analytical chemistry

Article

pubs.acs.org/ac

Validated High Resolution Mass Spectrometry-Based Approach for Metabolomic Fingerprinting of the Human Gut Phenotype

Julie Vanden Bussche,*[,†] Massimo Marzorati,[‡] Debby Laukens,[⊥] and Lynn Vanhaecke[†]

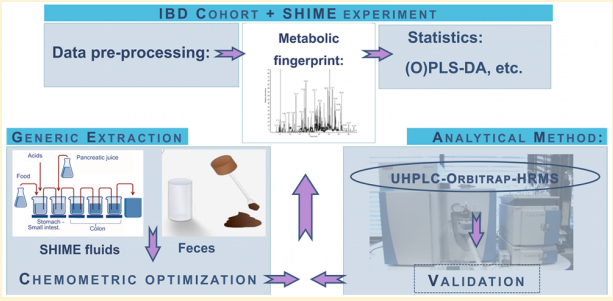
[†]Laboratory of Chemical Analysis, Ghent University, Merelbeke, 9820, Belgium

[‡]Laboratory of Microbial Ecology and Technology, Ghent University, Ghent, Belgium

[⊥]Department of Gastroenterology, Ghent University, Ghent, 9000, Belgium

Supporting Information

ABSTRACT: Fecal samples are an obvious choice for metabolomic approaches, since they can be obtained non-invasively and allow one to study the interactions between the gut microbiota and the host. The use of ultrahigh performance liquid chromatography hyphenated to Orbitrap high-resolution mass spectrometry (UHPLC-Orbitrap HRMS) in this field is unique. Hence, this study relied on Orbitrap HRMS to develop and validate a metabolic fingerprinting workflow for human feces and *in vitro* digestive fluids. After chemometric sample extraction optimization, an aqueous dilution appeared necessary to comply to the dynamic range of the MS. The method was proven "fit-for-purpose" through a validation procedure that monitored endogenous metabolites in quality control samples, which displayed in both matrices an excellent linearity (R^2 > 0.990), recoveries ranging from 93% to 105%, and precision with coefficients of variation (CVs) < 15%. Finally, feces from 10 healthy individuals and 13 patients diagnosed with inflammatory bowel disease were subjected to metabolomic fingerprinting. 9553 ions were detected, as well as differentiating profiles between Crohn's disease and ulcerative colitis by means of (orthogonal) partial least-square analysis ((O)PLS)-DA (discriminate analysis) models. Additionally, samples from the dynamic gastrointestinal tract simulator (SHIME (Simulator of the Human Intestinal Microbial Ecosystem) platform) were analyzed resulting in 6446 and 5010 ions for the proximal and distal colonic samples, respectively. Supplementing SHIME feed with antibiotics resulted in a significant shift (P < 0.05) of 27.7% of the metabolites from the proximal data set and 34.3% for the distal one. As a result, the presented fingerprinting approach provided predictive modeling of the gastrointestinal metabolome *in vivo* and *in vitro*, offering a window to reveal disease related biomarkers and potential insight into the mechanisms behind pathologies.



Metabolomics is a holistic method to acquire comprehensive insights in the functioning of a biological system, by unbiased analysis of as many small molecules as possible.[1] Its potential is reflected by recent reports in which it is applied to improve disease diagnosis or prognosis, understand the mechanisms behind pathologies, and increase the understanding toward individual therapeutic responses.[2,3] In this context, an increasing awareness exists that many diseases, which lead to a significant perturbation of metabolism, are intrinsically linked to the gut functional ecology.[4] Due to the close symbiotic association, the microbiome strongly impacts the metabolic phenotype of the host, which results in the hypothesis that metabolic readouts can give insights into functional metagenomic activity.[5,6]

These metabolic readouts can be performed trough a targeted (profiling) or an untargeted (fingerprinting) approach. Targeted metabolomics investigates the quantitative changes of predefined metabolites by employing authentic analytical standards, whereas untargeted metabolomics intends to reveal fingerprints of all metabolites monitored or biomarkers derived thereof. In both cases, the analytical technique requires high specificity and sensitivity. For this reason, high-resolution mass spectrometry-based techniques (HRMS) have become common practice in the field of metabolomics.[7] Moreover, the hyphenation of HRMS with high performance chromatography provides the benefit of reducing potential ion suppression.[3,8]

An important but often forgotten aspect in metabolomic analysis is the necessity of delivering reliable and reproducible results. Implementing a validation procedure for as many steps as possible in the presented workflow, i.e., sample pretreatment and analysis, data normalization, and biostatistics,[3] is therefore mandatory. Unfortunately, specific guidelines are lacking at the time.

Metabolomics can be applied to any biological matrix. The most frequently used specimens for exploring systemic alterations of metabolites in humans are serum,[9] urine,[10] and

Received: July 17, 2015
Accepted: October 9, 2015
Published: October 9, 2015

DOI: 10.1021/acs.analchem.5b02688
Anal. Chem. 2015, 87, 10927−10934

tissue.[11] More recently, fecal samples are considered as a valuable choice, since they can be obtained noninvasively. Its use circumvents the rather inaccessible nature of the intestinal habitat, and most importantly, it comprises a rich source of information that allows complex interactions between the gut microbiota and the host to be studied.[12−14] Moreover, this matrix enables one to account for the dietary input, as demonstrated by comparing the fecal metabolome from conventional and humanized mouse models.[4,15]

The experimental field investigating microbiome composition related to dysbiosis is shifting toward fecal metabolomics (colorectal cancer,[16,17] inflammatory bowel disease,[18] irritable bowel syndrome).[19] The use of HRMS is however scarce but could lead to the discovery of preliminary biomarkers employed for diagnosis or surveillance of disease. For deepening the knowledge on a suspected underlying physiological pathway, it is possible to target only a part of the metabolome (i.e., submetabolome). For example, Xu and colleagues managed to increase the LC-MS sensitivity of the amine/phenol submetabolome by means of isotopic labeling.[14]

Unfortunately, elevated colonic absorption of the bacterial produced metabolites can result in a loss of information, even when working with fecal samples.[20] Performing metabolomics on fluids derived from *in vitro* gastrointestinal digestion simulation models like static batch incubations[21] or dynamic models like the SHIME (Simulator of the Human Intestinal Microbial Ecosystem) or TIM (TNO Intestinal model) could allow a more in depth investigation of the impact of the gut microbiota and nutrients toward the host metabolism, although experimental data are limited up until now.

Therefore, the goal of this study was to develop and validate a metabolic fingerprinting method that enables one to produce reliable results for the selected biological matrices (human feces and *in vitro* simulated gastrointestinal fluids) in terms of robustness and repeatability. To this extent, ultrahigh performance liquid chromatography hyphenated to Orbitrap HRMS (UHPLC-HRMS) was employed. To ensure broad metabolite coverage, a set of analytical standards of preselected known gastrointestinal metabolites (metabolic profiling) characterized by a varying range of physicochemical properties were included. As part of the validation strategy, generic extraction procedures were developed for the selected biological matrices through a sequential strategy of experimental designs. Finally, fecal metabolic fingerprinting was applied on a cohort of healthy and diseased (inflammatory bowel disease, IBD) individuals to display the methods' discriminating potential toward metabolomic phenotypes. In addition, digestive fluids derived from an *in vitro* SHIME experiment underwent fingerprinting to ascertain the detection of the metabolic shifts upon antibiotic supplementation (i.e., inducing a shift in microbial diversity).

MATERIALS AND METHODS

Reagents and Chemicals. Analytical standards (Table S1, supplementary data) and the internal standard valine-d_8 (ISTD) were purchased from Sigma-Aldrich (St-Louis, Missouri, USA), ICN Biomedicals Inc. (Ohio, USA), or TLC Pharmchem (Vaughan, Ontario, Canada). Solvents were obtained from Fisher Scientific UK (Loughborough, UK) and VWR International (Merck, Darmstadt, Germany). For each standard, a primary stock solution was prepared in methanol or ultrapure water at a concentration of 10 mg/mL and stored at −20 °C.

Instrumentation. An Accela UHPLC system of a Thermo Fisher Scientific (San José, CA, USA) was used for the chromatographic separation of the gastrointestinal (GIT)-derived metabolites, this with an Acquity HSS T3 C18 column (1.8 μm, 150 mm × 2.1 mm, Waters) kept at 45 °C. A vanguard precolumn (1.8 μm, 5 mm × 2.1 mm, Waters) was attached to guarantee a longer column lifetime. As binary solvent system, ultrapure water (A) and acetonitrile (B) both acidified with 0.1% formic acid were used, this at a constant flow rate of 0.4 mL/min. A gradient profile with the following proportions (v/v) of solvent A was applied: 0−1.5 min at 98%, 1.5−7.0 min from 98% to 75%, 7.0−8.0 min from 75% to 40%, 8.0−12.0 min from 40% to 5%, 12.0−14.0 min at 5%, 14.0−14.1 min from 5% to 98%, followed by 4.0 min of re-equilibration. A 10 μL aliquot of each sample was injected for analysis.

HRMS analysis was performed on an Exactive stand-alone benchtop Orbitrap mass spectrometer (Thermo Fisher Scientific, San José, CA, USA), equipped with a heated electrospray ionization source (HESI), operating in polarity switching mode. Ionization source working parameters were optimized and were set to a sheath, auxiliary, and sweep gas of 50, 25, and 5 arbitrary units (au), respectively, heater and capillary temperature of 350 and 250 °C, and tube lens, skimmer, capillary, and spray voltage of 60 V, 20 V, 90 V, and 5 kV (±), respectively. A scan range of m/z 50−800 was chosen, and the resolution was set at 100 000 fwhm at 1 Hz. The automatic gain control (AGC) target was set at balanced (1×10^6 ions) with a maximum injection time of 50 ms.

Biological Samples. *Human Fecal Samples.* A pool of freeze-dried fecal samples ($n = 6$) was used for the development of the generic extraction procedure. Healthy volunteers (three males and three females) who were not subjected to any antibiotic treatment for at least six months prior to sample collection without any dietary restrictions donated these samples. After a maximum of 7 days of storage at −80 °C, the samples underwent 48 h of lyophilization, which resulted in the removal of 73.35 ± 7.38% of water.

To demonstrate the applicability of the fingerprinting approach, 23 additional fecal samples were analyzed (10 females and 13 males, EC 2010/116). This group consisted of 10 healthy volunteers, including four vegetarians, and 13 IBD patients (Crohn's disease: two endoscopically active and three in remission; ulcerative colitis: three endoscopically active and five in remission; remission was defined as clinical remission and extinguished endoscopic inflammation).

SHIME Fluids. A SHIME setup represents the GIT of the adult human.[22] For the sample extraction optimization, fluids of five different proximal colonic vessels ($n = 5$) were pooled after a stabilization period of 4 weeks (LabMET, Ghent University, Belgium). The freshly collected suspensions were preserved in aliquots at −80 °C.

To investigate whether our newly developed fingerprinting method could pick up relevant metabolic changes in this type of matrix, a SHIME experiment was conducted. Antibiotics (Ab) were added to the SHIME feed (i.e., simulation of oral delivery) to induce a shift in microbial diversity and thus in the gut metabolome. Two modified SHIME systems were run in parallel (i.e., "control arm" and "chronic stress arm"), each consisting of a succession of 3 reactors simulating stomach and small intestine (temporal succession), proximal colon (PC), and distal colon (DC) as described by Worametrachanon et al.[22] The colonic compartments were inoculated with fecal

DOI: 10.1021/acs.analchem.5b02688
Anal. Chem. 2015, 87, 10927−10934

Table 1. Statistical Outcome (*P*-value) of the Fractional Factorial Design for the Extraction of Human Feces and SHIME Suspensions Based on Quality Control Samples

	acetic acid	alanine	hexanoic acid	maltose	panthotenic acid	spermine	ursodeoxycholic acid
Human Feces							
feces (mg)	0.01	0.02	0.02	0.23	0.11	0.02	0.02
dilute with H_2O	0.74	0.01	0.07	0.02	0.45	0.27	0.49
extraction volume	0.01	0.58	0.04	0.85	0.08	0.03	0.08
solvent type	0.51	0.04	0.30	0.50	0.07	0.59	0.67
vortex + rotating	0.29	0.96	0.29	0.09	0.99	0.32	0.74
repeat extraction	0.03	0.81	0.04	0.09	0.03	0.04	0.01
filtrate aliquot	0.01	0.81	0.42	0.37	0.80	0.80	0.09
SPE	0.01	0.57	0.93	0.89	0.69	0.56	0.09
evaporate org. phase	0.24	0.01	0.08	0.43	0.85	0.65	0.06
evaporation (UV/N_2)	0.51	0.34	0.51	0.87	0.67	0.31	0.86

	acetic acid	acetylglucosamine	azelaic acid	cholic acid	hydroxypyridine	leucine	pentanoic acid
SHIME Fluids							
volume (mL)	0.79	0.71	0.76	0.02	0.43	0.46	0.86
centrifugation time (min)	0.67	0.25	0.68	0.40	0.49	0.77	0.74
filter type (PVDF/polyamide)	0.90	0.53	0.80	0.01	0.63	0.47	0.49
dilution with H_2O	0.00	0.00	0.08	0.00	0.00	0.00	0.00
% of dilution	0.01	0.07	0.15	0.01	0.01	0.01	0.04

microbiota of a healthy 35-year old volunteer who had no history of Ab use in the past six months. The following design was applied for 5 weeks: week 1−2 of stabilization; week 3−4 in which the "chronic stress arm" received twice a week a mild dose of a mix of amoxicillin, ciprofloxacine, and tetracycline (10, 10, and 2.5 ppm, respectively) while the "control arm" remained untreated; week 5 in which both arms received a single shock-dose of a mix of amoxicillin, ciprofloxacine, and tetracycline (40, 40, and 10 ppm, respectively). From each colonic compartment, samples were collected for the metabolic analyses (1.5 mL), twice a week during weeks 3 and 4 and 5 times per week during week 5.

Development and Chemometric Optimization of the Sample Extraction. To develop a generic metabolite extraction procedure for fecal and SHIME suspension samples, this study relied on a sequential strategy of experimental designs. The metabolite coverage and extraction efficiency during the experimental designs were assessed through the absolute peak areas of all detected known unknown endogenous metabolites (Table S1). To simplify the statistical evaluation carried out with Modde 5.0 (Umetrics, Umea, Sweden), the absolute peak areas of seven endogenous metabolites were taken into account. Each metabolite served as representative of a different compound class and was abundantly present. For human feces, alanine (Ala, amino acid), maltose (carbohydrate), panthotenic acid (*N*-compound), ursodeoxycholic acid (UDCA, bile acid), acetic acid (AcOH, short chain fatty acid), hexanoic acid (monocarboxylic acid), and spermine (polyamine) were selected, whereas for SHIME fluid leucine (amino acid), acetylglucosamine (carbohydrate), azelaic acid (multicarboxylic acid), cholic acid (bile acid), AcOH (short chain fatty acid), hexanoic acid (monocarboxylic acid), and hydroxypyridine (*N*-compound) were considered.

Human Fecal Samples. The elimination of microbial activity in fecal samples prior to metabolomic analysis was warranted, by storage of the fecal samples at −80 °C and lyophilization prior to extraction. For optimizing the generic extraction procedure, a pool of lyophilized feces (n = 6) was used. First, the dependent variables (n = 10, Table 1) that might significantly affect the extraction were screened with a fractional factorial design (FFD), requiring 16 experiments. Next, a response surface model (RSM) was applied to further optimize the statistical significant quantitative variables (i.e., mass of feces and volume extraction solvent).

SHIME Fluids. For ensuring a proper column lifetime, it was decided not to commence with a dilute-and-shoot approach immediately.[23] Five extraction parameters (Table 1) were compared during an FFD of 19 experiments for achieving a reliable, repeatable, and generic extraction procedure. No RSM could be performed, as the minimum of two quantitative variables was not reached.

Finalized Extraction Protocol. *Human Fecal Samples.* The generic extraction protocol started with mixing 200 mg of lyophilized homogenized feces with 4 mL of ultrapure water. After 30 s of thorough mixing, 1 mL of a mixture of ice-cold methanol and ultrapure water (80:20) was added. After 1 min of vortexing and 10 min of rotation, the sample was centrifuged at 13 300g for 10 min. The supernatant was collected and passed over a polyamide filter (diameter of 25 mm and pore size of 0.45 μm). After filtration, the extract was diluted (1:3) with ultrapure water and transferred to a glass HPLC-vial.

SHIME Fluids. The optimized protocol consisted of centrifugation (5 min at 13 300g) of 1.5 mL of suspension, followed by filtration of the supernatant over a PVDF filter (diameter of 13 mm and pore size of 0.22 μm). Finally, the extract was diluted (1:5) with ultrapure water and transferred to a glass HPLC-vial.

Validation. To incorporate matrix effects into the validation setup, biological samples, i.e., quality control (QC) samples, were used. These QC samples are considered as representative bulk control samples.[24] During the validation, different performance criteria (i.e., linearity, recovery, and precision) were assessed on the basis of the absolute peak areas of the ISTD and nine known unknown metabolites were endogenously present in the QC samples. The selected metabolites for feces/SHIME were Ala (amino acid), AcOH (short chain fatty acid), pipecolic acid (monocarboxylic acid), spermidine (polyamine), inositol (polyol), panthothenic acid (N-compound), dodecanedioic acid/azelaic acid (multicarboxylic acid), hexanoic acid (monocarboxylic acid)/acetylglucosamine (carbohy-

DOI: 10.1021/acs.analchem.5b02688
Anal. Chem. 2015, 87, 10927−10934

drate), UDCA (bile acid), and the ISTD valine-d_8. The selection of these "reference metabolites" was based on endogenous prevalence in the QC samples (Figure S1) and relevance toward the gut metabolome. For example, SCFAs play a key role in the prevention and treatment of metabolic syndrome, bowel disorders, etc.,[25] whereas bile acids have been recognized to play an important role in gut health.[26] Intestinal polyamines, synthesized by microbiota, are also thought to play a role in gut health.[27] Metabolomic studies investigating dysbiosis have shown alterations in the intestinal fatty acid[28] and amino acid metabolism[18] after decarboxylation produces amines.[29] Then again, carbohydrate intake has been mentioned in the etiology of IBD.[30]

For feces, QC samples were prepared by pooling all 23 human fecal extracts (derived from our IBD cohort) according to our newly developed extraction protocol. The SHIME QC samples consisted of the combined extracts derived from five different proximal colonic vessels.

Linearity. QC samples for both matrices were diluted serially (1, 2, 5, 10, 20, 50, 100, 200, and 500 times) with ultrapure water, and assessment was based on the obtained coefficients of determination (R^2).

Precision. Instrumental precision was investigated by repeatedly injecting a QC sample (n = 10), followed by evaluation of the obtained coefficients of variation (CVs). For the repeatability or intra-assay precision, multiple QC samples (n = 10) were extracted in parallel under the same experimental conditions. Interday assay precision (n = 20) included within-laboratory variation, such as different analyst, days, etc.

Recovery. The recovery was determined by adding 0, 20, 30, and 50 μL of an analytical standard mixture to a QC sample (50 μL), in triplicate. Each sample was standardized to a volume of 100 μL by adding ultrapure water. The mixture contained nine analytes at 5 ng/μL and AcOH at 20 ng/μL. Some compounds in the QC aliquot were however present at such high levels that fortification experiments would lead to a saturation of the mass detector. To overcome this problem and allow proper calculation of the recoveries in the fecal extracts, a 1:5 dilution sample was applied for inositol, UDCA, and Ala, whereas a 1:10 dilution was necessary for pipecolic acid. For the SHIME recovery, a 1:5 dilution was required for inositol and UDCA.

Chemometric Data Analysis. Extensive data preprocessing was conducted on the obtained full scan HRMS data files with Sieve 2.1 software (Thermo Fischer Scientific), including automated peak extraction, peak alignment, deconvolution, and noise removal. This differential analysis was performed separately for the negative and positive ionization mode.

Next, normalization of the data set was performed to take instrumental drift into account. For this purpose, QC samples (pool of all samples to be analyzed) were dispersed evenly across the sample batch and duplicate QC injections occurred after every ten samples. The average signal of those two injections was used for normalizing the ten preceding samples.[24] Samples were injected in a randomized order.

Finally, multivariate regression techniques (Simca 13.5.0, Umetrics, Sweden) were used to display the differentiation between the obtained fingerprints. This study relied on principal component analysis (PCA) to reveal outliers, groups, and trends, whereas (orthogonal) partial least-square analysis ((O)PLS) was used for constructing a prediction model that could explain and predict the Y-variable (disease phenotype) from the X-matrix (fecal fingerprint with metabolite abundances).

RESULTS AND DISCUSSION

Development of the Metabolomic Analysis Method. To ensure the holistic nature of the untargeted metabolic approach, a set of 120 known unknown metabolites was selected and incorporated (metabolic profiling). These metabolites were characterized by a broad range of physicochemical properties (Table S1). The selection of these compounds was based on previous reports on relevant gastrointestinal metabolites.[2,9,31−33]

To optimize the chromatographic separation of the metabolites, different sub-2 μm columns were tested. On the basis of baseline separation of the analytes of interest, their isomers, and the obtained retention times, the Acquity HSS T3 (1.8 μm, 150 × 2.1 mm) column was retained. The aqueous and organic solvents were both supplemented with 0.1% formic acid to avoid ionization of certain compounds prone to a high pH environment such as carboxylic acid.[34] As organic phase, it was decided to use acetonitrile as it allowed a better separation for some of the metabolites and ensured a more complete regeneration of the column.

For the optimization of the MS parameters, analytical standards of the known unknown metabolites were injected and chromatograms with corresponding mass spectra were recorded (Xcalibur qual browser 2.1 and ToxID 2.1.2 software, Thermo Fisher Scientific, San José, USA). The ionization of the nonpolar compound cholesterol was only possible with atmospheric pressure chemical ionization (APCI). Nevertheless, the goal of our metabolomics approach was to detect as many metabolites as possible within a single analytical run, resulting in the selection of HESI. The detection of short chain fatty acids (e.g., AcOH, propionic acid) and monocarboxylic acids (e.g., hexanoic acid, valeric acid) improved upon direct infusion. It appeared crucial for these compounds to lower the capillary temperature from 350 to 250 °C, this combined with an increase of the heater temperature to 350 °C.

After optimization, 111 of the predefined metabolites were detected via UHPLC-Orbitrap-HRMS, providing a 92.5% success rate. The following 9 compounds could not be detected: phenol, p-cresol, cholesterol, glycolic acid, oleic acid, linoleic acid, meso-erythrol, citrazinic acid, and piperidine. In general, phenol and p-cresol detection are described by GC/MS following derivatization.[35,36] Glycolic acid and meso-erythrol could only be detected upon direct infusion. These compounds were most likely too polar to be retained by the HSS T3 column. Oleic acid and linoleic acid, in contrast, are highly nonpolar compounds that were probably not dissolved in the mobile phase and therefore not detected.

Validation Strategy. *Chemometric Optimization of the Extraction.* To ensure the extraction of as many metabolites as possible,[8] an FFD was employed to statistically evaluate the effects of the 10 selected extraction variables (Table 1). In total, 16 experiments were required and analyzed in a random order. After evaluating the absolute peak areas of the seven selected compounds, the following parameters appeared to exert a significant (P-value < 0.05) positive influence on the extraction yield: mass of feces (200 mg), pre-extraction with ultrapure water, extraction volume (1 mL), and repeating extraction (no) (Table 1). The parameter evaporation of organic phase showed no statistical difference but enabled a better baseline separation for some isomeric compounds. This also influenced the number of extracted metabolites; e.g., some multicarboxylic acids were only detected upon inclusion of the evaporation.

DOI: 10.1021/acs.analchem.5b02688
Anal. Chem. 2015, 87, 10927−10934

In the RSM design, the following parameters were further optimized: mass of feces, pre-extraction with ultrapure water, solvent type, and dilution factor. The application of a dilution factor (1:3) was included as an alternative to the evaporation of the organic phase in order to circumvent potential matrix effects that caused improper detection of, e.g., multicarboxylic acids in the FFD. For the statistical evaluation of the extraction parameters, the same seven metabolites were considered. The statistically significant variables (P-value <0.05) resulted in the use of 200 mg feces and pre-extraction with 4 mL of UP water, according to the individual response plots. Evaporation as well as dilution with ultrapure water in combination with filtration (diameter of 25 mm, 0.45 μm) demonstrated a similar decrease in matrix effects. However, to ensure the generic character of the extraction procedure for metabolomic purposes and to avoid loss of volatile metabolites during evaporation, dilution of the extract with ultrapure water was preferred.

SHIME Fluids. Five selected variables (Table 1) were submitted to an FFD design. The following parameters appeared to exert a significant (P-value < 0.05) positive influence on the extraction yield: use of PVDF filter, no dilution with ultrapure water or, in case of dilution, a 50% dilution, while the use of a polyamide filter and dilution showed a negative effect (Table 1).

The negative effect of dilution was to be expected as Modde 5.0 incorporates absolute peak area values, implying an unambiguous decrease caused by the dilution. Nevertheless, this variable was important to suppress peak overload and matrix interferences and improve baseline separation from isomers. Therefore, it was decided to apply a 1:5 dilution of the filtrated sample. No RSM could be performed, as a response surface plot requires at least two quantitative variables.

Analytical Method. Untargeted metabolomics makes use of relative metabolite quantification based on absolute signal intensities for constructing multivariate statistical models such as (O)PLS-DA (discriminate analysis). A key factor in this approach is proving that the generated data are robust and reproducible enabling the correct biological interpretation of specific metabolic differences. As a result, a thoroughly controlled, validated analytical method is required. If the method can fulfill the performance criteria set out for targeted approaches, like linearity, precision (repeatability or intra-assay precision and intermediate precision or interassay precision), and recovery for the selected known unknowns, it can be concluded that the untargeted analytical method is "fit-for-its-purpose".[3] To consider potential (negative) matrix effects, biological samples in the form of QC samples were employed for this validation setup.

For determining the linearity of the fecal extracts, QC samples were serially diluted. All ten compounds showed an excellent linearity translated to coefficients of determination (R^2) >0.990. Spermidine could only be measured until a 5-fold dilution, resulting in a limited calibration curve of 3 points. Unfortunately, the endogenous levels of these metabolites were beyond our control. AcOH and the ISTD (valine-d_8) were detectable until a 50-fold dilution, whereas panthotenic acid, pipecolic acid, and inositol remained present until a 100-fold dilution. Ala, UDCA, 3-(4-hydroxyphenyl) propionic acid, and dodecanedioic acid were detected even at a 500-fold dilution. Despite their high endogenous levels, the R^2 indicated that we were still operating in the linear detection range of the instrument.

Injecting a QC sample derived from human feces multiple times (n = 10) allowed one to investigate the instrumental precision. The CVs for all ten metabolites ranged from 1.3% to 10.3%. Extracting ten QC samples in parallel led to the evaluation of the repeatability with CVs ranging from 3.8% to 4.8%. The interassay precision (n = 20) resulted in CVs from 3.9% to 13.8%. At this time, there are no general accepted criteria involving the repeatability of metabolomic data sets. Our obtained CVs however did meet the Food and Drug Administration (FDA) recommendations. According to the FDA, a CV of 15% from the nominal value is acceptable for a single bioanalytical test, and when operating closely to the limit of quantification, 20% remains acceptable.[37] The calculated recoveries were between 97.2% ± 7.1% and 104.8% ± 8.5% taking a CV range of 0.4−8.1% into account. The individual validation data are summarized in Table S2. The performance criteria for the SHIME matrix were also based on the absolute response of ten compounds (nine endogenous ones and an ISTD). Evidently, the metabolome of SHIME fluid differed from that of feces. For example, the multicarboxylic acid dodecanedioic acid was absent in SHIME fluid and subsequently replaced by azelaic acid while hexanoic acid was replaced by acetylglucosamine. As for the linearity, analyzing undiluted samples resulted in concentrations outside the dynamic range of the HRMS. This problem was resolved when all undiluted samples were excluded. For Ala and acetylglucsosamine, even the 5-fold dilution required exclusion. Finally, all ten compounds obtained R^2 > 0.990 for curves containing a minimum of five dilution levels. The instrumental precision resulted in CVs ranging from 0.4% to 6.7%. Repeatability (n = 10) and interassay precision (n = 20) resulted in CVs ranging from 0.2% to 11.5% and 0.7% to 12.9%, respectively. These performance criteria all meet the recommendations of the FDA (CV < 15%).[37] The recovery amounted between 93.0 ± 12.7% and 101.4 ± 2.9% with a CV ranging from 0.3% to 13.6% (Table S2). The current method validation for both matrices provided excellent performance criteria. Therefore, the data generated by our untargeted metabolomic method was considered reliable and suitable for relative metabolite quantification.[3]

***In Vitro* and *in Vivo* Metabolomic Fingerprinting.** *IBD Cohort.* To demonstrate the applicability of our validated analytical method, fecal samples from healthy individuals and patients with IBD were analyzed. Measuring a broad diversity of metabolites should allow differentiation among the different displayed phenotypes (diseased vs healthy). To increase this diversity, both ionization modes were considered separately. The positive mode allowed us to extrapolate 7002 ions from the raw files, while for the negative mode, 2551 ions were obtained after excluding the isotopic ions.

Prior to applying multivariate statistics, all data were normalized to correct for instrumental drift. Next, a logarithmic transformation and a Paretro scaling ($1/(SD)^{1/2}$, where SD is the standard deviation) were performed for inducing normality and standardizing the range of independent X-variables, respectively.[38] The PCA score plots displayed a good clustering of all QC samples in "both ionization modes" (bIM) (Figure S2). This indicated the reproducibility of our analytical method and strengthened the obtained validation results. Both PCA plots, one per ionization mode, displayed one to two outliers. The Hotelling's T2 plot revealed that for bIM only one of the two potential outliers was a true suspected outlier. For the negative and positive ionization mode, the outlier originated

DOI: 10.1021/acs.analchem.5b02688
Anal. Chem. 2015, 87, 10927−10934

from a healthy and a CD patient, respectively. Both values were located between the 95% and 99% confidence limit, an indication of a true potential outlier. For bIM, each outlier was excluded from the linear multivariate statistical approach.

After constructing the supervised PLS-DA model per ionization mode (Figures 1 and S3), the validity was assessed

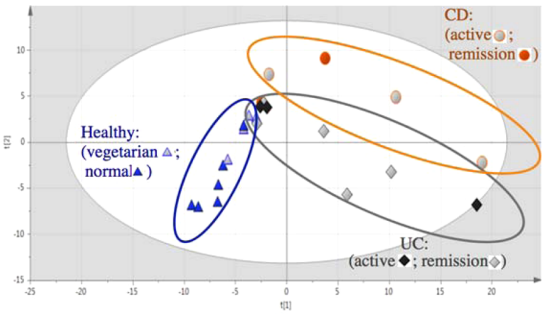


Figure 1. Score plot of a partial least-squares discriminant analysis for the positive ionization mode on a data set of healthy and IBD diagnosed individuals.

by means of an ANOVA test ($P_{CV\ ANOVA} < 0.05$) and the model parameters $R^2(X)$, $R^2(Y)$, $Q^2(Y)$. For bIM, the *P*-values were below 0.05 and the model characteristics were, respectively, 0.505, 0.843, and 0.527 for the positive and 0.511, 0.993, and 0.803 for the negative ionization mode. The significance thresholds for these parameters are considered application dependent, but the general consensus is a threshold of 0.50.[39] It could be concluded that our models displayed good discriminating capabilities for this data set. The OPLS-DA model was also investigated but did not enhance the model characteristics or provide better class separation. This indicated that the systematic variation residing in *X* was linearly correlated to the *Y* variable. Therefore, it was not necessary to cover the variation in *X* that is orthogonal to *Y* separately, as the OPLS model does.

The samples of all IBD patients and of three healthy persons were stored at −80 °C for 5 years. Therefore, this data set was merely seen as a training set for acknowledging the holistic and differentiation ability of our metabolomics workflow. Indeed, class separation occurred between the main different disease phenotypes in the predictive PLS-DA model (Figure 1).

To indicate the biological relevance of the detected metabolites, a putative identification was performed. All ions with a CV ≥ 50% in the healthy or diseased population were excluded, resulting in 2853 and 2372 ions for the positive and negative ionization mode, respectively. These metabolites were cross-referenced to the Human Metabolome Database (www.hmdb.ca). This search was based on the obtained accurate mass, and the mass window error was set at 10 ppm. In total, 84.2% (+) and 89.08% (−) metabolites were putatively identified. In 83.0% of signals, the match list could not be reduced to a single identity (Table S3). Nevertheless, this putative identification offers a clear indication that the detected peaks may be attributed to known metabolites and are not the result of random background noise.

SHIME Experiment. Since Ab treatment affects the microbial diversity and consequently the gut metabolome,[40] colonic samples were selected for metabolic fingerprinting. A first indication of a shift in the gut metabolome is provided through heat map visualization of the metabolic profile (the known unknowns) (Figure 2). The Ab treatment led to affected intensity levels of different compounds including bile acids, carbohydrates, purines, etc., in line with the literature.[40]

Additionally, the heatmap (Figure 2) clearly indicates the hierarchical clustering of proximal samples and the clustering of distal colon samples. This observation was expected as carbohydrate metabolism thrives in the proximal colonic vessel, while in the distal colonic vessel, it is the proteolytic activity that dominates.

Next, multivariate statistics were conducted on the positive ionization mode, enabling the monoisotopic extrapolation of 6446 and 5010 ions in total for the proximal and distal digestive fluids, respectively. Logarithmic transformation and Paretro scaling was applied. Both data sets, proximal and distal colonic fingerprints, did not display any outliers in the PCA score plots (Figure S4). As for the OPLS-DA, $R^2(X)$, $R^2(Y)$, and $Q^2(Y)$ values of 0.706, 0.929, and 0.726 for PC samples and 0.718, 0.946, and 0.794 for DC samples were obtained with a $P_{CV\ ANOVA} < 0.05$. In both data sets, three different clusters were observed, representing the control and mild and shock Ab (Figure 3) treatments. The Ab clearly altered the metabolic output, most likely due to a shift in microbial diversity.[40] This was acknowledged when calculating the cutoff value for the S-plot loading scores. A two-sided permutation was applied with a multiple comparison correction, which resulted in a significant alteration for 34.3% and 27.7% of the metabolites between the control and Ab treatment (shock and mild) ($P < 0.05$) for distal and proximal colon, respectively.

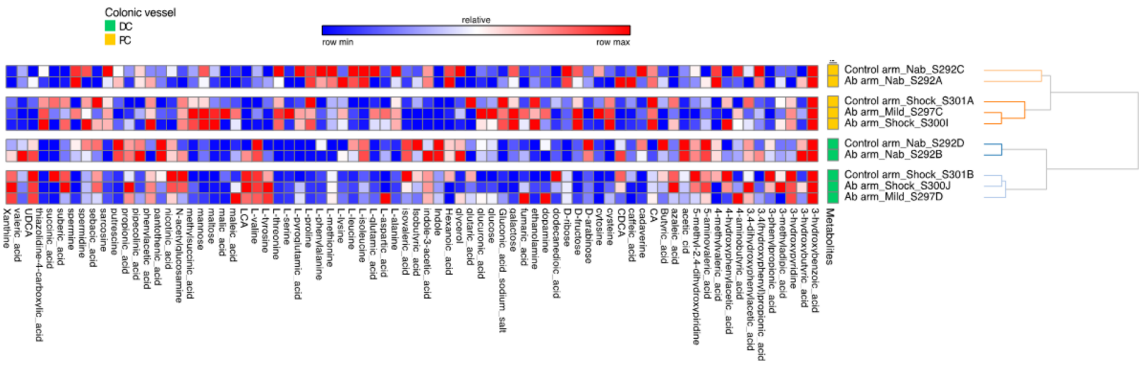


Figure 2. Heat map (GENE-E software, http://www.broadinstitute.org/cancer/software/GENE-E/index.html) visualizing a selected number of 76 known unknown metabolites detected in colonic SHIME suspension fluids with hierarchical clustering of the different samples (n = 10).

DOI: 10.1021/acs.analchem.5b02688
Anal. Chem. 2015, 87, 10927−10934

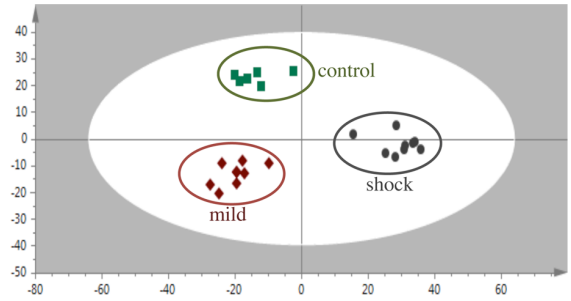


Figure 3. OPLS-DA score plot for the positive ionization mode derived from the distal colonic vessel of a SHIME experiment, with no addition of antibiotics and a mild and shock dosage.

CONCLUSIONS

This study showed that the use of a single analytical platform based on reversed phase UHPLC-Orbitrap HRMS provided a substantial coverage of the gastrointestinal metabolome. Validation of this analytical method was undertaken to produce "high quality" data, in terms of robustness and reliability, this for two biological matrices, human feces and suspensions derived from a validated dynamic *in vitro* gastrointestinal digestion simulator (SHIME). To this end, the intensity levels of known unknown metabolites endogenously present in QC samples were monitored. A chemometric extraction optimization was executed for both matrices, revealing that the mass and pre-extraction with H_2O were significant extraction parameters for feces, and the use of a PVDF filter for SHIME fluids. The performance criteria of the analytical method were excellent, i.e., recoveries between 93% and 105%, precision (instrumental, intra-assay, and interday assay) characterized by CVs below 15%, and linearity with R^2 higher than 0.990 upon exclusion of undiluted samples. Analysis of samples from IBD patients revealed that our developed metabolomics method was capable of detecting a vast amount of metabolites (9553) and provided differentiating profiles between Crohn's disease and ulcerative colitis by means of multivariate statistics. Analysis of digestive fluids derived from a SHIME reactor supplemented with antibiotics revealed that our analytical method is capable of displaying metabolic changes linked to a shift in microbial diversity. All of this ensures the sensitivity and suitability for using this newly developed metabolomic fingerprinting method for future measurements of gastrointestinal derived metabolites to assist in biomarker discovery (prognosis and diagnosis), to reveal underlying physiological pathways, and also to investigate the effect of therapeutic treatments and dietary input for different diseases *in vivo* and *in vitro*.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.5b02688.

Figure S1: UHPLC-HRMS derived XIC of the selected reference standards detected in a QC sample of human feces and of digestive fluids. Figure S2: Principal component analysis score plot for the positive ionization mode on a data set of healthy and IBD diagnosed individuals. Figure S3: Score plot of a partial least-square analysis for the negative ionization mode on a data set of healthy and IBD diagnosed individuals. Figure S4: Principal component analysis score plot for the positive ionization mode derived from the distal and proximal colonic vessel of a SHIME experiment, with no addition of antibiotics, a mild, and shock dosage. Table S1: Overview of the known unknown metabolites used for the profiling metabolomics method. Table S2: Overview of the individual validation results for fecal samples as well as digestive fluids. (PDF)

Table S3: The putative identification of the metabolites of the positive ionization database of the IBD cohort. (XLSX)

AUTHOR INFORMATION

Corresponding Author

*E-mail: Julie.VandenBussche@UGent.be.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

J.V.B. is supported by the BOF (Special Research Fund), Ghent University (BOF/13/PDO/117).The authors wish to thank Ine Opsteyn, Beata Pomain, Dirk Stockx, Mieke Naessens, and Eva Hyde for their technical contribution to this manuscript.

REFERENCES

(1) Idle, J. R.; Gonzalez, F. J. *Cell Metab.* **2007**, *6*, 348−351.
(2) Gao, X. F.; Pujos-Guillot, E.; Sebedio, J. L. *Anal. Chem.* **2010**, *82*, 6447−6456.
(3) Naz, S.; Vallejo, M.; Garcia, A.; Barbas, C. *J. Chromatogr. A.* **2014**, *1353*, 99−105.
(4) Martin, F. P.; Collino, S.; Rezzi, S.; Kochhar, S. *Front. Physiol.* **2012**, *3*, 113.
(5) Zheng, X.; Xie, G.; Zhao, A.; Zhao, L.; Yao, C.; Chiu, N. H.; Zhou, Z.; Bao, Y.; Jia, W.; Nicholson, J. K.; Jia, W. *J. Proteome Res.* **2011**, *10*, 5512−5522.
(6) Nicholson, J. K.; Holmes, E.; Kinross, J. M.; Darzi, A. W.; Takats, Z.; Lindon, J. C. *Nature* **2012**, *491*, 384−392.
(7) Lin, H.-M.; Helsby, N. A.; Rowan, D. D.; Ferguson, L. R. *Inflamm.Bowel Dis.* **2011**, *17*, 1021−1029.
(8) Roux, A.; Lison, D.; Junot, C.; Heilier, J. F. *Clin. Biochem.* **2011**, *44*, 119−135.
(9) Wikoff, W. R.; Anfora, A. T.; Liu, J.; Schultz, P. G.; Lesley, S. A.; Peters, E. C.; Siuzdak, G. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 3698−3703.
(10) Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A. C.; Wilson, M. R.; Knox, C.; Bjorndahl, T. C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; Dame, Z. T.; Poelzer, J.; Huynh, J.; Yallou, F. S.; Psychogios, N.; Dong, E.; Bogumil, R.; Roehring, C.; Wishart, D. S. *PLoS One* **2013**, *8*, e73076.
(11) Li, M.; Song, Y.; Cho, N.; Chang, J. M.; Koo, H. R.; Yi, A.; Kim, H.; Park, S.; Moon, W. K. *PLoS One* **2011**, *6*, e25563.
(12) Martin, F.-P. J.; Sprenger, N.; Montoliu, I.; Rezzi, S.; Kochhar, S.; Nicholson, J. K. *J. Proteome Res.* **2010**, *9*, 5284−5295.
(13) Marcobal, A.; Kashyap, P. C.; Nelson, T. A.; Aronov, P. A.; Donia, M. S.; Spormann, A.; Fischbach, M. A.; Sonnenburg, J. L. *ISME J.* **2013**, *7*, 1933−1943.
(14) Xu, W.; Chen, D.; Wang, N.; Zhang, T.; Zhou, R.; Huan, T.; Lu, Y.; Su, X.; Xie, Q.; Li, L.; Li, L. *Anal. Chem.* **2015**, *87*, 829−836.
(15) Hong, Y. S.; Ahn, Y. T.; Park, J. C.; Lee, J. H.; Lee, H.; Huh, C. S.; Kim, D. H.; Ryu, D. H.; Hwang, G. S. *Arch. Pharmacal Res.* **2010**, *33*, 1091−1101.
(16) Weir, T. L.; Manter, D. K.; Sheflin, A. M.; Barnett, B. A.; Heuberger, A. L.; Ryan, E. P. *PLoS One* **2013**, *8*, e70803.
(17) Goedert, J. J.; Sampson, J. N.; Moore, S. C.; Xiao, Q.; Xiong, J.; Hayes, R. B.; Ahn, J.; Shi, J.; Sinha, R. *Carcinogenesis* **2014**, *35*, 2089−2096.

DOI: 10.1021/acs.analchem.5b02688
Anal. Chem. 2015, 87, 10927−10934

(18) Bjerrum, J. T.; Wang, Y.; Hao, F.; Coskun, M.; Ludwig, C.; Günther, U.; Nielsen, O. H. *Metabolomics* **2015**, *11*, 122−133.
(19) Ponnusamy, K.; Choi, J. N.; Kim, J.; Lee, S. Y.; Lee, D. H. *J. Med. Microbiol.* **2011**, *60*, 817−827.
(20) Cummings, J. H.; Macfarlane, G. T. *J. Appl. Bacteriol.* **1991**, *70*, 443−459.
(21) Chow, J.; Panasevich, M. R.; Alexander, D.; Vester Boler, B. M.; Rossoni Serao, M. C.; Faber, T. A.; Bauer, L. L.; Fahey, G. C. *J. Proteome Res.* **2014**, *13*, 2534−2542.
(22) Worametrachanon, S.; Apichartsrangkoon, A.; Chaikham, P.; Van den Abbeele, P.; Van de Wiele, T.; Wirjantoro, T. I. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 5241−5250.
(23) Deventer, K.; Pozo, O. J.; Verstraete, A. G.; Van Eenoo, P. *TrAC, Trends Anal. Chem.* **2014**, *55*, 1−13.
(24) Kamleh, M. A.; Ebbels, T. M. D.; Spagou, K.; Masson, P.; Want, E. J. *Anal. Chem.* **2012**, *84*, 2670.
(25) Den Besten, G.; van Eunen, K.; Groen, A. K.; Venema, K.; Reijngoud, D. J.; Bakker, B. M. *J. Lipid Res.* **2013**, *54*, 2325−2340.
(26) Boesjes, M.; Brufau, G. *Curr. Med. Chem.* **2014**, *21*, 2822−2829.
(27) Kusano, T.; Suzuki, H. *Polyamines: A universal molecular nexus for growth, survival, and specialized metabolism*; Springer: Tokyo, 2015; p 260.
(28) Heimerl, S.; Moehle, C.; Zahn, A.; Boettcher, A.; Stremmel, W.; Langmann, T.; Schmitz, G. *Biochim. Biophys. Acta, Mol. Basis Dis.* **2006**, *1762*, 341−350.
(29) Blaut, M.; Clavel, T. *J. Nutr.* **2007**, *137*, 751S−755S.
(30) Chan, S. S.; Luben, R.; van Schaik, F.; Oldenburg, B.; Bueno-de-Mesquita, H. B.; Hallmans, G.; Karling, P.; Lindgren, S.; Grip, O.; Key, T.; Crowe, F. L.; Bergmann, M. M.; Overvad, K.; Palli, D.; Masala, G.; Khaw, K. T.; Racine, A.; Carbonnel, F.; Boutron-Ruault, M. C.; Olsen, A.; Tjonneland, A.; Kaaks, R.; Tumino, R.; Trichopoulou, A.; Hart, A. R. *Inflammatory Bowel Dis.* **2014**, *20*, 2013−2021.
(31) Le Gall, G.; Noor, S. O.; Ridgway, K.; Scovell, L.; Jamieson, C.; Johnson, I. T.; Colquhoun, I. J.; Kemsley, E. K.; Narbad, A. *J. Proteome Res.* **2011**, *10*, 4208−4218.
(32) Marchesi, J. R.; Holmes, H.; Khan, F.; Kochhar, S.; Scanlan, P.; Shanahan, F.; Wilson, I. D.; Wang, Y. *J. Proteome Res.* **2007**, *6*, 546−551.
(33) Jansson, J.; Willing, B.; Lucio, M.; Fekete, A.; Dicksved, J.; Halfvarson, J.; Tysk, C.; Schmitt-Kopplin, P. *PLoS One* **2009**, *4*, e6386.
(34) Huang, H. J.; Zhang, A. Y.; Cao, H. C.; Lu, H. F.; Wang, B. H.; Xie, Q.; Xu, W.; Li, L. J. *Dig. Liver Dis.* **2013**, *45*, 677−682.
(35) Boatto, G.; Nieddu, M.; Carta, A.; Pau, A.; Lorenzoni, S.; Manconi, P.; Serra, D. *Forensic Sci. Int.* **2004**, *139*, 191−194.
(36) Llompart-Vizoso, M. P.; Lorenzo-Ferreira, R. A.; Cela-Torrijos, R. *J. High Resolut. Chromatogr.* **1996**, *19*, 207−212.
(37) FDA. *Guidance for industry, bioanalytical method validation*; Food and Drug administration, CDER: Rockville, MD, 2001.
(38) Van den Berg, R. A.; Hoefsloot, H. C. J.; Westerhuis, J. A.; Smilde, A. K.; Van Der Werf, M. J. *BMC Genomics* **2006**, *7*, 142−156.
(39) Triba, M. N.; Le Moyec, L.; Amathieu, R.; Goossens, R.; Bouchemal, N.; Nahon, P.; Rutledge, D. N.; Savarin, P. *Mol. BioSyst.* **2015**, *11*, 13−19.
(40) Antunes, L. C. M.; Han, J.; Ferreira, R. B. R.; Lolic, P.; Borchers, C. H.; Finlay, B. B. *Antimicrob. Agents Chemother.* **2011**, *55*, 1494−1503.

DOI: 10.1021/acs.analchem.5b02688
Anal. Chem. 2015, 87, 10927−10934