Application Note

Deep analysis of microbiome diversity in stabilized stool samples

The human gut is the compartment of the body that contains the most microorganisms, with up to 10^{11} bacteria in just one gram of stool. Since the gut microbiome is importantly related to disease and health, its composition is of great interest in applied and basic research. Various diseases of the nervous system, metabolism or even cancer can be associated with the gut microbiome. Thus, stool samples are the subject in multiple studies applying different throughput, for invividual analysis as well as cohorts. Due to the high content of DNases/RNases and bacterial growth, stool samples must be treated with care to maintain bacterial composition. Conventional opinion holds that freezing stool samples at -80°C is the gold standard for specimen management. However, it is necessary to freeze or at least refrigerate the samples within a very short time, which is impractical for studies with home sampling. It is also often difficult and costly to maintain sufficient cooling of the sample until extraction.

A cost-saving and easy-to-perform alternative is the use of a sample stabilizer. The stabilizer in the Invitek Stool Collection Tube with DNA Stabilizer has lysis activity so that bacteria are inactivated, and the microbial titer is preserved at the time point of collection. This ensures stabilization of the sample for up to three months at room temperature.

Fig. 1: Stool Collection Tube with DNA Stabilizer

To assess sample integrity with the Stool Collection Tube with DNA Stabilizer, a study was performed to investigate different sample storage methods, extraction kits and deep analysis methods. Stool samples frozen at -80°C were extracted after 21 days and 90 days, and samples collected with the Stool Collection Tube with DNA Stabilizer were stored at room temperature and extracted after 3 days, 21 days and 90 days. Control samples were extracted shortly after collection without further storage. Two different extraction methods were used, the PSP® Spin Stool DNA Basic Kit (Invitek) and the QIAmp DNA Stool Mini Kit (Qiagen). For the latter ones the extractions were done on a Qiacube by the University of Kiel. In total, 55 samples from 4 individual donors were subject to analysis as described below.

Freenzing at -80°C
Extraction after 21 days, 90 days

Invitek Stabilization, RT
Extraction after 3 days, 21 days, 90 days

INFORMATICS PROCESSING

To generate data for analysis, microbiome sequencing data was processed as follows: PCR primers and adaptors were removed, and data were supplied to Eagle Genomics for analysis using their 16S rRNA microbiome processing pipeline and analytical tools incorporated into the e[datascientist]TM platform. Following quality control, sequences were imported into QIIME2 (v2020.2) [1] and denoising was performed using DADA2 [2], truncating forward and reverse reads at positions 280 and 180, respectively. Taxonomic annotation of the resultant amplicon sequence variants (ASVs) was performed using a machine learning classifier trained on version 138 of the SSU subset of the SILVA database [3]. Taxonomic count data was generated by combining the ASV counts with the taxonomic annotations generated from the classifier step.

DATA ANALYSIS

Diversity analyses were performed to compare the microbial community structure between groups. Alpha diversity analyses were performed on both the ASV and genus-level count data. Data were subsampled to a common read depth prior to analysis. ASV count data was subsampled to 7,200 reads, retaining 50% of features and discarding 4 samples that had read depths below the threshold. Genus-level count data was subsampled to 3,337 reads, retaining 45% of features and excluding 3 samples below the threshold. Four different alpha diversity measures were used: Chao1, Simpson's and Shannon-Wiener indices, and observed features. Beta-diversity analyses (Jaccard, Bray-Curtis and Aitchison distance-based) were performed on both ASV and genus level count data. For Jaccard and Bray-Curtis based analyses, subsampled data was used, as described above. Aitchison-distance based analysis was performed using centre-log ratio transformed count data. Differential abundance analysis was performed on both the ASV and genus-level counts data using fastANCOM [4], which is an optimised implementation of the ANCOM [5] approach for analysis of microbiome data that has been shown to perform well in controlling false discovery rates while maintaining high power.

RESULTS

Alpha diversity analysis revealed significant differences (Kruskall-Wallis, $p < 0.01$) between community structure across all different diversity measures using both ASV and genus-level count data when comparing samples grouped by proband (Figure 2). No significant differences were found between samples grouped according to preparation reagent, storage time or storage temperature using either ASV or genus-level counts.

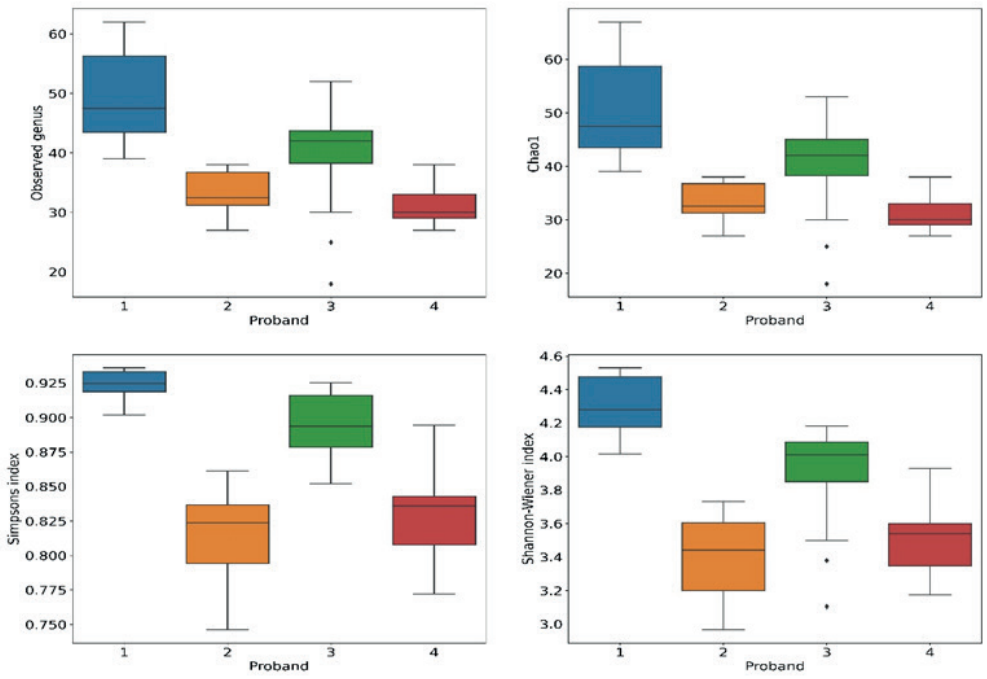

Fig. 2: Box and whisker plots showing differences in alpha diversity between samples grouped by proband. Results generated using genus-level count data and four different diversity measures are shown: observed genera (top left), Chao 1 Index (top right), Shannon-Wiener Index (bottom left) and Simpson's Index (bottom right).

Beta diversity analysis was performed to compare diversity between samples. As with the alpha diversity analysis results, significant differences (PERMANOVA, $p > 0.01$) were seen comparing samples from different probands according to all diversity measures and with both ASV and genus-level count data. No significant differences were observed comparing any other sample grouping. To illustrate differences in beta diversity, ordination plots showing PCA of centre-log ratio transformed count data are shown in Figure 3.

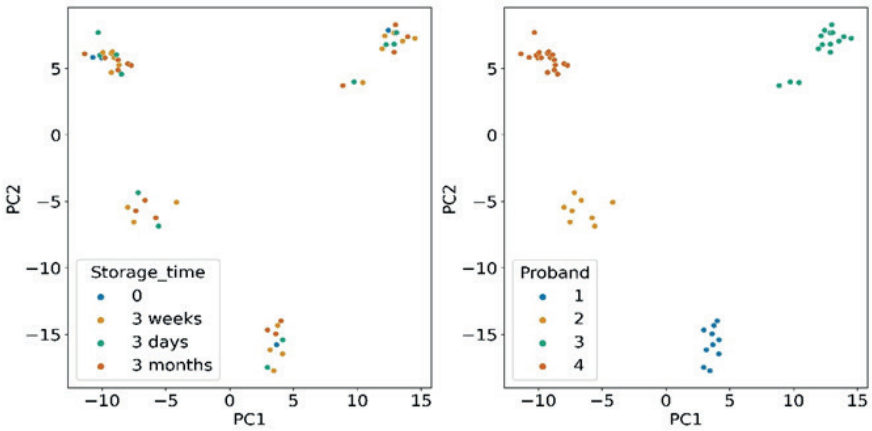

Fig. 3: Ordination plots showing PCA of centre-log ratio transformed genus-level taxonomic count data illustrating the effect of storage time (left) and proband (right).

To identify potential differences in the abundance of microbes between sample groups, fastANCOM was used to analyse the ASV and genus-level counts data. The effect of storage time and temperature was investigated, specifically samples stored at room temperature vs –80°C. Only one significantly differentially abundant genus (Defluviitaleaceae UCG-011) was observed between the different storage temperatures (see Figure 4). While this genus may warrant further investigation, the results indicate that the different storage factors generally do not affect microbial abundance in these samples.

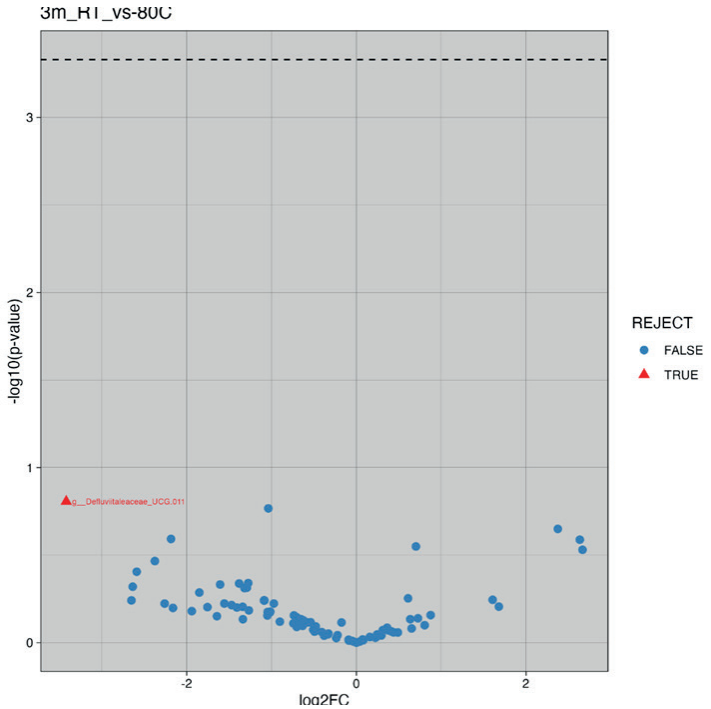

Fig. 4: Volcano plot showing differential abundance in genus level microbial count data from samples stored at RT vs –80°C after 3 months. The X-axis shows log fold changes between the two groups and the Y axis shows the –log10 adjusted p-value based on the output of fastANCOM. The legend shows whether the null hypothesis can be rejected for a specific taxon based on the fastANCOM analysis. Taxa can be considered differentially abundant where the rejection of the null hypothesis is shown as 'TRUE'.

CONCLUSIONS

Analysis of 16S NGS data showed that samples stored in the Stool Collection Tubes with DNA stabilizer are efficiently preserved over the whole 3-month timeframe analyzed in this study.

The different storage conditions, freezing at -80°C and stabilization at room temperature do not show any significant difference in bacterial composition or abundance, implying efficient stabilization of the samples due to storage in the DNA Stabilizer. Thus, sample stabilization in Stool Collection Tubes with DNA Stabilizer is a suitable and reliable alternative to freezing samples, which greatly simplifies sample handling in terms of sample collection, transport, storage and saving energy costs.

With the Stool Collection Tubes with DNA Stabilizer, samples have high integrity and analytical results are equivalent for all subjects at different time points and storage conditions. The results also indicate that these samples are extraction kit agnostic (including Qiagen* extraction kits).

The study was done based on NGS Data obtained by 16s rRNA Gene variable regions for bacterial composition typing. In perspective, it would be interesting to perform a shotgun analysis of the stool samples to learn more about species-level resolution and assessment of other taxa (e.g., viruses and fungi). In addition, it would be interesting to investigate whether the different storage conditions of samples, freezing at -80°C versus stabilization, results in a shift in microbial abundance. To investigate this, a larger test group and more data points at the different storage conditions would be required.

REFERENCES

1. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. (2019) 37(8):852-857. doi: 10.1038/s41587-019-0209-9.
2. Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods (2016) 13(7):581-3. doi: 10.1038/nmeth.3869.
3. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. (2013) 41 D590-6. doi: 10.1093/nar/gks1219.
4. Zhou C, Wang H, Zhao H and Wang T. fastANCOM: a fast method for analysis of compositions of microbiomes. Bioinformatics (2022) btac060. doi: 10.1093/bioinformatics/btac060.
5. Mandal S, Van Treuren W, White RA, et al. Analysis of composition of microbiomes: a novel method for studying microbial composition. Microb Ecol Health Dis. (2015) 26:27663. doi: 10.3402/mehd.v26.27663.

**** for further information on compatibility with other extraction kits you may reach out to techsupport@invitek.com***

INVITEK DIAGNOSTICS GERMANY
Invitek Molecular GmbH
Robert-Roessle-Str. 10
13125 Berlin, Germany

Phone: +49 30 9489 2908

info@invitek.com

INVITEK DIAGNOSTICS PORTUGAL
ALS Life Sciences, SA
Zona Industrial de Tondela, ZIM II, Lote 6
3460-070 Tondela, Portugal

Phone: +351 232 817 817

In Cooperation with Eagle Genomics

The Biodata Innovation Centre,
Wellcome Genome Campus
Hinxton, Cambridge,
CB10 1DR United Kingdom

Phone: +44 1223 654481
Email: contact@eaglegenomics.com