



Universiteit  
Leiden  
The Netherlands

## **Metabolomics approaches for unravelling host-gut microbiota changes in different stages of cognitive impairment**

Hosseinkhani, F.

### **Citation**

Hosseinkhani, F. (2024, December 5). *Metabolomics approaches for unravelling host-gut microbiota changes in different stages of cognitive impairment*. Retrieved from <https://hdl.handle.net/1887/4170946>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4170946>

**Note:** To cite this publication please use the final published version (if applicable).

## Towards standards for human fecal sample preparation in targeted and untargeted LC-HRMS studies

*Gut microbiota and their metabolic products are increasingly being recognized as important modulators of human health. The fecal metabolome provides a functional readout of the interactions between human metabolism and the gut microbiota in health and disease. Due to the high complexity of the fecal matrix, sample preparation often introduces technical variation, which must be minimized to accurately detect and quantify gut bacterial metabolites. Here, we tested six different representative extraction methods (single-phase and liquid–liquid extractions) and compared differences due to fecal amount, extraction solvent type and solvent pH. Our results indicate that a minimum fecal (wet) amount of 0.50 g is needed to accurately represent the complex texture of feces. The MTBE method (MTBE/methanol/water, 3.6/2.8/3.5, v/v/v) outperformed the other extraction methods, reflected by the highest extraction efficiency for 11 different classes of compounds, the highest number of extracted features (97% of the total identified features in different extracts), repeatability (CV <35%) and extraction recovery ( $\geq 70\%$ ). Importantly, optimization of the solvent volume of each step to the initial dried fecal material ( $\mu\text{L}/\text{mg}$  feces) offers a major step towards standardization, which enables confident assessment of the contributions of gut bacterial metabolites to human health.*

Publications associated to this chapter:

**F. Hosseinkhani**, AC. Dubbelman , N. Karu, A.C. Harms, and T. Hankemeier; Towards Standards for Human Fecal Sample Preparation in Targeted and Untargeted LC-HRMS Studies; *Metabolites* 2021, 11(6), 364; <https://doi.org/10.3390/metabo11060364>

## BACKGROUND

Gut microbial metabolites are rapidly becoming a key facet of human health and disease.<sup>1-3</sup> In this context, the fecal sample has emerged as the most accessible biological matrix, which reflects the functional readout of microbial activity in the gastrointestinal tract and probes the reciprocal regulation between gut microbiota and the physiology of the host.<sup>4</sup> Several previous studies have proposed the versatility of fecal metabolites in disease diagnosis. For example, elevated concentrations of amino acids, saturated fatty acids, and ursodeoxycholic acid were reported in fecal samples of colorectal cancer patients<sup>5</sup>, whereas increased levels of deoxycholic acid, L-tryptophan and putrescine were found in fecal samples of patients with systemic lupus erythematosus.<sup>6</sup> Short chain fatty acids, polyphenols and vitamins have been shown to decrease the risk of inflammatory bowel disease, cancer, diabetes and cardiovascular diseases.<sup>7,8</sup> Additional correlations will likely emerge soon. Thus, examining the fecal metabolome serves as a robust approach for understanding the interaction between human metabolism and the gut microbiota composition in health and disease.

Despite the rising popularity of fecal metabolomics, the methods for fecal sample preparation and analysis are still a long way from being standardized.<sup>9,10</sup> In recent years, various studies on human fecal metabolomics aimed to provide guidelines for fecal sample collection, storage and sample pre-treatment. Sample storage at  $-80^{\circ}\text{C}$  has been introduced as the “golden standard” for both microbiome and metabolomics studies, as the use of frozen samples in human studies is both more common and practical, while minimizing fecal microbial changes and metabolite variation.<sup>11-14</sup> For a uniform sample handling procedure in metabolomics studies, homogenization and aliquoting of the fecal sample prior to freezing has been suggested by Karu et al.<sup>9</sup>

However, controversy remains between studies regarding the critical parameters of fecal extraction such as appropriate sample size, extraction solvent type and solvent pH. One of the main challenges in fecal metabolomics is the high heterogeneity and complexity of fecal samples, which limits repeatability and efficiency of metabolite extraction. In addition, the fecal metabolome contains a wide range of metabolites with diverse structures and different chemical properties.<sup>15</sup> The selection of the extraction solvent can strongly bias the outcome of the study as it significantly affects the loss, underestimation or overestimation of the metabolites’ recovery. To obtain a thorough metabolite extraction that is analyzable with good coverage, efficiency and repeatability, compromises in the sample (pre)treatment have to be made. The lack of a well-defined and standardized protocol causes independent studies to report different metabolite coverages and distinct quantities of metabolite biomarkers, due to the utilization of different sample treatment platforms, which impedes biological interpretation.<sup>16</sup> This illustrates the importance of data generation and reporting

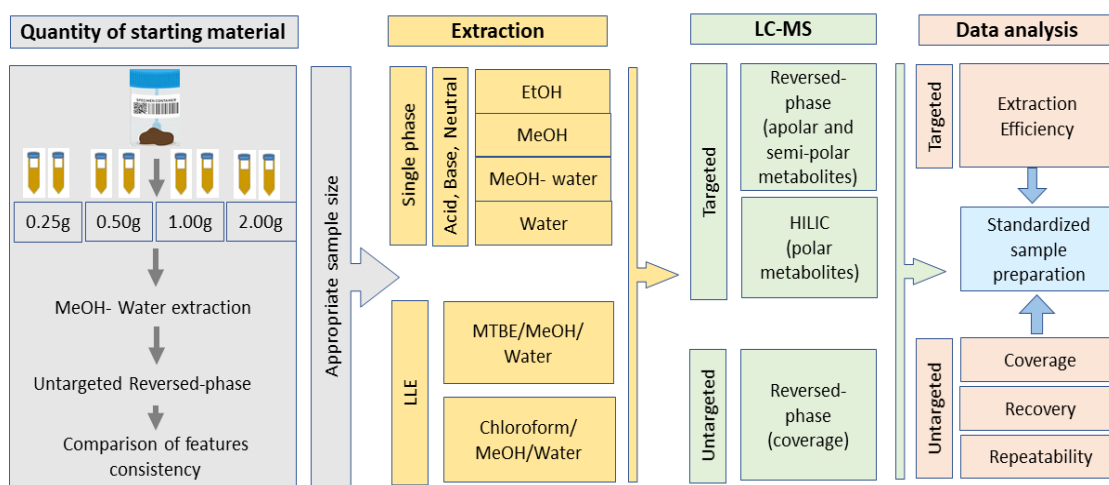
in a unified manner and highlights the need to standardize sample preparation and analysis workflow.

In recent years, liquid chromatography mass spectrometry (LCMS) has emerged as the most commonly used technique for global metabolic profiling, due to its sensitivity, and capability for resolving various classes of metabolites.<sup>17</sup> Therefore, in this study, we aimed at designing a standardized human fecal sample preparation methodology for targeted and untargeted metabolomic analysis using high resolution LCMS. We evaluated in particular the quantity of starting fecal material for extraction, the effect of solvent type and pH on the extraction efficiency, and the effect of the extraction methods on the extraction coverage, repeatability and recovery.

## RESULT

### *Sample Preparation Methods and Workflow*

Fecal samples from healthy volunteers ( $n = 3$ ) were collected for this study. The fresh samples were homogenized by stirring, aliquoted and stored directly at  $-80^{\circ}\text{C}$  until extraction. The experimental workflow is illustrated in Figure 1.



**Figure 1.** *Experimental workflow and sample treatment evaluation*

First, the effect of starting fecal amount on the extraction repeatability was assessed for four different sample sizes (in duplicate) using untargeted reversed-phase LC-MS (RPLC-MS); see method 1 in Table S2. This experiment was performed on aliquots from an individual fecal sample to exclude bias due to individual sample differences (Figure 2).

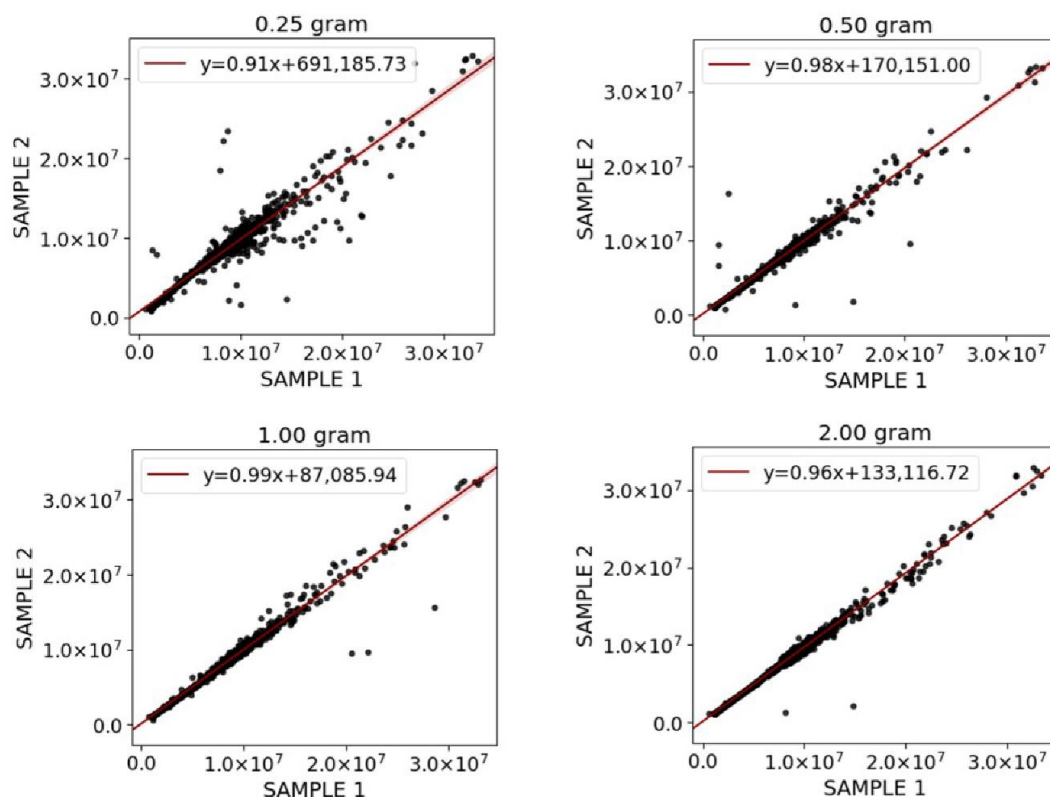
Next, side-by-side comparisons of four single-phase extractions (EtOH, MeOH, MeOH-water and water) in three different pH (acidic, neutral and basic), and two liquid-liquid extractions (MTBE and chloroform methods) were performed to evaluate the efficiency, coverage, recovery and repeatability of extraction of polar and apolar metabolites from fecal samples. For this experiment, samples from the three volunteers were pooled (see Materials and Methods Section 4.4). For extraction efficiency, the extracts were analyzed in a targeted manner using RPLC-MS (method 2 in Table S2) and hydrophilic interaction liquid chromatography (HILIC)-MS (method 3 in Table S2). Both platforms were operated only in negative mode (see Section 4). The choice of ionization mode was based on our laboratory R&D process and previously published data<sup>18,19</sup>, which showed that most polar compounds are ionized better in negative ion mode. Compounds such as acylcarnitines, which ionize only in positive ESI, are not included in our target list. Information on the performance of the targeted methods used can be found in Supplementary Material Tables S3 and S4 (excel files). These techniques targeted a total of 27 semi- and apolar compounds from three different classes (fatty acyls, steroids and steroids derivatives and glycerophospholipids), and 26 polar compounds from eight different classes (organooxygen compounds, hydroxy acids and derivatives, keto acids and derivatives, carboxylic acids and derivatives, diazines, purine nucleotides, imidazopyrimidines and organic sulfonic acids and derivatives). The target list was selected to include representative fecal metabolites from classes important for characterizing gut microbiota, diet and human health.<sup>3,20–23</sup> The classification of metabolites was based on the Human Metabolome Database (HMDB) [24]. The relative abundance of each detected metabolite (only the first extract) was used to evaluate the extraction efficiency for each corresponding metabolite class (Figure 3).

For extraction coverage, extracts were analyzed using untargeted RPLC-MS (method 1) and the total number of features from each extract was determined using an XCMS pipeline (Figure 4, Tables S5 and S6). Finally, extraction repeatability and recovery were examined for extraction methods with higher feature coverage using isotopically labeled standards and untargeted RPLCMS (method 1).

### **Quantity of Starting Material**

In this study, two aliquots of each size (0.25, 0.50, 1.00 and 2.00 g) were taken from the same homogenized fecal scoop. Each sample was measured in triplicate and the mean intensities of features from each sample size was compared and plotted against the mean intensity of features from the corresponding sample size (Figure 2). We observed that a minimum of 0.50 g wet feces from the whole scoop was needed as starting material, and smaller sample sizes ( $\leq 0.50$  g wet feces prior to freeze drying) tended to suffer from a

relatively higher bias. Although samples of 1.00 and 2.00 g exhibited the highest coefficient of determination ( $R^2 = 0.9804$  and  $0.9863$ , respectively), practical difficulties arose regarding the volume of solvent needed for extraction; in particular when the sample was treated with LLE. In this study, a sample size of 0.50 g was chosen for further steps as it showed negligible variance for replicate analysis ( $R^2 = 0.9629$ ).



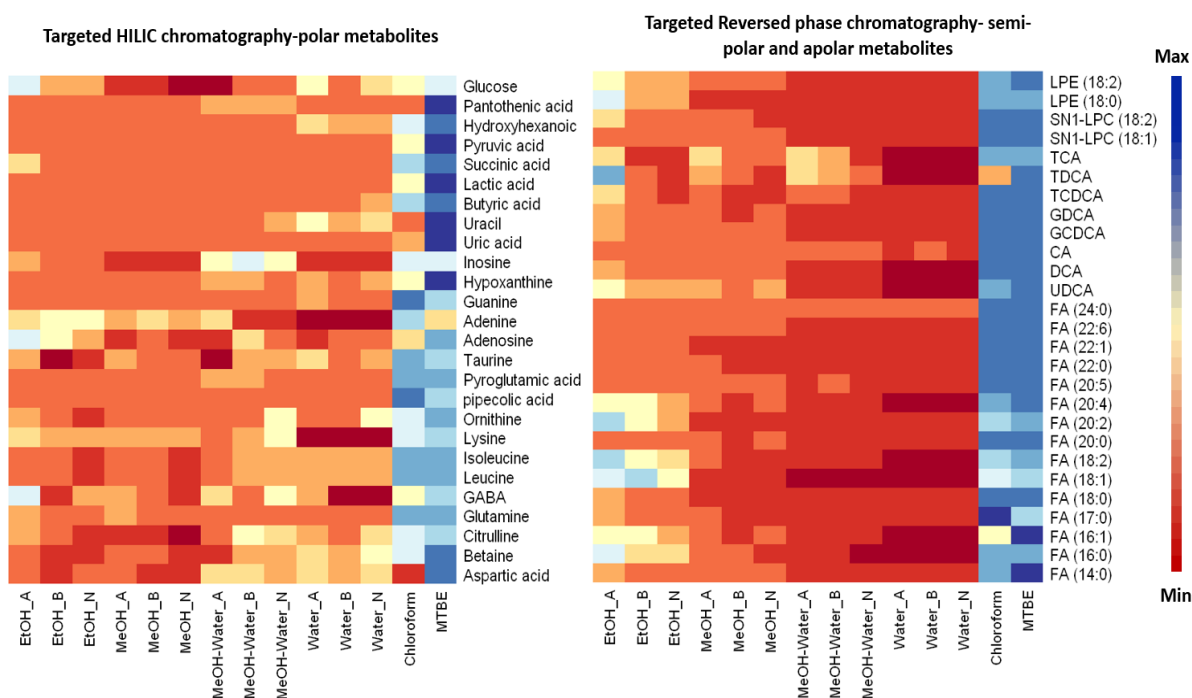
**Figure 2.** Effect of fecal amount (0.25, 0.50, 1.00 and 2.00 g) on feature intensities of two aliquots taken from the same homogenized fecal scoop. Black dots represent intensities of each feature. A linear relation between feature intensities of two subsamples from the same fecal scoop indicates a consistent representation of the sample.

#### *Assessment of Extraction Method, Solvent and pH for Optimal Extraction Efficiency*

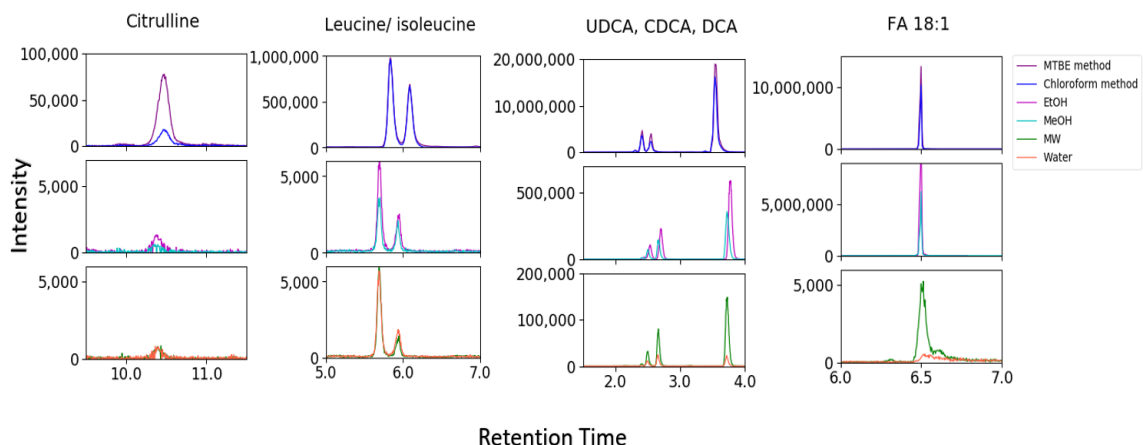
The extraction efficiency was affected by the type of solvent, polarity and pH value of extraction solvents. Owing to the diverse classes of compounds in our target list (53 compounds from 11 different biochemical classes) the obtained results can be extrapolated to vast numbers of metabolites with diverse chemical properties.

## pH Effect of Extraction Solvent

The effect of solvent pH on extraction efficiency was assessed by comparing three different pH values (acidic, basic and neutral) for solvents in single-phase extractions. Peak areas of 53 analytes of interest (measured in triplicate) were plotted in a heatmap (Figure 3a). Basic pH, independent of the solvent used, tended to provide the lowest efficiency across the wide range of metabolites. The extraction yield of fatty acyls, steroids and steroids derivatives, and glycerophospholipids was higher with acidic EtOH, while carboxylic acids such as ornithine, lysine and leucine had a higher yield in neutral pH; purine nucleotides and imidazopyrimidines showed a higher yield with basic pH. Due to the different charges of metabolites, we could not find one ideal pH value that was optimal for the full range of metabolites in our experiment.



**Figure 3.** The effect of the extraction solvent system on the extraction efficiency. (a) Heatmap of the peak area of the metabolites from different classes of compounds using single-phase (first cycle of extractions with different solvents and pH values) and two-phase liquid-liquid extraction (MTBE and chloroform methods). Dark blue and dark red colors represent the largest and smallest peak area, respectively.



**Figure 3.** (b) Peak shapes and intensity of metabolites extracted with different extraction solvents in neutral pH. -A: acidic, -B: basic, -N: neutral. FA: fatty acid; UDCA: ursodeoxycholic acid; CDCA: chenodeoxycholic acid; DCA: deoxycholic acid.

### Effect of Solvent Selection

The appropriate solvent selection for single-phase extractions is a subject of debate. Similar to the results obtained for pH, as expected, there is no ideal solvent for single-phase extractions with high efficiency for all types of metabolites (Figure 3a, columns with EtOH, MeOH, MeOH-water (MW) and water extraction under three different pH values). Specifically, a higher yield for fatty acids (fatty acyls), LPE, LPC (glycerophospholipids), bile acids (steroids and steroids derivatives), glucose (organooxygen compounds) and adenosine (purine nucleoside) was observed with acidified EtOH, whereas some carboxylic acids and derivatives (ornithine, lysine and GABA) showed a better extraction yield with MeOH-water. Others such as taurine, betaine, leucine and isoleucine were better extracted with water.

### LLE Solvent Effect

To circumnavigate the chemical limitations and coverage of single solvent extractions, we performed two different LLE methods using MTBE and chloroform. Slight modifications were implemented to obtain a better solvent ratio of MTBE/MeOH/water (3.6/2.8/3.5,



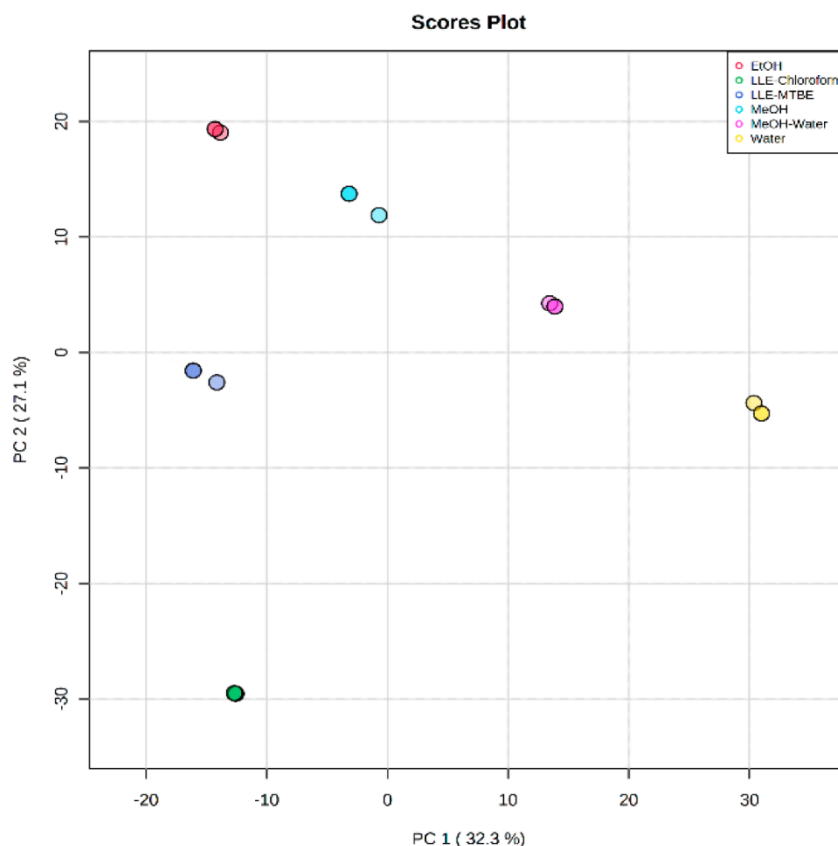
v/v/v) for extraction of both polar and apolar metabolites (see Materials and Methods Section). Although, the total volume of solvents in LLE is 10.2  $\mu$ L, the volume of polar solvent in LLE (MeOH and water) is similar to those in single-phase extractions. We considered a very low to no effect of MTBE/ chloroform on the extraction of polar metabolites. Therefore, to ensure a fair comparison, the volume collected from each layer of LLE to be dried and reconstituted was the same as the volume collected from single-phase extractions. Aqueous and organic layers of LLE were analyzed separately for assessing the extraction efficiency for polar and apolar metabolites. For a fair analysis, we first compared the efficiency based on the first cycle of single-phase extractions. In comparison to single-phase extractions, LLE with either of the methods showed a higher or comparable extraction efficiency (Figure 3a, columns entitled MTBE and chloroform). However, the MTBE method outperformed the chloroform method for polar metabolites, in terms of superior peak shapes and intensities (Figure 3b). In particular, the MTBE method demonstrated a higher yield of organic acids, such as pyruvic acid, succinic acid and lactic acid. Moreover, single-phase extractions such as with EtOH require at least three cycles of extraction, as well as solution pH optimization (dependent on the class of compounds) to obtain a comparable efficiency with LLE (Supplementary Figure S1).

### Metabolic Coverage

Untargeted metabolomics analysis of each extraction method was utilized to estimate and compare the metabolic coverage of a wide range of molecules with varying hydrophobicity and polarity. In addition to single-phase extractions, we also examined a mix of the polar and apolar phases of the LLE, aiming at increasing the metabolic coverage. Following untargeted RPLC-MS analysis, data were pre-processed using XCMS, MS-FLO and in-house tools, followed by strict rules of feature removal, to omit isotopes, adducts, by-products, background peaks (in procedure blank per method), and also unreliable retention areas (which restricted the inclusion of apolar and larger lipids). The different steps of the pre-processing and the feature yield per step and extraction method are summarized in Table S6. Limited feature annotation was conducted, to assess coverage similarities and differences between the extraction methods. Curation of features, their assessment and metabolite annotation are an ongoing effort as part of building our in-house database of metabolites per biospecimen, species and health condition.

A PCA plot provides an initial impression of the untargeted metabolomics pre-processed results (Figure 4). Apart from fine repeatability of replicate measurements in each extraction method, the PCA illustrates profound differences between the obtained metabolic profiles. PC1 (32%) reflects the polarity of the extraction solutions, separating water-containing solutions from those with higher organic solvent content (right to left,

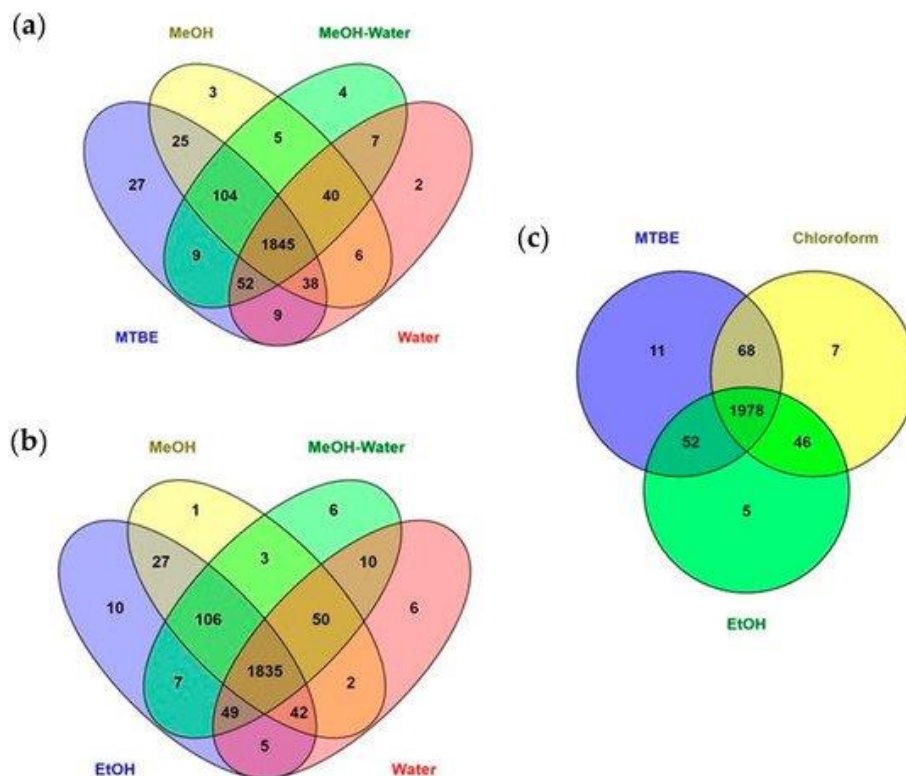
decreasing polarity index). PC2 (27%) is more affected by hydrophobicity and miscibility of the solutions, clearly isolating the LLE-chloroform phase from all others, and not surprisingly closest to the LLE-MTBE extraction. Hierarchical analysis of the untargeted metabolomics results also illustrates the similarities and clustering between the various extraction methods and replicates (Figure S2).



**Figure 4.** Principle component analysis (PCA) scores plot of pre-processed untargeted features ( $n = 2176$ ) measured in the different fecal extraction procedures ( $n = 3$  replicates).

Adding up all features in the different extracts, a total of 2176 features were detected and retained after data pre-processing, among which the highest coverage was observed for MTBE, chloroform and EtOH extractions (97%, 96% and 91% of total, respectively). Figure 5 summarizes the metabolic coverage overlap in Venn diagrams. Although MTBE and chloroform extractions represented 94% similarity in features, chloroform extractions represented 94% similarity in features, it is clear that MTBE extraction contributes the highest number of unique features, also compared to the more polar extracts. Moreover, MTBE extraction offers an advantage since the protein and cell debris layer is forced to the bottom of the extraction tube following centrifugation, simplifying the removal of both

solvent phases. Within the single-phase extraction methods (polar extraction), EtOH provided the highest coverage, probably due to its intermediate polarity.



**Figure 5.** Comparison of the mean numbers of features (triplicate of each extraction protocol); (a) comparison of single-phase extractions; (b) comparison of MTBE with single-phase extractions; (c) comparison of ethanol (which gave the highest number of features among the single-phase extractions) with liquid-liquid extractions.

To demonstrate the effect of the polarity and hydrophobicity of the different solutions on the efficiency of extraction, we present a few examples of standard-identified and putatively annotated features (level 5 identification, with supporting adducts, see Table S7 in Supplementary Excel file). The dominant feature of a metabolite in each extract was selected, if not identical between extracts (for example, the formate adduct of deoxycholate acid was higher only in aqueous extract, while [M-H] was the highest in all other extracts). A putative lignan (enterolactone, eluting at 5 min) recorded the highest peaks in MTBE and in chloroform extracts, followed by the increasing polarity: EtOH, MeOH, MeOH:water, water. A similar pattern was observed for a steroidal hormone (androsterone;  $rt = 5.5$  min) which gave the highest peak in MTBE extract, followed by a 50% decrease in chloroform

extract, and a further 100-fold decrease in EtOH and MeOH extracts, and with negligible peaks in MeOH:water and water extracts. Bile acids (rt = 5–6.7 min) were mostly higher in MTBE, lower in MeOH:water extract, and the lowest in water extracts. Variation in residues affected the performance of the chloroform, ethanol and methanol extracts to different degrees. A small polar lipid (e.g., LPC 12:0, rt = 6.5 min), showed the highest peak intensity in EtOH extract, followed by MeOH, then 2–3 fold lower peak intensity in chloroform and MTBE extracts, and 100–1000 fold lower peaks in MeOH-water and water extracts, respectively. More extremely, a putative triterpenoid (rt = 6.8 min) showed the highest peak area in the EtOH extract, a 5-fold decrease in the MeOH extract, and 100–1000 fold decrease in all other extraction methods. Bilirubin degradation products (urobilinogen etc.) showed the lowest peak area in chloroform extracts. Aromatic amino acids (Trp, Tyr, Phe; rt = 2.1–2.8 min) showed the highest peak area in the MTBE extract, and the lowest (5–20 fold decrease) in the chloroform extract. This also tended to be the picture for purines (uric acid, xanthine, inosine and others; rt 1.4–2.6 min), although the later eluters of the class (methyl- and dimethyl-urates) tended to show similar peak areas for all extraction methods. No dramatic differences between extraction methods were observed for polar benzenoids (such as phenylacetic acid and phenol glucuronide). The examples above highlight that no single extraction method provides the best sensitivity for all chemical classes, and it can be beneficial to tailor the method to the priorities of the study (lipids; natural products; various gut-bacteria metabolites etc.).

To further support the reliability of the untargeted measurements, Figure S3 presents the relative peak area of metabolites from different chemical classes (identity confirmed with authentic chemical standards). The same metabolites were analyzed by targeted methods to examine their recovery efficiencies, as discussed in the following section.

### **Extraction Repeatability and Recovery for Selected Methods**

Since MTBE, EtOH and chloroform yielded the highest metabolite coverage; extraction repeatability and recovery were evaluated for these methods only.

Within-day repeatability of the extraction methods was calculated based on the peak areas of 15 spiked stable-isotope labeled standards (Table 1). Overall, the CV values of the MTBE method (CV range 3–28.5%) were lower than those for the chloroform (CV range 1.5–42%) and EtOH (CV range 2–41%) methods. The overlaid total ion chromatogram and extracted ion chromatograms of three different MTBE extractions are illustrated in Supplementary Figure S4.

Table 1. The coefficient of variation (CV%) for peak areas of 15 stable isotope labeled standards spiked in fecal samples for MTBE, chloroform and ethanol extraction protocols in negative ESI using untargeted LCMS.

Extraction recoveries were calculated by the ratio of the stable-isotope labeled standard in the pre-extraction samples to the same-labeled standard in the post-extraction species. Similar to the trend observed for extraction efficiency of endogenous metabolites, MTBE outperformed other methods in extraction recovery. As illustrated in Supplementary Figure S5, D4-DCA, long chain fatty acids, U 13C6-lysine, D6-ornithine, D3-9-15N-aspartate and D2-glycine demonstrated high recovery ( $\geq 80\%$ ) in the chloroform and MTBE methods. D4- Succinate showed high recovery (70%) in MTBE method only. Overall, EtOH extraction showed a lower recovery for all tested compounds. This clearly indicates the importance of performing recovery tests prior to selecting a solvent for extraction.

The higher CV value and lower recovery in the chloroform method might be attributed to the presence of the interphase (cell debris and protein) between the aqueous and organic layer.

## DISCUSSION

The major objective of this study was to find an optimal and robust sample preparation method for human feces directed to both targeted and untargeted metabolomics analysis, through the application of an LC-MS-driven metabolomics workflow.

### Quantity of Starting Material

An important factor that affects the extraction efficiency and recovery is the quantity of starting material, especially for a complex matrix like feces. Unlike other biofluids, feces are not homogenous and contains different topographical locations. Recently, it has become apparent that the sampling region of the human fecal scoop significantly impacts metabolic profiles.<sup>14</sup> In this regard, many studies solely performed homogenization prior to aliquoting<sup>25,26</sup>, and the effect of fecal scoop size on the metabolite composition (i.e., what scoop size can be representative of the whole) has not yet led to a consensus. Gratton et al.<sup>27</sup> recommended 15 g of fecal sample as a representative amount. However, this amount was used for fecal water extraction and the sample size might be too large to be applicable for all cohort studies. Moreover, collected fecal samples from patients are often too large to be easily handled for homogenization. Our data revealed that, in case of homogenization with simple stirring, at least 0.50 g of wet feces from the whole feces scoop is needed as starting material to accurately represent the whole; smaller sample sizes ( $\leq 0.50$  g) suffered from relatively higher deviations in feature intensity. This might be attributed to the presence of fecal particulates and undigested material such as fibers in the feces. In addition, the effect of the sample type (e.g., diarrhea or constipated) on the required sample size needs to be explored.

### **Assessment of Extraction Method, Solvent and pH for Optimal Extraction Efficiency**

Fecal samples are typically prepared with non-selective sample pretreatment strategies to allow maximum coverage. While extraction with one solvent is common <sup>28</sup>, there is an inconsistency between many studies with regard to pH and solvent selection, and the resulting metabolic coverage and extraction efficiency. Deda et al. <sup>29</sup>, selected acidic or neutral acetonitrile as solvents with satisfactory extraction coverage for different classes of compounds in rat feces, while neutral propanol provided a better selectivity for compounds such as ornithine, lysine, hypoxanthine and tryptamine. Bascon et al. <sup>30</sup> used MeOH or MeOH followed by ethyl acetate for extraction of polar and apolar metabolites from pig feces. In the study by Turroni et al. <sup>31</sup>, who applied a targeted metabolomics approach to profile the human fecal metabolome, MeOH provided a higher yield for metabolites such as sphingomyelin, phospholipids and acylcarnitines, while for amino acids and hexose sugars, a mixture of MeOH and phosphate buffer resulted in a higher yield. Complementary to previous studies, our data revealed that there is no ideal pH value that suits the full range of metabolites with different chemical properties, and obtaining satisfactory extraction efficiency for all metabolites with single-phase extraction is unfeasible.

The potential of LLE for efficient extraction of a broad range of polar and apolar metabolites from different matrices became the focus of recent metabolomics studies. In this regard, the superiority of LLE extraction with MTBE was shown by Sostare et al. <sup>32</sup>, who reported a higher reproducibility and efficiency in human plasma and urine. In another study, Patterson et al. <sup>33</sup> demonstrated the strength of the Matyash method (based on MTBE) as compared to the Bligh and Dyer method (based on chloroform) for efficient and reproducible extraction of a broad range of polar metabolites and lipids from plasma. Similarly, in the presented experiment, the MTBE method outperformed LLE with chloroform and single-phase extractions in terms of extraction efficiency for polar and apolar metabolites in feces. Notably, solvents such as MTBE and chloroform can disrupt cells by dissolving membrane lipids, hence facilitating the release of intracellular metabolites and increasing metabolites yield. This is particularly important for complex biological matrices such as feces containing different types of cells. It has been shown that commonly used disruption methods such as sonication and bead beating are not sufficient to entirely lyse thick and waxy cell walls, especially for gram positive and fungal cells. <sup>34</sup> Although fecal metabolites of intact bacterial cells do not impact the human health state, fundamental bias between samples can be generated by partial cell lysis during sample pretreatment steps (i.e., freeze-thaw cycles, sonication, and vortexing). Therefore, a selection of solvents for fecal extraction, such as chloroform and MTBE, in combination with shearing forces can improve the comparability between samples. Overall, data must be interpreted with caution, acknowledging the effect of the release of intracellular metabolites on the final extraction yield.

## Extraction Coverage

Untargeted metabolomics aims to cover a broad range of metabolites from biological samples, which largely depend on the sample preparation protocols used for metabolite extraction. Our results demonstrated differences in metabolic coverage and sensitivity between the various extraction methods. Liu et al. reported improvements in metabolite coverage of plasma sample by using two-step LLE in comparison to conventional MeOH extraction.<sup>35</sup> Similarly, Whiley et al.<sup>36</sup> reported a drastic improvement in the coverage and reproducibility of plasma metabolites using LLE extraction with MTBE in comparison to MeOH/EtOH (1:1 v/v). To the best of our knowledge, we showed for the first time that mixing polar and apolar phases of MTBE extraction improves the coverage of fecal metabolites significantly. Although it is expected that water and acetonitrile (9:1 v/v) as reconstitution solvent reduce the amount of phospholipids in the final extract, we did not assess the matrix effect for each extraction method in this study. Further studies, which take the matrix effect into account, will need to be undertaken.

## Extraction Repeatability and Recovery for Selected Methods

Based on the criteria for repeatability ( $CV \leq 35\%$ ) and recovery ( $\geq 80\%$ ) of extraction suggested by previous studies<sup>37</sup>, our results suggest that good repeatability and recovery was obtained for the MTBE method, which was satisfactory for metabolomics extraction.

### *Standardization*

One aspect in fecal sample preparation, which has not been explicitly mentioned before is the way to report the amounts or volumes that are being used during sample preparation. Usually the solvent volume of the first step is reported as  $\mu\text{L}/\text{mg}$  feces, and for the solvents in further steps, it has been reported as v/v ratio<sup>16,29</sup>, which complicates calculating the actual used volumes. We propose to report the solvent volume of each step to the initial fecal material ( $\mu\text{L}/\text{mg}$  feces or  $\mu\text{L}/\text{mg}$  dried feces), as described in the Materials and Methods Section, to make it easier for others to repeat our method.

In conclusion, the MTBE extraction method exhibits several advantages, including superior extraction efficiency for different classes of compounds, high metabolic coverage (using mixed layers), high repeatability and satisfactory recovery. With interest in both polar and apolar metabolites, a robust biphasic extraction method such as MTBE for the complex matrix of feces can greatly reduce experimental time and cost. Importantly, reporting solvent volumes of each step to the initial dried fecal material ( $\mu\text{L}/\text{mg}$  feces) offers a major step towards standardization, which enables confident assessment of the contributions of gut bacterial metabolites to human health.

Despite some limitations (low number of samples and donors and incomplete chemical representation), this work shows how commonly applied feces sample preparation techniques compare in their performance for targeted and untargeted metabolomics. Additionally, it provides guidance on the optimization and standardization of fecal sample preparation, which is highly valuable in the further exploration of the effect of gut microbiota on human health and disease.

## **MATERIALS AND METHODS**

### **Chemical and Reagents**

Analytical grade solvents acetonitrile (MeCN), methanol (MeOH), ethanol (EtOH), chloroform and formic acid were purchased from Biosolve (Biosolve BV, Valkenswaard, The Netherlands), whereas ammonium formate, methyl tert-butyl ether (MTBE), and leucine-enkephalin were purchased from Sigma–Aldrich (Sigma–Aldrich, Burlington, WV, USA). Ultrapure water was obtained from an arium pro UF/VF water purification system with a Sartopore 0.2  $\mu\text{m}$  filter (Sartorius Stedim, Amersfoort, The Netherlands). The majority of standards and deuterated internal standards were purchased from Sigma Aldrich (Sigma-Aldrich, Zwijndrecht, The Netherlands), Avanti (Avanti Polar Lipids, Alabaster, AL, USA), and Fluka (Honeywell, Netherlands). More detailed information regarding the standard suppliers and utilized concentration can be found in the Supplementary Information (Supplementary Table S1).

### **LCMS Analysis**

Untargeted and targeted analysis were performed on an Acquity UHPLC system coupled to a Synapt\_G2S mass spectrometer (Waters Chromatography Europe BV, Etten-Leur, The Netherlands), collecting full scan MS data with a 40–1200 Da mass range. Electrospray ionization was applied in negative mode only. A concentration of 0.1 mg/L Leucine-Enkephalin in water/MeOH/formic acid (50/50/0.1 v/v/v) was used as reference for mass measurement with infusion flow rate of 10  $\mu\text{L}/\text{min}$ . MS data were collected in centroid mode. Targeted chromatographic separations were carried out on RP (Waters Acquity UPLC T3 column (2.1  $\times$  100 mm, 1.8  $\mu\text{m}$ )) and hydrophilic interaction liquid chromatography (SeQuant® ZIC-chILIC (2.1  $\times$  100 mm, 3  $\mu\text{m}$ )). Targeted RP (semi polar and apolar metabolites) followed a linear gradient from 0–99.9% mobile phase B (MeOH with 10 mM ammonium formate) over 11.5 min with an injection volume of 4  $\mu\text{L}$ . For HILIC separation (polar metabolites), a linear from 0–100% mobile phase B (10% MeCN in 5 mM ammonium formate) in 14 min with an injection volume of 2.5  $\mu\text{L}$ . Untargeted metabolomics was performed using RP chromatography which followed a linear gradient



from 0–99.9% mobile phase B (MeCN with 0.1% formic acid) in 8 min with an injection volume of 4  $\mu$ L. More details of separation conditions and mass spectrometry parameters are mentioned in Supplementary Table S2.

### **Quantity of Starting Material**

The adequacy of starting fecal material for the extraction was investigated for 4 different amounts ( $0.25 \pm 0.02$ ,  $0.50 \pm 0.02$ ,  $1.00 \pm 0.02$ , and  $2.00 \pm 0.02$  g). For each size, duplicates were weighed from the same homogenized sample. After freeze-drying (RVT400 Refrigerated Vapor Trap, Savant, Buckinghamshire, UK), extraction was performed for all different fecal aliquots in parallel using ice-cold MeOH and water (5:1 v/v) in a ratio of 6:1  $\mu$ L/mg feces. Every mixture was homogenized by sonication for 10 min in an ultrasonic bath (Ultrasound cleaning baths, USC-TH, VWR, Amsterdam, The Netherlands), followed by vortex mixing on high speed for 10 min. To maximize protein precipitation, samples were stored at  $-20$  °C for 20 min. Thereafter, samples were centrifuged for 20 min ( $14,700\times g$ ,  $10$  °C). Supernatants were filtered through syringe filters PTFE ( $0.22$   $\mu$ m) before untargeted RP LC-MS analysis. All samples were measured in triplicate (Supplementary Table S2) and the mean intensity of features was used for further analysis.

### **Assessment of Extraction Method, Solvent and pH for Optimal Extraction Efficiency**

To assess the extraction efficiency for chemically diverse metabolites, present in the fecal matrix, first, a standard quality control solution (with 53 compounds from 13 different classes) was used to test the performance of targeted analytical platforms. Identity of analytes was determined based on the selectivity of high-resolution mass-to-charge and the retention time of the authentic chemical standards. A pooled sample from 3 individual subjects was used. Firstly, each individual's whole fecal sample ( $\cong 7.00$  g) was aliquoted (1.00 g) and homogenized using sonication (10 min) and high-speed vortex (10 min); then aliquots of all three samples were mixed and homogenized (sonication: 5 min, high-speed vortex: 10 min, stirring with steel spatula). Next, aliquots of 0.50 g were prepared from the pool sample and randomly assigned to different extraction conditions. For single-phase extraction, four commonly used extraction solvents in metabolomics, namely EtOH, MeOH, MeOH: water (1:1 v/v), and water were utilized. Each extraction solvent was added in 4:1  $\mu$ L/mg of feces. To investigate the effect of pH, all above-mentioned solvents were used in acidic (0.1% formic acid), neutral, and basic (0.1% ammonium hydroxide) conditions. Further steps of extraction for sonication, centrifugation and filtration were the same as Section 4.3. Each sample pellet passed 3 rounds of extraction and peak areas of target analytes in each round were used for further interpretation of the extraction efficiency. LLE was performed using a MTBE/MeOH/water solvent system (here termed

MTBE method) and a chloroform/MeOH/water solvent system (here termed chloroform method) according to a previous study.<sup>32</sup> However, slight modifications were implemented to obtain a better solvent ratio (3.6/2.8/3.5, v/v/v) applicable for both polar and apolar metabolites in complex fecal texture. Briefly, 75% ice cold MeOH (5.4  $\mu$ L/mg feces of MeOH and 1.8  $\mu$ L/mg feces of ultrapure water) was added to 0.50 g of freeze-dried sample in a 15 mL Falcon tube (Corning® CentriStar™, VWR International B.V, Amsterdam, The Netherlands). Following homogenization (as in Section 4.2), for the MTBE method, 3  $\mu$ L/mg feces of MTBE were added, and samples were vortex-mixed for 2 min at high speed. Each monophasic supernatant was transferred to a clean tube, after which phase separation was induced by adding 4.2  $\mu$ L/mg feces of MTBE and 5  $\mu$ L/mg feces ultrapure water. Samples were vortex-mixed for 1 min, incubated at room temperature for 10 min and centrifuged for 20 min (14,700 $\times$  g, 10 °C). After precipitation of proteins and debris at the bottom of the tube, the polar (lower layer) and non-polar (upper layer) fractions were transferred into different tubes and dried using a nitrogen blow-down evaporator (Liebisch Labortechnik, Bielefeld, Germany). The residues of lower and upper layers were transferred to separate tubes then reconstituted with 2 mL acetonitrile: water (9:1 v/v) for both targeted reversed phase and HILIC chromatography. The chloroform method followed the MTBE method ratios with the difference of replacing MTBE by chloroform. Due to the higher density of chloroform, the polar fractions stayed at the top and non-polar fraction at the bottom, while protein and debris precipitated between the layers as an intermediate.

For untargeted metabolomics, single-phase extractions were used in neutral pH only, while LLE extraction polar and apolar layers were combined in a tube before being subjected to evaporation and the residue was reconstituted with 2 mL water and acetonitrile (9:1 v/v).

It is worth mentioning that in order to increase the quenching and minimize the residual enzymatic activity, all solvents for extraction in this study were used ice-cold.

### **Assessment of Extraction Repeatability and Recovery**

Only MTBE, chloroform and EtOH extraction methods that yielded the highest number of features in untargeted analysis were included in the repeatability and recovery assessment. To determine the extraction repeatability, the coefficient of variation (CV) was calculated for three replicates of the peak areas of 15 stable isotope labeled standards (Supplementary Table S1) spiked at a concentration of 5  $\mu$ M (apolar standards) and 10  $\mu$ M (polar standards) in 0.50 g of fecal sample.

The recovery was calculated according to the formula below:

$$\% \text{recovery} = (\text{Area A}) / (\text{Area B}) \times 100$$

(1) where A represents the peak area of a stable-isotope labeled standard spiked before extraction and B represents the peak area of a stable-isotope labeled standard in an extracted matrix sample spiked post-extraction. The results reported for the CV value and recovery were calculated using the mean value obtained across three different extractions tested for each method.

Before comparing the extraction protocols, five blanks and five standard quality control solutions were injected at the start of the analytical sequence to assure protocol reproducibility, controlling the performance of the UHPLC-MS system.

### **Data Pre-Processing and Software**

For targeted peak picking, the LC-MS data were processed with the TargetLynx application in the Masslynx software (Waters, version 4.1). For untargeted analysis, the raw data sets were converted to mzML format using ProteoWizard software version 3.0<sup>38</sup>; a retention time cut-off at 600 s and peak picking at MS level 1 was applied during the conversion. The converted data were then processed using XCMS (version 3.8.2) following an optimization of parameters (Supplementary Table S5). Adducts and isotopes were grouped into a single feature with the CAMERA package.<sup>39</sup> For recognition and removal of erroneous features in the datasets, MS-FLO (<http://msflo.fiehnlab.ucdavis.edu> (accessed on 9 May 2021)) was used. The resulting XCMS table was further processed using the steps explained in Table S4, to preserve only peaks eluting before 7 min, with  $s/n > 10$ , peaks present in at least 2/3 technical replicates per sample, and at least 5-fold higher intensity in biological samples than the blank procedure (per preparation method).

For data processing and visualization, Python (version 2.7.15) and R (version 3.6.0) were utilized. R packages “tidyverse” and “ggplot2” were used to format the data and plot the figures, respectively. To obtain an overview of the metabolomic data, abundance profiles of metabolites were log transformed and subjected to principal component analysis (PCA) and dendrogram tree clustering (Pearson distance and average linkage) using MetaboAnalyst version 5.0 (<https://www.metaboanalyst.ca> (accessed on 20 March 2021)). Venn diagrams were drawn via an online website at (<http://bioinfogp.cnb.csic.es/tools/venny/index.html> (accessed on 15 March 2021)).

### **FUNDING**

This project was cofounded by the University of Applied Sciences, Leiden. This research was part of the Netherlands X-omics Initiative and partially funded by NWO, project 184.034.019. This project was cofounded by the ZonMW Memorabel Grant 733050814 (Understanding the Role of the Gut Microbiome in the Pathogenesis and Prevention of

Dementia) and NIA U19 Grant 1U19AG063744-01 (Alzheimer's Gut Microbiome Project National Institute on Aging).

#### **INSTITUTIONAL REVIEW BOARD STATEMENT**

Ethical review and approval were waived for this study, since the collected samples are pooled and the data are reported in grouped format (counts and percentages), such that participants are not able to be identified from the results. Informed consent was obtained from all subjects at the time of original data collection.

#### **INFORMED CONSENT STATEMENT**

Informed consent was obtained from all subjects involved in the study.

#### **DATA AVAILABILITY STATEMENT**

The data presented in this study are available in Supplementary Materials.

#### **ACKNOWLEDGMENTS**

We would like to thank Michael van Vliet, Faiza Guled and Gerwin Spijksma for their technical support.

#### **CONFLICTS OF INTEREST**

The authors declare no conflict of interest.

## REFERENCES

1. Rooks, M. G. & Garrett, W. S. Gut microbiota, metabolites and host immunity. *Nat. Rev. Immunol.* **16**, 341–352 (2016).
2. Olive, A. J. & Sassetti, C. M. Metabolic crosstalk between host and pathogen: Sensing, adapting and competing. *Nat. Rev. Microbiol.* **14**, 221–234 (2016).
3. Hosseinkhani, F. *et al.* The contribution of gut bacterial metabolites in the human immune signaling pathway of non-communicable diseases. *Gut Microbes* **13**, 1–22 (2021).
4. Zierer, J. *et al.* The fecal metabolome as a functional readout of the gut microbiome. *Nat. Genet.* **50**, 790–795 (2018).
5. Weir, T. L. *et al.* Stool Microbiome and Metabolome Differences between Colorectal Cancer Patients and Healthy Adults. **8**, (2013).
6. Yan, R. *et al.* Fecal Metabolites Were Altered, Identified as Biomarkers and Correlated With Disease Activity in Patients With Systemic Lupus Erythematosus in a GC-MS-Based Metabolomics Study. *Front. Immunol.* **11**, 1–11 (2020).
7. Nagpal, R. *et al.* Human-origin probiotic cocktail increases short-chain fatty acid production via modulation of mice and human gut microbiome. *Sci. Rep.* **8**, 1–15 (2018).
8. Reddavid, R. *et al.* The role of diet in the prevention and treatment of inflammatory bowel diseases. *Acta Biomed.* **89**, 60–75 (2018).
9. Karu, N. *et al.* Analytica Chimica Acta A review on human fecal metabolomics : Methods , applications and the human fecal metabolome database. *Anal. Chim. Acta* **1030**, 1–24 (2018).
10. Fiori, J., Turrone, S., Candela, M. & Gotti, R. Journal of Pharmaceutical and Biomedical Analysis Assessment of gut microbiota fecal metabolites by chromatographic targeted approaches. *J. Pharm. Biomed. Anal.* **177**, 112867 (2020).
11. Blekhman, R. *et al.* Common methods for fecal sample storage in field studies yield consistent signatures of individual identity in microbiome sequencing data. *Sci. Rep.* **6**, 1–5 (2016).
12. Bundgaard-Nielsen, C., Hagstrøm, S. & Sørensen, S. Interpersonal Variations in Gut Microbiota Profiles Supersedes the Effects of Differing Fecal Storage Conditions. *Sci. Rep.* **8**, 1–9 (2018).
13. Carroll, I. M., Ringel-Kulka, T., Siddle, J. P., Klaenhammer, T. R. & Ringel, Y. Characterization of the Fecal Microbiota Using High-Throughput Sequencing Reveals a Stable Microbial Community during Storage. *PLoS One* **7**, 1–7 (2012).
14. Liang, Y. *et al.* Systematic Analysis of Impact of Sampling Regions and Storage Methods on Fecal Gut Microbiome and Metabolome Profiles. *mSphere* **5**, 1–13 (2020).
15. Smirnov, K. S. *et al.* Challenges of metabolomics in human gut microbiota research. *Int. J. Med. Microbiol.* **306**, 266–279 (2016).
16. Xu, J., Zhang, Q., Zheng, J., Yuan, B. & Feng, Y. Trends in Analytical Chemistry Mass spectrometry-based fecal metabolome analysis. *Trends Anal. Chem.* **112**, 161–174 (2019).
17. Wilson, A. *et al.* Trimethylamine-N-oxide: A Novel Biomarker for the Identification of Inflammatory Bowel Disease. *Dig. Dis. Sci.* **60**, 3620–3630 (2015).
18. Bobeldijk, I. *et al.* Quantitative profiling of bile acids in biofluids and tissues based on accurate mass high resolution LC-FT-MS: Compound class targeting in a metabolomics workflow. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **871**, 306–313 (2008).
19. Buijink, M. R. *et al.* The influence of neuronal electrical activity on the mammalian central clock metabolome. *Metabolomics* **14**, 1–11 (2018).
20. Connors, J., Dawe, N. & Van Limbergen, J. The role of succinate in the regulation of intestinal inflammation. *Nutrients* **11**, 1–12 (2019).
21. Venegas, D. P. *et al.* Short chain fatty acids (SCFAs) mediated gut epithelial and immune regulation and its relevance for inflammatory bowel diseases. *Front. Immunol.* **10**, (2019).
22. Lamichhane, S. *et al.* Linking Gut Microbiome and Lipid Metabolism: Moving beyond Associations.

- Metabolites* **11**, 55 (2021).
23. Jain, A., Li, X. H. & Chen, W. N. An untargeted fecal and urine metabolomics analysis of the interplay between the gut microbiome, diet and human metabolism in Indian and Chinese adults. *Sci. Rep.* **9**, 1–13 (2019).
  24. Wishart, D. S. *et al.* HMDB 4.0 : the human metabolome database for 2018. **46**, 608–617 (2018).
  25. Gao, X. *et al.* Metabolite analysis of human fecal water by gas chromatography/mass spectrometry with ethyl chloroformate derivatization. *Anal. Biochem.* **393**, 163–175 (2009).
  26. Garner, C. E. *et al.* Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. *FASEB J.* **21**, 1675–1688 (2007).
  27. Gratton, J. *et al.* Optimized Sample Handling Strategy for Metabolic Profiling of Human Feces. *Anal. Chem.* **88**, 4661–4668 (2016).
  28. Cajka, T. & Fiehn, O. Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics. *Anal. Chem.* **88**, 524–545 (2016).
  29. Deda, O. *et al.* Sample preparation optimization in fecal metabolic profiling. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1047**, 115–123 (2017).
  30. López-Bascón, M. A. *et al.* Comprehensive analysis of pig feces metabolome by chromatographic techniques coupled to mass spectrometry in high resolution mode: Influence of sample preparation on the identification coverage. *Talanta* **199**, 303–309 (2019).
  31. Turroni, S. *et al.* Fecal metabolome of the Hadza hunter-gatherers : a host- microbiome integrative view. *Sci. Rep.* 1–9 (2016) doi:10.1038/srep32826.
  32. Sostare, J. *et al.* Analytica Chimica Acta Comparison of modified Matyash method to conventional solvent systems for polar metabolite and lipid extractions. *Anal. Chim. Acta* **1037**, 301–315 (2018).
  33. Patterson, R. E. Comparison of blood plasma sample preparation methods for combined LC-MS lipidomics and metabolomics. *J Chromatogr B Anal. Technol Biomed Life Sci* **176**, 139–148 (2016).
  34. Starke, R. *et al.* Incomplete cell disruption of resistant microbes. *Sci. Rep.* **9**, 1–5 (2019).
  35. Liu, R. *et al.* Analytica Chimica Acta Evaluation of two-step liquid-liquid extraction protocol for untargeted metabolic profiling of serum samples to achieve broader metabolome coverage by UPLC-Q-TOF-MS. *Anal. Chim. Acta* **1035**, 96–107 (2018).
  36. Whiley, L., Godzien, J., Ruperez, F. J., Legido-quigley, C. & Barbas, C. In-Vial Dual Extraction for Direct LC-MS Analysis of Plasma for Comprehensive and Highly Reproducible Metabolic Fingerprinting. (2012).
  37. Sitnikov, D. G., Monnin, C. S. & Vuckovic, D. Systematic Assessment of Seven Solvent and Solid-Phase Extraction Methods for Metabolomics Analysis of Human Plasma by LC-MS. *Nat. Publ. Gr.* 1–11 (2016) doi:10.1038/srep38885.
  38. Chambers, M. C. *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).
  39. Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T. R. & Neumann, S. CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal. Chem.* **84**, 283–289 (2012).

Supplementary Table S1. Standards information							
Mix 1-HILIC				Mix 2-reversed phase			
Compound	CHEBI	Conc. (uM)	Supplier	Compound	CHEBI	Conc. (uM)	Supplier
Aspartic acid	17053	80.0	Sigma-Aldrich	FA (14:0)	140940	50	Merck
Betaine	17750	150	Sigma-Aldrich	FA (16:0)	140943	100	Sigma
Citrulline	16349	50	Sigma-Aldrich	FA (16:1)	140944	25	Fluka
Glutamic acid	16015	200	Sigma-Aldrich	FA (17:0)	140945	4	Sigma-Aldrich
Glutamine	17061	500	Fluka	FA (18:0)	140947	25	Sigma
Gamma-aminobutyric acid	16865	4	Sigma-Aldrich	FA (18:1)	140948	25	Sigma
Leucine	15603	100	Fluka	FA (18:2)	140949	25	Sigma
Isoleucine	17191	100	Fluka	FA (20:0)	140951	20	Sigma-Aldrich
Lysine	18019	400	Fluka	FA (20:2)	140952	10	Sigma-Aldrich
Ornithine	15729	100	Sigma-Aldrich	FA (20:4)	132539	25	Sigma
pipecolic acid	17964	10	Sigma-Aldrich	FA (20:5)	132540	15	Cayman chemicals
Pyroglutamic acid	16010	40	Sigma-Aldrich	FA (22:0)	140958	20	Sigma-Aldrich
Taurine	15891	200	Fluka	FA (22:1)	132541	20	Sigma-Aldrich
Adenosine	16335	0.5	Fluka	FA (22:6)	132544	20	Cayman chemicals
Adenine	16708	0.5	ALFA AESAR	FA (24:0)	155816	20	Sigma-Aldrich
Guanine	16235	4	Sigma-Aldrich	UDCA	9907	100	Sigma
Hypoxanthine	17368	100	Sigma-Aldrich	DCA	28834	100	Brunschwig
Inosine	17596	4	Sigma-Aldrich	CA	16359	100	Brunschwig
Uric acid	27226	1000	Sigma-Aldrich	GCDCA	36274	100	Sigma
Uracil	17568	4	Sigma-Aldrich	GDCA	27471	100	Sigma
Lactic acid	422	1500	Acros organics	TCDCa	16525	100	Sigma
Succinic acid	15741	50	Sigma-Aldrich	TDCA	9410	100	Sigma
Pyruvic acid	32816	100	Sigma-Aldrich	TCA	28865	100	Sigma
Pantothenic acid	7916	6	Sigma-Aldrich	SN1-LPC (18:1)	64566	40	Avanti
Butyric acid	30772	2	Sigma-Aldrich	SN1-LPC (18:2)	64549	40	Avanti
5-hydroxyhexanoic acid	131434	6	Sigma-Aldrich	LPE (18:0)	64576	20	Avanti
<b>LPE (17:1)</b>	—	5	Sigma-Aldrich	LPE (18:2)	91296	20	Avanti
<b>DCA-d4</b>	—	5	Sigma-Aldrich	<b>D2-Glycine</b>	—	10	CDN Isotopes
<b>CA-d4</b>	—	5	CDN Isotopes	<b>D3 Leucine</b>	—	10	CDN Isotopes
<b>FA 20(4)-d8</b>	—	5	Sigma Aldrich	<b>D4 succinate</b>	—	10	Cambridge isotope laboratories
<b>FA 22 (6)-d5</b>	—	5	Sigma-Aldrich	<b>U13-C5-valine</b>	—	10	Cambridge isotope laboratories
<b>D5-TUDCA</b>	—	5	Sigma-Aldrich	<b>D6 Ornithine</b>	—	10	CDN Isotopes
<b>D4-GDCA</b>	—	5	Sigma-Aldrich	<b>U 13C6- Lysine</b>	—	10	Cambridge isotope laboratories
<b>FA18(2) d4</b>	—	5	Cambridge isotope laboratories	<b>D3-9-15N-aspartate</b>	—	10	Cambridge isotope laboratories

<b>Supplementary Table S2. LC MS conditions</b>			
<b>Mode of separation</b>	<b>Analytical separation condition</b>	<b>Solvents</b>	<b>MS parameters</b>
Reversed phase-untargeted ( <b>method 1</b> ) Water ACQUITY UPLC T3 column (2.1 × 100 mm, 1.8 μm)	0-3.5 min: 0.1 to 30% B, 3.5-8 min: 99.9% B, 8-12 min: 99.9% B, 12.1-15 min: 0.1% B Flow rate: 0.4 mL/ min Column temperature: 40°C	A: water with 0.1% formic acid B: MeCN with 0.1% formic acid	Capillary voltage: 1500 V, sample cone voltage: 30 V, source temperature: 125 °C, desolvation temperature: 300 °C, desolvation gas: 500 L/h
Reversed phase- lipids ( <b>method 2</b> ) Water ACQUITY UPLC T3 column (2.1 × 100 mm, 1.8 μm)	0-7.5 min: 0.1 to 99.9% B, 7.5-11.5 min: 99.9% B, 11.6 - 15 min: 0.1% B Flow rate: 0.4 mL/ min Column temperature: 45°C	A: 5:95 MeCN: water with 10 mM Ammonium formate B: MeOH with 10 mM Ammonium formate	
HILIC-polar targeted ( <b>method 3</b> ) SeQuant® ZIC-cHILIC (2.1 × 100 mm, 3 μm)	0-9.16 min: 0.1 to 75% B, 9.16-14 min 100% B, 14.1-20 min: 0.1% B Flow rate: 0.25 mL/ min Column temperature: 30°C	A: 90% MeCN and 10% water with 5 mM ammonium formate B: 10% MeCN and 90% water with 5 mM ammonium formate	Capillary voltage: 2000 V, sample cone voltage: 45 V, source temperature: 80 °C, desolvation temperature: 250°C, desolvation gas: 500 L/h



Compounds	Platform	MS-ready monoisotopic mass (Da)	Theoretical mass [M-H]	Detected m/z	Retention time (min)	LOD (µM)	Correlation Coefficient	IS correction
Citrulline	Targeted-HILIC	175.0956878	174.0884	174.09	10.72	6.25	0.9915	U 13C6- Lysine
GABA		103.063325	102.0561	102.056	8.73	6.25	0.9934	
Glutamic acid		147.0531613	146.0459	146.047	8.8	12.5	0.9902	2,3,3-D3-leucine
Glutamine		146.0691457	145.0619	145.063	9.72	0.5	0.9924	U 13C6- Lysine
Isoleucine		131.0946251	130.0874	130.089	6.19	6.25	0.9964	2,3,3-D3-leucine
Leucine		131.0946251	130.0874	130.089	5.93	6.25	0.9964	
Lysine		146.1055242	145.0983	145.093	9.86	25	0.9922	U 13C6- Lysine
Ornithine		132.0898741	131.0826	131.085	10.73	6.5	0.995	
Taurine		125.014666	124.0074	124.007	8.49	6.25	0.9974	2,3,3-D3-leucine
Glucose		180.06339	179.0561	179.054	7.67	312.5	0.9979	
Guanine		151.04941	150.0421	150.041	5.59	1	0.9968	D2-Glycine
Inosine		268.080773	267.0735	267.073	5.5	0.5	0.9907	
Adenosine		267.0967504	266.0895	266.089	3.76	0.01	0.9968	2,2,3,3-D4-Succinate
Uric acid		168.0283365	167.0211	167.033	4.62	62.5	0.9915	
Aspartic acid		133.0375112	132.0302	132.03	8.82	20	0.9921	2,3,3-D3-leucine
Lactic acid		90.031695	89.0244	89.027	5.45	3	0.995	
Pyruvic acid		88.016045	87.0088	87.006	3.92	25	0.9984	FA (18:2)-D4
FA (16:0)	Targeted-Reversed phase	256.2402303	255.233	255.232	6.35	6.3	0.9979	
FA (18:0)		284.2715304	283.2643	283.262	7.31	6.3	0.9923	
FA (18:1)		282.2558803	281.2486	281.2485	6.61	6.3	0.9994	
FA (18:2)		280.2402303	279.233	279.2319	5.95	6.3	0.9963	
FA (20:4)		304.2402303	303.2329	303.2316	5.95	1.9	0.9917	
FA (20:5)		302.22458	301.2176	301.2159	5.35	0.3	0.9996	
FA (22:6)		328.24023	327.233	327.2362	5.86	0.3	0.9988	
DCA		392.292659	391.2859	391.2854	3.14	0.0018	0.9976	
CA		408.2875744	407.2809	407.2805	2.38	0.0018	0.9976	
GCDCA		449.3141234	448.3075	448.3133	2.11	0.014	0.9979	
GDCA		449.3141234	448.3075	448.3105	2.26	0.007	0.9989	
TDCA		499.296758	498.2893	498.2917	2.01	0.0018	0.9968	
TDCA		499.296758	498.2893	498.2907	2.14	0.0005	0.9978	
TCA		515.29167	514.2836	514.2863	1.63	0.00035	0.9964	

Supplementary Table S4. Calibration range used in developing targeted HILIC and Reversed phase platforms											
Targeted- HILIC			Calibration line concentration (µM)								
Metabolite	Formula	HMDB-Blood conc. Rounded (µM)	Cal-8	Cal-7	Cal-6	Cal-5	Cal-4	Cal-3	Cal-2	Cal-1	
Citrulline	C6H13N3O3	25	400	200	100	50	25	12.5	6.25	3.125	
GABA	C4H9NO2	0.25	32	16	8	4	2	1	0.5	0.25	
Glutamic acid	C5H9NO4	100	1600	800	400	200	100	50	25	12.5	
Glutamine	C5H10N2O3	500	8000	4000	2000	1000	500	250	125	62.5	
Isoleucine	C6H13NO2	50	800	400	200	100	50	25	12.5	6.25	
Leucine	C6H13NO2	50	800	400	200	100	50	25	12.5	6.25	
Lysine	C6H14N2O2	200	3200	1600	800	400	200	100	50	25	
Ornithine	C5H12N2O2	50	800	400	200	100	50	25	12.5	6.25	
Taurine	C2H7NO3S	100	1600	800	400	200	100	50	25	12.5	
Glucose	C6H12O6	5000	40000	20000	10000	5000	2500	1250	625	312.5	
Guanine	C5H5N5O	0.3	32	16	8	4	2	1	0.5	0.25	
Inosine	C10H12N4O5	0.3	32	16	8	4	2	1	0.5	0.25	
Adenosine	C10H13N5O4	0.3	16	8	4	2	1	0.5	0.25	0.01	
Uric acid	C5H4N4O3	500	8000	4000	2000	1000	500	250	125	62.5	
Aspartate	C4H7NO4	20	640	320	160	80	40	20	10	5	
Lactate	C3H6O3	1500	12000	6000	3000	1500	750	375	187.5	93.75	
Pyruvate	C3H4O3	50	1600	800	400	200	100	50	25	12.5	
Targeted- Reversed phase			Calibration line concentration (µM)								
Metabolite	Formula	HMDB-Blood conc. Rounded (µM)	Cal-8	Cal-7	Cal-6	Cal-5	Cal-4	Cal-3	Cal-2	Cal-1	
FA (16:0)	C16H32O2	100	800	400	200	100	50	25	12.5	6.25	
FA (18:0)	C18H36O2	50	400	200	100	50	25	12.5	6.25	3.125	
FA (18:1)	C18H34O2	100	800	400	200	100	50	25	12.5	6.25	
FA (18:2)	C18H32O2	100	800	400	200	100	50	25	12.5	6.25	
FA (20:4)	C20H32O2	15	120	60	30	15	7.5	3.75	1.875	0.938	
FA (20:5)	C20H30O2	0.5	4	2	1	0.5	0.25	0.125	0.0625	0.031	
FA (22:6)	C22H32O2	2	16	8	4	2	1	0.5	0.25	0.125	
DCA	C24H40O4	0.3	0.12	0.06	0.03	0.015	0.0075	0.0037	0.0018	9E-04	
CA	C24H40O5	0.7	0.12	0.06	0.03	0.015	0.0075	0.0037	0.0018	9E-04	
GCDA	C26H43NO5	0.06	0.9	0.45	0.225	0.127	0.063	0.0318	0.0159	0.007	
GDCA	C26H43NO5	NA	0.48	0.24	0.12	0.06	0.03	0.015	0.0075	0.004	
TCDA	C26H45NO6S	0.3	0.032	0.016	0.008	0.004	0.002	0.001	0.0005	3E-04	
TDCA	C26H45NO6S	0.06	0.12	0.06	0.03	0.015	0.0075	0.0037	0.0018	9E-04	
TCA	C26H45NO7S	0.3	0.05	0.025	0.0125	0.006	0.003	0.0015	0.0007	4E-04	

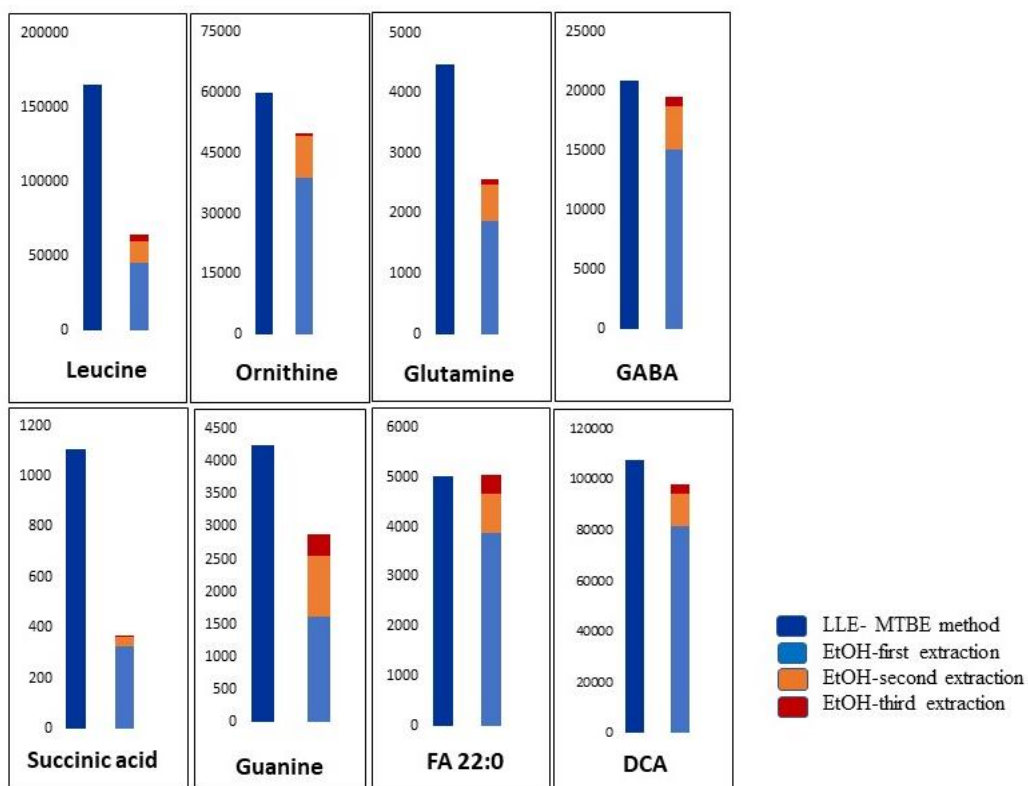
**Supplementary Table S5.** List of parameters used for untargeted data processing with XCMS

<b>Feature detection</b>	
Method	CentWave
Ppm	5
snthr	10
peakwidth	(3, 8)
mzdiff	0.01
prefilter peaks	3
prefilter intensity	1000
noise	1000
<b>Retention time correction</b>	
method	orbiwarp
profStep	1
<b>Grouping</b>	
method	density
bw	5
mzwid	0.01
minfrac	0.5
minsample	3
<b>Annotation</b>	
feature.Annotation.CAMERA.annotate	isotopes
feature.Annotation.CAMERA.mzabs	0.01
feature.Annotation.CAMERA.ppm	5
feature.Annotation.CAMERA.sigma	6
feature.Annotation.CAMERA.perfwhm	0.6
feature.Annotation.CAMERA.maxcharge	3
feature.Annotation.CAMERA.maxiso	5
feature.Annotation.CAMERA.intensity	Maxo

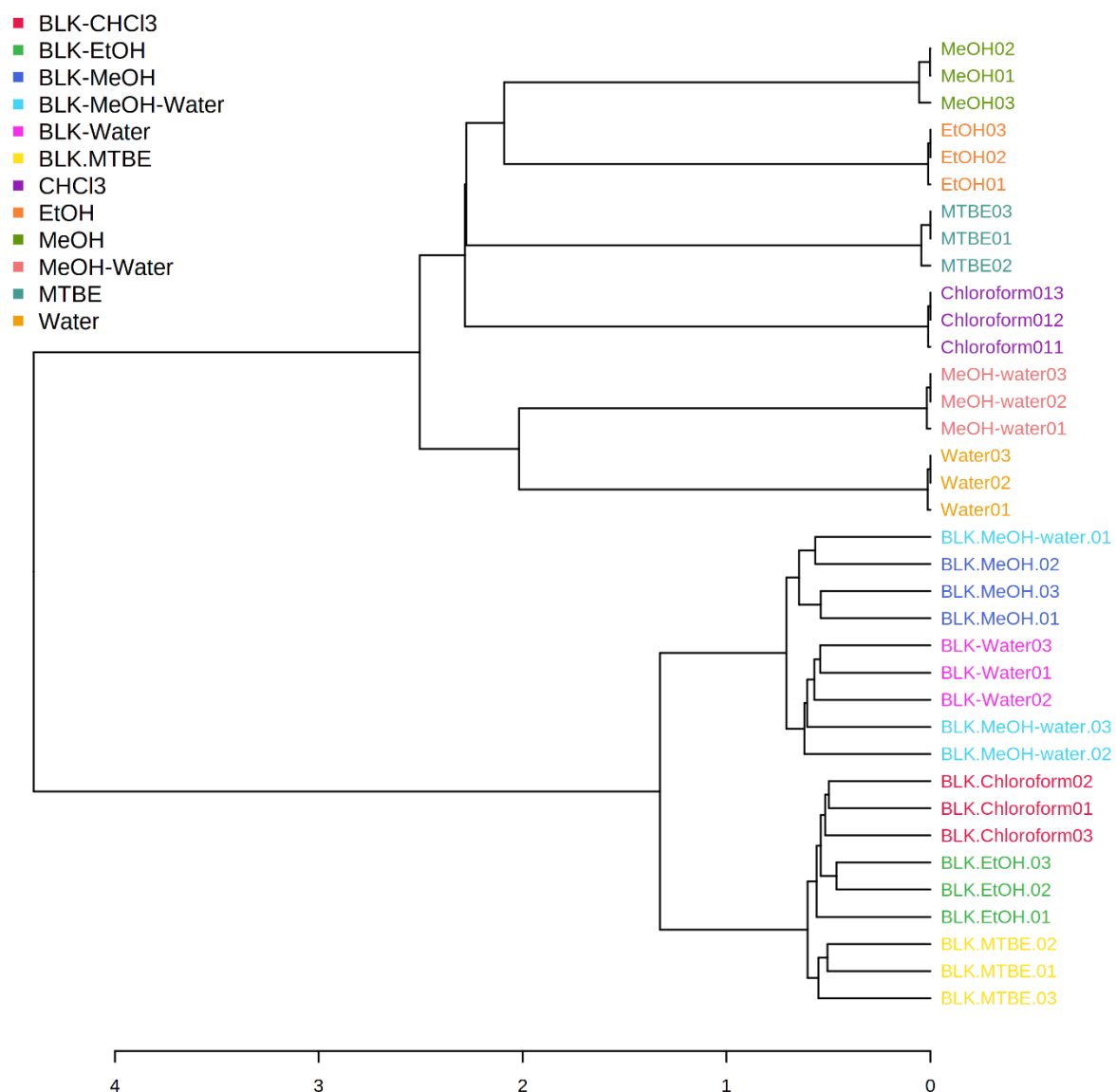
**Supplementary Table S6.** Process of untargeted feature filtering. The number detailed per extraction in each step denotes the features removed in that step.

Cleaning process	MTBE	Chloroform	EtOH	MeOH	MeOH -Water	Water
Total features detected by XCMS	9830					
Removal of features with $rt > 7\text{min}$	2088	2088	2086	2088	2086	2086
Removal of $m/z > 900$	718	718	718	718	718	718
Removal of isotopes	1237	1213	1216	1211	1179	1112
removal of non-biological peaks (by blank)	293	448	402	416	550	847
Features with $RSD > 35\%$ *	267	217	260	293	271	264
consolidating features per retention time	3208	3047	3067	3038	2960	3604
Final features	2109	2099	2081	2066	2066	1199

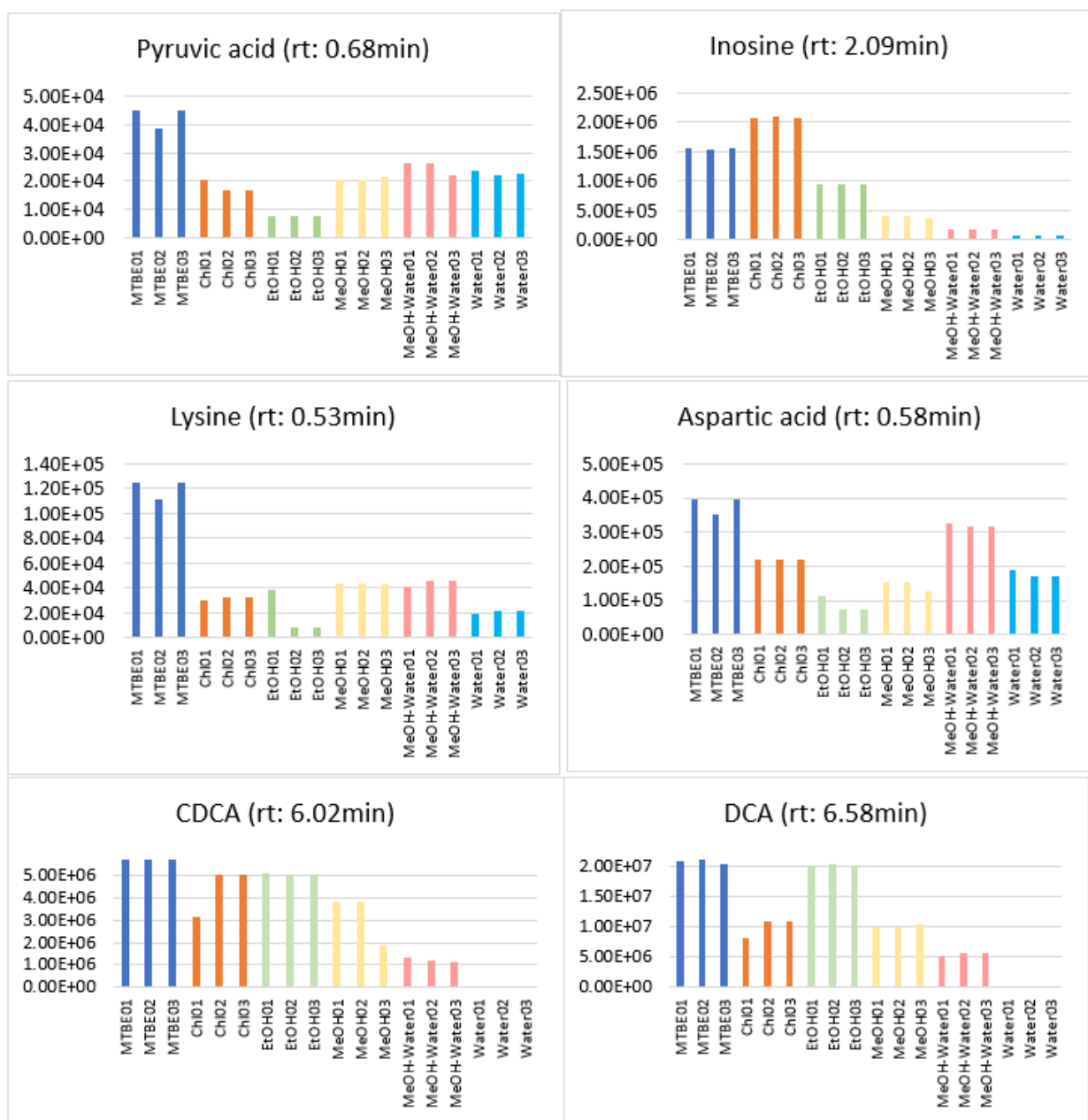
\*In a case of features with low intensity, no removal based on %RSD was applied



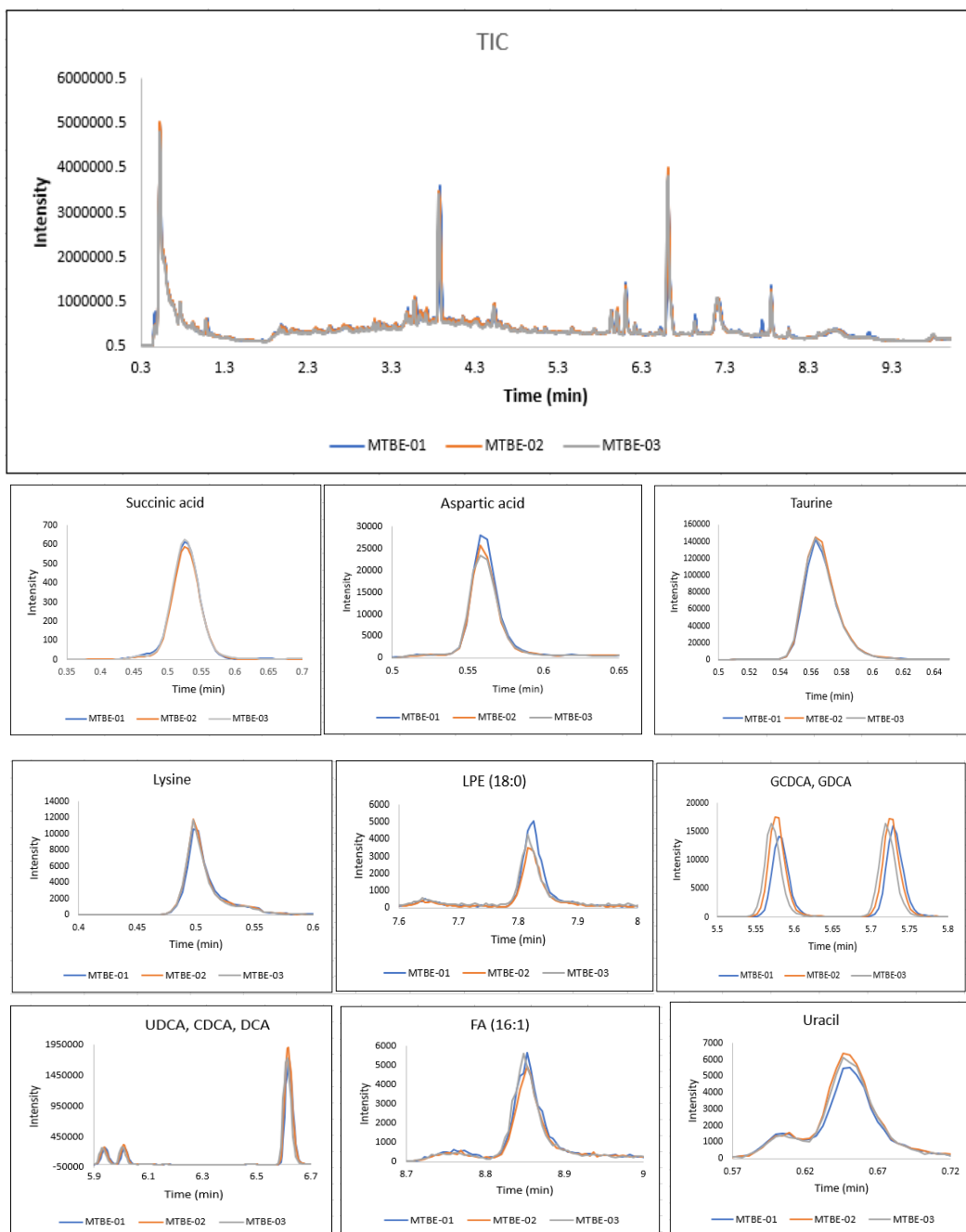
**Supplementary Figure S1.** Comparison of the LLE metabolic yield between MTBE and three cycles of single-phase ethanol extraction (accumulated bar).



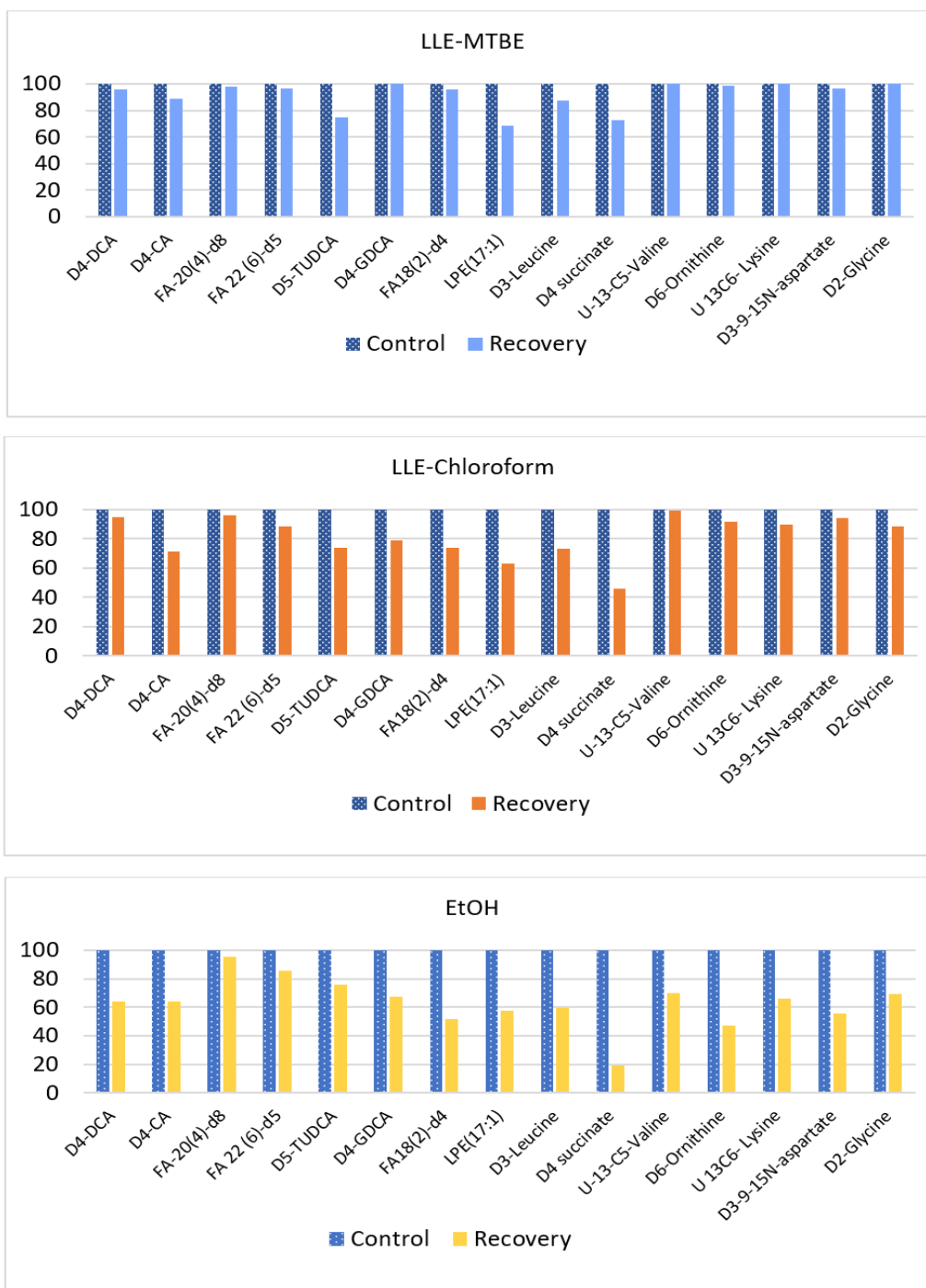
**Supplementary Figure S2.** Hierarchical analysis of the various extraction procedures, with  $n=3$  replicates each. Clustering was performed based on the Pearson distance and average linkage, on log-transformed and pareto-scaled data. BLK, blank.



**Supplementary Figure S3.** Example of the peak area between different extraction methods (LLE versus single phase extraction) in untargeted analysis. Organic solvents exhibit low yield of polar metabolites in the early part of chromatography, and high(er) yield of late eluters.



**Supplementary Figure S4.** The overlaid total ion chromatograms and extracted ion chromatograms of 3 different MTBE extractions on the same fecal sample



**Supplementary Figure S5.** Extraction recovery (% of starting material) for internal standards, following extraction with EtOH, chloroform and MTBE. The control sample represents the concentration of internal standard spiked post extraction.



