Analysing Biodiversity

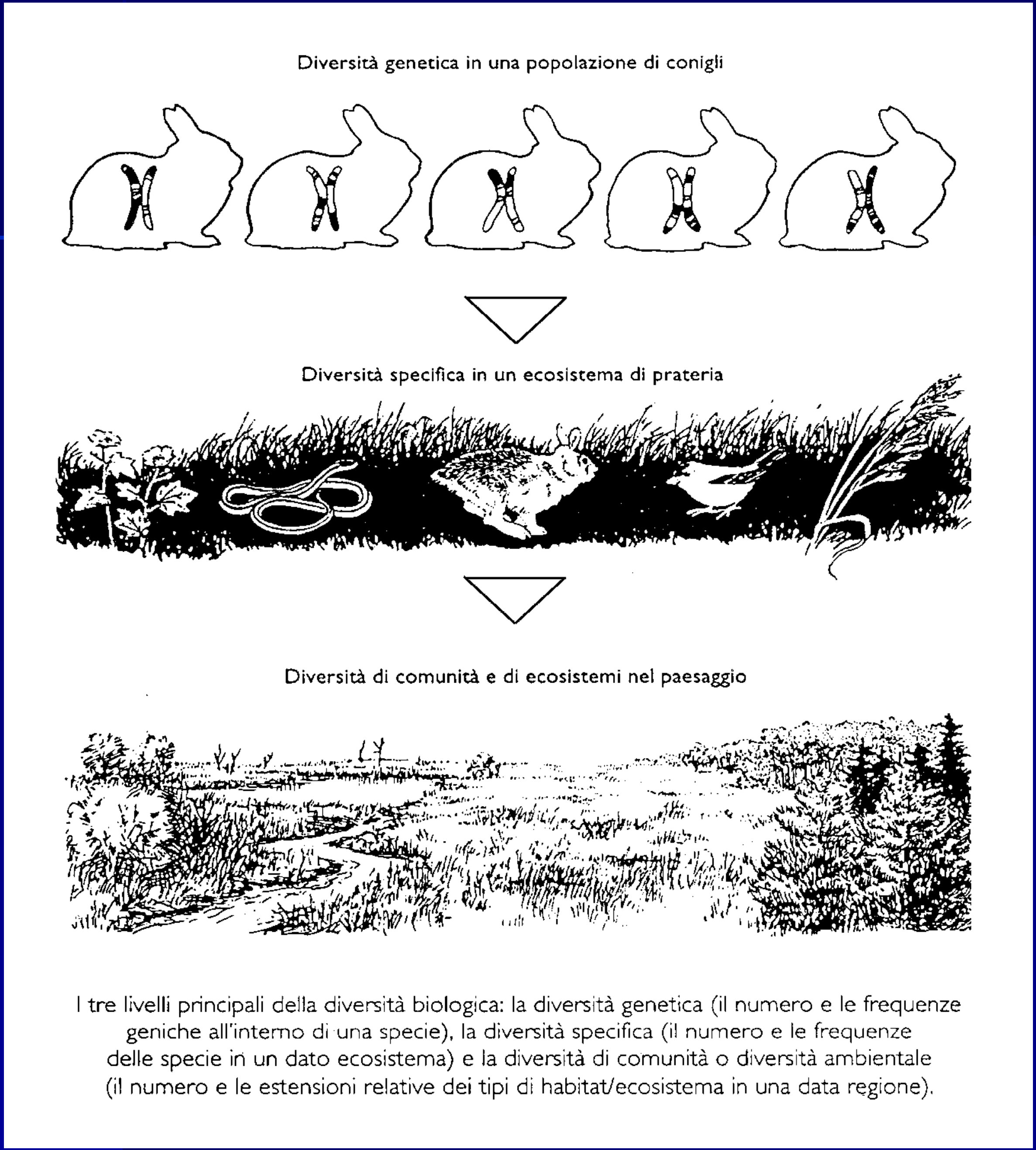
G. Bacaro

Design and Analysis of Environmental Monitoring and Experiments
Master Degree in Global Change Ecology
I Year, I term

What does biodiversity mean?

La diversità esprime la quantità di elementi diversi (tassonomici o morfologici) che costituiscono una comunità vegetale

Different level of organization



Diversità genetica in una popolazione di conigli

Diversità specifica in un ecosistema di prateria

Diversità di comunità e di ecosistemi nel paesaggio

I tre livelli principali della diversità biologica: la diversità genetica (il numero e le frequenze geniche all'interno di una specie), la diversità specifica (il numero e le frequenze delle specie in un dato ecosistema) e la diversità di comunità o diversità ambientale (il numero e le estensioni relative dei tipi di habitat/ecosistema in una data regione).

Genetic

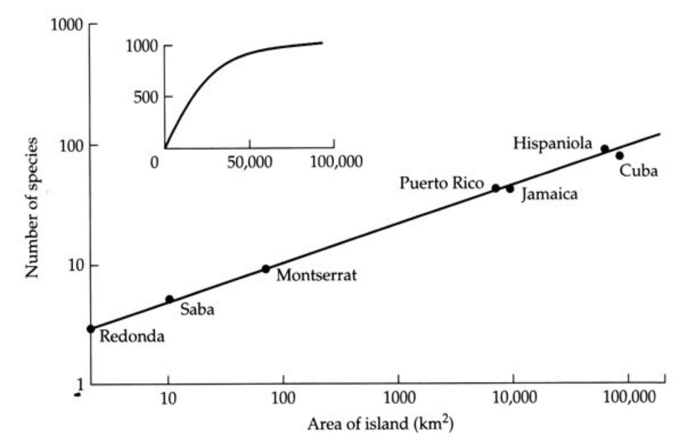
Specific

Community

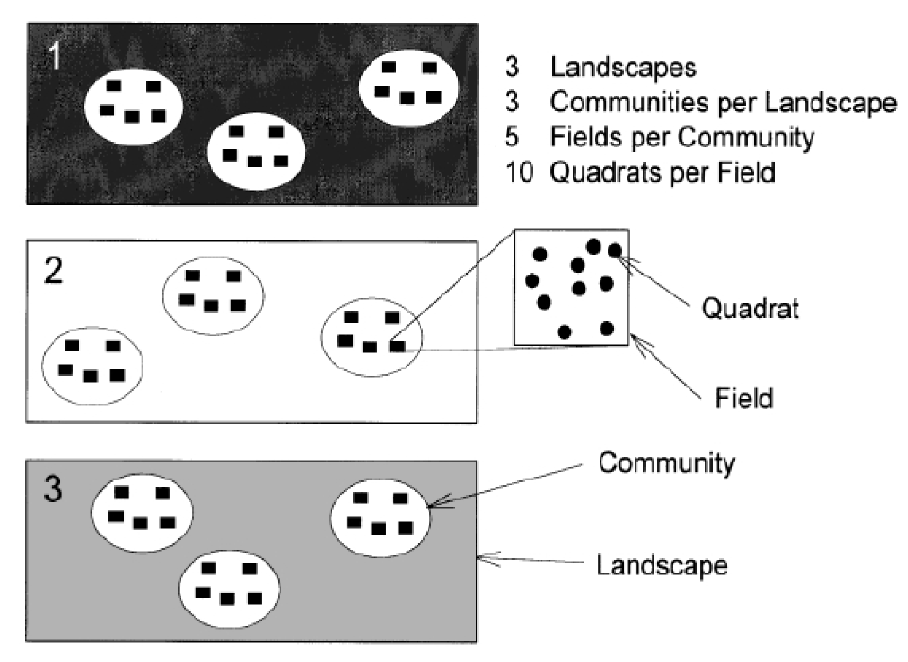
Increasing level of interaction

Definition of Spatial Scale

The effect of spatial scale on biodiversity is one of the oldest topics in ecology and biogeography



Spatial scale is not simply represented by increasing area, but also by interactions among the different scale components: grain and extent

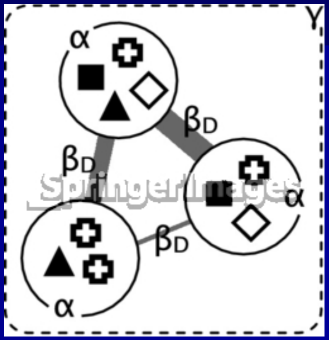


__Grain__ = standard spatial or temporal dimension of the collection unit

__Extent__ = geographic area of the samples, or the time span of the sample

Elements of Biodiversity in relation to Scale

Whittaker (1960) firstly propose the idea of partitioning diversity into spatial components



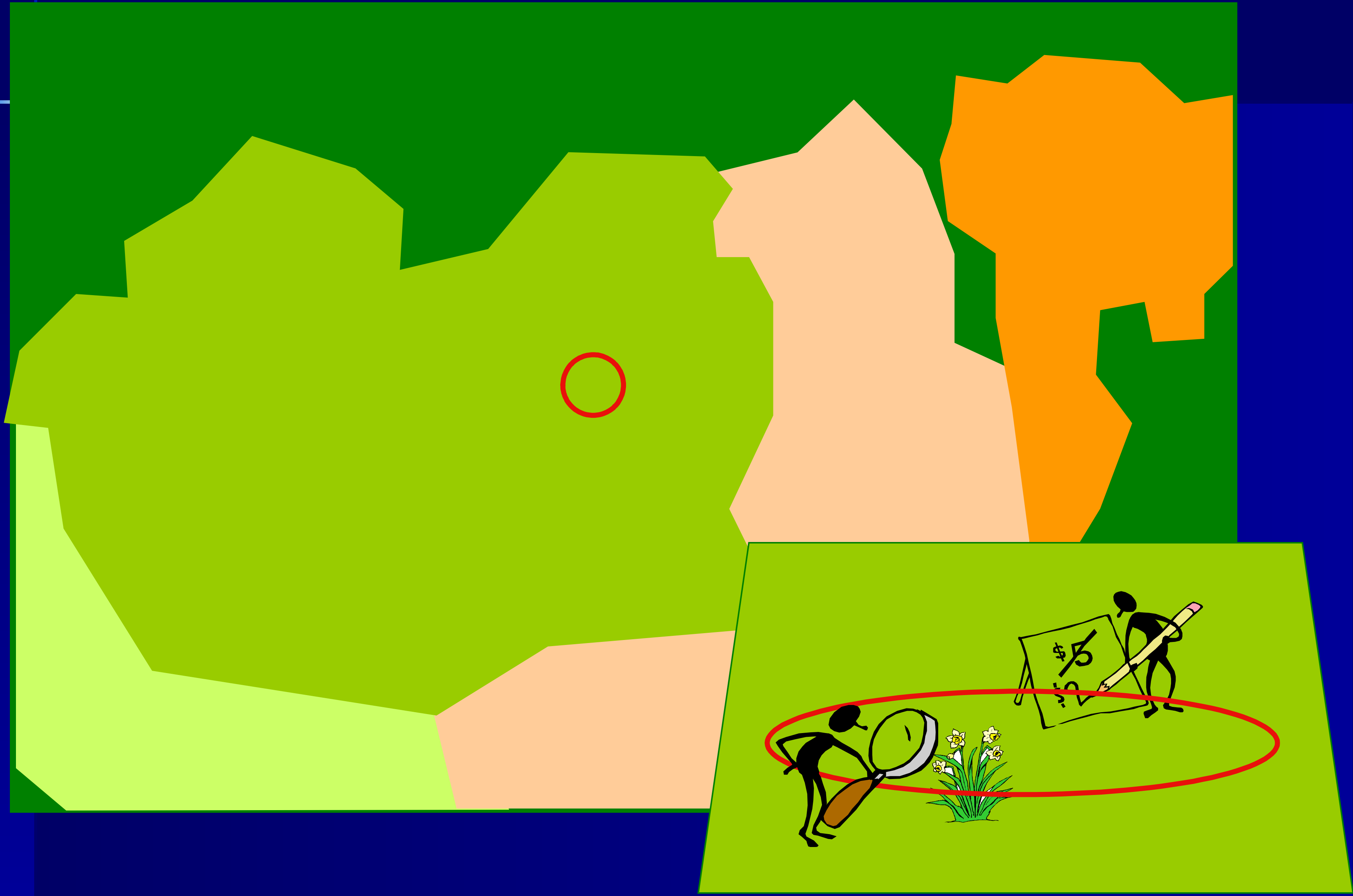
α-diversity -> *species diversity measured at community*
γ -diversity -> *species diversity measured at landscape extent*
β-diversity -> difference between γ and the average richness
within all communities (α-mean)

Additive partition	$\beta = \gamma - \alpha$-mean (Lande 1996)
Multiplicative partition	$\beta = \gamma / \alpha$-mean (Whittaker 192)

Additive partitioning is considered convenient: it provide measures of the different components based on the same metrics, i.e. species richness

Diversity partitioning across scales is fundamental to understand the spatial and ecological processes structuring biological communities

α-diversity



α-diversity

Plot	139
Numero di specie	2
Pseudotsuga menziesii (Mirbel) Franco	90
Abies alba Miller	10

Plot	156
Numero di specie	17
Pinus nigra Arnold	50
Castanea sativa Miller	30
Brachypodium rupestre (Host) R. et S.	2
Clematis vitalba L.	1
Rubus ulmifolius Schott	0.3
Euonymus europaeus L.	0.3
Rosa canina L. sensu Bouleng.	0.1
Vicia cracca L.	0.02
Cruciata glabra (L.) Ehrend.	0.02
Stellaria media (L.) Vill.	0.01
Lathyrus montanus Bernh.	0.01
Hypericum perforatum L.	0.01
Geum urbanum L.	0.01
Galium aparine L.	0.01
Galium album Miller	0.01
Crataegus monogyna Jacq.	0.01
Clinopodium vulgare L.	0.01

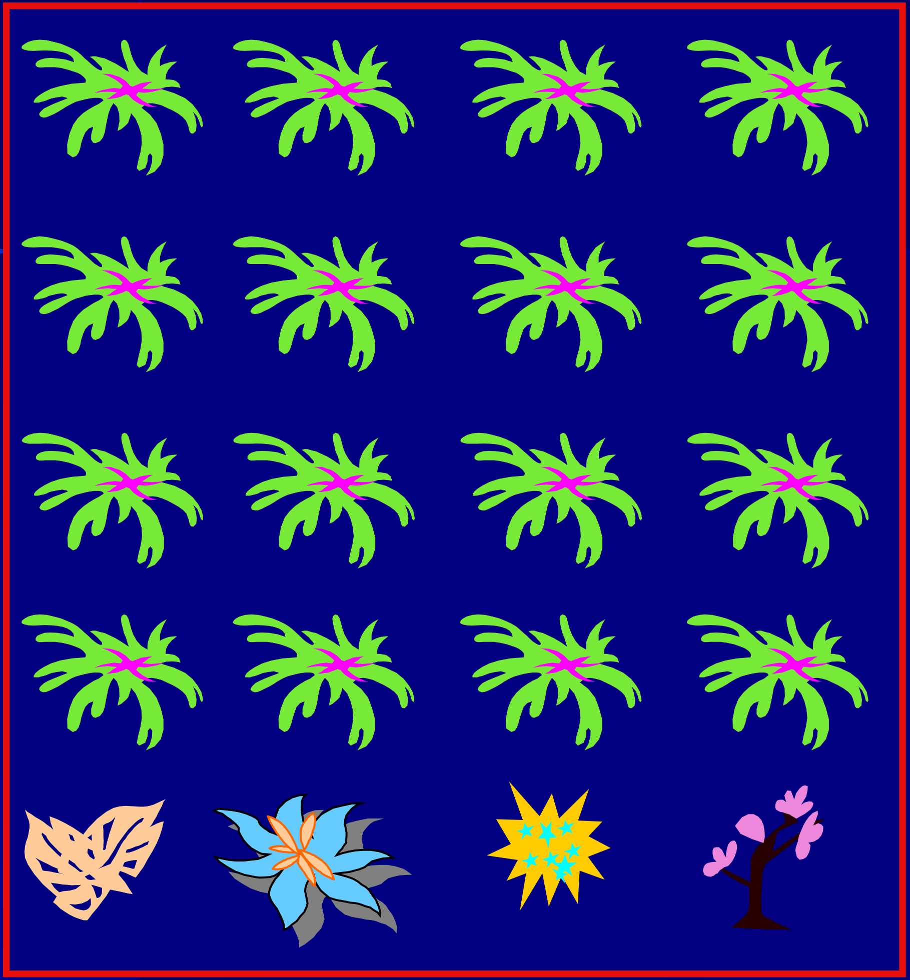
Plot	70
Numero di specie	44
Bromus erectus Hudson	40
Prunus spinosa L.	30
Trifolium incarnatum L.	5
Phleum pratense L.	5
Festuca arundinacea Schreber	5
Arrhenatherum elatius (L.) Presl	5
Vicia villosa Roth ssp. villosa	3
Thymus longicaulis Presl	3
Crataegus monogyna Jacq.	3
Brachypodium rupestre (Host) R. et S.	3
Rosa agrestis Savi	2
Dactylis hispanica Roth	2
Rhinanthus alectorolophus (Scop.) Pollich	1
Medicago sativa L. ssp. falcata (L.) Arcang.	1
Trifolium campestre Schreber	0.5
Medicago rigidula (L.) All.	0.5
Hypericum perforatum L.	0.5
Dorycnium hirsutum (L.) Ser.	0.5
Achillea collina Becker	0.5
Eryngium amethystinum L.	0.2
Dianthus carthusianorum L. ssp. carthusianorum	0.2
Helleborus foetidus L.	0.1
Geranium columbinum L.	0.1
Bunium bulbocastanum L.	0.05
Trifolium scabrum L.	0.03
Arabis hirsuta (L.) Scop.	0.03
Taraxacum laevigatum (Willd.) DC. (aggregato)	0.02
Sanguisorba minor Scop.	0.02
Plantago lanceolata L.	0.02
Lathyrus aphaca L.	0.02
Knautia purpurea (Vill.) Borbas	0.02
Euphorbia cyparissias L.	0.02
Erophila verna (L.) Chevall.	0.02
Cynosurus echinatus L.	0.02
Cerastium arvense L. ssp.arvense var etruscum Fiori	0.02
Bupleurum baldense Turra	0.02
Allium sphaerocephalon L.	0.02
Torilis japonica (Houtt.) DC.	0.01
Stachys salviifolia Ten.	0.01
Silene vulgaris (Moench) Garcke ssp. vulgaris	0.01

	Plot	139	156	70
Pinaceae	Pseudotsuga menziesii (Mirbel) Franco	90		
Pinaceae	Abies alba Miller	10		
Pinaceae	Pinus nigra Arnold		50	
Fagaceae	Castanea sativa Miller		30	
Graminaceae	Brachypodium rupestre (Host) R. et S.		2	3
Ranunculaceae	Clematis vitalba L.		1	
Celastraceae	Euonymus europaeus L.		0.3	
Rosaceae	Rubus ulmifolius Schott		0.3	
Rosaceae	Rosa canina L. sensu Bouleng.		0.1	
Rubiaceae	Cruciata glabra (L.) Ehrend.		0.02	
Leguminosae	Vicia cracca L.		0.02	
Rosaceae	Crataegus monogyna Jacq.		0.01	3
Guttiferae	Hypericum perforatum L.		0.01	0.5
Labiatae	Clinopodium vulgare L.		0.01	
Rubiaceae	Galium album Miller		0.01	
Rubiaceae	Galium aparine L.		0.01	
Rosaceae	Geum urbanum L.		0.01	
Leguminosae	Lathyrus montanus Bernh.		0.01	
Caryophyllacea	Stellaria media (L.) Vill.		0.01	
Graminaceae	Bromus erectus Hudson			40
Rosaceae	Prunus spinosa L.			30
Graminaceae	Arrhenatherum elatius (L.) Presl			5
Graminaceae	Festuca arundinacea Schreber			5
Graminaceae	Phleum pratense L.			5
Leguminosae	Trifolium incarnatum L.			5
Labiatae	Thymus longicaulis Presl			3
Leguminosae	Vicia villosa Roth ssp. villosa			3
Graminaceae	Dactylis hispanica Roth			2
Rosaceae	Rosa agrestis Savi			2
Leguminosae	Medicago sativa L. ssp. falcata (L.) Arcang.			1
Scrophulariacea	Rhinanthus alectorolophus (Scop.) Pollich			1
Compositae	Achillea collina Becker			0.5
Leguminosae	Dorycnium hirsutum (L.) Ser.			0.5
Leguminosae	Medicago rigidula (L.) All.			0.5
Leguminosae	Trifolium campestre Schreber			0.5
Caryophyllacea	Dianthus carthusianorum L. ssp. carthusianorum			0.2
Umbelliferae	Eryngium amethystinum L.			0.2
Geraniaceae	Geranium columbinum L.			0.1
Ranunculaceae	Helleborus foetidus L.			0.1
Umbelliferae	Bunium bulbocastanum L.			0.05
Cruciferae	Arabis hirsuta (L.) Scop.			0.03

	A_1	A_2	A_3	.	.	A_n
S_1	$X11$	$X12$	$X13$	.	.	$X1n$
S_2	$X21$	$X22$	$X23$	.	.	$X2n$
S_3	$X31$	$X32$	$X33$	.	.	$X3n$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
S_m	$Xm1$	$Xm2$	$Xm3$	.	.	Xmn

Plot	139	156	70
Numero di specie	2	17	44
Numero di Famiglie	1	11	16

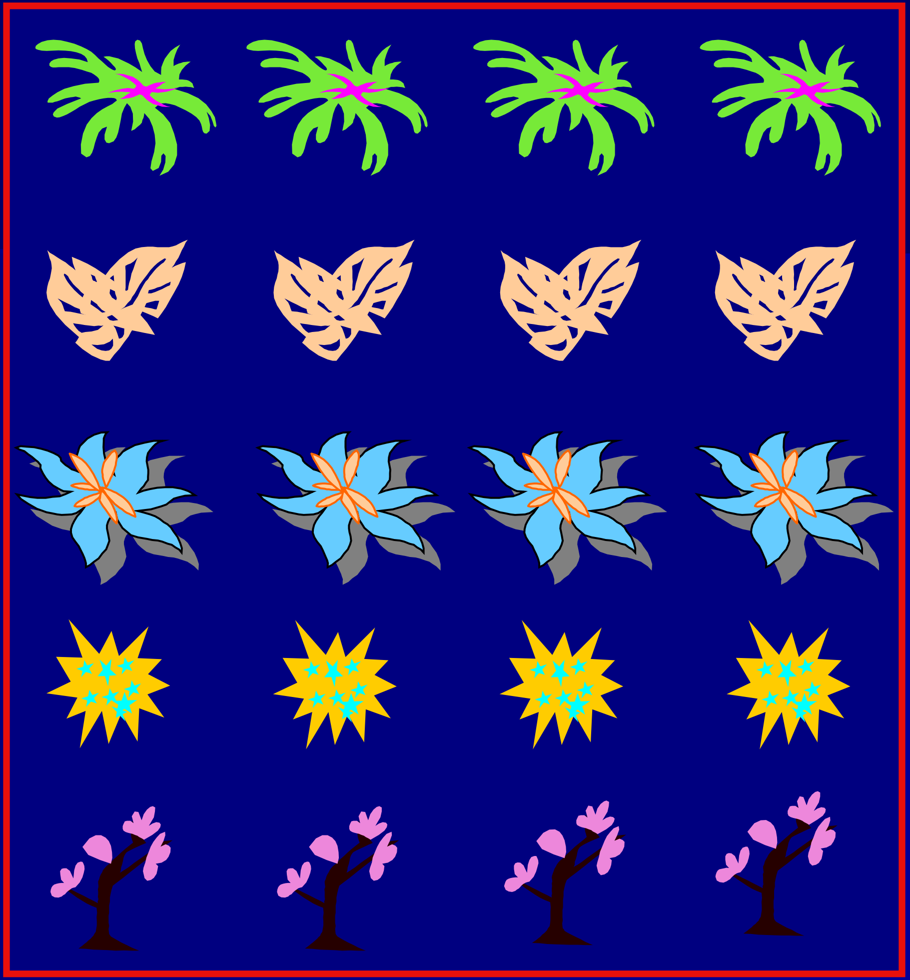
A



$S = 5$ $n = 20$

(16, 1, 1, 1,1)

B



$S = 5$ $n = 20$

(4, 4, 4, 4, 4)

Species richness and evenness: describing α-diversity

+ *<u>α-diversity</u>* (alpha) can easily measured considering the number of taxonomic elements (species, families, etc) and their abundance in a specified assemblage or delimited area

+ Species richness

+ Shannon Index

+ Simpson Index

+ Pielou Index

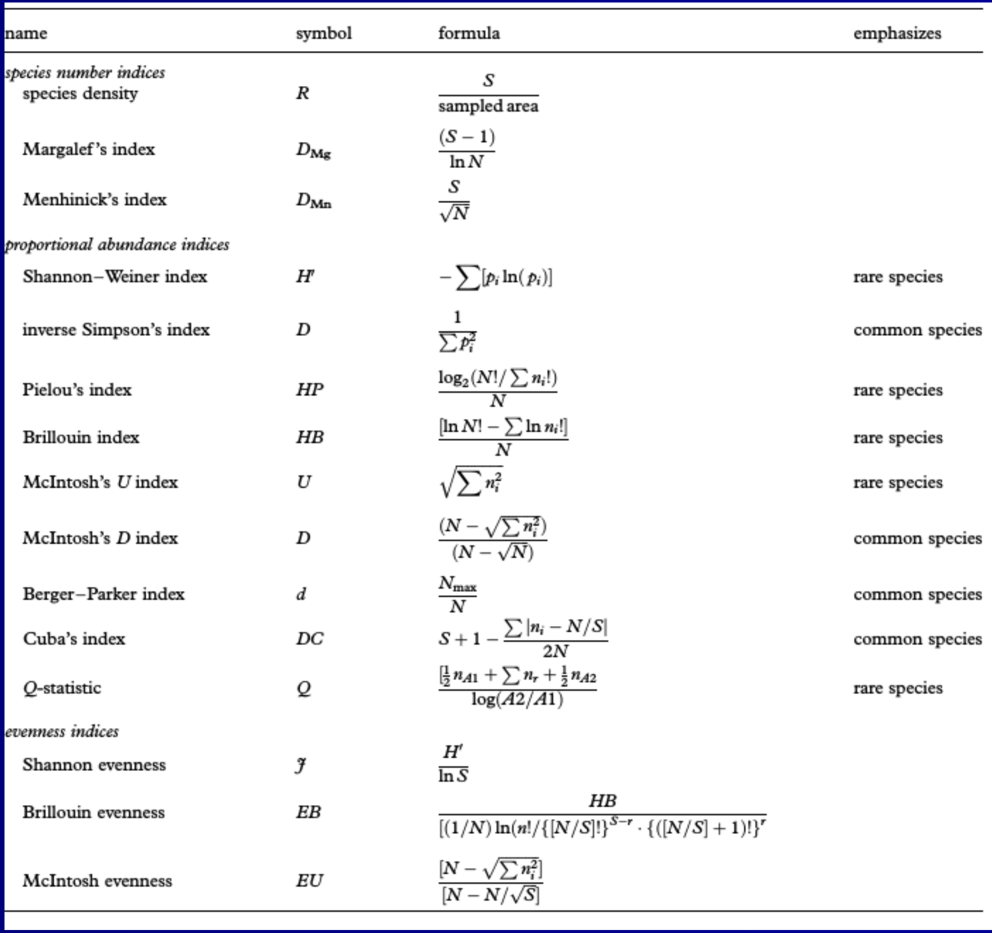
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2011) **366**, 2426–2437
doi:10.1098/rstb.2011.0065

Research

Old and new challenges in using species diversity for assessing biodiversity

Alessandro Chiarucci[1,*], Giovanni Bacaro[1] and Samuel M. Scheiner[2]



name	symbol	formula	emphasizes
species number indices			
species density	R	$\frac{S}{\text{sampled area}}$	
Margalef's index	D_{Mg}	$\frac{(S-1)}{\ln N}$	
Menhinick's index	D_{Mn}	$\frac{S}{\sqrt{N}}$	
proportional abundance indices			
Shannon–Weiner index	H'	$-\sum[p_i \ln(p_i)]$	rare species
inverse Simpson's index	D	$\frac{1}{\sum p_i^2}$	common species
Pielou's index	HP	$\frac{\log_2(N!/\sum n_i!)}{N}$	rare species
Brillouin index	HB	$\frac{[\ln N! - \sum \ln n_i!]}{N}$	rare species
McIntosh's U index	U	$\sqrt{\sum n_i^2}$	rare species
McIntosh's D index	D	$\frac{(N-\sqrt{\sum n_i^2})}{(N-\sqrt{N})}$	common species
Berger–Parker index	d	$\frac{N_{\max}}{N}$	common species
Cuba's index	DC	$S+1-\frac{\sum \lvert n_i - N/S \rvert}{2N}$	common species
Q-statistic	Q	$\frac{[\frac{1}{2} n_{A1} + \sum n_r + \frac{1}{2} n_{A2}}{\log(A2/A1)}$	rare species
evenness indices			
Shannon evenness	J	$\frac{H'}{\ln S}$	
Brillouin evenness	EB	$\frac{HB}{[(1/N)\ln(n!/\{[N/S]!\}^{S-r} \cdot \{([N/S]+1)!\}^{r}}$	
McIntosh evenness	EU	$\frac{[N-\sqrt{\sum n_i^2}]}{[N-N/\sqrt{S}]}$	

Shannon Diversity index:

$$H' = -\sum_{s=1}^{S} p_i * Ln(p_i)$$

Abbondanze assolute delle singole classi nelle due comunità:

	Comunità A	Comunità B
Specie 1	16	4
Specie 2	1	4
Specie 3	1	4
Specie 4	1	4
Specie 5	1	4
Totale	20	20

Abbondanze relative (p_i) delle singole classi nelle due comunità:

	Comunità A	Comunità B
Specie 1	0.80	0.20
Specie 2	0.05	0.20
Specie 3	0.05	0.20
Specie 4	0.05	0.20
Specie 5	0.05	0.20
Totale	1.00	1.00

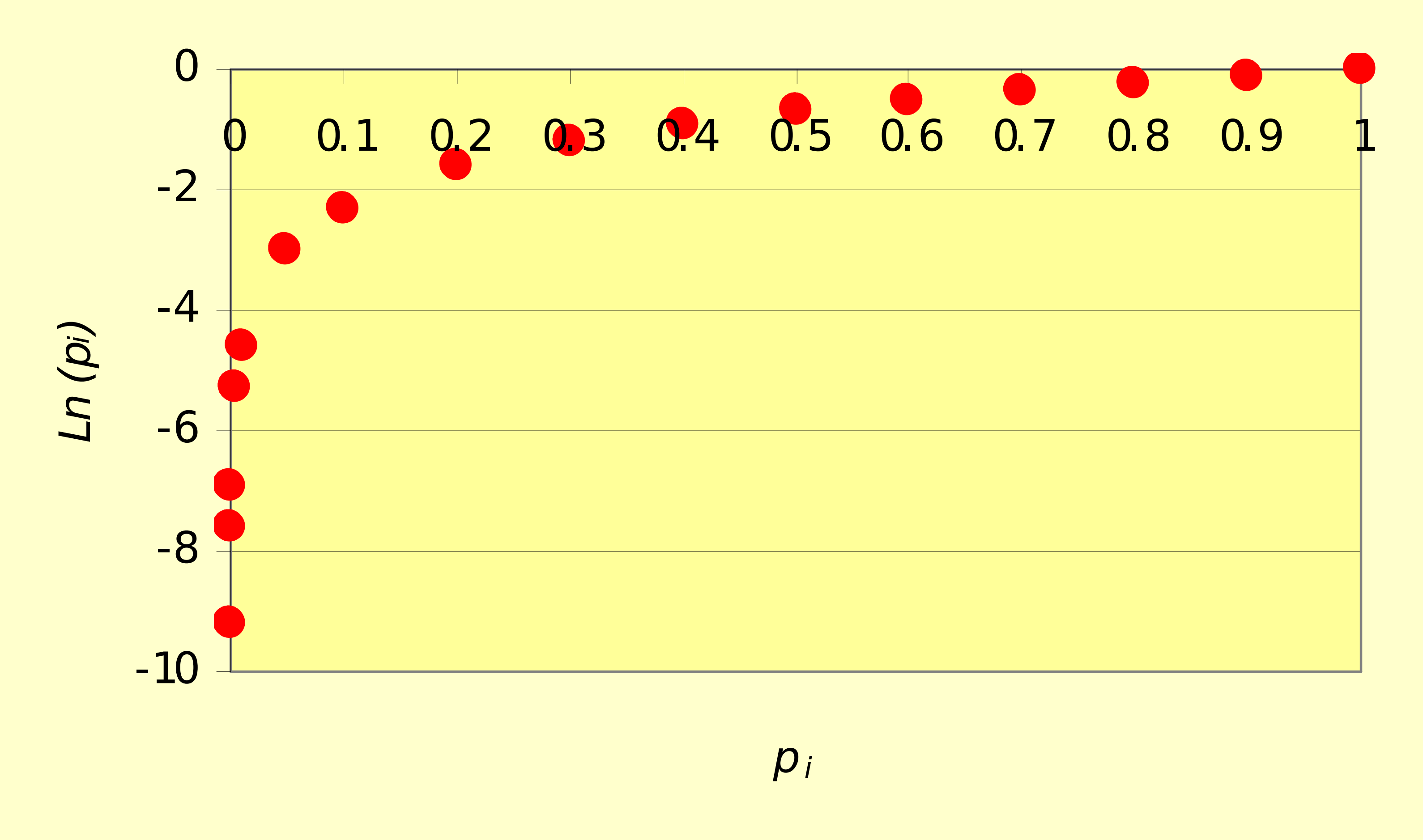
Passi per calcolare l’indice di diversità di Shannon *H’*:

1	Comunità A			Comunità B	
	p_i	$Ln(p_i)$		p_i	$Ln(p_i)$
Specie	0.80	-0.22		0.20	-1.61
Specie	0.05	-3.00		0.20	-1.61
Specie	0.05	-3.00		0.20	-1.61
Specie	0.05	-3.00		0.20	-1.61
Specie	0.05	-3.00		0.20	-1.61

2	Comunità A	Comunità B
	$pi*Ln(p_i)$	$pi*Ln(p_i)$
Specie	-0.18	-0.32
Specie	-0.15	-0.32
Specie	-0.15	-0.32
Specie	-0.15	-0.32
Specie	-0.15	-0.32

3	Comunità A	Comunità B
H'	0.78	1.61

Relazioni tra i valori di p_i e $Ln(p_i)$:



Massima diversità possibile:

$$H_{\max} = Ln(S)$$

Indice di equiripartizione (o uniformità):

$$E = \frac{H'}{H_{\max}}$$

Passi per calcolare l’indice di equiripartizione di Pielou E:

		Comunità A	Comunità B
1	H'	0.78	1.61
2	H_{max}	1.61	1.61
3	E	0.48	1.00

Indice di Dominanza di Simpson D:

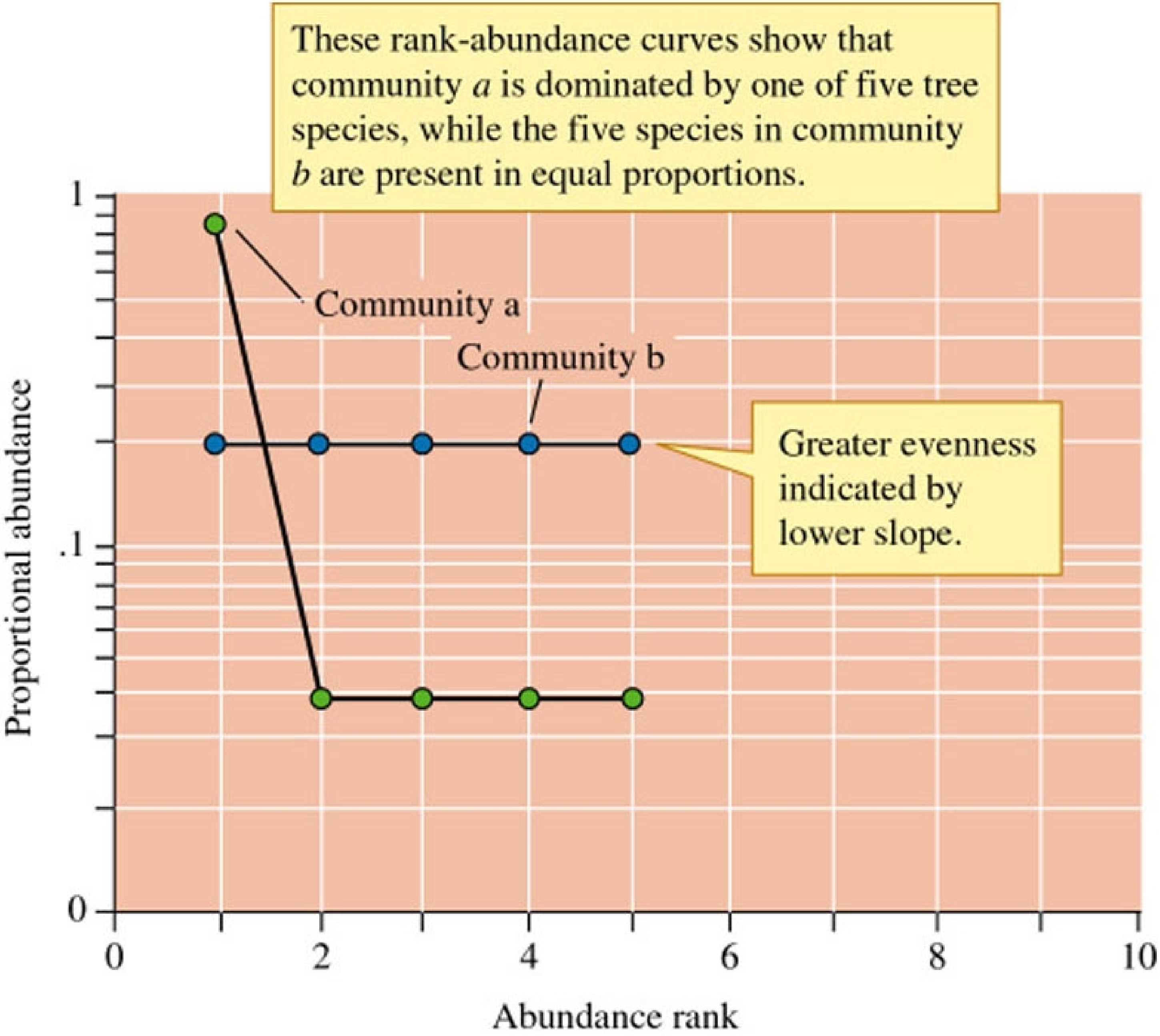
$$D = \sum_{s=1}^{S} p_i^2$$

Abbondanze relative (p_i) delle singole classi nelle due comunità:

	Comunità A	Comunità B
Specie 1	0.80	0.20
Specie 2	0.05	0.20
Specie 3	0.05	0.20
Specie 4	0.05	0.20
Specie 5	0.05	0.20
Totale	1.00	1.00

Passi per calcolare l’indice di dominanza di Simpson D:

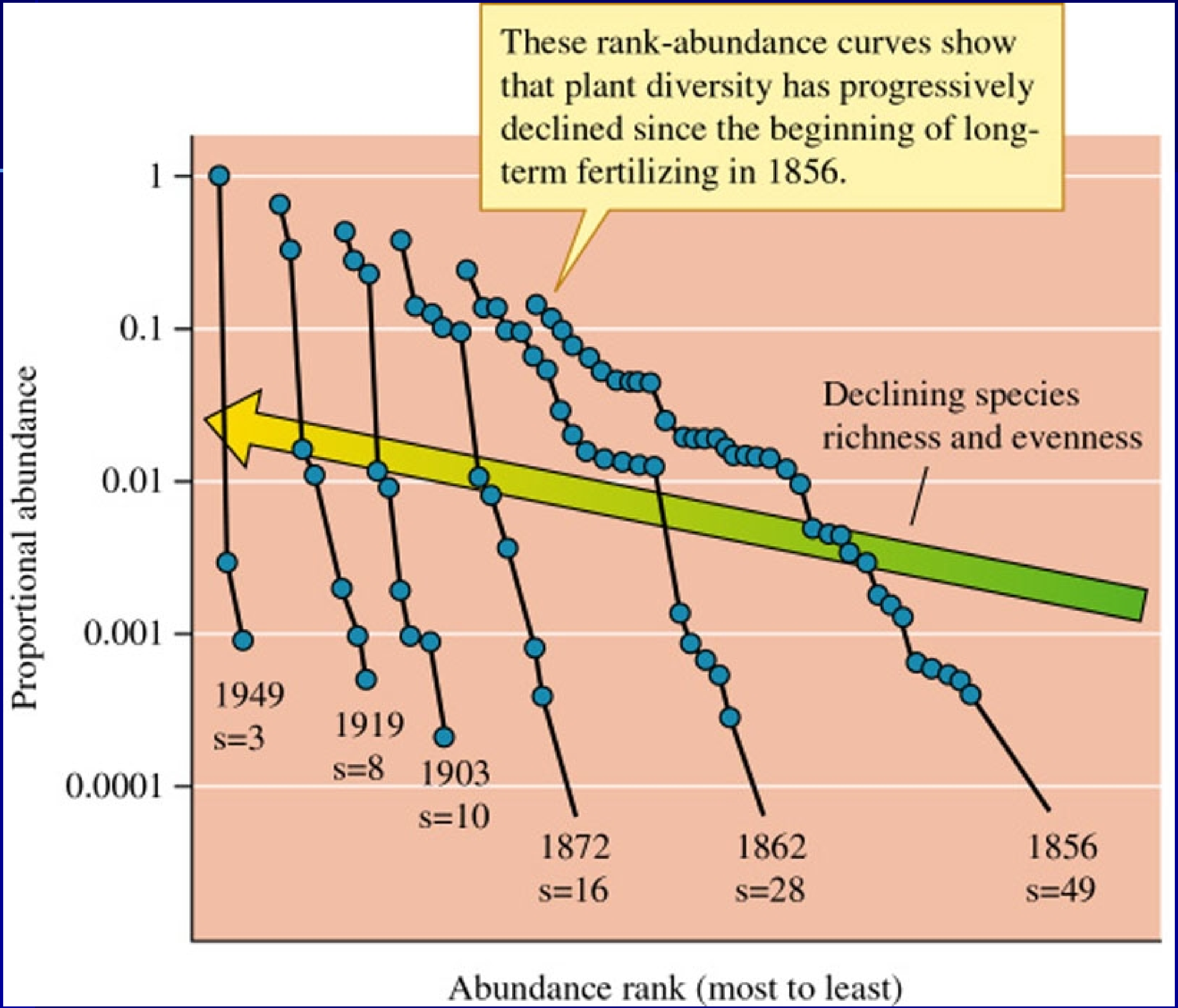
	Comunità A			Comunità B	
	p_i	$(p_i)^2$		p_i	$(p_i)^2$
Specie	0.80	0.640		0.20	0.04
Specie	0.05	0.003		0.20	0.04
Specie	0.05	0.003		0.20	0.04
Specie	0.05	0.003		0.20	0.04
Specie	0.05	0.003		0.20	0.04
	D	**0.650**			**0.200**



These rank-abundance curves show that community *a* is dominated by one of five tree species, while the five species in community *b* are present in equal proportions.

Greater evenness indicated by lower slope.

THE PARADOX OF ENRICHMENT

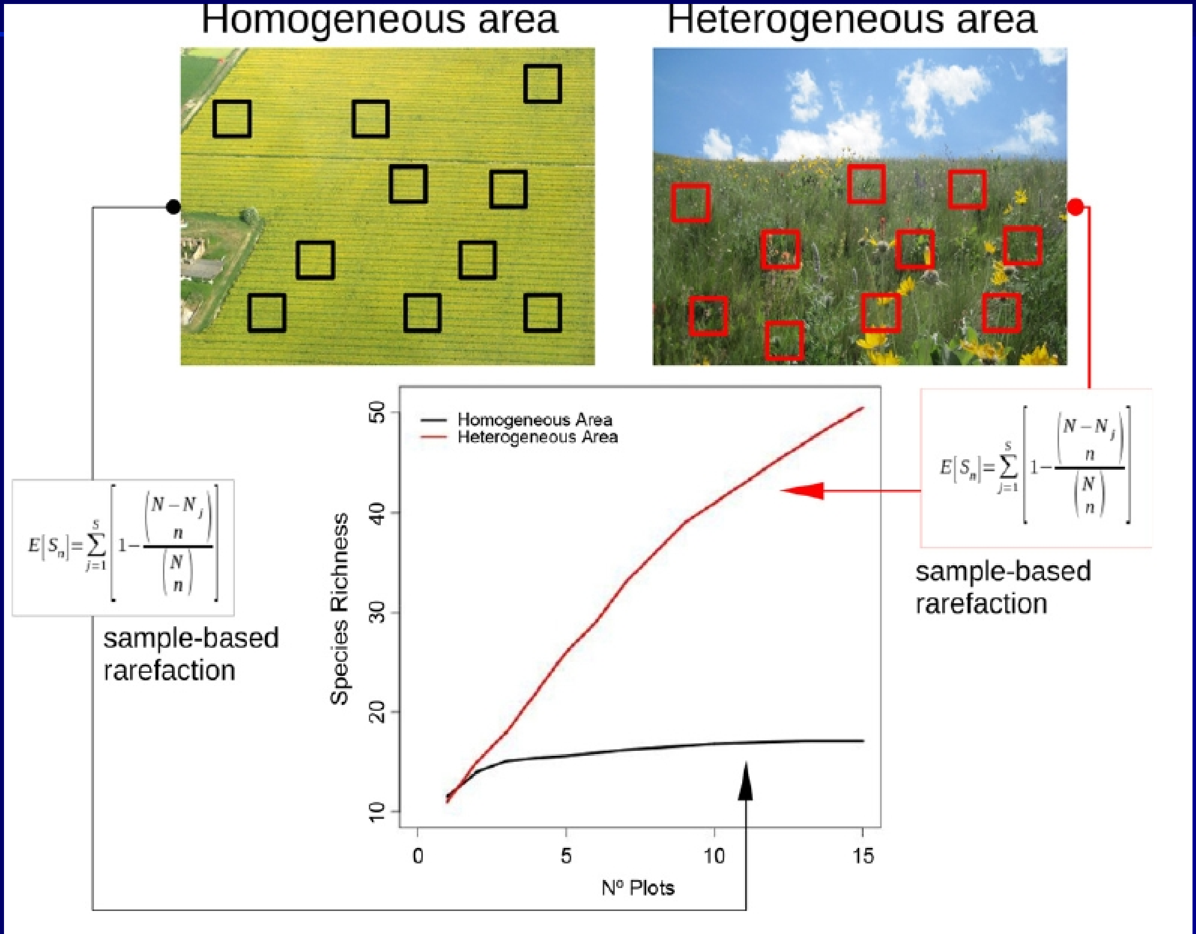


These rank-abundance curves show that plant diversity has progressively declined since the beginning of long-term fertilizing in 1856.

Declining species richness and evenness

β-diversity

Defined as the variability in community structure and composition among sites (Whittaker 1960, 1972; Legendre et al. 2005).

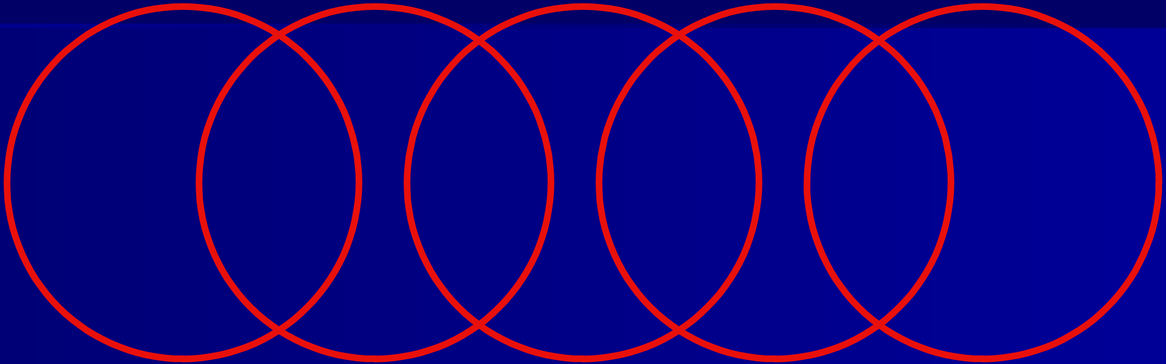


- Classically based only on species identities (similarity/dissimilarity index as Jaccard).

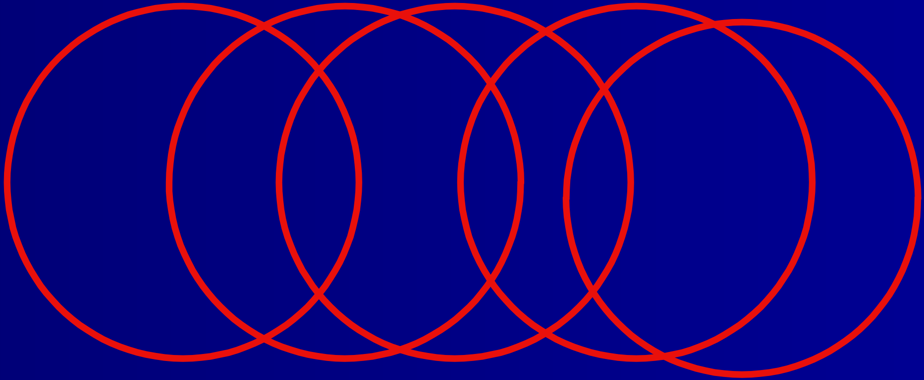
β-diversity

	A_1	A_2	A_3	.	.	A_n
S_1	X11	X12	X13	.	.	X1n
S_2	X21	X22	X23	.	.	X2n
S_3	X31	X32	X33	.	.	X3n
.	.	.	.	.	.	.
.	.	.	.	.	.	.
S_m	Xm1	Xm2	Xm3	.	.	Xmn

High β-Diversity



Low β-Diversity



Using compositional (dis)similarity to quantify β diversity

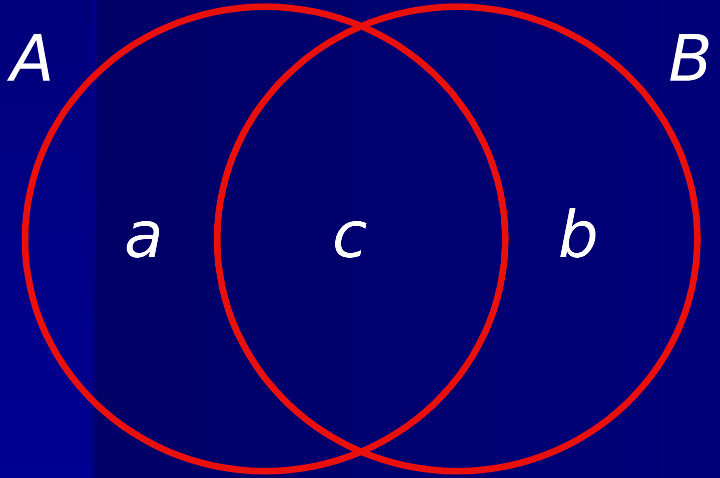
There are many coefficients that measure (dis)similarity in species composition between pairs of sites or plots (Legendre & Legendre 1998), and any one of these coefficients may be used as the starting point for quantifying *β*-diversity (Pielou 1975; Magurran 1988).

(dis)similarity in species composition is measured for each *pair* of plots separately, rather than for the whole set of plots simultaneously.

Jaccard Index of similarity

These indices can easily be remembered by using the set theory! (where a set is the list of species in one community or site)

$$J = \frac{c}{(a+b+c)}$$



a = Number of species only present in community (site) A

b = Number of species only present in community (site) B

c = Number of species present in both communities (sites) A and B

Sørensen Index of similarity

For a given pair of plots, the Sørensen coefficient expresses compositional similarity in a straightforward, easily interpretable way, and is calculated as follows:

$$S_s = c / \alpha$$

Where:

c is the number of species shared by both plots and *α* is average number of species in each plot

Values of S_s range from 0, indicating no species in common, to 1, indicating that every species found in one plot is also found in the other plot. Subtracting S_s from 1 yields a coefficient of dissimilarity (e.g. Colwell & Coddington 1994; Legendre & Legendre 1998).

Visualising differences: (dis)similarity matrix

A dissimilarity matrix is a matrix which reports the dissimilarity values (usually expressed as 1- similarity) between each pair of plots (communities) according to the selected index (both qualitative or quantitative).

The dissimilarity matrix below is calculated by using the following dissimilarity Index: 1 - Jaccard similarity index on the previous similarity matrix

	Sample / Communities				
	A	**B**	**C**	**D**	**E**
A	0	0.66	1	0.75	0.75
B		0	0.66	0.75	0.75
C			0	0.57	0.75
D				0	0.62
E					0

(dis)similarity matrix

A similarity matrix is a matrix which reports the similarity value between each pair of plots (communities) according to the selected similarity index.

Sample / Communities

	A	**B**	**C**	**D**	**E**
Species 1	1			1	
Species 2	1				1
Species 3	1	1			
Species 4	1	1		1	1
Species 5		1	1		
Species 6		1	1	1	1
Species 7			1	1	
Species 8			1	1	1
Species 9				1	1
Species 10					1

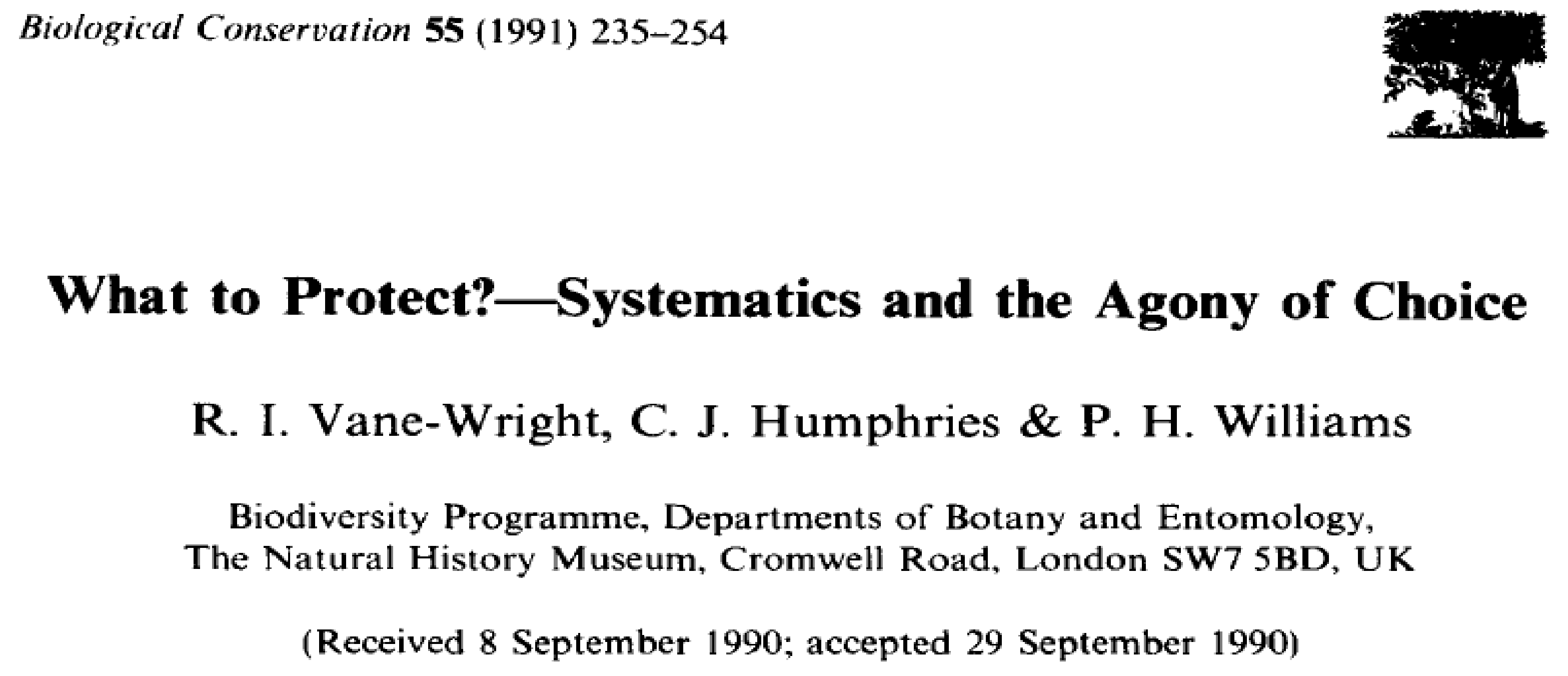
The matrix below is calculated by using the Jaccard similarity index

Sample / Communities

	A	**B**	**C**	**D**	**E**
A	1	0.33	0	0.25	0.25
B		1	0.33	0.25	0.25
C			1	0.43	0.25
D				1	0.38
E					1

More than species richness

- N° of species represents a good surrogate for biodiversity assessment.
- However, sometimes, it is not enough in order to select areas for nature conservation.....



Biological Conservation **55** (1991) 235–254

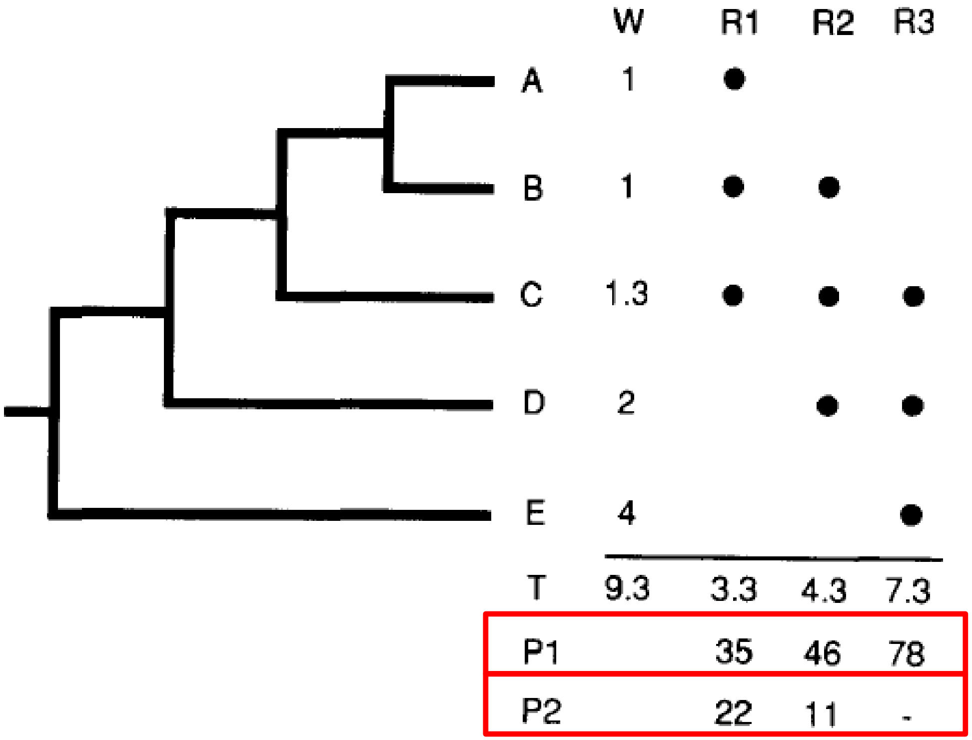
What to Protect?—Systematics and the Agony of Choice

R. I. Vane-Wright, C. J. Humphries & P. H. Williams

Biodiversity Programme, Departments of Botany and Entomology,
The Natural History Museum, Cromwell Road, London SW7 5BD, UK

(Received 8 September 1990; accepted 29 September 1990)

For instance....



	W	R1	R2	R3
A	1	●		
B	1	●	●	
C	1.3	●	●	●
D	2		●	●
E	4			●
T	9.3	3.3	4.3	7.3
P1		35	46	78
P2		22	11	-

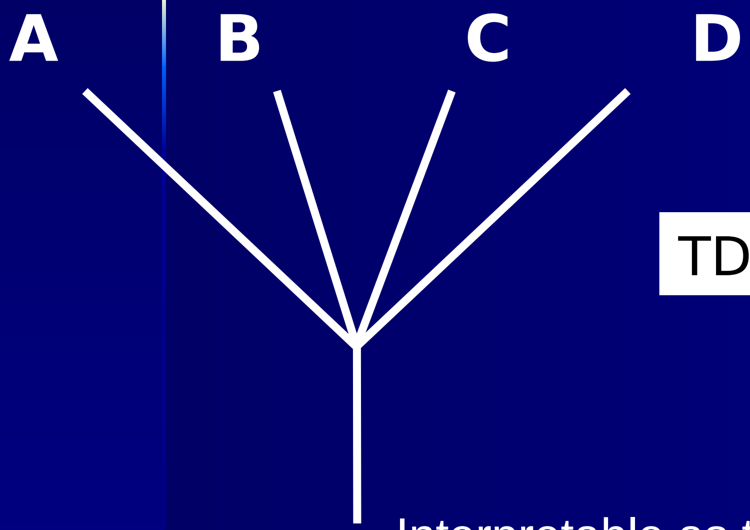
Fig. 5. Theoretical priority area analysis, based on topology and taxic weighting given in Fig. 2. Three of the five terminal taxa occur in each of three areas, R1–R3, according to the three-column matrix at the right. Column W gives the taxic weights (cf. Fig. 2). Row T gives the total (aggregate) scores for all five taxa, and for each of the three regions; row P1 gives the percentage diversity scores for each of the three regions at the first step, indicating that R3 is the top-priority region; row P2 gives the percentage diversity scores for the remaining two regions with respect to the taxa *complementary* to those occurring in R3, and indicates that R1 is the second priority; finally, the fact that the highest values of the first two steps sum to 100% indicates that the analysis is complete (Pressey & Nicholls' site efficiency value: 0·33).

Once you select the first

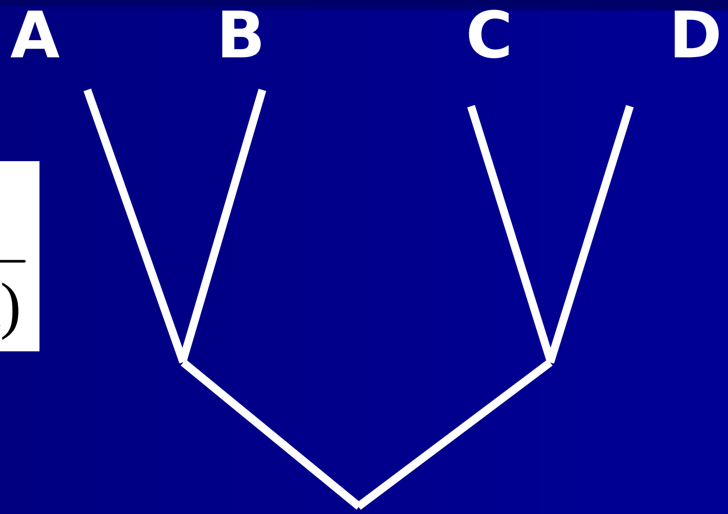
...you will use the complementarity principle to select the second!

Once you have constructed your (Linnean or Phylogenetic) tree…

as usual you can calculate the mean distance



$$\text{TD} = \frac{\sum \Delta_{i,j}}{N(N-1)}$$



Interpretable as the mean taxonomic distance between any two species in the considered taxonomc tree

TD = 2.00

TD = 3.33