ON ADAPTATION, EPISTASIS AND FITNESS LANDSCAPES

CLAUDIA BANK

INSTITUTO GULBENKIAN DE CIÊNCIA

EVOLUTION (& HIGH-SPEED INTRO TO POPULATION GENETICS)

- Change in heritable trait frequencies of a population over time

EVOLUTION (& HIGH-SPEED INTRO TO POPULATION GENETICS)

➤ Change in heritable trait frequencies of a population over time

Mutation *Genetic drift* *Migration* *Selection*

EVOLUTION (& HIGH-SPEED INTRO TO POPULATION GENETICS)

- Change in heritable trait frequencies of a population over time

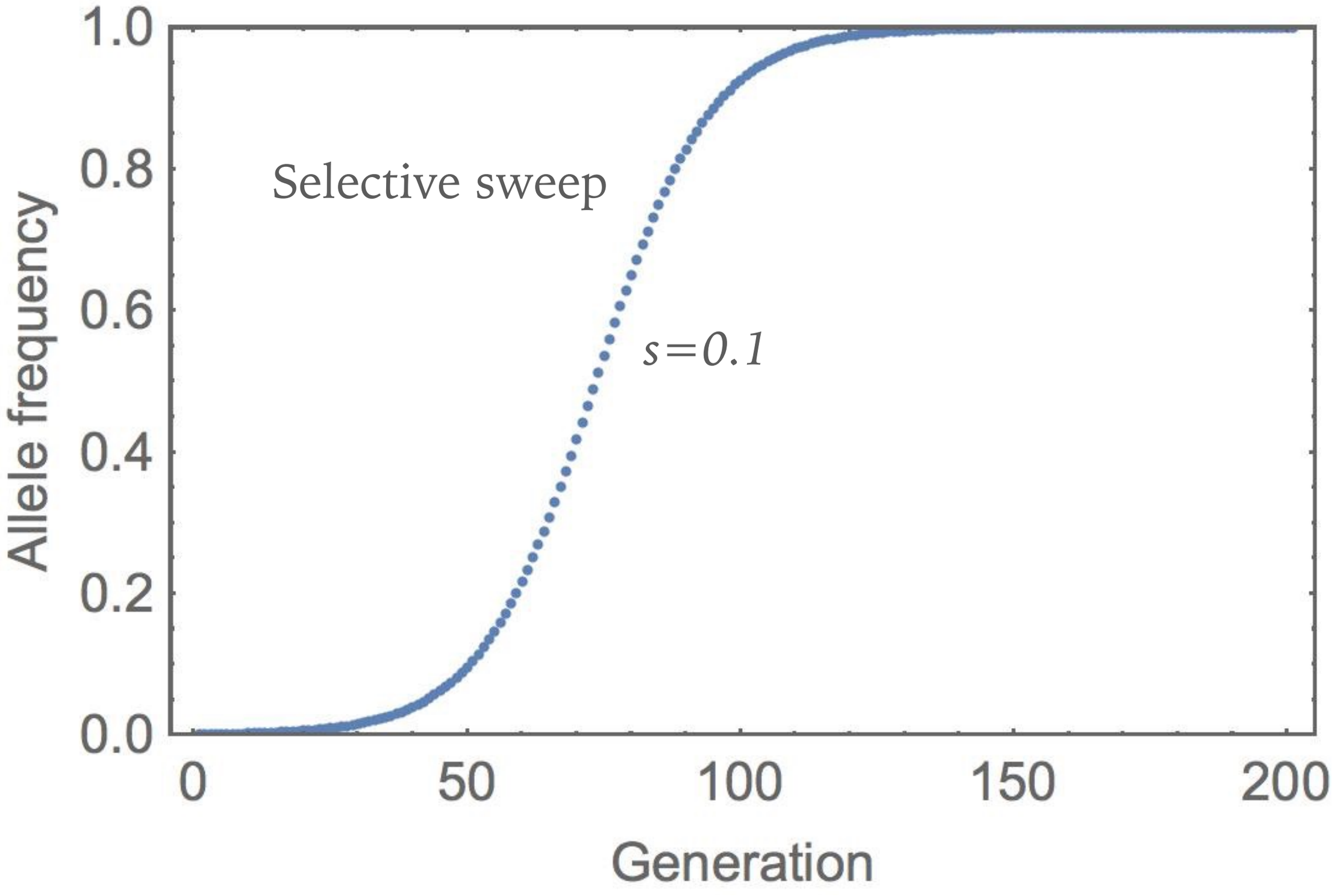
Mutation *Genetic drift* *Migration* *Selection*

What is your "favorite" evolutionary force, and why?

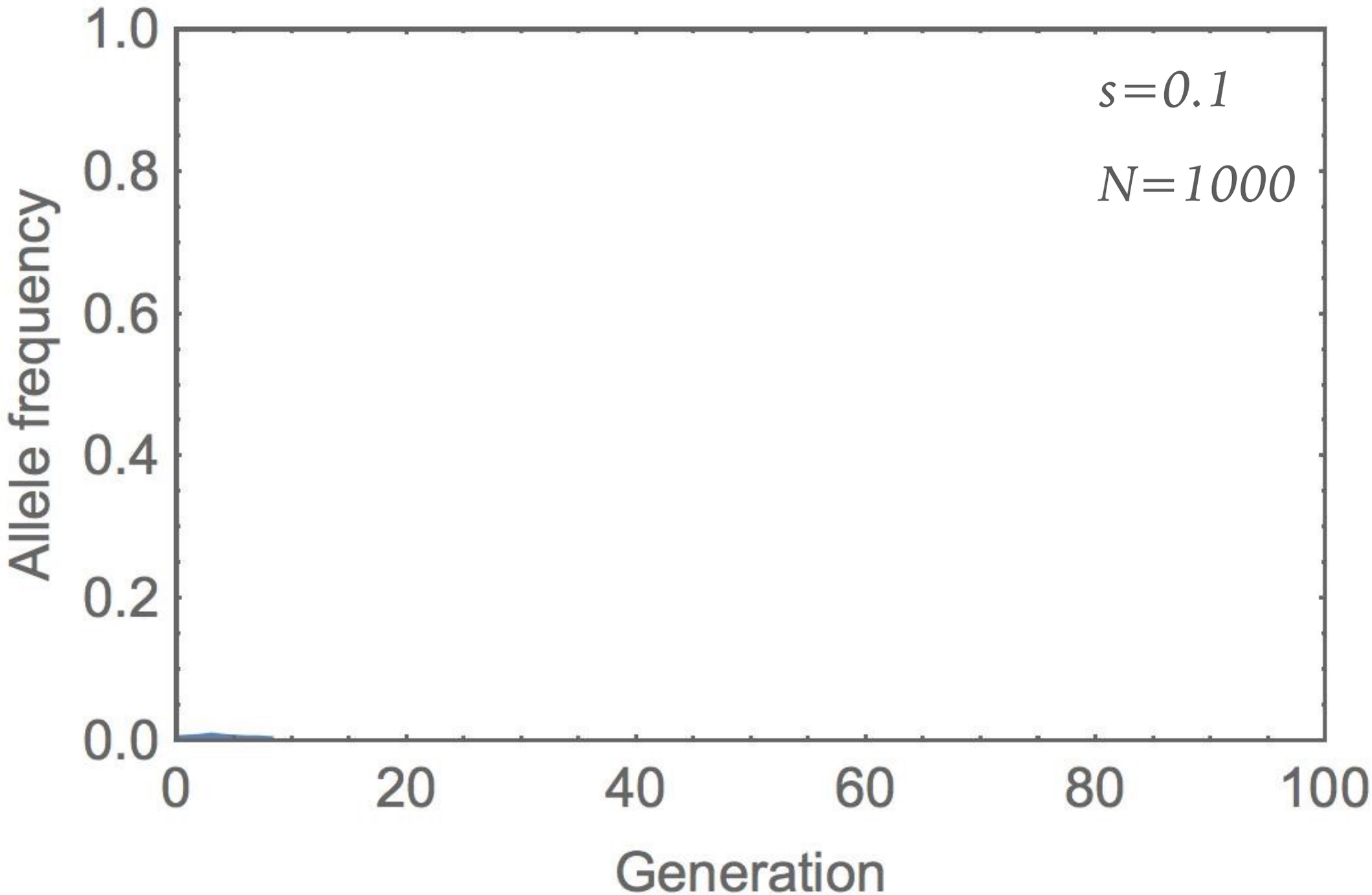
> Although many processes shape evolution, natural selection is special because it creates complex, functioning organisms. All other processes tend to degrade what has been built up by natural selection, simply because these processes act at random with respect to function.

-Barton et al. , Evolution (textbook)

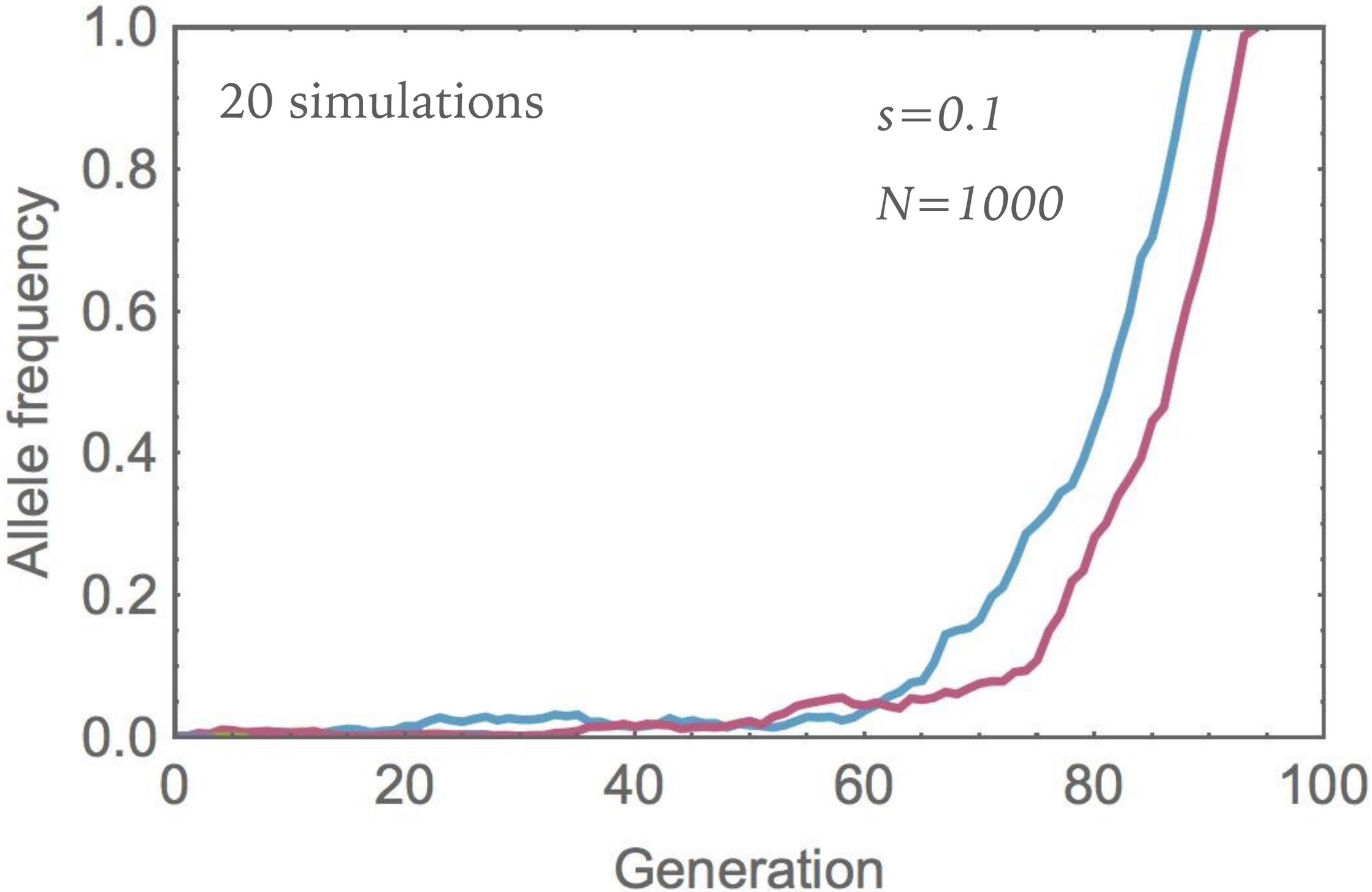
THEORETICALLY, SELECTION IS THE “EASIEST" EVOLUTIONARY FORCE



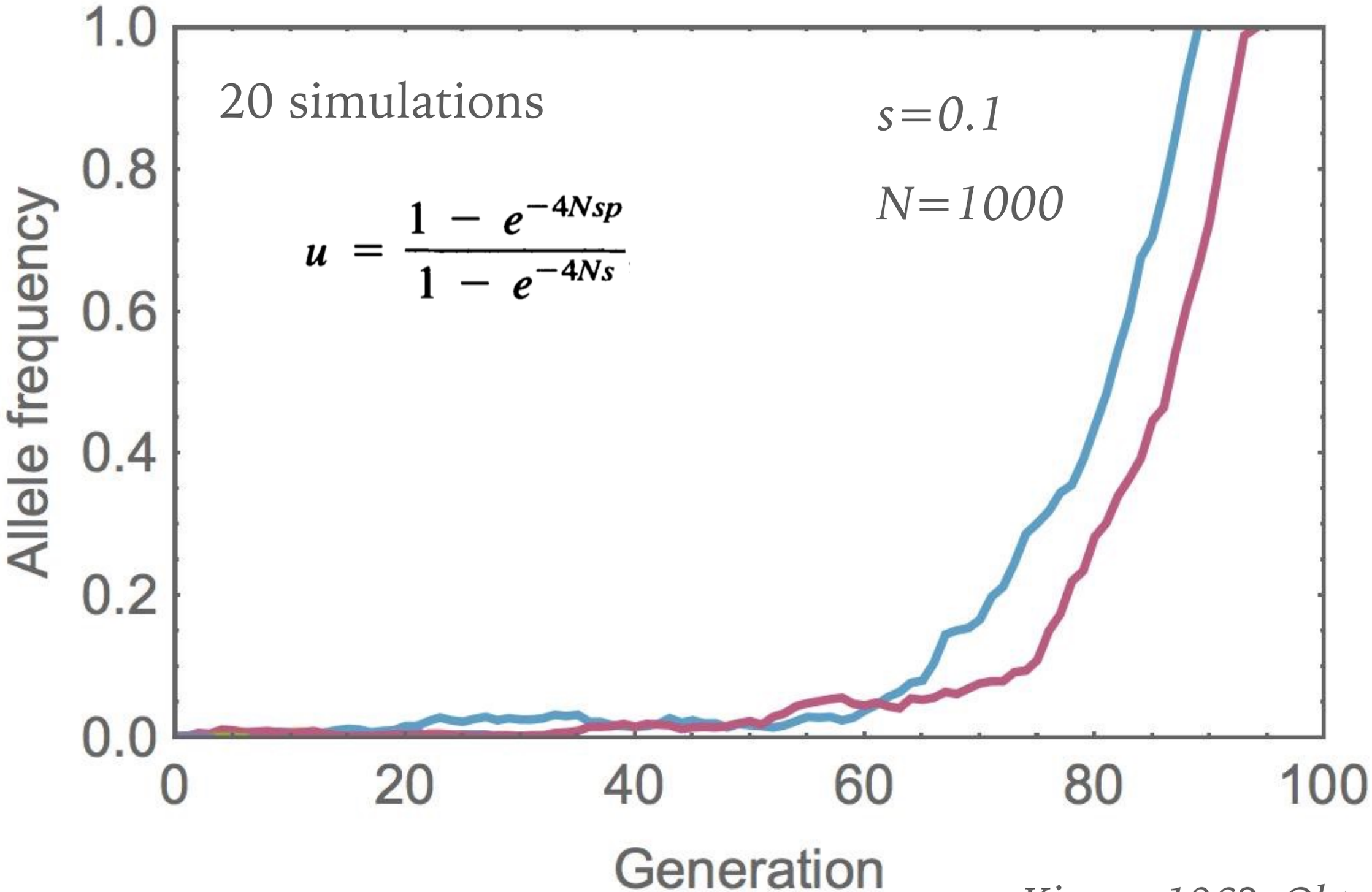
BUT THEN, GENETIC DRIFT COMES ALONG



BUT THEN, GENETIC DRIFT COMES ALONG

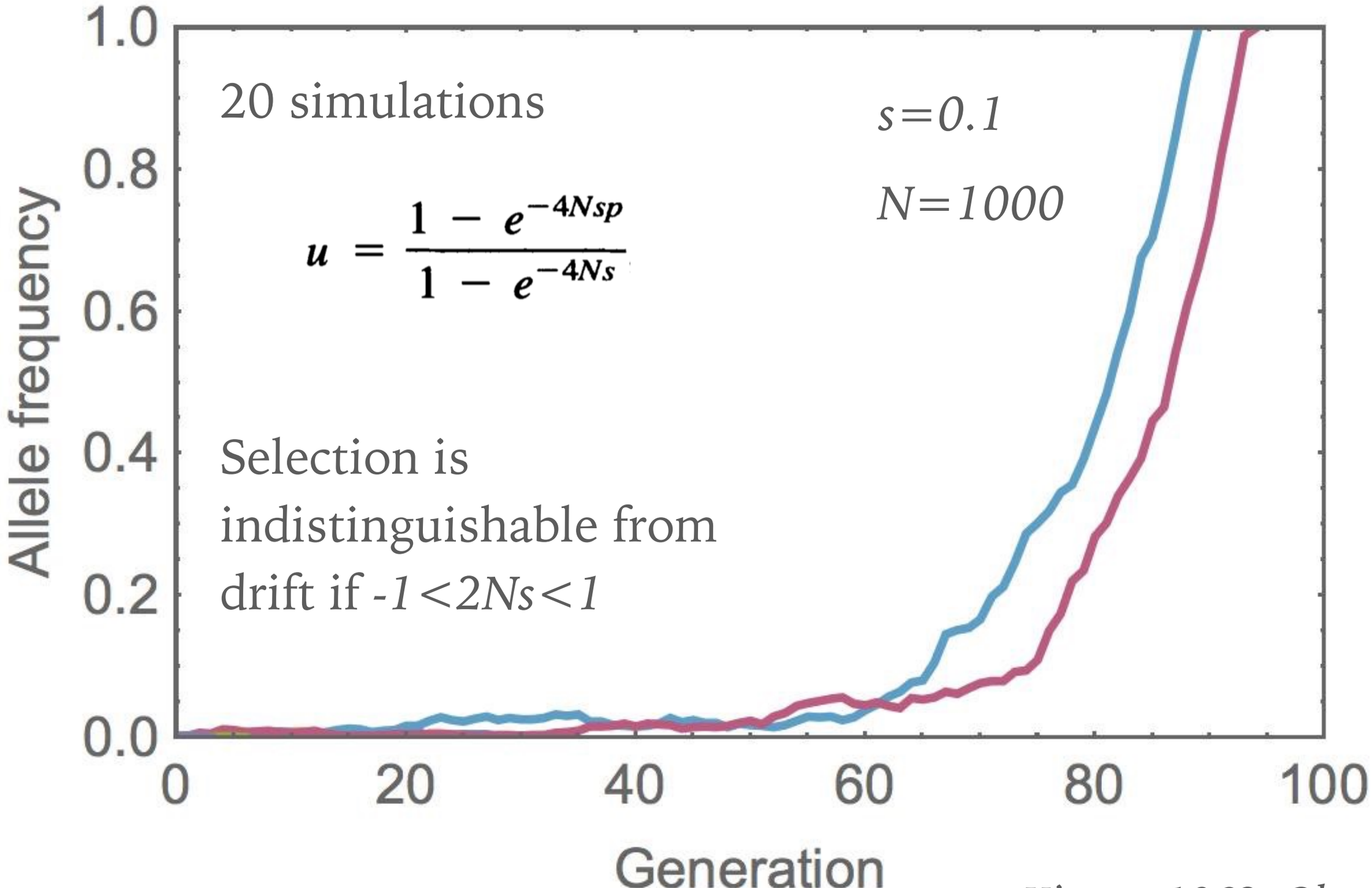


BUT THEN, GENETIC DRIFT COMES ALONG



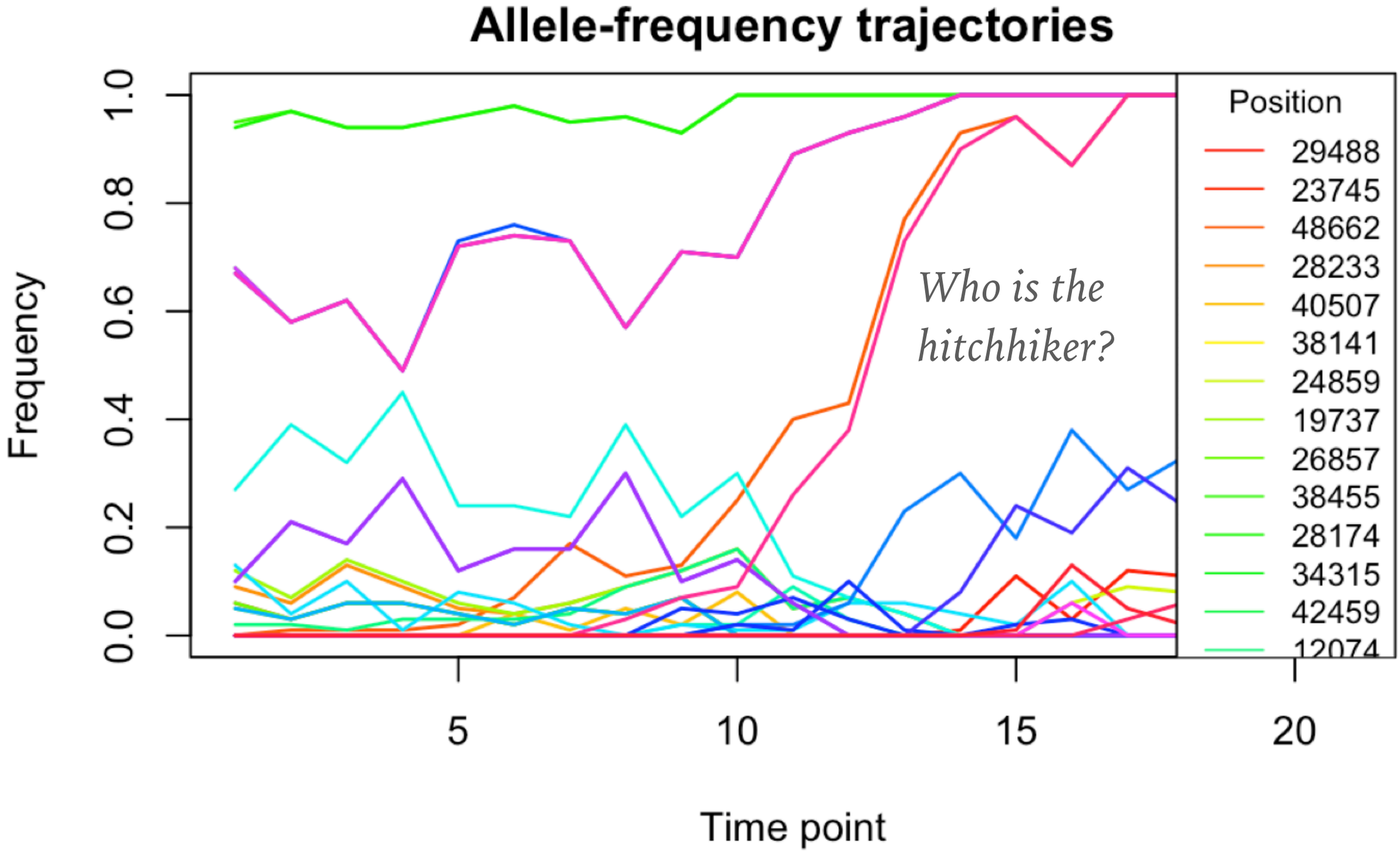
Kimura 1962, Ohta 1973

BUT THEN, GENETIC DRIFT COMES ALONG



Kimura 1962, Ohta 1973

OR LINKAGE


Allele-frequency trajectories
Frequency
0.0
0.2
0.4
0.6
0.8
1.0
5
10
15
20
Time point
Who is the hitchhiker?
Position
29488
23745
48662
28233
40507
38141
24859
19737
26857
38455
28174
34315
42459
12074

OR POPULATION STRUCTURE, OR EPISTASIS, OR [ADD YOUR FAVORITE HERE]

- Keep in mind that selection operates on phenotypic differences among individuals in a population; it does not act on a genotype, much less an allele.

WHAT POPULATION GENETICISTS WANT TO KNOW

WHAT POPULATION GENETICISTS WANT TO KNOW

- How many and which mutations were beneficial?

WHAT POPULATION GENETICISTS WANT TO KNOW

- How many and which mutations were beneficial?
- What are the proportions of beneficial, neutral, and deleterious mutations?

WHAT POPULATION GENETICISTS WANT TO KNOW

- How many and which mutations were beneficial?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- What is the relative role of selection vs. other evolutionary processes in shaping genomes?

WHAT POPULATION GENETICISTS WANT TO KNOW

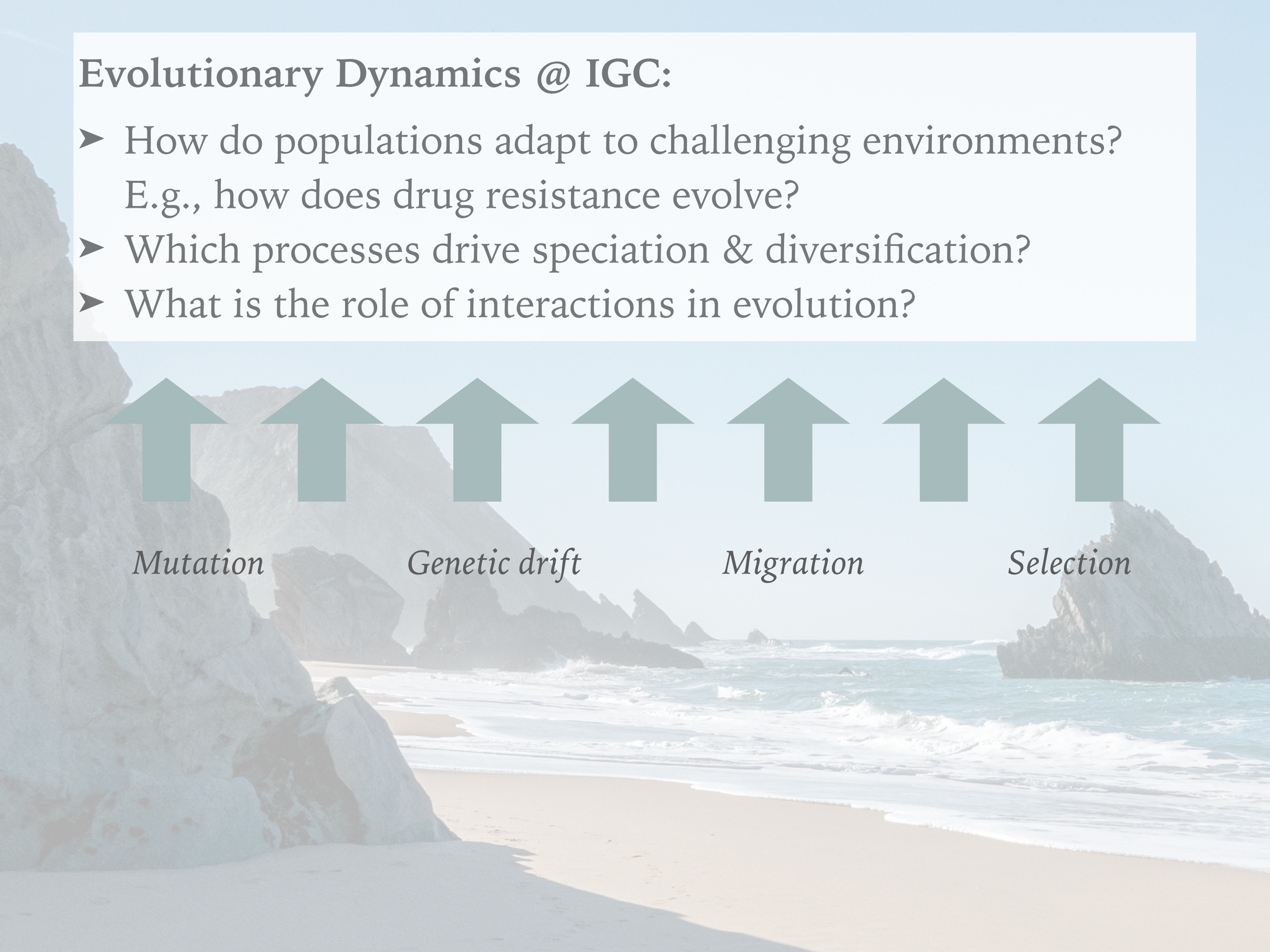
- How many and which mutations were beneficial?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- What is the relative role of selection vs. other evolutionary processes in shaping genomes?
- How can we infer the contribution of selection to molecular evolution?

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?

Evolutionary Dynamics @ IGC:

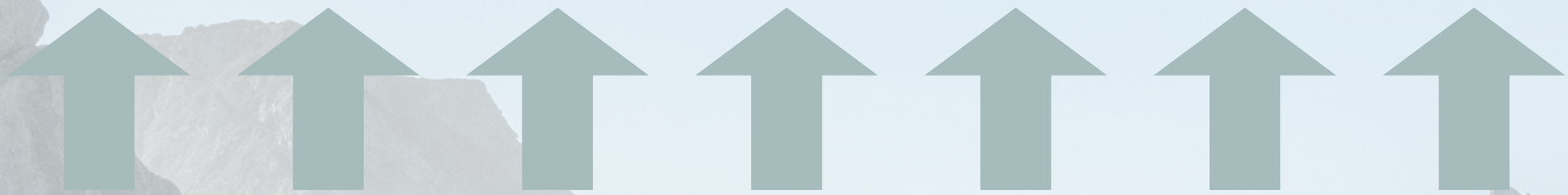
- How do populations adapt to challenging environments?
  E.g., how does drug resistance evolve?
- Which processes drive speciation & diversification?
- What is the role of interactions in evolution?



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?



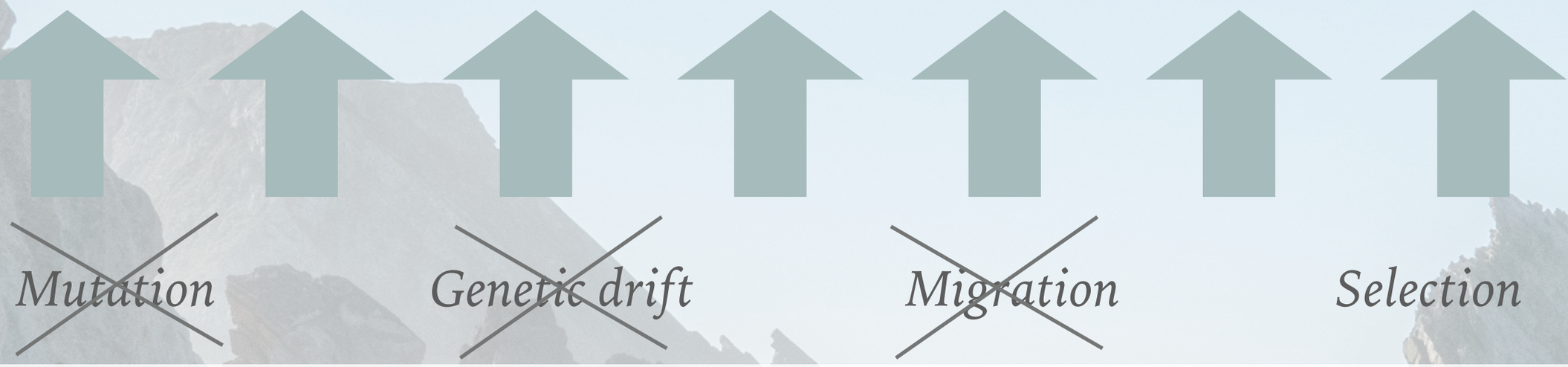
Mutation *Genetic drift* *Migration* *Selection*

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*

ADAPTATION VS. "AN ADAPTATION"

- Adaptation: the process of increasing (mean) fitness of a population in a given environment
- An adaptation: a trait that increases its carrier's fitness in a specific environment, and that has spread bc of of the direct action of natural selection for its function

ADAPTATION VS. "AN ADAPTATION"

- Adaptation: the process of increasing (mean) fitness of a population in a given environment
- An adaptation: a trait that increases its carrier's fitness in a specific environment, and that has spread bc of of the direct action of natural selection for its function

ADAPTATION VS. "AN ADAPTATION"

- Adaptation: the process of increasing (mean) fitness of a population in a given environment
- An adaptation: a trait that increases its carrier's fitness in a specific environment, and that has spread bc of of the direct action of natural selection for its function

From Nachmann et al., PNAS, 2003

WHAT WE WANT TO KNOW ABOUT ADAPTATION

WHAT WE WANT TO KNOW ABOUT ADAPTATION

- How big/small are adaptive steps?

WHAT WE WANT TO KNOW ABOUT ADAPTATION

- How big/small are adaptive steps?
- What are the proportions of beneficial, neutral, and deleterious mutations?

WHAT WE WANT TO KNOW ABOUT ADAPTATION

- How big/small are adaptive steps?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- How do mutational effects change dependent on the environment?

WHAT WE WANT TO KNOW ABOUT ADAPTATION

- How big/small are adaptive steps?
- 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 WE WANT TO KNOW ABOUT ADAPTATION

- How big/small are adaptive steps?
- 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 role of selection vs. other evolutionary processes in shaping genomes?
- How can we infer the contribution of selection to molecular evolution?

WHAT WE WANT TO KNOW ABOUT ADAPTATION

- How big/small are adaptive steps?
- 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 role of selection vs. other evolutionary processes in shaping genomes?
- How can we infer the contribution of selection to molecular evolution?

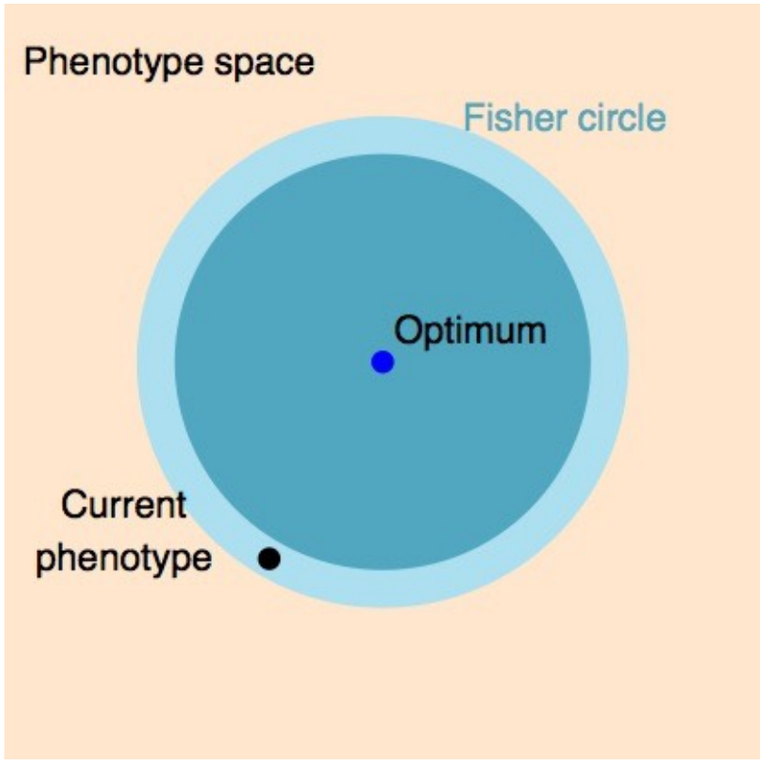
WHAT WE WANT TO KNOW ABOUT ADAPTATION

- How big/small are adaptive steps?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- How do mutational effects change dependent on the genetic background? (I.e., what is the role of epistasis?)
- What is the role of selection vs. other evolutionary processes in shaping genomes?
- How can we infer the contribution of selection to molecular evolution?

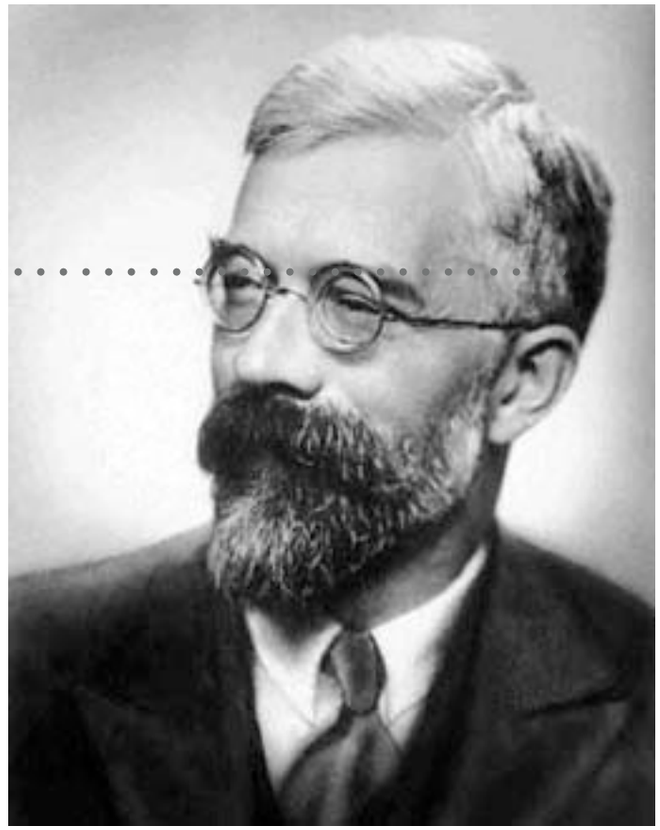
What do we expect adaptation to be like THEORETICALLY?

TWO MODELS OF ADAPTATION

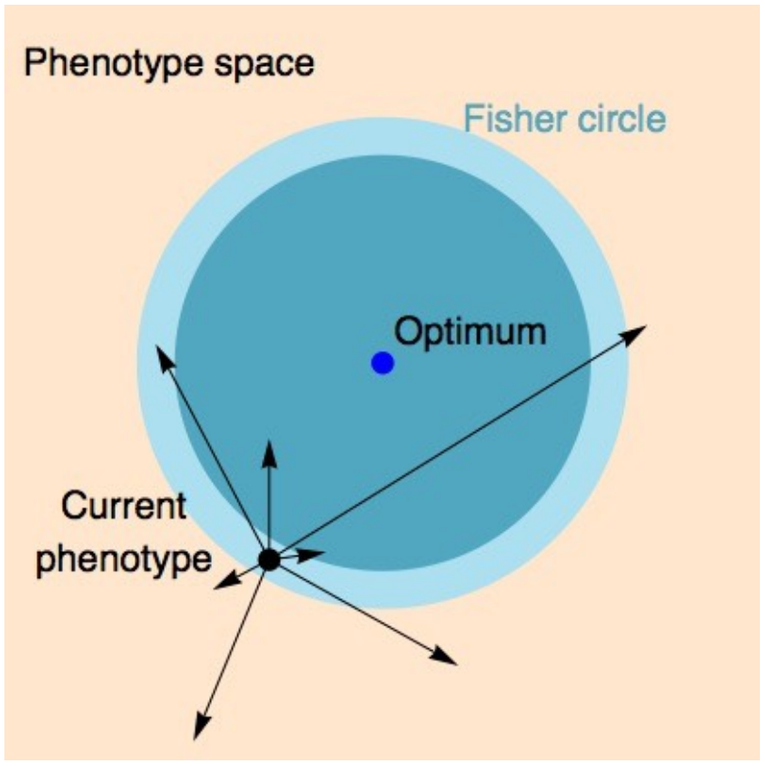
FISHER'S GEOMETRIC MODEL



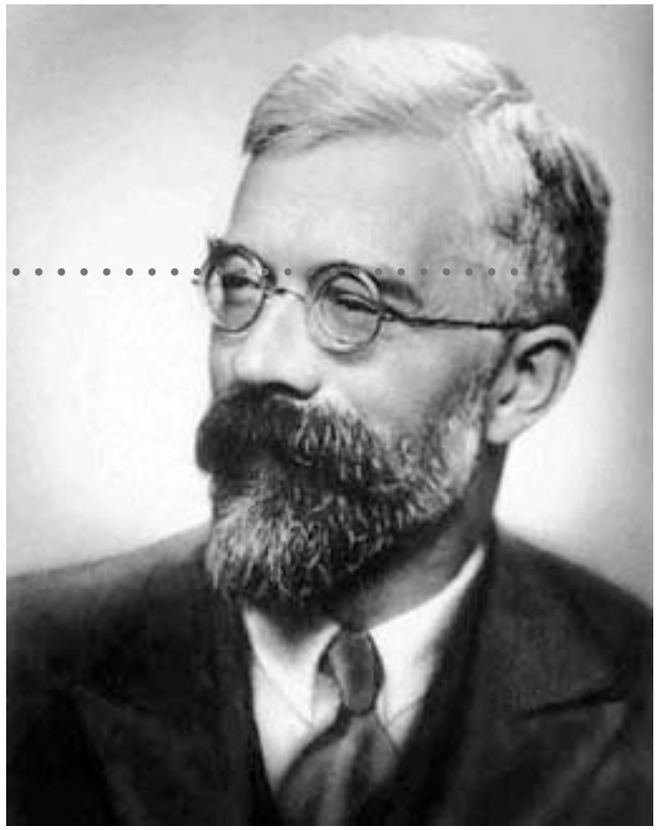
Fisher, 1930



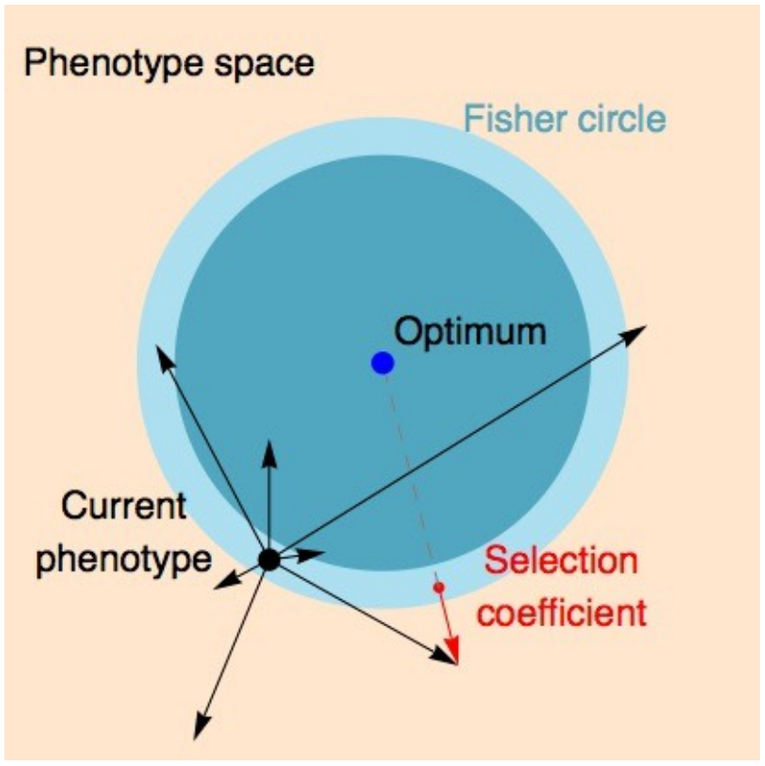
FISHER'S GEOMETRIC MODEL



Fisher, 1930

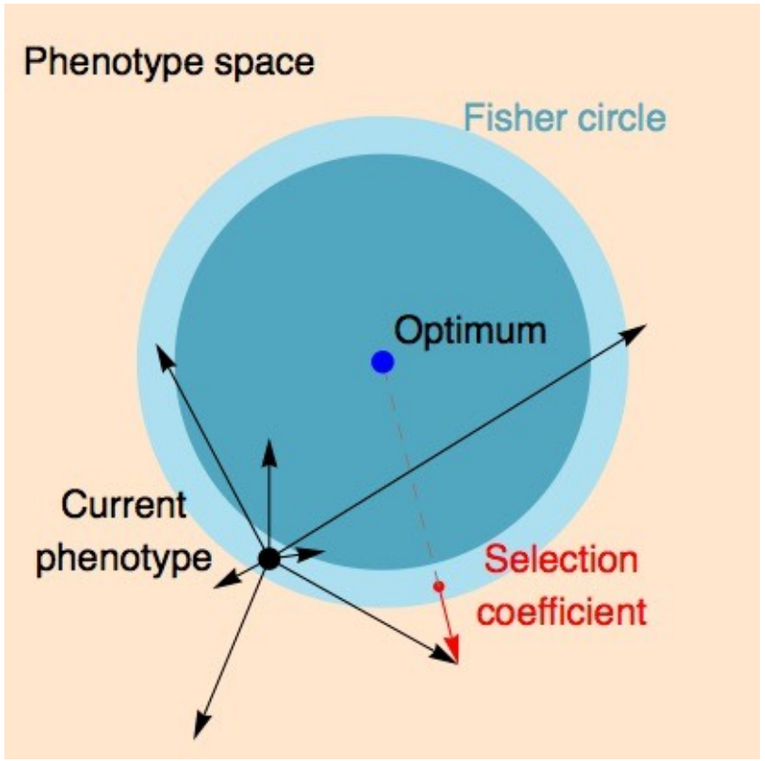


FISHER'S GEOMETRIC MODEL

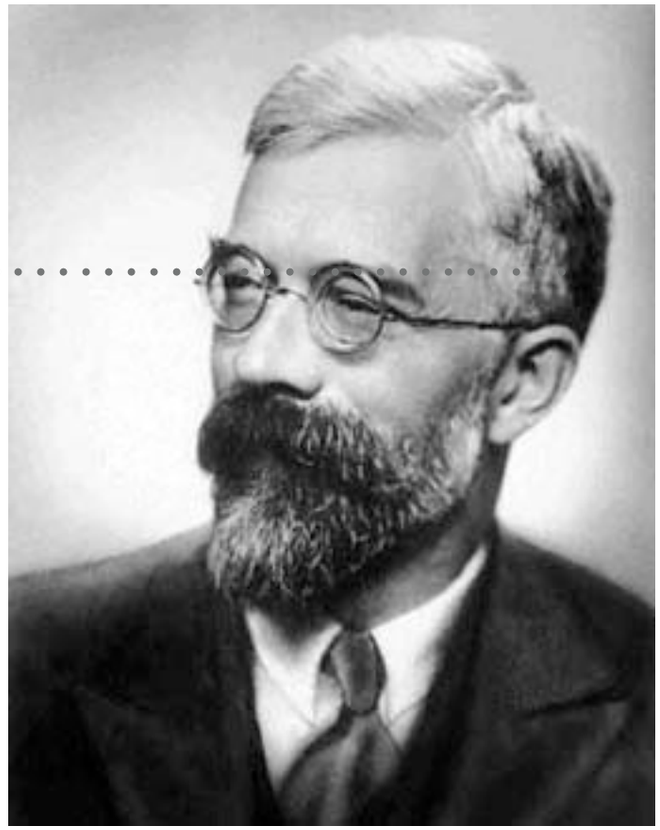


Fisher, 1930

FISHER'S GEOMETRIC MODEL



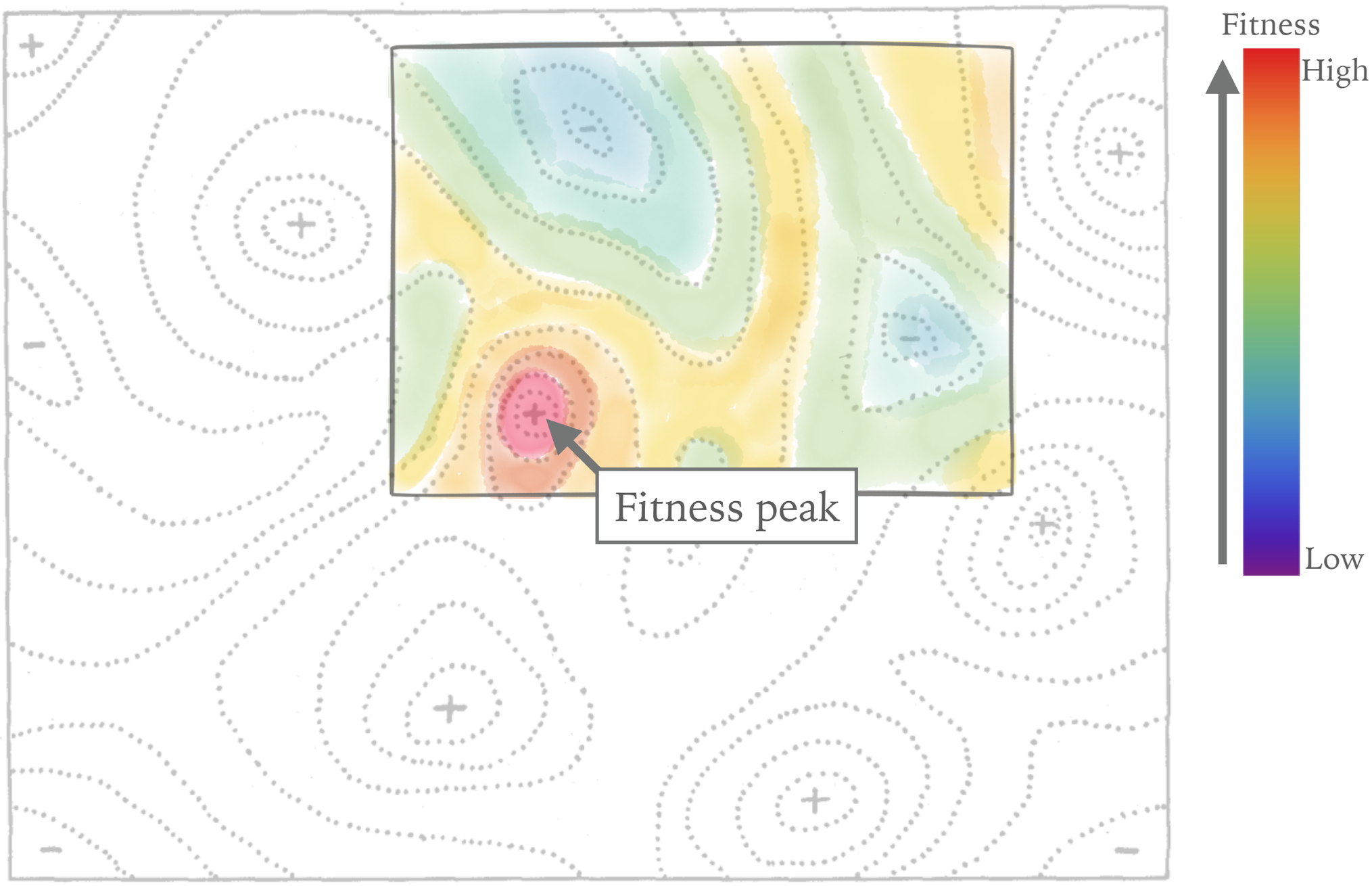
Fisher, 1930



- More challenging environment
  => more beneficial mutations

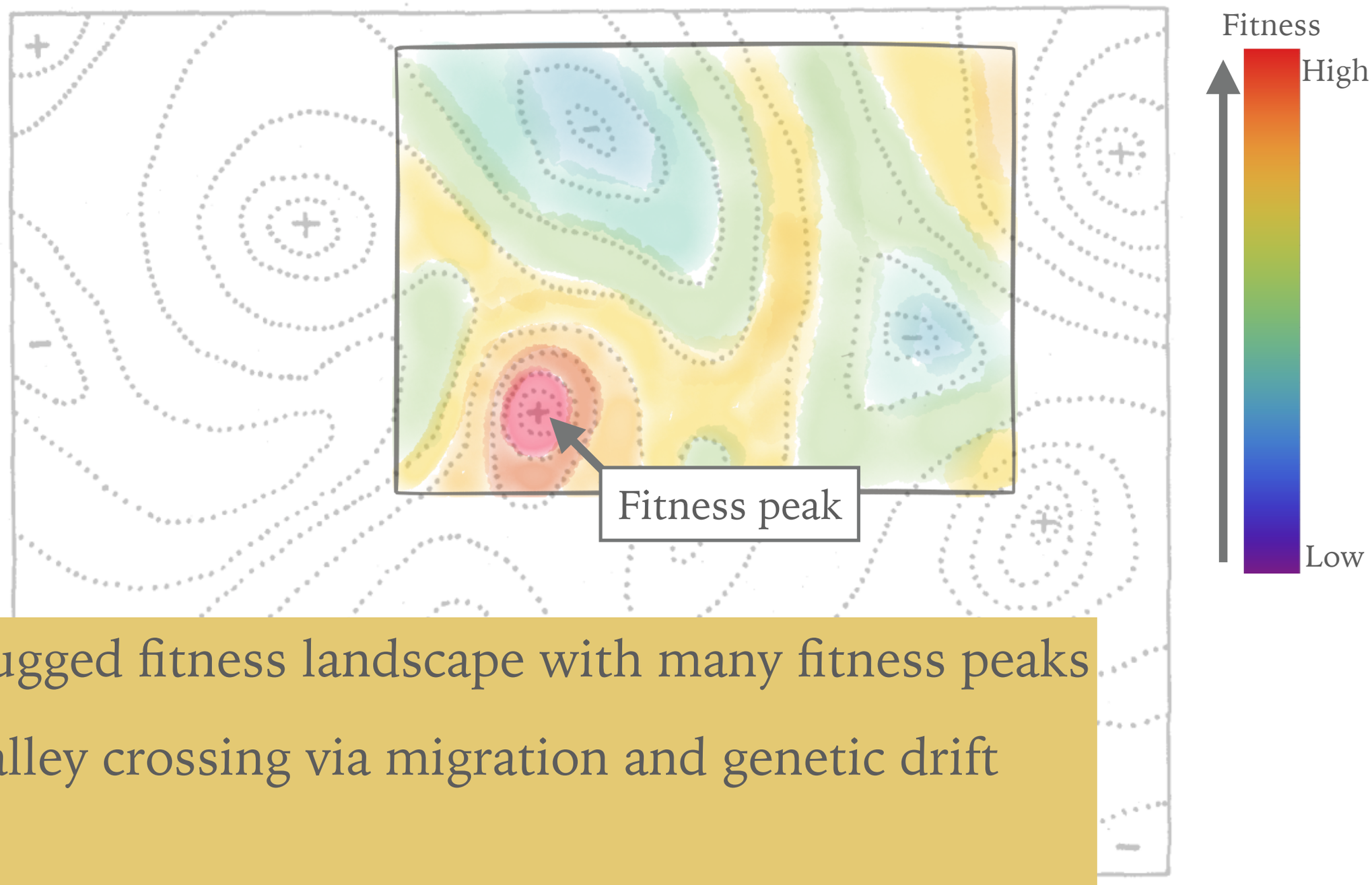
- Philosophy: Large populations, a single fitness optimum

WRIGHT'S SHIFTING BALANCE



Wright, 1932

WRIGHT'S SHIFTING BALANCE



- Rugged fitness landscape with many fitness peaks
- Valley crossing via migration and genetic drift
- Philosophy: Small, structured populations

Wright, 1932

Which team are you on,
Team Fisher or Team Wright, and why?

WHAT WE WANT TO KNOW ABOUT ADAPTATION

- How big/small are adaptive steps?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- How do mutational effects change dependent on the environment?

WHAT WE WANT TO KNOW ABOUT ADAPTATION

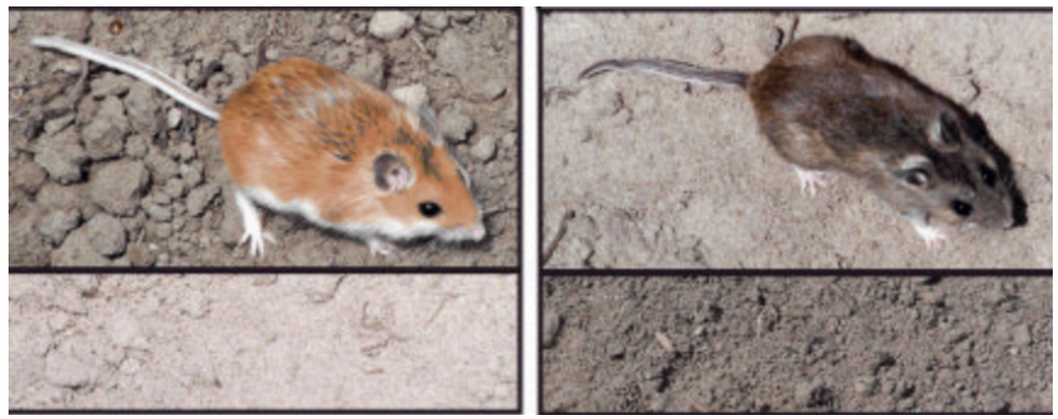
- How big/small are adaptive steps?
- What are the proportions of beneficial, neutral, and deleterious mutations?
- How do mutational effects change dependent on the environment?
- What is the shape of the distribution of fitness effects (DFE)?

ESTIMATES OF MEAN BENEFICIAL EFFECT SIZE FROM POLYMORPHISM DATA

- s=0.002 (Li and Stephan 2006; Jensen et al. 2008)
- s=0.01 (MacPherson et al. 2008)
- s=0.00001 (Andolfatto 2007)

ESTIMATES OF MEAN BENEFICIAL EFFECT SIZE FROM POLYMORPHISM DATA

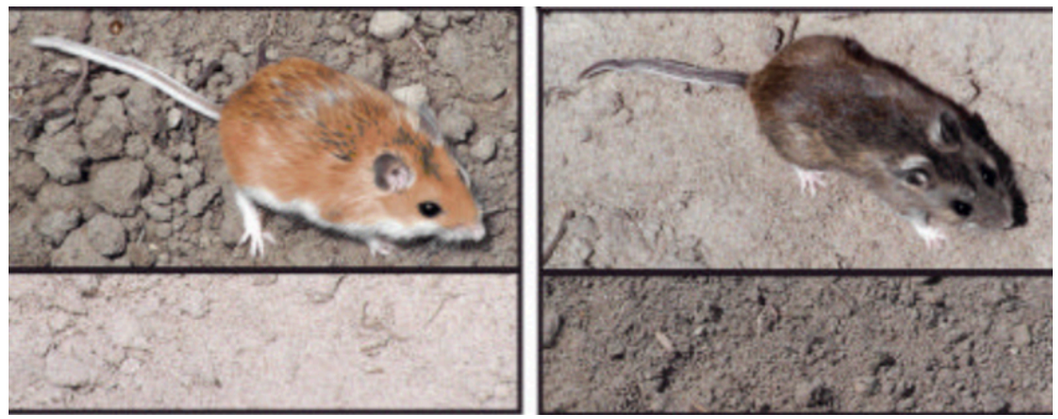
- s=0.002 (Li and Stephan 2006; Jensen et al. 2008)
- s=0.01 (MacPherson et al. 2008)
- s=0.00001 (Andolfatto 2007)



- Example for known phenotype: s=0.102 (Linnen et al. 2009)

ESTIMATES OF MEAN BENEFICIAL EFFECT SIZE FROM POLYMORPHISM DATA

- s=0.002 (Li and Stephan 2006; Jensen et al. 2008)
- s=0.01 (MacPherson et al. 2008)
- s=0.00001 (Andolfatto 2007)



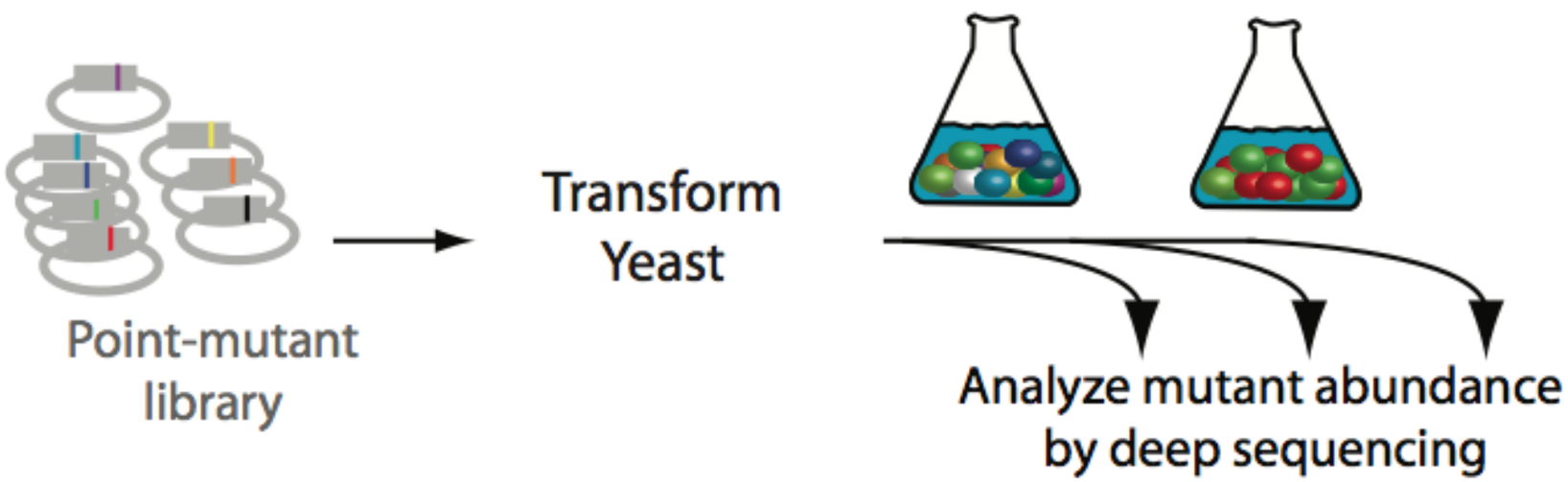
- Example for known phenotype: s=0.102 (Linnen et al. 2009)
- Use experimental approaches to get estimates for whole DFE

AN EXPERIMENTAL APPROACH TO THE DFE: DEEP MUTATIONAL SCANNING

Dan Bolon

Ryan Hietpas

Jeff Jensen



Hietpas, Jensen & Bolon, PNAS, 2011

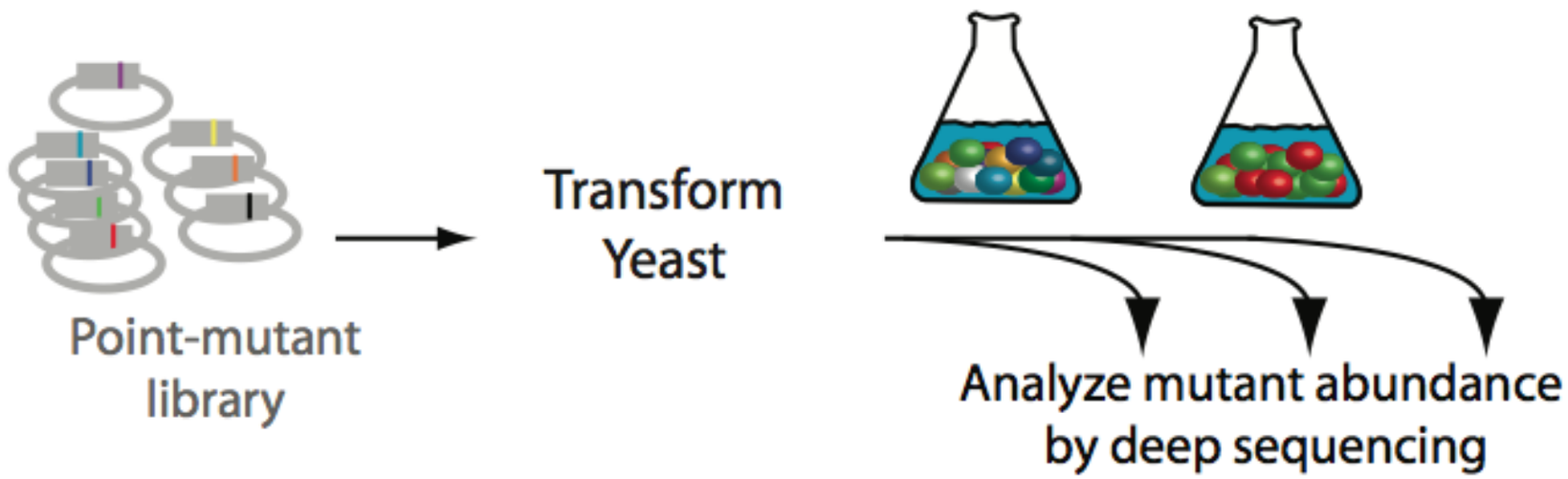
AN EXPERIMENTAL APPROACH TO THE DFE: DEEP MUTATIONAL SCANNING

- 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)

Dan Bolon

Ryan Hietpas

Jeff Jensen



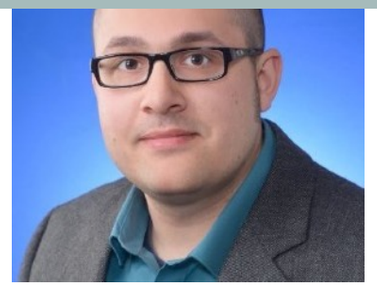
Hietpas, Jensen & Bolon, PNAS, 2011

AN EXPERIMENTAL APPROACH TO THE DFE: DEEP MUTATIONAL SCANNING

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

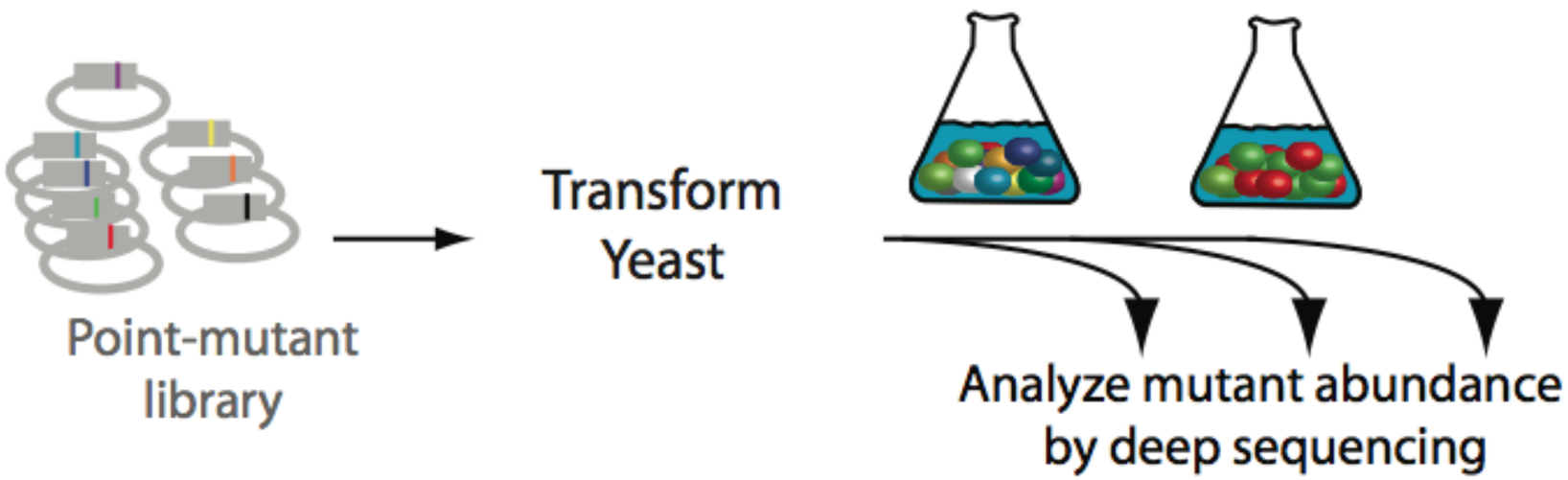
Deep mutational scanning results in a (almost "evolution-free") snapshot of the DFE

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



Ryan Hietpas

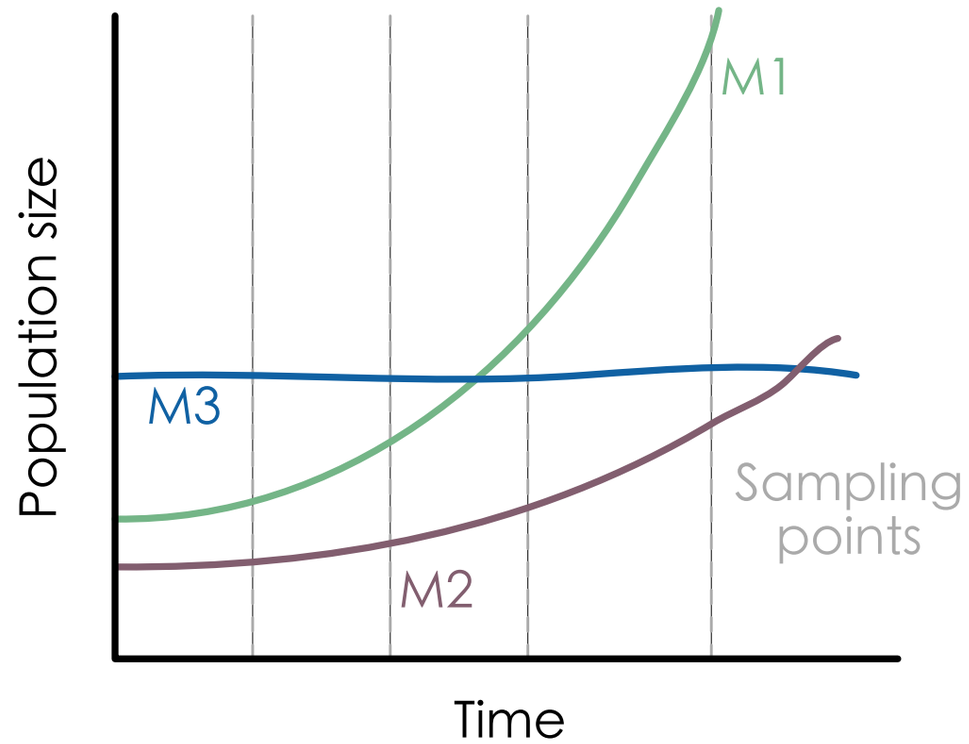
Jeff Jensen



Hietpas, Jensen & Bolon, PNAS, 2011

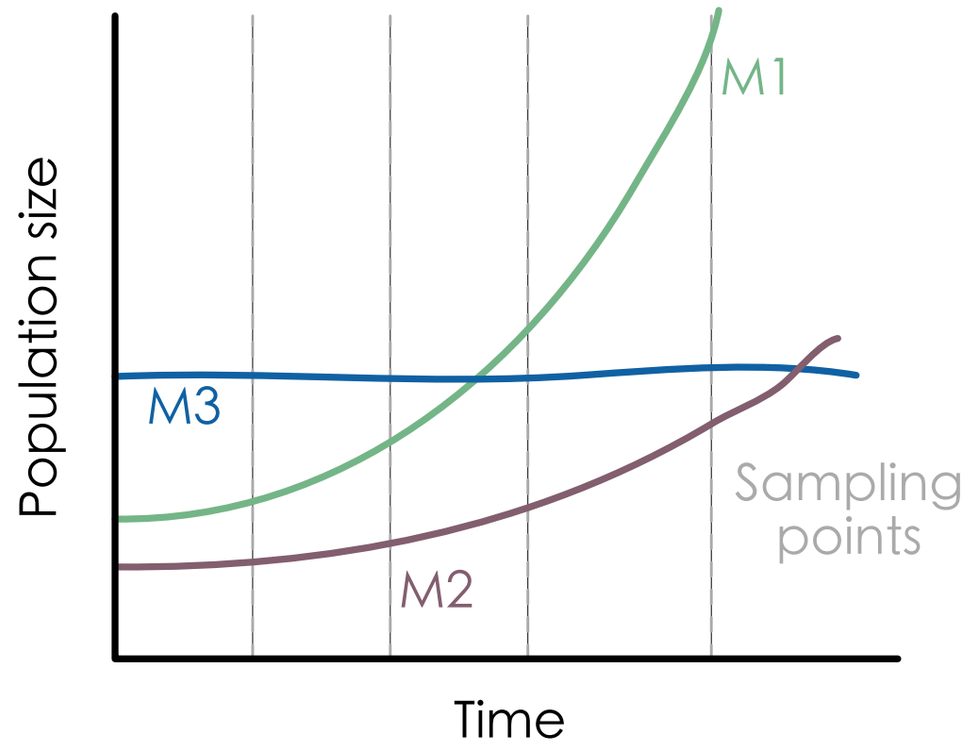
DEEP MUTATIONAL SCANNING FROM A MODELER'S POINT OF VIEW

- 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

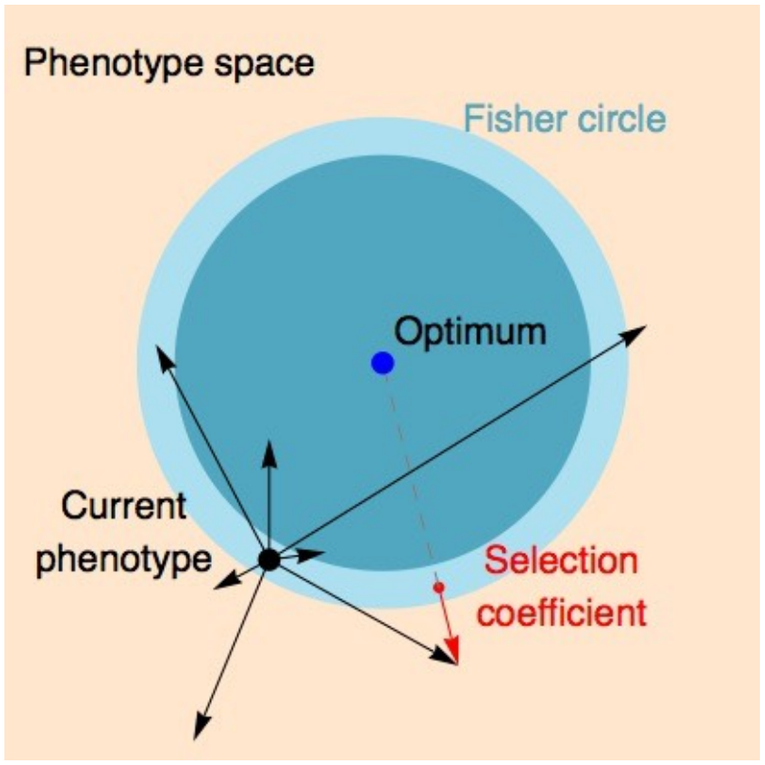
- Exponential growth of hundreds of mutants, each with its own growth rate/selection coefficient
- Sequencing corresponds to multinomial sampling of mutants independently at each sampling time



- <1% fitness differences detectable

For the “Fisherians”: the shape of the DFE across environments

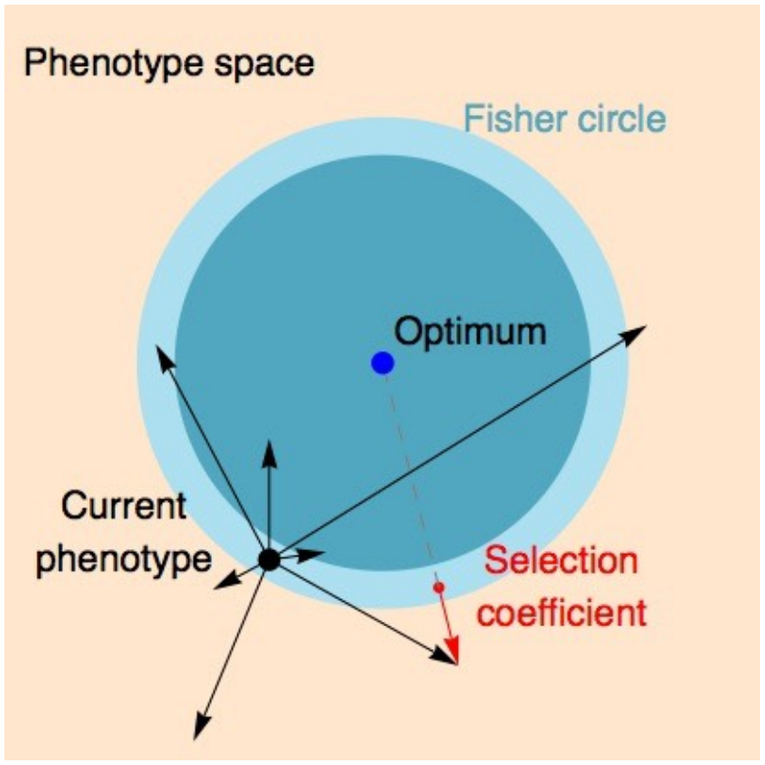
FISHER'S GEOMETRIC MODEL



Fisher, 1930

Hypotheses:

FISHER'S GEOMETRIC MODEL

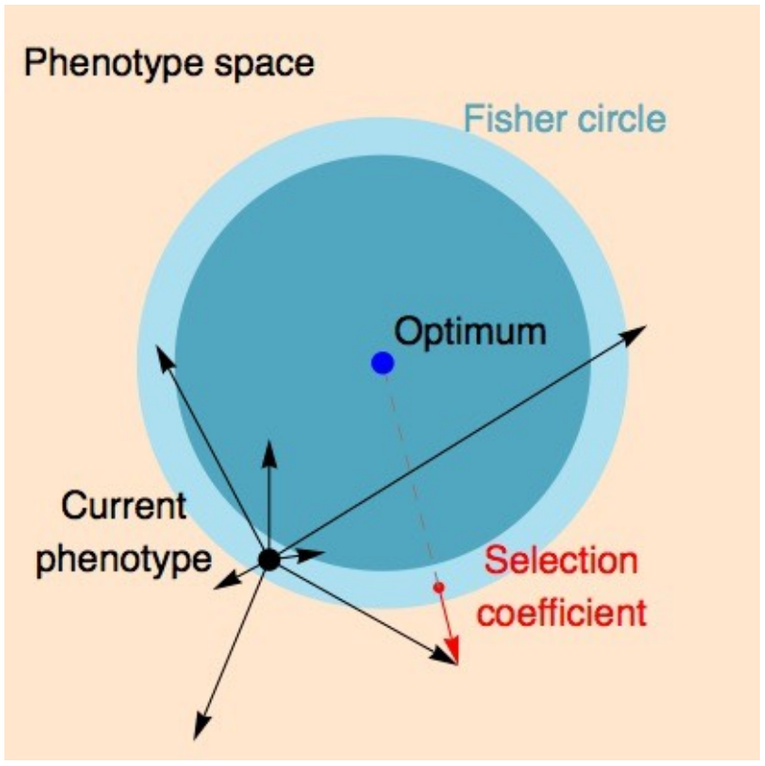


Fisher, 1930

Hypotheses:

- Relocation of the optimum or the current phenotype in a new environment can increase the distance to the optimum and hence the potential for beneficials.

FISHER'S GEOMETRIC MODEL



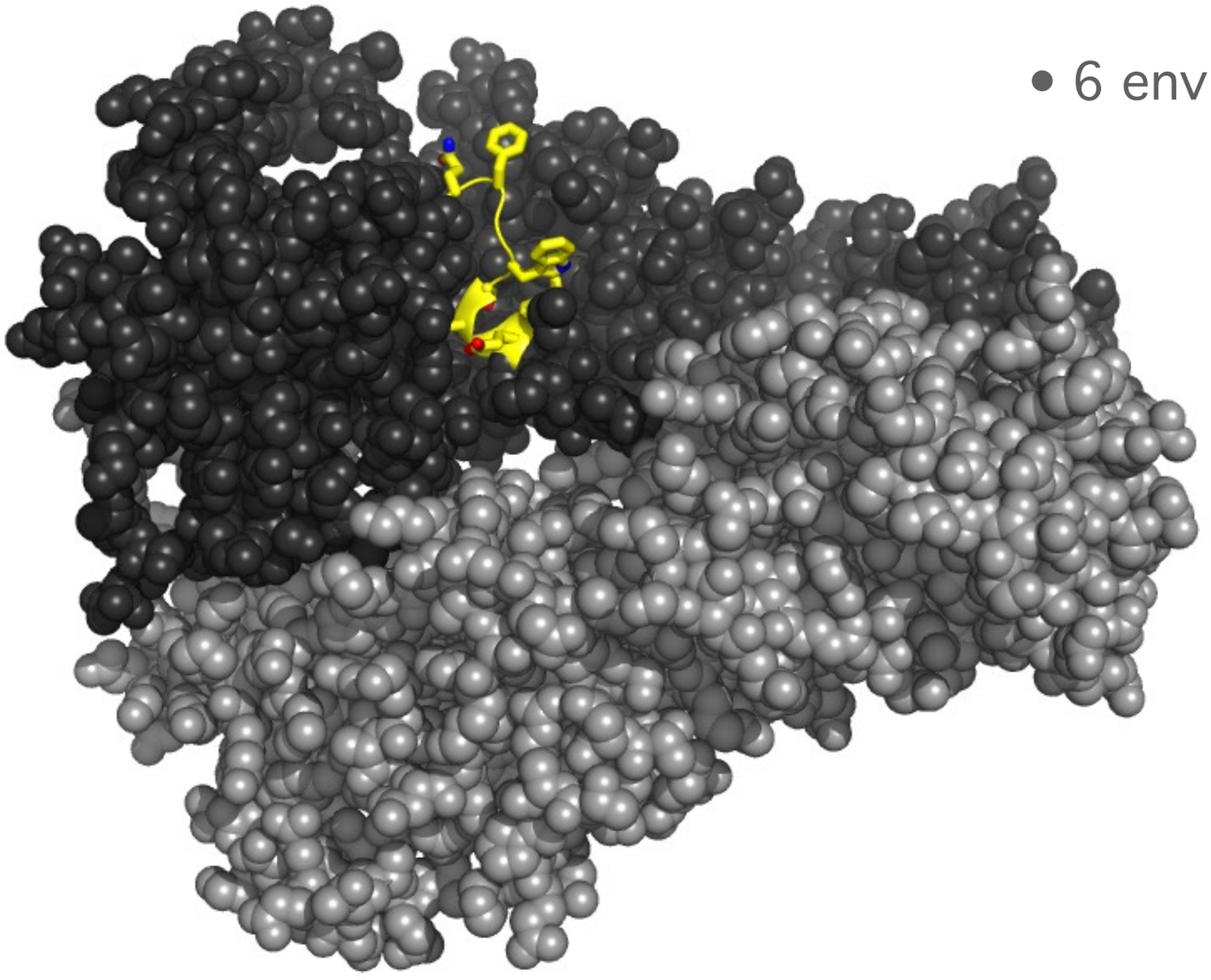
Fisher, 1930

Hypotheses:

- Relocation of the optimum or the current phenotype in a new environment can increase the distance to the optimum and hence the potential for beneficials.
- The distribution of beneficial mutations is bounded or exponential.

THE SHAPE OF THE DFE IN CHALLENGING ENVIRONMENTS

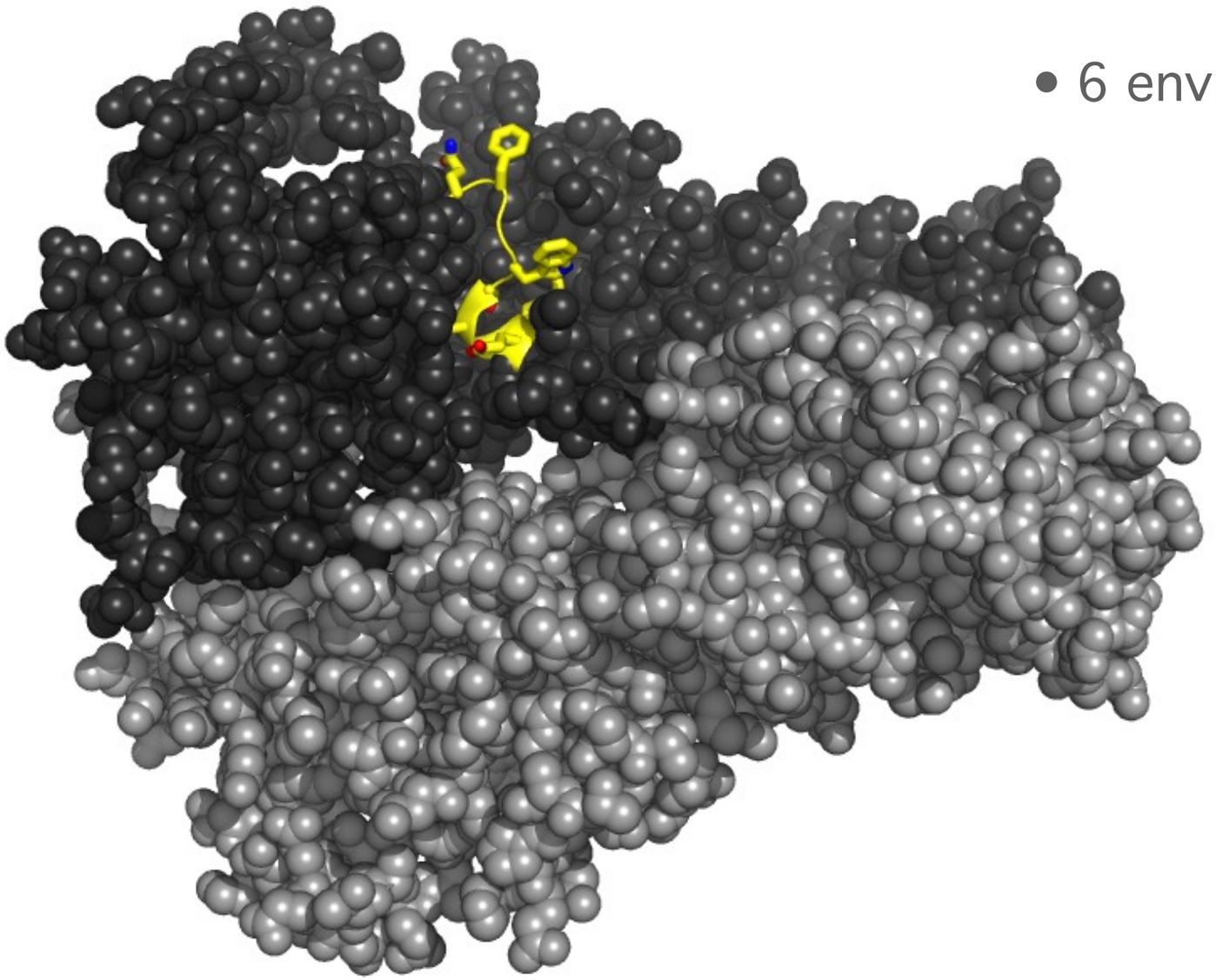
The data set



- 9 aa region from Hsp90 (aa positions 582-590) in Saccharomyces cerevisiae
- 6 environments:

Data obtained by Ryan Hietpas @ UMassMed

The data set



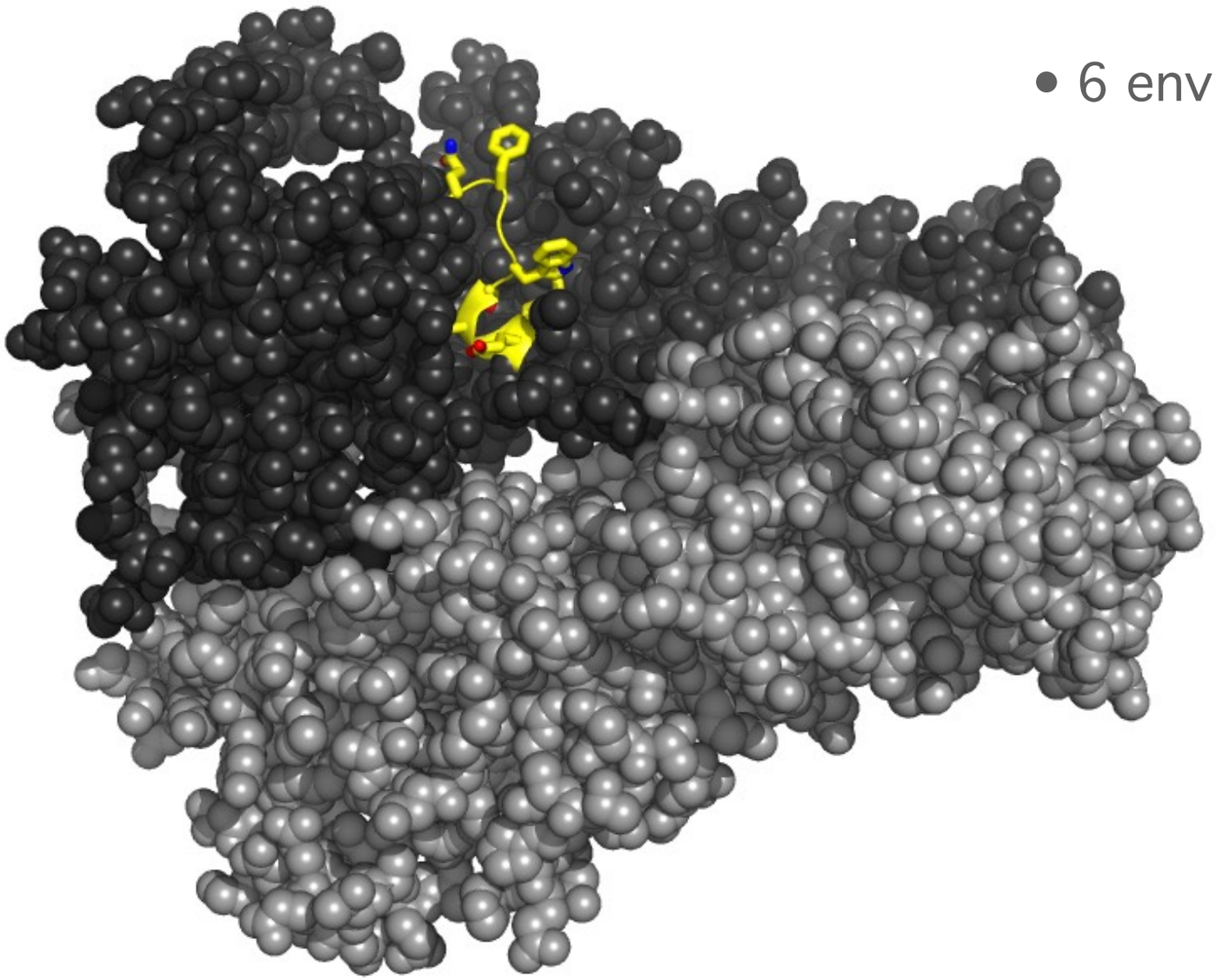
- 9 aa region from Hsp90 (aa positions 582-590) in Saccharomyces cerevisiae

- 6 environments:

30°C	30°C+0.5M NaCl
36°C	36°C+0.5M NaCl
25°C	25°C+0.5M NaCl

Data obtained by Ryan Hietpas @ UMassMed

The data set



- 9 aa region from Hsp90 (aa positions 582-590) in Saccharomyces cerevisiae

- 6 environments:

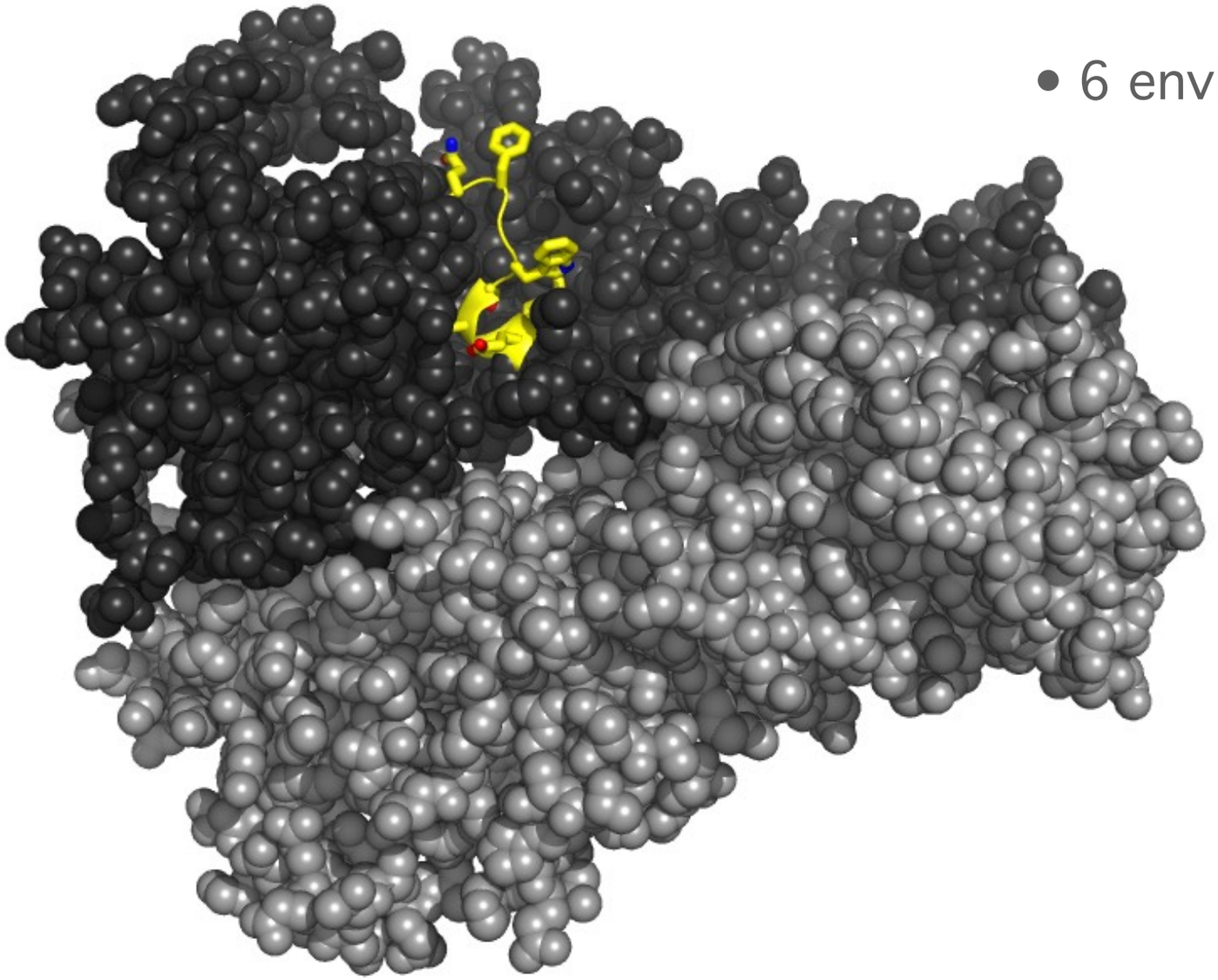
30°C	30°C+0.5M NaCl
36°C	36°C+0.5M NaCl
25°C	25°C+0.5M NaCl

Relative growth of wt:

1	0,45
0,83	0,33
0,63	0,3

Data obtained by Ryan Hietpas @ UMassMed

The data set



- 9 aa region from Hsp90 (aa positions 582-590) in Saccharomyces cerevisiae

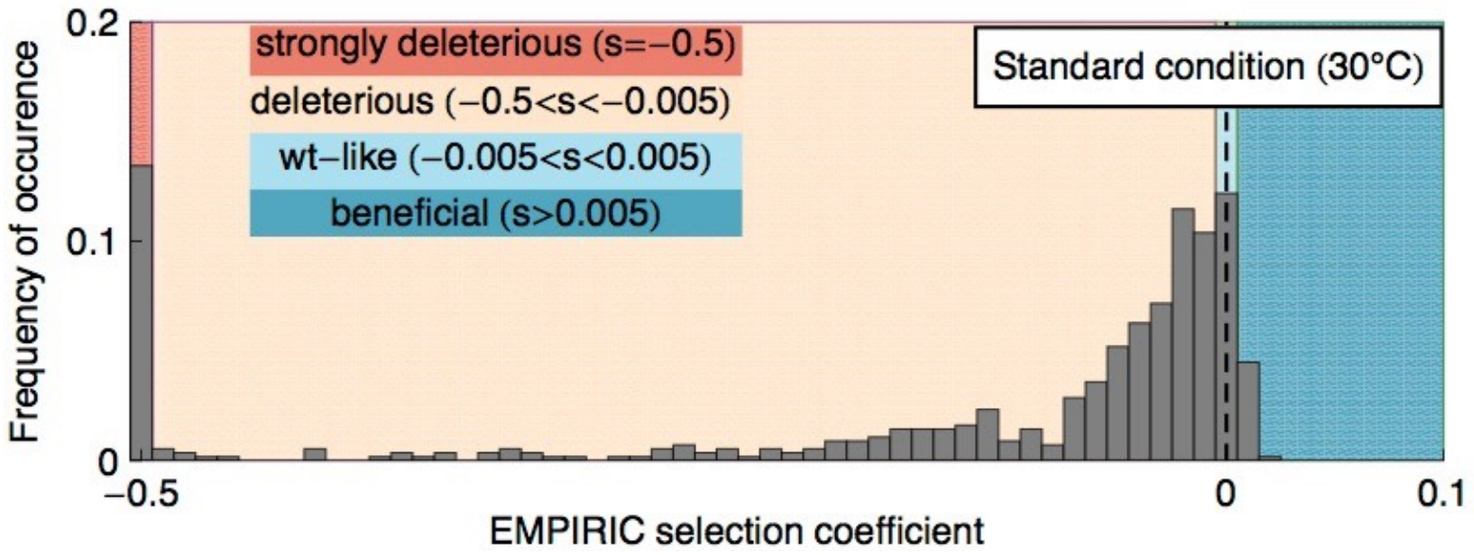
- 6 environments:

30°C	30°C+0.5M NaCl
36°C	36°C+0.5M NaCl
25°C	25°C+0.5M NaCl

- Fitness data for every possible codon at each aa position (i.e. the same 560 mutations per environment)

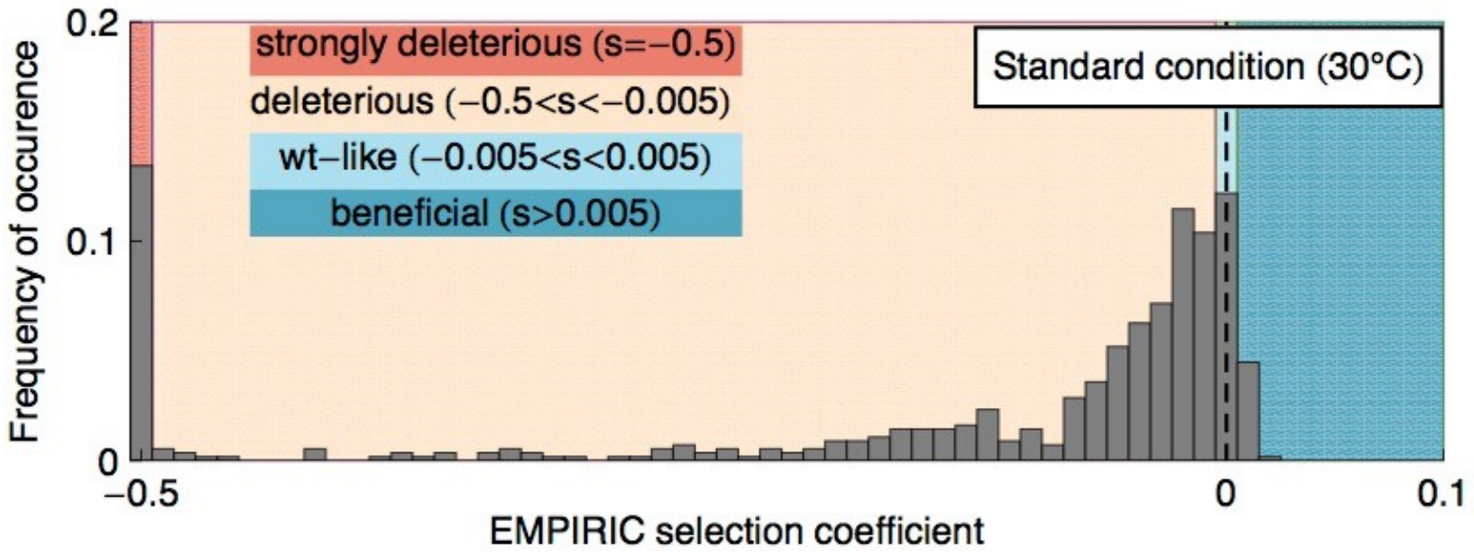
Data obtained by Ryan Hietpas @ UMassMed

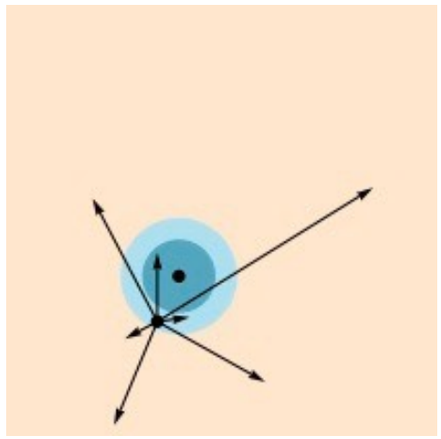
The shape of the full DFE



Hietpas, Bank et al. 2013

The shape of the full DFE

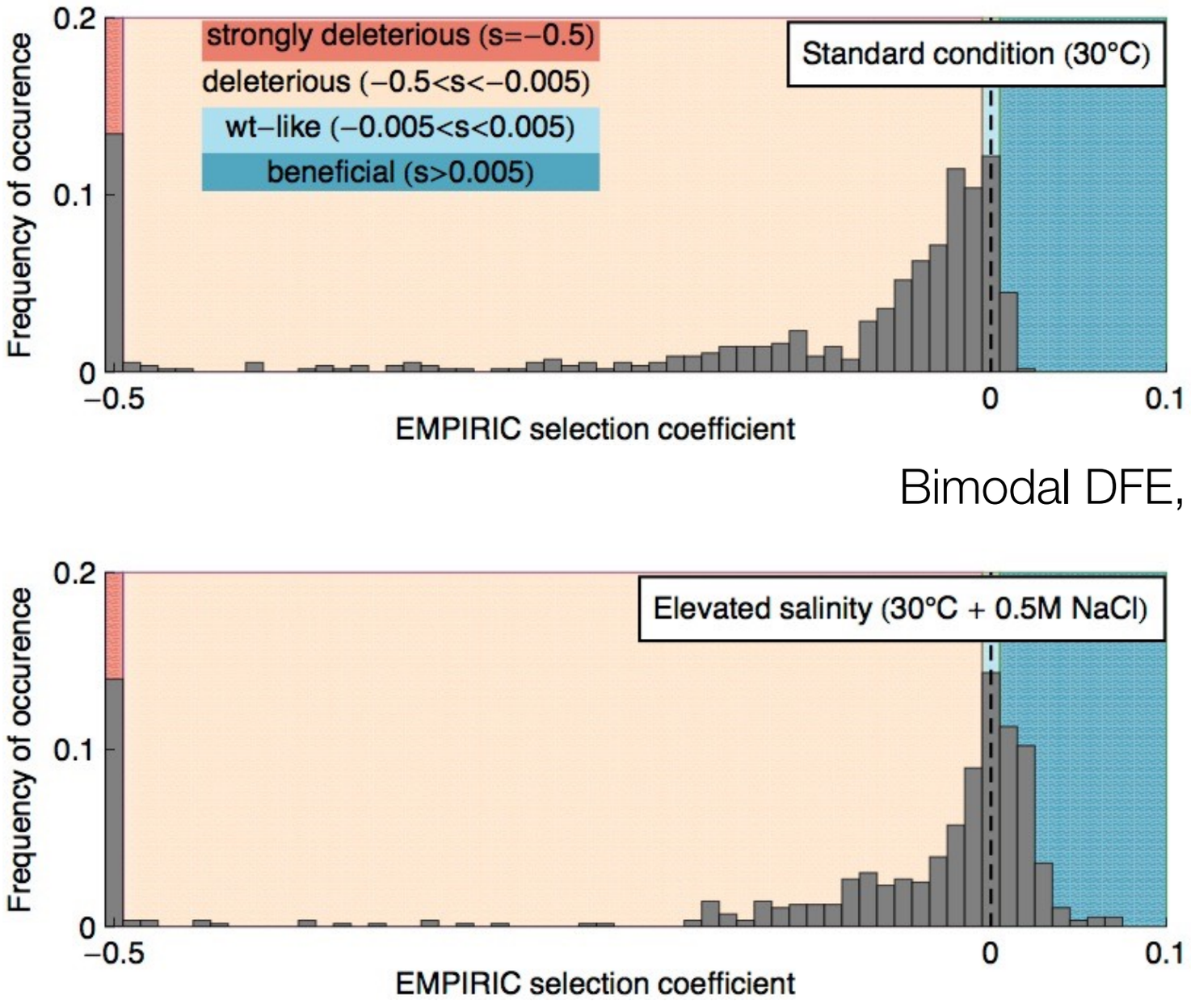


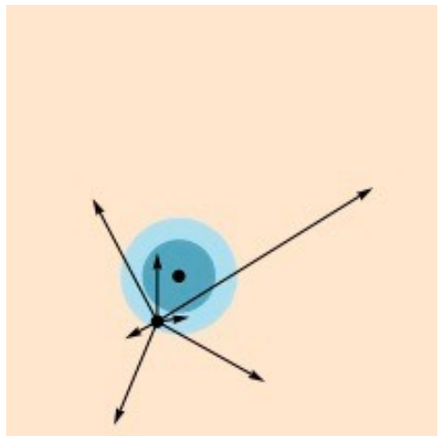


Bimodal DFE, few beneficials - close to optimum

Hietpas, Bank et al. 2013

The shape of the full DFE

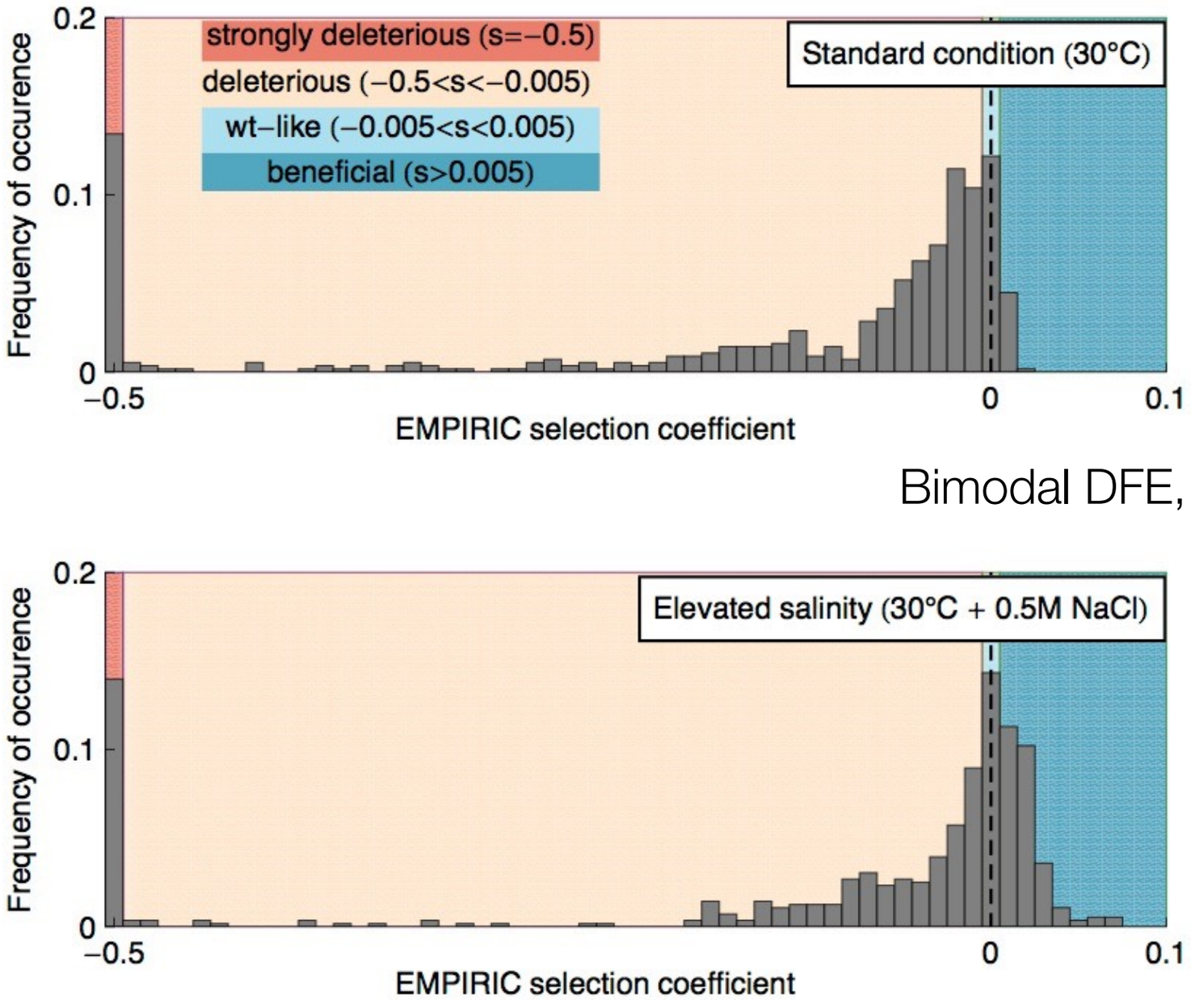


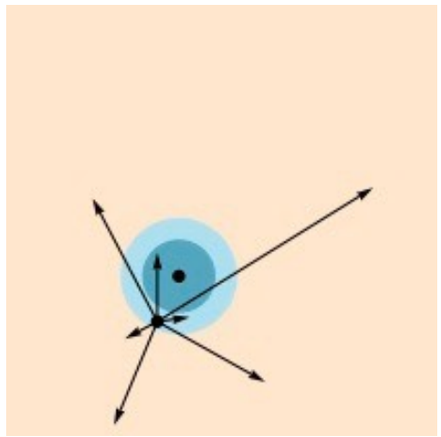


Bimodal DFE, few beneficials - close to optimum

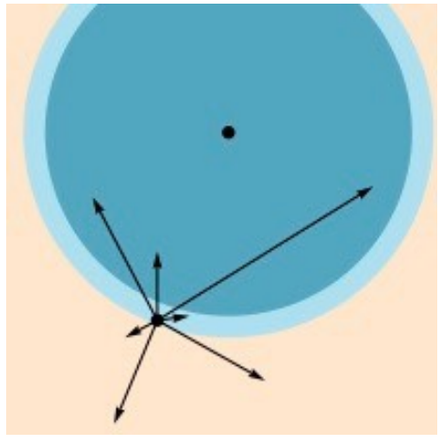
Hietpas, Bank et al. 2013

The shape of the full DFE





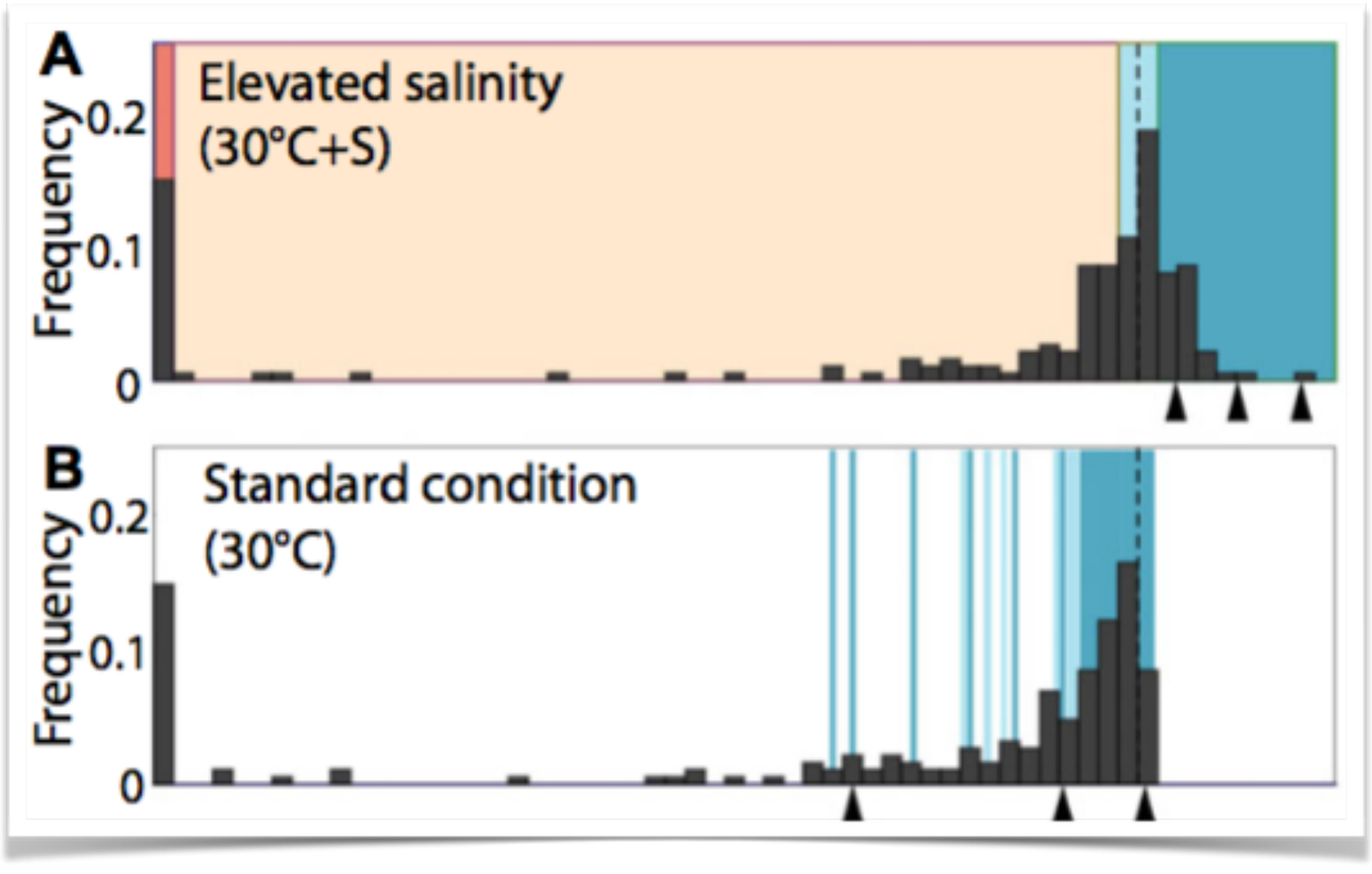
Bimodal DFE, few beneficials - close to optimum



Increased number of beneficials, increased variance - far from optimum

Hietpas, Bank et al. 2013

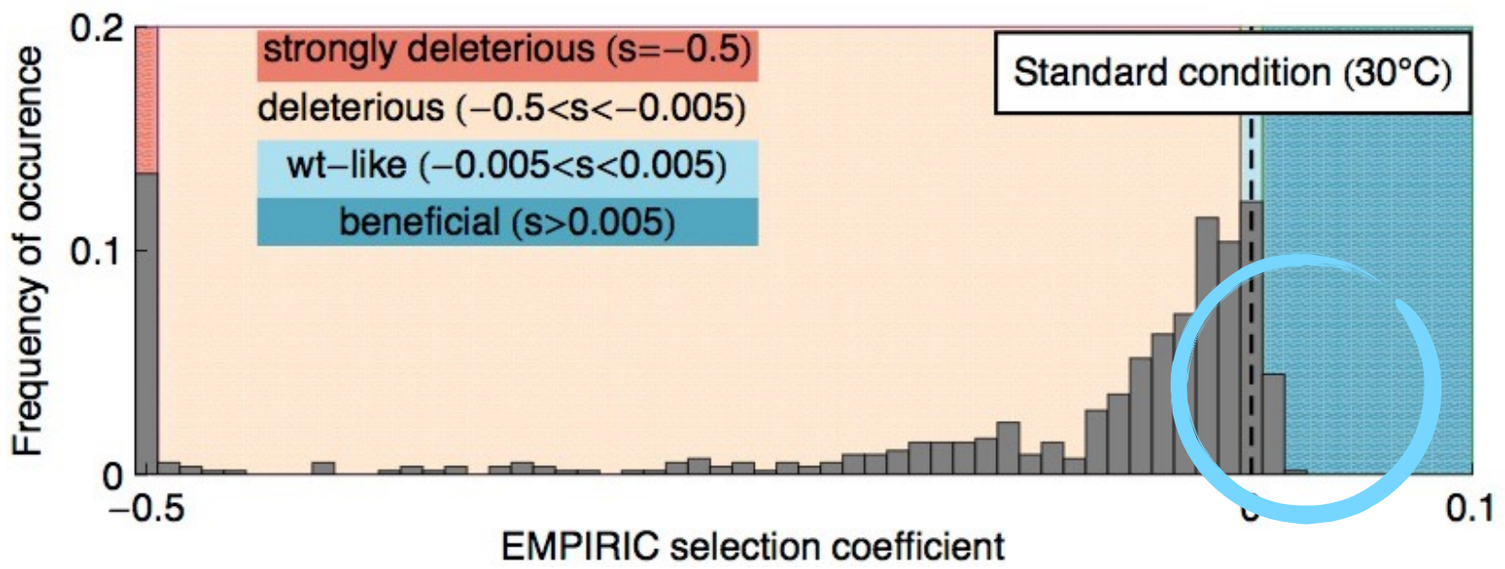
COSTS OF ADAPTATION



THE SHAPE OF THE BENEFICIAL TAIL OF THE DFE

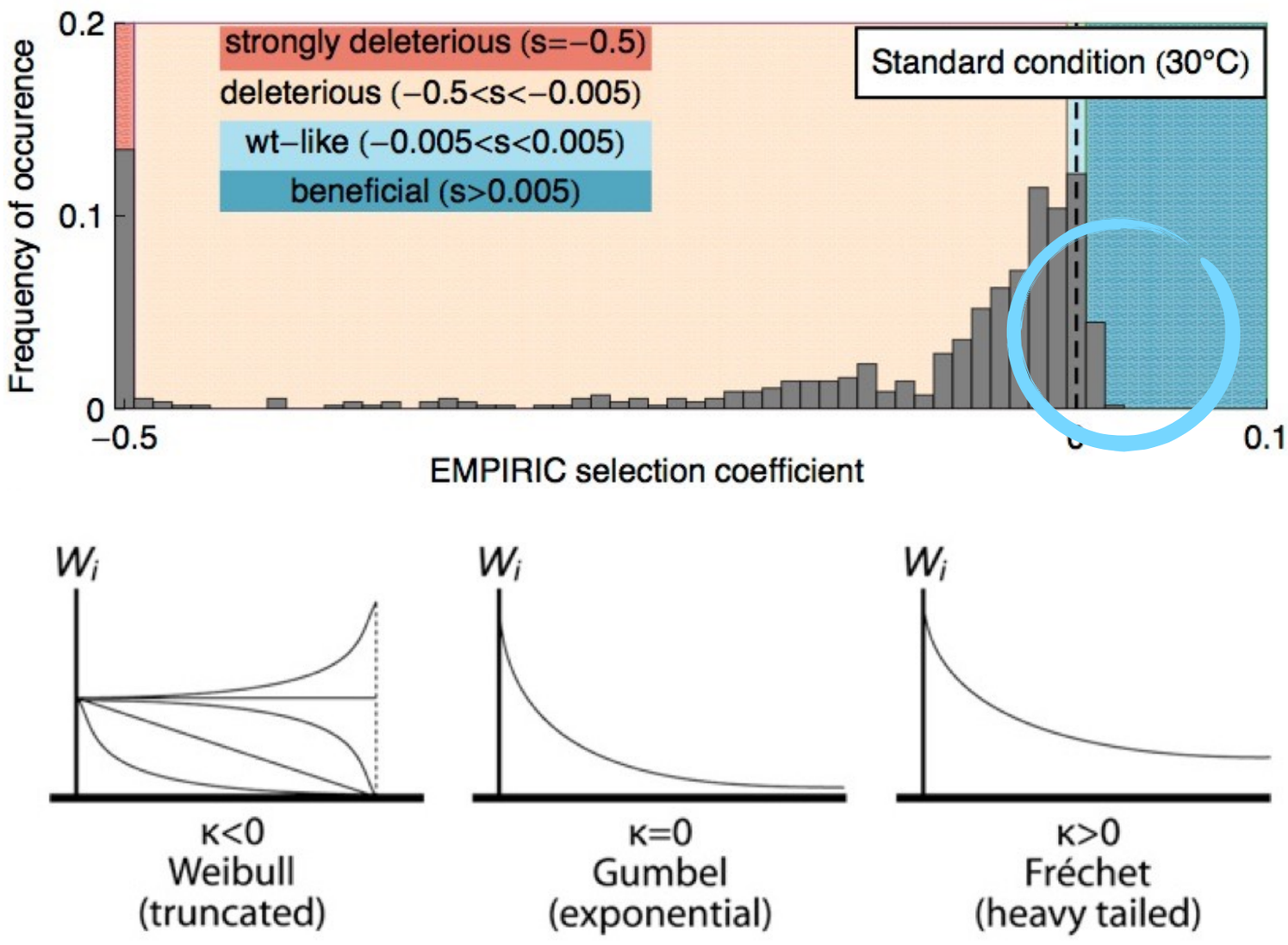
HOW PREDICTABLE IS ADAPTIVE EVOLUTION?

- Fit Generalized Pareto distribution to beneficial tail
- Kappa parameter determines tail shape



HOW PREDICTABLE IS ADAPTIVE EVOLUTION?

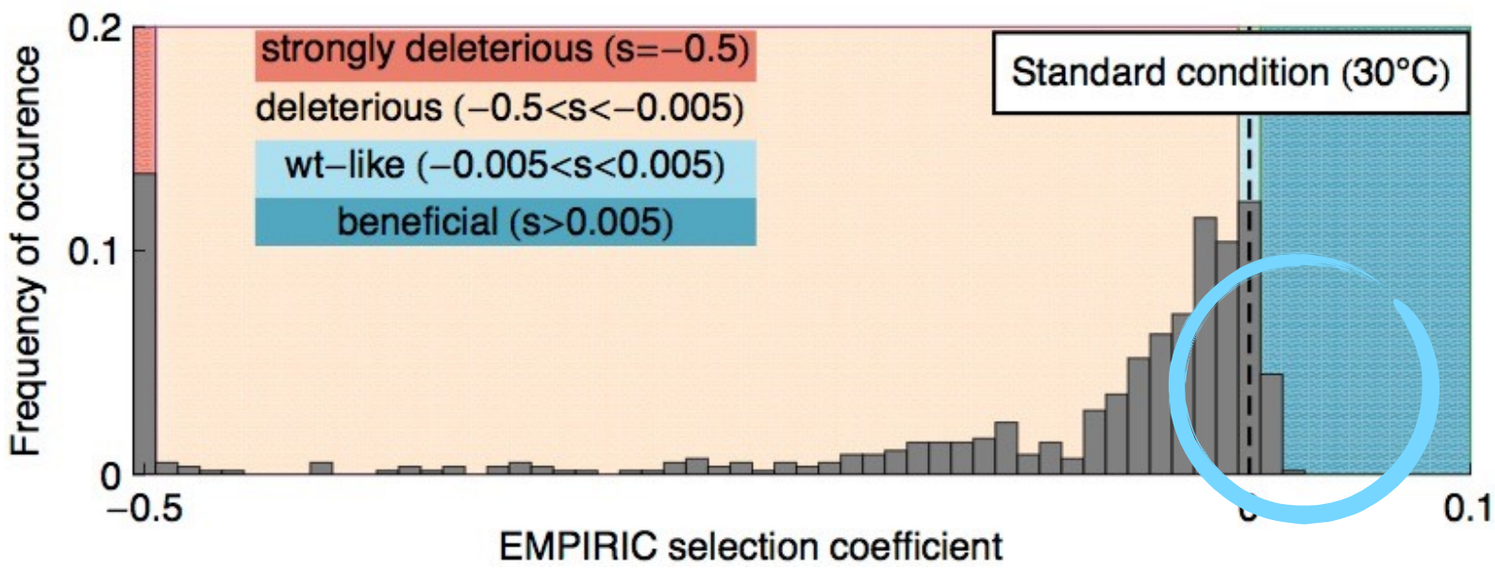
- Fit Generalized Pareto distribution to beneficial tail
- Kappa parameter determines tail shape



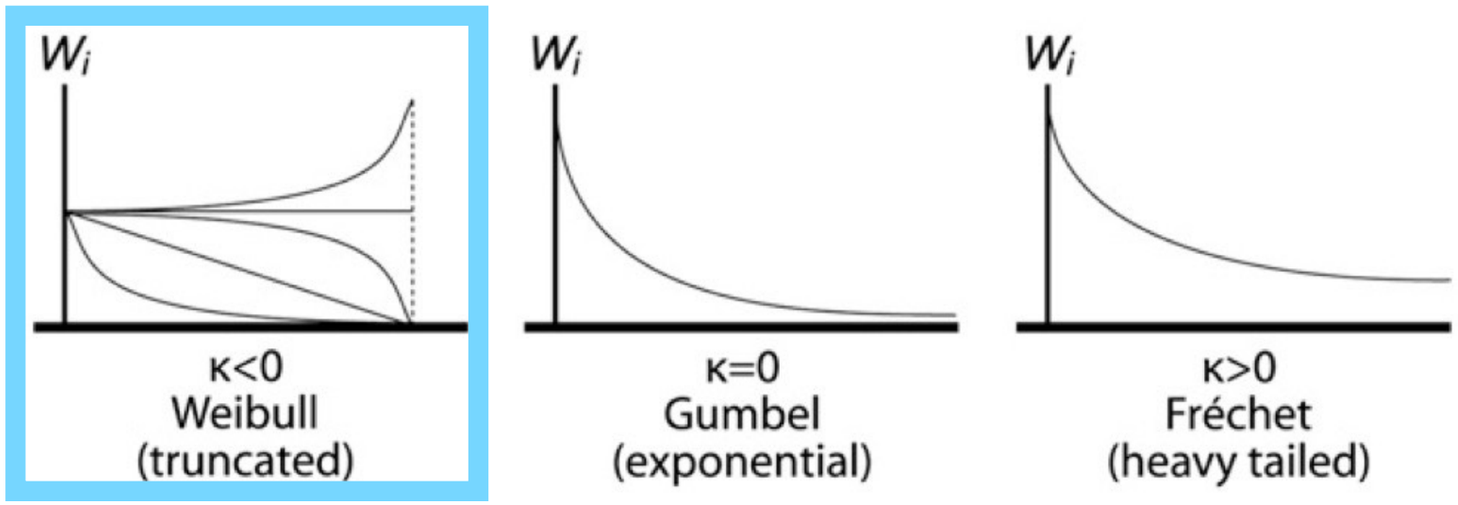
From Beisel et al., Genetics, 2007

HOW PREDICTABLE IS ADAPTIVE EVOLUTION?

- Fit Generalized Pareto distribution to beneficial tail
- Kappa parameter determines tail shape



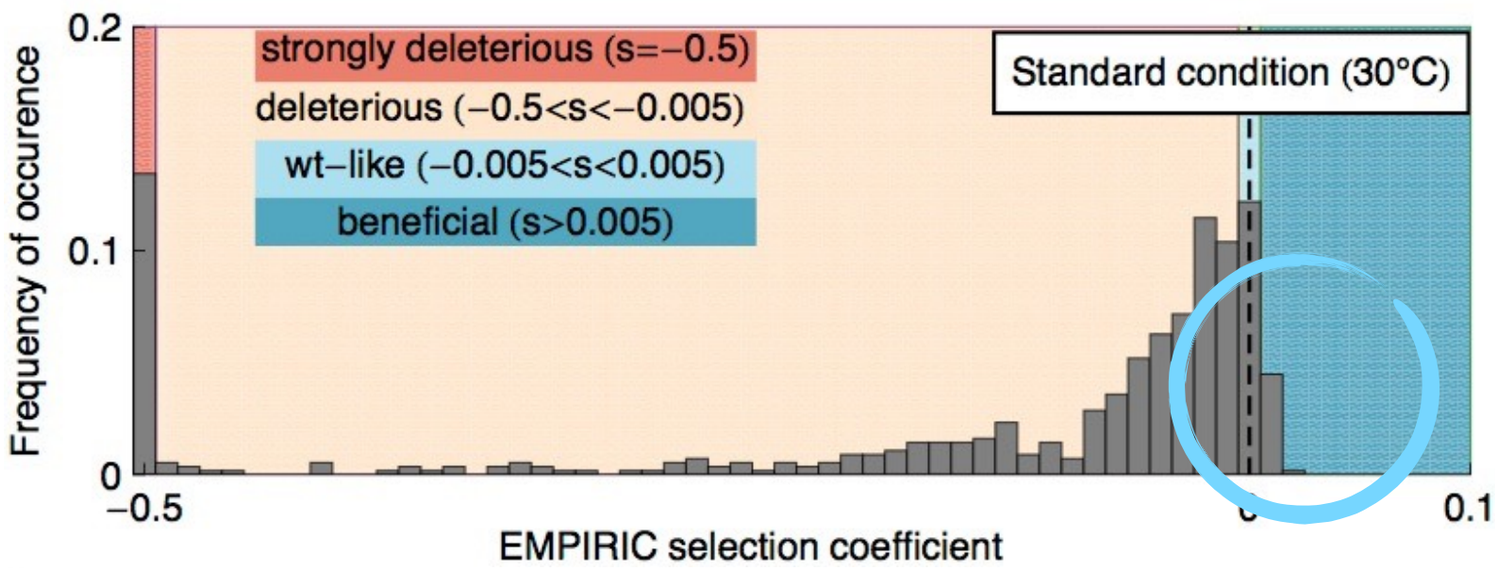
- Bounded DFE, limited potential for adaptation
- Consistent with FGM with close optimum

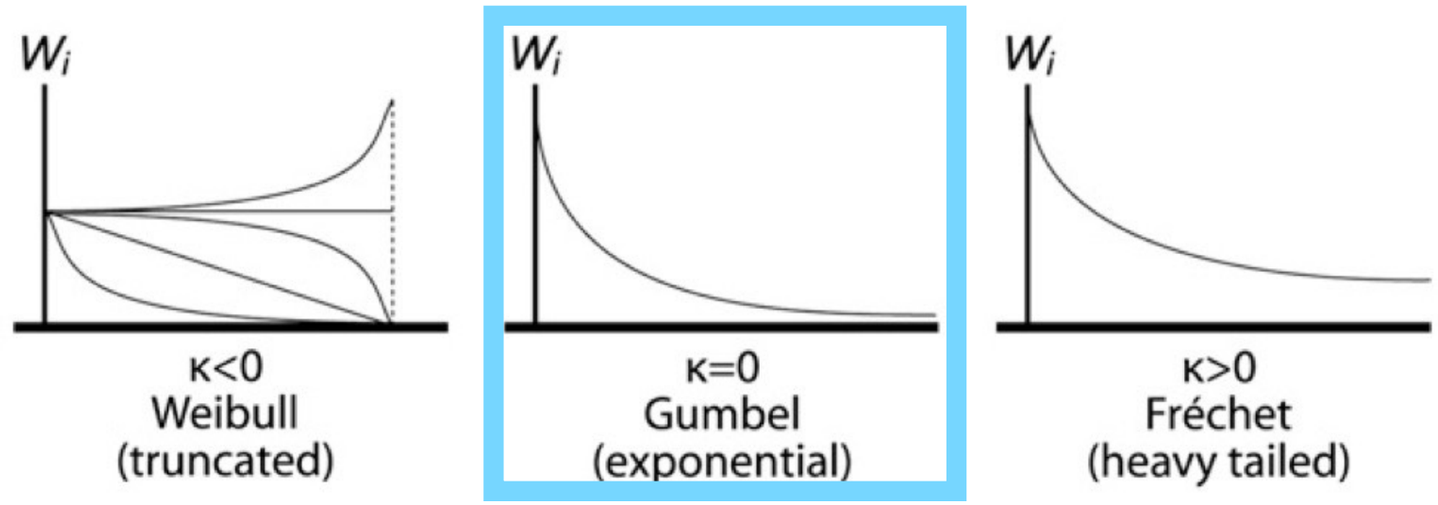


From Beisel et al., Genetics, 2007

HOW PREDICTABLE IS ADAPTIVE EVOLUTION?

- Fit Generalized Pareto distribution to beneficial tail
- Kappa parameter determines tail shape



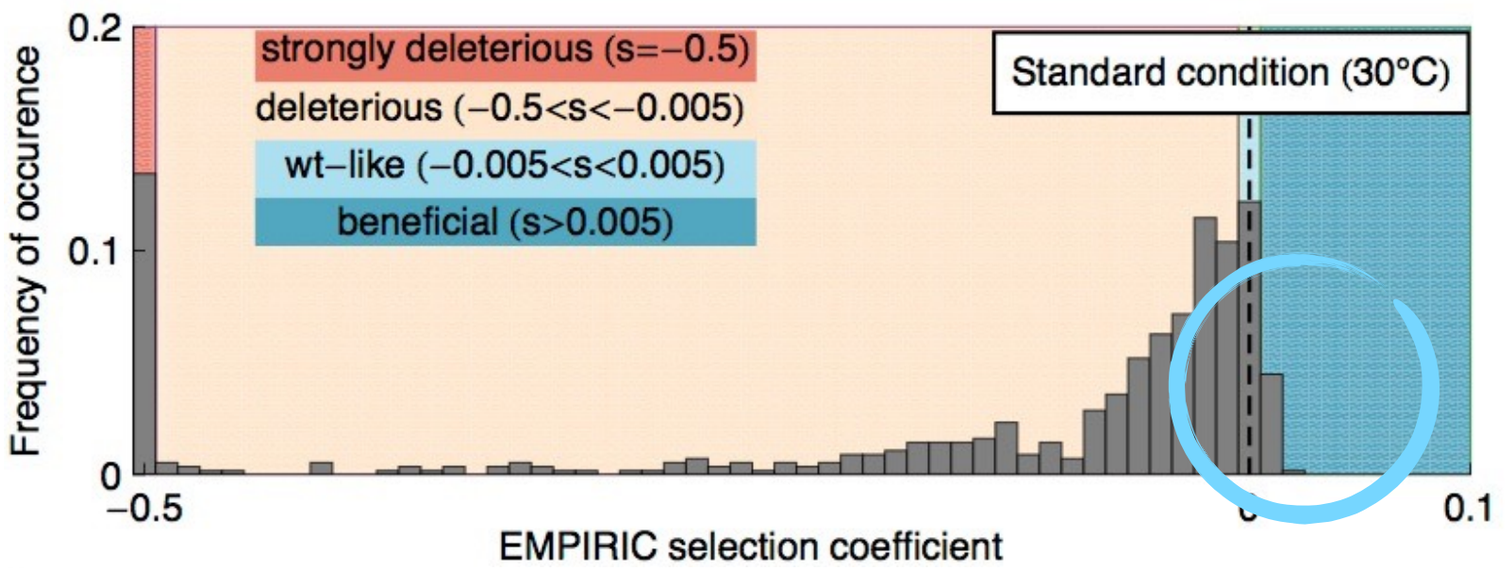


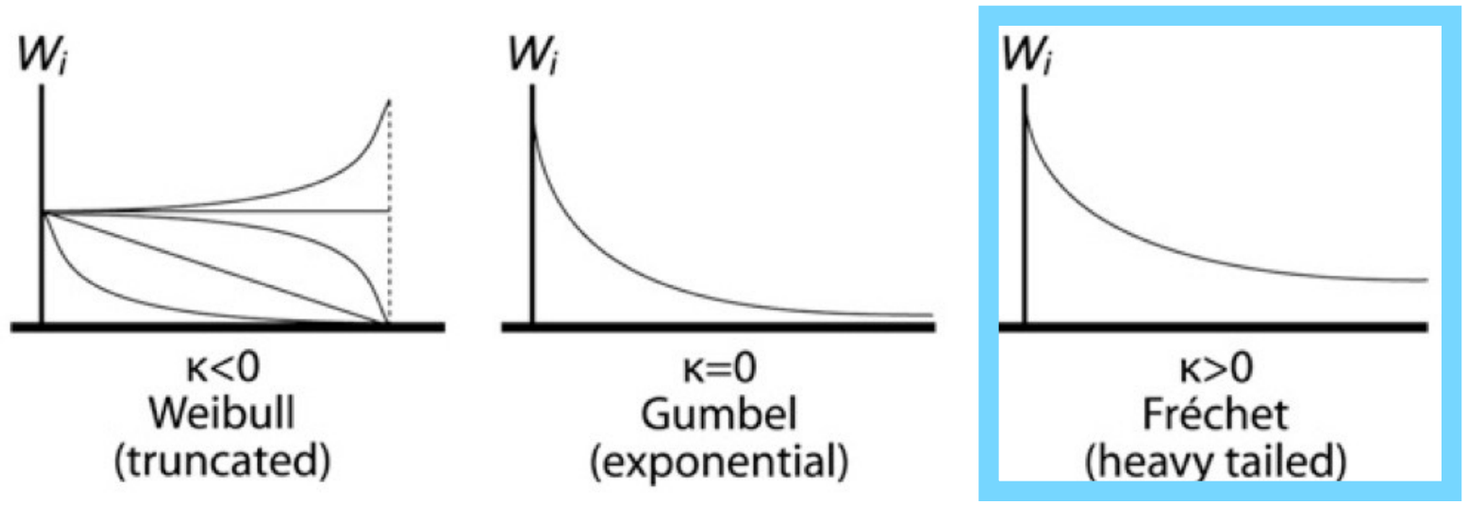
- Unbounded DFE, but low prob. of large-effect mutations
- Consistent with FGM with far optimum

From Beisel et al., Genetics, 2007

HOW PREDICTABLE IS ADAPTIVE EVOLUTION?

- Fit Generalized Pareto distribution to beneficial tail
- Kappa parameter determines tail shape



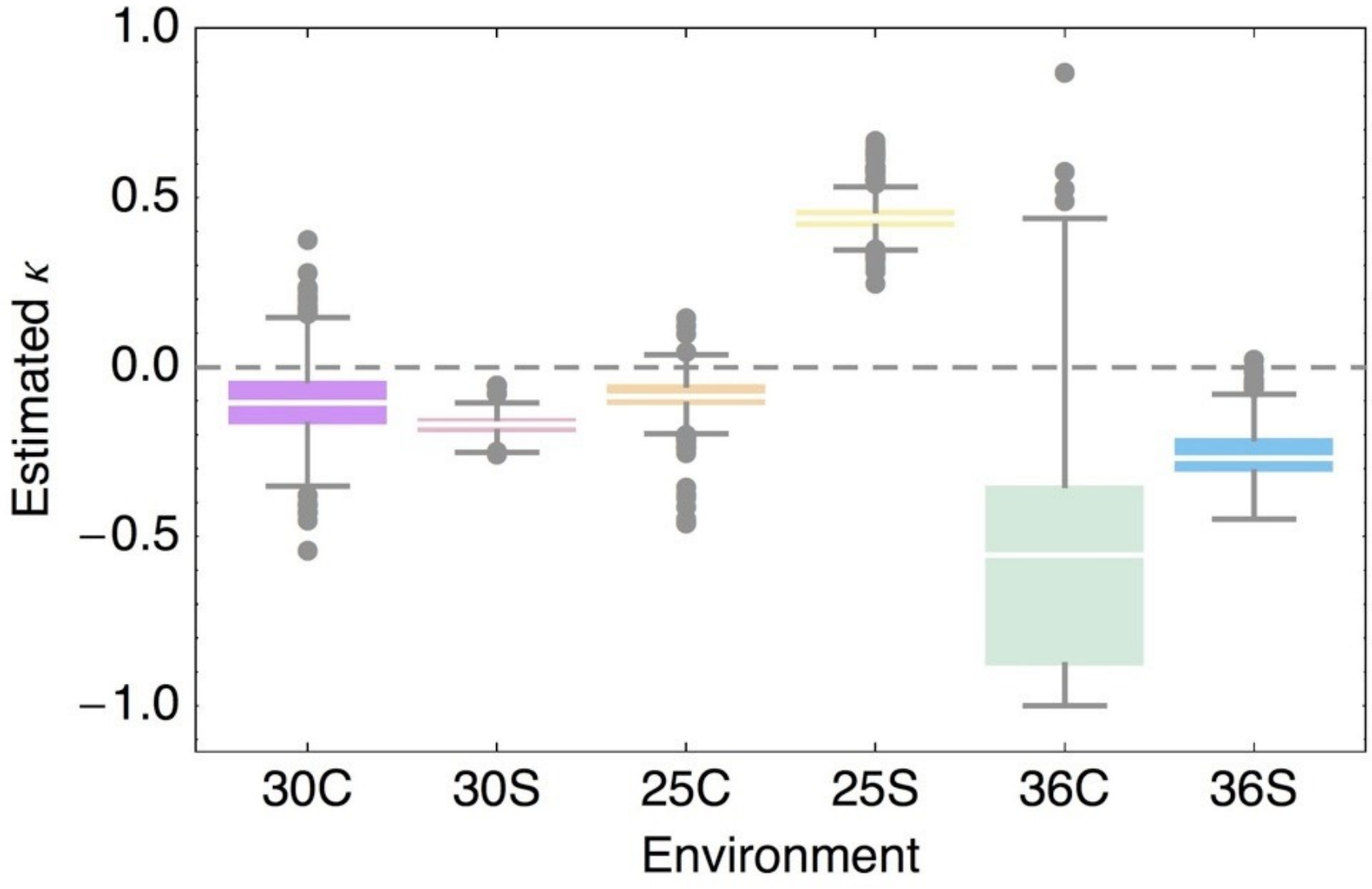


- Unbounded DFE, highly unpredictable mutational effects
- Not captured by FGM

From Beisel et al., Genetics, 2007

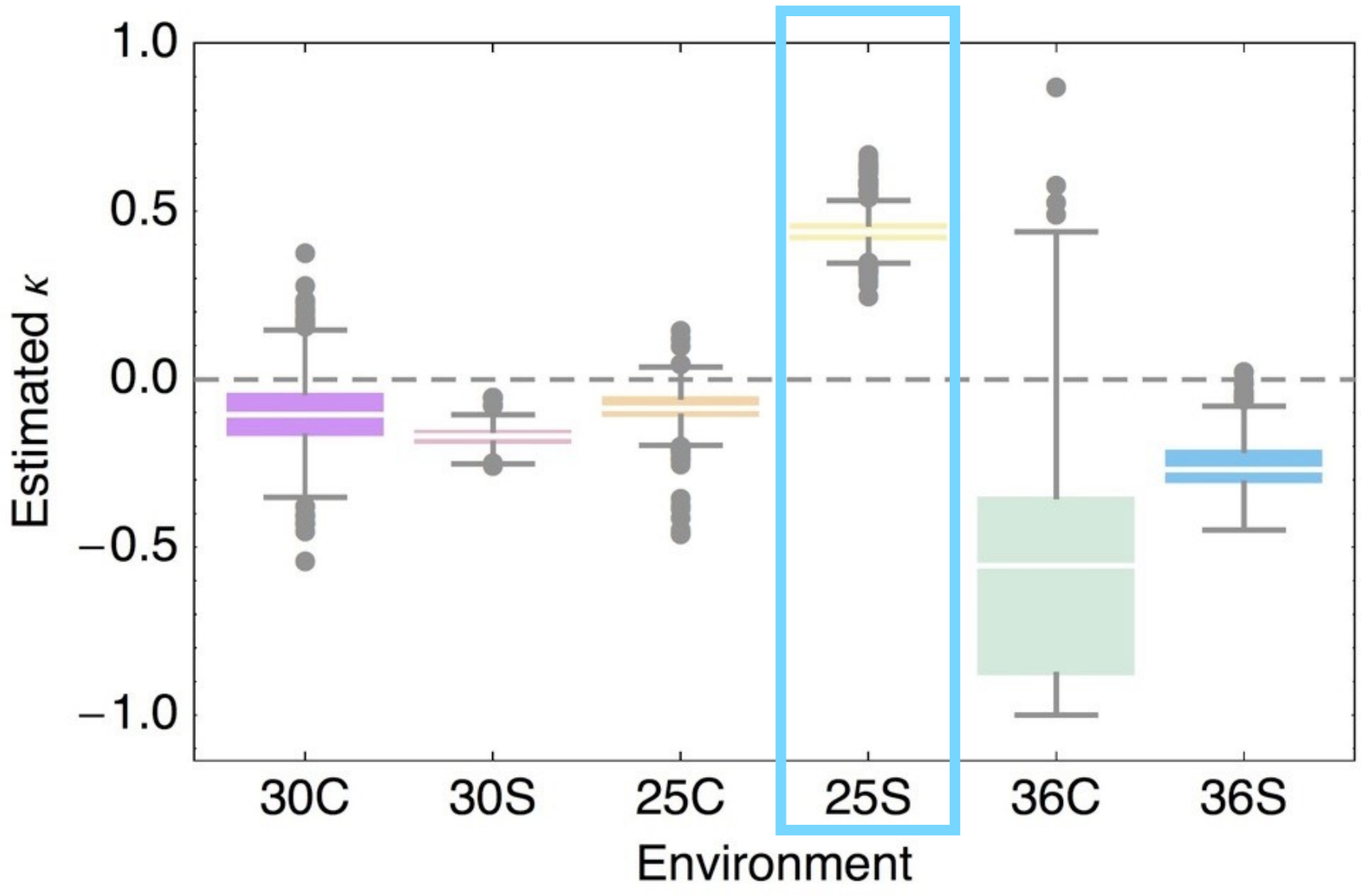
TAIL SHAPE PARAMETER IN CHALLENGING ENVIRONMENTS

S. cerevisiae EMPIRIC data from Hsp90



TAIL SHAPE PARAMETER IN CHALLENGING ENVIRONMENTS

S. cerevisiae EMPIRIC data from Hsp90



Tail shape parameter in challenging environments

An application to resistance evolution

Influenza H1N1 mutation-accumulation experiment

Tail shape parameter in challenging environments

An application to resistance evolution

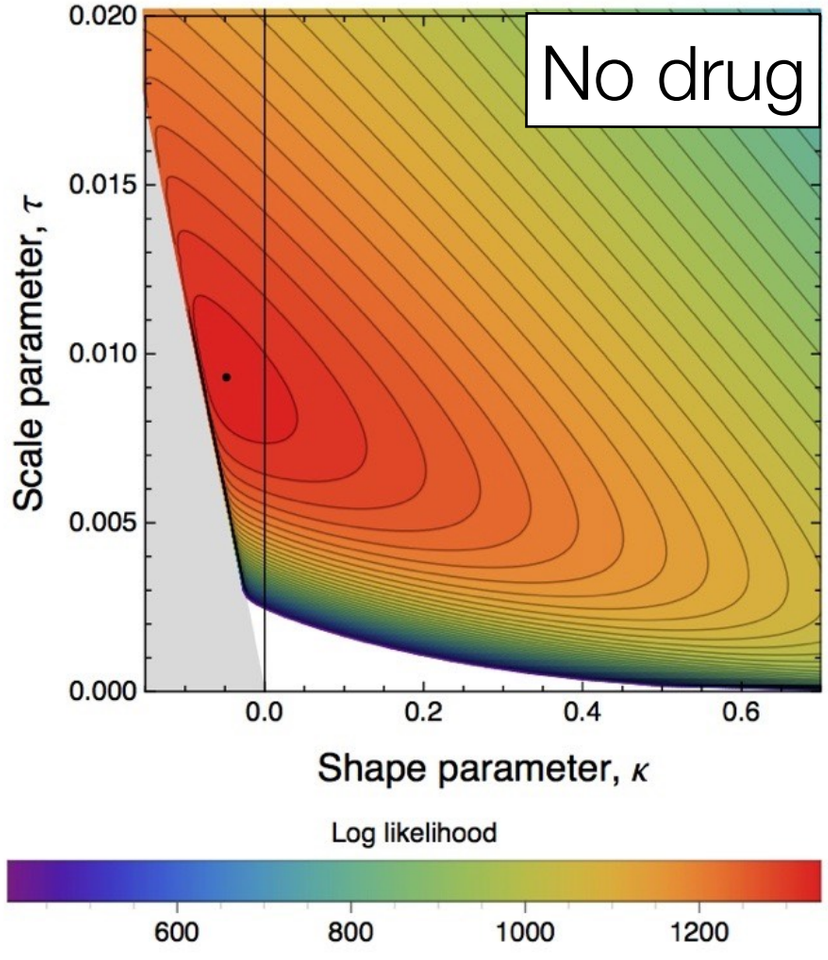
Influenza H1N1 mutation-accumulation experiment

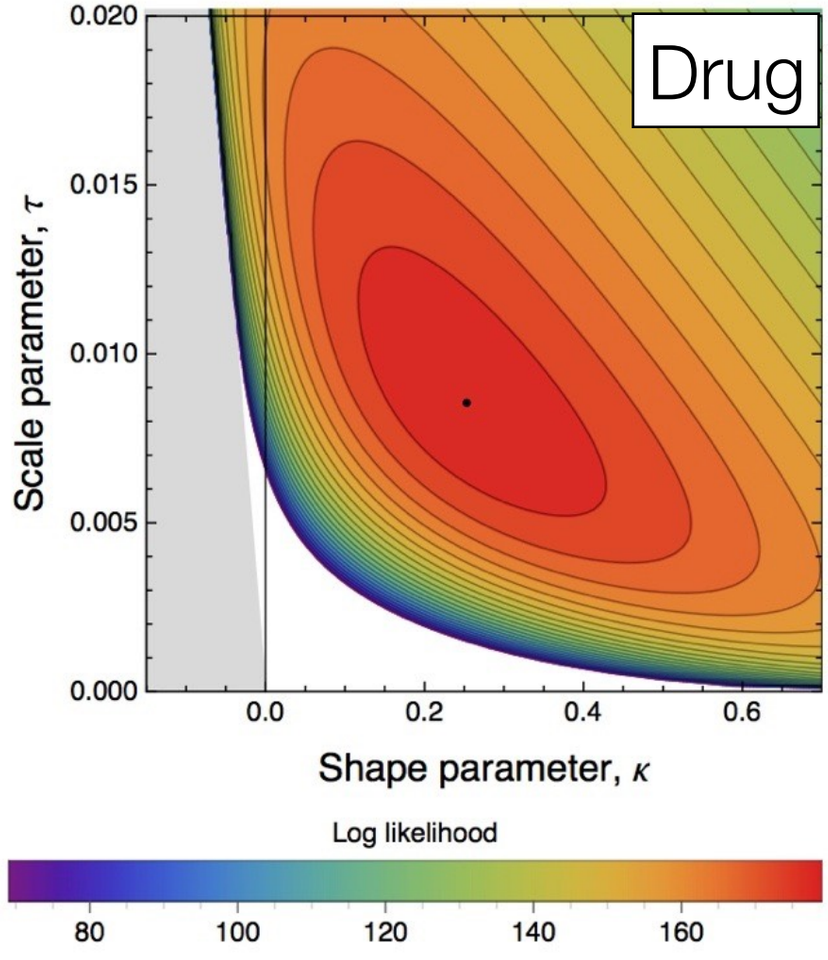
Approach: Serial passaging of influenza H1N1, assayed by time-sampled sequencing and inference via Approximate Bayesian Computation (WFABC; Foll*, Shim* & Jensen, *MER*, 2014)

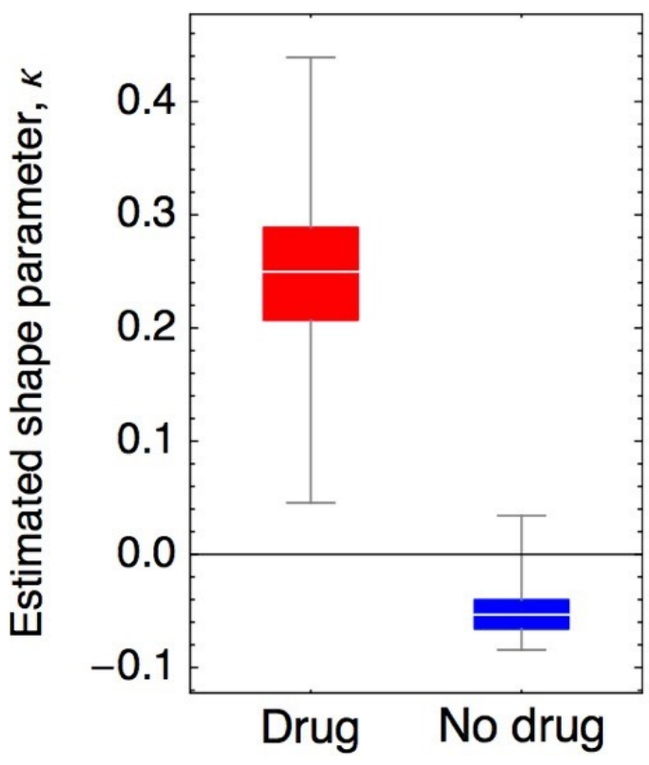
Environments:
MDCK (dog kidney cells) - **'standard' environment**,
MDCK + oseltamivir (Tamiflu) - **severe environmental challenge**

Tail shape parameter in challenging environments

Influenza H1N1 mutation accumulation experiment



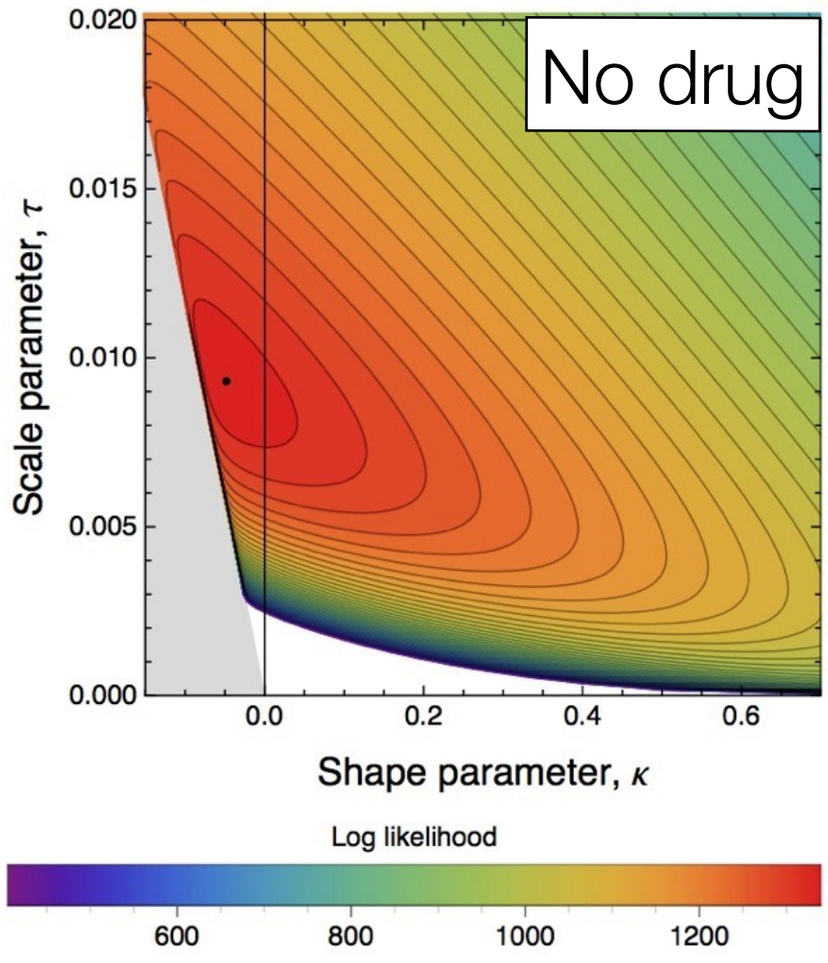


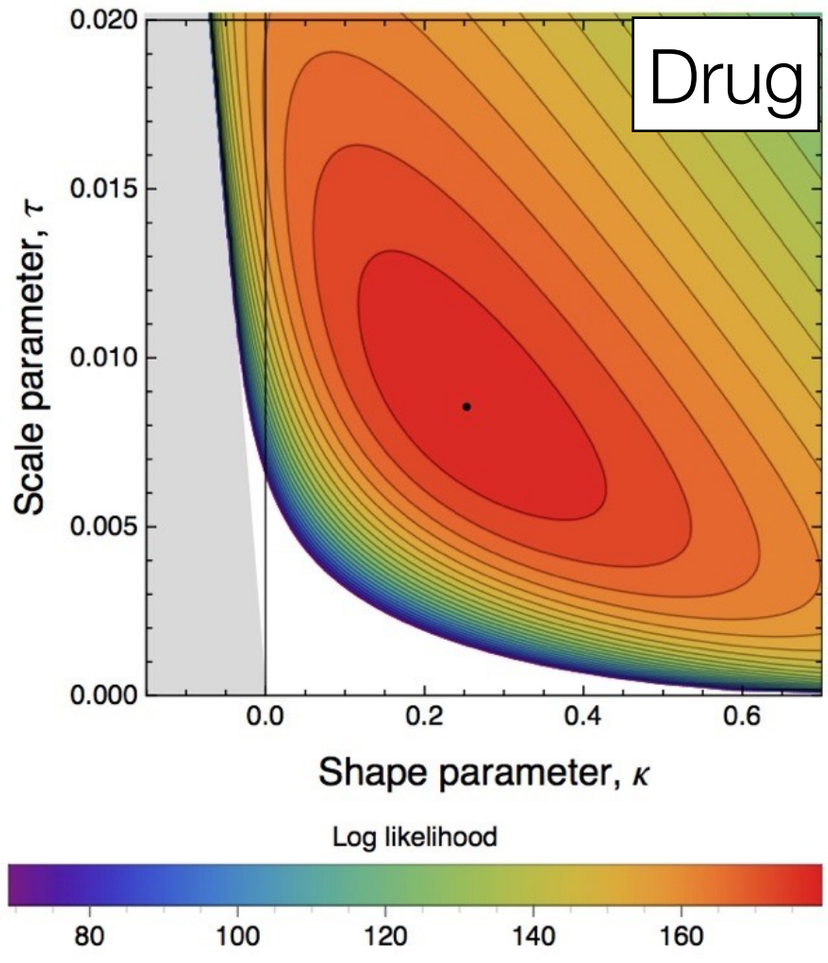


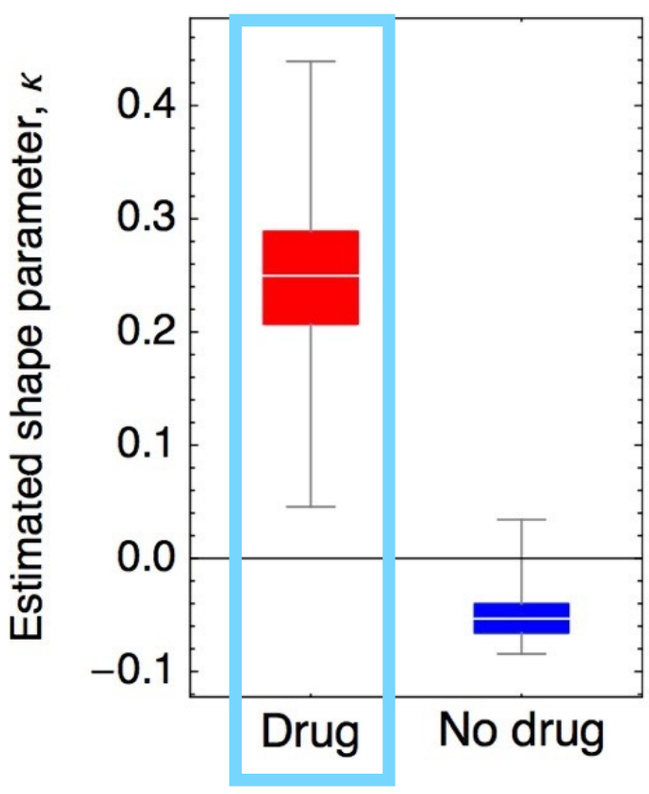
Foll, Poh, et al., *PLoS Genetics*, 2014

Tail shape parameter in challenging environments

Influenza H1N1 mutation accumulation experiment



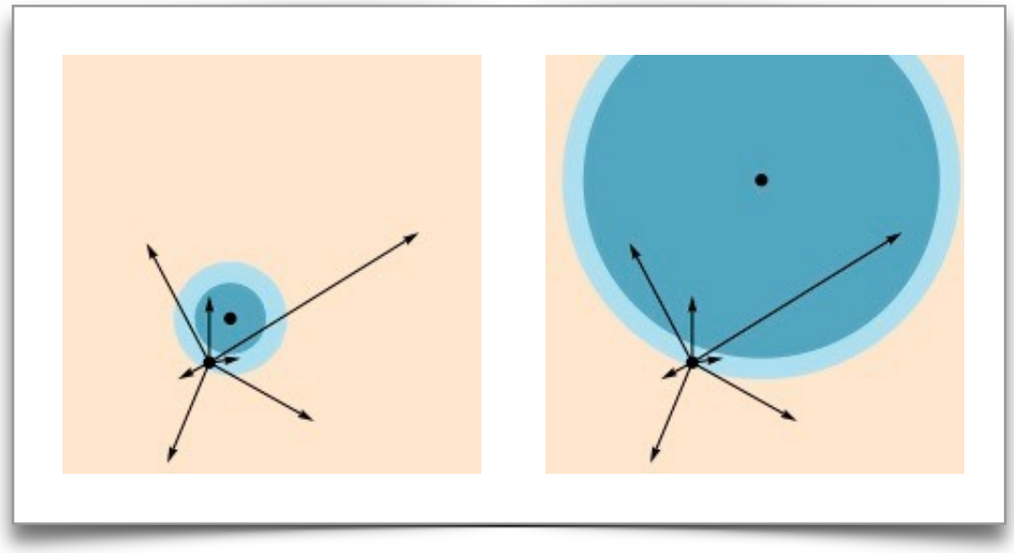




Foll, Poh, et al., *PLoS Genetics*, 2014

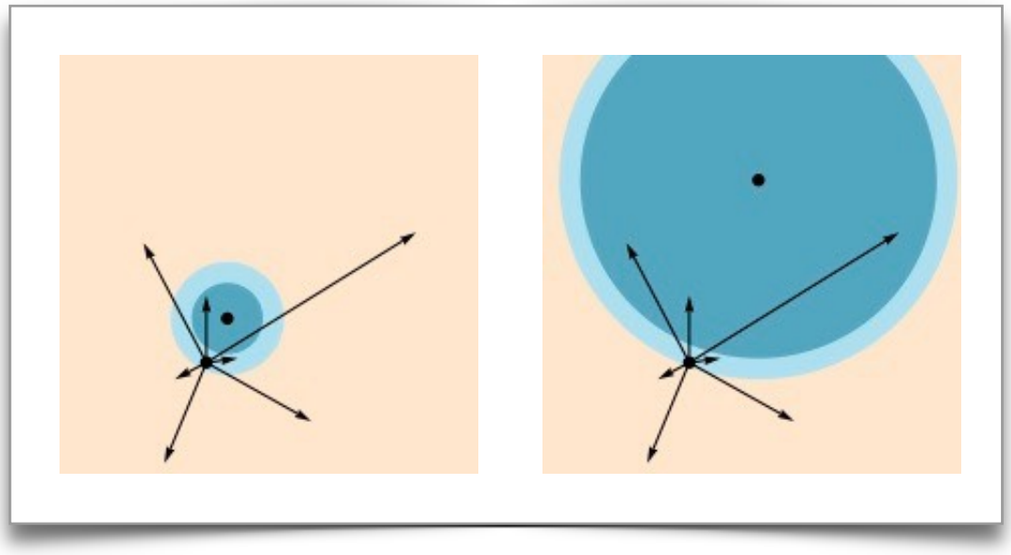
SUMMARY – TEAM FISHER

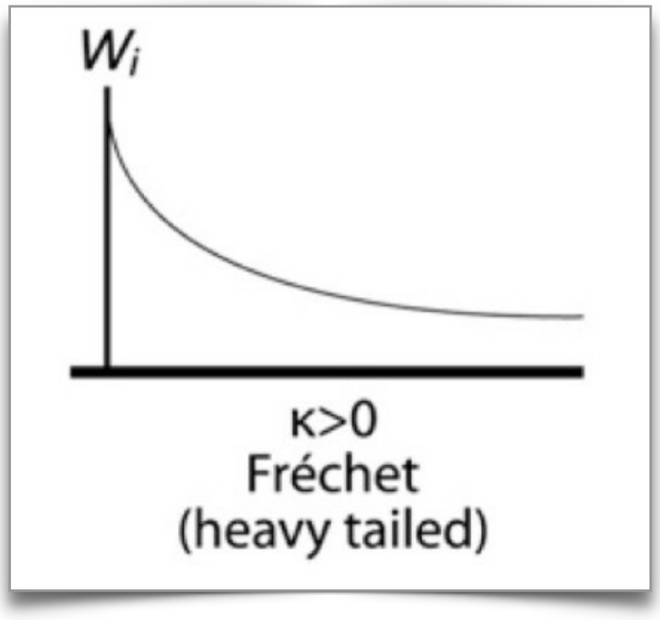
- In response to a novel environmental challenge the number and size of beneficial mutations increases, and costs of adaptation are observed - in agreement with predictions from Fisher's geometric model when the optimum is displaced.



SUMMARY – TEAM FISHER

- In response to a novel environmental challenge the number and size of beneficial mutations increases, and costs of adaptation are observed - in agreement with predictions from Fisher's geometric model when the optimum is displaced.
- Following severe environmental challenges, the step size of adaptive mutations might be highly unpredictable.

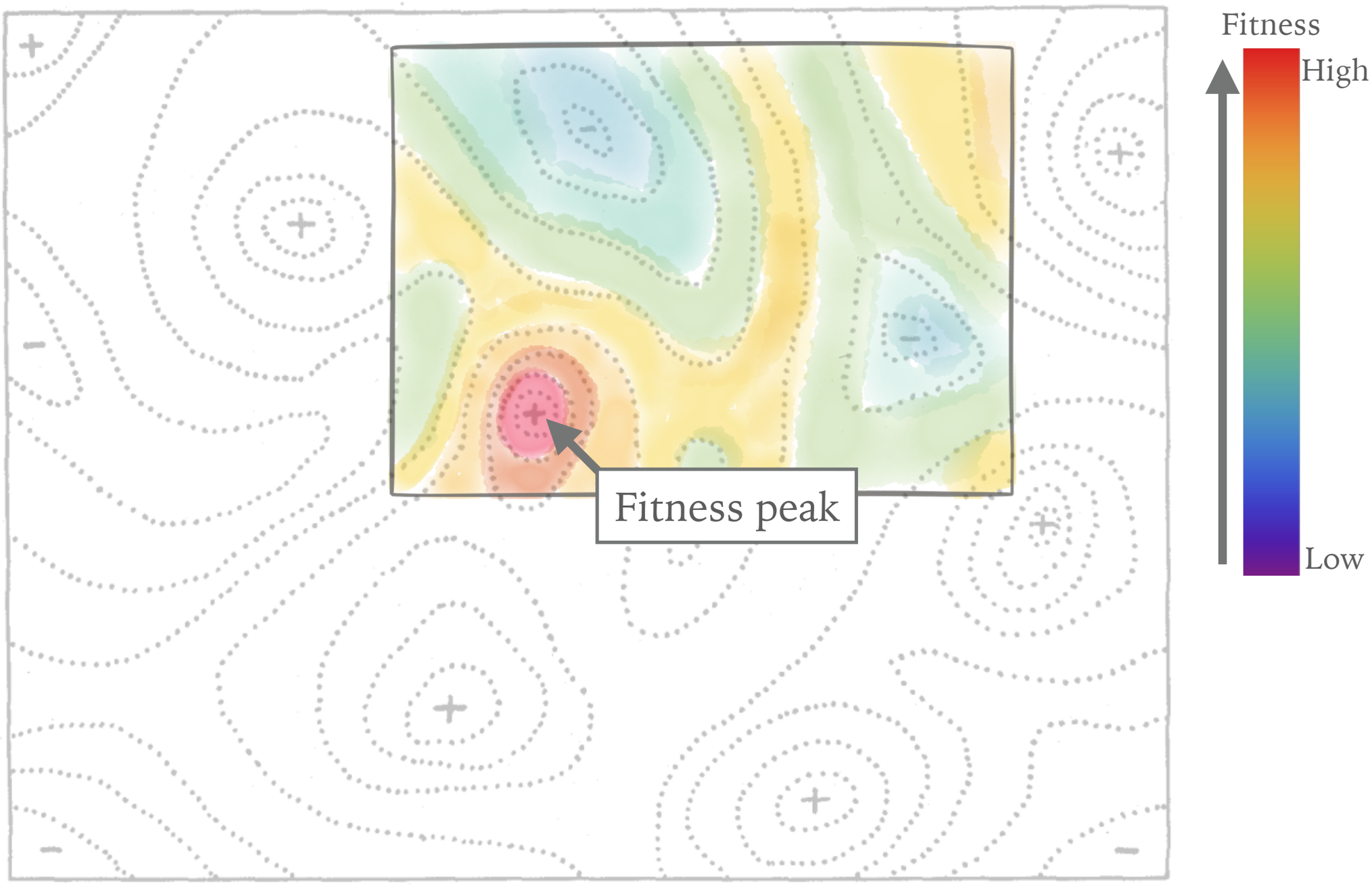




But what about epistasis?

(Spoiler: this is the part for the “Wrightians”)

WRIGHT'S SHIFTING BALANCE



Wright, 1932

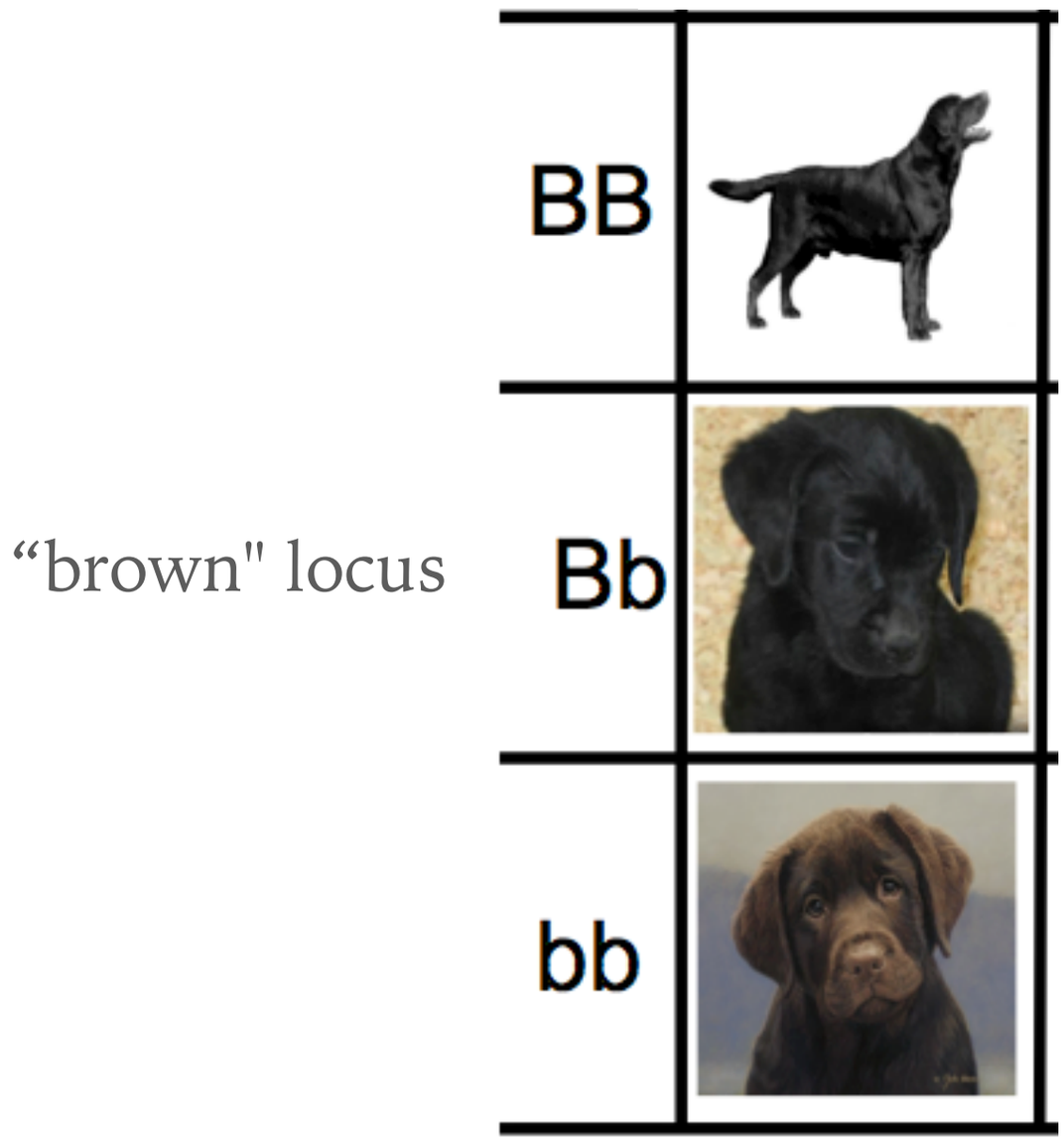
WHAT IS EPISTASIS?

Definition from Nature Reviews Genetics Glossary:

Epistatic interaction: any non-additive interaction between two or more mutations at different loci, such that their combined effect on a phenotype deviates from the sum of their individual effects.

WHAT IS EPISTASIS?

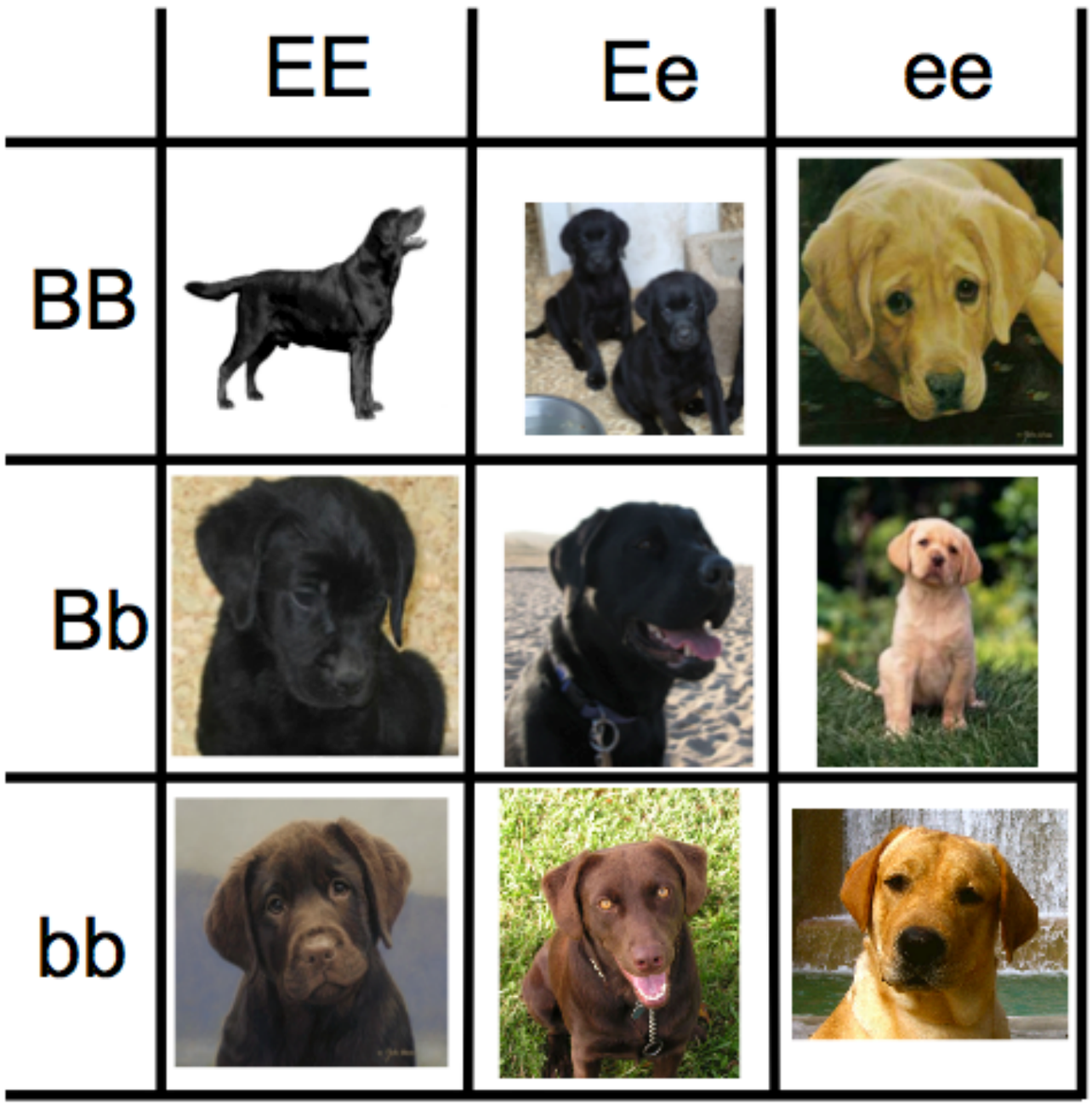
Classical genetics:



(Source: Google image search)

WHAT IS EPISTASIS?

Classical genetics:



(Source: Google image search)

WHAT IS EPISTASIS?

Classical genetics:

“extension" locus

	EE	Ee	ee
BB			
Bb			
bb			

“brown" locus

This guy should have a brown nose!

(Source: Google image search)

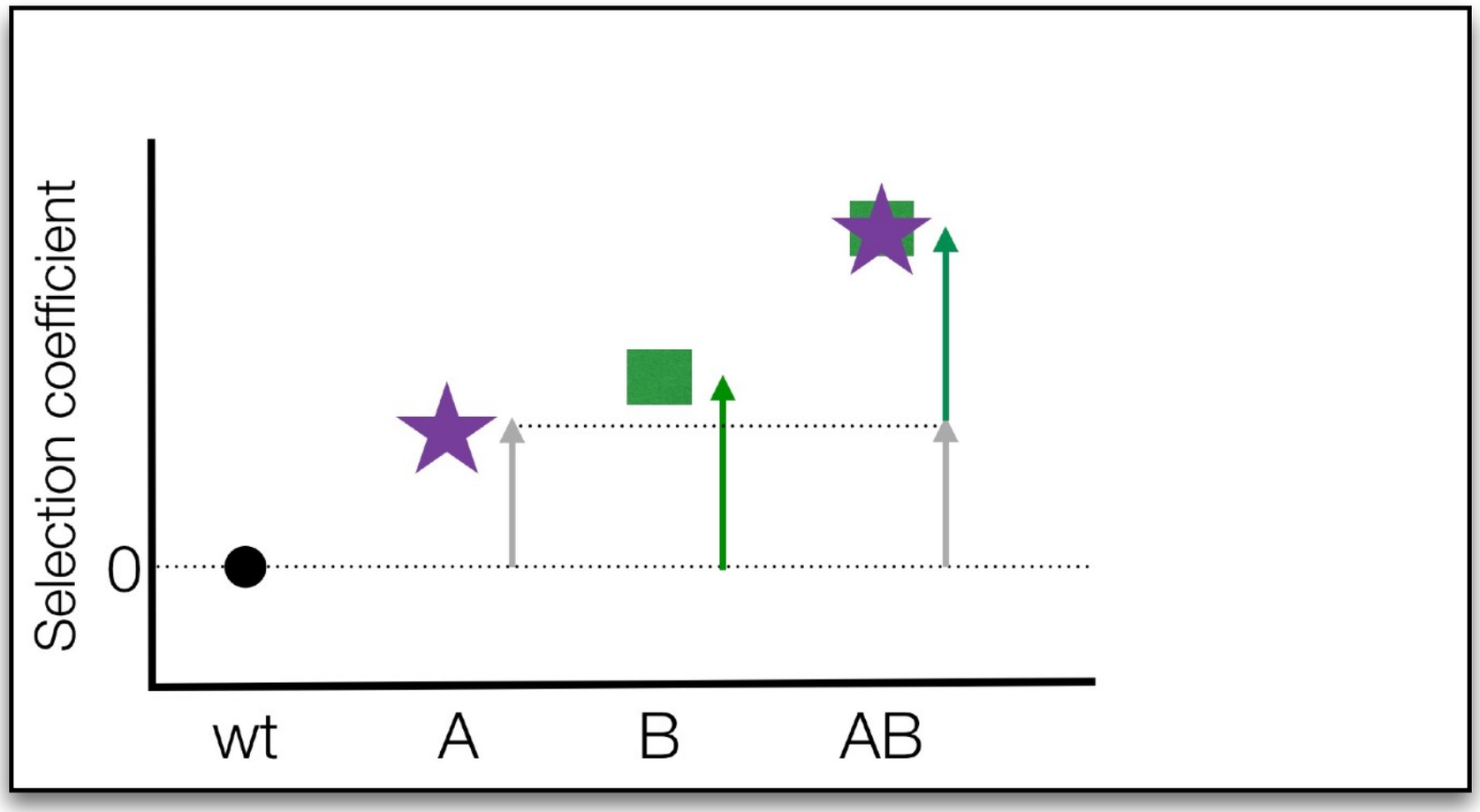
WHAT IS EPISTASIS?

Quantitative Genetics:

Interaction of genetic variants such that the net phenotypic effect of carrying more than one variant is different than would be expected by simply combining the effects of each individual variant.

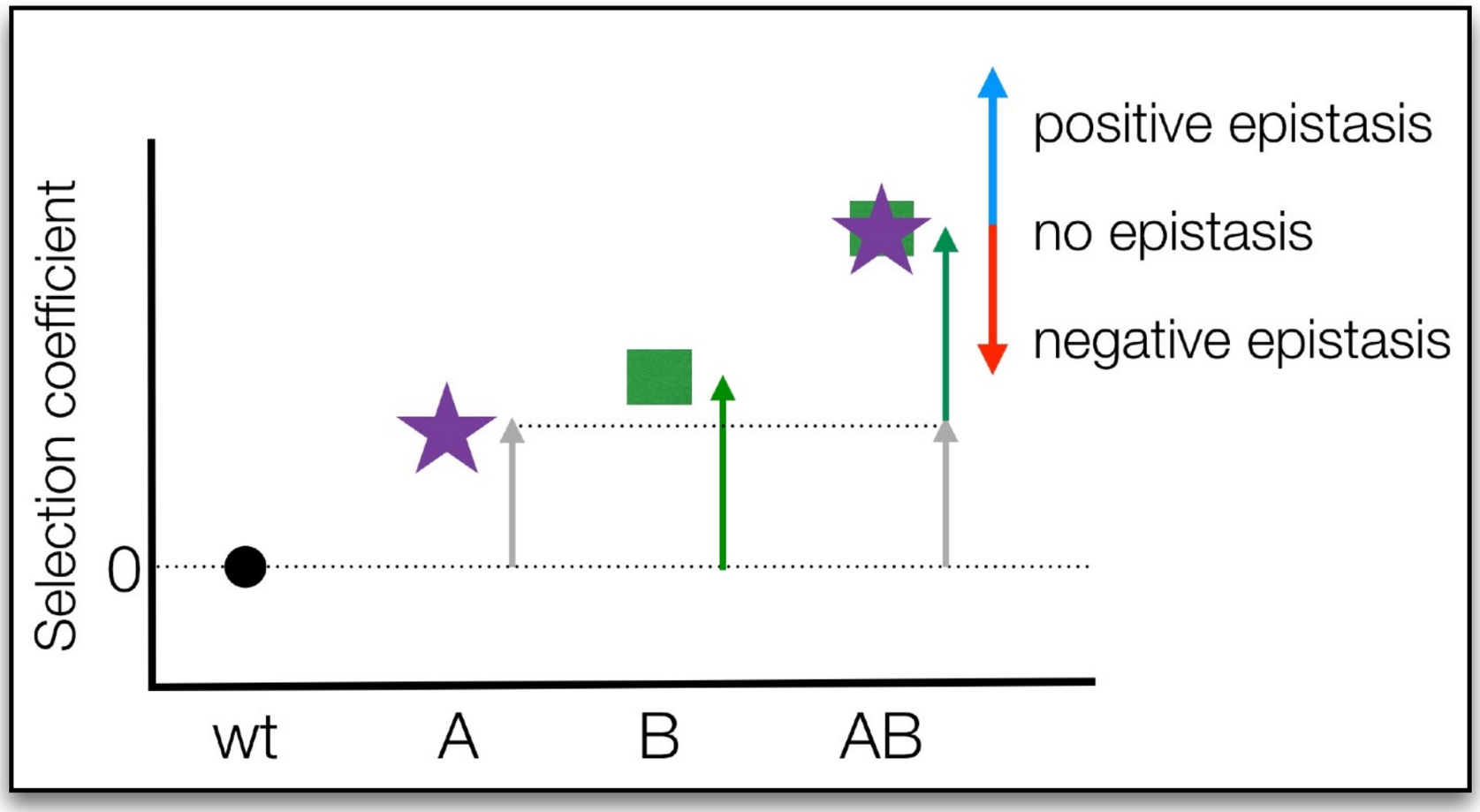
WHAT IS EPISTASIS?

Phenotype: fitness



WHAT IS EPISTASIS?

Phenotype: fitness



WHY SHOULD WE CARE?

Romero & Harnold, 2009

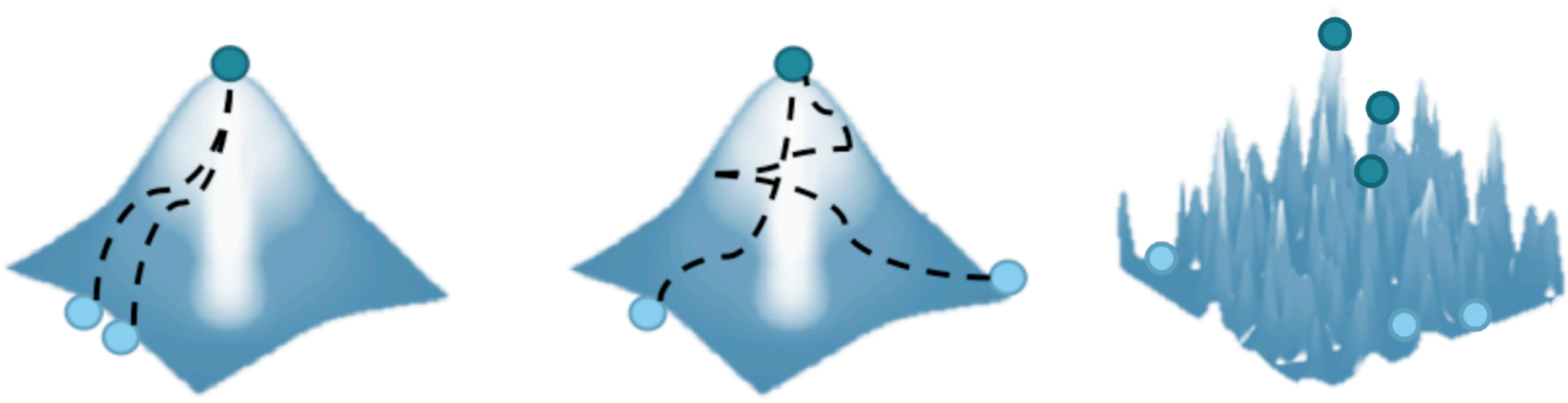
WHY SHOULD WE CARE?

- epistasis creates non-random associations between loci (linkage disequilibrium; LD)

Romero & Harnold, 2009

WHY SHOULD WE CARE?

- epistasis creates non-random associations between loci (linkage disequilibrium; LD)
- Ruggedness of fitness landscape is a determinant of predictability/repeatability of evolution



Predictable and Repeatable

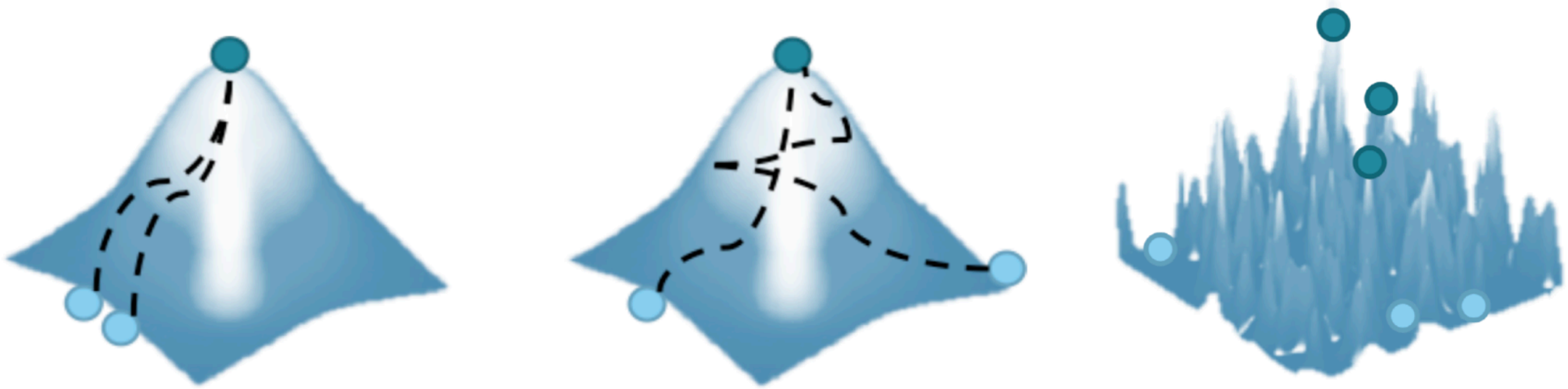
Predictable and Non-repeatable

Non-Predictable

Romero & Harnold, 2009

WHY SHOULD WE CARE?

- epistasis creates non-random associations between loci (linkage disequilibrium; LD)
- Ruggedness of fitness landscape is a determinant of predictability/repeatability of evolution
- accumulation of epistatic alleles is basis of the most widely accepted model for allopatric speciation



Predictable and Repeatable

Predictable and Non-repeatable

Non-Predictable

Romero & Harnold, 2009

WHERE DO WE EXPECT TO SEE EPISTASIS?

- Within proteins
- Between genomic regions involved in biological pathways
- Between species

WHERE DO WE EXPECT TO SEE EPISTASIS?

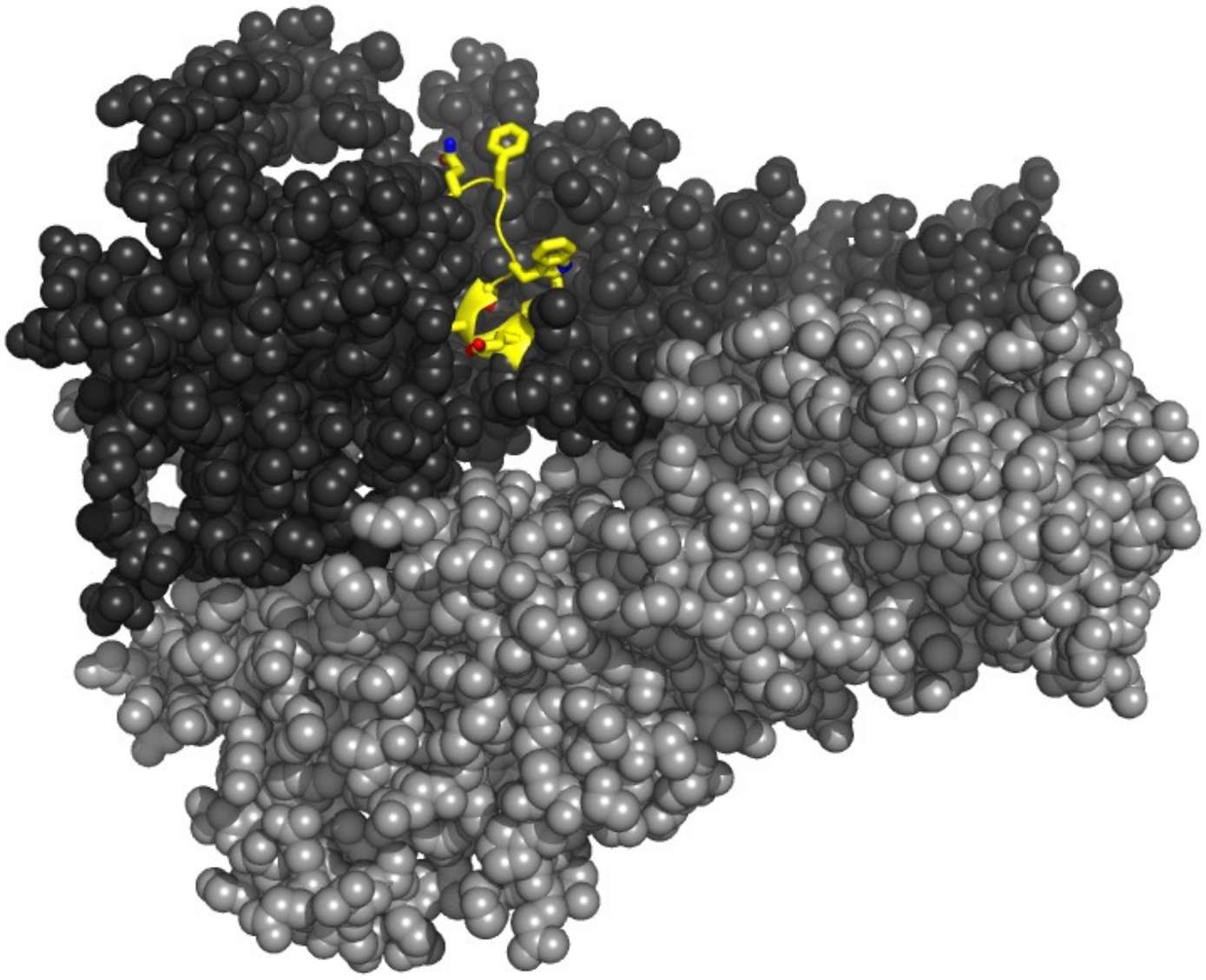
- Within proteins
- Between genomic regions involved in biological pathways
- Between species

Where do you expect to see the most epistasis, and why?

AN INTRAGENIC EPISTATIC LANDSCAPE

DATA SET

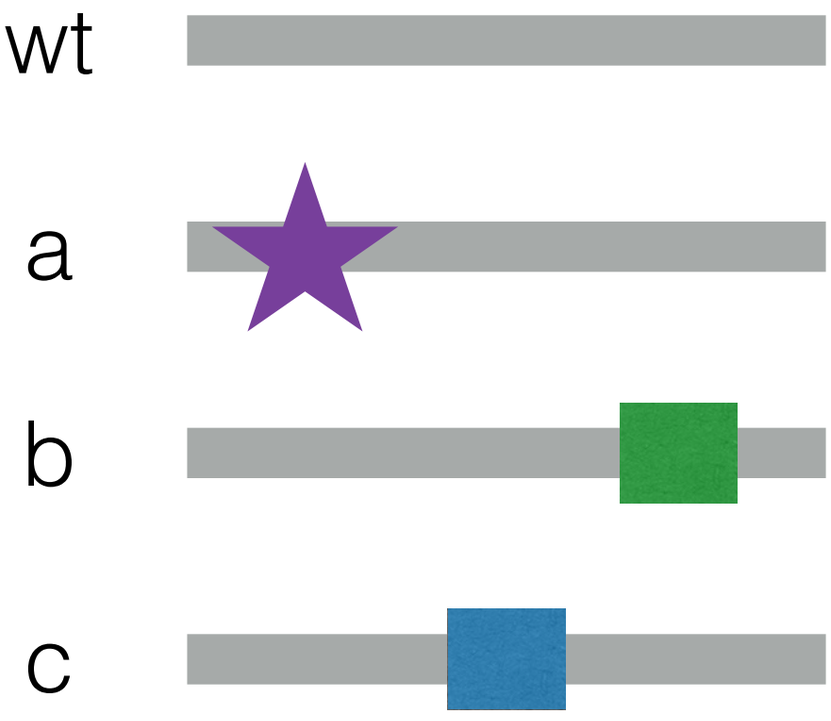
- 9 aa region from Hsp90 (aa positions 582-590) in Saccharomyces cerevisiae



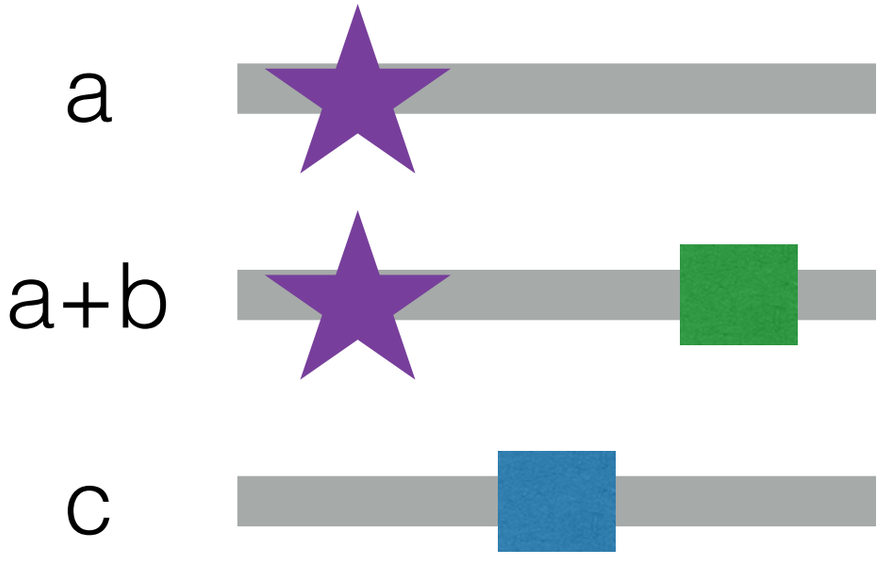
Data obtained by Ryan Hietpas @ UMassMed

Data set

3x 180 1-step mutations



7 anchors, each with
≈140 2nd steps,
≈20 1-step controls



Anchor fitness between 97.5% and 100% of wt

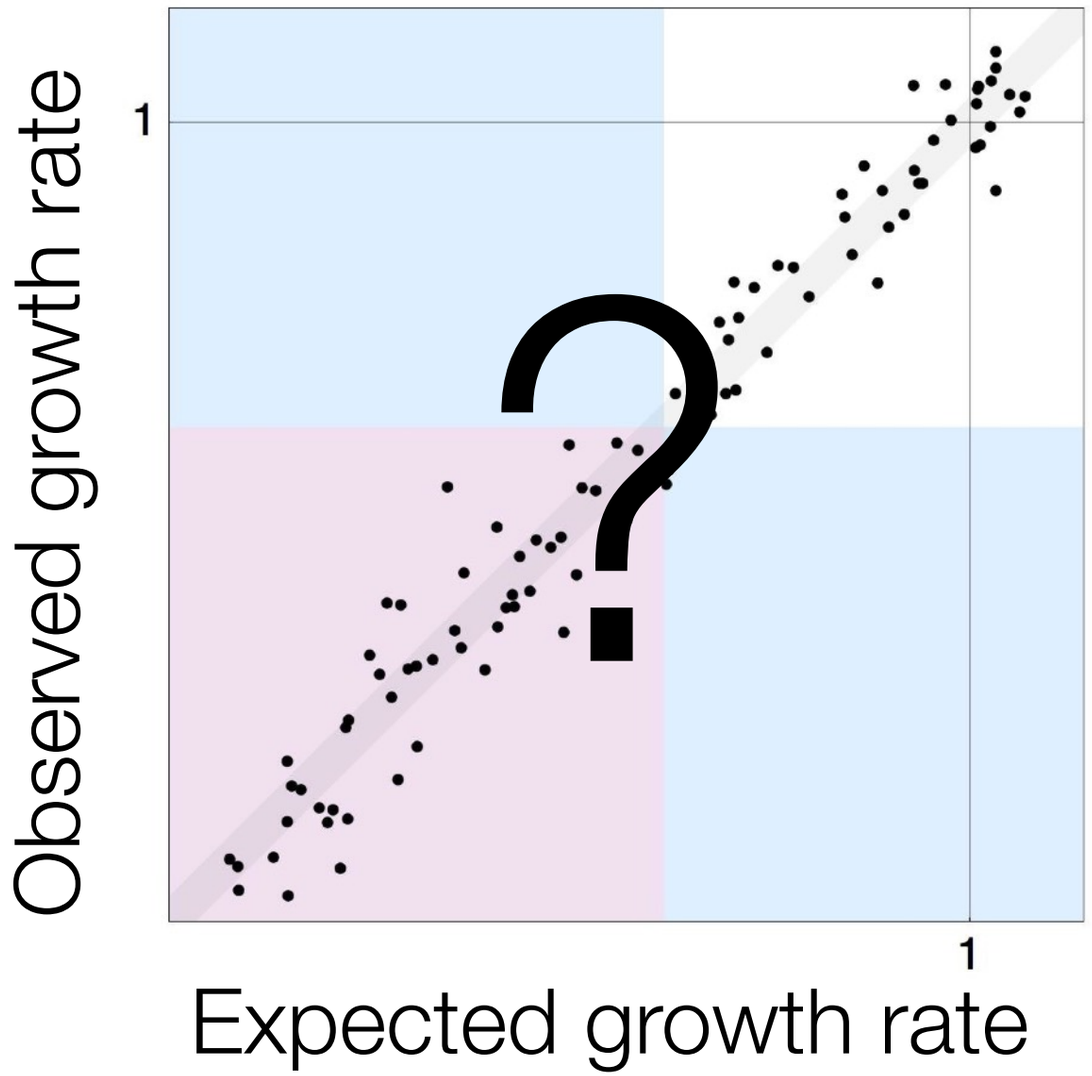
Data obtained by Ryan Hietpas @ UMassMed

Questions

- **What is the general pattern of epistasis?**

- How does “a step away” change the distribution of fitness effects?

- What is the shape of the local fitness landscape?



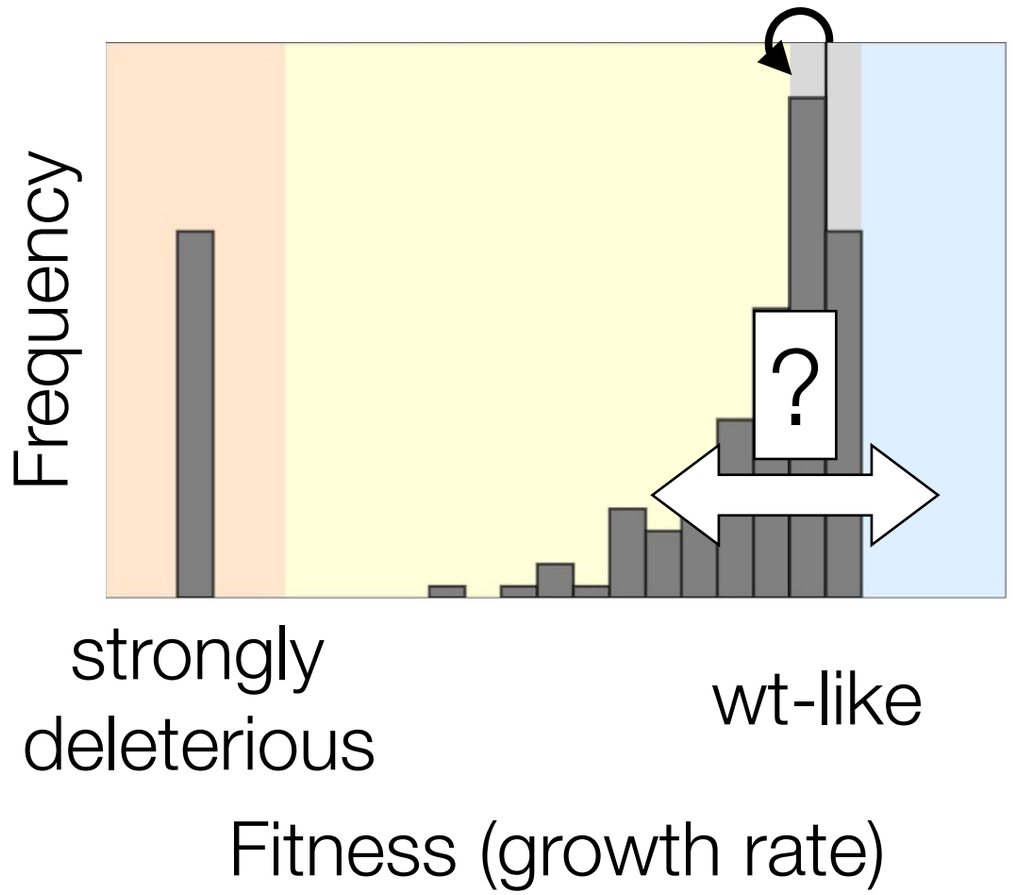
Questions

- What is the general pattern of epistasis?

- **How does “a step away” change the distribution of fitness effects?**

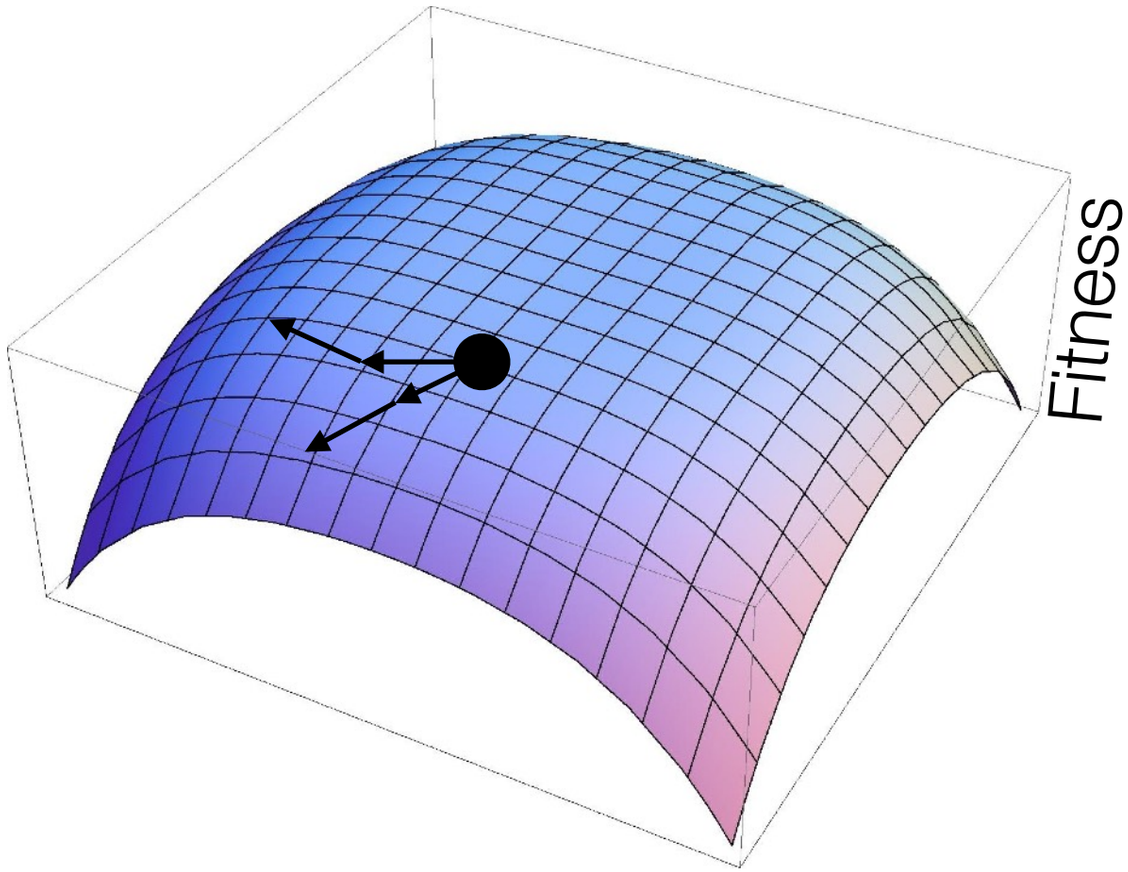
- What is the shape of the local fitness landscape?

The distribution of fitness effects (DFE)

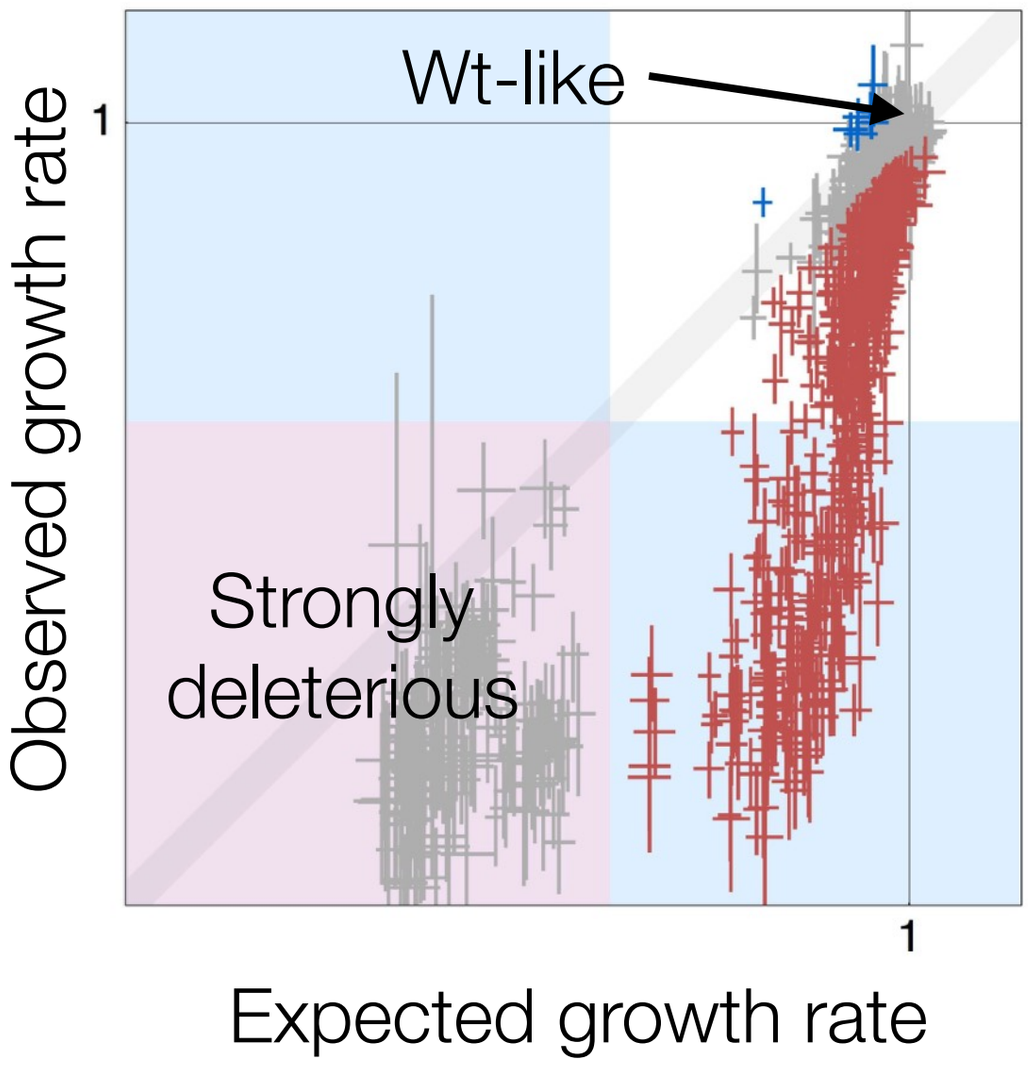


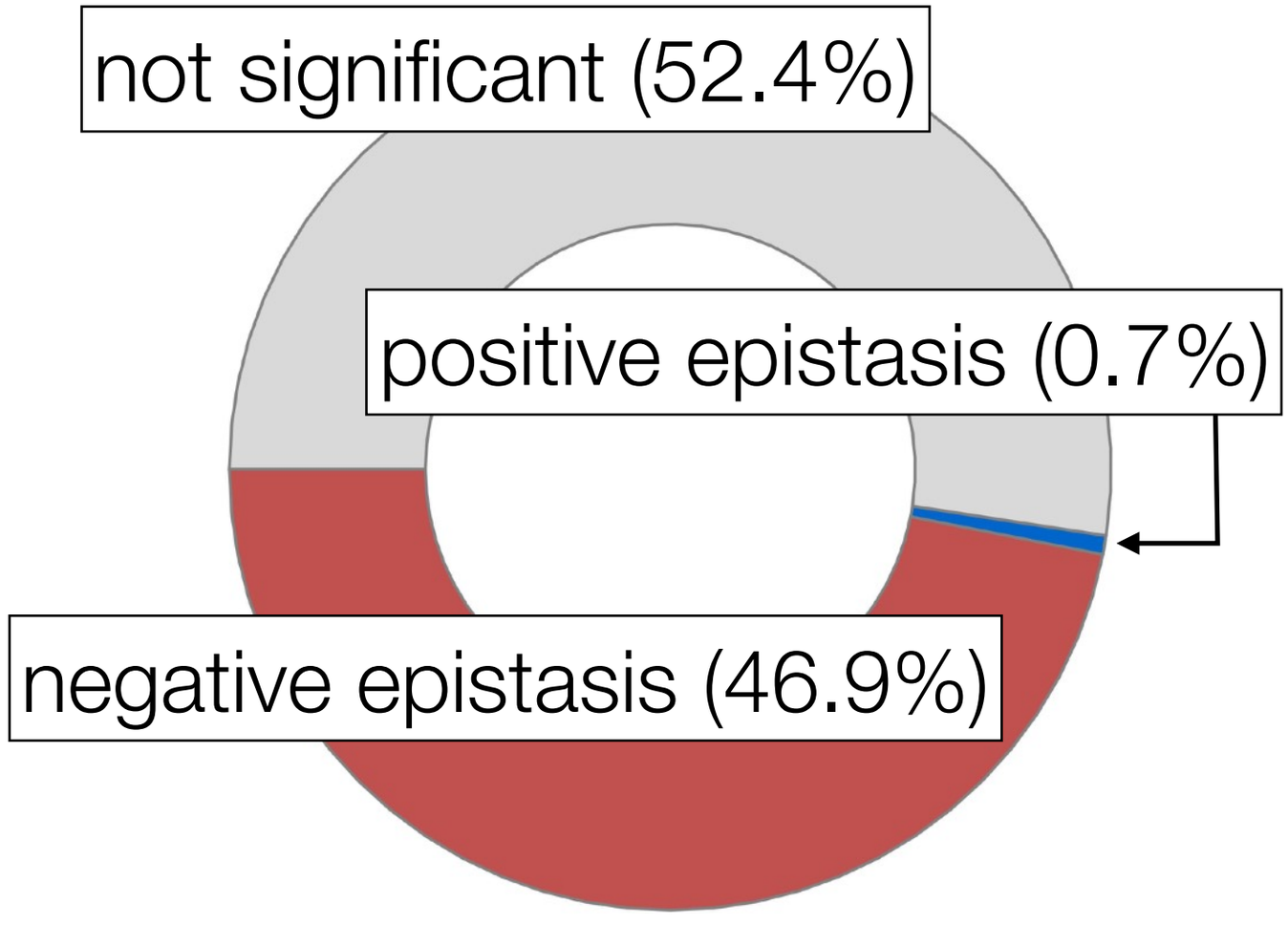
Questions

- What is the general pattern of epistasis?
- How does “a step away” change the distribution of fitness effects?
- **What is the shape of the local fitness landscape?**

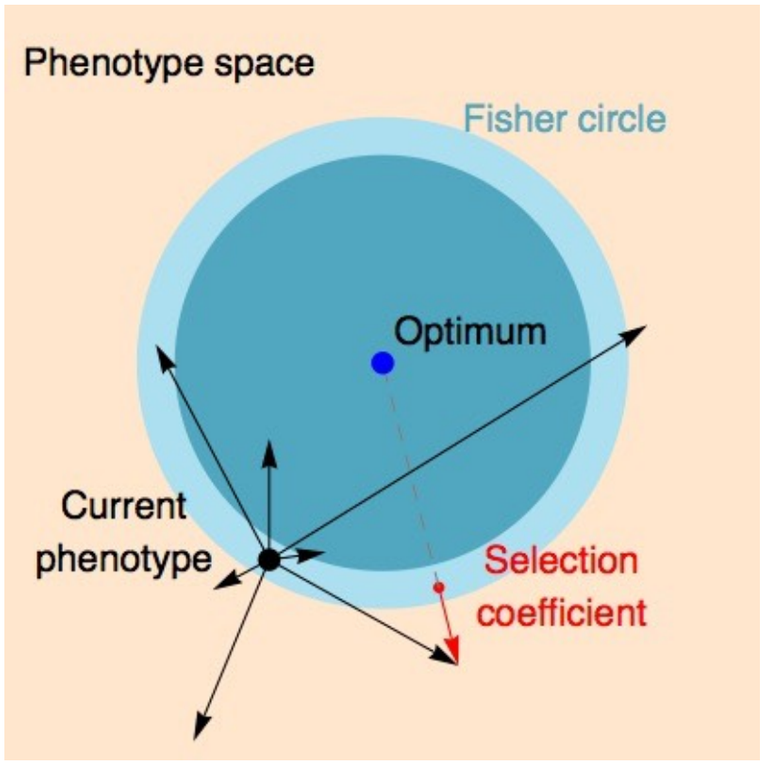


The pattern of epistasis for 1015 2-step mutations





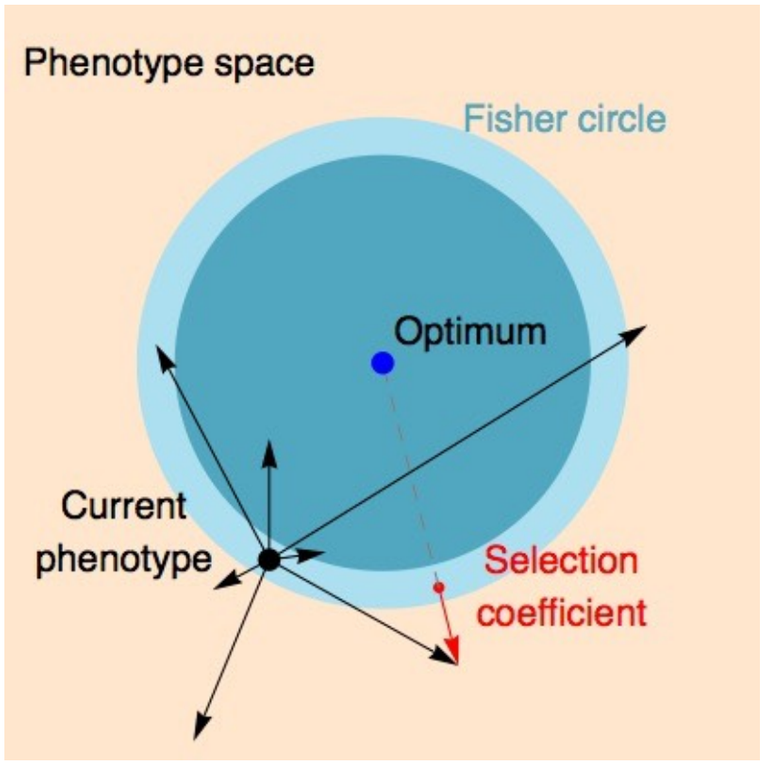
FISHER'S GEOMETRIC MODEL



Fisher, 1930

Hypothesis:

FISHER'S GEOMETRIC MODEL

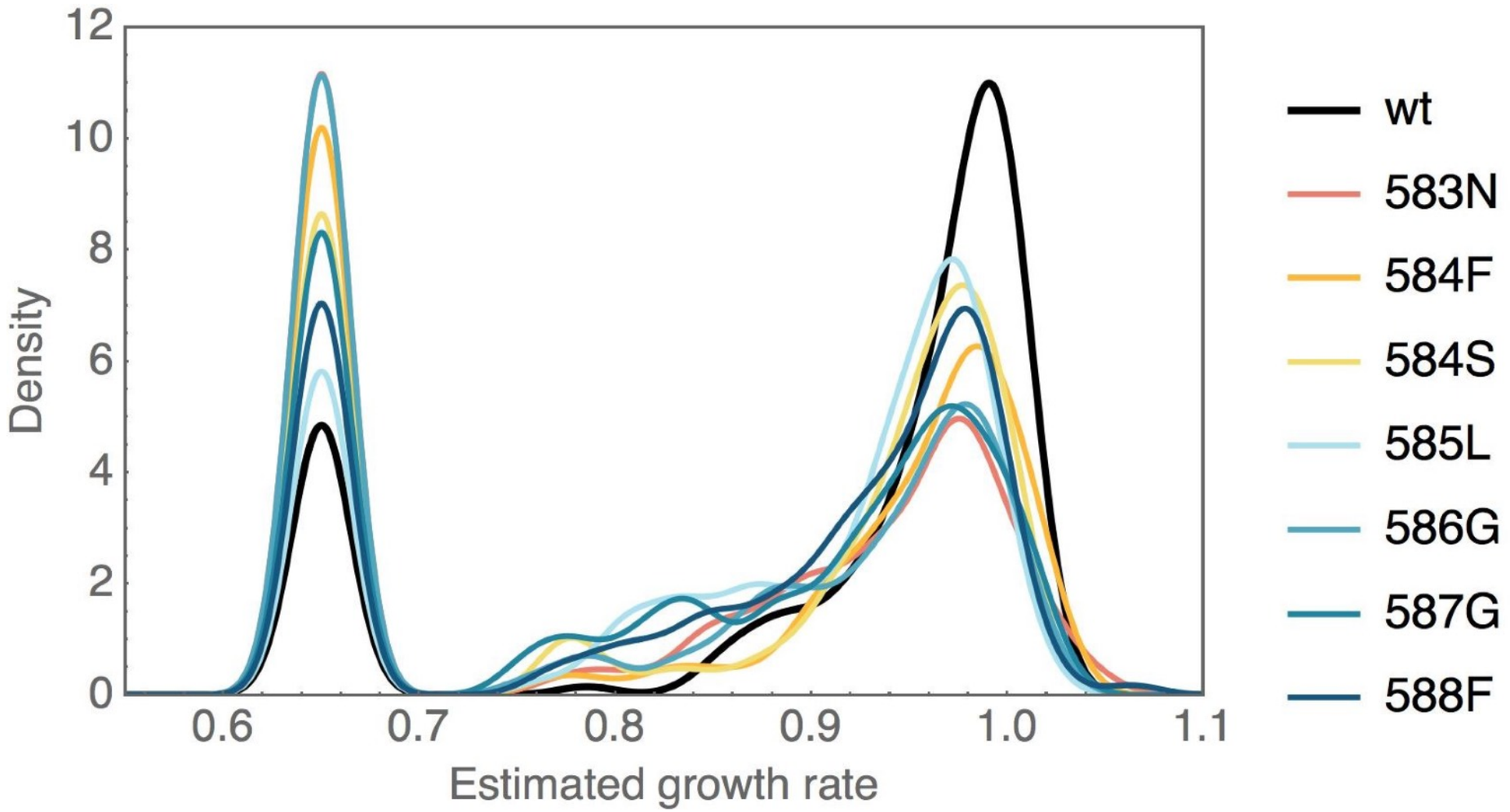


Fisher, 1930

Hypothesis:

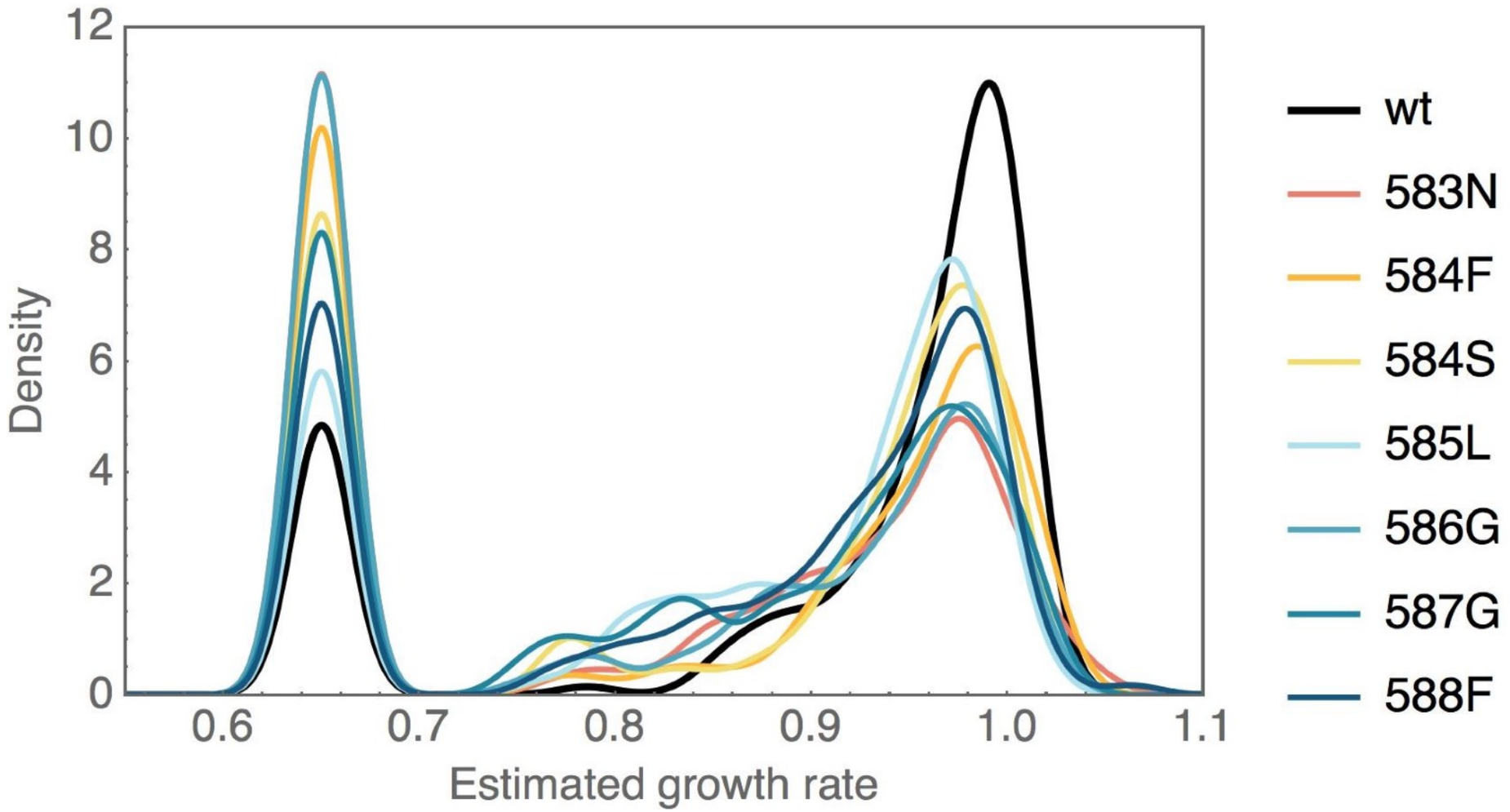
- Relocation of the phenotype can increase the distance to the optimum and hence the potential for beneficials.

Shape of the DFE one step away



- No higher potential for beneficial mutations one step away from the wild type
- 2nd step much more likely to be deleterious

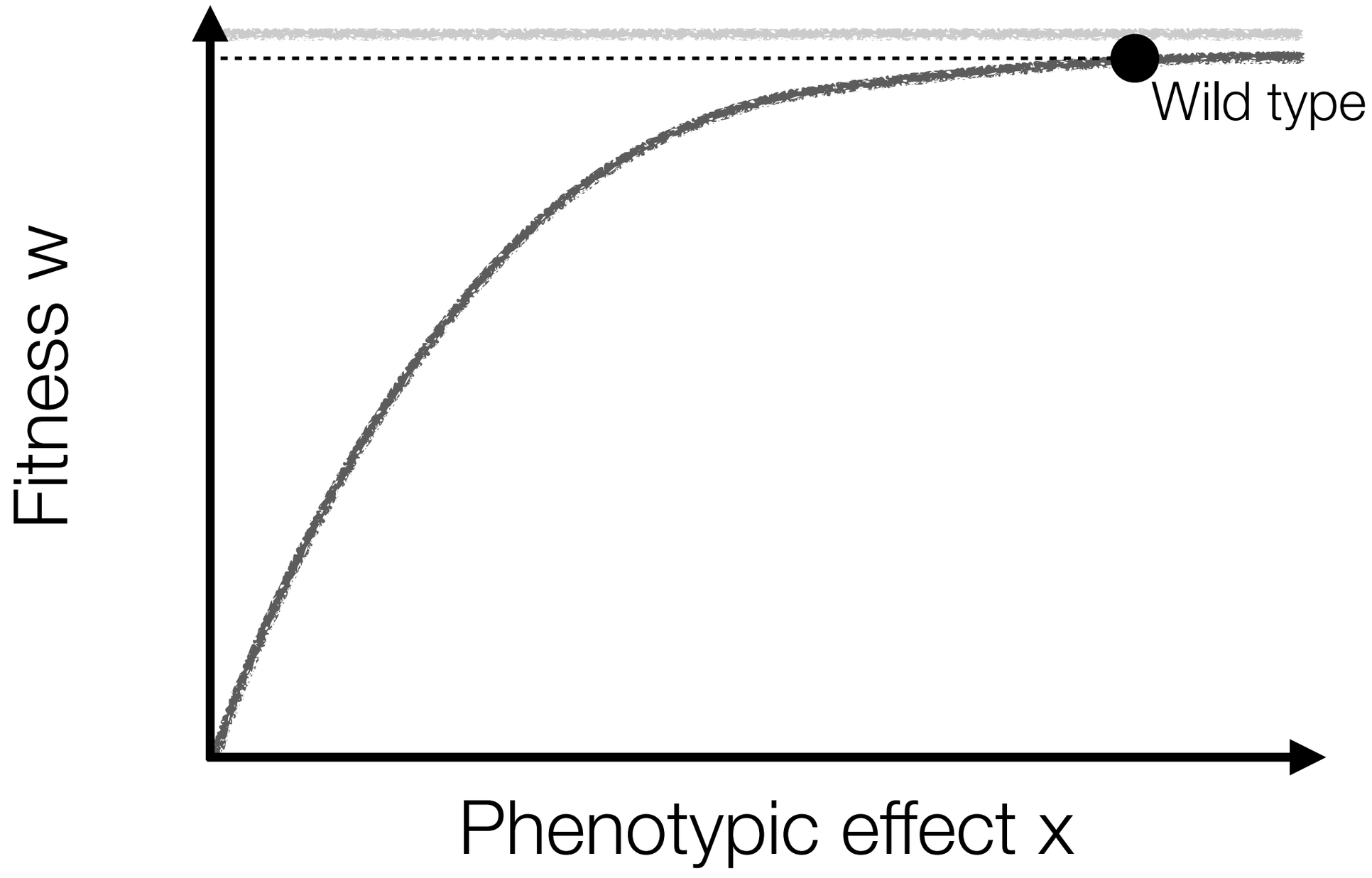
Shape of the DFE one step away



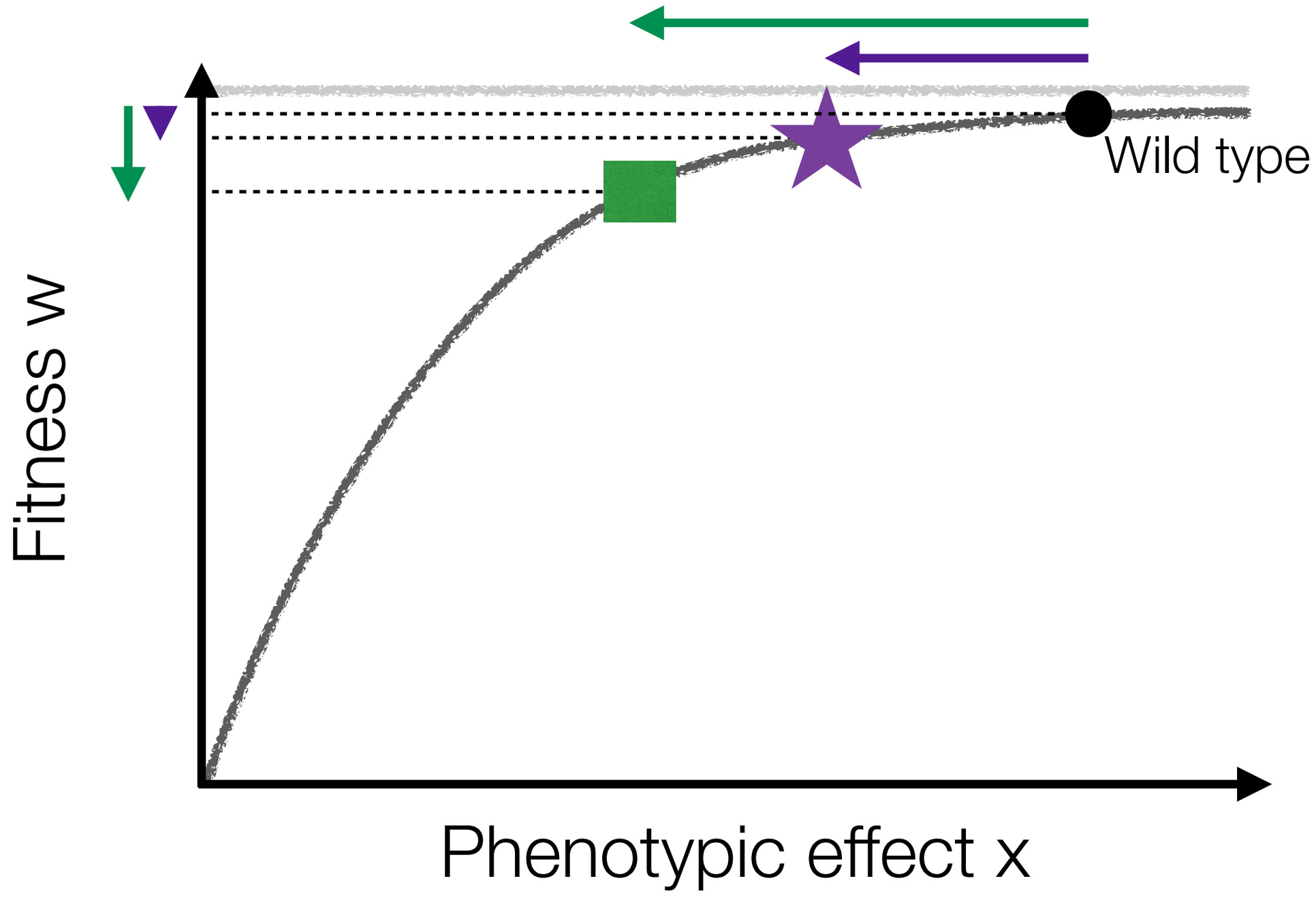
- No higher potential for beneficial mutations one step away from the wild type
- 2nd step much more likely to be deleterious

Robustness? - Concave fitness landscape?

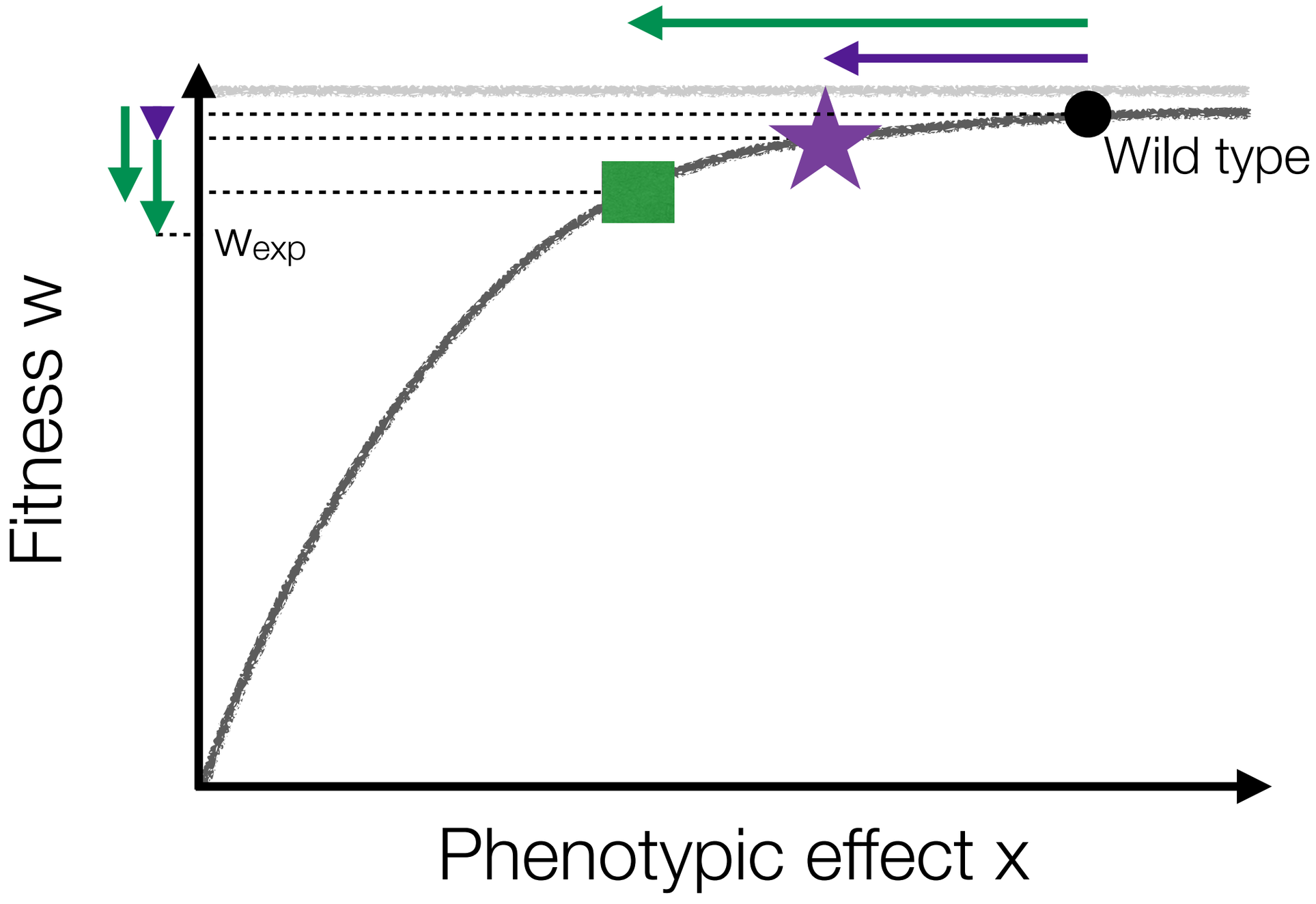
Characterizing the shape of the epistatic landscape



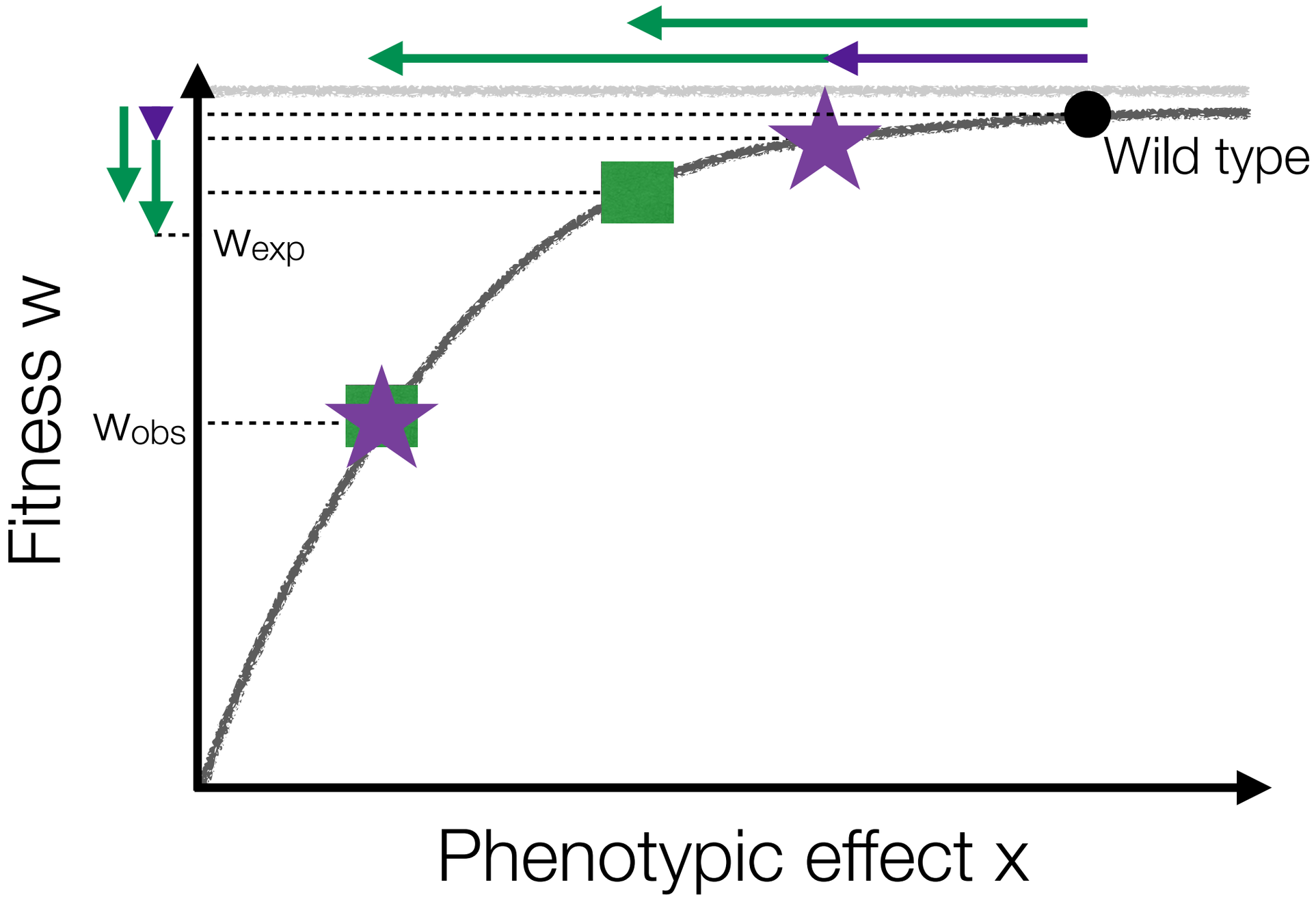
Characterizing the shape of the epistatic landscape



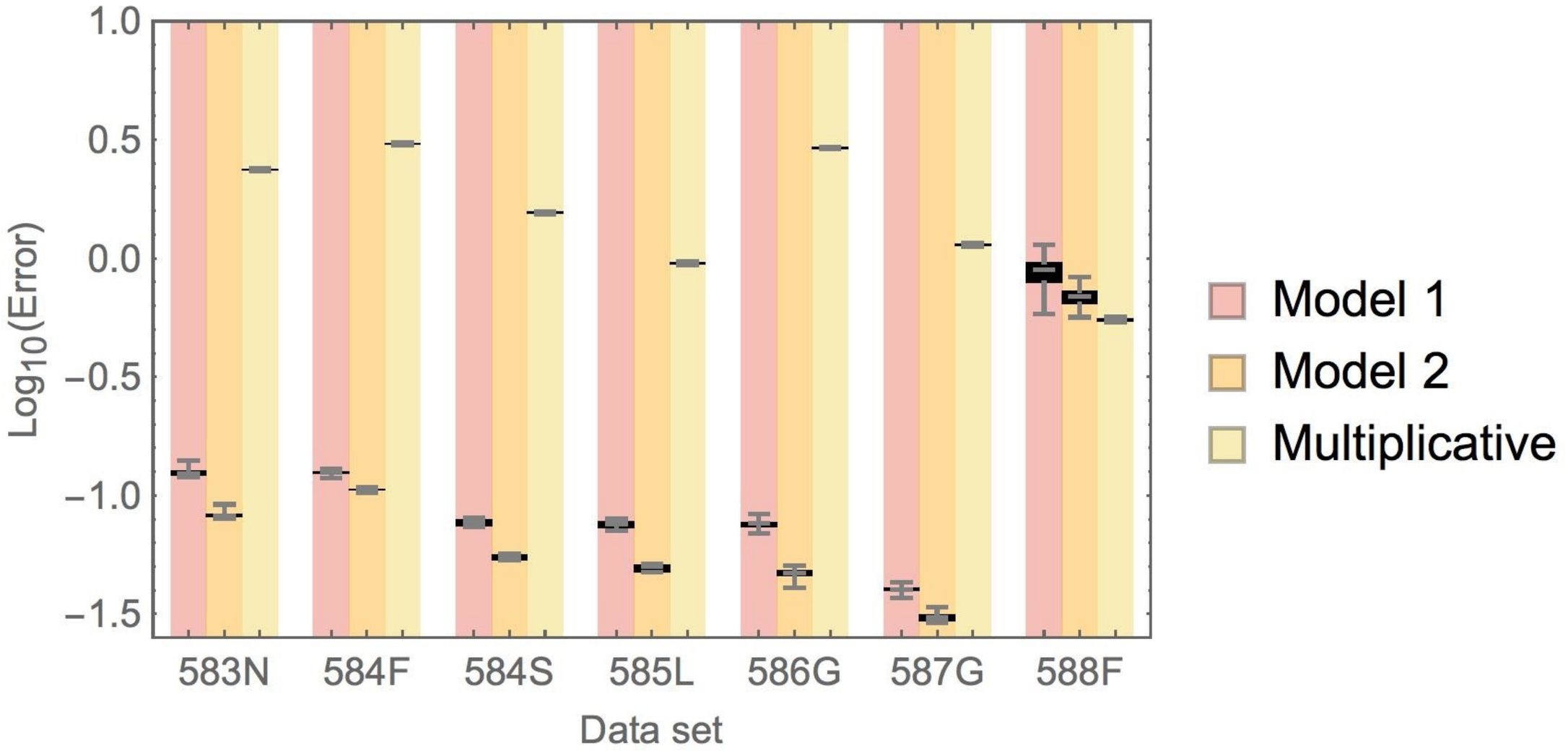
Characterizing the shape of the epistatic landscape



Characterizing the shape of the epistatic landscape

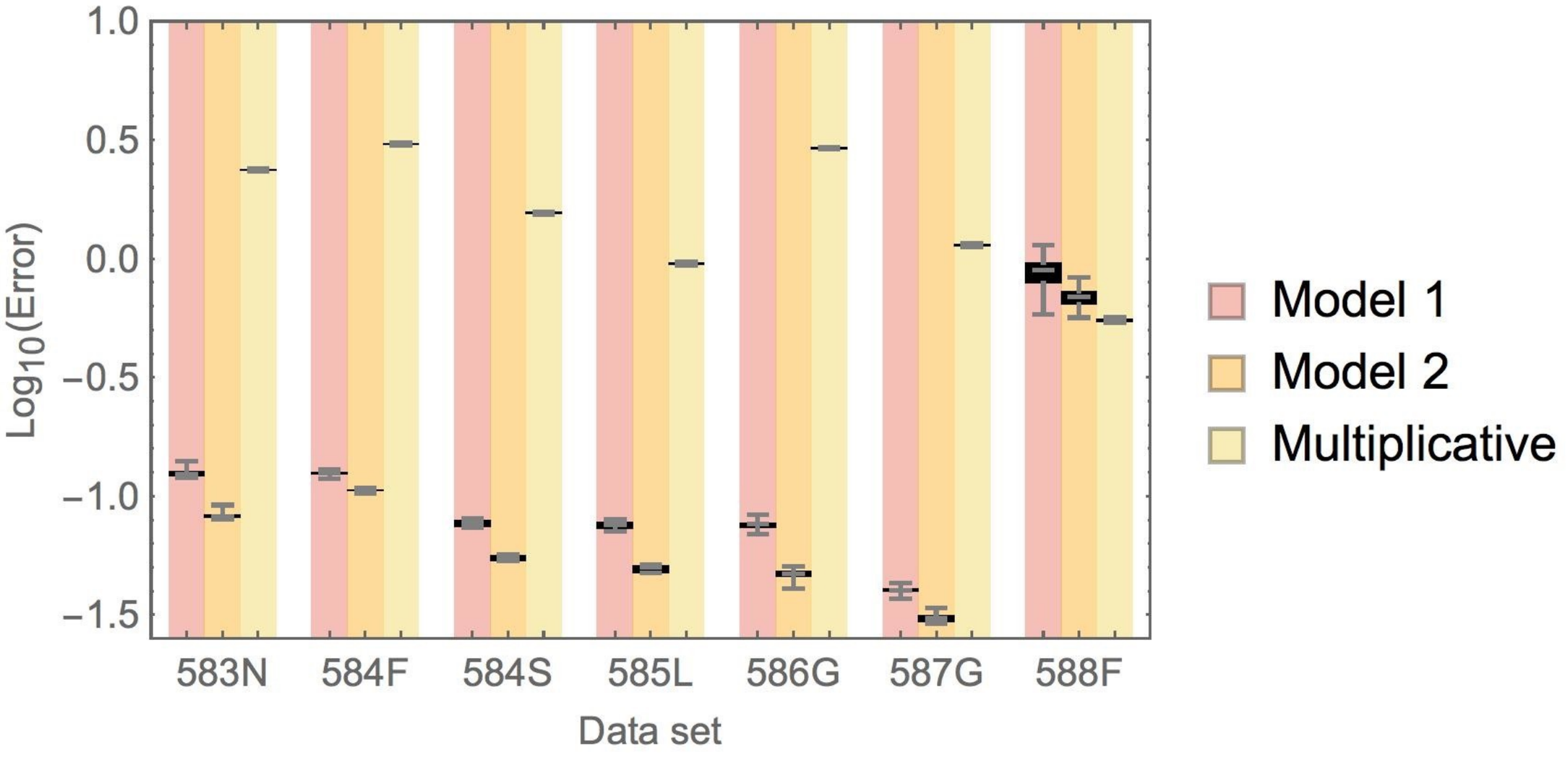


Characterizing the shape of the epistatic landscape



Bank, Hietpas, Jensen & Bolon, *MBE*, 2015

Characterizing the shape of the epistatic landscape

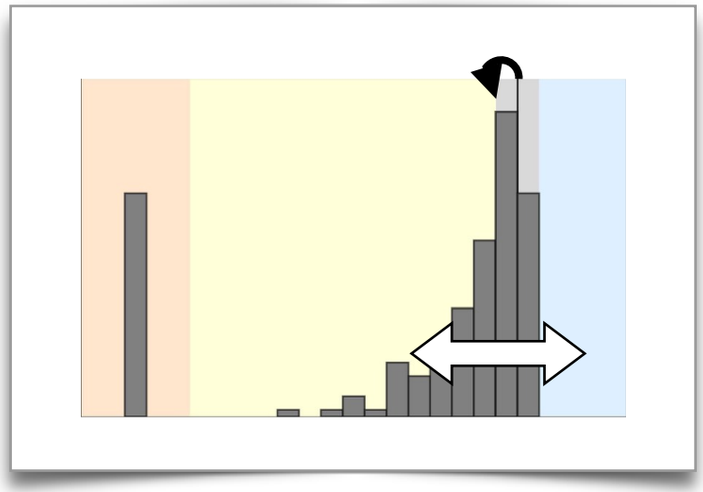


Models of concave fitness landscapes can explain the observed pattern of ubiquitous negative epistasis.

Bank, Hietpas, Jensen & Bolon, *MBE*, 2015

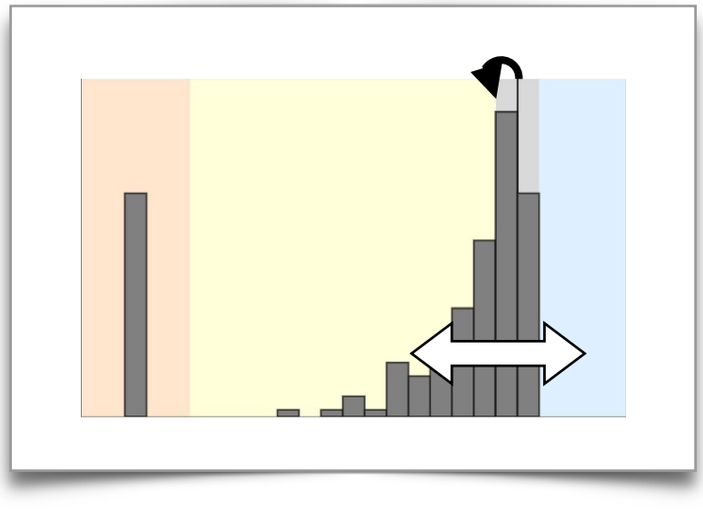
SUMMARY

- Intragenic epistasis is frequent. Changing the genetic background (i.e., dislocating the population from its current state) does not result in a higher potential for beneficial mutations.

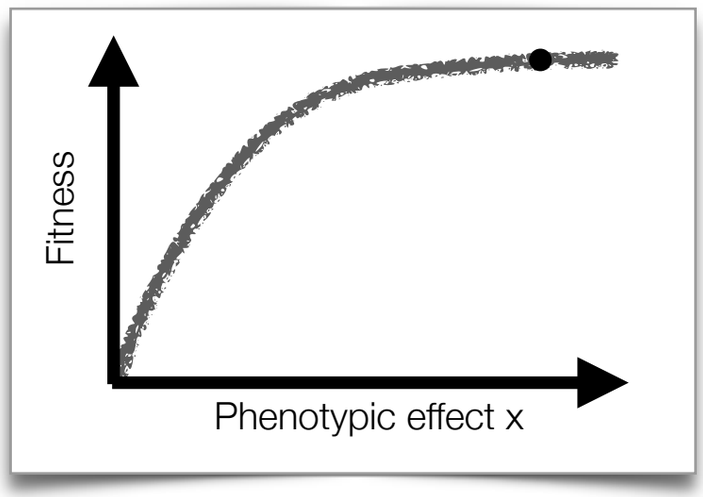


SUMMARY

- Intragenic epistasis is frequent. Changing the genetic background (i.e., dislocating the population from its current state) does not result in a higher potential for beneficial mutations.

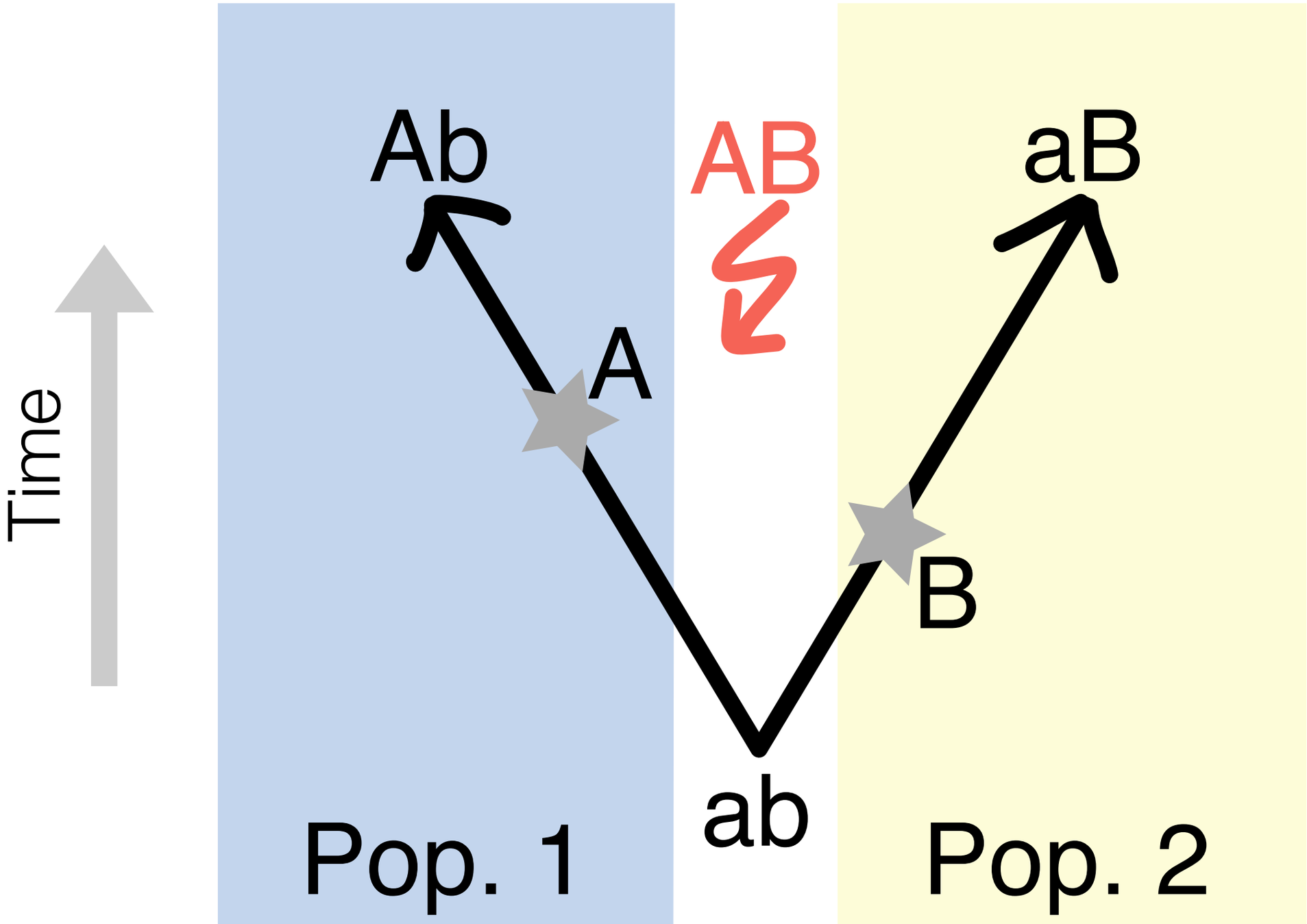


- Models of concave fitness landscape motivated by biophysical properties can explain the observed pattern of epistasis in the standard environment, supporting the idea of mutational robustness.

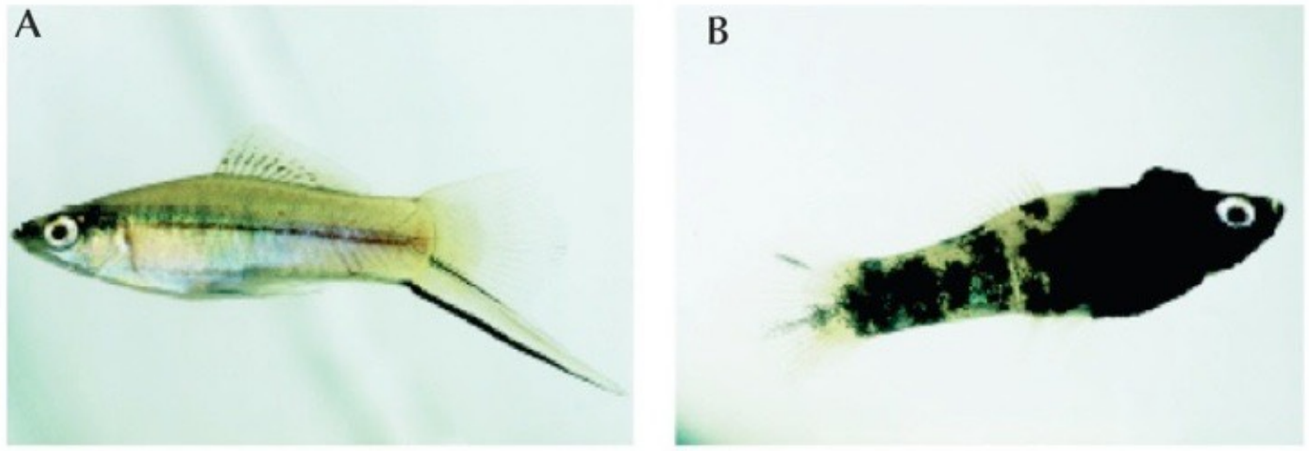


EPISTASIS BETWEEN SPECIES

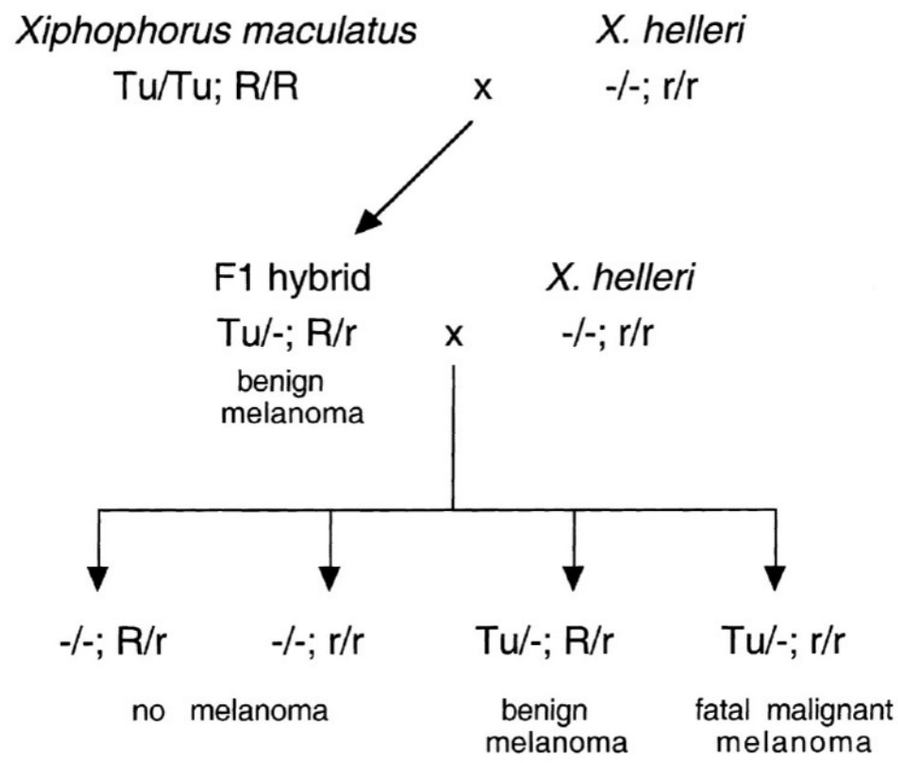
THE DOBZHANSKY-MULLER MODEL



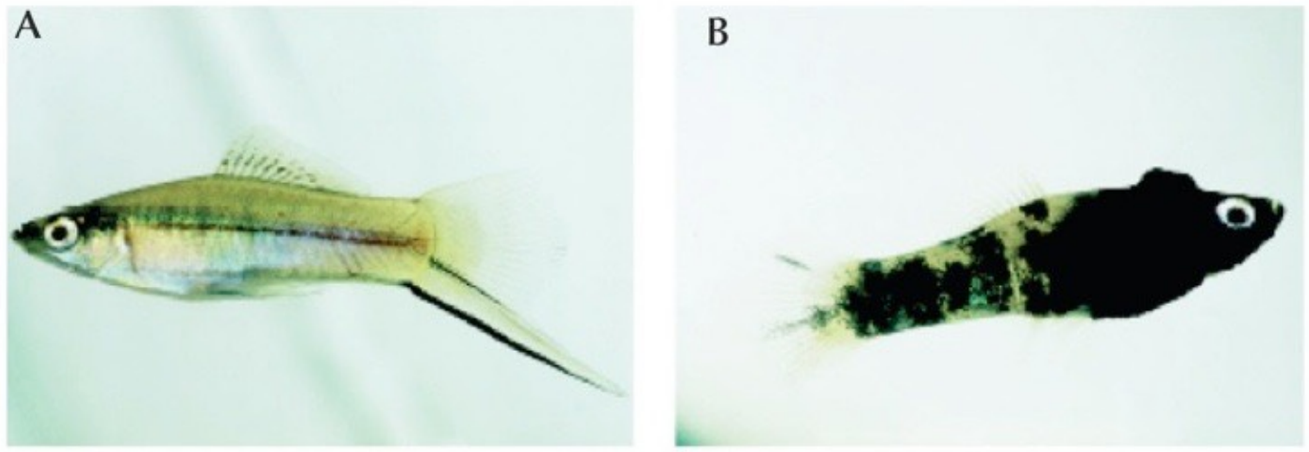
DMIS



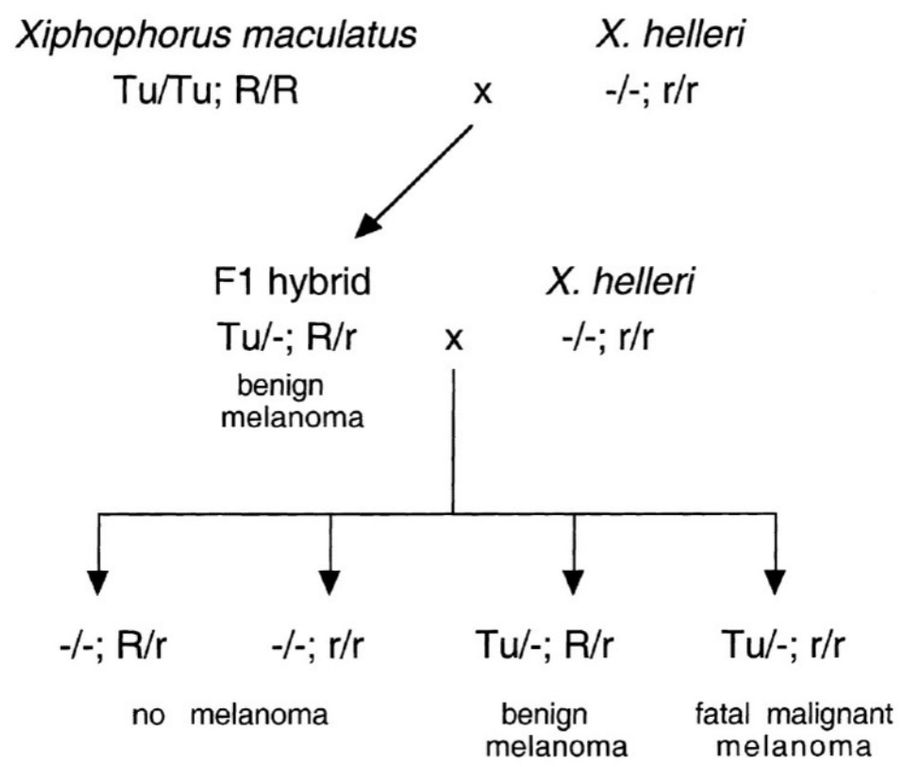
Orr & Presgraves, *Bioessays*, 2000



DMIS



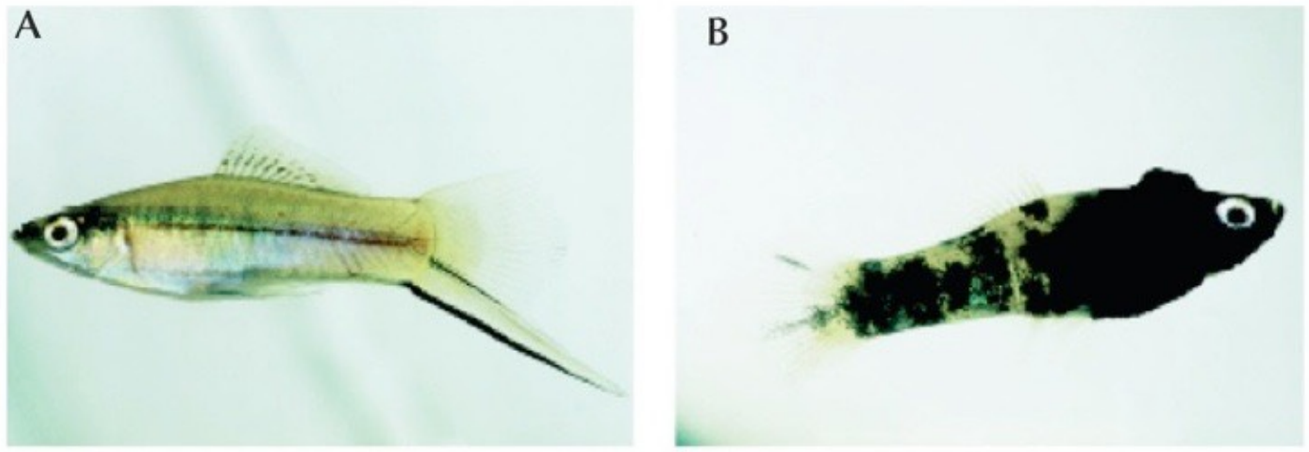
Orr & Presgraves, *Bioessays*, 2000



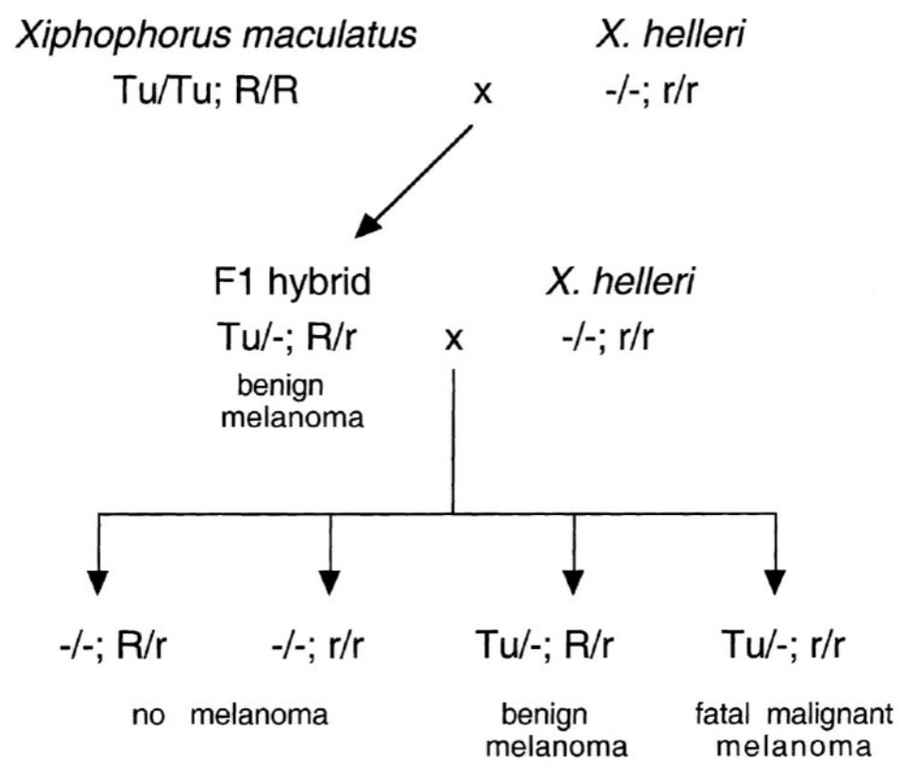
- sexual selection on tumor gene

Scarpino et al., MBE, 2013

DMIS



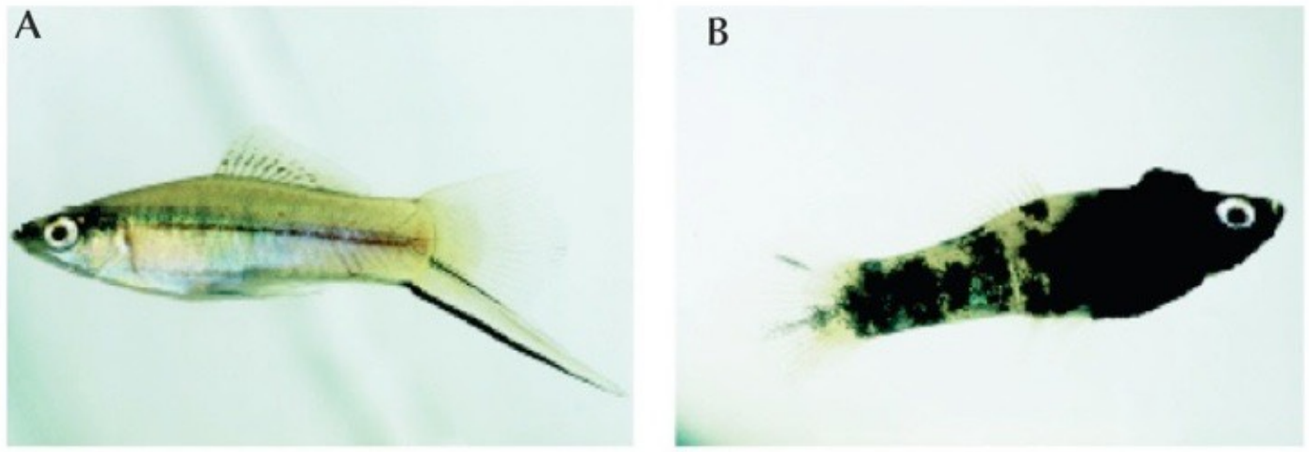
Orr & Presgraves, *Bioessays*, 2000



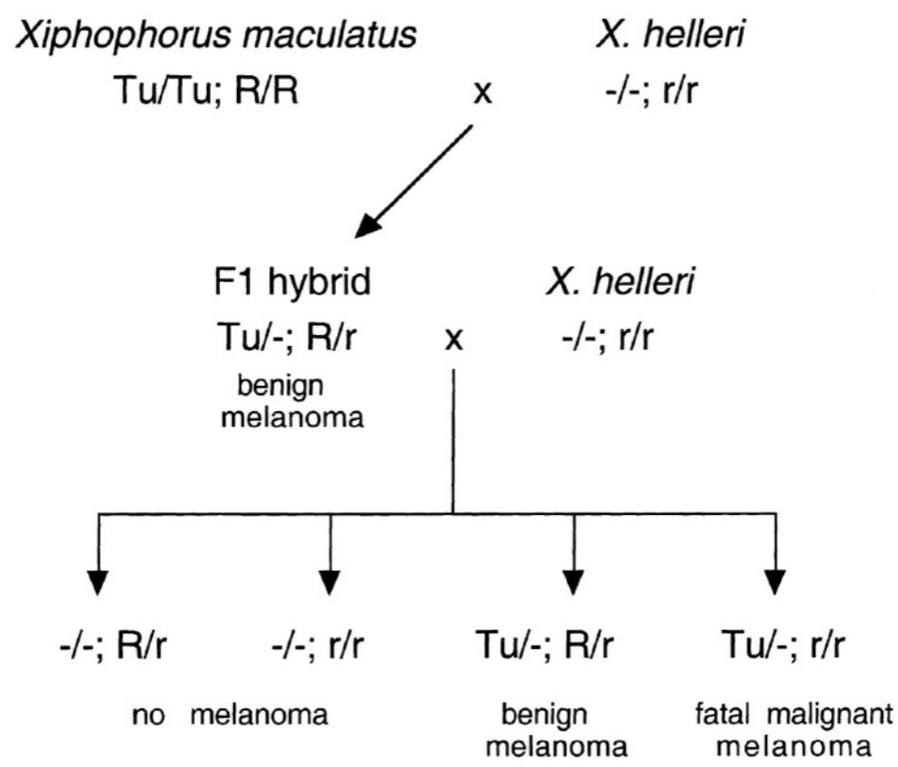
- sexual selection on tumor gene
- interaction with promoter of repressor gene

Scarpino et al., MBE, 2013

DMIS

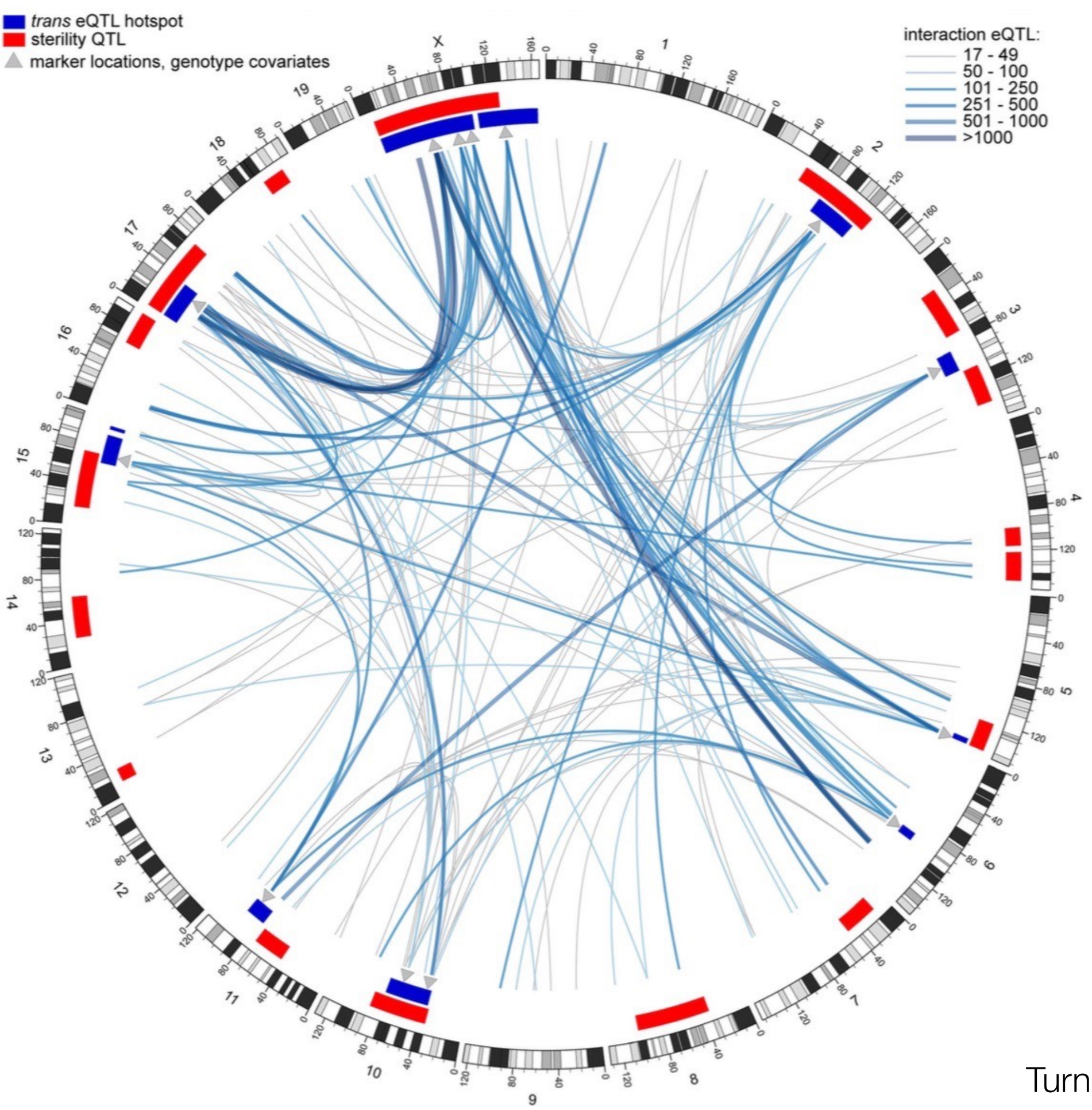


Orr & Presgraves, *Bioessays*, 2000



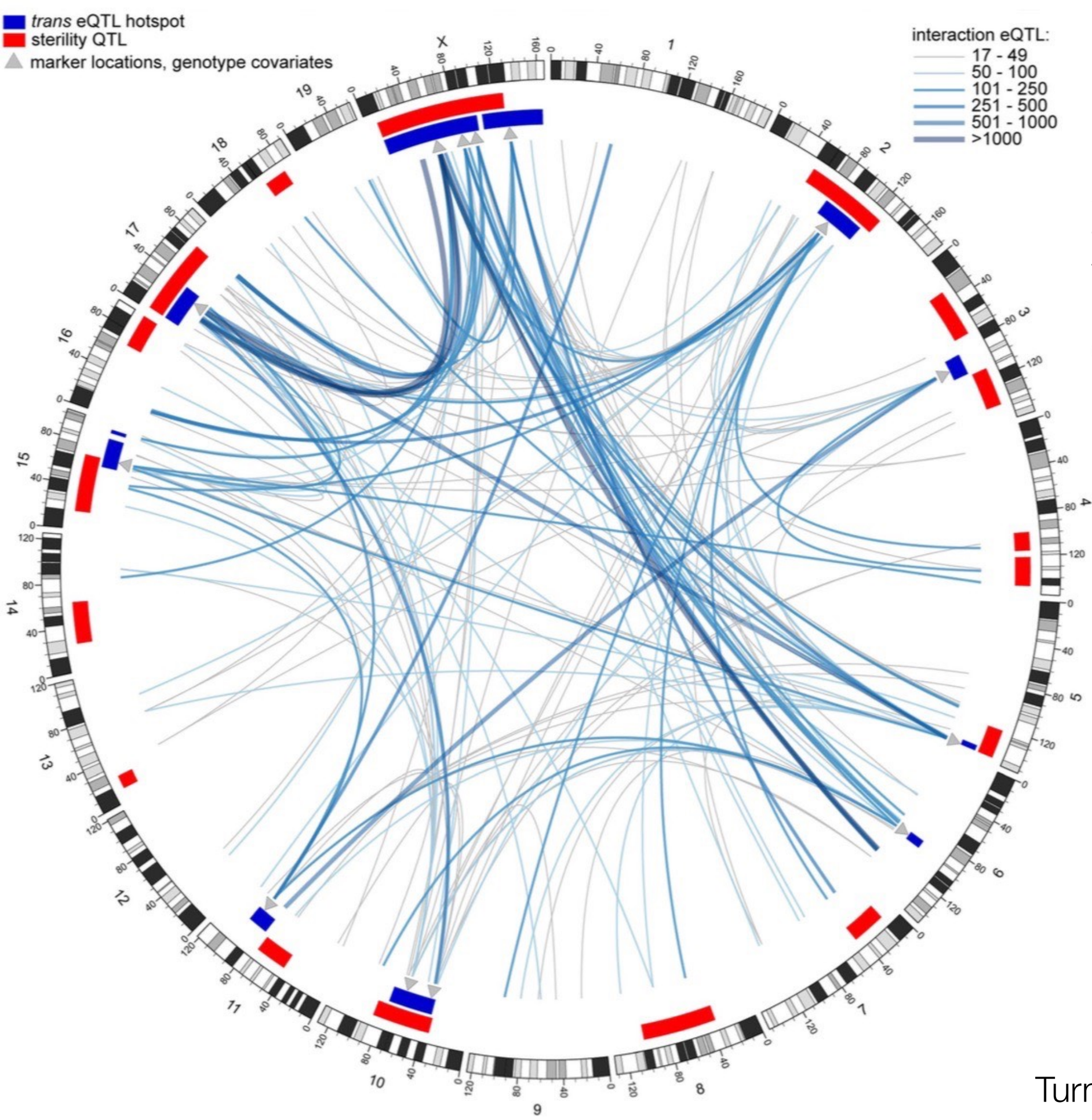
- sexual selection on tumor gene
- interaction with promoter of repressor gene
- ongoing gene flow

Scarpino et al., MBE, 2013



DMIS

Turner et al., *PlosGen*, 2014



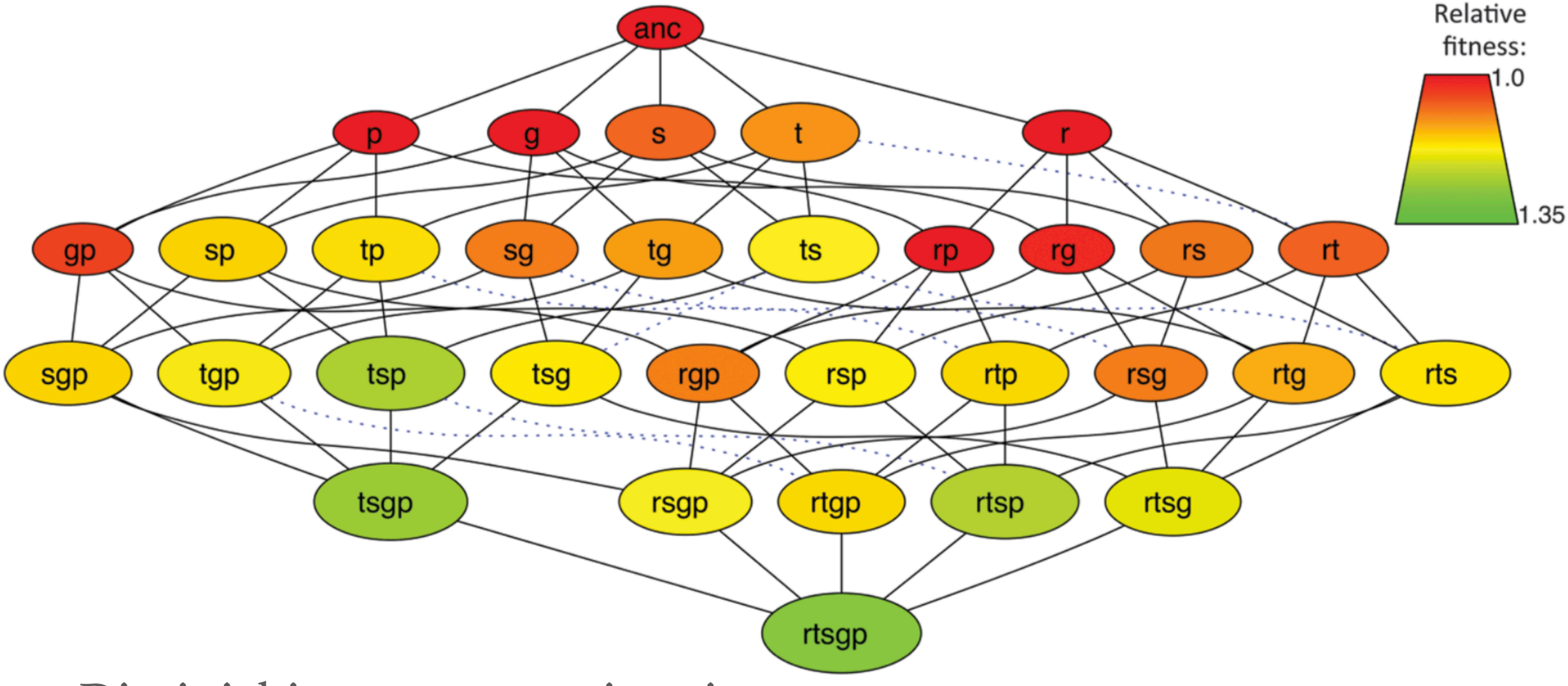
DMIS

In fact these were both "pathway" examples too…

Turner et al., *PlosGen*, 2014

TAKE A LOOK AT THE WHOLE FITNESS LANDSCAPE

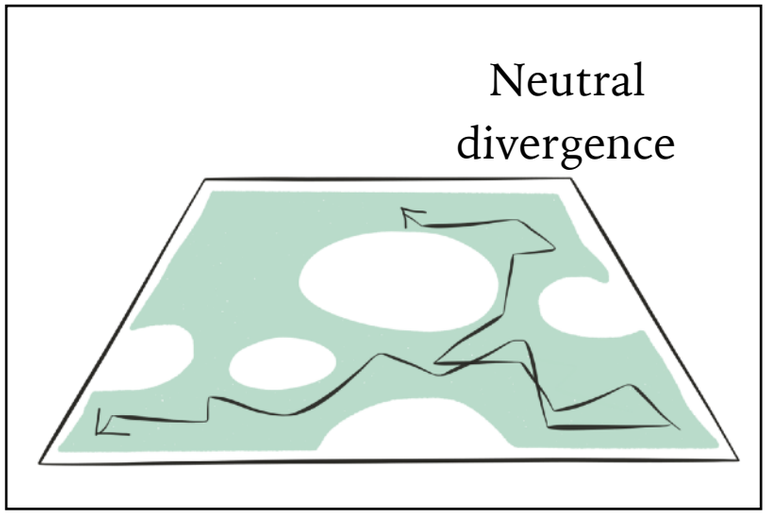
A FAMOUS EXAMPLE

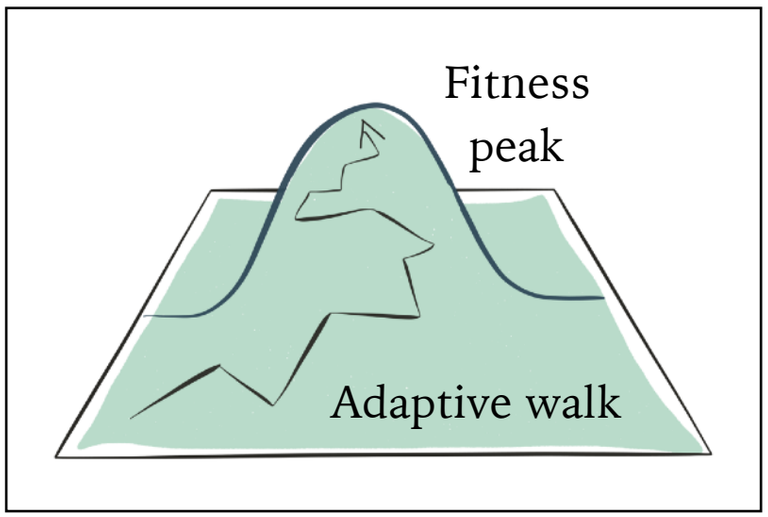


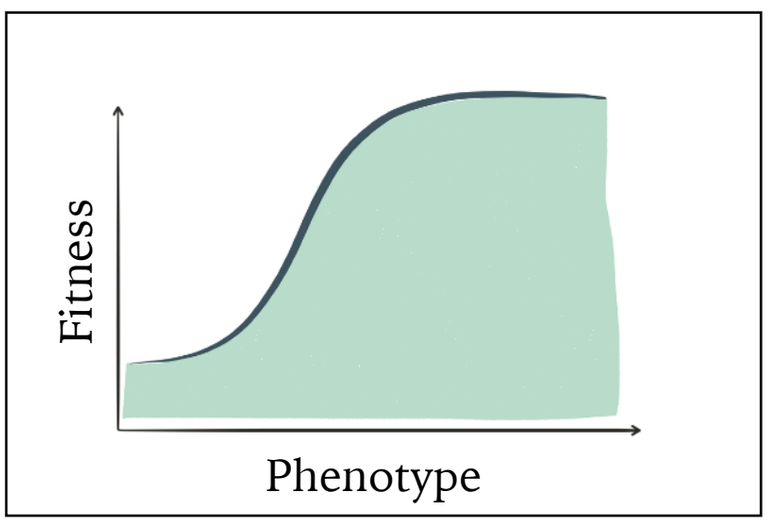
- Diminishing-returns epistasis
- Accessible but epistatic, single-peaked

Khan et al. 2011

WHY FITNESS LANDSCAPES ARE APPEALING

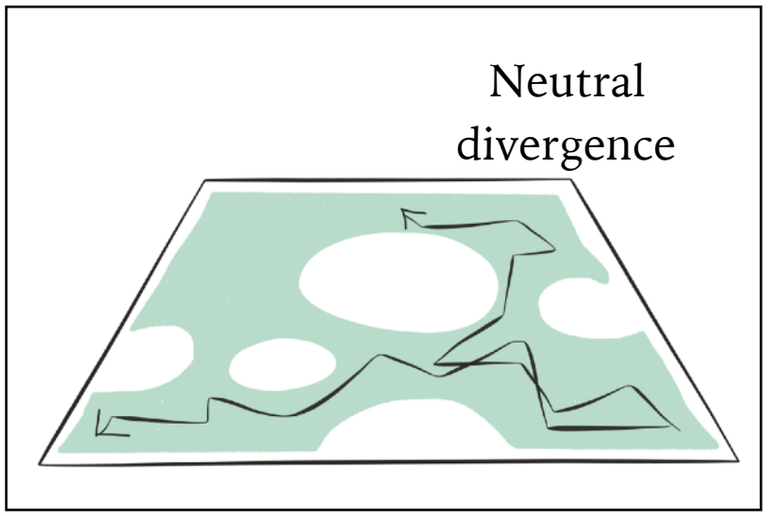


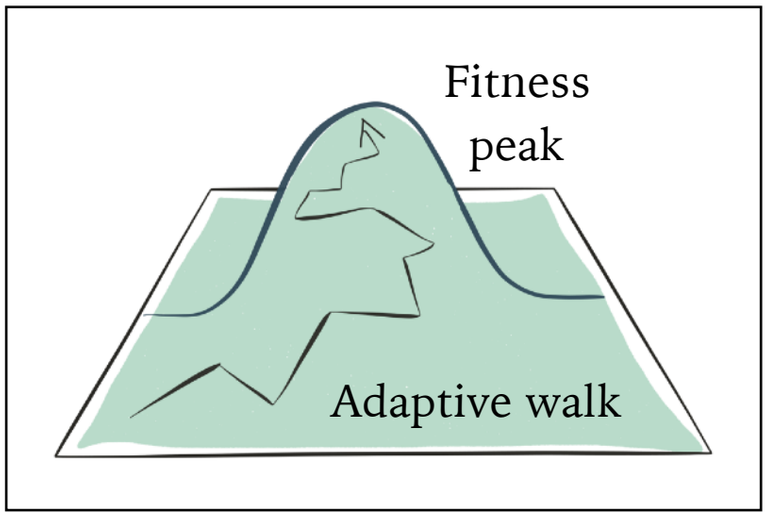


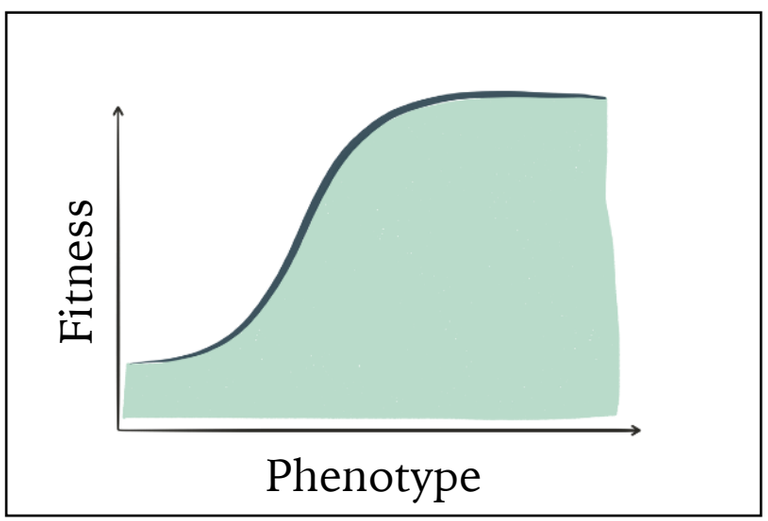


- fitness landscapes yield information on the predictability and repeatability of evolution

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

Comprehensive experimental fitness landscape and evolutionary network for small RNA

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

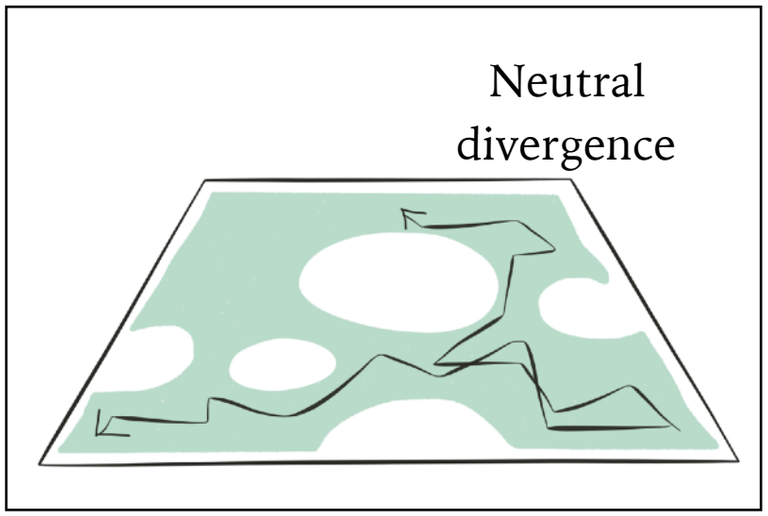
Biophysical principles predict fitness landscapes of drug resistance

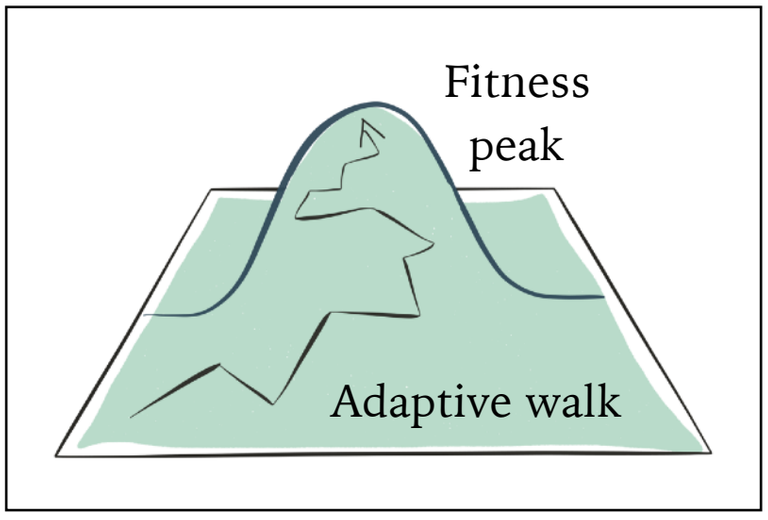
Mutational and fitness landscapes of an RNA virus revealed through population sequencing

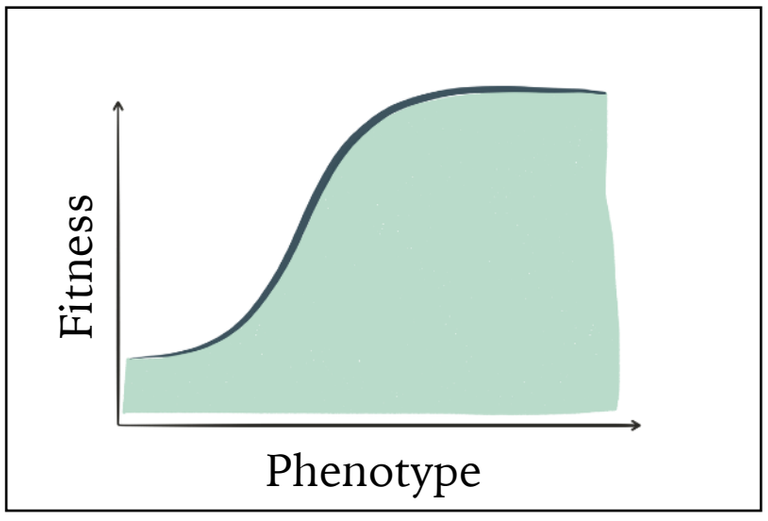
In-vivo mutation rates and fitness landscape of HIV-1

The fitness landscape of a tRNA gene

WHY FITNESS LANDSCAPES ARE DIFFICULT



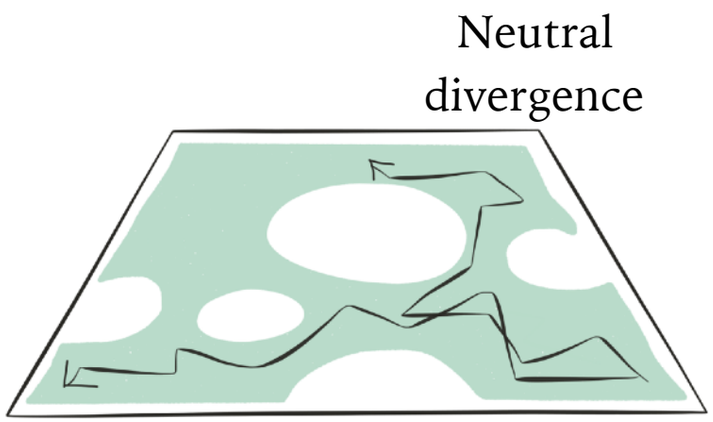


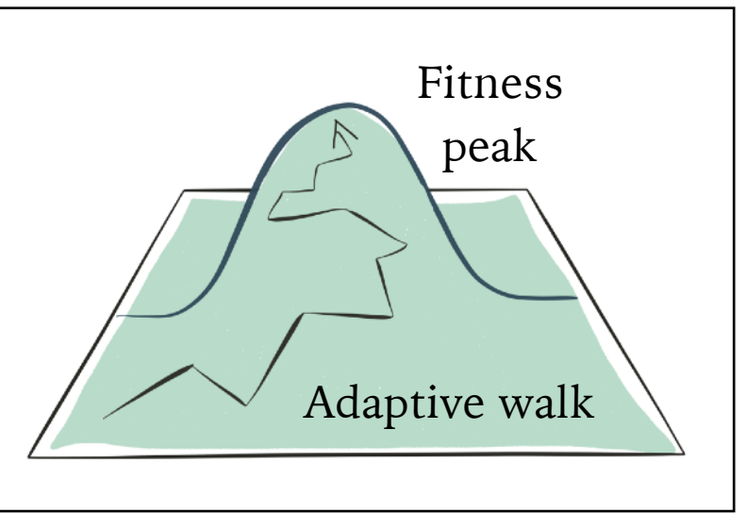


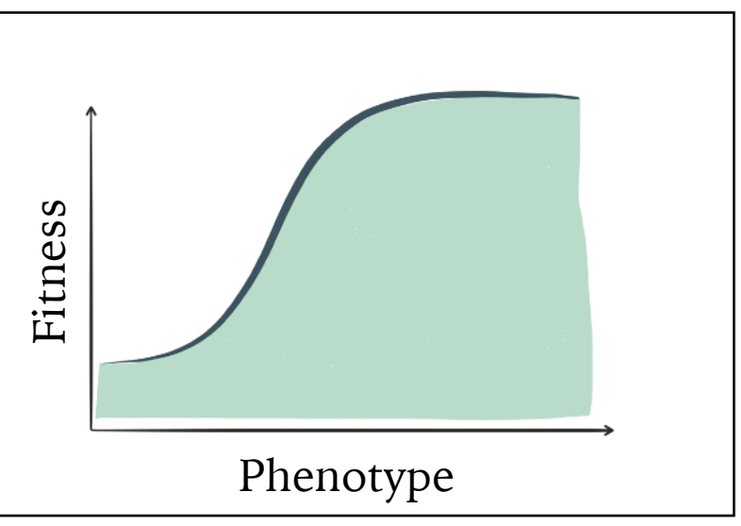
- fitness landscapes yield information on the predictability and repeatability of evolution
- it becomes increasingly simple to measure empirical fitness landscapes
- accumulating data on gene networks and pathways

But:

WHY FITNESS LANDSCAPES ARE DIFFICULT





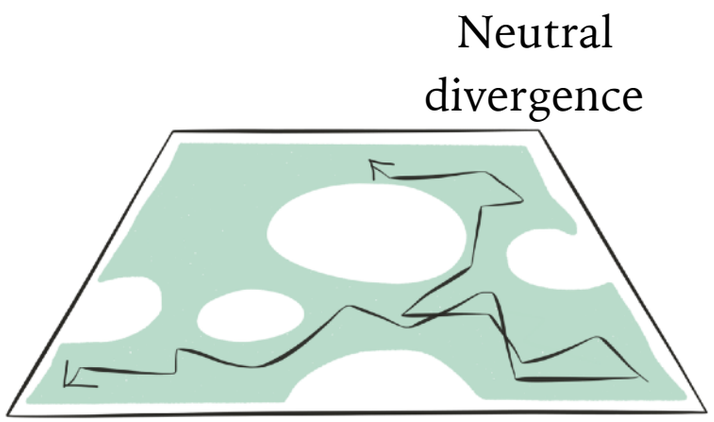


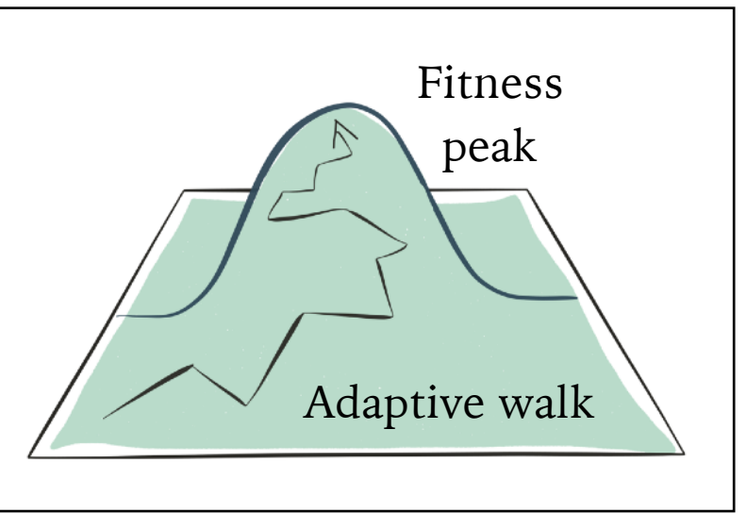
- fitness landscapes yield information on the predictability and repeatability of evolution
- it becomes increasingly simple to measure empirical fitness landscapes
- accumulating data on gene networks and pathways

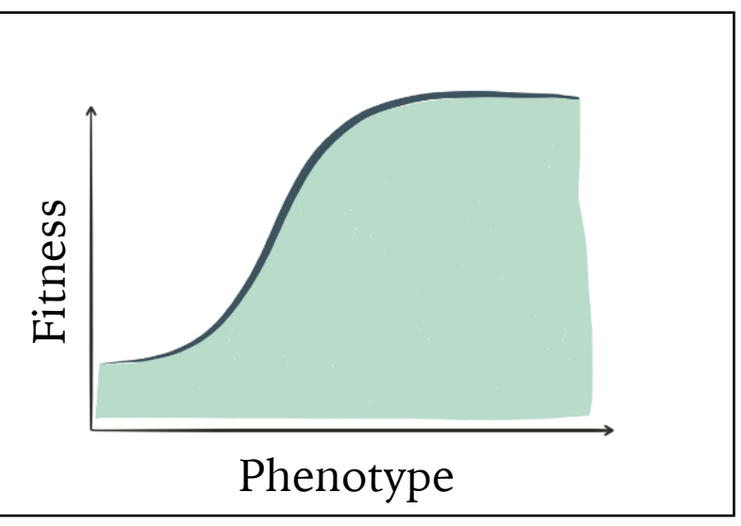
But:

- 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?

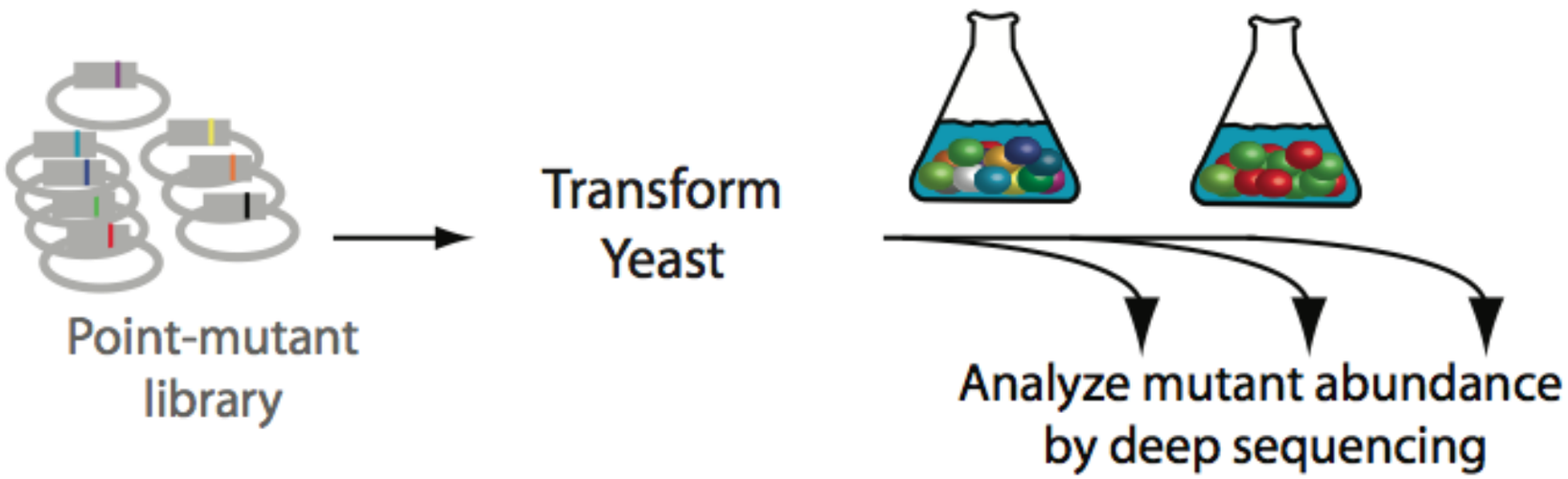
THE EXPERIMENTAL APPROACH: DEEP MUTATIONAL SCANNING

- 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)

Dan Bolon

Ryan Hietpas

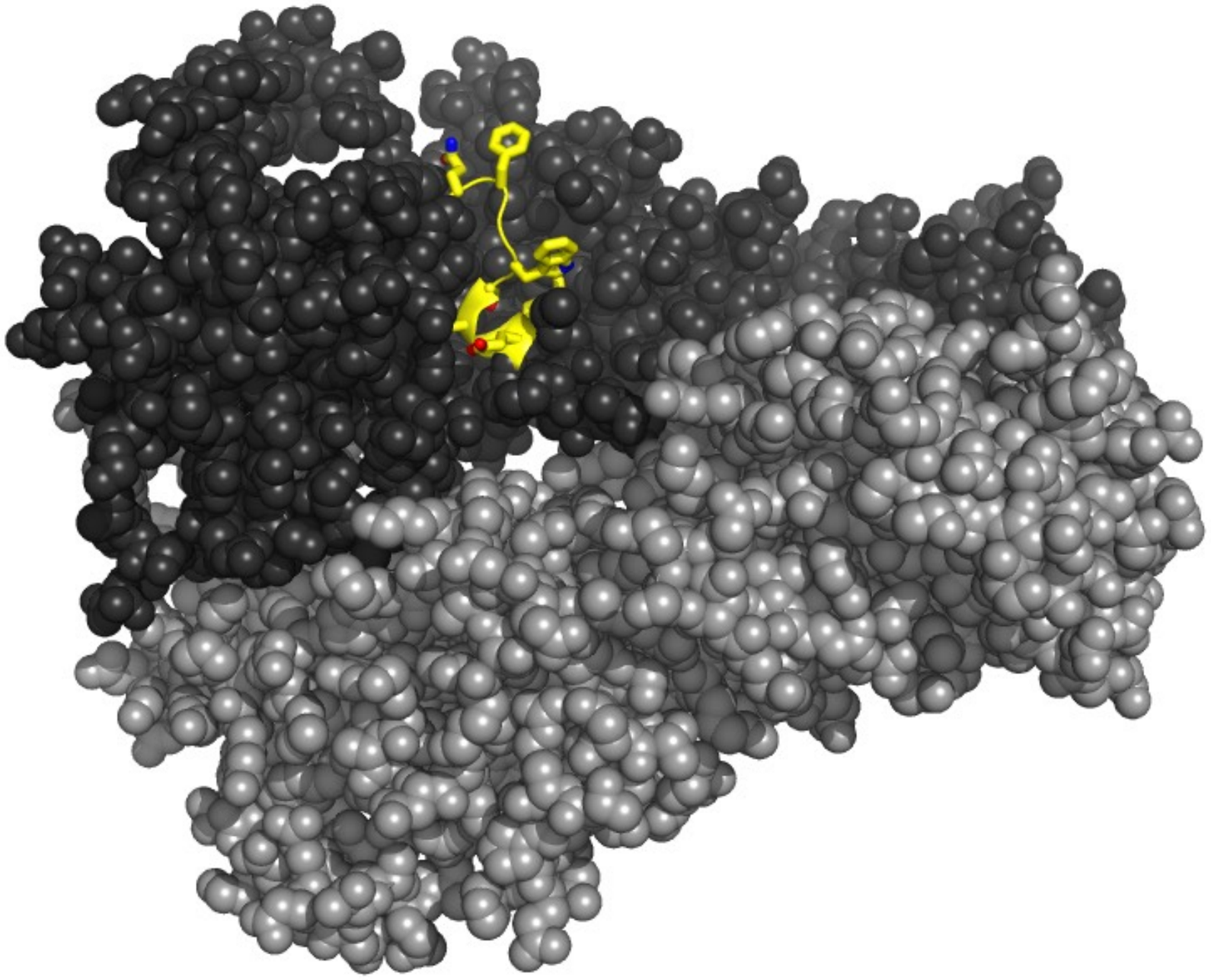
Jeff Jensen



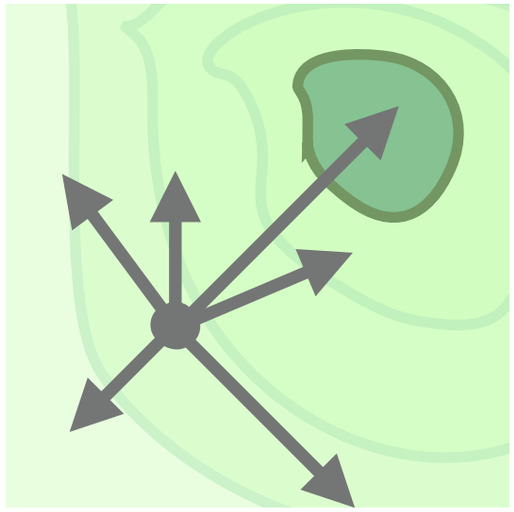
Hietpas, Jensen & Bolon, PNAS, 2011

WHAT WE DID SO FAR

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



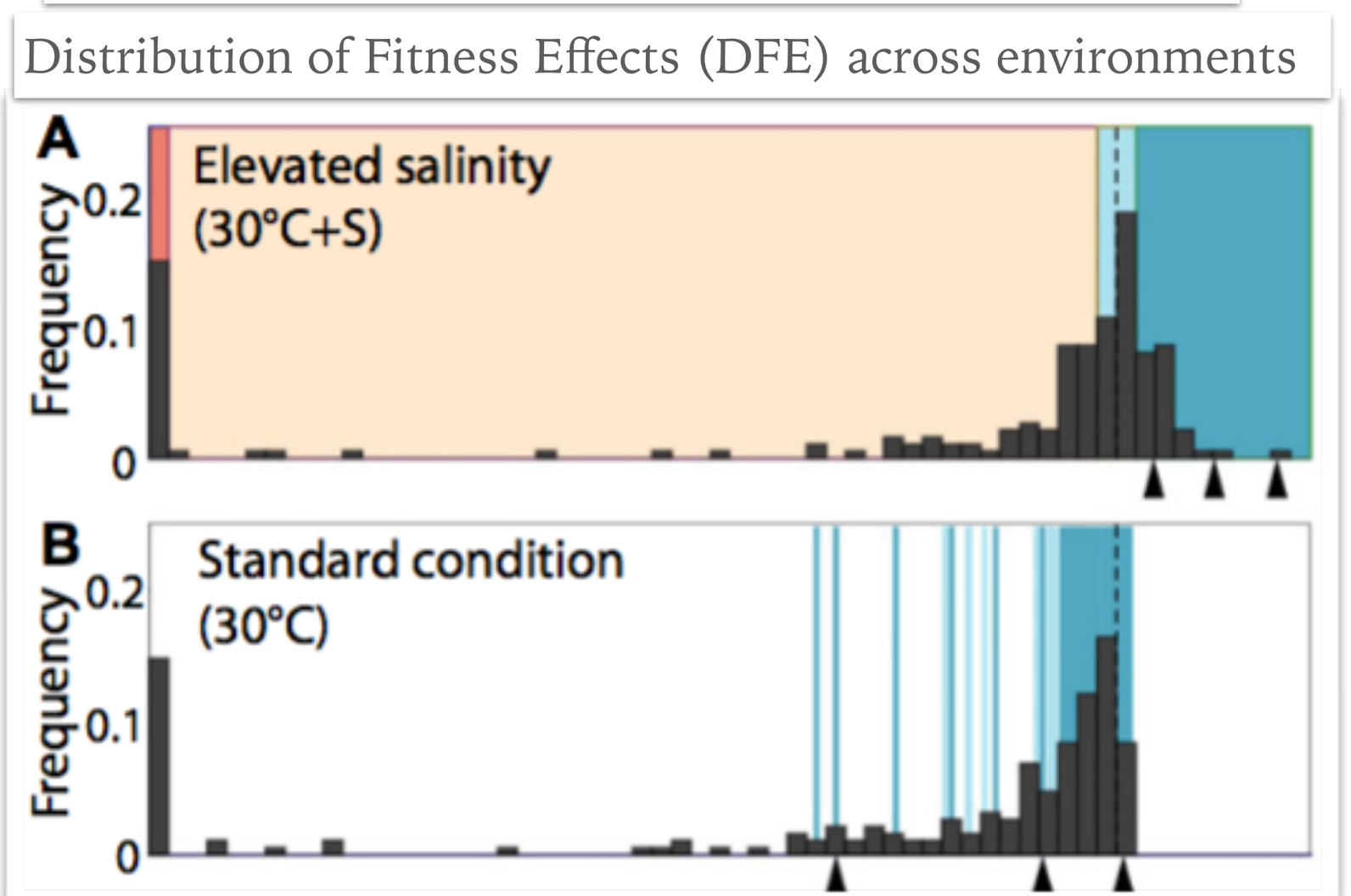
WHAT WE DID SO FAR



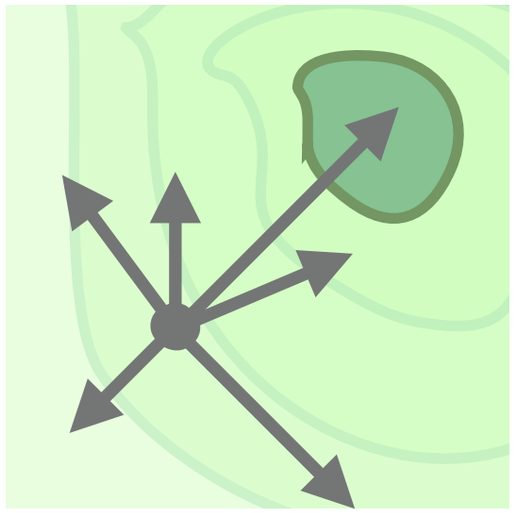
Engineered mutations from a 9 aa region from Hsp90 (aa positions 582-590) in *Saccharomyces cerevisiae*

Compare DFEs across 4 environments - high costs of adaptation

Distribution of Fitness Effects (DFE) across environments



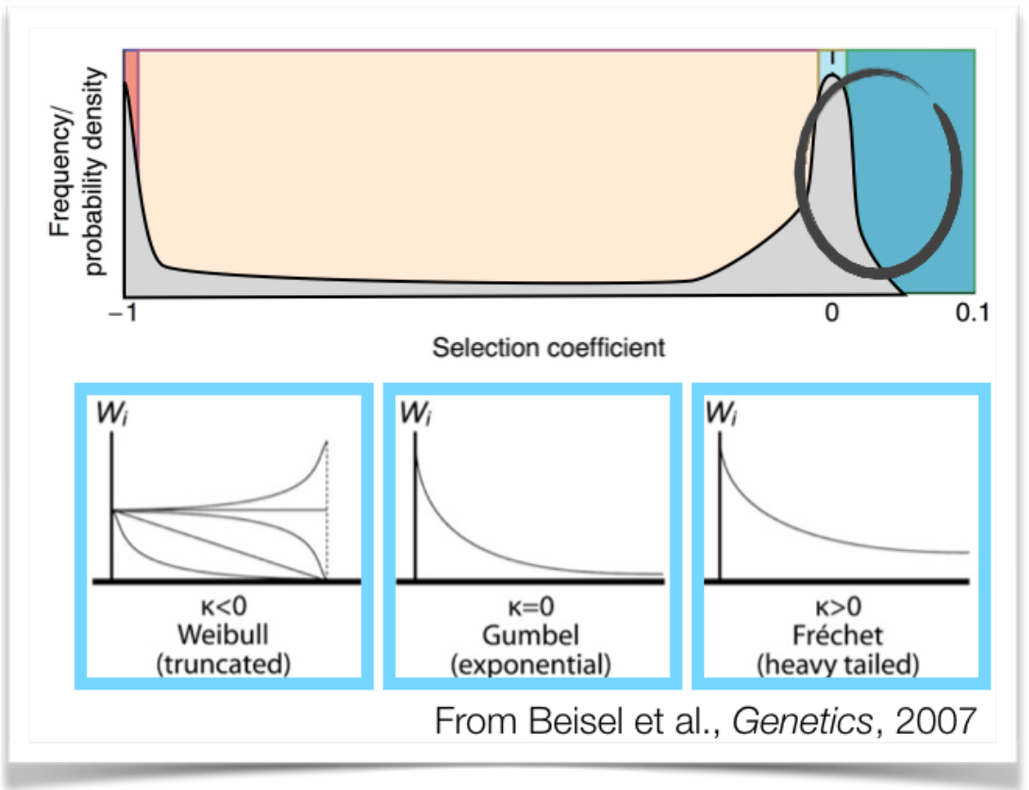
WHAT WE DID SO FAR



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

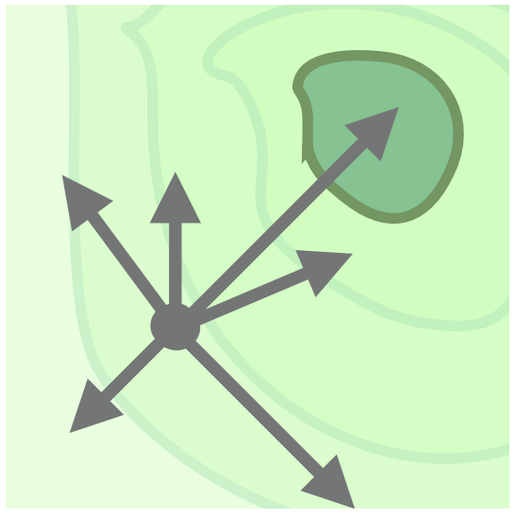
Compare DFEs across 4 environments - high costs of adaptation

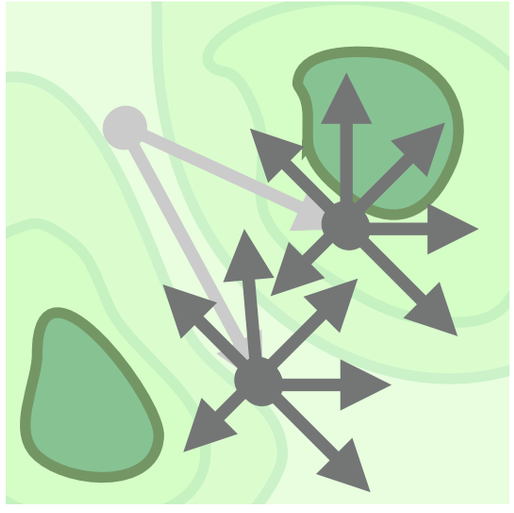
MCMC method to estimate selection coefficients, and DFEs across 6 environments - heavy-tailed DFE for most challenging environment

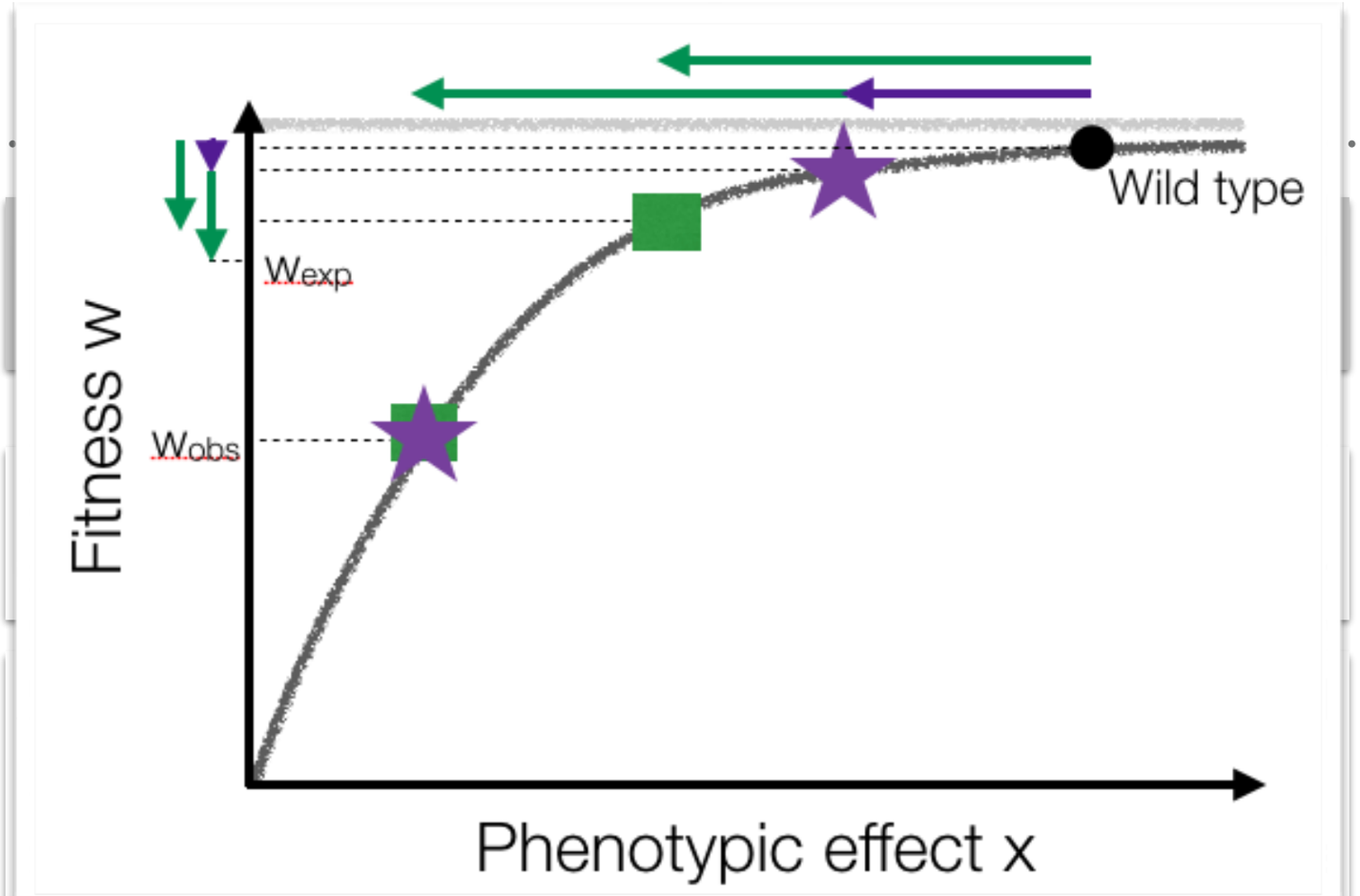


From Beisel et al., *Genetics*, 2007

WHAT WE DID SO FAR

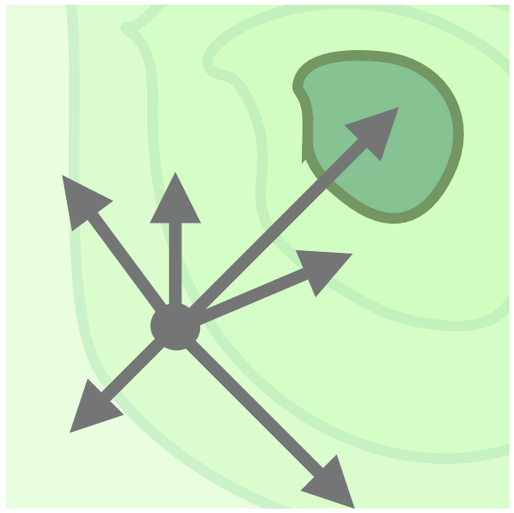


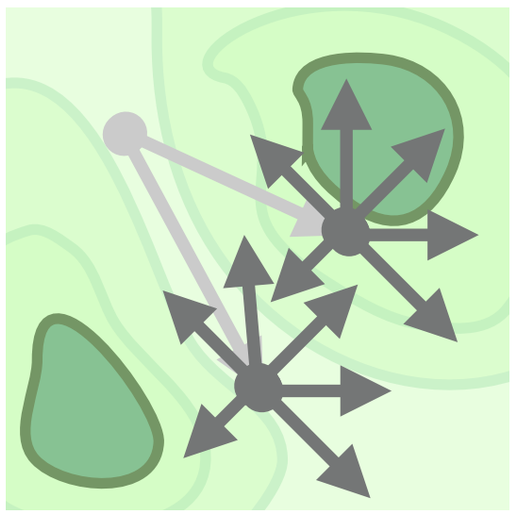


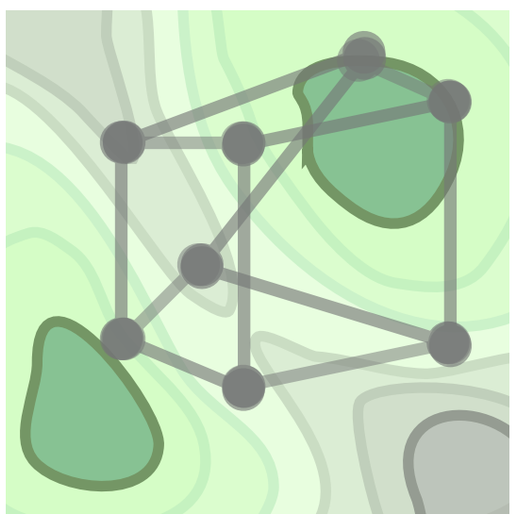


DFEs on 7 genetic backgrounds - ubiquitous negative epistasis indicating an underlying concave fitness landscape

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**

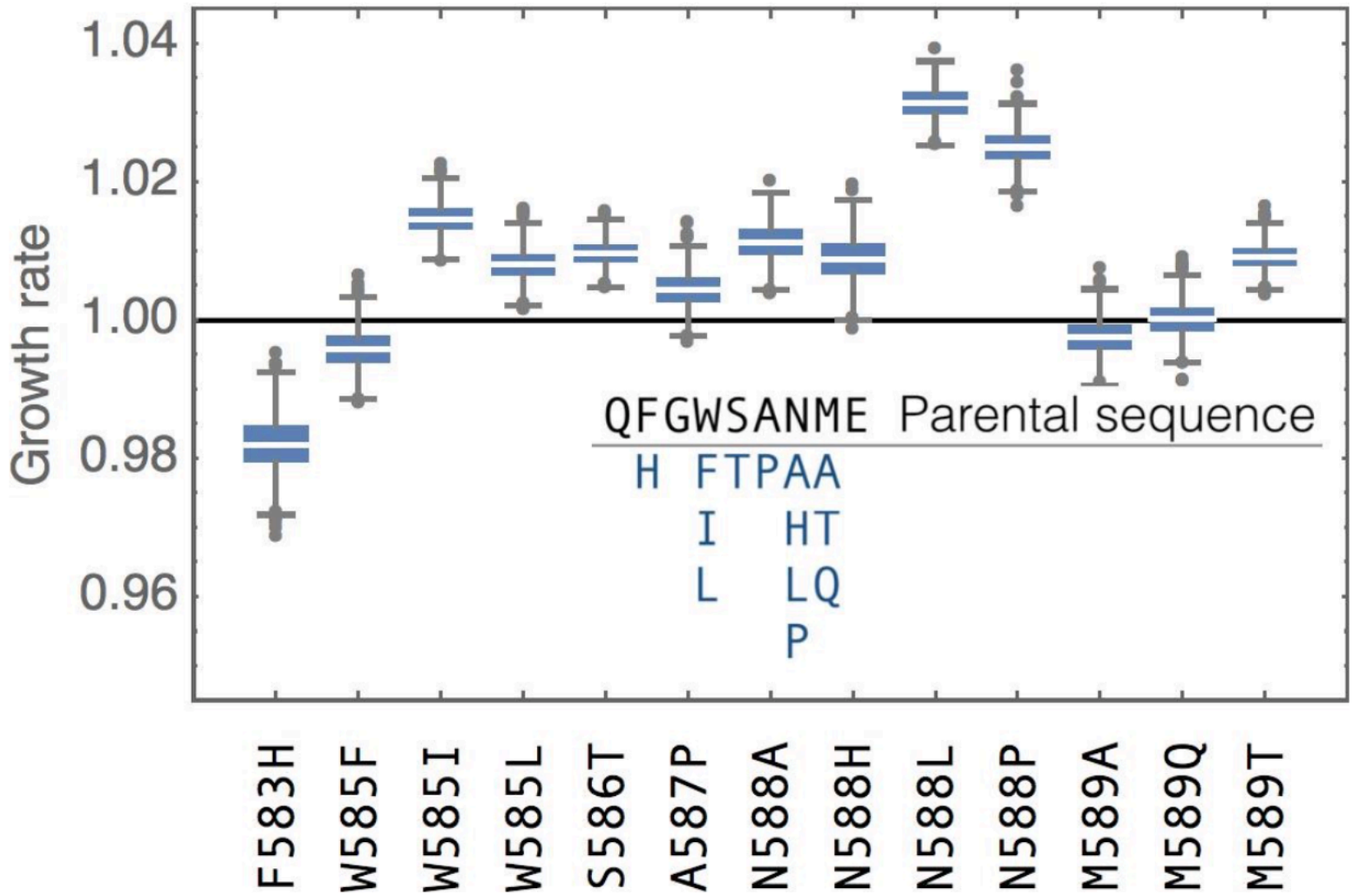
MCMC method to estimate selection coefficients, and DFEs across 6 environments - heavy-tailed DFE for most challenging environment **Bank et al., 2014, Genetics**

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

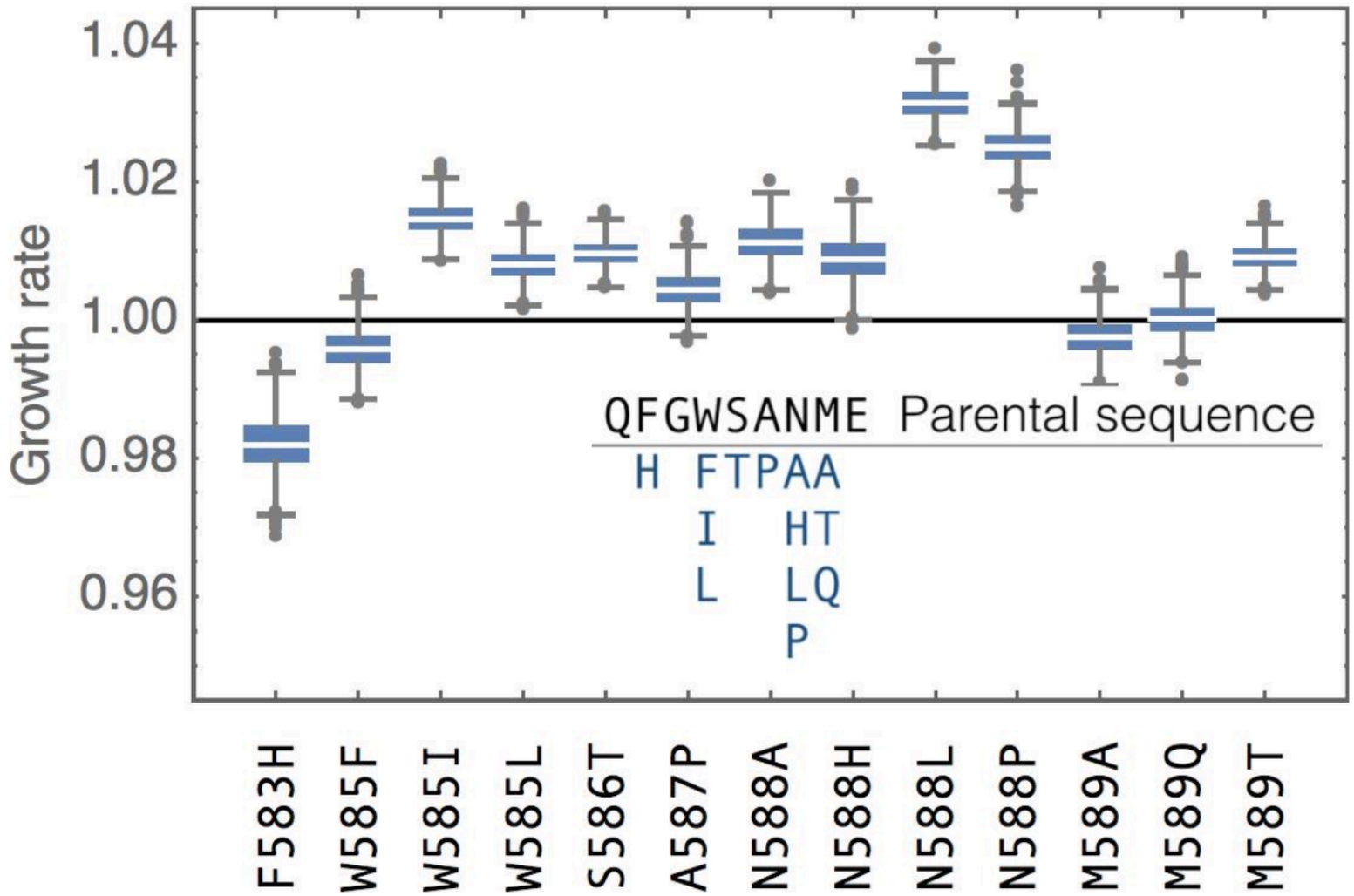
Guide to experimental design of deep mutational scanning studies

Complete fitness landscape of 640 combinations of mutations

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

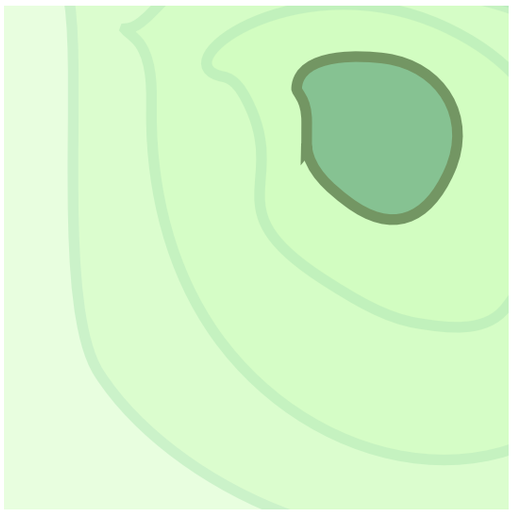


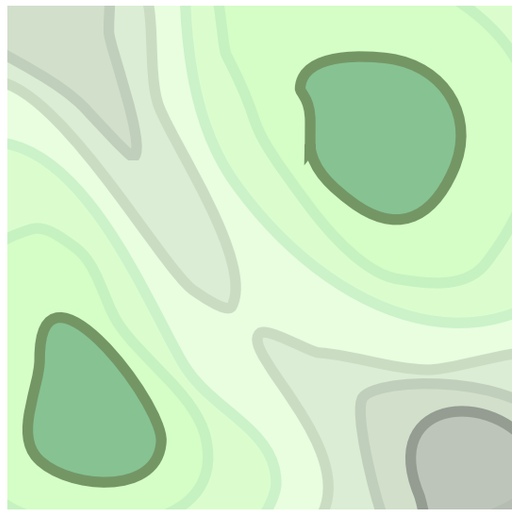
high salinity environment
13 single-aa mutations
2 replicates
all possible combinations of aa's
≈1600 nt mutations



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

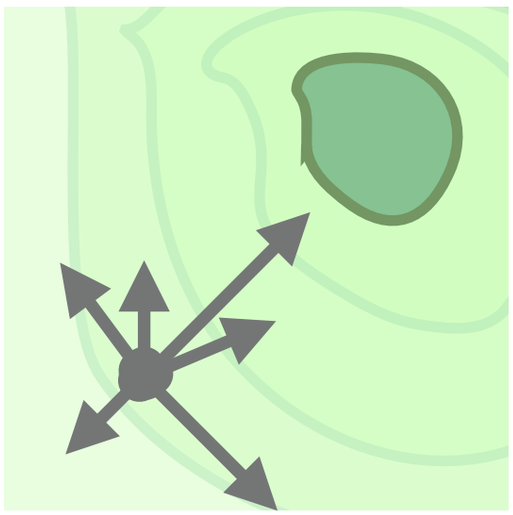
3 QUESTIONS

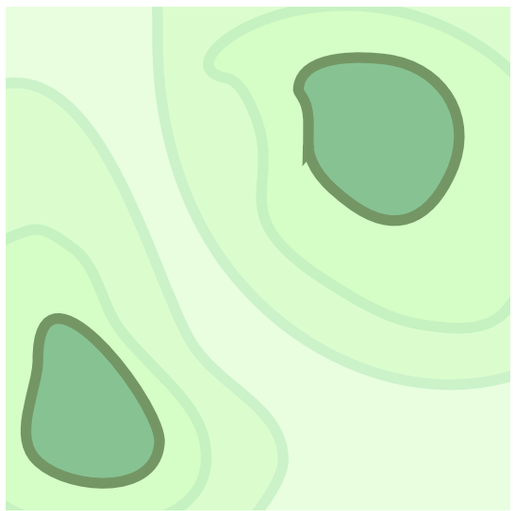


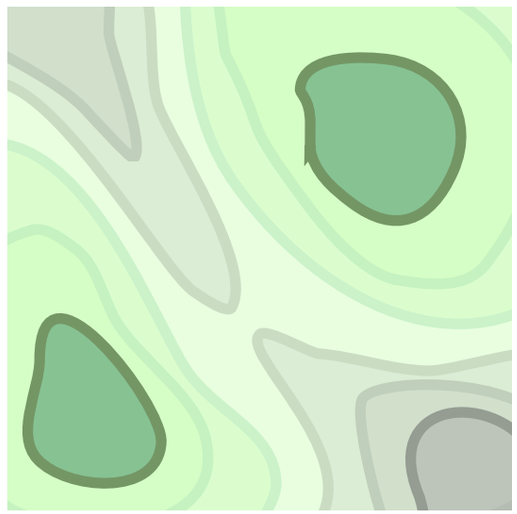


3 QUESTIONS

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

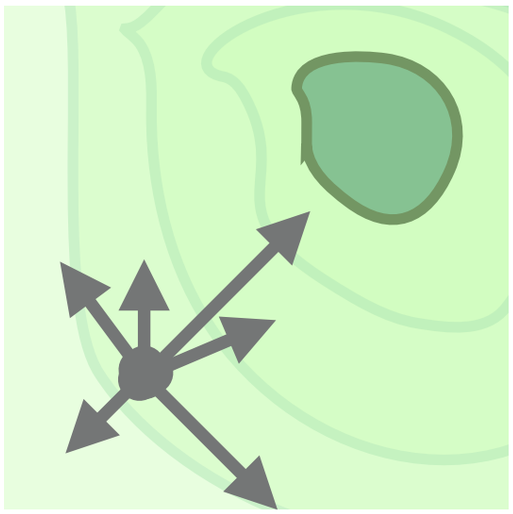


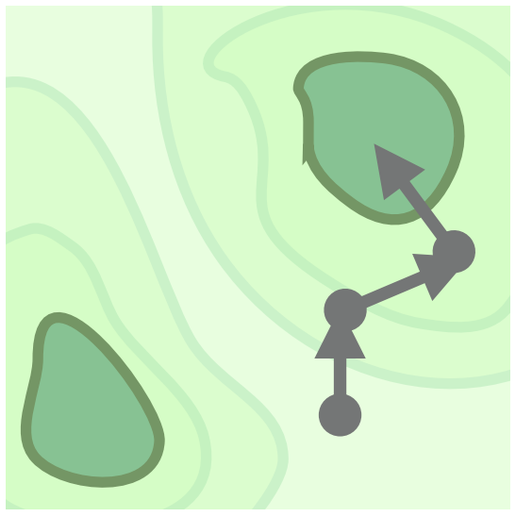


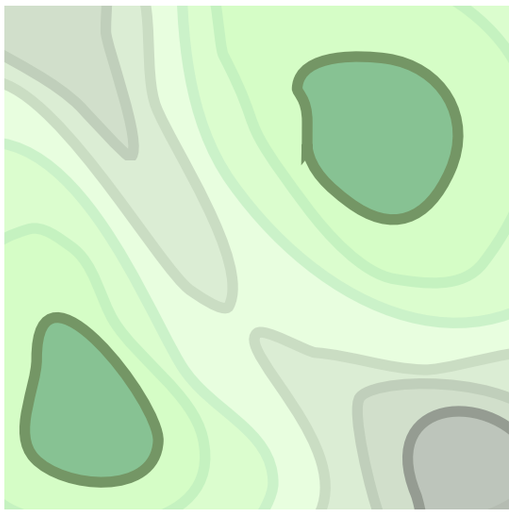


3 QUESTIONS

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

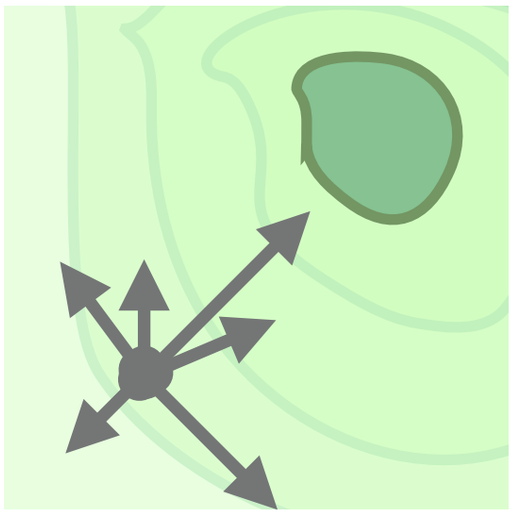


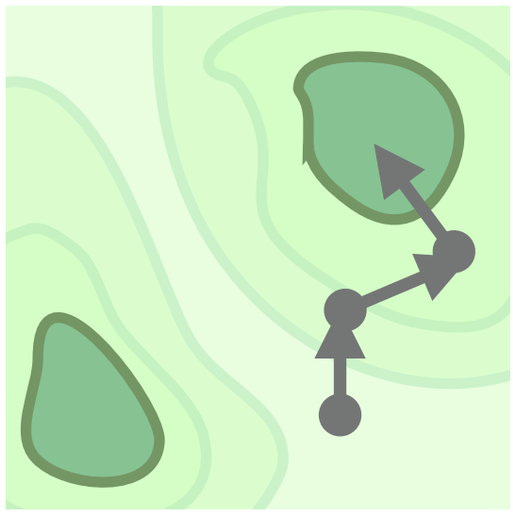


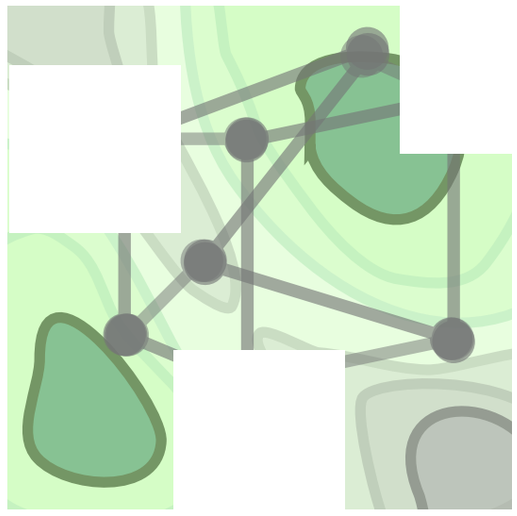


3 QUESTIONS

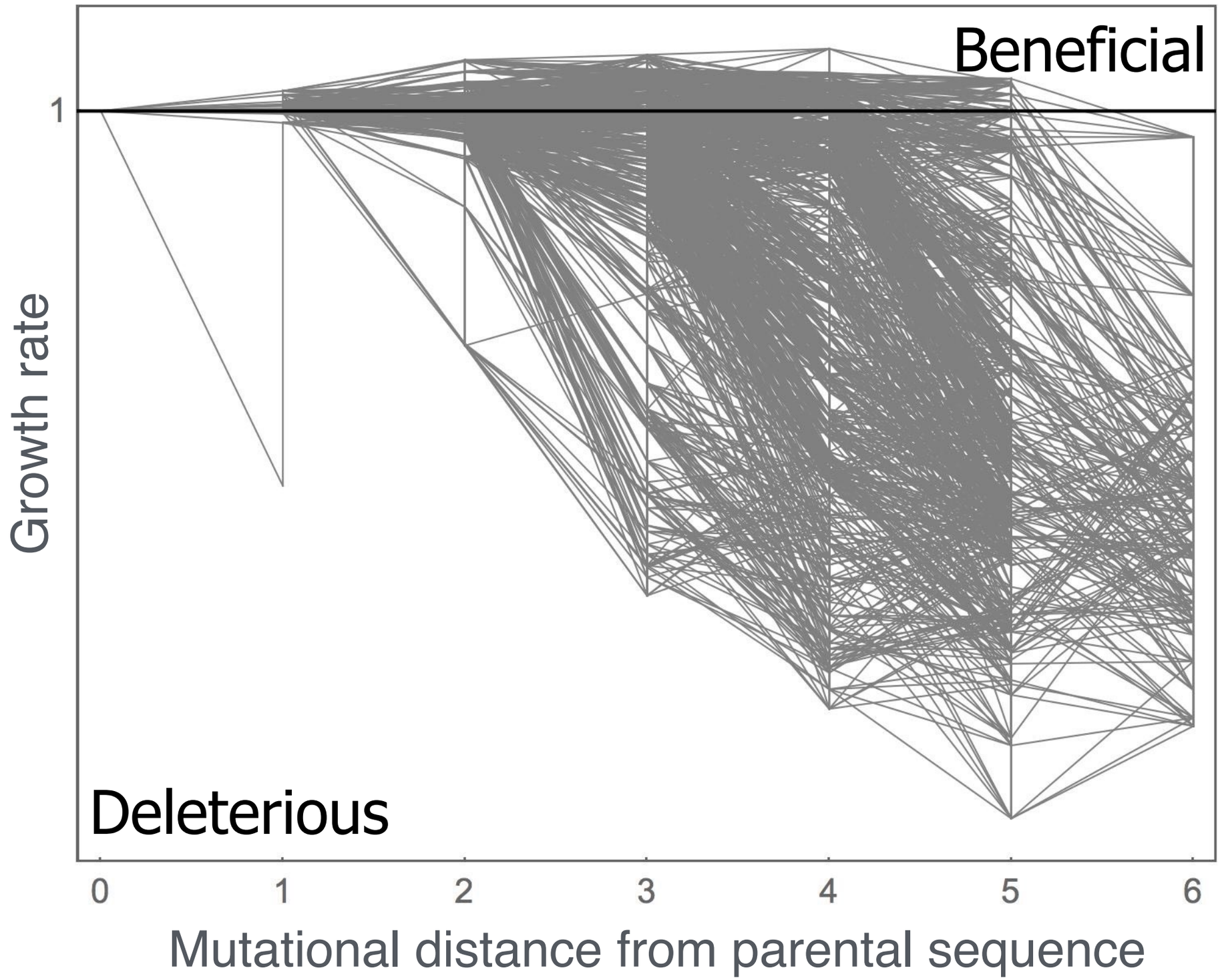
- Do single step mutations predict the way to the global optimum?
- Will adaptation take the population to the global optimum?
- Can we infer an unknown part of the fitness landscape?



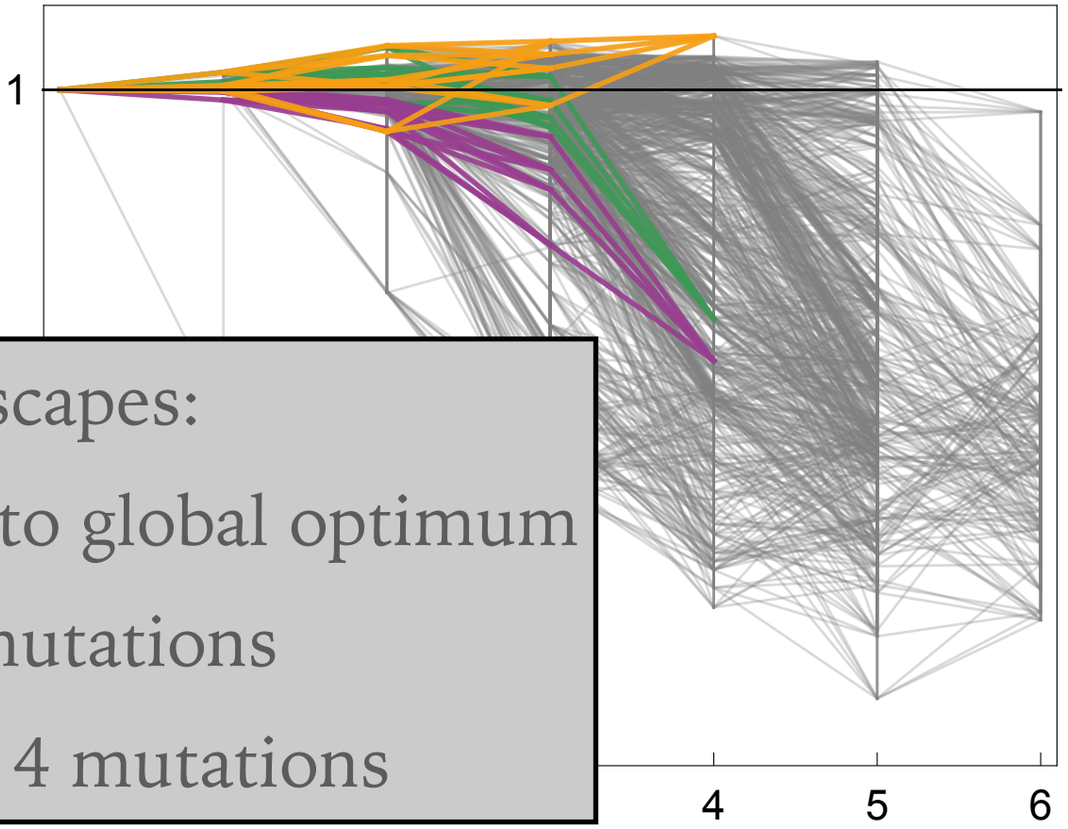




A PICTURE OF THE WHOLE LANDSCAPE



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

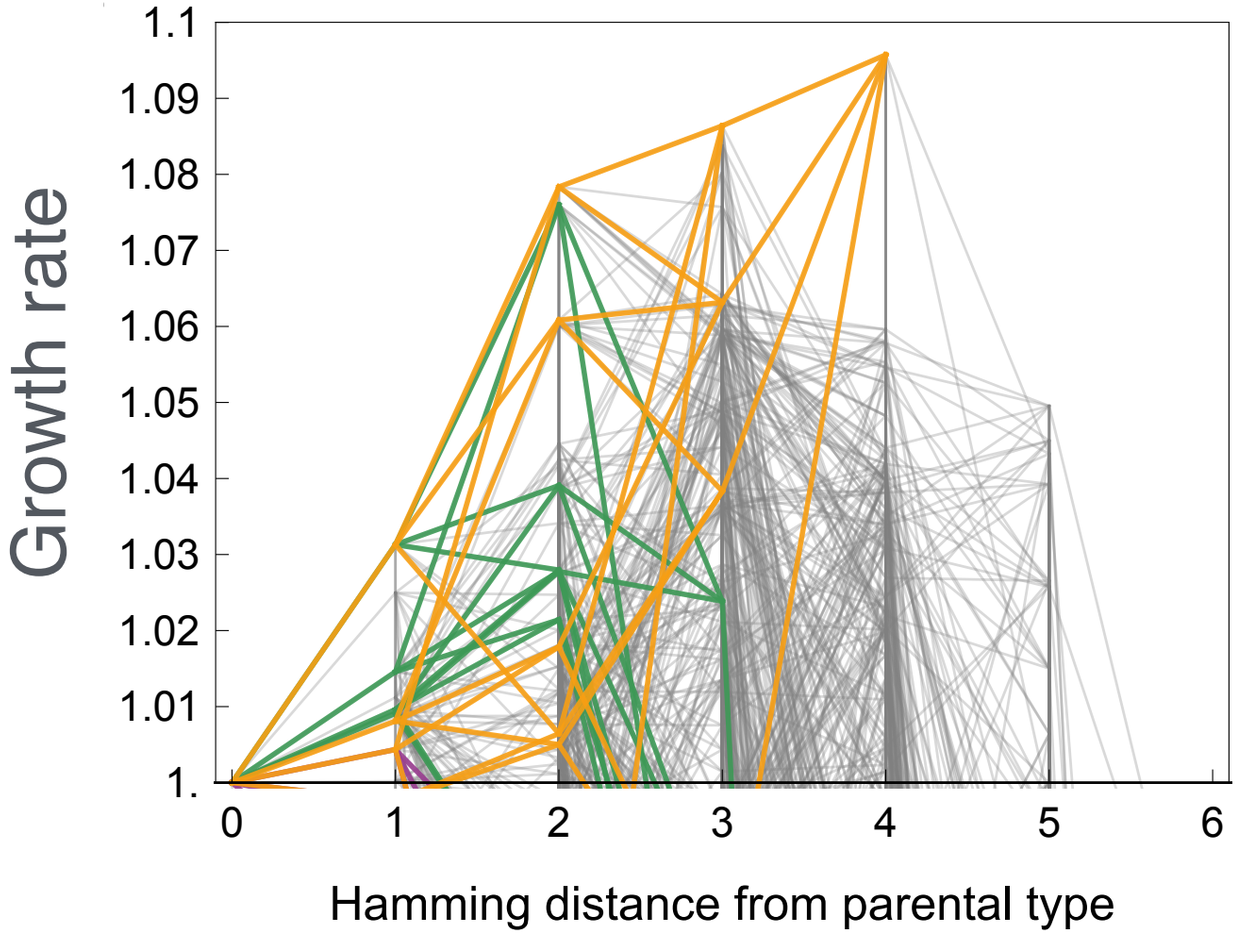


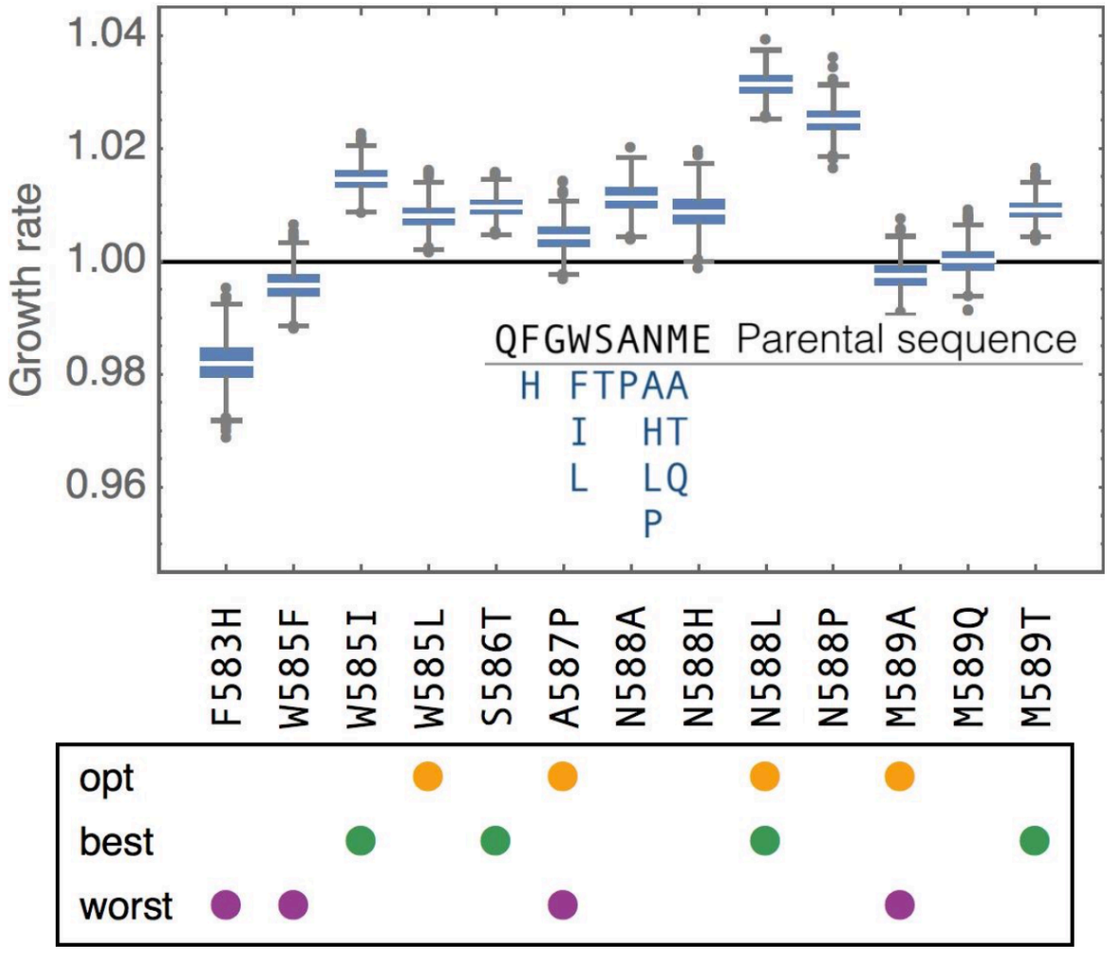
Focal landscapes:

1) leading to global optimum

2) best 4 mutations

3) "worst" 4 mutations





Growth rate

1
0 1 2 3 4 5 6

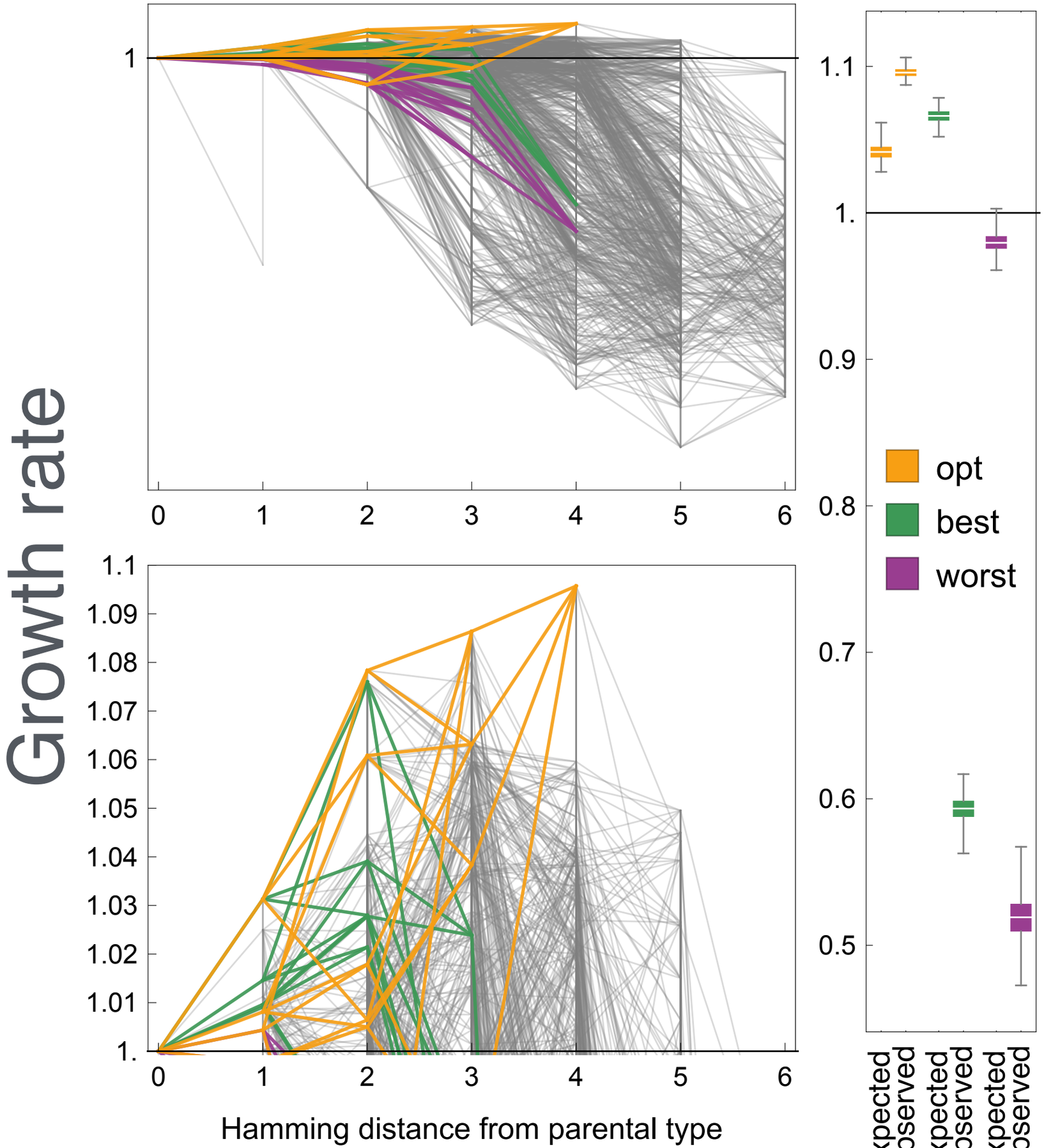
1.1
1.09
1.08
1.07
1.06
1.05
1.04
1.03
1.02
1.01
1.
0 1 2 3 4 5 6

Hamming distance from parental type

1.1
1.
0.9
0.8
0.7
0.6
0.5

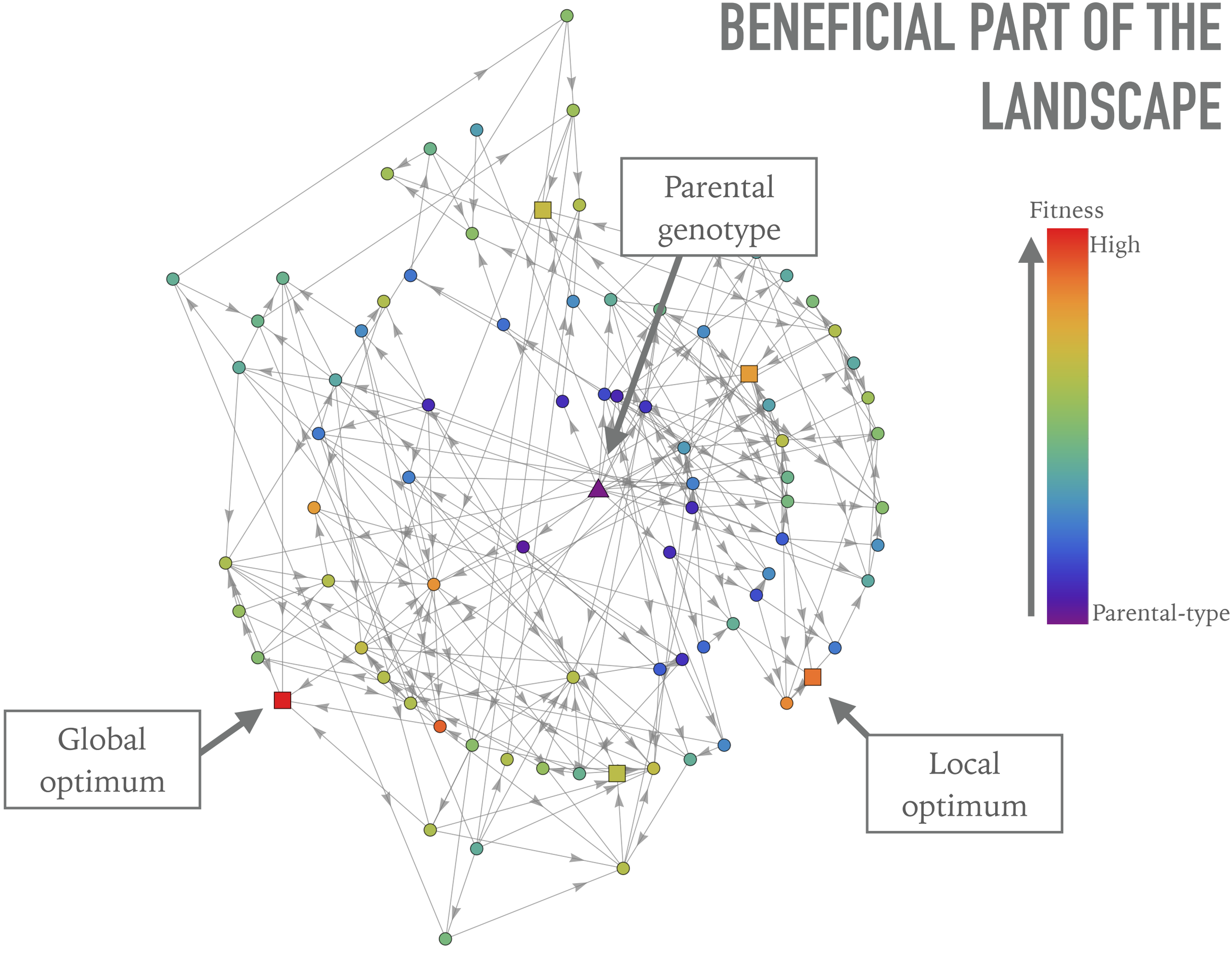
opt
best
worst

Expected
Observed
Expected
Observed
Expected
Observed

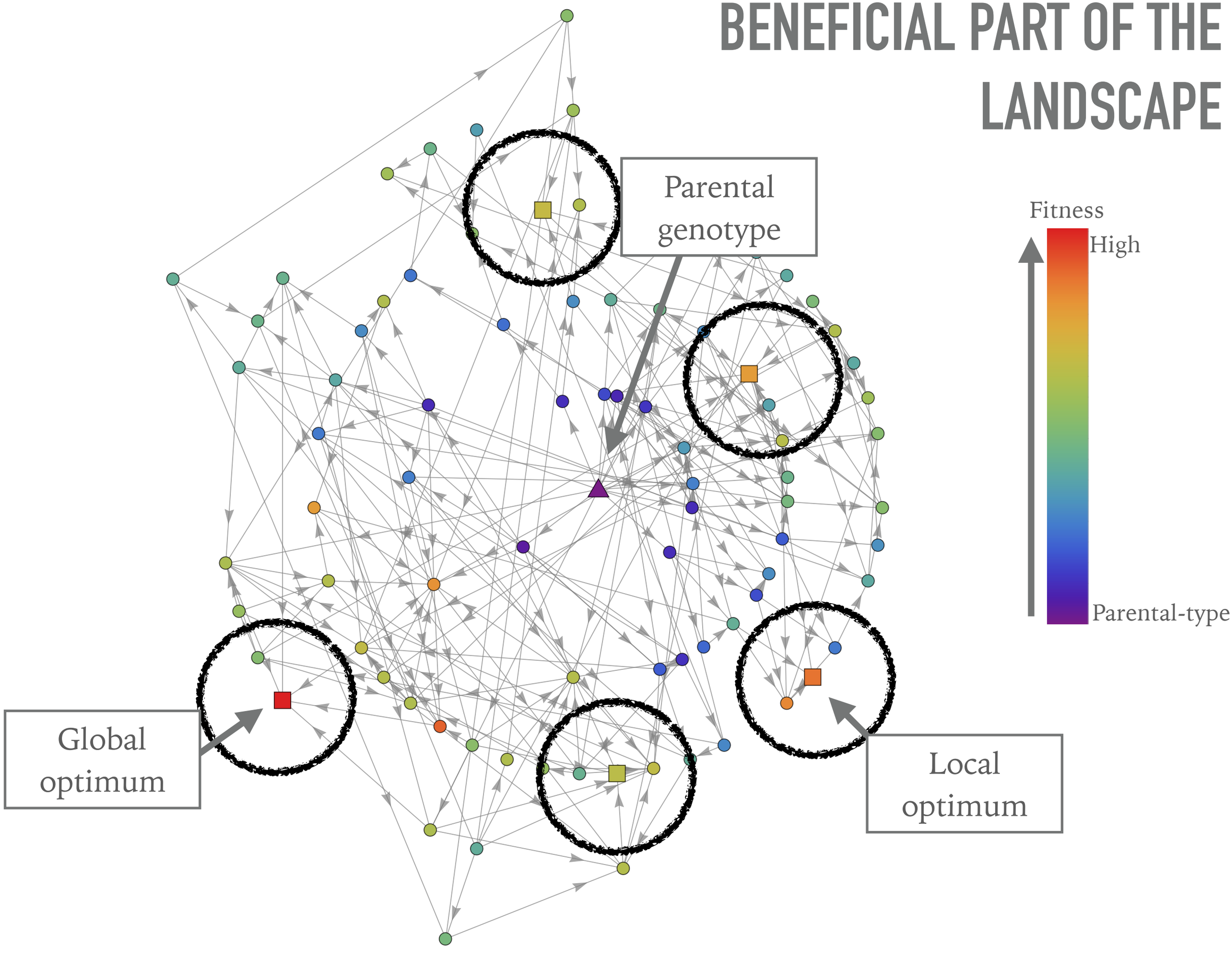


Strong positive and negative epistasis in the landscape.

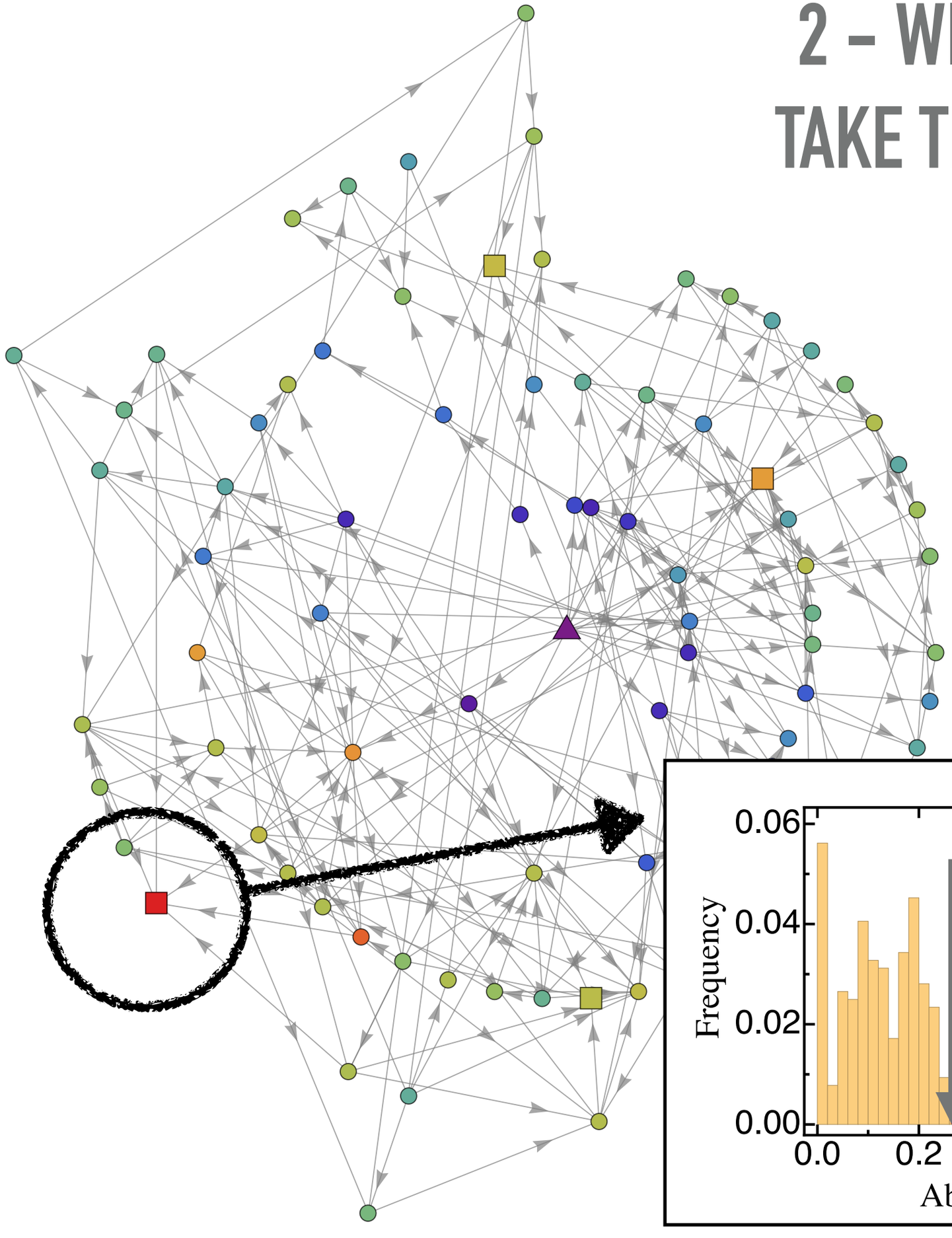
BENEFICIAL PART OF THE LANDSCAPE

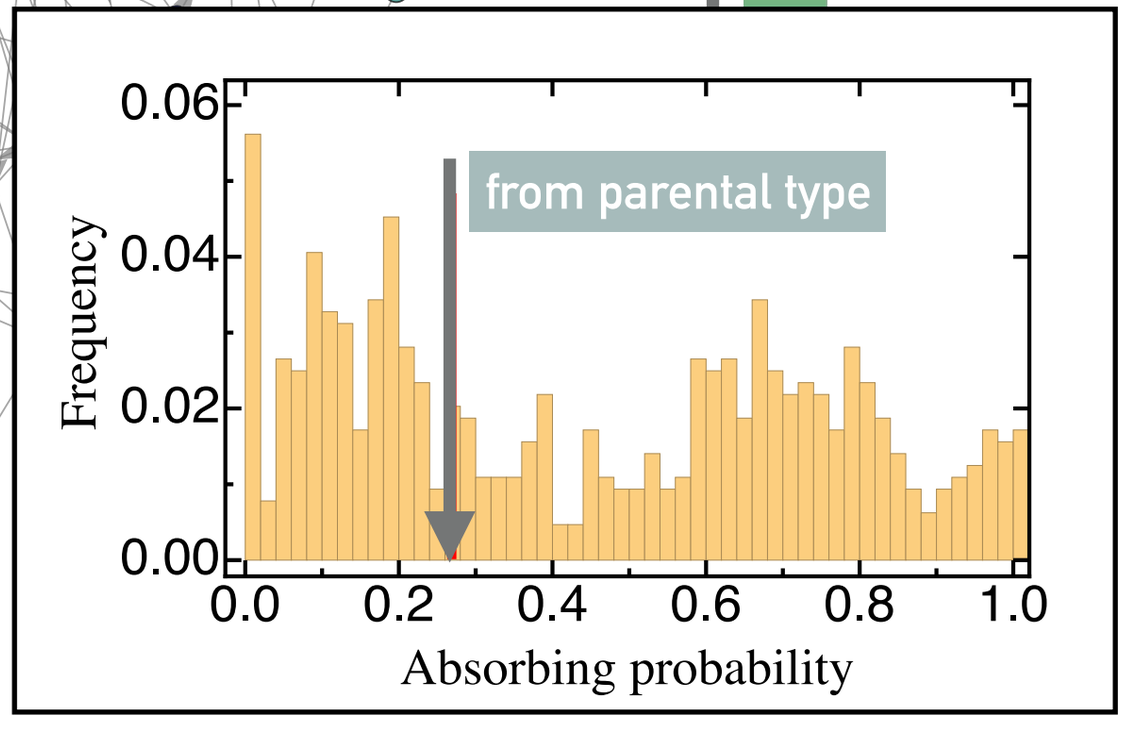


BENEFICIAL PART OF THE LANDSCAPE

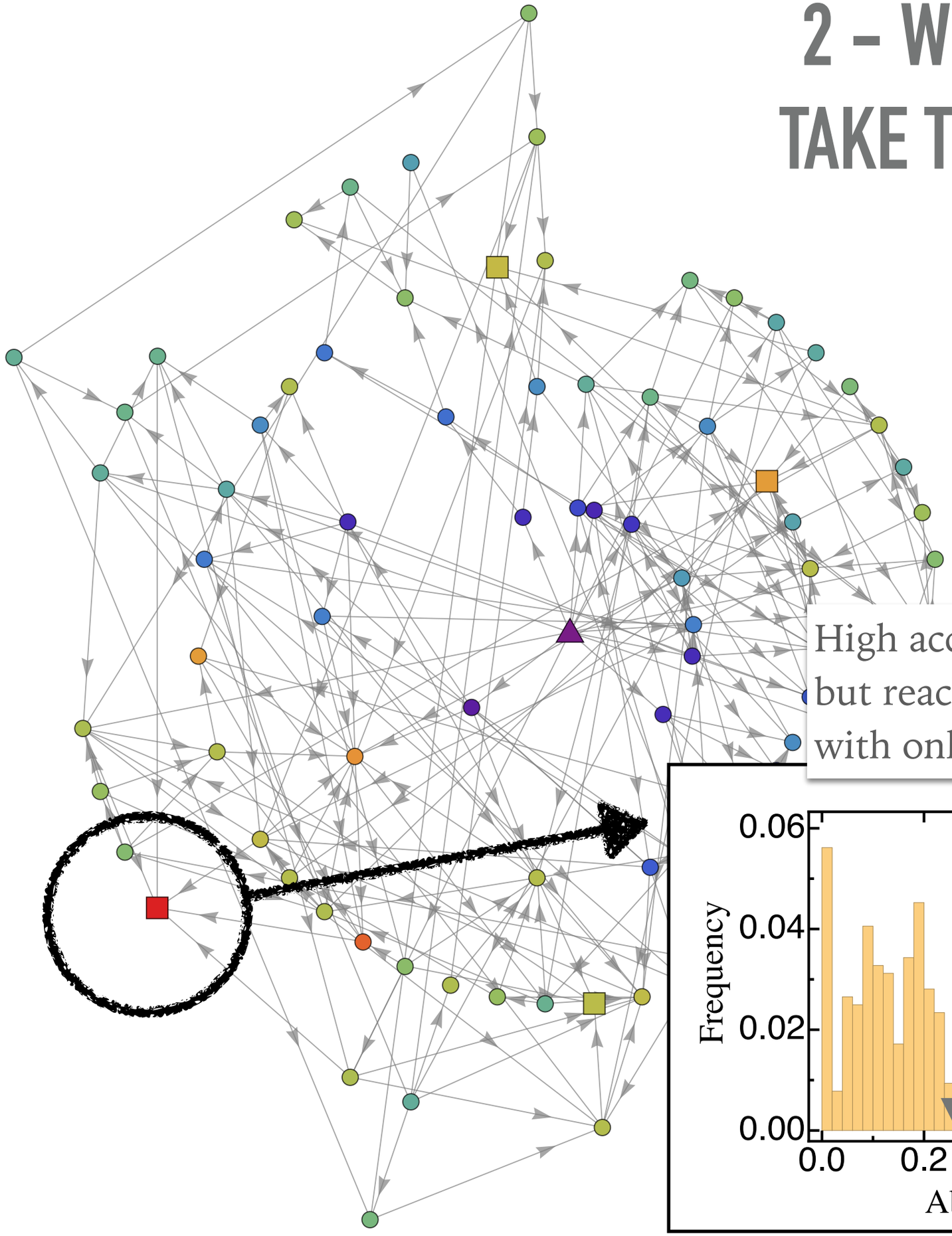


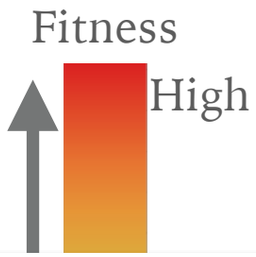
2 – WILL ADAPTATION TAKE THE POPULATION TO THE GLOBAL OPTIMUM?



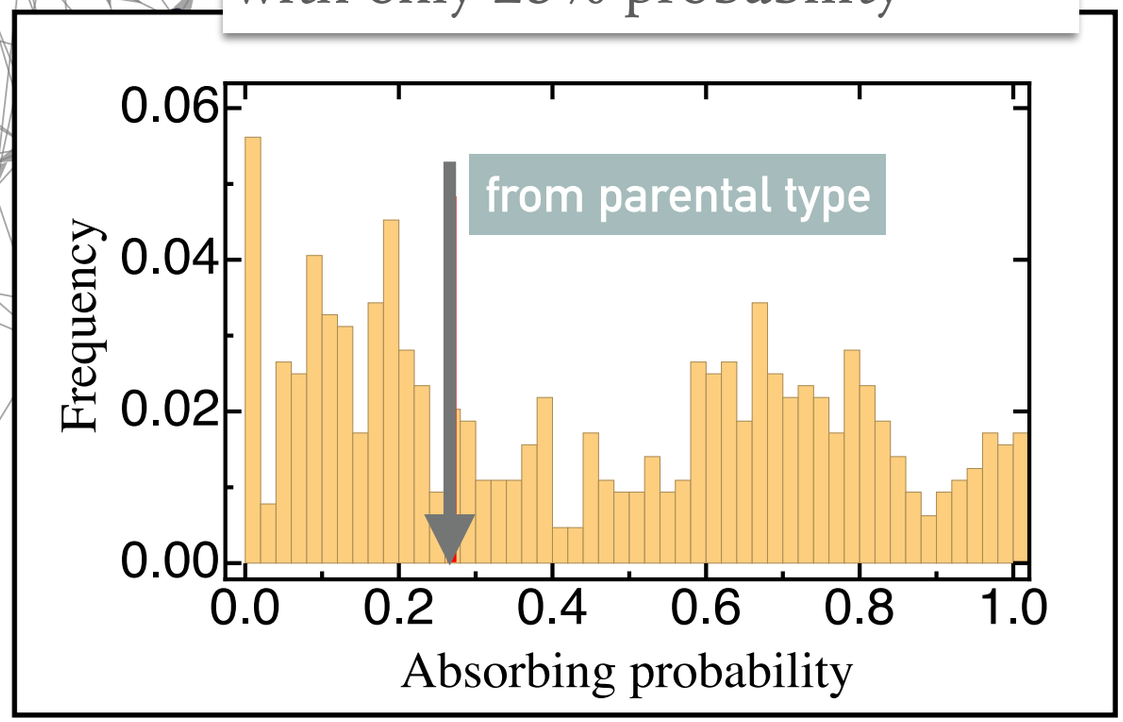


2 – WILL ADAPTATION TAKE THE POPULATION TO THE GLOBAL OPTIMUM?

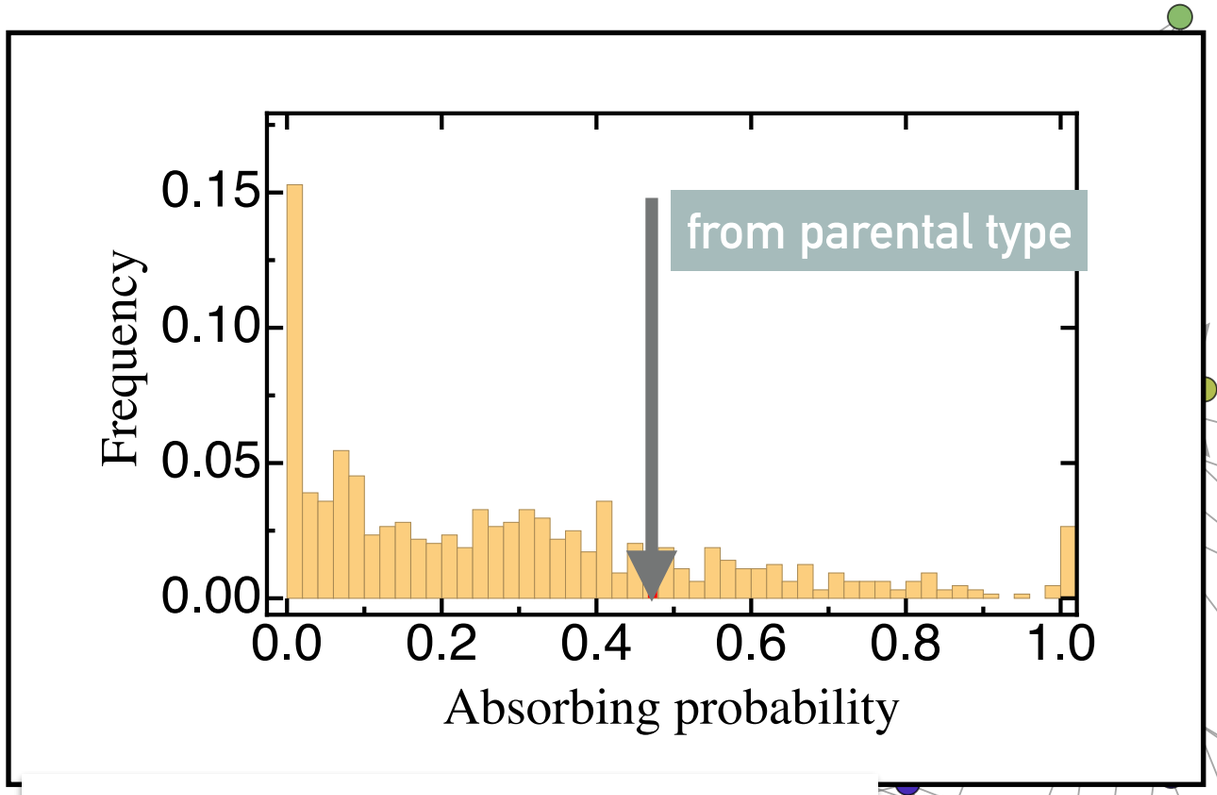




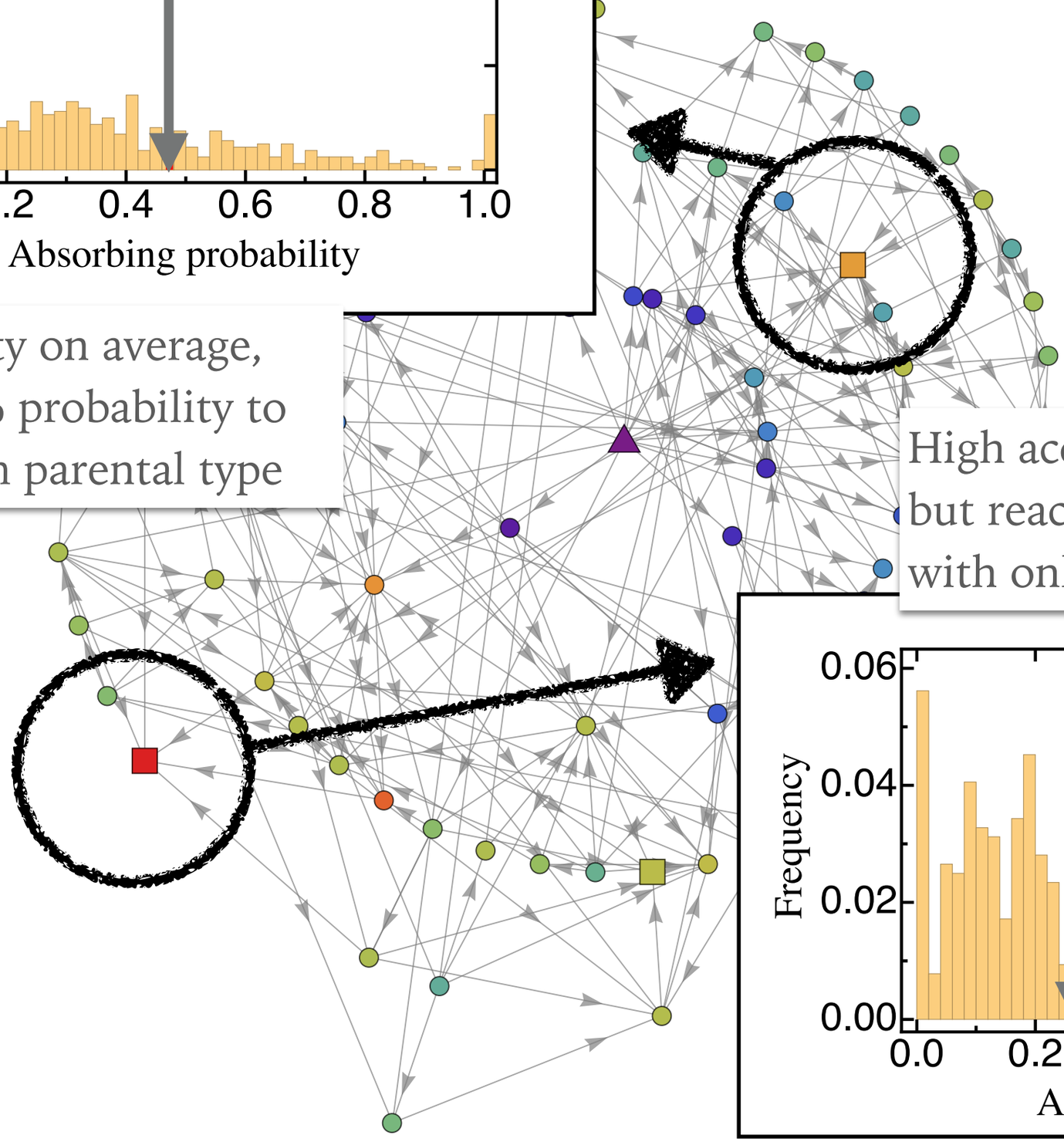
High accessibility on average, but reached from parental type with only 25% probability



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

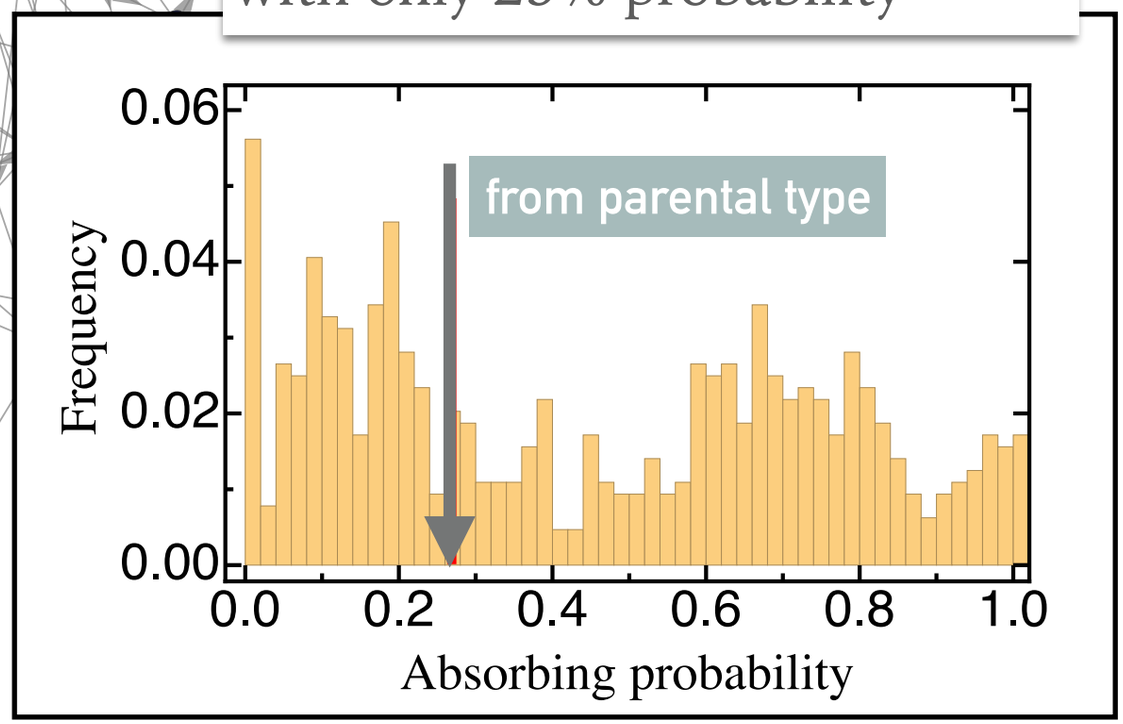


Low accessibility on average, but almost 50% probability to be reached from parental type

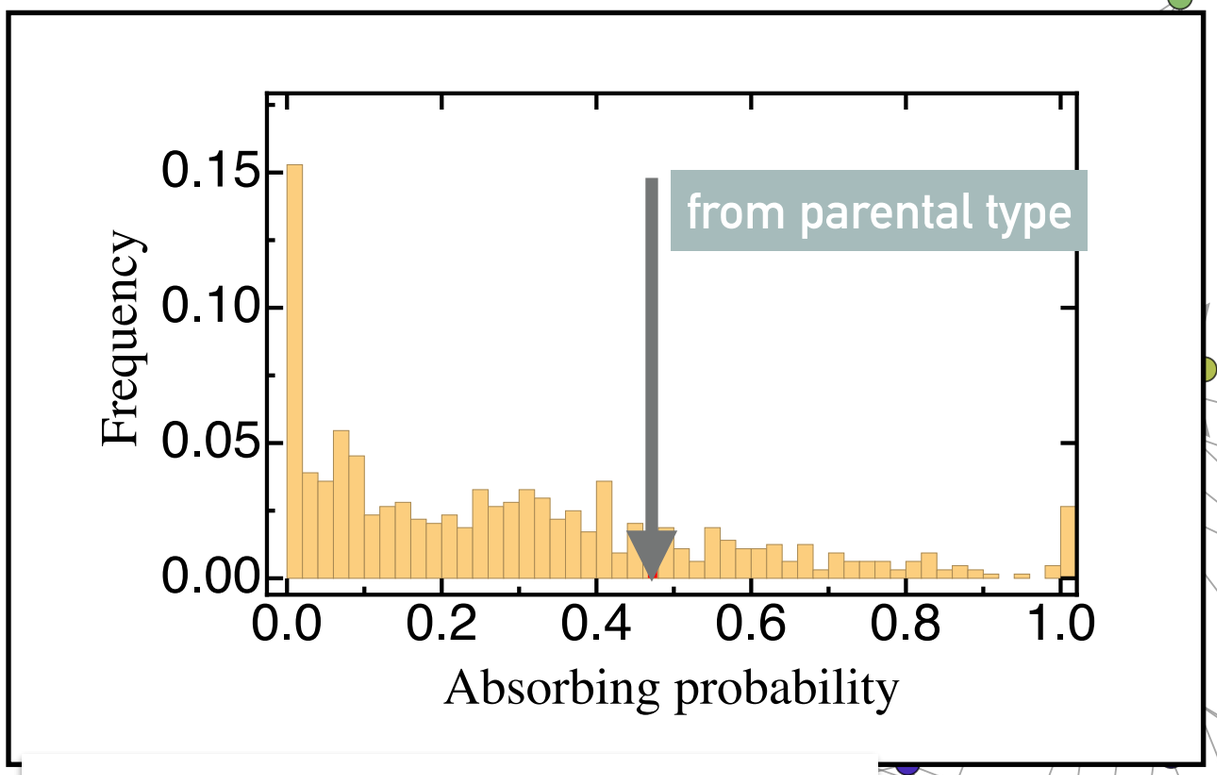


Fitness
High

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



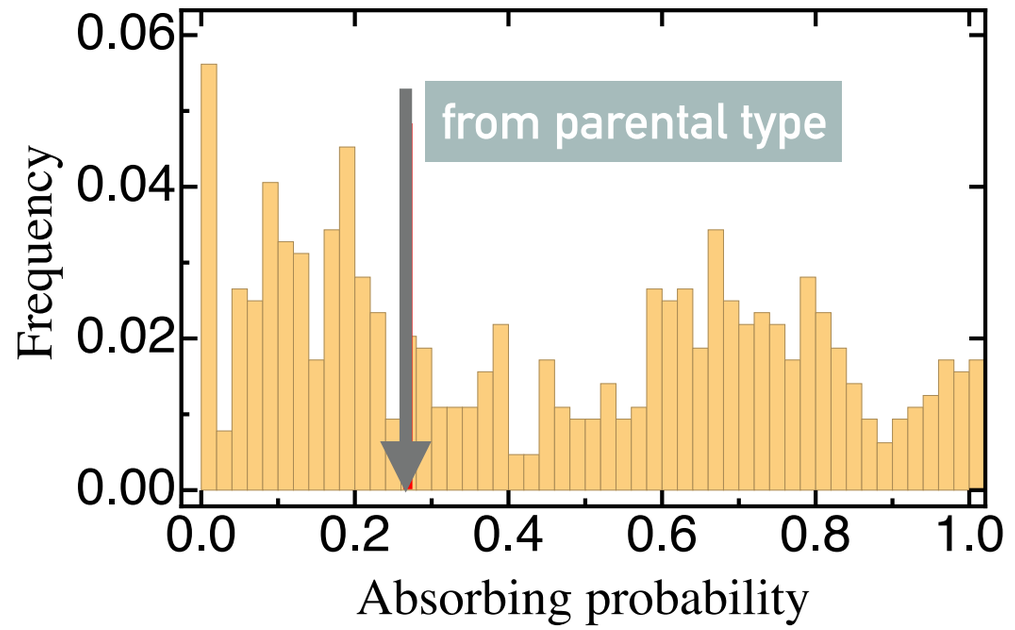
2 – WILL ADAPTATION TAKE THE POPULATION TO THE GLOBAL OPTIMUM?



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

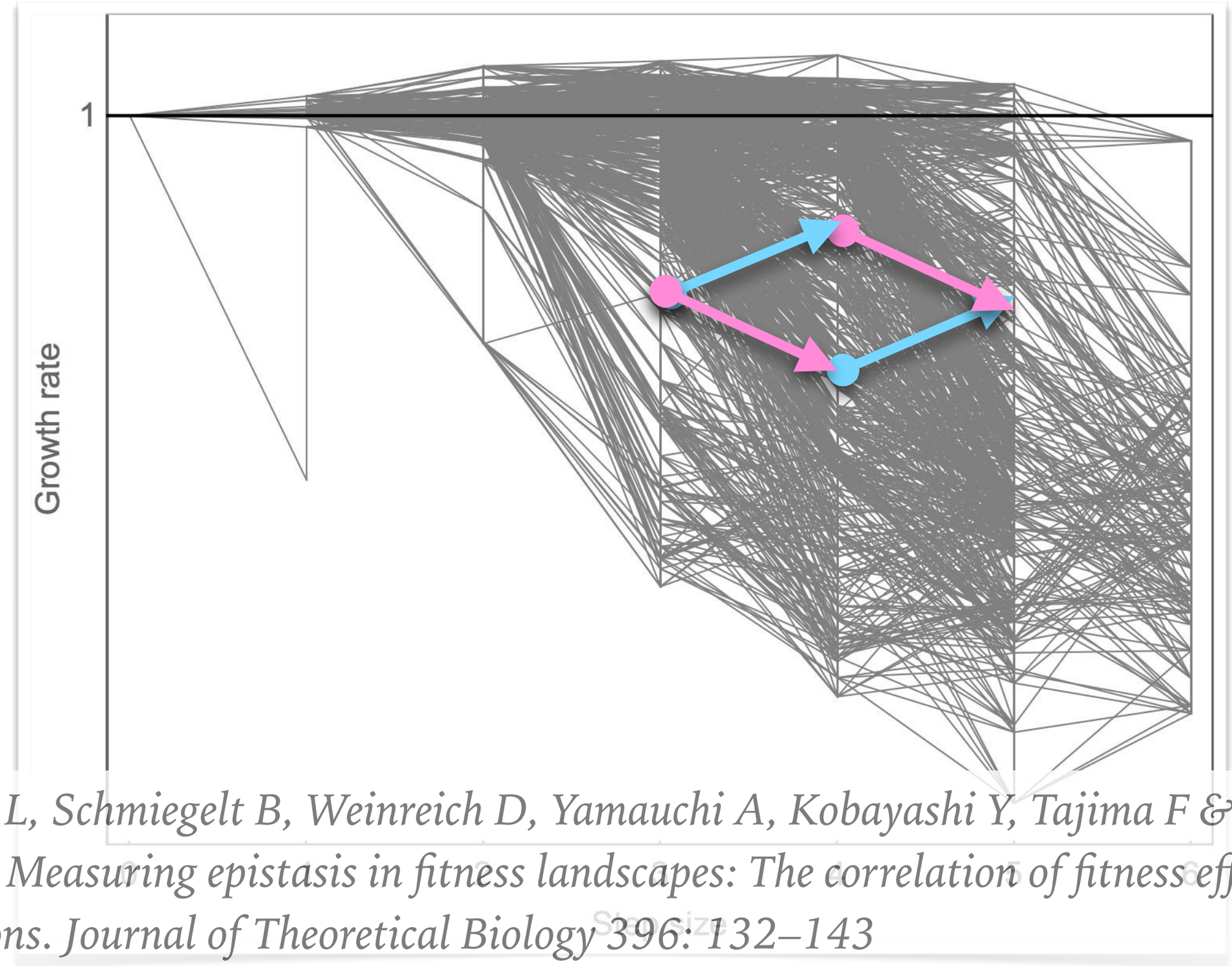
Fitness
High

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



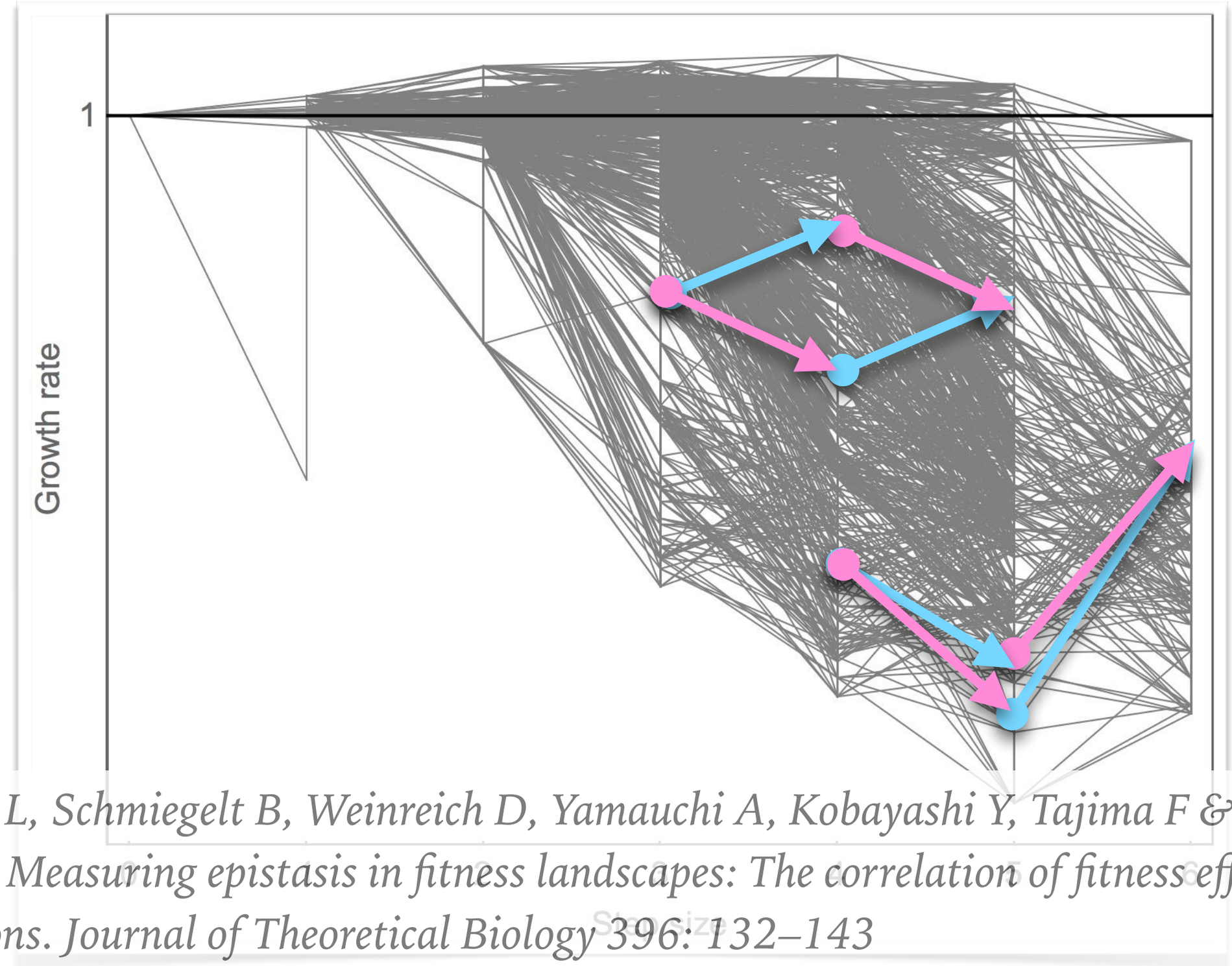
Evolution from parental type may stall at local optimum

LANDSCAPE STATISTICS INDEPENDENT OF REFERENCE



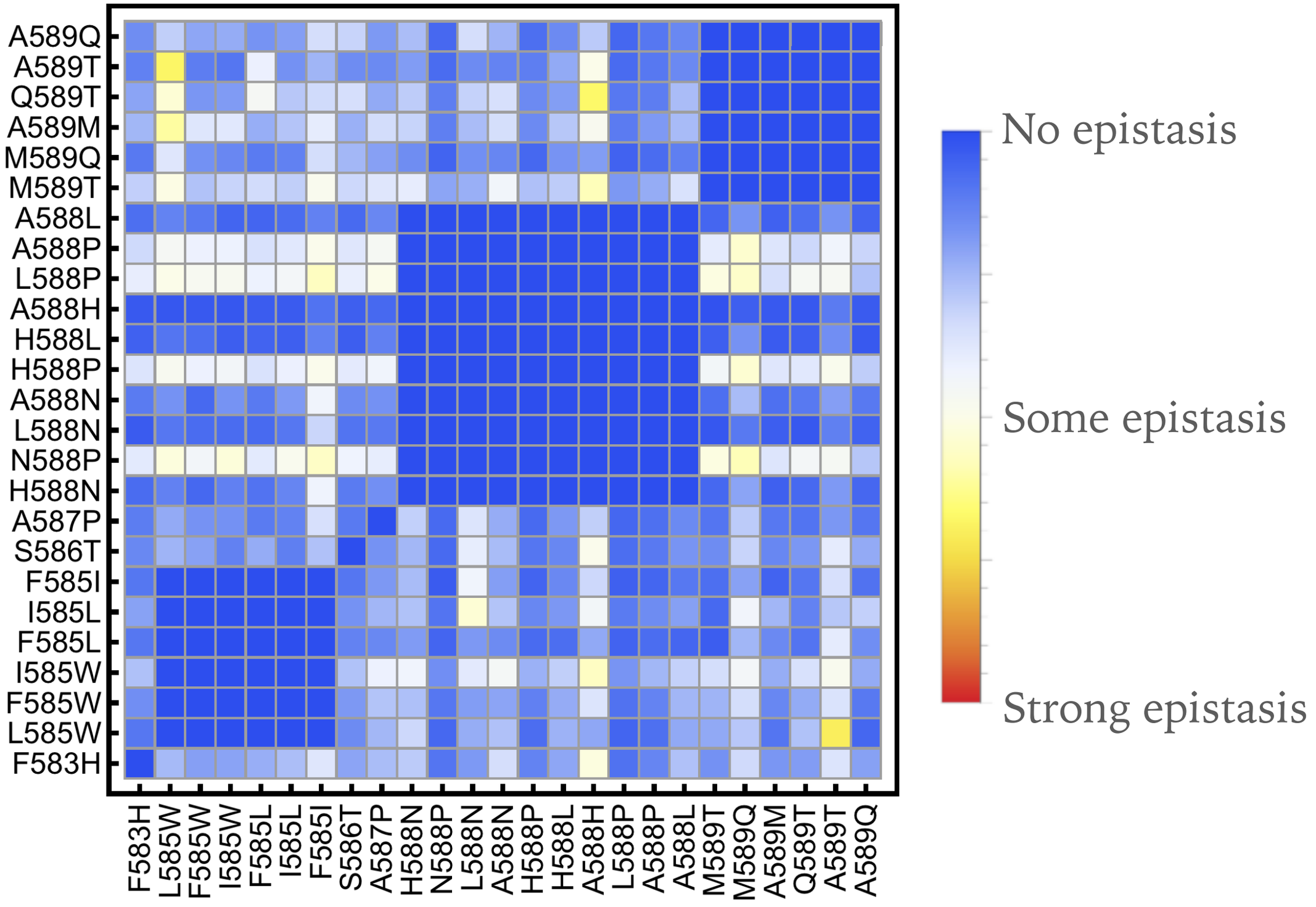
Ferretti L, Schmiegelt B, Weinreich D, Yamauchi A, Kobayashi Y, Tajima F & Achaz G (2016) Measuring epistasis in fitness landscapes: The correlation of fitness effects of mutations. Journal of Theoretical Biology 396: 132–143

LANDSCAPE STATISTICS INDEPENDENT OF REFERENCE

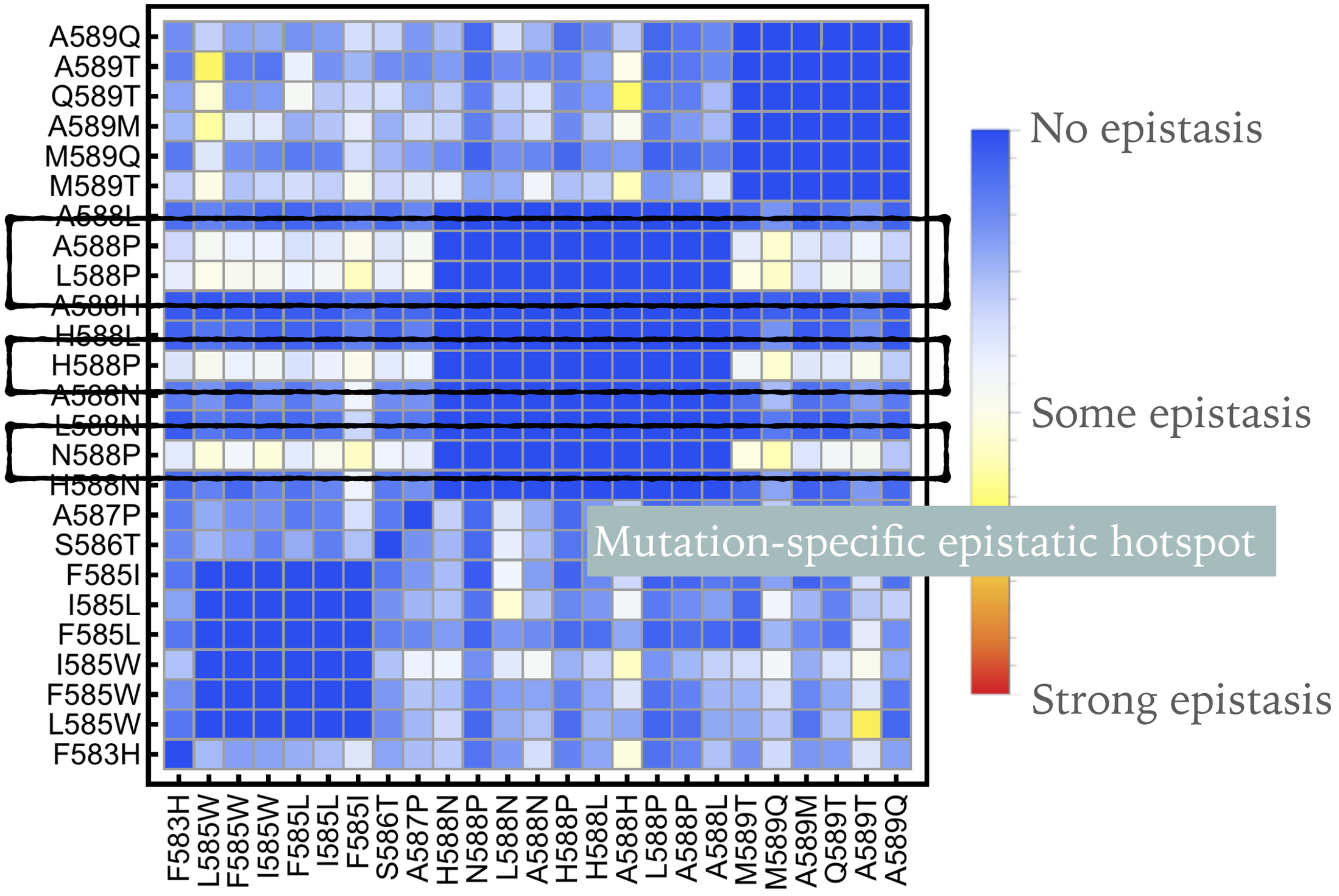


Ferretti L, Schmiegelt B, Weinreich D, Yamauchi A, Kobayashi Y, Tajima F & Achaz G (2016) Measuring epistasis in fitness landscapes: The correlation of fitness effects of mutations. Journal of Theoretical Biology 396: 132–143

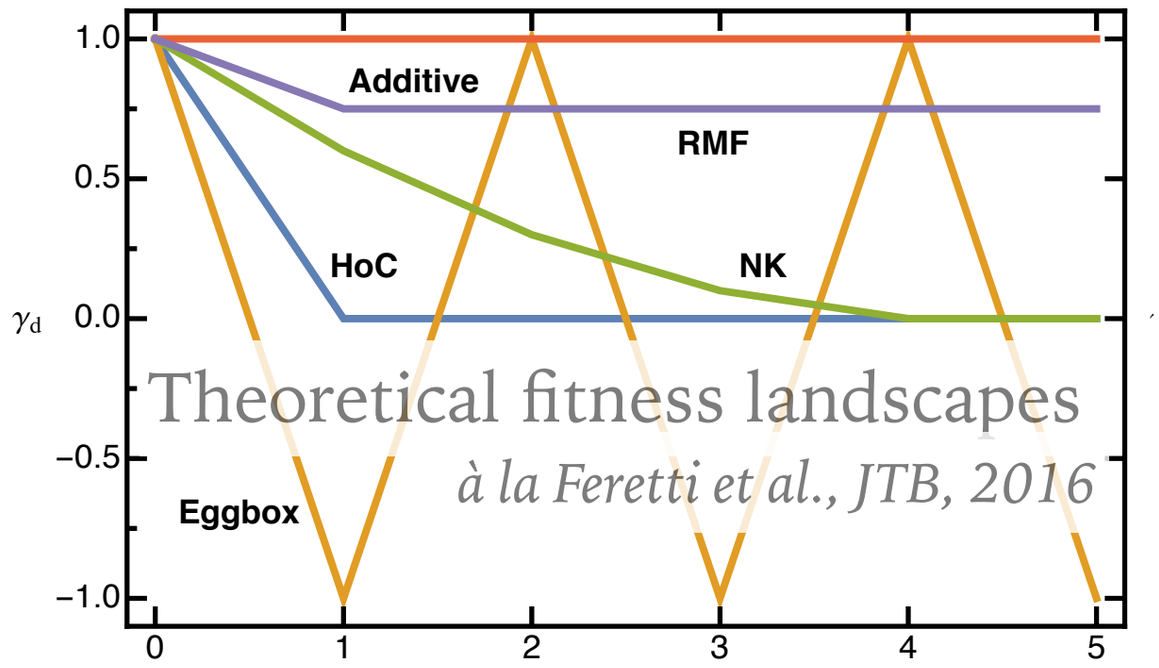
HOW CAN WE MEASURE FITNESS LANDSCAPES AND WHAT CAN WE LEARN FROM THIS EXERCISE?



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

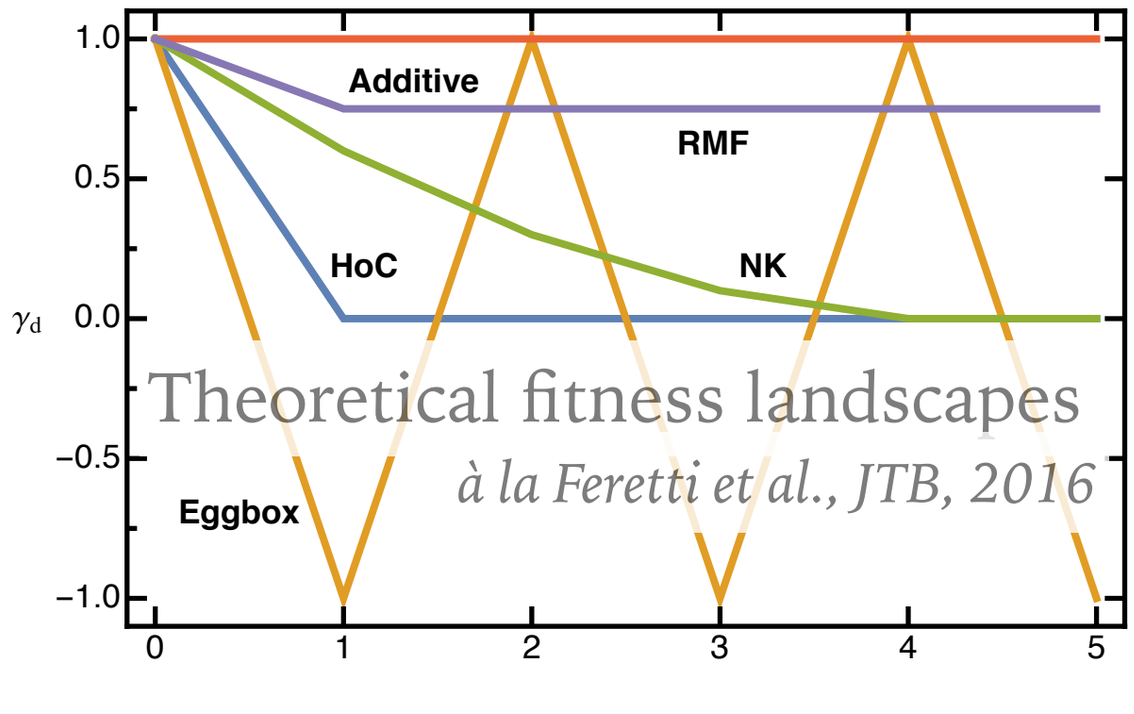


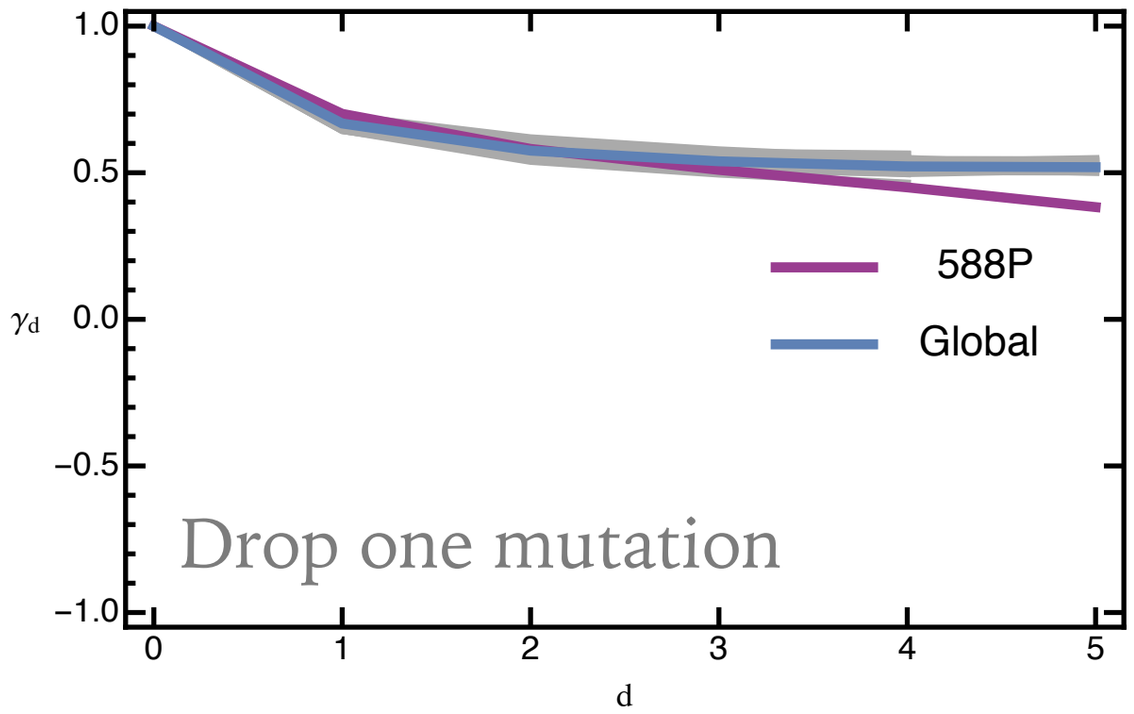
3 – CAN WE INFER AN UNKNOWN PART OF THE LANDSCAPE?



Epistasis measure $gamma_d$ with increasing Hamming distance d

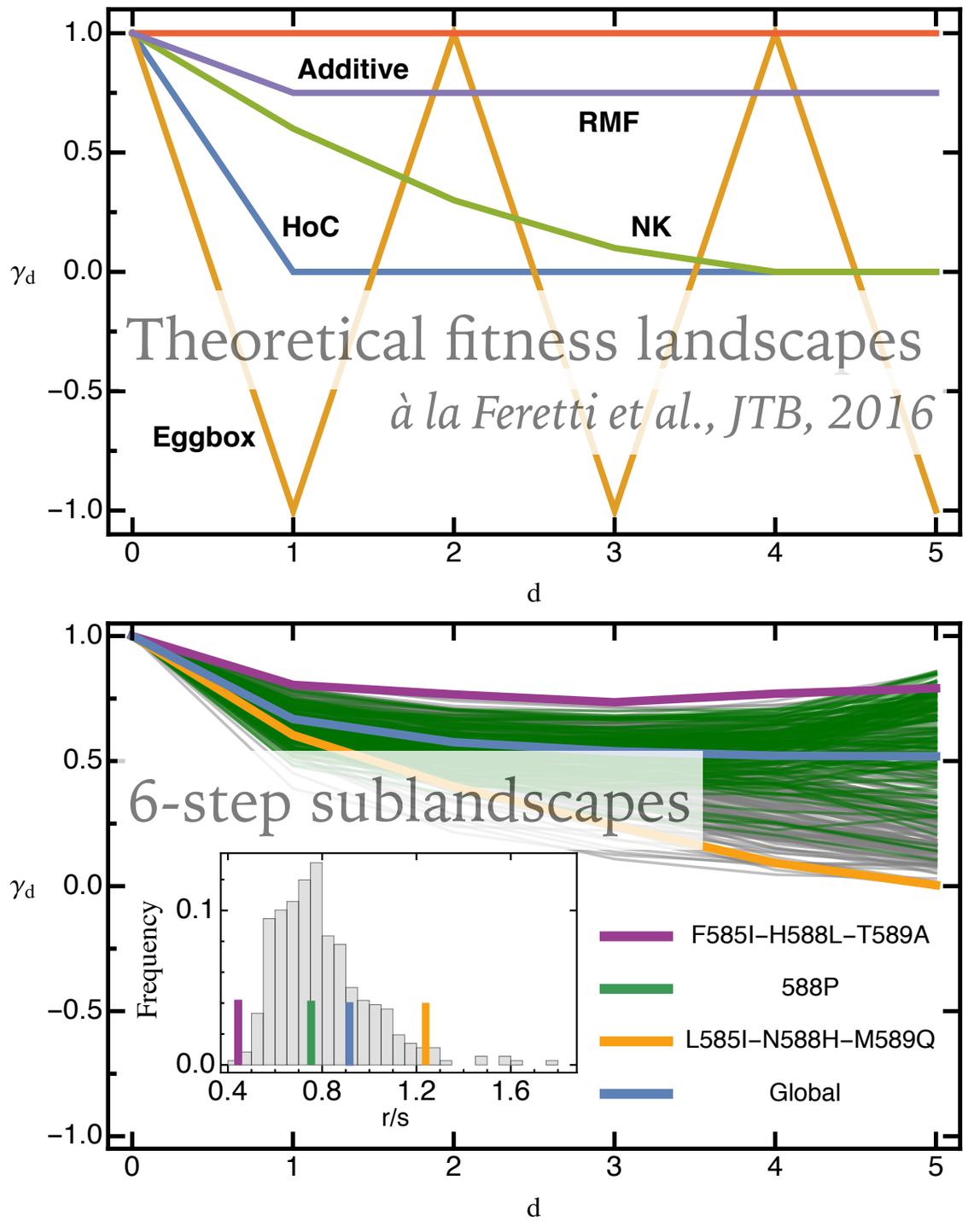
3 – CAN WE INFER AN UNKNOWN PART OF THE LANDSCAPE?

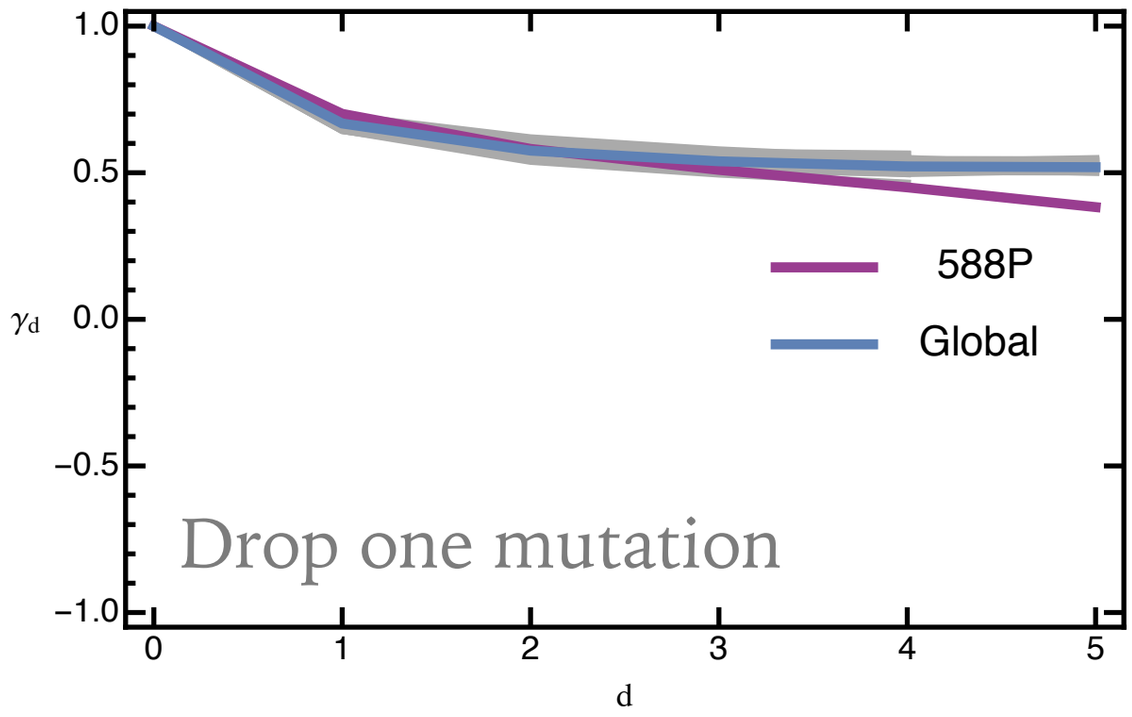




Epistasis measure gamma_d with increasing Hamming distance d

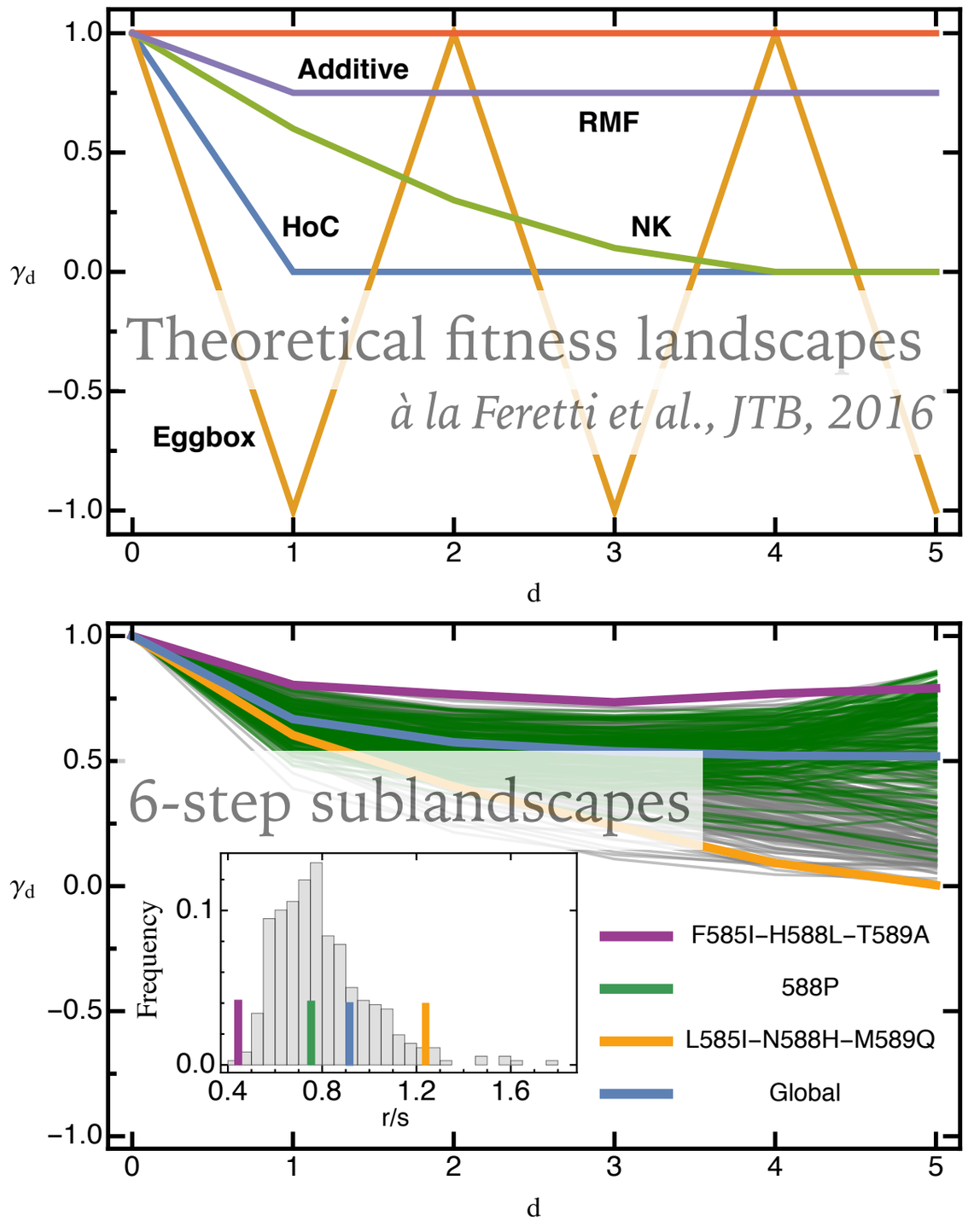
3 – CAN WE INFER AN UNKNOWN PART OF THE LANDSCAPE?

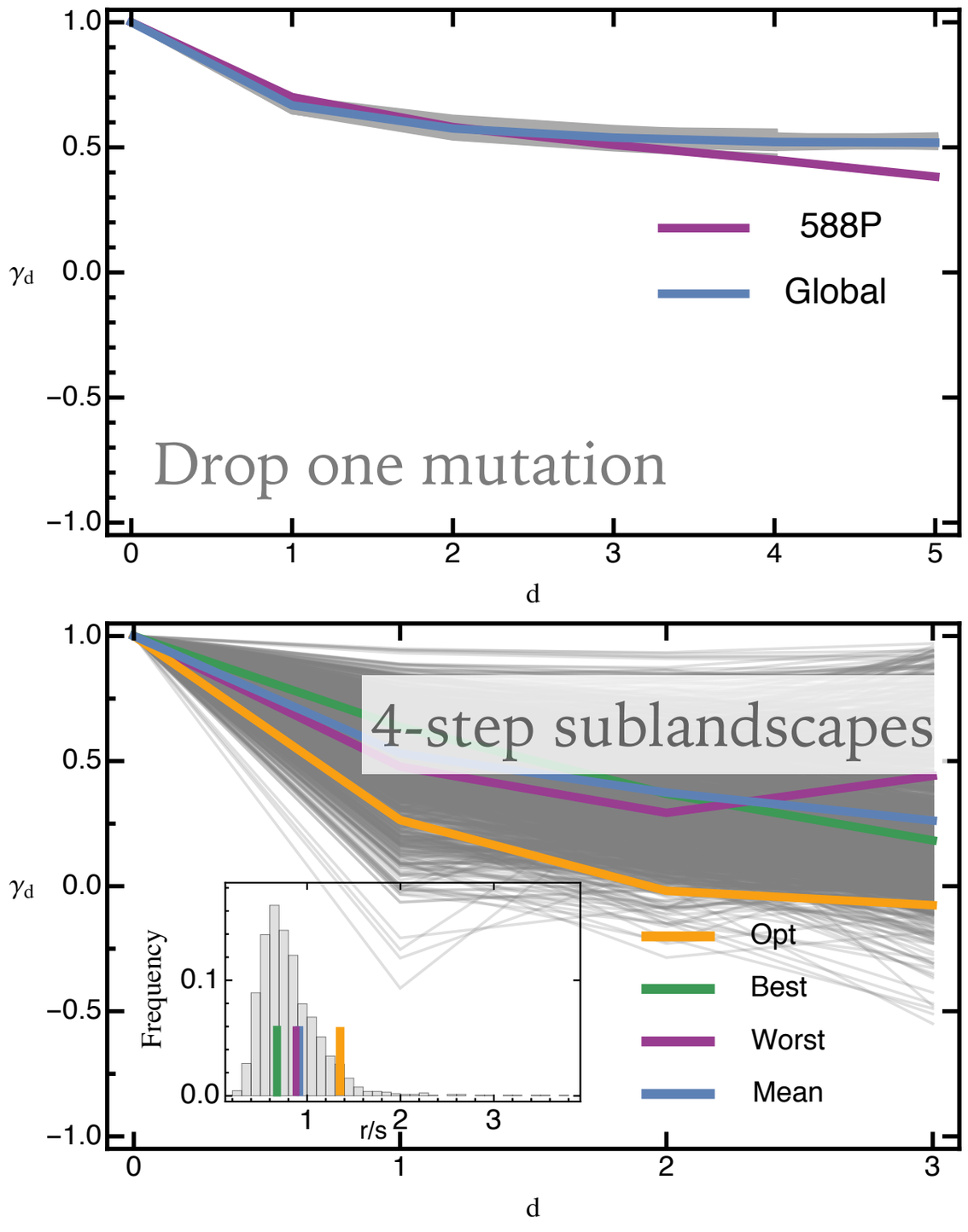




Epistasis measure $gamma_d$ with increasing Hamming distance d

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





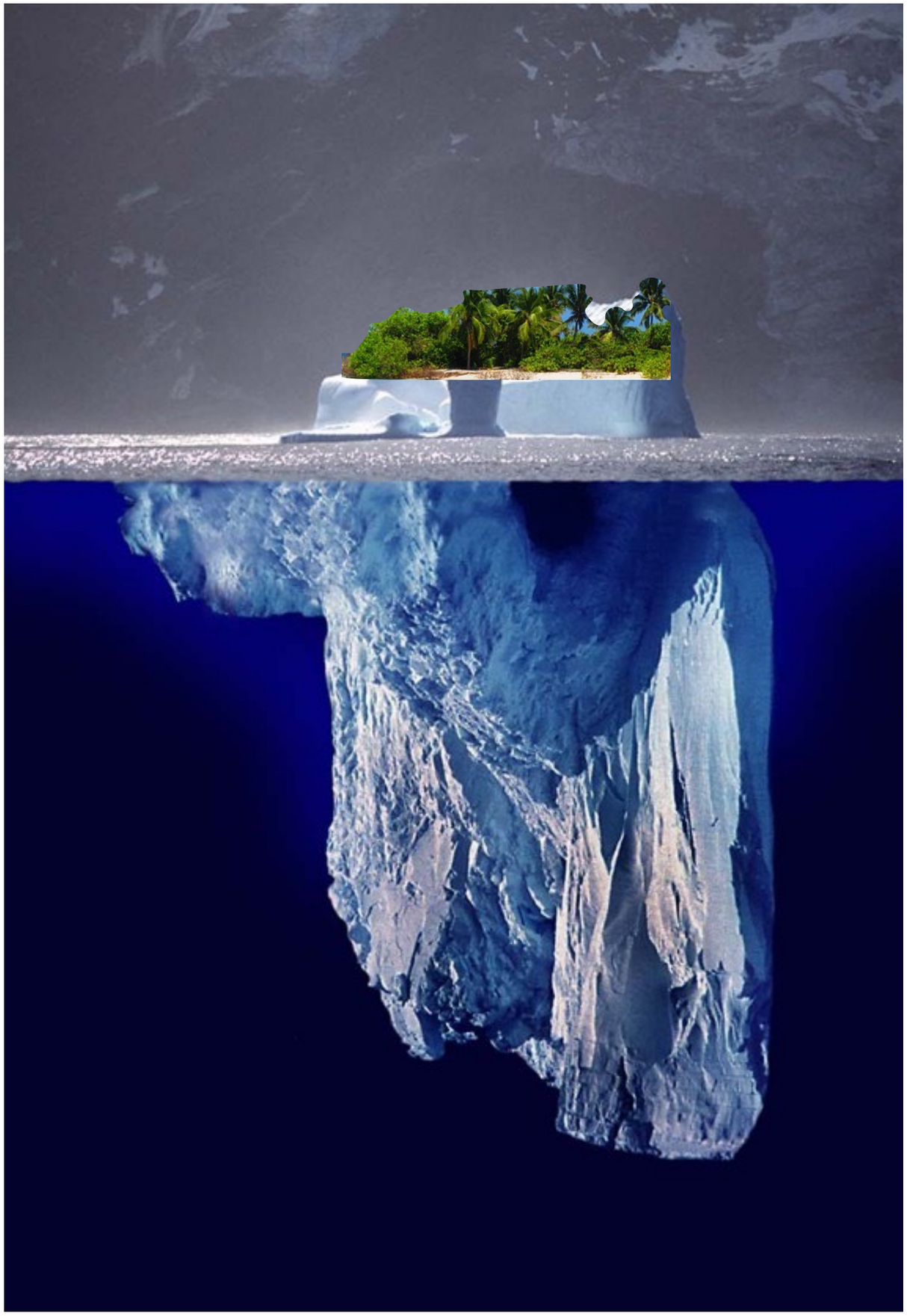
Epistasis measure gamma_d with increasing Hamming distance d

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.



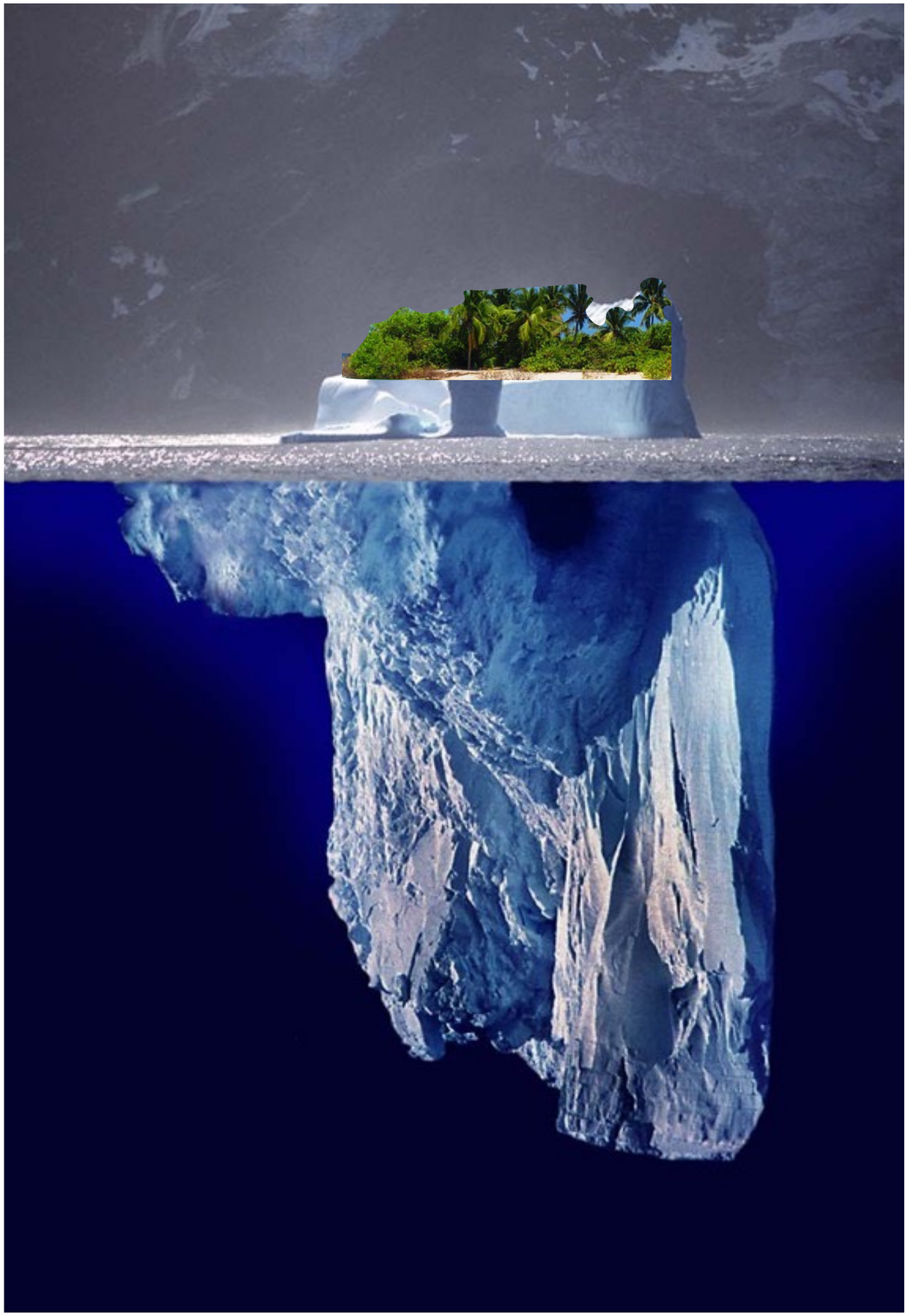
SUMMARY/CONCLUSION

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…

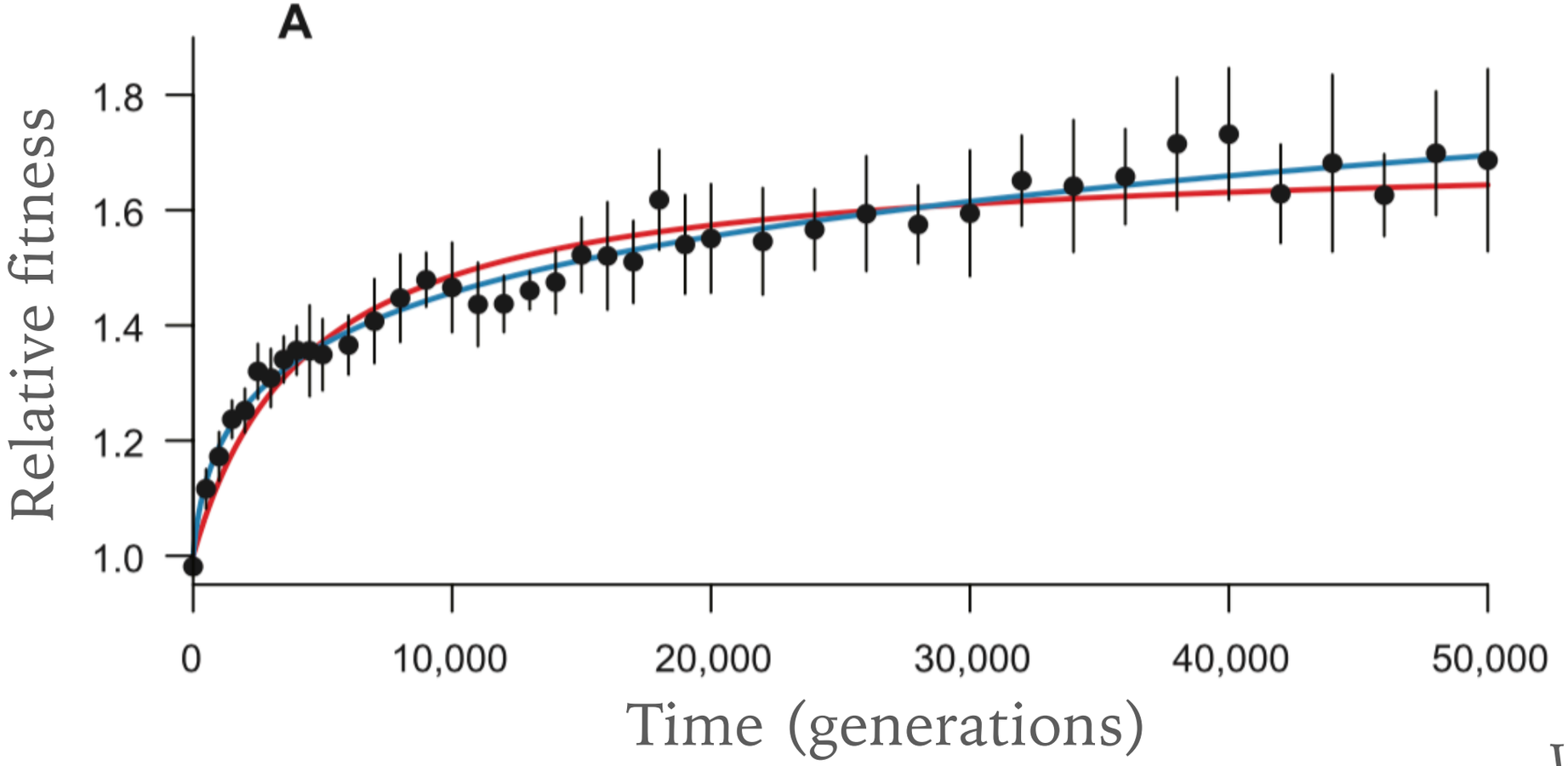


FISHER OR WRIGHT OR…?

- From an ecological point of view, frequent bottlenecks seem likely.

FISHER OR WRIGHT OR…?

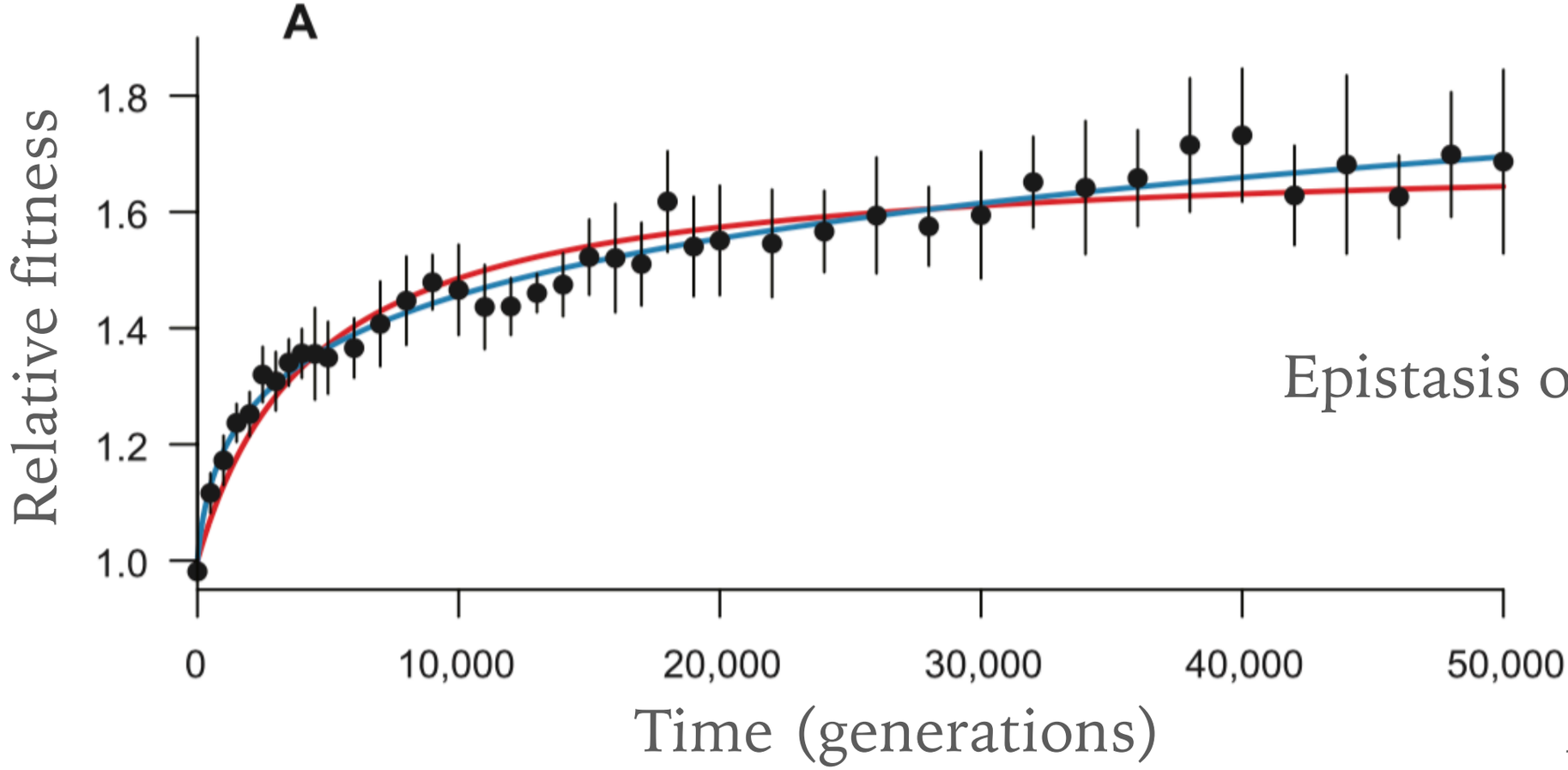
- From an ecological point of view, frequent bottlenecks seem likely.
- But adaptation is also miraculous in constant environments with high population sizes - how is that possible?



Wiser et al. 2013

FISHER OR WRIGHT OR…?

- From an ecological point of view, frequent bottlenecks seem likely.
- But adaptation is also miraculous in constant environments with high population sizes - how is that possible?



Epistasis or ecology?

Wiser et al. 2013

What would YOU like to know about fitness landscapes?

HTTPS://MICROBIAL.ECO-EVO.SCIENCE/

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

jpiamr
JOINT PROGRAMMING INITIATIVE ON ANTIMICROBIAL RESISTANCE