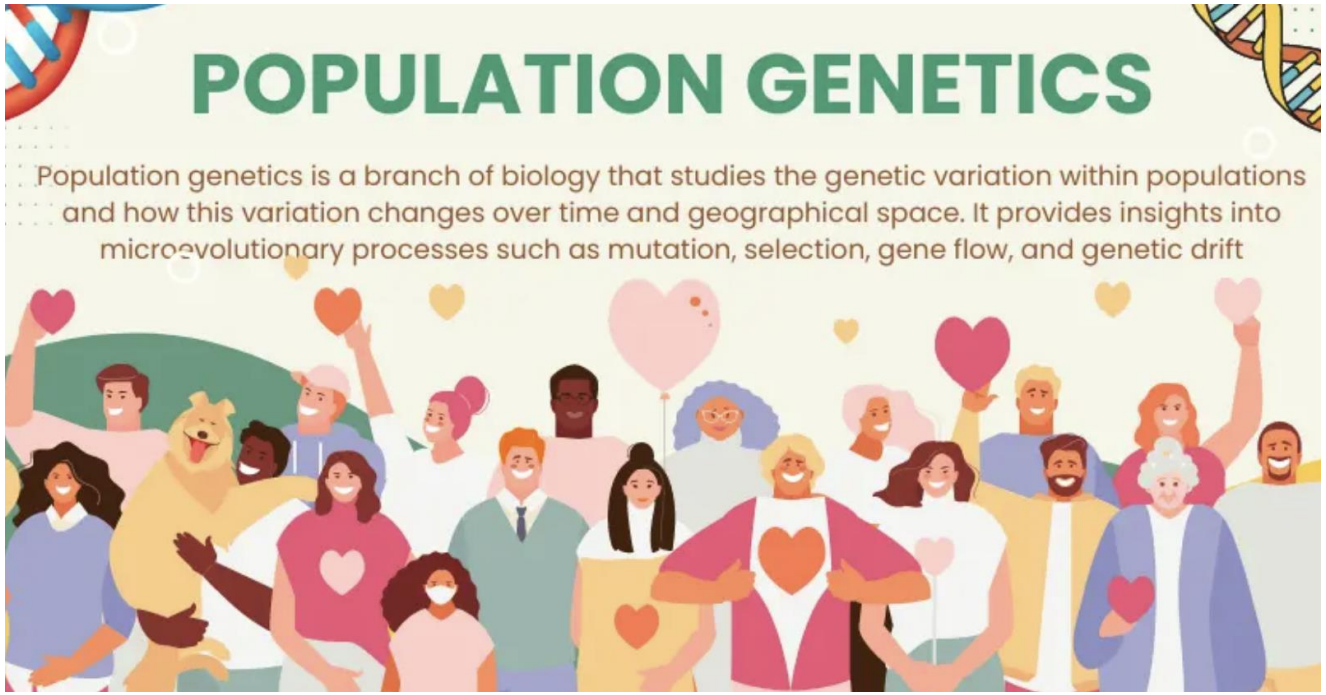
Population and Quantitative Genetics (**5 lectures**)

Prof. Schoen

Module: Population Genetics (first 2.5 lectures)
Lecture 10 — Genes in populations. Hardy-Weinberg theory.
Lecture 11 — Inbreeding. Mutation and migration. Genetic drift.
Lecture 12 — The genetics of natural selection. Intro to Quantitative genetics.



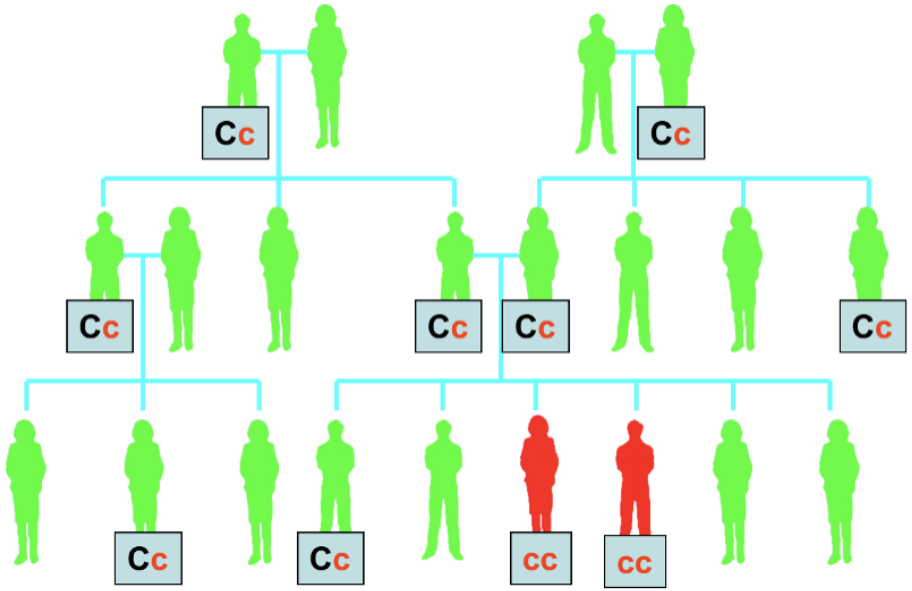
Module – Population Genetics

Lecture 10 — Genes in populations

1. Population genetic variation (Ch18 – section 1)
2. Allele and genotype frequencies in populations.
3. Hardy-Weinberg Theory (Ch18 – section 2)
PRIMER—Population Genetics … posted on MyCourses)
Textbook questions Chapter 18: 1, 3, 5, 7-14, 16, 19, 21, 22, 24, 26a, 27, 32a,b, 39 for upcoming conferences

From Mendelian Genetics *to* Population Genetics

Pedigree Analysis

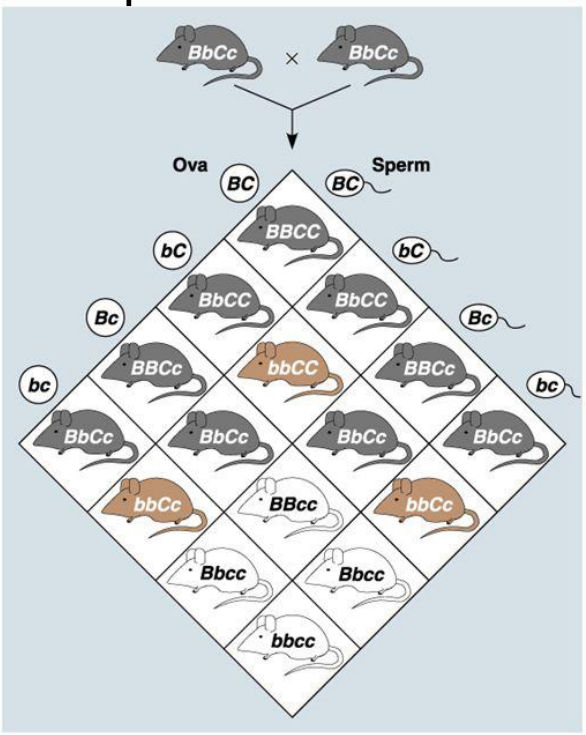


Epistasis

Normal dihybrid ratio is altered from 9:3:3:1 to 9:3:4

C and B gene have an epistatic interaction

Independent Assortment

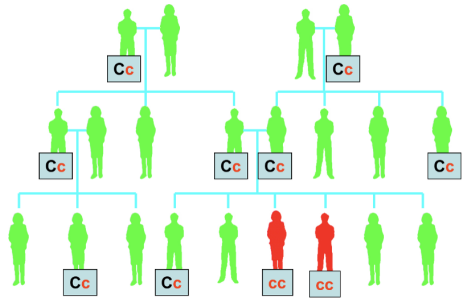


Moving from a few organisms → whole populations (*new genetic properties* emerge)

Populations

Emergent population level properties and processes

- Genotype and allele frequencies.
- Change in genotype and allele frequencies over time (evolution).
- Population to population variation (e.g., variation between populations in genotype and allele frequencies).



e.g., We can look at the frequencies of the alleles C and c in the population

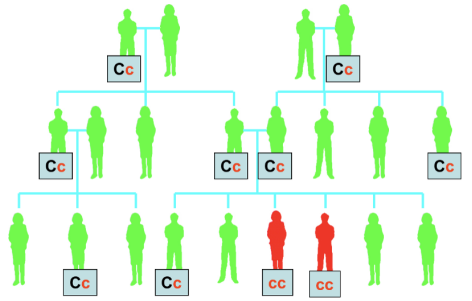
e.g., We can look at the frequencies of the genotypes CC, Cc, and cc in the population

Moving from a few organisms → whole populations (*new genetic properties* emerge)

Populations

Emergent population level properties and processes

- Genotype and allele frequencies.
- Change in genotype and allele frequencies over time (evolution).
- Population to population variation (e.g., variation between populations in genotype and allele frequencies).



e.g., We can look at the frequencies of the alleles C and c in the population

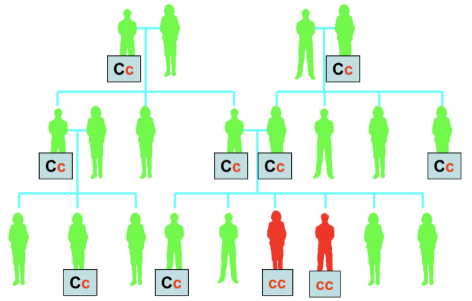
e.g., We can look at the frequencies of the genotypes CC, Cc, and cc in the population

Moving from a few organisms → whole populations (*new genetic properties* emerge)

Populations

Emergent population level properties and processes

- Genotype and allele frequencies.
- Change in genotype and allele frequencies over time (evolution).
- Population to population variation (e.g., variation between populations in genotype and allele frequencies).



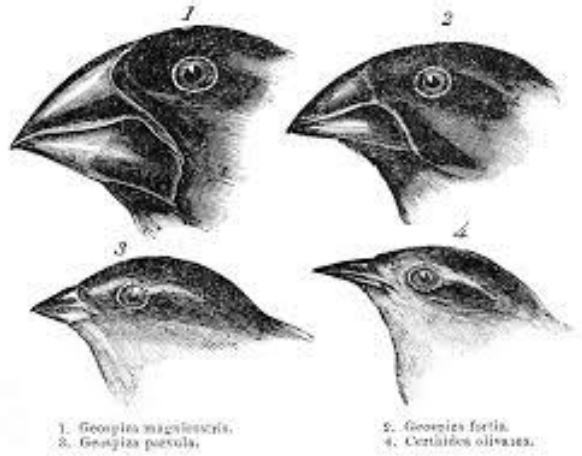
e.g., We can look at the frequencies of the alleles C and c in the population

e.g., We can look at the frequencies of the genotypes CC, Cc, and cc in the population

Moving from a few organisms → whole populations
(*new genetic **processes*** emerge)

Processes of evolutionary change

- Change in population level variation due to MUTATION
- Change in population level variation due to MIGRATION.
- Change in population level variation due to NATURAL SELECTION
- Change in population level variation due to chance (GENETIC DRIFT)



Darwin's Finches

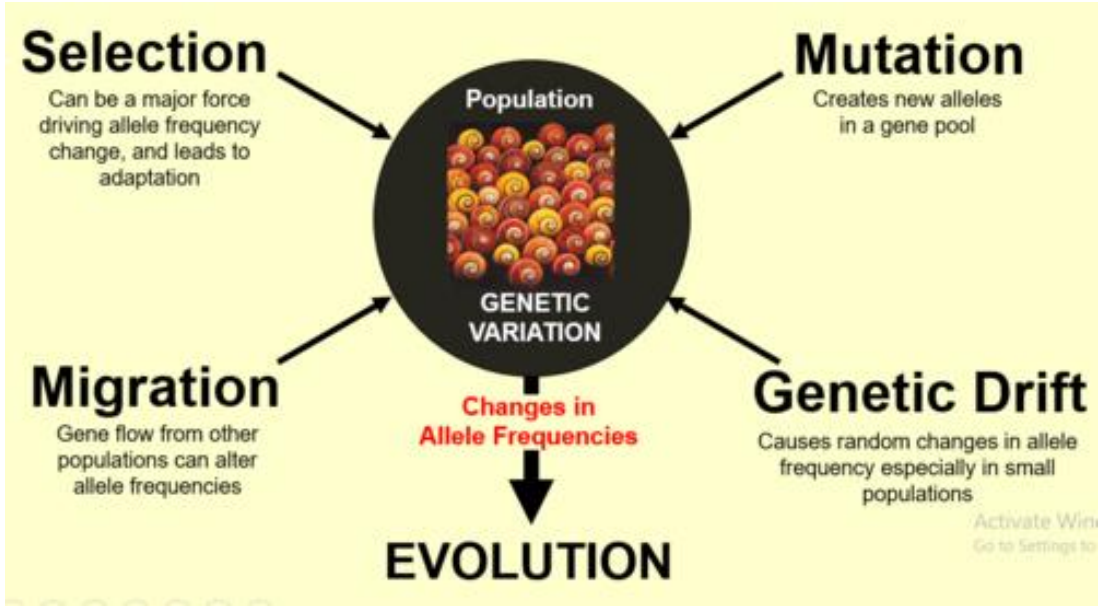
Population genetics ... some of the **many** emergent questions:

How and why do populations show a range of variation for traits (why are some populations more genetically variable than others)?

Population genetics ... some of the **many** emergent questions:

How and why do populations show a range of variation for traits (why are some populations more genetically variable than others)?

How do evolutionary forces such as selection, mutation, migration, and drift act and interact to determine genetic variation at the level of populations

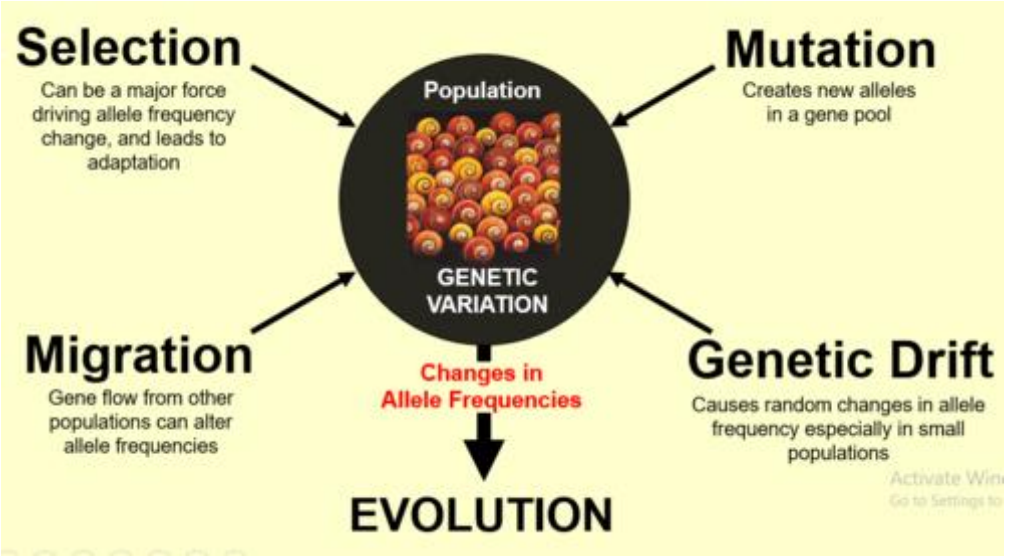


Population genetics ... some of the **many** emergent questions:

How and why do populations show a range of variation for traits (why are some populations more genetically variable than others)?

How do evolutionary forces such as selection, mutation, migration, and drift act and interact to determine genetic variation at the level of populations

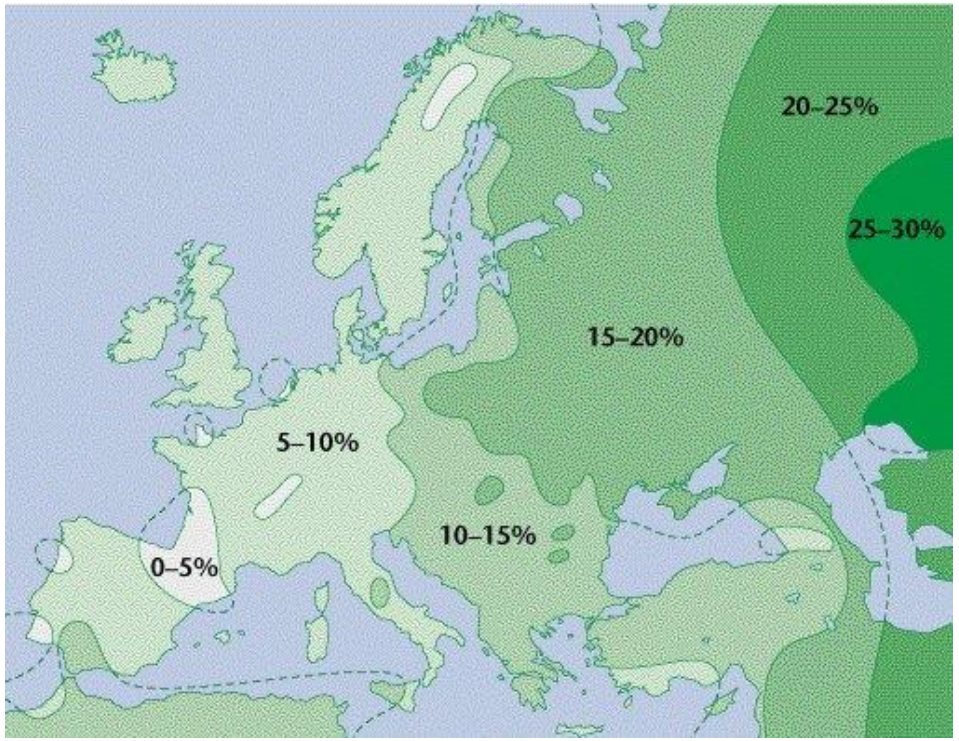
Which force(s) are the most important in determining the outcome of evolutionary change?



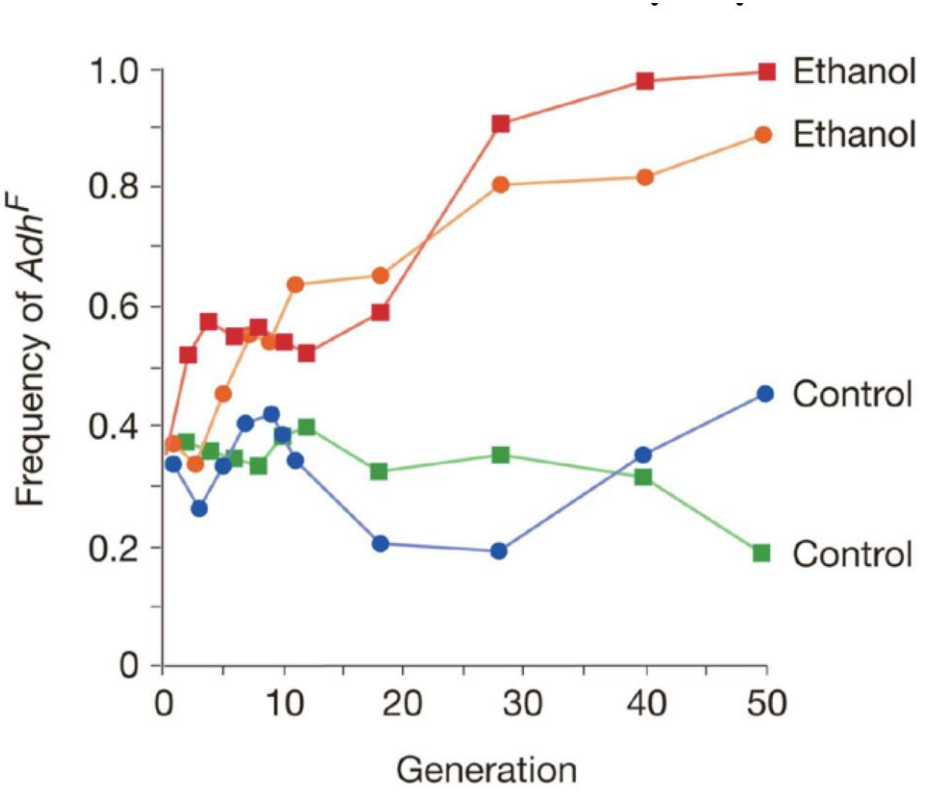
From Mendelian Genetics *to* Population Genetics

(examples of phenomena of interest)

Frequency of the *B* blood type across a geographical region



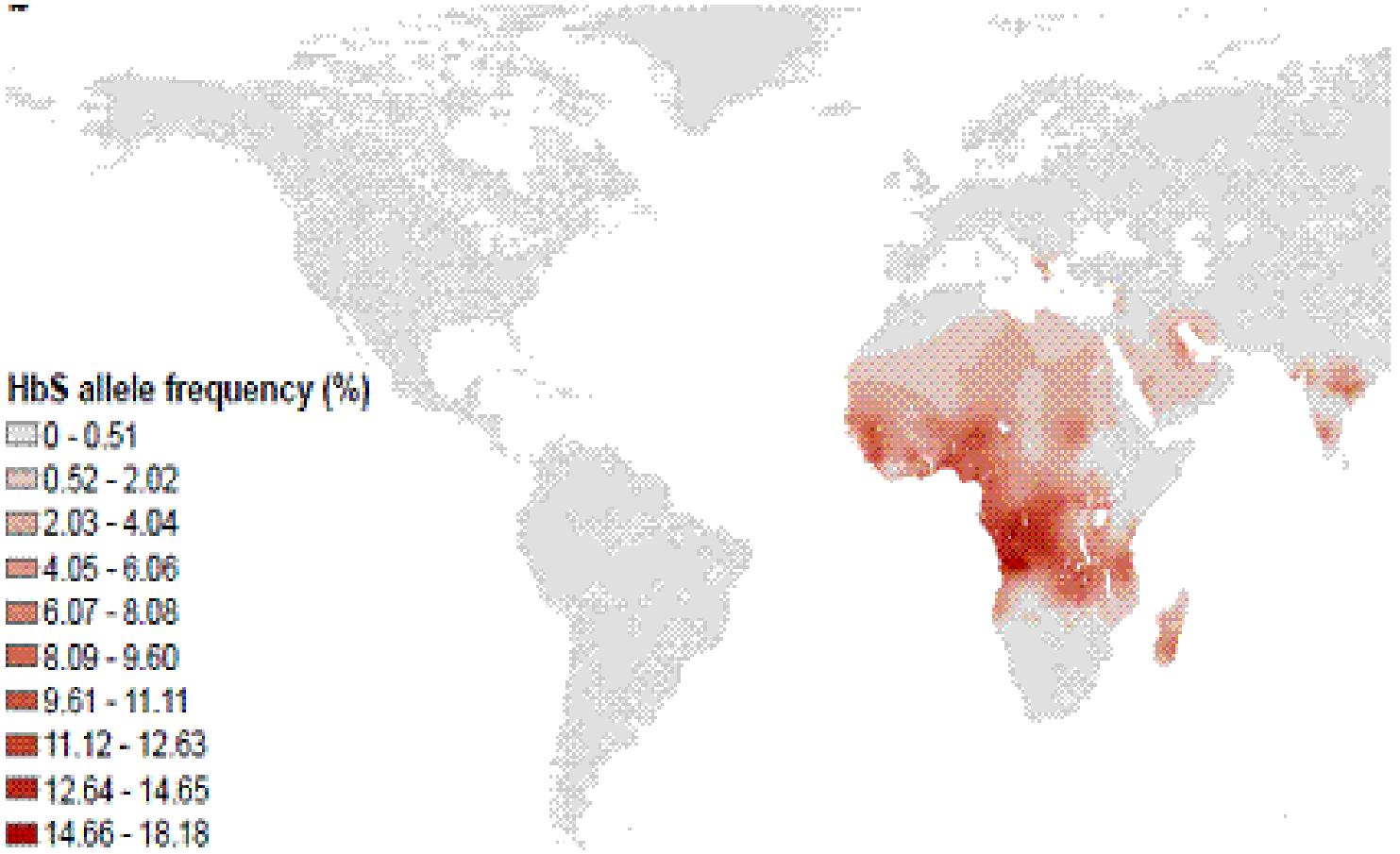
Frequency of the *Adh*(*F*) allele over time



Population level variation: example

Single gene (allele frequency) variation

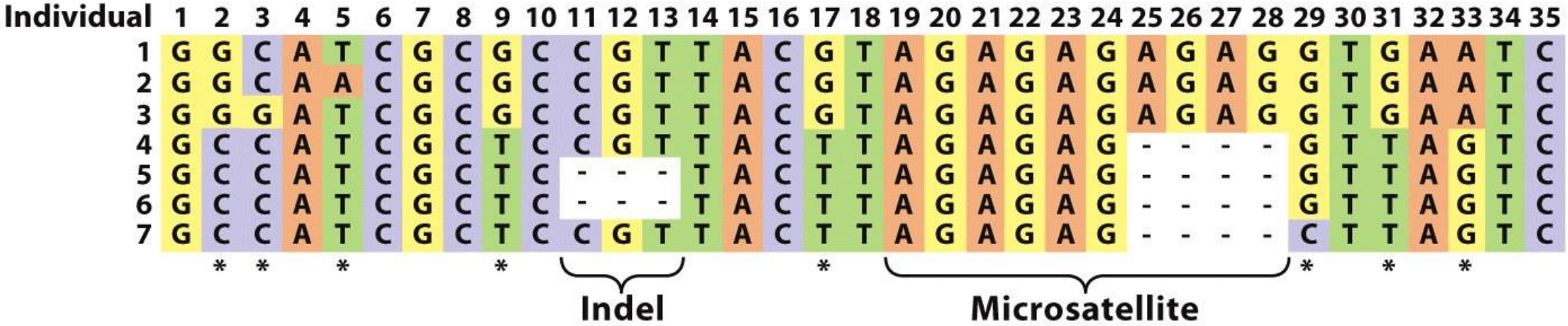
Sickle cell anemia
(when a person has two copies of the ***HbS* allele**)



Population level variation (yet another example)

DNA sequences in a SAMPLE (n=7 individuals) of a population

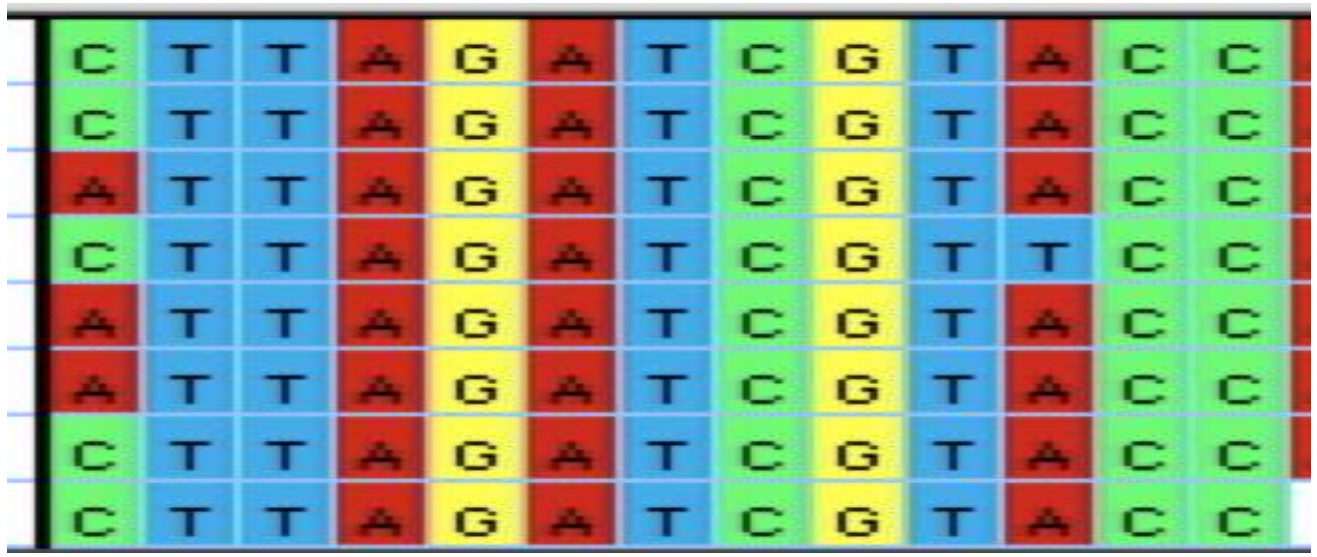
Variation in single nucleotide polymorphisms (SNPs), indels, microsatellites



SNP are denoted by "*" s

- These sequences are from a section of one of the two homologous chromosomes (e.g., chrm 12) of each of the 7 individuals
- …they are referred to as **haplotypes**

In the sample of haplotypes below, how many single locus polymorphisms (**SNPs**) are there?



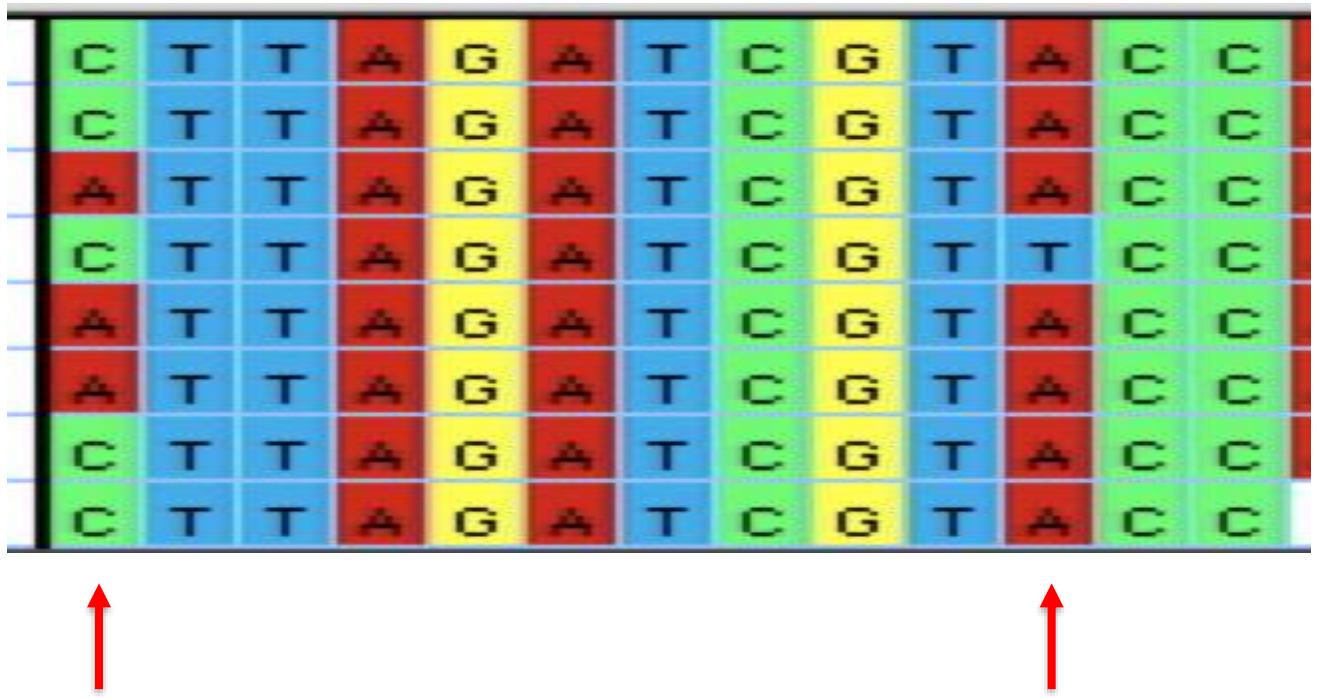
A. 2

B. 3

C. 4

D. 8

In the sample of haplotypes below, how many single locus polymorphisms (**SNPs**) are there?



A. 2

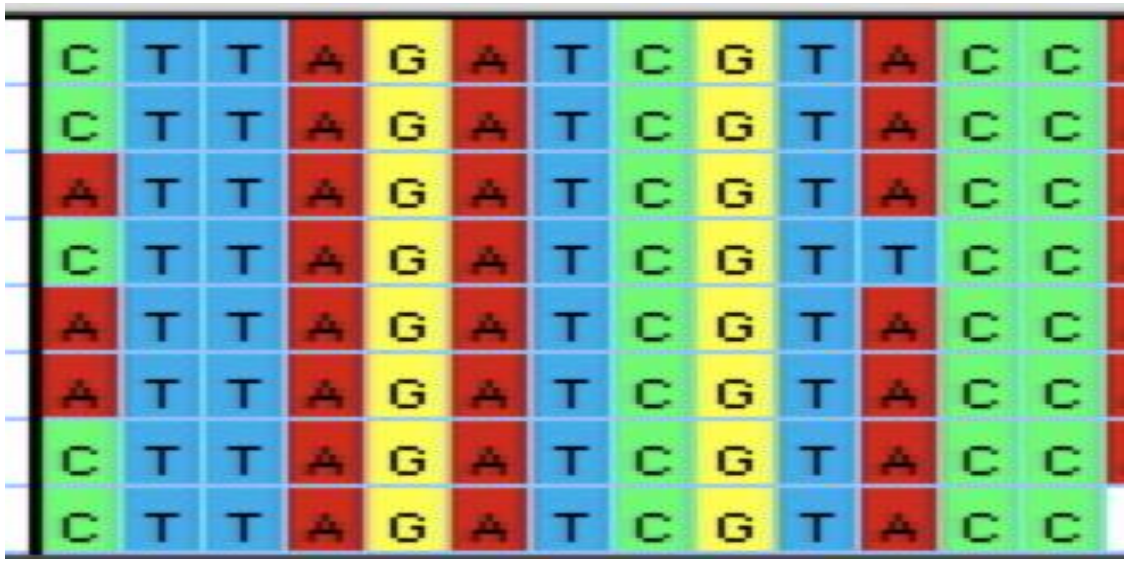
B. 3

C. 4

D. 8

Answer: To be considered a SNP, the nucleotide position must be variable ("polymorphic"). There are **2** such positions in thee sample of sequences

Below are a sample of DNA sequences, each from a portion of Chromosome 1 in a sample of 8 salmon. How many **uniquely different haplotypes** are there?



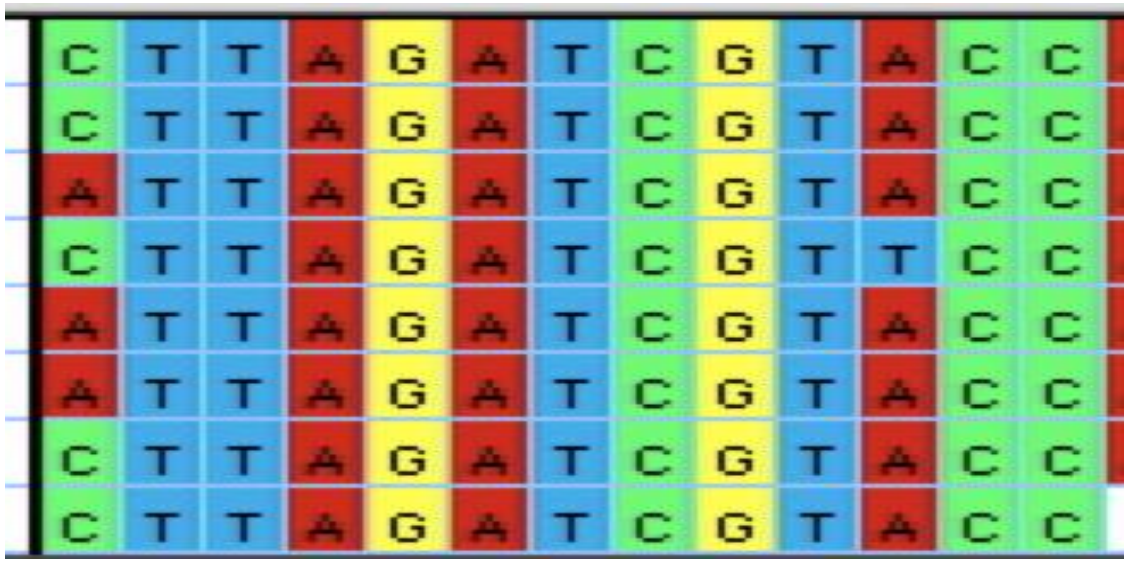
A. 2

B. 3

C. 4

D. 5

Below are a sample of DNA sequences, each from a portion of Chromosome 1 in a sample of 8 salmon. How many **uniquely different haplotypes** are there?



A. 2

B. 3

C. 4

D. 5

Answer: I have numbered the different haplotypes in the sample. Haplotype ”1” is the most common. Haplotypes 2 and 3 are less common than Haplotype 1.

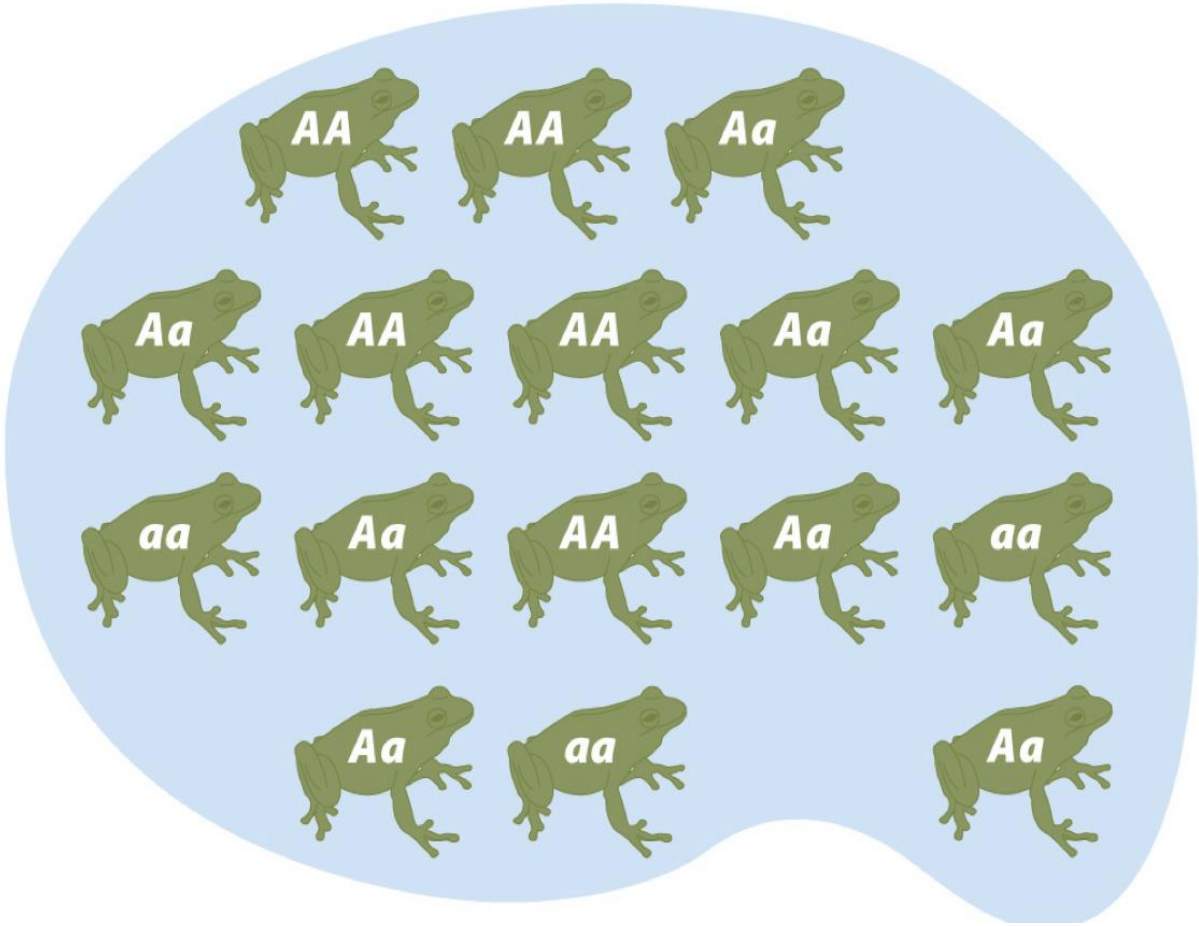
Genetics at the population level

1. Concentrates on *collections* of individuals and their genetic properties, especially the **frequencies of alleles, SNPs, genotypes, haplotypes in time and space (geography, habitats)**

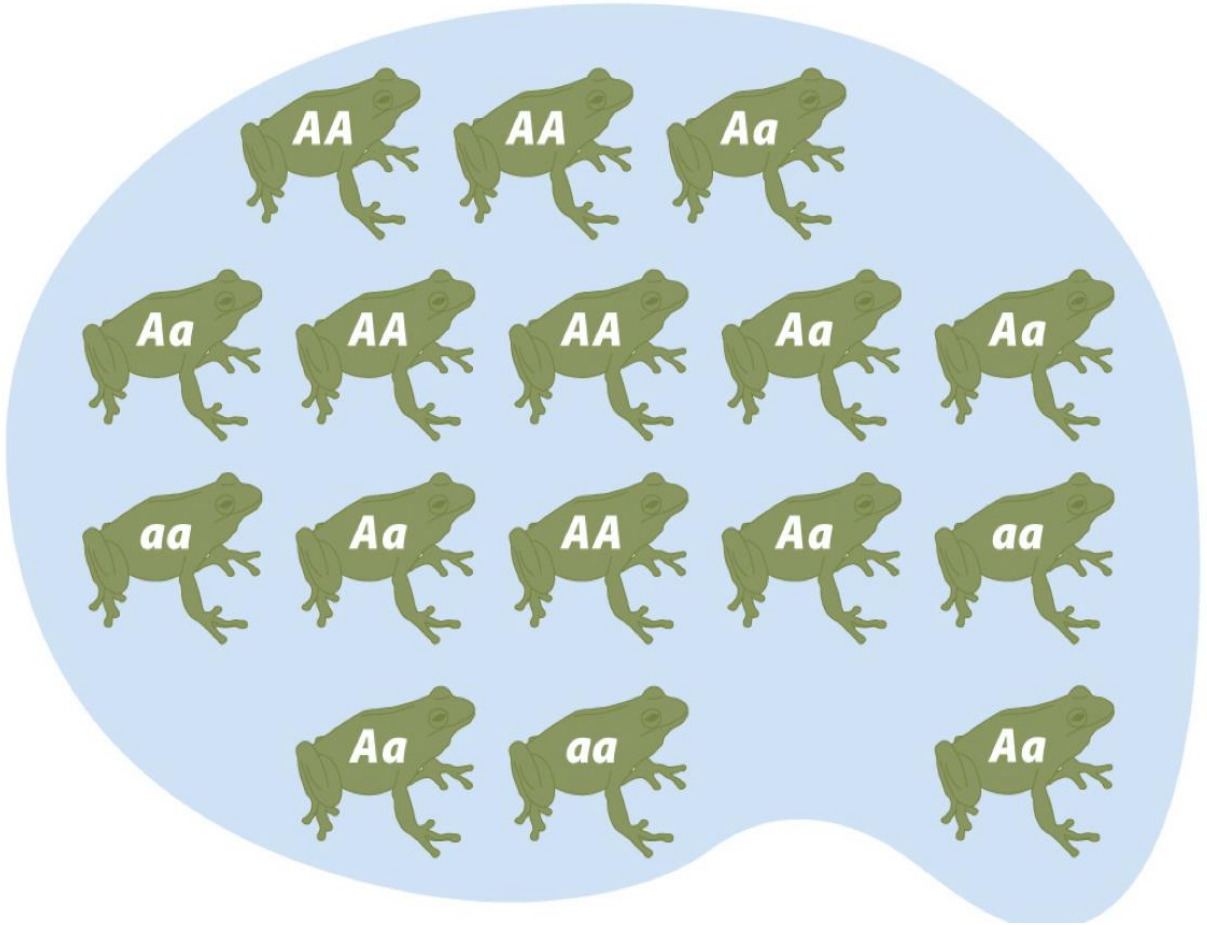
2. Studies the **origin, maintenance, and change** of allelic and genotypic variation in populations.

3. Makes use of **models** to study the processes that influence population genetic composition and make predictions about how it will change (e.g., in response to selection, migration, etc.).

Gene pool model



Characterizing the gene pool (genotype frequencies):



Genotypes	*AA*	*Aa*	*aa*
Number	5	8	3

16 frogs in the population

Notation … for genotype frequencies

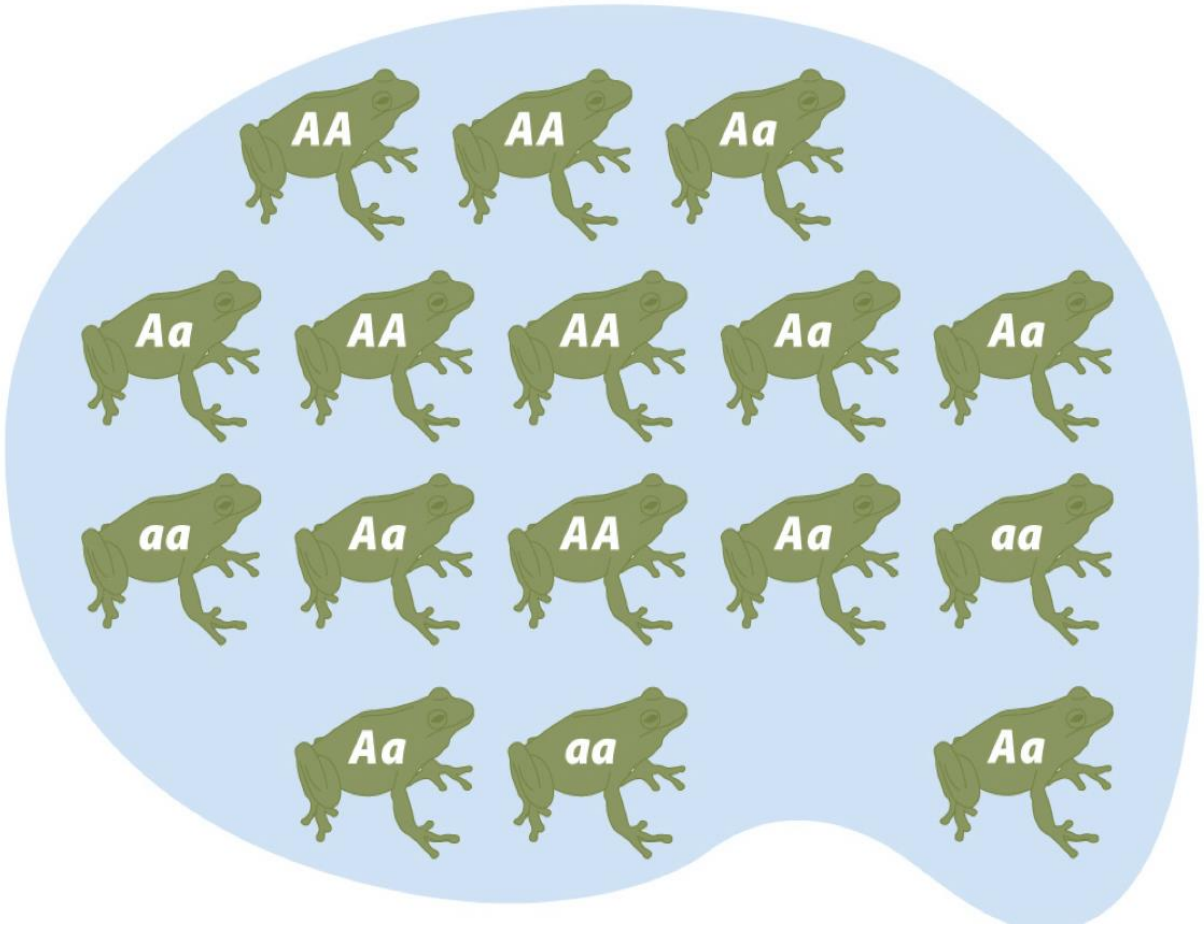
For a gene with two alleles *A* and *a*:

- The frequency genotypes *AA*, *Aa*, and *aa* can be denoted:

  f_{AA}, f_{Aa}, and f_{aa} *or* *f(AA), f(Aa), and f(aa)*

- *The sum:* $f_{AA} + f_{Aa} + f_{aa} = 1.0$

Characterizing the gene pool (genotype frequencies):



16 frogs in the population

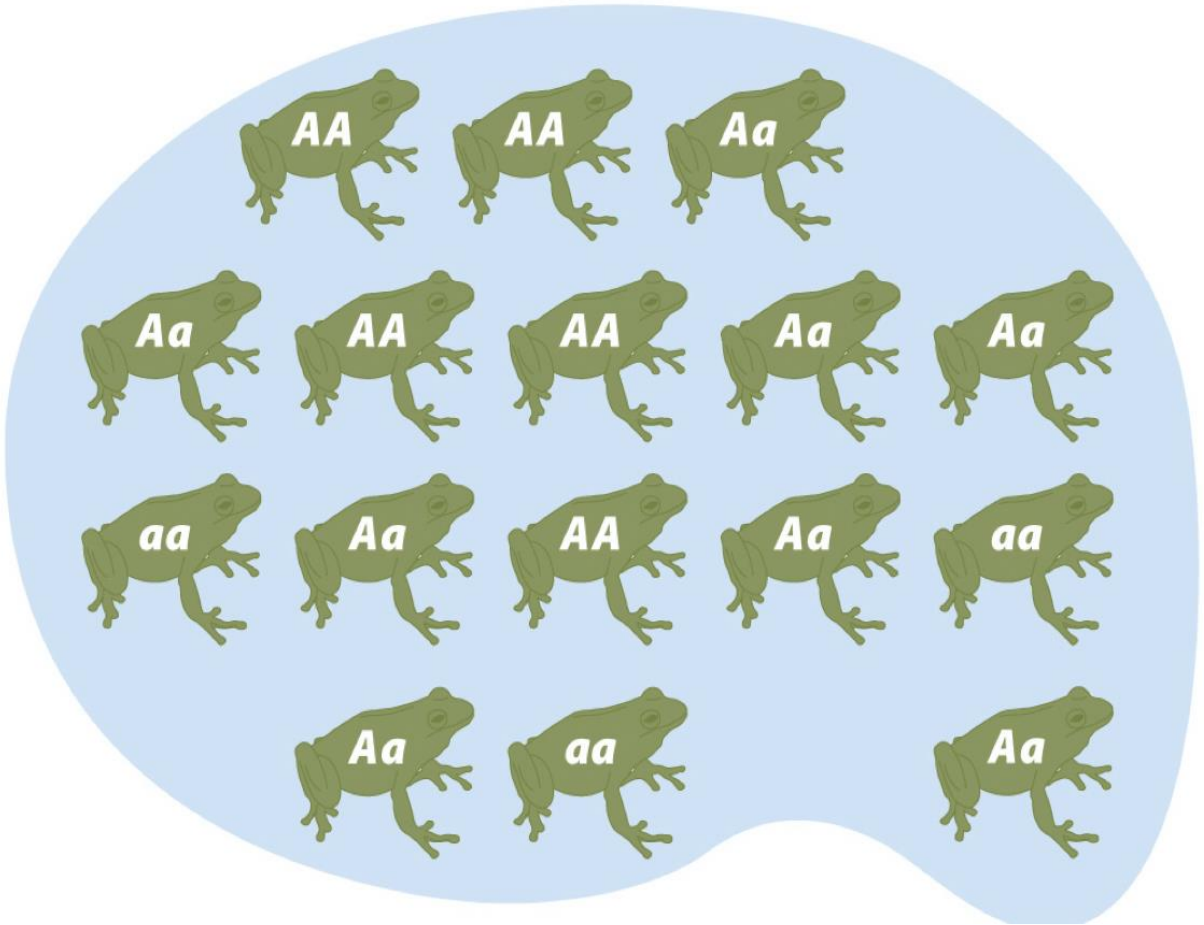
Genotypes	*AA*	*Aa*	*aa*
Number	5	8	3

f(*AA*) = 5/16 = 0.3125

f(*Aa*) = 8/16 = 0.5

f(*aa*) = 3/16 = 0.1875

Characterizing the gene pool (allele frequencies):



Alleles	*A*	*a*
Number	18	14

f(*A*) = 18/32 = 0.5625

f(*a*) = 14/32 = 0.4375

=

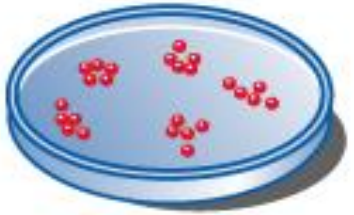
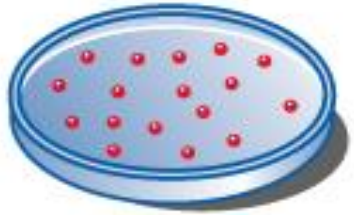
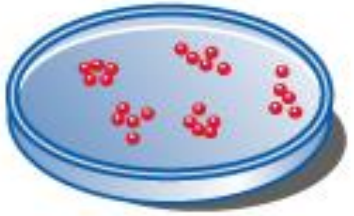
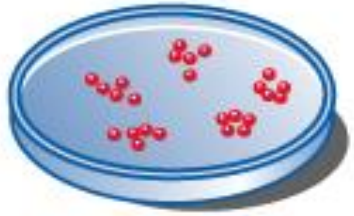
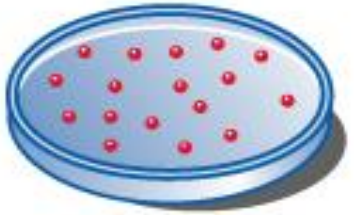
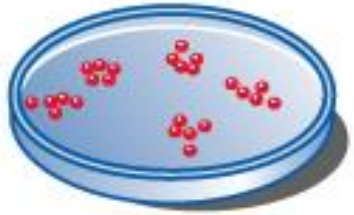
Notation for ALLELE frequencies (biallelic loci)

For a gene with two alleles *A* and *a* :

- The frequency of one allele (e.g., allele *A*) is denoted ***p***.
- The frequency of the other allele (e.g., allele *a*) is denoted ***q***.
- Note that since $p + q = 1$, we can solve for $q = 1-p$.

Population genetic variation

- e.g., MN blood group locus (2 alleles)

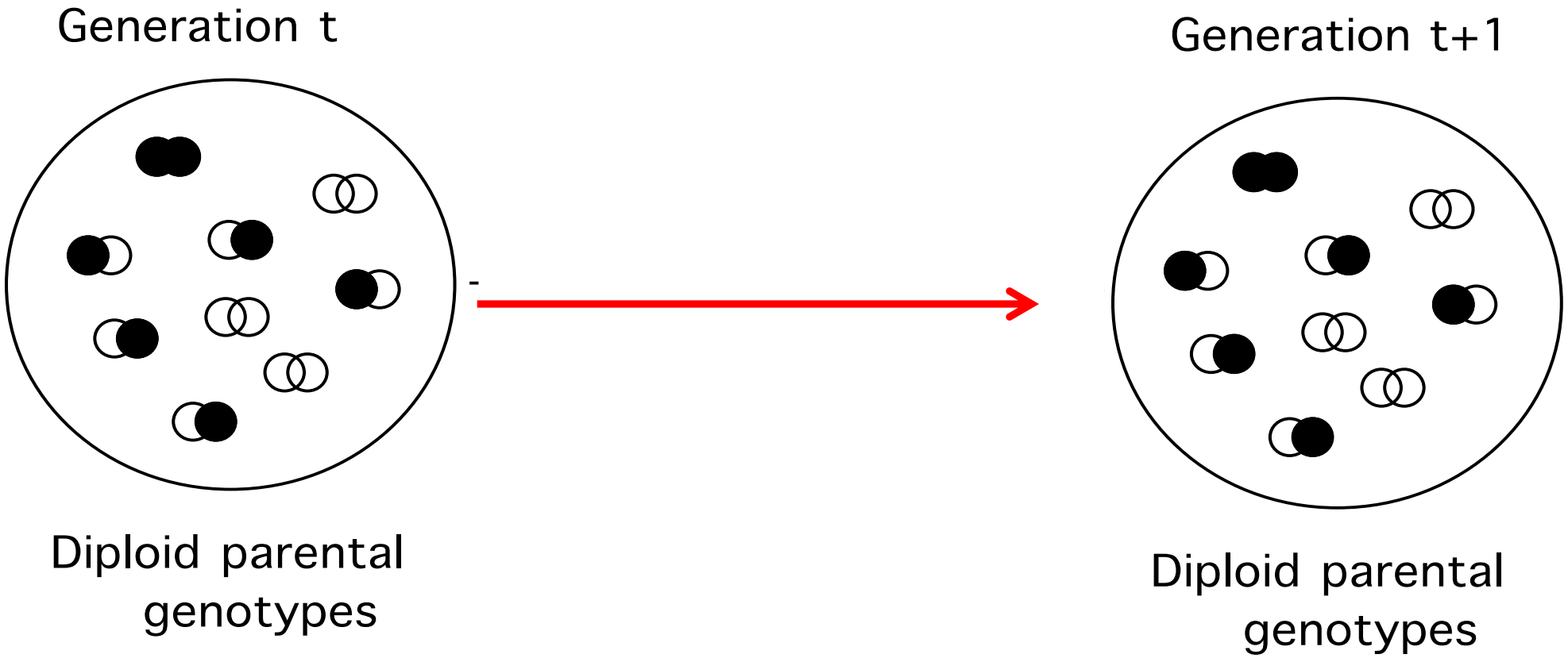
Genotype	Blood type (antigen present)	Reactions with anti-sera		In a population sample of 1000:
		Anti-M serum	Anti-N serum	
$L^M\ L^M$	M			820
$L^M\ L^N$	M N			140
$L^N\ L^N$	N			40

Calculation of allele frequencies at the *MN* locus in a human population from the genotype frequencies

Genotype:	*MM*	*MN*	*NN*
Frequency:	0.82	0.14	0.04

- $p = f_{MM} + (1/2)f_{MN} = 0.82 + 0.14/2 = 0.89$
- $q = f_{NN} + (1/2)\, f_{MN} = 0.04 + 0.14/2 = 0.11$

Can we calculate genotype frequencies in the <u>next generation</u> if we know the genotype frequencies in the current generation?



Freq(AA)=f_{AA}
Freq(Aa)= f_{Aa}
Freq(aa)= f_{aa}

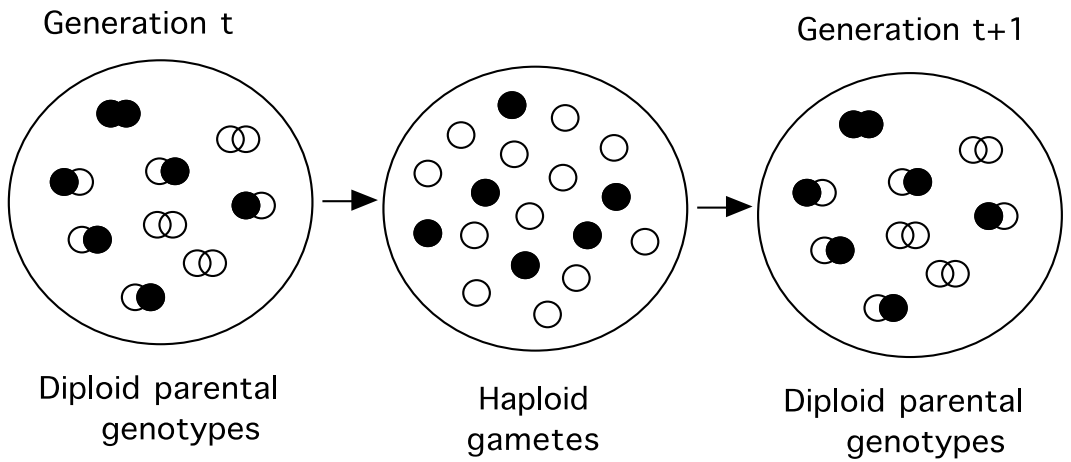
{

Freq(A)= p = f_{AA} +(1/2)f_{Aa}
Freq(a) = q = f_{aa} +(1/2)f_{Aa}

Freq(AA)=*?*
Freq(Aa)= *?*
Freq(aa)= *?*

If there is random mating, no mutation, no selection, no migration and no drift, we CAN calculate the genotype frequencies in the *next* generation from those in the *current one*

Hardy-Weinberg Principle: The constancy of allele frequencies when there are no acting evolutionary forces



Freq(AA)=f_{AA}

Freq(Aa)= f_{Aa}

Freq(aa)= f_{aa}

Step 1

Freq(A)= $p = f_{AA} + (1/2)f_{Aa}$

Freq(a) = $q = f_{aa} + (1/2)f_{Aa}$

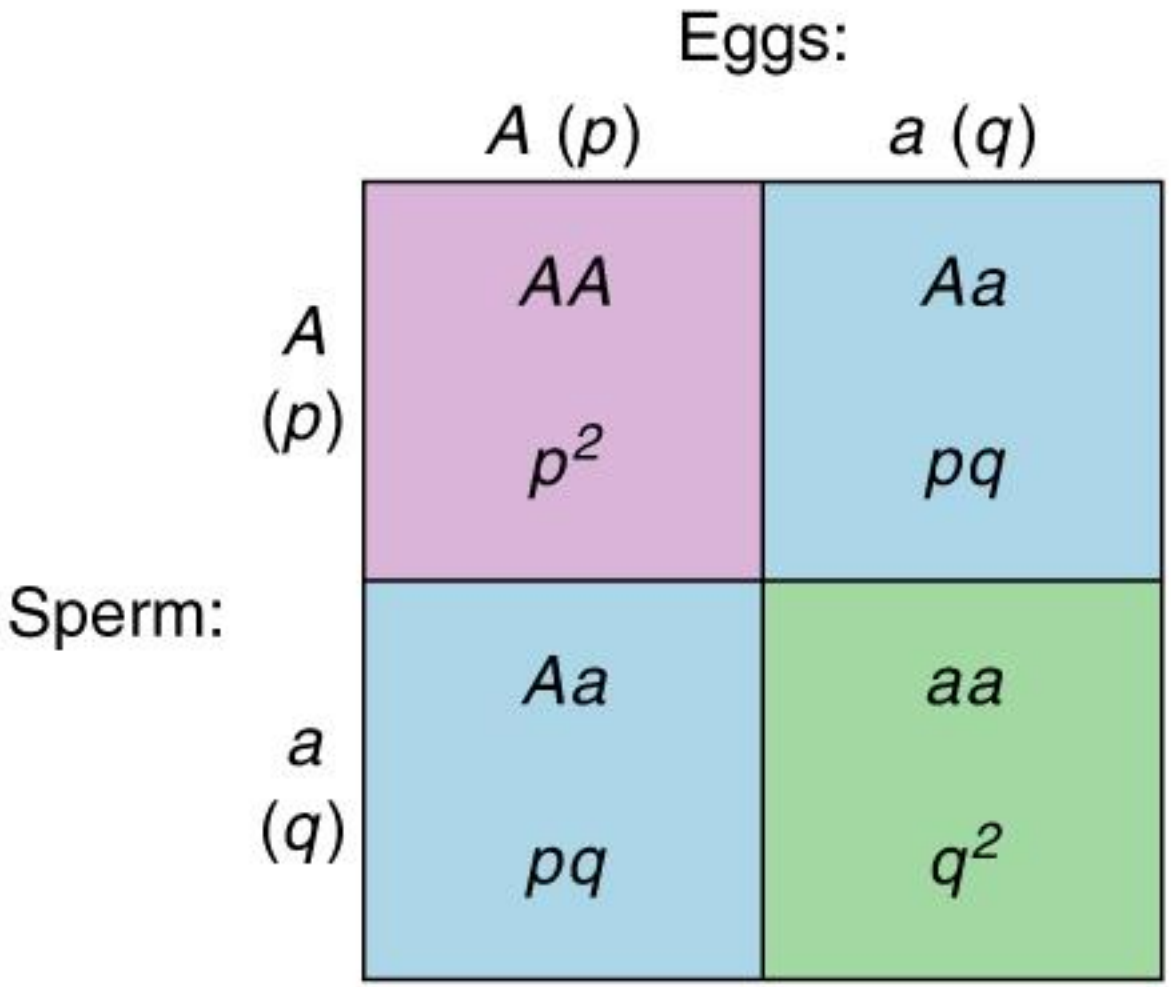
Step 2

Freq(AA)= p^2

Freq(Aa)= $2pq$

Freq(aa)= q^2

When there is **random mating, no mutation, no selection, no migration and no drift**, the genotype frequencies in the next generation are frequencies in the cells of this table:



These are the referred to as:

"Hardy-Weinberg proportions"

or

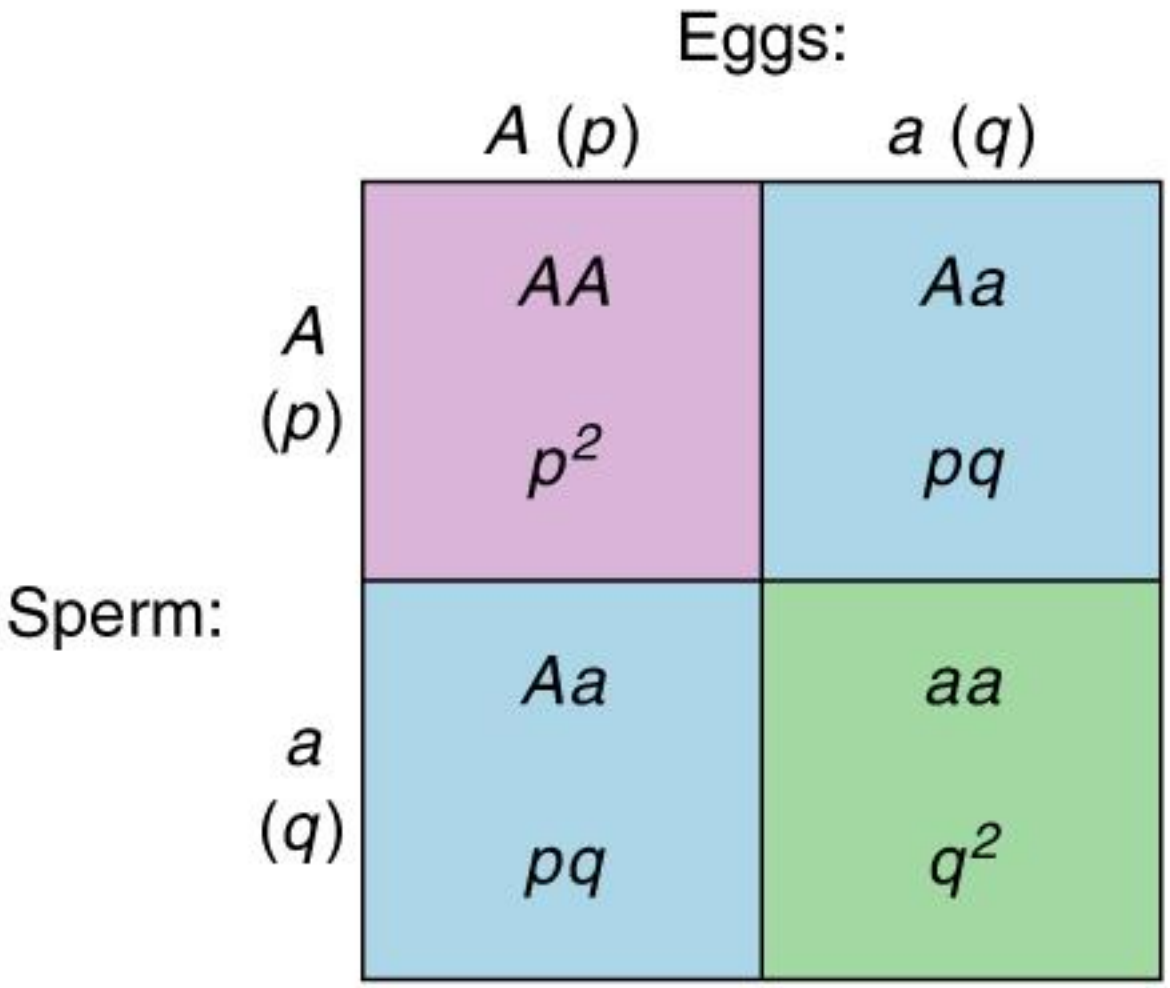
"Hardy-Weinberg genotype fqs"

Freq(*AA*)= p^2

Freq(*Aa*)= $2pq$

Freq(*aa*)= q^2

Hardy-Weinberg Genotype Frequencies



*And from then on,these frequencies will remain unchanged (provided there is continued random mating and absence of evolutionary forces)

Suppose a population has genotype frequencies f(A/A) = 0.1, f(A/a) = 0.4 and f(a/a) = 0.5. Are these genotypes in Hardy-Weinberg proportions?

A No.

B. Yes.

.

Suppose a population has genotype frequencies f(A/A) = 0.1, f(A/a) = 0.4 and f(a/a) = 0.5. Are these genotypes in Hardy-Weinberg proportions?

A No.

B. Yes.

.

p = f(A) = 0.1 + 0.4/2 = 0.3
q = f(a) = 0.5 + 0.4/2 = 0.7

Hardy-Weinberg genotype fqs:

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

Below are genotype frequencies for 6 major blood group genotypes. Calculate the allele frequencies of alleles A, B, and O.

Genotype	Frequency
A/A	0.1
A/0	0.1
B/B	0.2
B/0	0.2
A/B	0.3
O/O	0.1

A. 0.25, 0.35 0.40 (A,B,O)

B. 0.30, 0.45 0.25 (A,B,O)

C. 0.20, 0.40 0.20 (A,B,O)

D. 0.50, 0.40 0.10 (A,B,O)

A population has 3 alleles and genotype frequencies as follow. What are the allele frequencies?

Genotype	Frequency
A/A	0.1
A/0	0.1
B/B	0.2
B/0	0.2
A/B	0.3
O/O	0.1

A. 0.25, 0.35 0.40 (A,B,O)

B. 0.30, 0.45 0.25 (A,B,O)

C. 0.20, 0.40 0.20 (A,B,O)

D. 0.50, 0.40 0.10 (A,B,O)

Answer:

f(A) = 0.1 + 0.1/2 + 0.3/2 = 0.30

f(B) = 0.2 + 0.2/2 + 0.3/2 = 0.45

f(O) = 0.1 + 0.1/2 + 0.2/2 = 0.25

Hardy-Weinberg genotype frequencies with 3 alleles A_1, A_2, and A_3 having the frequencies p_1, p_2, and p_3, respectively (e.g., $p_1 = 0.5$, $p_2 = 0.3$, $p_2 = 0.2$)

Genotype	Expectation	Frequency
A_1A_1	p_1^2	0.25
A_2A_2	p_2^2	0.09
A_3A_3	p_3^2	0.04
A_1A_2	$2p_1p_2$	0.30
A_1A_3	$2p_1p_3$	0.20
A_2A_3	$2p_2p_3$	0.12
Sum		1.00

Cystic Fibrosis (CF), a recessive disease, occurs in 4 out of 10,000 persons. Assume that the genotypes at the locus underlying CF are present in Hardy-Weinberg proportions in the human population. What is the frequency of carriers (heterozygotes) in the population?

Cystic Fibrosis (CF) occurs in 4 out of 10,000 persons. Assume that the genotypes at the locus underlying CF are present in Hardy-Weinberg proportions in the human population.

What is the frequency of carriers (heterozygotes) in the population?

Genotype	Phenotype	Frequencies: Expected	Observed
AA	Normal	p^2	?
Aa	Normal (Carrier)	$2pq$	?
aa	CF	q^2	0.0004

If the freq(*aa*) = 0.0004, the frequency of the *a* allele must therefore be $q = (0.0004)^{0.5}$

That is, $q = 0.02$

Cystic Fibrosis (CF) occurs in 4 out of 10,000 persons. Assume that the genotypes at the locus underlying CF are present in Hardy-Weinberg proportions in the human population.

What is the frequency of carriers (heterozygotes) in the population?

Genotype	Phenotype	Frequencies:	
		Expected	*Observed*
AA	Normal	p^2	?
Aa	Normal (Carrier)	$2pq$	?
aa	CF	q^2	0.0004

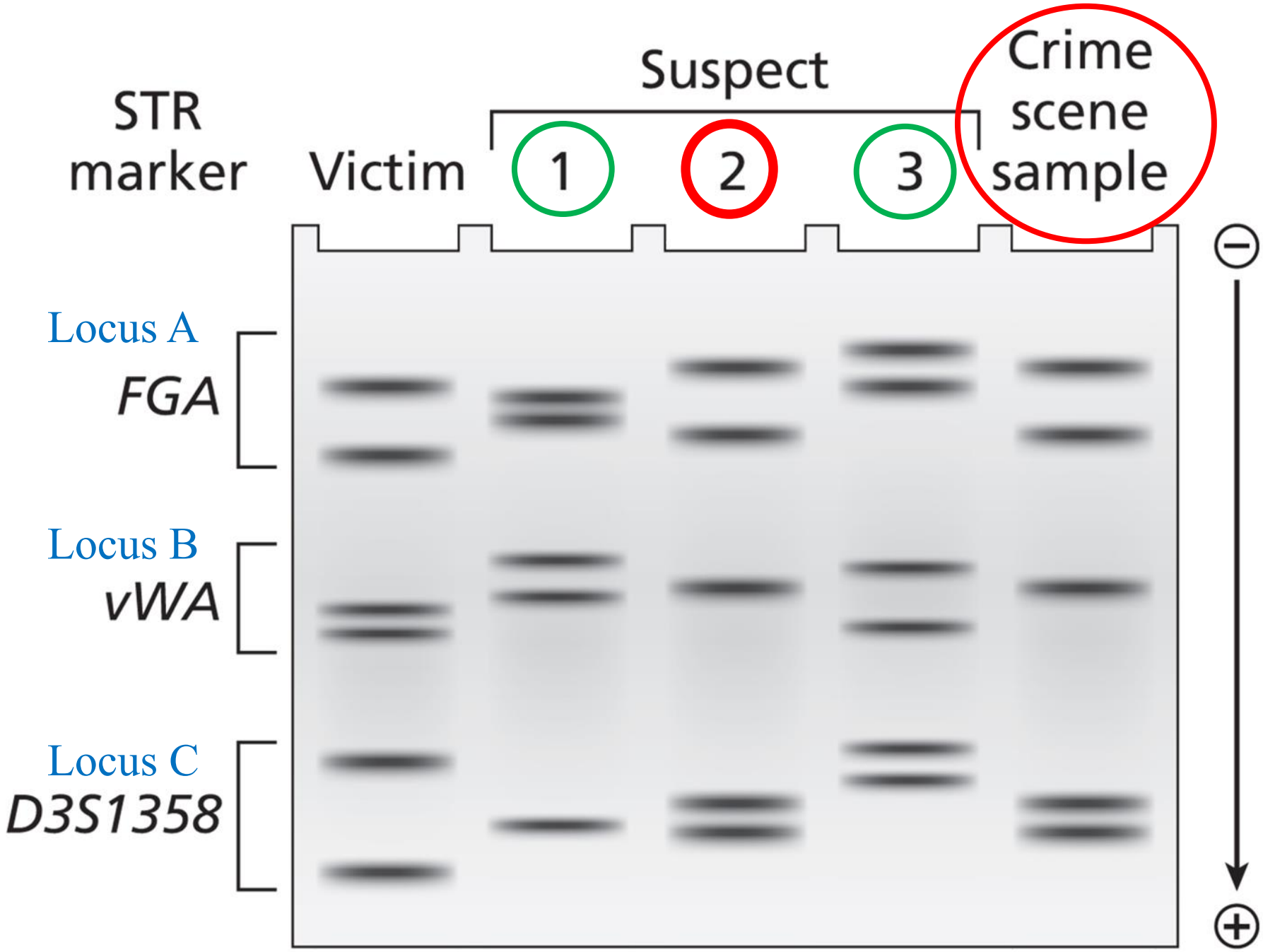
If the freq(*aa*) = 0.0004, the frequency of the *a* allele must therefore be $q = (0.0004)^{0.5}$

That is, $q = 0.02$

***So, $2pq = 2\ (0.98)\ (0.02) = 0.0392$**

Ratio of carriers:diseased

$= 2pq/q^2 = 0.0392/0.0004 = 98$



DNA fingerprinting

A suspect in a crime has the following genotype at three separate and unlinked biallelic loci, A, B and C: *A1/A2; B2/B2; C2/C2*. Blood found at the scene of the crime is seen to match this genotype. The allele frequencies of alleles A1, B1, C2 in the population are 0.9, 0.8, and 0.7, respectively.

What is the probability that this match occurs simply by chance alone? (Assume Hardy-Weinberg proportions for each locus)

A. 0.432001
B. 0.055440
C. 0.000648
D. 0.000493

A suspect in a crime has the following genotype at three separate and unlinked biallelic loci, A, B and C: A1/A2; B2/B2; C2/C2. Blood found at the scene of the crime is seen to match this genotype. The allele frequencies of alleles A1, B1, C1 in the population are 0.9, 0.8, and 0.7, respectively.

What is the probability that this match occurs in the general population? (Assume Hardy-Weinberg proportions for each locus)

A suspect in a crime has the following genotype at three separate and unlinked biallelic loci, A, B and C: *A1/A2; B2/B2; C2/C2*. Blood found at the scene of the crime is seen to match this genotype. The allele frequencies of alleles A1, B1, C1 in the population are 0.9, 0.8, and 0.7 respectively.

What is the chance that this match occurs in the general population? (Assume Hardy-Weinberg proportions for each locus)

Freq *(A1/A2)* = *2pq* = 2(0.9)(0.1) = 0.18

In-class Problem 2. A suspect in a crime has the following genotype at three separate and unlinked biallelic loci, A, B and C: *A1/A2; B2/B2; C2/C2*. Blood found at the scene of the crime is seen to match this genotype. The allele frequencies of alleles A1, B1, C2 in the population are 0.9, 0.8, and 0.7 respectively.

What is the chance that this match occurs in the general population? (Assume Hardy-Weinberg proportions for each locus)

Freq *(A1/A2)* = $2pq = 2(0.9)(0.1) = 0.18$

Freq *(B2/B2)* = $q^2 = (0.2)^2 = 0.04$

In-class Problem 2. A suspect in a crime has the following genotype at three separate and unlinked biallelic loci, A, B and C: *A1/A2; B2/B2; C2/C2*. Blood found at the scene of the crime is seen to match this genotype. The allele frequencies of alleles A1, B1, C1 in the population are 0.9, 0.8, and 0.7 respectively.

What is the chance that this match occurs in the general population? (Assume Hardy-Weinberg proportions for each locus)

Freq *(A1/A2)* = $2pq = 2(0.9)(0.1) = 0.18$

Freq *(B2/B2)* = $q^2 = (0.2)^2 = 0.04$

Freq *(C2/C2)* = $q^2 = (0.3)^2 = 0.09$

In-class Problem 2. A suspect in a crime has the following genotype at three separate and unlinked biallelic loci, A, B and C: *A1/A2; B2/B2; C2/C2*. Blood found at the scene of the crime is seen to match this genotype. The allele frequencies of alleles A1, B1, C1 in the population are 0.9, 0.8, and 0.7 respectively.

What is the chance that this match occurs in the general population? (Assume Hardy-Weinberg proportions for each locus)

Freq *(A1/A2)* = $2pq$ = 2(0.9)(0.1) = 0.18

Freq *(B2/B2)* = q^2 = $(0.2)^2$ = 0.04

Freq *(C2/C2)* = q^2 = $(0.3)^2$ = 0.09

Prob *(A1/A2; B2/B2; C2/C2)* = 0.18 x 0.04 x 0.09 = 0.000648

Lecture 10 (Main Points)

- Allele, genotype, haplotype frequencies can be used to characterize the genetic composition of populations.
- Evolution occurs when allele, genotype, and haplotype frequencies change.
- Hardy-Weinberg theory helps us to calculate genotype frequencies in populations. It assumes random mating with respect to the locus in question, AND the absence of evolutionary forces (mutation, migration, drift, selection).
- You should know how to apply Hardy-Weinberg theory, as in assigned problems.