

# Overview

---

- Catch-up from Monday:
- The exponential approximation in more detail
  - Sample TMRCA and population TMRCA
- Genetic drift as a forward process: the random walk
- PopG simulator
- Effective population size:
  - Basic definition
  - No self-fertilization
  - Separate sexes
- Parameters of the drift process

## From the one-minute responses

---

- Define terms and label graphs
- Show how we get the exponential approximation for the coalescent
- Be more precise about the role of mutation
- More class discussion and interaction (good point!)

## Working definitions

---

- Gene copy: one instance of a given sequence (it doesn't actually have to be a "gene")
  - Assume for the moment that if it changes several times we can always see all mutations—they don't overwrite or interfere
  - MRCA of two gene copies is most recent gene copy ancestral to both
  - We ignore recombination within gene copies for now
- Coalescence: looking back in time, the point at which two genetic lineages reach their common ancestor
- Coalescent: the expected distribution of coalescence times

## Derive the exponential approximation

---

- In discrete time:

- $Prob_c(t) = (\frac{1}{2N})(1 - \frac{1}{2N})^{t-1}$

- Continuous time approximation:

- $Prob_c(t) = \frac{1}{2N}e^{-\frac{t-1}{2N}}$

- Mean is  $2N$ , standard deviation is  $2N$  (big!)

## TMRCA – error in previous slides

---

- I claimed TMRCA was the same with same  $\theta$ 
  - NOT true in years or generations! Bigger  $N$  leads to bigger TMRCA
  - "Mutational time" (expected number of mutations) back to MRCA is the same

# Goals of studying genetic drift

---

- Predict:
  - Diversity of populations
  - Rate at which species diverge due to drift
  - Effect of demographics on diversity and divergence
- On the horizon: interactions between drift and natural selection

# Genetic drift

---

PopG demo goes here

## Effective population size $N_e$

---

- Informally, the size of a Wright-Fisher population with the same intensity of drift as the given population
- Can be calculated for a variety of situations:
  - Non-reproductive individuals
  - No selfing
  - Two sexes
  - Cycling population size (on Friday)
  - Overlapping generations, unequal reproductive success (difficult!)

## Flavors of $N_e$

---

- Possible to define  $N_e$  as the size of a Wright-Fisher population with the same:
  - Expected proportion of homozygotes
  - Expected rate of change in allele frequencies due to drift
- We are computing the first, but they are generally equal or close

## No self-fertilization

---

- Some hermaphrodites do not self-fertilize (especially plants)
- This means that 2 gene copies currently in the same individual did not coalesce last generation
- This increases  $N_e$  by approximately 1/2 an individual—generally trivial

## Separate sexes

---

- $N_e \approx 4N_f N_m / (N_f + N_m)$
- If  $N_f = N_m$  this is  $2N$
- As  $N_m$  decreases it approaches  $4N_m$
- (Felsenstein book says  $2N_m$  but he confirms this is wrong)
- The rare sex exerts a disproportionate influence on the population trajectory

## Derivation hints

---

- $1/2$  chance that two copies this generation were in same sex individual last generation
- Chance of coalescing in 1 generation is  $1/(2N_f)$  or  $1/(2N_m)$  at  $1/4$  each
- Overall chance therefore  $1/(8N_f) + 1/(8N_m)$
- Set  $1/2N_e$  equal to this and solve for  $N_e$

## Haplodiploids–Practice problem

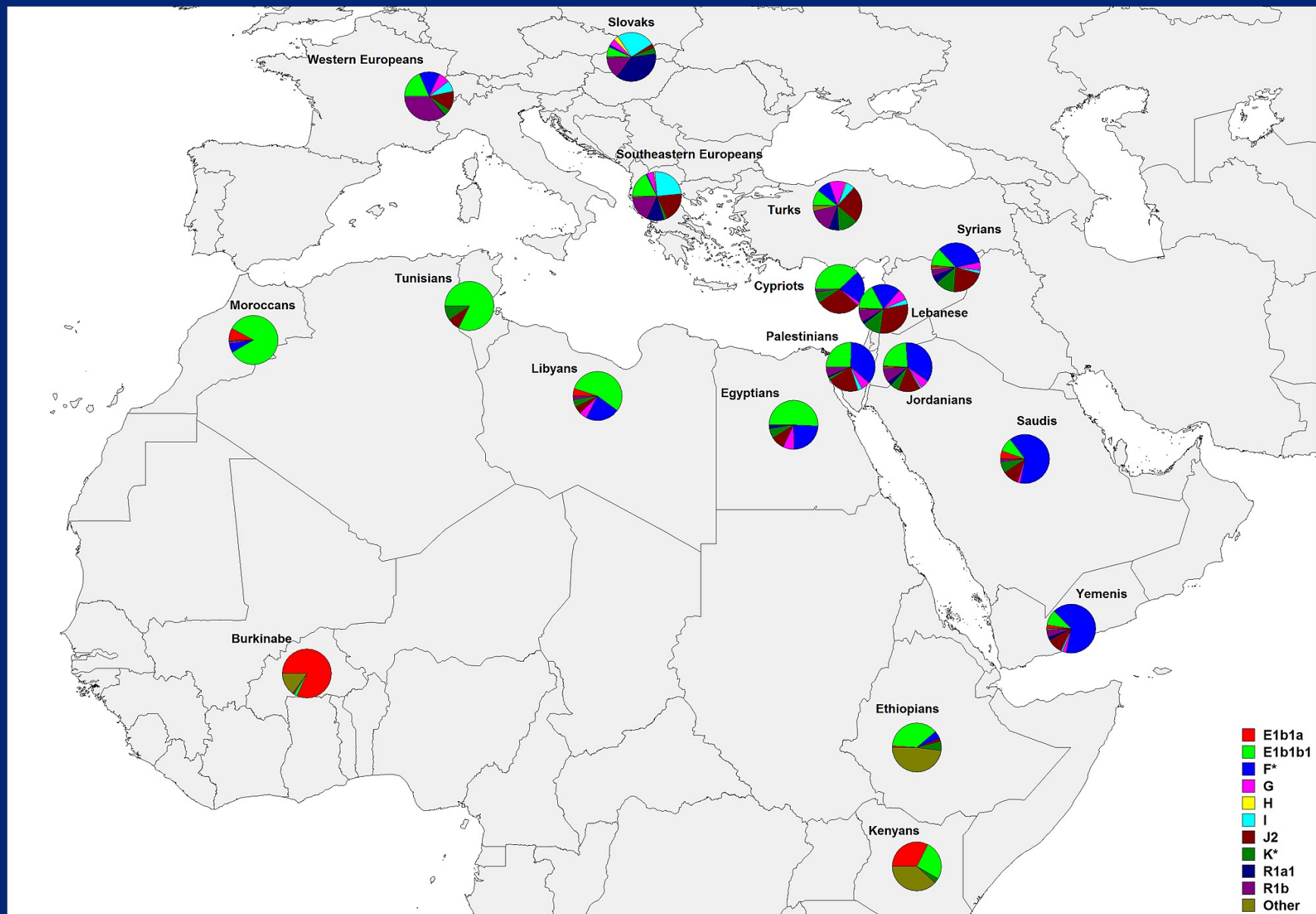
---

- Can we work out the same equation for haplodiploids?
  - Females have 2 copies of their genome
  - Males have only 1 copy which they give to all offspring
  - Obviously the effective size must be lower than  $2N$ , but how much lower?
- Try working in pairs

## Non-random reproductive success

---

- Human females have low variance in reproductive success (Wikipedia record: 69)
- Human males have much more variance (Wikipedia record: 860)
- Genetic drift is stronger in male-only than female-only DNA
- Could explain why Y chromosome shows more population-specific traits than mtDNA
- Alternative: sex-specific migration/dispersal



Y-chromosome variation in Europe and North Africa

## Why $N_e$ matters

---

- 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



## Why $N_e$ matters

---

- 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

## Who cares about $N_e$ ?

---

- Conservation biologists
  - How much diversity will a given population size/structure lose?
  - How can we minimize losses?
- Epidemiologists
  - What are a pathogen's likely resources for dealing with a new treatment?
  - How can we reduce them?

## Who cares about $N_e$ ?

---

- Animal breeders
  - How much trouble will we get in if we mainly breed from the few best males?
- Phylogeneticists
  - How does population size affect species divergence? (More on this later)

## Describing the drift process

---

Here are several related, but distinct, questions:

- What is the probability that a specific new mutant will eventually fix?
- How many mutants fix per generation?
  - How many new mutants destined to fix arise each generation?
  - How fast do two species diverge by drift?
- How long does it take a mutant to fix on average?
- How much variation will be present in a population on average?

## Describing the drift process

---

Taking them one at a time:

- What is the probability that a specific new mutant will eventually fix?  
 $\frac{1}{2N_e}$
- How many mutants fix per generation?  $\frac{2N_e\mu}{2N_e} = \mu$
- How long does it take a mutant to fix, on average? Approximately  $4N_e$  generations
- How much variation will be present in a population on average?

## Fraction of homozygotes

---

- Counting alleles is not a good way to quantify variation
  - Too sensitive to very rare alleles
- Measure variation as proportion of homozygotes—the fewer homozygotes, the more variation
  - Call the proportion of homozygotes  $F$
  - With two equally frequent alleles,  $F = 0.5$

## Fraction of homozygotes

---

- In cases with mutation and drift, an approximate formula is:

$$F \approx \frac{1}{1+4N_e\mu}$$

- This approximation assumes that every mutation is to a new allele. It is quite accurate in practice even when that's not true, as long as there are a decent number of different alleles possible.

## Fraction of homozygotes

---

$$F \approx \frac{1}{1+4N_e\mu}$$

Intuitive results of this equation:

- If the population is large, there will be fewer homozygotes (more diversity)
- If the mutation rate is large, there will be fewer homozygotes (more diversity)

(Always ask yourself—does this equation predict results that are in the right general direction?)

## Fraction of homozygotes—Practice problem

---

(Fictional problem inspired by real data of Potts et al.)

$$F \approx \frac{1}{1+4N_e\mu}$$

- We measure heterozygosity at one gene in the mouse MHC as 92%
- (Population: restaurant mice in Miami)
- Mutation rate (based on rat/mouse comparison) is around average for rodents:  $10^{-6}$  per gene per generation
- How many mice does this imply, if the MHC were non-selected?
- (You'll actually calculate  $N_e$ —that's okay)

## Fraction of homozygotes—Practice problem

---

(Fictional problem inspired by real data of Potts et al.)

- $F \approx \frac{1}{1+4N\mu}$
- $0.08 \approx \frac{1}{1+4Nx10^{-6}}$
- $N = 2,875,000$  mice
- That's probably too many mice. What might explain this?

# Summary

---

- Wright-Fisher model gives simple predictions for many aspects of the drift process:
  - Chance for a mutation to fix
  - Time it takes to fix
  - Diversity within a population
  - Divergence between populations
- These can often be adapted to a non-Wright-Fisher situation via the *effective population size*  $N_e$

## Next week

---

- What if population size is not constant?
  - Cycling population size
  - Exponential growth or shrinkage
  - Bottlenecks

## One-minute response

---

- Please:
  - Tear off a slip of paper
  - Give me one comment or question on something that worked, didn't work, needs elaboration, etc.