Chapter 3 – Genetic Drift & N_e

Genetic Drift

Effective Population Size

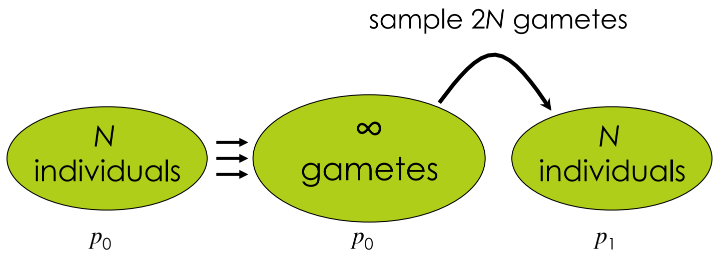
Relationship Between Drift and Inbreeding

The Coalescent

Genetic Drift

- What happens when we relax the assumption of infinite population size?
  - allele frequencies change due to random "genetic drift" (sampling)

The Wright-Fisher model



sample 2N gametes

N individuals

∞ gametes

N individuals

p_0

p_0

p_1

- random sampling of genotypes and random sampling of gametes essentially the same

http://psych.colorado.edu/~carey/hgss/hgssapplets/evolution/geneticdrift/GeneticDrift.html

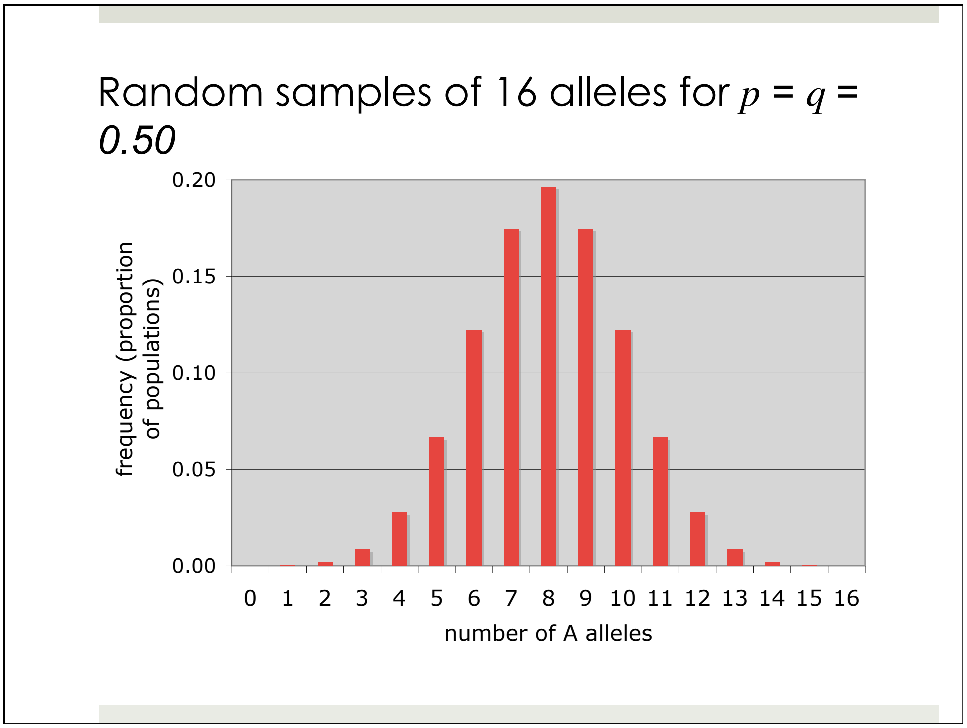
http://darwin.eeb.uconn.edu/simulations/drift.html

Genetic Drift

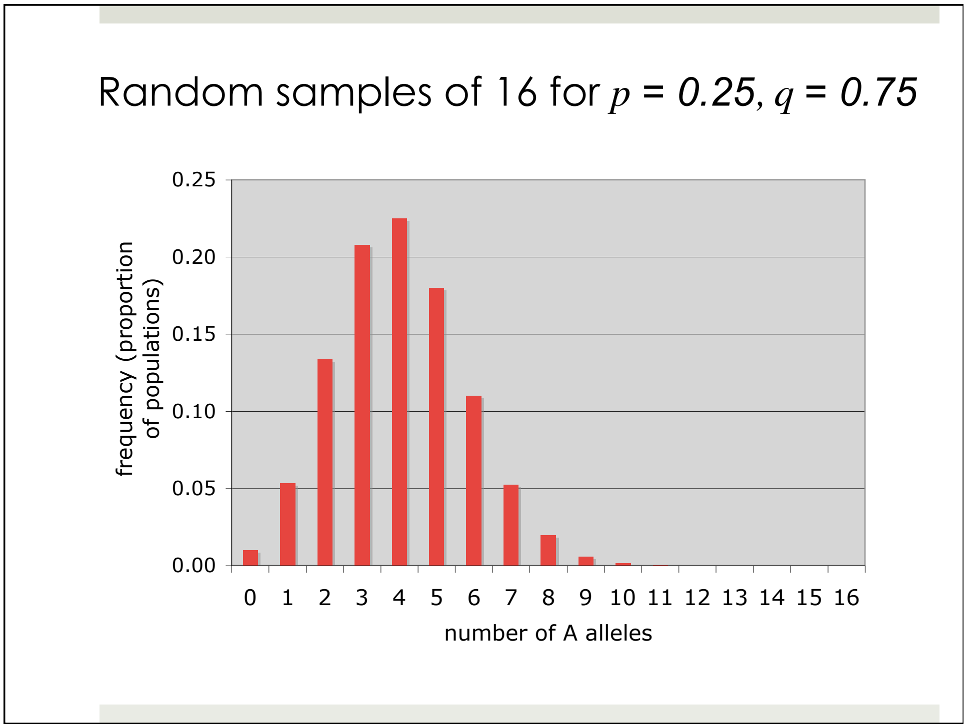
- Binomial probability

$$\Pr\left\{ j \text{ alleles of type } A \right\} =$$

$$\binom{2N}{j} p^j q^{2N-j} = \frac{(2N)!}{j!(2N-j)!} p^j q^{2N-j}$$



Random samples of 16 alleles for $p = q = 0.50$



Random samples of 16 for $p = 0.25$, $q = 0.75$

The Moran model

- derived for a haploid population
  - but results for 2*N* haploid individuals are essentially the same as for *N* diploid individuals when assuming neutrality
- randomly select an adult to produce 1 offspring, then randomly select another adult to die
- results in allele frequency increments of ±1

The Moran model

- transition probabilities

$$T_{ij} = \frac{i(2N-i)}{(2N)^2} = \frac{i}{2N}\frac{(2N-i)}{2N} = pq, \text{ for } j = i+1 \text{ or } j = i-1$$

$$T_{ii} = \frac{i^2 + (2N-i)^2}{(2N)^2} = p^2 + q^2$$

- repeat 2*N* times to approximate Wright-Fisher model

The Wright-Fisher model

- transition probabilities obtained directly from binomial probability distribution with p adjusted each generation

$$T_{ij} = \binom{2N}{j}\left(\frac{i}{2N}\right)^{j}\left(\frac{2N-i}{2N}\right)^{2N-j} = \frac{(2N)!}{j!(2N-j)!}p^{j}q^{2N-j}$$

- variance in allele frequency change for a single generation = $pq/2N$
  - (= $2pq/2N$ for 2N iterations of the Moran process)

Wright-Fisher transition probabilities

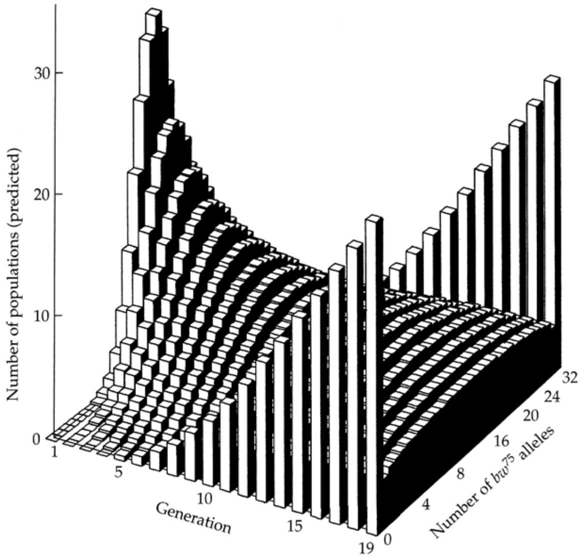
- binomial results in a matrix of transition probabilities that can be multiplied to predict the distribution of allele frequencies at any number of generations in the future

matrix for
2N = 8

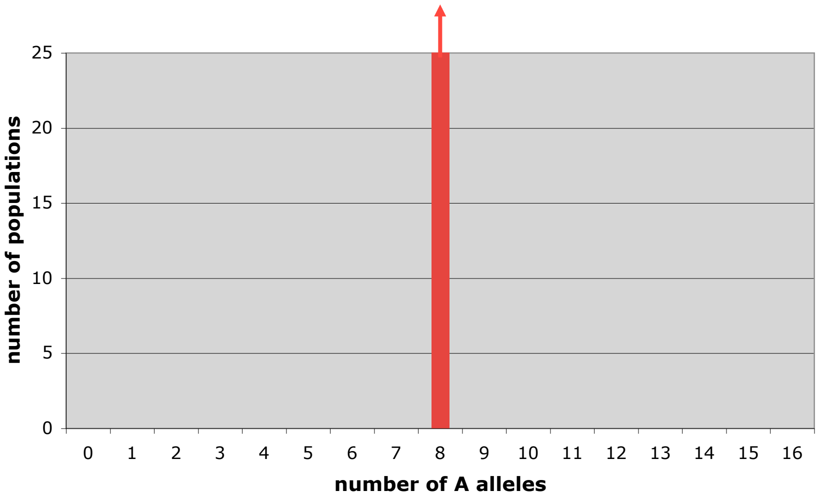
to j

	1	0	0	0	0	0	0	0	0
	0.344	0.393	0.196	0.056	0.010	0.001	0	0	0
	0.100	0.267	0.311	0.208	0.087	0.023	0.004	0	0
	0.023	0.112	0.235	0.282	0.211	0.101	0.030	0.005	0
from i	0.004	0.031	0.109	0.219	0.273	0.219	0.109	0.031	0.004
	0	0.005	0.030	0.101	0.211	0.282	0.235	0.112	0.023
	0	0	0.004	0.023	0.087	0.208	0.311	0.267	0.100
	0	0	0	0.001	0.010	0.056	0.196	0.393	0.344
	0	0	0	0	0	0	0	0	1

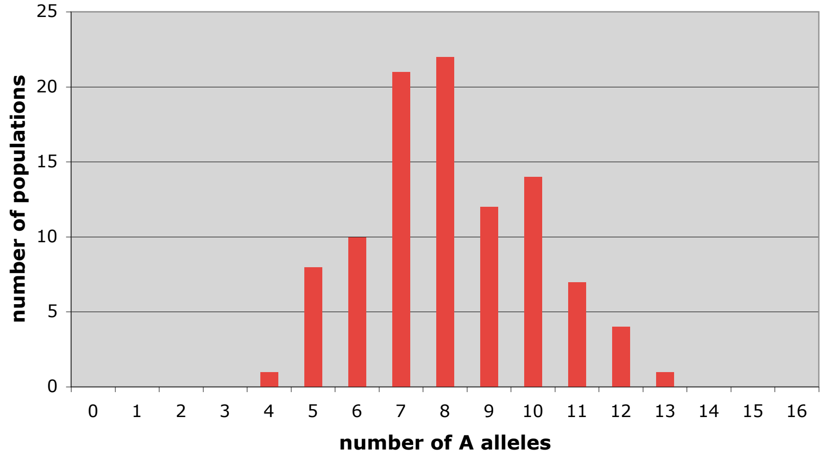
Predicted distributions of allele frequencies in replicate populations of N = 16



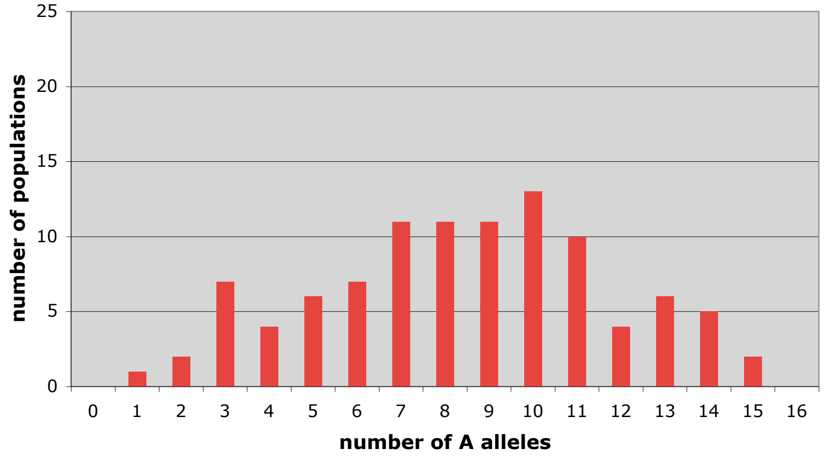
P = 100 populations

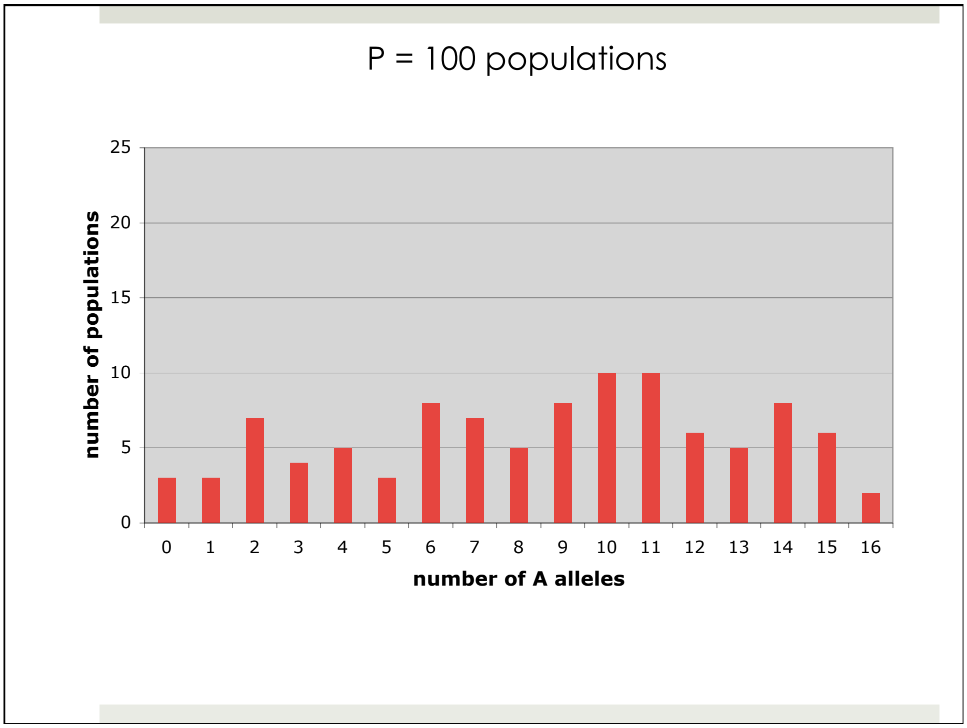


P = 100 populations

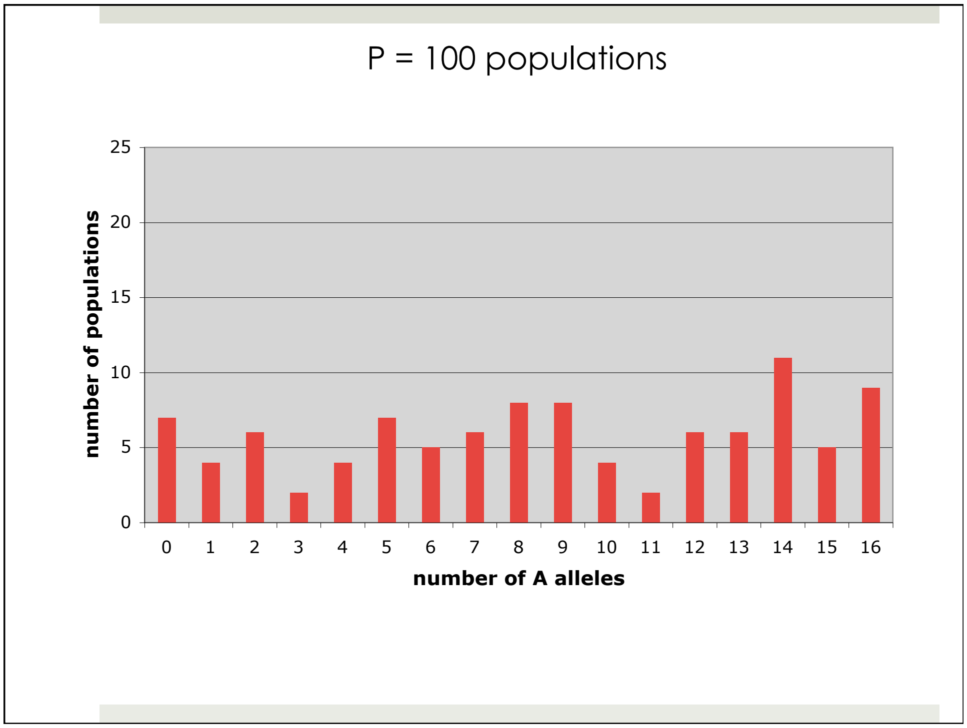


P = 100 populations

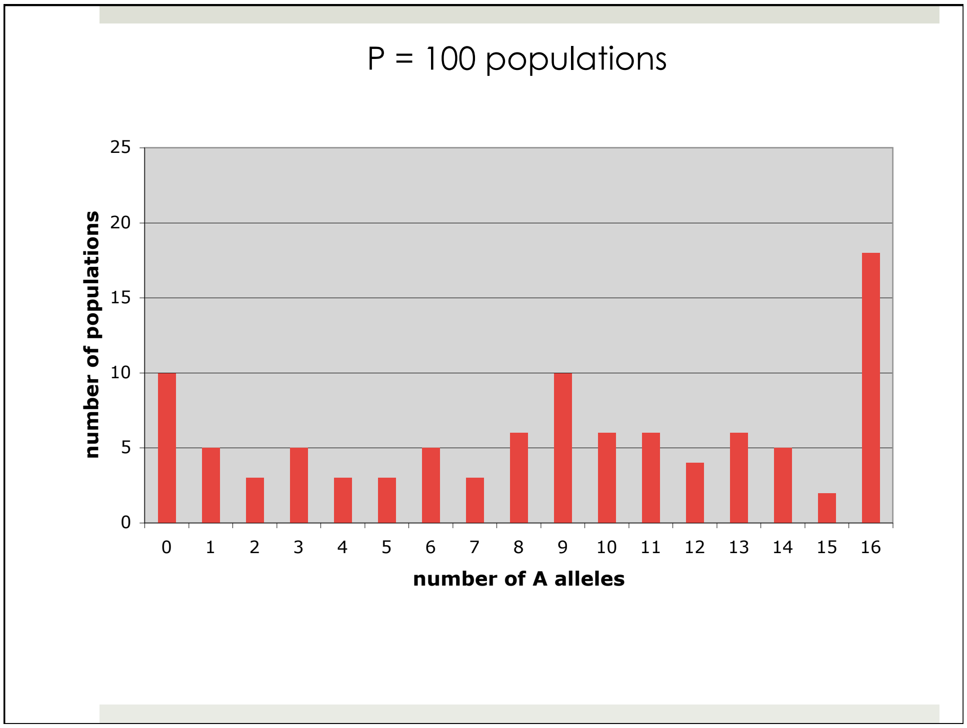




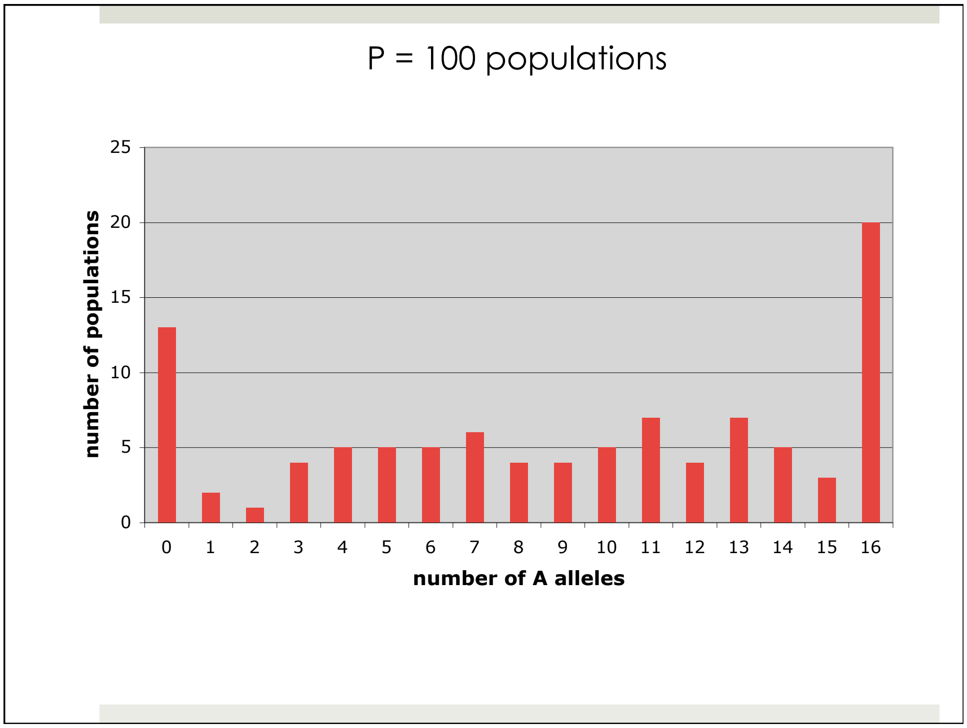
P = 100 populations



P = 100 populations

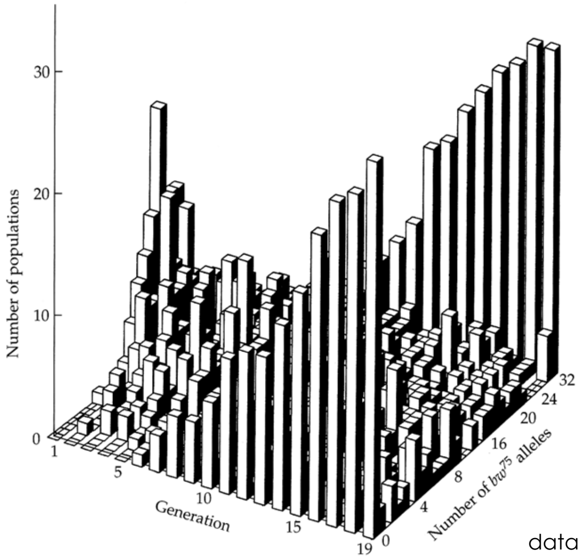


P = 100 populations



P = 100 populations

Random genetic drift in 107 *Drosophila* populations;
8 males, 8 females each generation



data from Buri (1956)

Questions

- How does time to fixation (or probability of fixation within a given number of generations) depend on population size?
- How does the probability of fixation for a given allele depend on its initial frequency?
- How likely is a new mutation to drift to fixation?