 201016_PoPhe_QC_v3	Data\Projects\2020\PoPhe\Data\201022_U_&_S_w_Stds_E1_v3_3 ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\201023_U_&_S_w_Stds_E1_v3_3	QTrap 6500+
201102_S_w_Stds_E1_v3_epi	Targeted method development - LLOQ conditions	201022_PoPhe_QC_v4 201102_mixE_06mL_30C_grad1_epi_part1 (and part2) 201102_mixE_06mL_30C_grad1_epi 201102_mixG_06mL_30C_grad1_epi_part1 (and part2) 201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\201102_S_w_Stds_E1_v3_epi ImcN40 MSC raw files\QTrap6500+\Analyst	QTrap 6500+
201104_U&S_w_Stds_E1_full_v1	Targeted method development - LLOQ conditions	201104_mixFull_06mL_30C_grad1 201022_mixE_m2_06mL_30C_grad1 201022_PoPhe_QC_v4	Data\Projects\2020\PoPhe\Data\201104_U&S_w_Stds_E1_full_v1	QTrap 6500+
201127_U_S_P_w_Stds_full_v2	Targeted method development - LLOQ conditions	201104_mixFull_06mL_30C_grad1 201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\201127_U_S_P_w_Stds_full_v2	QTrap 6500+
201214_U_S_P_w_Stds_full_v3	Targeted method development - LLOQ conditions	201214_mixFull_06mL_30C_grad1_v2 201214_mixFull_06mL_30C_grad2 201214_Delph_tester_06mL_30C_grad1	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2020\PoPhe\Data\201214_U_S_P_w_Stds_full_v3	QTrap 6500+
210111_U_S_P_w_Stds_Full_v4	Targeted method development - LLOQ conditions	201022_PoPhe_QC_v4 210111_mixFull_06mL_30C_grad2_v2	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210111_U_S_P_w_Stds_full_v4	QTrap 6500+
210120_U_testWash	Targeted method development - LLOQ conditions	210111_mixFull_06mL_30C_grad2_v2 210111_mixFull_06mL_30C_grad2_v2_1	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210120_TestWash	QTrap 6500+
210120_U_testWash	Targeted method development - LLOQ conditions	210122_mixH_grad2_v2_testHighC	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210122_TestConc	QTrap 6500+
210201_TestConc_and_Wash	Targeted method development - LLOQ conditions	210201_mixH_grad2_v2_testHighC_3 210201_mixFull_06mL_30C_grad2_v3 210308_mixH_grad2_v2_testHighC_4 210308_mixFull_06mL_30C_grad2_v3 210308_mixFull_06mL_30C_grad2_v3_Wash 210308_mixFull_06mL_30C_grad2_v3_clean	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210201_TestConc_and_Wash	QTrap 6500+
210308_TestCol_and_WashStep	Targeted method development - LLOQ conditions	210308_mixH_grad2_v2_testHighC_4_1 210308_mixFull_06mL_30C_grad2_v3 210308_mixFull_06mL_30C_grad2_v3_Wash 210308_mixFull_06mL_30C_grad2_v3_clean	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210308_TestCol_and_WashStep	QTrap 6500+
PoPhe analyte tuning	Targeted method development - MS conditions	Automatic tuning with direct injection	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\Tuning	QTrap 6500+
210210_TestNewTune	Targeted method development - MS conditions	210210_mixH_grad2_v2_testHighC_5	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210210_TestNewTune	QTrap 6500+
210215_U_S_P_LODs_and_Saliva	Targeted method development - Trial	210215_mixFull_06mL_30C_grad2_v5 201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210215_U_S_P_LODs_and_Saliva	QTrap 6500+
210223_TestEluent_and_Deconj	Targeted method development - LLOQ conditions and enzymatic deconjugation	210223_mixFull_06mL_30C_grad2_v5_1 210215_mixFull_06mL_30C_grad2_v5	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210223_TestEluent_and_Deconj	QTrap 6500+
210309_U_S_P_methVal_1	Targeted method validation	210309_PoPhe_method_5_MV1 201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210309_U_S_P_methVal_1	QTrap 6500+
210309_U_S_P_methVal_2	Targeted method validation	210403_PoPhe_method_5_MV2 201022_PoPhe_QC_v4 210403_PoPhe_method_5_MV2_Blank	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210403_U_S_P_methVal_2	QTrap 6500+
210309_U_S_P_methVal_2_IntraDay	Targeted method validation	210403_PoPhe_method_5_MV2 201022_PoPhe_QC_v4 210403_PoPhe_method_5_MV2_Blank	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210403_U_S_P_methVal_2_IntraDay	QTrap 6500+
210330_Test_CarryOver	Targeget method carryover measurements	210309_PoPhe_method_5_MV1 210330_PoPhe_method_5_v2_notFounds 210330_PoPhe_Inject_Wash_30s 210330_PoPhe_Inject_Wash_and_Eluent 210330_PoPhe_Inject_Wash_Double	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210330_Test_CarryOver	QTrap 6500+
210403_Test_Carryover_2	Targeget method carryover measurements	210403_PoPhe_method_5_MV1 210403_PoPhe_method_5_MV1_wash_inject	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210403_Test_CarryOver_2	QTrap 6500+
210429_S_T_HRT_M_Saliva_seq1	Bachelors' thesis of Madlen Meyer - Hormone replacement therapy: tests. etc: sample measurement	210403_PoPhe_method_5_MV2 210403_PoPhe_method_5_MV2_Blank 201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210429_S_T_HRT_M_Saliva_seq1	QTrap 6500+
210503_Test_QC	Measure QC's after instrument troubleshooting/cleaning	201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210503_Test_QC	QTrap 6500+
210504_Test_QC_v3	Measure QC's after instrument troubleshooting/cleaning	201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210504_Test_QC	QTrap 6500+
210506_M_seq2	Bachelor' s thesis of Madlen Meyer - Hormone replacement therapy: tests. etc: sample measurement	210403_PoPhe_method_5_MV2 210403_PoPhe_method_5_MV2_Blank 201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210506_M_seq2	QTrap 6500+
210506_Test_QC	Measure QC's after instrument troubleshooting/cleaning	201022_PoPhe_QC_v4	ImcN40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210506_Test_QC	QTrap 6500+

Batch Name	Purpose of Experiment	Methods	Raw Data Storage	Instrument
210517_Test_QC_v4	Measure QCs after instrument troubleshooting/cleaning	201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210517_Test_QC	QTrap 6500+
210520_Test_Don3GA	Measure Don3GA for sending to collaborator	210520_PoPhe_QC_v3_forDON 210520_PoPhe_QC_v3_2_forDON 190429_multi_DON_method_17	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210520_Test_Don3GA	QTrap 6500+
210521_U_S_P_methVal_3_and_PoP	Targeted method validation and pilot study	210521_PoPhe_method_5_MV3 210521_PoPhe_method_5_MV3_Blank 201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210521_U_S_P_methVal_3_and_PoP	QTrap 6500+
210527_Mushrooms	Test mushrooms extraction	210521_PoPhe_method_5_MV3	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\210527_Mushrooms	QTrap 6500+
211015_NigU_EPI_Berry_Mushrooms	Measure Nigerian urine samples, berries for Cornelia Schmutz, and mushroom samples	210521_PoPhe_method_5_MV3 210521_PoPhe_method_5_MV3_Blank 201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\211015_NigU_EPI_Berry_Mushrooms	QTrap 6500+
211015_EPI_seq2	EPI scan of pilot study	210521_PoPhe_method_5_MV3 210521_PoPhe_method_5_MV3_Blank 201022_PoPhe_QC_v4 211015_EPI_meth1 to 5	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2021\PoPhe\Data\211015_NigU_EPI_Berry_Mushrooms	QTrap 6500+
211118_PoPhe	Measurement to compare sensitivity trial	20211118_HSST3_FA_0_6mLmin_Grad_16min_60000_posneg 20211118_HSST3_FA_0_6mLmin_Grad_16min_120000_NEG 20211118_HSST3_FA_0_6mLmin_Grad_16min_120000_POS 20211118_HSST3_FA_0_6mLmin_Grad_16min_MS2_NEG 20211118_HSST3_FA_0_6mLmin_Grad_16min_MS2_POS 20211118_HSST3_FA_0_6mLmin_Grad_16min_wash_B3	lmch\40 MSC raw files\Q_Exactive_HF\2021\211118_PoPhe	QExactive HF
220120_BerrySamples_Diluted	Measure berry samples for Cornelia Schmutz	210521_PoPhe_method_5_MV3 210521_PoPhe_method_5_MV3_Blank 201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\220120_BerrySamples_Diluted	QTrap 6500+
DIA Tests	Trial of a method using DIA on QExactive HF	20220304_HSST3_FA_0_4mLmin_Grad_Abrev_DIA_NEG_v1 to 12	lmch\40 MSC raw files\Q_Exactive_HF\2022\KW-08\PoPhe\DIA Tests	QExactive HF
220308_BerrySamples_Diluted_v2	Remeasurement berry samples for Cornelia Schmutz	220307_PoPhe_final_method 220307_PoPhe_final_method_Blank 201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\220308_BerrySamples_Diluted_v2	QTrap 6500+
220413_Mushrooms_v1	Bachelors' thesis of Iris Neuburger - Measure mushroom samples (trial)	220307_PoPhe_final_method 220307_PoPhe_final_method_Blank 201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\220413_Mushrooms_v1	QTrap 6500+
220510_Mushrooms_all	Bachelors' thesis of Iris Neuburger - Measure mushroom samples	220307_PoPhe_final_method 220307_PoPhe_final_method_Blank 201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\220510_Mushrooms_all	QTrap 6500+
220607_Test_Cal	Measure QCs/standards for instrument troubleshooting/cleaning	201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\220607_Test_Cal	QTrap 6500+
220608_Mushrooms_All_Dilutions	Bachelors' thesis of Iris Neuburger - Measure mushroom samples	220307_PoPhe_final_method 220307_PoPhe_final_method_Blank 201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\220608_Mushrooms_All_Dilutions	QTrap 6500+
220711_PoPhe_Sens_v2	Measurement to compare sensitivity of untargeted method - urine	20220711_HSST3_FA_0_6mLmin_Grad_16min_120000_NEG 20220711_HSST3_FA_0_6mLmin_Grad_16min_120000_POS	lmch\40 MSC raw files\Q_Exactive_HF\2022\KW-27\220711_PoPhe_Sens_v2\Urine	QExactive HF
220711_PoPhe_Sens_v2	Measurement to compare sensitivity of untargeted method - serum	20220711_HSST3_FA_0_6mLmin_Grad_16min_120000_NEG 20220711_HSST3_FA_0_6mLmin_Grad_16min_120000_POS	lmch\40 MSC raw files\Q_Exactive_HF\2022\KW-27\220711_PoPhe_Sens_v2\Serum	QExactive HF
220711_PoPhe_Sens_v2	Measurement to compare sensitivity of untargeted method - plasma	20220711_HSST3_FA_0_6mLmin_Grad_16min_120000_NEG 20220711_HSST3_FA_0_6mLmin_Grad_16min_120000_POS	lmch\40 MSC raw files\Q_Exactive_HF\2022\KW-27\220711_PoPhe_Sens_v2\Plasma	QExactive HF
220713_Top10_DDA_Urine_Samples	Top 10 DDA measurement of pilot study samples	20220713_HSST3_FA_0_6mLmin_Grad_16min_MS2_DDA_Neg_stepNCE 20220713_HSST3_FA_0_6mLmin_Grad_16min_MS2_DDA_Pos_stepNCE 20220713_HSST3_FA_0_6mLmin_Grad_16min_MS2_DDA_Neg_stepNCE_smp1_run1 (20220713 to 20220722, smp 1 to 12, run 1 to 5) 20220713_HSST3_FA_0_6mLmin_Grad_16min_MS2_DDA_Pos_stepNCE_smp1_run1 (20220713 to 20220722, smp 1 to 12, run 1 to 5)	lmch\40 MSC raw files\Q_Exactive_HF\2022\KW-27\220713_Top10_DDA_Urine_Samples	QExactive HF
220722_Test_stds	Measure QCs/standards for instrument troubleshooting/cleaning	201022_PoPhe_QC_v4 220307_PoPhe_final_method_noAS	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\220722_Test_stds	QTrap 6500+
220803_Expomas_Urine_PoPhe	Measure EXPOSOMAS urine samples for polyphenols	201022_PoPhe_QC_v4 220307_PoPhe_final_method 220307_PoPhe_final_method_Blank	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\220803_Expomas_Urine_PoPhe	QTrap 6500+
220930_BM_Test_v1	Masters' thesis of Sabrina Berger - Measure breast milk samples (trial)	220307_PoPhe_final_method 201022_PoPhe_QC_v4	lmch\40 MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\220930_BM_Test_v1	QTrap 6500+
220930_Exposomas_IDA_v1	Measure EXPOSOMAS urine samples with untargeted (trial)	MS: TOF MS NEG, LC: Jamnik MS: TOF IDA NEG, LC: Jamnik MS: TOF MS POS, LC: Jamnik MS: TOF IDA POS, LC: Jamnik	lmch\40 MSC raw files\ZenoTOF\2022\20220930_PoPhe\Data\220930_Exposomas_IDA_v1	ZenoTOF

Batch Name	Purpose of Experiment	Methods	Raw Data Storage	Instrument
221021_BM_Test_v2	Masters' thesis of Sabrina Berger - Measure breast milk samples (troubleshooting)	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221021_BM_Test_v2	QTrap 6500+
221024_BM_Test_v2	Masters' thesis of Sabrina Berger - Measure breast milk samples (troubleshooting)	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221024_BM_Test_v2	QTrap 6500+
221024_QC_Tests	Measure QCs for troubleshooting	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221024_QC_Tests	QTrap 6500+
221025_BM_Test_v2	Masters' thesis of Sabrina Berger - Measure breast milk samples (troubleshooting)	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221025_BM_Test_v2	QTrap 6500+
221031_QC_Test_v1	Measure QCs for troubleshooting	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221031_QC_Test_v1	QTrap 6500+
221110_Mushrooms_All_Dilutions	Measure QCs/standards for instrument troubleshooting/cleaning	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221110_Mushrooms_All_Dilutions	QTrap 6500+
221111_QC_and_STDs_Test_v4	Measure QCs/standards for instrument troubleshooting/cleaning	220307_PoPhe_final_method 201022_PoPhe_QC_v4 220307_PoPhe_final_method_blank	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221111_QC_and_STDs_Test_v4	QTrap 6500+
221124_Contamination_Test	Measure QCs/standards for instrument troubleshooting/cleaning	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221124_Contamination_Test	QTrap 6500+
221129_Thermo_LC_Test	Measure QCs/standards for instrument troubleshooting/cleaning	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221129_Thermo_LC_Test	QTrap 6500+
221130_Solv_Test	Measure QCs/standards for instrument troubleshooting/cleaning	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221130_Solv_Test	QTrap 6500+
221130_Test_newPump	Measure QCs/standards for instrument troubleshooting/cleaning	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221130_Test_newPump	QTrap 6500+
221202_Test_Autosampler	Measure QCs/standards for instrument troubleshooting/cleaning	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221202_Test_Autosampler	QTrap 6500+
221205_Test_Bottles	Measure QCs/standards for instrument troubleshooting/cleaning	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221205_Test_Bottles	QTrap 6500+
221205_BM_Test_v3	Measure QCs/standards for instrument troubleshooting/cleaning	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\221205_BM_Test_v3	QTrap 6500+
20230116_Mushroom_Test_STDs	Measure mushrooms with untargeted (trial)	M5:20230116_DA_Top30_wInd_neg_0.05AcC_DP70_CE30op15, LC: 230116_PoPhe_final_Method M5:20230116 Swath_Neg, LC: 230116_PoPhe_Final_Method	ImchI40:MSC:raw files\ZeroTOF\20230323 PoPhe Optimization\Data\20230116_Mushroom_Test_STDs	Zero TOF
230209_BM_v6	Masters' thesis of Sabrina Berger - Measure breast milk samples (trial)	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2023\PoPhe\Data\230209_BM_v6	QTrap 6500+
230210_BM_v7	Masters' thesis of Sabrina Berger - Measure breast milk samples (troubleshooting)	220307_PoPhe_final_method 201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2023\PoPhe\Data\230210_BM_v7	QTrap 6500+
230210_Test	Measure QCs/standards for instrument troubleshooting/cleaning	201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2023\PoPhe\Data\230210_Test	QTrap 6500+
230213_Test_New_Capillary	Measure QCs/standards for instrument troubleshooting/cleaning	201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2023\PoPhe\Data\230213_Test_New_Capillary	QTrap 6500+
230213_Test_Cal_Curve	Measure QCs/standards for instrument troubleshooting/cleaning	201022_PoPhe_QC_v4	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2023\PoPhe\Data\230213_Test_Cal_Curve	QTrap 6500+
230221_BM_v8	Masters' thesis of Sabrina Berger - Measure breast milk samples (trial)	M5:220221_PoPhe_Final_Method_Transfered_LC: 230221_PoPhe_Final_Method_Transfered M5:220221_PoPhe_Final_Method_Transfered_noRT, LC: 230221_PoPhe_Final_Method_Transfered M5:220221_PoPhe_Final_Method_Transfered_changedRTs, LC: 230221_PoPhe_Final_Method_Transfered	ImchI40:MSC:raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230221_BM_v8	QTrap 7500
230221_BM_v9	Masters' thesis of Sabrina Berger - Measure breast milk samples (trial)	M5:220221_PoPhe_Final_Method_Transfered_changedRTs_v2, LC: 230221_PoPhe_Final_Method_Transfered	ImchI40:MSC:raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230221_BM_v9	QTrap 7500
20230222_Mushrooms	Measure mushroom samples with untargeted	M5:20230222_IDA_Top10_wInd_neg_0.05AcC_DP70_CE30op15_v2, LC: 230222_PoPhe_Final_Method_v2	ImchI40:MSC:raw files\ZeroTOF\20230323 PoPhe Optimization\Data\20230222_Mushrooms	Zero TOF
230227_BM_v9_remeasured	Masters' thesis of Sabrina Berger - Measure breast milk samples (trial)	M5:220221_PoPhe_Final_Method_Transfered_changedRTs_v2, LC: 230221_PoPhe_Final_Method_Transfered	ImchI40:MSC:raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230227_BM_v9_remeasurement	QTrap 7500
230301_BM_v10	Masters' thesis of Sabrina Berger - Measure breast milk samples (trial)	M5:220221_PoPhe_Final_Method_Transfered_changedRTs_v2, LC: 230221_PoPhe_Final_Method_Transfered_SuL	ImchI40:MSC:raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230301_BM_v10	QTrap 7500
230307_BM_v11	Masters' thesis of Sabrina Berger - Measure breast milk samples (trial)	M5:220221_PoPhe_Final_Method_Transfered_changedRTs_v2, LC: 230221_PoPhe_Final_Method_Transfered M5:220221_PoPhe_Final_Method_Transfered_changedRTs_v2, LC: 230221_PoPhe_Final_Method_Transfered_SuL	ImchI40:MSC:raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230307_BM_v11	QTrap 7500
230310_BM_v12	Masters' thesis of Sabrina Berger - Measure breast milk samples (trial)	220307_PoPhe_final_method	ImchI40:MSC:raw files\QTrap6500+\Analyst Data\Projects\2023\PoPhe\Data\230310_BM_v12	QTrap 6500+

Batch Name	Purpose of Experiment	Methods	Raw Data Storage	Instrument
230314_BM_Test_Guaiacol_HydroCinA_Tyrsl	Masters' thesis of Sabrina Berger - Measure breast milk samples (trial)	MS: 230314_PoPhe_Final_Method_Transfered_Guaiacol_HydroCinA_Tyrsl_only, LC: 230221_PoPhe_Final_Method_Transfered_5uL MS: 230314_PoPhe_Final_Method_Transfered_Guaiacol_HydroCinA_Tyrsl_only_noRT LC: 230221_PoPhe_Final_Method_Transfered_5uL	ImcN40:MSC raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230314_BM_Test_Guaiacol_HydroCinA_Tyrsl	QTrap 7500
230321_BM_val_1	Masters' thesis of Sabrina Berger - Breast milk method validation	MS: 230321_PoPhe_Method_BM_Valb, LC: 230221_PoPhe_Final_Method_Transfered_3uL MS: 230321_PoPhe_Method_BM_Valb, LC: 230221_PoPhe_Final_Method_Transfered	ImcN40:MSC raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230321_BM_val_1	QTrap 7500
230321_BM_val_1_rep1	Masters' thesis of Sabrina Berger - Breast milk method validation	MS: 230321_PoPhe_Method_BM_Valb, LC: 230221_PoPhe_Final_Method_Transfered_3uL	ImcN40:MSC raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230321_BM_val_1_intra_rep1	QTrap 7500
230321_BM_val_1_rep2	Masters' thesis of Sabrina Berger - Breast milk method validation	MS: 230321_PoPhe_Method_BM_Valb, LC: 230221_PoPhe_Final_Method_Transfered_3uL	ImcN40:MSC raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230321_BM_val_1_intra_rep2	QTrap 7500
230321_BM_val_1_rep3	Masters' thesis of Sabrina Berger - Breast milk method validation	MS: 230321_PoPhe_Method_BM_Valb, LC: 230221_PoPhe_Final_Method_Transfered_3uL	ImcN40:MSC raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230321_BM_val_1_intra_rep3	QTrap 7500
230324_BM_val_1	Masters' thesis of Sabrina Berger - Measure breast milk samples	220307_PoPhe_final_method	ImcN40:MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\230324_BM_val_1_intra_rep1	QTrap 6500+
230324_BM_val_1	Masters' thesis of Sabrina Berger - Measure breast milk samples	220307_PoPhe_final_method	ImcN40:MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\230324_BM_val_1_intra_rep2	QTrap 6500+
230324_BM_val_1	Masters' thesis of Sabrina Berger - Measure breast milk samples	220307_PoPhe_final_method	ImcN40:MSC raw files\QTrap6500+\Analyst Data\Projects\2022\PoPhe\Data\230324_BM_val_1_intra_rep3	QTrap 6500+
230404_test	Measure QCs/standards for instrument troubleshooting/cleaning	201022_PoPhe_QC_v4	ImcN40:MSC raw files\QTrap6500+\Analyst Data\Projects\2023\PoPhe\Data\230404_test	QTrap 6500+
230418_Feces_Trial_v1	Measure feces samples (trial to test sample preparation)	MS: 230321_PoPhe_Method_BM_Valb, LC: 230221_PoPhe_Final_Method_Transfered_5uL	ImcN40:MSC raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230418_Feces_Trial_v1	QTrap 7500
230421_Nigerian_BM	Measure Nigerian breast milk samples	MS: 20230421_TOFMS_neg_0.254cc_DP70, LC: 230222_PoPhe_Final_Method_v2 MS: 20230421_Swath_Optimized_Neg_10win, LC: 230222_PoPhe_Final_Method_v2	ImcN40:MSC raw files\ZenoTOF\2023\20230421_PoPhe_HBM\Data\230421_Nigerian_BM	ZenoTOF
230424_Nigerian_Feces	Measure Nigerian feces samples	MS: 20230421_TOFMS_neg_0.254cc_DP70, LC: 230222_PoPhe_Final_Method_v2 MS: 20230424_Feces_Swath_Optimized_Neg_10win, LC: 230222_PoPhe_Final_Method_v2	ImcN40:MSC raw files\ZenoTOF\2023\20230421_PoPhe_HBM\Data\230424_Nigerian_Feces	ZenoTOF
230424_BM_val_2	Masters' thesis of Sabrina Berger - Breast milk method validation	MS: 230321_PoPhe_Method_BM_Valb, LC: 230321_PoPhe Final Method Transfered_3uL	ImcN40:MSC raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230424_BM_val_2	QTrap 7500
230516_BM_val_3	Masters' thesis of Sabrina Berger - Breast milk method validation	MS: 230516_PoPhe_Method_BM_Valb, LC: 230516_PoPhe Final_Method Transfered_3uL	ImcN40:MSC raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230516_BM_val_3	QTrap 7500
230606_Nigerian_Feces_Xeno	Measure Nigerian feces samples with Xeno method (Jannik)	MS: 20230606_IDA_Top10_wrrcl_neg_0.054cc_DP80_CE30sp15_Xeno_v1, LC: 230606_Jannik_100percentACNin8_3uL MS: 20230606_IDA_Top10_wrrcl_neg_0.054cc_DP80_xeno, LC: 230606_Jannik_100percentACNin8_3uL MS: 20230606_Feces_Swath_Optimized_Neg_10win_Xeno, LC: 230606_Jannik_100percentACNin8_3uL	ImcN40:MSC raw files\ZenoTOF\2023\20230421_PoPhe_HBM\Data\230606_Nigerian_Feces_Xeno	ZenoTOF
230718_BM_Exposomas_JK	Measure EXPOSOMAS breast milk samples for Julia Kuligowski	MS: 230516_PoPhe_Method_BM_Valb, LC: 230516_PoPhe_Final_Method_Transfered_3uL	ImcN40:MSC raw files\QTrap7500\230720_SCIEX OS Data\2023 PoPhe\Data\230718_BM_Exposomas_JK	QTrap 7500

A.7. Standard Operating Procedures (SOP) of the Developed Workflows

SOP #1 - Extraction of Polyphenols in Human Urine, Serum, and Plasma

Version 1

By Ian Oesterle

09.11.2023

- **This workflow is used for the article**

  Ian Oesterle, Dominik Braun, Annette Rompel, Benedikt Warth. "Quantifying up to 90 polyphenols simultaneously in human bio-fluids by LC-MS/MS", *Analytica Chimica Acta*, **2022**, Vol. 1216:339977

- **General extraction procedure**

  1. Gently thaw a 1 mL aliquot of either urine, serum, or plasma, and vortex briefly to homogenize.
  2. Transfer 40 µL to a 0.5 mL clear micro-reaction tube.
  3. Add 160 µL of extraction solvent (acetonitrile: methanol (1:1) with 1% v/v formic acid).
  4. Briefly vortex.
  5. Sonicate on ice for 10 min.
  6. Place at -20°C for at least 2 h.
  7. Centrifuge at 4°C and 18'000 x g for 10 min.
  8. Transfer 75 µL of the supernatant to a 1.5 mL amber LC glass vial with a 300 µL glass insert.
  9. Add 225 µL dilution solvent (acetonitrile: water (1:2) with 1% v/v formic acid).
  10. Cap the amber HPLC vial.
  11. Analyze by LC-MS directly, ideally, or store at -20°C until analysis.

SOP #2 - Extraction of Polyphenols in Infant Stool

Version 1

By Ian Oesterle

09.11.2023

- **This workflow is used for the article**

  Ian Oesterle, Kolawole I. Ayeni, Chibundu N. Ezekiel, David Berry, Annette Rompel, Benedikt Warth. "Insights into the early-life chemical exposome of Nigerian infants and potential correlations with the developing gut microbiome", *bioRxiv*, **2023**

- **General extraction procedure**
  1. Weigh in around 100 mg of wet infant stool into a 2 mL clear micro-reaction tube.
  2. Dry the sample in a vacuum concentrator at 10°C overnight.
  3. Weigh the dried sample.
  4. Add water at a ratio of 40 µL per 20 mg of dry weight.
  5. Add acetonitrile: methanol (1:1) with 1% v/v formic acid at a ratio of 160 µL per 20 mg of dry weight.
  6. Vortex the sample.
  7. Ultrasonicate for 15 min on ice.
  8. If the pellet is not dissolved, repeat steps 6 and 7 as necessary.
  9. Place the sample at 20°C overnight.
  10. Centrifuge at 4°C and 18'000 x g for 10 min.
  11. Transfer 50 µL of the supernatant into a new 2 mL clear micro-reaction tube.
  12. Dilute by a factor of 10 by adding 450 µL of reconstitution solvent (composed of: 200 µL acetonitrile, 245.5 µL water, and 4.5 µL formic acid).
  13. Filter the sample through a PTFE filter into an amber HPLC vial.
  14. Cap the amber HPLC vial.
  15. Analyze by LC-MS directly, ideally, or store at -20°C until analysis.

SOP #3 - Extraction of Polyphenols in Plants and Mushrooms

Version 1

By Ian Oesterle

09.11.2023

- **This workflow is used for the article**

  Ian Oesterle, Mathias Pretzler, Annette Rompel, Benedikt Warth. "Polyphenolic Profiling of Plants and Edible Mushrooms to Aid Characterizing Polyphenol Oxidase Selectivity", *ChemRxiv*, **2023**

- **General extraction procedure**

  1. Freeze-dry the freshly acquired plant or mushroom samples.
  2. Take roughly 1 g of the freeze-dried sample into a mortar and pestle, and grind into a fine powder.
  3. Weigh 50 mg of the powdered sample into a 2 mL clear micro-reaction tube.
  4. Add 500 µL of extraction solvent (water: acetonitrile: methanol (1:1:1) with 1% v/v formic acid).
  5. Briefly vortex the sample.
  6. Place the sample in a thermoshaker at 75°C and 500 rpm for 1 h.
  7. Centrifuge the sample at 22°C and 18'000 x g for 10 min.
  8. Transfer the supernatant into an amber HPLC vial.
  9. Repeat steps 4 to 7 with the remaining solid sample in the micro-reaction tube, and add the supernatant to the amber HPLC vial.
  10. Cap the amber HPLC vial and briefly vortex it.
  11. Analyze by LC-MS directly, ideally, or store at -20°C until analysis.

SOP #4 - Analysis of Polyphenols by Targeted LC-MS/MS in MRM Mode on a QTrap 6500+

Version 1

By Ian Oesterle

09.11.2023

- **This workflow is used for the article**

  Ian Oesterle, Dominik Braun, Annette Rompel, Benedikt Warth. "Quantifying up to 90 polyphenols simultaneously in human bio-fluids by LC-MS/MS", *Analytica Chimica Acta*, **2022**, Vol. 1216:339977

- **Prepare the standard working mixes**
  - Use the individual stock solutions at the different concentrations to make Mix A (8Z), B (8X), C 8R, D 8Y, E 8S, F 8T, and G 8U (see file: 231109_Targeted_Working_Solutions_Urine_Serum_Plasma_v1.xlsx, located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents).
    - The stock solutions were prepared at different dates (see file: 231109_Stock_Solutions_v1.xlsx, located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents) at concentrations of 1 mg/mL and then serially diluted to make stocks with concentrations of: 100 µg/mL, 10 µg/mL, and 1 µg/mL. All stock solutions were created by diluting the compounds in methanol, and they were continuously stored at -20°C.
    - The stability of the individual stocks will depend on the compound, and the lifetimes can be determined with their respective Certificate of Analysis. Information to attain the Certificates of Analysis for each compound can be found in Table S1 of the relevant article. In addition, it will depend on the type of vial used and the location for storage.
  - With mixes A to G, prepare the seven working solutions by serial dilution as follows:

Working Mix	Quantity and Which Mix	Quantity of Methanol added
Working Mix 7 (highest concentrated)	200 µL of all 7 different mix 8s	100 µL
Working Mix 6	750 µL of working mix 7	750 µL
Working Mix 5	750 µL of working mix 6	750 µL
Working Mix 4	750 µL of working mix 5	750 µL
Working Mix 3	750 µL of working mix 4	750 µL
Working Mix 2	500 µL of working mix 3	1000 µL
Working Mix 1 (lowest concentrated)	500 µL of working mix 2	1000 µL

  - The concentrations of the different analytes in the seven different working mixes is listed in the *Excel*: 231109_Targeted_Working_Solutions_Urine_Serum_Plasma_v1.xlsx, located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents.

- **Sample preparation procedure**
  - For urine, serum, or plasma samples (which this method was originally optimized for), "SOP #1: Extraction of Polyphenols in Urine, Serum, and Plasma" should be followed.

- **Quality control**
  - It is recommended that triplicates at three different spiking levels (low, middle, high) are created along with the unknown samples in order to monitor the sample preparation for the various analytes and check for variances in extraction efficiencies compared to the validated method.
    - If it is not feasible to make all three different spiking levels, then it is suggested to make at least the middle level.
    - If variances in the extraction efficiencies are greater than 20% for individual analytes, it is suggested to investigate the cause of these variances.
  - These samples should be prepared in the same manner as the other unknown samples. However, instead of adding the extraction solvent from the SOP, the samples should be spiked with 52 µL of the working mixes listed below, and the extraction solvent should be adapted accordingly to retain the same ratios of formic acid, acetonitrile, and methanol.
    - Low level: working mix 2
    - Middle level: working mix 5
    - High level: working mix 7

- **Calibration curves**
  - In the best-case scenario, a neat solvent calibration curve and a matrix-matched calibration curve should be created.
    - The matrix-matched calibration curve would allow to account for signal suppression and enhancement effects directly, and it would allow to compare the chromatographic peak shapes in matrix to neat solvent.
    - If no matrix-matched calibration curve can be created, it is recommended to create triplicates of a sample enriched with the standard mixes at least in one concentration (ideally in three different concentrations). The sample enrichment should be done at the end of the sample preparation procedure, similarly to creating the calibrants. This will allow to monitor peak shapes and signal suppression and enhancement effects.
  - For creating the matrix-matched calibration, it is recommended to use a “blank matrix” that is minimally contaminated with polyphenols. During the method validation, “Day 3 Urine Blank” by MF was utilized (stored in the -80°C freezer).
  - For creating the neat solvent calibration, acetonitrile: methanol (1:1) with 1% v/v formic acid should be used instead of biological matrix.
  - The samples for the calibration curve should be prepared following the extraction procedure SOP, but after transferring the supernatant to an amber LC vial, 19.5 µL of the working mixes should be added according to the list below with 205.5 µL of the dilution solvent.
    - Standard blank: pure methanol
    - Standard level 1 (STD1) (lowest concentration): working mix 1
    - Standard level 2 (STD2): working mix 2
    - Standard level 3 (STD3): working mix 3
    - Standard level 4 (STD4): working mix 4
    - Standard level 5 (STD5): working mix 5
    - Standard level 6 (STD6): working mix 6
    - Standard level 7 (STD7) (highest concentration): working mix 7

- **LC-MS/MS Measurements**
  - Once the samples are prepared, the following LC-MS/MS method should be followed:

Instrument	Agilent Infinity 1290 II UHPLC with SCIEX QTrap 6500+

Acquisition method	220307_PoPhe_final_method and 220307_PoPhe_final_method_blank for blank samples to reduce carryover (a copy is located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Acquisition Methods)		
Injection volume	5 µL		
Autosampler temperature	7°C		
Column temperature	30°C		
Column	Acquity UPLC HSS T3 (2.1 x 100 mm, 1.8 µM, Waters) with a VanGuard precolumn (1.8 µM, Waters)		
Flow rate	0.6 mL/min		
Needle wash	Water: methanol: acetonitrile: isopropanol (1:1:1:1) with 1% v/v formic acid		
Eluent A	Water with 0.1% v/v formic acid		
Eluent B	Acetonitrile with 0.1% v/v formic acid		
Gradient	Time (min)	Eluent A (%)	Eluent B (%)
	0	95	5
	2	95	5
	12	37	63
	12.01	5	95
	14	5	95
	14	95	5
	16	95	5
Curtain gas	35 arb		
Ion source gas 1	90 arb		
Ion source gas 2	90 arb		
CAD gas	Medium		
Source temperature	550°C		
Entrance potential	+/- 10 V		
Ion spray voltage	+5500/-4500 V		
MRM Transitions	See **Table S2** in the supplementary information of the article: Oesterle et al. 2022 (*Analytica Chimica Acta)*		

- **Data analysis and reporting**
  - The data analysis software during method validation was SciexOS 2.0q, and the quantitation methods can be found under lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Quantitation Methods.
  - The quality controls should be compared with the neat solvent and matrix-matched calibrations to monitor the sample preparation.
  - The neat solvent and matrix-matched calibration curves should be compared to monitor the influence of the matrix to the analytes investigated.
  - Unknown samples should be deemed positive if the ion ratios between the quantifier and qualifier ions in the unknown samples differ by less than 20% compared to the matrix-matched calibrants. In addition, a retention time deviation of +/- 0.1 min is allowed.

SOP #5 - Untargeted LC-HRMS Workflow with Data-Dependent Acquisition (Iterative Exclusion) and Suspect Screening on a QExactive HF

Version 1

By Ian Oesterle

09.11.2023

- **This workflow is used for the article**

  Ian Oesterle, Manuel Pristner, Sabrina Berger, Mingxun Wang, Vinicius Verri Hernandes, Annette Rompel, Benedikt Warth. "Exposomic Biomonitoring of Polyphenols by Non-Targeted Analysis and Suspect Screening", *Analytical Chemistry*, **2023**, Vol. 95(28):10686-10694

- If using urine, serum, or plasma samples (as in the article), then the samples should be prepared following "SOP #1: Extraction of Polyphenols in Urine, Serum, and Plasma".

- **LC-HRMS Measurements**

Instrument		Thermo-Fisher Vanquish UHPLC with a Thermo-Fisher QExactive HF		
LC parameters	Injection volume	5 µL		
	Autosampler temperature	4°C		
	Column temperature	30°C		
	Column	Acquity UPLC HSS T3 (2.1 x 100 mm, 1.8 µM, Waters) with a VanGuard precolumn (1.8 µM, Waters)		
	Flow rate	0.6 mL/min		
	Needle wash	Water: methanol: acetonitrile: isopropanol (1:1:1:1) with 1% v/v formic acid		
	Eluent A	Water with 0.1% v/v formic acid		
	Eluent B	Acetonitrile with 3% v/v water and 0.1% v/v formic acid		
	Gradient	Time (min)	Eluent A (%)	Eluent B (%)
		0	94.8	5.2
		2	94.8	5.2
		12	34	66
		12.01	2.1	97.9
		14	2.1	97.9
		14	94.8	5.2
		16	94.8	5.2
MS method	Acquisition file	20220713_HSST3_FA_0_6mLmin_Grad_16min_MS2_DDA_Neg_stepNCE (20220713 to 20220722, smp 1 to 12, run 1 to 5, in Neg and Pos) (a copy is located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Acquisition Methods)		
	Acquisition Mode	DDA with top 10 and iterative exclusion list (from IE-Omics, DOI: 10.1007/s13361-017-1608-0) (a copy of the script is located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\R Scripts)		

	Scan range	*m/z* 100 to 1000
	Polarity	Negative and positive separately
	MS^1 resolution	60'000
	MS^2 resolution	30'000
	Normalized collision energy	Stepped of 10, 30, and 50 eV
	AGC of full scan	1E6
	Max. injection time of full scan	100 ms
	AGC of MS^2	5E5
	Max. injection time of MS^2	120 ms
	Isolation window	1.0 *m/z*
	Minimum AGC	4E3
	Dynamic exclusion	10 s

- **Data evaluation**
  1. Convert raw files into MGF files with MSConvert (v3.0.22067) (a copy of the software is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Softwares).
  2. Use MZMine (v3.1.0) for feature extraction and pre-processing with the parameters given in **Table S3** of the supplementary information in Oesterle et al. 2023 (*Analytical Chemistry*).
     a. A copy of the software is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Softwares.
  3. For suspect screening, two online databases were used: PhytoHub (https://phytohub.eu/) and Exposome-Explorer (http://exposome-explorer.iarc.fr/).
     a. An excel file containing all of the entries was created: 231109_Databases_Combined_PhytoHub_ExposomeExplorer_v2.xlsx and is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents
     b. From the excel file of the database, a csv file was created with [M-H]- masses (file: 231109_Databases_Combined_for_Suspect_Screening_v1.csv, located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents). This file can then be imported into MZMine in order to extract the features of interest by MS^1 matching. It is assumed in this case that the analytes will only undergo [M-H]- fragmentation, as it is the likely fragmentation of polyphenols due to their hydroxy groups.
  4. MZMine was then used for spectral library matching with a library curated by MS-DIAL (http://prime.psc.riken.jp/compms/msdial/main.html#MSP) called ESI(-)-MS/MS from Authentic Standards was used.
     a. A copy of the spectral library is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents. For the parameters used, see **Table S3** of the supplementary information in Oesterle et al. 2023 (*Analytical Chemistry*).
  5. The aligned feature list was exported from MZMine into SIRIUS (v5.5.7) for *in silico* fragmentation. See **Table S3** of the supplementary information in Oesterle et al. 2023 (*Analytical Chemistry*) for the parameters selected for *in silico* fragmentation.

   a. A copy of the software is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Softwares.
6. The aligned feature list exported from MZMine or raw acquisition files were imported into MassQL on the GNPS server for sulfate or glucuronide neutral loss queries. See **Table S3** of the supplementary information in Oesterle et al. 2023 (*Analytical Chemistry*) for parameters and queries.
7. Import the aligned feature list exported from MZMine into MetaboAnalyst (v5.0) (https://www.metaboanalyst.ca/) for statistical analysis.
8. Use R (v4.2.0) for Van Krevelen analysis with the exported aligned feature list. The code: VanKrevelen_Analysis.R, can be found under: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\R Scripts.

SOP #6 - Untargeted LC-HRMS Workflow with Data-Independent Acquisition using SWATH and Suspect Screening on a ZenoTOF 7600

Version 1

By Ian Oesterle

09.11.2023

- **This workflow is used for the article**

  Ian Oesterle, Kolawole I. Ayeni, Chibundu N. Ezekiel, David Berry, Annette Rompel, Benedikt Warth. "Insights into the early-life chemical exposome of Nigerian infants and potential correlations with the developing gut microbiome", *bioRxiv*, **2023**

- If using stool samples (as in the article), then the samples should be prepared following "SOP #2: Extraction of Polyphenols in Infant Stool".
  1. Besides preparing the "unknown" samples, several other samples should be prepared:
     - Pooled quality control (QC)
       1. This sample is prepared by mixing together small aliquots of each biological sample. This is used for quality control and assurance of the measurements.
     - QC dilution series
       1. Taking the pooled QC and create several serial dilutions using the same factor each time. This allows to check if features extracted decrease in peak intensity/area as the QC is diluted. If it does not, then the feature is most likely due to background noise.
       2. Example: QC -> 4x diluted QC -> 16x diluted QC -> 64x diluted QC -> 256x diluted QC
     - Neat solvent and QC enriched with reference standards
       1. See **Table A.5** of Supplementary Information from Oesterle et al. 2023 (*bioRxiv*)
       2. This allows for identification of features and for semi-quantification (especially if enriched at multiple concentrations).
     - Process blank
       1. This allows to control for and remove features that arise during sample preparation, e.g. from the solvents used.

- **LC-HRMS measurements**
  1. The batch for each individual biological matrix should be set-up as follows:
     a. The process blank sample, acquired in triplicate measurements.
     b. Neat solvent enriched with reference standards
     c. Several injections of the QC to condition the column, recommended 5-10 injections depending on the matrix used.
     d. QC dilution series samples measured from most diluted to the original (least diluted).
     e. QC enriched with reference standards.
     f. QC measured in triplicate measurements.
     g. Unknown samples with a QC injected between every five samples to check stability and see if there is any drift.
  2. The process blank and QCs for conditioning were acquired with only MS^1 scans. After conditioning the column, a QC was injected with MS^1 scan in order to optimized the SWATH windows using the TIC based on the *Excel* given by SCIEX. All subsequent measurements were measured with the SWATH method.
  3. The LC-HRMS parameters were as follows:

Instrument		Agilent Infinity 1290 II UHPLC with SCIEX ZenoTOF 7600		
Acquisition method		Breast Milk	MS: 20230421_Swath_Optimized_Neg_10win, LC: 230222_PoPhe_Final_Method_v2 (a copy is located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Acquisition Methods)	
		Stool	MS: 20230424_Feces_Swath_Optimized_Neg_10win, LC: 230222_PoPhe_Final_Method_v2 (a copy is located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Acquisition Methods)	
Injection volume		5 µL		
Autosampler temperature		4°C		
Column temperature		30°C		
Column		Acquity UPLC HSS T3 (2.1 x 100 mm, 1.8 µM, Waters) with a VanGuard precolumn (1.8 µM, Waters)		
Flow rate		0.6 mL/min		
Needle wash		Water: methanol: acetonitrile: isopropanol (1:1:1:1) with 1% v/v formic acid		
Eluent A		Water with 0.1% v/v formic acid		
Eluent B		Acetonitrile with 0.1% v/v formic acid		
Gradient		Time (min)	Eluent A (%)	Eluent B (%)
		0	95	5
		2	95	5
		12	37	63
		12.01	5	95
		14	5	95
		14	95	5
		16	95	5
Curtain gas		35 psi		
Ion source gas 1		50 psi		
Ion source gas 2		50 psi		
CAD gas		9 arb		
Source temperature		550°C		
Ion spray voltage		-4500 V		
TOF MS	Scan window	*m/z* 100 to 1000		
	Accumulation time	0.25 s		
	Declustering potential	-70 V		
	Collision energy	-10 V		
TOF MS/MS (with SWATH)	Scan window	*m/z* 100 to 1000		
	Accumulation time	0.05 s		
	Declustering potential	-70 V		

	Start *m/z*	End *m/z*	Collision energy (V)	Collision energy spread (V)
SWATH windows for the Stool samples	99.5	194.5	-15	5
	193.5	262.5	-15	5
	261.5	312.4	-20	10
	311.4	368.2	-20	10
	367.2	417.7	-35	15
	416.7	461.8	-35	15
	460.8	507.3	-35	15
	506.3	581.5	-35	15
	580.5	714.3	-35	15
	713.3	999.6	-35	15
	Start *m/z*	End *m/z*	Collision energy (V)	Collision energy spread (V)
SWATH windows for the Breast Milk samples	99.5	164.4	-15	5
	163.4	217.5	-15	5
	216.5	256.2	-15	5
	255.2	294.4	-20	10
	293.4	329.5	-20	10
	328.5	378.6	-20	10
	377.6	454.2	-35	15
	453.2	562.6	-35	15
	561.6	717.9	-35	15
	716.9	999.6	-35	15

- **Data evaluation**
  1. The raw data files were then converted to ABF files using ABFConverter (a copy of the software is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Softwares).
  2. The files were then processed using MSDial (v4.9) with the parameters given in **Table A.4** of the supplementary information in Oesterle et al. 2023 (*bioRxiv*).
     a. A copy of the software is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Softwares.
     b. Not so stringent parameters were used in order to get more features which are then filtered later by suspect screening (see future steps).
  3. Besides pre-processing, MSDial was also used for spectral library matching. For this, a spectral library made by MSDial (http://prime.psc.riken.jp/compms/msdial/main.html#MSP) called ESI(-)-MS/MS from Authentic Standards was used. The parameters for spectral library matching are given in the identification row of the tables above.
     a. A copy of the spectral library is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents.
     b. For the parameters used, see **Table A.4** of the supplementary information in Oesterle et al. 2023 (*bioRxiv*).

4. A txt file (230428_MSDIAL_PostidentificationSTDs_XenoPoPhe_v3.txt, located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents) was used to match features based on m/z and retention time with known reference standards.
5. After processing with MSDial, the entire alignment file was exported to MSFinder (v3.52) for *in silico* fragmentation (formula and structure elucidation) using "Compound annotation (batch analysis)" for the top 5 hits.
    a. These results were saved and then exported, this gave a file called "Structure result-2077.txt" which was opened in *Excel* and then saved as a csv after the following modifications:
    b. Deleting the column: File path
    c. Converting the column "File name" into three columns which are ID (feature ID), RT (retention time of the feature), and MZ (*m/z* of the feature). For ID, the letters ID were removed.
    d. A copy of the software is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Softwares.
6. After processing with MSDial, the alignment file was exported as "Raw data matrix (Area)". This gave an txt file which was then opened with Excel for modification and then saved as a csv. The file was then modified as follows:
    a. The following columns were deleted: Fill %, MS/MS assigned, Reference RT, Annotation tag, RT matched, m/z matched, MS/MS matched, Comment, Manually modified for quantification, Manually modified for annotation, Isotope tracking parent ID, Isotope tracking weight number, RT similarity, Fragment presence %, Spectrum reference file name, MS1 isotopic spectrum, MS/MS spectrum.
    b. The rows were modified:
        1. The first four rows for the sample results were:
            a. Sample names.
            b. Class -> ProcessBlank (if process blank), Dilution Series (if QC used for dilution series), QC rep (if triplicate QC measurements), QC Spiked (if spiked QCs with reference standards), Sample timepoints (month 1, 6 or 12), QC (if regular QC used for checking stability during entire sequence).
            c. File Type -> blank (if process blank), QC, or sample.
            d. Matrix -> feces or breast milk depending on which results file was being modified.
        2. The first four rows for the feature information were:
            a. name of each parameter -> ex. avg RT, avg m/z, adduct type.
            b. Just NAs.
            c. Just NAs.
            d. Matrix -> feces or breast milk depending on which results file was being modified.
    c. Then new columns were added which calculated the relative standard deviation based on the standard deviation and average area columns given automatically by MSDial.
7. The feature area results file for each matrix and their *in silico* fragmentation results were then imported into R for feature cleaning and suspect screening.
    a. This script cleaned up the features based on different aspects (see script) and filtered features if they had hits with the suspect screening list based on MS^1 mass. For suspect screening, the list of suspects was prepared from two online databases: PhytoHub (https://phytohub.eu/) and Exposome-Explorer (http://exposome-explorer.iarc.fr/)
        1. An excel file containing all of the entries was created: 231109_Databases_Combined_PhytoHub_ExposomeExplorer_v2.xlsx and is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents
        2. From the excel file of the database, a csv file was created with [M-H]- masses (file: 231109_Databases_Combined_for_Suspect_Screening_v1.csv, located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents). This file can then be imported into MZMine in order to extract the features of interest by MS^1 matching. It is

assumed in this case that the analytes will only undergo [M-H]- fragmentation, as it is the likely fragmentation of polyphenols due to their hydroxy groups.

b. The code: 231109_Feature_Cleaning_and_Suspect_Screening_v1.R, can be found under: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\R Scripts.

8. The feature annotation was then manually curated as follows:
   a. Features that had reference standards had the MS^2 of reference standards in the spiked pooled QCs compared to the MS^2 of the feature in unknown samples.
   b. Features that had a spectral library match were checked in MSDial to see if the annotation is correct.
   c. Features that did not have a spectral library match or had a bad spectral library match were then checked with their *in silico* fragmentation results in MSFinder.
   d. Features that did not have a good (or no) spectral library match and that did not have a good (or no) in silico results were simply labelled by their MS^1 match for putative annotation.
9. After feature annotation, several different workflows were or can be applied. For each workflow, it´s recommended to create new excels as they need to have the data prepared differently and this way there is always a "mother" datafile saved with all the information. Different workflows can include:
   a. Preparing the data to upload to MetaboAnalyst (v5.0) (https://www.metaboanalyst.ca/) for statistical analysis.
      1. If volcano plots/ANOVA analysis was used, the results downloaded from MetaboAnalyst giving the p-Values, fold changes, etc. can be matched with the full results file of the features in the following R markdown: 231109_Statistical_Combine_v1.Rmd (located in lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\R Scripts).
   b. Correlation studies with microbiome data.
      1. See 231109_Microbiome_Exposome_Correlations_v1.R located in lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\R Scripts
   c. Annotating the features so that they all have CAS, PubChemIDs, INCHIKEYs, and SMILES.
      1. See 231109_Annotation_MeSH_v1.Rmd file located in lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\R Scripts
   d. Chemical enrichment plots (https://www.nature.com/articles/s41598-017-15231-w) can be made (if INCHIKEYs, PubChemIDs, and SMILES are available for the features, otherwise it doesn't work).
      1. See paper or 231109_ChemRich_Chemical_Class_v1.Rmd file located in lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\R Scripts.
   e. Features annotated with reference standards were quantified using the calibration curves made (QC spiked series) simply in *Excel*. Moreover, these features were then correlated with one another and between the various matrices.
      1. See 231109_Exposure_Correlation_v1.R file located in lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\R Scripts

SOP #7 - Untargeted LC-HRMS Workflow with Data-Dependent Acquisition and Suspect Screening on a ZenoTOF 7600

Version 1

By Ian Oesterle

09.11.2023

- **This workflow is used for the article**
  - Ian Oesterle, Mathias Pretzler, Annette Rompel, Benedikt Warth. "Polyphenolic Profiling of Plants and Edible Mushrooms to Aid Characterizing Polyphenol Oxidase Selectivity", *ChemRxiv*, **2023**

- If using plant or mushrooms samples (as in the article), then the samples should be prepared following "SOP #3: Extraction of Polyphenols in Plants and Mushrooms".
  a. Besides preparing the "unknown" samples, several other samples should be prepared:
     - Pooled quality control (QC)
       a. This sample is prepared by mixing together small aliquots of each biological sample. This is used for quality control and assurance of the measurements.
     - QC dilution series
       a. Taking the pooled QC and create several serial dilutions using the same factor each time. This allows to check if features extracted decrease in peak intensity/area as the QC is diluted. If it does not, then the feature is most likely due to background noise.
       b. Example: QC -> 4x diluted QC -> 16x diluted QC
     - Neat solvent and QC enriched with reference standards
       a. See **Table S2** of Supplementary Information from Oesterle et al. 2023 (*ChemRxiv*)
       b. This allows for identification of features and for semi-quantification (especially if enriched at multiple concentrations).
     - Process blank
       a. This allows to control for and remove features that arise during sample preparation, e.g. from the solvents used.

- **LC-HRMS measurements**
  a. The batch should be set-up as follows:
     i. The process blank sample, acquired in triplicate measurements.
     ii. Neat solvent enriched with reference standards
     iii. Several injections of the QC to condition the column, recommended 5-10 injections depending on the matrix used.
     iv. QC dilution series samples measured from most diluted to the original (least diluted).
     v. QC enriched with reference standards.
     vi. QC measured in triplicate measurements.
     vii. Unknown samples with a QC injected between every five samples to check stability and see if there is any drift.
  b. The LC-HRMS parameters were as follows:

Instrument	Agilent Infinity 1290 II UHPLC with SCIEX ZenoTOF 7600
Acquisition method	MS: 20230222_IDA_Top10_wIncl_neg_0.05Acc_DP70_CE30sp15_v2, LC: 2302222_PoPhe_Final_Method_v2

		(a copy is located at: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Acquisition Methods)		
Injection volume		5 µL		
Autosampler temperature		4°C		
Column temperature		30°C		
Column		Acquity UPLC HSS T3 (2.1 x 100 mm, 1.8 µM, Waters) with a VanGuard precolumn (1.8 µM, Waters)		
Flow rate		0.6 mL/min		
Needle wash		Water: methanol: acetonitrile: isopropanol (1:1:1:1) with 1% v/v formic acid		
Eluent A		Water with 0.1% v/v formic acid		
Eluent B		Acetonitrile with 0.1% v/v formic acid		
Gradient		Time (min)	Eluent A (%)	Eluent B (%)
		0	95	5
		2	95	5
		12	37	63
		12.01	5	95
		14	5	95
		14	95	5
		16	95	5
Curtain gas		35 psi		
Ion source gas 1		50 psi		
Ion source gas 2		50 psi		
CAD gas		9 arb		
Source temperature		550°C		
Ion spray voltage		-4500 V		
TOF MS	Scan window	*m/z* 100 to 1000		
	Accumulation time	0.25 s		
	Declustering potential	-70 V		
	Collision energy	-10 V		
TOF MS/MS (with IDA)	Scan window	*m/z* 100 to 1000		
	Accumulation time	0.05 s		
	Declustering potential	-70 V		
	Collision energy	-30 V with 15 V spread		
	Zeno trap pulsing	On with threshold of 20000 counts		
	IDA Criteria	Top 10 with inclusion list (prepared from entries of PhytoHub and Exposome-Explorer)		
	Minimum threshold IDA acquisition	50 counts		
	Exclusion of candidate ions	6 s and after 1 occurrence		
	Dynamic background substraction	On		

- **Data evaluation**
  1. Convert raw files into MGF files with MSConvert (v3.0.22067) (a copy of the software is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Softwares).
  2. Use MZMine (v3.1.0) for feature extraction and pre-processing with the parameters given in **Table S7** of the supplementary information in Oesterle et al. 2023 (*ChemRxiv*).
     a. A copy of the software is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Softwares.
  3. For suspect screening, two online databases were used: PhytoHub (https://phytohub.eu/) and Exposome-Explorer (http://exposome-explorer.iarc.fr/).
     a. An excel file containing all of the entries was created: 231109_Databases_Combined_PhytoHub_ExposomeExplorer_v2.xlsx and is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents
     b. From the excel file of the database, a csv file was created with [M-H]- masses (file: 231109_Databases_Combined_for_Suspect_Screening_v1.csv, located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents). This file can then be imported into MZMine in order to extract the features of interest by MS^1 matching. It is assumed in this case that the analytes will only undergo [M-H]- fragmentation, as it is the likely fragmentation of polyphenols due to their hydroxy groups.
  4. MZMine was then used for spectral library matching with a library curated by MS-DIAL (http://prime.psc.riken.jp/compms/msdial/main.html#MSP) called ESI(-)-MS/MS from Authentic Standards was used.
     a. A copy of the spectral library is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Documents.
     b. For the parameters used, see **Table S7** of the supplementary information in Oesterle et al. 2023 (*ChemRxiv*).
  5. The aligned feature list was exported from MZMine into SIRIUS (v5.5.7) for *in silico* fragmentation. See **Table S7** of the supplementary information in Oesterle et al. 2023 (*ChemRxiv*) for the parameters selected for *in silico* fragmentation.
     a. A copy of the software is located in: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\Softwares.
  6. The feature results file and the *in silico* fragmentation results were then imported into R for feature cleaning.
     a. This script cleaned up the features based on different aspects (see script).
     b. The code: 231109_Feature_Cleaning_Mushrooms_v1.R, can be found under: lmch\12 AK Benedikt Warth\4 PhD and MSc\01_PhD\05. Ian Oesterle - PoPhe\11. Workflows\R Scripts.