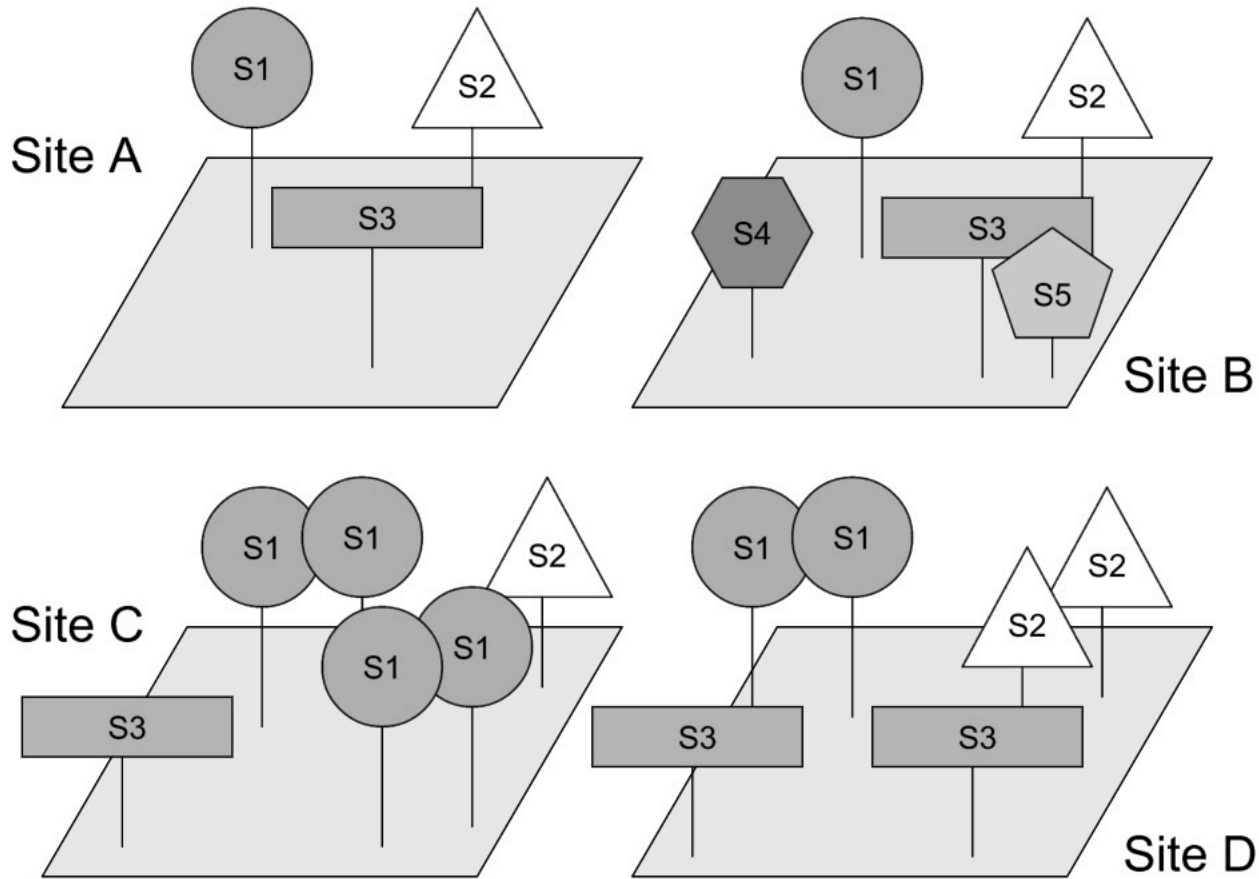


Diversity and community dissimilarity

Block Course

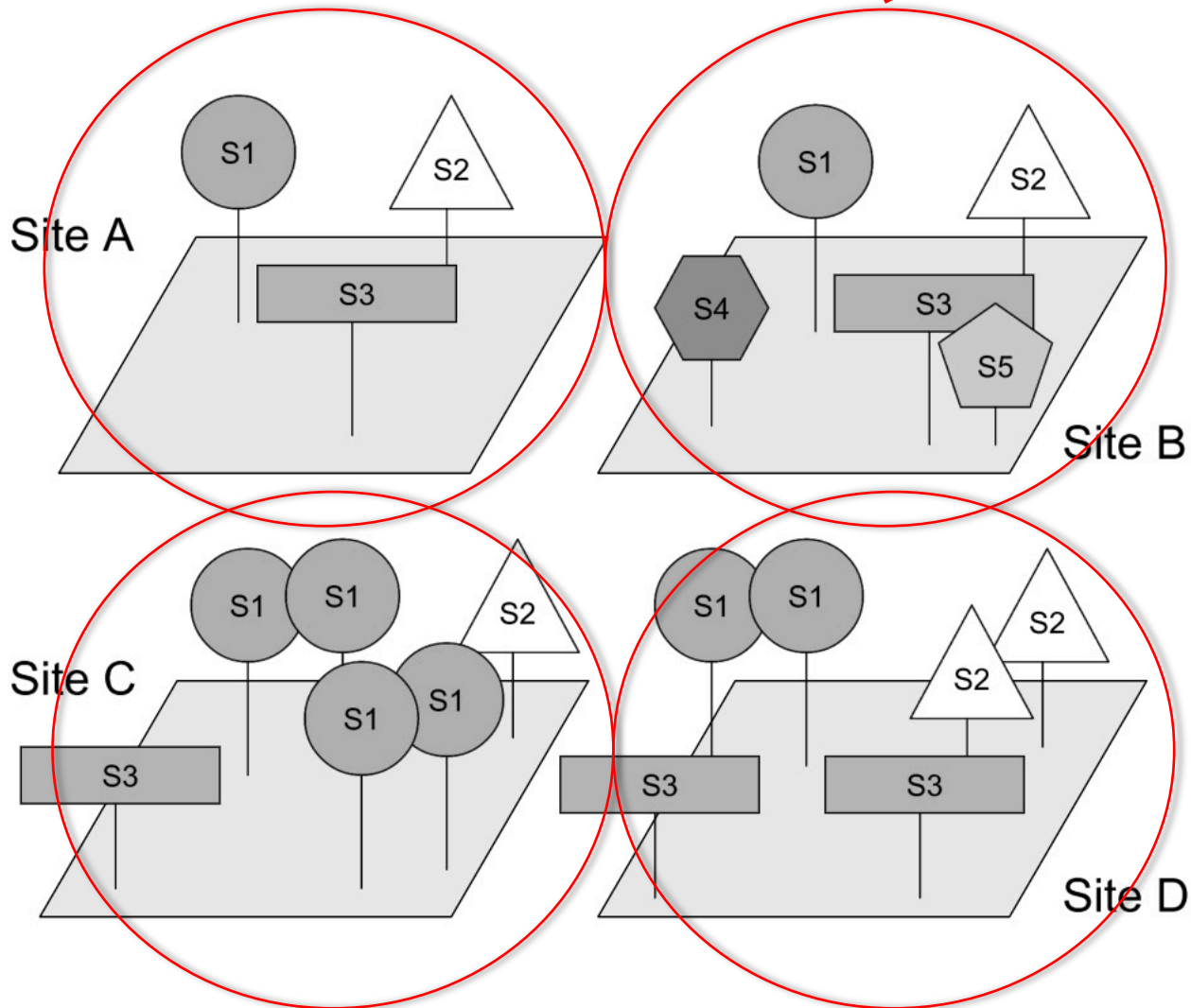
Guillem Salazar
(Sunagawa lab)

Diversity

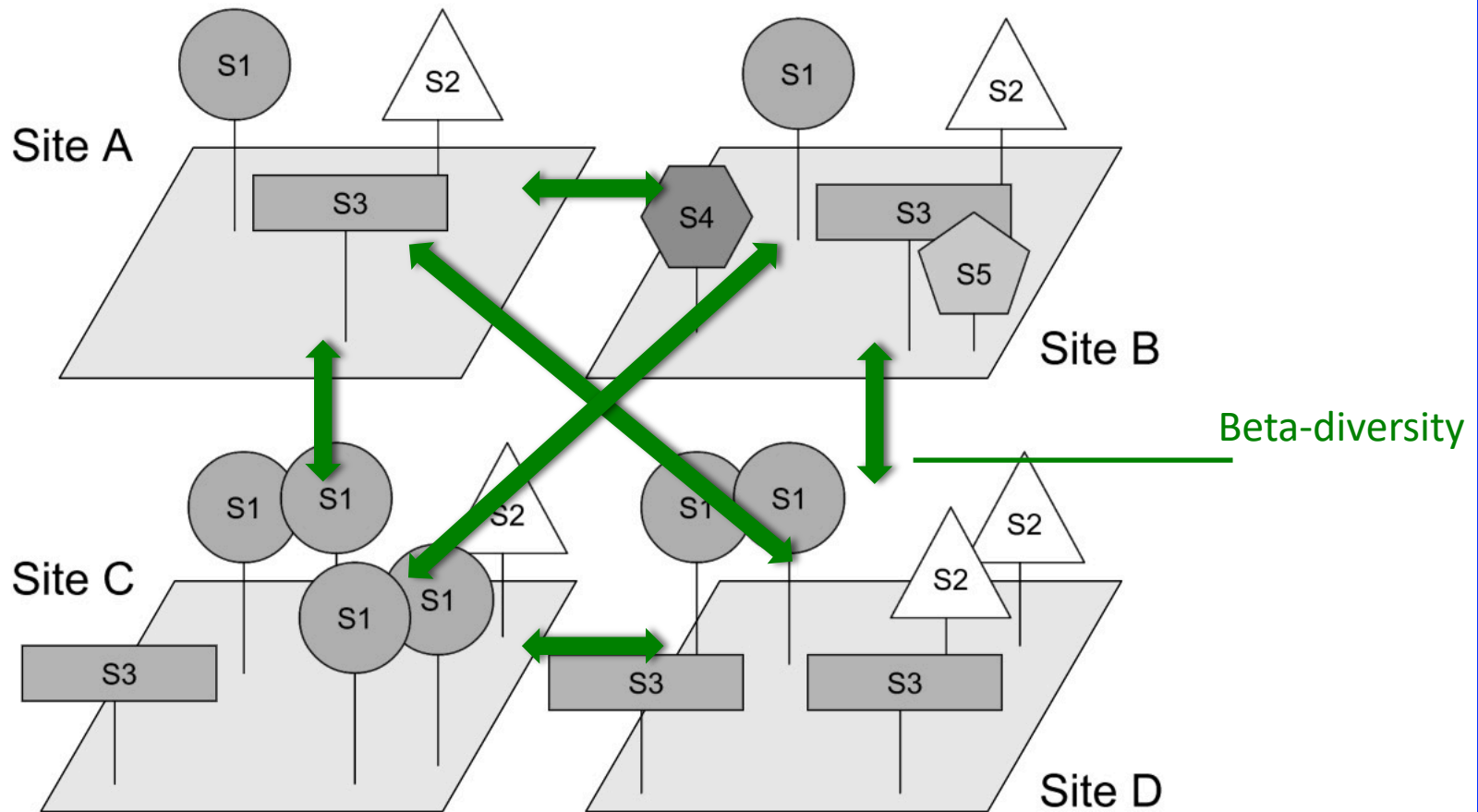


Diversity

Alpha-diversity

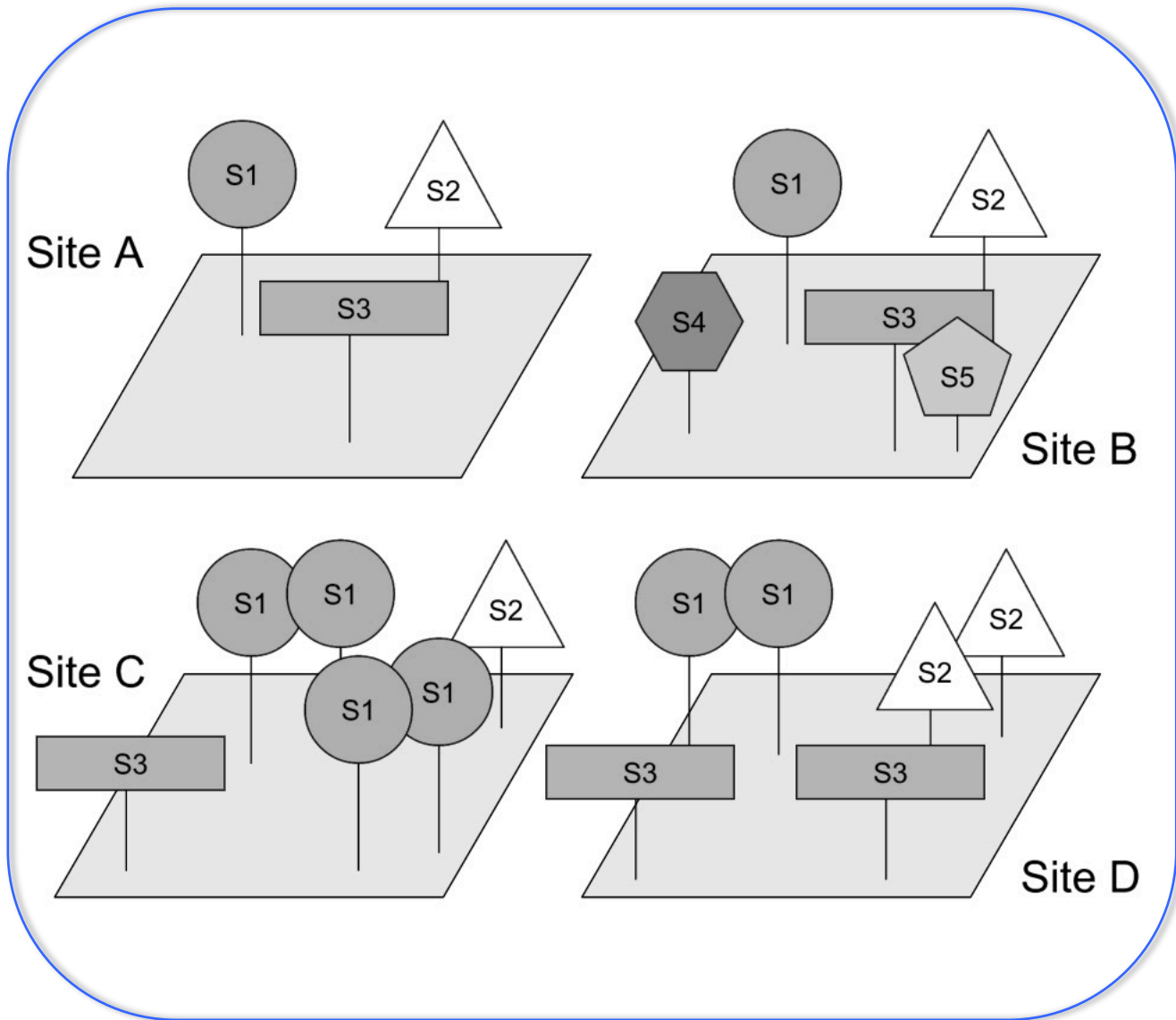


Diversity



Diversity

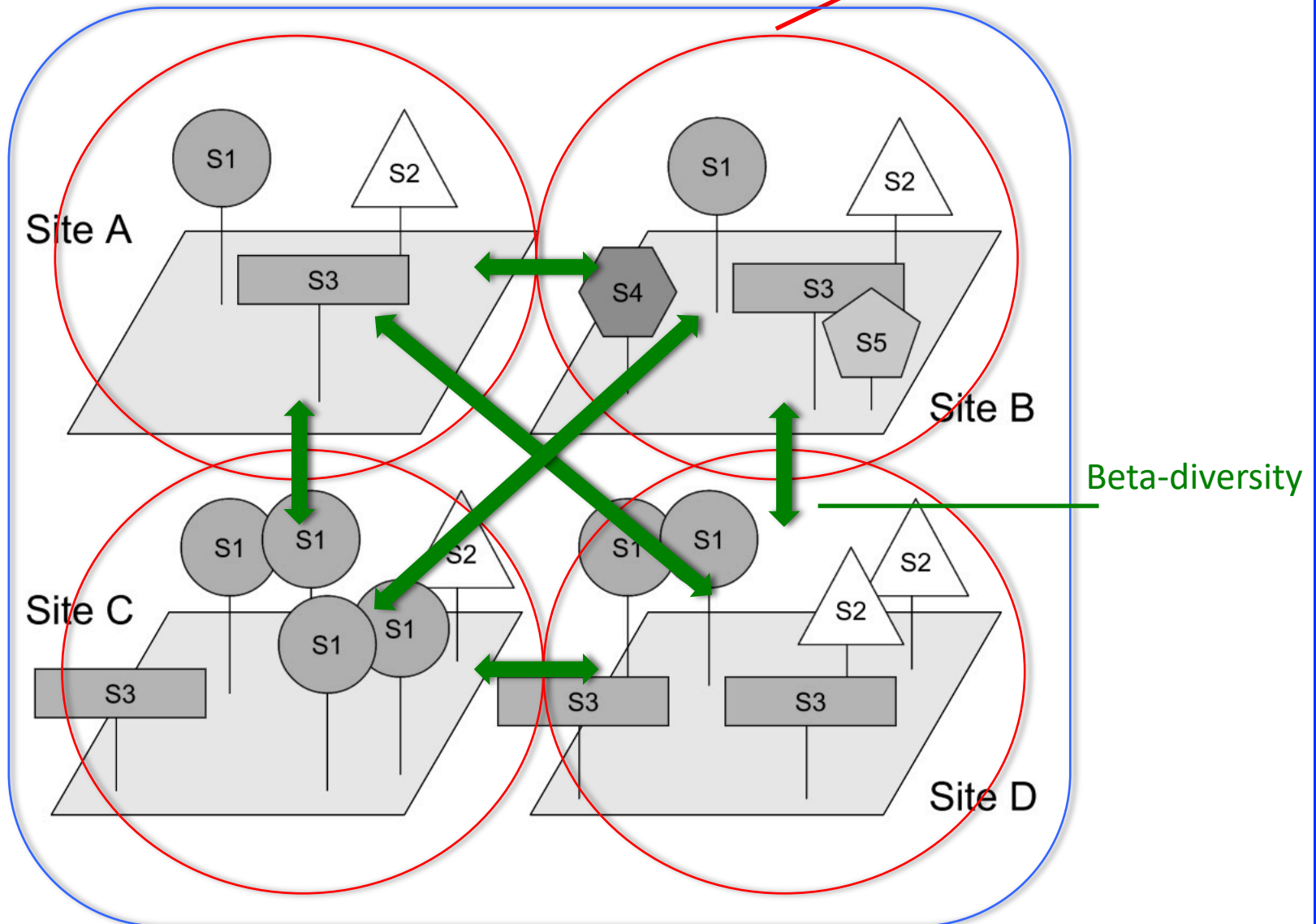
Gamma-diversity



Diversity

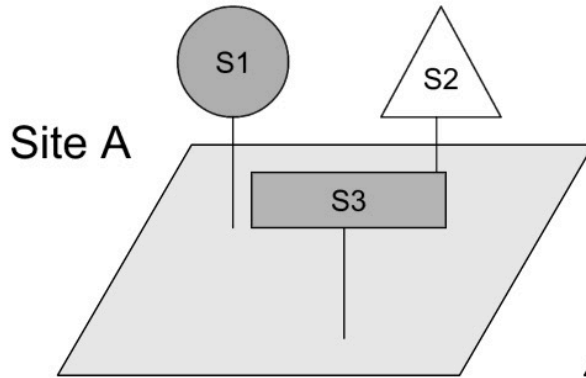
Gamma-diversity

Alpha-diversity

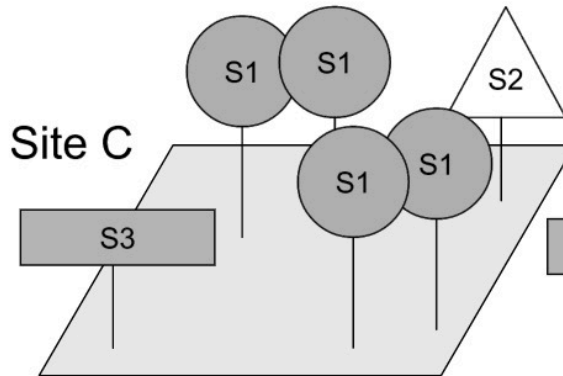
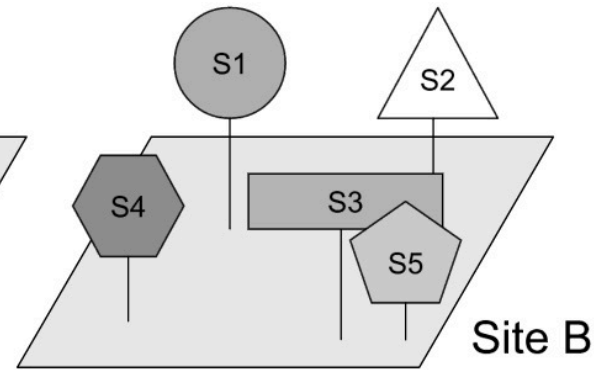


Alphadiversity \sim Richness + Evenness

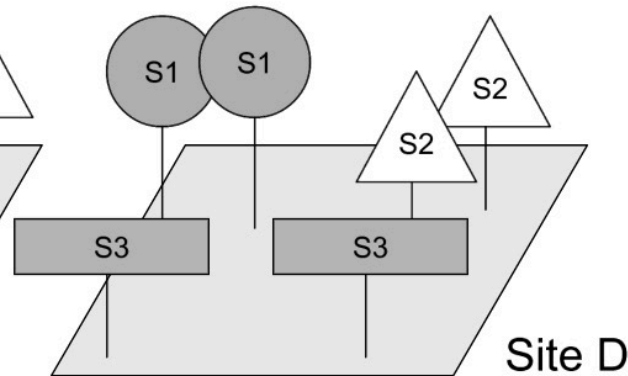
Richness = 3
High evenness



Richness = 5
High evenness



Richness = 3
Low evenness



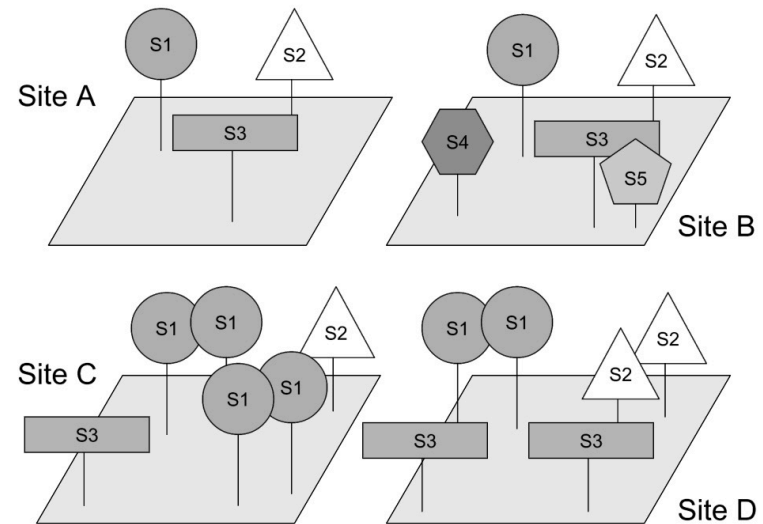
Richness = 3
High evenness

Alphadiversity ~ Richness + Evenness

Shannon diversity:

$$H' = - \sum_{i=1}^R p_i \ln p_i$$

where p_i is the relative abundance of species i



Site A:

$$H' = -(1/3 * \ln(1/3) + 1/3 * \ln(1/3) + 1/3 * \ln(1/3)) = \mathbf{1.0986}$$

Site B:

$$H' = -(1/5 * \ln(1/5) + 1/5 * \ln(1/5) + 1/5 * \ln(1/5) + 1/5 * \ln(1/5) + 1/5 * \ln(1/5)) = \mathbf{1.6094}$$

Site C:

$$H' = -(4/6 * \ln(4/6) + 1/6 * \ln(1/6) + 1/6 * \ln(1/6)) = \mathbf{0.8676}$$

Community dissimilarity

The formulae for calculating the ecological distances are:

$$\text{Bray-Curtis: } D = 1 - 2 \frac{\sum_{i=1}^S \min(a_i, c_i)}{\sum_{i=1}^S (a_i + c_i)}$$

$$\text{Kulczynski: } D = 1 - \frac{1}{2} \left(\frac{\sum_{i=1}^S \min(a_i, c_i)}{\sum_{i=1}^S a_i} + \frac{\sum_{i=1}^S \min(a_i, c_i)}{\sum_{i=1}^S c_i} \right)$$

$$\text{Euclidean: } D = \sqrt{\sum_{i=1}^S (a_i - c_i)^2}$$

$$\text{Chi-square: } D = \sqrt{\sum_{i=1}^S \frac{(a_i + c_i)}{(a_+ + c_+)} \left(\frac{a_i}{a_+} - \frac{c_i}{c_+} \right)^2} \text{ with } a_+ = \sum_{i=1}^S a_i$$

$$\text{Hellinger: } D = \sqrt{\sum_{i=1}^S \left(\sqrt{\frac{a_i}{a_+}} - \sqrt{\frac{c_i}{c_+}} \right)^2} \text{ with } a_+ = \sum_{i=1}^S a_i$$

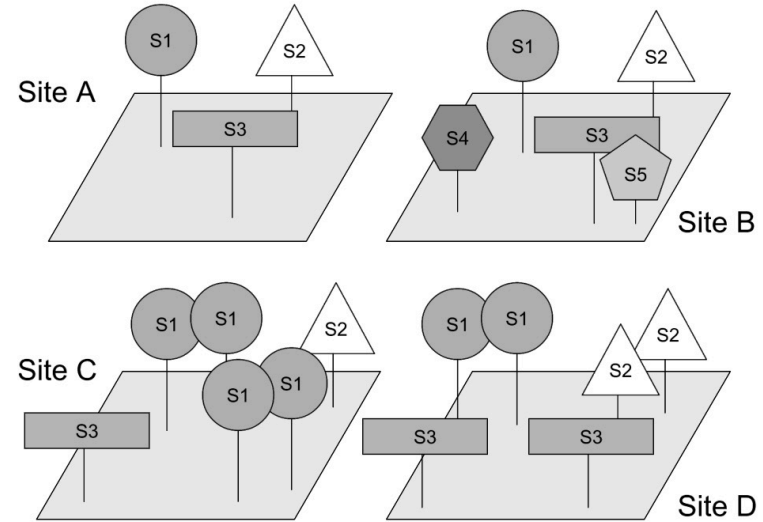
Community dissimilarity (beta-diversity)

Index	Equation	Properties
Jaccard	$S_7 = \frac{a}{a+b+c}$	Compares the number of shared species to the number of species in the combined assemblages placing more emphasis on taxa not shared between sites
Sørensen	$S_8 = \frac{2a}{(2a+b+c)}$	Compares the number of shared species to the mean number of species in a single assemblage placing more emphasis on similarity of samples owing to shared species

In the above table, a = the number of species shared between assemblages, b = the number of unique species in the first assemblage, and c = the number of unique species in the second assemblage.

Community dissimilarity (beta-diversity)

Index	Equation
Jaccard	$S_7 = \frac{a}{a+b+c}$
Sørensen	$S_8 = \frac{2a}{(2a+b+c)}$



In the above table, a = the number of species shared between assemblages, b = the number of unique species in the first assemblage, and c = the number of unique species in the second assemblage.

Similarity

Distance / Dissimilarity

Site A-B:

$$J = 3/(3+0+2) = \mathbf{0.6}$$

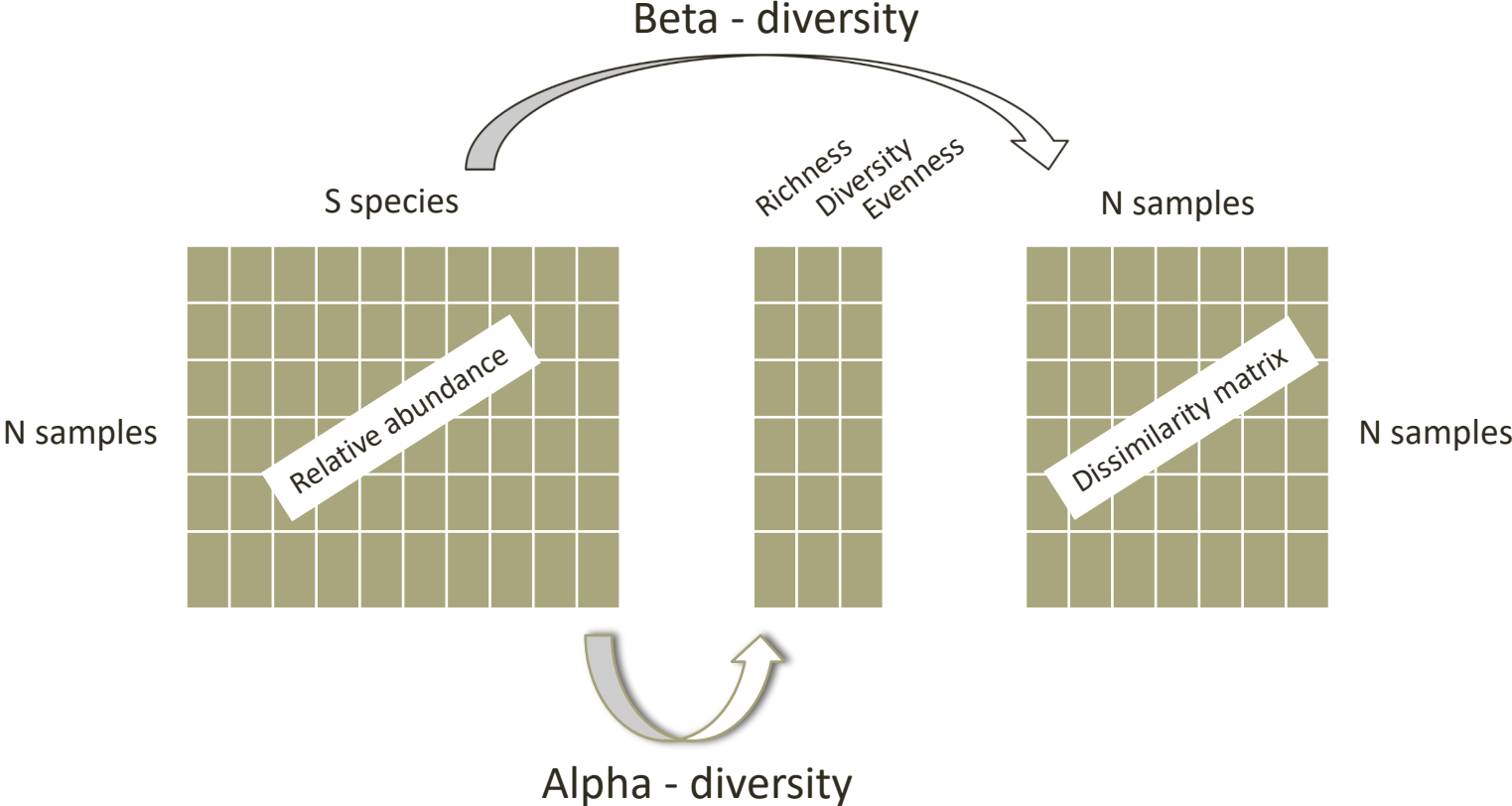
$$D = 1-J = \mathbf{0.4}$$

Site A-C:

$$J = 3/(3+0+0) = \mathbf{1}$$

$$D = 1-J = \mathbf{0}$$

Community dissimilarity



Community dissimilarity

