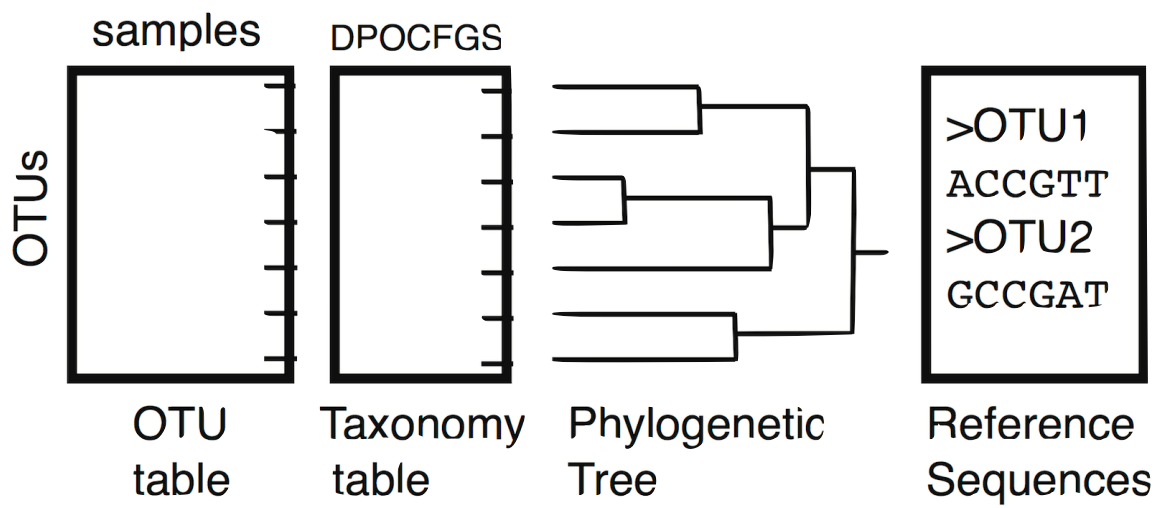
# phyloseq


samples
DPOCFGS
OTUs
>OTU1
ACCGTT
>OTU2
GCCGAT
OTU table
Taxonomy table
Phylogenetic Tree
Reference Sequences

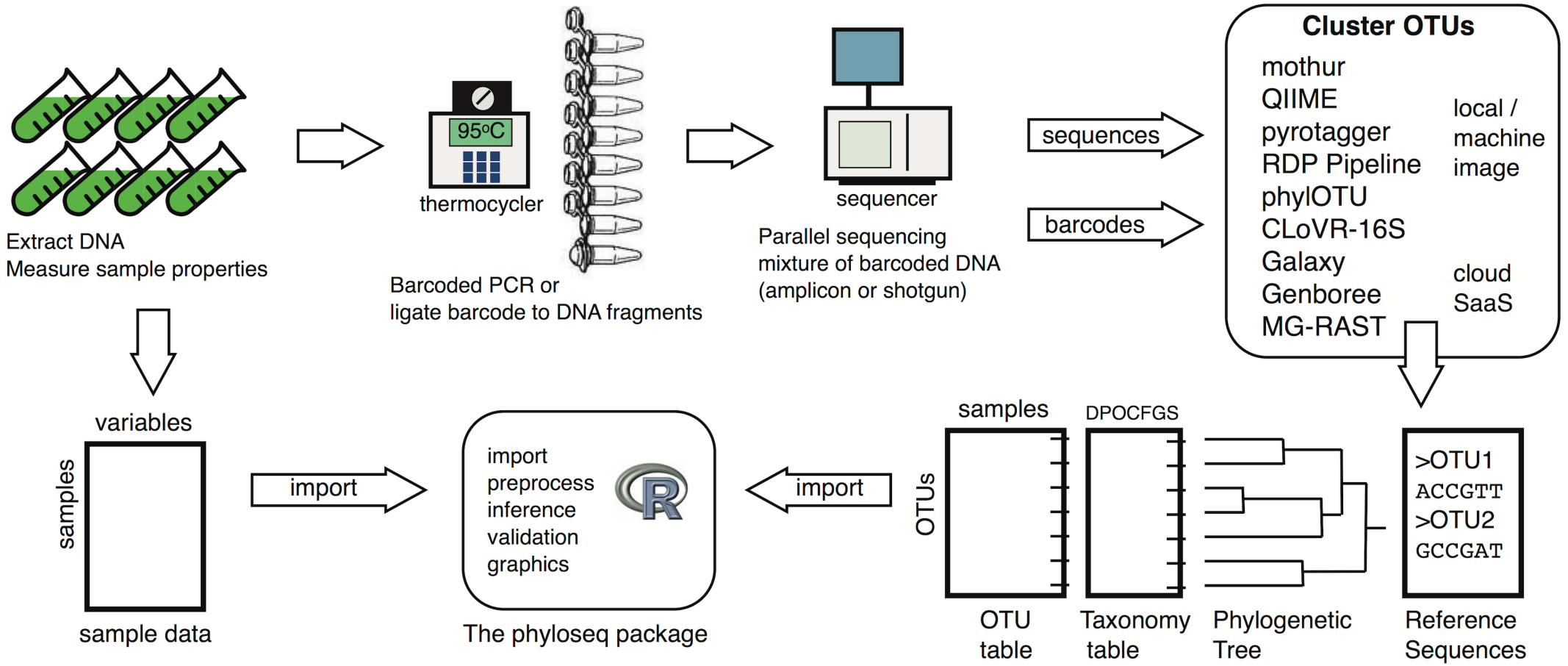
OPEN ACCESS Freely available online

PLOS | ONE

# phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data

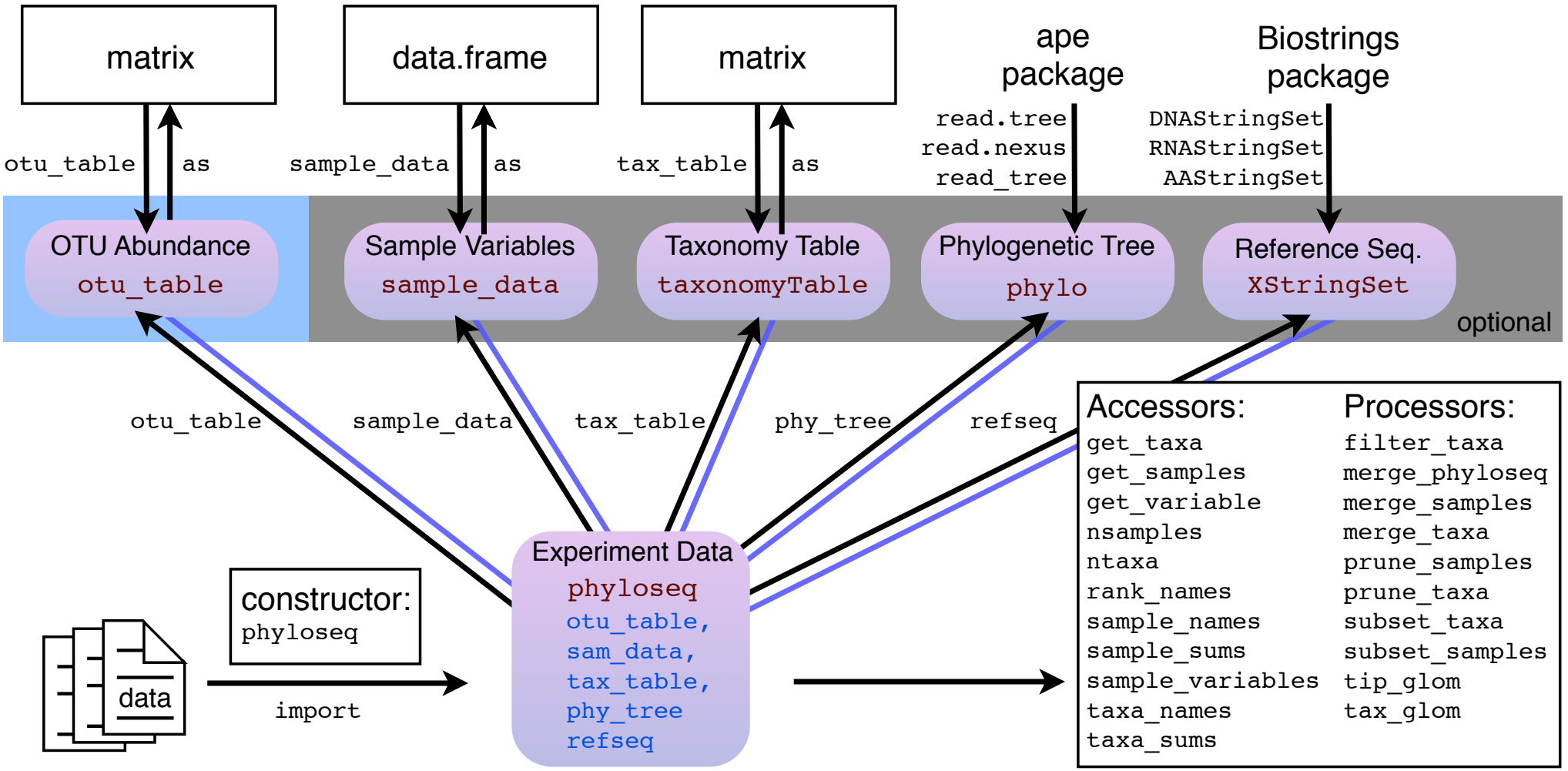
**Paul J. McMurdie, Susan Holmes***

Department of Statistics, Stanford University, Stanford, California, United States of America



Extract DNA
Measure sample properties
95°C
thermocycler
Barcoded PCR or
ligate barcode to DNA fragments
sequencer
Parallel sequencing
mixture of barcoded DNA
(amplicon or shotgun)
sequences
barcodes
Cluster OTUs
mothur
QIIME
pyrotagger
RDP Pipeline
phylOTU
CLoVR-16S
Galaxy
Genboree
MG-RAST
local / machine image
cloud SaaS
variables
samples
sample data
import
import
preprocess
inference
validation
graphics
The phyloseq package
samples
DPOCFGS
OTUs
>OTU1
ACCGTT
>OTU2
GCCGAT
OTU table
Taxonomy table
Phylogenetic Tree
Reference Sequences

# phyloseq

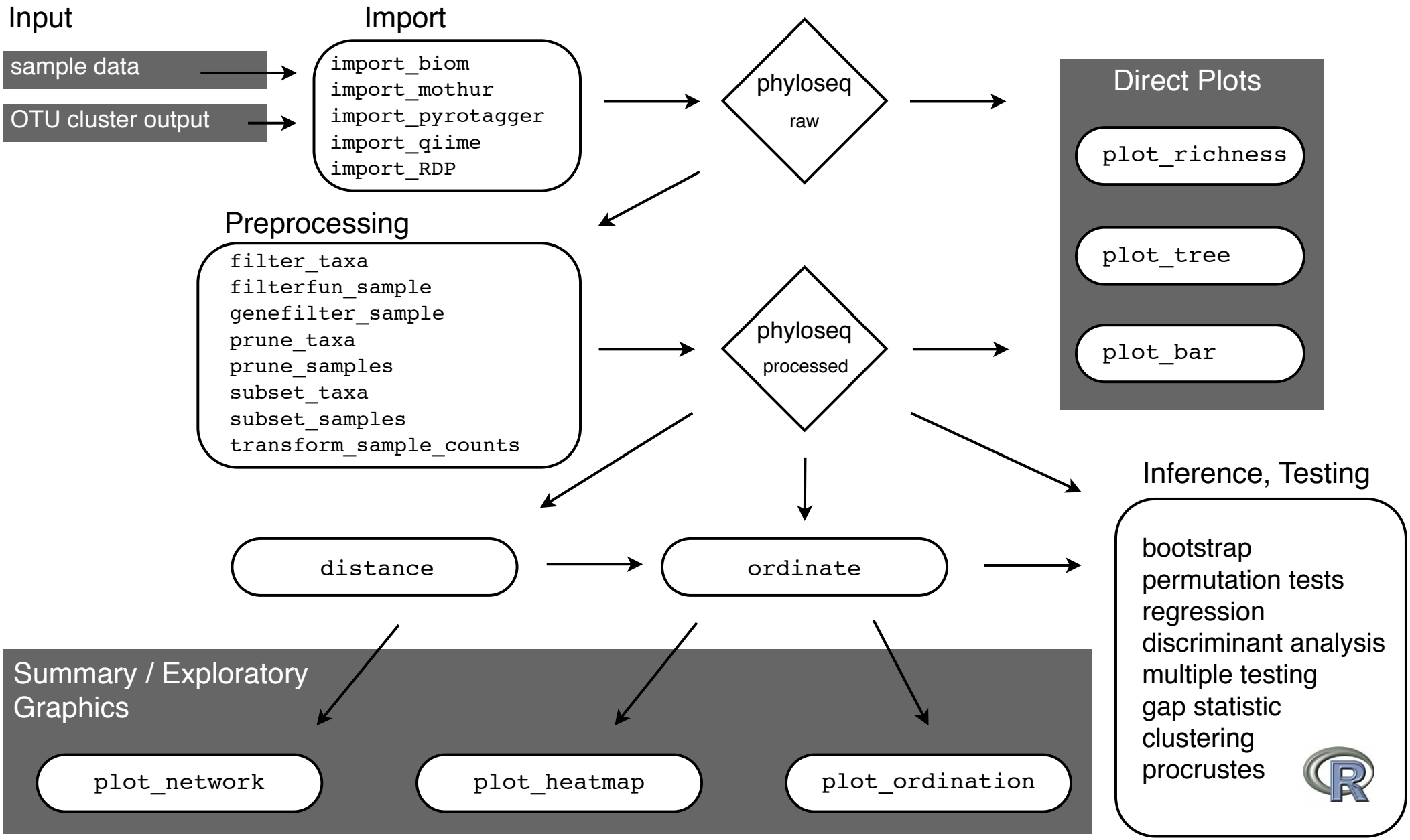
## data structure & API




http://joey711.github.io/phyloseq/

# phyloseq

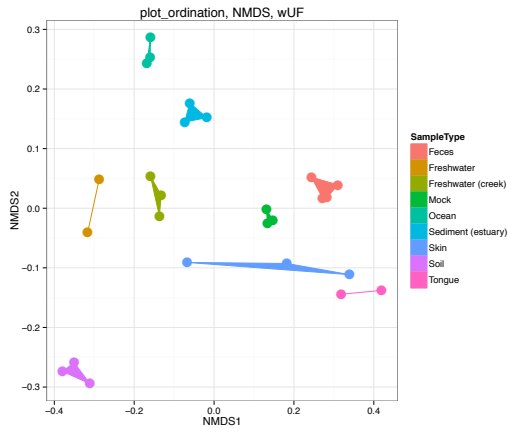
work flow

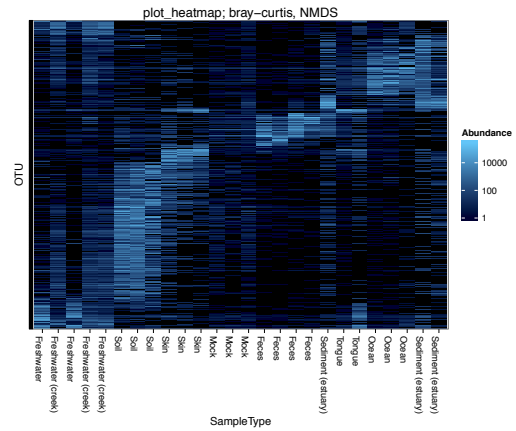


# phyloseq

graphics

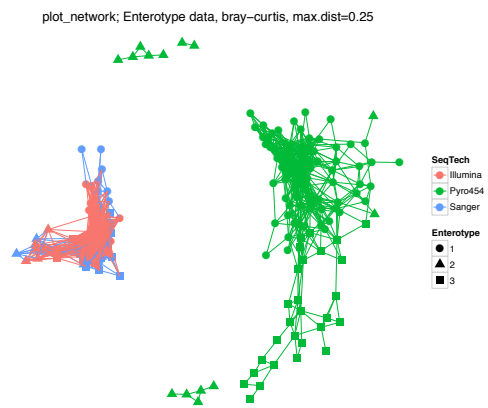
plot_ordination()

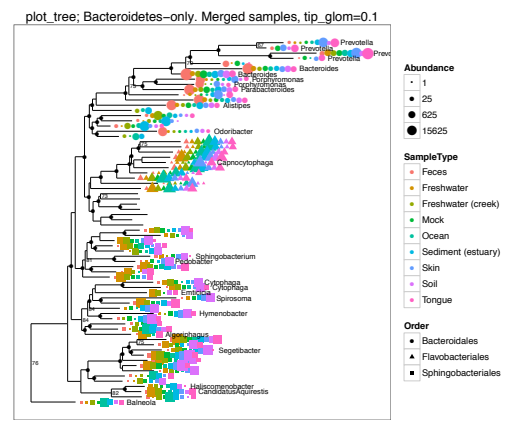



plot_heatmap()

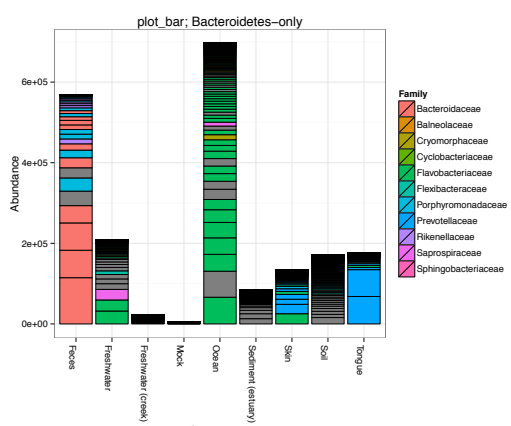
plot_network()

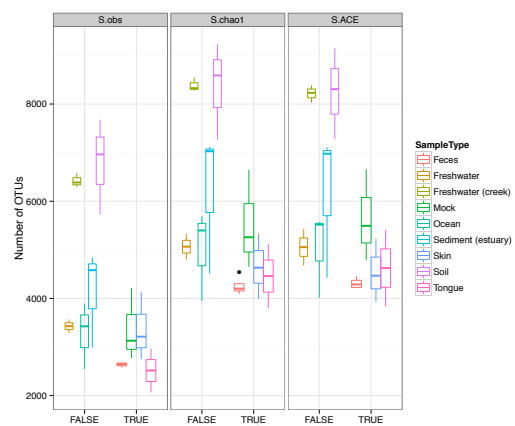



plot_tree()

plot_bar()



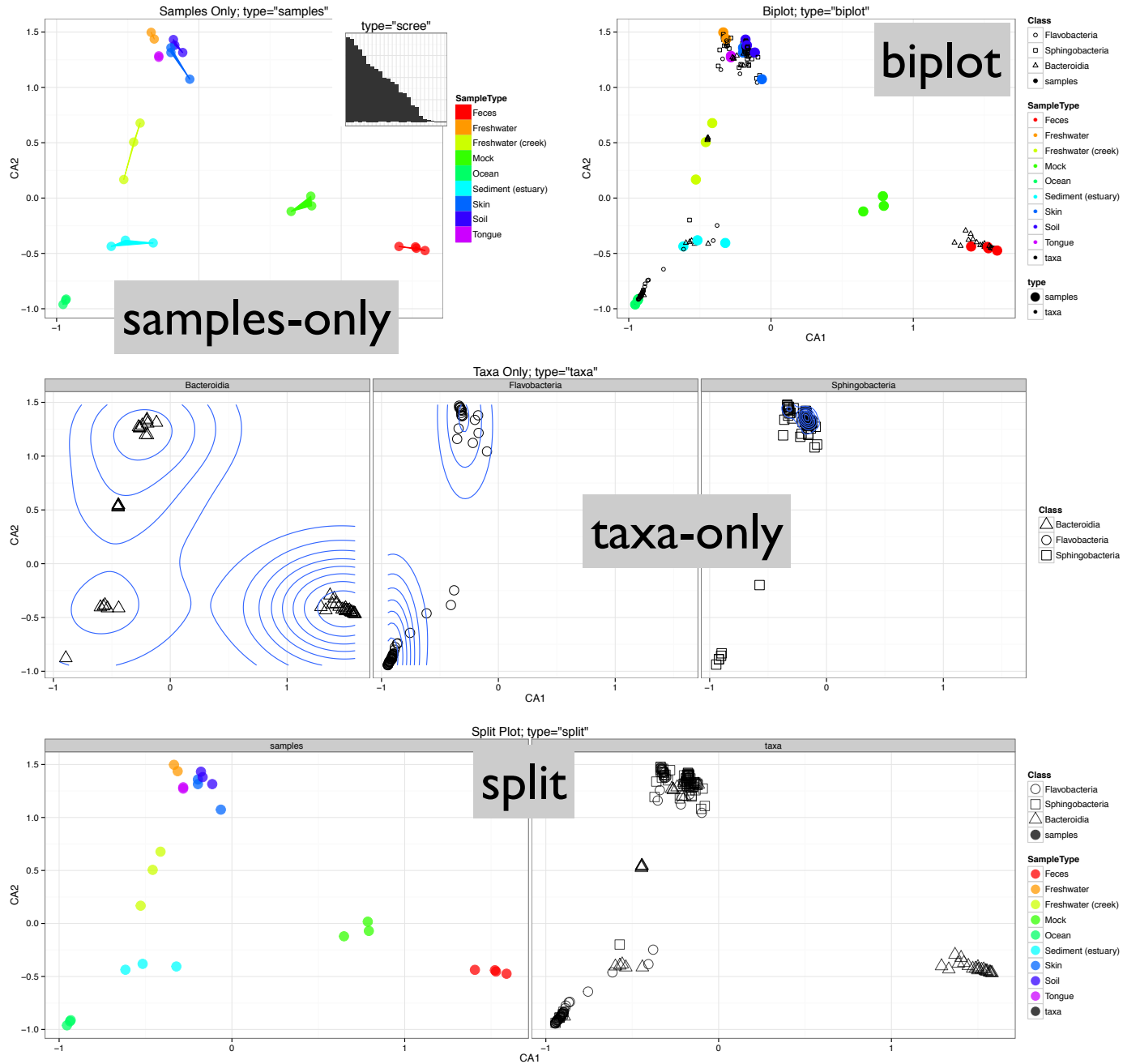




plot_richness()

# phyloseq

graphics

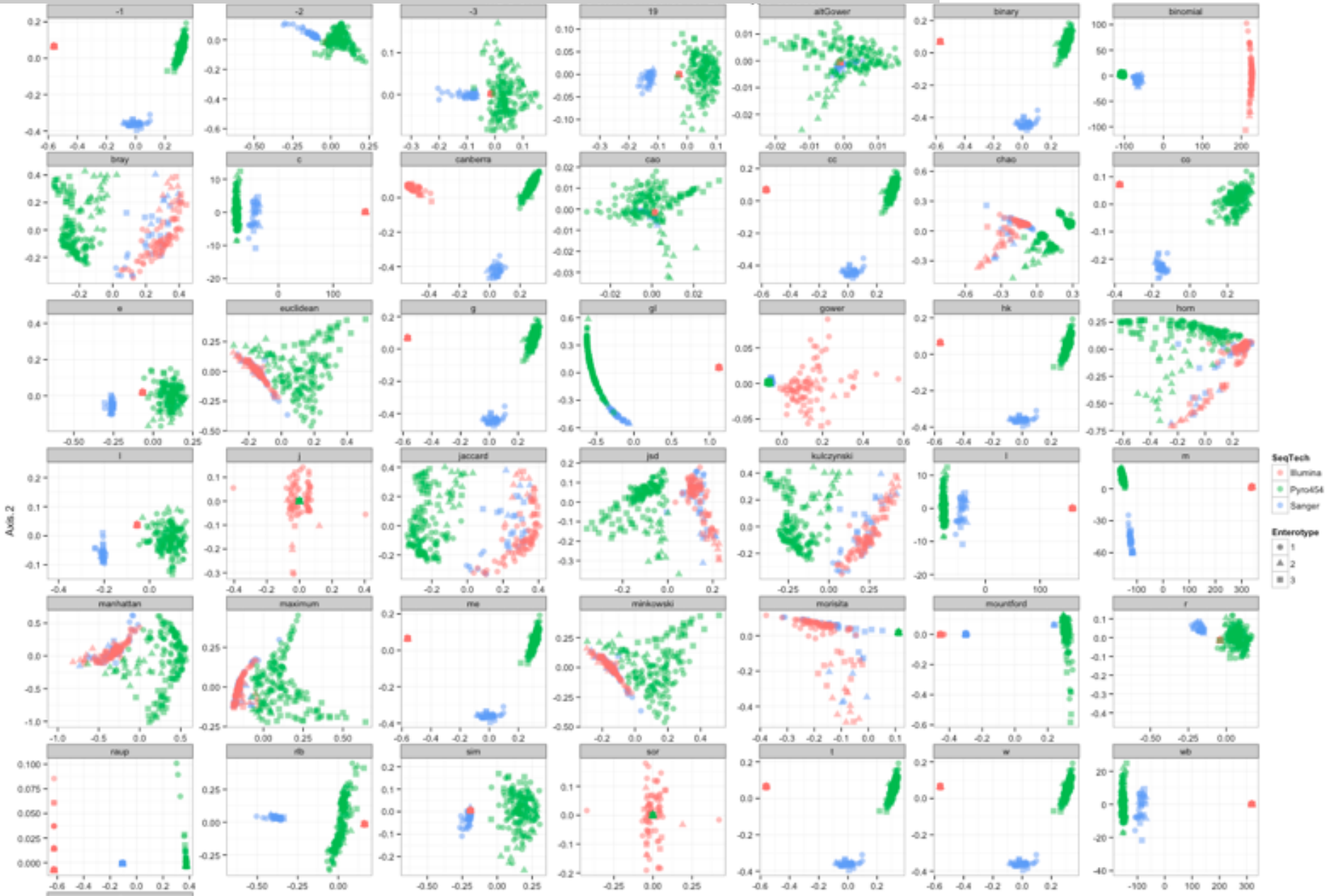
plot_ordination()



# phyloseq

supported distance methods

## MDS on supported distance metrics: enterotype data



# phyloseq: Analyze microbiome census data using R

The analysis of microbiological communities brings many challenges: the integration of many different types of data with methods from ecology, genetics, phylogenetics, network analysis, visualization and testing. The data itself may originate from widely different sources, such as the microbiomes of humans, soils, surface and ocean waters, wastewater treatment plants, industrial facilities, and so on; and as a result, these varied sample types may have very different forms and scales of related data that is extremely dependent upon the experiment and its question(s). The phyloseq package is a tool to import, store, analyze, and graphically display complex phylogenetic sequencing data that has already been clustered into Operational Taxonomic Units (OTUs), especially when there is associated sample data, phylogenetic tree, and/or taxonomic assignment of the OTUs. This package leverages many of the tools available in R for ecology and phylogenetic analysis (vegan, ade4, ape, picante), while also using advanced/flexible graphic systems (ggplot2) to easily produce publication-quality graphics of complex phylogenetic data. phyloseq uses a specialized system of S4 classes to store all related phylogenetic sequencing data as single experiment-level object, making it easier to share data and reproduce analyses. In general, phyloseq seeks to facilitate the use of R for efficient interactive and reproducible analysis of OTU-clustered high-throughput phylogenetic sequencing data.

**https://joey711.github.io/phyloseq/**