

Course mechanics

Evolutionary Genetics (GENOME 562)

MWF 11:30-12:20 am, S110 Foegen

Instructor: Dr. Mary Kuhner

Office: S420C Foegen

Hours: TBA, or by appointment

Phone: (206) 543-8751

Email: mkkuhner@uw.edu

Web: <http://courses.washington.edu/gs562/>

Lecture slides will be on the web page, usually by the morning of the lecture.

Textbook

- No textbook for this course
- Two potentially useful books:
 - Felsenstein, *Theoretical Evolutionary Genetics*, free on the web at <http://evolution.gs.washington.edu/pgbook/pgbook.html>
 - Nielsen and Slatkin, *An Introduction to Population Genetics*
- The syllabus lists possible readings from these books (mainly Felsenstein) for students who want additional readings; they are not required
- A valuable resource is the PopG simulator program:
<http://evolution.gs.washington.edu/popgen/popg.html>

Homework

- 8 assignments
- Handed out Monday, due following Monday
- Worth $1/3$ of course grade
- You may work together, but be sure you individually know how to solve the problems
- Electronic homework is fine, but NOT cell phone photos of handwritten homework!

Exams

- Midterm and non-comprehensive final count 1/3 each
- Closed book and notes; equation sheet will be provided
- Calculators strongly encouraged
- Homework is the best exam prep; Felsenstein text also has useful practice problems

Logistics

- Office Hours
 - Wednesday after class? Friday before class?
 - Additional times by appointment
- Black and white lecture notes?
- Programming level?

Course Plan

- Topics:
 - Coalescent theory and genetic drift
 - Natural selection
 - Linkage and recombination
 - Population subdivision and migration
 - Quantitative genetics
 - Sexual reproduction and its consequences
 - Chromosome and genome evolution
 - Phylogenies

Today's lecture

- A “default” population: Wright-Fisher model
- Times back to the common ancestor: the coalescent

- Tree of descent for gene copies
- Implications of the coalescent

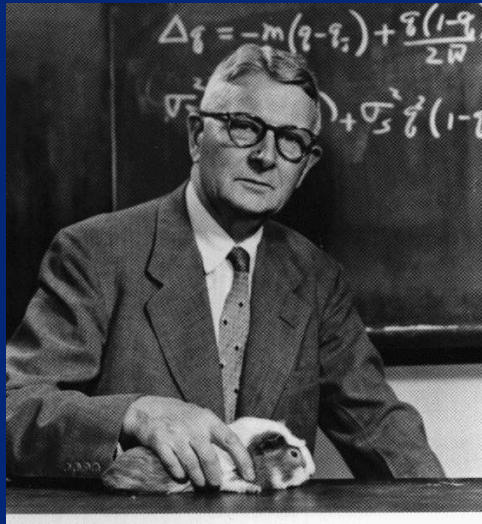
Forward vs. backward

- Classical population genetics looks forward:
 - Given these starting conditions, what might happen?
 - Useful in predicting results of experiments
- Coalescent theory looks backwards:
 - Given this outcome, how could it have arisen?
 - Useful in analyzing existing populations

Wright-Fisher population model

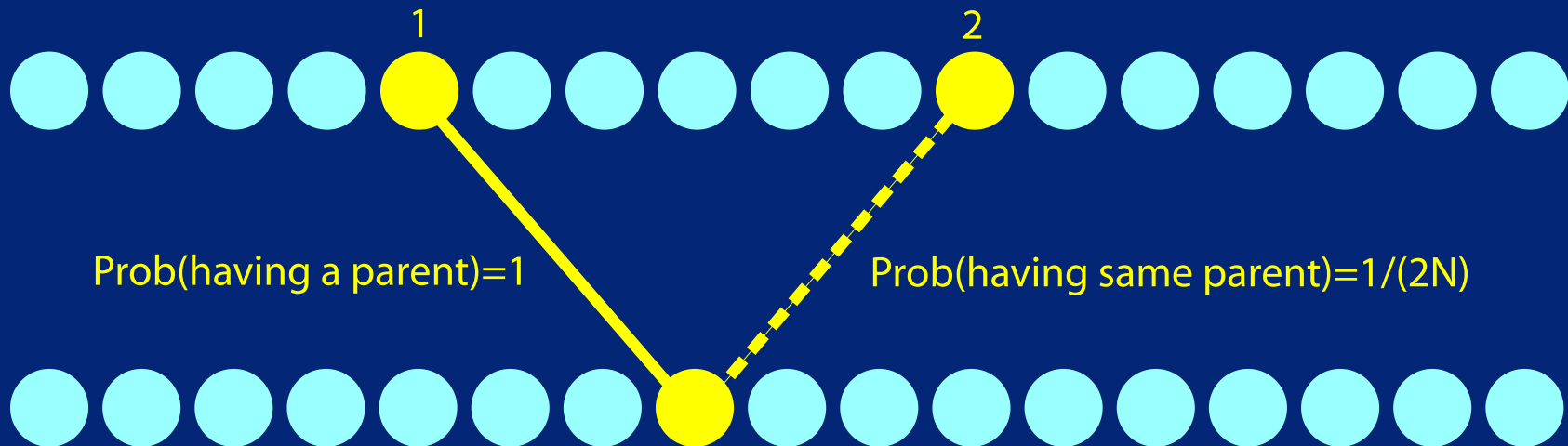
- The Platonic ideal of a population:
 - No natural selection
 - Random mating
 - Constant population size
 - Random reproductive success

Wright-Fisher population model



Sewall Wright showed that the probability that 2 gene copies come from the same gene copy in the preceding generation is

$$\text{Prob (two genes share a parent)} = \frac{1}{2N}$$



Wright-Fisher population model

- *Note that we are assuming diploids, hence number of copies = $2N$*
- For two chosen copies:
 - Distribution of time back to the most recent common ancestor – geometric distribution
 - Mean generations before the most recent common ancestor – mean of the geometric $(1-p)/p = (1-1/2N)/(1/2N) = 2N - 1$
 - MRCA therefore at $2N$ generations
- Painful to consider more than 2 copies due to combinatorics

The coalescent approximation

- In the limit as N gets big:
 - The chance that more than 2 copies reach a common ancestor at the same moment becomes very low
 - The geometric (discrete time) approaches an exponential (continuous time)
- This is JFC Kingman's *coalescent*
- Time is often rescaled so that $2N$ generations = 1 time unit

The coalescent approximation

- Number of ways to choose 2 of k lineages: $k(k-1)/2$
- Time until the first 2 of k copies reaches a common ancestor: exponential with expectation $2N/[(k(k-1))/2]$
- Then there are $k-1$ copies, and it repeats

Draw the coalescent tree

Limits of the approximation

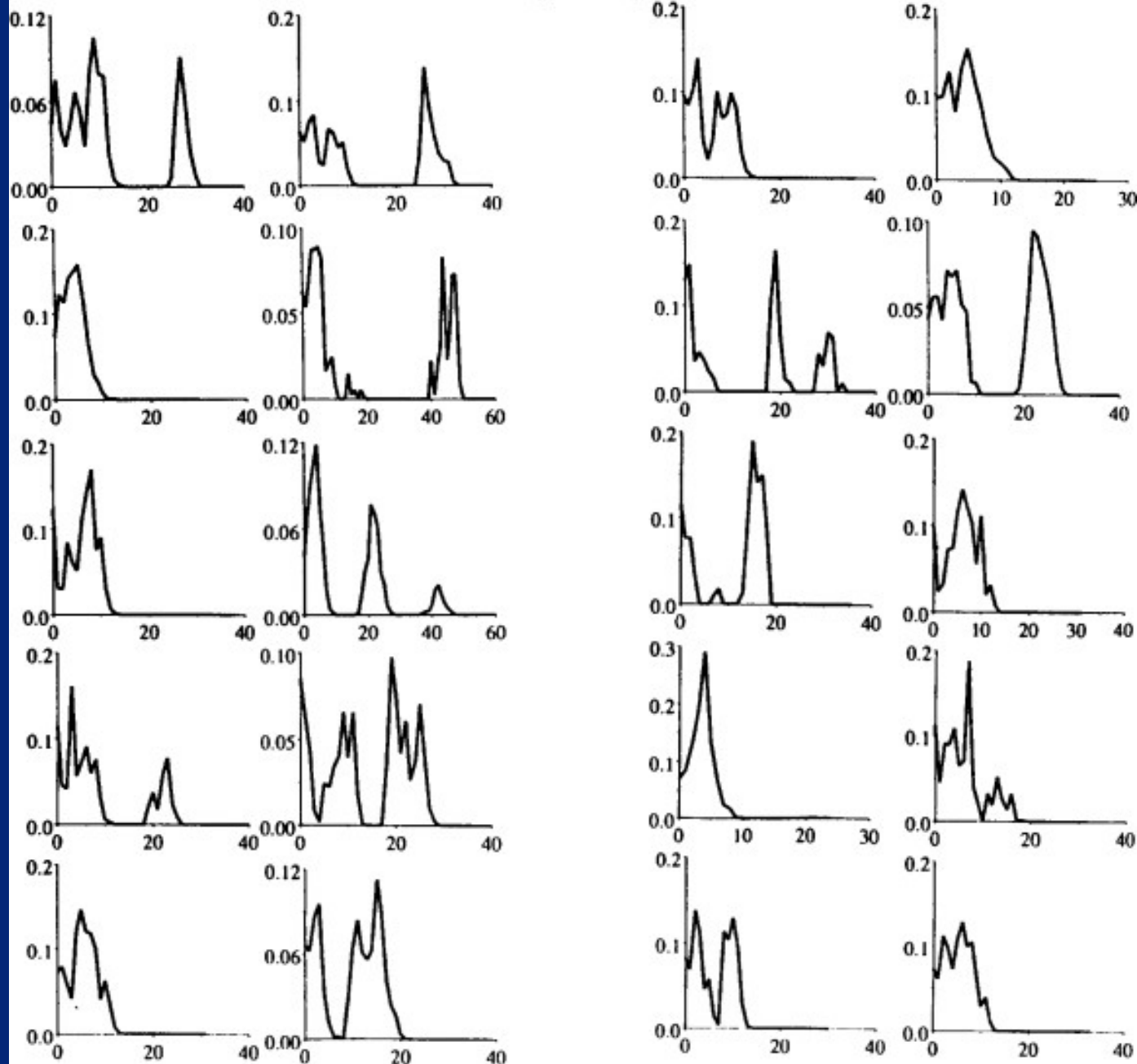
- N should not be tiny
- k should be much smaller than N
- Why? With very small N or large k :
 - Two coalescences at the same time can't be disregarded
 - Complications like two sexes, finite reproductive capability, discrete generations, etc. can't be disregarded
- Not appropriate math for sib-mating mouse lines

A cautionary tale

- In a constant size population, expected time to the common ancestor of two copies is exponential with mean $2N$
- Experiment:
 - Examine many pairs of copies of a given locus from a population
 - Estimate TMRCA based on mutational differences
 - Graph the distribution of TMRCA
- Results on human data:
 - Not exponential: has a pronounced peak in the past
 - This was interpreted as population growth
- Results with simulated data: Slatkin and Hudson (1991)

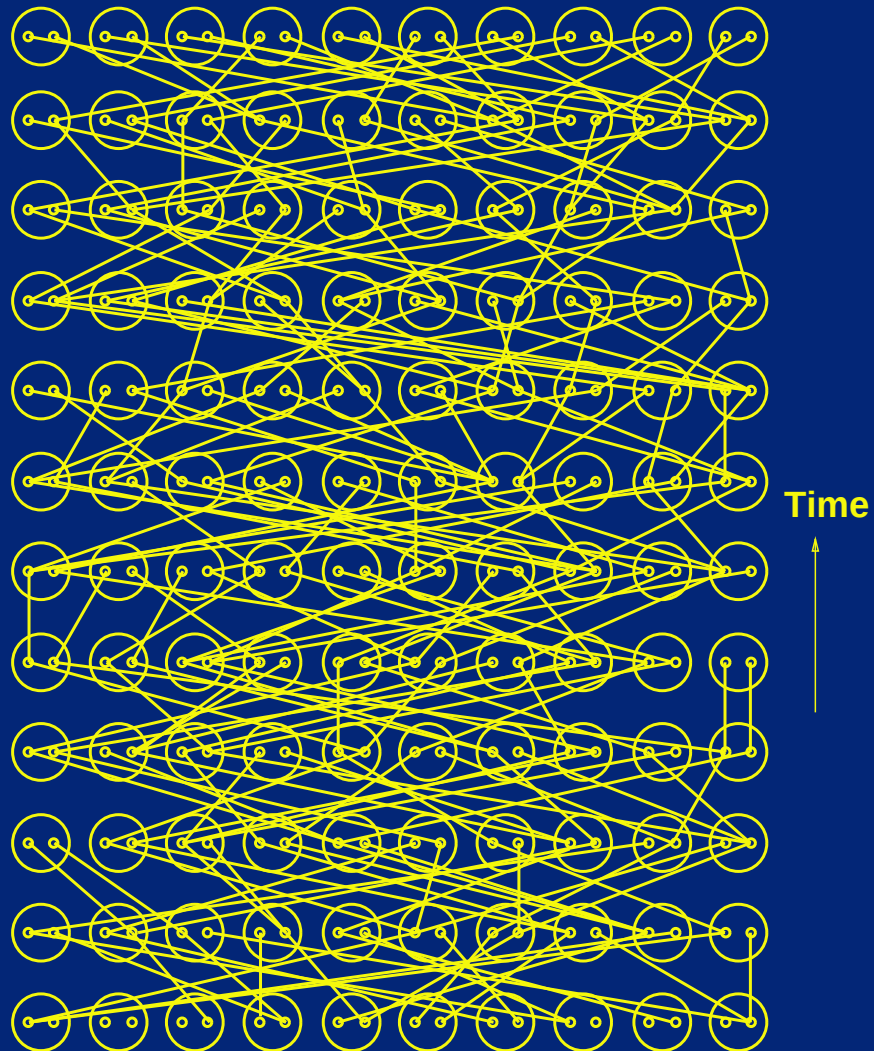
Pairwise Analysis of Sequence Data

33



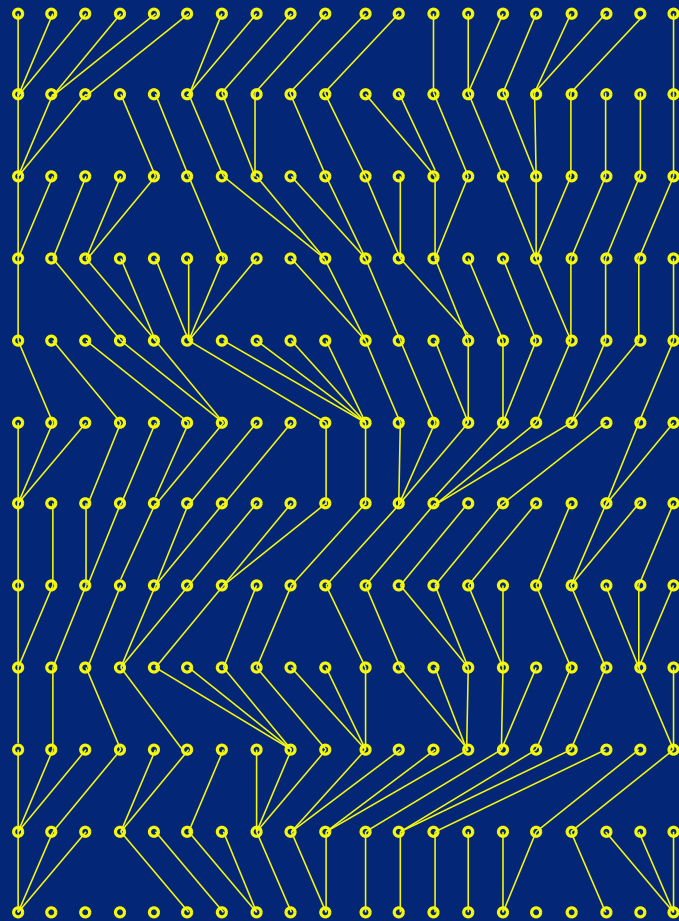
The coalescent is a tree

A random-mating population

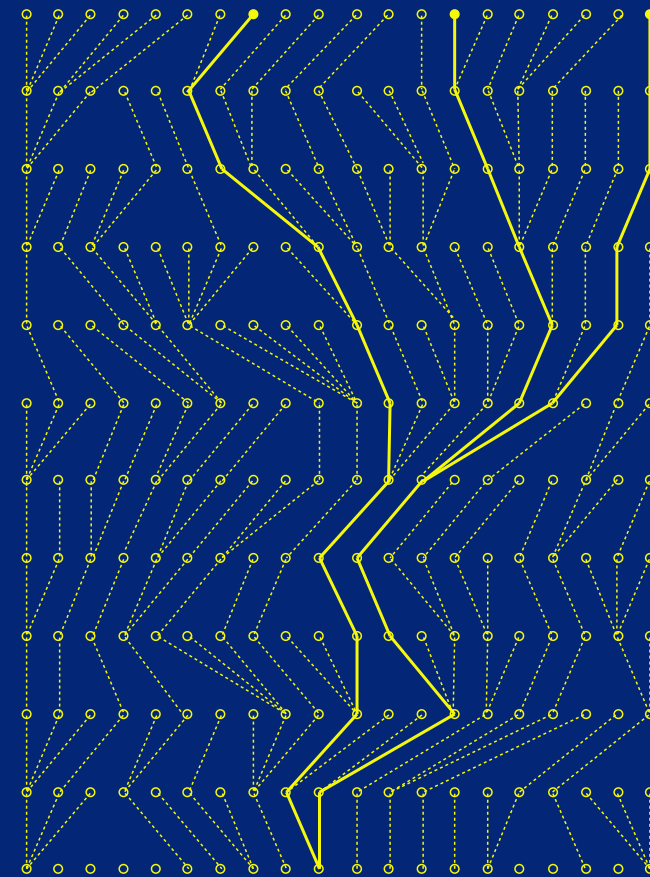


The coalescent is a tree

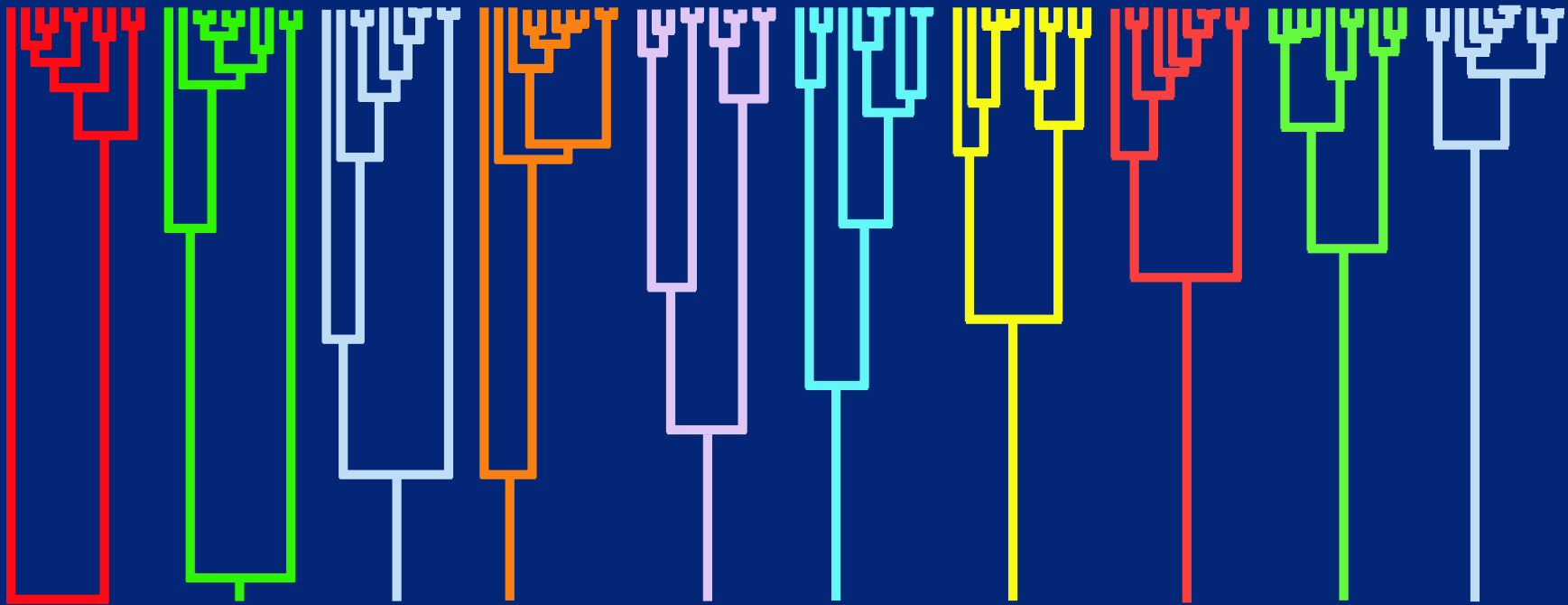
Genealogy of gene copies, after reordering the copies



Genealogy of a small sample of genes from the population

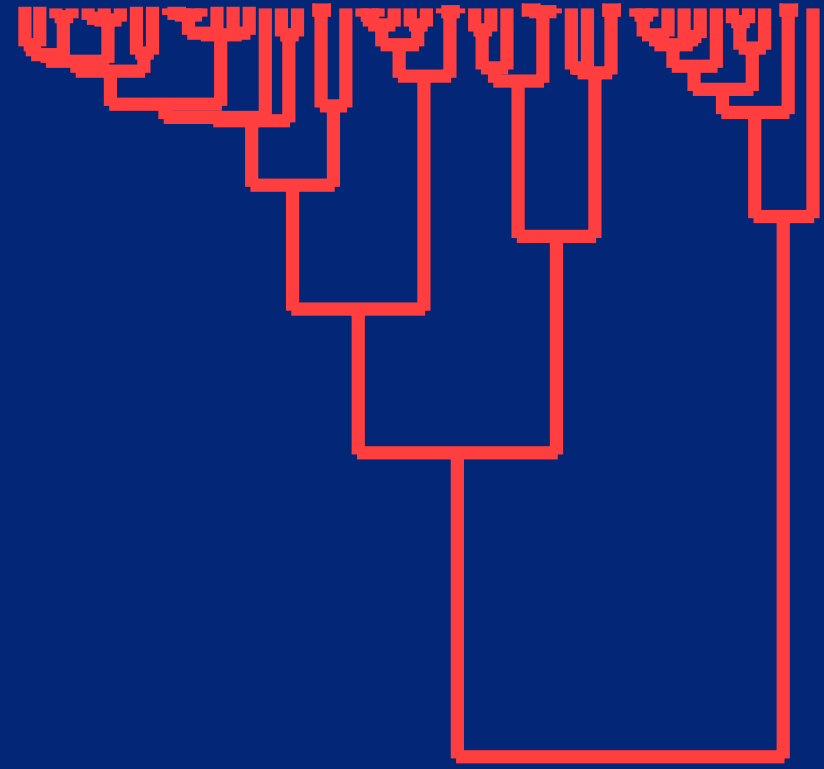
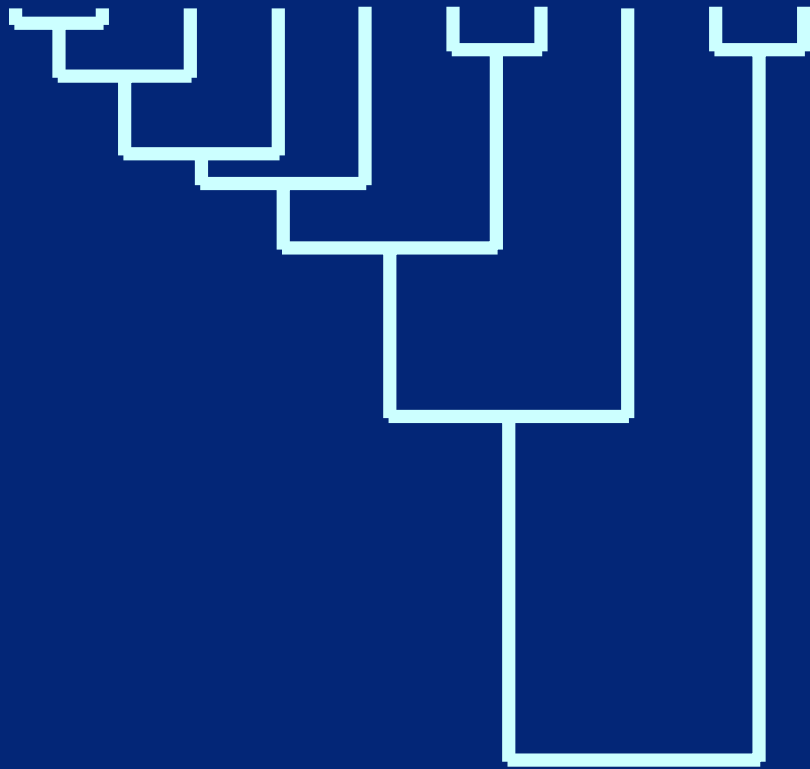


Variability of the coalescent



10 coalescent trees simulated with $N=10,000$

Behavior of additional samples



Sample MRCA vs. population MRCA

- TMRCA of a sample: $\sum_{k=2}^n \frac{2N}{k(k-1)/2}$
- TMRCA of the population: sum of the above for infinite n
- Approximately $4N$
- With even a small sample, the bottom population split is likely already observed

The parameter θ

- We want to know about time, but generally know only mutational distance
- These are related via the mutation rate μ
- $\theta = 4N\mu$ is a fundamental parameter
- Why 4?
 - One factor of 2 for diploidy
 - One factor of 2 because mutations accumulate on both lines

The parameter θ

- Two populations with the same θ :
 - Same amount of genetic variation
 - Same TMRCA
- *Remember this is a coalescent approximation and does not apply to very tiny populations*

The coalescent describes “genetic drift”

- Classical formulation of drift:
 - Force that causes random variation in allele frequencies
 - Strength varies inversely with population size
- Coalescent view:
 - Population size determines mean time back to ancestors
 - Reduced number of ancestors drives variation in allele frequencies

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.