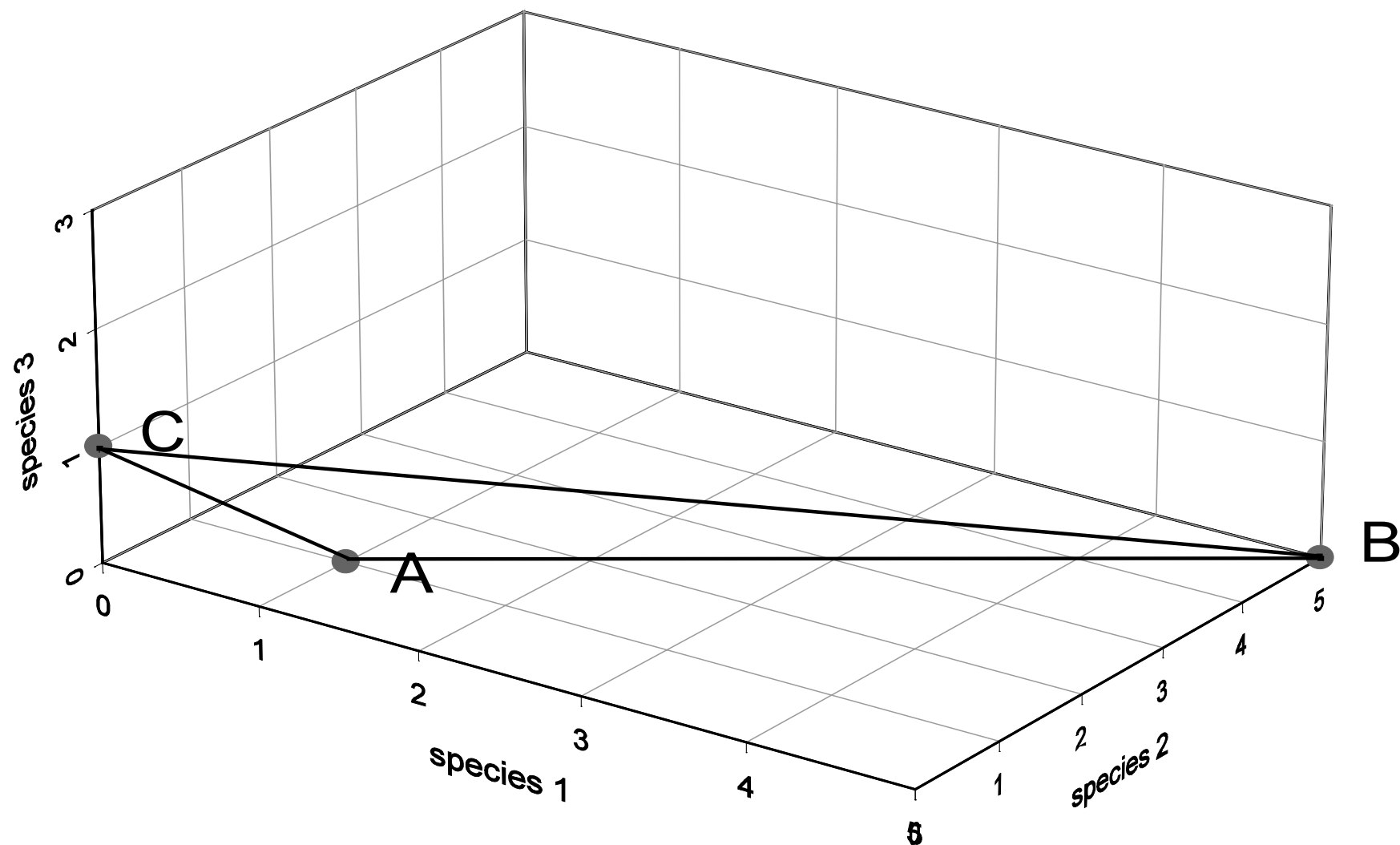# Statistical Analysis of Amplicon Data

Part 3
March 1, 2019

Beta Diversity:
How different are sites?
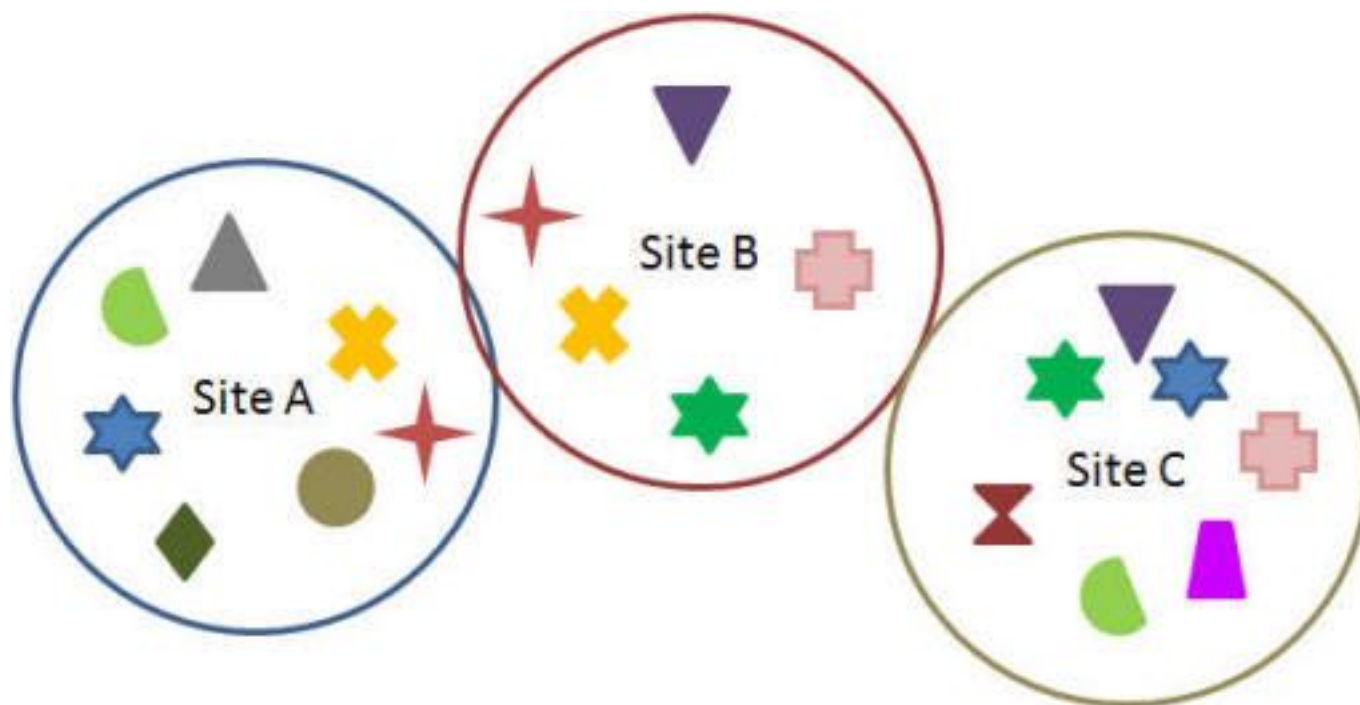
Beta Diversity Metrics

- Number of Taxa

- Bray-Curtis Dissimilarity
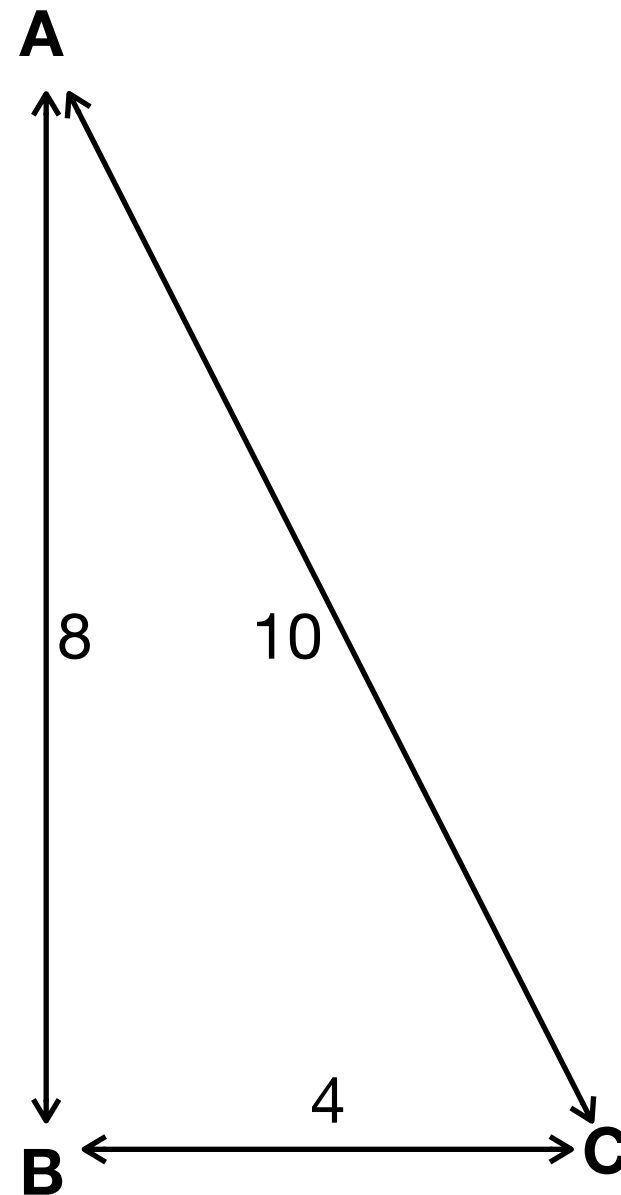
- UniFrac

- etc

Number of Taxa

- How different are samples?

Site
A
B
C

Species Differences

A vs B	8
B vs C	4
A vs C	10

Dimensionality reduction

Species
Differences

A vs B   8

B vs C   4

A vs C   10

A

8   10

B   4   C

Dimensionality reduction

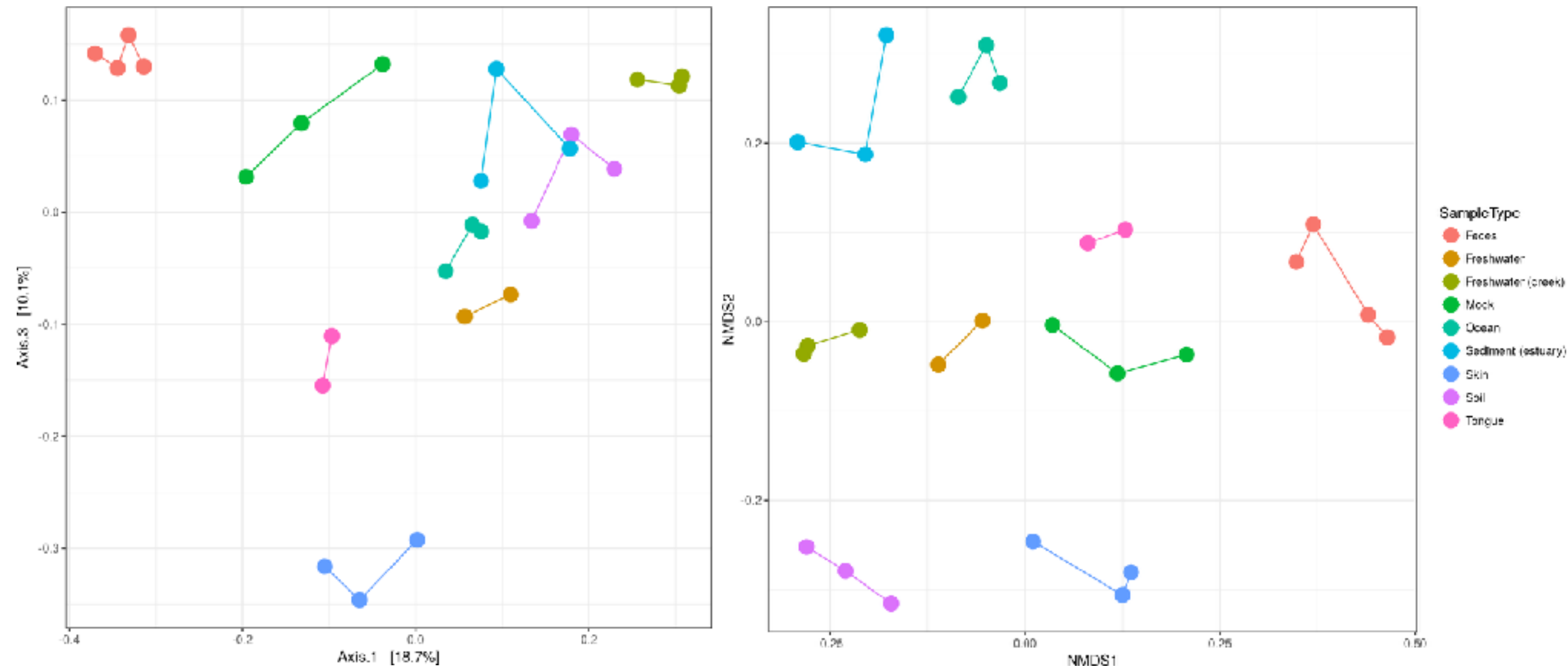| Species	Differences
A vs B	8
B vs C	4
A vs C	10

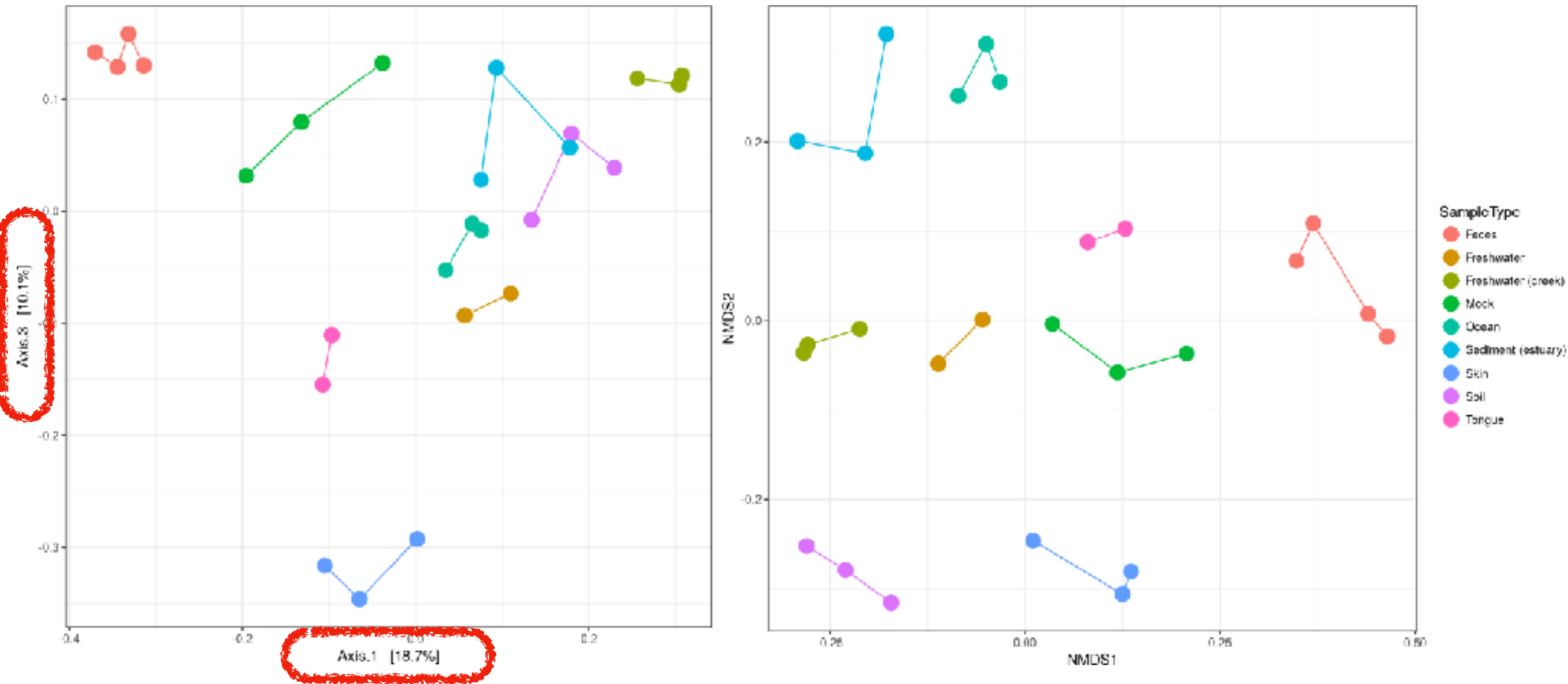# Dimensionality reduction

Stress = 0.16

Dimensionality reduction

- **NMDS (Nonmetric Multidimensional Scaling)**: attempts to maintain pairwise dissimilarity

- **PCoA (Principal coordinates analysis):** Determines axes along which variance is maximized
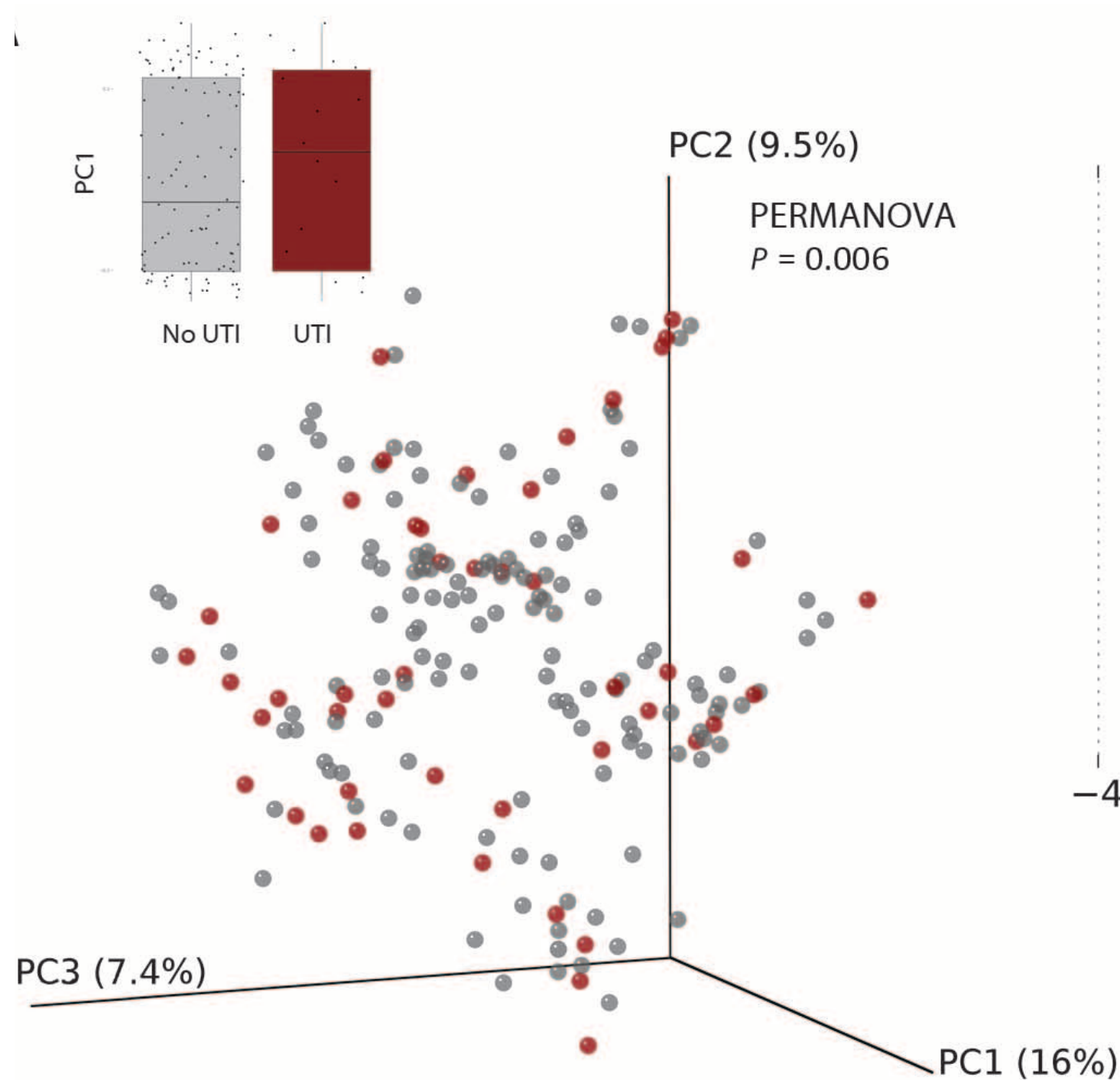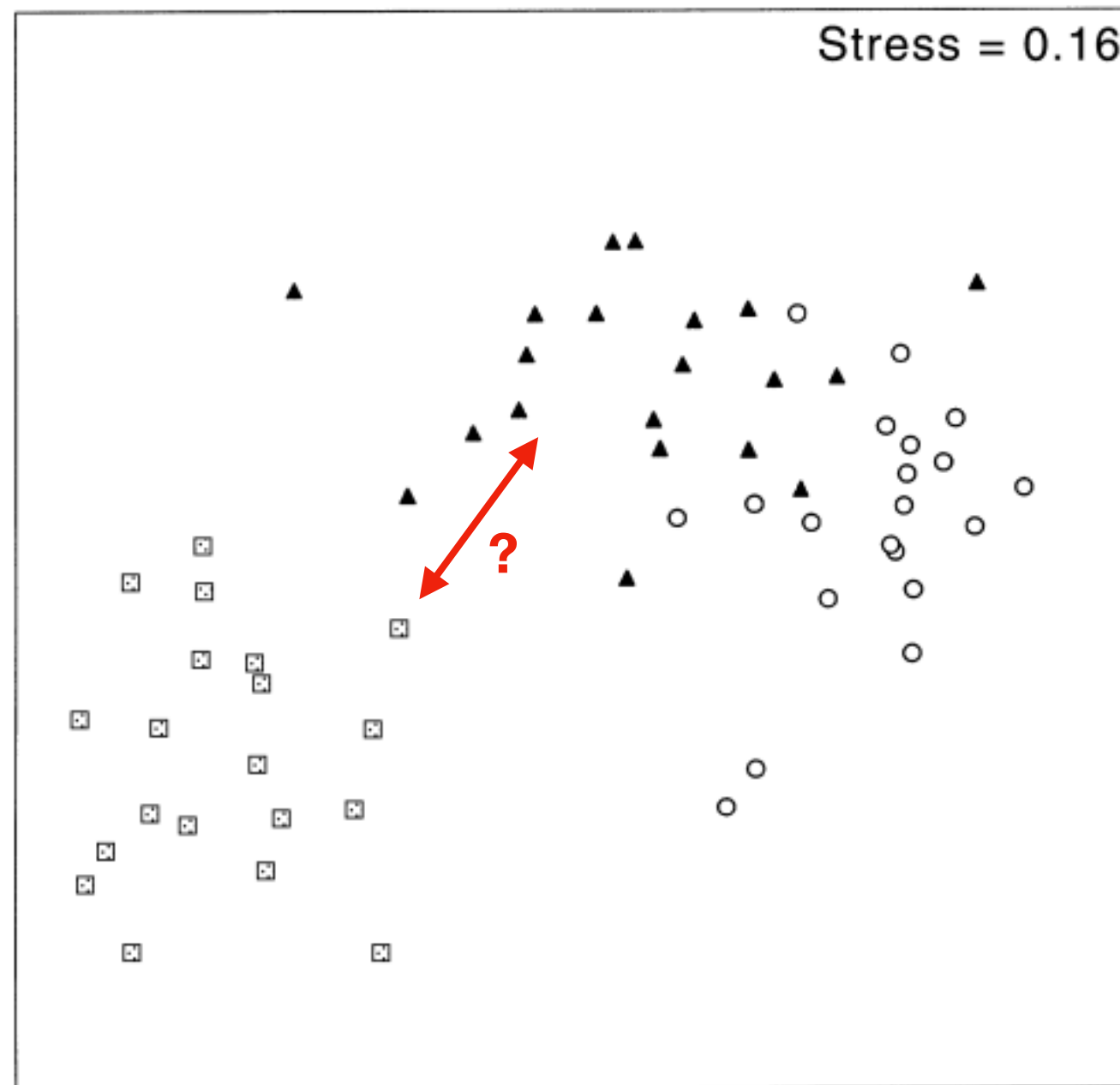
PCoA and NMDS
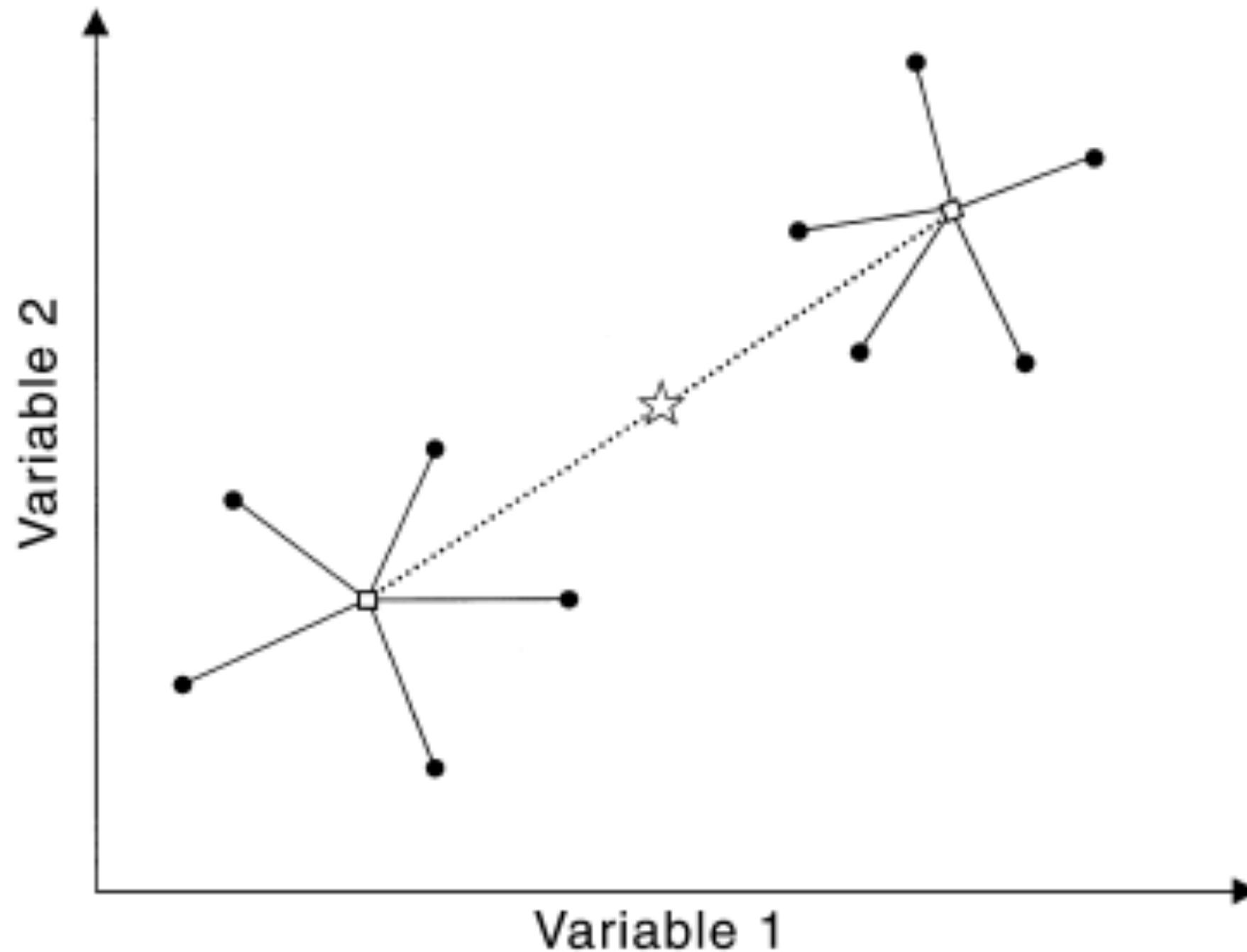
PCoA and NMDS

Figures Josh Hates: 3D PCoA Plots

Non-Parametric Multivariate Tests

PERMANOVA

PERMANOVA Caveat
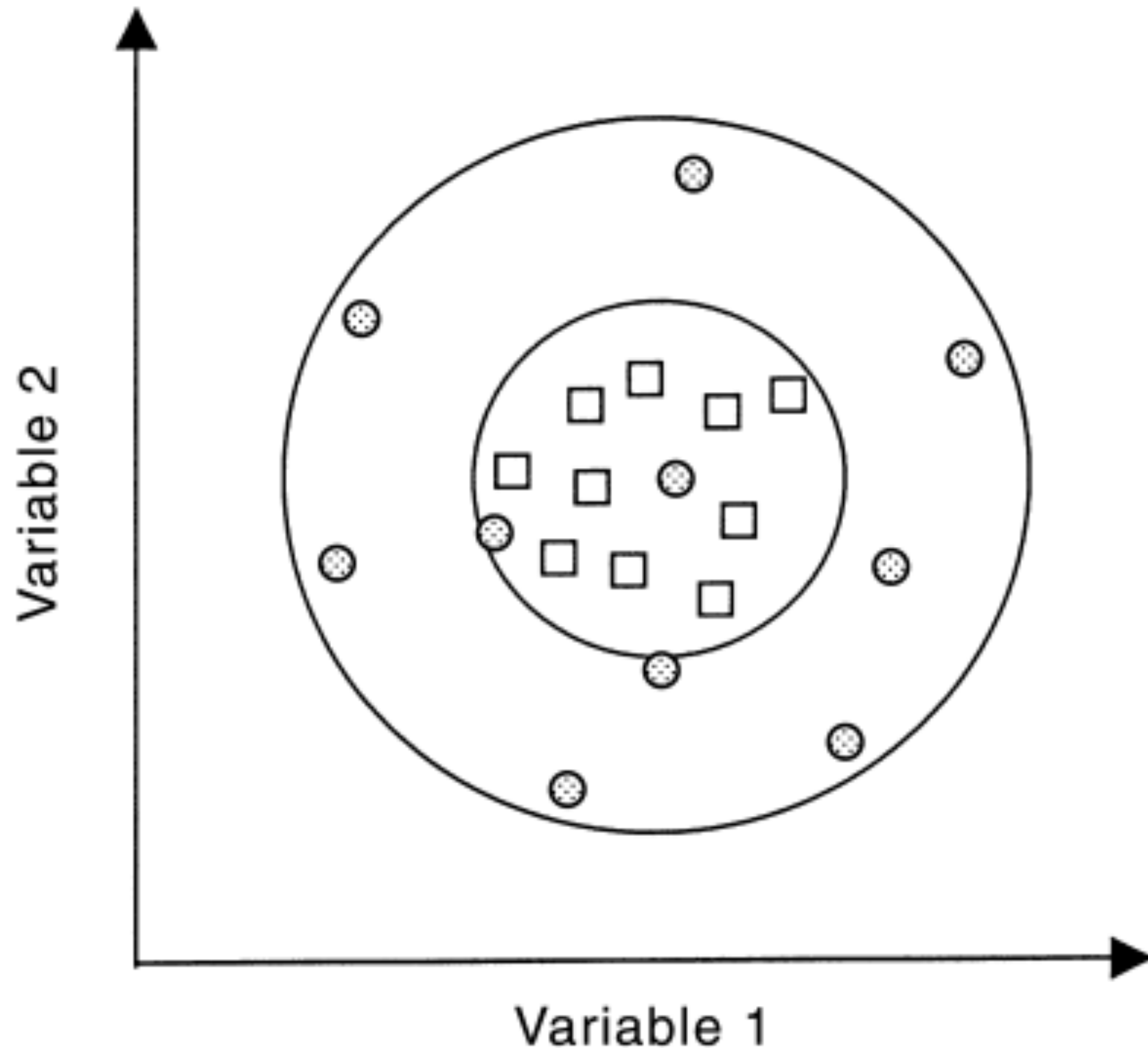
Phylogenetic Trees
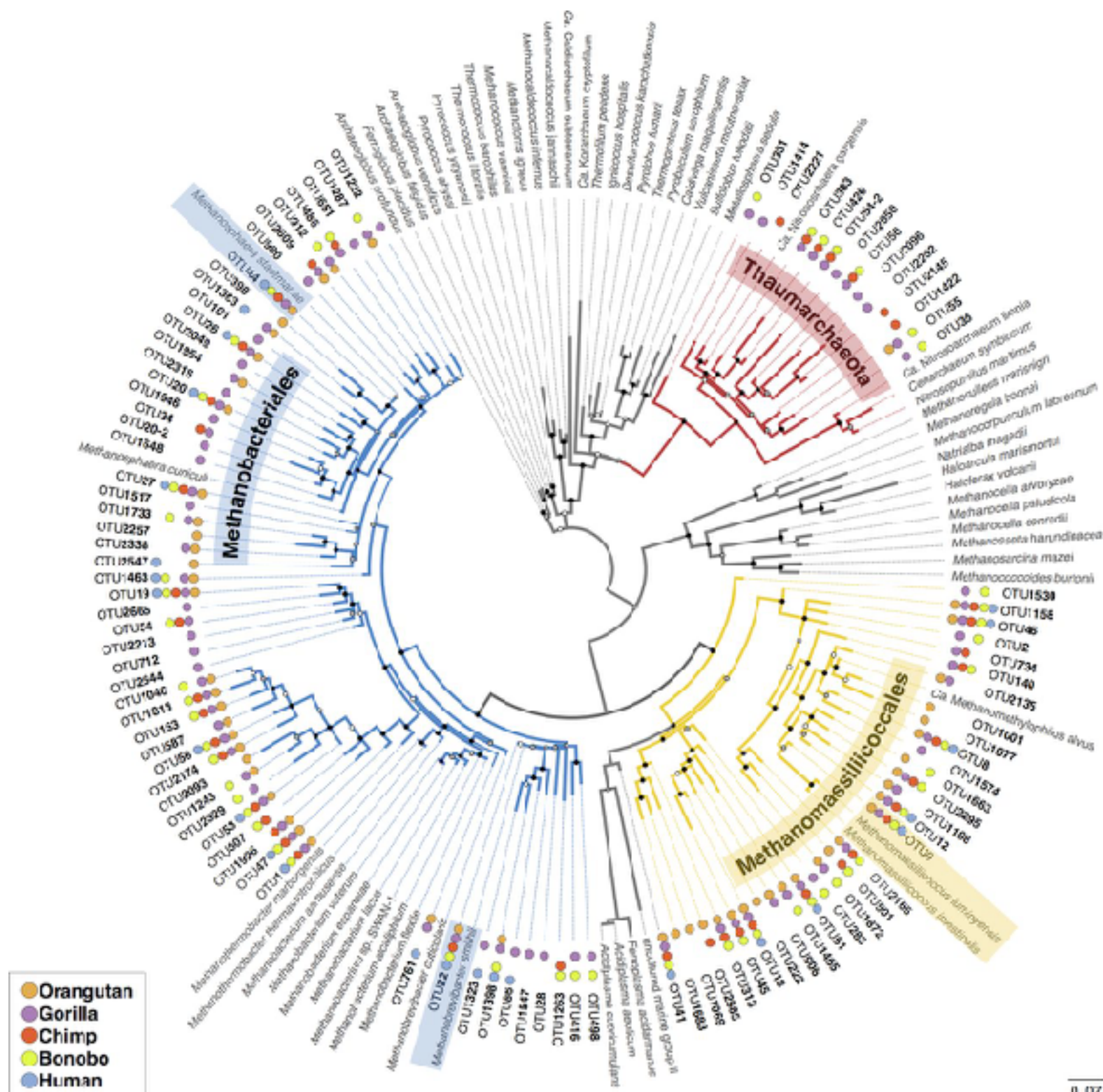
N >> P

	OTU 1	OTU 2	...	Meta 1	Meta 2	...
Sample 1						
Sample 2						
...						
Sample N						

N

P

Machine Learning

- Random Forests

- Support Vector Machines

- Kernel Ridge Regression

- Regularized Logistic Regression

- Bayesian Logistic Regression

- K-Nearest Neighbors

- Probabilistic Neural Networks

- etc