

Genetic Drift

1. Linkage disequilibrium (leftovers)
2. PopG program
3. Probability of fixation
4. Time to fixation

One-minute responses

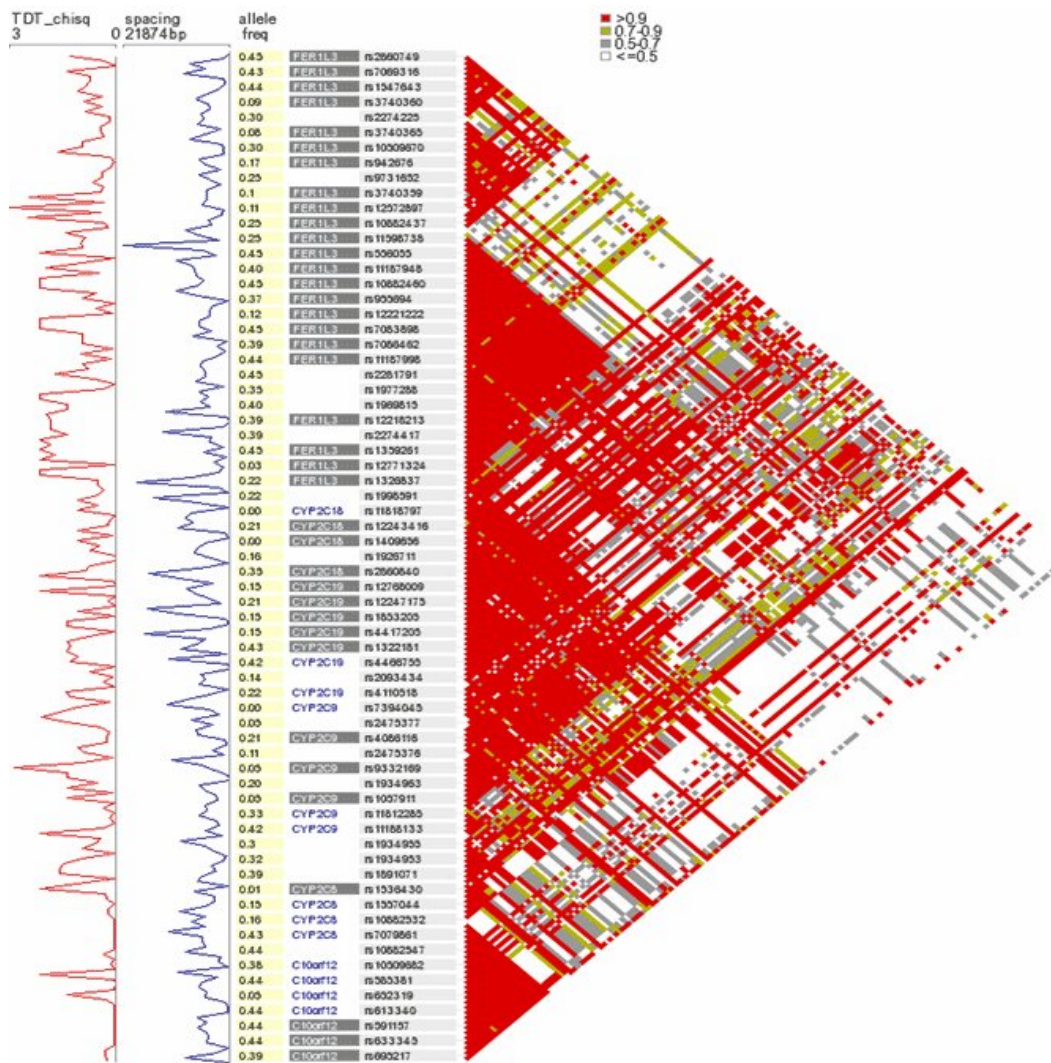
- Q: What statistical test is appropriate for testing for disequilibrium?
 - You can use χ^2 to test observed versus expected numbers of each chromosome
 - A significant χ^2 rejects linkage equilibrium
 - There is no rule for “this value of D is significant” as it depends on sample size and allele frequencies
- Q: I don't understand the disequilibrium coefficient D

Another look at D

- If there is no disequilibrium:
 - expected $p_{AB} = p_A \times p_B$
- If there is disequilibrium, measure the discrepancy:
 - observed $p_{AB} = \text{expected } p_{AB} + D$
- Example:
 - $p_A = 0.2, p_B = 0.3$
 - expected $p_{AB} = 0.2 \times 0.3 = 0.06$
 - observed $p_{AB} = 0.05$
 - $D = -0.01$

What forces can break down disequilibrium?

- Without recombination:
 - back mutation
 - repeated mutation
- With recombination and sexual reproduction:
 - chromosome segregation
 - recombination
 - back mutation
 - repeated mutation
- Disequilibrium areas are MUCH bigger and longer-lived without recombination



Blocks of LD around cytochrome P450 genes. Walton et al. Nature Genetics 2005.

LD graphs

- Graphs like this plot D' between sites
- In humans, conspicuous blocks of LD are usually seen
- Different ideas about their origin:
 - Recombination hotspots (at edges of blocks)
 - Population admixture followed by a few recombinations
 - Random chance
- I wish we could apply more statistics to this graph!

Genetic drift: A thought experiment

- The student on my left receives 1 red bean and 1 black bean
- They choose a bean at random and “reproduce” it
- They do this again
- They pass the new beans to the next student
- Repeat all the way around the table

What can we say about the color of the beans at the end?

A thought experiment

- Average over many trials is 50% red and 50% black
- Any single trial generally either 100% red or 100% black
- This is genetic drift in a tiny population

PopG program

This sort of experiment becomes cumbersome quickly with real beans. We will use computer simulation instead.

The program PopG is available from:

<http://evolution.gs.washington.edu/popgen/popg.html>

and will run on Windows, Mac and Linux/Unix machines. We will be using it for the next homework, so you will need a copy. Anyone who can't get one or doesn't have a computer to run it on can contact me for help.

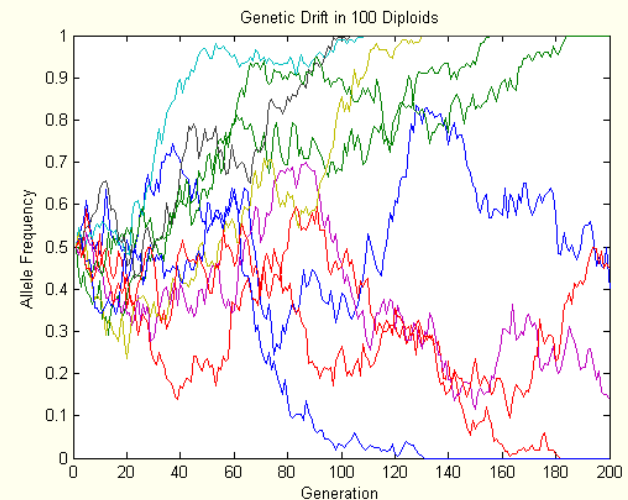
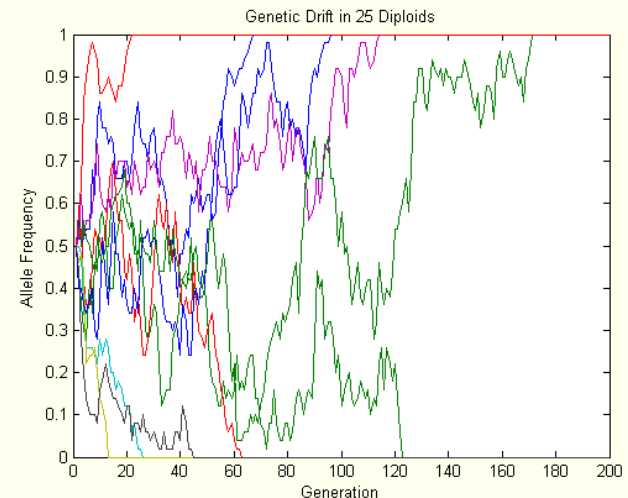
PopG program

PopG demonstration goes here!

Genetic Drift

Genetic drift: random change in allele frequencies due to finite population size

- Often called “sampling error” but it is not an experimental error
- The process of producing offspring IS a random sampling process
- The smaller the population, the more strongly it is affected by drift



Chance of fixation

What is the chance that drift will fix A as opposed to a ?

- Drift is a random walk that continues until it touches 1 or 0.
- The closer it is to 1, the more likely it is to reach 1 before it reaches 0.
- Probability of fixing A : p_A
- Probability of fixing a : p_a
- With no other forces, eventually one or the other must fix.

Chance of fixation

Imagine extending this to a population in which everyone started with two unique alleles.

N individuals would have 2N alleles.

After a few generations, some alleles will be lost.

Eventually all but 1 will be lost.

The chance that any given allele is a “winner” must be:

$$\frac{1}{2N}$$

Practice problem

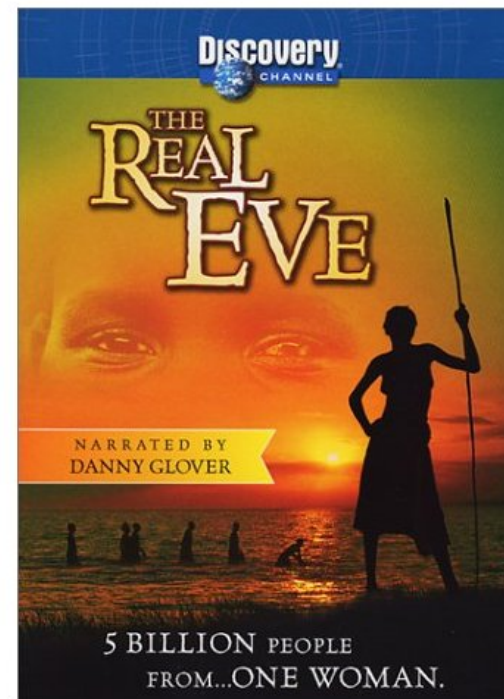
- Suppose you have a new, neutral mutant allele at your *ADH* locus.
- What is the chance that this allele will become the new “wild type”?
- Useful facts:
 - World population: 7 billion (7×10^9)
 - Don’t forget that you are diploid!
 - Assume the population size remains constant.

Practice problem

- World population contains 14×10^9 copies of this allele
- Your allele therefore has a chance of $1/14 \times 10^9$ to be the winner
- This is about 7×10^{-11} . You'd have better luck playing the lottery.
- If we ask if ANY gene of yours will win, it's about 0.00000035

Mitochondrial Eve

- Popular press was surprised that “mitochondrial Eve” existed
- Statisticians were surprised by:
 - Where she lived
 - When she lived
- For a non-recombining locus there MUST be a single common ancestor if you go back far enough



Time to fixation

How long will the “winner” take to win on average? This requires some assumptions:

- Diploid organism (two gene copies per individual)
- Random mating
- Non-overlapping generations
- Each individual has an equal chance to contribute to the next generation
- Constant population size

(Humans don't quite fit here....)

Time to fixation

With these assumptions:

- Average time to fixation is $2N$ generations.
- Variance is also $2N$ (uncertainty is high!)
- At any point, the chance that an allele will be the eventual winner is equal to its frequency
- mtDNA has an average fixation time of $N/2$ generations—why the difference?

Practice problem

- If one of your alleles will become the new wild type, on average how long will it take (in years)?
- Assume 20 years per generation
- World population: 7 billion (7×10^9) – assume it's constant

Practice problem

- If one of your alleles will become the new wild type, on average how long will it take (in years)?
- Assume 20 years per generation
- World population: 7 billion (7×10^9)
- 14×10^9 generations
- 2.8×10^{11} years
- Estimated future lifespan of Earth: 5×10^9 years

Drift in humans?

- At our current population size, humans can't fix rare alleles by drift in feasible time
- Things were different during most of our evolution, with an approximate population size around 5,000-50,000
- Humans **do** lose rare alleles by drift
- Fixation by drift is still significant for isolated human populations

South Sentinel Island

- Island in the Bay of Bengal
- Current population estimated at 50-400 adults
- Isolated for 60,000 years
- Violent rejection of outside contact
- Rate of genetic drift is probably VERY high in this population



Fixation of variants by drift?

- How fast does a population fix new variants by drift?
- Each generation:
 - Mutation rate μ per gene copy
 - $2N$ gene copies (diploid)
 - New mutations produced: $2N\mu$
- Exactly one copy out of $2N$ will be the long-term winner
- If the winner has a mutation, it will be fixed
- Rate of fixation of new mutants is $2N\mu/2N = \mu$

Rate of change in a species due to drift

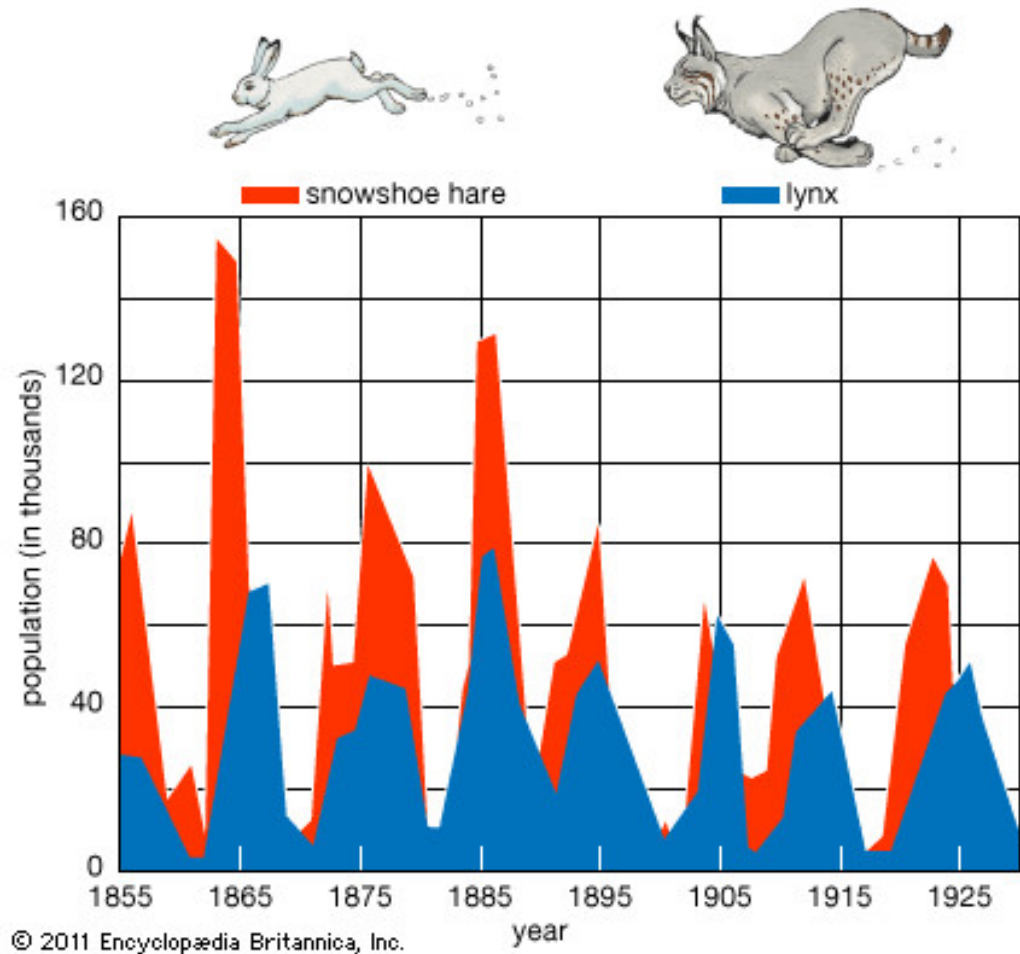
- Small and large populations will fix the same number of mutations per unit time
- Small population:
 - Few mutations arise
 - Those that arise fairly often fix
- Large population:
 - Many mutations arise
 - Each one has a very low chance to fix

Rate of change in a species due to drift

- Standing variation (mutations waiting to fix or be lost) is larger in a large population
- Divergence rate (rate at which two populations become genetically different) does not depend on population size
- (This isn't quite true due to non-neutral mutations, but it's close)

Effective population size

- Idealized population: constant size, random mating, no migration
- What about real populations?
- We can define an effective population size N_e which compares a real population to the idealized population



Effective size of cycling population

If a population grows and shrinks, N_e is the harmonic mean of the various sizes. For two generations:

$$\frac{1}{N_e} = \left(\frac{1}{N_1} + \frac{1}{N_2} \right) / 2$$

The generalized solution for k generations:

$$\frac{1}{N_e} = \left(\sum \frac{1}{N_i} \right) / k$$

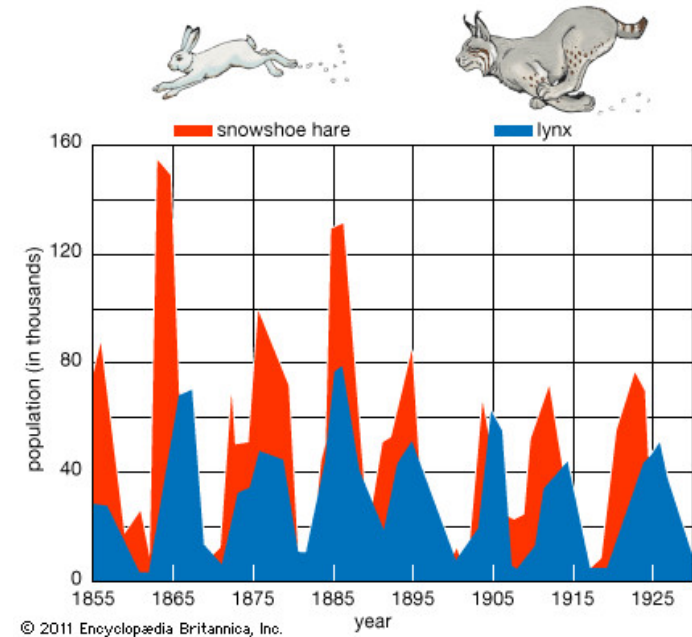
- Harmonic mean is close to minimum value
- Genetic diversity is lost when the population is small and is not rapidly regained when it is larger

Effective population size

Consider a rabbit population which cycles between size 40,000 and size 100,000:

$$\frac{1}{N_e} = (1/40000 + 1/100000)/2$$

$$N_e = 57,143$$



Effective population size

- Long-term human effective population size: 5,000 to 50,000
- We do not yet have the genetic diversity of a species of size 7 billion
- N_e is usually lower than N , often much lower
- Unusual mating systems can make N_e slightly higher than N but this is rare

Effective population size

- Red drum are large fish of the Gulf of Mexico
- Effective size 1000 times lower than census size
- This species has the numbers of a big population but the genetic drift of a small one
- Likely explanation is very unequal reproductive success



Effective population size

- Red drum spawn in very specific estuary environments
- A few lucky clutches have thousands of survivors; most have none
- Allele frequencies change substantially from one generation to the next, reflecting the few lucky individuals

Other non-ideal populations

- Unequal males and females (domestic cattle)
- Overlapping generations (redwoods)
- Highly nonrandom mating success (racehorses)

All of these tend to reduce N_e compared to N

One-minute responses

- Tear off a half-sheet of paper
- Write one line about the lecture:
 - Was anything unclear?
 - Did anything work particularly well?
 - What could be better?
- Leave at the back on your way out