β-Diversity: Basics

November 1, 2019

Useful Reference

- GUide to STatistical Analysis in Microbial Ecology (GUSTA ME)!: https://mb3is.megx.net/gustame
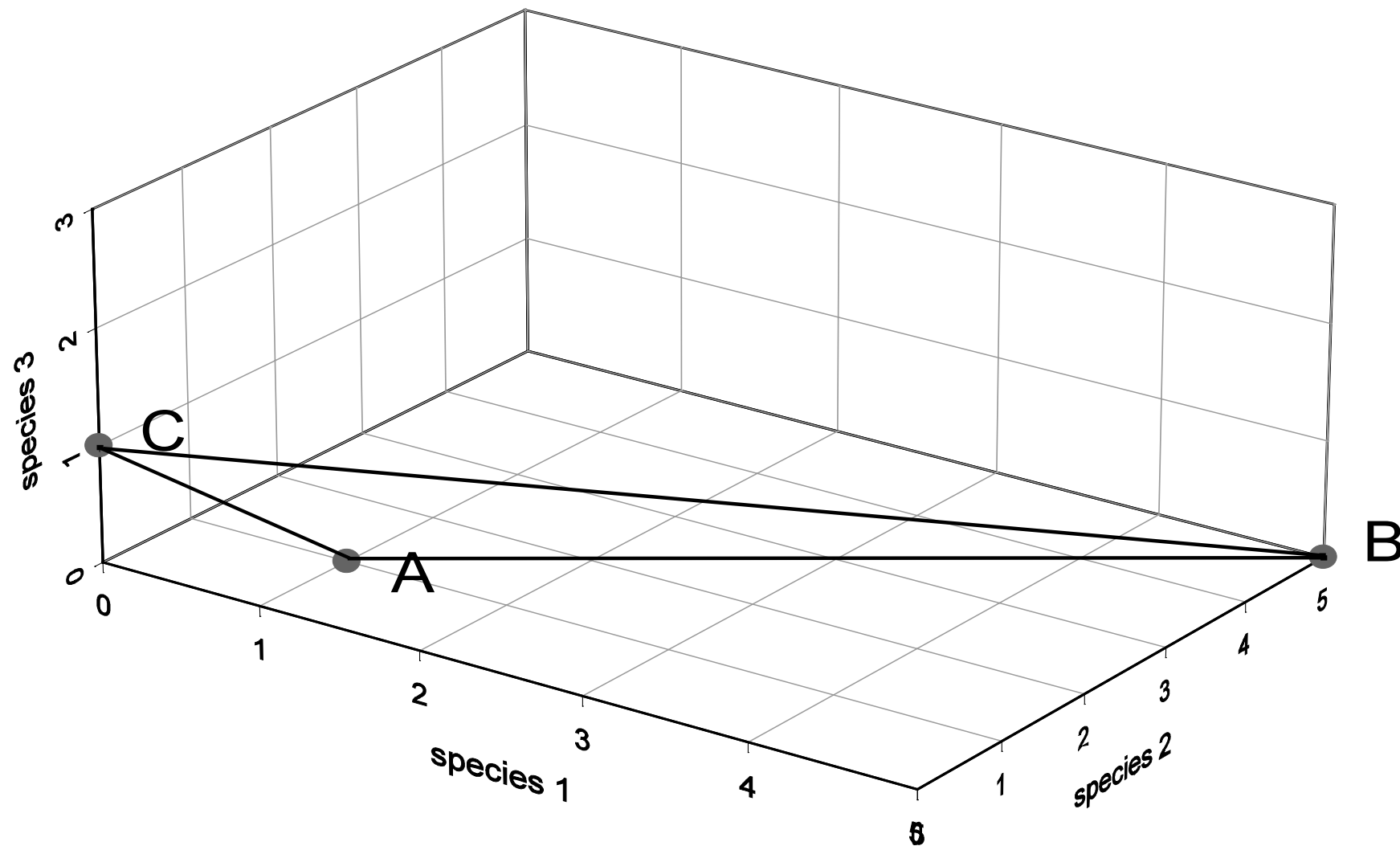
Beta Diversity:
How different are sites?
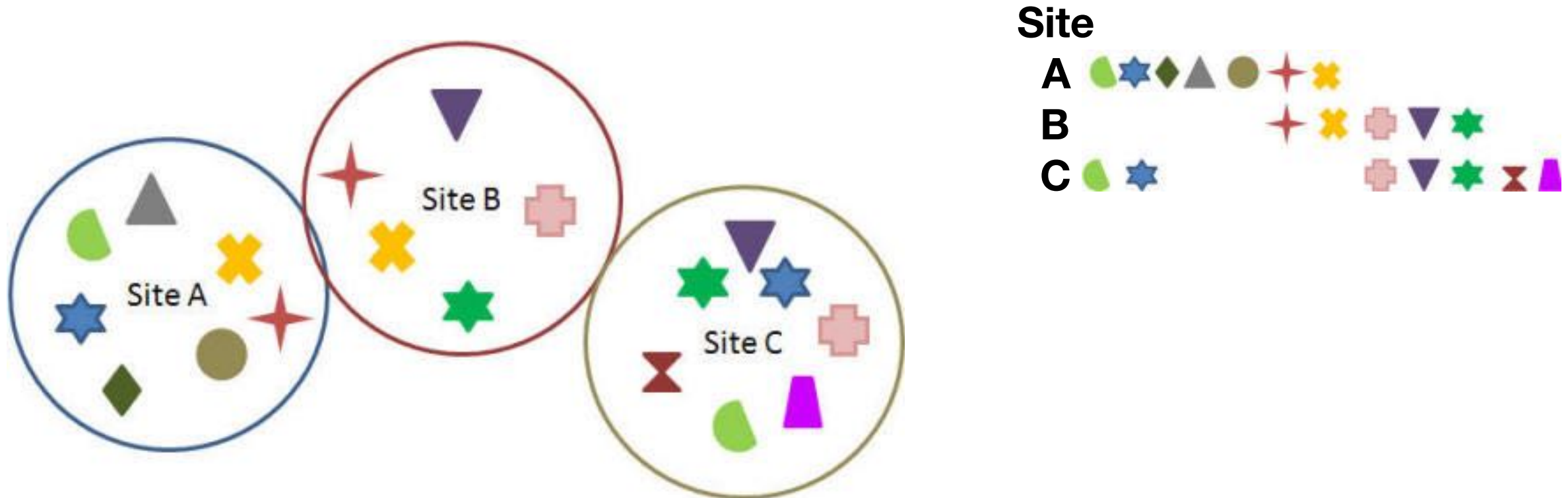
Beta Diversity:
How different are sites?

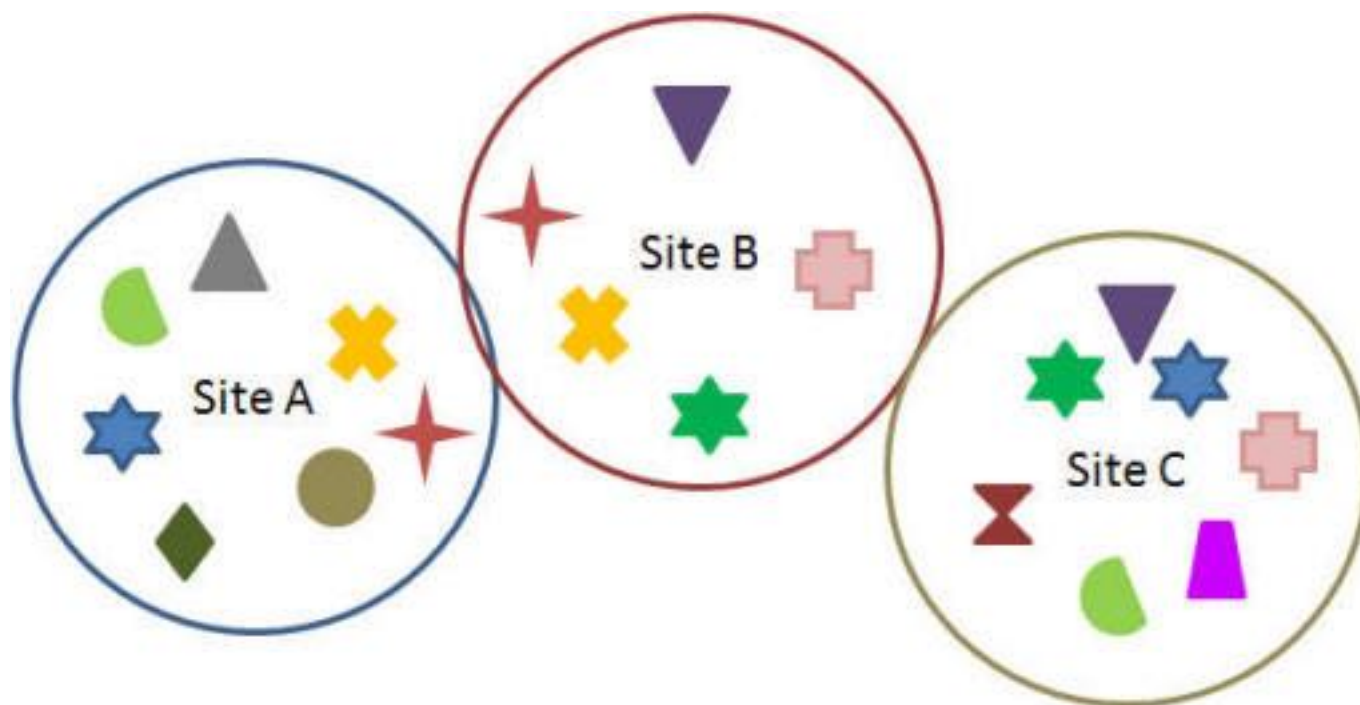| Species			
	1	2	3
A	1	1	0
B	5	5	0
C	0	0	1

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

Bray-Curtis Dissimilarity

$$D_{BC} = 1 - 2\frac{\sum_{i=1}^{S} min\left(a_i,\, b_i\right)}{\sum_{i=1}^{S} a_i + b_i}$$

	Species		
	1	2	3
A	1	1	0
B	5	5	0
C	0	0	1

Where x_i is the abundance of species i in sample X

Bray-Curtis Dissimilarity

```
> y
  housing foodstuffs alcohol other services
1     640        328     147   169      196
2    1800        484     515  2291      912
3     640        328     147   169      196
4    6400       3280    1470  1690     1960

> vegdist(y, method="bray")
          1         2         3
2 0.6043839
3 0.0000000 0.6043839
4 0.8181818 0.4807230 0.8181818
```
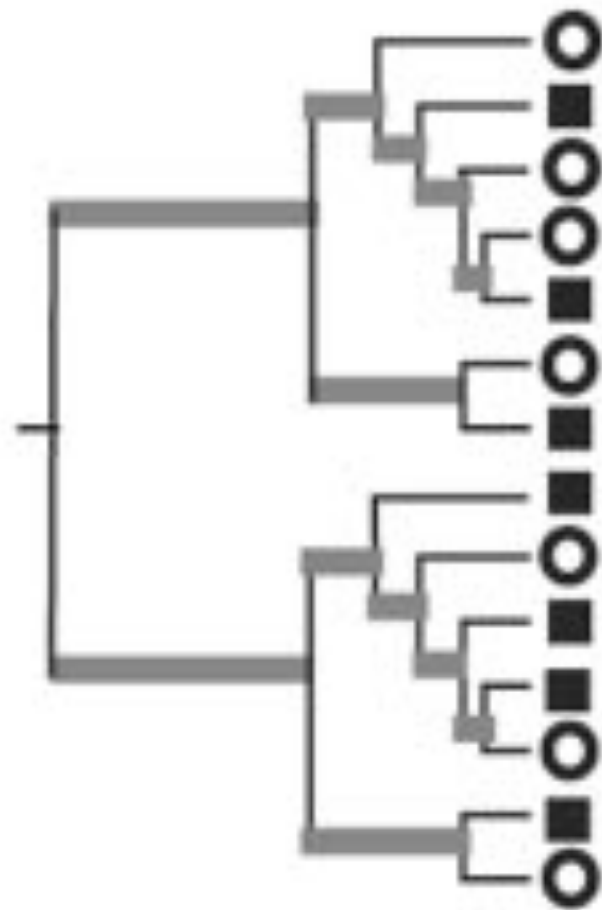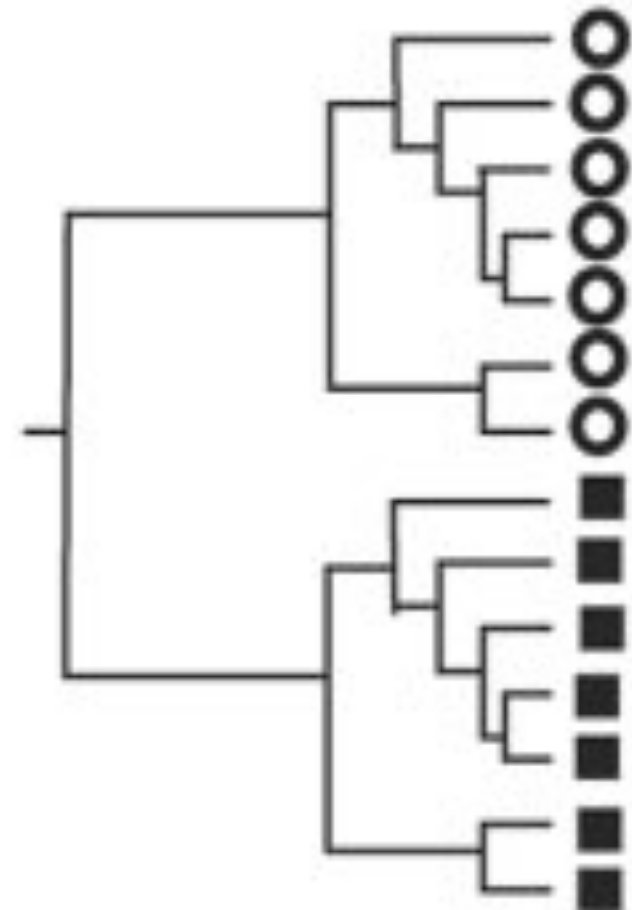
UniFrac

- The phylogenetic distance between sets of taxa in a phylogenetic tree as the fraction of the branch length of the tree that leads to descendants from either one environment or the other, but not both.
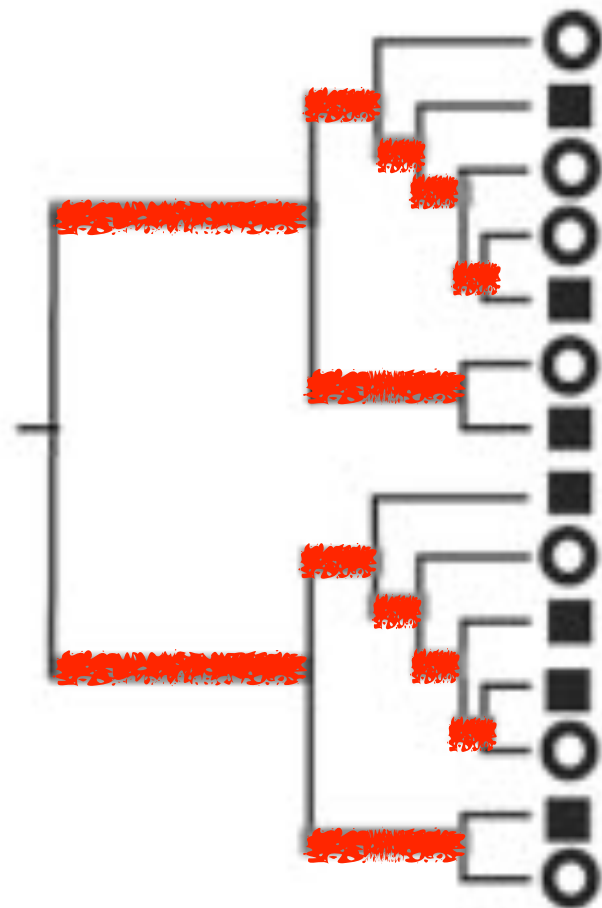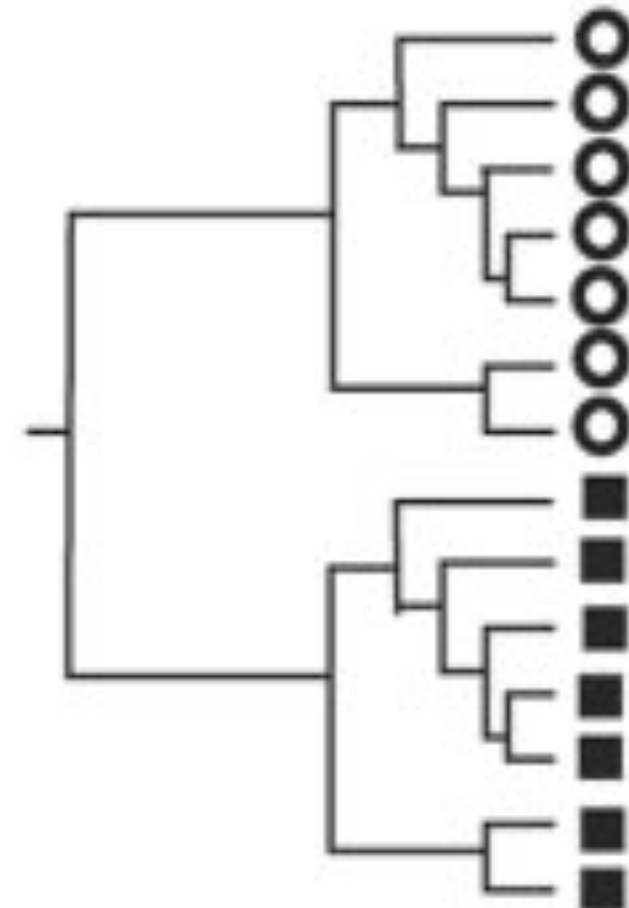
UniFrac

1. Label branches leading to taxa from both samples "shared"

2. Label branches leading to taxa which appears only in one sample "unshared".

3. Unifrac is the fraction of total branch length which is unshared. Alternatively, (the sum of "unshared" branch lengths)/(the sum of all tree branch lengths
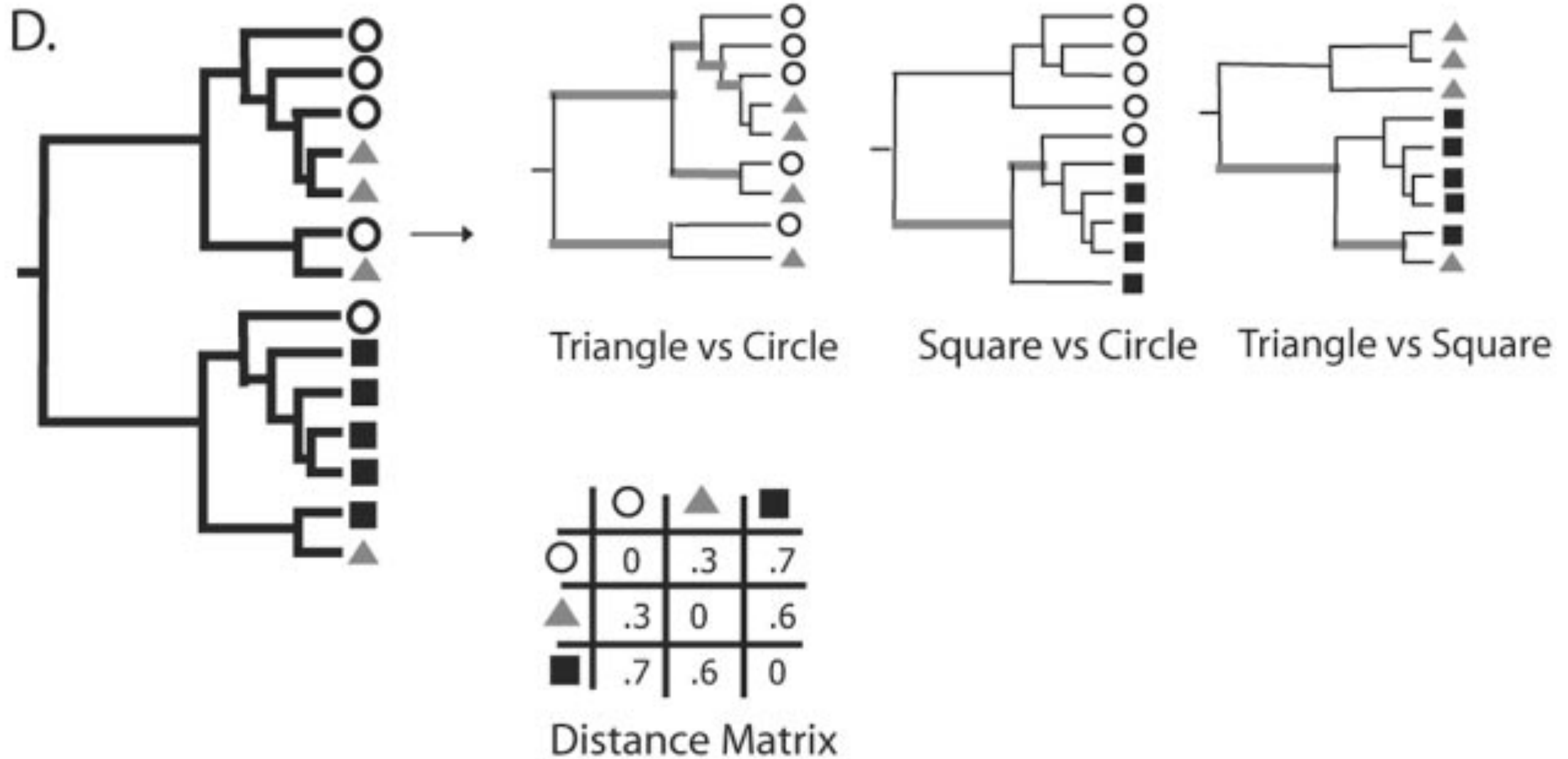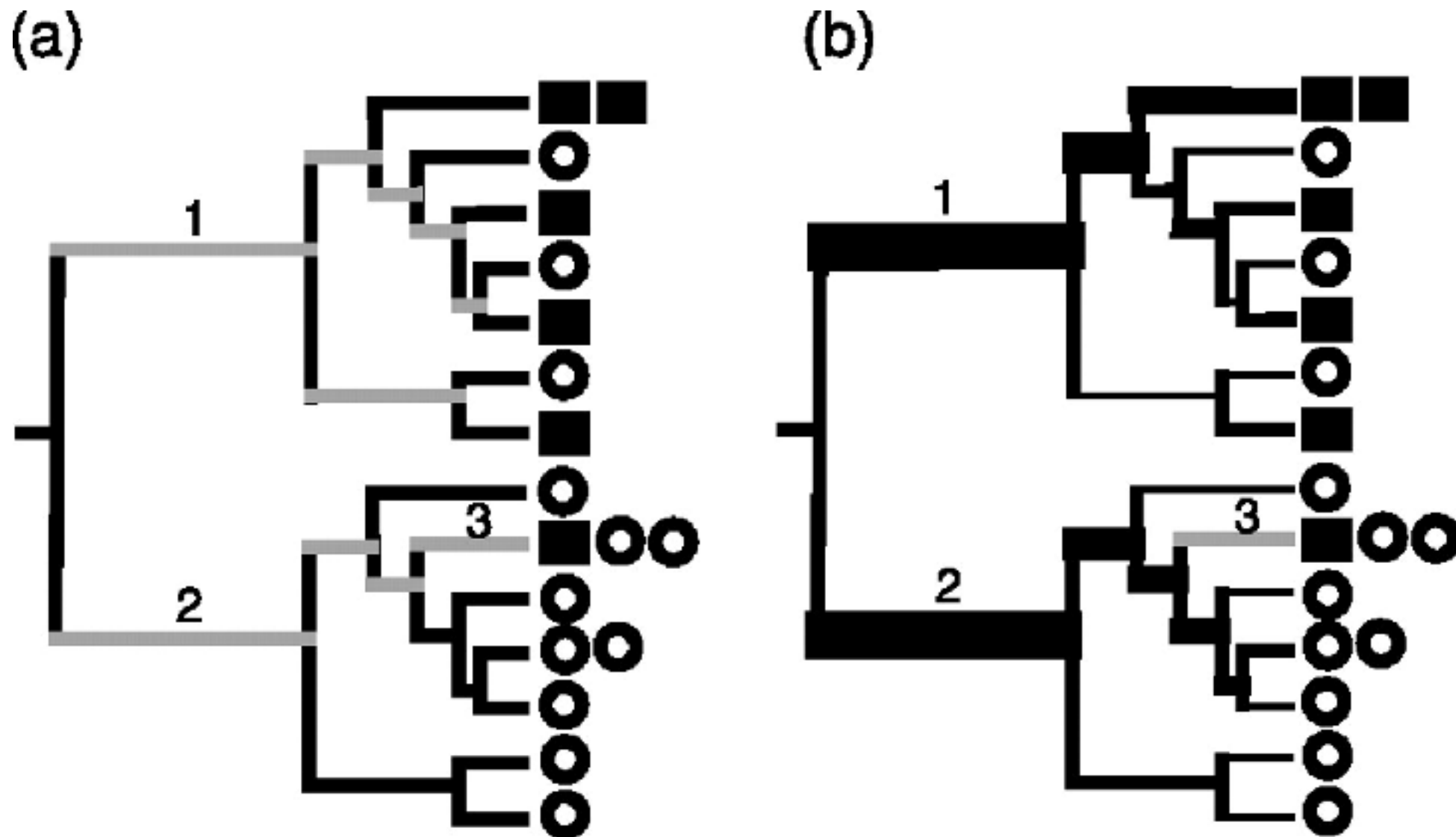
UniFrac

1. Label branches leading to taxa from both samples "shared"

2. Label branches leading to taxa which appears only in one sample "unshared".

3. Unifrac is the fraction of total branch length which is unshared. Alternatively, (the sum of "unshared" branch lengths)/ (the sum of all tree branch lengths

UniFrac

Triangle vs Circle

Square vs Circle

Triangle vs Square

Distance Matrix

	○	▲	■
○	0	.3	.7
▲	.3	0	.6
■	.7	.6	0

Weighted UniFrac

UniFrac: Unweighted vs Weighted

- Unweighted UniFrac

  - Qualitative

  - Sensitive to differences in overall community structure

  - Strongly influenced by differences in rare (low abundance) taxa

- Weighted UniFrac

  - Quantitative

  - Sensitive to differences in high abundance taxa

Generalized UniFrac

- Sensitive to differences in moderately abundant taxa

- Maintains ability to detect differences in high and low abundance taxa