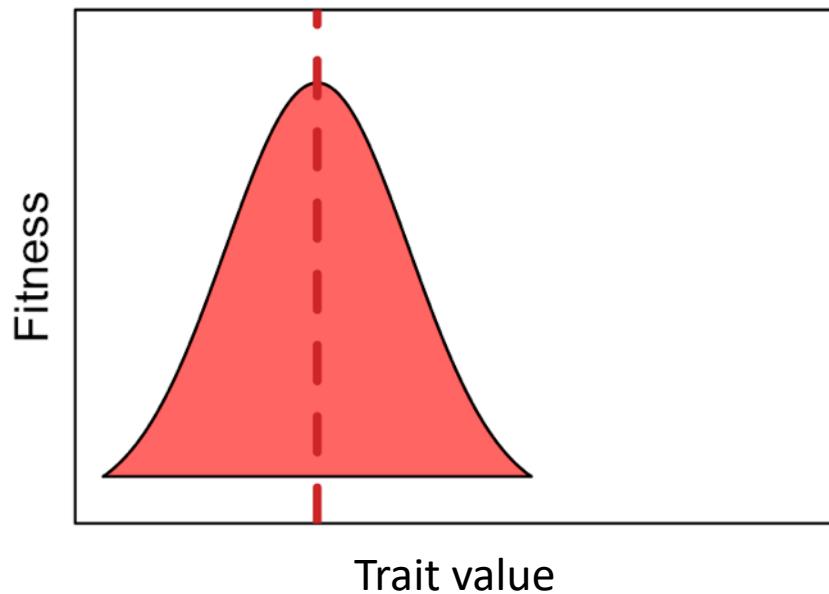
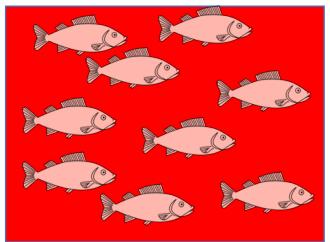
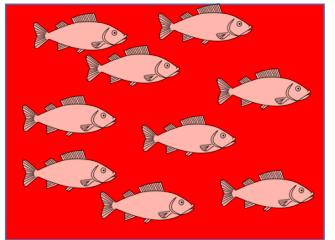


Selection

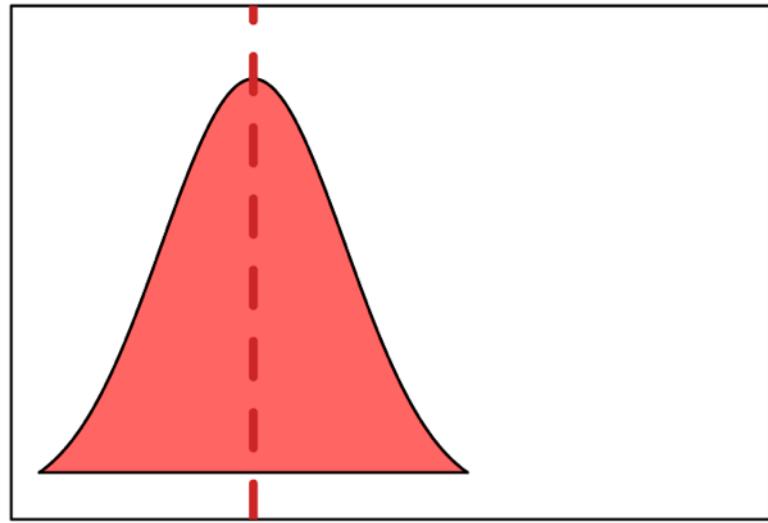
Week 5

- This week → selection at a single locus
- Next week → Effects of selection across the genome

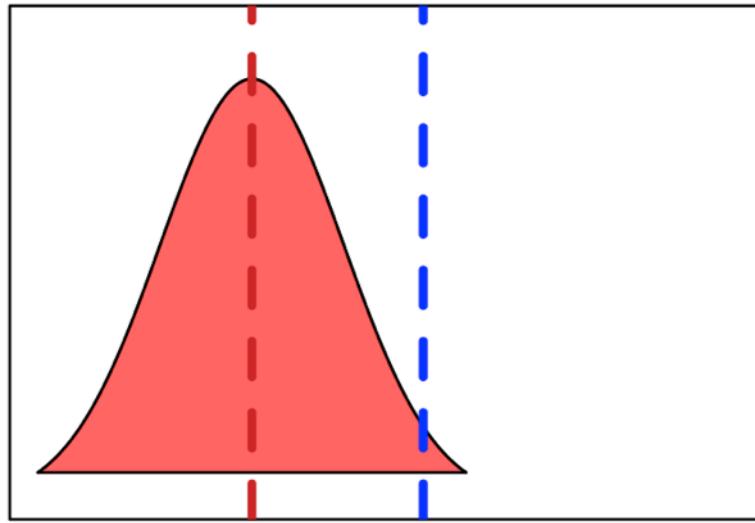
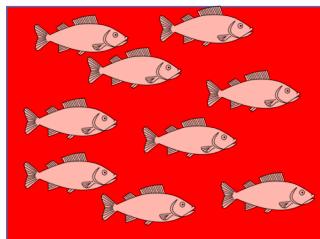




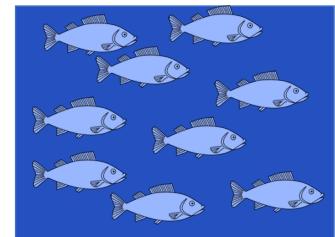
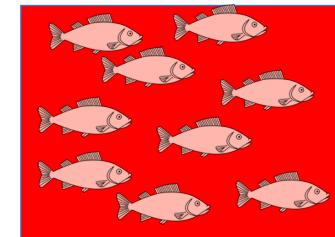
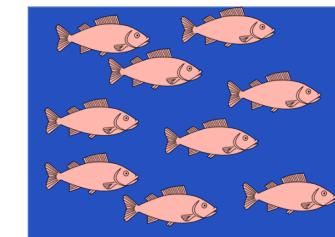
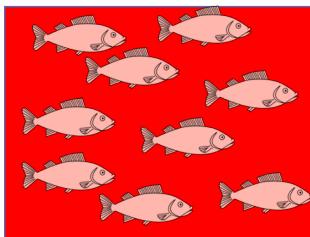
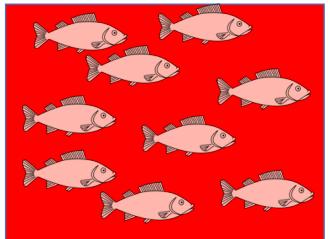
Fitness



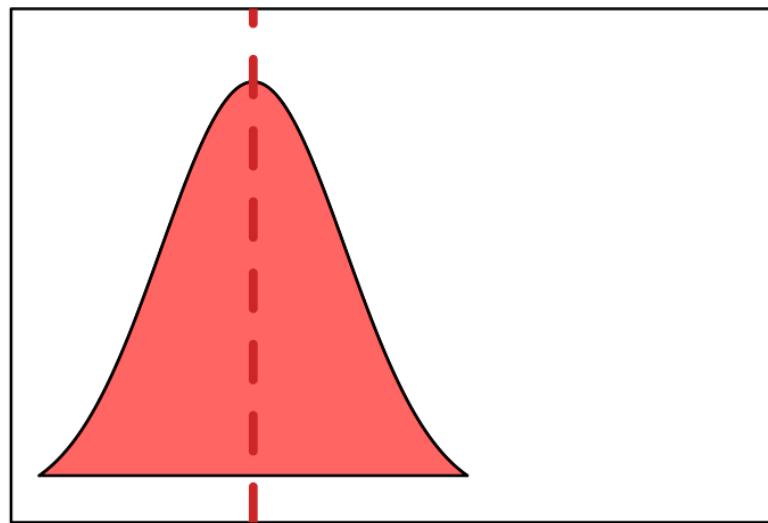
Trait value



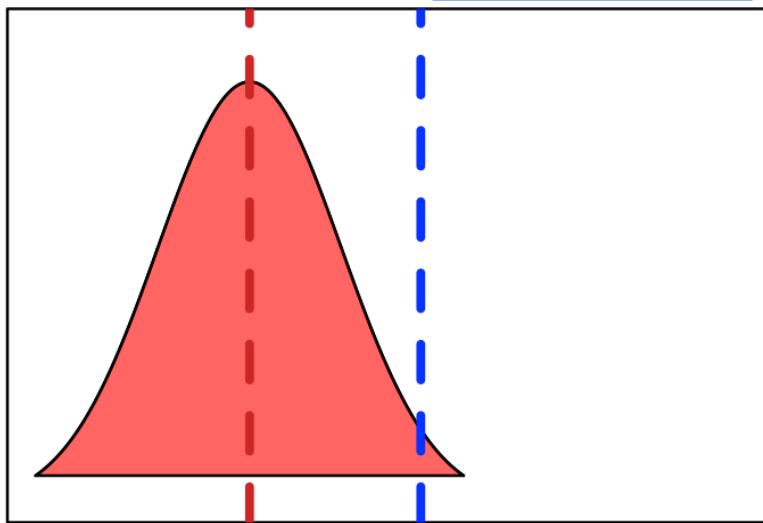
Trait value



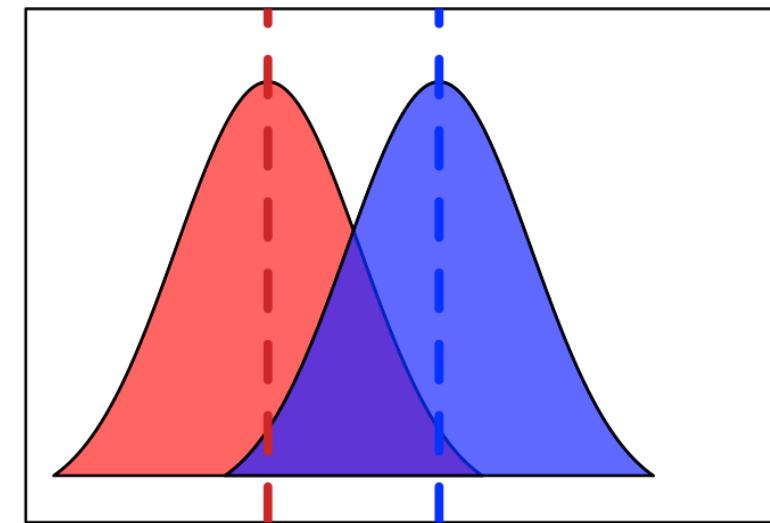
Fitness



Trait value



Trait value



Trait value

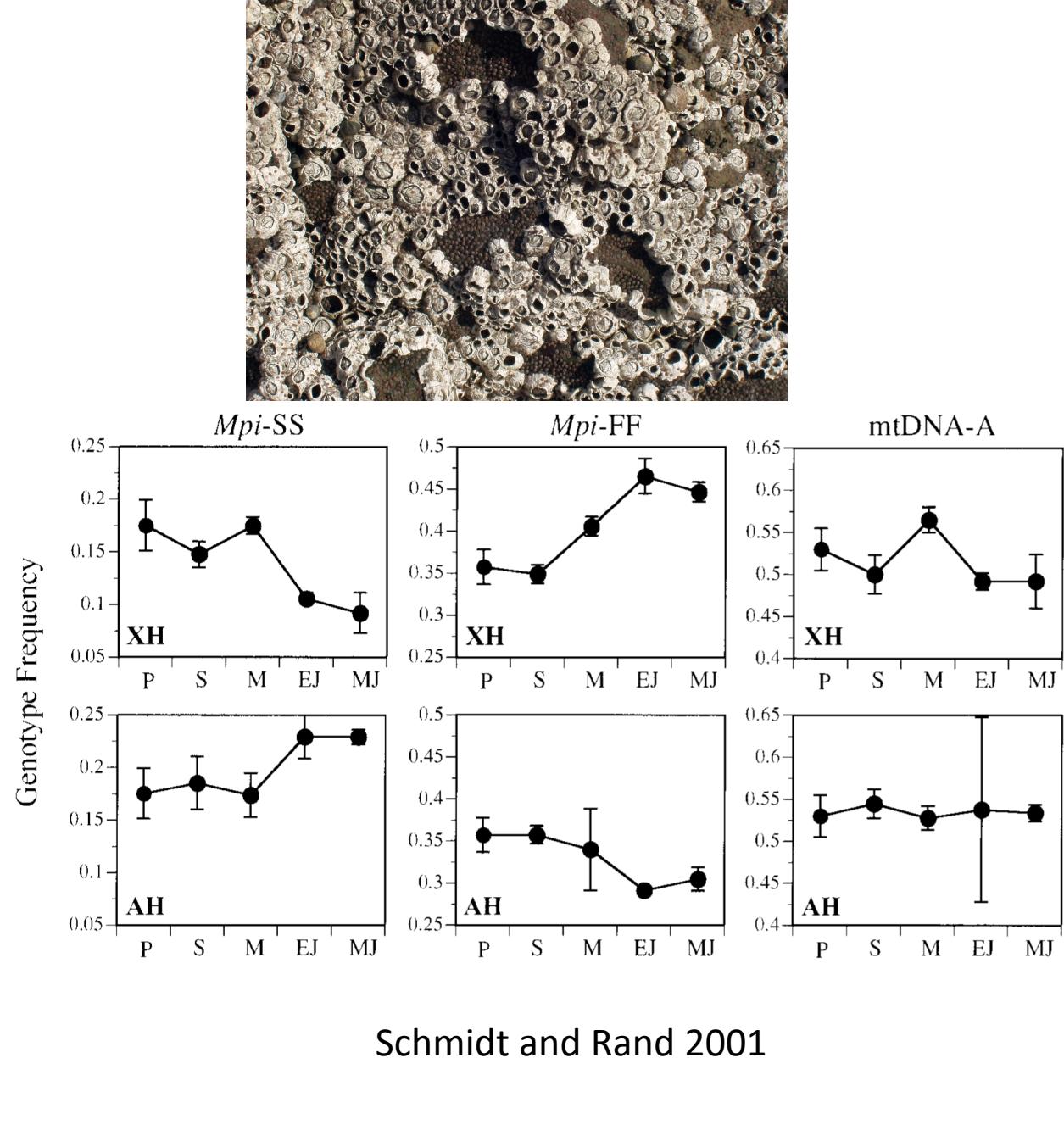
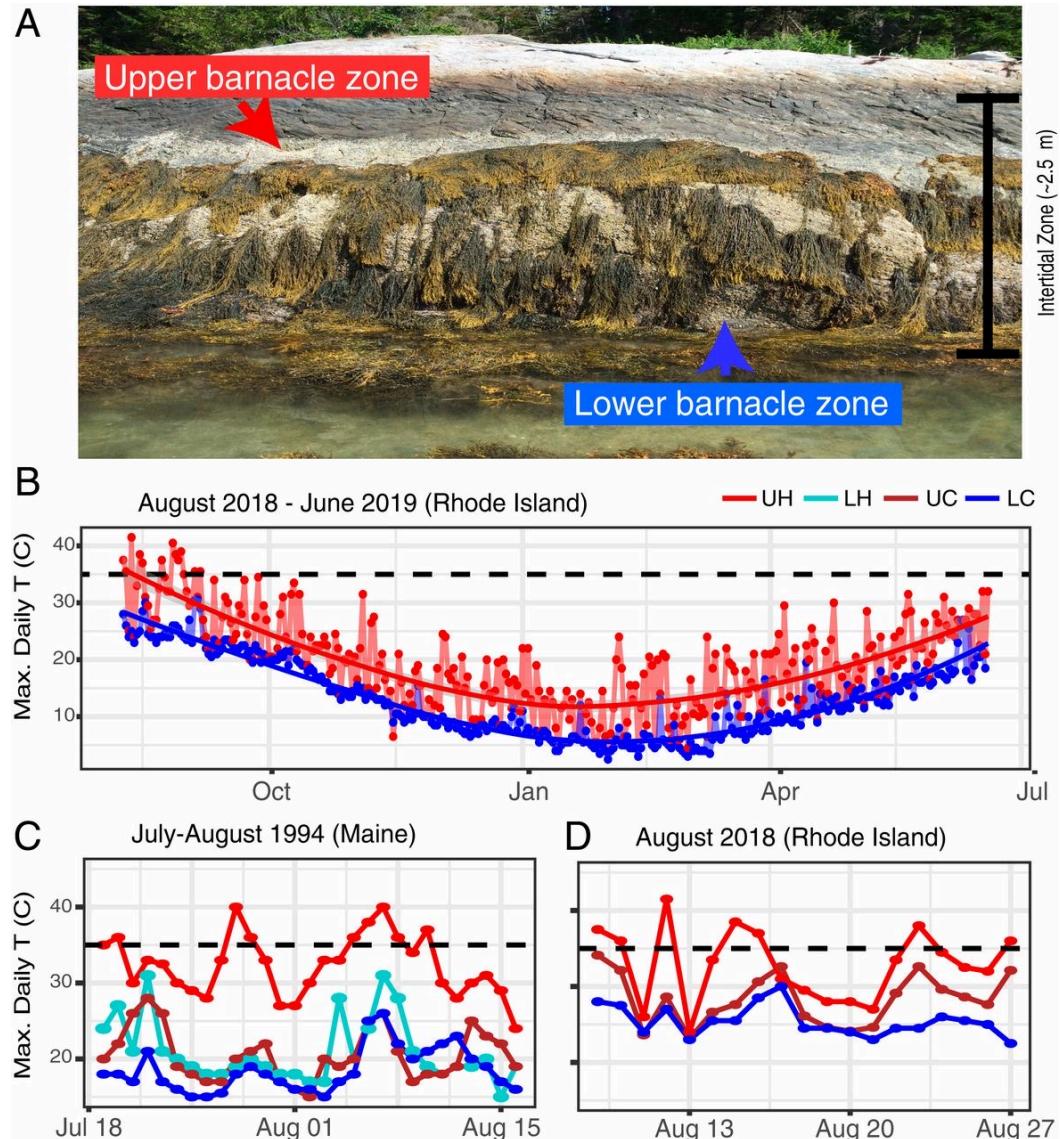
Adaptation

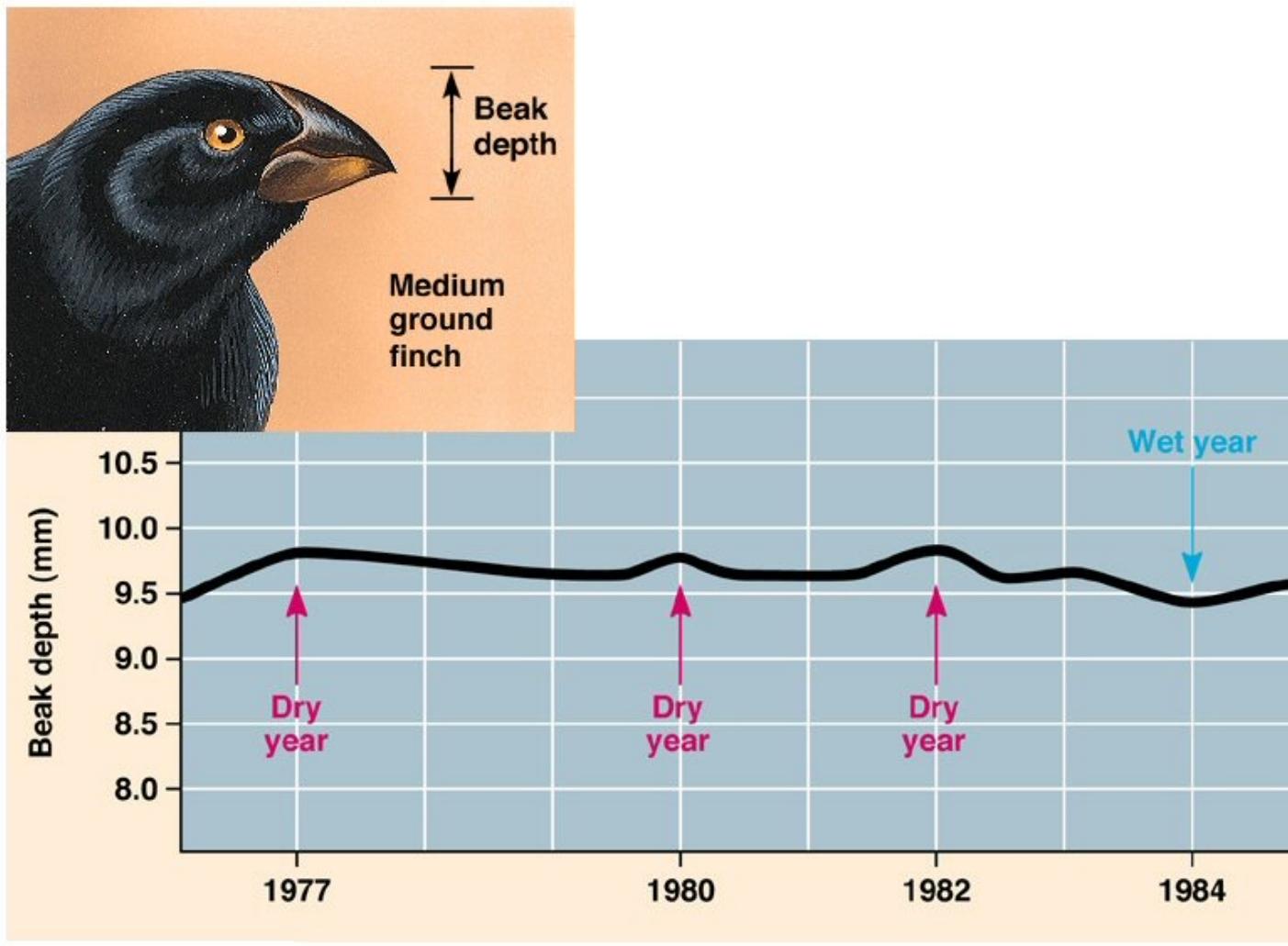
- Distinct environments
- Heritable variation
- Differential reproductive success
- Increase in population fitness across time
- At the population level, change in allele frequencies

Selection

- Differential survival and reproduction among genotypes
- Goes back to Darwin (1859)

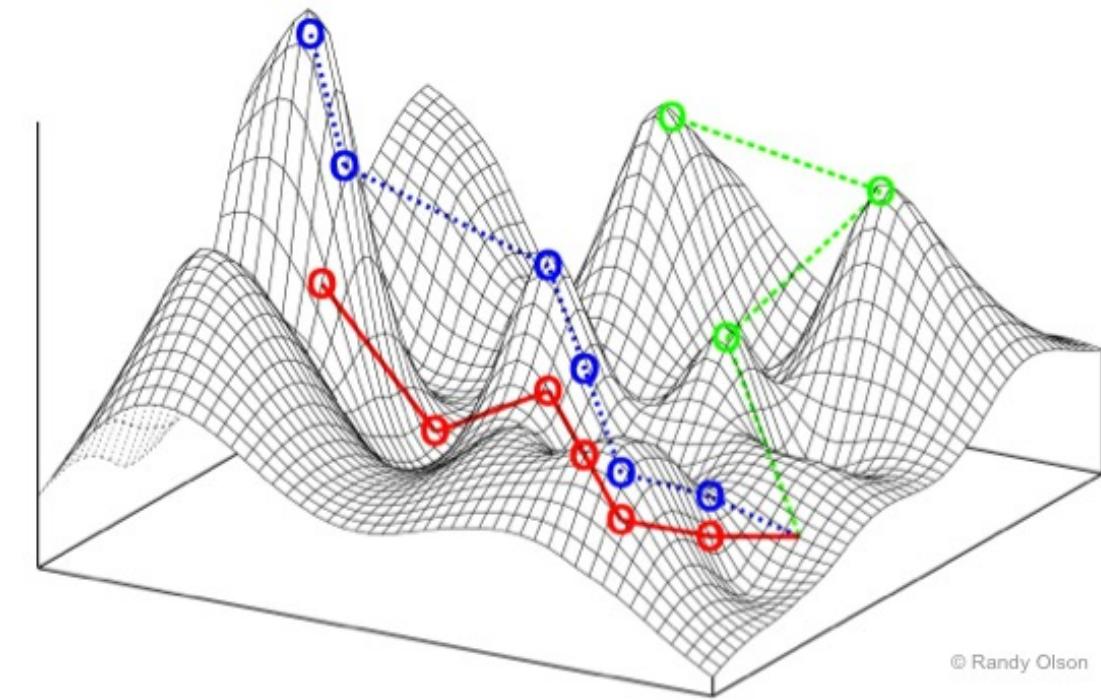
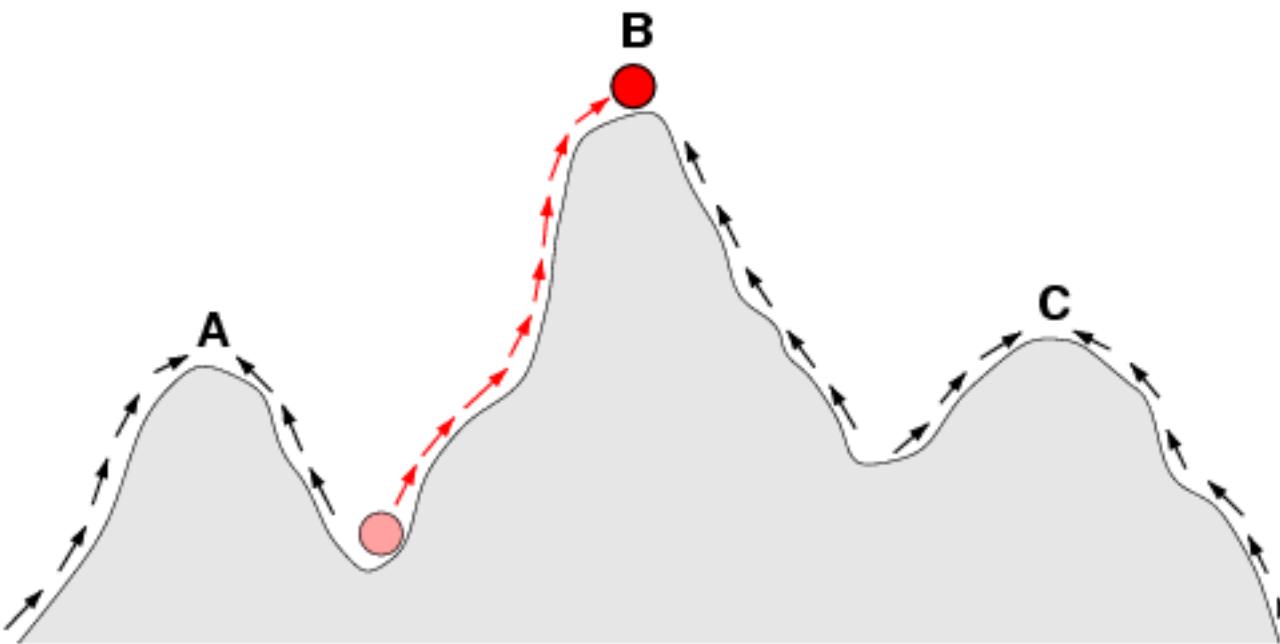
Selection on MPI in *Semibalanus balanoides*





©1999 Addison Wesley Longman, Inc.

Fitness landscapes



© Randy Olson

Local adaptation

- Conceptual issues in local adaptation, Kawecki and Ebert 2004
 - “*divergent selection* should cause each local population (deme) to evolve traits that provide an advantage under its local environmental conditions (which we refer to as its habitat), regardless of the consequences of these traits for fitness in other habitats.”
 - “resident genotypes in each deme would have on average a higher relative fitness in their local habitat than genotypes originating from other habitats.”
- Local vs. foreign:
 - in each habitat the local deme is expected to show higher fitness than demes from other habitats
- Home vs. away:
 - deme had a higher fitness in its own habitat (at home) than in other habitats (away)

A = both

B = local vs. foreign

C = home vs. away

Mexican cave fish



Surface



Micos



Pachón-surface
F₂ hybrid



Pachón

1 cm

Repeated loss of eyes and coloration

Why?

But you need to be careful...

- Easy to assume most traits are adaptations... is this true?

Proc. R. Soc. Lond. B **205**, 581–598 (1979)

581

Printed in Great Britain

The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme

BY S. J. GOULD AND R. C. LEWONTIN

*Museum of Comparative Zoology, Harvard University,
Cambridge, Massachusetts 02138, U.S.A.*



Dr. Pangloss

- Character in Voltaire's *Candide*
- We live in the best of all possible worlds
- *"Things cannot be other than they are... everything is made for the best purpose. Our noses were made to carry spectacles, so we have spectacles. Legs were clearly intended for breeches, and we wear them."*

Telling stories involving natural selection

If a trait seems adaptive it must be so
“Just so” stories

Not all traits result directly from selection

- Genetic drift
- Selection on other traits

1959 → now

Selected for tameness

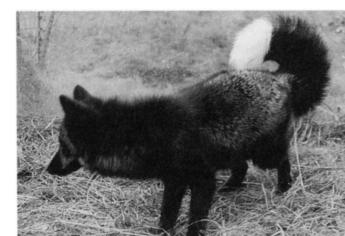
- Willingness to approach humans
- No biting



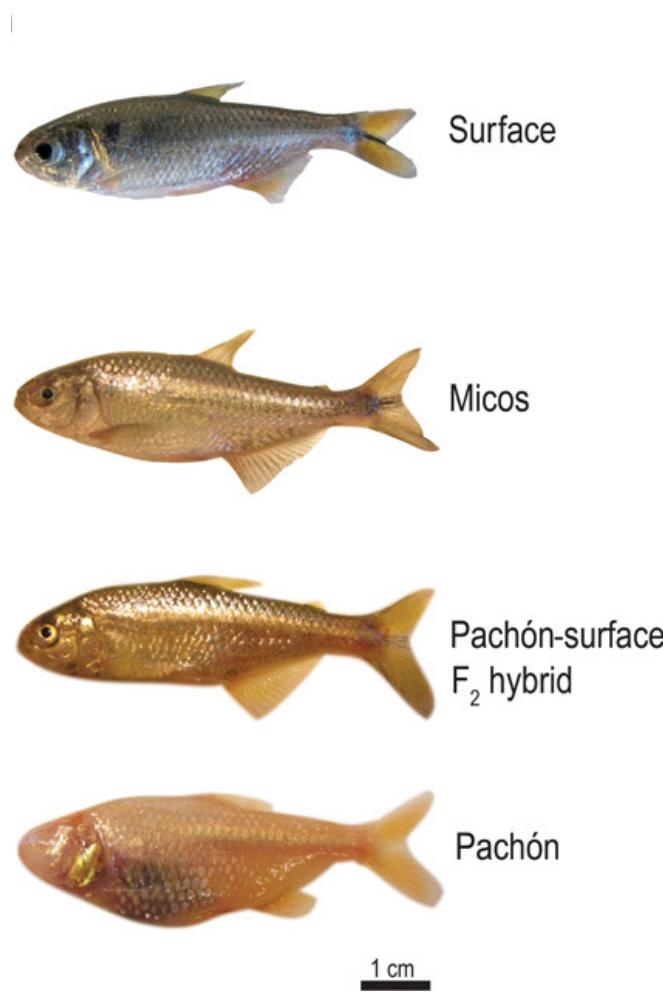
Dmitry Belyaev



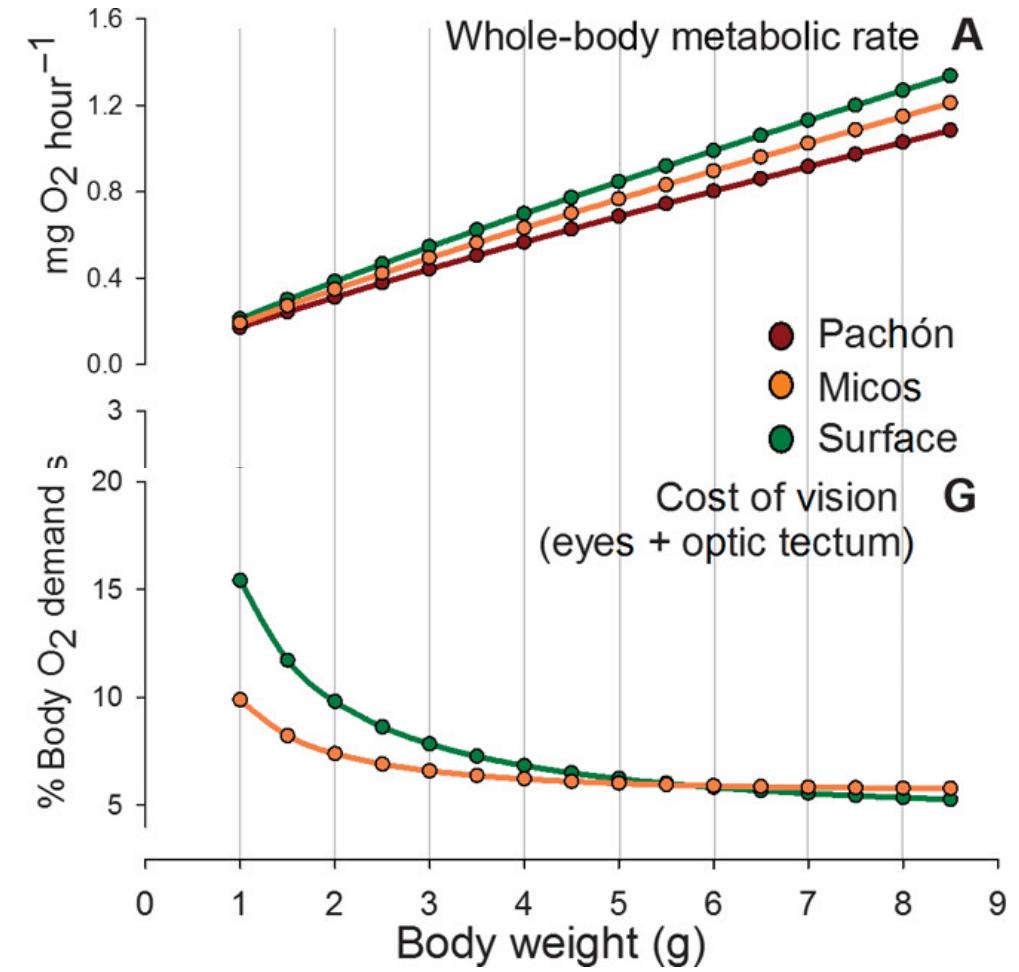
Lyudmila Trut



Trut 1999



Eyes are probably selected against



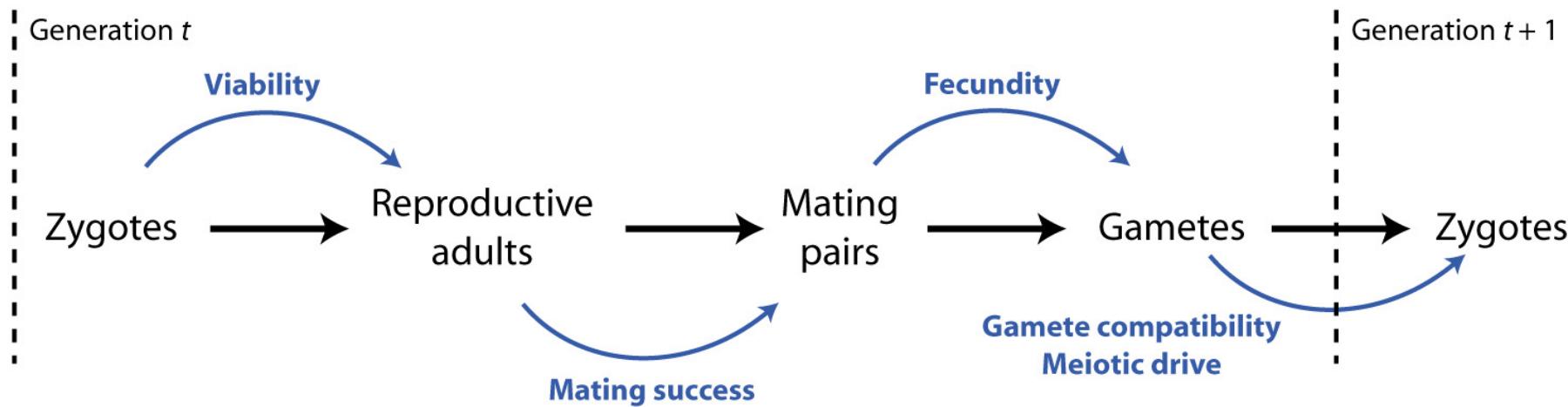
Selection

- Differential survival and reproduction among genotypes
- Goes back to Darwin (1859)

Challenge: develop expectations for the genetic changes that are caused by selection

SELECTION

May happen at different life stages



Building simple model

- Consider haploid first
- A and a alleles
- N_A and N_a are numbers of individuals carrying each allele
- w_A and w_a are the average number of descendants for A and a individuals

In the next generation: $f_A = \frac{w_A N_A}{w_A N_A + w_a N_a}$

Selection acts on relative fitness

- **Absolute Fitness**
 - The average number of offspring of a given type per parent of the given type.
- **Relative Fitness**
 - The average contribution to the offspring generation relative to the contribution of another type.

$$f_A(t+1) = \frac{w_A N_A}{w_A N_A + w_a N_a}$$

$$f_A(t+1) = \frac{w_A f_A}{w_A f_A + w_a f_a}$$

$$f_A(t+1) = \frac{f_A}{f_A + (\frac{w_a}{w_A}) f_a}$$

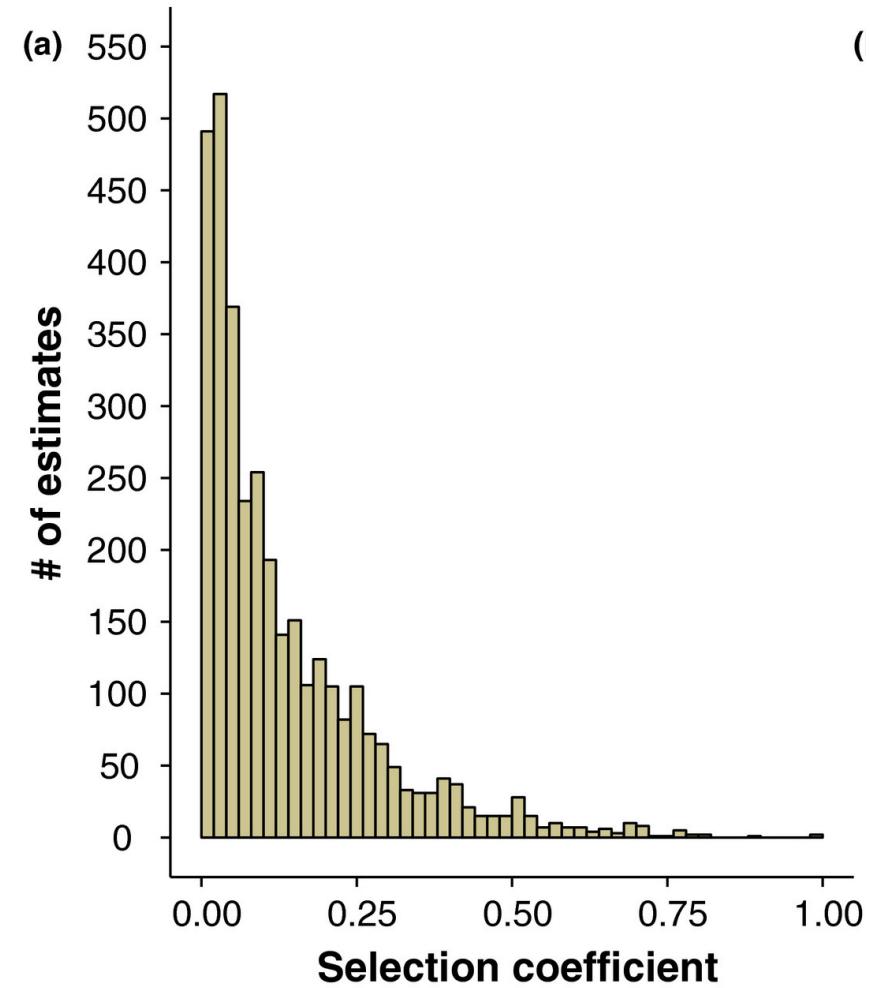
$$\begin{aligned}N_A &= 1000 \\N_a &= 3000\end{aligned}$$

$$\begin{aligned}w_A &= 17 & \text{or} & & w_A &= 1.7 \\w_a &= 15 & & & w_a &= 1.5\end{aligned}$$

Selection coefficient: s

$$w_a = 1 - s$$

$$w_A = 1$$



Thurman and Barrett 2016

Diploids

- Think about genotypes

AA	Aa	aa
v_{AA}	v_{Aa}	v_{aa}

Genetic

- Diploid individuals
- One locus with two alleles
- Obligate sexual reproduction

Reproduction

- Generations do not overlap
- Mating is random

Natural selection

- Mechanism of natural selection is genotype-specific differences in survivorship (fitness) that lead to variable genotype-specific growth rates, termed viability selection
- Fitness values are constants that do not vary with time, over space, or in the two sexes

Population

- Infinite population size so there is no genetic drift
 - No population structure
 - No gene flow
 - No mutation
-

Diploids

- Think about genotypes

AA	Aa	aa
v_{AA}	v_{Aa}	v_{aa}

- If allele frequency is f_A and f_a

- Number of indivs surviving to adulthood:

- $N * v_{AA} * f_A^2$

- $N * v_{AA} * 2f_A f_a$

- $N * v_{AA} * f_a^2$

- Average viability

- $v_{AA}f_A^2 + v_{AA}2f_A f_a + v_{AA}f_a^2 = \bar{v}$

- Total number of adults next gen

- $N * \bar{v}$

Genotype freqs in next gen:

$$f_{AA} = \frac{v_{AA}f_A^2}{\bar{v}}$$

$$f_{Aa} = \frac{2v_{Aa}f_A f_a}{\bar{v}}$$

$$f_{aa} = \frac{v_{aa}f_a^2}{\bar{v}}$$

Allele freqs in next gen:

$$f_A = \frac{v_{AA}f_A^2 + v_{Aa}2f_A f_a}{\bar{v}}$$

No change in allele frequencies when $pq = 0$ (fixation or loss)

No change in allele frequencies when all genotypes have the same fitness

Some terms

- Additive
- Dominant
- Recessive
- Heterozygote disadvantage
- Heterozygote advantage

ADDITIVE SELECTION

$$v_{AA} = 1$$

$$v_{Aa} = 1 - s$$

$$v_{aa} = 1 - 2s$$

Each copy of a reduces viability by S

$$\Delta f_A = \frac{sf_A f_a}{1 - 2sfa}$$

If S is small, then denominator is about 1

$$\Delta f_A = sf_A f_a$$

Can also think of relative fitness : $w_{AA} = \frac{v_{AA}}{v_{AA}}$; $w_{Aa} = \frac{v_{Aa}}{v_{AA}}$; $w_{aa} = \frac{v_{aa}}{v_{AA}}$

Category	Genotype-specific fitness		
	w_{AA}	w_{Aa}	w_{aa}
Selection against a recessive phenotype	1	1	$1 - s$
Selection against a dominant phenotype	$1 - s$	$1 - s$	1
General dominance (dominance coefficient $0 \leq h \leq 1$)	1	$1 - hs$	$1 - s$
Heterozygote disadvantage (underdominance for fitness)	1	$1 - s$	1
Heterozygote advantage (overdominance for fitness)	$1 - s$	1	$1 - t$

- Decrease in viability of a genotype compared to a relative fitness of 1

$$s_{xx} = 1 - w_{xx}$$

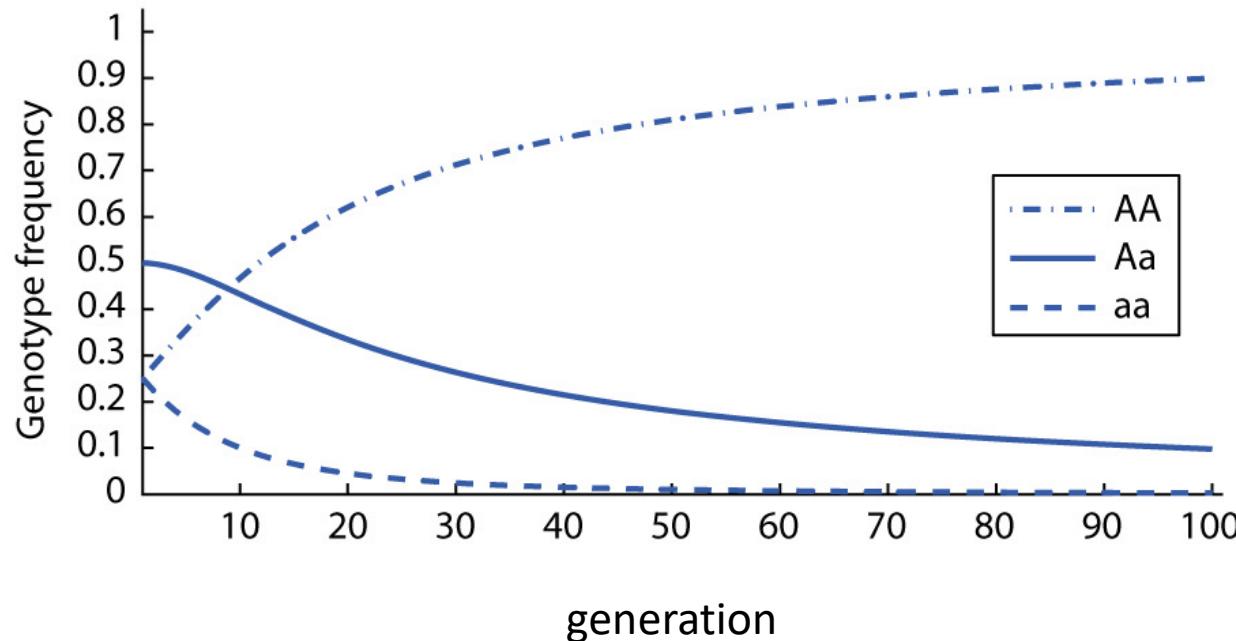
Recessive:

$$\frac{v_{AA}}{v_{AA}} = 1 \quad \text{and} \quad \frac{v_{Aa}}{v_{AA}} = 1 \quad \text{and.} \quad \frac{v_{aa}}{v_{AA}} = 1 - s_{aa}$$

Heterozygote advantage

$$\frac{v_{Aa}}{v_{Aa}} = 1 \quad \text{and} \quad \frac{v_{aa}}{v_{Aa}} = 1 - s_{aa} \quad \text{and} \quad \frac{v_{AA}}{v_{Aa}} = 1 - s_{AA}$$

Selection against a recessive phenotype



Deterministic (as opposed to genetic drift that is stochastic)!

Fixation?

No

Why?

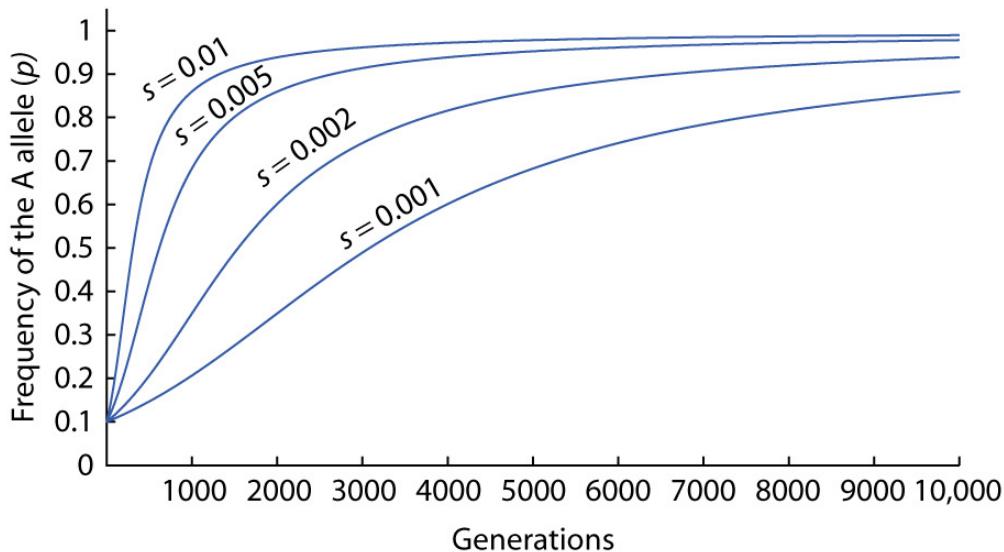
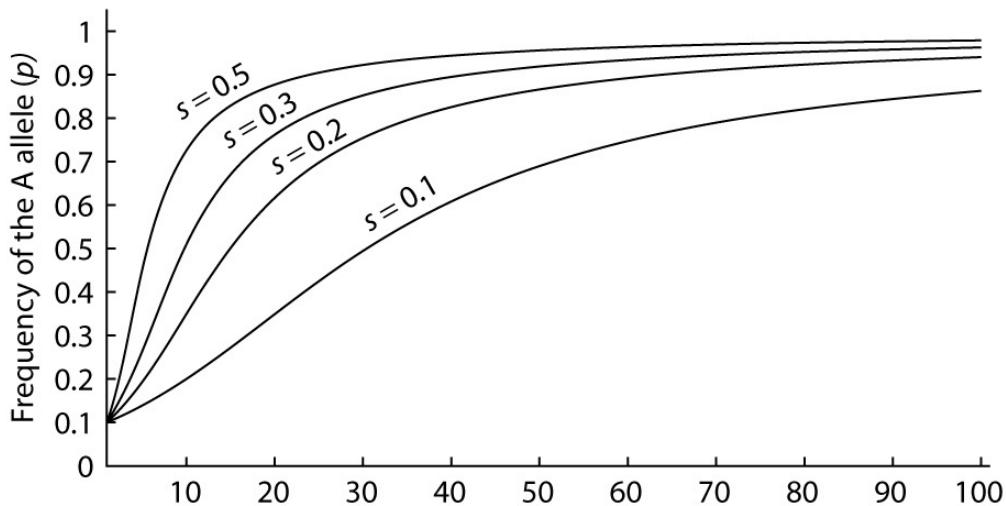
Selection against a recessive phenotype

Weaker/stronger selection:

- Same effect but on different timescales!

The same applies to all categories of selection

$$w_{AA} = w_{Aa} = 1 \quad W_{aa} = 0.8$$



Selection against a dominant phenotype

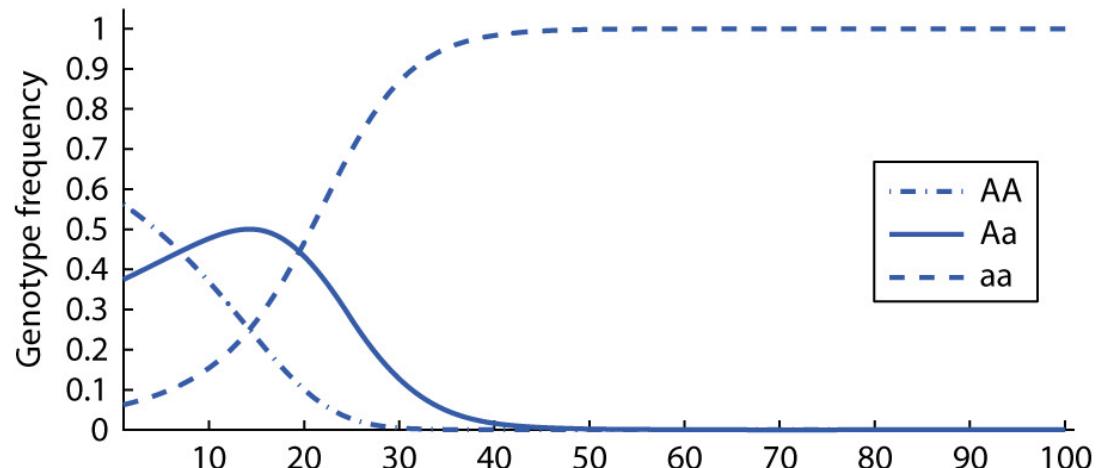
Heterozygote frequency initially increases until the two alleles have the same frequency of 0.5 (at which point the two homozygotes have a frequency of 0.25)

Fixation?

Yes!

Why?

Selection against both homozygotes and heterozygotes

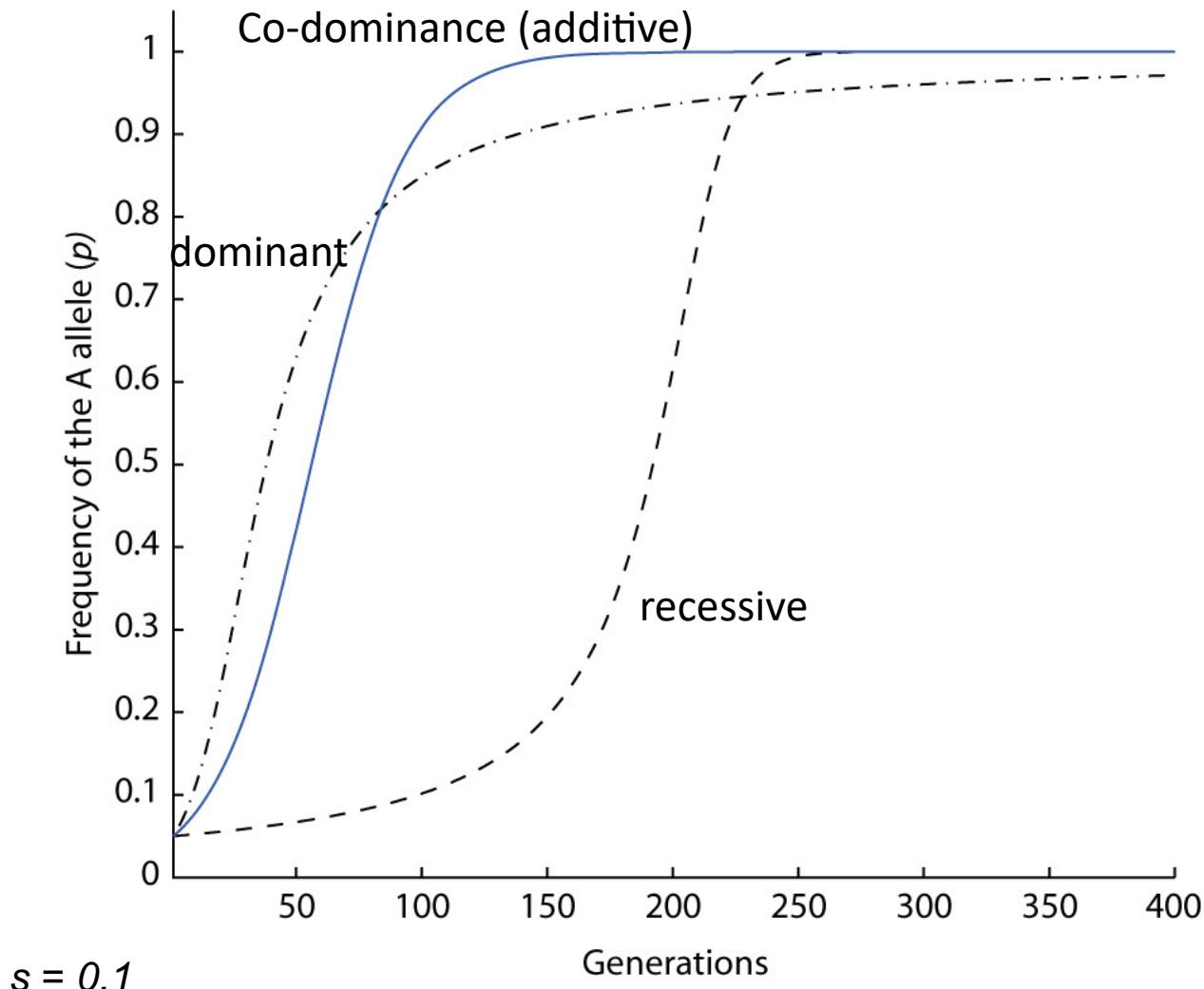


$$w_{AA} = w_{Aa} = 1 \quad W_{aa} = 0.8$$

General dominance

- $h = 0$: allele A is dominant and selected for
- $h = 1$: allele A is recessive and selected for
- $h = 0.5$: co-dominance
- Co-dominance combines the rapid initial change of the recessive case with fixation of the dominant case

Category	Genotype-specific fitness		
	w_{AA}	w_{Aa}	w_{aa}
Selection against a recessive phenotype	1	1	$1 - s$
Selection against a dominant phenotype	$1 - s$	$1 - s$	1
General dominance (dominance coefficient $0 \leq h \leq 1$)	1	$1 - hs$	$1 - s$
Heterozygote disadvantage (underdominance for fitness)	1	$1 - s$	1
Heterozygote advantage (overdominance for fitness)	$1 - s$	1	$1 - t$



Heterozygote disadvantage (underdominance, disruptive selection)

Category	Genotype-specific fitness		
	w_{AA}	w_{Aa}	w_{aa}
Selection against a recessive phenotype	1	1	$1 - s$
Selection against a dominant phenotype	$1 - s$	$1 - s$	1
General dominance (dominance coefficient $0 \leq h \leq 1$)	1	$1 - hs$	$1 - s$
Heterozygote disadvantage (underdominance for fitness)	1	$1 - s$	1
Heterozygote advantage (overdominance for fitness)	$1 - s$	1	$1 - t$

In exercise today

Heterozygote advantage (overdominance, balancing selection)

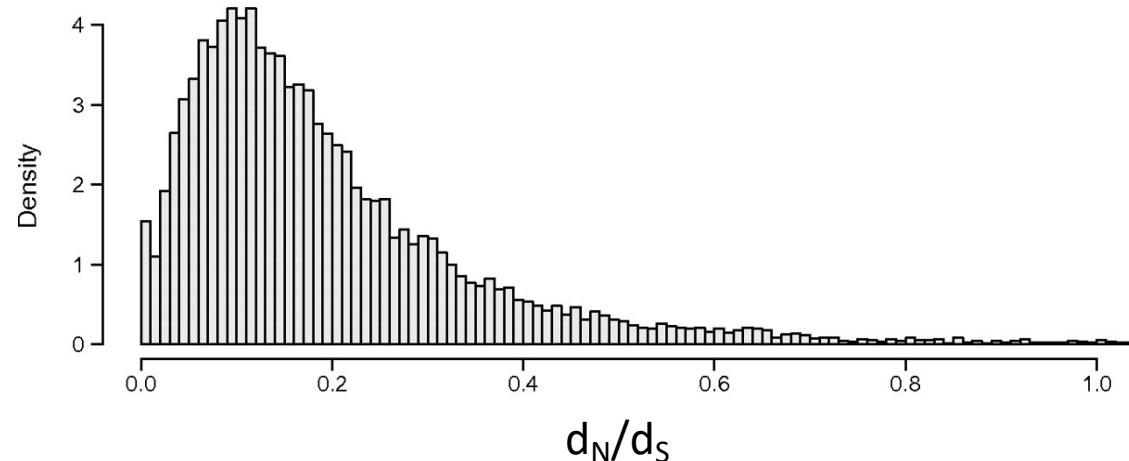
Category	Genotype-specific fitness		
	w_{AA}	w_{Aa}	w_{aa}
Selection against a recessive phenotype	1	1	$1 - s$
Selection against a dominant phenotype	$1 - s$	$1 - s$	1
General dominance (dominance coefficient $0 \leq h \leq 1$)	1	$1 - hs$	$1 - s$
Heterozygote disadvantage (underdominance for fitness)	1	$1 - s$	1
Heterozygote advantage (overdominance for fitness)	$1 - s$	1	$1 - t$

In exercise today

A few methods to detect selection

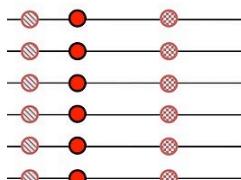
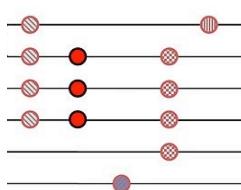
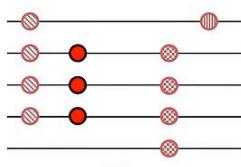
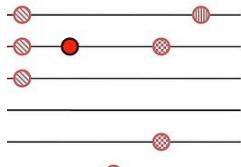
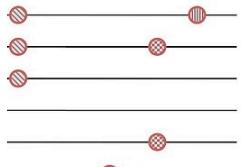
d_N/d_S ratios

- Compare synonymous to non-synonymous changes
- Neutrality
 - The majority of NS mutations are selected against
- Positive selection
- Purifying selection



Can't tell you about a single site.
Only works in coding regions

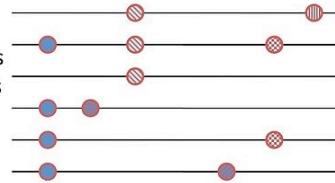
Selective sweeps and hitchhiking



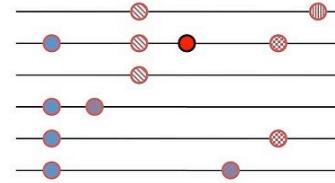
Selective sweeps and hitchhiking

Hard selective sweep with recombination

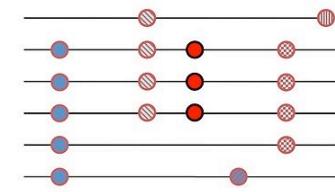
1. Six sequences from a population. Each line represents a DNA sequence and each dot is a neutral mutation, present in one or more sequences.



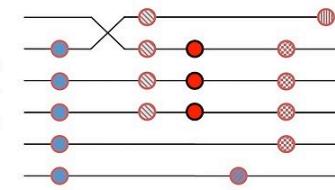
2. A beneficial mutation occurs (bright red dot on the second sequence).



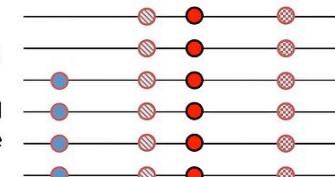
3. The beneficial mutation increases in frequency in the population, and so does the genomic background it is associated with.



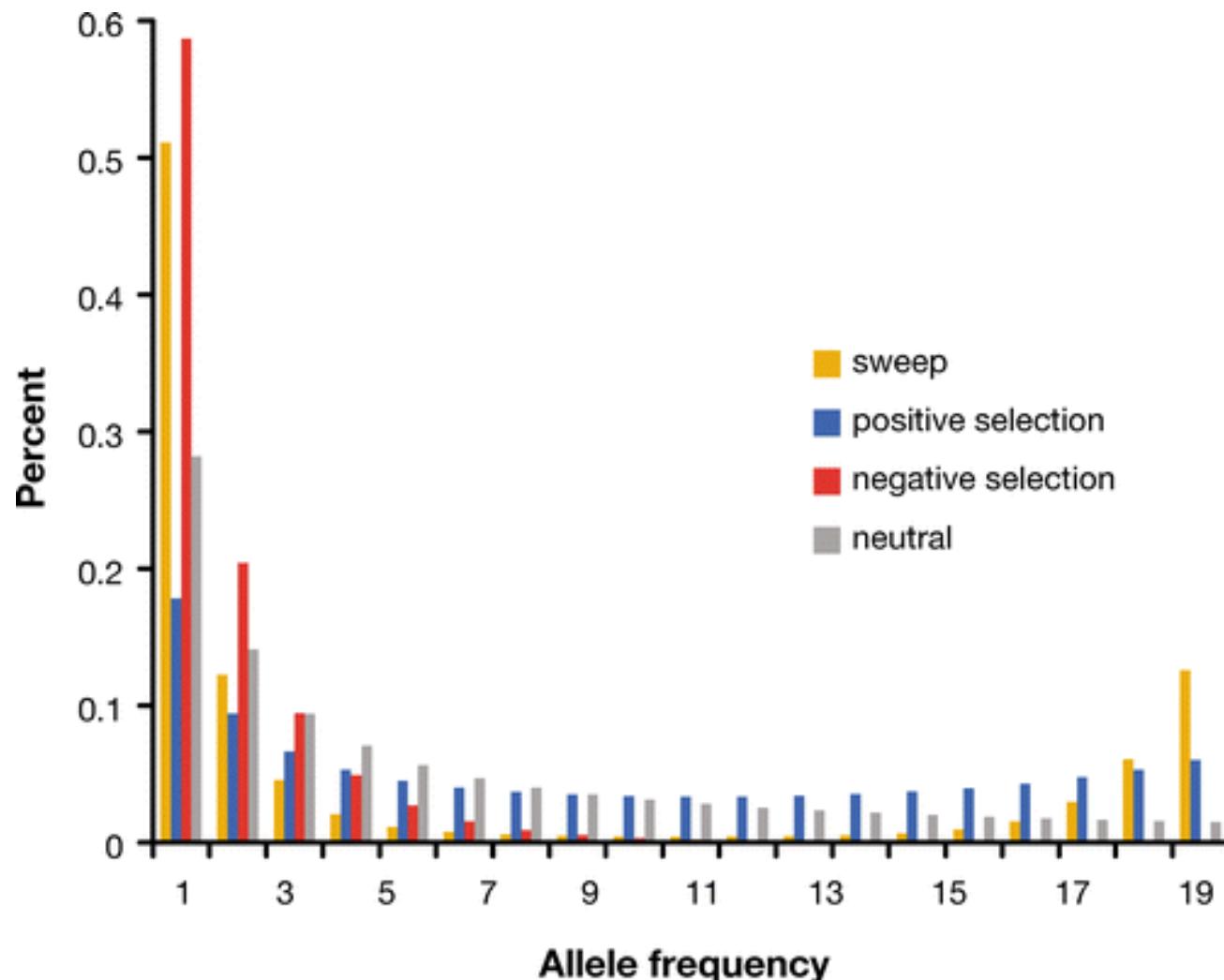
4. A recombination event creates a new combination (the beneficial mutation is no longer always associated with the blue neutral mutation).



5. The beneficial mutation is fixed in the population. Close to the beneficial mutation there is now no genetic variation left, all sequences look the same. Three neutral mutations have hitchhiked along with the beneficial mutation and reached high frequency.



Site frequency spectrum



Tajima's D

- Remember θ ...
 - θ_T : Tajima's estimator: pairwise number of differences
 - θ_W : Watterson's θ : number of segregating sites
- Tajima's $D = \frac{\theta_T - \theta_W}{\sqrt{\text{Var}(\theta_T - \theta_W)}}$
- The difference between θ_T and θ_W

Interpreting Tajima's D

- Think of bottleneck
 - Loss of variation
 - Most mutations would be rare → they arise after the bottleneck

Case 1:

---*--*---
-----*--*---
-----*-----
-----*---

Case 2:

-*-----
-----*-----
-----*-----
-----*-----

Interpreting Tajima's D

Case 1:

---*-*---
-----*-*---
-----*---
-----*---

$$\theta_T = \frac{2+1+3+1+1+2}{6} = 1.67$$

$$\theta_W: \frac{3}{1 + \frac{1}{2} + \frac{1}{3}} = 1.63$$

Case 2:

-*-----
---*-----
-----*---
-----*---

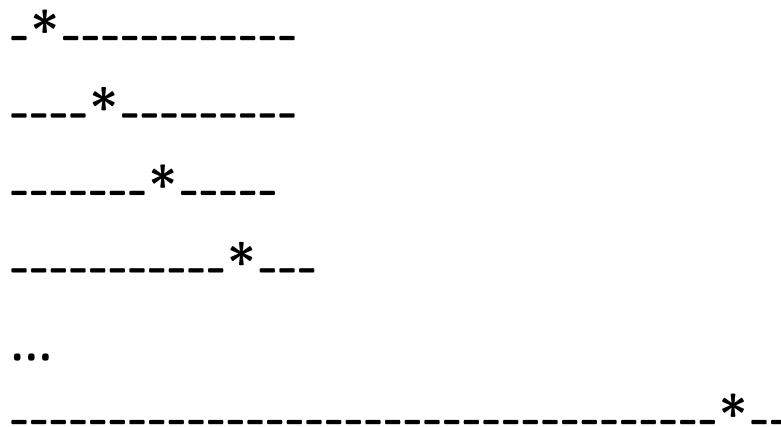
$$\theta_T = \frac{2+2+2+2+2+2}{6} = 2$$

$$\theta_W: \frac{4}{1 + \frac{1}{2} + \frac{1}{3}} = 2.2$$

θ_T underestimates
when variation is
rare

Interpreting Tajima's D

Case 3:

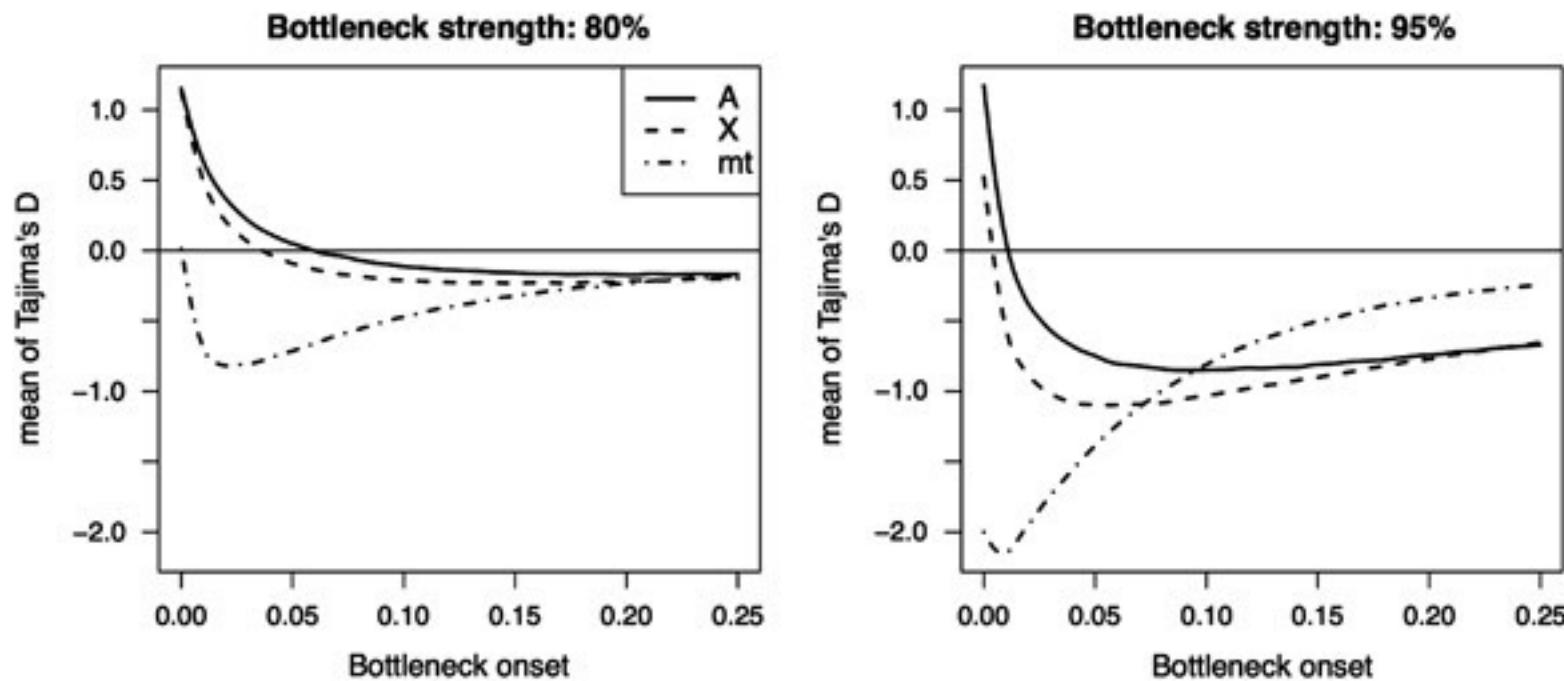


$$\theta_T = \frac{2*45}{45} = 2$$

$$\theta_W: \frac{10}{1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{9}} = 3.5$$

- rare variants across the sample, there is a huge difference between estimates.
- Most haplotypes the same after strong selection/bottleneck
- mutations that occur will be rare.
- Tajima's D will be negative.

Tajima's D following bottleneck



Paper for next week

Natural Selection on a Major Armor Gene in Threespine Stickleback

Rowan D. H. Barrett,* Sean M. Rogers, Dolph Schluter

Experimental estimates of the effects of selection on genes determining adaptive traits add to our understanding of the mechanisms of evolution. We measured selection on genotypes of the *Ectodysplasin* locus, which underlie differences in lateral plates in threespine stickleback fish. A derived allele (low) causing reduced plate number has been fixed repeatedly after marine stickleback colonized freshwater from the sea, where the ancestral allele (complete) predominates. We transplanted marine sticklebacks carrying both alleles to freshwater ponds and tracked genotype frequencies over a generation. The low allele increased in frequency once lateral plates developed, most likely via a growth advantage. Opposing selection at the larval stage and changing dominance for fitness throughout life suggest either that the gene affects additional traits undergoing selection or that linked loci also are affecting fitness.