

Lecture 5:
Selection 1

Neutral Theory

We have looked at models with mutation and drift

In those models, genetic diversity is explained by a balance of neutral mutation and drift (and migration)

The Neutral Theory makes the approximation that all mutations that are not neutral are deleterious (“dead-on-arrival”).

That is new mutations are either strongly negatively selected or neutral.

Selection

Besides neutral mutations, and strongly deleterious ones, some mutations must confer an advantage, and be positively selected.

How do these mutations spread throughout a population?

Allelic Selection - 3 general types

1. directional (positive or negative)

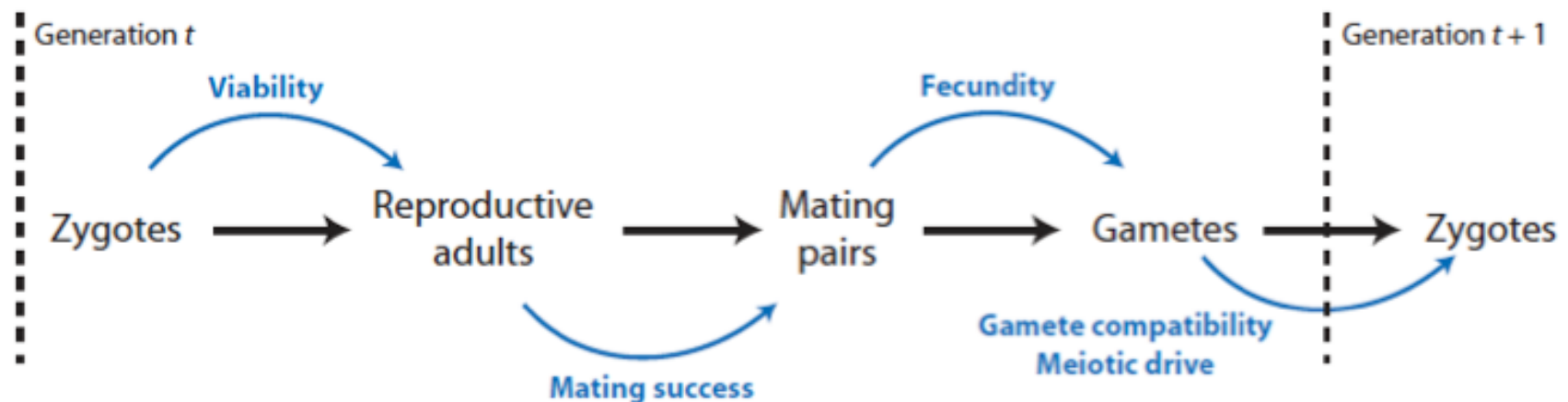
- selective sweeps (new allele advantageous, becomes fixed)
- purifying (new allele deleterious, will tend to disappear)

2. Heterozygous advantage (aka diversifying, balancing, frequency dependent or over-dominance)

3. Heterozygous disadvantage

Selection can act at many stages of the life cycle (filters)

- Natural selection depends on heritable variation in the number of offspring produced by individuals
- There are many events in a life history that affects offspring production



NOTE: In today's lecture, we'll introduce ourselves to models of natural selection by focusing on *viability selection*.

fitness

Definition

individual fitness: The reproductive contribution of an individual to the next generation.

fitness

Definition

individual fitness: The reproductive contribution of an individual to the next generation.

The population genetic theory of natural selection is mainly derived in terms of *genotypic fitnesses*

fitness

Definition

individual fitness: The reproductive contribution of an individual to the next generation.

Definition

genotypic fitness: The average fitness of all individuals who have the same genotype (e.g. w_{ij} , the genotypic fitness of A_iA_j).

The population genetic theory of natural selection is mainly derived in terms of *genotypic fitnesses*

absolute vs relative fitness

We can define genotypic fitnesses in absolute or relative terms:

Definition

absolute fitness: The expected number of offspring produced by a parent of a particular genotype.

Diploid selection

Genotype	Genotypic fitness	# at birth	# after selection
AA	w_{11}	$p^2 \times N$	$p^2 \times Nw_{11}$
Aa	w_{12}	$2pq \times N$	$2pq \times Nw_{12}$
aa	w_{22}	$q^2 \times N$	$q^2 \times Nw_{22}$
Total	-	N	$N(p^2 w_{11} + 2pqw_{12} + q^2 w_{22})$

Diploid selection

Genotype	Genotypic fitness	# at birth	# after selection
AA	w_{11}	$p^2 \times N$	$p^2 \times Nw_{11}$
Aa	w_{12}	$2pq \times N$	$2pq \times Nw_{12}$
aa	w_{22}	$q^2 \times N$	$q^2 \times Nw_{22}$
Total	-	N	$N(p^2 w_{11} + 2pqw_{12} + q^2 w_{22})$

Definition

$$\bar{w} = p^2 w_{11} + 2p(1 - p)w_{12} + (1 - p)^2 w_{22}.$$

We call \bar{w} the **mean population fitness**.

Mean Fitness affects population growth

Note from last slide:

$$N_{t+1} = \bar{w} N_t.$$

So the mean absolute fitness affects population growth. $\bar{w} > 1$ equals population growth, $\bar{w} < 1$ equals population decline.

Note

Our main interest is in how genotype and allele frequencies change through time...

relative fitness

Genotype	Genotypic fitness	# at birth	# after selection
AA	w_{11}	$p^2 \times N$	$p^2 \times Nw_{11}$
Aa	w_{12}	$2pq \times N$	$2pq \times Nw_{12}$
aa	w_{22}	$q^2 \times N$	$q^2 \times Nw_{22}$
Total	-	N	$N(p^2 w_{11} + 2pq w_{12} + q^2 w_{22})$

Genotype	Genotypic fitness	% at birth	% after selection
AA	w_{11}	p^2	$\frac{p^2 w_{11}}{\bar{w}}$
Aa	w_{12}	$2pq$	$\frac{2pq w_{12}}{\bar{w}}$
aa	w_{22}	q^2	$\frac{q^2 w_{22}}{\bar{w}}$
Total	-	1	1

Note

For studying genotypic frequencies, if we multiply w 's by a constant C , the scaling factor C cancels out in the ratio's $w_{..}/\bar{w}$. This is useful for us, because we can express work in terms of *relative fitnesses*.

relative fitness

Definition

relative fitness: Absolute fitnesses scaled in some way. (e.g. by dividing by the largest absolute fitness, such that the most fit genotype has relative fitness = 1).

Table 6.4 The general categories of relative fitness values for viability selection at a diallelic locus. The variables s and t are used to represent the decrease in viability of a genotype compared to the maximum fitness of 1 ($1 - w_{xx} = s$). The degree of dominance of the A allele is represented by h with additive gene action (sometime called codominance) when $h = 1/2$.

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	$1 - t$

Solving for the change in allele frequency per generation

From previous slide we can derive:

$$p' = \frac{p^2 w_{11} + pqw_{12}}{\bar{w}}$$

And solve for:

$$\Delta p = \frac{pq[p(w_{11} - w_{12}) + q(w_{12} - w_{22})]}{\bar{w}}$$

Solving for the change in allele frequency per generation

$$p' = \frac{p^2 w_{11} + pqw_{12}}{\bar{w}}$$

$$p' - p = \frac{p^2 w_{11} + pqw_{12}}{\bar{w}} - \frac{\bar{w}}{\bar{w}} p$$

$$\Delta p = \frac{p^2 w_{11} + pqw_{12} - \bar{w} p}{\bar{w}}$$

$$= \frac{1}{\bar{w}} [p^2 w_{11} + pqw_{12} - p(p^2 w_{11} + 2pqw_{12} + q^2 w_{22})]$$

$$= \frac{1}{\bar{w}} [p^2 w_{11} + pqw_{12} - p^3 w_{11} - 2p^2 q w_{12} - pq^2 w_{22}]$$

$$= \frac{1}{\bar{w}} [p^2 w_{11}(1 - p) + pqw_{12}(1 - p - p) - pq^2 w_{22}]$$

$$= \frac{1}{\bar{w}} [p^2 q w_{11} + pqw_{12}(q - p) - pq^2 w_{22}]$$

$$= \frac{pq}{\bar{w}} [p w_{11} + q w_{12} - p w_{12} - q w_{22}]$$

$$= \frac{pq [p(w_{11} - w_{12}) + q(w_{12} - w_{22})]}{\bar{w}}$$

Summary

The change in allele frequency in any generation:

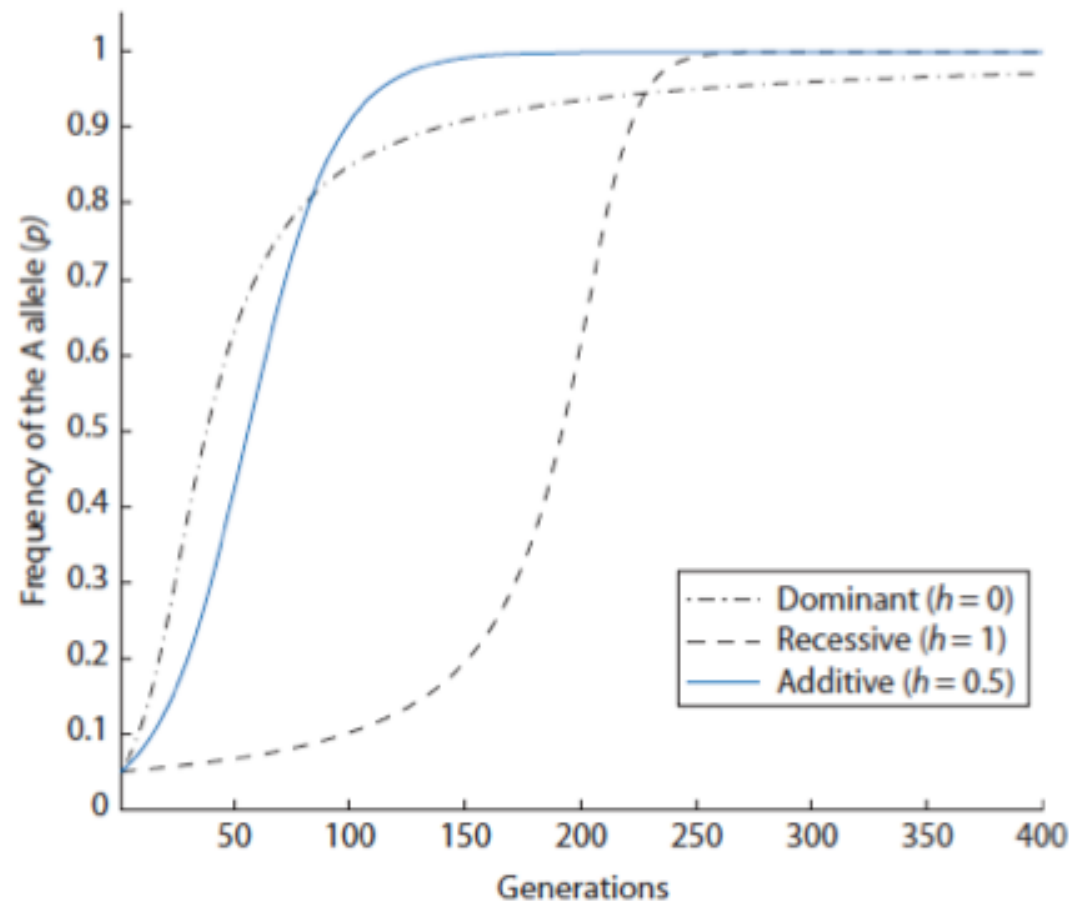
$$\Delta p = \frac{pq[p(w_{11} - w_{12}) + q(w_{12} - w_{22})]}{\bar{w}}.$$

Take-home points about the potential equilibria:

- Allele frequencies do not change when $pq = 0$. (i.e. $p = 0$ or $q = 0$, no genetic variation).
- Allele frequencies do not change when all fitnesses are equal
- Allele frequencies do not change when the difference in fitness weighted by allele frequencies $[p(w_{11} - w_{12}) + q(w_{12} - w_{22})] = 0$.

Change allele frequencies - advantageous recessive/dominant/additive when rare

We can use our recursion formula from above to solve for how allele frequency changes through time, and allow for dominance/recessiveness of the A allele:



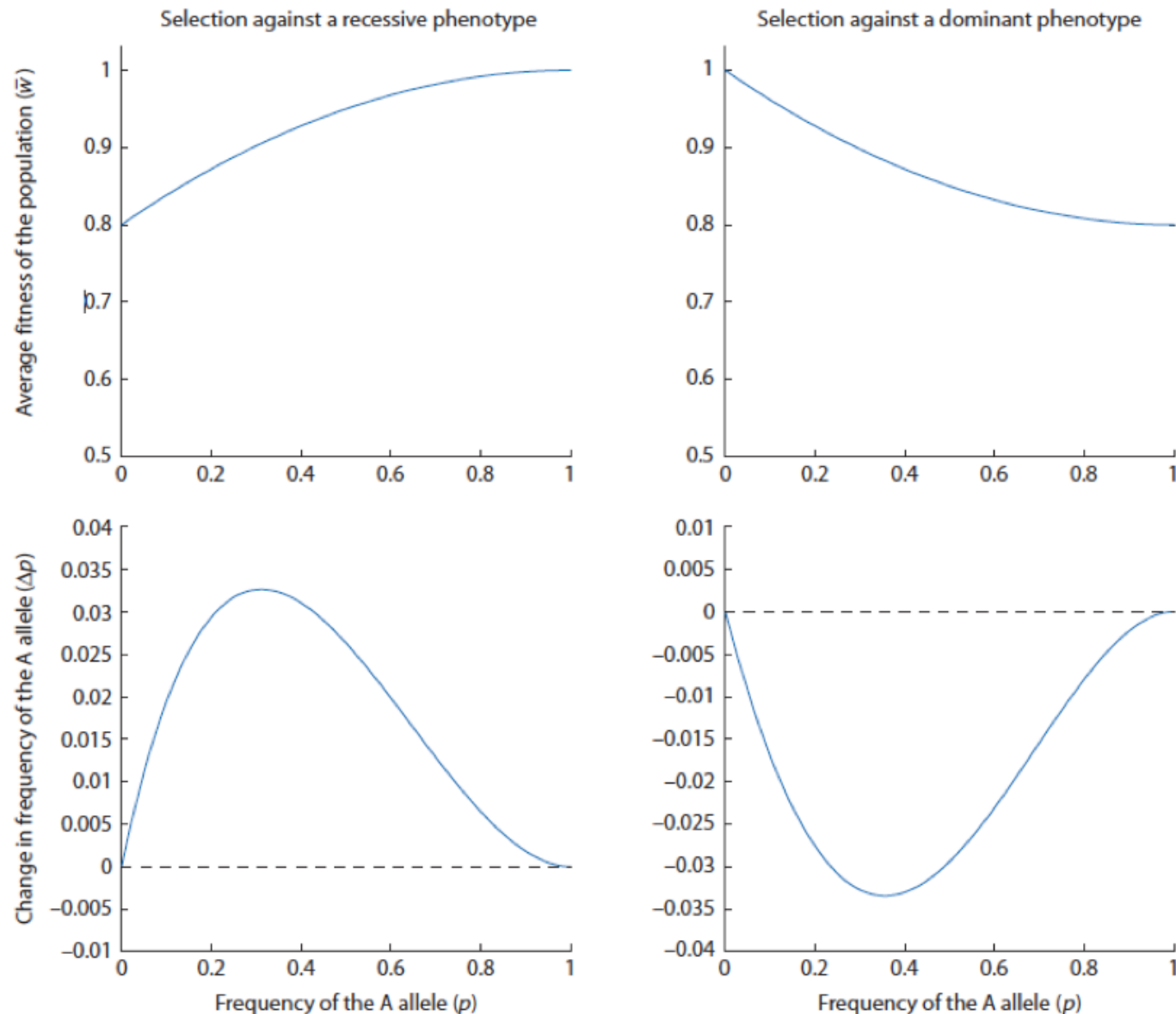
change in allele frequency per generation

Table 6.4 The general categories of relative fitness values for viability selection at a diallelic locus. The variables s and t are used to represent the decrease in viability of a genotype compared to the maximum fitness of 1 ($1 - w_{xx} = s$). The degree of dominance of the A allele is represented by h with additive gene action (sometime called codominance) when $h = 1/2$.

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	$1 - t$

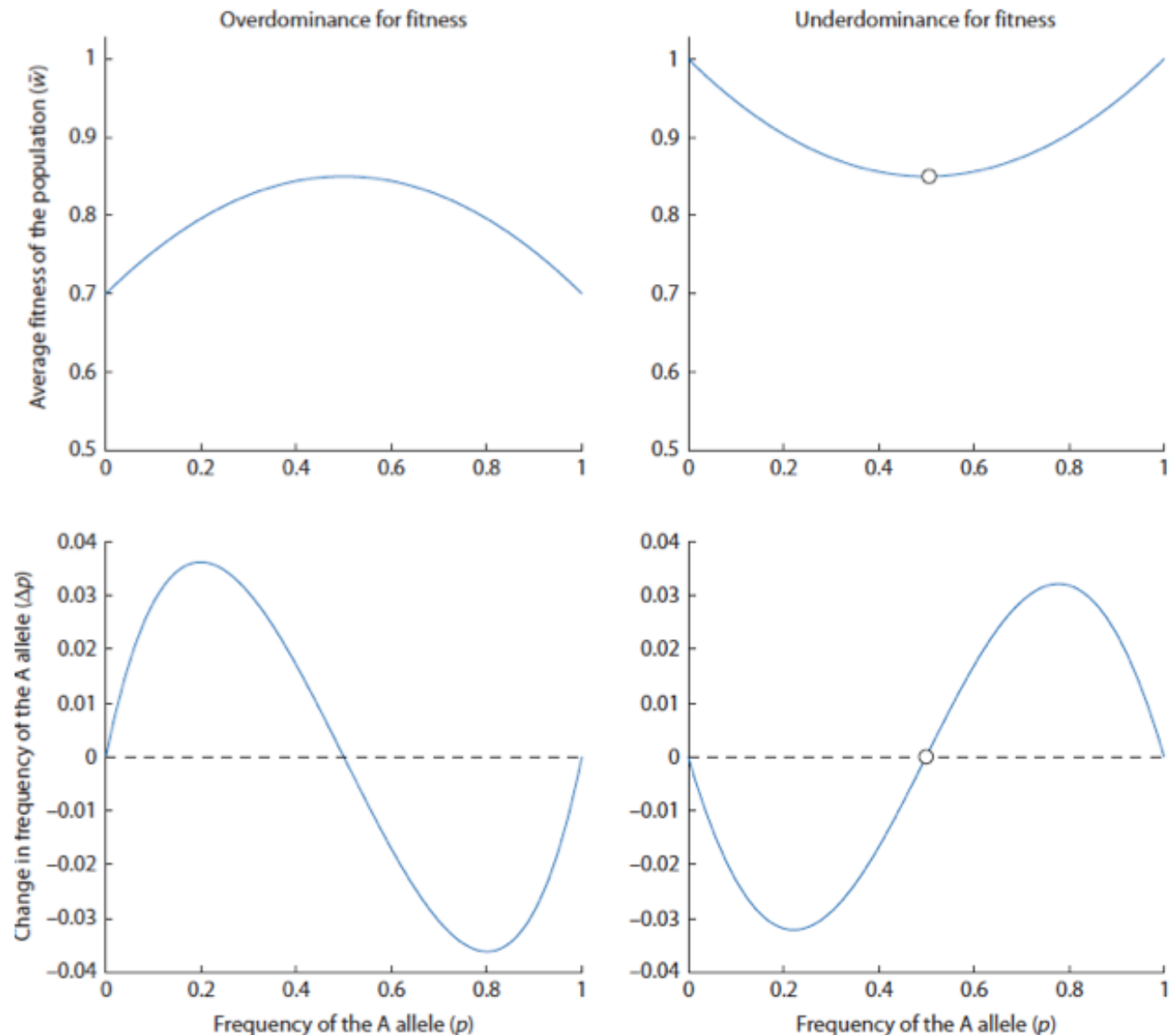
Fitness landscapes and the change in allele frequencies per generation

$$\bar{w} = w_{AA}p^2 + w_{Aa}2pq + w_{aa}q^2$$



Fitness landscapes and the change in allele frequencies per generation

For the overdominance and underdominance cases:



Over-dominance / heterozygote superiority

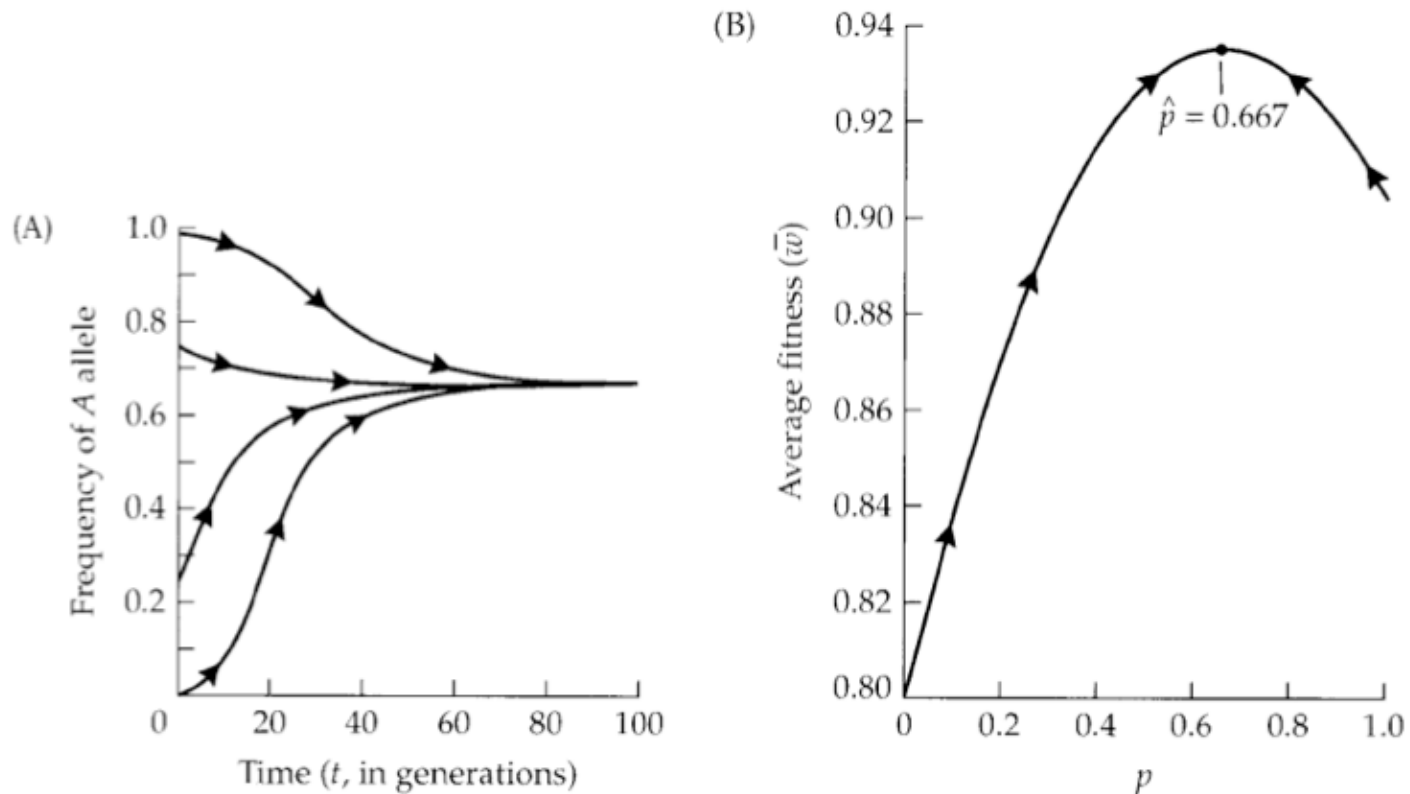
If $w_{11} = (1 - s)$, $w_{12} = 1$, $w_{22} = (1 - t)$, $p_{eq} = \frac{t}{s+t}$.

For $w_{11} = 0.9$, $w_{12} = 1$, and $w_{22} = 0.8$:

Over-dominance / heterozygote superiority

If $w_{11} = (1 - s)$, $w_{12} = 1$, $w_{22} = (1 - t)$, $p_{eq} = \frac{t}{s+t}$.

For $w_{11} = 0.9$, $w_{12} = 1$, and $w_{22} = 0.8$:



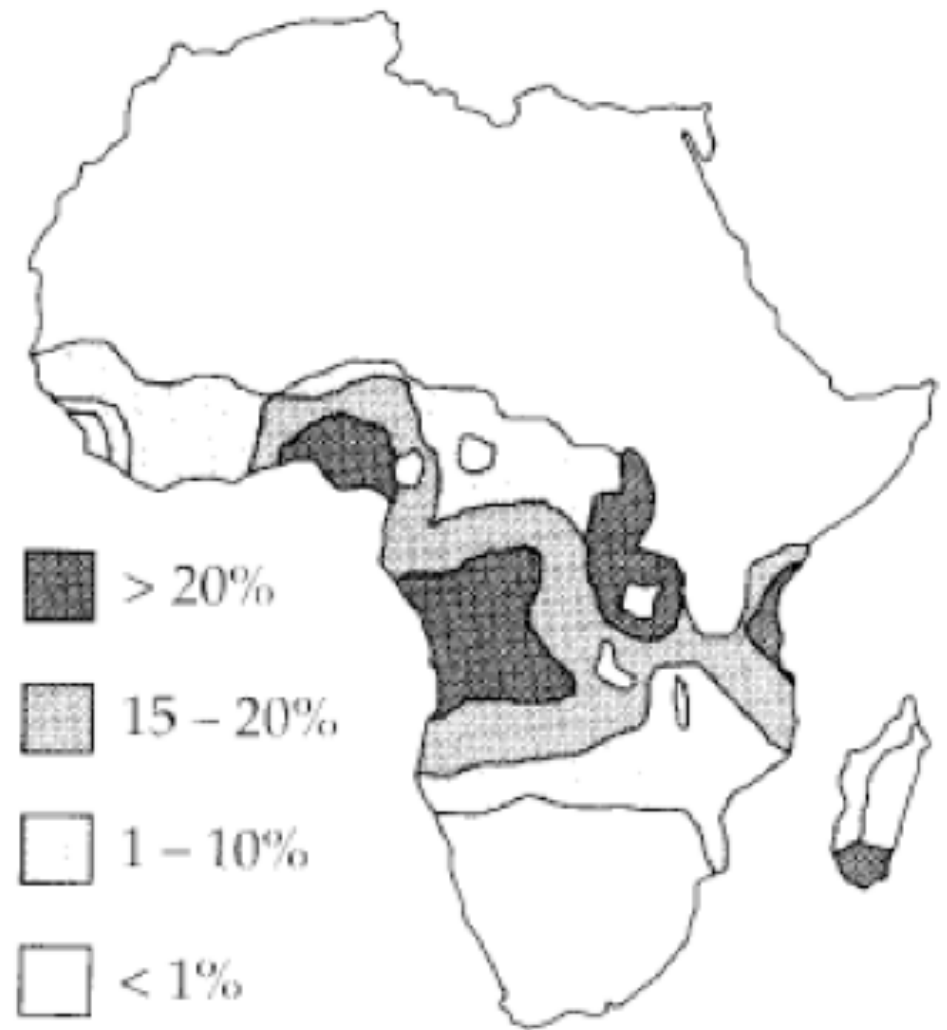
Note

Overdominance has a stable equilibrium that maintains variation in the variation.

Sickle-cell anemia & malaria



Malaria

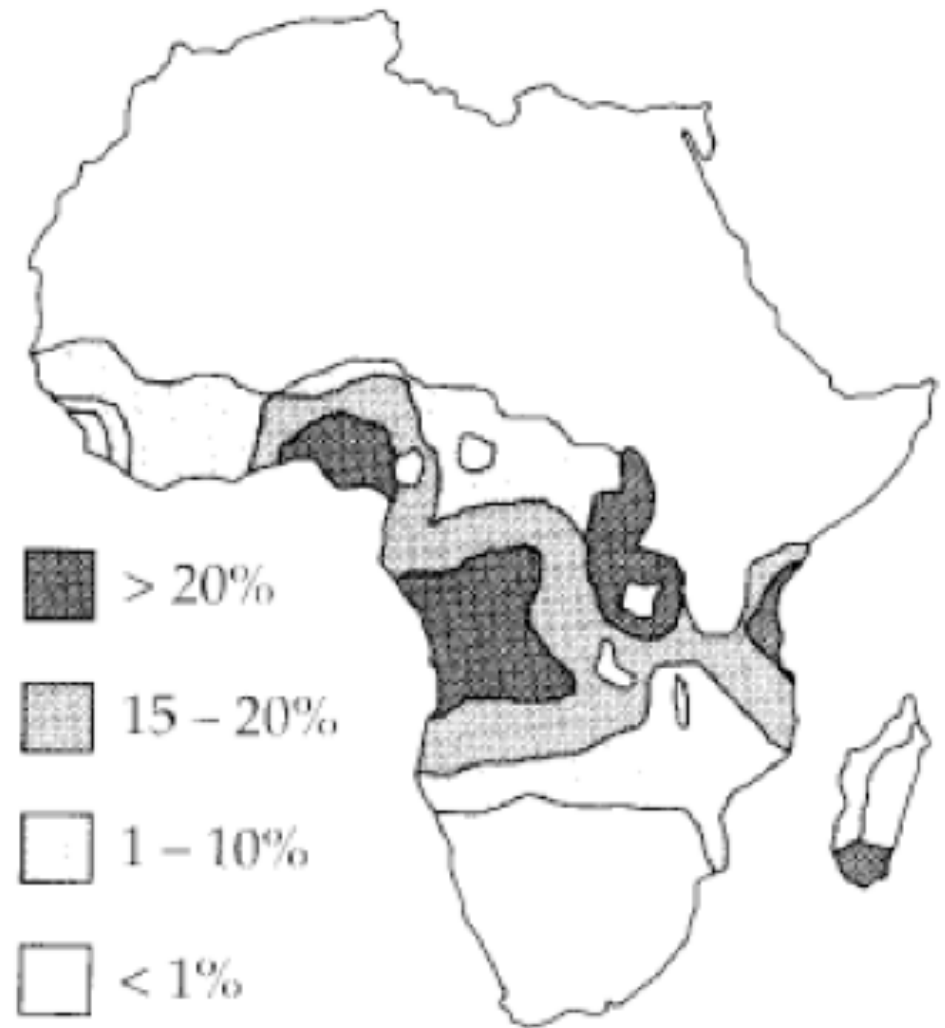


sickle-cell allele

Sickle-cell anemia & malaria



Malaria



sickle-cell allele

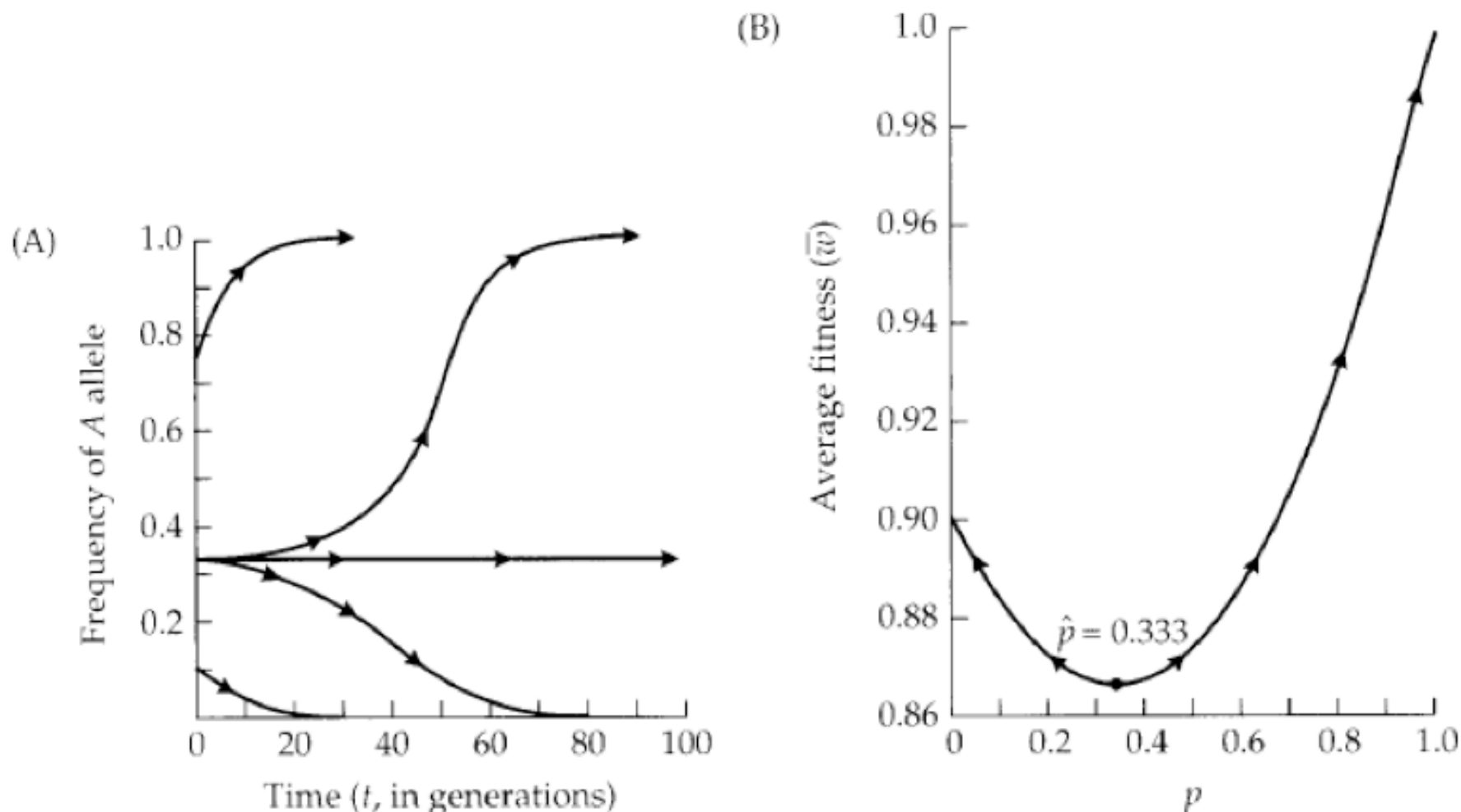
equilibrium frequency varies as a function of geographical space

Under-dominance

If $w_{11} = (1 + s)$, $w_{12} = 1$, $w_{22} = (1 + t)$, $p_{eq} = \frac{t}{s+t}$. For $w_{11} = 1.1, w_{12} = 1$, and $w_{22} = 1.2$:

Under-dominance

If $w_{11} = (1 + s)$, $w_{12} = 1$, $w_{22} = (1 + t)$, $p_{eq} = \frac{t}{s+t}$. For $w_{11} = 1.1, w_{12} = 1$, and $w_{22} = 1.2$:

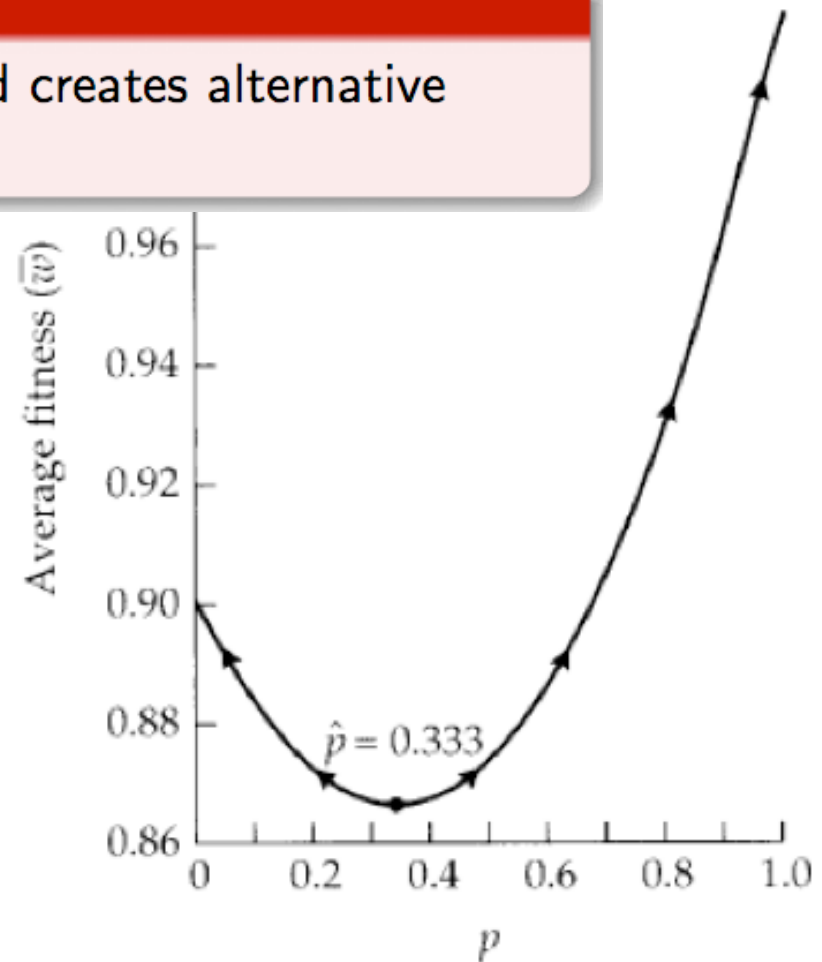
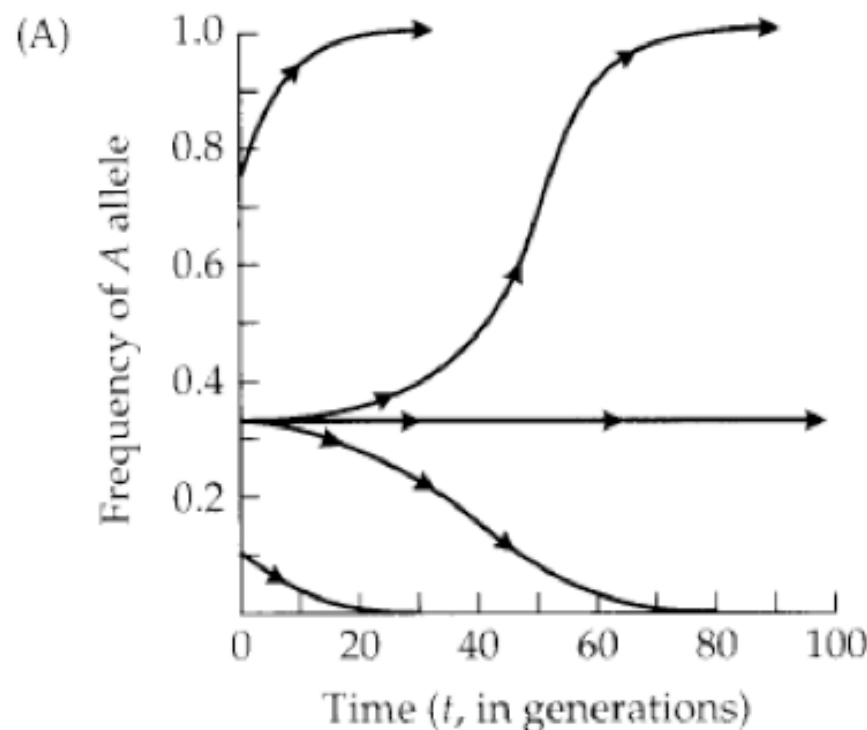


Under-dominance

If $w_{11} = (1 + s)$, $w_{12} = 1$, $w_{22} = (1 + t)$, $p_{eq} = \frac{t}{s+t}$. For $w_{11} = 1.1$, $w_{12} = 1$, and $w_{22} = 1.2$:

Note

Underdominance has an unstable equilibrium and creates alternative paths for evolution to take...



Strength and speed of selection

Top panel: Very strong selection; Bottom panel: Weaker selection.
Notice the different time-scales.

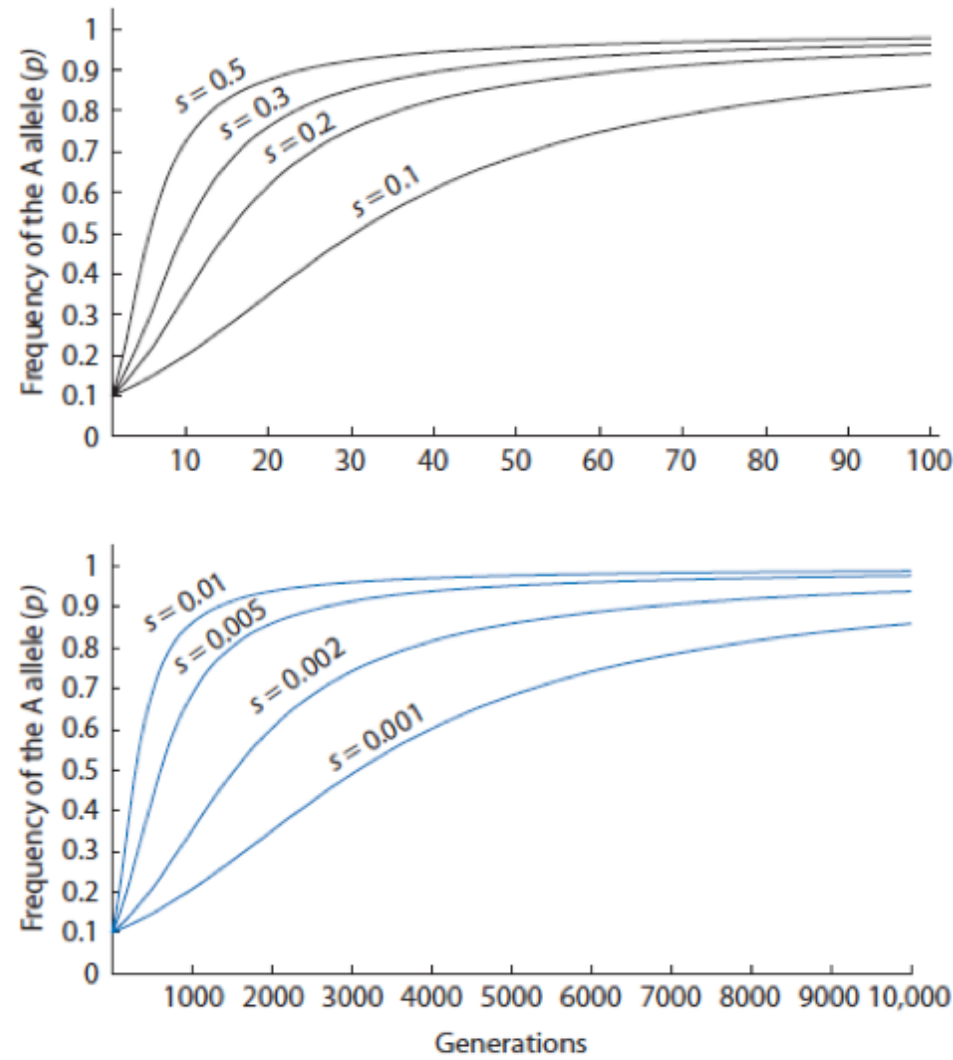


Figure 6.9 The strength of natural selection influences the rate of change in genotype and allele frequencies. In this

Conclusions

The speed at which selection increases an advantageous allele when rare depends on whether the allele is dominant or recessive, and the overall strength of selection.

Conclusions

The speed at which selection increases an advantageous allele when rare depends on whether the allele is dominant or recessive, and the overall strength of selection.

Selection can destroy variation (e.g. $p_{eq} = 0$ or 1 for directional selection)

Conclusions

The speed at which selection increases an advantageous allele when rare depends on whether the allele is dominant or recessive, and the overall strength of selection.

Selection can destroy variation (e.g. $p_{eq} = 0$ or 1 for directional selection)

Selection can also maintain variation (e.g. overdominance / also frequency-dependent selection).

Conclusions

The speed at which selection increases an advantageous allele when rare depends on whether the allele is dominant or recessive, and the overall strength of selection.

Selection can destroy variation (e.g. $p_{eq} = 0$ or 1 for directional selection)

Selection can also maintain variation (e.g. overdominance / also frequency-dependent selection).

A major game in evolutionary genetics today: Searching for the signatures of selection in genomes...

Conclusions

The speed at which selection increases an advantageous allele when rare depends on whether the allele is dominant or recessive, and the overall strength of selection.

Selection can destroy variation (e.g. $p_{eq} = 0$ or 1 for directional selection)

Selection can also maintain variation (e.g. overdominance / also frequency-dependent selection).

A major game in evolutionary genetics today: Searching for the signatures of selection in genomes...

Can we distinguish loci that have undergone recent positive selection from loci that have evolved according to the neutral theory (given the demographic history)

Next Week

Go over exam

Reading for Discussion

Evolutionary Modeling in SLiM 3 for Beginners

Benjamin C. Haller*,¹ and Philipp W. Messer¹

¹Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY

***Corresponding author:** E-mail: bhaller@benhaller.com.

Associate editor: Ryan Hernandez

Reading to keep up on (Selection)

Chapters 7 & 8

Simulation assignment

Population Structure and Selection

Next Week - Linked Selection

Simulation assignment

Use guiSLiM or SLiM w/ jupyter notebook

this exercise will allow some exploration of population structure and positive selection

Take the attached **SLiM code** and tweak the following parameters to explore the following parameter values. Describe the general behavior of the 1. allele frequency spectrum, 2. mutation fixation time spectrum, and 3. the mutation loss time spectrum. 4. When do “fixed differences arise between population p1 and p3?

parameter combinations to explore:

a.

m2 (MutationType) selection coefficient 0.002, 0.02, 0.2

b.

RecombinationRate() 1e-5, 1e-6, 1e-7

c.

all setMigrationRates() 0.5, 0.0005, 0.0000005