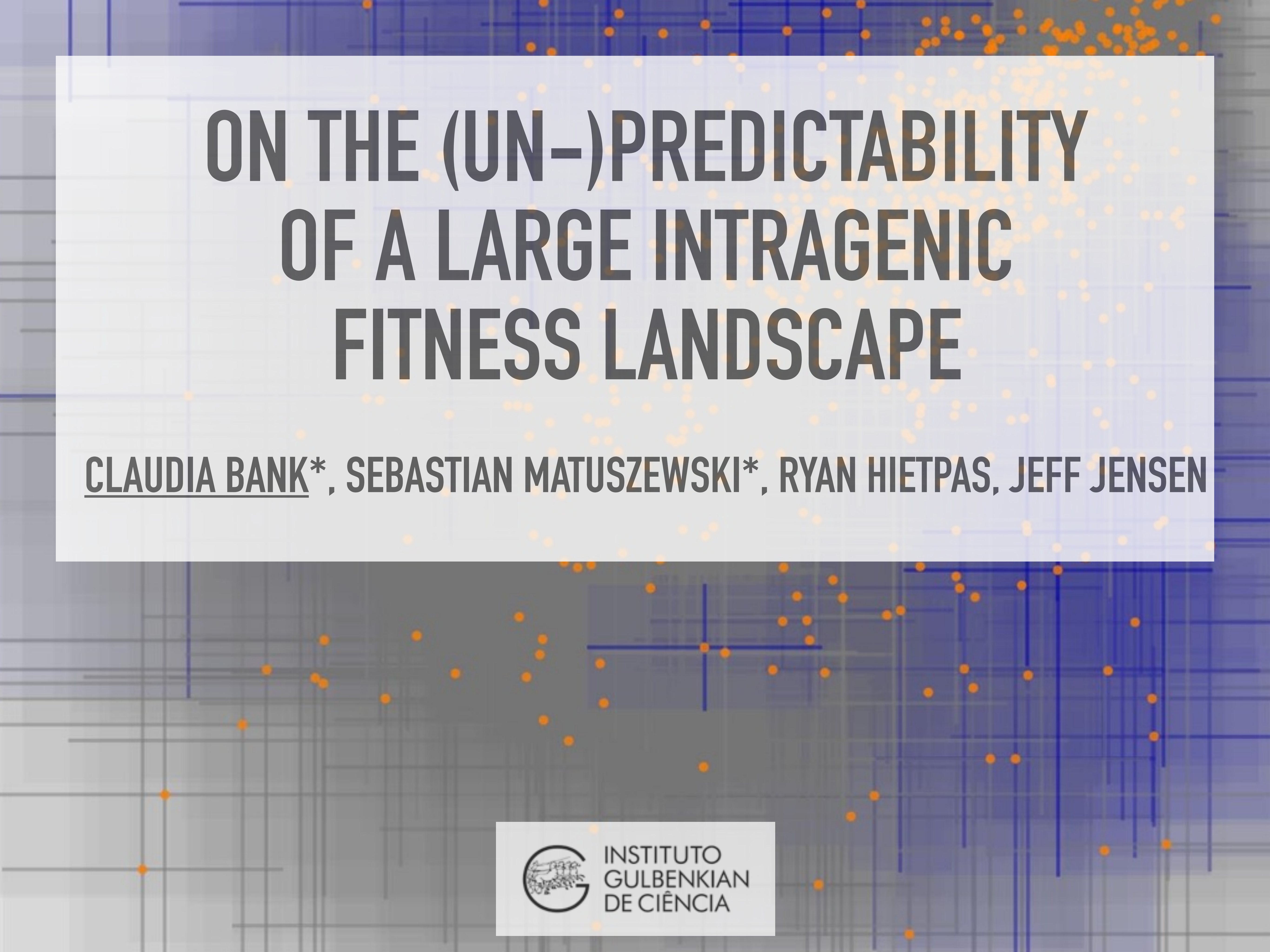


ON THE (UN-)PREDICTABILITY OF A LARGE INTRAGENIC FITNESS LANDSCAPE

CLAUDIA BANK*, SEBASTIAN MATUSZEWSKI*, RYAN HIETPAS, JEFF JENSEN

INSTITUTO
GULBENKIAN
DE CIÊNCIA

HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?

HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?

- How big/small are adaptive steps?

HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?

- How big/small are adaptive steps?
- What is the role of selection vs. genetic drift?

HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?

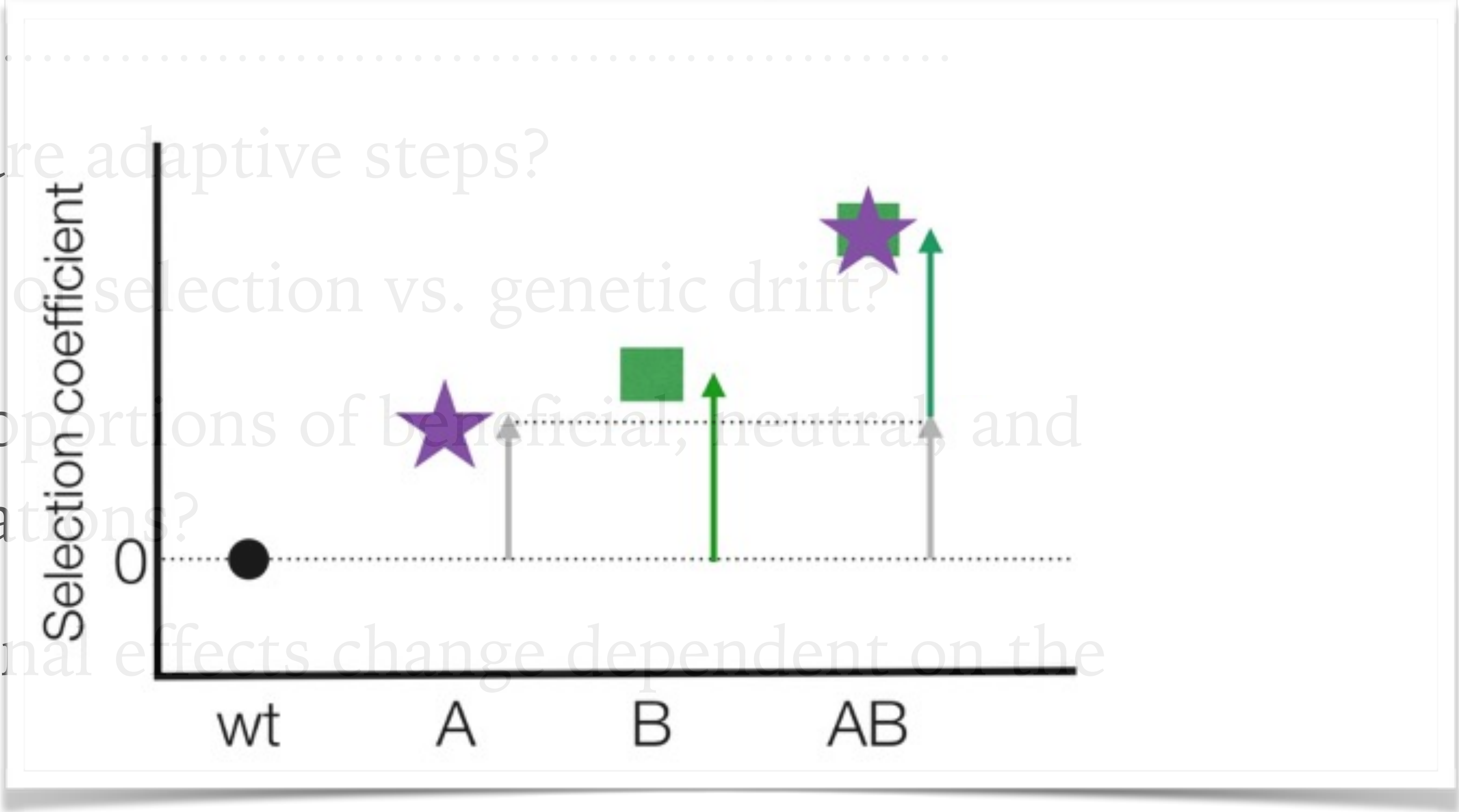
- How big/small are adaptive steps?
- What is the role of selection vs. genetic drift?
- What are the proportions of beneficial, neutral, and deleterious mutations?

HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?

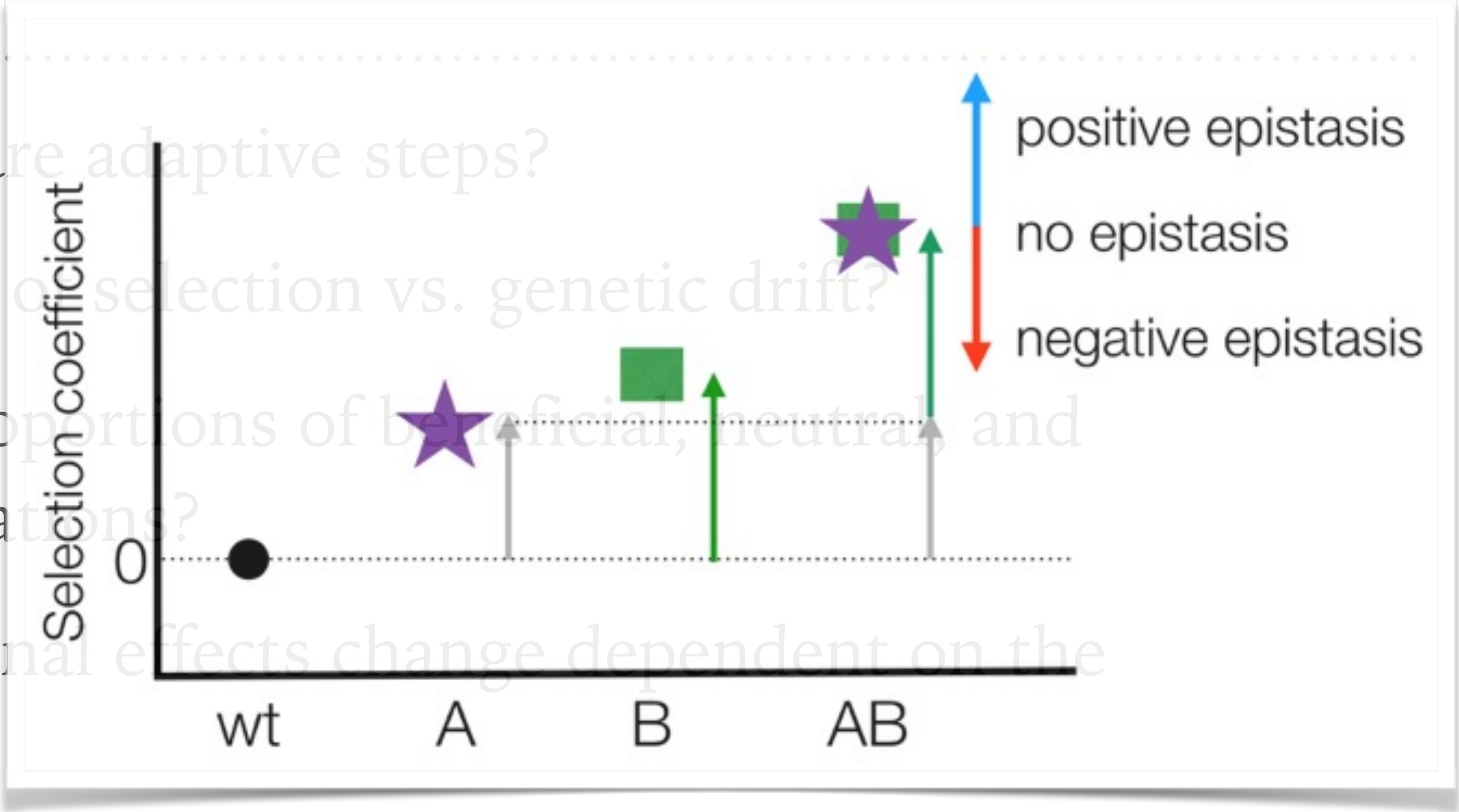
- How big/small are adaptive steps?
- What is the role of selection vs. genetic drift?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- How do mutational effects change dependent on the environment?

HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?

- How big/small are adaptive steps?
- What is the role of selection vs. genetic drift?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- How do mutational effects change dependent on the environment?
- How do mutational effects change dependent on the genetic background? (I.e., what is the role of epistasis?)

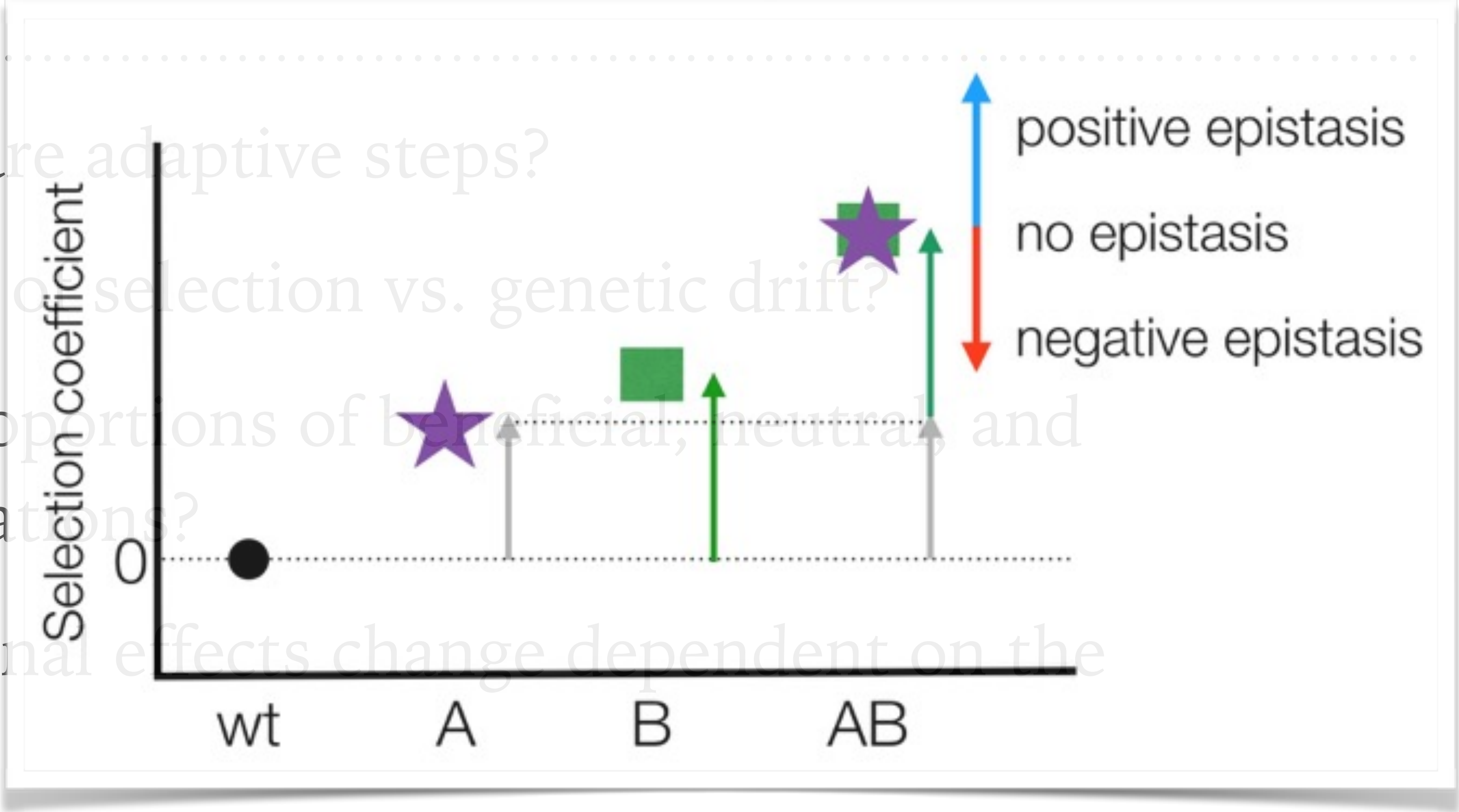


HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?



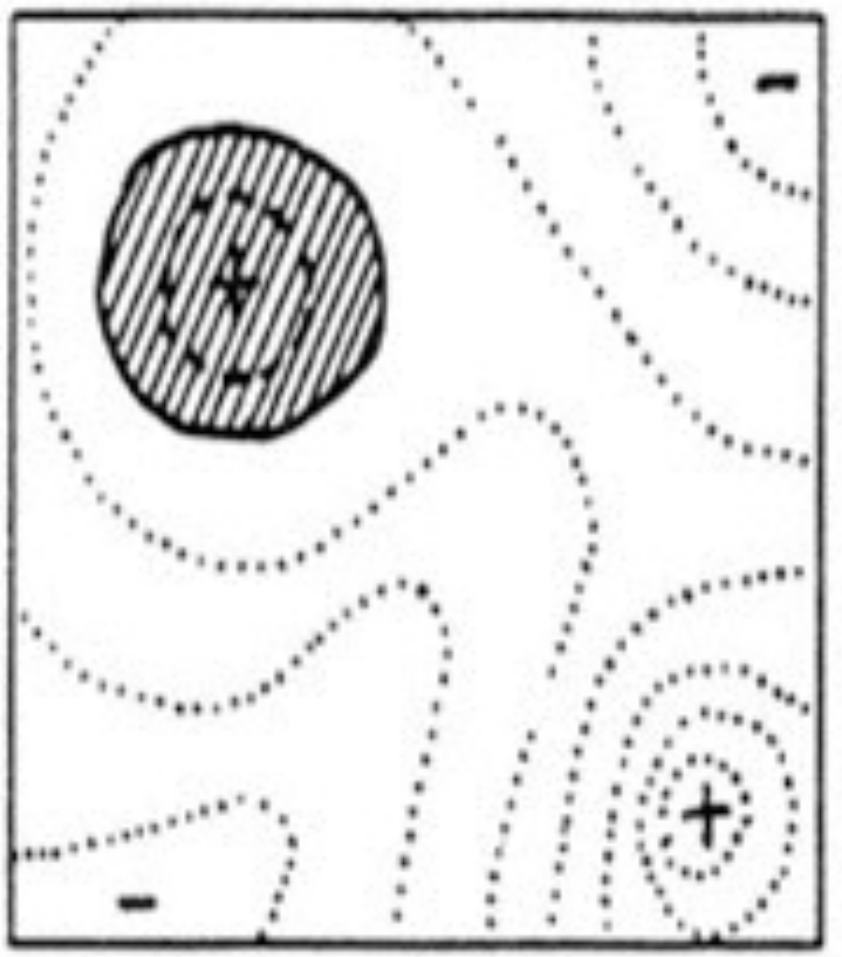
- How big/small are adaptive steps?
- What is the role of selection vs. genetic drift?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- How do mutational effects change dependent on the environment?
- How do mutational effects change dependent on the genetic background? (I.e., what is the role of epistasis?)

HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?



- How big/small are adaptive steps?
- What is the role of selection vs. genetic drift?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- How do mutational effects change dependent on the environment?
- How do mutational effects change dependent on the genetic background? (I.e., what is the role of epistasis?)
- What is the shape of the fitness landscape?

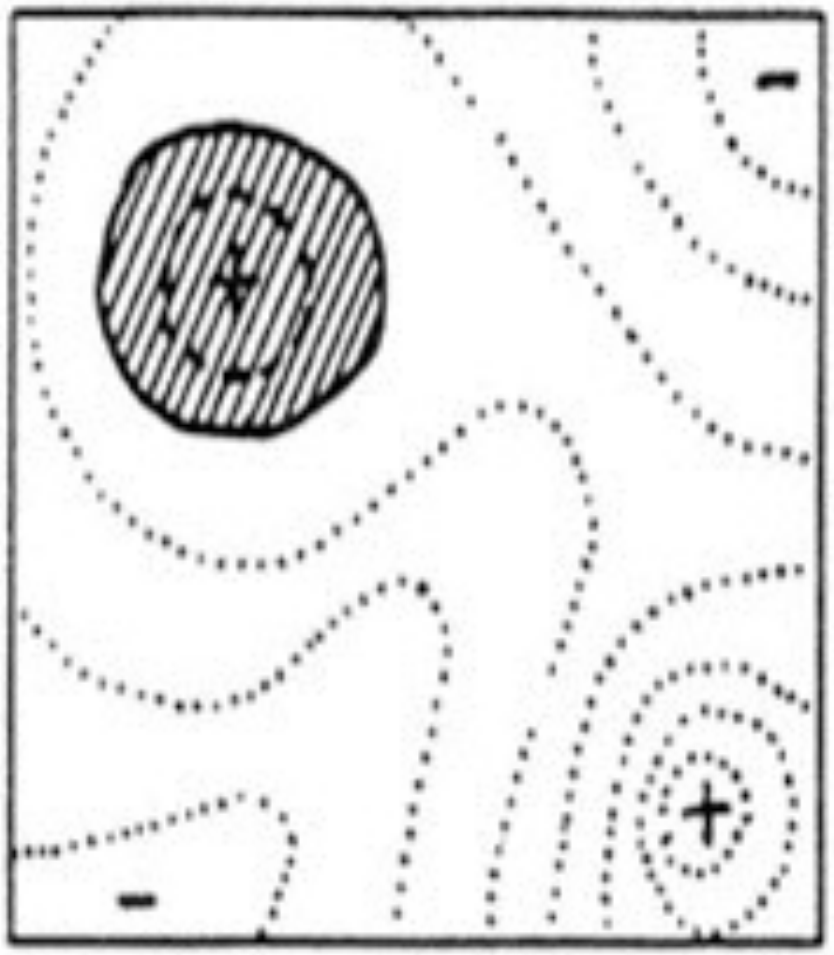
THE CONCEPT OF A FITNESS LANDSCAPE



The roles of mutation, inbreeding, crossbreeding, and selection in evolution

Wright, 1932

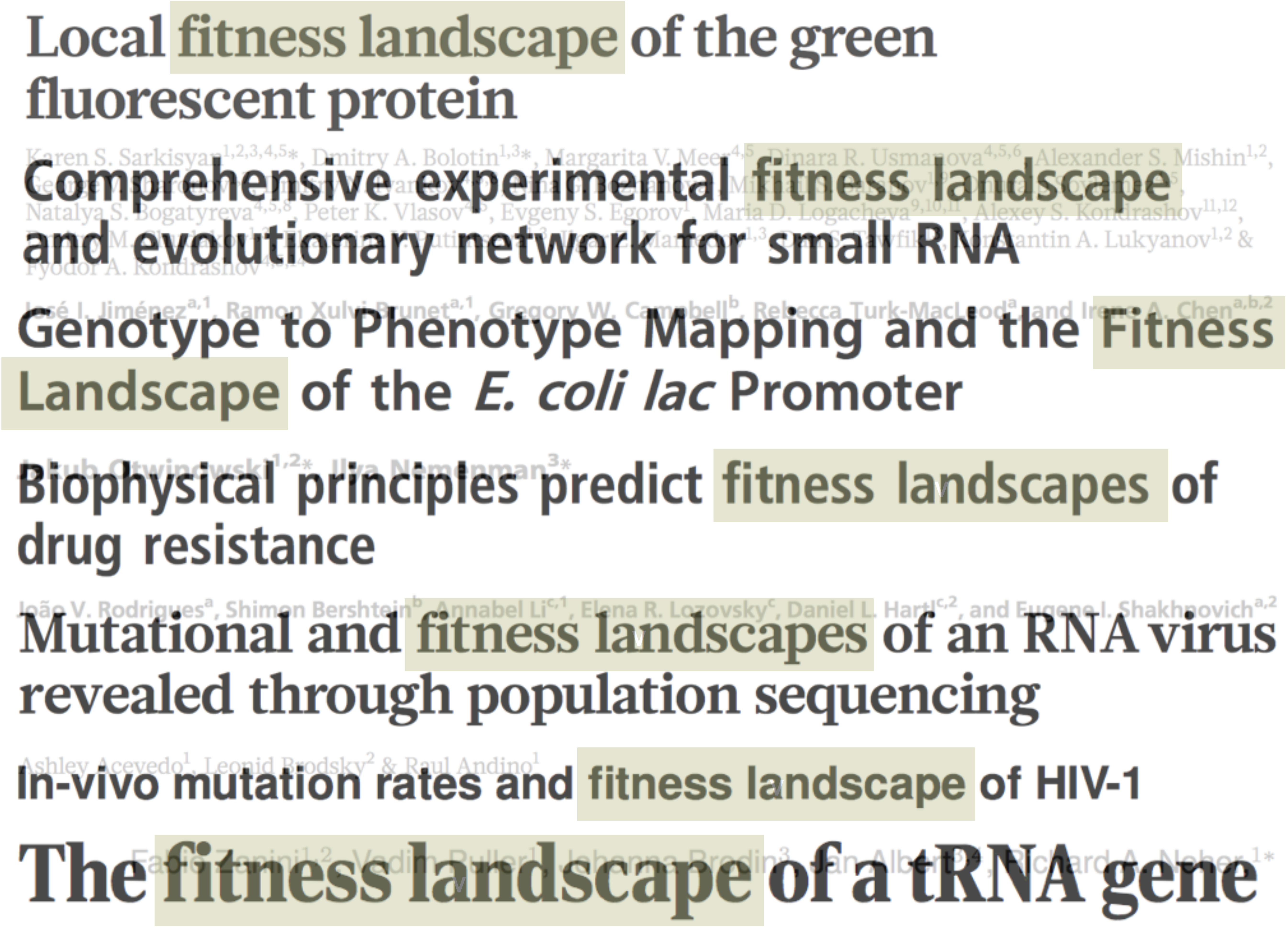
THE CONCEPT OF A FITNESS LANDSCAPE



The roles of mutation, inbreeding, crossbreeding, and selection in evolution

Wright, 1932

- fitness landscapes yield information on the predictability and repeatability of evolution
- it becomes increasingly simple to measure empirical fitness landscapes
- accumulating data on gene networks and pathways



Local fitness landscape of the green fluorescent protein

Karen S. Sarkisyan^{1,2,3,4,5*}, Dmitry A. Bolotin^{1,3*}, Margarita V. Meer^{4,5}, Dinara R. Usmanova^{4,5,6}, Alexander S. Mishin^{1,2}, [illegible]
Natalya S. Bogatyreva^{4,5,8}, Peter K. Vlasov^{4,5}, Evgeny S. Egorov^{1}, Maria D. Logacheva^{9,10,11}, Alexey S. Kondrashov^{11,12},
[illegible] Konstantin A. Lukyanov^{1,2} &
Fyodor A. Kondrashov [illegible]

Comprehensive experimental fitness landscape and evolutionary network for small RNA

José I. Jiménez^{a,1}, Ramon Xulvi-Brunet^{a,1}, Gregory W. Campbell^{b}, Rebecca Turk-MacLeod^{a}, and Irene A. Chen^{a,b,2}

Genotype to Phenotype Mapping and the Fitness Landscape of the *E. coli lac* Promoter

Jakub Otwinowski^{1,2*}, Ilya Nemenman^{3*}

Biophysical principles predict fitness landscapes of drug resistance

João V. Rodrigues^{a}, Shimon Bershtein^{b}, Annabel Li^{c,1}, Elena R. Lozovsky^{c}, Daniel L. Hartl^{c,2}, and Eugene I. Shakhnovich^{a,2}

Mutational and fitness landscapes of an RNA virus revealed through population sequencing

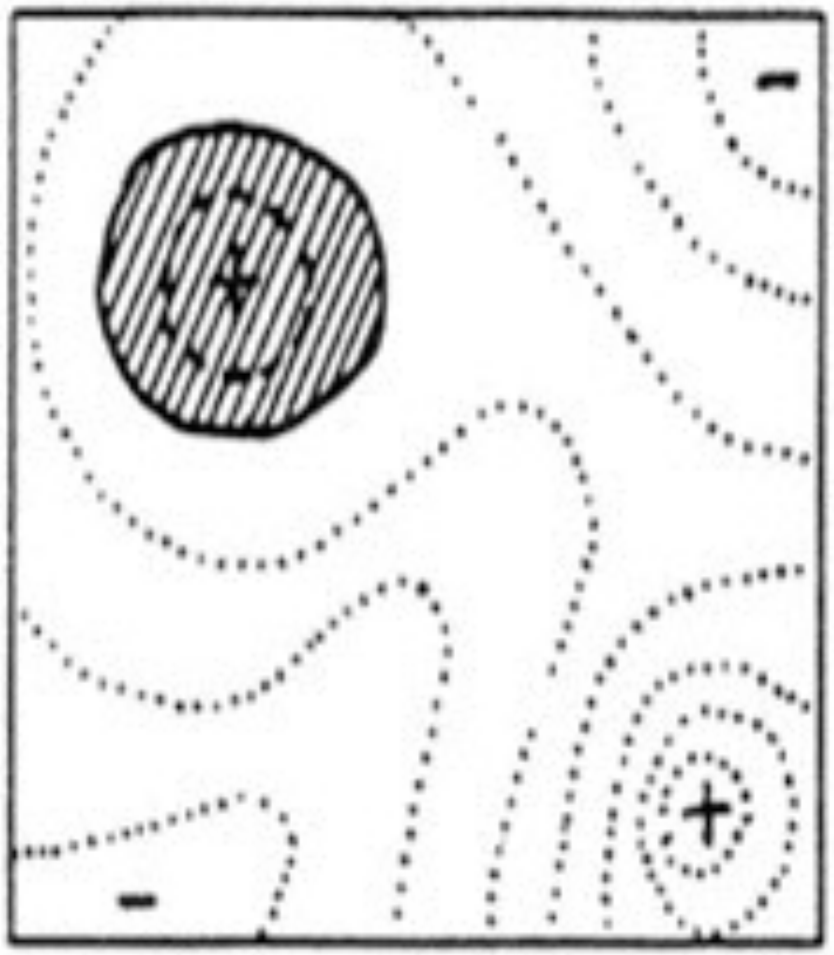
Ashley Acevedo^{1}, Leonid Brodsky^{2} & Raul Andino^{1}

In-vivo mutation rates and fitness landscape of HIV-1

Fabio Zanini^{1,2}, Vadim Puller^{1}, Johanna Brodin^{3}, Jan Albert^{3,4}, Richard A. Neher^{1*}

The fitness landscape of a tRNA gene

THE CONCEPT OF A FITNESS LANDSCAPE



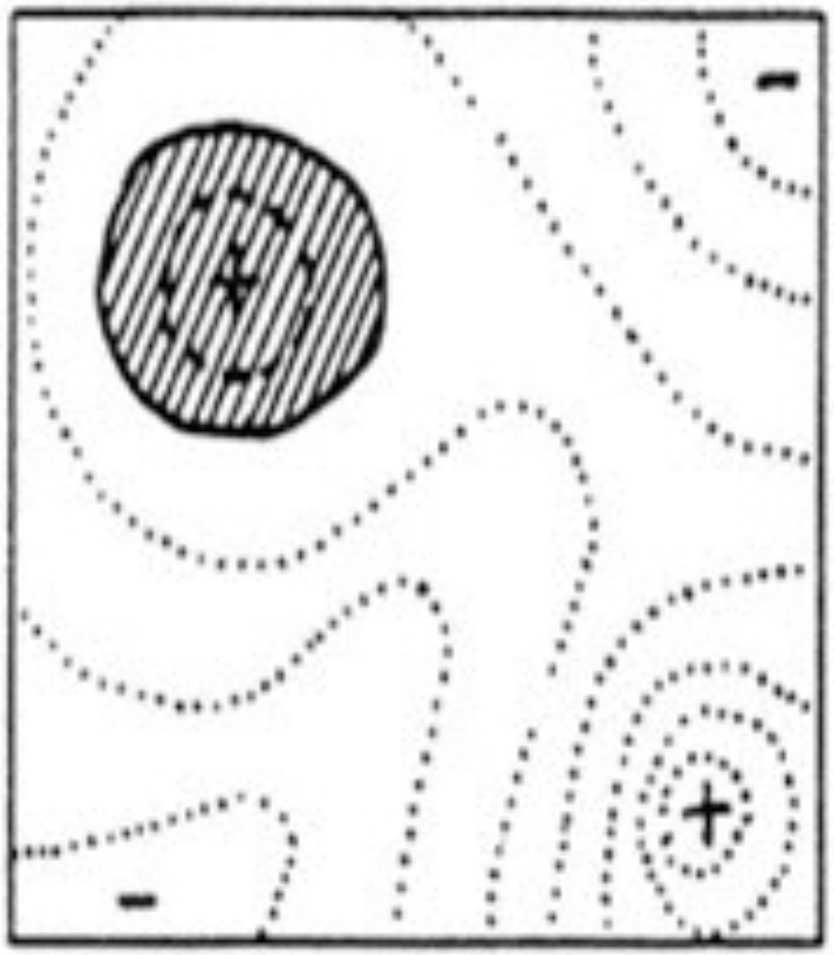
The roles of mutation, inbreeding, crossbreeding, and selection in evolution

Wright, 1932

- fitness landscapes yield information on the predictability and repeatability of evolution
- it becomes increasingly simple to measure empirical fitness landscapes
- accumulating data on gene networks and pathways

But:

THE CONCEPT OF A FITNESS LANDSCAPE



The roles of mutation, inbreeding, crossbreeding, and selection in evolution

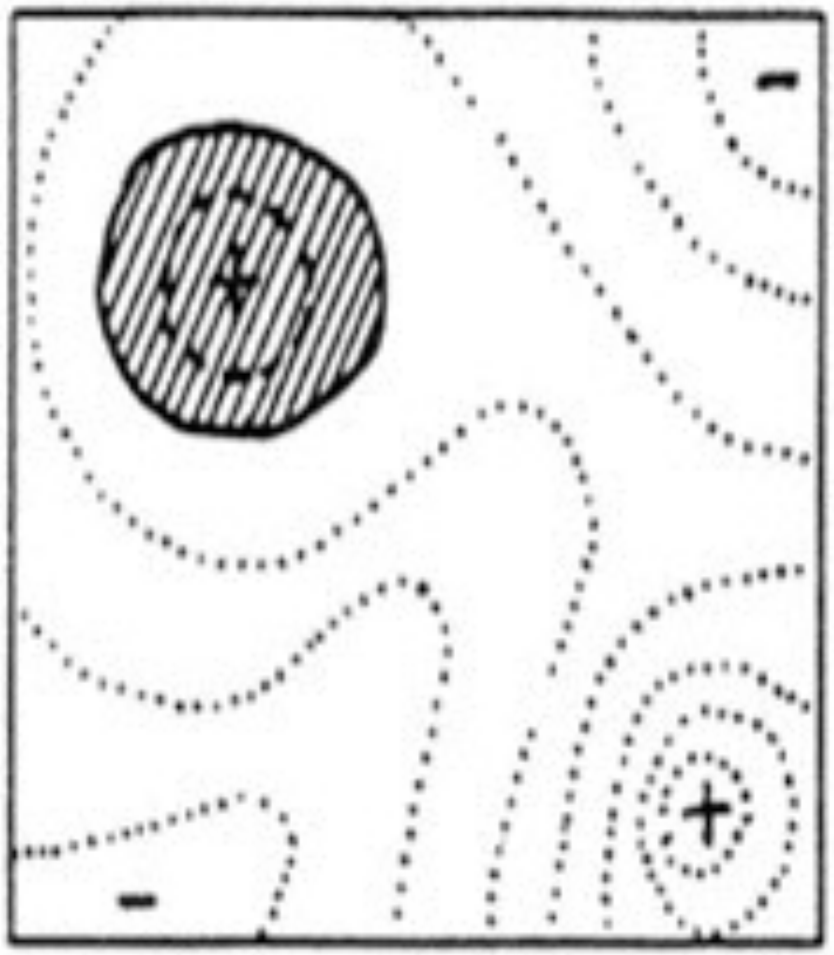
Wright, 1932

- fitness landscapes yield information on the predictability and repeatability of evolution
- it becomes increasingly simple to measure empirical fitness landscapes
- accumulating data on gene networks and pathways

But:

- very high complexity
- unclear whether there is predictive potential when combining theory and data

THE CONCEPT OF A FITNESS LANDSCAPE



The roles of mutation, inbreeding, crossbreeding, and selection in evolution
Wright, 1932

- fitness landscapes yield information on the predictability and repeatability of evolution
- it becomes increasingly simple to measure empirical fitness landscapes
- accumulating data on gene networks and pathways

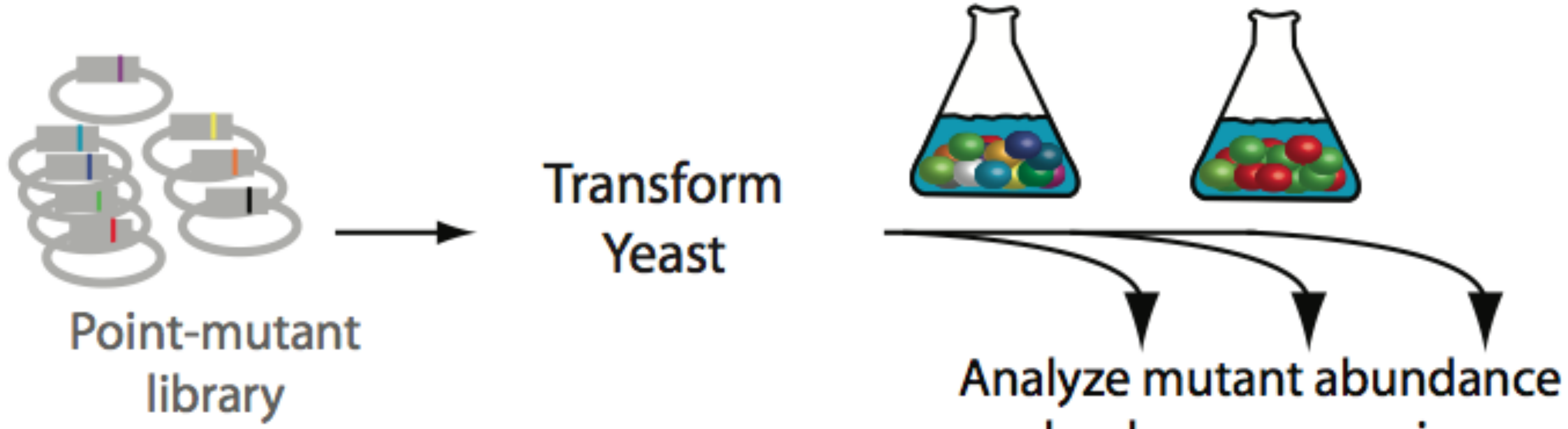
But:

- very high complexity
- unclear whether there is predictive potential when combining theory and data

E.g.: Can we predict costs of anti-microbial resistance across environments?

THE EMPIRIC APPROACH

- Systematic high-throughout sampling of hundreds of chosen mutations (including those that are strongly deleterious)
- Bulk competitions ensure identical conditions for all mutants
- Genetic background is precisely controlled (minimized potential for secondary mutations)



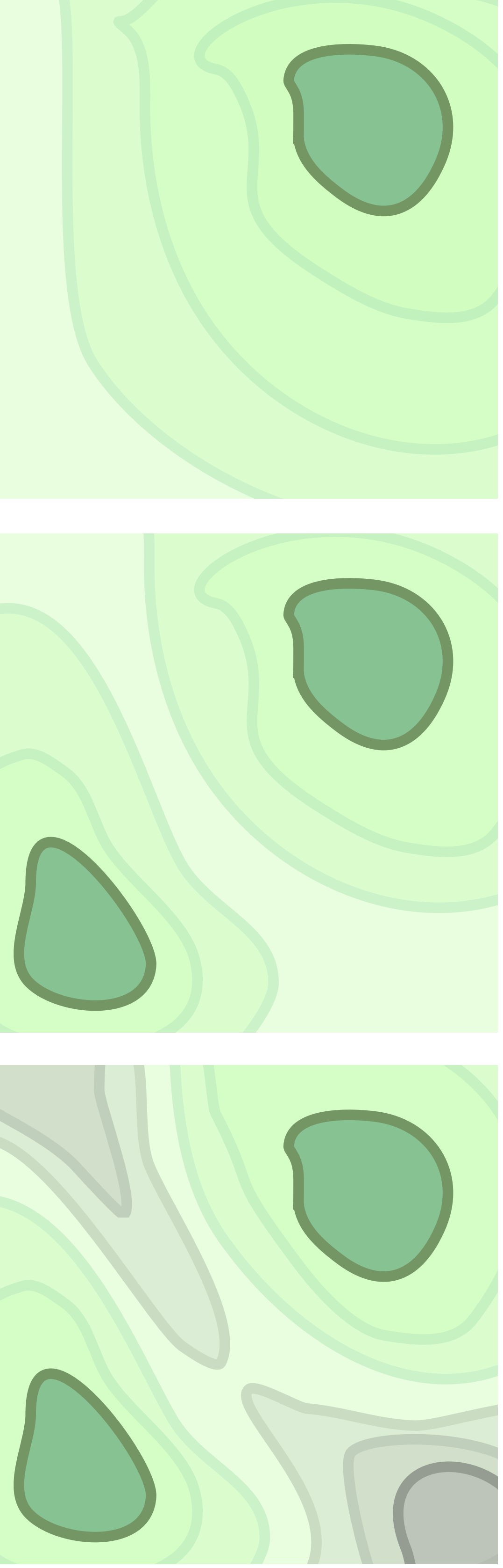
Point-mutant library

Transform Yeast

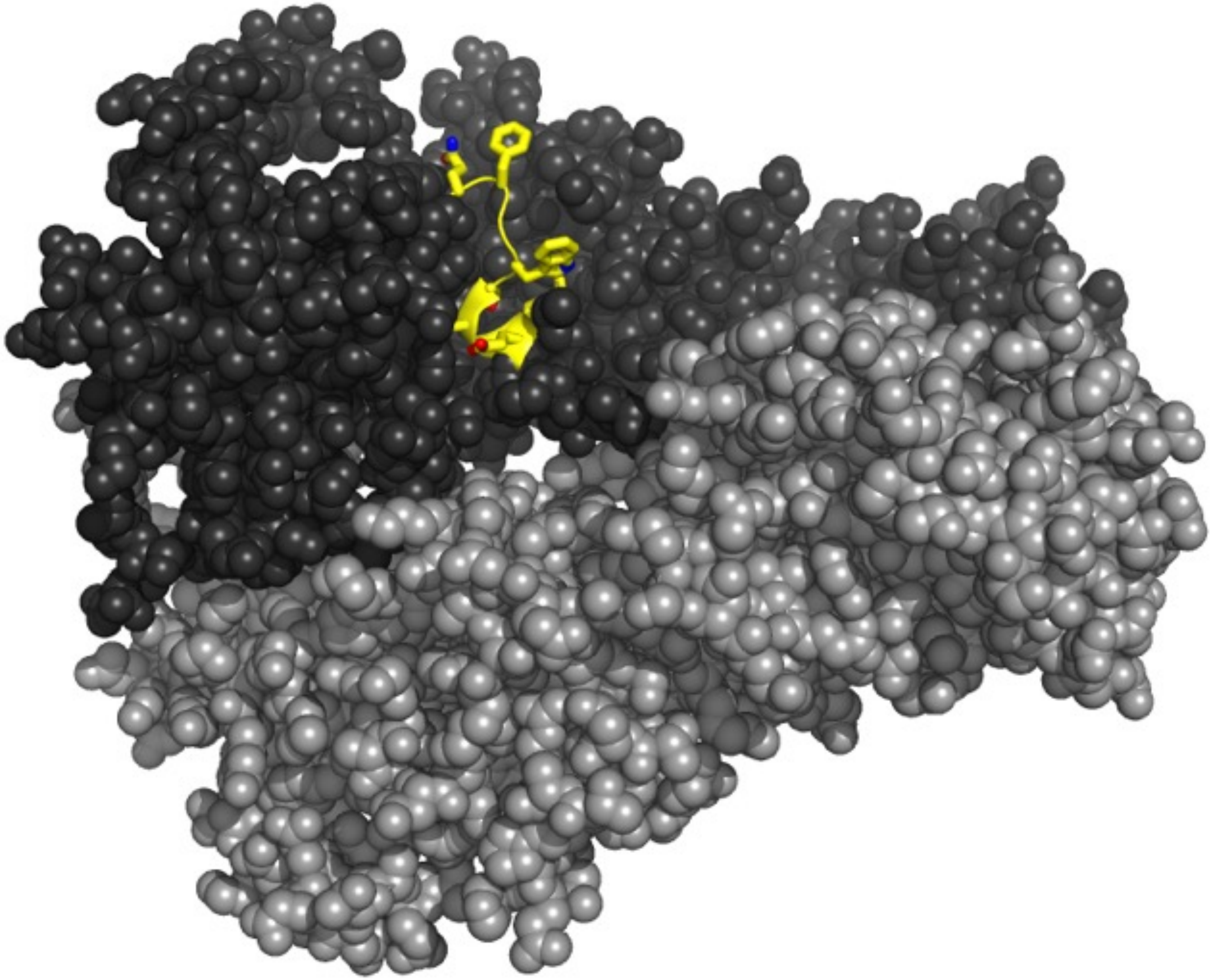
Analyze mutant abundance by deep sequencing

Hietpas, Jensen & Bolon, PNAS, 2011

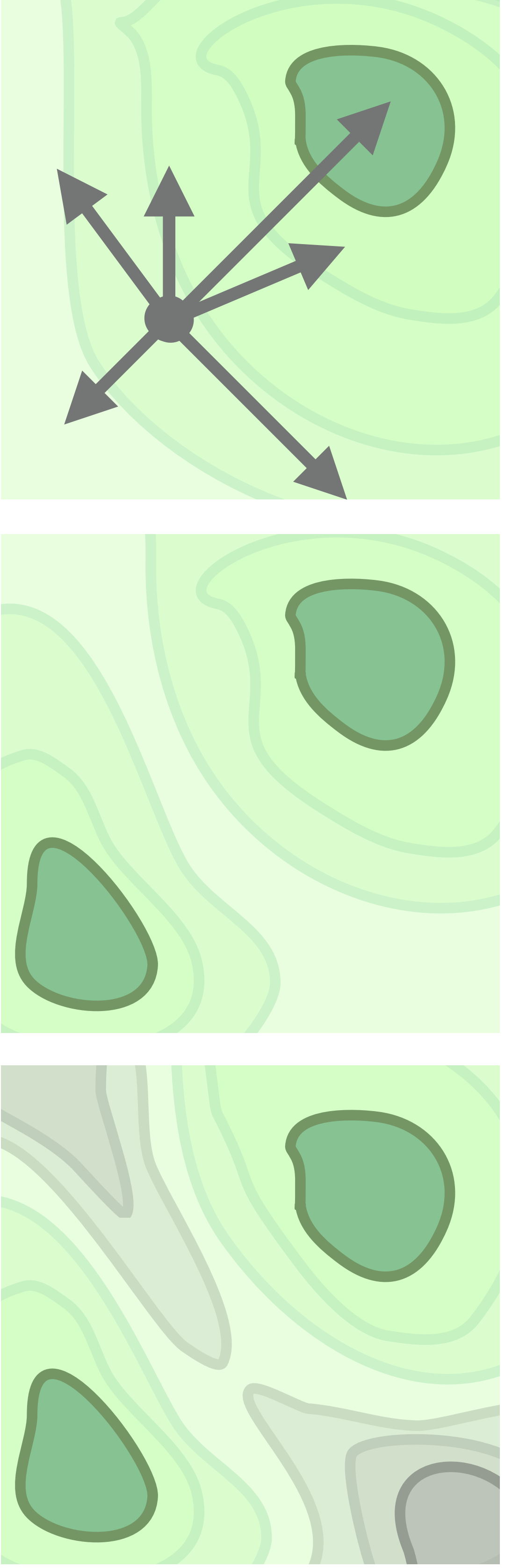
WHAT WE DID SO FAR



Engineered mutations from a 9 aa region from Hsp90 (aa positions 582-590) in *Saccharomyces cerevisiae*



WHAT WE DID SO FAR

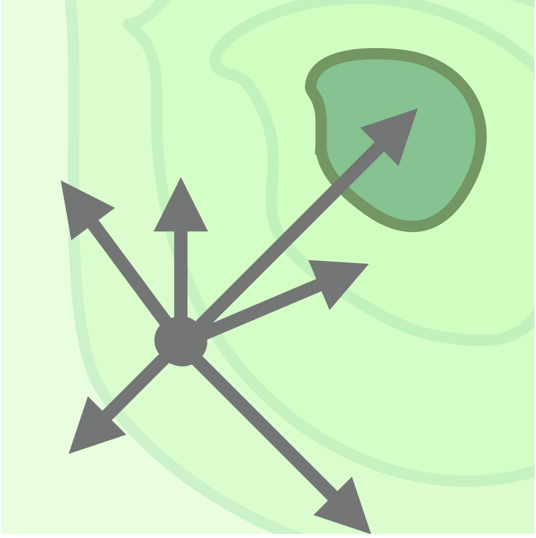


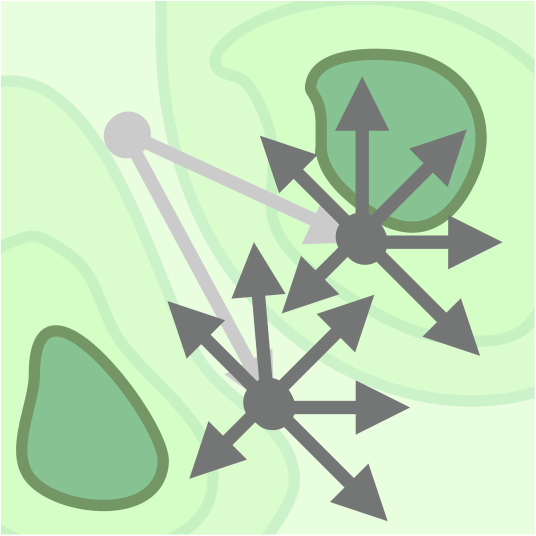
Engineered mutations from a 9 aa region from Hsp90 (aa positions 582-590) in *Saccharomyces cerevisiae*

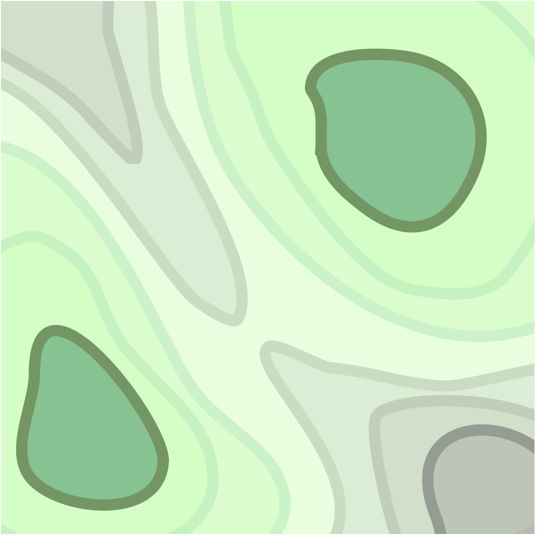
Compare DFEs across 4 environments - high costs of adaptation **Hietpas*, Bank* et al., 2013, Evolution**

MCMC method to estimate selection coefficients, and DFEs across 6 environments - heavy-tailed DFE for most challenging environment **Bank et al., 2014, Genetics**

WHAT WE DID SO FAR







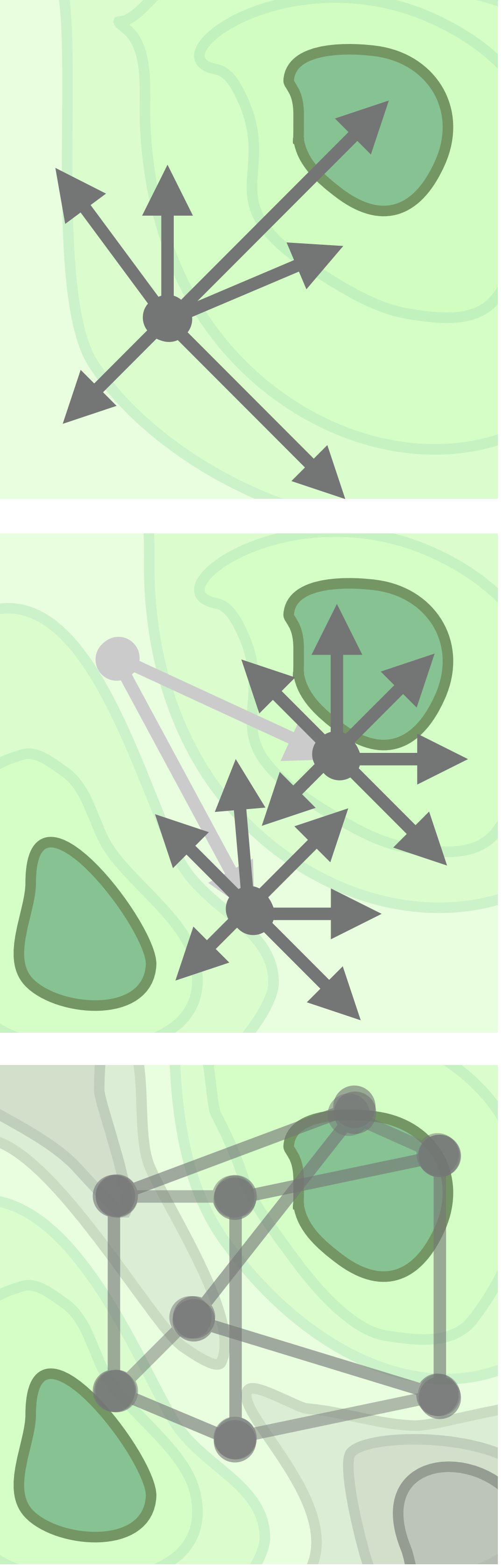
Engineered mutations from a 9 aa region from Hsp90 (aa positions 582-590) in *Saccharomyces cerevisiae*

Compare DFEs across 4 environments - high costs of adaptation **Hietpas*, Bank* et al., 2013, Evolution**

MCMC method to estimate selection coefficients, and DFEs across 6 environments - heavy-tailed DFE for most challenging environment **Bank et al., 2014, Genetics**

DFEs on 7 genetic backgrounds - ubiquitous negative epistasis indicating an underlying concave fitness landscape **Bank et al., 2015, MBE**

WHAT WE DID SO FAR



Engineered mutations from a 9 aa region from Hsp90 (aa positions 582-590) in *Saccharomyces cerevisiae*

Compare DFEs across 4 environments - high costs of adaptation **Hietpas*, Bank* et al., 2013, Evolution**

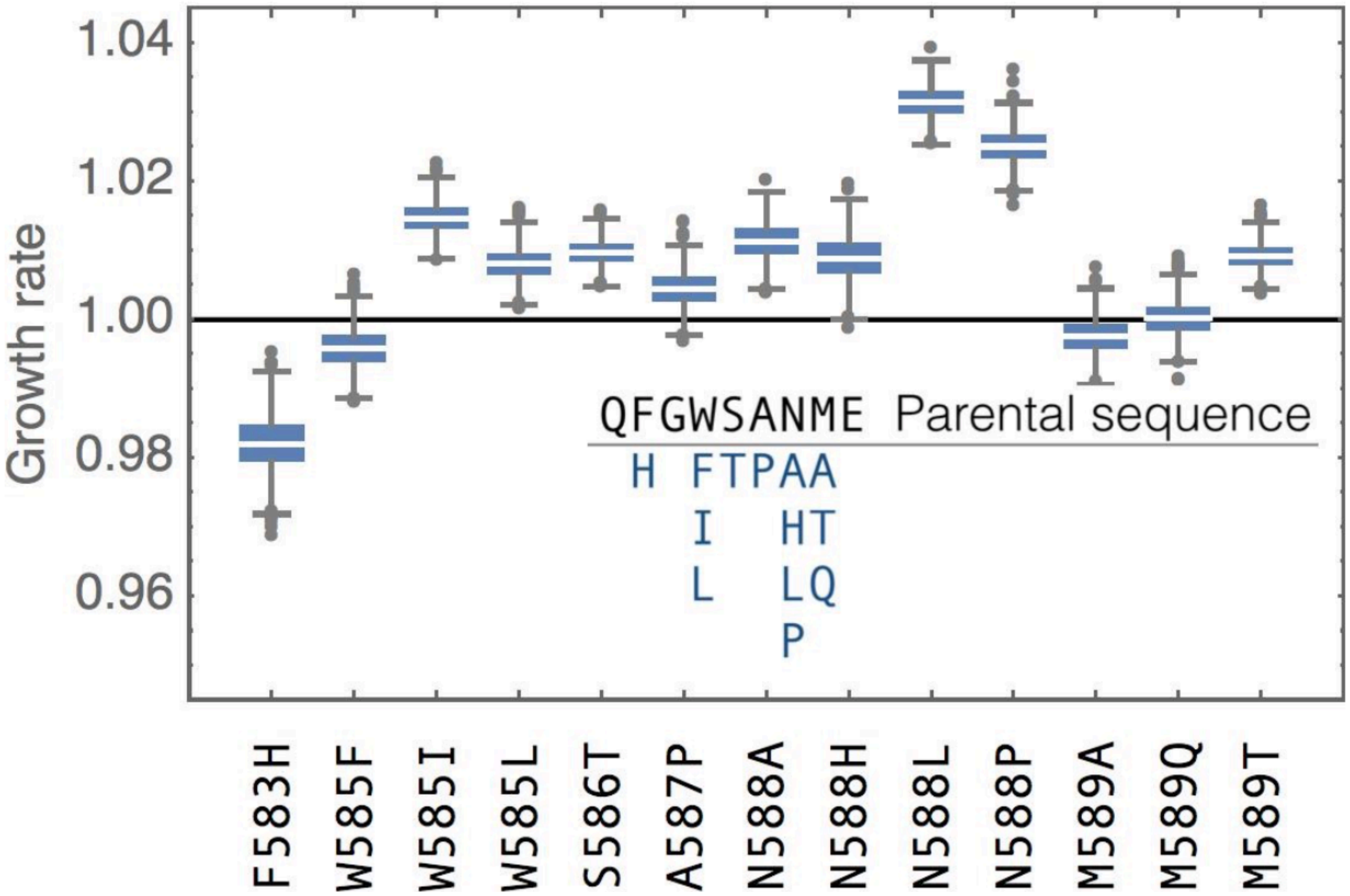
MCMC method to estimate selection coefficients, and DFEs across 6 environments - heavy-tailed DFE for most challenging environment **Bank et al., 2014, Genetics**

DFEs on 7 genetic backgrounds - ubiquitous negative epistasis indicating an underlying concave fitness landscape **Bank et al., 2015, MBE**

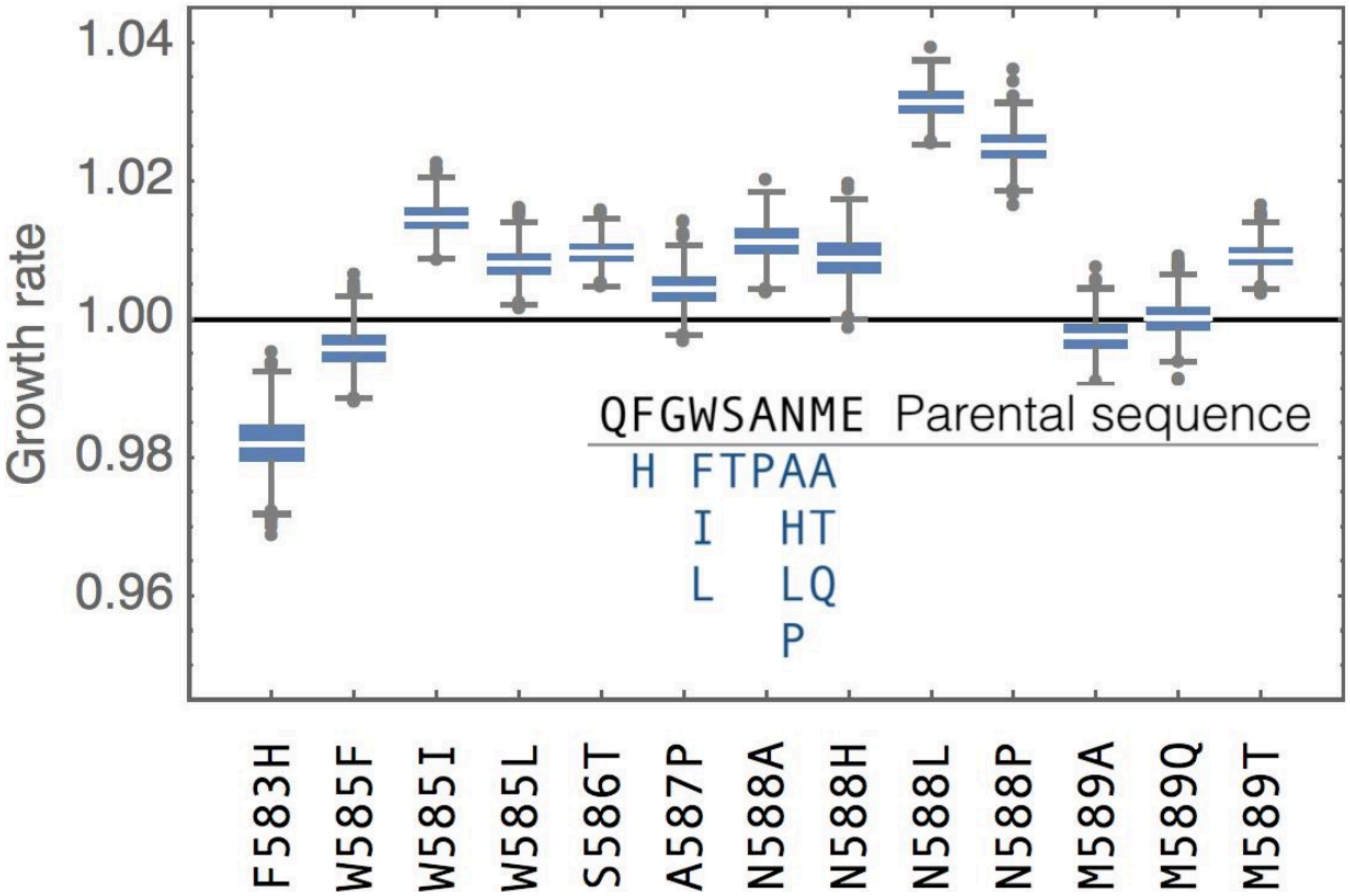
Guide to experimental design of deep mutational scanning studies **Matuszewski*, Hildebrandt* et al., 2016, Genetics**

Complete fitness landscape of 640 combinations of mutations **Bank*, Matuszewski* et al., BioRxiv**

high salinity environment
13 single-aa mutations
2 replicates
all possible combinations of aa's
≈1600 nt mutations

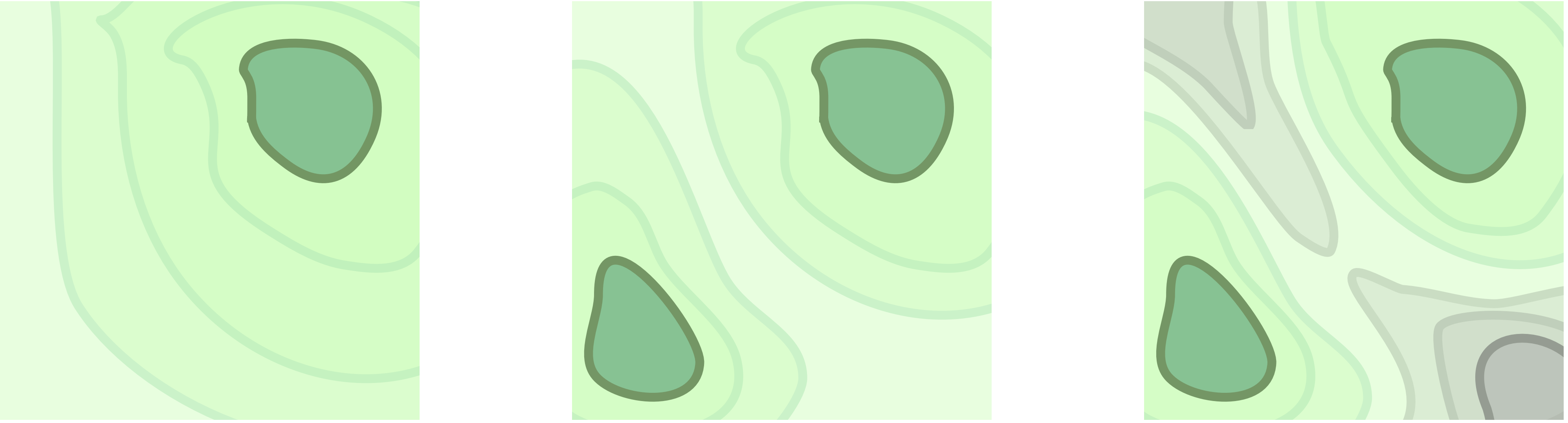


high salinity environment
13 single-aa mutations
2 replicates
all possible combinations of aa's
≈1600 nt mutations



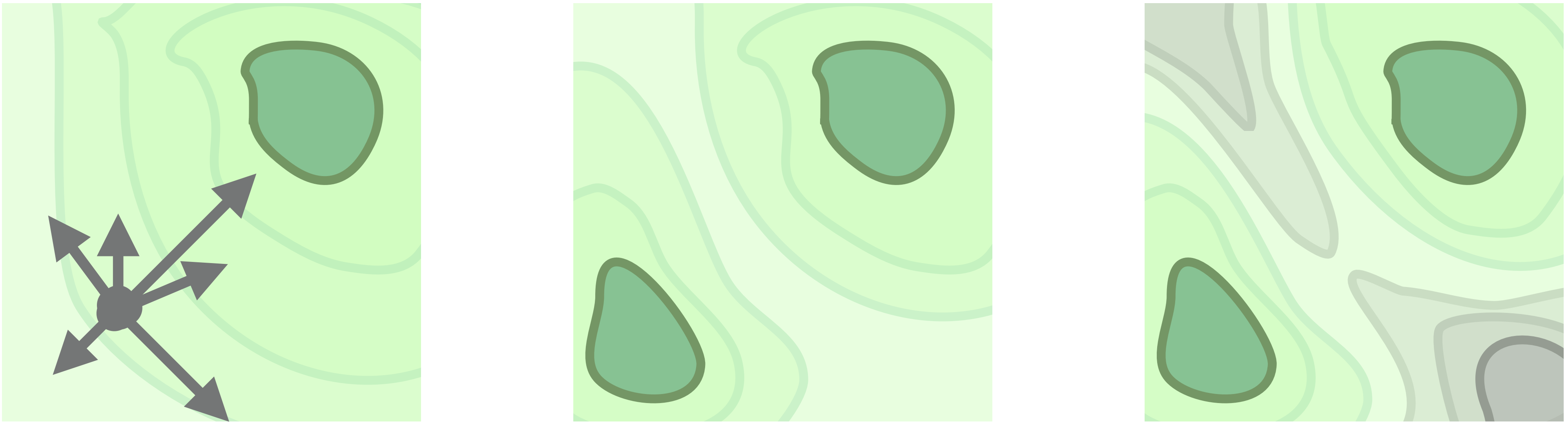
- relatively “unbiased" selection of mutations
- multi-allelic fitness landscape

3 QUESTIONS



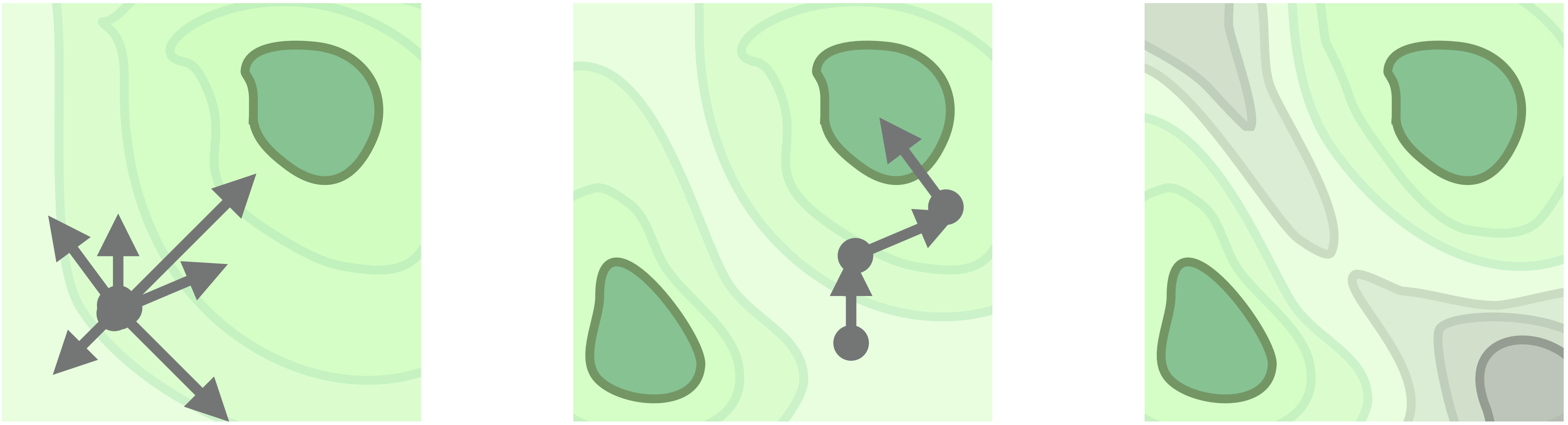
3 QUESTIONS

- Do single step mutations predict the way to the global optimum?



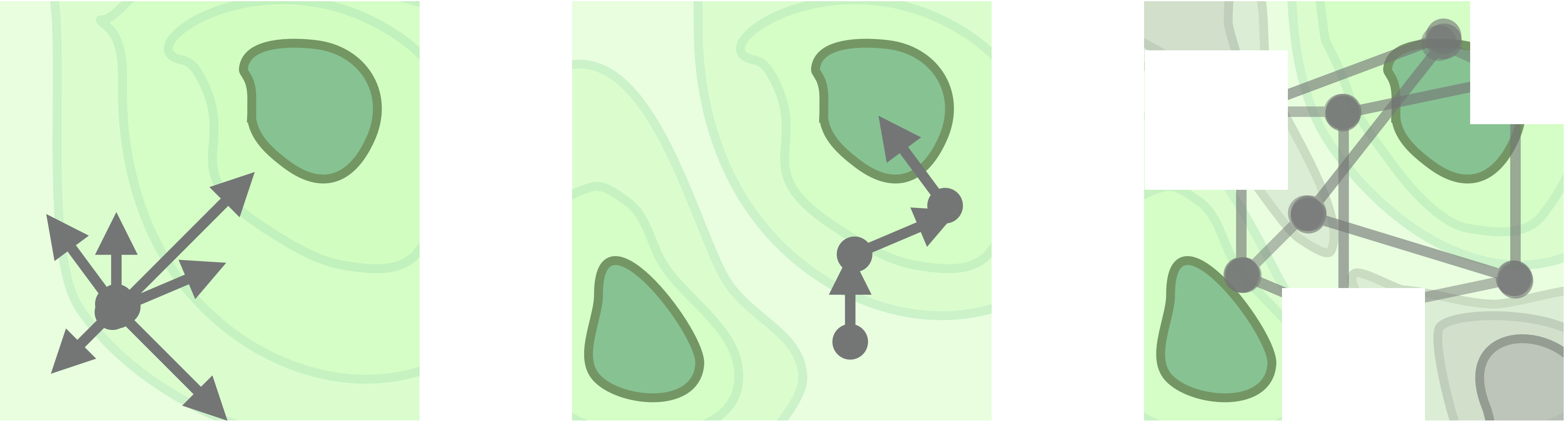
3 QUESTIONS

- Do single step mutations predict the way to the global optimum?
- Will adaptation take the population to the global optimum?

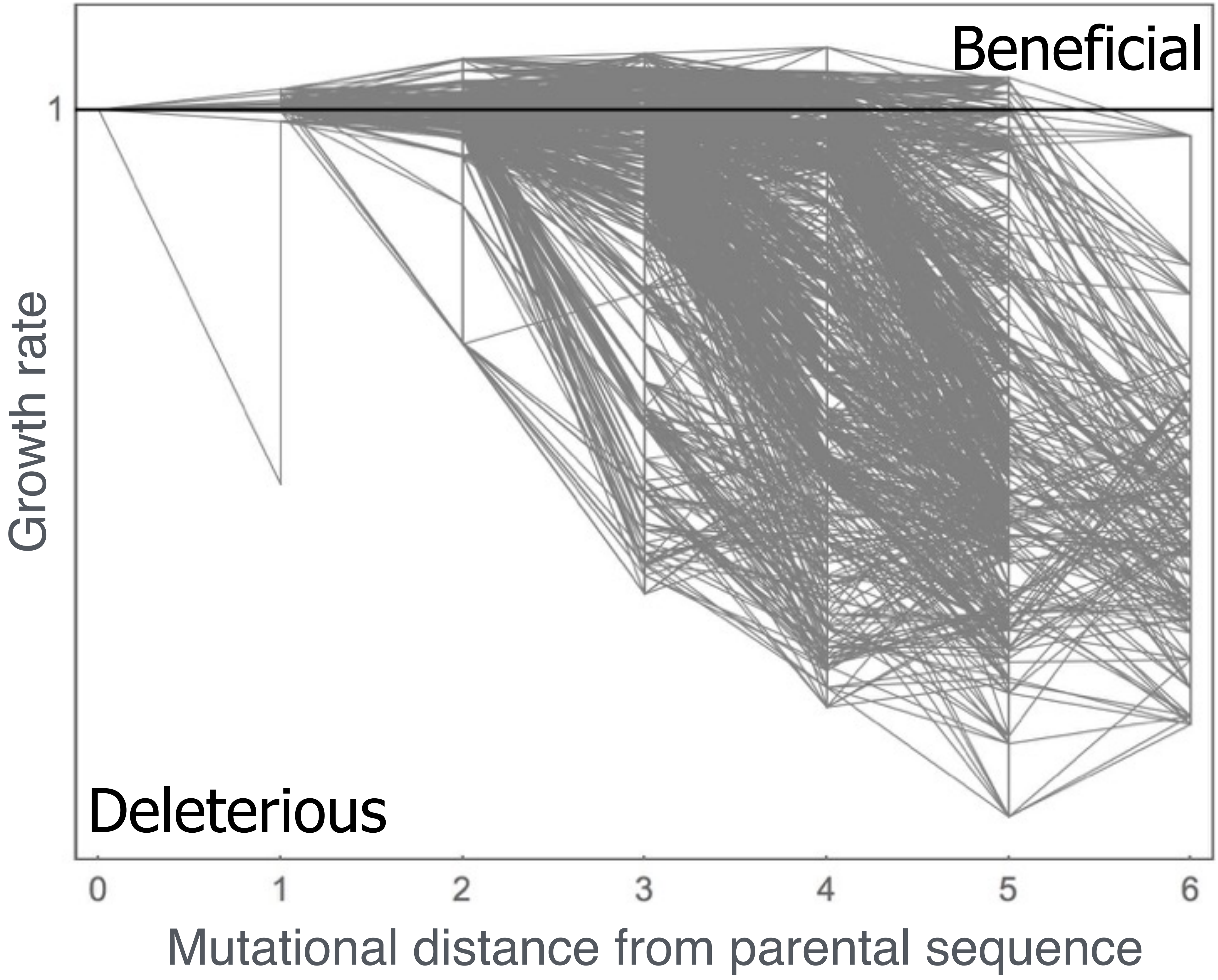


3 QUESTIONS

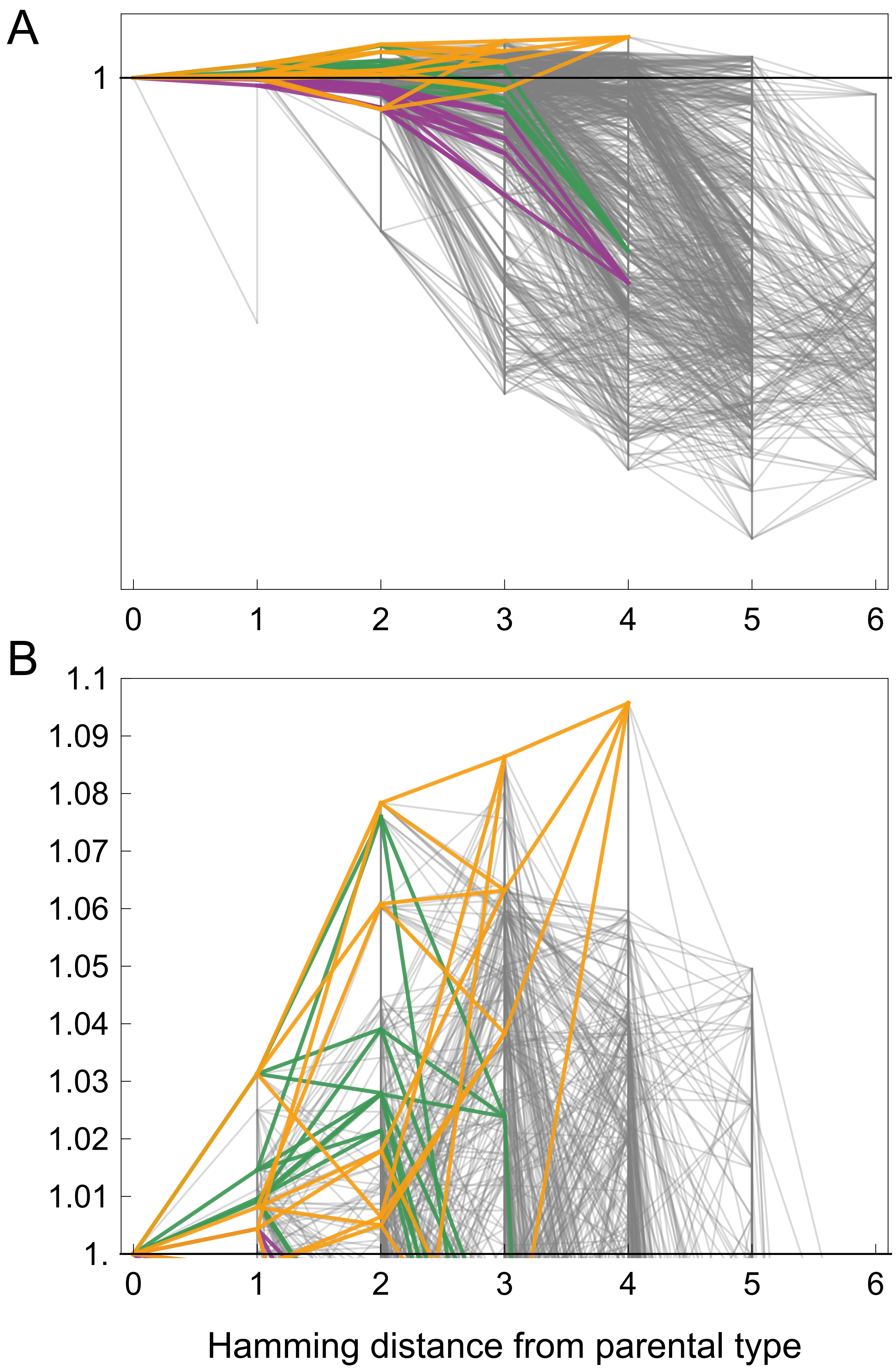
- Do single step mutations predict the way to the global optimum?
- Will adaptation take the population to the global optimum?
- Can we infer an unknown part of the fitness landscape?



A PICTURE OF THE WHOLE LANDSCAPE



1 – DO SINGLE STEP MUTATIONS PREDICT THE WAY TO THE GLOBAL OPTIMUM?

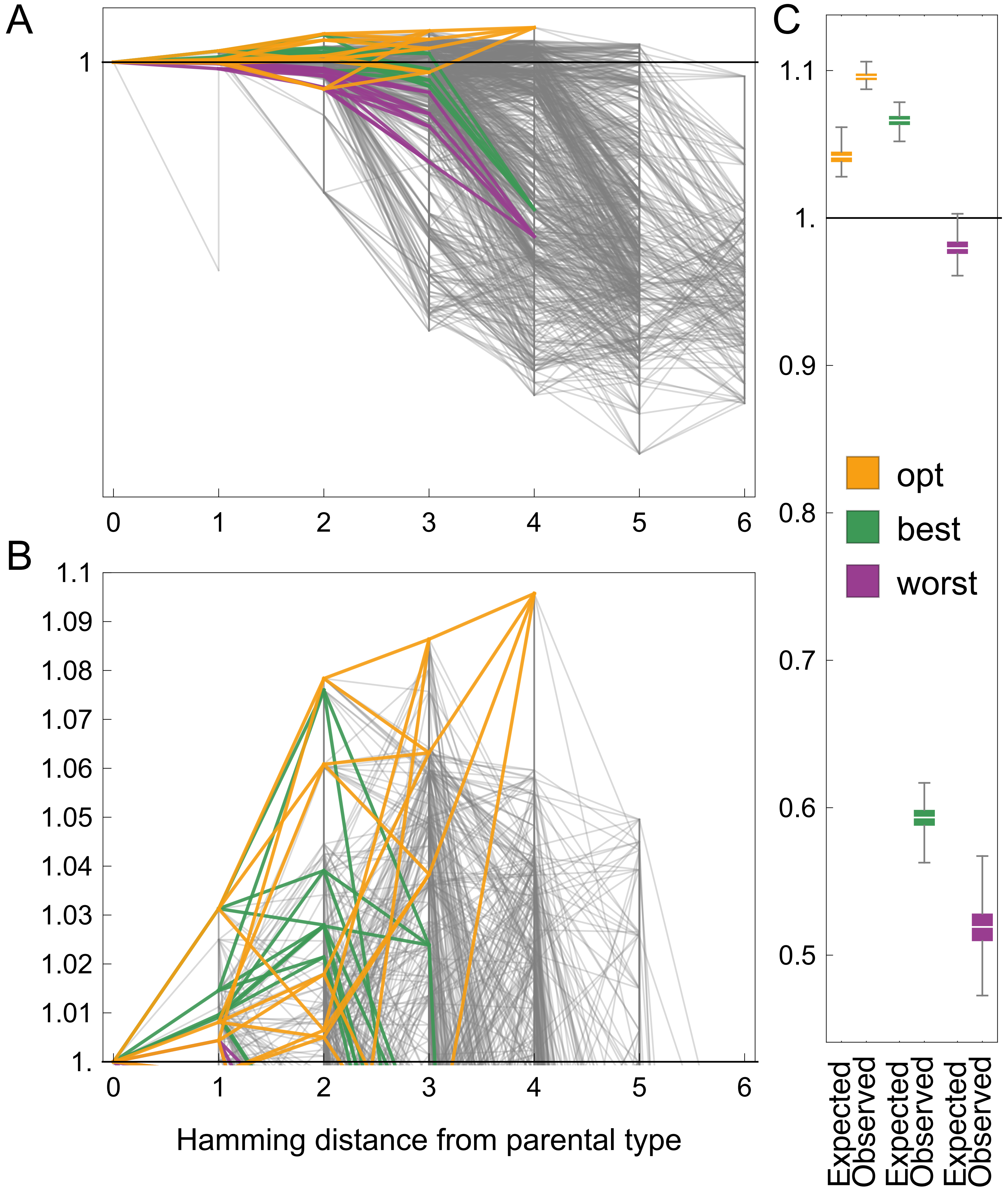


Focal landscapes:

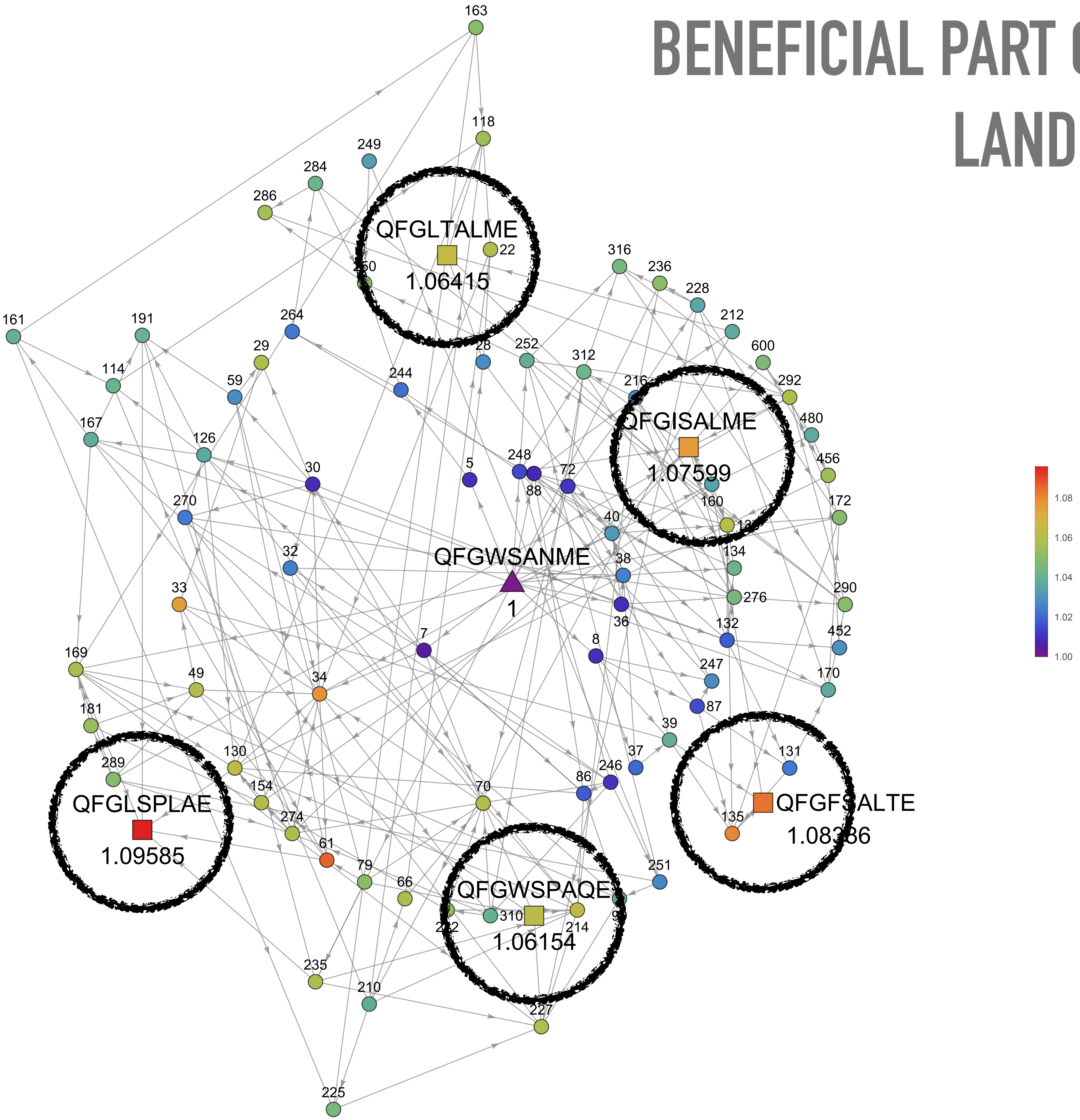
1) leading to global opt

2) best 4 mutations

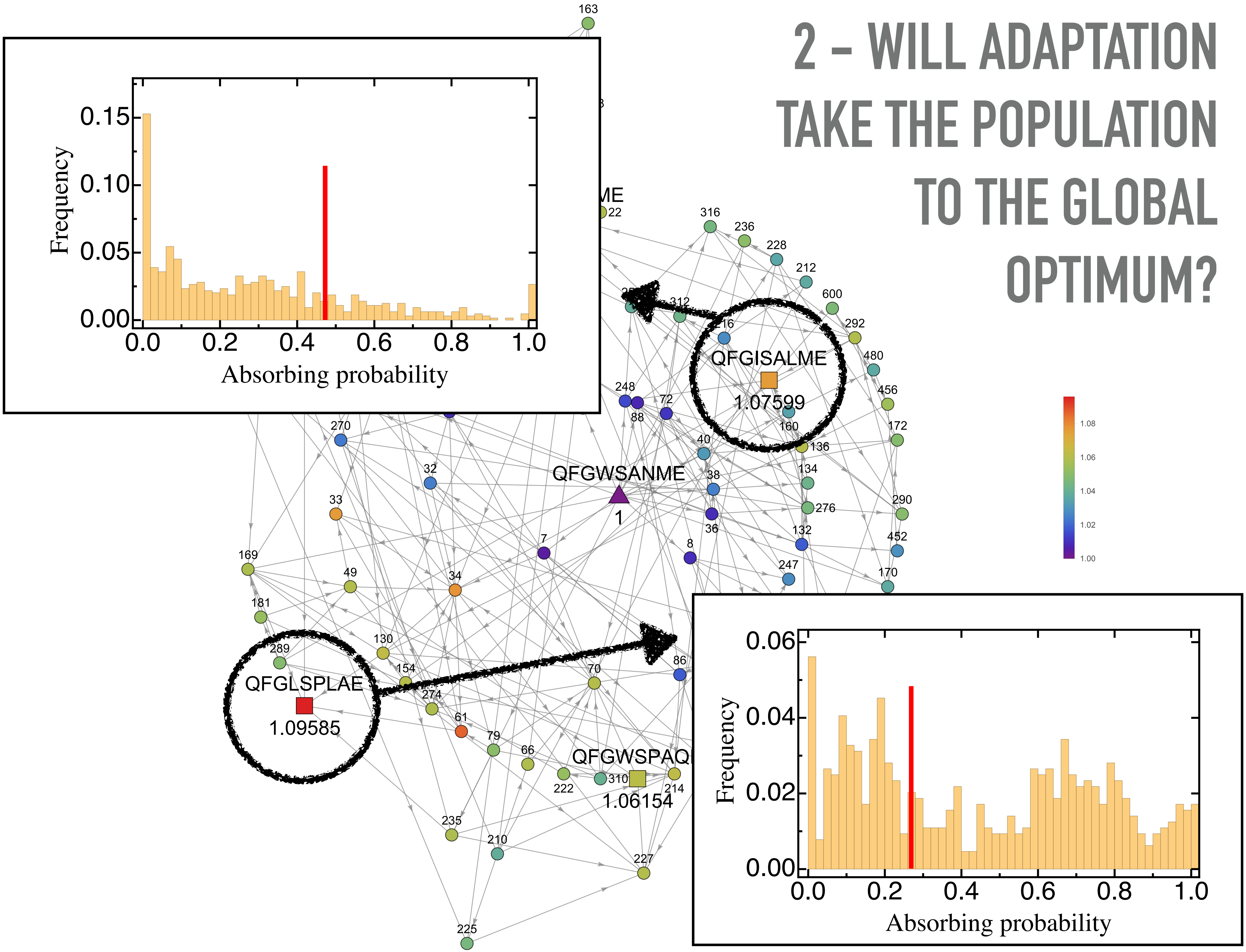
3) “worst” 4 mutations

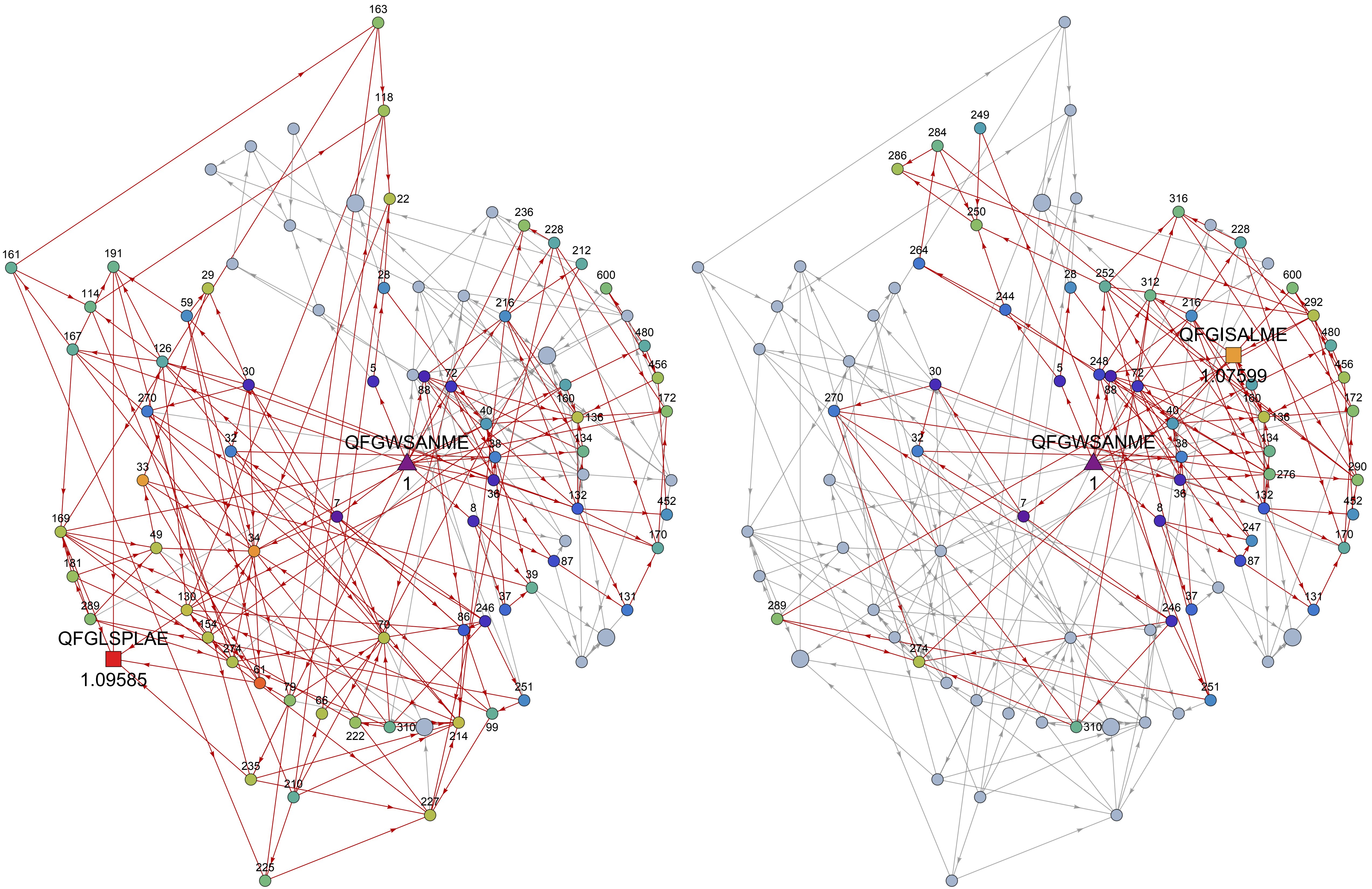

A
1
0
1
2
3
4
5
6
B
1.1
1.09
1.08
1.07
1.06
1.05
1.04
1.03
1.02
1.01
1.
0
1
2
3
4
5
6
Hamming distance from parental type
C
1.1
1.
0.9
0.8
0.7
0.6
0.5
opt
best
worst
Expected
Observed
Expected
Observed
Expected
Observed

BENEFICIAL PART OF THE LANDSCAPE



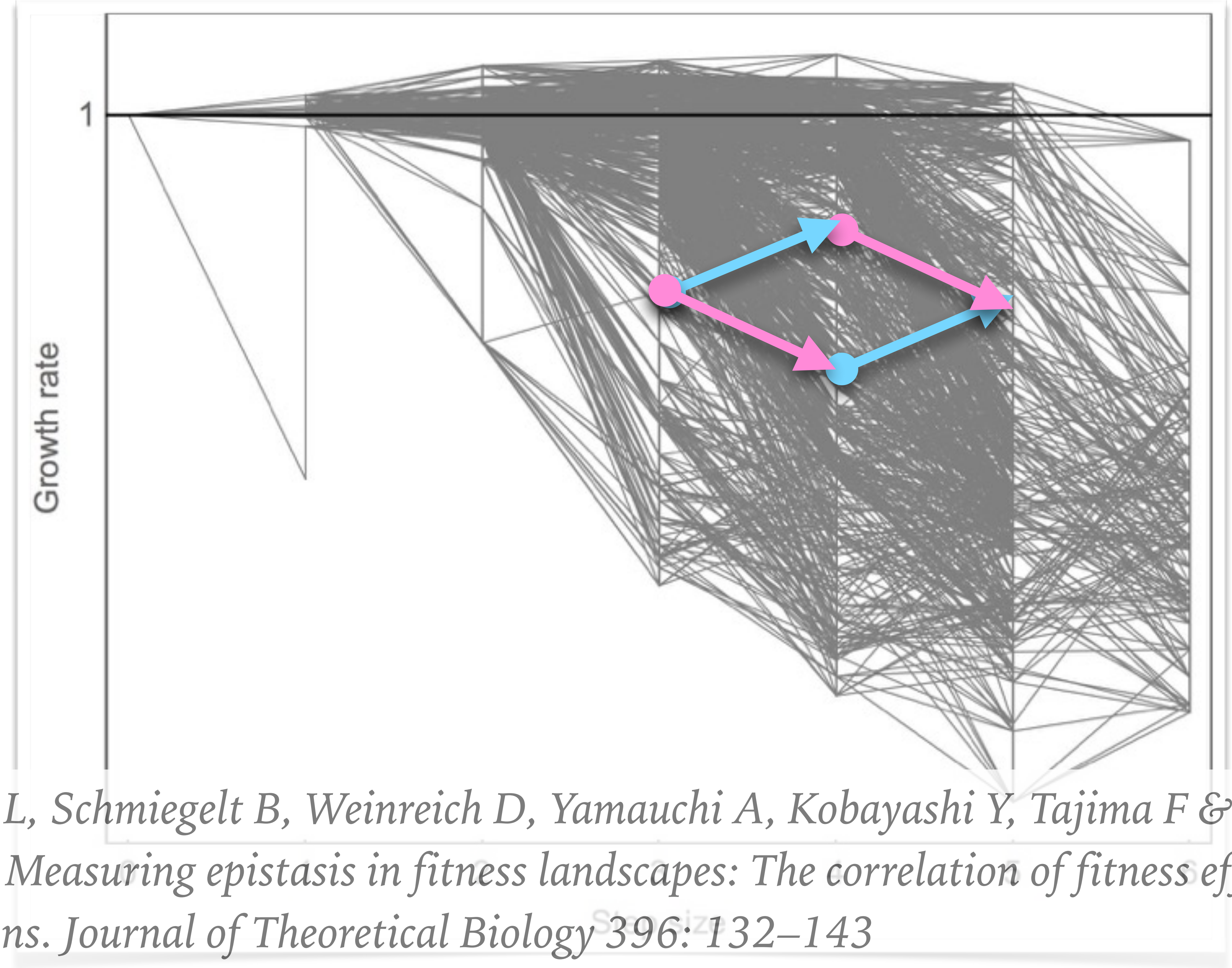
2 – WILL ADAPTATION TAKE THE POPULATION TO THE GLOBAL OPTIMUM?





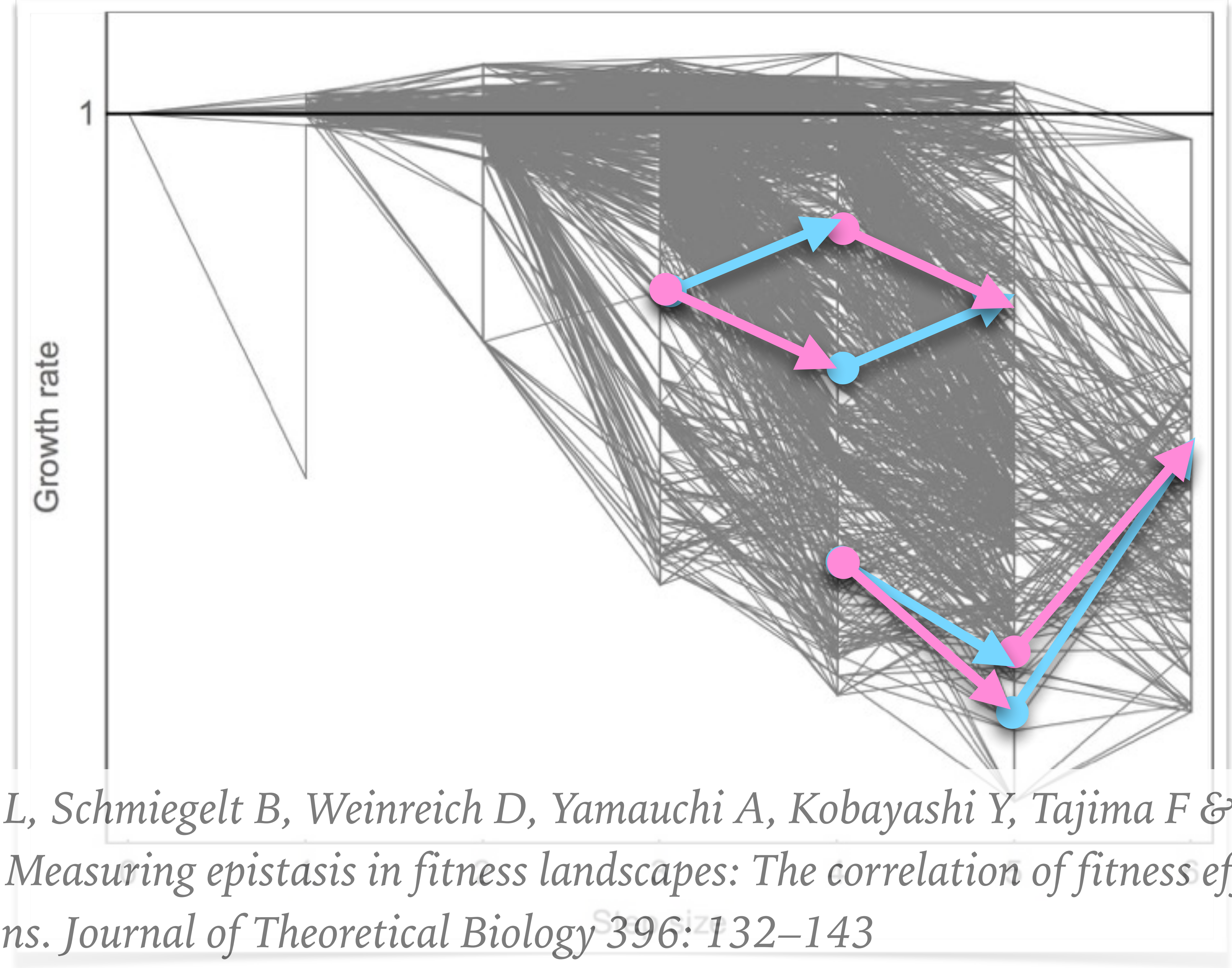
Evolution from parental type may stall at local optimum

LANDSCAPE STATISTICS INDEPENDENT OF REFERENCE



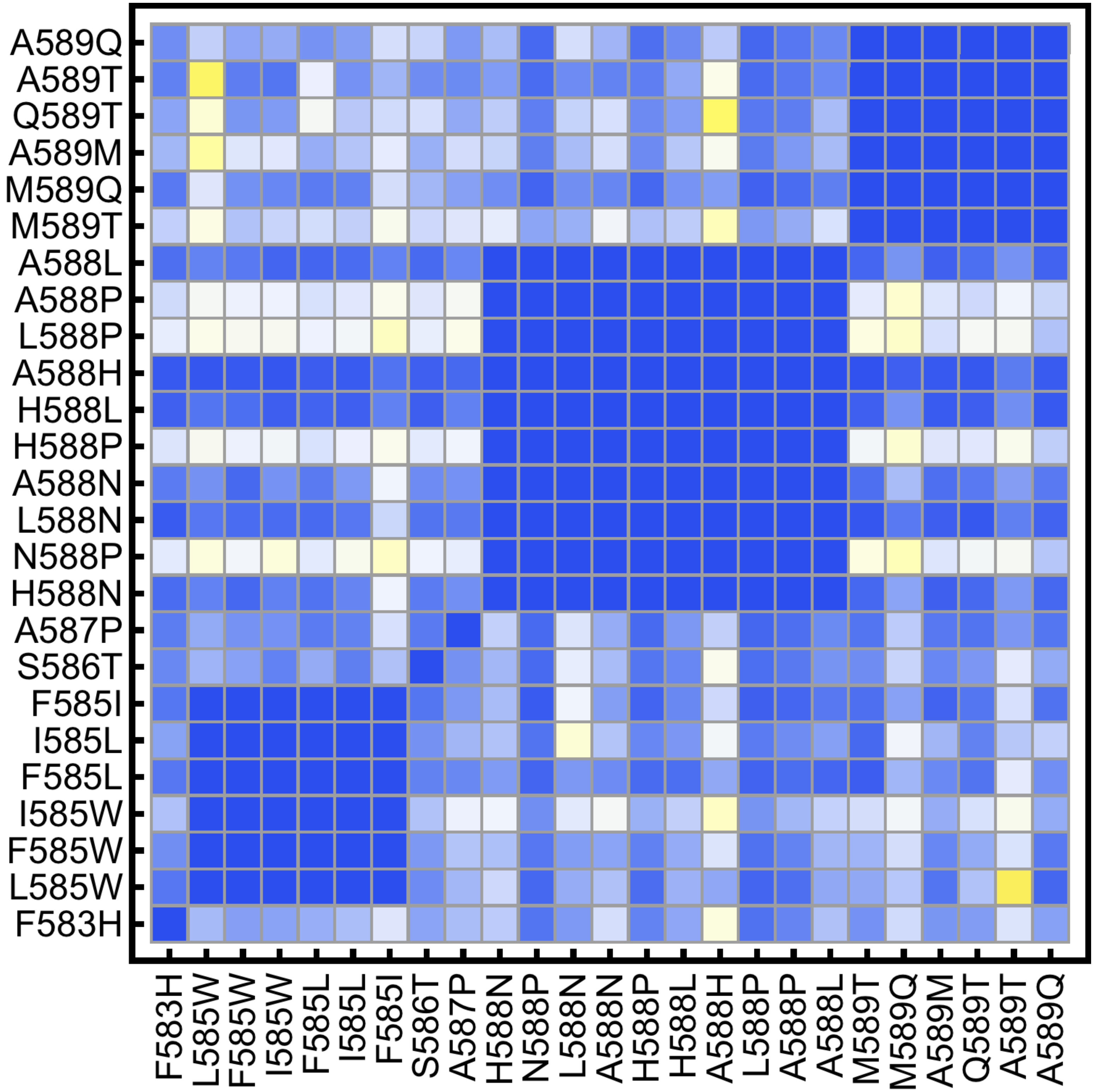
Ferretti L, Schmiegelt B, Weinreich D, Yamauchi A, Kobayashi Y, Tajima F & Achaz G (2016) Measuring epistasis in fitness landscapes: The correlation of fitness effects of mutations. Journal of Theoretical Biology 396: 132–143

LANDSCAPE STATISTICS INDEPENDENT OF REFERENCE

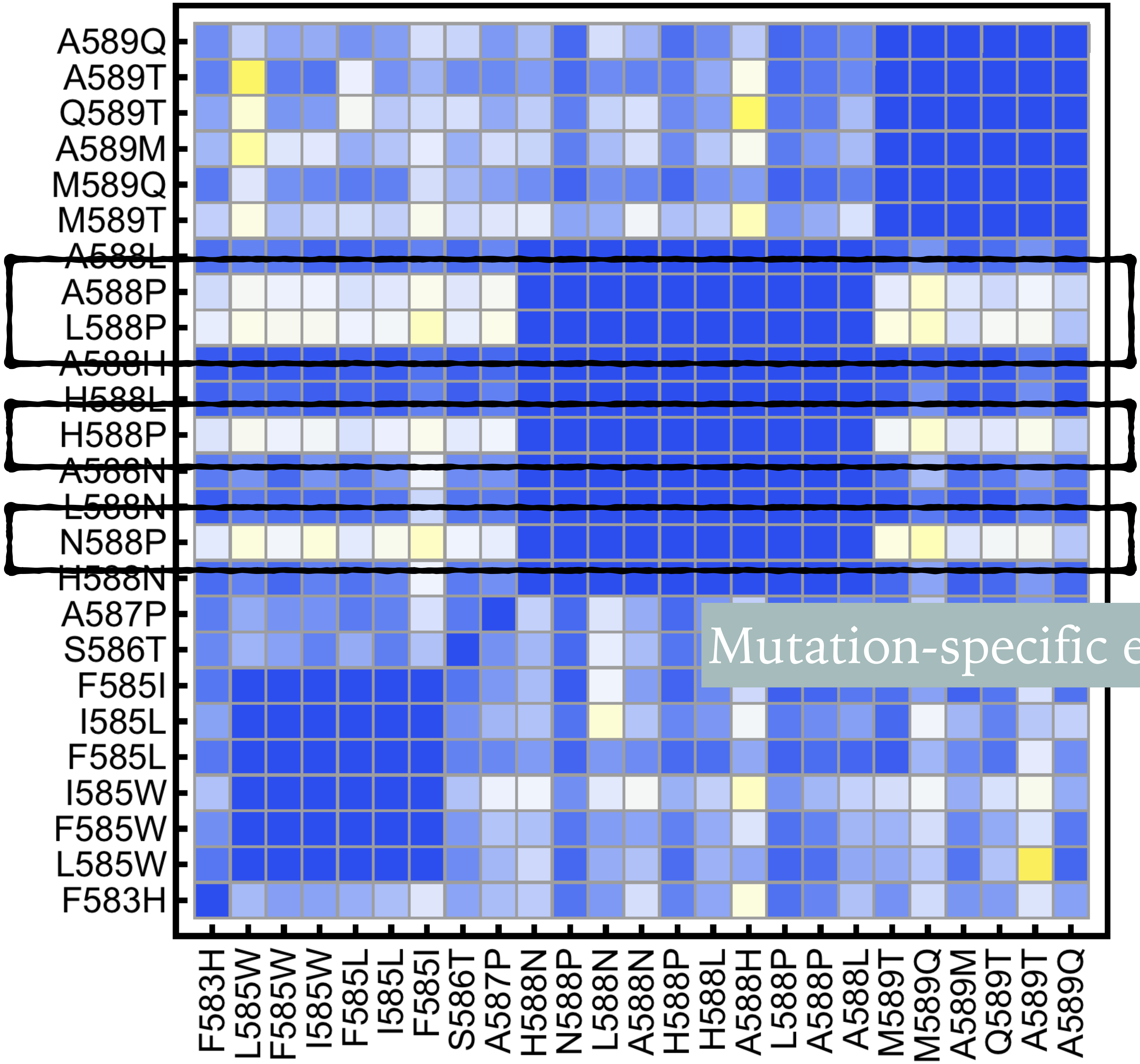


Ferretti L, Schmiegelt B, Weinreich D, Yamauchi A, Kobayashi Y, Tajima F & Achaz G (2016) Measuring epistasis in fitness landscapes: The correlation of fitness effects of mutations. Journal of Theoretical Biology 396: 132–143

HOW CAN WE MEASURE FITNESS LANDSCAPES AND WHAT CAN WE LEARN FROM THIS EXERCISE?

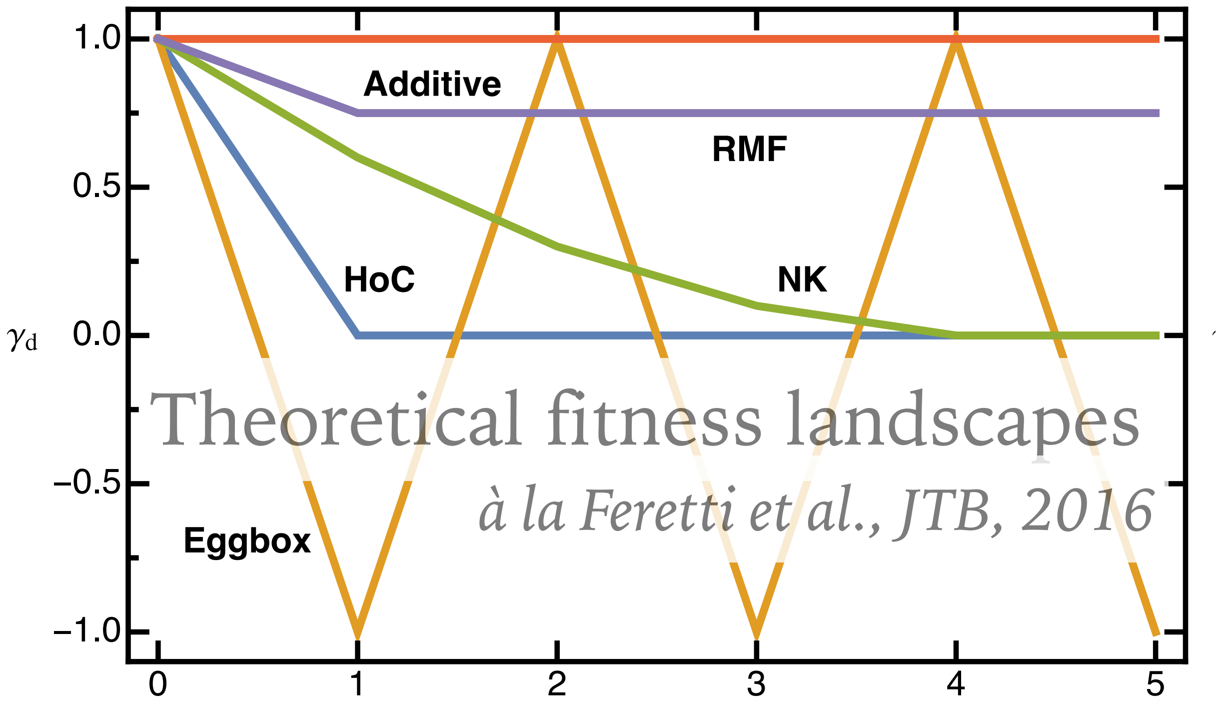


HOW CAN WE MEASURE FITNESS LANDSCAPES AND WHAT CAN WE LEARN FROM THIS EXERCISE?



Mutation-specific epistatic hotspot

3 – CAN WE INFER AN UNKNOWN PART OF THE LANDSCAPE?

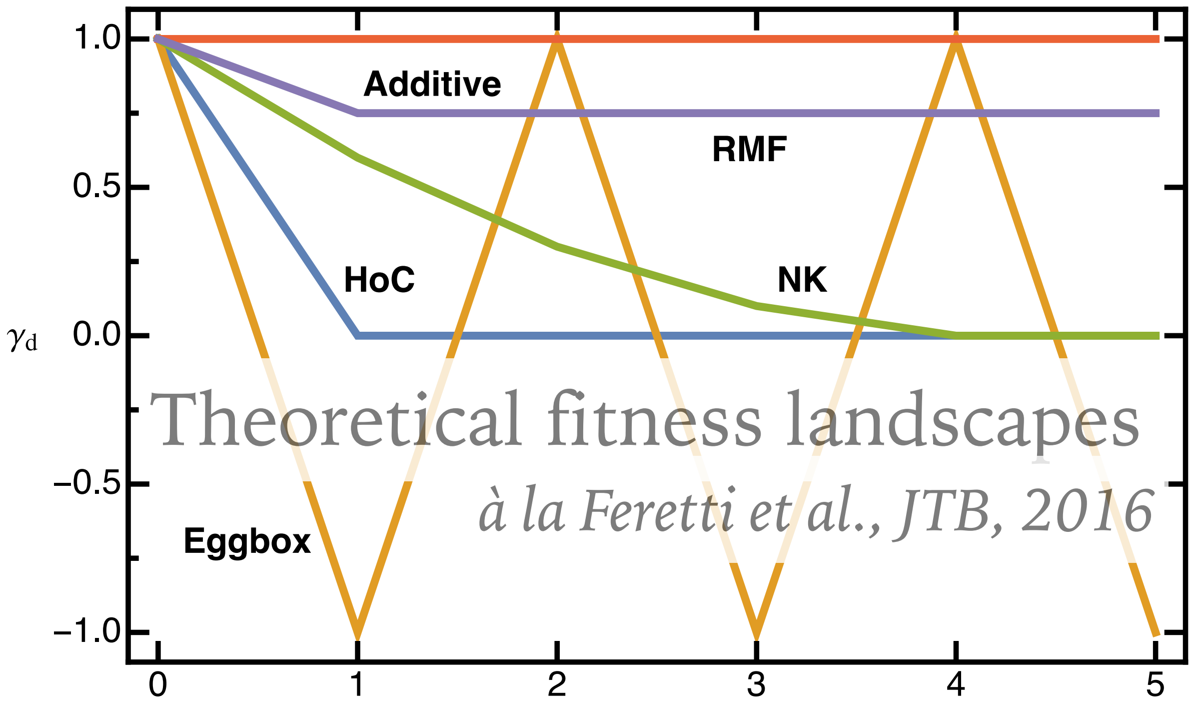


Theoretical fitness landscapes

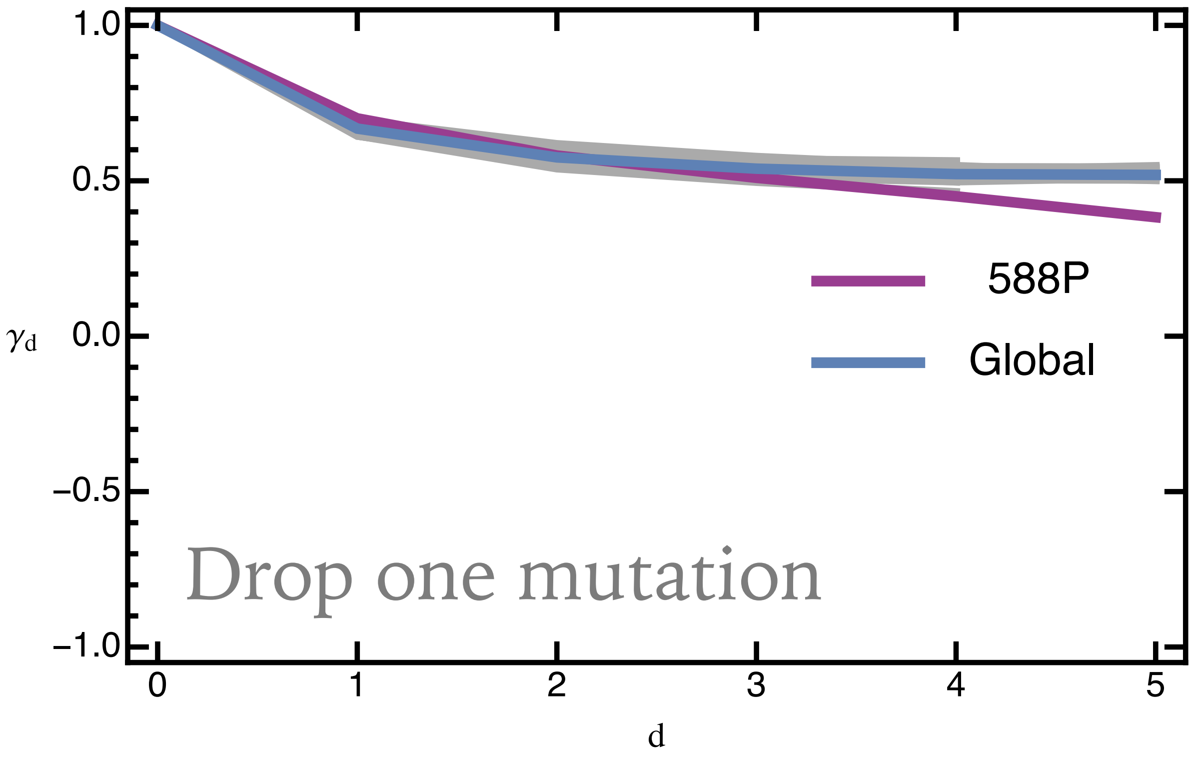
à la Feretti et al., JTB, 2016

Epistasis measure gamma$_d$ with increasing Hamming distance d

3 – CAN WE INFER AN UNKNOWN PART OF THE LANDSCAPE?

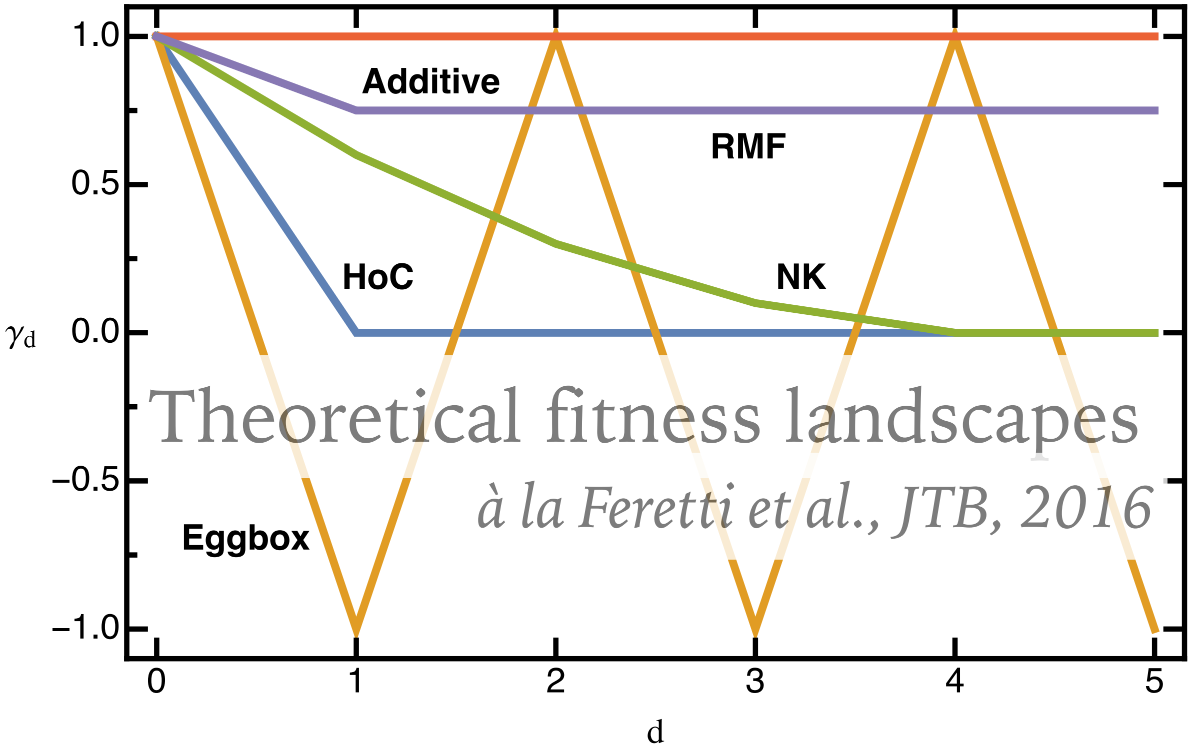

Theoretical fitness landscapes

à la Feretti et al., JTB, 2016


Drop one mutation

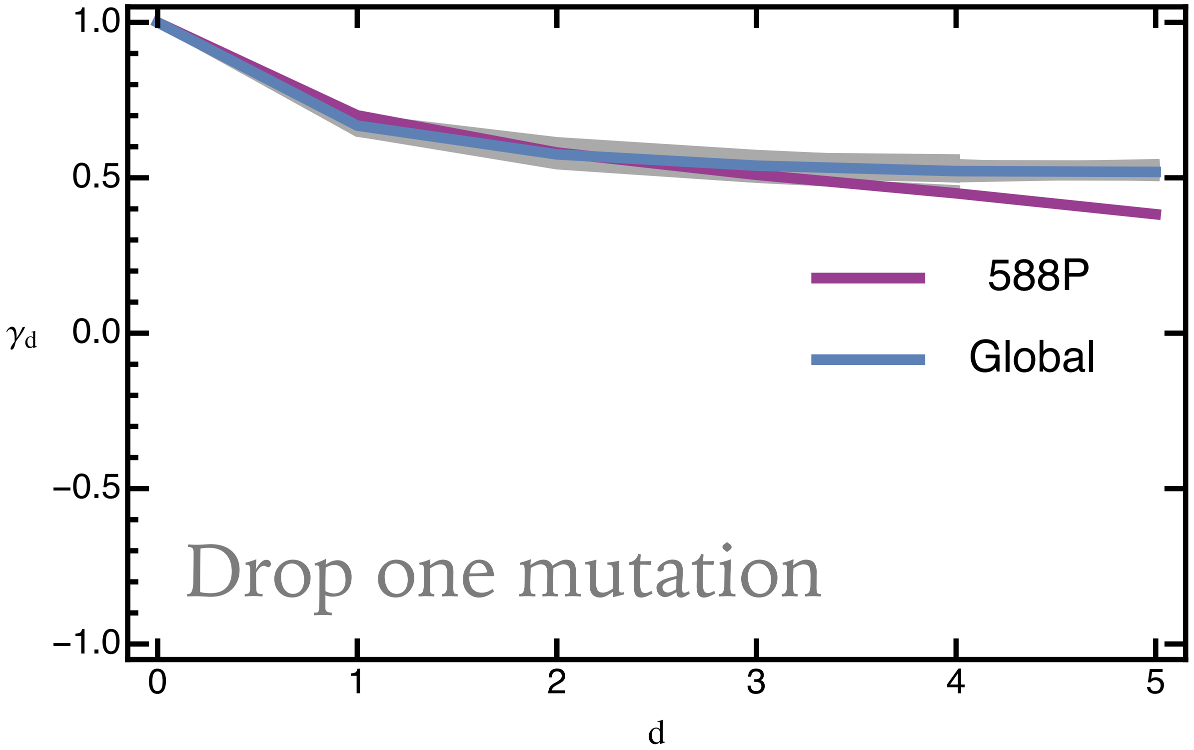
Epistasis measure gamma$_d$ with increasing Hamming distance d

3 – CAN WE INFER AN UNKNOWN PART OF THE LANDSCAPE?

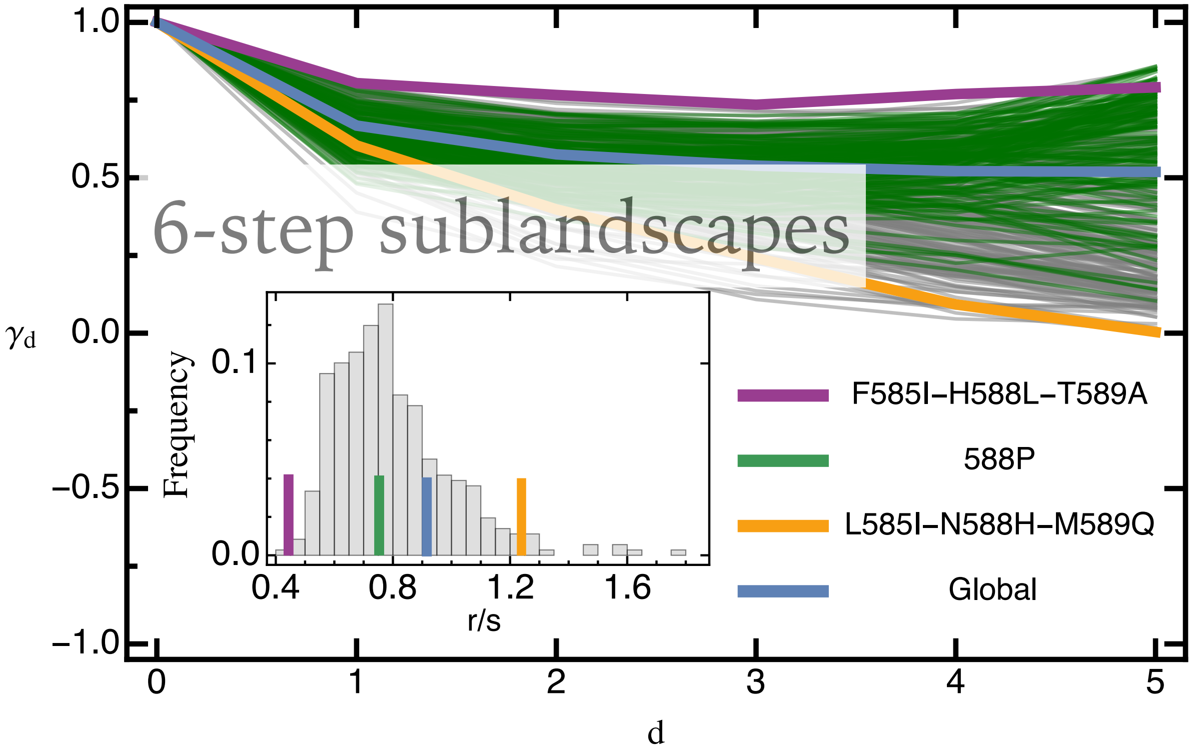


Theoretical fitness landscapes

à la Feretti et al., JTB, 2016



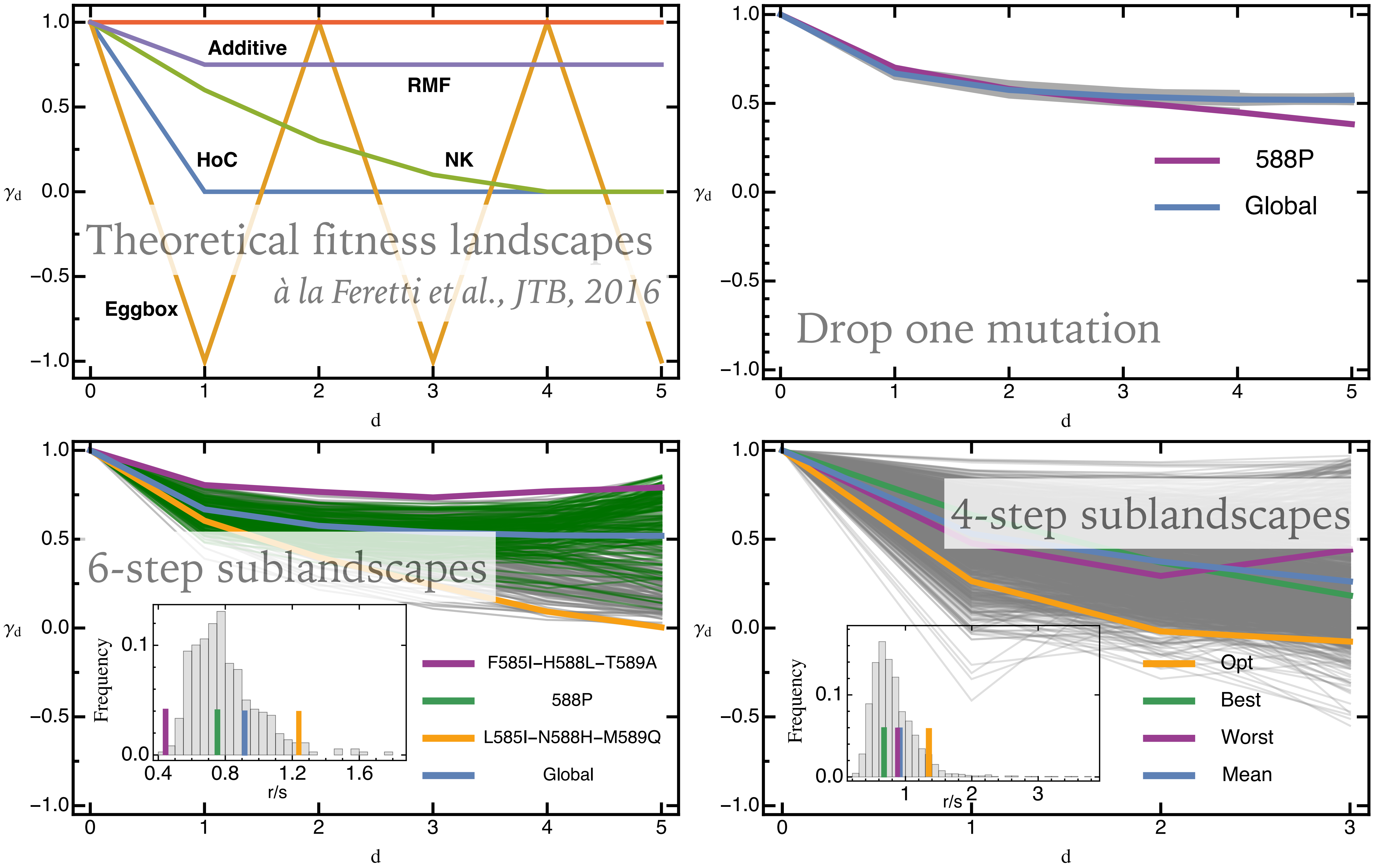
Drop one mutation



6-step sublandscapes

Epistasis measure gamma$_d$ with increasing Hamming distance d

3 – CAN WE INFER AN UNKNOWN PART OF THE LANDSCAPE?


Theoretical fitness landscapes

à la Feretti et al., JTB, 2016

Drop one mutation

6-step sublandscapes

4-step sublandscapes

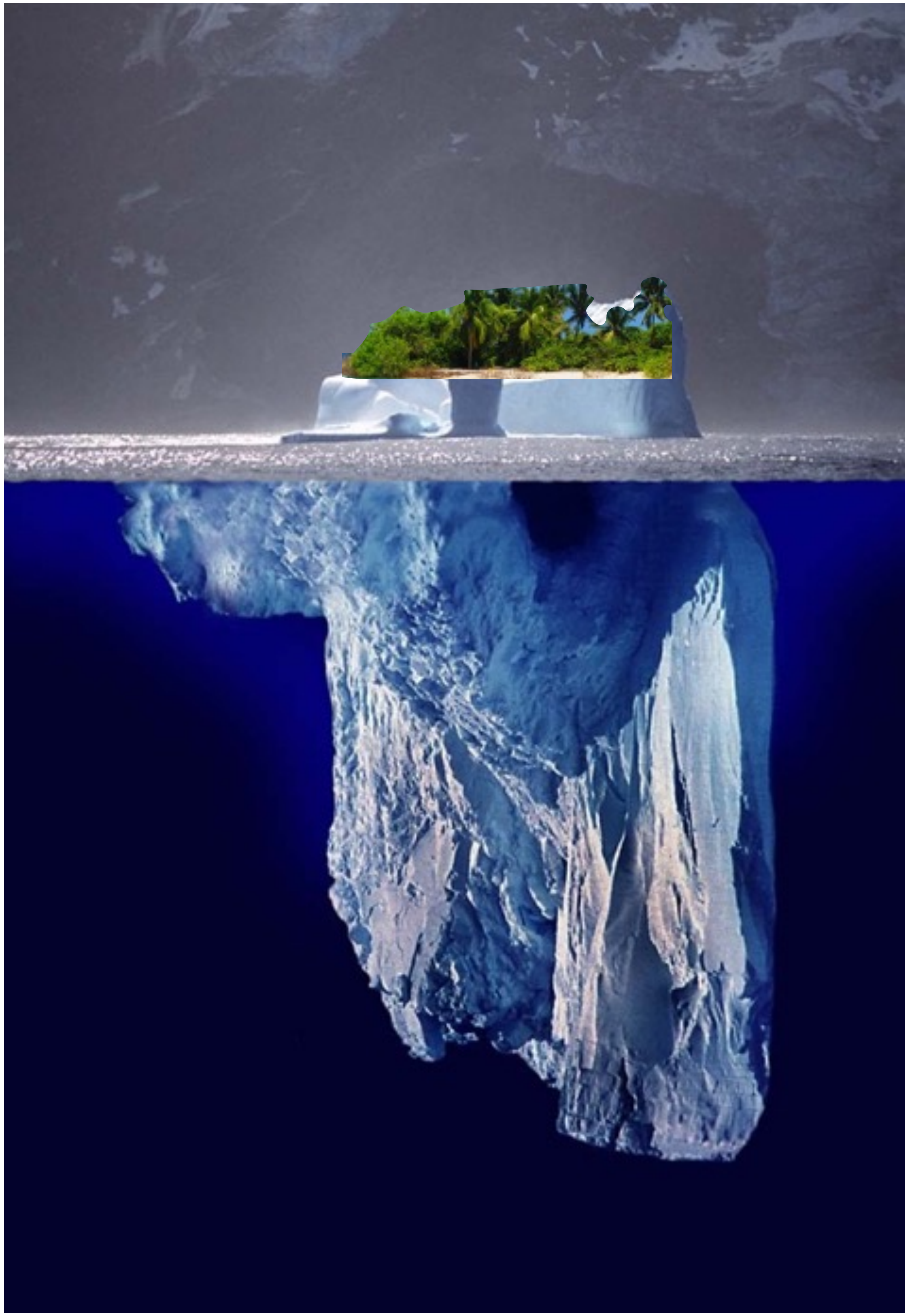
Epistasis measure gamma$_d$ with increasing Hamming distance d

SUMMARY/CONCLUSION

On average, our intragenic fitness landscape looks rugged and negative epistasis is common.

The global peak is accessible and reached via a highly synergistic combination of four mutations.

However, when evolving from parental type, adaptation may stall at a local peak.



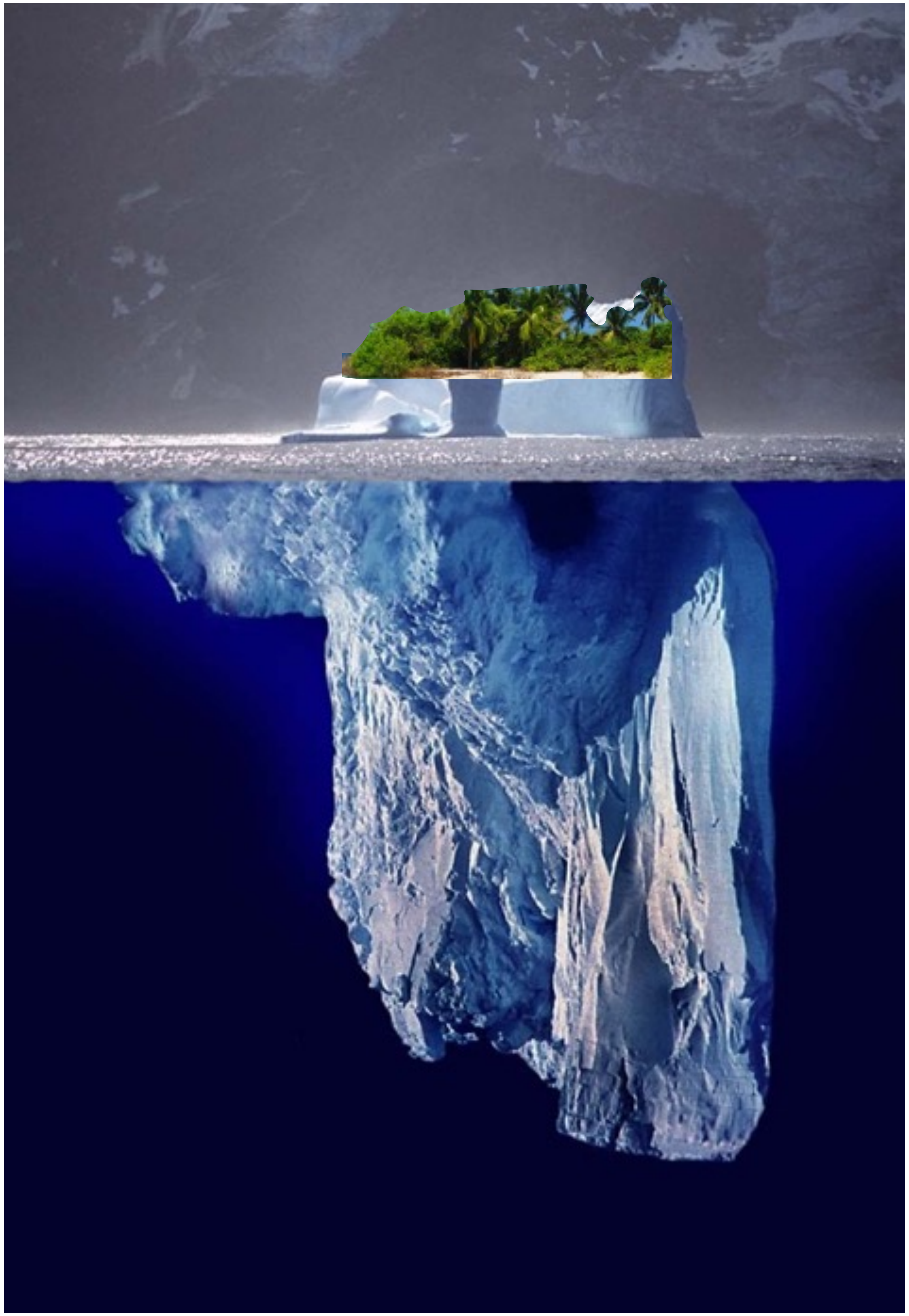
SUMMARY/CONCLUSION

On average, our intragenic fitness landscape looks rugged and negative epistasis is common.

The global peak is accessible and reached via a highly synergistic combination of four mutations.

However, when evolving from parental type, adaptation may stall at a local peak.

So far, limited predictive potential, but lots of ideas for the future…



ACKNOWLEDGEMENTS

Evolutionary Dynamics Group
Inês Fragata
Ana-Hermina Ghenu
Andreia Teixeira

www.evoldynamics.org

École Polytechnique Fédérale de Lausanne
Jeff Jensen
Sebastian Matuszewski
The Jensen lab

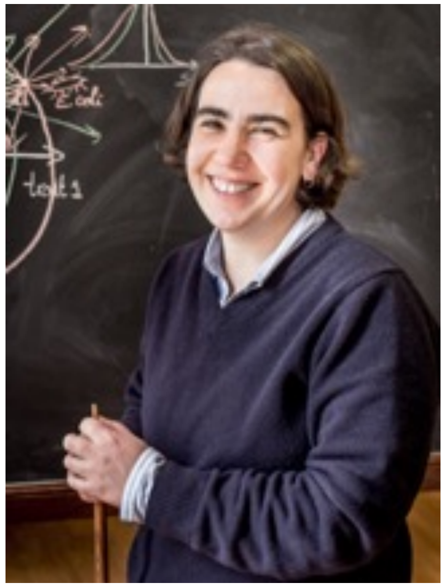
University of Massachusetts Medical School
Dan Bolon, Ryan Hietpas, Pamela Cote

INSTITUTO
GULBENKIAN
DE CIÊNCIA

FITNESS LANDSCAPES & ANTIBIOTIC RESISTANCE

- Costs of resistance, yet often no elimination of resistant types upon drug withdrawal
- What is the fitness landscape across environments? Are there refugia for resistant bacteria? How easily can resistant types invade novel environments?
- Is the architecture of resistance dependent on the bacterial strain? Can we predict resistance across genetic backgrounds?
- Are costs/routes to compensatory evolution predictable across genetic backgrounds and environments?

Alex Wong

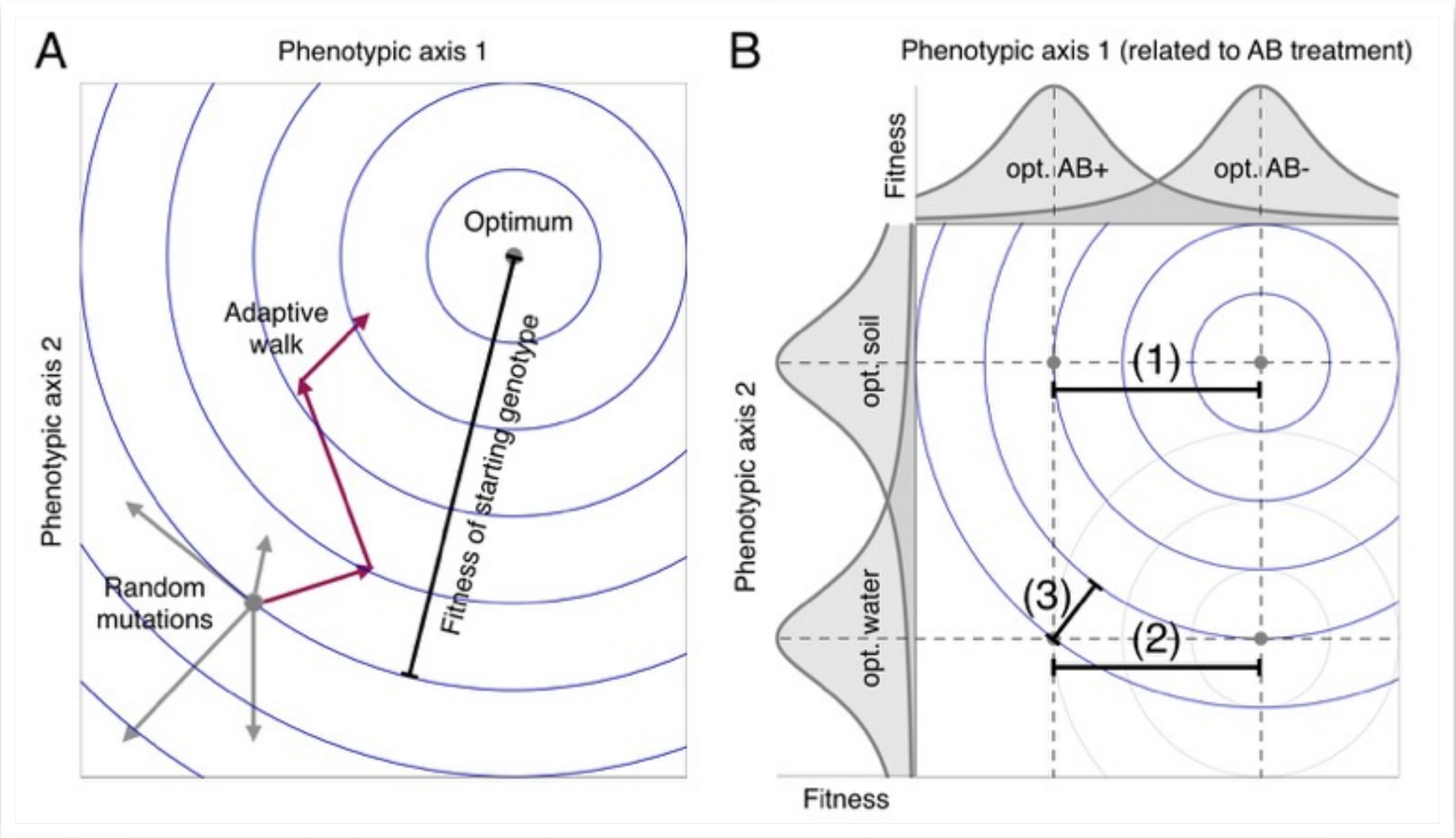


Isabel Gordo

Rees Kassen

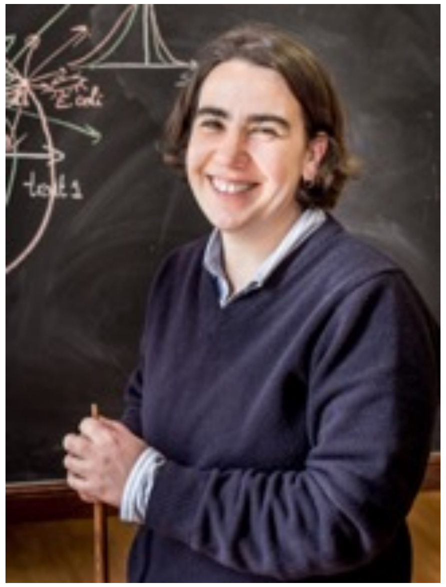
Thomas Bataillon

FITNESS LANDSCAPES & ANTIBIOTIC RESISTANCE



- Are costs/routes to compensatory evolution predictable across genetic backgrounds and environments?

Alex Wong

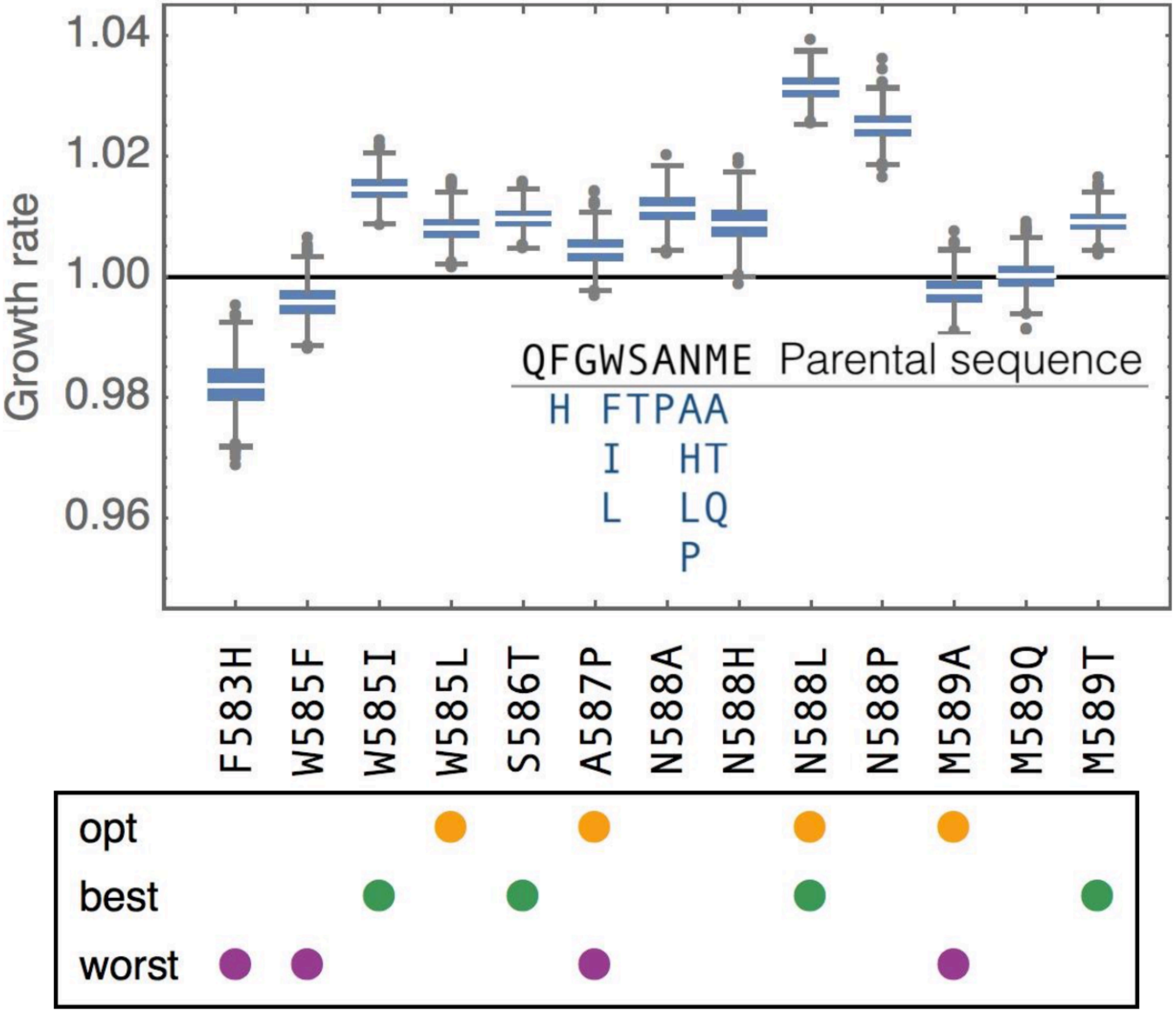


Isabel Gordo

Rees Kassen

Thomas Bataillon

high salinity environment
13 single-aa mutations
2 replicates
all possible combinations of aa's
≈1600 nt mutations



	F583H	W585F	W585I	W585L	S586T	A587P	N588A	N588H	N588L	N588P	M589A	M589Q	M589T
opt				●		●			●		●		
best			●		●				●				●
worst	●	●				●					●		