

Phyloseq and Microbiome

Anders K. Krabberød

Phyloseq

- Phyloseq is a set of **classes**, **wrappers**, and tools (in R) designed to make it easier to import, store, and analyse sequencing data that has been grouped (or clustered) into Operational Taxonomic Unts.
- The main page of phyloseq:
<http://joey711.github.io/phyloseq/>
- Published in PLOS ONE: Mcmurdi & Holmes 2013
- <https://journals.plos.org/plosone/article/file?type=printable&id=10.1371/journal.pone.0061217>
- Installation
 - `BiocManager::install("phyloseq")`

Phyloseq

- It is object-oriented and can store representations of microbiome census data in R.
- It supports importing data from a variety of common formats.
- Analyses include calibration, filtering, subsetting, agglomeration, multi-table comparisons, diversity analysis, parallelized Fast UniFrac, ordination methods, publication-quality graphics, etc.

Phyloseq

- Main advantage:
 - Integrates many kinds of data

An R Package for Microbiome Census Data

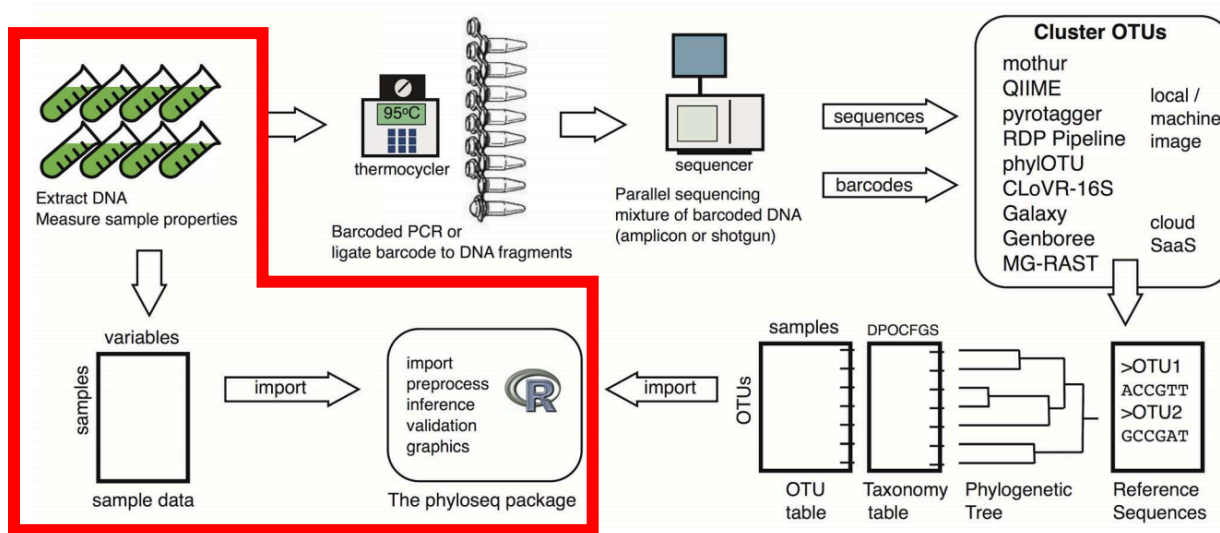


Figure 1. Example of a phylogenetic sequencing workflow. A diagram of an experimental and analysis workflow for amplicon or shotgun phylogenetic sequencing. The intended role for phyloseq is indicated.
doi:10.1371/journal.pone.0061217.g001

Phyloseq

- Content of a phyloseq object may contain:

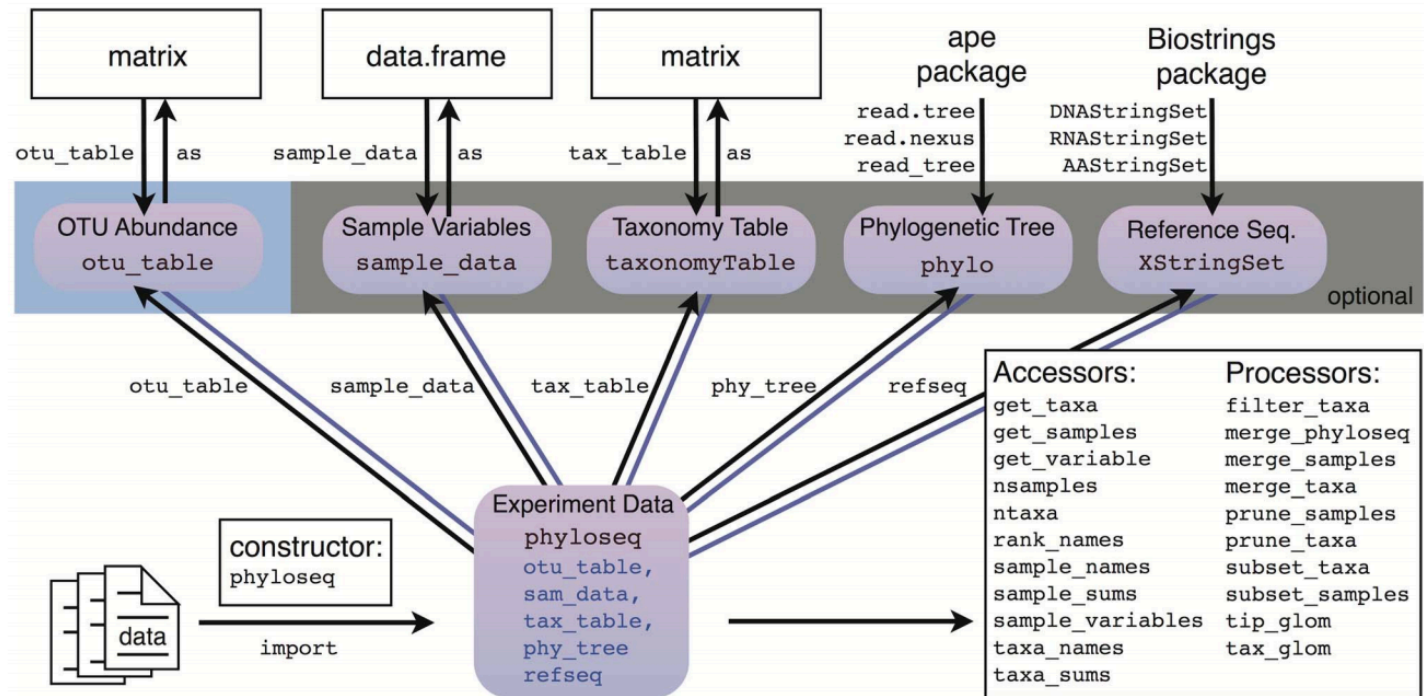


Figure 3. The “phyloseq” class. The phyloseq class is an experiment-level data storage class defined by the phyloseq package for representing phylogenetic sequencing data. Most functions in the phyloseq package expect an instance of this class as their primary argument. See the phyloseq manual [38] for a complete list of functions.

doi:10.1371/journal.pone.0061217.g003

Phyloseq

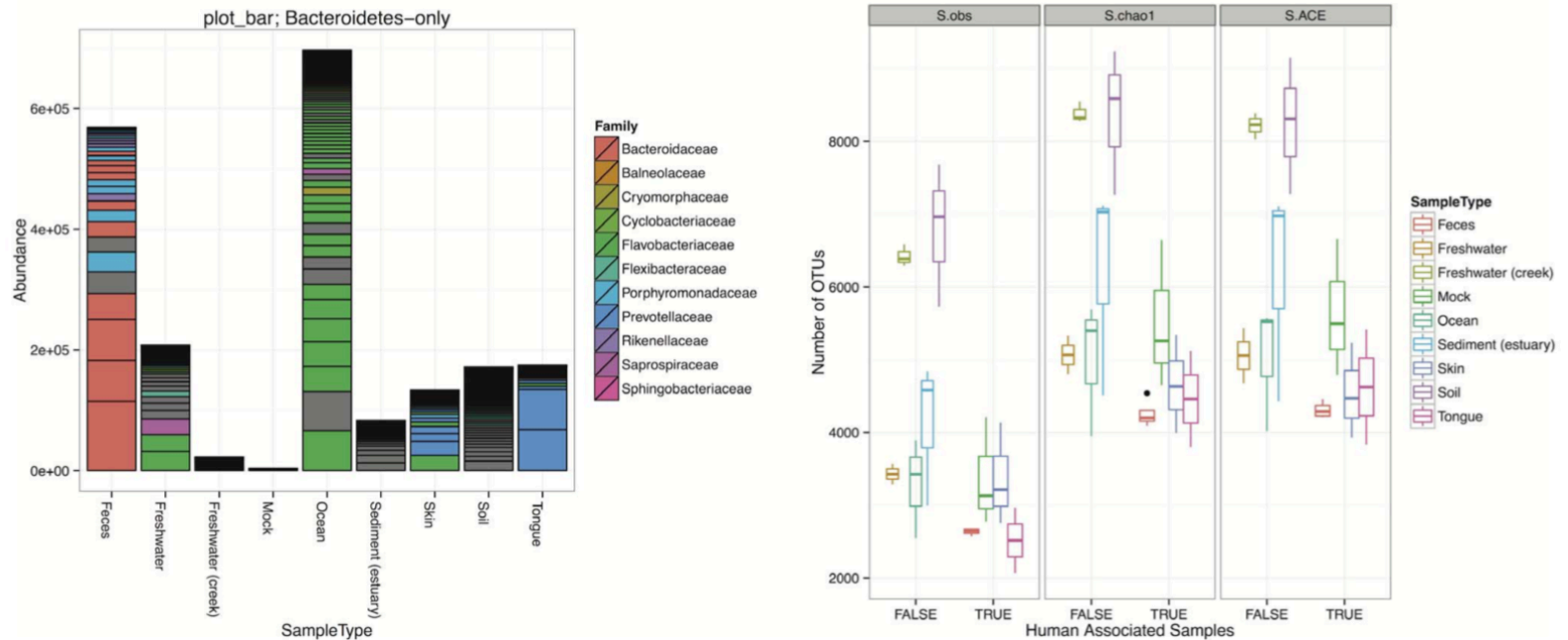
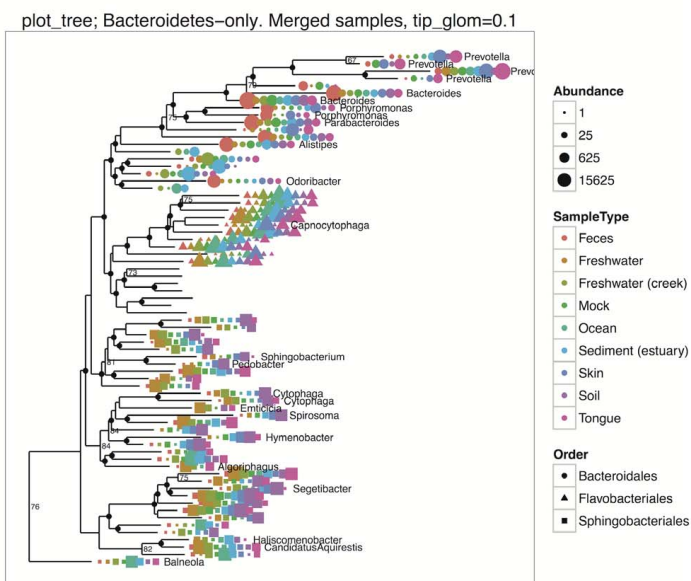
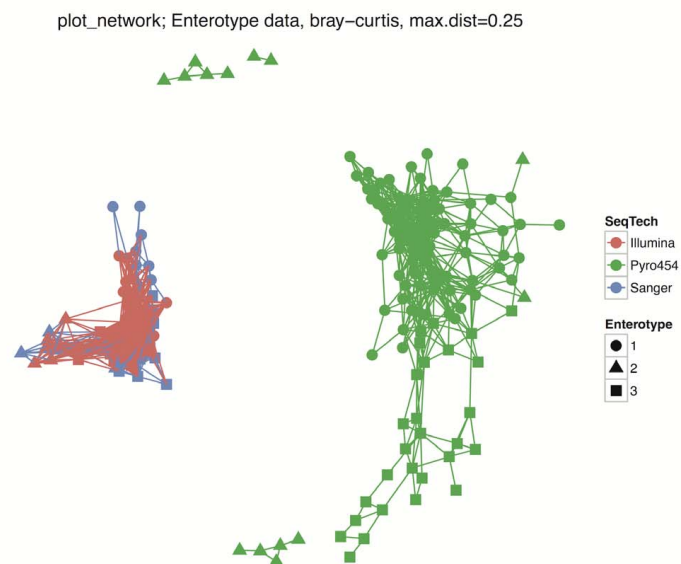
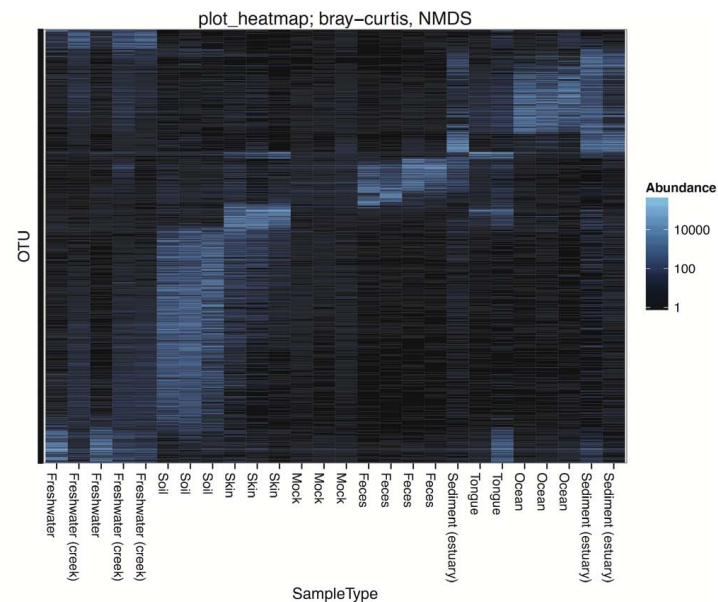
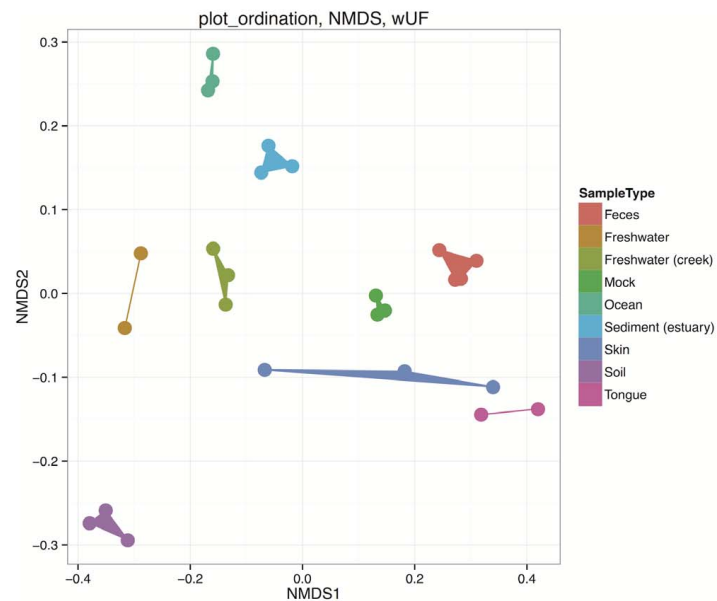


Figure 4. Graphic functions of the phyloseq package. The phyloseq class is an experiment-level data storage class defined by the phyloseq package for representing phylogenetic sequencing data. Most functions in the phyloseq package expect an instance of this class as their primary argument. See the phyloseq manual The Global Patterns [47] and Enterotypes [91] datasets are included with the phyloseq package. The Global



Phyloseq

- Many tutorials online:
 - General use with examples
 - <https://joey711.github.io/phyloseq/tutorials-index.html>
 - https://vaalot.github.io/tutorials/Phyloseq_tutorial.html
 - For data import
 - <https://joey711.github.io/phyloseq/import-data.html>
 - Plotting trees:
 - https://joey711.github.io/phyloseq/plot_tree-examples.html
 - Etc.

Microbiome

- Also a package in R.
- Developed by Leo Lahti, Sudarshan Shetty and others.
- <https://microbiome.github.io/>
- An extension of Phyloseq with a particular focus on amplicon sequencing data.
- Use the same phyloseq object, have included some extra functionality
 - Tutorials and examples
 - <https://microbiome.github.io/tutorials/>