

Forward in time models of allele frequency change under neutrality

Heterozygosity: probability that two randomly chosen gene copies are of different allelic types

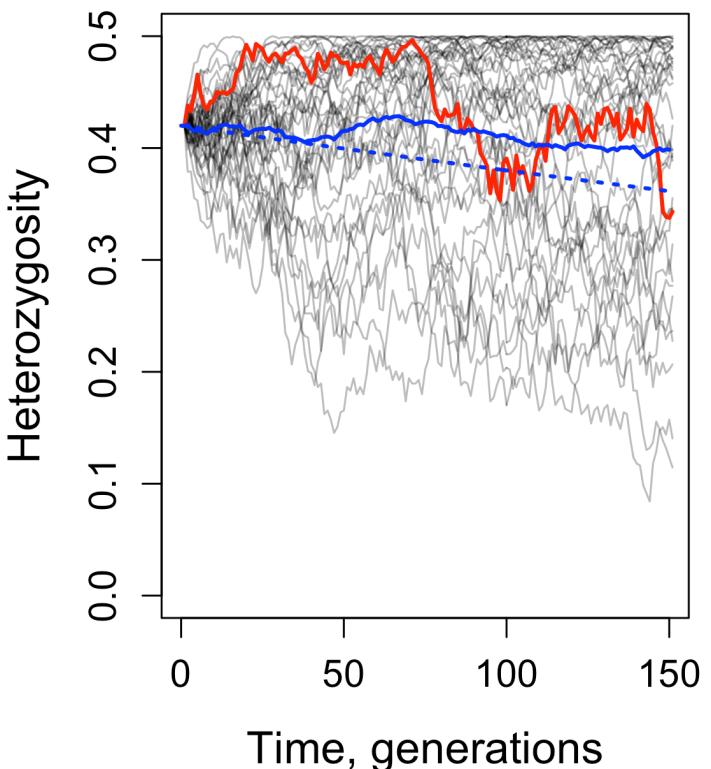
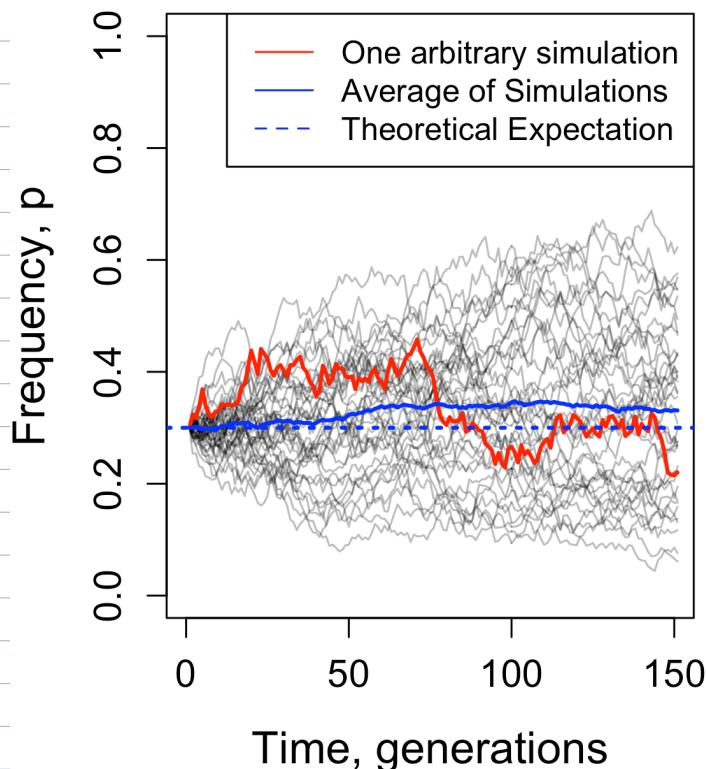
$$H_t = 2p_t q_t$$

Forward in time models of
allele frequency change under neutrality

Heterozygosity: probability that two randomly chosen
gene copies are of different allelic types

$$H_t = 2p_t q_t$$

$N=500$

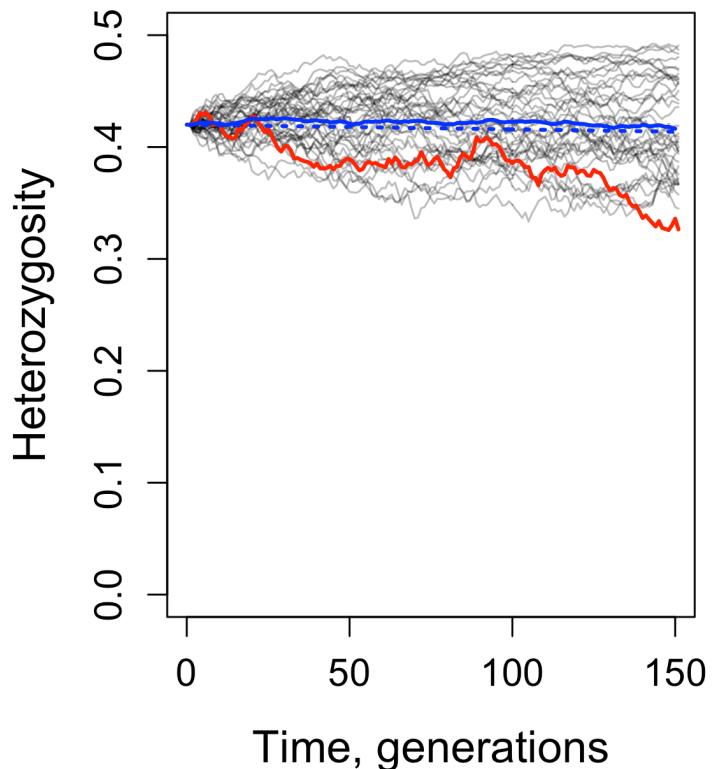
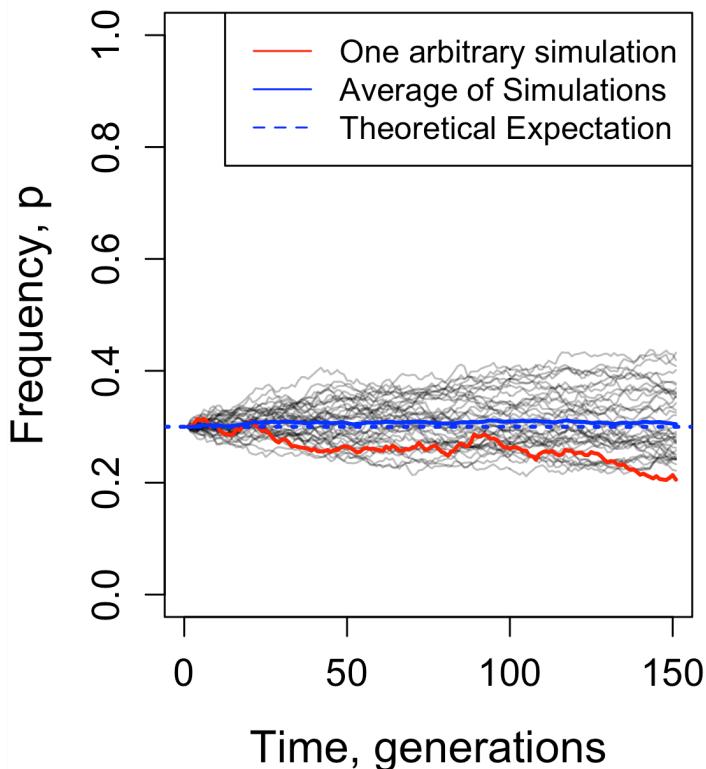


Forward in time models of
allele frequency change under neutrality

Heterozygosity: probability that two randomly chosen
gene copies are of different allelic types

$$H_t = 2p_t q_t$$

$$N=5000$$

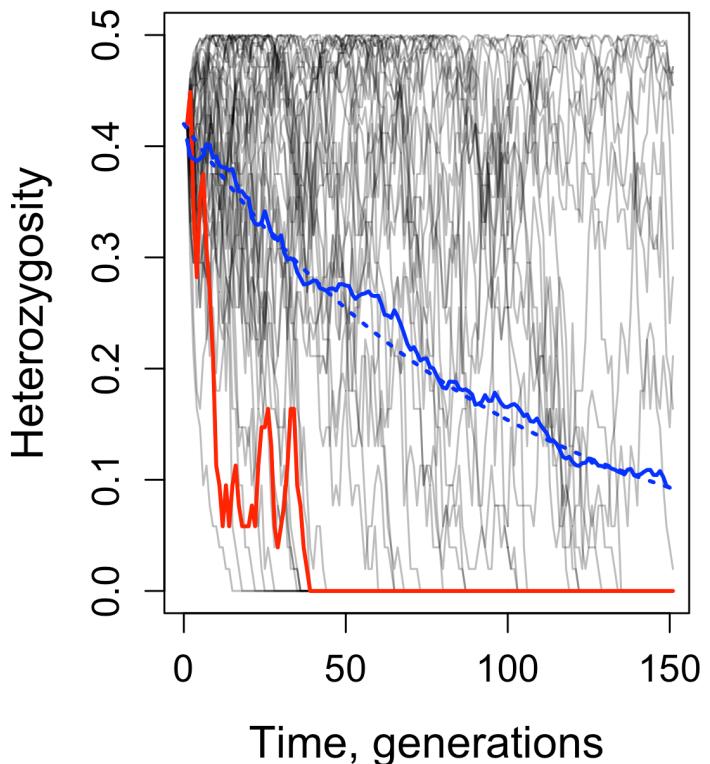
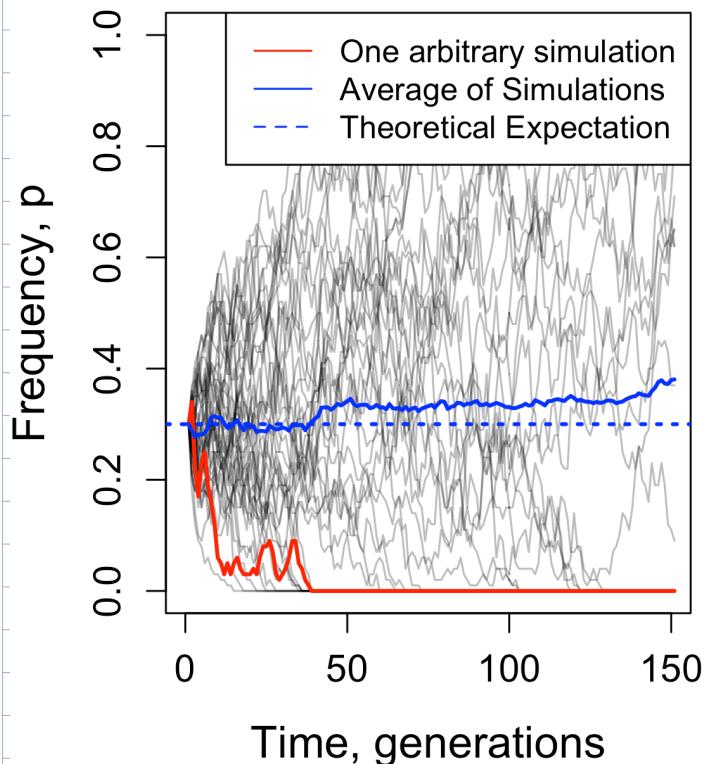


Forward in time models of allele frequency change under neutrality

Heterozygosity: probability that two randomly chosen gene copies are of different allelic types

$$H_t = 2p_t q_t$$

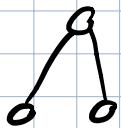
$$N=50$$



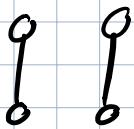
Genetic drift causes a decay in heterozygosity

Genetic drift causes a decay in heterozygosity

Case 1: two alleles
descended from same
allele in previous
generation

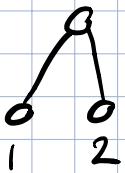


Case 2: two alleles
descended from different
alleles in the previous
generation

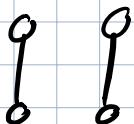


Genetic drift causes a decay in heterozygosity

Case 1: two alleles descended from same allele in previous generation



Case 2: two alleles descended from different alleles in the previous generation



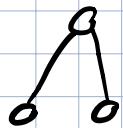
$$H_{++1} = \frac{1}{2N} \cdot 0 + \left(1 - \frac{1}{2N}\right) H_+$$

$$H_{++1} = \left(1 - \frac{1}{2N}\right) H_+$$

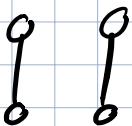
$$H_{++2} = \left(1 - \frac{1}{2N}\right) H_{++1}$$

Genetic drift causes a decay in heterozygosity

Case 1: two alleles descended from same allele in previous generation



Case 2: two alleles descended from different alleles in the previous generation



$$H_{t+1} = \frac{1}{2N} \cdot 0 + \left(1 - \frac{1}{2N}\right) H_t$$

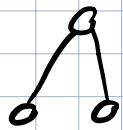
$$H_{t+1} = \left(1 - \frac{1}{2N}\right) H_t$$

After T generations:

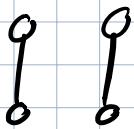
$$H_T = \left(1 - \frac{1}{2N}\right)^T H_0$$

Genetic drift causes a decay in heterozygosity

Case 1: two alleles descended from same allele in previous generation



Case 2: two alleles descended from different alleles in the previous generation



$$H_{t+1} = \frac{1}{2N} \cdot 0 + \left(1 - \frac{1}{2N}\right) H_t$$

$$H_{t+1} = \left(1 - \frac{1}{2N}\right) H_t$$

After T generations:

$$H_T = \left(1 - \frac{1}{2N}\right)^T H_0$$

$$\left(1 - \frac{1}{2N}\right)^T \approx e^{-\frac{T}{2N}}$$

$$H_T \approx e^{-\frac{T}{2N}} H_0$$

Equilibrium levels of heterozygosity are maintained by a balance between drift and mutation.

Equilibrium levels of heterozygosity are maintained by a balance between drift and mutation.

$$H_{t+1} = \left(1 - \frac{1}{2N}\right) H_t + \text{new variation due to mutation}$$

Equilibrium levels of heterozygosity are maintained by a balance between drift and mutation.

$$H_{t+1} = \left(1 - \frac{1}{2N}\right) H_t + \text{new variation due to mutation}$$

$$H_{eq} = \left(1 - \frac{1}{2N}\right) H_{eq} + (1 - H_{eq}) \cdot 2\mu$$

Equilibrium levels of heterozygosity are maintained by a balance between drift and mutation.

$$H_{t+1} = \left(1 - \frac{1}{2N}\right) H_t + \text{new variation due to mutation}$$

$$H_{eq} = \left(1 - \frac{1}{2N}\right) H_{eq} + (1 - H_{eq}) \cdot 2\mu$$



Do the algebra

$$H_{eq} = \frac{4N\mu}{4N\mu + 1}$$

Equilibrium levels of heterozygosity are maintained by a balance between drift and mutation.

$$H_{t+1} = \left(1 - \frac{1}{2N}\right) H_t + \text{new variation due to mutation}$$

$$H_{eq} = \frac{4N_u}{4N_u + 1}$$

$\approx 4N_u$ if $4N_u$ is small

Equilibrium levels of heterozygosity are maintained by a balance between drift and mutation.

$$H_{t+1} = \left(1 - \frac{1}{2N}\right) H_t + \text{new variation due to mutation}$$

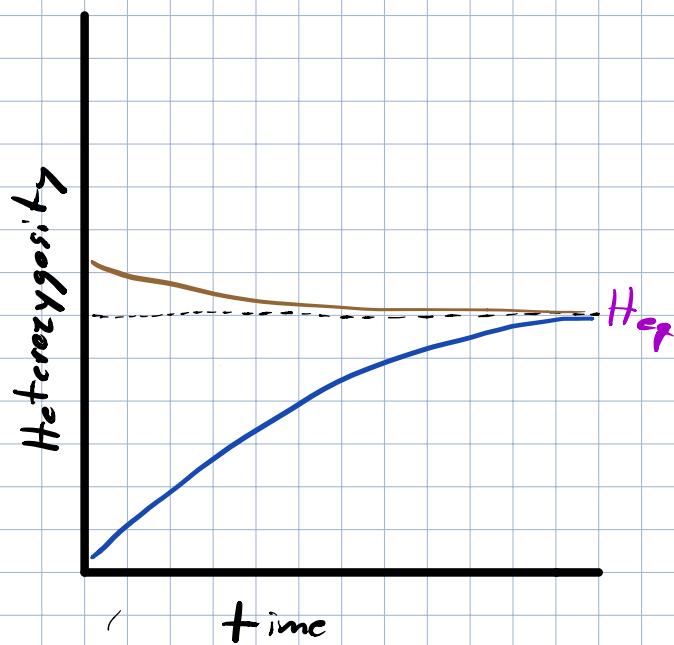
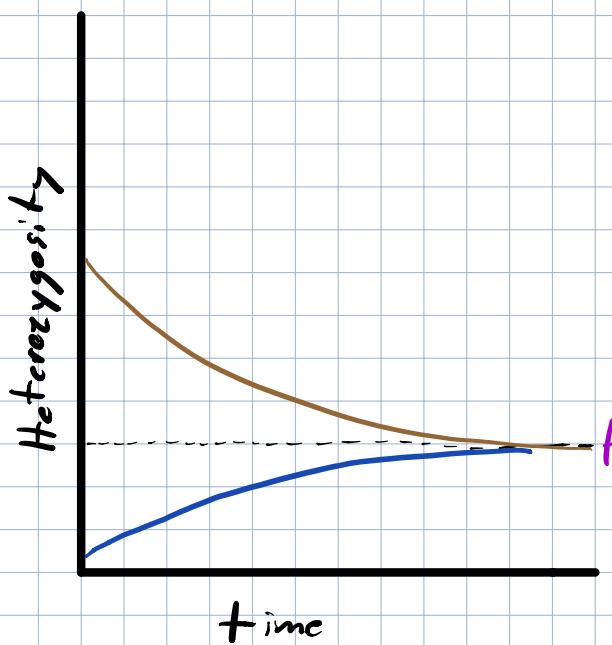
$$H_{eq} = \frac{4N_u}{4N_u + 1}$$

$\approx 4N_u$ if $4N_u$ is small

We will also show how to derive this from a backward in time perspective in the next lecture

Small population

Large population



Effective population size

Effective population size

- Average per base pair heterozygosity
in humans is roughly 1 in 1000

Effective population size

- Average per base pair heterozygosity in humans is roughly 1 in 1000
- If we are willing to make an assumption that current levels of heterozygosity are the result of an equilibrium between drift and mutation, then we can use the formula from before to estimate the "effective size" of the population

Effective population size

- Average per base pair heterozygosity in humans is roughly 1 in 1000
- If we are willing to make an assumption that current levels of heterozygosity are the result of an equilibrium between drift and mutation, then we can use the formula from before to estimate the "effective size" of the population

$$H = \frac{1}{1000} = qN\mu \quad \begin{matrix} \text{can be estimated} \\ \text{separately} \end{matrix}$$

$$N_e = \frac{1}{1000} \cdot \frac{1}{q\mu} \quad \mu = 1.3 \times 10^{-8}$$
$$\approx 29,000$$

Effective population size

- Average per base pair heterozygosity in humans is roughly 1 in 1000
- If we are willing to make an assumption that current levels of heterozygosity are the result of an equilibrium between drift and mutation, then we can use the formula from before to estimate the "effective size" of the population

$$H = \frac{1}{1000} = qN\mu \quad \begin{matrix} \text{can be estimated} \\ \text{separately} \end{matrix}$$
$$N_e = \frac{1}{1000} \cdot \frac{1}{q\mu} \quad \mu = 1.3 \times 10^{-8}$$
$$\approx 20,000$$

- In order for a Wright-Fisher population evolving at mutation-drift equilibrium to produce the same levels of heterozygosity as we observe in the human population, it would need to have a population size of 20,000. Therefore, the "effective size" of the human population is 20,000.

"Effective population size"

- Our first model of population history
- The model is obviously wrong, but useful because of its simplicity
 - If the population size changes over time (i.e. the population is not at equilibrium) there is still going to be some level of heterozygosity, and therefore an effective size
 - If the population is not actually a Wright-Fisher population, then the effective size will not be equal to the actual size, even if the actual size is constant.

typically, $N_e < N_c$

"Effective population size"

- Our first model of population history
- The model is obviously wrong, but useful because of its simplicity
 - If the population size changes over time (i.e. the population is not at equilibrium) there is still going to be some level of heterozygosity, and therefore an effective size
 - If the population is not actually a Wright-Fisher population, then the effective size will not be equal to the actual size, even if the actual size is constant.
typically, $N_e < N_c$
 - population structure can dramatically inflate the effective size