Computational Intractability Generates the Topology of Biological Networks

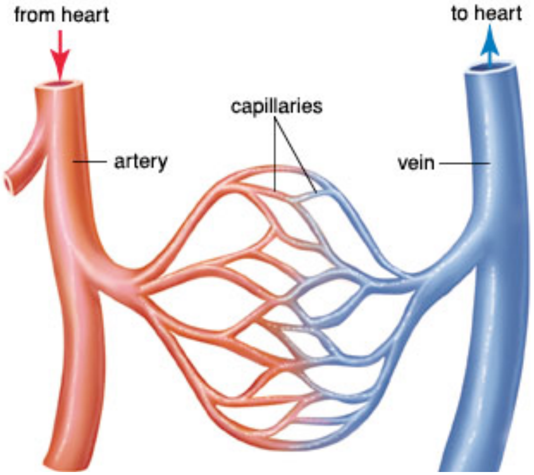
Ali Atiia

atiia@cs.mcgill.ca

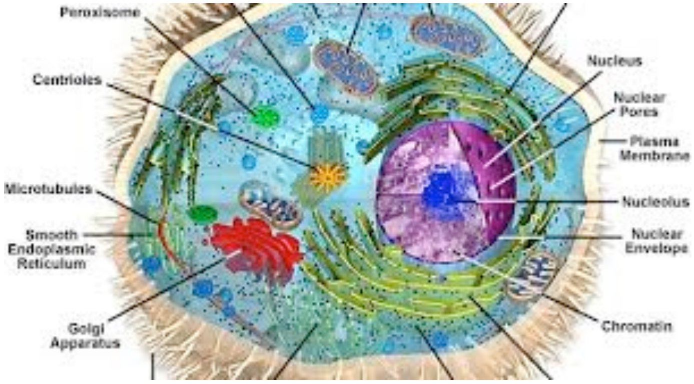
McGill University
Montreal, Canada

The signs of hardware natural selection are everywhere:

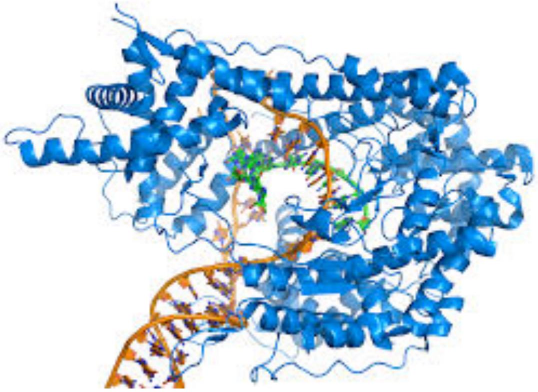
Allometry



Sturdy material

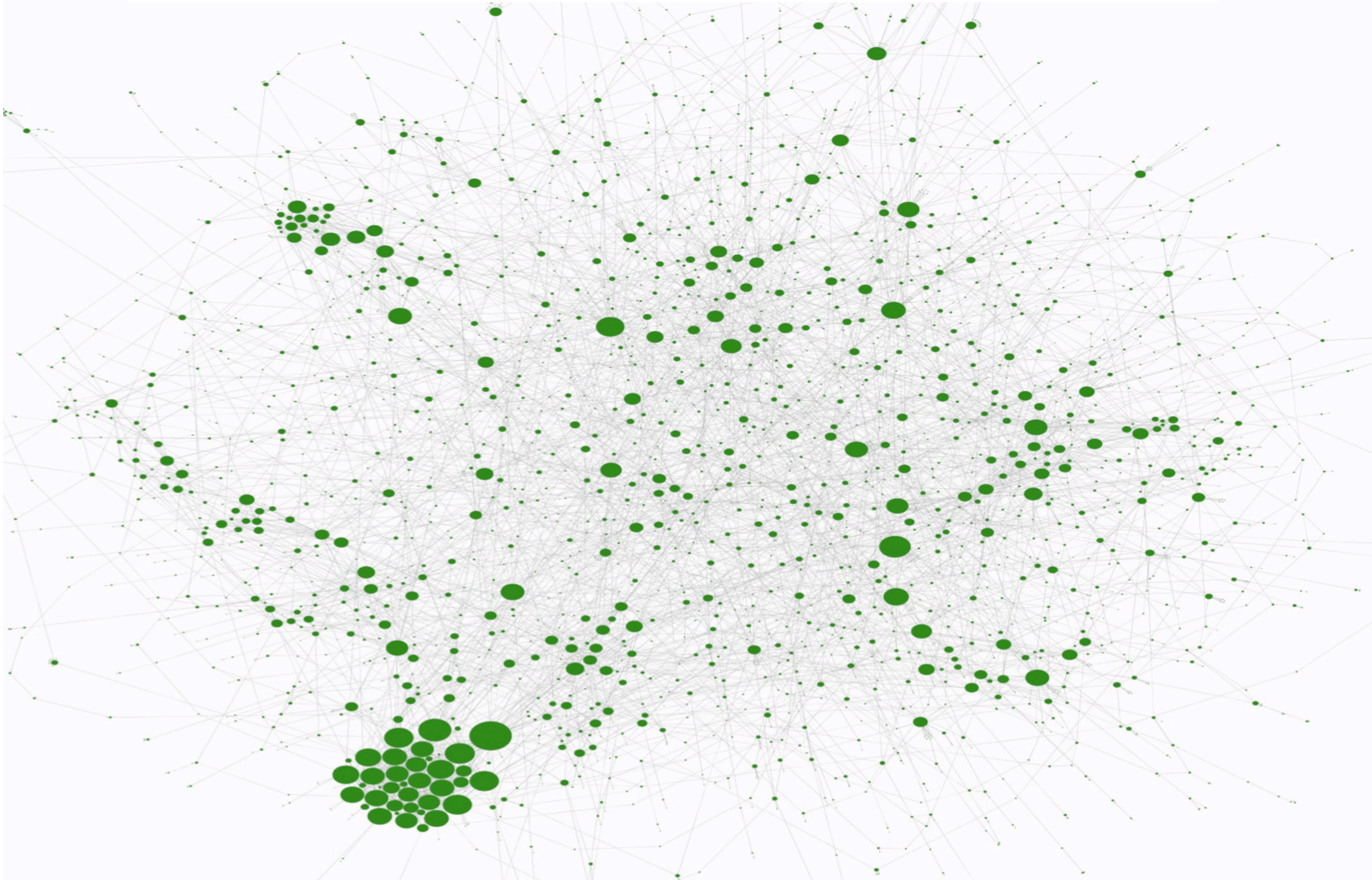


Cell diameter



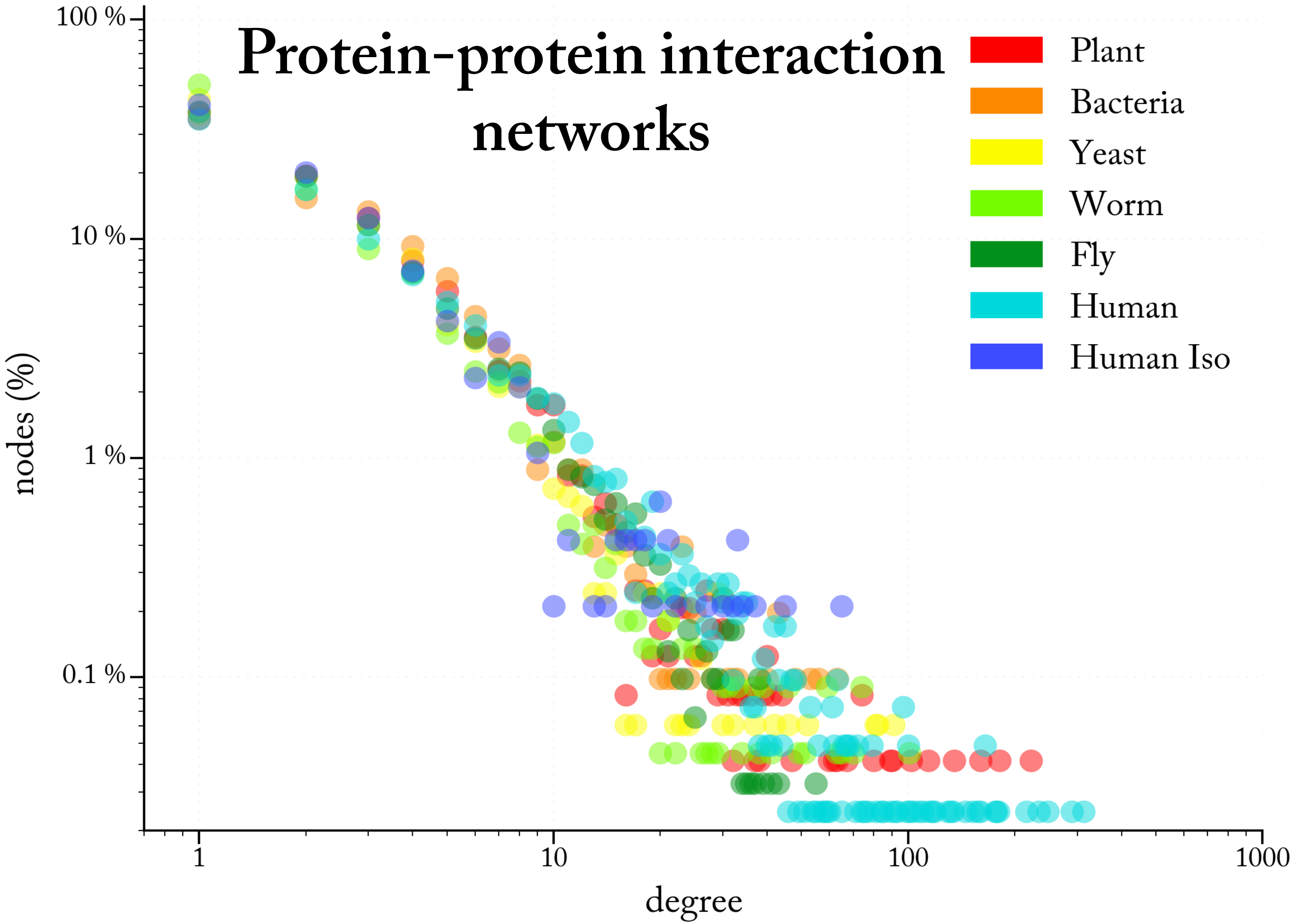
Stable folding at extreme temperature

Does natural selection act on software?

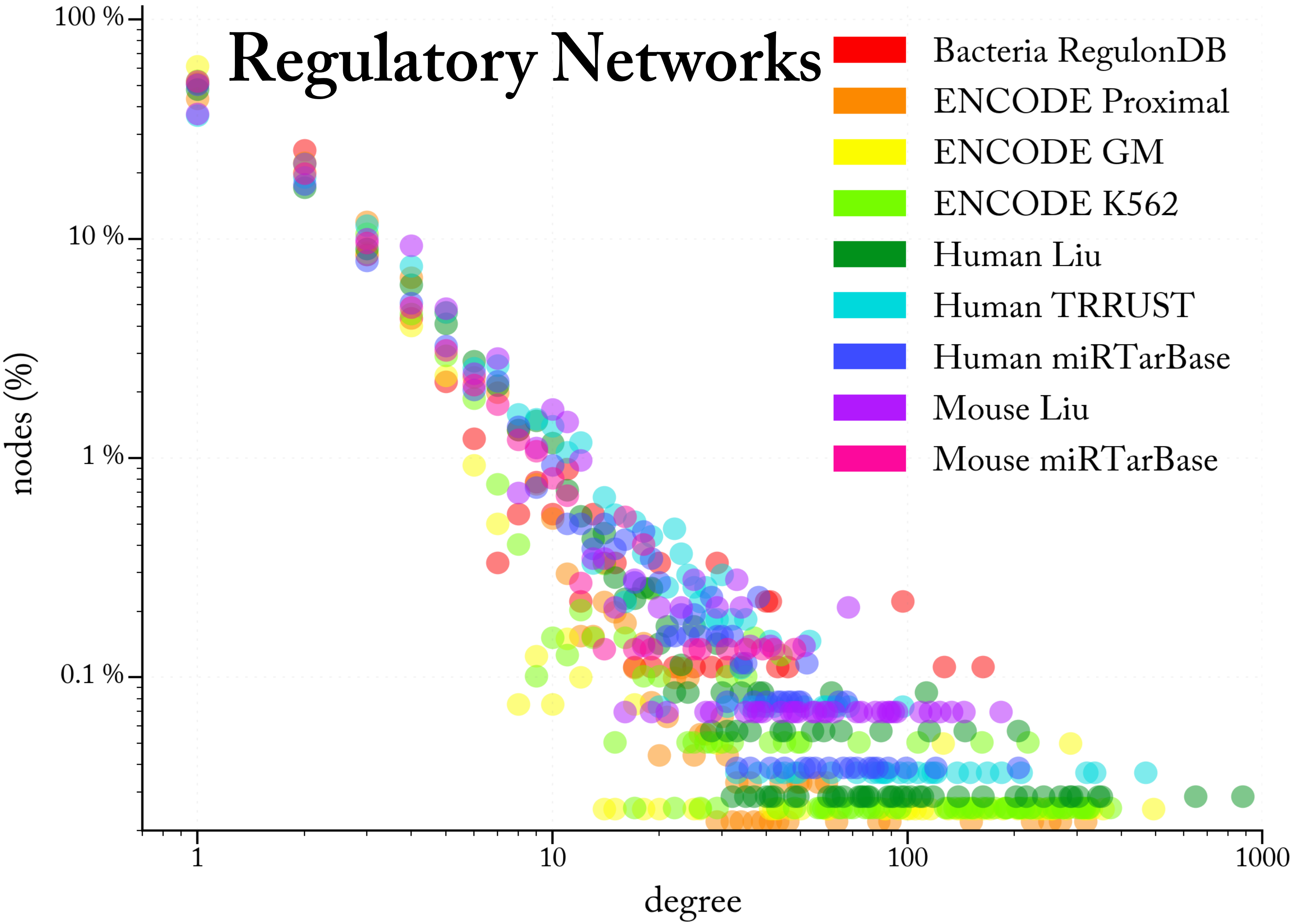


“software” = “connectivity” between bio-molecules

Majority-leaves Minority-hubs (**mLmH**) topology



Majority-leaves Minority-hubs (**mLmH**) topology



Majority-leaves Minority-hubs (**mLmH**) topology

Is natural selection behind the mLmH property?

Majority-leaves Minority-hubs (**mLmH**) topology

Is natural selection behind the mLmH property?

Our approach:

1. Model the evolution of biological networks as a computational problem.
2. This problem is fundamentally hard (NP-hard).
3. Show empirical evidence that mLmH produces easy instances of this problem

Majority-leaves Minority-hubs (**mLmH**) topology

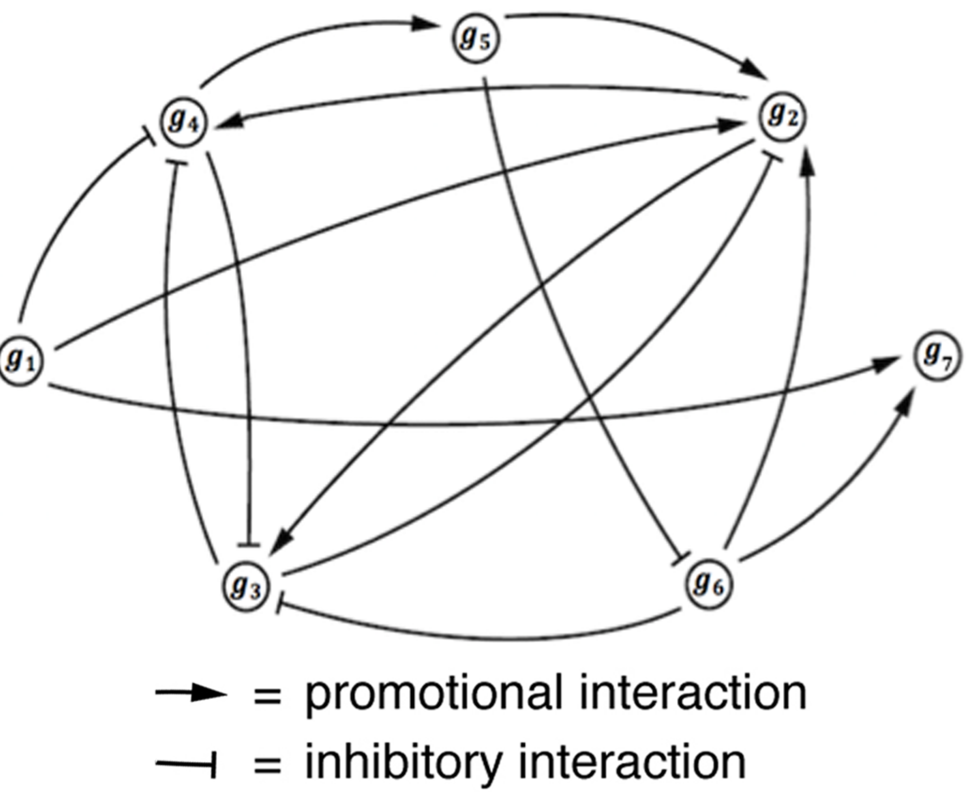
Is natural selection behind the mLmH property?

Our approach:

1. Model the evolution of biological networks as a computational problem.
2. This problem is fundamentally hard (NP-hard).
3. Show empirical evidence that mLmH produces easy instances of this problem

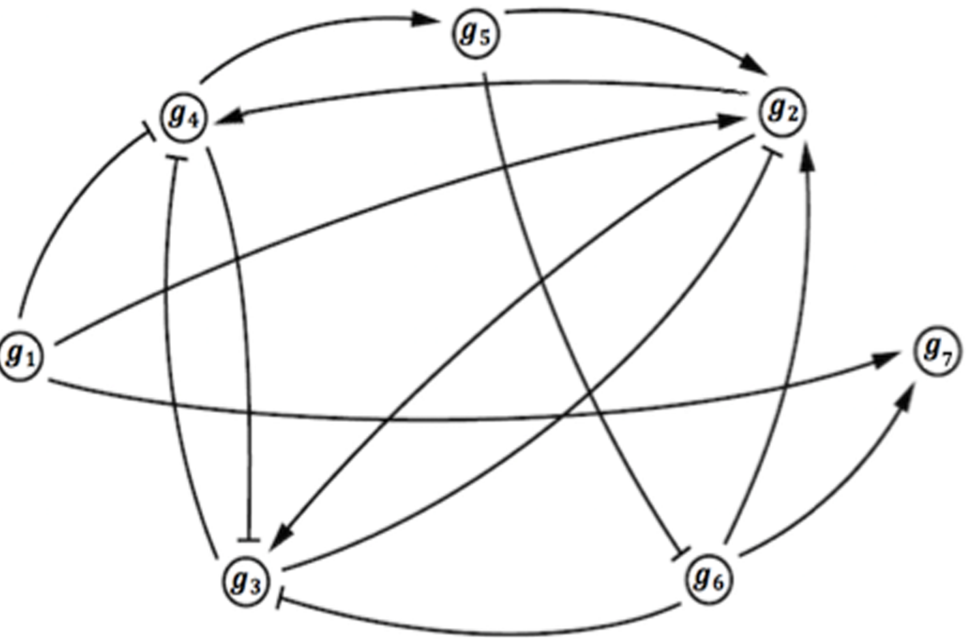
The Network Evolution Problem (NEP)

Consider this hypothetical biological network:



Nodes = biomolecules
Edges = interactions

The Network Evolution Problem (NEP)



→ = promotional interaction

⊣ = inhibitory interaction

=

Adjacency matrix

	g_1	g_2	g_3	g_4	g_5	g_6	g_7
g_1	0	+1	0	−1	0	0	+1
g_2	0	0	+1	+1	0	0	0
g_3	0	−1	0	−1	0	0	0
g_4	0	+1	−1	0	+1	0	0
g_5	0	+1	0	0	0	−1	0
g_6	0	+1	−1	0	0	0	+1
g_7	0	0	0	0	0	0	0

+1 = promotional interaction

−1 = inhibitory interaction

0 = no interaction

The Network Evolution Problem (NEP)

Adjacency matrix

	g_1	g_2	g_3	g_4	g_5	g_6	g_7
g_1	0	+1	0	−1	0	0	+1
g_2	0	0	+1	+1	0	0	0
g_3	0	−1	0	−1	0	0	0
g_4	0	+1	−1	0	+1	0	0
g_5	0	+1	0	0	0	−1	0
g_6	0	+1	−1	0	0	0	+1
g_7	0	0	0	0	0	0	0

beneficial interaction

detrimental interaction

The Network Evolution Problem (NEP)

Scores

	Benefit	Damage
g_1	1	2
g_2	4	2
g_3	2	3
g_4	2	2
g_5	1	0
g_6	2	1
g_7	0	2

Adjacency matrix

	g_1	g_2	g_3	g_4	g_5	g_6	g_7
g_1	0	+1	0	−1	0	0	+1
g_2	0	0	+1	+1	0	0	0
g_3	0	−1	0	−1	0	0	0
g_4	0	+1	−1	0	+1	0	0
g_5	0	+1	0	0	0	−1	0
g_6	0	+1	−1	0	0	0	+1
g_7	0	0	0	0	0	0	0

beneficial interaction

detrimental interaction

The Network Evolution Problem (NEP)

Scores

	Benefit	Damage
g_1	1	2
g_2	4	2
g_3	2	3
g_4	2	2
g_5	1	0
g_6	2	1
g_7	0	2

NEP

=

Which genes should ideally be conserved/ deleted so as to maximize the total benefits while the total damages <= some threshold?

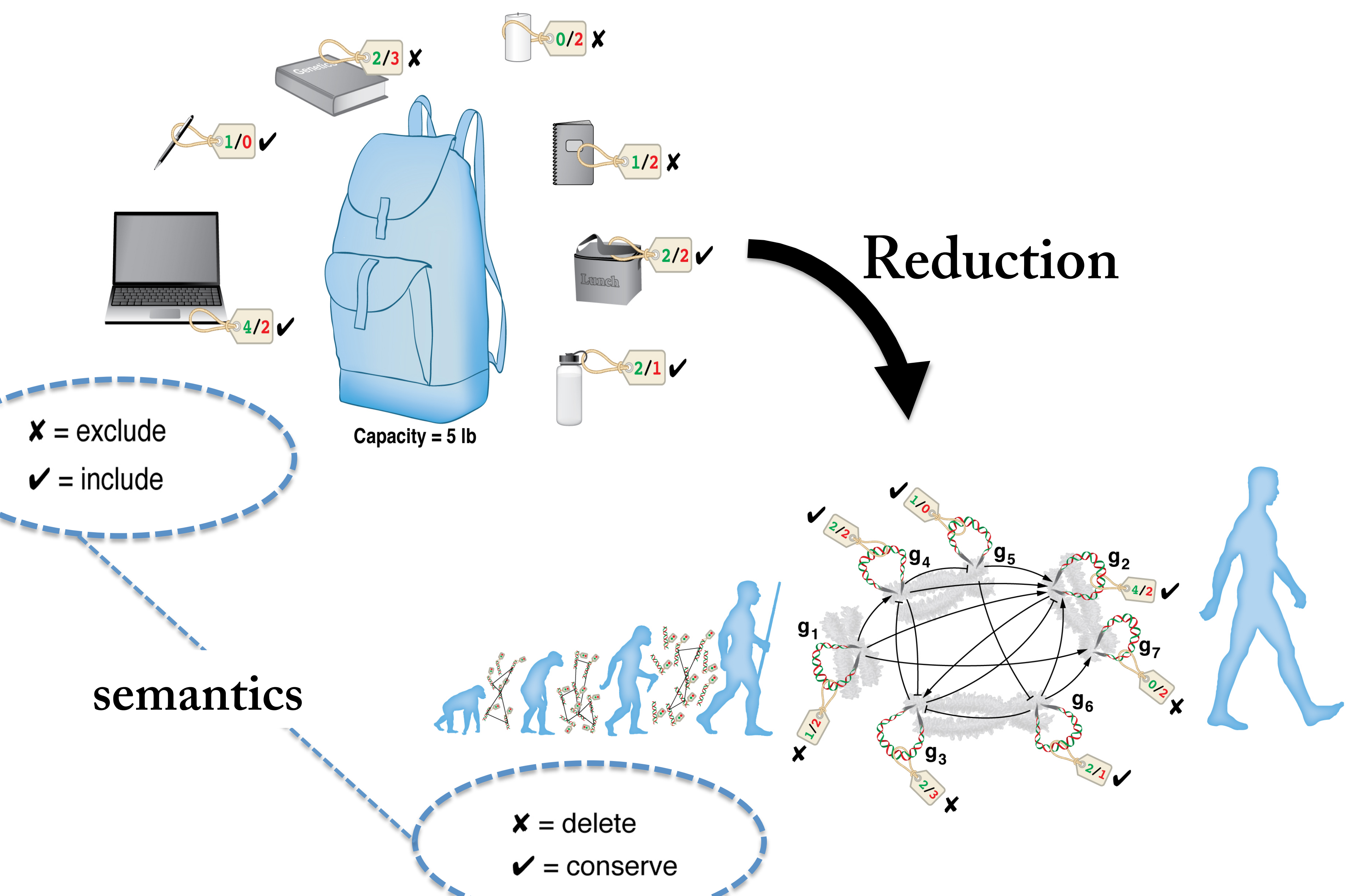
Majority-leaves Minority-hubs (**mLmH**) topology

Is natural selection behind the mLmH property?

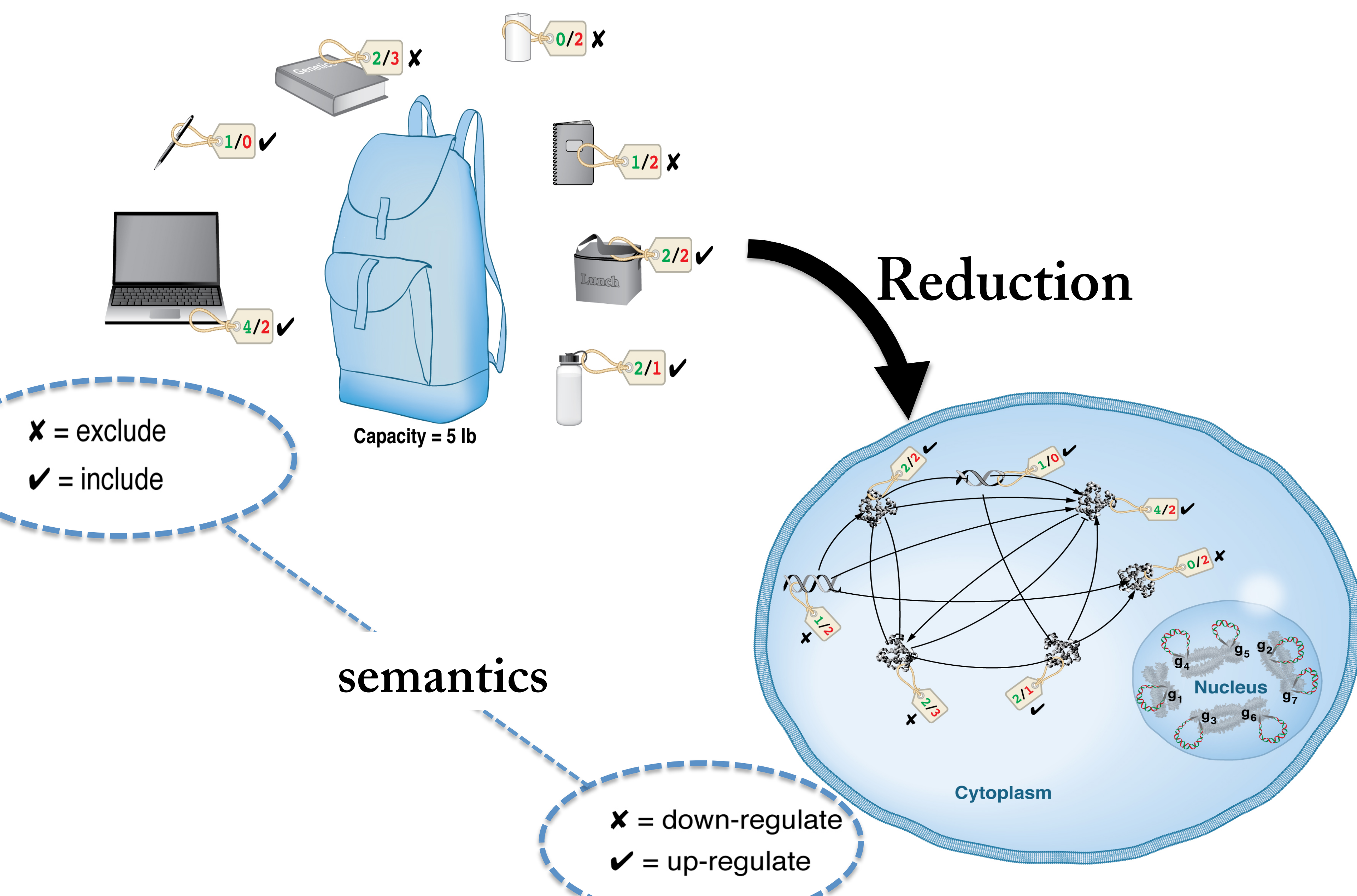
Our approach

1. Model the evolution of biological networks as a computational problem.
2. This problem is fundamentally hard (NP-hard).
3. Show empirical evidence that mLmH produces easy instances of this problem

NEP is NP-hard



NEP is NP-hard



Majority-leaves Minority-hubs (**mLmH**) topology

Is natural selection behind the mLmH property?

Our approach

1. Model the evolution of biological networks as a computational problem.

2. This problem is fundamentally hard (NP-hard).

3. Show empirical evidence that mLmH produces easy instances of this problem ➔ less iterations of random-variation non-random selection are needed to conserve(invent)/delete(mutate) the right set of genes

Basic idea

i. evolve a population of synthetic networks

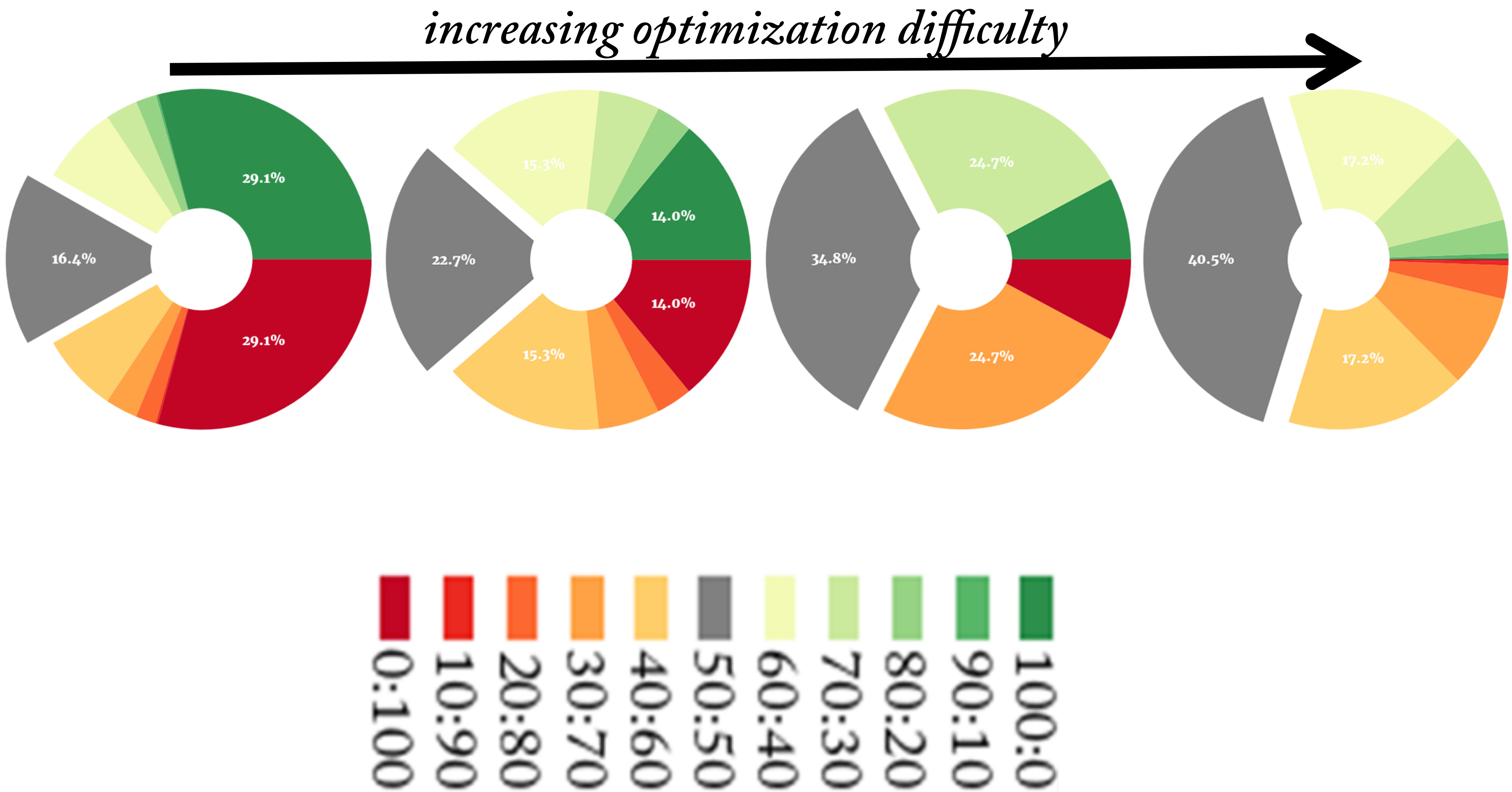
ii. select-and-breed from those that "circumvent" computational intractability more successfully

ii.

iii. How does the topology look after N generations

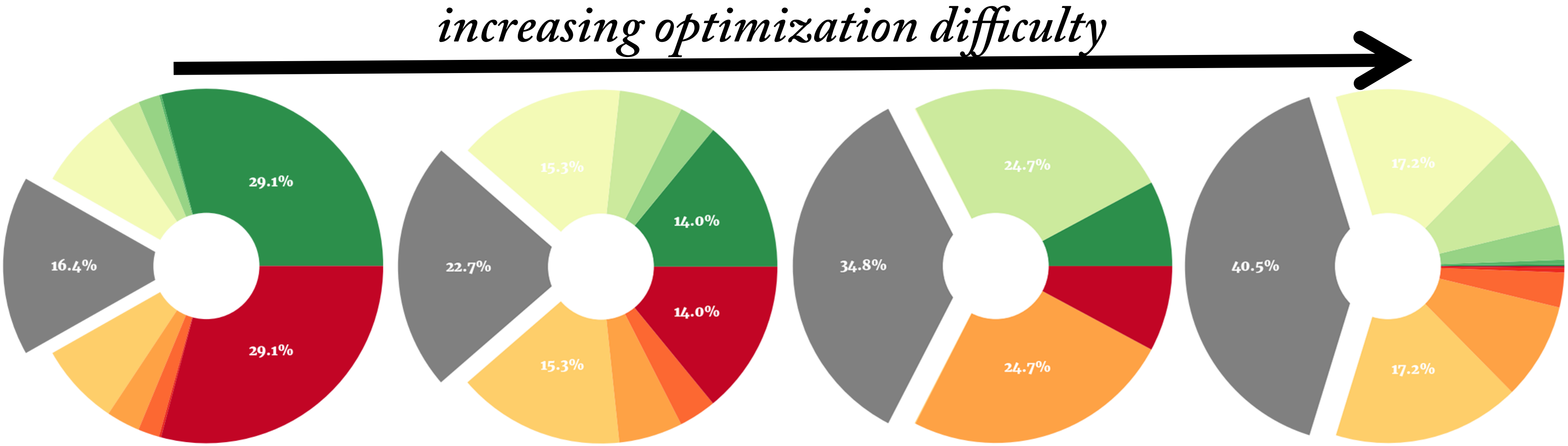
3. Show empirical evidence that mLmH produces easy instances of this problem ➔ less iterations of random-variation non-random selection are needed to conserve(invent)/delete(mutate) the right set of genes

fitness criteria (1)



benefit:damage ratio (%)

fitness criteria (1)



The smaller the number of "ambiguous" nodes, the less optimization needed

fitness criteria (2)

Effective Total Benefits :

Accumulating (cleansing) as many beneficial (detrimental) interactions with as few genes to conserve (delete) as possible

Algorithmic workflow: adaptation

1. Start with a population of N random networks whose edge:node ratio = that of a real biological network

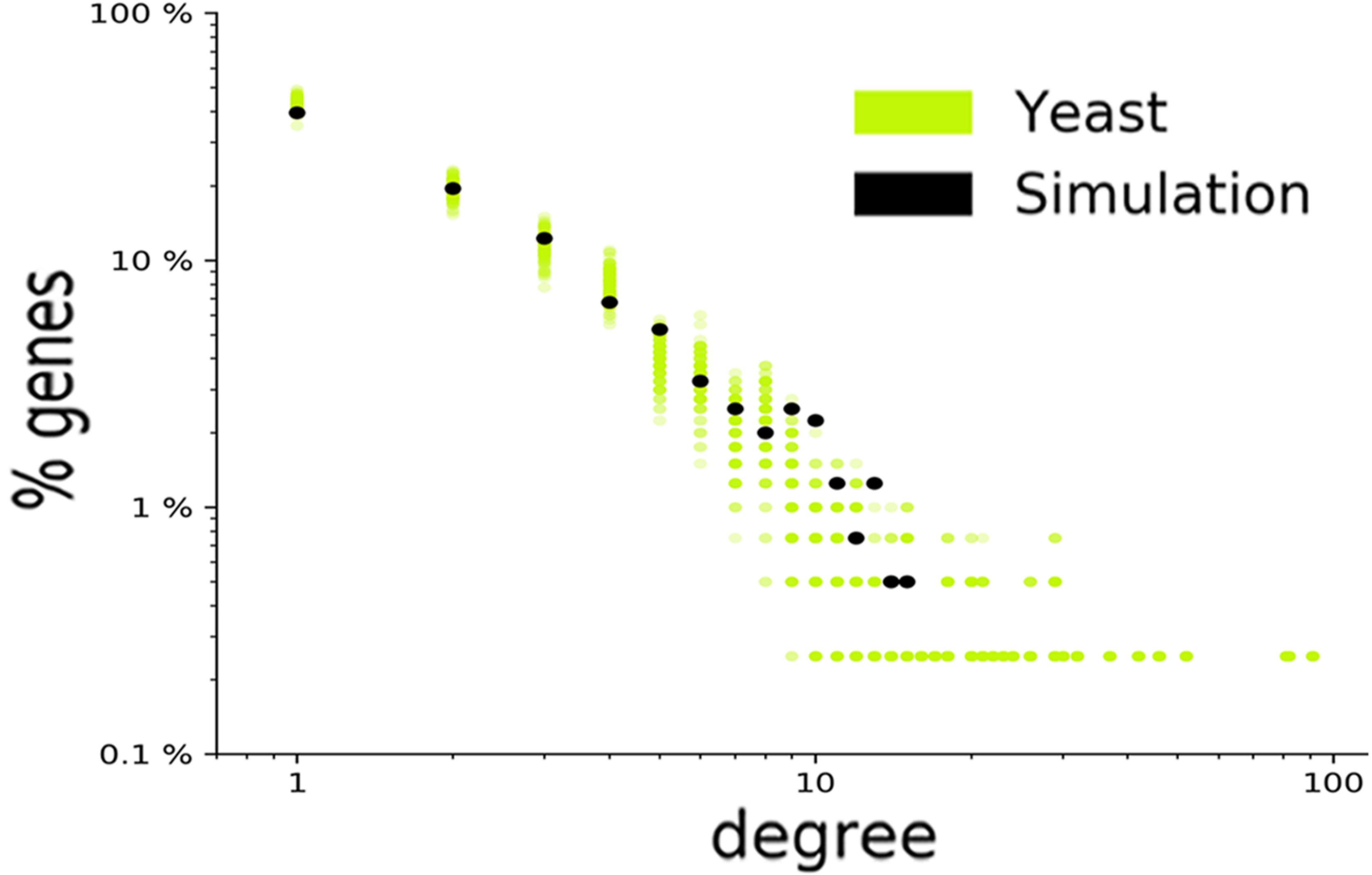
Algorithmic workflow: adaptation

1. Start with a population of N random networks whose edge:node ratio = that of a real biological network

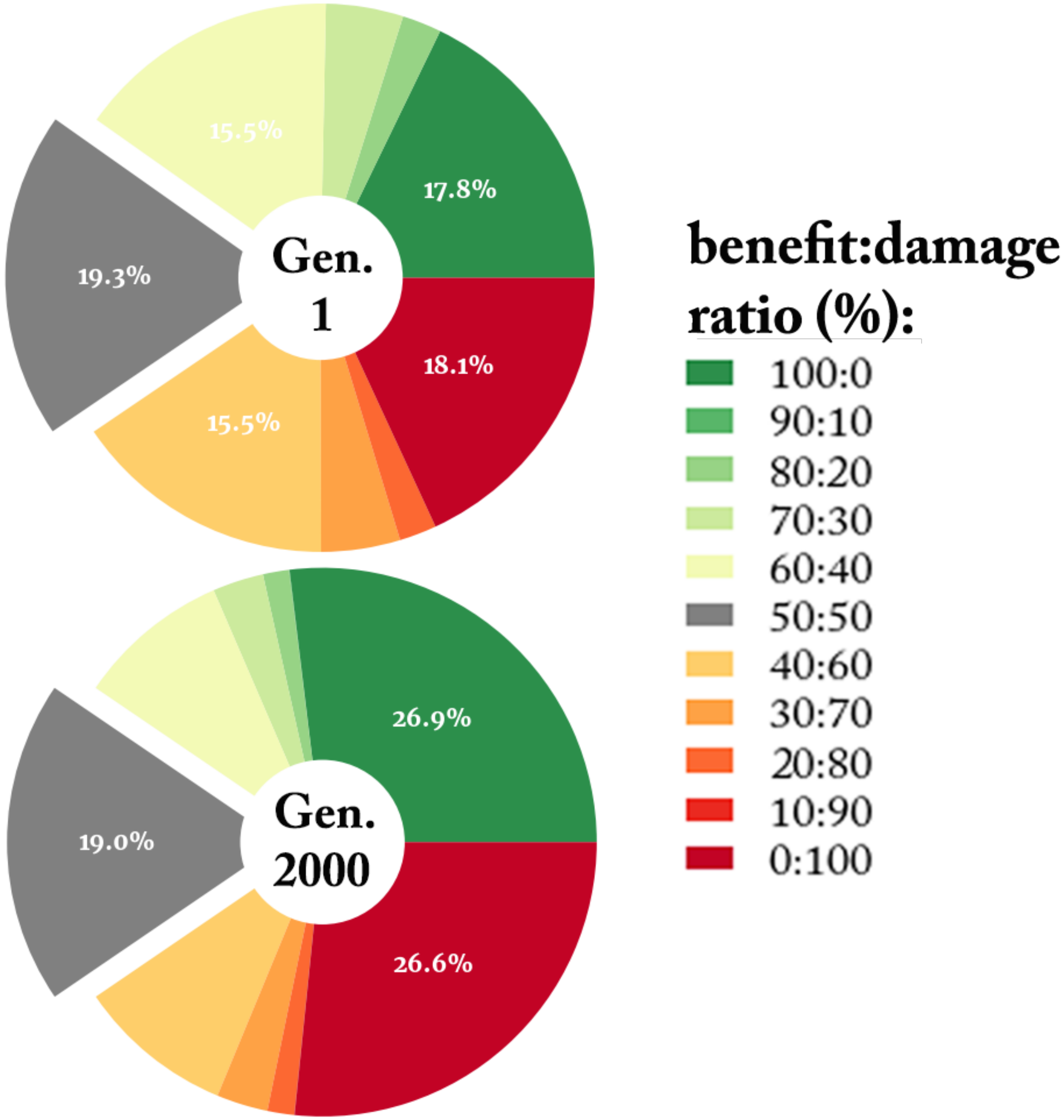
Loop: 2000 generations.

2. Mutate each one of those
3. Generate random NEP instances
4. Top 10% fittest networks per fitness criteria (1)x(2)
5. New population = 10 replicas of each mutated network

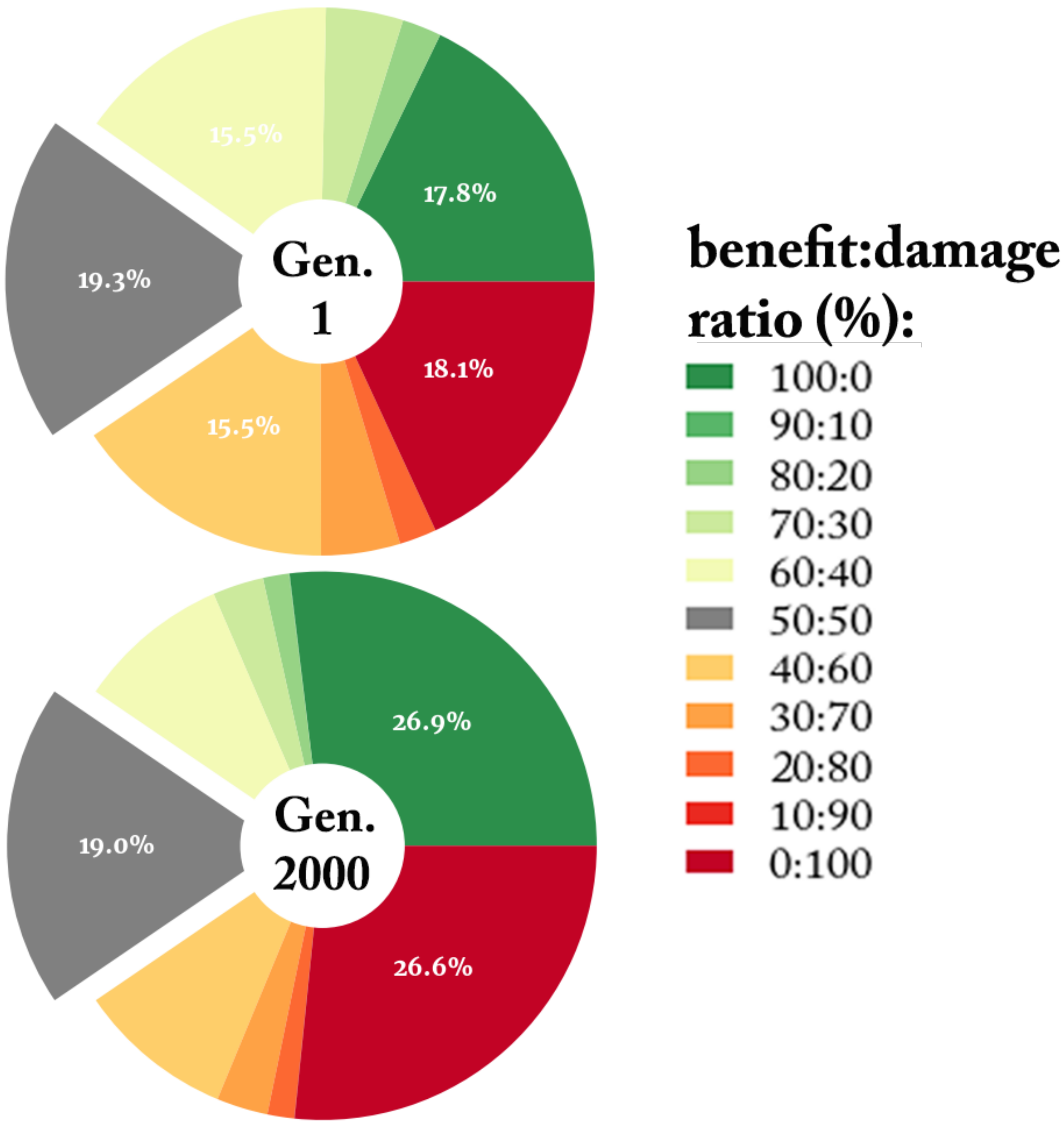
Results: adaptation

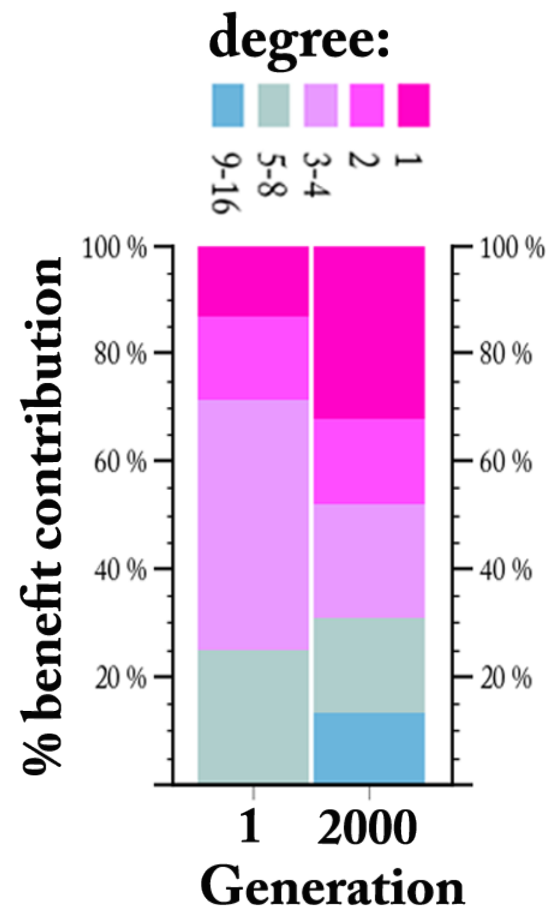


Results: adaptation

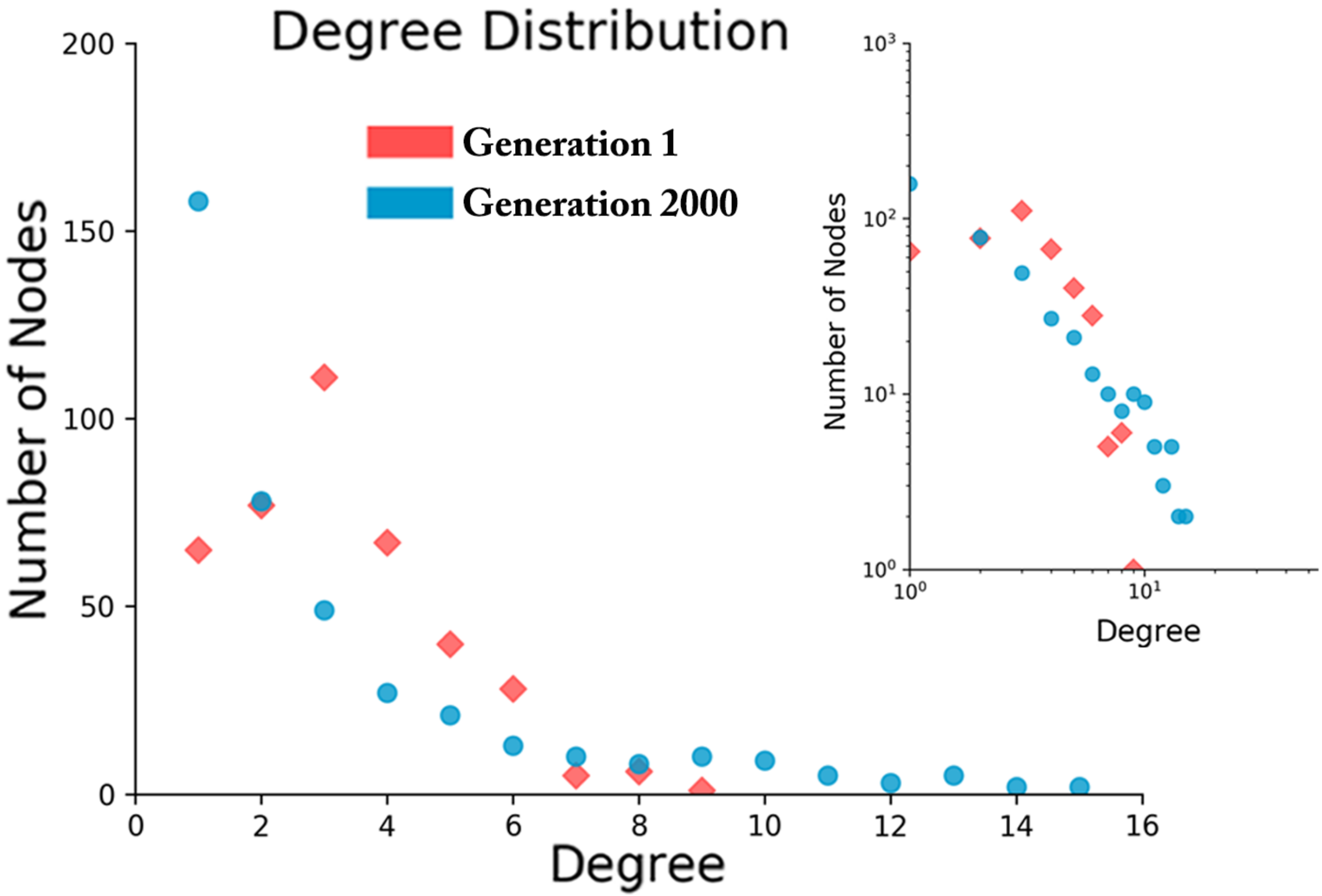


Results: adaptation





Results: adaptation



Results: **adaptation with growth**

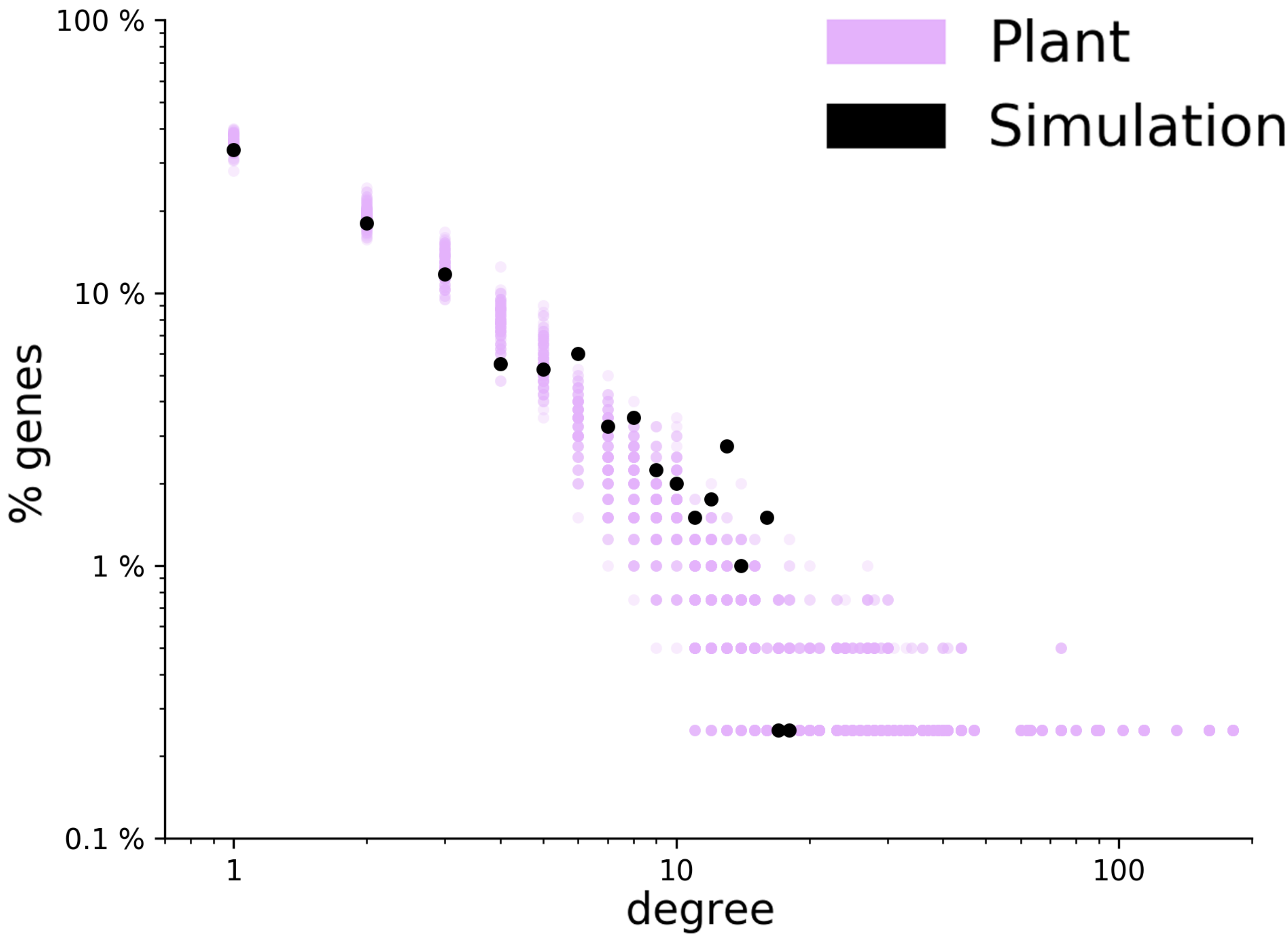
Results: **adaptation with growth**

In these simulations:

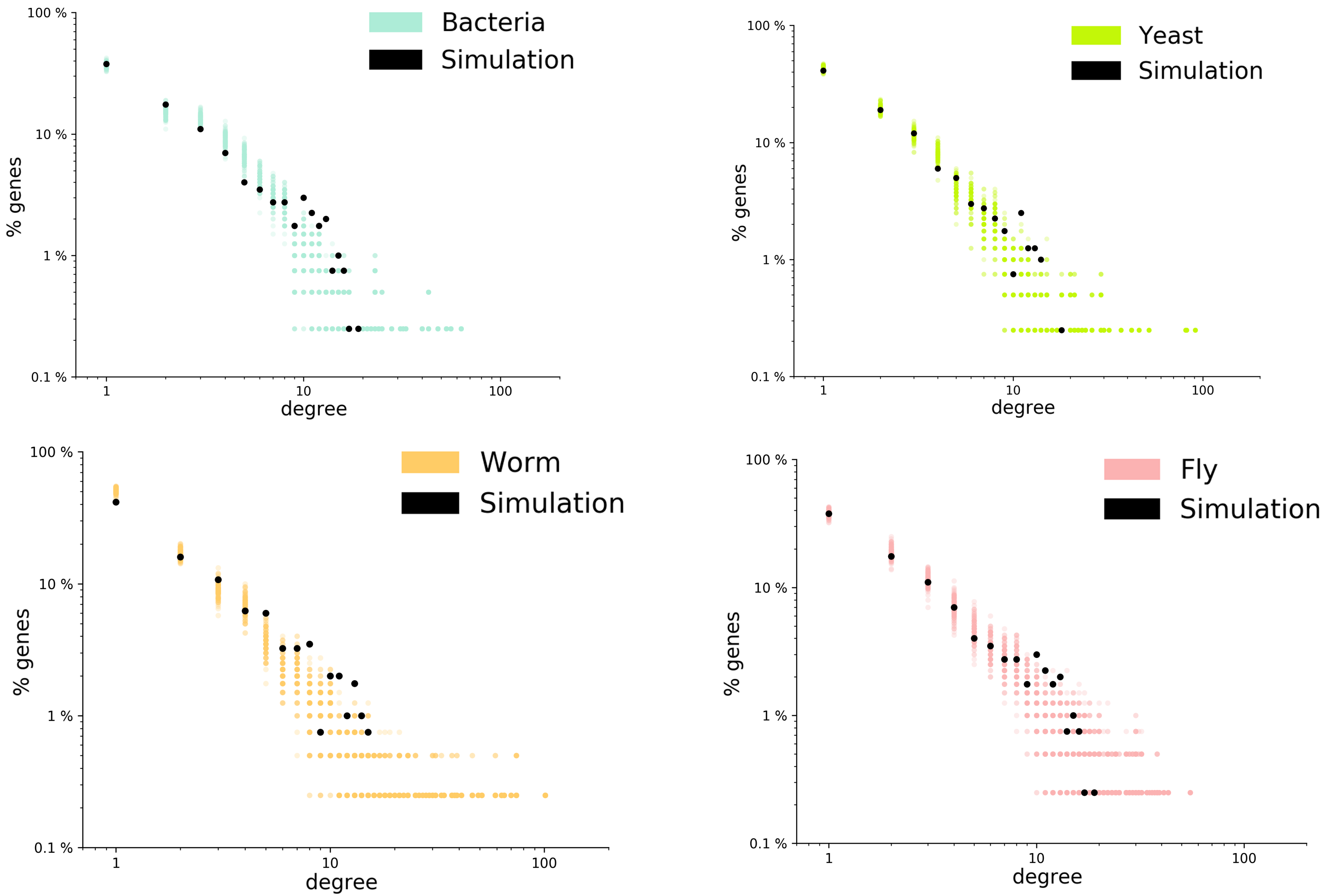
- start with <u>empty networks</u> that slowly <u>grow</u> in size over the generations.

- all else is the same as before

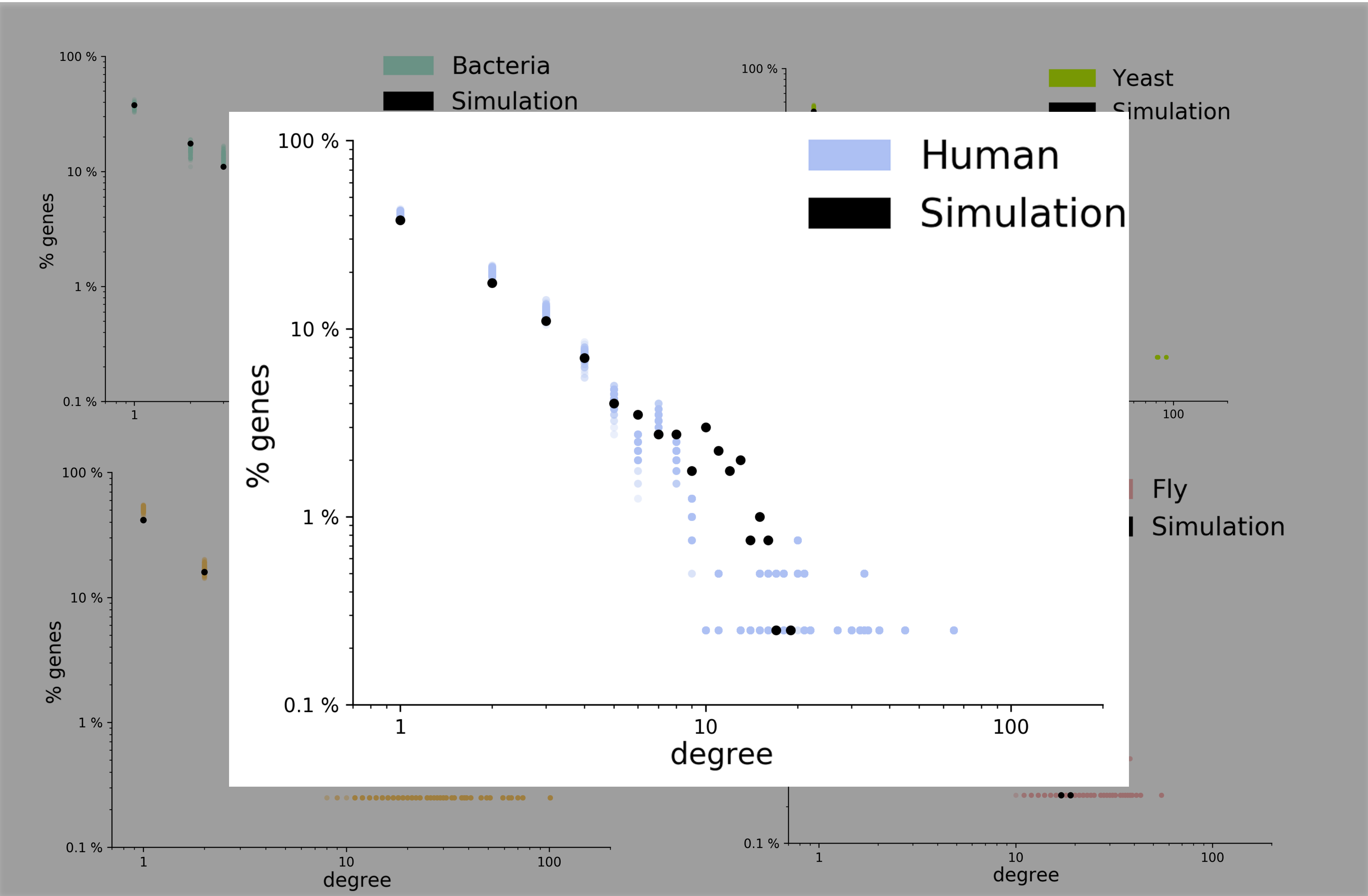
Results: adaptation with growth



Results: adaptation with growth



Results: adaptation with growth



Results: adaptation with growth

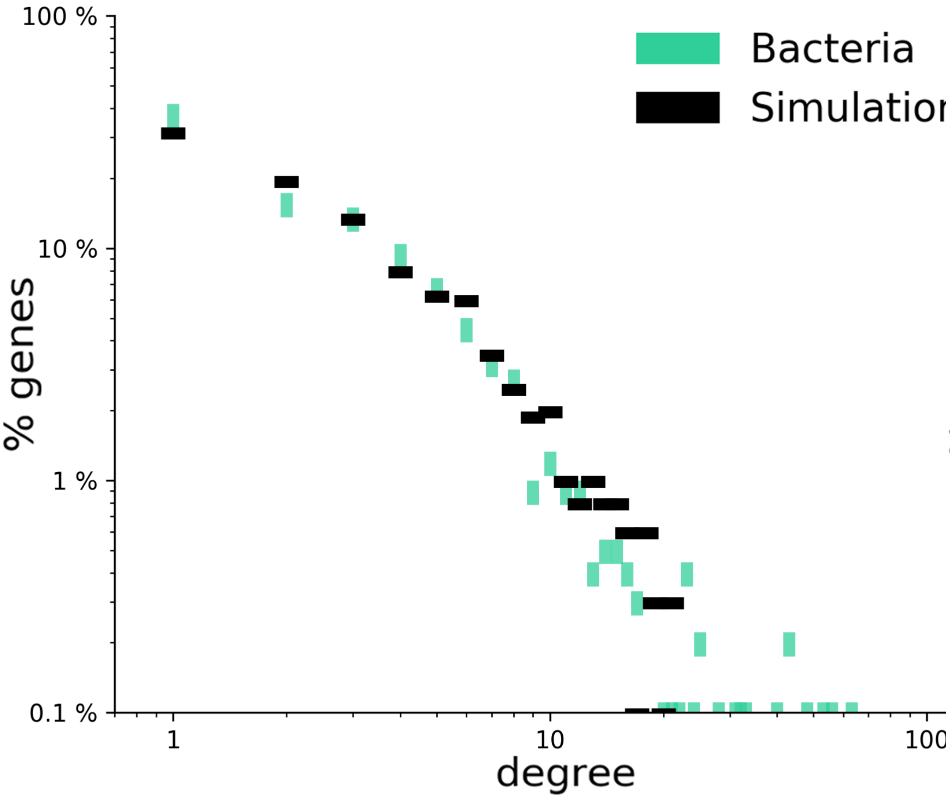
Results: **adaptation with growth**

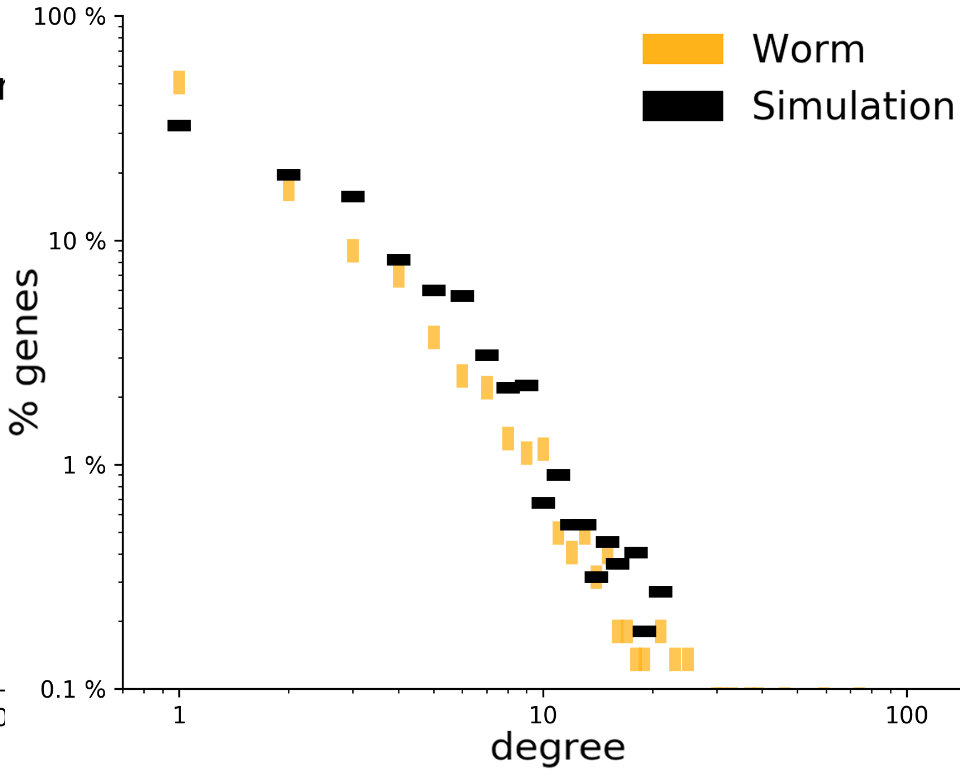
Next ...

- ➢ Evolve <u>larger</u> networks equal in size to real networks from <u>diverse</u> physiological contexts

Results: adaptation with growth

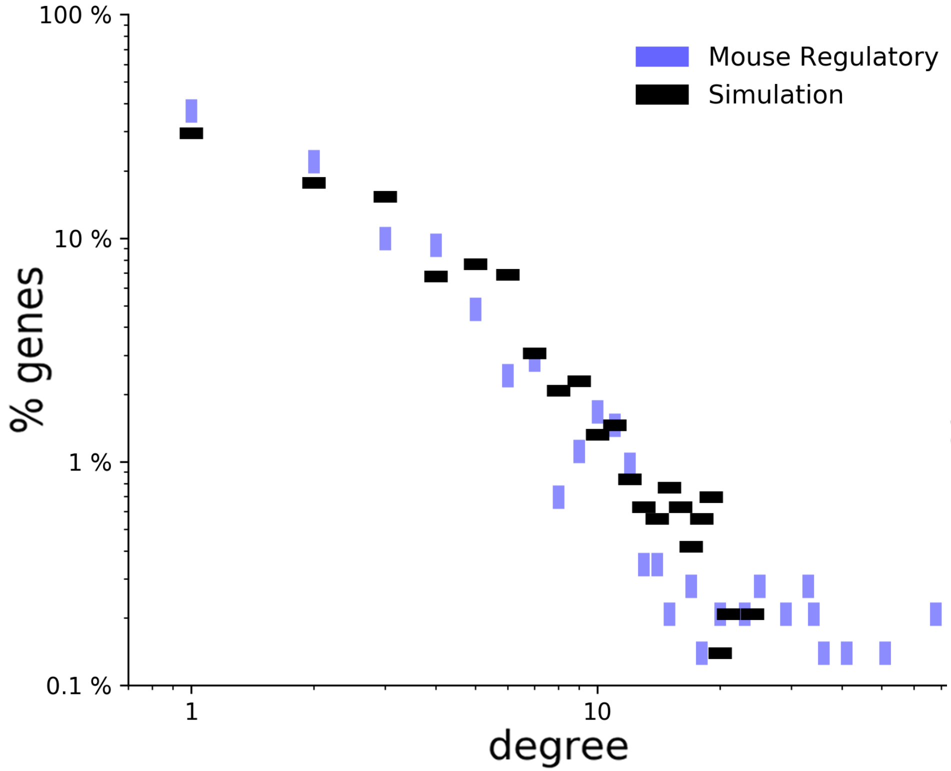
(larger and more diverse networks)

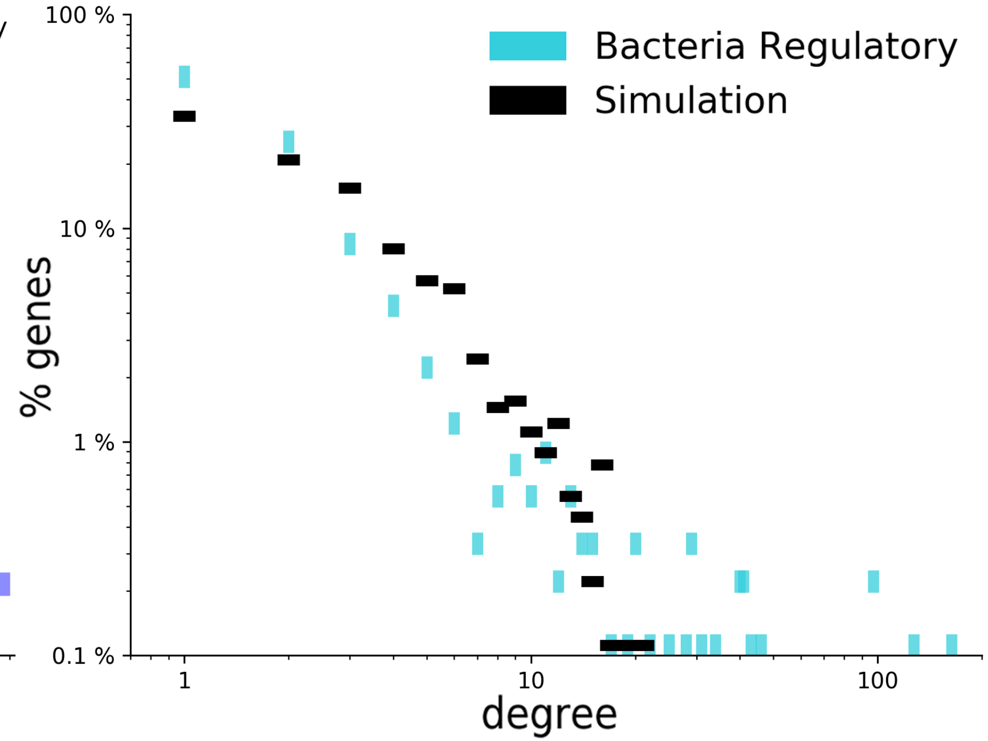




Protein-protein interaction networks

Results: adaptation with growth
(larger and more diverse networks)

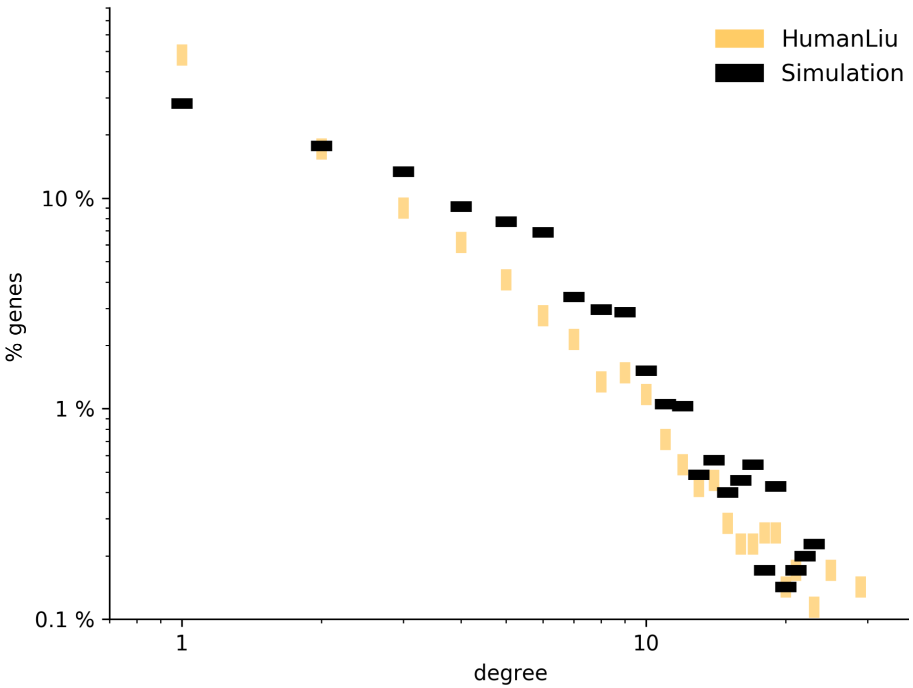


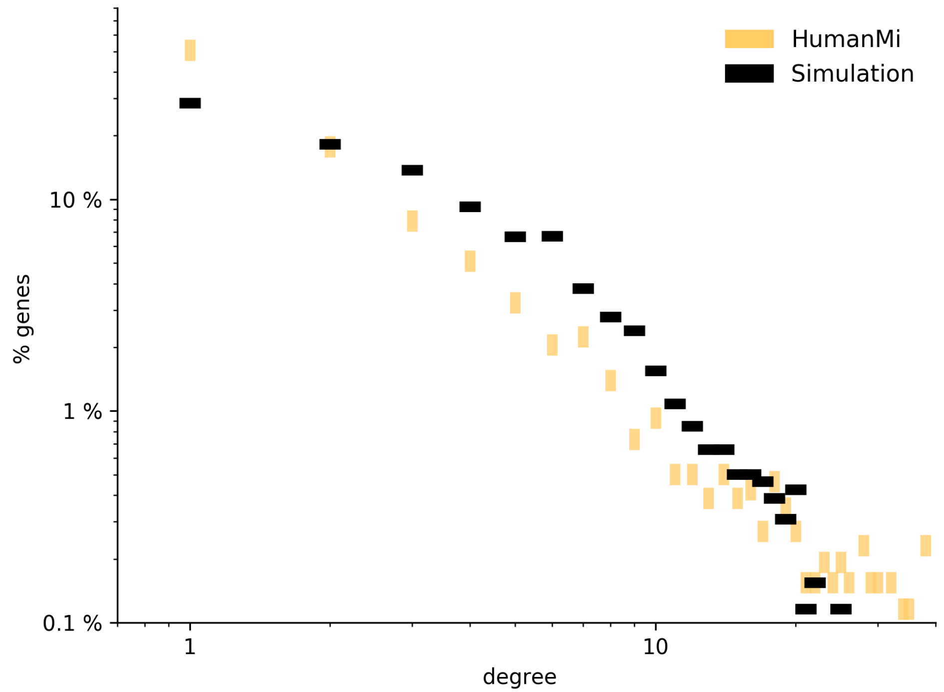


Regulatory networks

Results: adaptation with growth

(larger and more diverse networks)

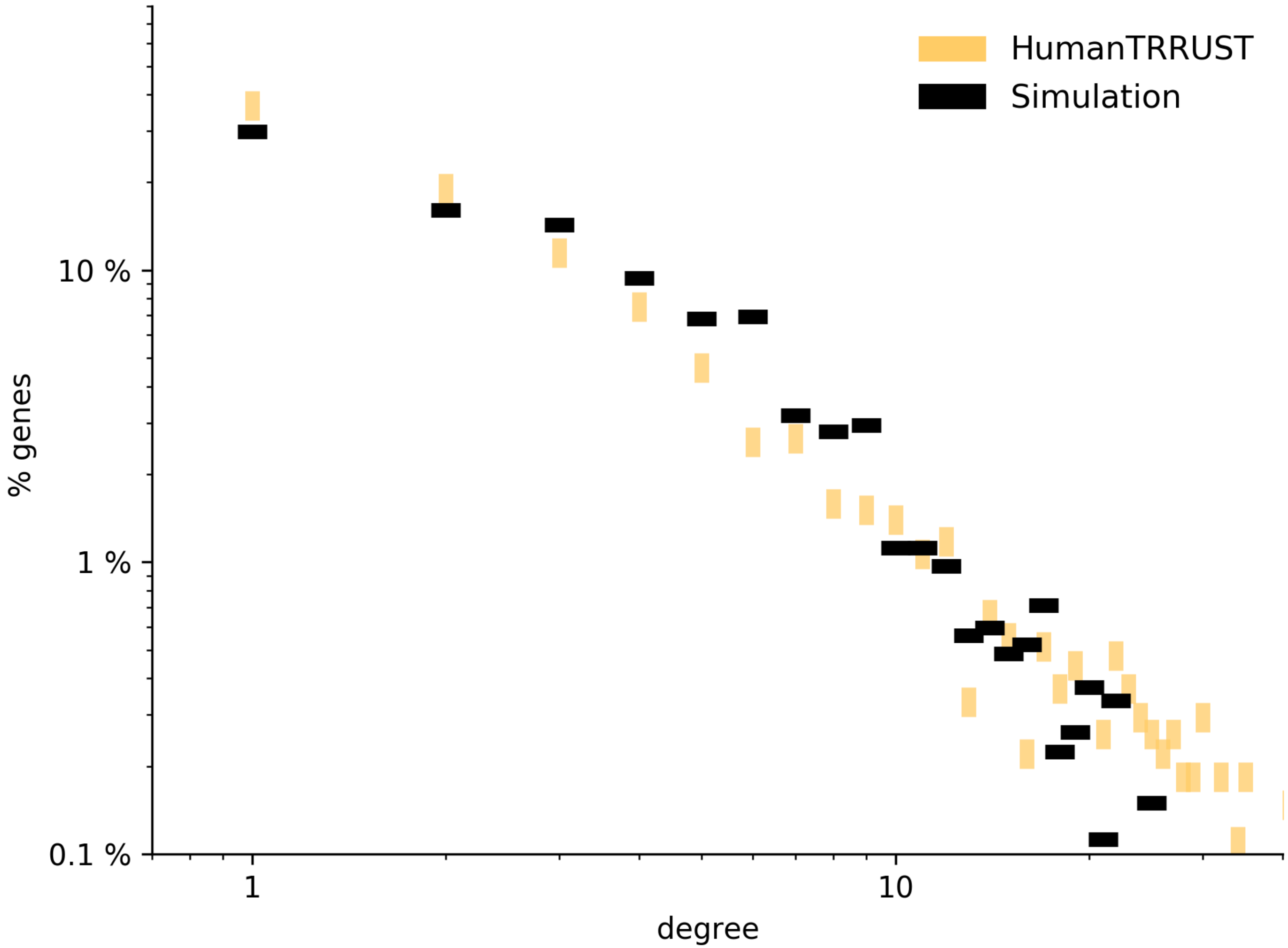




Regulatory networks

Results: adaptation with growth

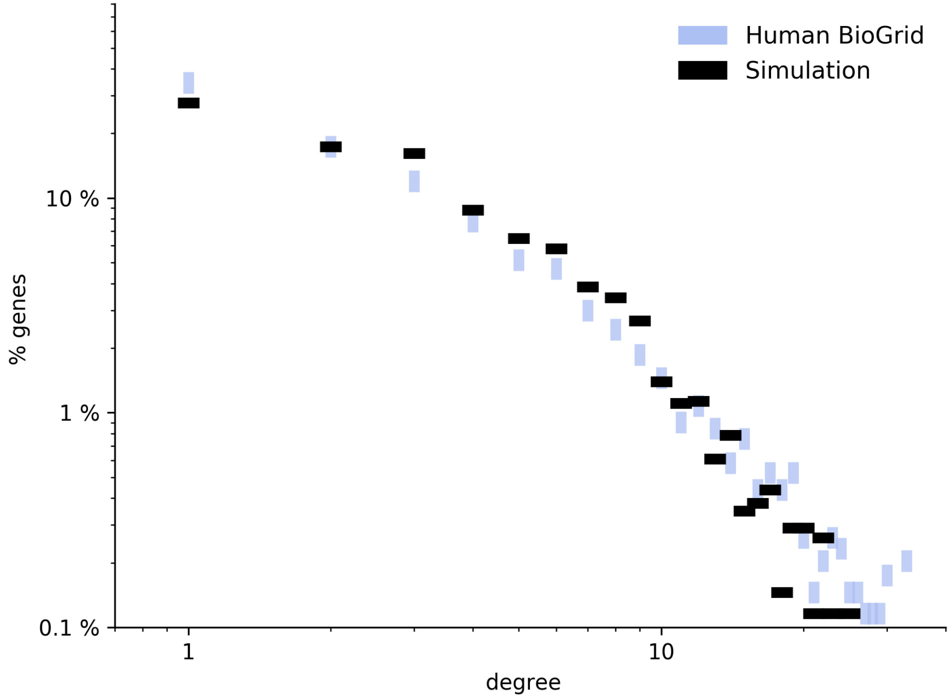
(larger and more diverse networks)

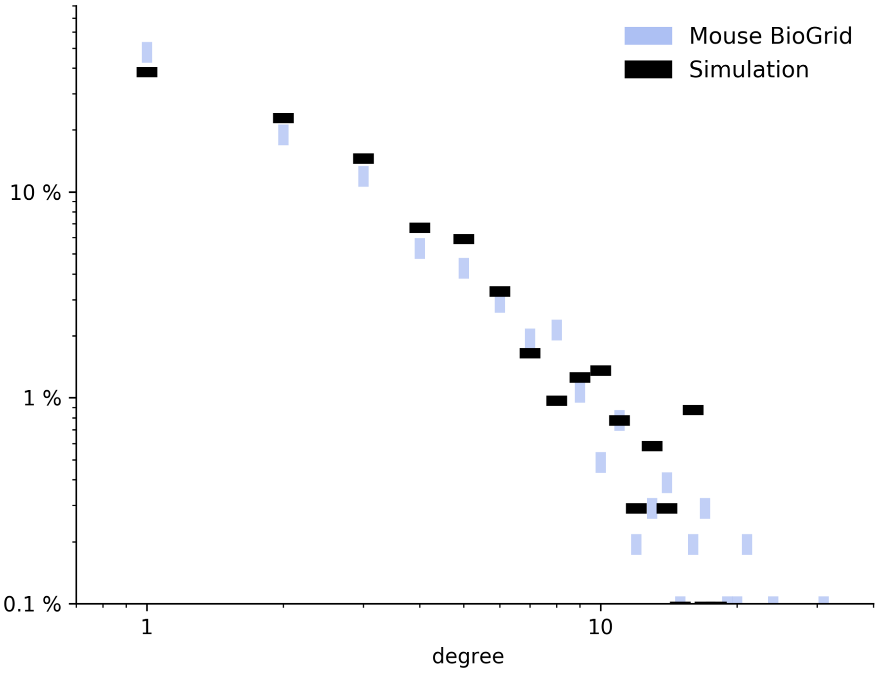


Regulatory networks

Results: **adaptation with growth**

(larger and more diverse networks)

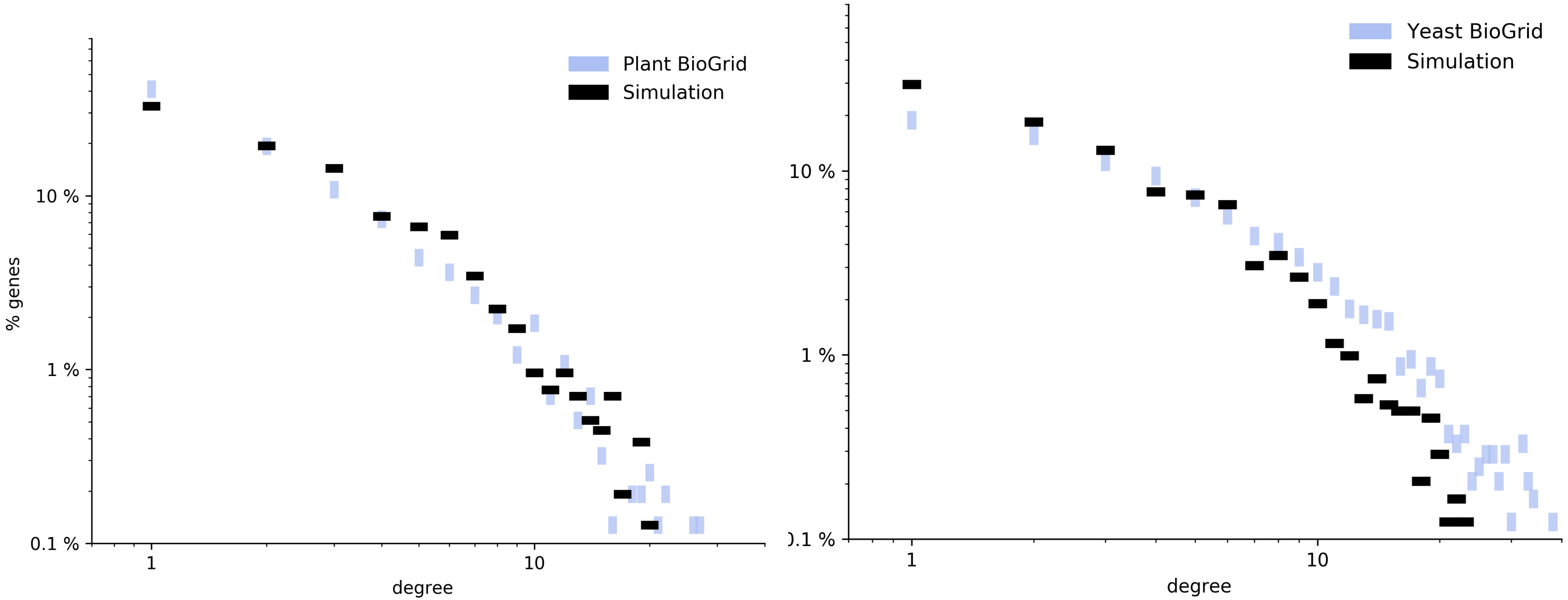




Database-sourced networks

Results: adaptation with growth

(larger and more diverse networks)



Database-sourced networks

Summary

- The topology of BNs is (at least in part) an adaptation to circumvent computational intractability.

Summary

- The topology of BNs is (at least in part) an adaptation to circumvent computational intractability.
- Sufficiency vs Necessity

Summary

- The topology of BNs is (at least in part) an adaptation to circumvent computational intractability.
- Sufficiency vs Necessity
- mLmH: concentrate **essential** functions in **hub** genes, and respond to evolutionary pressure by experimenting, on the **cheap**, with **leaf** nodes at the periphery of the network

Ongoing Work

- Use the model to explain the <u>modularity</u> of biological networks

- Intractability-based deterministic algorithm for <u>generating</u> mLmH-possessing synthetic networks

Near Future Work

- Is computational intractability behind cancer intractability?

Collaborators:

Dr. Jérôme Waldispühl

Corbin Hopper

School of Computer Science

McGill University, Montreal, Canada

http://csb.cs.mcgill.ca/

NSERC CRSNG

وزارة التعليم

Ministry of Education

Thanks

Results: adaptation