**Filtering and Analysis of Sediment Samples**

1. Extract the 4 sediment samples using the metadata file and filter sequences and taxonomy:

mkdir sediment-analysis

qiime feature-table filter-samples --i-table dada2-filtered/control-sample-filtered-table.qza --o-filtered-table sediment-analysis/sediment-filtered-table.qza --m-metadata-file metadata.txt --p-where 'Type="Sediment"'

qiime feature-table filter-seqs --i-data dada2/rep-seqs.qza --i-table sediment-analysis/sediment-filtered-table.qza --o-filtered-data sediment-analysis/sediment-filtered-rep-seqs.qza

biom summarize-table -i sediment-analysis/sediment-filtered-table.biom -o sediment-analysis/sediment-filtered-table-summary.txt

biom convert -i sediment-analysis/sediment-filtered-table.biom -o sediment-analysis/sediment-filtered-table.tsv --to-tsv

python taxonomy\_filter.py --input taxonomy/taxonomy.tsv -o sediment-analysis/sediment-filtered-taxonomy.tsv -f sediment-analysis/sediment-filtered-table.tsv -c 0

qiime tools import --input-path sediment-analysis/sediment-filtered-taxonomy.tsv --output-path sediment-analysis/sediment-filtered-taxonomy.qza --type FeatureData[Taxonomy]

qiime taxa barplot --i-table sediment-analysis/sediment-filtered-table.qza --i-taxonomy sediment-analysis/sediment-filtered-taxonomy.qza --m-metadata-file metadata.txt --o-visualization sediment-analysis/sediment-filtered-taxa-bar-plots.qzv

biom normalize-table -i sediment-analysis/sediment-filtered-table.biom -o sediment-analysis/sediment-filtered-table-rel.biom -r

2. Repeat the diversity analysis steps:

qiime alignment mafft --i-sequences sediment-analysis/sediment-filtered-rep-seqs.qza --o-alignment sediment-analysis/diversity-analysis/aligned-filtered-rep-seqs.qza

qiime alignment mask --i-alignment sediment-analysis/diversity-analysis/aligned-filtered-rep-seqs.qza --o-masked-alignment sediment-analysis/diversity-analysis/masked-aligned-filtered-rep-seqs.qza

qiime phylogeny fasttree --i-alignment sediment-analysis/diversity-analysis/masked-aligned-filtered-rep-seqs.qza --o-tree sediment-analysis/diversity-analysis/unrooted-tree.qza

qiime phylogeny midpoint-root --i-tree sediment-analysis/diversity-analysis/unrooted-tree.qza --o-rooted-tree sediment-analysis/diversity-analysis/rooted-tree.qza

qiime diversity core-metrics-phylogenetic --i-phylogeny sediment-analysis/diversity-analysis/rooted-tree.qza --i-table sediment-analysis/sediment-filtered-table.qza --p-sampling-depth 5800 --m-metadata-file metadata.txt --output-dir sediment-analysis/diversity-analysis/core-metrics-results

qiime diversity alpha-rarefaction --i-table sediment-analysis/sediment-filtered-table.qza --i-phylogeny sediment-analysis/diversity-analysis/rooted-tree.qza --p-max-depth 5800 --m-metadata-file metadata.txt --o-visualization sediment-analysis/diversity-analysis/alpha-rarefaction.qzv

qiime diversity alpha-group-significance --i-alpha-diversity sediment-analysis/diversity-analysis/core-metrics-results/observed\_features\_vector.qza --m-metadata-file metadata.txt --o-visualization sediment-analysis/diversity-analysis/observed-features-group-significance.qzv

3. Collapse feature table to family level as relative abundance for graphical analysis:

qiime tools import --input-path sediment-analysis/sediment-filtered-table-rel.biom --output-path sediment-analysis/sediment-filtered-table-rel.qza --type FeatureTable[Frequency]

qiime taxa collapse --i-table sediment-analysis/sediment-filtered-table-rel.qza --i-taxonomy sediment-analysis/sediment-filtered-taxonomy.qza --o-collapsed-table sediment-analysis/sediment-filtered-table-rel-6.qza --p-level 6

qiime tools export --input-path sediment-analysis/sediment-filtered-table-rel-5.qza --output-path sediment-analysis/

biom convert -i sediment-analysis/sediment-filtered-table-rel-5.biom -o sediment-analysis/sediment-filtered-table-rel-5.tsv --to-tsv

Note the abundance of Flavobacteriaceae and Halomonadaceae (families containing Arenibacter and Halomonas respectively) in the LBS and BHBS samples, as well as at low level in BPS. These sequences are likely due to barcode bleed from the corresponding (contaminated) water samples.

A graph with different colored bars

Description automatically generated

The main difference between the Blackford Pond sediment samples is the abundance of Sulfurimonadaceae in ISED but not in BPS. This corresponds to Sulfuricurvum sequences (sulphur oxidisers). ISED has much (>10 times) greater sequencing depth and more rare diversity.