From species coocurrences...

...to species interactions

Nerdchat (28/11/2016)

Part 1.
A historical perspective

Part 1.
A historical perspective

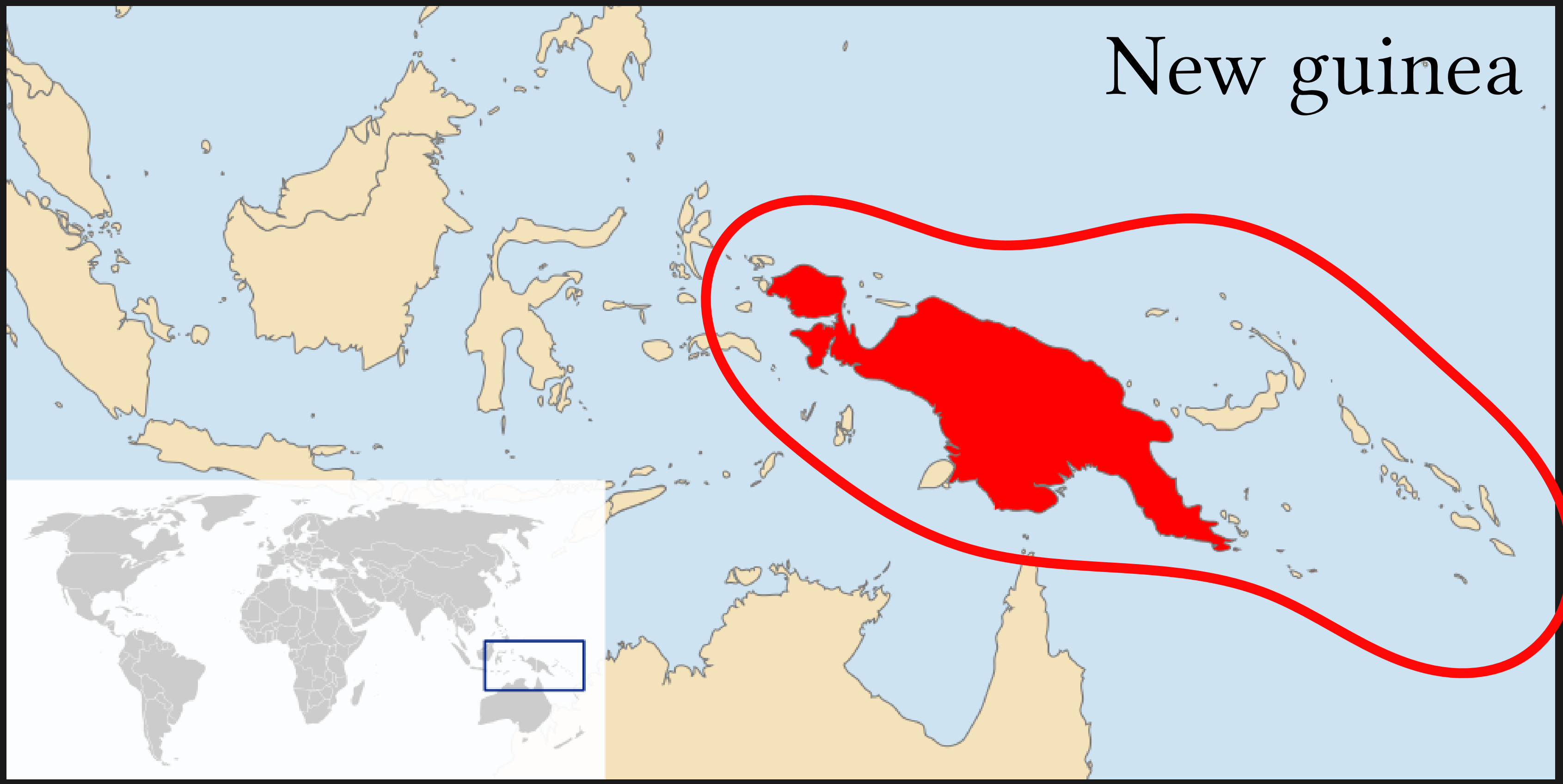
Part 2.
A "modern" permutation-based null model

Part 1.
A historical perspective

Part 2.
A "modern" permutation-based null model

Part 3.
Geographical scale and coocurrence patterns

jareddiamond.org



A Fruit Dove (Ptilinopus sp.)

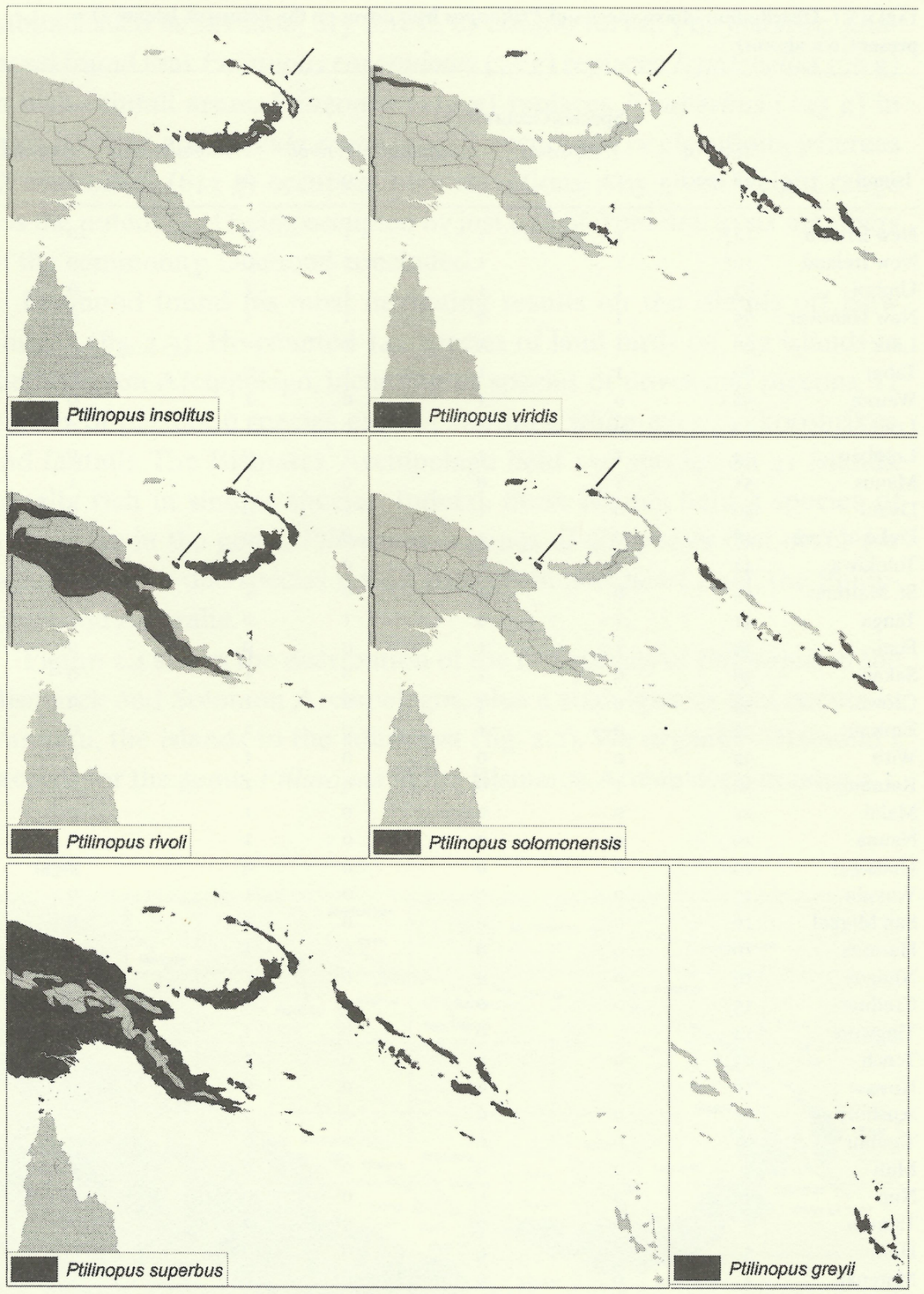


FIG. 2.4. The distribution of 6 species of *Ptilinopus* fruit dove. *Ptilinopus insolitus* and *P. viri-*

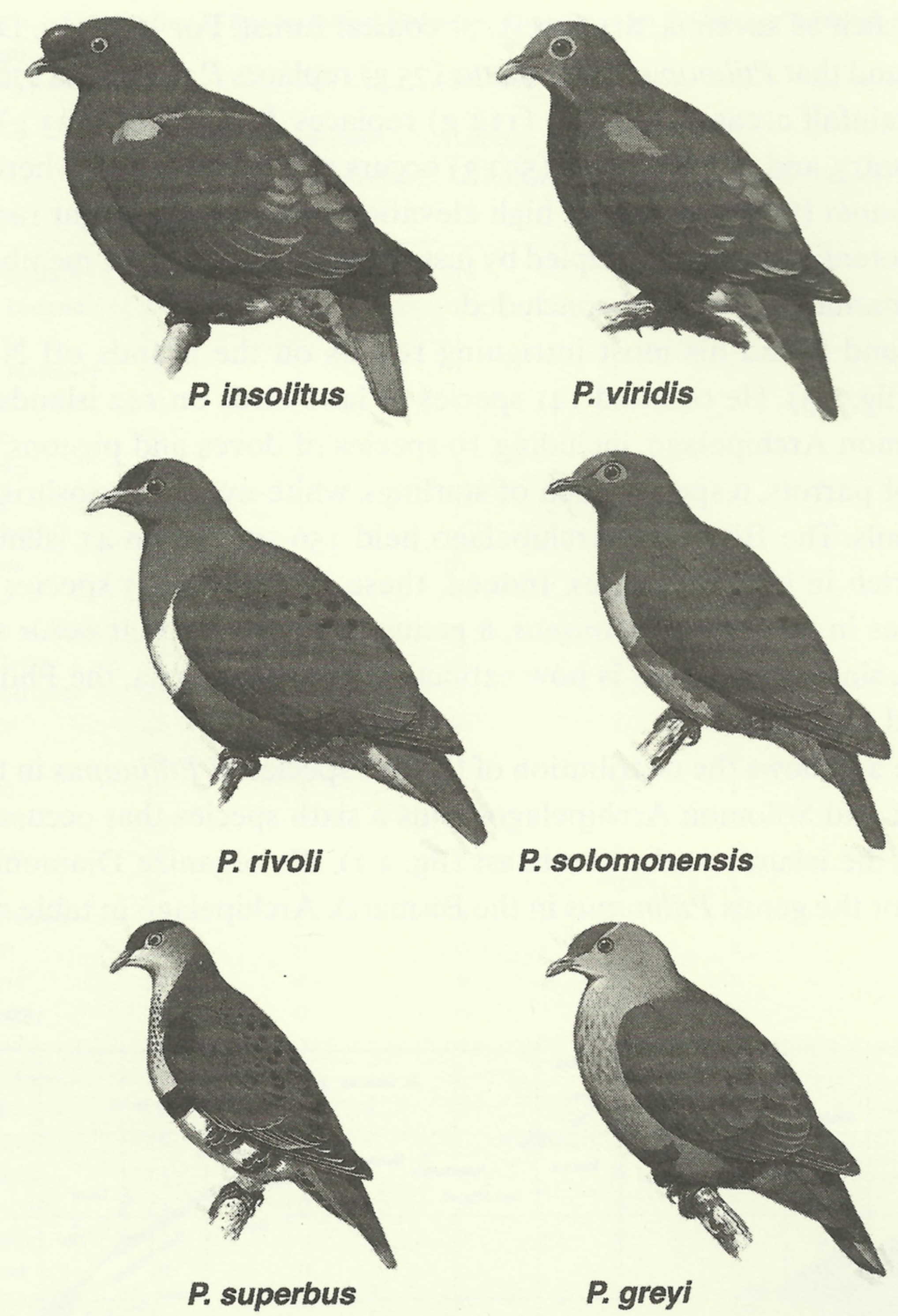


FIG. 2.2. The species of *Ptilinopus* fruit dove.

Checkerboards

Checkerboards

Evidence of competition shaping communities at large-scale ?

« In a few instances, competition expresses itself in 'simple' checkerboards distributions, by which species replace each other one-for-one »

Diamond, 1975 (In Cody and Diamond, 1975)

Assembly rules

In a widely cited new approach to the interpretation of biogeographic distributions, Diamond (1975) asserts that the assembly of bird communities manifests the following patterns:

a. "If one considers all the combinations that can be formed from a group of related species, only certain ones of these combinations exist in nature."

b. "Permissible combinations resist invaders that would transform them into forbidden combinations."

c. "A combination that is stable on a large or species-rich island may be unstable on a small or species-poor island."

d. "On a small or species-poor island, a combination may resist invaders that would be incorporated on a larger or more species-rich island."

e. "Some pairs of species never coexist, either by themselves or as a part of a larger combination."

f. "Some pairs of species that form an unstable combination by themselves may form part of a stable larger combination."

g. "Conversely, some combinations that are composed entirely of stable subcombinations are themselves unstable."

« We show that every assembly rule is either tautological, trivial, or a pattern expected were species distributed at random »
Connor and Simberloff, 1979

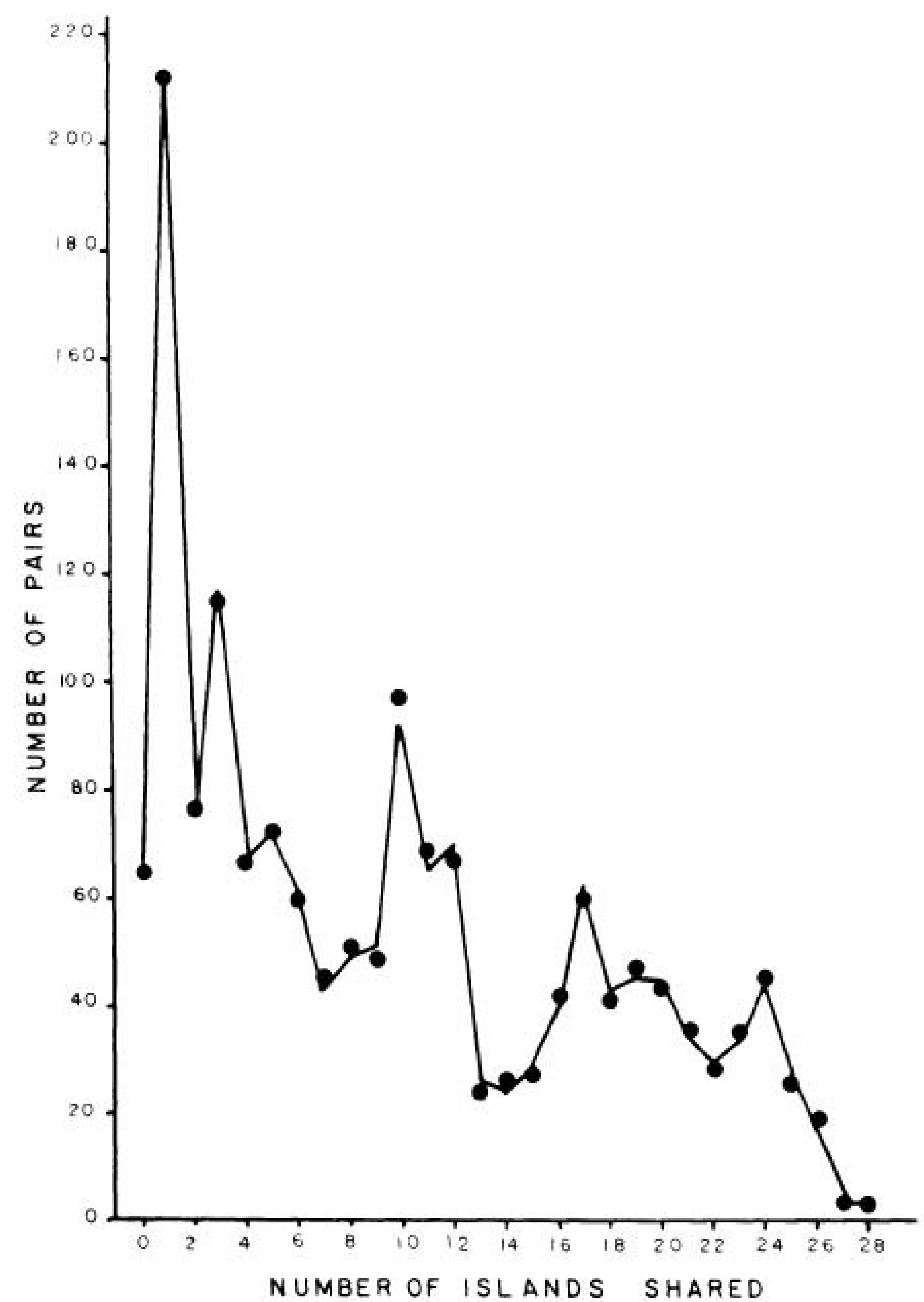


FIG. 1. Distribution of number of species pairs vs. number of islands shared for New Hebrides birds. Solid line is the expected distribution given the three constraints discussed in text. Dots are the observed values.

New Hebrides (Vanuatu) birds :

58 species x 28 islands

« We show that every assembly rule is either tautological, trivial, or a pattern expected were species distributed at random »
Connor and Simberloff, 1979

« Diamond (1975) assumed competition to be the primary determinant and then sought post facto to rationalize the observed data in the light of this assumption »
Connor and Simberloff, 1979

New Hebrides (Vanuatu) birds :

58 species x 28 islands

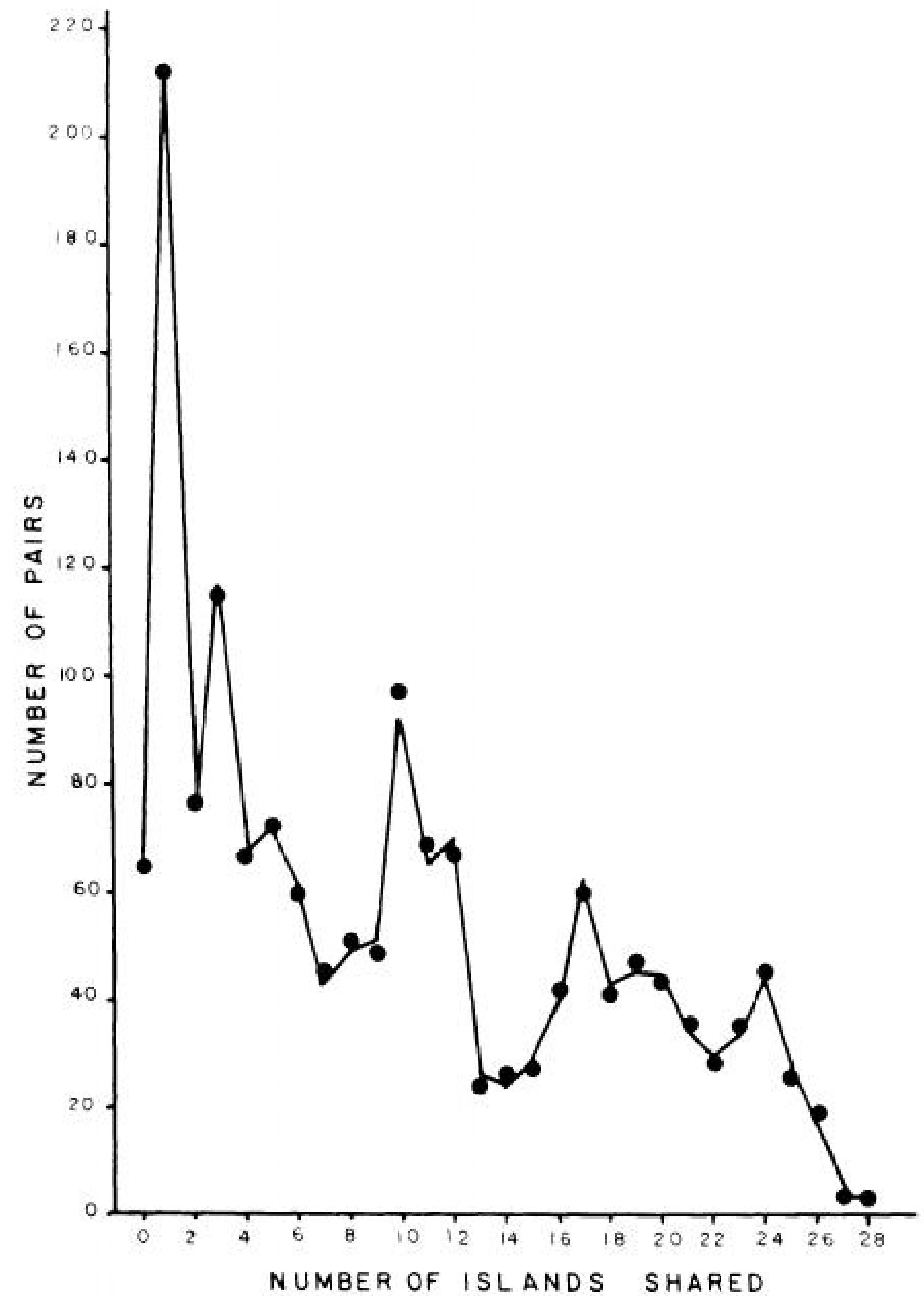


FIG. 1. Distribution of number of species pairs vs. number of islands shared for New Hebrides birds. Solid line is the expected distribution given the three constraints discussed in text. Dots are the observed values.

Connor and Simberloff's null model

A incidence matrix :

		Sites						
Species	0	0	0	0	0	1	0	1
	0	0	1	0	1	1	0	3
	1	0	1	1	1	0	0	4
	0	1	1	0	1	0	0	3
	1	1	3	1	3	2	0	

Connor and Simberloff's null model

A incidence matrix :

Sites

Species								
	0	0	0	0	0	1	0	1
	0	0	1	0	1	1	0	3
	1	0	1	1	1	0	0	4
	0	1	1	0	1	0	0	3
	1	1	3	1	3	2	0	

Islands can inherently be "bigger" (hold more species)

> the column sum is constrained

Species have an inherent dispersal ability

> the row sum is constrained

Species occur in a given "community size"

> if a species occur in a community with S species, then it can only occur with S species in the null model

Connor and Simberloff's null model

Sites

Species

$$\begin{array}{ccccccc|c} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 3 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 4 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 3 \\ 1 & 1 & 3 & 1 & 3 & 2 & 0 & \end{array}$$

$$\begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & \end{array} \longrightarrow \begin{array}{ccc} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & \end{array}$$

a Swap algorithm

Connor and Simberloff's null model

"several swaps", nobody knows (?)

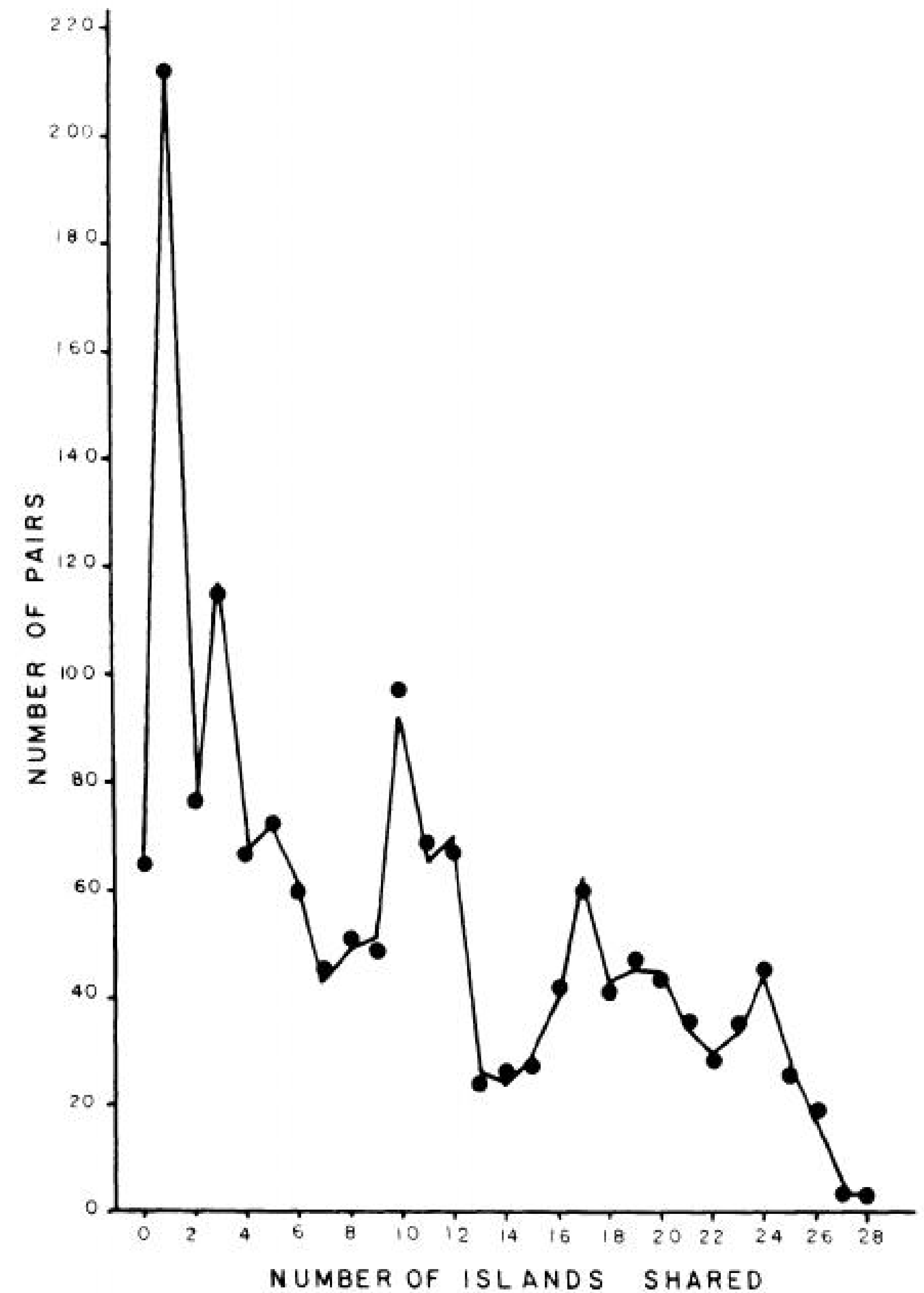


FIG. 1. Distribution of number of species pairs vs. number of islands shared for New Hebrides birds. Solid line is the expected distribution given the three constraints discussed in text. Dots are the observed values.

A "good" permutation-based null model

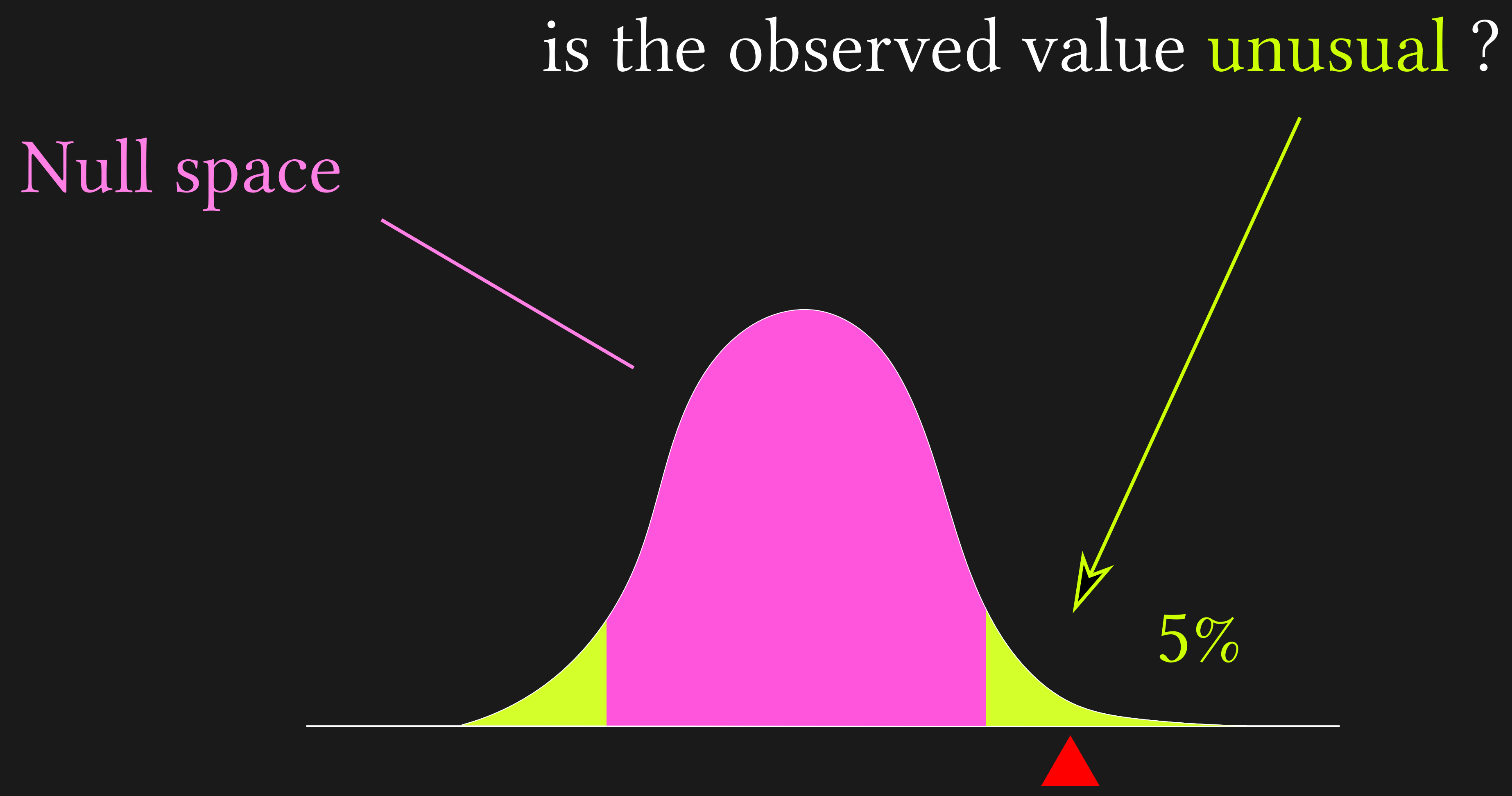
Apparent mutual exclusion, measured by the number of co-ocurrences

A pattern, measured by a metric

A "good" permutation-based null model

Apparent mutual exclusion, measured by the number of co-ocurrences

A pattern, measured by a metric



* The number of co-occurrences

S_{ij} — Species-pair level

$\overline{S_{ij}}$ — Matrix level

Metrics

* The number of co-occurrences

$$S_{ij} \qquad \overline{S_{ij}}$$

Species-pair level

Matrix level

* The checkerboard score

$$C_{ij} = (r_i - S_{ij})(r_j - S_{ij})$$

$$C_{score} = \frac{2}{n*(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} C_{ij}$$

* Togetherness, etc...

Generating the null space

Constraints

Row sums

Column sums

(Incidence) — The Narcissus effect
Colwell & Winkler 1984

↓

Many, many, many matrices

Sites

Species								
	0	0	0	0	0	1	0	1
	0	0	1	0	1	1	0	3
	1	0	1	1	1	0	0	4
	0	1	1	0	1	0	0	3
	1	1	3	1	3	2	0	

Generating the null space

Constraints

Row sums

Column sums

(Incidence) — The Narcissus effect
Colwell & Winkler 1984

↓

Many, many, many matrices

Sites

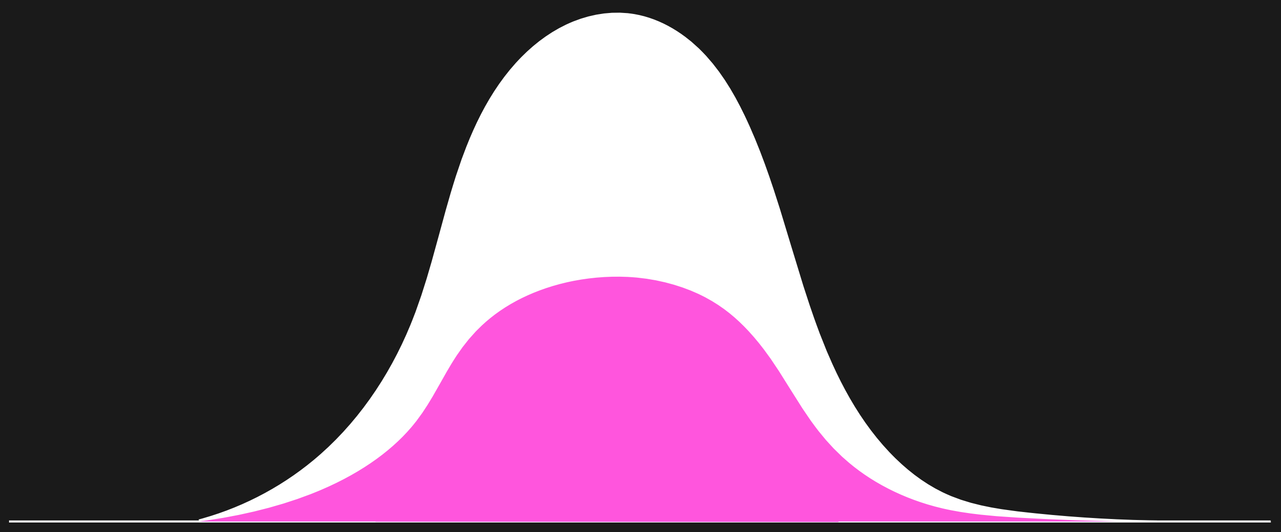
Species

0	0	0	0	0	1	0	1
0	0	1	0	1	1	0	3
1	0	1	1	1	0	0	4
0	1	1	0	1	0	0	3
1	1	3	1	3	2	0	

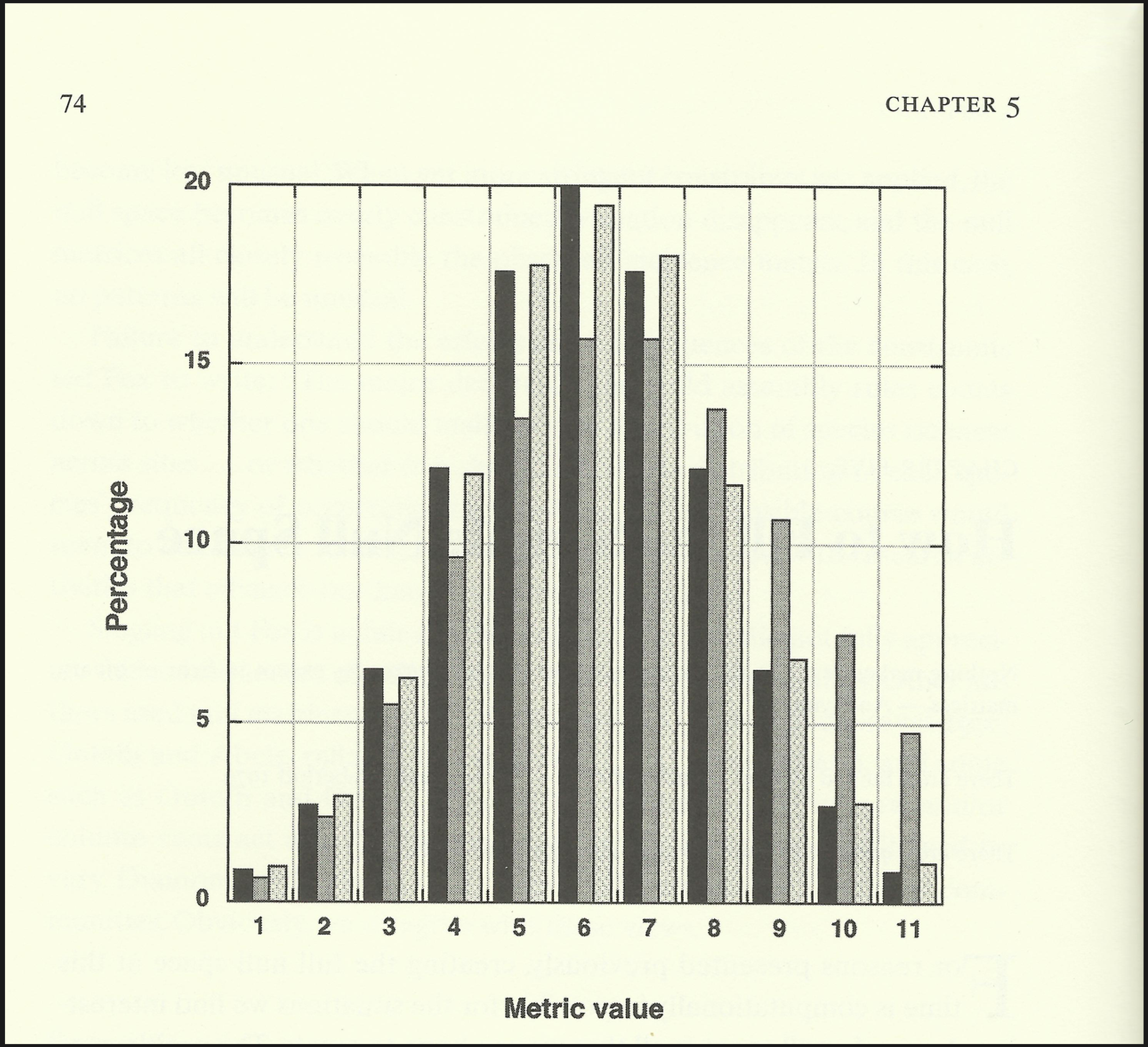
Very often, the full null space is impossible to compute !

↓

We compute a sample null space



Generating the null space



We compute a sample null space

Generating the null space

Enumeration algorithms — build all the possible matrices — expensive

Swap algorithms — start from a matrix, then swap submatrices — cheap

the trial-swap method (Miklos & Podani, 2004)

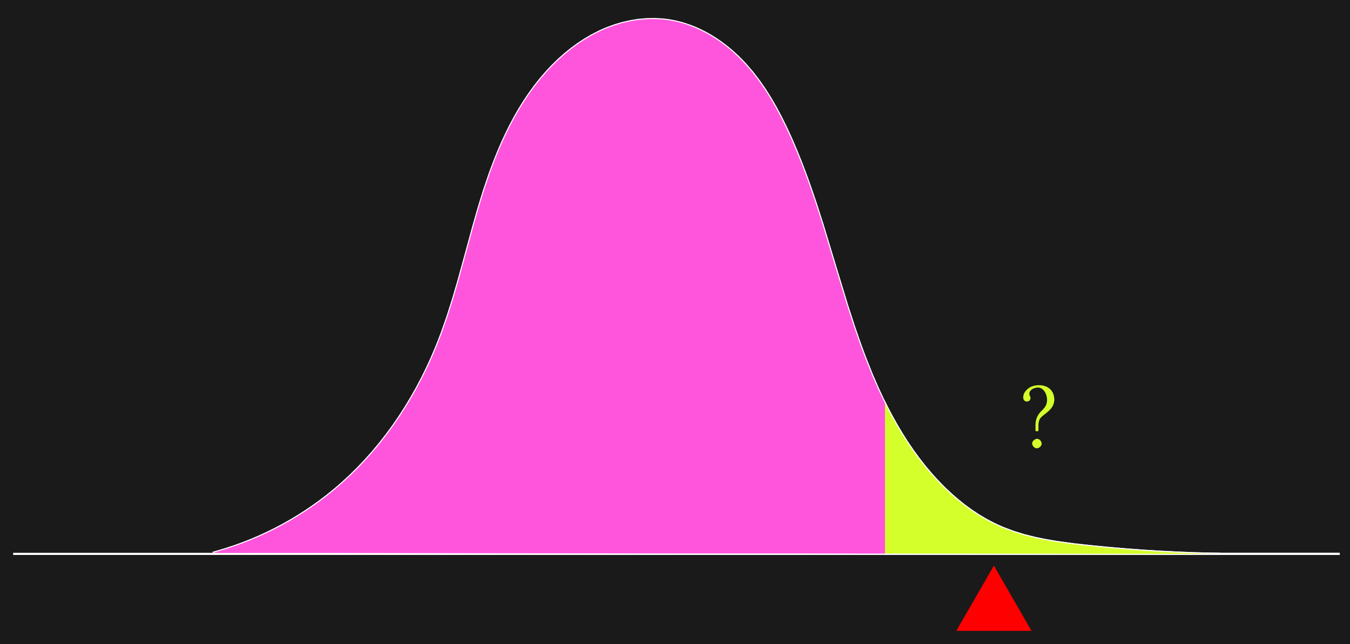
Construction algorithms — start from zeroes/random, then build a matrix satisfying the constraints — expensive

We compute a sample null space

Testing

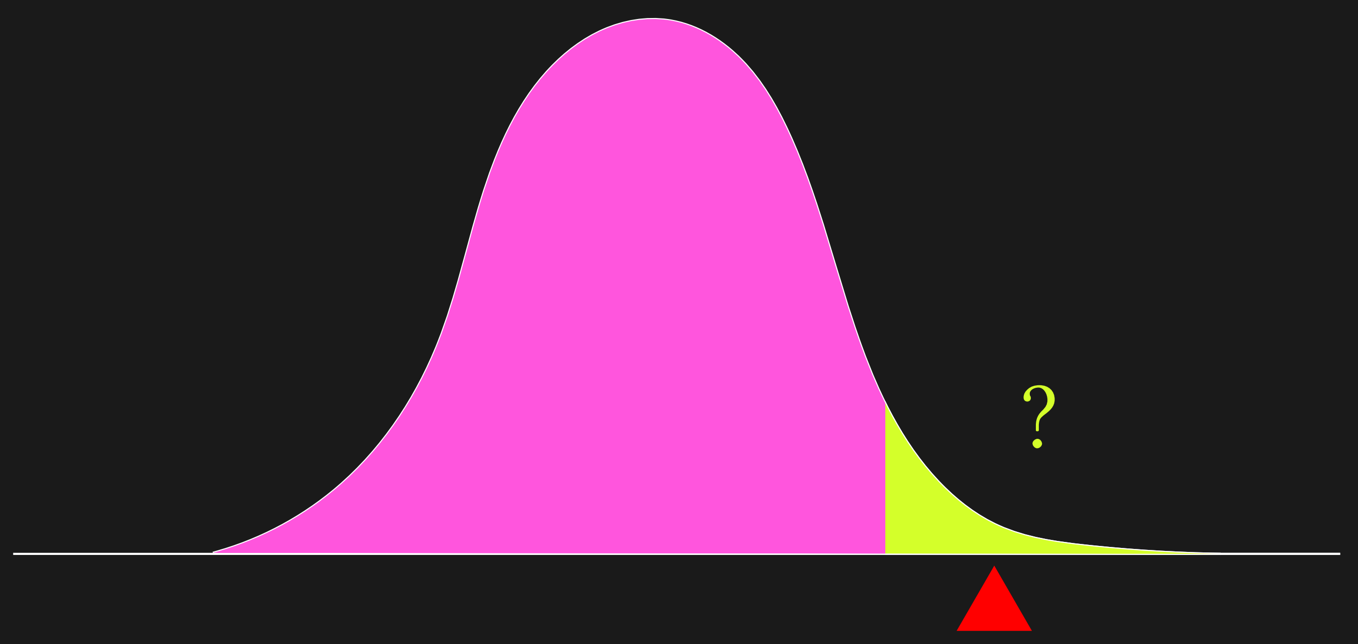
Does our metric value fall within the tails ?

(assume matrix-level metric -> one test)



Does our metric value fall within the tails ?

(assume matrix-level metric -> one test)



-> Your turn !

http://alex.lecairn.org/files/01_nerdchat.R

http://alex.lecairn.org/files/02_nerdchat.R

Problems

Problems: the pringle effect

we lump everything into one single measure

> use a different, pair-level measure

> do it

Problems: multiple tests corrections

Conservative: Bonferonni correction

at least one false discovery

m tests at $5\% \longrightarrow \dfrac{5\%}{\mathrm{m}}$

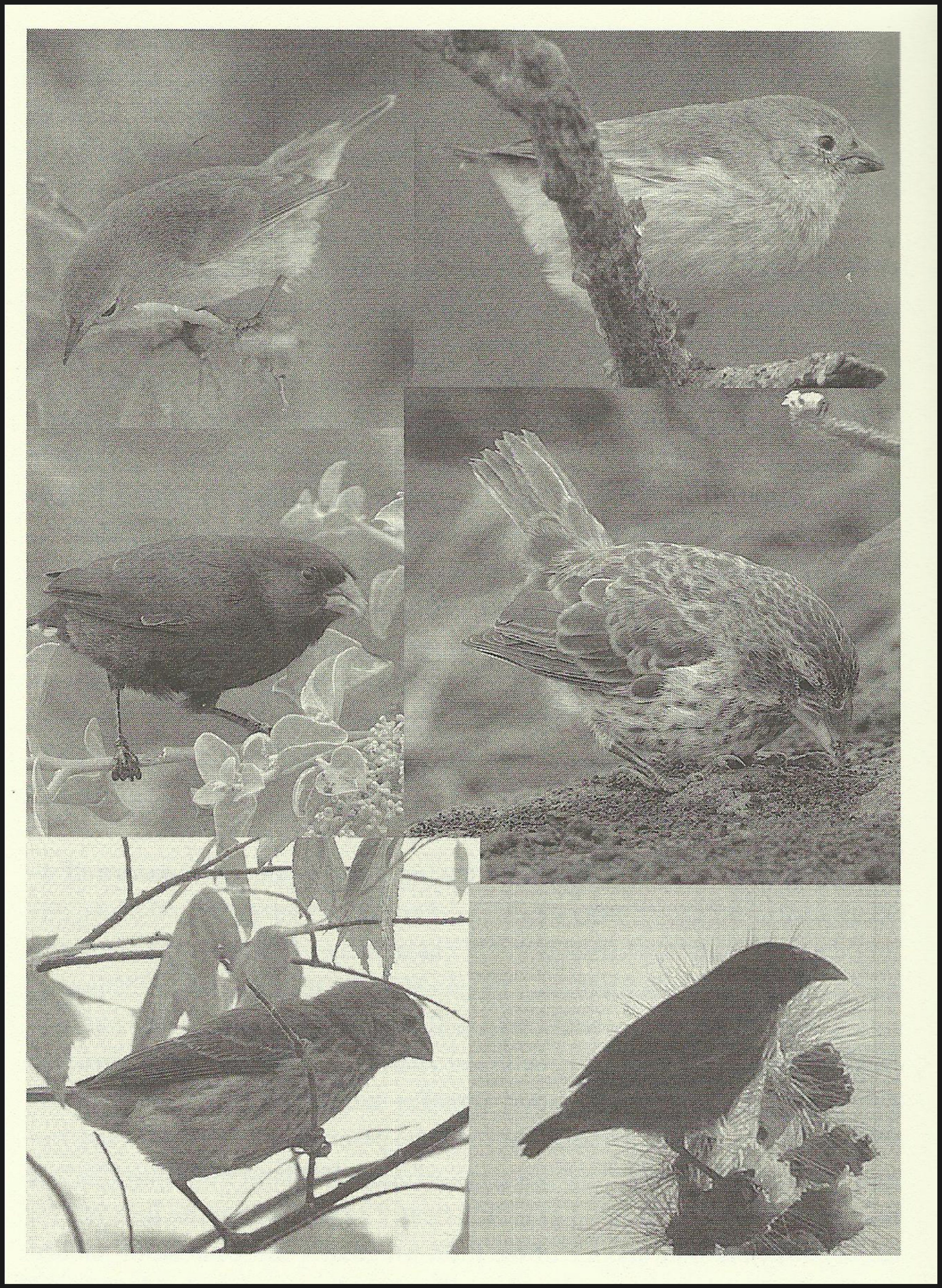
Reasonable: Benjamini-Hochberg procedure (1995)

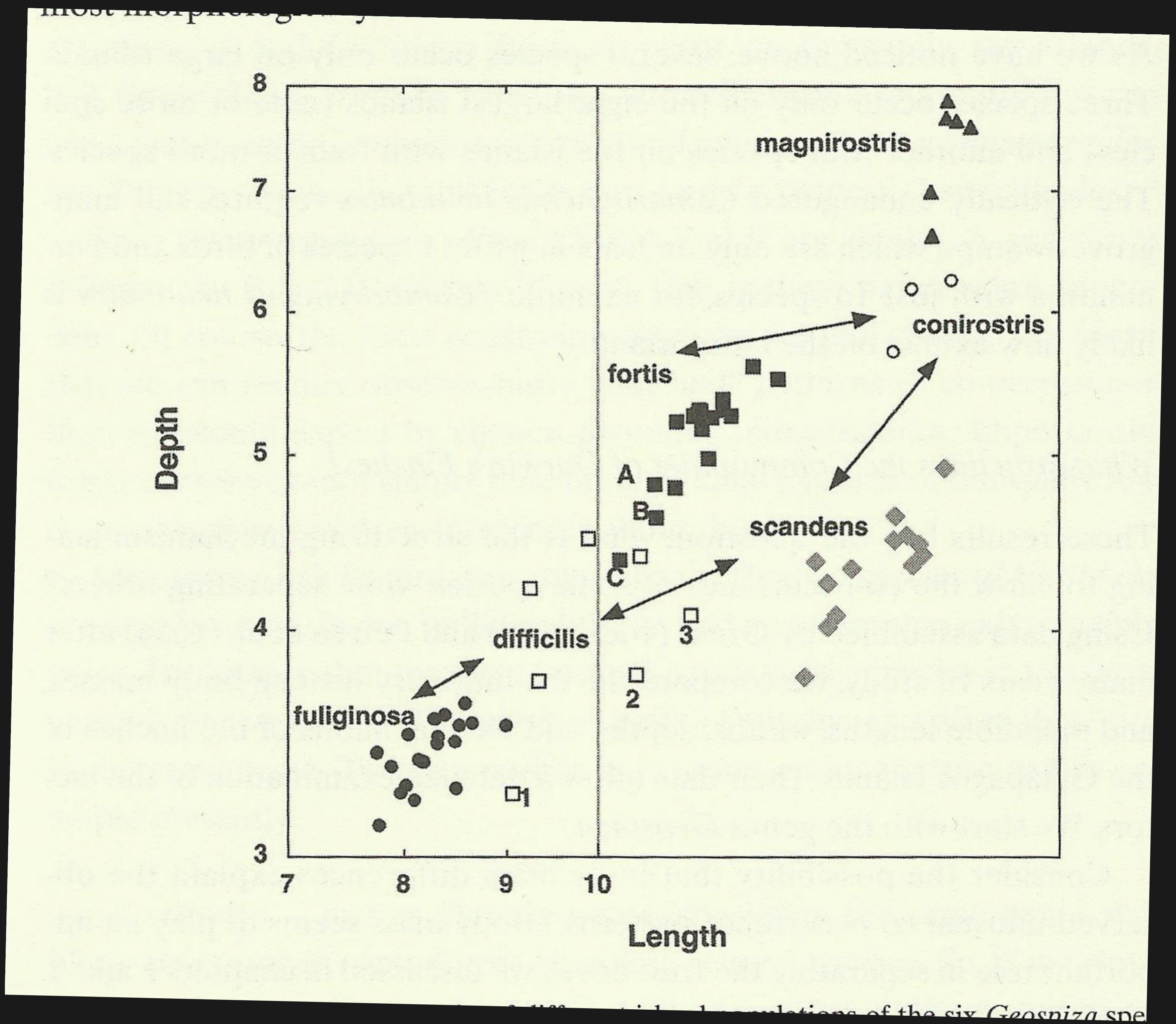
control rate of false discoveries

m tests

$$P_1, P_2, \ldots, P_m \qquad P_{(k)} \leqslant \frac{k}{m}\alpha$$

Case study - the darwin finches: is it a true interaction ?





Positive associations ?

« do we think that species that coocurr frequently are mutualists ? They could be, but there is a simpler explanation. [...] Some islands [...] have habitats not found on smaller ones »

The habitat problem

species that share habitat preferences will coocur more frequently

species that differ in habitat preferences will coocur less frequently

Case study: Morueta-Holme, 2016

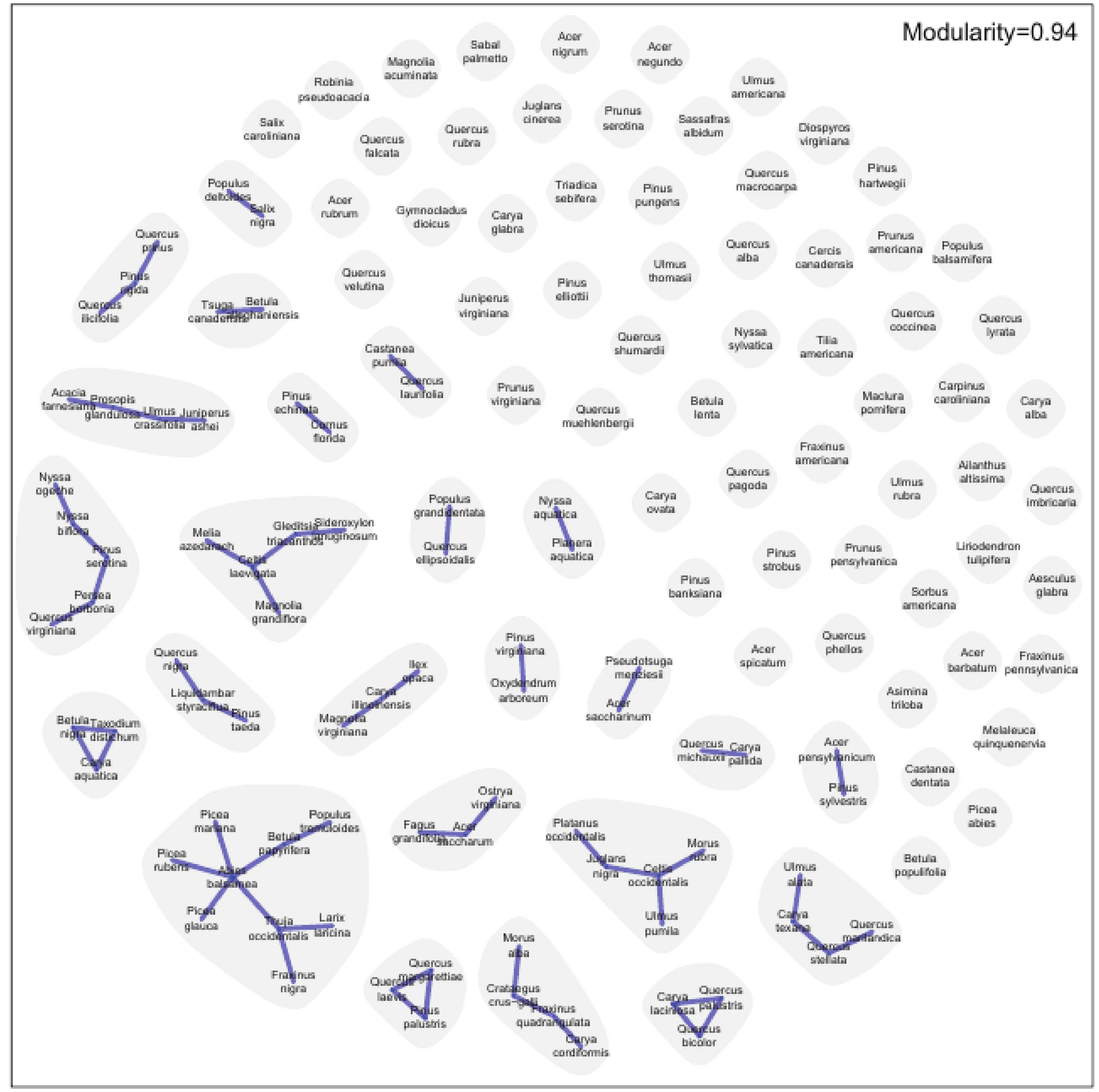


Figure 1. Empirical association network for North American trees. The network was constructed using the MaxEnt regional null model and a shrinkage inverse covariance estimator using 1000 null replicates and a false discovery rate of $\alpha = 0.05$. Gray envelopes indicate distinct modules. Positive associations are shown as blue lines; negative would be shown as red lines but none were found.

associations. For instance, *Carya aquatica* and *Taxodium distichum* are both species of Coastal Plain, strongly associated with large river backswamps that are periodically flooded. A similar habitat is preferred by *Nyssa aquatica* and *Planera aquatica*, although this species pair is found more up-stream than the first ones, in areas where flooding is less prolonged. *Quercus palustris*, *Q. bicolor* and *Carya laciniosa* also prefer swamp habitats, though more from interior flatlands on more calcareous soils. Habitat pref-

A network approach for inferring species associations from co-occurrence data

Naia Morueta-Holme, Benjamin Blonder, Brody Sandel, Brian J. McGill, Robert K. Peet, Jeffrey E. Ott, Cyrille Violle, Brian J. Enquist, Peter M. Jørgensen and Jens-Christian Svenning

Case study: Sfenthourakis 2005

Table 1 Plausible assumptions concerning causality of significant co-occurrence patterns, based on the simultaneous analysis of geographical and ecological data matrices at the species pair level. The geographical matrix contains species presence-absence data, and the ecological matrix contains binary data on species per ecological variable. Positive interaction means higher co-occurrence than expected by chance, negative means less co-occurrence than expected by chance, and none signifies the absence of deviating co-occurrence pattern. Here, allopatry is not necessarily meant in the strict sense, but denotes a tendency

	Ecological matrix			
	Interaction	**Positive**	**Negative**	**None**
Geographic matrix	**Positive**	common ecology and history — neutral/symbiotic	common history	common history
	Negative	competition — ecological 'allopatry'	irrelevance — historical 'allopatry'	historical 'allopatry'
	None	common ecology and history — neutral	different ecology	random occurrence

Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2005) **15**, 39–49

RESEARCH PAPER

Species co-occurrence: the case of congeneric species and a causal approach to patterns of species association

Spyros Sfenthourakis,* Evangelos Tzanatos and Sinos Giokas

Case study: Sfenthourakis 2005

Table 1 Plausible assumptions concerning causality of significant co-occurrence patterns, based on the simultaneous analysis of geographical and ecological data matrices at the species pair level. The geographical matrix contains species presence-absence data, and the ecological matrix contains binary data on species per ecological variable. Positive interaction means higher co-occurrence than expected by chance, negative means less co-occurrence than expected by chance, and none signifies the absence of deviating co-occurrence pattern. Here, allopatry is not necessarily meant in the strict sense, but denotes a tendency

	Ecological matrix			
	Interaction	**Positive**	**Negative**	**None**
Geographic matrix	**Positive**	common ecology and history — neutral/symbiotic	common history	common history
	Negative	competition — ecological 'allopatry'	irrelevance — historical 'allopatry'	historical 'allopatry'
	None	common ecology and history — neutral		

Table 6 Terrestrial isopod species pairs whose co-occurrence frequency deviates from that expected by chance in both matrices (species per sampling station and species per ecological variable), and the presumed causal explanation of the observed combination of interactions. The two pairs that are suspect for direct competition are given in boldface

Species pairs that deviate from random in both matrices	Type of interaction in each matrix: geographic	ecological	Causal explanation
Ligidium ghigii × *Paraschizidium album*	positive	negative	common history
Cretoniscellus strinatii × *Trichodillidium malickyi*	positive	positive	common ecology/history
Bathytropa granulata × *Rodoniscus anophthalmus*	positive	positive	common ecology/history
Bathytropa granulata × *Leptotrichus kosswigi*	positive	positive	common ecology/history
Bathytropa granulata × *Porcellio laevis*	positive	positive	common ecology/history
Bathytropa granulata* × *Proporcellio quadriseriatus	**negative**	**positive**	**competition**
Bathytropa granulata × *Protracheoniscus babori*	positive	positive	common ecology/history
Bathytropa granulata × *Schizidium hybridum*	positive	positive	common ecology/history
Chaetophiloscia elongata × *Armadillidium marmoratum*	positive	positive	common ecology/history
Chaetophiloscia elongata × *Armadillidium vulgare*	positive	positive	common ecology/history
Platyarthrus lindbergi* × *Armadillidium vulgare	**negative**	**positive**	**competition**
Porcellio werneri × *Armadillidium ameglioi*	positive	positive	common ecology/history
Trachelipus aegaeus × *Armadillidium vulgare*	positive	positive	common ecology/history
Paraschizidium aegaeum × *Schizidium tinum*	positive	positive	common ecology/history

Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2005) **15**, 39–49

RESEARCH PAPER

Species co-occurrence: the case of congeneric species and a causal approach to patterns of species association

Spyros Sfenthourakis,* Evangelos Tzanatos and Sinos Giokas

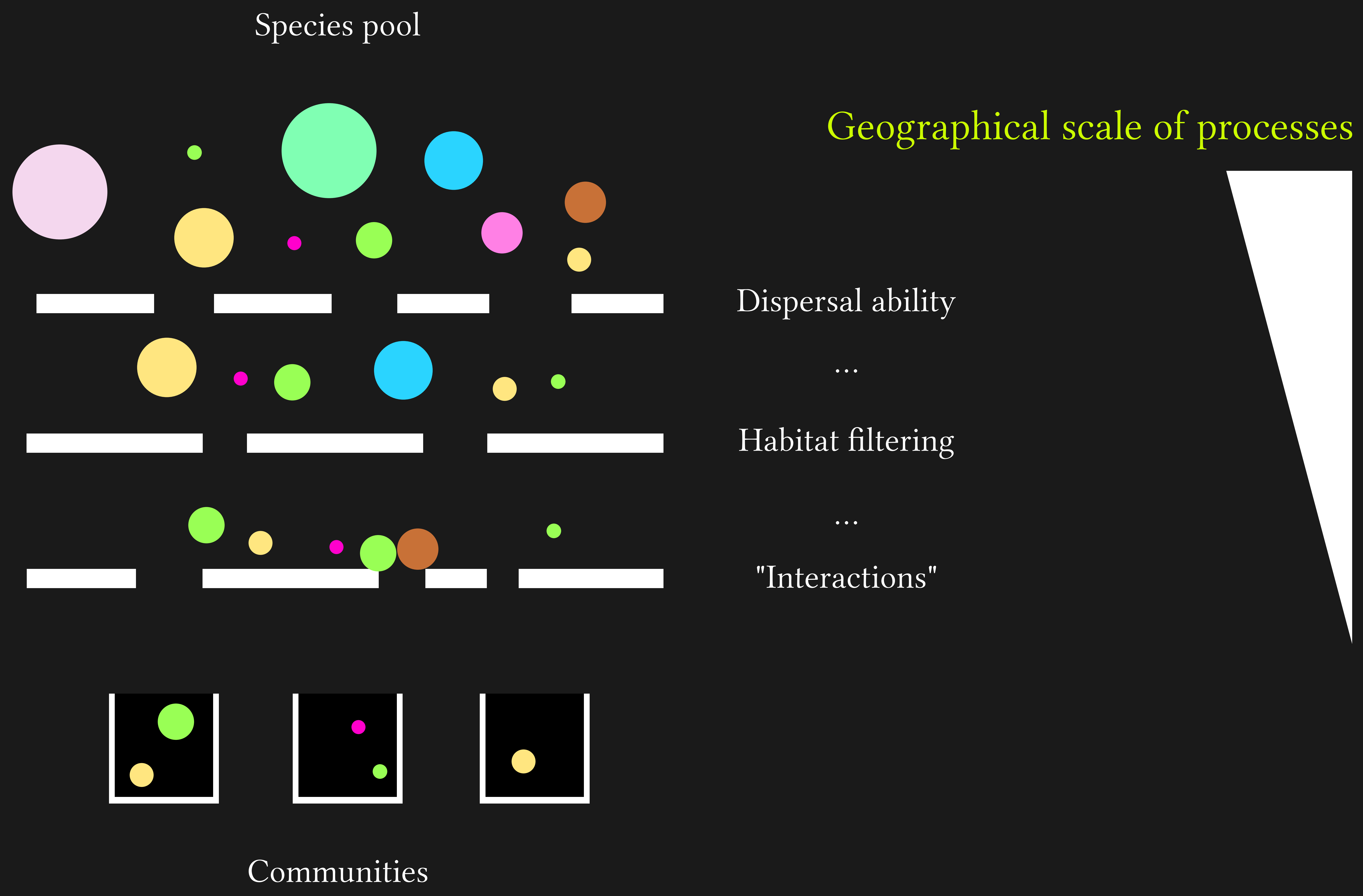
The habitat problem

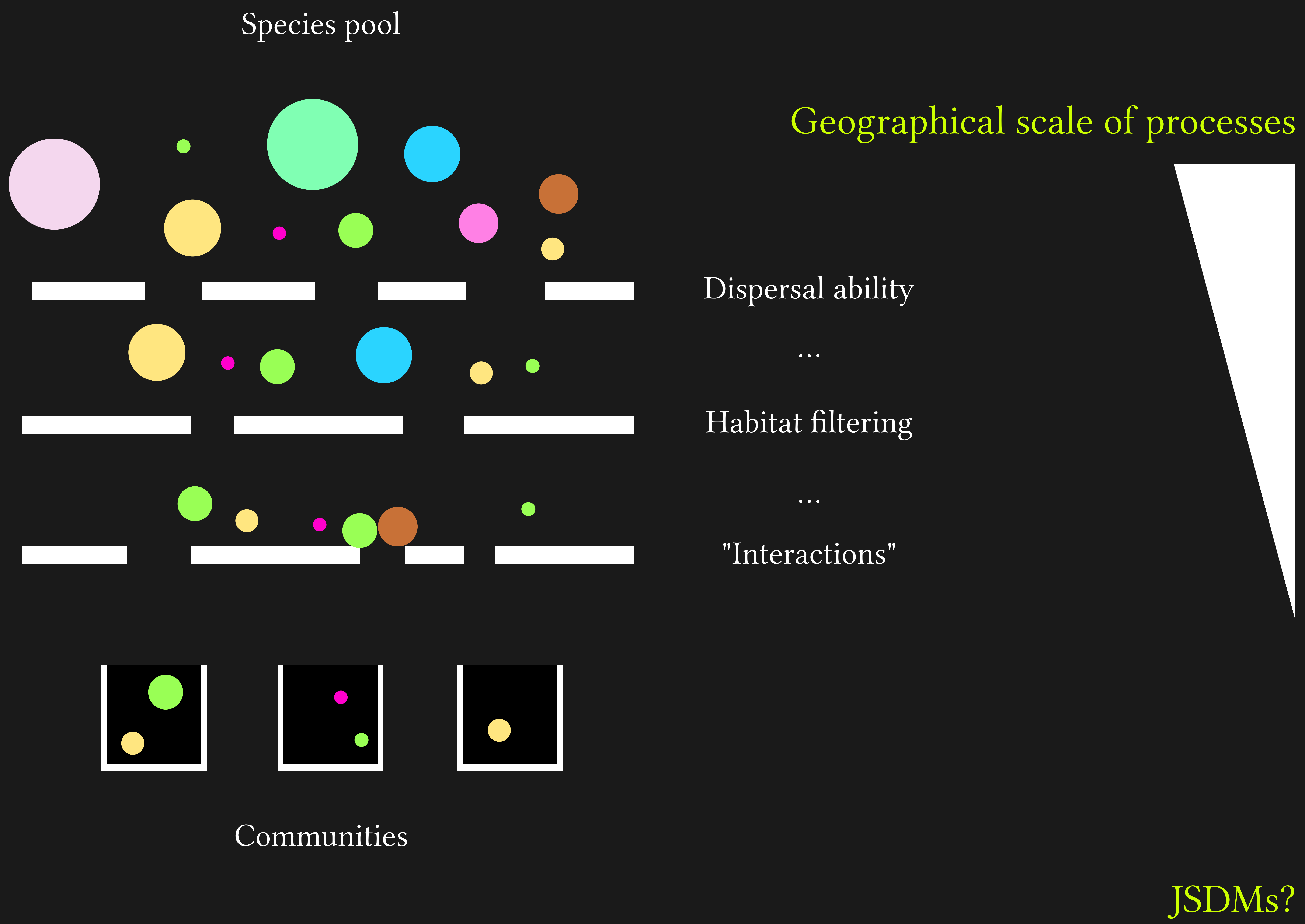
…how well do you describe your habitat ?

is there a hidden variable ?

The dispersal problem

The habitat problem


Species pool
Geographical scale of processes
Dispersal ability
...
Habitat filtering
...
"Interactions"
Communities


Species pool
Geographical scale of processes
Dispersal ability
...
Habitat filtering
...
"Interactions"
Communities
JSDMs?

JSDMs ?

Methods in Ecology and Evolution

Methods in Ecology and Evolution 2014, **5**, 397–406 doi: 10.1111/2041-210X.12180

Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM)

Laura J. Pollock[1]†, Reid Tingley[1]†, William K. Morris[1], Nick Golding[2], Robert B. O'Hara[3], Kirsten M. Parris[1], Peter A. Vesk[1] and Michael A. McCarthy[1]*

Frogs

shared responses to environmental variables. Facilitative interactions between frogs seem unlikely, but positive residual correlations between frog species could have been due to a shared response to an abiotic variable that was not considered in our model, such as the presence of fish (Hamer & Parris (2013).

JSDMs ?

Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2014, **5**, 397–406 doi: 10.1111/2041-210X.12180

Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM)

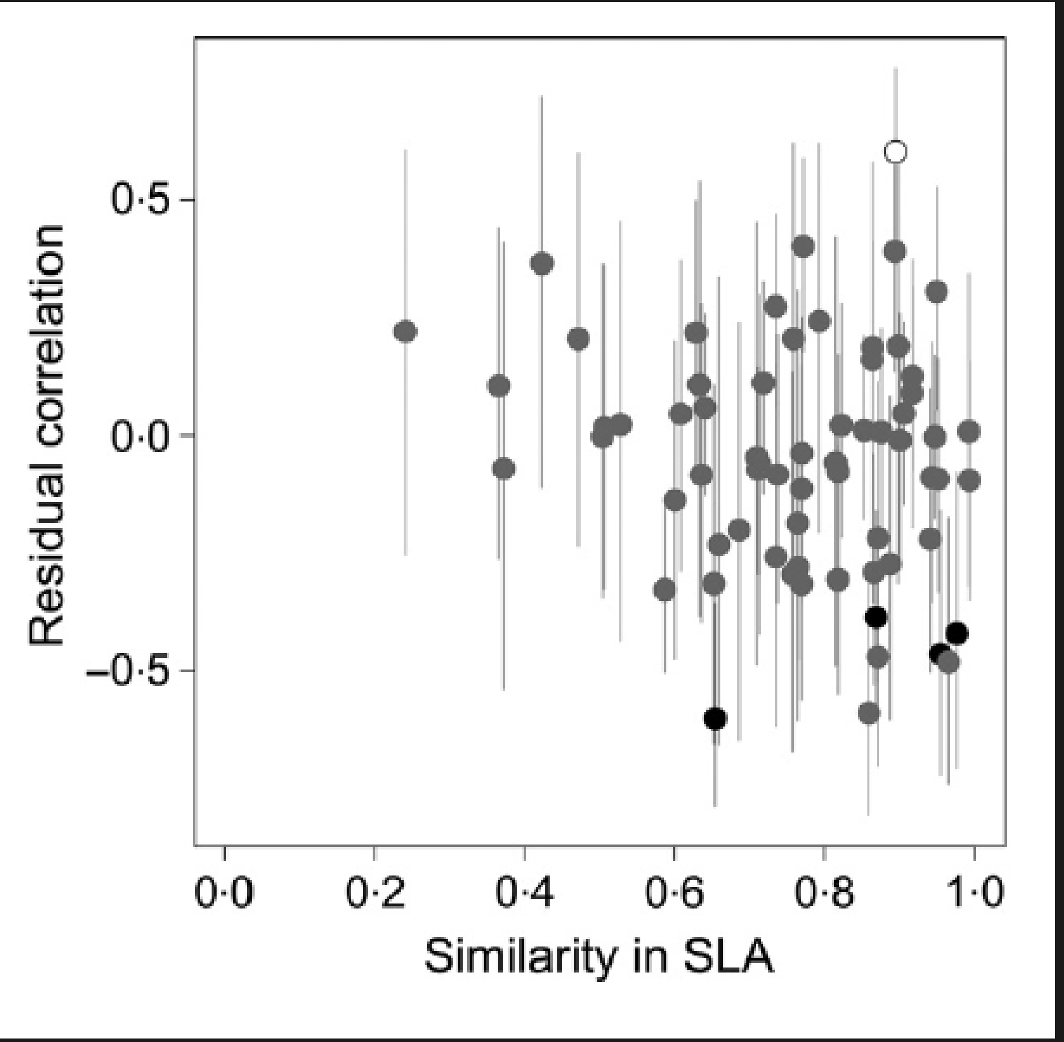
Laura J. Pollock[1]†, Reid Tingley[1]†, William K. Morris[1], Nick Golding[2], Robert B. O'Hara[3], Kirsten M. Parris[1], Peter A. Vesk[1] and Michael A. McCarthy[1]*

Frogs

shared responses to environmental variables. Facilitative interactions between frogs seem unlikely, but positive residual correlations between frog species could have been due to a shared response to an abiotic variable that was not considered in our model, such as the presence of fish (Hamer & Parris (2013).

as Pryor's rule (Pryor 1953). A possible explanation for this pattern is that species from different subgenera are able to differentiate resource use, thereby alleviating competition (Austin, Cunningham & Wood 1983). Another potential explanation is that species from different subgenera are not able to interbreed (Ellis, Sedgley & Gardner 1991). If species

Eucalyptus



Fine-scale coocurrence patterns

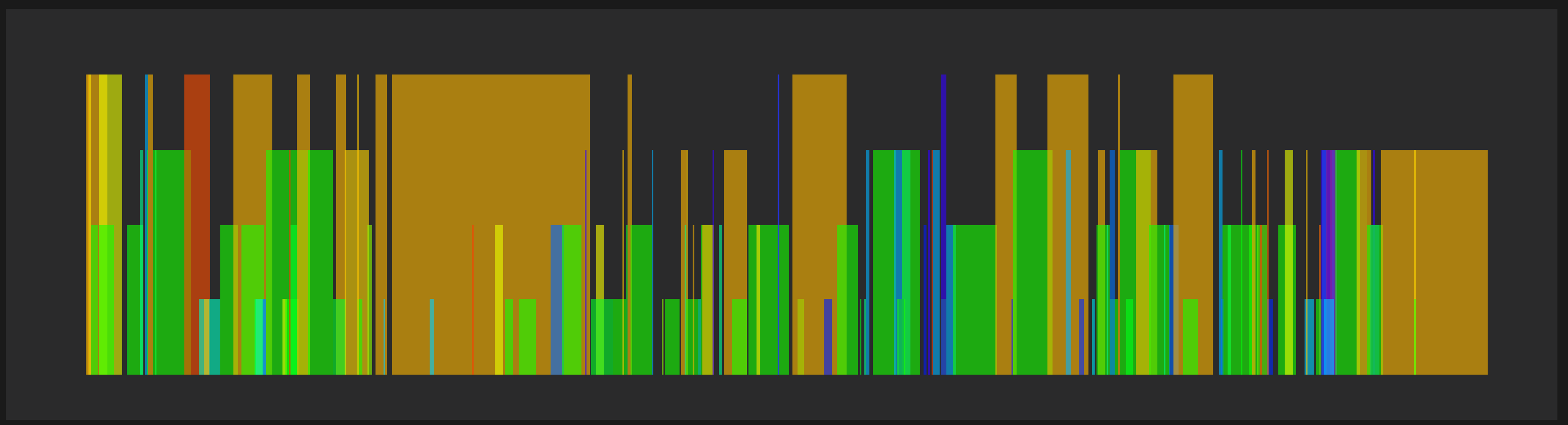
Fine-scale coocurrence patterns

because you believe they compete/facilitate each other

because they really occur in the same place

because they really share the same habitat

5 meters



500 meters

OPEN ACCESS Freely available online | PLoS one

Changes in Semi-Arid Plant Species Associations along a Livestock Grazing Gradient

Hugo Saiz*, Concepción L. Alados

Grupo de Conservación de la biodiversidad, Departamento de Biodiversidad y Restauración, Instituto Pirenaico de Ecología, Consejo Superior de Investigaciones Científicas, Zaragoza, Spain

Concluding remarks

Testing coocurrence through permutations is...

easy to get wrong !

metric
algorithm
tests

...but also very studied
(30+ years of research)

more so at large scales !

confounding factors in
the assembly process

Concluding remarks

Testing coocurrence through permutations is...

easy to get wrong !

metric
algorithm
tests

...but also very studied
(30+ years of research)

more so at large scales !

confounding factors in
the assembly process

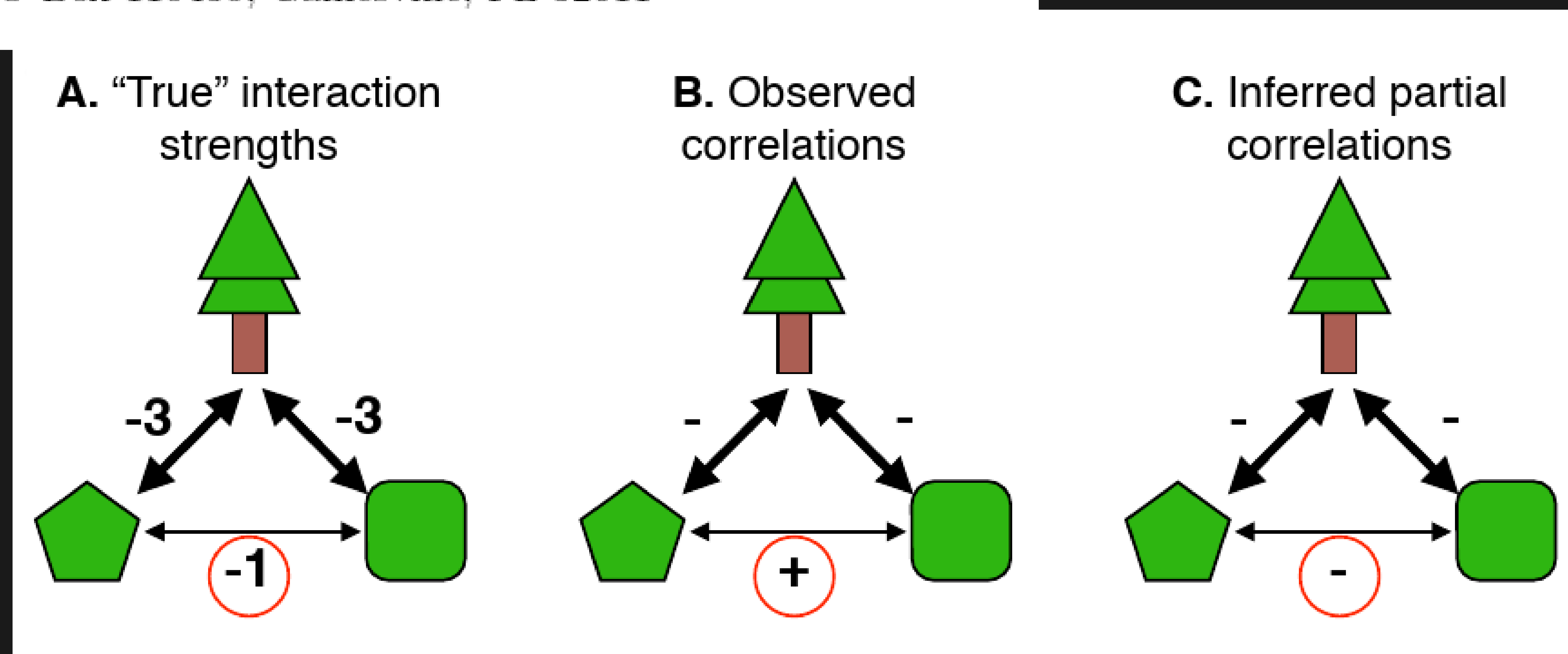
the unit size of sampling (one "community") should not exceed much the length of the interaction (?)

Concluding remarks

Title: Inferring species interactions from co-occurrence data with Markov networks

Author: David J. Harris: Department of Wildlife Ecology and Conservation, University of Florida, 110 Newins-Ziegler Hall PO Box 110430, Gainesville, FL 32611

Markov networks



...towards more efficient methods ?

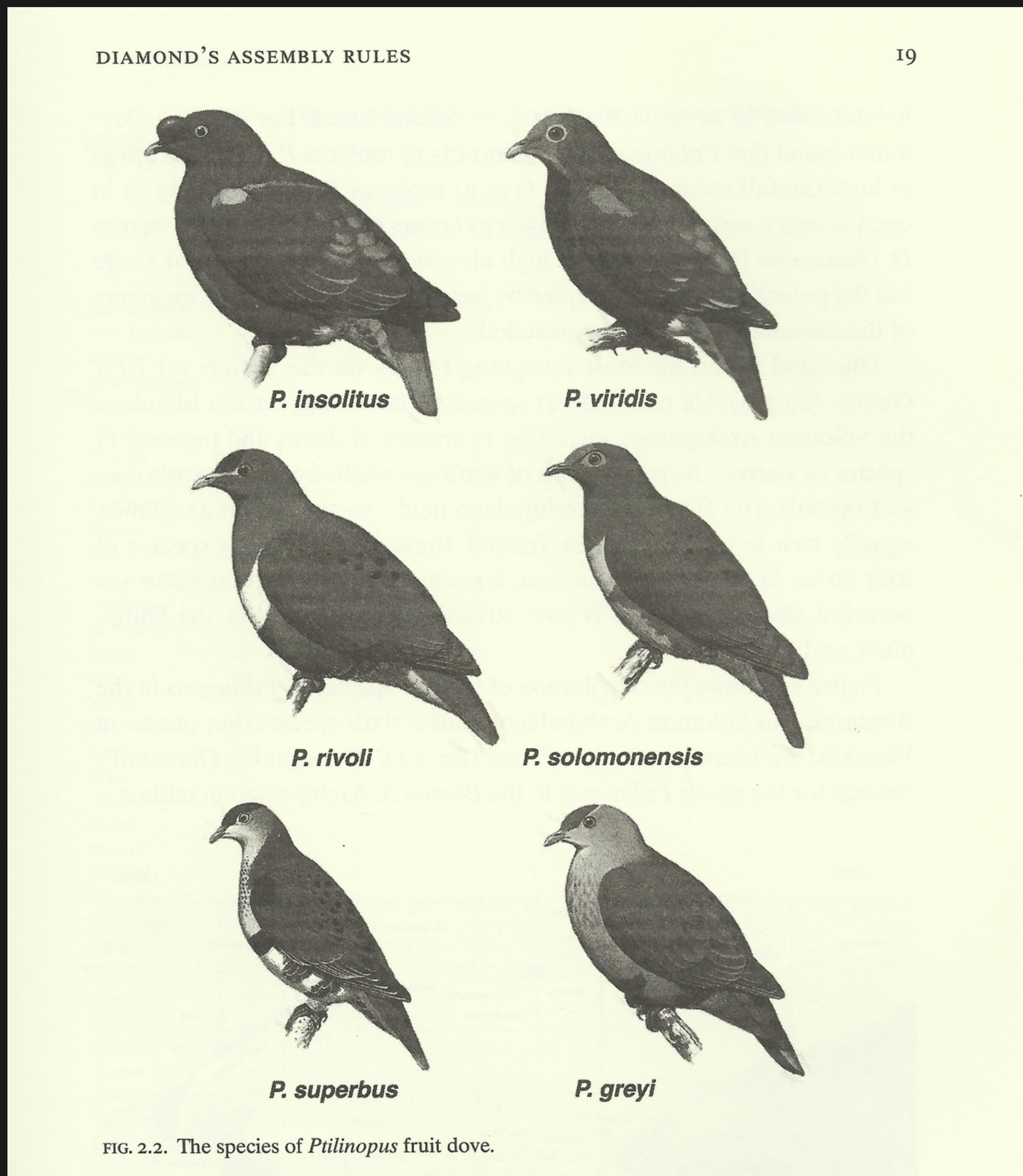

DIAMOND'S ASSEMBLY RULES 19

FIG. 2.2. The species of *Ptilinopus* fruit dove.

Does competition shape communities at large scale ?

or mutualism

Concluding remarks

Title: Inferring species interactions from co-occurrence data with Markov networks

Author: David J. Harris: Department of Wildlife Ecology and Conservation, University of Florida, 110 Newins-Ziegler Hall PO Box 110430, Gainesville, FL 32611

Markov networks