OPEN

Describing the fecal metabolome in cryogenically collected samples from healthy participants

Kajetan Trošt[1,9], Linda Ahonen[1,8,9], Tommi Suvitaival[1], Nina Christiansen[1], Trine Nielsen[4], Maja Thiele[2,3], Suganya Jacobsen[2,3], Aleksander Krag[2,3], Peter Rossing[1,7], Torben Hansen[4], Lars Ove Dragsted[5,10] & Cristina Legido-Quigley[1,6,10]*

The chemical composition of feces plays an important role in human metabolism. Metabolomics and lipidomics are valuable tools for screening the metabolite composition in feces. Here we set out to describe fecal metabolite composition in healthy participants in frozen stools. Frozen stool samples were collected from 10 healthy volunteers and cryogenically drilled in four areas along the specimen. Polar metabolites were analyzed using derivatization followed by two-dimensional gas chromatography and time of flight mass spectrometry. Lipids were detected using ultra high-performance liquid chromatography coupled with quadruple time-of-flight mass spectrometry. 2326 metabolic features were detected. Out of a total of 298 metabolites that were annotated we report here 185 that showed a technical variation of $x < 30\%$. These metabolites included amino acids, fatty acid derivatives, carboxylic acids and phenolic compounds. Lipids predominantly belonged to the groups of diacylglycerols, triacylglycerols and ceramides. Metabolites varied between sampling areas, some were broadly homogeneous, others varied 80%. A LASSO-computed network using metabolites present in all areas showed two main clusters describing the system, DAG lipids and phenyllactic acid. In feces from healthy participants, the main groups detected were phenolic compounds, ceramides, diacylglycerols and triacylglycerols.

The molecular composition of the intestinal contents is of particular importance in relation to gastrointestinal disorders[1]. The intestinal microbiota has been associated to metabolic syndrome and its complications (fatty liver disease, obesity and type 2 diabetes)[2,3] and linked to arthritis, gout, celiac disease and myalgic encephalomyelitis[4]. The intestinal microbiota affects the metabolism via secondary microbial metabolites and food nutrients[5]. The characterization of the microbiota is traditionally performed via metagenomic profiling[6]. An emerging approach to determine the influence of microbiota on the host organism is to combine metagenomic profiling with molecular phenotyping. Methods to assess the molecular composition of feces can be targeted and quantitative in concentration or untargeted for explorative purposes[7]. In contrast to other biofluids (e.g. plasma or urine), stool is highly heterogeneous because it is composed of living bacteria, food remains, nutrients such as lipids, fibers and non-digestible elements[8,9]. There is little agreement on sample collection techniques and the extent to which a technique can accurately reflect fecal composition[10–12]. To avoid fermentation, fecal samples can be frozen immediately after collection and stored at −80 °C until sample preparation[13].

Stool is frequently analyzed by targeted analysis focusing on abundant and well characterized metabolites like short chain fatty acids (SCFA) and bile acids. Using gas chromatography (GC), the detection of both volatile and non-volatile metabolites can be achieved[13]. In order to cover a broader range of metabolites, two-dimensional gas chromatography (GC × GC) can be employed[14]. These techniques use sample pre-derivatization and can

[1]Steno Diabetes Center Copenhagen, Gentofte, Denmark. [2]Department of Gastroenterology and Hepatology and Odense Patient Data Exploratory Network (OPEN), Odense University Hospital, Odense, Denmark. [3]Institute for Clinical Research, University of Southern Denmark, Odense, Denmark. [4]The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [5]Department of Nutrition, Exercise and Sports, Faculty of Science, University of Copenhagen, Copenhagen, Denmark. [6]Institute of Pharmaceutical Science, King's College London, London, UK. [7]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [8]Present address: Biosyntia ApS, Copenhagen, Denmark. [9]These authors contributed equally: Kajetan Trošt and Linda Ahonen. [10]These authors jointly supervised this work: Lars Ove Dragsted and Cristina Legido-Quigley. *email: cristina.legido.quigley@regionh.dk
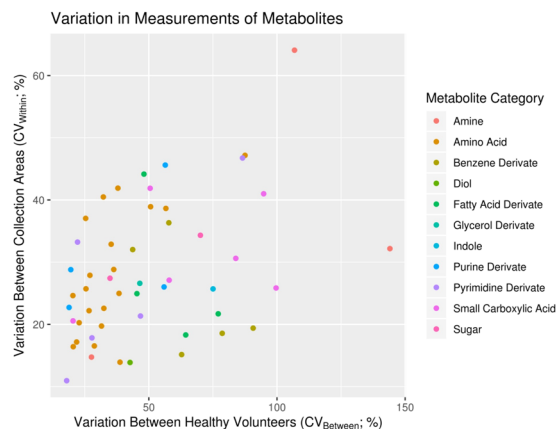
Figure 1. Variation between heathy individuals and sampling area in feces. Each dot represents a metabolite colored by their class. Most metabolites vary more than 30% between collection area and between volunteers making the fecal metabolome highly heterogenous.
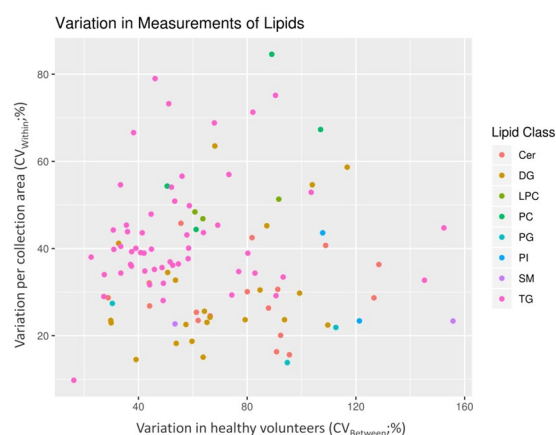

Figure 2. Variation between heathy individuals and sampling area in feces. Each dot represents a lipid colored by their class. Cer: ceramides, DG: diacylglyceride, LPC: lysophosphocholine, PC: phosphatidylcholine, PG: phosphoglycine, PI: phosphoinositol, SM: sphingomyelin, TG triacylglyceride.

detect many polar compounds, e.g. amino acids, small organic acids, phenols, phenolic acids, certain sugars and medium chain fatty acids. The characterization of lipids is an emerging field that is often studied for metabolic disorders as lipids can make up 15% of the feces composition[15–17]. Simplified metabolomic workflow is presented in Supplementary Fig. S1.

Several research groups have developed state-of-the-art fecal fingerprinting methods, including chemical labeling with isotopic reagents and high-performance liquid chromatography coupled with high resolution mass spectrometry (HPLC-HRESMS)[18–20], nuclear magnetic resonance (NMR) for compounds like SCFAs and BCAAs[21–23] and the use of LC-MS[24,25] and GC-MS[26,27]. In this study, we aim to measure molecules in feces with in-house analytical pipelines using cryogenically drilled samples[28–30], focusing on quality control, sampling procedure and biochemical pathways that can describe feces from healthy volunteers.

Results and Discussion

The metabolome is composed of molecules belonging to a myriad of chemical groups. Untargeted analytical methods are usually tailored to detect a broad chemical group, like lipids or polar metabolites, based on abundance, purity, polarity, volatility and on the availability of bioactive groups for derivatization[31]. In this study, the preprocessing of the data resulted in the detection of 2326 features in the polar metabolome and lipidome. Among these features, 182 polar and 116 lipid species were putatively identified. Metabolites with relative standard deviations (%RSDs) higher than 30% in the pooled samples were excluded from further statistical evaluation, thus resulting in 185 metabolites fulfilling this requirement (Figs. 1 and 2). Annotated features with pool RSD < 30% are presented in Supplementary Tables S1 and S2 for polar metabolites and lipids respectively.

To study the feasibility of cryogenically collected samples, we applied additional parameters and calculated metabolite variation in four sampling areas per specimen. As an example of the data collected in this study, the first metabolite in the Supplementary Tables S1 is 3-Hydroxyphenylacetic acid. The variation between participants for this metabolite exceeded 107% and the variation found in the four sampling area was also high at 64% meaning it is not found homogenously in the stool specimens. Since the technical variation of the measurement

is negligible at 4% we can be confident of the technical results. We then propose that this benzene derivative is found in a wide range of concentrations in the feces of the healthy population. The other benzene derivatives that were detected followed the same trend, apart from 3-Phenyllactic acid, 3-Phenylpropionic acid and 4-Hydroxyphenyllactic acid which were in similar concentrations in the four areas. "The most homogenous molecule between participants was uracil (11%) and throughout technical (7%) and between drills (18%). However, the interindividual metabolite variation is affected by water content differences between participants and it is difficult to normalize this with a limited number of samples."

The fecal metabolome and lipidome.

This study allowed for the reporting of 185 molecules, adding to previous research[32]. These metabolites included amino acids, fatty acid derivates, carboxylic acids, benzene compounds and indole compounds (Fig. 1). Fatty acids were mainly present as medium chain fatty acids (e.g. hydroxylated and dicarboxylic species). The group of small carboxylic acids included metabolites from the citric acid cycle as well as hydroxy butyric and propionic acids. Among the amino acids, both regular amino acids as well as branched chain amino acids were present in the feces. In addition to these, some amines, purine derivates and pyrimidine derivates were also annotated. Additionally, the results show that the compounds with the highest variation between participants were benzene derivates (e.g. hydroxyphenyllactic acid, 3-phenyllactic acid and 3-phenylpropionic acid), polyamines and fatty acid derivates (e.g. methylsuccinic acid and 2-hydroxyisocaproic acid), and some of the small carboxylic acids (e.g. tricarballylic acid).

Many of the polar metabolites found in healthy individuals can be associated with bacterial metabolism. A detailed review on microbiota metabolism is out of the scope of this report, but some examples are hydroxyphenyllactic acid which has been found in most anaerobic bacteria and is associated with tyrosine metabolism in humans[33]. In another study, 3-phenyllactic acid is thought to display antimicrobial properties and to be associated with phenylalanine metabolism[34]. In addition, the occurrence of 3-phenylpropionic acid is associated with microbially transformed plant polyphenols such as flavanols, flavanones and tannins[35]. Putrescin is a polyamine produced by the breakdown of amino acids and mainly produced by the microbiota[36]. Other metabolites such as the product of cellulose, cellobiose, has not been reported in feces in metabolomics analyses and is a microbe-produced metabolite.

Feces are also rich in BCAAs and these can be precursors of hydroxyisocaproic acid, a bacterial end product of leucine degradation[37]. BCAAs were in the low physiological variation group in this cohort, ie. isoleucine (technical variation 6%) varied less than 30% with area of collection and between participants. In comparison to other molecules such as fumaric acid (technical variation 8%) varied biologically over the 70% bracket. With this information in 185 molecules it is possible to ascertain if a molecule that might be relevant physiologically in the healthy was excreted homogenously.

Clinical studies using the lipidome in fecal samples are sparse in the literature. A study showed its clinical potential in five prematurely born infants[16]. Van Meulebroek et al. also reported the analysis of the fecal lipidomic profile in a small cohort of healthy volunteers and type 2 diabetic patients[15]. Using adult lyophilized feces, the team identified 127 lipids, out of which 54 were frequently present in feces.

In this study we observed that the fecal samples mainly consisted of ceramides, diglycerides and triglycerides (Fig. 2). Lipids, which are commonly detected in plasma (e.g. lysophosphocholines, glycerophosphocholines and sphingomyelins), were also detected in the fecal samples but with remarkably lower coverage[38]. In fact, only seven phosphatidylcholines could be annotated in the fecal samples while as many as 78 different phosphatidylcholines were identified[38]. Similarly, we detected two sphingomyelins while 30 different sphingomyelins were detected in plasma[38].

Lipids perform many functions, from storing energy to forming cell and organelle walls or as inflammation markers. Lipids are also a source of nutrition for the bacteria living in the gut, and a fatty diet can be a modifying factor of intestinal microbiota[39]. Interestingly, the lipid families that have been linked to metabolic syndrome were abundant in the feces of healthy volunteers. Ceramides which are metabolized from sphingomyelins in eukaryotic organisms were abundant in these participants' feces[40]. Ceramides are found in large quantities in cell membranes and contribute to cellular signaling mechanisms, particularly cell death, partially reproducing the effect of palmitic acid on insulin signaling[41]. Triacylglycerols and diacylglycerols are usually thought of as energy metabolites, although they can also be signaling molecules and have long been implicated in the occurrence of metabolic syndrome[42].

Correlation between fecal lipids and polar metabolites.

Two main clusters were assigned using the partial correlation network inferred with the graphical LASSO algorithm (Fig. 3). This machine learning algorithm will detect the metabolites that best describe the system out of the 185 metabolites and their relationships or correlations. One cluster was composed mostly of diacylglycerides, the other is composed mainly of amino acids and phenyl lactic acid. 3-phenyl lactic acid was the metabolite with most connections to others. Both clusters showed correlations to metabolites from different platforms. Phenyl lactic acid cluster corelated with certain ceramide species and diacylglyceride cluster correlates with a 3-hydroxy butyric acid and 3-phenyl propionic acid.

In the current screening it is not possible to estimate the source of the metabolites. They can derive from food, the microbiota converting food constituents, or they can be primary bacterial or host metabolites. According to the literature, Phenyl lactic acids are mostly associated with lactic acid bacteria and they have antifungal and antimicrobial properties[34]. Additionally, with a focus on integration with metagenomic analysis, the high variability seen in these ten healthy subjects is probably due to variations in both food intake and in microbiota composition[43]. On the other hand, ceramides are known to play a role in metabolic dysfunction[44]. The same is true for diacylglycerides, although their accumulation is thought to be less detrimental than that of triglycerides[45].
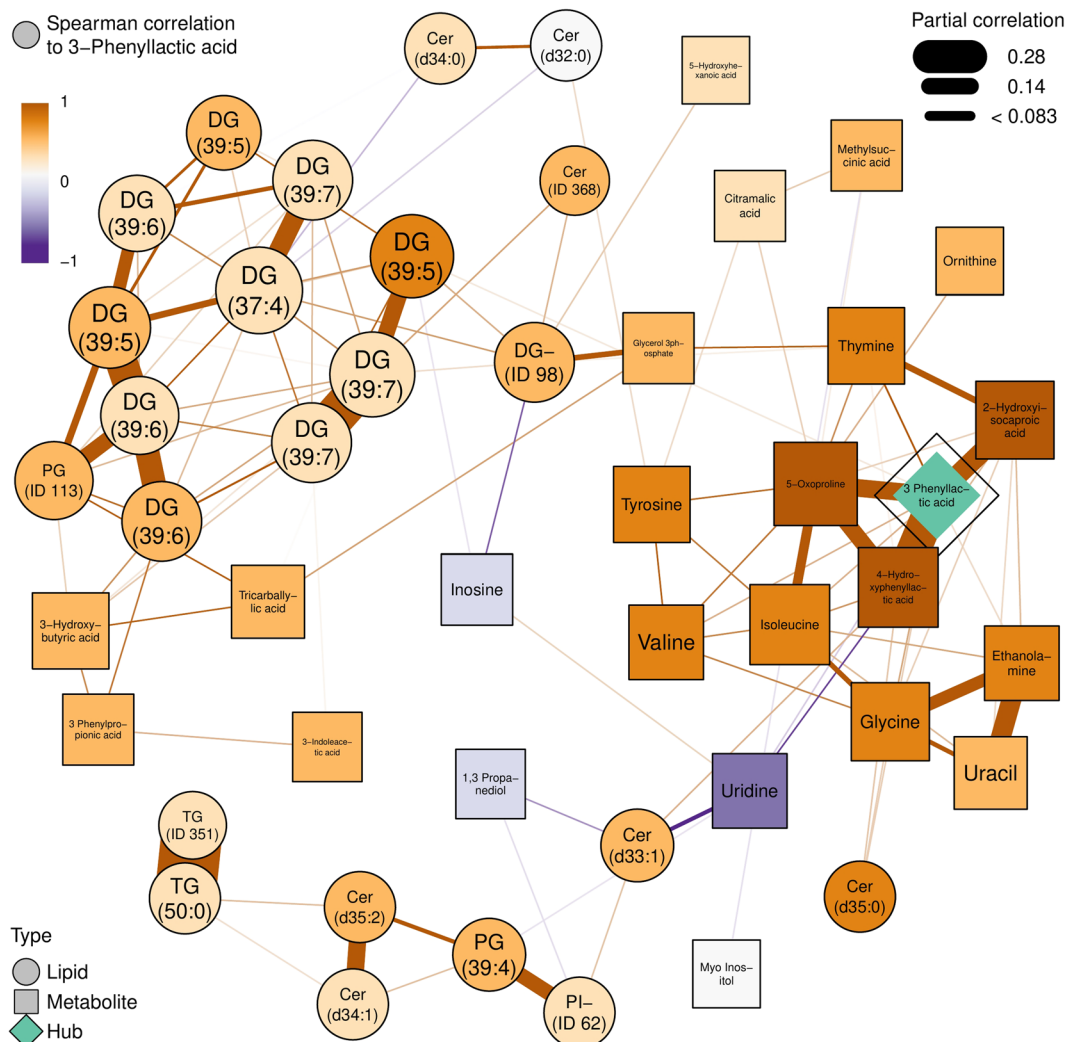
Figure 3. Partial correlation network of fecal polar metabolites and lipids, inferred with the graphical LASSO algorithm. In the figure, lipids were shown as circular and polar metabolites as rectangular nodes. Positive and inverse associations between nodes were shown as brown and purple lines respectively. Strength of the association was shown as width of the line. 3-Phenyl lactic is the main hub of the network with the most associations and is highlighted in green color. Other nodes were colored by the respective Spearman correlation to the hub, 3-Phenyl lactic acid: dark brown is positive correlation and purple is negative correlation. As an example, 4-Hydroxyphenyl lactic acid has a positive partial correlation to 3-Phenyl lactic acid (brown line) as well as a positive Spearman correlation to 3-Phenyl lactic acid (brown node color). The two compounds thus are proportional in abundance in the healthy participants. DGs with the same total number of carbons and double bonds have the same name in the network (ie. 39:7) are different lipids. Cer: ceramides, DG: diacylglyceride, LPC: lysophosphocholine, PC: phosphatidylcholine, PG: phosphoglycine, PI: phosphoinositol, SM: sphingomyelin, TG triacylglyceride.

There are several limitations in this study. The metabolites with good technical reproducibility were reported and analyzed further for correlations, but those classified as analytically inconsistent were not reported. Then, the differences in the metabolites by area of collection were reported, but the participants were not controlled for diet or time of food ingestion with traceable metabolites. The number of healthy participants was not representative of a healthy population, so the origin of the wide-range metabolite levels is unknown at this stage and this work will need replication in a larger cohort.

To conclude, the fecal metabolome composition was studied by applying two analytical platforms. The variation of metabolites was assessed for quality control, per area of collection and between healthy participants. Annotated metabolites included a myriad of metabolic pathways, molecules synthesized by the microbiome (phenol derivatives) and three main classes of lipids: ceramides, triacylglycerides and diacylglycerides in the feces of healthy individuals.

Materials and Methods

Samples. The samples analyzed were from healthy participants, recruited from the general population after a community call for volunteers from Odense University Hospital, under ethical approval from the regional ethics committee of Region of Southern Denmark (S-20160006G). All experiments were performed in accordance with relevant named guidelines regulations. We confirm that informed consent was obtained from all participants and that all methods involving human participants were carried out in accordance with the ethical principles of the Declaration of Helsinki.

Our inclusion criterium was age 40–75 years. We excluded volunteers in case of: (1) being on any medication, prescription or otherwise, at the time of inclusion, (2) having a chronic disease, whether medicated or not, (3) any use of antibiotics within the six months leading up to inclusion, (4) reported alcohol intake above the low-risk limit of 7 units of alcohol/week for females and 14 units/week for males, or binge drinking ($\geq$5 units at one event). Additionally, volunteers went through a sigmoidoscopy, abdominal ultrasonography and liver elastography to screen for existing gastrointestinal disease. We also performed routine blood tests to rule out diabetes, thyroid disease, dyslipidemia and liver disease. Finally, we checked whether any medical events had occurred during the six months following inclusion. In case of a positive finding, we excluded participants.

From February 2016 to December 2016 we included nine men and one female with a mean age of 51 years (range 42–72 years) and a mean BMI of 28 kg/m^2 (range 25–33). None of the volunteers smoked. One participant had a pacemaker due to arrhythmia, one participant consumed antihistamine for seasonal rhinitis, two consumed a daily vitamin supplement and two ingested fish oil. Volunteers sampled stool in their own home, within 24 hours of a scheduled visit to the research clinic. The volunteer stored the sample in his/her own freezer at −20 °C immediately after sampling. We instructed them to keep the sample on ice, in a cooling bag, during transportation. Upon arrival at the research clinic, we then transferred the stool to a −80 °C storage facility.

For the sampling analyses, each original fecal sample was cryogenically drilled four times at different positions along the specimen. Once the cryogenic drilling was finalized, each drill, 200 mg sample, was mixed with an equal amount of water (50:50, w/w), homogenized and distributed into 20 mg aliquots for further extraction procedures.

The chemicals, sample preparation and instrumental analyses for GC-GC-MS and Lipidomics used in this study are available as *Supplementary information*.

Data pre-processing GC $\times$ GC-MS. The detected and potentially identified peaks were aligned using the Gineu software[46]. The peak count filter was set to 5 in order to filter out features, which did not occur in all samples. Additionally, retention indexes were assigned using the NIST14 and GMD libraries[47]. The Golm grouping tool was also used for assessment of groups based on characteristic ions. All features that scored less than 850 of similarity score or had more than 35 units of retention index difference were annotated as unknowns. The resulting peak table was exported. Thereafter the data were post-processed and analyzed in R[55] as described later.

Lipidomics data pre-processing. Data processing was performed using MZmine 2.28[48] using internal mass spectral library and the Lipid Maps database[26]. The following steps were applied in the processing: 1) Crop filtering with a m/z range of 200–1700 m/z and a RT range of 2.4 to 13.6 min, 2) Mass detection with a noise level of 2500, 3) Chromatogram builder with a min time span of 0.04 min, a min height of 7500 and a m/z tolerance of 0.006 m/z or 10.0 ppm, 4) Chromatogram deconvolution using the local minimum search algorithm with a 70% chromatographic threshold, 0.05 min minimum RT range, 5% minimum relative height, 7500 minimum absolute height, a minimum ratio of peak top/edge of 1 and a peak duration range of 0.04–1.0, 5), Isotopic peak grouper with a m/z tolerance of 5.0 ppm, RT tolerance of 0.05 min, maximum charge of 2 and with the most intense isotope set as the representative isotope, 6) Peak filter with minimum 8 data points, a FWHM between 0.0 and 2.0, tailing factor between 0.36 and 2.78 and asymmetry factor between 0.33 and 3.00, 7) Peak list row filter keeping only peak with a minimum of 1 peak in a row, 8) Join aligner with a m/z tolerance of 0.006 or 10.0 ppm and a weight for of 2, a RT tolerance of 0.2 min and a weight of 1 and with no requirement of charge state or ID and no comparison of isotope pattern, 9) Peak list row filter with a minimum of 2 peak in a row, 10) Gap filling using the same RT and m/z range gap filler algorithm with an m/z tolerance of 0.006 m/z or 10.0 ppm, 11) Peak filter with minimum 8 data points, a FWHM between 0.0 and 0.2, tailing factor between 0.36 and 2.78 and asymmetry factor between 0.33 and 3.00 12) Peak list row filter with a minimum of 2 peaks in a row, 13) Identification of lipids using a custom database search with an m/z tolerance of 0.006 m/z or 10.0 ppm and a RT tolerance of 0.2 min, 14) Duplicate peak filter with a m/z tolerance of 0.006 m/z or 10.0 ppm and a RT tolerance of 0.1 min.

Data post-processing. A schematic figure of the post-processing steps can be seen in Fig. S2. The data from each of the platforms were post-processed in the same way in R[55]: Exported peak lists were imported to R[55] and each feature was normalized against the most-correlated internal standard. Annotated metabolites were assigned with a level 3[49]. The annotation included features which had equivalent standards injected during the sequence (Level 1) and structure information acquired previously with MS2 fragmentation (Level 2). Annotations were made using our in-house database, the human metabolome database[50] and Lipid Maps[51].

Features at level 4 (unknowns) and with over 30% coefficient of variation (CV; or, relative standard deviation) in pooled samples were discarded. Further, features with a missing value in more than 20% of samples were discarded, and remaining missing values were imputed with the k-nearest neighbor algorithm using the impute package[52].

Statistical analysis. Statistical analysis was done in R[55] (version 3.4.2, https://www.r-project.org/) with the package limma[53]. Each feature's variation was compared with a feature-wise F-test, where the feature is the dependent variable and the categorical variable representing the individual is the independent variable. The F-statistic, the associated p-value, and the multiple-testing corrected p-values (Benjamini-Hochberg) were reported.

Coefficients of variation (CV) were computed for each feature as follows, between pooled samples for technical variation, between sampling area for specimen heterogeneity and between participants. These feature-wise value pairs were visualized in two bubble plots for the lipidomics and metabolomics data using the ggplot2 package[54]. The metabolite/lipid category of each feature was shown by color of the data point to give an overview of the overall variation in the categories. In the integration step, the data were auto-scaled and averaged over replicates of the same individual as well as over features in each compound category. The two resulting individuals-by-compound categories data sets were then combined into a single data matrix, which was visualized as a rows-scaled heatmap using the ggplot2[54] package to show the relative abundances of each compound category between the individuals[55].

Data driven partial correlation network of lipids and polar metabolites was computed and visualized with R-package qgraph[56] using the graphical LASSO algorithm and RIC (Rotation Information Criterion). Data were imputed and auto-scaled prior to model-fitting. In the visualization, lipids were shown as circular and polar metabolites as rectangular nodes. Positive and inverse associations between nodes were shown as brown and blue lines, respectively. Strength of the association was shown as width of the line. 3-Phenyl lactic acid was selected as metabolite with most connections and highlighted with purple color. Other nodes were colored by the respective Spearman correlation to 3-Phenyl lactic acid.

Data availability

The datasets generated during and analyzed during the current study are available from the corresponding authors on reasonable request and according to the Danish legislature for clinical trials.

References

1. Matsuoka, K. & Kanai, T. The gut microbiota and inflammatory bowel disease. *Semin. In immunopath.* **37**, 47–55 (2015).
2. Marotz, C. A., Zarrinpar, A. & Focus Microbiome: Treating Obesity and Metabolic Syndrome with Fecal Microbiota Transplantation. *Yale J. Biol. Med.* **89**, 383–388 (2016).
3. Pedersen, H. K. *et al.* Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* **535**, 376–381 (2016).
4. Rosenbaum, J. T. & Asquith, M. J. The microbiome: a revolution in treatment for rheumatic diseases? *Curr. Rheumatol. Rep.* **18**, 62 (2016).
5. Nagpal, R. & Yadav, H. Bacterial translocation from the gut to the distant organs: an overview. *Ann. Nutr. Metab.* **71**(Suppl. 1), 11–16 (2017).
6. Sharon, G. *et al.* Specialized metabolites from the microbiome in health and disease. *Cell metab.* **20**, 719–730 (2014).
7. Matsuoka, K. *et al.* Fecal microbiota transplantation for gastrointestinal diseases. *Keio. J. Med.* **63**, 69–74 (2014).
8. De Paepe, E. *et al.* A validated multi-matrix platform for metabolomic fingerprinting of human urine, feces and plasma using ultra-high performance liquid-chromatography coupled to hybrid orbitrap high-resolution mass spectrometry. *Anal. Chim. Acta* **1033**, 108–118 (2018).
9. Cameron, S. J. & Takáts, Z. Mass spectrometry approaches to metabolic profiling of microbial communities within the human gastrointestinal tract. *Methods* **149**, 13–24 (2018).
10. Deda, O., Gika, H. G., Wilson, I. D. & Theodoridis, G. A. An overview of fecal sample preparation for global metabolic profiling. *J. Pharm. Biomed. Anal.* **113**, 137–150 (2015).
11. Lamichhane, S. *et al.* Optimizing sampling strategies for NMR-based metabolomics of human feces: pooled vs. unpooled analyses. *Anal. Meth.* **9**, 4476–4480 (2017).
12. Lamichhane, S., Sen, P., Dickens, A. M., Orešič, M. & Bertram, H. C. Gut metabolome meets microbiome: A methodological perspective to understand the relationship between host and microbe. *Methods.* **149**, 3–12 (2018).
13. Karu, N. *et al.* A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. *Analy. Chim. Acta* **1030**, 1–24 (2018).
14. Almstetter, M. F., Oefner, P. J. & Dettmer, K. Comprehensive two-dimensional gas chromatography in metabolomics. *Anal. Bioanal. Chem.* **402**, 1993–2013 (2012).
15. Van Meulebroek, L. *et al.* Holistic Lipidomics of the Human Gut Phenotype Using Validated Ultra-High-Performance Liquid Chromatography Coupled to Hybrid Orbitrap Mass Spectrometry. *Anal. Chem.* **89**, 12502–12510 (2017).
16. Gregory, K. E. *et al.* Method development for fecal lipidomics profiling. *Anal. Chem.* **85**, 1114–1123 (2012).
17. Han, X. Lipidomics for studying metabolism. *Nat. Rev. Endocrinol.* **12**, 668–679 (2016).
18. Yuan, B. F. *et al.* Comprehensive Profiling of Fecal Metabolome of Mice by Integrated Chemical Isotope Labeling-Mass Spectrometry Analysis. *Anal. Chem.* **90**, 3512–3520 (2018).
19. Xu, W. *et al.* Development of high-performance chemical isotope labeling LC–MS for profiling the human fecal metabolome. *Anal. Chem.* **87**(2), 829–836 (2014).
20. Zheng, S. J. *et al.* Establishment of Liquid Chromatography Retention Index Based on Chemical Labeling for Metabolomic Analysis. *Anal. Chem.* **90**, 8412–8420 (2018).
21. Gratton, J. *et al.* Optimized sample handling strategy for metabolic profiling of human feces. *Anal. Chem.* **88**, 4661–4668 (2016).
22. Lamichhane, S. *et al.* Strategy for nuclear-magnetic-resonance-based metabolomics of human feces. *Anal. Chem.* **87**, 5930–5937 (2015).
23. Dinges, M. M., Lytle, C. & Larive, C. K. 1H NMR-Based Identification of Intestinally Absorbed Metabolites by Ussing Chamber Analysis of the Rat Cecum. *Anal. Chem.* **90**, 4196–4202 (2018).
24. Vanden Bussche, J., Marzorati, M., Laukens, D. & Vanhaecke, L. Validated high resolution mass spectrometry-based approach for metabolomic fingerprinting of the human gut phenotype. *Anal. Chem.* **87**, 10927–10934 (2015).
25. Melnik, A. V. *et al.* Coupling targeted and untargeted mass spectrometry for metabolome-microbiome-wide association studies of human fecal samples. *Anal. Chem.* **89**, 7549–7559 (2017).
26. Gao, X., Pujos-Guillot, E. & Sébédio, J. L. Development of a quantitative metabolomic approach to study clinical human fecal water metabolome based on trimethylsilylation derivatization and GC/MS analysis. *Anal. Chem.* **82**, 6447–6456 (2010).
27. Zhao, L. *et al.* High throughput and quantitative measurement of microbial metabolome by gas chromatography/mass spectrometry using automated alkyl chloroformate derivatization. *Anal. Chem.* **89**, 5565–5577 (2017).
28. Whiley, L., Godzien, J., Ruperez, F. J., Legido-Quigley, C. & Barbas, C. In-vial dual extraction for direct LC-MS analysis of plasma for comprehensive and highly reproducible metabolic fingerprinting. *Anal. Chem.* **84**, 5992–5999 (2012).
29. Sen, A. *et al.* phenotype of the healthy rodent model using in-vial extraction of dried serum, urine, and cerebrospinal fluid spots. *Anal. Chem.* **85**, 7257–7263 (2013).

30. Wang, Y., Caldwell, R., Cowan, D. A. & Legido-Quigley, C. LC-MS-based metabolomics discovers purine endogenous associations with low-dose salbutamol in urine collected for antidoping tests. *Anal. Chem.* **88**, 2243–2249 (2016).
31. Gika, H. G., Theodoridis, G. A., Plumb, R. S. & Wilson, I. D. Current practice of liquid chromatography–mass spectrometry in metabolomics and metabonomics. *J. Pharm. Biomed. Anal.* **87**, 12–25 (2014).
32. Phua, L. C., Koh, P. K., Cheah, P. Y., Ho, H. K. & Chan, E. C. Y. Global gas chromatography/time-of-flight mass spectrometry (GC/TOFMS)-based metabonomic profiling of lyophilized human feces. *J. Chromatogr. B* **937**, 103–113 (2013).
33. Lam, V. *et al.* Intestinal microbial metabolites are linked to severity of myocardial infarction in rats. *PloS one* **11**, e0160840 (2016).
34. Zhang, Z., Lv, J., Pan, L. & Zhang, Y. Roles and applications of probiotic Lactobacillus strains. *Appl. Microbiol. Biotechnol.* **102**, 8135–8143 (2018).
35. Zhou, J., Tang, L., Shen, C. L. & Wang, J. S. Green tea polyphenols modify gut-microbiota dependent metabolisms of energy, bile constituents and micronutrients in female Sprague–Dawley rats. *J. Nutr. Biochem.* **61**, 68–81 (2018).
36. Matsumoto, M. *et al.* Impact of intestinal microbiota on intestinal luminal metabolome. *Sci. Rep.* **2**, 233 (2012).
37. Ehling, S. & Reddy, T. M. Investigation of the presence of β-hydroxy-β-methylbutyric Acid and α-hydroxyisocaproic acid in bovine whole milk and fermented dairy products by a validated liquid chromatography–mass spectrometry method. *J. Agric. Food Chem.* **62**, 1506–1511 (2014).
38. Bowden, J. A. *et al.* lipidomics: NIST interlaboratory comparison exercise for lipidomics using standard reference material 1950 metabolites in frozen human plasma. *J. Lipid Res.* **58**, 2275–2288 (2017).
39. Delzenne, N. M. & Cani, P. D. Interaction between obesity and the gut microbiota: relevance in nutrition. *Annu. Rev. Nutr.* **31**, 15–31 (2011).
40. Schmelz, E. M., Crall, K. J., Larocque, R., Dillehay, D. L. & Merrill, A. H. Jr. Uptake and metabolism of sphingolipids in isolated intestinal loops of mice. *J. Nutr.* **124**, 702–712 (1994).
41. Tran, T. T. *et al.* Short-term palmitate supply impairs intestinal insulin signaling via ceramide production. *Journal of Biol. Chem.* **291**, 16328–16338 (2016).
42. Iqbal, J. & Hussain, M. M. Intestinal lipid absorption. *Am. J. Physiol. Endocrinol. Metab.* **296**, E1183–1194 (2009).
43. Zierer, J. *et al.* The fecal metabolome as a functional readout of the gut microbiome. *Nature genetics* **50**, 790–795 (2018).
44. Jiang, C. *et al.* Intestine-selective farnesoid X receptor inhibition improves obesity-related metabolic dysfunction. *Nature commun.* **6**, 10166 (2015).
45. Nagao, T. *et al.* Dietary diacylglycerol suppresses accumulation of body fat compared to triacylglycerol in men in a double-blind controlled trial. *J. Nutr.* **130**, 792–797 (2000).
46. Castillo, S., Mattila, I., Miettinen, J., Orešič, M. & Hyötyläinen, T. Data analysis tool for comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry. *Anal. Chem.* **83**, 3058–3067 (2011).
47. Kopka, J. *et al.* GMD@ CSB. DB: the Golm metabolome database. *Bioinformatics* **21**, 1635–1638 (2004).
48. Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC bioinformatics* **11**, 395 (2010).
49. Sumner, L. W. *et al.* Proposed minimum reporting standards for chemical analysis. *Metabolomics* **3**, 211–221 (2007).
50. Wishart, D. S. *et al.* HMDB: the human metabolome database. *Nucleic Acids Res.* **35**, D521–526 (2007).
51. Fahy, E., Sud, M., Cotter, D. & Subramaniam, S. LIPID MAPS online tools for lipid research. *Nucleic Acids Res.* **35**, W606–612 (2007).
52. Hastie, T., Tibshirani, R., Narasimhan, B. & Chu, G. Impute: impute: Imputation for microarray data. R package version 1.54.0, https://doi.org/10.18129/B9.bioc.impute (2018).
53. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
54. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, http://ggplot2.org (2016).
55. R Core Team R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 Available at, http://www.R-project.org/ (2017).
56. Epskamp, S., Cramer, A. O., Waldorp, L., Schmittmann, V. & Borsboom, D. qgraph: Network visualizations of relationships in psychometric data. *J. Stat. Softw.* **48**, 1–18 (2012).

Acknowledgements

Author contributions

L.A., K.T., L.O.D. and P.R. designed the experiment. T.N. and T.H. designed sampling procedure. S.J., M.T. and A.K. conducted clinical trial. N.C. prepared the samples. T.S. contributed the statistical analysis. C.L.-Q., L.A. and K.T. wrote the main manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-57888-w.

Correspondence and requests for materials should be addressed to C.L.-Q.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.