RESOURCE ARTICLE OPEN ACCESS

Effects of Vacuum-Heat-Assisted Sample Desiccation on Microbiome Surveys

Stilianos Louca[1,2] 🔸 | Claire E. Mullin[1,2]

[1]Department of Biology, University of Oregon, Eugene, Oregon, USA | [2]Institute of Ecology and Evolution, University of Oregon, Eugene, Oregon, USA

Correspondence: Stilianos Louca (louca.research@gmail.com)

ABSTRACT

Sample preservation in the field and during transport can be a logistical challenge for microbiome surveys, particularly in remote areas. Sample desiccation eliminates the need for complicated cold chains and dangerous preservatives. However, the effects of desiccation on modern microbiome workflows such as gene-centric metagenomic profiling and metagenome-assembled genome (MAG) binning, remain poorly understood. In addition, most common desiccation tools such as lyophilisation cannot easily be deployed in the field. Here, we describe a proof-of-principle sample desiccator using vacuum and heat, specifically built for deployment in the field and exhibiting low power consumption and cost. We then test the effects of vacuum-heat-assisted sample desiccation followed by storage at room temperature, in comparison to conventional freezing, on multiple soil and animal faecal samples, via metagenomic and 16S rRNA amplicon sequencing. We consider multiple metrics related to the success of DNA extraction, sequencing, contig assembly, OTU clustering, gene annotation and MAG construction, as well as effects on inferred microbial community composition. We find that the impact of drying on considered success metrics was almost always either minor, non-significant or positive. For a subset of source materials we observed moderate but statistically significant differences in terms of inferred microbial taxonomic and genetic composition. We conclude that vacuum- and heat-assisted desiccation can be a useful, practical and cost-effective tool for microbiome field surveys, when a high consistency with frozen samples is not required.

1 | Introduction

The realisation that microorganisms strongly impact the functioning of nearly every ecosystem, ranging from the deep ocean and soils to the guts of animal hosts, has led to an intense interest in surveying microbial communities (Thompson et al. 2017). Modern techniques such as metabarcoding (marker gene amplicon sequencing) and metagenomics enable the detection and quantification of millions of microorganisms from a small amount of collected sample material such as wildlife faeces or a teaspoon of soil. However, sample preservation in the field and during transport remains a logistical challenge in microbiome surveys, particularly in remote areas or less developed countries. Cold chains, currently the golden standard of microbial sample preservation, are often impractical or outright impossible in the field. The most common alternative sample preservation methods include (a) preservation in ethanol or other fixatives and (b) desiccation (Hale et al. 2015; Mitchell and Takacs-Vesbach 2008;

Abbreviations: KO, KEGG Ortholog; MAG, metagenome-assembled genome; OTU, operational taxonomic unit.

Rissanen et al. 2010; Williams et al. 2016). Ethanol and many other fixatives are hazardous materials, which complicate sample storage and shipping, and fixatives tend to substantially increase weight (recommended added quantities typically range within 5–10 sample volumes). Further, fixatives can interfere with downstream molecular analyses and tend to only provide a short-term storage solution (a few weeks at most), as DNA degradation continues albeit at a slower rate.

Desiccation can be an effective short- and long-term sample preservation method at room temperatures for molecular applications, because in the absence of water most chemical reactions that would otherwise destroy DNA essentially come to a halt. Desiccation greatly facilitates transport in the field, shipping across borders and long-term archival of samples. Lyophilization (aka. freeze-drying), in particular, has long been used as a method for preserving viable bacterial cells over multi-year periods (Bolla et al. 2011; Heckly and Umbreit 1961) and has also been proposed as a means to preserve microbiome samples for sequencing (Bensch et al. 2022; Louca et al. 2016). Although lyophilization is a widely accepted method for biological sample preservation, it requires specialised equipment that is often not available even at institutions in developed countries and is impractical to carry on most expeditions. Approaches using desiccants (e.g., silica) at room temperatures are also problematic, because their slow action increases the risk of sample moulding and DNA degradation before desiccation (Murphy et al. 2000, 2002). Heat-based or vacuum-based desiccation, for example using food dehydrators or vacuum desiccators, is commonly used alternatives, but their effects on modern microbiome workflows, including gene-centric and genome-resolved metagenomics, are not well understood. Further, the combination of heating and vacuum in the field, which has the potential to reduce drying times and energy consumption, remains to our knowledge largely unexplored. Indeed, a reduction in pressure reduces the boiling temperature of water (Thomson 1946), thus decreasing the energy requirements for sample heating, and reciprocally, warming samples reduces the degree of vacuum that must be maintained.

Here, we examine vacuum- and heat-assisted desiccation as a practical technique for preserving samples collected in the field. Specifically, we present a proof-of-principle prototype desiccator using sample heating under vacuum, specifically built for use in expeditions, featuring sufficiently low power consumption for operation and transport in a vehicle. We use this desiccator to compare the effects of desiccation and long-term storage at room temperature relative to freezing on soil samples from three different locations and faecal samples from three different animals, at various stages of typical microbiome workflows. We emphasise that our objective is not to determine the separate effect of every possible factor such as storage duration, extraction protocol, storage temperature, bioinformatics approaches and so on. Instead, we focus on a few *typical* scenarios that environmental microbiologists can relate to in practice to test the feasibility of vacuum- and heat-assisted desiccation and to determine the extent to which it affects analysis outcomes compared to freezing and relative to other typical sources of error such as sample heterogeneities, stochasticity in PCR and sequencing. We also emphasise that we do not aim to determine whether freezing or drying is more accurate and instead focus on determining the

differences between the two. Workflows considered include 16S rRNA amplicon sequencing (aka. metabarcoding) (Thompson et al. 2017), gene-centric metagenomics for functional profiling (Tringe et al. 2005), and genome-resolved metagenomics for MAG construction (Parks et al. 2017). We consider 16 different metrics of 'success' such as DNA extraction yield, absorbance ratios (representing DNA purity) and peak fragment size, read counts, contig sizes and numbers of recovered operational taxonomic units (OTUs) from amplicons and MAGs from metagenomes. In addition, we examine whether sample desiccation affects inferred microbiome compositions relative to freezing, potentially due to taxonomic biases in either technique.

2 | Results and Discussion

2.1 | Sample Desiccation Using Vacuum and Heat

To explore the potential of vacuum- and heat-assisted sample desiccation as a practical tool in the field, we constructed a prototype desiccator using easily available, low-cost and lightweight materials (photo in Figure S1). We specifically wanted to explore the practical feasibility of such a device built for field work and determine its potential cost and effectiveness. The desiccator is essentially a modified vacuum degassing chamber with added thermostat-controlled heating blocks for sample heating and an added ceiling heater to restrict condensation of released moisture to the chamber's walls. Vacuum is generated using two serially connected small vacuum pumps. The desiccator can be powered by a 12 V car battery, weighs about 11 kg, and costs about $350 worth of materials to build. In this study, samples were kept at 40°C and < 0.01 bar during desiccation. Further technical details and a step-by-step demonstration are given in Data S1 and S2, respectively, and a schematic is given in Figure S2.

To test the effects of such a desiccation approach followed by sample preservation at room temperature on typical microbiome workflows relative to conventional freezing, we collected soil materials from 2 grasslands and 1 forest, as well as faecal materials from 3 different mammals (dog, cat, deer) in a municipal forest. An overview of these materials is given in Figure S3 and Table S1. From each material, we collected five samples to be desiccated and subsequently stored at room temperature, as well as five samples to be frozen at −80°C until further processing. Replication was necessary in order to account for inevitable random effects in sample collection, DNA extraction and sequencing. All samples were kept for 17 months until DNA extraction, followed by 16S rRNA gene amplicon sequencing as well as (shotgun) metagenomic sequencing.

2.2 | Effects on DNA Extraction

To examine the effects of sample desiccation relative to freezing on DNA extraction success, we extracted DNA from all samples using the same commercial kit, and measured DNA yield (ng), absorbance ratios at 260/280 nm and 260/230 nm as proxies for purity, as well as peak fragment size (bp). We then compared these four success metrics between the two treatments, whereas controlling for the source material. A visual overview of these

metrics, separately for each material and separately for dried and frozen samples, is shown in Figure 1. Although neither dried nor frozen samples achieved a better outcome in all cases, some general tendencies can be recognised. Notably, frozen samples generally yielded larger peak fragment sizes than their dried counterparts, regardless of material (Figure 1B), whereas for almost all materials dried samples yielded a higher average 260/280 ratio than their frozen counterparts (Figure 1C). The drop in peak fragment sizes under drying may be due to additional stress and shear experienced by DNA during desiccation, particularly due to the involved heating (Blake and Delcourt 1998). Results for DNA yields and 260/230 ratios were more ambiguous, with some materials yielding better outcomes when frozen and others yielding better outcomes when dried (Figure 1A,D).

To more quantitatively examine the effects of drying relative to freezing, we also performed statistical location tests for the above metrics using a null model that controls for source material. Specifically, for any given metric we computed its average value for each material and each treatment (dried vs. frozen), computed the ratio of these averages between treatments for each material (i.e., dried average divided by frozen average), then averaged those ratios across all materials, and finally compared that average ratio to that expected under a permutation null model. The null model assumes that for any given material both treatments are statistically equivalent, but allows differences between materials (see Section 2 for details, Table S2 for full results). Importantly, this model does not make any assumptions regarding the distribution of the data such as normality etc. We found moderately strong but significant differences between dried and frozen samples, albeit in mixed directions. Dried samples tended to yield significantly more DNA compared to frozen samples (on average 18% more, $p = 0.0041$, $n = 60$). These differences might be partly due to the higher biomass-to-water ratios of dried samples relative to their wet frozen counterparts, combined with the fact that the same start weight was used for all DNA extractions. Future follow-up studies could normalise extraction inputs by dry weight or include internal spike-ins to better distinguish true preservation effects from the effects of sample start quantity. Desiccation also tended to significantly reduce the peak fragment size (on average 20% smaller, $p = 0.0006$, $n = 60$). Although for short-read Illumina sequencing this reduction in peak fragment size is unlikely to be of concern (e.g., the lowest peak fragment size among all samples was still above 4.3 kbp, Figure 1B), it might impact long-read sequencing such as PacBio or Nanopore. Future studies should thus investigate the compatibility of desiccated samples with long-read sequencing technologies and resulting metagenomic assembly qualities. Although the observed decrease in DNA fragment size may not significantly affect short-read Illumina sequencing, future studies should investigate the compatibility of desiccated samples with long-read sequencing technologies such as PacBio or Nanopore, as well as their impact on metagenomic assembly quality.

A small but significant effect was also observed for 260/280 absorbance ratios, with a 1% higher value on average for dried samples ($p = 0.015$, $n = 60$), whereas no significant difference was observed for the 260/230 absorbance ratio. We point out that

260/280 and 260/230 absorbance ratios may not fully capture the effects of desiccation on DNA purity and quality, for example in terms of inhibitors potentially affecting molecular workflows. We thus also examined the downstream effects in two common molecular sequencing workflows, 16S metabarcoding and metagenomics, discussed in detail below.

2.3 | Effects on 16S rRNA Amplicon Analyses

To examine the impact of desiccation on typical 16S rRNA amplicon surveys (aka. metabarcoding), we performed 16S rRNA short amplicon paired-end Illumina sequencing (V4–V5 region) of DNA extracted from all samples. We then compared the number of reads, the average Phred quality score of forward and reverse reads, as well as the number of detected operational taxonomic units (OTUs, at 99% identity), between dried and frozen samples while controlling for source material. These metrics are shown separately for each material and each treatment in Figure 1E–H. Overall, dried and frozen samples yielded comparable outcomes, although some minor differences between treatments can be seen for some materials. Most notably, dried samples yielded on average substantially lower OTUs than their frozen counterparts for soil 1 and 2 (Figure 1H), although for other materials the numbers of OTUs were more similar between treatments. To more quantitatively examine the effects of drying relative to freezing, we also performed statistical location tests for the above metrics using the same null model described in the previous passage (detailed results in Table S2). We found no significant effect of drying compared to freezing in any of the considered metrics. Although on average dried samples yielded 13% fewer reads, 6% fewer OTUs, and 1% lower reverse read qualities compared to frozen samples, none of these differences were statistically significant ($p > 0.05$, $n = 60$).

To quantitatively examine how drying affects inferences of microbial community composition relative to freezing, we compared the proportions of OTUs between treatments, separately for each material, as follows. For any given material and any given OTU, we computed the OTU's average proportion across all 5 dried samples and separately across all 5 frozen samples, and then plotted these average proportions on two axes across all OTUs (Figure 2A–F). We measured the agreement between treatments in terms of the fraction of explained variance (R^2). Overall we observed a strong linear relationship between treatments for all materials, with the strongest agreement seen for faeces 1 and soil 3 ($R^2 = 0.87$) and the lowest agreement seen for faeces 2 ($R^2 = 0.42$), indicating that the two sample preservation methods are largely consistent in broad OTU abundance trends. In the theoretical ideal case of a perfect agreement, all OTUs should have the same average proportion in frozen as in dried samples, that is, all plotted points should be on the diagonal, and R^2 should be equal to 1. In reality, however, random effects in PCR and sequencing, as well as small microbial content differences between samples collected across a heterogeneous material, lead to variation in OTU proportions between samples, even within the same treatment. In other words, disagreements in average OTU proportions between treatments may or may not be due to the treatment itself. To resolve this ambiguity, we used a statistical test based on a permutation null model that
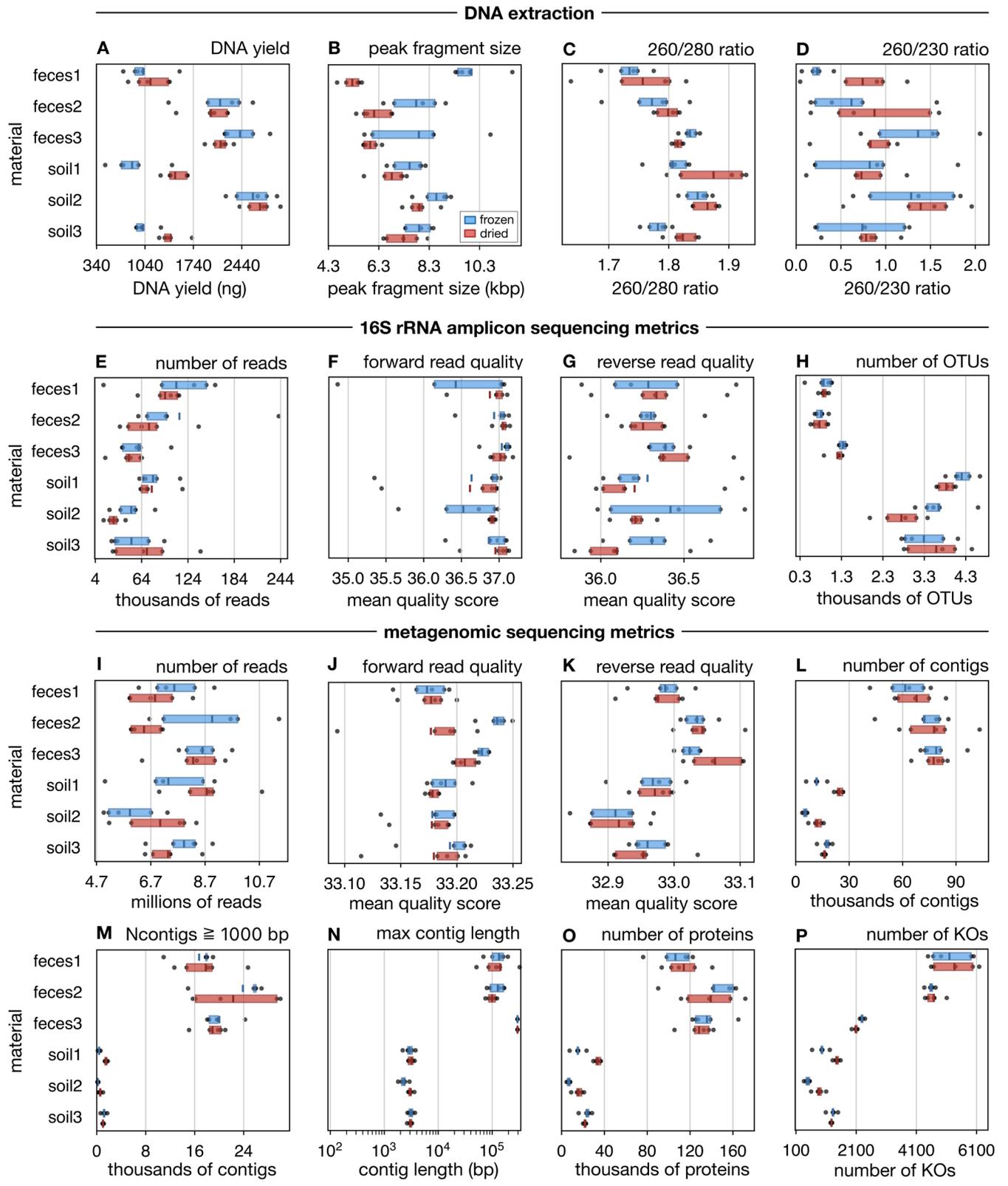
FIGURE 1 | Success metrics by material and treatment. (A) Box-plots of DNA extraction yields (ng) obtained for each sample (one point per sample), visually grouped by material (faeces 1–3, soils 1–3) and treatment (blue:Frozen, red:Dried). Boxes show interquartile ranges, small vertical bars show means. (B–P) Similar to A, but showing other metrics relevant to DNA extraction (top row), 16S rRNA amplicon sequencing (second row) and metagenomic sequencing, assembly and gene annotation (bottom two rows). For statistical tests quantifying the differences between dried and frozen samples see Table S2.
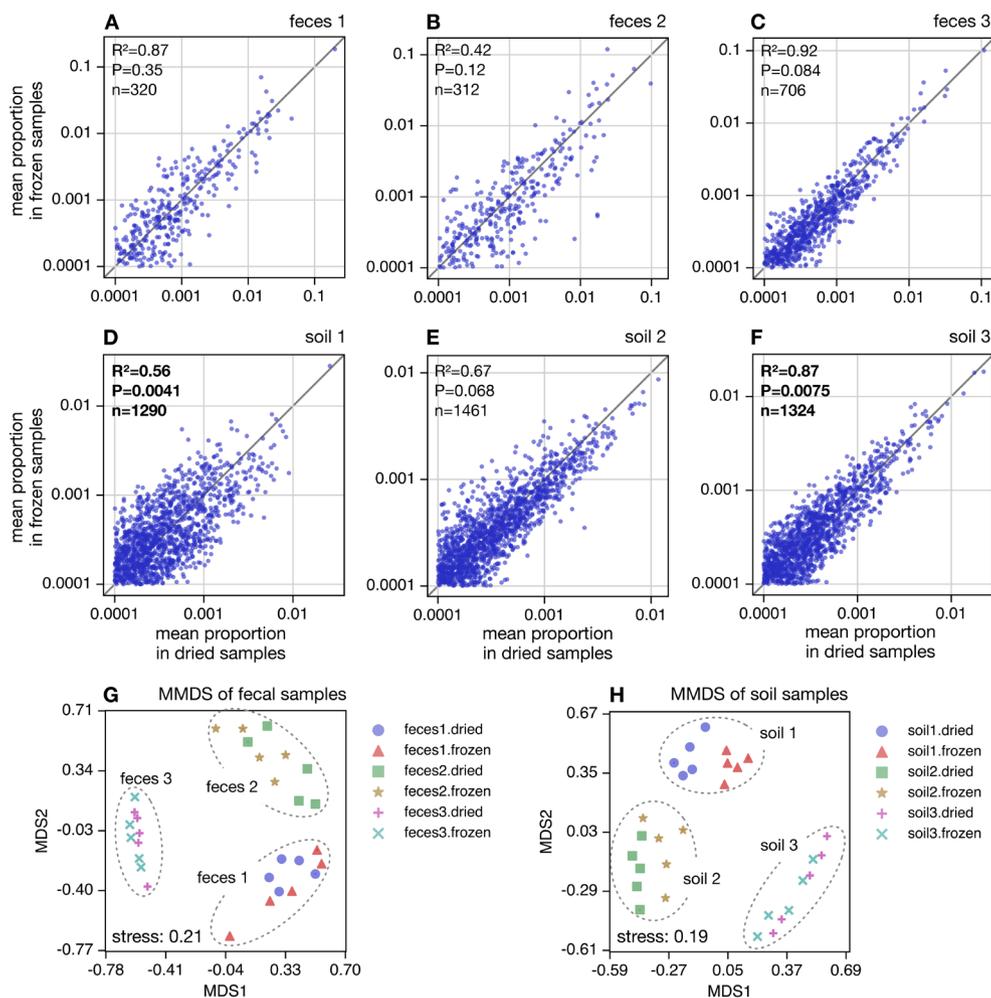
FIGURE 2 | OTU composition vs. treatment. (A) Mean OTU proportions in dried faecal 1 samples (horizontal axis) compared to mean OTU proportions in frozen faecal 1 samples (vertical axis, one point per OTU). Averaging of proportions was done among the 5 replicates in each treatment. The diagonal is shown for reference. Inscriptions show the fraction of variance in the vertical axis explained by the horizontal axis (R^2), the number of OTUs considered (n), and the statistical significance of R^2 compared to a permutation null model under which OTU proportions are statistically indistinguishable in the two treatments (P). A significantly low R^2 (i.e., $p < 0.05$) suggests dried samples tend to yield different OTU proportions compared to frozen samples. (B–F) Similar to A, but for the remaining samples. Statistically significant R^2 values are boldened. (G) Metric multidimensional scaling (MMDS) plot of weighted Bray-Curtis dissimilarities between faecal samples, based on OTU proportions. Points correspond to samples, and are shaped and coloured according to the material (faeces 1–3) and treatment (dried vs. frozen). The Kruskal stress is written in the plot. (H) Similar to G, but for soil samples. For similar plots using ASV or genus proportions see Figures S4 and S5, respectively. For MMDS plots based on rarefied data, or based on Jaccard dissimilarities, see Figure S7.

assumes no effects of treatment but controls for material. In this null model, samples from the same material are shuffled randomly regardless of treatment, average OTU proportions are subsequently re-computed for each treatment, and the R^2 is re-computed accordingly. The statistical significance of the observed R^2 was taken as the probability that the null model would generate a lower R^2 (i.e., lead to a stronger disagreement between treatments) simply by chance. We found that the R^2 was significantly low ($p < 0.05$) for 2 out of 6 materials (soil 1 and soil 3), whereas for all other materials R^2 was non-significant, that is, the disagreement between samples could plausibly be caused by chance (details in Figure 2A–F). Similar results were obtained at a greater taxonomic resolution of ASVs (roughly representing stains, Figure S4), where soil 1 and 3 were again the sole samples with statistically significant low R^2 (0.084 and 0.74, respectively). Comparable results were also obtained at the lower taxonomic resolution of

genera (Figure S5), where 3 out of 6 materials exhibited a statistically significant low R^2 (0.91 for faeces 3, 0.77 for soil 1 and 0.92 for soil 3). Such taxonomic differences may be caused, for example, by differential impacts of drying and freeze–thaw cycles on cell preservation and cell release from the material matrix, which in turn can impact the taxonomic distribution of extracted DNA within the relevant fragment size interval.

To assess whether differences in ASV proportions between treatments were due to relative changes among conspecific ASVs (i.e., changes in strain representations within a species) or due to changes at the species level or above, we also performed a modification of the above analysis. Specifically, for each OTU and for each sample, we computed the relative proportions of ASVs within that OTU, that is, such that the sum of ASV proportions in that OTU is 1. We then compared these within-OTU proportions of ASVs between treatments

5 of 13

in the same manner as described in the previous passage (Figure S6). We found no statistically significant low R^2 for any of the source materials, with R^2 ranging from 0.83 (faeces 1) to 0.94 (soil 1). This suggests that detection biases mostly occur at the species level or above, and not within a species, in other words most strains within a given species are affected similarly by such biases.

To evaluate the robustness of the above findings, we also considered pairwise abundance-weighted Bray-Curtis dissimilarities between samples based on OTU proportions (Legendre and Legendre 1998), and visualised those dissimilarities in the form of 2-dimensional metric multidimensional scaling (MMDS) plots (Borg and Groenen 2005) separately for faeces and soils (Figure 2G,H). Both for faeces and soils, the three materials are clearly discernible from each other; in other words, samples from each faecal material form a distinct cluster in the MMDS plot, and likewise for soil materials. To confirm the robustness of this observation, we repeated this analysis using abundance-weighted Jaccard dissimilarities, as well as using rarefied OTU tables in which each sample had the same number of reads (Figure S7). In all cases, we observed a clear and strong separation between materials. Thus, although differences exist between samples in each material, and while these differences may even be partly caused by the treatment, samples from different materials remained clearly distinguishable. In other words, the effects of desiccation—present or not—are much weaker than the differences between examined materials. When narrowing the scope of the MMDS plots to individual materials (Figure S8), for example only comparing dried and frozen samples from faeces 1, a separation between treatments becomes more apparent; however, this separation is generally comparable to the overall spread between samples within a given treatment. Thus, the effects of desiccation on inferred OTU composition seem comparable to those of other typical stochastic factors such as competition during PCR or randomness in the precise location sampled.

To more quantitatively assess if treatment has a detectable effect on pairwise dissimilarities, we used PERMANOVA tests (Anderson 2017), separately for each material and separately for each considered dissimilarity metric. We obtained mixed results depending on the material, the metric considered and the taxonomic level (OTU or genus), detecting a statistically significant separation by treatment in 23 out of 48 cases, mostly among soil materials (Tables S3 and S4). For example, when considering abundance-weighted Bray–Curtis dissimilarity, we found a significant separation by treatment for 1 out of 6 materials (soil 1) at the OTU as well as genus level. When considering abundance-weighted Jaccard dissimilarity we found a significant separation by treatment for 3 out of 6 materials (soils 1–3) at the OTU as well as genus level. Hence, desiccation appears to have had a detectable effect on inferred community composition in some materials (mostly soils), but this depended on the considered statistical test and taxonomic resolution. In fact, the fraction of variance in pairwise dissimilarities explained by treatment (R^2) ranged from 4.6% to 41% at the OTU level and from 3.5% to 58% at the genus level, and was almost always below 50% (Tables S3 and S4). Thus, residual factors other than treatment contributed substantially—typically more than treatment—to stochasticity in community composition.

2.4 | Effects on Metagenomics

To examine the impact of desiccation on metagenomic surveys, we performed short-read paired-end Illumina metagenomic sequencing. We then compared the number of reads, the average Phred quality score of forward and reverse reads, the number of assembled contiguous sequences (contigs), the number of contigs at least 1000 bp long, the maximum contig length, the number of predicted protein-coding sequences and the number of detected genes (KEGG Orthologs, or KOs), between dried and frozen samples while controlling for material. These metrics are shown separately for each material and each treatment in Figure 1I–P. In general, dried and frozen samples yielded comparable outcomes for these metrics. Some notable exceptions include faeces 3, where dried samples had substantially fewer reads and lower forward quality scores than frozen samples (Figure 1I,J), and soil 1, where dried samples yielded substantially more protein-coding sequences and KOs than their frozen counterparts (Figure 1O,P). To more quantitatively assess the effects of drying on these metrics, we performed statistical location tests using the same null model as described above (detailed results in Table S2). We found a minor but significant negative effect for forward read qualities (average dried/frozen ratio 0.99, $p = 0.013$, $n = 60$), and strong significant positive effects for the number of contigs (average ratio 1.4, $p < 0.001$, $n = 60$), the number of contigs at least 1000 bp long (average ratio 1.86, $p < 0.001$, $n = 60$), the number of protein-coding sequences (average ratio 1.41, $p < 0.001$, $n = 60$) and the number of detected KOs (average ratio 1.2, $p = 0.001$, $n = 60$). Hence, in nearly all statistically significant cases drying had a strong positive effect relative to freezing, and in all other cases the effects were either minor and/or insignificant. One reason for the increased number of long contigs in dried samples, and consequently the increased number of detected proteins and KOs, may be that desiccation somewhat lowers the observable species diversity (based on a lower OTU count, see previous section and Table S2), which generally facilitates long contig recovery.

To examine how drying affects inferences of the genetic composition of communities (gene-centric profiles), we compared the proportions of KOs between treatments, separately for each material (KEGG profiles shown in Figure S9). This comparison was done both visually (Figure 3A–F) as well as using a permutation null model, similarly to what we described earlier for OTU proportions. For all materials, we observed a strong linear relationship in average KO proportions between dried and frozen samples, with the fraction of explained variance (R^2) ranging from 0.78 for faeces 2 up to 0.98 for faeces 3. Hence, drying and freezing indicate similar broad trends in gene abundances. Nevertheless, for 5 out of 6 materials (i.e., except faeces 1) the R^2 was statistically significantly lower than expected by chance ($p < 0.05$, $n = 10$), indicating that treatment had a detectable impact on inferred genetic composition in these samples. Similar observation were made when KOs were grouped into higher-level KEGG-C groups (Figure S10).
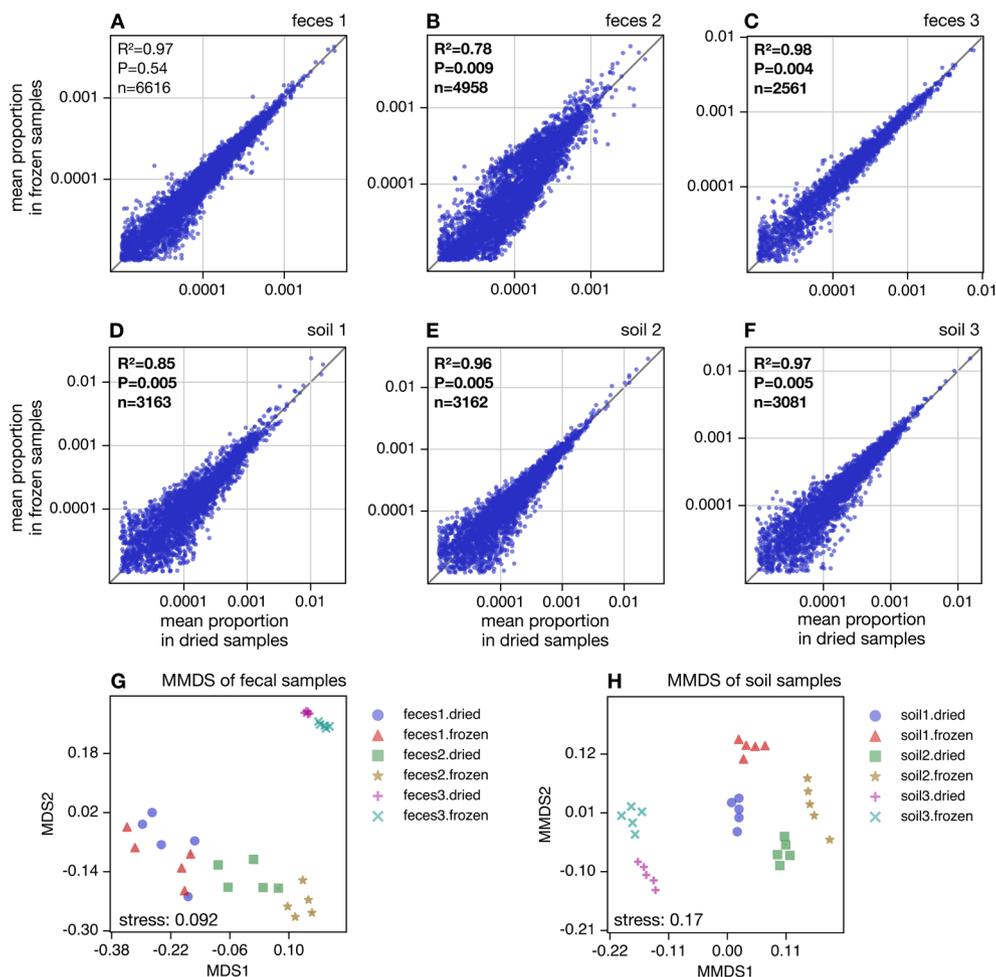
FIGURE 3 | Gene composition vs. treatment. (A) Mean gene (KO) proportions in dried faecal 1 samples (horizontal axis) compared to mean gene proportions in frozen faecal 1 samples (vertical axis, one point per gene). Averaging of proportions was done among the 5 replicates in each treatment. The diagonal is shown for reference. Inscriptions show the fraction of variance in the vertical axis explained by the horizontal axis (R^2), the number of genes considered (n), and the statistical significance of R^2 compared to a permutation null model under which gene proportions are statistically indistinguishable in the two treatments (P). A significantly low R^2 (i.e., $p < 0.05$) suggests dried samples tend to yield different gene proportions compared to frozen samples. (B–F) Similar to A, but for the remaining samples. Statistically significant R^2 values are boldened. (G) Metric multidimensional scaling plot of abundance-weighted Bray-Curtis dissimilarities between faecal samples, based on gene proportions. Points correspond to samples, and are shaped and coloured according to the material (faeces 1–3) and treatment (dried vs. frozen). The Kruskal stress is written in the plot. (H) Similar to G, but for soil samples. For similar plots using gene group proportions at KEGG level C see Figure S10.

To assess the robustness of the above findings, we also considered pairwise abundance-weighted Bray-Curtis dissimilarities between samples based on KO proportions (Legendre and Legendre 1998), and visualised those dissimilarities in the form of 2-dimensional metric multidimensional scaling (MMDS) plots (Figure 3G,H) separately for faeces and soils. Both for faeces and soils, the three materials are discernible from each other, in other words, samples from each material form a distinct cluster in the MMDS plot regardless of treatment, although the separation between materials is not as clear as what we observed for OTUs (Figure 2G,H). To more quantitatively compare dissimilarities between treatments to those within treatments, we performed PERMANOVA tests separately for each material (Tables S5 and S6), similar to what we described earlier for OTU dissimilarities. We found a statistically significant separation between treatments for 5 out of 6 materials (i.e., all except faeces 1, $p < 0.05$, $n = 10$ in each case), both when considering KO profiles (Table S5) or

when considering higher-level KEGG-C profiles (Table S6). Hence, drying has had a clearly detectable effect on gene-centric community profiles relative to freezing, for nearly all materials. This conclusion is further confirmed by statistical comparisons of individual KEGG C gene group abundances between treatments, which for nearly all materials (all except faeces 1) revealed several gene groups with statistically significant different proportions even after Bonferroni correction (Table S7, Data S4).

To assess the impact of drying on the ability to construct metagenome-assembled genomes (MAGs) (Sharon and Banfield 2013), we constructed a separate set of MAGs for each of the following subsets of metagenomes: dried faecal samples, frozen faecal samples, dried soil samples, and frozen soil samples. Splitting our samples in this way allowed us to examine the effects of drying relative to freezing while controlling for the broad material type (faeces vs. soil). Using multiple samples for

MAG construction (as opposed to constructing MAGs separately for each sample) allowed us to incorporate differential coverage profiles in the binning process, which is essential for obtaining high-quality MAGs (Albertsen et al. 2013; Chen et al. 2020). In total 302 MAGs were generated, with 281 and 21 obtained from faecal and soil samples, respectively (overview in Table S8 and Figure 4). The relatively low MAG yield from soil samples is not surprising, as the high microbial diversity found in soils generally makes contig assembly and MAG binning particularly challenging (Ma et al. 2023; Nayfach et al. 2021). For any given material type, the number, completeness and contamination level of MAGs was generally comparable between treatments, with no treatment yielding better results across the board. For example, dried faecal samples yielded slightly fewer MAGs than frozen faecal samples (135 vs. 146), but the situation was reversed for soil samples (14 vs. 7). The mean completeness of MAGs from dried faecal samples was slightly lower than from frozen faecal samples (55.5% vs. 56.9%), but the situation was again reversed for soil samples (35.0% vs. 26.4%). When considering only MAGs of at least medium quality, that is, with completeness $\geq 50\%$ and contamination $\leq 10\%$ (Bowers et al. 2017), we retain 61 MAGs from dried faeces and 62 MAGs from frozen faeces, whereas no MAGs are retained from soils.

2.5 | Limitations and Future Directions

Our results revealed that sample drying and freezing differ in how they impact the representation of taxonomic and functional groups in the generated sequences. Since both drying and freezing are expected to halt or at least dramatically slow down biological growth, these differences are unlikely to stem from actual microbial community changes, but instead are likely to stem from differential effects on DNA preservation during treatment, storage and DNA extraction. That is, some cell types may resist degradation during desiccation more than others, or may be preserved at room temperature after desiccation better than others, or may be easier to lyse in a desiccated state during DNA extraction than others. Given our data, it is impossible to dissect the contributions of each of these potential processes.

Further, comparative studies such as ours can only determine the effects of sample treatment on microbiome inferences relative to other treatments, but not in absolute terms. In our case, specifically, we determined how vacuum- and heat-assisted desiccation and storage at room temperature affect microbiome inferences relative to freezing, without clarifying whether desiccation or freezing yields the most accurate picture of the true microbial community composition. Indeed, the latter is hard to determine, given the many other biases involved, notably at the DNA extraction and PCR stages (Brooks et al. 2015) and due to variable 16S rRNA gene copy numbers (Louca et al. 2018). In fact, freezing and thawing itself can also introduce taxonomic biases (Cuthbertson et al. 2015; Poulsen et al. 2021).

Finally, we mention that apart from the specific examples examined here, we cannot extrapolate and provide general guidelines on required desiccation times, since these depend not just on initial water content but also on the specific material. Future multifactorial experiments might provide enough data to build a
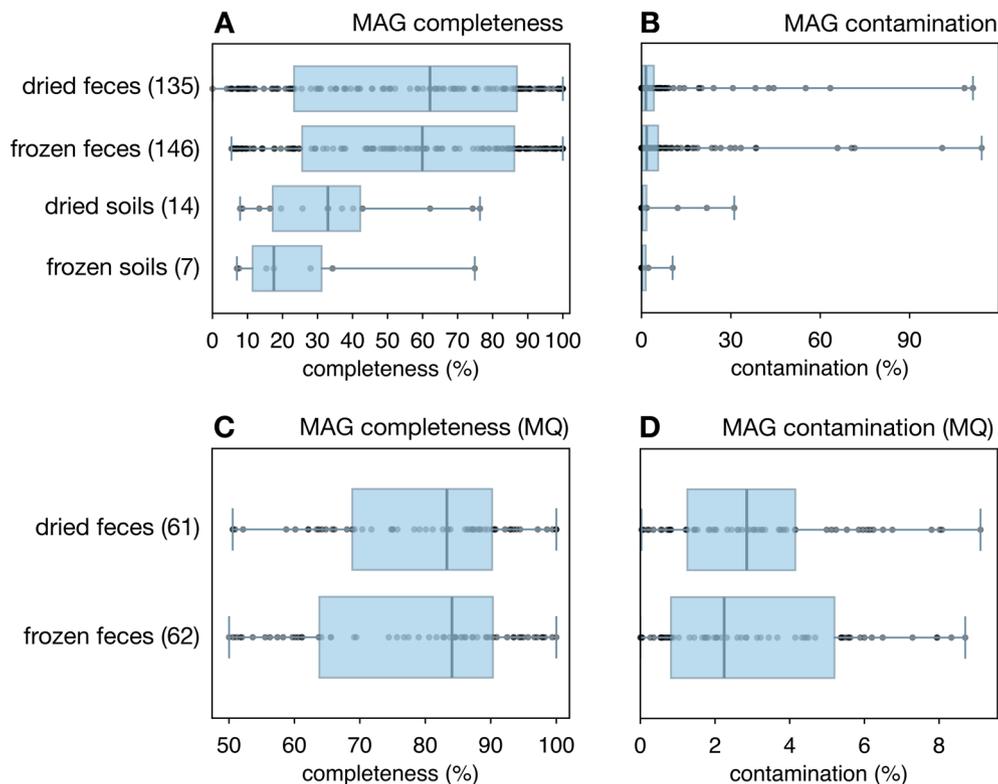
FIGURE 4 | MAG qualities. (A) Box-whisker plots of MAG completeness levels, separately for each combination of material type and treatment. The number of MAGs is written in the box labels. Boxes show interquartile ranges, whiskers show the full data range, points correspond to individual MAGs, and vertical bars correspond to medians. (B) Similar to A, but showing MAG contamination levels. (C, D) Similar to A, B, but focusing on MAGs of at least medium quality (completeness $\geq 50\%$, contamination $\leq 10\%$).

regression model for predicting desiccation times, but until then we think that trial-and-error and experience with their own specific types of samples will be the main approach for most researchers.

3 | Conclusions

We have explored sample desiccation using a combination of vacuum and heat, followed by storage at room temperature, as a practical tool for microbiome field surveys in remote regions. We systematically and quantitatively assessed the effects of sample desiccation on DNA extraction, 16S rRNA amplicon sequencing, and gene-centric and genome-resolved metagenomic sequencing, relative to conventional freezing at −80°C. Although our conclusions ultimately remain to be confirmed for other sample types, storage durations and temperatures, DNA extraction protocols, and so on, as well as different levels of microbial diversity, our work provides a demonstration of feasibility and a first assessment of possible impacts of vacuum- and heat-assisted desiccation in the field.

Overall, the observed effects of treatment on success metrics were statistically insignificant or moderate, and either favoured desiccation or freezing depending on the metric. The strongest significant effects were observed in DNA yield and peak DNA fragment size, where dried samples on average performed better and worse, respectively, as well as in the number of contigs, the number of contigs at least 1000 bp long, the number of detected protein-coding sequences, and the number of detected KOs, in all of which dried samples performed better on average.

For a subset of materials, we observed moderate but statistically significant differences in inferred microbiome compositions, both in terms of taxonomic group proportions and in terms of gene proportions, although samples from different source materials remained clearly distinguishable from each other. In many studies, especially those examining the effects of experimental treatments on microbial communities, it might be more important to maintain consistency in sample collection and processing than trying to infer the precise true composition of microbial communities or to maintain consistency with any particular preservation technique (e.g., freezing). Our results show that in cases where high consistency with freezing is not required, desiccation using vacuum and heat can be a practical sample preservation method in the field.

4 | Methods

4.1 | Sample Collection

Surface soil material was collected on 9 November 2022 from three different public locations in the Eugene area including from a municipal forest and two grasslands. Similarly, animal faecal materials were collected on 17 November 2022 around Spencer Butte municipal forest. Based on visual inspection as well as read mapping to reference genomes, faecal materials were identified as originating from dog, domestic cat and mule deer, respectively. An overview of material properties is given in Table S1, photos are provided in Figure S3. All materials were collected in sterile falcon tubes, and were kept at exterior temperatures in their original collection tube for about 1 day until further processing. Each material was split into 2×5 samples of roughly 500 mg each, of which 5 were frozen at −80°C until further processing. The remaining five samples were dried in our custom-built desiccator as described below, and subsequently stored in microcentrifuge tubes inside a carton box at 25°C until further processing. Note that the initial 1-day exposure to room temperature does not compromise our study's objectives, since it occurred prior to splitting, and thus its effects were shared by all samples from any given material. In fact, as all materials were collected from the natural environment, they had already been experiencing fluctuations in temperature, humidity and other environmental conditions, so one more day at room temperature merely extends that shared history. Replication was necessary in order to account for and quantify typical random effects encountered in the field including heterogeneities in sampled material. In fact, we intentionally did not homogenise the material prior to splitting, in order to fairly represent the inevitable stochasticity in typical field surveys stemming from local material heterogeneities, and to compare that stochasticity with the effects of our treatment (freezing vs. drying). Indeed, one of our goals was to determine how strongly sample preservation impacts scientific outcomes relative to other typical sources of error and bias. Freezing and drying was done on the same day, and thus this time point marks the beginning of divergent treatment between the two sets of samples. In total 60 samples were collected.

Desiccation was done under nearly completely vacuum (< 0.01 bar) at 40°C, using our custom-built desiccator (described in Data S1, photo in Figure S1, diagram in Figure S2). At this temperature, water boils once the pressure drops below 0.073 bar (Figure S11), which can easily be reached using the desiccator's vacuum pumps. Although a higher temperature could further shorten drying time, it would also increase the risk of DNA damage (Blake and Delcourt 1998). Soil and faecal samples were dried in two separate batches, that is, 30 soil samples were desiccated in one batch and 30 faecal samples desiccated in a separate batch. Samples were desiccated in 5 mL tubes without caps but protected with PTFE filters (pore size 0.22 μm) attached to the tube mouths using Parafilm. The weight loss of samples was periodically estimated using a precision scale and used to assess the progression of desiccation; desiccation was considered complete once the detectable rate of weight loss reached zero. The final fraction of weight lost was taken as an estimate of the material's original water content (% w/w). Soil pH was determined by mixing 20 g of residual soil into 50 mL of deionised water, mixing on a magnetic stirrer for 15 min, letting the slurry rest for another 15 min, then measuring using a West Tune digital pH probe, following common practice (Kirk et al. 2010; Liu et al. 2013). Total desiccation times, pH and water contents are listed in Table S1. In each desiccation batch, we included a negative control consisting of molecular-grade water in a 5 mL tube (covered with a filter similarly to samples), and placed centrally among the drying samples. Upon desiccation, the controls were rehydrated and vortexed vigorously, DNA was extracted from them as described below and quantified using a Qubit fluorometer and a high-sensitivity assay; none of the controls yielded any detectable DNA.

In April 2024, that is, about 17 months after sample collection and drying or freezing, all samples were visually inspected for

integrity and re-weighted to test for rehydration. No degradation or notable weight changes were observed, and DNA was then extracted from all samples. This duration was chosen as a representative of typical storage durations encountered in practice, where a considerable time might pass between sample collection and actual sequencing (1–2 years). DNA extractions were performed using the DNeasy PowerSoil Pro kit, following the manufacturer's protocol and always starting with 250 mg of material. Frozen and dried samples from the same material were always extracted on the same day, and all extractions were performed within a period of 1 week. DNA extraction yields were determined using a Qubit 4 fluorometer. Absorbance ratios (260/280 nm an 260/230 nm) were measured using an Implen Nanophotometer NP80. DNA fragment size distributions were measured using an Advanced Analytical fragment analyser. The fragment size at the maximum density, henceforth referred to as 'peak fragment size', is considered here as an extraction success metric. We did not quantitatively measure the spread (smear) of the distributions, however upon visual inspection we did not notice any obvious differences between treatments in that regard. The above metrics are visualised using standard box plots in Figure 1, which were generated using the python package `matplotlib`, functions `bxp` and `scatter`.

Shotgun metagenomic and 16S rRNA gene amplicon sequencing was performed for each of the 60 samples by the Integrated Microbiome Resource (IMR) in Dalhousie, Canada. Specifically, metagenomic libraries were prepared using the Illumina Nextera Flex kit and sequenced using a NextSeq2000 (2×150 bp paired ends). 16S rRNA gene amplicon fragments (V4–V5 region) were PCR-amplified using the Phusion Plus polymerase and 'universal' bacterial + archaeal primers (515FB = GTGYCAGCMGCCGCGGTAA, 926R = CCGYCAATTYMTTTRAGTTT; Parada et al. 2016; Walters et al. 2015), and sequenced on a MiSeq (2×300 bp paired ends).

4.2 | Statistical Tests of Success Metrics

To statistically examine the differences between dried and frozen samples in each considered success metric (e.g., DNA yield, read qualities etc.), whereas controlling for the source material, we used a permutation-based location test (Good 2013; Pesarin and Salmaso 2010), as follows. For any given metric, let $X_{m1}, \ldots, X_{m5}$ denote the metric's values among the 5 dried samples and $Y_{m1}, \ldots, Y_{m5}$ the metric's values among the 5 frozen samples, for the m-th material (faeces 1–3, soil 1–3). For each material m, we first computed the average value for the dried samples ($X_m := (X_{m1} + .. + X_{m5})/5$) and separately for the frozen samples ($Y_m := (Y_{m1} + .. + Y_{m5})/5$). We then compared these average values by forming their ratio for each material, $\rho_m := X_m / Y_m$. Hence, for example, a ratio $\rho_m > 1$ means that drying had an overall positive effect in the considered metric among samples of the m-th material. We then computed the average of these ratios, $\rho := (\rho_1 + .. + \rho_6)/6$ in order to quantify the overall effect of drying relative to freezing across all materials (Table S2). To assess the statistical significance of this ratio ρ, we compared it to ratios expected under a permutation null model that assumes no statistical differences between frozen and dried samples but still controls for the material. The null model randomly permutes the

samples within each material, thus breaking any associations between outcome and treatment but maintaining associations between outcome and material. We performed 10,000 such permutations, and for each permutation the X_m, Y_m, ρ_m and ρ were re-computed, thus yielding an empirical distribution of ρ under the null model. The statistical significance of the observed ρ was taken as the fraction of ρ generated by the null model that were at least as extreme as observed in either direction, that is, with an absolute distance from 1 at least as large as observed. One of the advantages of the above permutation approach, compared to for example conventional linear regression, is that we do not need to make any assumptions about the probability distribution of the data (Gaussian or not, constant variances or not; Good 2013). Further, it allowed us to conveniently implement a null model that directly controls for source material while assessing the statistical significance of differences between treatments.

4.3 | Analysis of 16S rRNA Gene Amplicons

On average 71,036 16S rRNA gene read pairs were obtained for each sample. Reads were quality-filtered and amplicon sequence variants (ASVs) were inferred and chimera-filtered using the R package `dada2` v1.28.0 (Callahan et al. 2016), as follows. Reads were quality-filtered using the `dada2` function `filterAndTrim`, with options 'truncLen = (250,200), maxEE=(1,1), truncQ=(0,0), trimLeft=(6,6), minLen = (100,100), maxLen = (100000,100000)', retaining on average 60,076 read pairs per sample. Error model calibration for ASV inference was performed jointly for all samples but separately for forward and reverse reads. Calibration was performed using the `dada2` function `learnErrors` with options 'nbases=1e8, randomize=TRUE, MAX_CONSIST=10, errorEstimationFunction=loessErrfun'. Reads were dereplicated using the `dada2` function `derepFastq`. ASVs were inferred from dereplicated sequences separately for forward and reverse reads, using the `dada2` function `dada` (options 'pool=TRUE, selfConsist=FALSE') and the previously calibrated error models. ASVs from forward and reverse reads were merged using the `dada2` function `mergePairs` with options 'minOverlap=12, maxMismatch=0, trimOverhang=TRUE'. Merged ASVs were chimera-filtered using the `dada2` function `removeBimeraDenovo` (option `method= 'consensus'`). ASVs were taxonomically classified based on a comparison to the SILVA database v138.1 (Glöckner et al. 2017), using a consensus approach (Soufi, Tran, and Louca 2024). Non-prokaryotic ASVs, including mitochondria and chloroplasts, were omitted from all subsequent analyses. Retained ASVs were then clustered de-novo into operational taxonomic units (OTUs) at 99% similarity, a commonly used species delineation threshold (Edgar 2018; Kim et al. 2014), using `vsearch --cluster_fast` with options '--iddef 2 --strand plus'. This yielded an OTU table with 17,680 OTUs representing 2,126,128 reads across 60 samples.

4.4 | Statistical Comparisons of Microbial Taxonomic Compositions

Pairwise dissimilarities between samples in terms of microbial taxonomic composition (OTU and genus level) were computed

using two alternative popular metrics, abundance-weighted Bray-Curtis and abundance-weighted Jaccard, both of which account for relative OTU/genus abundances (Chao et al. 2005; Legendre and Legendre 1998). We computed these metrics using the full data as well as using a rarefied subset where all samples were subsampled to the same number of reads. Metric multidimensional scaling (MMDS) embedding into 2 dimensions based on these pairwise dissimilarities was performed by minimising the Kruskal stress function (Borg and Groenen 2005) (Figure 2G,H, Figures S7 and S8). Principal Coordinates Analysis (PCoA) was also performed for comparison, shown in Figure S12 and confirming the separation between materials seen in the MMDS plots.

To test whether drying or freezing had any effects on inferred OTU or genus compositions, PERMANOVA tests were performed on these dissimilarities, using the pseudo F test statistic and estimating statistical significances via 10,000 permutations (Anderson 2017). A significantly high pseudo F statistic indicates that pairwise dissimilarities between treatments tend to be detectably larger than within each treatment (Tables S3 and S4).

To further visualise and quantify potential differences in OTU composition between treatments for any given material, we computed for each OTU its average proportion across all 5 dried samples (denoted X_i) and its average proportion across all 5 frozen samples (denoted Y_i), and then visually compared these average proportions across OTUs using the python function `pyplot.scatter` (Figure 2A–F). In addition, we computed for each material the fraction of variance in the Y_i explained by X_i, also known as coefficient of determination:

$$R^2 = 1 - \frac{\sum_i (Y_i - X_i)^2}{\sum_i (Y_i - \overline{Y})^2} \tag{1}$$

where i iterates over all OTUs and $\overline{Y}$ is the average of $Y_1, Y_2, \ldots$. A lower R^2 indicates a stronger disagreement between frozen and dried samples. The statistical significance of R^2 was estimated based on a permutation null model in which frozen and dried samples are statistically indistinguishable. In this null model, the treatments (frozen vs. dried) associated with samples were randomly permuted, thus breaking any statistical association between samples and treatment. For each permutation, average OTU proportions in each treatment and the corresponding R^2 were re-computed, and the statistical significance (p-value) was set to the fraction of all 10,000 permutations that yielded a lower R^2 than observed. Only OTUs with a mean proportion of at least 0.01% in both treatments were considered in this analysis, to avoid excessive noise associated with sampling stochasticity. A similar approach was used to also compare ASV proportions (Figure S4) and genus proportions (Figure S5) between treatments.

To assess whether differences in ASV proportions between treatments were due to relative changes among conspecific ASVs or due to changes at the species level, we proceeded as follows. For each OTU and for each sample, we computed the relative proportions of ASVs within that OTU, that is, such that

the sum of ASV proportions in that OTU is 1. Only ASVs also included in the previous analysis were considered (i.e., representing at least 0.01% of reads in a sample) to ensure comparability. We then compared these within-OTU proportions of ASVs between treatments in the same manner as described earlier for the within-sample proportions (Figure S6).

4.5 | Analysis of Metagenomes

On average 7,586,082 metagenomic read pairs were obtained per sample. Adapters were trimmed from reads using the tool skewer v0.2.2 (Jiang et al. 2014). Reads were then quality-filtered using vsearch v2.28.1 (Rognes et al. 2016) with options "–fastq_ascii 33 –fastq_maxee 1 –fastq_truncee 1 –fastq_qmax 64 –fastq_maxee_rate 0.005 –fastq_stripleft 0 –fastq_trunclen_keep 10000", retaining on average 5,997,909 high-quality read pairs per sample. Paired reads from each sample were assembled into longer contiguous sequences (contigs) using megahit v1.2.9 (Li et al. 2015) with option "–min-contig-len 500". On average 43,527 contigs were generated per sample. Gene-centric functional profiles were generated from assembled contigs similar to (Soufi, Porch, et al. 2024; Soufi, Tran, and Louca 2024). Here we thus only provide a brief summary. Contig coverages were computed for each sample by mapping the non-assembled reads to the contigs, then counting the number of reads mapped to each contig with a MAPQ score ≥ 30 and dividing that number by the contig length. Contig coverages were then normalised in each sample to sum 1, thus yielding contig "proportions". Protein-coding genes (PCGs) were predicted in the contigs using prodigal v2.6.3 with option "–p meta" and otherwise default options (Hyatt et al. 2010). PCGs were then mapped to KEGG gene orthologs (KOs) in the KOfam HMM database r105 (Aramaki et al. 2019) using hmmsearch v3.3 (Mistry et al. 2013). Only hits with an E-value below 10^{-10} were considered.

Gene-centric functional profiles were estimated for each sample as follows. Quality-filtered paired reads from multiple samples were coassembled separately for each combination of treatment and material type (i.e., dried faeces, frozen faeces, dried soils, frozen soils). Contig coverages in each sample were computed and PCGs were predicted as described earlier. Proportions of PCGs were computed in each sample by first associating with each PCG the proportion of its host contig, and then normalising those values in each sample to sum 1. The proportion of a given gene in a given sample was estimated by summing the proportions of all PCGs mapped to the gene. KOs were also grouped into standard high-level functional groups at KEGG level C; the proportion of each group in each sample was computed as the sum of proportions of all associated genes (Figure S9). The following KEGG groups were omitted, as they are not actually defined based on function: 'brite hierarchies', 'enzymes with ec numbers', 'not included in pathway or brite', 'poorly characterised', 'general function prediction only', 'others', 'unclassified viral proteins', 'function unknown'.

Metagenome-assembled genomes (MAGs) were constructed from the coassembled contigs as follows. Note that coassembling across multiple samples allowed us to use differential contig

coverages to improve MAG binning (Albertsen et al. 2013; Chen et al. 2020). For each sample, reads were mapped back to contigs using `bowtie2` v2.5.1 (Langmead and Salzberg 2012) with option '`-no-unal`'. SAM files generated by `bowtie2` were converted to BAM files using `samtools` v1.17. BAM files and contig fasta files were used as input to `MetaBAT2` v2.15-2 (Kang et al. 2019) for binning contigs into MAGs, with options '`-m 3000 -s 200000`'. This yielded a separate set of MAGs for each combination of treatment and material type. MAG qualities (completeness & contamination) were estimated using `checkM2` v1.0.1 (Chklovski et al. 2023). MAGs were taxonomically identified using the `GTDB-Tk` toolkit, workflow `classify_wf` v2.3.0 (Chaumeil et al. 2020). An overview of generated MAGs is given in Table S8. MAG completeness and contamination levels are visualised using standard box plots in Figure 4, which was generated using the python package `matplotlib`, functions `bxp` and `scatter`.

4.6 | Statistical Comparisons of Microbial Gene Composition

Pairwise dissimilarities in terms of gene (KO) or gene group (KEGG C) composition were computed using the abundance-weighted Bray-Curtis metric (Legendre and Legendre 1998). Metric multidimensional scaling and PERMANOVA analyses based on those dissimilarities were performed as described earlier for OTUs. Statistical comparisons of KEGG C gene group abundances between treatments (location test), separately for each material, were conducted using the Welch test statistic (Alekseyenko 2016) and a permutation null model (Table S7, Data S4). Further, gene and gene group proportions were visually and statistically compared between treatments using a similar approach as described above for OTUs (Figure 3). Only genes with a mean proportion of at least 0.001% in both treatments were considered in this analysis to avoid excessive noise associated with sampling stochasticity.

Author Contributions

S.L. designed the study, built the prototype desiccator, performed the sample desiccations and data analyses, and wrote a first version of the manuscript. C.E.M. performed the molecular work and contributed to the writing of the manuscript.

Disclosure

Benefit-sharing statement: Benefits from this research result from potential technological and methodological innovations inspired by our study, as well as the sharing of our data as described above.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Raw metagenomic and amplicon sequence data are available on the NCBI Sequence Read Archive under BioProject PRJNA1199176, BioSamples SAMN45866904–SAMN45866963, runs SRR31751350–SRR31751409 (metagenomes) and SRR31749508–SRR31749567 (16S rRNA gene amplicons). Sample metadata with 16S rRNA success metrics are available as Data S1. Sample metadata with metagenomic success metrics are available as Data S2. MAG details, including completeness, contamination, and taxonomic identity, are given in Data S3. Statistical comparison results (location tests) for individual KEGG-C group proportions are given in Data S4. Other intermediate data outputs, such as contigs or OTU representative sequences, are available from the corresponding author on reasonable request. All software used in this paper has been described in Section 2 and are freely available online.

References

Albertsen, M., P. Hugenholtz, A. Skarshewski, K. L. Nielsen, G. W. Tyson, and P. H. Nielsen. 2013. "Genome Sequences of Rare, Uncultured Bacteria Obtained by Differential Coverage Binning of Multiple Metagenomes." *Nature Biotechnology* 31: 533–538.

Alekseyenko, A. V. 2016. "Multivariate Welch t-Test on Distances." *Bioinformatics* 32: 3552–3558.

Anderson, M. J. 2017. "Permutational Multivariate Analysis of Variance (PERMANOVA)." In *Wiley StatsRef: Statistics Reference Online*, edited by N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, and J. Teugels, 1–15. John Wiley & Sons.

Aramaki, T., R. Blanc-Mathieu, H. Endo, et al. 2019. "KofamKOALA: KEGG Ortholog Assignment Based on Profile HMM and Adaptive Score Threshold." *Bioinformatics* 36: 2251–2252.

Bensch, H. M., C. Tolf, J. Waldenström, D. Lundin, M. Zöttl, and K. Kormas. 2022. "Freeze-Drying Can Replace Cold-Chains for Transport and Storage of Fecal Microbiome Samples." *PeerJ* 10: e13095.

Blake, R. D., and S. G. Delcourt. 1998. "Thermal Stability of DNA." *Nucleic Acids Research* 26: 3323–3332.

Bolla, P. A., M. de los Angeles Serradell, P. J. de Urraza, and G. L. De Antoni. 2011. "Effect of Freeze-Drying on Viability and In Vitro Probiotic Properties of a Mixture of Lactic Acid Bacteria and Yeasts Isolated From Kefir." *Journal of Dairy Research* 78: 15–22.

Borg, I., and P. J. F. Groenen. 2005. *Modern Multidimensional Scaling: Theory and Applications*. Springer Series in Statistics. 2nd ed. Springer.

Bowers, R. M., N. C. Kyrpides, R. Stepanauskas, et al. 2017. "Minimum Information About a Single Amplified Genome (MISAG) and a Metagenome-Assembled Genome (MIMAG) of Bacteria and Archaea." *Nature Biotechnology* 35: 725–731.

Brooks, J. P., D. J. Edwards, M. D. Harwich, et al. 2015. "The Truth About Metagenomics: Quantifying and Counteracting Bias in 16S rRNA Studies." *BMC Microbiology* 15: 66.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. "DADA2: High-Resolution Sample Inference From Illumina Amplicon Data." *Nature Methods* 13: 581–583.

Chao, A., R. L. Chazdon, R. K. Colwell, and T. J. Shen. 2005. "A New Statistical Approach for Assessing Similarity of Species Composition With Incidence and Abundance Data." *Ecology Letters* 8: 148–159.

Chaumeil, P. A., A. J. Mussig, P. Hugenholtz, and D. H. Parks. 2020. "GTDB-Tk: A Toolkit to Classify Genomes With the Genome Taxonomy Database." *Bioinformatics* 36: 1925–1927.

Chen, L. X., K. Anantharaman, A. Shaiber, A. M. Eren, and J. F. Banfield. 2020. "Accurate and Complete Genomes From Metagenomes." *Genome Research* 30: 315–333.

Chklovski, A., D. H. Parks, B. J. Woodcroft, and G. W. Tyson. 2023. "CheckM2: A Rapid, Scalable and Accurate Tool for Assessing Microbial Genome Quality Using Machine Learning." *Nature Methods* 20: 1203–1212.

Cuthbertson, L., G. B. Rogers, A. W. Walker, et al. 2015. "Implications of Multiple Freeze-Thawing on Respiratory Samples for Culture-Independent Analyses." *Journal of Cystic Fibrosis* 14: 464–467.

Edgar, R. C. 2018. "Updating the 97% Identity Threshold for 16S Ribosomal RNA OTUs." *Bioinformatics* 34: bty113.

Glöckner, F. O., P. Yilmaz, C. Quast, et al. 2017. "25 Years of Serving the Community With Ribosomal RNA Gene Reference Databases and Tools." *Journal of Biotechnology* 261: 169–176.

Good, P. 2013. *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. Springer.

Hale, V. L., C. L. Tan, R. Knight, and K. R. Amato. 2015. "Effect of Preservation Method on Spider Monkey (*Ateles geoffroyi*) Fecal Microbiota Over 8 Weeks." *Journal of Microbiological Methods* 113: 16–26.

Heckly, R. J., and W. W. Umbreit. 1961. *Preservation of Bacteria by Lyophilization*. Vol. 3, 1–76. Academic Press.

Hyatt, D., G. L. Chen, P. F. LoCascio, M. L. Land, F. W. Larimer, and L. J. Hauser. 2010. "Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification." *BMC Bioinformatics* 11: 119.

Jiang, H., R. Lei, S. W. Ding, and S. Zhu. 2014. "Skewer: A Fast and Accurate Adapter Trimmer for Next-Generation Sequencing Paired-End Reads." *BMC Bioinformatics* 15: 1–12.

Kang, D. D., F. Li, E. Kirton, et al. 2019. "Metabat 2: An Adaptive Binning Algorithm for Robust and Efficient Genome Reconstruction From Metagenome Assemblies." *PeerJ* 7: e7359.

Kim, M., H. S. Oh, S. C. Park, and J. Chun. 2014. "Towards a Taxonomic Coherence Between Average Nucleotide Identity and 16S rRNA Gene Sequence Similarity for Species Demarcation of Prokaryotes." *International Journal of Systematic and Evolutionary Microbiology* 64: 346–351.

Kirk, G. J. D., P. H. Bellamy, and R. M. Lark. 2010. "Changes in Soil pH Across England and Wales in Response to Decreased Acid Deposition." *Global Change Biology* 16: 3111–3119.

Langmead, B., and S. L. Salzberg. 2012. "Fast Gapped-Read Alignment With Bowtie 2." *Nature Methods* 9: 357–359.

Legendre, P., and L. Legendre. 1998. *Numerical Ecology. Developments in Environmental Modelling*. 2nd ed. Elsevier.

Li, D., C. M. Liu, R. Luo, K. Sadakane, and T. W. Lam. 2015. "MEGAHIT: An Ultra-Fast Single-Node Solution for Large and Complex Metagenomics Assembly via Succinct de Bruijn Graph." *Bioinformatics* 31: 1674–1676.

Liu, Z. P., M. A. Shao, and Y. Q. Wang. 2013. "Large-Scale Spatial Interpolation of Soil pH Across the Loess Plateau, China." *Environmental Earth Sciences* 69: 2731–2741.

Louca, S., M. Doebeli, and L. W. Parfrey. 2018. "Correcting for 16S rRNA Gene Copy Numbers in Microbiome Surveys Remains an Unsolved Problem." *Microbiome* 6: 41.

Louca, S., S. M. S. Jacques, A. P. F. Pires, et al. 2016. "High Taxonomic Variability Despite Stable Functional Structure Across Microbial Communities." *Nature Ecology & Evolution* 1: 0015.

Ma, B., C. Lu, Y. Wang, et al. 2023. "A Genomic Catalogue of Soil Microbiomes Boosts Mining of Biodiversity and Genetic Resources." *Nature Communications* 14: 7318.

Mistry, J., R. D. Finn, S. R. Eddy, A. Bateman, and M. Punta. 2013. "Challenges in Homology Search: HMMER3 and Convergent Evolution of Coiled-Coil Regions." *Nucleic Acids Research* 41: e121.

Mitchell, K. R., and C. D. Takacs-Vesbach. 2008. "A Comparison of Methods for Total Community Dna Preservation and Extraction From Various Thermal Environments." *Journal of Industrial Microbiology and Biotechnology* 35: 1139–1147.

Murphy, M. A., L. P. Waits, and K. C. Kendall. 2000. "Quantitative Evaluation of Fecal Drying Methods for Brown Bear DNA Analysis." *Wildlife Society Bulletin* 28: 951–957.

Murphy, M. A., L. P. Waits, K. C. Kendall, S. K. Wasser, J. A. Higbee, and R. Bogden. 2002. "An Evaluation of Long-Term Preservation Methods for Brown Bear (*Ursus arctos*) Faecal DNA Samples." *Conservation Genetics* 3: 435–440.

Nayfach, S., S. Roux, R. Seshadri, et al. 2021. "A Genomic Catalog of Earth's Microbiomes." *Nature Biotechnology* 39: 499–509.

Parada, A. E., D. M. Needham, and J. A. Fuhrman. 2016. "Every Base Matters: Assessing Small Subunit rRNA Primers for Marine Microbiomes With Mock Communities, Time Series and Global Field Samples." *Environmental Microbiology* 18: 1403–1414.

Parks, D. H., C. Rinke, M. Chuvochina, et al. 2017. "Recovery of Nearly 8,000 Metagenome-Assembled Genomes Substantially Expands the Tree of Life." *Nature Microbiology* 2: 1533–1542.

Pesarin, F., and L. Salmaso. 2010. "The Permutation Testing Approach: A Review." *Statistica* 70: 481–509.

Poulsen, C. S., R. S. Kaas, F. M. Aarestrup, and S. J. Pamp. 2021. "Standard Sample Storage Conditions Have an Impact on Inferred Microbiome Composition and Antimicrobial Resistance Patterns." *Microbiology Spectrum* 9: e0138721.

Rissanen, A. J., E. Kurhela, T. Aho, T. Oittinen, and M. Tiirola. 2010. "Storage of Environmental Samples for Guaranteeing Nucleic Acid Yields for Molecular Microbiological Studies." *Applied Microbiology and Biotechnology* 88: 977–984.

Rognes, T., T. Flouri, B. Nichols, C. Quince, and F. Mahé. 2016. "VSEARCH: A Versatile Open Source Tool for Metagenomics." *PeerJ* 4: e2584.

Sharon, I., and J. F. Banfield. 2013. "Genomes From Metagenomics." *Science* 342: 1057–1058.

Soufi, H. H., R. Porch, M. V. Korchagina, et al. 2024. "Taxonomic Variability and Functional Stability Across Oregon Coastal Subsurface Microbiomes." *Communications Biology* 7: 1663.

Soufi, H. H., D. Tran, and S. Louca. 2024. "Microbiology of Big Soda Lake, a Multi-Extreme Meromictic Volcanic Crater Lake in the Nevada Desert." *Environmental Microbiology* 26: e16578.

Thompson, L. R., J. G. Sanders, D. McDonald, et al. 2017. "A Communal Catalogue Reveals Earth's Multiscale Microbial Diversity." *Nature* 551: 457–463.

Thomson, G. W. 1946. "The Antoine Equation for Vapor-Pressure Data." *Chemical Reviews* 38: 1–39.

Tringe, S. G., C. von Mering, A. Kobayashi, et al. 2005. "Comparative Metagenomics of Microbial Communities." *Science* 308: 554–557.

Walters, W., E. R. Hyde, D. Berg-Lyons, et al. 2015. "Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys." *Msystems* 1: e00009-15.

Williams, K. E., K. P. Huyvaert, and A. J. Piaggio. 2016. "No Filters, No Fridges: A Method for Preservation of Water Samples for eDNA Analysis." *BMC Research Notes* 9: 298.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Appendix S1:** men70020-sup-0001-AppendixS1.pdf. **Data S1:** men70020-sup-0002-DataS1.tsv. **Data S2:** men70020-sup-0003-DataS2.tsv. **Data S3:** men70020-sup-0004-DataS3.tsv. **Data S4:** men70020-sup-0005-DataS4.xlsx.

Effects of vacuum-heat-assisted sample desiccation on microbiome surveys
- Supplemental Materials -

Stilianos Louca[1,2,*] & Claire E. Mullin[1,2]

[1]Department of Biology, University of Oregon, Eugene, USA
[2]Institute of Ecology and Evolution, University of Oregon, Eugene, USA
[*]Corresponding author

S.1 Description of desiccator

In this section we describe the vacuum-heat-assisted desiccator built and used in this study (photo in Fig. S1, diagram in Fig. S2). We hope that the details provided will help others replicate or further improve upon our prototype for their own purposes. The desiccator consists of a vacuum chamber with internal temperature-controlled electric heating. The vacuum chamber is a modified cylindrical 1.5 gallon BACOENG™ vacuum degassing chamber with stainless steel walls, glass lid, silicone sealing gasket and pre-attached vacuum gauge. Vacuum is established using two small serially connected oilless diaphragm vacuum pumps (brand Maisi, model TC-150C, max flow 15 L/min), which combined can achieve nearly complete vacuum (<0.01 bar) in less than 5 minutes. Heating of samples is done in two conventional rectangular dry bath aluminum heating blocks, which can accommodate up to 24 centrifuge tubes of 5 mL capacity. Each block is heated using two ceramic Positive Temperature Coefficient (PTC) elements, with a terminal temperature of 60°C at 12 V under standard conditions. To allow for arbitrary target temperatures, the PTC elements of each block are powered through a digital adjustable temperature controller, with the controller's temperature sensor being inserted in the block center. In our case, each controllers' target temperature was set between 40°C (low) and 41°C (high). At this temperature, water begins to boil when the chamber's pressure drops below 0.073 bar or 55.12 mm Hg (Fig. S11). If desiccation is to be performed at lower temperatures (for example to reduce power consumption), a greater vacuum is needed.

Water removal: The water released by the drying samples is partly removed through the pumping progress, however some residual moisture remains and condensates inside the chamber. To prevent condensation on the glass ceiling, which would risk water dropping back into the open sample tubes, we also attached two ceramic PTC heating elements on the ceiling's interior surface. Heating the ceiling ensures that condensation is restricted to the colder steel walls of the chamber. Water droplets forming on the walls ultimately descend to the chamber's metallic bottom, where they can be wiped off with a cloth once desiccation is complete. The ceiling PTC elements have a terminal temperature of 70°C at 12V, although PTCs with a lower terminal temperature would probably suffice as well.

Power consumption: The desiccator is powered by 12V DC, which is available in most personal vehicles or may be delivered by a deep-cycle battery. Under standard conditions, the maximum power consumption

is 40 W for the two pumps, 40W for the four sample-heating PTC elements, and 77 W for the ceiling PTC elements. The PTC elements' power consumption rapidly decreases as their temperature increases. At around 40°C, the sample-heating PTC elements consume only 20 W (when active) and the ceiling PTC elements consume only 18 W. Similarly, the power consumption of the pumps decreases as vacuum is established, and near vacuum the pumps consume only 27 W. Thus, overall power consumption gradually drops during the desiccator's startup. Each of these three parts (block PTCs, ceiling PTCs, pumps) can be powered on and off independently, to limit the initial peak power consumption if necessary. Hence, for example, one may first establish the vacuum, then turn the pumps off, activate the ceiling heating until power consumption drops, and lastly activate the sample heating. Based on our experience, the vacuum pumps only need to be run for a few minutes every 2–3 hours to maintain adequate vacuum throughout the desiccation process. Once the target temperature of 40°C is reached, the long-term average power consumption to needed maintain this temperature is 21 W. Hence, over a 24 hour operation the prototype consumes approximately 0.5 kWh.

Wiring: All necessary wires entering the vacuum chamber, i.e., for powering the PTC elements and for the temperature sensors, are driven through small holes drilled into the steel walls and sealed using marine-grade adhesive sealant. It was important to also seal the ends of all wires to prevent air from leaking into the vacuum chamber and eroding the vacuum. Indeed, during early trials we observed that air could enter some open-ended wires at the ends and be sucked through those wires into the chamber.

Size and cost: The desiccator is mounted in a sturdy plastic Hudson Exchange™ container, with dimensions 24×15×11 inches and partly padded with a rubber mat for sound insulation. The desiccator's total weight, including sound insulation material and carrying case, is about 11 kg. The total cost of construction materials was around $350 USD in 2022; industrial production of such a device at higher quantities is expected to push this cost down significantly.

Operational stability: During our sample desiccations, vacuum pressure varied slightly during the first 5 minutes of the desiccation phase due to the release of relatively large amounts of vapor from the samples, but never exceeded 0.01 bar. Temperature distribution inside the vacuum chamber was assessed using two thermocouples spaced 10 cm apart, with reading differences between them generally being below 1°C except for the initial warm-up phase (up to 3°C difference in the first 10 minutes) due to slight manufacturing differences between nominally identical heating elements. While we do not consider these heterogeneities to be of major importance in most practical applications, if desired one might achieve improvements using higher quality heating elements and more efficient vacuum pumps.

Major envisioned improvements: Based on our experience, we can particularly recommend the following future improvements:

- Install an internal scale for continuously measuring weights and thus monitoring desiccation progress without the need to open the chamber.

- Increase the thermal insulation of the chamber in order to reduce power consumption.

- Use PTCs with a lower terminal temperature to heat the chamber's ceiling, to reduce power consumption.

- Increase the sample capacity relative to the device's footprint by using heating blocks shaped to fit and occupy the full chamber.

- Use wider and shallower sample wells and tubes to increase active evaporation surface and accelerate desiccation.

S.2 Usage of desiccator

This section demonstrates the use of our desiccator step by step. We start with two soil samples collected into sterile 5 mL centrifuge tubes, one from a muddy puddle (S1) and one from a dry garden lawn (S2).

Remove the cap of each sample tube, cover the opening with a PTFE or PES filter membrane (pore size 0.2 μm, diameter 2-3 cm), firmly held in place using some parafilm™. Ensure that the parafilm fully seals the gap along the mouth's perimeter, while minimizing the amount of parafilm covering the mouth so as to allow the escape of vapor.

Place the sample tubes into the heating blocks inside the desiccator's vacuum chamber. Also ensure that the temperature sensors are inserted into the proper slot in the heating blocks.
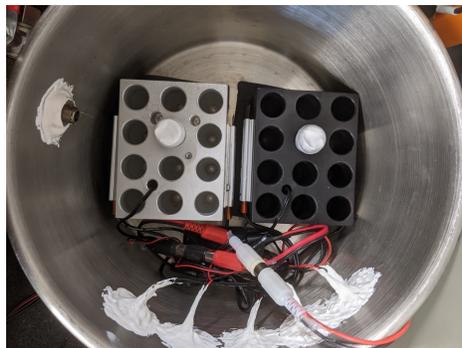
Cover the vacuum chamber with its lid, ensuring that the lid is centered to avoid any gaps. Close the vacuum chamber's vent valve and open its connection valve to the pump.

Note that our desiccator has two PTC heating elements glued to the lid's interior, visible above, to prevent water condensation on the lid during desiccation. Next, connect the desiccator to a 12V DC power supply, for example a portable deep cycle battery, a car's "cigarette lighter" power supply, or through an AC-DC converter to a residential power outlet. Our desiccator uses a female barrel jack (5.5×2.1 mm) for power input. Please see Supplement S.1 regarding the desiccator's adjustable power consumption.
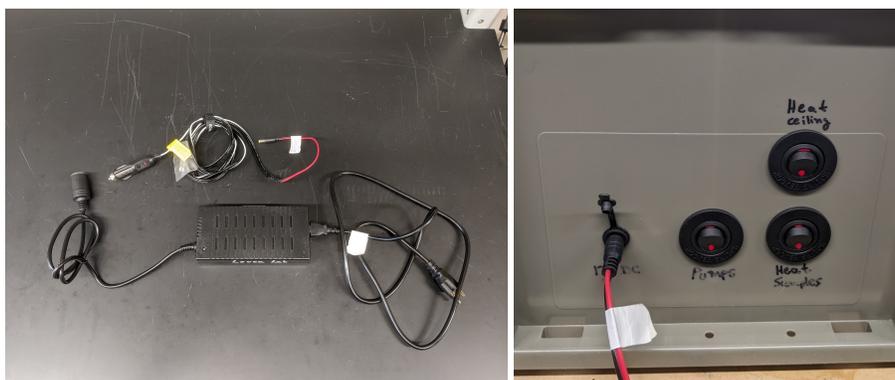
The following electric activation sequence is designed to reduce peak power consumption. Turn on the ceiling heaters and wait for 10 minutes until they approach their terminal temperature, at which point their power consumption will be substantially reduced. Then turn on the thermostat-controlled PTC elements on the heating blocks, ensuring that each thermostat's target temperature range is 40–41°C. The current temperature of the heating blocks can be seen on the thermostat displays.
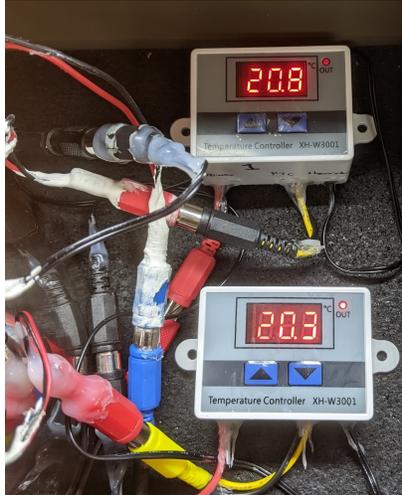
Once the target temperature has been reached for both heat blocks, turn on the vacuum pumps. Wait for a few minutes until nearly complete vacuum is reached and pressure has stabilized (<0.01 bar). The rate of vapor release from the samples will now typically be at its highest, and will gradually decrease thereafter. Some of the vapor will condensate on the chamber's walls, and some of it will be removed via the pumps. Close the lid of the desiccator's container to minimize noise from the pumps.

From this point forward you may keep the vacuum pumps running, or you may turn them off and only turn them on occasionally briefly to restore vacuum as necessary. If you have more samples, or samples with higher water content, more frequent pumping will be necessary. Further, if you plan to occasionally weigh the samples to monitor desiccation progress, you will need to restore vacuum as well. For this demonstration we kept samples in the desiccator for a predetermined 12 hours. Repeat pumping was only necessary after an initial 3 hours, at which point pressure had risen to about 0.03 bar. The dried samples are shown below:

Table S1: **Materials overview.** Overview of materials considered. Note that from each material we collected 2×5 samples, of which 5 were frozen and 5 were dried.

Material	description	pH	water content (% w/w)	hours dried
feces 1	dog feces	NA	81.9%	21.05
feces 2	feline feces	NA	68.8%	21.05
feces 3	mule deer feces	NA	48.1%	21.05
soil 1	surface soil, urban forest	6.59	18.1%	7.55
soil 2	surface soil, urban grassland	6.86	18.7%	7.55
soil 3	surface soil, urban park	7.90	23.1%	7.55

Table S2: **Statistical comparisons of success metrics.** Overview of location tests of various success metrics, including the average ratio of dried over frozen (averaged over all materials), standardized effect size (SES) and two-sided statistical significance (P-value). P-values below 0.05 are bolded. All tests are based on 60 samples, and P-values are estimated using a permutation null model described in the main article.

Metric	Average ratio dried/frozen	SES	P-value
DNA extraction			
DNA yield (ng)	1.18	+3.0	**0.0041**
peak fragment size (bp)	0.80	-3.82	**0.0006**
260/280 ratio (nm/nm)	1.01	+2.42	**0.0146**
260/230 ratio (nm/nm)	1.37	+1.79	0.10
16S rRNA amplicons			
number of reads	0.87	-0.89	0.39
forward read quality	1.01	+1.41	0.16
reverse read quality	0.99	-1.22	0.22
number of OTUs (α-diversity)	0.94	-1.36	0.18
metagenomes			
forward read quality	0.99	-2.37	**0.013**
reverse read quality	1.00	+0.64	0.53
number of reads	0.98	-0.38	0.71
number of contigs	1.40	+4.93	**<0.001**
number of contigs $\geq$1000 bp	1.86	+6.18	**<0.001**
max contig length (bp)	1.04	+0.50	0.62
number of proteins	1.41	+4.84	**<0.001**
number of KOs	1.20	+4.02	**0.0001**

Table S3: **PERMANOVA analysis, OTU composition.** PERMANOVA test results, comparing various dissimilarity metrics in OTU composition between dried and frozen samples, separately for each material. Considered metrics include abundance-weighted Bray-Curtis with and without rarefaction and abundance-weighted Jaccard with and without rarefaction. Shown are the pseudo F-statistic, the standardized effect size (SES), the statistical significance (P) and the fraction of variance in pairwise distances explained by the treatment (R^2). A P value below 0.05 (shown in bold) indicates systematic differences in inferred OTU composition between dried and frozen samples. Each test is based on 10 samples (5 dried, 5 frozen).

Material	F-statistic	SES	P	R^2
Bray-Curtis, non-rarefied				
feces 1	0.85	-0.36	0.57	0.046
feces 2	1.66	1.19	0.11	0.13
feces 3	1.09	0.041	0.28	0.075
soil 1	4.81	6.10	**0.0082**	0.41
soil 2	1.92	1.62	0.064	0.15
soil 3	1.10	0.011	0.25	0.068
Bray-Curtis, rarefied				
feces 1	1.04	-0.14	0.41	0.074
feces 2	1.89	1.19	0.090	0.14
feces 3	2.39	3.22	**0.0077**	0.23
soil 1	3.67	6.54	**0.0081**	0.37
soil 2	1.49	2.11	**0.044**	0.15
soil 3	1.87	4.73	**0.0044**	0.20
Jaccard, non-rarefied				
feces 1	0.089	-0.67	0.69	0.0083
feces 2	1.12	0.043	0.36	0.076
feces 3	1.46	1.35	0.12	0.093
soil 1	4.64	6.39	**0.0031**	0.39
soil 2	2.39	2.92	**0.026**	0.16
soil 3	1.84	3.66	**0.0017**	0.11
Jaccard, rarefied				
feces 1	0.59	-0.52	0.68	0.037
feces 2	2.29	1.06	0.12	0.10
feces 3	1.58	2.61	**0.013**	0.16
soil 1	2.75	6.19	**0.0006**	0.29
soil 2	1.47	2.62	**0.029**	0.15
soil 3	1.51	5.28	**0.0079**	0.16

Table S4: **PERMANOVA analysis, genus composition.** PERMANOVA test results, comparing various dissimilarity metrics in genus composition between dried and frozen samples, separately for each material. Considered metrics include abundance-weighted Bray-Curtis with and without rarefaction and abundance-weighted Jaccard with and without rarefaction. Shown are the pseudo F-statistic, the standardized effect size (SES), the statistical significance (P) and the fraction of variance in pairwise distances explained by the treatment (R^2). A P value below 0.05 (shown in bold) indicates systematic differences in inferred genus composition between dried and frozen samples. Each test is based on 10 samples (5 dried, 5 frozen).

Material	F-statistic	SES	P	R^2
Bray-Curtis, non-rarefied				
feces 1	0.73	-0.62	0.67	0.035
feces 2	2.1	1.8	0.059	0.15
feces 3	0.74	-0.31	0.46	0.044
soil 1	5.2	5.4	**0.0066**	0.36
soil 2	2.6	1.3	0.079	0.13
soil 3	0.98	-0.12	0.28	0.043
Bray-Curtis, rarefied				
feces 1	1.06	0.019	0.36	0.072
feces 2	2.87	1.94	**0.042**	0.19
feces 3	5.36	5.17	**0.0058**	0.40
soil 1	8.26	7.65	**0.008**	0.58
soil 2	2.09	1.72	0.067	0.17
soil 3	3.61	5.59	**0.0082**	0.33
Jaccard, non-rarefied				
feces 1	3.91	1.35	0.091	0.045
feces 2	2.65	1.08	0.19	0.11
feces 3	1.44	-0.023	0.37	0.023
soil 1	6.62	5.60	**0.0057**	0.28
soil 2	5.21	3.12	**0.013**	0.18
soil 3	2.11	3.69	**0.004**	0.10
Jaccard, rarefied				
feces 1	0.96	-0.13	0.44	0.045
feces 2	1.57	0.57	0.21	0.065
feces 3	1.38	0.22	0.22	0.12
soil 1	4.59	6.50	**0.0054**	0.37
soil 2	2.29	2.49	**0.029**	0.18
soil 3	1.72	4.57	**0.0043**	0.18

Table S5: **PERMANOVA analysis, gene composition.** PERMANOVA test results, comparing weighted Bray-Curtis dissimilarities in gene (KO) composition between dried and frozen samples, separately for each material. Shown are the pseudo F-statistic, the standardized effect size (SES), the statistical significance (P) and the fraction of variance in pairwise distances explained by the treatment (R^2). A P value below 0.05 (shown in bold) indicates systematic differences in inferred gene composition between dried and frozen samples. Each test is based on 10 samples (5 dried, 5 frozen).

Material	F-statistic	SES	P	R^2
feces 1	0.79	-0.23	0.38	0.042
feces 2	9.3	6.9	**0.003**	0.44
feces 3	20.4	11.1	**0.002**	0.79
soil 1	16.8	10.7	**0.005**	0.80
soil 2	6.1	8.5	**0.003**	0.52
soil 3	7.0	8.8	**0.005**	0.59

Table S6: **PERMANOVA analysis, KEGG C composition.** PERMANOVA test results, comparing weighted Bray-Curtis dissimilarities in KEGG-C gene group composition between dried and frozen samples, separately for each material. Shown are the pseudo F-statistic, the standardized effect size (SES) and the statistical significance (P). A P value below 0.05 (shown in bold) indicates systematic differences in inferred gene group composition between dried and frozen samples. Each test is based on 10 samples (5 dried, 5 frozen).

Material	F-statistic	SES	P	R^2
feces 1	0.19	-0.42	0.85	0.012
feces 2	13.6	8.6	**0.003**	0.54
feces 3	46.2	12.9	**0.004**	0.87
soil 1	36.7	12.5	**0.004**	0.88
soil 2	11.4	9.9	**0.005**	0.69
soil 3	10.9	9.7	**0.001**	0.71

Table S7: **Location tests for KEGG C group abundances.** Overview of statistical location tests of KEGG C gene group proportions, i.e., comparing proportions of individual gene groups between treatments and separately for each material. Shown are the total number of gene groups considered, the number of groups exhibiting a statistically significant difference between the two treatments ($P < 0.05$), and the number of significant groups after Bonferroni correction. Each test is based on 10 samples (5 dried, 5 frozen).

Material	Ngroups	Nsignificant	Nsignificant Bonferroni
feces 1	487	7	0
feces 2	376	210	57
feces 3	306	250	120
soil 1	360	247	99
soil 2	364	165	41
soil 3	362	151	35

Table S8: **Overview of MAGs.** Number of MAGs constructed for each combination of material type and treatment. Also shown are the number of MAGs of at least medium quality (completeness $\geq$50%, contamination $\leq$10%).

Material type	treatment	NMAGs	Mean completeness (%)	Mean contamination (%)	NMAGs at least medium quality
feces	dried	135	55.5	6.34	61
feces	frozen	146	56.9	7.35	62
soil	dried	14	35.0	4.92	0
soil	frozen	7	26.4	1.90	0
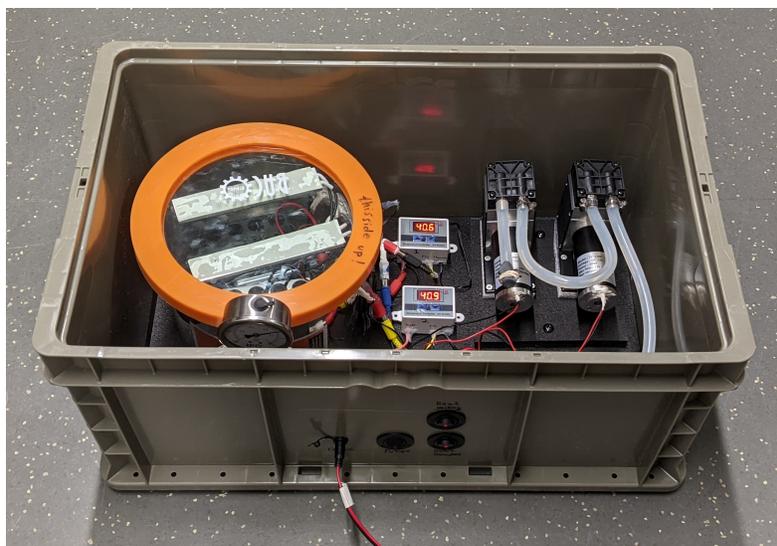
Figure S1: **Desiccator prototype.** Vacuum- and heat-assisted sample desiccator developed and used in this study.
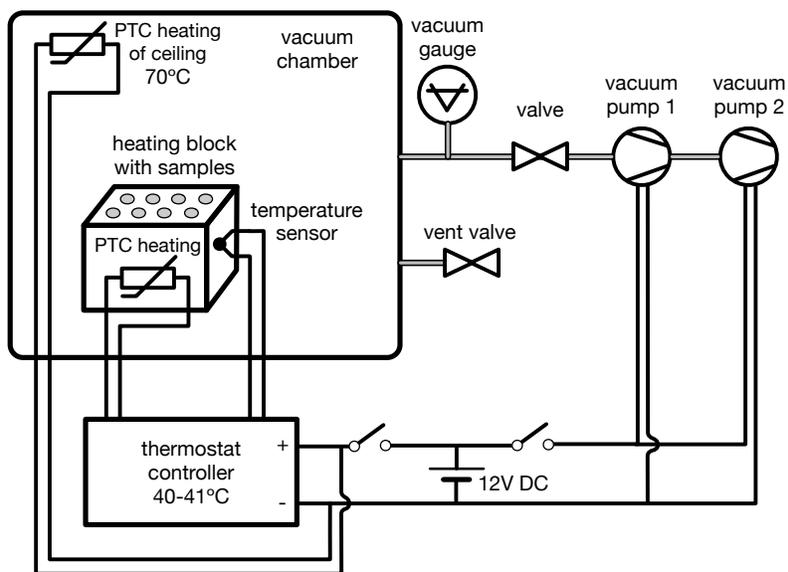
Figure S2: **Desiccator schematic.** Schematic diagram of the prototype desiccator developed and used in this study. Only one of two thermostat-controlled heating blocks is shown here for simplicity.
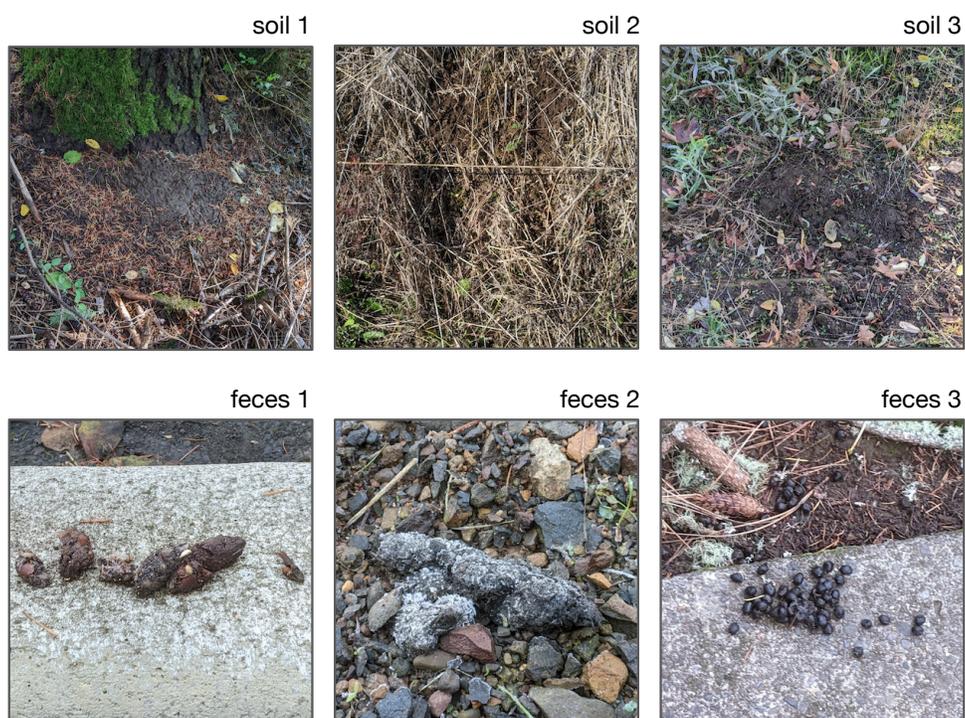
Figure S3: **Photos of samples prior to collection.** Photos of soil samples (top row) and fecal samples (bottom row) prior to collection.
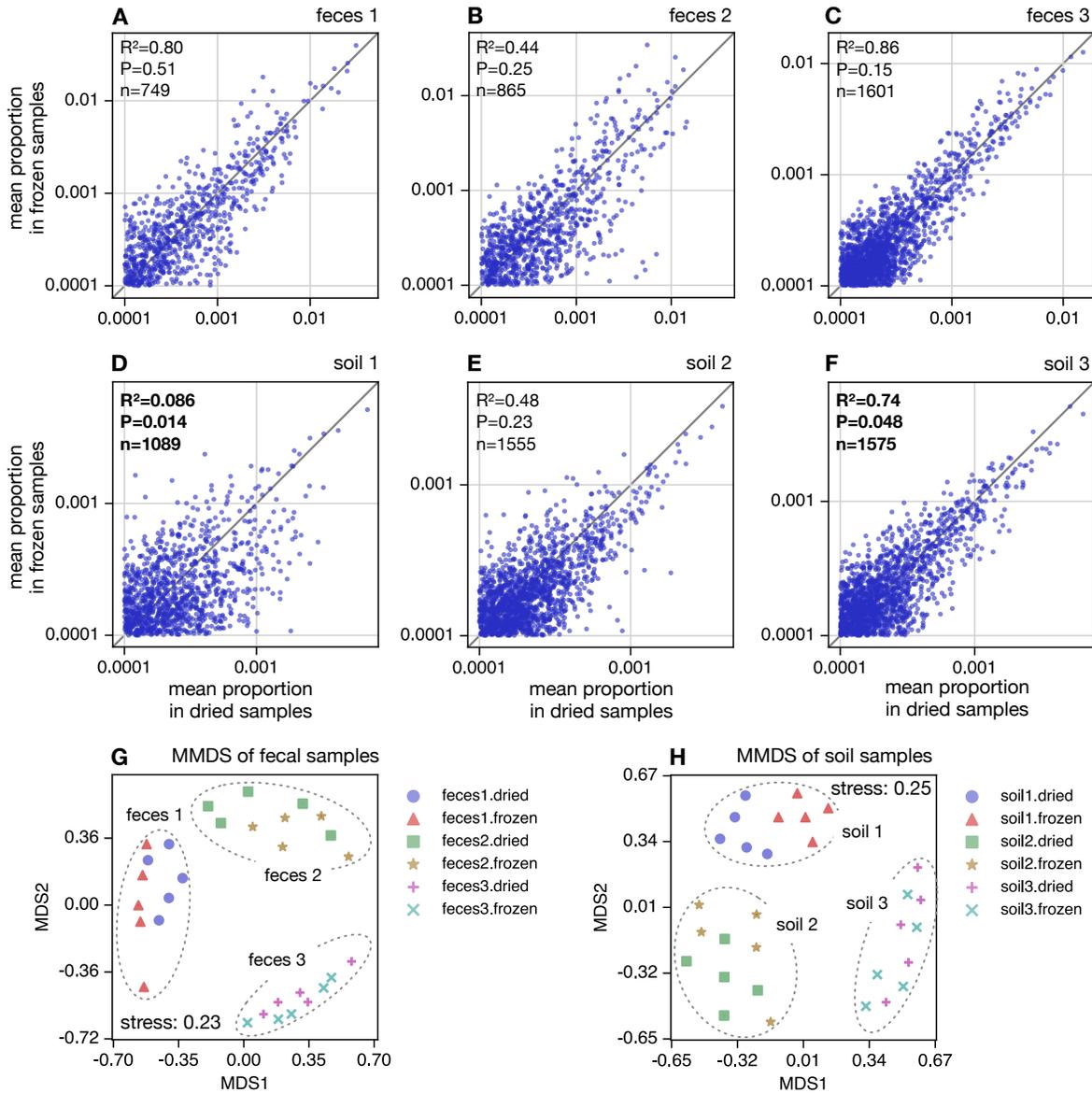
Figure S4: **ASV composition vs. treatment.** (A) Mean ASV proportions in dried fecal 1 samples (horizontal axis) compared to mean ASV proportions in frozen fecal 1 samples (vertical axis, one point per ASV). Averaging of proportions was done among the 5 replicates in each treatment. The diagonal is shown for reference. Inscriptions show the fraction of variance in the vertical axis explained by the horizontal axis (R^2), the number of ASVs considered (n), and the statistical significance of R^2 compared to a permutation null model under which ASV proportions are statistically indistinguishable in the two treatments (P). A significantly low R^2 (i.e., P<0.05) suggests that R^2 is lower than expected by chance, and that dried samples tend to yield different ASV proportions compared to frozen samples. (B–F) Similar to A, but for the remaining samples. Statistically significant R^2 values are boldened. (G) Metric multidimensional scaling plot of abundance-weighted Bray-Curtis dissimilarities between fecal samples, based on ASV proportions. Points correspond to samples, and are shaped and colored according to the material (feces 1–3) and treatment (dried vs frozen). The Kruskal stress is written in the plot. (H) Similar to G, but for soil samples. For similar plots using OTU or genus proportions see Figures 2 and S5, respectively.

14
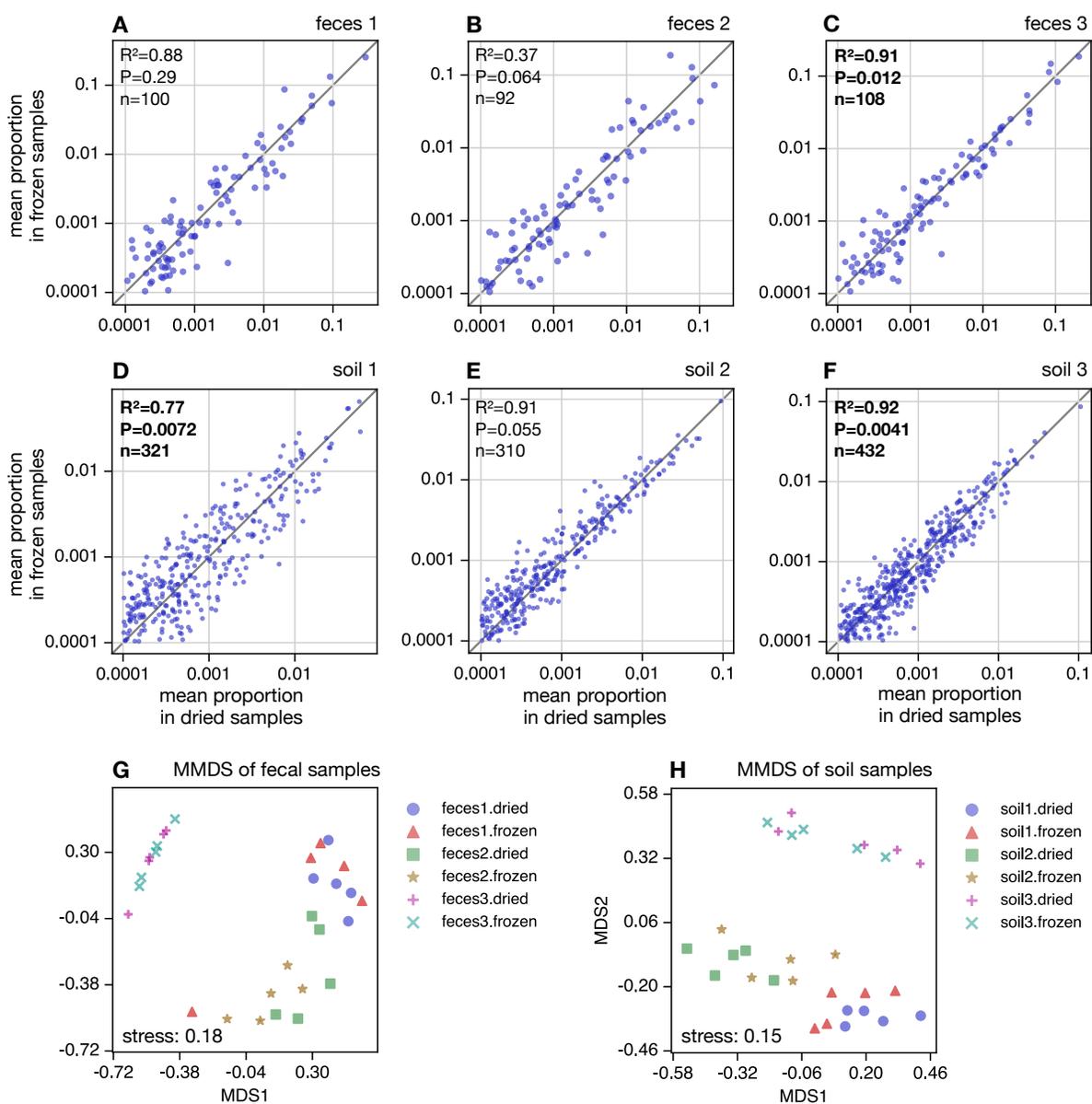
Figure S5: **Genus composition vs. treatment.** (A) Mean genus proportions in dried fecal 1 samples (horizontal axis) compared to mean genus proportions in frozen fecal 1 samples (vertical axis, one point per genus). Averaging of proportions was done among the 5 replicates in each treatment. The diagonal is shown for reference. Inscriptions show the fraction of variance in the vertical axis explained by the horizontal axis (R^2), the number of genera considered (n), and the statistical significance of R^2 compared to a permutation null model under which genus proportions are statistically indistinguishable in the two treatments (P). A significantly low R^2 (i.e., P<0.05) suggests that R^2 is lower than expected by chance, and that dried samples tend to yield different genus proportions compared to frozen samples. (B–F) Similar to A, but for the remaining samples. Statistically significant R^2 values are boldened. (G) Metric multidimensional scaling plot of abundance-weighted Bray-Curtis dissimilarities between fecal samples, based on genus proportions. Points correspond to samples, and are shaped and colored according to the material (feces 1–3) and treatment (dried vs frozen). The Kruskal stress is written in the plot. (H) Similar to G, but for soil samples. For similar plots using ASV or OTU proportions see Figures S4 and 2, respectively.

15
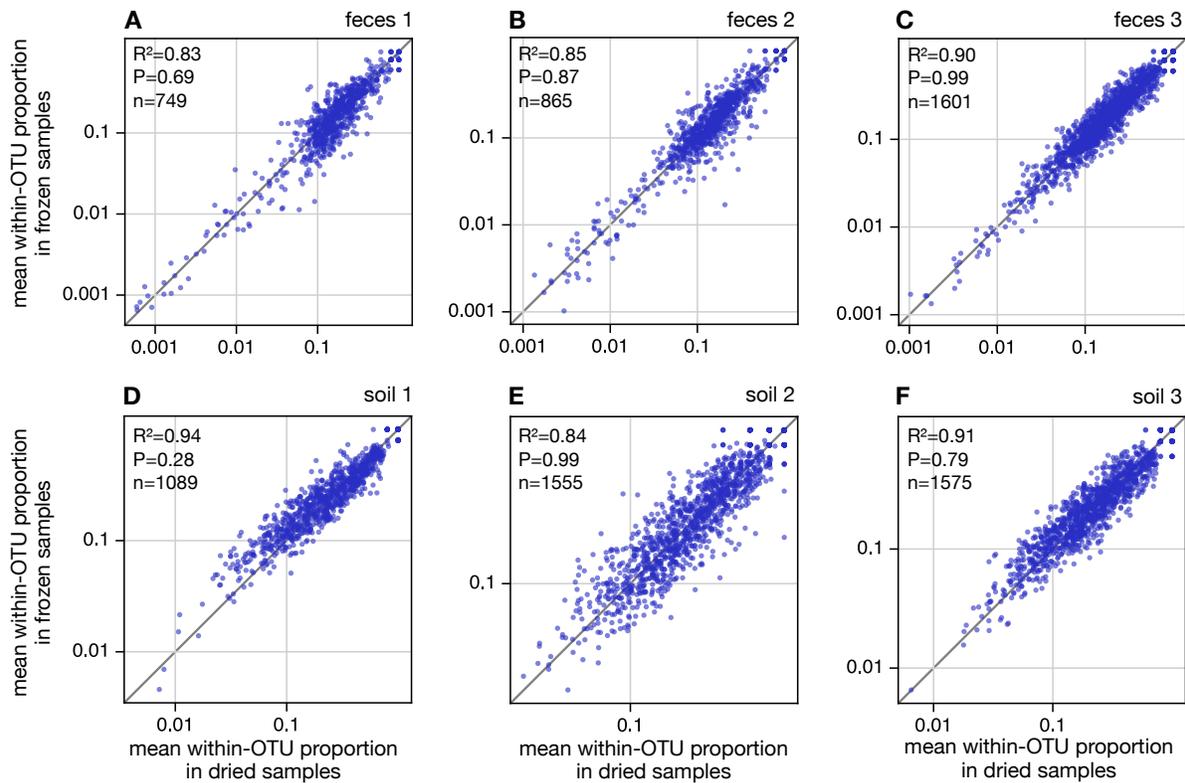
Figure S6: **ASV composition within OTUs vs. treatment.** (A) Mean ASV proportions within their respective OTU in dried fecal 1 samples (horizontal axis) compared to mean ASV proportions within their respective OTU in frozen fecal 1 samples (vertical axis, one point per ASV). Averaging of proportions was done among the 5 replicates in each treatment. The diagonal is shown for reference. Inscriptions show the fraction of variance in the vertical axis explained by the horizontal axis (R^2), the number of ASVs considered (n), and the statistical significance of R^2 compared to a permutation null model under which ASV proportions within their respective OTUs are statistically indistinguishable in the two treatments (P). A significantly low R^2 (i.e., P<0.05) suggests that R^2 is lower than expected by chance, and that dried samples tend to yield different ASV proportions compared to frozen samples. (B–F) Similar to A, but for the remaining samples. Statistically significant R^2 values are boldened.
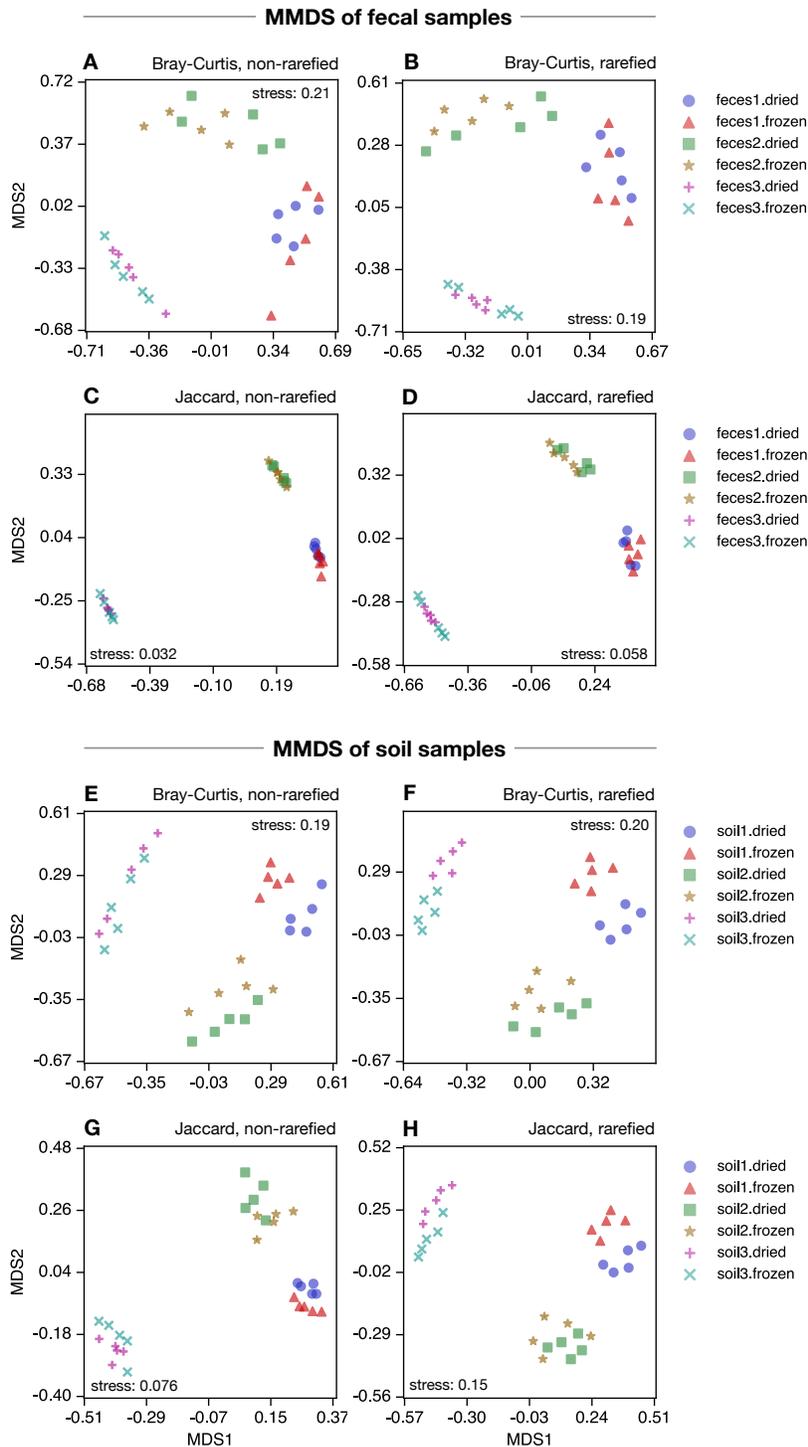
Figure S7: **Metric multidimensional scalings (OTU composition).** (A) Metric multidimensional scaling plot of abundance-weighted Bray-Curtis dissimilarities between fecal samples, based on OTU proportions. Points correspond to samples, and are shaped and colored according to the material (feces 1–3) and treatment (dried vs frozen). The Kruskal stress is written in the plot. (B) Similar to A, but based on a rarefied OTU table where all samples have equal numbers of reads. (C, D) Similar to A and B, but using abundance-weighted Jaccard dissimilarities. (E–H) Similar to A–D, but focusing on soil samples. For analogous PCoA plots see Fig. S12.
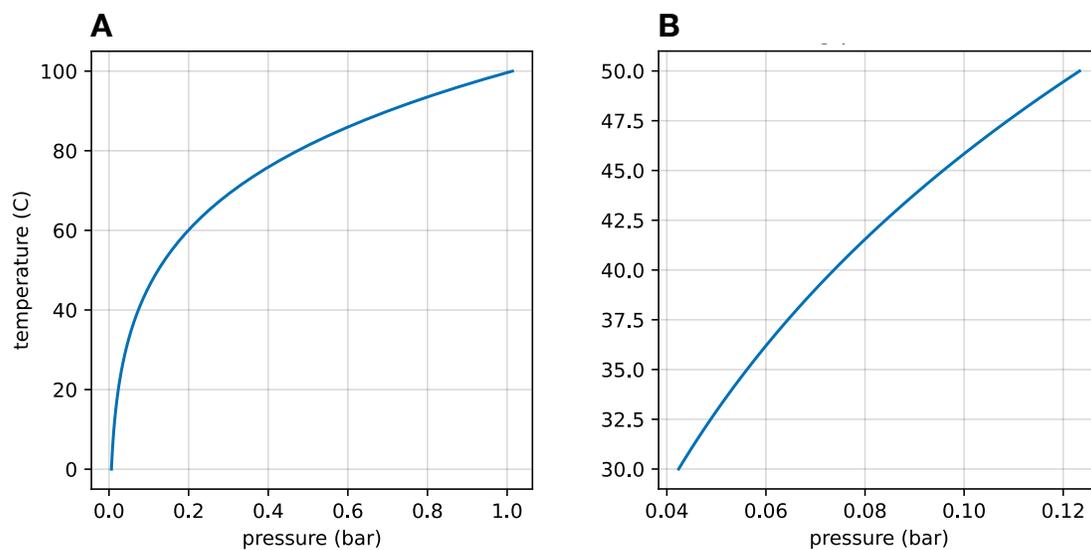
17
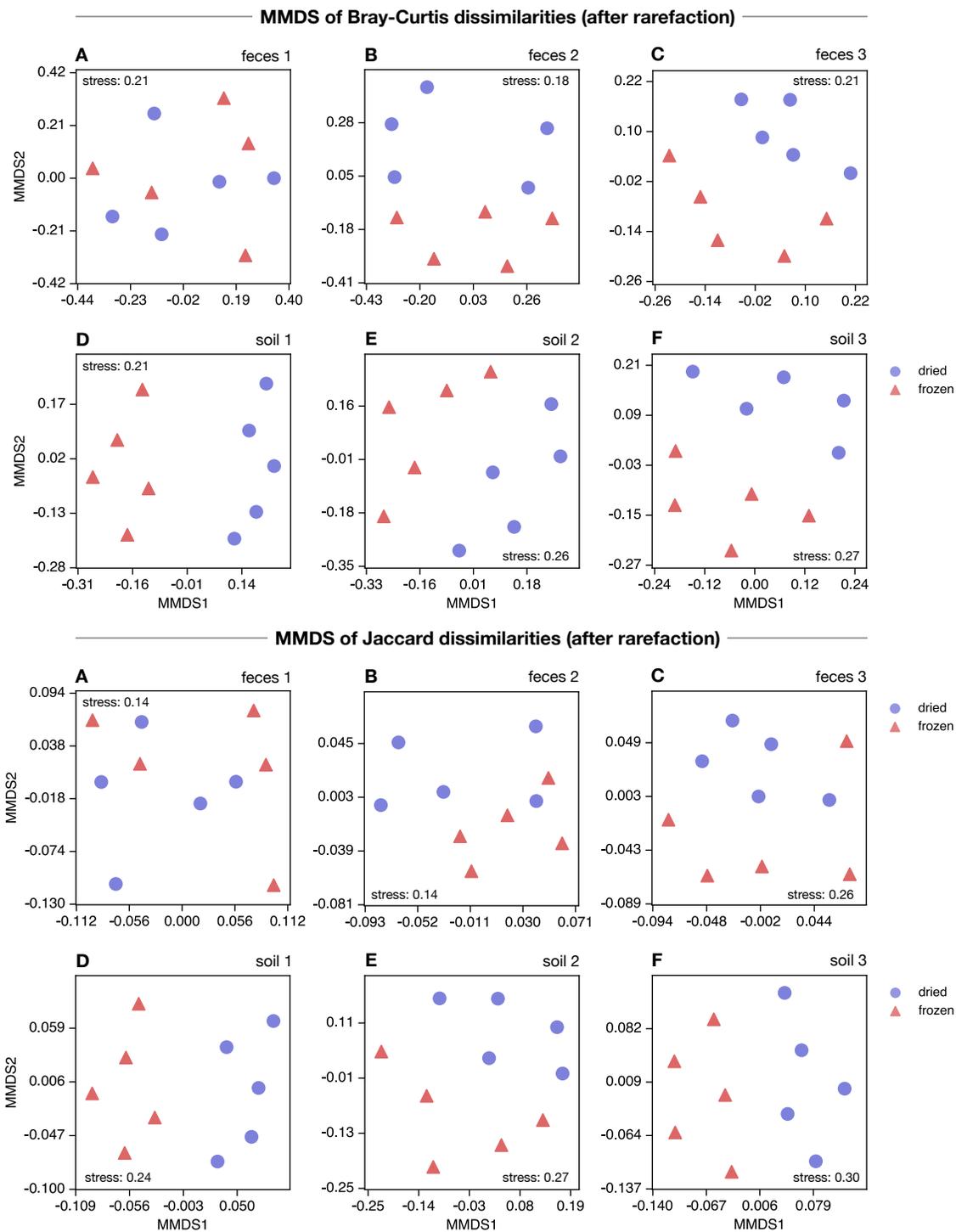
Figure S8: **Metric multidimensional scalings, separately for each material (OTU composition).** (A) Metric multidimensional scaling plot of abundance-weighted Bray-Curtis dissimilarities between "feces 1" samples, based on OTU proportions after rarefaction. Points correspond to samples derived from feces 1 material, and are shaped and colored according to treatment (dried vs frozen). The Kruskal stress is written in the plot. (B–F) Similar to A, but for feces 2–3 and soils 1–3. For PERMANOVA tests quantitatively examining the separation by treatment see Table S3.
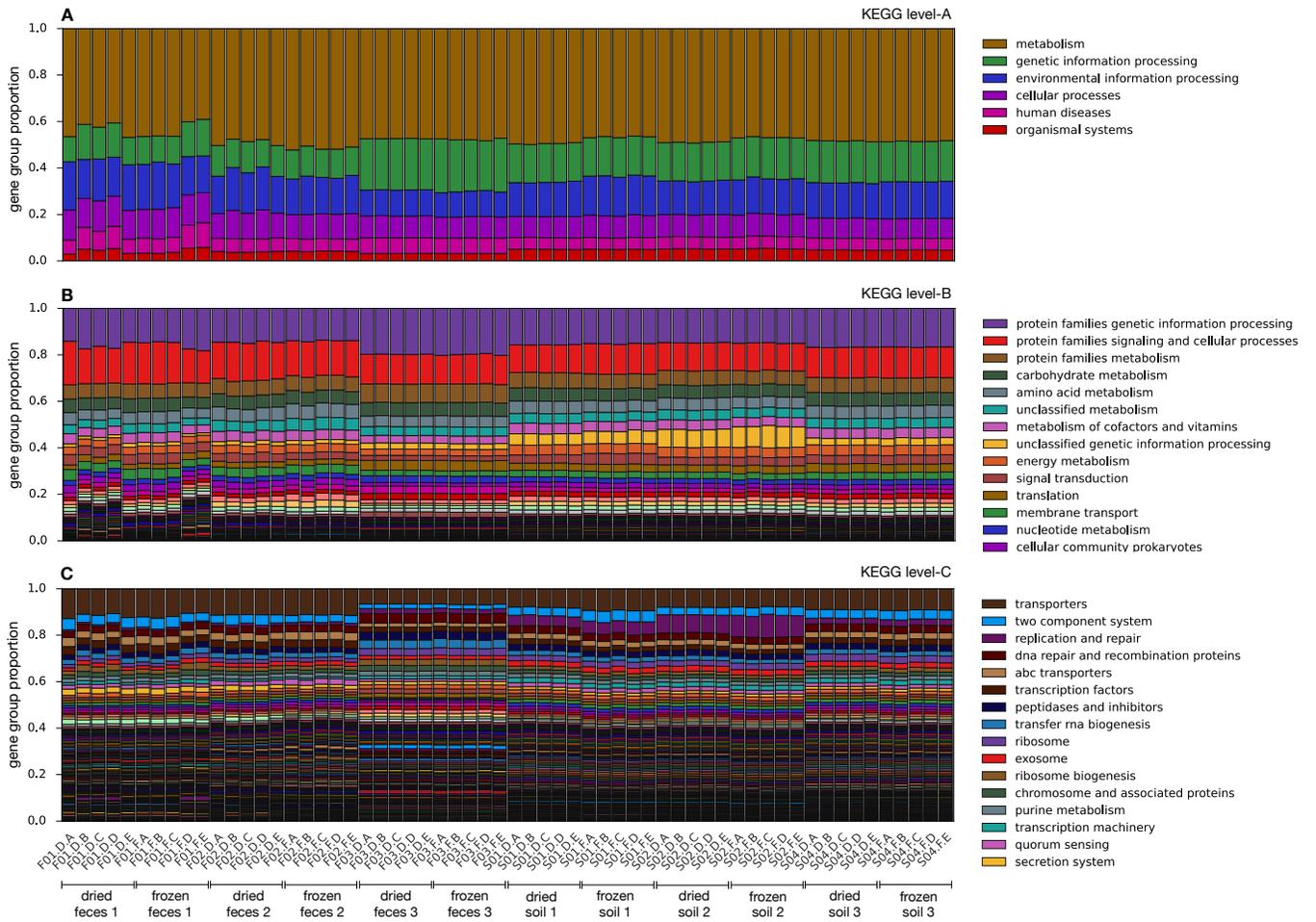
Figure S9: **KEGG gene group profiles.** (A) Estimated proportions of KEGG-level-A gene groups in each sample. Each column corresponds to a sample, and bar segment heights correspond to group proportions, which are based on numbers of reads mapped. (B, C) Similar to A, but showing KEGG level-B and KEGG level-C groups, respectively. In B and C, inly the top 100 groups are shown and only a subset of these is included in the legend to avoid cluttering the figure.
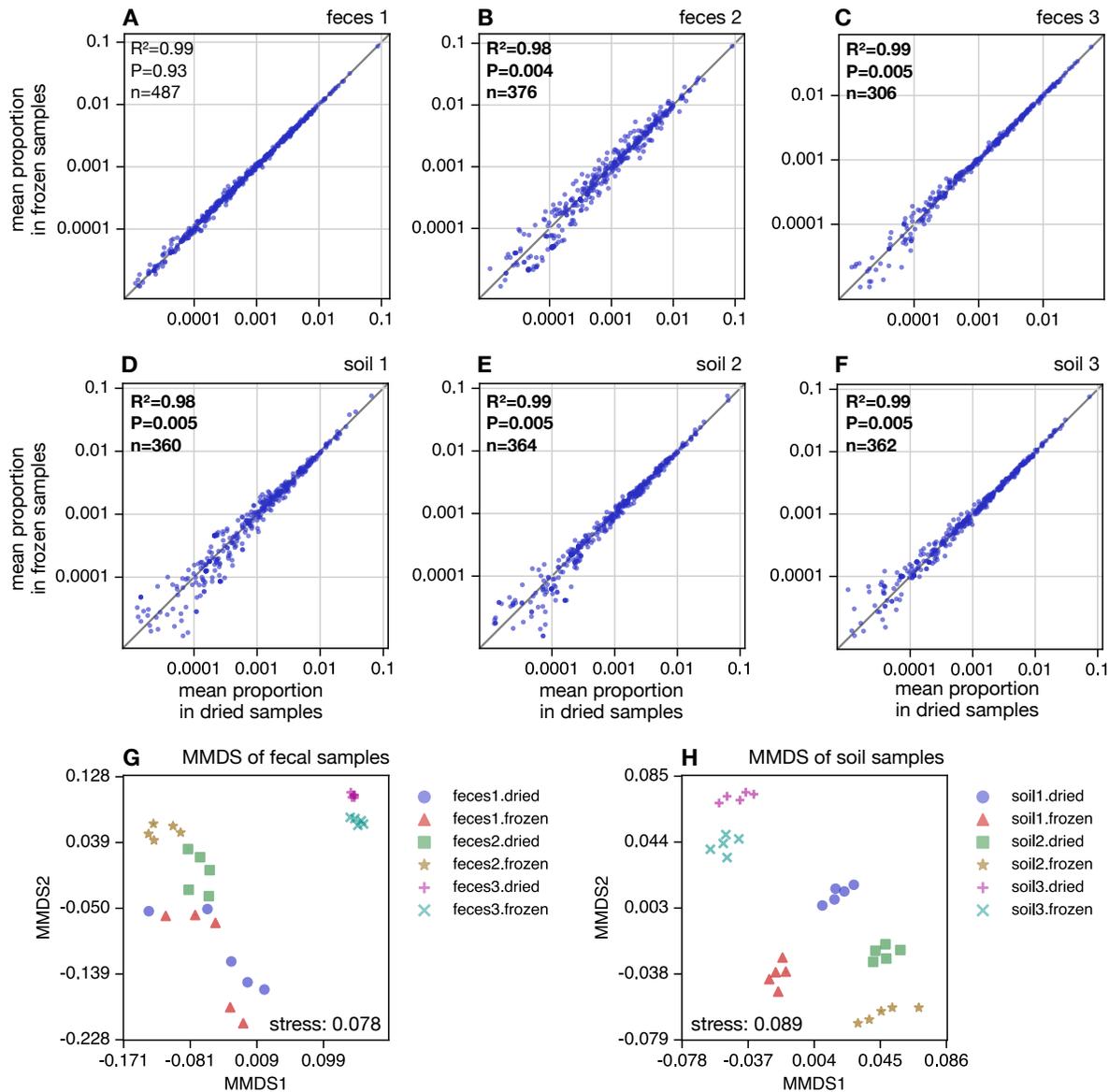
Figure S10: **KEGG C composition vs. treatment.** (A) Mean KEGG C gene group proportions in dried fecal 1 samples (horizontal axis) compared to mean gene group proportions in frozen fecal 1 samples (vertical axis, one point per gene group). Averaging of proportions was done among the 5 replicates in each treatment. The diagonal is shown for reference. Inscriptions show the fraction of variance in the vertical axis explained by the horizontal axis (R^2), the number of gene groups considered (n), and the statistical significance of R^2 compared to a permutation null model under which gene group proportions are statistically indistinguishable in the two treatments (P). A significantly low R^2 (i.e., P<0.05) suggests dried samples tend to yield different gene group proportions compared to frozen samples. (B–F) Similar to A, but for the remaining samples. Statistically significant R^2 values are boldened. (G) Metric multidimensional scaling plot of weighted Bray-Curtis dissimilarities between fecal samples, based on gene group proportions. Points correspond to samples, and are shaped and colored according to the material (feces 1–3) and treatment (dried vs frozen). The Kruskal stress is written in the plot. (H) Similar to G, but for soil samples. For similar plots using gene (KO) proportions see Fig. 3.

Figure S11: **Boiling point of water.** (A) Boiling temperature of water at pressures 0–1 bar. (B) Boiling temperature of within a narrower pressure range, covering temperatures 30–50°C. Based on the empirical formula in (Thomson, 1946, Eq. 29).
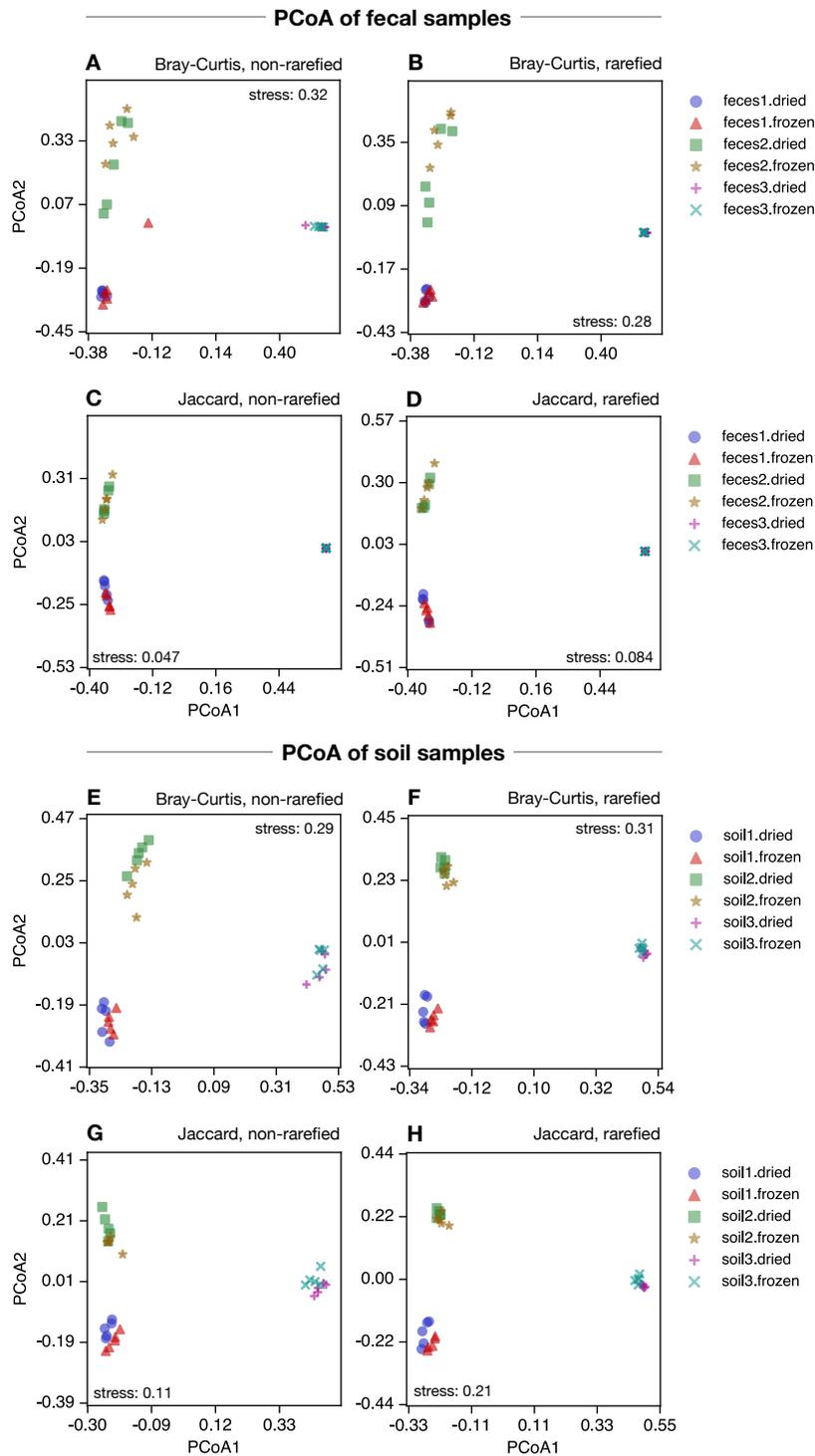
Figure S12: **PCoA (OTU composition).** (A) Principal coordinates analysis (PCoA, aka. classical multidimensional scaling) of abundance-weighted Bray-Curtis dissimilarities between fecal samples, based on OTU proportions. Points correspond to samples, and are shaped and colored according to the material (feces 1–3) and treatment (dried vs frozen). The Kruskal stress is written in the plot. (B) Similar to A, but based on a rarefied OTU table where all samples have equal numbers of reads. (C, D) Similar to A and B, but using abundance-weighted Jaccard dissimilarities. (E–H) Similar to A–D, but focusing on soil samples. For analogous metric multidimensional scaling plots see Fig. S7.

References

Thomson, G.W. (1946) The Antoine equation for vapor-pressure data. *Chemical Reviews*, 38, 1–39.