University of Glasgow

A general framework for measuring diversity

Tom Leinster and Christina Cobbold

University of Glasgow

Diversity measurement spans disciplines

- Ecology (degree of variation of life)
- Population genetics (expected heterozygosity)
- Information theory (amount of uncertainty in a message)
- Physics (entropy in thermodynamics)
- Economics (amount of income concentrated in highest earners)

University of Glasgow

Outline

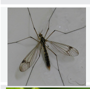
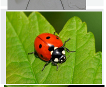
- Part 1: Measuring diversity, ignoring species similarity
- Part 2: Measuring diversity, incorporating species similarity

Outline

- Part 1: Measuring diversity, ignoring species similarity
- Part 2: Measuring diversity, incorporating species similarity

Quantifying diversity

- Richness: How many species?

	3	100	1000
	3	10	1
	3	1	1
	3	1	1

- All sites have 4 species – how do we distinguish them?
- Ans: Include relative contribution of a species to the community
- Diversity is a relationship between richness and abundance
- Key ideas: richness, eveness and dominance

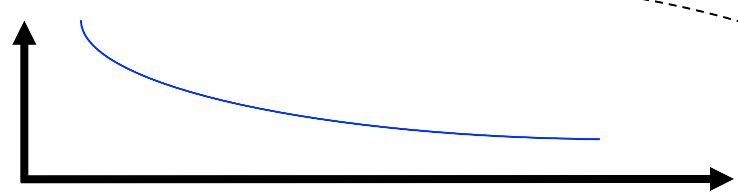
Quantifying diversity

Relative abundance data, **p**

model of community

Formula

measure of diversity



Quantifying diversity

Relative abundance data, **p**

$$\mathbf{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$$

p_i = relative frequency, or relative abundance, of the ith species

$$p_i \geq 0, \quad \sum p_i = 1$$

University of Glasgow

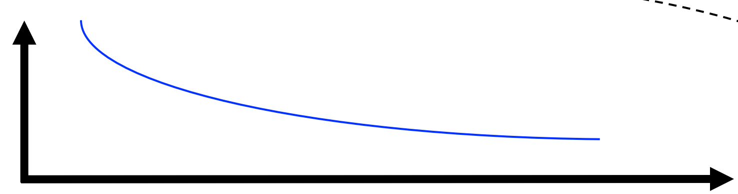
Quantifying diversity

Relative abundance data, **p**

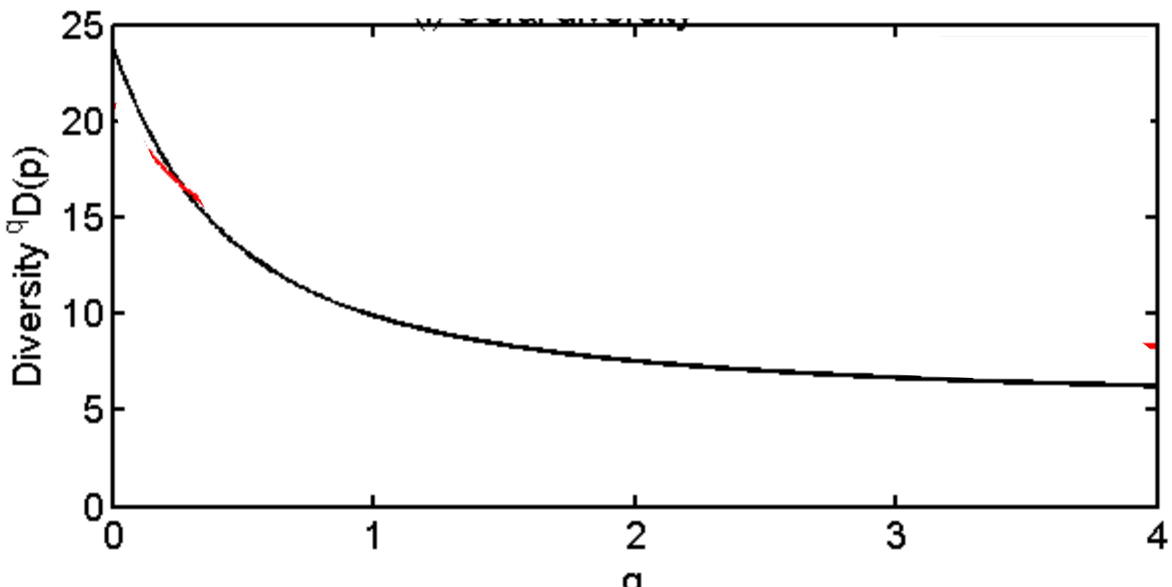
model of community

Formula (Hill numbers)

measure of diversity



Diversity profile



'view point parameter'

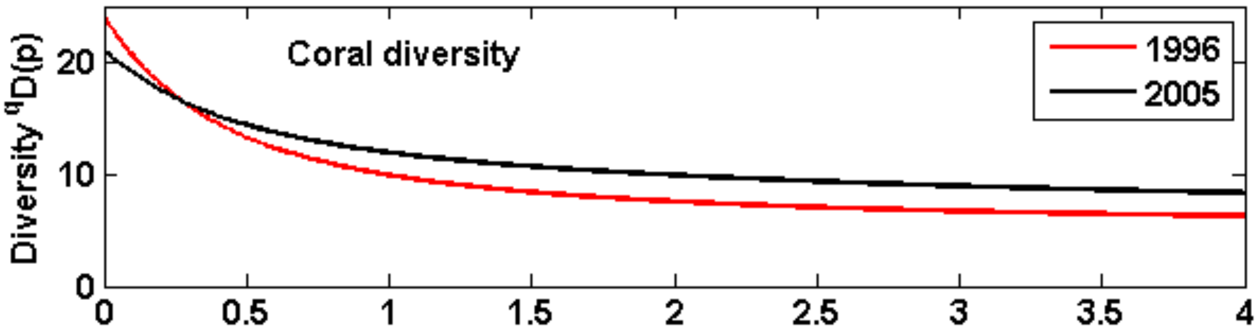
q=0, species richness
rare species are important

q=infinity, 1/Berger-Parker
rare species are unimportant

'Importance of evenness'

Comparing communities using diversity profiles

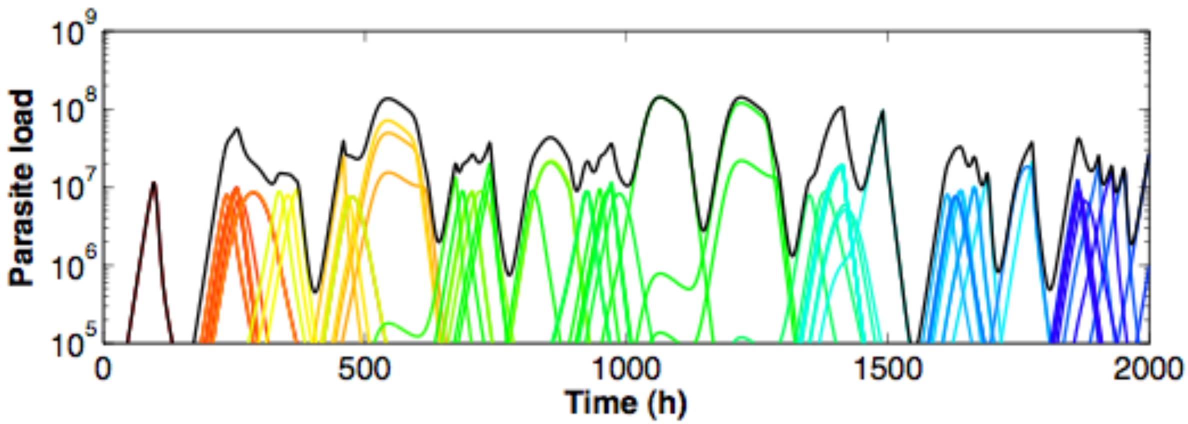
1996 community was more rich in species, but less even

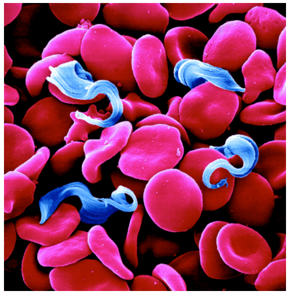


Understorey is more diverse, unless we care only about dominance.

Sensitivity parameter q

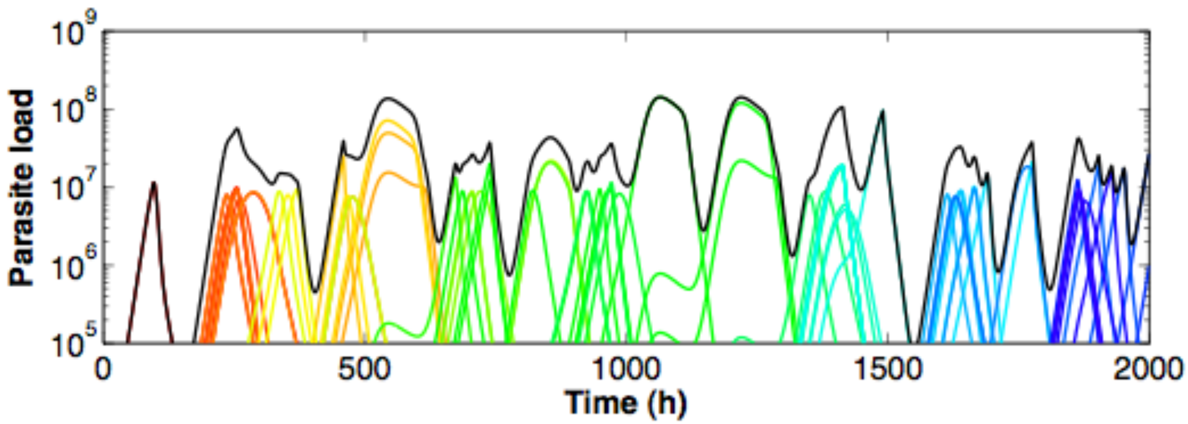
Trypanosome infections and the role of antigenic diversity?

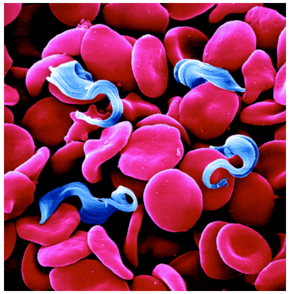


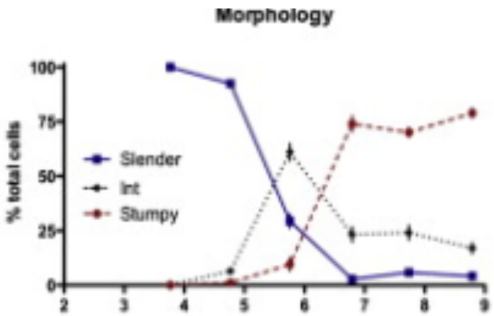


How is antigenic diversity expressed and how does it power trypanosnome persistence and survival?

Trypansome infections and the role of antigenic diversity?







University of Glasgow

Outline

- Part 1: Measuring diversity, ignoring species similarity
- Part 2: Measuring diversity, incorporating species similarity

What are some of the limitations of Hill numbers

- "…associated with the idea of diversity is the concept of 'distance' i.e. some measure of the dissimilarity of the resources in question"

OECD Handbook on biodiversity valuation: A guide for policy makers

- "A community of ten species of barnacle should clearly be less diverse than a community of ten very different species"

The essence of a point made by E.C. Pielou, Ecological Diversity

- There is not always a good notion of what constitutes a species e.g. microbial communities

Quantifying diversity

Similarity data, **Z**   Abundance data, **p**

model of community

n × n matrix (n = number of species)

Z_{ij} = similarity between ith and jth species

totally dissimilar → $0 \le Z_{i,j} \le 1$ ← *identical*

Example

This model assumes that distinct species are totally dissimilar

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

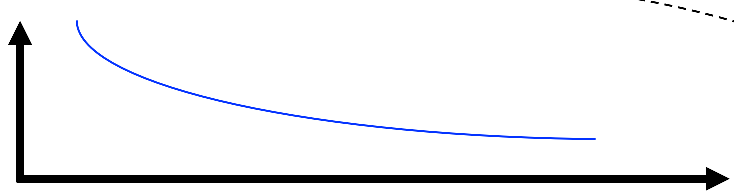
Quantifying diversity

Similarity data, **Z**   Abundance data, **p**

model of community

$$ {}^{q}D^{\mathbf{Z}}(\mathbf{p}) = \left(\sum_{i:p_i>0} p_i (\mathbf{Zp})_i^{q-1} \right)^{\frac{1}{1-q}} $$

measure of diversity



Example: Butterfly data in the Ecuadorian rainforest

Abundance data of Charaxinae Butterflies at a rainforest site in Ecuador

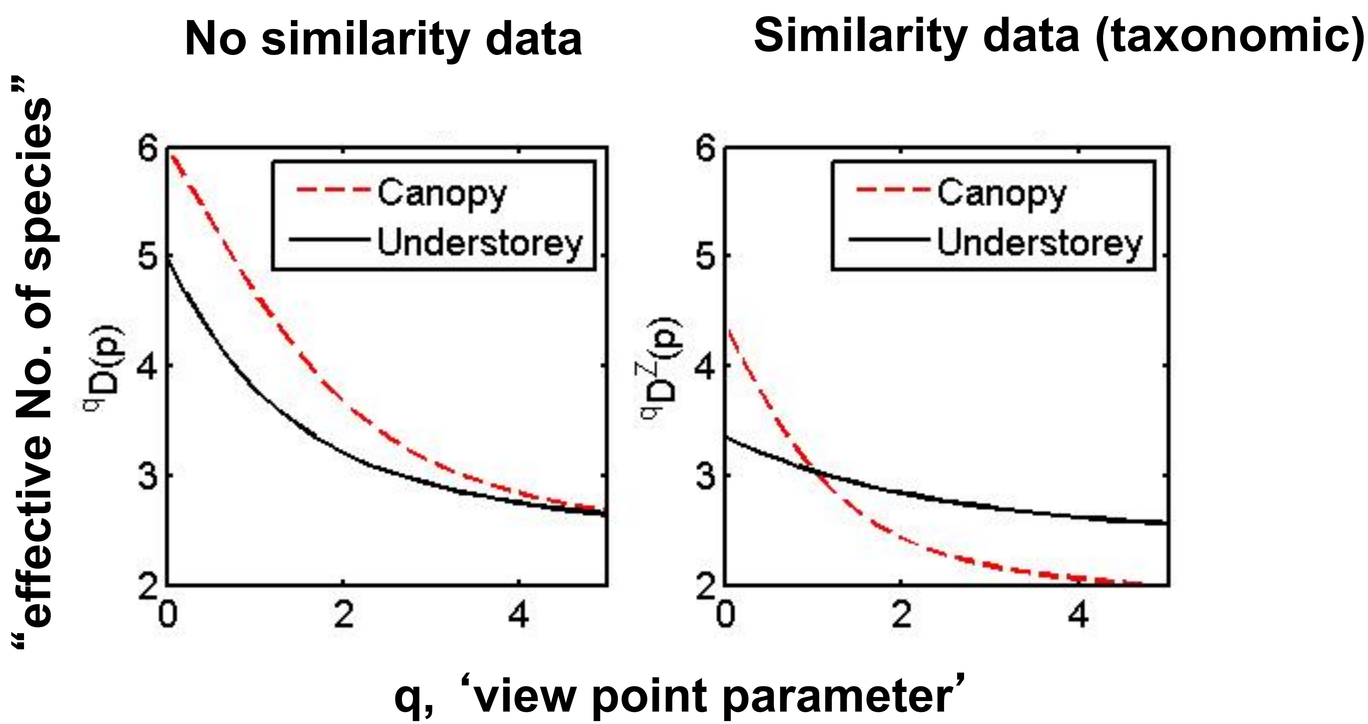
Species	Canopy	Understorey
Prepona laertes	15	0
Archaeoprepona demophon	14	37
Zaretis itys	25	11
Memphis arachne	89	23
Memphis offa	21	3
Memphis xenocles	32	8

Taxonomic similarity matrix

$$Z_{ij} = \begin{cases} 0 & \text{if of different genera} \\ 0.5 & \text{if different but congeneric} \\ 1 & \text{if i = j,} \end{cases}$$

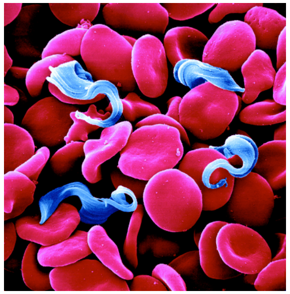
Data from DeVries et al (1997)

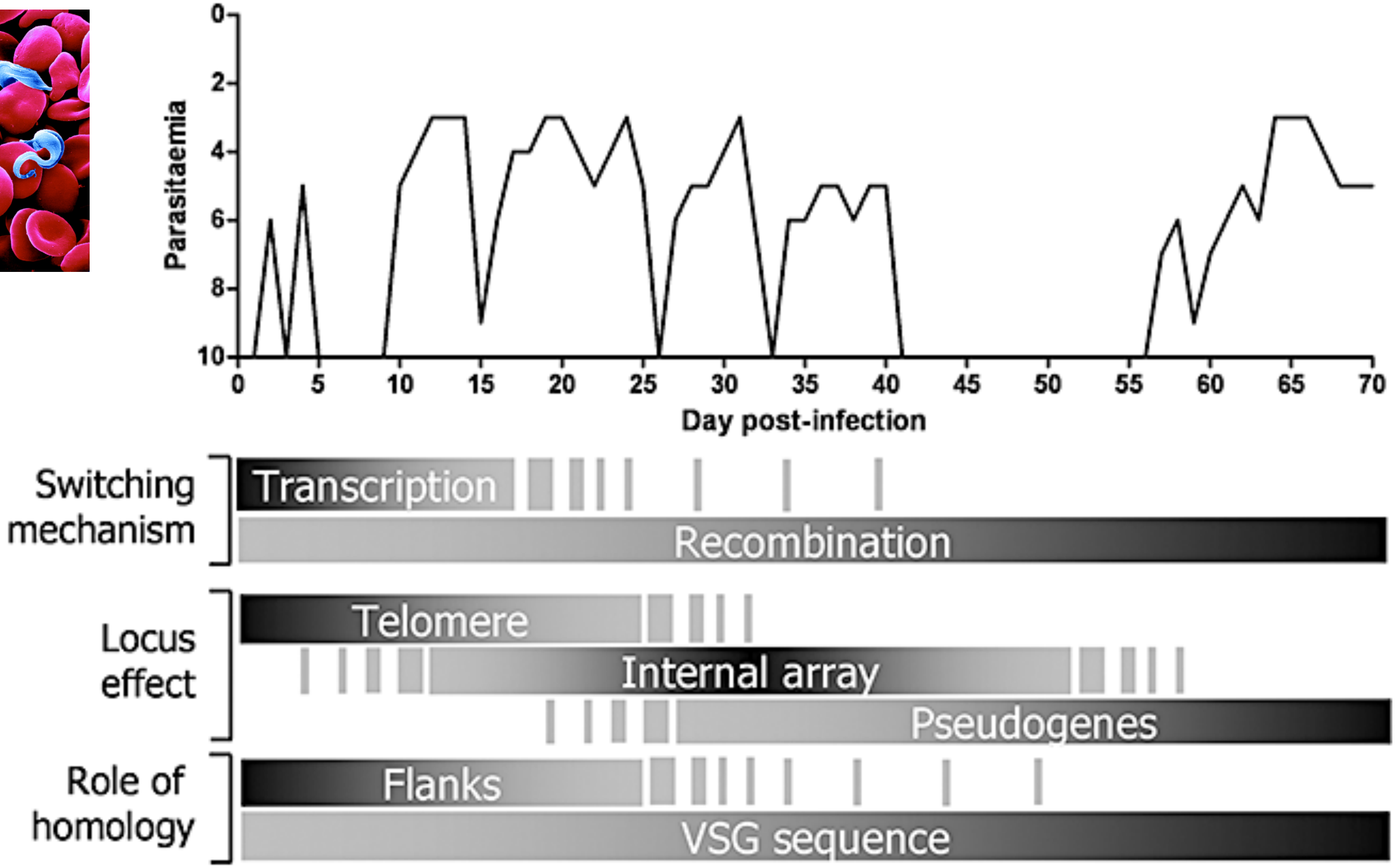
Butterfly data in the Ecuadorian rainforest



Data from DeVries et al (1997)

Trypanosome infections and similarity





University of Glasgow

Microbe communities: notion of a species is problematic

- Expected similarity between a randomly chosen pair of individuals:

$$1/{}^2D^Z(p) = \mu_2$$

- Now consider q>2 (whole number). Given q individuals of species $i_1, i_2 \ldots, i_q$ *then a measure of group similarity is:*

$$Z_{i_1 i_2} Z_{i_1 i_3} \cdots Z_{i_1 i_q}$$

- *Let* μ_q *be the expected similarity of a randomly chosen group of q individuals (sampled with replacement). Then*

$${}^qD^Z(p) = \mu_q^{1/(1-q)}$$

University of Glasgow

Answering the common criticisms

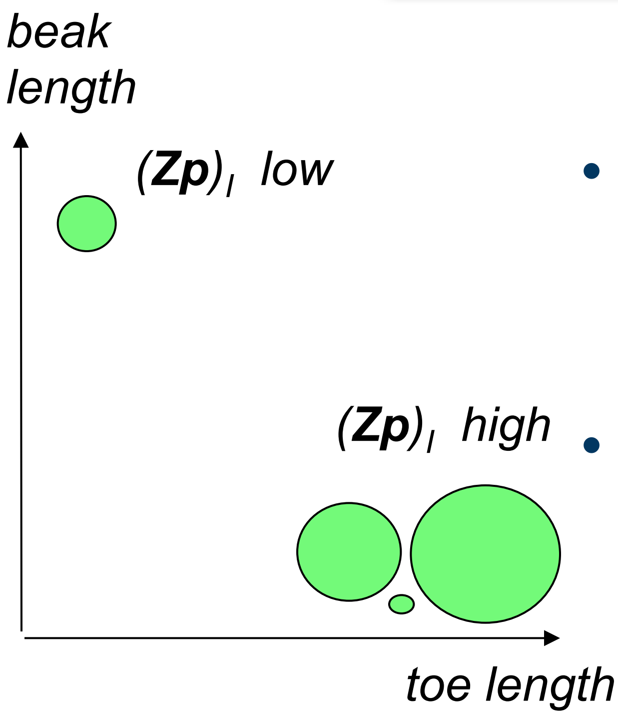
- Diversity can mean too many different things
  - We can separate meanings by choosing different Z
- Too many diversity measures have been proposed
  - Many of them are unified under the umbrella of ${}^{q}D^{Z}$
- Diversity measures produce meaningless numbers
  - Effective numbers produce meaningful numbers
- A single number carries little information
  - Draw a diversity profile
- Diversity measures are too dependent on the notion of species
  - These behave proportionately when species boundaries are changed
- The varying differences between species are ignored
  - Not here

Acknowledgements to….

Tom Leinster
School of Mathematics
and Statistics
University of Edinburgh

University of Glasgow

Similarity-sensitive diversity measures



- The ordinariness of the *i*-th species is

$$(Zp)_i = \sum_{j=1}^{S} Z_{ij} p_j$$

- Average ordinariness of an individual within the community

$$\sum_{i=1}^{S} p_i (Zp)_i$$

which measures lack of diversity. So one measure of diversity is

$$1 \Big/ \sum_{i=1}^{S} p_i (Zp)_i$$