



Review Article

Mass spectrometry-based analysis of gut microbial metabolites of aromatic amino acids

Narumol Jariyasopit^{a,b}, Sakda Khoomrung^{a,b,c,*},¹^a Siriraj Center of Research Excellence in Metabolomics and Systems Biology (SiCORE-MSB), Faculty of Medicine Siriraj Hospital Mahidol University, Bangkok 10700, Thailand^b Siriraj Metabolomics and Phenomics Center, Faculty of Medicine Siriraj Hospital Mahidol University, Bangkok 10700, Thailand^c Department of Biochemistry, Faculty of Medicine Siriraj Hospital Mahidol University, Bangkok 10700, Thailand

ARTICLE INFO

Keywords:

Gut microbial metabolite
Metabolomics
Aromatic amino acid
Mass spectrometry
Quantitative analysis

ABSTRACT

Small molecules derived from gut microbiota have been increasingly investigated to better understand the functional roles of the human gut microbiome. Microbial metabolites of aromatic amino acids (AAA) have been linked to many diseases, such as metabolic disorders, chronic kidney diseases, inflammatory bowel disease, diabetes, and cancer. Important microbial AAA metabolites are often discovered via global metabolite profiling of biological specimens collected from humans or animal models. Subsequent metabolite identity confirmation and absolute quantification using targeted analysis enable comparisons across different studies, which can lead to the establishment of threshold concentrations of potential metabolite biomarkers. Owing to their excellent selectivity and sensitivity, hyphenated mass spectrometry (MS) techniques are often employed to identify and quantify AAA metabolites in various biological matrices. Here, we summarize the developments over the past five years in MS-based methodology for analyzing gut microbiota-derived AAA. Sample preparation, method validation, analytical performance, and statistical methods for correlation analysis are discussed, along with future perspectives.

1. Introduction

Gut microbiota, the predominant microbial community in the human body, is highly diverse and comprises up to 1150 species [1]. Gut microbiota play important roles in symbiotic conditions and disease pathogenesis [2,3]. Investigating gut microbial metabolites (molecules < 1500 Da) [4] can provide deep insights into their connections to human health and diseases. Typically, these metabolites are synthesized by gut microbes themselves or from dietary molecules, or co-metabolized by the host and gut microbes [1].

Among the gut microbial metabolites, amino acids (AA) and their metabolites were found to substantially affect the gut and distant organs, and contribute to disease pathogenesis [5–11]. Aromatic amino

acids (AAA), including phenylalanine (Phe), tryptophan (Trp), and tyrosine (Tyr), are essential nutrients for humans and are naturally produced by plants and microorganisms [12–14]. Humans can obtain Phe and Trp from dietary proteins while Tyr can be produced endogenously from Phe via phenylalanine hydroxylase in the liver [15,16]. After entering the body, AAA are metabolized by host and gut microbiota [13,16,17]. A list of gut microbial AAA metabolites, bacterial species, and enzymes involved in Phe, Tyr, and Trp metabolism has been summarized [16].

The interplay between host and gut microbiota with regard to AAA has been linked to the pathogenesis of metabolic disorders [8], type 2 diabetes (T2D) [11], liver diseases [7], chronic kidney diseases [6,10], and neurological diseases [5,9]. The gut microbial metabolism in AAA

Abbreviations: AA, amino acids; AAA, aromatic amino acids; Phe, phenylalanine; Trp, tryptophan; Tyr, tyrosine; T2D, type 2 diabetes; NMR, nuclear magnetic resonance; MS, mass spectrometry; LC-MS, liquid chromatography-MS; GC-MS, gas chromatography-MS; CE-MS, capillary electrophoresis; DE, diethyl ether; MEPS, microextraction by packed sorbent; SPE, solid-phase extraction; PBS, phosphate-buffered saline; TMS, trimethylsilyl; SQ, single quadrupole; LLOQ, limit of quantification; IT, ion trap; TOF, time-of-flight; CCA, canonical correlation analysis.

* Corresponding author at: Siriraj Center of Research Excellence in Metabolomics and Systems Biology (SiCORE-MSB), Faculty of Medicine Siriraj Hospital Mahidol University, Bangkok 10700, Thailand.

E-mail address: Sakda.kho@mahidol.edu (S. Khoomrung).

¹ <https://orcid.org/0000-0001-9461-8597>

<https://doi.org/10.1016/j.csbj.2023.09.032>

Received 4 May 2023; Received in revised form 24 September 2023; Accepted 24 September 2023

Available online 26 September 2023

2001-0370/© 2023 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

and its association with disease have recently been reviewed by [16].

Functional roles of the gut microbiota have been investigated through gut microbial metabolomics using technologies that have improved over the past decades. Two main analytical platforms, nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS), are often recommended for metabolomic analyses [18]. Because of minimal sample pretreatment requirements, NMR facilitates high-throughput metabolomic analysis, but at the expense of sensitivity. Nonetheless, NMR is suitable for both ionizable and non-ionizable metabolites and can provide definite structures for the concerned metabolites. Alternatively, the compatibility of MS with chromatography systems, together with sample pretreatment, considerably enhances the analytical performance via multi-dimensional separation channels. Typically, the detection limits of MS are approximately 1–2 orders of magnitude lower than those of NMR [19]. This high sensitivity makes it suitable for both untargeted and targeted metabolomic analyses.

Generally, for a given cohort study (human or animal), MS-based untargeted analysis of biological specimens is applied to preliminarily search for potential biomarkers (Fig. 1). Though the potential biomarker identified through the untargeted analysis can be limited by the choice of analytical techniques and availability of mass spectral libraries, it offers overall metabolite profiles without prior knowledge of the metabolites involved in the research questions. The candidate metabolites identified at this stage require further confirmation and quantification with reference standards using a targeted approach [20]. The validity of biomarkers should be confirmed using various types of biological validation, such as animal models, bacterial cultures, and larger cohorts [21, 22]. Absolute quantitative analysis of validated biomarkers allows for comparison across different studies, which eventually leads to the establishment of threshold concentrations and facilitates the translation of potential metabolite biomarkers into clinical practice. This underscores the importance of quantitative metabolomics in bridging scientific outcomes with clinical applications.

However, major challenges in quantitative metabolomics include the availability of authentic reference standards and lack of standardized methods [23]. MS-based quantitative methods are typically optimized according to sample matrices and available MS platforms. Therefore,

different methods can measure similar compounds, posing difficulties in the standardization of metabolomic protocols across different laboratories. Nonetheless, the development of a reliable quantitative MS-based method usually adheres to analytical guidelines that include a series of validation tests: method recovery, inter- and intra-day variations, linearity, specificity, matrix effects, analyte stability, carryover, and determination of the limit of detection and quantification [20].

In response to the increasing interests in gut microbial metabolites, analytical methods for analyzing short-chain fatty acid, bile acids, and methylamines have been widely developed and summarized [24,25]. In contrast, MS-based metabolomic methods for qualitative and quantitative analyses of gut microbial AAA metabolites have not been reviewed. Thus, the scope of this review is to summarize the developments over the past five years in MS-based methodologies for profiling and quantifying gut microbial metabolites of AAA. We discuss the general guidelines of sample preparation, method validation, and analytical performance for quantitative analysis. Finally, we summarize the statistical methods often used for metabolome-microbiome correlation analyses and offer future perspectives.

2. Metabolite profiling and targeted analysis

Table 1 summarizes gut microbial metabolites of AAA identified in various types of biological specimens collected from rats, mice, and humans using gas chromatography – MS (GC-MS) and liquid chromatography – MS (LC-MS).

2.1. GC-MS

2.1.1. Sample preparation

2.1.1.1. Serum. The volume of serum used range between 40 and 200 μL . Serum preparation often begins with protein precipitation using either an organic solvent (e.g., methanol (MeOH)) or an acidified solution (e.g., H_2SO_4) after diluting the serum with water [26–28]. Protein precipitation is a sample pretreatment used to remove proteins without eliminating the metabolites of interest. This process decreases the

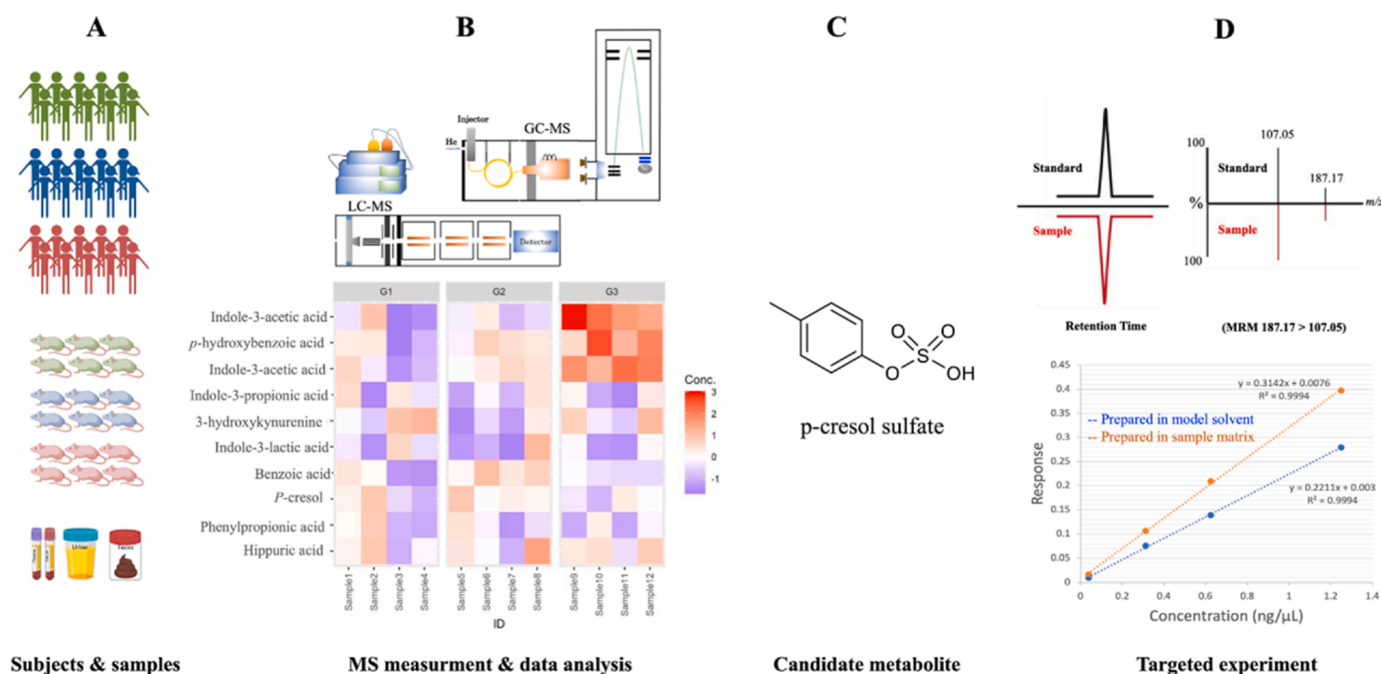


Fig. 1. Overview of workflow for discovery and validation of gut microbial metabolites in clinical research. A: study design and sample collection, B: sample preparation and metabolite profiling via untargeted or semi-targeted analysis using MS-based methods followed by substantial statistical analysis, C: proposed candidate metabolites, D: Targeted analysis to confirm metabolite identity and determine their absolute concentrations.

Table 1
 Summary of gut microbial AAA metabolite analytical methods. ESI modes of analysis are provided in parenthesis. Abbreviations: IT, Ion-trap; TOF, Time-of-flight; Q, quadrupole; SQ, single quadrupole; TRAP, orbitrap; QTRAP, quadrupole and orbitrap.

Tyrosine metabolites	Tryptophan metabolites	Phenylalanine metabolites	Model	Matrix quantified	Analytical Platform	Column Mobile phases	Derivatizing agent	Ref
ferulic acid (-)4-hydroxycinnamic acid (-) 4-hydroxyphenylacetic acid (-) 3-hydroxybenzoic acid (-) 4-hydroxyphenylpropionic acid (-)	indole-3-acetic acid (+)indole-3-aldehyde (-)indole-3-acrylic acid (+) indole-3-propionic acid (+)	benzoic acid (-)hippuric acid (-) phenylpyruvic acid (-)phenyllactic acid (-) phenylacrylic acid (-)phenylpropionic acid (-)phenylacetylglutamine phenylacetic acid (-)	Rat (T2D)	Serum, urine, feces	LC-MSMS (TQ)	Phenyl-Hexyl column (150 × 2.1 mm, 1.7 μm) Solvent A: water 2 mM ammonium formate Solvent B: methanol		Zou et al. 2021
	3-hydroxykynurenine (+) serotonin (+) 5-hydroxytryptophan (+) 5-hydroxyindoleacetic acid (+) tryptamine (+) indole-3-acetic acid (+) indoxyl sulfate (+)indole-3-acetamide (+) indole-3-lactic acid (+) indole-3-carboxaldehyde (+) indole-3-acetic acid (+) tryptophol (+)		Mouse (healthy)	Serum, urine, intestinal contents, liver	UPLC-MS/MS (HRMS, QTRAP)	XB – C18 (150 mm × 2.10 mm, 1.7 μm, 100 Å) Solvent A: 0.5% (v/v) formic acid in water Solvent B: methanol with 0.5% (v/v) formic acid		Lefevre et al. 2019
p-cresol (-)p-cresyl sulfate (-)	indole (+)indoxyl sulfate (-)	p-cresol (-)p-cresyl sulfate (-)	Rat (Crohn's disease)	Serum, Urine	HPLC-MS/MS (TQ)	Phenyl-Hexyl column (50 × 2.1 mm, 1.7 μm) Solvent A: water with 5 mmol/L ammonium formate and 0.05% formic acid Solvent B: methanol		Zeng et al. 2017
p-hydroxybenzoic acid (-)p-hydroxyphenylacetic acid (-)	indole-3-acetic acid (-)indole-3-propionic acid (-)5-hydroxyindoleacetic acid (-)	benzoic acid (-)phenylacetic acid (-) phenylpropionic acid (-)	Rat (high-fat diet-induced obesity)	Serum	HPLC-MS/MS (TQ)	BDS Hypersil C18 column (50 × 2.1 mm, 2.4 μm) Solvent A: 0.05% formic acid in water Solvent B: methanol		Zeng et al. 2019
p-hydroxyphenylpropionic acid (+)p-hydroxybenzoic acid (+)p-hydroxyphenyllactic acid (+) p-hydroxyphenylacetic acid (+) p-ethylphenol (+) p-coumaric acid (+) phenol (+)tyramine (+)p-cresol glucunoride (+) p-cresol (+)p-cresol sulfate (+)		p-cresol (+)p-cresol sulfate (-)	Mouse (high tyrosine diet)	Plasma, urine	UPLC-MSMS (TQ)	BEH C8 (1 mm × 100 mm, 1.7 μm) Solvent A: 0.1% (v/v) formic acid in water Solvent B: 0.1% (v/v) formic acid in acetonitrile: isopropanol (1:1 v/v)	Dansyl chloride	Letertre et al. 2021
	Quantitative analysis indole-3-acetic acid (+) methyl indole-3-acetic acid (+)methyl indole-3-propionic acid (+) Semi-quantitative analysis 5-hydroxyindoleacetic acid (+)5-hydroxy-L-tryptophan (+)indole-3-acetamide (+)tryptophol, indole-3-ethanol (+)indole-3-propionic acid (+) serotonin (+)tryptamine (+)		Human (pregnant women)	Urine	UHPLC-MS/MS (TQ)	C18 CSH (2.1 mm i.d. 100 mm, 1.7 μm) Solvent A: water Solvent B: acetonitrile with 0.1% formic acid		Pavlova et al. 2017

(continued on next page)

Table 1 (continued)

Tyrosine metabolites	Tryptophan metabolites	Phenylalanine metabolites	Model	Matrix quantified	Analytical Platform	Column Mobile phases	Derivatizing agent	Ref
ferulic acid (-)4-hydroxycinnamic acid (-)3-hydroxybenzoic acid (-)4-hydroxybenzoic acid (-)4-hydroxyphenylacetate (-)p-hydroxyphenylpropionic acid (-)		phenylacetic acid (-)phenylpropionic acid (-)hippurate (-)	Human (healthy subjects after the consumption of orange juice)	Urine	UHPLC-MS (TRAP)	C18 (150 × 4.6 mm i.d., 5 μm, 100 Å) 0.1% acidic methanol in 0.1% aqueous formic acid		Ordonez et al. 2018
	indoxyl sulfate	p-cresol sulphatehippuric acidphenylacetylglutamine	Human (end state renal disease)	Serum	UHPLC-MS/MS (TQ)	Discovery HS F5-3 column (150 × 2.1 mm, 3 μm) Mobile phases: not provided.		Wang et al. 2020
	indoxyl sulfate (-)	phenylacetylglutamine (-) hippuric acid (-)	Human (postmenopause female subjects taking a green tea polyphenol supplement)	Feces, urine	UPLC-MS/MS (QTOF)	BEH C18 column Fecal: Solvent A: 0.05% (v/v) aqueous acetic acid with 5 mM ammonium acetate Solvent B: 95% (v/v) aqueous acetonitrile with 0.05% acetic acid and 5 mM ammonium acetate Urine: Solvent A: 0.1% formic acid Solvent B: acetonitrile with 0.1% formic acid		Zhou et al. 2019
4-hydroxyphenylpyruvic acid (-)4-hydroxyphenyllactic acid (-)4-hydroxyphenylacrylic acid (-) 4-hydroxyphenylpropionic acid (-)4-hydroxyphenylacetic acid (-)	indolepyruvic acid (-)indole-3-acetic acid (-)indole-3-acrylic acid (-)indole-3-propionic acid (-)indole-3-lactic acid (-)	phenylpyruvic acid (-)phenyllactic acid (-) phenylacrylic acid (-)phenylpropionic acid (-)phenylacetetic acid (-)	Mouse, Bacterial culture	Serum, cecal contents	LC-MSMS (TQ)	poroshell 120 C18 column (4.6 × 50 mm) Solvent A: 0.1% formic acid (v/v) in water Solvent B: acetonitrile		Dodd et al. 2017
	indole-3-propionic acid (+)indole-3-acetic acid (+) indole-3-carboxaldehyde (indole-3-aldehyde) (+) indole-3-lactic acid (+)indole-3-acetamide (+) indoxyl sulfate (+)tryptamine (+)serotonin (+)5-hydroxytryptophan (+)			Human (healthy)	Plasma (adults), urine (children)	LC-MS/MS (TQ)	ACQUITY HSST3 (1.8 μm, 2.1 × 150 mm) Solvent A: 0.1% formic acid in water Solvent B: acetonitrile with 0.1% formic acid	
ferrulic acid4-hydroxycinnamic acid4-hydroxybenzoic acid3-hydroxybenzoic acid4-hydroxyphenylacetic acid3-(p-hydroxyphenyl) propionate, p-hydroxyphenylpropionic acid		phenylacetic acidhippuric acid	Human (healthy subjects after the consumption of orange juice)	Urine	GC-MS (IT)	ZB-5MS (30 m × 0.25 mm × 0.25 μm film thickness)	MSTFA	Ordonez et al. 2018

(continued on next page)

Table 1 (continued)

Tyrosine metabolites	Tryptophan metabolites	Phenylalanine metabolites	Model	Matrix quantified	Analytical Platform	Column Mobile phases	Derivatizing agent	Ref
4-hydroxycinnamic acid-p-hydroxybenzoic acid-p-hydroxyphenylacetic acid-p-hydroxyphenyllactic acid-p-cresolphenol4-Hydroxyphenylpyruvic acid	indole-3-acetic acidindole-3-propionateindole-3-acrylic acid5-hydroxytryptophanserotonintryptamineindole	phenyllactic acidphenylacetic acidhippuric acidp-cresolphenylethylaminephenylpyruvic acid	Human (healthy), Bacterial cell	Serum, urine, feces, Escherichia coli cell samples	GC-MS (TOF)	DB-5 MS capillary column (30 m × 0.25 mm i.d., 0.25 µm film thickness)	Alkyl chloroformate	Zhao et al. 2017
p-hydroxyphenylacetic acidp-hydroxyphenyllactic acidp-hydroxyphenylpropionic acid4-hydroxybenzoic acid		benzoic acidphenylpropionic acidphenyllactic acid	Human (acute critical ill patients with nosocomial pneumonia, chronically critically ill patients)	Serum, feces	GC-MS (Q)	TR-5MS capillary column (30 m × 0.25 mm, 0.25 µm film thickness)	BSTFA	Chernevskaya et al. 2020
p-hydroxyphenylacetic acidp-hydroxyphenylpropionic acid4-hydroxybenzoic acid4-hydroxyphenyllactic acid		benzoic acidphenylpropionic acidphenyllactic acid	Human (healthy)	Serum	GC-MS (Q)	TR-5 ms quartz capillary column (30 m × 0.25 mm, 0.25 µm film thickness)	BSTFA	Pautova et al. 2018
	indole-3-acetic acidindole-3-propionic acidIndole-3-lactic acid5-hydroxyindole-3-acetic acid		Human (central nervous system disease)	Serum, cerebrospinal fluid	GC-MS (Q)	TR-5MS capillary column (30 m × 0.25 mm, 0.25 µm film thickness)	BSTFA, MTBSTFA	Pautova et al. 2020
4-hydroxyphenylacetic acid	indole-3-acetic acidindole-3-lactic acidindole-3-propionic acid	hippuric acidbenzoic acid phenylethylamine	Rat (Rhubarb-treated)	Feces	GC-MS (TOF)	DB-5 ms capillary column (30 m, 0.25 mm i.d. 0.25 mm 0.25 µm film thickness)	BSTFA	Yin et al. 2017

Table 2
Comparison of lower limit of quantitation (LLOQ) values (nM) of selected gut microbial AAA metabolites analyzed in more than two studies. Abbreviations: IT, Ion-trap; TOF, Time-of-flight; Q, quadrupole; SQ, single quadrupole; TRAP, orbitrap; QTRAP, quadrupole and orbitrap.

Standard solution matrix	GC-MS (IT) [30]	GC-MS (TOF) [26]	GC-MS (SQ) [27]	GC-MS (TOF) [31]	GC-MS (SQ) [9]	GC-MS (TOF) [47]	LC-MS/MS (TQ) [51]	LC-MS/MS (TQ) [11]	LC-MS/MS (TQ) [58]	LC-MS/MS (TQ) [59]	LC-MS/MS (TQ) [48]	LC-MS (TRAP) [30]	LC-MS/MS (TQ) [46]
Hippuric acid	58,036	12,837		125				2260		14	140		38
p-hydroxybenzoic acid	2919	3620	500	1250				284		205	182		25
p-hydroxyphenylacetic acid	5980	3286	500								1324		
Indole-3-acetic acid		65,074		625	400	56	1	123		22	10		
Indole-3-propionic acid		119,970		438	400	3		285		41			
Phenylacetic acid	19,833	3672						397		57		5924	
Phenylpropionic acid	9987		500					719		416		1006	
5-hydroxyindole acetic acid						20	11			10			
Indoxyl sulfate						366	8						240

interference from unwanted species and extends the lifetime of the column and mass detector. However, protein precipitation can impair the reproducibility or accuracy of analytical methods if the metabolites of interest have high affinities for proteins [29]. Subsequent to protein precipitation, simple extraction with MeOH or diethyl ether (DE) using liquid-liquid extraction or centrifugation is often recommended. However, Pautova et al. extracted indolic acids, which are gut microbial metabolites of Trp, from serum and cerebrospinal fluid without protein precipitation using microextraction by packed sorbent (MEPS) [9]. MEPS is a solid-phase extraction (SPE) technique that is used for small-volume extraction and purification. Unlike SPE columns, MEPS is reusable and can be fully automated followed by GC-MS or liquid chromatography (LC)-MS. This technique reduces the serum matrix effect on indolic acid analysis. However, it adds additional steps to the conditions, including washing and drying of the MEPS column, which requires additional development and validation.

2.1.1.2. Urine. Urinary AAA metabolites can be extracted by SPE using reversed-phase SPE cartridges after acidification of urine with phosphoric acid [30]. Another study lyophilized urine samples and reconstituted them in 1 M sodium hydroxide (NaOH) and MeOH or ethanol before derivatization [26]. The same study also reported that lyophilization of urine samples could enhance the intensities of target metabolites across the chromatogram and increase the number of identified metabolites.

2.1.1.3. Feces. Owing to the heterogeneity of fecal samples, the sample pretreatment protocol is important, and the amount of sample must be optimized to avoid column overloading and detector saturation. Zhao et al. homogenized 10 mg of freeze-dried human fecal samples with NaOH solution before centrifugation [26]. After collecting the supernatant, the sample was re-extracted with cold MeOH and the two extracts were combined. Yin et al. [31] performed a series of experiments to optimize the amount of rat feces and extraction solvent combinations. Multivariate analysis of fecal metabolite profiles suggested that a higher proportion of MeOH in the MeOH/chloroform extraction mixture was better for gut microbial metabolites of AAA, but a lower proportion of MeOH was more suitable for metabolites in the tricarboxylic cycle. However, a combination of MeOH, chloroform, and water (225:75:300, v/v/v) was chosen in their study because it is suitable for a wide range of metabolites; however, this could compromise the sensitivity of this method.

2.1.1.4. Bacterial cultures. Following untargeted or metabolite profiling analyses, key AAA metabolites identified in human samples can be validated using bacterial cultures to determine their roles in metabolic pathways. To accurately measure intracellular metabolites from bacteria or other microorganisms, samples must be quenched to prevent metabolite turnover and minimize leakage of intracellular metabolites [32,33]. The most common quenching methods for bacterial metabolomics are cold MeOH solution, cold glycerol-saline solution, and liquid nitrogen [34]. Most of previous quenching methods were optimized for untargeted metabolomic analysis, not specifically for the analysis of AAA metabolites. However, a fast-quenching method using liquid nitrogen in combination with a filter-based sampling was shown to be suitable for intracellular amino acid quantification in bacteria [35]. After collecting bacterial cells and removing the culture medium, Zhao et al. quenched the bacterial cells by washing with phosphate-buffered saline (PBS) before reconstituting them in PBS [26]. Cell lysates were homogenized with water and extracted with cold MeOH via centrifugation. Although this high-throughput GC-MS based method was able to measure AAA metabolites in serum, urine, and feces, the metabolites were not detected in the bacterial cells used in the study.

Typically, polar and nonvolatile metabolites, such as AA, free fatty acids, and AAA microbial metabolites, require chemical derivatization

to modify their structures for GC-MS analysis. Silylation-based derivatization reagents such as *N,O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA), *N*-Methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA), and *N*-Methyl-*N*-tert-butyl dimethylsilyl trifluoroacetamide (MTBSTFA) are often used [36]. The silylation reaction replaces the labile hydrogen (i.e., -COOH, -OH, -NH, NH₂, and -SH) with a trimethylsilyl (TMS) moiety, producing less polar and more volatile TMS derivatives. Based on the previous literature and our experience, the important drawbacks of silylation using TMS reagents are as follows: I) TMS reagents and derivatives are sensitive to moisture, II) TMS derivatives are stable for a short period of time, and III) compounds containing more than one labile hydrogen can form multiple TMS derivatives. Owing to these drawbacks, MTBSTFA is sometimes preferred, because it produces tert-butyl dimethylsilyl (TBDMS) derivatives that are more stable and less sensitive to water [37]. Although TBDMS derivatives are more stable, the formation of TBDMS silylation requires a longer incubation time than TMS silylation. Moreover, owing to steric hindrance, the bulky TBDMS moiety can cause the incomplete derivatization of large molecules such as sugars [38].

Most studies on gut microbial AAA metabolites used TMS to derivatize the analytes [27,28,30]. Pautova et al. reported that the TMS derivatization of indolic acids yielded slightly better and more consistent recoveries than TBDMS derivatization [9]. As a result, TMS was chosen because of its high reproducibility and good recovery of most indolic acids in the model solution. However, TMS derivatization of a few compounds yielded low recoveries ($\leq 50\%$) in sample matrices: 5-hydroxyindole-3-acetic acid in pooled cerebrospinal fluid and serum, indole-3-propionic acid and indole-3-lactic acid in serum. Another derivatization method involves alkyl chloroformate using reagents such as methyl chloroformate and ethyl chloroformate [39]. With regard to AAA, Zhao et al. preferred alkyl chloroformate derivatization because of its faster and milder reaction, producing more stable derivatives within six days [26].

For absolute quantification, the matrix effect must be evaluated to ensure that the measured metabolite concentrations represent actual concentrations. The matrix effect can be minimized using several approaches, including sample dilution [40], standard addition, and matrix-matched calibration [41], and compensated using isotope-labeled internal standards [42]. Pautova et al. evaluated the matrix effect on eight targeted phenylcarboxylic acids and reported a matrix effect ranging between 5% and 30% in serum; therefore, matrix-matched calibration solutions were prepared in the serum to account for the matrix effect [27].

Because GC-MS analysis typically uses electron ionization, a hard ionization technique in which the ionization efficiency is less likely to be affected by the presence of co-eluting molecules [43], the matrix effect evaluation is sometimes overlooked when validating GC-MS methods. Although sample preparation for GC analysis can inherently reduce the matrix effect by sample dilution or can be accounted for by the addition of isotope-labeled internal standards, the inclusion of matrix effect evaluation in GC-MS development should be encouraged to ascertain the accuracy of the quantitative method.

2.1.2. GC-MS and analytical performance

The chromatographic separation of AAA metabolites often relies on reverse-phase chromatography using a capillary column coated with a polysiloxane backbone modified with methyl and 5% phenylarylene as the stationary phase. A single quadrupole (SQ) MS instrument appears to be adequate for the quantitative analysis of gut microbial AAA metabolites in biological matrices, with a lower limit of quantification (LLOQ) of 400–500 nM for the measurement of phenylcarboxylic acids and indole metabolites in serum and cerebrospinal fluid (Table 2) [9,27]. Other low-resolution MS techniques, including ion trap (IT) and time-of-flight (TOF), have also been employed to quantify a comprehensive set of host gut microbiota metabolites, including AAA metabolites [26,31]. To assess the sensitivity of the methods, we compared the

LLOQ values of metabolites measured in at least three studies. The SQ method outperformed the IT and TOF methods (Table 2). Overall, the dynamic range of the GC-MS methods ranged from one to two orders of magnitude but reached three orders of magnitude for some compounds using the TOF method (Table 2).

Recently, two-dimensional GC-MS (GC \times GC-MS) has been shown to significantly increase the peak capacity, number of metabolite detections, and metabolite identification accuracy [44]; however, compared to conventional GC-MS, it has not been widely used to quantify AAA metabolites.

2.2. LC-MS

Untargeted analysis of gut microbial metabolites often employs LC-MS [10,17,45], which provides better metabolite coverage than GC-MS or capillary electrophoresis (CE)-MS, particularly for semi-volatile and non-volatile metabolites. Nonetheless, LC-MS can be biased by sample preparation, column selection, or ionization modes because no single condition can provide a complete picture of the metabolome of a given sample. For metabolite separation, a nonpolar column is often used to screen AAA metabolites [10,17,45]. Wang et al. used two non-polar columns with different bonded phases, a C8 column for positive (ESI⁺) and a C18 column for negative (ESI⁻) ionization modes. The study yielded 6600 features (unidentified metabolites) in serum samples, of which 180 were identified using an in-house database [10]. Consequently, four gut microbial metabolites of AAA, including *p*-cresol, indoxyl sulfate, hippuric acid, and phenylacetyl glycine, were quantified using another reversed phase, pentafluorophenylpropyl, which provides different retention and selectivity compared to the C18 and phenyl columns.

2.2.1. Sample preparation

2.2.1.1. Serum/plasma. Starting volumes of serum/plasma and urine range between 20 and 100 μ L and 10–150 μ L, respectively. Such variations in sample volume depend on the existing concentrations of targeted gut microbial AAA in the biological fluids. Similar to the sample preparation for GC-MS analysis, preparing serum/plasma and urine for LC-MS analysis begins with protein precipitation using MeOH, followed by acetonitrile (ACN) or a mixture of ACN and isopropanol [46]. Anesi et al. precipitated proteins and removed phospholipids from human plasma using commercial Ostro 96-well plates and ACN containing 1% formic acid [47]. Dodd et al. used 6% aqueous sulfosalicylic acid to precipitate proteins from serum samples, followed by incubation at room temperature for 5 min [17].

2.2.1.2. Urine. Lyophilization appears unnecessary for urine samples; only one study used freeze-dried urine samples before extracting the urine residue with MeOH [48]. Simple extraction methods involving vortexing and centrifugation are commonly used after protein precipitation. After extraction, the supernatant was either dried and reconstituted in solvents or directly analyzed using LC-MS. Ordonez et al. developed an LC-MS method to analyze a wide range of urinary microbial metabolites produced after orange juice consumption [30]. The study compared the recoveries of targeted metabolites derived from simple urine extraction (i.e., dilution and centrifugation) with those derived from the SPE method. With regard to gut microbial AAA metabolites, the simple extraction method provided a performance comparable to that of the SPE method, suggesting that sample purification may not be necessary.

2.2.1.3. Solid samples. Solid samples such as feces, intestinal contents, and tissue samples require additional sample preparation steps to homogenize and prevent the transformation of metabolites after sample collection (i.e., microbial fermentation and oxidation). Homogenization

and cold storage are recommended to reduce sampling bias and preserve the original structures and concentrations of the metabolites [49]. Lyophilization is a common pretreatment for removing or controlling the water content of solid samples, and is a crucial step in absolute quantification. Fecal extraction often uses a mixture of MeOH and water with varying mixing ratios or 50% aqueous ACN [50]. Brief vortexing and centrifugation seemed adequate for extracting AAA metabolites from the fecal matrix. Alternatively, Dodd et al. extracted cecal content with ACN using a bead beater homogenizer to disrupt microbial cells and tissues before centrifugation [17]. For tissue samples, sample pretreatment includes snap freezing or lyophilization, homogenization, followed by centrifugation [51,52]. Pure MeOH, or a mixture of MeOH and water, is commonly used to extract AAA metabolites from tissue samples.

2.2.1.4. Bacterial cultures. Sample preparation for the bacterial culture supernatant was similar to that for the biological fluids. Dodd et al. precipitated proteins from a bacterial culture supernatant using MeOH, which was then incubated at room temperature, followed by centrifugation, evaporation of the supernatant to dryness, and reconstitution in organic solvents and 0.5% BSA [17].

Many LC-MS methods do not require derivatization for analyzing gut microbial AAA, thereby making them suitable for high-throughput analysis. However, a quantitative analysis of Tyr host microbial metabolites in plasma and urine developed by Letertre et al. adopted derivatization using a mixture of ammonium hydroxide and dansyl chloride solutions [46]. Dansylation is commonly used for the fluorescent detection of AA and the conversion of phenols and amines to dansylated derivatives that are easily ionizable by LC-MS [53]. Although dansylation helps improve the detection of phenols, the stability of the derivatized samples is a concern when stored for longer than one day [46]. Moreover, dansyl chloride can react violently with water, releasing toxic fumes of hydrogen chloride [54]. These issues must be considered when using this method.

2.2.2. LC Column

Reverse-phase chromatography based on a C18 column [55] is often employed to analyze a wide range of metabolites, including AAA. A phenyl-hexyl column, which contains silica particles bonded to a phenyl-hexyl group, is another frequently used column. The presence of the phenyl group promotes pi-pi interactions with analytes containing an aromatic moiety, providing different retention selectivity for aromatic compounds compared with the C18 column [56]. Indeed, the phenyl-hexyl column was proven to provide better chromatographic separation and optimal retention for most of the target metabolites, especially Trp metabolites, when compared to the C18 and amide columns [11]. Hydrophilic interaction liquid chromatography (HILIC) is also used to separate polar metabolites that may not be well-resolved on a C18 column [57].

The mobile phase for reversed-phase LC generally consists of water and an organic eluent, usually MeOH or ACN. For the analysis of gut microbial AAA metabolites, a combination of organic solvents (MeOH, formic acid in MeOH, or formic acid in ACN) and basic or acidic aqueous solutions is commonly used. Although ACN generates a lower back pressure than MeOH, it provides a higher elution strength than MeOH, which may not resolve highly polar compounds such as indole derivatives [58]. In the same study, nine AAA metabolites were eluted during a stable 30:70 (v/v) water: MeOH gradient with the total run time of 12 min [59]. On the other hand, Tyr and Trp metabolites were eluted at lower mixing ratio of water: ACN, ~ 25–50% [46,48]. Zeng et al. evaluated the chromatographic performance of different aqueous eluents containing formic acid, ammonium formate, or both at different mixing ratios [59]. They reported that 0.05% formic acid yielded better peak shapes and optimal retention times for the analysis of nine microbially derived AAA metabolites (Table 1) in the ESI⁻ mode.

2.2.3. Mass spectrometry

LC-MS/MS, particularly triple quadrupole (TQ) MS, has become the standard platform for quantitative analysis owing to its excellent sensitivity, selectivity, and wide dynamic range. Most LC-MS quantitative methods have been developed using TQ in the multiple reaction monitoring (MRM) mode (MRM-TQ) (Table 1). Dodd et al. used LC-Q-TOF for global analysis and LC-MRM-TQ in the ESI⁻ for quantification [17]. To date, only one study has performed MRM analysis using LC coupled with high-resolution tandem MS, quadrupole and orbitrap (Q-OT), to quantify Trp and its microbial indole derivatives [51].

To increase metabolite coverage, a combination of ESI⁺ and ESI⁻ analyses is frequently used to profile gut microbial metabolites. Typically, the ESI⁻ analysis mode is preferred for Tyr metabolites, but ESI⁺ was only used by Letertre et al. who performed dansyl chloride derivatization [46]. The study reported an enhanced sensitivity for most phenolic compounds except for *p*-hydroxyphenylpyruvic acid and *p*-cresol sulfate. As a result, *p*-hydroxyphenylpyruvic acid was not quantified, and *p*-cresol sulfate was analyzed in ESI⁻ instead [46]. Most of Trp metabolites are analyzed in ESI⁺, but ESI⁻ is also applicable for indole-3-acetic acid, indolepropionic acid, 5-hydroxyindoleacetic acid, indoxyl sulfate, indolepyruvic acid, indoleacetic acid, indoleacrylic acid, and indolepropionic acid [17,45,50,58,59]. Phe metabolites appeared to be favorably ionized in ESI⁻. A study by Zeng et al. (2019) targeting nine gut microbial AAA metabolites in serum samples (Table 1) reported higher intensities and stabilities of their deprotonated ions [59].

The structures of AAA gut microbial metabolites that contain a carboxylic acid group allow them to readily ionize in negative ion mode forming deprotonated species. For the molecules that can ionize in both ESI modes, mode selection should take into consideration the quality of quantitative signals, e.g., chromatographic peak shape and signal-to-noise ratio, as well as other target analytes needed to be quantified in the same run.

2.2.4. Analytical method performance

The convenience of sample pretreatment associated with the LC-MS methodology comes at the cost of sensitivity and column lifetime because of the overload of highly complex matrices. Because ESI is a competitive process, target analytes can suffer from the presence of multiple components in the sample matrix that are co-eluted from the column. Therefore, the evaluation of the matrix effect is crucial for developing and validating LC-MS methods. The magnitude of the matrix effect was highly dependent on the types of samples and analytical protocols used for sample preparation. AAA metabolites have been shown to exhibit varying degrees of matrix effect. A comparison of the slopes derived from the neat solvent and matrix-matched calibration curves suggested that methyl indole-3-acetate would have suffered greatly from the matrix effect if the neat solvent calibration curve had been used while the other two metabolites (indole-3-acetate and methyl indole-3-propionate) exhibited minimal matrix effects [48]. Lefèvre et al. reported poor recoveries of indole derivatives in solid samples collected from mice [51]; four known microbial metabolites of Trp including indole-3-carboxaldehyde, indole-3-acetic acid, indole-3-lactic acid, and tryptophol yielded low recoveries ranging between 48% and 60% in the intestinal matrix, indicating a strong matrix effect. Their serum and urine recoveries were satisfactory (100 ± 20%), but marginal in the liver. In the same study, indoxyl sulfate, another known microbial metabolite of Trp, exhibited good recovery from all matrices (serum, urine, liver, and intestinal contents). The authors used isotope-labeled internal standards to correct for matrix effects.

In contrast, acceptable matrix effect levels have been reported for some AAA metabolites. For example, Anesi et al. observed minimal matrix effects of 80–120% for metabolites in Trp and Tyr metabolites and other gut microbial metabolites in human plasma and urine [47]. The good recoveries of 20 Phe, Tyr, and Trp metabolites from serum, urine, and fecal matrices justified the use of calibration curves derived from the neat solvent in a study by Zou et al. [11]. Two studies by Zeng

et al. also reported no significant matrix effects in rat serum, urine, or feces [58,59].

Because LOD values were not available in many studies, we compared the LLOQ values of the selected AAA measured in at least three studies to assess the sensitivity of the method (Table 2). The LLOQ can be dependent on the biological matrices in which the analytes are quantified; however, most studies reported LLOQ values derived from calibration curves prepared from solvents. The high-resolution MRM-Q-OT method showed 1–2 orders of magnitude higher sensitivity than the MRM-TQ methods. The superior performance of high-resolution MRM-Q-OT analysis is in line with the results of our recent study that compared the analytical performance of MRM-TQ and high-resolution MRM-QTOF in a fungal matrix [60]. Our study demonstrated that high-resolution MRM-QTOF is more accurate in identifying and quantifying trace levels of triterpenoids in a complex matrix. Collectively, these results highlight the superior selectivity and sensitivity of alternative platforms, such as high-resolution Q-OT and high-resolution MRM-QTOF, compared with conventional MRM-TQ.

Ordóñez et al. evaluated two quantitative GC-MS and LC-MS methods for analyzing microbial-derived phenolic acids and aromatic metabolites in human urine [30]. They reported that LC-MS provided higher coverage of microbial metabolites, especially phenolic sulfate and glucuronide metabolites, and lower LOD and LLOQ values [30]. However, LC was coupled to a high-resolution Orbitrap MS, while GC was coupled with IT-MS. The different types of MS could partly contribute to the different analytical performances.

2.3. Capillary electrophoresis-mass spectrometry

The CE-MS technique is suitable for polar and charged metabolites and can be used as a complementary method to LC-MS and GC-MS analyses [61]. Although CE-MS is an attractive tool for metabolomics analysis because of its ability to simultaneously analyze the cationic and anionic forms of metabolites, the applications of CE-MS in metabolomics research are still limited compared to those of NMR, LC-MS, and GC-MS [62]. A known disadvantage of CE-MS is its low sensitivity compared to GC-MS and LC-MS; however, technical developments have been made to improve instrumental sensitivity over the years [63]. Few studies have employed CE-MS to analyze gut microbial metabolites. Most of the currently used CE-MS methods were developed and validated in the early 2000 s [64–66]. For example, Soga et al. developed a CE-Q-MS method and applied it to obtain the global profiles of AA and other polar metabolites from *Bacillus subtilis* extracts [66]. The samples were extracted using a mixture of MeOH and chloroform, and the MeOH fraction was precipitated. Cationic metabolites such as AA, amines, and nucleosides were analyzed in positive mode on a fused silica capillary (50 μm i.d. \times 100 cm), where formic acid was used as the electrolyte and ammonium acetate MeOH-water was used as the sheath liquid. Overall, the method detected 1692 features, with 352 metabolites identified using authentic standards. The analysis time for each run was approximately 16 h. In 2006, Soga et al. modified a previous protocol using TOF-MS. This method has been used to quantify AA and other metabolites in mouse liver tissue and serum samples [67]. Accurate mass measurement and sensitivity provided using CE-TOFMS were improved by supplementing the sheath solution with reserpine (mass recalibration standard). This allowed the identification of all the targeted metabolites (338 metabolites) with a mass tolerance of less than 10 ppm. This method exhibited good reproducibility and linearity. One of the main factors contributing to the improved analytical performance is the TOF instrument used in this study, which provides a faster scanning rate than a quadrupole. In 2017, Mishima et al. showed a good example of the use of CE-TOFMS to analyze gut microbial metabolites [68]. This study used a CE-TOFMS method originally developed by Soga et al. in 2003 [66] to quantify gut microbial metabolites in plasma, feces, and urine samples from adenine-induced renal failure and control mice under germ-free or specific pathogen-free conditions. The metabolites from plasma, urine,

and fecal samples were extracted using MeOH, similar to the sample preparation for the GC-MS and LC-MS analyses. Two gut-derived microbial AAA, indoxyl sulfate and *p*-cresyl sulfate, were detected among the 183 metabolites identified in this study.

3. Linking metabolites to gut microbiota

Identification of the microbial origin of metabolites or associations between gut microbes and their metabolites is frequently accomplished using a multi-method approach that includes cohort studies, animal experiments, fecal microbiota transplantation, in vitro studies, bacterial cultures, and genomics. Particularly for AAA metabolites, in 2017, Dodd and colleagues used LC-MS/MS metabolite profiling in combination with bacterial culture, genetic data, and a mouse model to characterize the metabolic pathways of AAA metabolites from *Clostridium sporogenes*. The team discovered that the reductive AAA pathway in *C. sporogenes* could generate 12 AAA metabolites, nine of which are known circulating metabolites in the host serum [17]. In 2018, Koh et al. identified imidazole propionate as a microbially produced metabolite that may contribute to the development of insulin resistance. The authors performed an untargeted analysis using LC-MS/MS and identified four plasma metabolites, dopamine sulfate, glutamate, imidazole propionate, and N-acetylputrescine, that were significantly elevated in subjects with T2D compared to those in BMI-matched subjects without T2D. Cross-validation of the elevated metabolites in the mouse model showed that only imidazole propionate exhibited higher plasma concentrations in conventionalized mice than in germ-free mice. To validate these results, the authors quantified plasma imidazole propionate in a second cohort of 649 individuals. Combining the results from an in vitro experiment (gut simulator), bacterial culture, metagenomic data, animal experiments, and human cohorts, the authors concluded that imidazole propionate is a microbially produced metabolite that may contribute to the development of T2D [69]. In 2017, Saito et al. screened over hundred bacterial strains in culture media fortified with Tyr and other microbial intermediate metabolites to identify phenol-producing strains [70]. Although the validation of microbial metabolites using a mouse model and bacterial culture provided substantial evidence of microbial metabolite production, the experiment was costly and time-consuming [71]. In addition, whether the experimental results can be translated into humans remains questionable.

Alternatively, metabolome-microbiome integrative analysis using advanced statistical and bioinformatic tools is a crucial step in identifying gut microbes and host interplay, which can unravel the functional diversity of the gut microbiome. A summary workflow of the data input and handling processes for the metabolome and microbiome before correlation analysis is shown in Fig. 2. While metabolomic data are expressed as relative abundance (semi-quantitative analysis) or absolute concentrations (quantitative analysis) of detected metabolites, gut microbiota data can be investigated using various proxies, such as the relative abundance of Operational Taxonomic Units, read count, gut microbiome diversity, metagenomic cluster, or microbial function. One of the challenges regarding microbiome-metabolome correlation analysis is the different characteristics of the data. Microbiome data are unique due to their high dimensionality, compositionality, and sparsity, which should be considered when applying statistical analysis [72]. Microbe-metabolite correlations can be explored using classical univariate analysis such as parametric relationships (Pearson's correlation) [73–75] and non-parametric tests (Spearman's correlation) [10,76,77]. For example, Wang and colleagues characterized the gut microbiome and determined the associations between gut microbiome composition and uremic toxin metabolites (serum and fecal) in patients with end-stage renal disease and in a rat model [10]. In this study, Spearman's correlation method was used to analyze the correlation coefficients between the functional modules of the gut microbiota and serum or fecal metabolites, whereas the random forest model was used to evaluate the correlation between the concentration of each

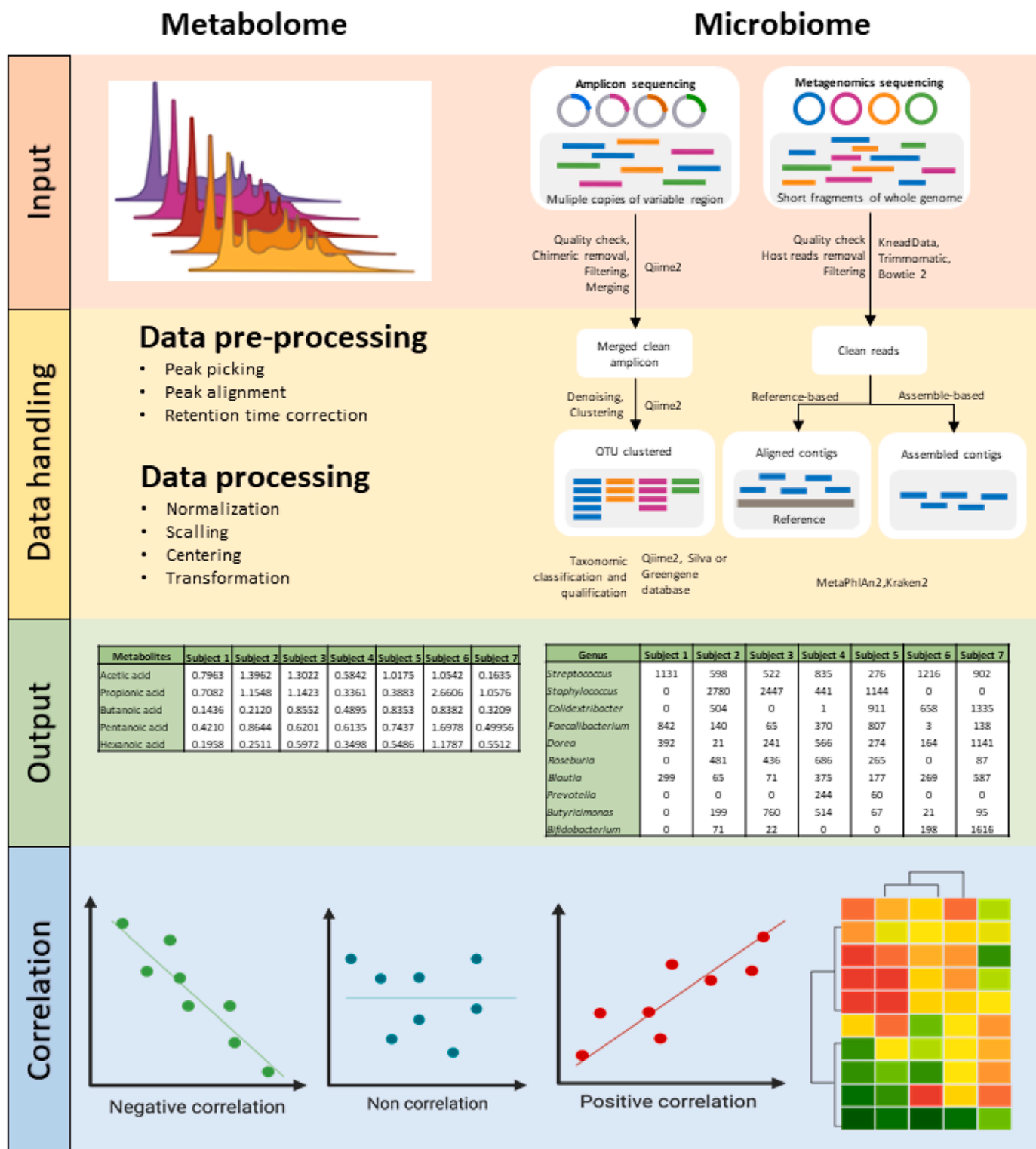


Fig. 2. Summary workflow of data input and data handling of metabolome and microbiome before correlation analysis.

metabolite (in serum or feces) and abundance of microbial species containing metabolite synthetase-encoding genes. Using these models, the authors identified a group of microbial species associated with uremic toxin and secondary bile acid metabolite production.

Multivariate analysis is predominantly used in the context of microbiome-metabolome investigations. Previously used multivariate analysis methods include sparse principal component analysis (PCA) [78], Co-Inertial Analysis (CIA) [79–81], Canonical correlation analysis

(CCA) [78,79], sparse CCA [82,83], two-way orthogonal partial least squares (O2-PLS) [84,85], and Procrustes analysis [86]. Compared to univariate analysis, multivariate analysis reduces the false discovery rate and considers the relationships between and within data matrices [86]. Liu et al. used the CIA and CCA methods to determine the associations between metagenomic linkage groups and the serum metabolome corresponding to lean and obese groups of Chinese individuals [79].

In addition to statistical correlation methods, a number of emerging bioinformatics tools enabling metabolome-microbiome or multi-omics integrative analyses have been recently developed. Incorporating multiple methods, a generalized correlation analysis for the Metabolome and Microbiome (GRaMM) was designed specifically for microbe-metabolite correlation detection. GRaMM provides data preprocessing for metabolomic and microbiome data before identifying linear and nonlinear correlations with confounder adjustment and *p*-value correction [87]. A neural network model, microbe-metabolite vectors (mmvec) deduces a microbe-metabolite interaction from the co-occurrence probability of a microbe-metabolite pair [88]. Another neural network-based model, the Microbiome-Metabolome Network (MiMeNet), uses paired microbiome and metabolomic datasets to train the model, which is then used to predict the comprehensive metabolomic profile of a given microbiome sample [89]. These newly developed tools are independent of the existing knowledge regarding microbial and host metabolic pathways, making them suitable for exploratory studies. Another approach is to exploit existing knowledge of microbial and host metabolic pathways to predict metabolic profiles based solely on microbiome sequencing data, which is beyond the scope of this review. Tools for predicting the metabolite profiles from microbiome sequencing data were evaluated and discussed [90].

You et al. evaluated six widely used correlation detection methods for integrative analysis of the metabolome microbiome: Pearson's correlation, Spearman's correlation, Sparse Correlations for Compositional data (SparCC), correlation inference for compositional data through Lasso (CCLasso), Mutual Information Coefficient (MIC), and cosine similarity methods [91]. They demonstrated better performance using Spearman's correlation and MIC methods. The latter is an information-based non-parametric correlation method that measures linear and non-linear associations between variables in large datasets [92] and is also embedded in GRaMM.

4. Future perspectives

Owing to growing interest in the human gut microbiome and its functions, gut microbial AAA metabolites have been increasingly investigated. MS is a powerful tool not only for identifying gut microbial metabolites, but also for quantitatively analyzing complex biological samples. Over the past decade, one of the major advancements in metabolomics analysis using MS has been the addition of ion mobility spectrometry (IMS) to MS platforms. This combination has been proven to enhance the accuracy of compound identification using a rotationally averaged collision cross-section (CCS), a molecular descriptor that is derived from the time the gaseous ion travels through the drift tube and is independent of the matrices. IMS can potentially assist in novel gut microbial metabolite discovery, especially with the introduction of high-resolution cyclic IMS to enhance the separation capacity [93]. The development of CCS databases and prediction models has supported the incorporation of IMS into conventional MS-based metabolomic workflows. Current CCS databases are associated with a wide range of compound classes, but not with gut microbial metabolites. Therefore, the development of a CCS database of known and unknown gut microbial metabolites can be beneficial for large-scale profiling of metabolites and aid in novel metabolite identification. However, the use of IMS in quantitative metabolomics is limited. This is mainly due to the robustness and convenience of conventional techniques, such as LC-TQ.

While novel gut microbial metabolite discovery often exploits the power of cutting-edge mass spectrometry technology and advanced data analysis tools, quantitative analysis of known AAA microbial metabolites can be accomplished using simpler but robust measurement platforms, either LC or GC coupled with a low-resolution MS, such as TQ, Q, or TOF, which provides excellent sensitivity. In particular to the known AAA microbial metabolites, thoroughly validated quantitative methods for various biological matrices are available for both GC-MS and LC-MS platforms. But neither inter-laboratory comparison nor certified

reference material has been completed for these metabolites. This concern will need to be addressed to support the implementation of these metabolites in clinical setting as accumulated evidence has demonstrated significant connections of these metabolites to multiple disease pathogenesis. Without certified reference materials, the best effort to evaluate the accuracy of a developed method is to determine recoveries of target analytes spiked in a biological matrix. Such recoveries represent the efficiency of the analytical method in recovering the free form of the analytes from the particular matrix. This may not be suitable for determining recoveries of matrix-bound analytes because the spike-in experiment may not completely mimic the actual interactions between analytes and matrix. The development of stool and urine reference materials has been initiated by the National Institute of Standards and Technology [94]. Once available, the adoption of reference materials should be encouraged, in addition to the general validation tests, to ensure the quality and accuracy of quantitative data.

CRediT authorship contribution statement

NJ: Conceptualization, write the original draft, edit the manuscript.
SK: Conceptualization, write the original draft, edit the manuscript.

Declaration of Competing Interest'

None.

Acknowledgments

This project is funded by National Research Council of Thailand (NRCT) and Mahidol University: N42A650370. S.K. was supported by Mahidol University, Bangkok, Thailand (grant no. (IO) R016420001. This research received funding from the National Science, Research, and Innovation Fund (NSRF) via the Program Management Unit for Human Resources & Institutional Development, Research, and Innovation, Grant no. B36G660007. This project is partially supported by the Research Excellence Development (RED) program, Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand. We acknowledge Dr. Kassaporn Duangkumpha, Dr. Krittima Anekthanakul, Weerawan Manokasemsan, Patcha Pungsombat, Suphitcha Limjiasahapong, and Ammarin In-on for their constructive discussions during the preparation of the manuscript.

References

- [1] Postler TS, Ghosh S. Understanding the holobiont: how microbial metabolites affect human health and shape the immune system. *Cell Metab* 2017;26:110–30.
- [2] McCarville JL, Chen GY, Cuevas VD, et al. Microbiota metabolites in health and disease. *Annu Rev Immunol* 2020;38:147–70.
- [3] Fan Y, Pedersen O. Gut microbiota in human metabolic health and disease. *Nat Rev Microbiol* 2021;19:55–71.
- [4] Wishart DS. Metabolomics for investigating physiological and pathophysiological processes. *Physiol Rev* 2019;99:1819–75.
- [5] Rothhammer V, Mascanfroni ID, Bunse L, et al. Type I interferons and microbial metabolites of tryptophan modulate astrocyte activity and central nervous system inflammation via the aryl hydrocarbon receptor. *Nat Med* 2016;22:586–97.
- [6] Kikuchi M, Ueno M, Itoh Y, et al. Uremic toxin-producing gut microbiota in rats with chronic kidney disease. *Nephron* 2017;135:51–60.
- [7] Wang J, Wang Y, Zhang X, et al. Gut microbial dysbiosis is associated with altered hepatic functions and serum metabolites in chronic hepatitis B patients. *Front Microbiol* 2017;8:2222.
- [8] Arnoriaga-Rodríguez M, Mayneris-Perxachs J, Burokas A, et al. Obesity impairs short-term and working memory through gut microbial metabolism of aromatic amino acids. *Cell Metab* 2020;32:548–60. e547.
- [9] Pautova A, Khesina Z, Getsina M, et al. Determination of tryptophan metabolites in serum and cerebrospinal fluid samples using microextraction by packed sorbent, silylation and GC-MS detection. *Molecules* 2020;25:3258.
- [10] Wang X, Yang S, Li S, et al. Aberrant gut microbiota alters host metabolome and impacts renal failure in humans and rodents. *Gut* 2020;69:2131–42.
- [11] Zou B, Sun Y, Xu Z, et al. Rapid simultaneous determination of gut microbial phenylalanine, tyrosine, and tryptophan metabolites in rat serum, urine, and faeces using LC-MS/MS and its application to a type 2 diabetes mellitus study. *Biomed Chromatogr* 2021;35:e4985.

- [12] Tzin V, Galili G. New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. *Mol Plant* 2010;3:956–72.
- [13] Maeda H, Dudareva N. The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annu Rev Plant Biol* 2012;63:73–105.
- [14] Rodriguez A, Martinez JA, Flores N, et al. Engineering *Escherichia coli* to overproduce aromatic amino acids and derived compounds. *Microb Cell Fact* 2014;13(1):15.
- [15] Fitzpatrick PF. Tetrahydropterin-dependent amino acid hydroxylases. *Annu Rev Biochem* 1999;68:355–81.
- [16] Liu Y, Hou Y, Wang G, et al. Gut microbial metabolites of aromatic amino acids as signals in host–microbe interplay. *Trends Endocrinol Metab* 2020;31:818–34.
- [17] Dodd D, Spitzer MH, Van Treuren W, et al. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. *Nature* 2017;551:648–52.
- [18] Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Discov* 2016;15:473.
- [19] Emwas A-H, Roy R, McKay RT, et al. NMR spectroscopy for metabolomics research. *Metabolites* 2019;9:123.
- [20] Khoomrung S, Wanichthanarak K, Nookaew I, et al. Metabolomics and integrative Omics for the development of Thai traditional medicine. *Front Pharmacol* 2017;8:474.
- [21] Wang Z, Klipfell E, Bennett BJ, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 2011;472:57–63.
- [22] Witkowski M, Nemet I, Alamri H, et al. The artificial sweetener erythritol and cardiovascular event risk. *Nat Med* 2023;29:710–8.
- [23] Vinaixa M, Schymanski EL, Neumann S, et al. Mass spectral databases for LC/MS- and GC/MS-based metabolomics: State of the field and future prospects. *TrAC Trends Anal Chem* 2016;78:23–35.
- [24] Zhou L, Yu D, Zheng S, et al. Gut microbiota-related metabolome analysis based on chromatography–mass spectrometry. *TrAC Trends Anal Chem* 2021;143:116375.
- [25] Luan H, Wang X, Cai Z. Mass spectrometry-based metabolomics: Targeting the crosstalk between gut microbiota and brain in neurodegenerative disorders. *Mass Spectrom Rev* 2019;38:22–33.
- [26] Zhao L, Ni Y, Su M, et al. High throughput and quantitative measurement of microbial metabolome by gas chromatography/mass spectrometry using automated alkyl chloroformate derivatization. *Anal Chem* 2017;89:5565–77.
- [27] Pautova A, Bedova AY, Sarshor YN, et al. Determination of aromatic microbial metabolites in blood serum by gas chromatography–mass spectrometry. *J Anal Chem* 2018;73:160–6.
- [28] Chernenkaya E, Beloborodova N, Klimenko N, et al. Serum and fecal profiles of aromatic microbial metabolites reflect gut microbiota disruption in critically ill patients: a prospective observational pilot study. *Crit Care* 2020;24(1):13.
- [29] Boernsen KO, Gatzek S, Imbert G. Controlled protein precipitation in combination with chip-based nanospray infusion mass spectrometry. An approach for metabolomics profiling of plasma. *Anal Chem* 2005;77:7255–64.
- [30] Ordóñez JL, Pereira-Caro G, Ludwig I, et al. 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* 2018;1575:100–12.
- [31] Yin S, Guo P, Hai D, et al. Optimization of GC/TOF MS analysis conditions for assessing host-gut microbiota metabolic interactions: Chinese rhubarb alters fecal aromatic amino acids and phenol metabolism. *Anal Chim Acta* 2017;995:21–33.
- [32] Winder CL, Dunn WB, Schuler S, et al. Global metabolic profiling of *Escherichia coli* cultures: an evaluation of methods for quenching and extraction of intracellular metabolites. *Anal Chem* 2008;80:2939–48.
- [33] Pinu FR, Villas-Boas SG, Aggio R. Analysis of intracellular metabolites from microorganisms: quenching and extraction protocols. *Metabolites* 2017;7:53.
- [34] Mohd Kamal K, Mahamad Maifiah MH, Abdul Rahim N, et al. Bacterial metabolomics: sample preparation methods. *Biochem Res Int* 2022;2022.
- [35] Lei T, Wang L, Chen C, et al. Metabolomic investigation of methicillin-resistant *Staphylococcus aureus*. *Methicillin-Resist Staphylococcus Aureus (MRSA) Protoc* 2014:251–8.
- [36] Halket JM, Waterman D, Przyborowska AM, et al. Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. *J Exp Bot* 2005;56:219–43.
- [37] Khoomrung S, Martinez JL, Tippmann S, et al. Expanded metabolite coverage of *Saccharomyces cerevisiae* extract through improved chloroform/methanol extraction and tert-butylidimethylsilyl derivatization. *Anal Chem Res* 2015;6:9–16.
- [38] Fiehn O. Metabolomics by gas chromatography–mass spectrometry: combined targeted and untargeted profiling. *Curr Protoc Mol Biol* 2016;114. 30.34. 31–30.34. 32.
- [39] Smart KF, Aggio RB, Van Houtte JR, et al. Analytical platform for metabolome analysis of microbial cells using methyl chloroformate derivatization followed by gas chromatography–mass spectrometry. *Nat Protoc* 2010;5:1709–29.
- [40] Tippmann S, Nielsen J, Khoomrung S. Improved quantification of farnesene during microbial production from *Saccharomyces cerevisiae* in two-liquid-phase fermentations. *Talanta* 2016;146:100–6.
- [41] Limjiasahapong S, Kaewnarin K, Jariyasopit N, et al. UPLC-ESI-MRM/MS for absolute quantification and MS/MS structural elucidation of six specialized pyranonaphthoquinone metabolites from *ventilago harmandiana*. *Front Plant Sci* 2021;11:2038.
- [42] Panuwet P, Hunter Jr RE, D'Souza PE, et al. Biological matrix effects in quantitative tandem mass spectrometry-based analytical methods: advancing biomonitoring. *Crit Rev Anal Chem* 2016;46:93–105.
- [43] Rigano F, Tranchida PQ, Dugo P, et al. High-performance liquid chromatography combined with electron ionization mass spectrometry: a review. *TrAC Trends Anal Chem* 2019;118:112–22.
- [44] Duangkumpha K, Jariyasopit N, Wanichthanarak K, et al. GC× GC-TOFMS metabolomics analysis identifies elevated levels of plasma sugars and sugar alcohols in diabetic mellitus patients with kidney failure. *J Biol Chem* 2022;298.
- [45] Xu T, Lu C, Feng L, et al. Liquid chromatography–mass spectrometry-based urinary metabolomics study on a rat model of simulated microgravity-induced depression. *J Pharm Biomed Anal* 2019;165:31–40.
- [46] Letertre MP, Myridakis A, Whitley L, et al. A targeted ultra performance liquid chromatography–Tandem mass spectrometric assay for tyrosine and metabolites in urine and plasma: Application to the effects of antibiotics on mice. *J Chromatogr B* 2021;1164:122511.
- [47] Anesi A, Rubert J, Oluwagbemigun K, et al. Metabolic profiling of human plasma and urine, targeting tryptophan, tyrosine and branched chain amino acid pathways. *Metabolites* 2019;9:261.
- [48] Pavlova T, Vidova V, Bienertova-Vasku J, et al. Urinary intermediates of tryptophan as indicators of the gut microbial metabolism. *Anal Chim Acta* 2017;987:72–80.
- [49] Gojda J, Cahova M. Gut microbiota as the link between elevated BCAA serum levels and insulin resistance. *Biomolecules* 2021;11:1414.
- [50] Zhou M, Shao J, Wu C-Y, et al. Targeting BCAA catabolism to treat obesity-associated insulin resistance. *Diabetes* 2019;68:1730–46.
- [51] Lefevre A, Mavel S, Nadal-Desbarats L, et al. Validation of a global quantitative analysis methodology of tryptophan metabolites in mice using LC-MS. *Talanta* 2019;195:593–8.
- [52] Alexeev EE, Lanis JM, Kao DJ, et al. Microbiota-derived indole metabolites promote human and murine intestinal homeostasis through regulation of interleukin-10 receptor. *Am J Pathol* 2018;188:1183–94.
- [53] Tang Z, Guengerich FP. Dansylation of unactivated alcohols for improved mass spectral sensitivity and application to analysis of cytochrome P450 oxidation products in tissue extracts. *Anal Chem* 2010;82:7706–12.
- [54] Riahi S, Ganjali MR, Hariri M, et al. Determination of the formation constant for the inclusion complex between Lanthanide ions and Dansyl chloride derivative by fluorescence spectroscopy: Theoretical and experimental investigation. *Spectrochim Acta Part A: Mol Biomol Spectrosc* 2009;74:253–8.
- [55] Žuvela P, Skoczylas M, Jay Liu J, et al. Column characterization and selection systems in reversed-phase high-performance liquid chromatography. *Chem Rev* 2019;119:3674–729.
- [56] Nakamura K, Saito S, Shibukawa M. Intrinsic difference between phenyl hexyl- and octadecyl-bonded silicas in the solute retention selectivity in reversed-phase liquid chromatography with aqueous mobile phase. *J Chromatogr A* 2020;1628:461450.
- [57] Buszewski B, Noga S. Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Anal Bioanal Chem* 2012;402:231–47.
- [58] Zeng Y, Luo L, Hou W, et al. Targeted metabolomics analysis of aromatic amino acids and their gut microbiota–host metabolites in rat serum and urine by liquid chromatography coupled with tandem mass spectrometry. *J Sep Sci* 2017;40:3221–30.
- [59] Zeng Y, Lin Y, Li L, et al. Targeted metabolomics for the quantitative measurement of 9 gut microbiota–host co-metabolites in rat serum, urine and feces by liquid chromatography–tandem mass spectrometry. *J Chromatogr B* 2019;1110:133–43.
- [60] Kaewnarin K, Limjiasahapong S, Jariyasopit N, et al. High-resolution QTOF-MRM for highly accurate identification and quantification of trace levels of triterpenoids in *ganoderma lucidum* mycelium. *J Am Soc Mass Spectrom* 2021.
- [61] Soga T. Advances in capillary electrophoresis mass spectrometry for metabolomics. *TrAC Trends Anal Chem* 2022;116883.
- [62] Migglioli P, Wouters B, van Westen GJ, et al. Novel technologies for metabolomics: More for less. *TrAC Trends Anal Chem* 2019;120:115323.
- [63] Zhang W, Ramautar R. CE-MS for metabolomics: Developments and applications in the period 2018–2020. *Electrophoresis* 2021;42:381–401.
- [64] Soga T, Heiger DN. Amino acid analysis by capillary electrophoresis electrospray ionization mass spectrometry. *Anal Chem* 2000;72:1236–41.
- [65] Soga T, Ueno Y, Naraoka H, et al. Simultaneous Determination of Anionic Intermediates for *Bacillus subtilis* Metabolic Pathways by Capillary Electrophoresis Electrospray Ionization Mass Spectrometry. *Anal Chem* 2002;74:2233–9.
- [66] Soga T, Ohashi Y, Ueno Y, et al. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *J Proteome Res* 2003;2:488–94.
- [67] Soga T, Baran R, Suematsu M, et al. Differential metabolomics reveals ophthalmic acid as an oxidative stress biomarker indicating hepatic glutathione consumption. *J Biol Chem* 2006;281:16768–76.
- [68] Mishima E, Fukuda S, Mukawa C, et al. Evaluation of the impact of gut microbiota on uremic solute accumulation by a CE-TOFMS–based metabolomics approach. *Kidney Int* 2017;92:634–45.
- [69] Koh A, Molinaro A, Ståhlman M, et al. Microbially produced imidazole propionate impairs insulin signaling through mTORC1. *Cell* 2018;175:947–61. e917.
- [70] Saito Y, Sato T, Nomoto K, et al. Identification of phenol- and p-cresol-producing intestinal bacteria by using media supplemented with tyrosine and its metabolites. *FEMS Microbiol Ecol* 2018;94:fiy125.
- [71] Piacentino D, Grant-Beurmann S, Vizioli C, et al. Gut microbiome and metabolome in a non-human primate model of chronic excessive alcohol drinking. *Transl Psychiatry* 2021;11(1):15.
- [72] Xia Y. Correlation and association analyses in microbiome study integrating multomics in health and disease. *Prog Mol Biol Transl Sci* 2020;171:309–491.
- [73] Yang Y, Misra BB, Liang L, et al. Integrated microbiome and metabolome analysis reveals a novel interplay between commensal bacteria and metabolites in colorectal cancer. *Theranostics* 2019;9:4101.

- [74] Chen Y-T, Yang N-S, Lin Y-C, et al. A combination of *Lactobacillus mali* APS1 and dieting improved the efficacy of obesity treatment via manipulating gut microbiome in mice. *Sci Rep* 2018;8(1):14.
- [75] Turnbaugh PJ, Ley RE, Hamady M, et al. The human microbiome project. *Nature* 2007;449:804–10.
- [76] Ammons MCB, Morrissey K, Tripet BP, et al. Biochemical association of metabolic profile and microbiome in chronic pressure ulcer wounds. *PLoS One* 2015;10:e0126735.
- [77] McHardy IH, Goudarzi M, Tong M, et al. Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite interrelationships. *Microbiome* 2013;1(1):19.
- [78] Smolinska A, Tedjo DI, Blanchet L, et al. Volatile metabolites in breath strongly correlate with gut microbiome in CD patients. *Anal Chim Acta* 2018;1025:1–11.
- [79] Liu R, Hong J, Xu X, et al. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat Med* 2017;23:859–68.
- [80] Hill CJ, Lynch DB, Murphy K, et al. Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort. *Microbiome* 2017;5(1):18.
- [81] Claesson MJ, Jeffery IB, Conde S, et al. Gut microbiota composition correlates with diet and health in the elderly. *Nature* 2012;488:178–84.
- [82] Nguyen QP, Karagas MR, Madan JC, et al. Associations between the gut microbiome and metabolome in early life. *BMC Microbiol* 2021;21(1):19.
- [83] Lee-Sarwar KA, Kelly RS, Lasky-Su J, et al. Integrative analysis of the intestinal metabolome of childhood asthma. *J Allergy Clin Immunol* 2019;144:442–54.
- [84] Mayneris-Perxachs J, Cardellini M, Hoyles L, et al. Iron status influences non-alcoholic fatty liver disease in obesity through the gut microbiome. *Microbiome* 2021;9(1):18.
- [85] El Aidy S, Derrien M, Merrifield CA, et al. Gut bacteria–host metabolic interplay during conventionalisation of the mouse germfree colon. *ISME J* 2013;7:743–55.
- [86] Wheelock AM, Wheelock CE. Trials and tribulations of ‘omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Mol Biosyst* 2013;9:2589–96.
- [87] Liang D, Li M, Wei R, et al. Strategy for intercorrelation identification between metabolome and microbiome. *Anal Chem* 2019;91:14424–32.
- [88] Morton JT, Aksenov AA, Nothias LF, et al. Learning representations of microbe–metabolite interactions. *Nat Methods* 2019;16:1306–14.
- [89] Reiman D, Layden BT, Dai Y. MiMeNet: exploring microbiome-metabolome relationships using neural networks. *PLoS Comput Biol* 2021;17:e1009021.
- [90] Yin X, Altman T, Rutherford E, et al. A comparative evaluation of tools to predict metabolite profiles from microbiome sequencing data. *Front Microbiol* 2020;11:3132.
- [91] You Y, Liang D, Wei R, et al. Evaluation of metabolite-microbe correlation detection methods. *Anal Biochem* 2019;567:106–11.
- [92] Reshef DN, Reshef YA, Finucane HK, et al. Detecting novel associations in large data sets. *science* 2011;334:1518–24.
- [93] Giles K, Ujma J, Wildgoose J, et al. A cyclic ion mobility-mass spectrometry system. *Anal Chem* 2019;91:8564–73.
- [94] Lippa KA, Aristizabal-Henao JJ, Beger RD, et al. Reference materials for MS-based untargeted metabolomics and lipidomics: a review by the metabolomics quality assurance and quality control consortium. *mQACC* 2022;18:24.