



Workshop report: Toward the development of a human whole stool reference material for metabolomic and metagenomic gut microbiome measurements

Rupasri Mandal¹ · Raul Cano² · Cindy D. Davis³ · David Hayashi⁴ · Scott A. Jackson⁵ · Christina M. Jones⁵ · Johanna W. Lampe⁶ · Marie E. Latulippe⁷  · Nancy J. Lin⁵ · Katrice A. Lippa⁵ · Paulina Piotrowski⁵ · Sandra M. Da Silva⁵ · Kelly S. Swanson⁸ · David S. Wishart¹

Received: 30 May 2020 / Accepted: 29 October 2020 / Published online: 8 November 2020
© The Author(s) 2020

Abstract

Introduction To date, there has been little effort to develop standards for metabolome-based gut microbiome measurements despite the significant efforts toward standard development for DNA-based microbiome measurements.

Objectives The National Institute of Standards and Technology (NIST), The BioCollective (TBC), and the North America Branch of the International Life Sciences Institute (ILSI North America) are collaborating to extend NIST's efforts to develop a Human Whole Stool Reference Material for the purpose of method harmonization and eventual quality control.

Methods The reference material will be rationally designed for adequate quality assurance and quality control (QA/QC) for underlying measurements in the study of the impact of diet and nutrition on functional aspects of the host gut microbiome and relationships of those functions to health. To identify which metabolites deserve priority in their value assignment, NIST, TBC, and ILSI North America jointly conducted a workshop on September 12, 2019 at the NIST campus in Gaithersburg, Maryland. The objective of the workshop was to identify metabolites for which evidence indicates relevance to health and disease and to decide on the appropriate course of action to develop a fit-for-purpose reference material.

Results This document represents the consensus opinions of workshop participants and co-authors of this manuscript, and provides additional supporting information. In addition to developing general criteria for metabolite selection and a preliminary list of proposed metabolites, this paper describes some of the strengths and limitations of this initiative given the current state of microbiome research.

Conclusions Given the rapidly evolving nature of gut microbiome science and the current state of knowledge, an RM (as opposed to a CRM) measured for multiple metabolites is appropriate at this stage. As the science evolves, the RM can evolve to match the needs of the research community. Ultimately, the stool RM may exist in sequential versions. Beneficial to this evolution will be a clear line of communication between NIST and the stakeholder community to ensure alignment with current scientific understanding and community needs.

Keywords Metabolites · Metabolomics · Metagenomics · Nutrition · Diet

Disclaimer: Certain commercial equipment, instruments, or materials are identified in this paper to specify adequately the experimental procedures. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology; nor does it imply that the materials or equipment identified are necessarily the best for the purpose.

✉ Marie E. Latulippe
mlatulippe@ilsil.org

Extended author information available on the last page of the article

Abbreviations

CRM	Certified Reference Material
GC	Gas chromatography
ILSI North America	North American Branch of the International Life Sciences Institute
LC	Liquid chromatography
MS	Mass spectrometry
NGS	Next generation sequencing
NIST	National Institute of Standards and Technology
NMR	Nuclear magnetic resonance
PCA	Principal component analysis

QA/QC	Quality assurance and quality control
RM	Reference material
SCFA	Short-chain fatty acid
TBC	The BioCollective

1 Introduction and background

Over the past decade, it has become evident that the human gut microbiome and its metabolic by-products play an important role in a vast and disparate set of health and disease states, including inflammatory bowel disease, diabetes, obesity, cancer, and depression (Ahmed et al. 2013; Anand and Mande 2018; Wishart 2019a). A complex interrelationship of metabolic, hormonal, neurological, and immunological systems exists between the gut microbiome and the host. This molecular cross-talk is critical in regulating many physiological processes. Changes in the composition or function of the gut microbiome and its microbially derived metabolome can have profound consequences, both positive and negative, for the host (de Vos and de Vos 2012; Shreiner et al. 2015; Wishart 2019a).

1.1 Challenges for gut microbiome and gut metabolome research

Traditionally, microbiome research focuses on the characterization and enumeration of microbial species using DNA analysis (Arndt et al. 2012; Knight et al. 2018). However, to truly understand the gut microbiome and the influence it has on human health, it is critical to look beyond identification of microbes. Among other important elements to examine is the chemistry that different gut microbes perform on our food as well as microbially produced metabolites more broadly. The human gut microbiome essentially functions as a metabolic organ that is vital to the digestion of food. Indeed, many of the chemical by-products arising from gut microbiome digestion play important roles in human metabolism, health, and disease. Short-chain fatty acids (SCFAs), such as butyric acid, are generated via the digestion of dietary fiber. These SCFAs have many positive health benefits, including anti-cancer and anti-inflammatory effects (Hoving et al. 2018; Primec et al. 2017). In contrast, many uremic toxins, such as indoxyl sulfate and hippuric acid, are produced by microbial biotransformation of dietary amino acids or dietary polyphenols (Shaw 2010; Tanaka et al. 2015; Wishart 2019a). Thus, the gut metabolome reflects the complex interaction of gut microbial activity and host genetics, diet, demographics, and health (Shah et al. 2020; Zierer et al. 2018).

A key challenge in both gut microbiome and gut metabolome research is collecting and analyzing samples that most accurately reflect microbial composition and metabolism and are of highest relevance to host health. It is well known that the microbial communities of fecal samples differ from those taken from the intestinal lumen or mucosa of healthy and diseased patients (Chen et al. 2014; Durbán et al. 2011; Rangel et al. 2015; Stanley et al. 2015). Because many of the fermentative metabolites are rapidly absorbed by the large intestinal mucosa, fecal samples are not a good measure of production and are a poor measure of microbial fermentative activity (Topping and Clifton 2001; Verbeke et al. 2015). The lack of intestinal sampling has hampered our ability to understand the influence of SCFAs on gut biological functions and host health (Dalile et al. 2019). This problem is likewise applicable to other metabolite classes produced by intestinal microbiota. Given the differences that exist among feces and the material present across the various intestinal regions, sampling the digesta or mucosa of the intestinal region of interest is preferred. However, although non-invasive collection devices may be available in the future (Hällgren 2003; Sinaiko 2004), feces are most commonly collected because invasive techniques (e.g., biopsies) are the only means by which intestinal samples may be collected currently and such sampling requires colon cleansing which can perturb the microbiome. Because of the invasiveness, these samples are usually limited to patients with disease or suspected disease so true controls are rarely available. Moreover, sample numbers are typically limited so studies require long periods of time.

Another key challenge is determining which gut microbial reactions are responsible for producing helpful or harmful metabolites and from which food products. Microbial metabolism is not only dependent on the foods that the host consumes, but also to what level the food components are digested and if/how gastrointestinal secretions (e.g., enzymes, mucus) have been affected. Because most of the microbial activity occurs in the large intestine, the metabolites produced by bacteria are primarily derived from the substances escaping host digestion in the stomach and small intestine. Currently, most investigations do not consider or estimate the host vs. microbial contribution to nutrient digestion, with very little being known about these aspects of gut microbiome metabolism. Using metabolomics, information has been compiled (Wishart 2019a; Wishart et al. 2018) about the bacteria that act on food components and which helpful/harmful chemicals are produced. This “catalogue,” however, is still very sparse. Furthermore, there has been very little action toward the development of standards for functional/metabolic measurements of gut microbiomes or gut-derived metabolomes.

1.2 Value of a reference material for the field of fecal metabolomics

To understand biologically relevant properties of the human gut microbiome, validated analytical measurements that accurately describe various properties of the microbial community, both quantitatively and qualitatively, are needed. To date, DNA-based metagenomic measurements have been the go-to method for most microbiome studies (Knight et al. 2018; Riesenfeld et al. 2004). As a result, there have been significant efforts within the scientific community to increase confidence in metagenomic-based microbiome measurements. A few examples of these efforts include the development of reference materials (RMs) (mock communities and DNA mixtures), the launch of various interlaboratory comparison studies and measurement challenges to identify bias and assess reproducibility (e.g., <https://platform.mosaicbiome.com/challenges/8>), the organization of workshops and consortia that focus on standards development, and the development of better in vitro model systems that allow systematic studies across laboratories (Allen-Vercoe et al. 2019; Fritz et al. 2019; Hardwick et al. 2018; Li et al. 2019a; Sczyrba et al. 2017).

Despite the significant efforts devoted to the development of standards for DNA-based microbiome (especially gut microbiome) measurements, there has been little effort to develop standards for metabolome-based gut microbiome measurements. As a result, metabolomic data collected by different laboratories at different times cannot be readily compared, replicated, or studied in a systematic way. This has created a barrier to understanding links between the diet, the gut microbiome, and health (Karu et al. 2018).

With this need in mind, the National Institute of Standards and Technology (NIST), The BioCollective (TBC), and the North American Branch of the International Life Sciences Institute (ILSI North America) are collaborating to address the need for developing a Human Whole Stool RM. The intent of this RM is to provide method harmonization and eventual quality assurance and quality control (QA/QC) for measurements intended to capture gut microbiome and gut metabolome function. While this RM will be applicable across many scientific disciplines, it will be rationally designed and characterized to be especially well-suited to support those in the scientific community studying the impact of diet and nutrition on functional aspects of the host microbiome as well as the relationship between the host microbiome and human health. To identify which metabolites deserve priority in their value assignment, the three parties jointly conducted a workshop on September 12, 2019 at the NIST campus in Gaithersburg, Maryland. The objectives of this workshop were to identify metabolites for which evidence indicates relevance to health and disease and to decide on the appropriate course of action

to develop fit-for-purpose RM(s). This document represents the consensus opinions of the workshop participants and co-authors of this manuscript. In addition, this paper describes the goals of the whole stool RM project, it offers a consensus approach for NIST's development of the whole stool RM, and it addresses some of the strengths and limitations of going forward with this initiative given the current state of microbiome research.

Overall, the goals of this RM project are to (1) develop a human gut microbiome RM prepared from human stools, (2) promote the adoption of this RM by the scientific community to validate and benchmark current protocols, and (3) promote adoption of the RM as the "gold standard" for QC of metagenomic and metabolomic measurements of clinically relevant stool materials.

2 Current knowledge of human fecal sample collection, preparation, and analysis

2.1 Sample collection and storage

Due to the complexity of the human stool matrix, stool samples reflect more than the environment of the host gut. Stool samples are susceptible to environmental effects. Significant post-collection metabolite deterioration occurs due to exposure to aerobic conditions and to temperature changes, mainly caused by microbial fermentation at room temperature. In many of the reviewed studies, it has been indicated that stool samples were immediately stored at 4 °C or lower until processing (Gratton et al. 2016). For a metabolomics study design, the collection methodology should be kept as consistent as possible across all samples in the experimental classes to avoid bias.

Sample collection is a large source of variability across studies. A comparison between sample preparation conditions and their effects on stool metabolite detection is given in a recent review paper (Karu et al. 2018). The use of frozen stool samples in human studies is both more common and somewhat more practical. However, freeze-drying is also an option but may result in the loss of volatile compounds, many of which have been linked to physiological benefits and are thus of research interest. For fecal collections at home or in locations where immediate freezing is not possible, OMNIgene Gut Stool Microbiome Kits (Anderson et al. 2016) and others like it may be used. These kits contain reagents to stabilize and preserve samples at ambient temperatures, allowing for short-term storage and shipping without compromising sample and data integrity (Liang et al. 2020; Wang et al. 2018). Homogenizing and aliquoting samples prior to freezing can minimize unintended or unnecessary freeze-thaw cycles. As with other biofluids or tissue specimens in general, it is important to minimize handling time

and to use uniform sample handling/storage procedures across the sample set. Reporting an averaged value for specific metabolite mass fractions or concentrations for samples collected over multiple days is also a common strategy for reducing the apparent variability.

Spot versus total fecal sampling is also an important consideration because it influences how data are interpreted. While concentrations (e.g., $\mu\text{mol/g}$ digesta or feces) are usually the focus, the total metabolite pools (e.g., total μmol in colon; total μmol in body) are just as, or even more important. Exercise, dietary intake (e.g., fiber), obesity, and other environmental factors may impact intestinal transit time/defecation frequency, total volume of digesta/feces excreted, fecal moisture content, and microbial density. These variables may consequently affect metabolite production rates, digesta/fecal concentrations, and total exposure to the host (Falony et al. 2018). From a host energetics and health perspective, for example, the total SCFA produced and utilized by the body has much greater relevance than a fecal concentration (Bergman 1990; Dalile et al. 2019). Thus, calculating the total gastrointestinal SCFA pool from the SCFA concentration and the total volume of digesta/feces will provide greater insight than concentration alone. Likewise, the impact of bile acids on the body is likely not well described by fecal concentrations because the body's pool includes the serum, liver, gallbladder (bile), gastrointestinal tract (duodenum, jejunum, ileum, cecum, and colon), and fecal components (Chiang and Ferrell 2020). Lastly, from a gastrointestinal disease perspective (e.g., colorectal cancer; inflammatory bowel diseases), the digesta/fecal concentrations of toxic compounds (e.g., phenols and indoles; hydrogen sulfide; heterocyclic amines; branched-chain fatty acids), transit time, and digesta/fecal volume may be used to estimate total exposure to the body (Ikeda et al. 1994; Le Gall et al. 2018). Given the number of factors involved and their influences on health, total fecal collections over a period of a few days, in addition to the fresh samples used for metabolite concentration measurement, are suggested so that total exposures may be estimated.

2.2 Challenges in fecal sample preparation

As human stool samples contain an extraordinary variety of chemical classes, a true representation of the metabolic profile requires broad and consistent metabolite recovery during the extraction process. However, by maximizing metabolite recovery, the extraction process should not be destructive, nor should it alter the bacterial content of human stools (Gratton et al. 2016). The risk for this is reduced when samples are filtered, dried, or undergo sterilization by means of sonication (Saric et al. 2008). However, these procedures

may increase the risk of metabolite degradation or loss, depending on the chemical or physical "aggressiveness" of the applied extraction method. The optimal extraction processes also depend on the study objectives. If the focus of a study is on intracellular bacterial metabolites and membrane lipids, a more vigorous homogenization may be required to lyse the bacterial cells. This may also be promoted via multiple freeze–thaw cycles (Shao et al. 2016, 2017). The use of a multi-step extraction procedure is another source of variation in metabolite recovery. It is recommended that the reproducibility of a multi-step recovery can be partly assessed (per chemical class) by the addition of isotopically labeled internal standards at the start of each sample extraction, along with appropriate sample homogenization to mix the standards into the matrix (Saric et al. 2008).

The quantity of starting material is another important factor that affects the reproducibility of the extraction recovery. Samples less than 150 mg prior to freeze-drying tend to suffer from relatively higher deviations in measured weight as well as chemical perturbations due to the introduction of external contaminants (Cesbron et al. 2017; Deda et al. 2015). It is generally recommended that blank extractions be added to the set of actual sample extractions to prevent misinterpretation of the collected NMR spectra (Marchesi et al. 2007).

The issue of water content in stool samples also influences the extraction recovery and metabolite concentrations reported. First, if appropriate care is not taken to achieve good extraction reproducibility, it is possible to have situations where the technical variance is greater than the biological variance, leading to incorrect conclusions. Second, water content of feces may vary considerably. Human feces commonly contain 60% to 80% water (by mass), which can decrease or increase beyond this range in cases of constipation and diarrhea, respectively (Nishimuta et al. 2006). Given the high water content of feces and the wide range by which it may fluctuate within and among study participants, representing fecal metabolite data on a wet or dry basis is not a minor consideration.

2.3 Analysis of human stool samples

The success of a stool metabolomics study depends on having an appropriate analytical approach and access to suitable analytical platforms in addition to adequate sample collection and preparation protocols. As highlighted in a recent review, there are no best practices for metabolomic analyses of human stool material (Karu et al. 2018). The lack of consensus best practices can lead to tremendous variability in reported results. In addition, for many metabolomic studies, the metabolite sensitivity and coverage can vary

tremendously between different types of instruments or different types of platforms. It is highly recommended that more than one analytical platform is used because (1) different platforms often have complementary sensitivity to different classes of metabolites (gas chromatography [GC]-mass spectrometry [MS] for volatiles and organic acids, nuclear magnetic resonance [NMR] for very polar compounds, and liquid chromatography [LC]-MS for more hydrophobic molecules); (2) using more than one platform can greatly increase the breadth of metabolite coverage; and (3) detection and/or quantification of a metabolite on one or more platform helps with confirmation and instrument calibration.

LC-MS methods for fecal metabolomics provide relatively wide metabolite coverage compared to that of NMR and GC-MS assays. This can be further extended by the application of careful LC separations coupled with multiple ionization technologies and polarities (Nordstrom et al. 2008). High-throughput approaches in LC-MS are advancing quickly owing to the development of ultra-performance LC systems and polarity switching, which is now widely available on many high-resolution MS instruments. These advantages have already contributed to the dramatic increase in utilization of LC-MS in stool analysis or fecal metabolomics over the past 3 years to 4 years (Karu et al. 2018).

3 NIST's intention to produce a metagenomic and metabolomic reference material

NIST has a unique global role as a National Metrology Institute for the United States but also as a non-regulatory agency in the U.S. Department of Commerce with a mission to promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology. While serving as the national reference laboratory for measurements in the chemical, biological, and material sciences, NIST has developed RMs that are designed to validate analytical methods used in the determination of nutritional and health status markers, contaminant exposure components, and clinically relevant metabolites in human serum, plasma, and urine (National Institute of Standards and Technology 2020a, b; Rasberry 2003). One of the more notable is Standard Reference Material 1950 – Metabolites in Frozen Human Plasma, which has been value assigned (i.e., certified and non-certified mass fraction content) for nearly 100 electrolytes, vitamins, hormones, fatty acids, and amino acids, among other chemical species (Simon-Manso et al. 2013). Currently, suites of pooled human plasma and urine RMs for untargeted metabolome analysis are under development. These material suites will be characterized

with respect to both the chemical annotation/identification of the predominant metabolites and their respective fold changes/percent differences for purposes of underpinning differential metabolomics and lipidomics studies.

NIST is presently developing a candidate Human Gut Microbiome (Whole Stool) RM for measurement harmonization and QA/QC for stool-based metagenomics, metabolomics, and clinical measurements. The most common measurements currently being used in gut microbiome research and sample testing laboratories are next generation sequencing (NGS)-based metagenomics and MS-based metabolomics. However, in either case, no fit-for-purpose RMs exist that enable researchers to compare results generated across different laboratories and to assess the impact of the multitude of methodological variables that exist in either measurement platform. To begin to understand the biologically relevant properties of the human gut microbiome, the community needs such RMs for confident identification of relevant biomarkers that may serve as health or disease indicators while supporting the validation of analytical measurements (mass fraction or mass concentration) for clinically relevant metabolites and/or nutritional assessment metabolic markers.

This candidate Human Whole Stool RM will comprise two types of human whole stool materials—one from healthy donors who are omnivores and one from healthy vegan and vegetarian donors. The homogeneity and stability, with respect to the metabolomic and metagenomic profiles of their complex microbial communities, of both materials will be thoroughly assessed. Similar to the NIST metabolomics plasma and urine RMs (in the development pipeline), the predominant metabolites will be annotated/identified before the RM is made publicly available. The metabolites annotated or identified will primarily be those for which current evidence supports a link to physiological effects or better health outcomes, and those which are consistently present in human stool at high levels. Additionally, DNA will be quantified and metagenomic profiles will be evaluated via NGS. Certification of any preliminary value assignments may not be necessary for the initial harmonization of metabolomics-based whole stool measurements. However, NIST can evolve its RMs toward Certified Reference Materials (CRMs) when deemed necessary by stakeholders. Thus, if a targeted metabolic biomarker is identified as a critical measure or indicator for stool-based measurements, options are available to “upgrade” the Human Whole Stool RM to include certified values. This generally involves conducting further measurements exclusively at NIST using reference measurement procedures or additional higher-order measurement procedures that are often metrologically traceable to the SI.

NIST has already begun working with TBC to evaluate their human whole stool preparation and preservation methodologies to determine how best to stabilize and store candidate RMs for both metagenomic and metabolomic measurements over an extended period. Additionally, NIST is working with ILSI North America to determine metabolites of interest (i.e., linked to health outcomes) that should be annotated/identified in the RM even if they are not the most dominant biochemical species present.

4 NIST pilot study on stool stability and homogeneity

As a first step to development of the Human Whole Stool RM, the candidate materials need to be stable and homogeneous for specified targeted measurands. We performed a pilot study to investigate stool sample storage conditions that could potentially stabilize the targeted metabolites present in the material prepared by TBC using its proprietary technologies for processing, stabilization, and storage. The production of candidate stool RMs will be based largely on

the manufacturing process illustrated in Fig. 1. A similar process was used to produce the material employed in this pilot study.

Metabolites were extracted from stool samples that were either homogenized in water and stored at -80°C or 4°C , lyophilized and stored at 4°C , or homogenized in a preservation buffer (two buffers were used, defined as 1 and 2) and stored at 4°C . In total, five different conditions were analyzed as shown in Fig. 2. The ideal preservation technique would be amenable to various measurement platforms (fit-for-purpose). For the purposes of this pilot study, microbial taxa and metabolite patterns across multiple aliquots from the same storage condition were evaluated and used to estimate homogeneity within that storage condition. Extensive homogeneity studies will be performed on the candidate RM in the future, once preservation methods have been optimized. The stool samples were stored for 4 months under the above conditions prior to analysis. This pilot study did not include any analyses of the freshly prepared material ($t=0$ timepoint), therefore stability over time could not be assessed. Only the homogeneity and the impact of the various storage conditions was assessed.

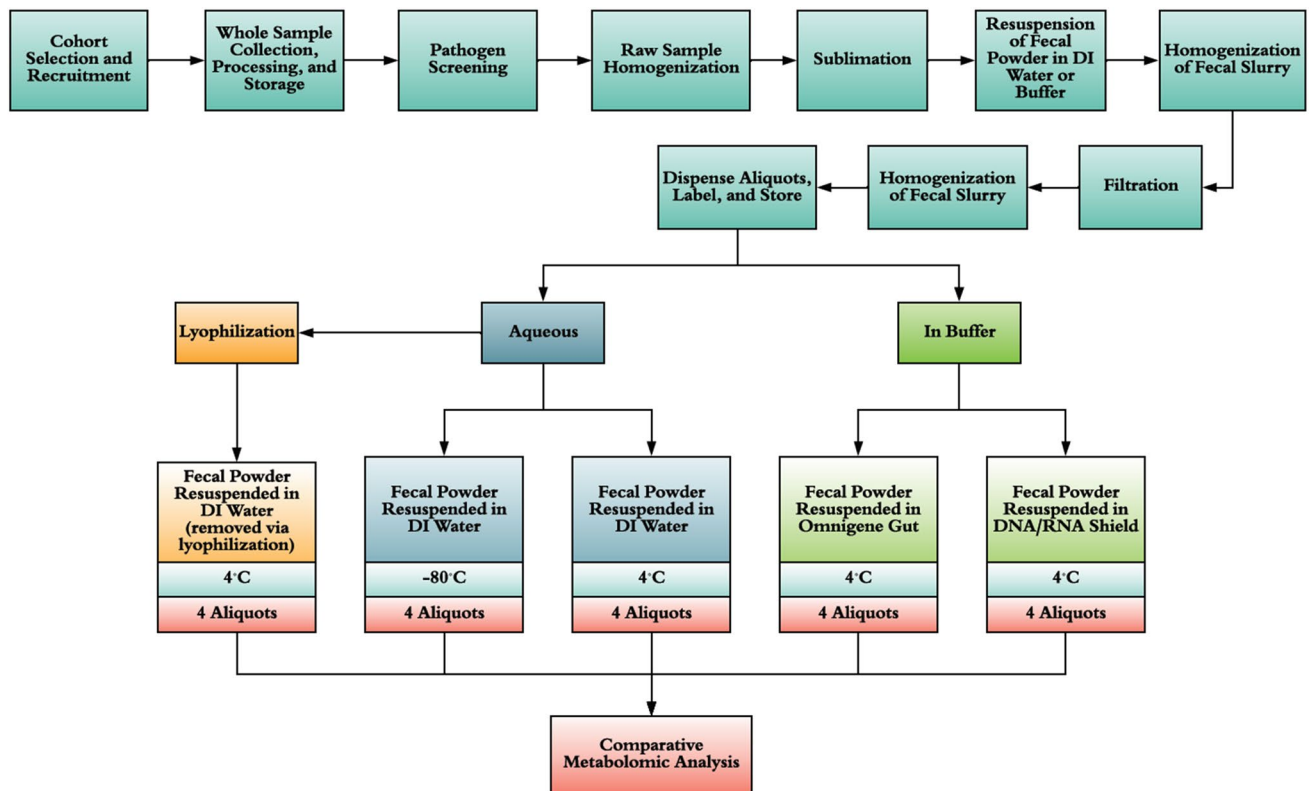
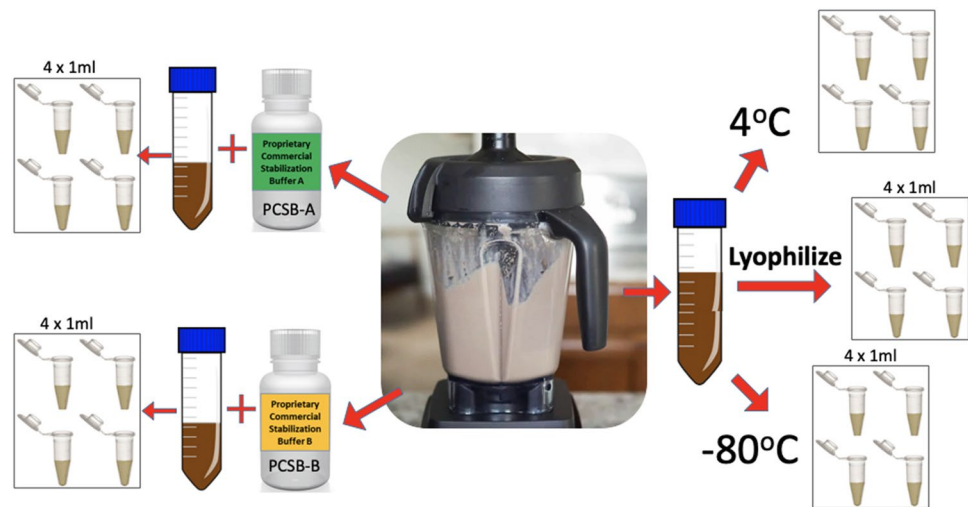


Fig. 1 Manufacturing flow of the TruMatrix™ (The BioCollective, LLC) Microbiome Reference Material

Fig. 2 Manufacturing flow for the five different stabilization methods used for the homogenized stool sample



Samples were sent to NIST from TBC and analyzed via NMR-based metabolomics, MS-based metabolomics, and 16S rRNA amplicon DNA sequencing. These three analytical approaches provide a picture of the taxonomic and metabolic profiles of the microbial community (Butlen and Jard 1972; Wishart 2019b). MS and NMR are complementary approaches in the detection of small molecules (< 1.5 kDa) and widely used analytical platforms for untargeted metabolome characterization. In terms of RM development, MS-based metabolomics covers a large range of metabolites due to its analytical sensitivity, thus it constitutes an ideal technique for quantifying low-level biomarkers. MS-based analyses often require a separation on the front end of the workflow. GC is ideal for the detection of volatile and semi-volatile compounds, which are highly abundant in whole stool samples. LC covers a wide range of compounds and is regarded as the gold standard separation technique for metabolomics. Both GC-MS and LC-MS analyses are prone to reproducibility challenges due to batch effects and the innate instability of MS, which can be mitigated through careful sample preparation and the use of labeled internal standards. Utilizing NMR for gut microbiome metabolomic studies is advantageous as this technique offers high reproducibility and absolute quantification. Nearly all protons have an equivalent response across molecules. Moreover, it offers simplicity in sample preparation. Metabolite profiles obtained by NMR are virtually independent of the operator and instrument parameters, which provides a high degree of reliability to the results. Additionally, since samples do not interact directly with the instrument, NMR analysis is free from batch effects such as carryover. However, NMR does not provide a global analysis due to low sensitivity, which is typically in the micromolar range as opposed to the nanomolar range provided by MS. NMR is an ideal analysis for monitoring the stability

and homogeneity of the metabolites present in a fecal RM as well as absolute quantification of some metabolites.

To keep within the scope of this manuscript, which is to summarize the September 2019 workshop, we focus here on a global discussion of the sample analysis. The five preservation methods presented unique challenges for the characterization of the samples. In the NMR analysis, pH differences among the samples led to peak shifting of acidic metabolites such as acetic acid. Additionally, the use of ethanol in the proprietary commercial preservation solutions significantly hindered spectral analysis by introducing strong peaks in multiple regions of the spectrum that were much larger than neighboring metabolite peaks. In LC-MS analysis, one of the proprietary buffers contained extremely high levels of the DNA stabilizer, N-dodecanoyl-N-methylglycine, which led to a large interference in the analysis and instrument contamination. Overall, the results of all analyses were consistent with each other in that within a preservation technique, the metabolite profiles appear to be homogeneous. However, different preservation techniques introduce varying metabolite biases. 16S rRNA amplicon gene sequencing indicated changes in the relative abundance of the microbial taxa that presumably occurred during storage under different conditions. The microbial community profile of the samples homogenized in water and stored at 4 °C was most different from all other preservation and storage conditions, likely indicating relative abundance shifts in the microbial population due to potential growth during storage. In fact, many bacterial species can sustain growth at 4 °C, particularly given the abundance of nutrients in fecal material. The results of the combined metagenomic and metabolomic pilot study

Table 1 Potential list of candidate fecal metabolites of health relevance*Candidate metabolites*

Bilirubin

Butyric acid

Enterolactone (bacterial metabolite of plant lignans) and/or its glucuronide and sulfate conjugates

Odd chain fatty acid such as 15:0 or 17:0

Lithocholic acid (secondary bile acid) and/or glycine and taurine conjugates (or other secondary bile acid)

Hippuric acid

Indole

Isoflavone derivatives (for complex ring structures)

p-Cresol

Putrescine

Phenyl lactic and p-hydroxyphenyl lactic acid

Trimethylamine oxide (bacterial trimethylamine oxidized to TMAO in liver)

TMAO trimethylamine N-oxide

This list was generated by the planning committee and brought to the workshop for participant reaction

suggest that lyophilized samples as well as samples stored at $-80\text{ }^{\circ}\text{C}$ are the most suitable for RM development.

5 Criteria for reference material metabolite selection

One of the key objectives of the September 2019 meeting was to develop a consensus for the stool RM metabolite selection. Metabolites for further vetting were initially generated using three approaches: (i) expert consensus on metabolites of health relevance (Table 1); (ii) microbially produced compounds as catalogued in the Fecal Metabolome Database (www.fecalmetabolome.ca) (Table 2); and (iii) a PubMed search with chemical synonyms (Table 3). Tables in this paper were generated through an analysis of a number of online databases as of September 2019, such as the HMDB (Wishart et al. 2018; www.hmdb.ca), the fecal metabolome database (www.fecalmetabolome.ca), as well as through literature reviews conducted by the authors and feedback provided by workshop participants.

The following general principles were proposed as criteria for selecting metabolites to be eventually characterized in the RM: The metabolites should (1) have health or disease relevance, (2) be identifiable as part of a known biological pathway, (3) be common in a diversity of fecal samples, (4) have demonstrated robustness in the analytical method used for measurement, (5) be considered in light of the variety of analytical platforms that may be used and research foci, (6) include specific analytes that

are challenging to measure (considering retention time, increased volatility, trace concentration levels, etc.), and finally, (7) include analytes with a range of chemical stability that are known to change over time frames relevant to RM production.

A modified list of potential candidate metabolites was generated through discussion groups and report-back during the September 19 event (Table 4) from a careful review of Tables 1, 2 and 3 and considering the criteria outlined above. This list and the accompanying criteria are being provided to NIST as a guide as they engage in the whole stool RM development process. At present, knowledge is limited with respect to how sample collection affects the presence and measurement of these metabolites, how well metabolite profiles in fecal samples represent digesta in the gastrointestinal tract, and which metabolites are linked to measurable physiological effects or health benefits. It is recognized that the metabolites characterized in the stool RM can, and will, evolve in parallel with progressive developments in stool metabolomic approaches as well as advances in nutrition and gut microbiome science, as linked to health and disease. Future opportunities include understanding the full spectrum of health relevance for various metabolite classes, and consideration of individuals consuming various diets (e.g., high fiber, strict vegan, high animal protein/fat) and with various disease states/conditions (e.g., irritable bowel disease/syndrome [active], obese vs. lean). Prior to the creation of a stool RM, a typical requirement is knowledge of the dynamic range of metabolites in a complex matrix.

Table 2 List of candidate microbially produced compounds from the Fecal Metabolome Database (www.fecalmetabolome.ca)

Compound name ^a	InChi key
SCFAs	
Propionic acid	XBDQKXXYIPTUBI-UHFFFAOYSA-N 1
Acetic acid	QTBSBXVTEAMEQO-UHFFFAOYSA-N 1
Formic acid	BDAGIHXXWWSANSR-UHFFFAOYSA-N 1
Isobutyric acid	KQNPFQTWMSNSAP-UHFFFAOYSA-N 1
Valeric acid	NQPDZGKIBAWPEJ-UHFFFAOYSA-N 1
Isovaleric acid	GWYFCOCPABKNJV-UHFFFAOYSA-N 1
(S)-3,4-Dihydroxybutyric acid	DZAIQXUZHHTJKN-UHFFFAOYSA-N 1
1,2,3-Propanetricarboxylic acid	KQTIIECAUMSDG-UHFFFAOYSA-N 1
1,2,3-Trihydroxybenzene	WQGWDDDDVZFFDIG-UHFFFAOYSA-N 1
1-Butanol	LRHPLDYGYMQRHN-UHFFFAOYSA-N 1
1-Deoxy-D-xylulose 5-phosphate	AJPADPZSRRUGHI-RFZPGFLSSA-N 1
1H-Indole-3-carboxaldehyde	OLNJUISKUQQNIM-UHFFFAOYSA-N
2,3-Butanediol	OWBTYPJTUOEWEK-UHFFFAOYSA-N 1
2,5-Furandicarboxylic acid	CHTHALBTIRVDBM-UHFFFAOYSA-N 1
2-Hydroxycaproic acid	NYHNVHGFPZAZGA-UHFFFAOYSA-N 1
2-Hydroxyglutarate	HWXBTVNAVRSUOJR-UHFFFAOYSA-N 1
2-Isopropylmalic acid	BITYXLXUCSKTJS-ZETCQYMHSA-N 1
2-Keto-3-deoxy-6-phosphogluconic acid	OVPRPPOVAXRCED-WVZVXSGGSA-N 1
2-Oxo-3-phenylpropanoic acid (Mixture oxo and keto)	BTNMPGBKDVTSJY-UHFFFAOYSA-N 0
2-oxoglutarate(2-)	KPGXRSRHYNQIFN-UHFFFAOYSA-L 0
2-Phenylpropionic acid	YPGCWEMNNLXISK-UHFFFAOYSA-N
3-(3-Hydroxyphenyl)-3-hydroxypropanoic acid	KHTAGVZHYUZYM-F-UHFFFAOYSA-N 1
3-(3-Hydroxyphenyl)propanoic acid	QVWAEZJXDYOKEH-UHFFFAOYSA-N 1
3-(Methylthio)-1-propene	NVLPQIPTCCLBEU-UHFFFAOYSA-N 1
3,4-Dihydroxybenzeneacetic acid	CFDZDZCDUFOSFZ-UHFFFAOYSA-N 1
3,4-Dihydroxyhydrocinnamic acid	DZAUWHJDUNRCTF-UHFFFAOYSA-N 1
3-Hydroxybenzoic acid	IJFXRHURBJZNAO-UHFFFAOYSA-N 1
3-Hydroxybutyric acid	WHBMMWSBFZVSSR-GSVOUGTGSA-N 1
3-Hydroxyhippuric acid	XDOFWFNMYJRHEW-UHFFFAOYSA-N 1
3-Hydroxyphenylacetic acid	FVMDYYGIDFPZAX-UHFFFAOYSA-N 1
3-Indolebutyric acid	JTEDVYBZBROSJT-UHFFFAOYSA-N 1
3-Methylindole	ZFRKQXVRDFCRJG-UHFFFAOYSA-N 1
3-O-Methylgallic acid	KWCCUYSXAYTNKA-UHFFFAOYSA-N 0
4-Hydroxybenzoic acid	FJKROLUGYXJWQN-UHFFFAOYSA-N 1
4-Hydroxybutyric acid	SJZRECIHVVDYJC-UHFFFAOYSA-N 1
4-Hydroxycyclohexylcarboxylic acid	HCFRWBBJISAZNK-UHFFFAOYSA-N 1
4-Hydroxyvaleric acid	FMHKPLXYWVCLME-UHFFFAOYSA-N 0
4-O-Methylgallic acid	UBXDWYFLYIJQFR-UHFFFAOYSA-N 1
5-(3',4'-Dihydroxyphenyl)-gamma-valerolactone	ZNXXWTPQHVLMT-MRVPVSSYSA-N 1
5-Amino-6-(5'-phosphoribitylamino)uracil	RQRINYSISXYAZKL-RPDRRWSUSA-N 1
5beta-Coprostanol	QYIXCDOBOSTCEI-NWKZBHTNSA-N 1
5-Hydroxyhexanoic acid	YDCRNMIJQROAWFT-UHFFFAOYSA-N 1
5-Hydroxyindoleacetic acid	DUUGKQCEGZLZNO-UHFFFAOYSA-N 1
5-Hydroxypentanoic acid	PHOJOSOUIAQEDH-UHFFFAOYSA-N 1
5-Keto-D-gluconate	IZSRJDGCGRAUAR-UHFFFAOYSA-N 1
6,7-Dimethyl-8-(1-D-ribityl)lumazine	SXDXRJZUAJBNFL-XKSSXDPKSA-N 1
Acetaldehyde	IKHGUXGNUITLKF-UHFFFAOYSA-N 1
Acetone	CSCPPACGZOOZCGX-UHFFFAOYSA-N 1
Acetylcholine	OIPILFWXSMYKGL-UHFFFAOYSA-N 1
ADP-glucose	WFPZSXYXPSUOPY-ROYWQJLOSA-N 1
Allantoin	POJWUDADGALRAB-UHFFFAOYSA-N 1
Ascorbic acid	CIWBSHSHKDKBQ-JLAZNSOCSA-N 1
Benzoic acid	WPYMKLBDIGXBTP-UHFFFAOYSA-N 1

Table 2 (continued)

Compound name ^a	InChi key
Benzoyl-CoA	VEVJTUNLALKRNO-TYHXJLICSA-N 1
Cadaverine	VHRGRCVQAFMJIZ-UHFFFAOYSA-N 1
Caffeic acid	QAIPRVGONGVQAS-DUXPYHPUSA-N 1
cis,cis-Muconic acid	TXXHDPDFNKHGGW-CCAGOZQPSA-N 1
Citramalic acid	XFTRTWQBIOMVPK-UHFFFAOYSA-N 1
D-Alanine	QNAYBMKLOCPYGJ-UWTATZPHSA-N 1
D-Arabinose	SRBFZHDQGSBBOR-SQOUGZDYSA-N 1
D-Arabitol	HEBKCHPVOIAQTA-QWWZWVQMSA-N 1
Desaminotyrosine	NMHMNPHRMNGLLB-UHFFFAOYSA-N 1
D-Glutamic acid	WHUUTDBJXRKMK-GSVOUGTGSA-N 1
Diaminopimelic acid	GMKMEZVLHJARHF-WHFBIAKZSA-N 1
Dimethyl sulfone	HHVIBTZHLRERCL-UHFFFAOYSA-N 1
Dimethyl trisulfide	YWHLKYXPLRWGSE-UHFFFAOYSA-N 1
Dimethylamine	ROSDSFDQCJNGOL-UHFFFAOYSA-N 1
Dimethylsulfide	QMMFVYPAHWCMCS-UHFFFAOYSA-N 1
D-Lactic acid	JVTAAEKCZFNVCJ-UWTATZPHSA-N 1
Dopamine	VYFYTYLLBUKUH-UHFFFAOYSA-N 1
D-Ribose	HMFHBZSHGGEWLO-SOOFDHNKSA-N 1
D-Ribulose	LQXVFWRQNMEDEE-UOWFLXDJSA-N 1
Ethanol	LFQSCWFLJHTTHZ-UHFFFAOYSA-N 1
Etiocholanedione	RAJWOBJTTGJROA-QJISAEMRSA-N 1
Folic acid	OVBPIULPVIDEAO-LBPRGKRZSA-N 1
Galactan	PTHCMJGKKRQCBF-ICIGWGHZSA-N 1
Gallic acid	LNTHITQWFMADLM-UHFFFAOYSA-N 1
gamma-Aminobutyric acid	BTCSSZJGUNDROE-UHFFFAOYSA-N 1
Gentisic acid	WXTMDXOMEHJXQO-UHFFFAOYSA-N 1
Glutaric acid	JFCQEDHGNNZCLN-UHFFFAOYSA-N 1
Glyceric acid	RBNPOMFGQQGHHO-UWTATZPHSA-N 1
Glycocholic acid	RFDAIACWWDREDC-FRVQLJSFSA-N 1
Glycolic acid	AEMRFAOFKBGASW-UHFFFAOYSA-N 1
Homogentisic acid	IGMNYECMUMZDDF-UHFFFAOYSA-N 1
Homovanillic acid	QRMZSPFSDQBLIX-UHFFFAOYSA-N 1
Hydrocinnamic acid	XMIIGOLPHOKFCH-UHFFFAOYSA-N 1
Hydroxyisocaproic acid	LVRFTAAXQPQHI-YFKPBYRVSA-N 1
Hydroxyphenyllactic acid	JVGVDSSUAVXRDY-UHFFFAOYSA-N 1
Hydroxypropionic acid	ALRHLSYJTWAHJZ-UHFFFAOYSA-N 1
Hydroxytyrosol	JUUBCHWRXWPFH-UHFFFAOYSA-N 1
Indole	SIKJAJRHWYJAI-UHFFFAOYSA-N 1
Indole-3-methyl acetate	KTHADMDGDNYQRX-UHFFFAOYSA-N 1
Indole-3-propionic acid	GOLXRNDWAUTYKT-UHFFFAOYSA-N 1
Indoleacetic acid	SEOVTRFCIGRIMH-UHFFFAOYSA-N 1
Indoleacrylic acid	PLVPLCLBIEYEA-AATRIKPKSA-N 1
Indolepyruvate	RSTKLPZEZYQPY-UHFFFAOYSA-N 1
Indoxyl sulfate	BXFFHSIDQOFMLE-UHFFFAOYSA-N 1
Isobutanol	ZXEKIIBDNHEJCQ-UHFFFAOYSA-N 1
Isopropyl alcohol	KFZMGEQAYNKOFK-UHFFFAOYSA-N 1
L-Gulose	WQZGKKKJIFFOK-QRQFDPRISA-N 1
Lipid A	GZQKNULLWNGMCW-PWQABINMSA-N 1
L-Lactic acid	JVTAAEKCZFNVCJ-REOHCLBHSAN 1
L-Rhamnulose	QZNPKNKJXABGCRC-OTWZMJIISA-N 1
L-Sorbose	LKDRXBCSQODPBY-BGPJRJDNSA-N 1
Mannitol	FBPFZTCFMRRESA-KVTDHHQDSA-N 1
Mannitol 1-phosphate	GACTWZZMVMUKNG-KVTDHHQDSA-N 1
m-Cresol	RLSSMJSEOOYNOY-UHFFFAOYSA-N 1

Table 2 (continued)

Compound name ^a	InChi key
Melibiose	DLRVVLDZNNYCBX-CQHUIXDMSA-N 1
Methylisocitric acid	HHKPKXCESHMJWCF-UHFFFAOYSA-N 1
Monodehydroascorbate	PIECNOFONCLXPG-SAGVERJBASA-N 0
N-Acetylmannosamine	OVRNDRQMDRJTHS-OZRXBAMMSA-N 1
N-Acetylputrescine	KLZGKIDSEJWEDW-UHFFFAOYSA-N 1
N-benzoylglycinate	QIAFMBKCNZACKA-UHFFFAOYSA-N 0
Nonadecanoic acid	ISYWECDZWTKFF-UHFFFAOYSA-N 1
Norepinephrine	SFLSHLFXELFNJZ-QMMMGOBSA-N 1
o-Cresol	QWVGKYWNOKOFNN-UHFFFAOYSA-N 1
Oxalic acid	MUBZPKHOEPUJKR-UHFFFAOYSA-N 1
Oxoglutaric acid	KPGXRSRHYNQIFN-UHFFFAOYSA-N 1
p-Aminobenzoic acid	ALYNCZNDIQEVRV-UHFFFAOYSA-N 1
p-Cresol sulfate	WGNAKZGUSRVWRH-UHFFFAOYSA-N 1
Phenol	ISWSIDIOOJBQZ-UHFFFAOYSA-N 1
Phenylacetylglutamine	JFLIEFSWGNOPJJ-JTQLQIEISA-N 1
Phenylethylamine	BHHGXPLMPWCGHP-UHFFFAOYSA-N 1
Phenyllactic acid	VOXXWSYKYCBWHO-UHFFFAOYSA-N 1
Phloretin	VGEREEWJJVICBM-UHFFFAOYSA-N 1
p-Hydroxyphenylacetic acid	XQXPVVBIMDBYFF-UHFFFAOYSA-N 1
Propyl alcohol	BDERNNFJNPAEC-UHFFFAOYSA-N 1
Protocatechuic acid	YQUVCSBJEUQKSH-UHFFFAOYSA-N 1
Putrescine	KIDHWZJUCRJVML-UHFFFAOYSA-N 1
Pyrocatechol	YCIMNLLNPGFGHC-UHFFFAOYSA-N 1
Pyrrolidine	RWRDLPLDKQPQOW-UHFFFAOYSA-N 1
Pyruvic acid	LCTONWCANYUPML-UHFFFAOYSA-N 1
Rhamnose	SHZGCJCMOBCMCKK-JFNONXLTSA-N 1
Salicylic acid	YGSDEFSMJLZEOE-UHFFFAOYSA-N 1
Serotonin	QZAYGJVTTNCVMB-UHFFFAOYSA-N 1
Sorbose 1-phosphate	ZKLLSNQJRLJIGT-OTWZMJIIISA-N 1
Stercobilin	DEEUSUJLZQESV-BQUSTMGCSA-N 1
Succinic acid	KDYFGRWQOYBRFD-UHFFFAOYSA-N 1
Sulfoglycolithocholate(2-)	FHXBAFXQVZOILS-UHFFFAOYSA-N 1
Sumiki's acid	PCSKKIURRTAEM-UHFFFAOYSA-N
Syringic acid	JMSVCTWVEWCHDZ-UHFFFAOYSA-N 1
Tartaric acid	FEWJPZIEWOKRBE-JCYAYHJZSA-N 1
Taurodeoxycholic acid	AWDRATDZQPNJFN-VAYUFCLWSA-N 1
trans-Cinnamic acid	WBYWAXJHAXSJNI-VOTSOKGWSA-N 1
trans-Ferulic acid	KSEBMYQBYZTDHS-HWKANZROSA-N
Trehalose	HDTRYLNUVZCQOY-LIZSDCNHSA-N 1
Trehalose 6-phosphate	LABSPYBHMPDTEL-LIZSDCNHSA-N 1
Tryptamine	APIYDQYYACXCRM-UHFFFAOYSA-N 1
Tryptophanol	MBBOMCVGYCRMEA-UHFFFAOYSA-N
Tuberculostearic acid	BEOUGZFCUMNGOU-UHFFFAOYSA-N 1
Tyraminium	DZGWFCGJZKJUFU-UHFFFAOYSA-O 0
Tyrosol	YCCILVSKPBXVIP-UHFFFAOYSA-N 1
Vanillic acid	WKOLLVMJNQIZCI-UHFFFAOYSA-N 1
Zeaxanthin	JKQXZKUSFCKOGQ-QAYBQHTQSA-N 1

^aCompounds are listed in alphabetical order from top to bottom, left to right. The international chemical identifier (InChi) keys can be found at www.hmdb.ca

Four different routes that can be used to determine if a metabolite is microbially or host produced: (1) literature review, (2) published pathways, (3) genomic reconstruction evidence, and (4) evidence from experimental work for robogut (a bioreactor that allows researchers to study the colon) and SHIME (The Simulator of the Human Intestinal Microbial Ecosystem) studies

6 Conclusions and next steps for creating and evolving the material

Given the rapidly evolving nature of gut microbiome science and the current state of knowledge, an RM (as opposed to a CRM) measured for multiple metabolites is appropriate at this stage. As the science evolves, the RM can evolve to match the needs of the research community. Ultimately, the stool RM may exist in sequential versions. Beneficial to this evolution will be a clear line of communication between NIST and the stakeholder community to ensure alignment with current scientific understanding and community needs.

This document represents the consensus opinions of participants and co-authors of this manuscript, as captured in a September 12, 2019 workshop to explore creation of a human whole stool RM. ILSI North America, NIST, and TBC have engaged in Phase II of this effort which consists of a stability study of human stool material prepared from omnivore and vegan cohorts, each treated by freezing at $-80\text{ }^{\circ}\text{C}$ or lyophilized. NIST will perform the metabolomics stability assessment on cohort aliquots prepared by TBC, including the following: (1) identifying “conserved” metabolites that can be used to assess homogeneity and stability, (2) assessing the effects of freeze–thaw cycles, (3) assessing microheterogeneity with respect to subsampling size, (4) assessing the effects of different extraction methods and solvents, (5) evaluating the presence of intracellular vs. extracellular metabolites, and (6) assessing the reproducibility across 5 to 10 laboratories through an interlaboratory study. With respect to metagenomic measurements, NIST will also assess the utility of spike-in organisms that can serve as pseudo-ground truth and provide an understanding of the analytical sensitivity and specificity of metagenomic measurements.

While there may be questions about whether this is the “right time” to develop a RM, it is important to remember that in much the same way that the metagenomics community has strived to develop and adopt standards, the functional microbiome and microbial metabolomic communities must also begin to develop and adopt standards for metabolomic measurements. In doing so, these measurements can be employed to make informed, clinically relevant, actionable decisions regarding the state of an individual’s gut microbiome, including disease predispositions and possible therapeutic or dietary intervention strategies aimed to benefit health.

Table 3 Typical microbe-source compounds: based on PubMed search with chemical synonyms

Metabolite	Reference count
Ethanol	3233
Phenol	2444
Folic acid	2253
L-Lactic acid	2020
Acetic acid	2006
D Lactic acid	1512
Butyric acid	1426
Benzoic acid	1172
Indole	1080
Propionic acid	911
Formic acid	699
Bilirubin	586
D-Alanine	569
D-Glutamic acid	547
Acetaldehyde	531
2-Hydroxyglutarate	485
Galactan	437
Serotonin	407
Pyrocatechol	385
Lithocholic acid	319
Pentadecanoic acid	278
Ascorbic acid	256
Succinic acid	238
Mannitol	225
Pyruvic acid	221
Oxalic acid	190
Propyl alcohol	184
Tryptamine	177
Valeric acid	177
Phenylethylamine	167
Dopamine	134
Acetone	131
Acetylcholine	128
p-Cresol	126
Hippuric acid	121
Norepinephrine	120
Gallic acid	117
Isobutyric acid	108
Rhamnose	102

Search results through August 12, 2019 are included. Only metabolites with 100 or more indicated references were included. Searches were conducted using PubMed with chemical synonyms and were based on abstracts only

Table 4 Modified list of potential metabolites for value assignment in a NIST Reference Material

Candidate metabolite	Alignment with proposed selection criteria ^a	Example references (alignment if specific)
SCFAs (butyric acid, propionic acid, acetic acid, formic acid)	1, 2, 3, 4, 6, 7, 8	Alexander et al. (2019) and Zhang et al. (2019) (4)
Succinic acid	1, 2, 3, 6	Fernández-Veledo and Vendrell (2019) and Xu and Wang (2016)
Lipids (e.g. linoleic acid; sphingomyelins [initially select based on abundance])	1, 2, 3, 6	Wang et al. (2016)
Deoxycholic acid	1, 2	Hullar and Fu (2014) (1) and Jewison et al. (2014) (2)
Enterohepatic metabolites (bacterial metabolites that are modified in the liver [TMAO])	1, 3, 6	Karu et al. (2018) (3) and Xu et al. (2015) (1, 6)
Tryptamine	1, 2, 6	Jewison et al. (2014) (2) and Williams et al. (2014) (1, 2)
Gallic acid	1, 6	Li et al. (2019b)
Tryptophan, serotonin (neurotransmitters)	1, 3, 2, 6	Karu et al. (2018) (3 for tryptophan), Jewison et al. (2014) (2) and Strandwitz (2018) (1)
Enterolactone (bacterial metabolite of plant lignans) and/or its glucuronide and sulfate conjugates	1	Halldin et al. (2019)
GABA	1, 3, 2, 6	Karu et al. (2018) (3), Strandwitz (2018) (1) and Jewison et al. (2014) (2)
Branched-chain fatty acids (isovaleric acid; isobutyric acid)	1, 2, 3	Abdallah et al. (2020) and Canfora et al. (2019)
Salicylic acid	1	Damman (2013)
Phenols and indoles (phenol, indole, p-cresol)	1, 3	Abdallah et al. (2020)
LPS with o-antigen (byproduct of Gram-negative bacteria)	1	Reyes et al. (2012)
Trans-ferulic acid	1	Xu and Wang (2016) and Agulló et al. (2020)
Mannitol	1	Xu and Wang (2016)
4-Pyridoxic acid	1	Obeid et al. (2019) (1)
Carnosine and trans-ferulic acid	1, 2, 6	Jewison et al. (2014) (2 for carnosine) and Schon et al. (2019) (1 for carnosine)
Theophylline	1	Cyr et al. (2008)

GABA gamma-aminobutyric acid, *LPS* lipopolysaccharide, *SCFA* short-chain fatty acid, *TMAO* trimethylamine N-oxide

This list is neither exhaustive nor definitive, but rather includes some options discussed by workshop participants

^aProposed criteria noted include that the metabolite: 1, is of health or disease relevance; 2, is identifiable as part of a known biological pathway; 3, is common in a diversity of fecal samples; 4, demonstrates confidence in the analytical method used for measurement; 5, is considered in light of the variety of analytical platforms that may be used and research foci; 6, is selected considering both technical and biological aspects; 7, is challenging to capture; 8, is an analyte of substantial stability or is known to change over time

Acknowledgements Participants present for the September 12, 2019 workshop held on the NIST campus that have agreed to be acknowledged for their contributions to the discussion follow:

Olle de Bruin: DNA Genotek, Inc.; Oliver Chen: Biofortis, Merieux NutriSciences; Jo May Chow: Abbott; Latha Devareddy: The Coca-Cola Company; Dylan Dodd: Stanford University; Oliver Fiehn: University of California – Davis; Leigh A. Frame: George Washington University; Michael Gidding: Kaleido Biosciences, Inc.; Meredith Hullar: Fred Hutchinson Cancer Research Center; Curtis Huttenhower: Harvard University; Bruce S. Kristal: Harvard University; Tracey Schock: NIST; Emily Vogtmann: DGEG/NCI/NIH; Syeda Fatima Zaidi: George Mason University. Individuals that did not respond to the acknowledgement request, but that did attend the event follow: Clary Clish: The Broad Institute; Samuel Forry: NIST; Benjamin Scott: NIST; Sonia Hartunian-Sowa : DSM; Rima Kaddurah-Daouk: Duke University; Sun Eun Lee: Johns Hopkins University; Mark Lowenthal: NIST; Tytus Mak: NIST; Guilherme Pinheiro: NIST; Rebecca Ann Replogle: PepsiCo, Inc.; Benjamin Scott: NIST; Kieran Tuohy: Edmund Mach Foundation of San Michele.

Author contributions PP and SMDS analyzed the samples prepared by The BioCollective, LLC to generate the data presented in this report. RM, RC, CDD, DH, SAJ, CMJ, JWJ, MEL, NJL, KAL, KSS, and DSW conceived and designed the project. All authors contributed to the manuscript. All authors read and approved the manuscript.

Funding This work was partially supported by the ILSI North America Gut Microbiome Committee. ILSI North America is a public, non-profit science foundation that provides a forum to advance understanding of scientific issues related to the nutritional quality and safety of the food supply. ILSI North America is funded primarily by its industry membership. RM, RC, JWJ, and KSS received travel funding to attend the September 2019 event.

Data availability Not applicable.

Code availability Not applicable.

Compliance with ethical standards

Conflict of interest RC is employed by The BioCollective, LLC, a company that specializes in the preparation of fecal samples for a variety of research applications. DH is employed by Mondelez, a global snacking company. RM, CDD, SAJ, CMJ, JWL, MEL, NJL, KAL, PP, SMDS, KSS, and DSW have no conflicts to declare.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Consent to participate The BioCollective collects samples using an institutional review board (IRB)-approved protocol that includes consent of the donor. All human stool samples were received at NIST after informed consent under approved IRB protocols reviewed by the NIST Human Subjects Protection Office.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdallah, A., Elemba, E., Zhong, Q., & Sun, Z. (2020). Gastrointestinal interaction between dietary amino acids and gut microbiota: With special emphasis on host nutrition. *Current Protein & Peptide Science*. <https://doi.org/10.2174/1389203721666200212095503>
- Agulló, V., Villaño, D., García-Viguera, C., & Domínguez-Perles, R. (2020). Anthocyanin metabolites in human urine after the intake of new functional beverages. *Molecules*, *25*(2), 371.
- Ahmed, I., Greenwood, R., Costello Bde, L., Ratcliffe, N. M., & Probert, C. S. (2013). An investigation of fecal volatile organic metabolites in irritable bowel syndrome. *PLoS ONE*, *8*(3), e58204.
- Alexander, C., Swanson, K. S., Fahey, G. C., Jr., & Garleb, K. A. (2019). Perspective: Physiologic importance of short-chain fatty acids from nondigestible carbohydrate fermentation. *Advances in Nutrition*, *10*(4), 576–589.
- Allen-Vercoe, E., Carmical, J. R., Forry, S. P., Gail, M. H., & Sinha, R. (2019). Perspectives for consideration in the development of microbial cell reference materials. *Cancer Epidemiology, Biomarkers and Prevention*, *28*(12), 1949–1954.
- Anand, S., & Mande, S. S. (2018). Diet, microbiota and gut-lung connection. *Frontiers in Microbiology*, *9*, 2147.
- Anderson, E. L., Li, W., Klitgord, N., Highlander, S. K., Dayrit, M., Seguritan, V., et al. (2016). A robust ambient temperature collection and stabilization strategy: Enabling worldwide functional studies of the human microbiome. *Scientific Reports*, *6*, 31731.
- Arndt, D., Xia, J., Liu, Y., Zhou, Y., Guo, A. C., Cruz, J. A., et al. (2012). METAGENassist: A comprehensive web server for comparative metagenomics. *Nucleic Acids Research*, *40*(Web Server Issue), W88–W95.
- Bergman, E. N. (1990). Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. *Physiological Reviews*, *70*(2), 567–590.
- Butlen, D., & Jard, S. (1972). Renal handling of 3'-5'-cyclic AMP in the rat. The possible role of luminal 3'-5'-cyclic AMP in the tubular reabsorption of phosphate. *Pflügers Archiv: European Journal of Physiology*, *331*(2), 172–190.
- Canfora, E. E., Meex, R. C. R., Venema, K., & Blaak, E. E. (2019). Gut microbial metabolites in obesity, NAFLD and T2DM. *Nature Reviews Endocrinology*, *15*(5), 261–273.
- Carbonero, F., Benefiel, A., Alizadeh-Ghamsari, A., & Gaskins, H. R. (2012). Microbial pathways in colonic sulfur metabolism and links with health and disease. *Frontiers in Physiology*, *3*, 448.
- Cesbron, N., Royer, A. L., Guitton, Y., Sydor, A., Le Bizec, B., & Dervilly-Pinel, G. (2017). Optimization of fecal sample preparation for untargeted LC-HRMS based metabolomics. *Metabolomics*, *13*(8), 99.
- Chen, L., Wang, W., Zhou, R., Ng, S. C., Li, J., Huang, M., et al. (2014). Characteristics of fecal and mucosa-associated microbiota in Chinese patients with inflammatory bowel disease. *Medicine*, *93*(8), e51.
- Chiang, J. Y. L., & Ferrell, J. M. (2020). Bile acid receptors FXR and TGR5 signaling in fatty liver diseases and therapy. *American Journal of Physiology—Gastrointestinal and Liver Physiology*, *318*(3), G554–G573.
- Cyr, M. C., Beauchesne, M. F., Lemiere, C., & Blais, L. (2008). Effect of theophylline on the rate of moderate to severe exacerbations among patients with chronic obstructive pulmonary disease. *British Journal of Clinical Pharmacology*, *65*(1), 40–50.
- Dalile, B., Van Oudenhove, L., Vervliet, B., & Verbeke, K. (2019). The role of short-chain fatty acids in microbiota-gut-brain communication. *Nature Reviews Gastroenterology and Hepatology*, *16*(8), 461–478.
- Damman, C. J. (2013). Salicylates and the microbiota: A new mechanistic understanding of an ancient drug's role in dermatological and gastrointestinal disease. *Drug Development Research*, *74*(6), 344–352.
- de Vos, W. M., & de Vos, E. A. (2012). Role of the intestinal microbiome in health and disease: From correlation to causation. *Nutrition Reviews*, *70*(Suppl 1), S45–S56.
- Deda, O., Gika, H. G., Wilson, I. D., & Theodoridis, G. A. (2015). An overview of fecal sample preparation for global metabolic profiling. *Journal of Pharmaceutical and Biomedical Analysis*, *113*, 137–150.
- Durbán, A., Abellán, J. J., Jiménez-Hernández, N., Ponce, M., Ponce, J., Sala, T., et al. (2011). Assessing gut microbial diversity from feces and rectal mucosa. *Microbial Ecology*, *61*(1), 123–133.
- Falony, G., Vieira-Silva, S., & Raes, J. (2018). Richness and ecosystem development across faecal snapshots of the gut microbiota. *Nature Microbiology*, *3*(5), 526–528.
- Fernández-Veledo, S., & Vendrell, J. (2019). Gut microbiota-derived succinate: Friend or foe in human metabolic diseases? *Reviews in Endocrine and Metabolic Disorders*, *20*(4), 439–447.
- Fritz, A., Hofmann, P., Majda, S., Dahms, E., Droge, J., Fiedler, J., et al. (2019). CAMISIM: Simulating metagenomes and microbial communities. *Microbiome*, *7*(1), 17.
- Gratton, J., Phetcharaburanin, J., Mullish, B. H., Williams, H. R., Thursz, M., Nicholson, J. K., et al. (2016). Optimized sample handling strategy for metabolic profiling of human feces. *Analytical Chemistry*, *88*(9), 4661–4668.
- Halldin, E., Eriksen, A. K., Brunius, C., da Silva, A. B., Bronze, M., Hanhineva, K., et al. (2019). Factors explaining interpersonal

- variation in plasma enterolactone concentrations in humans. *Molecular Nutrition and Food Research*, 63(16), e1801159.
- Hällgren, R. (2003). *Apparatus for intestinal sampling and use thereof*. U.S. Patent No. US6576429B1. U.S. Patent and Trademark Office. <https://patents.google.com/patent/US6576429B1/en>.
- Hardwick, S. A., Chen, W. Y., Wong, T., Kanakamedala, B. S., Devesson, I. W., Ongley, S. E., et al. (2018). Synthetic microbe communities provide internal reference standards for metagenome sequencing and analysis. *Nature Communications*, 9(1), 3096.
- Hertel, J., Harms, A. C., Heinken, A., Baldini, F., Thinnies, C. C., Glaab, E., et al. (2019). Integrated analyses of microbiome and longitudinal metabolome data reveal microbial–host interactions on sulfur metabolism in Parkinson’s disease. *Cell Reports*, 29(7), 1767–1777.
- Hoving, L. R., Heijink, M., van Harmelen, V., van Dijk, K. W., & Giera, M. (2018). GC-MS analysis of short-chain fatty acids in feces, cecum content, and blood samples. *Methods in Molecular Biology*, 1730, 247–256.
- Hullar, M. A. J., & Fu, B. C. (2014). Diet, the gut microbiome, and epigenetics. *Cancer Journal (Sudbury, Mass.)*, 20(3), 170–175.
- Ikeda, N., Saito, Y., Shimizu, J., Ochi, A., Mizutani, J., & Watabe, J. (1994). Variations in concentrations of bacterial metabolites, enzyme activities, moisture, pH and bacterial composition between and within individuals in faeces of seven healthy adults. *Journal of Applied Bacteriology*, 77(2), 185–194.
- Jewison, T., Su, Y., Disfany, F. M., Liang, Y., Knox, C., Maciejewski, A., et al. (2014). SMPDB 2.0: Big improvements to the small molecule pathway database. *Nucleic Acids Research*, 42(Database Issue), D478–D484.
- Karu, N., Deng, L., Slae, M., Guo, A. C., Sajed, T., Huynh, H., et al. (2018). A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. *Analytica Chimica Acta*, 1030, 1–24.
- Knight, R., Vrbanc, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., et al. (2018). Best practices for analysing microbiomes. *Nature Reviews Microbiology*, 16(7), 410–422.
- Le Gall, G., Guttula, K., Kellingray, L., Tett, A. J., Ten Hoopen, R., Kemsley, E. K., et al. (2018). Metabolite quantification of faecal extracts from colorectal cancer patients and healthy controls. *Oncotarget*, 9(70), 33278–33289.
- Li, L., Abou-Samra, E., Ning, Z., Zhang, X., Mayne, J., Wang, J., et al. (2019). An in vitro model maintaining taxon-specific functional activities of the gut microbiome. *Nature Communications*, 10(1), 4146.
- Li, Y., Xie, Z., Gao, T., Li, L., Chen, Y., Xiao, D., et al. (2019). A holistic view of gallic acid-induced attenuation in colitis based on microbiome-metabolomics analysis. *Food and Function*, 10(7), 4046–4061.
- Liang, Y., Dong, T., Chen, M., He, L., Wang, T., Liu, X., et al. (2020). Systematic analysis of impact of sampling regions and storage methods on fecal gut microbiome and metabolome profiles. *mSphere*, 5(1), e00763-e1719.
- Marchesi, J. R., Holmes, E., Khan, F., Kochhar, S., Scanlan, P., Shanahan, F., et al. (2007). Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *Journal of Proteome Research*, 6(2), 546–551.
- National Institute of Standards and Technology. (2020a). Standard Reference Materials. Retrieved April 1, 2020, from <https://www.nist.gov/srm>.
- National Institute of Standards and Technology. (2020b). Standard Reference Materials SRM Order Request System. Retrieved April 1, 2020, from <https://www-s.nist.gov/srmors/>.
- Nishimuta, M., Inoue, N., Kodama, N., Morikuni, E., Yoshioka, Y. H., Matsuzaki, N., et al. (2006). Moisture and mineral content of human feces—high fecal moisture is associated with increased sodium and decreased potassium content. *Journal of Nutritional Science and Vitaminology*, 52(2), 121–126.
- Nordstrom, A., Want, E., Northen, T., Lehtio, J., & Siuzdak, G. (2008). Multiple ionization mass spectrometry strategy used to reveal the complexity of metabolomics. *Analytical Chemistry*, 80(2), 421–429.
- Obeid, R., Geisel, J., & Nix, W. A. (2019). 4-pyridoxic acid/pyridoxine ratio in patients with type 2 diabetes is related to global cardiovascular risk scores. *Diagnostics (Basel)*, 9(1), 28.
- Primec, M., Mičetić-Turk, D., & Langerholc, T. (2017). Analysis of short-chain fatty acids in human feces: A scoping review. *Analytical Biochemistry*, 526, 9–21.
- Rangel, I., Sundin, J., Fuentes, S., Reptsilber, D., de Vos, W. M., & Brummer, R. J. (2015). The relationship between faecal-associated and mucosal-associated microbiota in irritable bowel syndrome patients and healthy subjects. *Alimentary Pharmacology and Therapeutics*, 42(10), 1211–1221.
- Rasberry, S. D. (2003). *Standard Reference Materials--the First Century. No. Special Publication (NIST SP)-260-150*. Gaithersburg, MD: National Institute of Standards and Technology.
- Reyes, R. E., González, C. R., Jiménez, R. C., Herrera, M. O., & Andrade, A. A. (2012). Mechanisms of O-antigen structural variation of bacterial lipopolysaccharide (LPS). In *The complex world of polysaccharides* (pp. 71–98). West Palm Beach, FL: InTech.
- Riesenfeld, C. S., Schloss, P. D., & Handelsman, J. (2004). Metagenomics: Genomic analysis of microbial communities. *Annual Review of Genetics*, 38, 525–552.
- Saric, J., Wang, Y., Li, J., Coen, M., Utzinger, J., Marchesi, J. R., et al. (2008). Species variation in the fecal metabolome gives insight into differential gastrointestinal function. *Journal of Proteome Research*, 7(1), 352–360.
- Schon, M., Mousa, A., Berk, M., Chia, W. L., Ukropec, J., Majid, A., et al. (2019). The potential of carnosine in brain-related disorders: A comprehensive review of current evidence. *Nutrients*, 11(6), 1196.
- Sczyrba, A., Hofmann, P., Belmann, P., Koslicki, D., Janssen, S., Droge, J., et al. (2017). Critical assessment of metagenome interpretation—a benchmark of metagenomics software. *Nature Methods*, 14(11), 1063–1071.
- Shah, R. M., McKenzie, E. J., Rosin, M. T., Jadhav, S. R., Gondalia, S. V., Rosendale, D., et al. (2020). An integrated multi-disciplinary perspective for addressing challenges of the human gut microbiome. *Metabolites*, 10(3), 94.
- Shao, T., He, Z., Xie, Z., Li, H., Wang, M., & Wen, C. (2016). Characterization of ankylosing spondylitis and rheumatoid arthritis using 1H NMR-based metabolomics of human fecal extracts. *Metabolomics*, 12(4), 70.
- Shao, T., Shao, L., Li, H., Xie, Z., He, Z., & Wen, C. (2017). Combined signature of the fecal microbiome and metabolome in patients with gout. *Frontiers in Microbiology*, 8, 268.
- Shaw, W. (2010). Increased urinary excretion of a 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA), an abnormal phenylalanine metabolite of *Clostridia* spp. in the gastrointestinal tract, in urine samples from patients with autism and schizophrenia. *Nutritional Neuroscience*, 13(3), 135–143.
- Shreiner, A. B., Kao, J. Y., & Young, V. B. (2015). The gut microbiome in health and in disease. *Current Opinion in Gastroenterology*, 31(1), 69–75.
- Simon-Manso, Y., Lowenthal, M. S., Kilpatrick, L. E., Sampson, M. L., Telu, K. H., Rudnick, P. A., et al. (2013). Metabolite profiling of a NIST Standard Reference Material for human plasma (SRM 1950): GC-MS, LC-MS, NMR, and clinical laboratory analyses, libraries, and web-based resources. *Analytical Chemistry*, 85(24), 11725–11731.

- Sinaiko, R. J. (2004). *Externally controlled intestinal content sampler*. U.S. Patent No. US5316015A. U.S. Patent and Trademark Office. <https://patents.google.com/patent/US5316015A/en>.
- Stanley, D., Geier, M. S., Chen, H., Hughes, R. J., & Moore, R. J. (2015). Comparison of fecal and cecal microbiotas reveals qualitative similarities but quantitative differences. *BMC Microbiology*, *15*(1), 51.
- Strandwitz, P. (2018). Neurotransmitter modulation by the gut microbiota. *Brain Research*, *1693*(Pt B), 128–133.
- Tanaka, H., Sirich, T. L., Plummer, N. S., Weaver, D. S., & Meyer, T. W. (2015). An enlarged profile of uremic solutes. *PLoS ONE*, *10*(8), e0135657.
- Topping, D. L., & Clifton, P. M. (2001). Short-chain fatty acids and human colonic function: Roles of resistant starch and nonstarch polysaccharides. *Physiological Reviews*, *81*(3), 1031–1064.
- Verbeke, K. A., Boobis, A. R., Chiadini, A., Edwards, C. A., Franck, A., Kleerebezem, M., et al. (2015). Towards microbial fermentation metabolites as markers for health benefits of prebiotics. *Nature Research Reviews*, *28*(1), 42–66.
- Wang, Z., Koonen, D., Hofker, M., & Fu, J. (2016). Gut microbiome and lipid metabolism: From associations to mechanisms. *Current Opinion in Lipidology*, *27*(3), 216–224.
- Wang, Z., Zolnik, C. P., Qiu, Y., Usyk, M., Wang, T., Strickler, H. D., et al. (2018). Comparison of fecal collection methods for microbiome and metabolomics studies. *Frontiers in Cellular and Infection Microbiology*, *8*, 301.
- Williams, B. B., Benschoten, V., Andrew, H., Cimermanic, P., Donia, M. S., Zimmermann, M., et al. (2014). Discovery and characterization of gut microbiota decarboxylases that can produce the neurotransmitter tryptamine. *Cell Host & Microbe*, *16*(4), 495–503.
- Wishart, D. S. (2019a). Metabolomics for investigating physiological and pathophysiological processes. *Physiological Reviews*, *99*(4), 1819–1875.
- Wishart, D. S. (2019b). NMR metabolomics: A look ahead. *Journal of Magnetic Resonance*, *306*, 155–161.
- Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vazquez-Fresno, R., et al. (2018). HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Research*, *46*(D1), D608–D617.
- Xu, R., & Wang, Q. (2016). Towards understanding brain-gut-microbiome connections in Alzheimer's disease. *BMC Systems Biology*, *10*(3), 63.
- Xu, R., Wang, Q., & Li, L. (2015). A genome-wide systems analysis reveals strong link between colorectal cancer and trimethylamine N-oxide (TMAO), a gut microbial metabolite of dietary meat and fat. *BMC Genomics*, *16*(7), S4.
- Zhang, S., Wang, H., & Zhu, M. J. (2019). A sensitive GC/MS detection method for analyzing microbial metabolites short chain fatty acids in fecal and serum samples. *Talanta*, *196*, 249–254.
- Zierer, J., Jackson, M. A., Kastenmüller, G., Mangino, M., Long, T., Telenti, A., et al. (2018). The fecal metabolome as a functional readout of the gut microbiome. *Nature Genetics*, *50*(6), 790–795.

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

Affiliations

Rupasri Mandal¹ · Raul Cano² · Cindy D. Davis³ · David Hayashi⁴ · Scott A. Jackson⁵ · Christina M. Jones⁵ · Johanna W. Lampe⁶ · Marie E. Latulippe⁷  · Nancy J. Lin⁵ · Katrice A. Lippa⁵ · Paulina Piotrowski⁵ · Sandra M. Da Silva⁵ · Kelly S. Swanson⁸ · David S. Wishart¹

¹ Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada

² The BioCollective, LLC, 5650 N Washington St, Denver, CO 80216, USA

³ Office of Dietary Supplements, National Institutes of Health, Bethesda, MD 20852, USA

⁴ Mondelez Global, LLC, Chicago, IL 60607, USA

⁵ Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

⁶ Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M4-B802, PO Box 19024, Seattle, WA 98109, USA

⁷ North American Branch of the International Life Sciences Institute (ILSI North America), 740 15th Street NW, Suite 600, Washington, DC 20005, USA

⁸ University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA