Exercise: 16S rRNA gene sequencing – Quantative Analysis

Q1: In broad terms, what is the application of 16S rRNA gene amplicon sequencing analyses?

16S rRNA gene amplicon analysis is used to investigate microbial communities from various environments

Q2: Please describe the computational pipeline used to process 16S rRNA gene amplicon sequencing data until the creation of the OTU table and representative sequences.

- Merge paired reads (if applicable)
- Quality filter the reads
- Remove Chimeras
- Cluster reads on a 97% (or other) similarity level
- Remap the raw reads to the representative sequences from the clusters to asses the abundance of an OTU
- Create abundance OTU table
- Annotate the representative sequences to a database to extract the taxonomy
- Optional: Create a phylogenetic tree

Q3: Describe with your own words, what is alpha diversity?

Alpha diversity describes the richness and evenness of species in a community

Q4: Describe with your own words, what is beta diversity?

Beta diversity describes the similarity between environments

Q5: Explain with your own words the importance of considering sample sizes

Differences in sample sizes is a technical parameter which can affect the comparison between samples and in the end lead to false biological conclusions

Q6: Did we cover the whole microbial diversity in the sample, with the current sequencing depth? Please explain how you came to that conclusion.

The histogram shows unnatural high number of OTUs which are observed twice and indicates erroneous reads. These error reads will prevent the rarefaction curves from reaching a complete plateau. However the rarefaction curves display a nice trajectory towards a plateau indicating that most of the diversity has been covered.

Q7: Is there a difference in richness and alpha diversity between the two individuals?

Yes. ID1 has a higher richness and diversity compared to ID2. However, the Shannon Index displays a more stable measure due to the weighting of the higher abundant OTUs.

Q8: What is the effect of the different methods to account for differences in sample sizes?

Normalization does not affect richness compared to downsizing.

Q9: In the case where you want to compare samples, which normalization method would you choose and why?

The plots of the richness of ID1 and sample size display a dependency of sample size and richness, when normalization is applied. The richness of the rarefied samples shows less dependency, so I would choose rarefication.

Q10: How do the samples cluster? What does it tell you about the dynamics of the microbial community structure in the individuals?

Samples from the same person are clustering together regardless of the diet.

Q11: Are there any differences between the two individuals?

Yes, the individuals have unique gut microbiomes.