Published in final edited form as:
Anal Chem. 2013 January 15; 85(2): 1114–1123. doi:10.1021/ac303011k.

Method Development for Fecal Lipidomics Profiling

Katherine E. Gregory[1,2,#], **Susan S. Bird**[3,#], **Vera S. Gross**[4], **Vasant R. Marur**[3], **Alexander V. Lazarev**[4], **W. Allan Walker**[5], and **Bruce S. Kristal**[3,*]

[1]Boston College, W.F. Connell School of Nursing, 140 Commonwealth Avenue, Chestnut Hill, Massachusetts 02467

[2]Department of Nursing, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115

[3]Department of Neurosurgery, Brigham and Women's Hospital and, Department of Surgery, Harvard Medical School, 221 Longwood Avenue, LMRC-322, Boston, Massachusetts 02115

[4]Pressure BioSciences Inc, 14 Norfolk Ave, South Easton, MA, 02375

[5]Department of Pediatrics, Massachusetts General Hospital for Children, and Division of Nutrition, Harvard Medical School, Boston, Massachusetts, 02115

Abstract

Robust methodologies for the analysis of fecal material will facilitate the understanding of gut (patho)physiology, its role(s) in health and disease and help improve care for individual patients, especially high-risk populations such as premature infants. Because lipidomics offers a biologically and analytically attractive approach, we developed a simple, sensitive, and quantitatively precise method for profiling intact lipids in fecal material. The method utilizes two separate, complementary extraction chemistries, dichloromethane (DCM) and a methyl tert-butyl ether/hexafluoroisopropanol (MTBE) mixture, alone or with high pressure cycling. Extracts were assessed by liquid chromatography-high resolution mass spectrometry-based profiling with all ion higher energy collisional dissociation fragmentation in both positive and negative ionization modes. This approach provides both class-specific and lipid-specific fragments, enhancing lipid characterization. Solvents preferentially extracted lipids based on hydrophobicity. More polar species preferred MTBE; more hydrophobic compounds preferred DCM. Pressure cycling differentially increased the yield of some lipids. The platform enabled analysis of >500 intact lipophillic species with over 300 lipids spanning 6 LIPID MAPS categories identified in the fecal matter from premature infants. No previous report exists that provides these data; thus, this study represents a new paradigm for assessing nutritional health, inflammation, and infectious disease in vulnerable populations.

Keywords

fecal analysis; lipidomics; mass spectrometry; liquid chromatography; high pressure cycling; LC-MS; premature infant; neonate

INTRODUCTION

The relevance of the gastrointestinal (GI) system in health and disease is increasingly evident with advances in immunology and microbiology. In addition to GI morbidities such

*Corresponding author: bkristal@partners.org.
#Co-first authors, contributed equally

as colon cancer[1], chronic diarrhea[2, 3] and inflammatory bowel disease,[4, 5] the gut has been implicated in the pathophysiology of obesity,[6–9] autism,[10, 11] and allergy.[12] Therefore, more sophisticated tools are needed for understanding and evaluating gut function for both scientific and clinical investigation, and to query the health of individuals, especially during infancy and early childhood when few safe assessment modalities are available.

Intestinal function needs to be assessed broadly, quantitatively, and non-invasively to determine gut integrity and its digestive and absorptive functions. The ability to characterize intestinal function would facilitate attempts to recognize and remedy existing pathology and to reduce susceptibilities to future disease. Infants born prematurely provide a pertinent example; they are ill-equipped for autonomous function, partly because of the major organ immaturity and often aberrant microbial colonization within the GI tract, which can lead to many clinical morbidities. These challenges, along with the premature infant's complex nutritional needs, hinder growth and neurodevelopment. Improved strategies for monitoring the gut would help optimize care for vulnerable patients, which also include those with inflammatory bowel diseases such as Crohn's and ulcerative colitis, and GI cancers.

Lipidomics analysis of fecal matter appears to fit the necessary requirements for assessing intestinal function broadly, quantitatively, and non-invasively. For example, fecal samples are collected non-invasively and should reflect intestinal health and function. Similarly, lipidomics, the study of lipids in biological systems, offers a potentially powerful approach in the monitoring of the gut physiology. This field has great promise because lipids are central to intestinal biology, more stable than many metabolites, and more conserved (e.g., across microbiota) than proteins. Lipidomics will complement microbiome studies,[6–10] merging information from microbial genomes, the host, and environmental factors. Notably, lipid absorption is among the most complex processes in the gut; therefore, lipid malabsorption may be a sensitive indicator of GI disease. The lack of established experimental methods for in-depth fecal lipid analysis has limited the study of lipid absorption, lipid (mal)absorption disorders, and associated treatment strategies. Thus, lipidomics offers both biological and analytical advantages for the study of GI function.

The first step in a lipidomics profiling study is the efficient extraction and recovery of lipophilic species from the biological sample. To date, little is known about the extraction and measurement of lipids from fecal material. Thus, lipid extraction techniques must be tested and validated to study this biospecimen type -- beginning with an evaluation of the ability of solvents to extract different classes of lipids efficiently and reproducibly. Bligh and Dyer[13] extractions generally utilize chloroform (CF) as the extraction solvent.[14] Cequier-Sanchez and colleagues modified this approach, replacing CF with dichloromethane (DCM)[15] and showed enhanced coverage of lipid classes in a human serum and rat liver study.[16, 17] The extraction efficiency was compared by assessing the total amount of lipids extracted and by evaluating the number of lipid classes covered from a variety of samples including plants, plant seeds and animal tissues using thin layer chromatography. These results show that DCM yields broader lipid class coverage, suggesting that DCM is an appropriate solvent to replace CF. Alternatively, studies in *E. coli*, mouse brain, *C. elegans*, and human blood plasma suggest that methyl-tert-butyl ether also yields the same extraction efficiency as CF[18], with the added benefit that methyl-tert-butyl ether improves the recovery of certain phosphatidylinositol (PI) and ceramide (Cer) compounds. In addition, the low density of methyl-tert-butyl ether means that the lipid-enriched layer is the upper phase of the solvent system, potentially yielding a cleaner extraction because all insoluble proteins reside at the bottom of the container. With DCM, lipids are extracted into the bottom phase.

Here we directly compare DCM with a methyl tert-butyl ether/hexafluoroisopropanol (MTBE) mixture for extracting lipids from human fecal samples. In the MTBE system, lysis

and extraction occur in the hexafluoroisopropanol phase, with MTBE partitioning lipophilic species away from polar species (proteins, etc.) and water providing phase separation. The potential advantage of this approach is that it also extracts useable nucleic acids and proteins for microbiome studies and may correlate with the inflammatory processes of the host. We further investigated whether the use of pressure cycling could enhance extraction precision and reproducibility. Pressure cycling technology (PCT) applies hydrostatic pressure that rapidly alternates between ambient and very high levels (e.g. 35,000 psi in this study). Such cycling may also improve the yield of some lipids by improved homogenization of the fecal material with the solvent and improved lysis and extraction of bacterial constituents.

MATERIALS AND METHODS

Clinical methods

Approximately 1 g of stool sample was collected from the diapers of five premature infants who were born prior to 32 weeks gestation and cared for in the Neonatal Intensive Care Unit at Brigham and Women's Hospital. Samples were pooled together and homogenized by mixing to form one sample. This pooled sample was used to represent the possible diversity of the lipidome and create a representative mixture for method development. The sample was stored at −80°C until lipid extraction.

Analytical methods

Chemicals and Supplies—All water used for the samples was HPLC-grade water from VWR (Alfa Aesar). For the DCM method of extraction, DCM (HPLC-grade) and methanol (MeOH; HPLC-grade) from VWR were used. For the MTBE method of extraction, we used hexafluoroisopropanol from DuPont and methyl-tert-butyl ether (BDH brand) from VWR. The FT 500-ND PULSE Tubes were from Pressure BioSciences (South Easton, MA). All PULSE tubes were washed with MeOH and dried prior to the experiment. For the DCM method, the PULSE Tubes were modified with (DCM-resistant) ultrahigh purity fluoropolymer O-rings.

LC-MS grade acetonitrile (ACN), MeOH, and isopropanol (IPA), as well as HPLC grade DCM and dimethyl sulfoxide, were purchased from Fisher Scientific (Pittsburg, PA) and ammonium formate was purchased from Sigma-Aldrich (St. Louis, MO). A detailed list of all lipid standards purchased, as well as their abbreviations, preparation conditions and vendor sources can be found in the supplemental information of our previously published LC-MS method.[19]

Preparation of Lipid Standards—Stock solutions were prepared by dissolving lipid standards in DCM:MeOH (2:1 v/v) at concentrations ranging from 10–50 mg/ml and were stored at −20°C. Working solutions were diluted in ACN/IPA/H_2O (65:30:5 v/v/v) to 1 μg/ml prior to spiking studies or LC-MS analysis.

Lipid Extraction—The initial sample was thawed and a slurry of pooled fecal material in water was made by combining 150 mg solid fecal material per 3 ml water, then vortexing well to form a slurry. To test the effect of sample dilution on lipid extraction efficiency, the slurry was also diluted 1:3 and 1:10 in water. Five aliquots of each dilution (full strength, 1:3 and 1:10) were used for each method (n=5). 180 μl of an I.S. consisting of 1,2,3-tripentadecanoylglycerol [$TG(15:0)_3$], 1,2-diheptadecanoyl-*sn*-glycero-3-phosphocholine [PC(17:0/17:0], 1,2-dimyristoyl-*sn*-glycero-3-[phospho-rac-(1-glycerol)] [PG(14:0/14:0], 1-O-Hexadecanoyl-*sn*-glycero-3-phosphocholine [lysoPC(16:0)] and 1,2-dipalmitoyl-*sn*-glycero-3-phospho-L-serine [PS(16:0/16:0)], each at the concentration of 5 μg/ml, was added to 1020 μl of each dilution. 50 μl of fecal suspension was used per sample, resulting in

2, 0.67 or 0.2 mg fecal material per replicate. Two extraction chemistries, each used in combination with either pressure cycling or conventional extraction, were applied to each of the three dilutions, yielding a total of 60 samples for analysis. The experimental plan is outlined in Figure 1.

DCM Method (Bligh-Dyer)—Fecal slurry, 50 µl per sample, was placed into ND PULSE tubes for pressure cycling, or into 2 ml centrifuge tubes for conventional extraction. To each sample 316 µl MeOH was added and vortexed for 30 seconds before the addition of 633 µl DCM. All samples were then vortexed for an additional 30 seconds to ensure that all components were well mixed. The conventional (non-PCT) samples were vortexed intermittently for 15 minutes, whereas the PCT samples were placed in the Barocycler at 35,000 psi for 30 cycles (each cycle consisting of 20 seconds at high pressure and 10 seconds at ambient pressure for a total run time of ~15 mins). The PCT samples were then transferred to 2 ml centrifuge tubes. All samples (PCT and conventional) were centrifuged at 12,000g for 10 minutes to pellet insoluble fecal material. The supernatants were transferred to new 2 ml tubes and 200 µl dH_2O was added to induce phase separation. All samples were then vortexed for 20 seconds, allowed to equilibrate at room temperature for 10 minutes and centrifuged at 8,000g for 10 minutes at 10°C. The lipid-containing DCM phase was carefully aspirated from the bottom of the tubes. To avoid protein contamination, only 600 µl of the DCM phase was recovered for LC-MS analysis.

MTBE Method—As above, 50 µl of slurry was placed into ND PULSE tubes or 2 ml tubes. 500 µl HFIP was added to each tube and vortexed for 30 seconds to ensure that all components were well mixed. The conventional (non-PCT) samples were vortexed intermittently for 15 minutes, whereas the PCT samples were placed in the Barocycler as above. The PCT samples were then transferred to 2 ml centrifuge tubes. All samples (PCT and conventional) were centrifuged at 12,000g for 10 minutes to pellet insoluble fecal material. (Unlike methanol, HFIP efficiently solubilizes both lipids and proteins, but nucleic acids remain insoluble. Therefore this first pellet contains DNA and RNA that can be recovered and analyzed. The pellet was saved and stored at −80 for subsequent DNA isolation. See supplementary figure 1.) The supernatant, containing proteins and lipids dissolved in HFIP, was then transferred to 15 ml polypropylene centrifuge tubes. MTBE, 2 ml per sample, was added to each tube, vortexed to mix, and incubated at room temperature for 10 minutes to allow the proteins to precipitate. All samples were then centrifuged again at 5,000g for 5 minutes to pellet the proteins (these pellets were saved and stored at −80 for subsequent protein isolation; see supplementary figure 2.) The lipid-containing supernatants were transferred to new 15 ml tubes, and 1.5 ml dH_2O was added to induce phase separation. Samples were vortexed for 30 seconds to mix, then centrifuged at 5,000g for 5 minutes. After phase separation, the lipid-containing MTBE phase forms the top layer, whereas the HFIP and water from the lower phase. The MTBE was aspirated and 2 ml transferred to fresh tubes for LC-MS analysis. The lower phase which contains any residual proteins that did not precipitate was also saved at −80 for subsequent protein isolation (see supplementary figure 2).

LC-MS Conditions and Experiments

Details of the LC-MS method and SIEVE analysis have been described previously[19, 20]. Lipid extracts were first evaporated under vacuum and then resuspended in 300 µl of ACN/IPA/H_2O (65:30:5 v/v/v) containing PG (17:0/17:0) at a concentration of 5 µg/ml. They were separated on an Ascentis Express C_{18} 2.1 × 150 mm 2.7 µm column (Sigma-Aldrich, St. Louis, MO) connected to a Thermo Fisher Scientific PAL autosampler, Accela quaternary HPLC pump, and an Exactive benchtop Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, CA) equipped with a heated electrospray ionization (HESI)

probe. Separations ran for 30 minutes with mobile phase A and B. Phase A consisted of 60:40 H_2O:ACN in 10 mM ammonium formate and 0.1% formic acid and phase B of 90:10 IPA:ACN also with 10 mM ammonium formate and 0.1% formic acid. The gradient was as follows: start at 32% B (0 to 1.5 min); increase to 45% B (from 1.5 to 4 min); increase to 52% B (from 4 to 5 min); increase to 58% B (from 5 to 8 min); increase to 66% B (from 8 to 11 min); increase to 70% B (from 11 to 14 min); increase to 75% B (from 14 to 18 min); increase to 97% B (from 18 to 21 min); maintain 97% B (from 21 to 25 min); decrease to 32% B (from 25 to 30 min); maintained at 32% B (remainder of experiment). The column oven temperature was maintained at 45°C and the temperature of the autosampler was set to 4°C. The same LC conditions and buffers were used for all MS experiments; the flow rate was 260 μl/min and the scan range was between *m/z* 120–2000.

For profiling experiments, the MS was run in high resolution mode, corresponding to a resolution of 60k and a 2 Hz scan speed with the spray voltage set to 3.5 kV. The heated capillary and the HESI probe were held at 250°C and 350°C, respectively. The sheath gas flow was set to 25 units and the auxiliary gas set to 15 units. These conditions were held constant for both positive and negative ionization mode full scan acquisitions. The instrument was tuned by direct infusion of PG (17:0/17:0) in both positive and negative mode and external mass calibration was performed using the standard calibration mixture and protocol from Thermo Fisher approximately every five days. A total of 60 samples representing the 5 replicate samples from both extraction techniques and the 3 sample dilutions with and without PCT were analyzed in randomized order by single injections. Additionally, a total study POOL sample was made by combining 10 ul aliquots from each lipid extract. This pool was injected 7 times over the course of the LC-MS sequence, or approximately every 8 samples.

For lipid identification studies, HCD experiments were performed by alternating between full scan acquisitions and HCD scans, both run at 2 Hz. Three different HCD energies, 30, 60 and 100 eV, were used in separate experiments in both positive and negative mode.

All LC-MS profiling samples were analyzed using the MS label free differential analysis software package SIEVE v 1.3 (Thermo Fisher Scientific and Vast Scientific, Cambridge, MA). The framing parameters in these experiments were set at 0.01 Daltons for the *m/z* window and 1.00 minute for the RT window; 1000 was used as the intensity threshold. A pool sample from the middle of the sequence was used as a qualitative reference and for relative quantitation. Frames built off the reference were then applied to all samples in the experiment.

All analyses were done using peak areas determined from SIEVE on all 10,000 frames, and all graphs were made using these raw peak area values. When addressing lipidome coverage, the total number of frames observed was used, whereas our reproducibility assessments were done by calculating the CV of the measurement intensity area, determined by SIEVE, across the 5 sequential sample preparations of each method.

Data Analysis and Lipid Identification

For lipid identifications, the frame *m/z* values were used to do batch searches on the Metlin database [21], the human metabolome database [22] (HMDB), and the LIPID MAPS database.[23] Those matches were confirmed using the molecule's exact mass observed during the analysis, RT regions based on the IS elution times and HCD fragmentation. The data processing, statistical analysis and visualization were done in R 2.15.0[24] and the ggplot2 0.9.1 package[25].

RESULTS AND DISCUSSION

Overview

In fecal lipid analysis, three extraction parameters must be optimized. (i) Quantity of material: Enough starting material is required to provide sufficient lipid for detection; however, too much material may prohibit proper stool homogenization and/or lead to suboptimal solvent/sample ratios, which may impede efficient and unbiased extraction. (ii) Sample homogenization: To achieve reproducible and robust lipid extraction, sample preparation must ensure proper homogenization and -- depending on study goals -- must consistently lyse cells (eukaryotic, eukaryotic and prokaryotic, or neither). The PCT protocol here in conjunction with hexafluroisopropanol lyses both eukaryotic and prokaryotic cells as well as most prokaryotic spores[26, 27]; (iii) Solvent choice: The solvent must address the complexity of fecal samples (e.g., mixed polar, non-polar and neutral lipids in a matrix of water soluble and insoluble material). Both solvent chemistries were tested with and without PCT, on three dilutions of starting material.

Significant pros and cons are associated with each method in terms of extraction efficiency, analytical reproducibility, and lipidome coverage. Because of their different polarities, the two solvent systems should provide complementary coverage of the lipidome; PCT should improve the efficiency and precision of extraction for certain lipid species. The potential advantages of PCT must be balanced against factors such as the need for additional sample handling, which is time consuming and introduces the possibility of greater error, reduced throughput, and the expense related to specialized tubes used with PCT. Additional concerns are non-specific analyte absorption to the container and/or contamination due to leaching.

When the methods are compared, positive and negative LC-MS are presented separately because significant differences were observed that warrant discussion. All analyses were done using peak areas determined from SIEVE v1.3 (see Methods), and all graphs were made using these raw peak area values. SIEVE's unit of measurement is termed a frame. Each frame represents a unique mass/charge (m/z) and retention time pair that could represent either a unique lipid monoisotopic peak, an additional isotope peak of each unique lipid, an ionization adduct (e.g., $[M+H]+$, $[M+Na]+$ or $[M+NH_4]+$), or noise such as a background ion. Thus, a lipid may be represented by more than one frame. We begin by following frames and transition to specific lipids at the end of the report. To guide the reader, we have organized each discussion section similarly, with identical figure designs. Specifically, Figures 2–4 contain multiple panels designated as I, II, III, and IV. Within each panel, the six plots or graphs are labeled A–F and the three concentrations studied increase in dilution from left to right (2.0 mg, 0.67 mg, 0.20 mg). The top and bottom sets of panels represent either the solvents tested (Figures 2 and 3) or the presence or absence of PCT (Figure 4).

Extraction Method: Pressure and Dilution: Positive Ion Mode

Initial results obtained from positive ionization LC-MS established an upper boundary for sample concentration and support the utility of PCT. Figure 2, Panel I.A shows a few compounds extending along the x-axis at y=~0. These are not seen in Figure 2 (Panels I.B,C), however, suggesting that the 2.0 mg sample is too concentrated for effective extraction of these lipids, which are abundant in the less concentrated samples. The vertical line at x=~0 in Figure 2, Panel I.D-F shows that, when MTBE is the solvent, PCT extracts a series of high abundance compounds from fecal material that are either poorly or not extracted in the absence of pressure. Their abundance is roughly equal in all three panels, suggesting that the effect is largely dilution independent, and the related possibility that these signals represent contaminants (e.g., from the PULSE tubes) or molecules that saturate

the extraction system. The utility of PCT is further supported by the data shown in Figure 2, Panel II.A–F, which highlight the relative distribution of lipids measured in each of the four quadrants. In all cases, distributions in the lower left and upper right quadrants are roughly equal, but there is a decidedly greater occupation of lipids in the upper left versus the lower right quadrants. Thus, the use of PCT during extraction appears to consistently increase access to specific lipid subsets.

Precision analyses clarify and support the above findings relevant to dilution and PCT. Regarding sample dilution, a comparison of the CV curves in Figure 2, Panels IV.A and IV.B shows that the CV curve values are much higher (worse) with a 2.0 mg dilution than a 0.67 mg dilution, and, more importantly, the number of useful peaks (i.e., low CV) is not greater. Thus, the 2.0 mg dilution again appears too concentrated. The 0.67 mg and 0.20 mg curves are essentially indistinguishable in these analyses. Regarding PCT, the CV curves for the 0.67 mg and 0.20 mg dilutions (Figure 2, Panels III.B,C,E,F and IV.B,C,E,F) suggest that extraction with PCT is equal to or better than atmospheric pressure extraction both quantitatively (overall CV curve is lower) and qualitatively (curves' right-hand sharp rise occurs later, giving more features for analysis). This improved precision is consistent with greater and more robust isolation efficiency, as noted above.

Extraction Method: Pressure and Dilution: Negative Ion Mode

Initial results obtained from negative ionization LC-MS also support the utility of PCT and establish an upper boundary for sample concentration. The vertical line at x=~0 in Figure 3, Panel I.A–C indicates that when DCM is the solvent, PCT extracts high abundance compounds that are either poorly or not extracted in the absence of pressure. The compounds are roughly equal in abundance in all three plots (A–C) within Panel I, suggesting that the effect is largely dilution independent. The same effect of PCT is further supported by the findings presented in Figure 3, Panel II.A–C (c.f., the relative distribution in the four quadrants). In all DCM extracts, distributions in the lower left and upper right quadrants are roughly equal, but there is a decidedly greater occupation of the upper left quadrants versus the lower right quadrants, suggesting that PCT consistently increases access to lipid subsets, regardless of dilution. These negative ion DCM data are qualitatively similar to data from positive ion mode (Figure 2) with MTBE.

Again, precision analyses clarify and support the above findings on dilution and PCT. Regarding sample dilution: Comparing the CV curves in Figure 3, Panels IV.A and IV.B shows that the CVs are much greater at the 2.0 mg dilution than at 0.67 mg, with no more useful peaks. These data suggest that the 2.0 mg dilution is too concentrated, yielding a less reproducible extraction. With regard to PCT, for DCM with PCT, the 0.67 mg dilution yields ~25% more peaks at a given CV than the 0.20 mg dilution. For MTBE with PCT, the 0.67 mg and 0.20 mg curves are essentially indistinguishable. As noted above, however, when DCM is used, PCT consistently increases access to additional lipid subsets. In sum, when considering only the higher quality peaks at the population level in the DCM-extracted samples, the first 4,000 features are essentially independent of PCT at 0.67 mg; the advantage of PCT for the high quality features becomes apparent only in the 0.20 mg sample. This is again consistent with the primary effect of PCT being to increase yield/sensitivity. With MTBE, PCT has much less effect on the lipid extraction CV curves.

Overall, the data in Figures 2 and 3 show that the 0.67 mg sample dilution is the most appropriate for fecal lipid analysis. Additionally, the use of PCT is at least neutral and usually beneficial to the extraction protocol. The data on MTBE and DCM are nearly opposite in the positive and negative ion mode, suggesting that MTBE is more appropriate for lipid extracts to be analyzed by positive ion LC-MS and DCM best for those species that prefer negative mode. We directly test this below.

Extraction Method: Solvent: Positive Ion Mode

Data comparing MTBE and DCM differ qualitatively. The vertical line at X=~0 in all plots of Figure 4, Panel I shows that, regardless of PCT or dilution, some highly abundant lipids/artifacts are poorly extracted by MTBE. The isolated dots high above the overall lipid population shown in Figure 4, Panel II indicate that this group represents a relatively specific, yet minor, subset of the fecal lipidome. Similarly, the horizontal line -- prominent with PCT and less so in its absence -- shows that there are a series of frames, potentially representing one or more lipids/artifacts that are extracted more efficiently by MTBE with PCT. These data are generally supported by Figure 4, Panel II. The CV curves further clarify and support the findings relevant to sample dilution and extraction solvent (Figure 4, Panels III and IV). Again, the 2.0 mg dilution appears too concentrated (most clearly seen in Figure 4, Panel IV.A versus IV.B). Together, these data indicate a requirement for both MTBE and DCM extractions for studying fecal lipids.

Extraction Method: Solvent: Negative Ion Mode

The negative ion mode data broadly support and highlight the positive ion mode data (supplementary figure 3).

Endogenous Fecal Lipids

The data above speak to the effects of pressure, dilution and solvent chemistry at the lipid population level, defined as frames identified in the LC-MS analysis. This overview indicates that there are clear "groups" of compounds that are more efficiently measured when PCT is used during extraction (i.e., those found along the vertical line at X=0 in Figures 2 and 3, Panel I) or with either MTBE (those found at Y=~0 in Figure 4, Panel I A–C) or DCM (those found at X=~0 in Figure 4, Panel I). To clarify these results, we analyzed the internal standard mixture containing 5 individual lipids (see methods) each at 5 μg/ml. Signal differences for these species based on the different extraction methods (supplementary figure 4) suggest that the endogenous lipids that comprise these populations probably belong to similar lipid classes with shared chemistries.

Endogenous fecal lipids can originate from sources such as the gut flora, epithelial cells from the stomach and intestinal tract, and undigested food. These species have been analyzed to study cholesterol metabolism[28, 29] and excretion, as well as fat (mal)absorption, and can reveal health/disease of both the pancreas and digestive tract. To appreciate the effects that sample dilution, extraction solvent, and pressure have on the fecal lipidome, these endogenous species must be identified and then assessed. Ultimately they partially represent the diversity of the fecal lipidome in terms of molecule chemistry and concentration, and ensure that the patterns we observe when examining SIEVE frames were not misrepresented by background species, lipid adducts, or contamination. These species vary by: (i) ionization potential, (ii) intensity over four orders of magnitude, and (iii) retention time, spanning the entire length of the chromatogram.

In a preliminary positive ion analysis, we identified 212 unique lipids from 6 LIPID MAPS categories, including fatty acyls (FA), glycerolipids, glycerophospholipids, sphingolipids, sterols, and prenols. These identifications, based only on the monoisotopic mass, are expected to include false positives (resolved below). The selected species, profiled to understand the effect of pressure, dilution and chemistry, are graphed in Figure 5, in which the bars represent the average SIEVE area from all 5 replicates for both extraction methods and all 3 sample dilutions. From this analysis, several species exhibited no change between any of the extraction methods, such as the Cer and diglycerides (DGs). A few trends were determined; for example, the oligosaccharides and LysoPE yield higher intensity frames after extraction with MTBE. The PE species were preferentially extracted by DCM, and the

triglyceride (TG) and cholesterol ester (CE) show preference toward MTBE without PCT. Furthermore, the dilution effect on lipid intensity was as expected, with the 2.0 mg dilution having the highest signals, and the 0.20 mg the lowest with the exception of the oligosaccharide, suggesting a problem with solubility.

In the preliminary negative ion analysis, we identified 187 unique lipids (again focused on monoisotopic mass) from 4 LIPID MAPS categories, including FA, glycerolipids, glycerophospholipids, and sphingolipids. The selected species, displayed in Figure 5, show similar trends as observed in the positive ion analysis and emphasize the need for using both DCM and MTBE for comprehensive lipid analysis. Just as in the positive ion analysis, lysoPE and the oligosaccharide are more efficiently extracted by MTBE, regardless of PCT. In contrast, the DCM extraction results in better detection of the PE and PG species (unique to negative ion mode), whereas the Cer signal in negative ion mode is enhanced in MTBE without PCT.

The analyses indicate that the more hydrophilic (i.e., less lipophilic) the molecule, the more it favors MTBE versus DCM. This concept is best highlighted when comparing the lysoPE versus intact PE lipid, in which the loss of a FA side-chain (lyso) decreases the hydrophobicity of the molecule enough to alter its extraction efficiency to favor MTBE. Generally, the species that elute earliest in the chromatogram are the most hydrophilic; however, these separations are dictated by both overall polarity and FA side-chain length. The oligosaccharide compounds, which elute immediately after the void volume and overwhelmingly prefer MTBE extraction, also exemplify this trend.

We utilized such extraction preference trends to eliminate false positive identifications and highlight background ions identified in the preliminary analysis. The species shown in Figure 5 represent the trends seen for each lipid class. For example, all DGs and Cers analyzed in positive ion mode showed little signal variation between extraction protocols, with three distinct exceptions that were favored in DCM with PCT (supplementary figure 5); these were eliminated as suspected background species pending ongoing analysis. Using this method, we confirmed, using exact mass measurements and chromatographic retention times, the structural identification of 304 of the original 337 species. The remaining 33 frames/peaks of interest were eliminated as possible background, unidentified lipid species, or reinvestigated as new identifications based on these trends. The bar graphs for all 304 species identified in this analysis are shown in supplementary figure 6. Additionally, the true number of unique lipids extracted and analyzed from fecal samples is more likely to be well over 500 intact lipophillic species. This number represents all unique species extracted into the lipophillic phase of our liquid-liquid extraction method and detected by LC-MS analysis that are not artifacts of the extraction method. Because not all online databases are comprehensively populated and these samples definitely contain species with unusual side-chains and head groups not commonly analyzed, not all species detected were able to be structurally identified.

Table 1 lists the 304 unique fecal lipids classified by 6 lipid MAPS categories, class, and subclasses. It should be noted that there is overlap between positive and negative ionization mode species and we only count each lipid once even if it was observed in both analyses. From this initial study, we found a prevalence of DGs and Cers, specifically the glycoceramides, in comparison to analyses in sera/plasma and mitochondria.[30] Additionally, we observe an increased incorporation of long, odd chained FAs, such as FA 17:0, 17:1, 15:0, and 19:1, into lipid side chains, most likely because of the dairy-based[31] premature infant diet.

The 304 endogenous lipids identified are listed in supplementary figure 7, which shows under what conditions each was observed Additionally in the supplement, we include a spreadsheet for both the positive and negative analysis each lipid identified containing the RT, *m/z* value observed, error of the mass measurement and the average intensity and measurement CV of each species across all 6 pool samples. These pool samples represent a post sample prep mixture containing an aliquot from each method and dilution tested during analysis. When possible, the side-chains of each species are highlighted, indicating that these lipids were identified using higher energy collisional dissociation (HCD) fragmentation and chromatographic alignment of those fragments. Examples of unknown fecal lipid characterization from the more unique lipid categories of plasmalogens, gangliosides, hydroxy-containing FAs, hydroxy-containing Cers and ether-linked TGs can be found in the supplement (supplementary figures 8a–e). Supplemental data includes the chromatographic alignment of full scan MS and HCD fragmentation data. We have previously described the characterization of Cers, prenols and phospholipids in rat liver mitochondria[19, 30] and TGs in rat serum[20] and those examples directly describe the methods used in this application.

CONCLUSION

To our knowledge this is the first report of full fecal lipidome extraction and analysis of intact species via LC-MS. Previous studies focused on fatty acid analysis using GC-MS,[32, 33] wherein much of the lipidome diversity is lost. We analyzed the lipidome as intact species to better capture biological information. By applying analytical LC-MS methods previously used to characterize the rat liver mitochondrial lipidome[30], we identified 304 endogenous lipid species in fecal material. These species cover 6 of the LIPID MAPS established lipid categories as well as various related classes and subclasses. Our results suggest that the amount of starting material, extraction chemistry, and LC-MS analysis polarity can greatly affect the results. For discovery phase analyses in which the greatest lipidome coverage is desired, both positive and negative ionization should be utilized. In addition, both DCM and MTBE should be used as extraction solvents together with PCT to yield the broadest, most precise coverage of fecal lipid species. We detect two limits: (i) saturation in the extraction phase and (ii) low abundance of some lipids. Therefore, follow-up studies may be useful to optimize sample size, additional solvents for various lipids, and/or concentration protocols (e.g., resuspension in lower volumes) to focus on less abundant lipids.

The techniques developed herein open the door to the use of fecal lipid profiling for both scientific and clinical applications. The method provides a reproducible, generally-applicable, broad based, quantitatively driven, and non-invasive technique for establishing gut function in patients regardless of age or disease. Using fecal lipidomic profiling to define gut function in health and disease will help physicians assess intestinal function, determine gut response to nutrition, and recognize existing inflammation/pathology. Finally, we note that because no previous report provides fecal lipidome data for premature infants, this study represents a new paradigm for determining the impact of initial intestinal colonization and assessing nutritional, infectious, and inflammatory status in this vulnerable patient population. In future work, we intend to examine the origin of the identified intact fecal lipids as either microbial or human, pursue the characterization of the still unknown biological lipids found in fecal matter and use this profiling method to further investigate patient populations by assessing the differences in their fecal lipidome and how these differences may relate to overall gut function and health.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The research reported was supported by K23 NR011320 (PI-KEG), U01-ES16048 (PI-BSK), the Harvard Clinical Nutrition Research Center, Pilot/Feasibility Funding (P30-DK040561 PI-Walker), Brigham and Women's Hospital Department of Neurosurgery Lamas Hydrocephalus Research Initiative (BSK) and Intramural Funding from Boston College (KEG). The authors (SSB, BSK), thank ThermoFisher for the loan of an Exactive Benchtop orbitrap for demonstration testing and financial support for scientific meeting attendance.

REFERENCES

1. McGarr SE, Ridlon JM, Hylemon PB. J Clin Gastroenterol. 2005; 39:98–109. [PubMed: 15681903]
2. Zella GC, Israel EJ. Pediatr Rev. 2012; 33:207–217. quiz 217–208. [PubMed: 22550264]
3. Mathai J, Raju B, Bavdekar A. Indian Pediatr. 2011; 48:37–42. [PubMed: 21317467]
4. Hume G, Radford-Smith GL. Pathology. 2002; 34:561–567. [PubMed: 12555995]
5. Dharmani P, Chadee K. Curr Mol Pharmacol. 2008; 1:195–212. [PubMed: 20021434]
6. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Proc Natl Acad Sci U S A. 2005; 102:11070–11075. [PubMed: 16033867]
7. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. Nature. 2006; 444:1027–1031. [PubMed: 17183312]
8. Kootte RS, Vrieze A, Holleman F, Dallinga-Thie GM, Zoetendal EG, de Vos WM, Groen AK, Hoekstra JB, Stroes ES, Nieuwdorp M. Diabetes Obes Metab. 2012; 14:112–120. [PubMed: 21812894]
9. Tsai F, Coyle WJ. Curr Gastroenterol Rep. 2009; 11:307–313. [PubMed: 19615307]
10. Finegold SM, Molitoris D, Song Y, Liu C, Vaisanen ML, Bolte E, McTeague M, Sandler R, Wexler H, Marlowe EM, Collins MD, Lawson PA, Summanen P, Baysallar M, Tomzynski TJ, Read E, Johnson E, Rolfe R, Nasir P, Shah H, Haake DA, Manning P, Kaul A. Clin Infect Dis. 2002; 35:S6–S16. [PubMed: 12173102]
11. Wang L, Christophersen CT, Sorich MJ, Gerber JP, Angley MT, Conlon MA. Dig Dis Sci. 2012; 57:2096–2102. [PubMed: 22535281]
12. Macdonald TT, Monteleone G. Science. 2005; 307:1920–1925. [PubMed: 15790845]
13. Bligh EG, Dyer WJ. Can J Biochem Physiol. 1959; 37:911–917. [PubMed: 13671378]
14. Iverson SJ, Lang SL, Cooper MH. Lipids. 2001; 36:1283–1287. [PubMed: 11795862]
15. Cequier-Sanchez E, Rodriguez C, Ravelo AG, Zarate R. J Agric Food Chem. 2008; 56:4297–4303. [PubMed: 18505264]
16. Carlson LA. Clin Chim Acta. 1985; 149:89–93. [PubMed: 4028437]
17. Hu C, van Dommelen J, van der Heijden R, Spijksma G, Reijmers TH, Wang M, Slee E, Lu X, Xu G, van der Greef J, Hankemeier T. J Proteome Res. 2008; 7:4982–4991. [PubMed: 18841877]
18. Matyash V, Liebisch G, Kurzchalia TV, Shevchenko A, Schwudke D. J Lipid Res. 2008; 49:1137–1146. [PubMed: 18281723]
19. Bird SS, Marur VR, Sniatynski MJ, Greenberg HK, Kristal BS. Anal Chem. 2011; 83:940–949. [PubMed: 21192696]
20. Bird SS, Marur VR, Sniatynski MJ, Greenberg HK, Kristal BS. Anal Chem. 2011; 83:6648–6657. [PubMed: 21774539]
21. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G. Ther Drug Monit. 2005; 27:747–751. [PubMed: 16404815]
22. Wishart DS, Knox C, Guo AC, Eisner R, Young N, Gautam B, Hau DD, Psychogios N, Dong E, Bouatra S, Mandal R, Sinelnikov I, Xia J, Jia L, Cruz JA, Lim E, Sobsey CA, Shrivastava S, Huang P, Liu P, Fang L, Peng J, Fradette R, Cheng D, Tzur D, Clements M, Lewis A, De Souza A, Zuniga A, Dawe M, Xiong Y, Clive D, Greiner R, Nazyrova A, Shaykhutdinov R, Li L, Vogel HJ, Forsythe I. Nucleic Acids Res. 2009; 37:D603–D610. [PubMed: 18953024]

23. Fahy E, Sud M, Cotter D, Subramaniam S. Nucleic Acids Res. 2007; 35:W606–W612. [PubMed: 17584797]
24. Team, RDC. 2.15.0 ed. R Foundation for Statistical Computing; 2012.
25. Wickham, H. ggplot2: elegant graphics for data analysis. New York: Springer; 2009.
26. Gross V, Carlson G, Kwan AT, Smejkal G, Freeman E, Ivanov AR, Lazarev A. J Biomol Tech. 2008; 19:189–199. [PubMed: 19137106]
27. Powell, BS.; Cybulski, RJIn. Sample Preparation in Biological Mass Spectrometry. Ivanov, AR.; Lazarev, AV., editors. Springer Science+Business Media B.V.; 2011. p. 941-977.
28. Grundy SM, Ahrens EH Jr, Miettinen TA. J Lipid Res. 1965; 6:397–410. [PubMed: 14336211]
29. Miettinen TA, Ahrens EH Jr, Grundy SM. J Lipid Res. 1965; 6:411–424. [PubMed: 14340676]
30. Bird S, Marur V, Stavrovskaya I, Kristal B. Metabolomics. 2012 Accepted Online February 21st.
31. Wolk A, Vessby B, Ljung H, Barrefors P. Am J Clin Nutr. 1998; 68:291–295. [PubMed: 9701185]
32. Birk JJ, Dippold M, Wiesenberg GL, Glaser B. J Chromatogr A. 2012; 1242:1–10. [PubMed: 22560452]
33. Sannolo N, Vajro P, Caporale C, Acampora A, Malorni A. Boll Soc Ital Biol Sper. 1981; 57:118–124. [PubMed: 7248104]

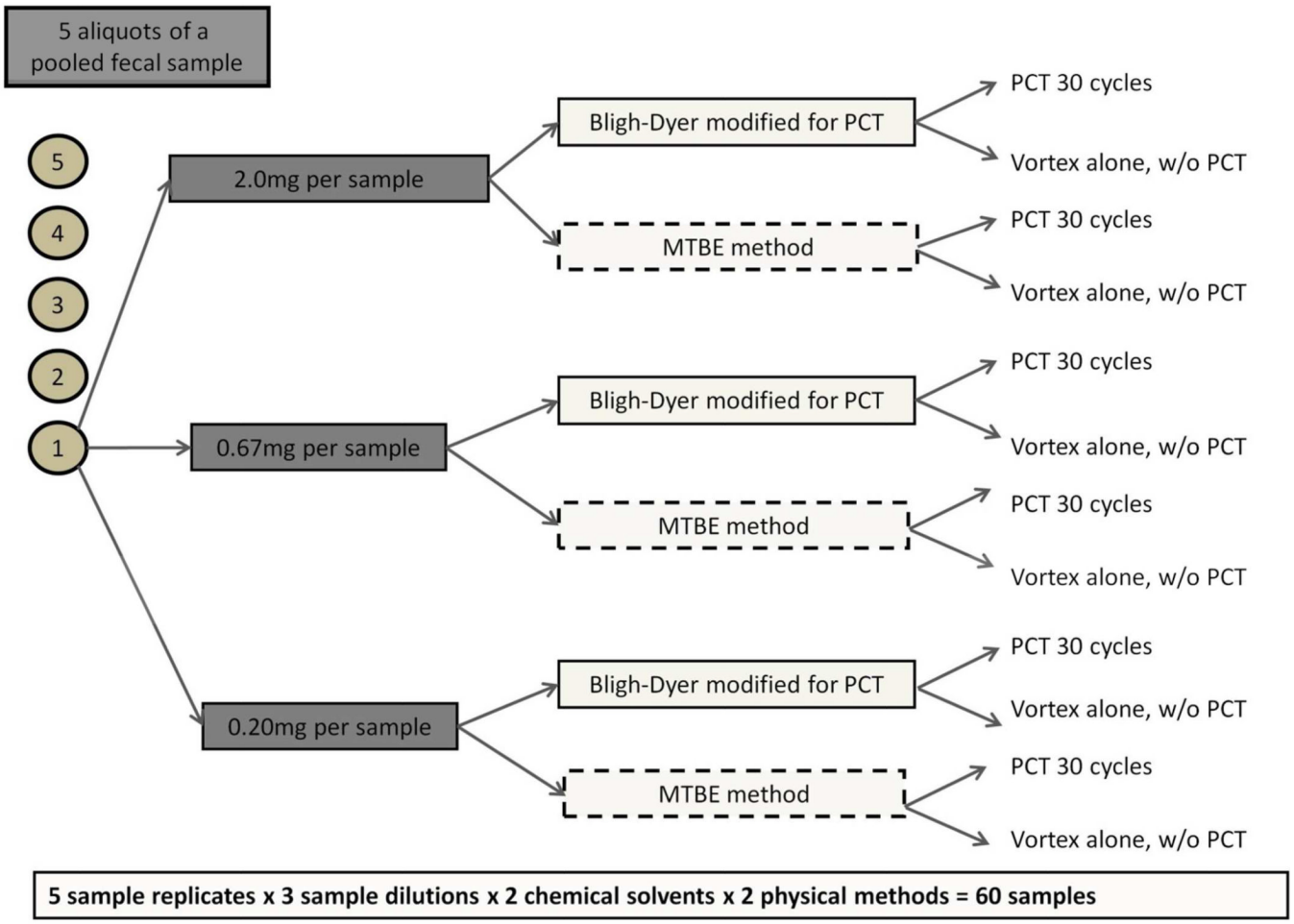

Figure 1. Experimental Design

The pooled fecal slurry was made by starting with well-mixed, pooled sample. 150 mg solid fecal sample per 3 ml water was combined and vortexed well to form a slurry. The slurry was diluted 1:3 and 1:10 in water. 180 μl of I.S. was added to 1020 μl of stock and each dilution. This yielded 2 mg, 0.6 mg, and 0.2 mg fecal material and 7.5 μl I.S. per 50 μl sample aliquot. Lipids were extracted using either a modified Bligh and Dyer method with DCM in place of CF, or an MTBE/HFIP mixture. The use of pressure at 35,000 psi for 30 cycles was applied to half of the samples, whereas the other half were vortexed extensively for approximately 10 minutes without the use of PCT.

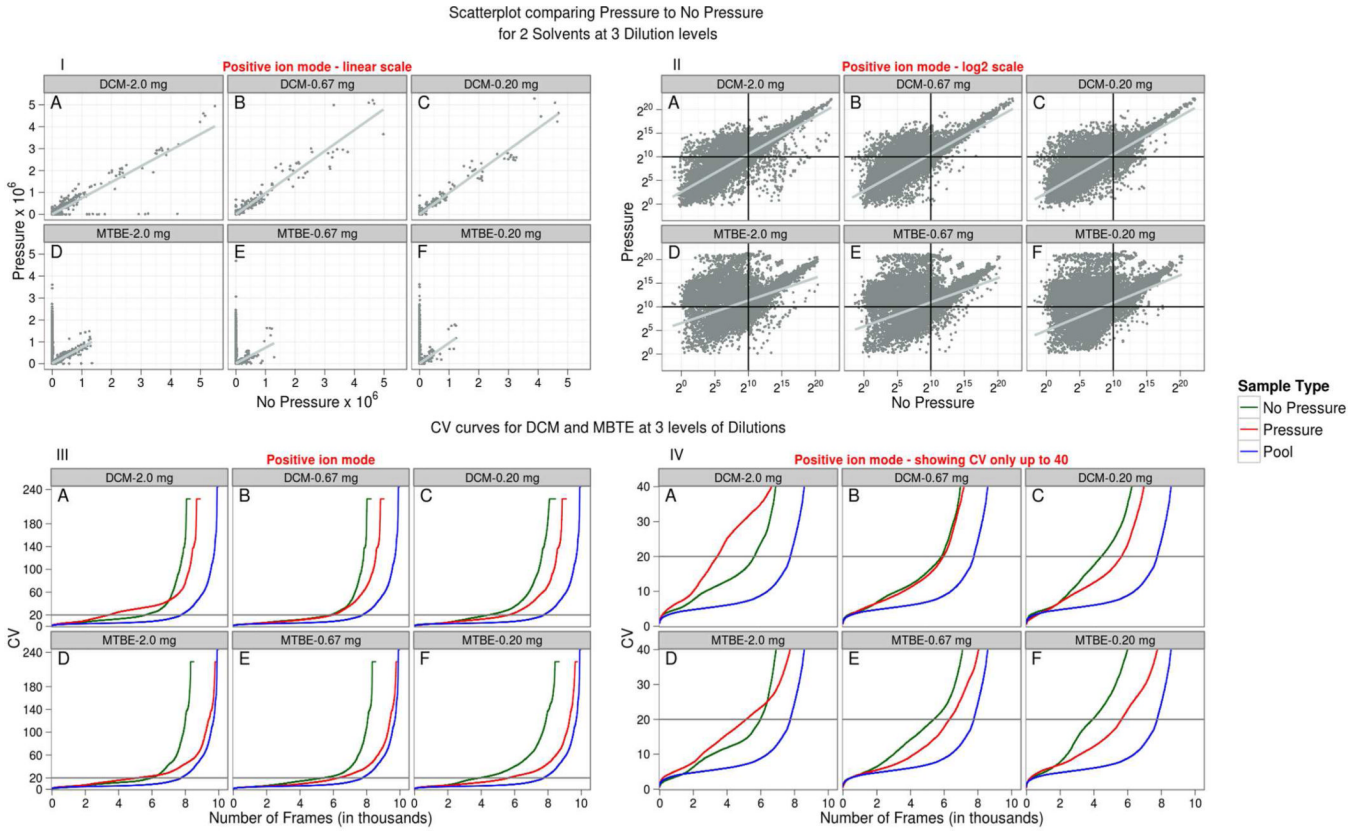

Figure 2. Lipid Abundance and Precision Comparisons: Pressure vs No Pressure, Positive ion Mode

Each panel presents the following data:

I. Top left set of panels: The linear scale of lipid abundance as defined by MS relative abundance. This panel emphasizes the most abundant lipids in the sample.

II. Top right set of panels: The log scale of abundance as defined by MS relative abundance. This panel emphasizes the overall distribution of lipids in the sample. Within these panels, the lower left quadrant is populated by lipids whose overall signal has a relative abundance less than 1000 for either the X or Y condition. We expect that these lipids would be quantitatively unreliable for the LC-MS platform used. In contrast, lipids in the upper right would be expected to be quantitatively reliable, regardless of extraction approach. In contrast, lipids identified in the upper left or lower right panels would be expected to be much more analytically reliable under the "Y" or "X" extraction protocol, respectively.

III. Bottom left set of panels: The overall CV scale of SIEVE data series. This panel shows 10,000 frames for positive ion mode and 7,000 frames for negative ion mode. These panels emphasize the total lipid yield based on the specific approach.

IV. Bottom right set of panels: The CV scale truncated at 40% CV. This panel emphasizes the "useful peaks" of lipids measured.

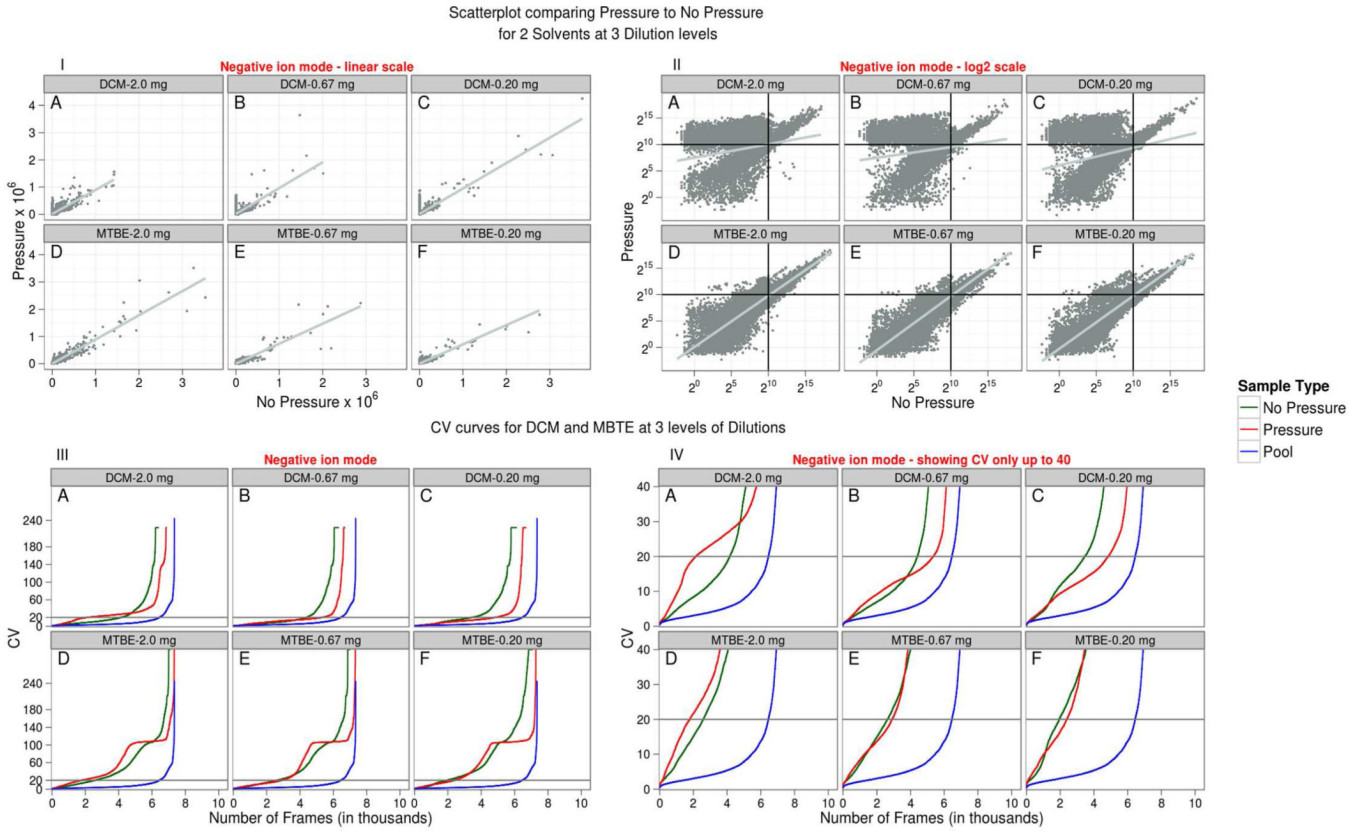

Figure 3. Lipid abundance and Precision Comparisons: Pressure vs No Pressure, Negative ion Mode

Details as in Figure 2

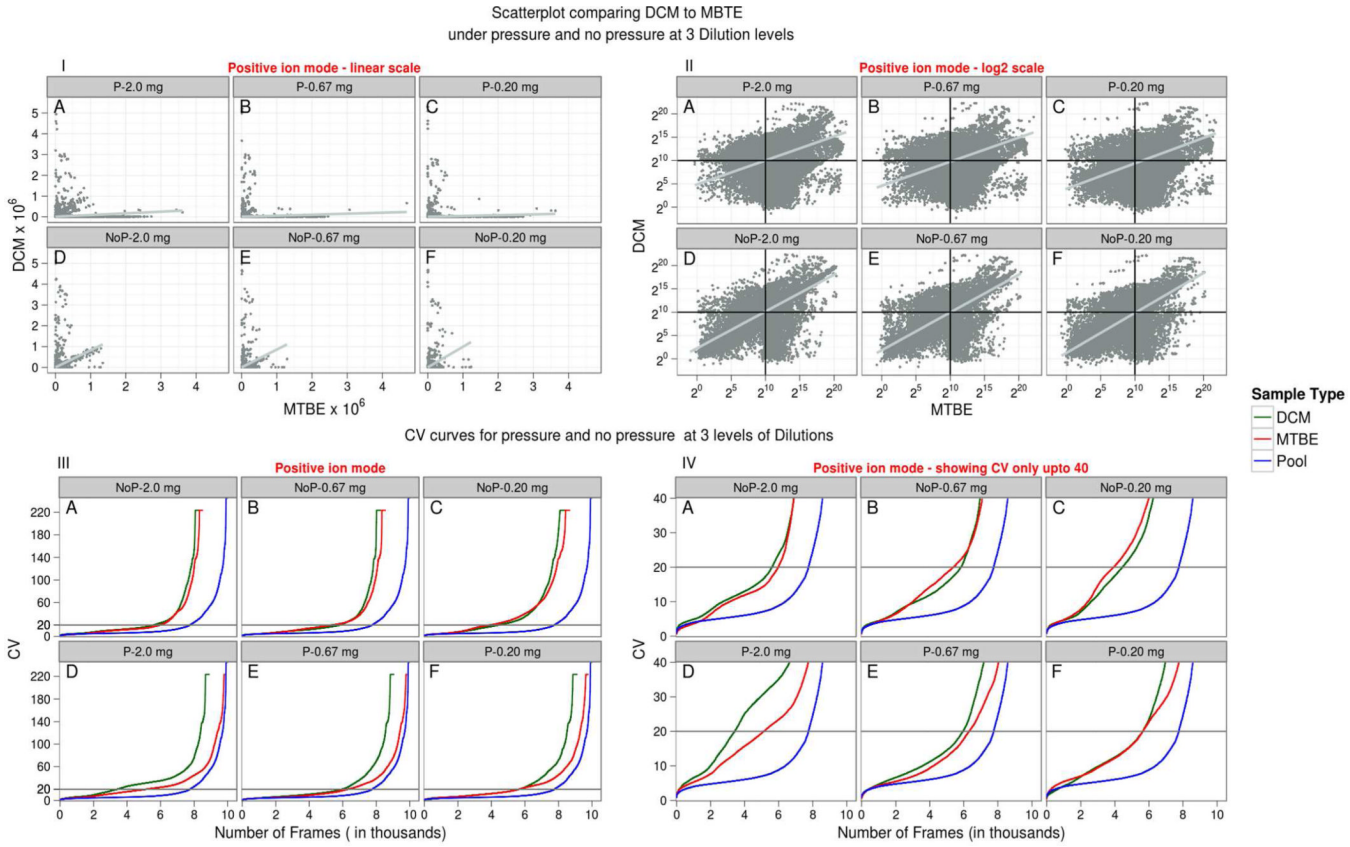

Figure 4. Lipid abundance and Precision Comparisons: DCM vs MTBE, <u>Positive ion Mode</u>
Details as in Figure 2

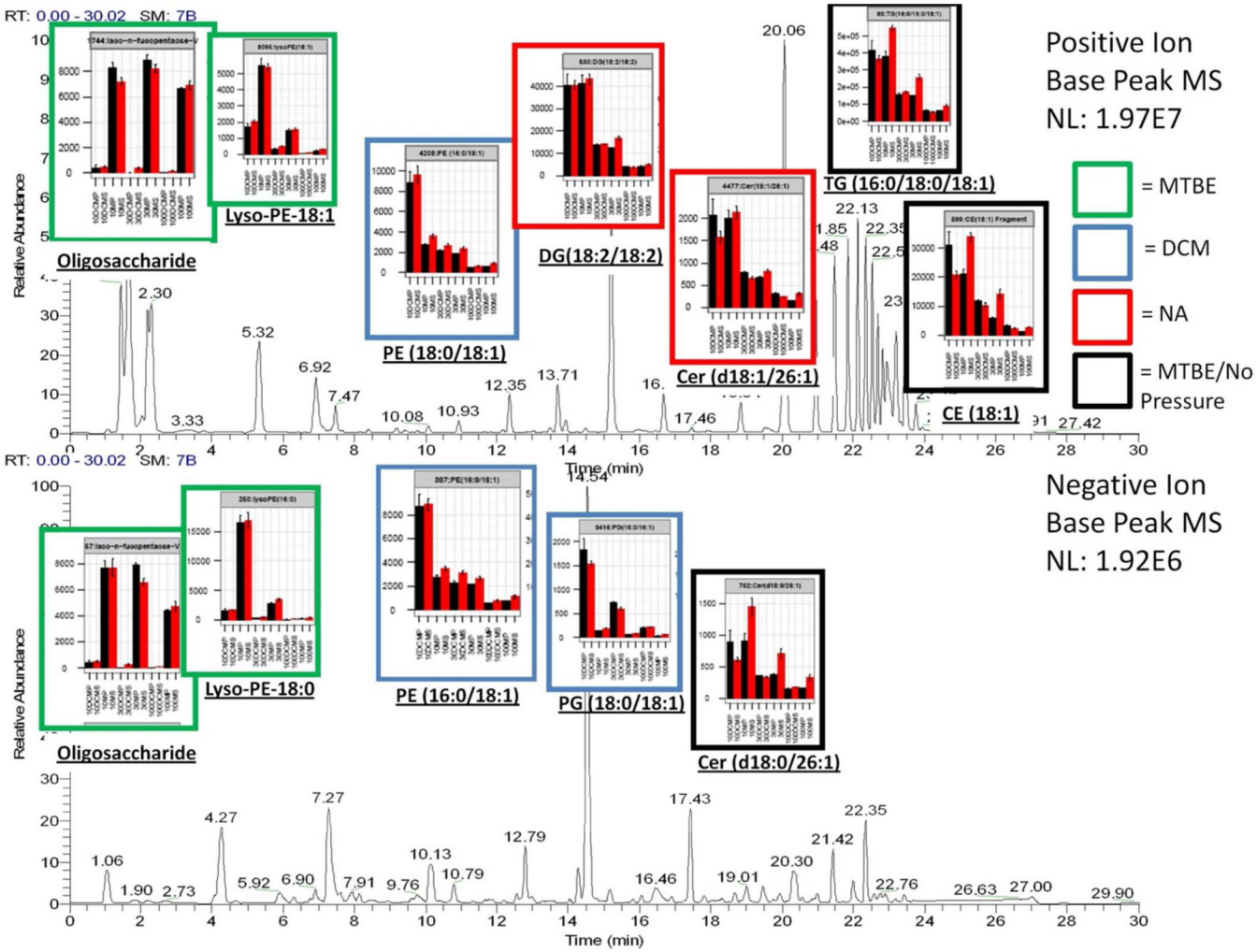

Figure 5. Schematic Showing Relationship between Fecal Lipid Extraction and Elution Characteristics

Panels A and B show two base peak chromatograms from the fecal lipid POOL sample, indicating the difference in signal based on ionization polarity and the need to study these samples in both modes, positive (A) and negative (B). Panel A shows 7 species and Panel B shows 5 species and their individual bar graphs representing their mean signal (±sem) across the 5 samples analyzed across all 3 dilutions for DCMP (DCM, pressure) and DCMS (DCM, no pressure), MTBEP (MTBE, pressure) and MTBES (MTBE, no pressure). The boxes around each graph represent the individual species' extraction preference, and indicate the need for both extraction solvents and ionization modes to be used to achieve comprehensive lipidome analysis. Equivalent graphs of all 304 identified lipids are in the supplement.

Table 1

LIPIDMAPS Categories, Classes, Subclasses of the 304 Unique Fecal Lipids Identified. IDs, including fatty acid moieties, were identified by monoisotopic mass and fragmentation; IDs, not including fatty acid moieties, were identified by monoisotopic mass, chromatographic retention time and quantitation pattern across the different extraction methods and dilutions.

Lipid Class	Number of Unique Lipids**	− Ionization	+ Ionization
Glycerophospholipids (GP)			
Phosphocholine (PC)	22	[M+FormAcid]−	[M+H]+
plasmalogenphosphocholine (plsPC)	4	[M+FormAcid]−	[M+H]+
lysoPhosphocholine (LysoPC)	6	[M+FormAcid]−	
Phosphoethanolamine (PE)	29	[M−H]−	[M+H]+
plasmalogenphosphoethanolamine (plsPE)	7	[M−H]−	[M+H]+
lysoPhosphoethanolamine (LysoPE)	3	[M−H]−	[M+H]+ / [M+NH4]+
Phosphoserine (PS)	6	[M−H]−	
Phosphoinositol (PI)	9	[M−H]−	
Phosphoglycerol (PG)	14	[M−H]−	
lysoPhosphoglycerol (LysoPG)	2	[M−H]−	
PhosphatidicAcid (PA)			
lysoPhosphatidic Acid (LysoPA)	2	[M−H]−	
FattyAcyls (FA)			
Free Fatty Acids (FFA)	19	[M−H]−	[M+NH4]+
Sphingolipids (SL)			
Sphingomyelins (SM)	13	[M+FormAcid]−	[M+H]+
Ceramides (Cer)	50	[M−H]−	[M+H]+
Gangliosides (Gan)	5	[M−H]−/[M-H20-H]−	
Glycerolipids (GL)			
Diacylglycerol (DG)	19		[M+NH4]+
Triacylglycerol (TG)	88		[M+NH4]+
Sterol Lipids (ST)			
Bile acids and derivatives	2		[M+H]+ / [M+NH4]+
Cholesterol Esters (CE)	2		[M+H]+ / [M+NH4]+
Prenol Lipids (PL)			
coenzyme Q (CoQ)	2		[M+H]+ / [M+NH4]+

** defined as having a unique mass and retention time pair that represents a distinct species. Those species that a re observed in both ionization modes are only counted once.