

## A wipe-based stool collection and preservation kit for microbiome community profiling

Hui Hua<sup>1,+,\*</sup>, Cem Meydan<sup>1,+</sup>, Evan E. Afshin<sup>1</sup>, Loukia Lili<sup>1</sup>, Christopher R. D'Adamo<sup>2</sup>, Joel Dudley<sup>1</sup>, Nathan D. Price<sup>1</sup>, Bodi Zhang<sup>1</sup>, Christopher E. Mason<sup>1,\*</sup>

### Affiliations:

<sup>1</sup>Thorne HealthTech, 152 W 57th, New York, NY 10019, USA

<sup>2</sup>Department of Family and Community Medicine, University of Maryland School of Medicine, 520 W. Lombard Street, Baltimore, MD 21201, USA

<sup>+</sup>These authors contributed equally to this work

<sup>\*</sup>To whom correspondence should be addressed

### Contact information:

Hui Hua <[hui.hua@oxfordalumni.org](mailto:hui.hua@oxfordalumni.org)>, Christopher E. Mason <[cmason@onegevity.com](mailto:cmason@onegevity.com)>

### Abstract

While a range of methods for stool collection exist, many require complicated, self-directed protocols and stool transfer. In this study, we introduce and validate a novel, wipe-based approach to fecal sample collection and stabilization for metagenomics analysis. A total of 72 samples were collected across four different preservation types: freezing at -20°C, room temperature storage, a commercial DNA preservation kit, and DESS (dimethyl sulfoxide, ethylenediaminetetraacetic acid, sodium chloride) solution. These samples were sequenced and analyzed for taxonomic abundance metrics, metabolic pathway classification, and diversity analysis. Overall, the DESS wipe results validated the use of a wipe-based capture method to collect stool samples for microbiome analysis, showing an  $R^2$  of 0.96 for species across all kingdoms, as well as exhibiting a maintenance of Shannon diversity (3.1-3.3) and species richness (151-159) compared to frozen samples. Moreover, DESS showed comparable performance to the commercially available preservation kit ( $R^2$  of 0.98), and samples consistently clustered by subject across each method. Future studies will be needed to further explore sample processing options and their applications in non-healthy subjects, particularly patients with irritable bowel syndrome, inflammatory bowel disease, and colorectal cancer, but these data suggest the DESS wipe method can be used for stable, room temperature collection and transport of human stool specimens.

**Keywords:** microbiome, wipe, metagenomics, DESS. stool collection

### Introduction

The detection and identification of biomolecules in microbial communities from samples is widely used for monitoring disease and overall health. Stable transportation and delivery of

biomolecules is generally required for such analysis. As such, low cost and efficient collection, storage, and delivery of biomolecules are critical for the field of medical diagnosis.

For human gut microbiome analysis, recent advances in sequencing techniques and bioinformatics have increased our knowledge of the complex microbial communities and their interactions. It is now well established that these microbes play important roles in relation to inflammation [1], metabolic disease [2, 3], mental disorders [4, 5], aging [6, 7] and several other diseases and health conditions [8-12]. However, different approaches of sample processing can introduce human error variability or technical biases through inappropriate sample handling or storage. For example, fecal microbiota sequencing profiles have been shown to change significantly during ambient temperature storage after 48 hours [13,14]. While performing nucleic acid extraction on fresh samples immediately after collection is impractical, freezing and storing samples at  $-20^{\circ}\text{C}$ ,  $-80^{\circ}\text{C}$ , or below, is widely considered to be best practice when preserving microbial composition for sequence-based analysis [15–17]. However, this is difficult to achieve in many situations, such as sampling in remote areas, and thus may dramatically increase the costs of such studies. While some studies have investigated in detail the rapid deterioration of fecal samples that have been stored at room temperature for several days prior to lab processing [13, 18–20], there are few methods to address such issues.

Moreover, stool specimen collection using most methods can be challenging, and many individuals find the process difficult and not user-friendly [21]. Challenges include embarrassment, fear of results, concerns around hygiene and contamination, discretion and privacy, and lack of information. A 3-year randomized trial of 997 participants found that discomfort with the collection of a stool sample is the most frequently cited barrier for participation in fecal test-based screening. Furthermore, the study found that having a choice of screening methods significantly increases (13% vs. 43%) patient adherence [22].

A 2016 study, spanning 15 individuals and over 1,200 samples, provided the most comprehensive view to date of storage and stabilization effects on stool [23]. It suggested that 95% ethanol can preserve samples sufficiently well at ambient temperatures for periods of up to 8 weeks, and include the types of variation often encountered under field conditions, such as freeze-thaw cycles and high temperature fluctuations [23]. In addition, a solution containing dimethyl sulfoxide, disodium EDTA, and saturated NaCl (DESS) was originally used for various applications in the preservation of nematodes for combined morphological and molecular analyses, has also been used to preserve entire soil/sediment samples, or as a storage medium for microbial community analysis [24-26]. Such preserved material can be easily stored for months at room temperature, shipped by mail, or carried in luggage, which provides an efficient, cost-effective method with widespread applications for microbiome studies.

To address such technical errors and biases in sample collection methods, as well as to enhance the user experience of stool sample collection, we have designed a practical and user-friendly

fecal sample collection kit that includes a dissolvable wipe (e.g., ethanol-soluble film), in which the biological sample is dissolved in a DNA stabilizing solution (DESS). The film, solution, and biological sample are disposed of in a sealable container. Therefore, people can collect their fecal samples as easily as using toilet paper after defecation. The DNA stabilizing solution ensures that the microbiomic community structure is well-preserved during ambient temperature transportation and storage, and the microbiome DNA is extracted from the fixed microbes and used for further laboratory analysis when the samples arrive at the laboratory. In addition, the dissolvable characters of the wipe ensure that all the microbes contact the stabilizing solution adequately when the sample is collected. The primary objective of this study is to assess the extent to which our novel approach to fecal sample collection and stabilization could maintain microbiota composition relative to immediate freezing (-20°C), preservation with a commercially available kit, and storage at room temperature (RT).

## Results

### *Study Summary*

**Figure 1** summarizes the overall study design. Six subjects were recruited to participate in the study to validate the wipe capture method. Two males and four females enrolled in the study. The median age was 42 years old (range: 31-60 years). Three subjects were white, non-Hispanic and three subjects were black, non-Hispanic. The average BMI across the cohort was 27.8 kg/m<sup>2</sup> (range: 20.3-40.7 kg/m<sup>2</sup>). Four preservation methods were used to process the samples for metagenomics sequencing: freezing (-20°C), room temperature storage, a commercial preservation kit (room temperature), and DESS DNA preservation (room temperature). Three replicates per subject for each preservative were collected for a total of 72 samples. A total of 71 samples were successfully sequenced with an average of 8 million sequencing reads and a range of 3 to 38 million reads. One sample was removed from the analysis due to sequencing failure (no library). The amount of DNA captured by wipe was found to be comparable to other collection and extraction methods. The average DNA yields from extraction were 98.2ng/uL for DESS, 44.6ng/uL for the commercial preservation kit, 286ng/uL for the -20°C samples, and 122.3ng/uL for the room temperature samples.

### *Sample Similarity*

A t-SNE analysis across sample types and subjects showed clear clustering by each individual across the cohort, wherein each subject was isolated and separated from one another based on their unique microbiome signature (**Figure 2**). Interestingly, wipe samples in the DESS clustered more closely with the frozen samples and the commercial preservation kit's samples. Meanwhile, in most subjects, the negative control samples that were stored at room temperature cluster together separately from the other preservation types. **Supplemental Figure 1** further demonstrates this finding in a dendrogram showing clustering by subject and divergence of room temperature samples compared to the other sample types.

### *Taxonomic Profiles*

Taxonomic assignment of reads to each domain of life were then examined for their relative distributions across the sample types. As expected, Bacteria was the predominant domain (>99% relative abundance) captured by the microbiome analysis across all subjects (**Figure 3A**). Subject 1 had some more hits to Archaea (<1%) than others, Subjects 4 and 6 had some samples with Eukaryota hits, and Subject 5 had some samples with viral hits (most <1%) (**Figure 3B**). Commensal gut flora including *Firmicutes*, *Bacteroidetes*, and *Actinobacteria*, were the top phyla across all samples, with some room temperature samples also having *Proteobacteria*, particularly in Subject 4. **Figures 3C** and **3D** highlight the correlation of the subject's microbiome profile across different domains comparing wipe DESS preparation to frozen and the commercial kit to frozen, respectively. There is an increased relative abundance of human DNA seen in wipe samples compared to the commercial kit (**Figure 3C**), however, this is still negligible compared to the predominance of reads matching to bacteria (**Figure 3A**).

### *Diversity Metrics*

The metagenomic data were then examined for two metrics of species diversity (**Figure 4**). The Shannon index metric showed a similar range (2.7-3.8) across all sample types, but the wipe in DESS (median 3.2) had more comparable levels to the commercial preservation method (3.1) and gold standard frozen samples (3.2), than the room temperature storage (2.9). The median species richness (151-159), however, was more comparable across all preservation techniques (**Figure 4B**).

### *Intra- and Inter-sample Comparisons*

Taxonomic profiles comparing the different preservation methods showed that DESS has a very strong correlation with the frozen samples and is comparable to the commercial kit (**Figure 5**). The Pearson correlation of taxa log abundance with intra-group and inter-group comparisons. DESS was found to be very similar to the -20°C frozen samples when considering replicate-to-replicate variability (positive-to-positive correlation = 0.92, positive-to-DESS correlation = 0.91) (**Figure 5A**). Furthermore, **Figures 5B** and **5C** highlight Pearson correlations calculated by median log<sub>10</sub> relative abundances and median HUMAnN functional pathway scores, respectively (**Supplemental Figures 2** and **3**). These analyses demonstrate strong intra-sample correlation across the wipe samples, as well as strong inter-sample correlation between the wipe and frozen samples. These correlations are even comparable to the correlation found between the commercial preservation and frozen samples (**Figures 5B, 5C**). **Figures 5D** and **5E** highlight each subject's taxonomic relative abundance comparing wipe to frozen and commercial preservation to frozen samples respectively. They further demonstrate that the wipe in DESS preservation method has a high positive correlation with the taxonomic profiles of the gold standard frozen preservation, and is comparable to the commercial preservation method.

## **Discussion**

This study demonstrates the use of a wipe-based capture method to collect stool samples for microbiome/metagenomics analysis. The DESS wipe preservation method showed comparable performance to a commercial DNA/RNA preservation kit, and was also very similar ( $R^2 > 0.96$ ) to the gold standard frozen samples for most metrics (i.e. taxonomic classification, diversity, functional pathway classification, and abundances). Both the DESS and the commercial preservation method showed significant diversity compared to the room temperature negative control.

Although the quality and significance of standard microbiome metrics are comparable across the wipe method, gold standard, and other commercial methods, further validation and better understanding of the bacterial to human DNA ratio in a broader population can be addressed in future studies. This will involve including non-healthy subjects such as samples from people with gut conditions (i.e. bloody diarrhea, IBS, blood in the stool, colorectal cancer, hemorrhoids, etc.) where human DNA is more present in the stool [27, 28], to further assess the performance of the wipe and integrity of microbiome analysis. Recruiting subjects with GI conditions such as constipation and diarrhea will further test the efficacy of the wipe. Moreover, subjects with different disease statuses and infections will be important to test, specifically patients with irritable bowel syndrome, IBD, *Clostridium difficile* infection, etc. Finally, RNA preservation and isolation for metatranscriptomics analysis poses its own set of unique challenges [29] and future studies will be needed to assess the wipe capture in DESS preservation for RNAseq analysis.

However, this study shows evidence of validation for a wipe-based collection and RT transport method for gut microbiome sampling and metagenomics sequencing analysis. Such a method may enable easier access to sampling, testing, and metagenomics implementation in clinical trials, home use, or even in remote environments, especially given the stability of the method. Indeed, wipe-based collection and processing offers a more user-friendly approach to collecting stool samples for microbiome analysis. Its ease-of-use design and simple instructions (just wipe and place into the tube) should enable easy integration with commercial stool collection kits and future biomedical studies and trials. Indeed, tools and methods such as these can be applied to help deploy metagenomics tools and methods for a wide range of both research and clinical applications.

## **Methods**

### *Study Design*

A total of 6 subjects were enrolled in this pilot study. The inclusion criteria to be enrolled in the study included: age  $\geq 18$ ; Bristol Stool Scale type 3 and 4 (normal), agree to collect and donate the feces, and the ability to understand and write English. Exclusion criteria included people with constipation, slightly dry, or diarrhea feces (Bristol Stool Scale types 1-2, 5-7), pregnant or breastfeeding females, history of alcohol, drug, or medication abuse, known allergies to any

substance in the study product, current diagnosis of inflammatory bowel disease (Crohn's Disease or Ulcerative Colitis), and currently taking any medication that may interfere with defecation.

### *Sample Collection and Processing*

Fecal collection kits were created and mailed to enrolled subjects with clear instructions on sample collection. A total of 12 samples were collected by each subject yielding a total of 72 samples to be processed. Four preservation methods were used to process the samples for metagenomics sequencing: freezing (-20°C), room temperature storage, Zymo DNA shield kit (room temperature), and DESS DNA preservation (room temperature). All the samples are shipped at room temperature except the samples meant for freezing which were shipped on dry ice. It took several days to up to a week for the shipment. When the samples arrived in the lab, all the samples were put into a fridge (4°C) except the frozen samples which were put in a -20°C freezer.

### *Microbiome Sequencing and Analysis*

DNA was extracted from all samples using QIAgen PowerSoil Pro Kit, libraries were prepared with Illumina Nextera FLEX, and samples were sequenced on the NextSeq500 platform. Samples were sequenced as paired-end 150bp for a mean depth of 8.0 million reads per sample (min: 2.9M, max:38.9M). Resulting sequences were trimmed by Trimmomatic [30], and then aligned to human genome reference using bwa [31]. Taxonomic annotation was performed by utilizing KrakenUniq [32] and subsequently Bracken [33] on a database that includes all bacterial, archaeal, viral, fungal references from RefSeq along with human reference. The lowest common ancestor taxonomic annotations were adjusted within the lineage until at least 10% of the unique k-mers belong to a specific clade and not its parent, then filtered for at least 10 reads and a minimum Bracken adjusted relative abundance of 0.005%. The pathway annotations were performed by using HUMAnN3 with the UniRef90 clusters [34].

### **IRB**

The study protocol was approved by the Institutional Review Board of the University of Maryland Baltimore (Protocol # HP-00087571). The study participants provided their written informed consent before enrolling in the study.

### **Author Contributions**

HH, CM, BZ, JTD, and CEM designed the study. HH, BZ, and CRD coordinated the IRB and subject recruitment. HH, CM, EA, LL, NDP, BZ, CEM were involved in data analysis, interpretation, and preparation of the manuscript. All authors approved and contributed to writing the manuscript.

### **Competing Interests**



All authors except for CRD are employees of Thorne HealthTech and the Onegevity Health platform.

## References

1. He Q, Li X, Liu C, Su L, Xia Z, Li X, Li Y, Li L, Yan T, Feng Q, et al. Dysbiosis of the fecal microbiota in the TNBS-induced Crohn's disease mouse model. *Appl Microbiol Biotechnol*. 2016;100(10):4485–94.
2. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature*. 2012;490(7418):55–60.
3. Tremaroli V, Backhed F. Functional interactions between the gut microbiota and host metabolism. *Nature*. 2012;489(7415):242–9.
4. Desbonnet L, Garrett L, Clarke G, Kiely B, Cryan JF, Dinan TG. Effects of the probiotic *Bifidobacterium infantis* in the maternal separation model of depression. *Neuroscience*. 2010;170(4):1179–88.
5. Mayer EA. Gut feelings: the emerging biology of gut-brain communication. *Nat Rev Neurosci*. 2011;12(8):453–66.
6. Wilmanski T, Diener C, Rappaport N, Patwardhan S, Wiedrick J, Lapidus J, Earls JC, Zimmer A, Glusman G, Robinson M, Yurkovich JT, Kado DM, Cauley JA, Zmuda J, Lane NE, Magis AT, Lovejoy JC, Hood L, Gibbons SM, Orwoll ES, Price ND. Gut microbiome pattern reflects healthy ageing and predicts survival in humans. *Nat Metab*. 2021 Feb;3(2):274-286.
7. Bana B, Cabreiro F. The Microbiome and Aging. *Annu Rev Genet*. 2019 Dec 3;53:239-261.
8. Chen H, Jiang W. Application of high-throughput sequencing in understanding human oral microbiome related with health and disease. *Front Microbiol*. 2014;5:508.
9. Stewart CJ, Embleton ND, Marrs ECL, Smith DP, Fofanova T, Nelson A, Skeath T, Perry JD, Petrosino JF, Berrington JE, et al. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. *Microbiome*. 2017;5(1):75.
10. Gruber J, Kennedy BK. Microbiome and longevity: gut microbes send signals to host mitochondria. *Cell*. 2017;169(7):1168–9.
11. Manor O, Dai CL, Kornilov SA, Smith B, Price ND, Lovejoy JC, Gibbons SM, Magis AT. Health and disease markers correlate with gut microbiome composition across thousands of people. *Nat Commun*. 2020 Oct 15;11(1):5206.
12. Diener C, Qin S, Zhou Y, Patwardhan S, Tang L, Lovejoy JC, Magis AT, Price ND, Hood L, Gibbons SM. Baseline Gut Metagenomic Functional Gene Signature Associated with Variable Weight Loss Responses following a Healthy Lifestyle Intervention in Humans. *mSystems*. 2021 Oct 26;6(5):e0096421.

13. Carroll IM, Ringel-Kulka T, Siddle JP, Klaenhammer TR, Ringel Y. Characterization of the fecal microbiota using high-throughput sequencing reveals a stable microbial community during storage. *PLoS One*. 2012;7(10):e46953.
14. Roesch LF et al. Influence of fecal sample storage on bacterial community diversity. *The open microbiology journal*. 2009; 40–46.
15. Wu GD et al. Sampling and pyrosequencing methods for characterizing bacterial communities in the human gut using 16S sequence tags. *BMC microbiology*. 2010; 10, 206.
16. Fouhy F et al. The effects of freezing on faecal microbiota as determined using MiSeq sequencing and culture-based investigations. *PloS one*. 2015; 10, e0119355.
17. Bahl MI, Bergstrom A & Licht TR Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis. *FEMS microbiology letters*. 2012; 329, 193–197.
18. Lauber CL, Zhou N, Gordon JI, Knight R, Fierer N. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. *FEMS Microbiol Lett*. 2010;307(1):80–6.
19. Guo Y, Li SH, Kuang YS, He JR, Lu JH, Luo BJ, Jiang FJ, Liu YZ, Papasian CJ, Xia H-M, et al. Effect of short-term room temperature storage on the microbial community in infant fecal samples. *Sci Rep*. 2016;6:26648.
20. Song SJ, Amir A, Metcalf JL, Amato KR, Xu ZZ, Humphrey G, Knight R, Dearing MD. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. *mSystems*. 2016;1(3):e00021–16.
21. Lecky DM, Hawking MKD, McNulty CAM. Patients’ perspectives on providing a stool sample to their GP: a qualitative study. *Br J Gen Pract*. 2014; 64(628): e684–e693.
22. Liang PS, Wheat CL, Abhat A, Brenner AT, Fagerlin A, Hayward RA, Thomas JP, Vijan S, Inadomi JM. Adherence to Competing Strategies for Colorectal Cancer Screening Over 3 Years. *Am J Gastroenterol*. 2016 Jan;111(1):105-14.
23. Song SJ, Amir A, Metcalf JL, Amato KR, Xu ZZ, Humphrey G, Knight R. Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies. *mSystems*. 2016 May 3;1(3).
24. Melissa Y, et al. DESS: a versatile solution for preserving morphology and extractable DNA of nematodes *Nematology*, 2006, Vol. 8(3), 367-376
25. Beknazarova M, Millsteed S, Robertson G, Whiley H, Ross K. Validation of DESS as a DNA Preservation Method for the Detection of Strongyloides spp. In *Canine Feces*. *Int J Environ Res Public Health*. 2017 Jun 9;14(6).
26. Lee KM, Adams M, Klassen JL. Evaluation of DESS as a storage medium for microbial community analysis. *PeerJ*. 2019 Feb 5;7:e6414.
27. Sidransky D, Tokino T, Hamilton SR, Kinzler KW, Levin B, Frost P, Vogelstein B. Identification of ras oncogene mutations in the stool of patients with curable colorectal tumors. *Science*. 1992; 256: 102-105.



28. Kisiel, John B, and David A Ahlquist. Stool DNA testing for cancer surveillance in inflammatory bowel disease: an early view. *Therapeutic advances in gastroenterology* vol. 6,5 (2013): 371-80.
29. Reck M, Tomasch J, Deng Z, Jarek M, Husemann P, Wagner-Döbler I; COMBACTE Consortium. Stool metatranscriptomics: A technical guideline for mRNA stabilisation and isolation. *BMC Genomics*. 2015 Jul 4;16(1):494.
30. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014 Aug 1;30(15):2114-20.
31. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009 Jul 15;25(14):1754-60.
32. Breitwieser FP, Baker DN, Salzberg SL. KrakenUniq: confident and fast metagenomics classification using unique k-mer counts. *Genome Biol*. 2018 Nov 16;19(1):198.
33. Lu J, Breitwieser FP, Thielen P, Salzberg SL. 2017. Bracken: estimating species abundance in metagenomics data. *PeerJ Computer Science* 3:e104
34. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, Mailyan A, Manghi P, Scholz M, Thomas AM, Valles-Colomer M, Weingart G, Zhang Y, Zolfo M, Huttenhower C, Franzosa EA, Segata N. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife*. 2021 May 4;10:e65088.

## Figure Legends

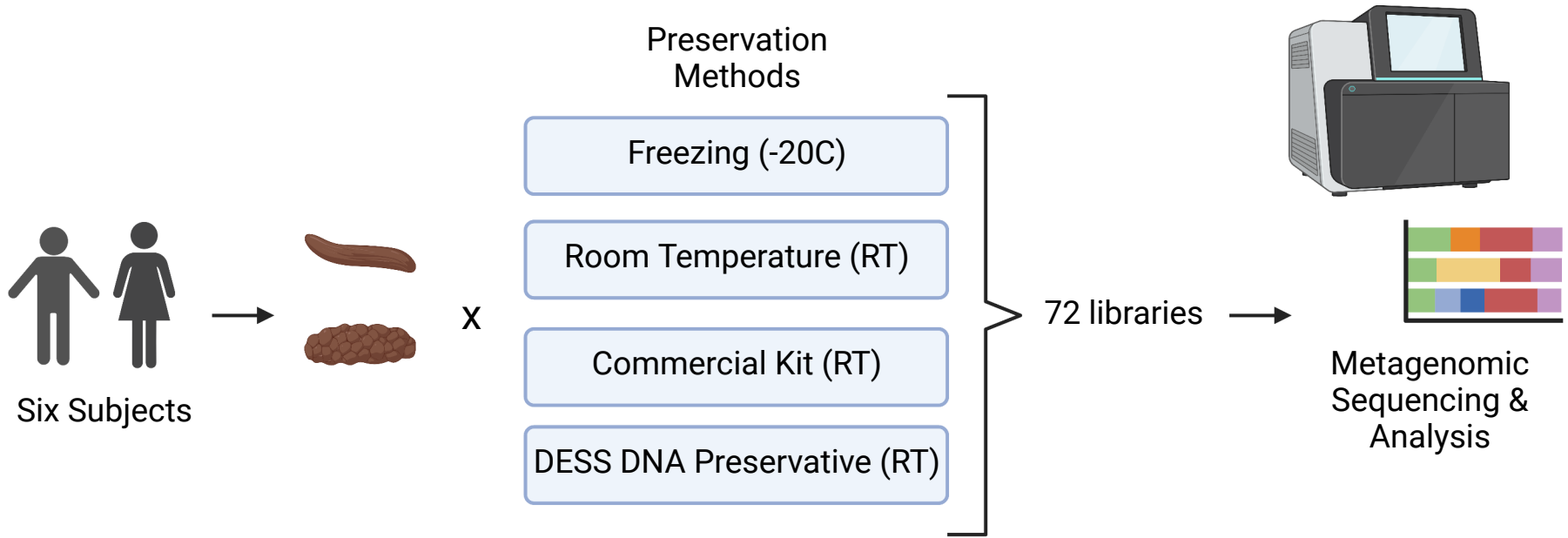
**Figure 1. Study Design.** Six participants were enrolled in the study and collected stool samples (Bristol Scale Type 3 and 4) for metagenomics/microbiome analysis. The samples were processed using four different preservation techniques: freezing at  $-20^{\circ}\text{C}$ , stored at room temperature (RT), Zymo DNA Shield (RT), and DESS (RT). A total of 72 samples were then sequenced with next-generation sequencing and analyzed for taxa and metabolic profiles. Created with BioRender.com.

**Figure 2. Sample Similarity.** A t-SNE plot displaying sample comparisons and clustering. Sample types are denoted by different colors and subjects by different shapes. Six distinct clusters are shown, one for each subject, and froze, wipe in DESS, and commercially processed samples cluster together while room temperature samples cluster separately from the other preservation types.

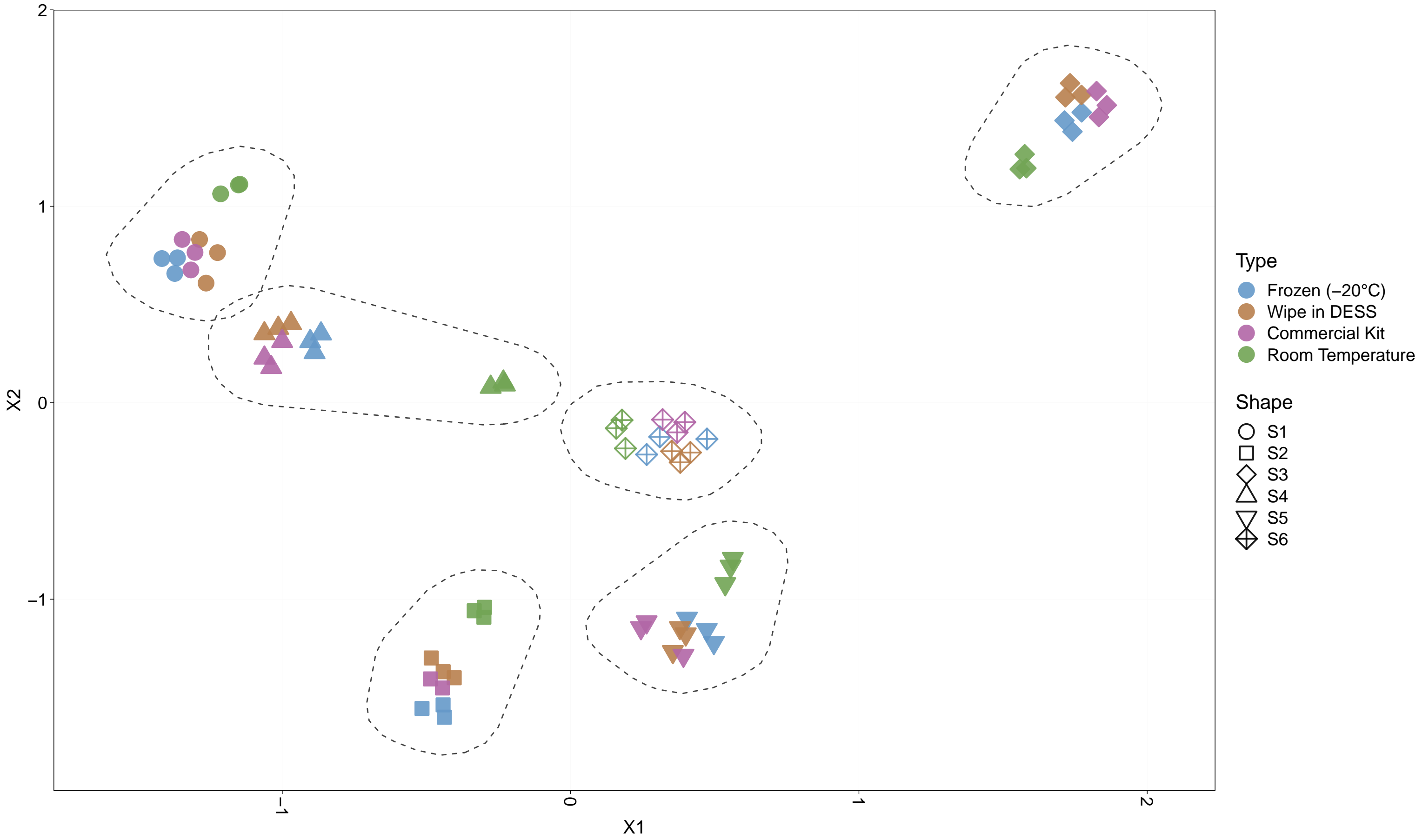
**Figure 3. Taxonomic Profiles.** Relative abundances of (A) Domains and (B) Phyla across the different subjects and sample types. Correlation plots comparing the relative abundances of wipe in DESS vs frozen samples (C) and commercial DNA preservation vs frozen samples (D).

**Figure 4. Diversity Metrics.** (A) Shannon index and (B) Species richness diversity metrics across sample types.

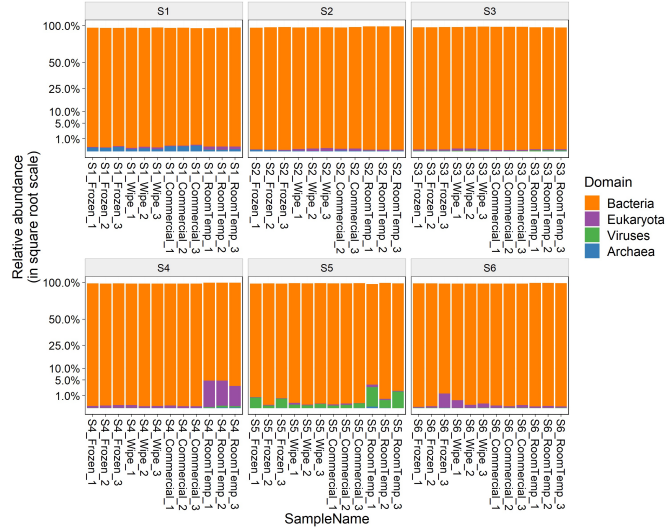
**Figure 5. Intra- and Inter-sample Sample Comparisons.** Pearson correlation by (A) Taxa log abundance with intra- and inter-group comparisons, (B) Median log<sub>10</sub> relative abundances, (C) Median HUMAnN pathway scores, and correlation plots comparing the relative abundances found in (D) Wipe vs Frozen and (E) Commercial preservation vs Frozen samples.



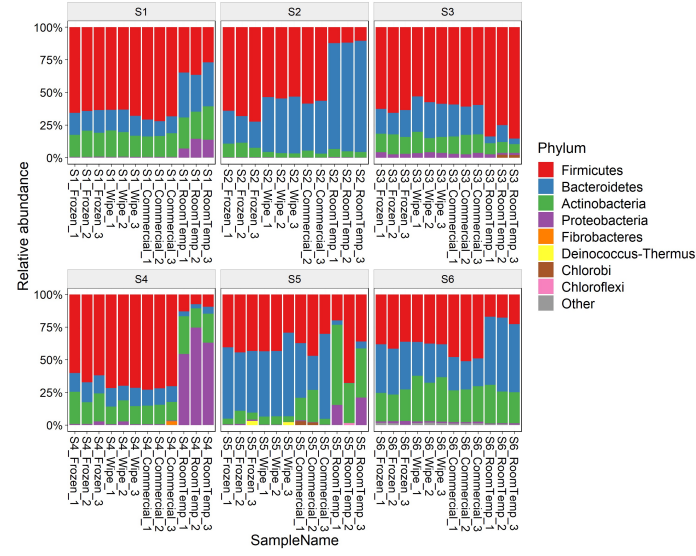
t-SNE of Wipe Comparison, LogAbBrackenFiltered, All, by Type, SD>0.5  
(perplexity=10)



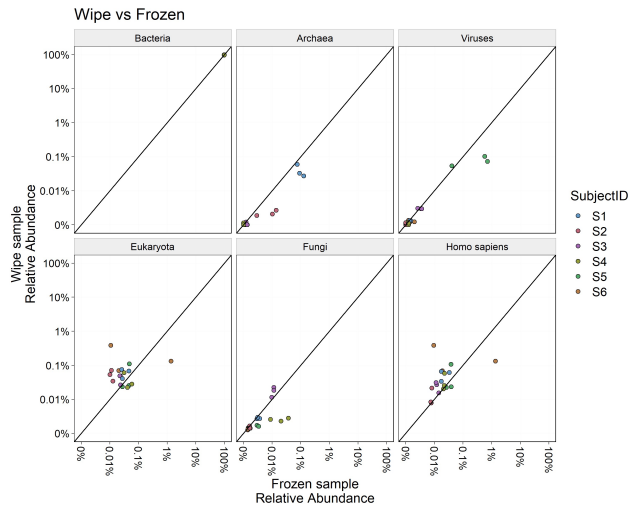
(A)



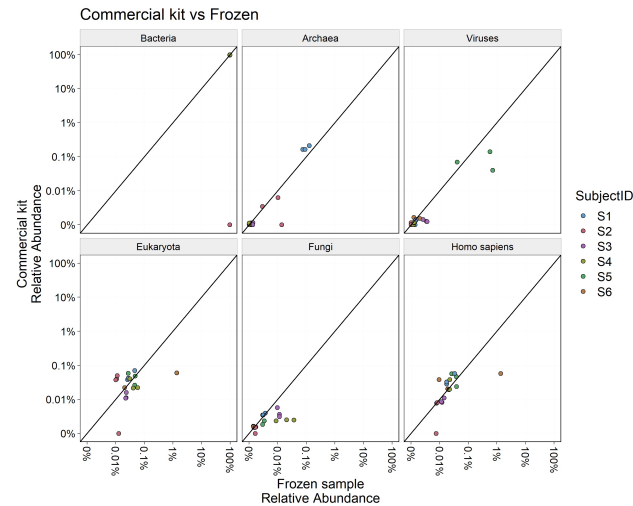
(B)



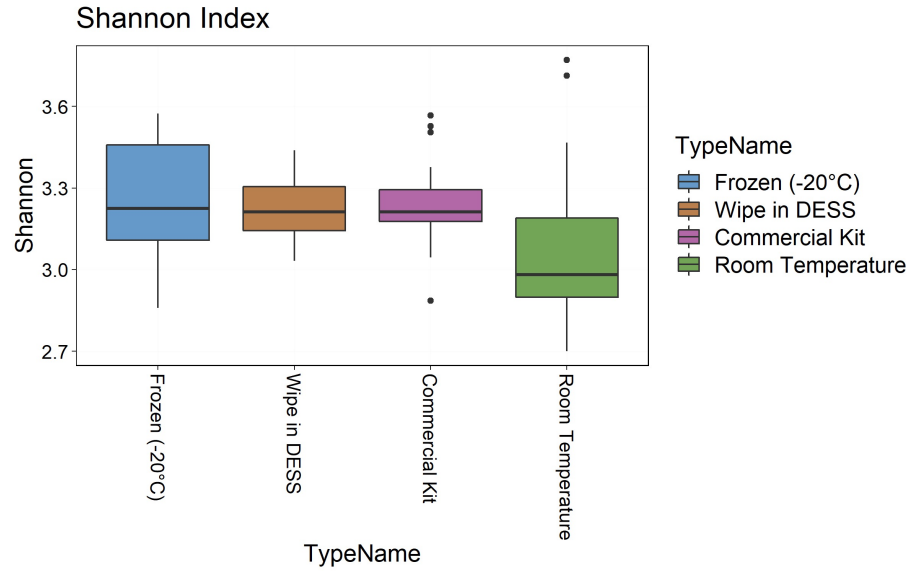
(C)



(D)



(A)



(B)

