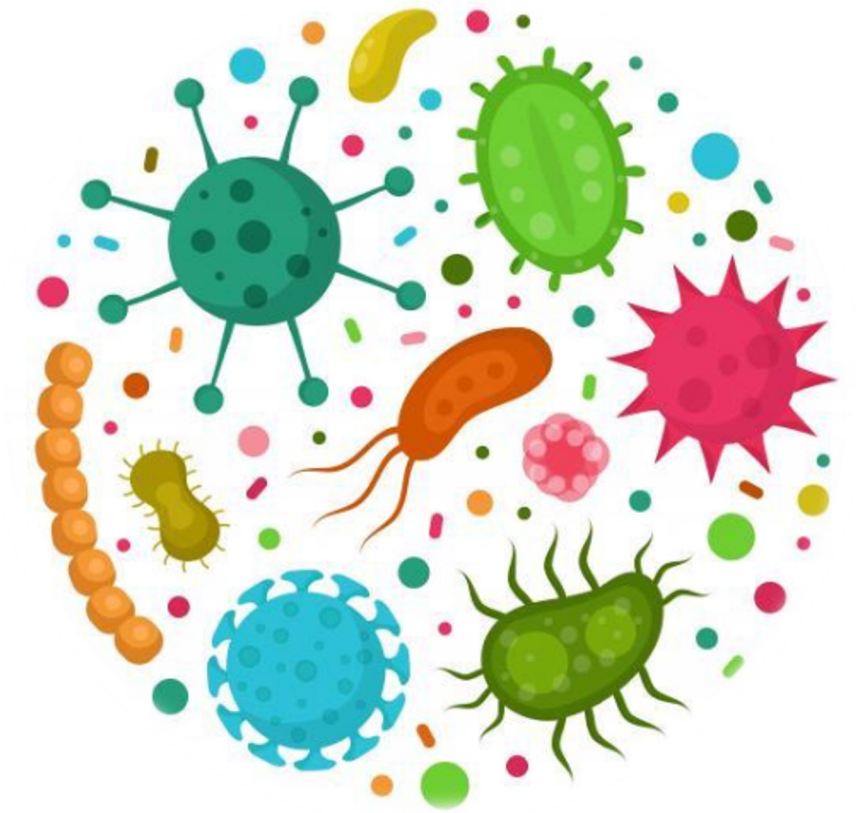
α-diversity

j3
– 01.10.25 –



ANF METABIODIV

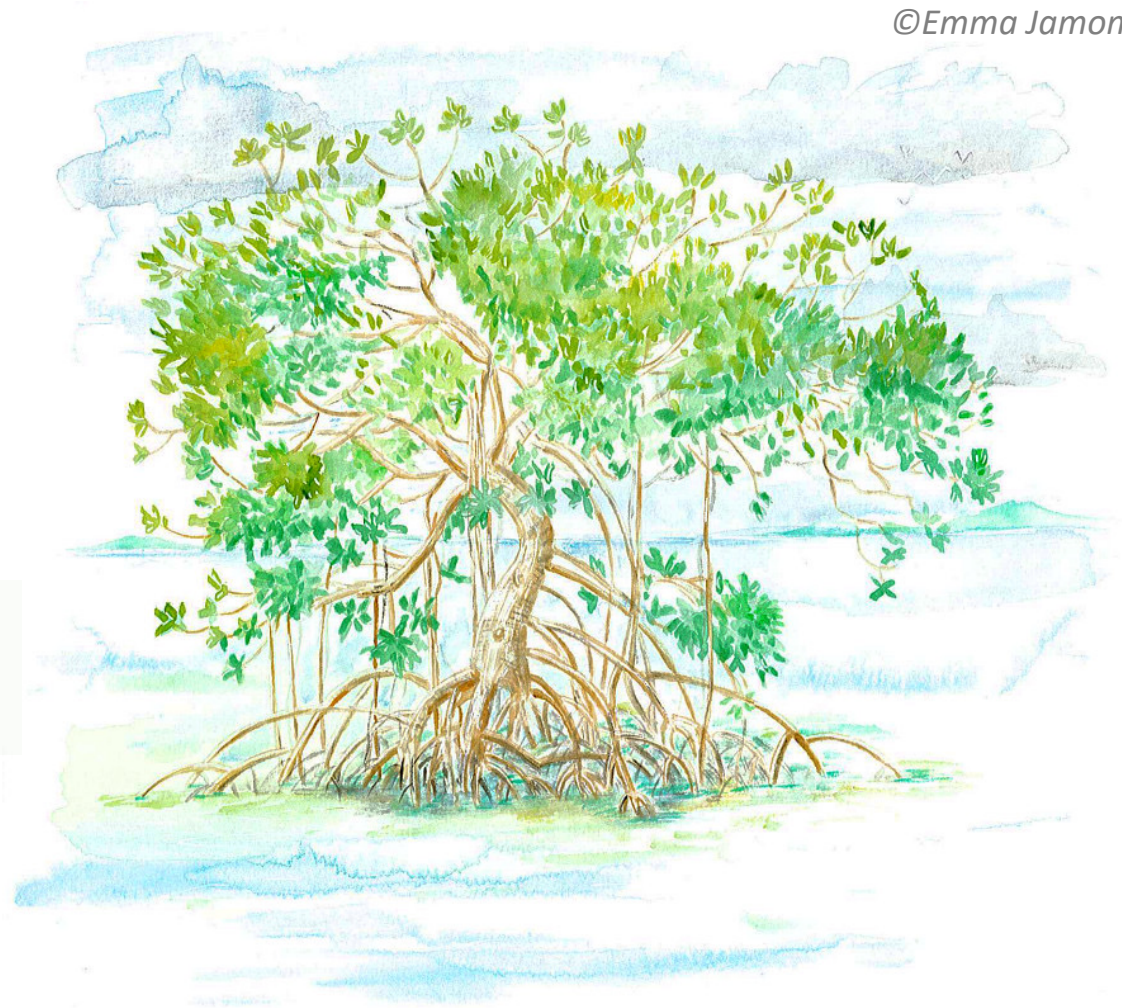
Bio-informatique & Sciences de l'Environnement : Exploration de la Diversité Taxonomique des Ecosystèmes par Metabarcoding

Institut Méditerranéen d'Océanologie

Institut de Recherche pour le Développement FRANCE

cnrs

What is Biodiversity?



The **variety** and **variability** among living organisms on site, ecosystem and their **interractions** between them

Diversity can be use to describe **variation** in several forms:

- **Genetic**
- **Taxonomic**
- **Functional** group (e.g. nitrogen-fixing)

Why measuring Biodiversity?

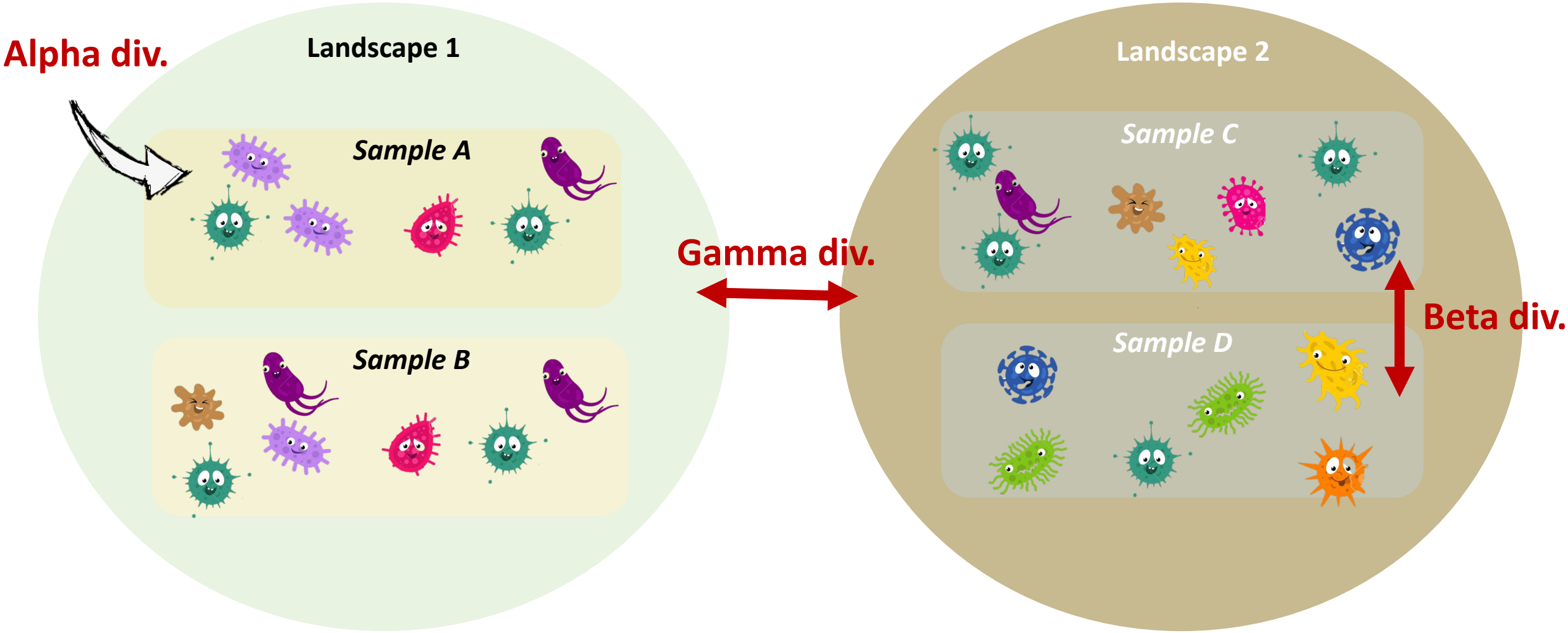
proxy

Biodiversity → **Ecosystem health**

In general, diverse communities are believed to have increased stability, increased productivity, and resistance to invasion and other disturbances

Alpha vs Beta vs Gamma Diversity (Whittaker, 1972)

Alpha, beta and gamma diversity are three types of biodiversity measures described over a special scale



α-Biodiversity ?

α-diversity is local diversity, measured within a closed system

→ **The diversity within an habitat of fixed size**

α-biodiversity has **two components**

- **Richness**
- **Evenness**

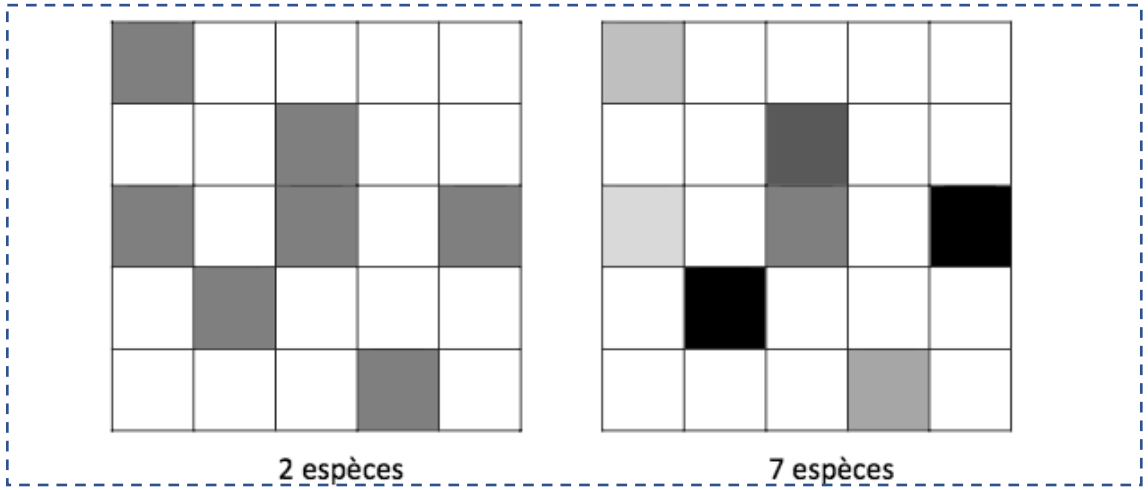
And another one…

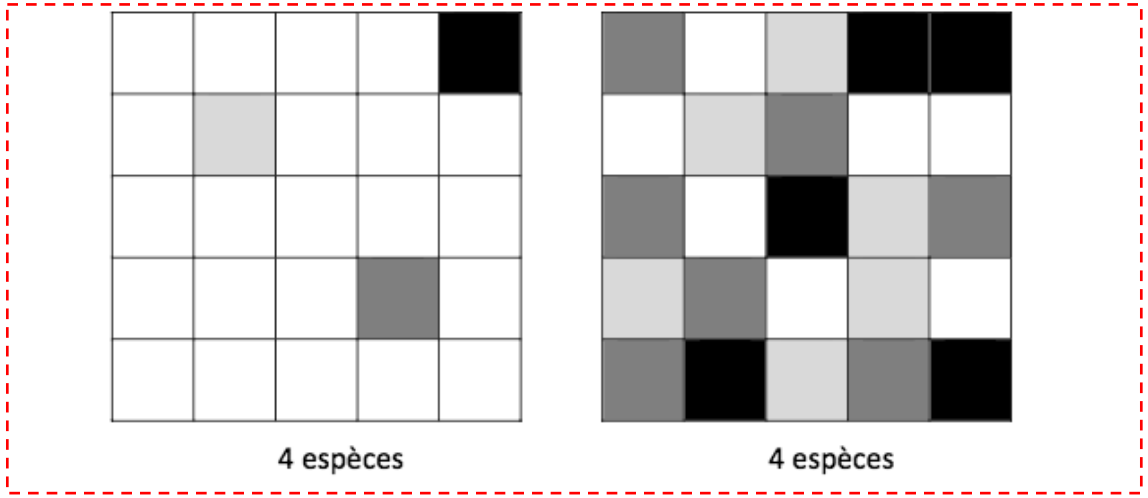
- **Disparity (A. Stirling)**

→ **Follow the evolution of populations over time, but also to compare them between the stations studied**

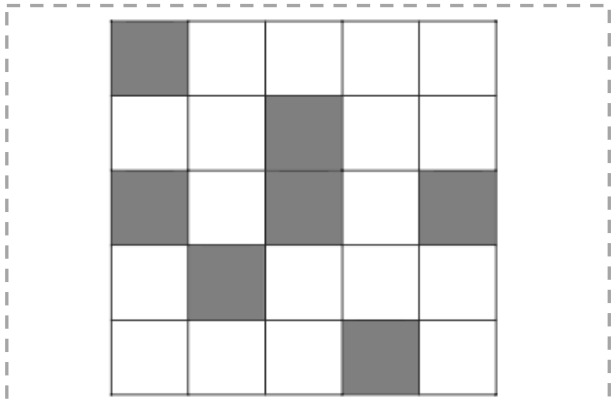
Components of Diversity

Species number

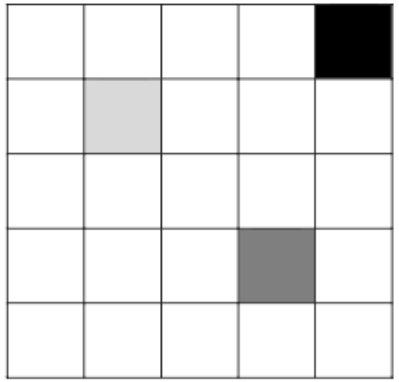




Species evenness

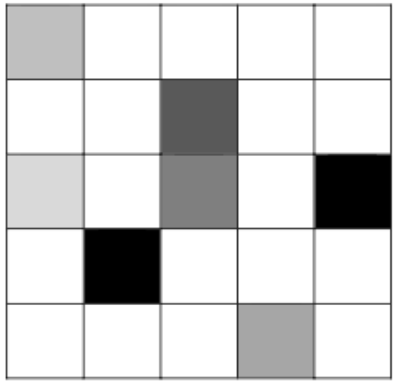


2 espèces

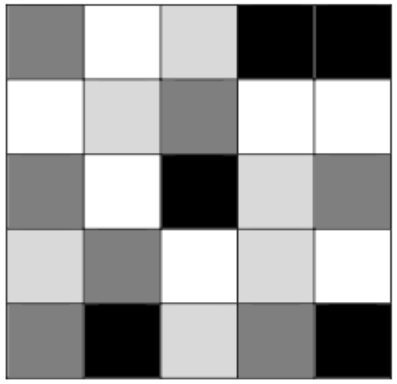


4 espèces

?



7 espèces



4 espèces

?

Two components are necessary for describing diversity

- **Richness**
- **Evenness**

Definition

Specific Richness (**S**) = The measurement of the **number of species** present in a location/studied system

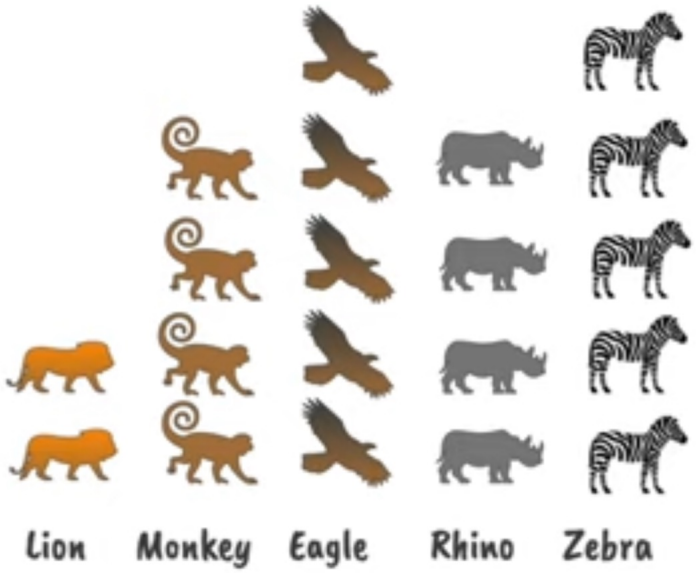
- The more species present, the richer sample is
- Gives equal weight to species which have many/few individuals

Evenness (equitability) = **Relative population of each species**

- Species represented by many individuals or by few ones do not give the same contribution

Evenness index is independent of Richness!!

Higher biodiversity

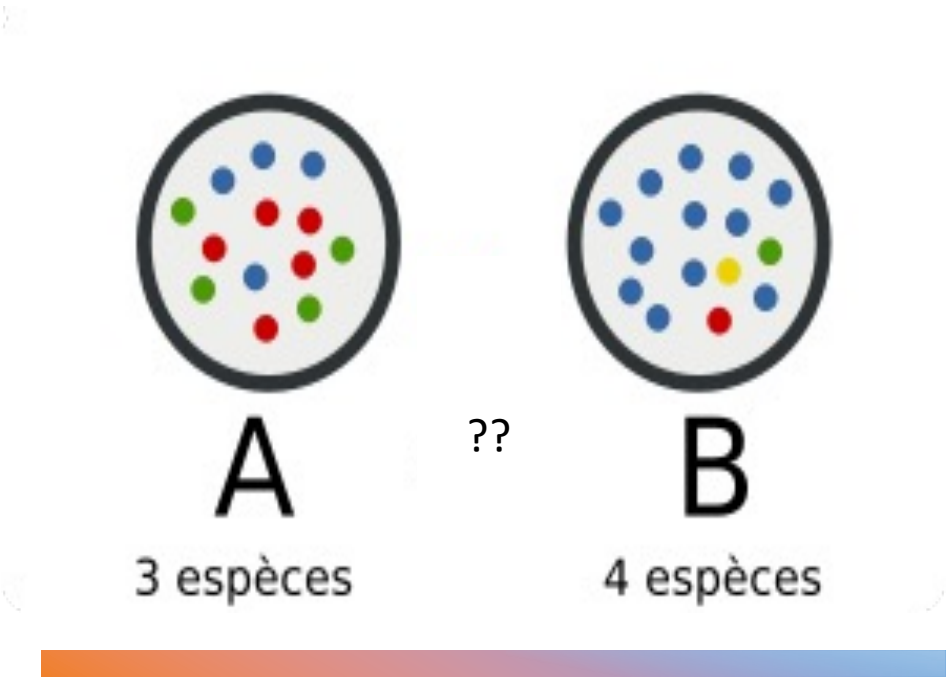


Wide range of species = High richness
Relatively **equal** proportion = high evenness

Lower biodiversity

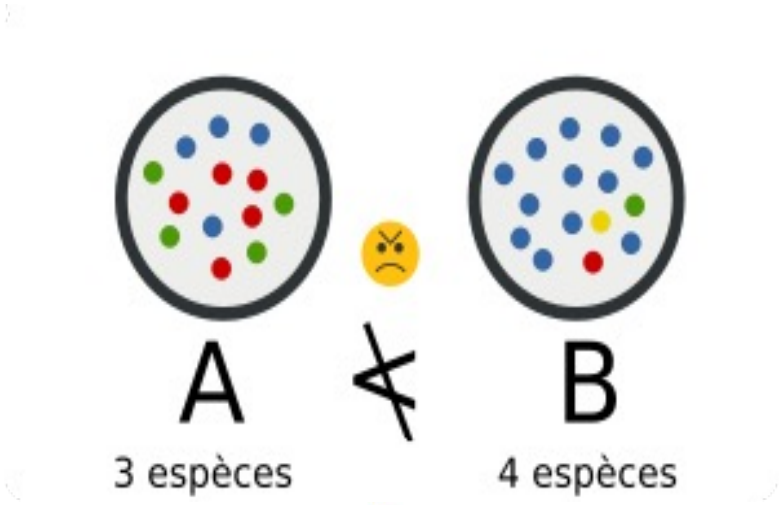
Few species = low richness
Unequal proportion = low evenness

Which one is more diversified ?



Difficult to judge by only looking the numbers !!

Use α-**diversity indices to quantify the biodiversity** of an habitat/sample



Neutral Diversity Indices (=Taxonomic)

Why neutral?

- Species are considered as equidistant (= do not consider species relatedness)

Common indices are Shannon, Simpson and specific richness

- **Subjected to estimation bias because of species not sampling**

- Suppose that the delineation of « species » is clear!

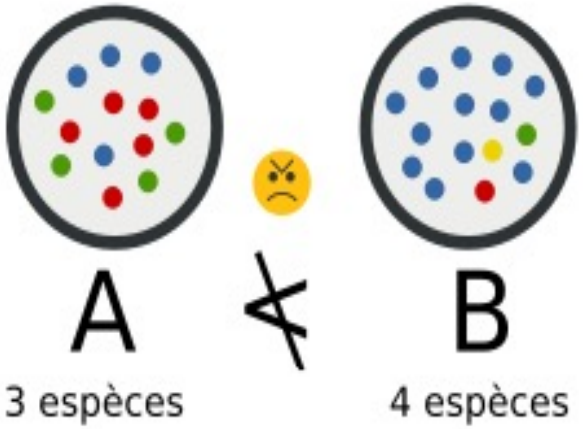
→ It's not!

- Species concept definition (phylogenetic, biological, ecological niche)
- Methodological issues

Shannon-Weaver Index: Combine Richness & Evenness

$$H(X) = H_2(X) = -\sum_{i=1}^{n} P_i \log_2 P_i.$$

Pi : relative abundance of the species i

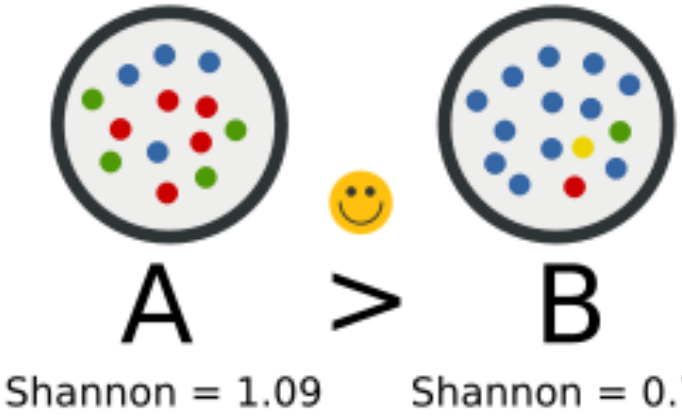


- A: consists of 3 species, of which 4 green, 5 red & 4 blue

The Shannon index will be :

$$-(\frac{4}{13}log(\frac{4}{13}) + \frac{5}{13}log(\frac{5}{13}) + \frac{4}{13}log(\frac{4}{13})) = 1.09$$

Finally, after estimating Shannon for B sample …



- B consist of 4 species, of which 1 green, 1 red, 1 yellow & 11 blue

Shannon value: Influenced by richness

H is generally between 1.5 - 4

H gets higher when :

- There are **more species**
- There is **greater evenness**

There's no upper limit to the index

The maximum : All species have the same number of individuals (perfect evenness)

H max = log(S)

100 species : log(100) = 4.6

1000 species : log(1000) = 6.9

Equitability Pielou Index… with Shannon Scores

Shannon is dependent on species richness!

→ Pielou index is independent of species richness

$$\textit{Pielou index} = \frac{\textit{Shannon Index } (H)}{\log(S)}$$

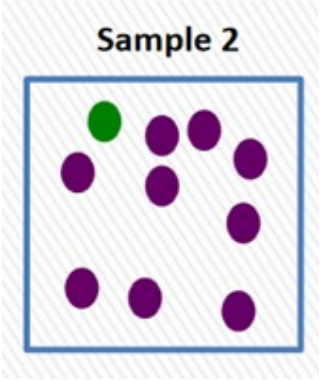
S is Species Richness

Pielou index is a way to measure how the species are evenly distributed in a community

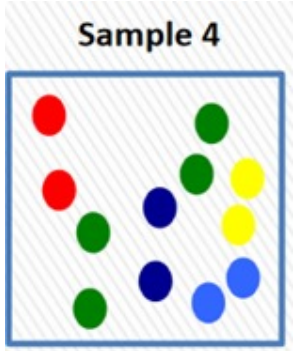
Range 0 to 1 (max equitability, habitat diversity increase)

Simpson's Index of Diversity

Idea : Indicates the taxa dominance and gives the probability of two individuals that belong to the same taxa being randomly chosen



A value of 0.8 …
2 sequences randomly selected
have 80% chance to belong to the same ASV!



Simpson index = D

$$D = \sum_{i=1}^{S} p_i^2$$

P = proportion of the species

Gini-Simpson = 1 - D

$$E = 1 - \sum_{i=1}^{S} p_i^2$$

- Influenced by highly abundant Taxa
- Greater weight on evenness
- Range 0 to 1 (high)

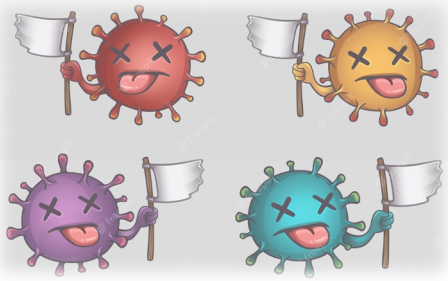
The closer the value to 1, the more diverse the habitat is

0 — **No diversity**

1 — **Infinite Diversity**

Low sp. Diversity :

- **Few successful species** in the habitat
- **Env stressful**, few niches, few organisms well adapted to env
- Any **change** in env may have **serious effect on ecosystem**



High sp. Diversity :

- **Great number** of successful species, more **stable ecosystem**
- Env is less likely to be hostile
- Complex food
- Env change is **less likely to dammage** the whole ecosystem

Diversity Estimators

- Chao1 & ACE are non-parametric estimators of taxa richness

- Sampling at infinity

- Good sampling gives you a total number of ASV/OTU observed not far from the Chao1 / ACE value (predicted for the sampled environment)

Chao1= S_{obs} + Adjustment (linked to the rare)

Chao1 adjustment

$$\frac{F_1(F_1-1)}{2(F_2-1)'}$$

Singletons = F1

Doubletons = F2

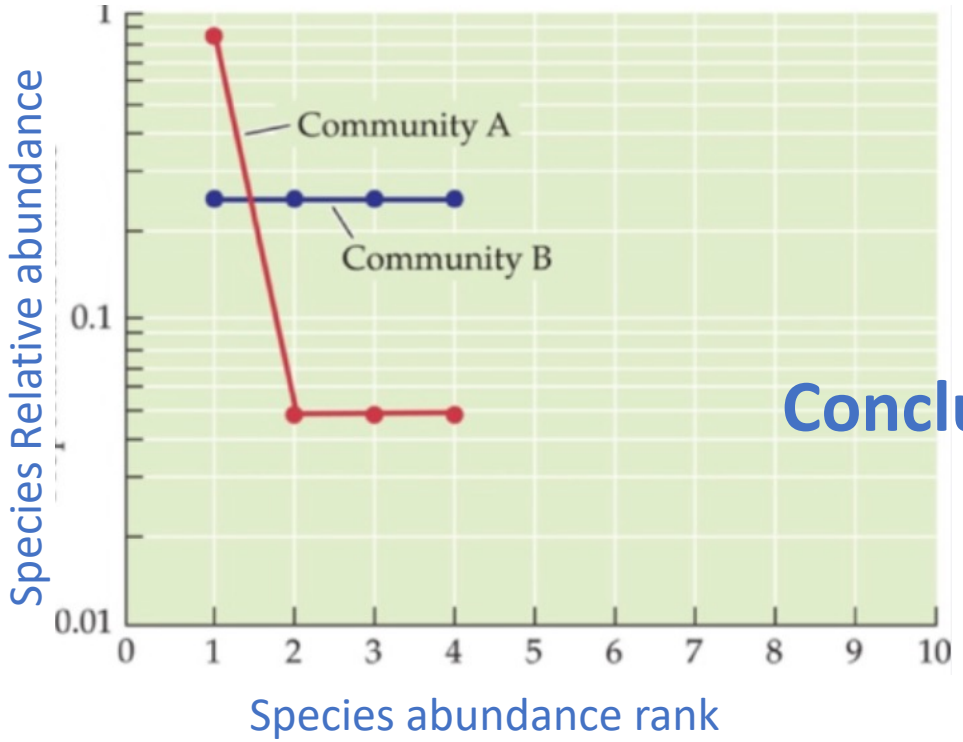
Idea : Rare taxa bring most information about the number of missing taxa

→ Issue : DADA2 remove singletons !!!

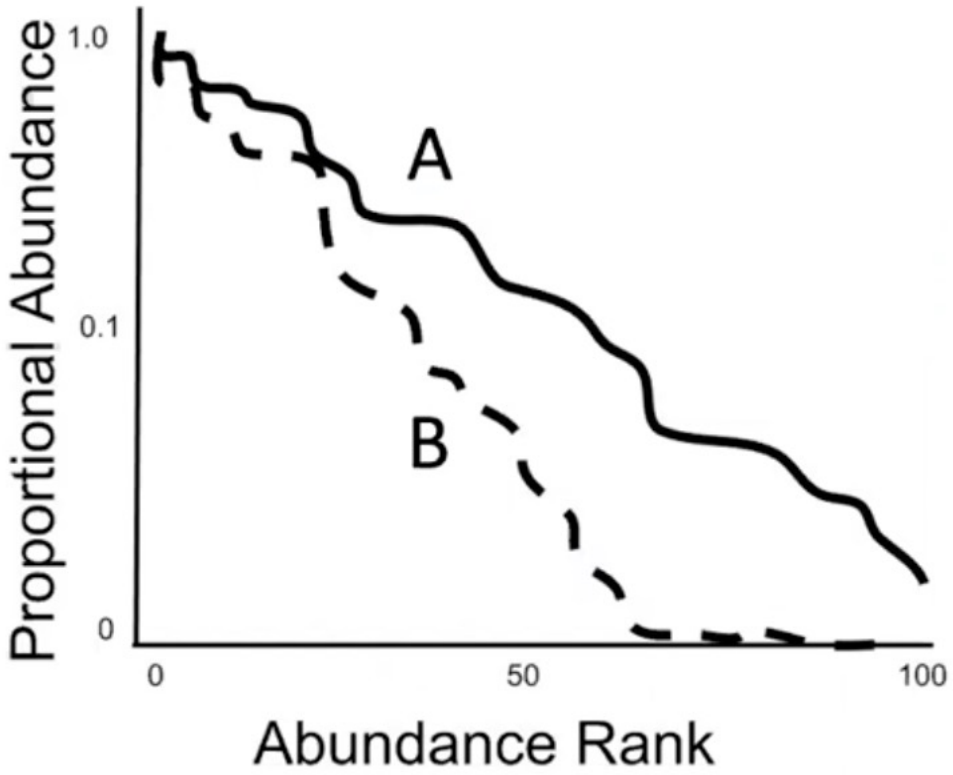
Rank Abundance Curves

Plot relative abundance of species against their rank in abundance

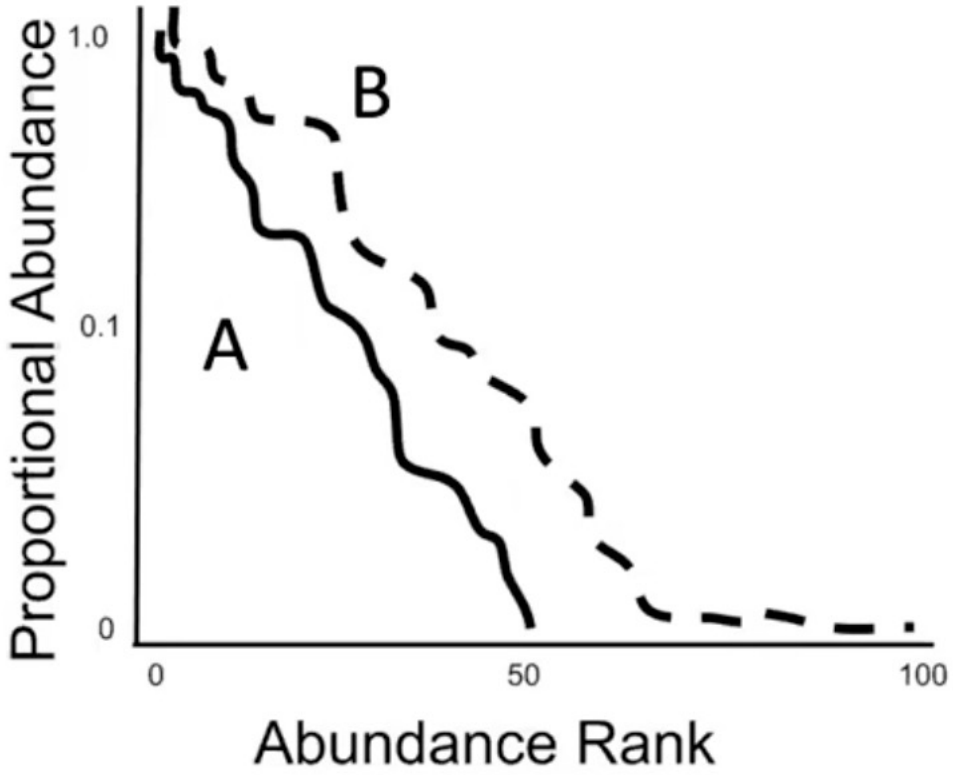
Dominance and diversity… graphically



Conclusions?



- **Which one has the greatest evenness?**

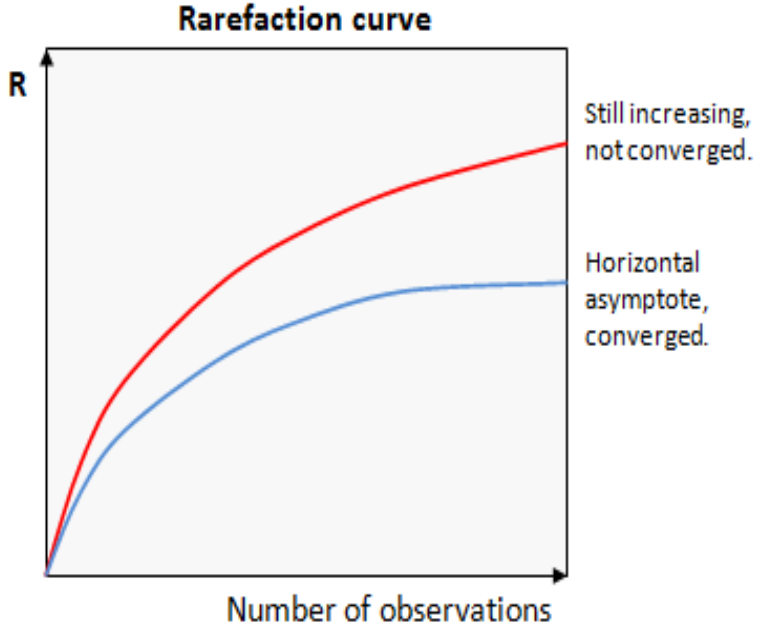


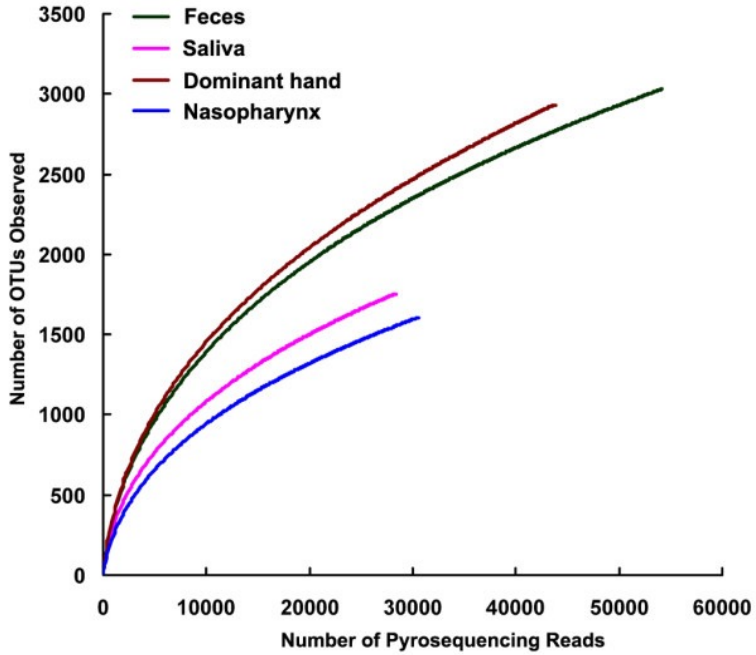
Which one has :

- **Greatest evenness?**
- **Greatest richness?**

Rarefaction Curves

« Is the sequencing effort performed (sequencing depth) for a sample (s) sufficient for the number of species observed ? »





→Reach the asymptote ???

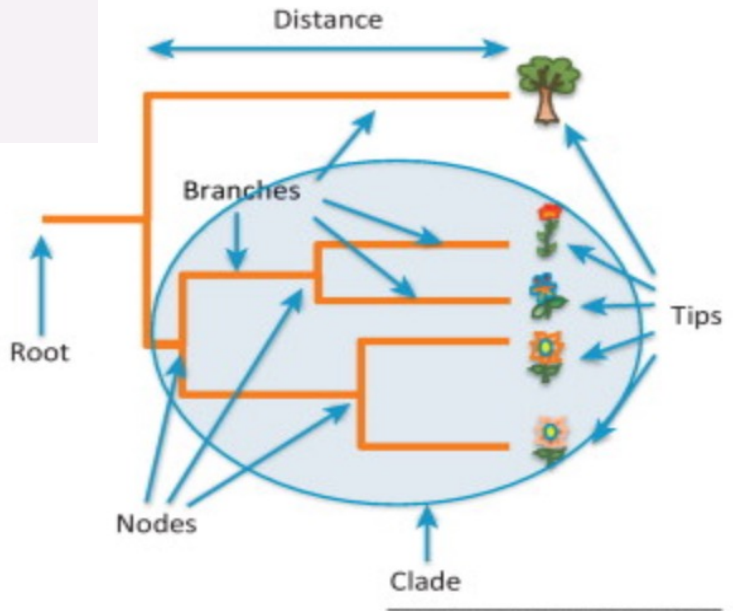
Asymptote means that sequencing more (depth), will not increase your number of OTU/ASVs observed

Phylogenetic Indices

Phylogenetic Diversity (PD) measure of the evolutionary history within a set of species :

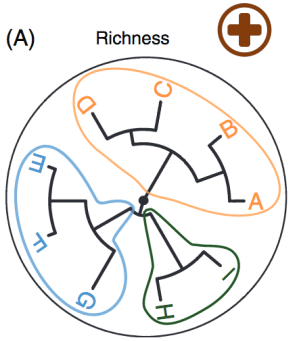
- Relatedness, speciation, events …

→ describes a fundamental aspect of biodiversity



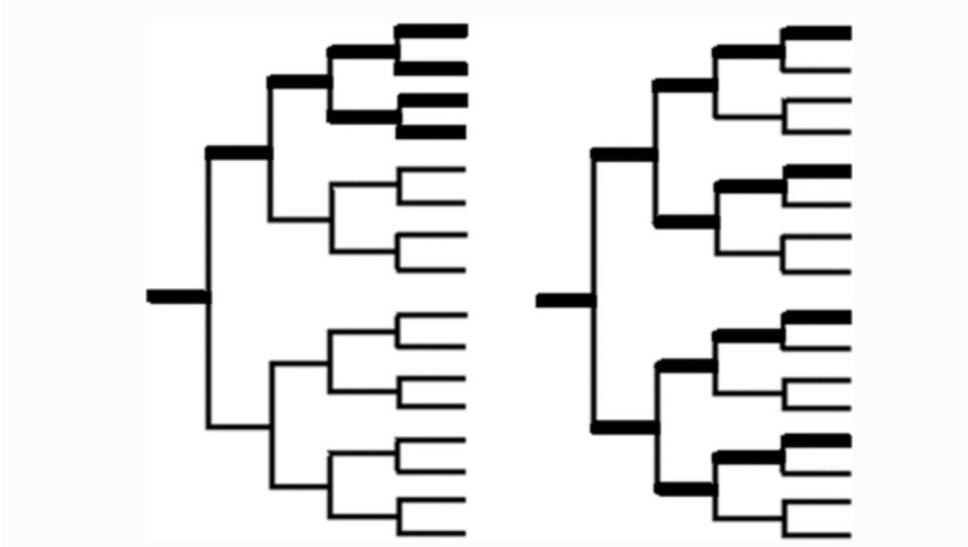
A better predictor of ecosystem function than species richness & evenness

Richness = How much ?



- **PD** $_{faith}$ (Faith's Phylogenetic Diversity)

→ As the minimal total length of all the branches required to span a given set of taxa to the root (Related to SR)

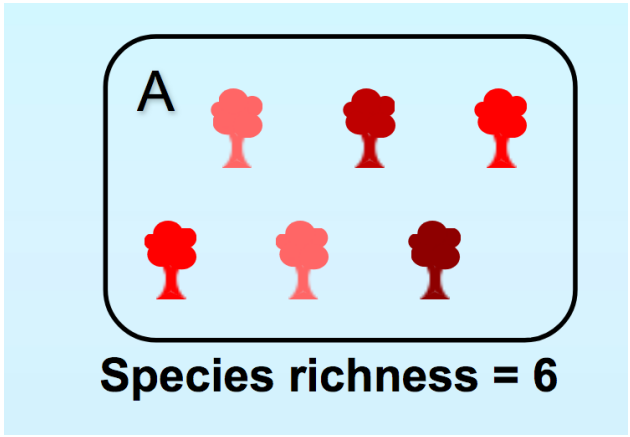


Low PD

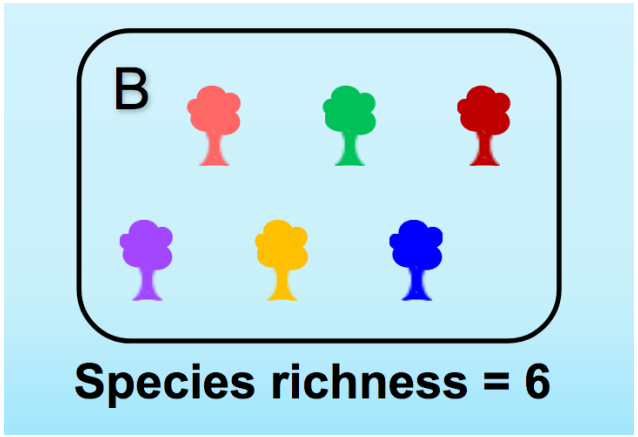
Assemblage of taxa with more recent evolutionary history

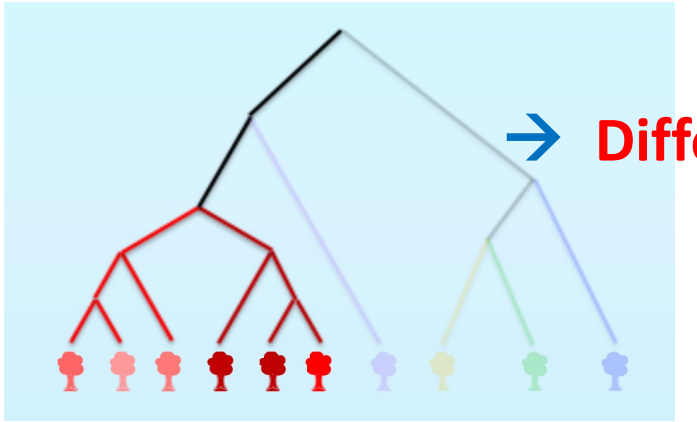
High PD

Communities with more evolutionary divergent taxa and older history

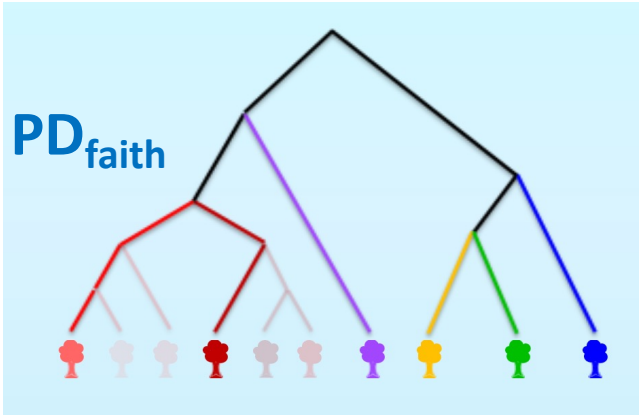


Same SR





→ Different PD_{faith}



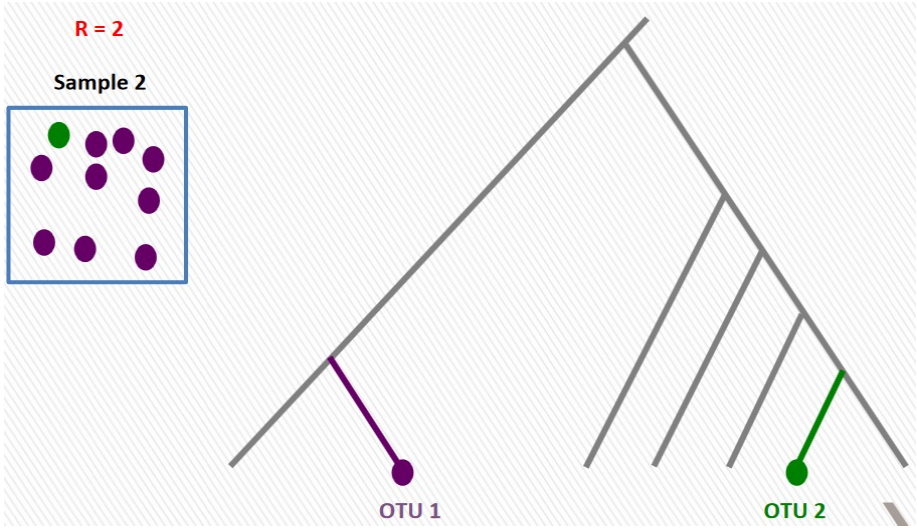
EnvA : Sum up branch lengths

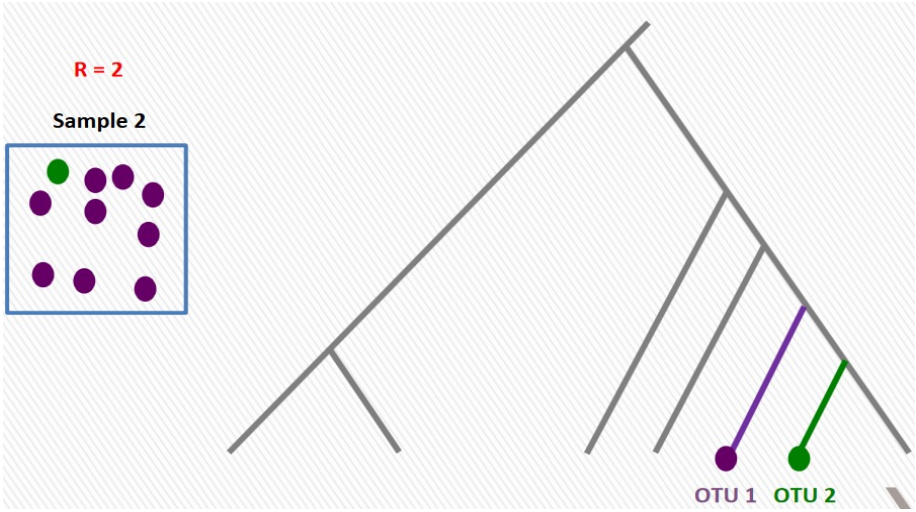
EnvB : Sum up branch lengths

If not using phylogenetic indices : lost of this information !

Higher PD

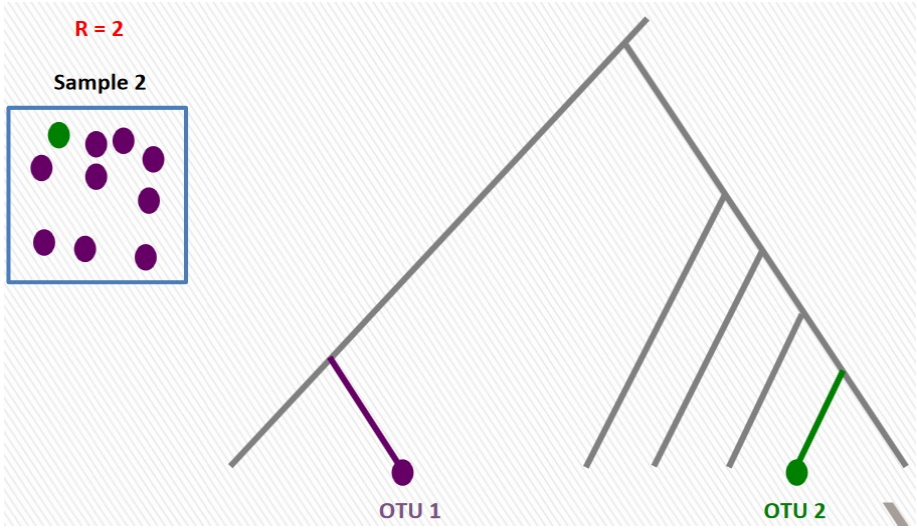
Same SR & same evenness = Same Shannon/simpson

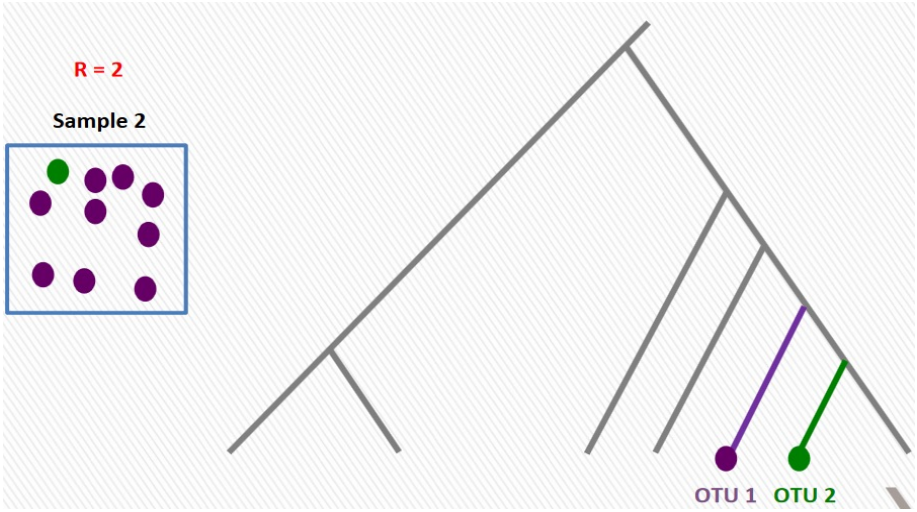


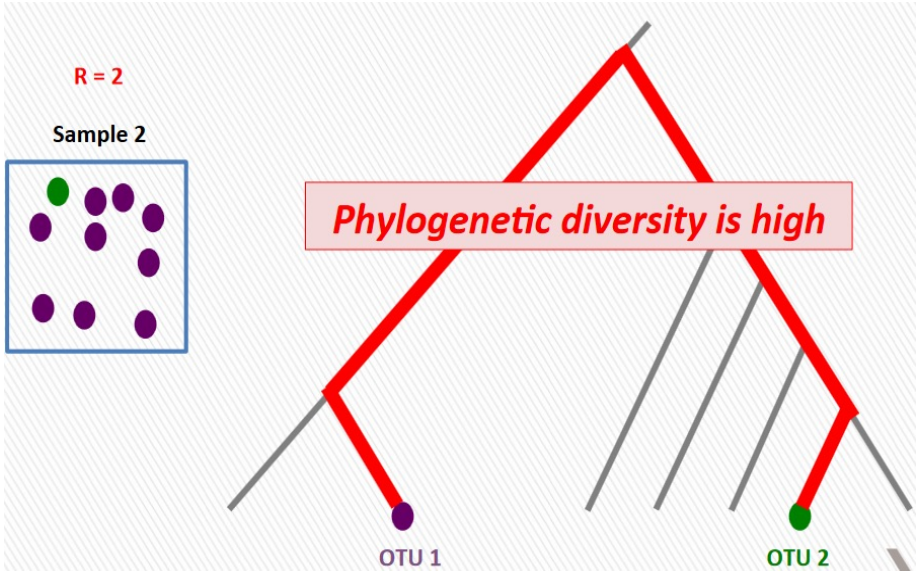


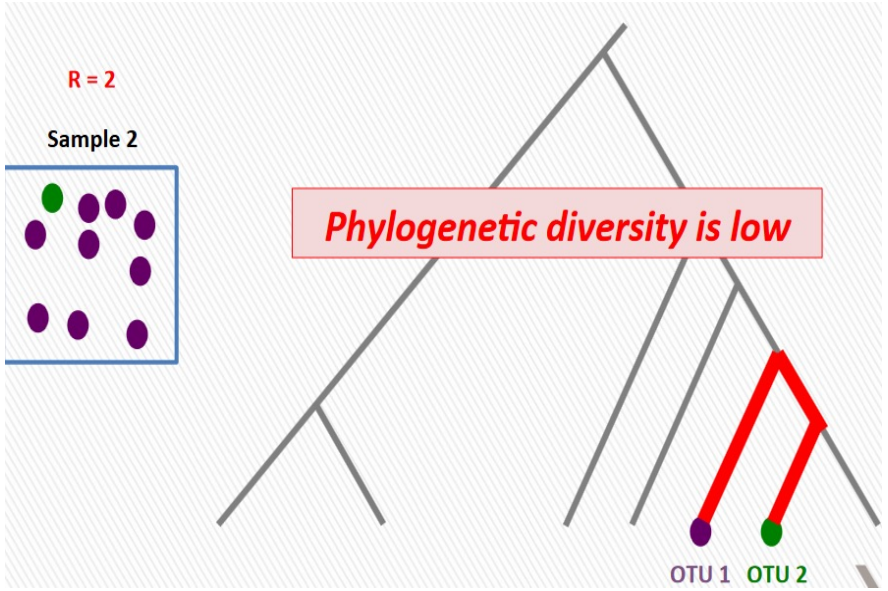
Same phylogenetic diversity ???

Same SR & same evenness = Same Shannon/simpson



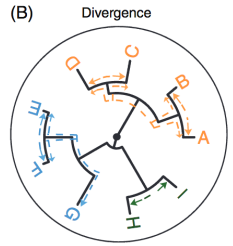






BUT different PD_{faith}!!!!

Divergence : Quantify the phylogenetic difference…



Why PD is a proxy of functional diversity, niche/community dissimilarity :

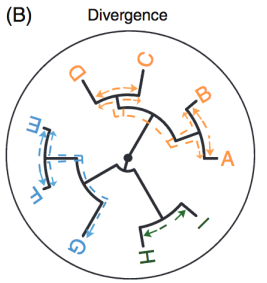
- **Closely related species tend to have similar functions/traits** (similar habitat requirement)
- **Distant related species tend to have greater complementary functions**

Meaning for the community assemblage :

Dominance of **closely related** species → **Clustering pattern**
Dominance of **distant related** species → **Overdispersion pattern**

Given species richness,
does the **phylogenetic diversity** in **AN** assemblage
is greater or less than that expected?

Divergence : How different?



Two commonly used metrics were used to quantify:
the Net Relatedness Index (NRI) and Nearest Taxon Index (NTI)

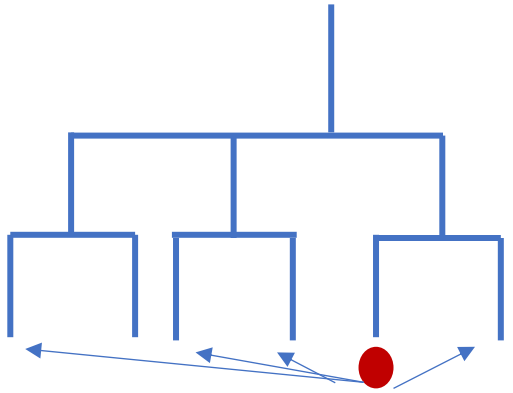
Highlight phylogenetic structure of assemblages **at different evolutionary depths**

NRI : Net Relatedness Index

- Based on the **Mean Phylogenetic Distances (MPD)** in each community.

→ **Average** phylogenetic **distance** of species (**to every other species)**

→ « Basal measure » : Clade representation. **Strongly influenced by the 'basal' structure** of the phylogenetic tree



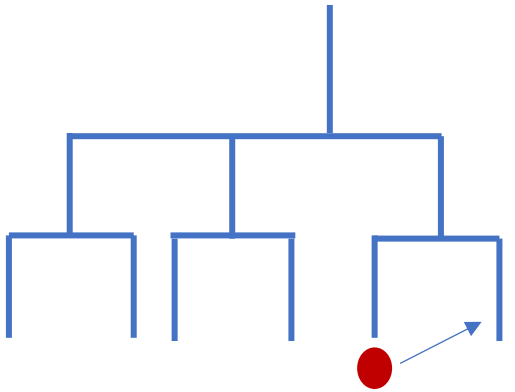
« Global »

NTI : Nearest Taxonomy Index

- Based on **MNTD** the Mean Nearest phylogenetic Neighbor Distance

→ **Average** phylogenetic **distance** to the **nearest neigbour**

→ **Reflect Phylogenetic structure of the tree tips**



« Local »

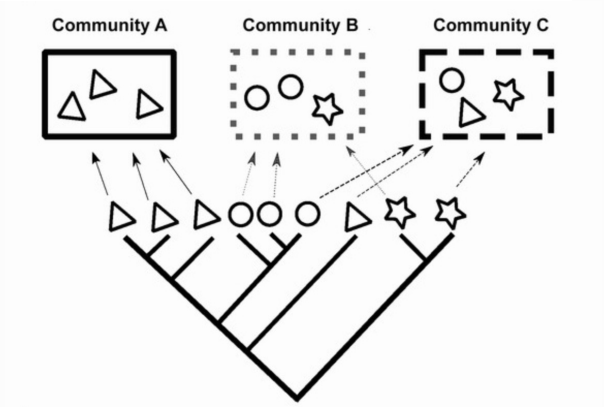
So get NRI /NTI values .. And so what ????

The « Null model » : Phylogeny randomization…

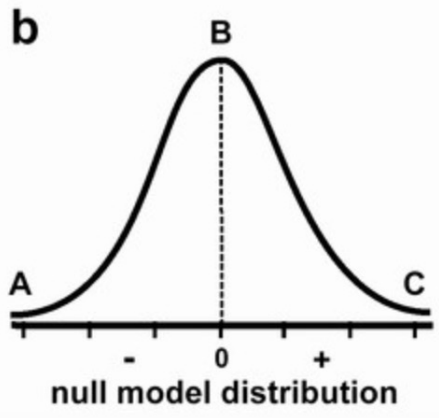
Need a **reference** for comparison → Absence/overdispersion/Clustering!

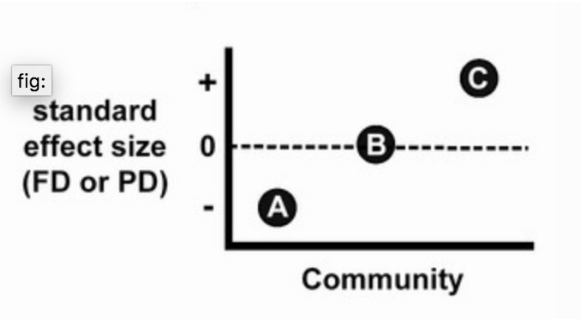
→ A distribution « Null Model » based on random taxa positions within tree

→Is the measure for a specific community is more or less expected by chance ?



Randomization





SES for your values

NTI and NRI are Z scores!!

Interpretation

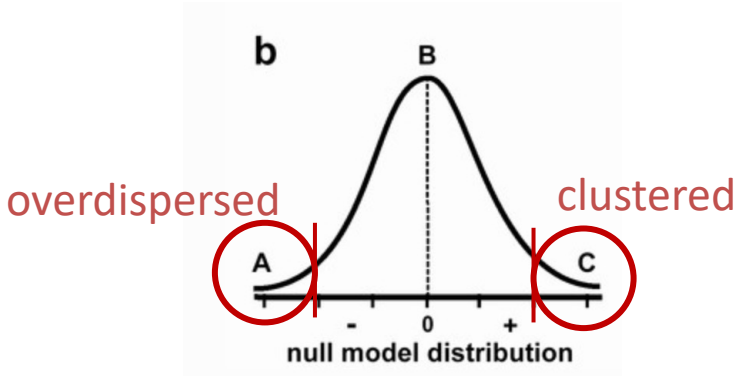
- A negative NRI/NTI value indicates an overdispersed phylogeny where taxa are less related to each other than expected by chance

→ Significance < -1.96

- Positive NRI/NTI values indicate a clustered phylogeny

where taxa are more related to each other than expected by chance

→ Significance > 1.96



Community Assembly : Spatial and Temporal processes

- Niche-driven = Deterministic
- **Selection**

- Biotic interactions (taxa interactions)

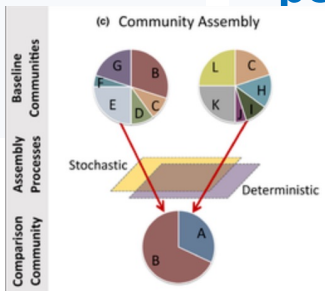
- Environmental filtering
(=Abiotic conditions, physico-chemical)

- Neutral = Stochastic process
- **Unpredictable**

- Random profileration, dispersal

- Random birth-death events

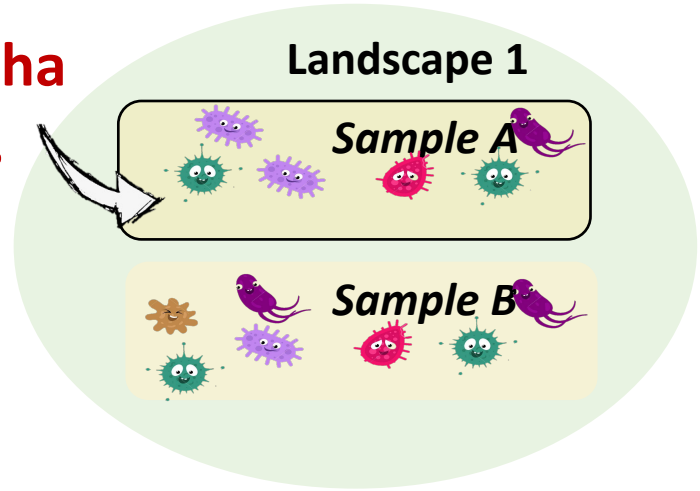
Ecological drift (loss diversity, small pop)



→ See β NTI/NRI

Sum up …

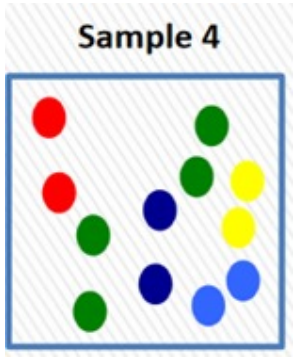
Alpha div.



Diversity within an habitat of fixed size

Use α-diversity indices to quantify the biodiversity of an habitat/sample

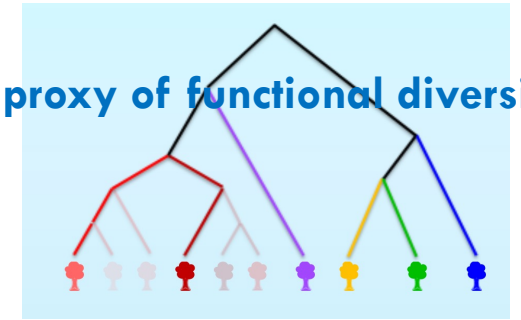
Neutral diversity indices (taxonomic)



Richness & equitability

Shannon, Simpson , pielou …

Phylogenetic indices

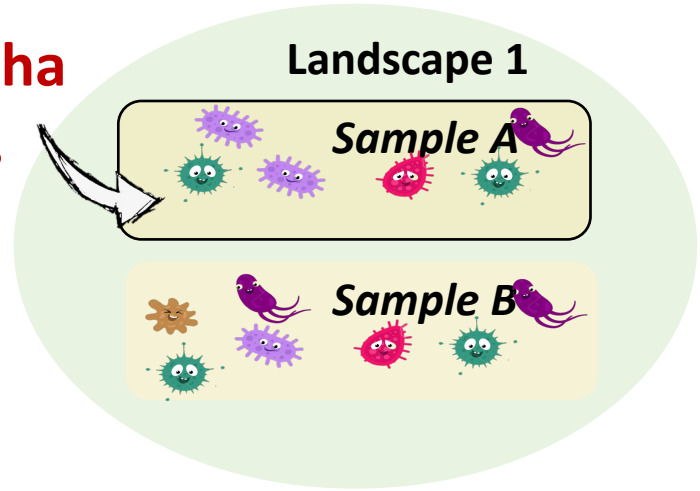


NTI/NRI

Sum up …

Be aware of the biases/limitations associated with the indices you use when interpreting your results !!

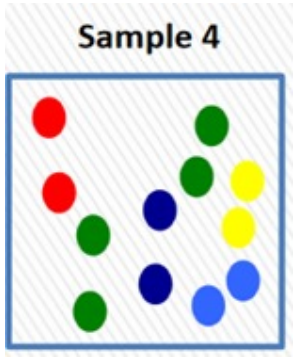
Alpha div.



Diversity within an habitat of fixed size

Use α-diversity indices to quantify the biodiversity of an habitat/sample

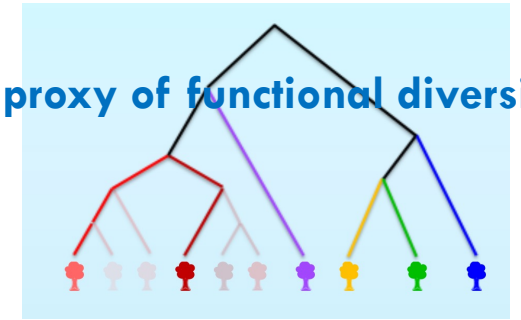
Neutral diversity indices (taxonomic)



Richness & equitability

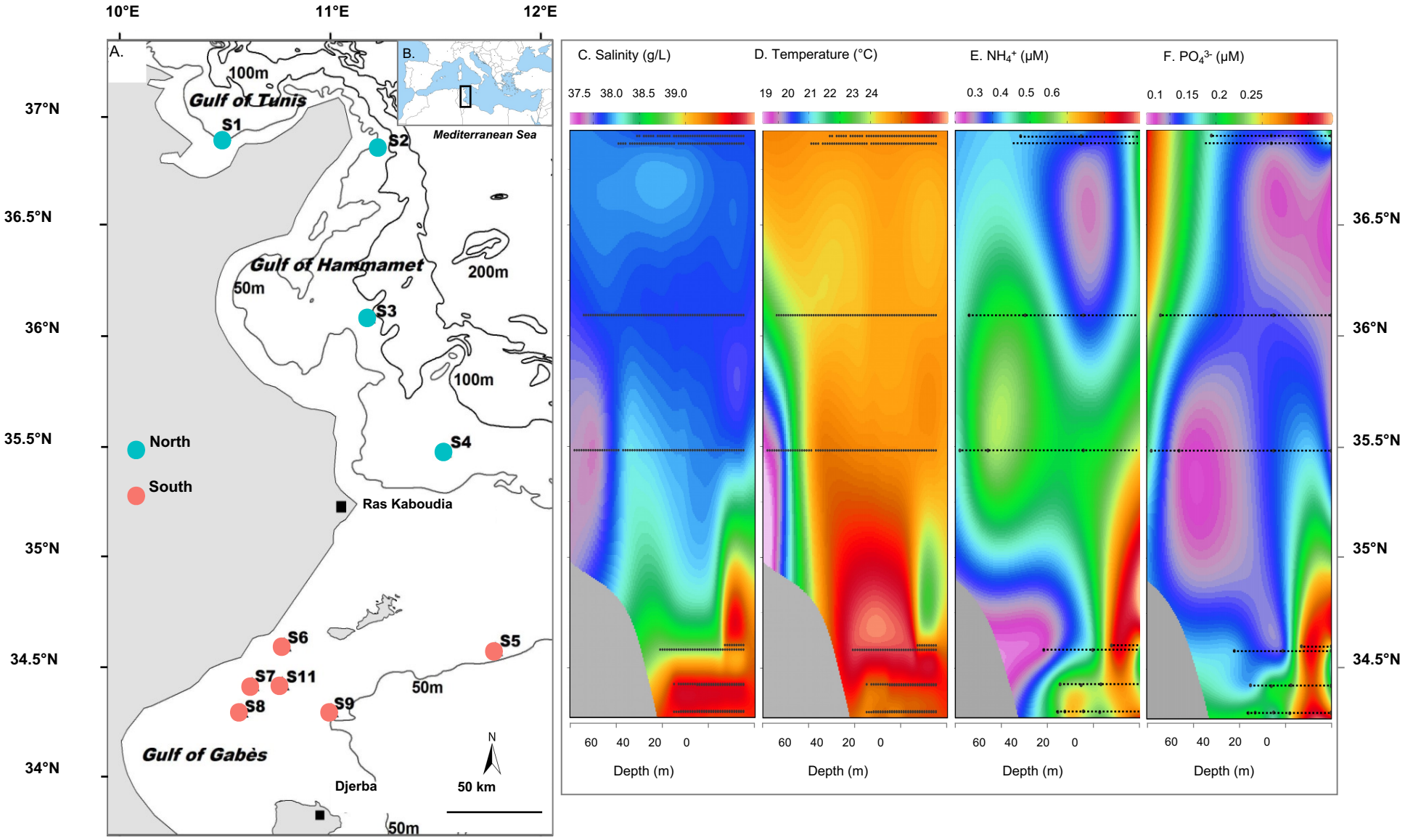
Shannon, Simpson , pielou …

Phylogenetic indices



NTI/NRI

MetaData


10°E
11°E
12°E
37°N
36.5°N
36°N
35.5°N
35°N
34.5°N
34°N
A.
B.
100m
Gulf of Tunis
S1
S2
Mediterranean Sea
Gulf of Hammamet
50m
200m
S3
100m
North
South
S4
Ras Kaboudia
S6
S5
S7
S11
50m
S8
S9
Gulf of Gabès
Djerba
N
50 km
50m
C. Salinity (g/L)
37.5 38.0 38.5 39.0
D. Temperature (°C)
19 20 21 22 23 24
E. NH4+ (µM)
0.3 0.4 0.5 0.6
F. PO43- (µM)
0.1 0.15 0.2 0.25
60 40 20 0
Depth (m)