PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

Santa Fe Institute

CMM Centro de Modelamiento Matemático

INSTITUTO DE Sistemas Complejos DE VALPARAÍSO

Centro UC Cambio Global

# The emergence, maintenance and dynamics of biodiversity

Pablo A. Marquet
Facultad de Ciencias Biológicas
P. Universidad Católica de Chile

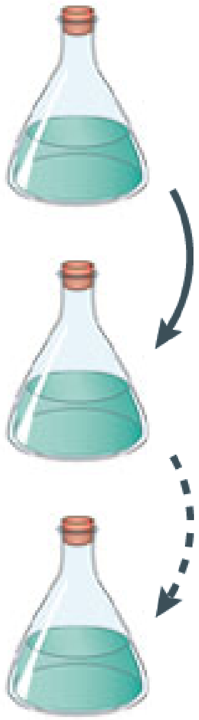
Wissenschaftskolleg zu Berlin
INSTITUTE FOR ADVANCED STUDY

Workshop: Mathematical Modelling and Biodiversity, Paris, MNHN, April 8, 2026

# The emergence and maintenance of biodiversity

**Chemostat**

**Serial transfer**



**Long-term**

Modified from Stams, A. and Plugge, C. (2009) Electron transfer in syntrophic communities of anaerobic bacteria and archaea. Nature Reviews Microbiology

# *E. coli* Long-term starvation experiment (Finkel, Kolter 1999)

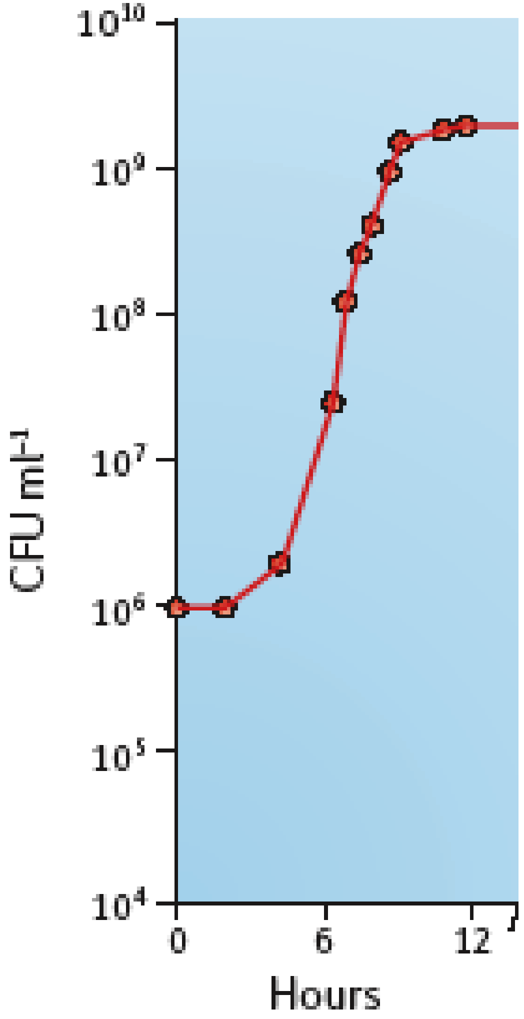
*Proc. Natl. Acad. Sci. USA*
Vol. 96, pp. 4023–4027, March 1999
Microbiology, Evolution

**Evolution of microbial diversity during prolonged starvation**

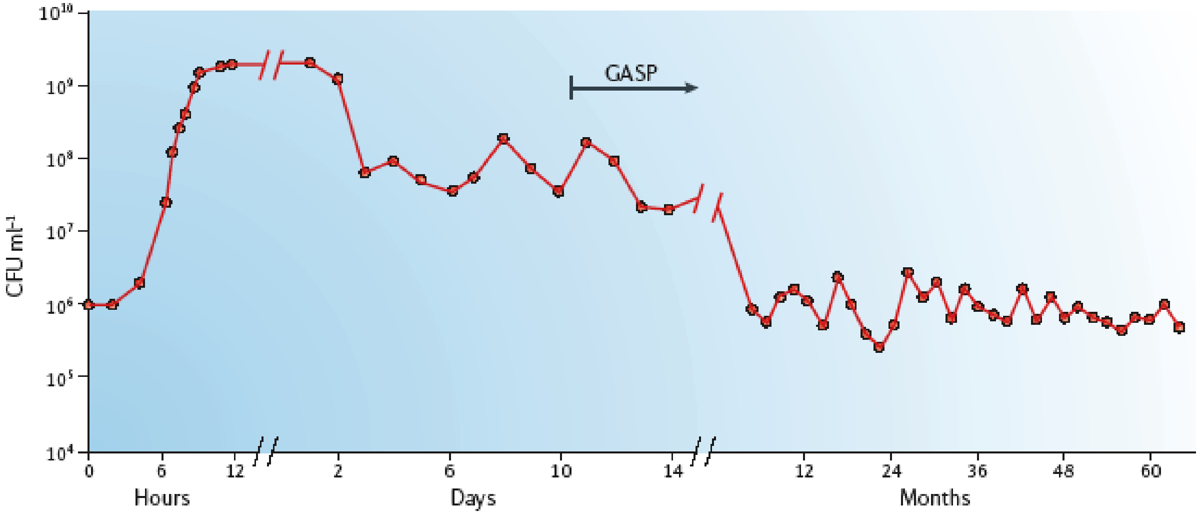
STEVEN E. FINKEL AND ROBERTO KOLTER*

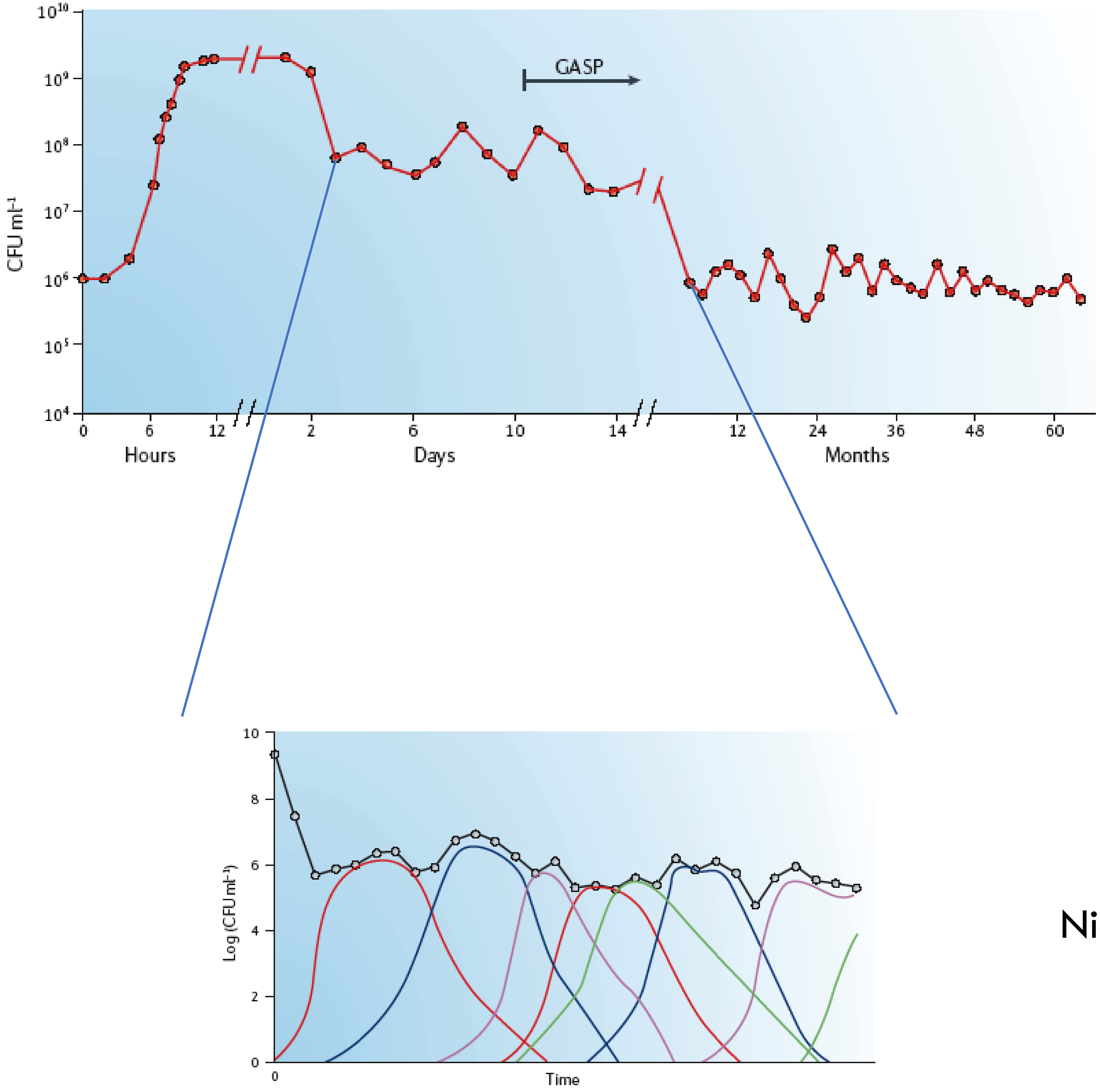
Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115

*Communicated by Stephen Jay Gould, Harvard University, Cambridge, MA, February 8, 1999 (received for review June 15, 1998)*



# *E. coli* Long-term starvation experiment
(Finkel, Kolter 1999)

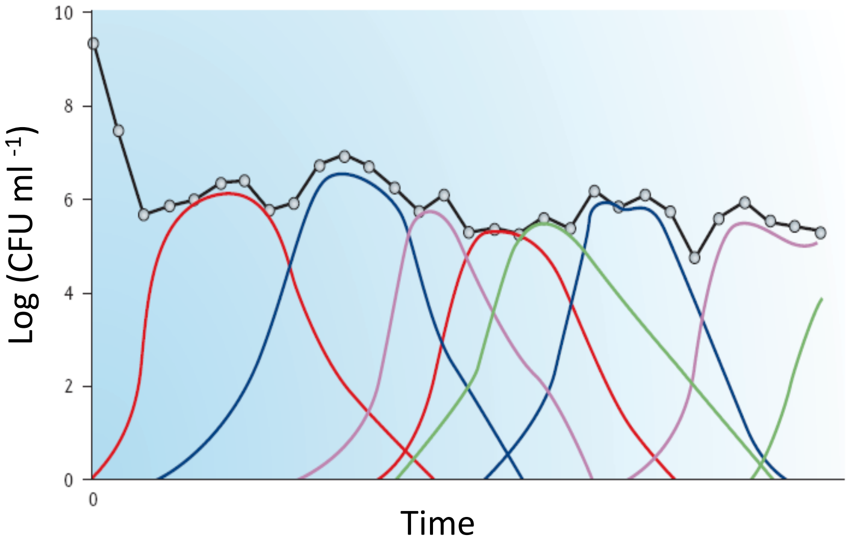




Niche construction

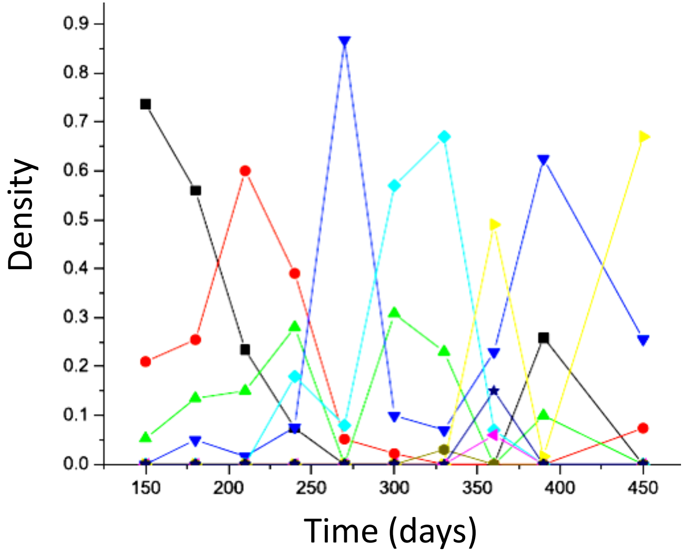
Finkel, S.E. & R. Kolter (1999) PNAS

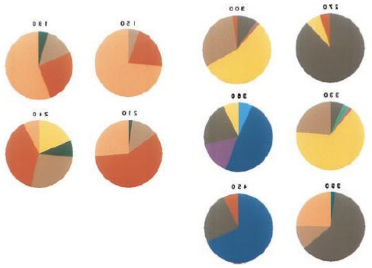
# Periodic Selection



3 months

# Coexistence and neutrality



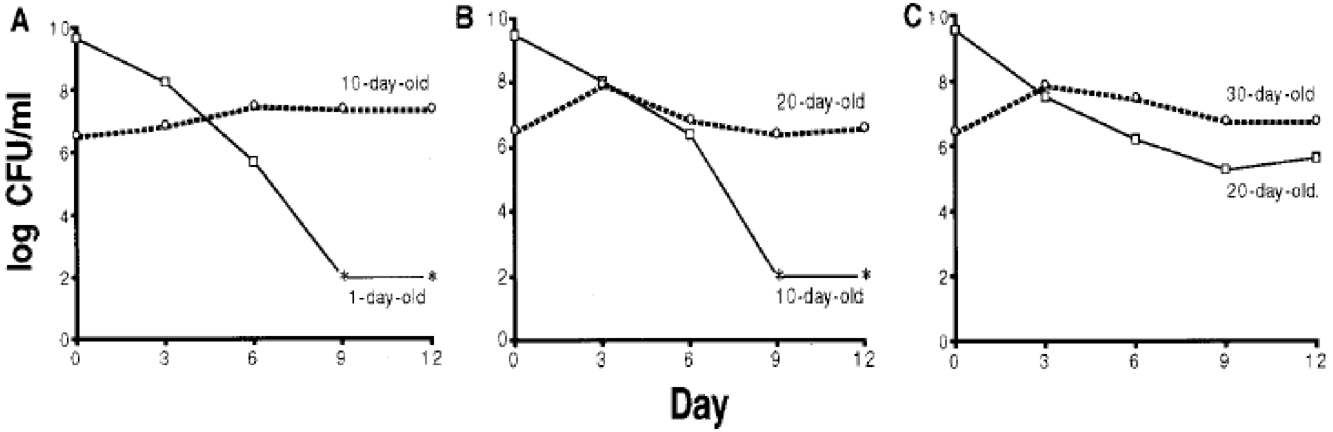


**Finkel, S.E. y R. Kolter** (1999) Evolution of microbial diversity during prolonged starvation. PNAS

**Finkel, S.E.** (2006) Long-term survival during stationary phase: evolution and the GASP phenotype. Nature Rev. Microbiol.

**Keymer et al**.( accepted) Diversity emerging: From competitive exclusion to neutral coexistence in ecosystems.

Older mutants exclude younger ones but older ones coexist.



# A simple model

$$\frac{d\phi}{dt} = r\phi(1-\phi)$$

$$r(\omega) = f\omega - m$$

$$\frac{d\omega}{dt} = S - C$$

$$S = \lambda(1-\omega)$$

$$C = \epsilon\phi\omega f$$

$\phi = Biomass\ proportion$

$r = Growth\ rate$

$m = Mortality$

$\omega = Resource\ concentration$

$f = Fecundity$

$\lambda = Resource\ replenishing\ rate$

$\epsilon = Convertion\ efficiency$

Keymer, Fuentes & Marquet (2012)
*Theoretical Ecology 5*(3), 457-463

$$\frac{d\phi}{dt} = r\phi(1-\phi) \tag{1}$$

$$\frac{d\omega}{dt} = \lambda(1-\omega) - \epsilon\phi\omega f \tag{2}$$

$$\omega^* = \frac{m}{f}$$

Normalizing time

$$\tau = \frac{1}{f}$$

$$\frac{d\phi}{d\tau} = (\omega - \omega^*)\phi(1-\phi) \tag{3}$$

$$\frac{d\omega}{d\tau} = \frac{\lambda}{f}(1-\omega) - \epsilon\phi\omega \tag{4}$$

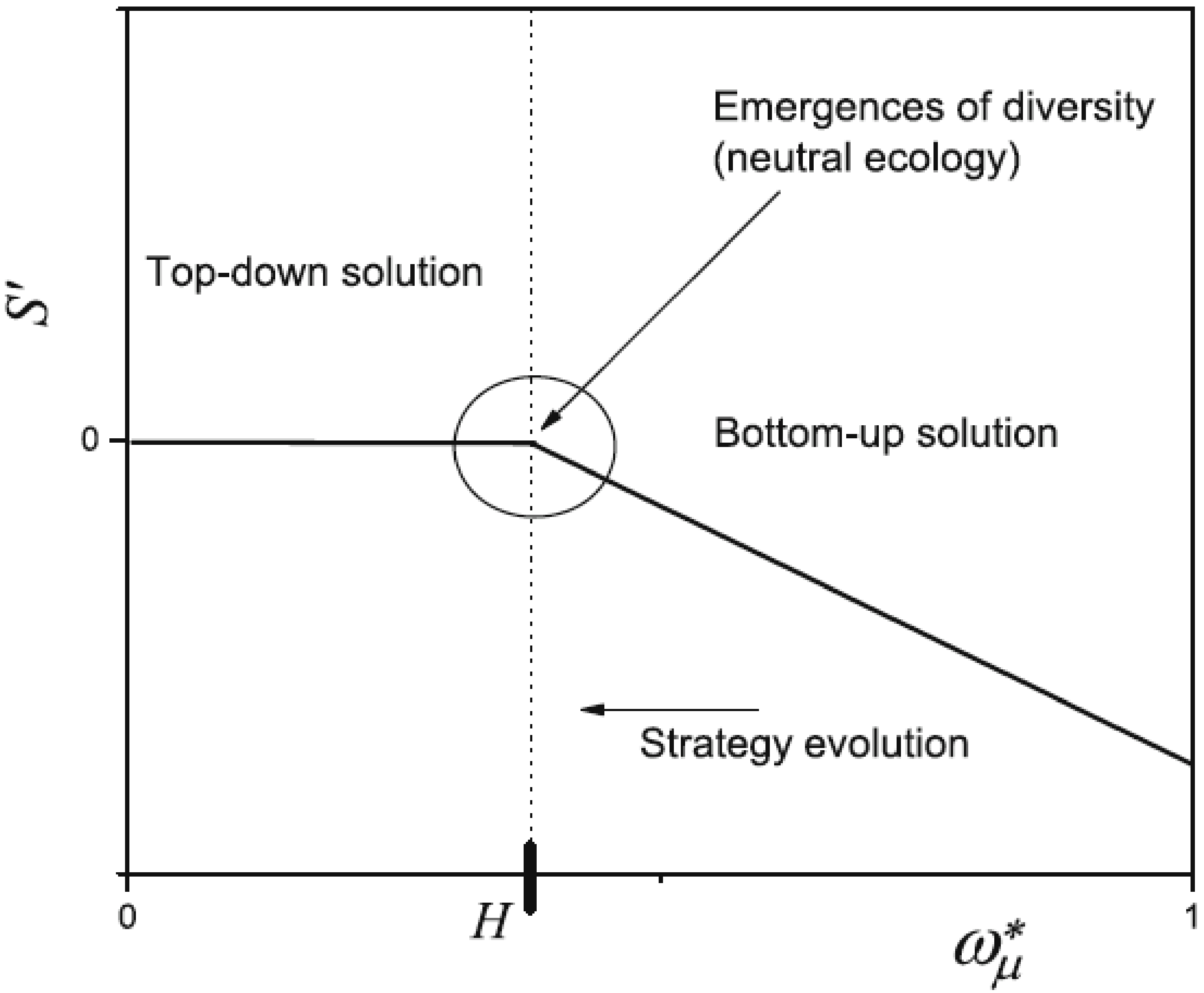
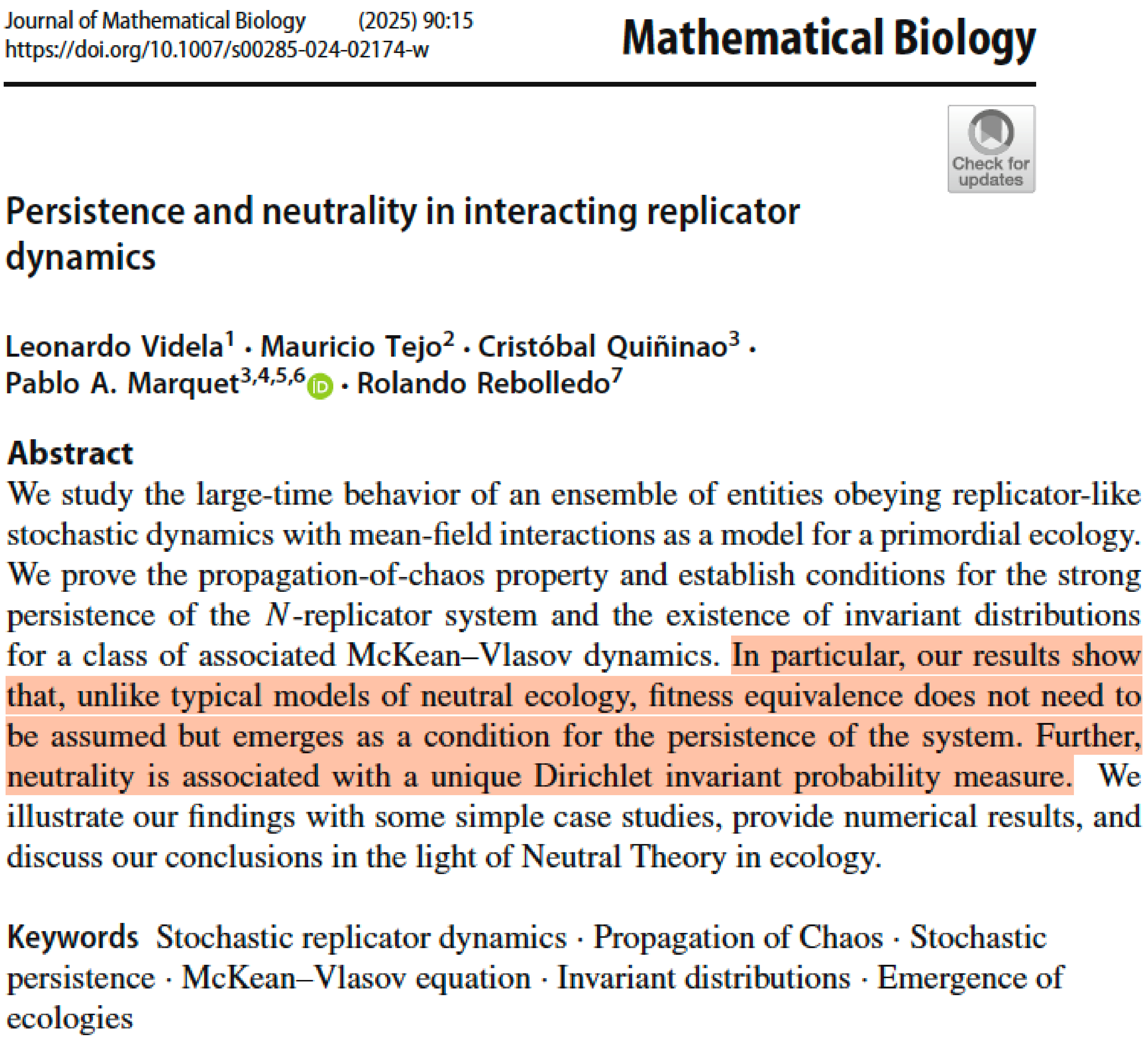


**Fig. 3** Selection gradient as a function of $\omega^*$. The threshold value $H$, is where the top-down solution becomes stable. This is the boundary at which a neutral ecology emerges



Journal of Mathematical Biology (2025) 90:15
https://doi.org/10.1007/s00285-024-02174-w

**Mathematical Biology**

# Persistence and neutrality in interacting replicator dynamics

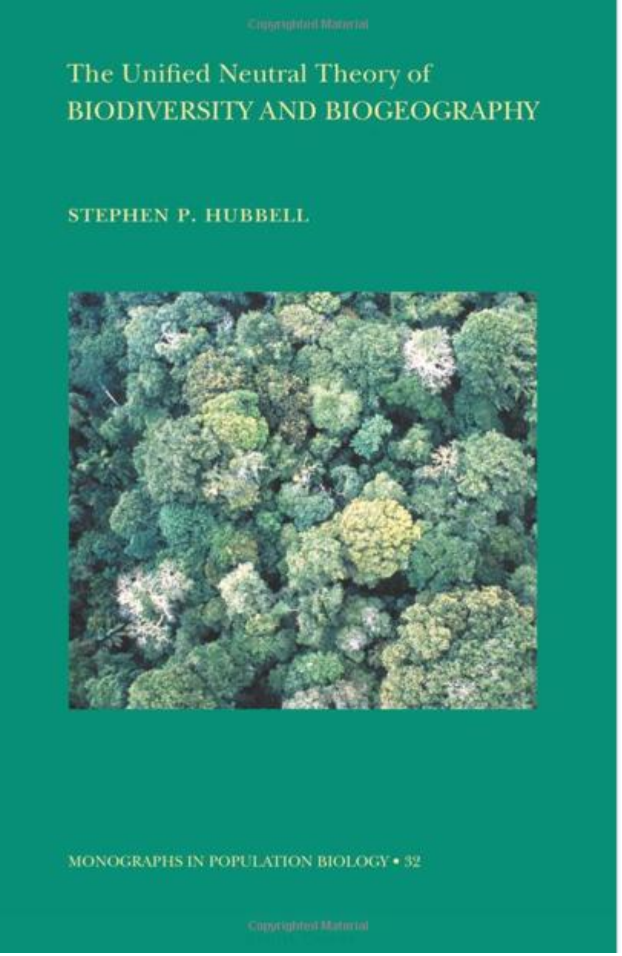
Leonardo Videla[1] · Mauricio Tejo[2] · Cristóbal Quiñinao[3] ·
Pablo A. Marquet[3,4,5,6] · Rolando Rebolledo[7]

## Abstract

We study the large-time behavior of an ensemble of entities obeying replicator-like stochastic dynamics with mean-field interactions as a model for a primordial ecology. We prove the propagation-of-chaos property and establish conditions for the strong persistence of the $N$-replicator system and the existence of invariant distributions for a class of associated McKean–Vlasov dynamics. In particular, our results show that, unlike typical models of neutral ecology, fitness equivalence does not need to be assumed but emerges as a condition for the persistence of the system. Further, neutrality is associated with a unique Dirichlet invariant probability measure. We illustrate our findings with some simple case studies, provide numerical results, and discuss our conclusions in the light of Neutral Theory in ecology.

**Keywords** Stochastic replicator dynamics · Propagation of Chaos · Stochastic persistence · McKean–Vlasov equation · Invariant distributions · Emergence of ecologies

# Neutral theory of Ecology and Biogeography



The Unified Neutral Theory of
BIODIVERSITY AND BIOGEOGRAPHY

STEPHEN P. HUBBELL

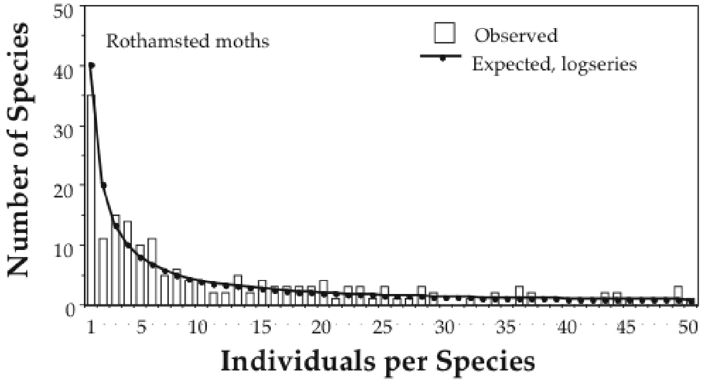
MONOGRAPHS IN POPULATION BIOLOGY • 32

## Master equation for the probability of observing k species with n individuals

$$\boxed{\frac{dP_{n,k}(t)}{dt} = P_{n+1,k}(t)d_{n+1,k}(t) + P_{n-1,k}(t)b_{n-1,k} - P_{n,k}(t)(b_{n,k} + d_{n,k})}$$

Assumptions:

i) The species are assumed to be demographically identical, i.e. $b_{n,k} = b_n$ and $d_{n,k} = d_n$.
ii) Density independent case, i.e. $b_n = b * n$ and $d_n = d * n$ $(n > 0)$



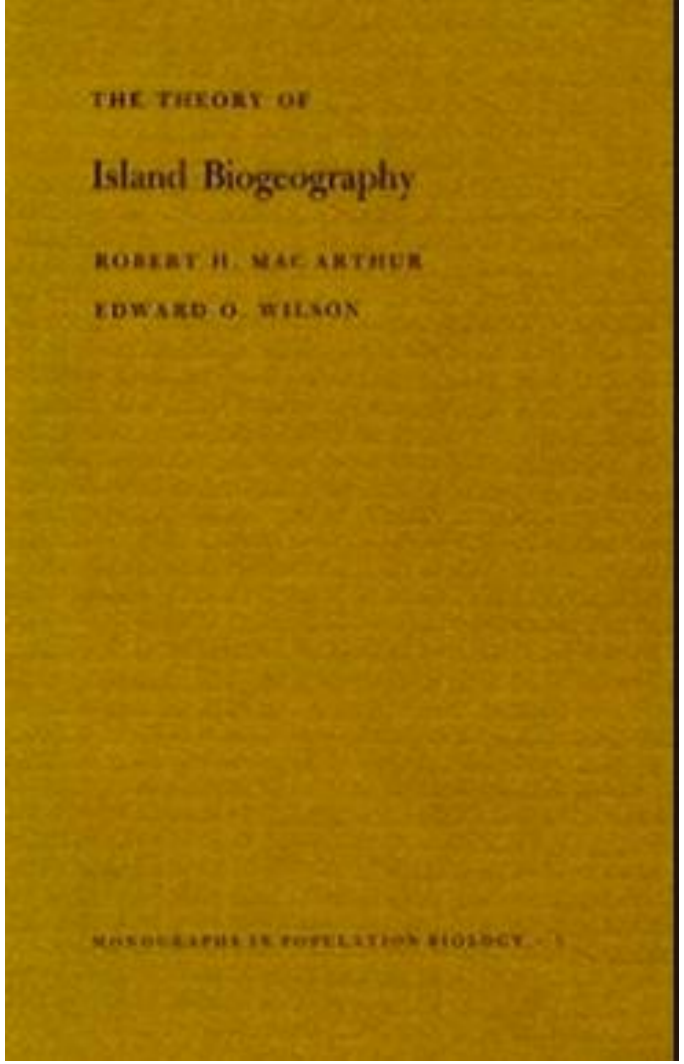
Fisher's Log-series distribution

$$< \Phi_n >= \theta \frac{x^n}{n}$$

where $x = b/d$ and $\theta = SP_0\nu/d$ biodiversity parameter.


**Neutral theory and relative species abundance in ecology**

Igor Volkov[1], Jayanth R. Banavar[1], Stephen P. Hubbell[2,3] & Amos Maritan[4,5]



THE THEORY OF

Island Biogeography

ROBERT H. MAC ARTHUR

EDWARD O. WILSON

MONOGRAPHS IN POPULATION BIOLOGY · 1

$$(3\text{-}4)\quad \frac{dP_S(t)}{dt} = -(\lambda_S + \mu_S)P_S(t) + \lambda_{S-1}P_{S-1}(t) + \mu_{S+1}P_{S+1}(t).$$

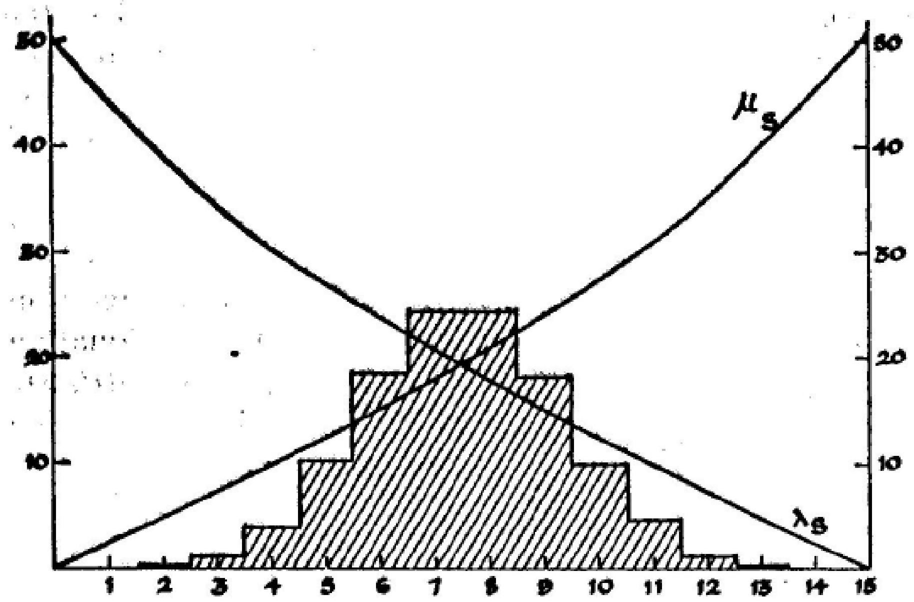


FIGURE 19. A particular case of a predicted distribution of numbers of species on a family of island biotas all with identical extinction and immigration curves and all having had time to reach equilibrium. The histogram represents the number of islands with each number of resident species in an equilibrium situation. The species pool from which the biotas were assembled contained 15 species. If the immigration and extinction curves were straighter, the variance of equilibrial species numbers would be even greater; yet this large variance is still consistent with the equilibrial condition.

*“In principle one could solve eq. 3-4….for our purpose is more useful to find the mean M(t) and the variance, var(t), of the number of species at time t. These can be estimated in nature by measuring the mean and variance in number of species of a series of islands of about the same distance and are and hence of the same λ and μ." pp.33-34*

ME for the number of species within communities (islands)

$$\frac{dP_s(t)}{dt} = P_{s+1}(t)\mu_{s+1} + P_{s-1}(t)\lambda_{s-1} - P_s(t)(\lambda_s + \mu_s),$$

ME for the number of individuals within species

$$\frac{dP_{n,k}(t)}{dt} = P_{n+1,k}(t)d_{n+1,k}(t) + P_{n-1,k}(t)b_{n-1,k} - P_{n,k}(t)(b_{n,k} + d_{n,k})$$

They cannot be true at the same time!

for any $t \in \mathbb{R}_+$, let $(N_1, ..., N_K)(t)$ be the abundance vector of the local community, i.e., $\{N_k(t)\}_{t \in \mathbb{R}_+}$ is the stochastic process accounting for the number of individuals of the species $k$ present in the focal island through time.

let $\{\mathcal{S}(t)\}_{t \in \mathbb{R}_+}$ be the stochastic process accounting for the number of species present in a focal island $A$, say, and consider $K =$ pool.

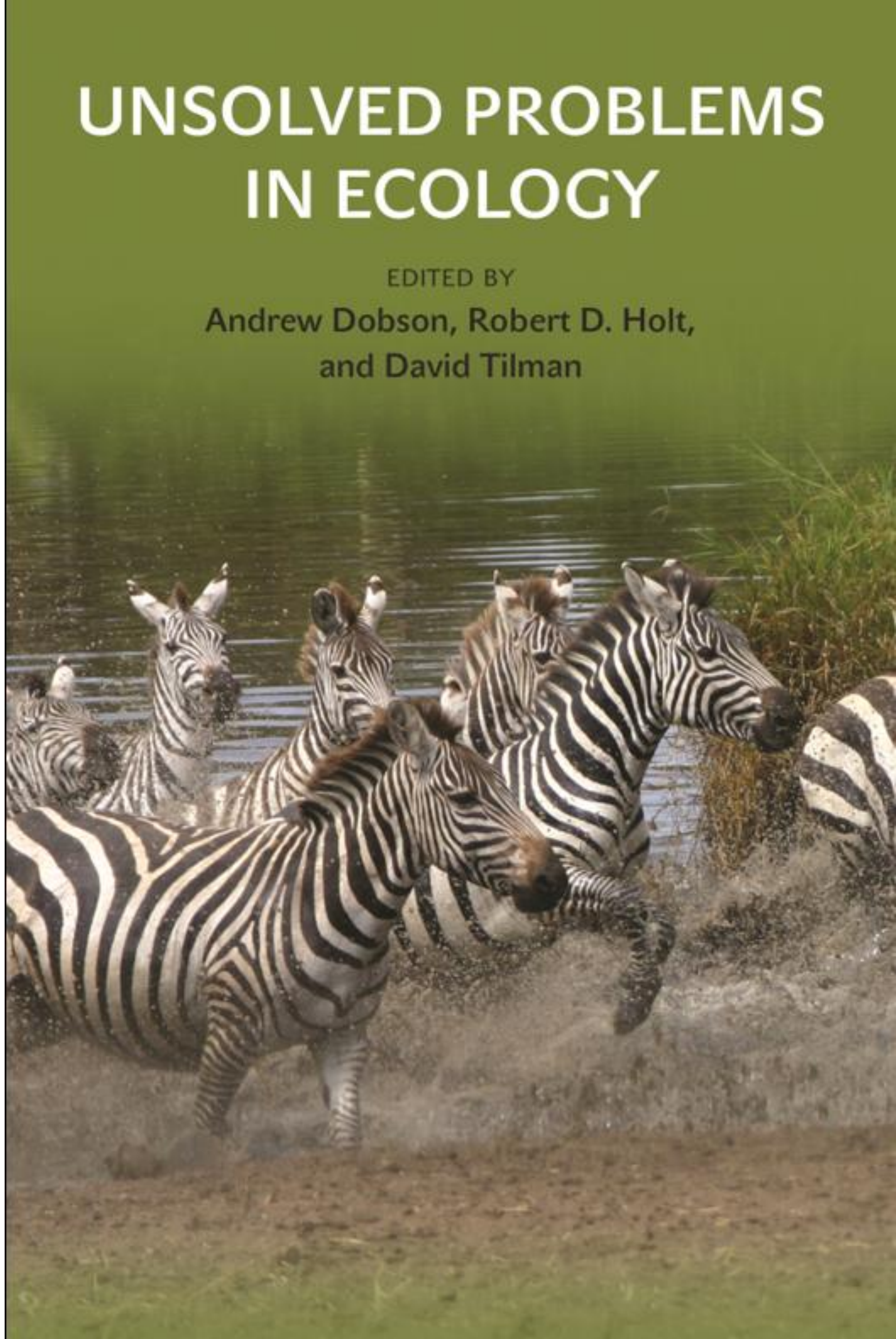
Then, $S(t) = \sum_{k \in A} \mathbf{1}_{[1,\infty[}(N_k(t))$, where $\mathbf{1}_{[0,\infty[}$ denotes the indicator function of $[1, \infty[$ describes the number of species living (or dying) inside the island $A$.

Then, the event $\{s \text{ species at time } t\}$ now depends upon a large number of possibilities of the abundance vector $(N_1, ..., N_K)(t)$:

$$\{S(t) = s\} = \bigcup_{k_1, \ldots, k_s} \{(N_{k_1}(t), \ldots, N_{k_s}(t)) \in [1, \infty[^s\}$$

so that times between transitions $\mathcal{S} \longrightarrow \mathcal{S} \pm 1$ are not longer exponential distributed.

That is, $S(t)$ is a function is a function of a Markov process, but it is not Markov itself.



UNSOLVED PROBLEMS
IN ECOLOGY

EDITED BY
Andrew Dobson, Robert D. Holt,
and David Tilman

# Diffusion processes

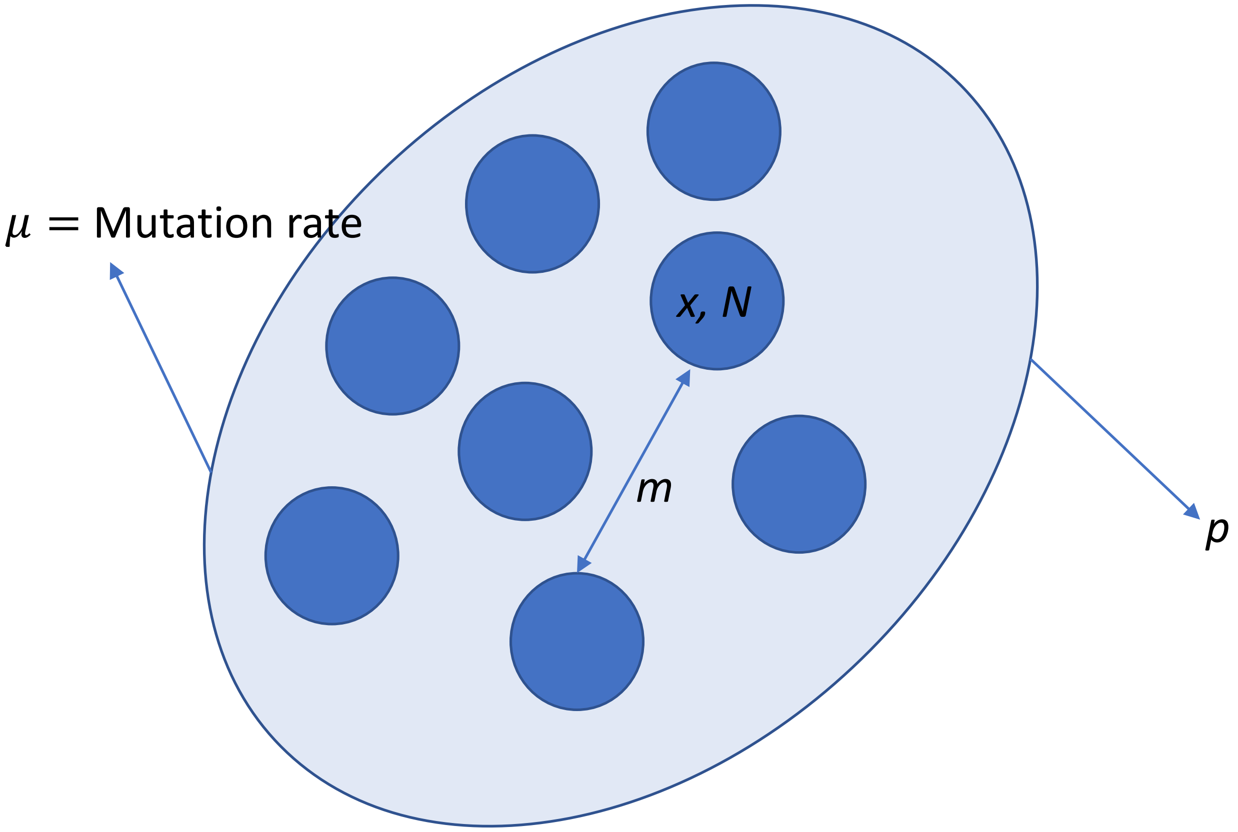
Sewall Wright

Andrey Kolmogorov

Ronald Fisher

- The frequency of genes in a structured population

$$\rho_{\infty}(x) = \frac{\Gamma(4Nm)}{\Gamma(4Nmp)\Gamma(4Nmq)} x^{4Nmq-1}(1-x)^{4Nmp-1}.$$



$x$ = Frequency of a given allele in a local population

N= Effective population size

$m$ = Proportion of migrating individuals among population each generation.

p= Frequency of a given allele in the total population

# Kolmogorov (1935) Dokl. Akad. Nauk SSSR 3 (1935), 129-132.

Large population of ***N*** individuals  consisting of ***s*** partial populations with ***n*** individuals each (*N=sn*)

Each generation ***k*** individuals disperse randomly across the partial populations.

$\bar{p}$ = gene frequency in the large population

$p$ = gen frequency in a local population

$\Delta p$ =change in the frequency of $p$ over one generation

$q = 1 - p$

Following Wright and Fisher:

$$M(p) = E(\Delta p) = \frac{k}{n}(\bar{p} - p) \qquad Var(p) = \ E(\Delta p)^2 = \frac{pq}{2n}$$

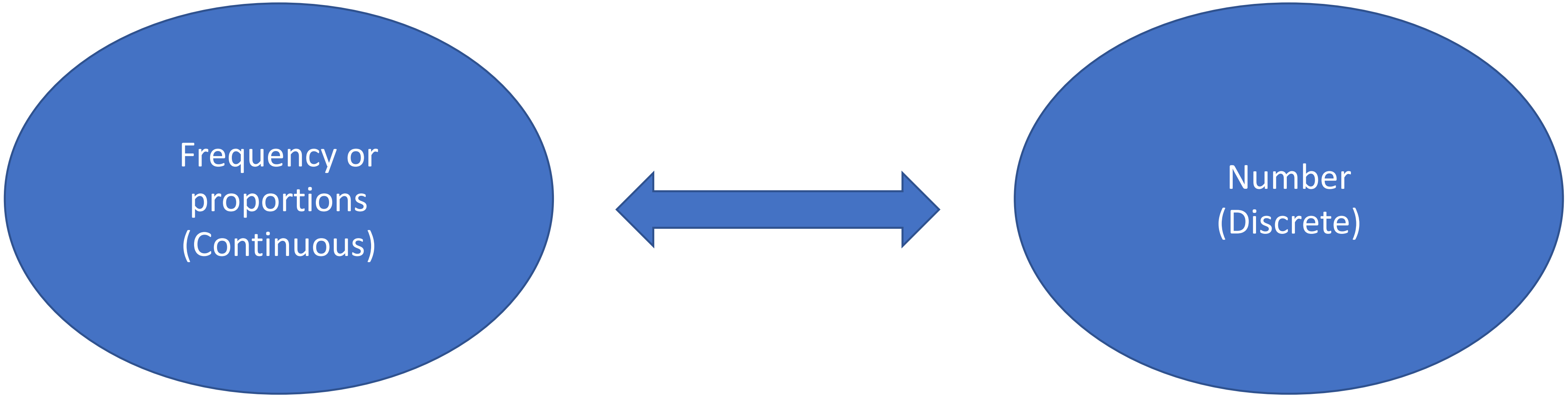
"Since *s* is large, the variation of the total concentration $\bar{p}$ will proceed much more slowly than those of the partial concentrations *p*. Therefore, $\bar{p}$ can temporarily be taken constant. The concentrations *p* in partial populations deviate from $\bar{p}$ in either direction. After sufficiently long time the fluctuations of *p* around $\bar{p}$ result in a certain stationary probability distribution for the concentrations *p*." (Kolmogorov 1935)

This stationary distribution satisfies the Kolmogorov forward or Fokker-Planck equation:

$$\frac{1}{2}\frac{\partial^2}{\partial p^2}(\mathrm{Var(p)\mu}) - \frac{\partial}{\partial x}(M(p)\mu) = 0$$

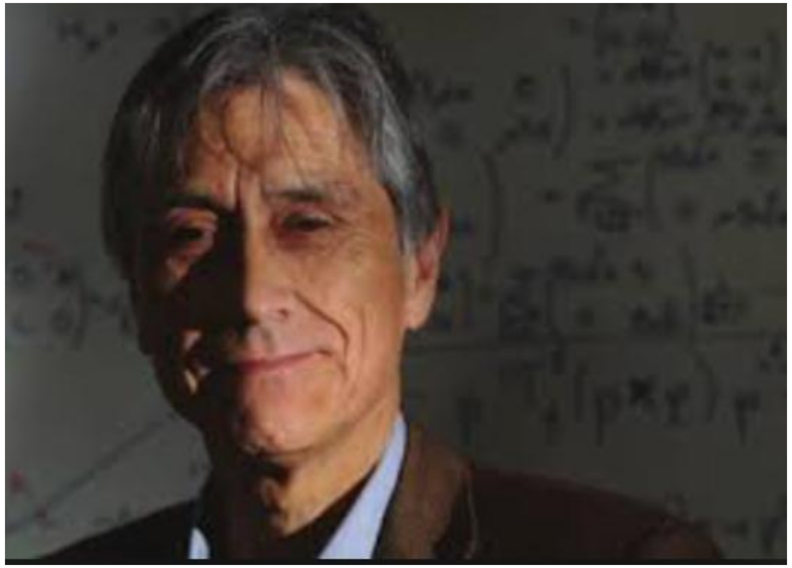
Whose solution $\mu(p)$ is:

$$\mu(p) = \frac{1}{\mathrm{B}(4k\bar{p}, 4k\bar{q})}\ p^{4k\bar{p}-1} q^{4k\bar{q}-1}$$

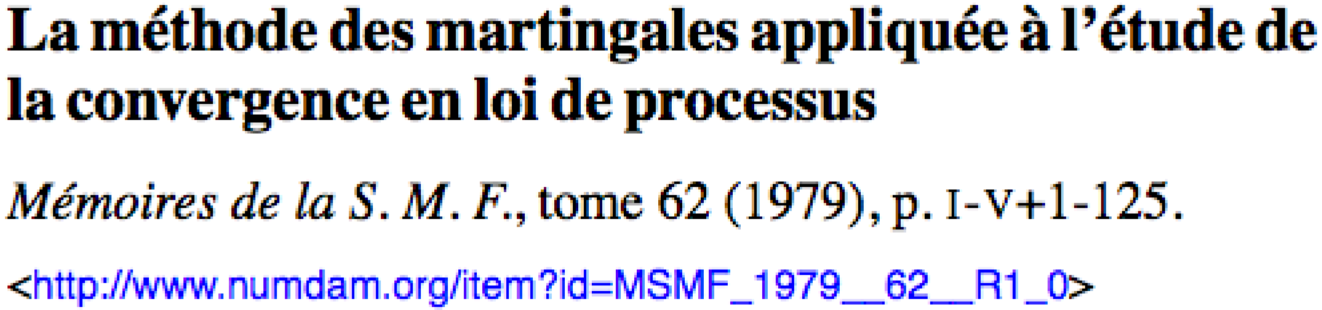


Frequency or
proportions
(Continuous)

Number
(Discrete)



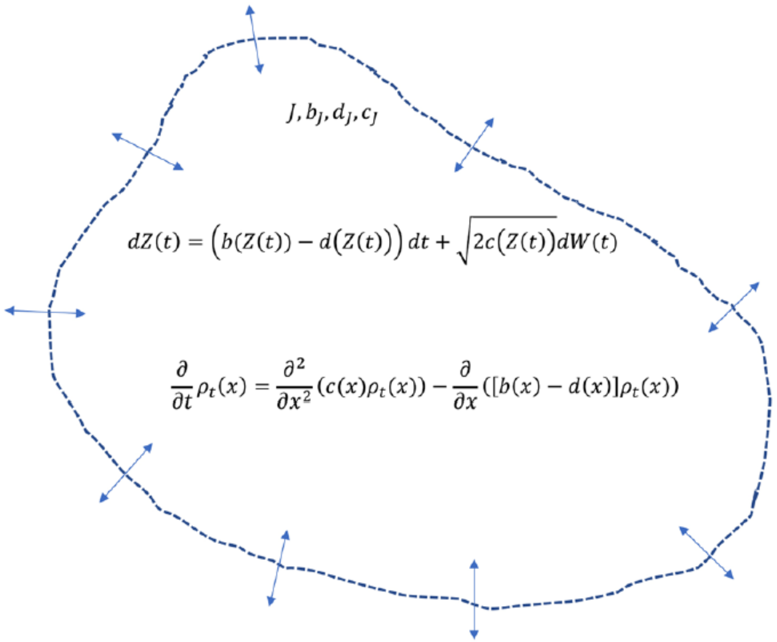
**Prof. Rolando Rebolledo**



**La méthode des martingales appliquée à l'étude de la convergence en loi de processus**

*Mémoires de la S. M. F.*, tome 62 (1979), p. I-V+1-125.

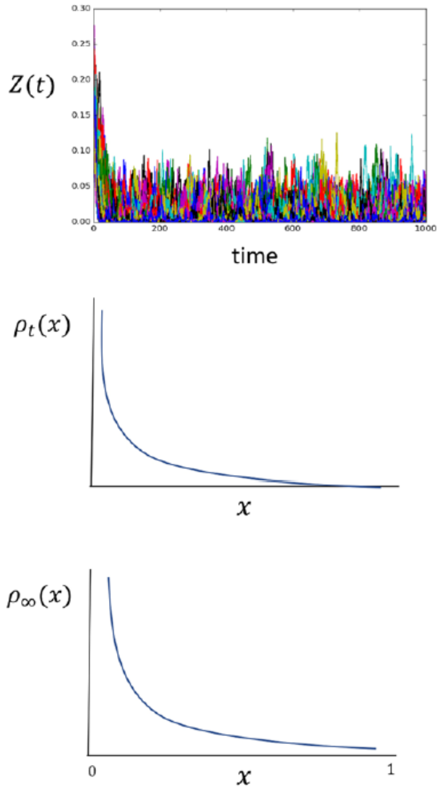
<http://www.numdam.org/item?id=MSMF_1979__62__R1_0>



$J, b_J, d_J, c_J$

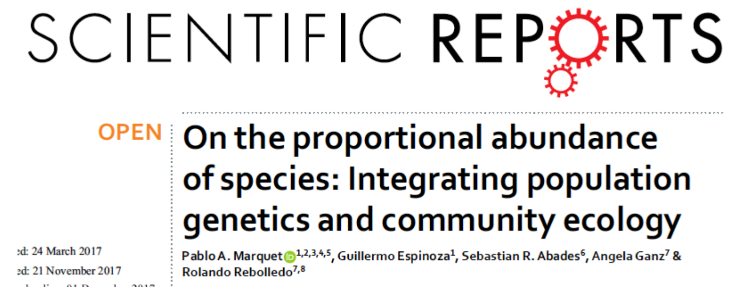
$$dZ(t) = \big(b(Z(t)) - d(Z(t))\big)\,dt + \sqrt{2c(Z(t))}\,dW(t)$$

$$\frac{\partial}{\partial t}\rho_t(x) = \frac{\partial^2}{\partial x^2}(c(x)\rho_t(x)) - \frac{\partial}{\partial x}([b(x) - d(x)]\rho_t(x))$$



$$\rho_\infty(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

$$\begin{aligned} b(x) &= b_0 + b_1 x \\ d(x) &= d_0 + d_1 x \\ c(x) &= \gamma x(1-x), \end{aligned}$$



SCIENTIFIC REPORTS

OPEN

On the proportional abundance of species: Integrating population genetics and community ecology

Pablo A. Marquet[1,2,3,4,5], Guillermo Espinoza[1], Sebastian R. Abades[6], Angela Ganz[7] & Rolando Rebolledo[7,8]

ed: 24 March 2017
ed: 21 November 2017

$$\rho_\infty(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

The Proportional Species Abundance Distribution (PSAD)

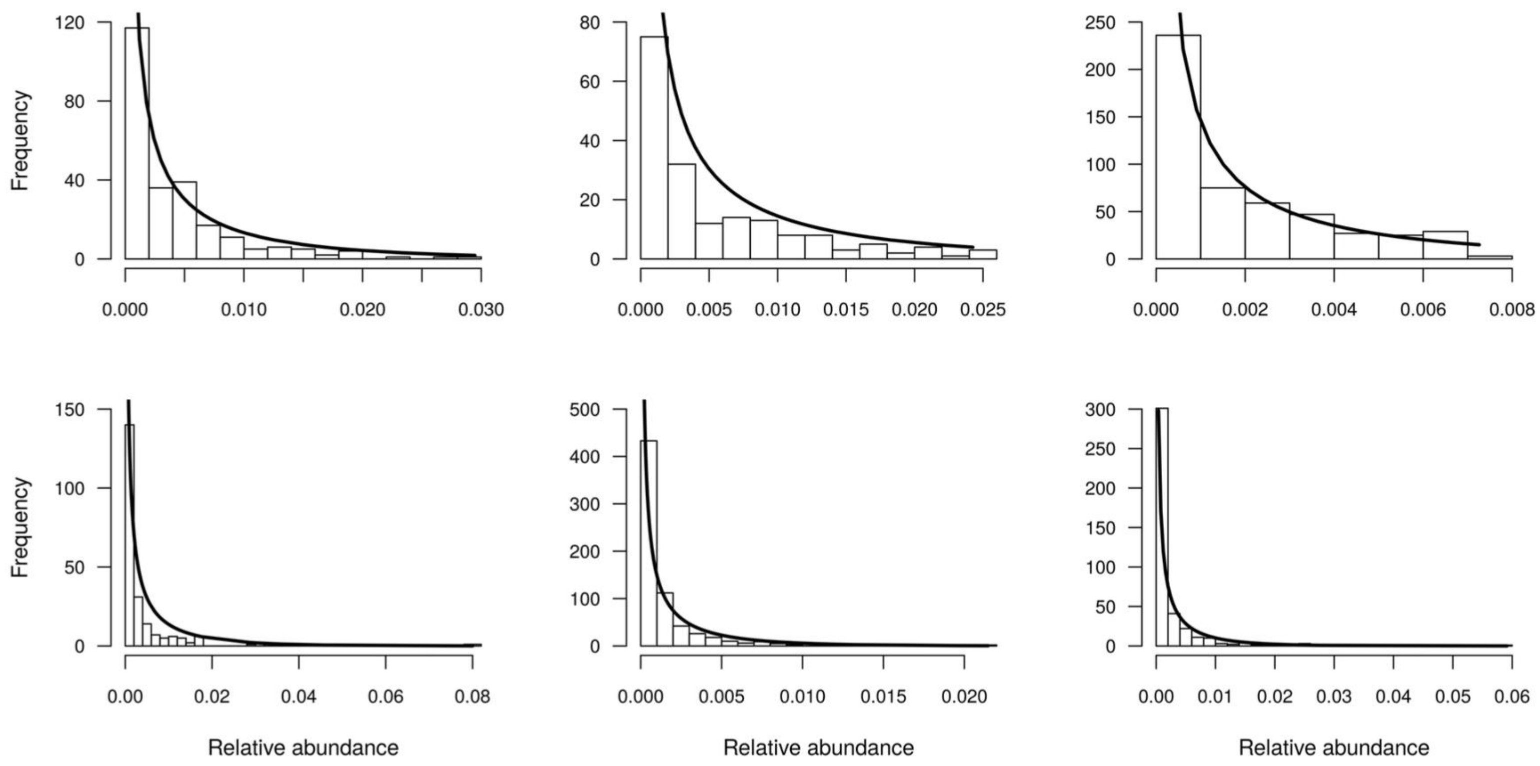
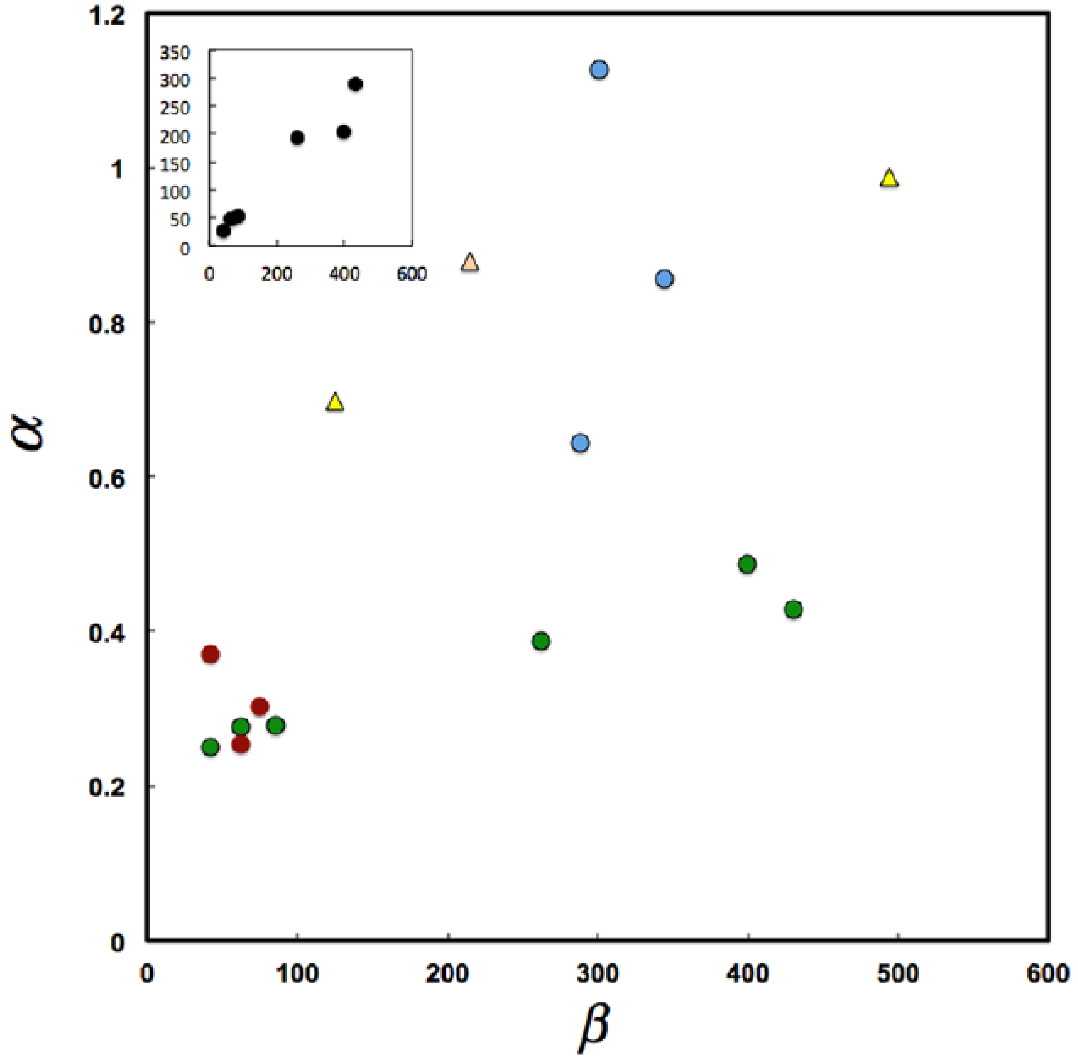


**Figure 1.** Fit of the Beta distribution to different animal and plant communities. First row, from left to right Amazon birds (community 10 in Table 1), Lepidoptera (12 in Table 1), butterflies (11 in Table 1), second row from left to right Tropical trees (6 in Table 1), Tropical trees (2 in Table 1) and Coral reefs (14 in Table 1)

$$\rho_\infty(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}$$



● Marine communities

● Shrublands

● Tropical Forest

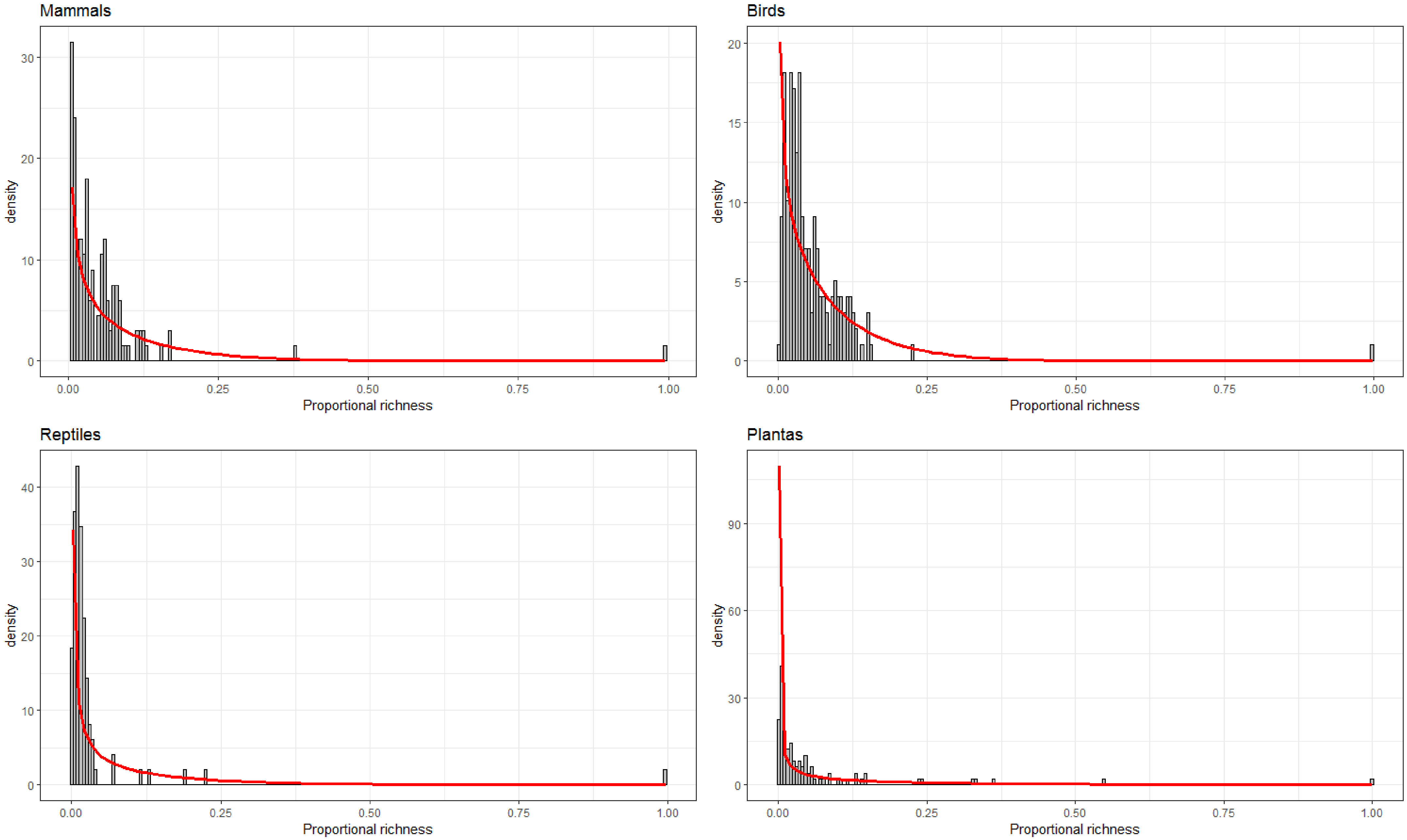
▲ Bird communities

$\alpha \sim$ Dispersal

$\beta \sim$ Speciation

# Some Sources of Stochasticity at Different Levels of Organization

| Frequency of Gene families within Genomes | Frequency of genes within and among population | Frequency of individuals among species | Frequency of species Among communities |
|---|---|---|---|
| • Mutation<br>• Lateral gene transfer<br>• Gene duplication | • Mutation<br>• Drift<br>• Migration<br>• Loss | • Birth<br>• Death<br>• Migration | • Speciation<br>• Extinction<br>• Migration |


Mammals
density
30
20
10
0
0.00
0.25
0.50
0.75
1.00
Proportional richness
Birds
density
20
15
10
5
0
0.00
0.25
0.50
0.75
1.00
Proportional richness
Reptiles
density
40
30
20
10
0
0.00
0.25
0.50
0.75
1.00
Proportional richness
Plantas
density
90
60
30
0
0.00
0.25
0.50
0.75
1.00
Proportional richness

# Abundance of gene families within genomes

- Birth-death processes are applicable to understand genome evolution

- Driving processes are similar

Duplication

Mutation

Lateral gene transfer and retroviral infection

Plasmids

Gene extinction

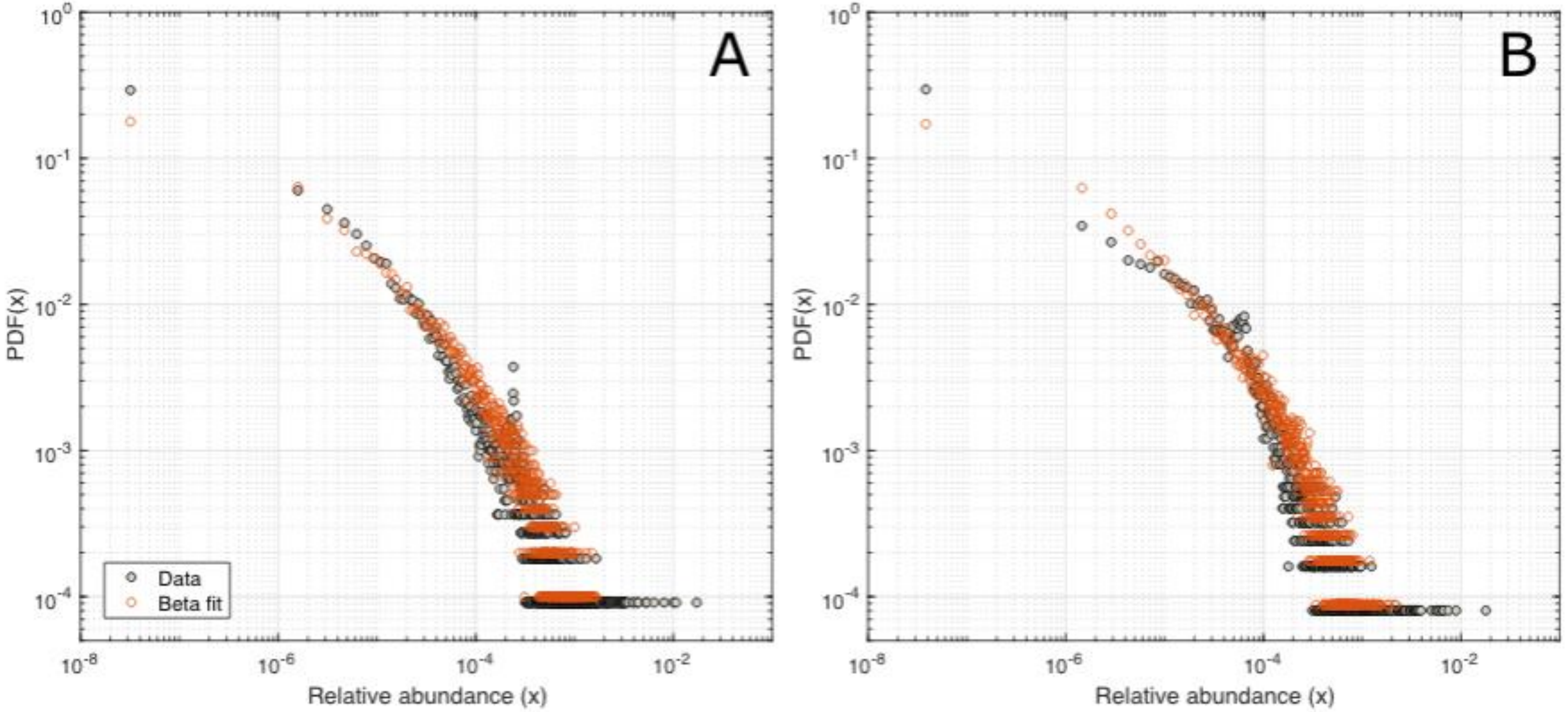


**Figure 1**. The relative abundance of protein families follows a Beta distribution. The relative abundance of protein families in bacteria and eukarya according to the Pfam classification. **A**. Data from 7,694 bacterial proteomes; 11,106 families; and 30,743,438 genes. Best fit beta distribution parameters: $\alpha$=0.273 (0.264, 0.283), $\beta$=3,031.9 (2823.4, 3240.4). **B**. Data from 1,496 eukaryotic proteomes; 12,579 families; and 25,625,956 genes. Best fit beta distribution parameters: $\alpha$=0.291 (0.281, 0.300), $\beta$=3,655.2 (3424.3, 3886.2).

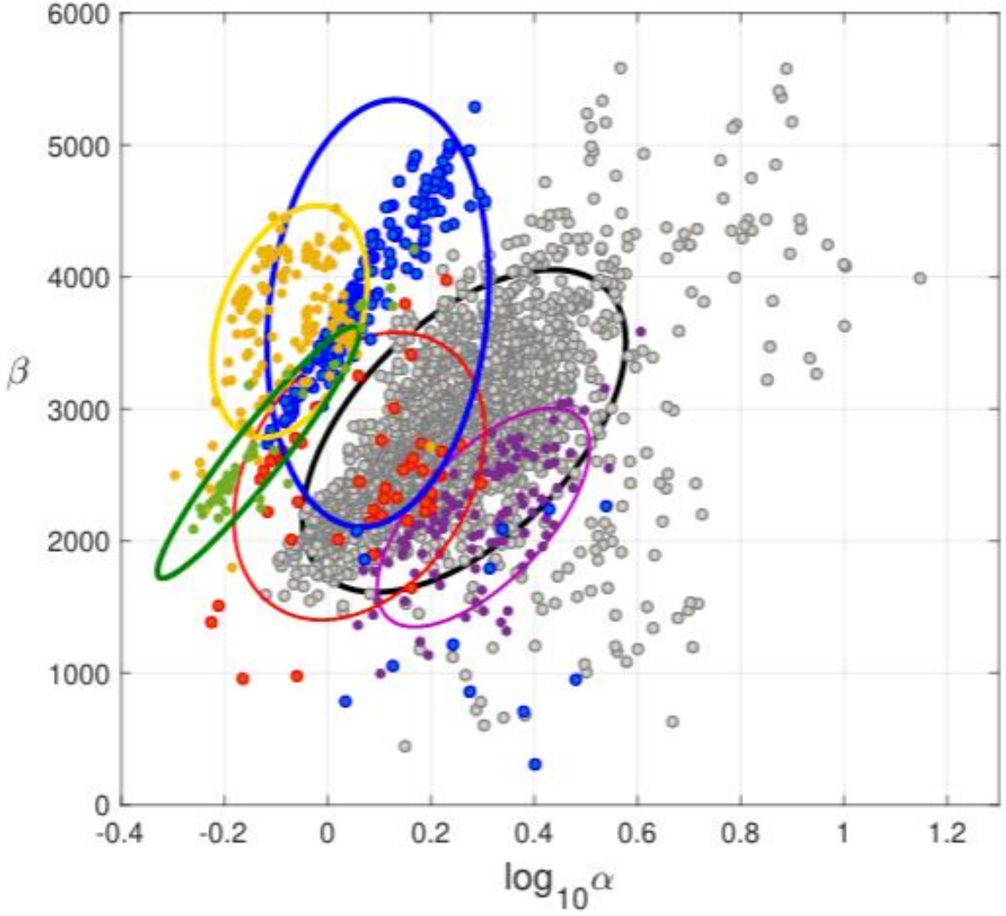


**Figure 2**. Shape parameters of Beta distributions arising from the relative abundance of gene families in genomes. Annotations of protein families (Pfam v32) in Uniprot reference proteomes (n=17,543). Reference proteomes include Archea (n=285, magenta), Bacteria (n=6,554, grey); unicellular eukarya (n=1,202, red); plants (n=, green); metazoa (n=, yellow); fungi (n=, blue). The total number of protein families classified by Pfam is 15,964.

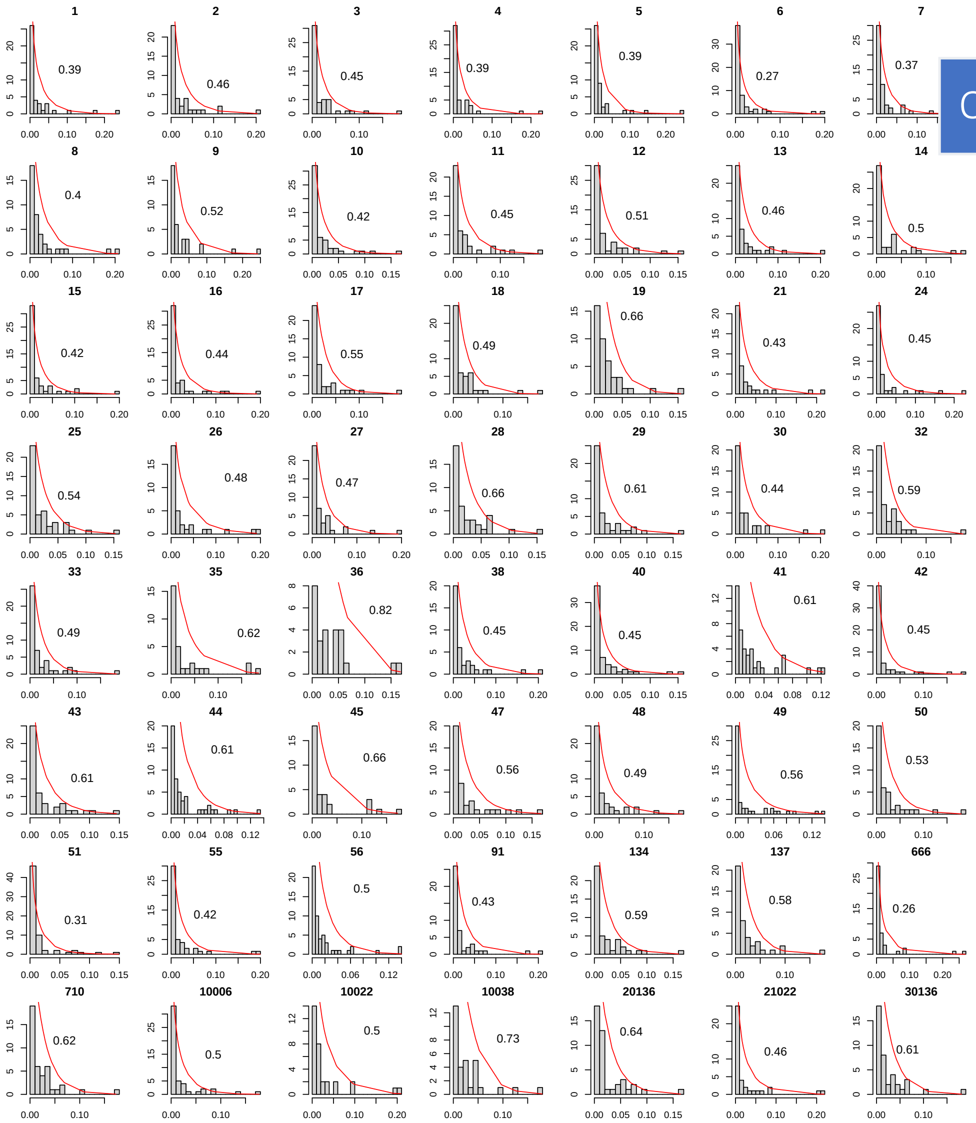
# Empirical data

123 soecies, 21 years

Wide changes in area, heterogeneity, productivity, connectivity, diversity…

From: Matías Arim

# The BeTa distribution provide a good fit to the relative abundance within each community



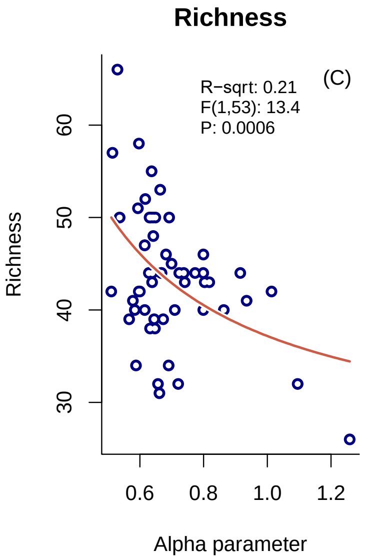
OK Beta

$(0.26 < R^2 < 0.82)$

Arim et al. (ms)

# The parameters $\alpha$ y $\beta$

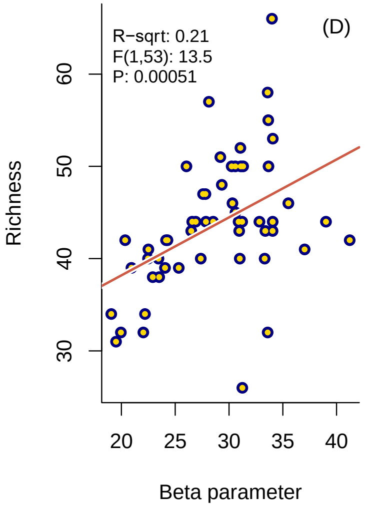
$$\rho_\infty(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$



$\alpha$ Reflects:

Performance of resident species. Favors dominance

$$\langle S \rangle = 1 + \frac{\beta}{\alpha}$$



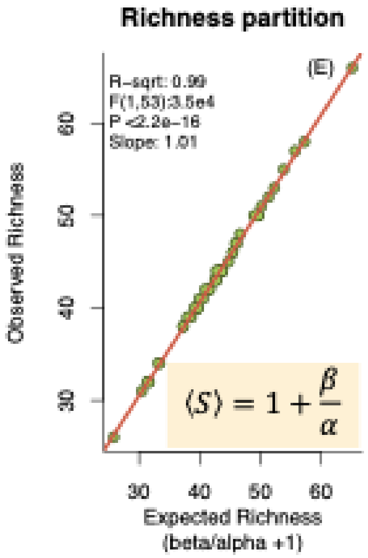
$\beta$ Reflects:

· Dispersal of non residents. Favors coexistence.

$$\langle x \rangle = \frac{\alpha}{(\beta+\alpha)}$$

Arim et al. (ms)

# Expected Richness $S = 1 + \frac{\beta}{\alpha}$



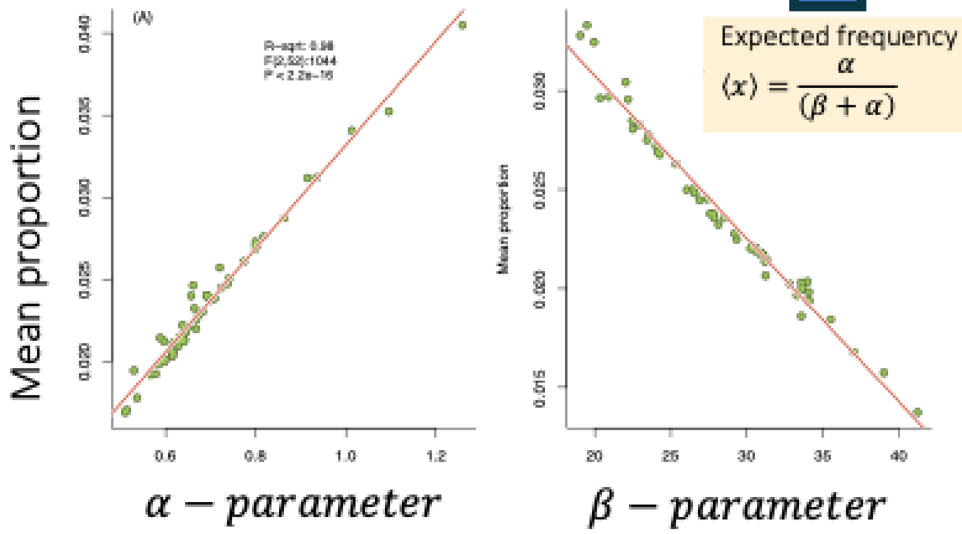
$$\langle S \rangle = 1 + \frac{\beta}{\alpha}$$

The Beta distribution breaks down observed species richness into the ratio of the intensity of processes that promote ($\beta$) and limit ($\alpha$) diversity.

The same applies to the expected frequency, which is a function of the ratio between the intensity of the processes that promote ($\alpha$) or limit ($\beta$) species dominance in communities.

Expected frequency

$$\langle x \rangle = \frac{\alpha}{(\beta + \alpha)}$$

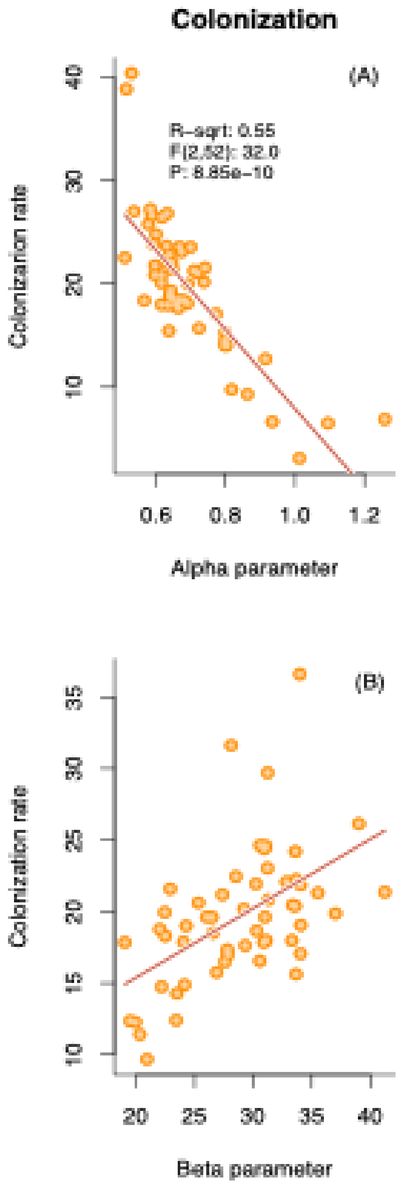


Arim et al. (ms)

$\alpha$ :
Reflects performance of resident species.

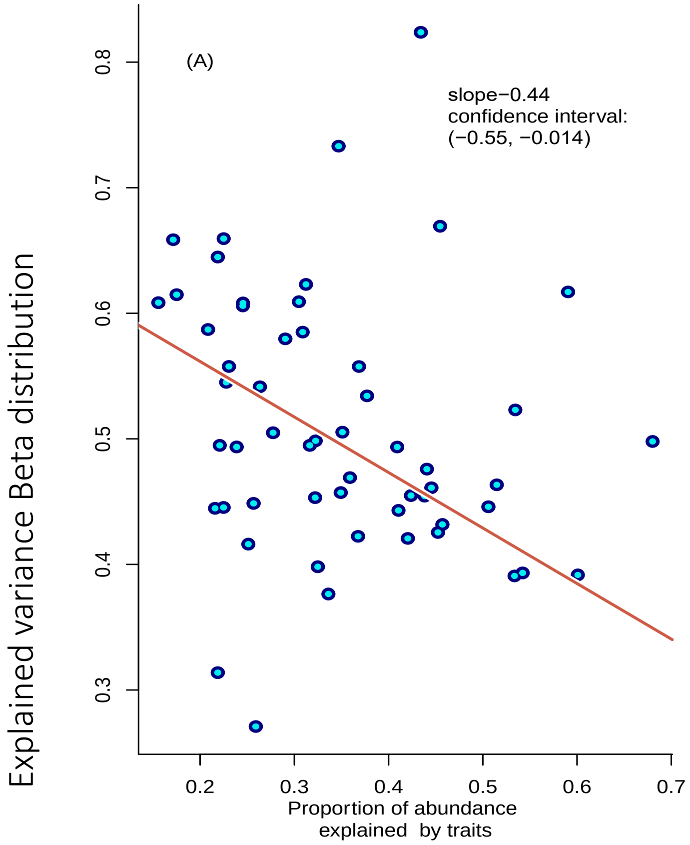
Determines low colonization

$\beta$ :
Performance of non-residents
Increase colonization



Arim et al. (ms)

# Degrees of neutrality….



Arim et al. (ms)

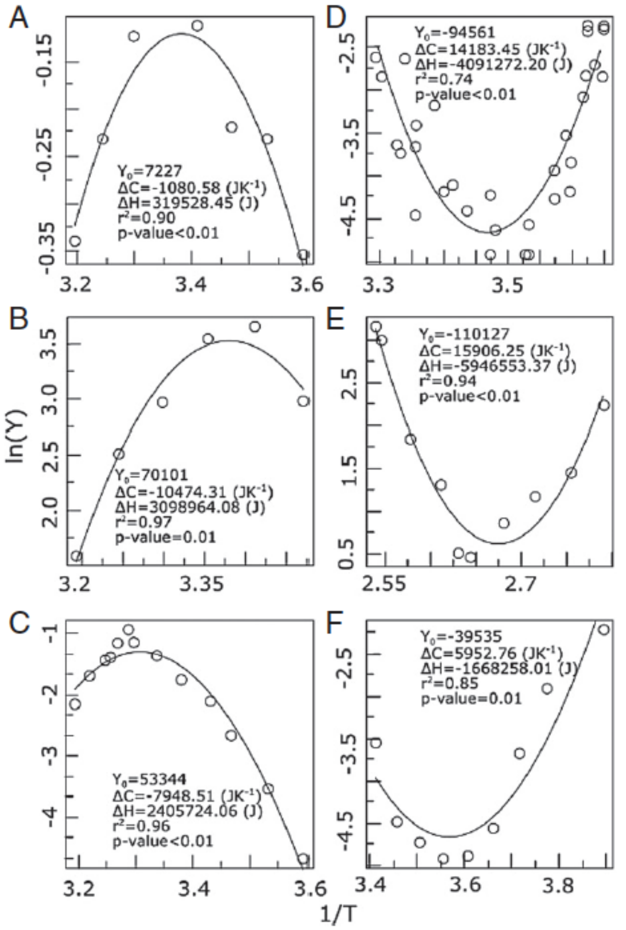
Special thanks to:

Rolando Rebolledo
Mauricio Tejo
Evandro Ferrada
Octavio Reyes
Simón Castillo
Ignacio Arroyo
Lucas Gallart

- Thanks
- Obrigado
- Gracias
- *Ckantatur*
- Chaltu may
- Yuspagara

# Integration

## Across disciplines and levels of organization



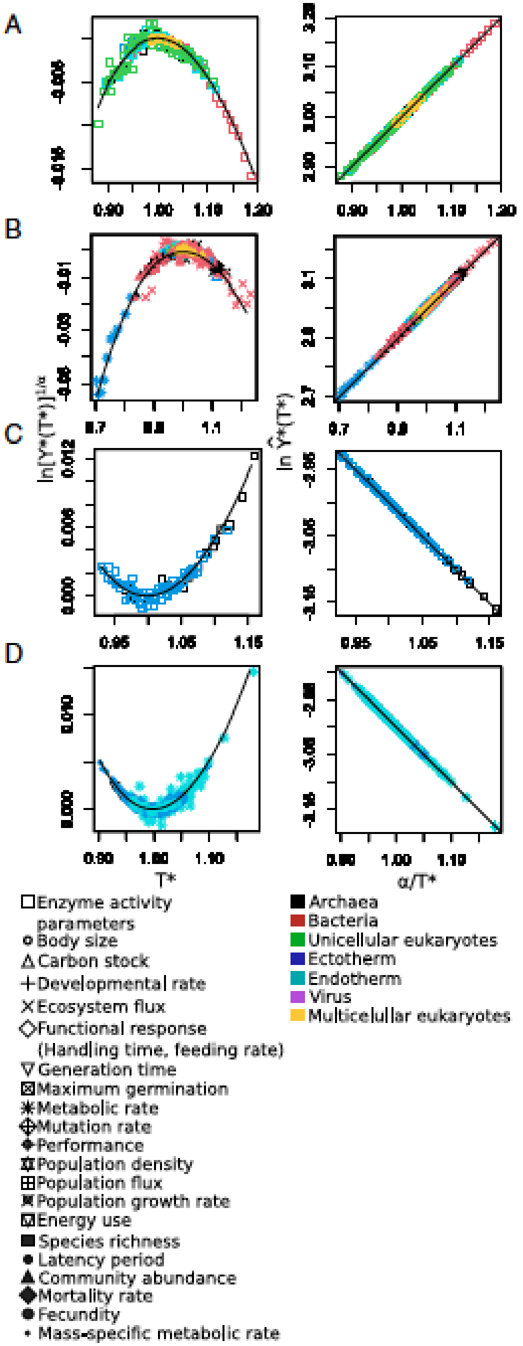
The Eyring–Evans–Polanyi (EEP) transition state theory (TST)

$$k = \frac{k_B}{h} T e^{\Delta S/R} e^{-\Delta H/RT}.$$

**PNAS** RESEARCH ARTICLE | ECOLOGY

# A general theory for temperature dependence in biology

José Ignacio Arroyo[a,b,1,2], Beatriz Díez[c,d,e], Christopher P. Kempes[b], Geoffrey B. West[b], and Pablo A. Marquet[a,b,f,g,h,l,2]



$$Y(T) \approx Y_0 \left(\frac{1}{T}\right)^{\frac{-\overline{\Delta C}}{R} - \alpha} e^{\frac{-\overline{\Delta H}}{RT}}.$$

$$Y^{*1/a} = T^* e^{1/T^* - 1}$$

$$\hat{Y}^*(T^*) \equiv (e/T^*)^a \, Y^*(T^*) = e^{a/T^*}$$

Arroyo et al. (2022) PNAS

- Integration among theories is a first step towards a more logically consistent ecological theory

- The theory of island biogeography, given its quantitative character provided a foundational model to foster integration in ecology (although the was wrong!)

- Genes, individuals and species seem to obey the same probability law.