Chapter 20

Variation and Selection in Populations

Sections to study

20.1 The Hardy-Weinberg law: Predicting genetic variation in "ideal" populations

20.2 What causes allele frequency to change in real populations?

~~**20.3 Ancestry and the evolution of modern humans**~~

- Prior to the 20th century, many people thought that “recessive” phenotypes such as naturally blonde or red hair would become extinct over time in human population.
  - The blended inheritance theory claims that the information for blonde and red hair would be lost gradually when people with lighter hair color mated with people with darker hair.
- In 1908, Geoffrey H. Hardy and Wilhelm Weinberg proposed that if certain assumptions were met, these phenotypes would remain constant over time and between generations.

- **Population genetics**: The scientific discipline that studies what happens in whole populations at the genetic level.
  - Population geneticists rely on mathematical models in predicting a population's potential for stasis or change.
  - Simple models help clarifying the questions about frequency of genetic diseases or rate of spread of pathogens, as well as making predictions about future populations.

20.1 The Hardy-Weinberg law: Predicting genetic variation in “ideal” populations

- **Population**: A group of interbreeding individuals of the same species that inhabit the same space at the same time. Also called ***Mendelian population***.

Girrafe herd on the African savanna

- **Gene pool: The sum total of all alleles carried in all members of a population.**
  - **A gene pool represents all of the alleles present on the chromosomes of all members of a population and the relative prominence or rareness of each allele.**
- **A gene pool can change over time.**
  - **New alleles arise by mutation or are introduced by immigration.**
  - **Rare pre-existing alleles disappear when all individuals carrying them leave the population or die.**

Microevolution: Alterations of a population's gene pool.

A population is defined by its genotype frequencies and allele frequencies, which together make up a gene pool.

- **Phenotype frequency: The proportion of individuals in a population that express a particular phenotype.**
- **Genotype frequency: The proportion of total individuals in a population that carry a particular genotype.**
- **Allele frequency: The proportion of gene copies in a whole population that are of a given allele type.**

- **Phenotype frequency: The proportion of individuals in a population that express a particular phenotype.**

Example: In a population of 20 human individuals, 4 people have blue-colored eyes because they are homozygous for the recessive *B* allele at a particular "blue eyes" locus, where the alternative allele is *A*. Molecular analyses showed that 12 individuals are of genotype *AA*, 4 are of genotype *AB*.

Phenotype frequency of blue eyes: 4/20 = 0.2

dark eyes: (20-4)/20 = 0.8

- **Genotype frequency: The proportion of total individuals in a population that carry a particular genotype.**

Example: In a population of 20 human individuals, 4 people have blue-colored eyes because they are homozygous for the recessive *B* allele at a particular "blue eyes" locus, where the alternative allele is *A*. Molecular analyses showed that 12 individuals are of genotype *AA*, 4 are of genotype *AB*.

Genotype frequency of *AA*: 12/20 = 0.6
_____ *AB*: 4/20 = 0.2
_____ *BB*: 4/20 = 0.2

- **Allele frequency: The proportion of gene copies in a whole population that are of a given allele type.**

Example: In a population of 20 human individuals, 4 people have blue-colored eyes because they are homozygous for the recessive *B* allele at a particular "blue eyes" locus, where the alternative allele is *A*. Molecular analyses showed that 12 individuals are of genotype *AA*, 4 are of genotype *AB*.

genotype	"AA"	"AB"	"BB"	Total
number of individuals	12	4	4	20

allele	"A"	"B"	Total
number of chromosomes	28	12	40
allele frequency	0.7	0.3	1.0

Allele frequency:

***A*: (12×2 + 4)/(20×2) = 0.7**

***B*: (4 + 4×2)/(20×2) = 0.3**

Fig. 20.2

Calculating allele frequencies from genotypes frequencies

Genotypes in first generation: RR Rr rr

Number of individuals in first generation: N_{RR} N_{Rr} N_{rr}

Allele types in first generation

Allele frequencies in first generation

$$p = \text{frequency of } R \text{ gametes} = \frac{2N_{RR} + N_{Rr}}{2N_{\text{total}}}$$

$$q = \text{frequency of } r \text{ gametes} = \frac{N_{Rr} + 2N_{rr}}{2N_{\text{total}}}$$

$N_{\text{total}} = N_{RR} + N_{Rr} + N_{rr}$

$2N_{\text{total}}$ = total chromosomes

The Hardy-Weinberg law correlates allele and genotype frequencies

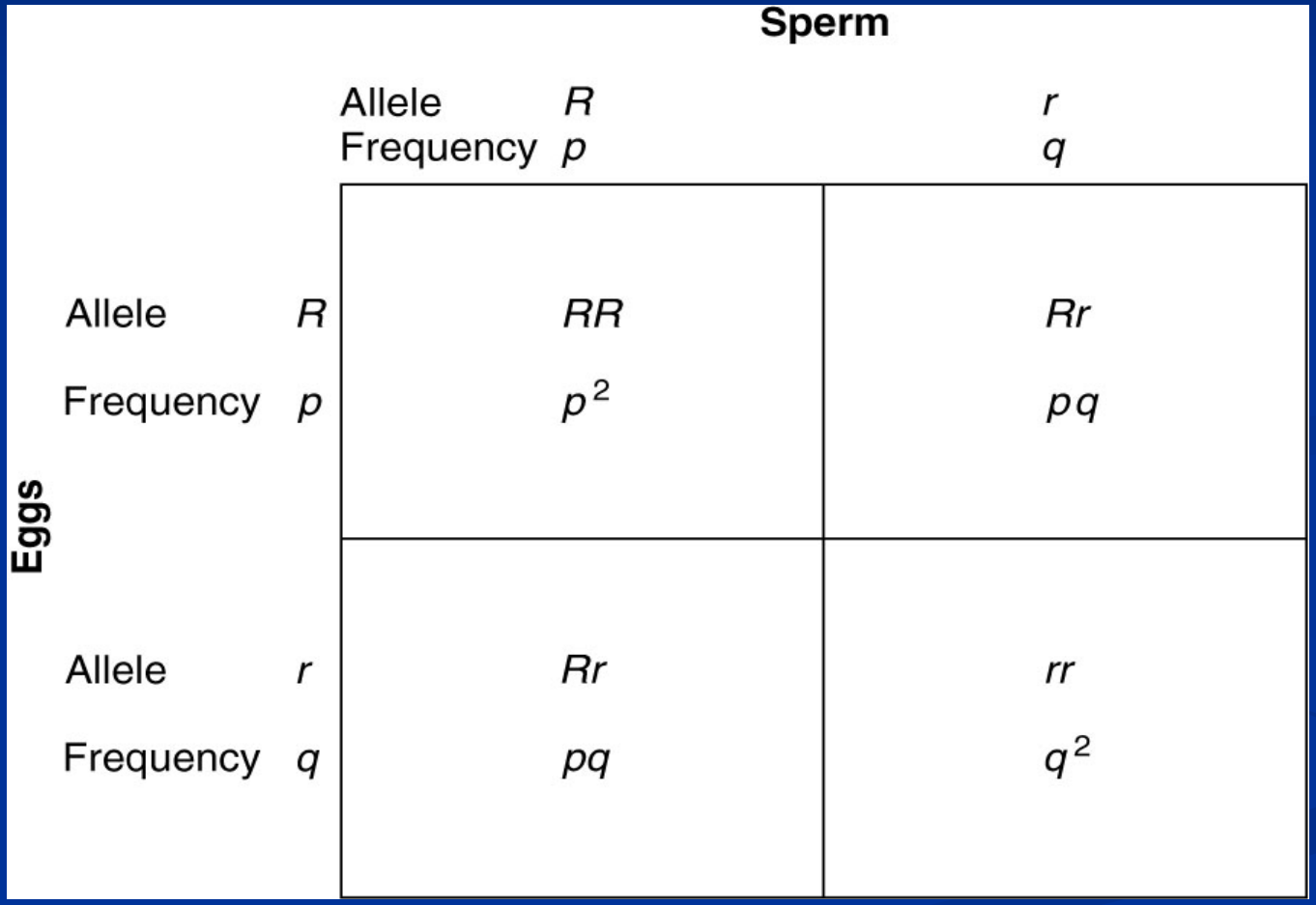
- A model for understanding allele, genotype, and phenotype frequencies for single gene traits in a genetically stable population.
- Developed independently in 1908 by

Godfrey H. Hardy

Wilhelm Weinberg

- **Five key assumptions of population**
  1. **Populations large enough to minimize randomized events.**
  2. **Individuals mate at random.**
  3. **No new mutations appear in gene pool.**
  4. **No migration in or out of population**
  5. **All genotypes have equal chances of survival and reproduction.**
- **If all assumptions hold, population is in Hardy-Weinberg equilibrium.**
  - **Allele frequencies do not change.**
  - **Genotype frequencies can be predicted.**

Using gamete allele frequencies to calculate genotype frequencies in the next generation



Example: In a population of 20 human individuals, 4 people have blue-colored eyes because they are homozygous for the recessive *B* allele at a particular "blue eyes" locus, where the alternative allele is *A*. Molecular analyses showed that 12 individuals are of genotype *AA*, 4 are of genotype *AB*.

Genotype frequencies of the next generation:

(p): frequency of allele "A" = 0.7

(q): frequency of allele "B" = 0.3

			eggs: p (0.7)	eggs: q (0.3)
sperm	p	0.7	"AA" 49%	"AB" 21%
sperm	q	0.3	"AB" 21%	"BB" 9%

homozygote "AA" = p^2 = 0.49

homozygote "BB" = q^2 = 0.09

heterozygote "AB" = 2(pq) = 0.42

1.00

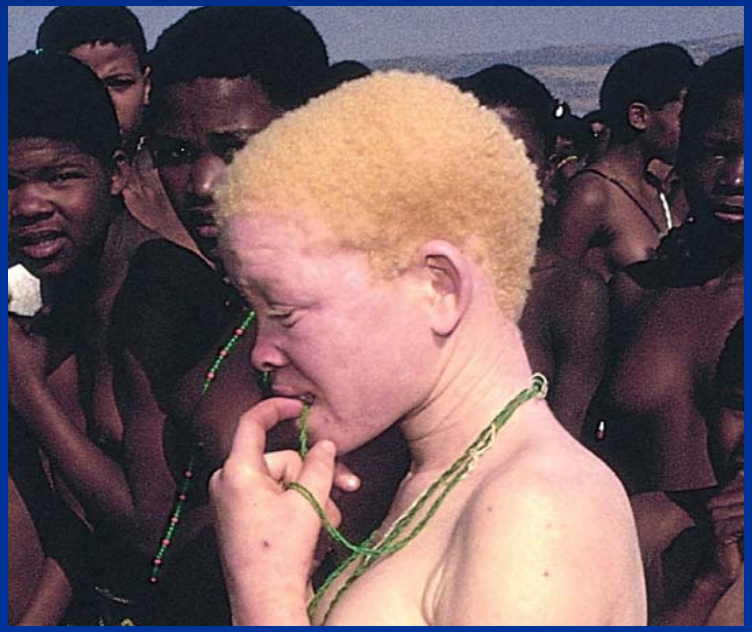
Fig. 20.3

The genotype frequencies of zygotes arising in a large population of sexually reproducing diploid organisms are p^2 for *AA*, $2pq$ for *AB*, and q^2 for *BB*.

Allele frequency equation: $p + q = 1$

Genotype frequency equation: $p^2 + 2pq + q^2 = 1$

An example: Predicting the frequency of albinism



An albino African girl

- **Population of 100,000 people**
  - **98,100 *AA* individuals**
  - **1,800 *Aa* carriers**
  - **100 *aa* albinos**
- **The frequency of alleles**
  - ***A* allele frequency: (98,100 × 2 + 1,800)/200,000 = 0.99**
    - **$p = 0.99$**
  - ***a* allele frequency: (1,800 + 100 × 2)/200,000 = 0.01**
    - **$q = 0.01$**

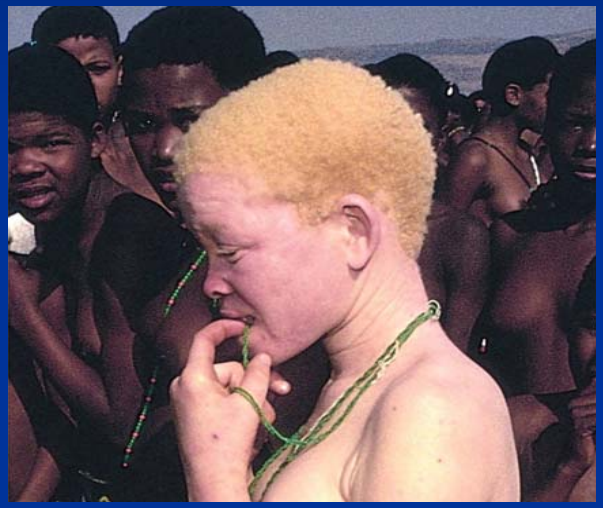
- **Hardy-Weinberg equation for the albino gene in the population is:**
  - $p^2 + 2pq + q^2 = (0.99)^2 + 2\times(0.99 \times 0.01) + (0.01)^2 =$
  - **$0.9801 + 0.0198 + 0.0001 = 1$**
- **Next generation of 100,000 people:**
  - **100,000 × 0.9801 *AA* individuals = 98,010**
  - **100,000 × 0.0198 *Aa* individuals = 1,980**
  - **100,000 × 0.0001 *aa* individuals = 10**

Frequency of *p* and *q* allele in next generation

- $p + q = 1$
  - Thus the frequency of the p allele in the next generation is
    - $p^2 + ½[2p(1-p)] = p^2 + p(1-p) = p^2 + p – p^2 = p$
  - and the frequency of the q allele in the next generation is
    - $q^2 + ½[2q\ (1-q)] = q^2 + q\ (1-q) = q^2 + q – q^2 = q$
- For albinism where $p = 0.99$ and $q = 0.01$
  - The frequency of *A* allele in second generation is
    - 0.98 + 0.99 – 0.98 = 0.99
  - The frequency of the *a* allele is
    - 0.0001 + 0.01 – 0.0001 = 0.01
- Genotype frequencies changed, but allele frequencies stay the same for both dominant and recessive alleles.

- **Allele frequencies do not change from generation to generation in a population at Hardy-Weinberg equilibrium.**
- **A Hardy-Weinberg population achieves the genotype frequencies of p^2, $2pq$, and q^2 in just one generation and maintains them in subsequent generations.**

It is difficult to eliminate recessive disease alleles from the population



An albino African girl

A population of 1,000,000 people in Hardy-Weinberg equilibrium:

- 980,100 *AA* individuals
- 19,800 *Aa* carriers
- 100 *aa* albinos

Albino allele frequency: $q = 0.01$

If all albinos (*aa*) do not reproduce,

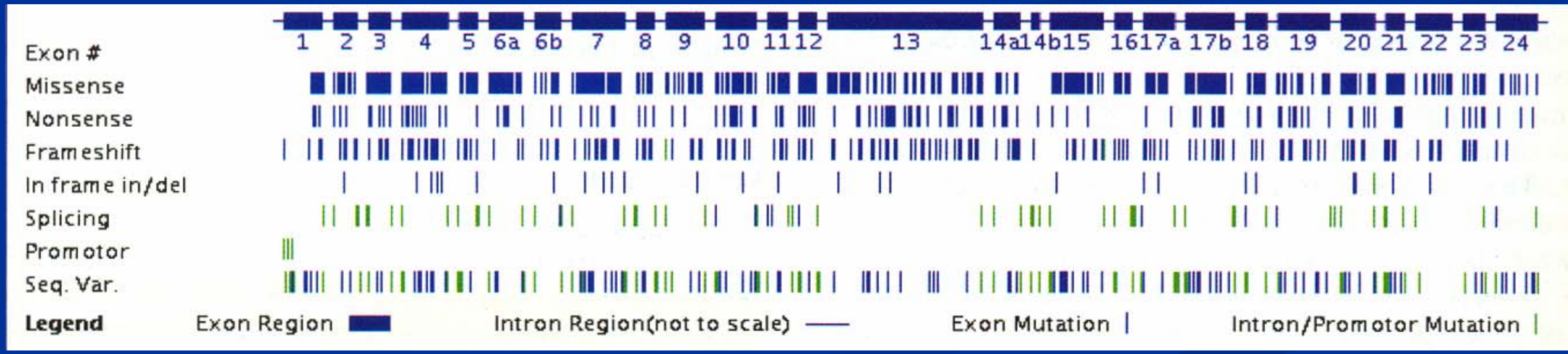
- After 100 generations, $q = 0.005$, 25 albinos/million
- After 9900 generations, $q = 0.0001$, 0.01 albinos/million

Using genotype frequencies to calculate allele frequencies

- Frequency of recessive genetic disease: q^2
- The disease allele frequency: $q = \sqrt{q^2}$
- $p + q = 1$, So, wild-type allele frequency: $p = 1 - q$
- Genotype frequencies
  - Homozygous dominant = p^2
  - Heterozygotes = $2pq$
  - Homozygous recessive = q^2

Predicting the frequency of CFTR alleles

- Cystic fibrosis
- Caused by recessive mutations in CFTR gene.
- Transmembrane receptor
- Over 1600 disease-causing mutations identified.



- ~ 1/2500 Caucasian newborns develop the disease.

$q^2 = 1/2500 = 0.0004$

So, the disease allele frequency: $q = \sqrt{q^2} = 0.02$

Wild-type allele frequency: $p = 1-q = 0.98$

Heterozygote genotype frequency: $2pq = 2 \times 0.98 \times 0.02 = 0.0392$

1 in 25 Caucasian people are carriers of the diseased allele.

20.2 What causes allele frequencies to change in real populations?

- Mutations introduce new genetic variation.
  - New mutations appear occasionally at every locus. Rate in mammals is 10^{-4} – 10^{-6} per gene per generation.
  - ~ 0.02 – 2 new mutations in each human gamete.
- Small groups of individuals migrate out and found a new population.
- Separate populations merge together.
- Individuals do not mate at random.
- Natural selection acts on differences in fitness to alter allele frequencies.

Blue eyes

- Blue eyes first appeared between 6,000 and 10,000 years ago in a population living near the north shore of the Black Sea.

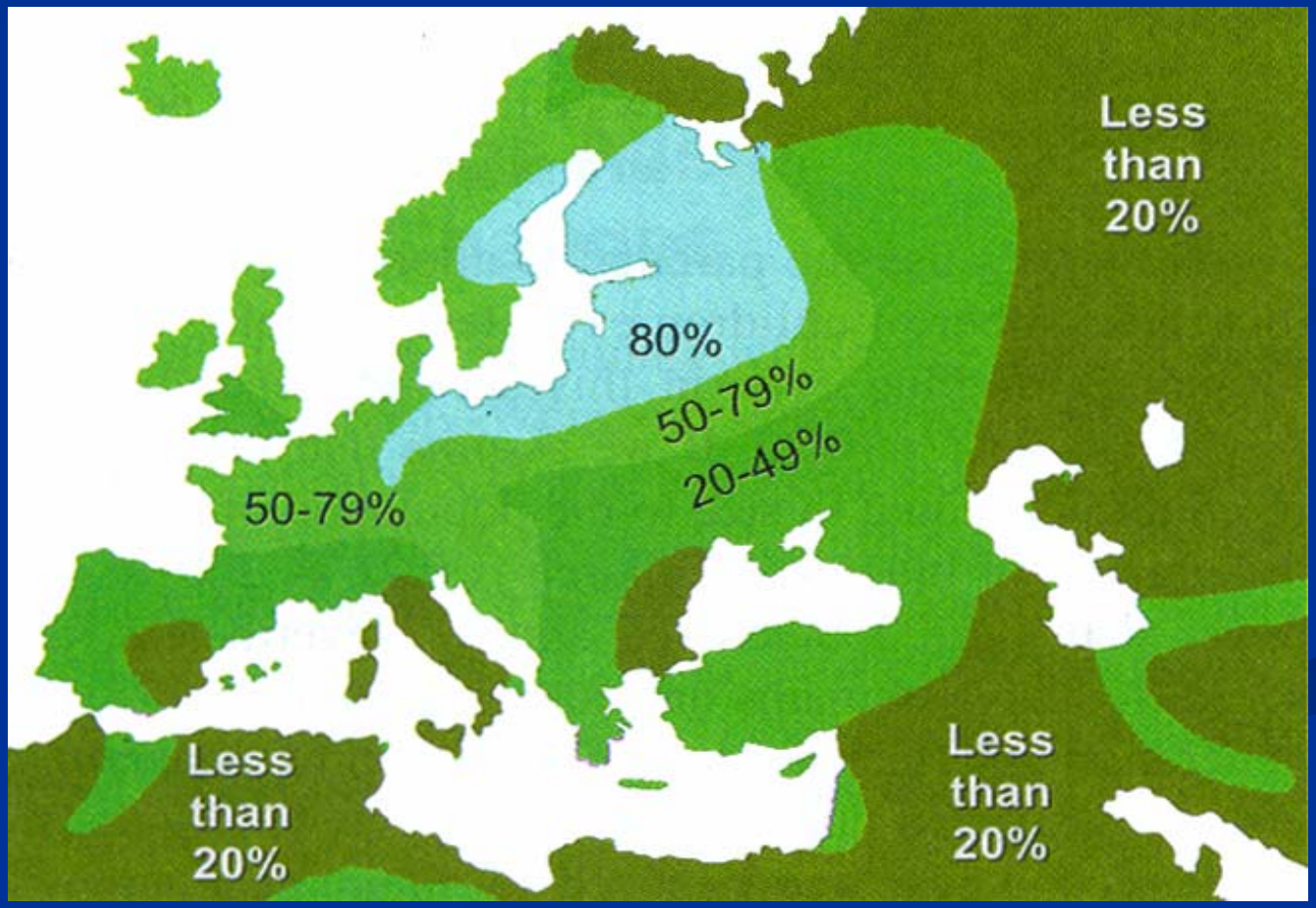
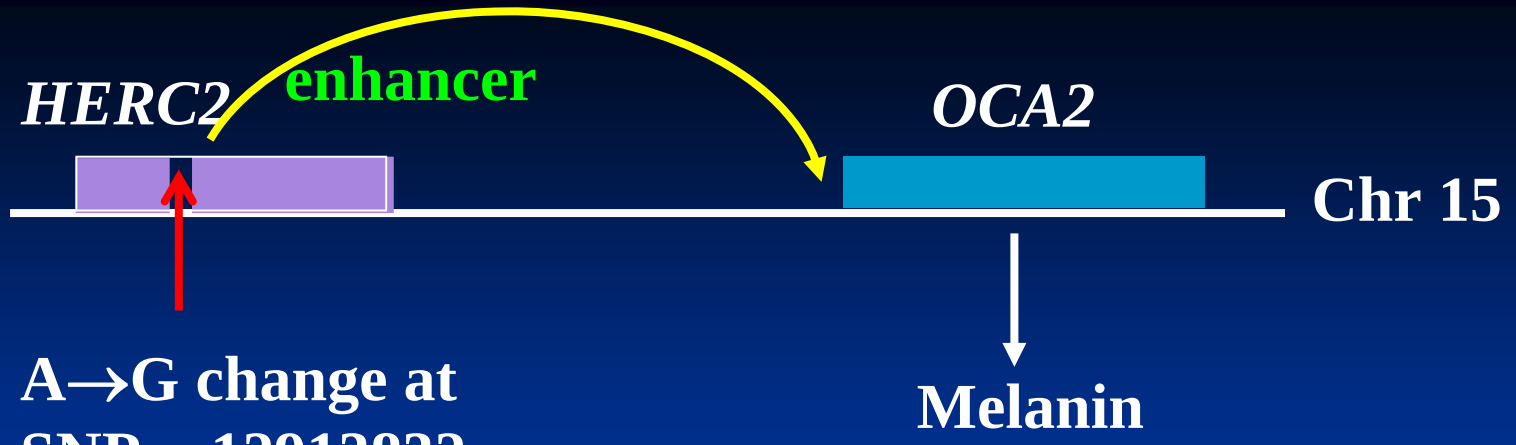
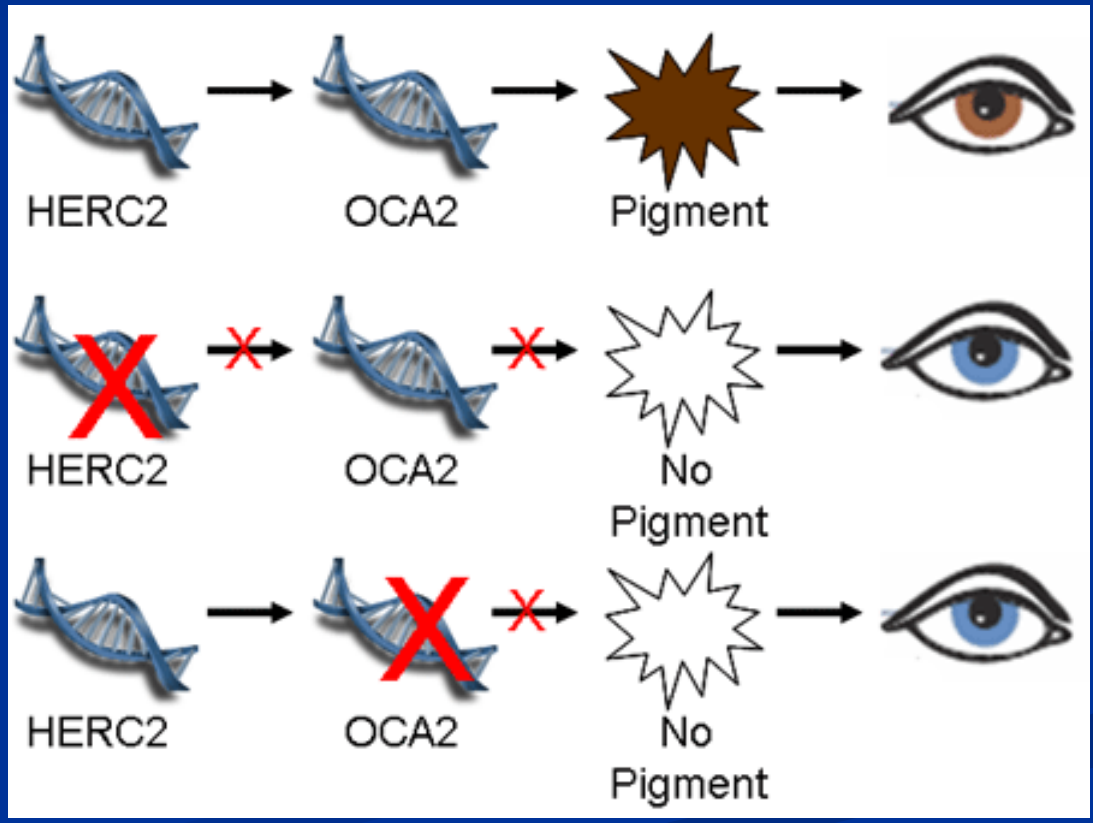


Fig. 20.7



- **Blue eyes are associated with a A to G base substitution at the SNP locus rs12913832 located in an intron of the *HERC2* gene on chromosome 15.**
- **rs12913832 is inside a highly conserved enhancer of the *OCA2* gene, which plays a critical role in the biosynthesis of the melanin.**



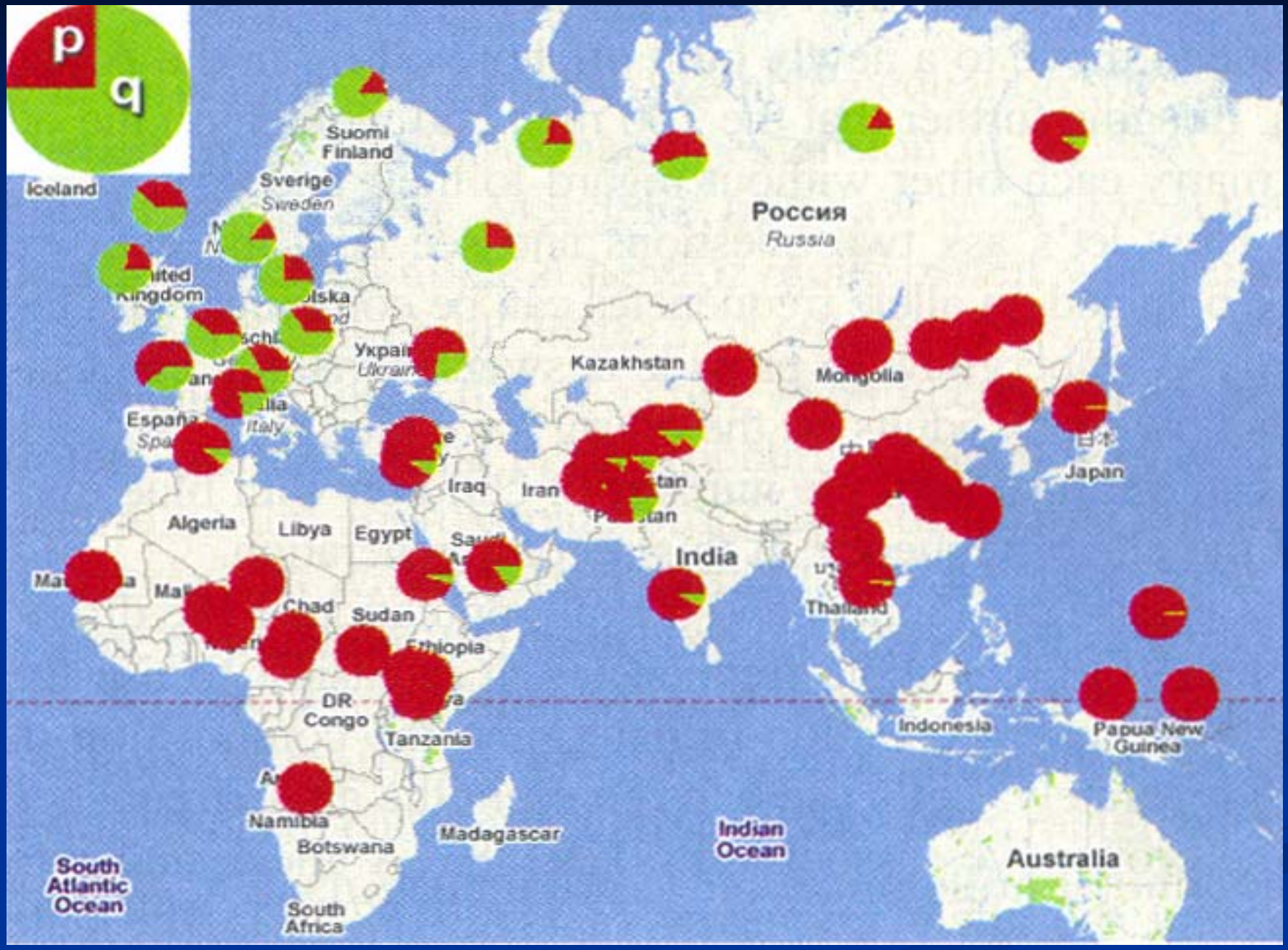

p
q
Iceland
Suomi
Finland
Sverige
Sweden
Россия
Russia
Kazakhstan
Mongolia
Japan
Iraq
Iran
India
Algeria
Libya
Egypt
Chad
Sudan
Ethiopia
DR
Congo
Tanzania
Namibia
Botswana
Madagascar
South
Africa
South
Atlantic
Ocean
Indian
Ocean
Indonesia
Papua New
Guinea
Australia
Thailand

Fig. 20.7

The wide-ranging variation in a population is the result of a balance between three forces

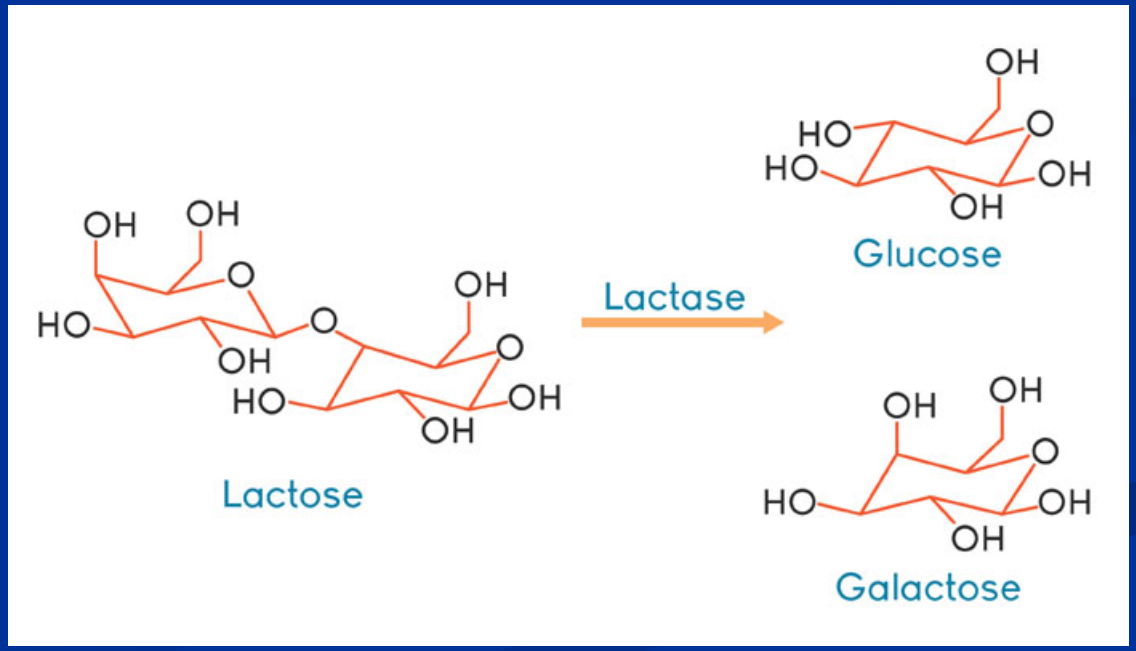
1. Continuous introduction of new mutations.
2. Loss of deleterious mutations because of the selective disadvantage they impose on individual carrying them.
3. Increase in frequency of rare mutations that either provide a selective advantage to the individuals carrying them or that spread through a population by other means.

p.215 (Chapter 7)

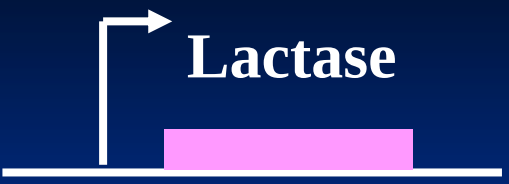
- **Evolutionary equilibrium:** A balance between mutation to a new allele and selection against the allele.

Lactase persistence and milk drinking in humans

- 65% of the human population today has **lactose intolerance**: drinking milk makes them ill, with symptoms including cramps (腹痛) and bloating (腹胀) .
- **Lactase (乳糖酶)** is required for breaking down lactose in milk.



- Humans express lactase only in infants but not in adults.
- **Lactase persistence** allows adults to drink milk.
  - 35% of the humans carry a mutation that caused constitutive expression of lactase (**lactase persistence**).



Lactase persistence

- **Lactase persistence** allows adults to drink milk.
  - The mutation occurred between 7,000-9,000 years ago among several dairying communities in places like northern Europe and eastern Africa.

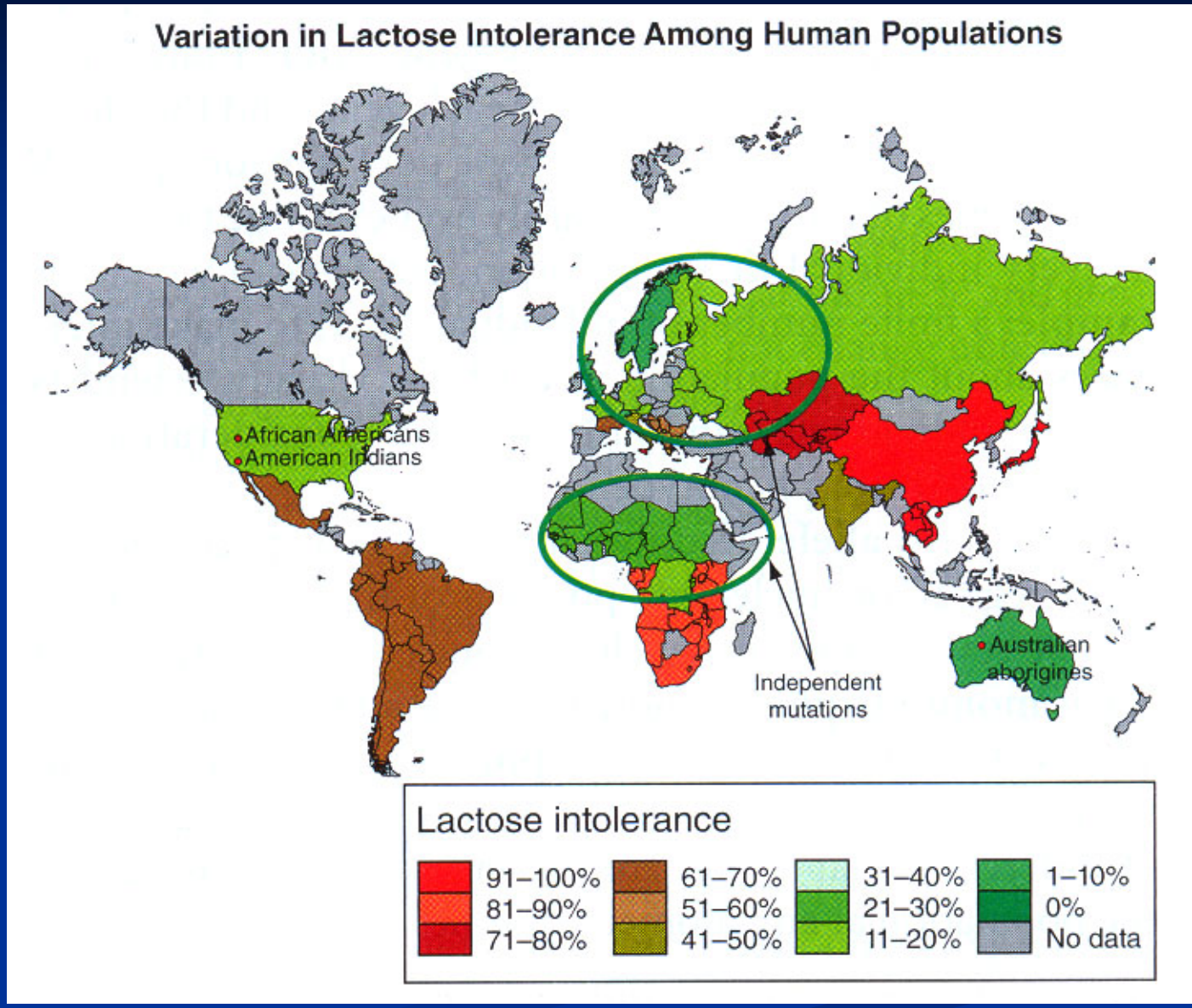
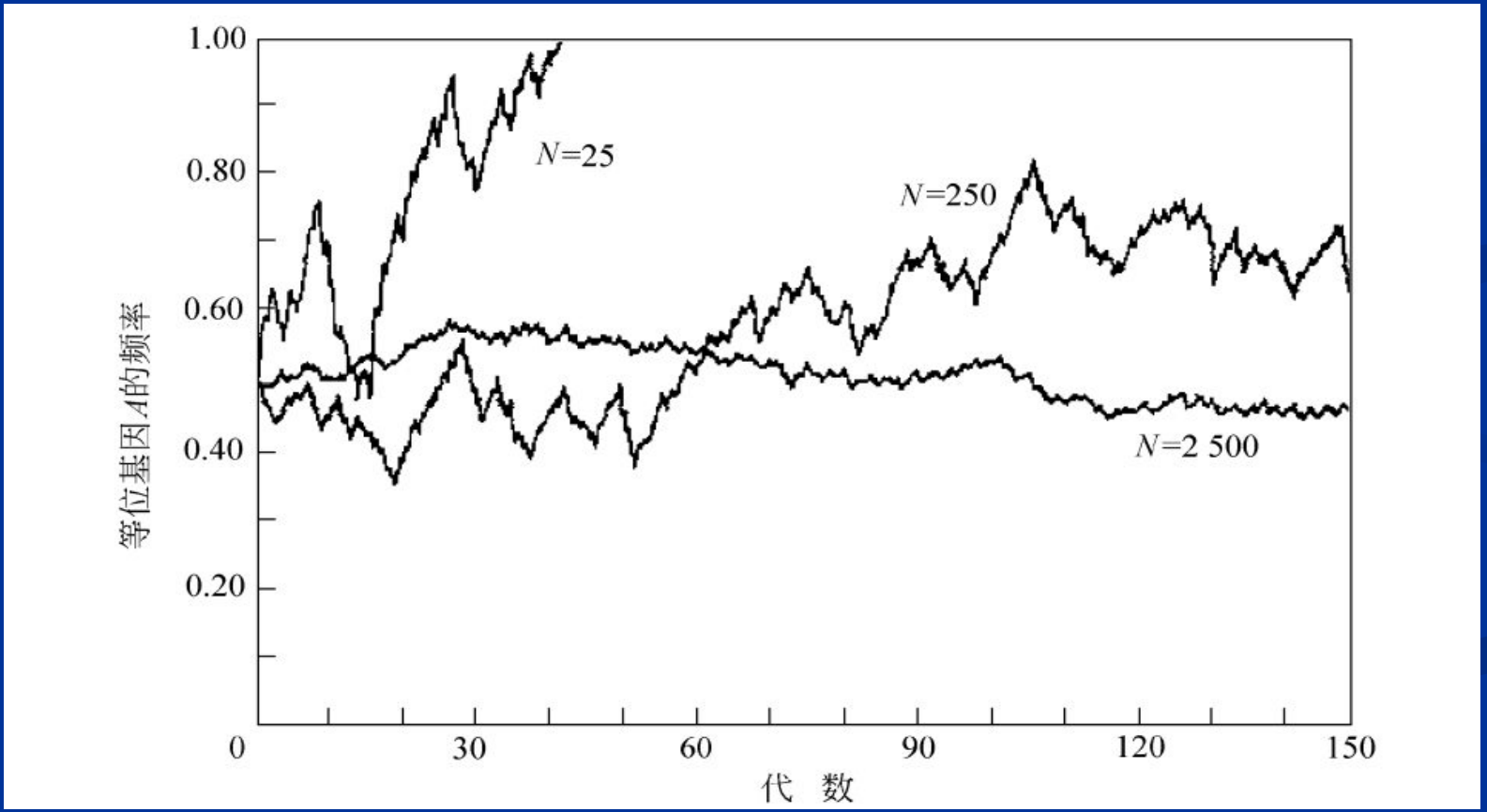


Fig. 20.1

Genetic drift

Genetic drift (遗传漂变): A change in allele frequencies as a consequence of the randomness of inheritance from one generation to the next.

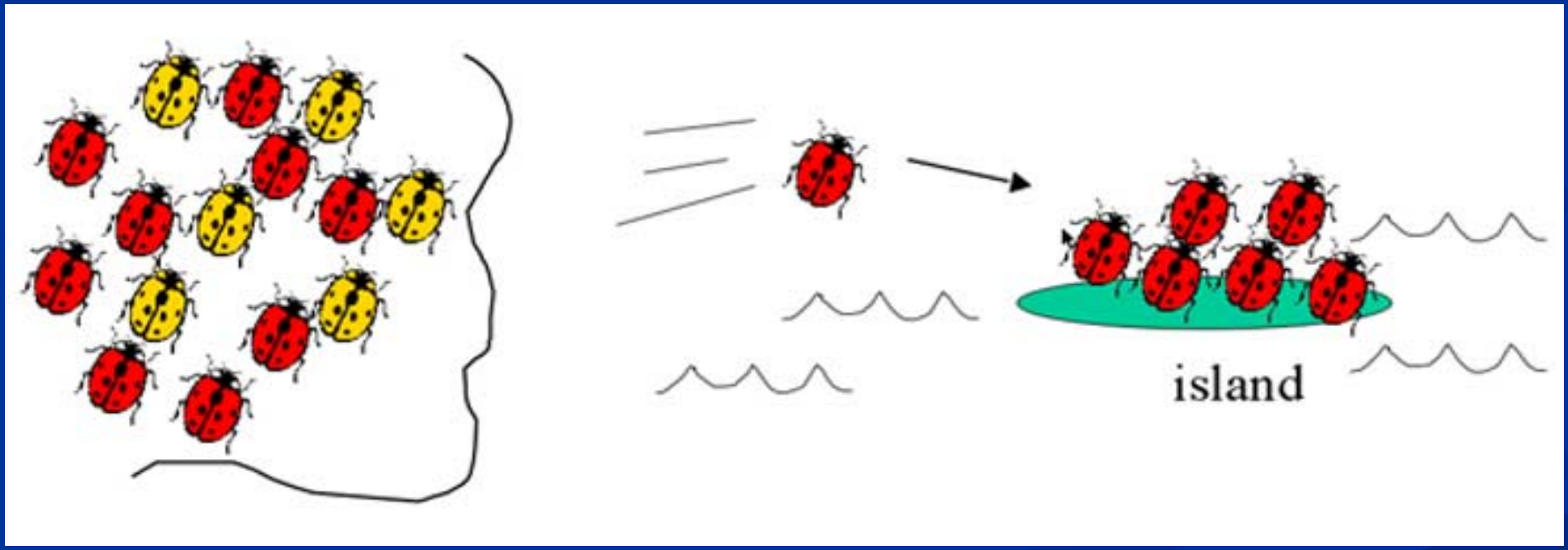
- **More dramatic in small populations.**



Genetic drift:

- Founder effect
- Bottleneck effect

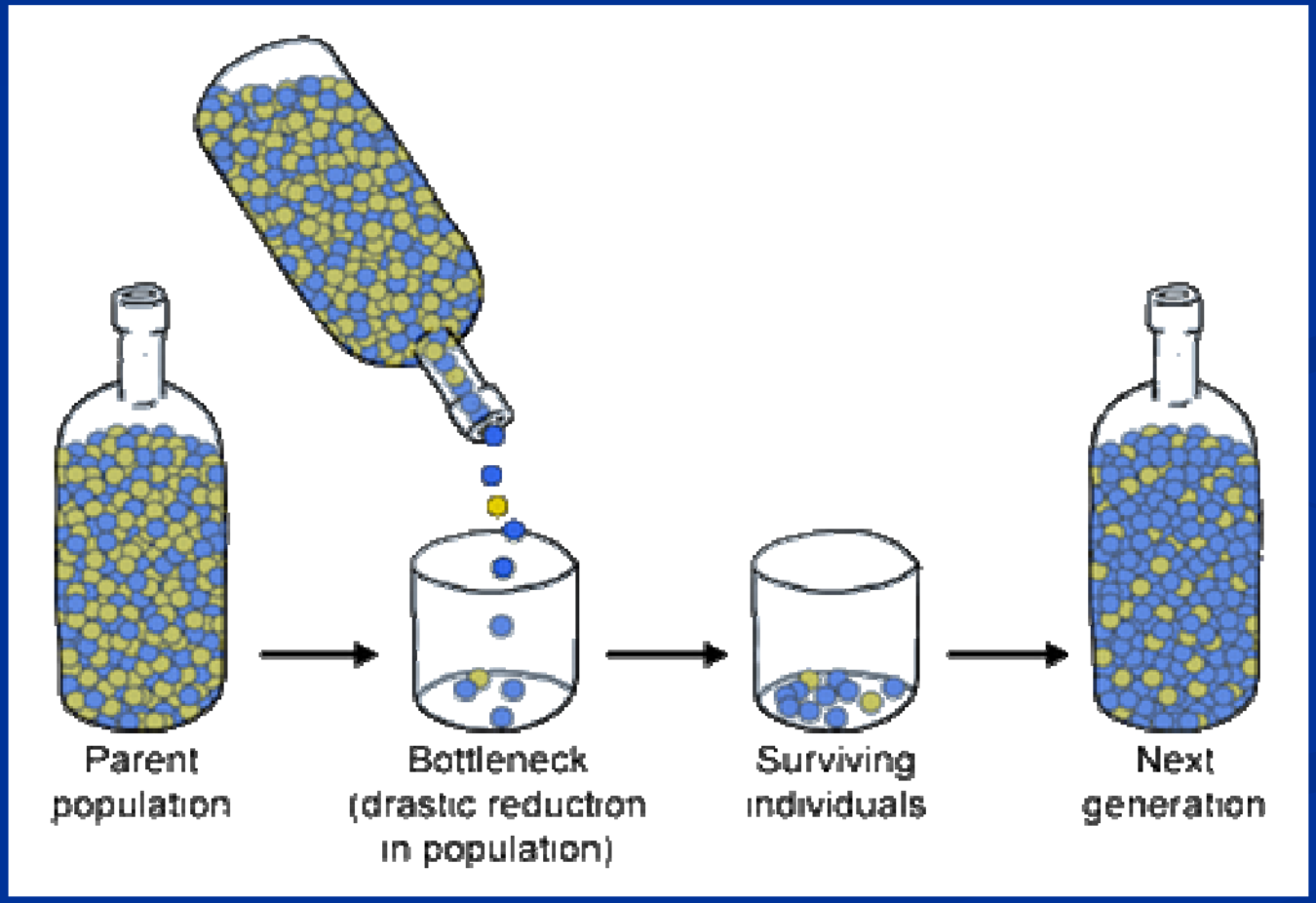
- **Founder effect (建立者效应): It occurs when a few individuals separate from a large population and establish a new one, causing altered allele frequency in the new population.**
  - **Higher incidence of manic-depressive illness in ~14,000 Amish people in eastern Pennsylvania than their European relatives.**



Founder Populations

Population	Number of Founders	Number of Generations	Population Size Today
Costa Rica	4,000	12	2,500,000
Finland	500	80–100	5,000,000
Hutterites	80	14	36,000
Japan	1,000	80–100	120,000,000
Iceland	25,000	40	300,000
Newfoundland	25,000	16	500,000
Quebec	2,500	12–16	6,000,000
Sardinia	500	400	1,660,000

- **Bottleneck effect (瓶颈效应)**: Changes of allele frequencies in a large population due to drastic reduction of population size, often as a consequence of environmental disturbances.
  - It reduces genetic diversity.



Natural selection acts on differences in fitness to alter allele frequencies

- **Fitness**: The relative advantage or disadvantage in reproduction that a particular genotype provides to members of a population in comparison to alternative genotypes at the same locus.
  - Fitness has two basic components: **viability** and **reproductive success**.
- **Fitness cost**: The effect of deleterious allele it is selected against.
- **Natural selection**: In nature, the process that progressively eliminates individuals whose fitness is low and choose individuals of high fitness to survive and reproduce.

Pocket mice from New Mexico living on sandy soils and very black volcanic rock have diverged by natural selection to match their substrate.

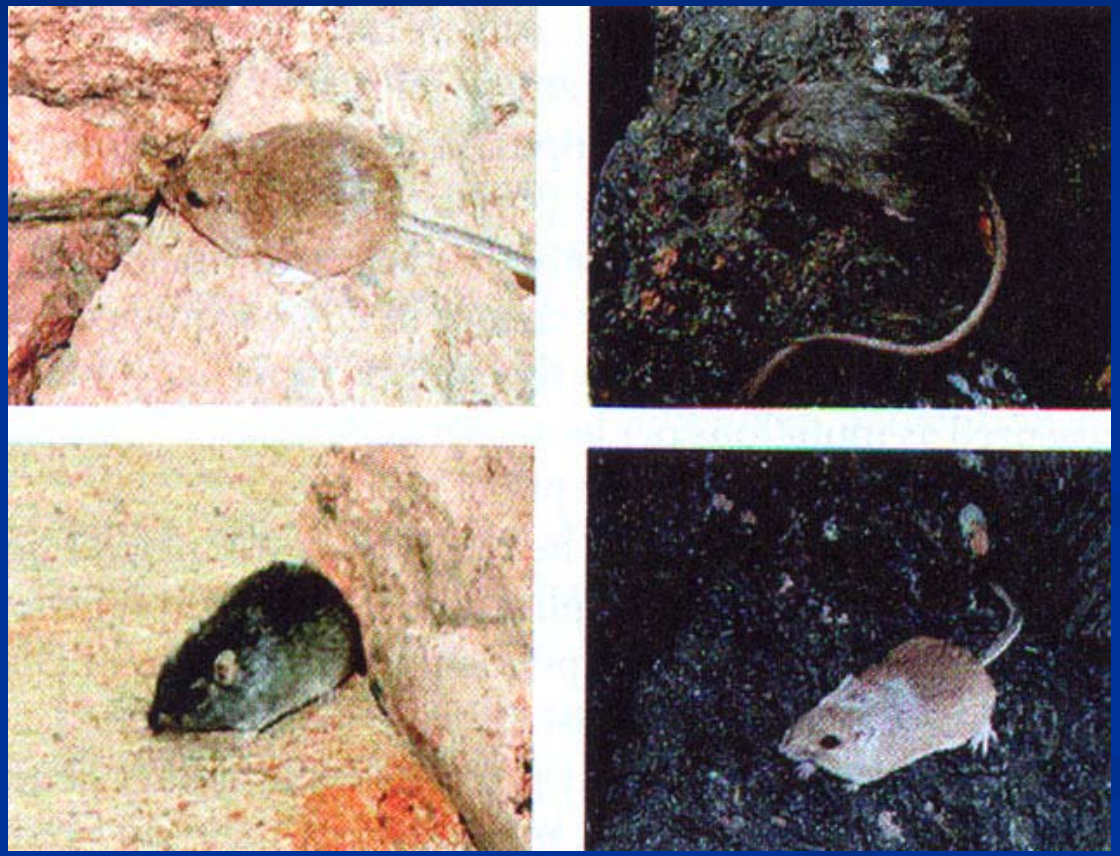
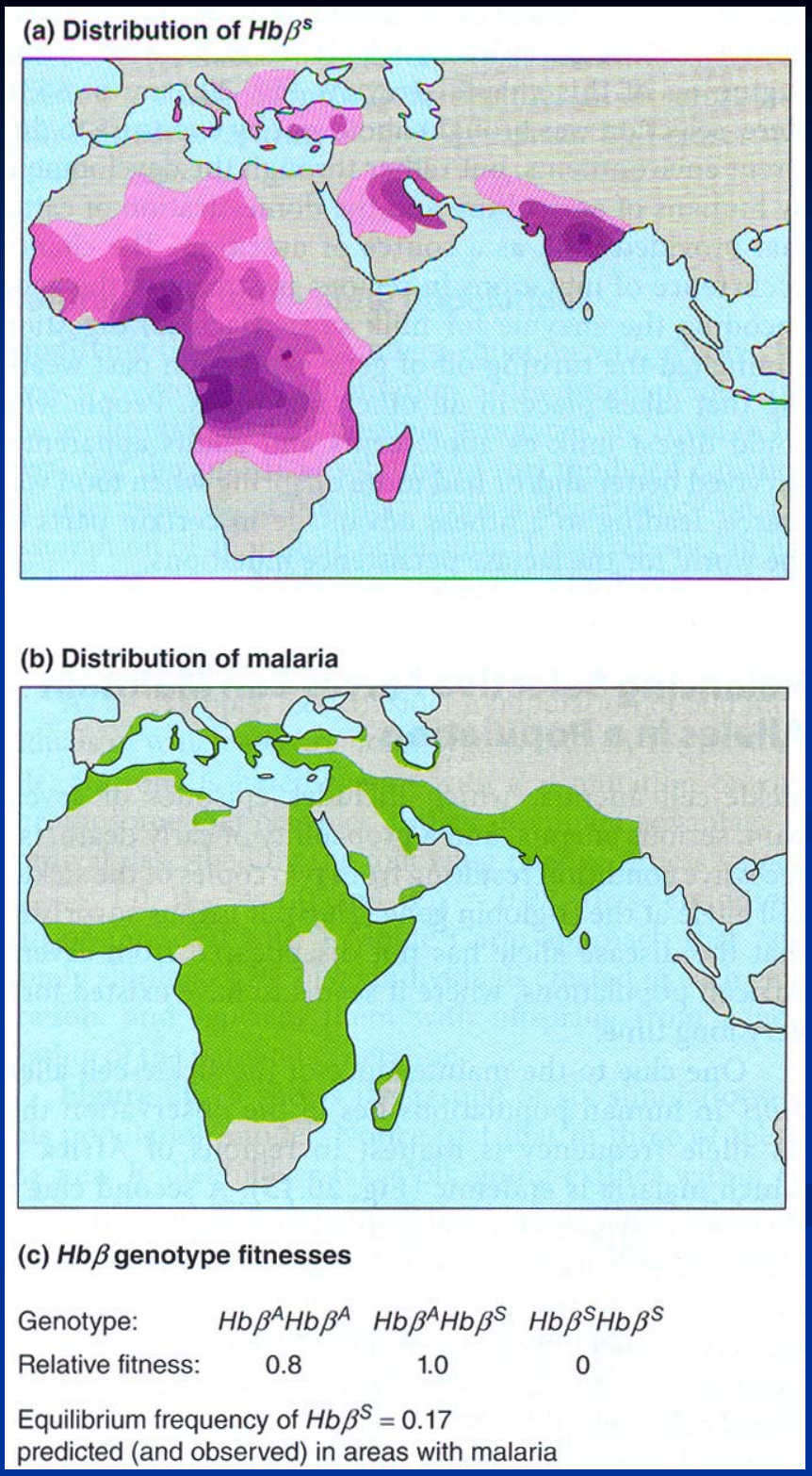


Fig. 20.10

Balancing selective forces can maintain alleles in a population



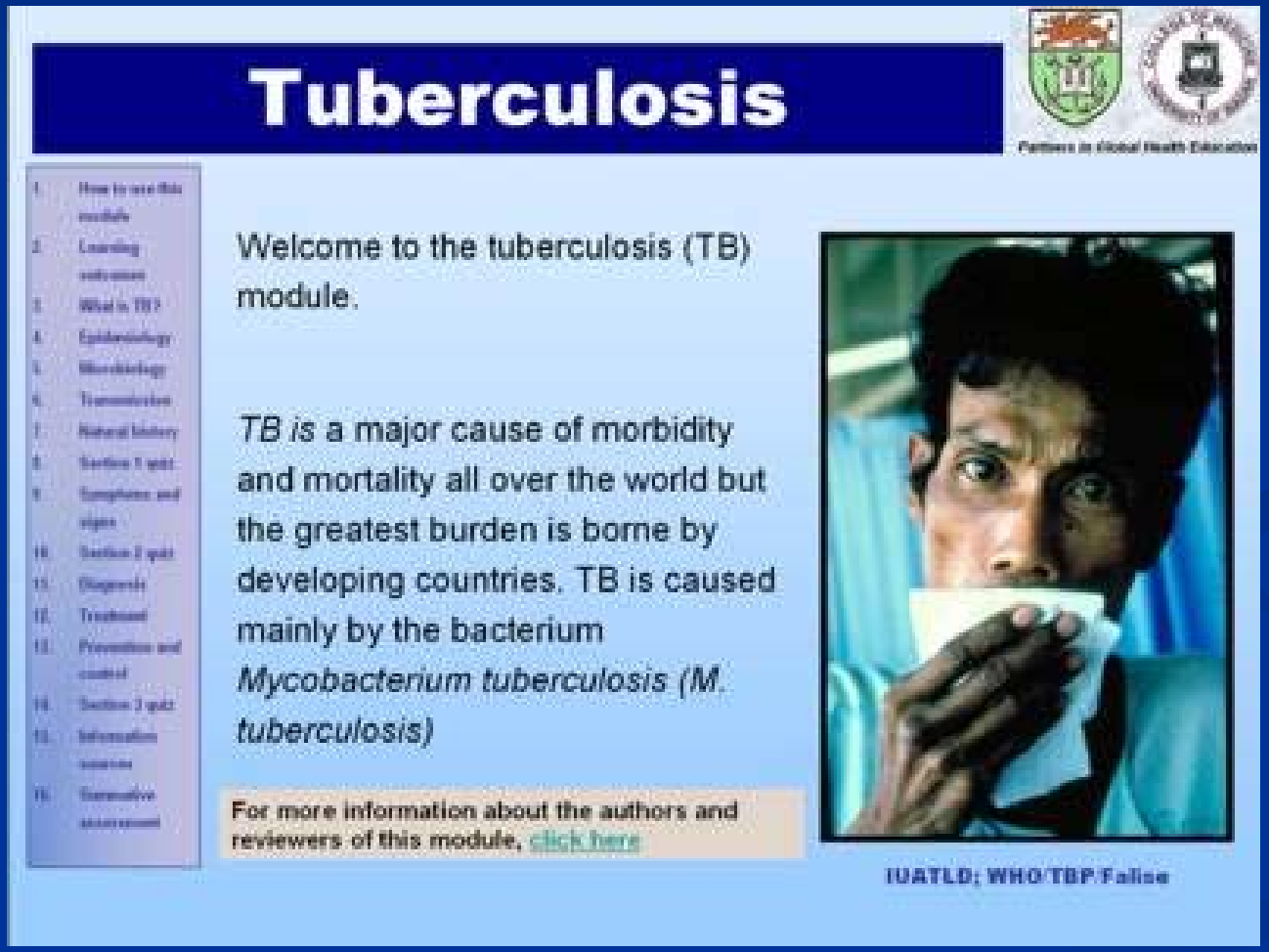
High frequency of the sickle-cell allele $Hb\beta^S$ in regions of Africa where malaria is prevalent.

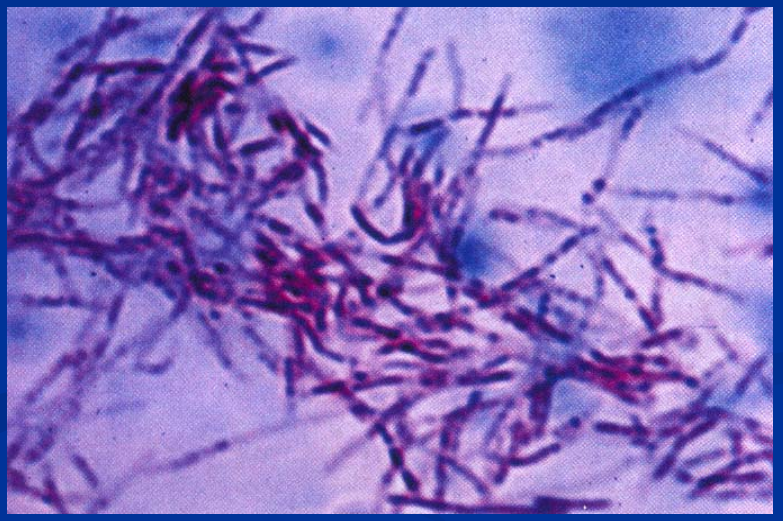
Fig. 20.15

Human behavior can affect evolution of pathogens and pests

Example 1: The evolution of drug resistance in pathogens

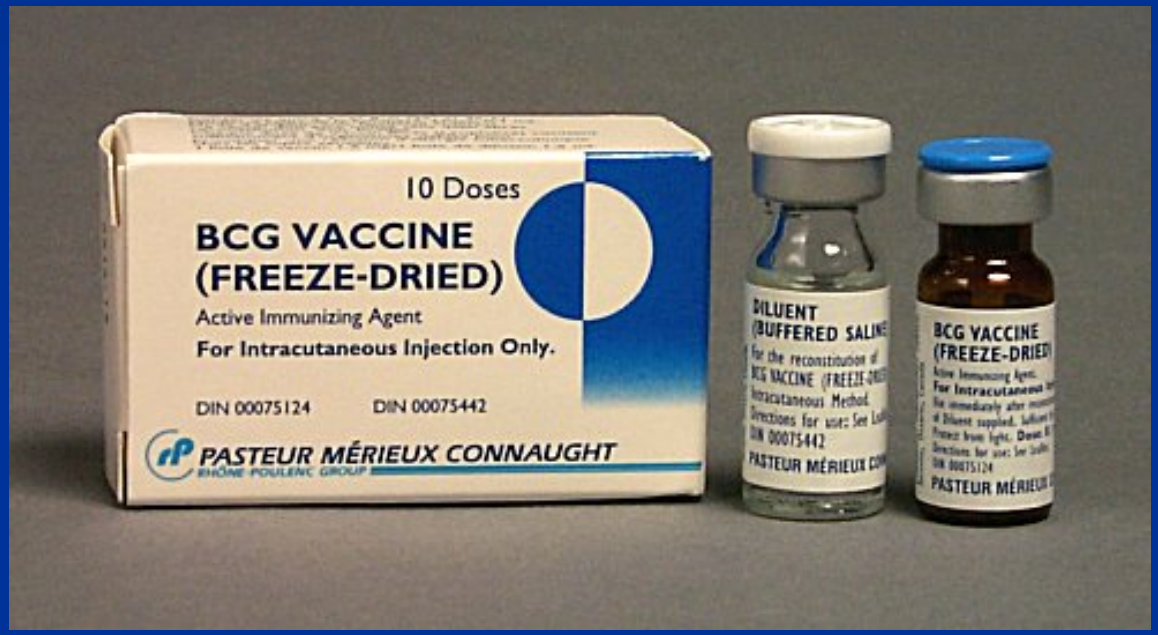
Tuberculosis (TB, 肺结核病)





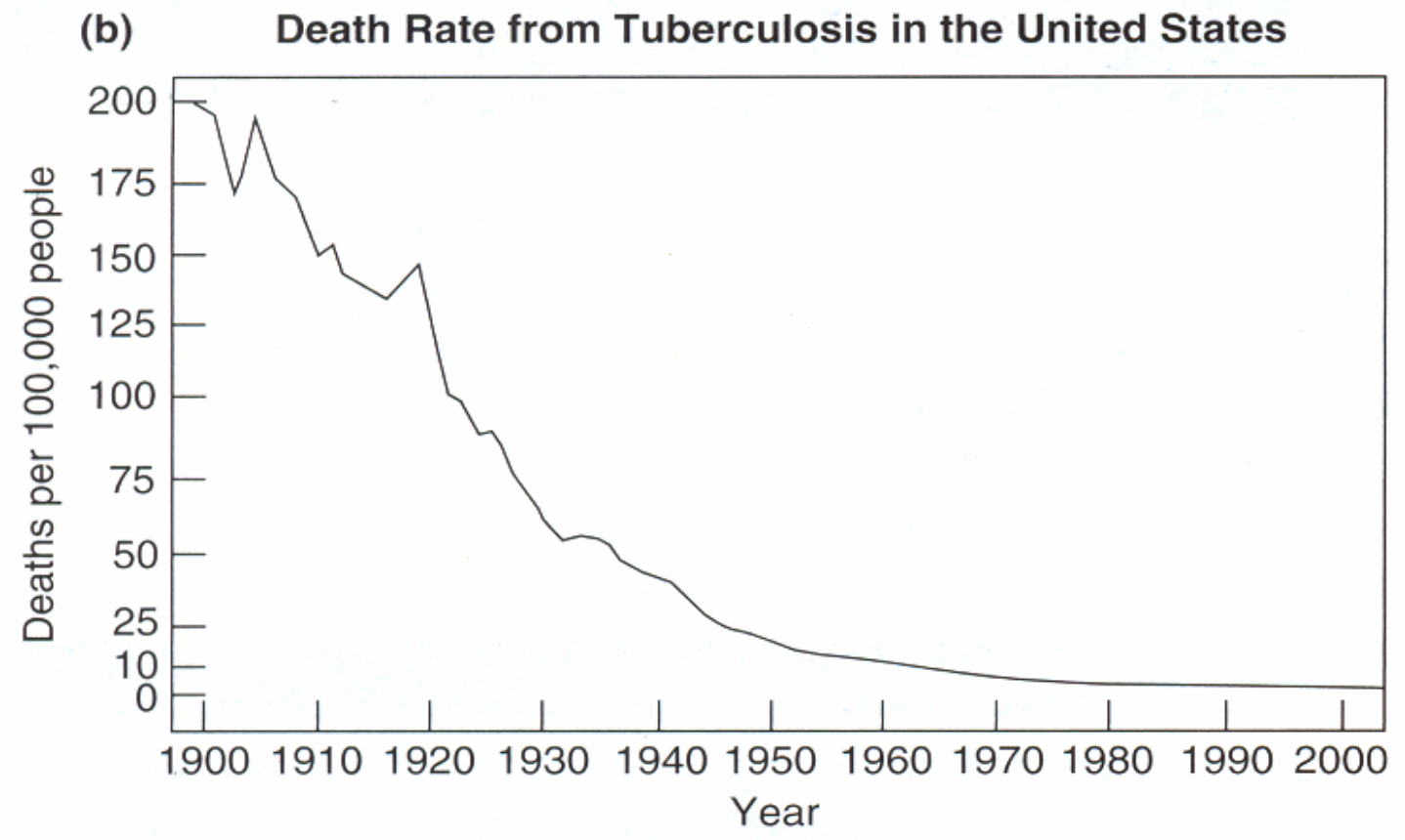
Mycobacterium tuberculosis **colonizing the lungs and bones**

BCG vaccine (Bacille Calmette-Guérin,卡介苗): Used since 1921.



Drug (year of discovery)	MIC μg/ml	
Isoniazid (1952)	0.02–0.2	(异烟肼)
Rifampicin (1966)	0.05–1	(利福平)
Pyrazinamide (1952)	16–50 (pH 5.5)	
Ethambutol (1961)	1–5	
Streptomycin (1944)	2–8	
Amikacin/kanamycin (1957)	2–4	
Capreomycin (1960)		
Quinolones (1963)	0.5–2.5	
Ethionamide (1956)	2.5–10	
PAS (1946)	1–8	

MIC = minimum inhibitory concentrati


(b) Death Rate from Tuberculosis in the United States
Deaths per 100,000 people
200
175
150
125
100
75
50
25
10
0
1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
Year

Drug-resistant TB in populations from New York City - 1991

- 23% of strains isolated from previously treated patients resistant to one or more antibiotics.
- 7% resistant to both **isoniazid**（异烟肼）and **rifampicin (**利福平**)**.
- 44% isolates from relapsed patients resistant to one or more drugs.
- 30% isolates resistant to isoniazid and rifampicin.

- How did drug-resistant TB arise?
- Why does it persist?
- Why does it resurge in frequency even though it has been long under control?

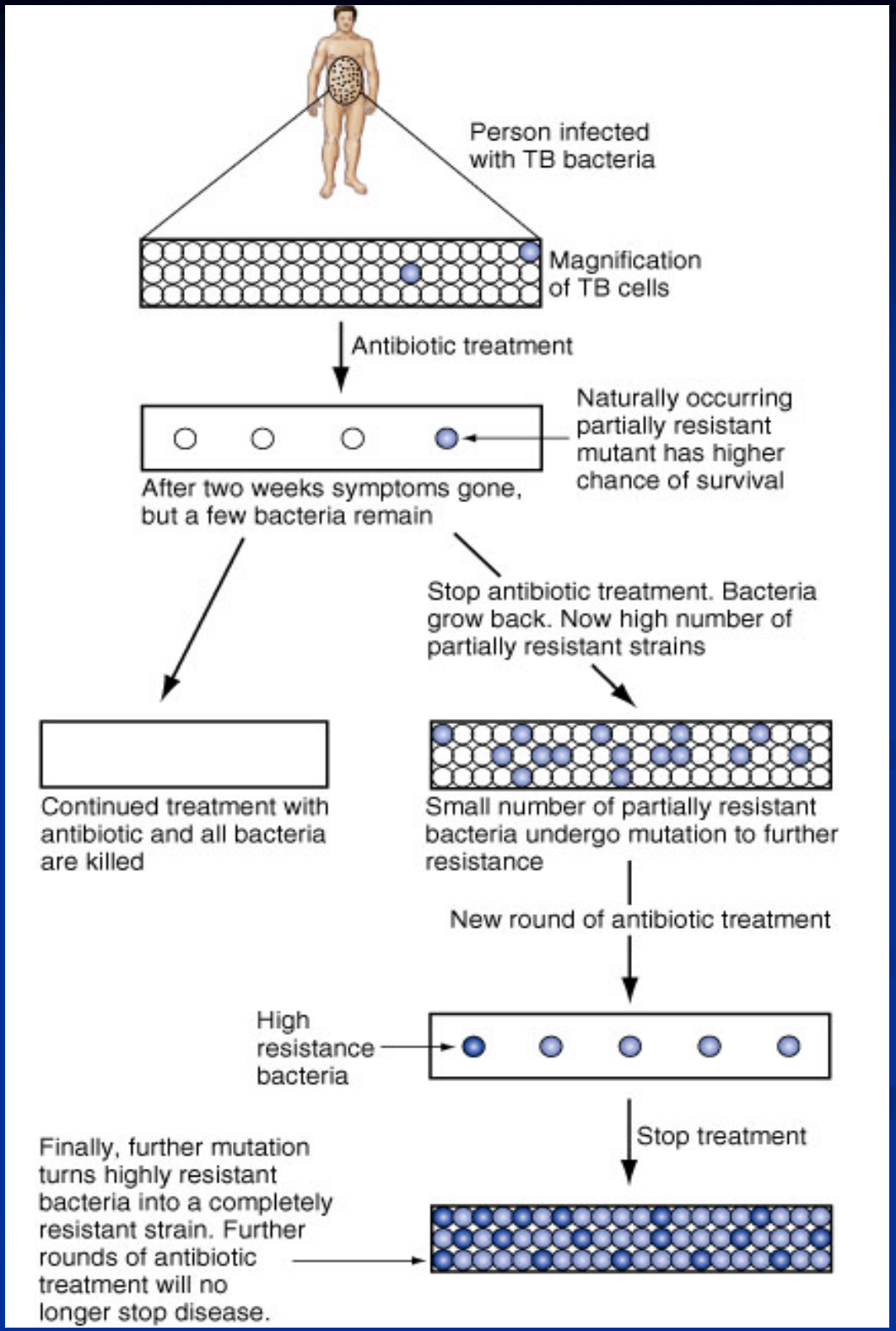
- **How did drug-resistant TB arise?**
  - **By mutation and selection.**
- **Why does it persist?**
  - **Changes in allele frequency reach an evolutionary equilibrium.**
- **Why does it resurge in frequency even though it has been long under control?**
  - **Human behavior tips the balance in the interaction of microbes with hosts.**

Table Mechanisms of drug resistance in *M. tuberculosis*

Drug (year of discovery)	MIC µg/ml	Gene(s) involved in resistance	Gene function	Role	Mechanism of action	Mutation frequency %
Isoniazid (1952)	0.02–0.2	*kat*G *inh*A	Catalase-peroxidase Enoyl ACP reductase	Pro-drug conversion Drug target	Inhibition of mycolic acid biosynthesis and other multiple effects	50–95 8–43
Rifampicin (1966)	0.05–1	*rpo*B	β subunit of RNA polymerase	Drug target	Inhibition of RNA synthesis	95
Pyrazinamide (1952)	16–50 (pH 5.5)	*pnc*A	Nicotinamidase/pyrazinamidase	Pro-drug conversion	Depletion of membrane energy	72–97
Ethambutol (1961)	1–5	*emb*B	Arabinosyl transferase	Drug target	Inhibition of arabinogalactan synthesis	47–65
Streptomycin (1944)	2–8	*rps*L *rrs* *gid*B	S12 ribosomal protein 16S rRNA rRNA methyltransferase (G527 in 530 loop)	Drug target Drug target Drug target	Inhibition of protein synthesis	52–59 8–21 ?
Amikacin/kanamycin (1957)	2–4	*rrs*	16S rRNA 16S rRNA	Drug target	Inhibition of protein synthesis	76
Capreomycin (1960)		*tly*A	2′-O-methyltransferase			
Quinolones (1963)	0.5–2.5	*gyr*A *gyr*B	DNA gyrase subunit A DNA gyrase subunit B	Drug target	Inhibition of DNA gyrase	75–94
Ethionamide (1956)	2.5–10	*eta*A/*eth*A *inh*A	Flavin monooxygenase	Prodrug conversion Drug target	Inhibition of mycolic acid synthesis	37 56
PAS (1946)	1–8	*thy*A	Thymidylate synthase	Drug activation?	Inhibition of folic acid and iron metabolism?	36

MIC = minimum inhibitory concentration; ACP = acyl carrier protein; PAS = para-aminosalicylic acid.

Patient noncompliance with drug treatments is a major factor in evolution of antibiotic resistance

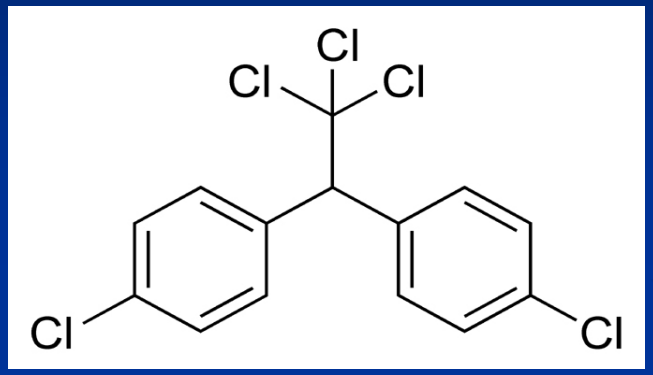


Factors contributing to rapid evolution of resistance in bacterial pathogens

- Short generation times and rapid rate of reproduction
  - Evolution proceeds quickly relative to human life span.
- Large population densities
  - Ensure resistance mutations will appear in population
- Strong selection imposed by antibiotics
  - Increases rate of evolution in each generation unless bacterial population is entirely eliminated
- Variety of ways bacteria acquire genes speeds evolution
  - Plasmids which carry resistance genes multiply rapidly.
  - Plasmids are readily exchanged among bacterial populations and among different species through transformation, conjugation, and transduction.

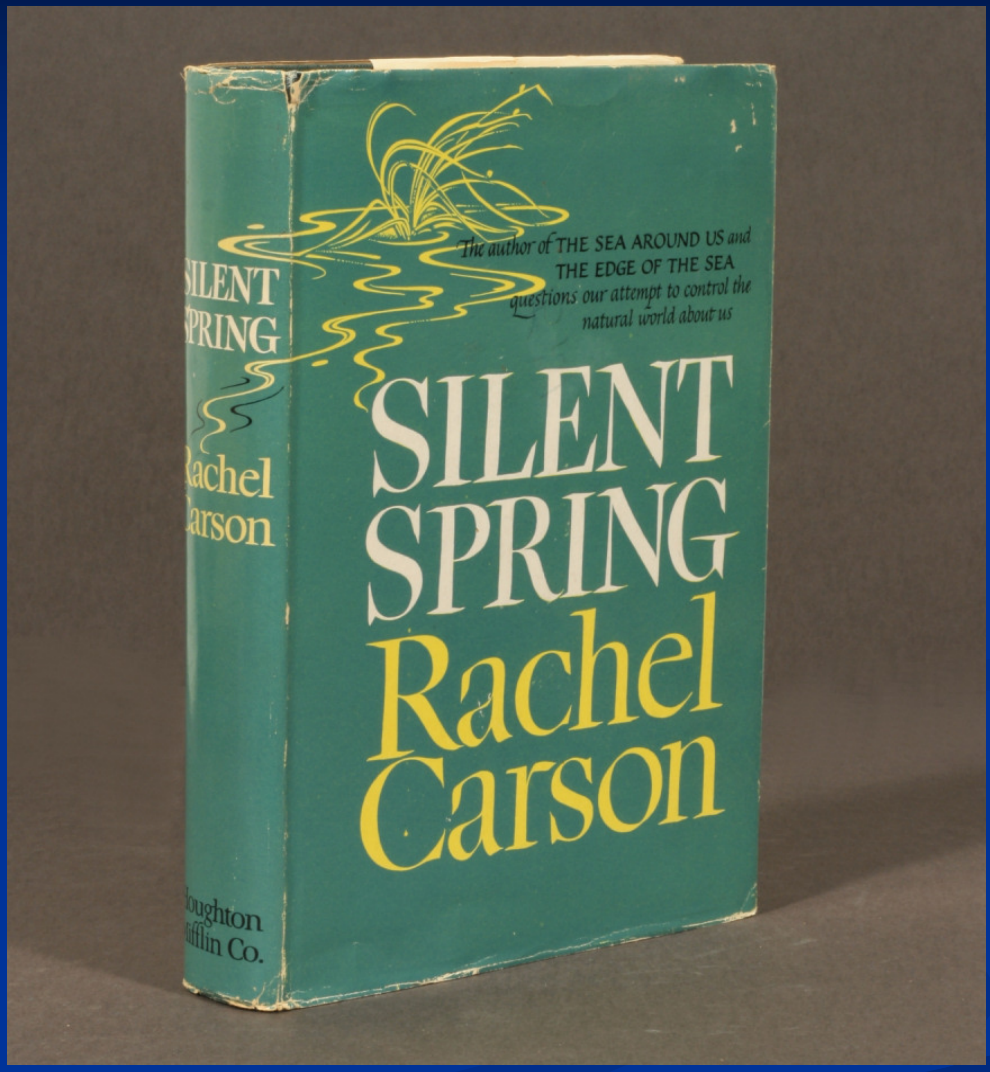
Example 2: Evolution of pesticide resistance

- DDT and other chemicals used on large scale in 1940s highly successful.



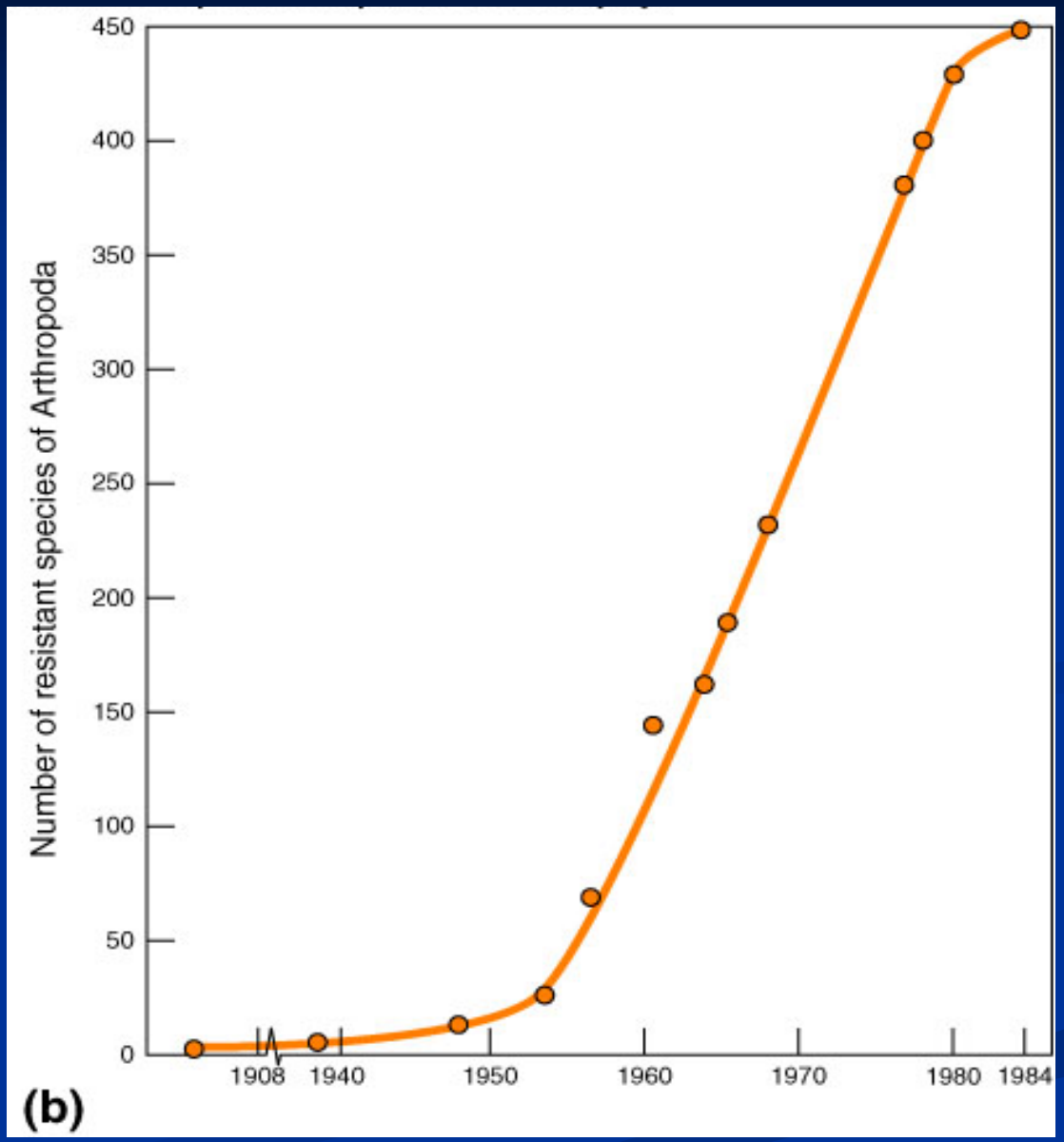
Soft egg shell

Rachel Carson



The first edition in 1962

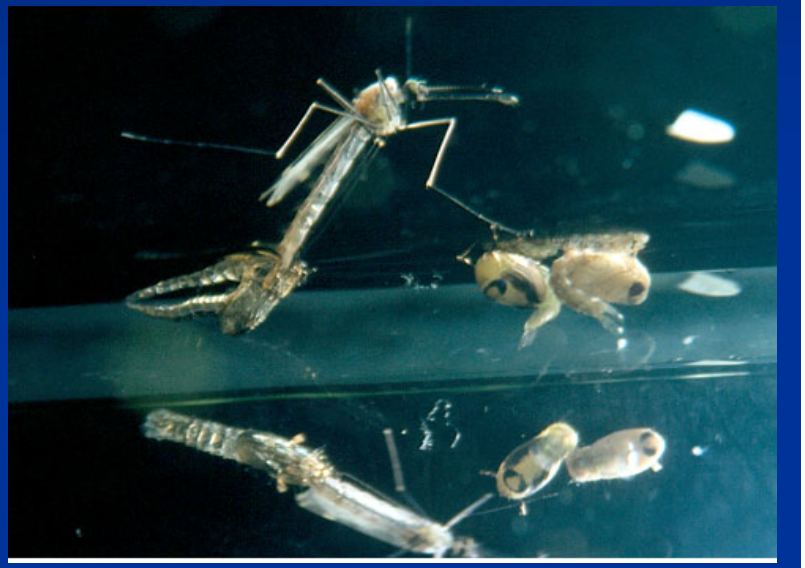
- Since 1950s, resistance to every known insecticide has evolved within 10 years.
- 1984 - >450 resistant species of insects and mites.

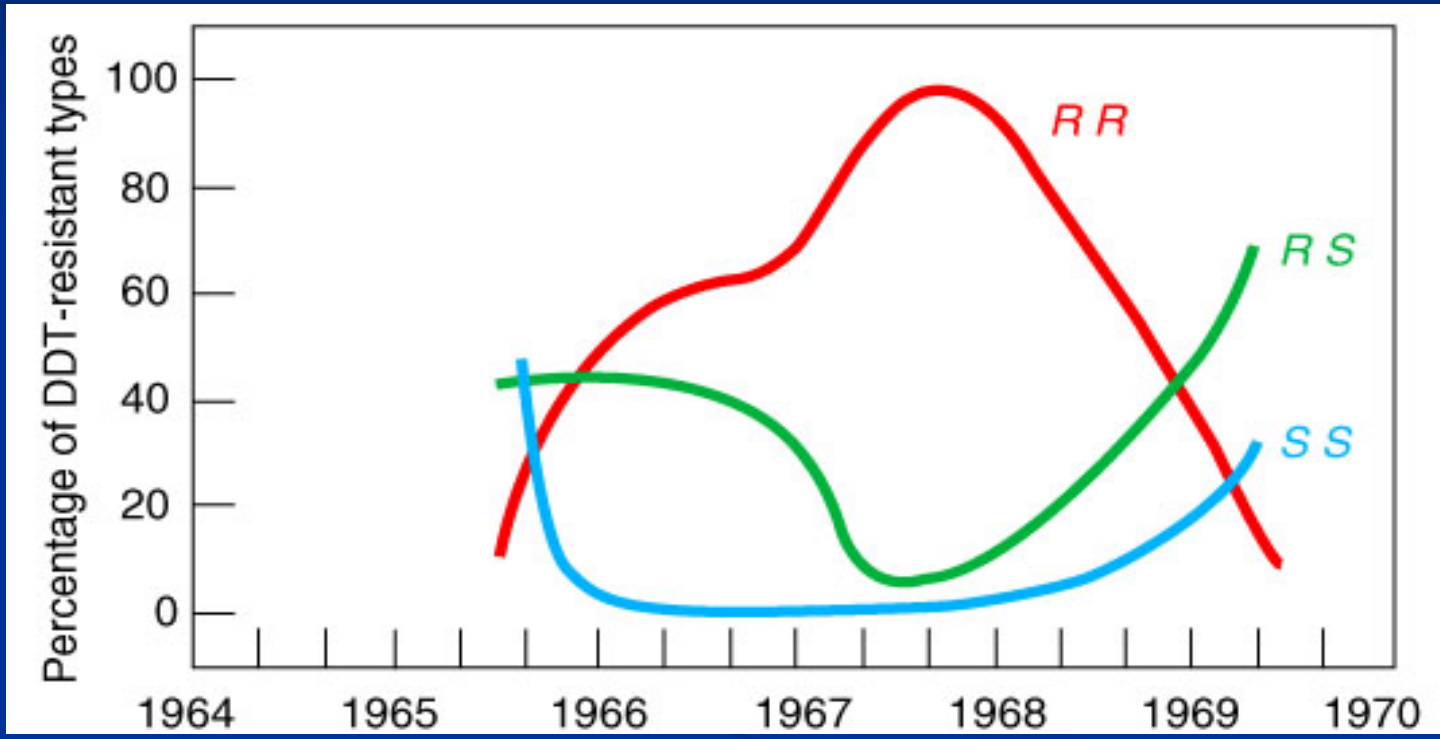


How does resistance evolve?

- Often evolves from mutations in a single, enzyme-encoding gene.
- DDT is a nerve toxin.
  - Resistance develops from dominant mutations in a single gene that detoxifies DDT.
  - Dominant alleles can experience strong selection because of heterozygous advantage, even at low frequency in a population.

Genotype frequencies among populations of mosquito larvae change in response to insecticide





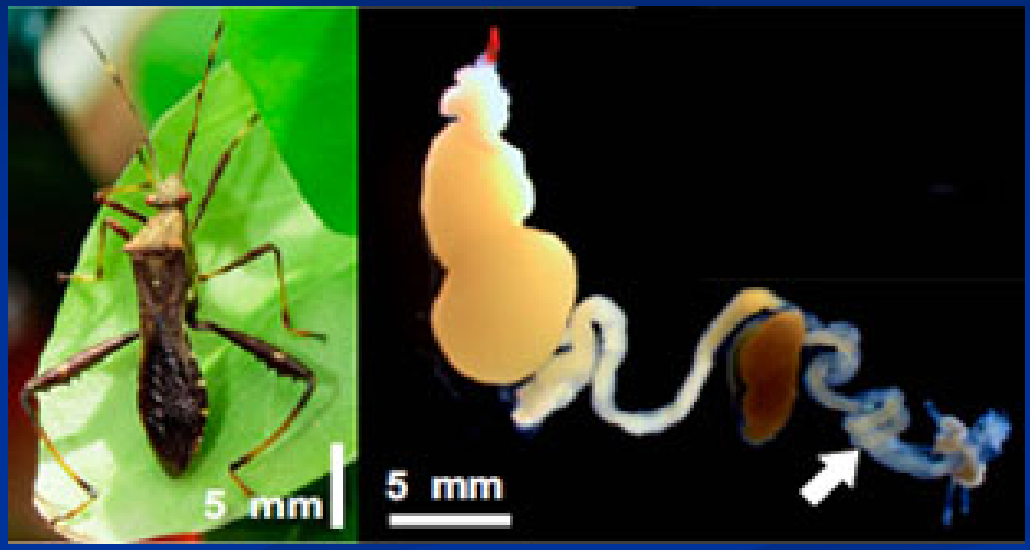
Changing proportions of resistance genotypes of *A. aegypti* (larvae) under selection with DDT (1964-1967), and after selection was relaxed (1968), in a suburb of Bangkok, Thailand.

Fig. 20.16

- **The homozygous resistance genotype imposes a fitness cost on individuals such that in the absence of insecticide, resistance is subject to a negative selection that decreases the frequency of *R* allele in the population.**

The bean bug develops insecticide resistance by swallowing bacteria that breakdown it

- Fenitrothion has been widely used to kill bean bugs, the major pest of soybeans.
- As nymphs, bean bugs swallow *Burkholderia* bacteria from the surrounding soil.



A bean bug, and its digestive system, showing where the symbiotic bacteria live.

Kikuchi Y *et al*. (2012) Symbiont-mediated insecticide resistance. *PNAS* 109:8618-8622

Example 3: Industrial melanism (工业黑化)

Peppered moth *Biston betularia*

Lichen-covered tree in Dorset, England

Soot-covered tree near Birmingham, England