Drift and Selection

BIOL 434/509

Wright-Fisher model

Neutral case:

$$T_{i,j} = \Pr[j\ copies\ |i\ copies] = \binom{2N}{j}\left(\frac{i}{2N}\right)^{j}\left(1-\frac{i}{2N}\right)^{2N-j}$$

i is the number of A alleles in the parent generation
j is the number of A alleles in the offspring generation

Wright-Fisher

Thus, for example, if a population of 10 diploid individuals has an allele frequency of 0.2, the probability that it will have 6 copies of this allele in the next generation (i.e. a new allele frequency of 0.3) is

$$T_{4,6} = \binom{20}{6}\left(\frac{4}{20}\right)^{6}\left(1-\frac{4}{20}\right)^{20-6} = 0.1091.$$

Wright-Fisher with selection

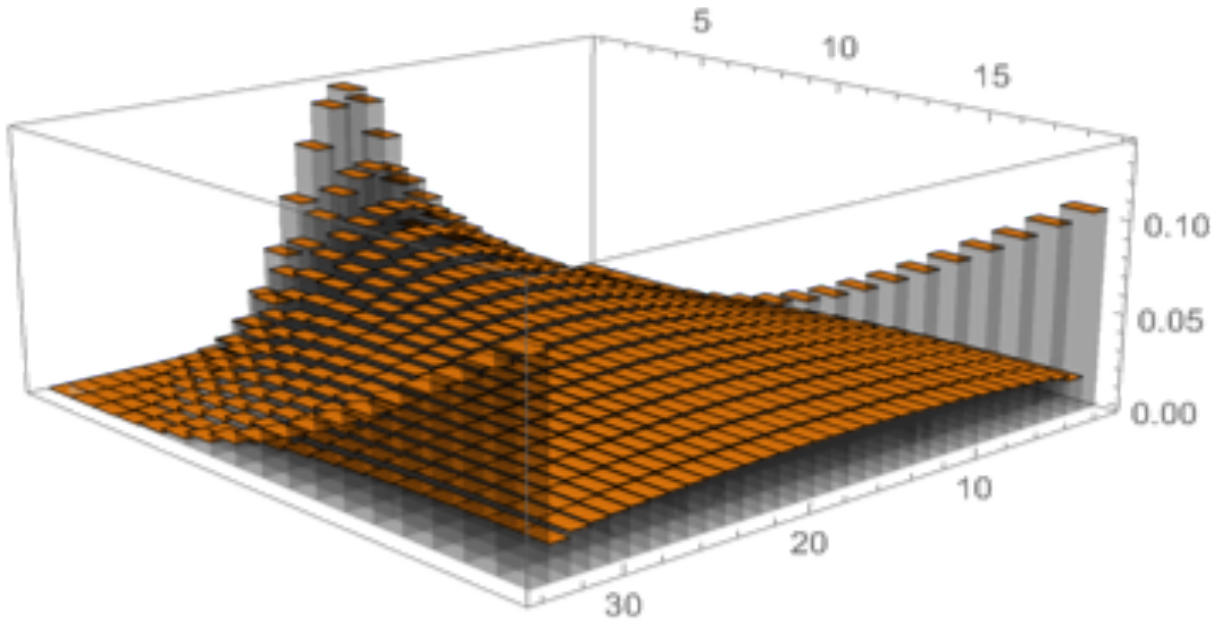
Use $p'[i]$ for the expected allele frequency after deterministic changes.

E.g., with selection $p'[i] = \left(\frac{i}{2N}\right)\frac{w_A}{\bar{w}}$

$$T_{i,j} = \Pr[j\ copies\ |i\ copies]$$
$$= \binom{2N}{j}(p'[i])^j(1-p'[i])^{2N-j}$$

With uncertainty about starting frequency :

$$Pr[X_{t+1} = j] = \sum_i Pr[X_t = i] T_{i,j}$$



Probability of fixation of new alleles

Neutral alleles

The probability of fixation of neutral alleles is simply their starting allele frequency.

For a new allele, this is $\frac{1}{2N}$

Probability of fixation of new alleles

Beneficial alleles (ideal population)

$$\Pr[fixation] = 2s$$

where s is the benefit of the heterozygote

Most beneficial alleles are lost by chance.

Haldane

Probability of fixation of new alleles

Beneficial alleles, non-ideal population

$$\Pr[fixation] = 2s\frac{N_e}{N}$$

where s is the benefit of the heterozygote

Most beneficial alleles are lost by chance.

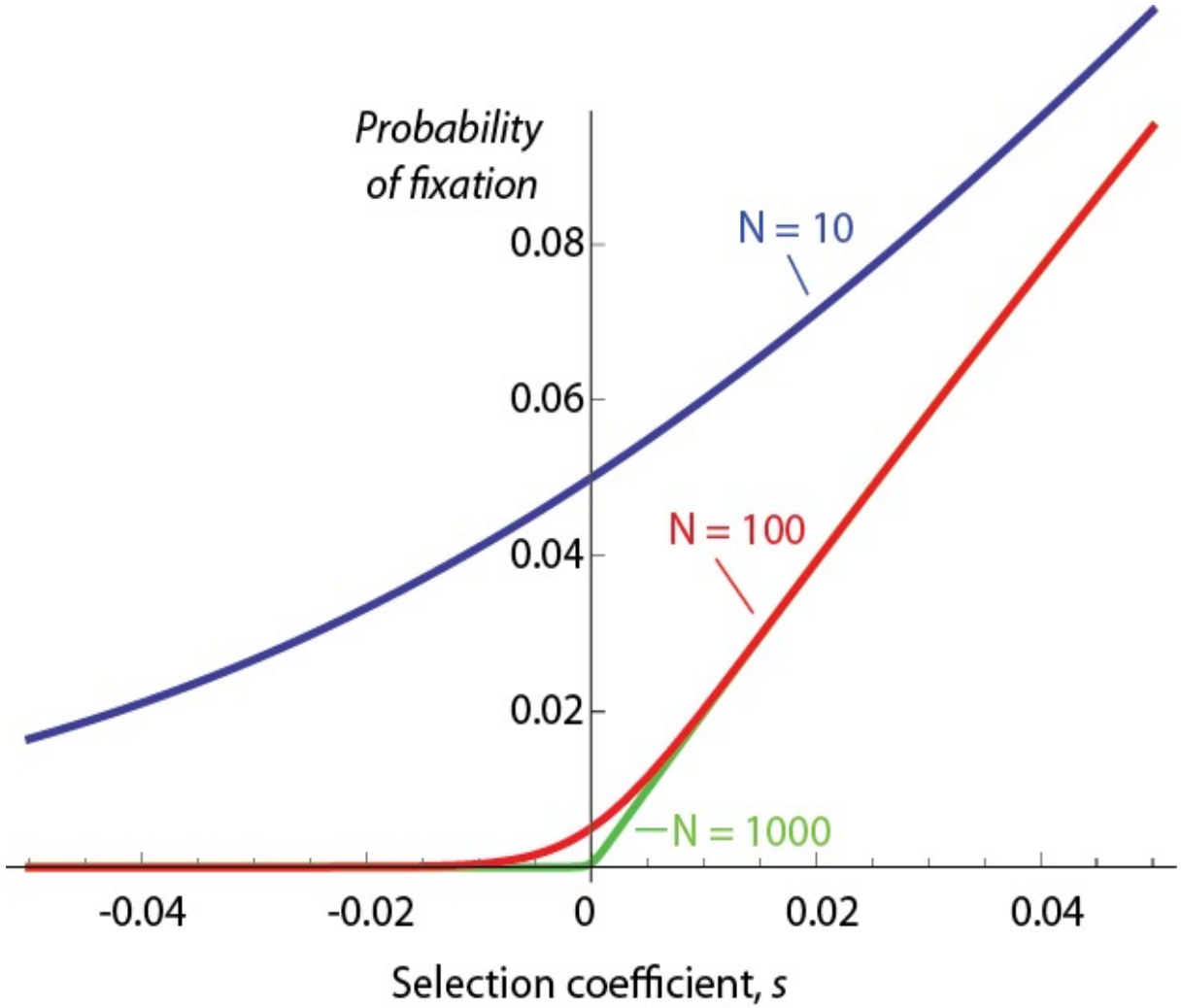
Kimura

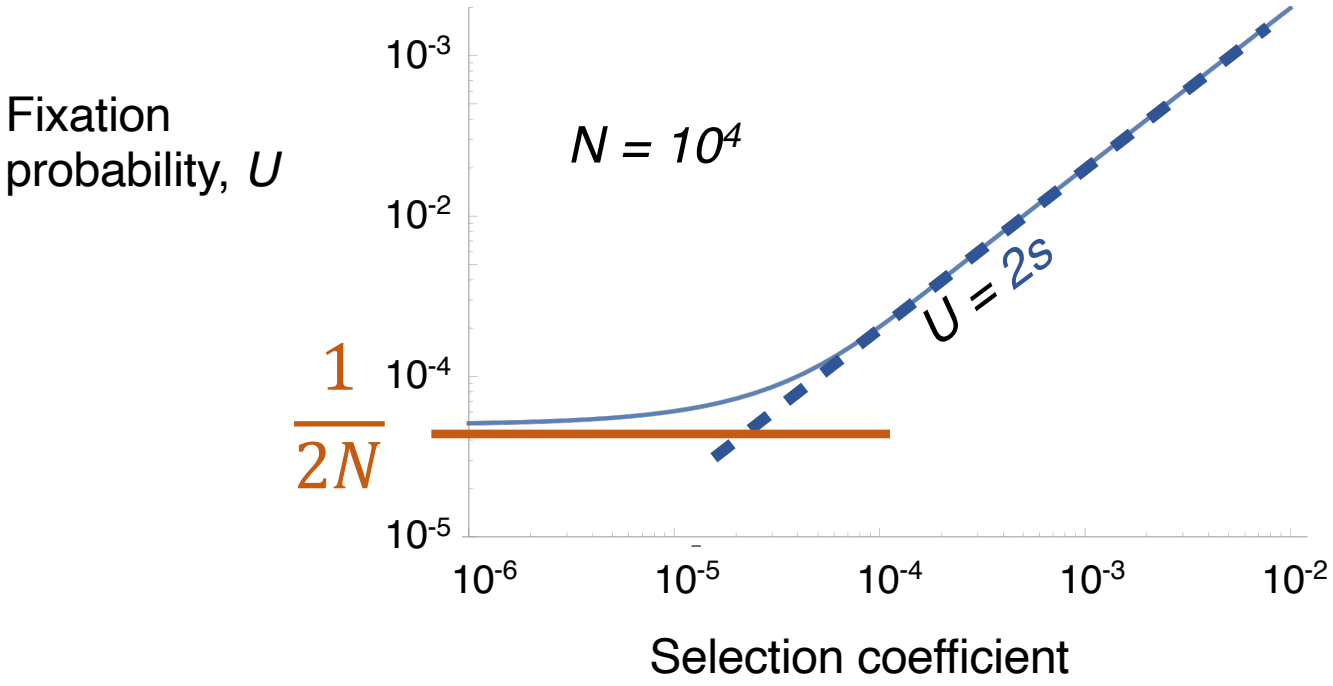
Fixation probability with arbitrary fitness and $N_e \neq N$

$$\Pr[fixation] = \frac{1 - \exp\left[-\frac{2sN_e}{N}\right]}{1 - \exp[-4sN_e]}$$

Kimura

Fixation probability with arbitrary fitness and $N_e \neq N$




Fixation probability, U
10^{-3}
10^{-2}
10^{-4}
10^{-5}
$N = 10^4$
$U = 2s$
$\frac{1}{2N}$
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
Selection coefficient

Drift load

The reduction in mean fitness caused by fixation of deleterious alleles in a finite population.

Effects of selection on linked loci

Allele frequencies at one locus can be influenced by indirect selection at other loci.

Genetic hitchhiking

Selection for one allele increases the frequency of alleles at linked sites. This is called genetic hitchhiking.

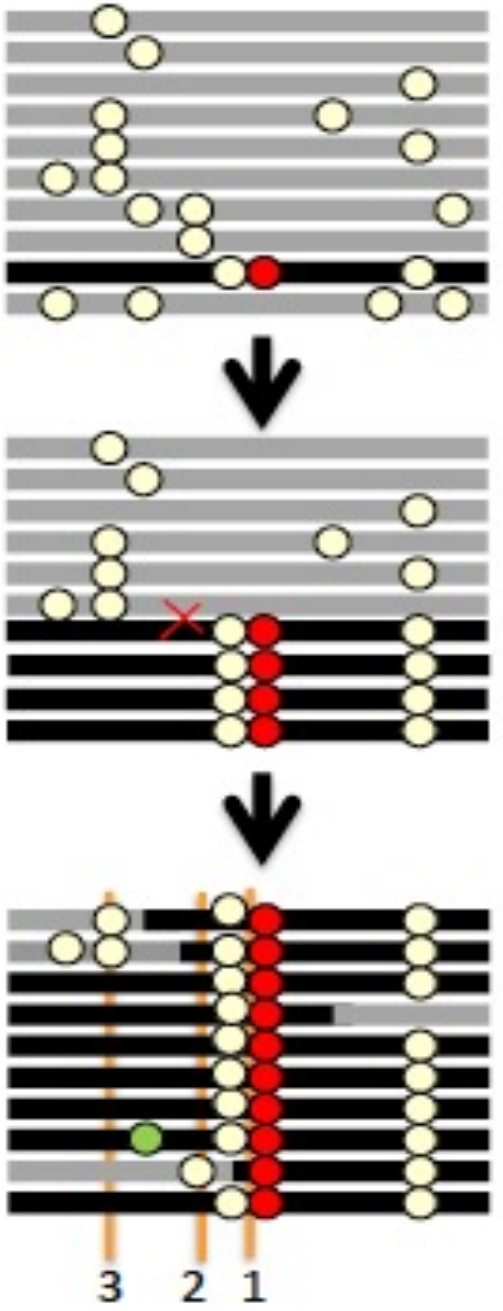


Figure from Coop

Selective sweeps

A selective sweep is the reduction in diversity at linked loci caused by linked selection at new beneficial mutation.

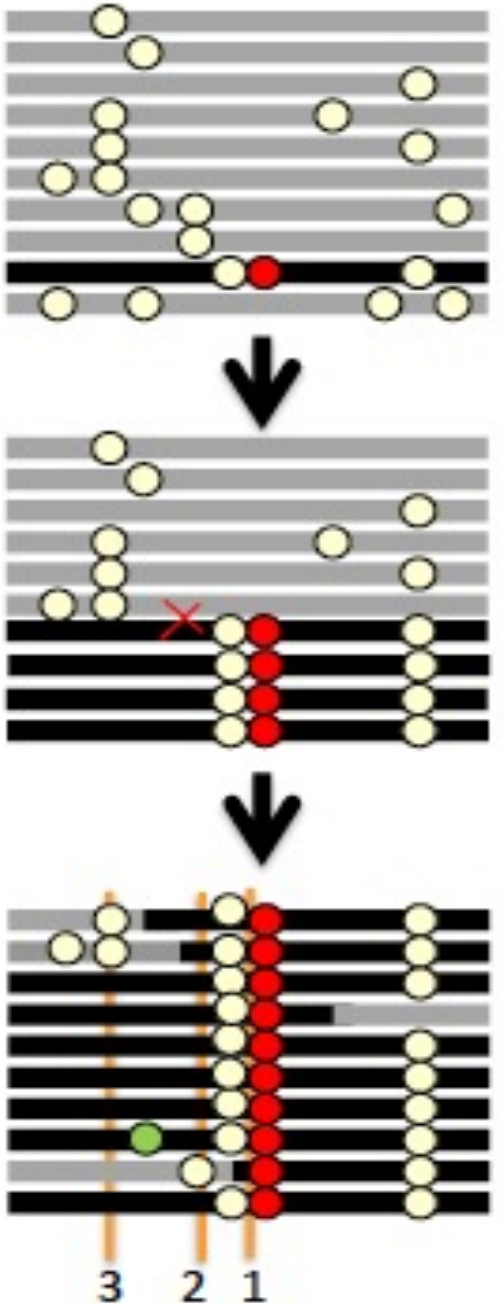
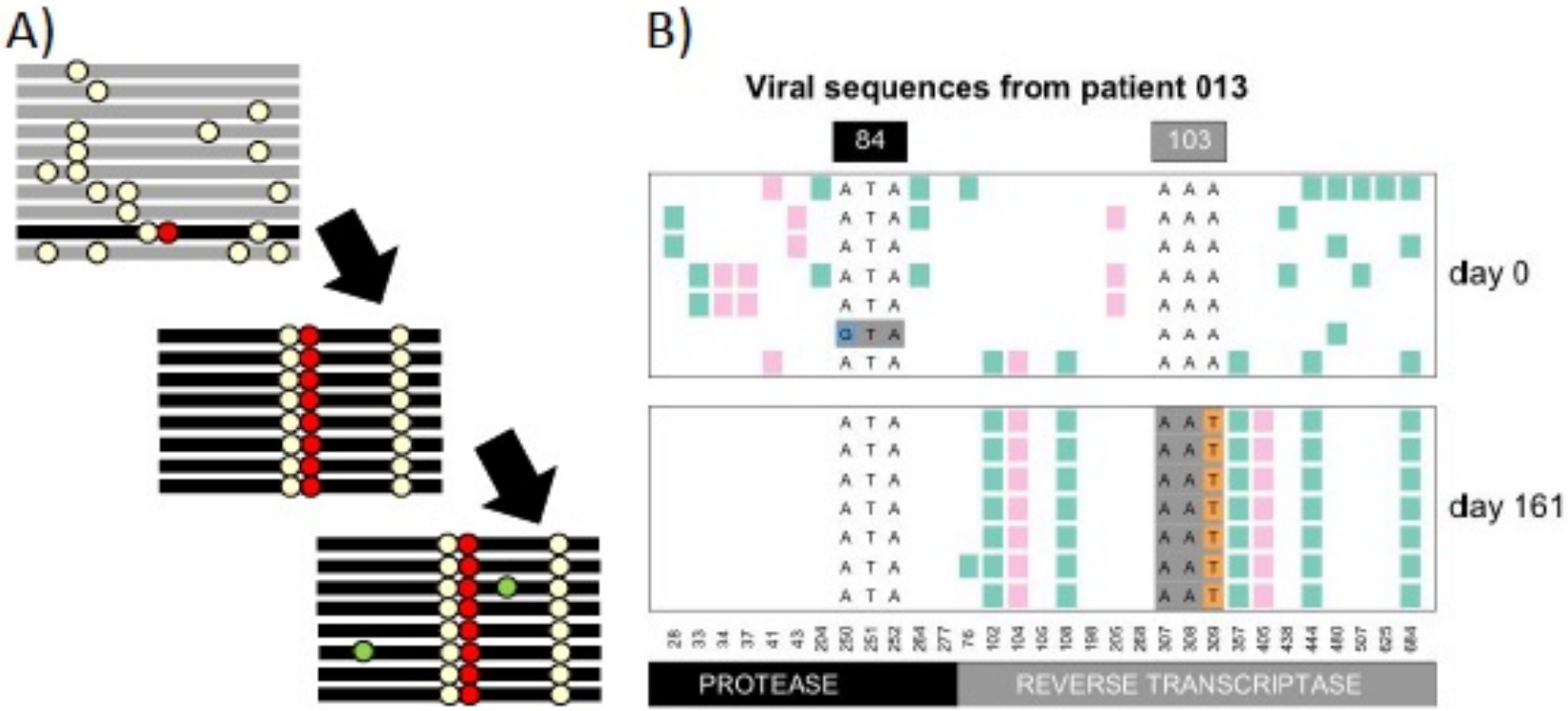


Figure from Coop

Selective sweeps in HIV



Selective sweeps and diversity at linked sites

The selective sweep takes about $\tau = \frac{4 \ln[2N]}{s}$

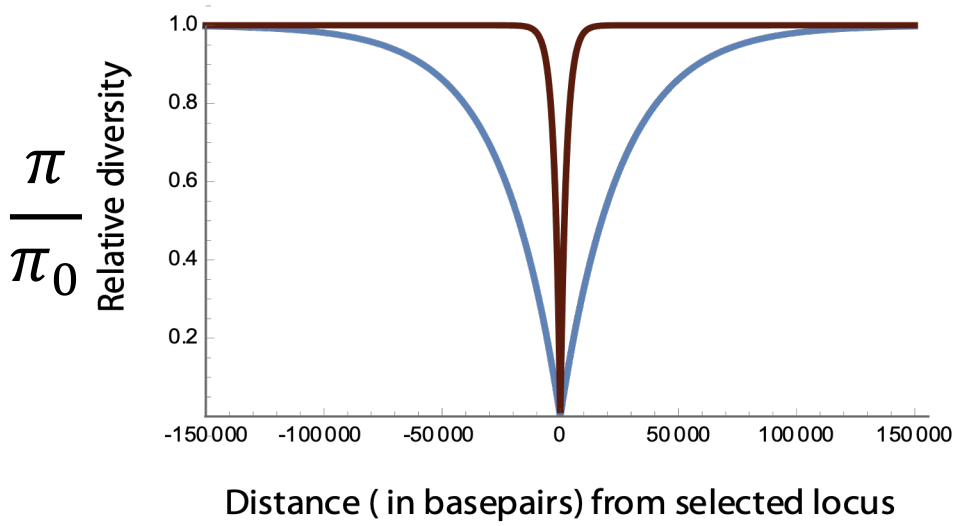
For alleles completely linked to the selected allele ($r = 0$), all alleles in the population will coalesce in the generation that the beneficial mutation appeared generations.

Selective sweeps and diversity at linked sites

The allelic diversity at a site which recombines with the selected locus at rate *r* is predicted by:

$$E[\pi_r] = \pi_0(1 - e^{-r\tau})$$

where π_0 is the diversity at that locus before the selective sweep.



Indirect selection on modifier alleles

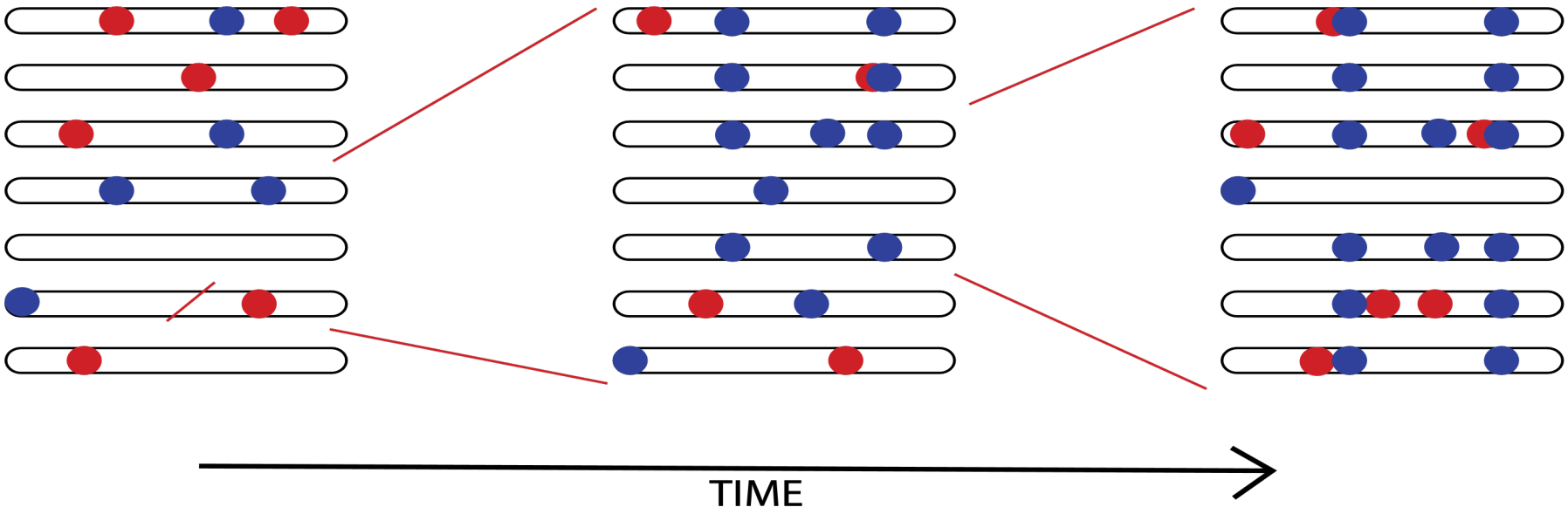
A modifier locus is a site that might affect some genetic property of the genome (or of the genomic region), such as a change in the recombination rate or the mutation rate.

Often such changes do not increase fitness on their own, but these sites might still evolve if the changes they cause are (1) beneficial and (2) linked to the modifier locus.

Example: mutator genes in cancer tumors

Background selection

The effects of deleterious mutations at linked sites on the diversity of a focal locus is called background selection.



Background selection reduces diversity

The reduction in diversity at a locus caused by a selection against deleterious mutation at a linked locus (*r* distance away in the genome) is approximated by:

$$E[\pi] = \pi_0 \left(1 - \frac{\mu s h}{2(r + sh)^2}\right)$$

Over multiple sites

$$E[\pi] = \pi_0 \prod_{i=1}^{L} \left(1 - \frac{\mu_i s_i h_i}{2(r_i + s_i h_i)^2}\right) \cong \pi_0 \exp\left(\sum_{i=1}^{L} \frac{-\mu_i s_i h_i}{2(r_i + s_i h_i)^2}\right)$$

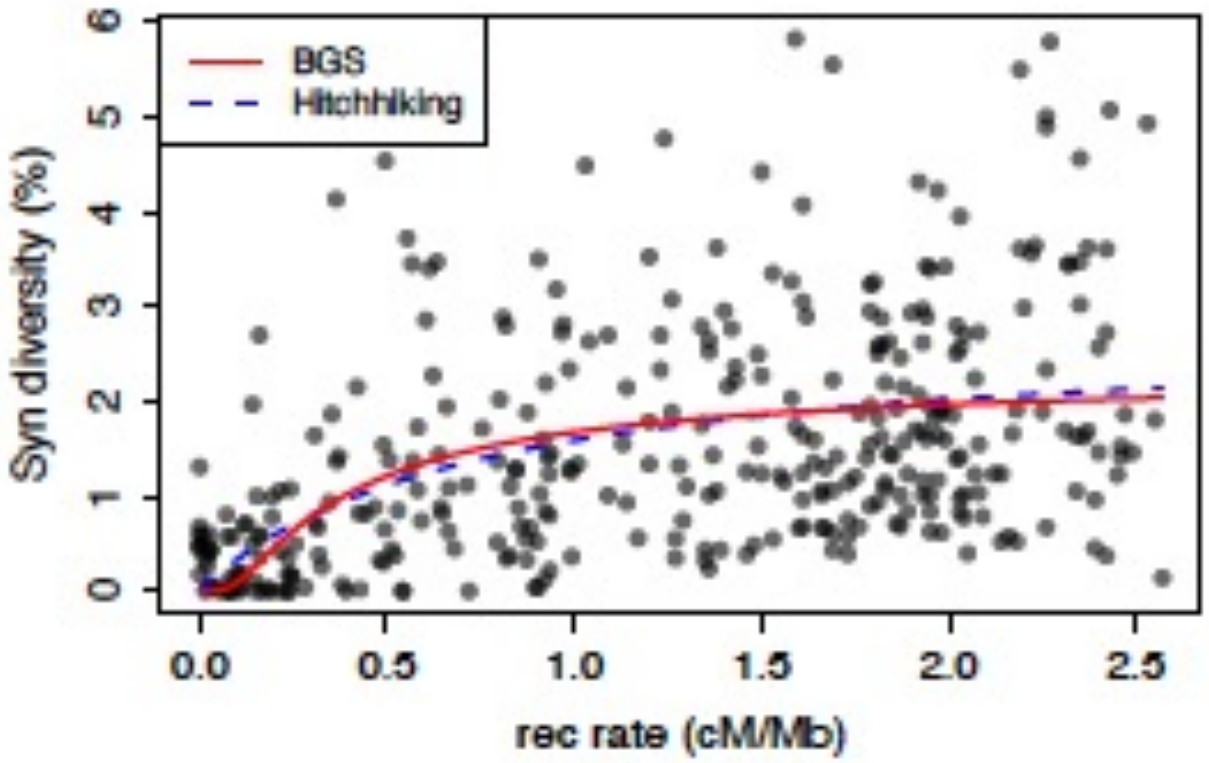


Figure 13.16 from Coop
Recombination rate and diversity in *Drosophila melanogaster*

Evolution of sex

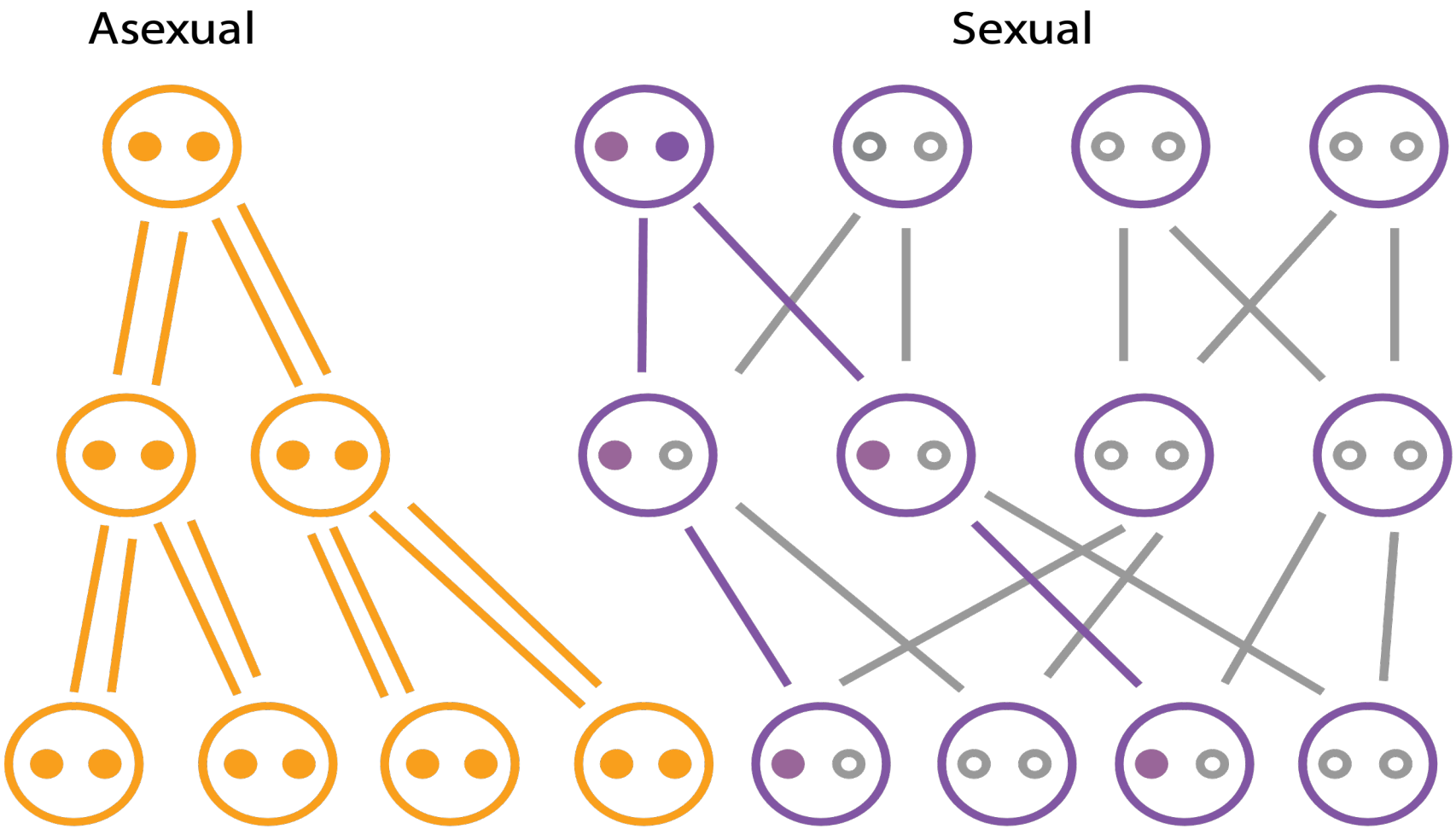
Sexual reproduction occurs when new organisms are formed by fusion of gametes produced by meiosis.

Sex allows recombination and segregation.

Disadvantages of sex

1. Sexual reproduction often incurs costs (searching for mates, sexually transmitted disease, sexual conflict).
2. Recombination and segregation can cause beneficial combinations of alleles (beneficial heterozygotes; epistatically beneficial combinations) to be broken up.
3. Sex creates the cost of males (also known as the twofold cost of sex).

Two-fold cost of sex



Advantages of sex

Sex is advantageous when recombinant genotypes have a higher mean fitness than the currently occurring distribution. (This occurs when breaking up negative associations has a bigger effect on fitness than breaking up positive associations.)

Clonal interference

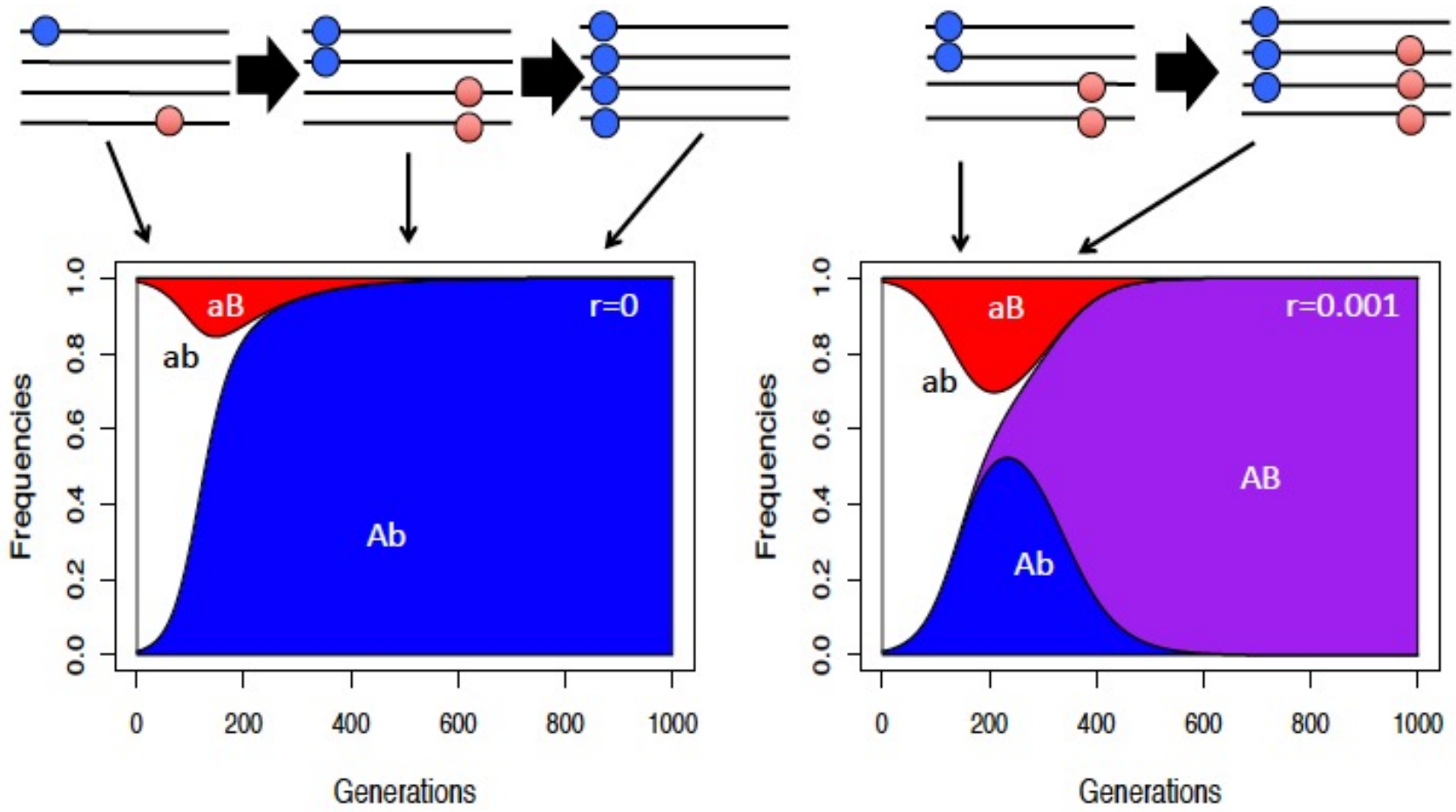
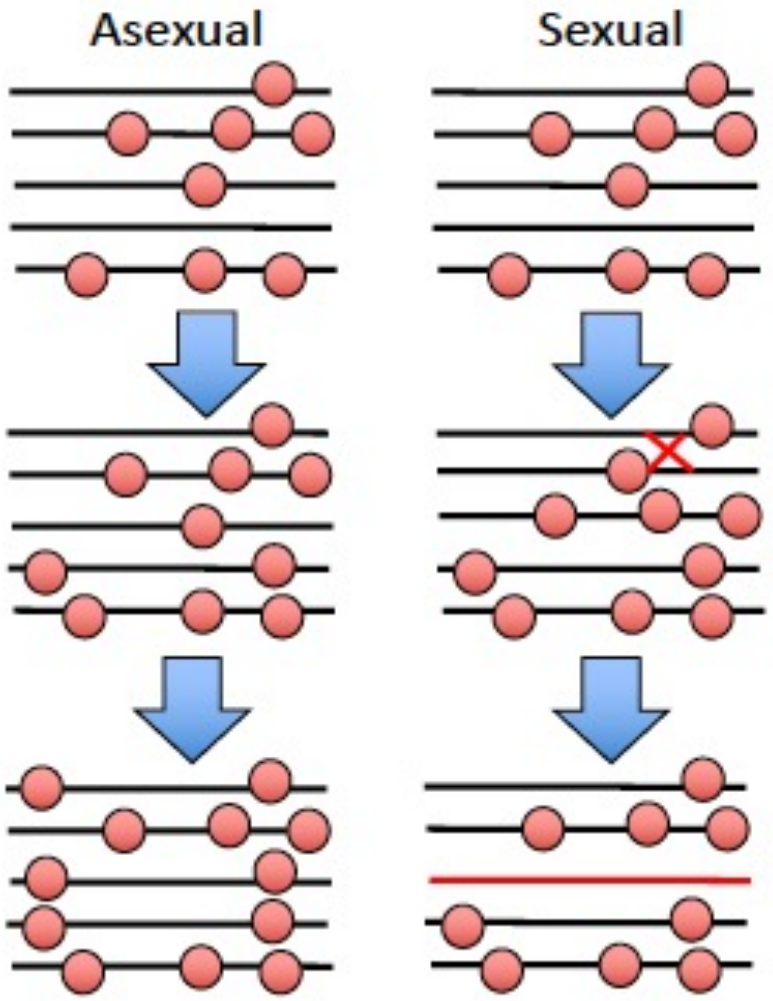


Figure 1.3 from Coop 2020

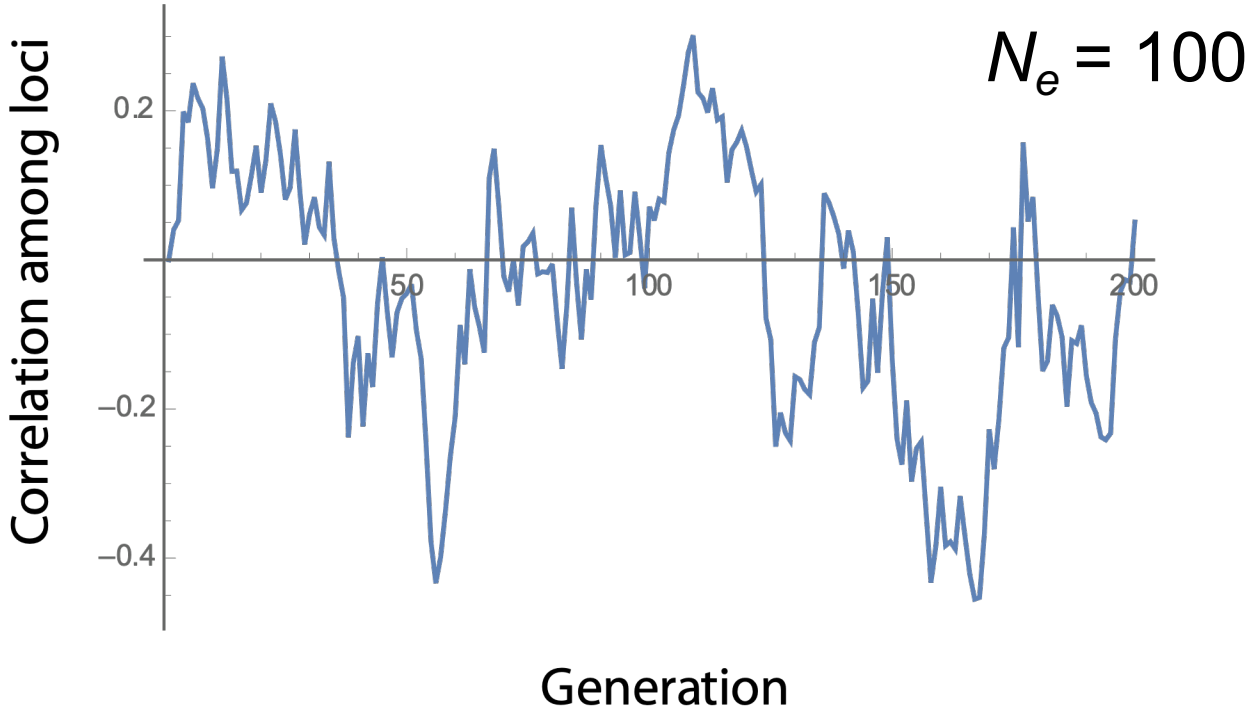
Muller's ratchet

In an asexual population the genotype with the fewest deleterious mutations is rare and can be lost by chance. Then there is no way to recover that high fitness genotype.



Drift, selection, and sex

Drift can generate linkage disequilibrium. On average drift does not lead to LD, but in any given population drift may randomly create an association between alleles.



Drift, selection, and sex

When LD is positive (between favorable alleles) selection quickly acts to remove deleterious combinations.

When LD is negative (association between beneficial and deleterious alleles), selection is less effective. (Both Good-Bad and Bad-Good combinations may have similar fitness, and selection cannot eliminate the bad alleles.)

As a result, in finite populations the combination of drift and selection generates negative associations, meaning that increased recombination can be beneficial to create fit offspring.