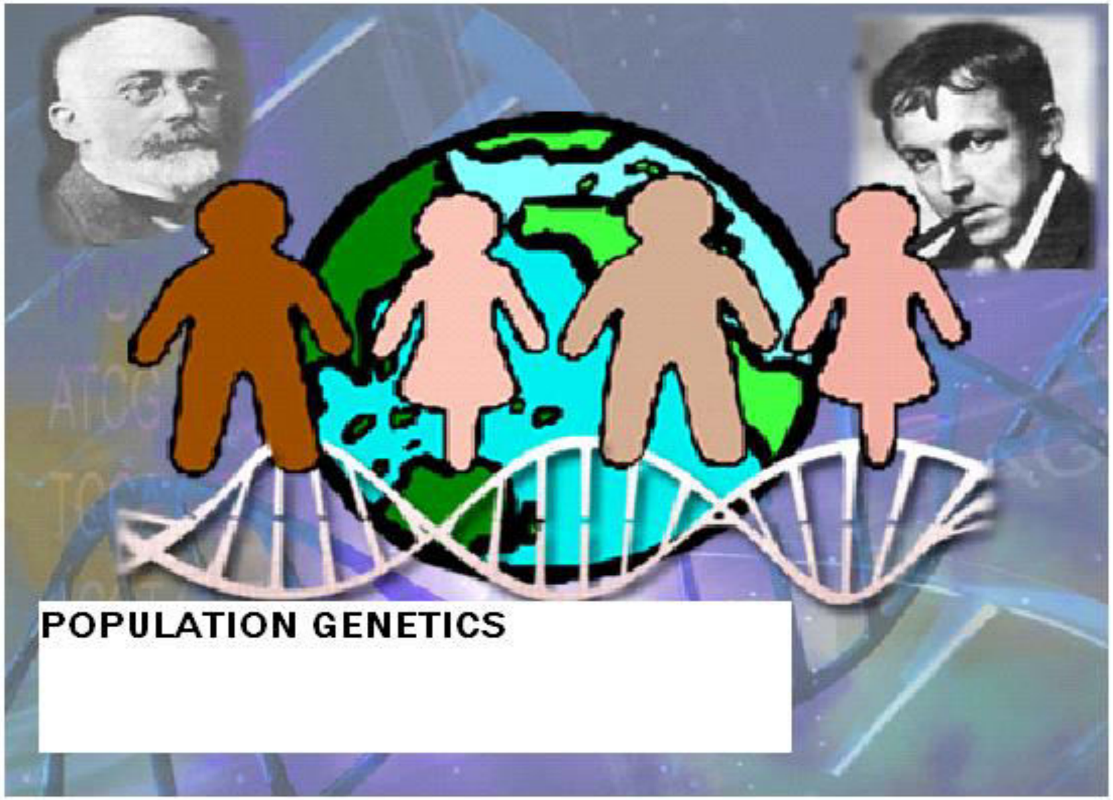
POPULATION GENETICS



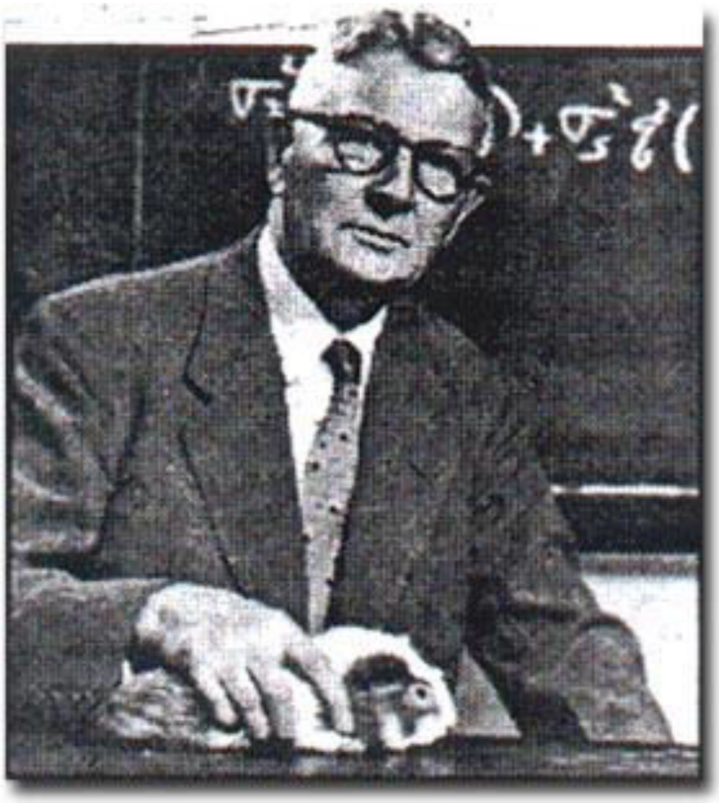
Dr. Monisha Banerjee, Professor, Molecular & Human Genetics
Lab, Department of Zoology, University of Lucknow, Lucknow.

CONTENTS

FOUNDERS OF POPULATION GENETICS

Ronald Fisher

J. B. S. Haldane



Sewall Wright

Population genetics is important because <u>populations</u> evolve (individuals do not)

- **population**: localized group of individuals belonging to same species

Population genetics

- genetic structure of a population

Population genetics

- genetic structure of a **population**

group of individuals
of the same species
that can interbreed

Population genetics

- **genetic structure of a population**
  - alleles
  - genotypes

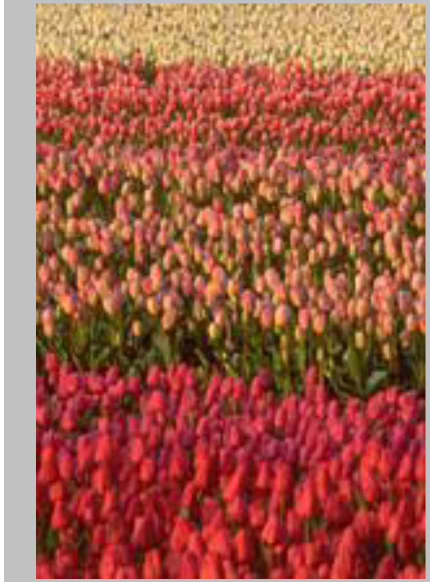
group of individuals
of the same species
that can interbreed

Patterns of genetic variation in populations

Changes in genetic structure through time

Describing genetic structure

- genotype frequencies
- allele frequencies



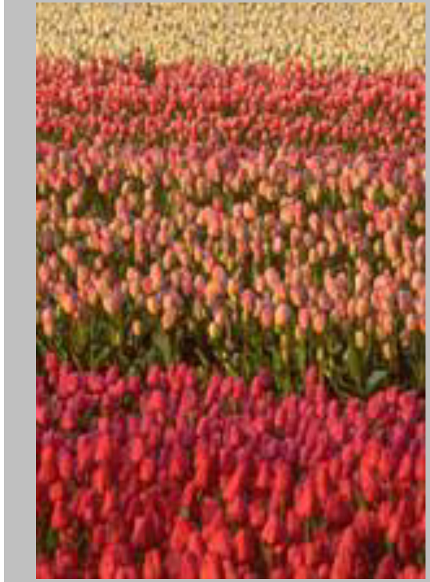
rr = white

Rr = pink

RR = red

Describing genetic structure

- **genotype frequencies**
- **allele frequencies**



200 white

500 pink

300 red

total = 1000 flowers

genotype frequencies:

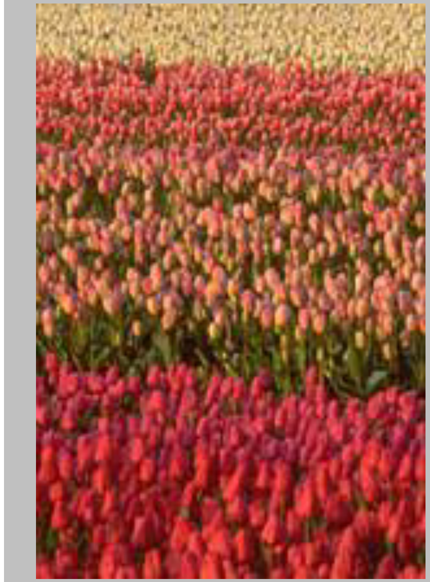
200/1000 = 0.2 rr

500/1000 = 0.5 Rr

300/1000 = 0.3 RR

Describing genetic structure

- genotype frequencies
- **allele frequencies**



200 rr = 400 r

500 Rr = 500 r
500 R

300 RR = 600 R

allele frequencies:

900/2000 = 0.45 r

1100/2000 = 0.55 R

total = 2000 alleles

for a population with genotypes:

100 GG

160 Gg

140 gg

calculate:

Genotype frequencies

Phenotype frequencies

Allele frequencies

for a population with genotypes:

100 GG

160 Gg

140 gg

calculate:

Genotype frequencies

$$260\left\{\begin{array}{l}100/400 = 0.25\ GG\\ 160/400 = 0.40\ Gg\\ 140/400 = 0.35\ gg\end{array}\right\}0.65$$

Phenotype frequencies

260/400 = 0.65 green

140/400 = 0.35 brown

Allele frequencies

360/800 = 0.45 G

440/800 = 0.55 g

another way to calculate allele frequencies:

100 GG

160 Gg

140 gg

Genotype frequencies

0.25 GG	→	G	0.25
0.40 Gg	↗	G	0.40/2 = 0.20
0.35 gg	↘	g	0.40/2 = 0.20
	→	g	0.35

Allele frequencies

360/800 = 0.45 G
440/800 = 0.55 g

OR [0.25 + (0.40)/2] = 0.45
[0.35 + (0.40)/2] = 0.65

Population genetics – Outline

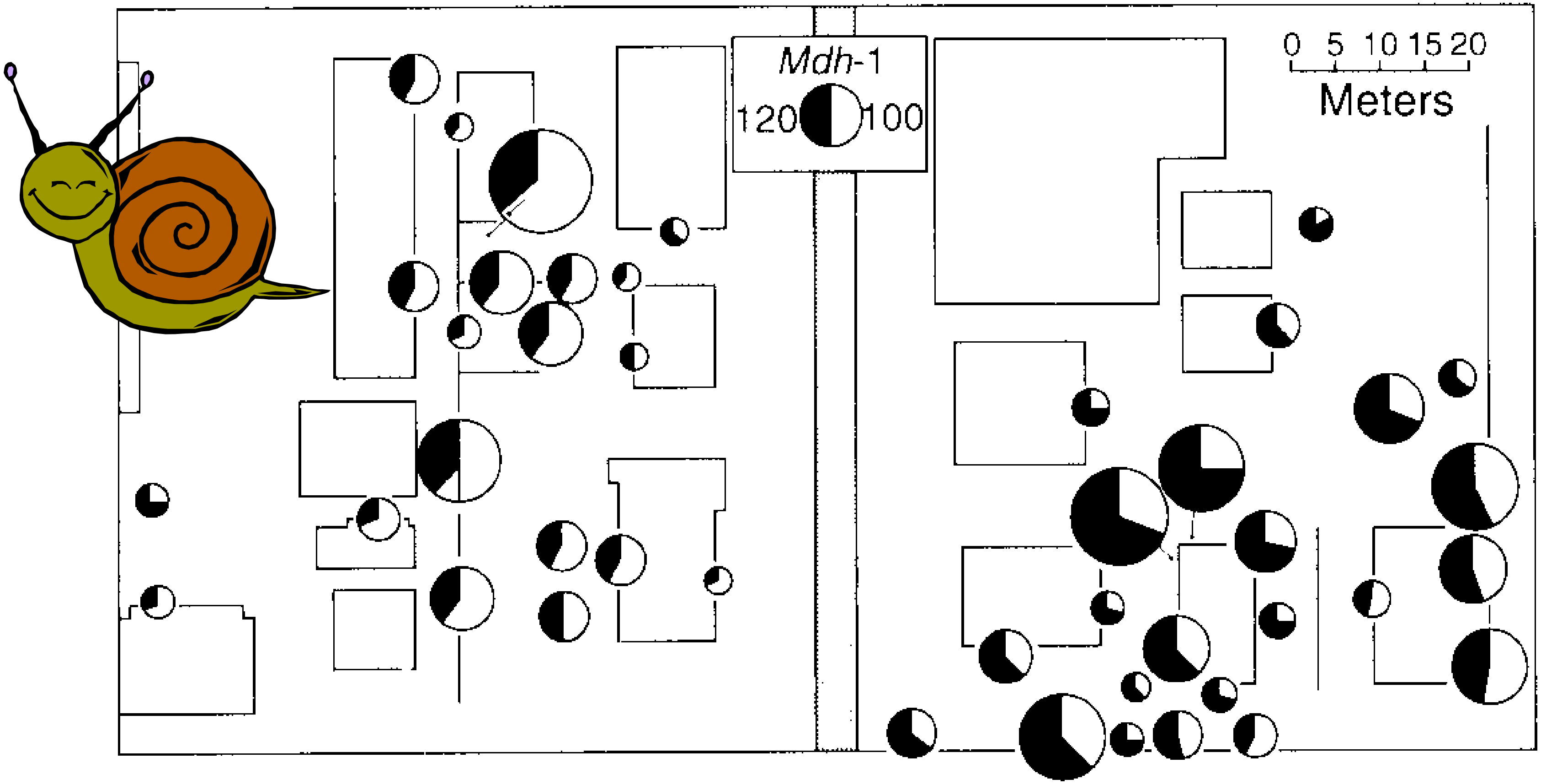
✓ What is population genetics?

✓ Calculate - genotype frequencies
- allele frequencies

Why is genetic variation important?

How does genetic structure change?

Genetic variation in space and time



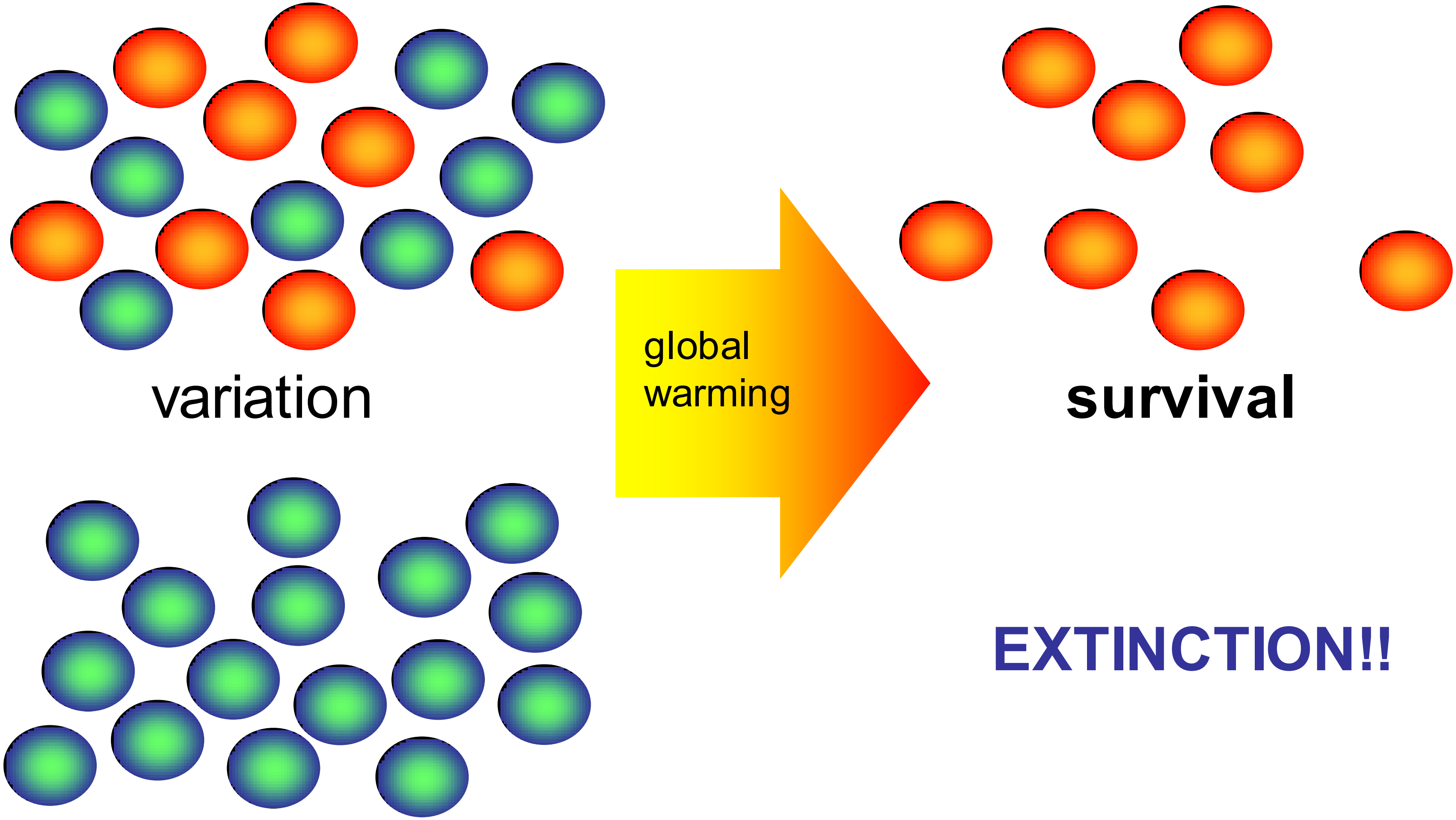
Frequency of *Mdh-1* alleles in snail colonies in two city blocks

Why is genetic variation important?
potential for change in genetic structure

- adaptation to environmental change
  - conservation

- divergence of populations
  - biodiversity

Why is genetic variation important?



variation

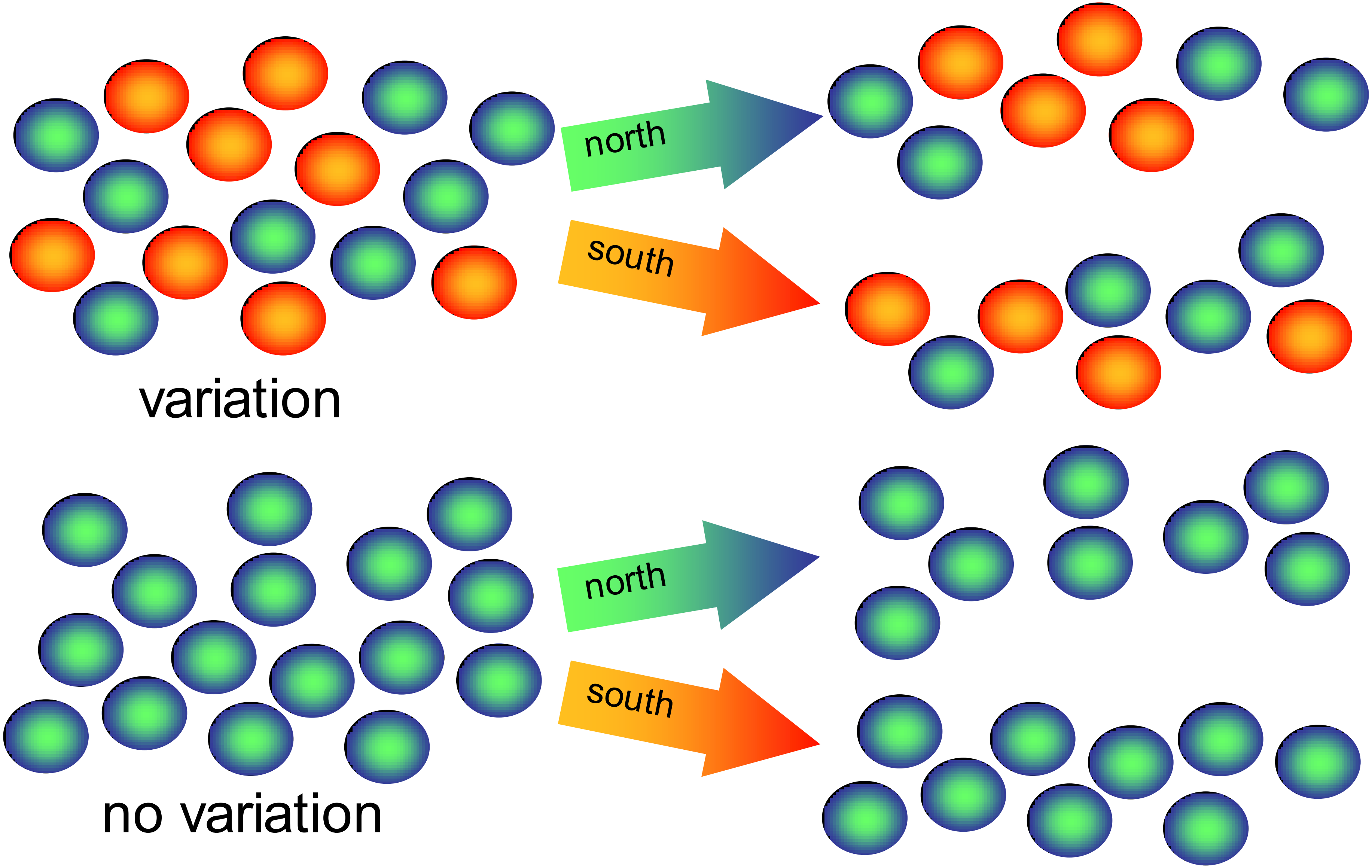
global warming

survival

no variation

EXTINCTION!!

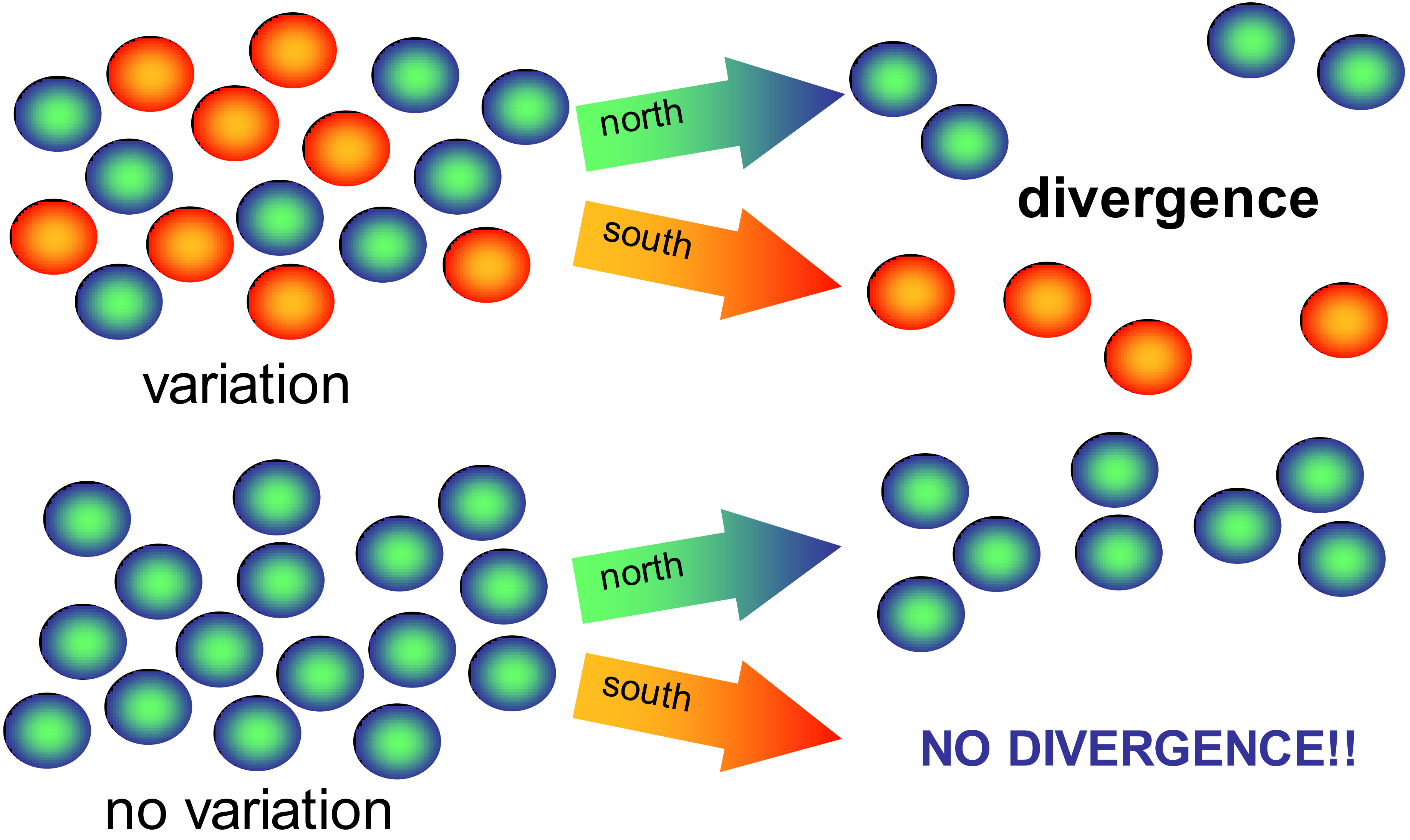
Why is genetic variation important?



variation

no variation

Why is genetic variation important?



variation

north

south

divergence

no variation

north

south

NO DIVERGENCE!!

How does genetic structure change?

How does **genetic structure** change?

changes in allele frequencies and/or genotype frequencies through time

How does genetic structure **change?**

changes in allele frequencies and/or genotype frequencies **through time**

- mutation
- migration
- natural selection
- genetic drift
- non-random mating

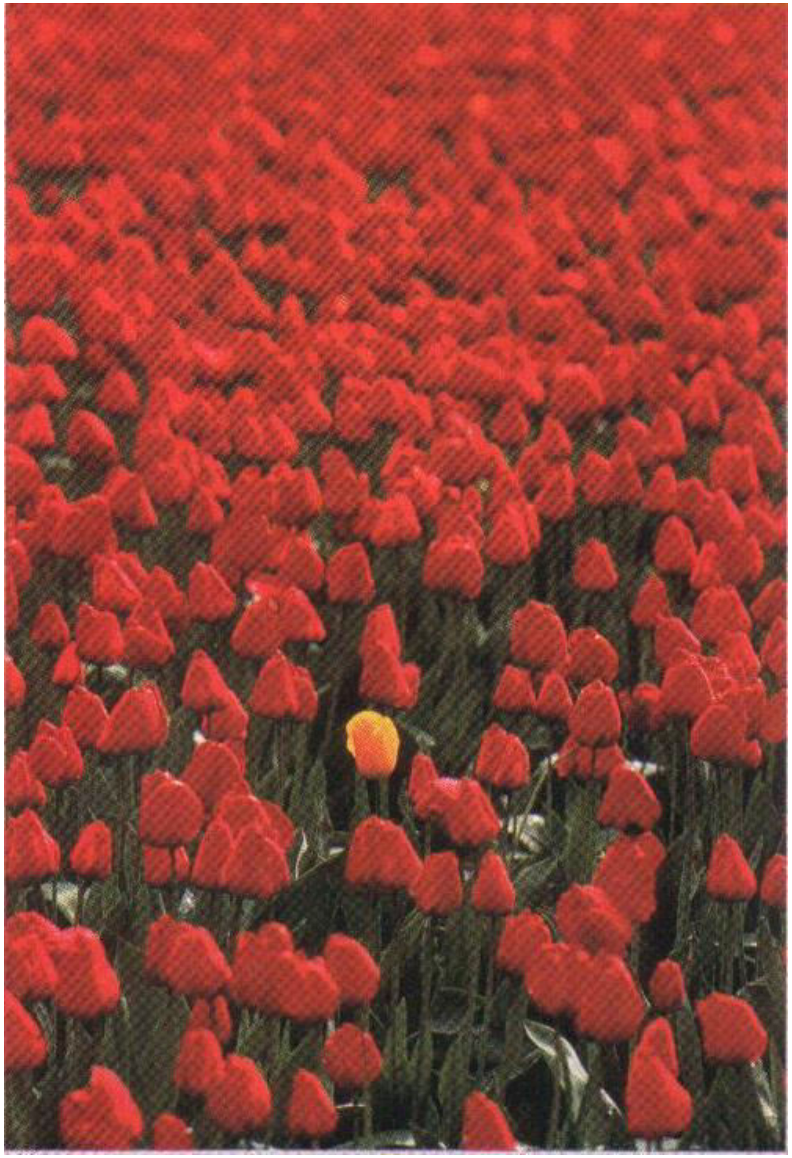
How does genetic structure change?

- **mutation**
- migration
- natural selection
- genetic drift
- non-random mating

spontaneous change in DNA

- creates new alleles
- ultimate source of all genetic variation

Mutations are the engines of variation



Without new variations, evolution would come to a halt:
all individuals would become the same

How does genetic structure change?

- mutation
- **migration**

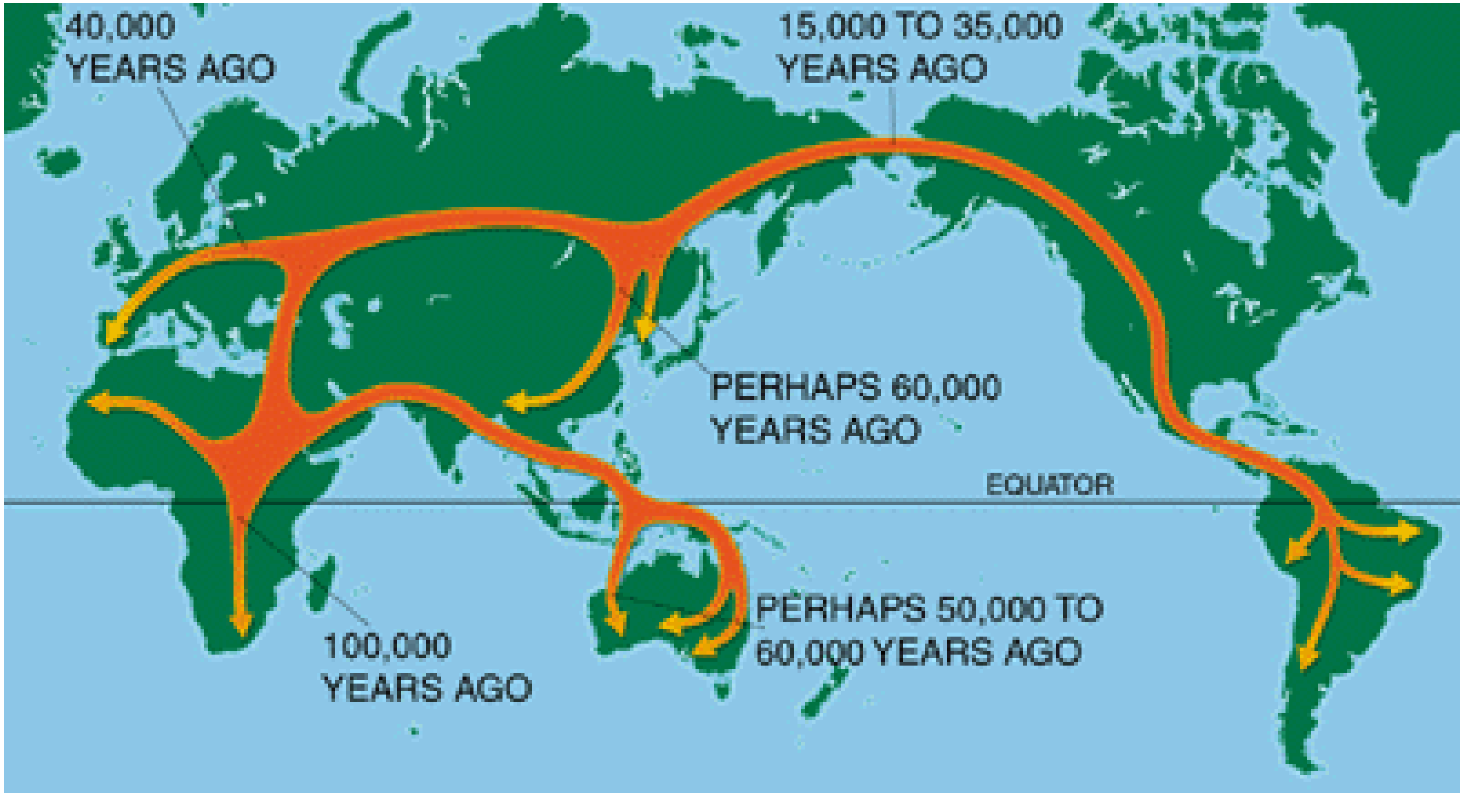
individuals move into population

- introduces new alleles

“gene flow”

- natural selection
- genetic drift
- non-random mating

Out of Africa



Modern humans arose in Africa and replaced other human species across the globe.

How does genetic structure change?

- mutation
- migration
- **natural selection**
- genetic drift
- non-random mating

certain genotypes produce more offspring

- differences in survival or reproduction

differences in "fitness"

- leads to adaptation

Natural Selection

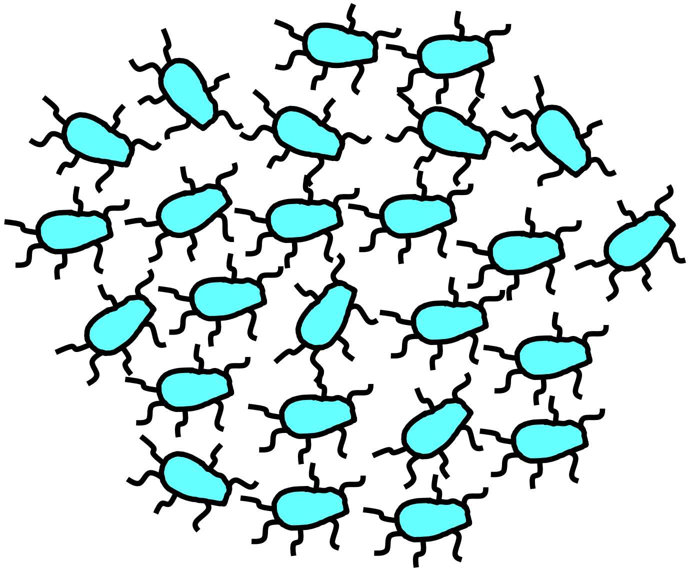
- Natural Selection refers to unequal success in terms of reproduction (a.k.a. survival of the fittest)
- Adaptations are inherited characteristics that enhance an organism's ability to survive and reproduce in a certain environment

Natural selection

Resistance to antibacterial soap

Generation 1:	1.00	not resistant
	0.00	resistant

Natural selection



Resistance to antibacterial soap

Generation 1: 1.00 not resistant
0.00 resistant

Natural selection

mutation!

Resistance to antibacterial soap

Generation 1: 1.00 not resistant
0.00 resistant

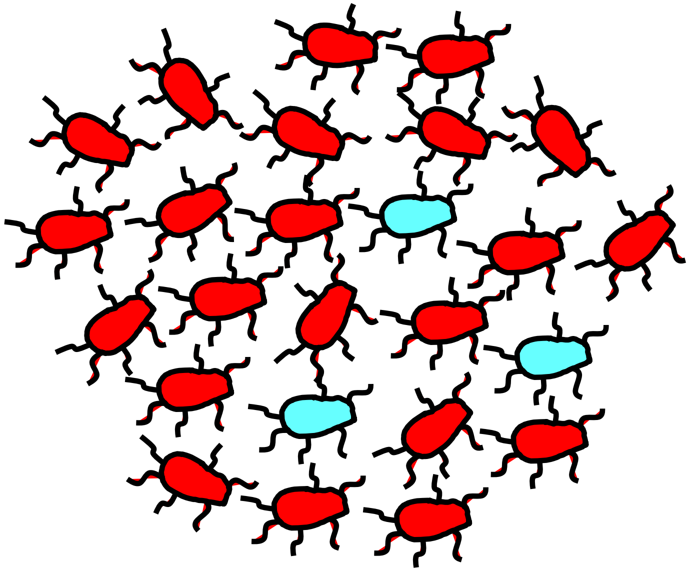
Generation 2: 0.96 not resistant
0.04 resistant

Natural selection

Resistance to antibacterial soap

Generation 1:	1.00	not resistant
	0.00	resistant
Generation 2:	0.96	not resistant
	0.04	resistant
Generation 3:	0.76	not resistant
	0.24	resistant

Natural selection



Resistance to antibacterial soap

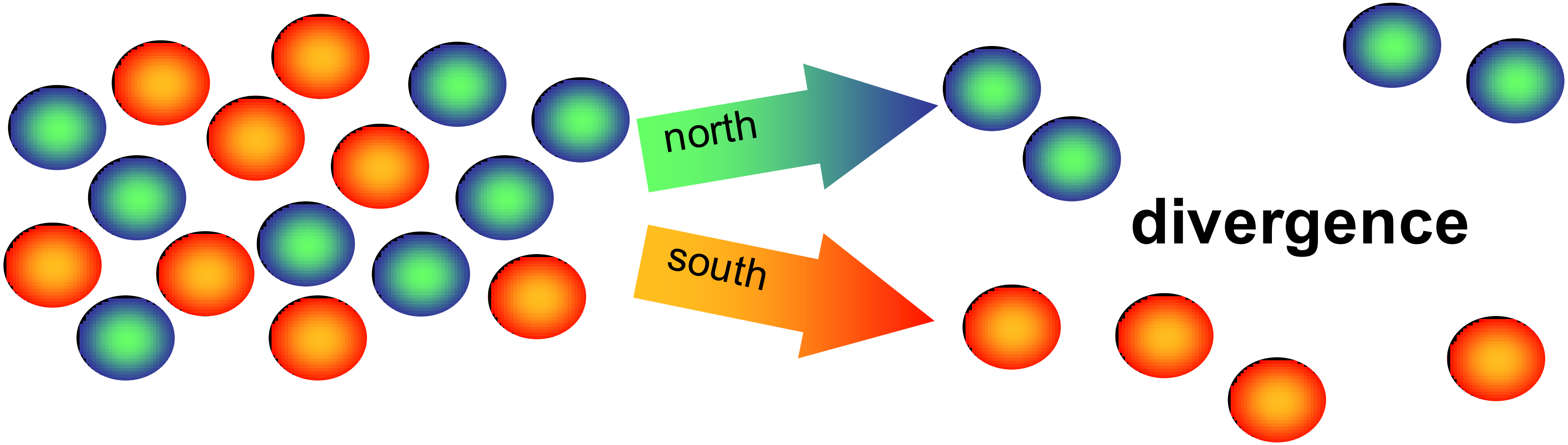
Generation 1: 1.00 not resistant
0.00 resistant

Generation 2: 0.96 not resistant
0.04 resistant

Generation 3: 0.76 not resistant
0.24 resistant

Generation 4: 0.12 not resistant
0.88 resistant

Natural selection can cause populations to diverge



Variation in Disease Alleles between Populations

Allele/Disease	Population Variation (allele frequency)
β^s allele of β-globin gene (sickle cell anemia)	1/20 in African Americans <1/200 in Hispanic Americans
Cystic fibrosis	1/40-50 in European Americans very low in Asian and African Americans
Tay-Sachs disease	1/60 in Ashkenazi Jews <1/6,000 in other groups
Myotonic dystrophy	1/50,000 in Europeans non-existent in Africans 1/950 in regions of Quebec

Selection on sickle-cell allele

aa – abnormal ß hemoglobin sickle-cell anemia	**very low fitness**
AA – normal ß hemoglobin vulnerable to malaria	**intermed. fitness**
Aa – both ß hemoglobins resistant to malaria	**high fitness**

Selection favors heterozygotes (Aa).

Both alleles maintained in population (a at low level).

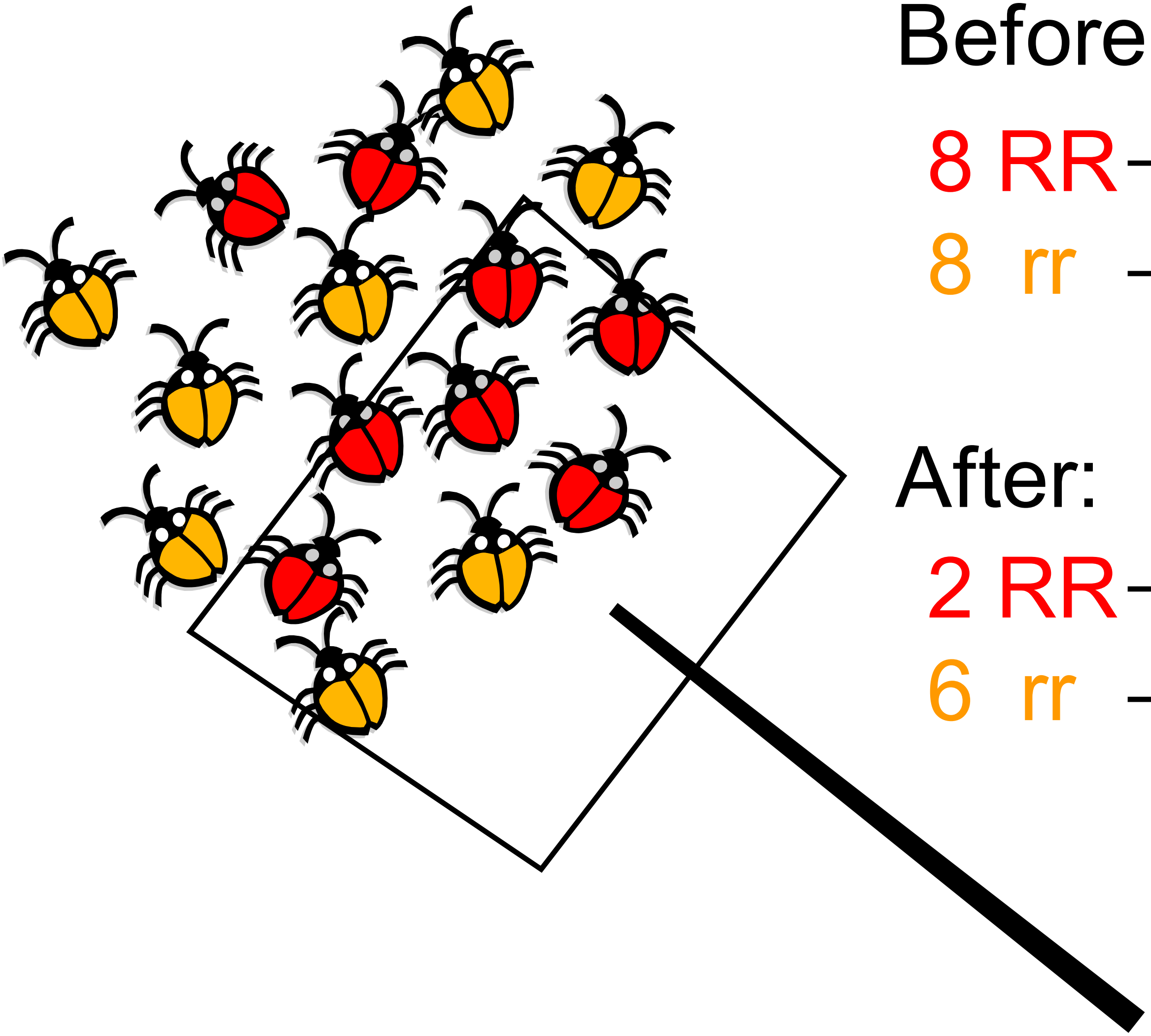
How does genetic structure change?

- mutation
- migration
- natural selection
- **genetic drift**
- non-random mating

genetic change by chance alone

- sampling error
  - misrepresentation
  - small populations

Genetic drift



Before:

8 RR → 0.50 R

8 rr → 0.50 r

After:

2 RR → 0.25 R

6 rr → 0.75 r

How does genetic structure change?

- mutation
- migration
- natural selection
- genetic drift

cause changes in allele frequencies

- non-random mating

How does genetic structure change?

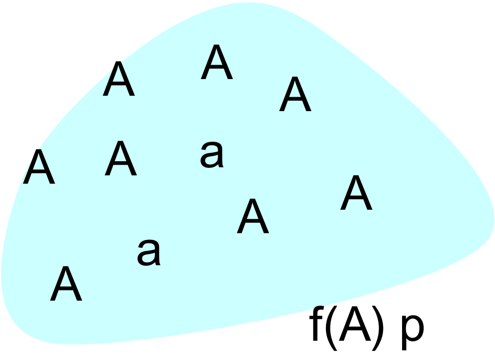
- mutation
- migration
- natural selection
- genetic drift
- **non-random mating**

mating combines alleles into genotypes

- non-random mating

↓

non-random
allele combinations



aa x aa AA x AA

aa AA

	f(A) p 0.8	f(a) q 0.2
A 0.8	AA (p^2) 0.8 x 0.8	Aa (pq) 0.8 x 0.2
a 0.2	aA (qp) 0.2 x 0.8	aa (q^2) 0.2 x 0.2

allele frequencies:
f(A) p= 0.8
f(a) q= 0.2

genotype frequencies:
f(AA) p^2 = 0.8 x 0.8 = 0.64
f(Aa) 2pq= 2(0.8 x0.2) = 0.32
f(aa) q^2 = 0.2 x 0.2 = 0.04

Conditions for Hardy Weinberg Equilibrium

1. Population size must be very large (infinite)
2. Mating is random
3. No gene flow: migration of alleles into or out of the population does not occur.
4. No mutations to alter the gene pool
5. No selection: no selective force favors one individual over another.

What is the distribution of genotypes?

For two alleles (🟠 A_1 and 🔵 A_2) we have three genotypes:

🟠🟠 A_1A_1 🟠🔵 A_1A_2 🔵🔵 A_2A_2

Let p = the frequency of A_1 and

q = the frequency of A_2 🟠

$p + q = 1$ 🔵

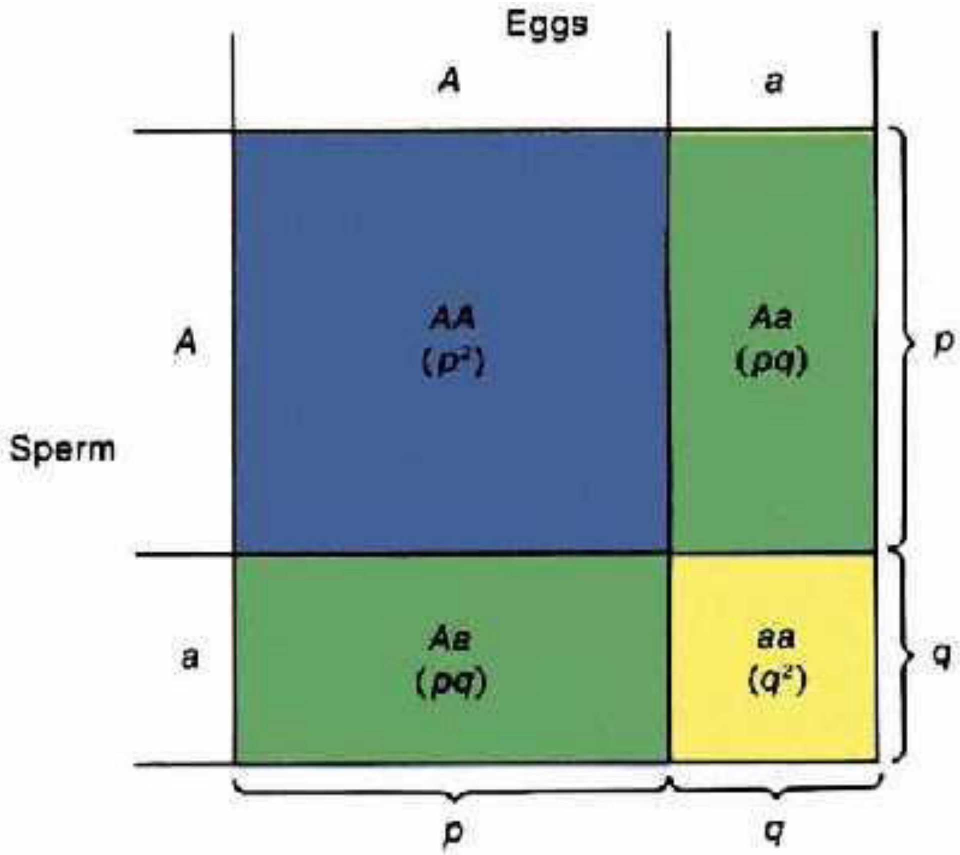
What is the frequency of 🟠🟠 A_1A_1 🟠🔵 A_1A_2 🔵🔵 A_2A_2 ?

🟠🟠 A_1A_1 $= p^2$

🔵🔵 A_2A_2 $= q^2$

🟠🔵 A_1A_2 $= pq + qp = 2pq$

The Hardy-Weinberg equilibrium



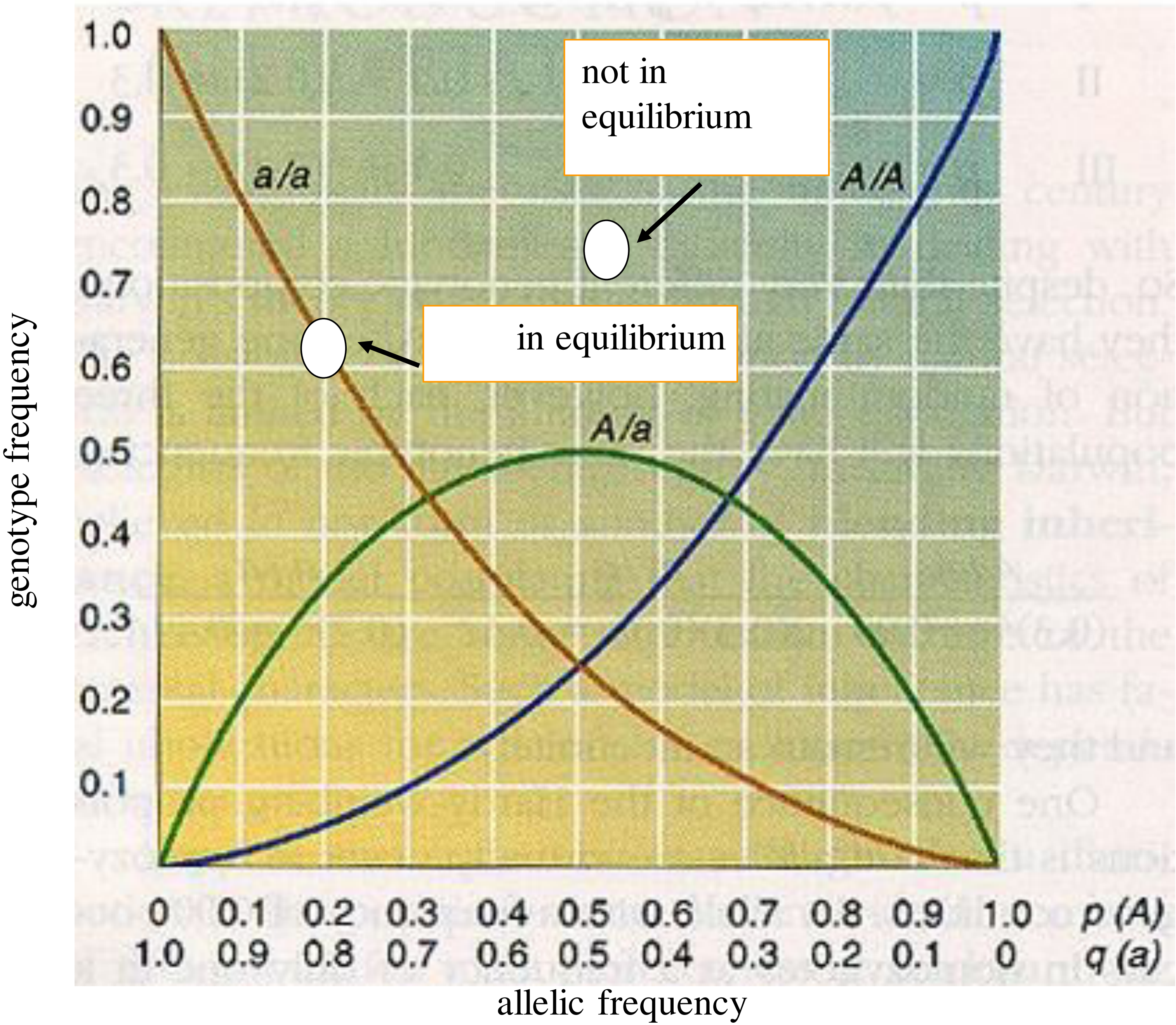
$p^2 + 2pq + q^2 = 1$

$(p + q)^2 = 1$

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

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

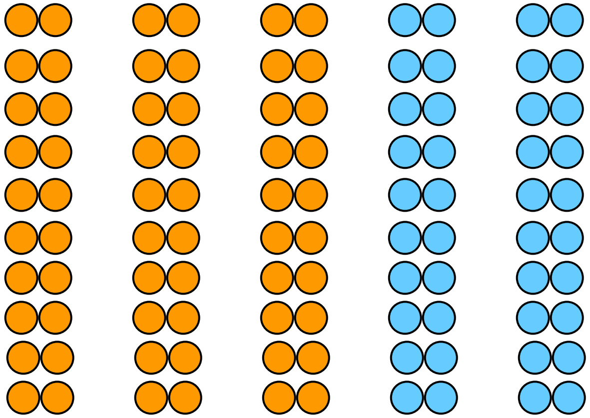
Hardy-Weinberg equilibrium



not in equilibrium

in equilibrium

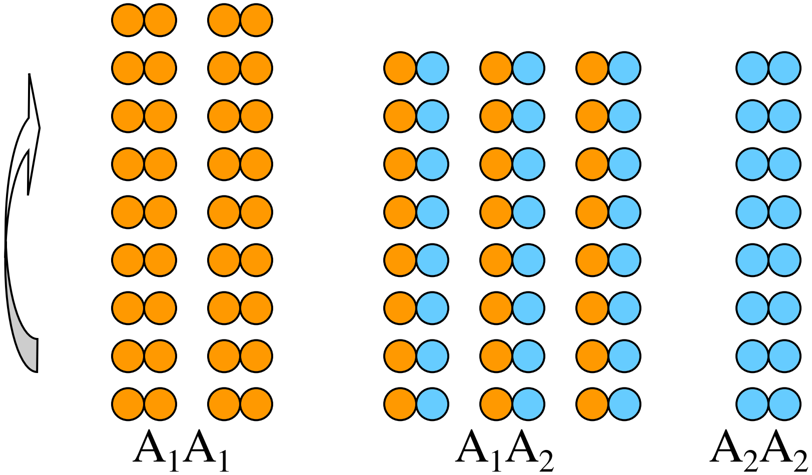
Population achieves Hardy-Weinberg equilibrium after one generation



Not in equilibrium

A_1A_1	A_1A_2	A_2A_2
0.6	0	0.4

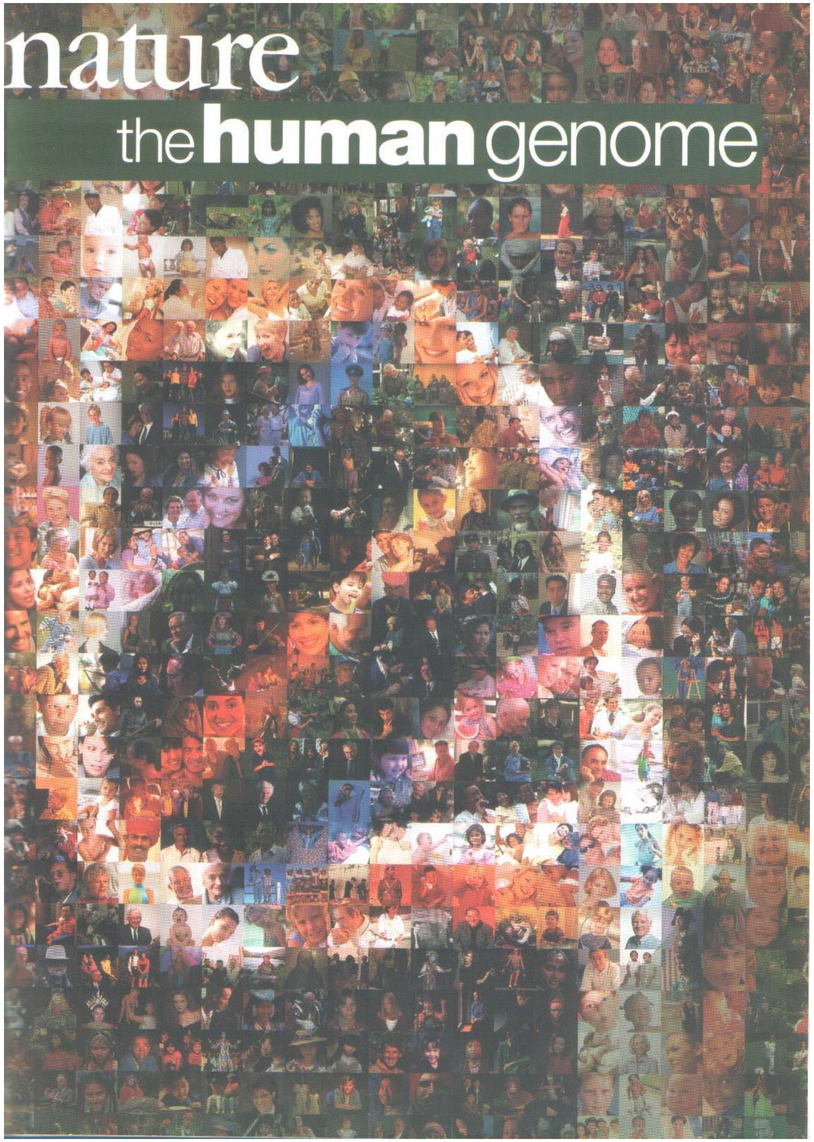
Next generation

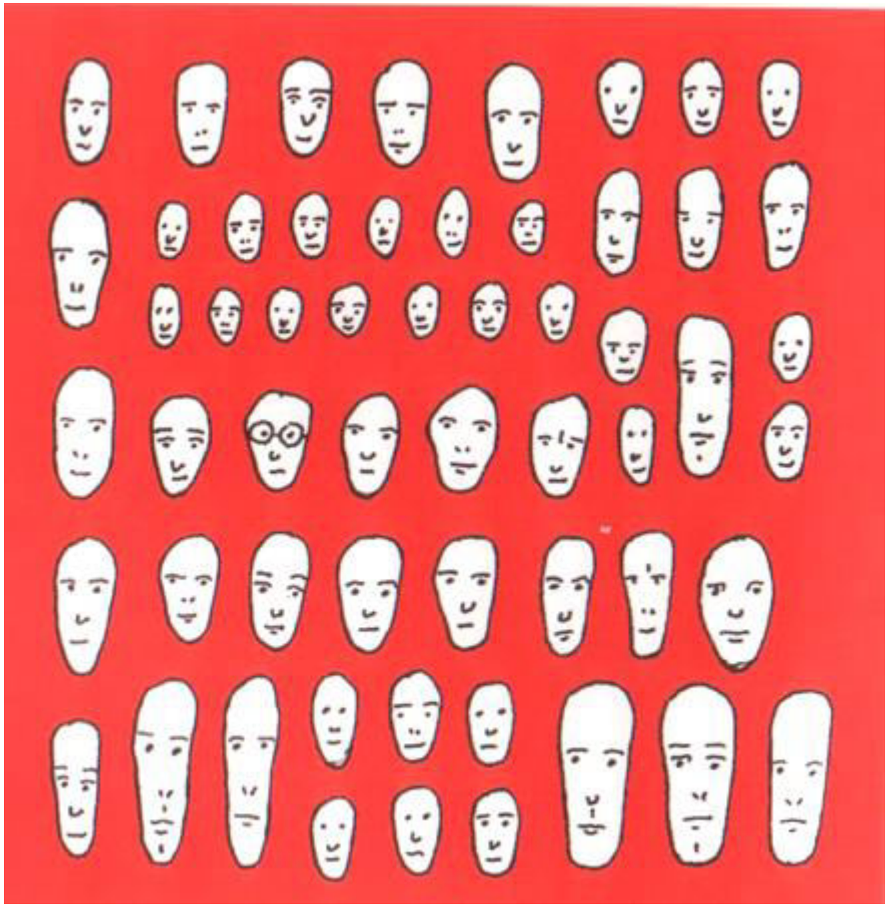


In equilibrium

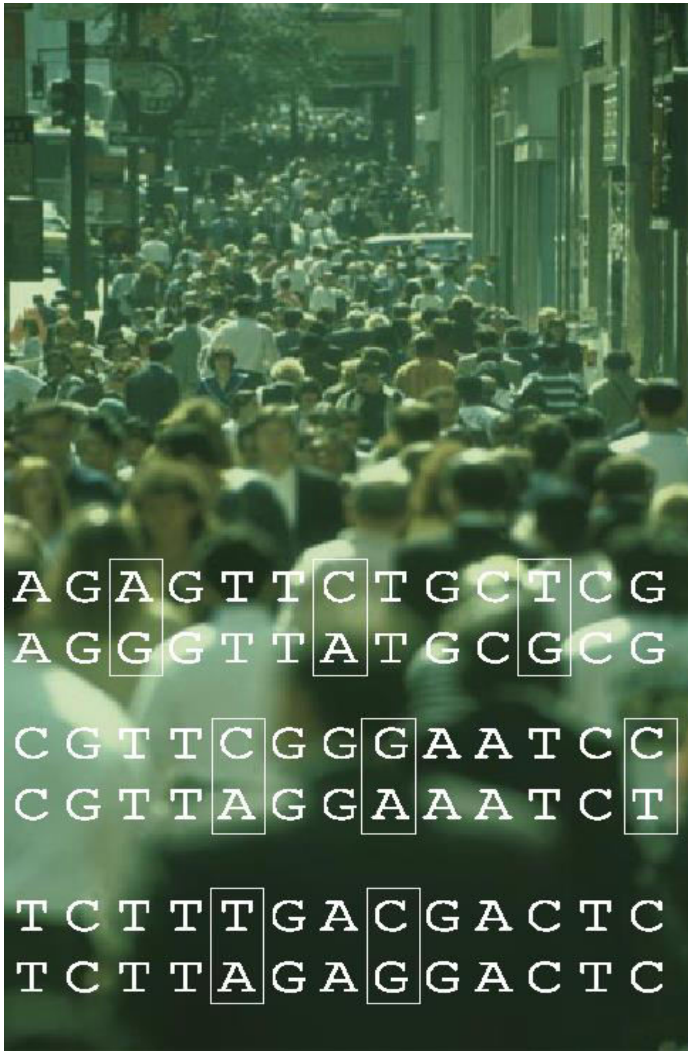
A_1A_1	A_1A_2	A_2A_2
0.36	0.48	0.16

DNA POLYMORPHISM IN POPULATIONS





How are genomes of individuals different?



Most of the differences take the form of substitutions at a single base.
Single Nucleotide Polymorphism (SNP)

Variation is the splice of life

If :

1. Any two genomes are roughly 99.9% identical and

2. A genome is 3.2 billion base pairs long

Then:

Every two genomes have 3.2 million differences (SNPs)

(remember also that each individual has two genomes)

How many SNPs are there in human populations?

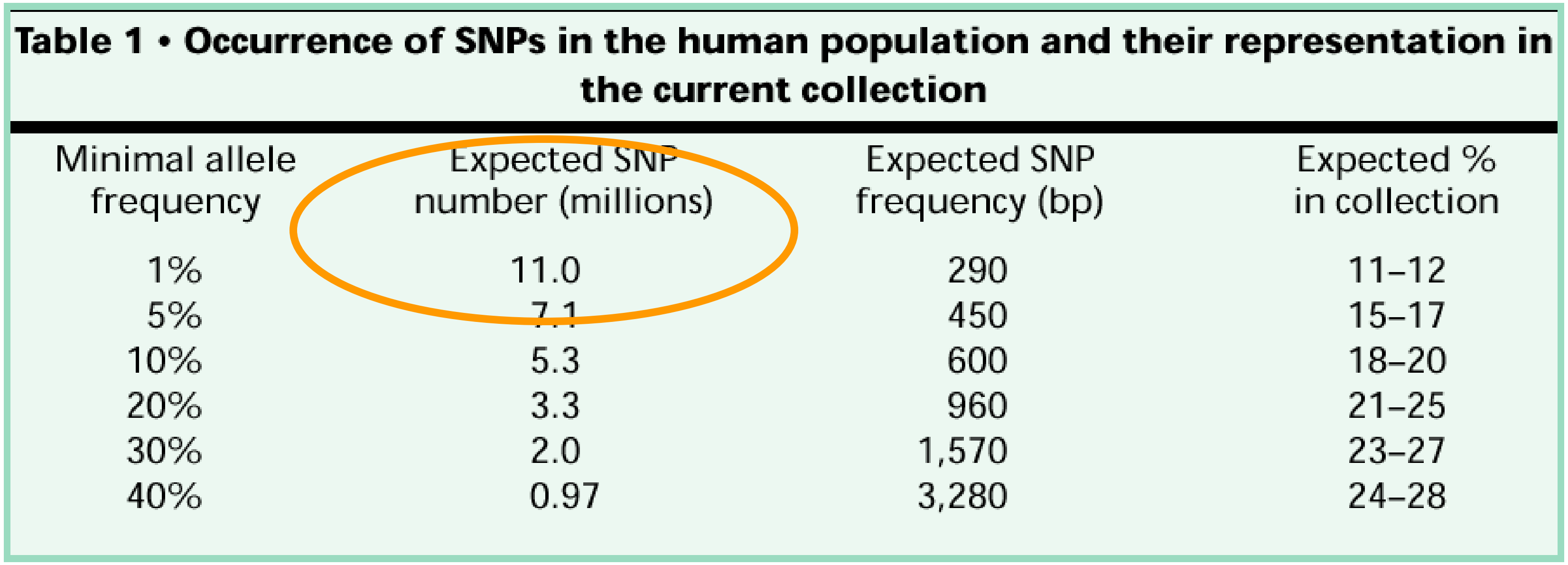
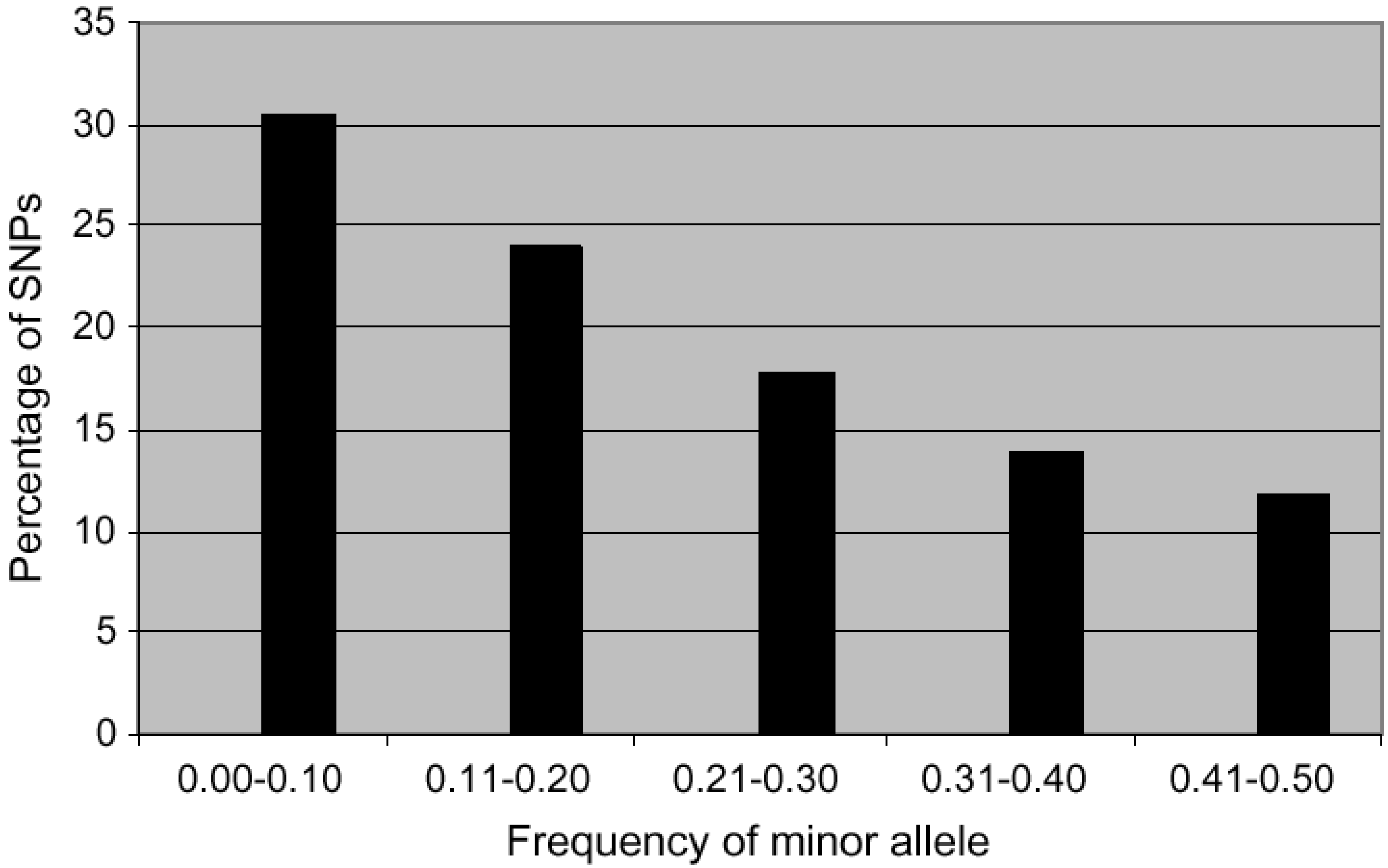


Table 1 • Occurrence of SNPs in the human population and their representation in the current collection

Minimal allele frequency	Expected SNP number (millions)	Expected SNP frequency (bp)	Expected % in collection
1%	11.0	290	11–12
5%	7.1	450	15–17
10%	5.3	600	18–20
20%	3.3	960	21–25
30%	2.0	1,570	23–27
40%	0.97	3,280	24–28

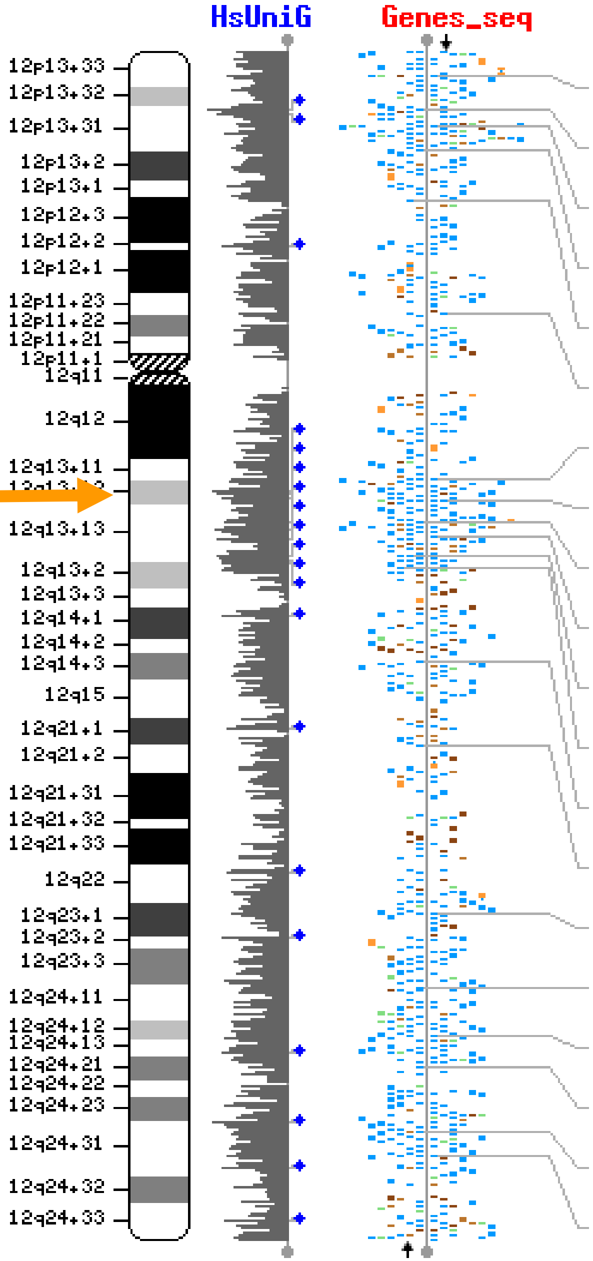
(based upon the classical neutral theory of population genetics)

How are the frequencies of SNPs distributed?



Identified 35,989 SNPs in a sample of 20 copies of chromosome 21.

A locus is a region on the chromosomes



One locus

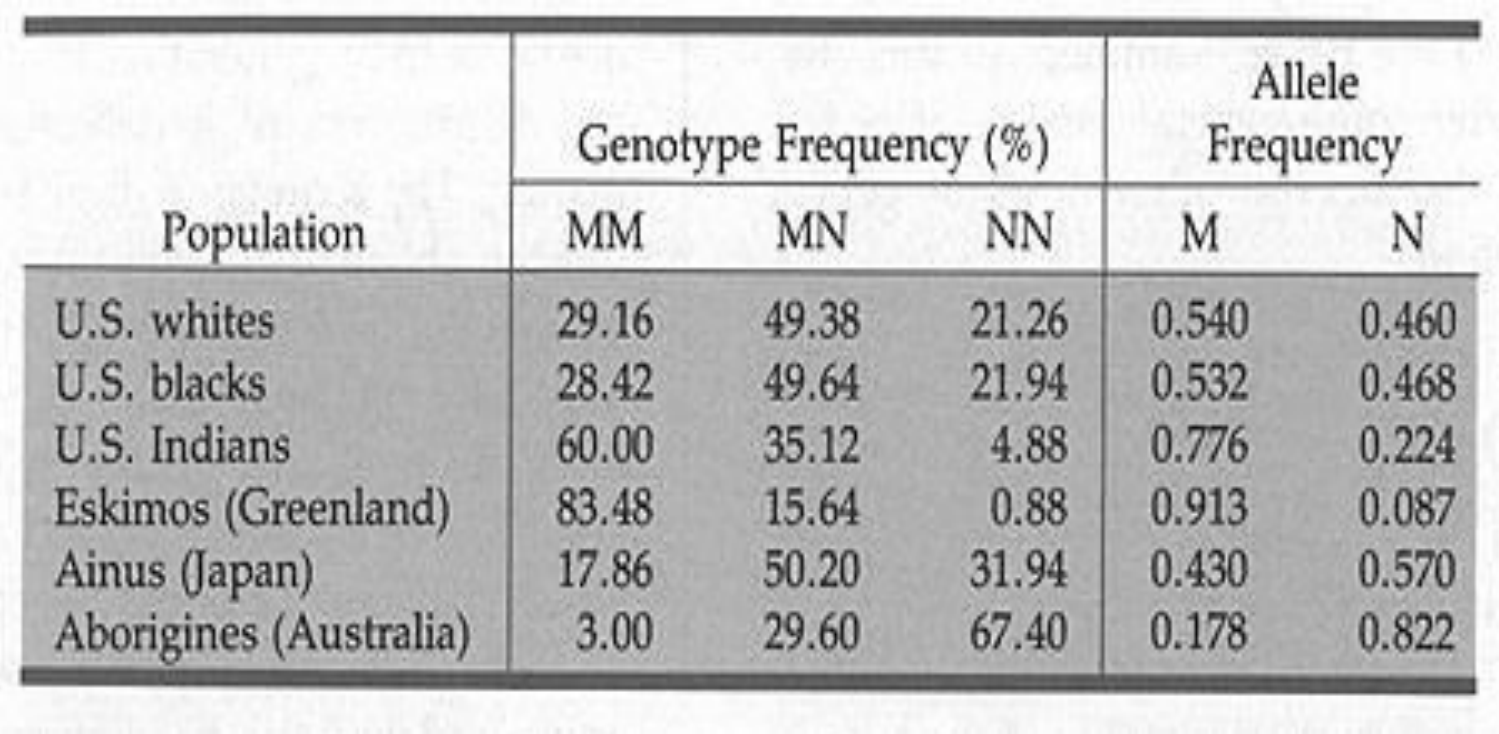
An **allele** is a variation at a locus

For example in the same locus we may have:

Allele 1: GTTTCTGATTTTTTGATGTCTTCATCCATCACTGTCCTTGTCAAATAGTTT……

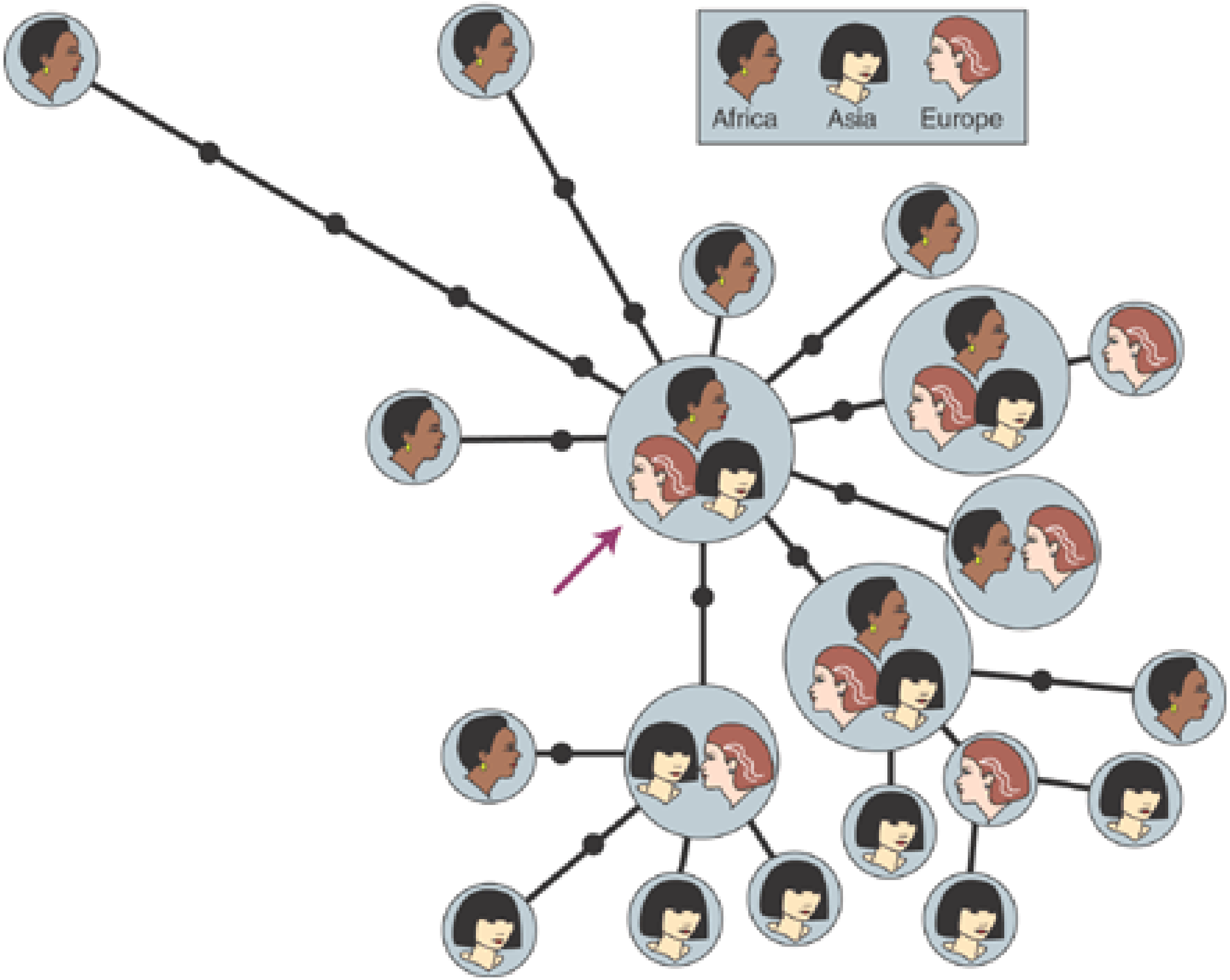
Allele 2: GTTTCTGATTTTTTGATGTCTTCAGCCATCACTGTCCTTGTCAAATAGTTT……

allele frequency is the proportion
of a certain allele within a population.

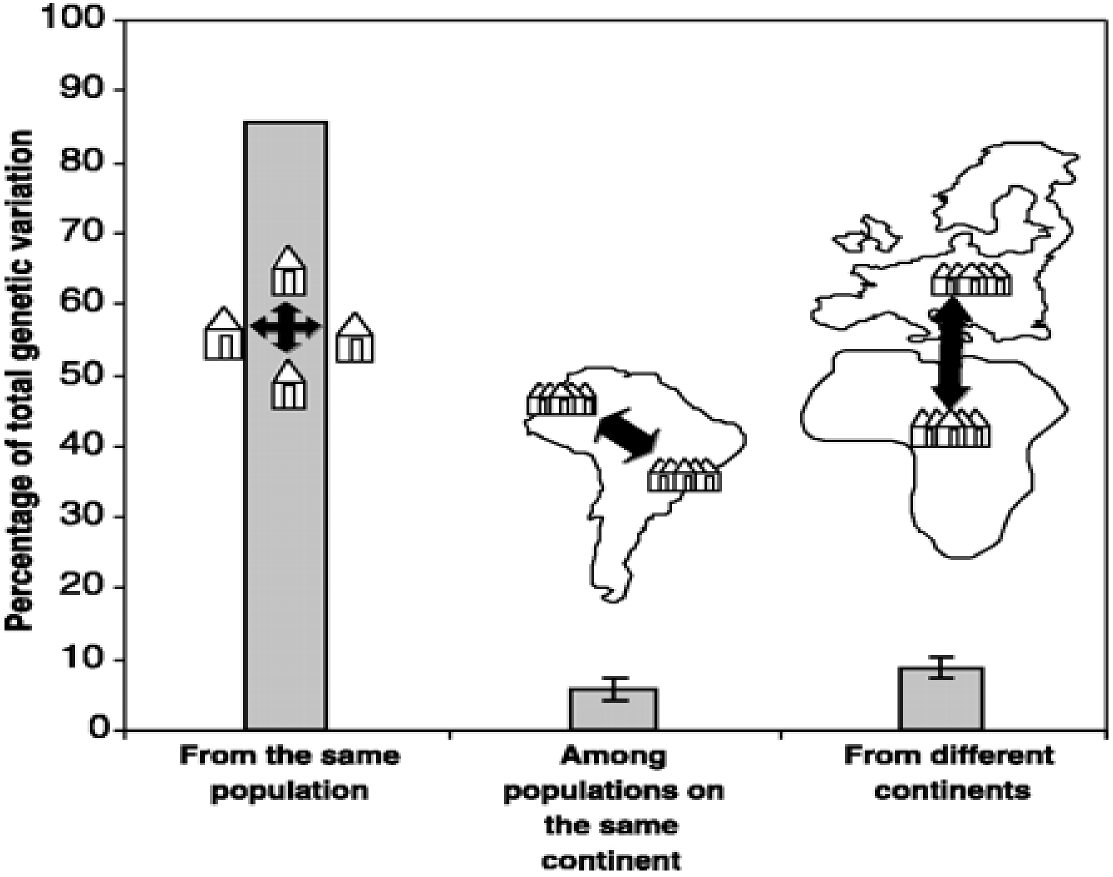


Population	Genotype Frequency (%)			Allele Frequency	
	MM	MN	NN	M	N
U.S. whites	29.16	49.38	21.26	0.540	0.460
U.S. blacks	28.42	49.64	21.94	0.532	0.468
U.S. Indians	60.00	35.12	4.88	0.776	0.224
Eskimos (Greenland)	83.48	15.64	0.88	0.913	0.087
Ainus (Japan)	17.86	50.20	31.94	0.430	0.570
Aborigines (Australia)	3.00	29.60	67.40	0.178	0.822

People from different continents often carry identical DNA sequences



Percentage of human genetic variation within and between populations.



An average population from anywhere in the world includes 85% of all human variation at autosomal loci and 81% of all human variation in mtDNA sequences. Differences among populations from the same continent contribute another 6% of variation; only 9 to 13% of genetic variation differentiates populations from different continents