Table of Contents

[ASSIGNMENT | Application of Bioinformatics in Metagenomics 2](#_Toc188186031)

[1. Download FASTQ files, which correspond to the given twelve SRA IDs 3](#_Toc188186032)

[2. Check quality of the source data using FastQC and MultiQC 3](#_Toc188186033)

[3. Import FASTQ files into QIIME2 file format, pre-process and QC the data with QIIME2 tools 3](#_Toc188186034)

[4. Make ASVs feature table using DADA2 algorithm 3](#_Toc188186035)

[5. Evaluate phylogeny of identified microbial sequences 3](#_Toc188186036)

[6. Make rarefaction plots, select an appropriate rarefaction threshold 3](#_Toc188186037)

[7. Evaluate and interpret Alpha and Beta diversity in the dataset 3](#_Toc188186038)

[8. Evaluate the taxonomy of the soil microbial communities in the studied grassland sites 3](#_Toc188186039)

# ASSIGNMENT | Application of Bioinformatics in Metagenomics

Matthew Spriggs

Applied Bioinformatics

ESTIMATED WORD COUNT (EXCLUDING FIGURES AND QUESTIONS):

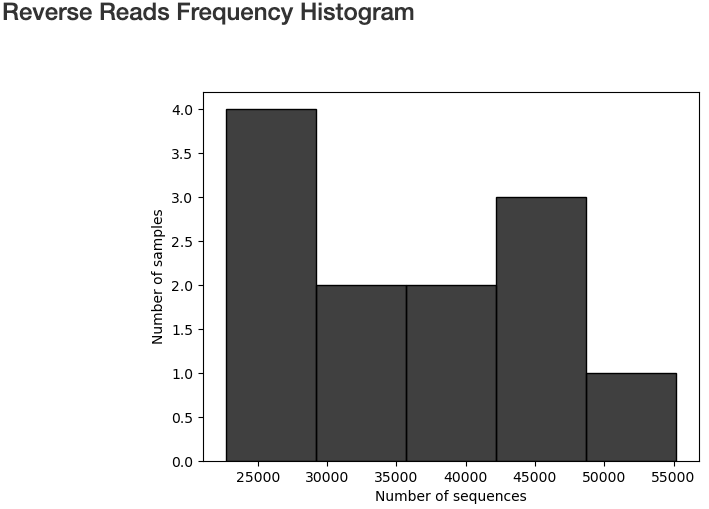
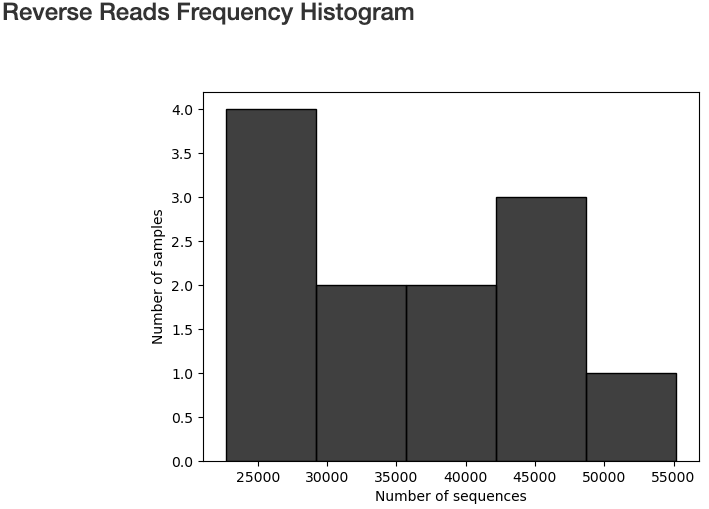
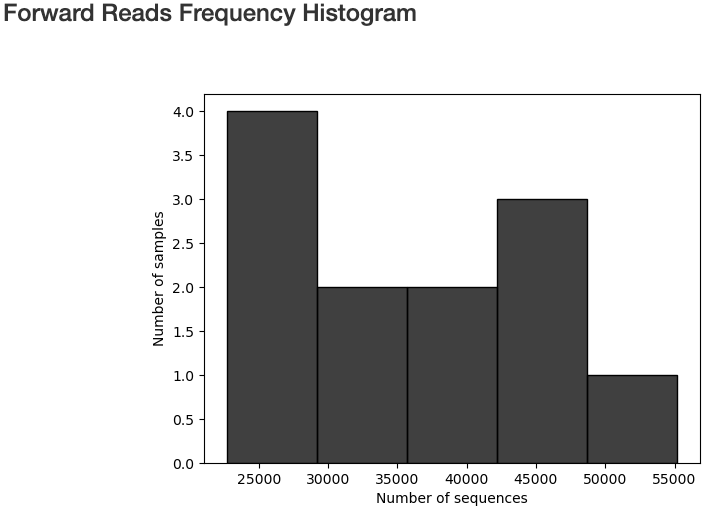
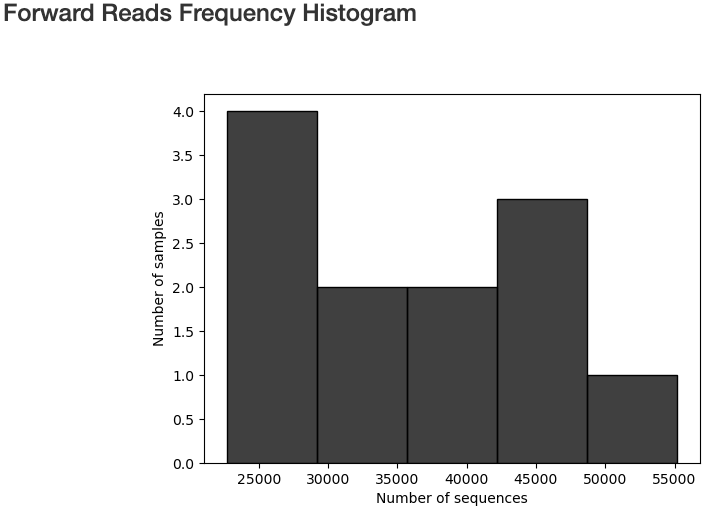
## Download FASTQ files, which correspond to the given twelve SRA IDs

### PASS

## Check quality of the source data using FastQC and MultiQC

### PASS

## Import FASTQ files into QIIME2 file format, pre-process and QC the data with QIIME2 tools



### A screenshot of a number AI-generated content may be incorrect.

## Make ASVs feature table using DADA2 algorithm

### A white text box with a black text box AI-generated content may be incorrect.

### A screenshot of a computer AI-generated content may be incorrect.

### A graph of a number of blue bars AI-generated content may be incorrect.

## Evaluate phylogeny of identified microbial sequences

### See (supporting\_materials/s05\_tree\_nwk.pdf)

### <COME BACK TO THIS WITH THE R SCRIPT s11 sample dendrogram>

## Make rarefaction plots, select an appropriate rarefaction threshold

### Max-depth: Taken as max value for `*non-chimeric*` in supporting\_materials/metadata.tsv.

### `awk 'NR > 2 {print $8}' metadata.tsv | sort -r | head -n1`

### SEE s06a\_alpha\_rarefaction.qzv

## Evaluate and interpret Alpha and Beta diversity in the dataset

### PASS

## Evaluate the taxonomy of the soil microbial communities in the studied grassland sites

### PASS