Macroecological laws and emerging biological patterns

A conceptual approach

Emanuele Pigani – PhD student SZN
Naples and LIPh Lab Padova (Italy)

Towards a modern analysis of omics data of the Ocean - workshop

Valparaiso (Chile)
May 18th, 2023

Macroecological laws on the seascape

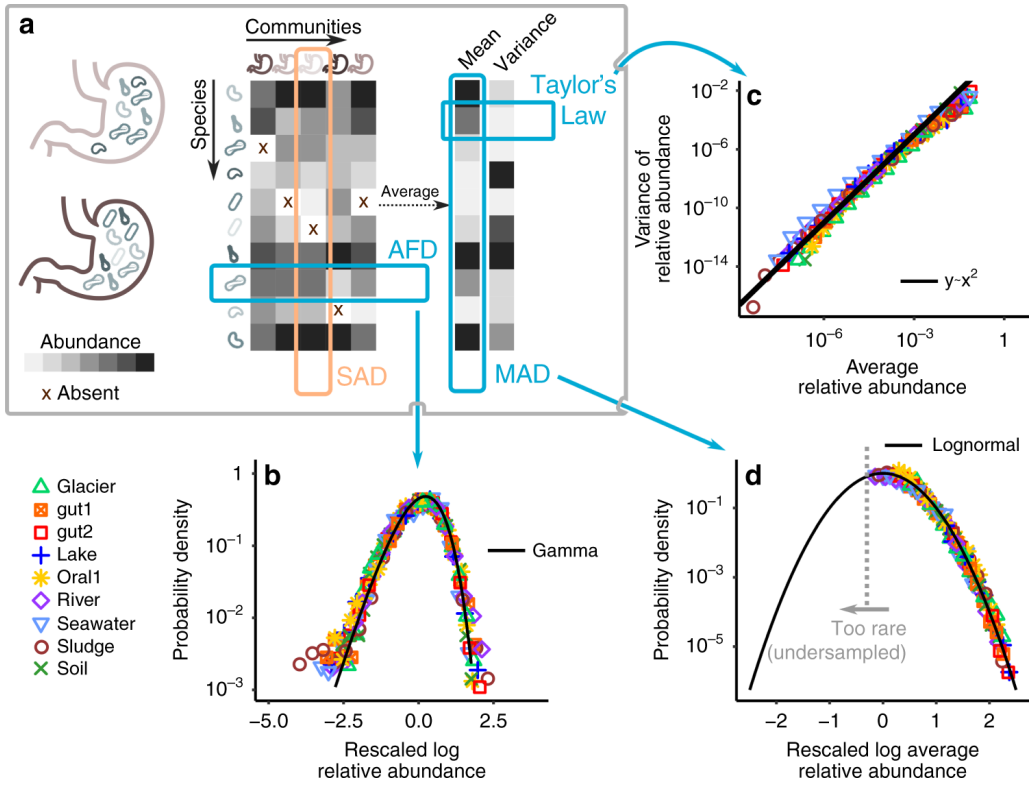
ARTICLE

https://doi.org/10.1038/s41467-020-18529-y OPEN

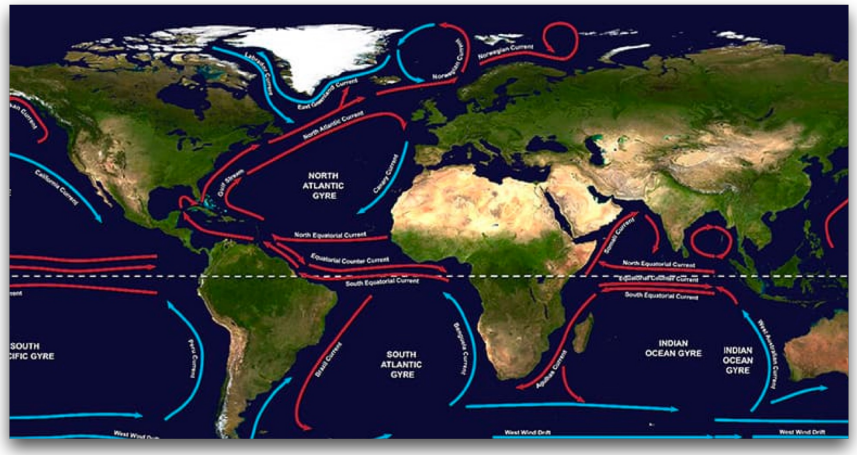
Macroecological laws describe variation and diversity in microbial communities

Jacopo Grilli [1,2]✉

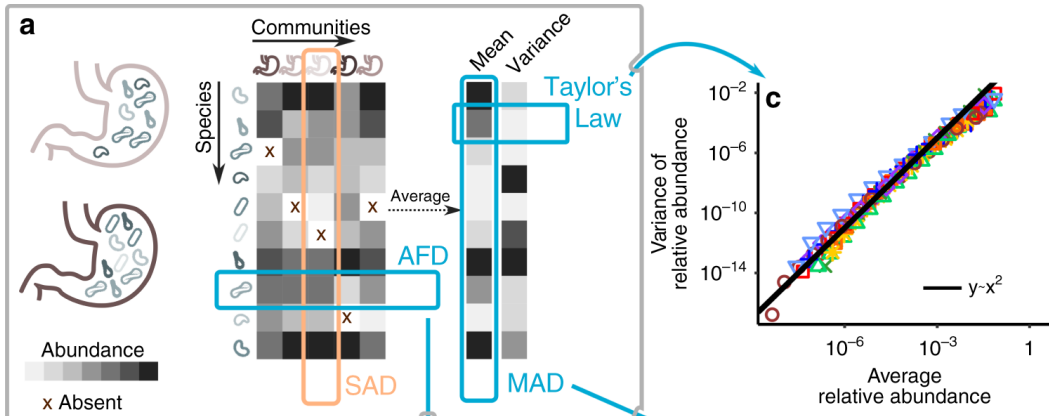
nature COMMUNICATIONS

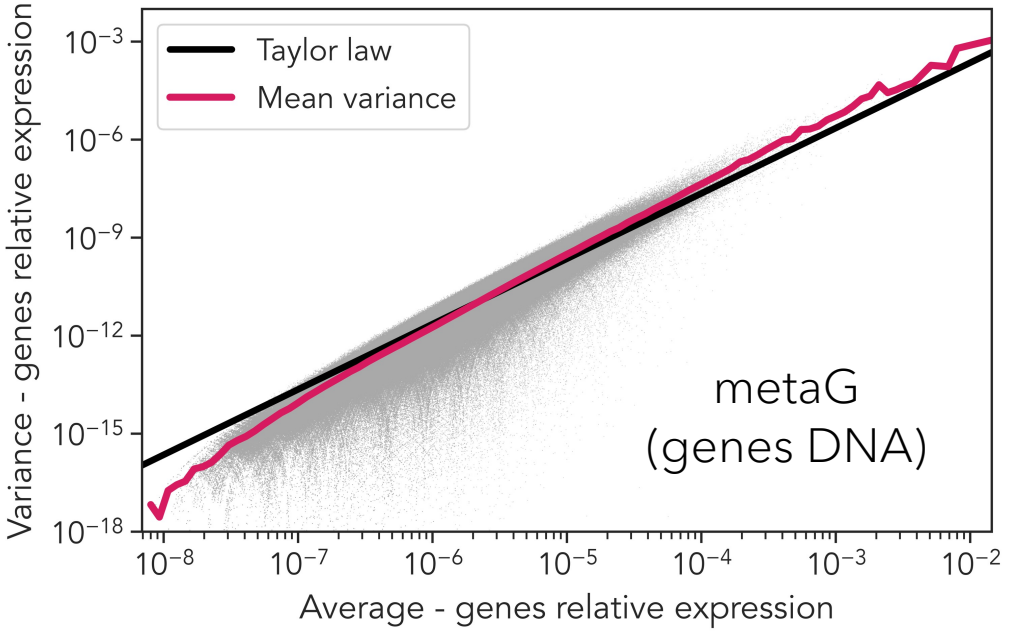


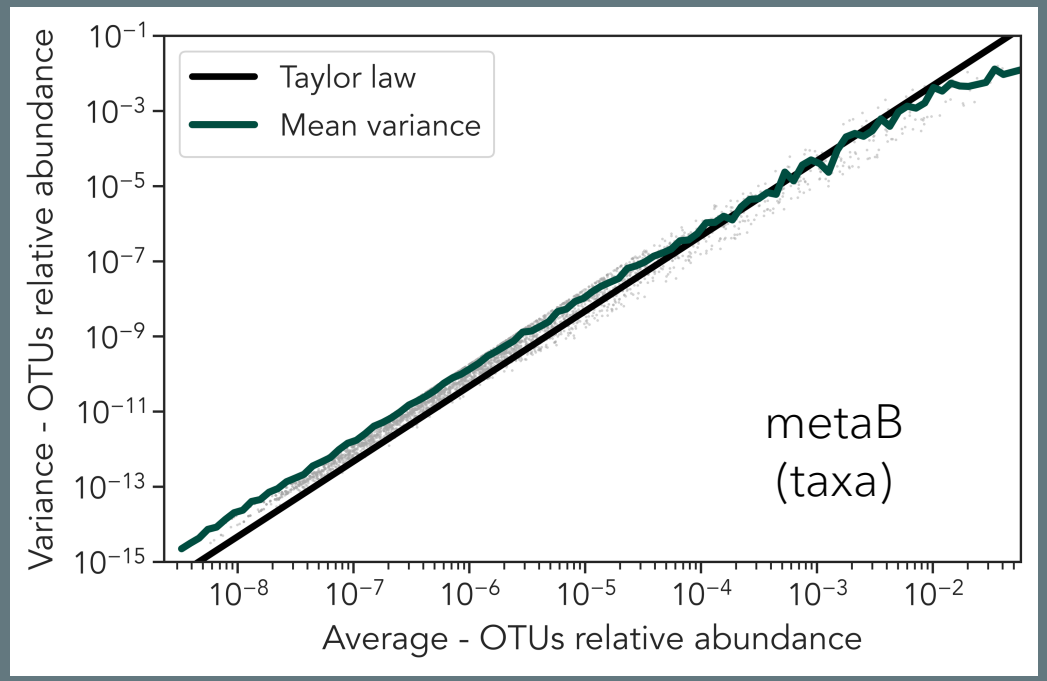
What is the **taxonomical** and **functional impact** of the **seascape** to the communities at **macroecological** level?

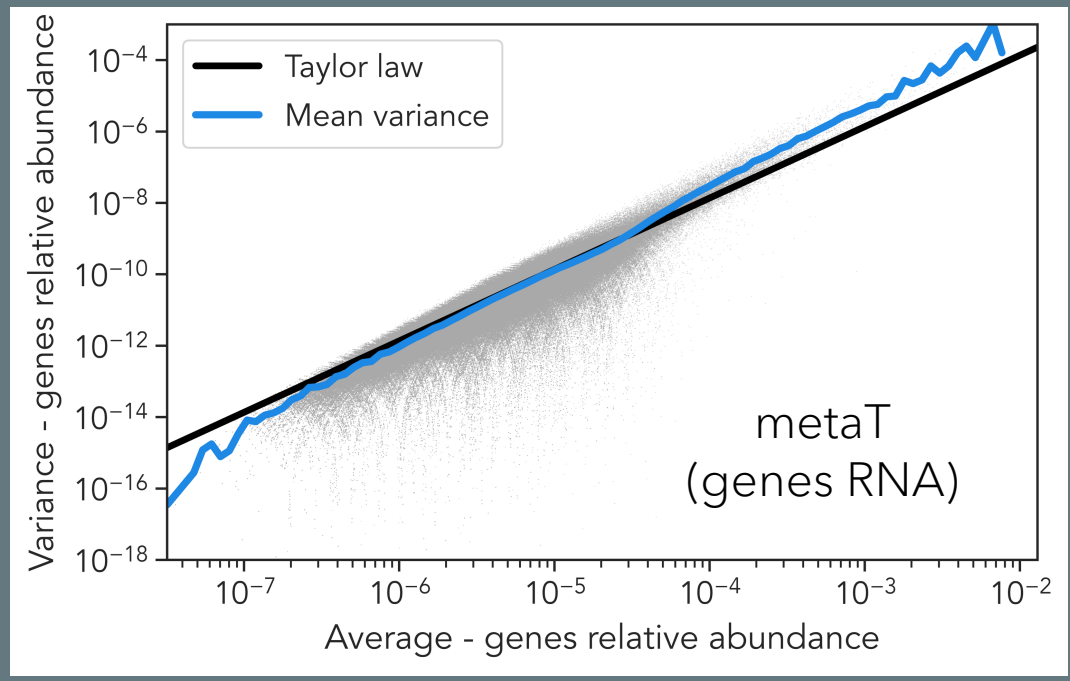


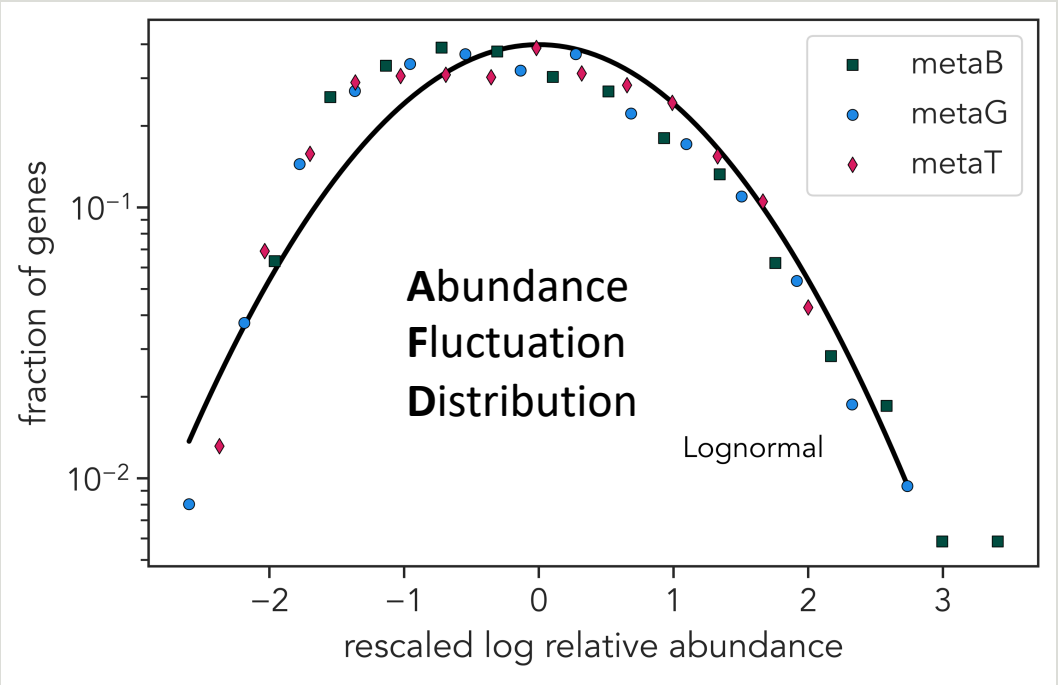
Taylor's law for diatoms meta-omics

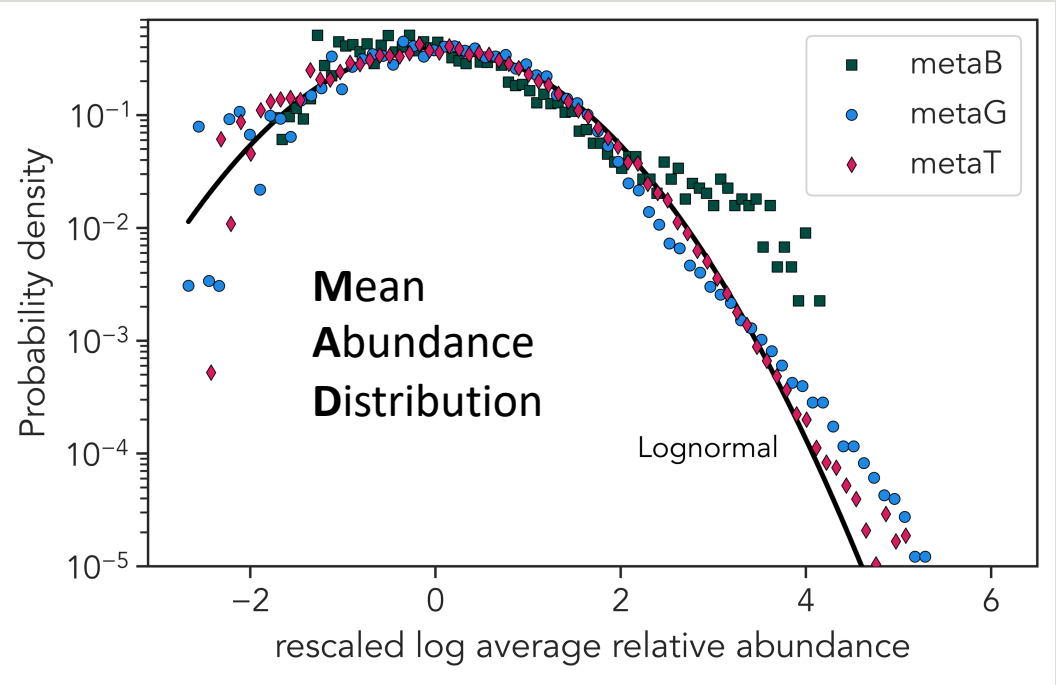




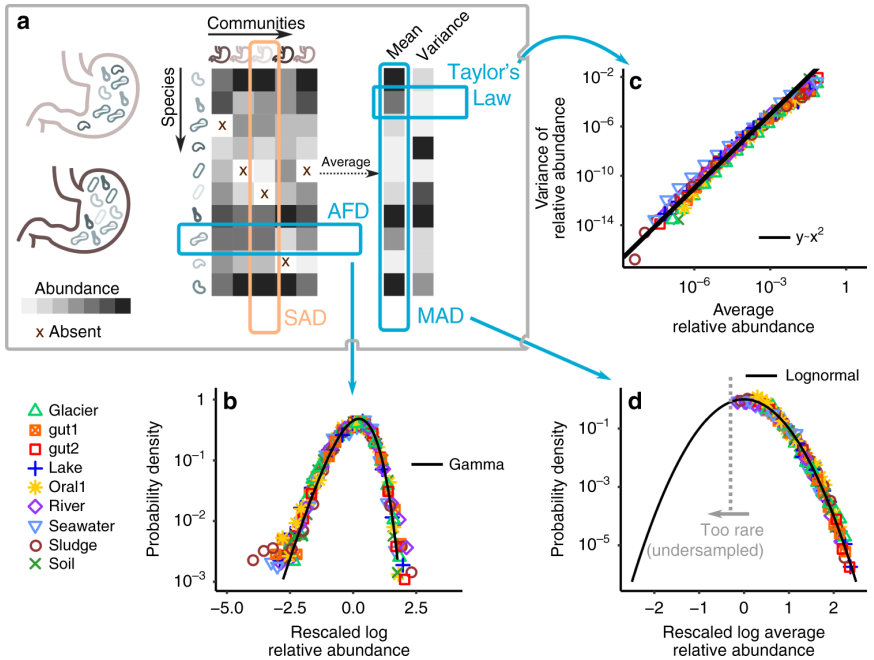




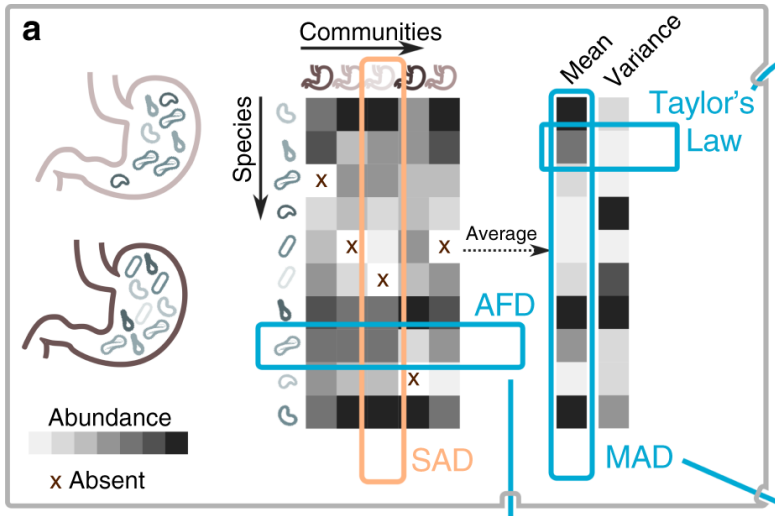

metaB
metaG
metaT
Abundance
Fluctuation
Distribution
Lognormal
fraction of genes
rescaled log relative abundance


metaB
metaG
metaT
Mean
Abundance
Distribution
Lognormal
Probability density
rescaled log average relative abundance

AFD & MAD


a
Communities
Species
Mean
Variance
Taylor's Law
Average
AFD
SAD
MAD
Abundance
x Absent
b
Gamma
Probability density
Rescaled log relative abundance
c
Variance of relative abundance
Average relative abundance
y~x²
d
Lognormal
Too rare (undersampled)
Rescaled log average relative abundance
Glacier
gut1
gut2
Lake
Oral1
River
Seawater
Sludge
Soil

SAD & RAD

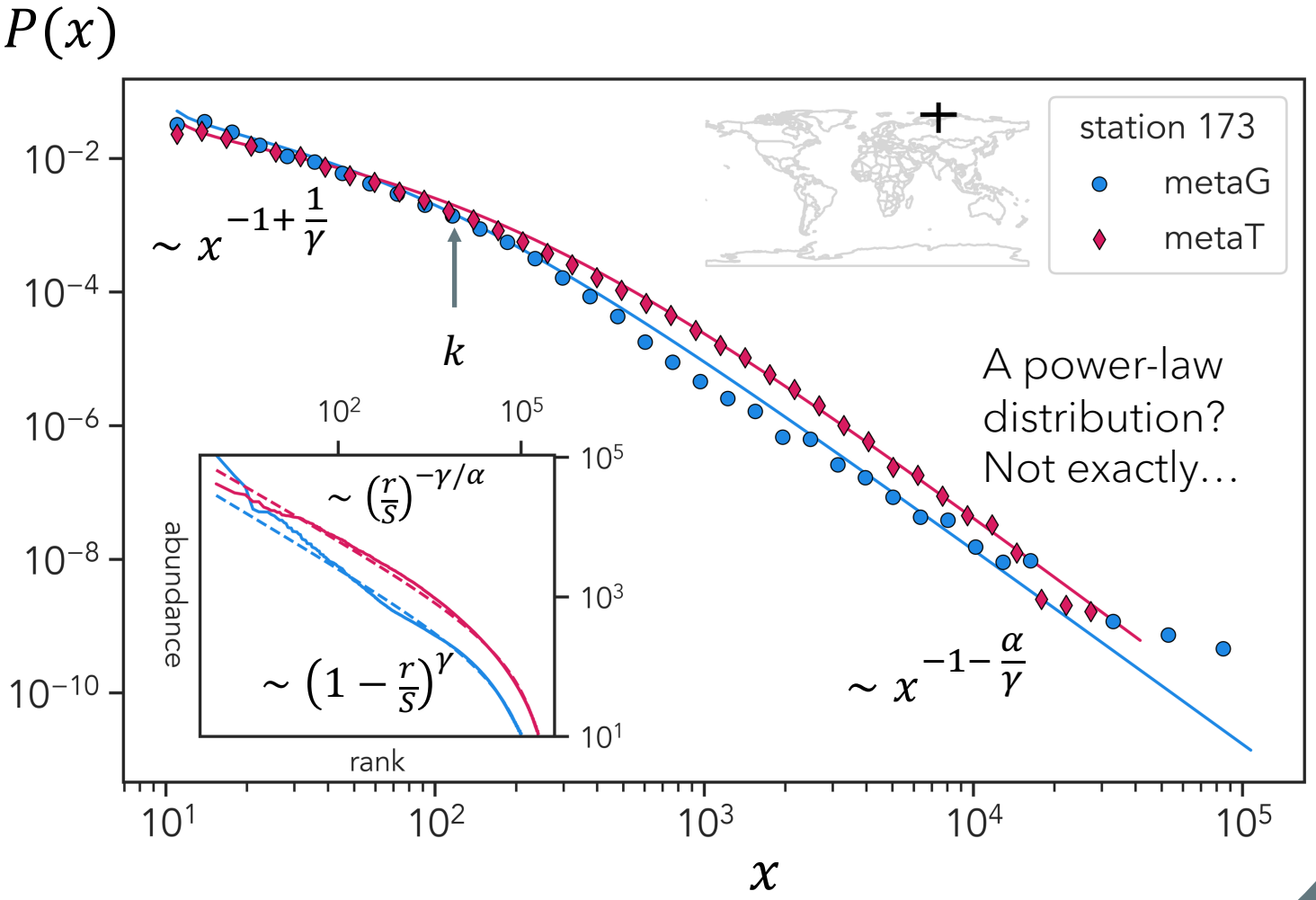


Properties:

- Two power-law regimes separated by a scale k;
- Simple power-law distribution (Pareto) in the limit $\gamma = 1$ and $\mu = 0$.

Pareto IV distribution:

$$P(x|k,\alpha,\gamma,\mu) = \frac{\alpha}{\gamma k}\left[1+\left(\frac{x-\mu}{k}\right)^{1/\gamma}\right]^{-1-\alpha}\left(\frac{x-\mu}{k}\right)^{-1+1/\gamma}$$



- Analytical formula for the Rank Abundance distribution: $RAD(r) = \mu + k\left[\left(\frac{r}{S}\right)^{-1/\alpha} - 1\right]^{\gamma}$

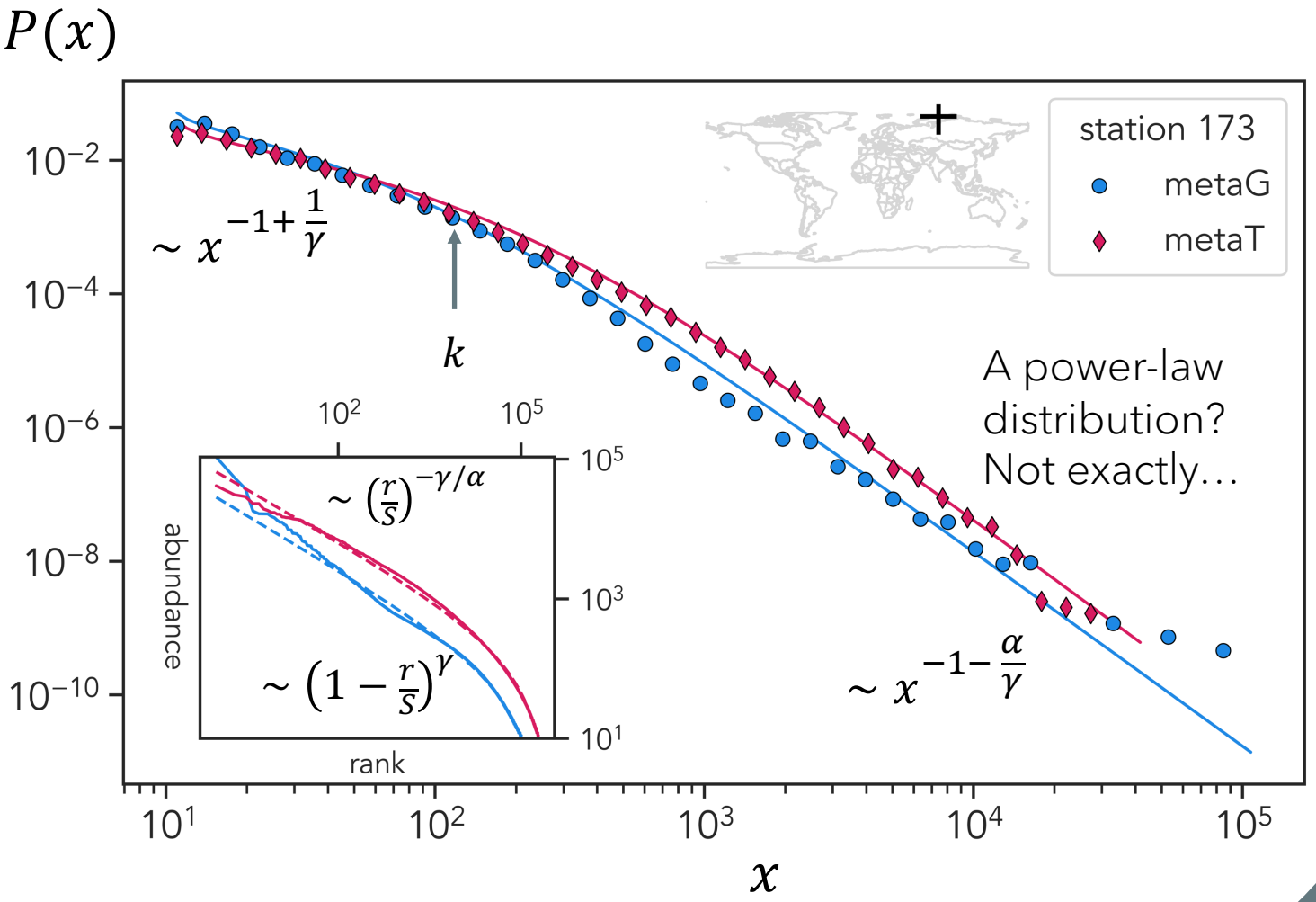
SAD & RAD

Open questions:

- Is there a **relation** between **metaG** and **metaT**?
- Why the Pareto IV? Is there a **generative model**?
- What is the **biological** meaning of the **parameters**?
- Does it work also for **transcriptomics** data?
- Do we have to take it into account in the **data normalization**?

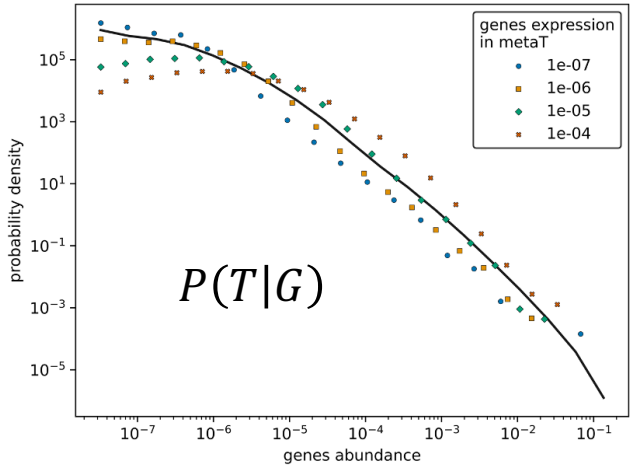
Pareto IV distribution:

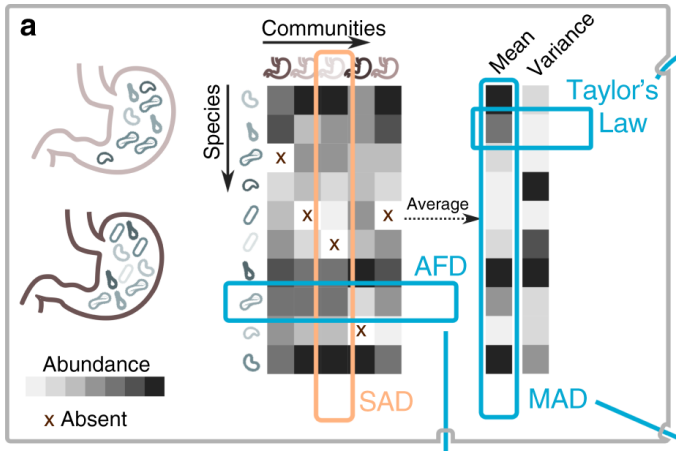
$$P(x|k,\alpha,\gamma,\mu) = \frac{\alpha}{\gamma k}\left[1+\left(\frac{x-\mu}{k}\right)^{1/\gamma}\right]^{-1-\alpha}\left(\frac{x-\mu}{k}\right)^{-1+1/\gamma}$$



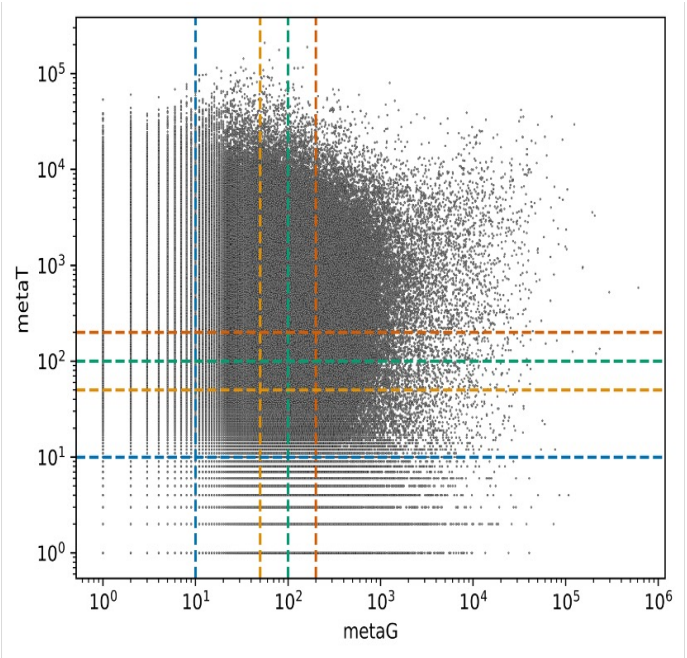
- Analytical formula for the Rank Abundance distribution: $RAD(r) = \mu + k\left[\left(\frac{r}{s}\right)^{-1/\alpha} - 1\right]^{\gamma}$

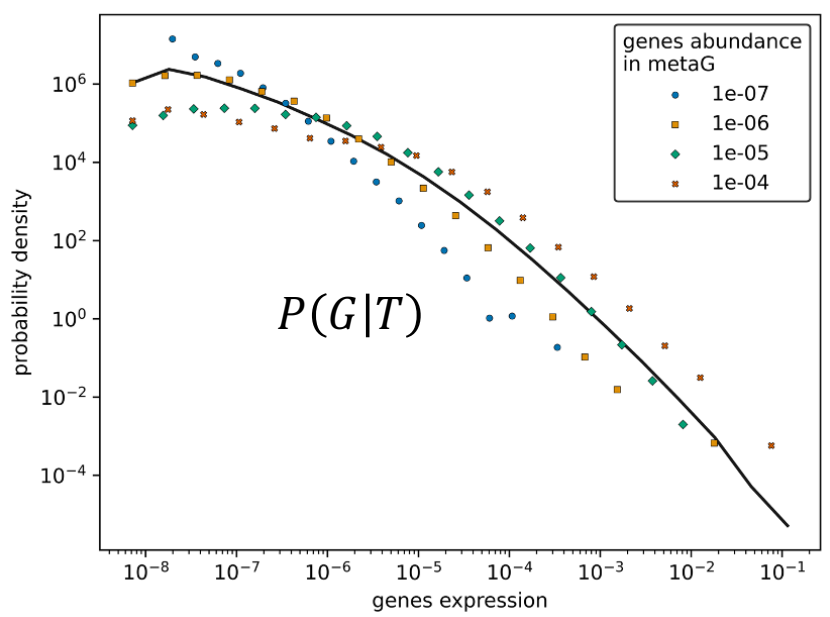
A relation between metaG and metaT?





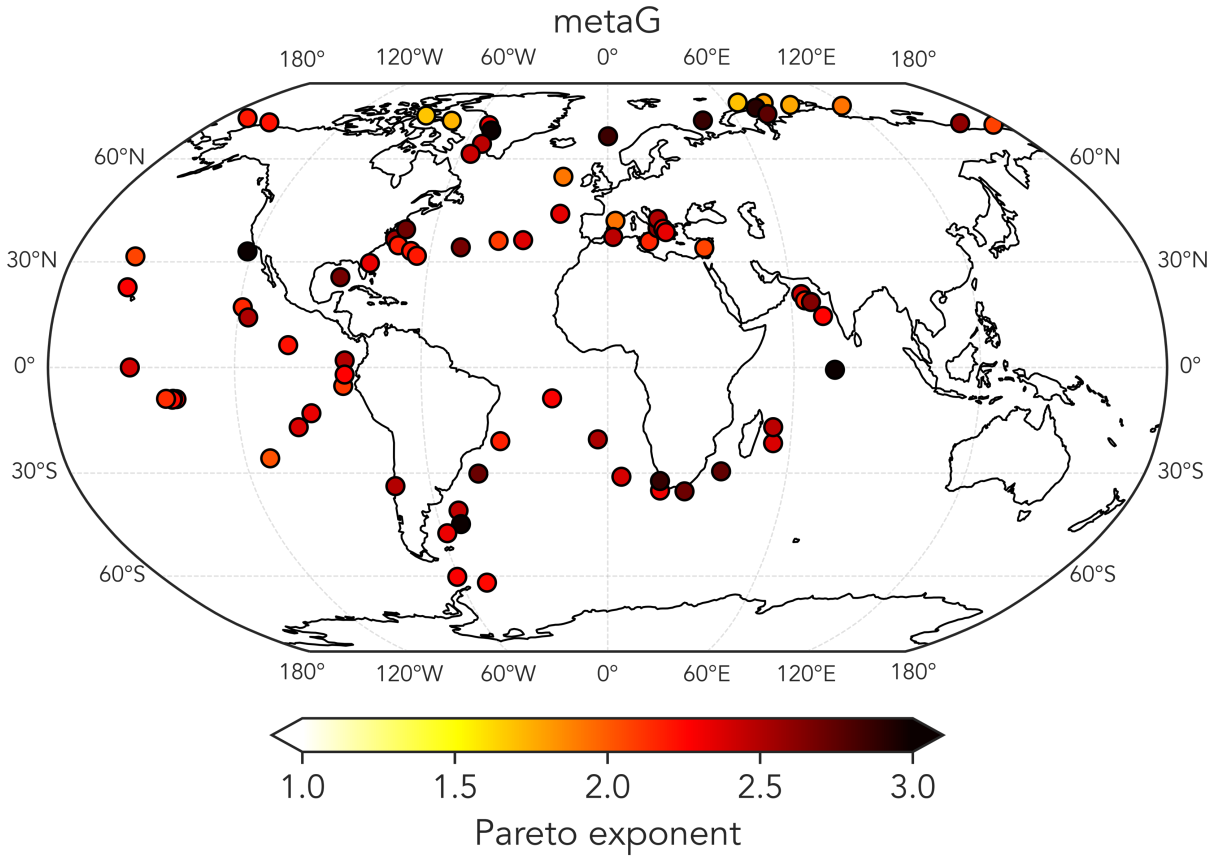
Is a convolution possible?
It seems difficult…

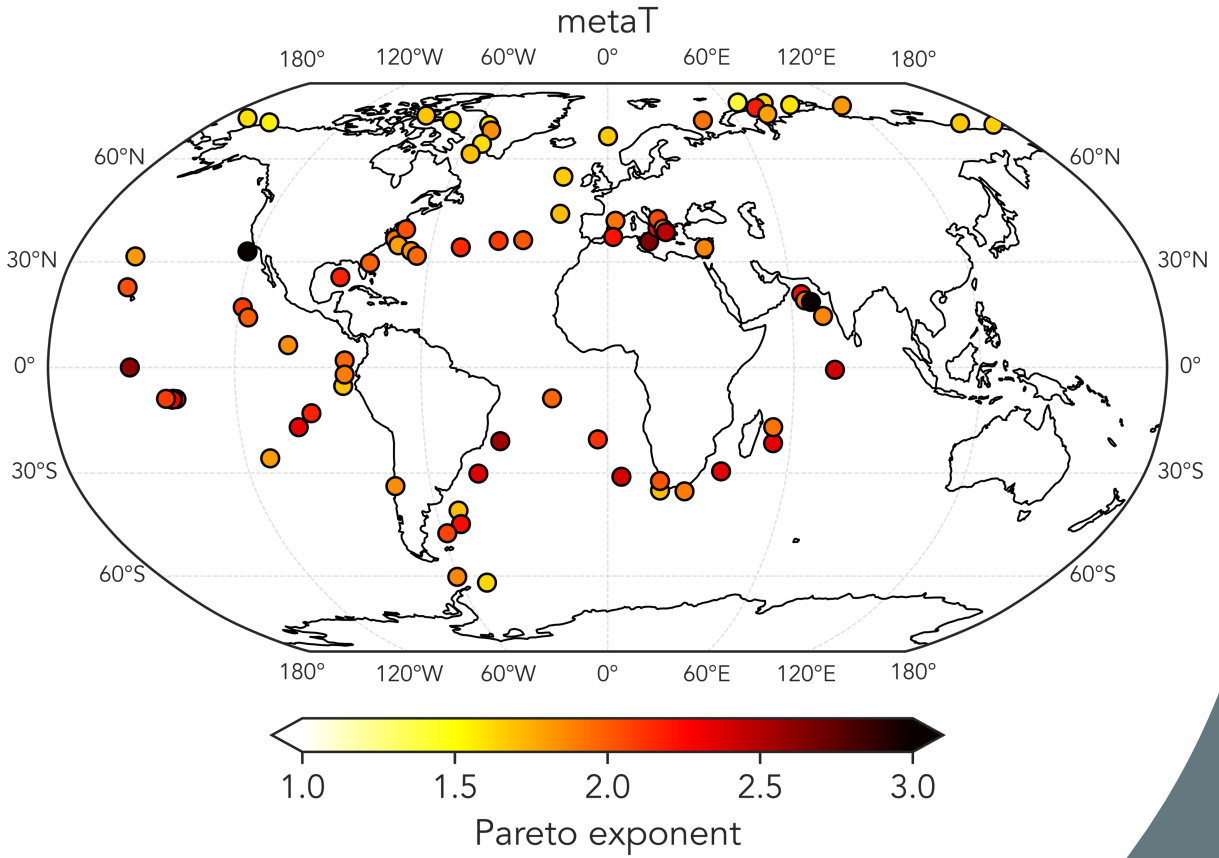




A relation between metaG and metaT?

$$P(x \gg k) \sim x^{-1-\frac{\alpha}{\gamma}}$$





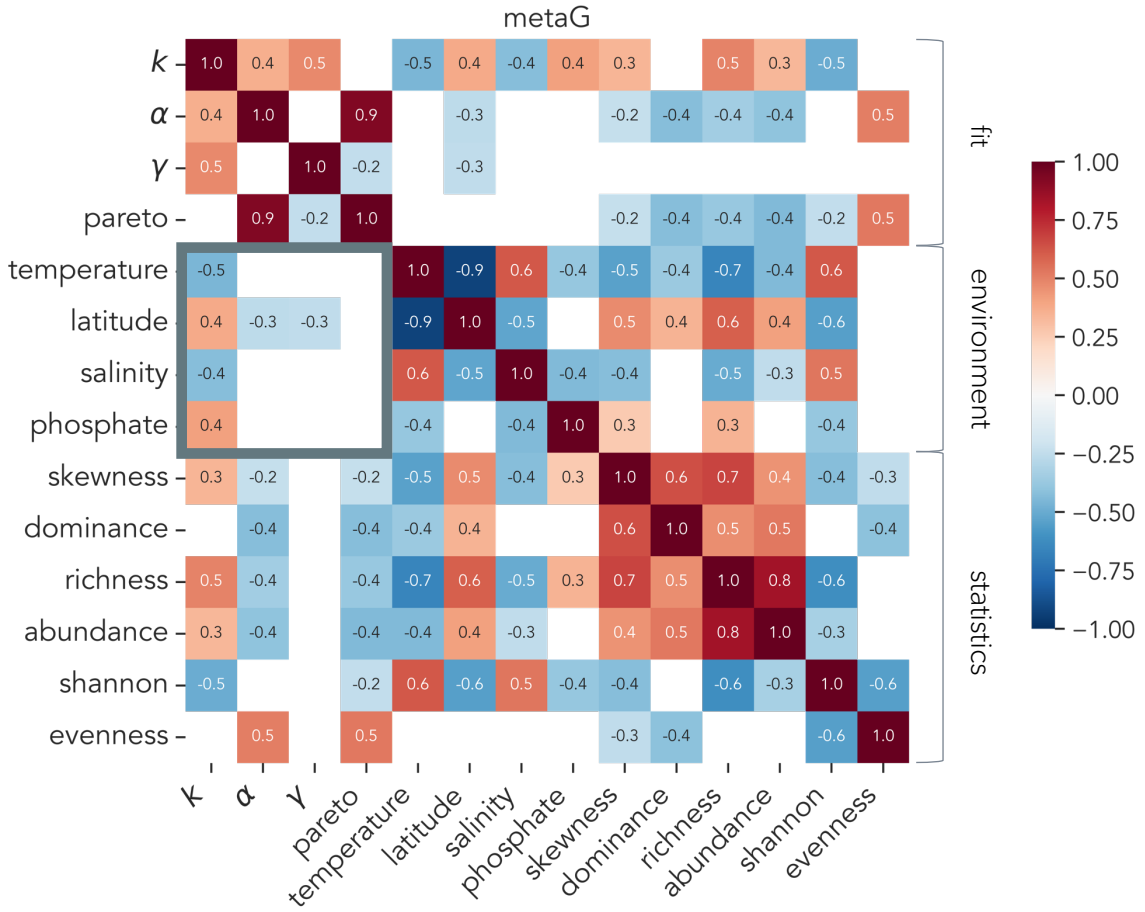
metaT exponents display more **heterogeneity** and **regionality**

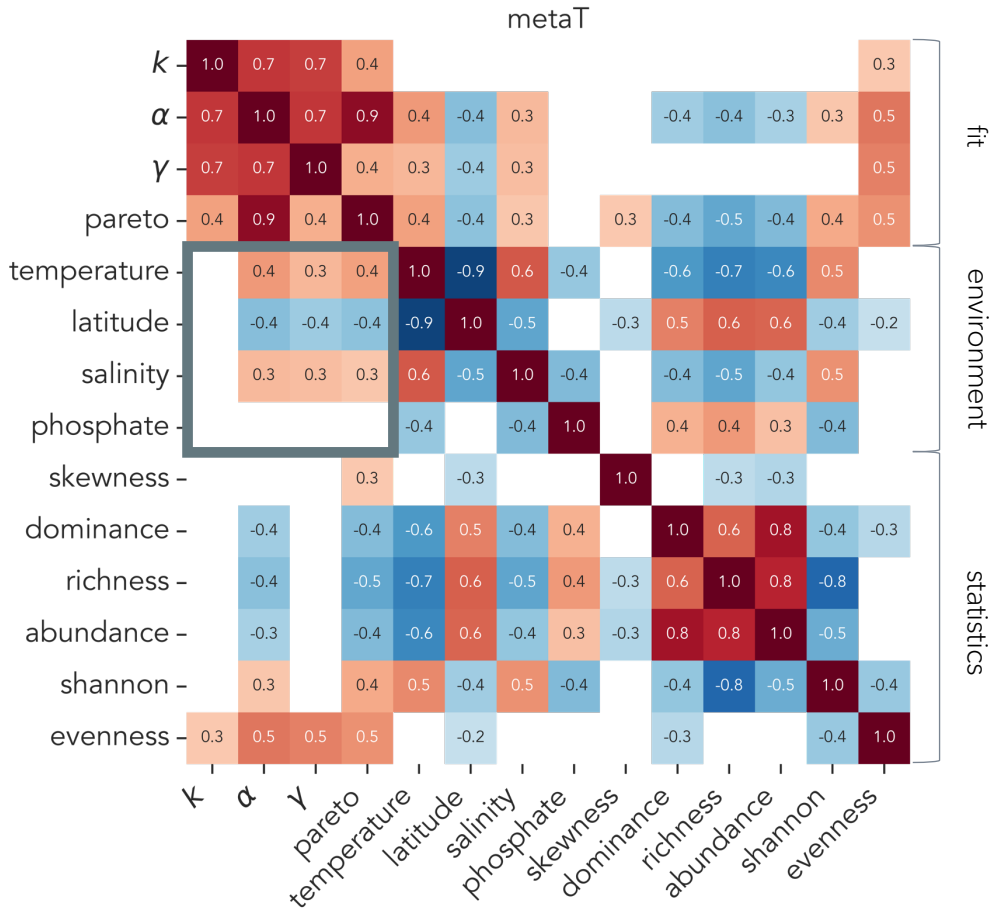
metaG exponents display **higher** and more **homogenous** exponents

A relation between metaG and metaT?

$$P(x \gg k) \sim x^{-1-\frac{\alpha}{\gamma}}$$

$$P(x \ll k) \sim x^{-1+\frac{1}{\gamma}}$$





metaT exponents display more **heterogeneity** and **regionality**

metaG exponents display **higher** and more **homogenous** exponents

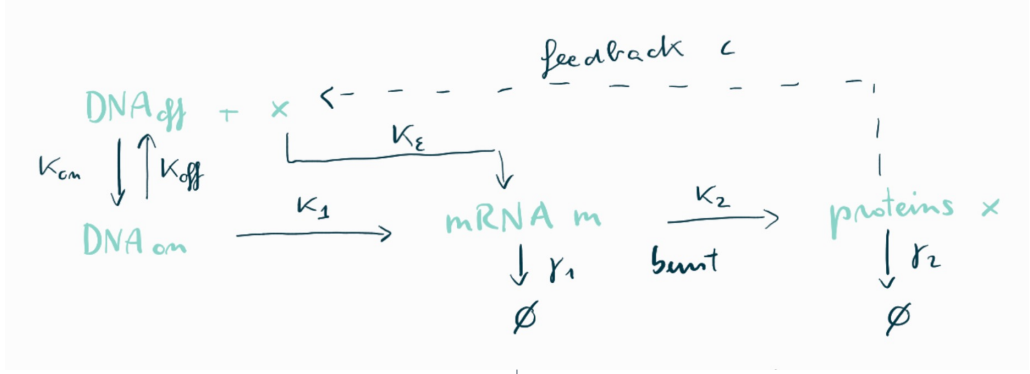
An analytical framework for gene expression?

How to derive the Pareto distribution

PRL 97, 168302 (2006) PHYSICAL REVIEW LETTERS week ending 20 OCTOBER 2006

Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression

Nir Friedman, Long Cai, and X. Sunney Xie
Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford St., Cambridge, Massachusetts 02138, USA
(Received 9 June 2006; published 19 October 2006)



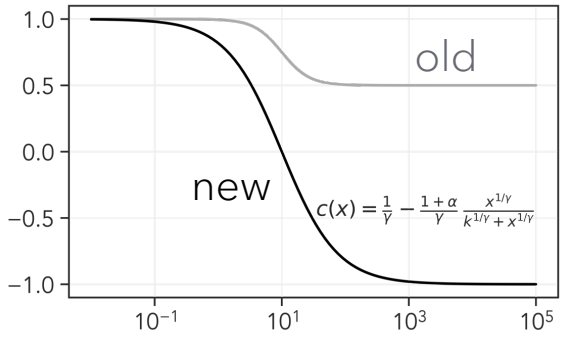
From the Master Equation

assuming **burst** and **feedback** (Hill function)

$$p(x) = Ax^{a(1+\varepsilon)-1}e^{-x/b}[1 + (x/k)^H]^{-a/H}.$$

Removing bursts + slightly **different feedback**: Pareto IV distribution!

$$P(x) = \frac{\alpha}{\gamma k}\left[1 + \left(\frac{x}{k}\right)^{1/\gamma}\right]^{-1-\alpha}\left(\frac{x}{k}\right)^{-1+1/\gamma}$$

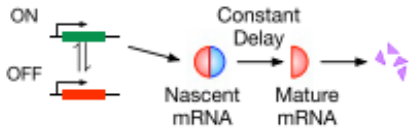


A **biological explanation** for the Pareto exponents…

But… we impose the feedback to be a negative function + the model is for proteins, not RNA… still working on it!

Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells

Zhixing Cao[a,b] and Ramon Grima[b,1]



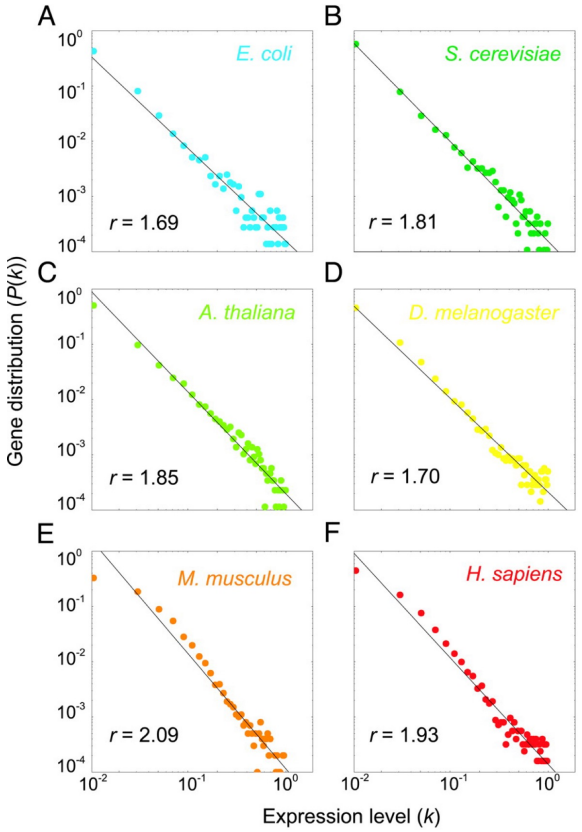
"Universal" distribution for gene expression?

RESEARCH ARTICLE | BIOCHEMISTRY |

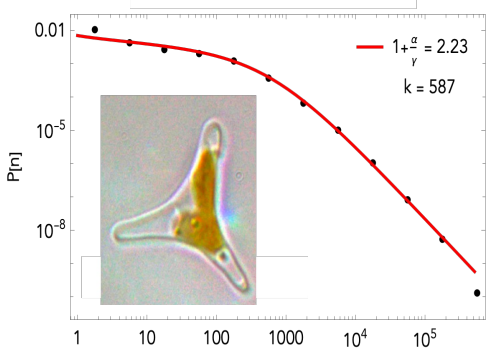
Universality and flexibility in gene expression from bacteria to human

Hiroki R. Ueda, Satoko Hayashi, Shinichi Matsuyama, Tetsuya Yomo, Seiichi Hashimoto, Steve A. Kay, John B. Hogenesch, and Masamitsu Iino -4 Authors Info & Affiliations

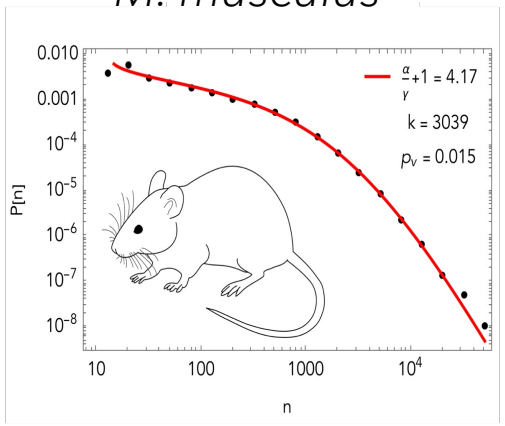
March 3, 2004 | 101 (11) 3765-3769 | https://doi.org/10.1073/pnas.0306244101



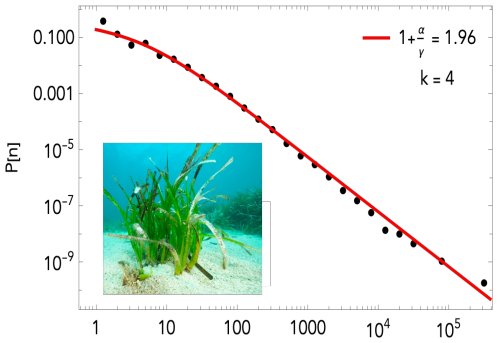
P. tricornutum

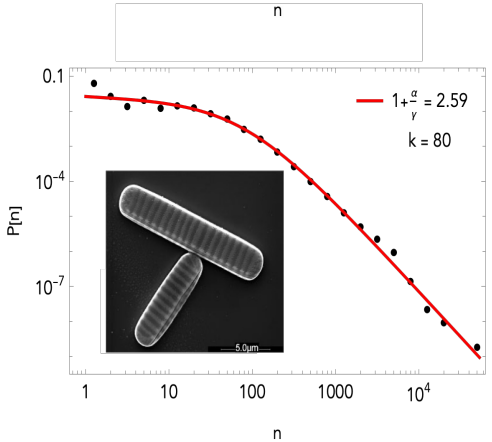


M. musculus

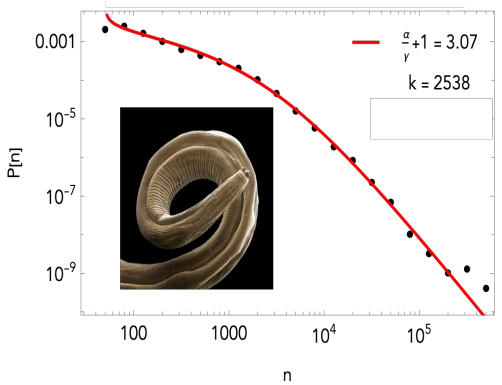


P. oceanica

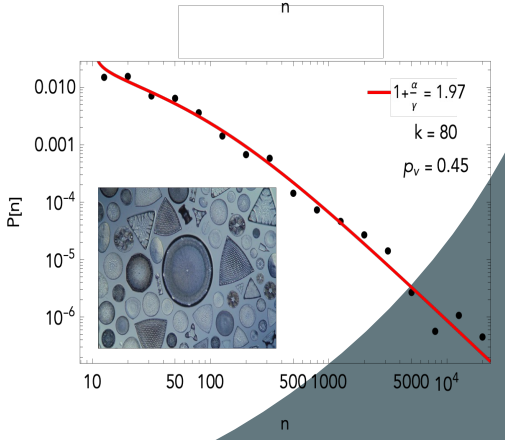




F. cylindrus



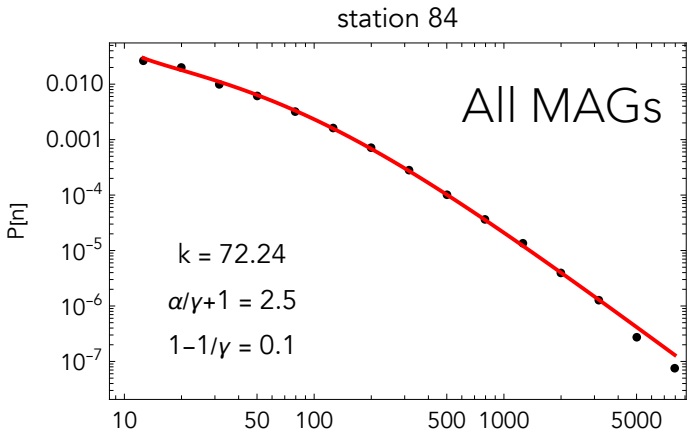
C. elegans

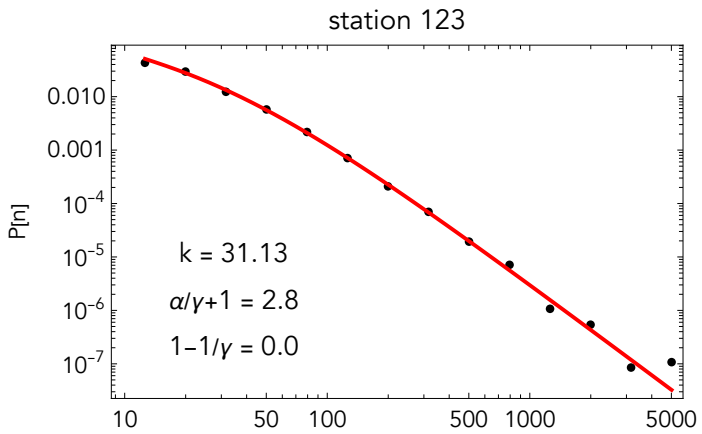


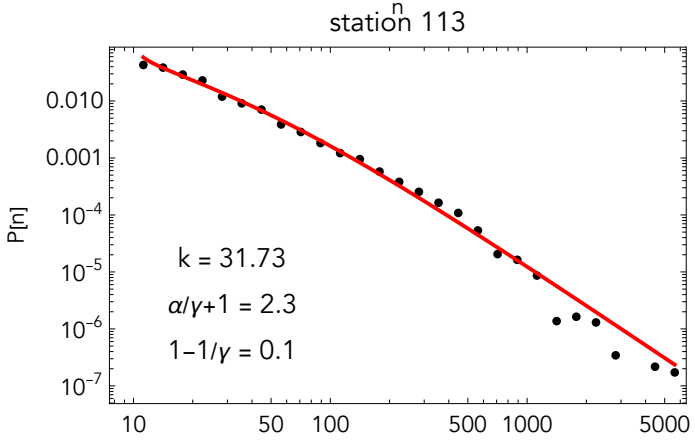
Diatoms

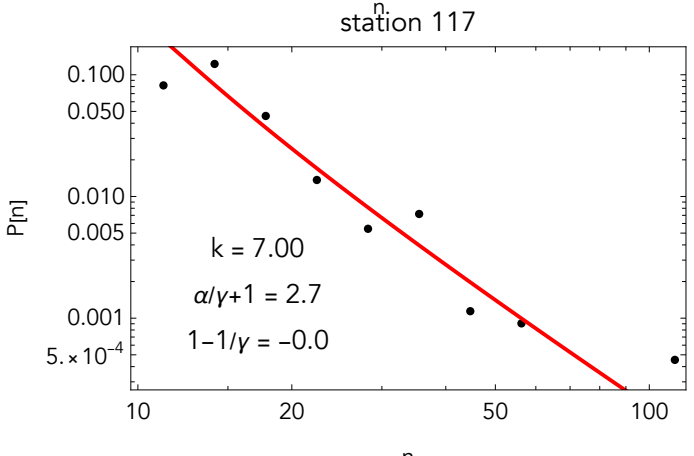
"Universal" distribution for gene expression?

Looking at MAGs

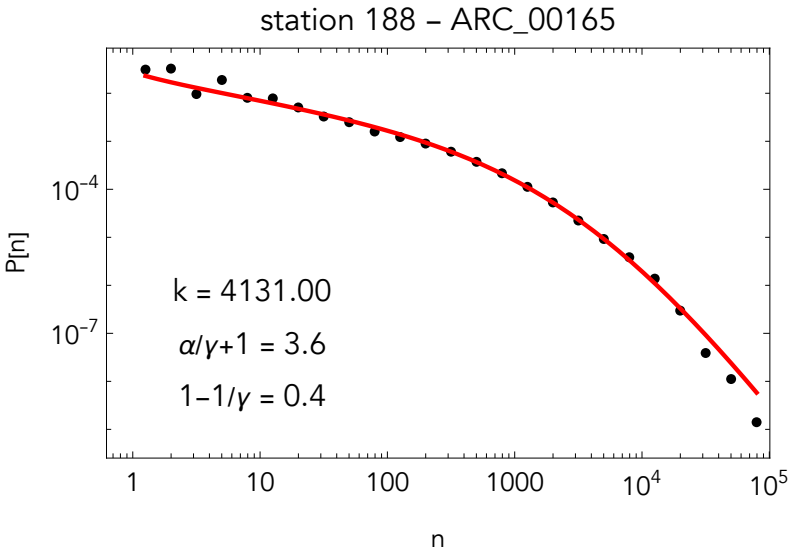






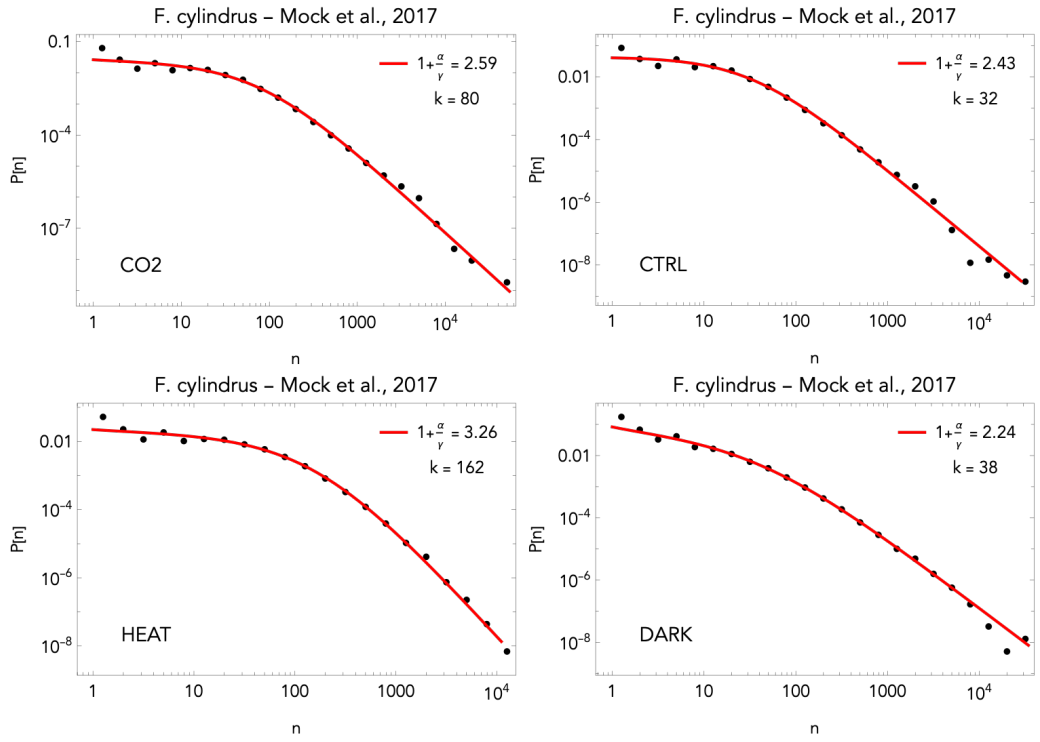


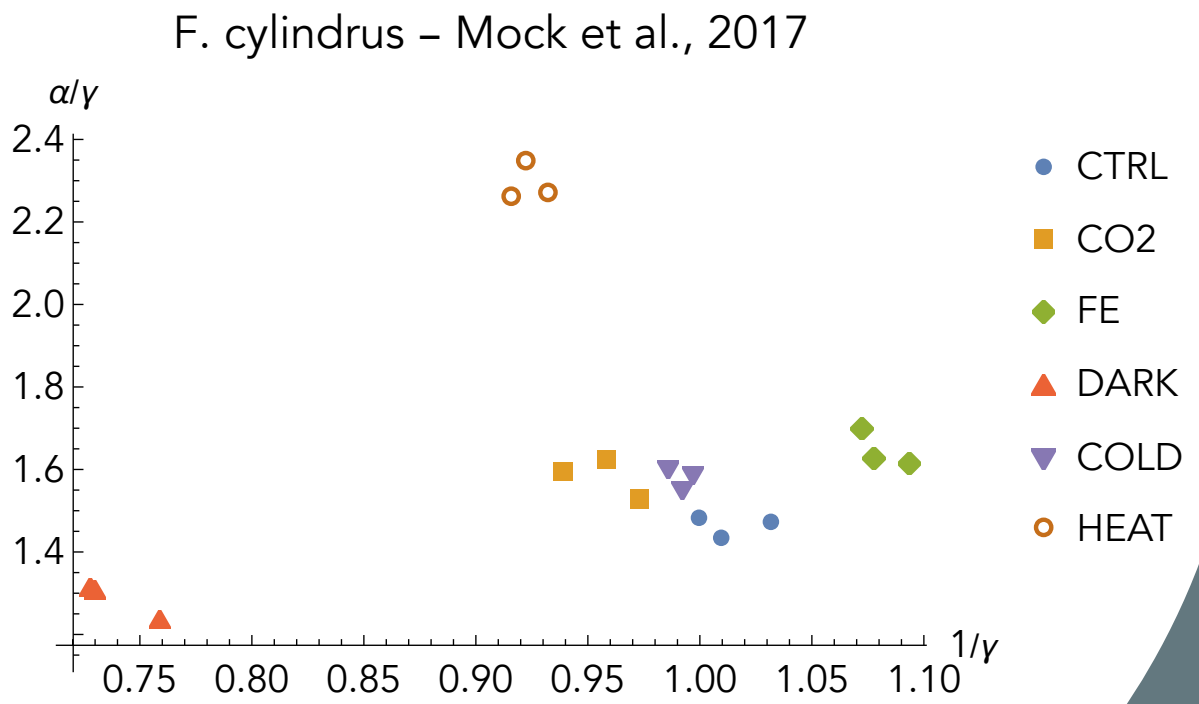
Focusing on only 1 MAG



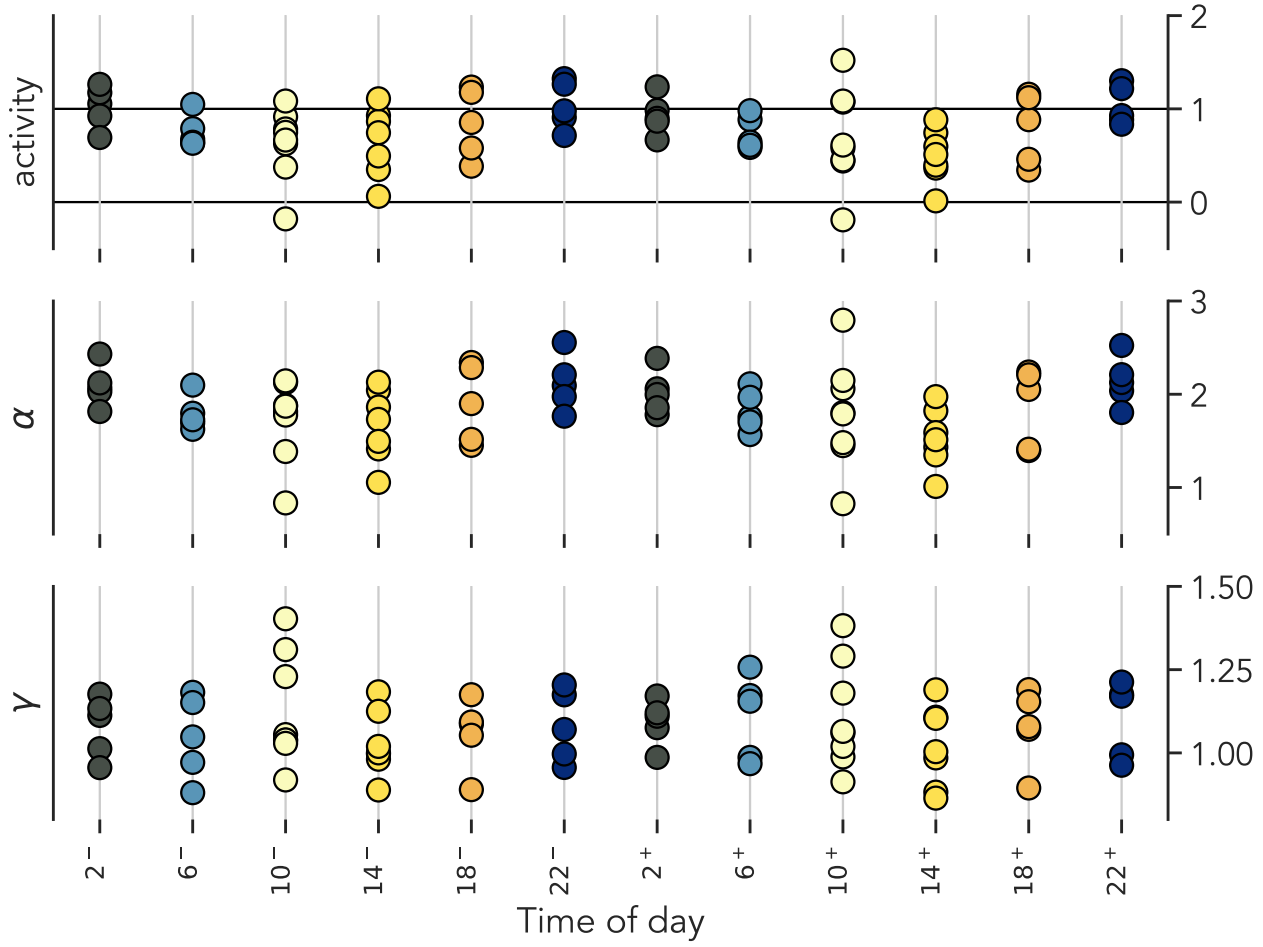
It usually works, when the MAGs are enough…

Assessing the role of biology





Assessing the role of biology



PLOS GENETICS

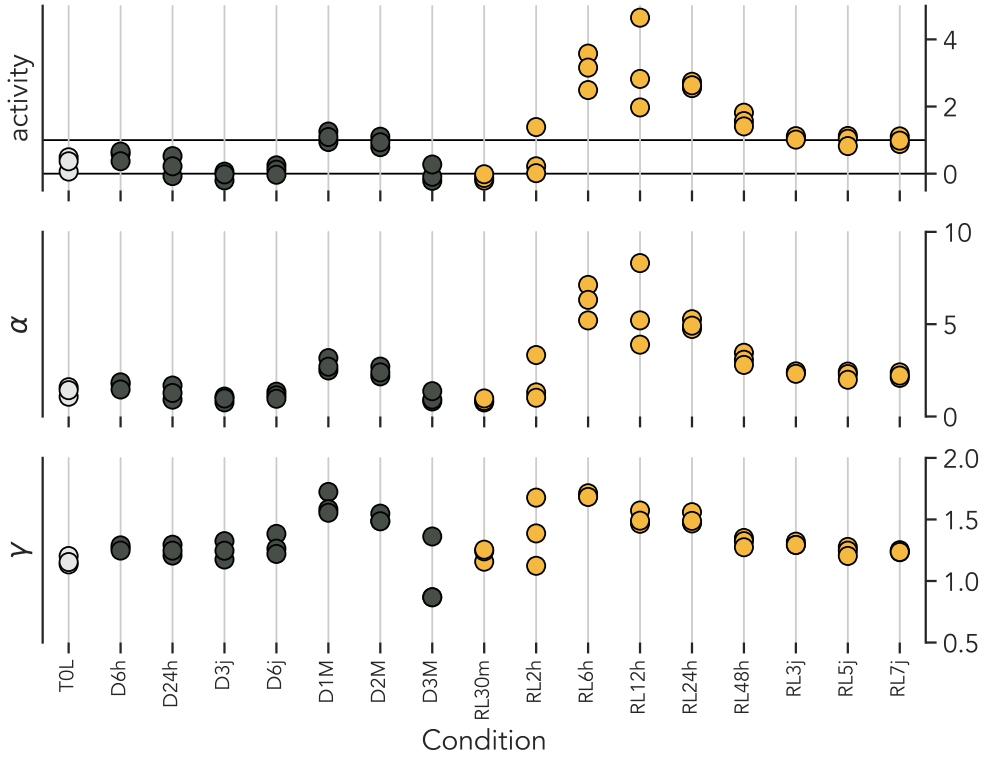
Transcriptional Orchestration of the Global Cellular Response of a Model Pennate Diatom to Diel Light Cycling under Iron Limitation

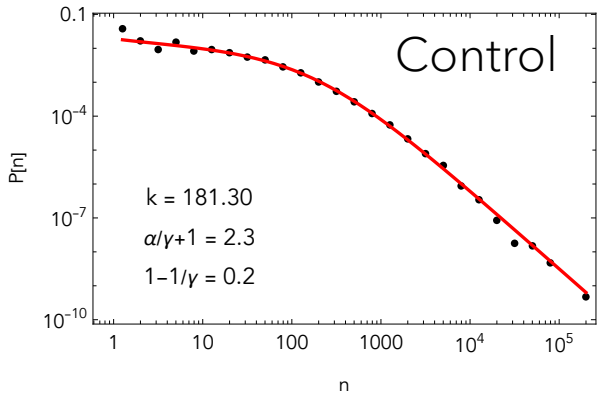
Sarah R. Smith, Jeroen T. F. Gillard, Adam B. Kustka, John P. McCrow, Jonathan H. Badger, Hong Zheng, Ashley M. New, Chris L. Dupont, Toshihiro Obata, Alisdair R. Fernie, Andrew E. Allen

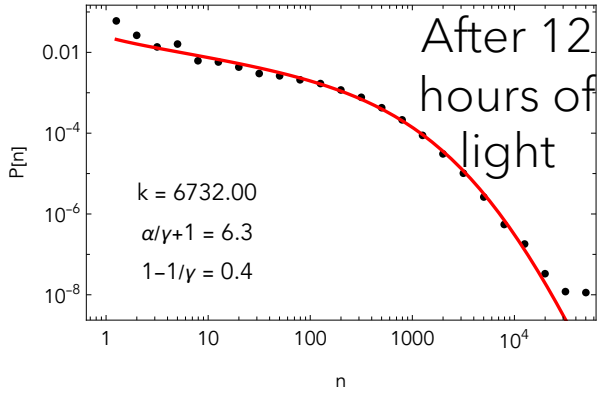
We can derive the mean activity per cell from the exponents of the Pareto IV!

$$\text{activity} = \frac{1}{\gamma} + \frac{\alpha}{\gamma}$$

Assessing the role of biology





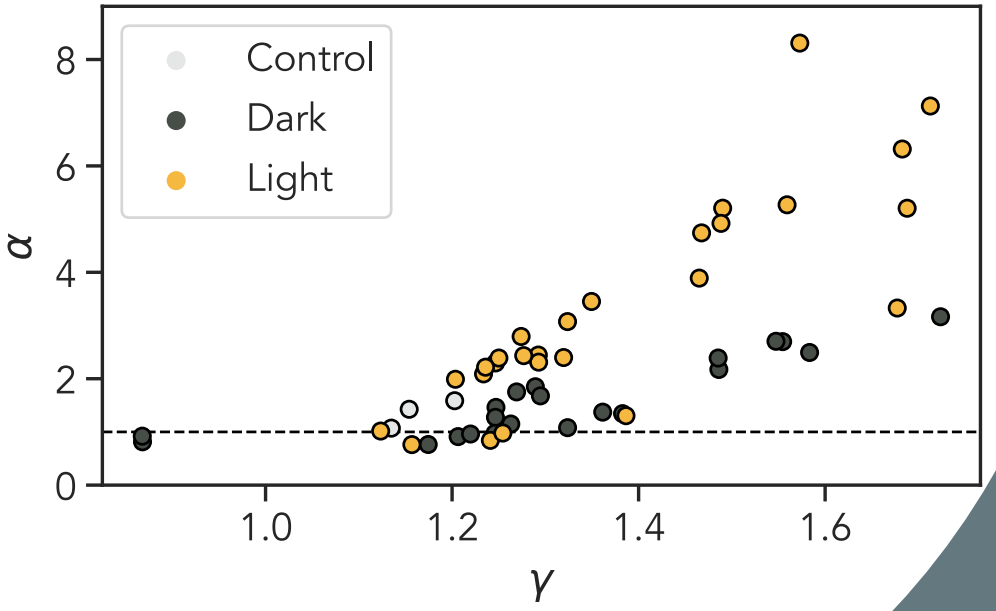


Open Access | Published: 16 January 2017

Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*

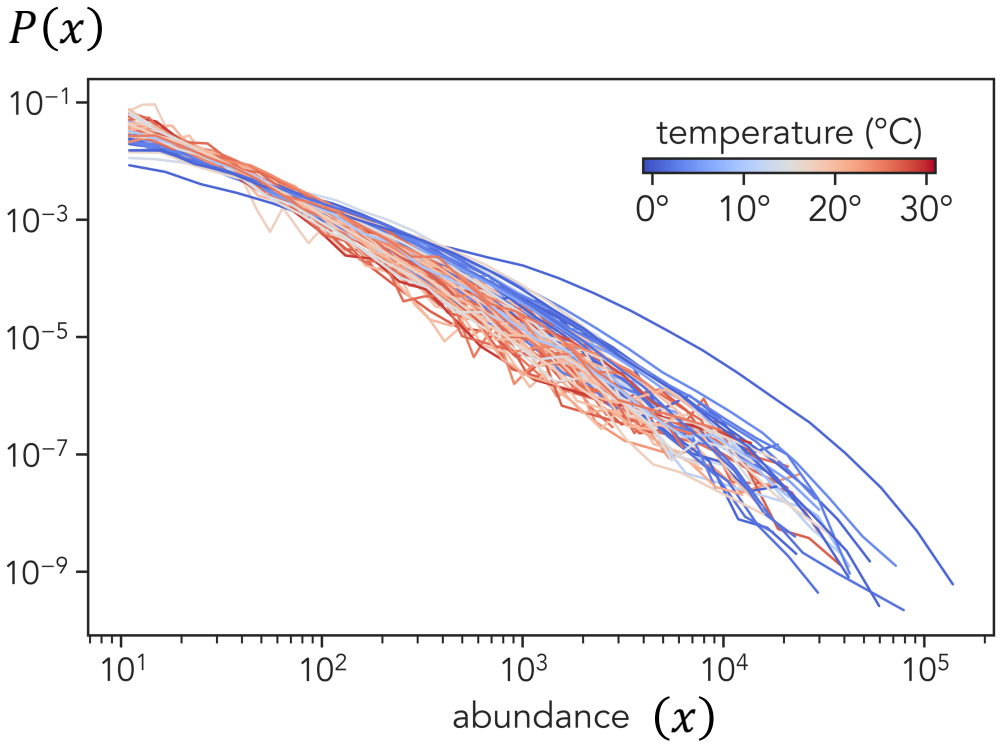
Thomas Mock, Robert P. Otillar, Jan Strauss, Mark McMullan, Pirita Paajanen, Jeremy Schmutz, Asaf Salamov, Remo Sanges, Andrew Toseland, Ben J. Ward, Andrew E. Allen, Christopher L. Dupont, Stephan Frickenhaus, Florian Maumus, Alaguraj Veluchamy, Taoyang Wu, Kerrie W. Barry, Angela Falciatore, Maria I. Ferrante, Antonio E. Fortunato, Gernot Glöckner, Ansgar Gruber, Rachel Hipkin, Michael G. Janech, ... Igor V. Grigoriev + Show authors

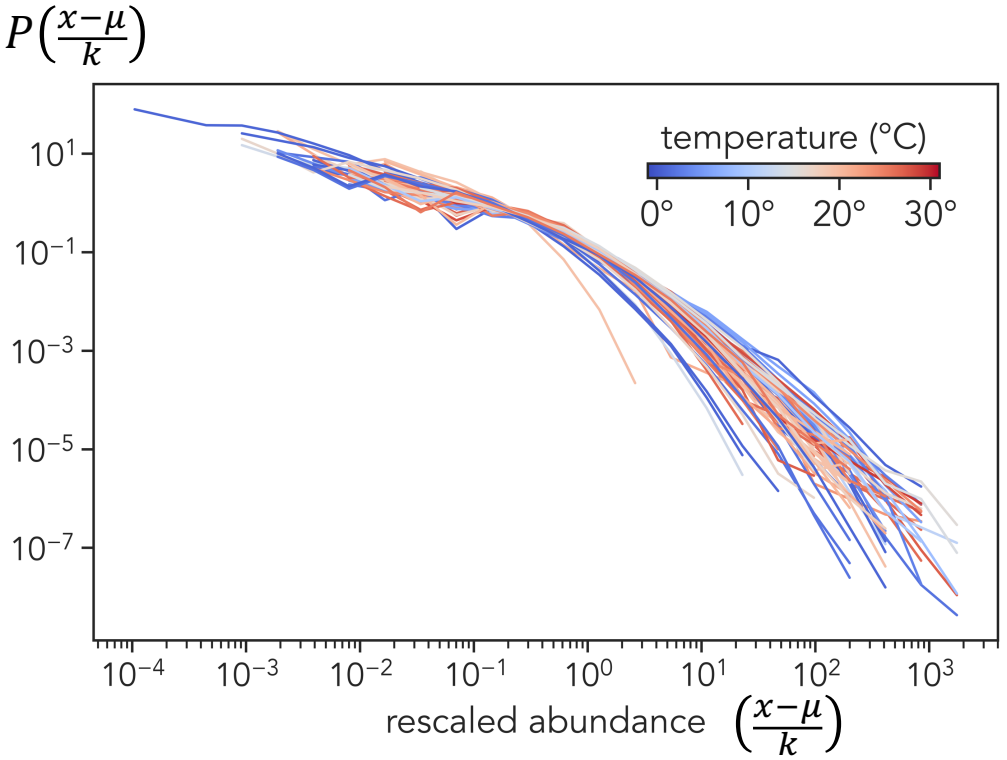
Nature **541**, 536–540 (2017) | Cite this article



$$\text{activity} = \frac{1}{\gamma} + \frac{\alpha}{\gamma}$$

SAD normalization?





Pareto IV distribution:

$$P(x|k,\alpha,\gamma,\mu) = \frac{\alpha}{\gamma k}\left[1+\left(\frac{x-\mu}{k}\right)^{1/\gamma}\right]^{-1-\alpha}\left(\frac{x-\mu}{k}\right)^{-1+1/\gamma}$$

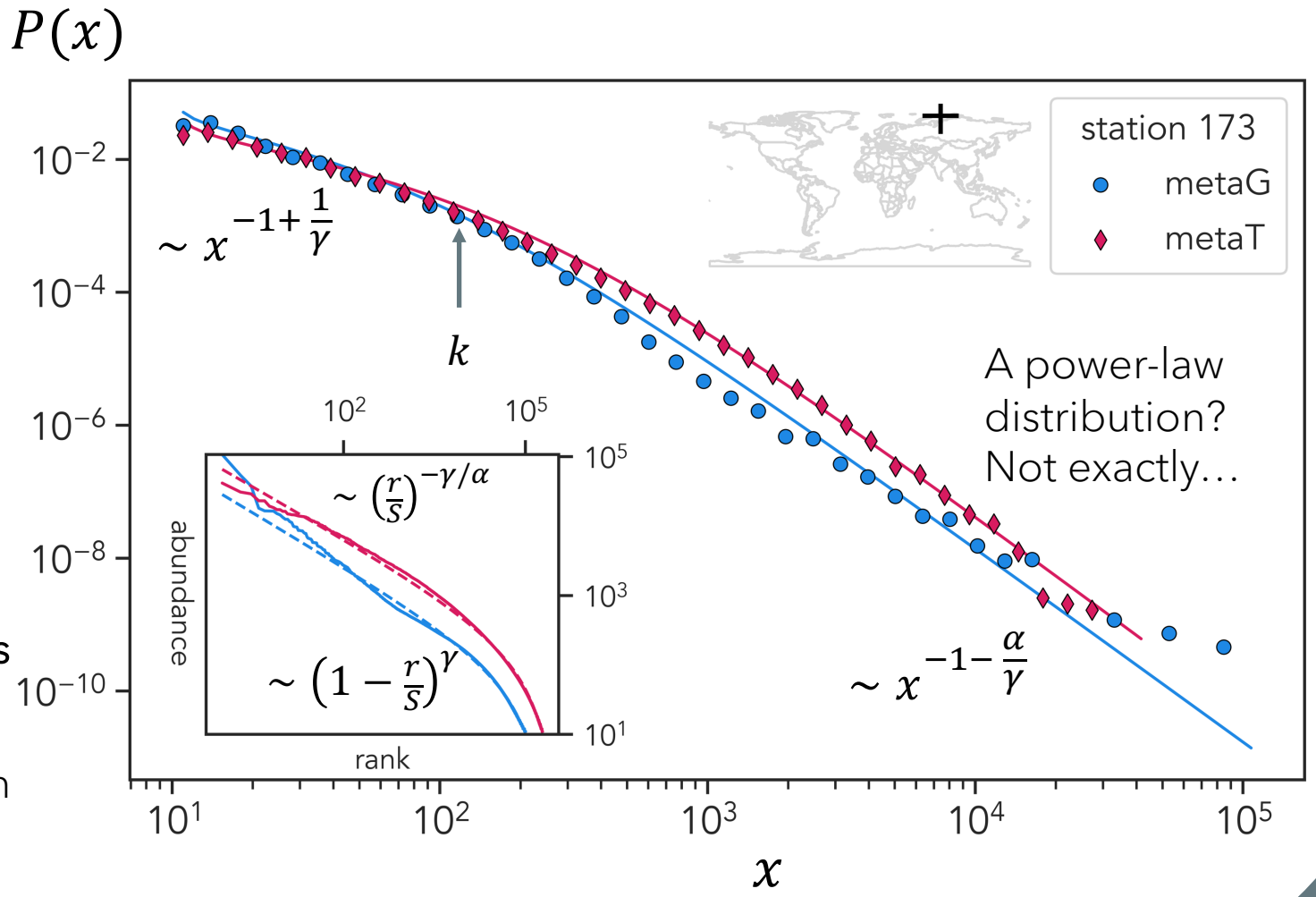
SAD & RAD

Pareto IV distribution:

$$P(x|k,\alpha,\gamma,\mu) = \frac{\alpha}{\gamma k}\left[1+\left(\frac{x-\mu}{k}\right)^{1/\gamma}\right]^{-1-\alpha}\left(\frac{x-\mu}{k}\right)^{-1+1/\gamma}$$

Open questions:

- Is there a **relation** between **metaG** and **metaT**? Probably no
- Why the Pareto IV? Is there a **generative model**? Probably yes
- What is the **biological** meaning of the **parameters**? Yes, model
- Does it work also for **transcriptomics** data? Yes
- Do we have to take it into account in the **data normalization**? Probably



- Analytical formula for the Rank Abundance distribution: $RAD(r) = \mu + k\left[\left(\frac{r}{s}\right)^{-1/\alpha} - 1\right]^{\gamma}$

Thank you!

Daniele Iudicone

Samir Suweis

Emanuele Pigani

Maurizio Ribera d'Alcalà

Sandro Azaele

Lucia Campese

Collaborative LIPh Lab!

Time for the ring model?

Research Article

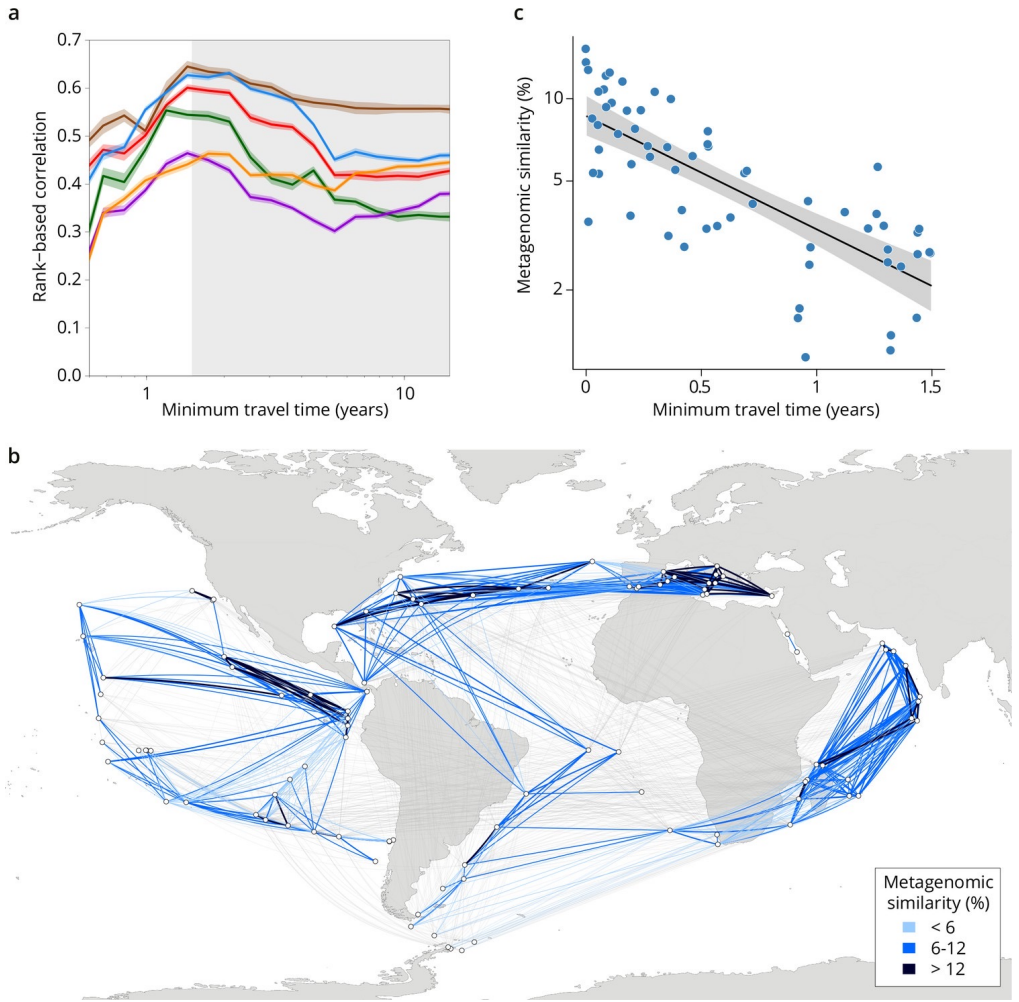
Ecology, Genetics and Genomics

Genomic evidence for global ocean plankton biogeography shaped by large-scale current systems

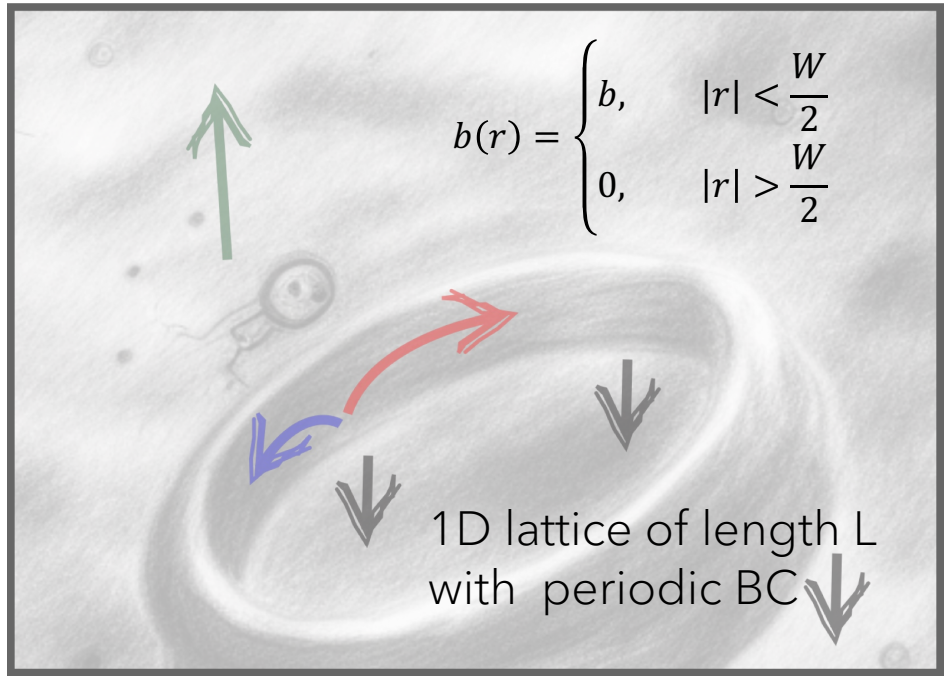
Daniel J Richter, Romain Watteaux, Thomas Vannier, Jade Leconte, Paul Frémont, Gabriel Reygondeau, Nicolas Maillet, Nicolas Henry, Gaëtan Benoit, Ophélie Da Silva, Tom O Delmont, Antonio Fernàndez-Guerra, Samir Suweis, Romain Narci, Cédric Berney, Damien Eveillard, Frederick Gavory, Lionel Guidi, Karine Labadie, Eric Mahieu, Julie Poulain, Sarah Romac, Simon Roux, Céline Dimier, Stefanie Kandels, Marc Picheral, Sarah Searson, Tara Oceans Coordinators, Stéphane Pesant, Jean-Marc Aury, Jennifer R Brum, Claire Lemaitre, Eric Pelletier, Peer Bork, Shinichi Sunagawa, Fabien Lombard, Lee Karp-Boss, Chris Bowler, Matthew B Sullivan, Eric Karsenti, Mahendra Mariadassou, Ian Probert, Pierre Peterlongo, Patrick Wincker, Colomban de Vargas, Maurizio Ribera d'Alcalà, Daniele Iudicone, Olivier Jaillon « see less

The metagenomic similarity **decay exponentially** with travel time (**Lagrangian distance**)

Why?



The Ring model

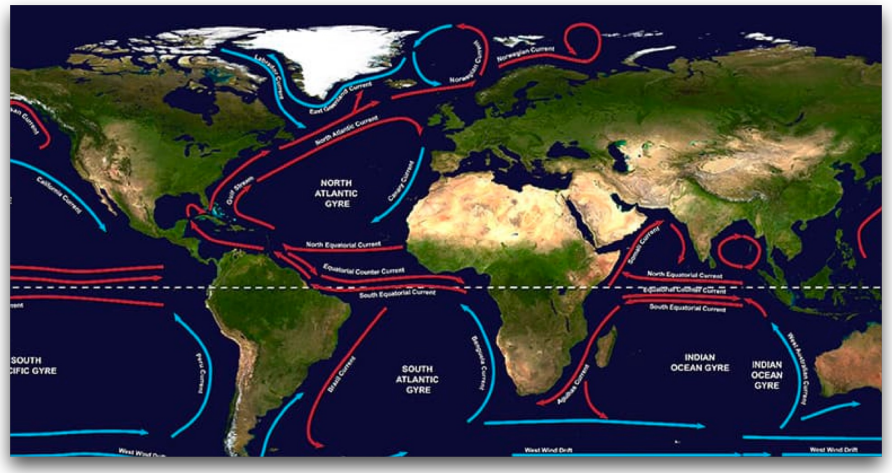


$$X_i \xrightarrow{b(r)} X_i + X_i$$

$$X_i \xrightarrow{d} \emptyset$$

$$X_i \xrightarrow{p_+} X_{i+1}$$

$$X_i \xrightarrow{p_-} X_{i-1}$$

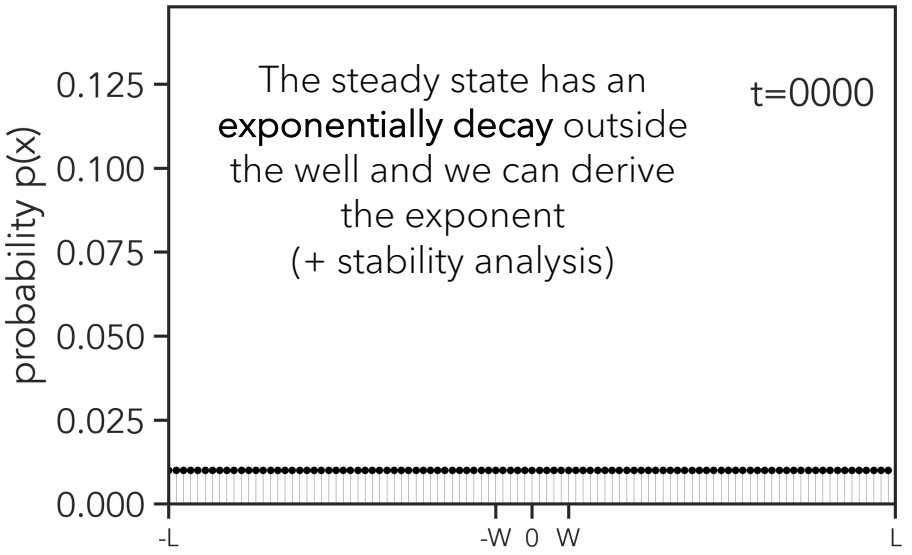


Continuum limit of the deterministic part of the Langevin equation:

$$\frac{\partial x}{\partial t} = D\frac{\partial^2 x}{\partial r^2} - v\frac{\partial x}{\partial r} + \mu(r)x$$

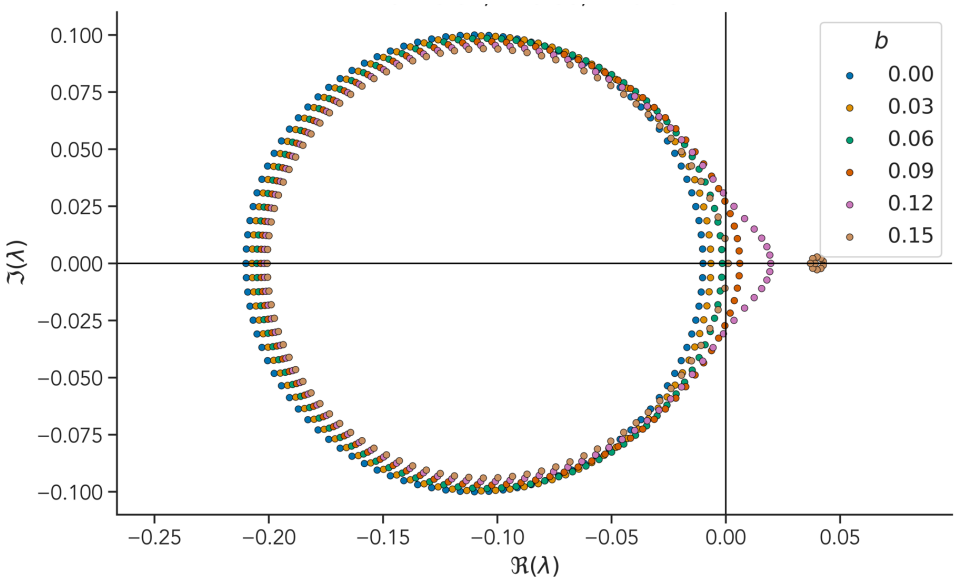
(~ particle in a potential well)

$$D = p_+ \; ; \; v = p_+ - p_- \; ; \; \mu(r) = b(r) - d$$



By means of the stochastic terms we want to analytically investigate the decay in similarity…

The Ring model

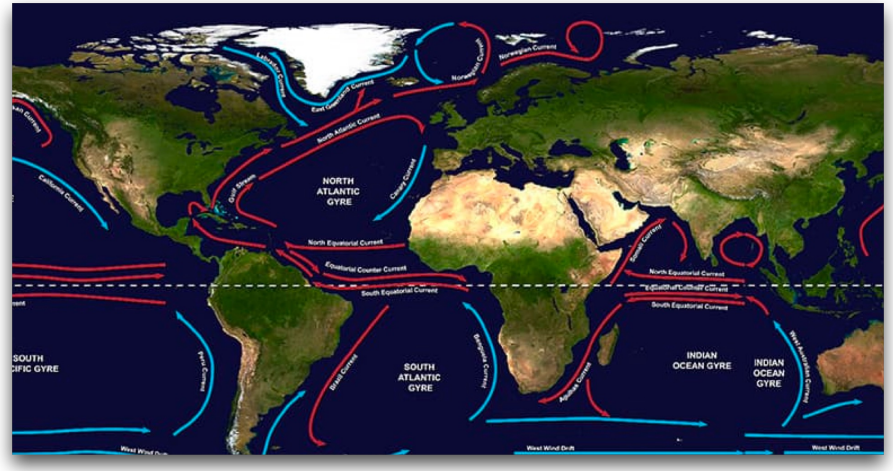


$$X_i \xrightarrow{b(r)} X_i + X_i$$

$$X_i \xrightarrow{d} \emptyset$$

$$X_i \xrightarrow{p_+} X_{i+1}$$

$$X_i \xrightarrow{p_-} X_{i-1}$$

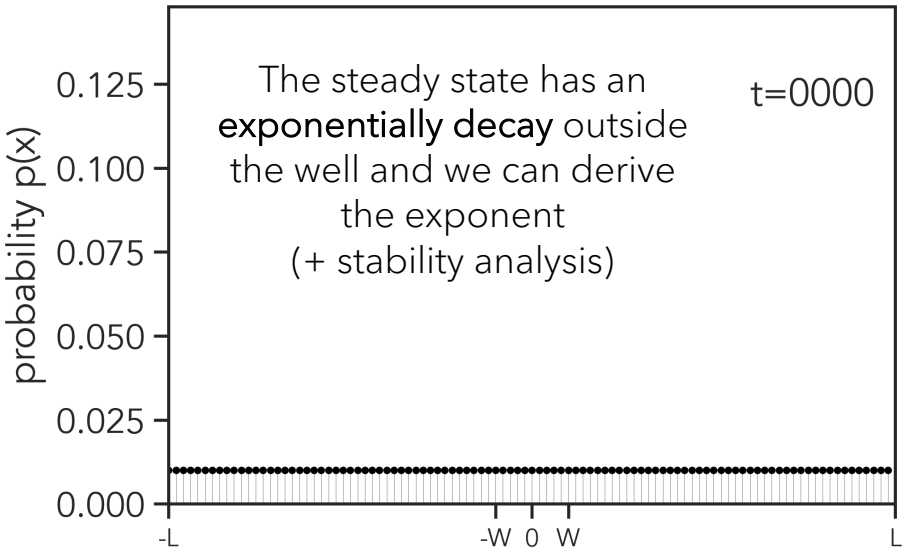


Continuum limit of the deterministic part of the Langevin equation:

$$\frac{\partial x}{\partial t} = D\frac{\partial^2 x}{\partial r^2} - v\frac{\partial x}{\partial r} + \mu(r)x$$

(~ particle in a potential well)

$$D = p_+ \; ; \; v = p_+ - p_- \; ; \; \mu(r) = b(r) - d$$



By means of the stochastic terms we want to analytically investigate the decay in similarity…

Thank you again!

Daniele Iudicone

Samir Suweis

Emanuele Pigani

Maurizio
Ribera d'Alcalà

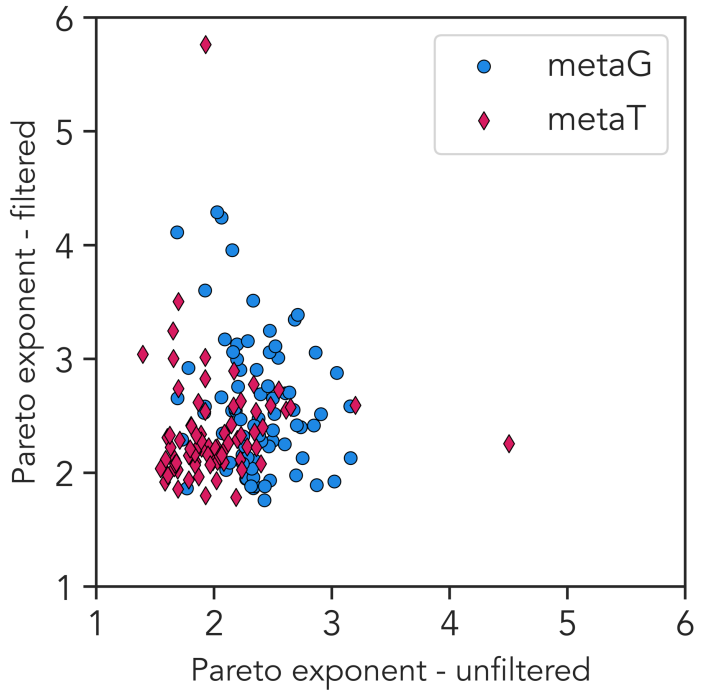
Sandro Azaele

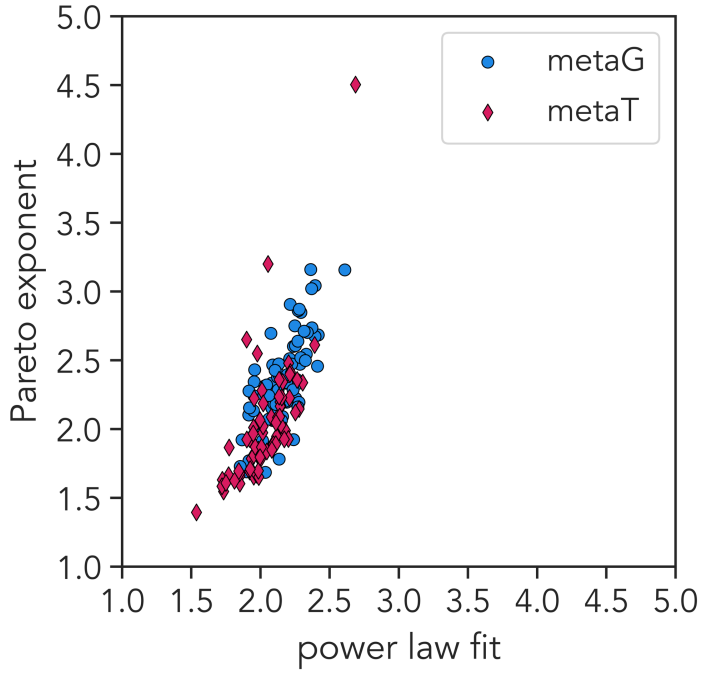
Lucia Campese

Collaborative
LIPh Lab!

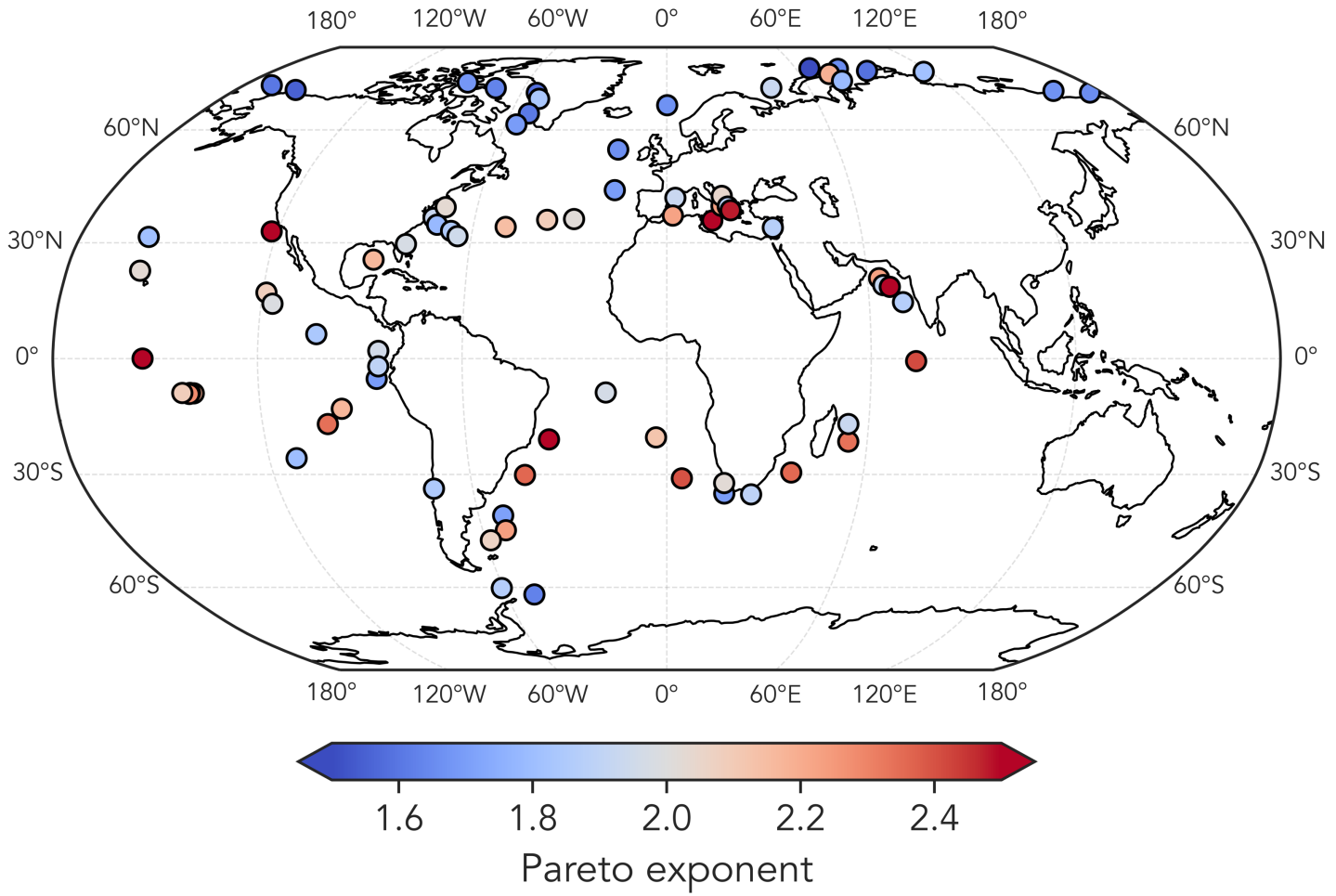
That's all!

BACKUP SLIDES

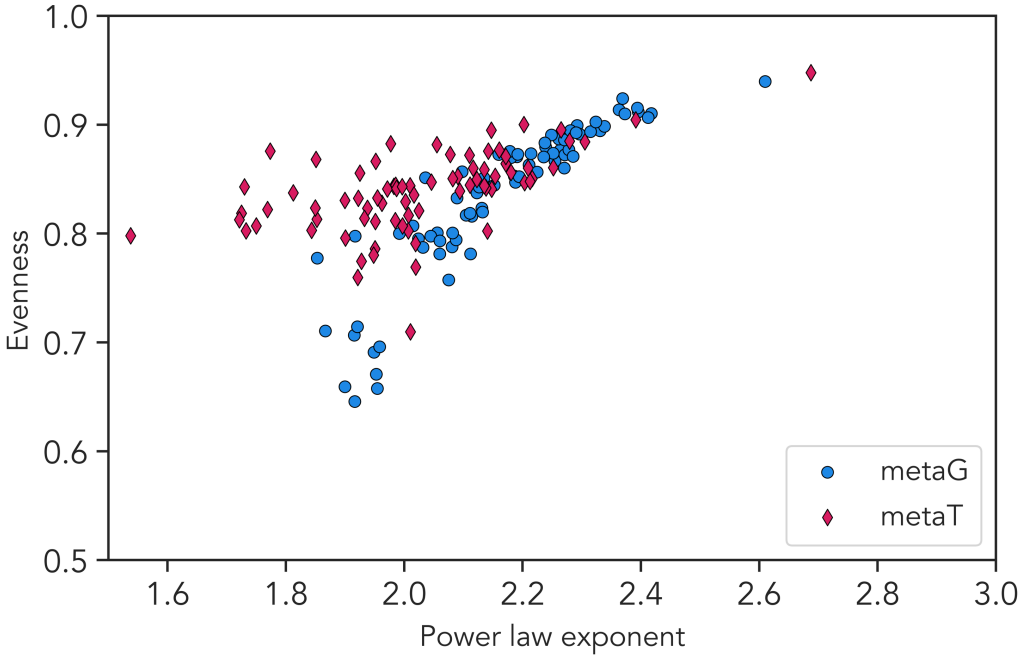

metaG
metaT
Pareto exponent - filtered
Pareto exponent - unfiltered
1
2
3
4
5
6

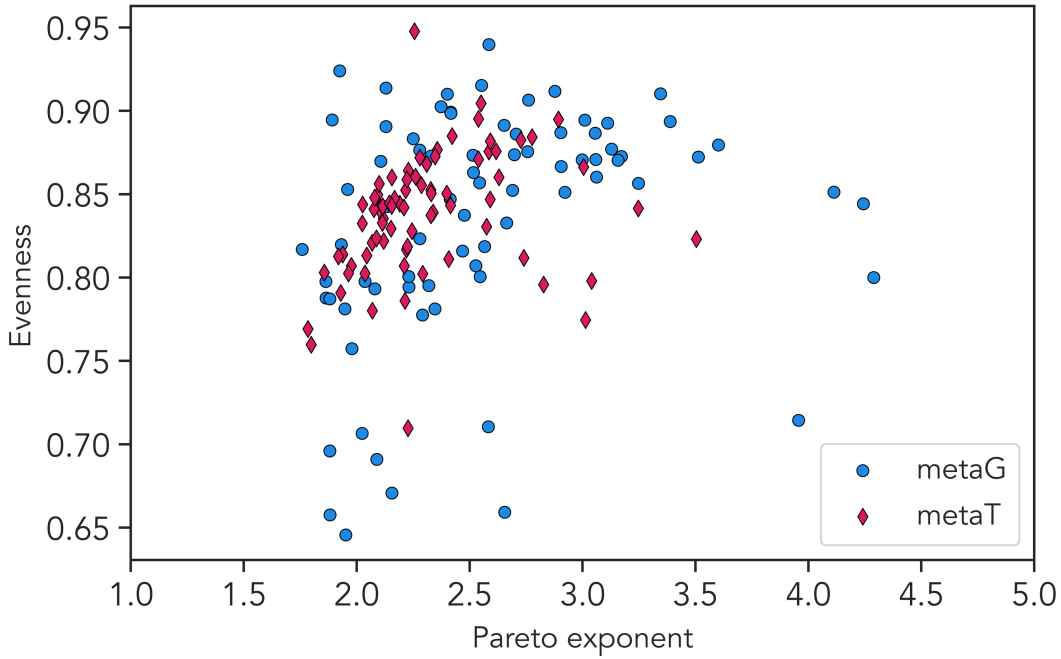

metaG
metaT
Pareto exponent
power law fit
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0

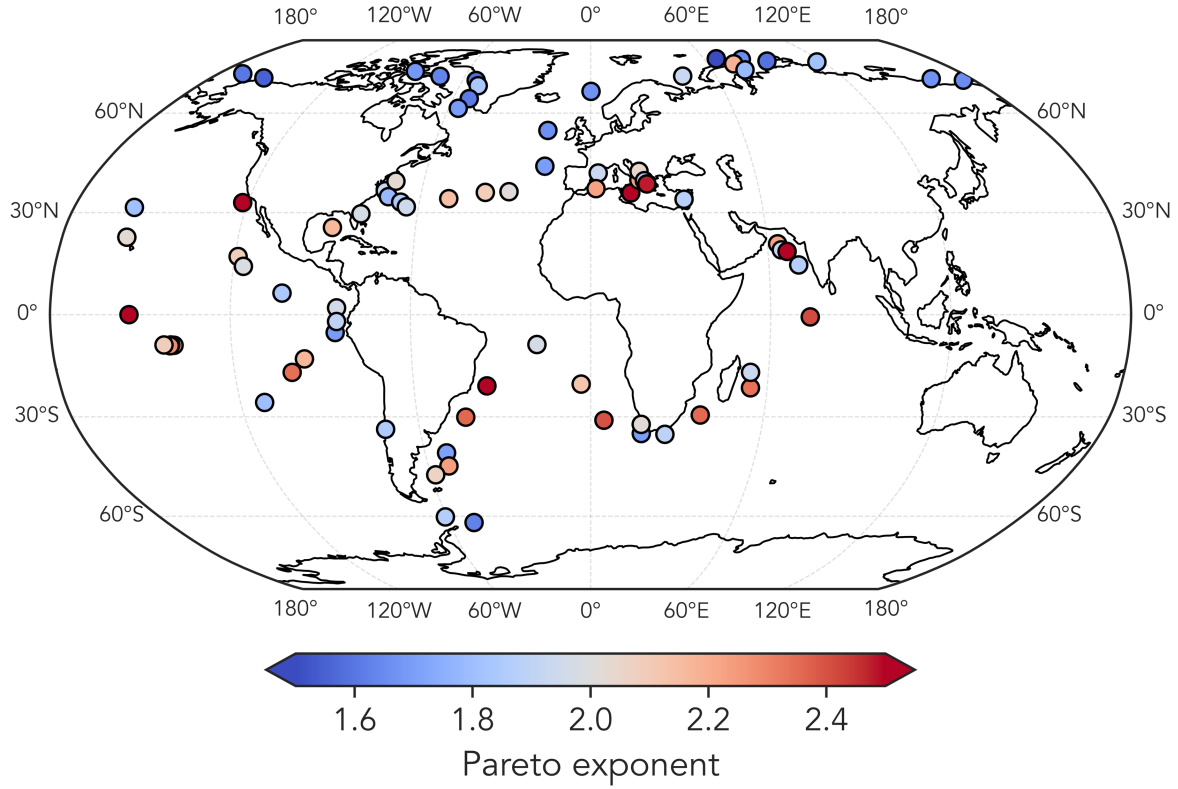
metaT


180°
120°W
60°W
0°
60°E
120°E
180°
60°N
30°N
0°
30°S
60°S
1.6
1.8
2.0
2.2
2.4
Pareto exponent

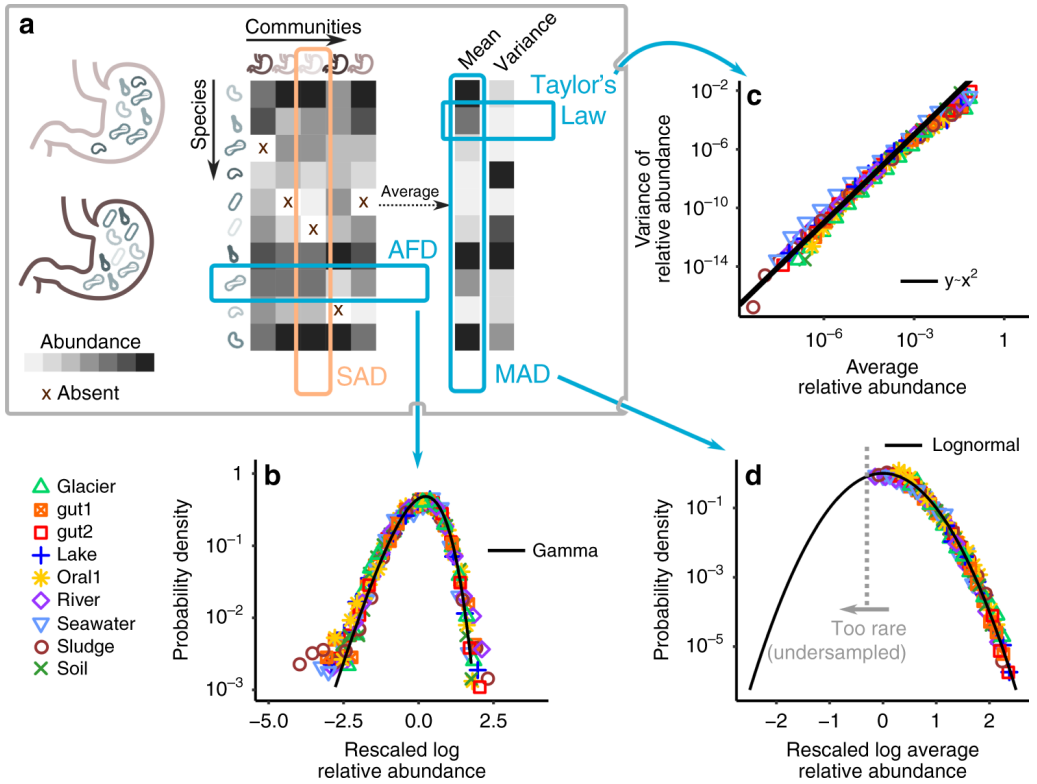
metaT

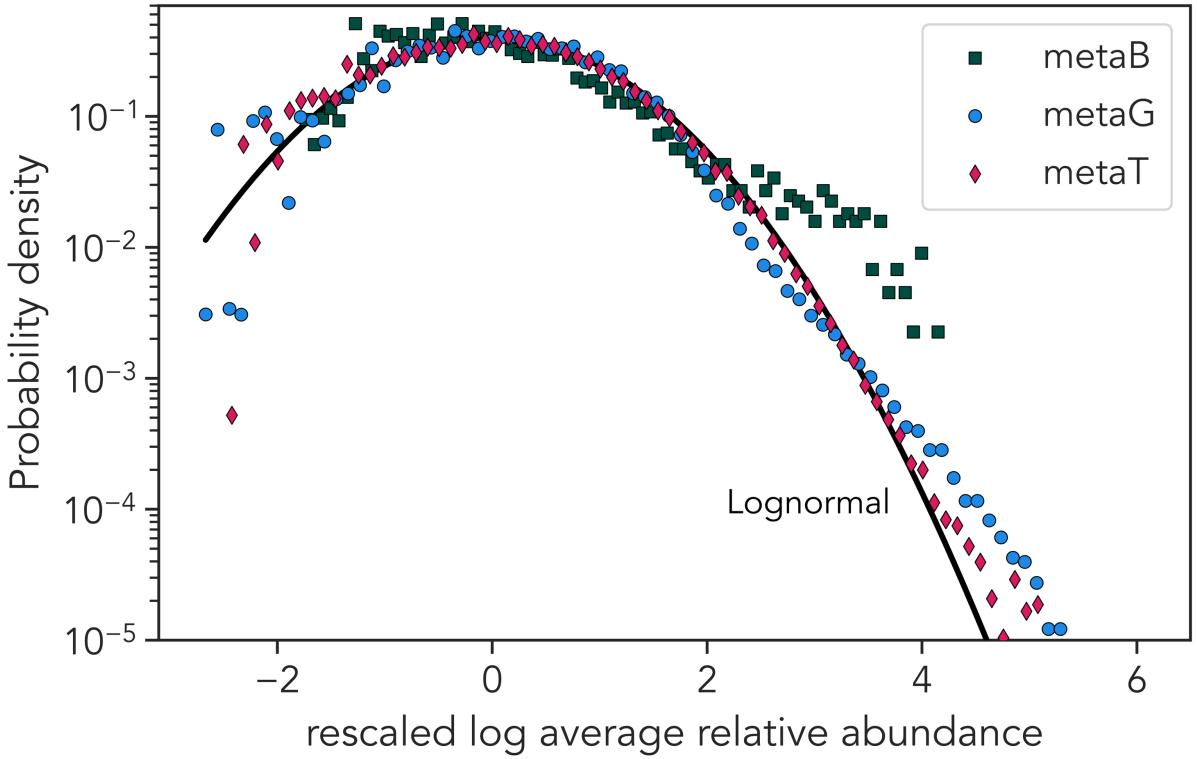

Evenness
Power law exponent
metaG
metaT


Evenness
Pareto exponent
metaG
metaT

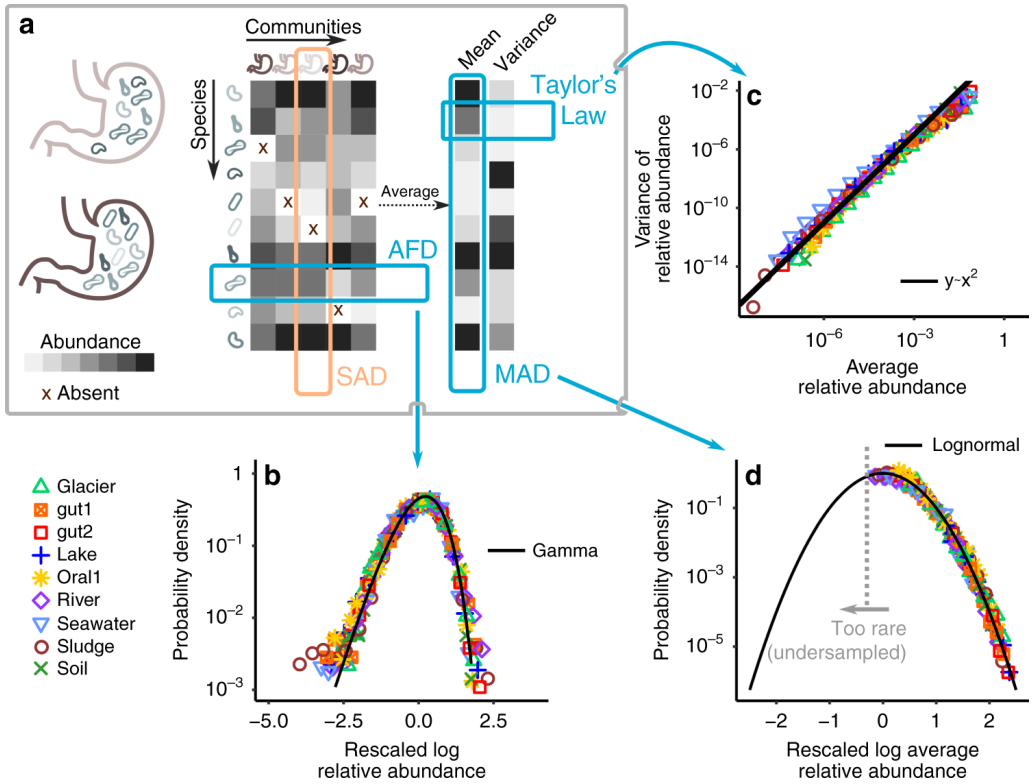

180°
120°W
60°W
0°
60°E
120°E
180°
60°N
30°N
0°
30°S
60°S
1.6
1.8
2.0
2.2
2.4
Pareto exponent

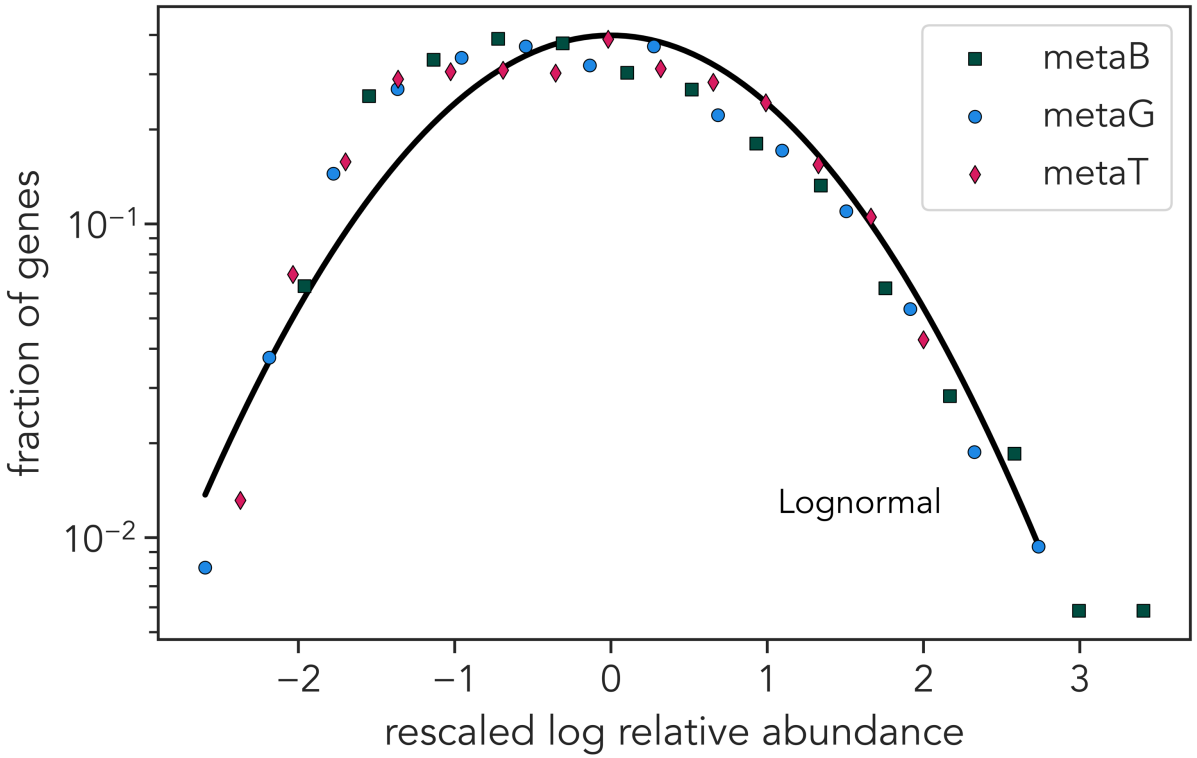
MAD


a
Communities
Species
Mean
Variance
Taylor's Law
Average
AFD
SAD
MAD
Abundance
x Absent
b
Probability density
Gamma
Rescaled log relative abundance
c
Variance of relative abundance
y~x²
Average relative abundance
d
Lognormal
Too rare (undersampled)
Probability density
Rescaled log average relative abundance
Glacier
gut1
gut2
Lake
Oral1
River
Seawater
Sludge
Soil

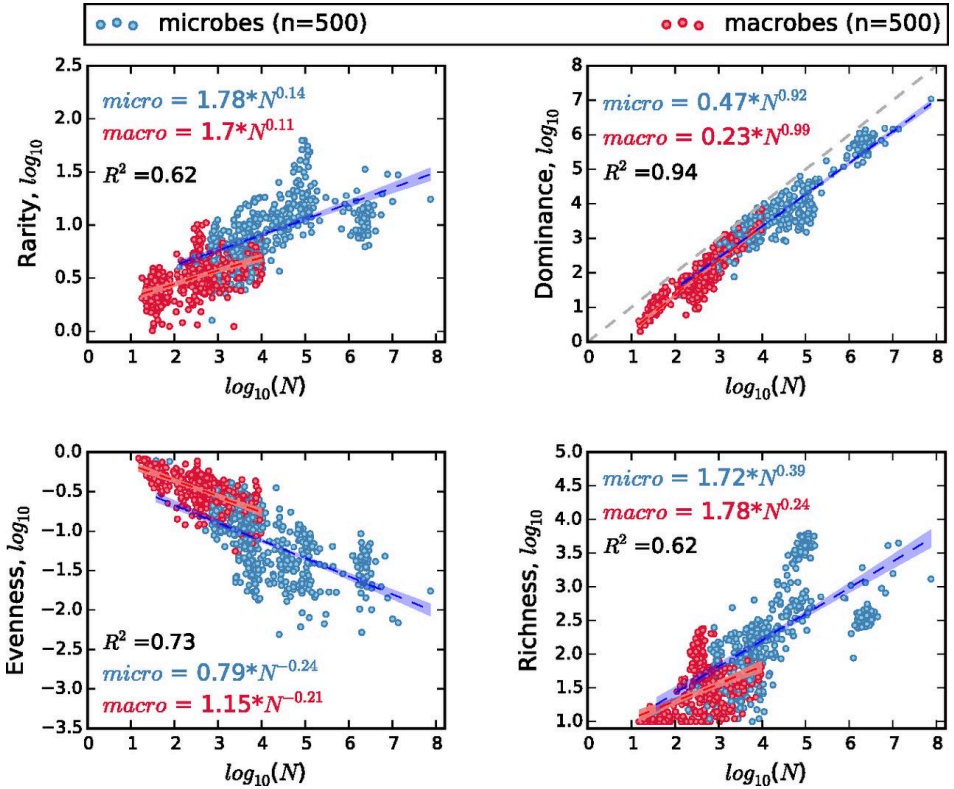

metaB
metaG
metaT
Lognormal
Probability density
rescaled log average relative abundance

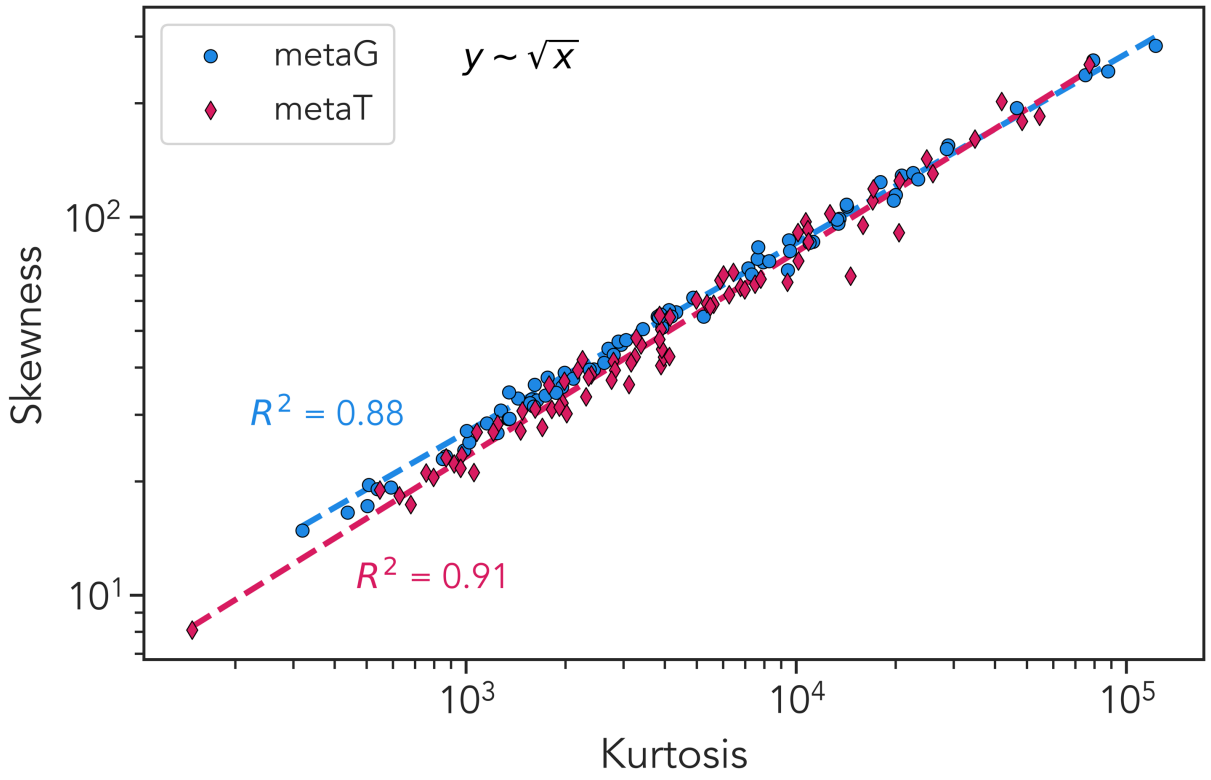
AFD


a
Communities
Species
Mean
Variance
Taylor's Law
Average
AFD
SAD
MAD
Abundance
x Absent
Glacier
gut1
gut2
Lake
Oral1
River
Seawater
Sludge
Soil
b
Probability density
Gamma
Rescaled log relative abundance
c
Variance of relative abundance
y~x²
Average relative abundance
d
Lognormal
Too rare (undersampled)
Probability density
Rescaled log average relative abundance

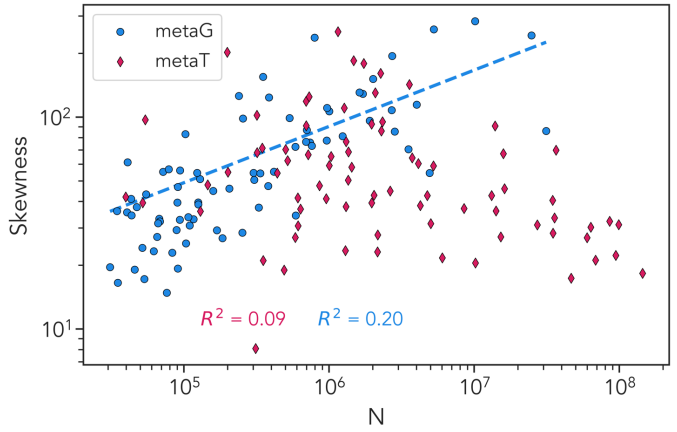

metaB
metaG
metaT
Lognormal
fraction of genes
rescaled log relative abundance

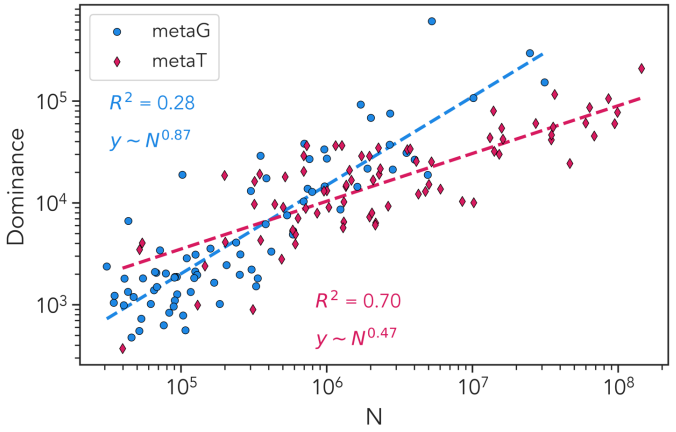
Macroecological patterns

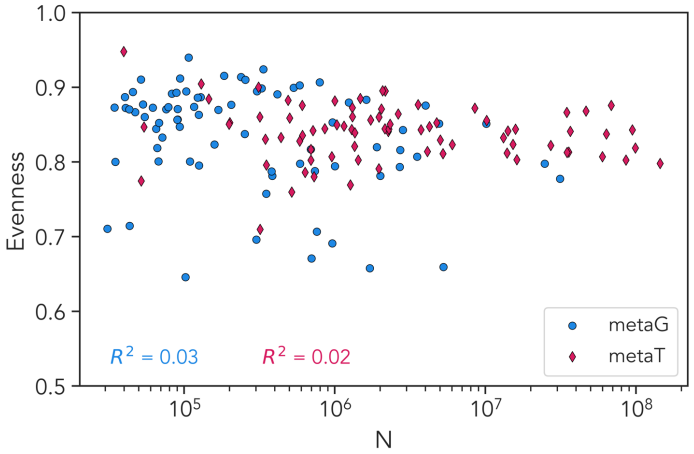


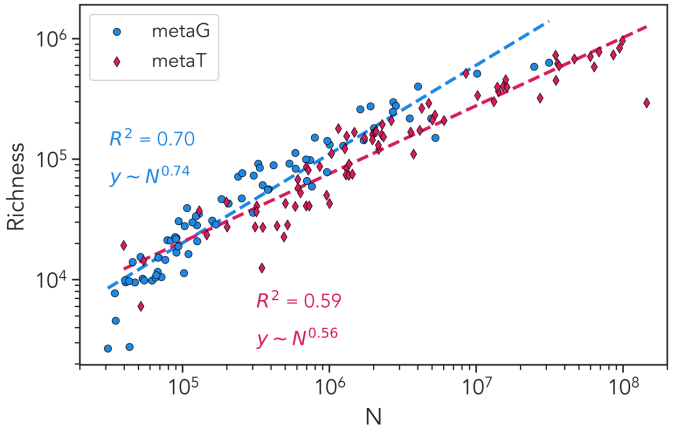


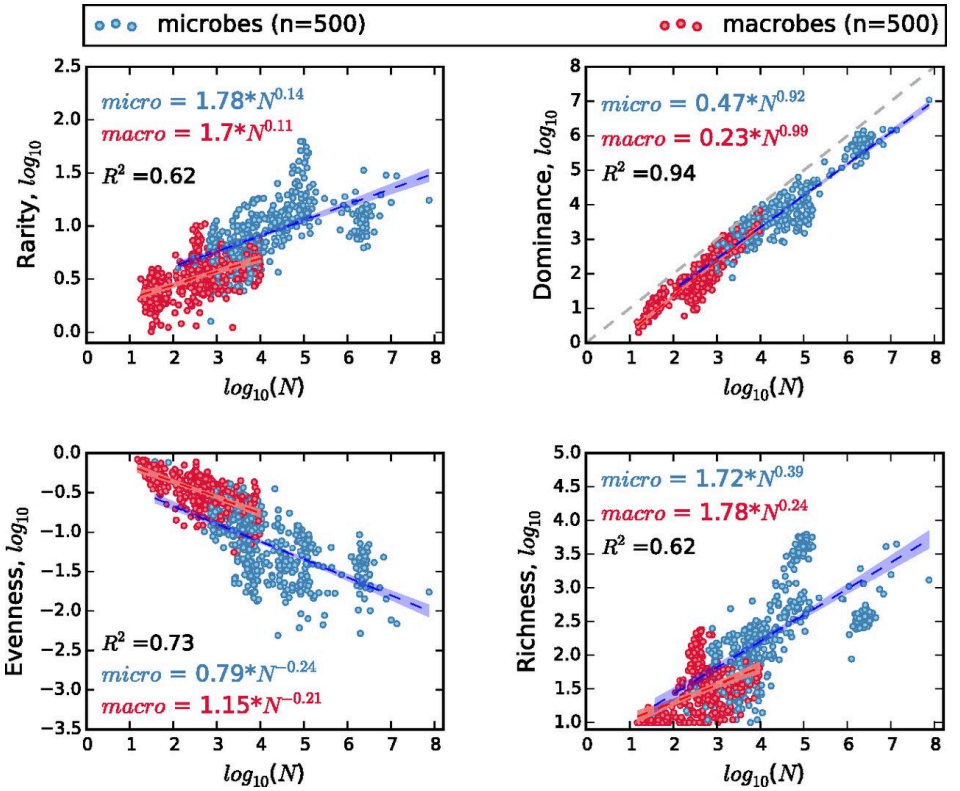
Macroecological patterns



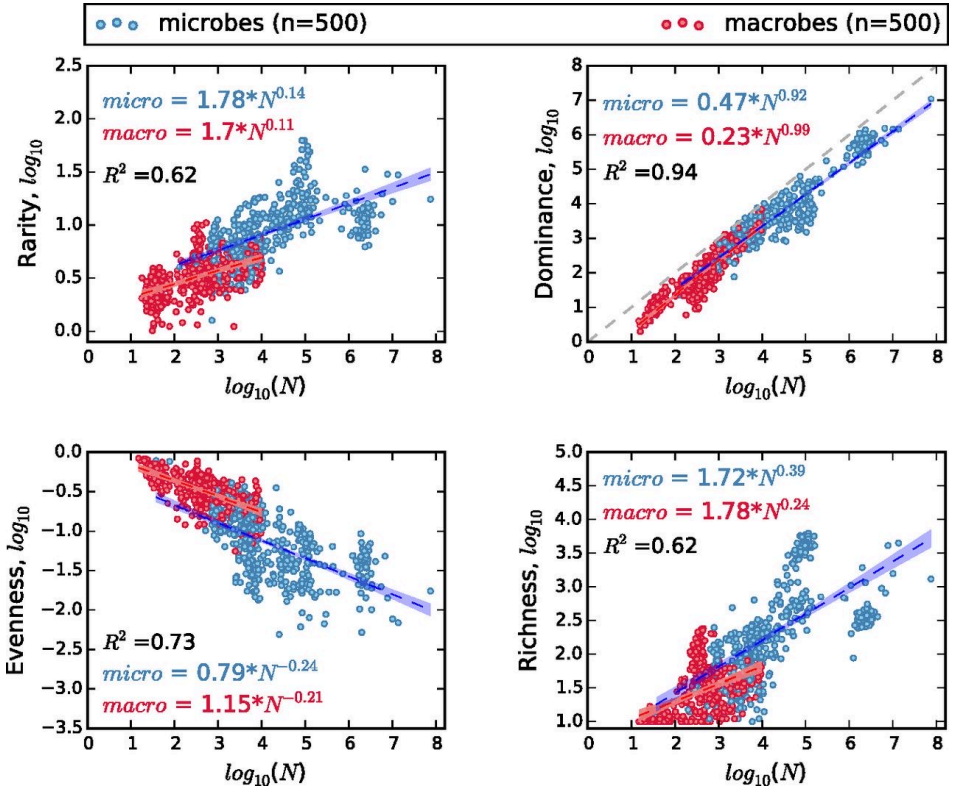


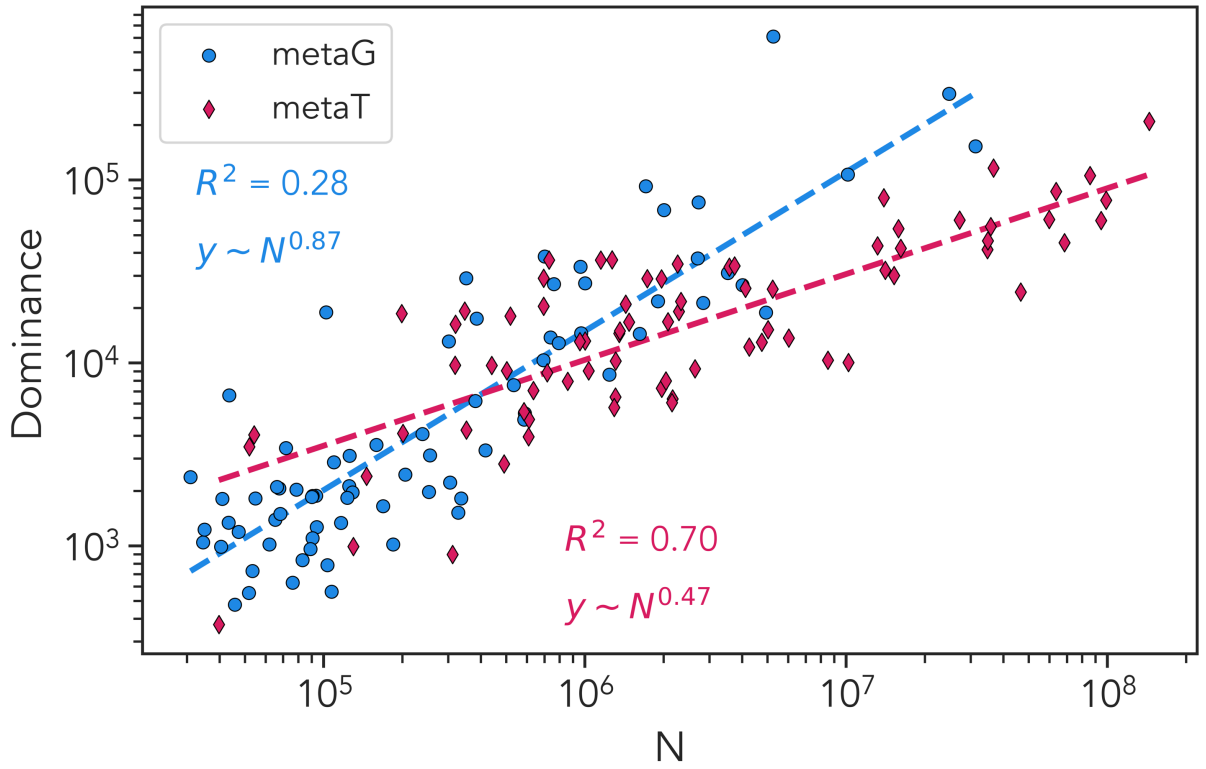




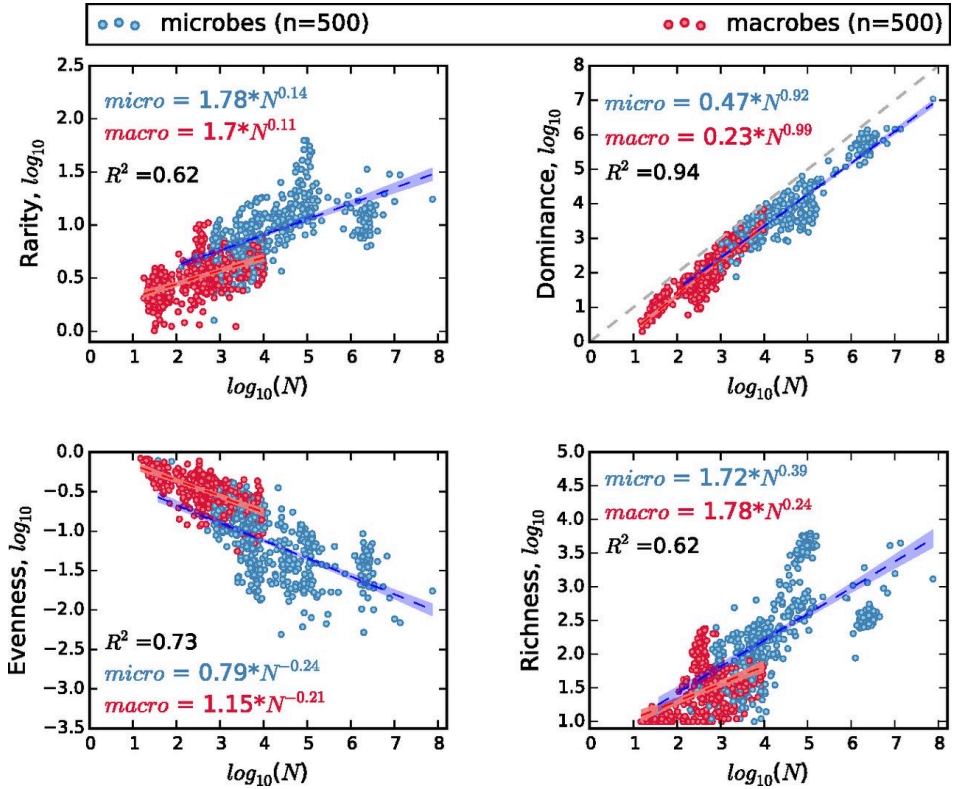


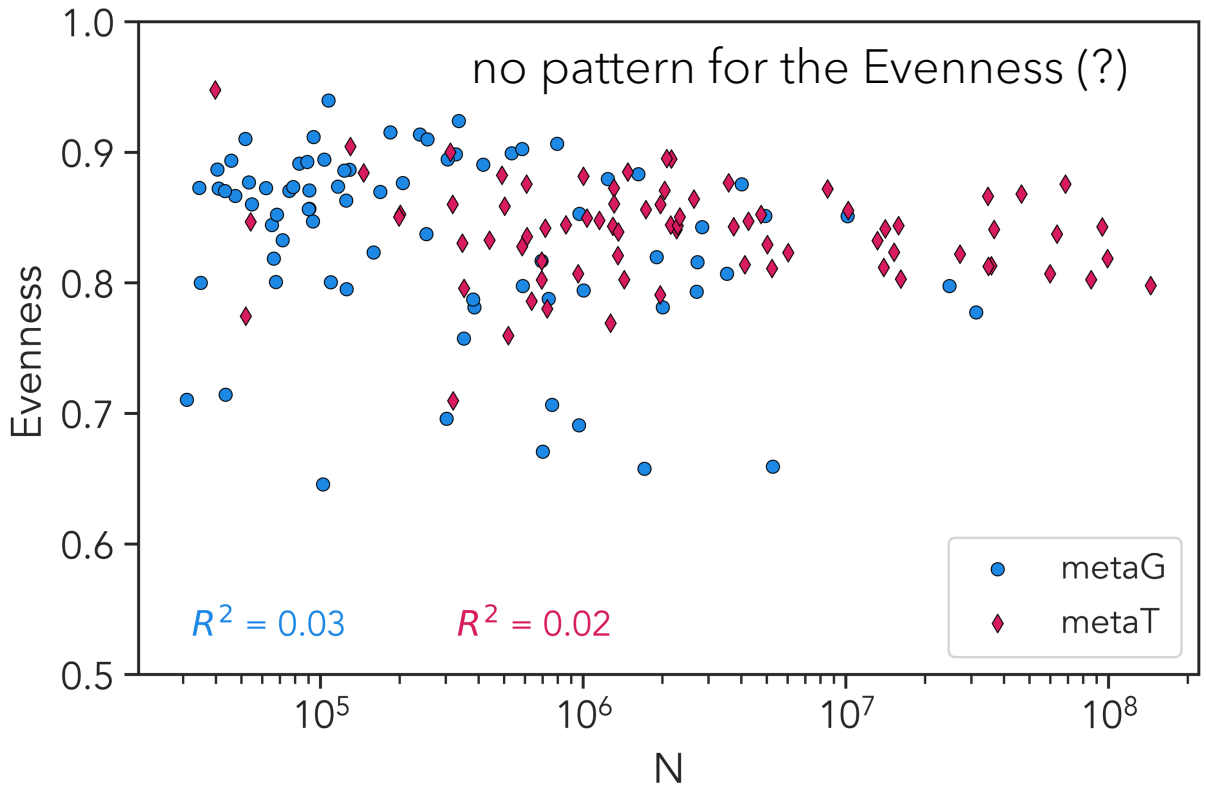
Macroecological patterns



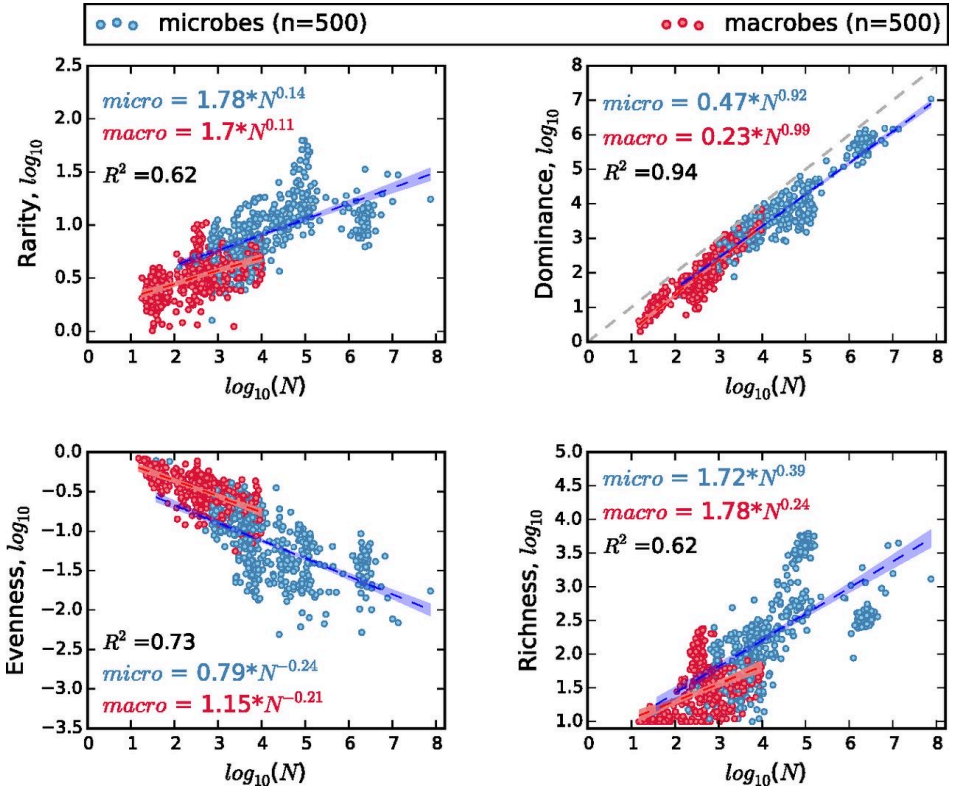


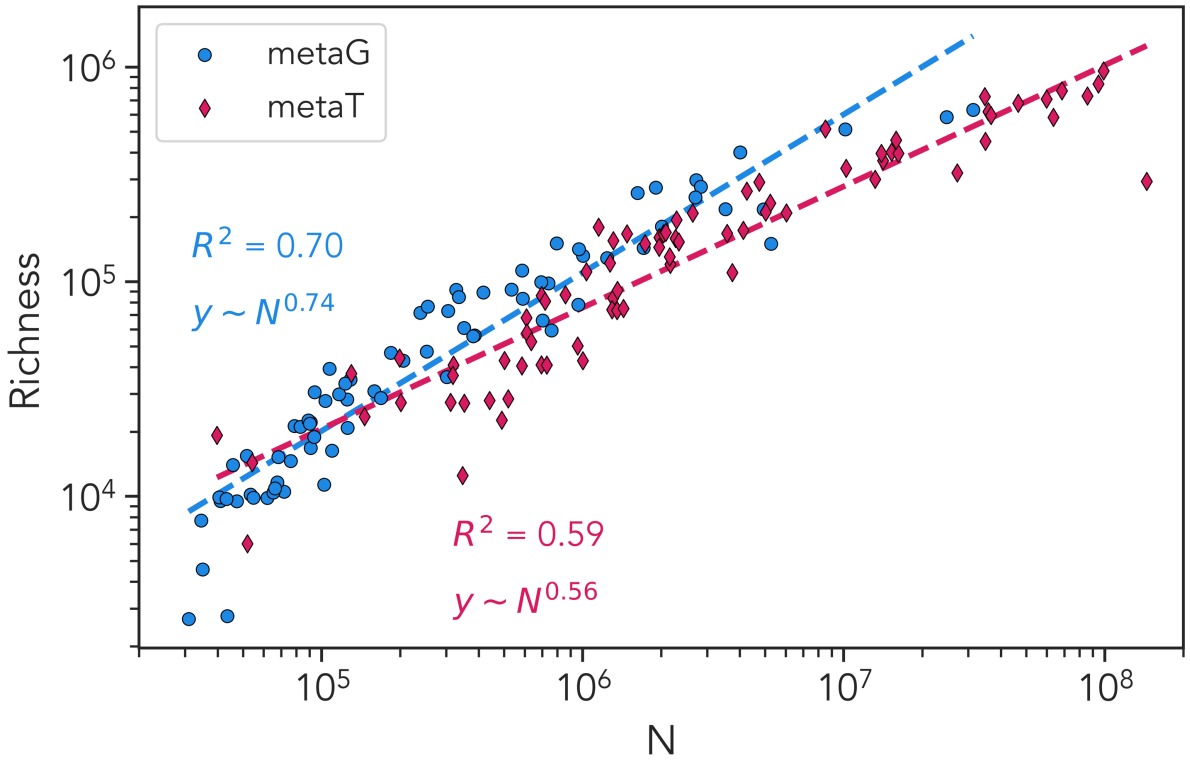
Macroecological patterns





Macroecological patterns





Macroecological laws on the seascape

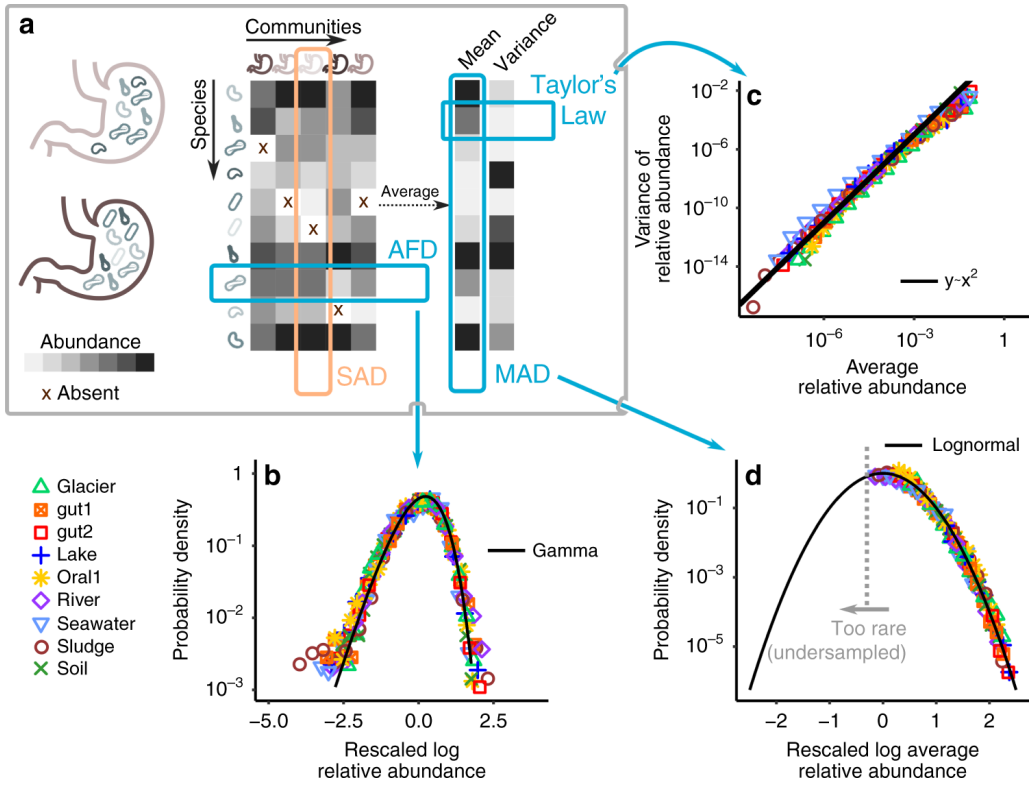
ARTICLE

https://doi.org/10.1038/s41467-020-18529-y OPEN

Macroecological laws describe variation and diversity in microbial communities

Jacopo Grilli [1,2]✉

nature COMMUNICATIONS



RESEARCH ARTICLE | ECOLOGY

Scaling laws predict global microbial diversity

Kenneth J. Locey and Jay T. Lennon Authors Info & Affiliations

PNAS