

Universidad de  
los Andes

# INDICES DE DIVERSIDAD EN ECOLOGIA MICROBIANA

Profesor: Diego J. Jimenez PhD.



# ALGUNAS DEFINICIONES Y CONCEPTOS EN ECOLOGÍA MICROBIANA

Ecosistema

Abundancia

Coexistencia

Riqueza

Sucesión

Diversidad

Nicho

Cepa

Especie

Biosfera

Población

Comunidad

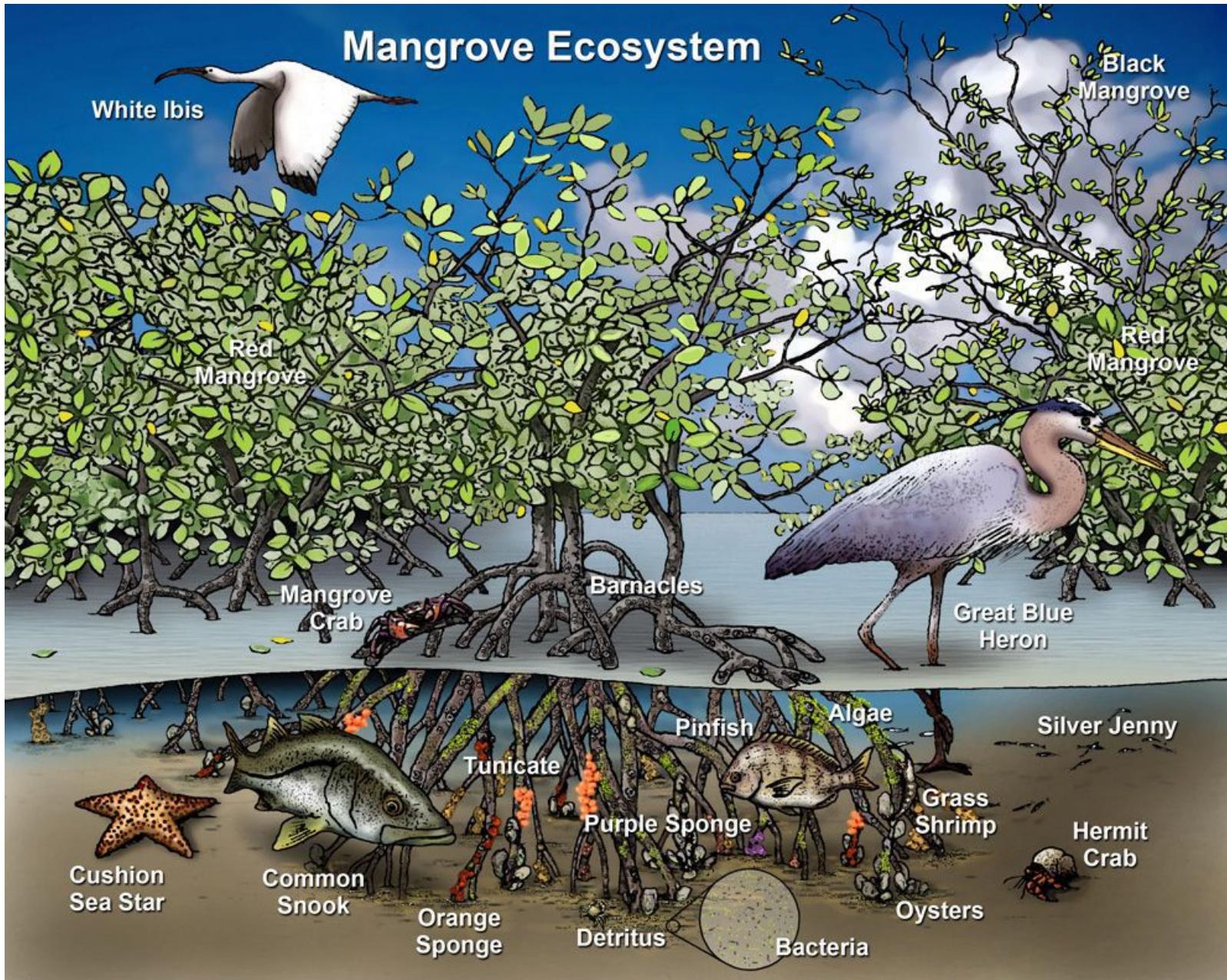
Metataxonómica

Biomasa

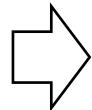
Procariontas

Biogeografía

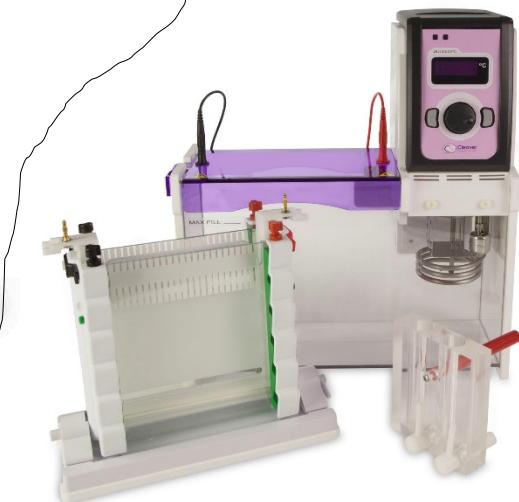
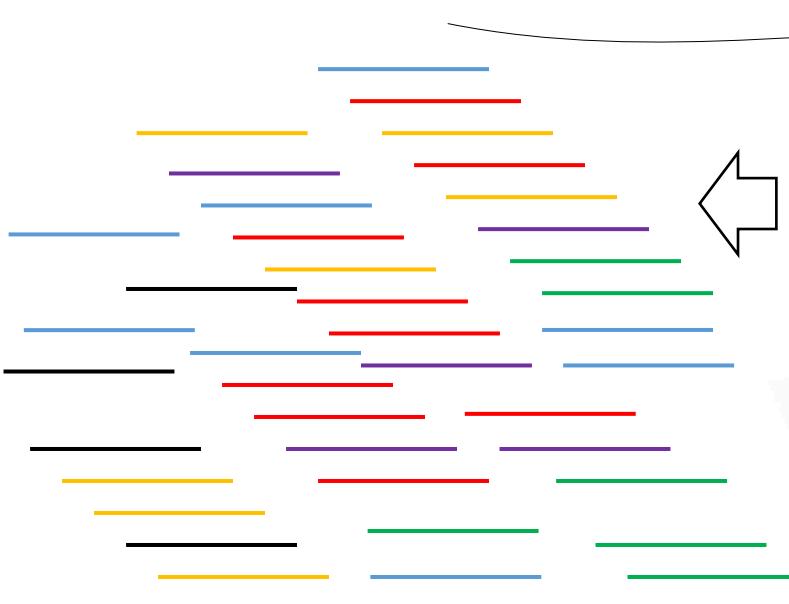
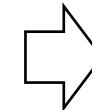
OTUs



# Y como investigo la diversidad y taxonomía de una comunidad microbiana?



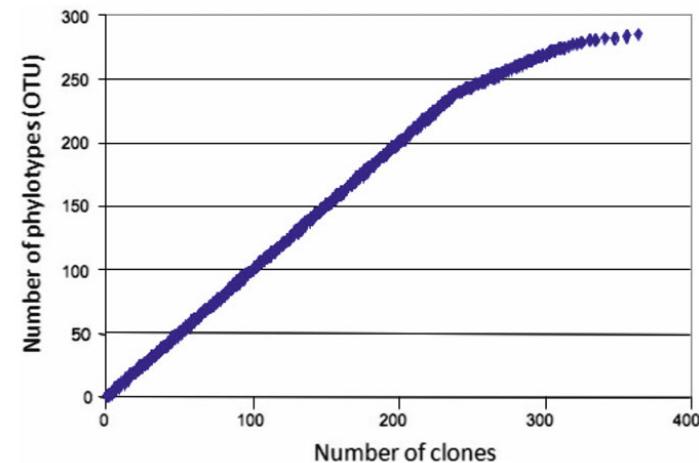
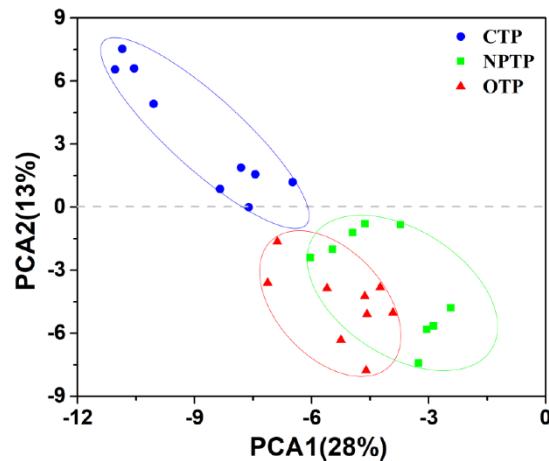
16S rRNA  
ITS1



# Como medir la diversidad microbiana?

1. Numero de OTUs (riqueza) y frecuencia
2. Abundancia absoluta / relativa (proporción)
3. Curvas de rarefacción
4. Alpha diversidad
5. Beta diversidad

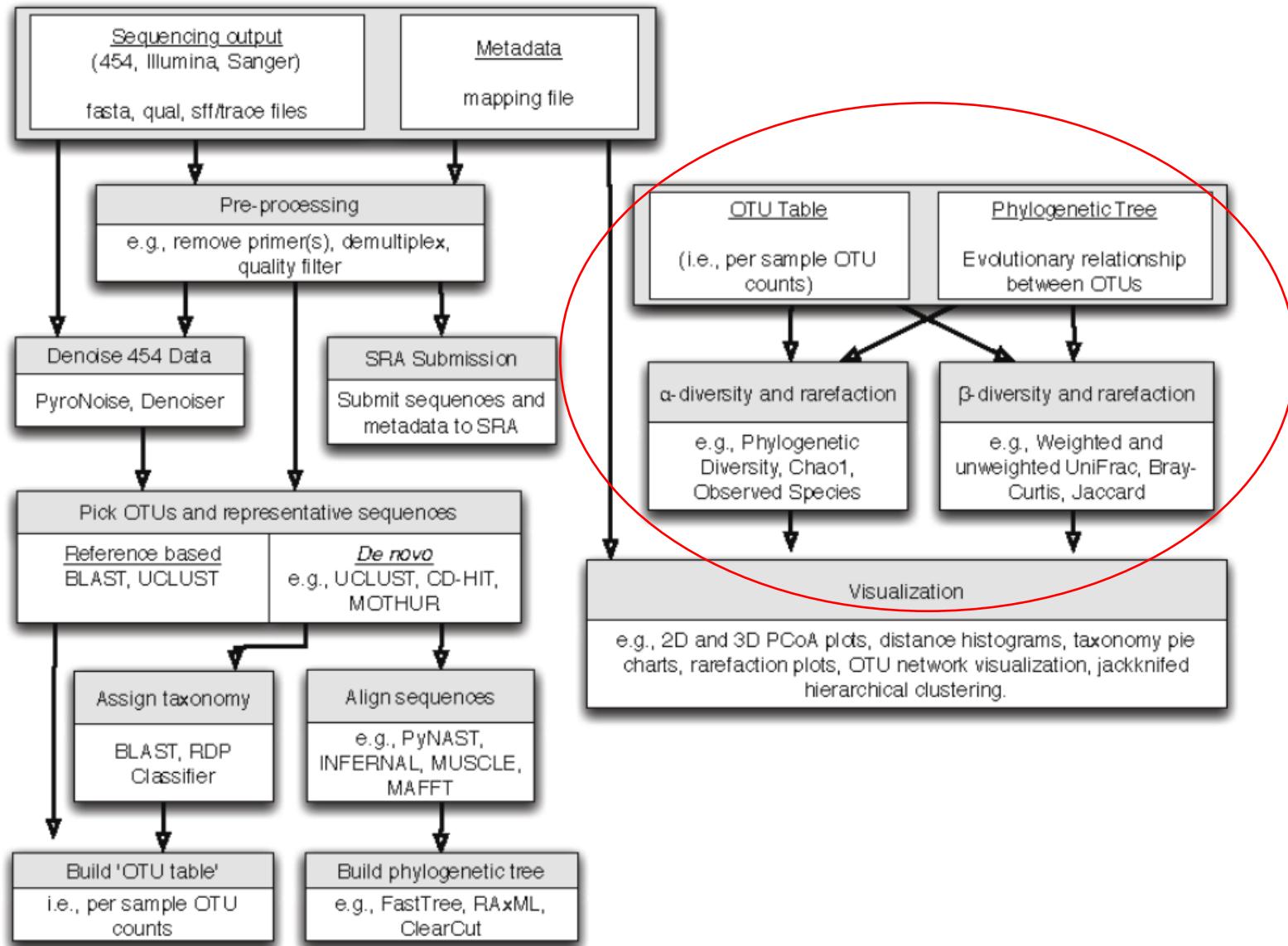
$$H' = - \sum_{i=1}^S p(i) \ln(p(i))$$



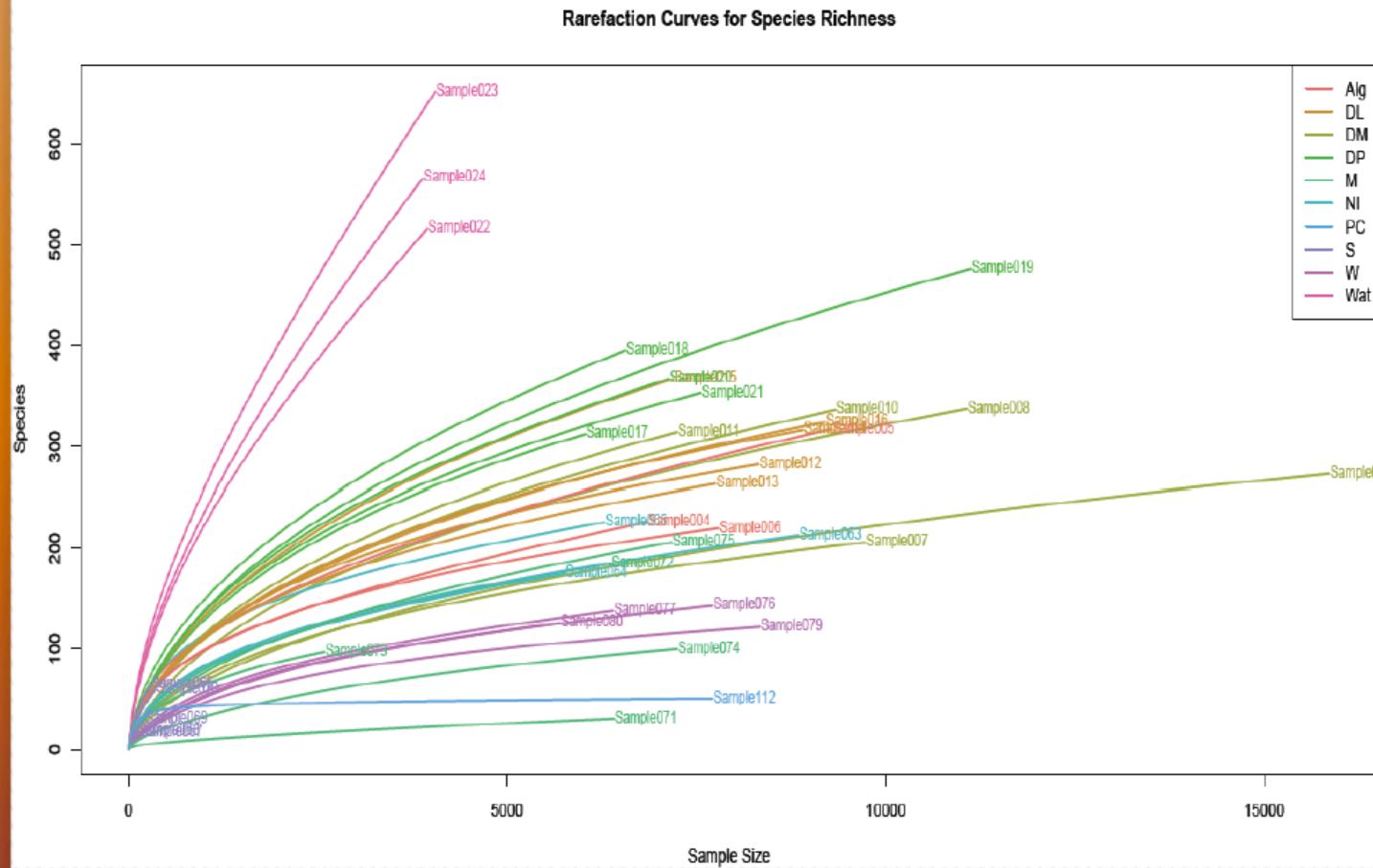
**OTU:** a group of phylogenetically related organisms used in a study without specifying its taxonomic rank



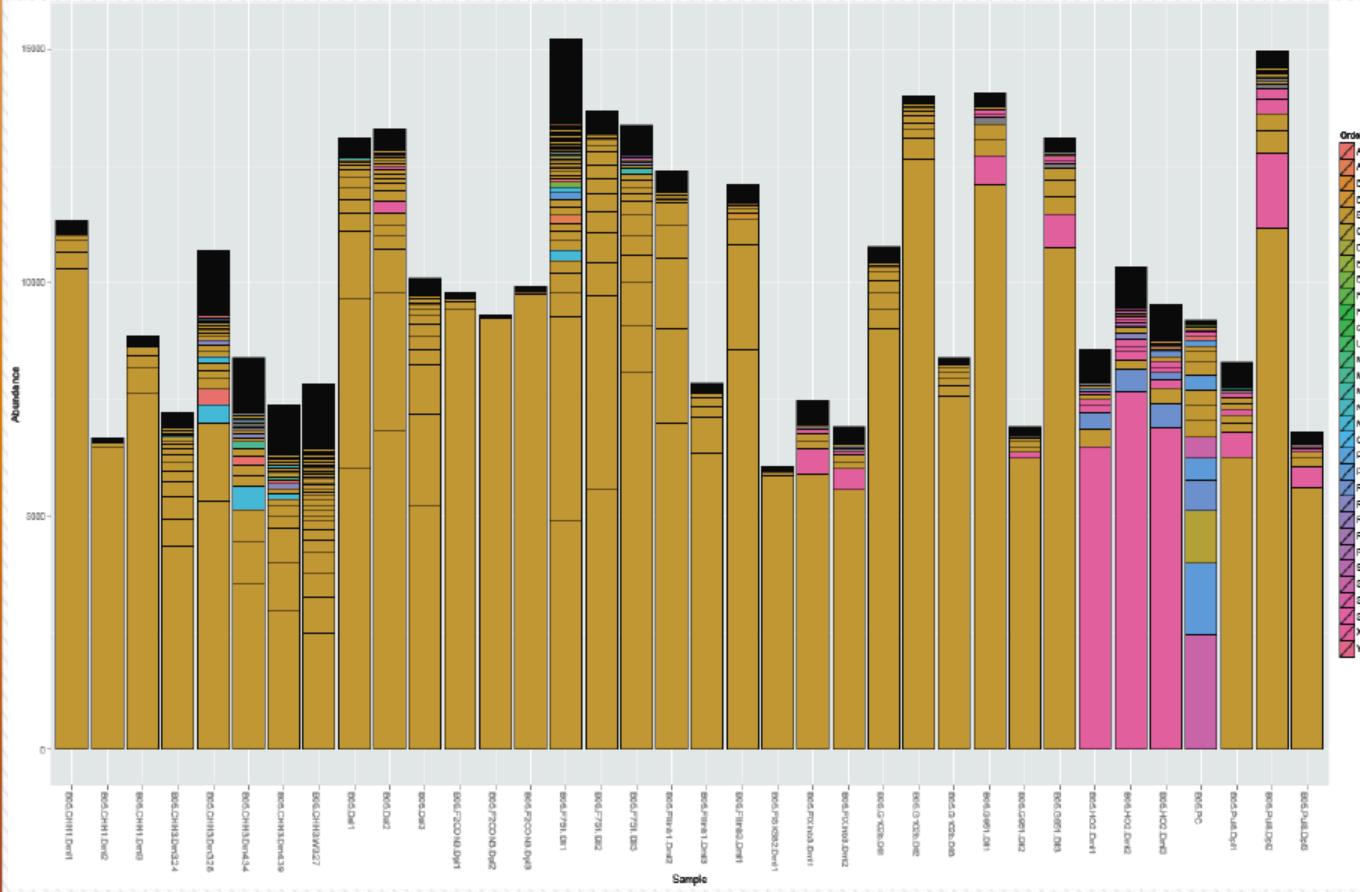
Quantitative Insights Into Microbial Ecology



# Rarefaction curves



# Variation in reads counts



# Rarefaction

Rarefying was first recommended for microbiome counts in order to moderate differences in the presence of rare OTUs

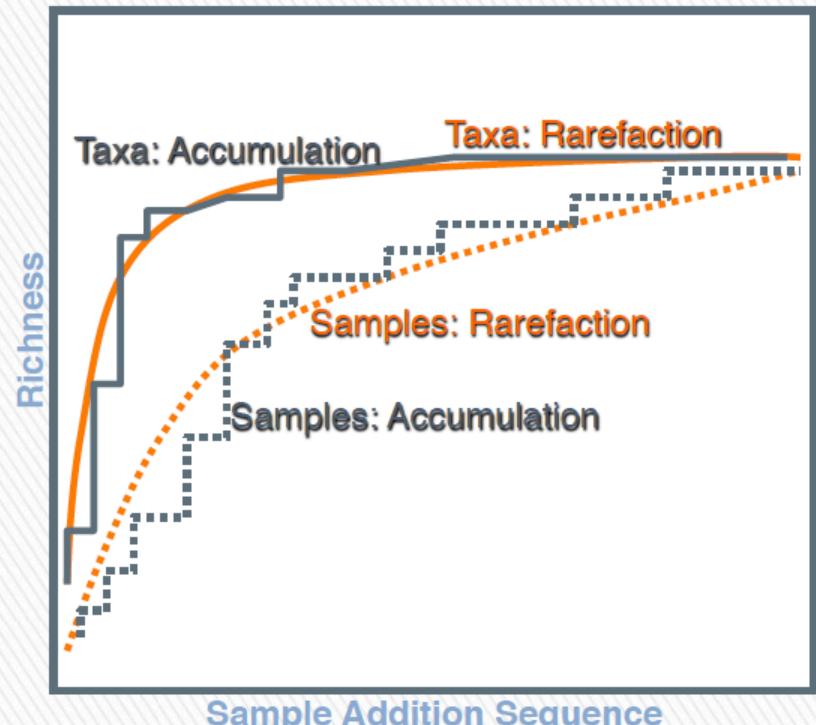
(Lozupone *et al.* 2011 ISME J)

## Goal:

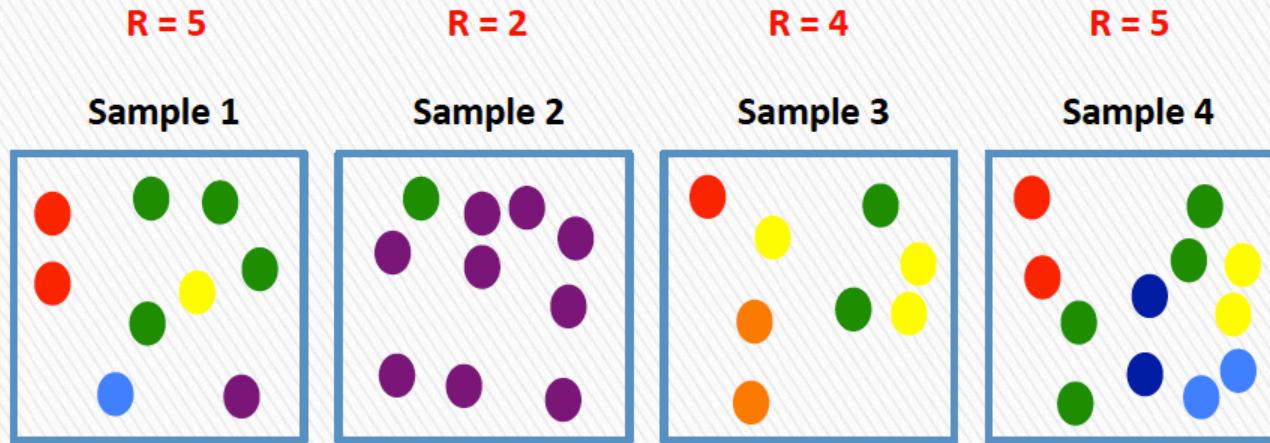
- Standardize unequal sequencing effort.
- Enable similarity comparisons along a range of samples or a gradient.
- Enable comparison of different runs or replicates.

## Procedure:

- Determine the minimum sequencing depth.
- Subsample without replacement sequences from the larger libraries so that all have the same smallest size.
- Note that the term is a bit misleading as this step should really be called "subsampling to a given depth".



# Alpha Diversity: richness ( $R$ )



## SPECIES RICHNESS ( $S$ ) ESTIMATORS:

- **OTU richness** – count of different species/OTUs
- **Observed Species** – count of unique OTUs in each sample
- **Chao1 index** – estimate diversity from abundance data (importance of rare OTUs)
- **ACE index** ...



# Species richness indices

Let  $S_0$  be the number of taxa observed at least once in a sample,

$a_0$  the unknown number of species present in the community but not observed.

1. OTU richness:  $R = S_0$  (no correction for taxa not observed);

2. Chao-1 index: assumes that the number of observations for a taxa has a Poisson distribution and corrects for variance;

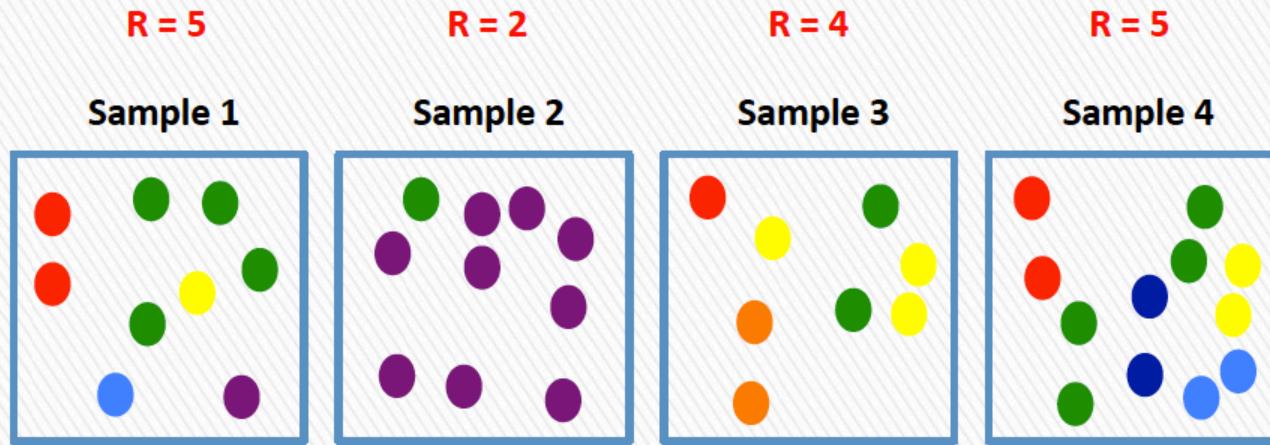
$$R = S_0 + a_0 \text{ i.e. } S_0 + a_1(a_1 - 1)/(2a_2 + 1)$$

3. ACE (abundance-based coverage estimators): involves an arbitrary abundance threshold to label  $S_{\text{abun}}$  as the number of abundant taxa,  $S_{\text{rare}}$  as the number of rare taxa; The expression basically inflates the number of rare taxa and inflates again the number of taxa with abundance 1.

$$S_{\text{abun}} + S_{\text{rare}}/C_{\text{ace}} + a_1/C_{\text{ace}} * \gamma^2$$

$$R = S_0 + a_0 \text{ i.e. } \gamma^2 = \max S_{\text{rare}}/C_{\text{ace}} \sum_{i=1}^{10} (i(i-1)a_i/(N_{\text{rare}}(N_{\text{rare}}-1)) - 1), 0$$

# Alpha Diversity: within sample diversity



## SPECIES RICHNESS ESTIMATORS

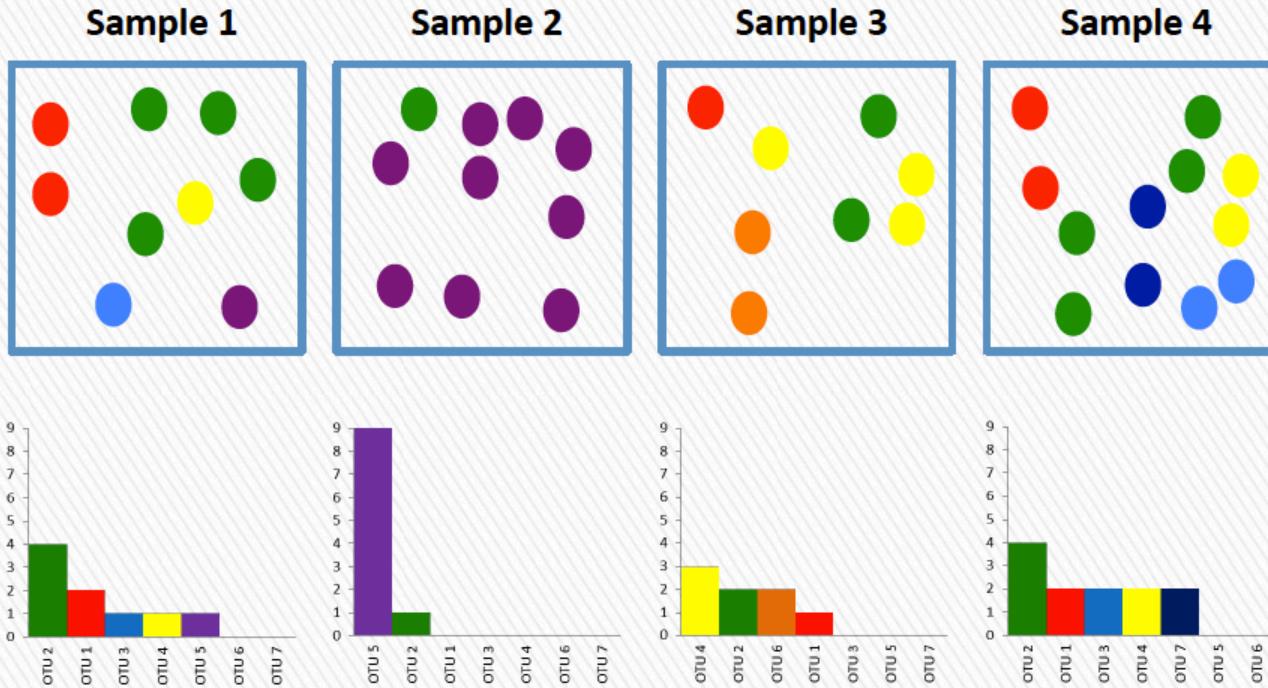
**RICHNESS and EVENNESS ESTIMATORS:** the calculated value of diversity increases both when the number of species increases and when evenness increases.

- **Information statistics:** Shannon-Wiener, Shannon-Weaver, Shannon entropy
- **Dominance indices:** Inverse Simpson, Gini–Simpson, Berger–Parker index



# Alpha Diversity: relative abundances

Refers to how common or rare a species is relative to other species in a community.

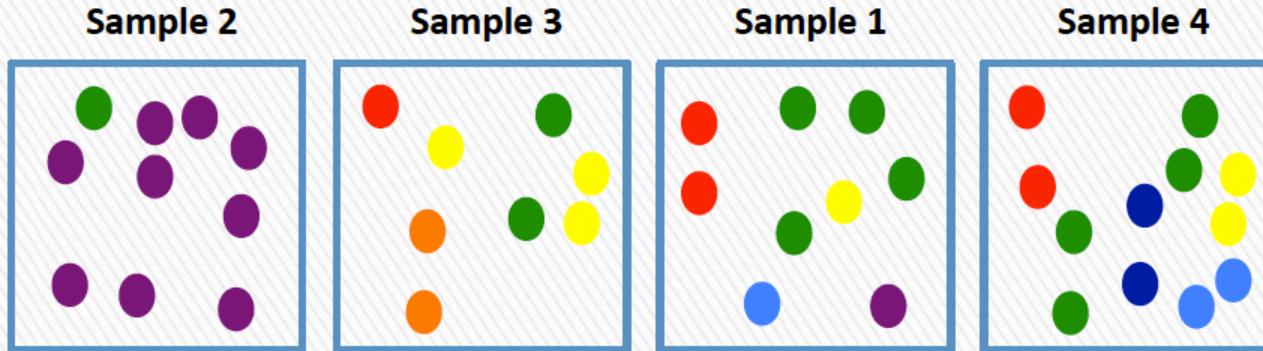


Relative species abundance distributions are graphed as rank-abundance diagrams.

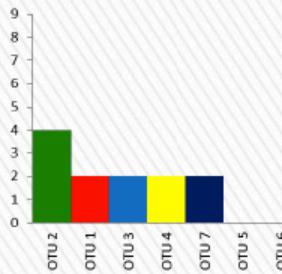
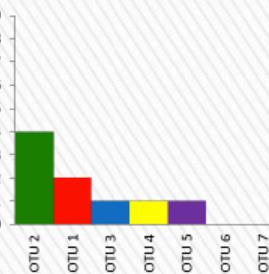
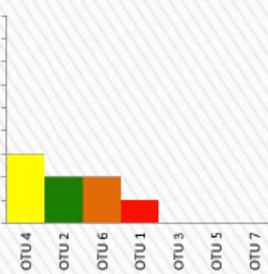
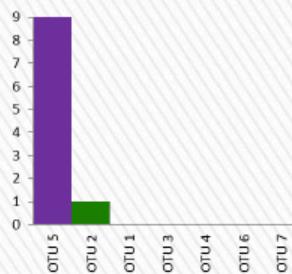


# Alpha Diversity: species evenness

**Species evenness** refers to how equally abundant species in an environment are.



----- Increasing richness and abundance ----- ➡



# **Shannon Alpha Diversity Index (H)**

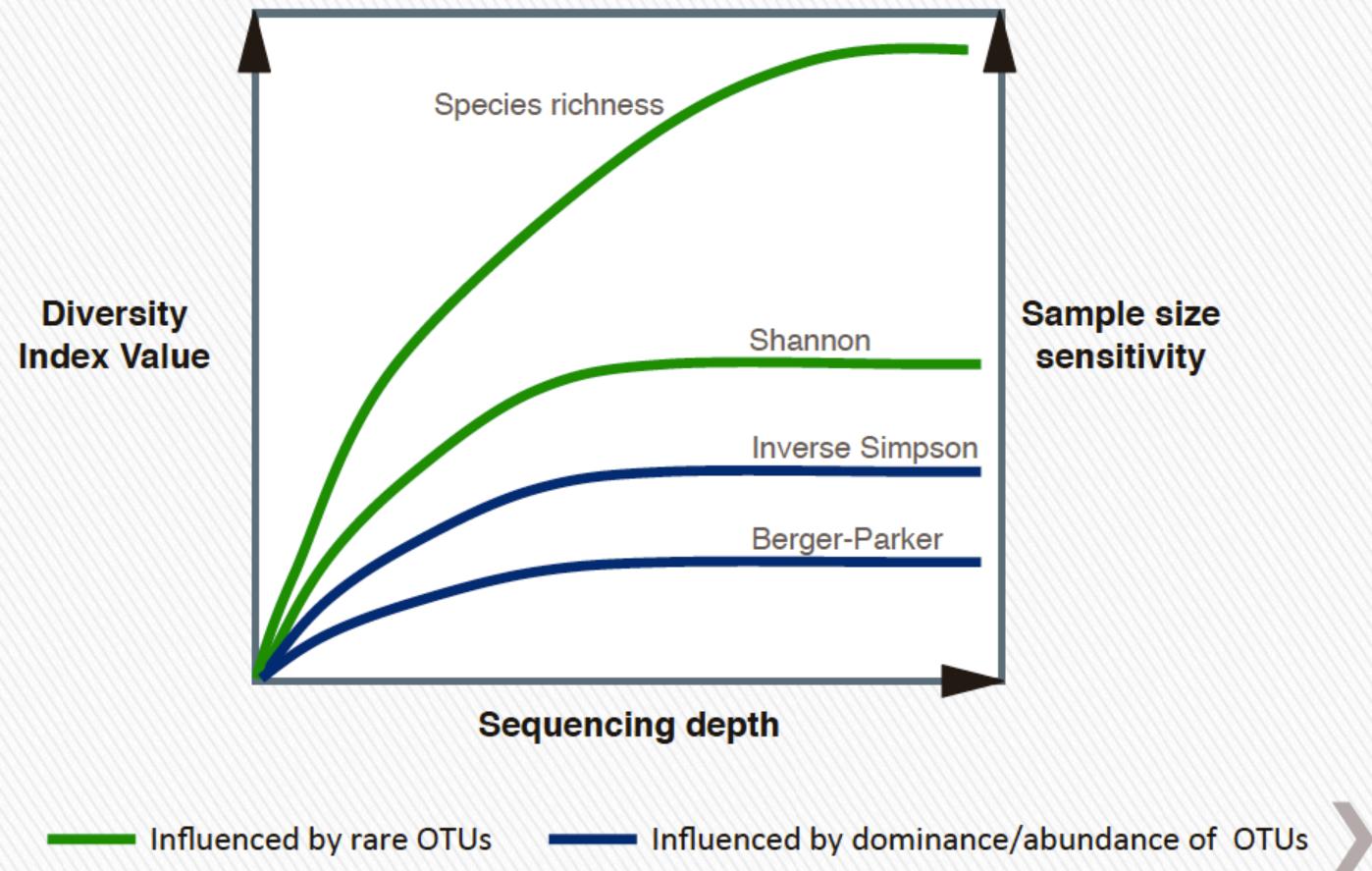
$$H' = - \sum_{i=1}^S p(i) \ln(p(i))$$

- Entre mayor sea el valor, la comunidad es mas diversa
- El valor es 0 si una sola especie esta presente
- No tiene en cuenta la taxonomía de las especies

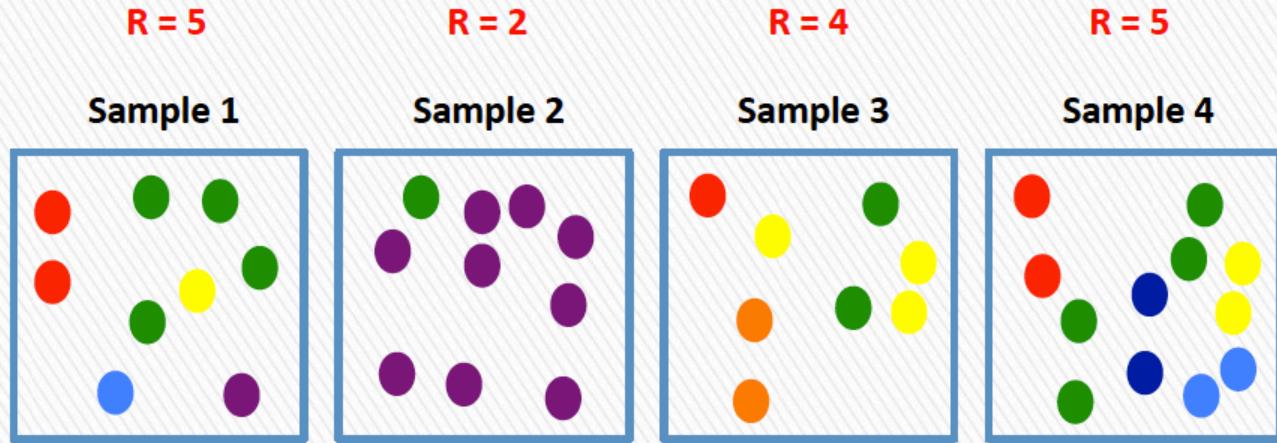
Especie	Abundancia	p(i)	In p(i)	p(i) x In p(i)
Sp 1	13	0,26	-0,58	-0,15
Sp 2	20	0,4	-0,39	-0,16
Sp 3	2	0,04	-1,39	-0,05
Sp 4	5	0,1	-1	-0,1
Sp 5	10	0,2	-0,69	-0,139
<b>N</b>	<b>50</b>			<b>H = -0,60....</b>

**H = 0,60**

# Alpha Diversity: indices' performance



# Alpha Diversity: phylogenetic diversity



SPECIES RICHNESS ESTIMATORS

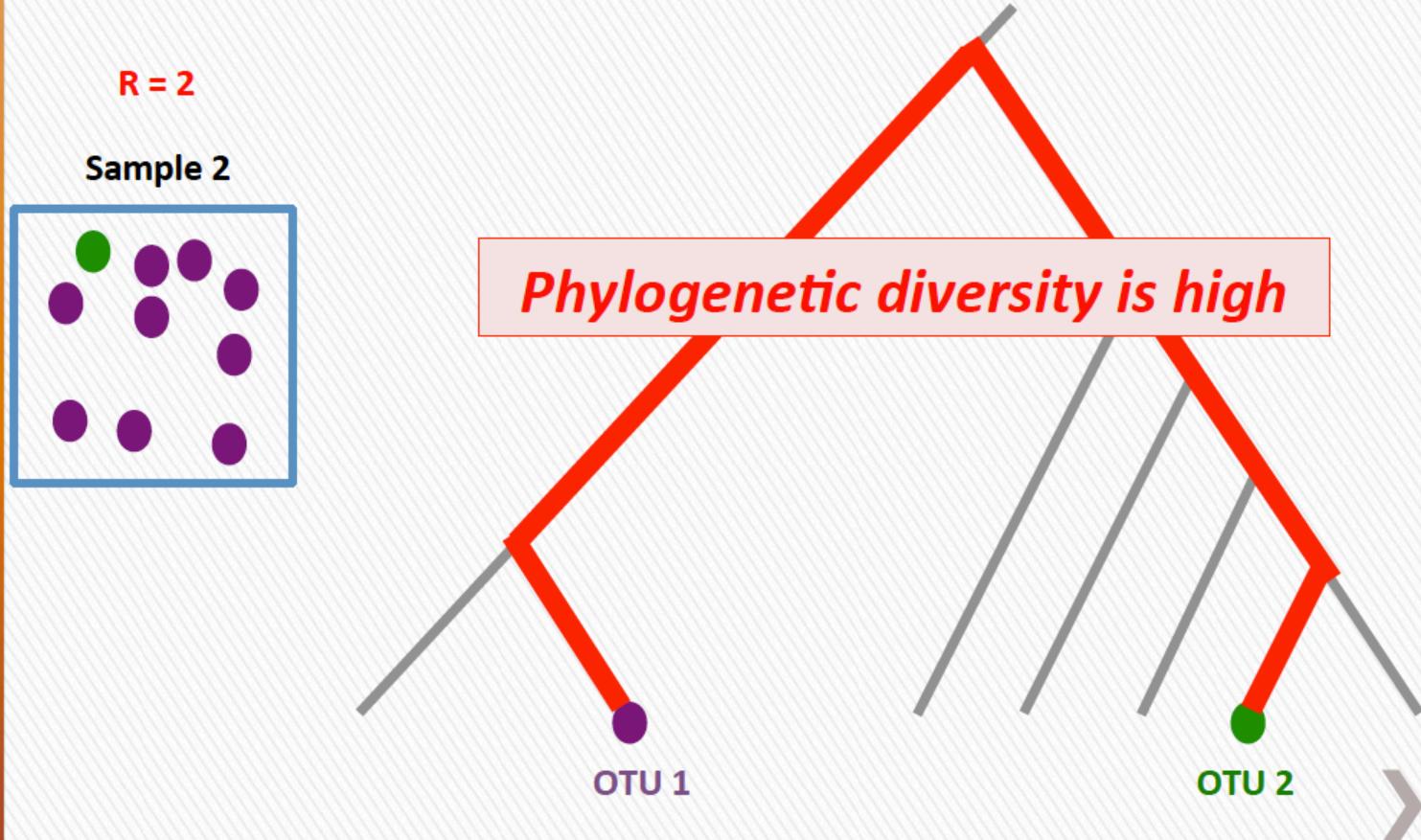
RICHNESS and EVENNESS ESTIMATORS

PHYLOGENETIC RICHNESS ESTIMATOR:

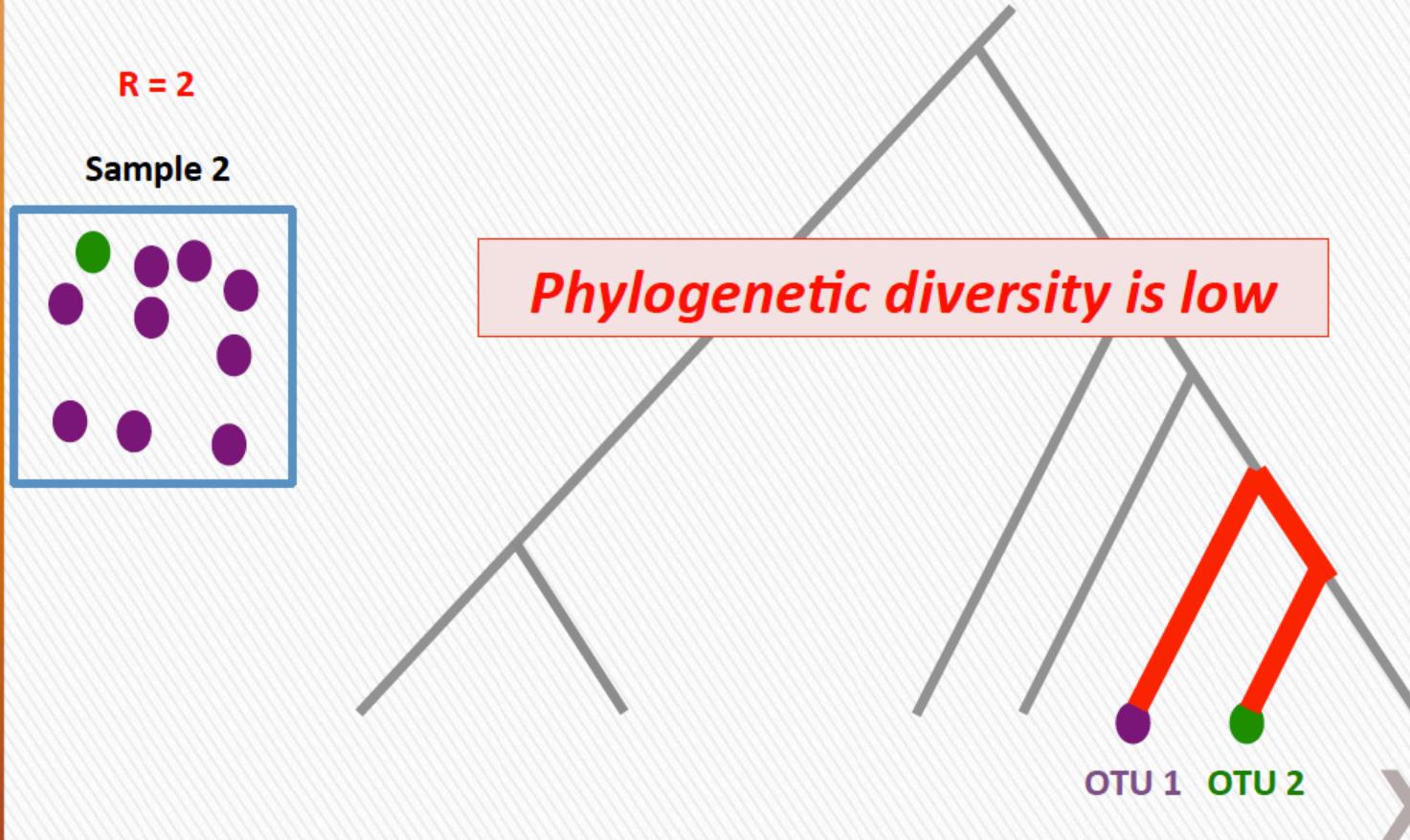
- **Phylogenetic diversity (PD)** – takes into consideration the phylogeny of microbes to estimate diversity across a tree



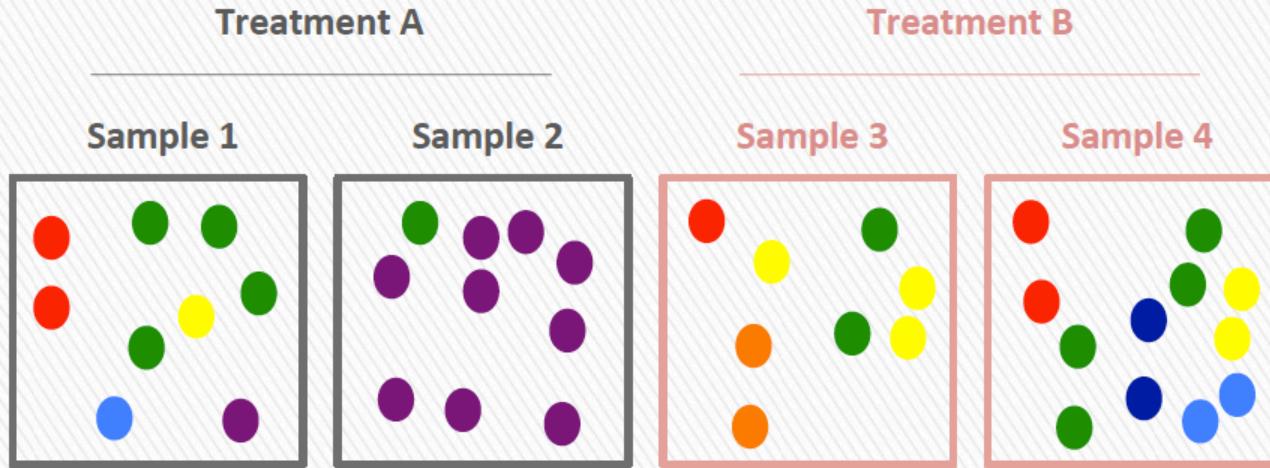
# Alpha Diversity: phylogenetic diversity



# Alpha Diversity: phylogenetic diversity



# Beta Diversity: between sample diversity



**Main concepts of beta-diversity:**

Question – What is the influence of treatments A & B?

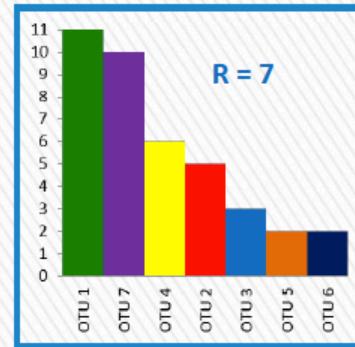
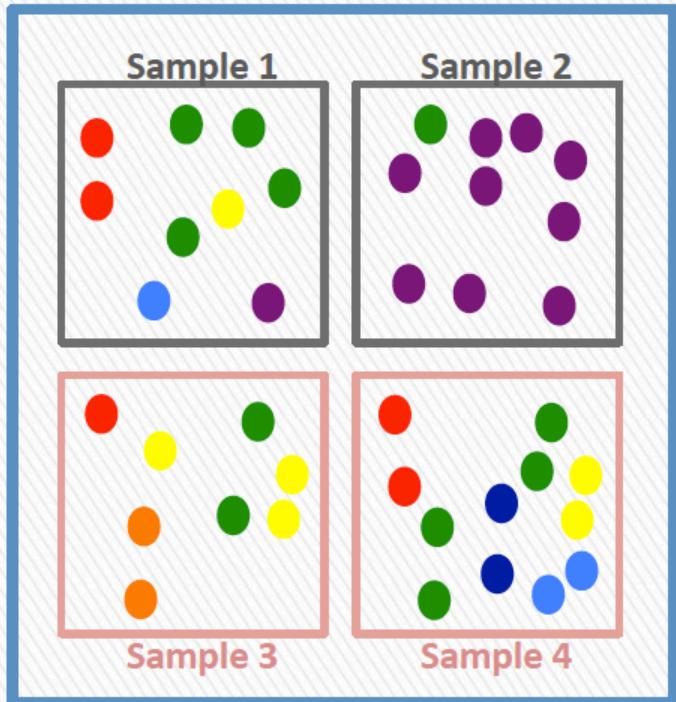
What is the species diversity along transects & gradients?

**Notion of similarity:**

Species differences among samples are positively correlated to  $\beta$ -diversity values and inversely correlated to similarity.

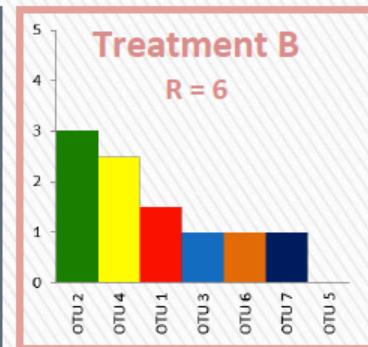
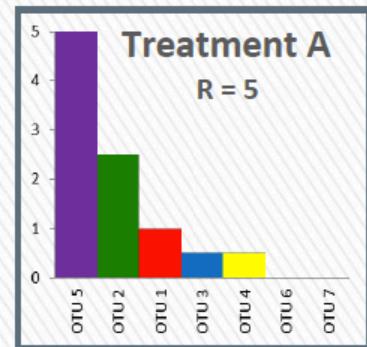


# Beta Diversity *stricto sensu*



$$\beta = \gamma / \alpha$$

Mean diversity  
per treatment



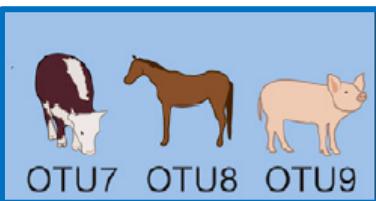
# Beta Diversity: between sample diversity

	Presence/Absence	Abundance
<b>Without a phylogenetic tree</b>	<ul style="list-style-type: none"><li>• Jaccard</li></ul>	<ul style="list-style-type: none"><li>• Bray-Curtis (PCoA)</li><li>• Euclidean (PCA)</li></ul>
<b>With a phylogenetic tree</b>	<ul style="list-style-type: none"><li>• Unweighted UniFrac</li><li>• Comdist</li></ul>	<ul style="list-style-type: none"><li>• Weighted UniFrac</li><li>• Comdist</li></ul>

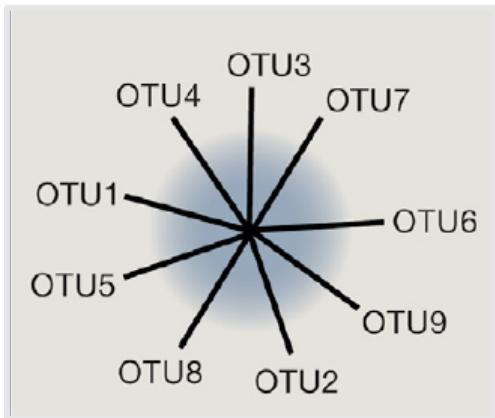


# Operational Taxonomic Units binning

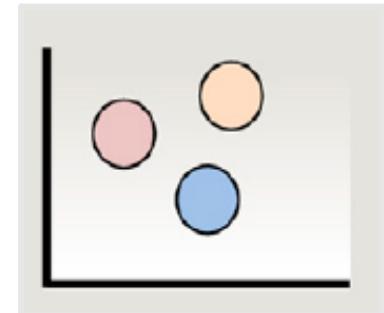
Sample



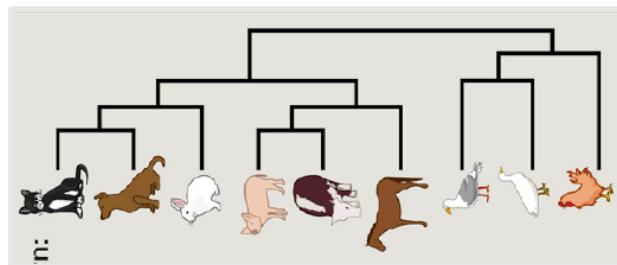
OTU unknown phylogeny



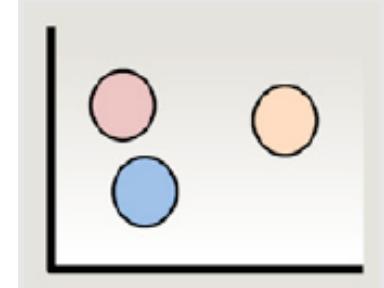
Equidistant group

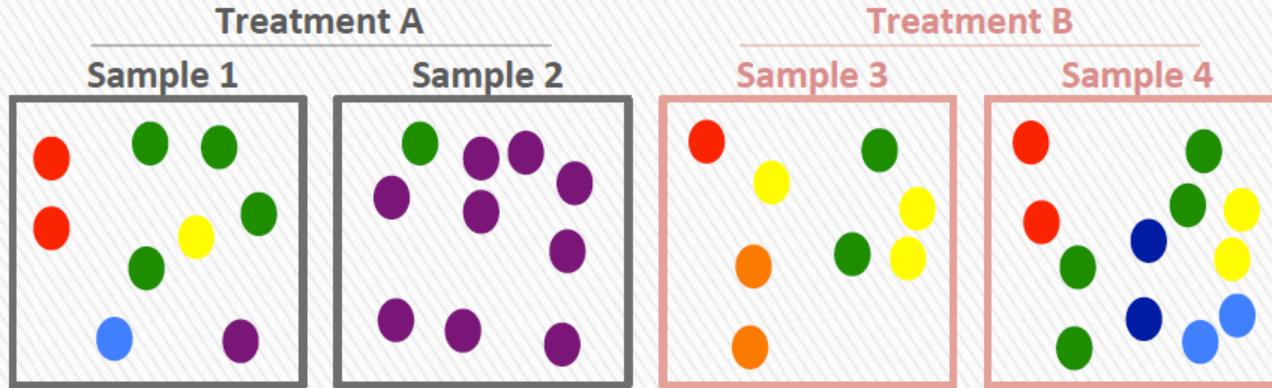


OTU known phylogeny



A more distant group





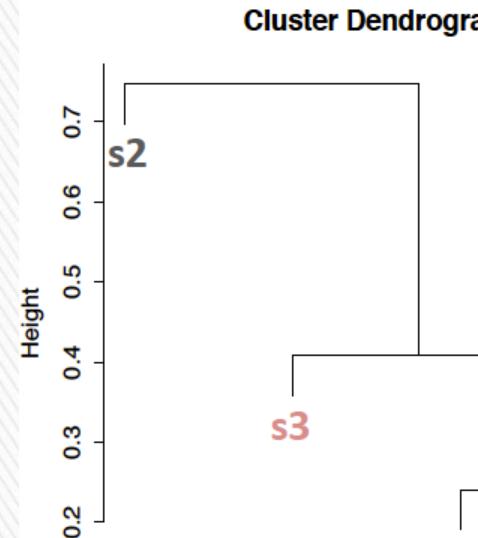
OTU\_table

	Sample1	Sample2	Sample3	Sample4
OTU 1	2	0	1	2
OTU 2	4	1	2	4
OTU 3	1	0	0	2
OTU 4	1	0	3	2
OTU 5	1	9	0	0
OTU 6	0	0	2	0
OTU 7	0	0	0	2

MapFile

#SampleID	Treatment
Sample1	A
Sample2	A
Sample3	B
Sample4	B

Marker-based metagenomic tutorial



**Hierarchical clustering**  
tutorial.data.bray  
hclust (\*, "average")

# Beta Diversity: visualization / ordination

