# Project bionfo report: MEF-24-01, 16S and ITS data - fermented product analyses

A. Zaiko

January 18, 2025

## Sample processing, sequencing and bionformatics outputs (Client - Memelio Fermentuotas)

Background

The Client Memelio fermentuotas has engaged SEQUENCH to provide the following services:

- Extract DNA from 48 samples (DNA extracts) from fermented products of 7 different manufacturers, sampled at min. 2 different time points.
- PCR amplifications of two markers: **16S** for general bacterial diversity and **ITS** for fungi diversity assessment in samples.
- High-throughput sequencing (HTS) library preparation (clean-up, indexing, QA/QC), HTS setup and analysis of the positive amplicons on an **Illumina sequencing platform**.
- QA of the HTS results and bioinformatic analysis of the derived HTS data using established protocol.

Below we provide a brief report on the outputs of the analysis for the both markers.

## Sample processing - DNA extraction and PCR amplifications

At the laboratory, samples were processed for DNA extractions and polymerase chain reactions (PCRs) to amplify COI and 16S rRNA gene regions. The quality of the PCR products was assessed using gel electrophresis. Following that, the PCR products were purified to remove unincorporated nucleotides and reagents, indexed, quantified and QA-ed, and processed for high-throughput sequencing on the NextSeq sequencing platform, using 2 × 300 base paired-end sequencing chemistry. All samples produced good quality amplification products and were successfully processed for sequencing.

Negative controls (molecular grade distilled water) was added to the library to control for potential lab-based contamination during the library preparation steps (DNA extraction, PCRs, clean-ups,pooling, etc.)

## Sequencing outputs

The sequencing run was completed without errors, meeting all specs criteria as defined by Illumina. The raw sequence files were bioinformatically processed to remove primers from the sequencing reads, filter, trim and de-noise, merge forward and reverse reads and assign the resulting unique sequences or amplicon sequence variants (ASVs) to a species by comparing the ASV to the reference taxonomy file.

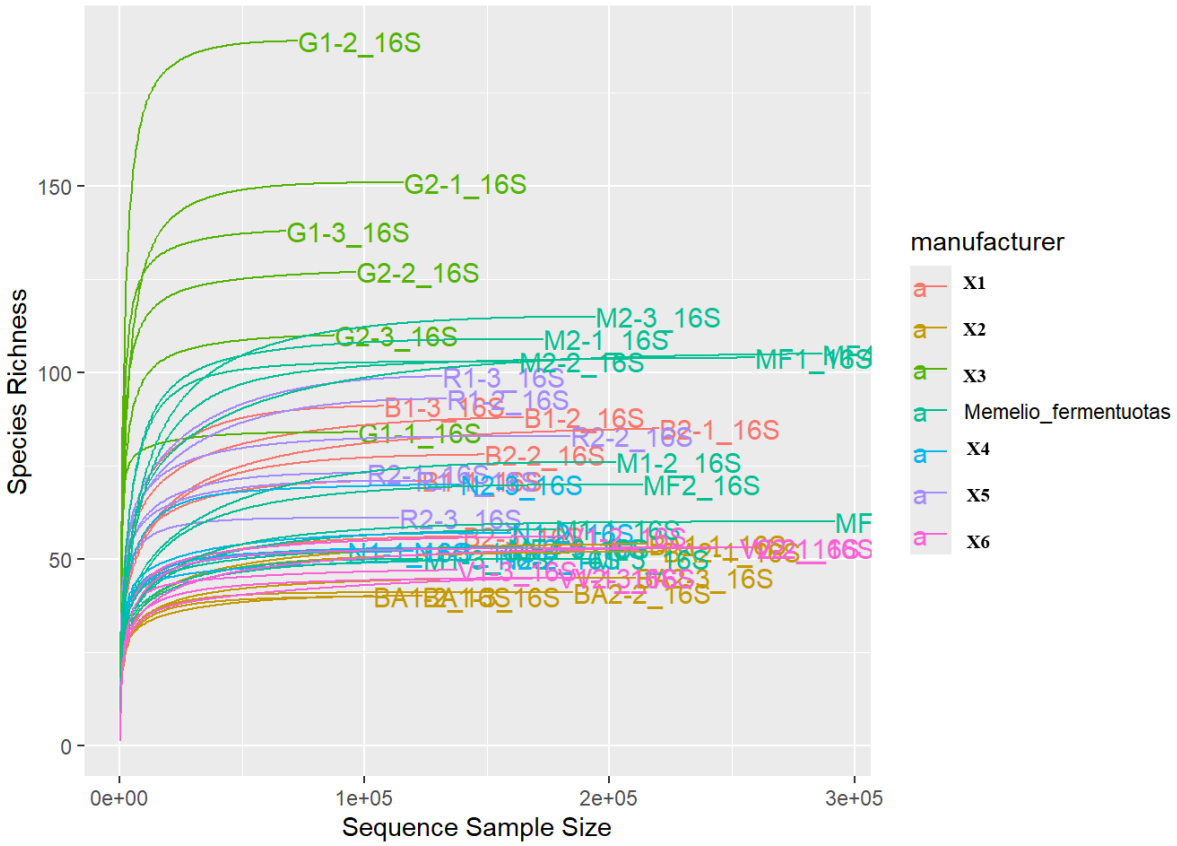
Below we present the results of data exploration for 16S (bacterial) marker used.

## 16S (Bacterial diversity)

The total number of sequenc reads retained in the dataset after all bioinformatical processing and denoising **7 756 626**, min = 679 820, average = 161 596 sequence reads per sample.

A rarefaction plot was created to look at how well diversity has been captured in samples. This graph confirms the adequate sequencing coverage across the samples: as all curves plateaued - this means that the majority of the targeted taxa diversity is captured well in the samples.

Note: Samples from **X3** manufacturer showed substantially higher diversity of detected species, as indicated by the rarefaction curves being higher along the Y axis.




Minor processing contamination was observed in blanks DNA extraction and PCR blanks. A decontamination step (proportional removal of sequences found in blanks) was undertaken, the outcome is:

- Number of ASVs totally removed: 0
- Percent of ASVs removed: 0 %
- Total number of reads removed: 0
- Percent of reads removed: 0 %

For follow-up data analyses and biodiversity comparison between samples, samples were rarefied to equal sequence number (67 900 sequences) to eliminate bias caused by differences in sequencing depth across samples. This data processing step allows to compare biodiversity across samples more fairly, no matter how much data each sample started with. This prevents one sample from looking more diverse just because it had more sequencing reads.

We also removed very rare ASVs (unique sequences) - those that are represented by less than 50 sequence reads across the entire dataset, as very rare bacterial sequences are likely to be

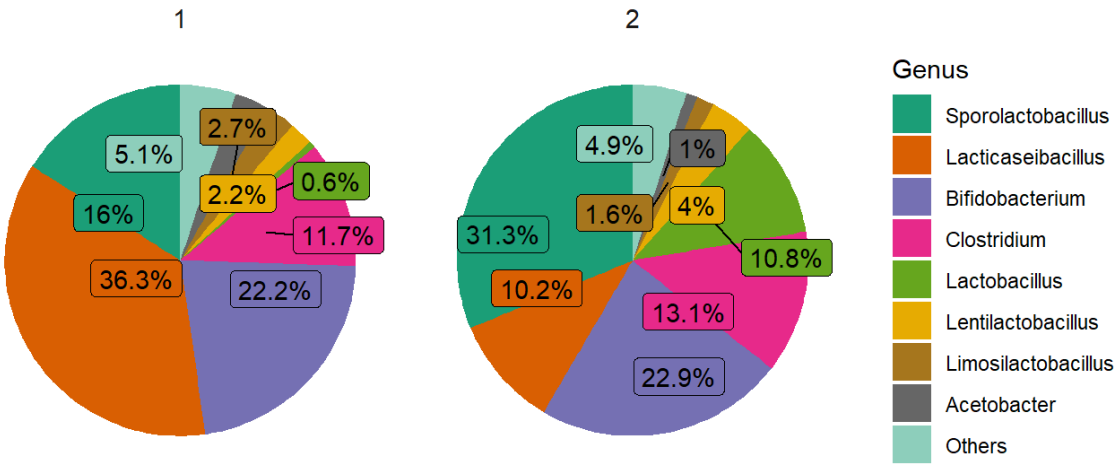
environmental noise or legacy signal from product handling and packaging, rather than inherent microbiome of the product.

Below we present the graphical summary of the core bacterial diversity (percentage abundance of 8 dominant genera visualized, with “Others” lumped into one group) the analyzed samples, presented by manufacturer and the time replicate.

**Note:** A total of 4 different batches of "Memelio fermentuotas" samples were analysed. The first two (1 and 2) are of standard composition and quality, sample 3 (kept at very warm temperatures for 1.5 months), where the cork popped upon opening, and sample 4 – made using a low-quality batch of oats.

**X1**

Substantial variation between time replicates observed. In both batches dominant bacterial genera were *Sporolactobacillus*, *Lacticaseibacillus*, *Bifidobacterium* and *Clostridium*, but in the second batch there was also substantial amount of *Lactobacilus*, and in the first one - *Bacteroides*. In these samples, 80 ASVs were detected overall, which were classified to 43 bacterial species.

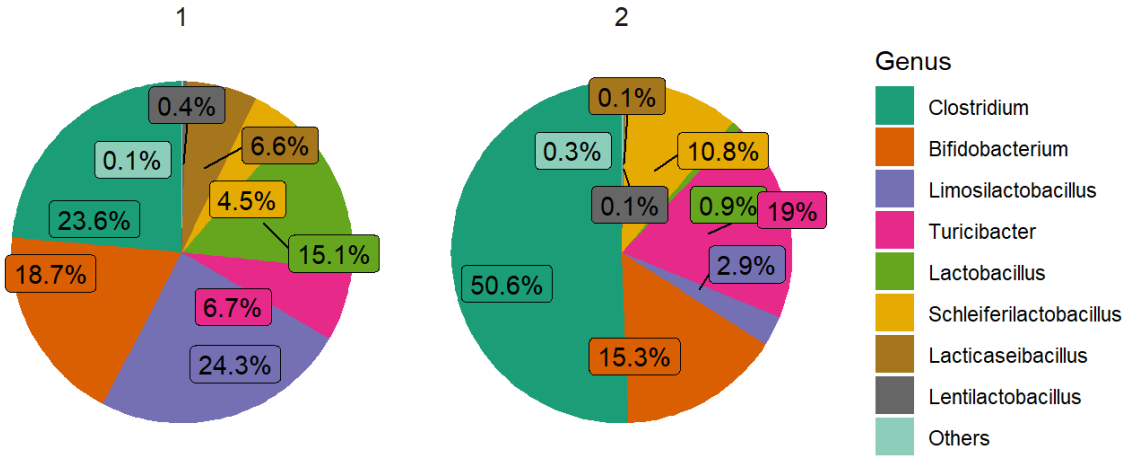


**X2**

Substantial variation between time replicates observed as well. High levels of *Clostridium* were deected in both batches. Other abundant genera
were *Bifidobacterium*, *Limosilactobacillus*, *Turicibacter*, *Lactobacillus* and *Schleiferilactobacillus*.

*Bifidobacterium* and *Clostridium*, but in the second batch there was also substantial amount of , and in the first one - *Bacteroides*.

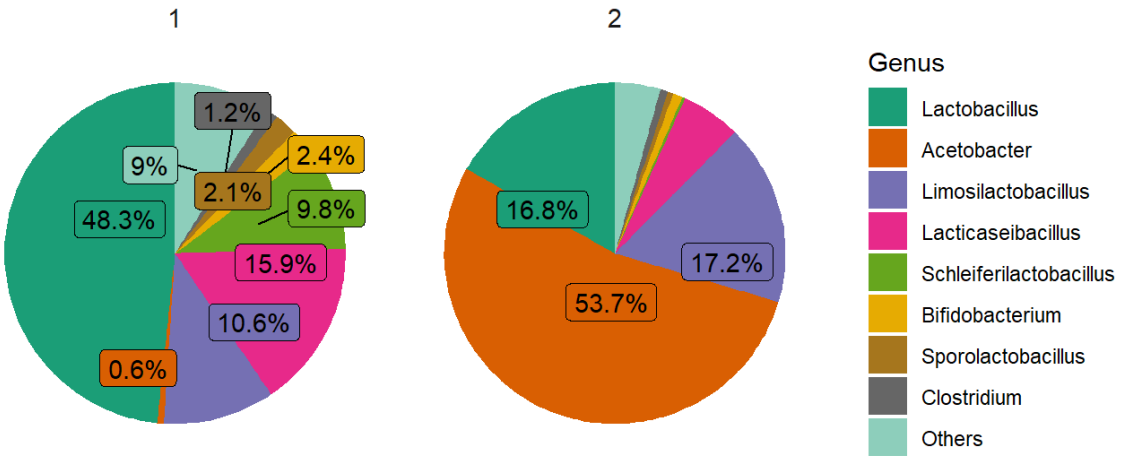
In these samples, 54 ASVs were detected overall, which were classified to 25 bacterial species.




**X3**

Substantial variation between time replicates observed as well. In the first batch, the microbiome was dominated by bacterial
genera *Lactobacilus*, *Lacticaseibacillus*, *Limosilactobacillus* and *Schleiferilactobacillus*. In the second batch, the most dominant genus was **Acetobacter**, followed
by *Limosilactobacillus* and *Lactobacillus*. *Bifidobacterium* and *Clostridium*, but in the second batch there was also substantial amount of , and in the first one - *Bacteroides*.

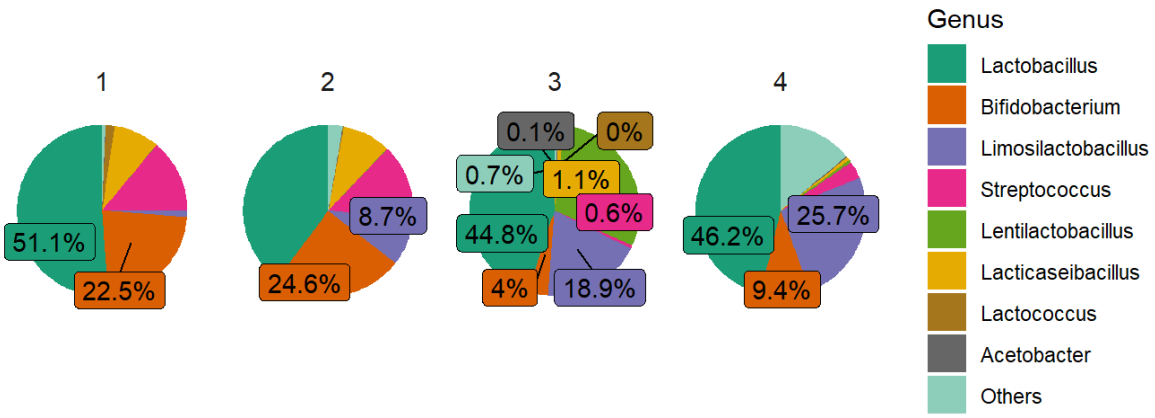
In these samples, 241 ASVs were detected overall, which were classified to 115 bacterial species.



**MEMELIO FERMENTUOTAS**

Variability between time replicates was observed as well. However batches 1 and 2 looked more consistent, while batches 3 & 4 differed substantially. The dominant genera in batches 1 & 2 were *Lactobacillus, Bifidobacterium, Limosilactobacillus,Streptococcus* and *Lacticaseibacillus*. In batch 3, lower levels of *Streptococcus* were detected. In batch 3, there was Lentilactobacillus present.

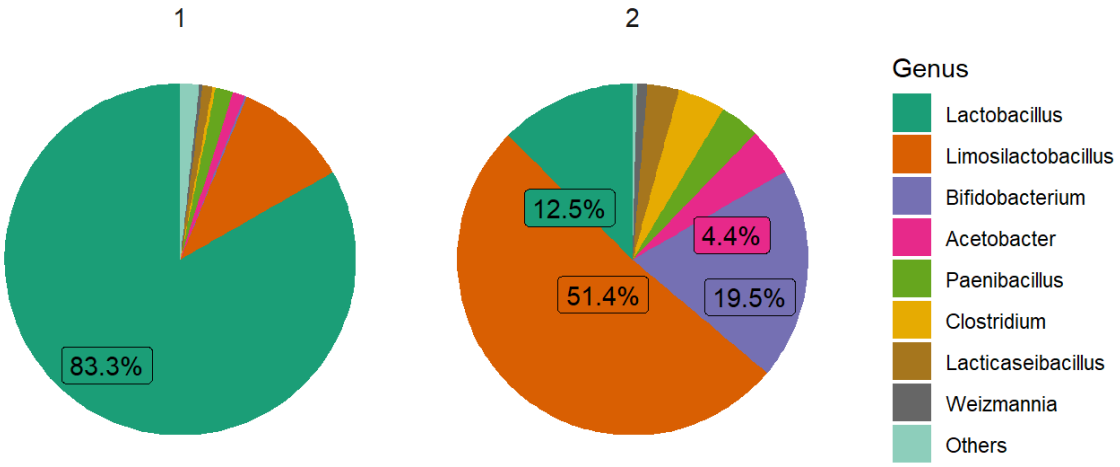
In Memelio Fermentuotas samples, 119 ASVs were detected overall, which were classified to 68 bacterial species.




**X4**

Again, bacterial diversity was substantially different between time replicates. The first batch was highly dominated by *Lactobacillus*, followed by *Limosilactobacillus*, while in the second batch *Limosilactobacillus* was the prevalent one with substantial contribution from *Lactobacillus* and *Bifidobakterium*. In the second batch there were also noticeable quantities of *Acetobacter*, *Paenibacillus*, *Clostridium* and *Lacticaseibacillus*.

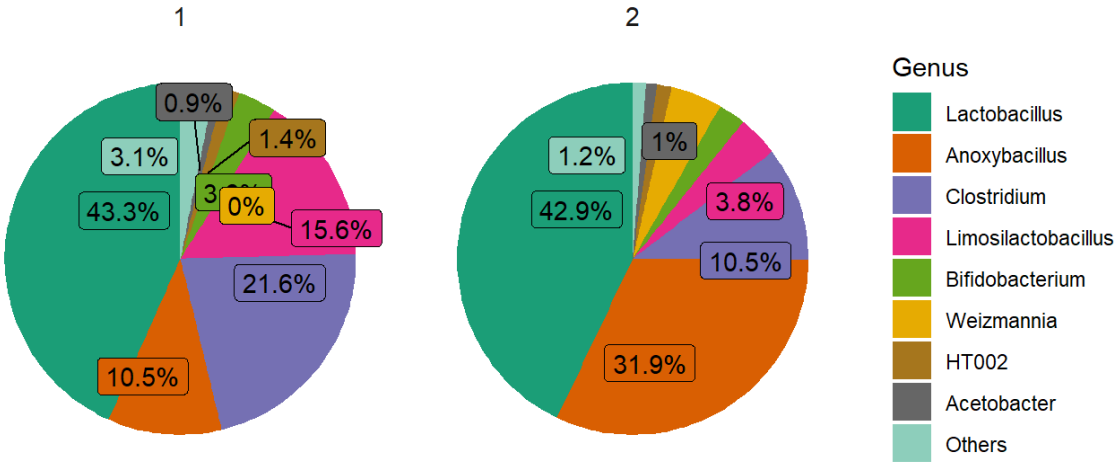
In these samples, 76 ASVs were detected overall, which were classified to 39 bacterial species.



**X5**

Variability in bacterial diversity was oberved between time replicates. The first batch was highly dominated by *Lactobacillus*, followed by *Clostridium*, *Anoxybacillus* and *Limosilactobacillus* with lower, but still substantial abundance of *Bifidobakterium*. In the second batch, *Lactobacillus* was sill the predominant genus, but *Anoxybacillus* was more abundant than in the first one, while *Limosilactobacillus* was less prevalent. Other important genera in the second batch were *Clostridium*, *Weizmannia* and *Bifidobakterium*.

In these samples, 95 ASVs were detected overall, which were classified to 47 bacterial species.




**X6**

Some variability between time replicates was observed. Most dominant bacteria was *Weizmannia*, followed by *Clostridium*, *Lactobacillus* and *Bifidobakterium*.

In these samples, 55 ASVs were detected overall, which were classified to 29 bacterial species.