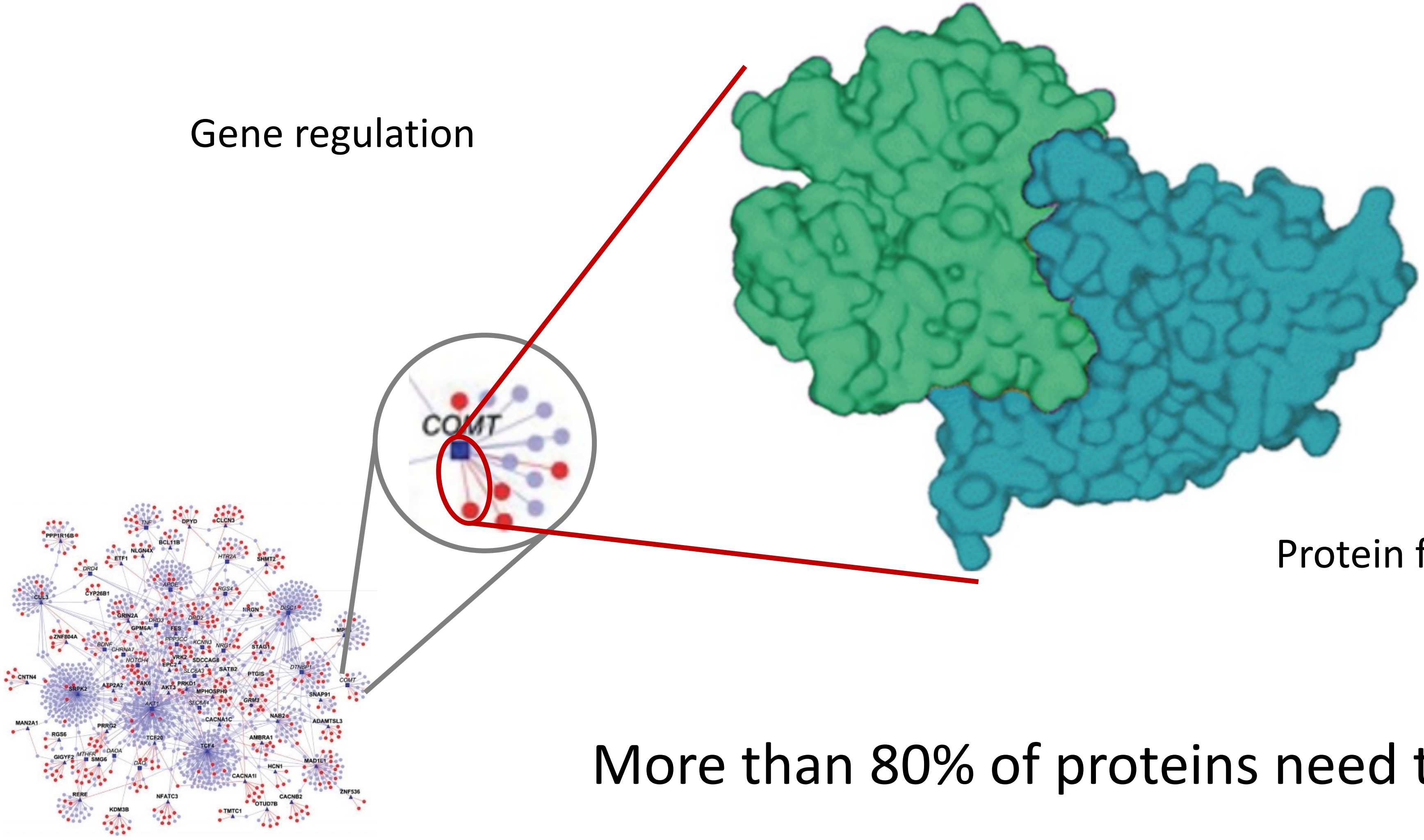
Evolution of Networks: From Ecology to Proteins

2025.03.28

Zhengzhang Zhu

Backgrounds about PPI

PPI network consists of protein-protein interaction



Signal transduction

Gene regulation

Immune response

Cell cycle control

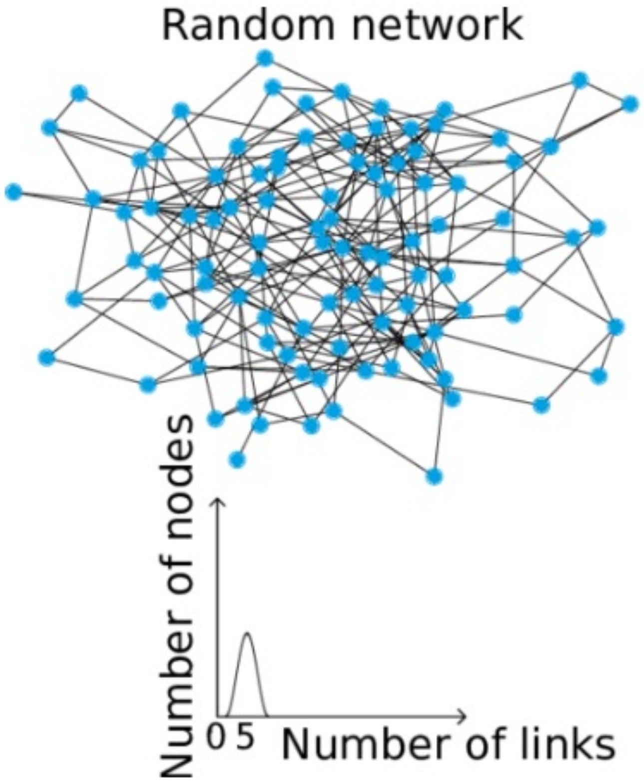
Metabolic pathways

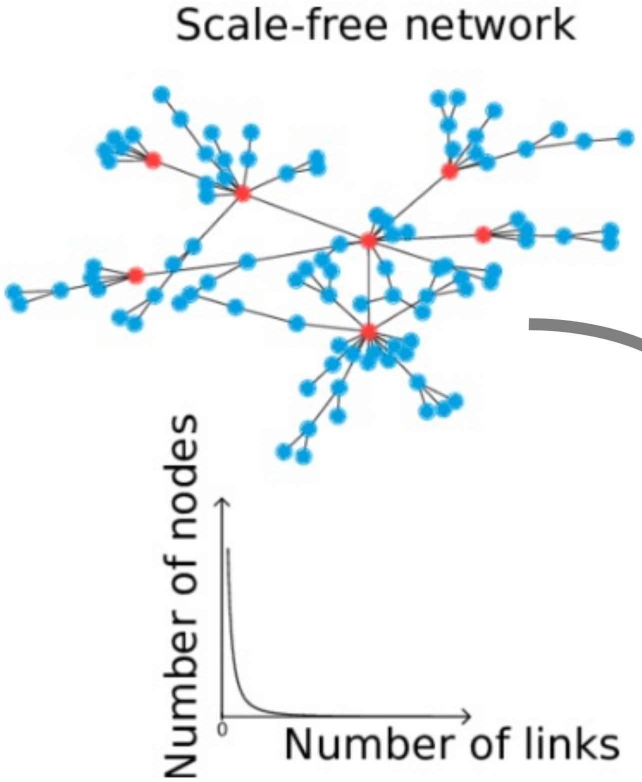
Protein folding and degradation

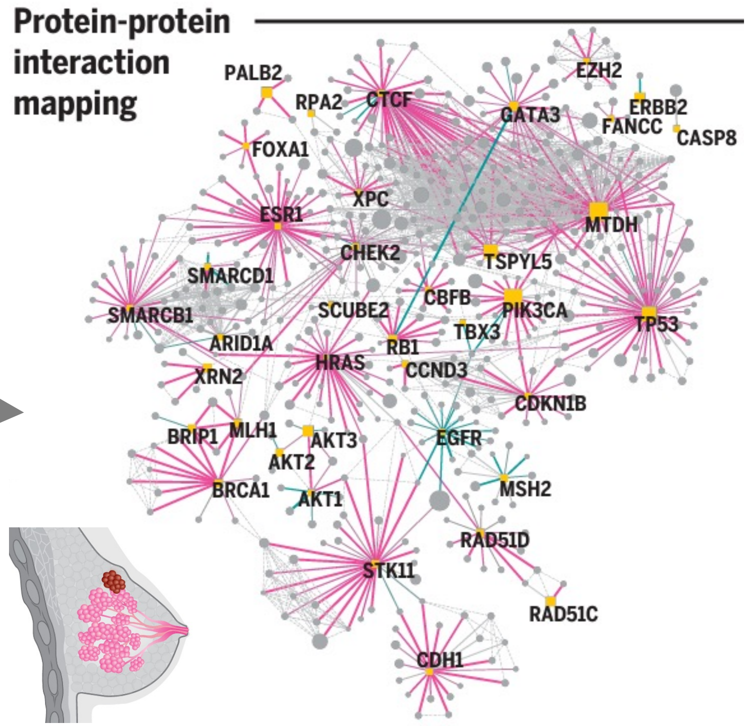
More than 80% of proteins need to interact with others

PPI-networks are scale-free

A few nodes with many connections, while most nodes have only a few connections. Like breast cancer related PPI network

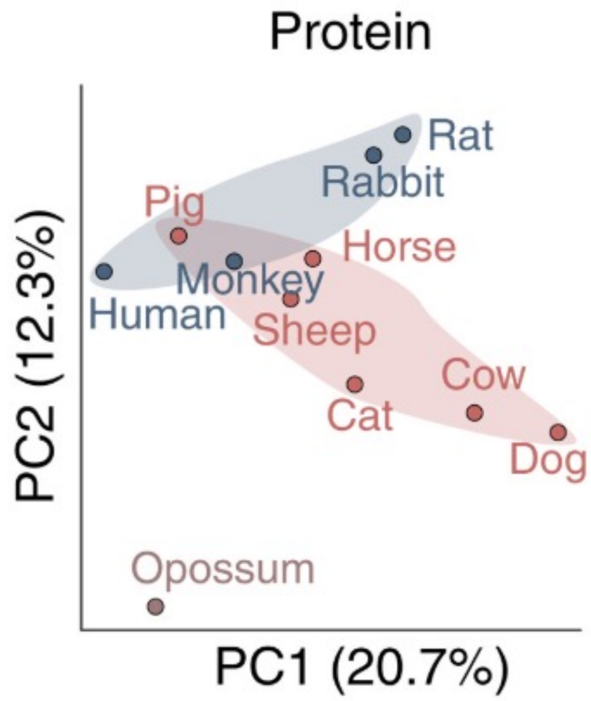




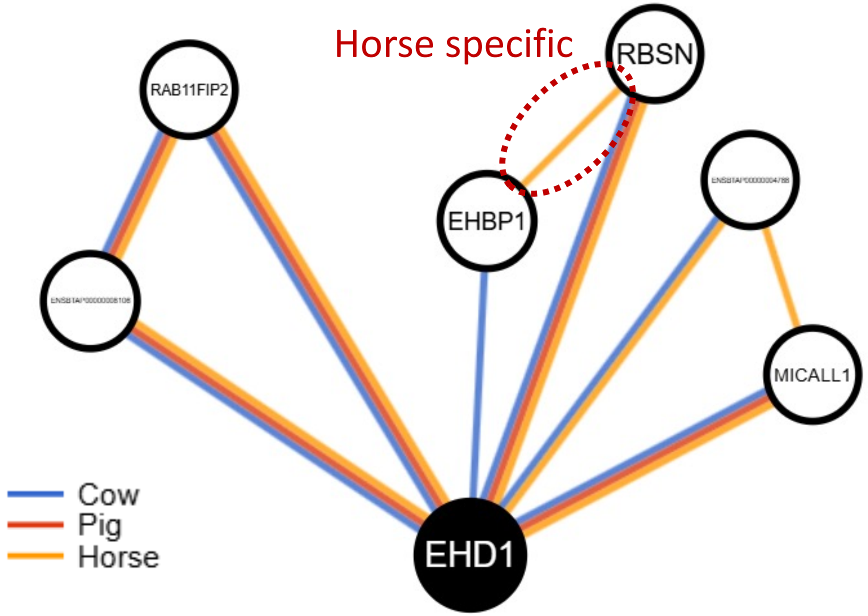


Minkyu Kim *et al.*, A protein interaction landscape of breast cancer.*Science* **374**,eabf3066(2021).

Different species, Different networks



Proteome of different species are quite different



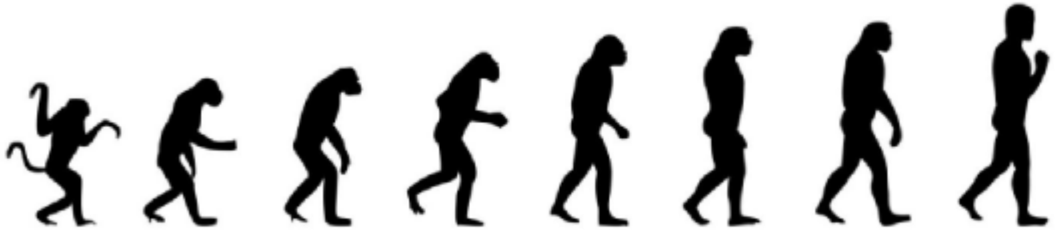
PPI network comparison

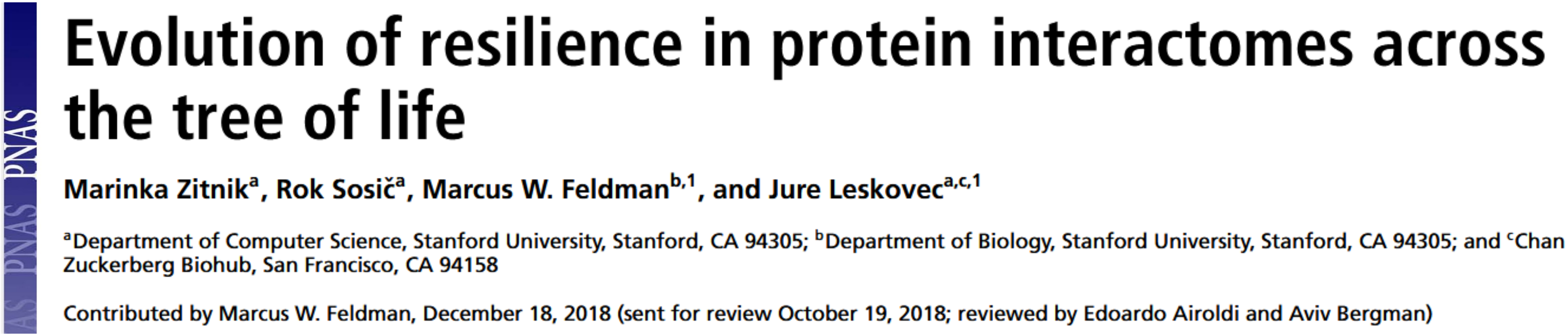
Qian Ba et al., Proteotype coevolution and quantitative diversity across 11 mammalian species. Sci. Adv.(2022).
Daehong Kwon et al., INTERSPIA: a web application for exploring the dynamics of protein-protein interactions among multiple species, Nucleic Acids Research.(2018).

PPI-network evolution:

Are PPI networks becoming more resilient during evolution?

How does the network structure change?



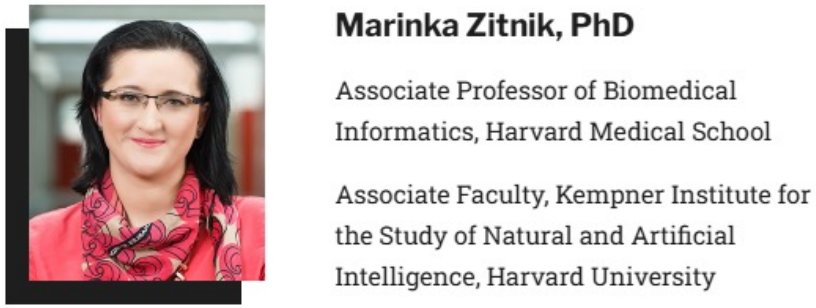


Evolution of resilience in protein interactomes across the tree of life

Marinka Zitnik[a], Rok Sosič[a], Marcus W. Feldman[b,1], and Jure Leskovec[a,c,1]

[a]Department of Computer Science, Stanford University, Stanford, CA 94305; [b]Department of Biology, Stanford University, Stanford, CA 94305; and [c]Chan Zuckerberg Biohub, San Francisco, CA 94158

Contributed by Marcus W. Feldman, December 18, 2018 (sent for review October 19, 2018; reviewed by Edoardo Airoldi and Aviv Bergman)



Marinka Zitnik, PhD

Associate Professor of Biomedical Informatics, Harvard Medical School

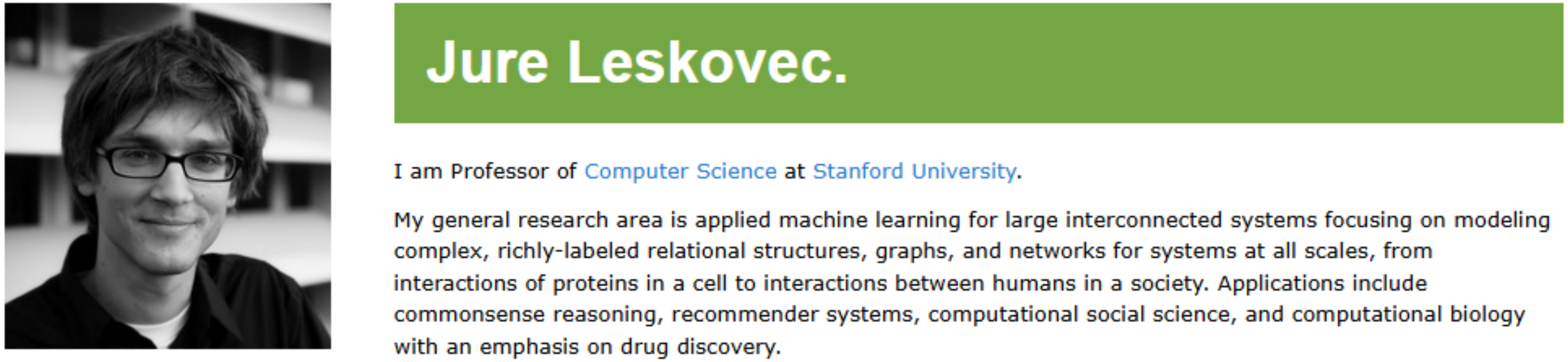
Associate Faculty, Kempner Institute for the Study of Natural and Artificial Intelligence, Harvard University

Marcus W. Feldman, MS, PhD

- Burnet C. and Mildred Finley Wohlford Professor of Biological Sciences
- Director of the Morrison Institute for Population and Resource Studies
- Stanford Health Policy Associate

428 Herrin Labs
Department of Biological Sciences
Stanford University
Stanford, CA 94305-5020

marc@charles.stanford.edu

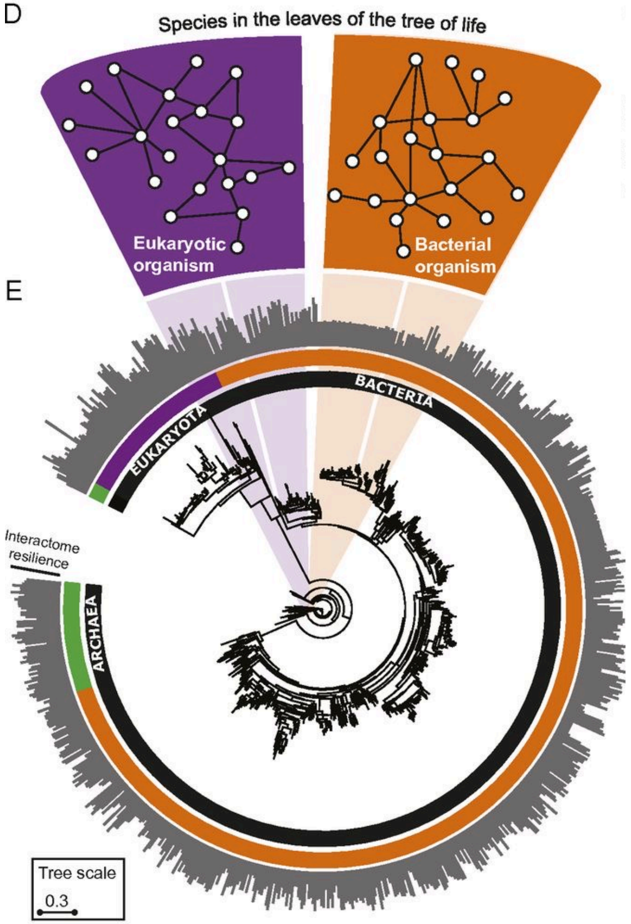


Jure Leskovec.

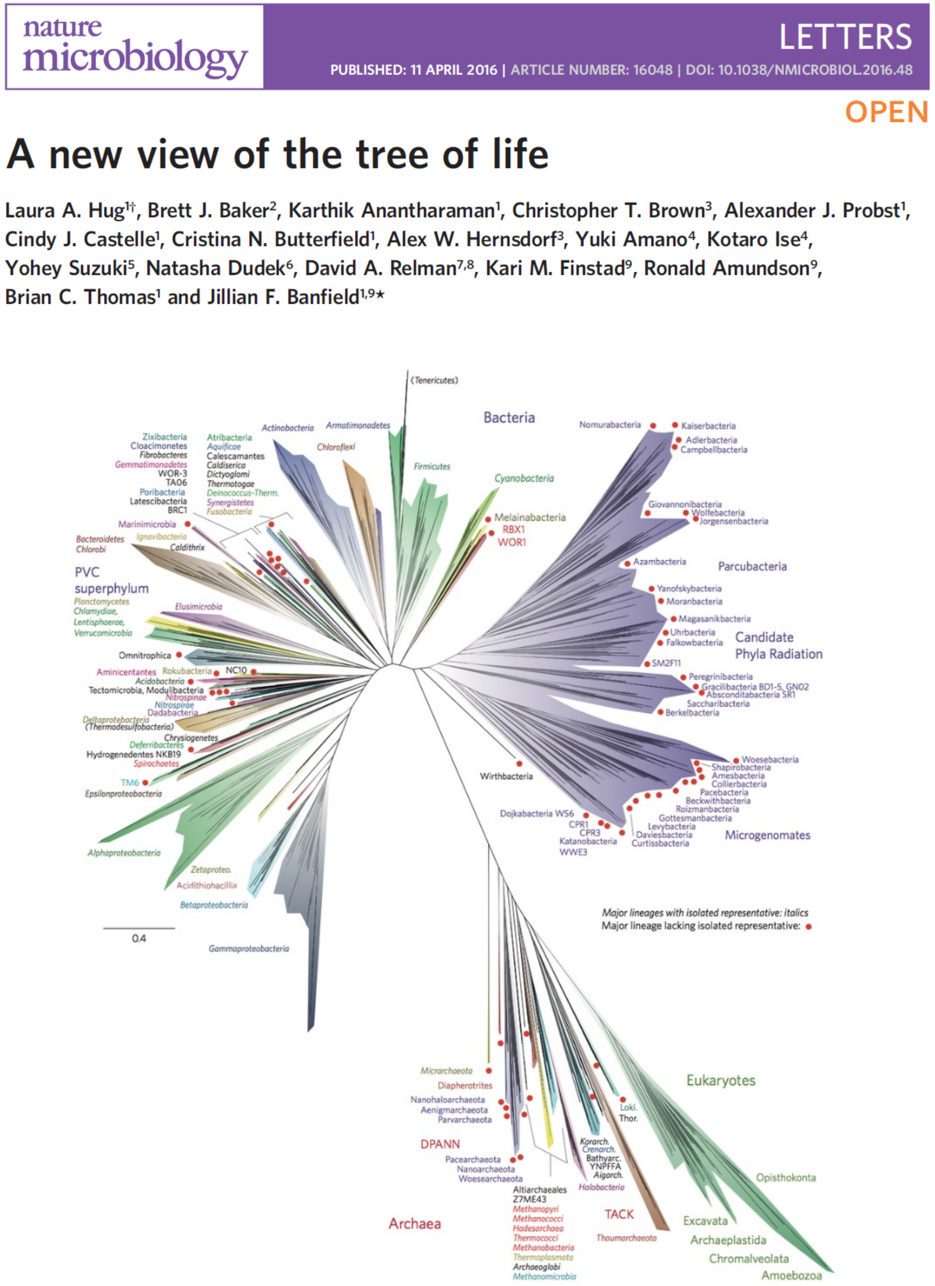
I am Professor of Computer Science at Stanford University.

My general research area is applied machine learning for large interconnected systems focusing on modeling complex, richly-labeled relational structures, graphs, and networks for systems at all scales, from interactions of proteins in a cell to interactions between humans in a society. Applications include commonsense reasoning, recommender systems, computational social science, and computational biology with an emphasis on drug discovery.

1. Modeling Resilience of the Interactome



For each species (1,840 in total) we have a PPI network


nature microbiology

LETTERS

PUBLISHED: 11 APRIL 2016 | ARTICLE NUMBER: 16048 | DOI: 10.1038/NMICROBIOL.2016.48

OPEN

A new view of the tree of life

Laura A. Hug[1†], Brett J. Baker[2], Karthik Anantharaman[1], Christopher T. Brown[3], Alexander J. Probst[1], Cindy J. Castelle[1], Cristina N. Butterfield[1], Alex W. Hernsdorf[3], Yuki Amano[4], Kotaro Ise[4], Yohey Suzuki[5], Natasha Dudek[6], David A. Relman[7,8], Kari M. Finstad[9], Ronald Amundson[9], Brian C. Thomas[1] and Jillian F. Banfield[1,9*]

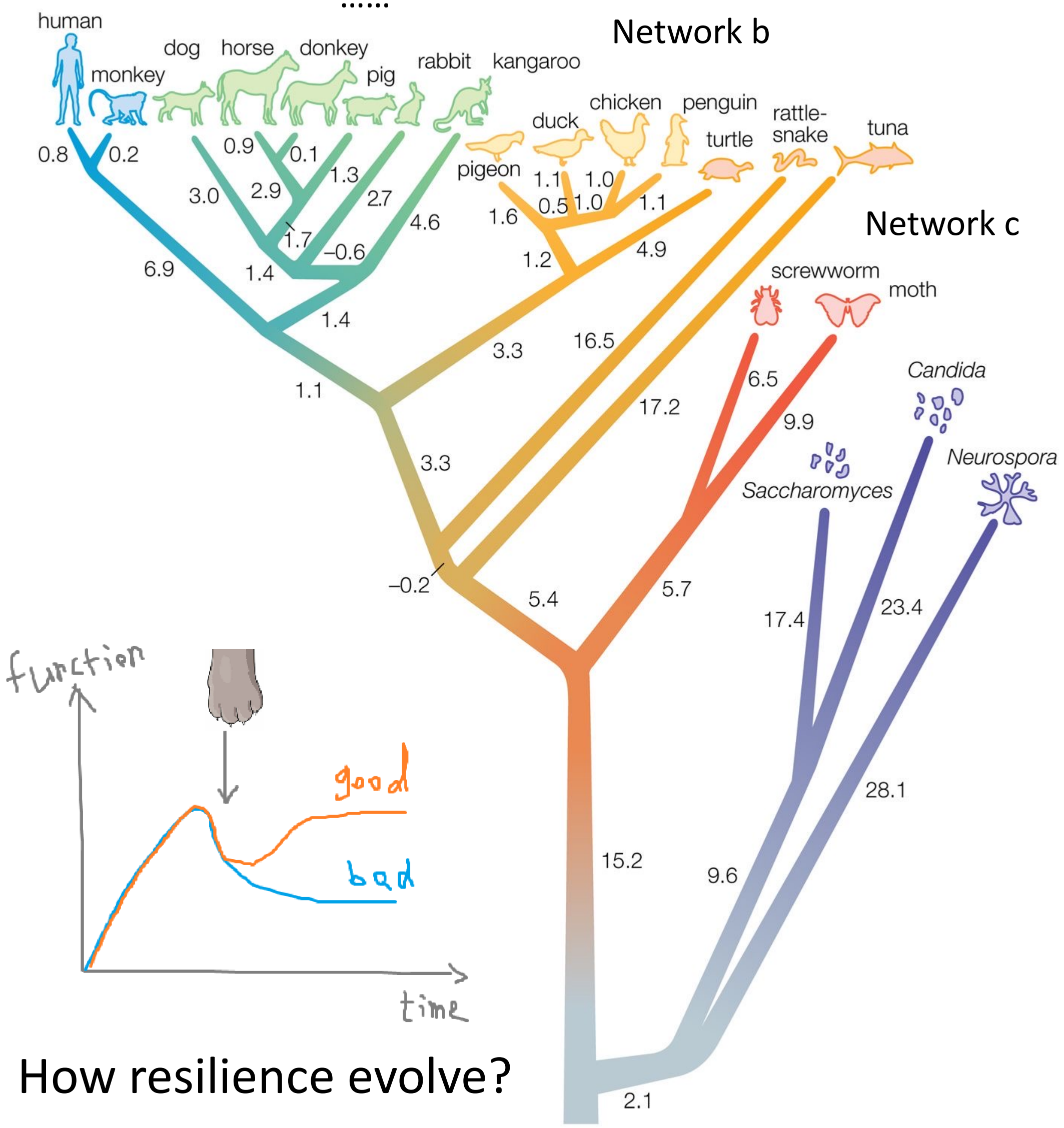
How to measure the resilience of a network?

Network a

......

Network b

Network c

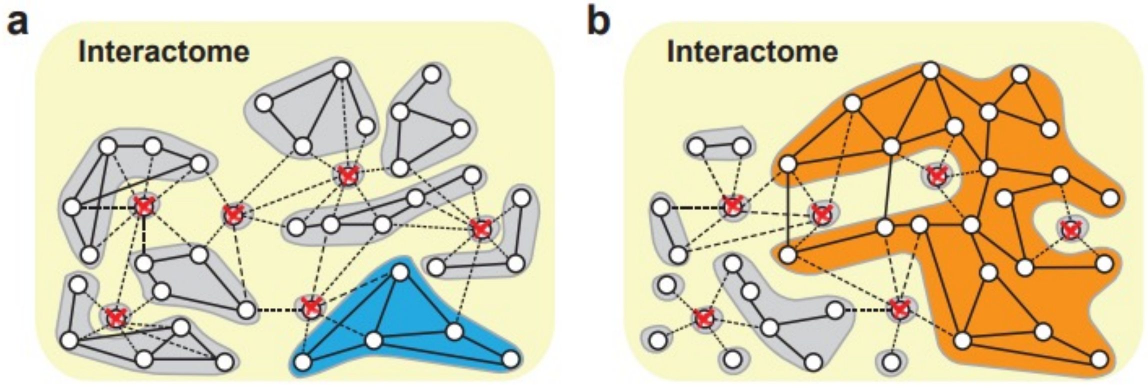


How resilience evolve?

Network a

Network b

remove some nodes



High Shannon diversity

A large number of small components, all of approximately equal size

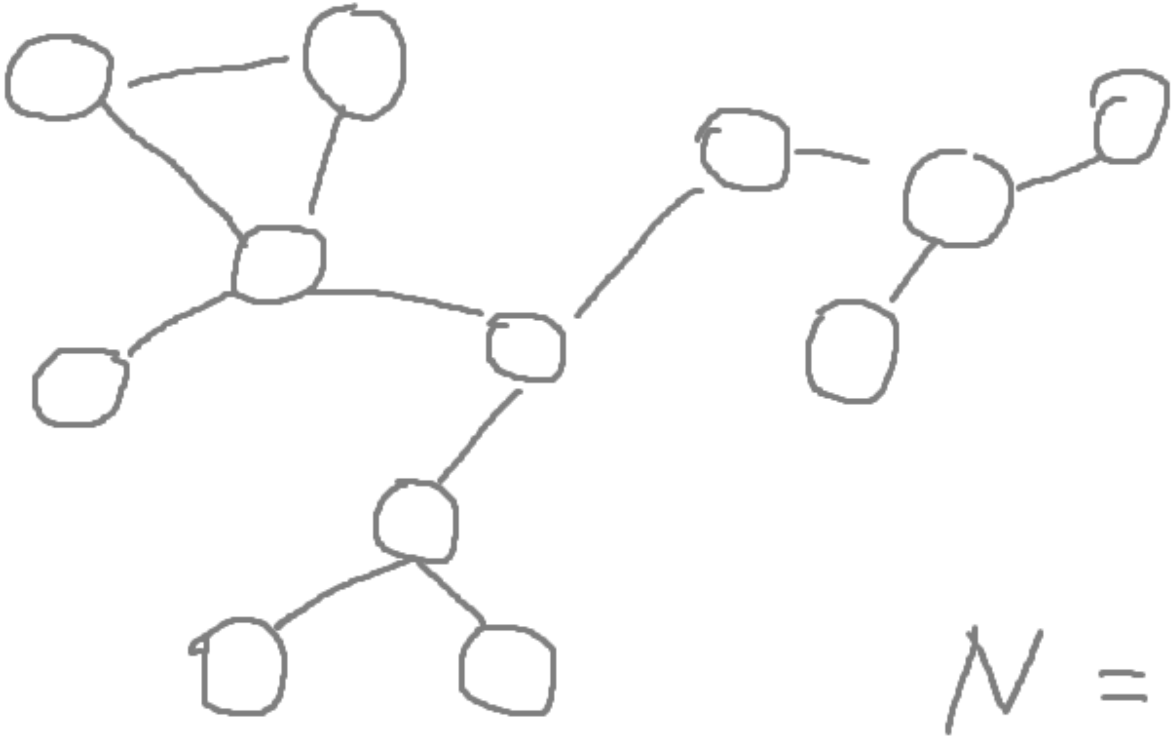
Low Shannon diversity

A few large components and and only a few small broken-off components

Low resilience

High resilience

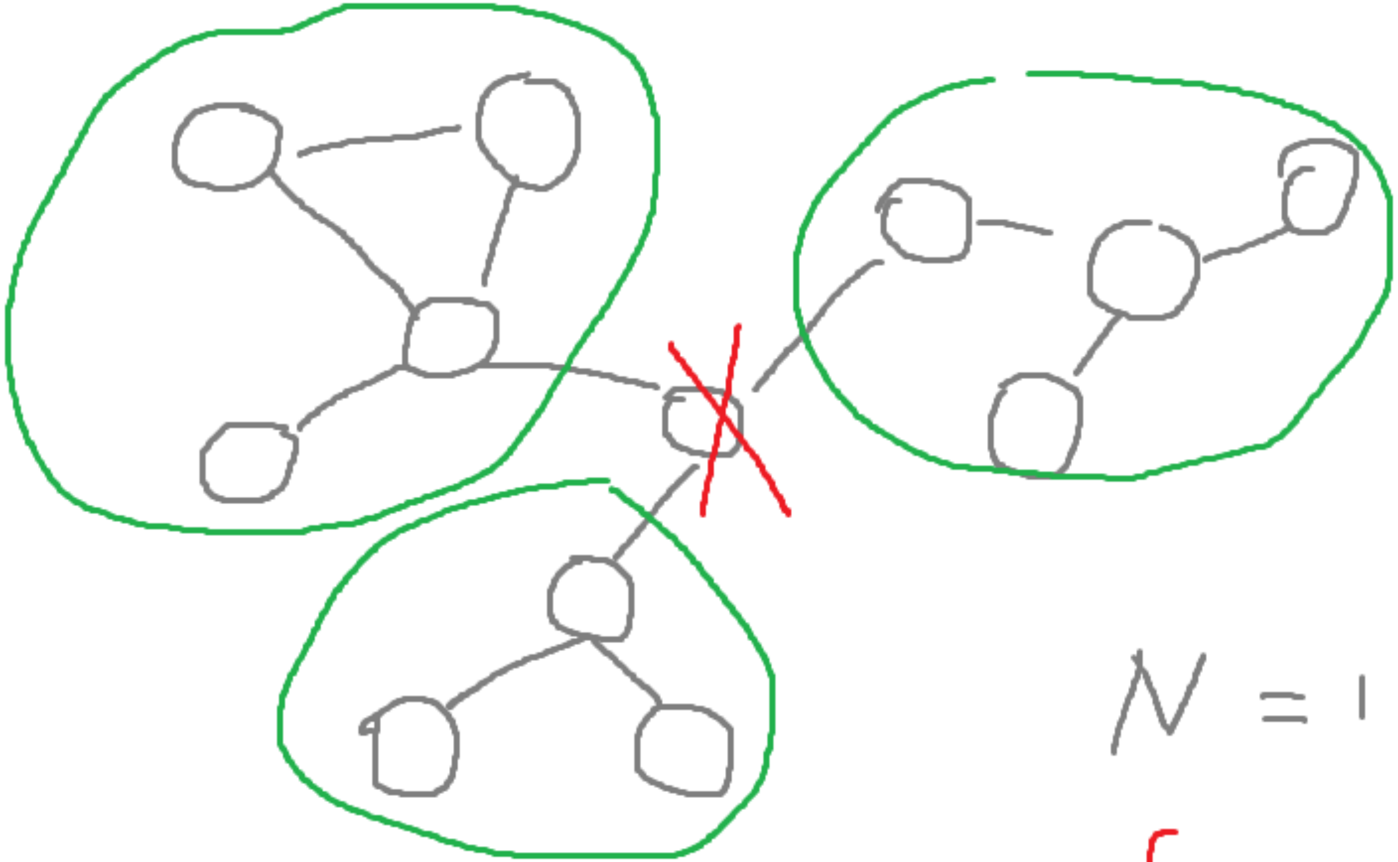
How to measure the resilience of a network?



$N = 12$

Given a species s, whose network has N nodes

How to measure the resilience of a network?

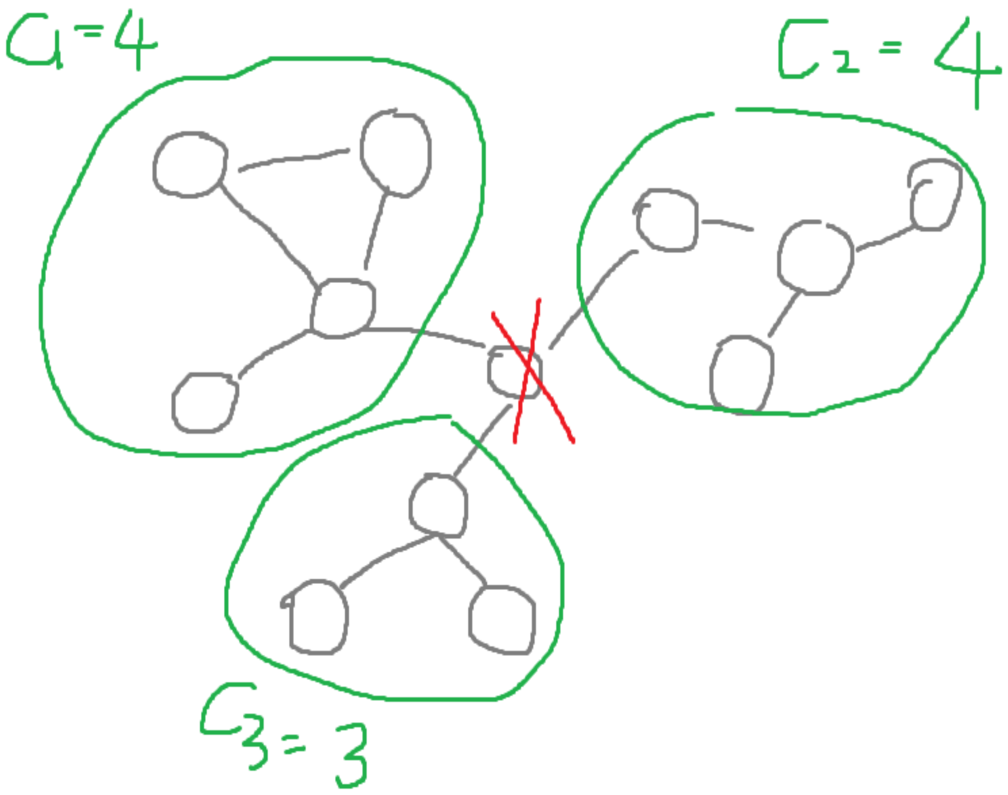


$N = 12$

$f = \frac{1}{12}$

Given a species s, whose network has N nodes
f ∈ [0, 1] denote network failure rate

How to measure the resilience of a network?



$N = 12$

$f = \frac{1}{12}$

$p_1 = \frac{4}{12}$

$p_2 = \frac{4}{12}$

$p_3 = \frac{3}{12}$

$$H(G_f^{(s)}) = -\sum_{i=1}^{k} p_i \log p_i$$

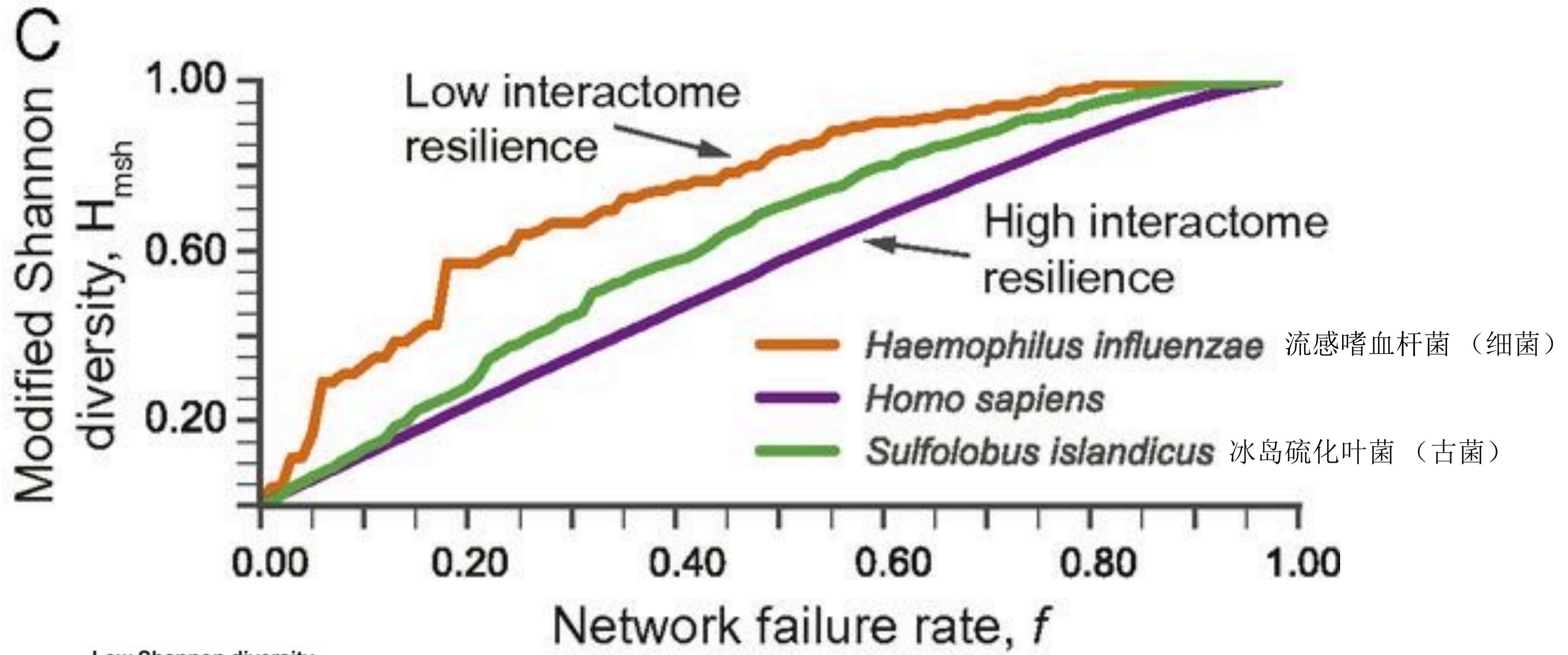
$$H_{\text{msh}}(G_f^{(s)}) = H(G_f^{(s)}) / \log N.$$

Resilience

Given a species s, whose network has N nodes.
$f \in [0, 1]$ denote network failure rate
$\{C_1, C_2, \ldots, C_k\}$ be k isolated components in the fragmented network $G_f^{(s)}$
$p_i = C_i/N$ is the proportion of nodes belonging to component C_i

Then repeat it many times
Use every possible f
And do it for every species

1. 1 Modeling Resilience of the Interactome



High Shannon diversity

A large number of small components, all of approximately equal size

↓

Low resilience

Low Shannon diversity

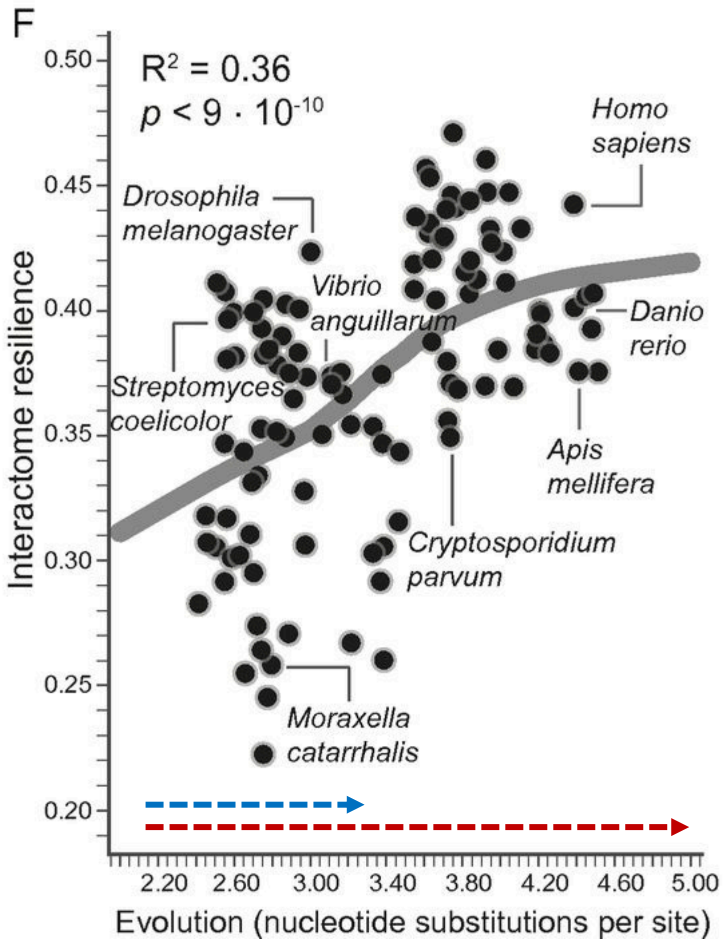
A few large components and and only a few small broken-off components

↓

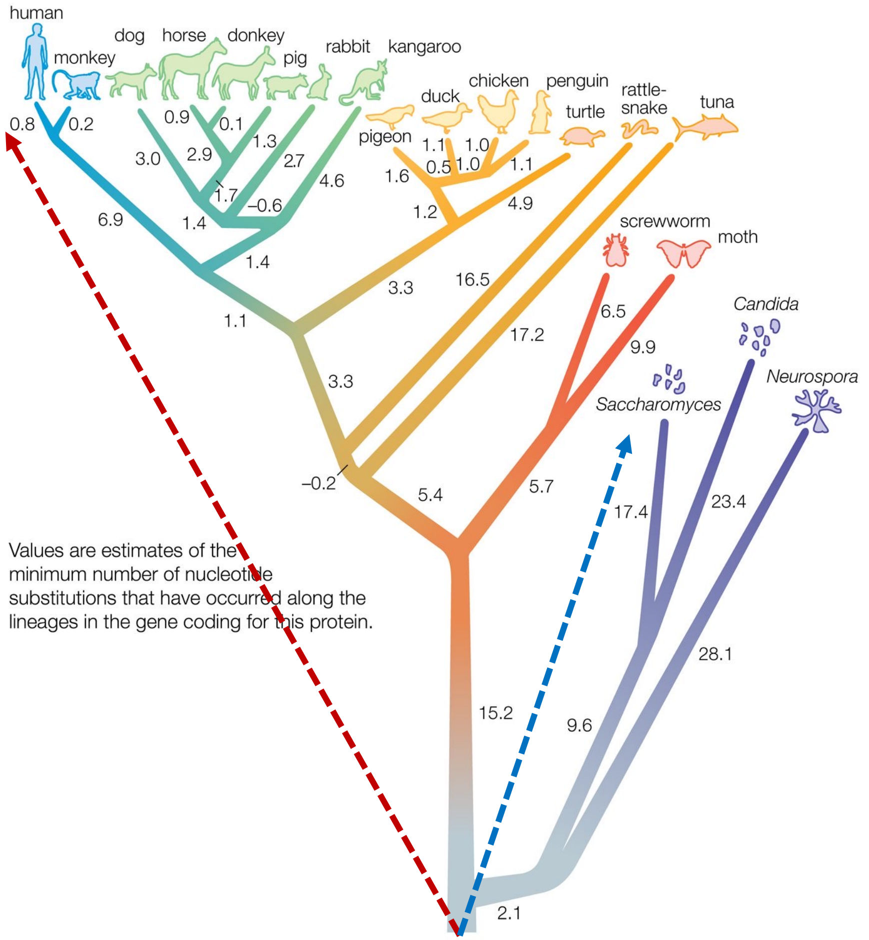
High resilience

(C) The resilience of the interactome integrates modified Shannon diversity H across all possible failure rates f. Resilience value 1 indicates the most resilient interactome, and resilience value 0 indicates a complete loss of the connectivity of the interactome

1.2 Resilience of Interactomes Throughout Evolution



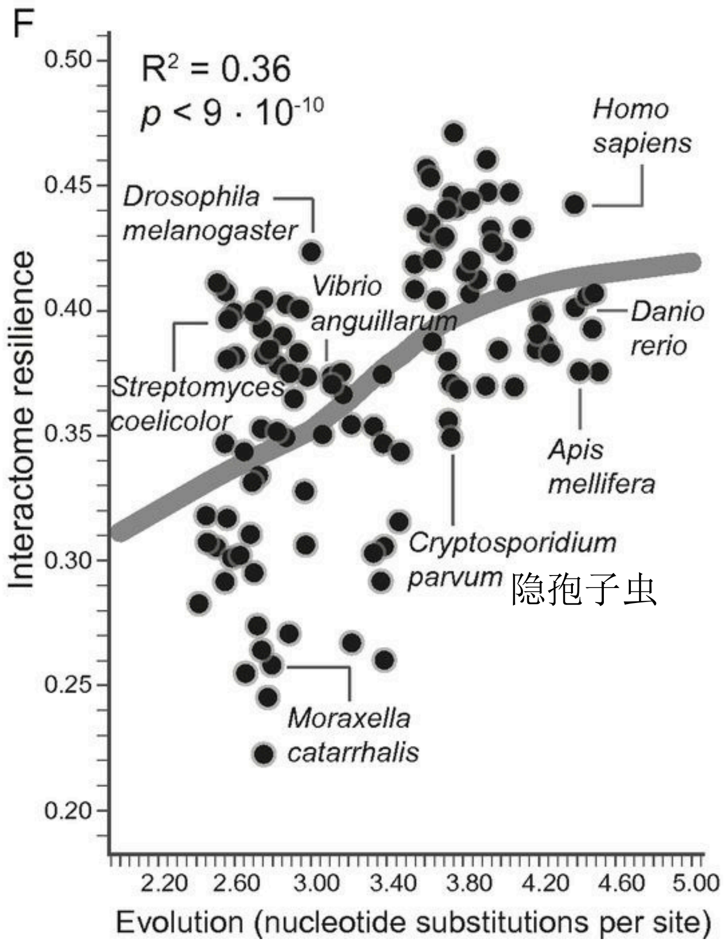
(F) interactome resilience for 171 species. More genetic change implies a more resilient interactome



Values are estimates of the minimum number of nucleotide substitutions that have occurred along the lineages in the gene coding for this protein.

“Given a species s, its evolution ts is calculated as the total **branch** **length** (i.e., nucleotide substitutions per site) from the root of the tree to the leaf representing species s.”

1.2 Resilience of Interactomes Throughout Evolution



(F) interactome resilience for 171. More genetic change implies a more resilient interactome

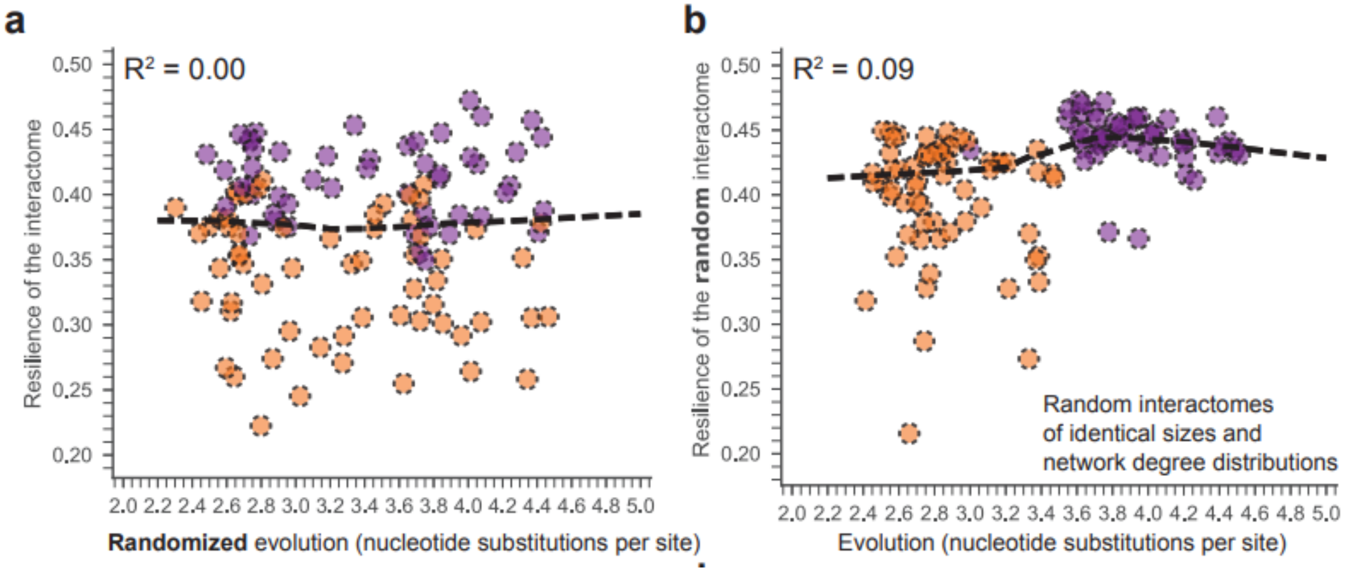
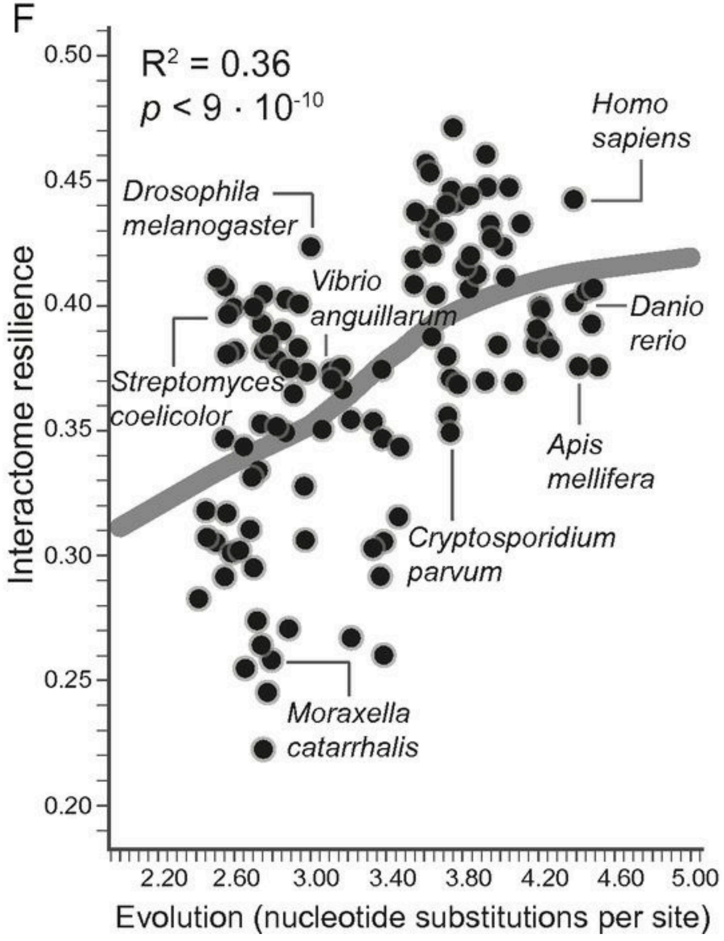


Figure S9: Relationship between evolution and interactome resilience under random expectation.

1.2 Resilience of Interactomes Throughout Evolution



(F) interactome resilience for 171. More genetic change implies a more resilient interactome

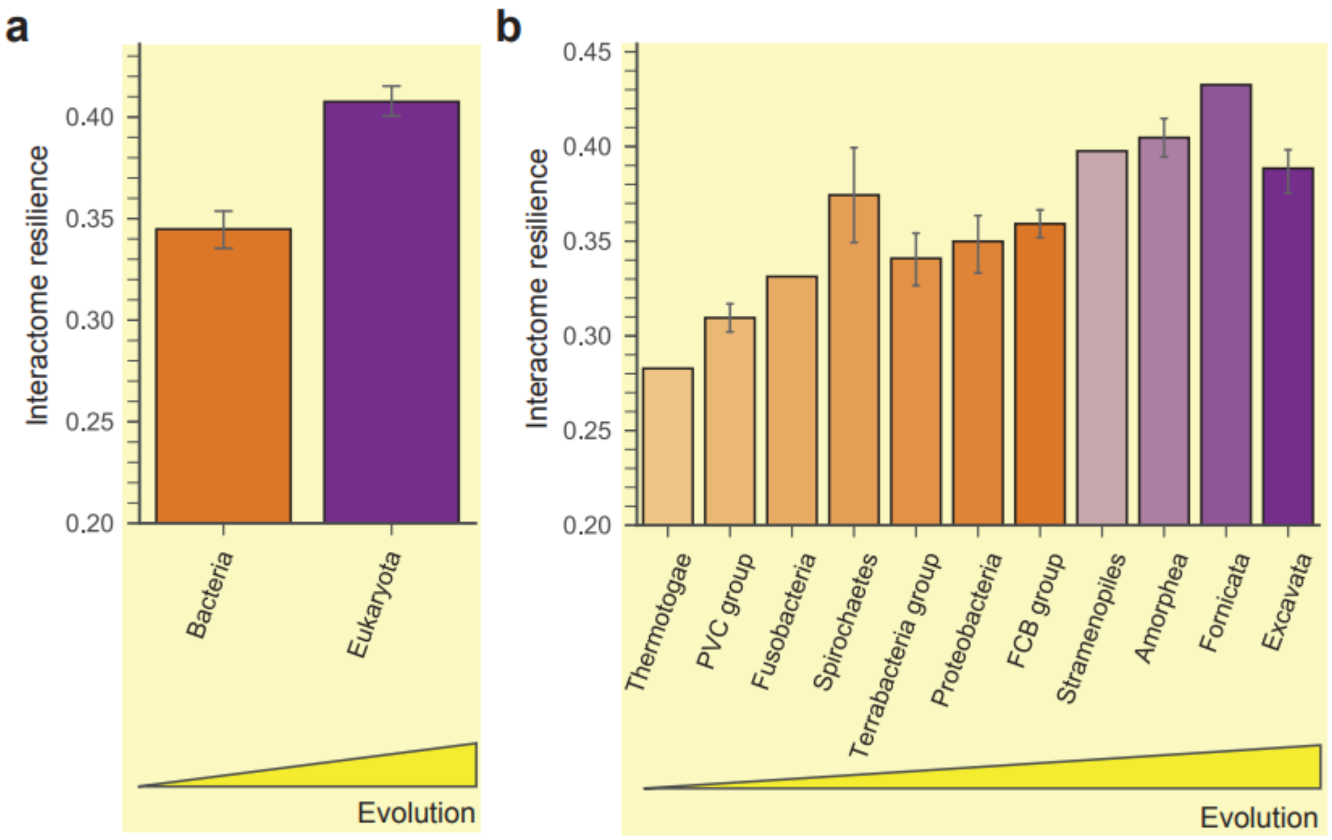
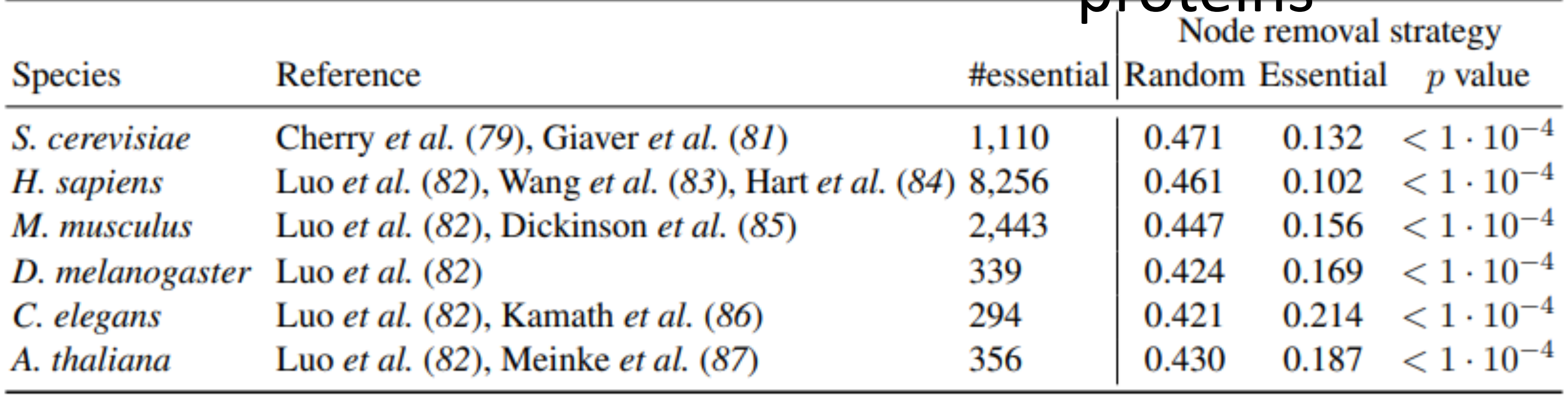


Figure S10: Interactome resilience for species from the same taxonomic groups.

1.3 Resilience of Interactomes depend highly on essential proteins



Species	Reference	#essential	Node removal strategy Random	 Essential	 p value
S. cerevisiae	Cherry *et al.* (*79*), Giaver *et al.* (*81*)	1,110	0.471	0.132	$< 1 \cdot 10^{-4}$
H. sapiens	Luo *et al.* (*82*), Wang *et al.* (*83*), Hart *et al.* (*84*)	8,256	0.461	0.102	$< 1 \cdot 10^{-4}$
M. musculus	Luo *et al.* (*82*), Dickinson *et al.* (*85*)	2,443	0.447	0.156	$< 1 \cdot 10^{-4}$
D. melanogaster	Luo *et al.* (*82*)	339	0.424	0.169	$< 1 \cdot 10^{-4}$
C. elegans	Luo *et al.* (*82*), Kamath *et al.* (*86*)	294	0.421	0.214	$< 1 \cdot 10^{-4}$
A. thaliana	Luo *et al.* (*82*), Meinke *et al.* (*87*)	356	0.430	0.187	$< 1 \cdot 10^{-4}$

Table S3: Resilience of species' interactomes to network failure of essential protein-coding genes

~~Random~~ Guided removal / add :
Some nodes are more important

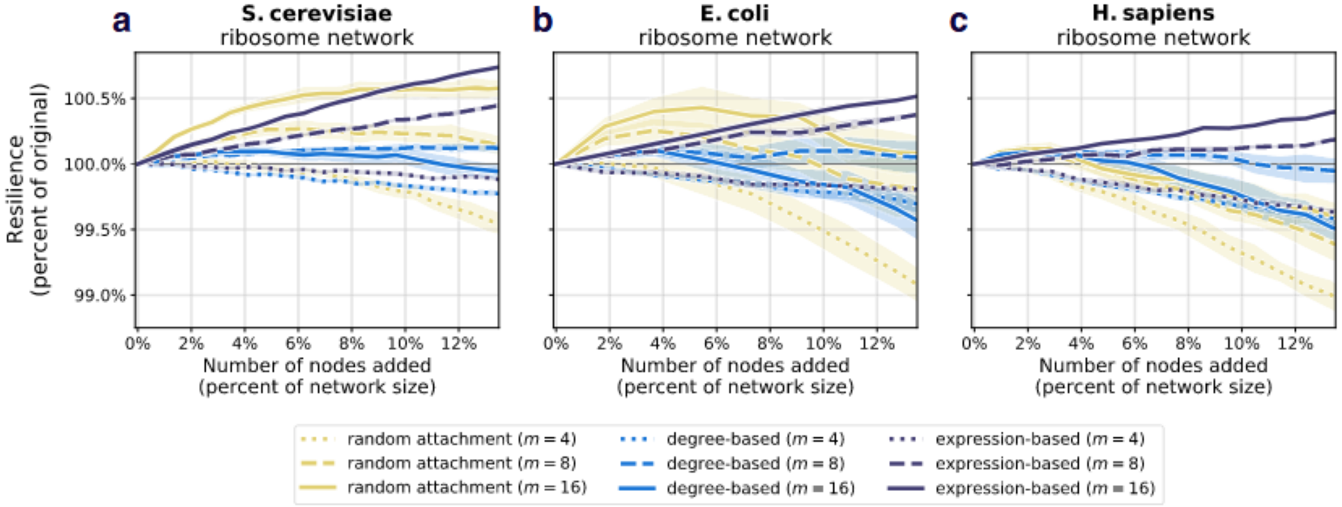


Fig. 4 Prospective resilience of three ribosomal networks. As more nodes are added (horizontal axes), the resilience of the resulting network changes (vertical axes). The color of each curve corresponds to the number of new links that each new node enters the network with, and the line style (solid, dashed, or dotted) corresponds to the three different node attachment mechanisms. **a** Prospective resilience of *S. cerevisiae* ribosomal network. **b** Prospective resilience of *E. coli* ribosomal network. **c** Prospective resilience of *H. sapiens* ribosomal network. Ribbons around each curve correspond to their 95% confidence intervals.

communications biology

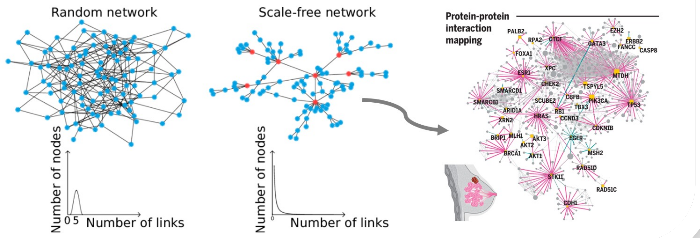
https://doi.org/10.1038/s42003-021-02867-8 OPEN

A computational exploration of resilience and evolvability of protein–protein interaction networks

Brennan Klein[1,2]✉, Ludvig Holmér[3], Keith M. Smith[4]✉, Mackenzie M. Johnson[5], Anshuman Swain[6], Laura Stolp[7], Ashley I. Teufel[5,8,9] & April S. Kleppe[10,11]✉

PPI-networks are scale-free

A few nodes with many connections, while most nodes have only a few connections. Like breast cancer related PPI network



1.4 Relationship Between Resilience and Ecology

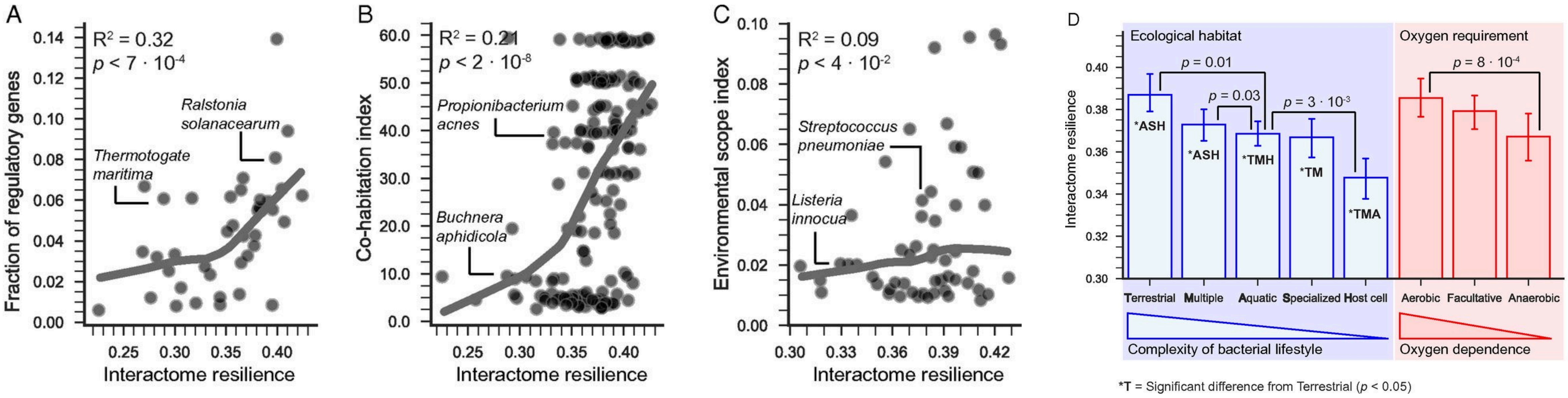
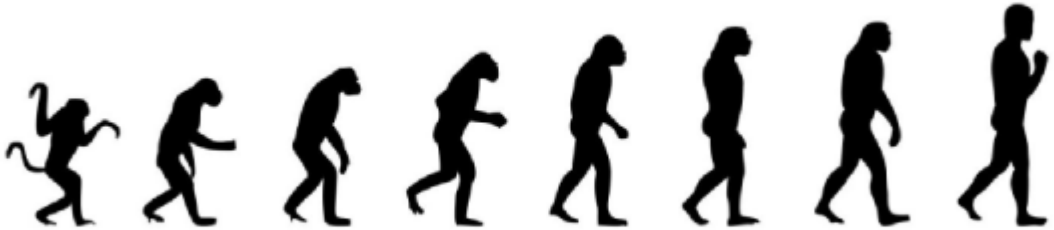


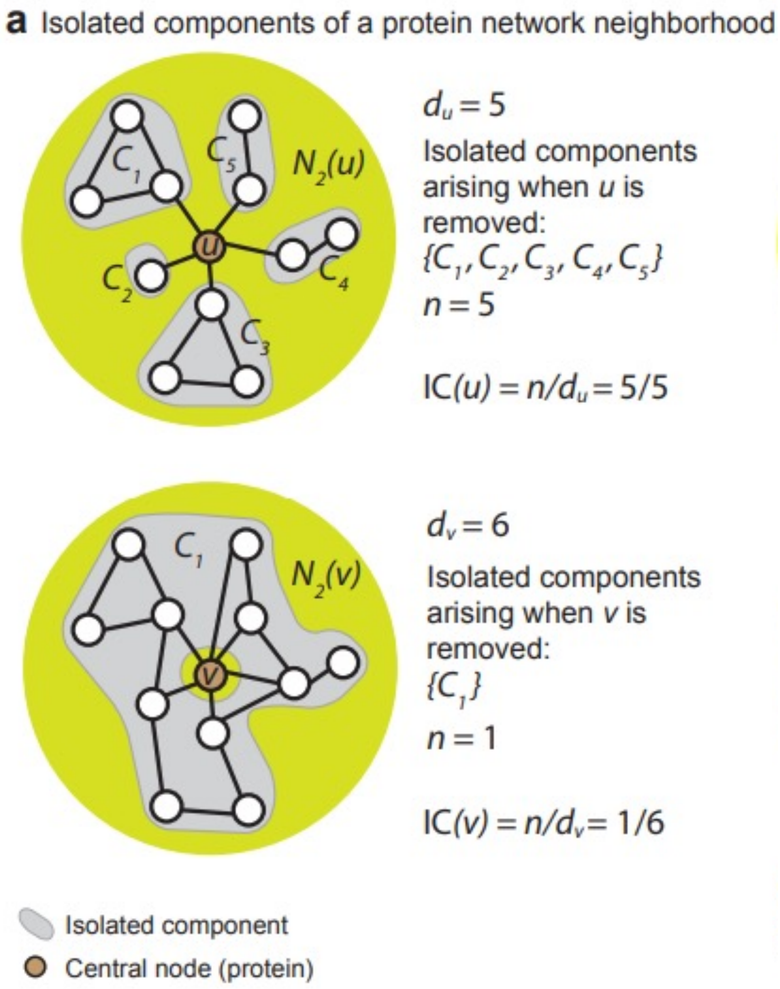
Figure 2.Bacteria with more resilient interactomes survive in more complex, variable, and competitive environments.

More regulatory genes (A), more cohabitation (B), wider range (C), more complex lifestyle (D)
More resilient

PPI-network evolution:
Are PPI networks becoming more resilient during evolution?
How does the network structure change?



2.1 Structural Changes of Protein Network Neighborhoods



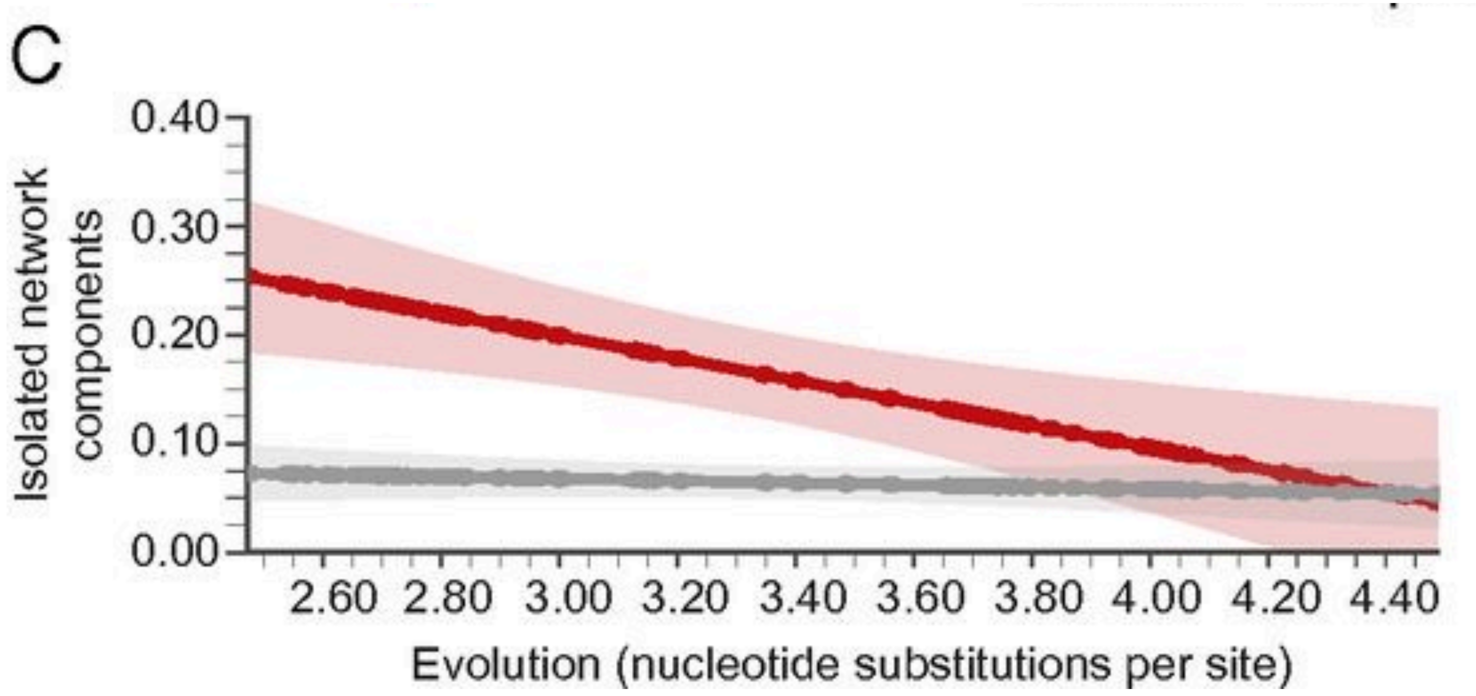
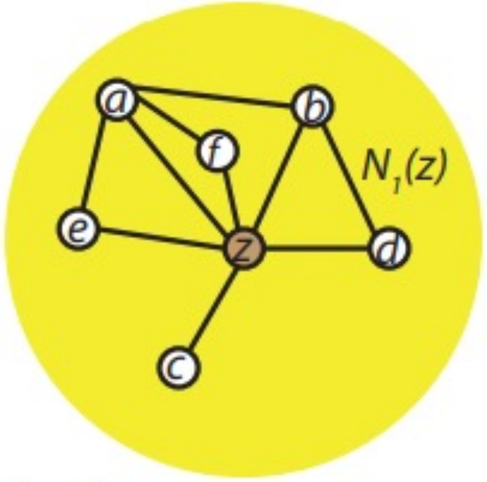


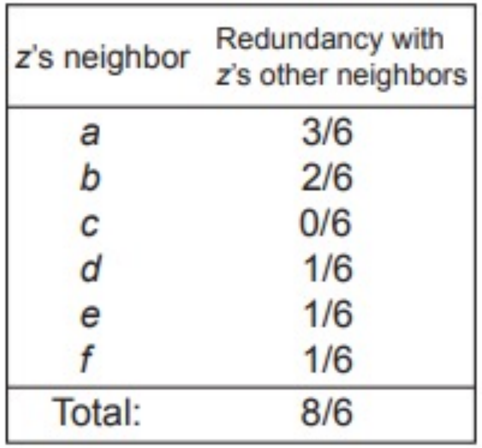
Figure 3C : The number of isolated network components decrease with evolution

2.1 Structural Changes of Protein Network Neighborhoods



$d_z = 6$

$ES(z) = \sum_{i \in N_1(z)} (1 - 1/d_z \sum_{j \in N_1(z)} e_{ij})$



z's neighbor	Redundancy with z's other neighbors
a	3/6
b	2/6
c	0/6
d	1/6
e	1/6
f	1/6
Total:	8/6

$ES(z) = 6 - 8/6 = 4.67$

$ES(z)/d_z = 0.78$

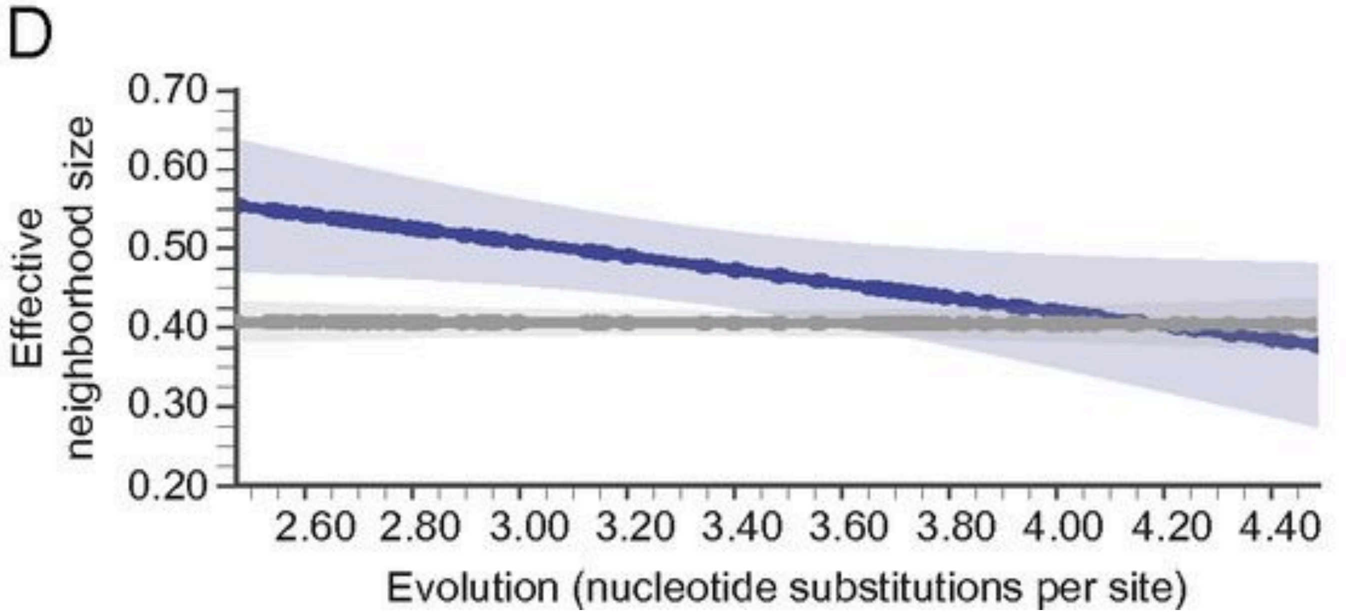
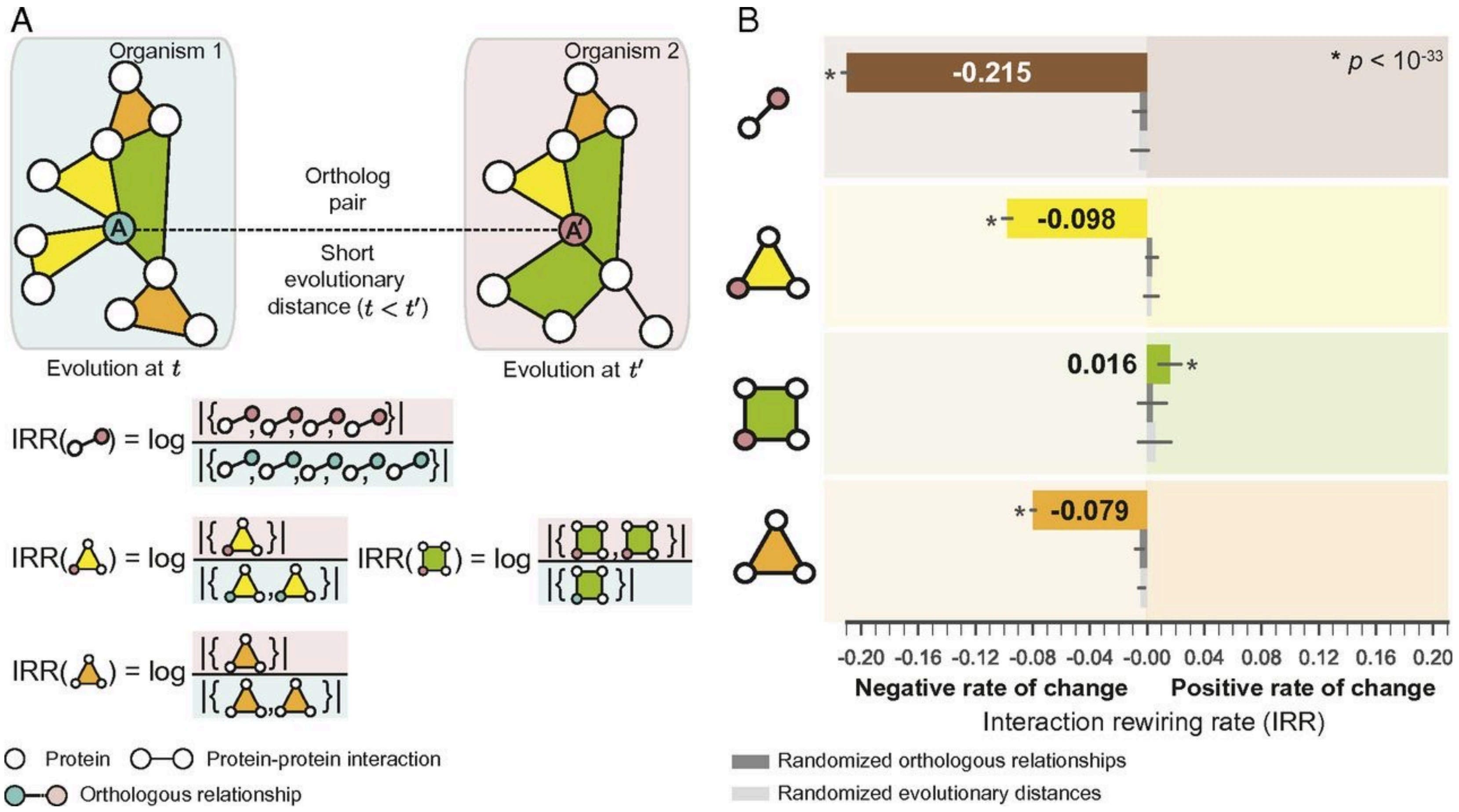


Figure 3D: The number of the effective size of protein neighborhoods decrease with evolution

The effective size of z's network neighborhood is the sum of the non-redundant portion of z's connections over all z's neighbors

2.2 Network Rewiring of Protein–Protein Interactions



The rewiring rate of interactions in local protein neighborhoods varies with the topology of network motifs.

network motifs rewire during evolution

Summary

1. Species that have undergone more genetic changes tend to have more resilient interactomes.

2. This resilience is positively correlated with the complexity of their ecological environments.

3. PPI network structure changed through gradual rewiring and became more efficient and compact.

Comments

1. Unbalanced data may be a problem.

2. The calculated resilience is a little bit far from the “resilience” in reality. Are there more biological metrics available to use to describe the possible changes of a network? “evolution of networks” seems still a topic for physicist instead of biologist.

3. They treat proteins in different species as if they were the same(it’s inevitable when you try to compare the networks). But is it possible to take more detailed information (eg. Mutations) into account? Additionally, the weight or importance of different edges can also be considered in the analysis.

PPI-data are extremely unbalanced

Organism	Experiment Type	Raw Interactions
Homo sapiens	PHYSICAL	1,293,309
Mus musculus	PHYSICAL	100,464
Rattus norvegicus	PHYSICAL	10,590
Bos taurus	PHYSICAL	653
Canis familiaris	PHYSICAL	558
Oryctolagus cuniculus	PHYSICAL	385
Macaca mulatta	PHYSICAL	316
Sus scrofa	PHYSICAL	118
Cricetulus griseus	PHYSICAL	67
Chlorocebus sabaeus	PHYSICAL	60
Pan troglodytes	PHYSICAL	49
Cavia porcellus	PHYSICAL	10
Equus caballus	PHYSICAL	4
Monodelphis domestica	PHYSICAL	4
Ovis aries	PHYSICAL	1
Myotis lucifugus	PHYSICAL	1
Felis Catus	PHYSICAL	5
All mammals	PHYSICAL	1,406,594
All Organisms	PHYSICAL	1,886,614

Thank you
Q & A time