



Freeze-drying can replace cold-chains for transport and storage of fecal microbiome samples

Hanna M. Bensch^{1,2}, Conny Tolf¹, Jonas Waldenström¹, Daniel Lundin¹ and Markus Zöttl^{1,2}

¹Centre for Ecology and Evolution in Microbial Model Systems (EEMIS), Department of Biology and Environmental Science, Linnaeus University, Sweden

²Kalahari Research Centre, Kuruman River Reserve, Van Zylsrus, South Africa

ABSTRACT

Background. The transport and storage of samples in temperatures of minus 80 °C is commonly considered as the gold standard for microbiome studies. However, studies conducting sample collection at remote sites without a reliable cold-chain would benefit from a sample preservation method that allows transport and storage at ambient temperature.

Methods. In this study we compare alpha diversity and 16S microbiome composition of 20 fecal sample replicates from Damaraland mole-rats (*Fukomys damarensis*) preserved in a minus 80 °C freezer and transported on dry ice to freeze-dried samples that were stored and transported in ambient temperature until DNA extraction.

Results. We found strong correlations between relative abundances of Amplicon Sequence Variants (ASVs) between preservation treatments of the sample, no differences in alpha diversity measures between the two preservation treatments and minor effects of the preservation treatment on beta diversity measures. Our results show that freeze-drying samples can be a useful method for cost-effective transportation and storage of microbiome samples that yields quantitatively almost indistinguishable results in 16S microbiome analyses as those stored in minus 80 °C.

Subjects Biodiversity, Bioinformatics, Ecology, Microbiology, Zoology

Keywords Microbiome, 16S, Fecal samples, DNA metabarcoding, *Fukomys damarensis*, Damaraland mole-rat, Amplicon, Freeze-drying

INTRODUCTION

The transport and storage of microbiome samples on dry ice and in minus 80 °C freezers (*i.e.*, super freezers) is commonly considered the gold standard ([Vandeputte et al., 2017](#)). However, for research carried out in remote locations, at field stations or when the analytical laboratory work is carried out at laboratories far away from the collection site, it can be challenging to provide super freezers and maintain uninterrupted cold chains ([Kim et al., 2017](#); [Vandeputte et al., 2017](#)). Super freezers require a reliable source of electricity and are expensive to purchase and maintain. Ordinary freezers may be a cost efficient alternative and increasing evidence suggests that DNA storage of microbiome samples in those freezers yield similar 16S rRNA gene sequencing data profiles as samples stored in super freezers ([Song et al., 2016](#); [Gavriiliuc et al., 2021](#)). However, because the microbiota composition

Submitted 30 November 2021

Accepted 20 February 2022

Published 15 March 2022

Corresponding author

Hanna M. Bensch,
hanna.bensch@lnu.se

Academic editor

Konstantinos Kormas

Additional Information and
Declarations can be found on
page 12

DOI 10.7717/peerj.13095

© Copyright
2022 Bensch et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

in frozen samples is sensitive to thawing (*Cardona et al., 2012; Song et al., 2016*), freezing requires an uninterrupted cold chain from sampling to library preparation and samples are therefore often transported on dry ice or in liquid nitrogen from the field sites to the laboratory. Transportation of frozen samples hence introduces the risk of thawing and incurs additional costs of cold chain logistics. This is especially problematic when the sampling location is in a remote area or in another country than the laboratory facility and commercial couriers are needed.

When freezing is not an available option, the most commonly used and tested alternative storage of samples include ethanol, RNAlater or FTA-cards (*Hale et al., 2015; Song et al., 2016; Vogtmann et al., 2017; Wang et al., 2018*). Although not requiring immediate freezing, these sample treatments are commonly frozen after a couple of hours or days at ambient temperatures for long-term storage prior to library preparation. Out of these alternative methods, ethanol is the cheapest and has been shown to be the most effective for preserving the microbiome community structure when freezers are not available (*Hale et al., 2015*). However, the use of ethanol comes with drawbacks, for example the additional step of ethanol removal prior to DNA extraction, as well as risk of evaporation. Furthermore, ethanol is both volatile and flammable, which increases shipping costs and restricts export transport options (*Song et al., 2016; Vandeputte et al., 2017*). Consequently, methods where samples can be transported at room temperature and without hazardous liquids could be of great value to research projects working at remote locations.

In this study, we evaluate freeze-drying as an alternative method for the preservation of fecal samples prior to transport, storage and 16S sequencing. Although freeze-drying has rarely been used in microbiome studies (*Gavriliuc et al., 2021*), it offers considerable advantages. No chemicals are added and the samples can be stored and transported at ambient temperature, eliminating the need for cold chain logistics and hazardous material handling. Previous studies have offered some support for this preservation technique by showing an overlap in 16S data from freeze-dried human stool samples with data from frozen samples (*Kia et al., 2016*) and little influence of freeze-drying on homogenized neonatal fecal samples (*Shen et al., 2021*). However experimental work evaluating the effects of freeze-drying on field samples and samples from non-human animals is still lacking. To evaluate whether freeze-drying, subsequent long-distance transport, and storage at ambient conditions affects the results of 16S sequencing, we conducted an experiment where we investigated the gut microbiome diversity and composition of 20 replicate fecal samples from Damaraland mole-rats (*Fukomys damarensis*) by comparing freeze-drying with super freezers.

METHODS

Sample collection and preservation treatment

Fecal samples were collected from 20 wild-caught Damaraland mole-rats shortly after capture at the Kuruman River Reserve, South Africa, between 1st of April and 15th of October 2019 as part of a long-term study on social behaviours and ecology of this subterranean rodent (e.g., *Zöttl et al., 2016; Thorley et al., 2021*). The animals were from

five randomly selected family groups (four individuals per group). Each sample was split into two replicate tubes with 1–3 fecal pellets per tube and stored in a minus 80 °C freezer at the field site. After all samples had been collected, one replicate per sample was thawed, and subsequently freeze-dried for 48 h at -40 °C using an ALPHA 1-2 LDplus Freeze Dryer following the manufacturer's protocol. Freeze-dried samples were then stored and transported at ambient temperature while frozen replicates were transported on dry ice with a commercial courier service to Sweden.

The animal captures and sample collection protocol used in this study was approved by the animal ethics committee of the University of Pretoria (EC050-16) and by Northern Cape Nature Conservation (ODB 1859-2016).

Library preparation and sequencing

The 20 samples were randomly placed on three 96-well plates together with other fecal samples from Damaraland mole-rats (as part of a larger microbiome study). Freeze-dried samples had been stored at room temperature (4 to 10 months) while frozen samples were thawed shortly before extraction. Each plate included four negative controls without a fecal sample and one mock community standard (25 μl ZymoBIOMICS Microbial Community DNA Standard). DNA from samples were extracted inside a UV-hood using the DNeasy PowerSoil Pro Kit (Qiagen) following the manufacturer's protocol. DNA concentration was quantified using a NanoDropTM1000 spectrophotometer (Thermo Fisher Scientific), and 2.5 ng extracted DNA from each sample was amplified in an initial PCR using the primer 341F (5'-CCTACGGGNGGCWGCAG-3') and 805R (5'-GACTACHVGGGTATCTAATCC-3') targeting the V3-V4 region of the 16S rRNA gene ([Herlemann et al., 2011](#); [Hugerth et al. 2014](#)). The primers included adapter sequences for Illumina n5/n7 index primers used in a second PCR. DNA samples were amplified in 25 μl reactions containing 0.5 μM of each primer and 12.5 μl Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific). Thermal PCR cycling conditions used were as follows: 30 s at 98 °C, followed by 20 cycles consisting of 10 s at 98 °C, 30 s at 58 °C and 15 s at 72 °C, and a final 2 min elongation step at 72 °C. The PCR products were purified using AMPure XP magnetic beads (Becker Coulter, USA) and used as template in the second PCR where each sample within a sequencing plate was amplified with a unique combination of Illumina n5/n7 index primers. The 50 μl reaction mix contained 23 μl purified DNA from the first PCR, 0.2 μM index primers and 25 μl Phusion High-Fidelity PCR Master Mix. For the PCR, the following thermal cycling steps were used: 30 s at 98 °C, followed by 12 cycles consisting of 10 s at 98 °C, 30 s at 62 °C and 5 s at 72 °C, and a final 2 min elongation at 72 °C. After the second PCR and another round of purification, DNA concentrations were measured using a Qubit fluorometer (Thermo Fisher Scientific), and equimolar amounts of each sample library from individual sample plates combined into pools with a final concentration of 4 ng/ μl . The pools were 300-bp paired end sequenced following standard Illumina sequencing protocols on an Illumina MiSeq platform at the Swedish National Genomics Infrastructure (NGI) at SciLifeLab in Uppsala, Sweden.

Bioinformatics and sequencing filtering

The quality of the reads was checked with FastQC v0.11.8 and MultiQC v1.9 (Andrews, 2010; Ewels et al., 2016). Raw reads from FastQ inputs were processed using the Ampliseq workflow v1.2.0dev (<https://mf-co.re/ampliseq/1.2.0>, Straub et al., 2020) which uses Cutadapt v.2.8 (Marin, 2011) to identify sequences with primers, remove sequences without primers and delete primers. Sequences passing the primer identification are denoised with the QIIME2 v2019.10.0 (Bolyen et al., 2019) implementation of DADA2 v.1.10.0 (Callahan et al., 2016) and Amplicon sequence variants (ASVs) are created with taxonomy assigned using the SILVA database v.132 (Quast et al., 2013) and QIIME2's Bayesian classifier. We used the default parameters, besides specifying our own primer sequences and trimming lengths (259 for forward, and 199 for reverse reads) so that sequences were trimmed to equal lengths before the actual denoising, as suggested by the DADA2 protocol. A phylogenetic tree of the ASV sequences was estimated using SEPP on the Greengenes 16S rRNA gene reference tree (McDonald et al., 2012; Mirarab, Nguyen & Warnow, 2012; Janssen et al., 2018).

Quality check and filtering of NGS data

Analyses post Ampliseq were conducted in R version 4.0.1, using mainly functions within the tidyverse, phyloseq and vegan packages (McMurdie & Holmes, 2013; Wickham et al., 2019; Oksanen et al., 2020), and figures were created with ggplot2 (Wickham & Chang, 2016). We combined reads of samples on plates that had been sequenced twice (plate 2 and 3) and removed all contaminant ASVs identified as more prevalent in the negative control samples than true fecal samples by the prevalence method in the decontam package v1.8.0, with a threshold of 0.5 and plate number as batch argument (Davis et al., 2018).

Statistical analysis

ASV richness, Shannon index and Faith's phylogenetic diversity (PD) were calculated on subsampled ASV counts rarefied to the minimum library size (39323 reads) with phyloseq (McMurdie & Holmes, 2013) which removed 144 ASVs. The effect of sample preservation treatment on library sizes was analysed using a linear mixed model, fitting library size as response variable, including the sample preservation treatment as fixed factor and sample identity and plate number as random factors (Bates et al., 2015). Subsequently, we tested if any of the alpha diversities were associated with sample library size using Pearson correlations. For those alpha diversity measures with a significant association between library size and alpha diversity (ASV richness and PD), we controlled for the effects of library size by calculating the residual alpha diversity from a linear regression between alpha diversity and library size. We assessed the effect of sample preservation treatment on (residual) alpha diversity measures fitting a linear mixed model with (residual) alpha diversity as response variable, the sample preservation treatment as fixed factor and sample identity and plate number as random factors. All mixed models were fitted using the R package lme4 (Bates et al., 2015). Finally, we tested Pearson correlations between alpha diversity measures from sample replicates.

To investigate beta diversity of samples, we performed two principal component analyses (PCA) with the prcomp function on two phylogeny-independent Euclidean

distance matrices based on two different transformation methods of raw counts of ASVs: Hellinger transformation (Rao, 1995) using the `decostand` function in the `vegan` package (Oksanen et al., 2020) and centered log-ratio (CLR) transformation (Aitchison, 1982) using the `clr` function in the `compositions` package. These methods deal with skewed abundance distribution within amplicon microbiome data in different ways. The Hellinger transformation reduces the influence of uncommon ASVs and weighs heavier on the more common ASVs (McMurdie & Holmes, 2013), while the CLR weighs heavier on the rare ASVs (Aitchison, 1982). Additionally, we calculated weighted and unweighted UniFrac distances (Lozupone & Knight, 2005), based on a phylogeny built with SEPP (Janssen et al., 2018). Non-metric multidimensional scaling (NMDS) using the `ordinate`-function in the `phyloseq`-package was applied on UniFrac distances (McMurdie & Holmes, 2013). One sample pair which was dominated by the family *Enterobacteriaceae* (sample 4, see Fig. S1) strongly reduced distances between the other samples in the NMDS (see Fig. S2) and was excluded for further analysis and visualizations.

The amount of variation explained by preservation treatment and sample identity for each beta diversity measure was tested with a Permutational Multivariate Analyses of Variance (PERMANOVA) with the `adonis2`-function from the `vegan` package with sample preservation treatment, sample library size and sample identity as fixed factor, by-argument set as “margin”, and testing with and without plate number as the strata argument to control for variation between sequencing plates. Likewise for the NMDS on phylogenetic beta diversity measures, we excluded the outlier sample pair (sample 4, see Fig. S1) in the PERMANOVAs to resemble that of the visualizations of the NMDS. To evaluate multivariate homogeneity of group dispersion of sample preservation treatment and sequencing plates we ran beta disperser tests on each of the beta diversity measures.

Correlations of relative abundances of ASVs within sample pairs were analysed with Pearson correlations across all ASVs and separated by the 8 most common phyla. Lastly, we tested for differences in relative abundances of families between sample pairs, and ran paired Wilcoxon signed-rank tests on families with a mean relative abundance of >1% as relative abundances were non-normally distributed, and *p*-values were adjusted with Bonferroni correction for multiple testing.

RESULTS

Preservation treatment did not bias library size

We obtained a total of 5,010,719 raw sequence reads from our 40 samples containing 1914 unique ASVs. After removing 146 contaminant ASVs from our samples, we obtained a total of 3,626,584 sequence reads and 1768 unique ASVs. The mean number of reads per sample was 90665 (range freeze-dried = 47510 to 141211, range frozen = 39323 to 183409) and treatment types did not differ significantly in number of reads per sample (*p* = 0.63).

Preservation treatment did not bias alpha diversity

Although there was substantial between-sample variation, the preservation treatment did not significantly bias measures of alpha diversity. For both sample treatments, we found large variation in ASV richness (freeze-died: mean = 304.55, range = 202–405;

frozen: mean = 303.30, range = 202–402), Shannon index (freeze-dried: mean = 3.96, range = 3.20–4.55; frozen: mean = 3.97, range = 2.42–4.54) and Faith's PD (freeze-dried: mean = 38.24, range = 27.41–47.63; frozen: mean = 37.13, range = 24.86–47.44) among samples. Some of that variation was explained by library size (ASV richness: $R = 0.68$, $p = <0.001$; Shannon index: $R = 0.12$, $p = 0.48$; Faith's PD: $R = 0.6$, $p < 0.001$). However, after controlling for the effect of library size, the preservation treatment of the sample replicates did not significantly bias ASV richness, Shannon index and Faith's PD (Fig. 1A residual ASV richness LMM (Estimate \pm Std. Error): Intercept = 1.73 ± 8.09 , estimate Treatment = 5.05 ± 7.542 , $p = 0.09$; Fig. 1B Shannon index LMM: Intercept = 3.94 ± 0.10 , estimate Treatment = 0.004 ± 0.07 , $p = 0.96$; Fig. 1C residual Faith's PD LMM: Intercept = 0.69 ± 0.98 , estimate Treatment = -1.46 ± 0.87 , $p = 0.10$), and the correlations between replicate samples were positive (Fig. 1D residual ASV richness: $R = 0.64$, $p = 0.003$; Fig. 1E Shannon index $R = 0.7$, $p < 0.001$; Fig. 1F residual Faith's PD $R = 0.62$, $p < 0.001$).

Preservation treatment has minor effect on beta diversity

Compositional differences between samples of the preservation treatments were minor. PCAs of Euclidean distances on both Hellinger and CLR transformed counts of ASVs revealed no clear clustering of sample treatment (Fig. 2). Similarly, NMDS on phylogenetic distance measures did not reveal clustering by sample treatment (Fig. 3).

Consistent with the interpretation of the PCAs and the NMDS, a PERMANOVA analysis revealed that sample treatment explained only a small proportion of the variation among the samples (1.5–2.2%) whereas the sample identity was identified as the main source of variation (73.2–86.6%, Table 1). Library size was significant for Euclidean distances on CLR-transformed data and unweighted UniFrac, but only explained 1.8–3.1% (Table 1). The PERMANOVA analysis further confirmed that plate number did not significantly explain any of the variation in microbiome composition of any of the four distance measures (Table 1). Finally, we did not detect any significant differences between sample treatments with Beta dispersion tests ($p = 0.08$ – 0.94 , Table 1).

Preservation treatment has a minor effect on compositional differences

Out of the 1768 ASVs, 98.5%, 91.9% and 73.7% were assigned to a phylum, family or genus, respectively. The dominating phyla were *Bacteroidetes* (mean relative abundance 72.9%; range 31.4–88.9%) and *Firmicutes* (mean relative abundance 16.4%; range 4.9–47.8%). However, within one sample, one ASV of the family *Enterobacteriaceae* dominated the community composition in both sample treatments with as much as 43.7% and 58.6% of the relative abundance of the freeze-dried and frozen replicates, respectively (see sample 4, Fig. S1). The ASV was not unique to the sample replicates, but prevalent in five other samples but in much lower abundance. Among the families with $>1\%$ mean relative abundance, only three families within the phylum *Firmicutes*, *Christensenellaceae*, *Ruminococcaceae* and *Lachnospiraceae*, was significantly different in relative abundances between treatments and was underrepresented within freeze-dried samples (Fig. 4, adjusted $p < 0.001$, Table S1).

About half (54%) of the ASVs were shared between the two sample treatments (Fig. S3) and these reads summed to a total of 99.5% of the reads of the full dataset. The two

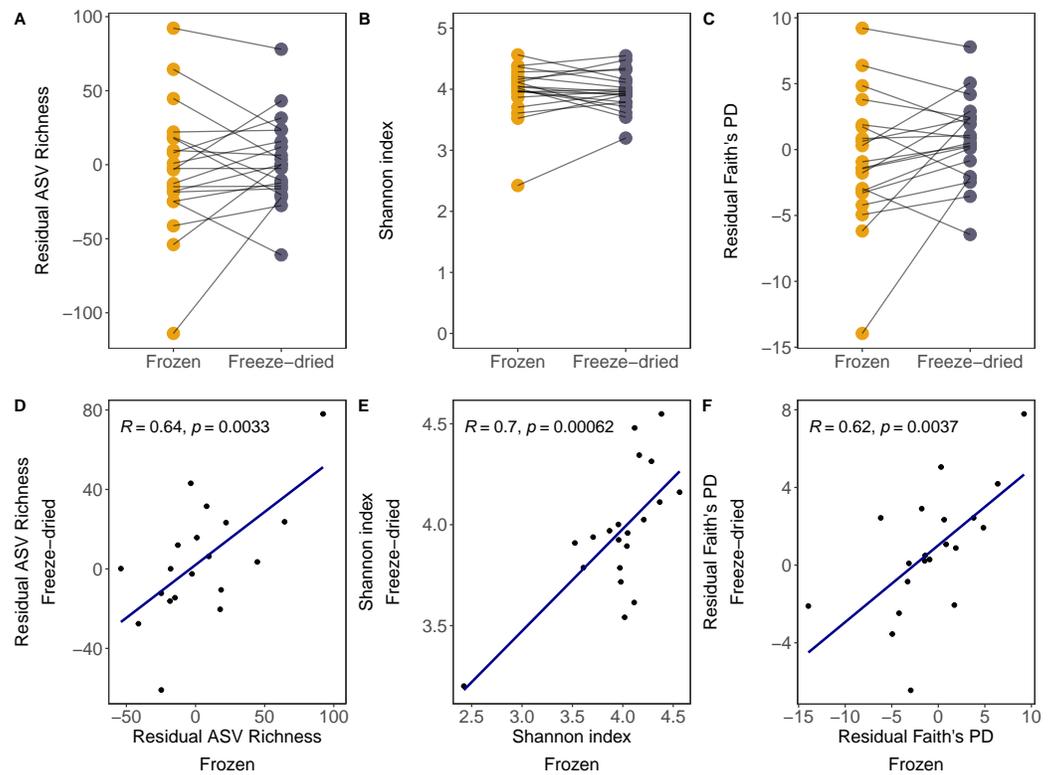


Figure 1 Alpha diversity measures of freeze-dried and frozen fecal samples. Comparison of (A) Residual ASV richness (number of ASVs), (B) Shannon index, and (C) residual Faith's phylogenetic diversity (PD) between sample replicates from different treatments. Lines between points in A–C connect sample replicates of the same original sample identity and point colour represent the two sample preservation treatments: freeze-dried (yellow); frozen (blue-grey). Correlations (Pearson's R) between alpha diversity measures from sample replicates of different treatments for (D) residual ASV richness, (E) Shannon index, (F) residual Faith's PD. When excluding the outlier in (E), the results did not change qualitatively ($R = 0.51$, $p = 0.027$).

Full-size DOI: 10.7717/peerj.13095/fig-1

preservation treatment types had very similar numbers of unique ASV, and the majority of these were also unique to a single sample (Fig. S3). Overall, we found a strong correlation between relative abundances of ASVs between the two types of treatments ($R = 0.87$, $p < 0.001$). When analysing the phyla separately, the more common phyla still showed strong correlations (Fig. 5). The phyla *Lentisphaerae* and *Spirochaetes* showed weaker trends between treatment types (Fig. 5), and *Lentisphaerae* was the only phylum that had overrepresentation of unique ASVs in one of the sample pairs (the freeze-dried treatment, Fig. 5).

DISCUSSION

Fecal samples from wild animal populations can yield important insights into the role of the gut microbiome for the fitness, health and survival of their hosts (Suzuki, 2017). However, the transport of samples in stable cold chains from remote field collection sites to laboratories can be challenging and ways of preserving fecal samples for microbiome

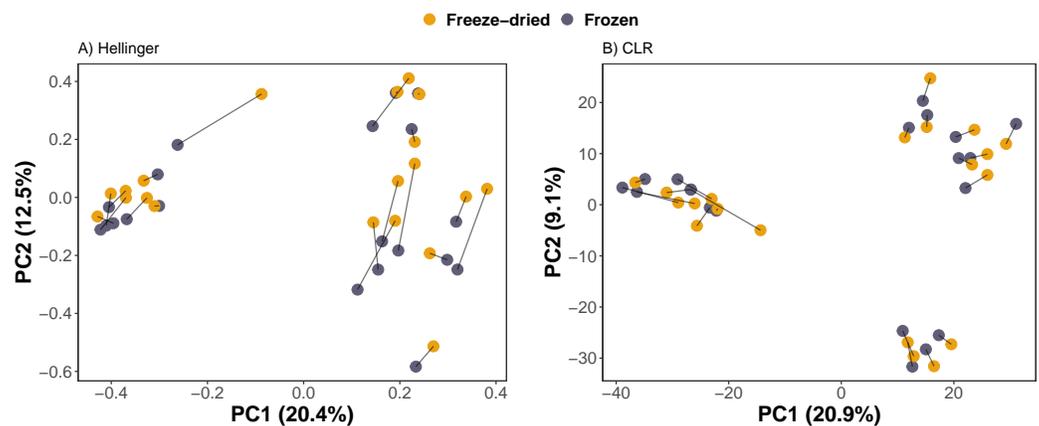


Figure 2 Clustering of freeze-dried and frozen fecal samples by Euclidean distance measures. Principal Component Analyses (PCA) of (A) Euclidean distances of Hellinger and (B) of centered log-ratio (CLR) transformed counts. Lines between points in A and B pair replicates of the same original sample and point colour represent the two sample preservation treatments: freeze-dried (yellow); frozen (blue-grey).

Full-size [DOI: 10.7717/peerj.13095/fig-2](https://doi.org/10.7717/peerj.13095/fig-2)

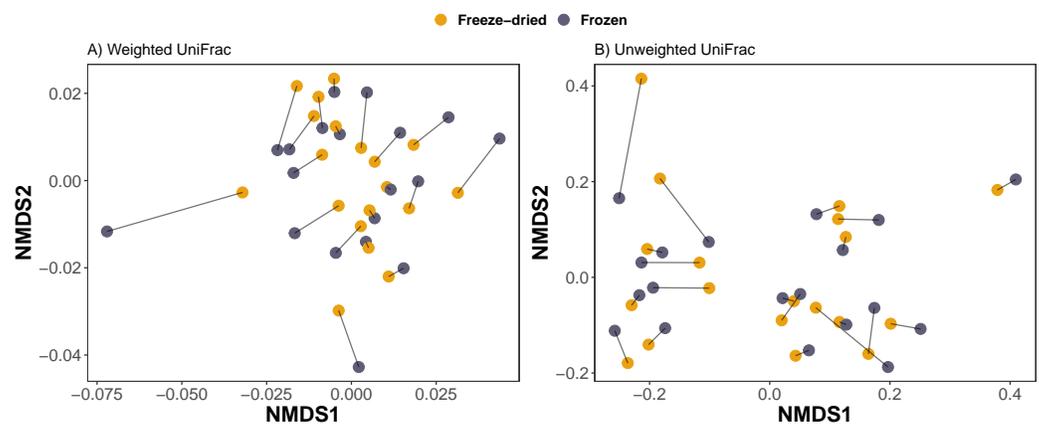


Figure 3 Clustering of freeze-dried and frozen fecal samples by phylogenetic distance measures. Non-metric multidimensional scaling (NMDS) on (A) weighted and (B) unweighted UniFrac distances. Stress weighted = 0.168, unweighted = 0.165. Outlier sample identity 4 was excluded, see Fig. S2 for NMDS on UniFrac distances including all 20 sample pairs. Lines between points in figure A and B pair replicates of the same original sample and point color represent the two sample preservation treatments: freeze-dried (yellow); frozen (blue-grey).

Full-size [DOI: 10.7717/peerj.13095/fig-3](https://doi.org/10.7717/peerj.13095/fig-3)

analyses without freezing or chemical preservatives would facilitate the development of this research field. In this study we show that frozen and freeze-dried samples do not differ in alpha diversity measures and the variation induced by the treatment amounted to less than 2.2% of the variation across the samples.

Alpha diversity measures did not differ between the two preservation methods, similar to other studies comparing alternative sample methods to deep-frozen microbiome samples (Song et al., 2016; Gavriiliuc et al., 2021). Because freeze-drying did not bias any of the three alpha diversity measures we investigated, contamination or shifts in the sample

Table 1 PERMANOVA results table. Models on dissimilarity matrices of Euclidean distances (of Hellinger or centered log-ratio (CLR) transformed counts) and phylogenetic distances (Weighted and Unweighted UniFrac). Models on UniFrac distance matrices excluding outlier sample 4. *p*-values < 0.05 are highlighted in bold.

Permanova	Dissimilarity matrix	Factor	F	R ²	<i>p</i> (PERMANOVA)	<i>p</i> (Beta dispersion)
1	Unweighted UniFrac	Library size	2.601	0.031	<0.001	
1	Unweighted UniFrac	Sample identity	3.437	0.732	<0.001	
1	Unweighted UniFrac	Treatment	1.549	0.018	0.03	0.937
2	Unweighted UniFrac	Plate number	1.221	0.065	0.101	0.117
3	Weighted UniFrac	Library size	0.813	0.004	0.536	
3	Weighted UniFrac	Sample identity	9.081	0.854	<0.001	
3	Weighted UniFrac	Treatment	4.166	0.022	<0.001	0.085
4	Weighted UniFrac	Plate number	0.854	0.047	0.652	0.117
5	Hellinger	Library size	1.347	0.006	0.139	
5	Hellinger	Sample identity	9.916	0.866	<0.001	
5	Hellinger	Treatment	4.737	0.022	<0.001	0.269
6	Hellinger	Plate number	0.922	0.047	0.563	0.269
7	CLR	Library size	2.272	0.018	0.007	
7	CLR	Sample identity	5.207	0.795	<0.001	
7	CLR	Treatment	1.926	0.015	0.021	0.563
8	CLR	Plate number	1.128	0.057	0.235	<0.001

community composition during the additional freeze-drying procedure are likely rare. Compared to freeze-drying, some other preservation methods that successfully eliminate the necessity of freezing sometimes come with the drawback of introducing the risk of altering sample alpha diversity through either contamination by additional bacterial exposure or change how easily different taxa are extracted. For example, *Song et al. (2016)* found that FTA-cards had increased Shannon diversity compared to other sample methods and fresh samples. Furthermore, the library sizes in our study did not differ systematically between the treatment types (Fig. S4) which suggests that rare ASVs are sequenced with similar probabilities (*Wu et al., 2010*), and both methods perform equally using the same DNA-extraction and library preparation protocol. Our data suggests that none of these possible biases were detectable in the comparison between frozen and freeze-dried samples.

For all beta diversity measures, sample preservation treatment explained a minor proportion of the variation (1.5–2.1%) within the microbiome community composition whereas sample identity explained a much larger part of the variation (73.2–86.6%). Sample identity explained more of the variation in beta diversity calculated on Hellinger-transformed than CLR-transformed data. Hellinger transformation puts weight on abundant ASVs (*Legendre & Legendre, 2012*) and might hence be less sensitive to random sampling of rare ASVs which might explain the higher impact of sample identity for this transformation. Alternatively, the CLR-transformed data was better at picking up differences between pellets from the same animal as each replicate tube from an animal contained one to two pellets and replicate samples did not originate from a homogenized

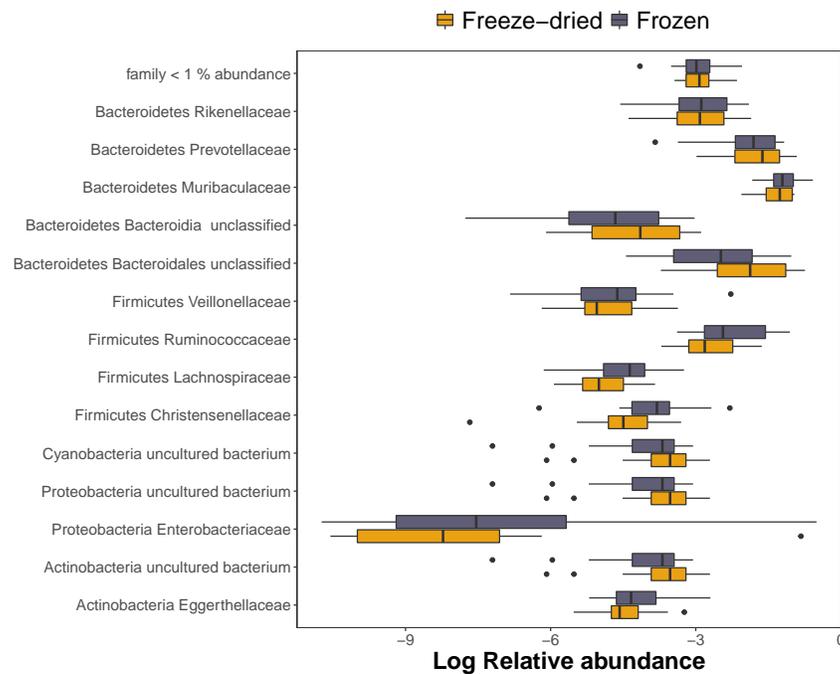


Figure 4 Log-transformed relative abundances of families with > 1% mean relative abundance. Amplicon Sequence Variants (ASVs) belonging to other families are combined within “family < 1% abundance”. Colour represent the two sample preservation treatments: freeze-dried (yellow); frozen (blue-grey).

Full-size DOI: [10.7717/peerj.13095/fig-4](https://doi.org/10.7717/peerj.13095/fig-4)

mix of pellets. While the variation explained by treatment was still small compared to the sample identity, it suggests that taxonomic groups are most often prevalent in both sample pairs but that some taxonomic groups vary slightly in abundance between treatments (see further discussion below). Sample library size explained a similar proportion of the variation as sample treatment, and because sample treatment did not bias library sizes, this suggest that the amount of variation caused by sample treatment was similar to random variation in library sizes between samples.

The relative abundances of ASVs were overall strongly correlated between the two treatments, suggesting that freeze-died samples reliably reflect the prevalence and abundance of bacterial populations in deep-frozen samples. Notably, one sample pair was dominated by the same single ASV in both sample treatments, showing that both treatments succeeded at capturing “extremes”. Given that this ASV had unusually high abundance within both treatments of the sample and was prevalent in other samples too, it was likely a truly dominating ASV within the gut microbiome of the individual and not a contamination. Furthermore, both treatments had similar proportions of sample-pair unique ASV abundances of the more common phyla and these were in general lower in relative abundance than ASVs shared within sample pairs, demonstrating that both treatments do equally well with sequencing less dominant ASVs. The phylum with the lowest correlation was *Lentisphaerae*, in which ASVs unique to one of the samples were overrepresented within freeze-dried samples (Fig. 4). Although correlations among

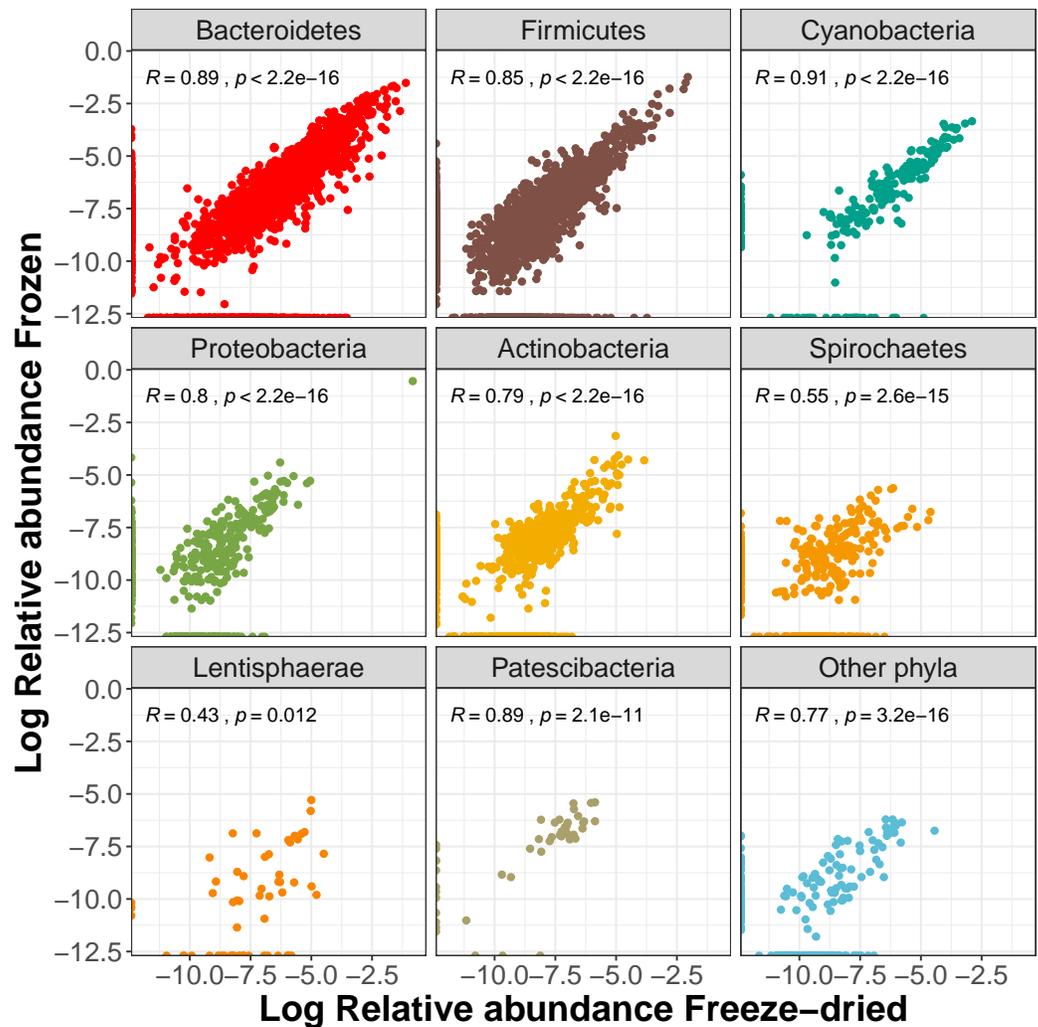


Figure 5 Correlations of relative Amplicon Sequence Variant (ASV) abundances between the two sample treatments for the eight most common phyla (freeze-dried or stored in minus 80 freezer). The R value is the Pearson correlation across ASVs within each phylum.

Full-size DOI: [10.7717/peerj.13095/fig-5](https://doi.org/10.7717/peerj.13095/fig-5)

ASVs were in general strong within the dominating phyla, we found that among families with >1% mean relative abundance, three families within the second most common phylum *Firmicutes* differed significantly between treatments. It is likely that differences in abundance of these taxa were responsible for the small effect of treatment on variation in microbiome community structure. Given that this was still small compared to the effect of sample identity in all four beta diversity measures, our conclusion is that freeze-drying has only minor influences on the composition and reflect the overall diversity of the fecal microbiome well. It is however important to note that fecal microbiome samples are widely known to slightly change in composition with time and sample preservation method and the importance of not mixing different preservation methods to avoid batch effects is widely acknowledged (Song et al., 2016). Because our results suggests that there may be

a minor but significant bias in the community composition of freeze dried samples of a magnitude of 1–2% percent of the variation, we suggest that researcher considering mixing both frozen and freeze dried samples should consider whether the benefits of mixing outweigh the costs of introducing this small bias.

CONCLUSION

Our results show that freeze-drying can be a suitable method when microbiome samples have to be exported over long-distances and stable cold chains are not available. Alpha diversity measures were not biased between frozen and freeze-dried samples and the sample preservation treatment had only a minor influence on the community composition whereas sample identity explained a large proportion of the variance. Together, our findings suggest that already collected samples stored in freezers can be thawed and freeze-dried to be transported and stored at ambient temperatures prior to 16S microbiome analyses.

ACKNOWLEDGEMENTS

We are grateful to Tim Clutton-Brock for his support and advice during this project. We thank Marta Manser and the Kalahari Research Trust for their tireless effort organizing the research site and we are grateful to Walter Jubber and Dave Gaynor for managing the research site. We thank Yannick Francioli, Andreas Friedrichs, Megha Rao, Johanna Sunde and Daniël Wille for help with the fieldwork. We thank the Mammals Research Institute of the University of Pretoria for their support of the research in the Kalahari.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was supported by grants from Vetenskapsrådet (2017-05296) and Crafoordska Stiftelsen (2018-2259 & 2020-0976). The facilities of the of the mole-rat project and the long-term study are supported by European Research Council grants 294494 and 742808 to T. H. Clutton-Brock. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Vetenskapsrådet (2017-05296).
Crafoordska Stiftelsen (2018-2259 & 2020-0976).

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Hanna M. Bensch conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

- Conny Tolf performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Jonas Waldenström conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Daniel Lundin analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Markus Zöttl conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Animal Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The animal ethics committee of the University of Pretoria (EC050-16).

Data Availability

The following information was supplied regarding data availability:

The R-scripts, asv-table and metadata are available at GitHub: <https://github.com/HannaBensch/FreezeDriedVSFrozen>.

The raw 16S sequences are available at NCBI short read archive (SRA): [PRJNA781121](https://www.ncbi.nlm.nih.gov/sra/PRJNA781121).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.13095#supplemental-information>.

REFERENCES

- Aitchison J. 1982.** The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological)* **44**:139–160
DOI [10.1111/j.2517-6161.1982.tb01195.x](https://doi.org/10.1111/j.2517-6161.1982.tb01195.x).
- Andrews S. 2010.** *FastQC: a quality control tool for high throughput sequence data.* Available at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Bates D, Mächler M, Bolker B, Walker S. 2015.** Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**:1–48 DOI [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).
- Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Bin KK, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciulek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D,**

- Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, Van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, Von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* 37:852–857 DOI 10.1038/s41587-019-0209-9.
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods* 13:581–583 DOI 10.1038/nmeth.3869.
- Cardona S, Eck A, Cassellas M, Gallart M, Alastrue C, Dore J, Azpiroz F, Roca J, Guarner F, Manichanh C. 2012. Storage conditions of intestinal microbiota matter in metagenomic analysis. *BMC Microbiology* 12:158 DOI 10.1186/1471-2180-12-158.
- Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6:1–14 DOI 10.1186/s40168-017-0383-2.
- Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32:3047–3048 DOI 10.1093/bioinformatics/btw354.
- Gavriliuc S, Stothart MR, Henry A, Poissant J. 2021. Long-term storage of feces at –80 C versus –20 C is negligible for 16S rRNA amplicon profiling of the equine bacterial microbiome. *PeerJ* 9:e10837 DOI 10.7717/peerj.10837.
- Hale VL, Tan CL, Knight R, Amato KR. 2015. Effect of preservation method on spider monkey (*Ateles geoffroyi*) fecal microbiota over 8 weeks. *Journal of Microbiological Methods* 113:16–26 DOI 10.1016/j.mimet.2015.03.021.
- Herlemann DPR, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF. 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME Journal* 5:1571–1579 DOI 10.1038/ismej.2011.41.
- Hugerth LW, Wefer HA, Lundin S, Jakobsson HE, Lindberg M, Rodin S, Engstrand L, Andersson AF. 2014. DegePrime, a program for degenerate primer design for broad-taxonomic-range PCR in microbial ecology studies. *Applied and Environmental Microbiology* 80:5116–5123 DOI 10.1128/AEM.01403-14.
- Janssen S, McDonald D, Gonzalez A, Navas-molina JA, Jiang L, Xu Z. 2018. Phylogenetic placement of exact amplicon sequences. *MSystems* 3:e00021-18 DOI 10.1128/mSystems.00021-18.
- Kia E, MacKenzie BW, Middleton D, Lau A, Waite DW, Lewis G, Chan YK, Silvestre M, Cooper GJS, Poppitt SD, Taylor MW. 2016. Integrity of the human faecal microbiota following long-term sample storage. *PLOS ONE* 11:1–8 DOI 10.1371/journal.pone.0163666.
- Kim D, Hofstaedter CE, Zhao C, Mattei L, Tanes C, Clarke E, Lauder A, Sherrill-Mix S, Chehoud C, Kelsen J, Conrad M, Collman RG, Baldassano R, Bushman FD,

- Bittinger K.** 2017. Optimizing methods and dodging pitfalls in microbiome research. *Microbiome* 5:52 DOI 10.1186/s40168-017-0267-5.
- Legendre P, Legendre L.** 2012. Ecological resemblance. In: *Numerical ecology*. Amsterdam: Elsevier, 265–335 DOI 10.1016/B978-0-444-53868-0.50007-1.
- Lozupone C, Knight R.** 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Environmental Microbiology* 71:8228–8235 DOI 10.1128/AEM.71.12.8228-8235.2005.
- Marin M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* 17:10–12 DOI 10.14806/ej.17.1.200.
- McDonald D, Price MN, Goodrich J, Nawrocki EP, Desantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P.** 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal* 6:610–618 DOI 10.1038/ismej.2011.139.
- McMurdie PJ, Holmes S.** 2013. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE* 8:e61217 DOI 10.1371/journal.pone.0061217.
- Mirarab S, Nguyen N, Warnow T.** 2012. SEPP: SATé-enabled phylogenetic placement. *Pacific Symposium on Biocomputing* 24:7–258 DOI 10.1142/9789814366496_0024.
- Oksanen AJ, Blanchet FG, Friendly M, Kindt R, Legendre P, Mcglinn D, Minchin PR, Hara RBO, Simpson GL, Solymos P, Stevens MHH, Szoecs E.** 2020. vegan: community ecology package. Available at <https://cran.r-project.org/package=vegan>.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO.** 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41:D590–D596 DOI 10.1093/nar/gks1219.
- Rao C.** 1995. A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance. *Questiò: Quaderns d'Estadística, Sistemes, Informàtica I Investigació Operativa* 19:23–63.
- Shen W, Wu D, Qiu W, Yi X.** 2021. Evaluation of freeze-drying for quantification of the microbiome and metabolome in neonatal faecal samples. *Medicine in Microecology* 8:100044 DOI 10.1016/j.medmic.2021.100044.
- Song SJ, Amir A, Metcalf JL, Amato KR, Xu ZZ, Humphrey G, Knight R.** 2016. Preservation methods differ in fecal microbiome stability. *MSystems* 1:e00021-16 DOI 10.1128/mSystems.00021-16.
- Straub D, Blackwell N, Langarica-Fuentes A, Peltzer A, Nahnsen S, Kleindienst S.** 2020. Interpretations of environmental microbial community studies are biased by the selected 16S rRNA (gene) amplicon sequencing pipeline. *Frontiers in Microbiology* 11:1–18 DOI 10.3389/fmicb.2020.550420.
- Suzuki TA.** 2017. Links between natural variation in the microbiome and host fitness in wild mammals. *Integrative and Comparative Biology* 57:756–769 DOI 10.1093/icb/icx104.

- Thorley J, Bensch HM, Finn K, Clutton-Brock TH, Zöttl M. 2021.** Fitness of breeders in social Damaraland mole-rats is independent of group size. *BioRxiv*. DOI [10.1101/2021.12.08.471794](https://doi.org/10.1101/2021.12.08.471794).
- Vandeputte D, Tito RY, Vanleeuwen R, Falony G, Raes J. 2017.** Practical considerations for large-scale gut microbiome studies. *FEMS Microbiology Reviews* **41**:S154–S167 DOI [10.1093/femsre/fux027](https://doi.org/10.1093/femsre/fux027).
- Vogtmann E, Chen J, Amir A, Shi J, Abnet CC, Nelson H, Knight R, Chia N, Sinha R. 2017.** Comparison of collection methods for fecal samples in microbiome studies. *American Journal of Epidemiology* **185**:115–123 DOI [10.1093/aje/kww177](https://doi.org/10.1093/aje/kww177).
- Wang Z, Zolnik CP, Qiu Y, Usyk M, Wang T, Strickler HD, Isasi CR, Kaplan RC, Kurland IJ, Qi Q, Burk RD. 2018.** Comparison of fecal collection methods for microbiome and metabolomics studies. *Frontiers in Cellular and Infection Microbiology* **8**:1–10 DOI [10.3389/fcimb.2018.00301](https://doi.org/10.3389/fcimb.2018.00301).
- Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Golemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019.** Welcome to the Tidyverse. *Journal of Open Source Software* **4**:1686 DOI [10.21105/joss.01686](https://doi.org/10.21105/joss.01686).
- Wickham H, Chang W. 2016.** Create elegant data visualisations using the grammar of graphics.
- Wu JY, Jiang XT, Jiang YX, Lu SY, Zou F, Zhou HW. 2010.** Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. *BMC Microbiology* **10**:255 DOI [10.1186/1471-2180-10-255](https://doi.org/10.1186/1471-2180-10-255).
- Zöttl M, Vullioud P, Mendonça R, Torrents Ticó M, Gaynor D, Mitchell A, Clutton-Brock T. 2016.** Differences in cooperative behavior among Damaraland mole rats are consequences of an age-related polyethism. *Proceedings of the National Academy of Sciences of the United States of America* **113**:10382–10387 DOI [10.1073/pnas.1607885113](https://doi.org/10.1073/pnas.1607885113).