وراثة العشائر والوراثة الكمية
Population & Quantitative Genetics

Dr. Nabil MS Amor
KSU 2018/2019

Presentation of the Unit and context within the syllabus

- The first part of the unit includes an introduction to population genetics, analyzing the genetic constitution of a population, the changes in allele and genotype frequencies, the population structure, the genetic distance between populations and the linkage disequilibrium.

- The second part of the unit is devoted to the understanding of continuous variation of phenotype trait expression, the analysis of the components of phenotypic and genotypic variance and the different models for assessing genotypic values.

Learning outcomes

Students will be able to

- determine the genetic structure of a population and is aware of the changes that can be produced in it throughout time.
- assess the available phenotypic and genotypic variability and to determine which are the components of variation.
- calculate and explain genetic diversity estimates, alleles frequency
- integrate theoretical and practical knowledge in detecting loci involved in quantitative traits
- integrate advanced statistics, bioinformatics and genome data

Learning activities

- **Learning activity 1:** Lectures combined with examples
- **Learning activity 2:** Practical work on solving exercises and problems on population genetics and quantitative genetics for determining genetic distances, population structure, variance components and basic genetic parameters such as heritability and response to selection. The exercises will be carried out in pairs, manually, and using specific software.

Resources /Textbooks

iGenetics: A Molecular Approach 3rd Edition. Russell, Peter J. published by Benjamin Cummings (2009). ISBN-13: 978-0321569769.

GENETICS ESSENTIALS Concepts and Connections. Benjamin A. Pierce Southwestern University. First edition. ISBN-13: 978-1429230407 2010 pp 536

Principles of Genetics, Binder Ready Version, 6th Edition D. Peter Snustad, Michael J. Simmons ISBN: 978-0-470-90359-9 - 2011 pp 784.

Schaum's Outline of Genetics, Fifth Edition (Schaums Outline Series) Susan Elrod and William Stansfield5th Edition,Kindle Edition ISBN-13: 978-0071625036 - 2010 pp414.

Course topics

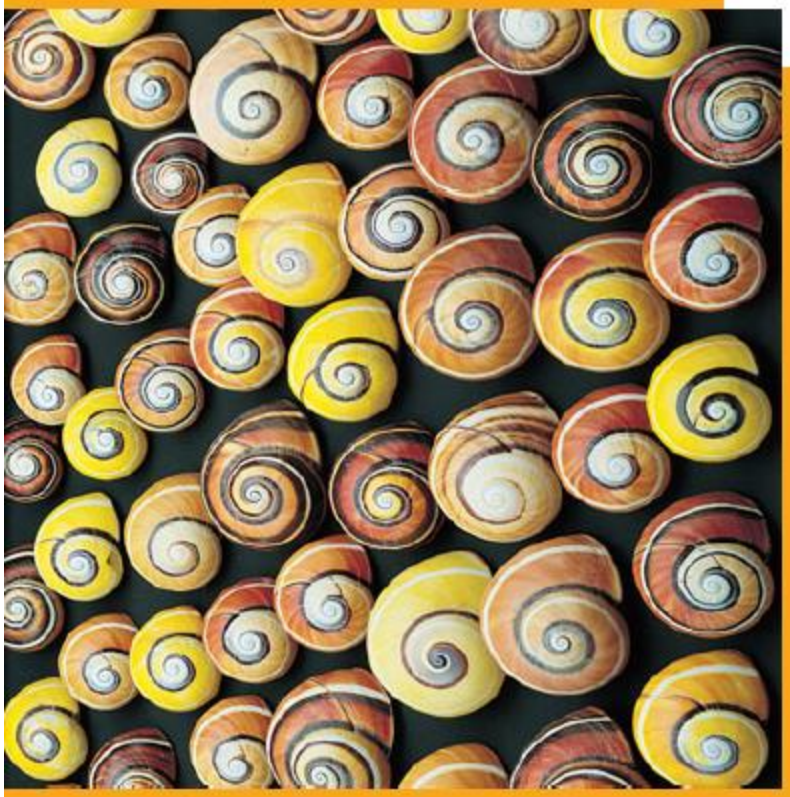
1. Population Genetics (p603)
2. The Hardy–Weinberg Law
3. Estimating Genetic variation
4. Deviations from Hardy-Weinberg equilibrium
5. Quantitative Genetics (p650)
6. The Inheritance of Continuous Traits
7. Polygene Hypothesis for Quantitative
8. Inheritance
9. Statistical Tools
10. Heritability

iGenetics: A Molecular Approach 3rd Edition. Russell, Peter J. published by Benjamin Cummings (2009). ISBN-13: 978-0321569769.

Assessment

- Lecture
  - 2 lecture exams (30%)
  - Cumulative final exam (40%)
  - 70% of total grade
- Laboratory
  - 2 laboratory exams
  - 30% of total grade

Lecture 1: Introduction



Extensive phenotypic variation in the color patterns of the Cuban tree snail.

I. Introduction

- **What is genetics?**
  - "Genetics is the study of **heredity**, the process in which a parent passes certain **genes** onto their children."

- **What does that mean?**
  - Children **inherit** their biological parents' genes that express specific **traits**, such as some physical characteristics, natural talents, and genetic disorders.

- **Genetic Concepts**

Heredity describes how some traits are passed from parents to their children.

The traits are expressed by **genes**, which are small sections of DNA that are coded for specific traits.

Genes are found on **chromosomes**.

Humans have two sets of **23** chromosomes—one set from each parent

I. Introduction

- Alleles: alternate forms or varieties of a gene.
- Evolution: genetic change in a population of organisms that occurs over time.
- Genes: units of inheritance usually occurring at specific locations, or loci, on a chromosome.
- Gene flow: the transference of genes from one population to another, usually as a result of migration.
- Gene pool: all of the genes in all of the individuals in a breeding population.
- Genome: the full genetic complement of an individual (or of a species).
- Genotype: the combination of alleles for a particular gene or locus.
- Heterozygous: a genotype consisting of two different alleles of a gene for a particular trait (Aa). Individuals who are heterozygous for a trait are referred to as heterozygotes (≠homozygotes).

I. Introduction

- Mendelian genetics : inheritance patterns which can be explained by simple rules of dominance and recessiveness of genes.
- Phenotype: the observable or detectable characteristics of an individual organism--the detectable expression of a genotype.
- Species: a group of organisms that can reproduce and produce fertile offspring.

For example, these happy face spiders *look* different, but since they can interbreed, they are considered the same species: *Theridion grallator*.

- Population: a group of organisms of the same species living in an area at the same time.

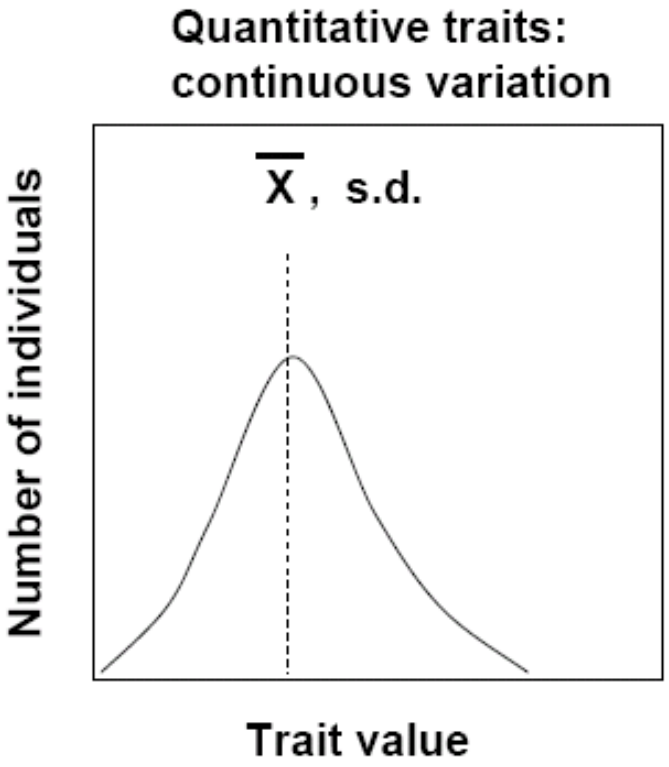
I. Introduction

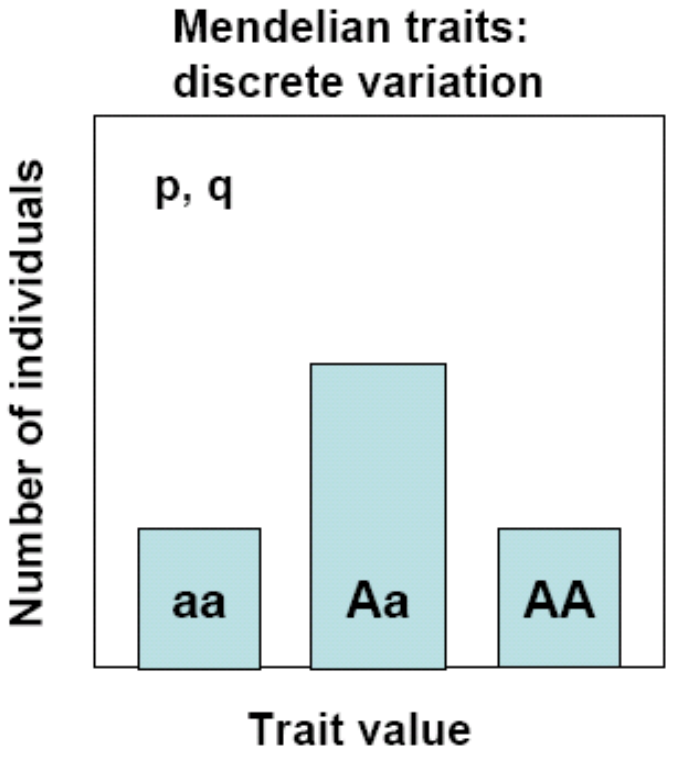
Qualitative genetics versus quantitative genetics

	Qualitative genetics	Quantitative genetics
Nature of traits	concerned with traits that have Mendelian inheritance and can be described according to kind	traits are described in terms of the degree of expression of the trait, rather than the kind
Scale of variability	provide discontinuous phenotypic variation	produce phenotypic variation that spans the full spectrum (continuous)
Number of genes	single genes	Poly-genic
Statistical analysis	is quite straightforward, and is based on counts and ratios	estimates of population parameters (attributes of the population from which the sample was obtained).

I. Introduction

- **Quantitative versus qualitative inheritance**





II. Population Genetics

•The study of the rules governing the maintenance and transmission of genetic variation in natural populations.

•Investigates genetic variation among individuals within groups (populations, gene pools).

•Usually in Natural Populations, examining geographic variation and through time.

•Different types of population genetics:

Empirical population genetics: measures and quantifies aspects of genetic variation in populations.

Theoretical population genetics: explains variation in terms of mathematical models of the forces that change allele frequencies (genetics drift, selection, gene flow, etc.).

II. Population Genetics

Types of questions studied by population geneticists:

How much variation occurs in natural populations, and what processes control the variation observed?

How does geography and dispersal behavior shape population structure?

What evolutionary forces are responsible for population differentiation, and the speciation process?

How do demographic factors such as breeding system, fecundity, changes in population size, and age structure influence the gene pool in the population?

II. Population Genetics

Important topics within population genetics:

- Genetic structure of populations
- Hardy-Weinberg Equilibrium
- Genetic variation in space and time
- Variation in natural populations, Wright's Fixation Index (F_{ST})
- Forces that change gene frequencies

II. Population Genetics

- The study of populations is intimately related to the study of evolution.
- It is the population, not the individual that evolves.
- Evolution can be defined in terms of what happens to the genetic structure of a population over time.
- Typically populations undergo changes in:
  - **Size** i.e. number of individuals
  - **Composition** i.e. the extent of phenotypic variation
  - **Behaviour** i.e. mating behaviour

  Charles Darwin saw evolution in these terms.

II. Population Genetics

A. Genetic structure of the parent population's gene pool

Ways to describe genetic structure of populations:

- We can define the genetic structure of the population i.e. the composition of the gene pool in terms of:

  (1) the genotype frequencies i.e. frequency of AA, Aa and aa

or

  (2) the allele frequencies i.e. Frequency of A Frequency of a

Genotypic frequency

- Count individuals with one genotype and divide by total number of individuals. Repeat for each genotype in the population.

***f(AA)* = count of *AA*/ total number of individuals**

***f(Aa)* = count of *Aa*/ total number of individuals**

***f(aa)* = count of *aa*/ total number of individuals**

II. Population Genetics

A. Genetic structure of the parent population's gene pool

Ways to describe genetic structure of populations:

Allelic frequency

- **Allelic frequencies offer more information than genotypic frequencies and can be calculated in two different ways:**

  1. **Allele (gene) counting method:**

     ***p* = *f(A)* = (2 x count of *AA*) + (1 count of *Aa*)/ 2 x total number of individuals**

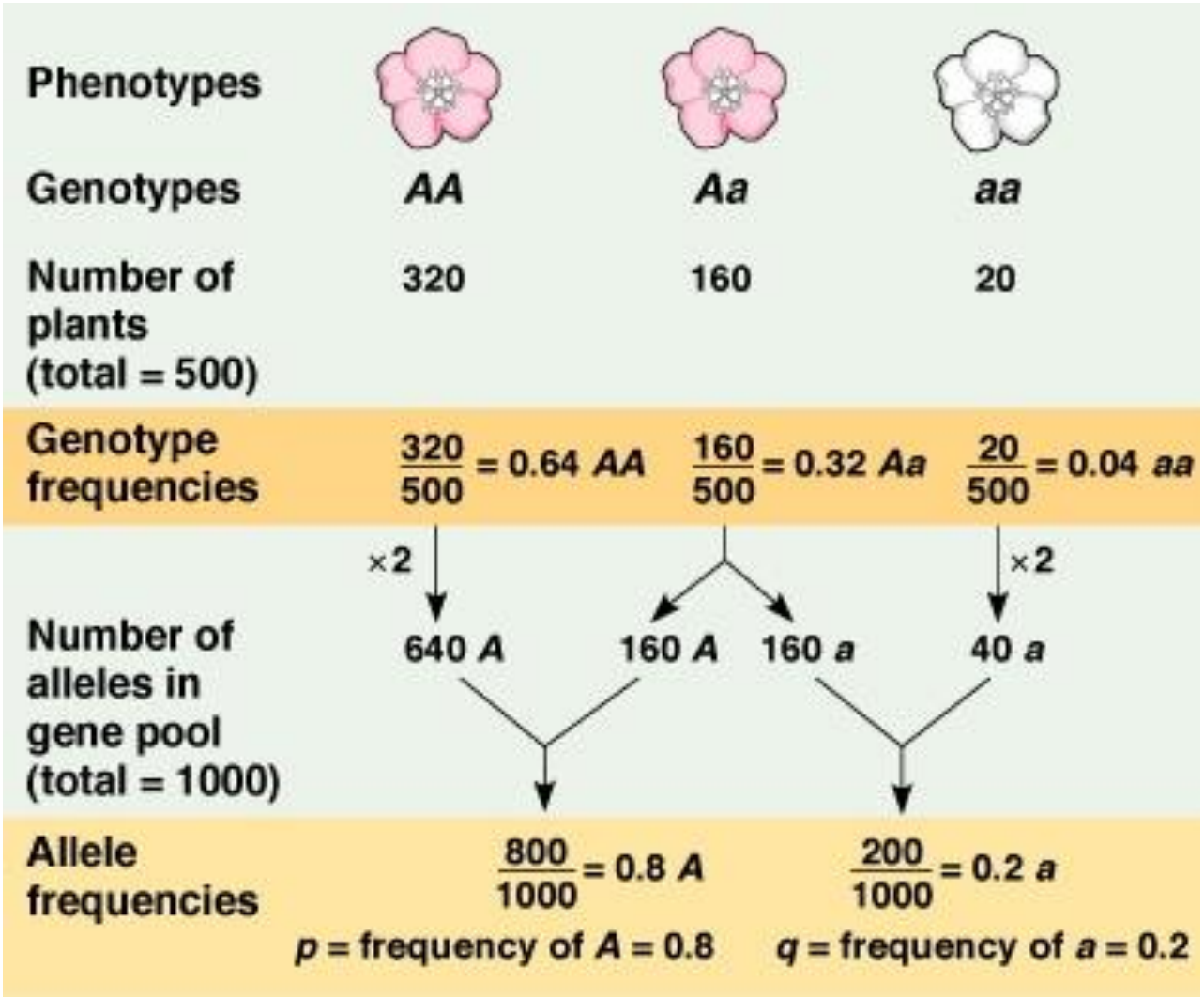
  2. **Genotypic frequency method:**

     ***p* = *f(A)* = (frequency of the *AA* homozygote) + (1/2 x frequency of the *Aa* heterozygote)**

     ***q* = *f(a)* = (frequency of the *aa* homozygote) + (1/2 x frequency of the *Aa* heterozygote)**

II. Population Genetics

A. Genetic structure of the parent population's gene pool



II. Population Genetics

A. Genetic structure of the parent population's gene pool

A frequency is a proportion with a range of 0–1. If 43% of population has a trait, the frequency of that trait is 0.43. For any given trait, the sum of the genotypic frequencies in a population should be 1.

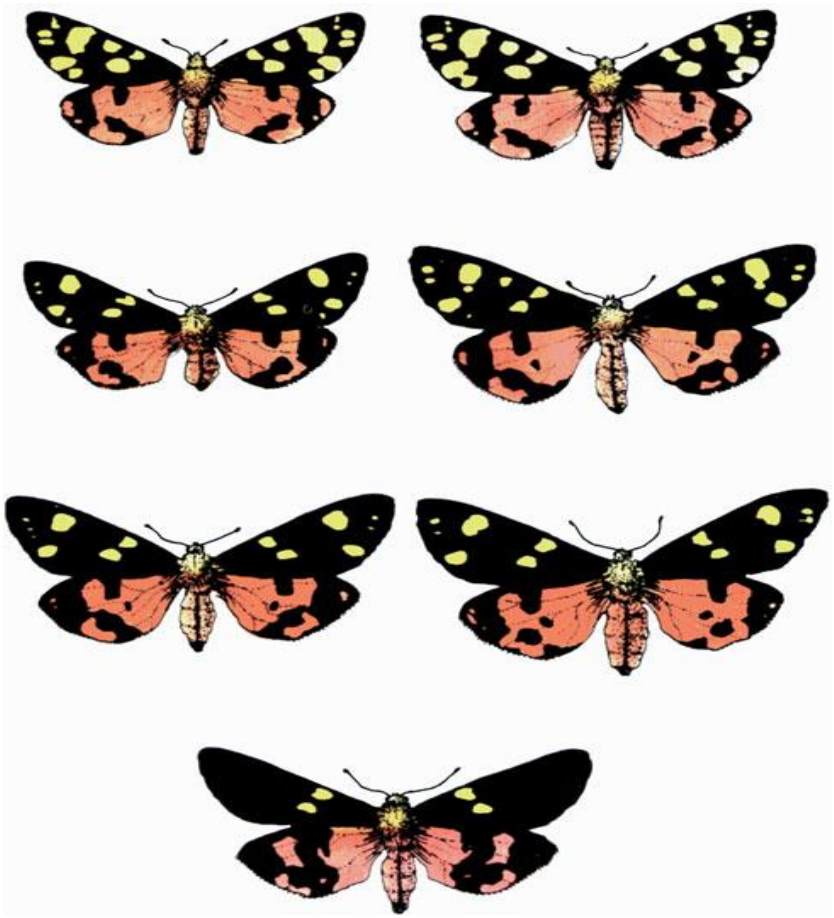


Figure 21.2 *Panaxia dominula*, the scarlet tiger moth. The top two moths are normal homozygotes (*BB*), those in the middle two rows are heterozygotes (*Bb*), and the bottom moth is the rare homozygote (*bb*).

Box 21.1 Sample Calculation

Box 21.1 Sample Calculation of Genotype and Allele Frequencies for Hemoglobin Variants Among Nigerians Where Multiple Alleles Are Present

Hemoglobin Genotypes

AA	AS	SS	AC	SC	CC	TOTAL
2017	783	4	173	14	11	3002

Calculation of Genotype Frequencies

Genotype frequency=Number of individuals with the genotype / Total number of individuals

Calculation of Allele Frequencies from the Number of Individuals with a Particular Genotype

Allele frequency=Number of copies of a given allele in the population / Sum of all alleles in the population

Calculation of Allele Frequencies from the Frequencies of Particular Genotypes

II. Population Genetics

A. Genetic structure of the parent population's gene pool

Allelic frequencies at X-linked loci:

Females have 2 X-linked alleles, and males have 1 X-linked allele.

$$p = f(X^A) = \frac{(2 \times X^A X^A \text{ females}) + (X^A X^a \text{ females}) + (X^A Y \text{ males})}{(2 \times \#\text{ females}) + (\#\text{ males})}$$

$$q = f(X^a) = \frac{(2 \times X^a X^a \text{ females}) + (X^A X^a \text{ females}) + (X^a Y \text{ males})}{(2 \times \#\text{ females}) + (\#\text{ males})}$$

If number of females and males are equal:

$$p = f(X^A) = 2/3[f(X^AX^A) + 1/2f(X^AX^a)] + 1/3f(X^AY)$$

$$q = f(X^a) = 2/3[f(X^aX^a) + 1/2f(X^AX^a)] + 1/3f(X^aY)$$

Lecture 2: Hardy-Weinberg law

II. Population Genetics

B. Hardy-Weinberg law

1. Basis of the Hardy-Weinberg law:

- **Independently discovered by Godfrey H. Hardy (1877-1947) and Wilhelm Weinberg (1862-1937).**
- **Explains how Mendelian segregation influences allelic and genotypic frequencies in a population.**

Assumptions:

1. **Population is infinitely large, to avoid effects of genetic drift (= change in genetic frequency due to chance).**
2. **Panmictic population: Mating is random (with regard to traits under study).**
3. **No natural selection (for traits under study).**
4. **No mutation.**
5. **No migration.**

II. Population Genetics

B. Hardy-Weinberg law

1. Basis of the Hardy-Weinberg law:

- **If assumptions are met, population will be in genetic equilibrium.**

Two expected predictions:

1. **Allele frequencies do not change over generations.**

2. **After one generation of random mating, genotypic frequencies will remain in the following proportions:**

   p^2 (frequency of AA)

   2pq (frequency of Aa)

   q^2 (frequency of aa)

   ***p = allelic frequency of A**
   ***q = allelic frequency of a**
   ***$p^2 + 2pq + q^2 = 1$**

II. Population Genetics

B. Hardy-Weinberg law

1. Basis of the Hardy-Weinberg law:

Hardy-Weinberg state: $p^2 + 2pq + q^2 = 1$ at equilibrium

1. **Zygotes form by random combinations of alleles, in proportion to the abundance of the alleles in the population.**

2. **If *f(p)* = 0.5 and *f(q)* = 0.5, outcome is as follows:**

	A(p)	a(q)
A(p)	AA(p^2) 0.5 x 0.5 = 0.25	Aa(pq) 0.5 x 0.5 = 0.25
a(q)	Aa(pq) 0.5 x 0.5 = 0.25	aa(q^2) 0.5 x 0.5 = 0.25

Table 21.1 Possible Combinations of A and a Gametes from Gametic Pools for a Population

3. **When population is at equilibrium:**

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

Algebraic Proof of Genetic Equilibrium

Table 21.2 Algebraic Proof of Genetic Equilibrium in a Randomly Mating Population for One Gene Locus with Two Alleles

Type of Mating ♀ ♂	Mating Frequency	Offspring Frequencies Contributed to the Next Generation by a Particular Mating: *AA*	*Aa*	*aa*
$p^2\,AA \times p^2\,AA$	p^4	p^4	—	—
$p^2\,AA \times 2pq\,Aa$ $2pq\,Aa \times p^2\,AA$ [a]	$4p^3q$	$2p^3q$	$2p^3q$	—
$p^2\,AA \times q^2\,aa$ $q^2\,aa \times p^2\,AA$	$2p^2q^2$	—	$2p^2q^2$	—
$2pq\,Aa \times 2pq\,Aa$	$4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
$2\,pq\,Aa \times q^2\,aa$ $q^2\,aa \times 2\,pq\,Aa$	$4pq^3$	—	$2pq^3$	$2pq^3$
$q^2aa \times q^2aa$	q^4	—	—	q^4
Totals	$(p^2 + 2pq + q^2)^2 = 1$	$p^2(p^2 + 2pq + q^2) = p^2$	$2pq(p^2 + 2pq + q^2) = 2pq$	$q^2(p^2 + 2pq + q^2) = q^2$
Genotype frequencies $= (p + q)^2 = p^2 + 2pq + q^2 = 1$ in each generation afterward.				
Gene (allele) frequencies $= p(A) + q(a) = 1$ in each generation afterward.				

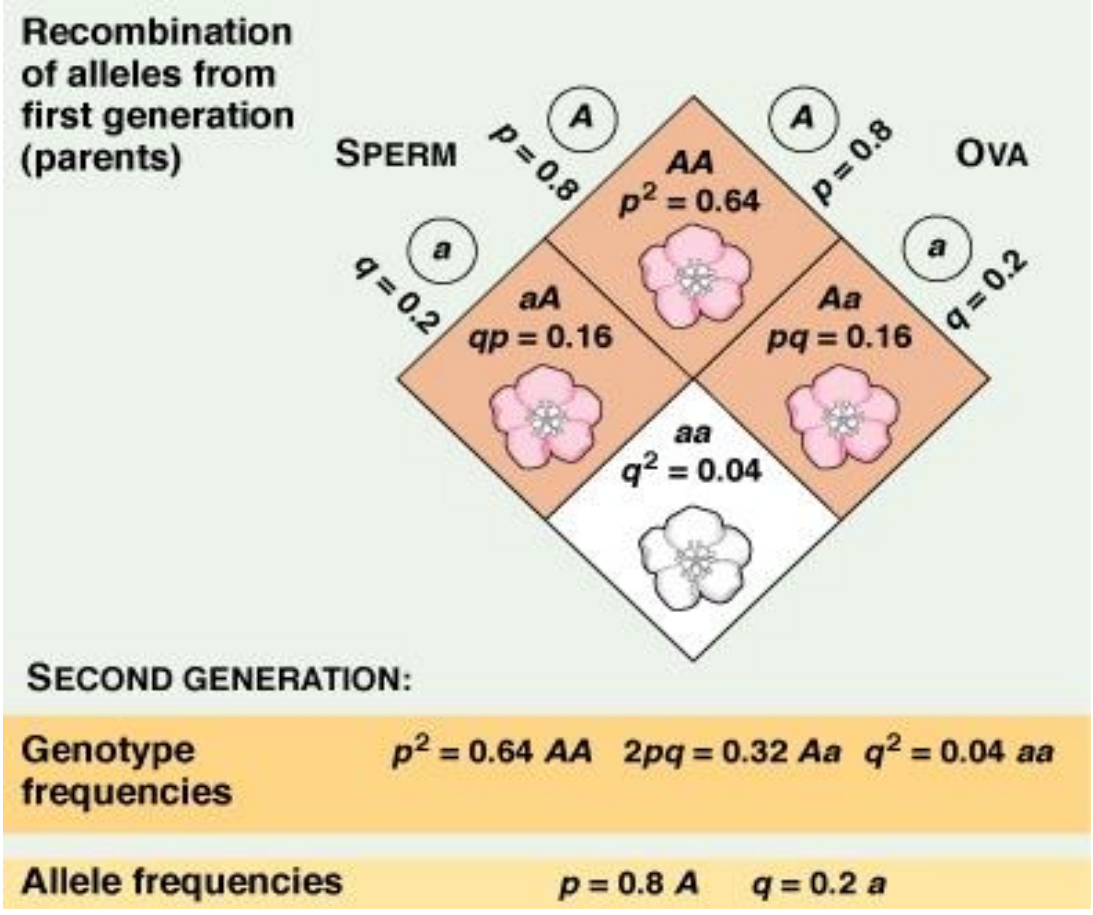
[a] For example, matings between *AA* and *Aa* will occur at $p^2 \times 2pq = 2p^3q$ for $AA \times Aa$ and at $2pq \times p^2 = 2p^3q$ for $Aa \times AA$ for a total of $4p^3q$. Two progeny types, *AA* and *Aa*, result in equal proportions from these matings. Therefore, offspring frequencies are $2p^3q$ (i.e., $\frac{1}{2} \times 4p^3q$) for *AA* and for *Aa*.

II. Population Genetics

B. Hardy-Weinberg law

1. Basis of the Hardy-Weinberg law:

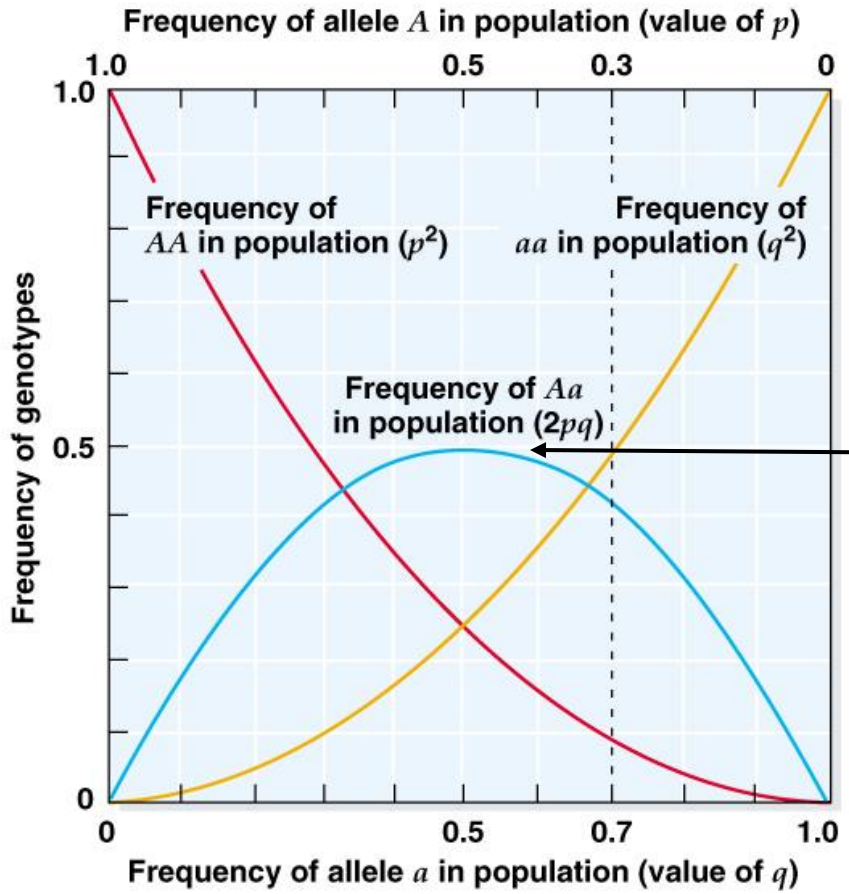
- No change in allele frequency
- Thus, no evolution
- Therefore, all five Hardy–Weinberg equilibrium conditions must have been met!



II. Population Genetics

B. Hardy-Weinberg law

1. Basis of the Hardy-Weinberg law:



Max. heterozygosity @ p = q = 0.5

Fig. 22.3 Relationship of the frequencies of the genotypes *AA*, *Aa*, and *aa* to the frequencies of alleles *A* and *a* in populations in Hardy-Weinberg equilibrium

II. Population Genetics

B. Hardy-Weinberg law 1. Basis of the Hardy-Weinberg law:

Some facts about assumptions of the Hardy-Weinberg law:

1. Population is infinitely large.

- Assumption is unrealistic.
- Large populations are mathematically similar to infinitely large populations.
- Finite populations with rare mutations, rare migrants, and weak selection generally fit Hardy-Weinberg proportions.

2. Mating is random: Panmixia .

- Few organisms mate randomly for all traits or loci.
- Hardy-Weinberg applies to any locus for which mating occurs randomly, even if mating is non-random for other loci.
- This works because different loci assort independently due to recombination.

II. Population Genetics

B. Hardy-Weinberg law

3. **No natural selection**
4. **No mutation**
5. **No migration**

- Gene pool must be closed to the addition/subtraction of new alleles.
- Selection can subtract alleles or cause some alleles to increase in frequency.
- Mutation always adds to variation (generates novel alleles).
- Effects of mutation can be accommodated with a model (e.g., infinite alleles model).
- Migration can either add or subtract variation depending on which alleles migrants carry and whether they immigrate or emigrate.
- Like random mating, condition applies only to the locus under study. Genes are unlinked because alleles sort independently on different chromosomes due to recombination.

II. Population Genetics

B. Hardy-Weinberg law

2. H W for loci with more than two alleles:

For three alleles (A, B, and C) with frequencies p, q, and r:

Binomial expansion ⇒

$$(p + q + r)^2 = p^2(AA) + 2pq(AB) + q^2(BB) + 2pr(AC) + 2qr(BC) + r^2(CC)$$

Blue mussel population of Long Island Sound (Figure 22.6) is an example.

For four alleles (A, B, C, and D) with frequencies p, q, r, and s:

$$(p + q + r + s)^2 = p^2(AA) + 2pq(AB) + q^2(BB) + 2pr(AC) + 2qr(BC) + r^2(CC) + 2ps(AD) + 2qs(BD) + 2rs(CD) + s^2(DD)$$

II. Population Genetics

B. Hardy-Weinberg law 3. H W for X-linked alleles:

e.g., Humans and *Drosophila* (XX = female, XY = male)

	$X^A(p)$	$X^a(q)$	Y
$X^A(p)$	X^AX^A p^2	X^AX^a pq	X^AY p
$X^a(q)$	X^AX^a qp	X^aX^a q^2	X^aY q

Females

- **Hardy-Weinberg frequencies are the same for any other locus:** $p^2 + 2pq + q^2 = 1$

Males

- **Genotype frequencies are the same as allele frequencies:** $p + q = 1$
- **Recessive X-linked traits are more common among males.**

II. Population Genetics

B. Hardy-Weinberg law 3. H W for X-linked alleles:

If alleles are X-linked and sexes differ in allelic frequency, Hardy Weinberg equilibrium is approached over several generations.

Allelic frequencies oscillate each generation until the allelic frequencies of males and females are equal.

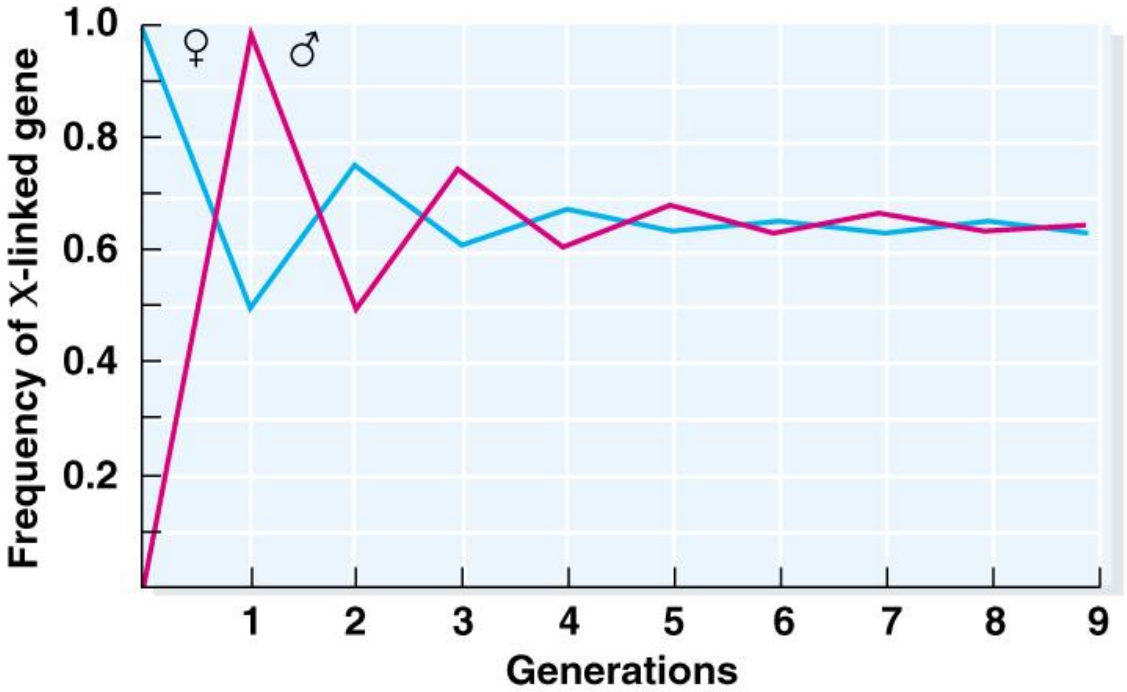


Fig. 22.4 Representation of the gradual approach

Lecture 3: Estimating Genetic variation

II. Population Genetics

B. Hardy-Weinberg law

4. Estimating frequencies & Testing HW assumptions :

Estimating allelic frequencies:

- **If one or more alleles are recessive, can't distinguish between heterozygous and homozygous dominant individuals.**
- **Use Hardy-Weinberg to calculate allele frequencies based on the number of homozygous recessive individuals.**

If $q^2 = 0.0043$,

then $q = 0.065$;

$p = 1 - q = 0.935$

$p^2 = 0.8742$, $2pq = 0.1216$

Figure 21.5

Three Hopi girls, photographed about 1900. The middle child has albinism, an autosomal recessive disorder that occurs with high frequency among the Hopi Indians of Arizona.

II. Population Genetics

B. Hardy-Weinberg law

4. Estimating frequencies & Testing HW assumptions :

Testing Hardy-Weinberg assumptions:

- **Data from real populations rarely match Hardy-Weinberg proportions.**
- **Test observed and expected proportions with a goodness of of fit (GF) test such as a chi-square test.**
- **If deviation is larger than expected, begin to determine which assumptions are violated (this is where the real work of population genetics begins).**
- **Factors that contribute to non-equilibrium:**
  - **Population differentiation (through drift and mutation)**
  - **Fluctuations in population size**
  - **Selection & migration**

II. Population Genetics

B. Hardy-Weinberg law 5. Important measures of genetic variation:

- **Polymorphism = % of loci or nucleotide positions showing more than one allele or base pair.**
- **Heterozygosity (*H*) = % of individuals that are heterozygotes.**
- **Allele/haplotype diversity = measure of # and diversity of different alleles/haplotypes within a population (note---it is important to correct for sample size, because larger samples are expected to harbor more greater allelic variation).**
- **Nucleotide diversity = measure of number and diversity of variable nucleotide positions within sequences of a population.**
- **Genetic distance = measure of number of base pair differences between two homologous sequences.**
- **Synonomous/nonsynonomous substitutions = % of nucleotide substitutions that do not/do result in amino acid replacement.**

II. Population Genetics

B. Hardy-Weinberg law

5. Important measures of genetic variation:

- **Useful to partition genetic variation into components:**

  within populations / between populations / among populations

- **Sewall Wright's Fixation index (F_{ST}) is a useful index of genetic differentiation and comparison of overall effect of population substructure.**

  - **Measures reduction in heterozygosity (H) expected with non-random mating at any one level of population hierarchy relative to another more inclusive hierarchical level.**

  - $F_{ST} = (H_{Total} - H_{subpop})/H_{Total}$

  - F_{ST} **ranges between minimum of 0 and maximum of 1:**

    = 0 ⇒ **no genetic differentiation**
    << 0.5 ⇒ **little genetic differentiation**
    >> 0.5 ⇒ **moderate to great genetic differentiation**
    = 1.0 ⇒ **populations fixed for different alleles**

Recap---methods used to measure genetic variation:

- **1960s-1970s: genetic variation was first measured by protein electrophoresis.**
  - **Allozymes (protein electrophoresis)**
  - **Restriction Fragment Length Polymorphisms (RFLPs)**
- **1980s-2000s: genetic variation measured directly at the DNA level**
  - **Minisatellites (VNTRs)**
  - **Microsatellites (STRs)**
  - **DNA sequence**
  - **Single Nucleotide Polymorphisms (SNPs)**
  - **Single-stranded Conformation Polymorphism (SSCP)**
  - **Random Amplified Polymorphic DNAs (RAPDs)**
  - **Amplified Fragment Length Polymorphisms (AFLPs)**
- **2000s+: genetic variation measured at whole genome level**
  - **Whole genome shotgun sequencing**
  - **RNA-Seq (sequence complete transcriptome)**
  - **Restriction-site associated DNA markers (RAD-Seq)**
  - **Sequence Capture (targeting specific regions)**

DNA from individual 1 and individual 2 differ in one nucleotide, found within the sequence recognized by the restriction enzyme *Bam*HI

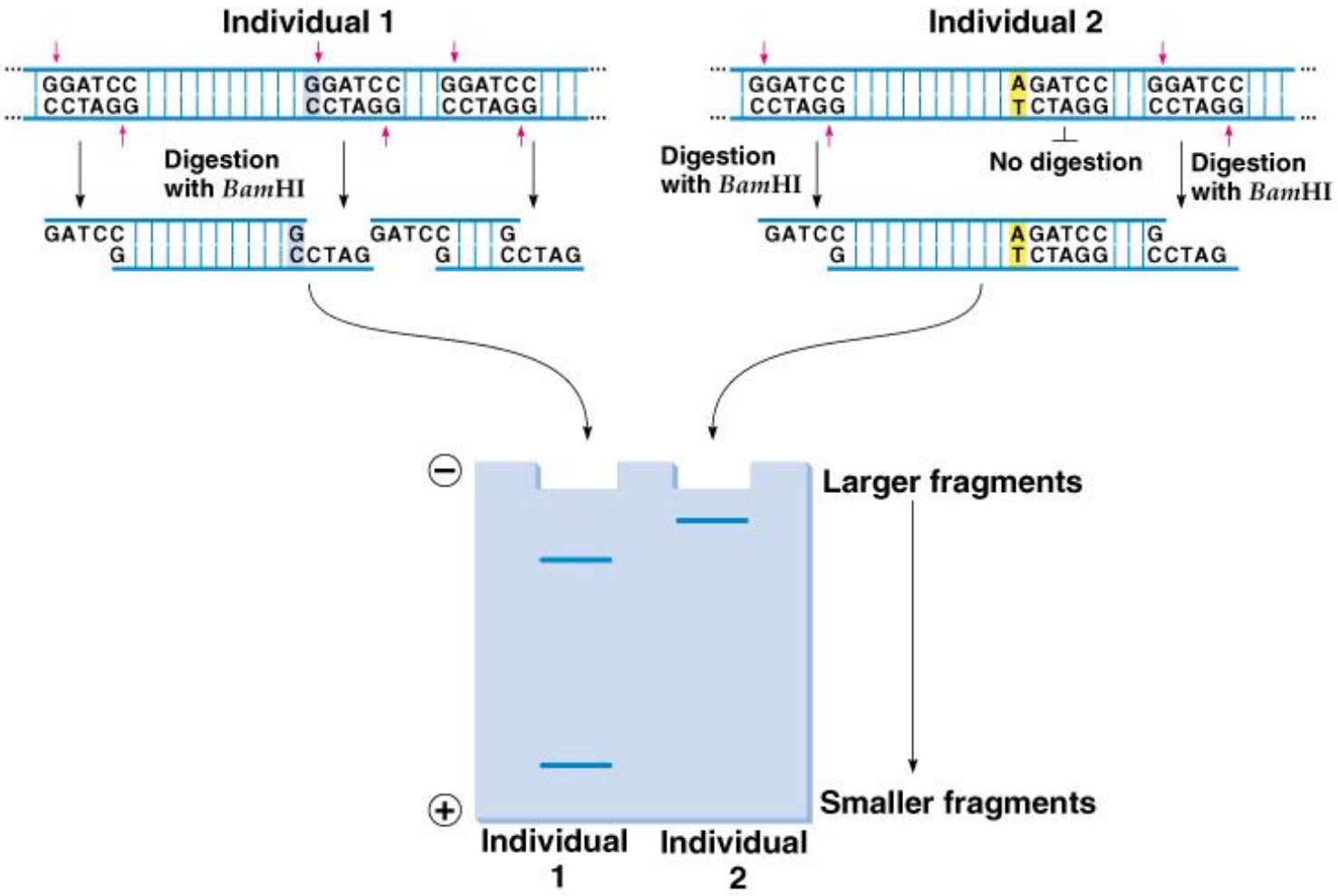
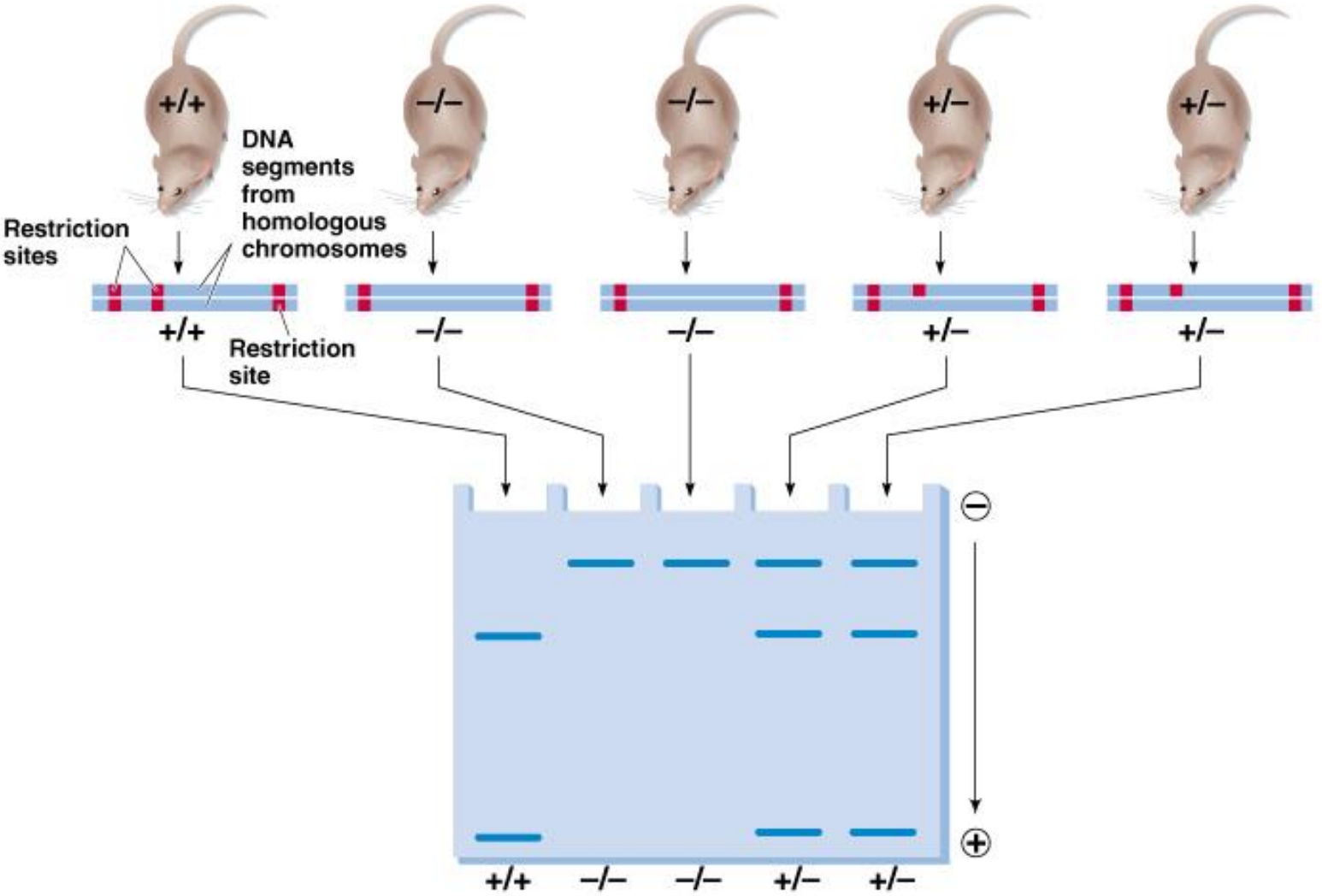


Fig. 21.8 Restriction patterns from five mice



KEY CONCEPT

Hardy-Weinberg equilibrium provides a framework for understanding how populations evolve.

Lecture 4: Deviations from Hardy-Weinberg equilibrium(1)

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

- if a population is NOT in HWE, then one of the assumptions must be violated.

Sources of Variation

Mutation

Recombination

- crossing over

- independent assortment

VARIATION

Agents of Change

N.S.

Drift

Migration

Mutation

Non-random Mating

So, if NO AGENTS are acting on a population, then it will be in equilibrium and WON'T change.

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

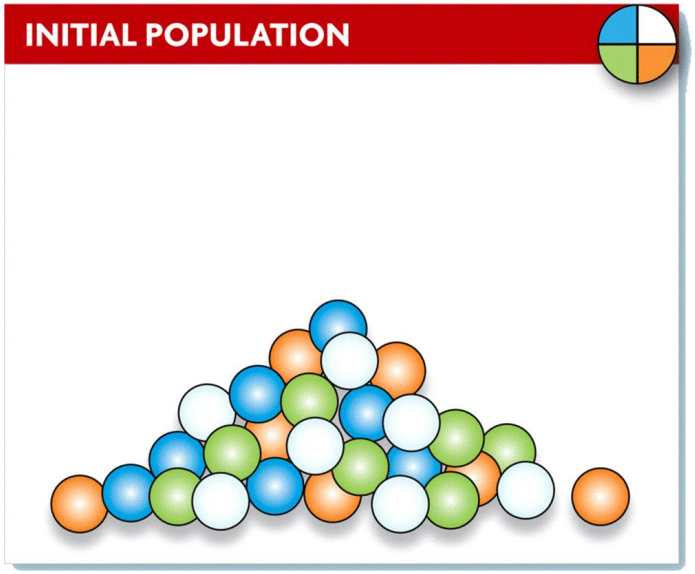
- The Hardy-Weinberg theory provides a baseline against which we can compare the allele and genotype frequencies of an evolving population.
- We can define **microevolution** as generation-to-generation change in a population's frequencies of alleles.
  - Microevolution occurs even if the frequencies of alleles are changing for only a single genetic locus in a population while the others are at equilibrium.

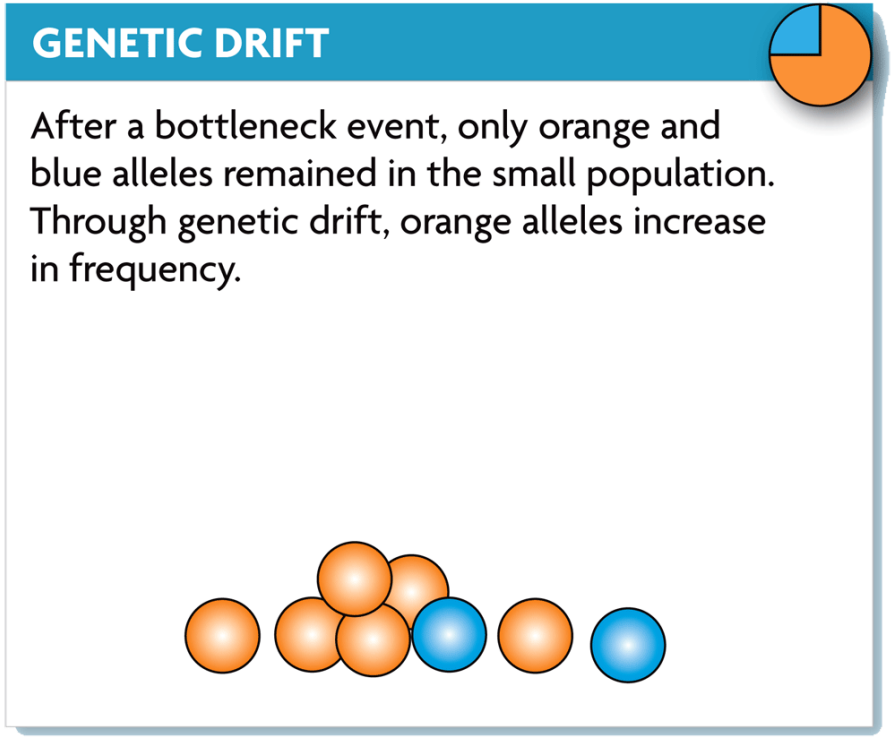
II. Population Genetics

C.Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

- Small Pop= Genetic drift changes allele frequencies due to chance alone.





II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

The organisms that actually reproduce in a population may not be representative of the genetics structure of the population; they may vary just due to *sampling error*

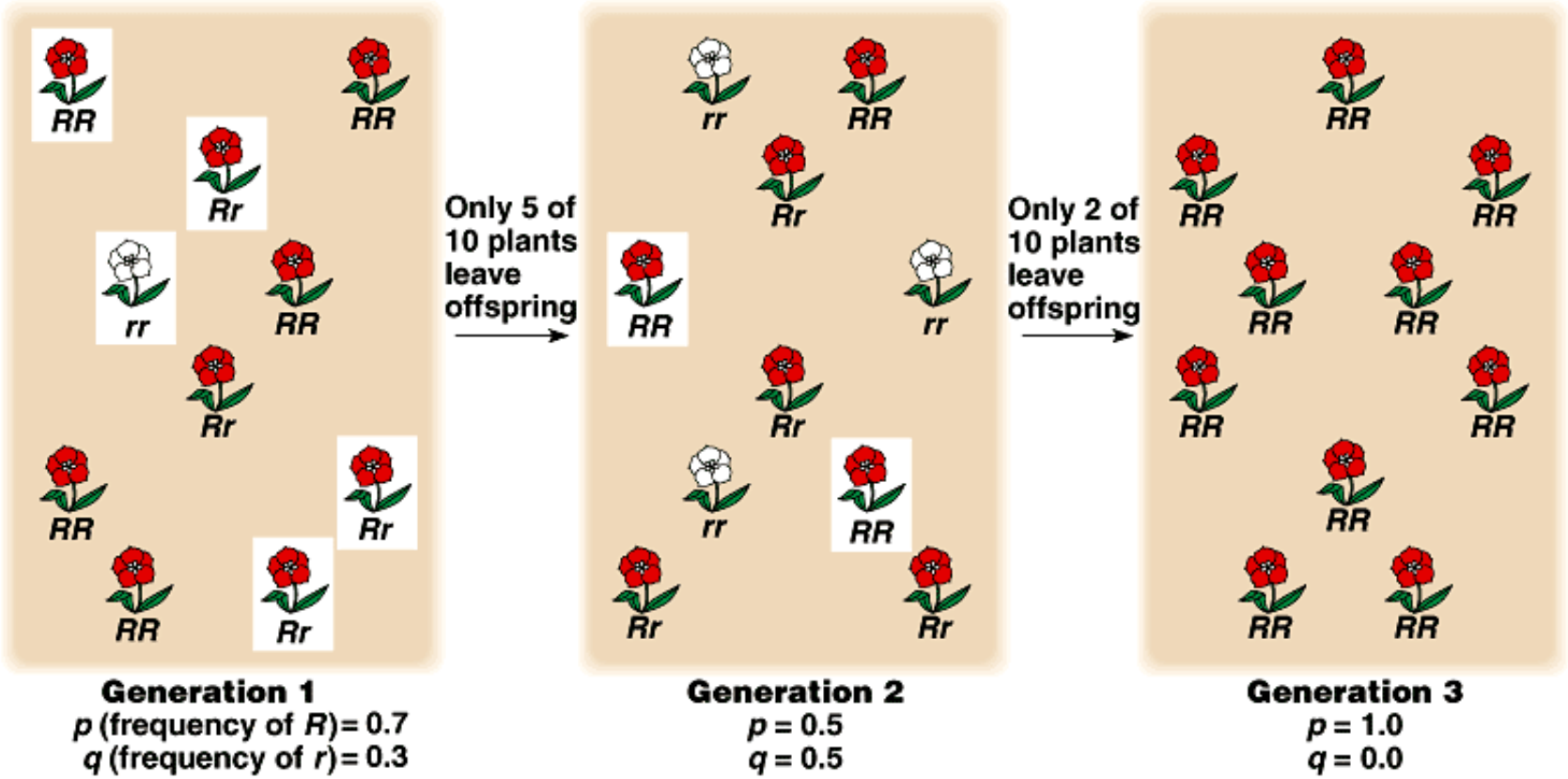
- ***Genetic drift*** occurs when changes in gene frequencies from one generation to another occur because of chance events (sampling errors) that occur when populations are finite in size.
  - For example, one would not be too surprised if a coin produced seven heads and three tails in ten tosses, but you would be surprised if you saw 700 heads and 300 tails in 1000 tosses - you expect 500 of each.
  - The smaller the sample, the greater the chance of deviation from an idealized result.
  - Genetic drift at small population sizes often occurs as a result of two situations: the bottleneck effect or the founder effect.

II. Population Genetics

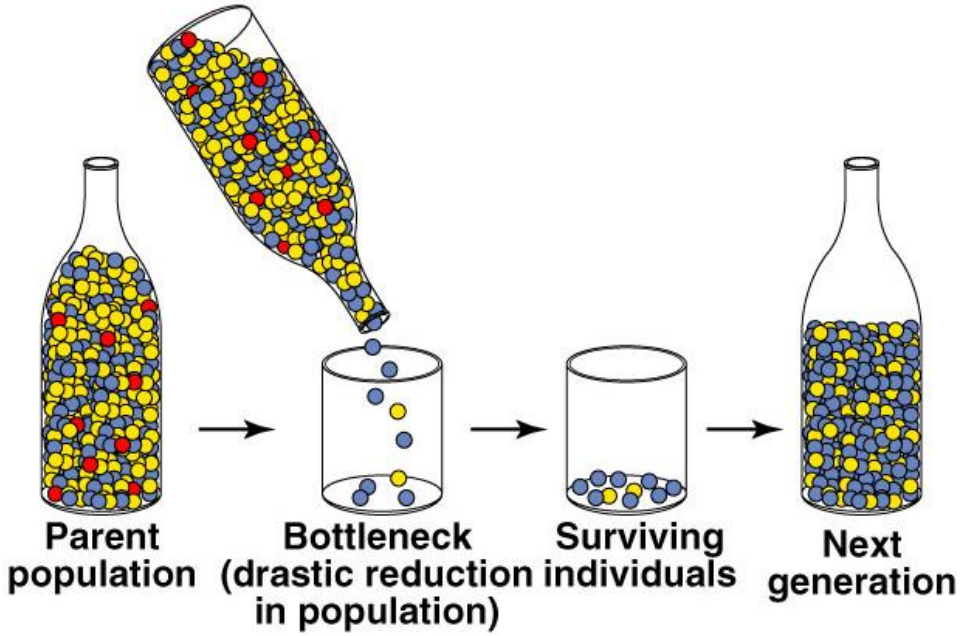
C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

- For example, in a small wildflower population with a stable size of only ten plants, genetic drift can completely eliminate some alleles.



- The **bottleneck effect** occurs when the numbers of individuals in a larger population are drastically reduced by a disaster.
  - By chance, some alleles may be overrepresented and others underrepresented among the survivors.
  - Some alleles may be eliminated altogether.
  - Genetic drift will continue to impact the gene pool until the population is large enough to minimize the impact of sampling errors.



- "Genetic Bottleneck"

If a population crashes (perhaps as the result of a plague) there will be both selection and drift. There will be selection for those resistant to the disease (and correlated selection for genes close to the genes conferring resistance), but there will also be drift at other loci simply by reducing the size of the breeding population.

European Bison, hunted to 12 individuals, now number over 1000.

Cheetah have very low genetic diversity, suggesting a severe bottleneck in the past. They can even exchange skin grafts without rejection…

Elephant seals fell to 100's in the 1800s, now in the 100,000's

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

- The **founder effect** occurs when a new population is started by only a few individuals that do not represent the gene pool of the larger source population.
  - At an extreme, a population could be started by single pregnant female or single seed with only a tiny fraction of the genetic variation of the source population.
- Genetic drift would continue from generation to generation until the population grew large enough for sampling errors to be minimal.
  - Founder effects have been demonstrated in human populations that started from a small group of colonists.

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

GENETIC DRIFT AS A CAUSE OF INBREEDING

As we have seen, **inbreeding results from drift** because alleles become **identical by descent** (**IBD**). We can therefore measure drift in terms of our inbreeding coefficient, **F**, and hence how the fraction of heterozygosity, **Het**, declines with time.

We can show: $$Het_t = Het_0\left(1-\frac{1}{2N}\right)^t$$

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

LOSS OF HETEROZYGOSITY IN A RANDOM MATING POPULATION OF N ADULTS

- The rate of loss of heterozygosity per generation is equal to the probability that a newborn contains two alleles at a locus that are identical-by-descent from the previous generation,

= 1/(2N)

- Heterozygosity after 1 generation at size N,

$$H_1 = \left[1 - \frac{1}{2N}\right] * H_0$$

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

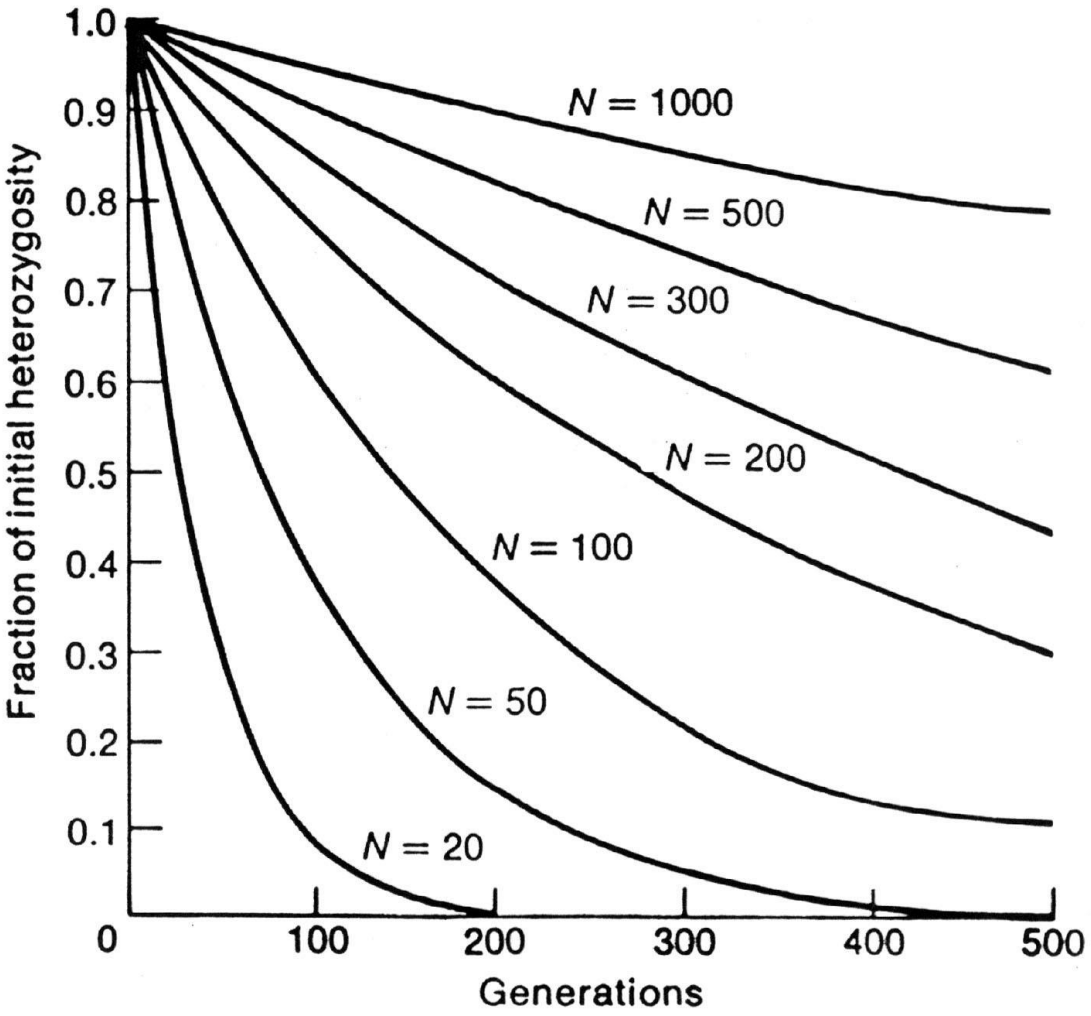
1. Finite Population Sizes: Genetic Drift

- Heterozygosity after **t** generations at size N,

$$H_t = \left[1 - \frac{1}{2N}\right]^t * H_0 \cong H_0 e^{-t/(2N)}$$

- Example: After t = 6N generations, $e^{-t/(2N)} = 0.05$, implying that 95% of the original heterozygosity has been lost. This is 60 generations for a population size of 10 breeding adults.

LOSS OF HETEROZYGOSITY VS. POPULATION SIZE



EFFECT OF GENETIC DRIFT ON GENETIC DIVERSITY

$N_e = 16$

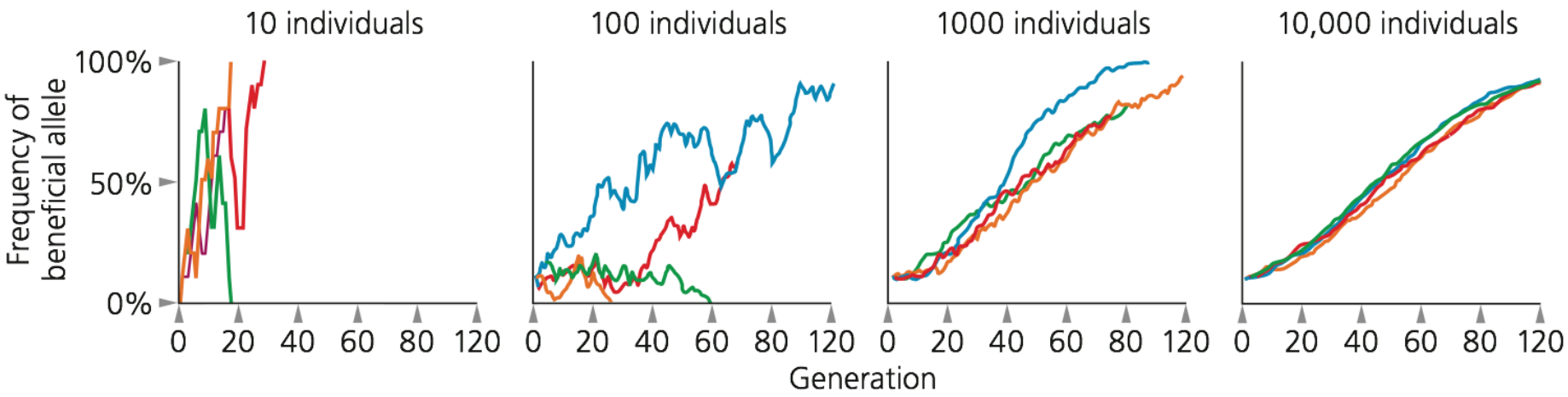
$N_e = 9$

- Genetic drift leads to a reduction in heterozygosity.

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift



- Drift weaker in large populations
- Small advantages in fitness can lead to large changes over the long term

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

RATE OF GENETIC DRIFT AND FLUCTUATIONS IN POPULATION SIZE

- Effective population size (N_e): the number of individuals in an ideal population (in which every individual reproduces) in which the rate of genetic drift would be the same as it is in the actual population.
- The rate of genetic drift is highly influenced by the lowest population size in a series of generations.
- The effective population size (N_e) over multiple generations is best represented by the harmonic mean not the arithmetic mean.

$$\frac{1}{N_e} = \frac{1}{t}\left(\frac{1}{N_0} + \frac{1}{N_1} + \cdots + \frac{1}{N_{t-1}}\right)$$

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

RATE OF GENETIC DRIFT AND FLUCTUATIONS IN POPULATION SIZE

Example:

- Suppose a population went through a bottleneck as follows:

$$N_0 = 1000,\ N_1 = 10,\ N_2 = 1000$$

- What is the effective size (Ne) of this population across all three generations?

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

RATE OF GENETIC DRIFT AND FLUCTUATIONS IN POPULATION SIZE

HARMONIC MEAN:

$$1/N_e = (1/3)(1/1000 + 1/10 + 1/1000) = 0.034$$

$$N_e = 29.4$$

ARITHMETIC MEAN:

$$N_e = (1/3)(1000 + 10 + 1000) = 670$$

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

1. Finite Population Sizes: Genetic Drift

GENETIC EFFECTIVE POPULATION SIZE (N_E)

- The effective population size is often << than the actual census size

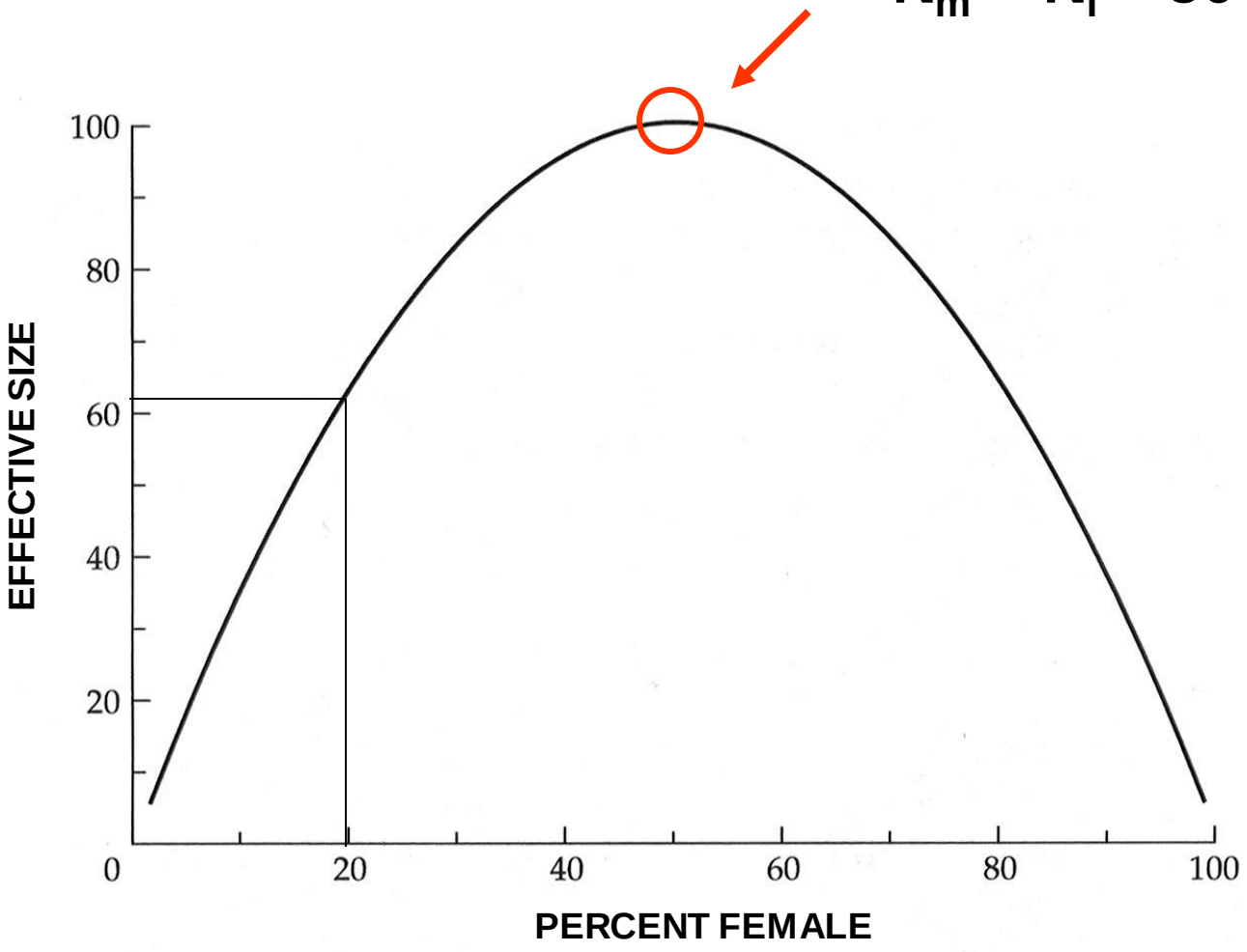
$$N_e << N_a$$

- Consider a sexual population consisting of N_m males and N_f females
- The actual size is, $N_a = N_m + N_f$ but,
- The effective size is,

$$N_e = \frac{4N_m N_f}{N_m + N_f}$$

- For a population of $N_a = 100$

Equal Sex Ratio
$N_m = N_f = 50$



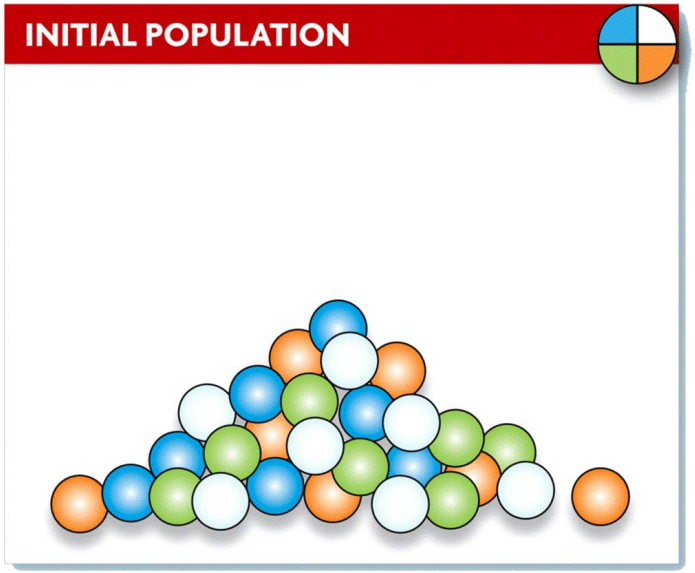
Lecture 5: Deviations from Hardy-Weinberg equilibrium(2)

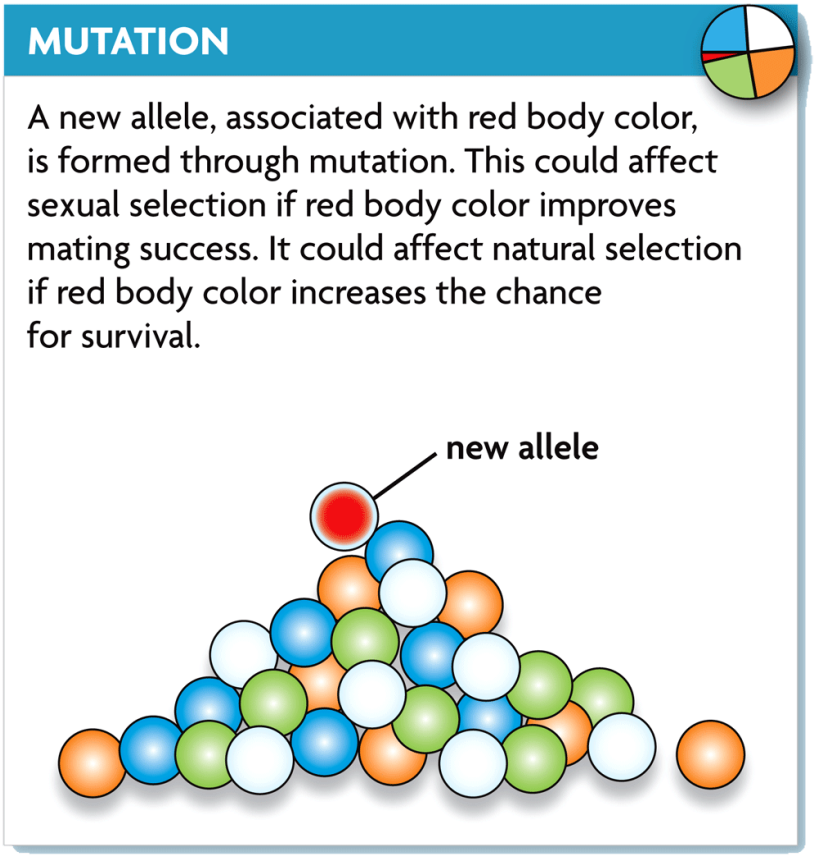
II. Population Genetics

C.Deviations from Hardy-Weinberg equilibrium

2. mutation

- Mutations produce the genetic variation needed for evolution.





II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

2. mutation

- A ***mutation*** is a change in an organism's DNA.
- A new mutation that is transmitted in gametes can immediately change the gene pool of a population by substituting the mutated allele for the older allele.
  - For any single locus, mutation alone does not have much quantitative effect on a large population in a single generation.
  - An individual mutant allele may have greater impacts later through increases in its relative frequencies as a result of natural selection or genetic drift.

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

2. mutation

- While mutations at an individual locus is a **rare event**, the cumulative impact of mutations at *all* loci can be significant.
  - Each individuals has thousands of genes, any one of which could experience a mutation.
  - Populations are composed of thousands or millions of individuals that may have experienced mutations.
- Over the long term, mutation is very important to evolution because it is the original source of genetic variation that serves as the raw material for natural selection.

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

2. mutation

1. Consider a population with:

$f(A) = p = 0.6 \qquad f(a) = q = 0.4$

2. Suppose 'a' mutates to 'A' at a realistic rate of:

$\mu = 1 \times 10^{-5}$

3. Well, what fraction of alleles will change?

'a' will decline by: $qm = .4 \times 0.00001 = 0.000004$

'A' will increase by the same amount.

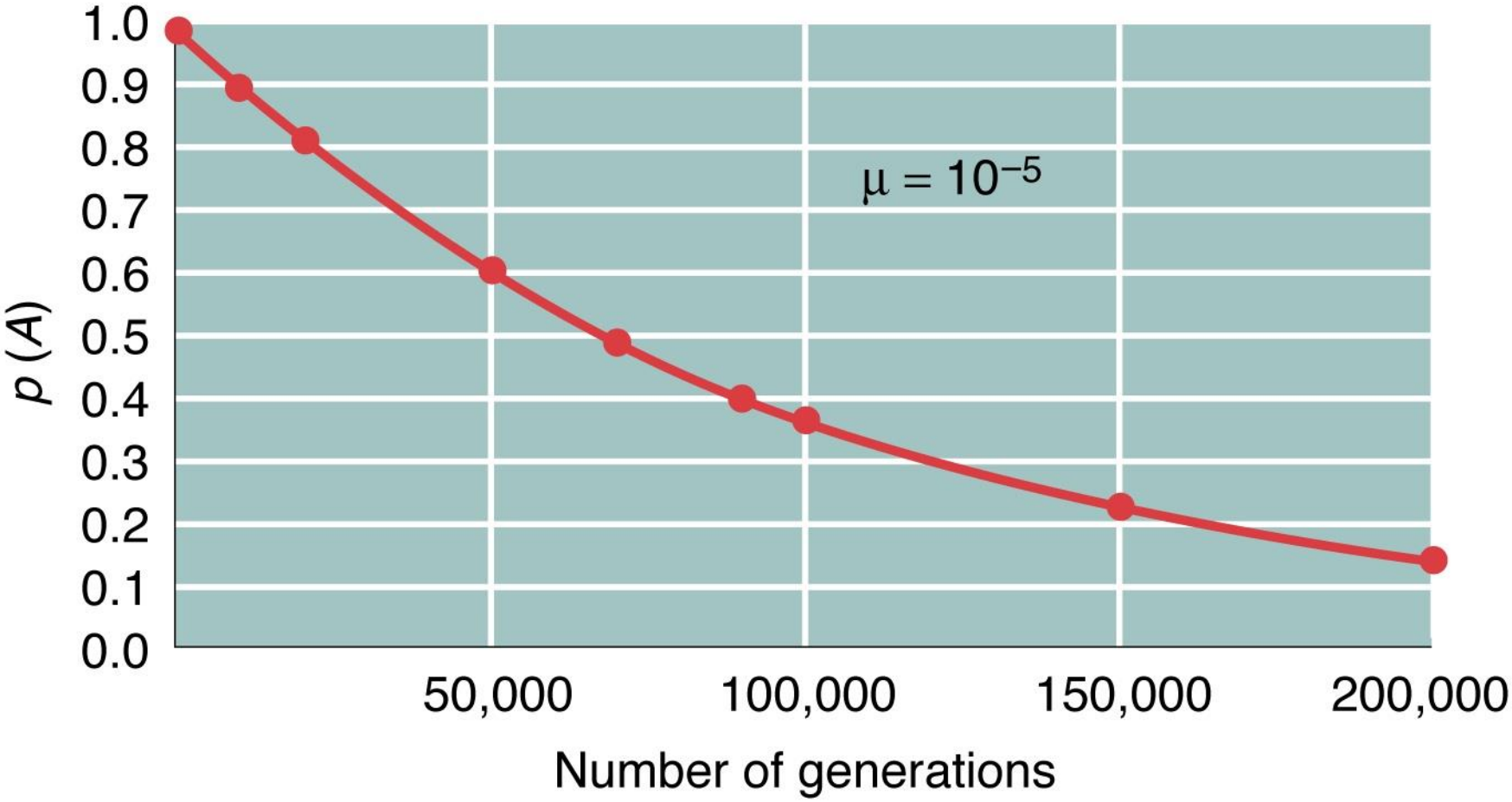
$f(A) = p1 = 0.600004 \qquad f(a1) = q = 0.399996$

1 generation

$\Delta p = -\mu p$

n generations

$p_n = p_{(0)}e^{-\mu n}$

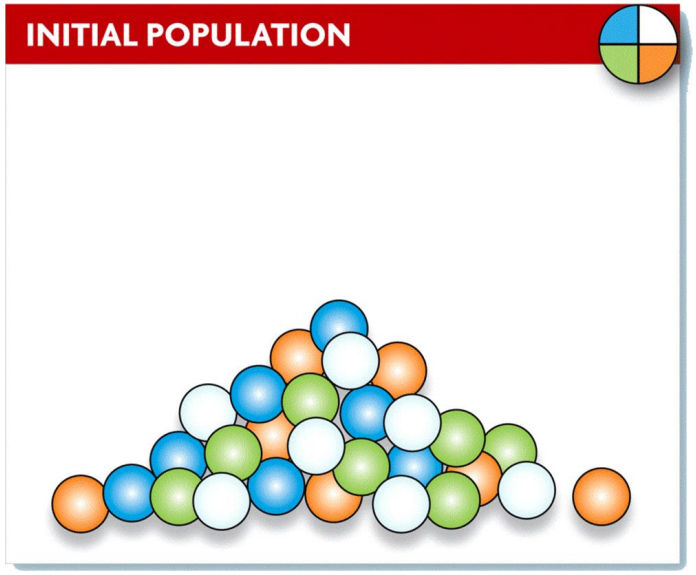


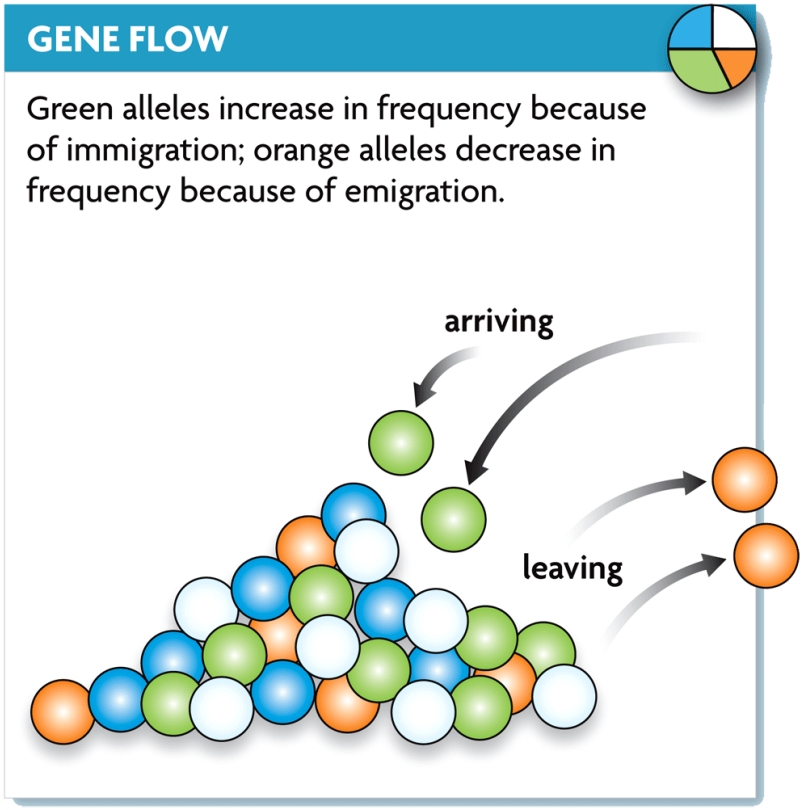
II. Population Genetics

C.Deviations from Hardy-Weinberg equilibrium

3. migration

- Migration= Gene flow moves alleles from one population to another.





II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

3. migration

Gene flow is genetic exchange due to migration of fertile individuals or gametes between populations.

Gene flow tends to reduce differences between populations.

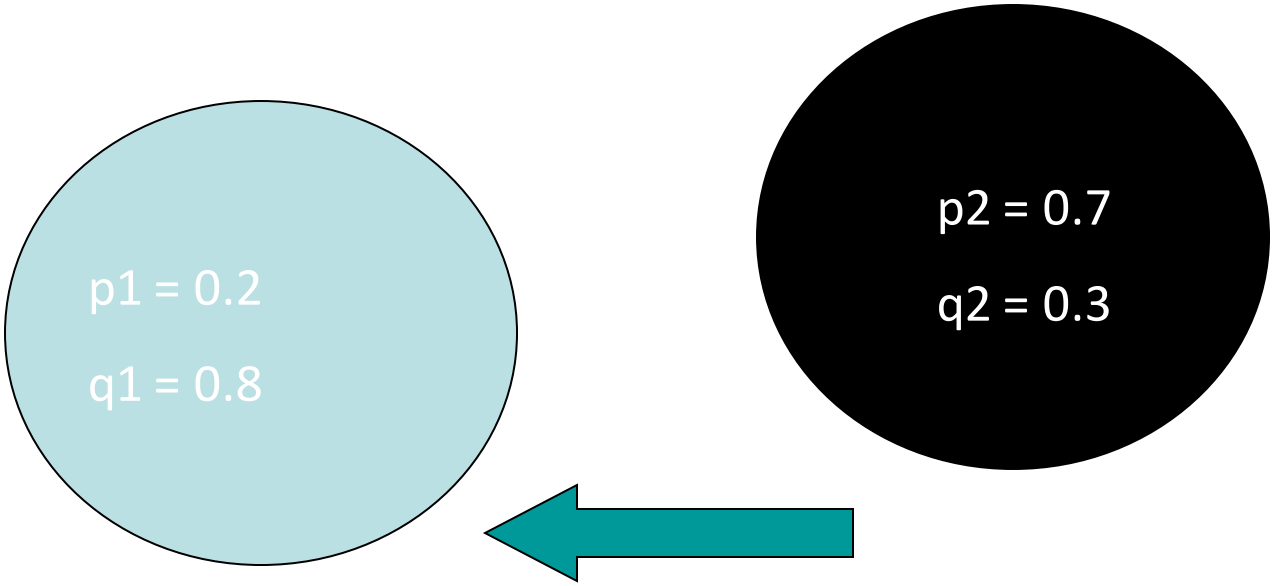
If extensive enough, gene flow can amalgamate neighboring populations into a single population with a common genetic structure.

The migration of people throughout the world is transferring alleles between populations that were once isolated, increasing gene flow

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

3. migration

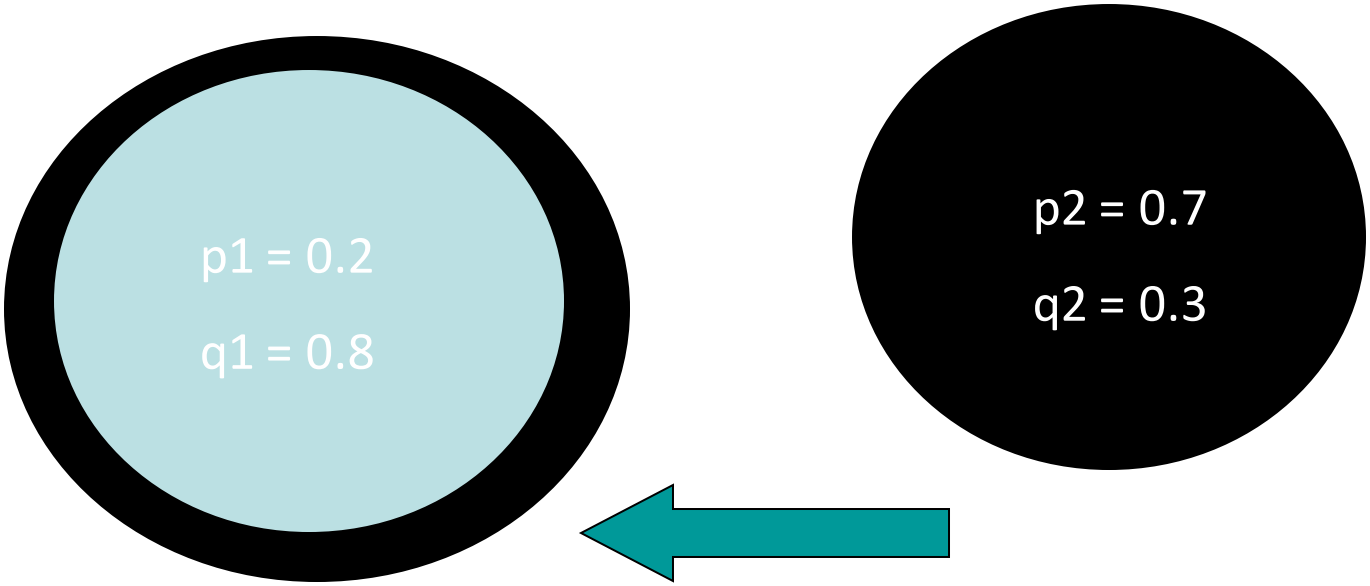


suppose migrants immigrate at a rate such that the new immigrants represent 10% of the new population

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

3. migration



M = 10%

p(new) = p1(1-m) + p2(m)

= 0.2(0.9) + 0.7(0.1)

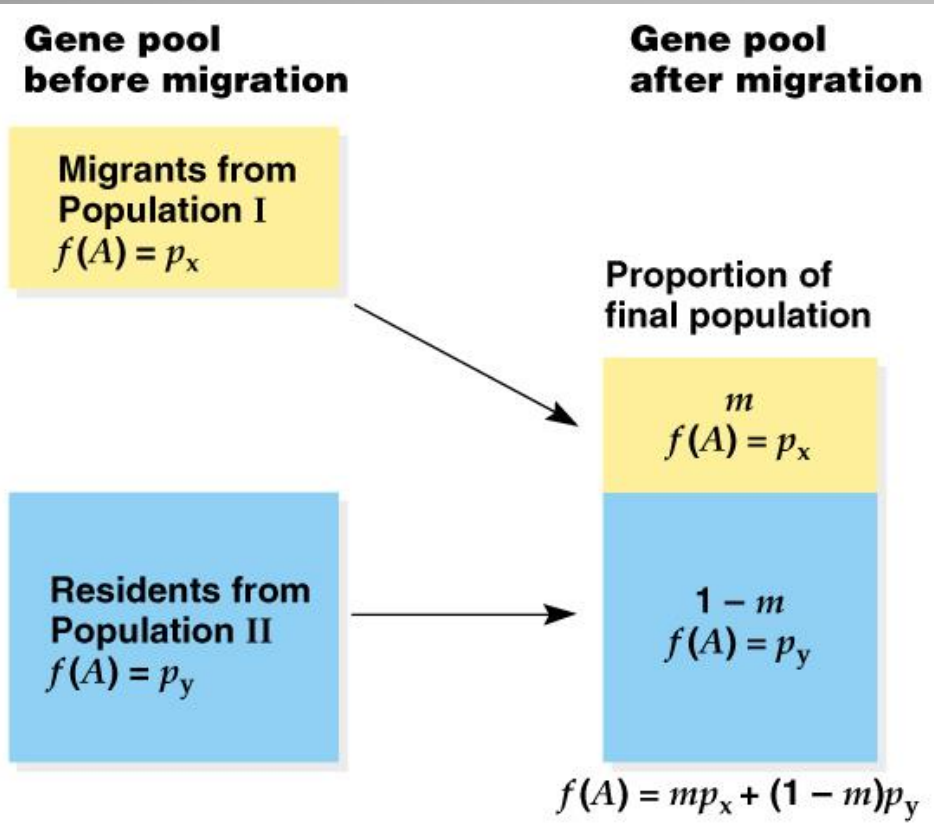
= 0.18 + 0.07 = 0.25

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

3. migration

Fig. 21.13 Theoretical model illustrating the effect of migration on the gene pool of a population



II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

3. migration

CONTINENT-ISLAND MODEL OF MIGRATION

Population on island (i) has migrants (m) from the mainland

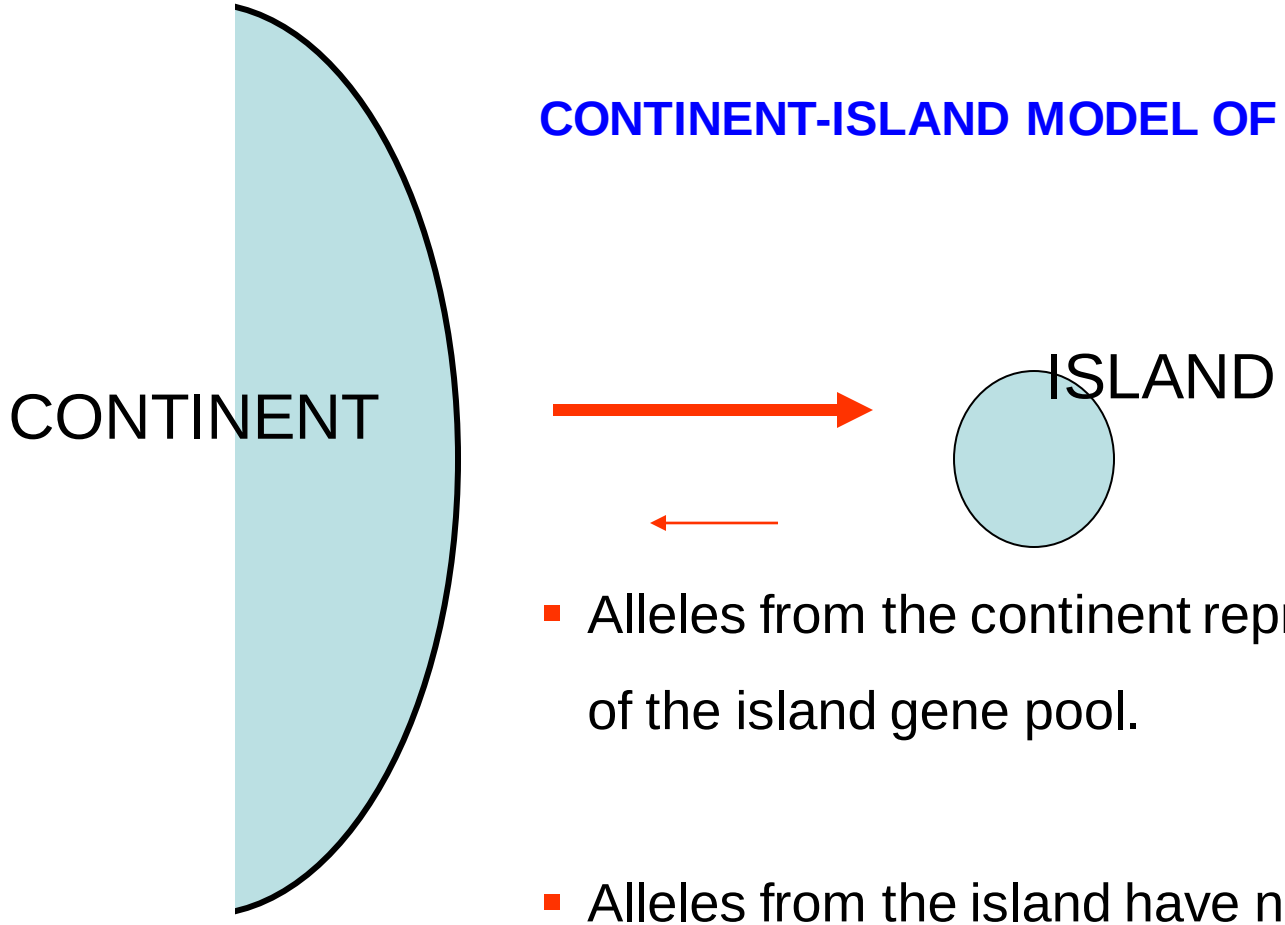
Allele frequency in the next generation is the weighted avergare of the two populations

$$p'_{(i)} = (1\text{-}m)p_{(i)} + (m)p_{(m)}$$

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

3. migration



- Alleles from the continent represent a large fraction of the island gene pool.

- Alleles from the island have negligible effect on gene frequencies in the continent.

THE HOMOGENIZING EFFECT OF MIGRATION IN A CONTINENT-ISLAND SYSTEM

Let:

Frequency of A on island = $\mathbf{p_I}$

Frequency of A on continent = $\mathbf{p_C}$

Proportion of island population from the continent = **m**

- Frequency of A on the island after migration:

$$\mathbf{p_I^* = (1-m)p_I + mp_C}$$

- Change on island from one generation to the next:

$$\mathbf{\Delta p_I = p_I^* - p_I = (1-m)p_I + mp_C - p_I}$$

- At equilibrium: $\mathbf{p_I = p_C}$

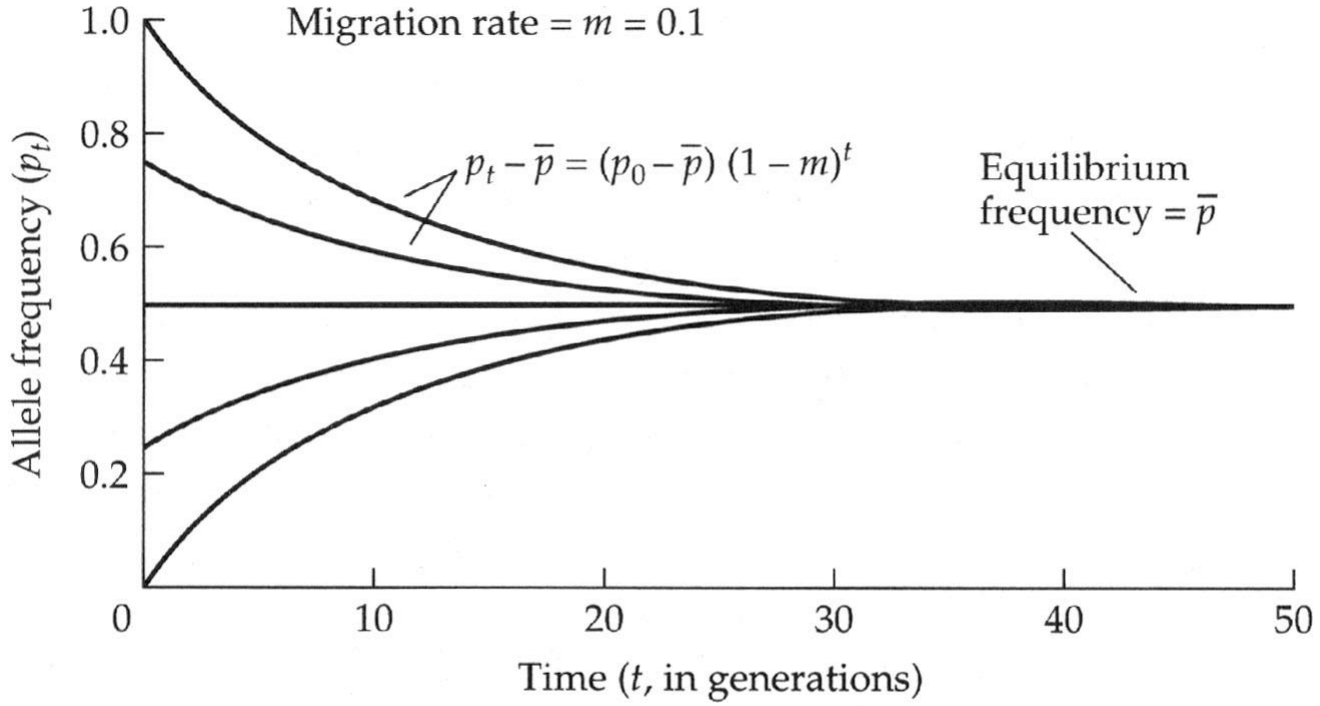
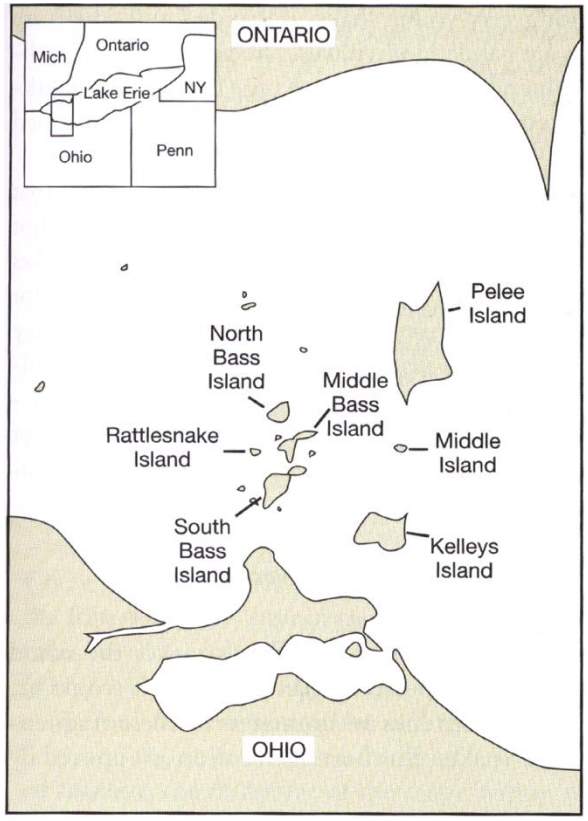


Figure 5.16 Change of allele frequency with time in five subpopulations exchanging migrants at the rate $m = 0.1$ per generation. Note the rapid convergence to a common equilibrium frequency.

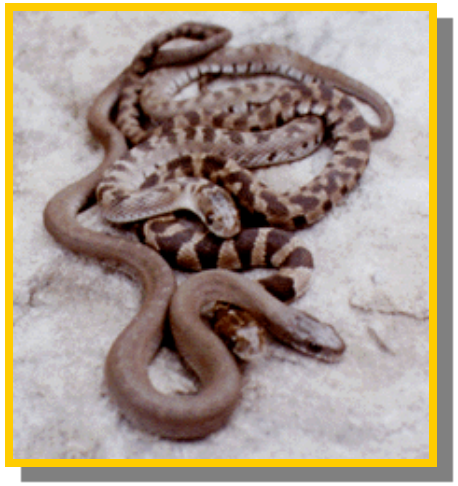
Case Study: Lake Erie Water Snakes

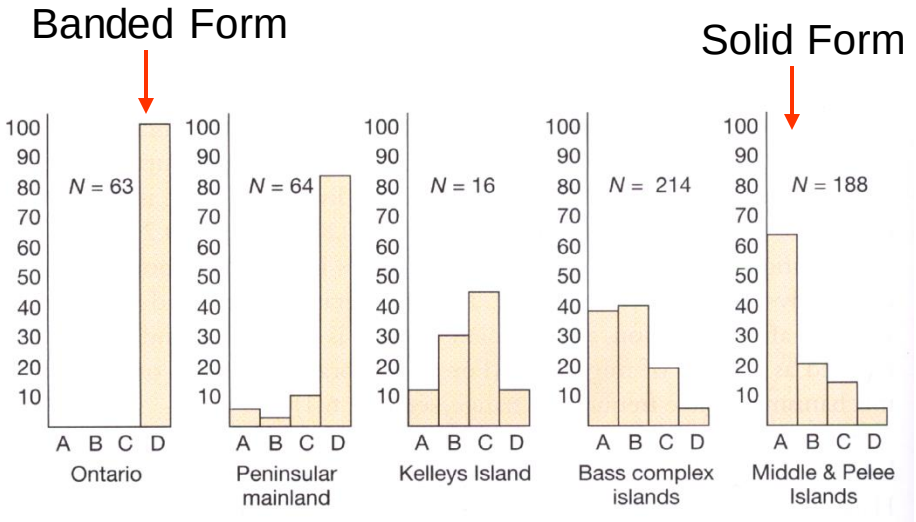
FROM: King & Lawson (1995)



Distribution on islands in Lake Erie

Nerodia sipedon





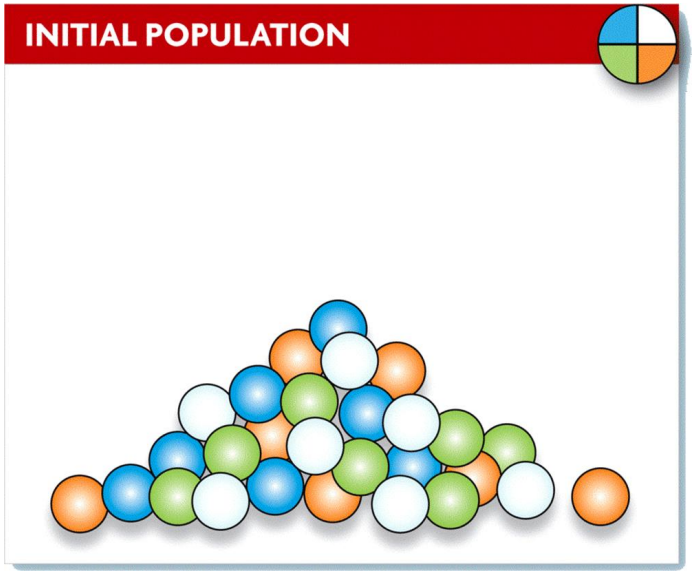
Frequency of color patterns on mainland and offshore islands.

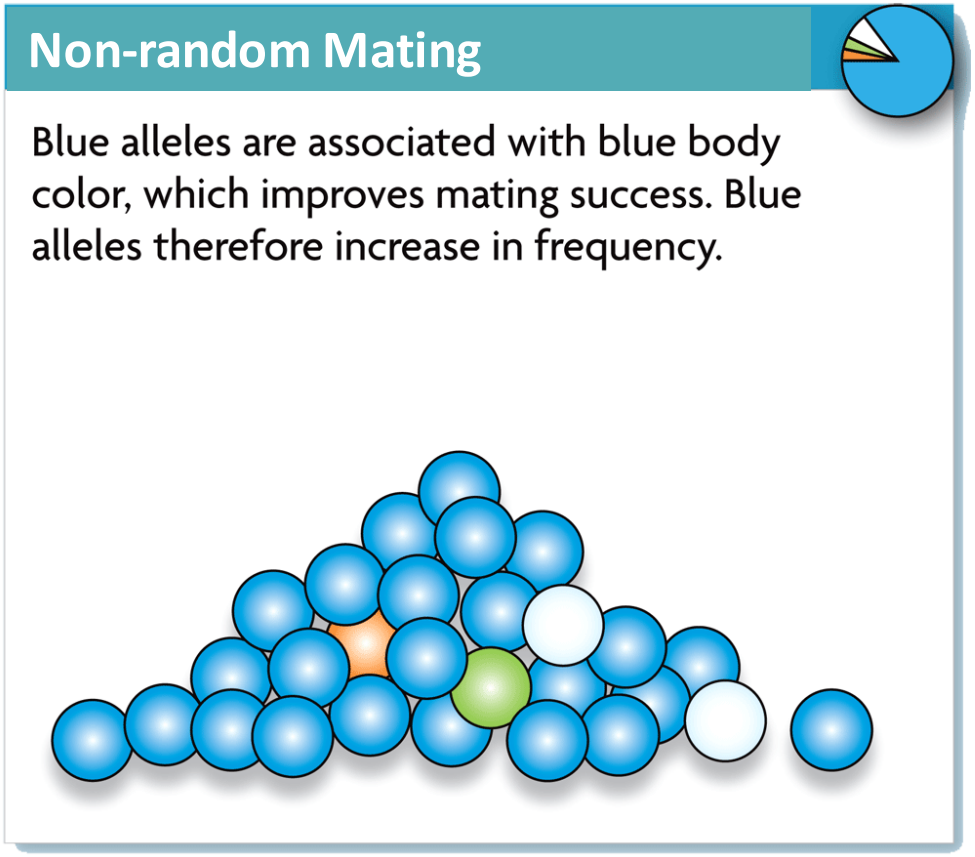
Lecture 6: Deviations from Hardy-Weinberg equilibrium(3)

II. Population Genetics

C.Deviations from Hardy-Weinberg equilibrium

- Sexual selection selects for traits that improve mating success.





II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

4. Non-random Mating

a. Positive Assortative Mating – "Like mates with Like"

	AA	Aa	aa
	0.2	0.6	0.2
offspring			
F1			

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

4. Non-random Mating

a. Positive Assortative Mating – "Like mates with Like"

	AA	Aa	aa
	0.2	0.6	0.2
offspring	ALL AA	1/4AA:1/2Aa:1/4aa	ALL aa
F1			

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

4. Non-random Mating

a. Positive Assortative Mating – "Like mates with Like"

	AA	Aa	aa
	0.2	0.6	0.2
offspring	ALL AA	1/4AA:1/2Aa:1/4aa	ALL aa
	0.2	0.15 + 0.3 + 0.15	0.2
F1	0.35	0.3	0.35

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

4. Non-random Mating

b. Inbreeding: Mating with Relatives

Decreases heterozygosity across the genome, at a rate dependent on the degree of relatedness among mates.

INBREEDING COEFFICIENT, *F*,

F = probability that two alleles in an individual are **identical by descent** (**IBD**).

Identical by descent vs. **identical in state**

Identity in state (homozygosity) does *not* necessarily imply recent **identity by descent**. (Conversely ...)

F for **fix**ation index:
homozygosity, or "**fixation**", results from inbreeding.

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

4. Non-random Mating

How does inbreeding affect the numbers of heterozygotes?

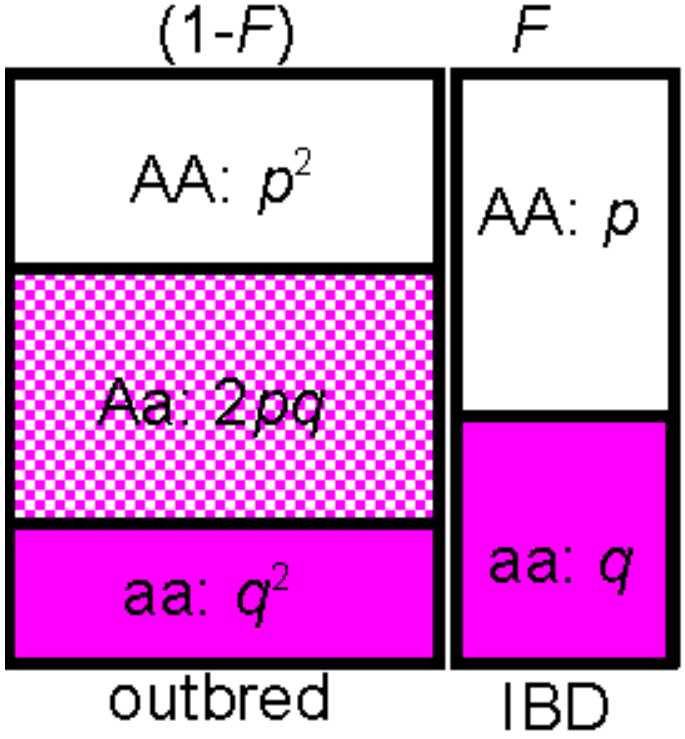
Consider alleles, **A**, and **a** with freqs p,q and inbreeding (**IBD**) at rate ***F***:
Frequency of homozygotes:

AA = $(1\text{-}F)p^2$ [outbred] + Fp [inbred]

$= p^2 + F(p\text{-}p^2)$

$= p^2 + Fp(1\text{-}p)$

$= p^2 + Fpq$



Similarly, frequency of other homozygotes,

aa = $q^2 + Fpq$

All genotype frequencies must add to 1
so the extra $2Fpq$ **AA** and **aa** homozygotes
must have come from the heterozygotes

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

4. Non-random Mating

genotype	*AA*	*Aa*	*aa*	(Sum)
frequency	p^2+Fpq	$2pq(1-F)$	q^2+Fpq	$p^2+Fpq+2pq-2Fpq+Fpq+q^2 = 1$

Inbreeding leads to a **reduction in heterozygosity**.
Heterozygosity (***Het***, i.e. fraction that are heterozygotes under inbreeding) is reduced by a fraction ***F*** compared with the outbred (Hardy-Weinberg) expectation $Het_{HW} = 2pq$:

$$Het = Het_{HW}(1 - F)$$

F measures reduction of heterozygosity, or ***heterozygote deficit*** compared to Hardy-Weinberg,
as well as ***probability of identity by descent***!

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

4. Non-random Mating

Problems of inbreeding:

Deleterious recessive alleles in most populations.

Few deleterious recessives per gene (usually $<< 10^{-3}$)

…but many deleterious alleles per genome.

You and I each carry about 1 strongly deleterious recessive mutation,

or "**lethal equivalent**".

When homozygous, these mutations cause problems,

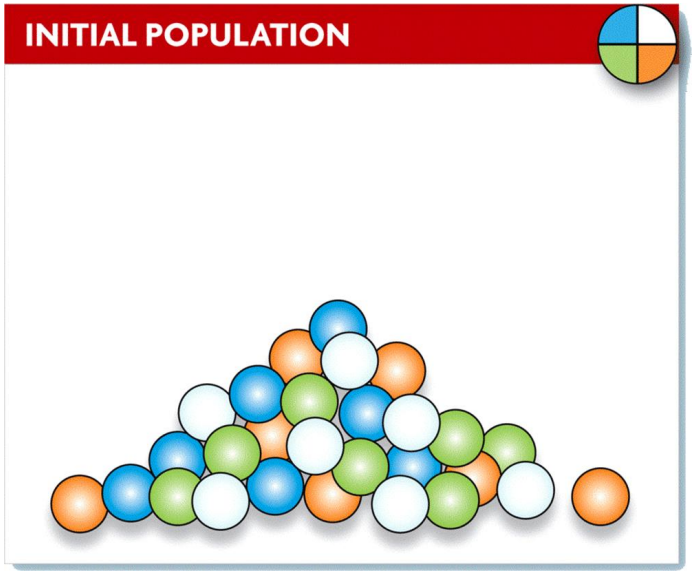
(**inbreeding depression**).

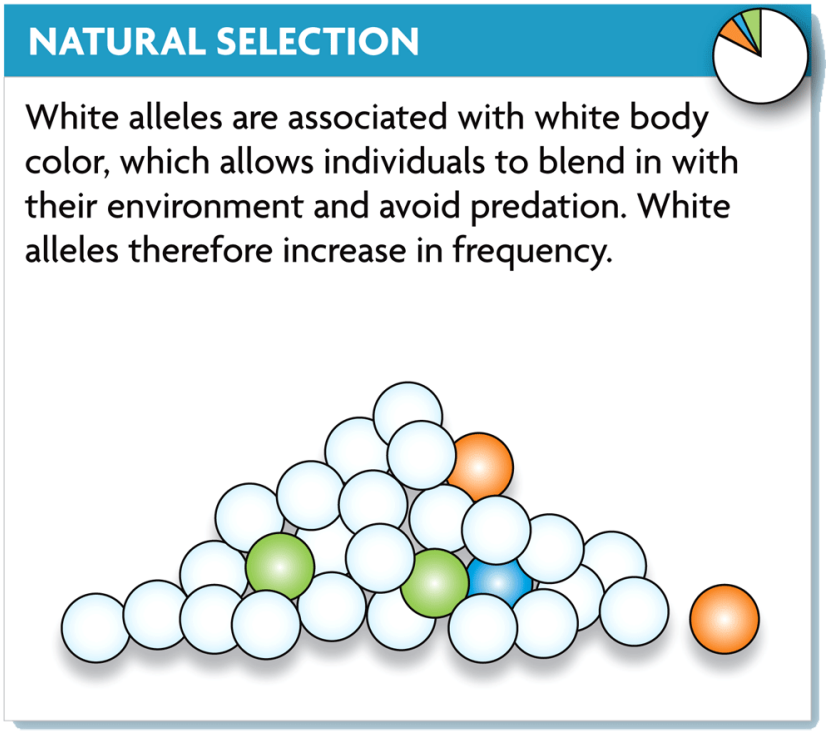
II. Population Genetics

C.Deviations from Hardy-Weinberg equilibrium

5. Natural Selection

- Natural selection selects for traits advantageous for survival.





II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

5. Natural Selection

- *Natural selection* is clearly a violation of the conditions necessary for the Hardy-Weinberg equilibrium.
  - The later expects that all individuals in a population have **equal ability to survive and produce** viable, fertile offspring.
  - **Fitness = Survival x Reproductive success**
  - However, in a population with variable individuals, **natural selection** will lead some individuals to leave more offspring than others.

- Natural selection maintains favorable genotypes in a population.

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

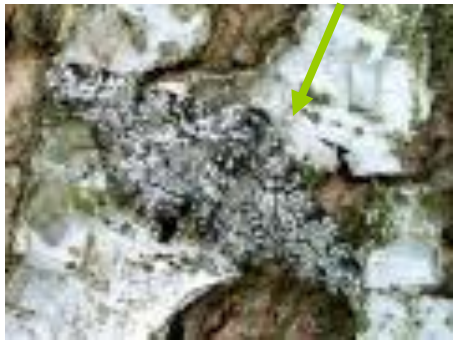
5. Natural Selection

Figure 21.16 Industrial Melanism in the peppered moth

Biston betularia - the peppered moth

Dark morph on sooty tree

Grey moth on lichen



Each morph on the "wrong" tree

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

5. Natural Selection

Released in rural woodland

	grey	melanic
# released	496	473
# recaptured	62	30
% recaptured	**12.5**	**6.3**

Released in industrial region

	grey	melanic
# released	137	447
# recaptured	18	123
% recaptured	**13.1**	**27.5**

II. Population Genetics

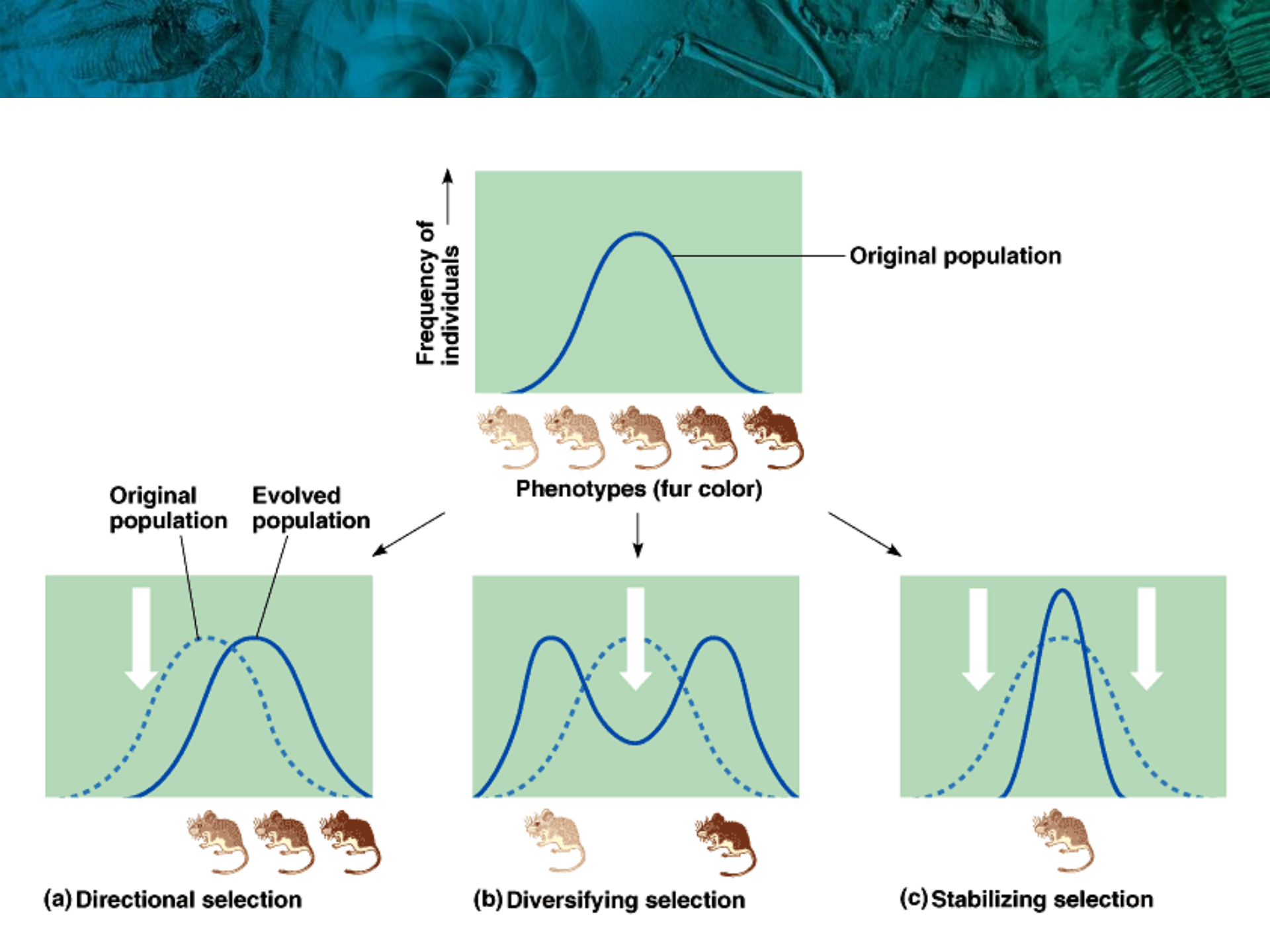
C. Deviations from Hardy-Weinberg equilibrium

5. Natural Selection

Table 21.10 General Method of Determining Change in Allele Frequency Caused by Natural Selection

	Genotypes		
	A^1A^1	A^1A^2	A^2A^2
Initial genotype frequencies	p^2	$2pq$	q^2
Fitness[a]	w_{11}	w_{12}	w_{22}
Frequency after selection	p^2w_{11}	$2pqw_{12}$	q^2w_{22}
Relative genotype frequency after selection[b]	$P' = \frac{p^2\overline{w}_{11}}{\overline{w}}$	$H' = \frac{2pq\overline{w}_{12}}{\overline{w}}$	$Q' = \frac{q^2\overline{w}_{22}}{\overline{w}}$
Allele frequency after selection $= p' = P' + {}^1/_2(H')$			
	$q' = 1 - p'$		
Change in allele frequency caused by selection $= \Delta p = p' - p$			

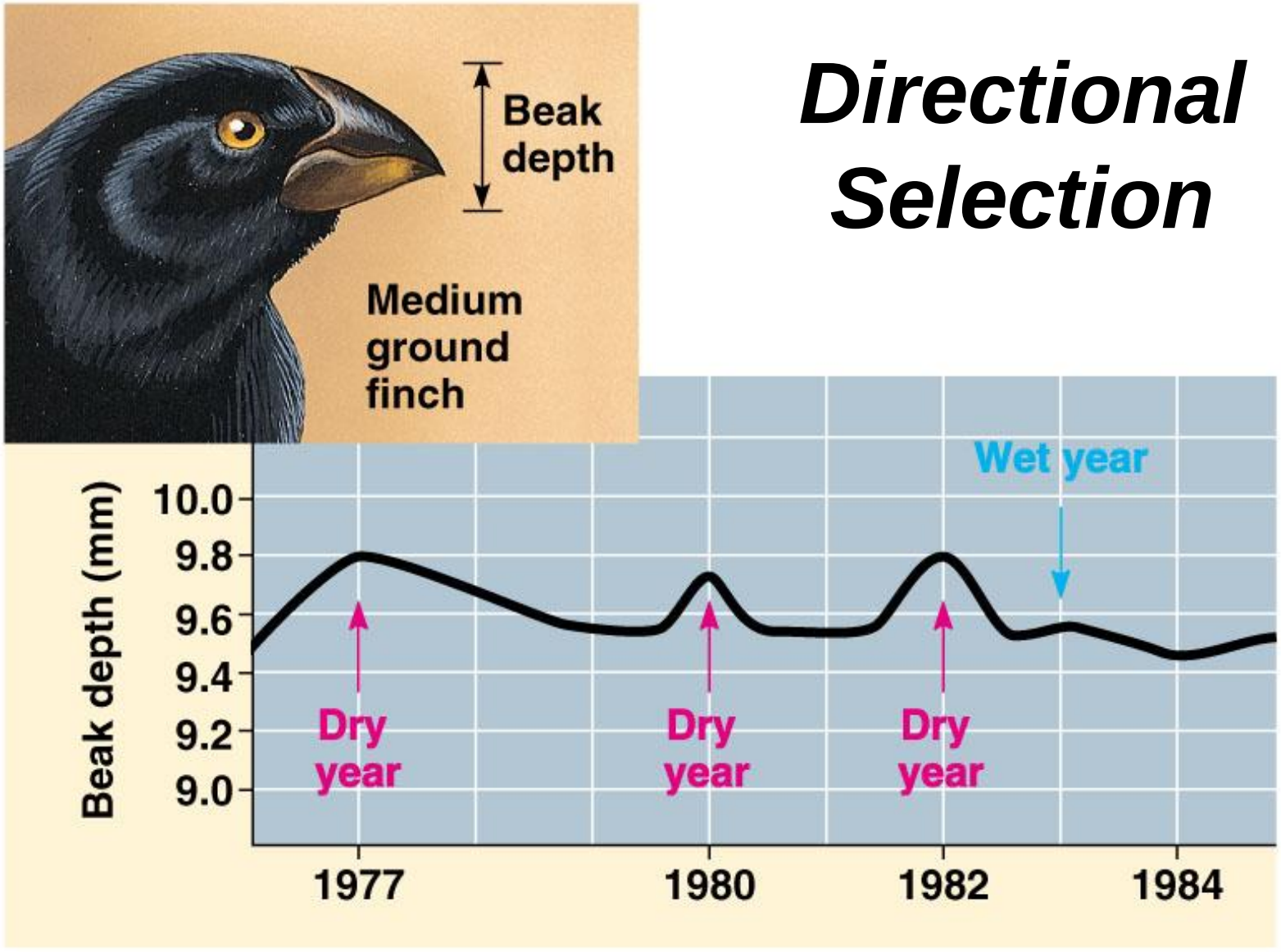
[a]For simplicity, fitness in this example is considered to be the probability of survival. Change in allele frequency caused by differences in the number of offspring produced by the genotypes is calculated in the same manner.
[b]$\overline{w} = p^2w_{11} + 2pqw_{12} + q^2w_{22}$


Frequency of individuals
Original population
Phenotypes (fur color)
Original population
Evolved population
(a) Directional selection
(b) Diversifying selection
(c) Stabilizing selection

Modes of Natural Selection

- ***Directional Selection***
  - *Favors individuals at one end of the phenotypic range*
  - *Most common during times of environmental change or when moving to new habitats*
- ***Diversifying selection***
  - *Favors extreme over intermediate phenotypes*
  - *Occurs when environmental change favors an extreme phenotype*

Directional Selection



Modes of Natural Selection

- *Stabilizing Selection*
  - *Favors intermediate over extreme phenotypes*
  - *Reduces variation and maintains the current average*
  - *Example: Human birth weight*

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

6. Heterozygote Advantage

- *Favors heterozygotes (Aa)*
- *Maintains both alleles (A,a) instead of removing less successful alleles from a population*
- *Sickle cell anemia*

  > Homozygotes exhibit severe anemia, have abnormal blood cell shape, and usually die before reproductive age.

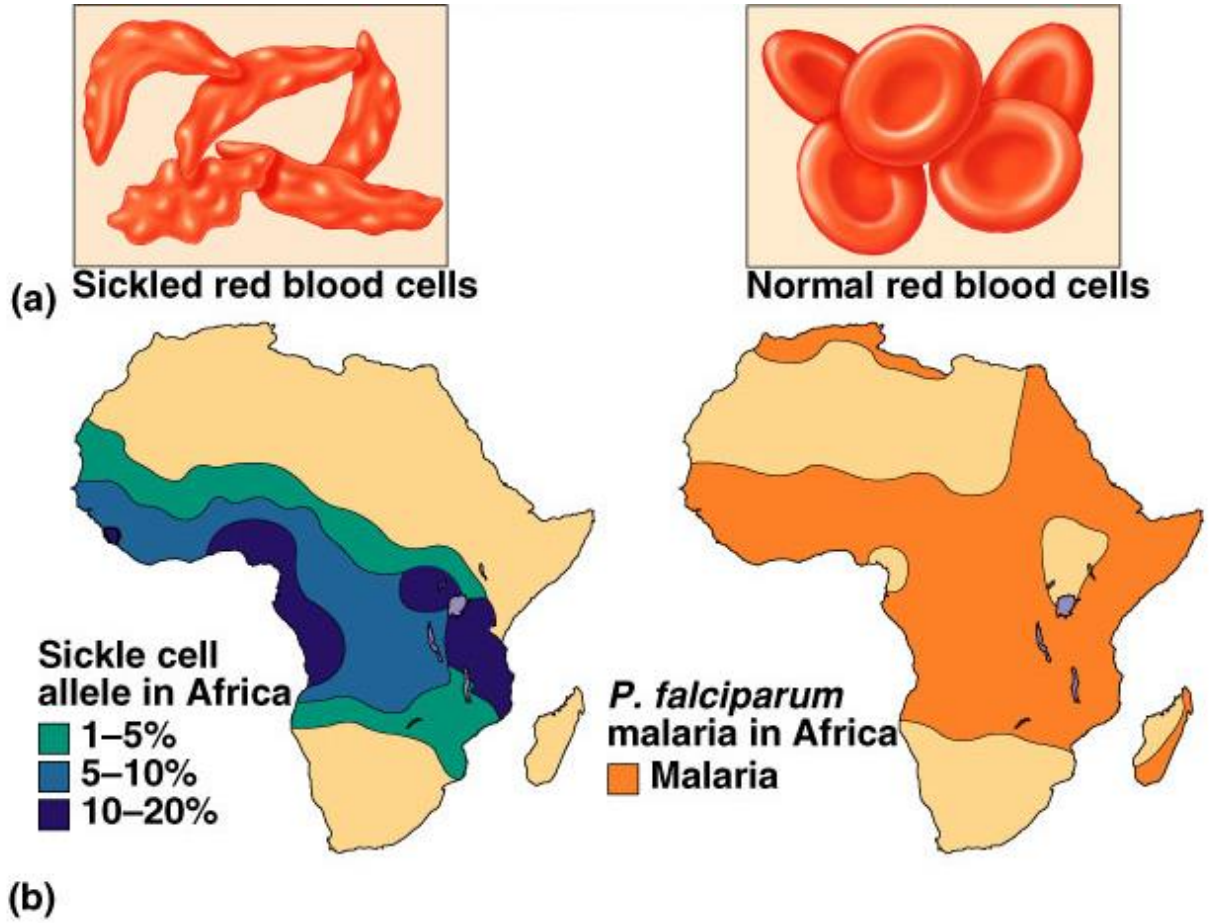
  > Heterozygotes are less susceptible to malaria

II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

6. Heterozygote Advantage

- **Sickle Cell and Malaria**



II. Population Genetics

C. Deviations from Hardy-Weinberg equilibrium

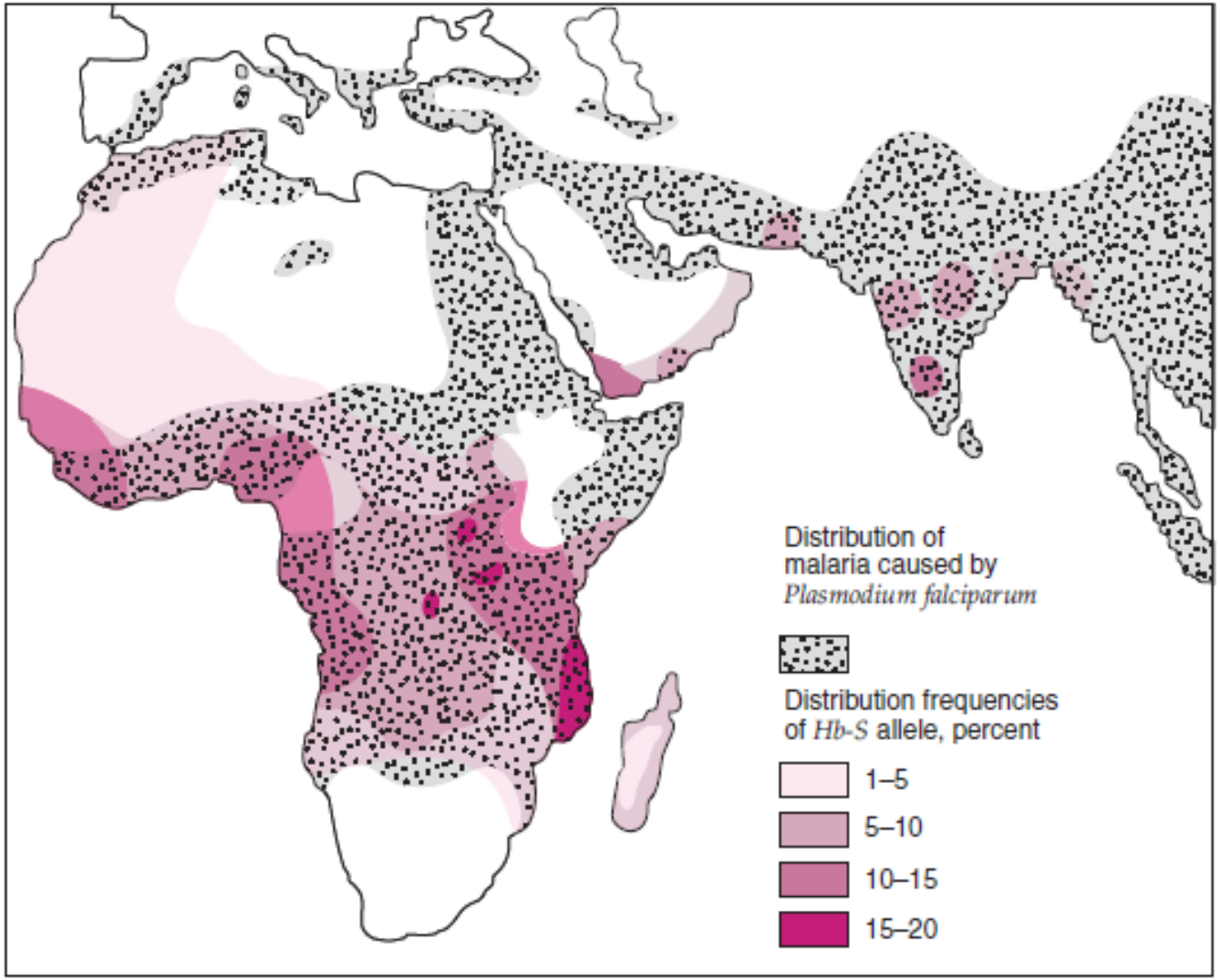


Figure 21.19

The distribution of malaria caused by the parasite *Plasmodium falciparum* coincides with distribution of the *Hb-S* allele for sickle-cell anemia. The frequency of *Hb-S* is high in areas where malaria is common because *Hb-A/Hb-S* heterozygotes are resistant to malarial infection.

Lecture 7: Quantitative Genetics

Various human eye colors.

The Nature of Continuous Traits

Fig. 22.1 Discontinuous distribution of shell color in the snail *Cepaea nemoralus* from a population in England

- For discontinuous traits, a simple relationship usually exists between the genotype and the phenotype.
-
- In most cases, the effects of variant alleles at the single locus are observable at the level of the organism, so the phenotype can be used as a quick assay for the genotype.

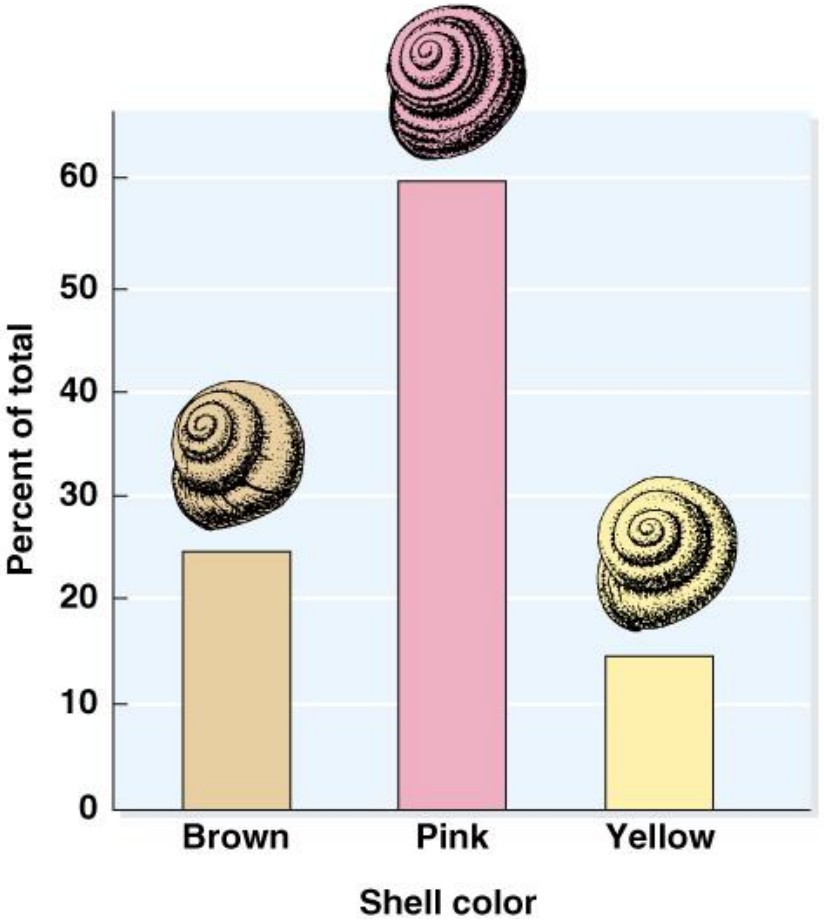
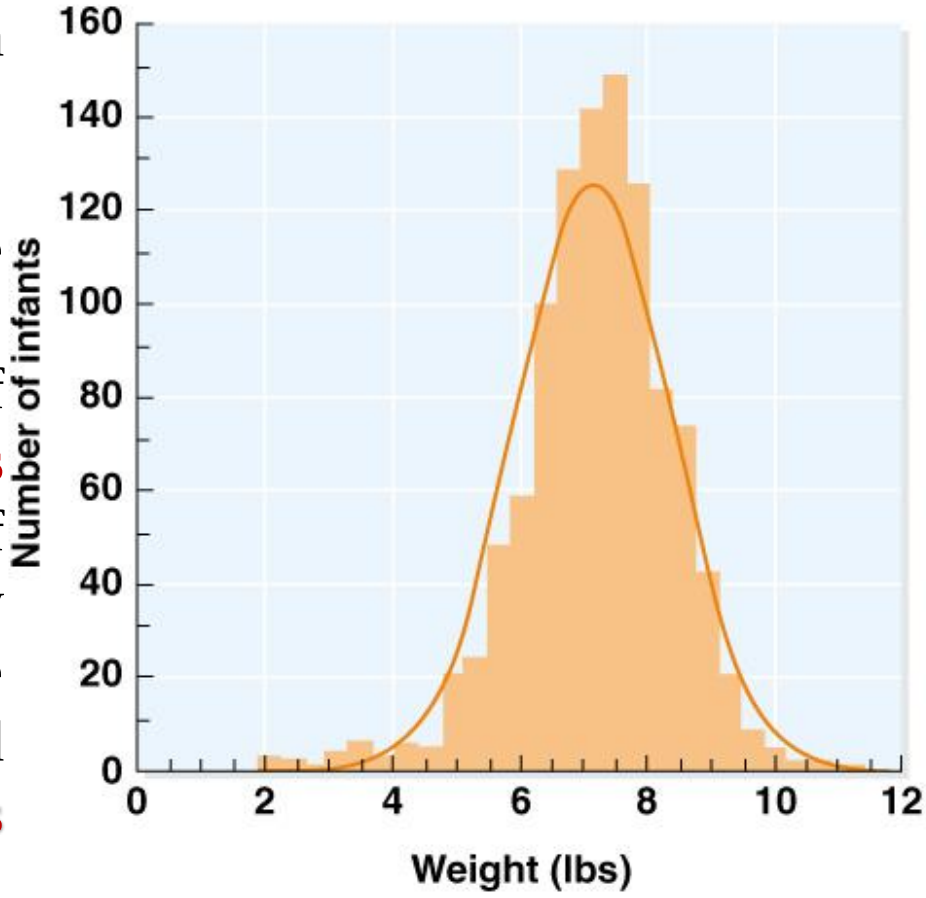


Fig. 22.2 Distribution of birth weight of babies (males + females) born to teenagers in Portland, Oregon, in 1992

Many traits, such as human birth weight (Fig. 22.2) and adult height, protein
content in corn…, exhibit a wide range of possible phenotypes.

Traits with a continuous distribution of phenotypes, are called **continuous traits**. Since the phenotypes of continuous traits must be described by quantitative measures, such traits are also known as quantitative traits, and the field of quantitative genetics studies the inheritance of these traits.



The Nature of Continuous Traits

1. Statistical study of continuous traits began with human traits such as height, weight and mental traits, even before Mendel's principles were understood.
2. Galton and Pearson (late 1800s) showed that these traits are statistically linked between parents and offspring, but could not determine the mode of transmission.
3. Johannsen (1903) showed that continuous variation in bean seed weight is partly genetic and partly environmental.
4. Nilsson-Ehle proposed that continuous variation in wheat results from multiple genes segregating by Mendelian principles.
5. Fisher demonstrated that mathematical models of Mendelian and population genetics also apply to traits controlled by multiple gene loci.

Why Some Traits Have Continuous Phenotypes

1. Multiple traits arise in several ways:
   a. When a trait is influenced by many loci (polygenic), a range of phenotypes results from the numerous genotypes. For example:
      i. When a single locus with two alleles determines a trait, there are three possible genotypes (AA, Aa and aa).
      ii. With two loci, each with two alleles, there are nine possible genotypes.
      iii. In general, the number of genotypes is $\mathbf{3^n}$, where **n** is the number of loci with two alleles. If the number of alleles is higher, the number of genotypes will be even larger.
      iv. If every genotype in a polygenic trait produces a different phenotype and the differences between phenotypes are slight, the trait appears to be continuous.
      v. More often, several genotypes produce the same phenotype for a polygenic trait. Reasons for this include:
         (1) **Dominance**, producing the same phenotype in both heterozygous and homozygous dominant individuals.
         (2) **Epistasis**, resulting from control of the expression of other loci.

Why Some Traits Have Continuous Phenotypes

2. A range of phenotypes is also produced when environmental factors affect the trait. Each genotype will have a range of possible phenotypes, the norm of reaction.

3. Most traits are influenced by both multiple genotypes and environmental factors, and are thus **multifactorial**. The rules of transmission genetics and gene function still apply. Understanding the role of each gene is difficult, and so quantitative genetics is employed to analyze continuous traits.

Questions Studied in Quantitative Genetics

1. What role is played by genetics, and what role by environment?
2. How many genes are involved in producing phenotypes of the trait?
3. Do some genes play a major role in determining phenotype, while others modify it only slightly, or are the contributions equal?
4. Do the alleles interact with each other to produce additive effects?
5. What changes occur when there is selection for a phenotype, and do other traits also change?
6. What method of selecting and mating individuals will produce desired phenotypes in the progeny?

Statistical Tools

1. Genes are always expressed in an environmental **(E)** context, and without genes **(G)** there would be nothing to express.
2. When multiple genes and environmental factors influence a trait, the relationship between individual loci and their contribution to the phenotype **(P)** may be obscured.
3. The same rules of transmission genetics and gene function still apply,
4. How much of a variation in phenotype **(V_P)** is due to genetic variation **(V_G)** and how much to environmental variation **(V_E)**?

This can be expressed: **$V_P = V_G + V_E$.**

1. To work this equation, variation must be measured and then partitioned into genetic and environmental components.

Samples and Populations

1. It is difficult to collect data for each individual in a large population.
2. Sampling of a subset is an alternative method.
    a. The sample must be large enough to minimize chance differences between the sample and the population.
    b. The sample must be a random subset of the population.
3. Birth weight in humans is an example (Figure 22.2).

Distributions

1. Phenotypes are not easily grouped into classes when a continuous range occurs. Instead, a frequency distribution is commonly used, showing the **proportion** of individuals that fall within a **range of phenotypes**.
2. In a frequency distribution, **the classes** consist of **specified ranges** of the phenotype, and **the number of individuals in each class is counted**.
   a. An example is Johannsen's study of seed weight in the dwarf bean (*Phaseolus vulgaris*).
   b. A histogram is used **to show** the distribution of individuals into phenotypic classes.
   c. Tracing the outline of the histogram gives a **curve** characteristic of the frequency distribution.

3. Continuous traits often show a **bell-shaped curve (normal distribution)**, due to influences of multiple genes and environmental factors.

Table 22.2 Weight of 5,494 F_2 Beans (Seeds of *Phaseolus vulgaris*) Observed by Johannsen in 1903

Weight (mg)	50–150	150–250	250–350	350–450	450–550	550–650	650–750	750–850	850–950
(Midpoint of range)	(100)	(200)	(300)	(400)	(500)	(600)	(700)	(800)	(900)
Number of beans	5	38	370	1,676	2,255	928	187	33	2

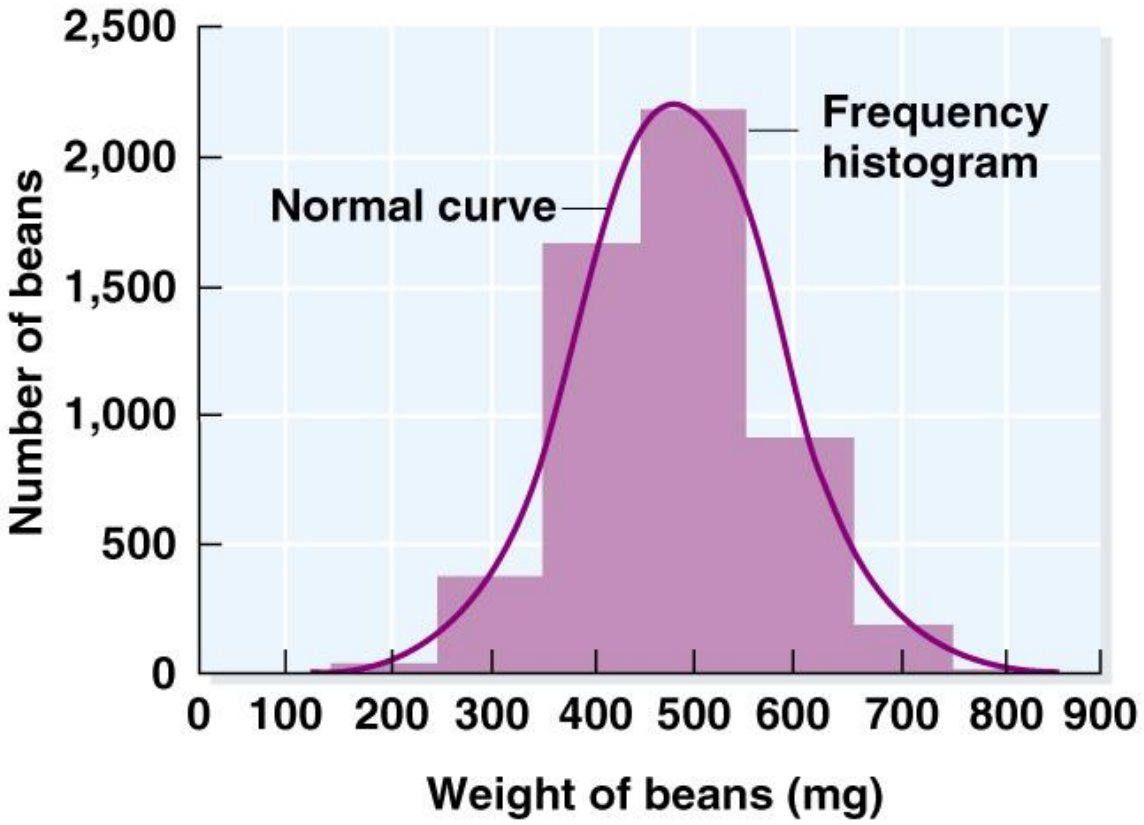


Fig. 22.3 Frequency histogram for bean weight in *Phaseolus vulgaris*

Lecture 8: Analysis of Variance

The Mean $\bar{x}$

1. Frequency distribution of a phenotypic trait can be summarized with two statistics, the mean and the variance.
2. **The mean (average)** represents the center of the phenotype distribution, and is calculated simply by adding all individual measurements and then dividing by the number of measurements added.

$$\bar{x} = (\Sigma x_i) / n$$

3. An example is body length of spotted salamanders (Table 22.3).
4. Another example is East's study of tobacco flower length in genetic crosses.
   a. He crossed a short-flowered strain (mean length of 40.4 mm) with a long-flowered strain (mean length of 93.1 mm).
   b. The F_1 progeny (173 plants) had a mean flower length of 63.5 mm.

Table 22.3 Sample Calculations of the Mean, Variance, and Standard Deviation for Body Length of 10 Spotted Salamanders from Penobscot County, Maine

Body Length (x_i) (mm)	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
65	$(65 - 57.1) = 7.9$	$7.9^2 = 62.41$
54	$(54 - 57.1) = -3.1$	$-3.1^2 = 9.61$
56	$(56 - 57.1) = -1.1$	$-1.1^2 = 1.2$
60	$(60 - 57.1) = 2.9$	$2.9^2 = 8.41$
56	$(56 - 57.1) = -1.1$	$1.1^2 = 1.21$
55	$(55 - 57.1) = -2.1$	$2.1^2 = 4.41$
53	$(53 - 57.1) = -4.1$	$-4.1^2 = 16.81$
55	$(55 - 57.1) = -2.1$	$-2.1^2 = 4.41$
58	$(58 - 57.1) = 0.9$	$0.9^2 = 0.81$
59	$(59 - 57.1) = 1.9$	$1.9^2 = 3.61$
$\sum x_i = 571$		$\sum (x_i - \bar{x})^2 = 112.9$

$$\text{Mean} = \bar{x} = \frac{\sum x_i}{n} = \frac{571}{10} = 57.1$$

$$\text{Variance} = s_x^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} = \frac{112.9}{9} = 12.54$$

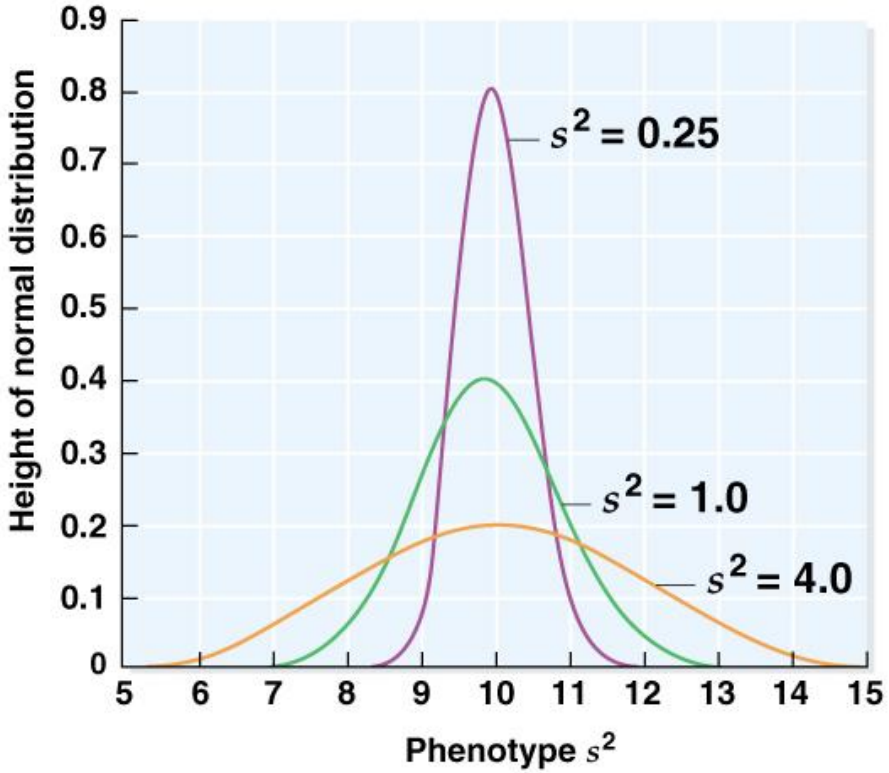
$$\text{Standard deviation} = s_x = \sqrt{12.54} = 3.54$$

The Variance (s^2) and the Standard Deviation

1. Variance is the measure of how much the individual measurements spread out around the mean (how variable they are).
   a. Two sets of measurements may have the same mean, but different variances (Figure 22.4).
   b. The variance (s^2) is the average squared deviation from the mean. To calculate s^2:
   c. Standard deviation is used more often than variance, because it shares the same units as the original measurements (rather than units^2 as in variance). Standard deviation is the square root of the variance.
   d. Table 23.2 shows sample calculations for variance and standard deviation.

$$\text{Variance} = s^2 = \frac{\sum(x_i - \bar{x})^2}{n - 1}$$

Fig. 22.4 Graphs showing three distributions with the same mean but different variances



A broad curve indicates a large variability in the measurements and a large standard deviation.

A narrow curve implies little variability and a small standard deviation.

e. When mean and standard deviation are known, a **theoretical normal distribution** is specified. Its shape is shown in Figure 22.5. In a theoretical normal distribution:

   i. One standard deviation above or below the mean (± 1s) includes **66%** of the individual observations.

   ii. Two standard deviations (± 2s) includes 95% of the individual values.

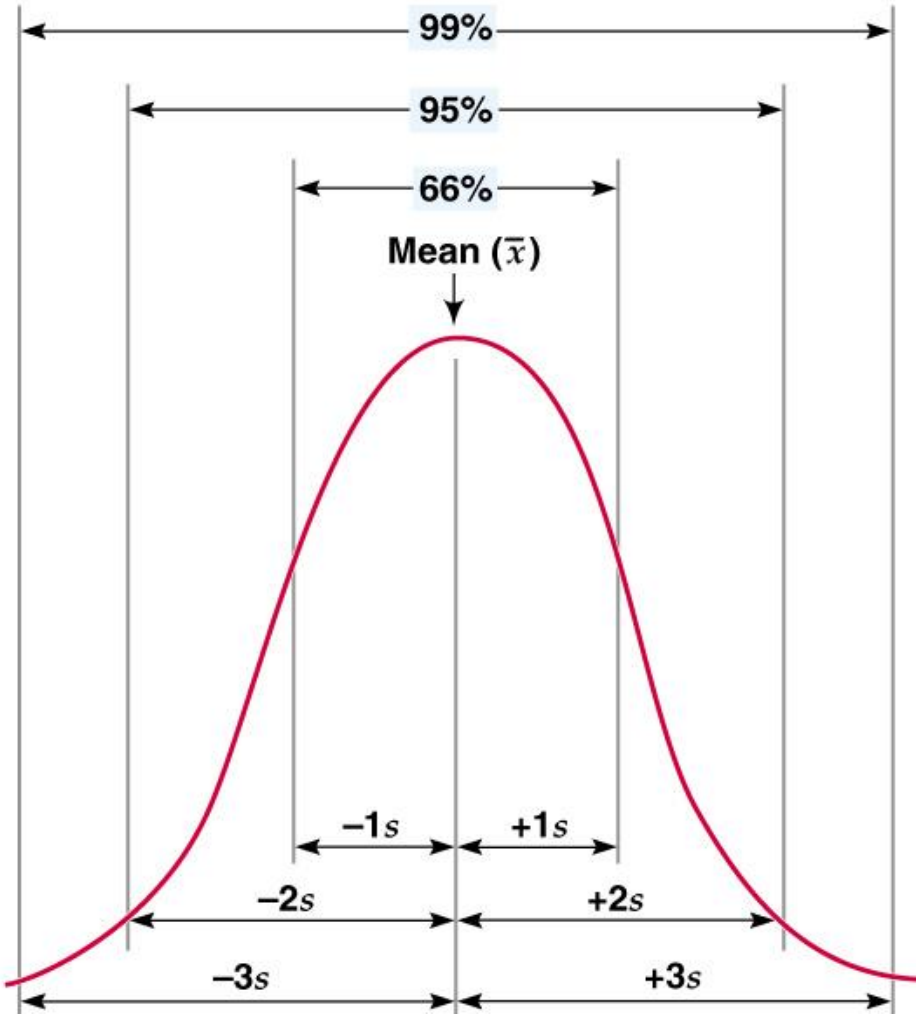
f. Analysis of variance is a statistical technique used to help partition variance into components.

2. Variance and standard deviation provide information about the phenotypes of a group. In the tobacco flower example:

   a. The original cross of short-flowered with long-flowered produced an F_1 with a mean flower length of 63.5 mm, intermediate to the parents.

   b. The F_2 had a mean of 68.8 mm, very similar to the F_1. But the F_2 had a variance of 42.2 mm^2, while F_1 variance was only 8.6 mm^2, indicating that more phenotypes occur among the F_2 than among the F_1.

Fig. 22.5 Normal distribution curve showing proportions of the data in the distribution that are included within certain multiples of standard deviation



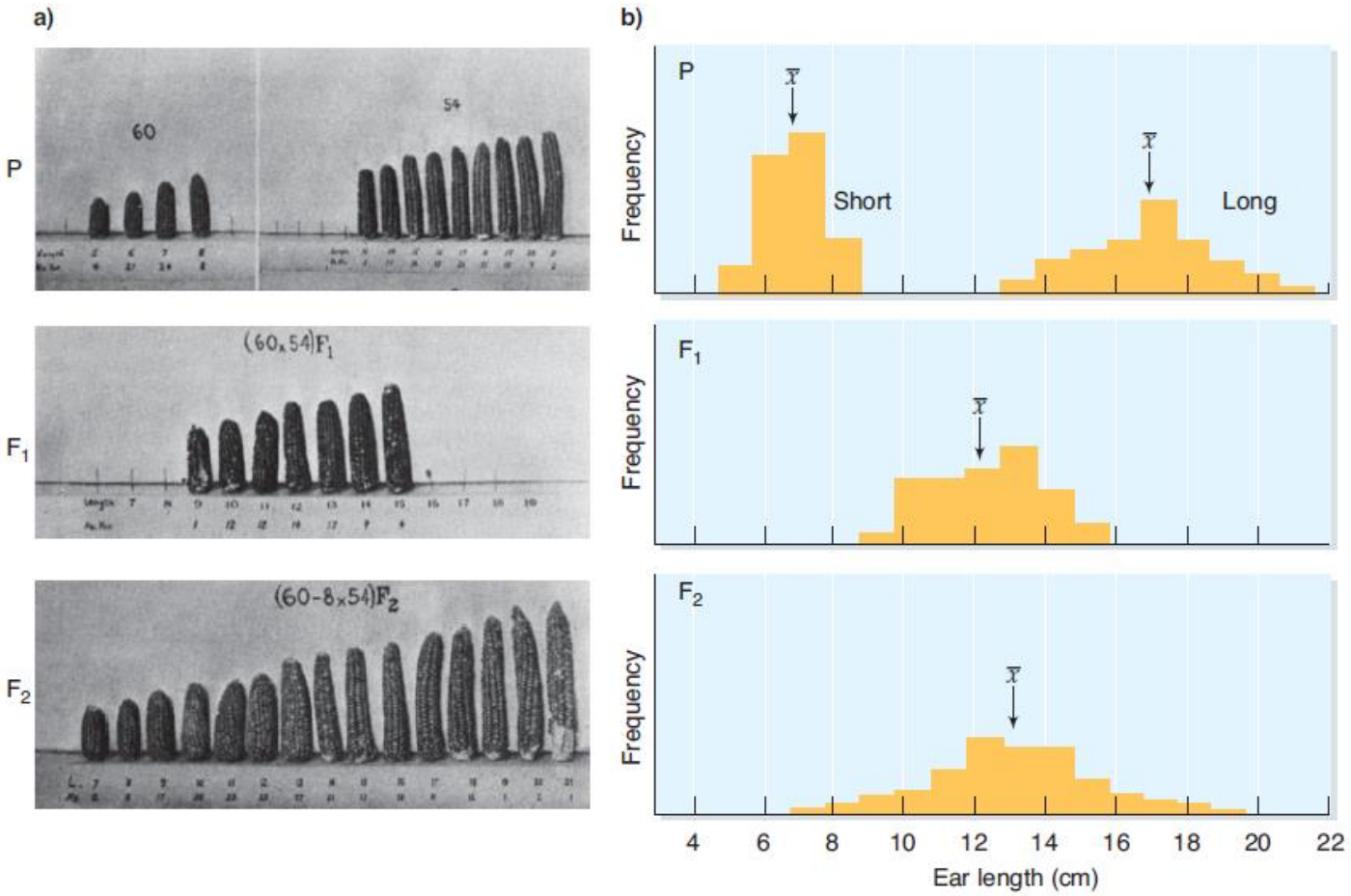
Lecture 9: Polygenic Inheritance

Inheritance of Ear Length in Corn

1. Emerson and East (1913) experimented with two **pure-breeding** strains of corn.
   a. Each strain shows little variation in ear length.
      i. The Black Mexican sweet corn variety has short ears (mean length 6.63 cm) with a standard deviation (s) of 0.816.
      ii. Tom Thumb popcorn has long ears (mean length 16.80 cm), and s = 1.887.
   b. The two strains were crossed, and the F_1 plants interbred (Figure 22.9).
      i. The mean ear length in the F_1 is 12.12 cm, approximately intermediate, and s = 1.519.
      ii. Since both parents were true-breeding, all F_1 plants should have the same heterozygous genotype, and any variation in length would be due to environmental factors.
      iii. The mean ear length of the F_2 is 12.89 cm, very similar to the F_1, but in the F_2, s = 2.252, reflecting its greater variability.
      iv. It is expected that the environment would have the same effect on the F_2 that it had on the P and F_1 plants, but it would not be expected to have more effect.
      v. The increased variability in the F_2 most likely results from its greater genetic variation.

Inheritance of Ear Length in Corn

Fig. 22.9 Inheritance of ear length in corn



Inheritance of Ear Length in Corn

2. Aside from the environmental influence, four observations emerge that apply generally to similar quantitative-inheritance studies:
   a. The F_1 will have a mean value for the trait intermediate between the means of the two true-breeding parental lines.
   b. The mean value in the F_2 is about the same as that for the F_1.
   c. F_2 shows more variability around the mean than the F_1 does.
   d. Extreme values for the trait in the F_2 extend farther into the parental range than the extreme values for the F_1.
3. The data are not consistent with a single Mendelian locus, because the discrete classes expected do not occur.

Polygene Hypothesis for Quantitative Inheritance

Animation: Polygenic Hypothesis for Wheat Kernel Color

1. The simplest explanation for the corn ear length data is the polygene (multiple-gene) hypothesis for quantitative inheritance, which says that these traits are controlled by many genes.
2. Nilsson-Ehle (1909) studied kernel color in wheat. He crossed true-breeding red kernel wheat with true-breeding white.
   a. The F_1 were all the same intermediate color between red and white. Incomplete dominance was a possibility.
   b. Intercross of F_1 gave an F_2 with four discrete shades of red (ranging from parental red to very light) plus white, in a ratio of 1 : 4 : 6 : 4 : 1 (15 red : 1white). 1⁄16 of the F_2 are parental red, and 1⁄16 are parental white.

c. A 15:1 pattern is typical of a trait that results from the interactions of the products of two pairs of alleles. Both genes affect the same trait, and so are duplicate genes.

i. In the case of wheat, there appear to be two pairs of alleles that segregate independently. Both control red pigment.

ii. Alleles R (red) and C (crimson) result in red pigment, while r and c do not produce pigment.

iii. In the cross RRCC (dark red) 3 rrcc (white), the F_1 are all RrCc (intermediate red).

iv. Interbreeding the F_1 produces an F_2 with genotypes distributed as in a typical dihybrid cross.

v. The numbers in the phenotypic ratio are the same as the coefficients in the binomial expansion of $(a + b)^4$.

vi. The range of phenotypes results from incomplete dominance of the alleles, with each copy of R and C acting as a contributing allele to produce more red pigment. The r and c alleles are noncontributing alleles.

d. Some F_2 populations show only three phenotypic classes (3 red : 1 white), while others show a ratio of 63 red : 1 white, with many shades of red.

i. The 3:1 ratio is consistent with a single-gene system with two contributing alleles.

ii. The 63:1 ratio is consistent with a polygene series with six contributing alleles, $(a + b)^6$.

3. This multiple-gene hypothesis has been applied to other traits, including corn ear length. It proposes that some attributes of quantitative inheritance result from the action and segregation of a number of allelic pairs (polygenes), each making a small but additive contribution to the phenotype.

4. Quantitative trait inheritance appears to be complex, however, and molecular aspects are often not yet well understood.

Lecture 10: Heritability

Heritability

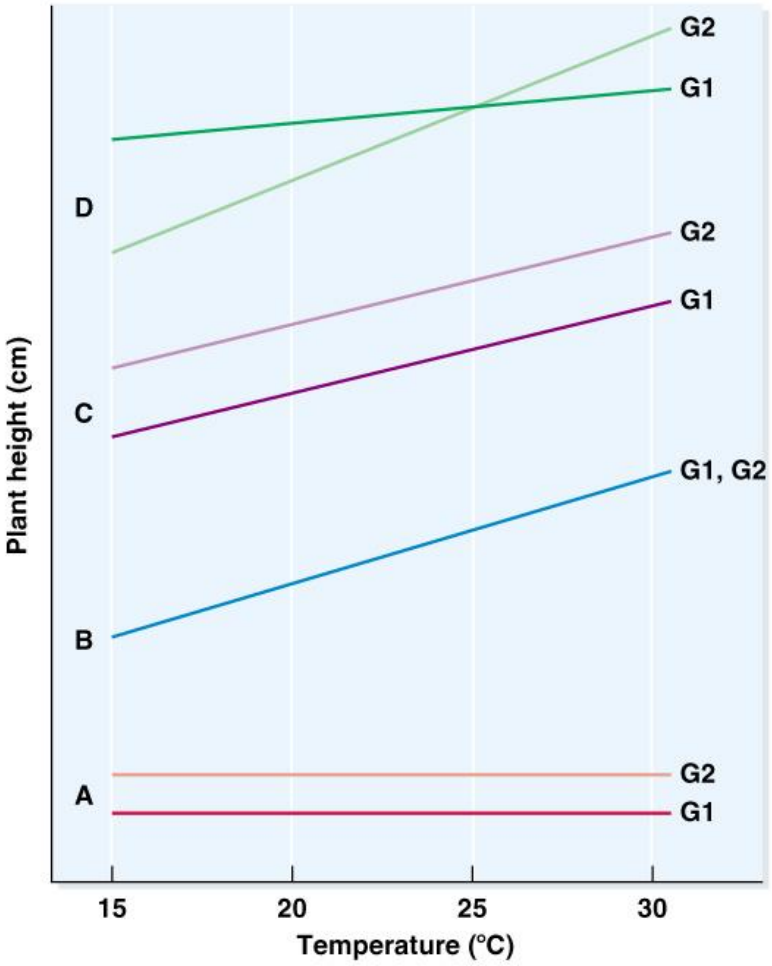
1. Heritability is the proportion of a population's phenotypic variation attributable to genetic factors.
2. Continuous traits are often determined by multiple genes and by environmental factors.
   a. Many ecologically important traits (e.g., body size, fecundity, development rate) are polygenic, and understanding them will contribute to knowledge of how natural populations evolve.
   b. Understanding the roles of genetics and environment in human health (e.g., blood pressure, birth weight) will lead to better health care.
   c. Human social behaviors (e.g., alcoholism, criminality) may also have genetic components, and scientific information has the potential to be misused in constructing social policy.

3. Heritability can be divided into two types, broad-sense and narrow-sense. To assess heritability:
   a. Measure the variation in the trait.
   b. Partition the variance into components attributable to different causes.

Components of the Phenotypic Variance

1. Phenotypic variance (V_P) is the measure of all variability observed for a trait.(Figure 22.12)
   a. The portion of phenotypic variance caused by genetic factors is the genetic variance (V_G).
   b. Nongenetic sources of variation (e.g., temperature, nutrition, parental care) constitute environmental variance (V_E).
   c. The relationship is $V_P = V_G + V_E$.

Fig. 22.12 Hypothetical example of the effects of genes and environments on plant height



Components of the Phenotypic Variance

2. Genetic variance **(V_G)** can be subdivided into three components arising from different gene actions and interactions between genes.

   a. Some genetic variance results from average effects of the different alleles, creating additive genetic variance **(V_A).** Example:

      i. Allele g contributes 2 cm to plant height, while G contributes 4 cm.

      ii. A gg homozygote would receive 4 inches of height, a Gg heterozygote 6 inches, and a GG homozygote 8 inches.

      iii. This effect is added to the height effects produced by alleles at other loci.

      iv. Nilsson-Ehle's experiments with wheat kernel color are another example of additive genetic variance.

Components of the Phenotypic Variance

b. Dominance variance **(V_D)** occurs when one allele masks the effect of another allele at the same locus, and prevents the effects from being strictly additive.

   i. If G is a dominant allele and g a recessive one, the Gg genotype would make the same contribution to phenotype as GG.

   ii. As dominance decreases, genotypic differences become phenotypic differences, turning dominance variance into additive genetic variance.

c. Epistatic interactions occur among alleles at different loci, adding another source of genetic variation, the epistatic or interaction variance **(V_I)**.

d. Genetic variance is the sum of these three factors. **($V_G = V_A + V_D + V_I$)**

Broad-Sense and Narrow-Sense Heritability

1. The amount of variation among individuals resulting from genetic variance (V_G) is the broad-sense heritability of a phenotype.

   a. **Broad-sense heritability = $H^2_B = V_G/V_P$** (h^2 is heritability, and $_B$ designates "broad-sense").

   b. Heritability ranges from 0–1, with 0 meaning no variation from genetic differences, and 1 meaning that all variation is genetically based.

   c. Broad-based heritability:

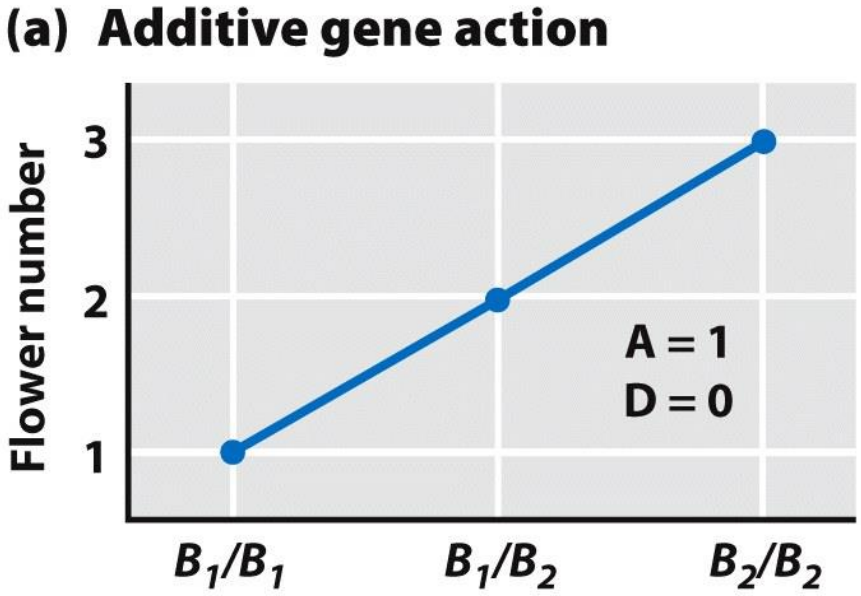
      i. Includes all types of genes and gene actions.

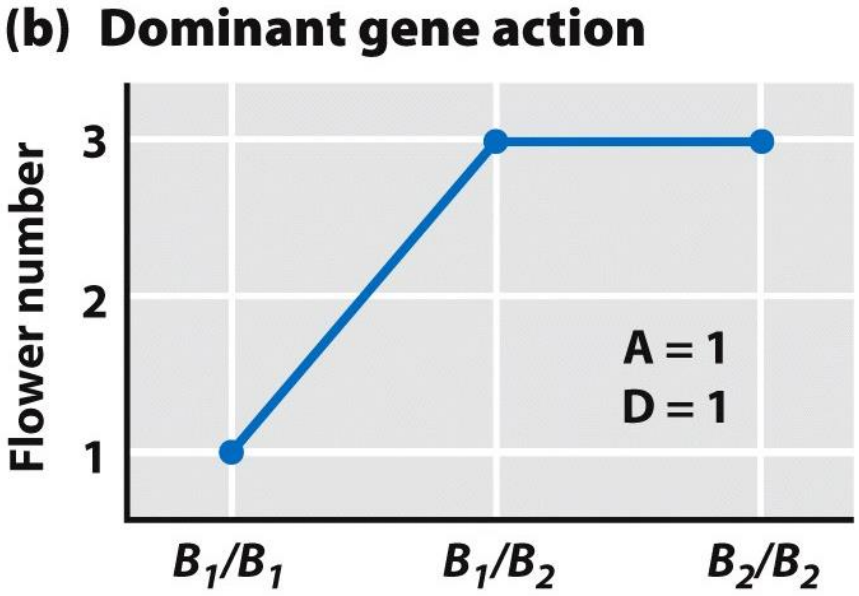
      ii. Does not distinguish between additive, dominance and interactive genetic variance.

2. Additive genetic effects are more often used, because this component allows prediction of the average phenotype of the offspring when phenotypes of the parents are known.
   a. For example, in a cross for a trait involving a single locus:
      i. One parent might be 10 cm tall, with the genotype A^1A^1, and the other parent 20 cm tall, with the genotype A^2A^2.
      ii. If the alleles are additive, the F_1 (A^1A^2) will be 15 cm tall, while if one allele is dominant, the F_1 will resemble one of the parents.
   b. In the same way, epistatic genes will not always contribute to the resemblance between parents and offspring.

3. Narrow-sense heritability is the proportion of the variance resulting from additive genetic variance.
   a. **Narrow-sense heritability = $h^2_N = V_A/V_P$** (h^2 is heritability, and $_N$ designates "narrow-sense").
   b. V_A determines resemblance across generations, and responds to selection in a predictable way.

- The difference between additive and dominant gene action





Understanding Heritability

1. Heritability estimates have limitations that are often ignored, leading to misunderstanding and abuse. Important qualifications and limitations of heritability:

    a. Broad-sense heritability does not indicate the extent to which a trait is genetic. Rather, it measures the proportion of the phenotypic variance in a population resulting from genetic factors.

    b. Heritability does not indicate what proportion of an individual's phenotype is genetic. Heritability is a characteristic of a population, not an individual.

c. Heritability is not fixed for a trait. It depends on the genetic makeup and environment of a population, and so a calculation for one group may not hold true in another. An example is human height.

i. In a population with a uniformly high quality diet, differences in height are likely to be due to genetic factors, especially if there is high ethnic diversity.

ii. In a population where quality of diet varies widely, height differences will be due less to genetic factors, and relatively more to environmental factors.

d. High heritability for a trait does not imply that a population's differences in the same trait are genetically determined. An example is diet in mice.

i. Genetically diverse mice were randomly split into two groups. Both groups received the same space, water and other environmental necessities. The only difference was diet.

(1) Mice in the group receiving nutritionally rich food grew large. Heritability of adult body weight was calculated at 0.93.

(2) Mice in the group receiving an impoverished diet lacking calories and essential nutrients were smaller. Heritability of adult body weight was 0.93.

(3) Both large size and small size were calculated to be highly heritable in these mice. This is a contradiction, since both groups are from the same genetic stock.

How Heritability Is Calculated

1. Many of the methods used compare related and unrelated individuals, or compare individuals with different degrees of relatedness. When environmental conditions are identical:

   a. Closely related individuals have similar phenotypes when the trait is genetically determined.

   b. Related individuals are no more similar in phenotype than unrelated ones when the trait is environmentally determined.

2. In humans, environmental conditions are complex in structure, and extended parental care means that it is difficult to separate the effects of genetics from those of nurture and culture. Methods of calculating heritability in humans include:

   a. Comparison of parents and offspring.

   b. Comparison of full and half siblings.

   c. Comparison of identical and nonidentical twins.

   d. Response-to-selection data.

3. Heritability from parent-offspring regression is calculated by correlation and regression of data measuring the phenotypes of parents and offspring in a series of families.

a. The mean phenotype of the parents (mid-parent value) and mean phenotype of offspring are plotted with each point on the graph representing one family (Figure 22.13).

i. Random scatter across the plot indicates no relationship between the traits of parents and offspring, and thus low heritability.

ii. Linear relationships between phenotypes of parents and offspring indicate that heritabifity is high (unless environmental effects have influenced the trait).

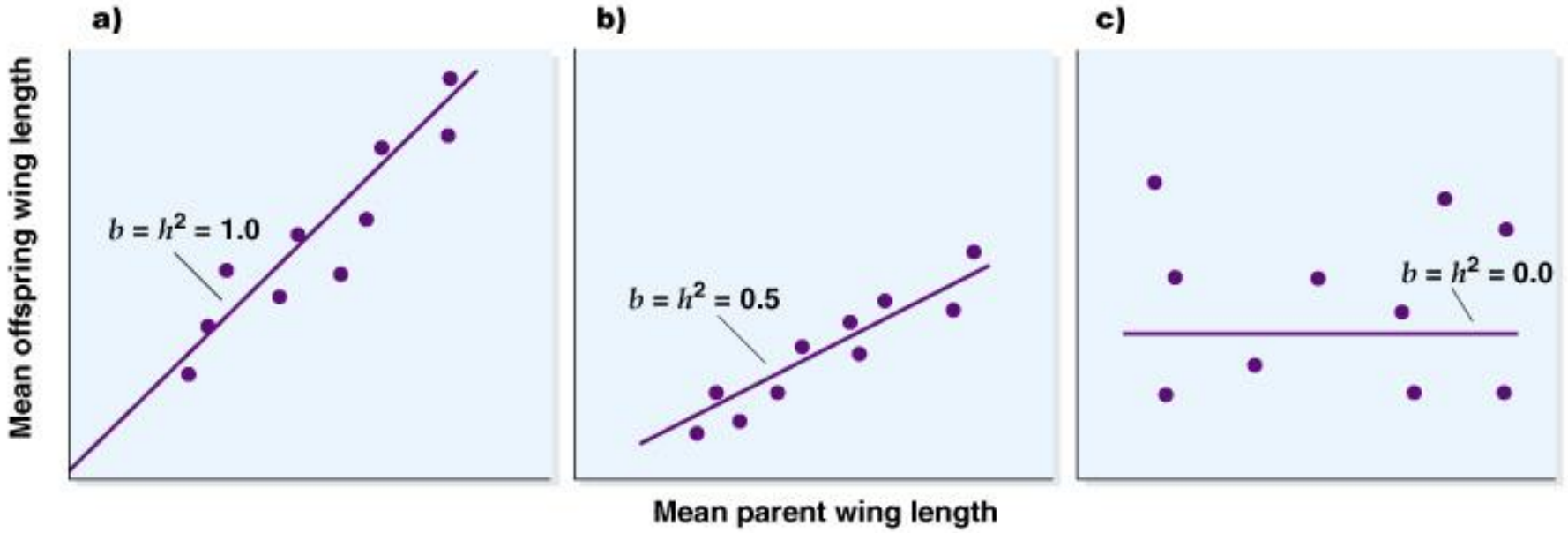
iii. Slope of the parent-offspring regression line reflects the magnitude of heritability.

(1) If the slope is 0, narrow-sense heritability (h^2_N) is also 0.

(2) If the slope is 1, offspring have a phenotype exactly intermediate between the two parents, and additive gene effects account for the entire phenotype.

(3) If the slope is between 0 and 1, both additive genes and nonadditive factors (dominant and epistatic genes, environment) affect the phenotype.

Fig. 22.13 Three hypothetical regressions of mean parental wing length on mean offspring wing length in *Drosophila*



b. When the mean phenotype of the offspring is regressed against the phenotype of only one parent, the narrow-sense heritability is twice the slope, because the offspring shares only 1/2 its genes with the parent.

c. The factor by which the slope is multiplied to obtain heritability increases as the distance between relatives increases.

4. Heritability values have been calculated for many traits in a variety of species and populations, using several methods (Table 22.5).

a. Heritability estimates are not precise and may vary widely for the same trait in the same organism.

b. Human heritability values are especially difficult, because genetic and environmental factors are hard to separate.

Table 22.5 Heritability Values for Some Traits in Humans, Domesticated Animals, and Natural Populations[a]

Organism	Trait	Heritability
Humans	Stature	0.65
	Serum immunoglobulin (IgG) level	0.45
Cattle	Milk yield	0.35
	Butterfat content	0.40
	Body weight	0.65
Pigs	Back-fat thickness	0.70
	Litter size	0.05
Poultry	Egg weight	0.50
	Egg production (to 72 weeks)	0.10
	Body weight (at 32 weeks)	0.55
Mice	Body weight	0.35
Drosophila	Abdominal bristle number	0.50
Jewelweed	Germination time	0.29
Milkweed bugs	Wing length (females)	0.87
	Fecundity (females)	0.50
Spring peeper (frog)	Size at metamorphosis	0.69
Wood frog	Development rate (mountain population)	0.31
	Size at metamorphosis (mountain population)	0.62

[a]The estimates given in this table apply to particular populations in particular environments; heritability values for other individuals may differ.