

# FITNESS LANDSCAPES AND THE PREDICTABILITY OF EVOLUTION

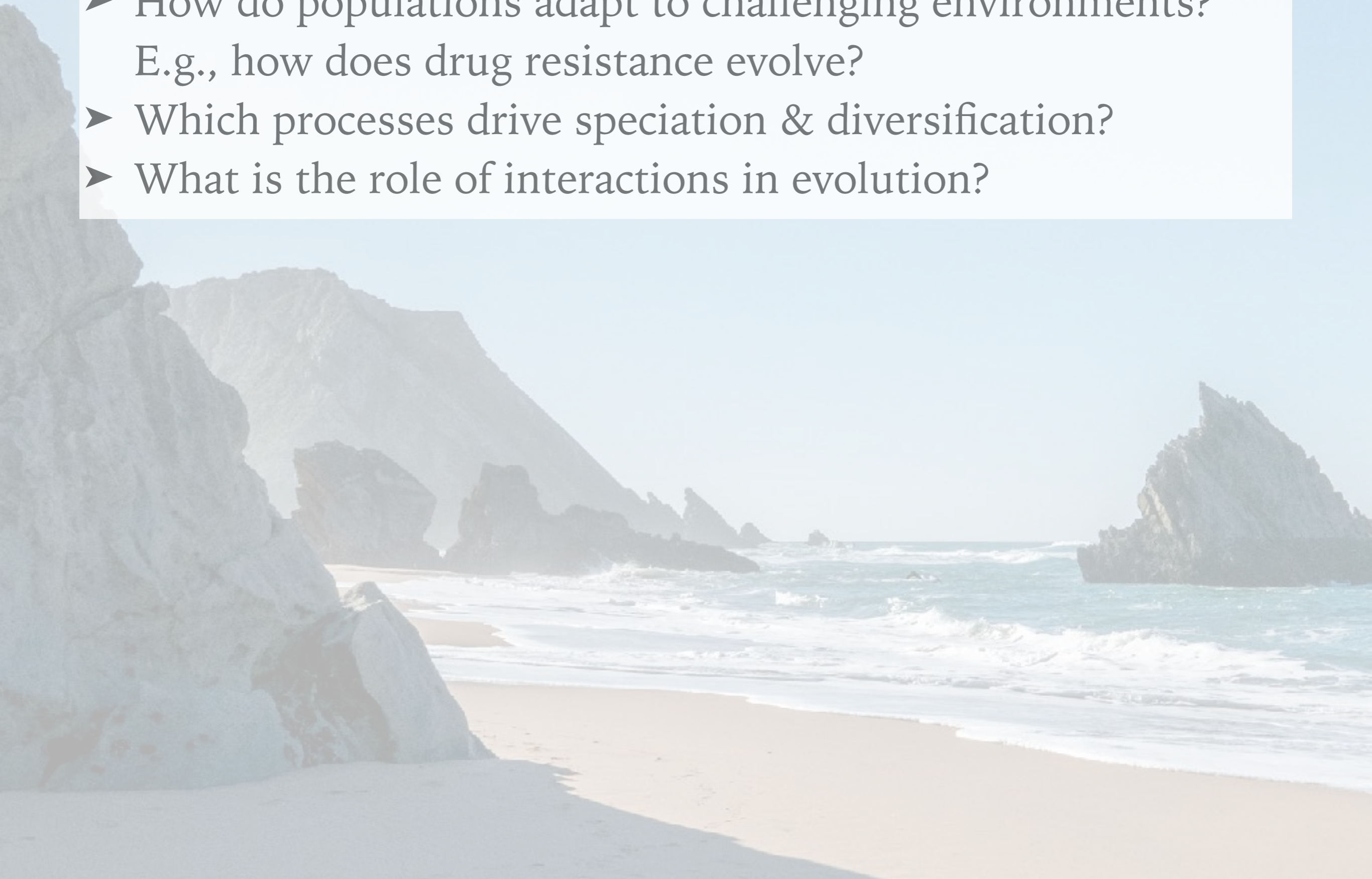
CLAUDIA BANK  
EVOLUTIONARY DYNAMICS GROUP



INSTITUTO  
GULBENKIAN  
DE CIÊNCIA

## Evolutionary Dynamics @ IGC:

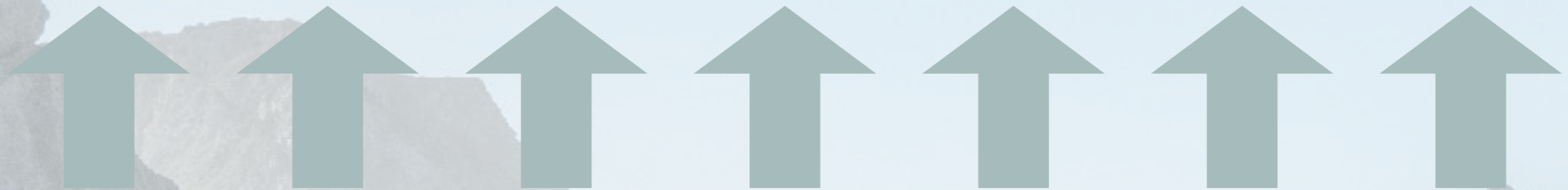
- How do populations adapt to challenging environments?  
E.g., how does drug resistance evolve?
- Which processes drive speciation & diversification?
- What is the role of interactions in evolution?





## Evolutionary Dynamics @ IGC:

- How do populations adapt to challenging environments?  
E.g., how does drug resistance evolve?
- Which processes drive speciation & diversification?
- What is the role of interactions in evolution?



*Mutation*

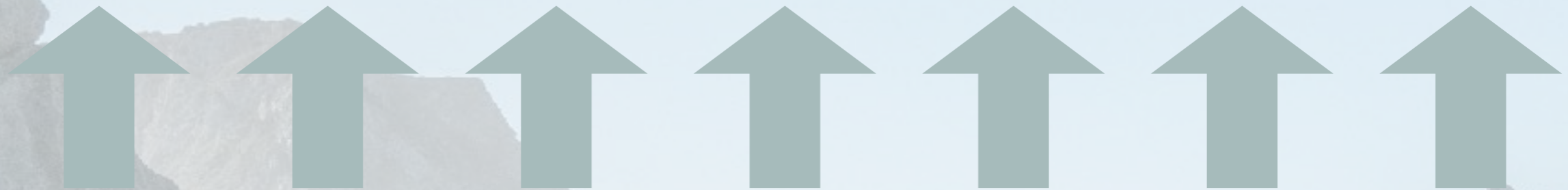
*Genetic drift*

*Migration*

*Selection*

## Evolutionary Dynamics @ IGC:

- How do populations adapt to challenging environments?  
E.g., how does drug resistance evolve?
- Which processes drive speciation & diversification?
- What is the role of interactions in evolution?



*Mutation*

*Genetic drift*

*Migration*

*Selection*

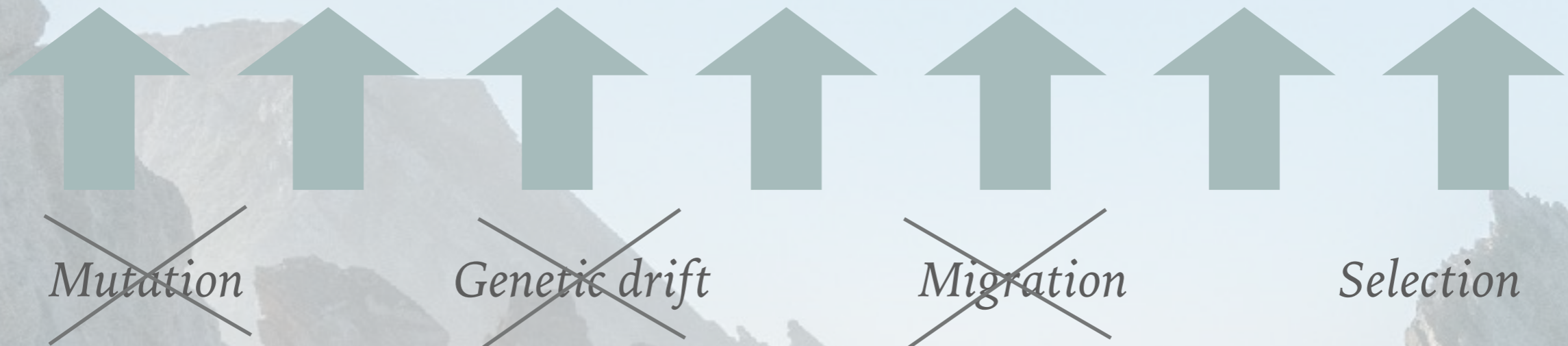
## What we do

- Study evolutionary processes using simple models
- Evaluate these models using empirical and simulated data
- Use modeling to inform experimental design *a priori*



## Evolutionary Dynamics @ IGC:

- How do populations adapt to challenging environments?  
E.g., how does drug resistance evolve?
- Which processes drive speciation & diversification?
- What is the role of interactions in evolution?



## What we do

- Study evolutionary processes using simple models
- Evaluate these models using empirical and simulated data
- Use modeling to inform experimental design *a priori*

# 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?

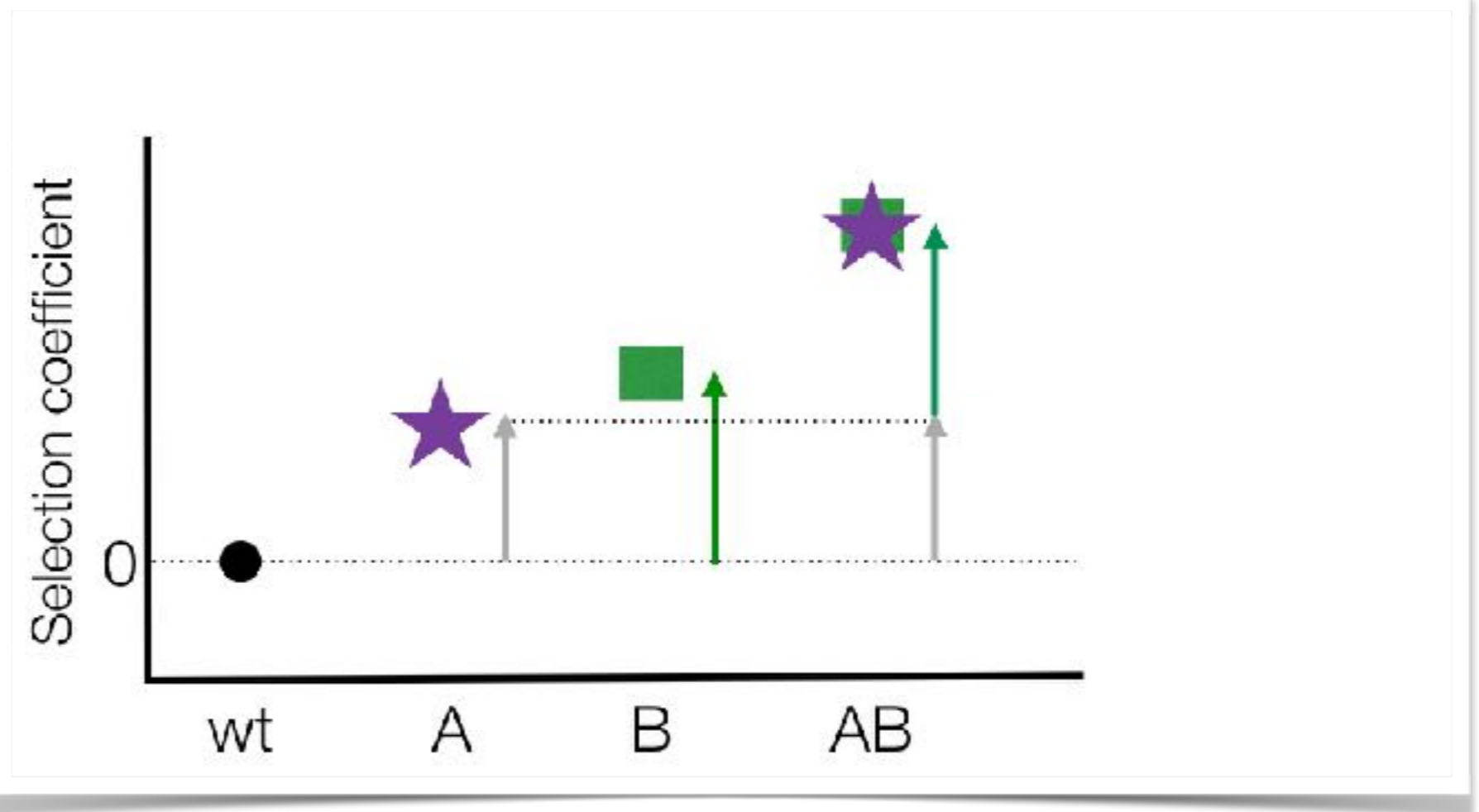
# HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?

---

- 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 the effects?
- What is the role of the environment?
- What are the probabilities of deleterious mutations?
- How do mutations affect fitness in a challenging environment?

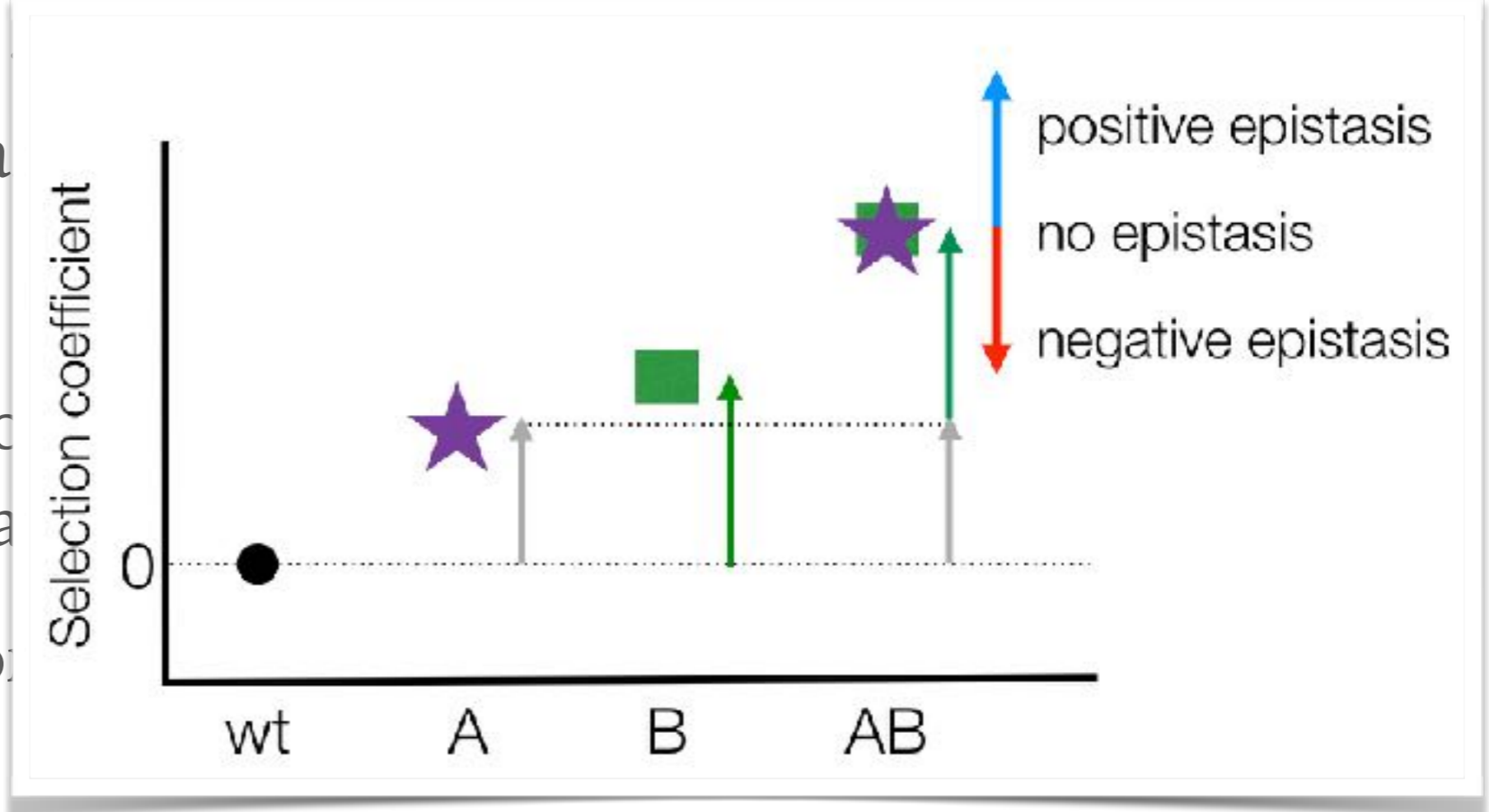


- How do mutational effects change dependent on the genetic background? (I.e., what is the role of epistasis?)



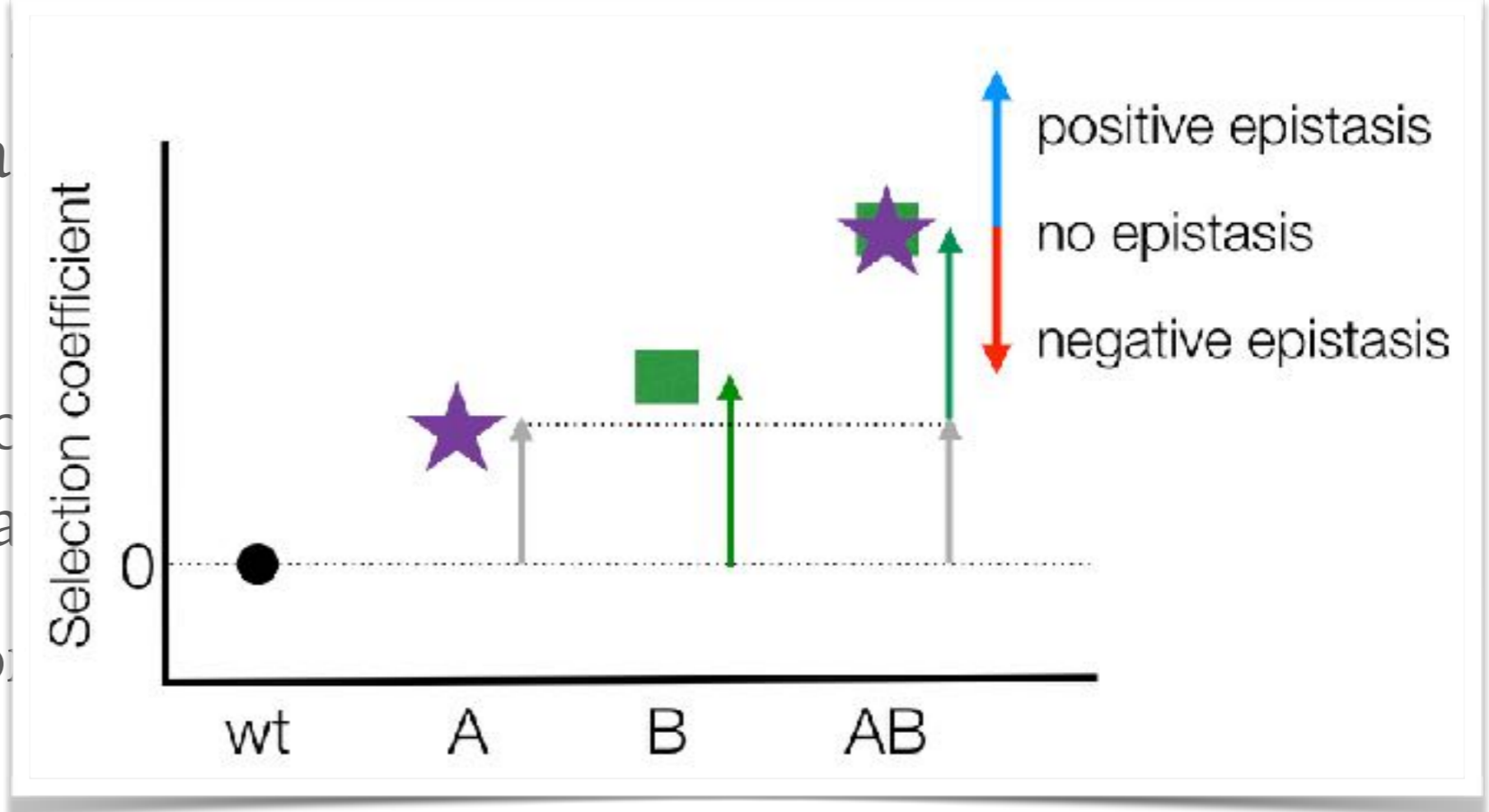
# HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?

- How big/small are the effects of mutations?
- What is the role of deleterious mutations?
- What are the probabilities of deleterious mutations occurring in a challenging environment?
- How do mutational effects change dependent on the genetic background? (I.e., what is the role of epistasis?)



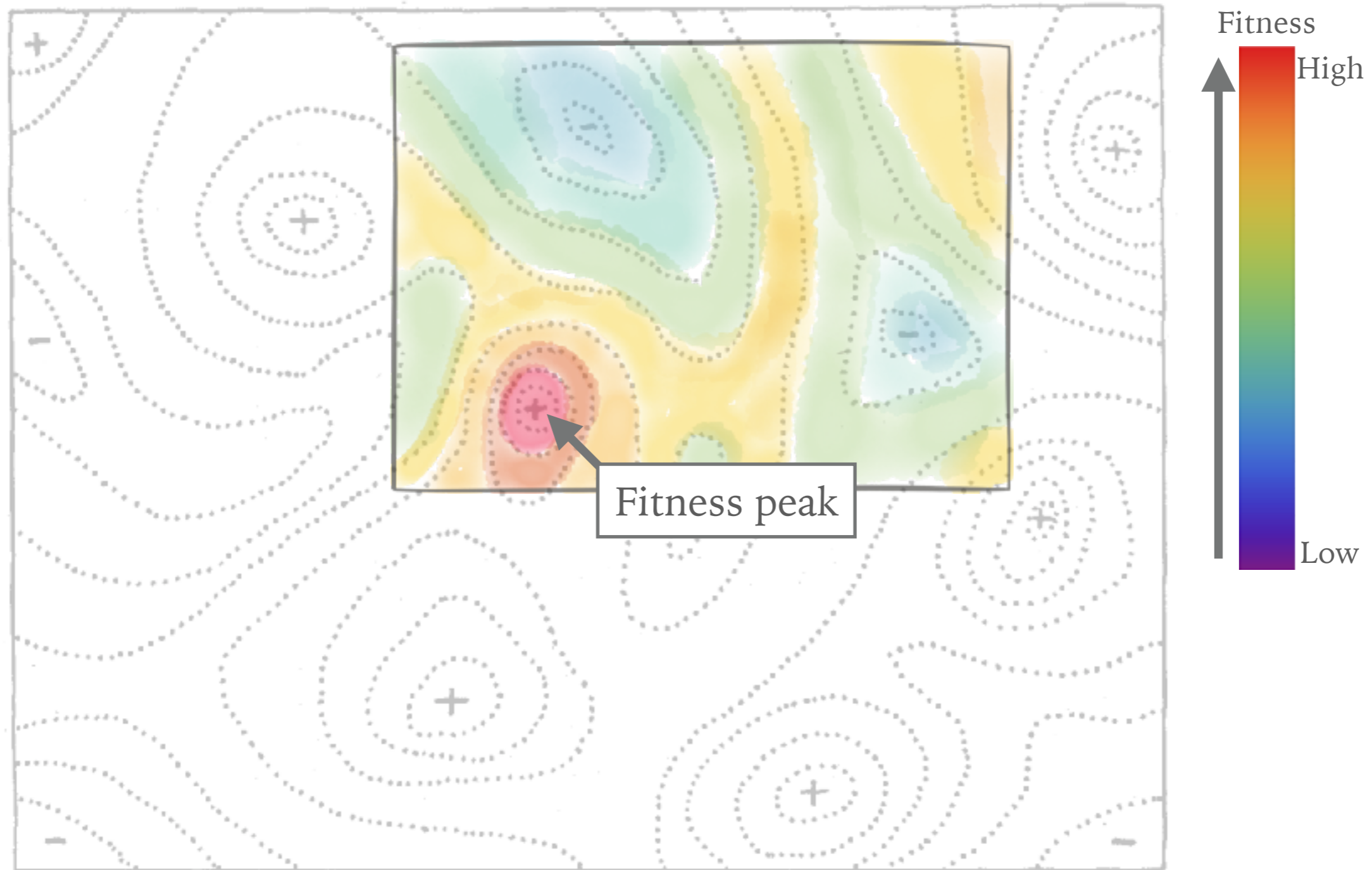
# HOW DO ORGANISMS ADAPT TO CHALLENGING ENVIRONMENTS?

- How big/small are the challenges?
- What is the role of the environment?
- What are the probabilities of deleterious mutations?
- How do mutations affect 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?



# UNDERSTANDING EVOLUTION THROUGH FITNESS LANDSCAPES

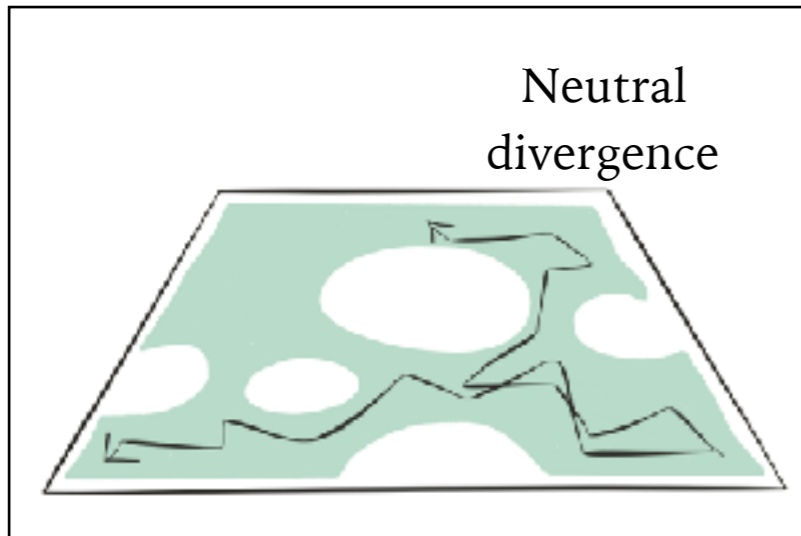
---



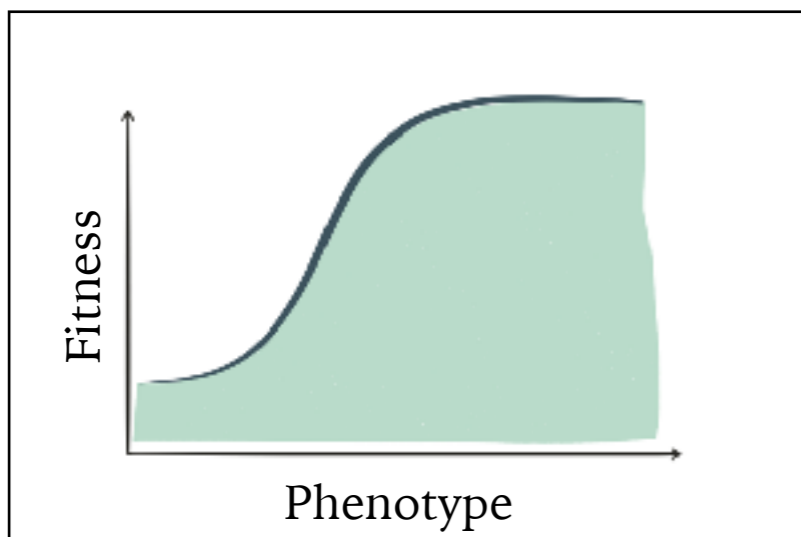
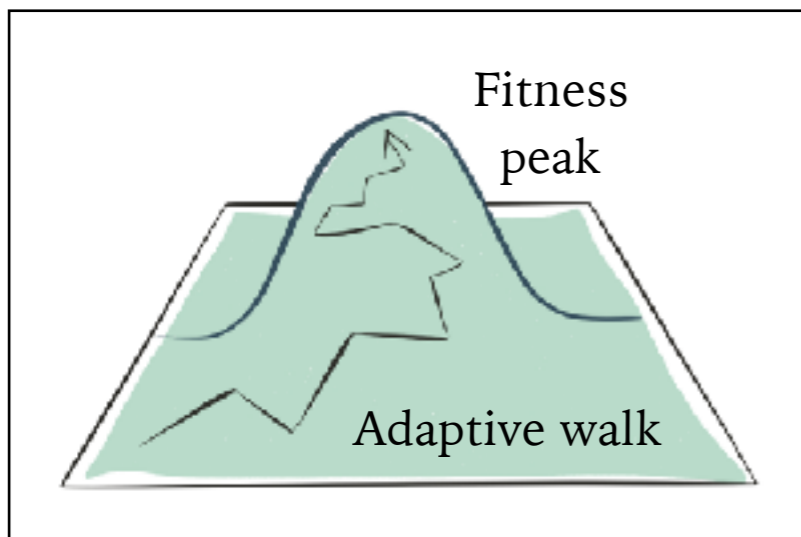


# WHY FITNESS LANDSCAPES ARE APPEALING

---

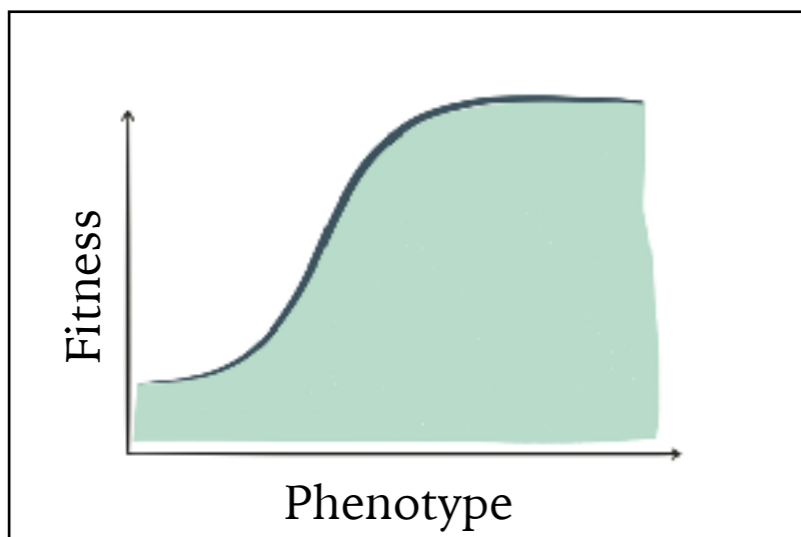
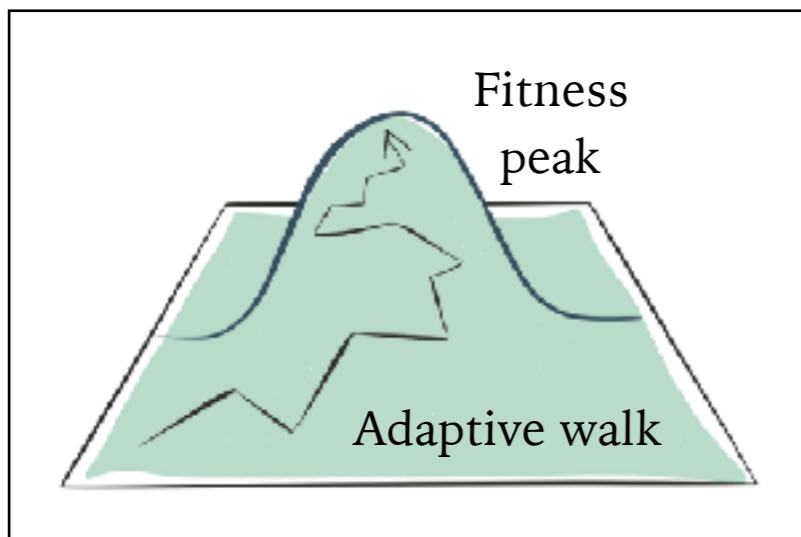
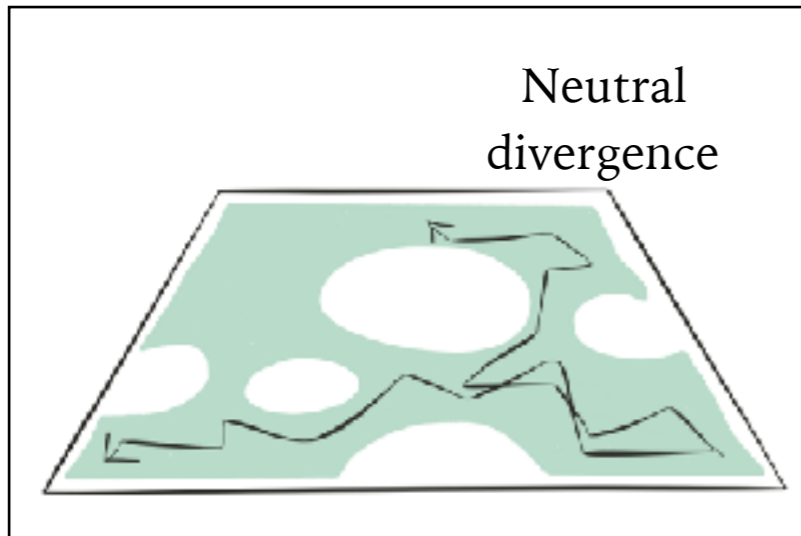


- fitness landscapes yield information on the predictability and repeatability of evolution



# WHY FITNESS LANDSCAPES ARE APPEALING

---



- 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<sup>1,2,3,4,5\*</sup>, Dmitry A. Bolotin<sup>1,3\*</sup>, Margarita V. Meer<sup>4,5</sup>, Dinara R. Usmanova<sup>4,5,6</sup>, Alexander S. Mishin<sup>1,2</sup>, George M. Skafarou<sup>7</sup>, Andrey V. Frankov<sup>4,8</sup>, Elena G. Ivanova<sup>4,9</sup>, Mikhail S. Kassov<sup>4,9</sup>, Konstantin A. Lukyanov<sup>1,2</sup>, Natalya S. Bogatyreva<sup>4,5,8</sup>, Peter K. Vlasov<sup>4,5</sup>, Evgeny S. Egorov<sup>1</sup>, Maria D. Logacheva<sup>9,10,11</sup>, Alexey S. Kondrashov<sup>11,12</sup>, Dmitry M. Shvachkin<sup>1,2</sup>, Alexander P. Putintsev<sup>1,3</sup>, Igor E. Markevich<sup>1,3</sup>, Dan S. Rayfield<sup>13</sup>, Konstantin A. Lukyanov<sup>1,2</sup> & Pyodor A. Kondrashov<sup>4,5,14</sup>

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

José I. Jiménez<sup>a,1</sup>, Ramon Xulvi Brunet<sup>a,1</sup>, Gregory W. Campbell<sup>b</sup>, Rebecca Turk-MacLeod<sup>a</sup>, and Irene A. Chen<sup>a,b,2</sup>

# Biophysical principles predict fitness landscapes of drug resistance

Ioão V. Rodrigues<sup>a</sup>, Shimon Bershtein<sup>b</sup>, Annabel Li<sup>c,1</sup>, Elena R. Lozovsky<sup>c</sup>, Daniel L. Hartl<sup>c,2</sup>, and Eugene I. Shakhovich<sup>a,2</sup>

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

Ashley Acevedo<sup>1</sup>, Leonid Brodsky<sup>2</sup> & Raul Andino<sup>1</sup>

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

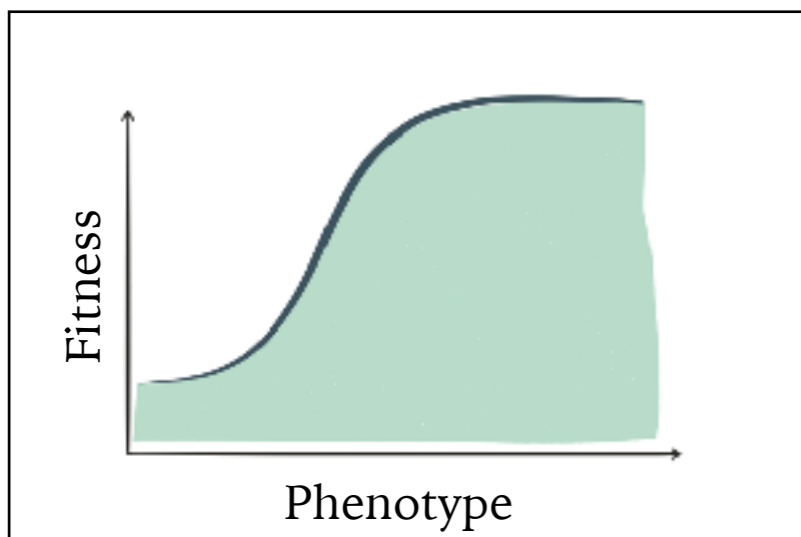
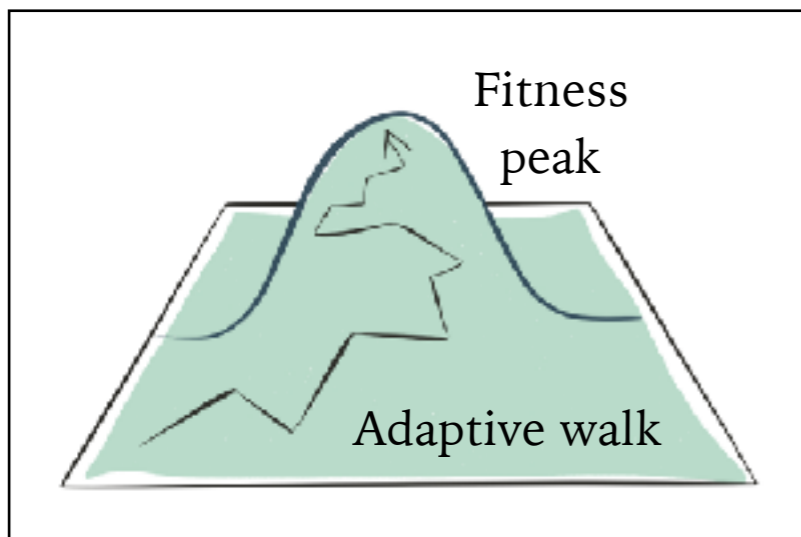
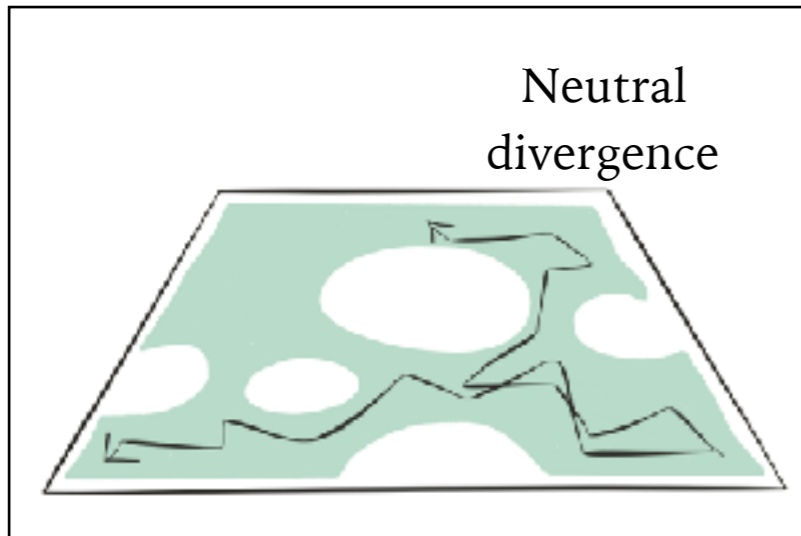
Fabio Zanini<sup>1,2</sup>, Vedim Buller<sup>1</sup>, Johanna Bredin<sup>3</sup>, Jan Albert<sup>4</sup>, Richard A. Noher<sup>1\*</sup>

# The fitness landscape of a tRNA gene



# WHY FITNESS LANDSCAPES ARE DIFFICULT

---

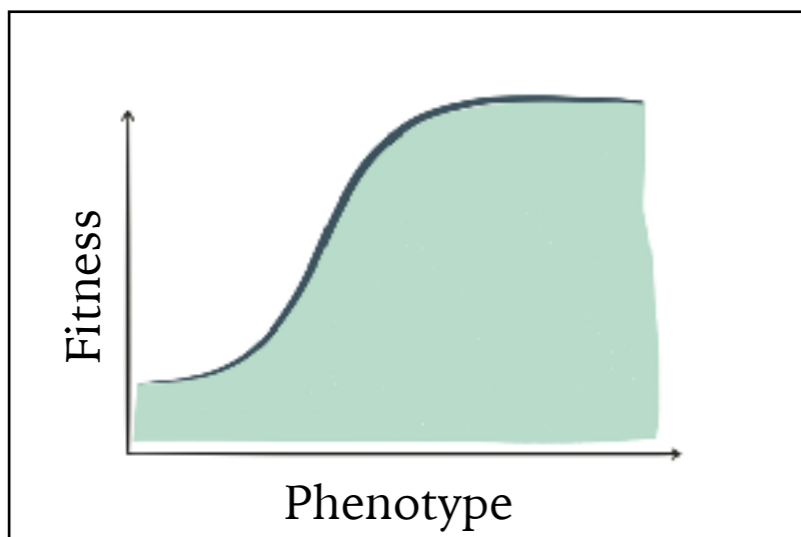
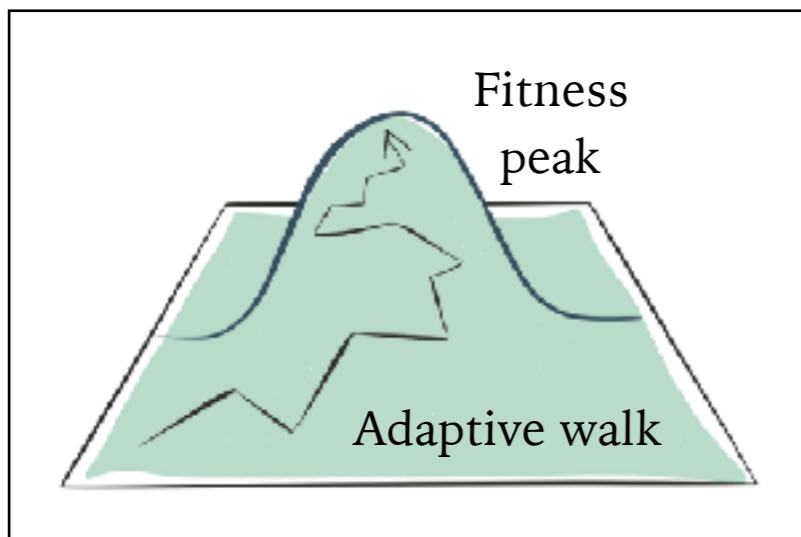
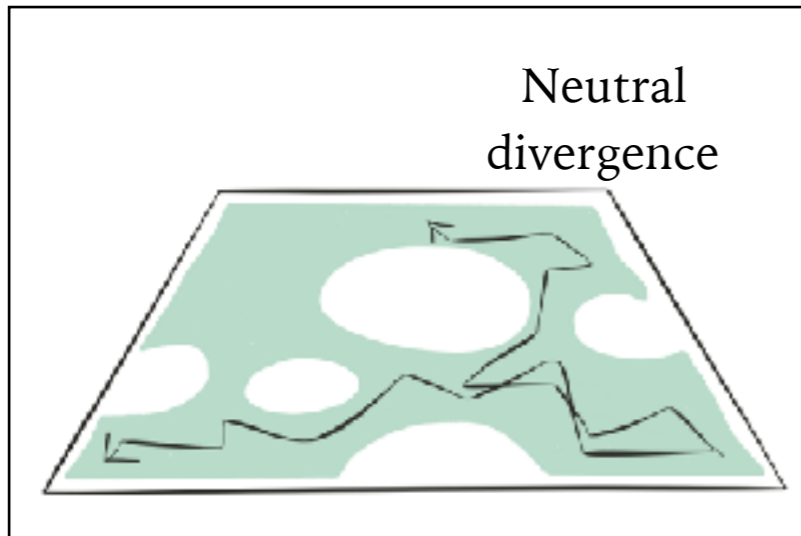


- fitness landscapes yield information on the predictability and repeatability of evolution
- it becomes increasingly simple to measure empirical fitness landscapes
- accumulating data on gene networks and pathways

But:

# WHY FITNESS LANDSCAPES ARE DIFFICULT

---



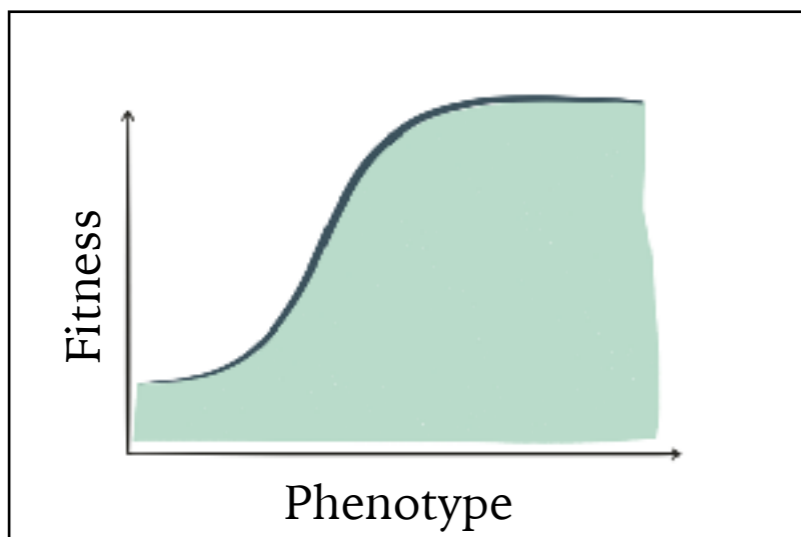
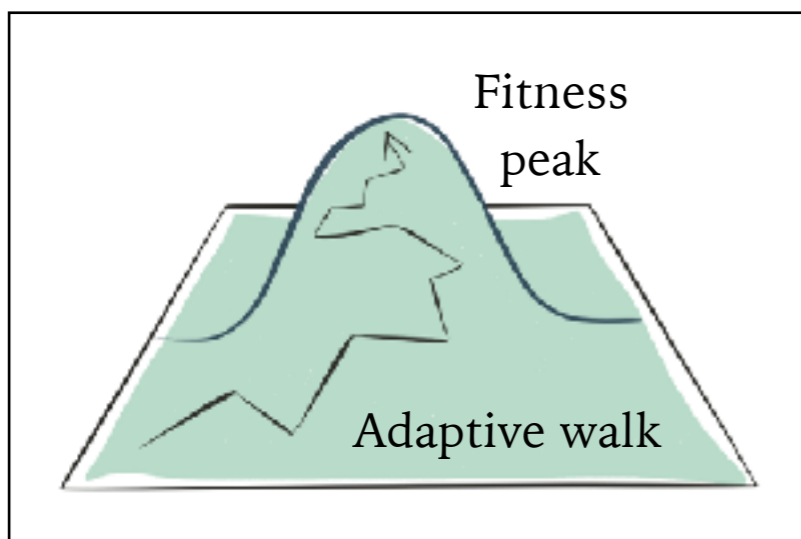
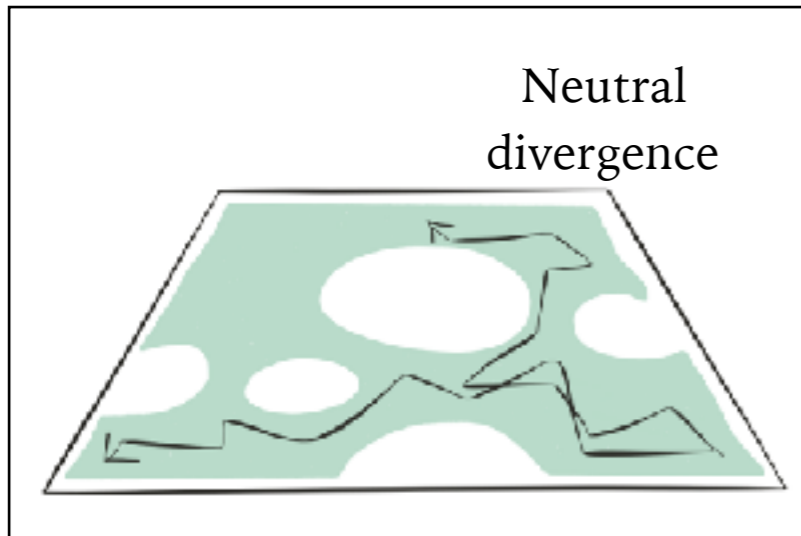
- fitness landscapes yield information on the predictability and repeatability of evolution
- it becomes increasingly simple to measure empirical fitness landscapes

- accumulating data on gene networks and pathways

But:

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

# WHY FITNESS LANDSCAPES ARE DIFFICULT



- 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:

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

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

# PROJECTS IN THE LAB

---

- Fitness landscapes in **adaptation**
  - How can we measure fitness landscapes?
  - Is evolution on a fitness landscape predictable?



# PROJECTS IN THE LAB

---

- Fitness landscapes in **adaptation**
  - How can we measure fitness landscapes?
  - Is evolution on a fitness landscape predictable?
- Fitness landscapes in **speciation**
  - What is the role of genetic incompatibilities in speciation and hybridization?
  - What are the dynamics and the genomic signature of epistatic interactions?

# PROJECTS IN THE LAB

---

- Fitness landscapes in **adaptation**
  - How can we measure fitness landscapes?
  - Is evolution on a fitness landscape predictable?
- Fitness landscapes in **speciation**
  - What is the role of genetic incompatibilities in speciation and hybridization?
  - What are the dynamics and the genomic signature of epistatic interactions?
- Fitness landscapes in **disease**
  - Which forces are responsible for the distribution of centriole numbers in cancer?
  - Can we use fitness landscapes to predict the persistence of antibiotic resistance?
  - How can viruses escape mutagenic drugs?

# PROJECTS IN THE LAB

---

- **Fitness landscapes in adaptation**
  - How can we measure fitness landscapes?
  - Is evolution on a fitness landscape predictable?
- **Fitness landscapes in speciation**
  - What is the role of genetic incompatibilities in speciation and hybridization?
  - What are the dynamics and the genomic signature of epistatic interactions?
- **Fitness landscapes in disease**
  - Which forces are responsible for the distribution of centriole numbers in cancer?
  - Can we use fitness landscapes to predict the persistence of antibiotic resistance?
  - How can viruses escape mutagenic drugs?

# THE EXPERIMENTAL APPROACH: DEEP MUTATIONAL SCANNING

---

- Systematic high-throughput 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)



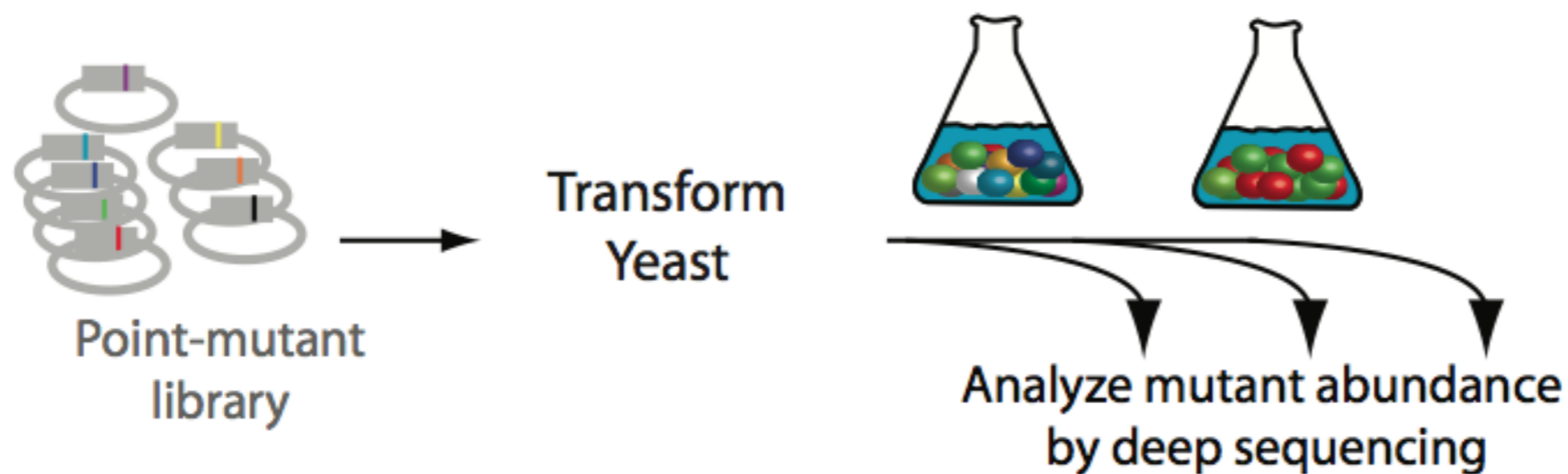
*Dan Bolon*



*Ryan Hietpas*



*Jeff Jensen*





# THE EXPERIMENTAL APPROACH: DEEP MUTATIONAL SCANNING

- Systematic high-throughput sampling of hundreds of chosen mutations (including those that are strongly deleterious)



Deep mutational scanning results in a (almost “evolution-free”) snapshot of the fitness landscape.

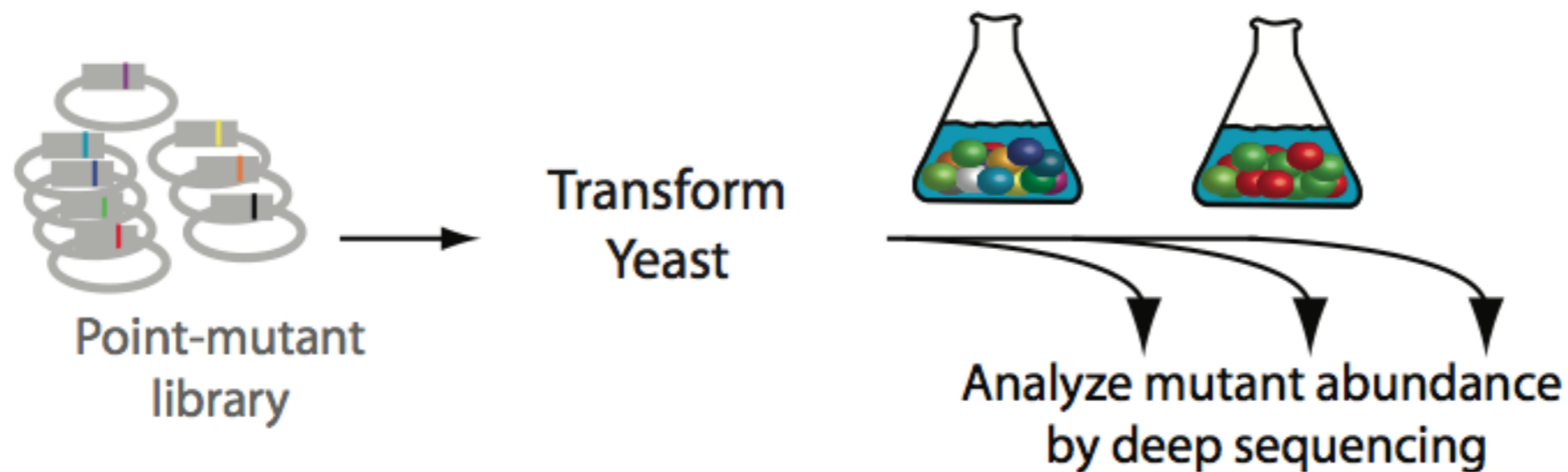
- Genetic background is precisely controlled (minimized potential for secondary mutations)



*Ryan Hietpas*



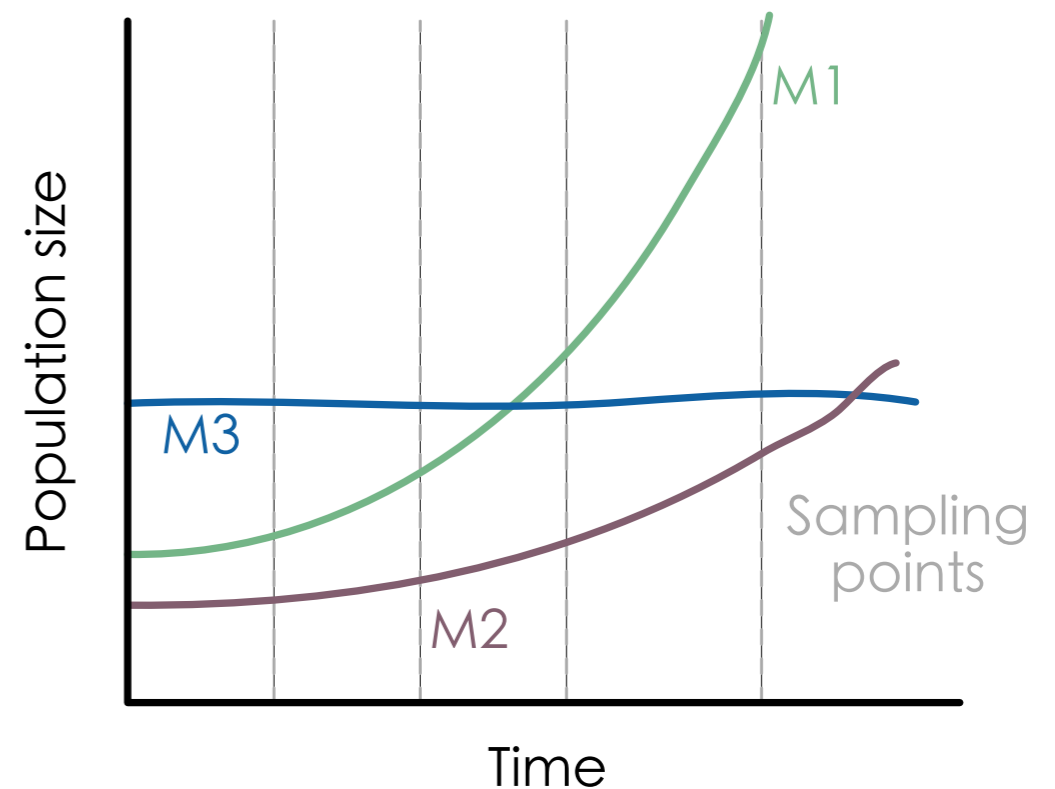
*Jeff Jensen*



# DEEP MUTATIONAL SCANNING FROM A MODELER'S POINT OF VIEW

---

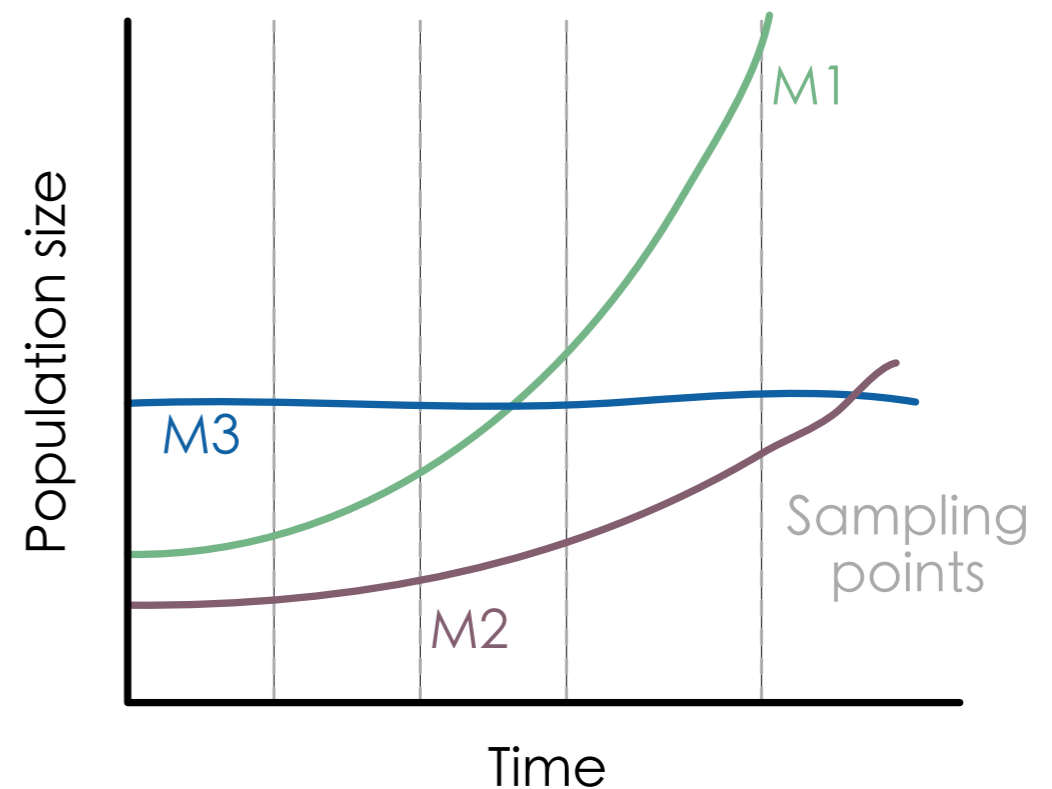
- Exponential growth of hundreds of mutants, each with its own growth rate/selection coefficient
- Sequencing corresponds to multinomial sampling of mutants independently at each sampling time



# DEEP MUTATIONAL SCANNING FROM A MODELER'S POINT OF VIEW

---

- Exponential growth of hundreds of mutants, each with its own growth rate/selection coefficient
- Sequencing corresponds to multinomial sampling of mutants independently at each sampling time



- What is the attainable experimental accuracy?
- How do experimental details (number and timing of samples, sequencing coverage, number of mutants, etc.) affect the outcome?

# A statistical guide to the design of deep mutational scanning experiments

Sebastian Matuszewski<sup>\*, †, §, §§</sup>, Marcel E. Hildebrandt<sup>\*, §</sup>, Ana-Hermina Ghenu<sup>‡</sup>, Jeffrey D. Jensen<sup>\*, †</sup> and Claudia Bank<sup>\*, †, ‡, §§, 1</sup>

<sup>\*</sup>School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, <sup>†</sup>Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland, <sup>‡</sup>Instituto Gulbenkian de Ciência, Oeiras, Portugal, <sup>§</sup>equal contribution, <sup>§§</sup>co-corresponding author

➤ optimize experimental setup for deep mutational scanning approaches

➤ statistical predictions

➤ applicable via interactive web tool:

[www.evoldynamics.org/tools](http://www.evoldynamics.org/tools)

[www.evoldynamics.org/tools](http://www.evoldynamics.org/tools)

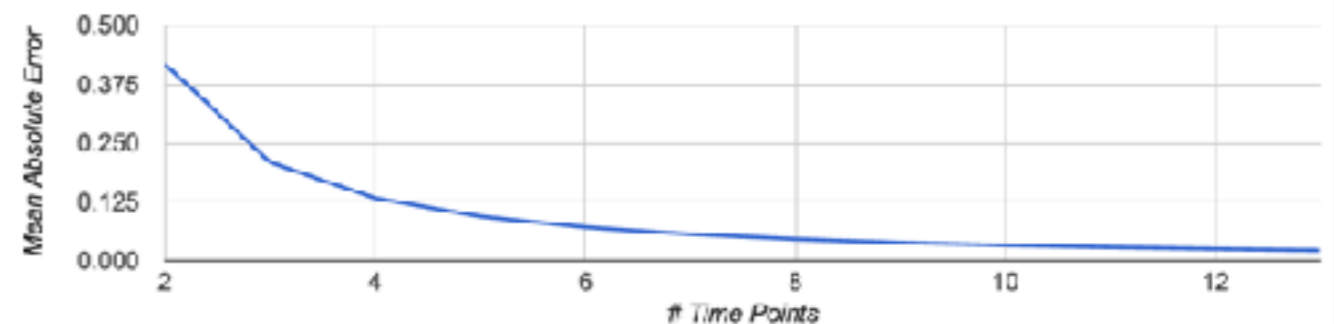


Hermina Ghenu

## Statistical equations for experimental design of high-throughput bulk competition experiments

This page allows you to improve the experimental design of your high-throughput bulk competition experiments by looking at the results from the equations of Matuszewski et al. 2016 for values that are relevant to your proposed experiment(s). The publication can be found here: [insert link].

Use the sliders to select the number of reads (S) and the number of mutants (K):



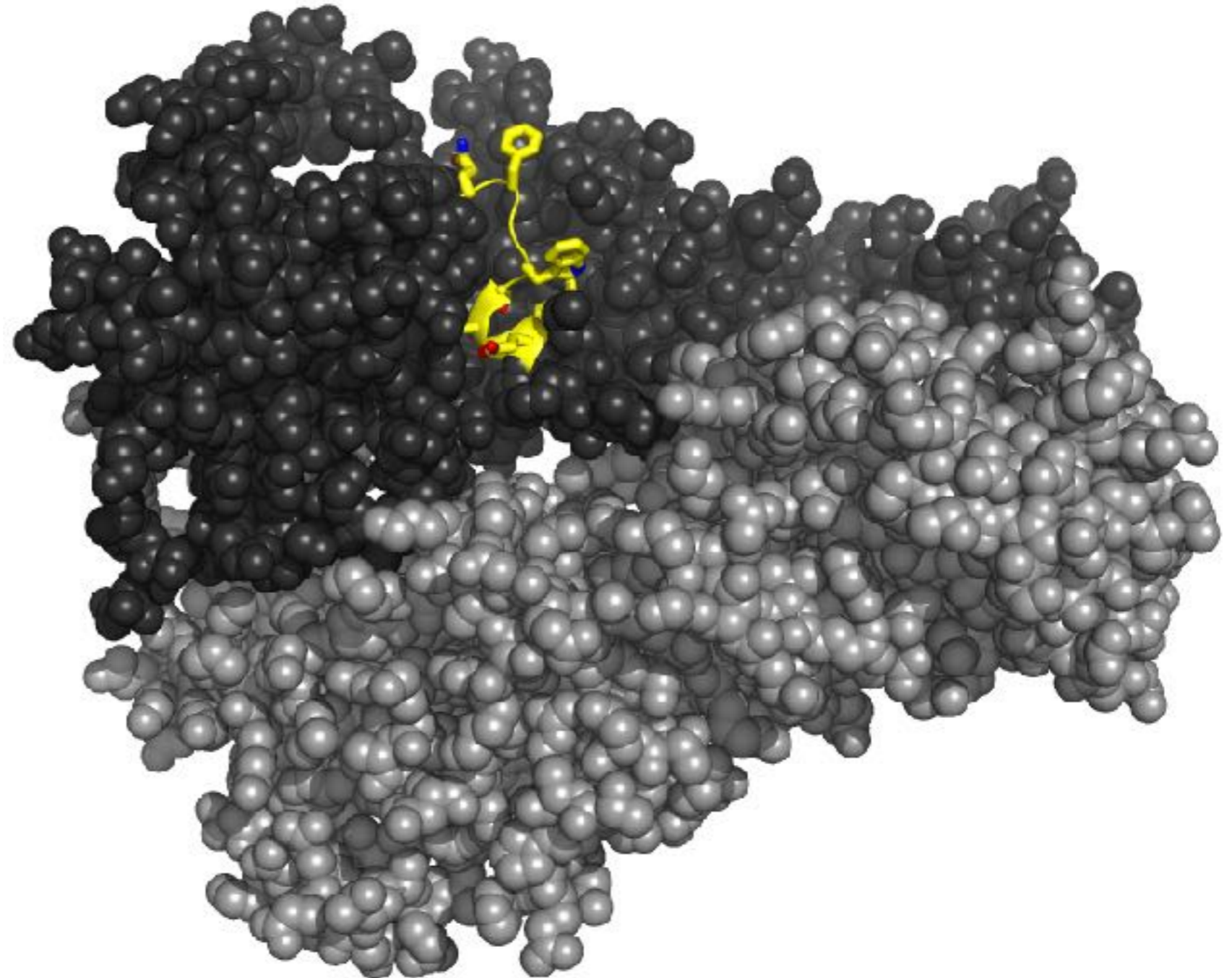
# Time Points	Mean Absolute Error	Mean Squared Error	$L_{0.95}$	$\tau(T)$
2	4.18336e-1	1.75005e-1	2.50593e+0	1
3	2.09168e-1	4.37512e-2	1.25296e+0	4
4	1.32289e-1	1.75005e-2	7.92443e-1	10
5	9.35427e-2	8.75024e-3	5.60342e-1	20
6	7.07116e-2	5.00014e-3	4.23579e-1	35
8	4.56442e-2	2.08339e-3	2.73419e-1	84
10	3.25674e-2	1.05063e-3	1.95086e-1	165



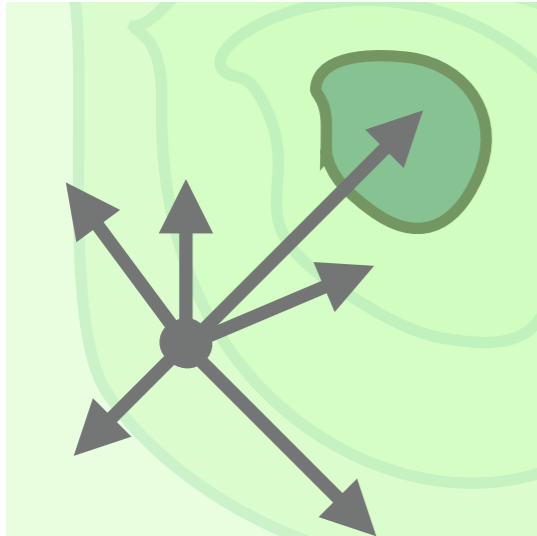
# WHAT WE DID SO FAR

---

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



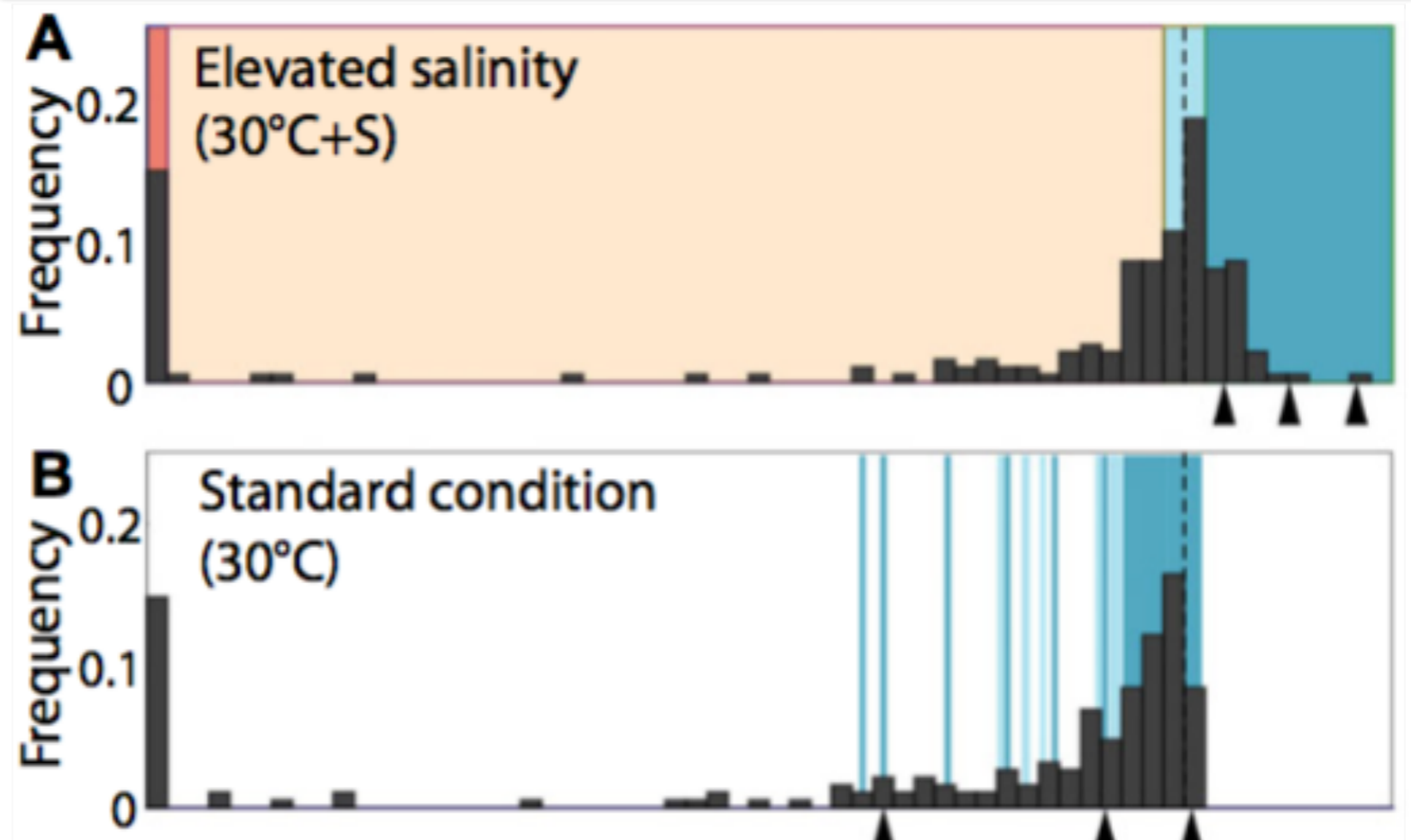
# WHAT WE DID SO FAR



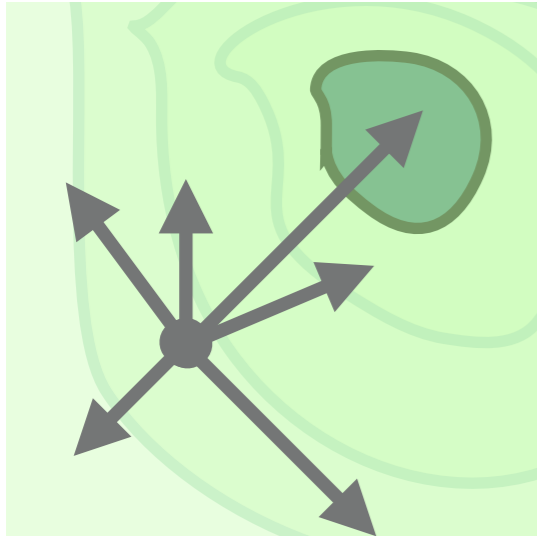
Engineered mutations from a 9 aa region from Hsp90 (aa positions 582-590) in *Saccharomyces cerevisiae*

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

Distribution of Fitness Effects (DFE) across environments



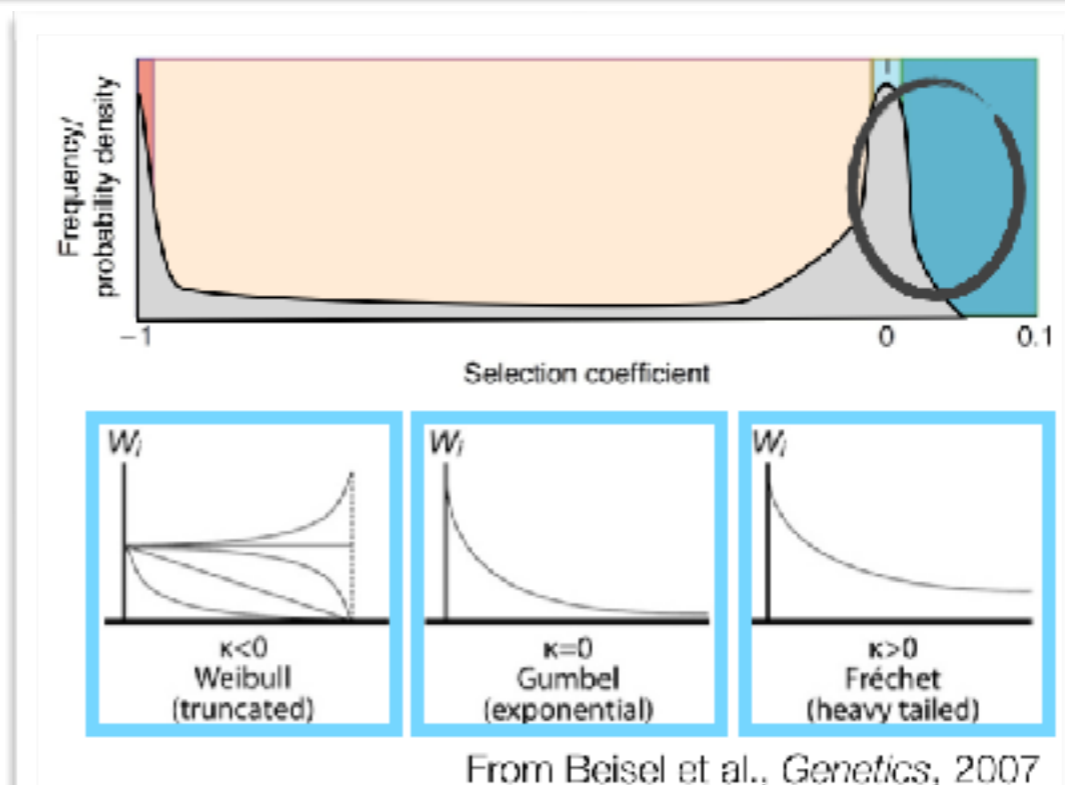
# WHAT WE DID SO FAR



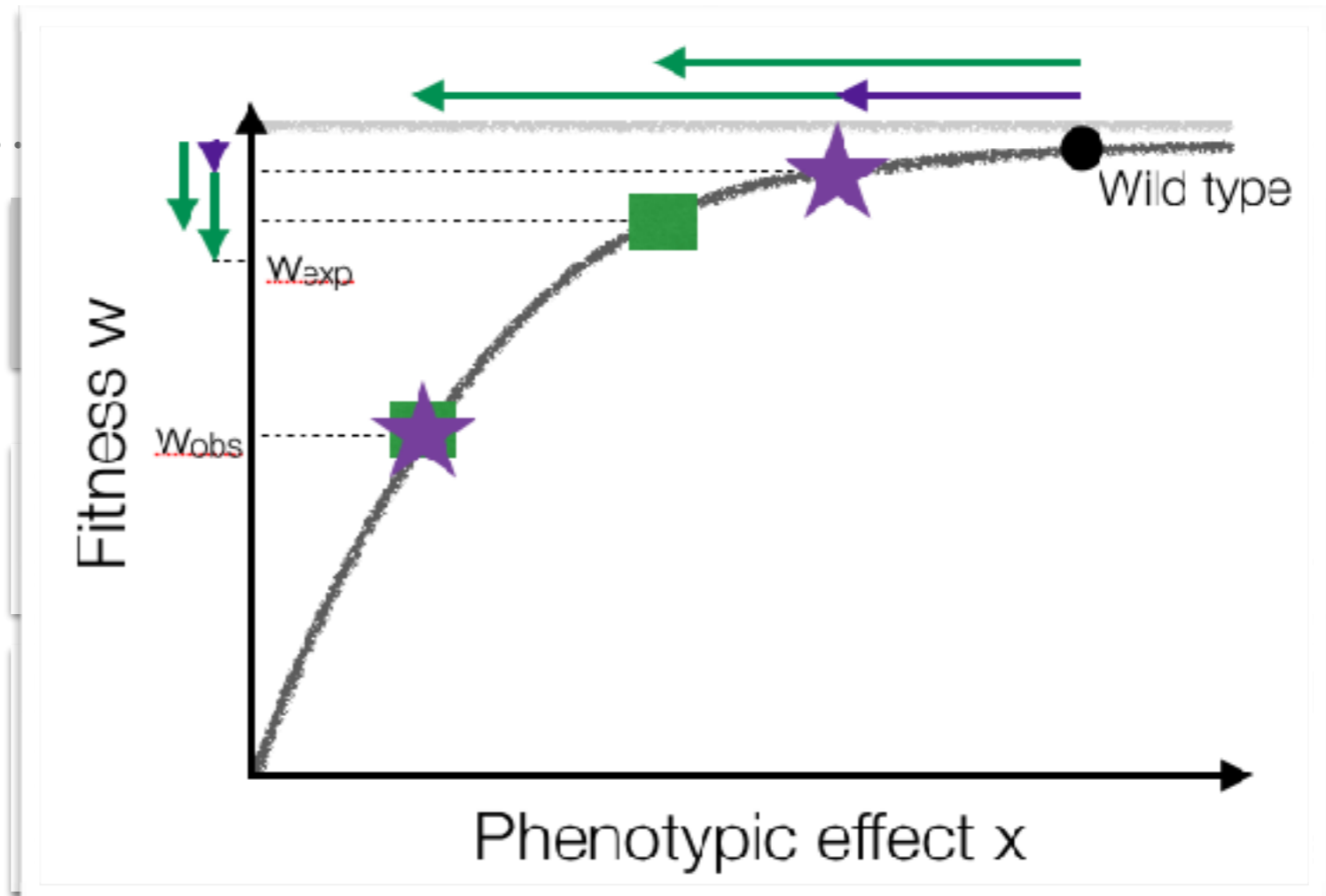
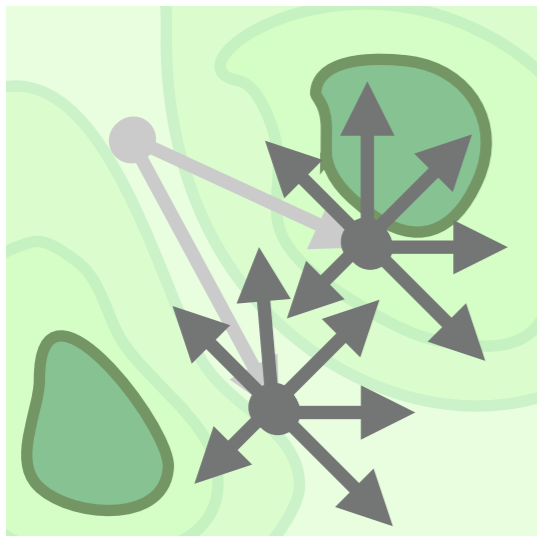
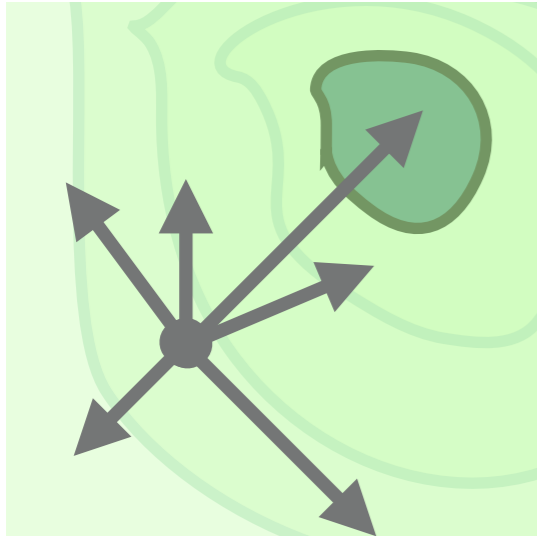
Engineered mutations from a 9 aa region from Hsp90 (aa positions 582-590) in *Saccharomyces cerevisiae*

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

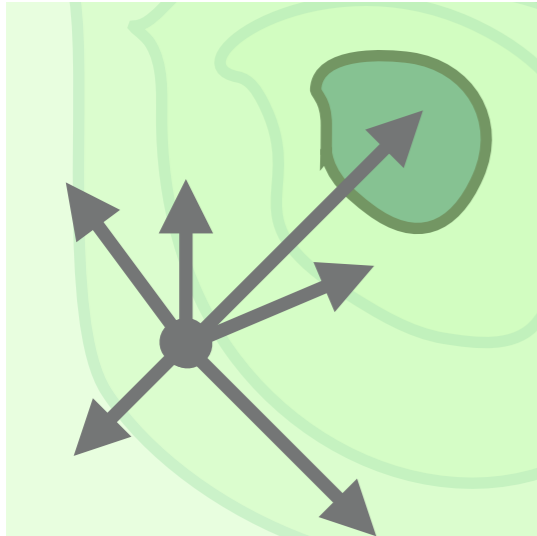


DFEs on 7 genetic backgrounds - ubiquitous negative epistasis indicating an underlying concave fitness landscape

Bank et al., 2015, MBE

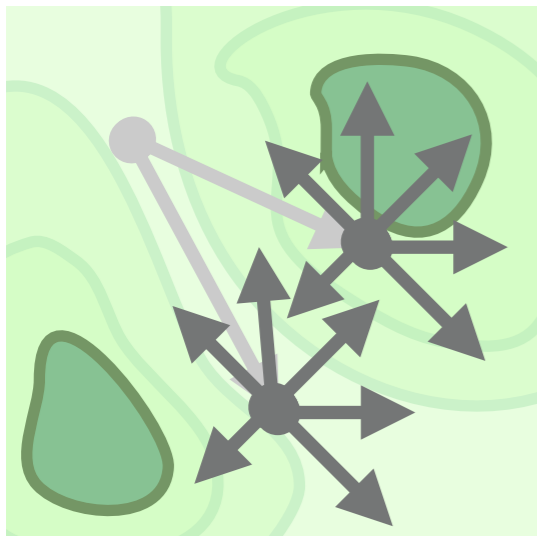
# WHAT WE DID SO FAR

---



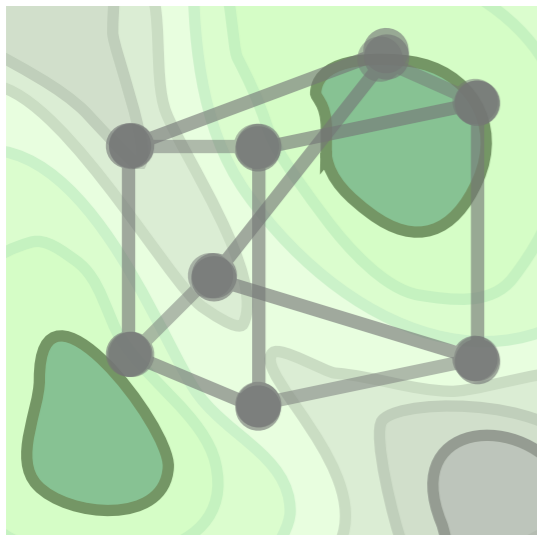
Engineered mutations from a 9 aa region from Hsp90 (aa positions 582-590) in *Saccharomyces cerevisiae*

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**

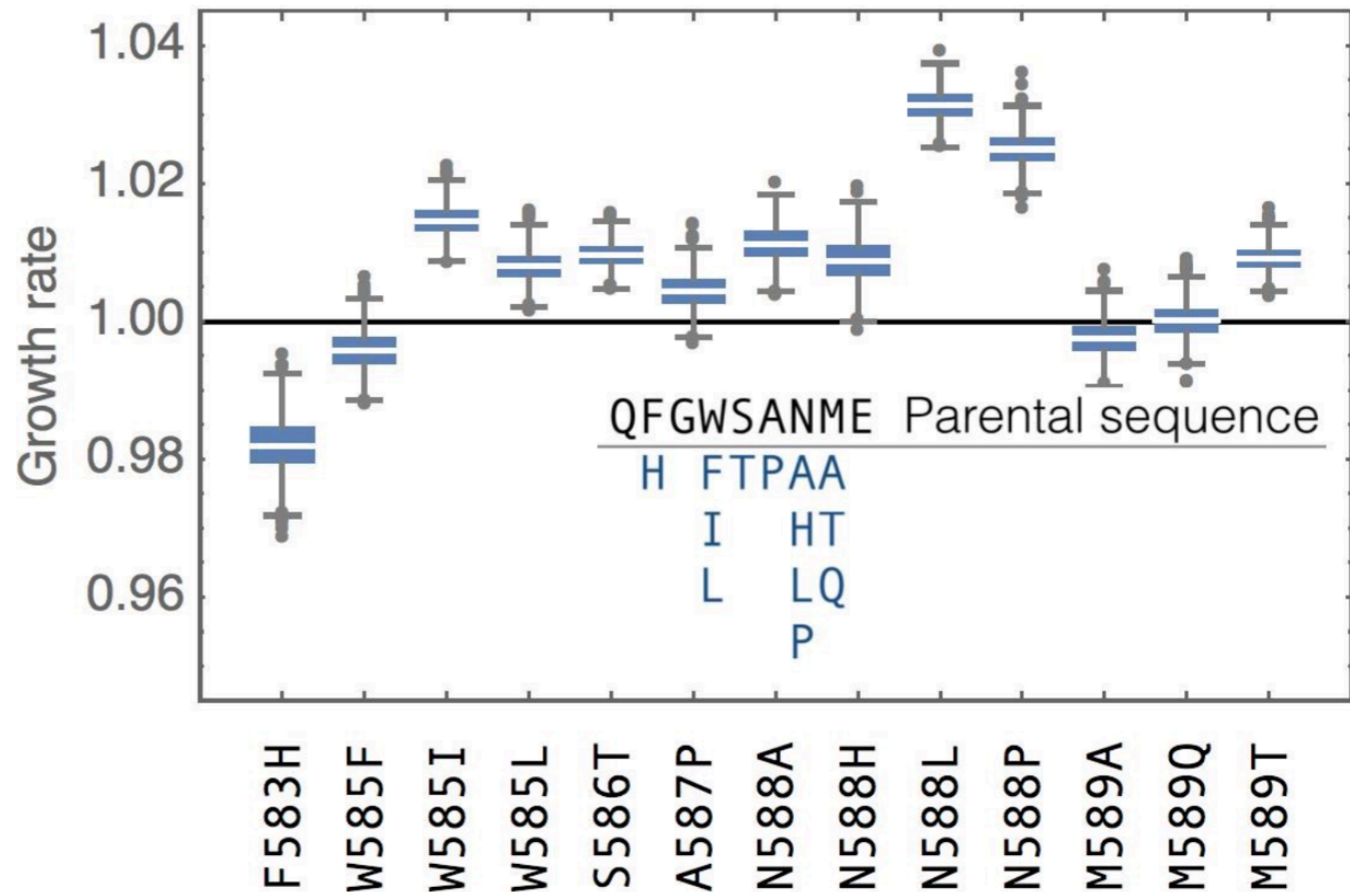


Guide to experimental design of deep mutational scanning studies  
**Matuszewski\*, Hildebrandt\* et al., 2016, Genetics**

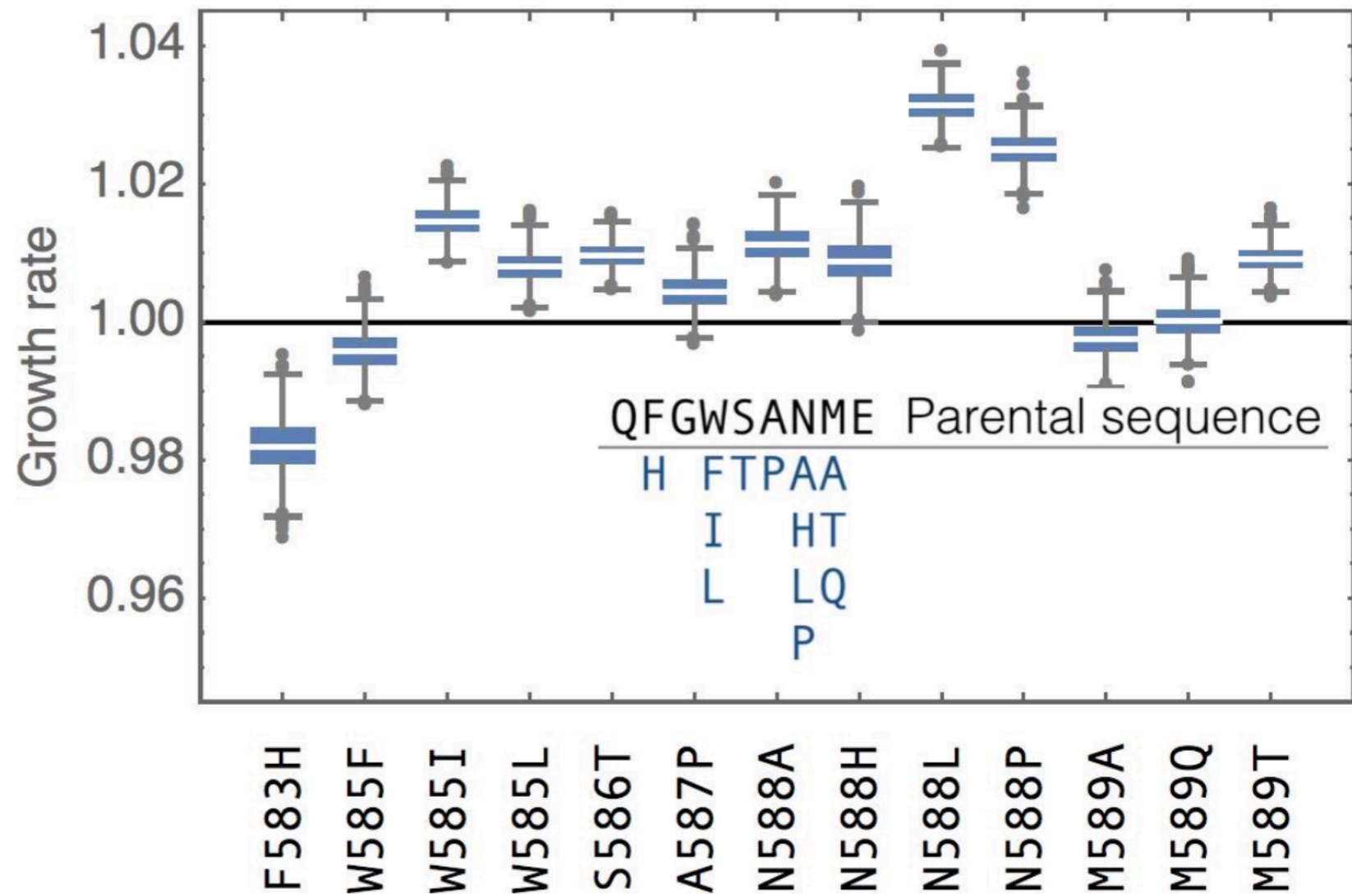
Complete fitness landscape of 640 combinations of mutations  
**Bank\*, Matuszewski\* et al., 2016, PNAS**



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



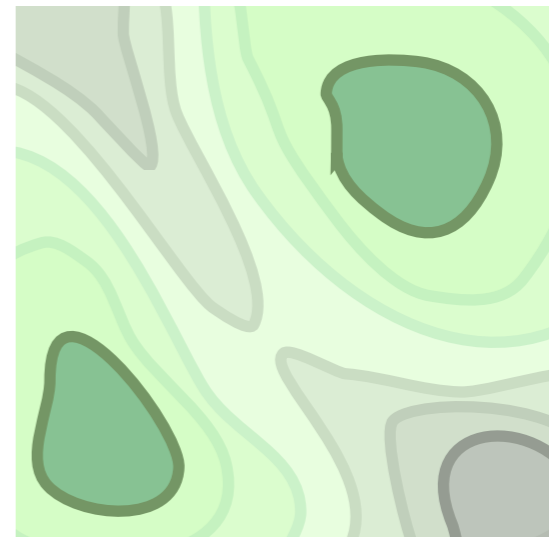
high salinity environment  
13 single-aa mutations  
2 replicates  
all possible combinations of aa's  
≈1600 nt mutations



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

# 3 QUESTIONS

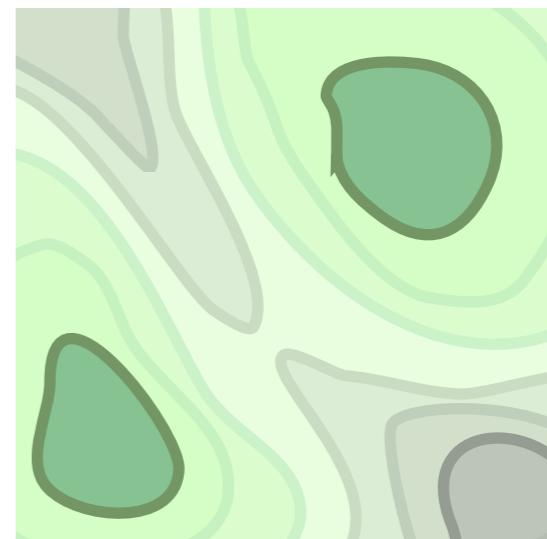
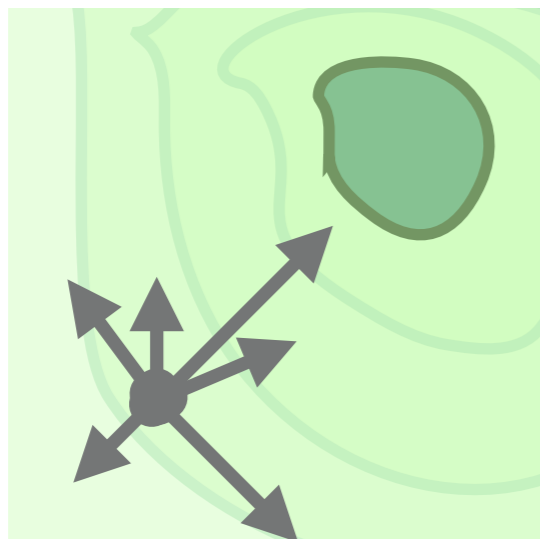
---



# 3 QUESTIONS

---

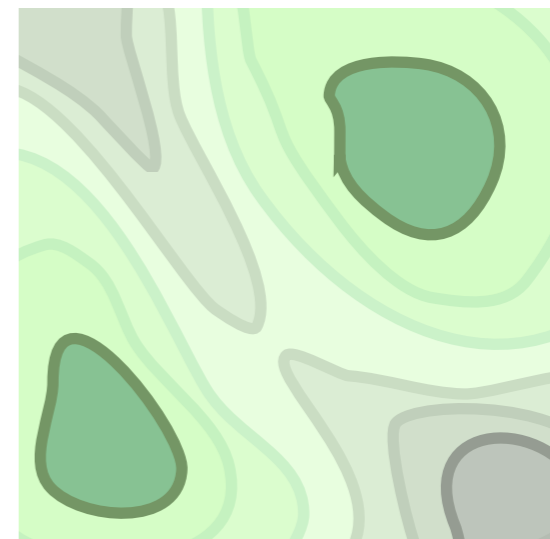
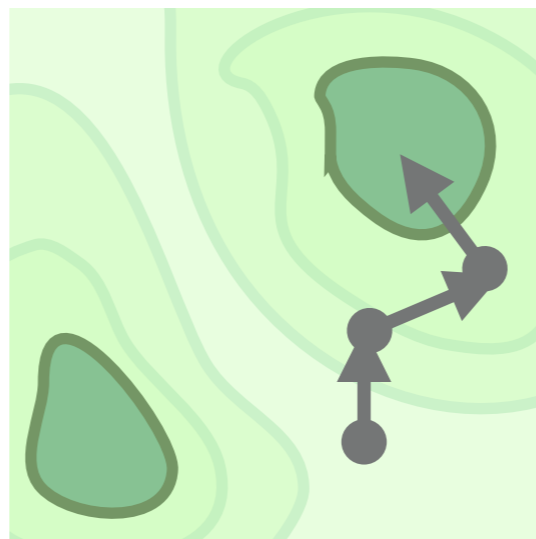
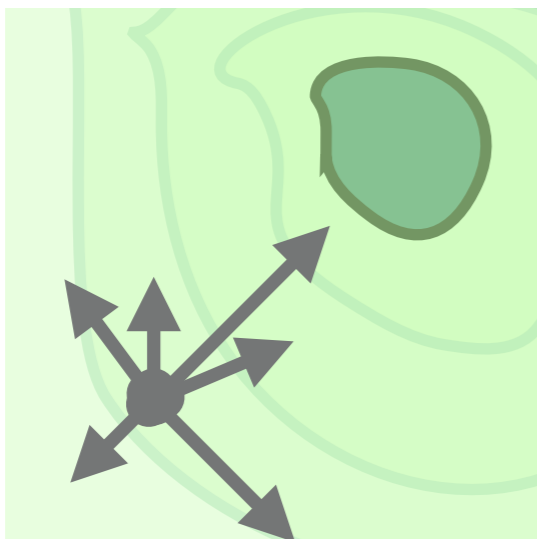
- ▶ Do single step mutations predict the way to the global optimum?



# 3 QUESTIONS

---

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

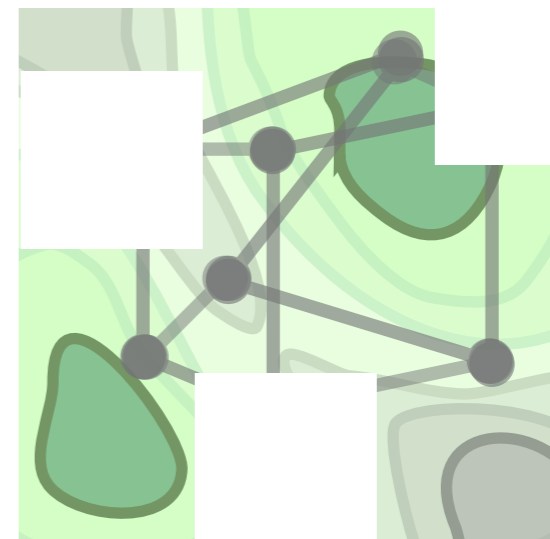
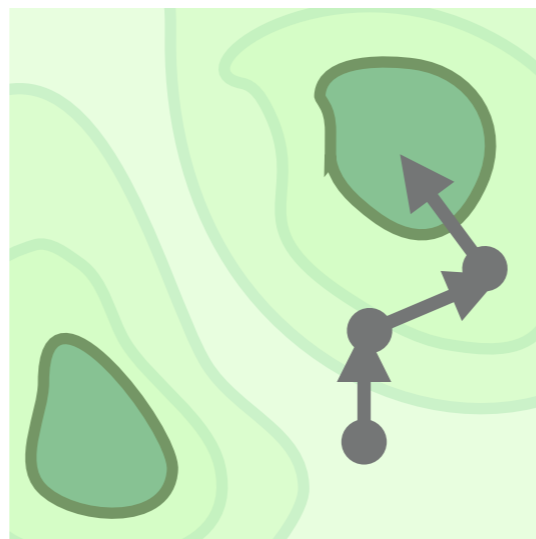
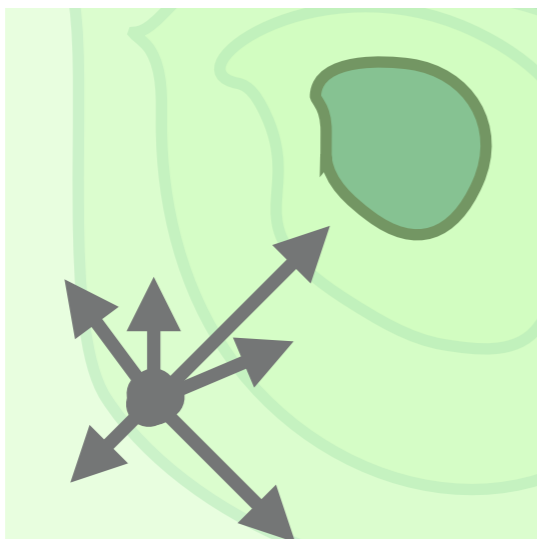




# 3 QUESTIONS

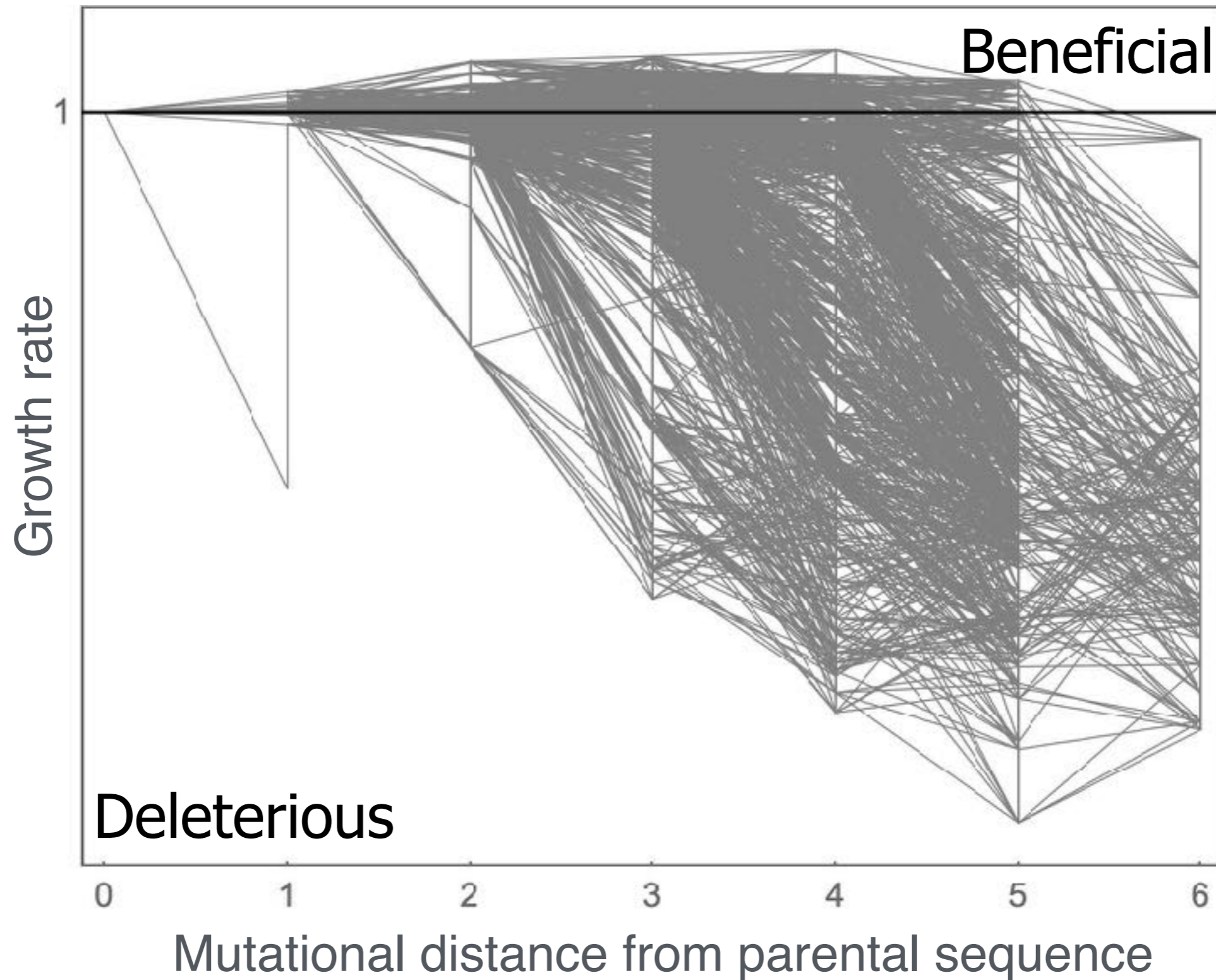
---

- 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?



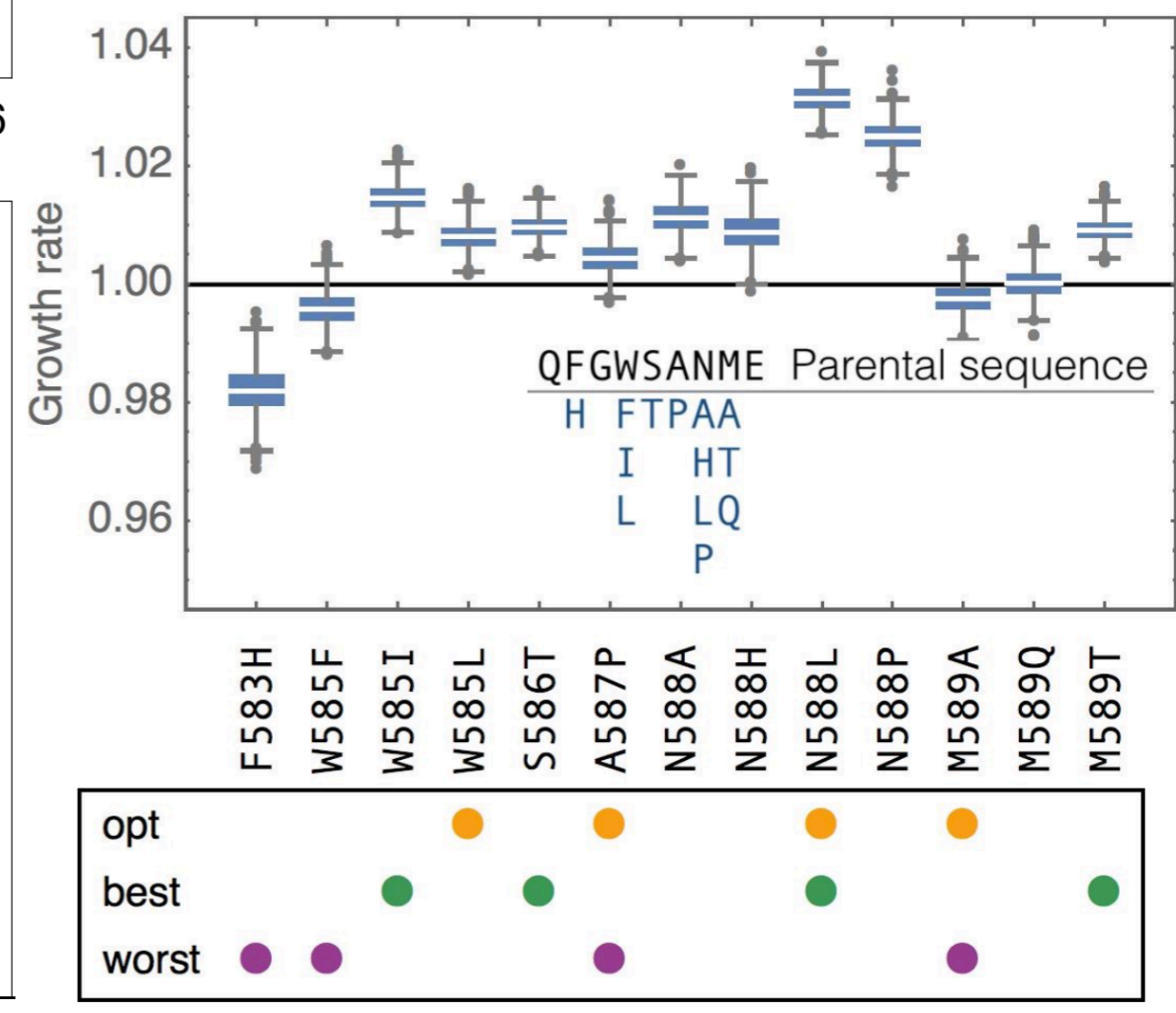
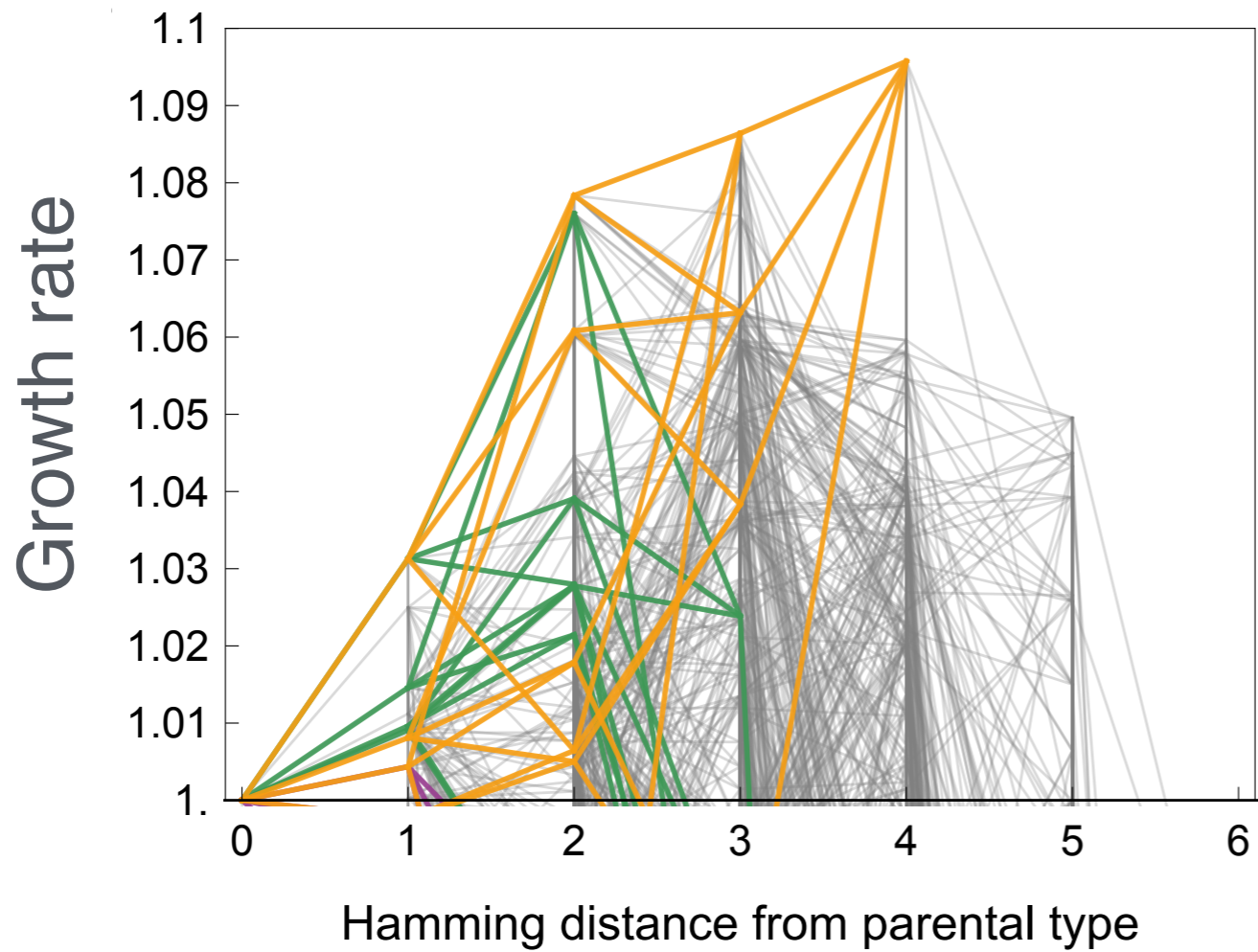
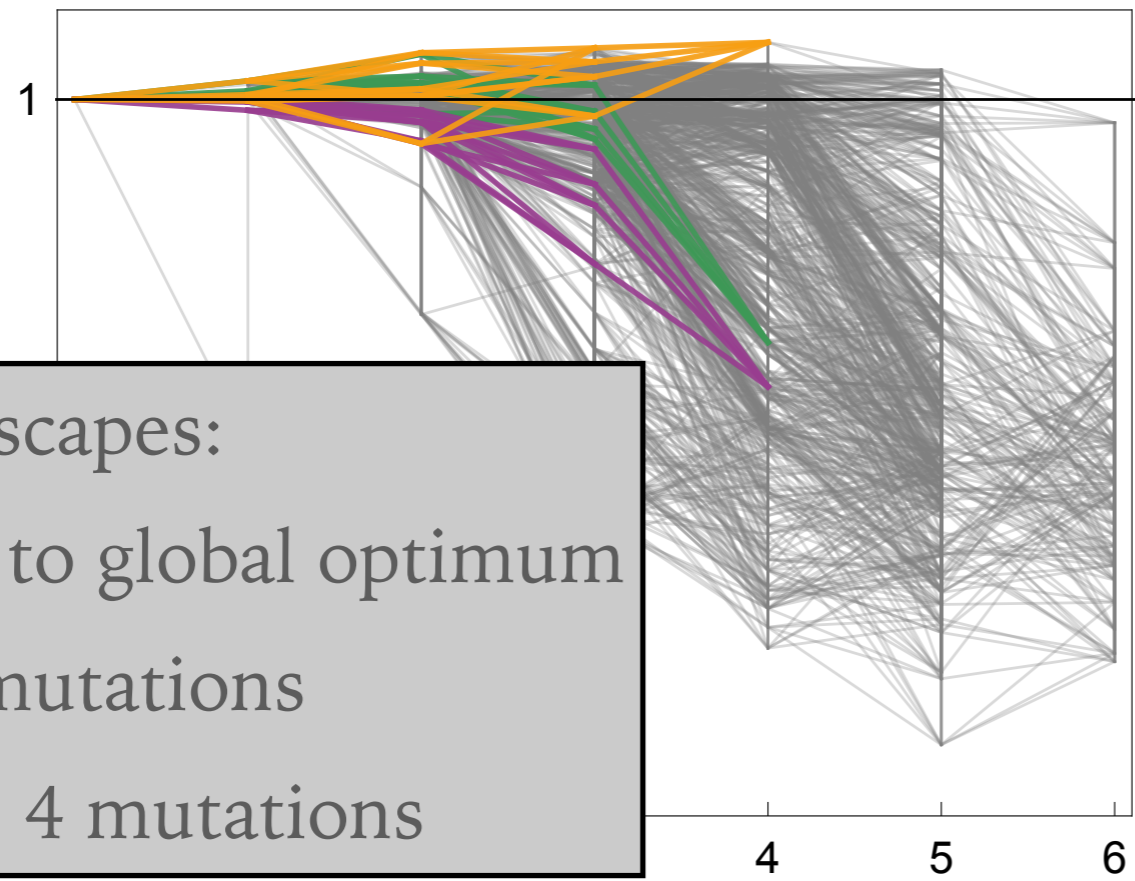
# A PICTURE OF THE WHOLE LANDSCAPE

---



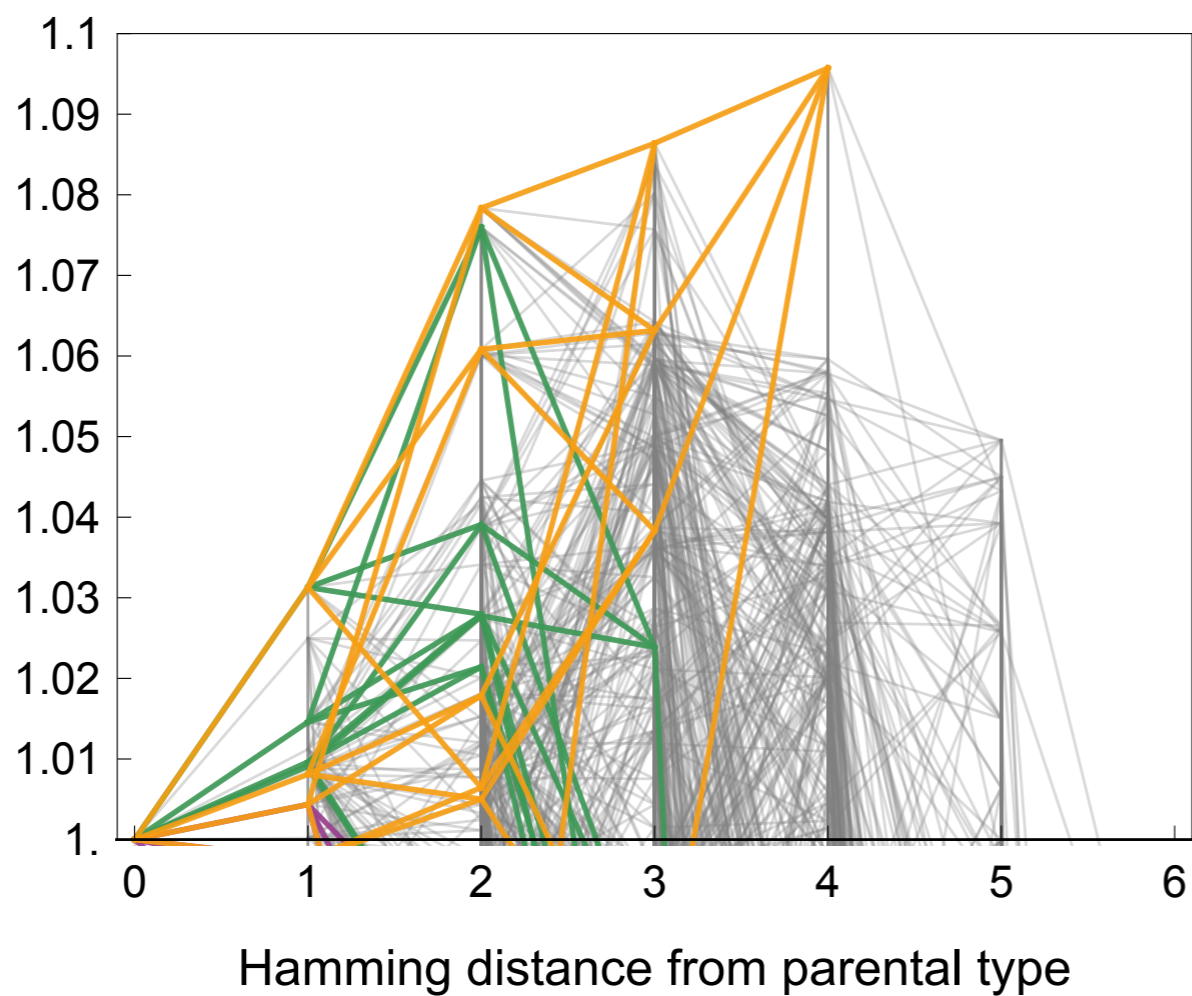
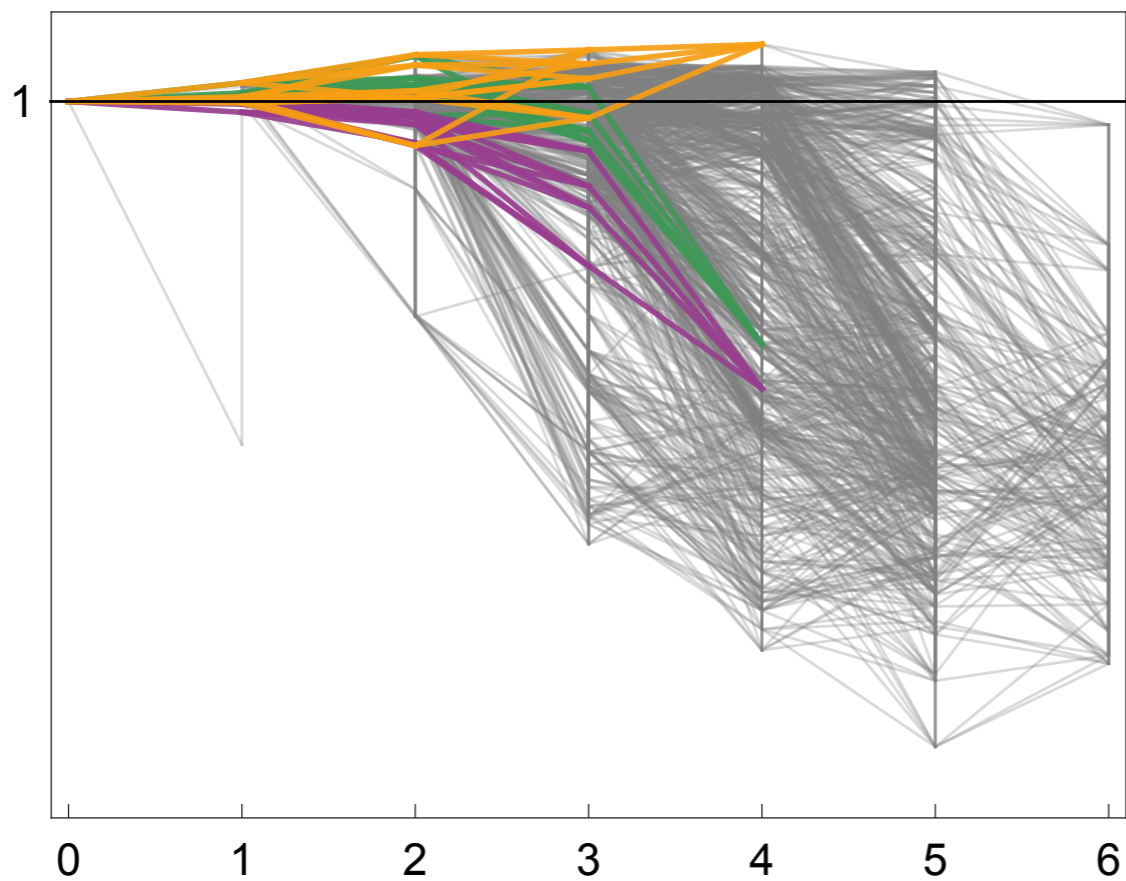
# 1 - DO SINGLE STEP MUTATIONS PREDICT THE WAY TO THE GLOBAL OPTIMUM?

Focal landscapes:  
 1) leading to global optimum  
 2) best 4 mutations  
 3) "worst" 4 mutations

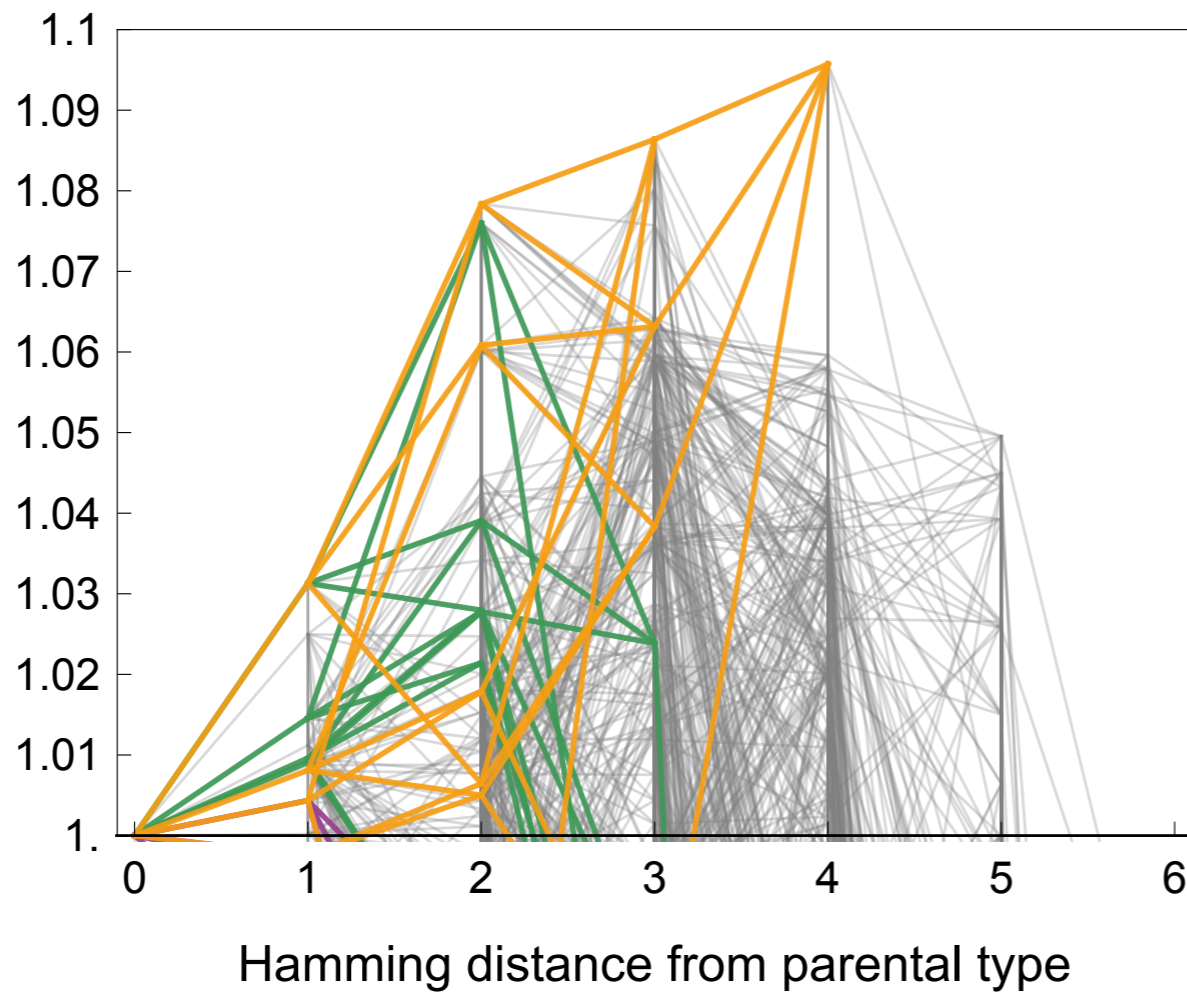
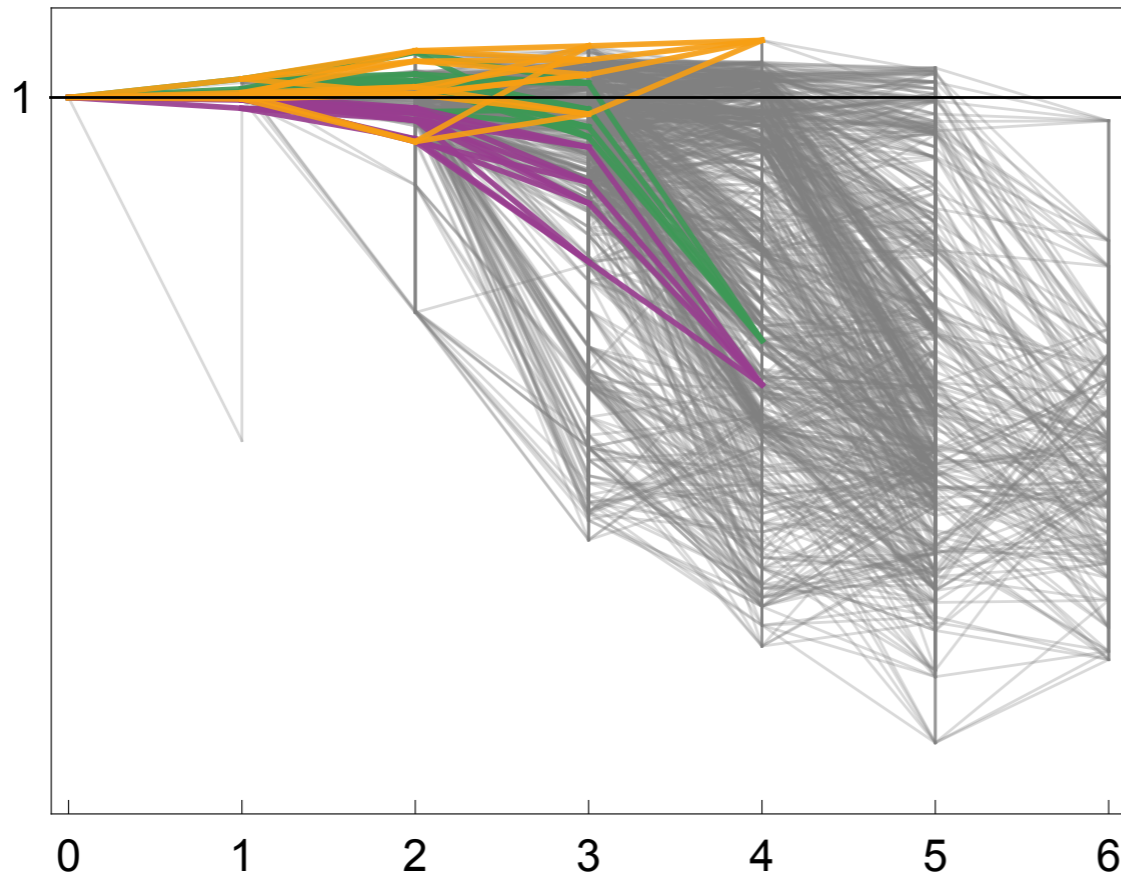




# Growth rate



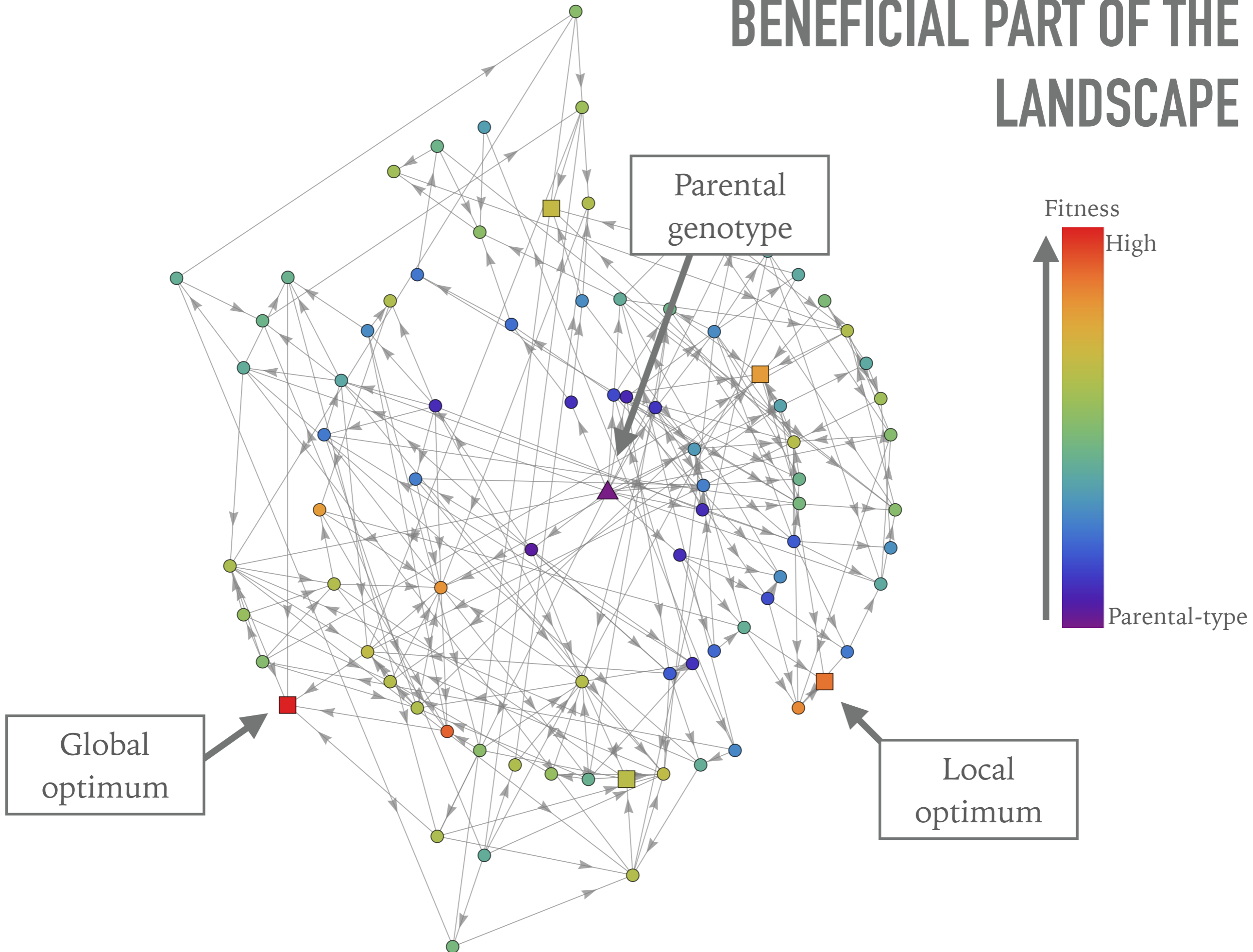
# Growth rate



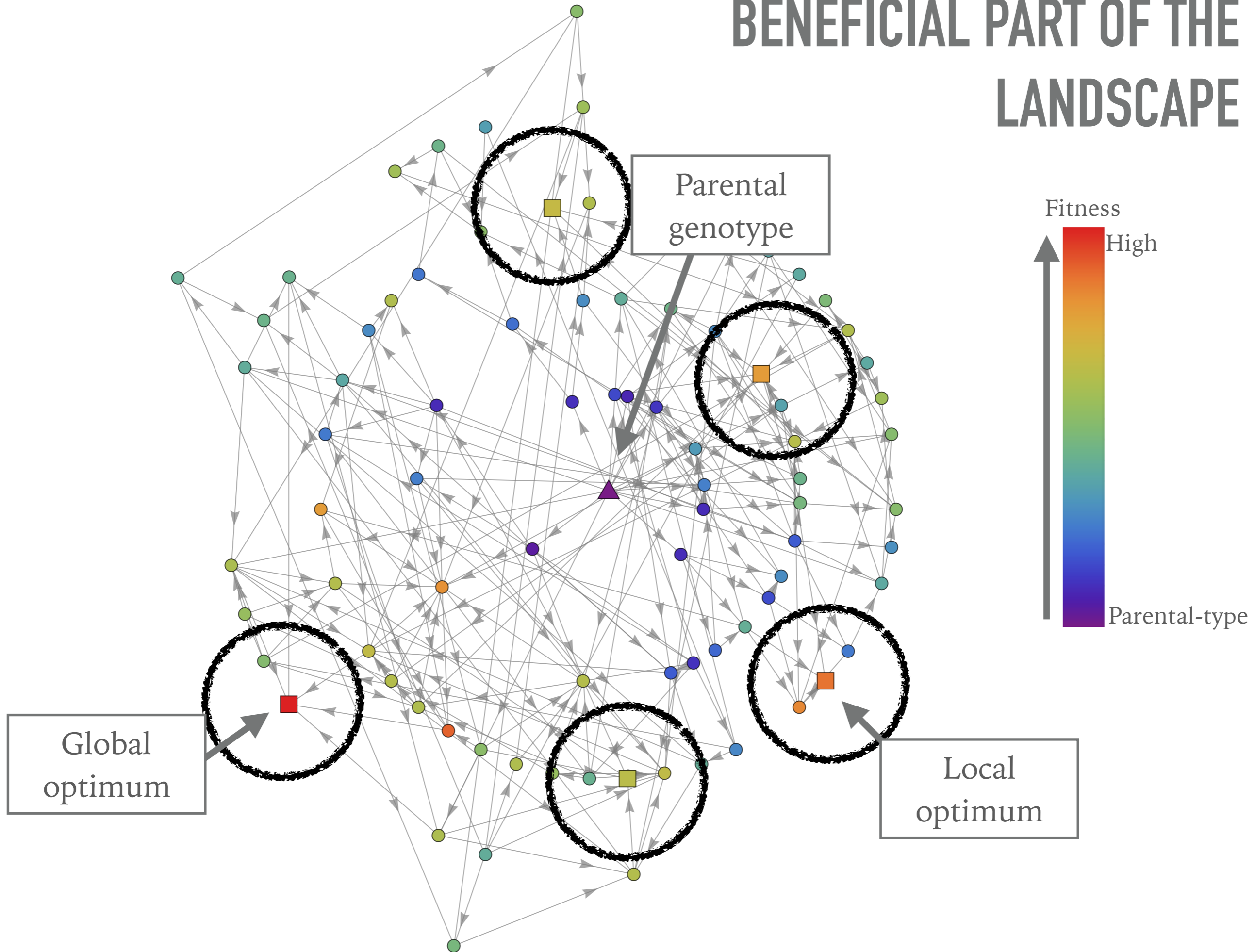
Strong positive and negative epistasis in the landscape.



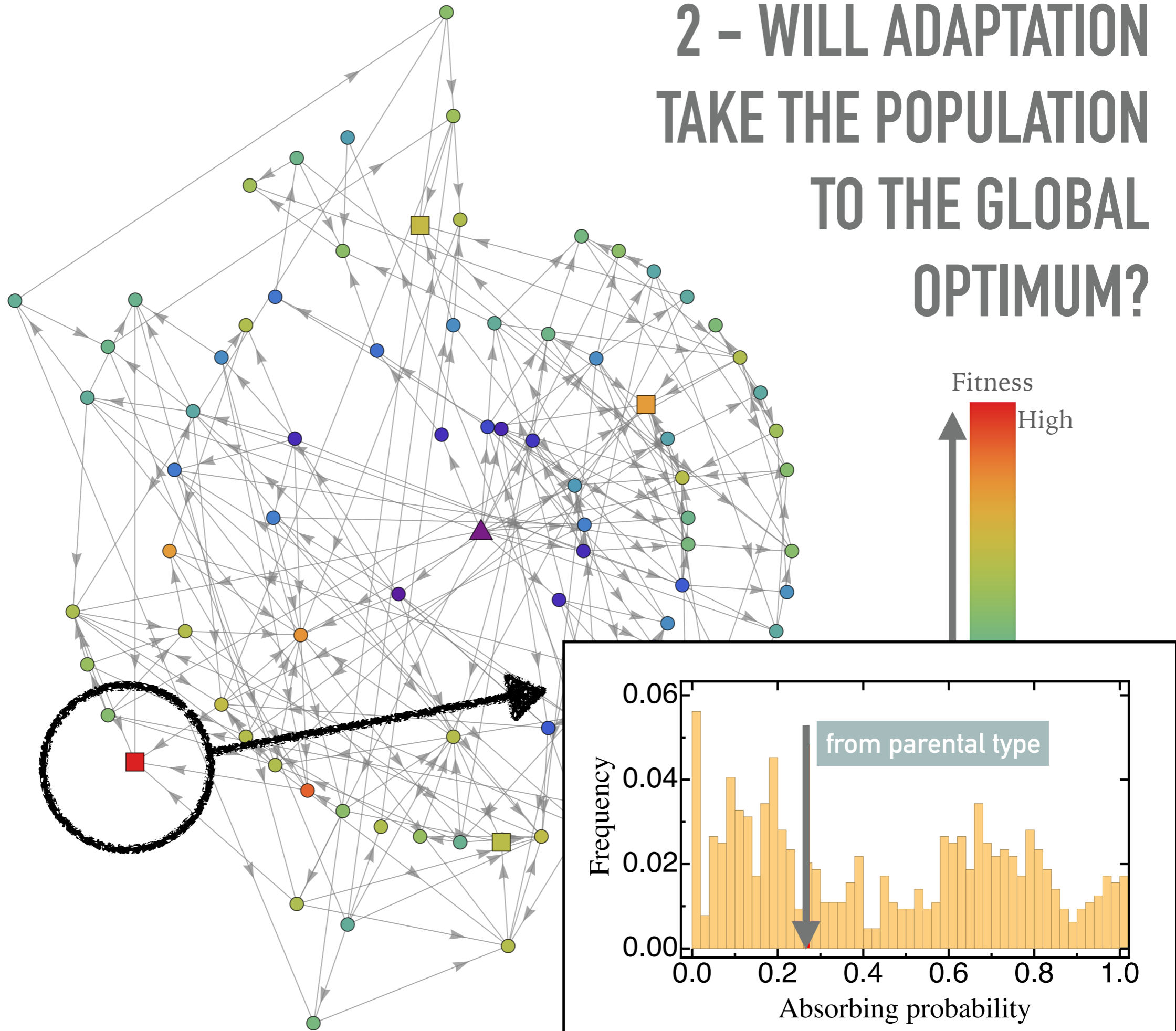
# BENEFICIAL PART OF THE LANDSCAPE



# BENEFICIAL PART OF THE LANDSCAPE

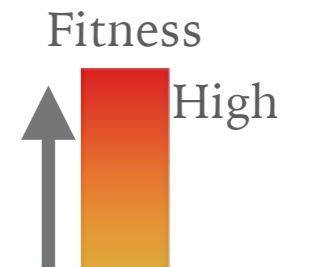
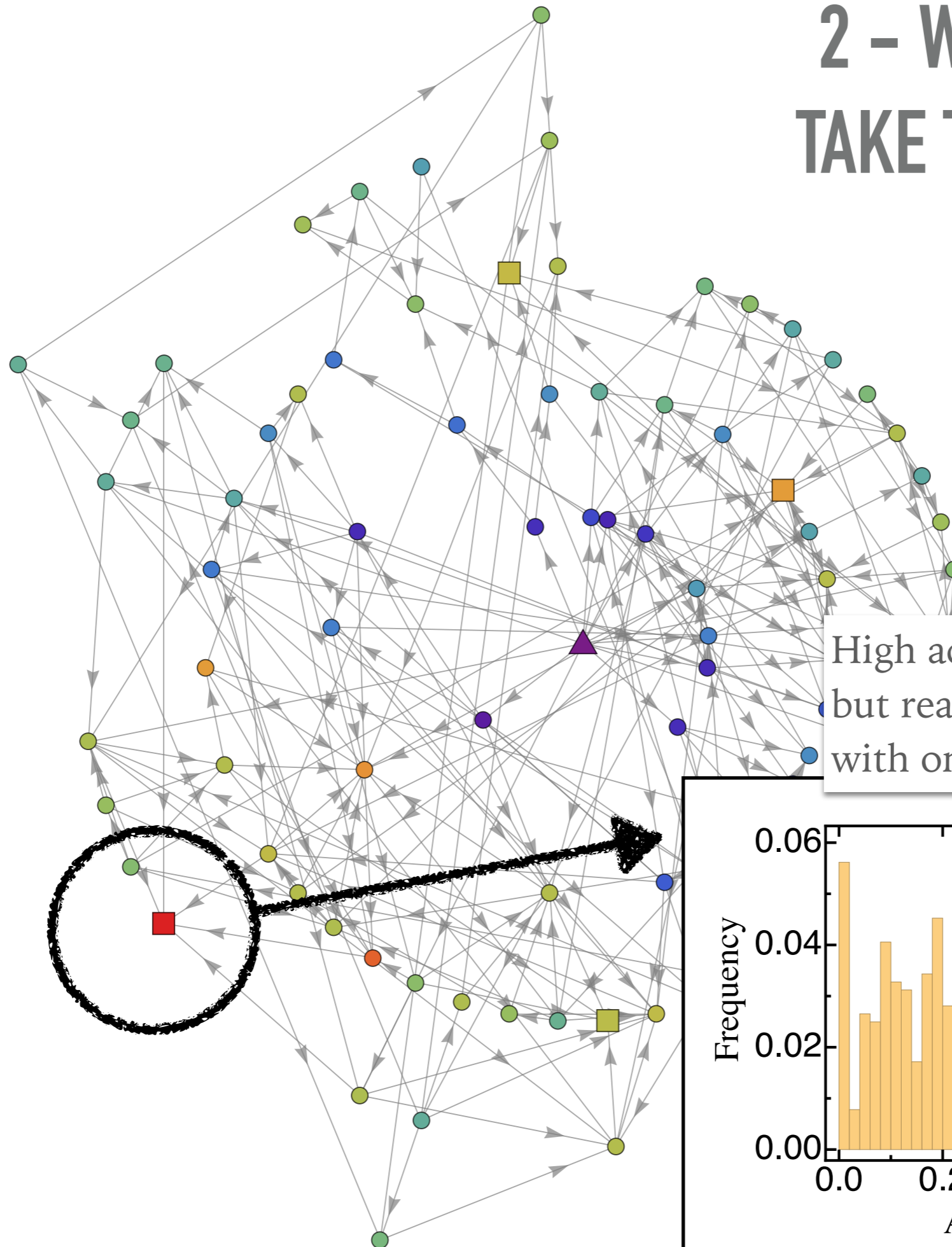


# 2 - WILL ADAPTATION TAKE THE POPULATION TO THE GLOBAL OPTIMUM?

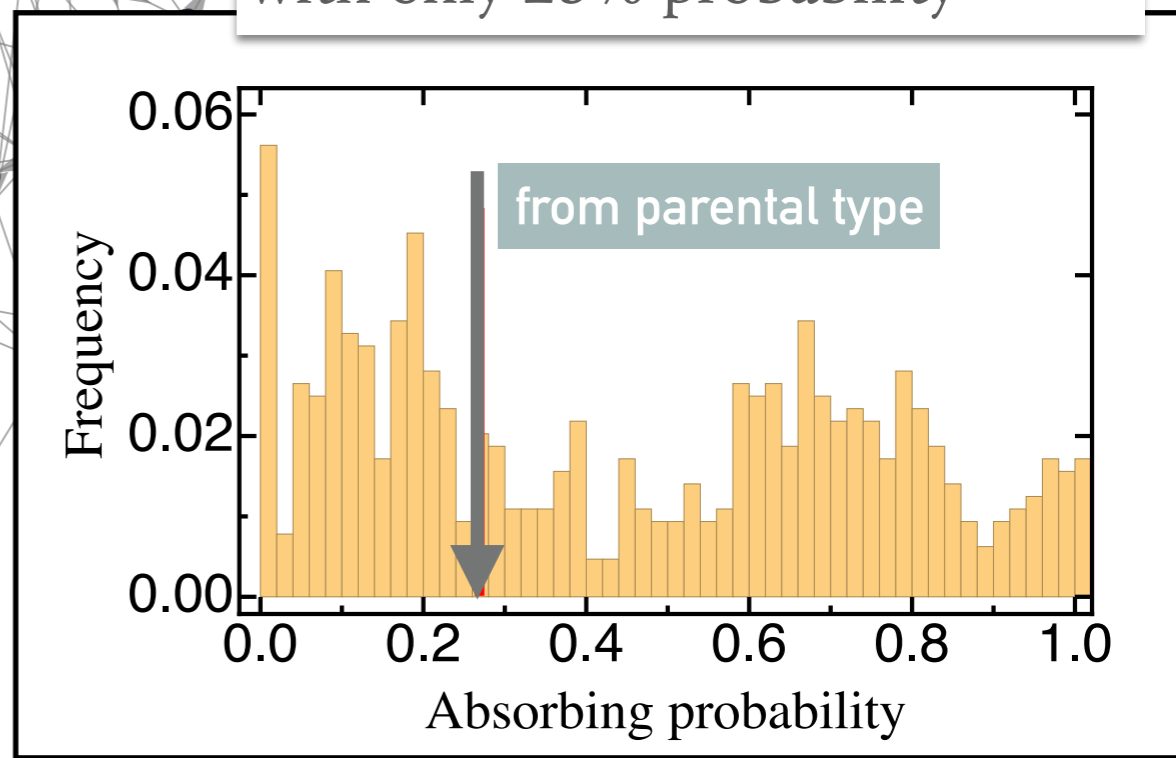




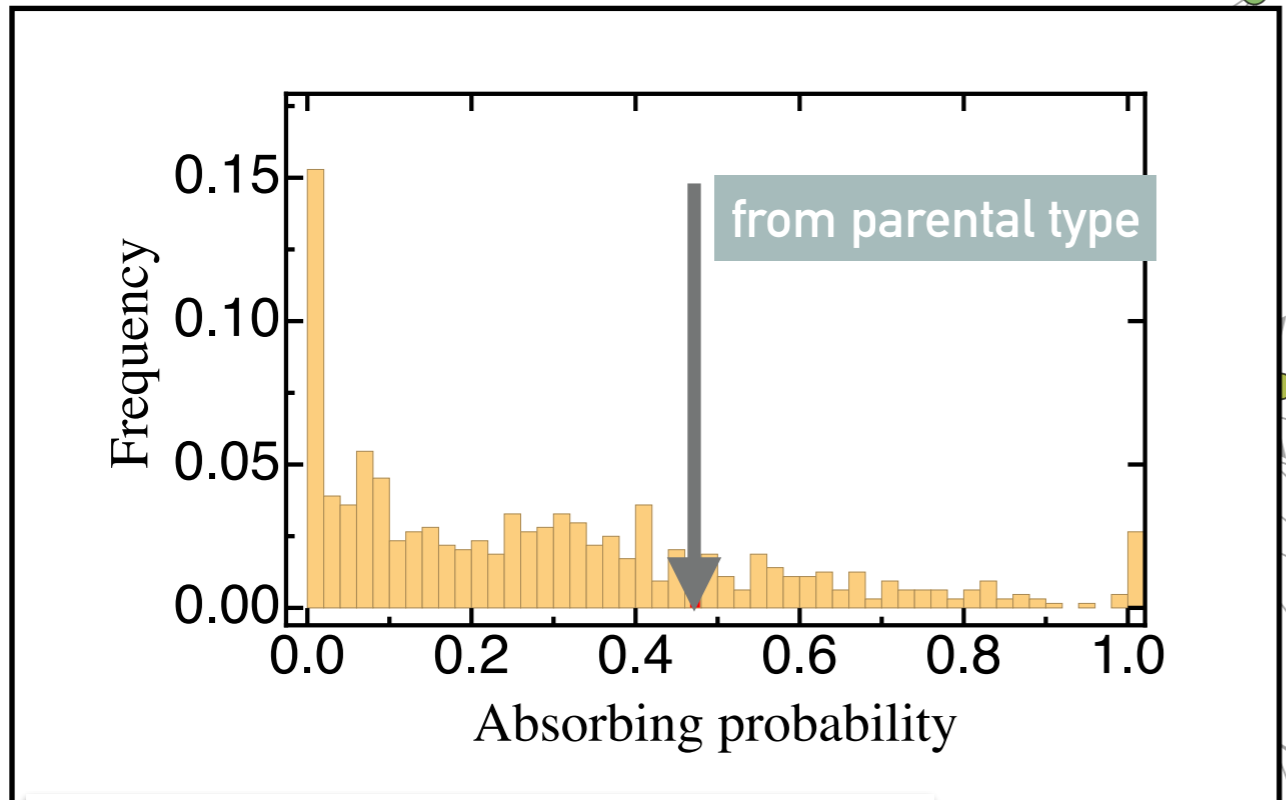
# 2 - WILL ADAPTATION TAKE THE POPULATION TO THE GLOBAL OPTIMUM?



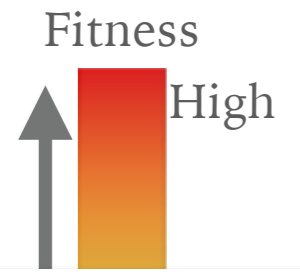
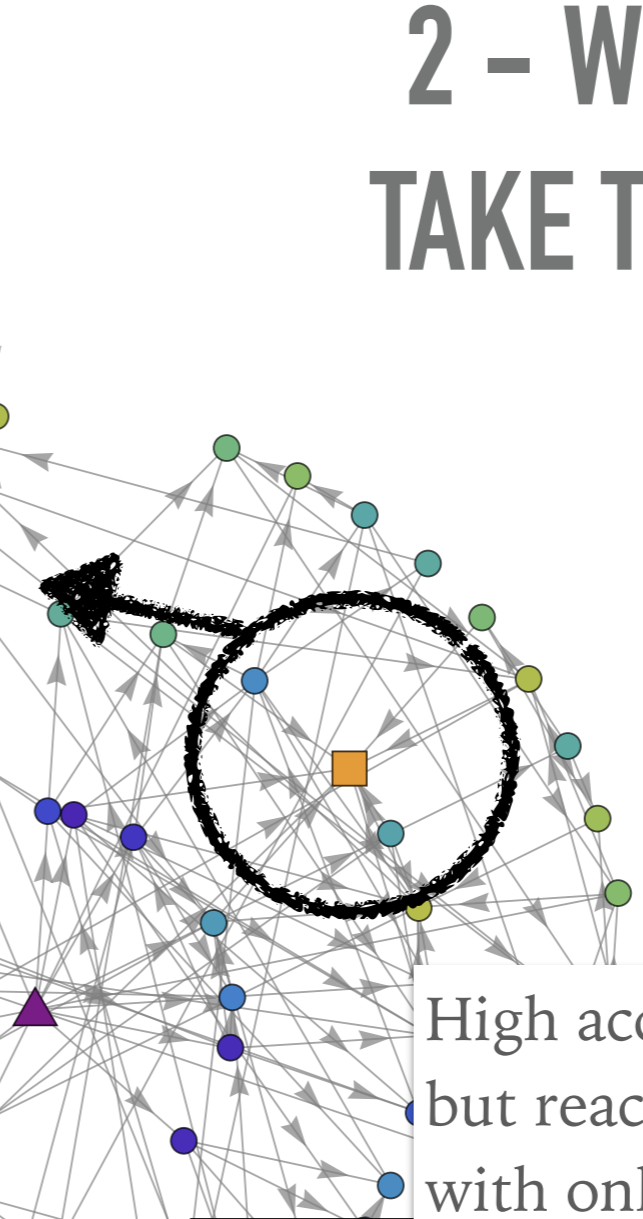
High accessibility on average,  
but reached from parental type  
with only 25% probability



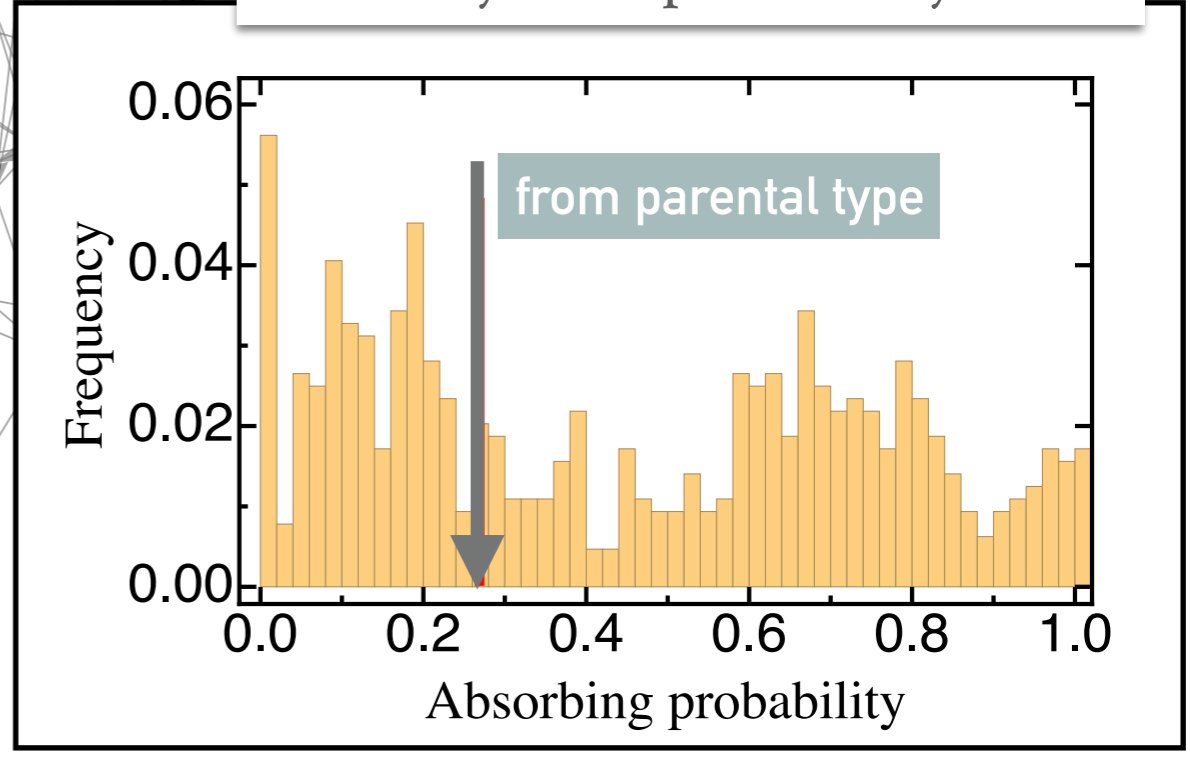
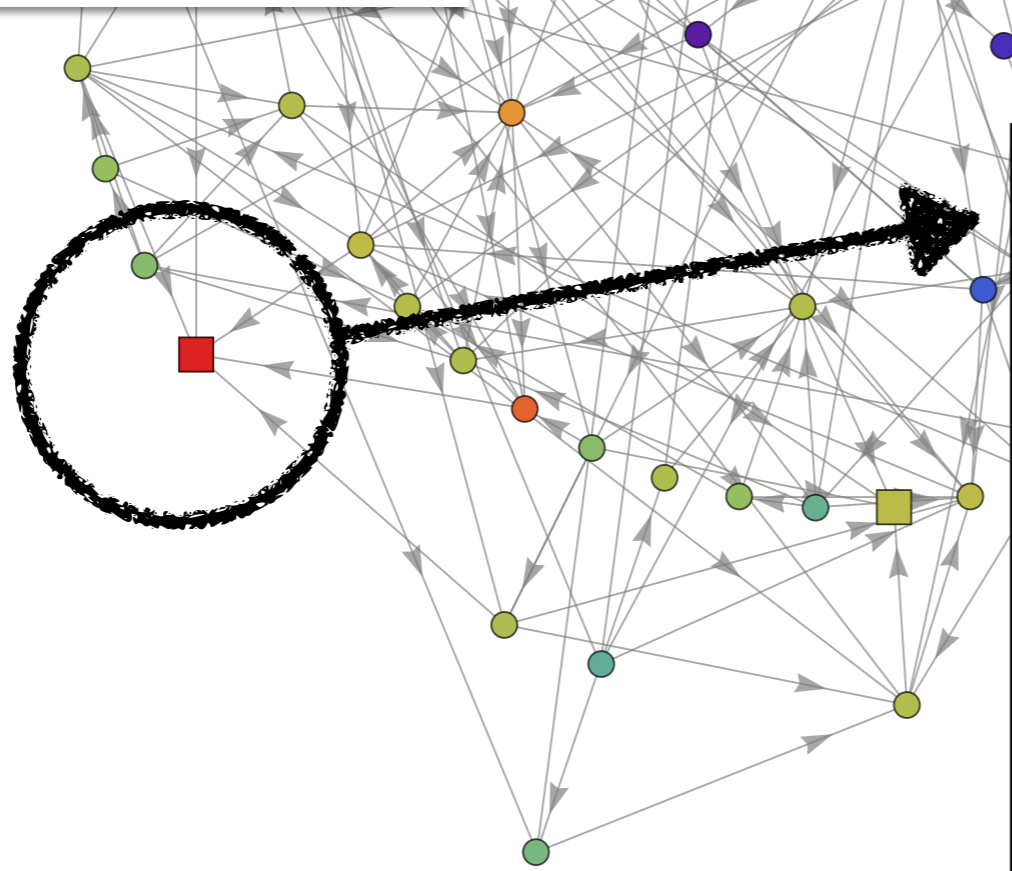
# 2 - WILL ADAPTATION TAKE THE POPULATION TO THE GLOBAL OPTIMUM?



Low accessibility on average, but almost 50% probability to be reached from parental type

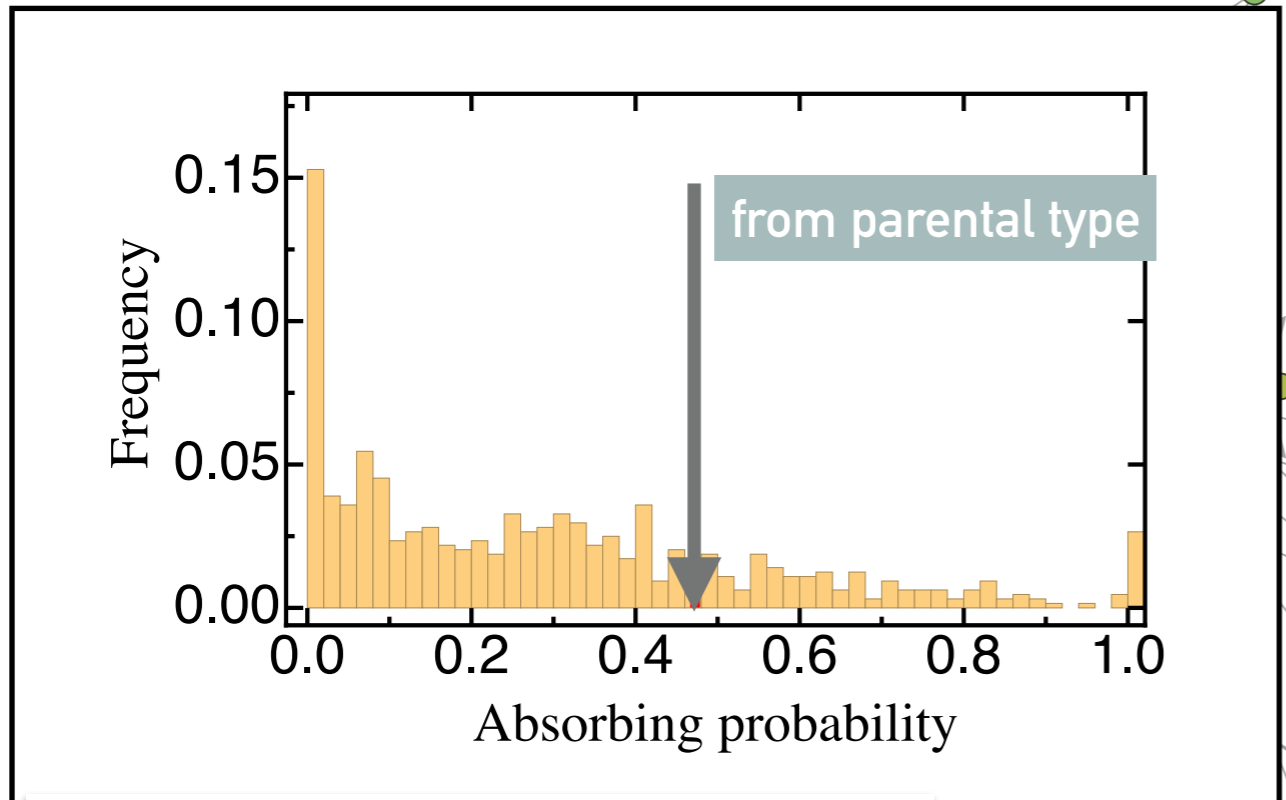


High accessibility on average, but reached from parental type with only 25% probability

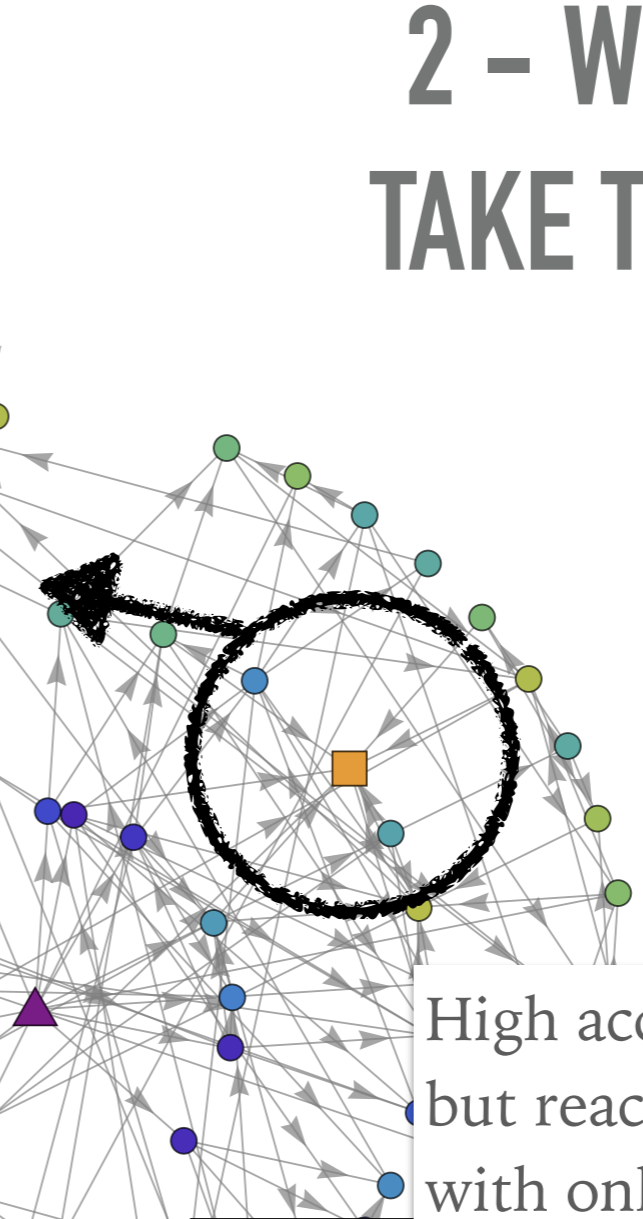




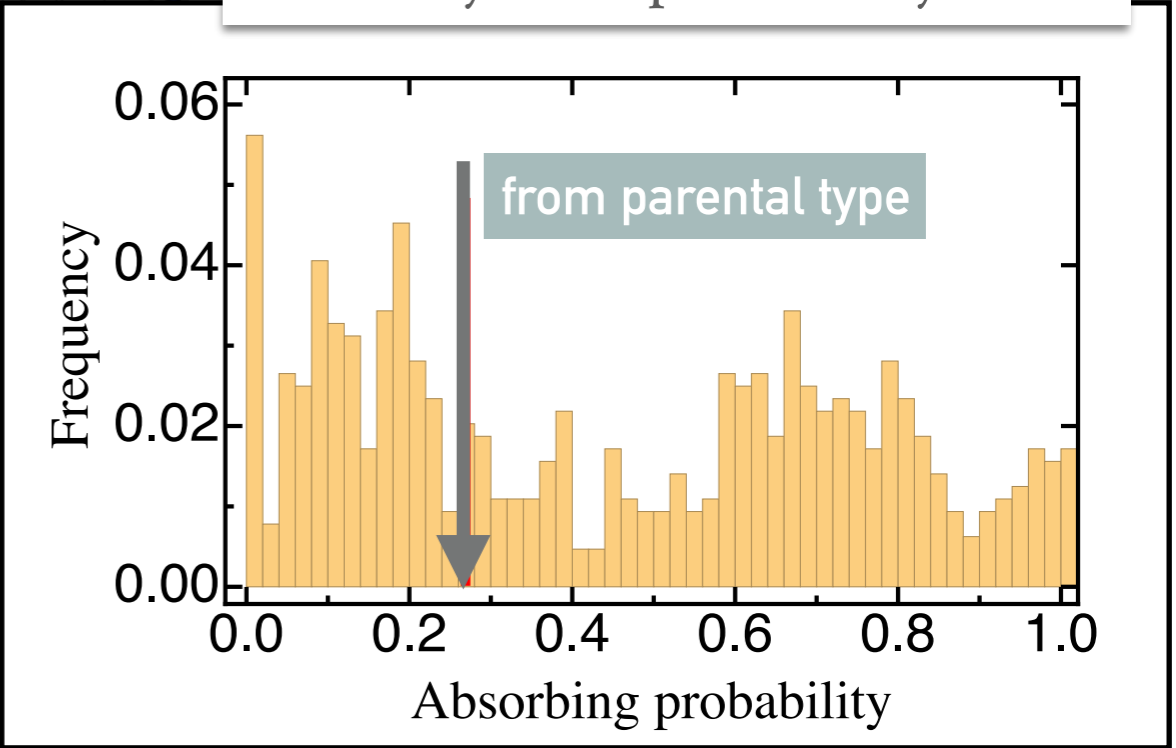
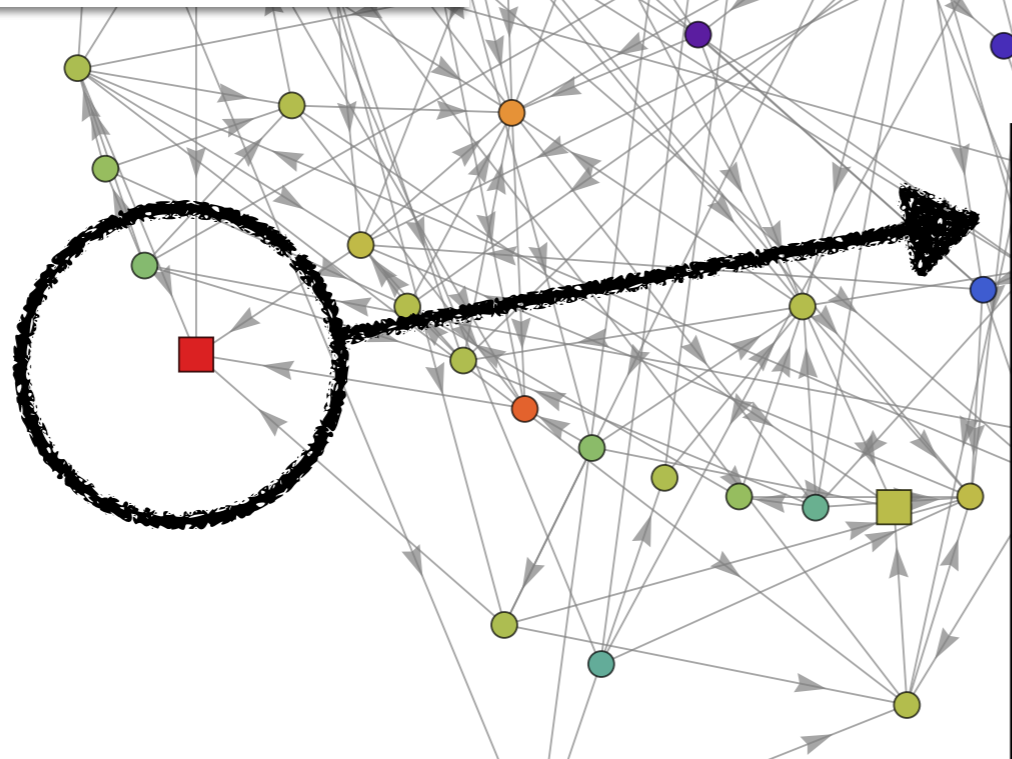
# 2 - WILL ADAPTATION TAKE THE POPULATION TO THE GLOBAL OPTIMUM?



Low accessibility on average, but almost 50% probability to be reached from parental type



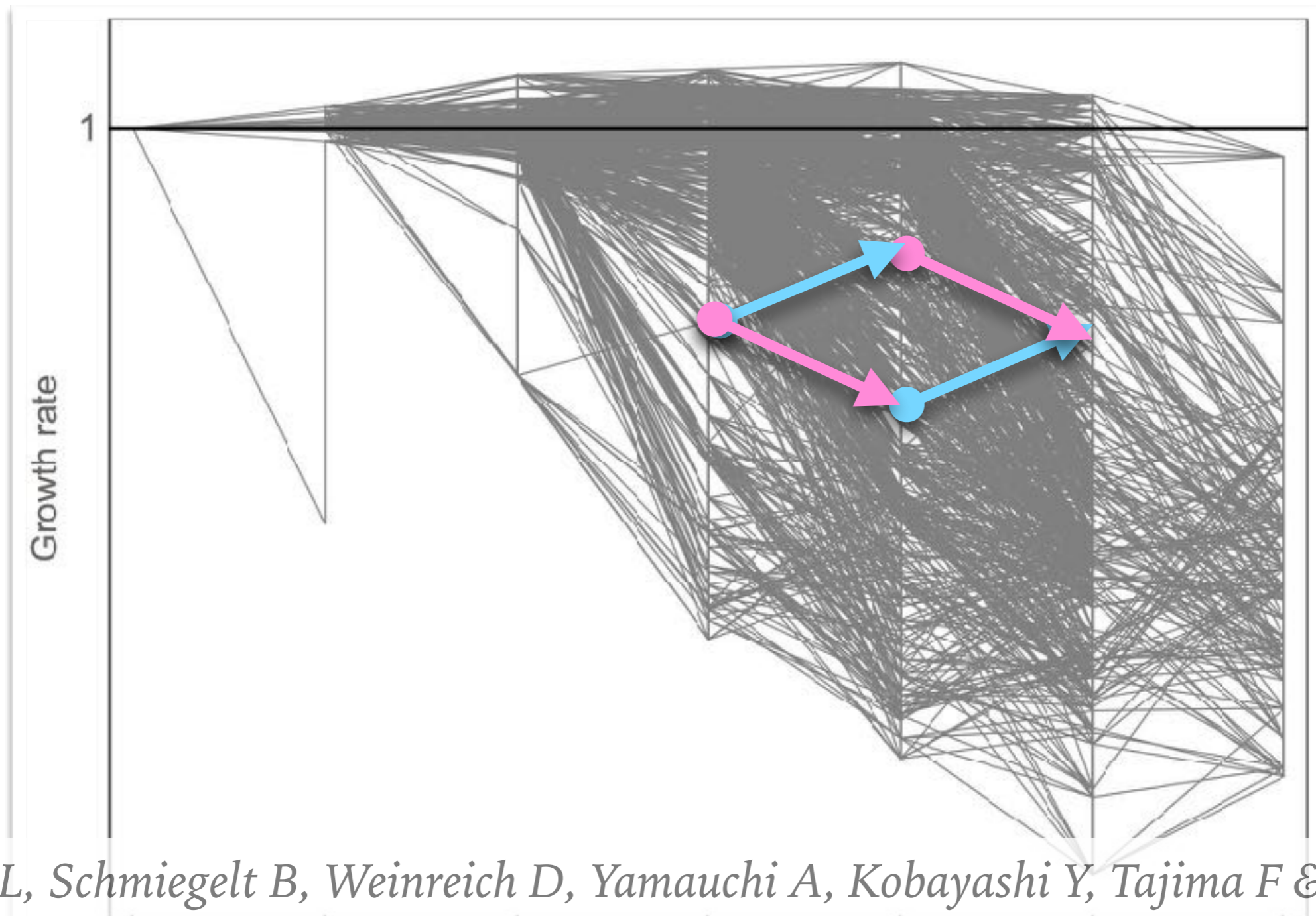
High accessibility on average, but reached from parental type with only 25% probability



Evolution from parental type may stall at local optimum

# LANDSCAPE STATISTICS INDEPENDENT OF REFERENCE

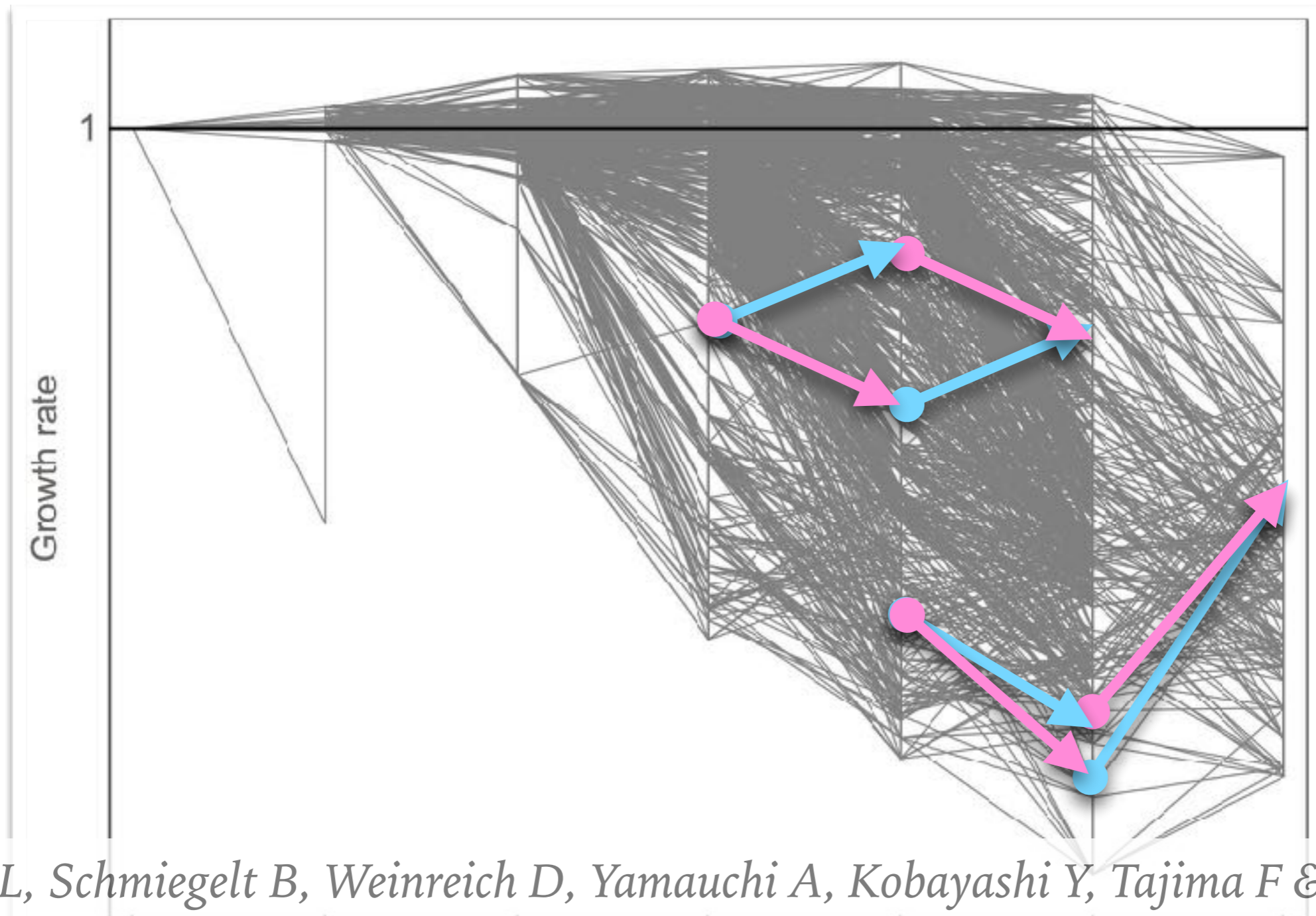
---



*Ferretti L, Schmiegel 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, Schmiegel 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*



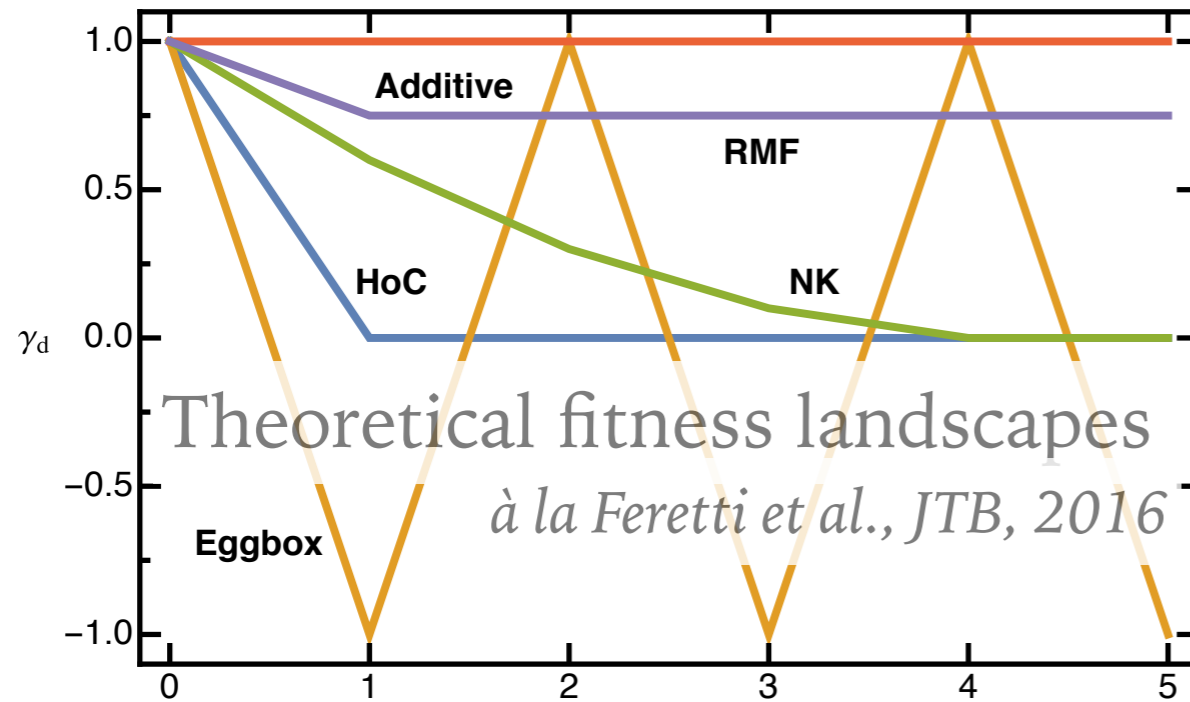






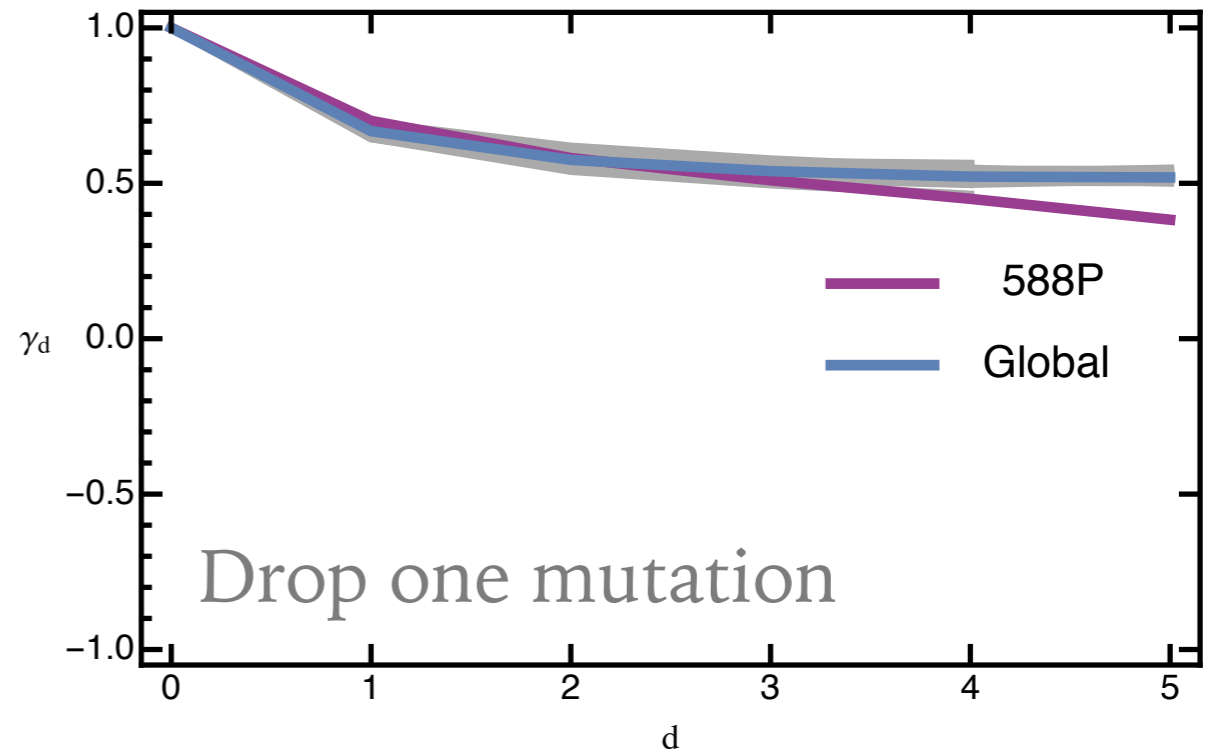
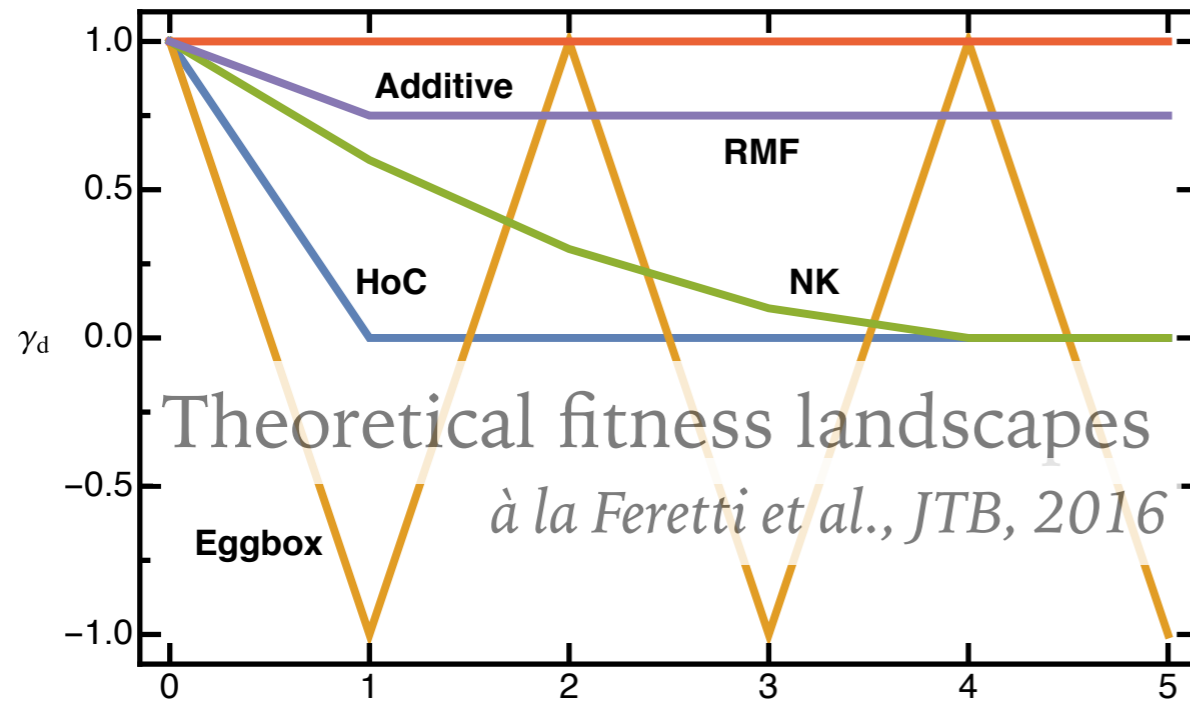
# 3 – CAN WE INFER AN UNKNOWN PART OF THE LANDSCAPE?

---



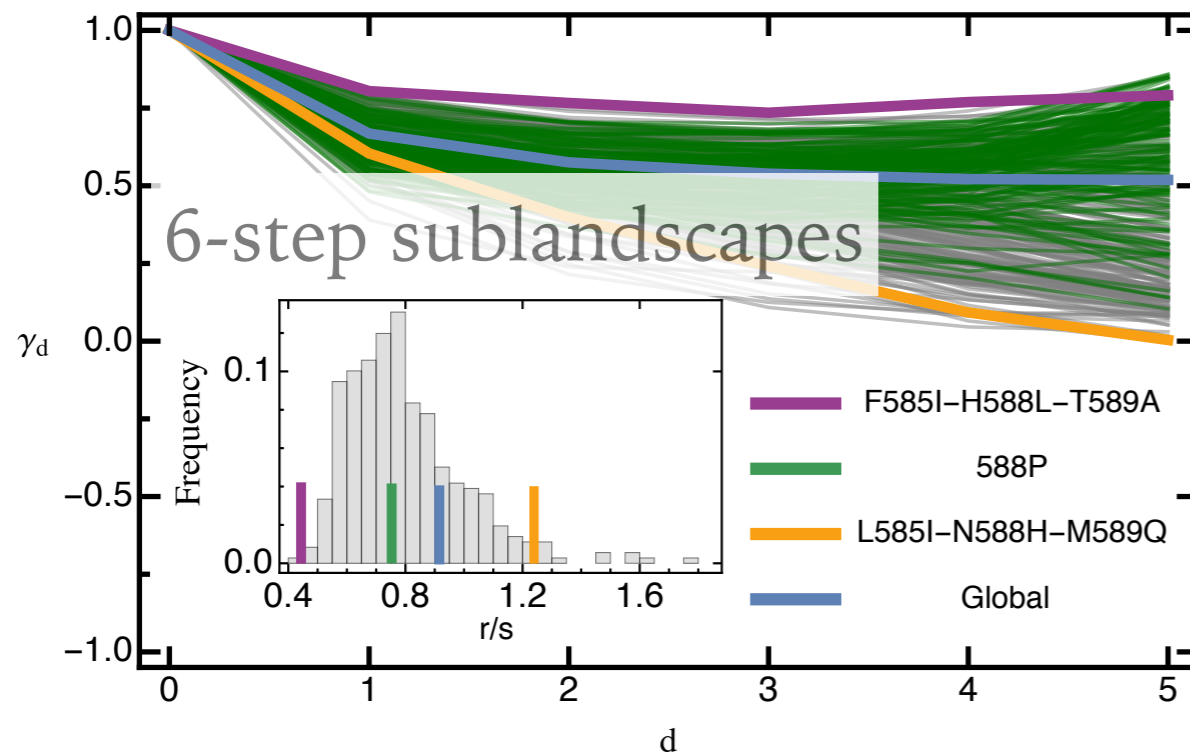
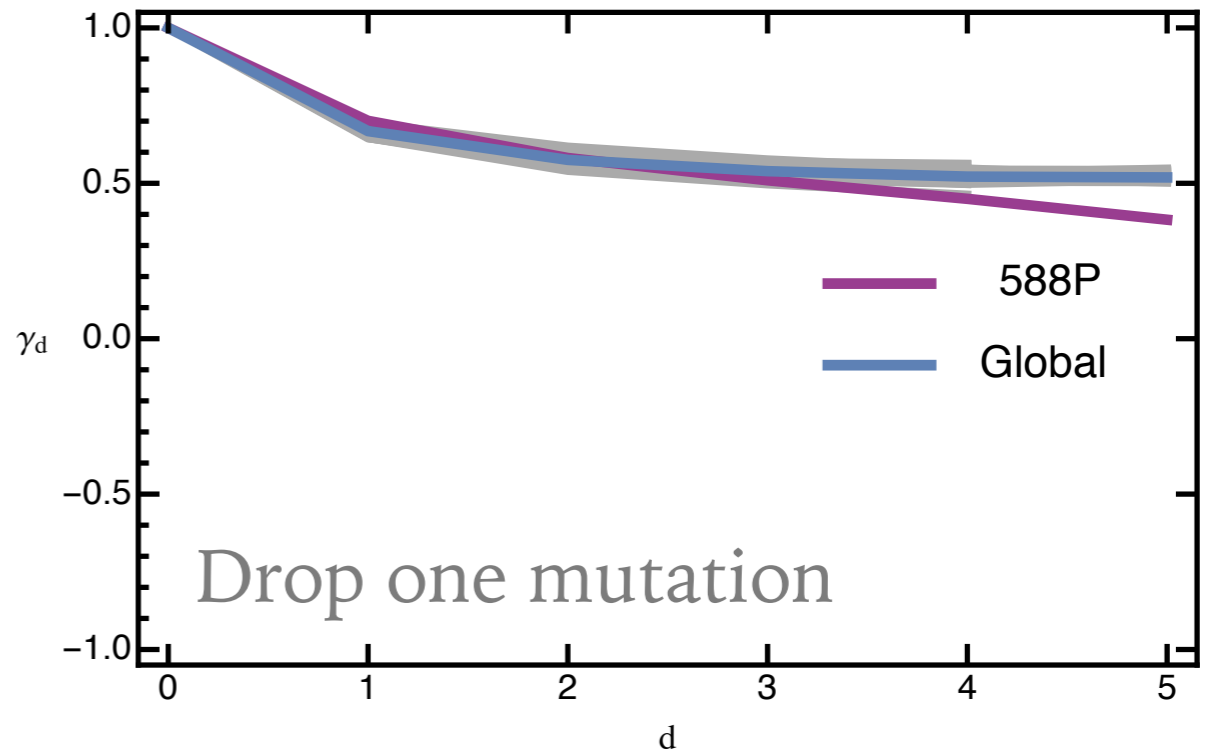
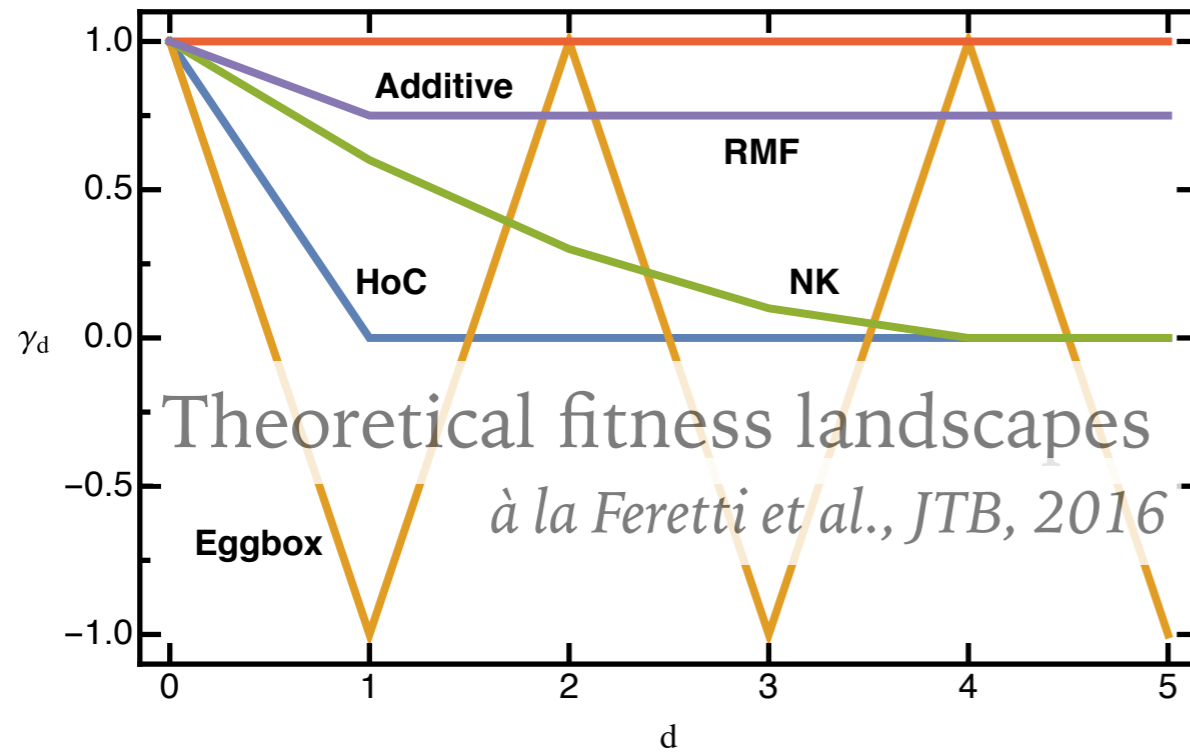
Epistasis measure  $\gamma_d$  with increasing Hamming distance  $d$

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



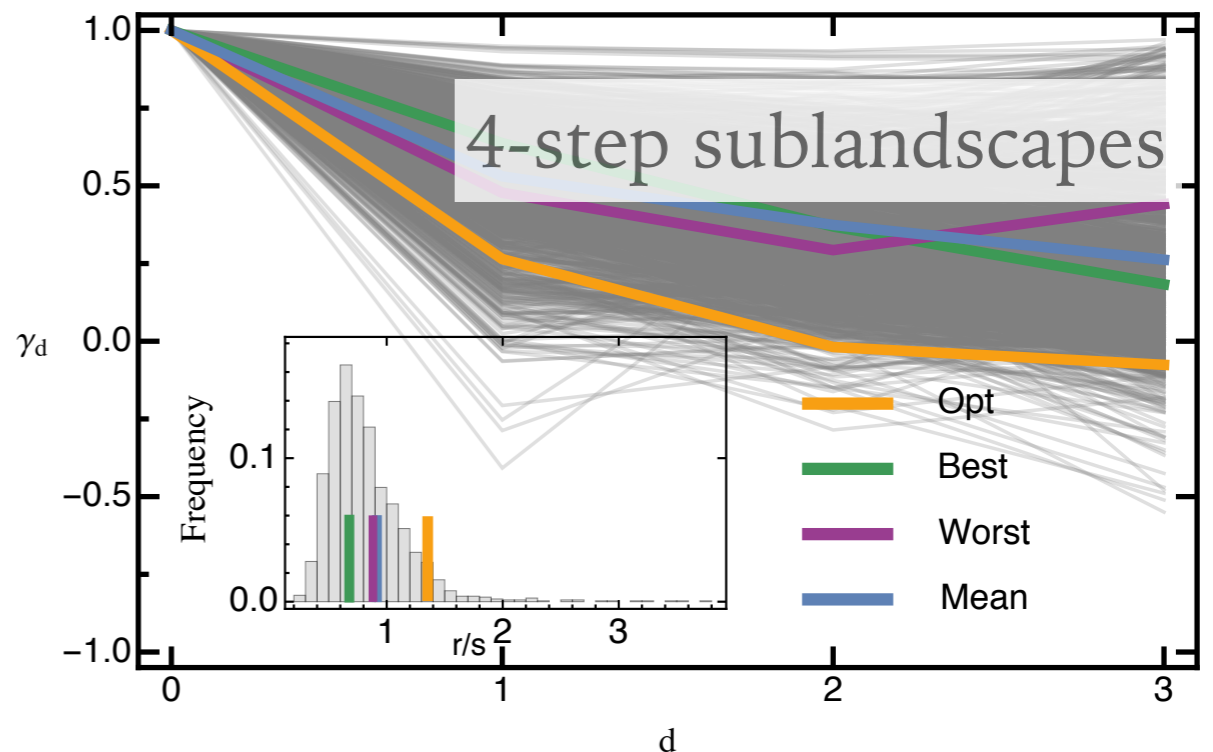
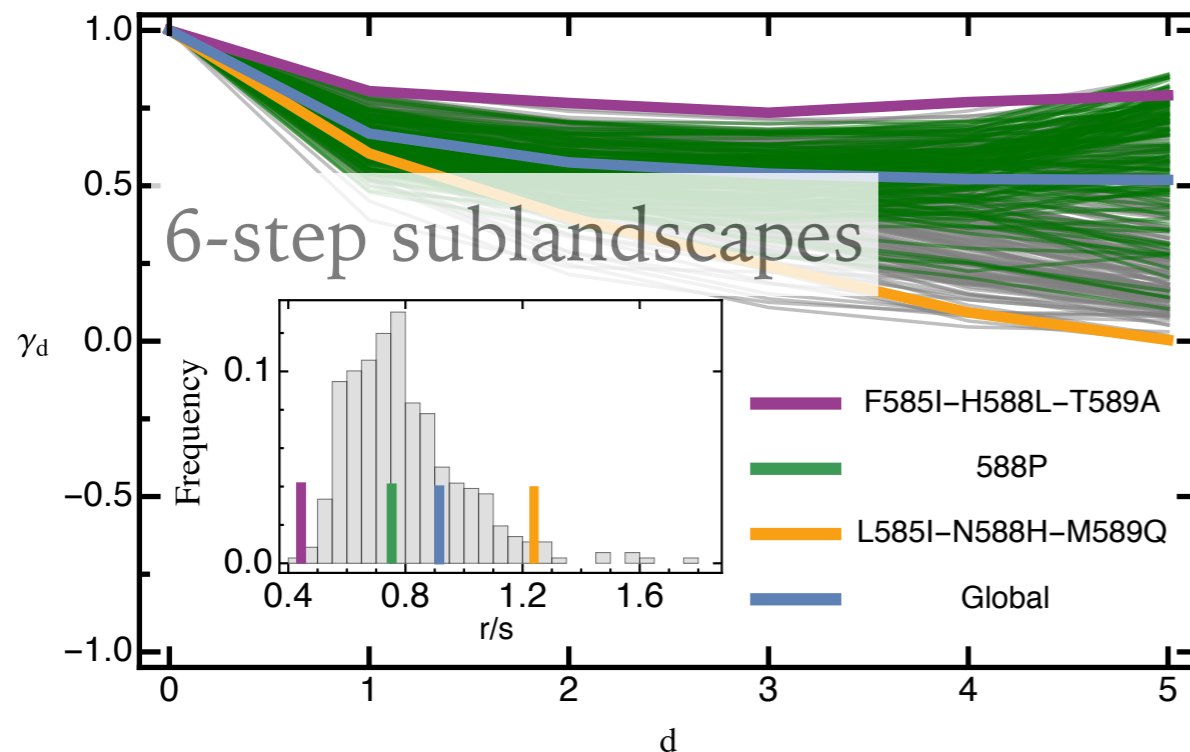
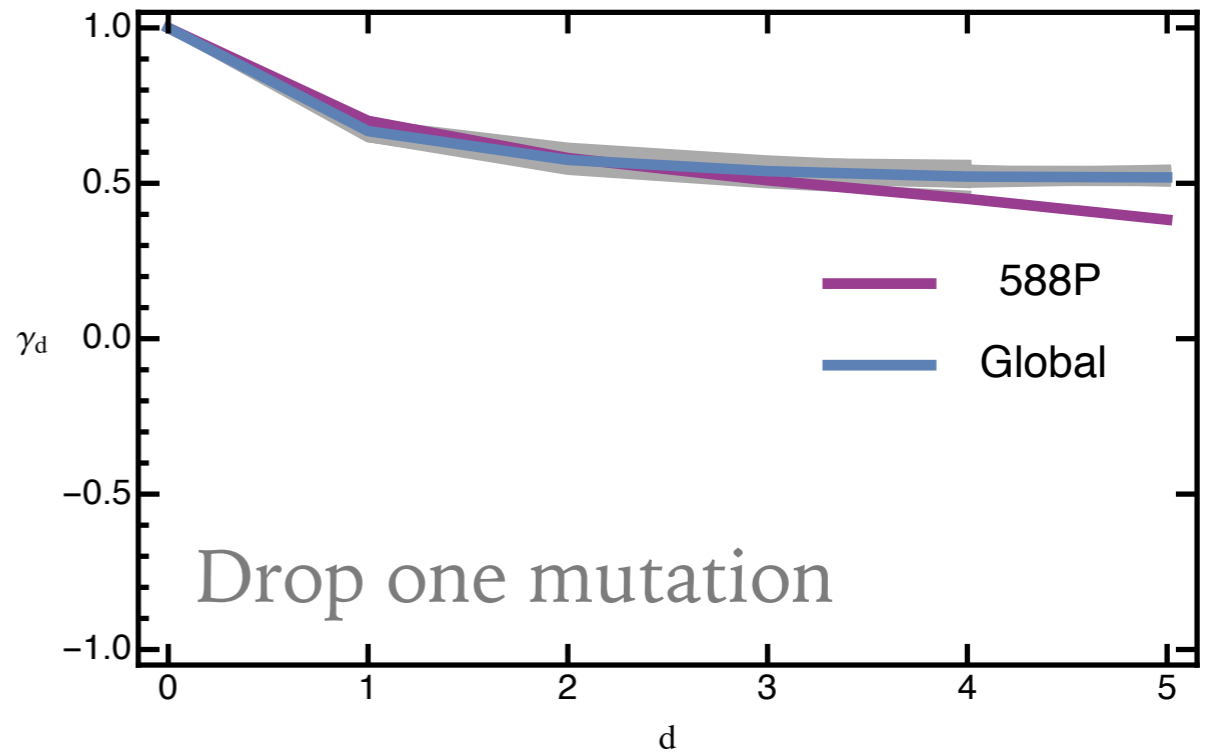
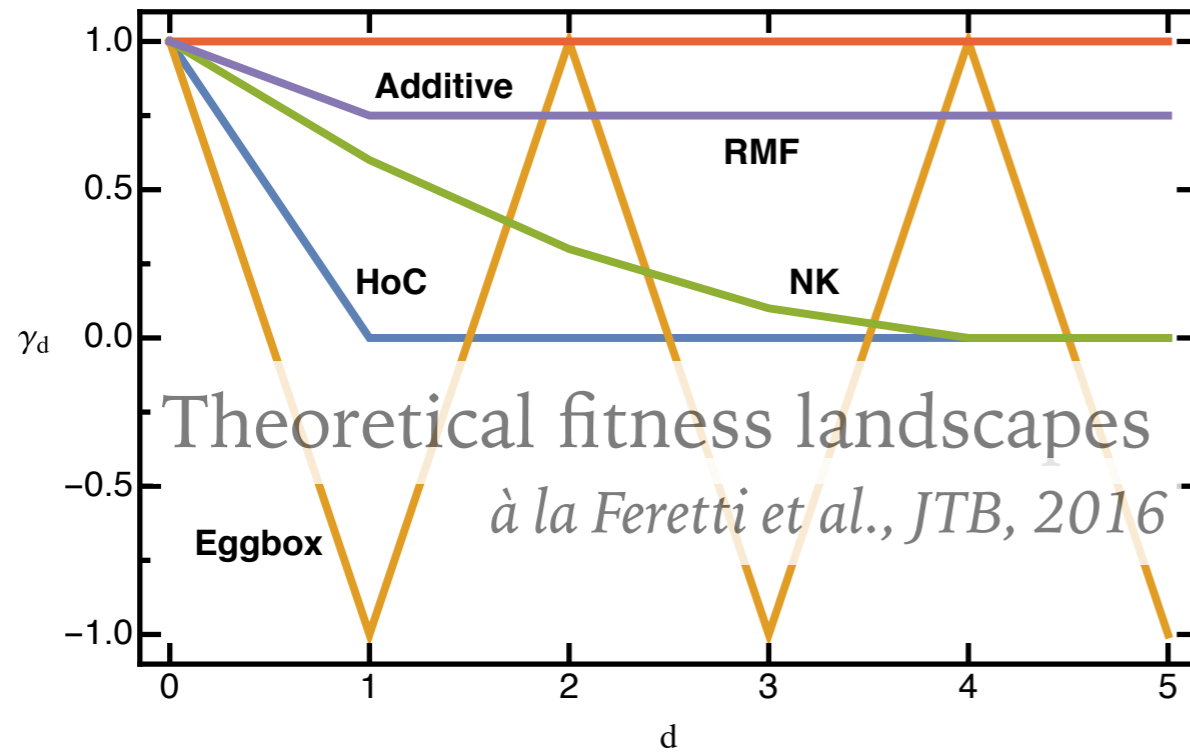
Epistasis measure  $\gamma_d$  with increasing Hamming distance  $d$

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



Epistasis measure  $\gamma_d$  with increasing Hamming distance  $d$

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



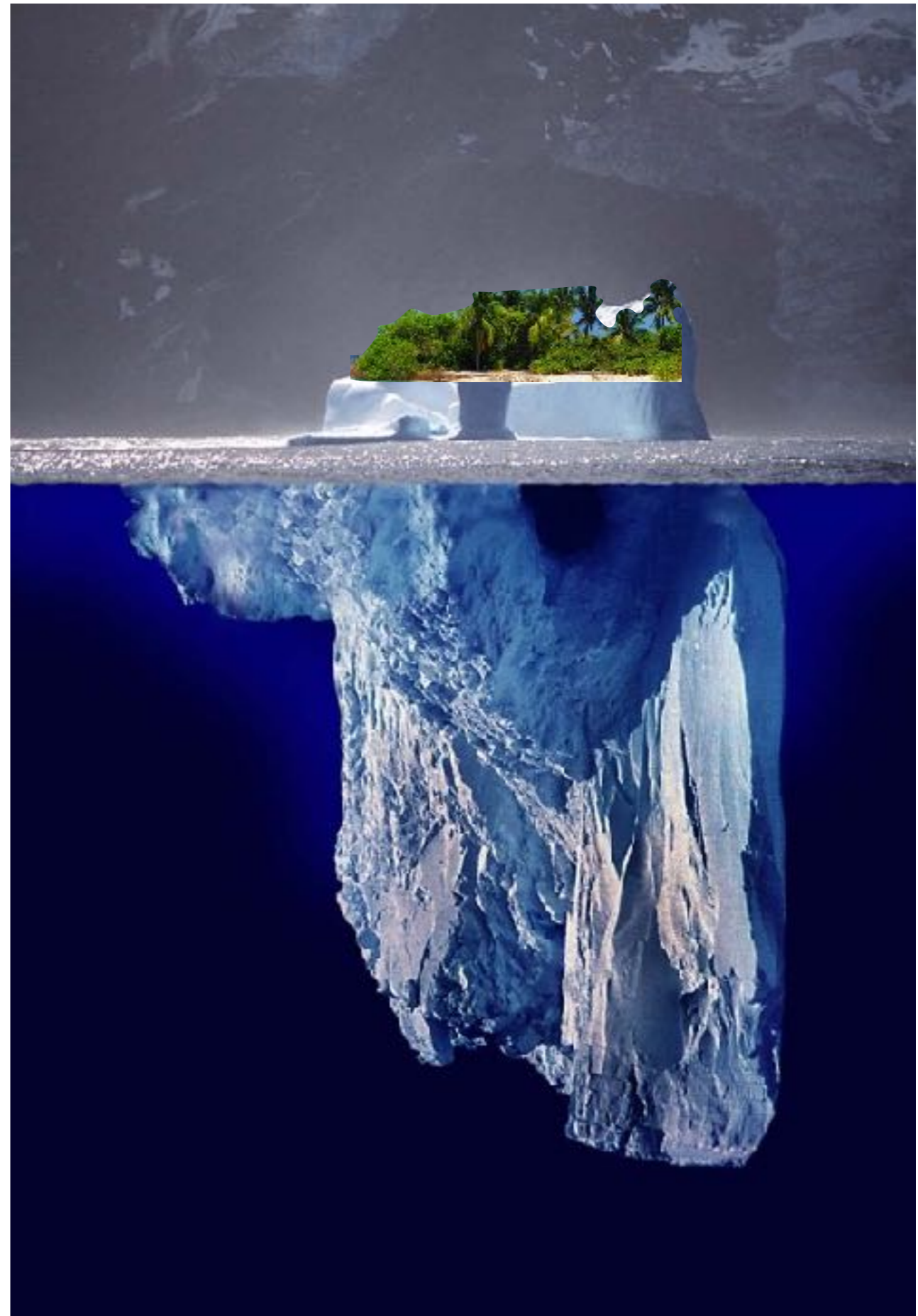
Epistasis measure  $\gamma_d$  with increasing Hamming distance  $d$

# 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...



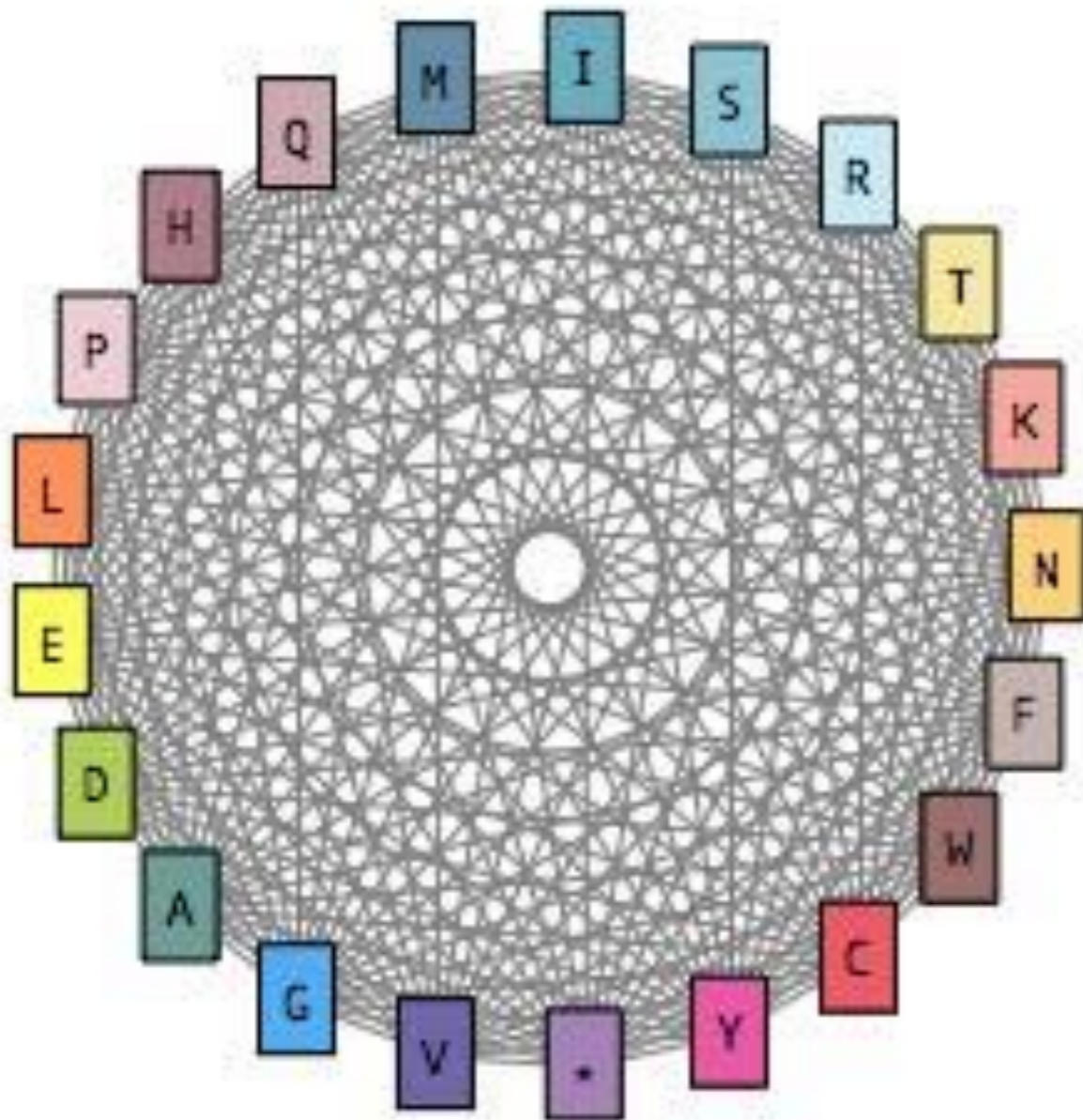
# FUTURE ENDEAVORS

---

- identify/develop informative measures and (mechanistic) models to characterize empirical fitness landscapes
- understand how fitness landscapes change across environments and genetic backgrounds
- study the role of epistasis and fitness landscapes across levels of organization and time scales

# THE FITNESS LANDSCAPE OF SYNONYMOUS MUTATIONS

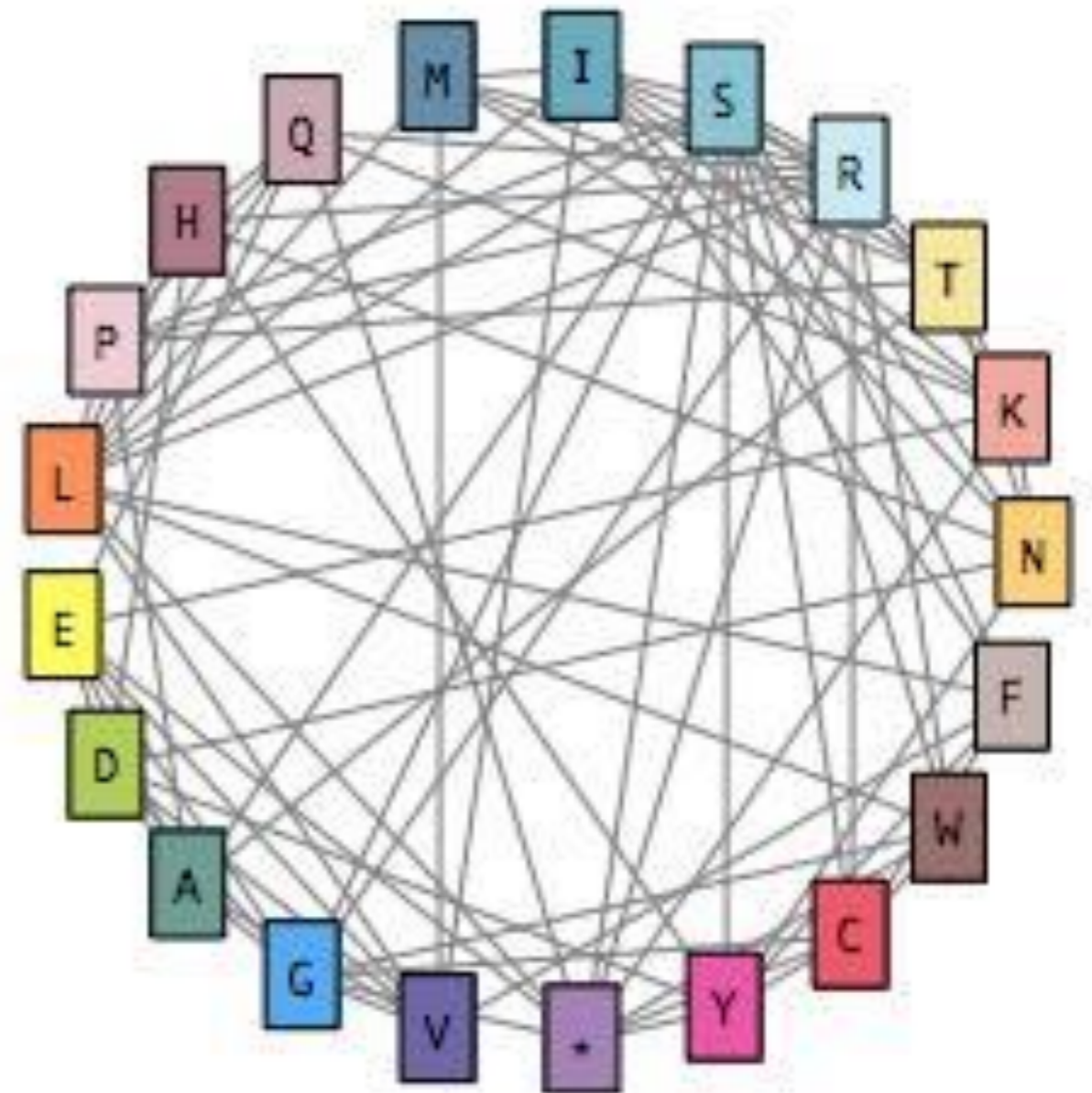
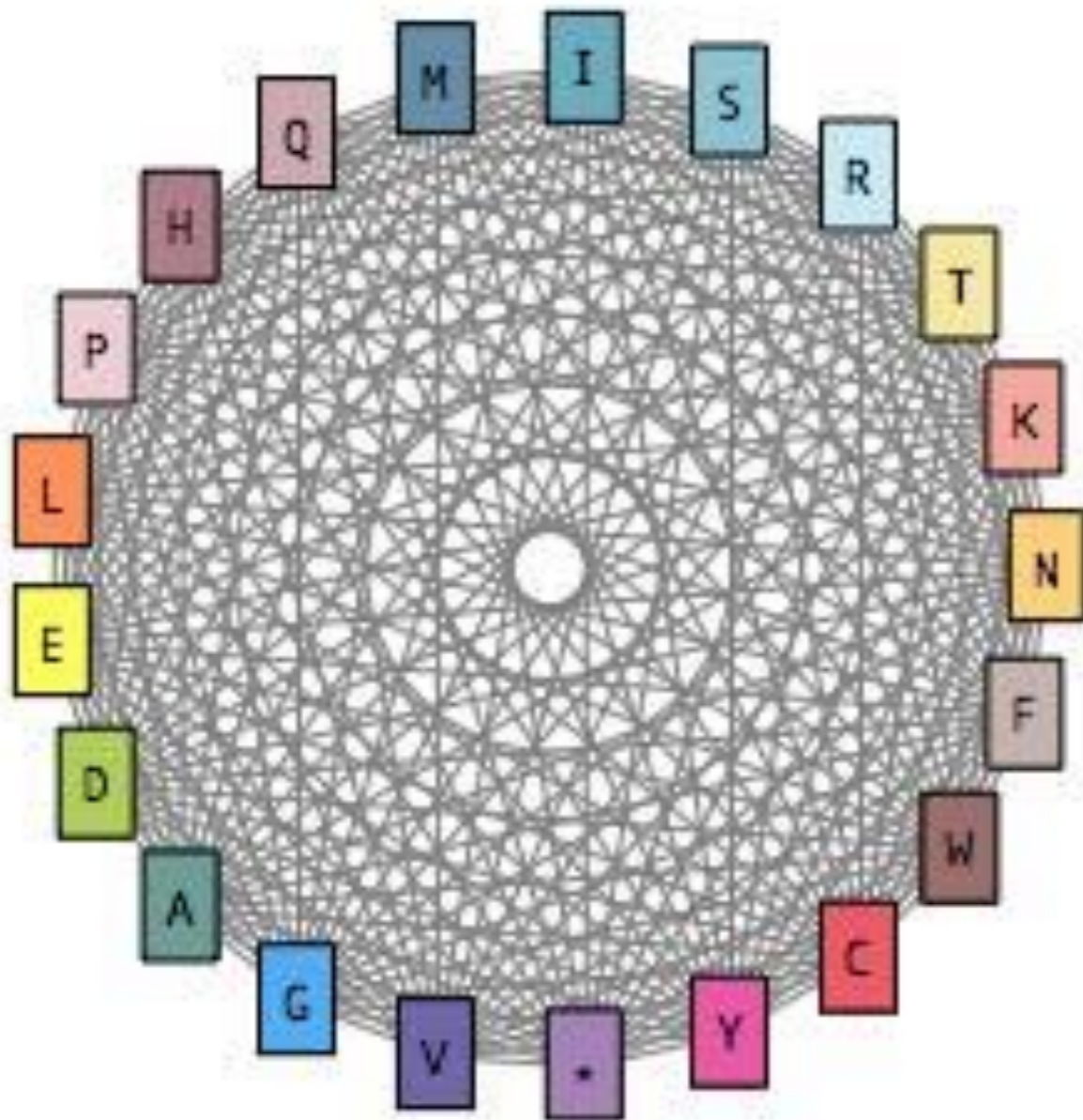
---





# THE FITNESS LANDSCAPE OF SYNONYMOUS MUTATIONS

---





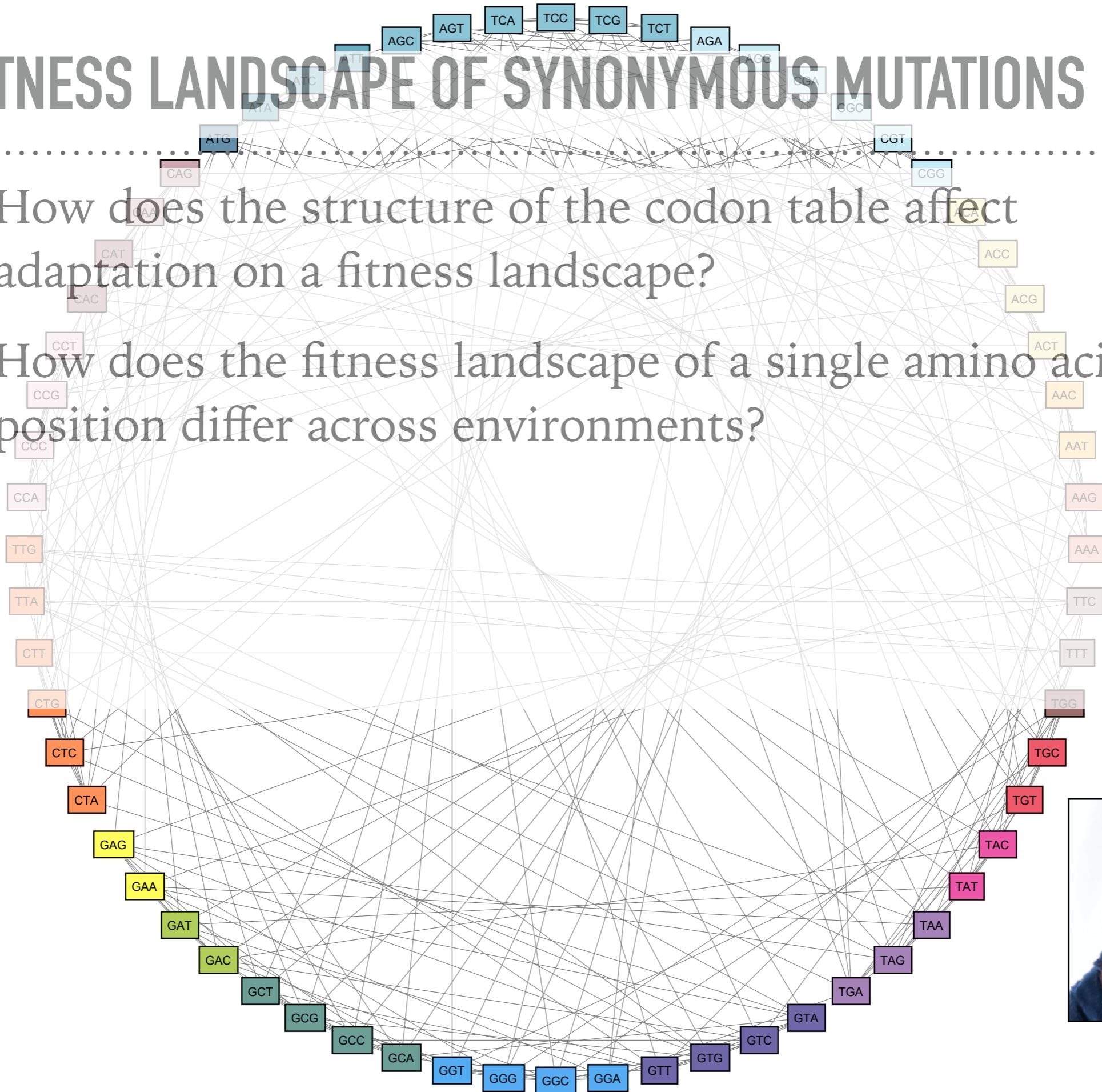






# THE FITNESS LANDSCAPE OF SYNONYMOUS MUTATIONS

- How does the structure of the codon table affect adaptation on a fitness landscape?
- How does the fitness landscape of a single amino acid position differ across environments?



Inês Fragata

# THE FITNESS LANDSCAPE OF SYNONYMOUS MUTATIONS

- How does the structure of the codon table affect adaptation on a fitness landscape?
- How does the fitness landscape of a single amino acid position differ across environments?
- How strong are the effects of synonymous mutations, and do they matter?



Inês Fragata



# THE FITNESS LANDSCAPE OF SYNONYMOUS MUTATIONS

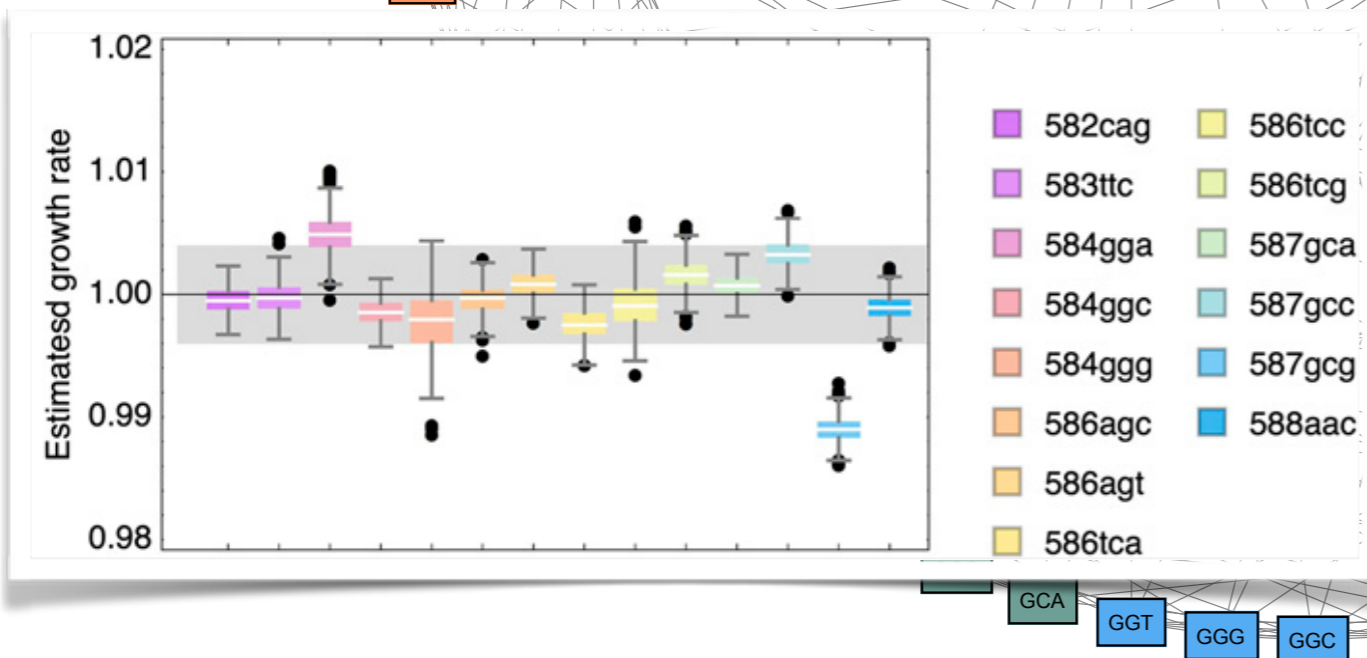
- How does the structure of the codon table affect adaptation on a fitness landscape?
- How does the fitness landscape of a single amino acid position differ across environments?
- How strong are the effects of synonymous mutations, and do they matter?
- Which factors can explain synonymous effects?



Inês Fragata

# THE FITNESS LANDSCAPE OF SYNONYMOUS MUTATIONS

- How does the structure of the codon table affect adaptation on a fitness landscape?
- How does the fitness landscape of a single amino acid position differ across environments?
- How strong are the effects of synonymous mutations, and do they matter?
- Which factors can explain synonymous effects?

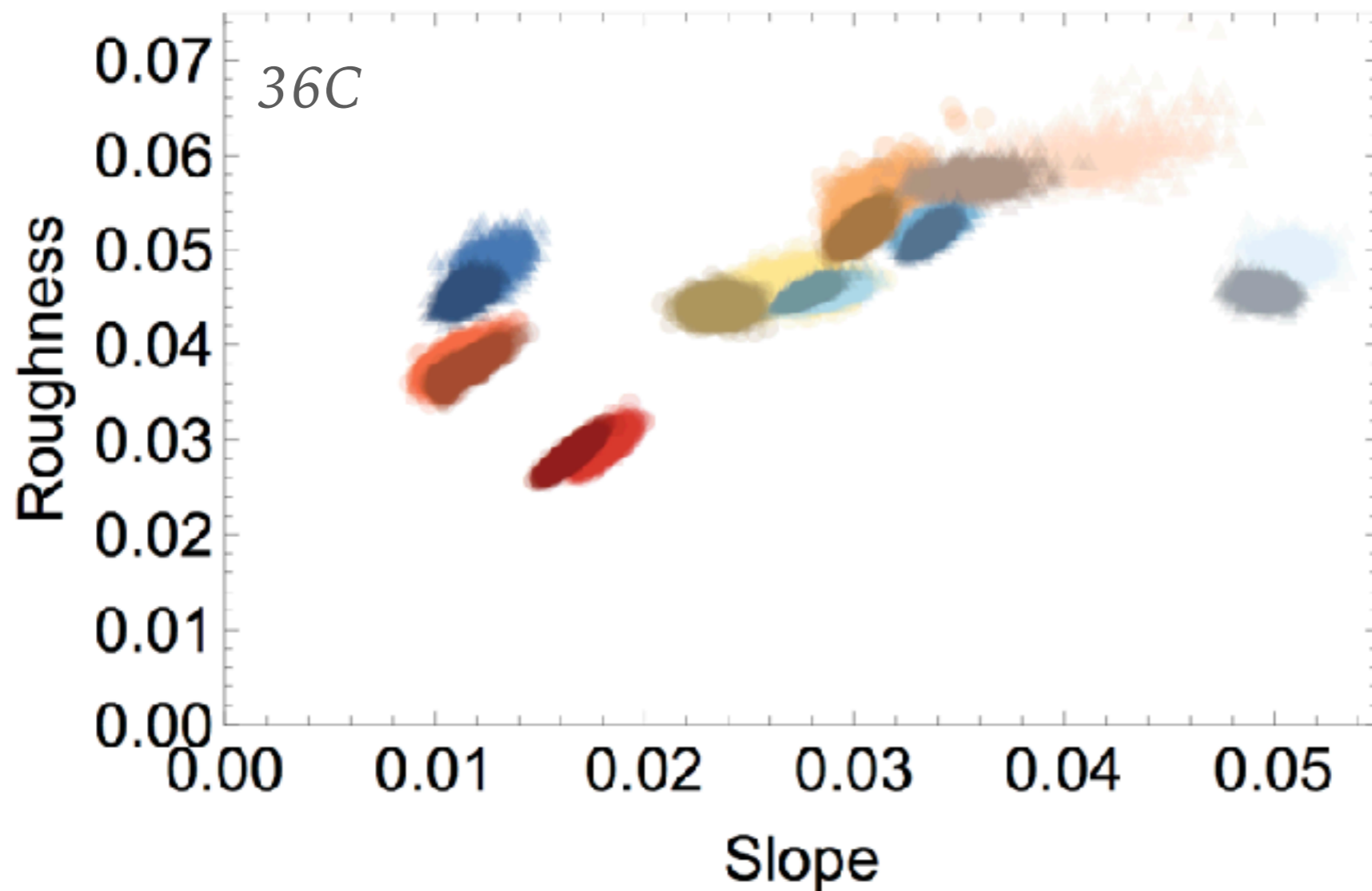


Inês Fragata



# THE FITNESS LANDSCAPE OF SYNONYMOUS MUTATIONS

- Compare landscape without synonymous effects with a landscape that has synonymous effects



Inês Fragata

## Evolutionary Dynamics @ IGC:

- How do populations adapt to challenging environments?  
E.g., how does drug resistance evolve?
- Which processes drive speciation & diversification?
- What is the role of interactions in evolution?



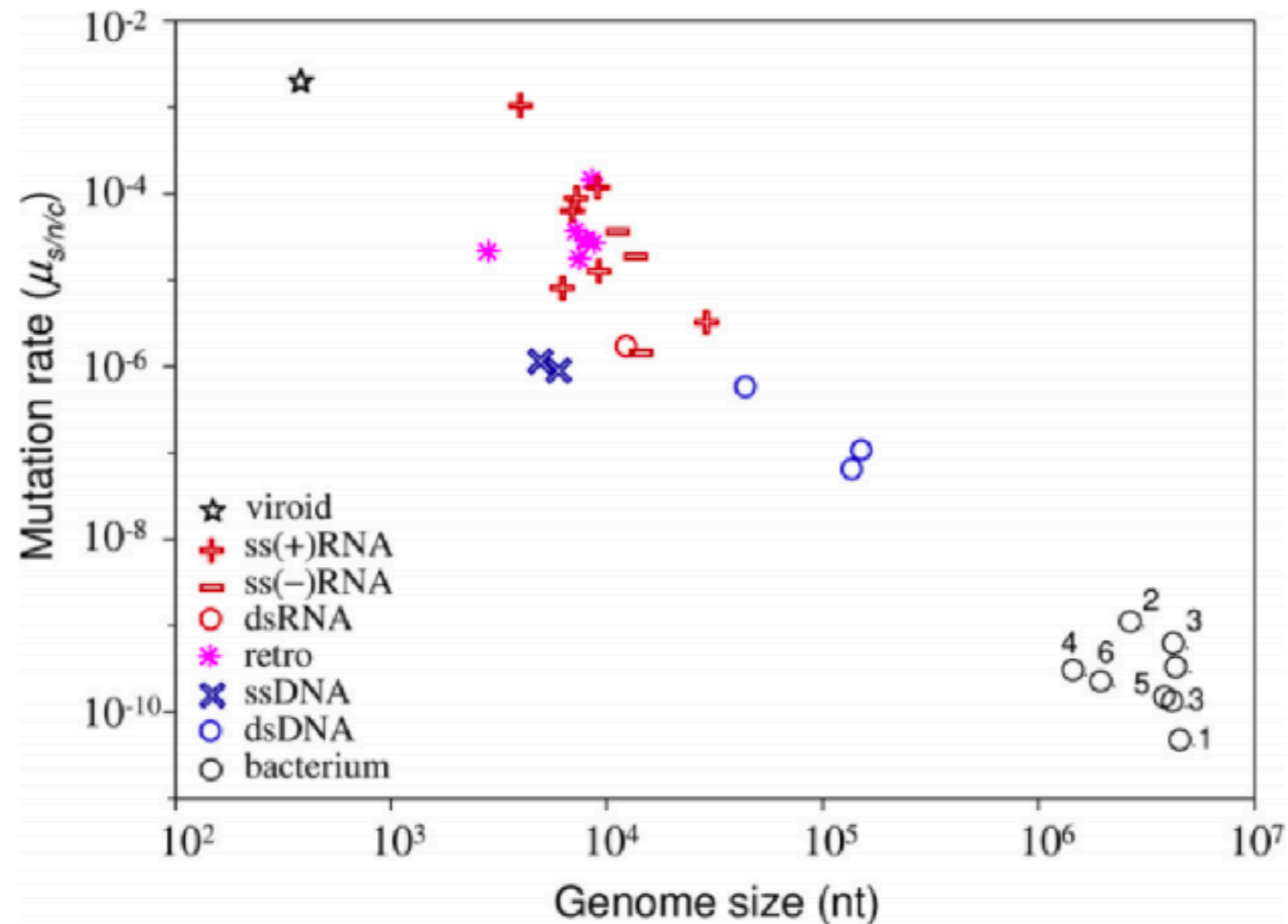
## What we do

- Study evolutionary processes using simple models
- Evaluate these models using empirical and simulated data
- Use modeling to inform experimental design *a priori*



# MUTAGENIC DRUGS AGAINST RNA VIRUS INFECTIONS

---



- RNA viruses have exceptionally large mutation rates.

# MUTATIONAL MELTDOWN/LETHAL MUTAGENESIS

---

Muller's ratchet: the step-wise loss of the fittest genotype due to accumulation of deleterious mutations in asexual populations



# MUTATIONAL MELTDOWN/LETHAL MUTAGENESIS

---

- a population goes extinct because it accumulates too many deleterious mutations (such that the absolute growth rate becomes  $< 1$ ) - this can be caused by mutation pressure or random genetic drift (or both)

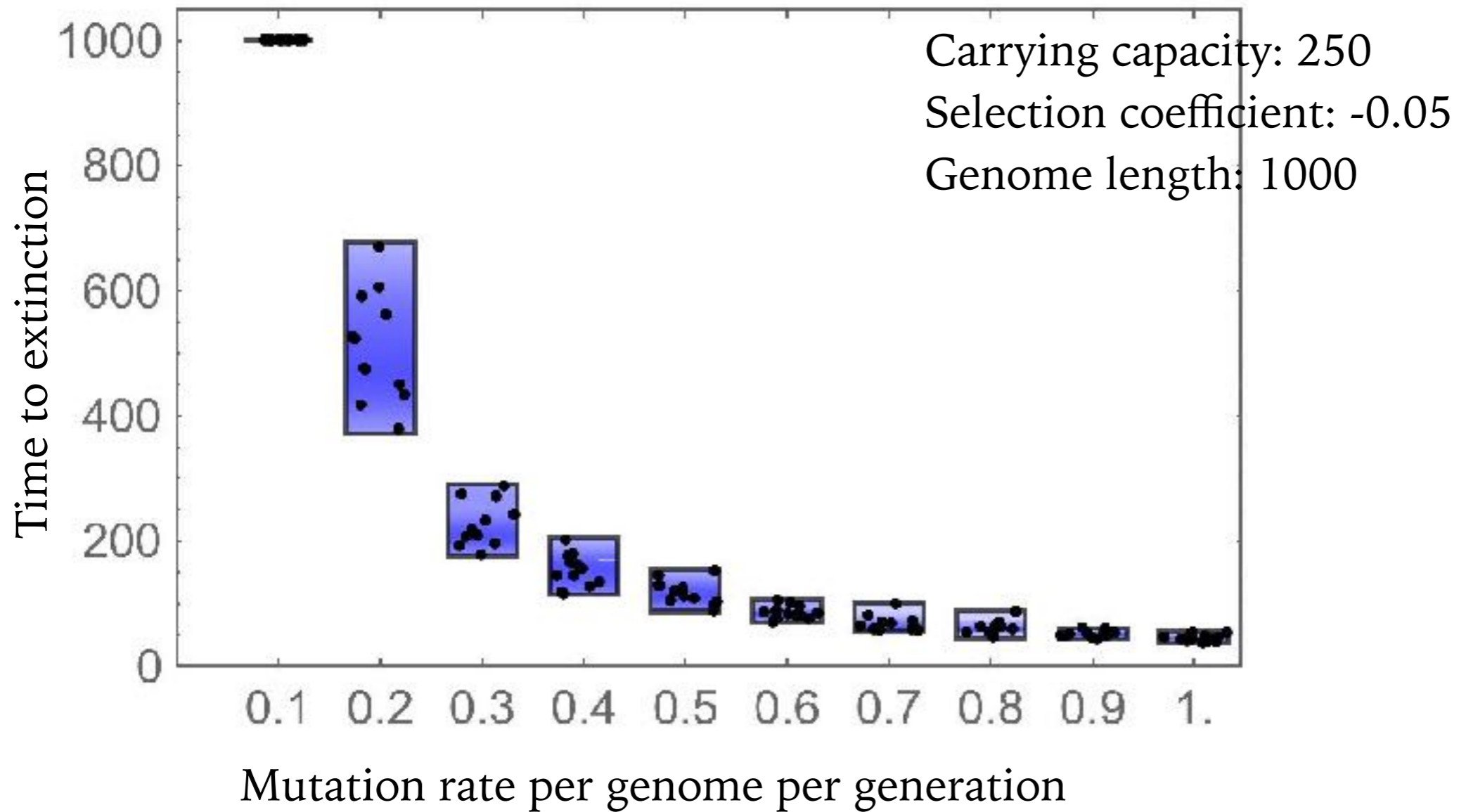
Muller's ratchet: the step-wise loss of the fittest genotype due to accumulation of deleterious mutations in asexual populations



# MUTAGENIC DRUGS AGAINST RNA VIRUS INFECTIONS

---

## Simulated example

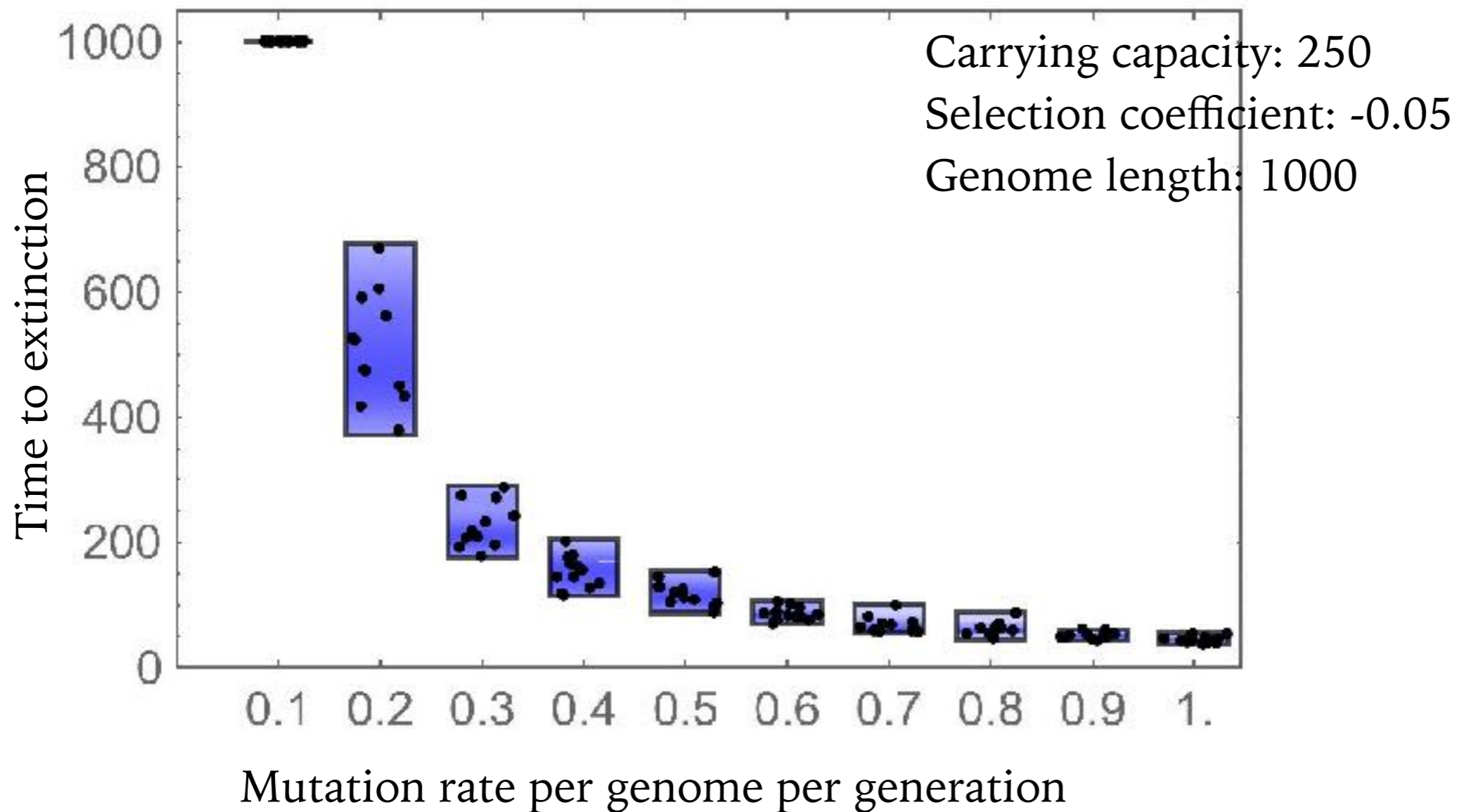




# MUTAGENIC DRUGS AGAINST RNA VIRUS INFECTIONS

---

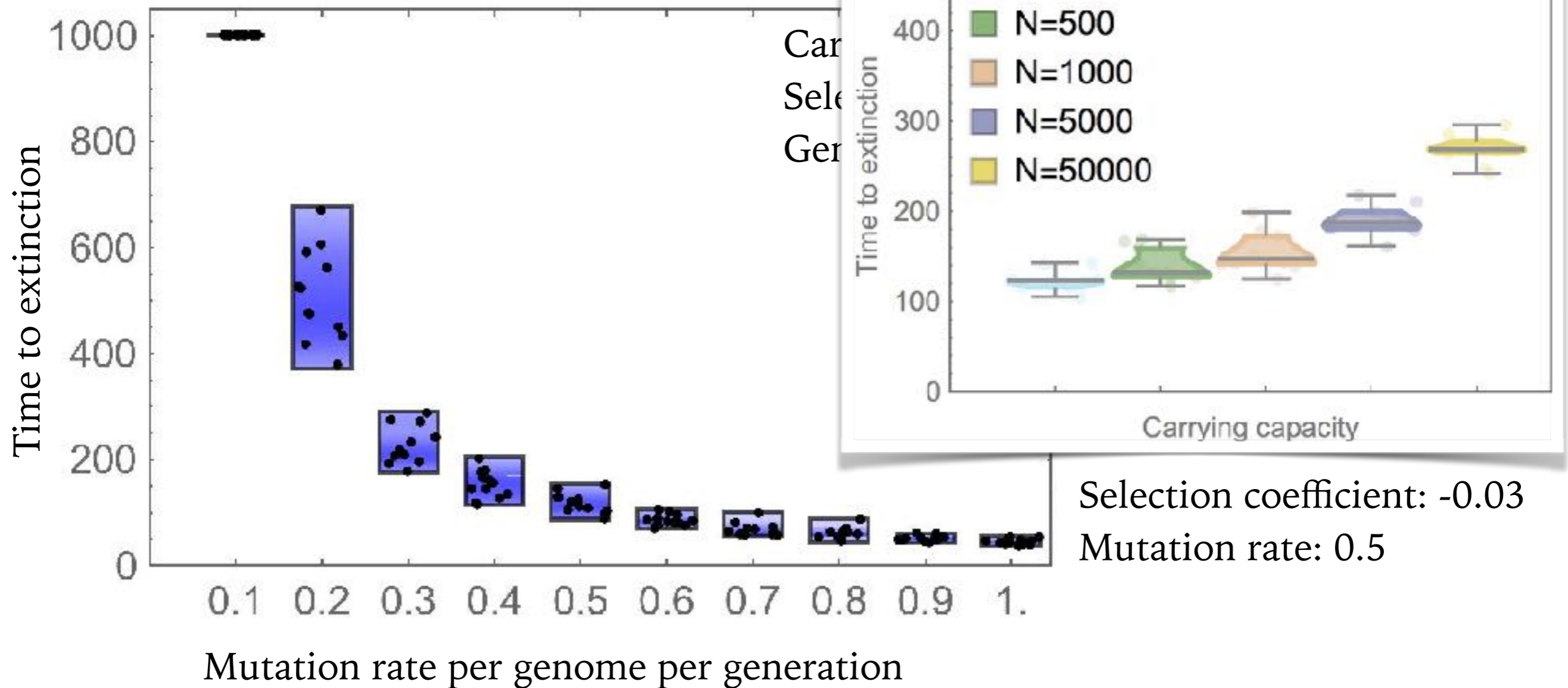
## Simulated example



- Mutagenic drug favipiravir approved for use against influenza in Japan and discussed as promising candidate drug against various RNA viruses.

# MUTAGENIC DRUGS AGAINST RNA VIRUS INFECTIONS

## Simulated example

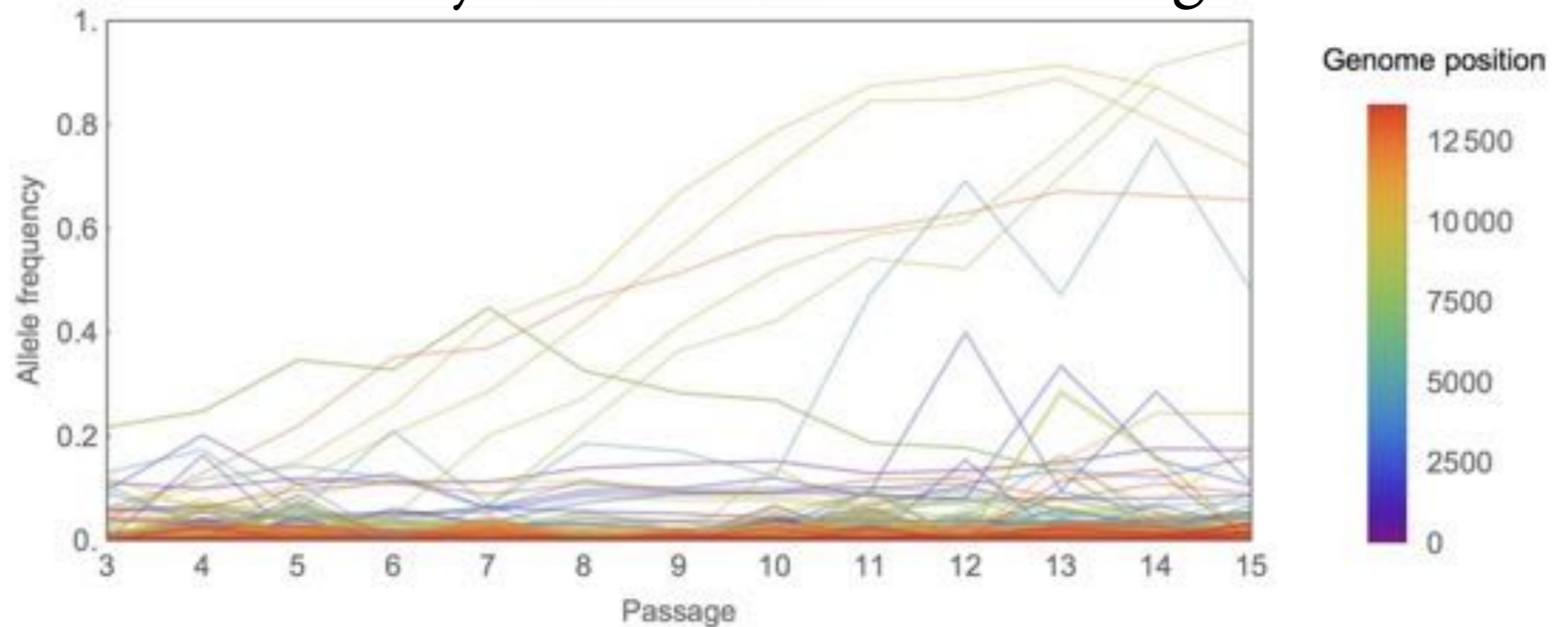


- Mutagenic drug favipiravir approved for use against influenza in Japan and discussed as promising candidate drug against various RNA viruses.

# INFLUENZA A LABORATORY EVOLUTION UNDER MUTAGENIC DRUG TREATMENT

---

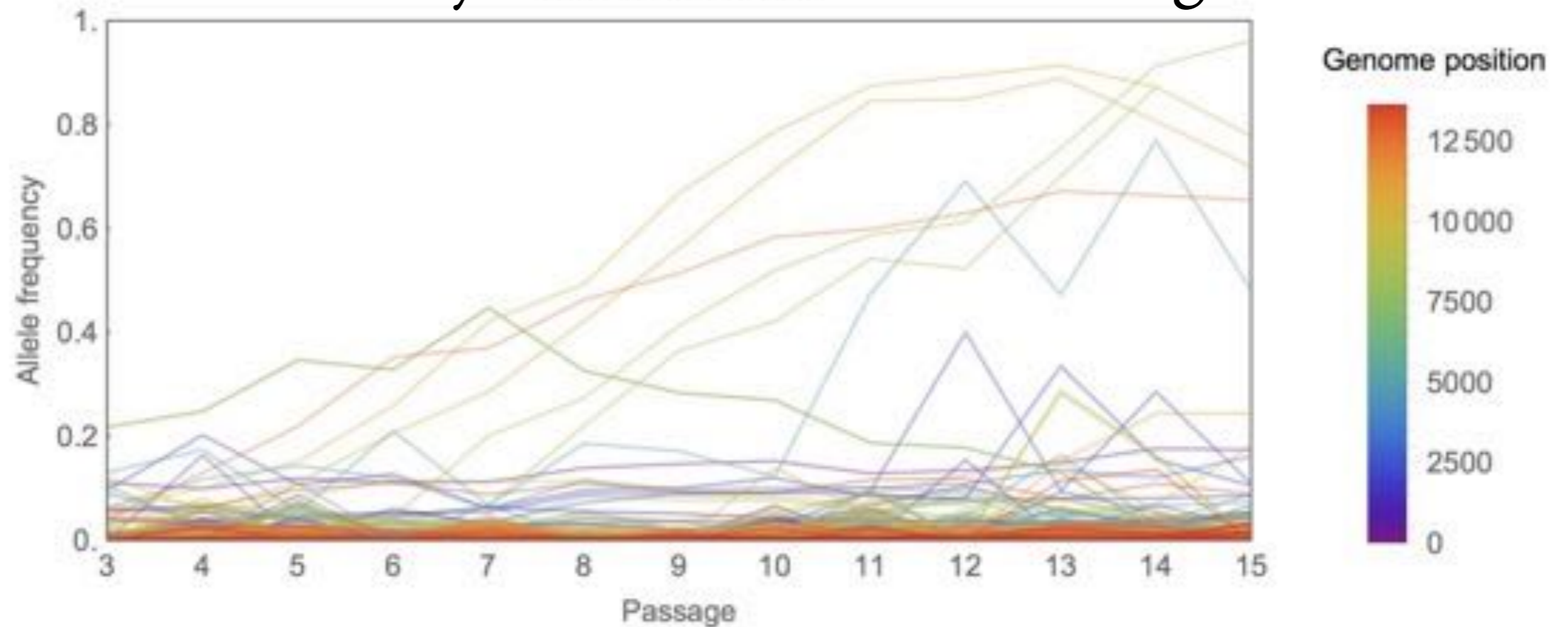
## Laboratory evolution without drug



# INFLUENZA A LABORATORY EVOLUTION UNDER MUTAGENIC DRUG TREATMENT

---

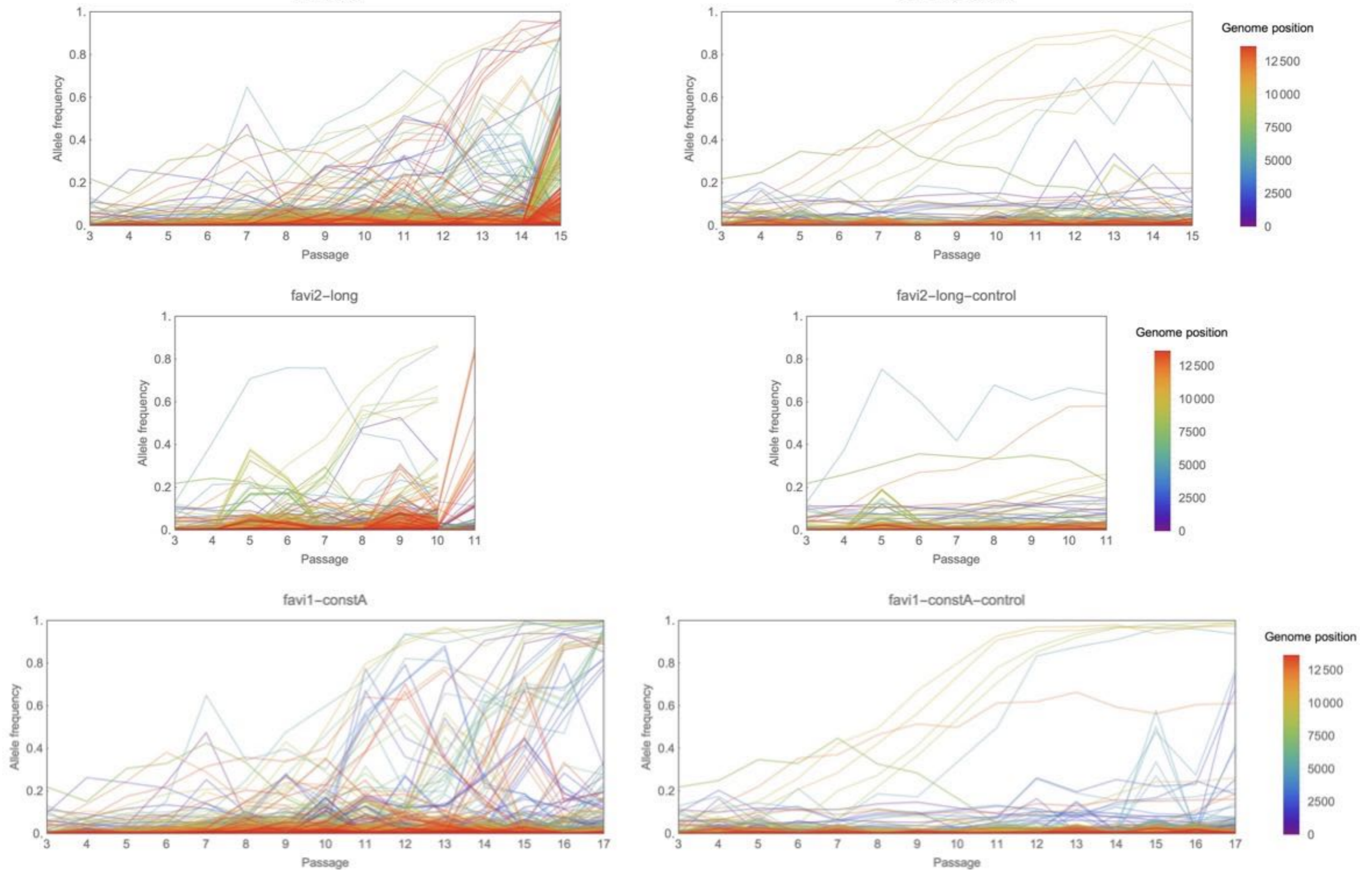
## Laboratory evolution without drug



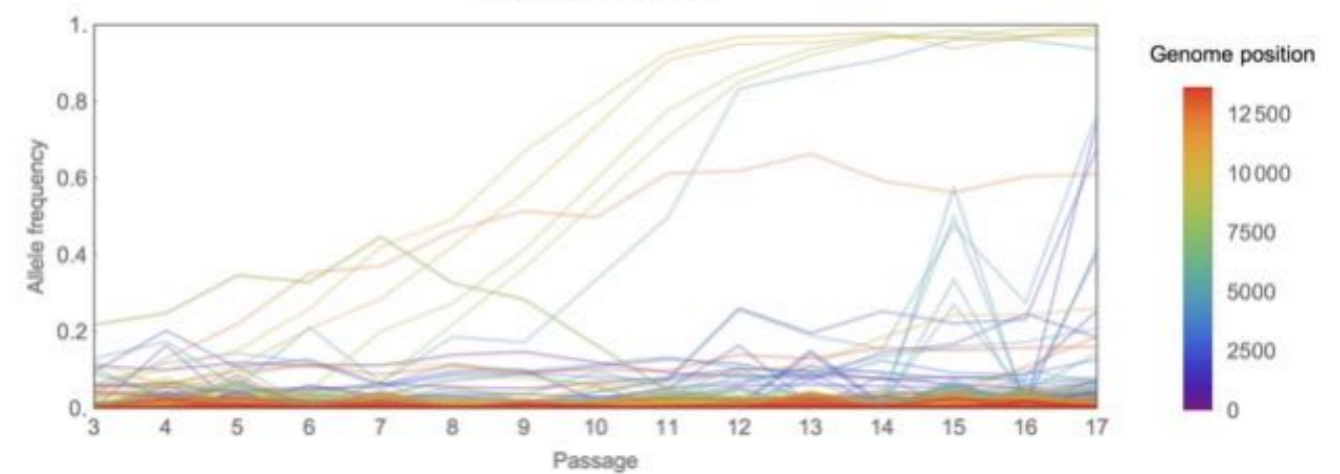
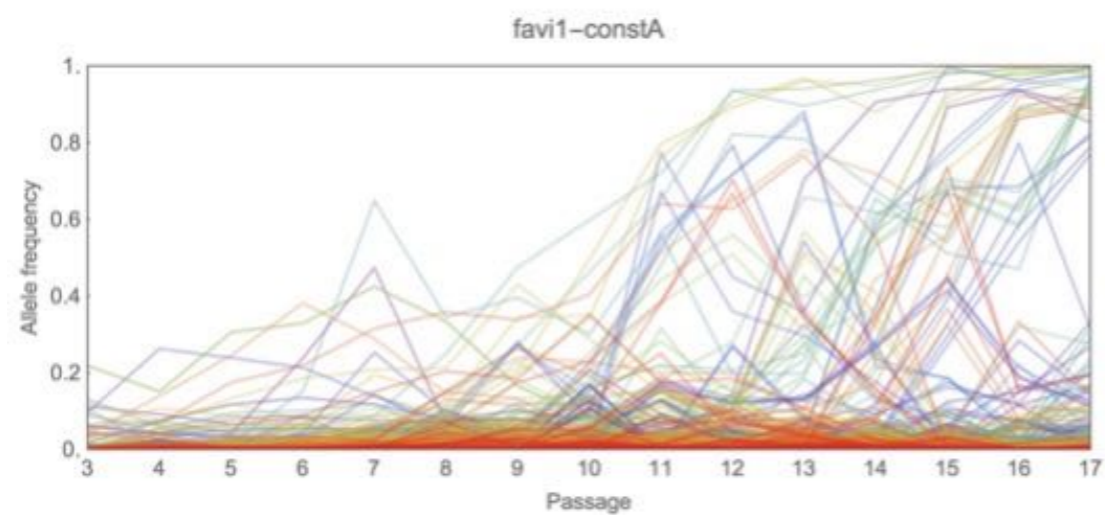
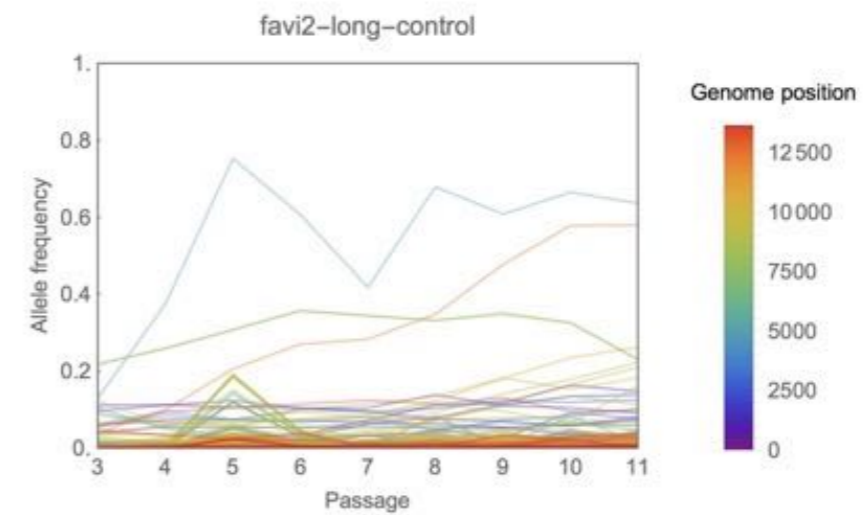
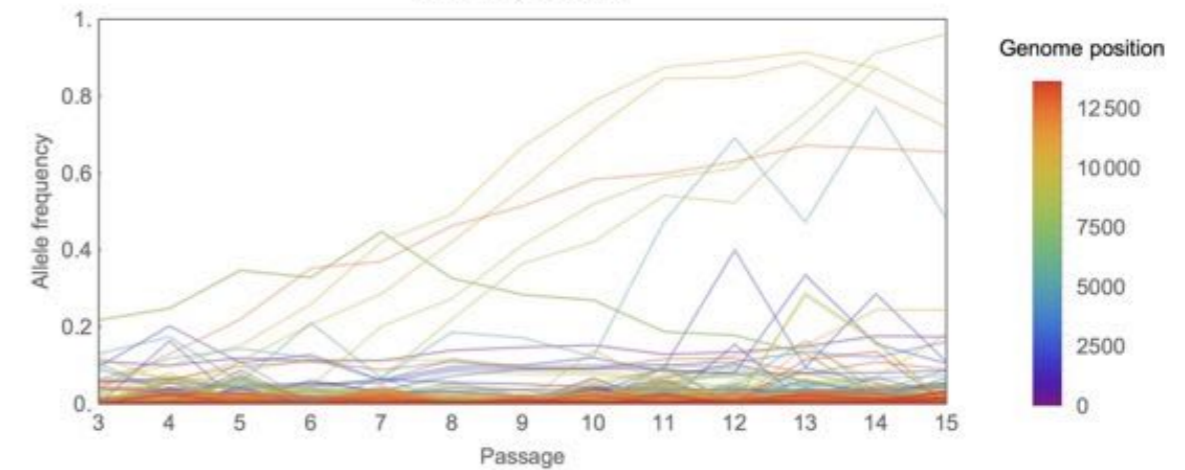
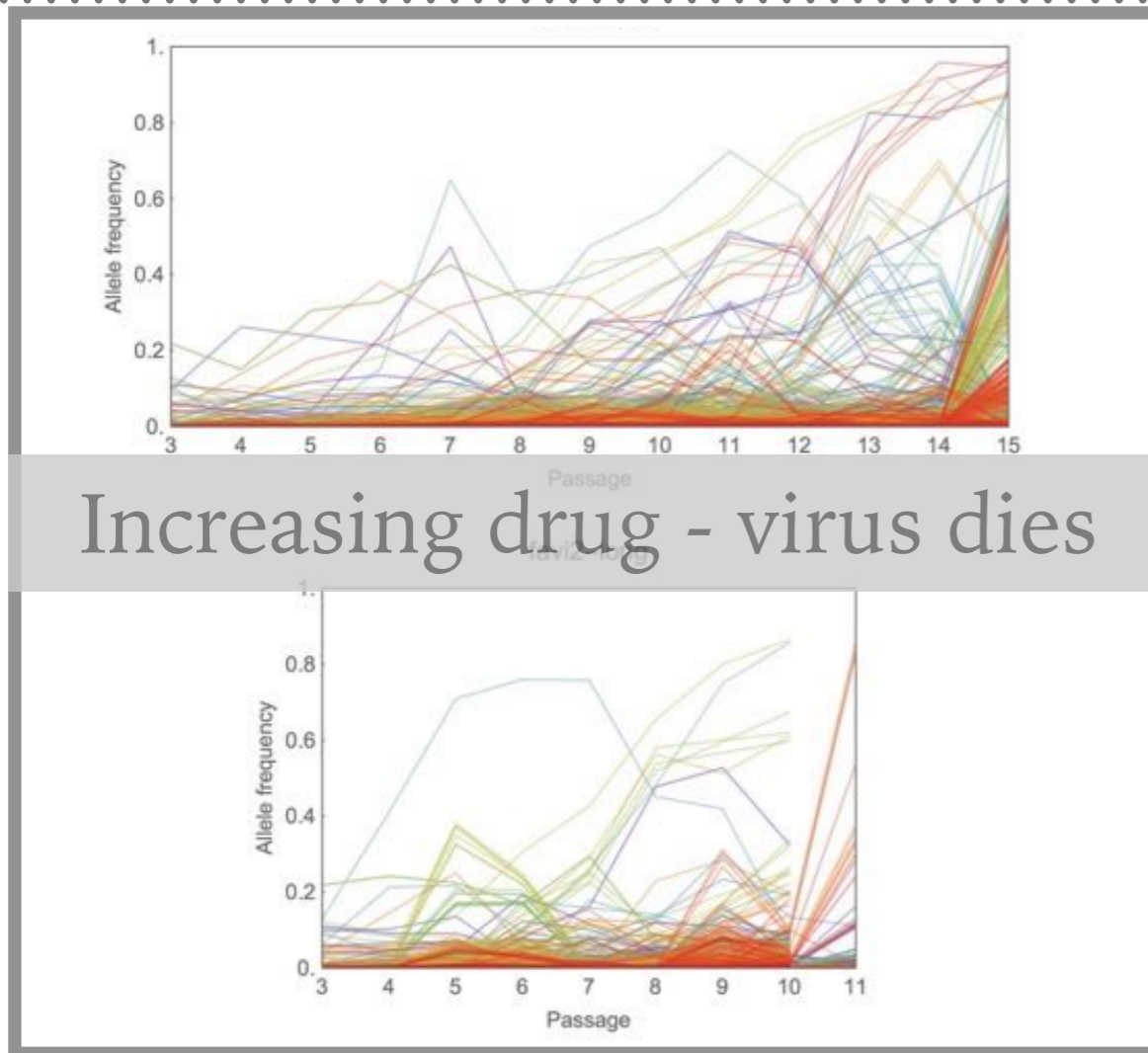
- Using laboratory evolution combined with time-serial genome-wide sequencing we can quantify the evolutionary dynamics of influenza virus.



# INFLUENZA A LABORATORY EVOLUTION UNDER MUTAGENIC DRUG TREATMENT

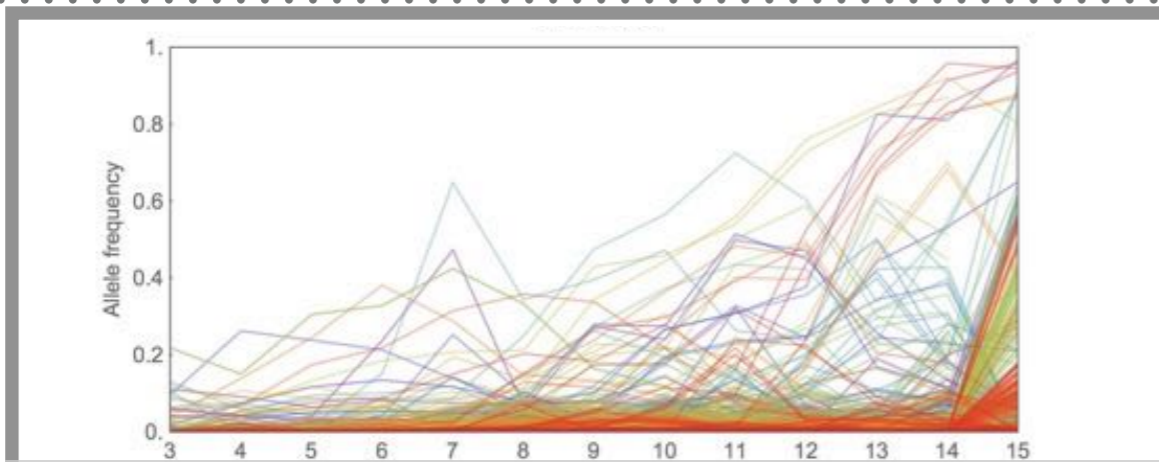


# INFLUENZA A LABORATORY EVOLUTION UNDER MUTAGENIC DRUG TREATMENT

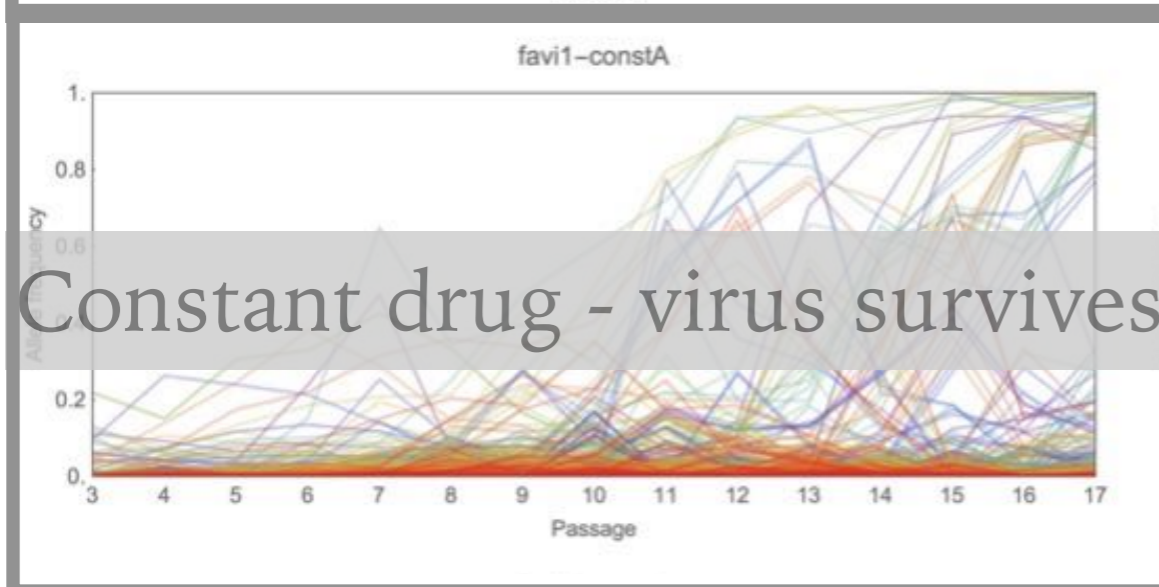
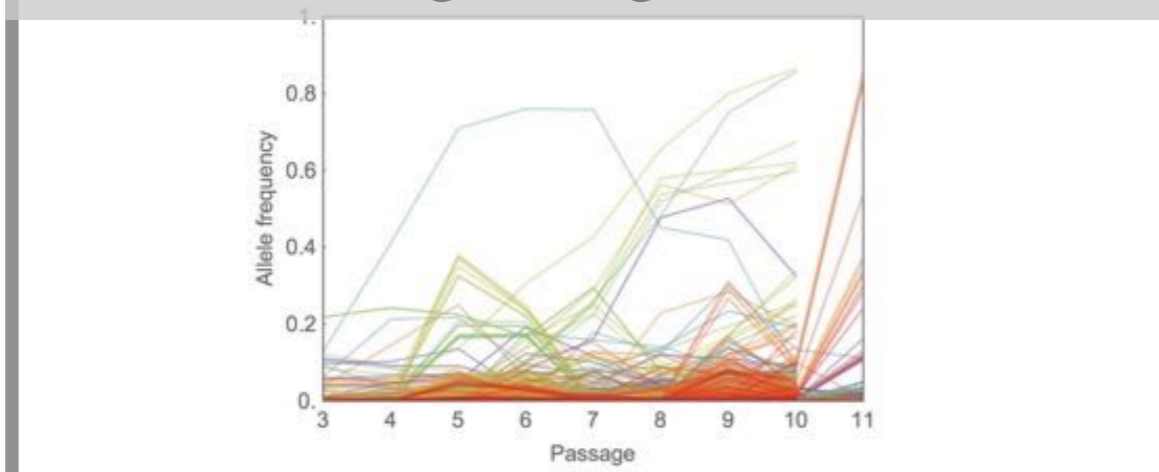




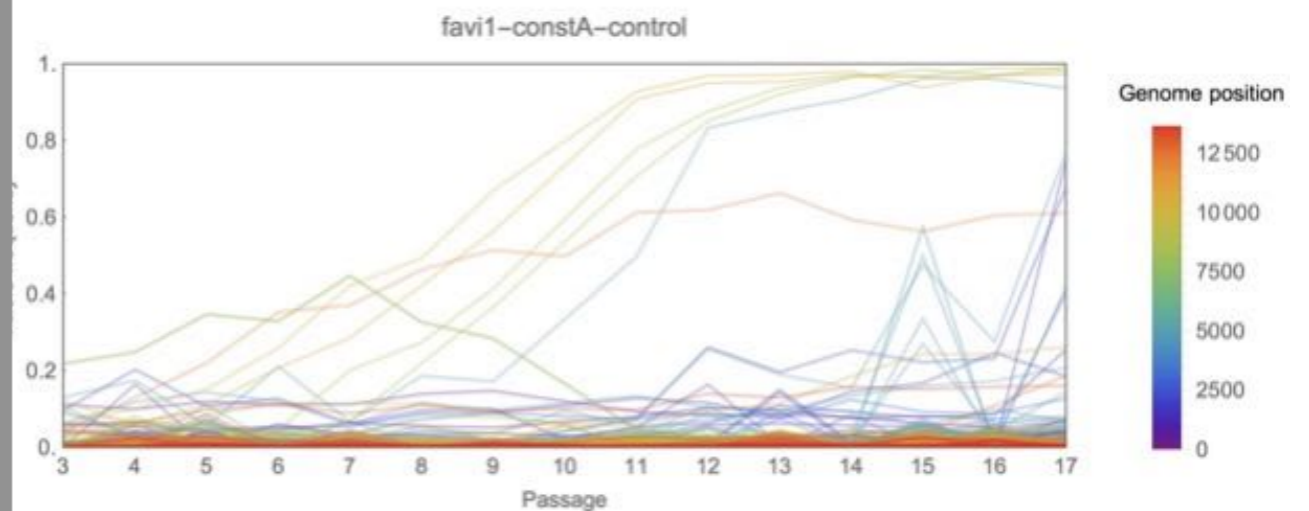
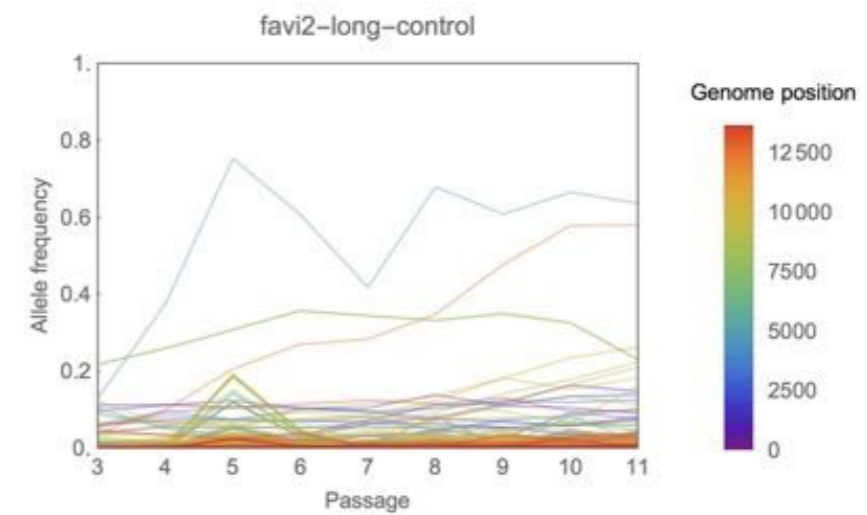
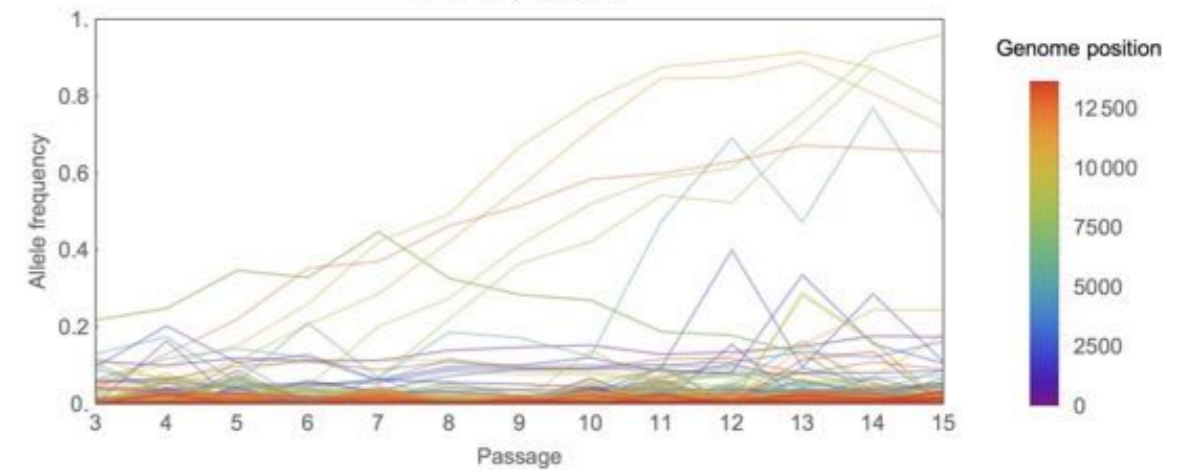
# INFLUENZA A LABORATORY EVOLUTION UNDER MUTAGENIC DRUG TREATMENT



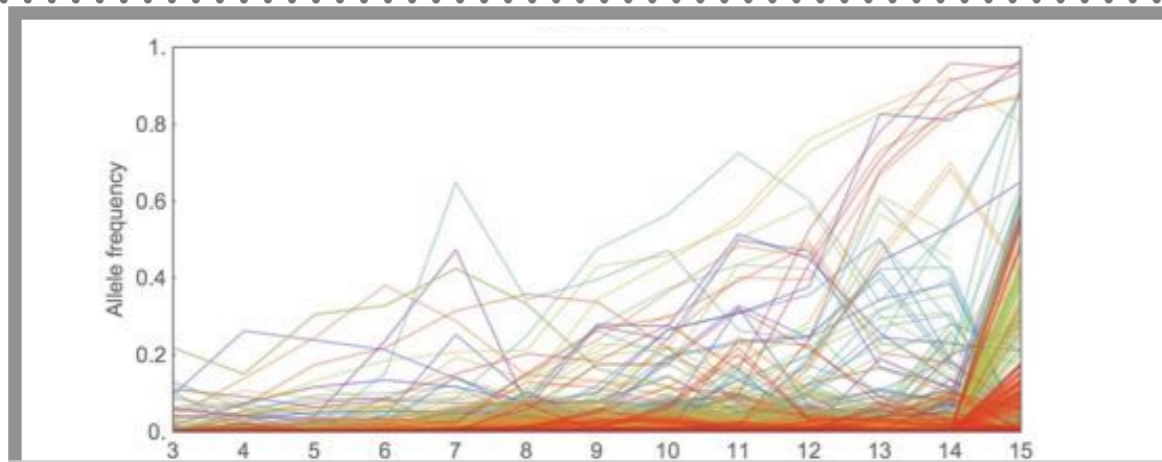
Increasing drug - virus dies



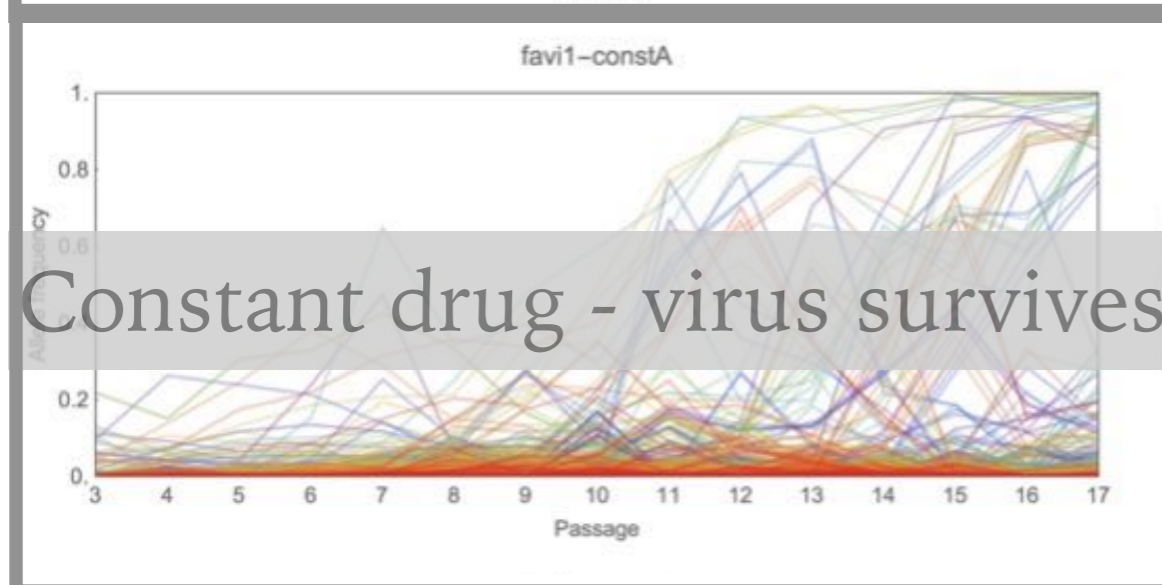
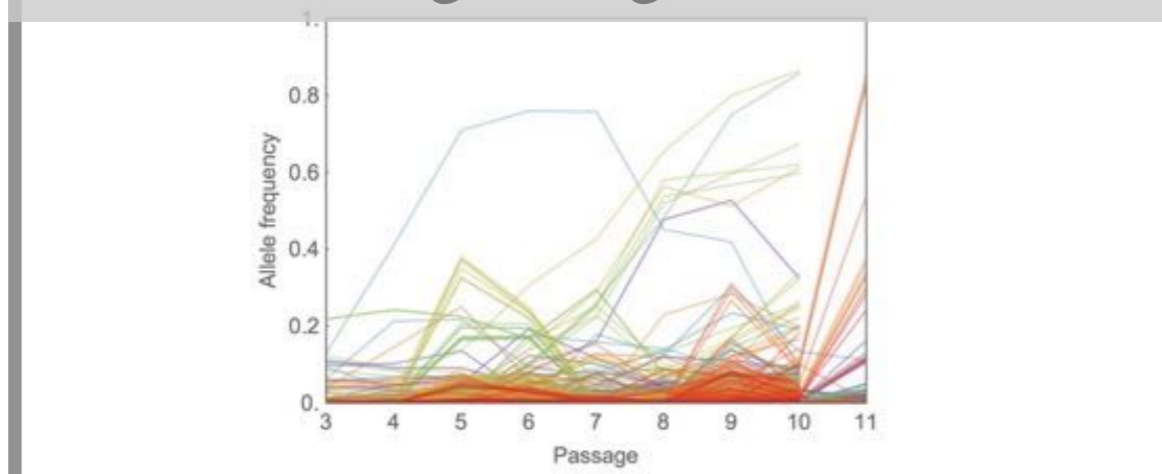
Constant drug - virus survives



# INFLUENZA A LABORATORY EVOLUTION UNDER MUTAGENIC DRUG TREATMENT



Increasing drug - virus dies



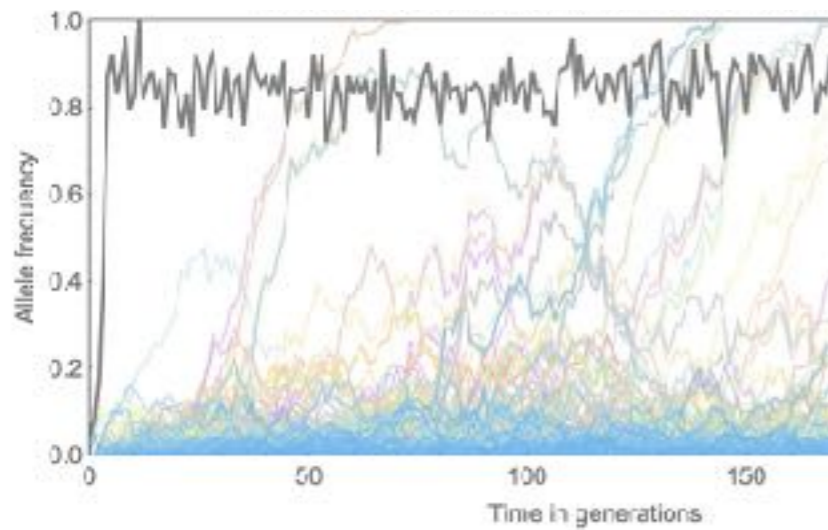
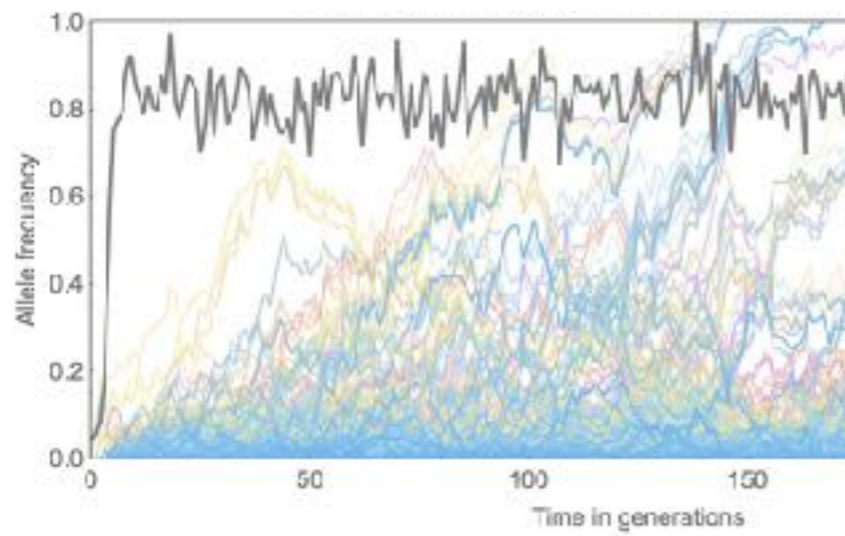
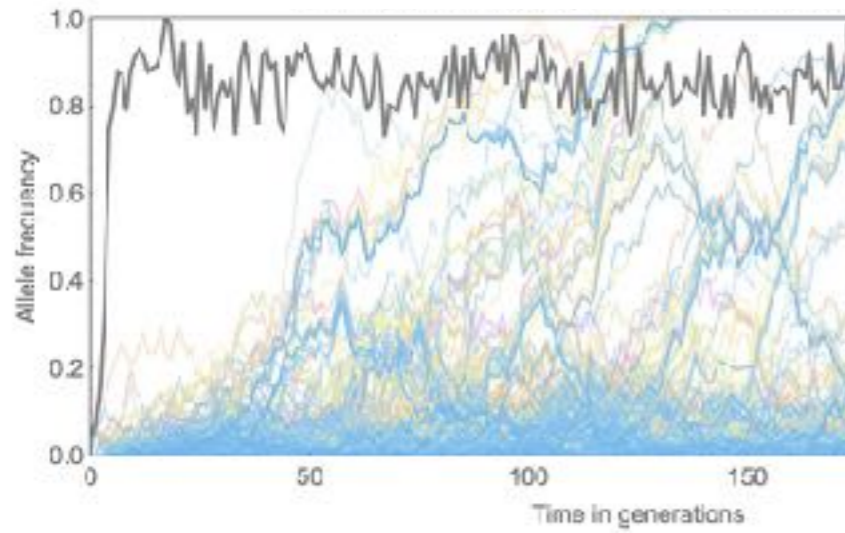
Constant drug - virus survives



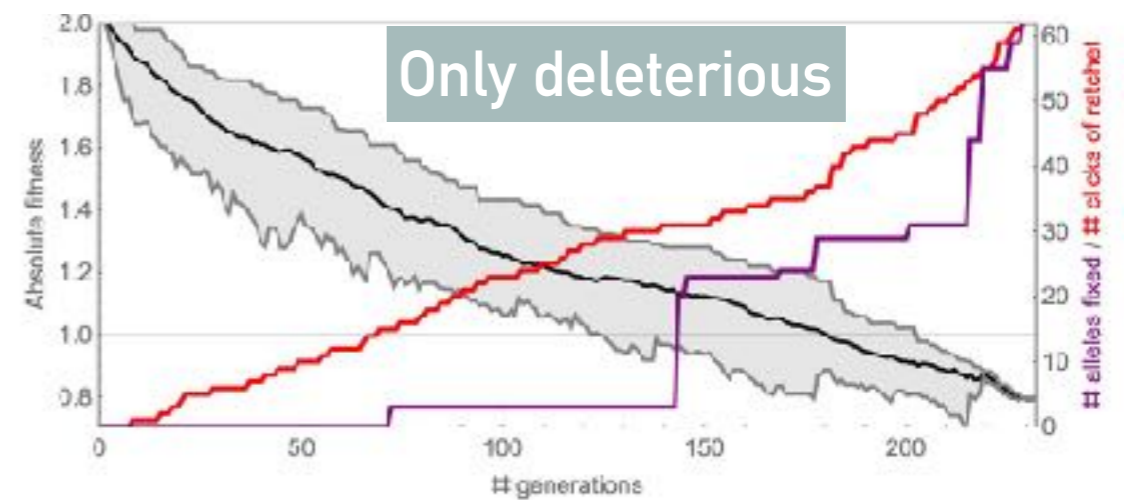
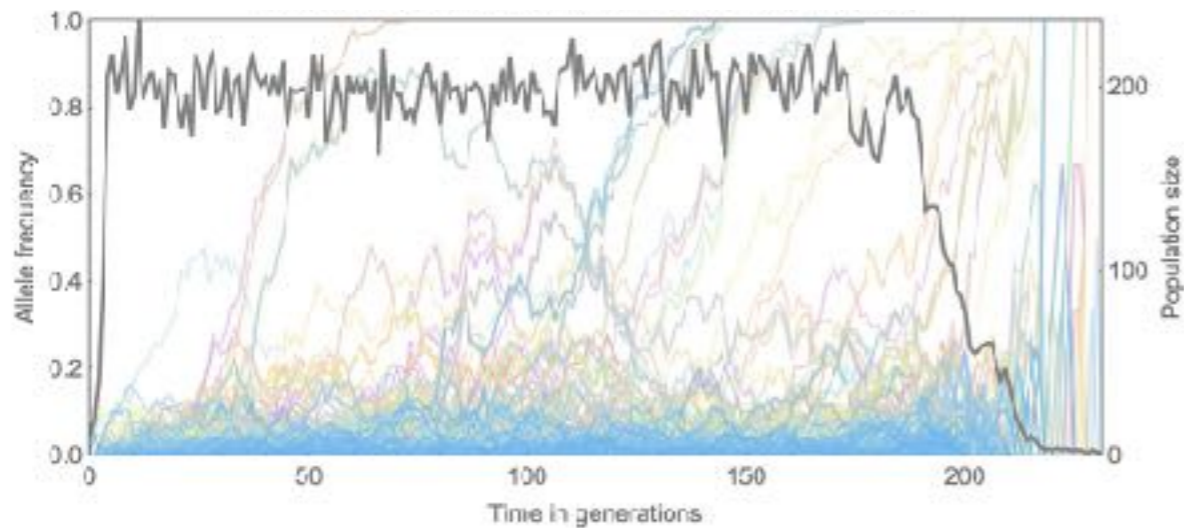
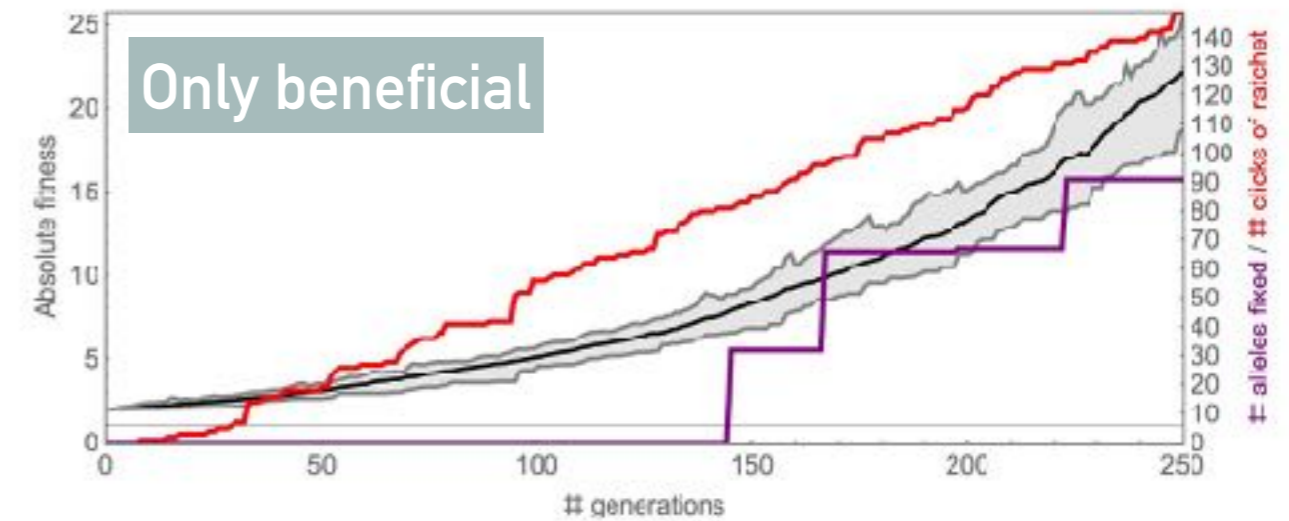
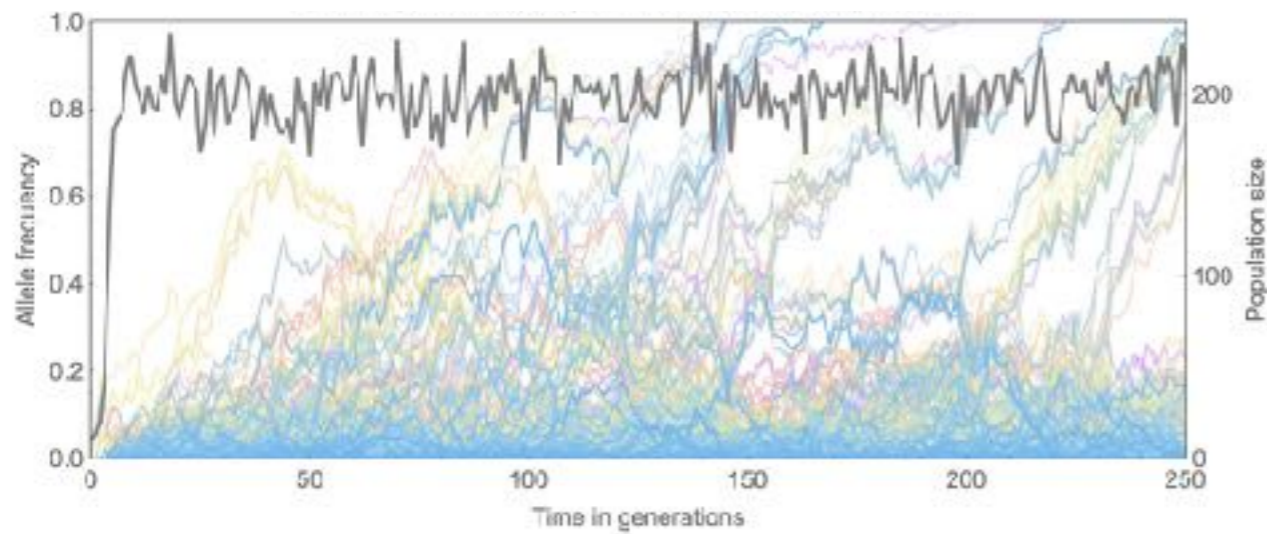
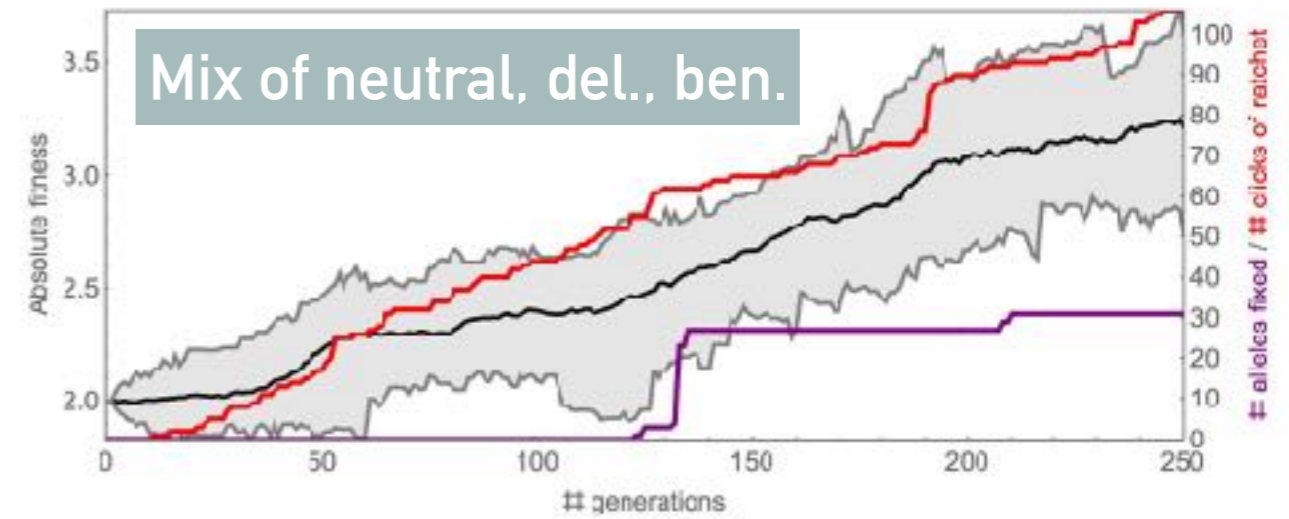
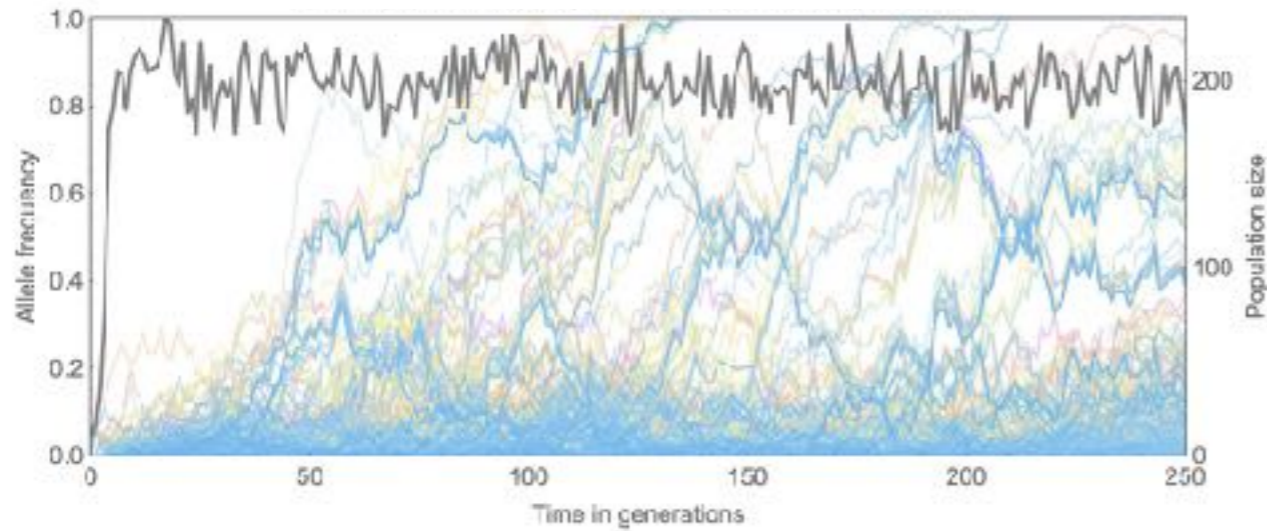
- Favipiravir effective at high/ increasing concentration, but indication of population recovery (i.e., adaptation to drug treatment) at constant (intermediate) concentrations.
- What is the signature of different adaptation mechanisms? How good are our methods for detection of candidate loci? How informative are allele frequencies?



# SIMULATE EVOLUTION OF A CLONAL POPULATION WITH HIGH MUTATION RATES



# SIMULATE EVOLUTION OF A CLONAL POPULATION WITH HIGH MUTATION RATES



# CAN VIRUSES ADAPT TO MUTAGENIC DRUG TREATMENTS?

---

- "Adaptation" in this context means survival/persistence of a pathogen or other health threat despite exposure to drug, immune system, novel environments, etc.



# CAN VIRUSES ADAPT TO MUTAGENIC DRUG TREATMENTS?

---

- "Adaptation" in this context means survival/persistence of a pathogen or other health threat despite exposure to drug, immune system, novel environments, etc.
- By which mechanisms can viruses escape from mutagenic drug treatment? Can we detect the signatures of such adaptation? What are the dangers of mutagenic drugs?



# CAN VIRUSES ADAPT TO MUTAGENIC DRUG TREATMENTS?

---

- "Adaptation" in this context means survival/persistence of a pathogen or other health threat despite exposure to drug, immune system, novel environments, etc.
- By which mechanisms can viruses escape from mutagenic drug treatment? Can we detect the signatures of such adaptation? What are the dangers of mutagenic drugs?
- An example of **evolutionary rescue**: an adaptation spreads in a population that is otherwise doomed to extinction due to a change in the environment

# POTENTIAL MECHANISMS OF RESCUE FROM INCREASED MUTATION RATES



# POTENTIAL MECHANISMS OF RESCUE FROM INCREASED MUTATION RATES

---

- "traditional" beneficial mutations that increase growth rate:  
only a temporary fix because they will not stop the ratchet

# POTENTIAL MECHANISMS OF RESCUE FROM INCREASED MUTATION RATES

---

- "traditional" beneficial mutations that increase growth rate: only a temporary fix because they will not stop the ratchet
- a mutation rate modifier that reduces the mutation rate below the critical level: **evolution of drug resistance**



# POTENTIAL MECHANISMS OF RESCUE FROM INCREASED MUTATION RATES

---

- "traditional" beneficial mutations that increase growth rate: only a temporary fix because they will not stop the ratchet
- a mutation rate modifier that reduces the mutation rate below the critical level: **evolution of drug resistance**
- a modifier of the fitness distribution, i.e. a mutation that changes mutational effects genome-wide: **evolution of drug tolerance**

# POTENTIAL MECHANISMS OF RESCUE FROM INCREASED MUTATION RATES

---

- "traditional" beneficial mutations that increase growth rate: only a temporary fix because they will not stop the ratchet
- a mutation rate modifier that reduces the mutation rate below the critical level: **evolution of drug resistance**
- a modifier of the fitness distribution, i.e. a mutation that changes mutational effects genome-wide: **evolution of drug tolerance**

Important to note: both weaker and stronger effects of (deleterious) mutations can slow down the ratchet (Gordo & Charlesworth 2000)

# POTENTIAL MECHANISMS OF RESCUE FROM INCREASED MUTATION RATES

---

- "traditional" beneficial mutations that increase growth rate: only a temporary fix because they will not stop the ratchet
- a mutation rate modifier that reduces the mutation rate below the critical level: **evolution of drug resistance**
- a modifier of the fitness distribution, i.e. a mutation that changes mutational effects genome-wide: **evolution of drug tolerance**

Important to note: both weaker and stronger effects of (deleterious) mutations can slow down the ratchet (Gordo & Charlesworth 2000)

Tolerance could be the most dangerous mechanism of adaptation to mutagenic drugs because it allows the virus to propagate at high mutation rates, which may allow rare/unseen/complex beneficial mutations to invade subsequently.

# TODAY'S QUESTIONS

---

- How does the availability of “traditional” beneficials prolong extinction times?
- When does a mutation rate modifier invade?
- In which conditions does a modifier of the distribution of fitness effects (DFE) invade?



# SIMULATION DETAILS

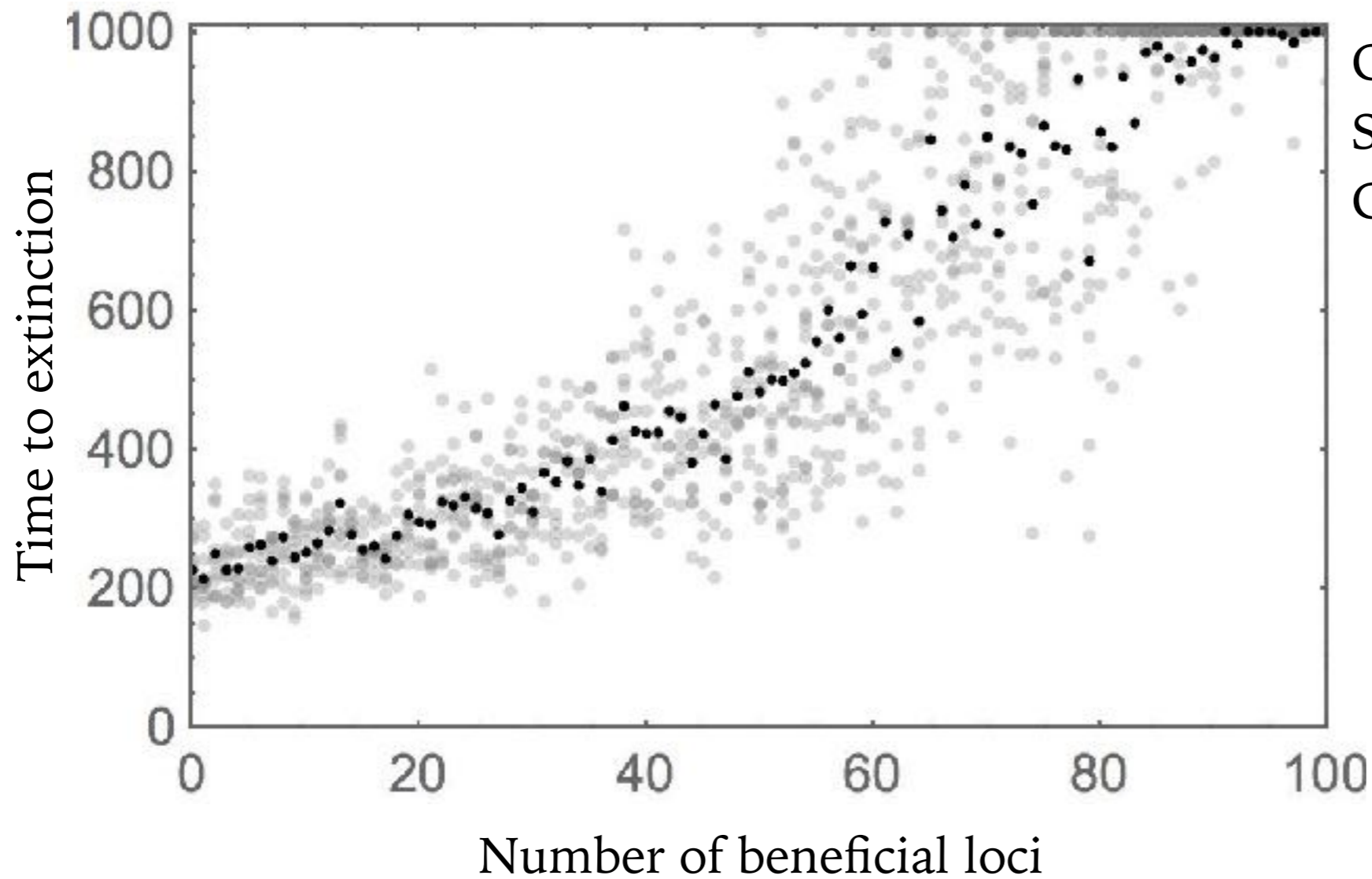
---

- Genome with  $L$  di-allelic loci [1000]
- Carrying capacity  $C$  of the clonal population [250], initial population size  $C_0$  [invasion size: 10]
- Initial absolute growth rate  $R$  [2]
- Arbitrary distribution of fitness effects [-0.05; multiplicative]
- Mutation rate  $\mu$  per genome per generation [0.3]
- Record haplotypes in each generation, stop if no extinction has occurred after 1000 generations (transmission/immune reaction)
- $l$  loci with “adaptive” mutations; either beneficial, mutation rate modifier, or DFE modifier

**Today: focus on extinction time & “rescue” probability**

# EXTINCTION TIMES WITH BENEFICIAL MUTATIONS

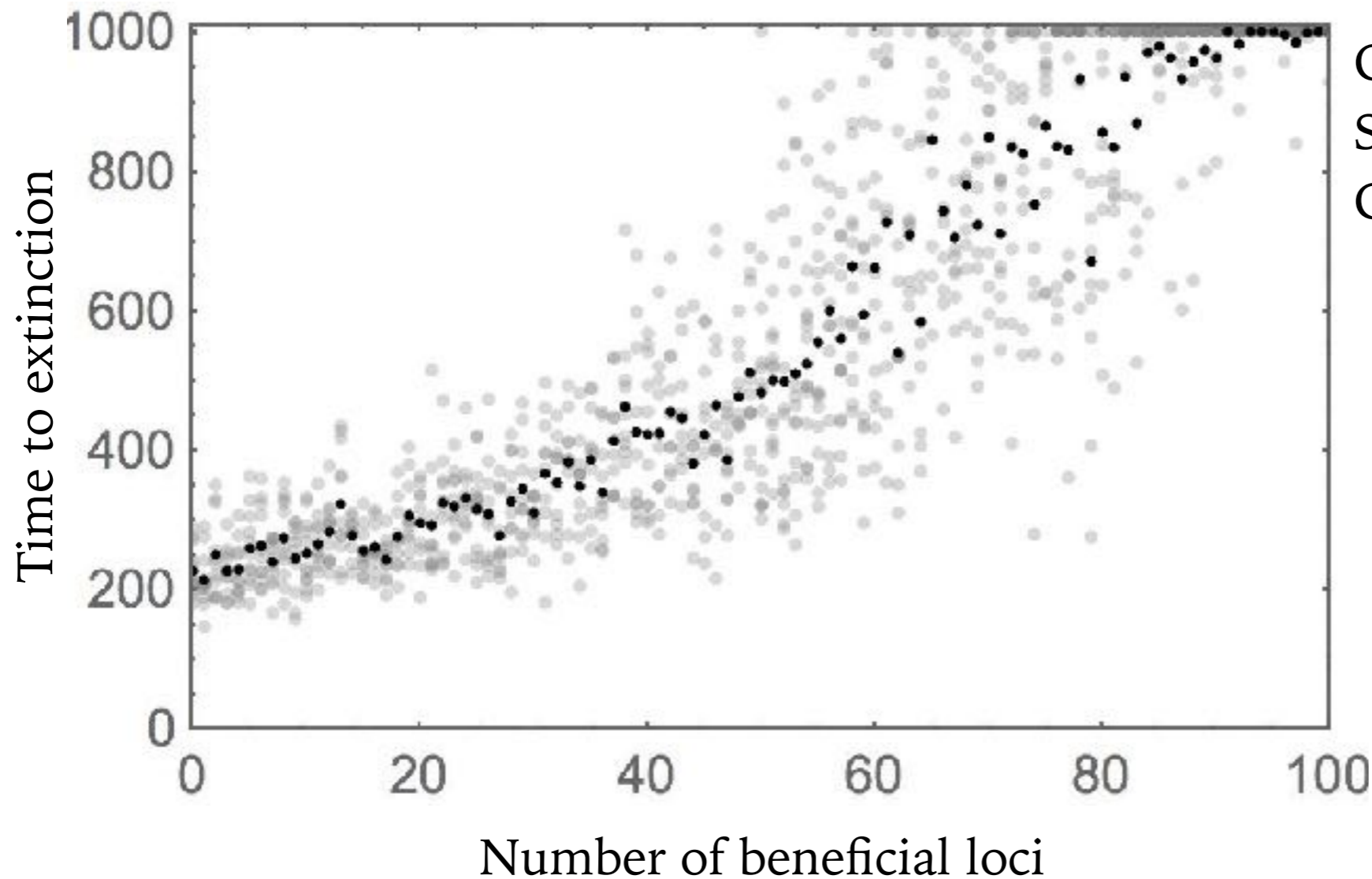
---



Carrying capacity: 250  
Selection coefficient: +/-0.05  
Genome length: 1000

# EXTINCTION TIMES WITH BENEFICIAL MUTATIONS

---

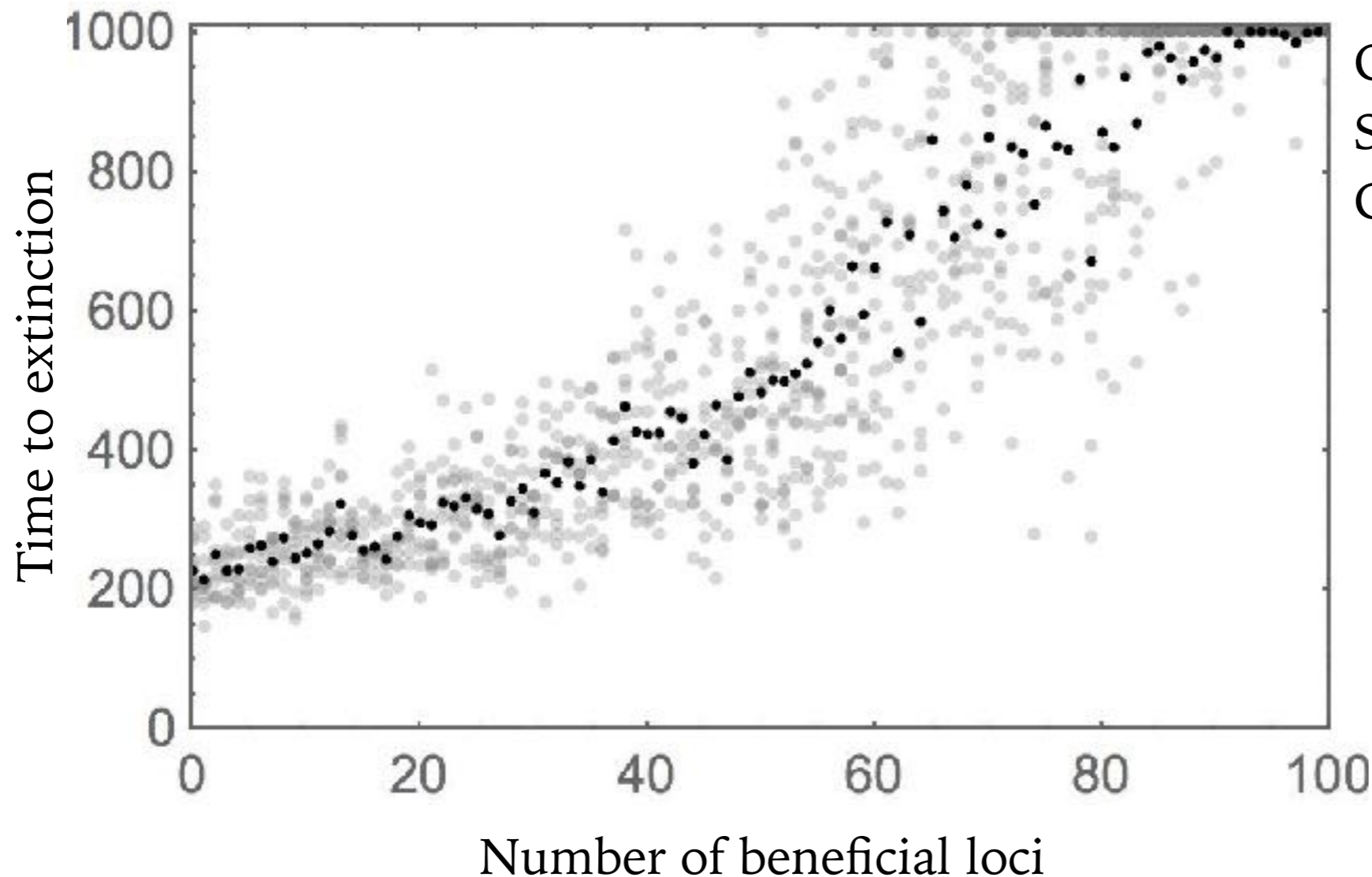


Carrying capacity: 250  
Selection coefficient: +/-0.05  
Genome length: 1000

- Many beneficials necessary to allow for significantly prolonged time to extinction.

# EXTINCTION TIMES WITH BENEFICIAL MUTATIONS

---



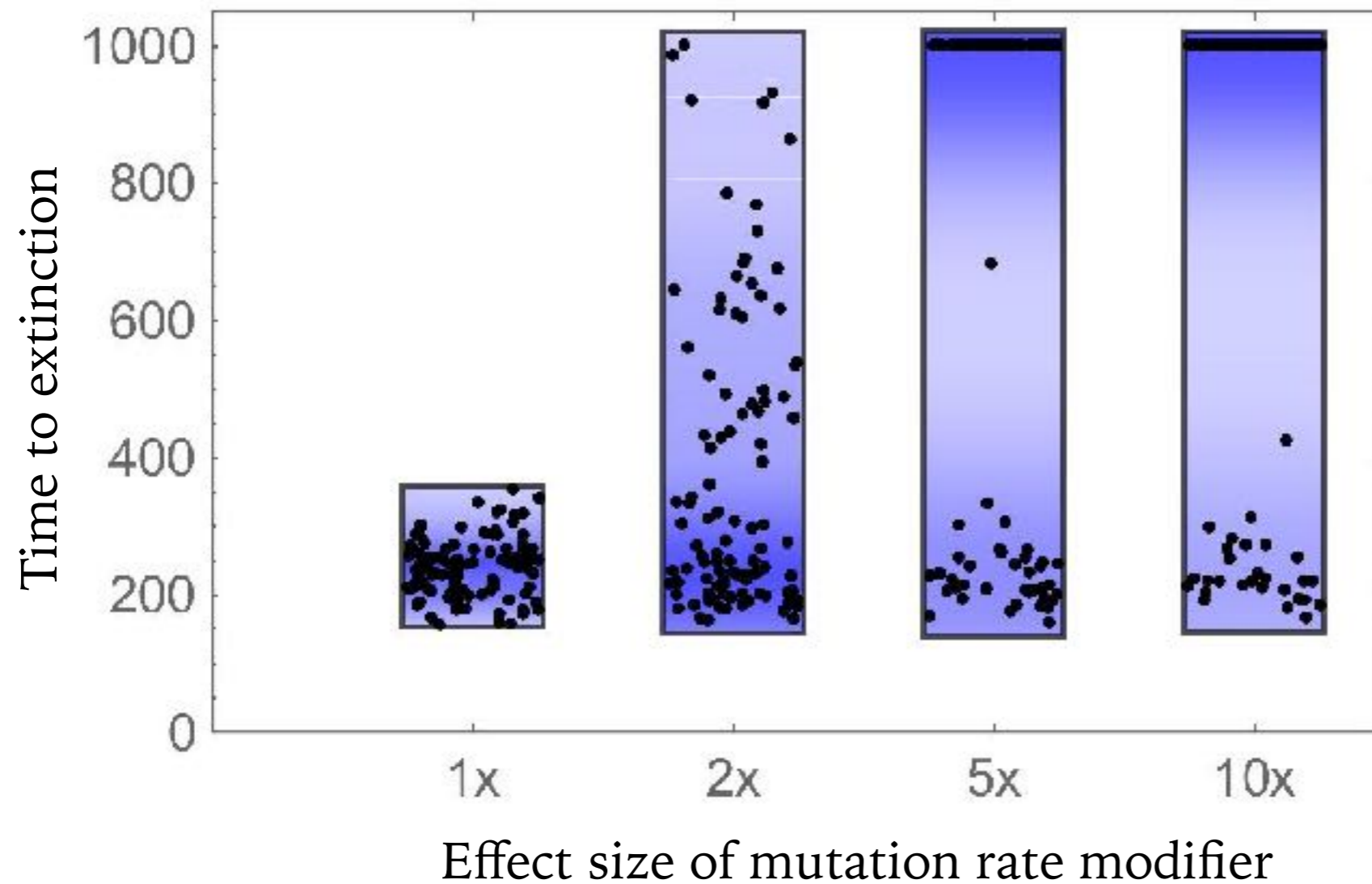
Carrying capacity: 250  
Selection coefficient: +/-0.05  
Genome length: 1000

- Many beneficials necessary to allow for significantly prolonged time to extinction.
- Clonal interference impedes efficient spread of multiple beneficials and increases variance in extinction times.



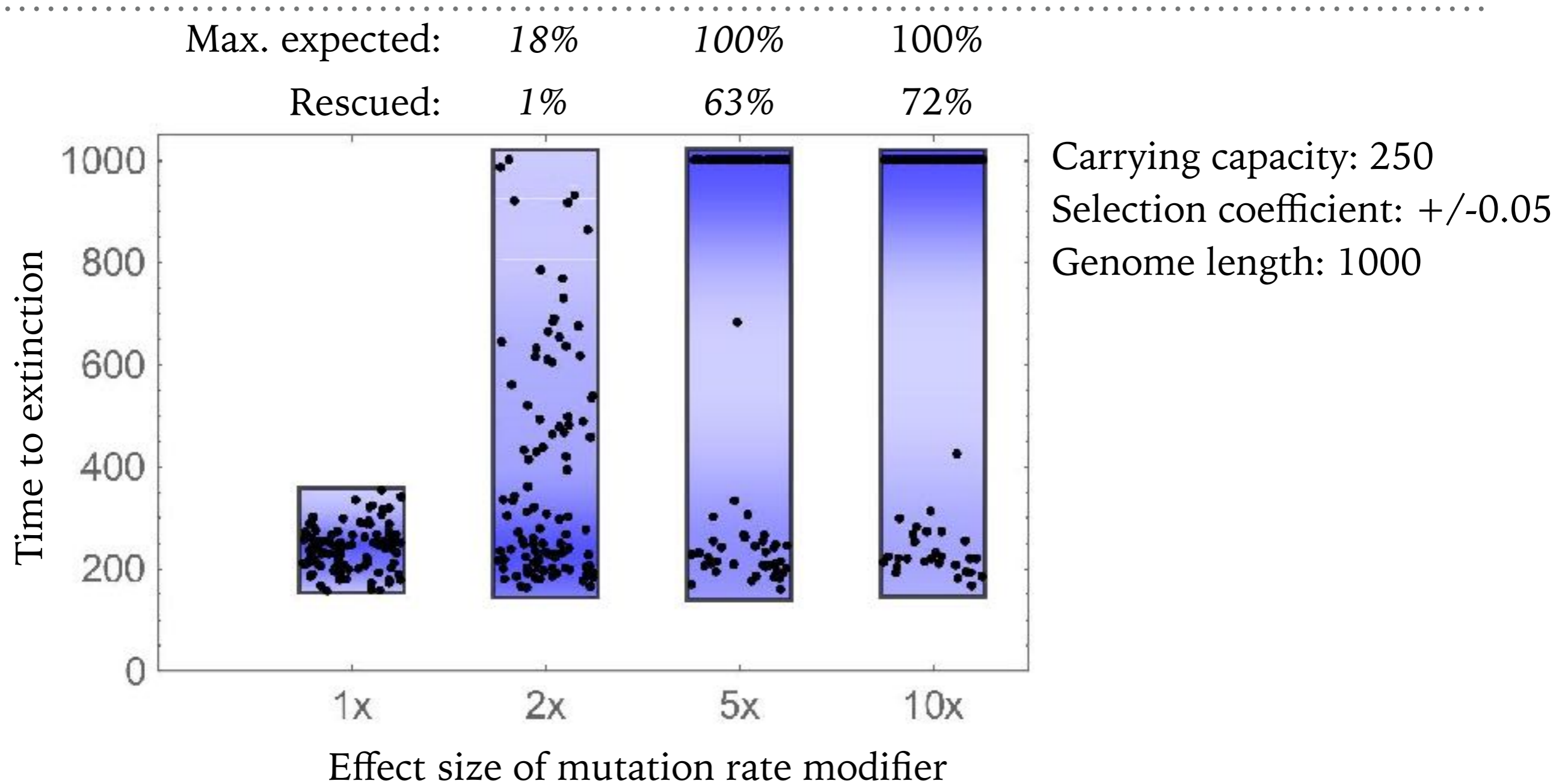
# INVASION OF A MUTATION RATE MODIFIER

---

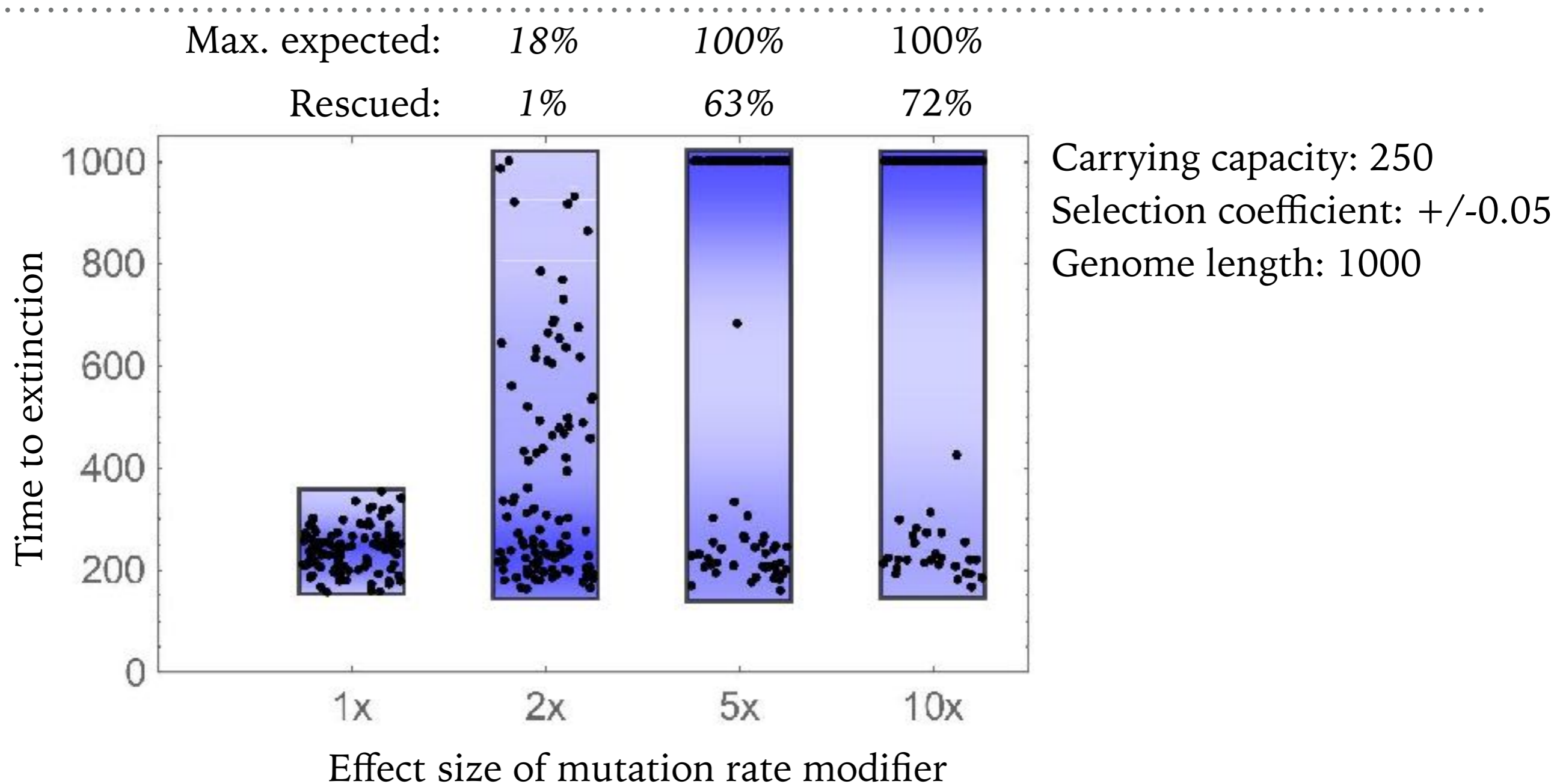


Carrying capacity: 250  
Selection coefficient: +/-0.05  
Genome length: 1000

# INVASION OF A MUTATION RATE MODIFIER



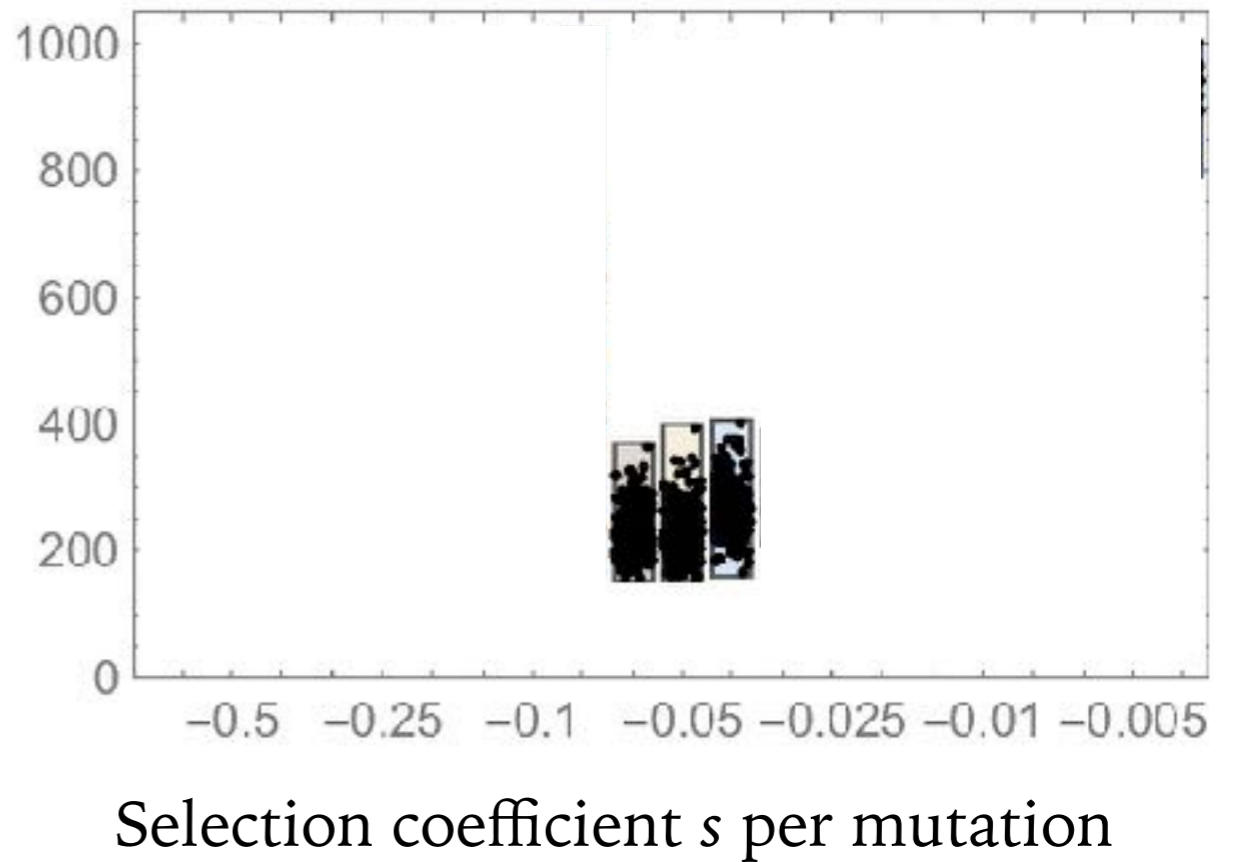
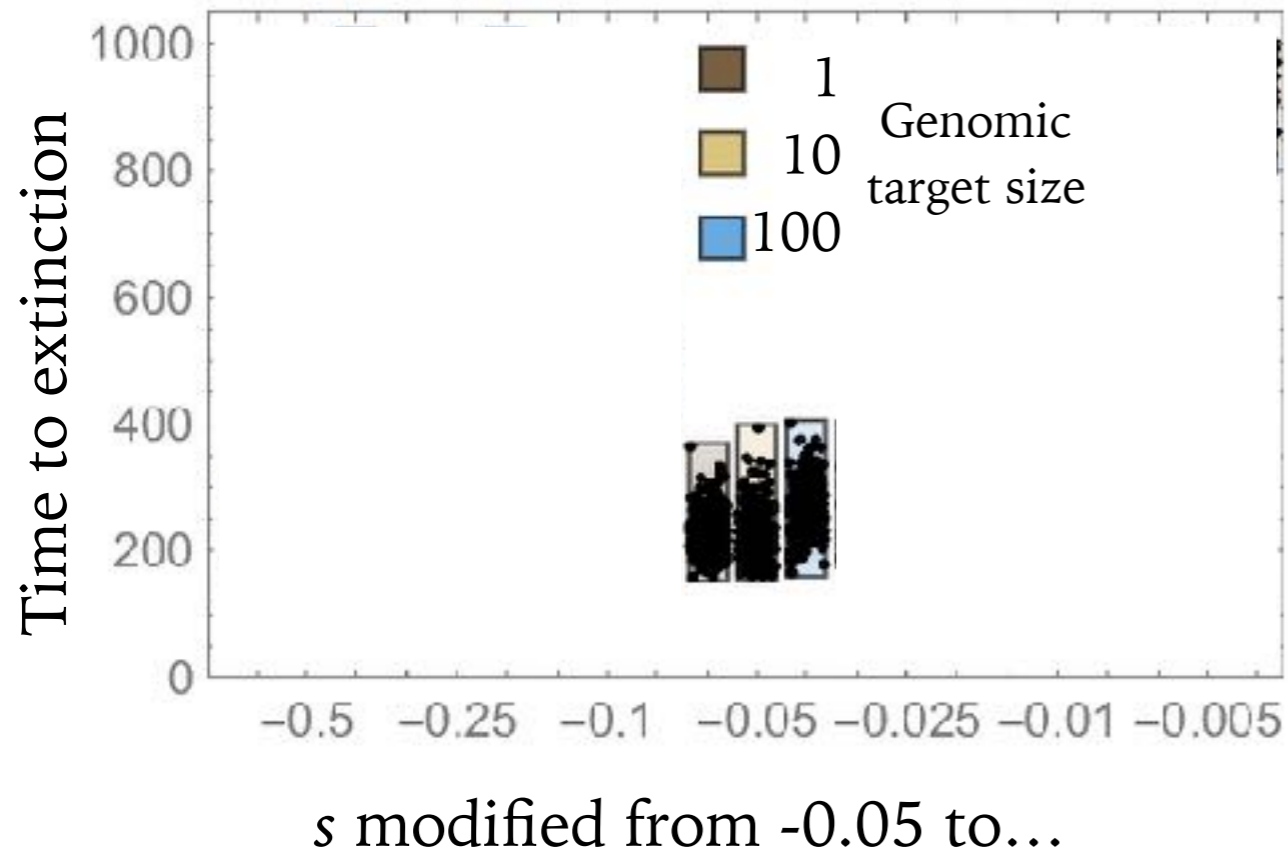
# INVASION OF A MUTATION RATE MODIFIER



- Mutation rate modifier of sufficient strength readily invades and rescues the population with high probability.

# INVASION OF A DFE MODIFIER

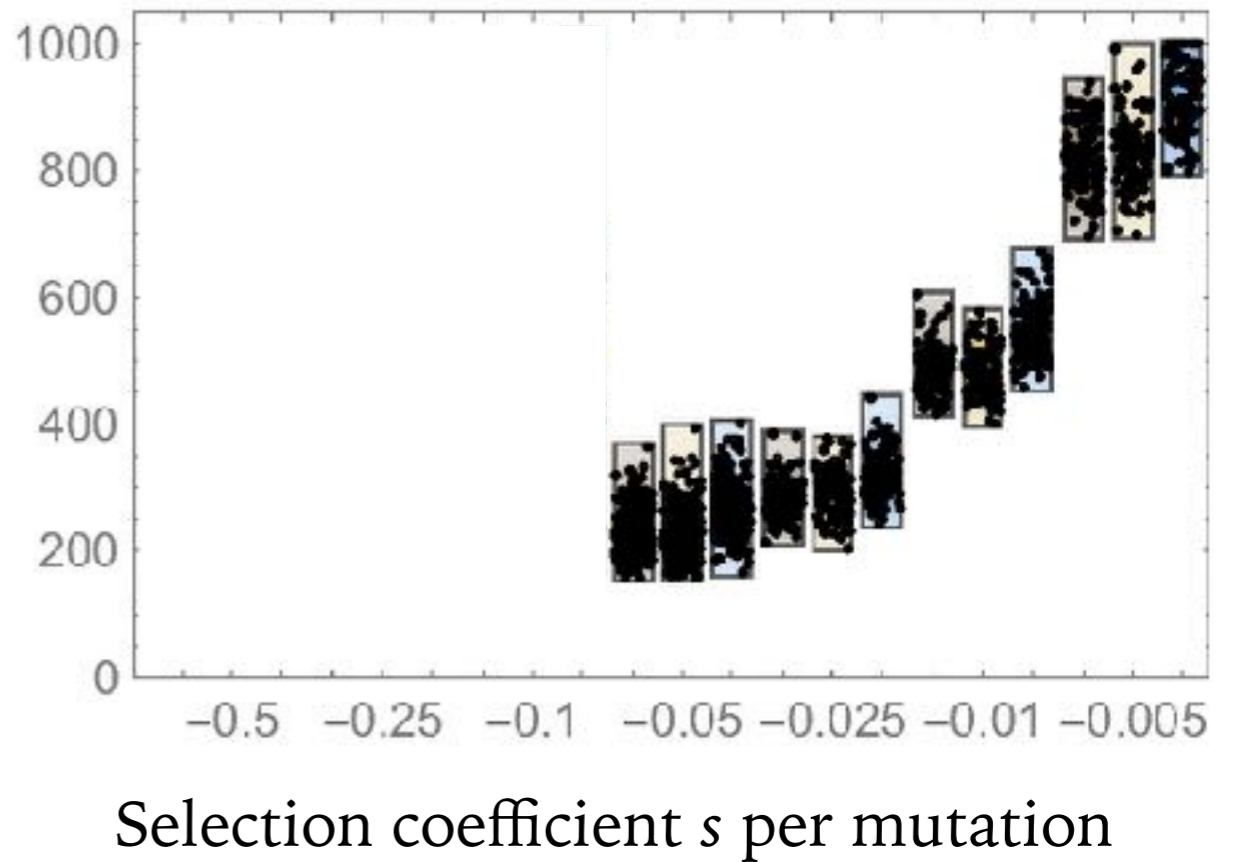
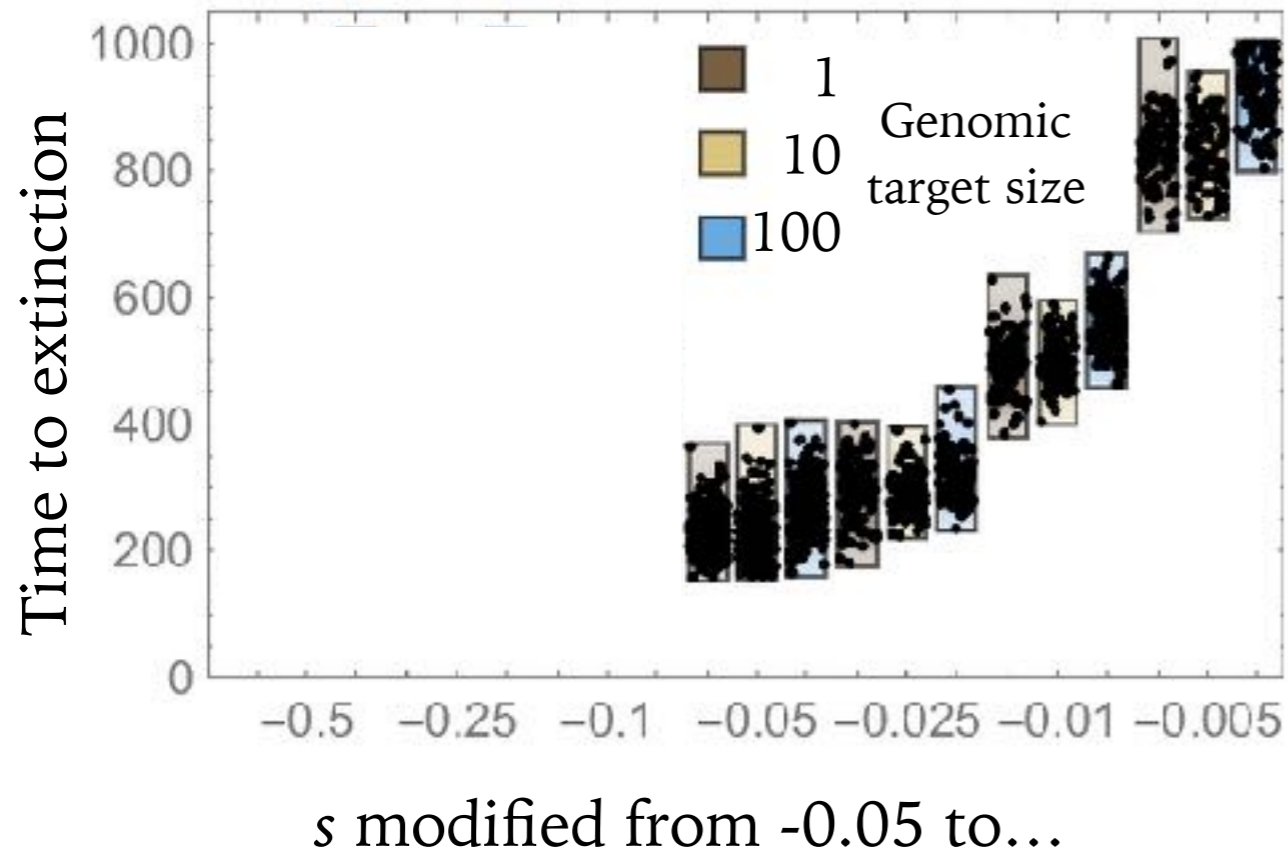
---





# INVASION OF A DFE MODIFIER

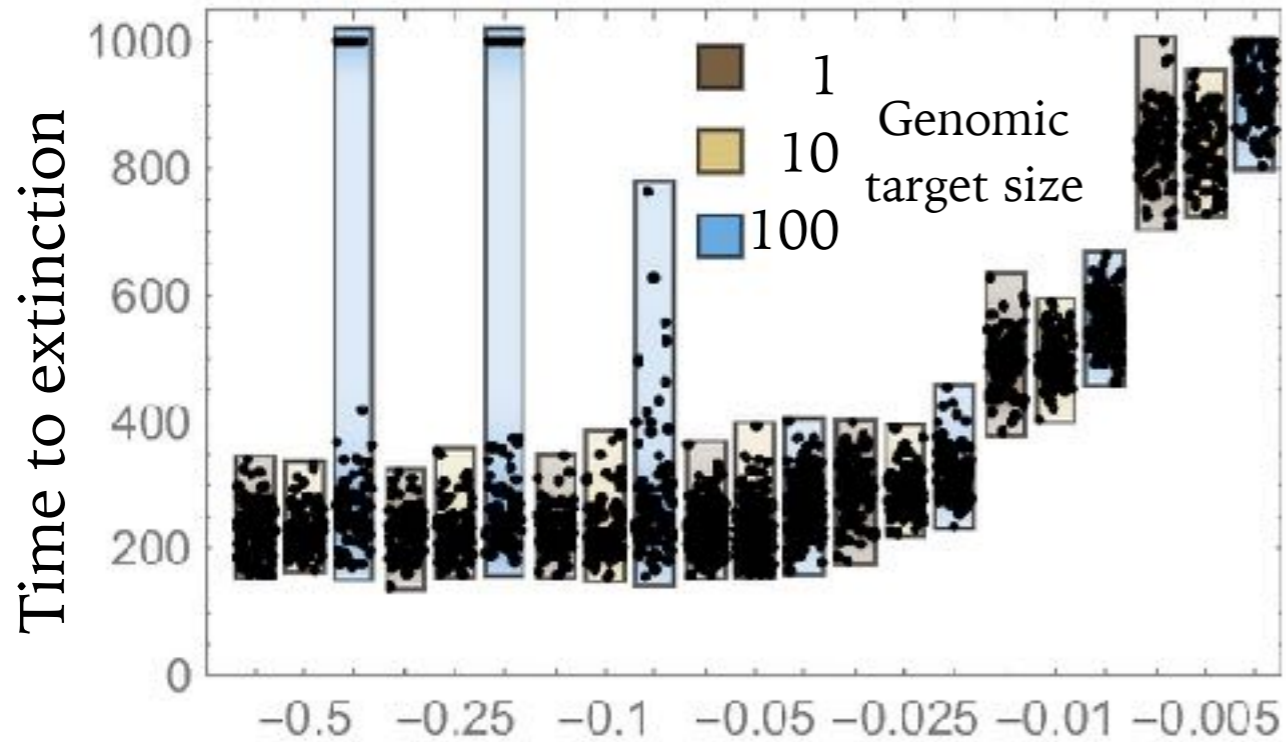
---



# INVASION OF A DFE MODIFIER

Rescued: 21% 22%

1% 14%

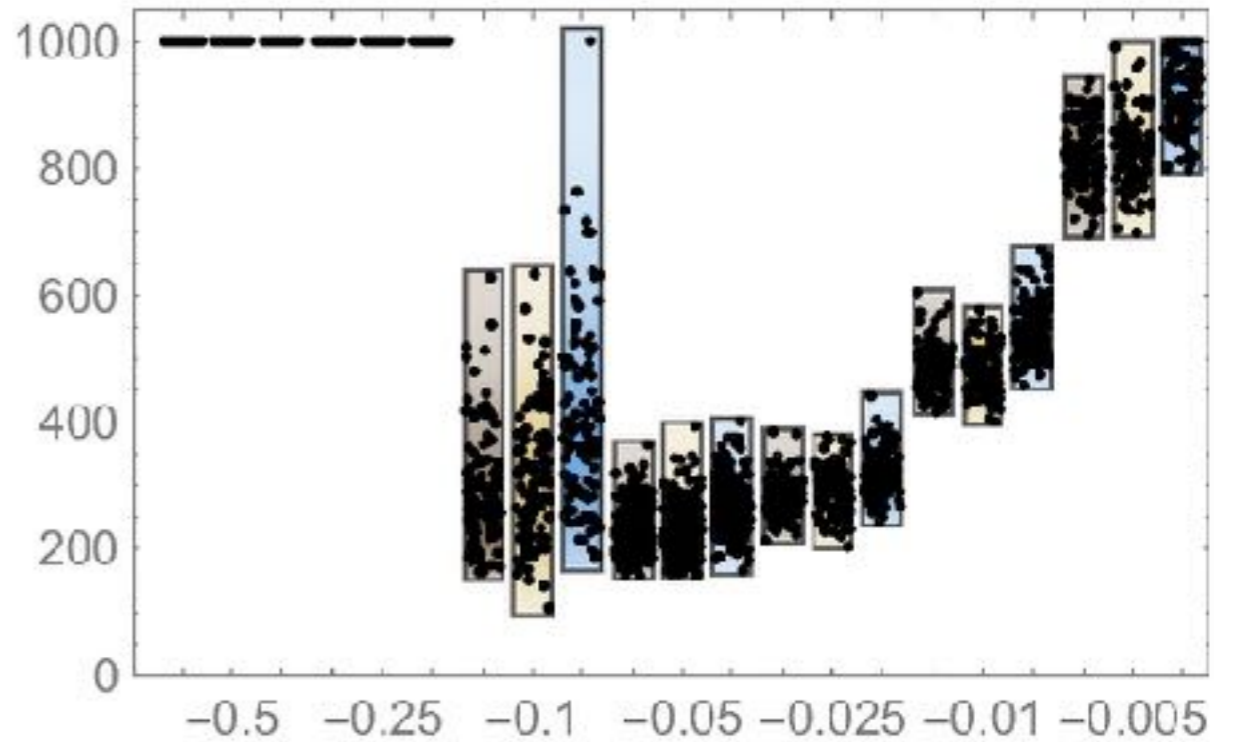


s modified from -0.05 to...

100%

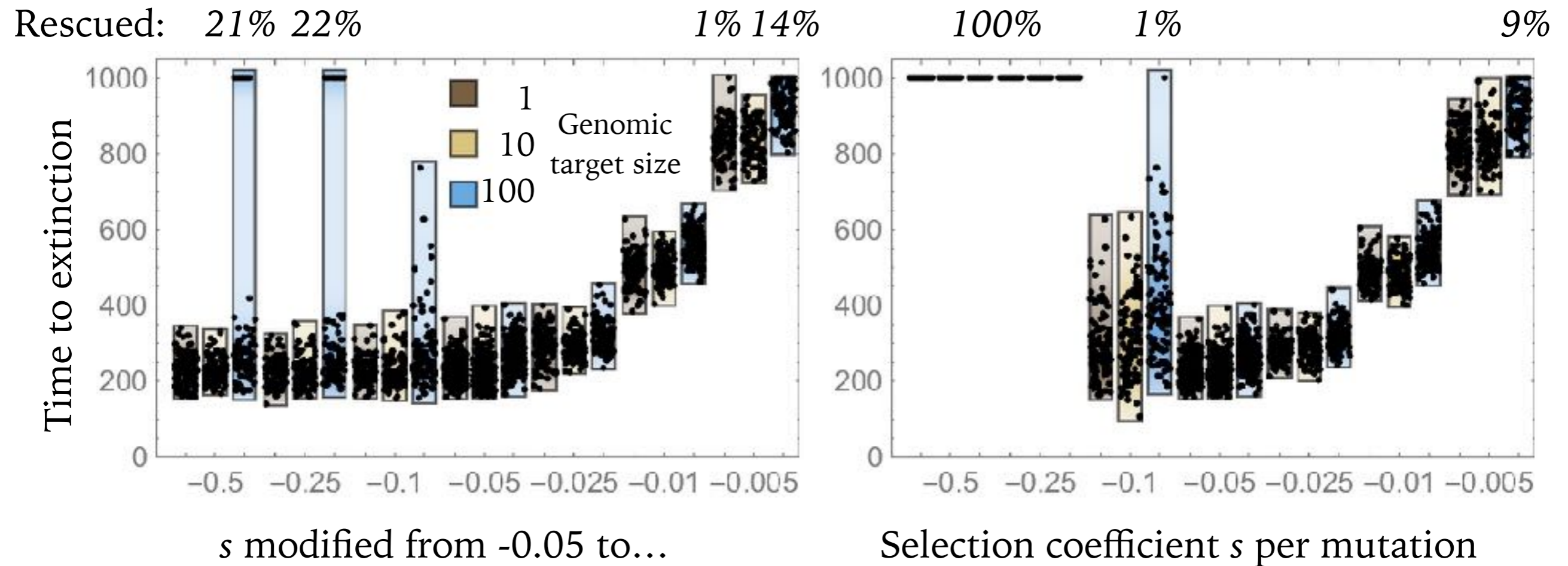
1%

9%



Selection coefficient  $s$  per mutation

# INVASION OF A DFE MODIFIER



- Both types of modifiers can invade; “chaperone” modifier invades easily but rarely rescues; “negative” modifier only invades under specific conditions but then rescues reliably.

# CONCLUSIONS





# CONCLUSIONS

---

- Extinction process is rather deterministic over a large range of the parameter space.

# CONCLUSIONS

---

- Extinction process is rather deterministic over a large range of the parameter space.
- Many available beneficials are needed to prolong the extinction time (e.g., to successful transmission of the virus).

# CONCLUSIONS

---

- Extinction process is rather deterministic over a large range of the parameter space.
- Many available beneficials are needed to prolong the extinction time (e.g., to successful transmission of the virus).
- If available, mutation rate modifiers readily invade and make the population resistant to mutagenic treatment.

# CONCLUSIONS

---

- Extinction process is rather deterministic over a large range of the parameter space.
- Many available beneficials are needed to prolong the extinction time (e.g., to successful transmission of the virus).
- If available, mutation rate modifiers readily invade and make the population resistant to mutagenic treatment.
- DFE modifiers in both directions can invade and make the virus tolerant to high mutation rates. This is possibly the most dangerous adaptation mechanism, because it could modify virus evolution also in absence of the drug.



# ACKNOWLEDGEMENTS

## TEAM



Alex Blanckaert



Inês Fragata



Hermina Ghenu



Marco Louro



Mark Schmitz

## COLLABORATORS

Maria João Amorim, IGC  
Thomas Bataillon, Aarhus University  
Daniel Bolon, Pam Cote, Ryan Hietpas,  
UMass Medical School  
Roger Butlin, University of Sheffield  
Mónica Bettencourt-Dias, IGC  
Isabel Gordo, IGC  
Jeffrey Jensen, Arizona State University  
Rees Kassen, University of Ottawa  
Jonna Kulmuni, Helsinki University  
David Liberles, Temple University  
Sebastian Matuszewski, EPFL  
Vitor Sousa, University of Lisbon  
Alex Wong, Carleton University

## FUNDING



FUNDAÇÃO  
CALOUSTE  
GULBENKIAN

**FCT**

Fundação para a Ciência e a Tecnologia  
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR



JOINT PROGRAMMING  
INITIATIVE ON ANTIMICROBIAL  
RESISTANCE