

# Community analysis II

Sarah Faulwetter, Pieter Provoost, Sander Wijnhoven

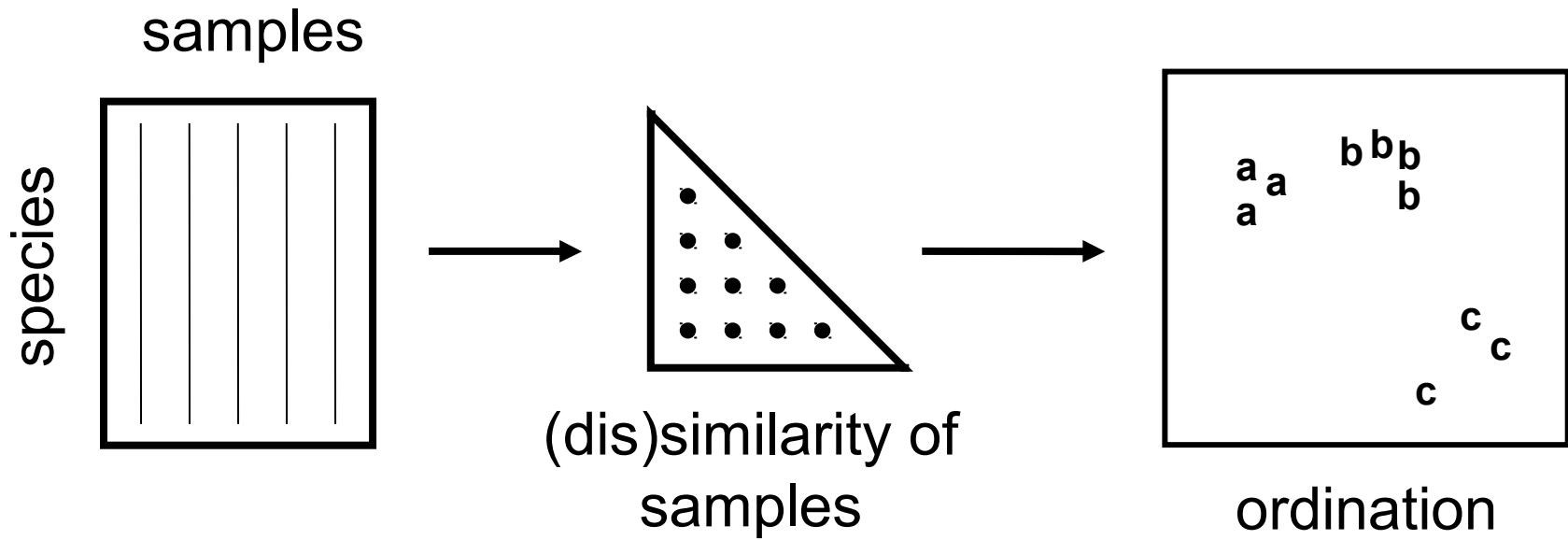
A Practical introduction to acquisition, validation,  
quality control and access to (biodiversity) data

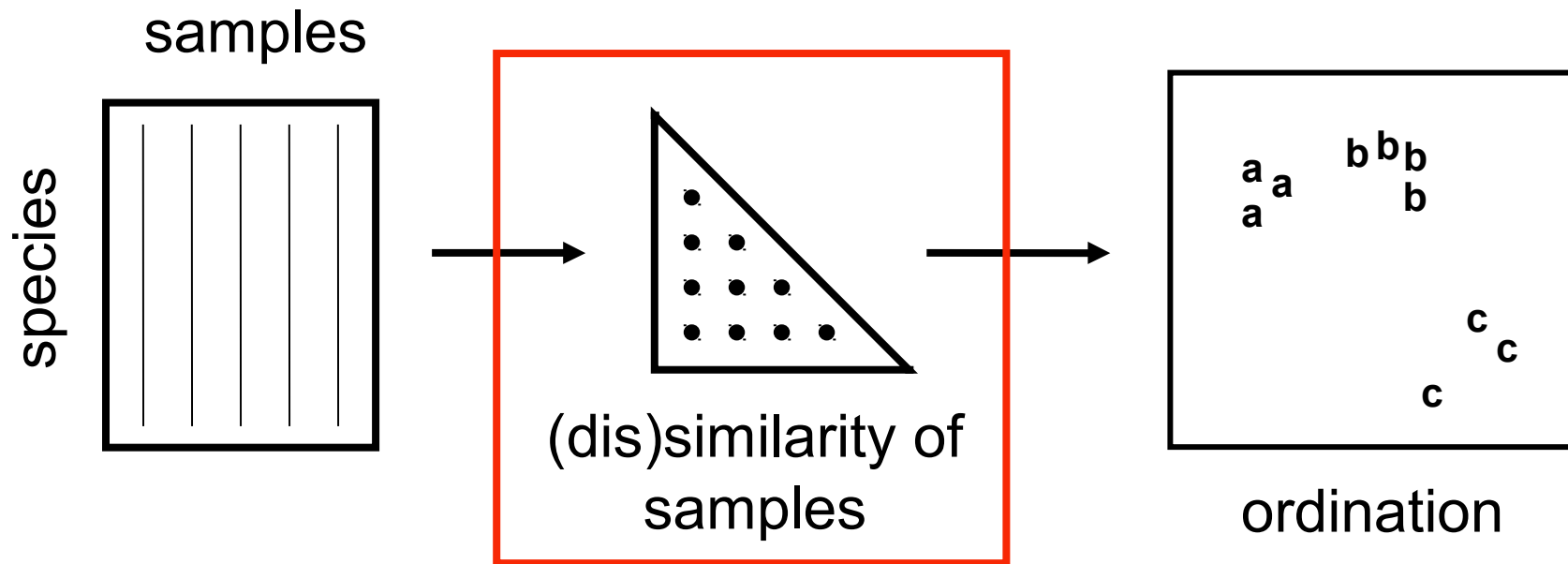
**Vlissingen, Netherlands, 15.06.2015**



# Ordination

Non-metric multidimensional  
scaling, nMDS

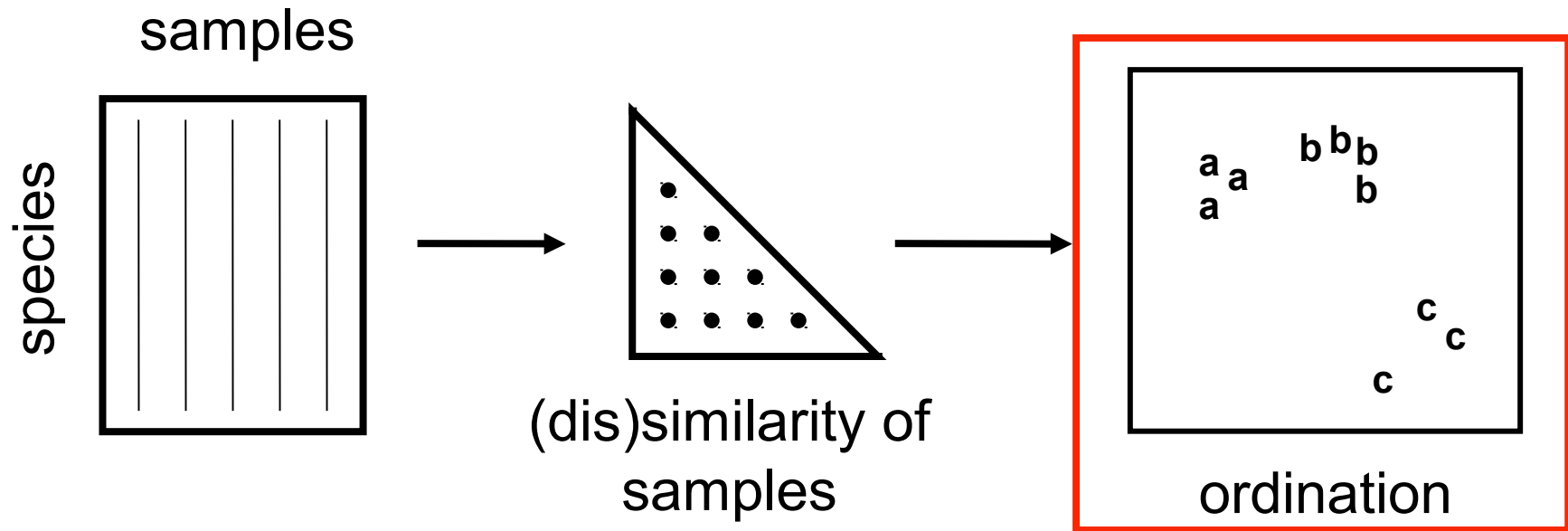




Bray-Curtis coefficient

# Why Bray-Curtis?

- most commonly used dissimilarity measure for biocommunities
- used for data from non-normal distributions
- values between 0 and 100%
  - 0 = samples differ completely (no species in common)
  - 100 = samples are identical
- does not take joint absences of species into account



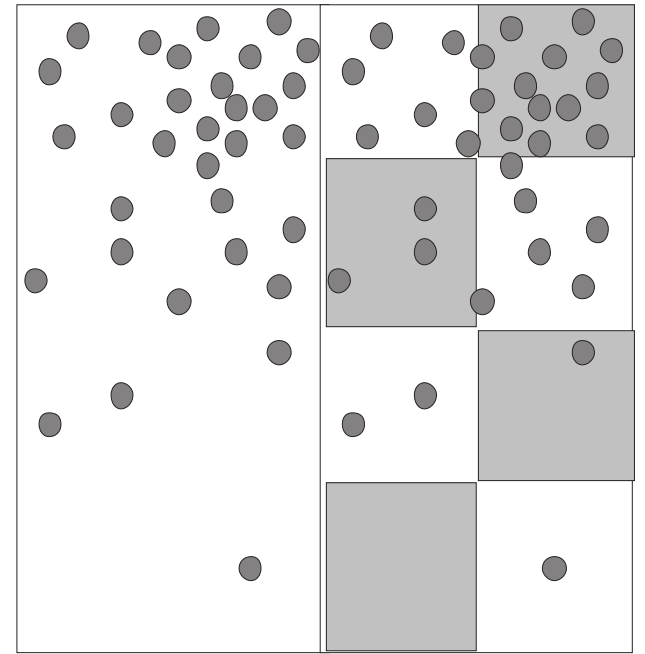
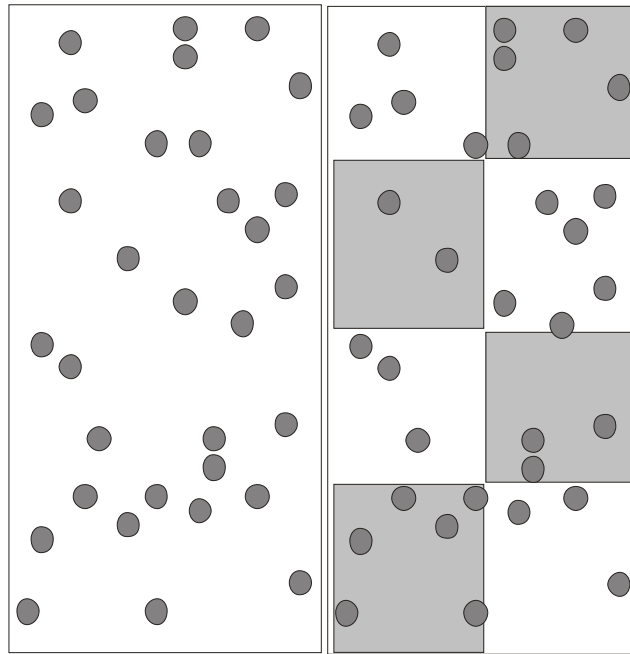
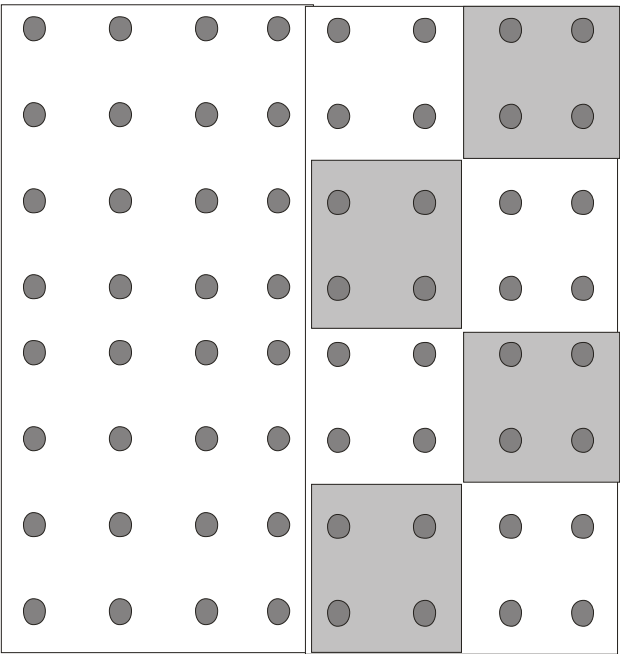
Creation of an MDS plot

# Distribution & Sampling

**uniform/ homogeneous**

**random**

**clumped / contagious**



# Transformations

**no transformation:** Similarity takes only most abundant species into account

**square root:** takes also species with moderate abundances into account

**fourth root:** takes also rare species into account

**log(1+y):** for large abundance values: even stronger than 4<sup>th</sup> root, for low abundances: less strong than 4<sup>th</sup> root

**presence-absence:** only for specific analyses, e.g. samples from very different sampling campaigns, large scale analyses etc.

→ Depends on question / hypothesis!



# non-metric MDS

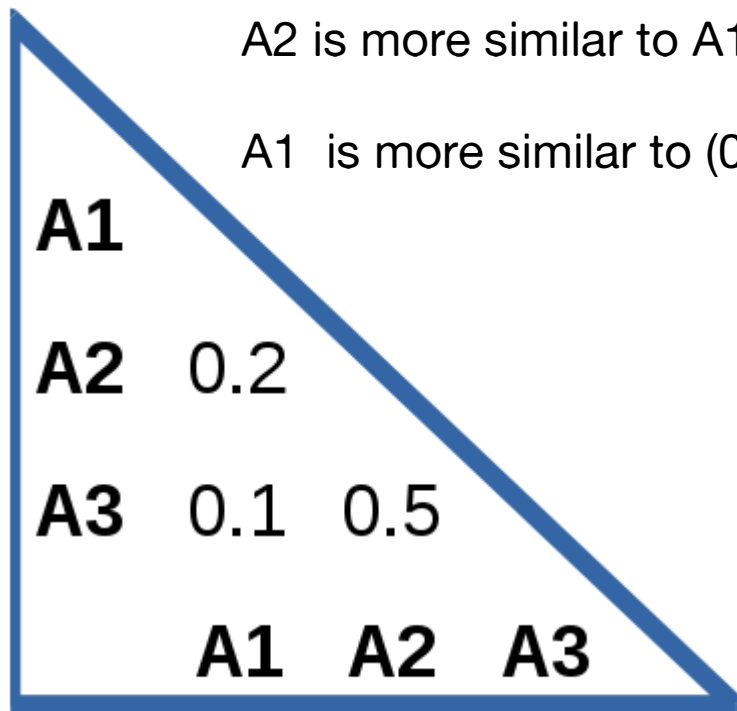
**Nonmetric:** no axes

**Multidimensional:** depicts relationships between multiple variables in 2 or 3 dimensions

**Scaling:** ratio between reality and its visualisation

# How does the nMDS work?

Triangular matrix = (dis)similarities

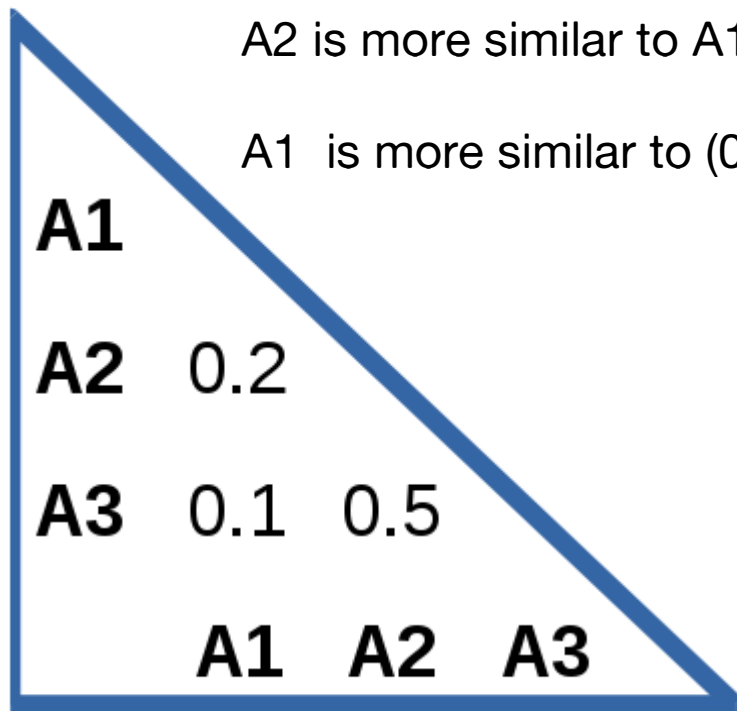


A2 is more similar to A1 (0.2) than to A3 (0.5)

A1 is more similar to (0.1) than to A2 (0.2)

# How does the nMDS work?

Triangular matrix = (dis)similarities

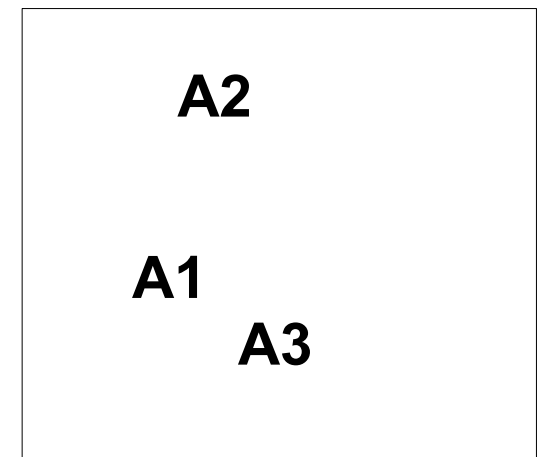


A2 is more similar to A1 (0.2) than to A3 (0.5)

A1 is more similar to (0.1) than to A2 (0.2)

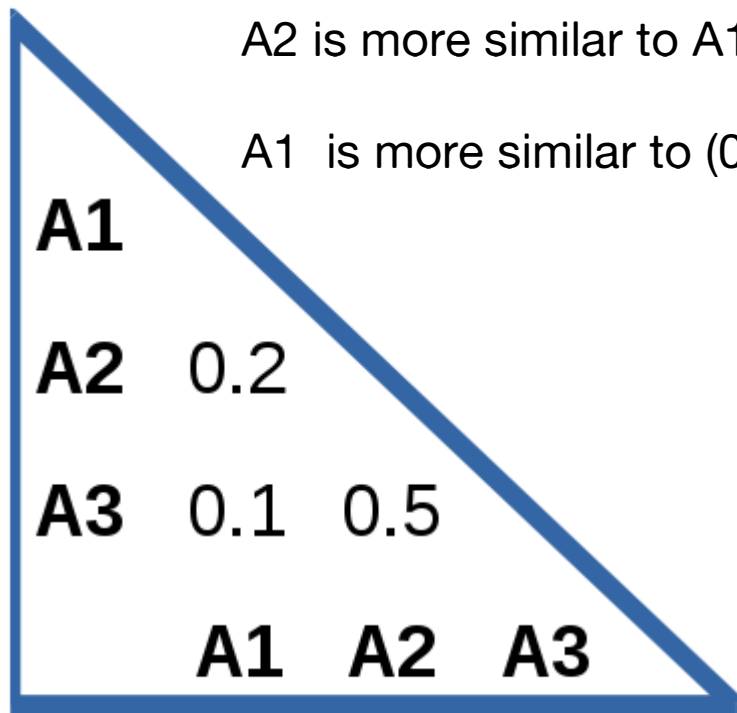
A2 is closer to A1 than to A3

A1 is closer to A3 than to A2



# How does the nMDS work?

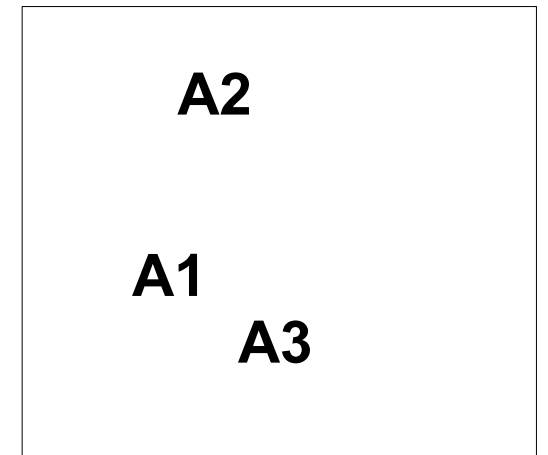
Triangular matrix = (dis)similarities



A2 is more similar to A1 (0.2) than to A3 (0.5)

A1 is more similar to (0.1) than to A2 (0.2)

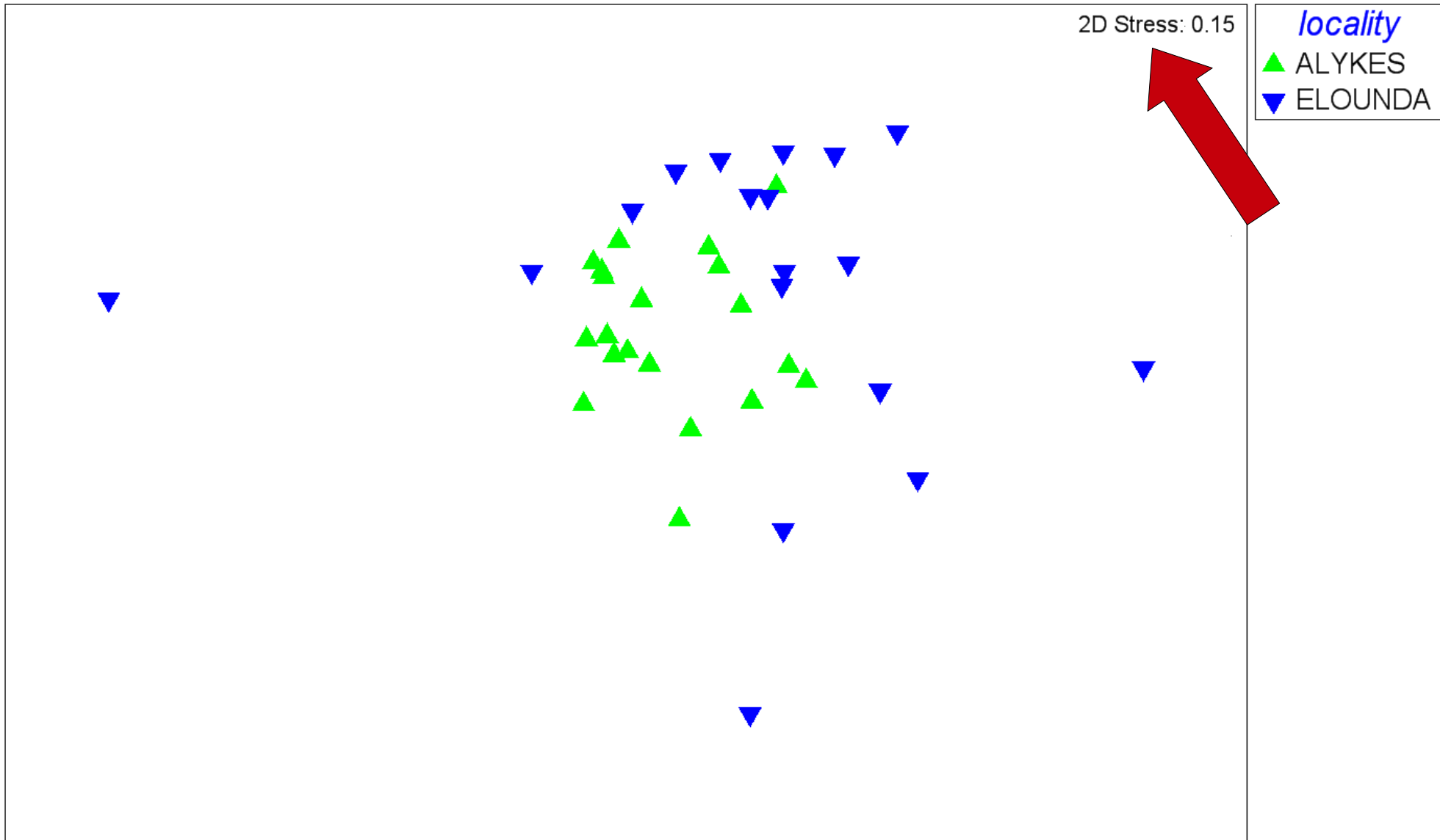
Hierarchic order, not absolute  
Relative relationships  
No axes (non-metric)



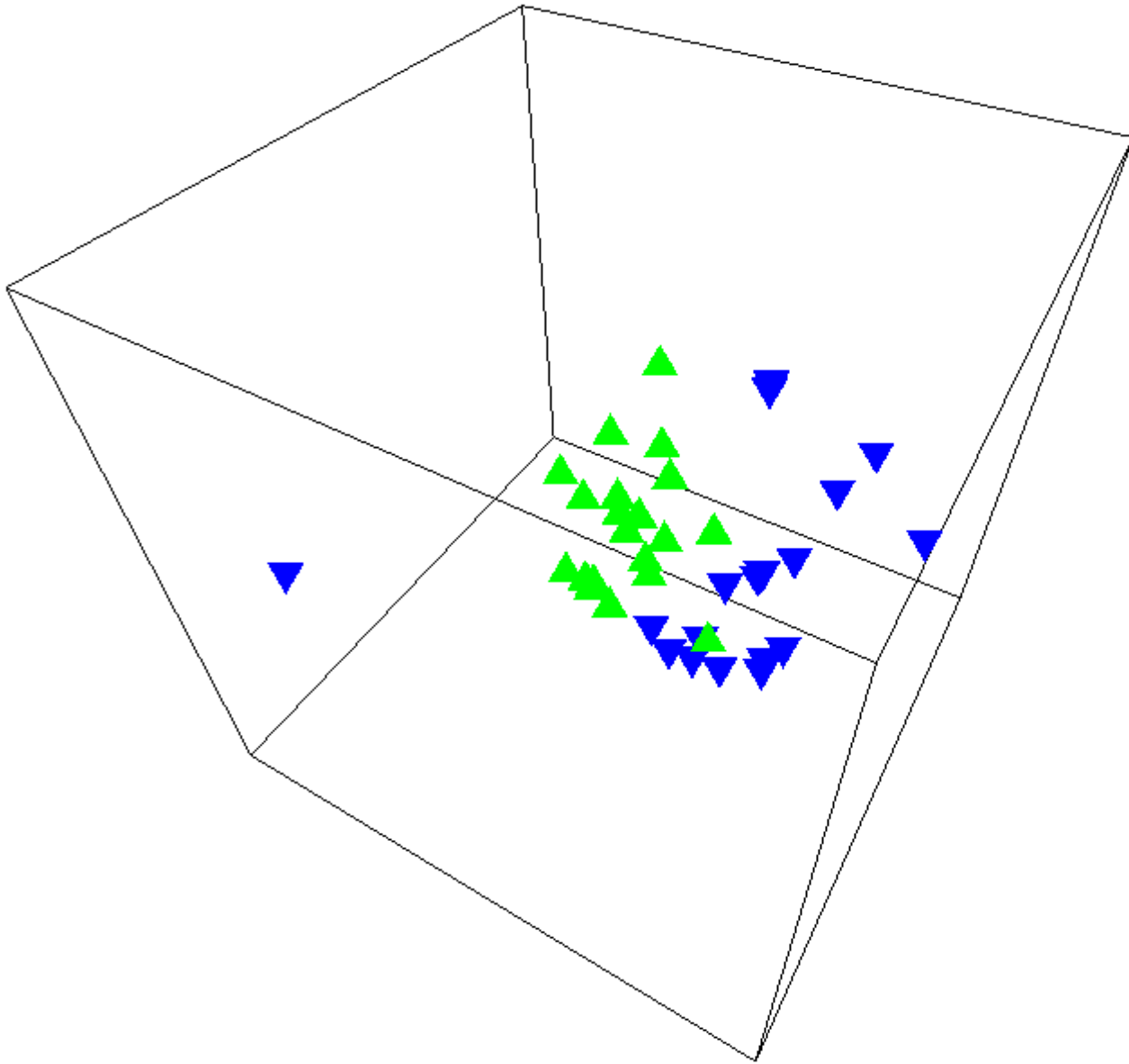
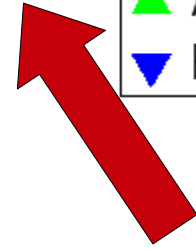
A2 is closer to A1 than to A3

A1 is closer to A3 than to A2

# Stress



3D Stress: 0.09

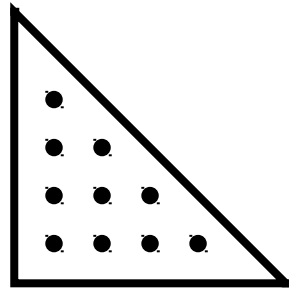
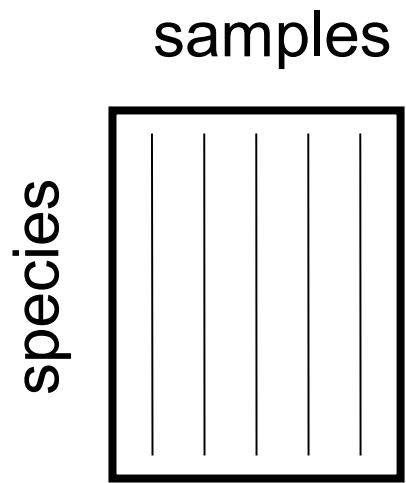


# Stress Value

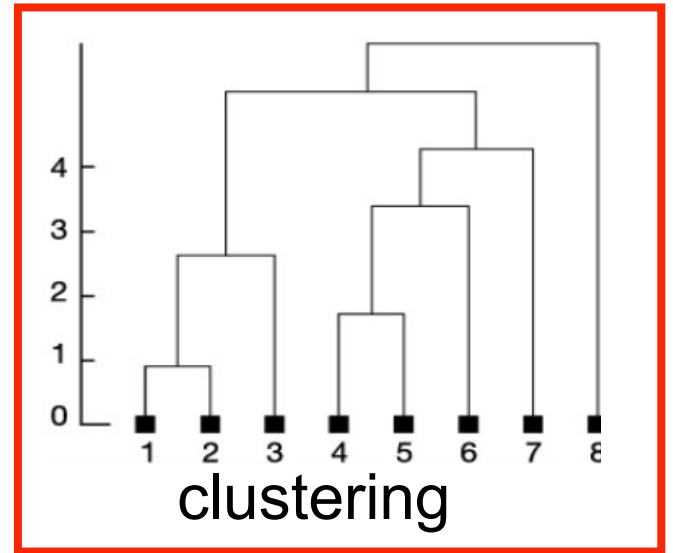
- 0.0 : perfect representation
- > 0 – 0.1 : good representation
- > 0.1 – 0.2 : acceptable representation
- > 0.2 : non-acceptable representation, information is lost in the reduced dimensions of the plot

# Hierarchical Clustering



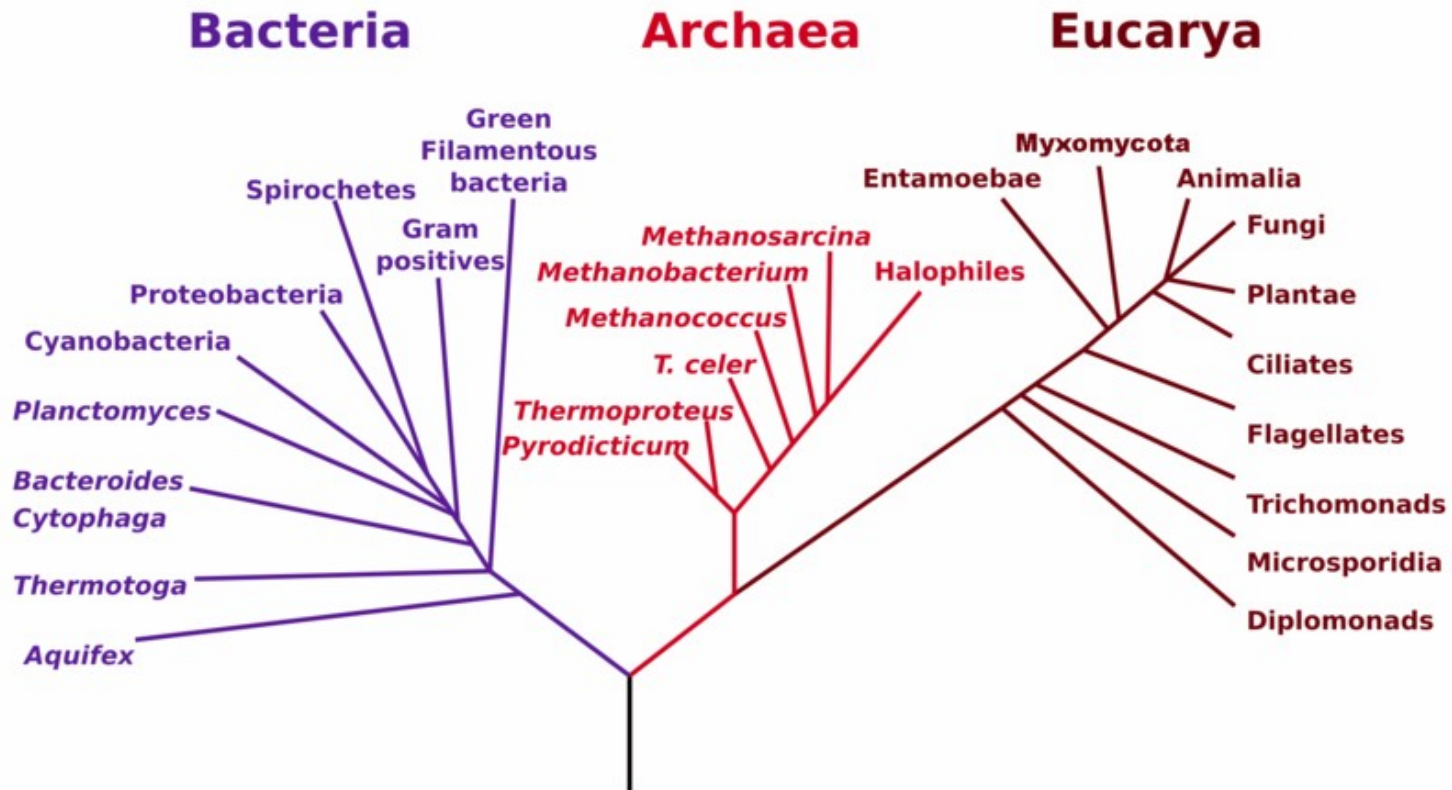


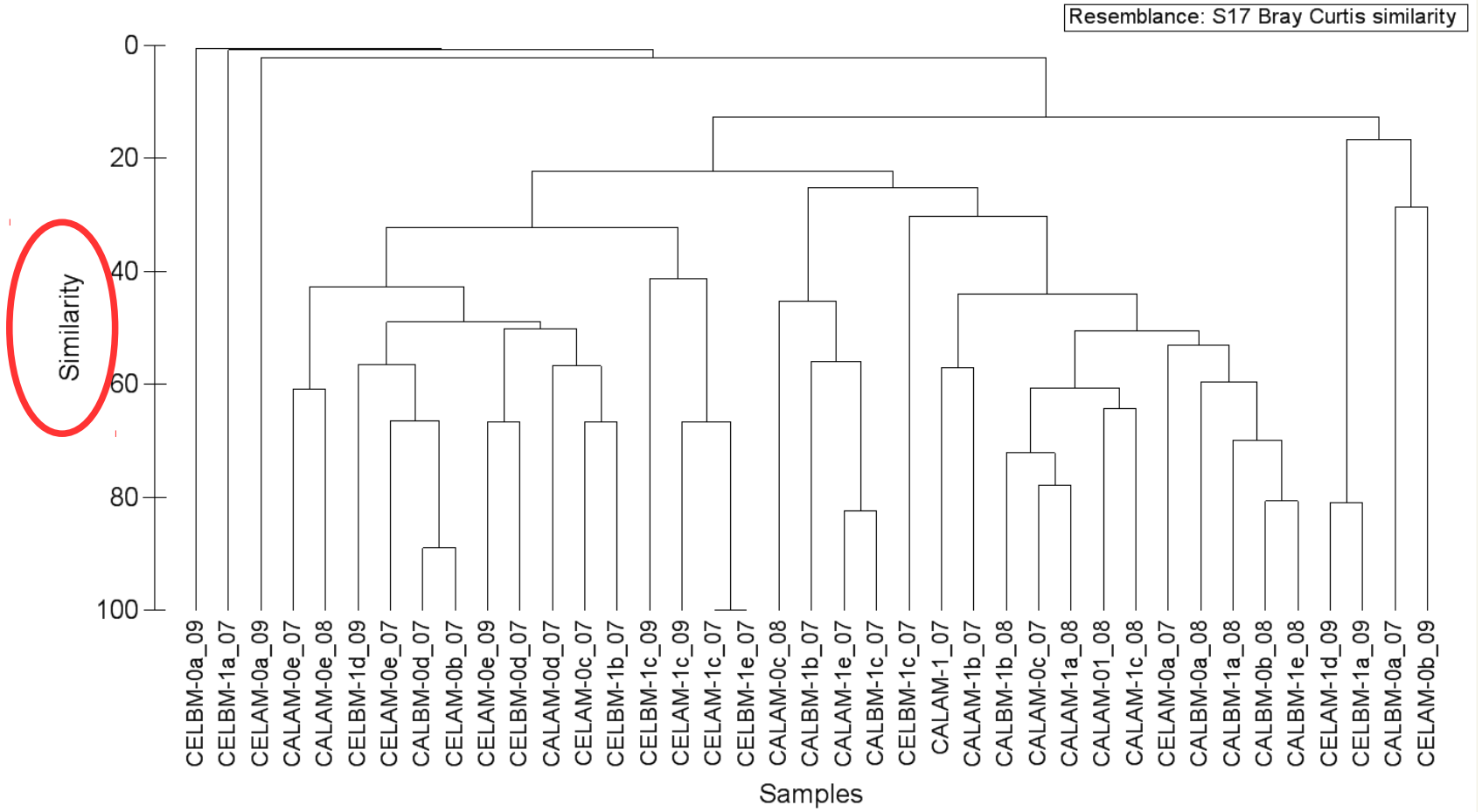
(dis)similarity of samples



Finds groupings so that samples in one group are more similar to each other than to other samples outside the group

# Phylogenetic Tree of Life





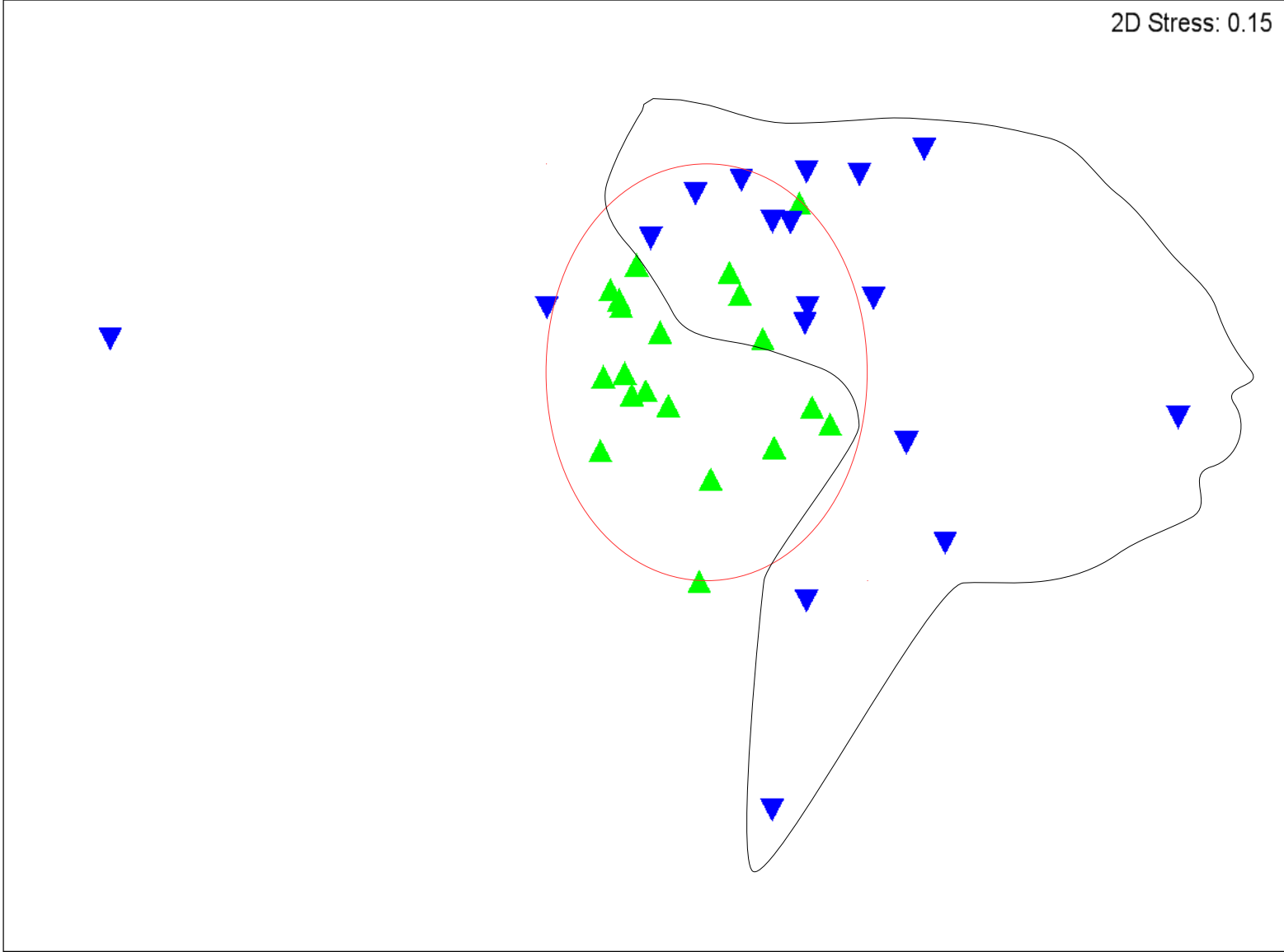
# Linkage options

- Single linkage
- Complete linkage
- Group average → often the most balanced

# Combination Cluster - MDS

- MDS and Cluster are visual and relative methods to assess the relationship between samples
- No information on whether groups show statistically significant differences

2D Stress: 0.15



# ANOVA

parametric, univariate



# ANOSIM

(Analysis of similarities)

# PERMANOVA

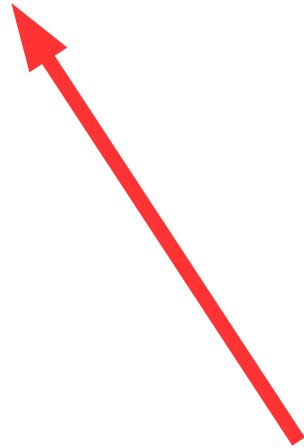
(Permutational Multivariate Analysis of Variance)

non-parametric, multivariate



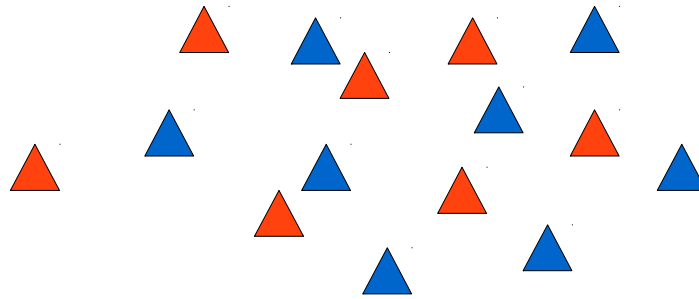
**ANOSIM**

$$R = \frac{(r_{\text{between}} - r_{\text{within}})}{\text{standardizing factor}}$$



Level of differentiation between samples

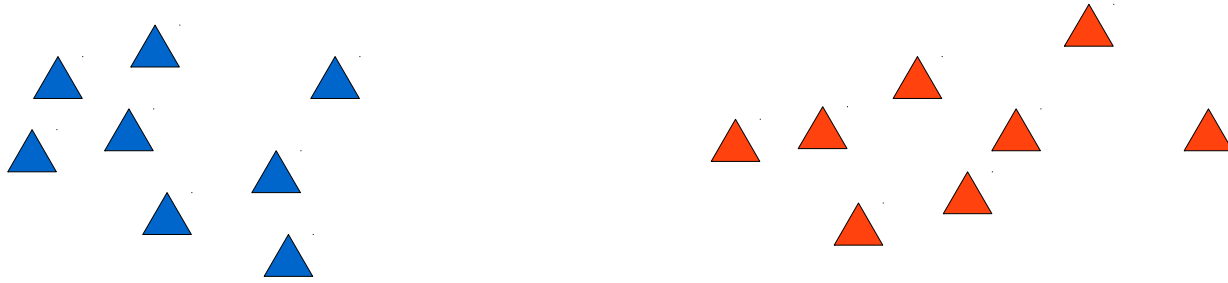
$H_0 =$  Groups are the same



Similarity within = Similarity between

$$R = \frac{(r_{\text{between}} - r_{\text{within}})}{\text{standardizing factor}} = 0$$

$H_a =$  Groups are different

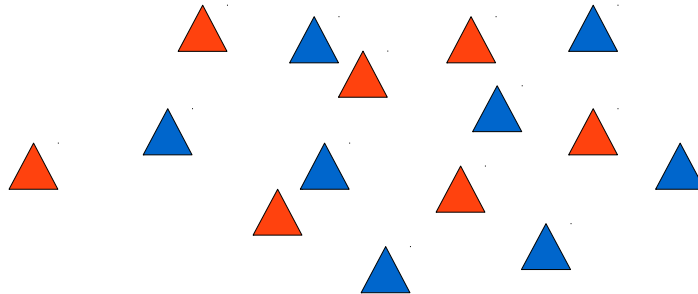


Similarity within > Similarity between

$$R = \frac{(r_{\text{between}} - r_{\text{within}})}{\text{standardizing factor}} = 1$$

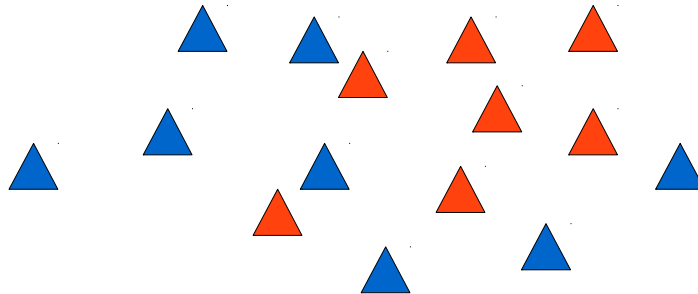
# Simulation of the null distribution (no differentiation between groups)

1. randomisation of samples
2. Calculation of R
3. 1000 repetitions
4. plotting of R values



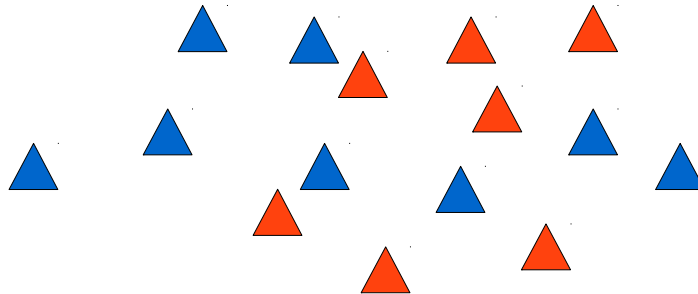
# Simulation of the null distribution (no differentiation between groups)

1. randomisation of samples
2. Calculation of R
3. 1000 repetitions
4. plotting of R values



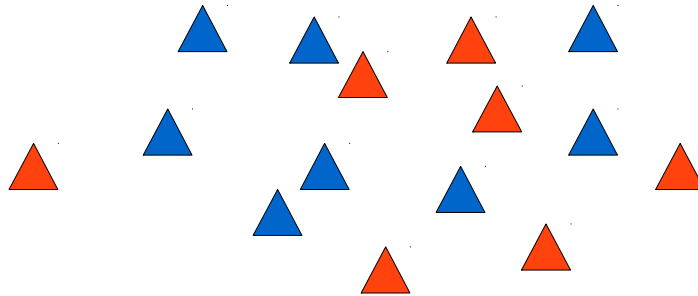
# Simulation of the null distribution (no differentiation between groups)

1. randomisation of samples
2. Calculation of R
3. 1000 repetitions
4. plotting of R values



# Simulation of the null distribution (no differentiation between groups)

1. randomisation of samples
2. Calculation of R
3. 1000 repetitions
4. plotting of R values





*Global Test*

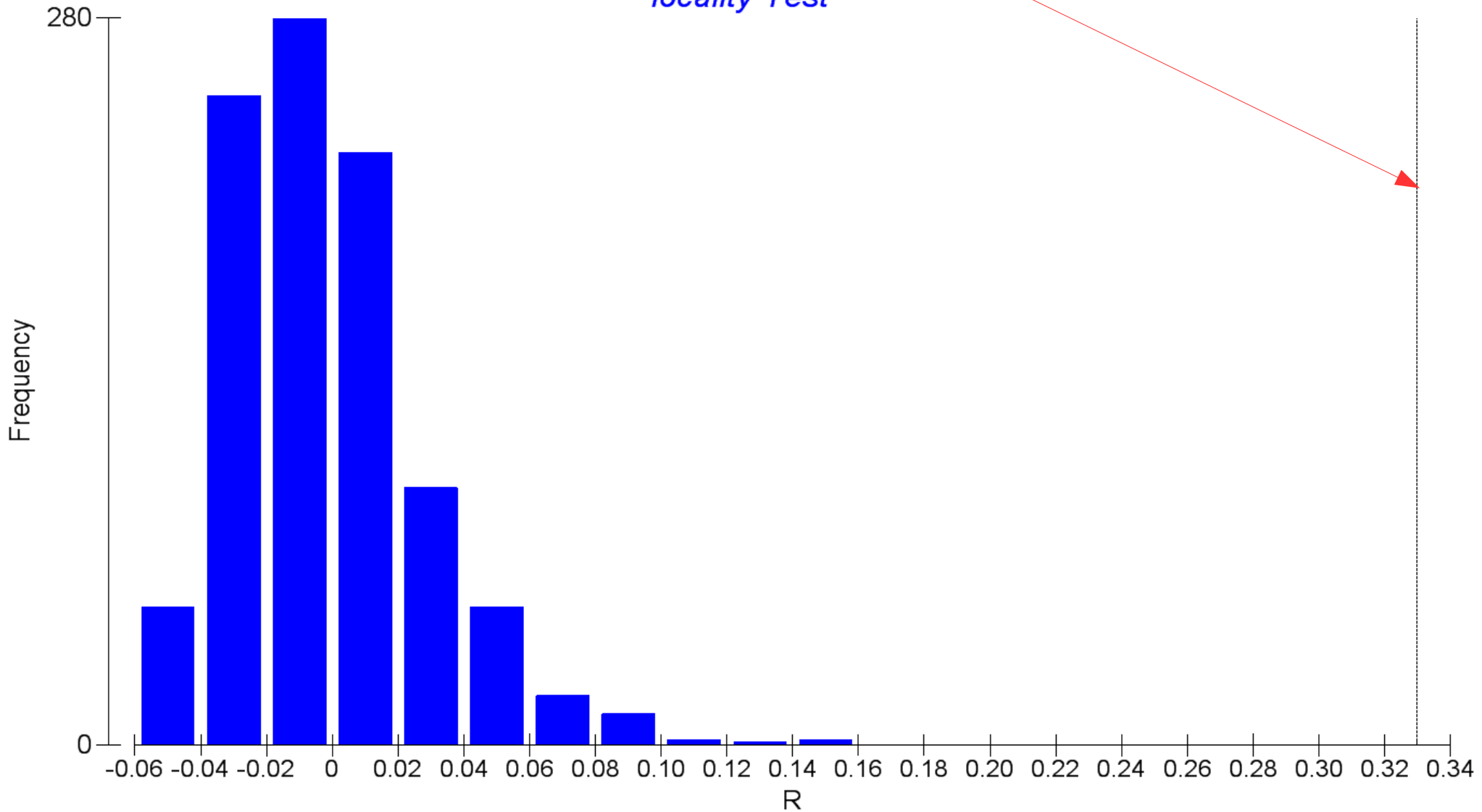
Sample statistic (Global R): 0.33

Significance level of sample statistic: 0.1%

Number of permutations: 999 (Random sample from a large number)

Number of permuted statistics greater than or equal to Global R: 0

*locality Test*



Why are groups different?

# SIMPER

(Similarity percentages)

identifies species that are responsible  
for differences between groups

Group ALYKES

Average similarity: 37.40

Species	Av. Abund	Av. Sim	Sim/SD	Contrib%	Cum. %
Platynereis dumerilii (Audouin & Milne Edwards, 1834)	7.40	15.13	1.18	40.45	40.45
Nereis pulsatoria (Savigny, 1822)	4.55	10.26	1.43	27.44	67.88
Syllis corallicola Verrill, 1900	5.10	7.72	0.86	20.64	88.52
Syllis gracilis Grube, 1840	0.85	1.24	0.50	3.32	91.84

Group ELOUNDA

Average similarity: 21.04

Species	Av. Abund	Av. Sim	Sim/SD	Contrib%	Cum. %
Platynereis dumerilii (Audouin & Milne Edwards, 1834)	3.79	16.39	0.80	77.87	77.87
Syllis corallicola Verrill, 1900	0.58	1.01	0.32	4.79	82.67
Amphiglana mediterranea (Leydig, 1851)	1.79	0.92	0.19	4.39	87.06
Opisthosyllis brunnea Langerhans, 1879	1.05	0.82	0.27	3.87	90.93

Groups ALYKES & ELOUNDA

Average dissimilarity = 81.79

Species	Group ALYKES		Group ELOUNDA		Contrib%	Cum. %
	Av. Abund	Av. Sim	Av. Abund	Av. Diss		
Platynereis dumerilii (Audouin & Milne Edwards, 1834)	7.40	15.13	3.79	19.82	24.23	24.23
Syllis corallicola Verrill, 1900	5.10	7.72	0.58	13.70	16.75	40.98
Nereis pulsatoria (Savigny, 1822)	4.55	10.26	0.16	13.39	16.37	57.35
Amphiglana mediterranea (Leydig, 1851)	0.85	1.24	1.79	4.24	5.18	62.53
Perinereis cultrifera (Grube, 1840)	0.70	1.01	0.71	3.71	4.54	67.07
Lysidice ninetta Audouin & Milne-Edwards. 1833	1.00	0.50	0.00	2.94	3.60	70.67

