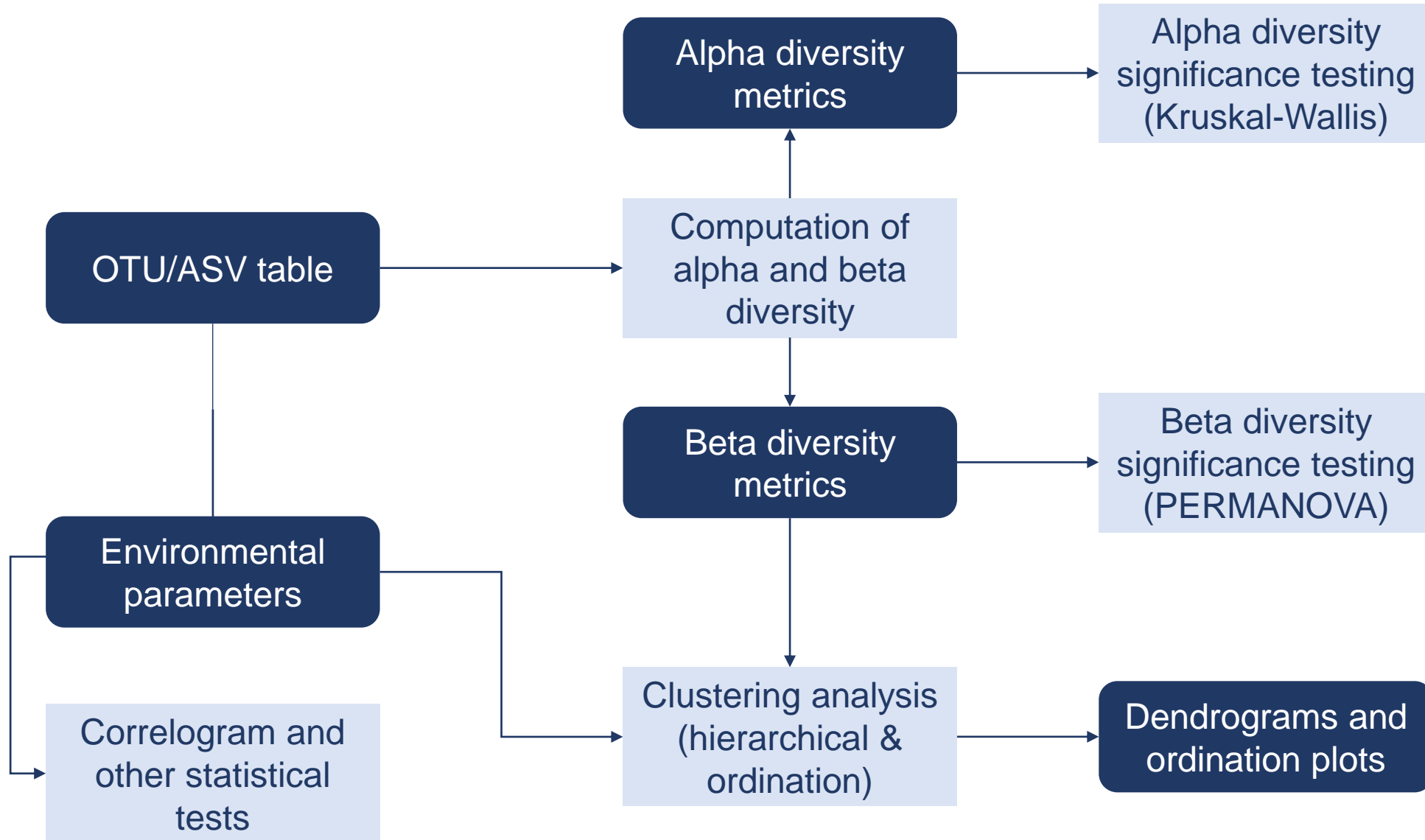


Ecological statistics, reiteration



Ecological statistics workflow



Diversity metrics

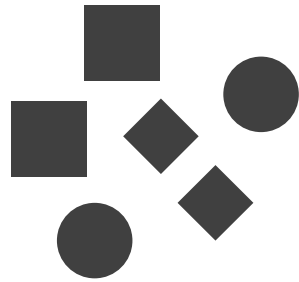
Two types of diversity metrics are commonly explored in microbial community analysis

- Alpha diversity – diversity within a sample
- Beta diversity – difference between samples

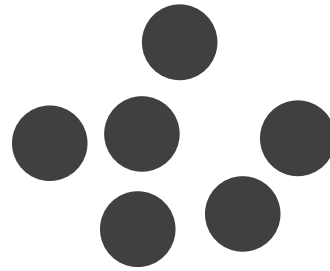


Alpha diversity

How diverse is a sample?



Diverse



Not diverse

- Richness – how many taxa? (Observed features/taxa, Simpson index)
- Evenness – how evenly are the abundances distributed across taxa? (Evenness index)
- Richness and evenness – alpha diversity metric that combines both (Shannon index)



Beta diversity

How different are two or more samples/sites from one another?



- Unweighted – how many taxa are shared? (Jaccard index, unweighted Unifrac)
- Weighted – do shared taxa have similar abundances? (Bray-Curtis dissimilarity, weighted Unifrac)



Clustering methods

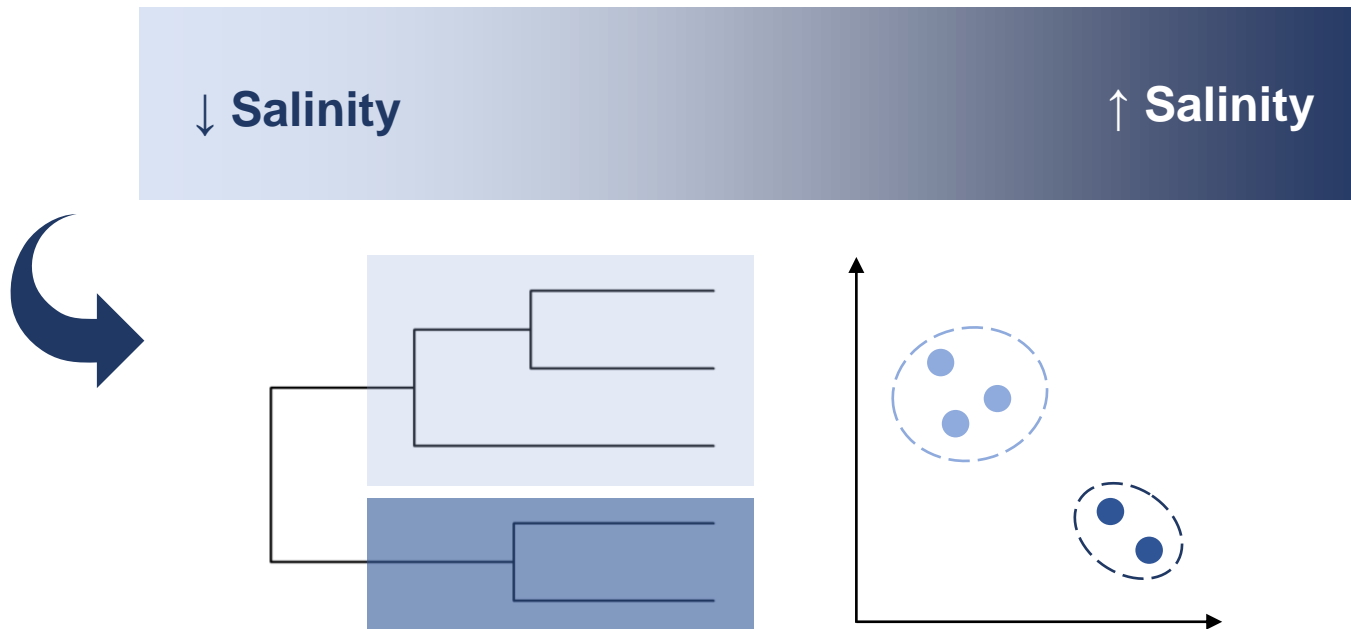
These methods differ by their algorithms. The following are the most commonly used clustering methods:

- Hierarchical clustering
- Ordination methods



Why identify clusters?

Sampling in a gradient



Hierarchical clustering

Members of inferior-ranking members cluster to become members of higher-ranking clusters

Common hierarchical clustering methods:

- Single linkage
- Complete linkage
- Average linkage
- Ward's minimum variance clustering



Ordination methods

Statistical techniques used to extract main trends in data in the form of continuous axes

Two types:

- Unconstrained ordination
- Constrained ordination



Unconstrained ordination

Relies on the information of a single dataset (similarity or dissimilarity metric)

Common types:

- PCA
- PCoA
- NMDS



Constrained ordination

Uses the matrix of the one being clustered and a matrix of explanatory variables (2 matrices) providing a constraint where to group or divide the data of the first matrix

Common types:

- RDA
- db-RDA
- CCA



Statistical tests/data exploration

for alpha diversity

Alpha diversity metrics provide a single value for each sample

These metrics can be viewed/explored with boxplots

This is suitable for classic univariate tests (Kruskal-Wallis)



Statistical tests/data exploration

for beta diversity

Beta diversity metrics are more complicated, data is usually not normal

Can be visualized using dendrograms and ordination plots

Permutational ANOVA (PERMANOVA) is commonly used

