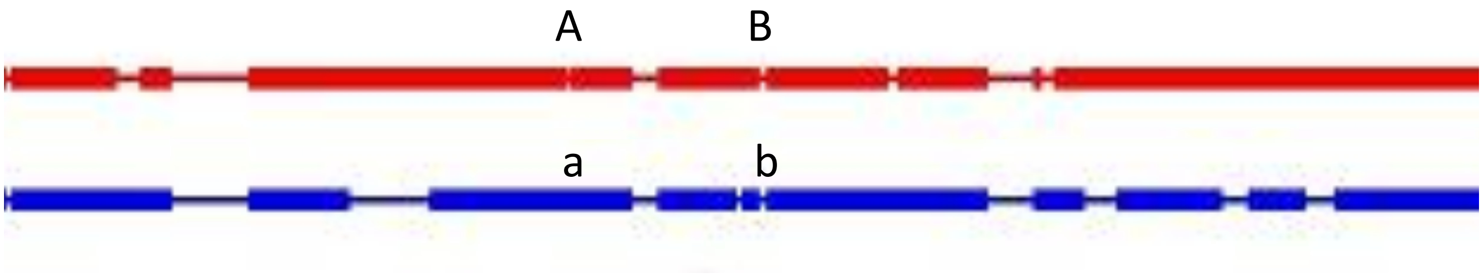
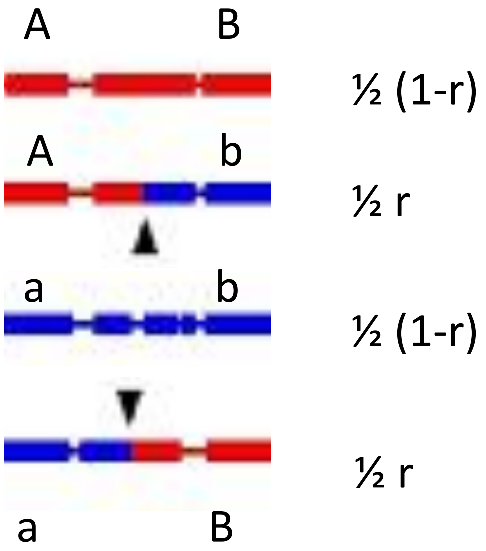
Recombination and Linkage Disequilibrium (LD)

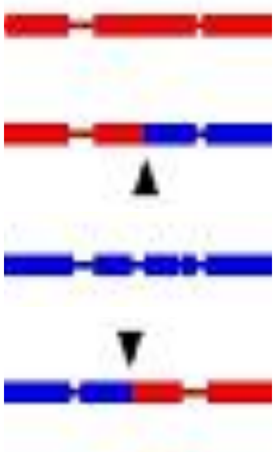


r = recombination fraction
probability of an odd
Number of crossovers occur
Between our markers

$0 < r < \frac{1}{2}$



Linkage disequilibrium: The non-random association of alleles at different sites in the genome in a population.



p_{AB} = frequency of AB

p_{ab} = frequency of ab

p_{Ab} = frequency of Ab

p_{aB} = frequency of aB

If independent the expected

frequency of gametes (haplotypes

$p_A \text{ X } p_B$

$p_a \text{ X } p_b$

$p_A \text{ X } p_b$

$p_a \text{ X } p_B$

Define “D”

$D_{AB} = p_{AB} - p_A p_B$

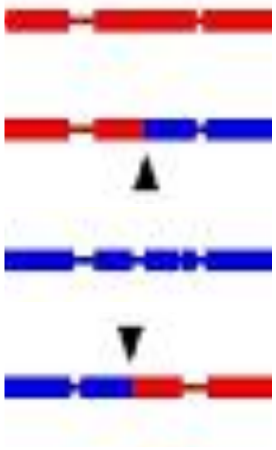
$D_{ab} = p_{ab} - p_a p_b$

$D_{Ab} = p_{Ab} - p_A p_b$

$D_{aB} = p_{aB} - p_a p_B$

The covariance of A and B.

Linkage disequilibrium: The non-random association of alleles at different sites in the genome.



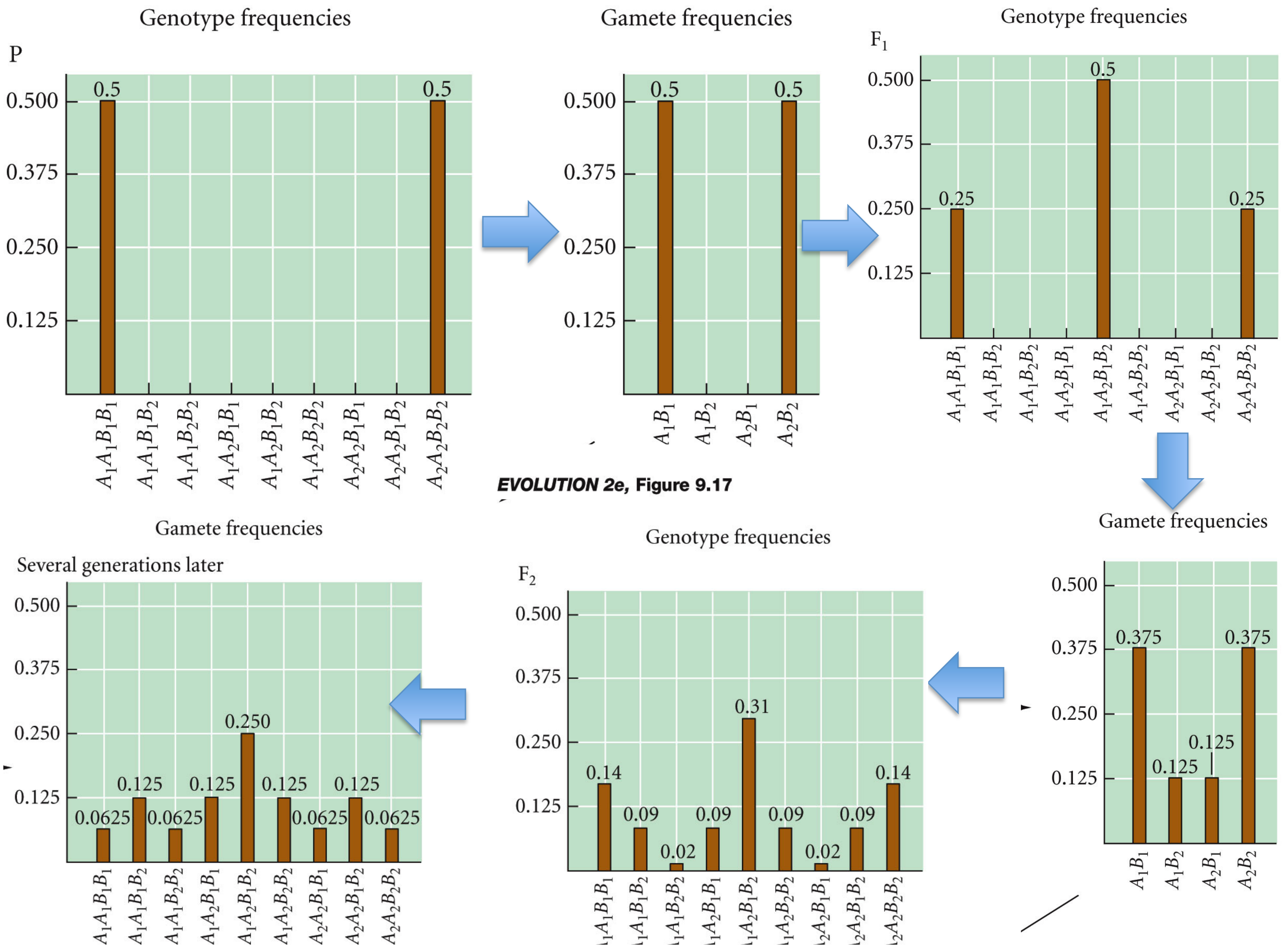
Define "D"

$D_{AB} = p_{AB} - p_A p_B$
$D_{ab} = p_{ab} - p_a p_b$
$D_{Ab} = p_{Ab} - p_A p_b$
$D_{aB} = p_{aB} - p_a p_B$

$D_{AB} = - D_{Ab}$
$D_{AB} = D_{ab}$ and $D_{Ab} = D_{aB}$

(so, knowing D_{AB} is enough - call this "D")
If O = E, then D = 0
If D >0 (or D<0) then there is "linkage disequilibrium (LD)"

Note: you can also write $p_{AB} = p_A p_B + D$



***EVOLUTION 2e*, Figure 9.17**

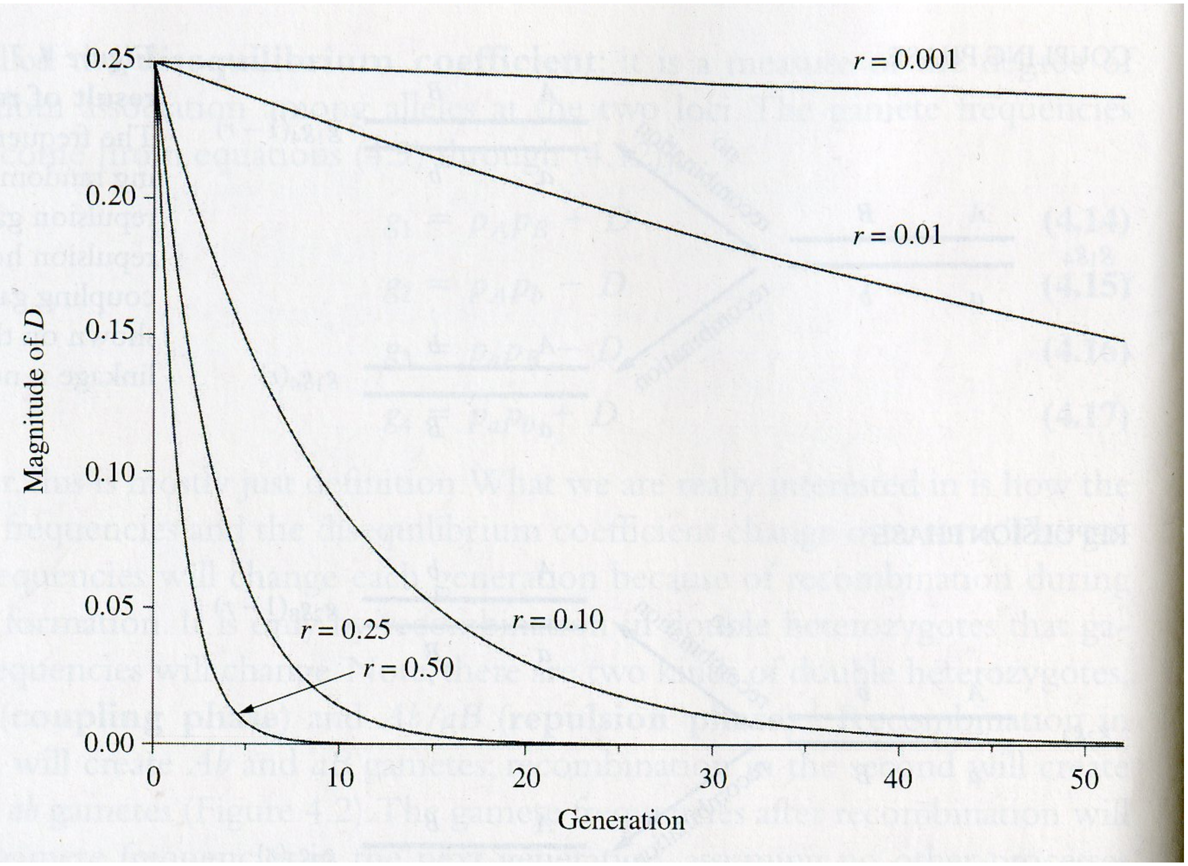
Decay of LD in a very large boring randomly mating population

$$D_t = (1- r)^t D_o$$

With inbreeding coefficient f replace r with r(1-f) linkage disequilibrium

How does LD change over time due to recombination?

$D_t = (1- r)^t D_o$



Note: more distant markers recombine more!

So eventually recombination leads to D=0.

Even with free recombination (r=0.5), it isnt instantaneous

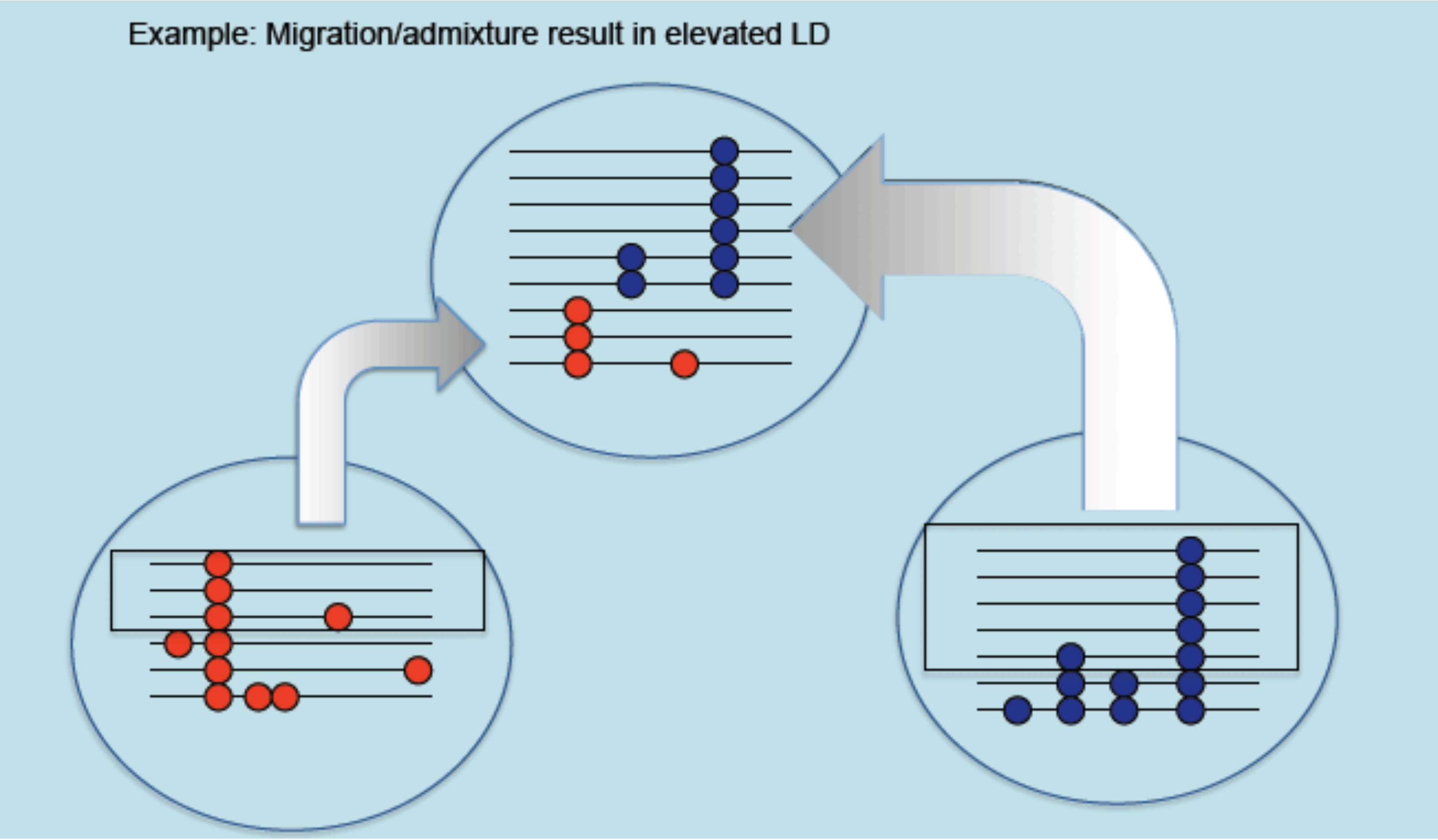
What creates LD

- Mutational origin
- Genetic drift (and Hitchhiking)
- Epistatic selection*
- Assortative mating.
  - Inbreeding
  - Population structure and admixture
  - Assortative mating by phenotype*

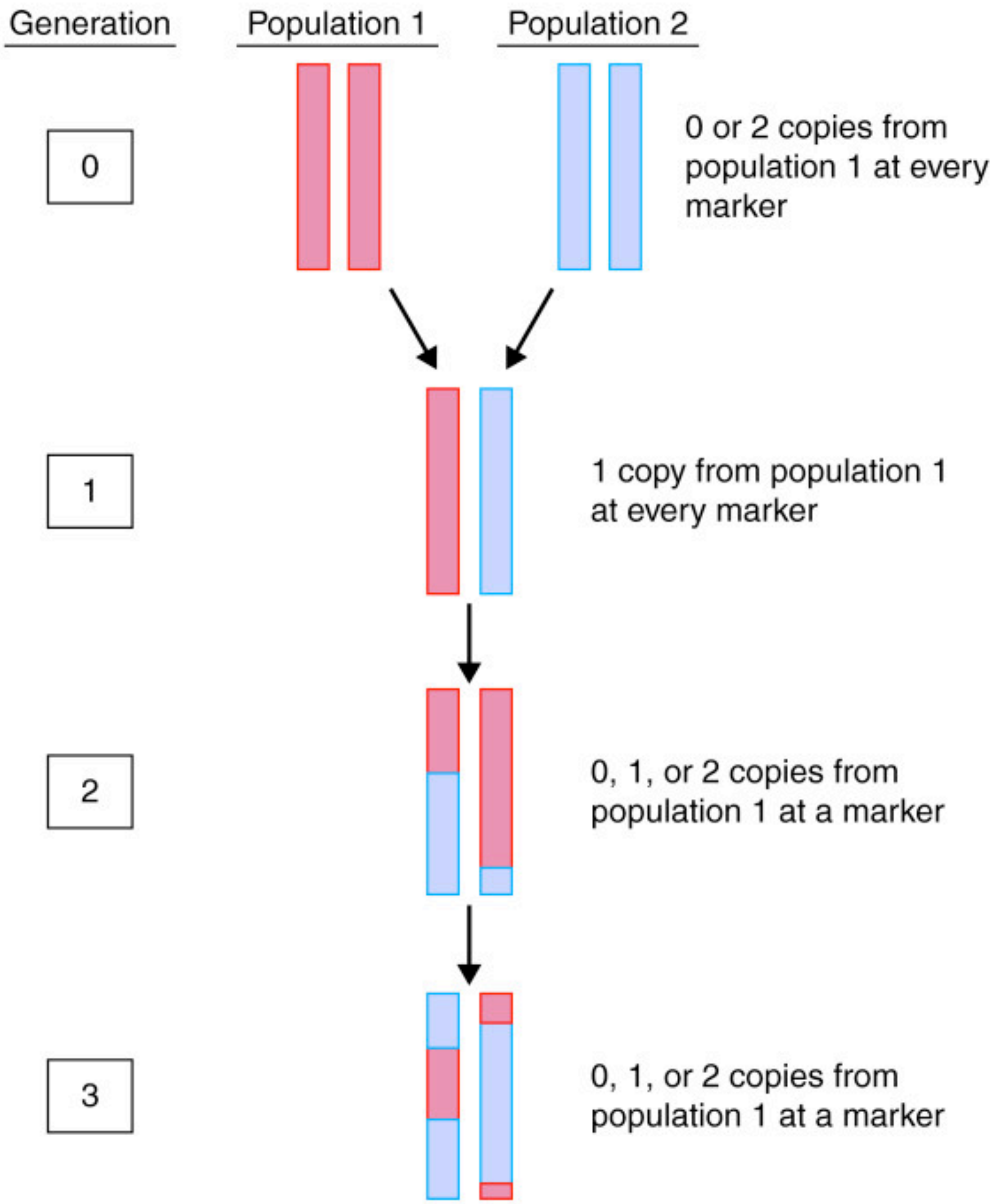
- *only for specific markers

Countervailing forces that increase LD

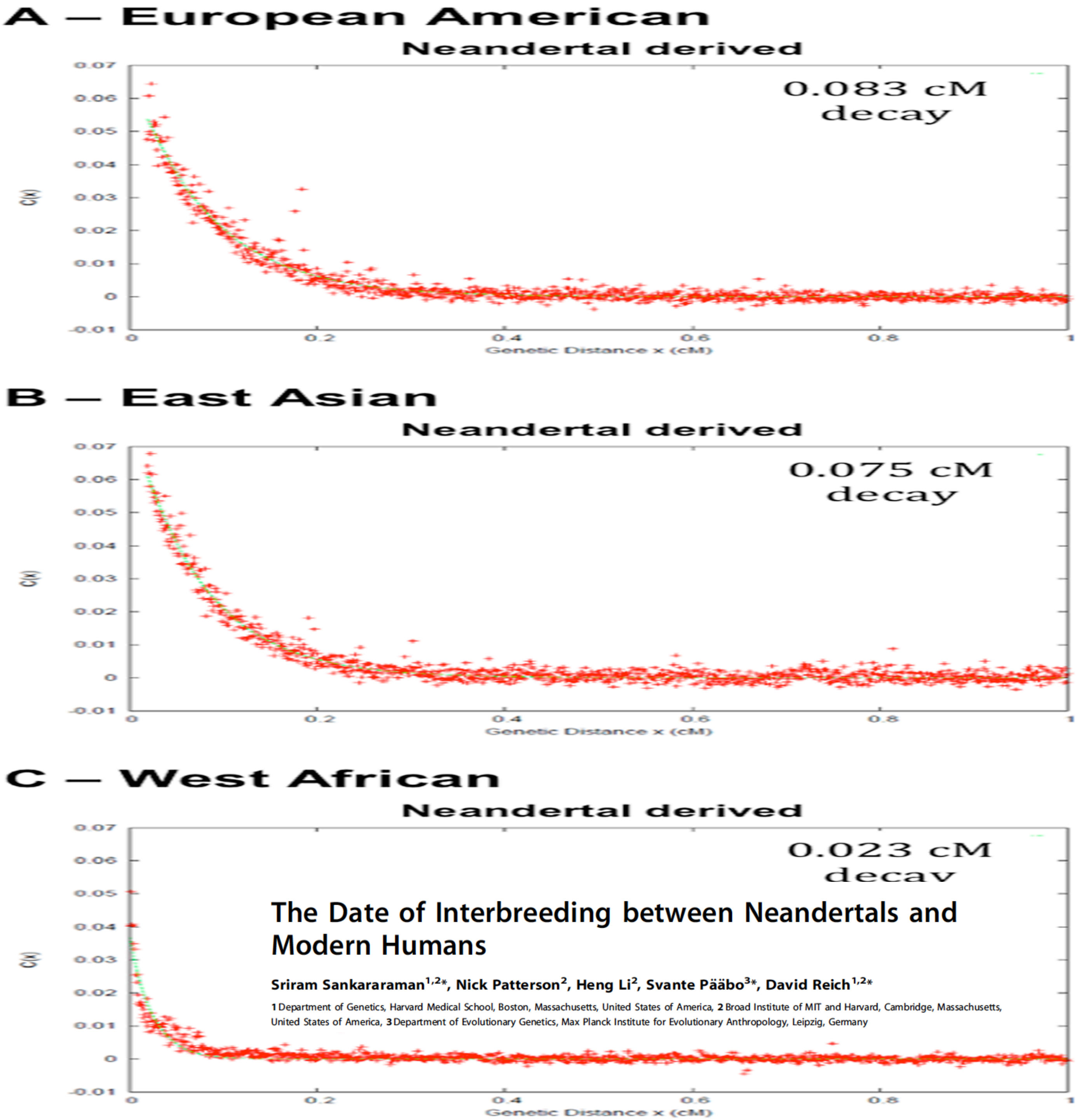
Population structure can increase LD if allele frequencies differ among populations



Pretty pictures courtesy of P. Andolfatto (Princeton)


Generation
Population 1
Population 2
0
0 or 2 copies from population 1 at every marker
1
1 copy from population 1 at every marker
2
0, 1, or 2 copies from population 1 at a marker
3
0, 1, or 2 copies from population 1 at a marker

LD between Neandertal alleles in modern human populations



The Date of Interbreeding between Neandertals and Modern Humans

Sriram Sankararaman[1,2]*, Nick Patterson[2], Heng Li[2], Svante Pääbo[3]*, David Reich[1,2]*

1 Department of Genetics, Harvard Medical School, Boston, Massachusetts, United States of America, 2 Broad Institute of MIT and Harvard, Cambridge, Massachusetts, United States of America, 3 Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Evolution by genetic drift

- Evolution by Genetic drift: a change in allele frequency because individuals carry the allele by chance produce more / less offspring in any given generation. *
- *in sexual populations….

  - Genetic drift can affect selected alleles but only if they are very weakly selected (except when they are rare).

- A neutral allele: An allele with no effect on fitness from other alleles at that locus.

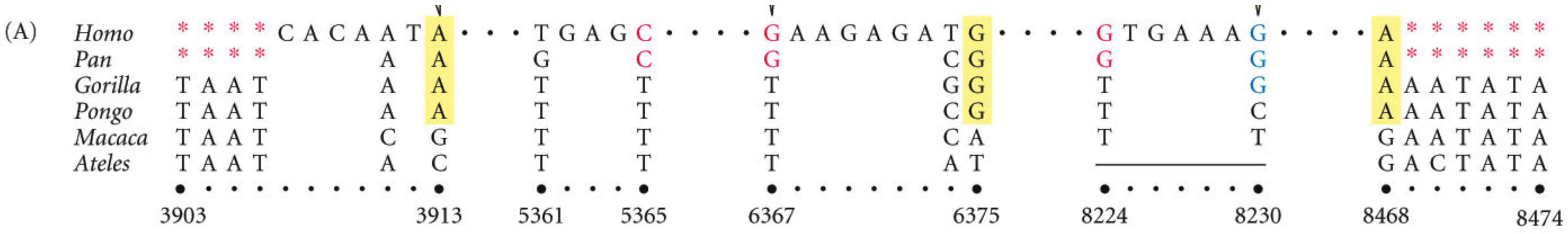
- Neutral polymorphism/alleles:

- Only 2% of our genome encodes for proteins
- Changes outside exons may be completely neutral if they do not disrupt regulatory sites.

- Examples of potentially neutral alleles:
- A synonymous change in a codon.
- A non-synonymous change that replaces one amino-acid with a functionally similar one.
- A non-synonymous change which produces a large change in a phenotype on which selection no longer acts.

How much of genetic divergence between species is neutral.

- ~36 million substitutions have occurred since human and chimp last shared a common ancestor.
- How many of these substitutions fix due to selection?
- How much of polymorphism is neutral?



Why is there so much polymorphism?

- The paradox of variation in population genetics:

Selection quickly fixes alleles that are beneficial so why is there so much genetic polymorphism within natural populations?

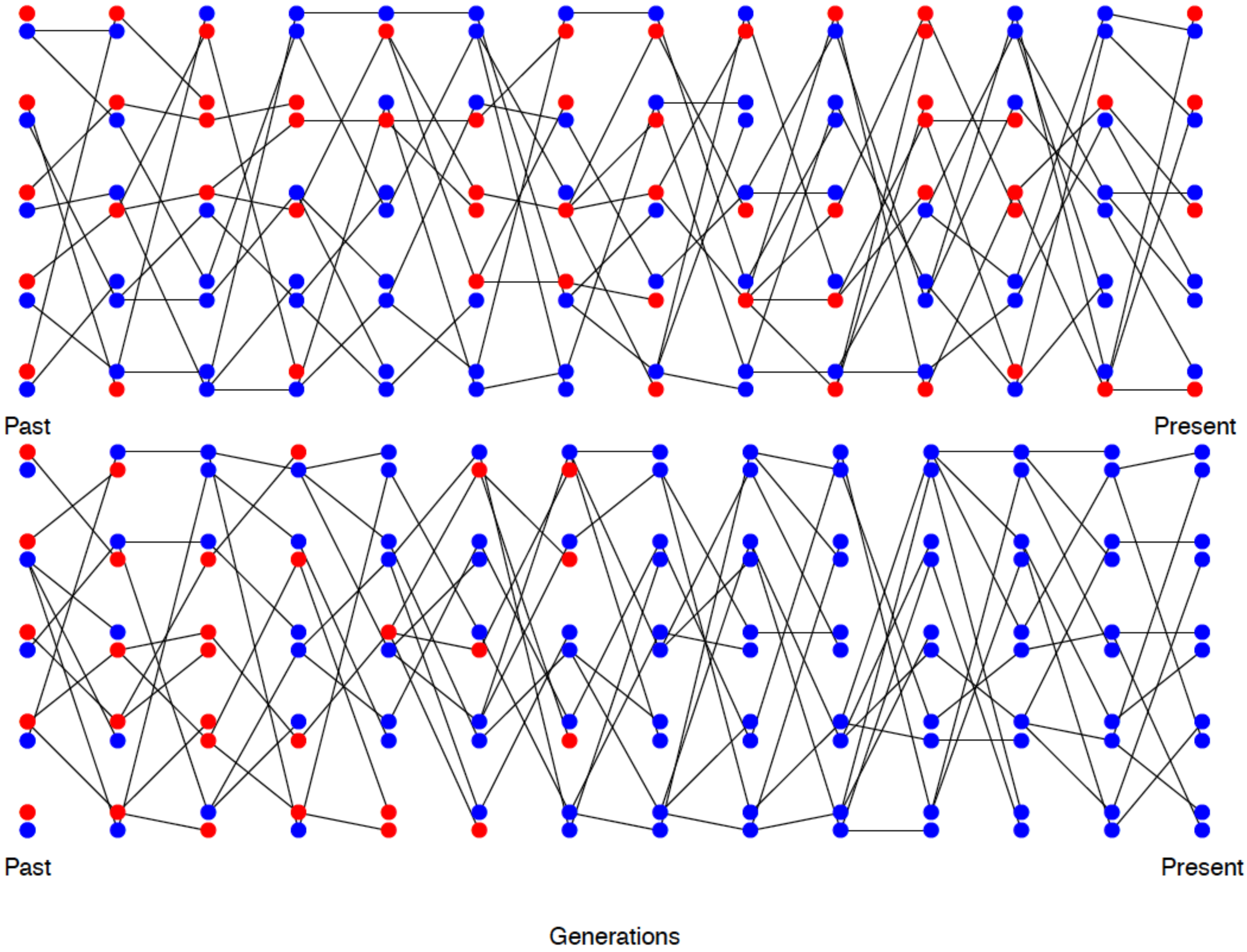
Three explanations:

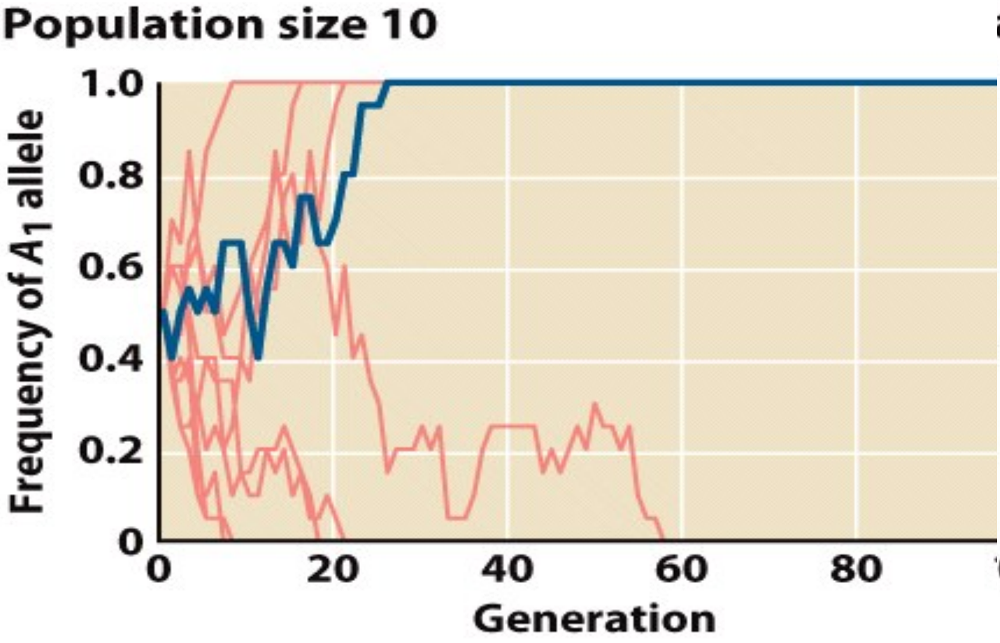
- Balancing selection
- Mutation-selection balance
- Mutation-genetic drift balance (Neutral theory).

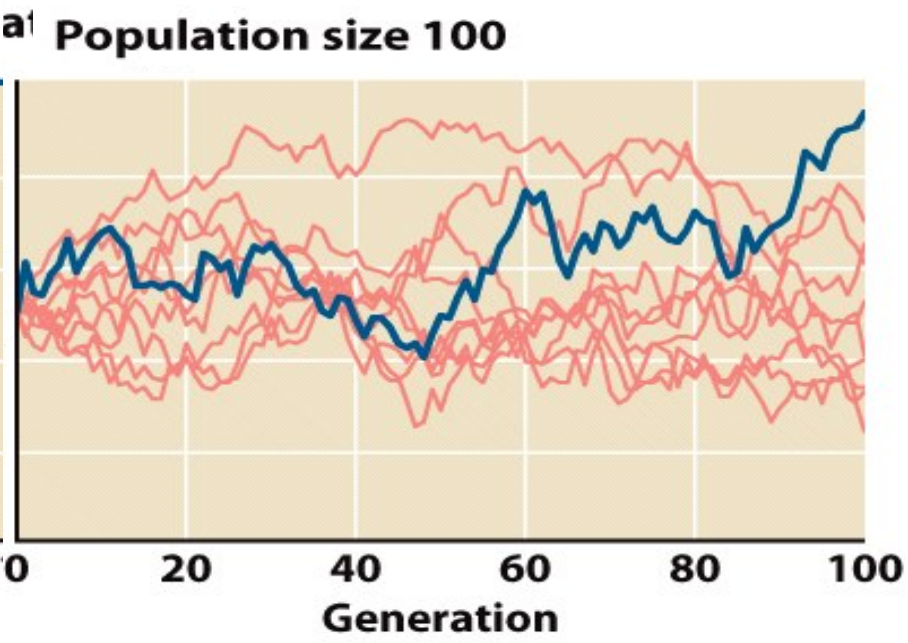
Neutral theory of molecular evolution

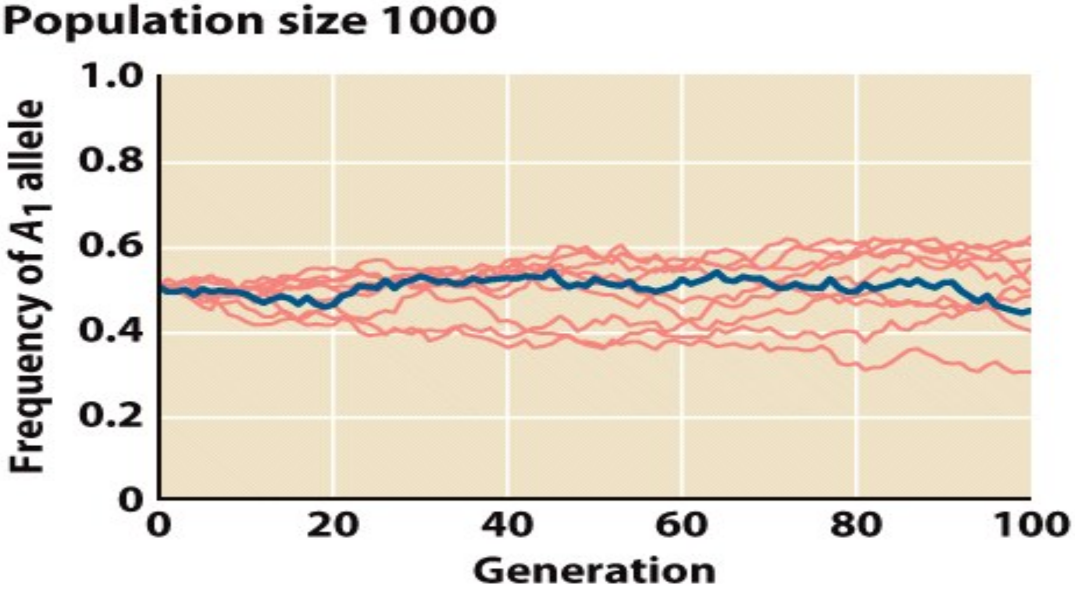
- Kimura 1968; King and Jukes 1969
- Claimed:
  - Most new mutations are deleterious and are lost immediately
  - Most of the observed molecular polymorphism and substitutions are neutral

Claimed that this is consistent with:
-High levels of genetic polymorphism
-The molecular clock


Past
Present
Past
Present
Generations


Population size 10


Population size 100


Population size 1000

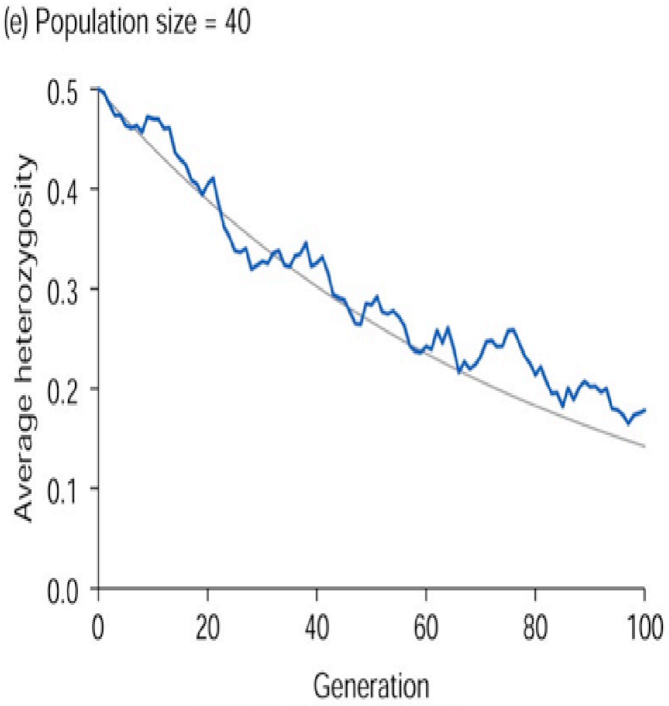
Evolution, 1/e Figure 8.3

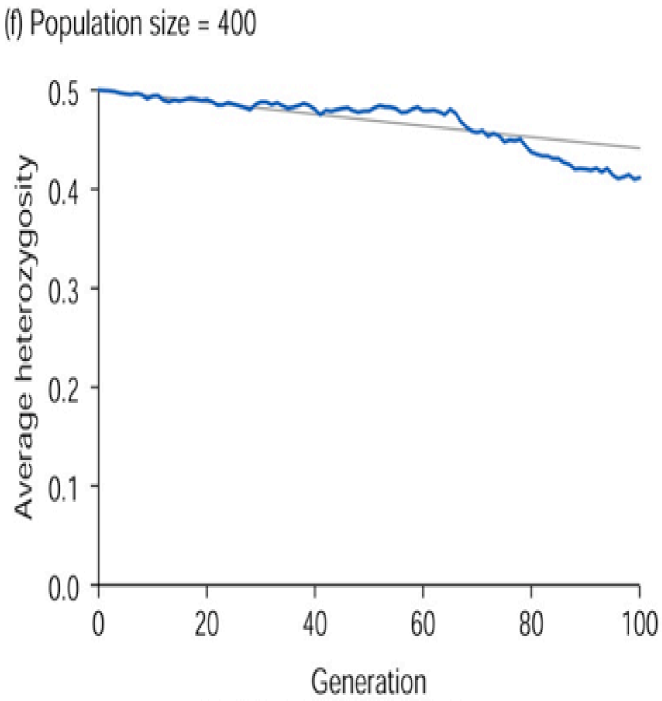
- The rate of genetic drift is higher in smaller populations Allele frequencies do drift in very large populations but at a very slow rate.

Loss of heterozygosity

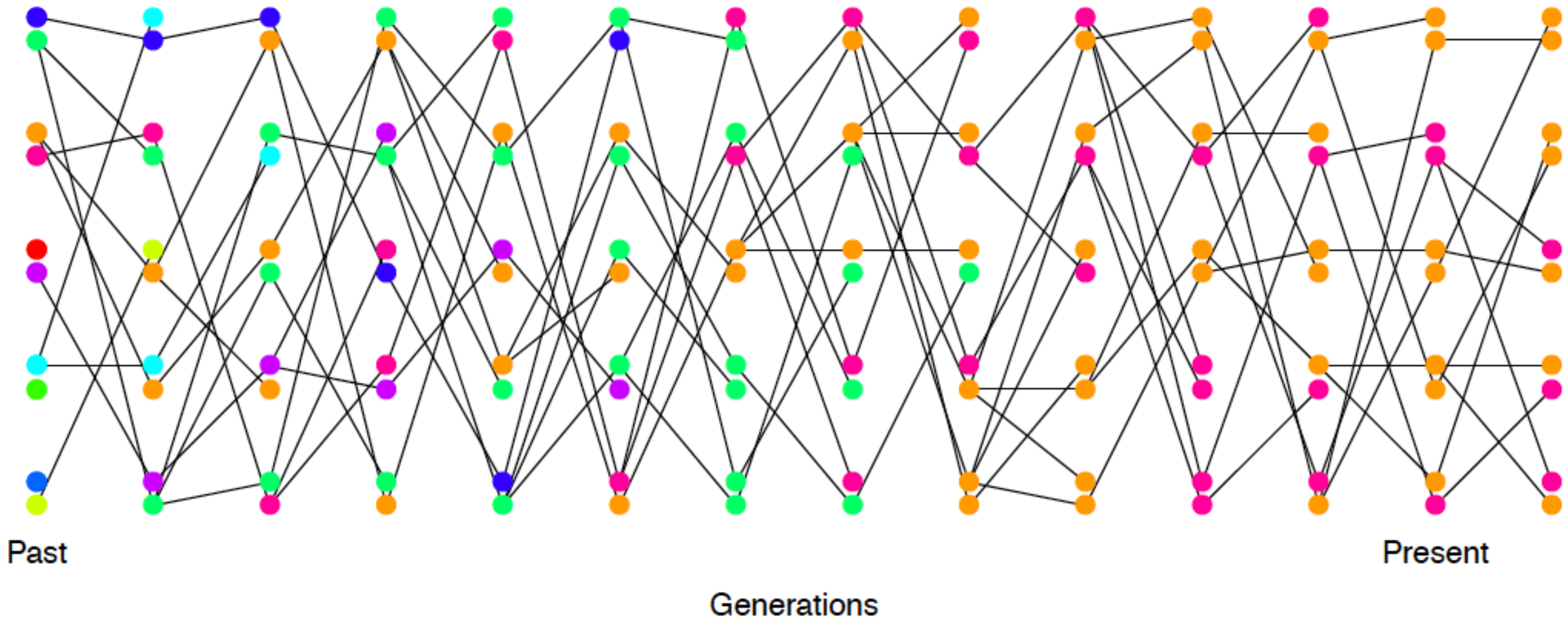
- In the absence of new mutations, alleles drift to either loss or fixation. Thus the amount of heterozygosity within the population decreases a rate inversely proportional to population size

Heterozygosity =
Fraction of sites that
are heterozygous

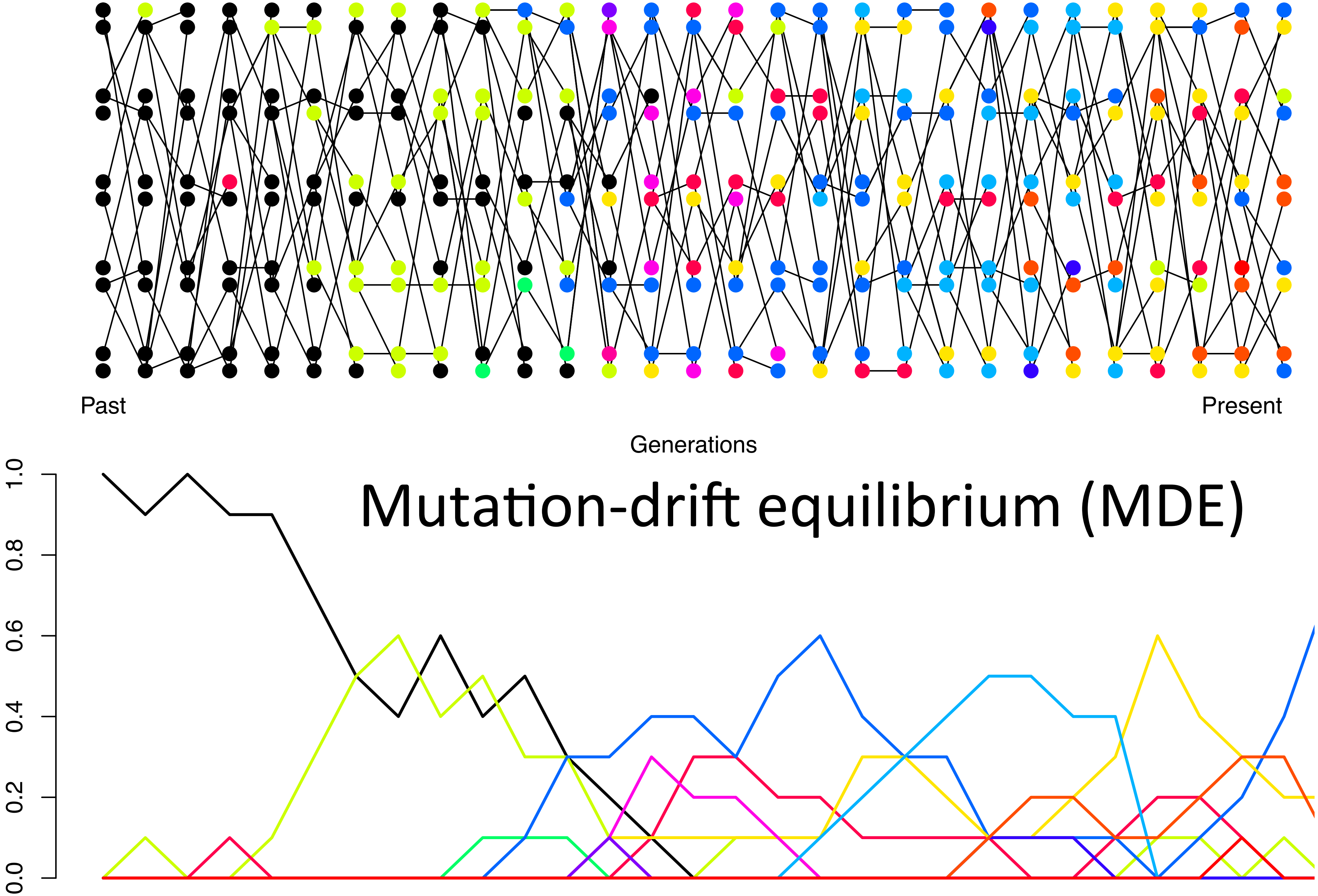

(e) Population size = 40


(f) Population size = 400

Loss of heterozygosity



Mutation-drift equilibrium (MDE)



Mutation-drift equilibrium (MDE)

Mutation-drift equilibrium (MDE)

- Variation lost by drift = variation introduced by mutation
- N_e = effective population size,

μ = mutation rate to new neutral alleles

At MDE, heterozygosity = H = $4 N_e\mu$

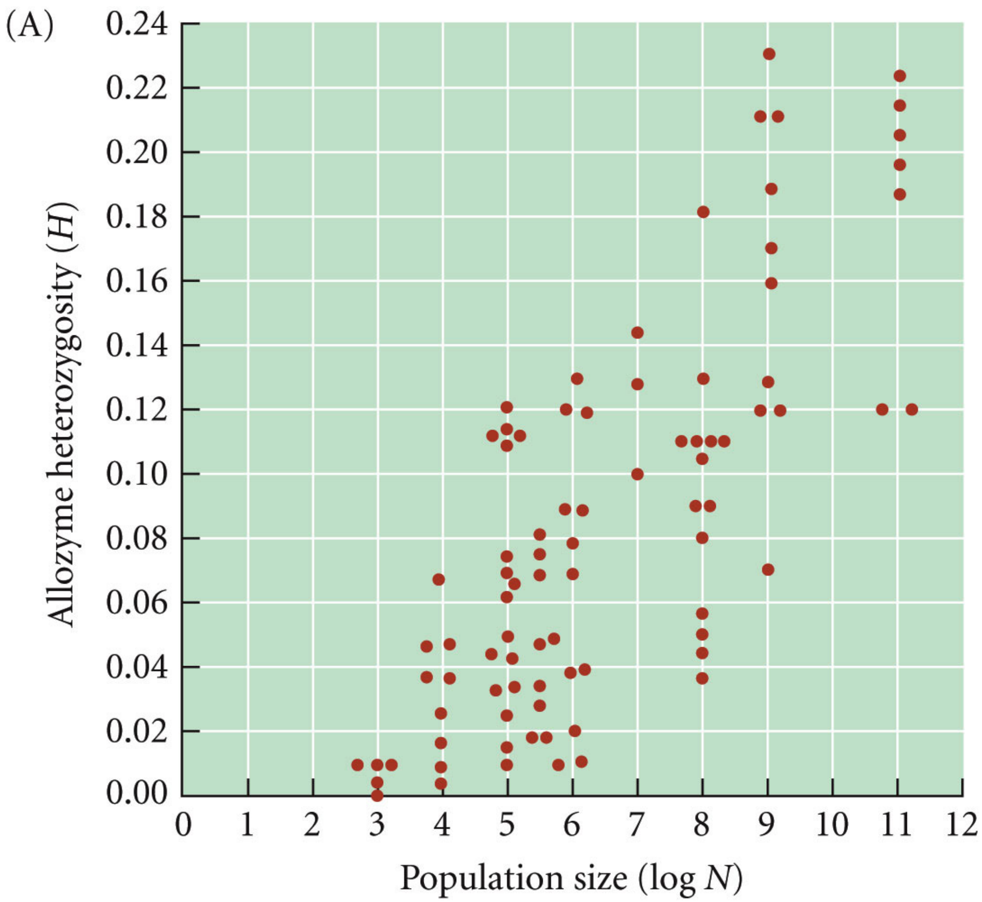
The rate of genetic drift is slower in larger populations and the input of new mutations is higher. Large populations have a higher level of neutral polymorphism

Neutral theory of molecular evolution

- Kimura 1968; King and Jukes 1969
- Claim: Most of the observed molecular polymorphism is neutral

Consistent with Neutral theory

Levels of genetic variability within Populations (i.e. heterozygosity) are high and are broadly correlated with population size



EVOLUTION 2e, Figure 10.9 (Part 1)