How To Navigate Your BiomCare Folder

January 9, 2020

Overview over your folders

Figure 1 shows an overview over all the folders in your root folder. The folders described in figure 1 will be described in more detail below. Each section will detail one main folder and its subfolders.
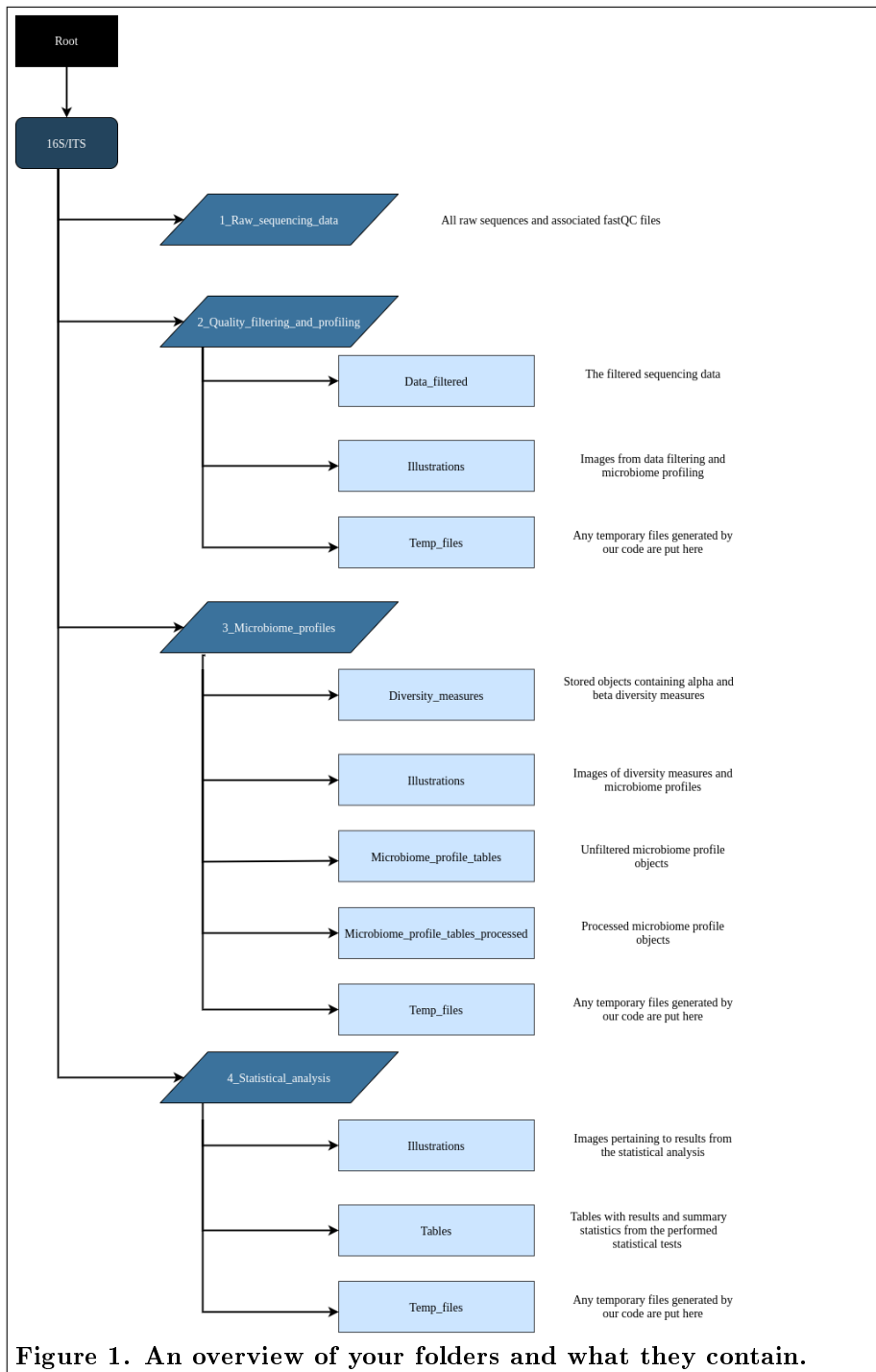
Figure 1. An overview of your folders and what they contain.

Root

16S/ITS

1_Raw_sequencing_data — All raw sequences and associated fastQC files

2_Quality_filtering_and_profiling

Data_filtered — The filtered sequencing data

Illustrations — Images from data filtering and microbiome profiling

Temp_files — Any temporary files generated by our code are put here

3_Microbiome_profiles

Diversity_measures — Stored objects containing alpha and beta diversity measures

Illustrations — Images of diversity measures and microbiome profiles

Microbiome_profile_tables — Unfiltered microbiome profile objects

Microbiome_profile_tables_processed — Processed microbiome profile objects

Temp_files — Any temporary files generated by our code are put here

4_Statistical_analysis

Illustrations — Images pertaining to results from the statistical analysis

Tables — Tables with results and summary statistics from the performed statistical tests

Temp_files — Any temporary files generated by our code are put here

16S / ITS

Data and results from each performed sequencing method has its own folder. Each folder hold all files and images from the performed sequencing, data processing and analysis. If you only ordered one type of sequencing run, e.g. only bacterial, you will only have a folder named 16S in your project folder.

1_Raw_sequencing_data

This is where your raw sequence data is stored. Each sample has eight parts, four parts for forward and four parts for reverse reads. The four files are:

1. a md5 checksum in order to verify file integrity

2. the sequencing data itself ending with fastq.gz

3. a zipped FastQC file containing various files on sequence quality

4. a FastQC report in html format.

2_Quality_filtering_and_profiling

Data_filtered

This folder contains the quality filtered sequencing data that is used in the downstream microbiome profiling and analysis.

Illustrations

This folder contains all images from data filtering and microbiome profiling. This includes the following:

1. Error plot from microbiome profiling (see report 1 for details).

2. Stacked barplot of the read counts that remains after quality filtering, and how many reads are removed at each filtering step. This is also discussed in Report 1.

3. Filter table. A numerical version of the filter plot, this table demonstrates how much is removed at each quality control step and also how much remains at the end of the process.

4. Quality profiles of the sequencing data. There are four plots in total: two before and two after quality filtering for forward and reverse reads. F indicates forward reads and R indicates reverse reads. Each sample has its own panel with read-position on the x-axis and base-quality score on the y-axis. The plotted lines show summary statistics for each position across the read length: green lines show the mean quality score, orange lines show the median quality, and the dashed orange lines show the 25th and 75th quality quantiles.

Temp_files

Any temporary files generated for use later are stored here.

3_Microbiome_profiles

Diversity_measures

Here file objects containing the calculated diversity measures are stored.

Illustrations

Here you will find plots of alpha diversity (illustrated across all sample) and beta diversity (illustrated in ordination plots as the distance between samples).

Microbiome_profile_tables

Here is your un-processed ASV table and a table over taxonomy annotations. These together describe the total microbiome community across all samples.

Microbiome_profile_tables_processed

Here you find your ASV table and taxa tables combined into a single RDS object. Each object is processed to different extents as indicated by the file name. For example, you might have a rarefyed (normalised) version, a version with calculated relative abundances (relab) or an object not normalized. Objects subject to filtering of samples e.g. due to low read count are annotated with 'SampleFilter'. However, these files may contain the same number of samples as unfiltered if no single samples failed quality evaluations. These are examples and your folder might contain objects that was subject to other processing steps that was deemed necessary.

Temp_files

Again, any temporary files for later use are stored here.

4_Statistical_analysis

Illustrations

Images pertaining results from the statistical analysis. These will be presented and discussed in Report 3.

Tables

Tables with results and summary statistics from the performed statistical tests. These statistics are often used as input for the illustrations of results, such as boxplots of differently abundant taxa based on a Wilcoxon signed-rank test comparing cases and controls.

Temp_files

Storage for any files generated and perhaps needed later.