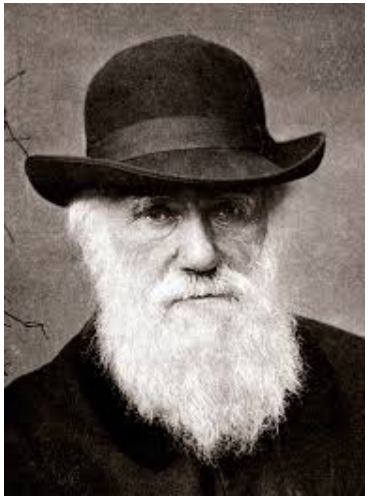
Population genetics

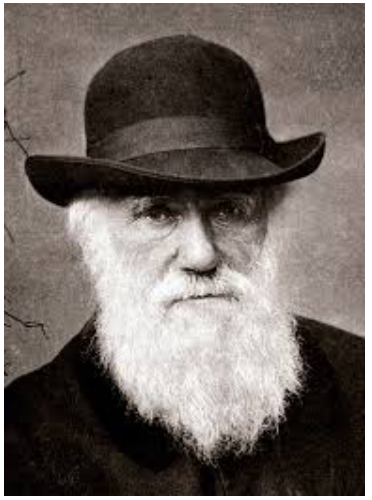
One big challenge for early 20th century biology: how to reconcile Mendelian genetics with Darwinian evolution?

??



One big challenge for early 20th century biology: how to reconcile Mendelian genetics with Darwinian evolution?

??

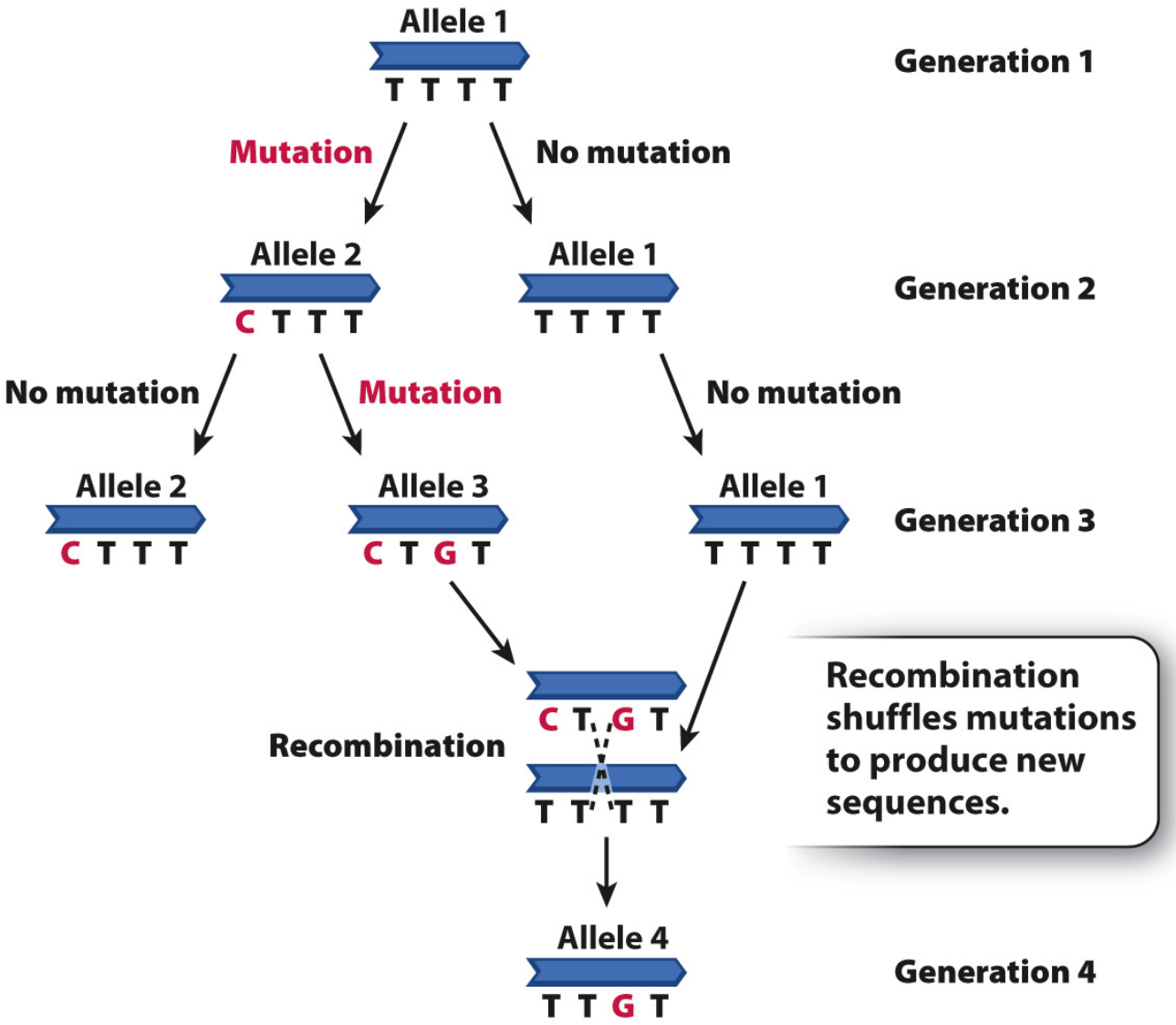


The field of **population genetics** was a part of the solution.

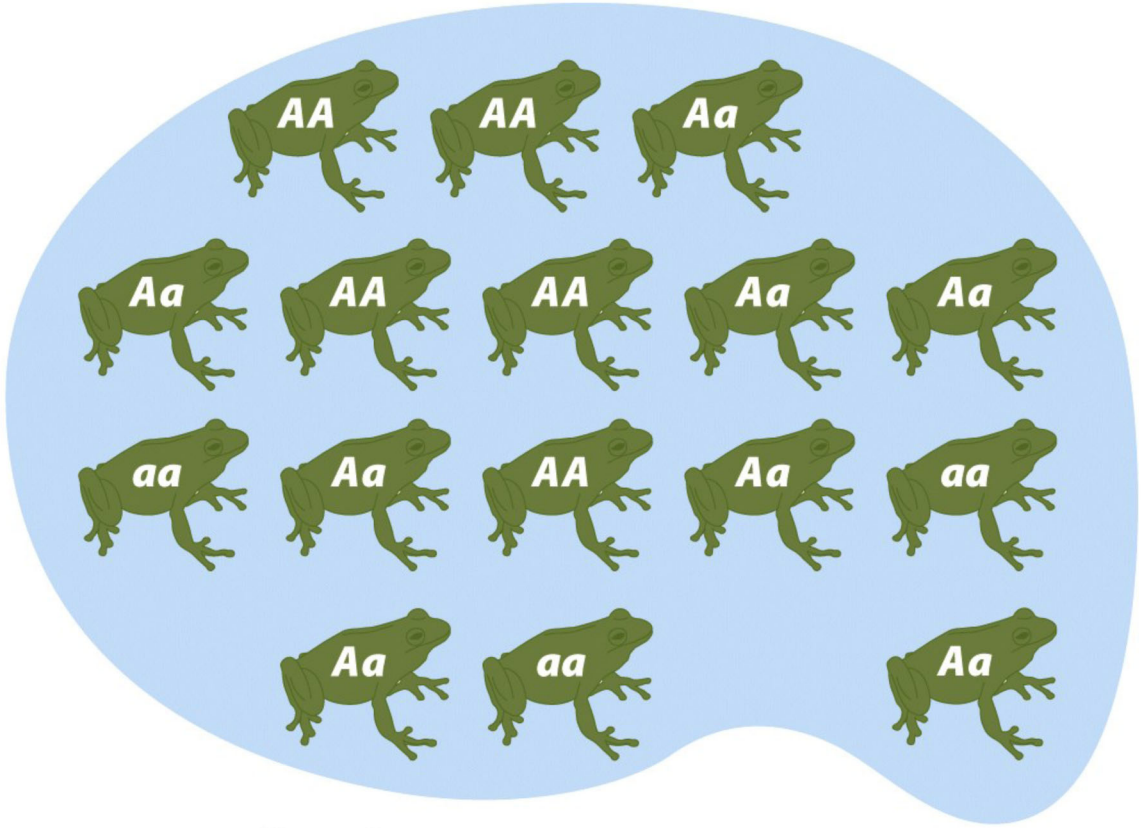
Population genetics: causes and consequences of genetic variation within a population

Population: a group of individuals that mate with one another to produce the next generation.

Mutation and recombination are the two sources of genetic variation.



Genetic polymorphism: Within a population, a single gene has more than one allele



AA	*Aa*	*aa*
5	8	3

Early population geneticists relied on observable traits and gel electrophoresis to measure variation.

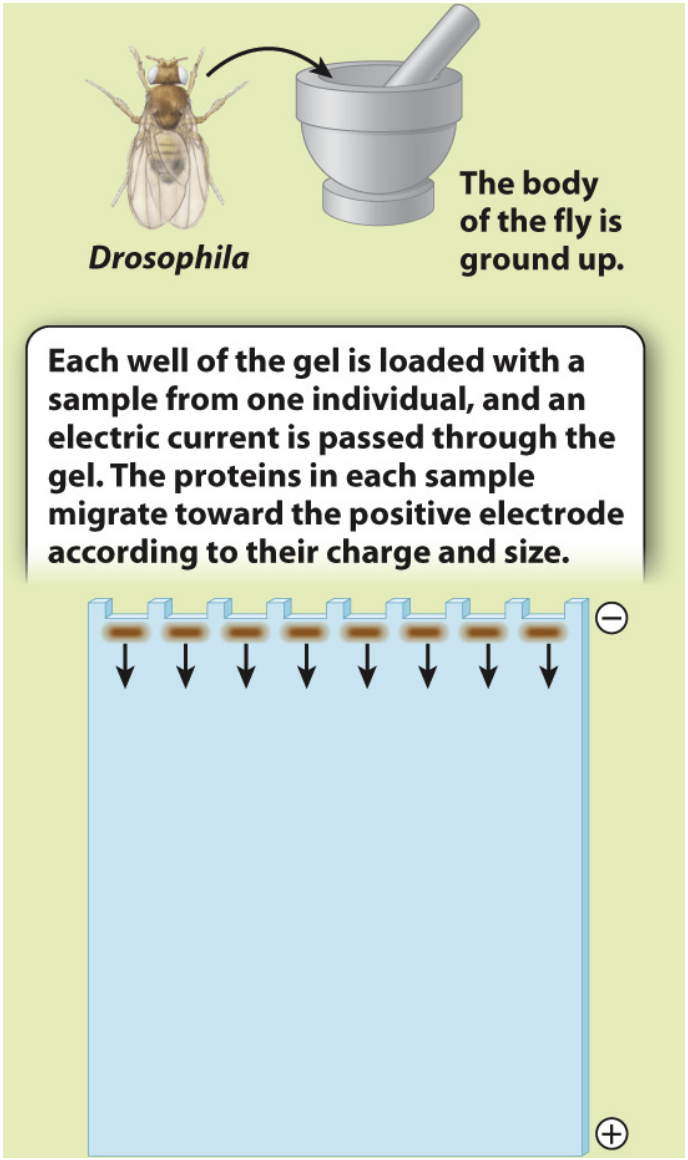
(top) © Biopix: G. Drange; (bottom) Howard Marsh/Shutterstock

TABLE 21.1 The ABO blood system.

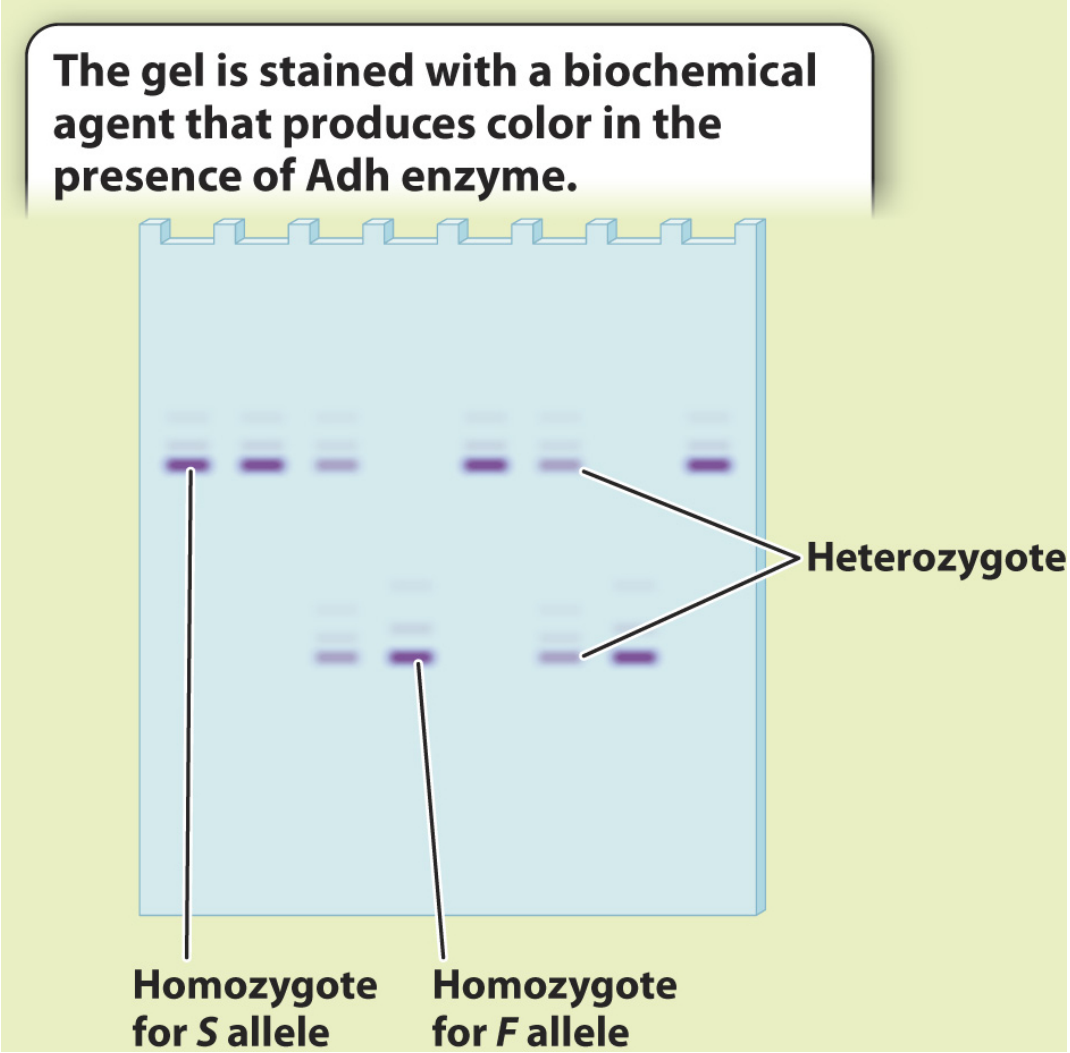
PHENOTYPE	GENOTYPE
A	*AA* or *AO*
B	*BB* or *BO*
AB	*AB*
O	*OO*

Single-gene variation. A genetic difference in color in the two-spot ladybug, *Adalia bipunctata,* results from variation in a single gene.

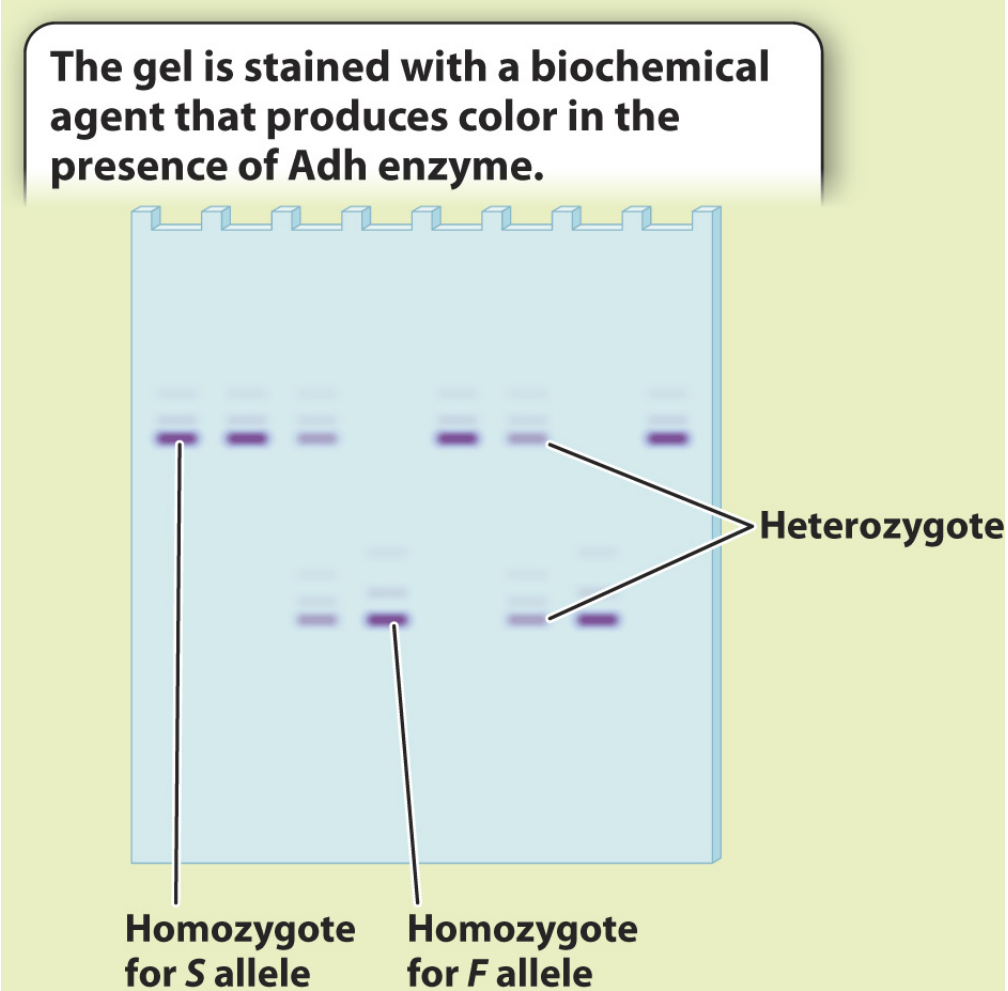
How is genetic variation measured?



How is genetic variation measured?

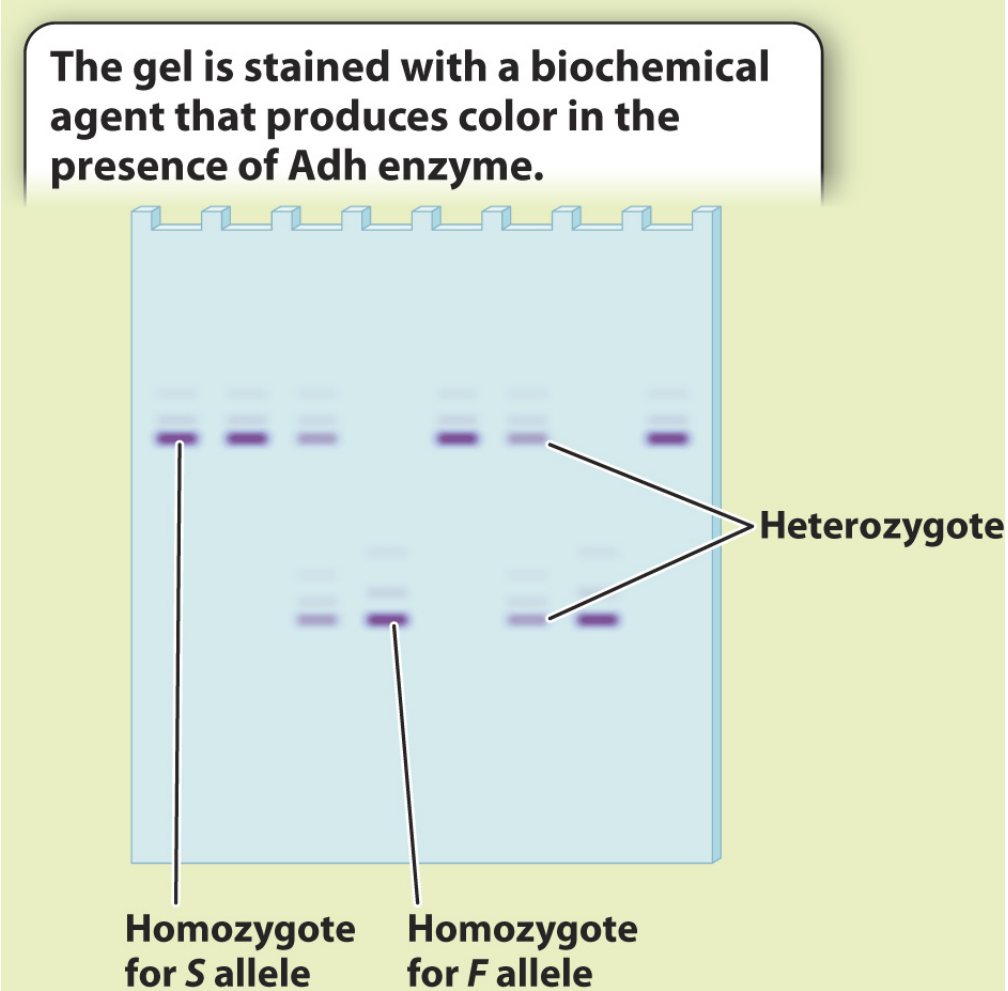


1. How many individuals does this gel represent?



A. 8 B. 10 C. 16 D. 20

2. How many alleles does this gel represent?



A. 8 B. 10 C. 16 D. 20

Profiling genetic variation in a population:

8 individuals represented
16 alleles (homozygotes are brighter)

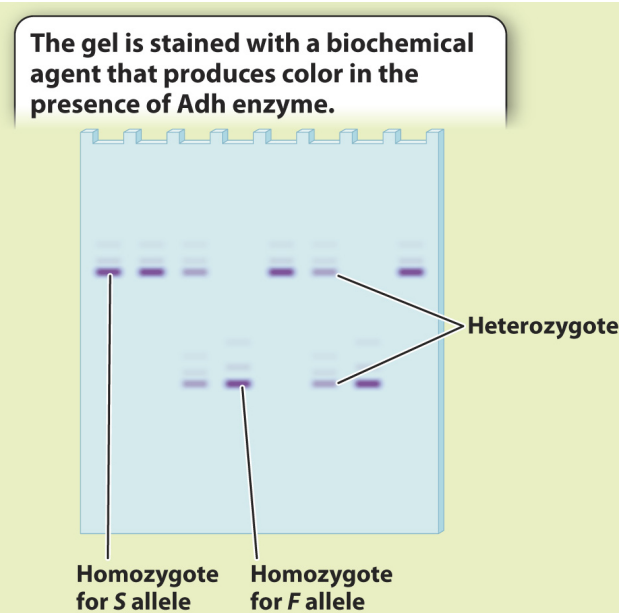
Number of *S* in the population = 2 × (number of *S* homozygotes) + (number of heterozygotes) = 8 + 2 = 10

Frequency of S = 10/16 = 5/8

Number of *F* in the population = 2 × (number of *F* homozygotes) + (number of heterozygotes) = 4 + 2 = 6

Frequency of F = 6/16 = 3/8

*Note that the two allele frequencies add to 1 (5/8 + 3/8 = 1)



Example of genetic polymorphism

The MN blood group in humans is caused by a single gene with two alleles, M and N.

Individuals within a population vary, because they can have one of three different genotypes:

Population	Genotype		
	MM	MN	NN
German	**0.297**	**0.507**	**0.196**

Within a species, different populations can have different allele frequencies

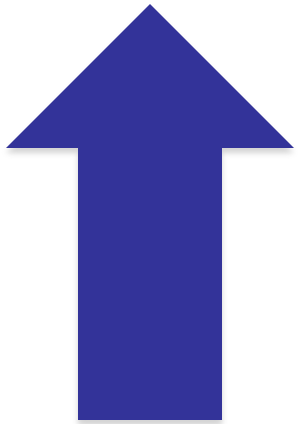
Population	Genotype		
	MM	MN	NN
German	**0.297**	**0.507**	**0.196**
Australian Aborigine	**0.024**	**0.304**	**0.672**

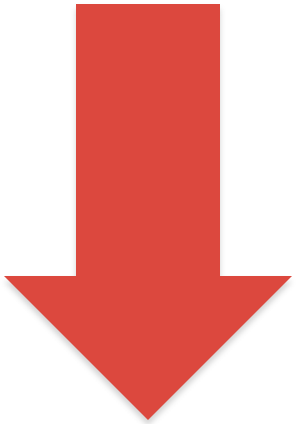
What determines how much polymorphism a population has?

Forces creating polymorphism

Mutation

Migration



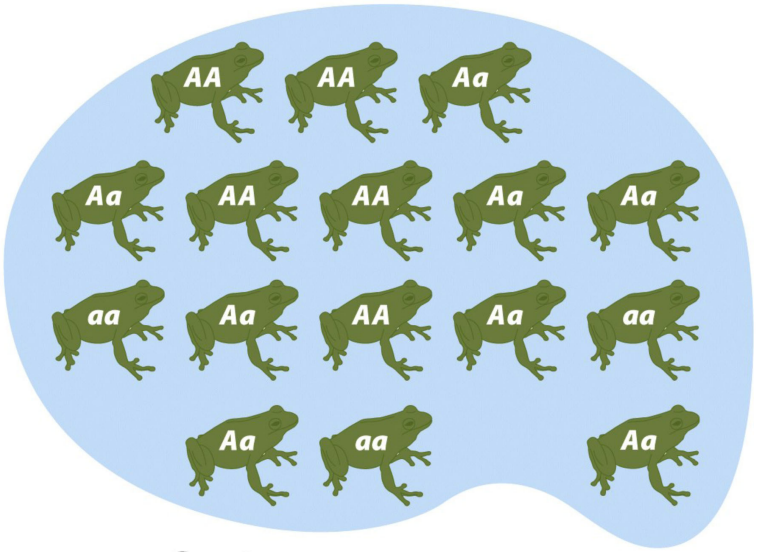


Forces destroying polymorphism

Selection

Drift

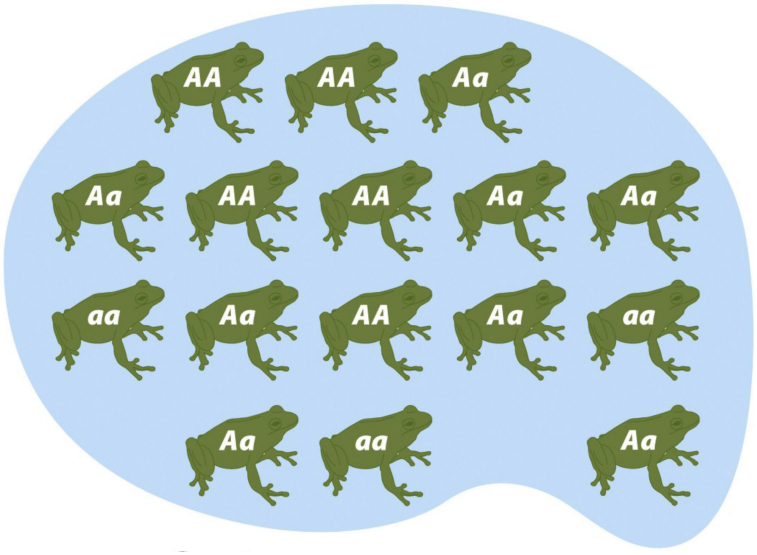
A population's **gene pool** is the sum total of all alleles in all breeding members



Population size (**N**) = 16

Number of alleles (**2N**) = 32

The **genotype frequencies** give the relative number of each genotype in the population



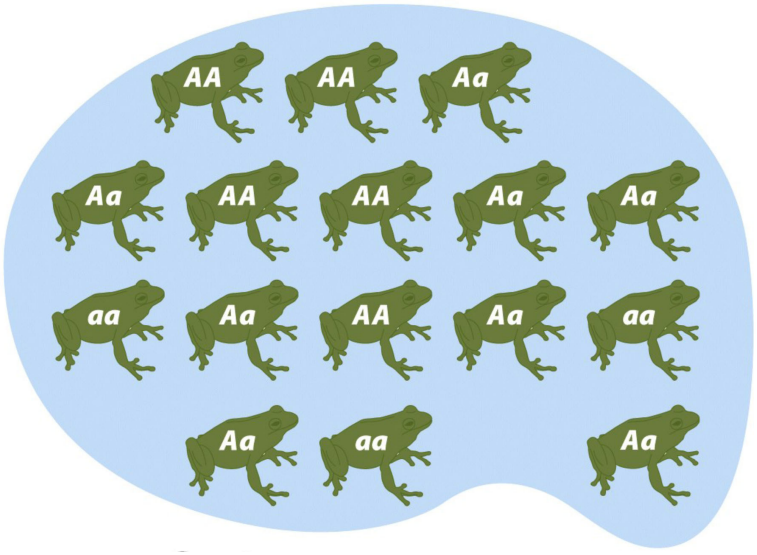
AA	***Aa***	***aa***
5	**8**	**3**

Frequency of ***AA*** = $\frac{5}{16}$ = 0.31

Frequency of ***Aa*** = $\frac{8}{16}$ = 0.50

Frequency of ***aa*** = $\frac{3}{16}$ = 0.19

The **allele frequencies** give the relative number of each allele in the population



AA	***Aa***	***aa***
5	**8**	**3**

$$p = \text{Frequency of } A = \frac{10+8}{32} = 0.56$$

$$q = \text{Frequency of } a = \frac{6+8}{32} = 0.44$$

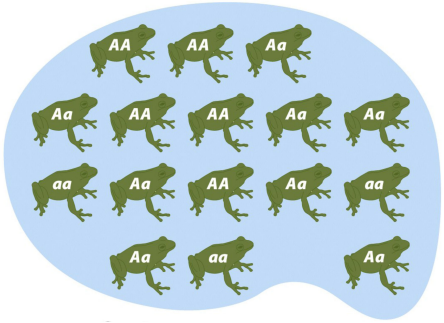
3. What would the genotype frequencies be?

AA	*Aa*	*aa*
4	**7**	**5**

Answers

	A.	B.
Freq. AA	0.25	0.20
Freq. Aa	0.44	0.65
Freq. aa	0.31	0.20

The **genotype frequencies** give the relative number of each genotype in the population



AA	***Aa***	***aa***
4	**7**	**5**

Frequency of ***AA*** = $\frac{4}{16}$ = 0.25

Frequency of ***Aa*** = $\frac{7}{16}$ = 0.44

Frequency of ***aa*** = $\frac{5}{16}$ = 0.31

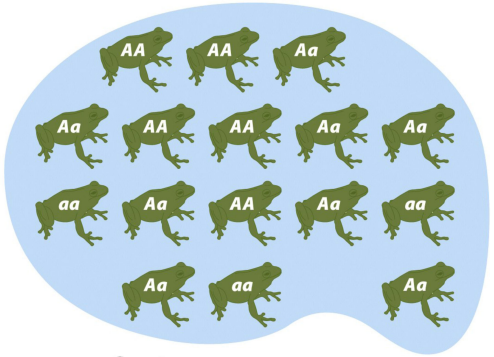
4. What would the allele frequencies be?

AA	***Aa***	***aa***
4	**7**	**5**

p = A
q = a

A. p = 0.56 q = 0.44
B. p = 0.53 q = 0.47
C. p = 0.60 q = 0.40
D. p = 0.47 q = 0.53

The **allele frequencies** give the relative number of each allele in the population



AA	***Aa***	***aa***
4	**7**	**5**

There are 16 individuals = 32 alleles

p =Frequency of ***A*** = $\frac{8+7}{32}$ = 0.47

q = Frequency of ***a*** = $\frac{10+7}{32}$ = 0.53

How polymorphic is a locus?

- This is given by its **heterozygosity**: the probability that two alleles chosen at random from the population are different.
- The higher this number, the greater the polymorphism.
  - One very common allele = low heterozygosity
  - Many common alleles = high heterozygosity.

Calculating allele frequencies from genotype frequencies

AA	Aa	aa
0.04	**0.32**	**0.64**

$$p = \text{freq}(A) = \text{freq}(AA) + \tfrac{1}{2}\,\text{freq}(Aa)$$

$$= 0.04 + \tfrac{1}{2}\,(0.32) = \mathbf{0.2}$$

$$q = \text{freq}(a) = \text{freq}(aa) + \tfrac{1}{2}\,\text{freq}(Aa)$$

$$= 0.64 + \tfrac{1}{2}\,(0.32) = \mathbf{0.8}$$

*Note: If there are only two alleles, **p** + **q** always equals 1.0*

If allele frequencies are known, can genotype frequencies be estimated?

Yes, if certain assumptions are made:

- **Large population size.**
- **All genotypes have equal survival and reproduction.**
- **Mating is random with respect to genotype.**
- **(essentially all of this above means that evolution is NOT happening)**

If these assumptions hold, the population is at **Hardy-Weinberg equilibrium**

1. **Genotype frequencies remain the same generation after generation.**
2. **Genotype frequencies can be easily predicted from allele frequencies**.

Hardy-Weinberg equilibrium frequencies are given by this formula:

$$(p+q)^2 = p^2 + 2pq + q^2 = 1$$

Where ***p*** is the frequency of allele 1 and ***q*** the frequency of allele 2.

For two alleles, the equilibrium frequencies are:

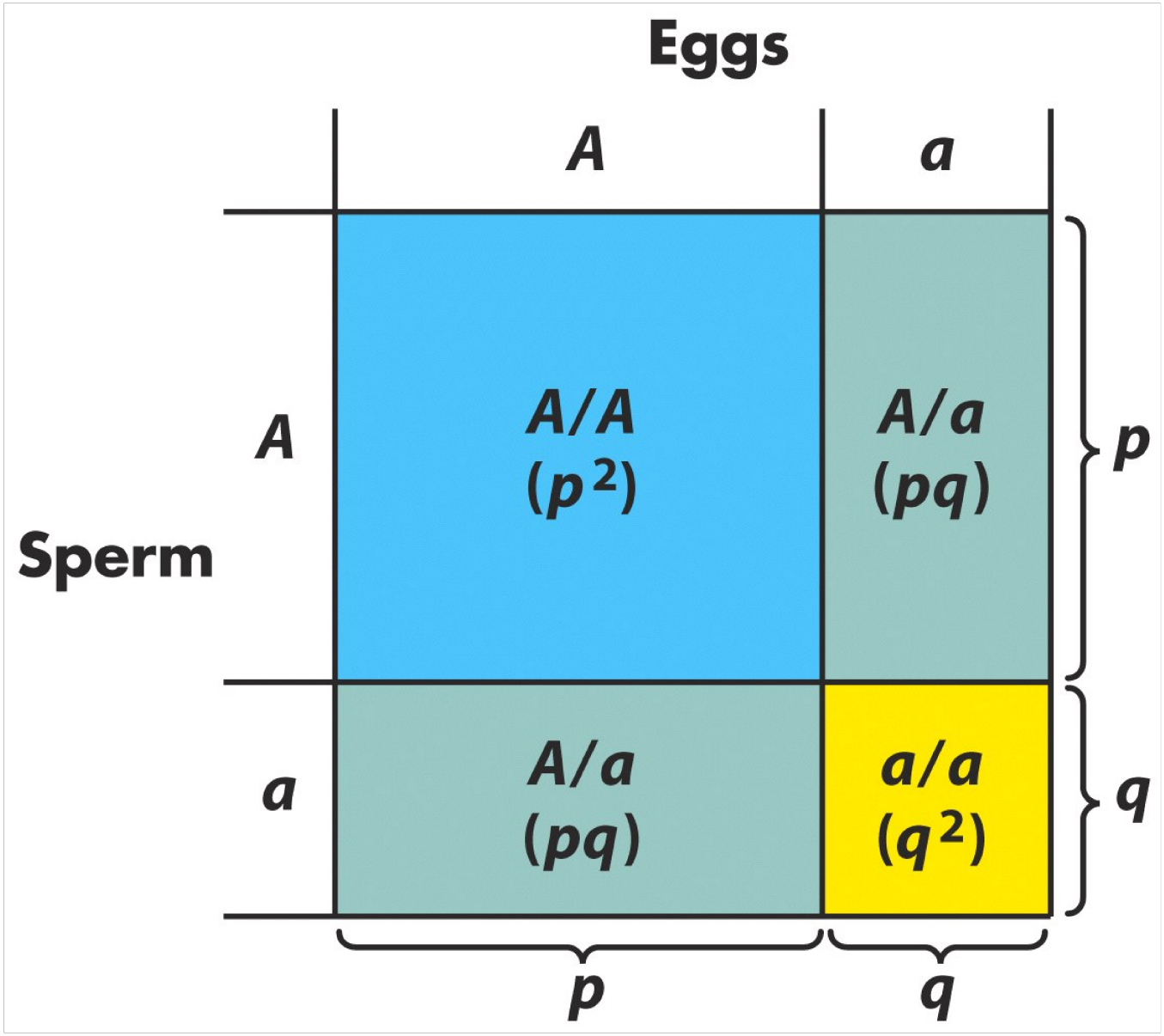
AA	Aa	aa
p^2	$2pq$	q^2

If Hardy-Weinberg assumptions hold…

Each offspring can be viewed as the combination of a random egg and a random sperm drawn from the gene pool

Egg	Sperm	Offspring	Probability	
A	*A*	*AA*	$p \times p$	p^2
A	*a*	*Aa*	$p \times q$	$2pq$
a	*A*	*Aa*	$q \times p$	
a	*a*	*aa*	$q \times q$	q^2

Graphic representation of Hardy-Weinberg equilibrium



The Hardy–Weinberg equilibrium is the starting point for population genetic analysis.

If we can find a population whose allele or genome frequencies are not in Hardy-Weinberg equilibrium, we can infer that evolution has occurred.

If a population is at Hardy-Weinberg equilibrium, carrier frequency can be estimated from disease frequency

Phenotype:	**Normal**		**X**
Genotype:	***AA***	***Aa***	***aa***
Phenotype frequency:	**0.99**		**0.01**

p is the frequency of ***A***, and ***q*** is the frequency of ***a***

$$Freq(aa) = q^2 = 0.01$$

$$q = \sqrt{0.01} = 0.1$$

$$p = 1 - q = 0.9$$

$$Freq(Aa) = 2pq = 2(0.9)(0.1) = 0.18$$

For a rare allele, how common are homozygotes compared to heterozygotes?

From previous question:

Genotype:	***Aa***	***aa***
Frequency:	0.18	0.01

Proportion of carriers who are heterozygotes:

$$\frac{0.18}{0.18+0.01}=0.95$$

Key point: For a rare recessive allele, heterozygotes are far more common than homozygotes.

Most populations are not at Hardy-Weinberg equilibrium

- Why not?
- Assumptions are often not met:
  - Mating is often non-random.
  - Many populations are quite small.
  - Not all alleles are equally viable.

Causes of departure from Hardy-Weinberg

- **Non-random mating**
- Random changes in allele frequency (also known as **drift**)
- Differences in allele effects on fitness (that is, **selection**)

Departures from Hardy-Weinberg:
Non-random mating

Random mating means that individuals do not choose their mates *on the basis of a particular heritable character*

Nonrandom mating:

Positive assortative mating: Bias toward phenotypically similar mates

Negative assortative mating: Bias toward phenotypically different mates

Inbreeding: Bias toward mating with relatives.

Departures from Hardy-Weinberg:
Drift

First, an important point about mutations

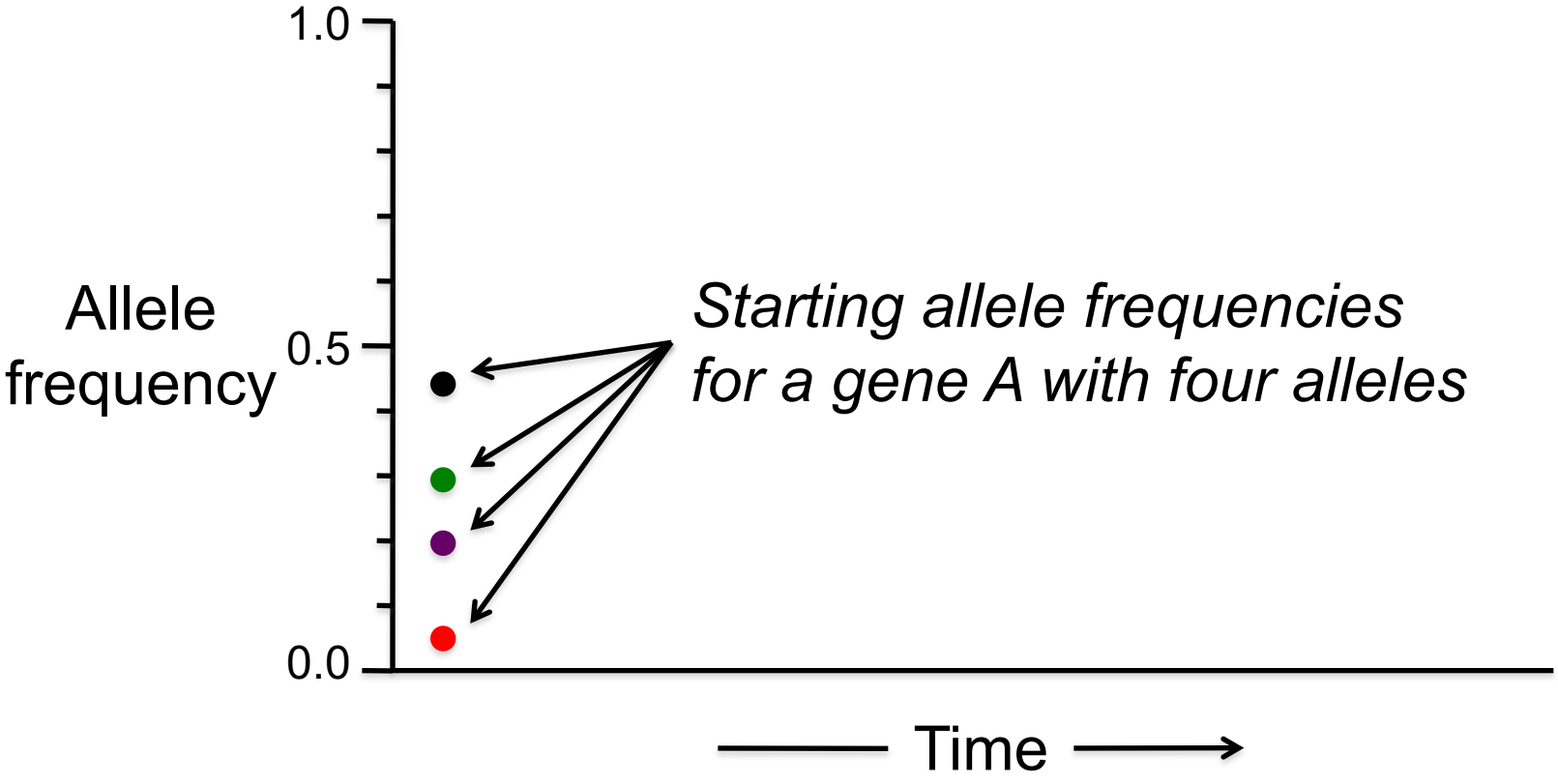
Most mutations are **neutral**

This means that there is no difference in fitness (survival and reproduction) between individuals with different alleles.

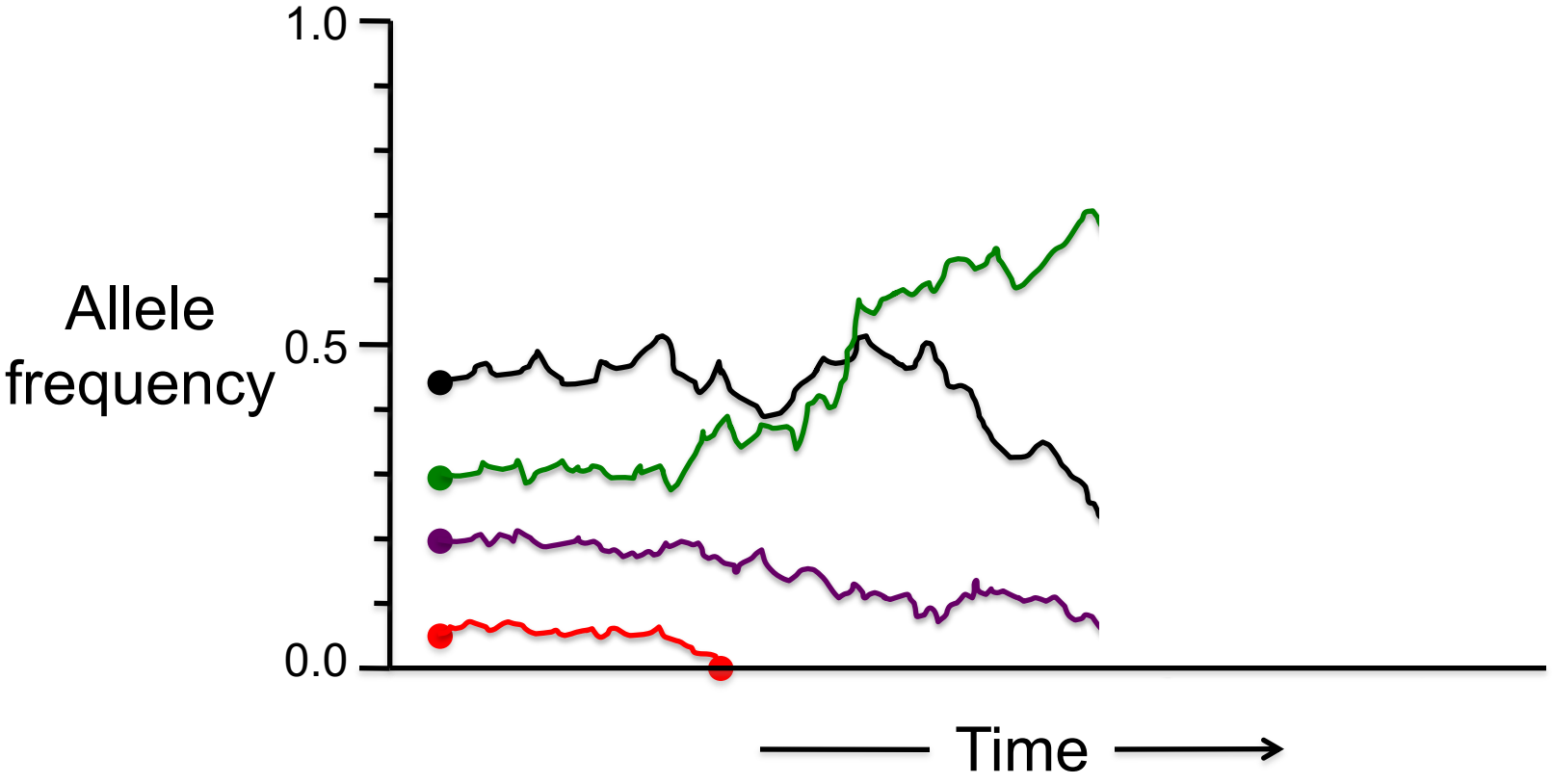
Examples of mutations causing neutral polymorphism

- Sequence changes in intergenic regions or introns
- Synonymous changes in protein coding sequences
- Nonsynonymous changes that replace one amino acid with a chemically similar one.

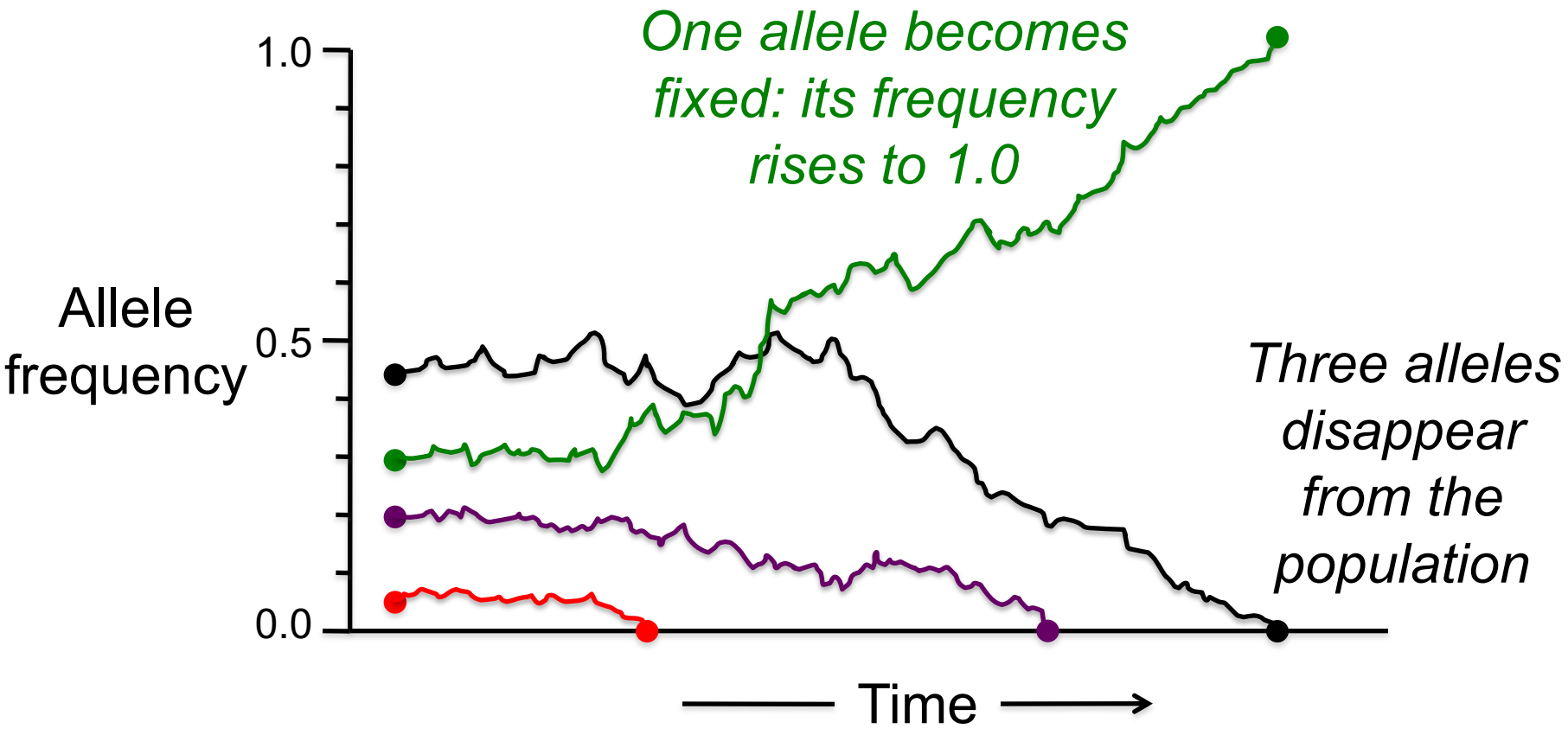
Genetic drift: Allele frequency can change over time due to random fluctuations



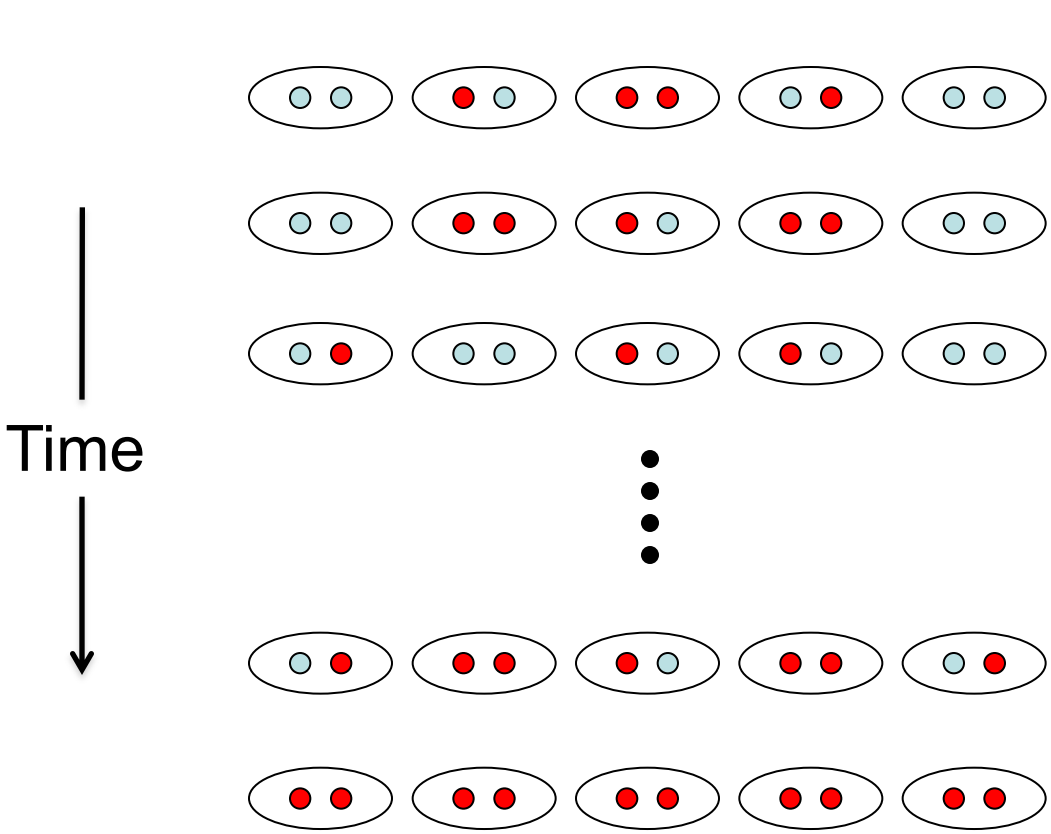
Allele frequencies fluctuate randomly over time



Genetic drift can eventually lead to the loss of all but one allele of a gene, termed "**fixation**"



The probability that a particular allele is eventually fixed by genetic drift equals its frequency in the population



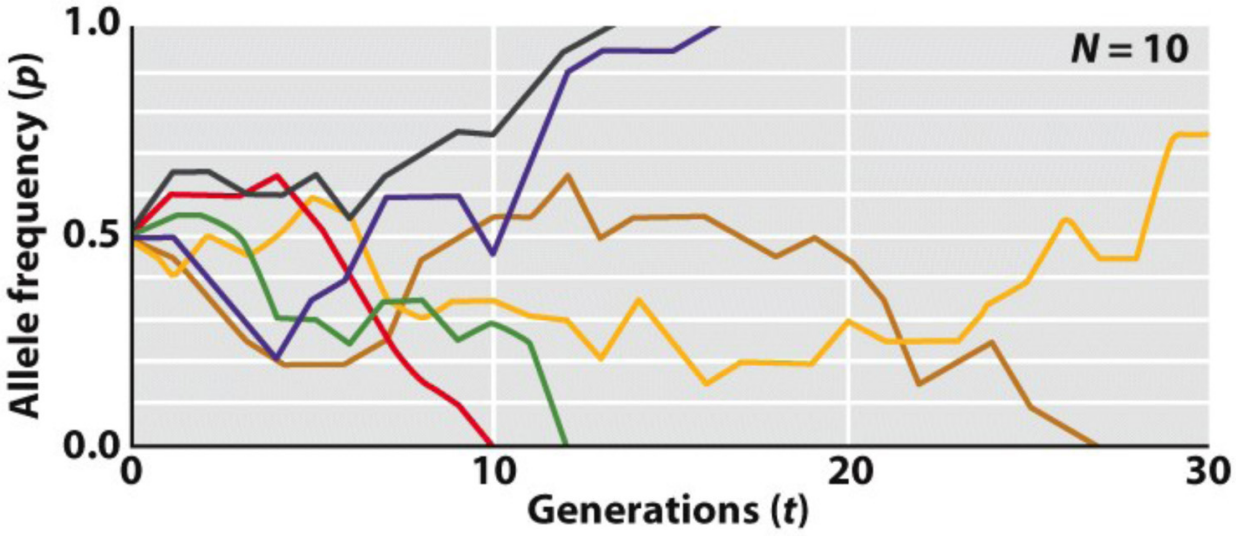
N individuals (**2N** alleles)

p = 0.4 (red)

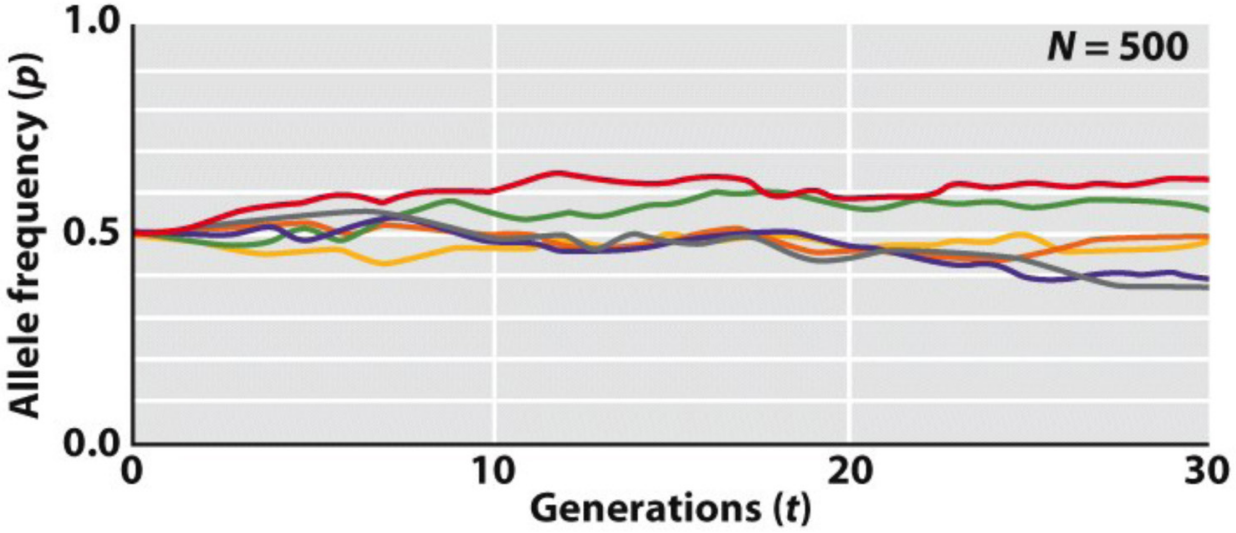
q = 0.6 (blue)

The red allele has become fixed. The probability of this outcome is 0.4.

Random genetic drift is weakest in large populations

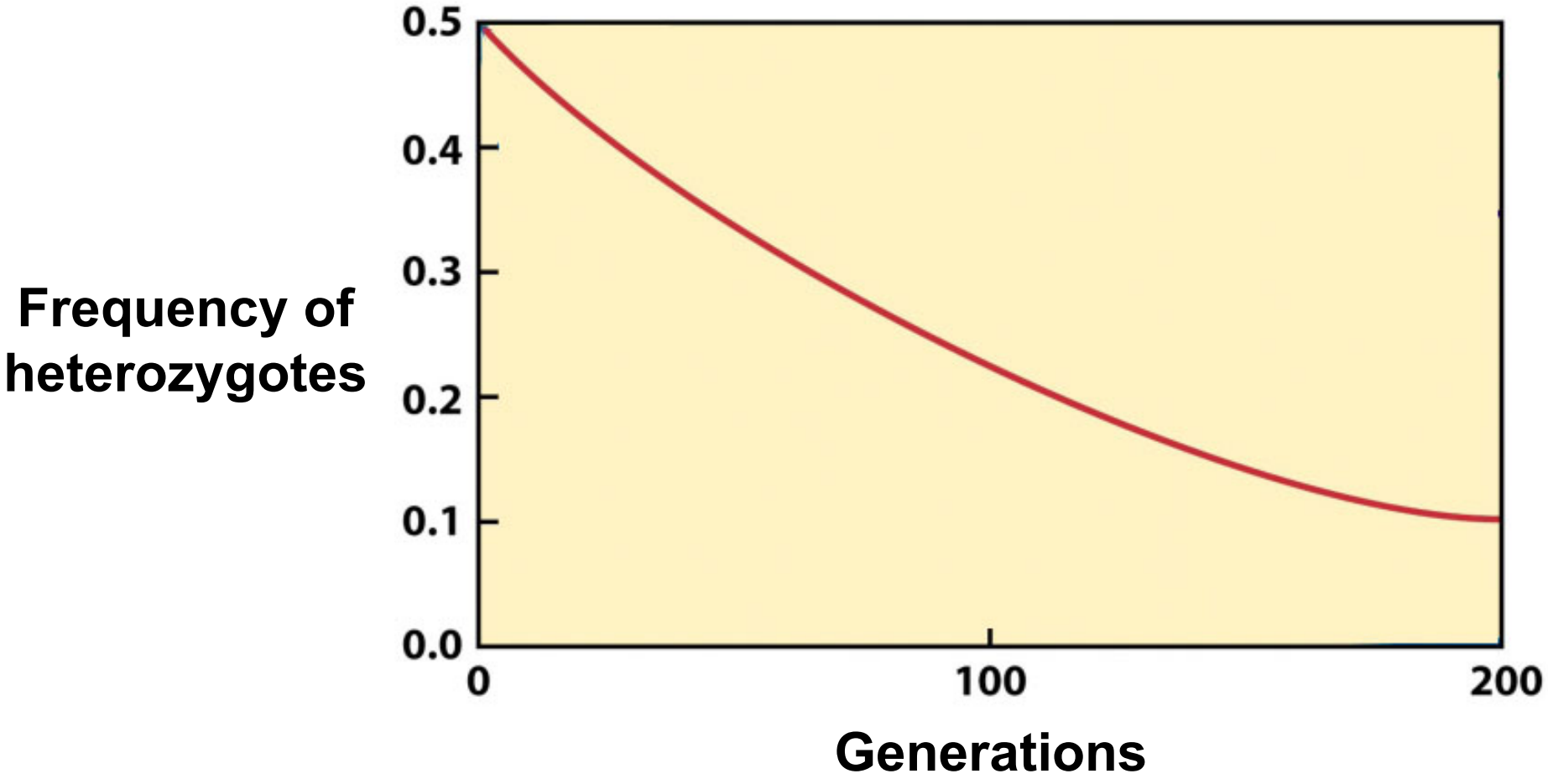


Population = 10



Population = 500

Over time, genetic drift reduces the amount of allelic variation in a population

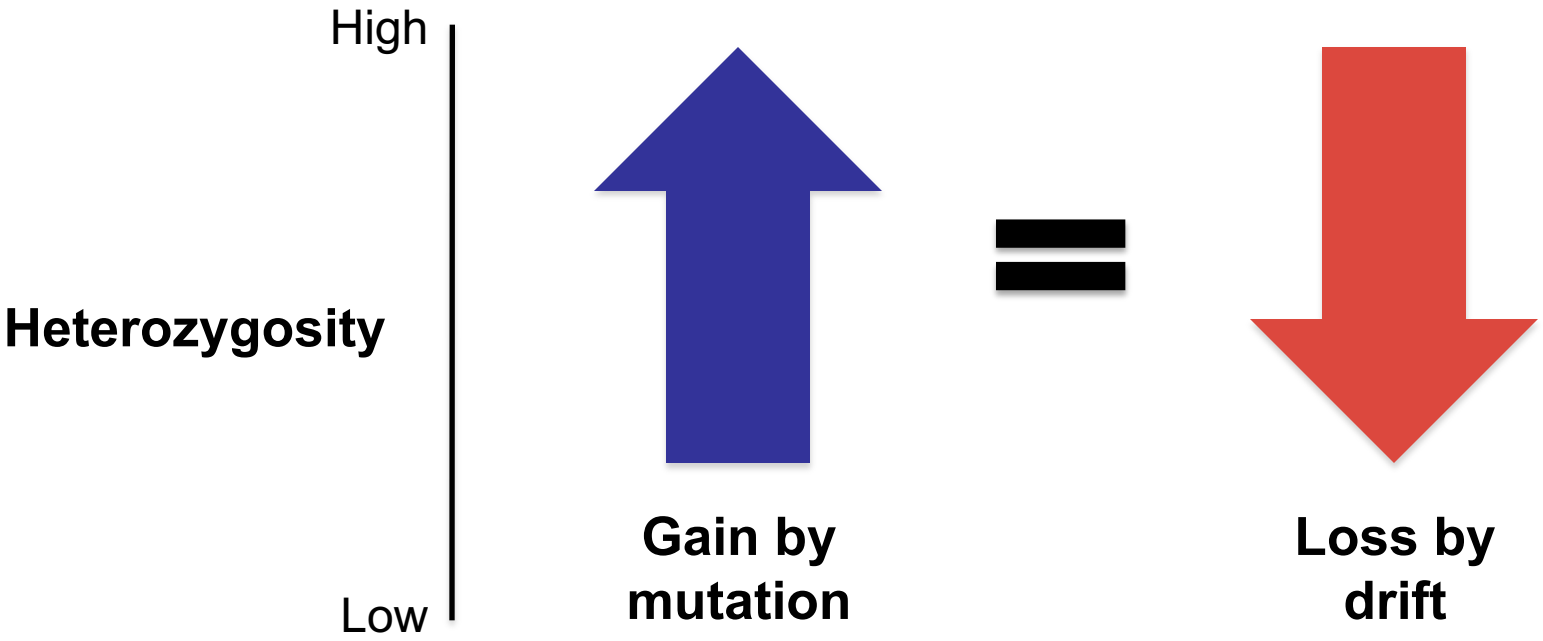


Drift is a one way street: it removes alleles, but it does not restore them

Ultimately, drift will *always* lead to fixation of only one allele.

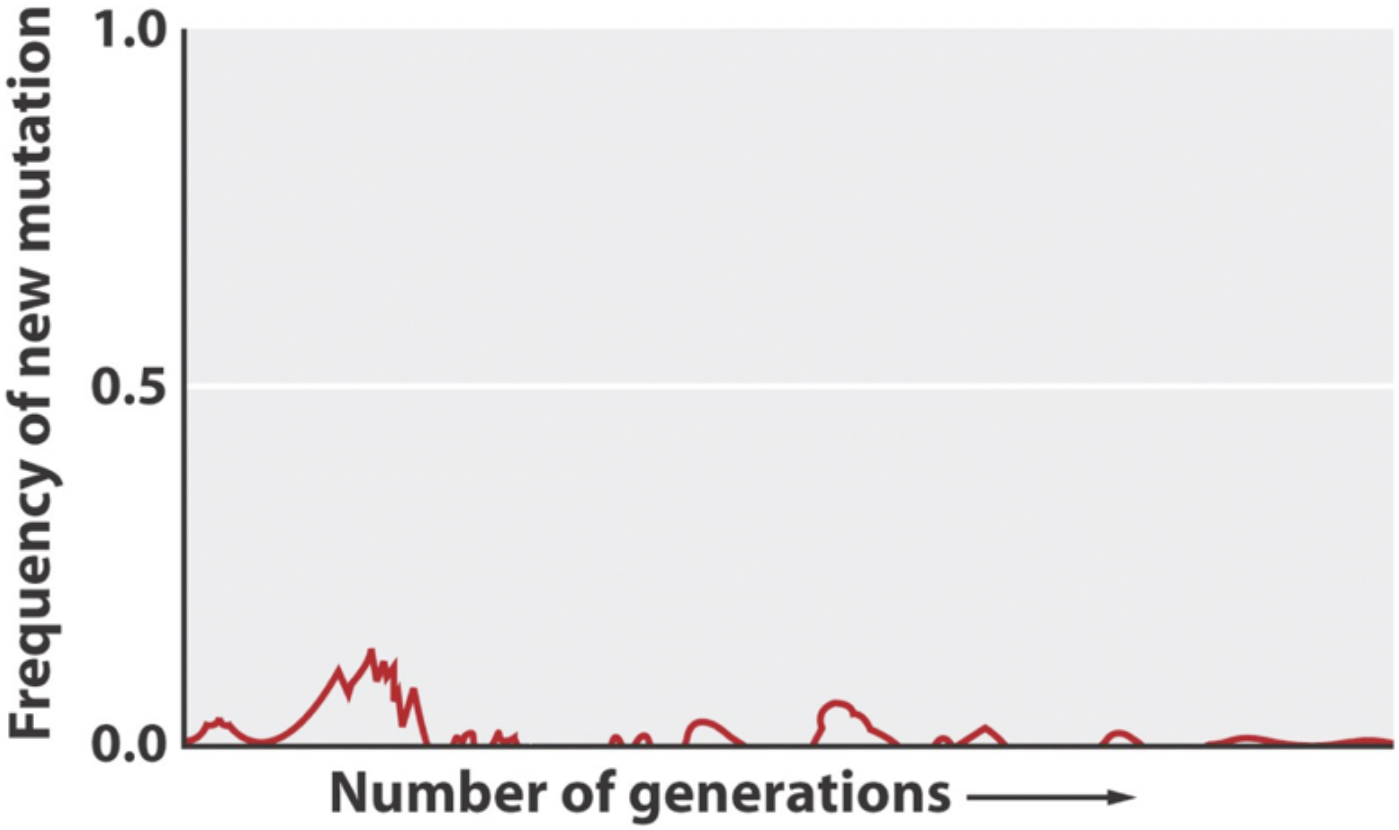
Why doesn't drift eliminate all neutral polymorphism?

The answer is ***mutation***

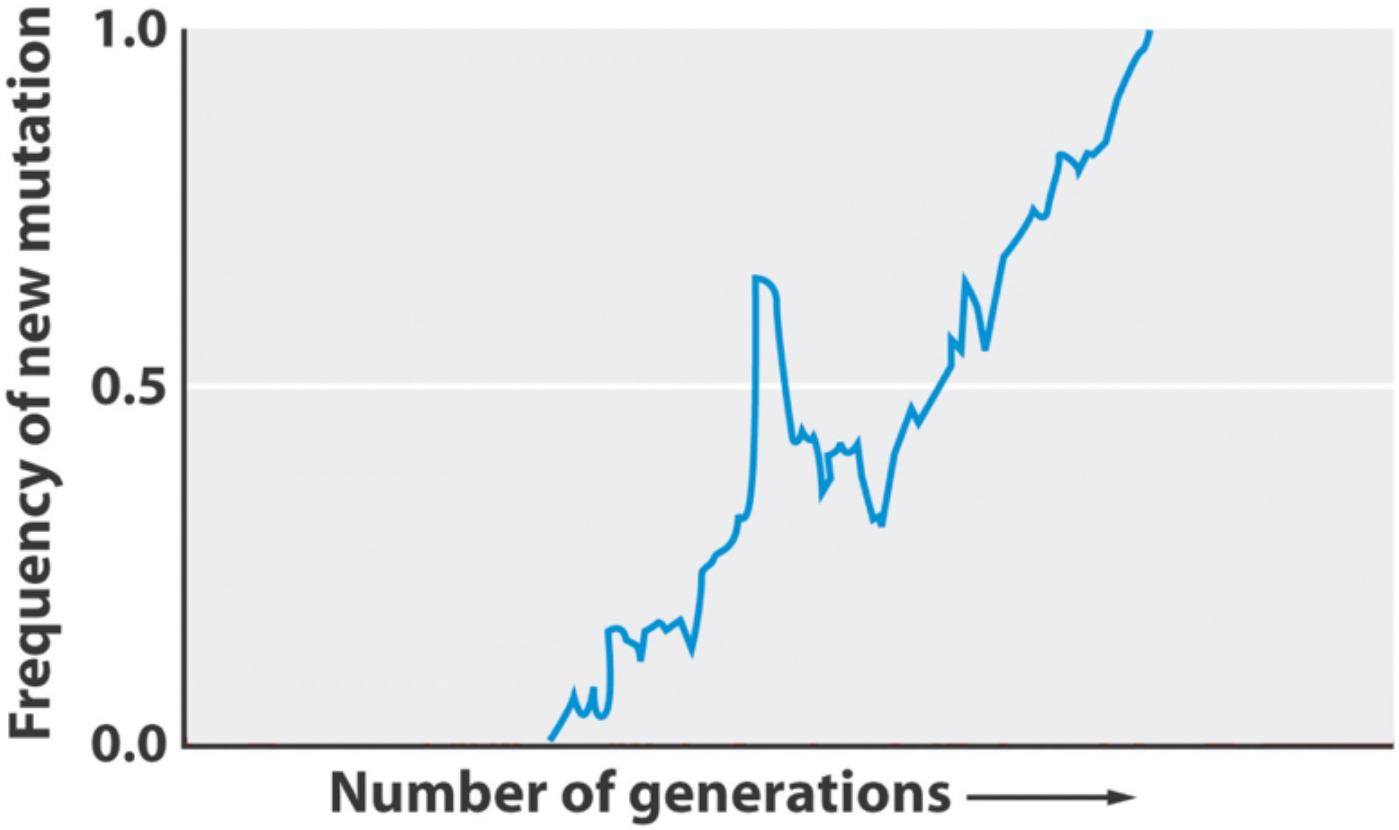


Heterozygosity ends up at a level where mutation adds alleles as fast as drift removes them.

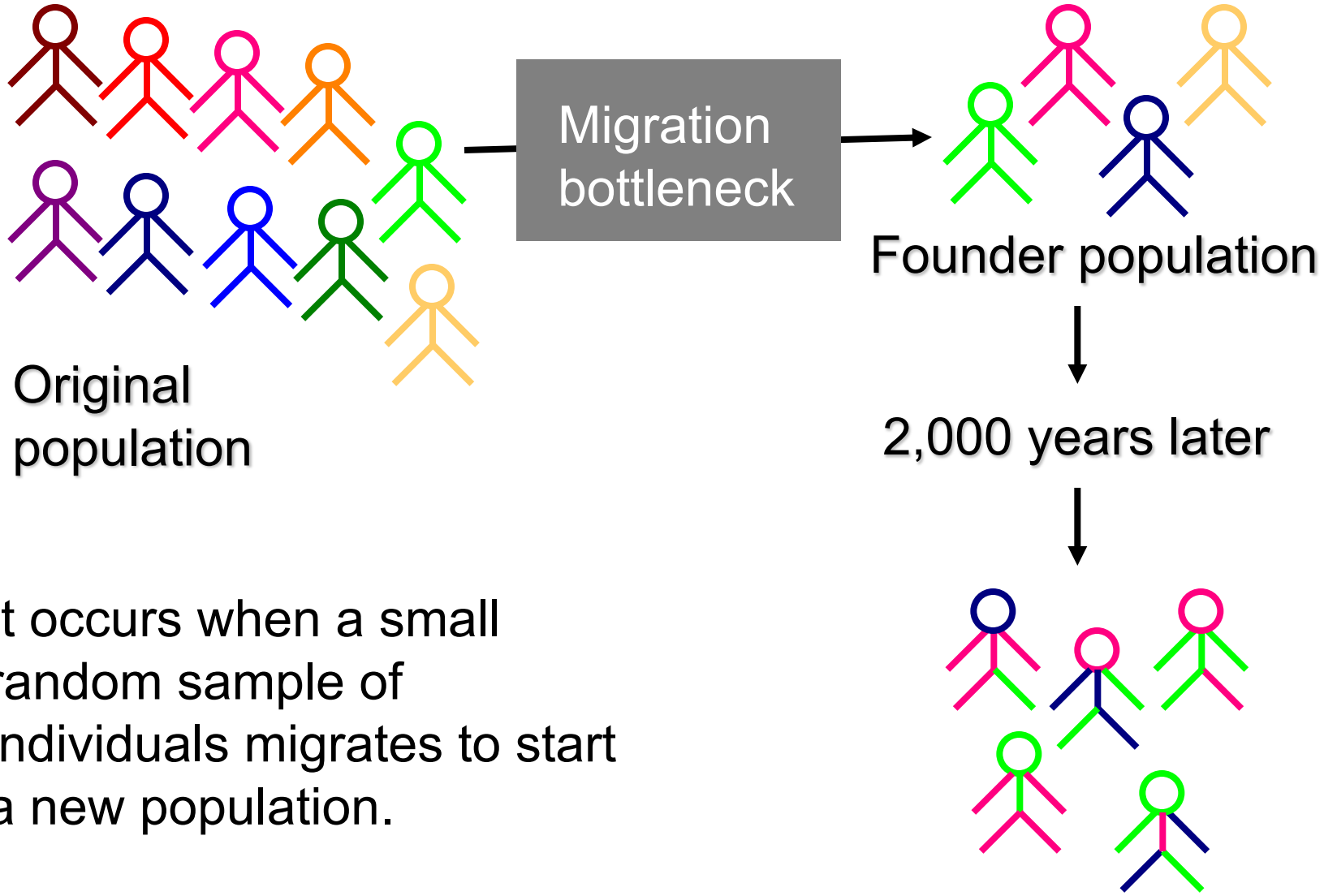
All new mutations start off rare in the population, and most soon disappear



Rarely, a new mutation spreads through the population and becomes fixed



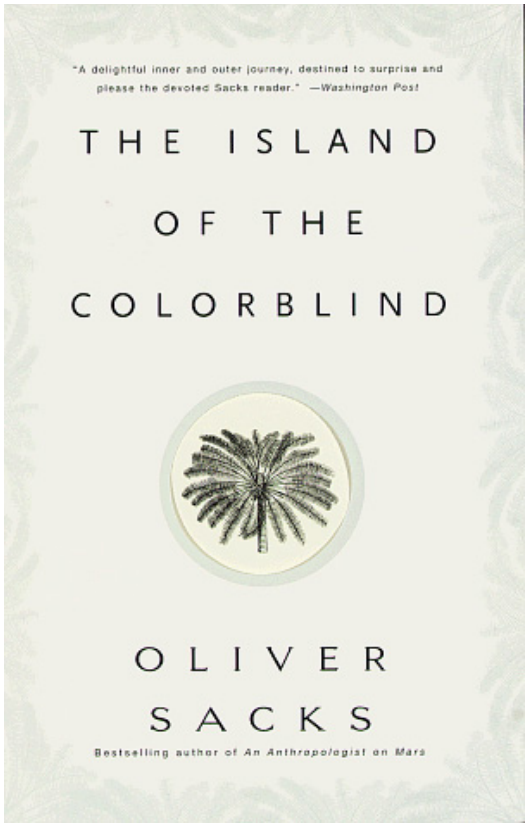
The **founder effect** is a particular kind of genetic drift



It occurs when a small random sample of individuals migrates to start a new population.

The founder effect can also lead to large differences in allele frequencies between populations

An example of the founder effect: the island of the colorblind



Residents
Of the island of Pingelap
With achromatopsia

Departures from Hardy-Weinberg:
Selection

Hardy-Weinberg equilibrium assumes that all genotypes have equal **fitness**

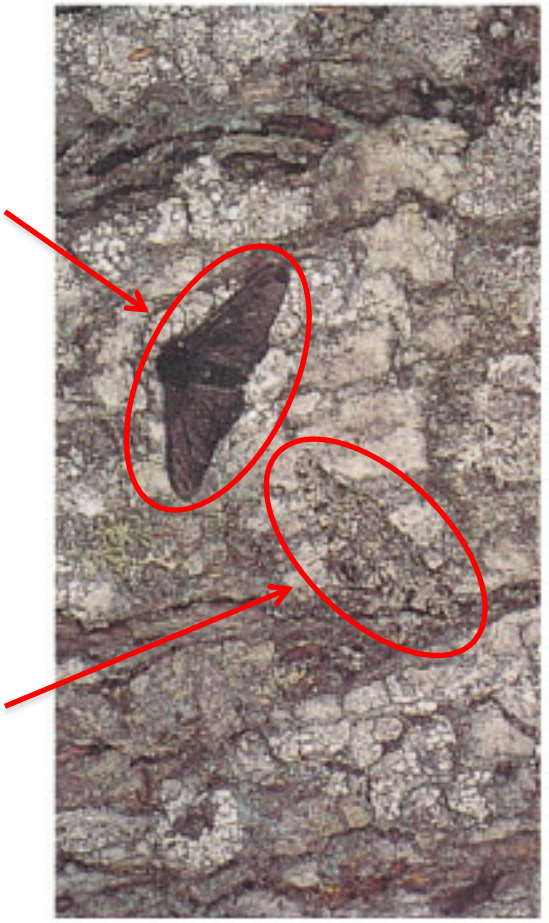
Genotype:	**AA**	**Aa**	**aa**
Fitness:	1	1	1

Fitness here means the average number of offspring produced by an individual.

If genotypes have unequal fitness, then allele frequencies will change over time (**natural selection**)

Color polymorphism in the moth *Biston betularia* is due to a single gene with two alleles.

The light morph has higher fitness when trees are clean and covered with lichens.

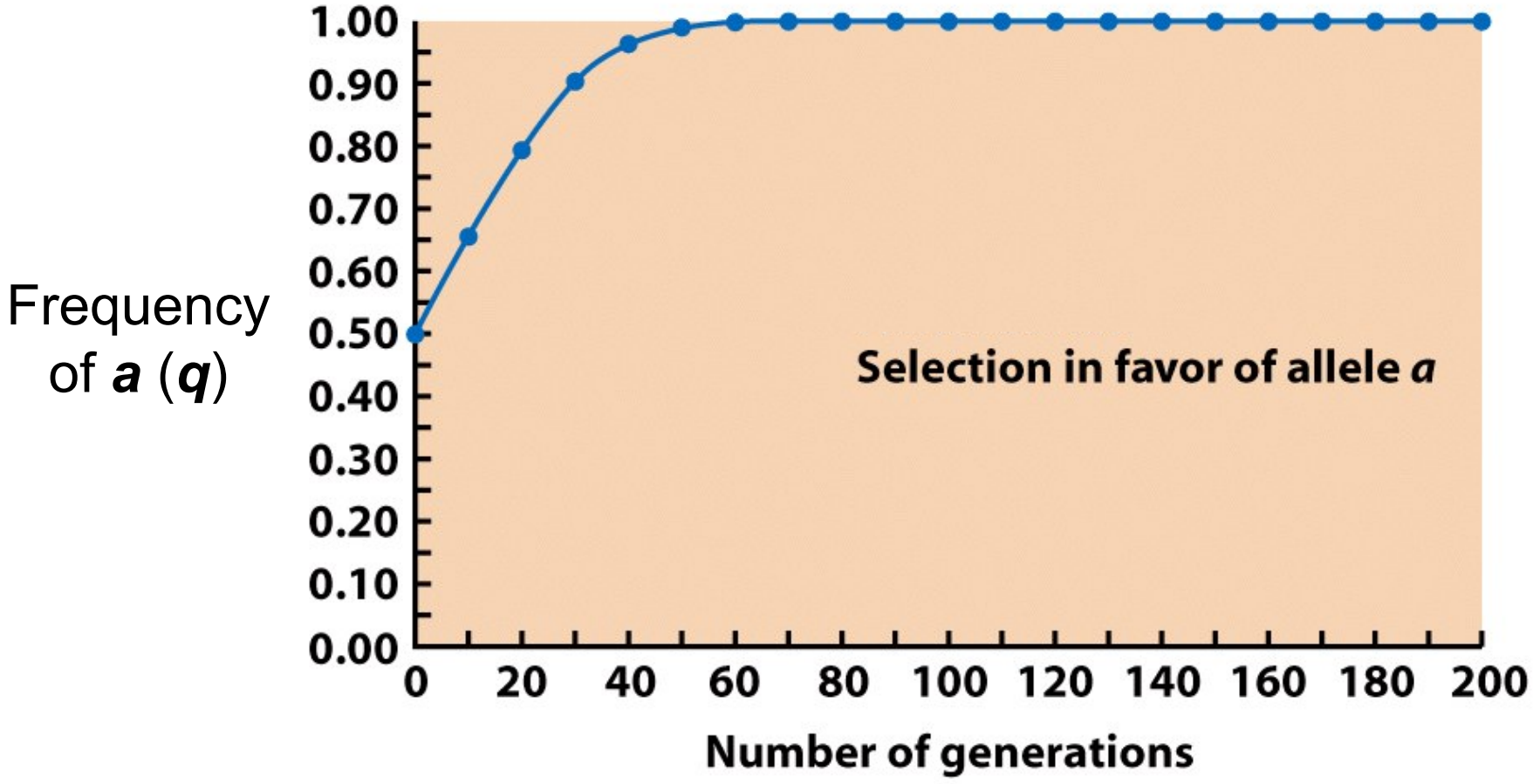


Lower fitness of the inferior genotype is measured by the **selection coefficient** (***s***)

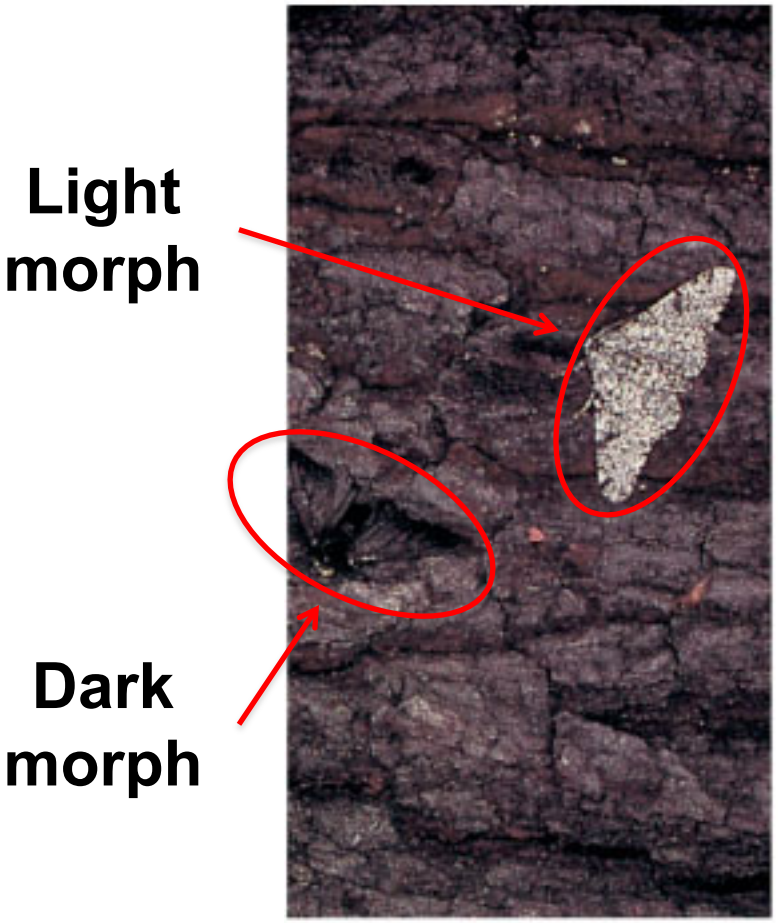
Genotype:	AA	Aa	aa
Phenotype:	Dark	Dark	Light
Fitness in clean forest:	1-s	1-s	1

The fitness of the inferior genotype (dark) is less than the fitness of the superior genotype (light).
The value of ***s*** shows how much less.

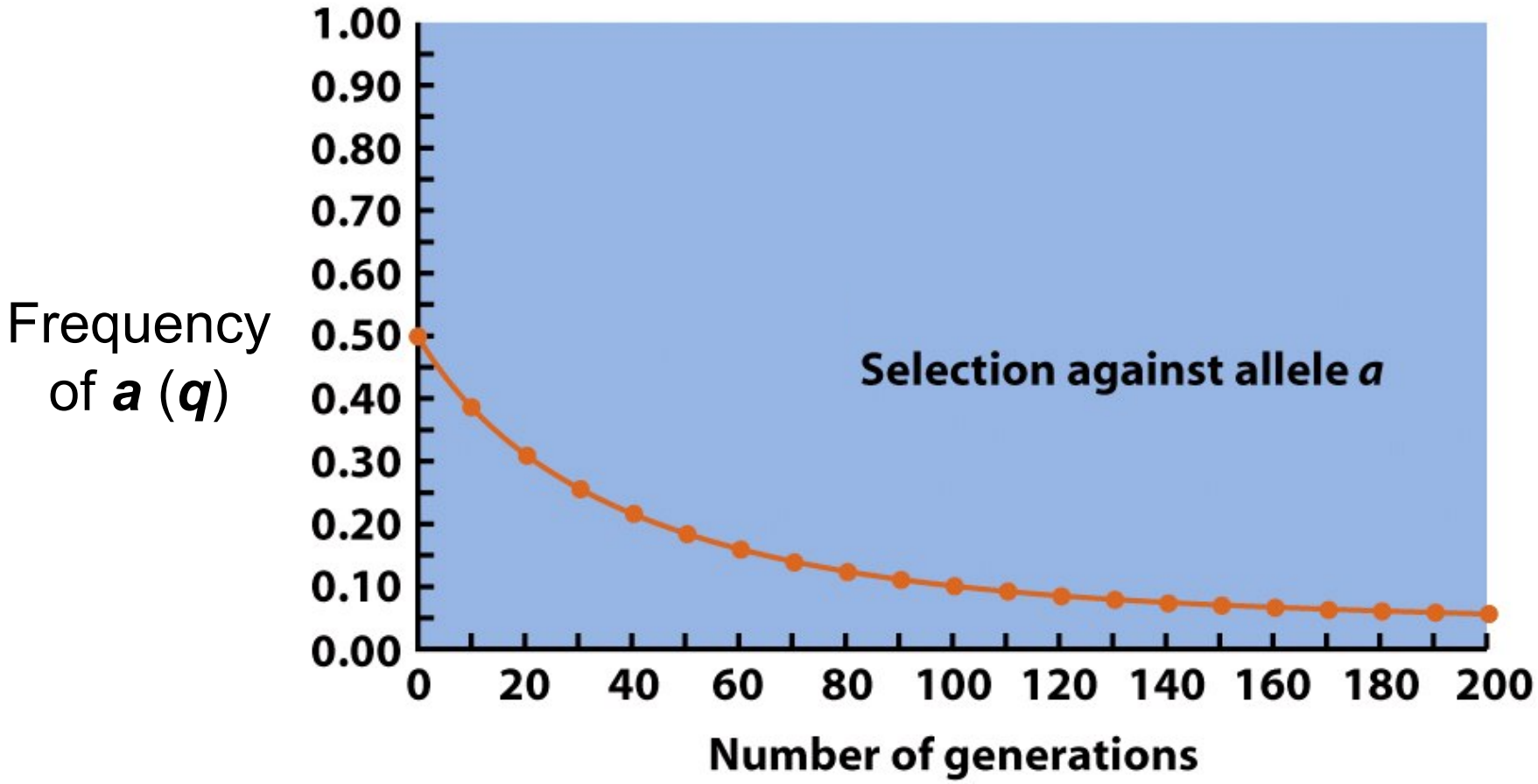
Selection against a deleterious dominant allele eventually eliminates it from the population



In polluted regions, where trees are dark, the dark morph has higher fitness than the light morph

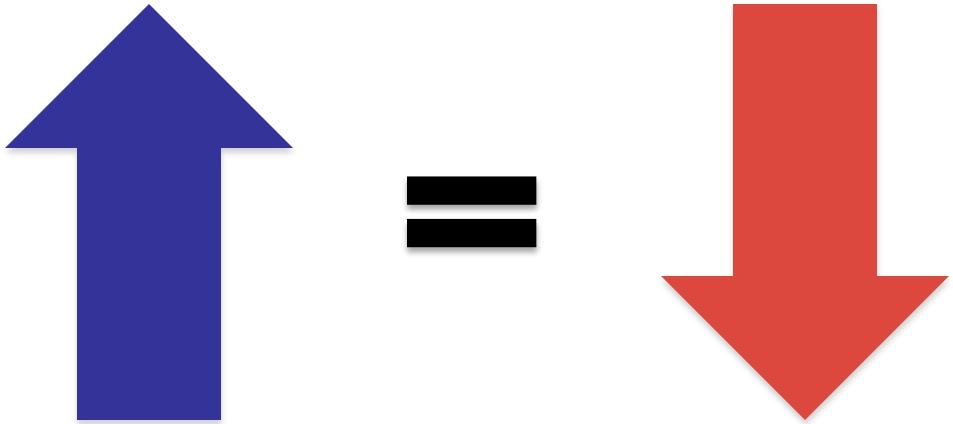


Selection in favor of a dominant allele gradually reduces the frequency of the recessive allele, ***but does not eliminate it***



Why doesn't selection eliminate all deleterious alleles?

Mutation



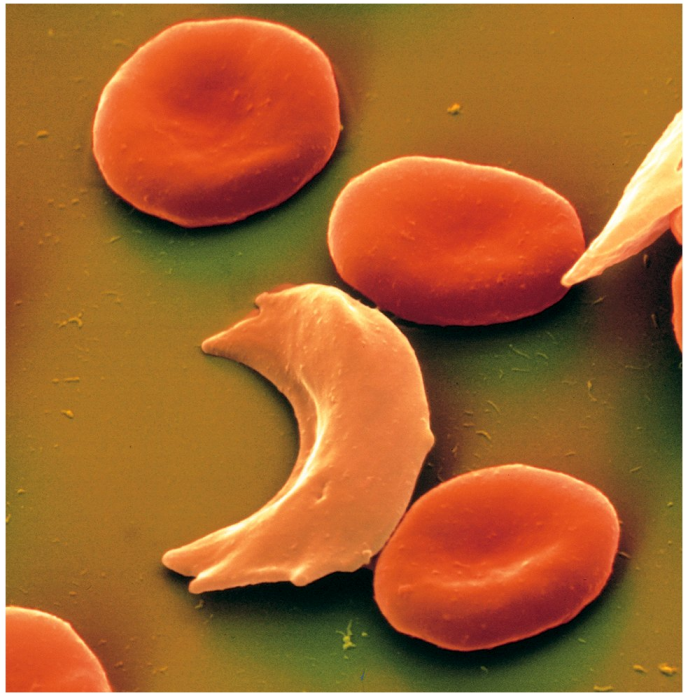
Gain by mutation

Loss by selection

Deleterious recessive alleles end up at a frequency where gain by mutation balances loss by selection

In some cases, selection can ***maintain*** polymorphism rather than reduce it

Sickle cell anemia is caused by a mutation in the beta-hemoglobin gene



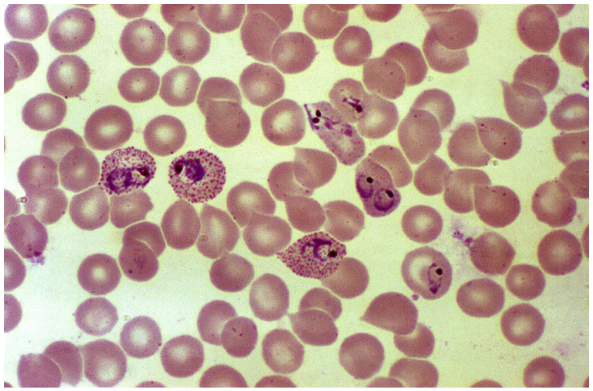
Homozygous mutant individuals have misshaped red blood cells leading to circulatory problems and shortened lifespan.

Hb^AHb^A	Hb^AHb^S	Hb^SHb^S
Normal	Sickle cell trait	Sickle cell anemia

Generally normal phenotype, but some cells may be partially sickle-shaped

Despite high mortality of homozygotes, *Hb*^{S} allele is very common in *some* populations

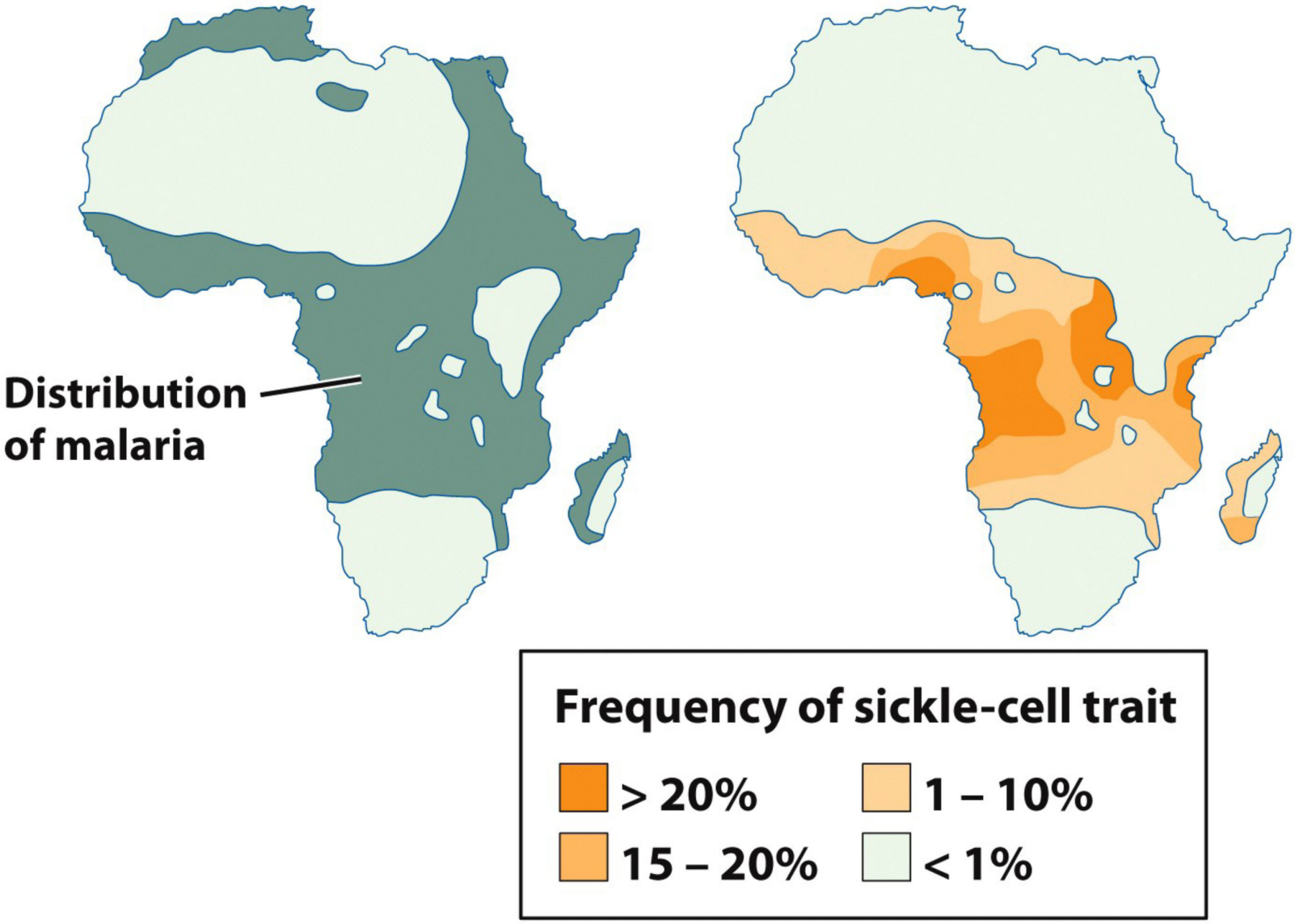
Region	Allele frequency of Hb^S
Highland Kenya	< 1%
Lowland Kenya	10 – 40%



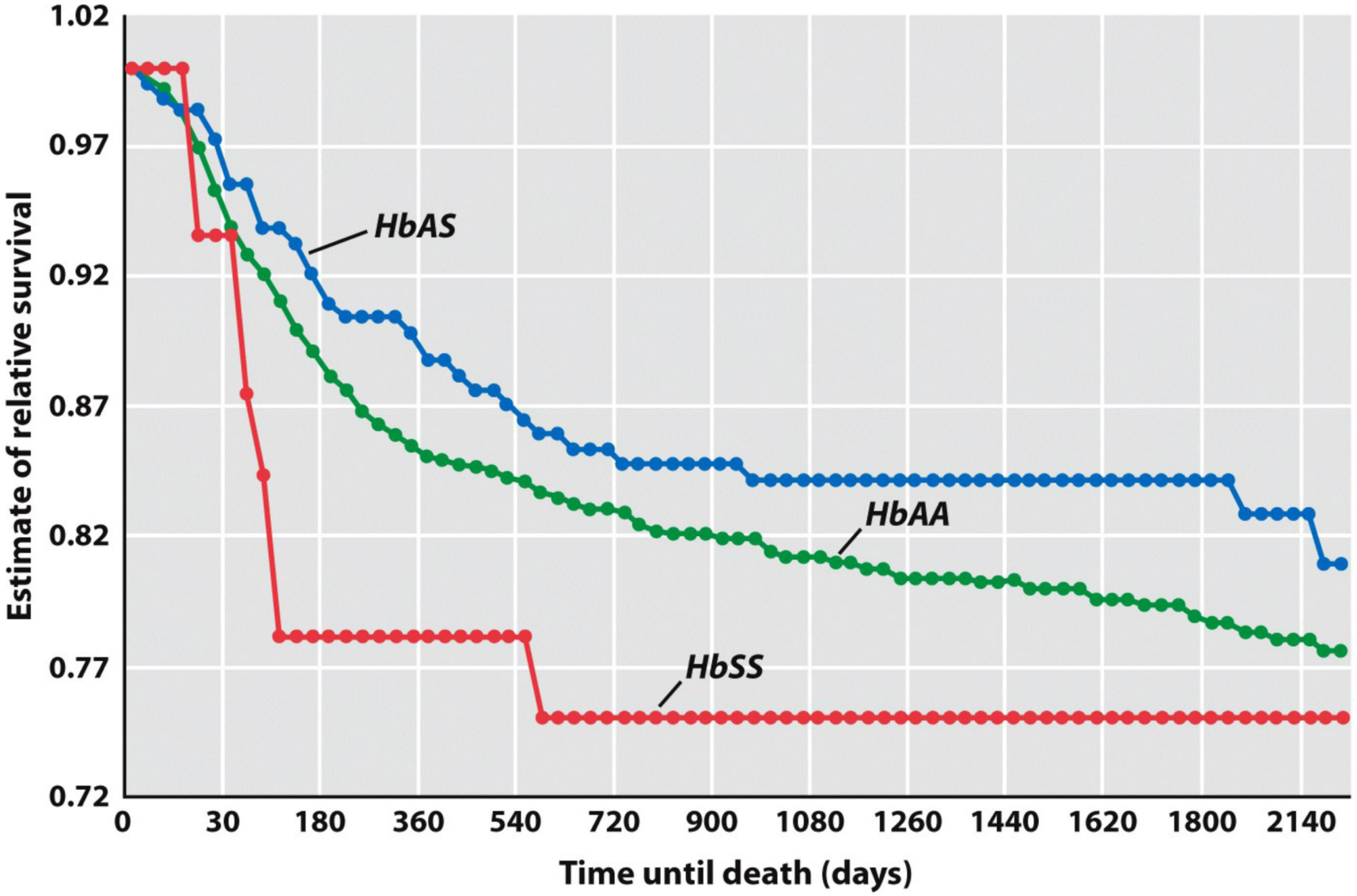
What's the difference?

Malaria is very common in lowland areas, but rare in highland areas

Distribution of malaria correlates with distribution of sickle cell trait



Heterozygotes have lower childhood mortality



Sickle cell trait is an example of **heterozygote advantage**

Heterozygotes have higher survival and reproduction than ***either*** homozygote

This causes both alleles to be maintained in the population, rather than either one becoming fixed.

5) Is this population at Hardy-Weinberg equilibrium?

AA	Aa	aa
0.2	0.2	0.6

A) Yes **B) No**

Observed frequencies:

AA	Aa	aa
0.2	0.2	0.6

p = **Freq(*A*)** = 0.2 + ½ (0.2) = **0.3**

$q = 1 - p$ = **0.7**

Expected frequencies, under Hardy-Weinberg:

Freq(AA) = p^2 = $(0.3)^2$ = **0.09**

Freq(Aa) = $2pq$ = 2(0.3)(0.7) = **0.42**

Freq(aa) = q^2 = $(0.7)^2$ = **0.49**

Observed frequencies strongly depart from Hardy-Weinberg expectation.

6) A population has four alleles at a gene A, in the frequencies shown below. After many generations, genetic drift eliminates all but one of these alleles. What is the probability that A_2 is the allele that goes to fixation?

Allele:	A_1	A_2	A_3	A_4
Frequency:	0.41	0.3	0.24	0.05

A. 0.09
B. 0.25
C. 0.3
D. 0.5