

# Evolutionary trees and population genetics: a family reunion

9 October 2009.

Joe Felsenstein

500th anniversary (or something) of the University of Chicago

# The modern synthesis, part 1



R. A. Fisher



J. B. S. Haldane



Sewall Wright

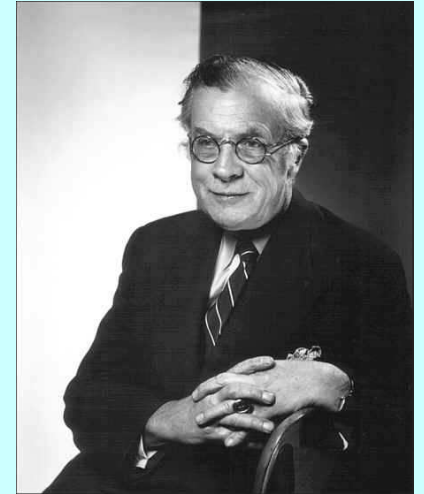
## The modern synthesis, part 2



Ernst Mayr



George Gaylord Simpson



Sir Julian Huxley

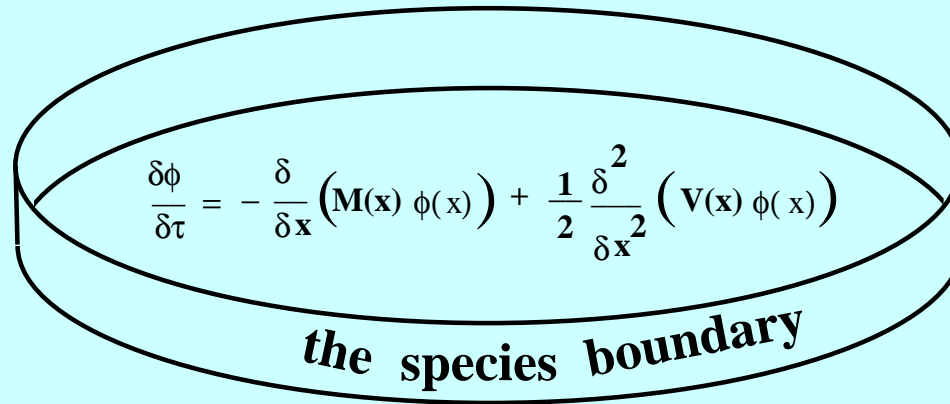


G. Ledyard Stebbins



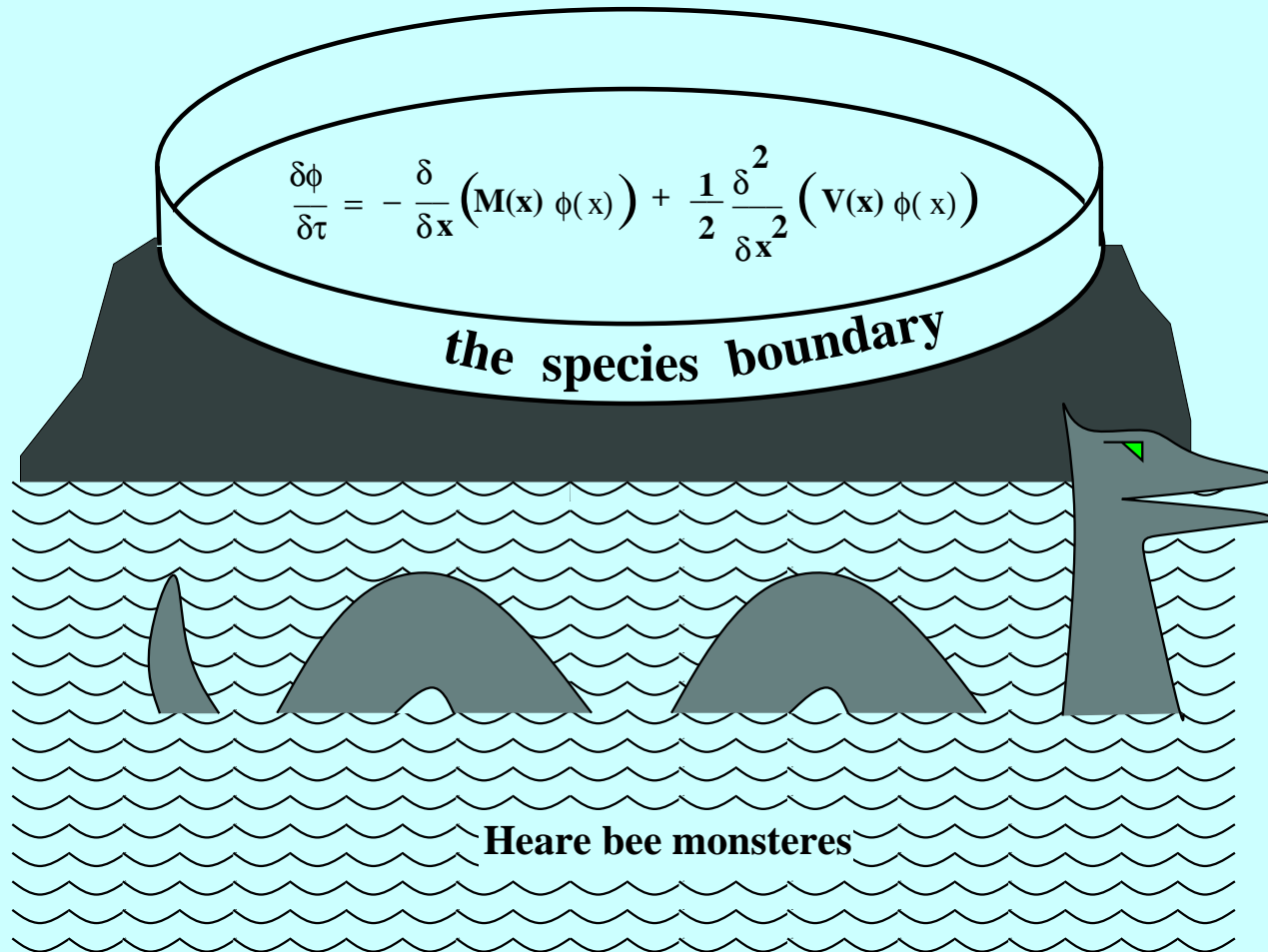
Theodosius Dobzhansky

# Population genetics, around 1970


$$\frac{\delta\phi}{\delta\tau} = -\frac{\delta}{\delta\mathbf{x}} \left( \mathbf{M}(\mathbf{x}) \phi(\mathbf{x}) \right) + \frac{1}{2} \frac{\delta^2}{\delta\mathbf{x}^2} \left( \mathbf{V}(\mathbf{x}) \phi(\mathbf{x}) \right)$$

*the species boundary*

# Population genetics, around 1970



## Finding molecular variation



Richard Lewontin and Jack Hubby's 1966 paper on protein variation (using gel electrophoresis) found many loci to show variation at the molecular level. It was not obvious that this variation affected fitness. Lewontin pointed out that this "neutral mutation" might account for much of the molecular variation within populations.

# The neutral mutation theory



Motō Kimura with his family in Mishima, Japan in the 1960s. The greatest theoretical population geneticist of the late 1900s, he was the chief advocate for the neutral mutation theory and worked out many of its consequences.

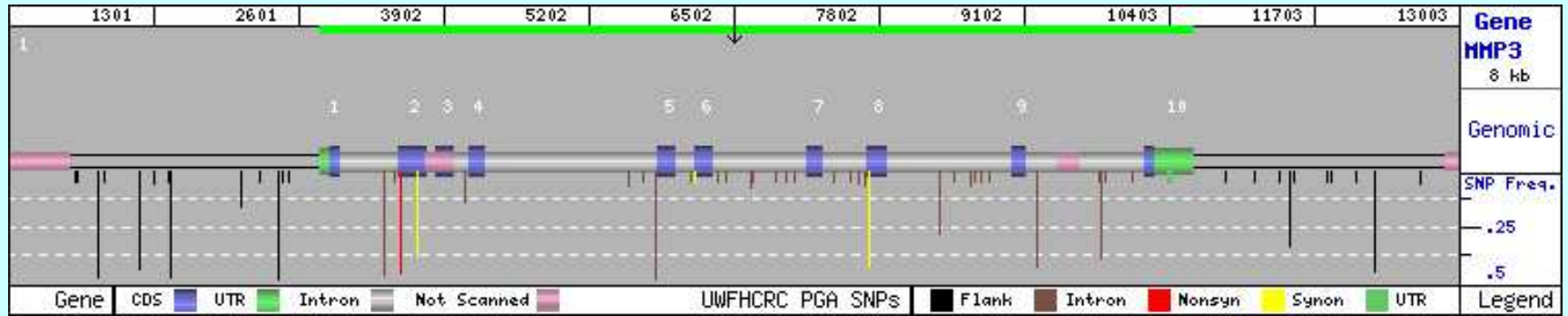
## Lots of variation at the DNA level



Marty Kreitman, as a student of Lewontin's in the early 1980s, used early sequencing methods to look for variation in DNA sequences. Result: you are heterozygous about every 1500 nucleotides.



# A typical locus showing SNP variation



(From Debbie Nickerson's SeattleSNPs project). Single-nucleotide polymorphisms (SNPs) at the Matrix Metalloproteinase 3 locus.

## Molecular evolution (1963 on)

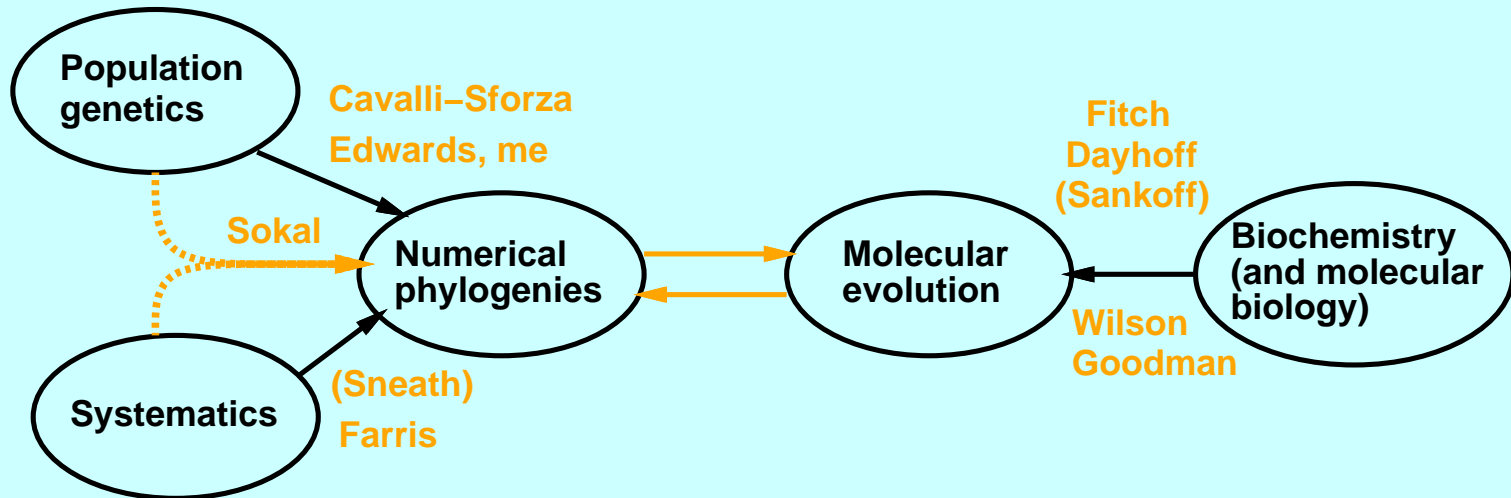


Linus Pauling in 1963



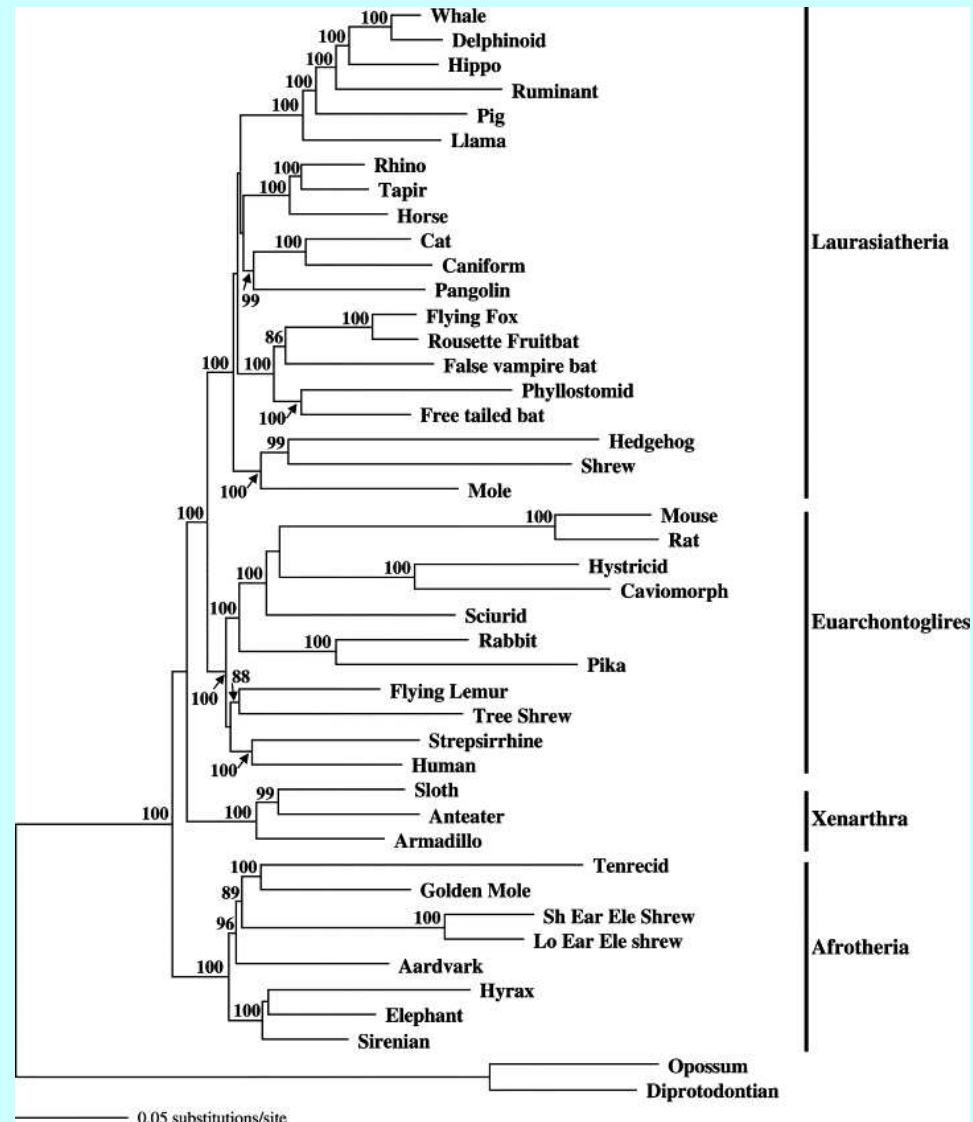
Emile Zuckerkandl, more recently

# Molecular evolution and phylogeny methods



People who pioneered in phylogeny methods and the analysis of molecular evolution data with them.

# An example: who is most closely related to whales?



from Amrine-Madsen, H. et al., 2003, *Molecular Phylogenetics and Evolution*

# Molecular phylogenies

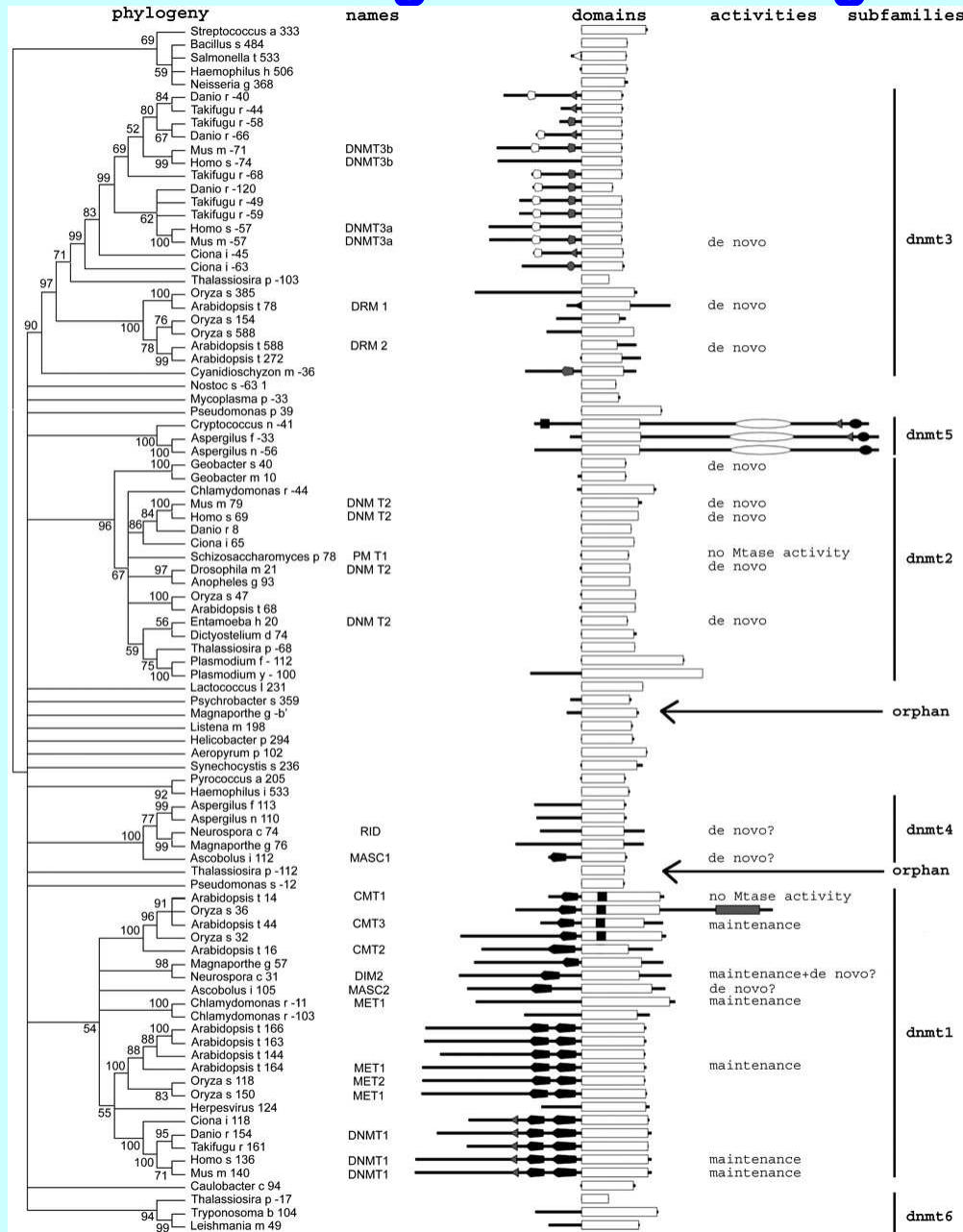
Some examples of other important conclusions from molecular phylogenies:

- Using immunological distances, Morris Goodman (1962 on) and later Wilson and Sarich (1966) show that humans, gorilla, and chimps were a clade.
- Wilson and Sarich (in that work, 1967) date the divergence of humans to 5 million years.
- Charles Sibley and Jon Ahlquist (1984) use DNA hybridization to argue for the clade humans-chimps.
- Carl Woese (1978) uses rRNA trees to introduce evolution into microbiology, argue for the domain Archaea.
- Much progress on early radiation of angiosperms
- Protostome-deuterostome tree of metazoans (more or less) replaced by deuterostome-lophotrichozoa-ecdysozoa tree.
- Fungi closer to animals than either is to plants.
- Symbiotic origin of mitochondria and of chloroplasts verified.
- Amphioxus diverged before split of tunicates from craniate chordates.
- Lots of horizontal gene transfer in prokaryotes, almost not a tree.

# Wen-Hsiung Li



# Wen-Hsiung Li's work on gene duplication



C5-cytosine methyltransferase gene family tree. Where are humans? (from Ponger and Li, 2005, *Molecular Biology and Evolution*)

## The “mitochondrial Eve” study in 1987



Rebecca Cann, Mark Stoneking, and the late Allan Wilson. In 1987 they made a molecular tree of mitochondria from humans.



# One female ancestor? of what? When? Where?

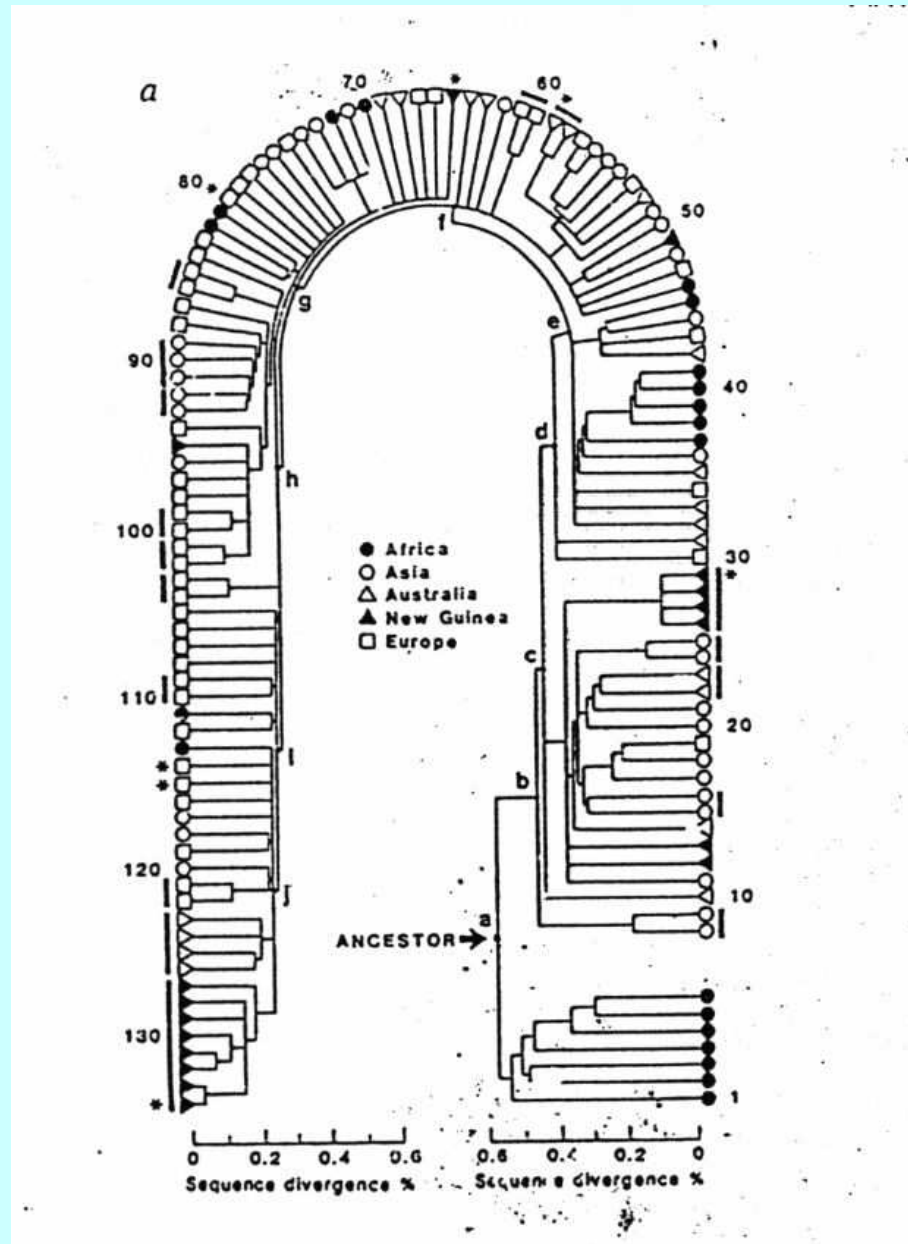


Fig. 3 a, Genealogical tree for 134 types of human mtDNA (133 restric

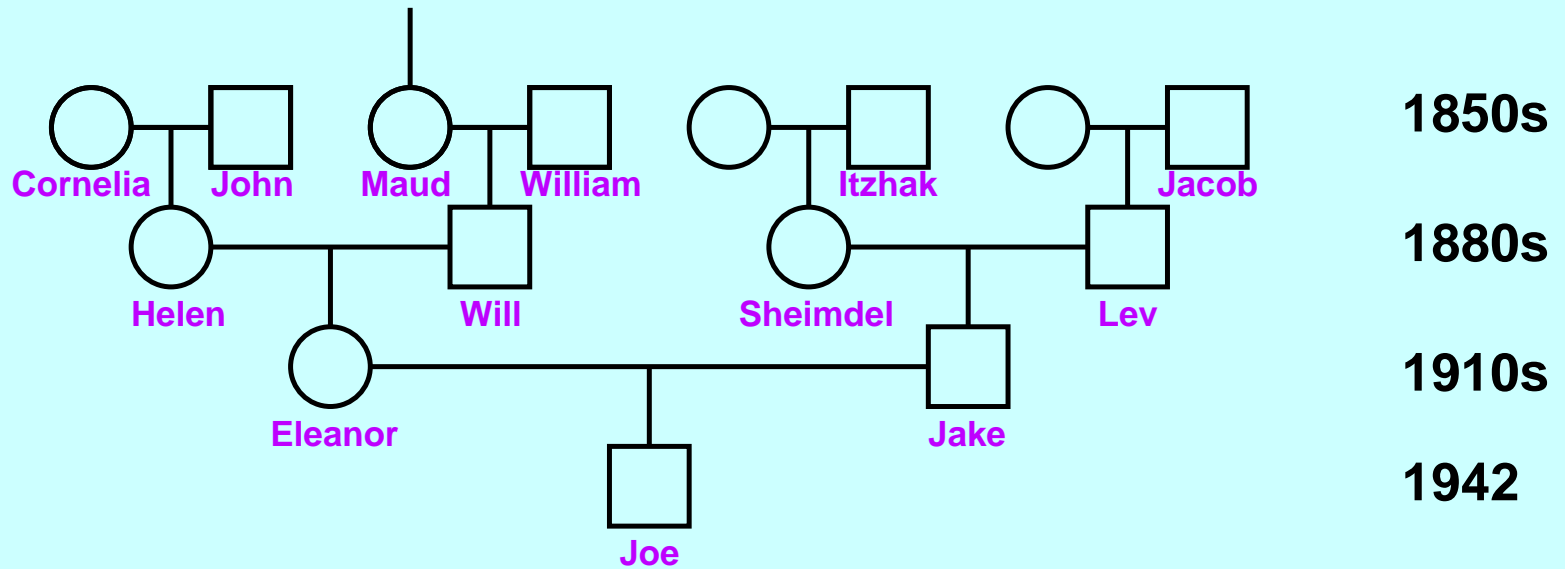
# My ancestor?

Charles the Great  
(Charlemagne)

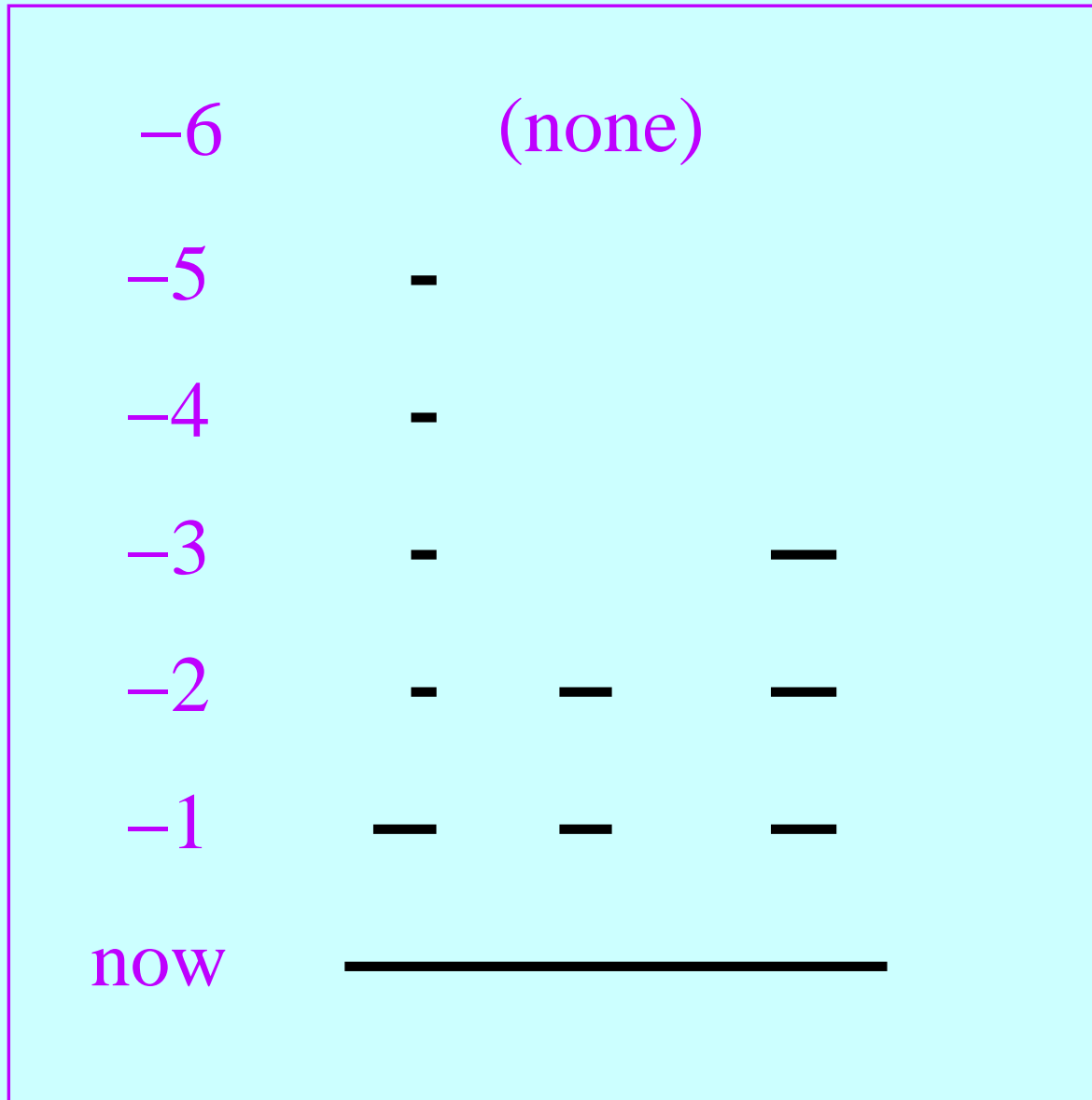


born  
747

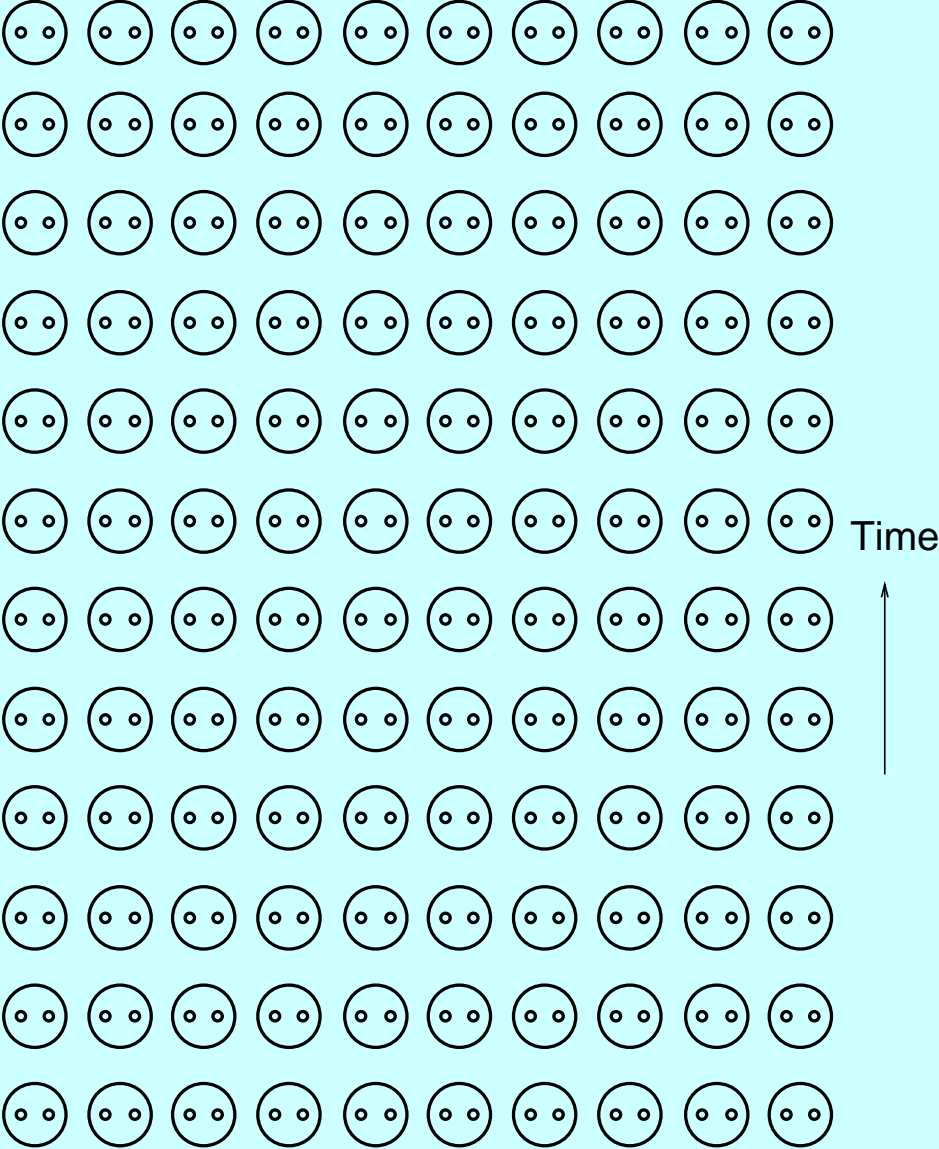
about 44 more generations



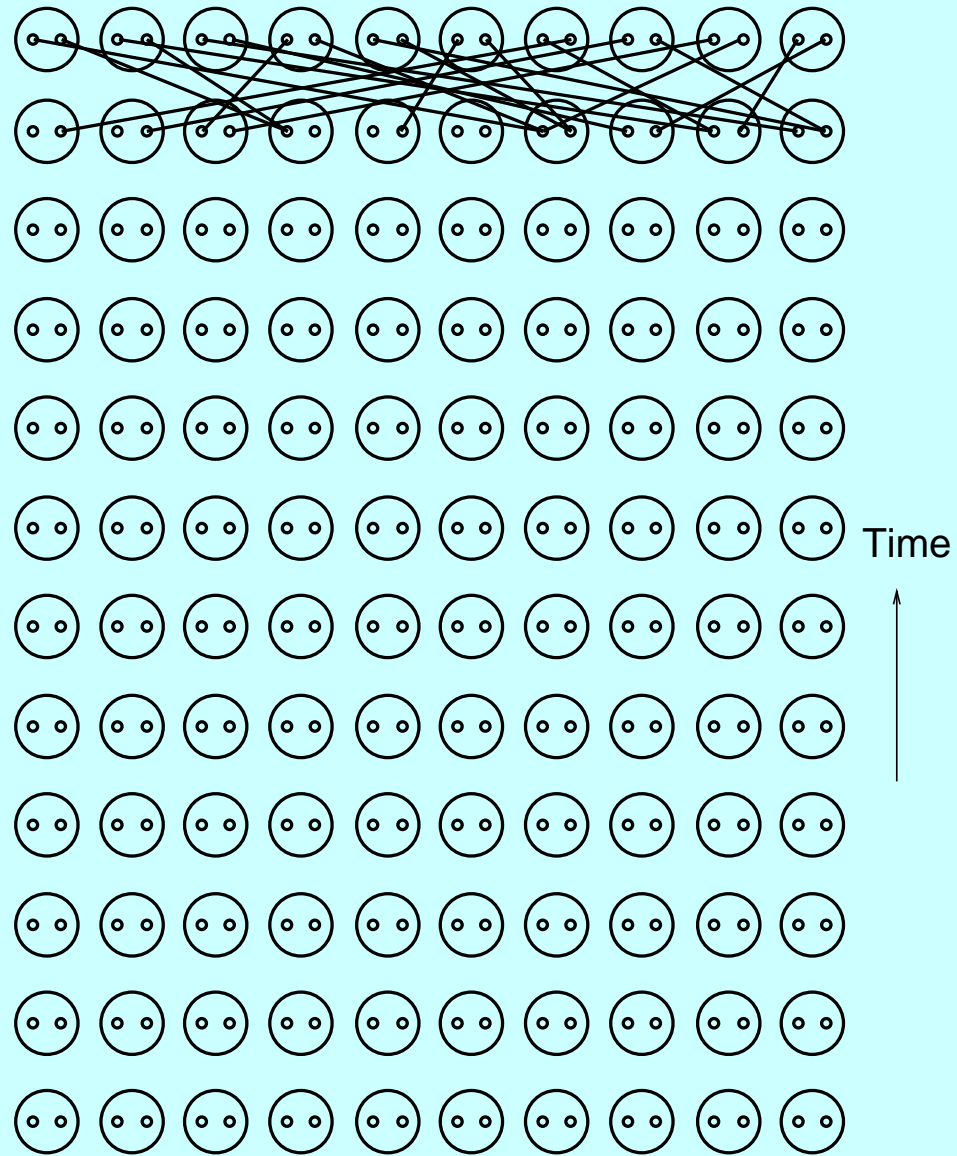
# Chromosome 1, back up one lineage



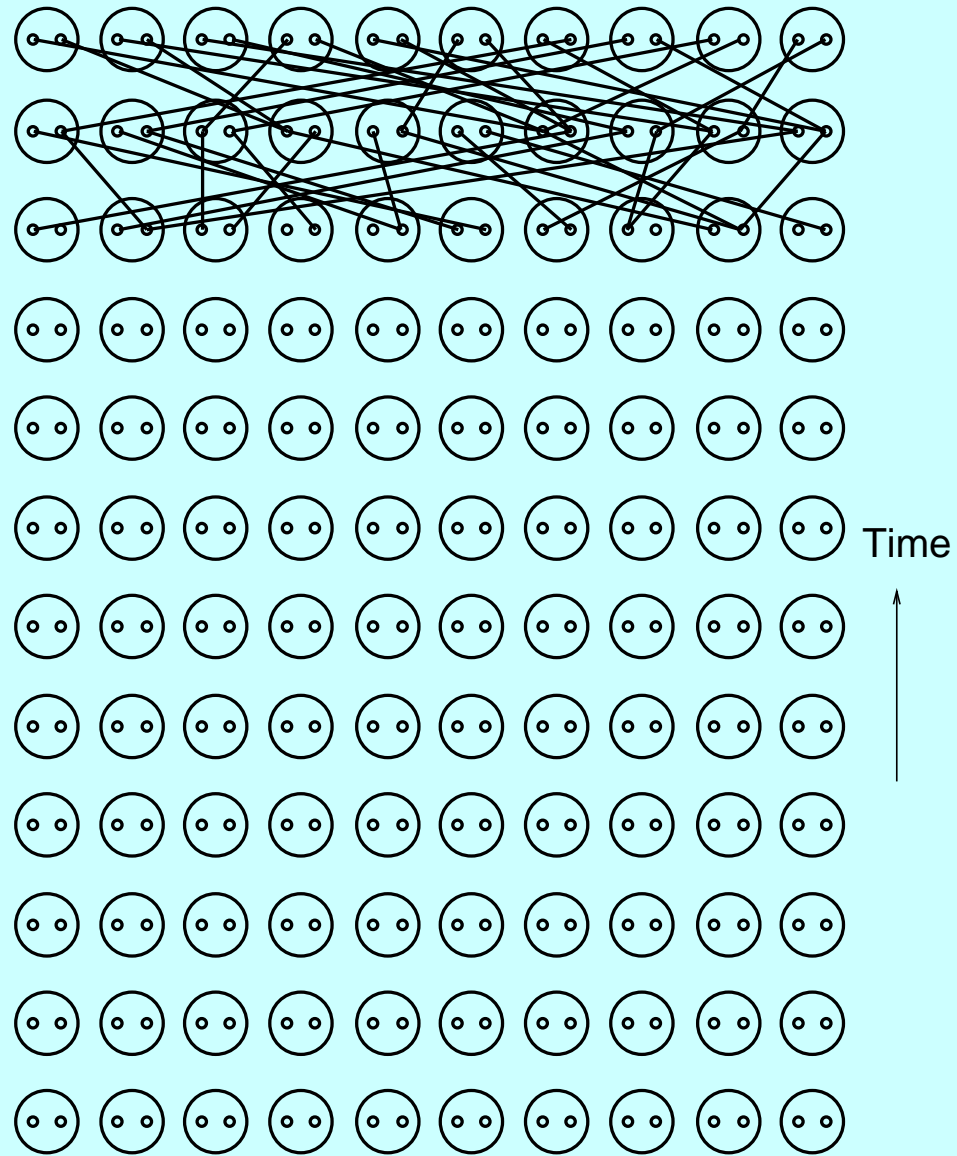
# Coalescent genealogy for one gene



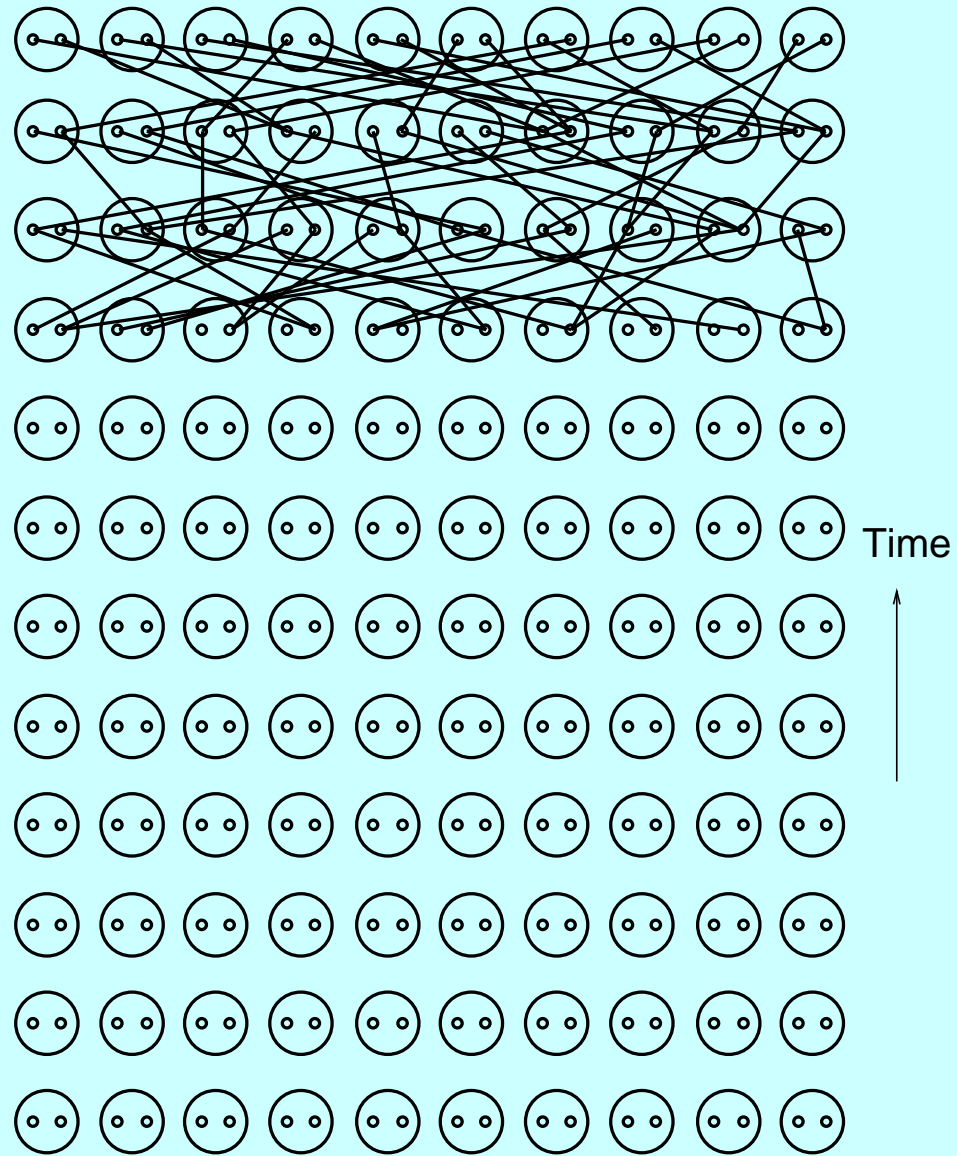
# Coalescent genealogy for one gene



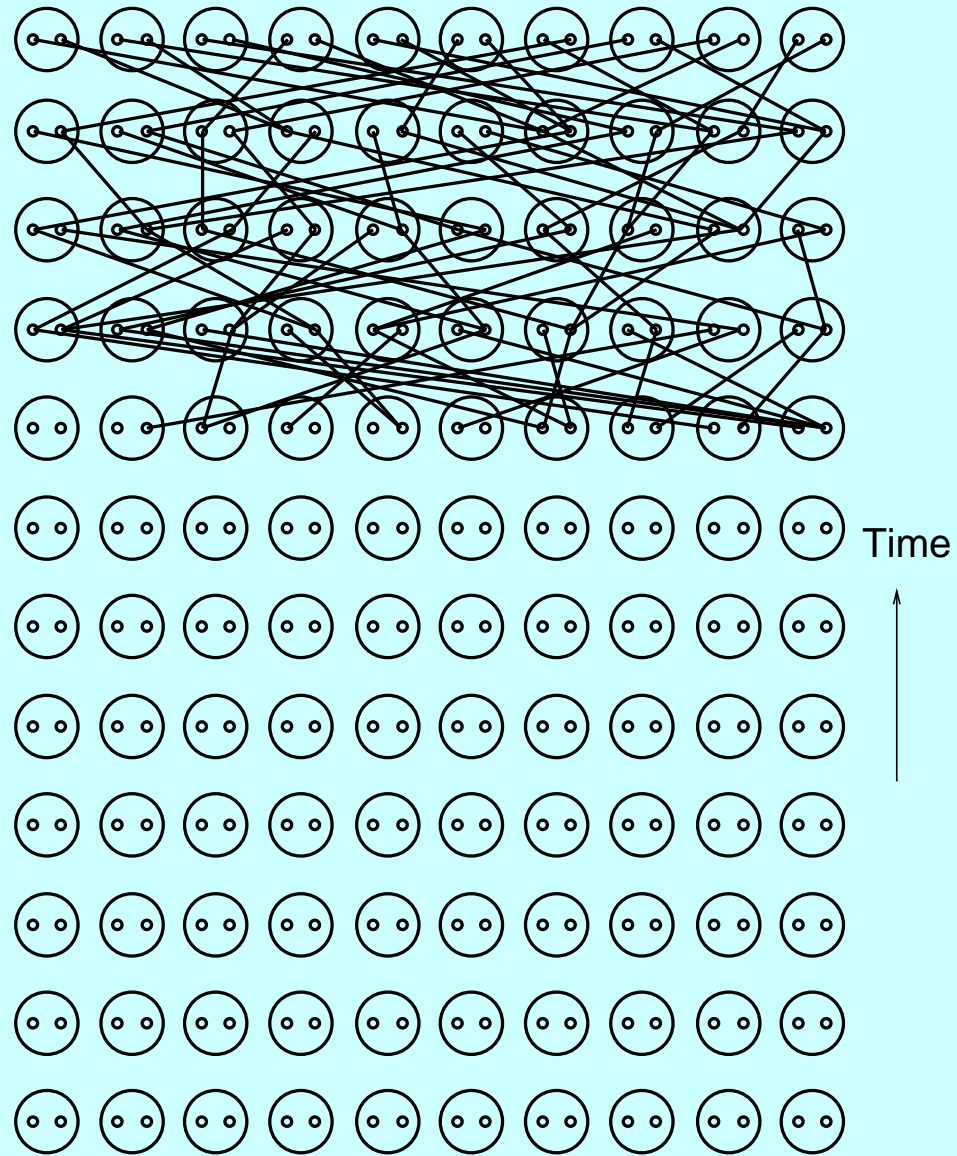
# Coalescent genealogy for one gene



# Coalescent genealogy for one gene

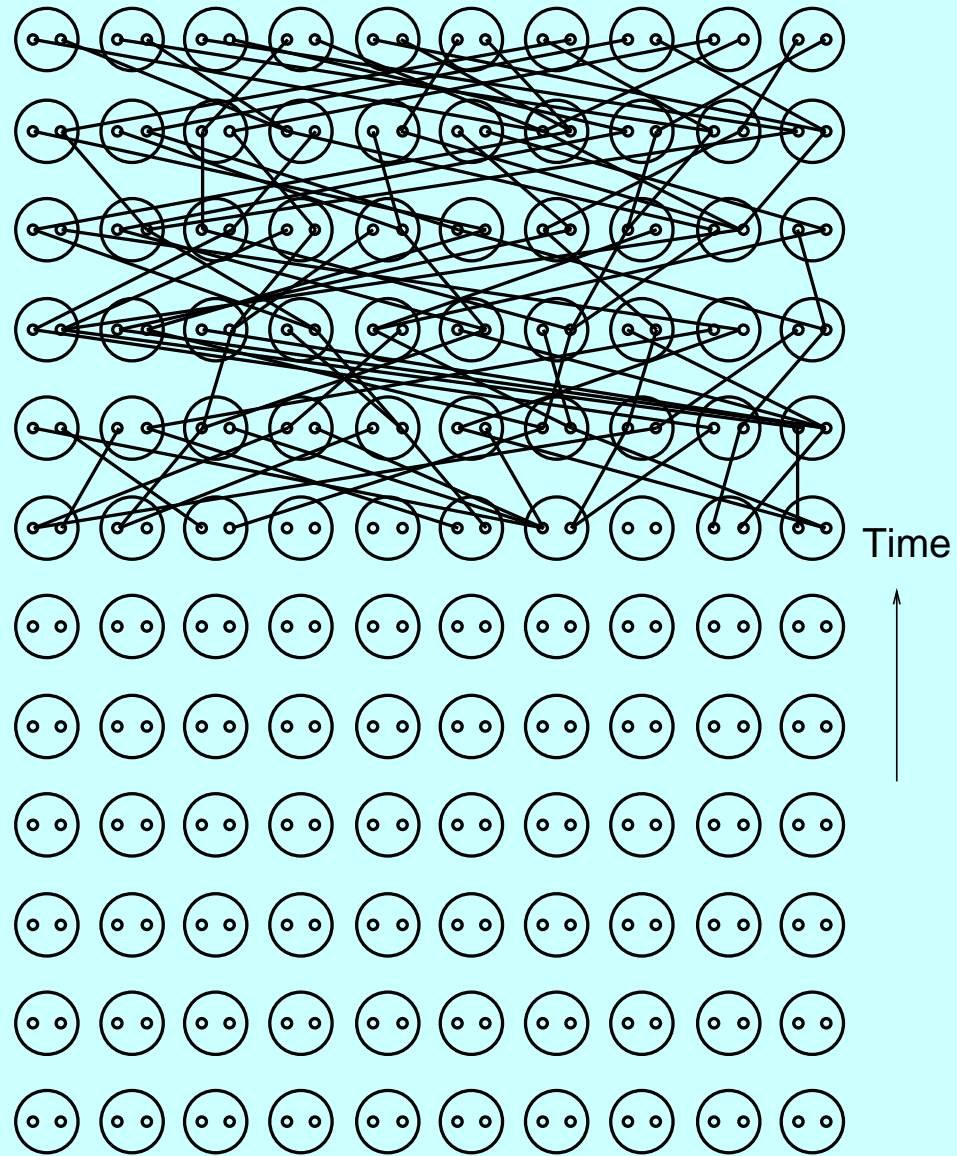


# Coalescent genealogy for one gene

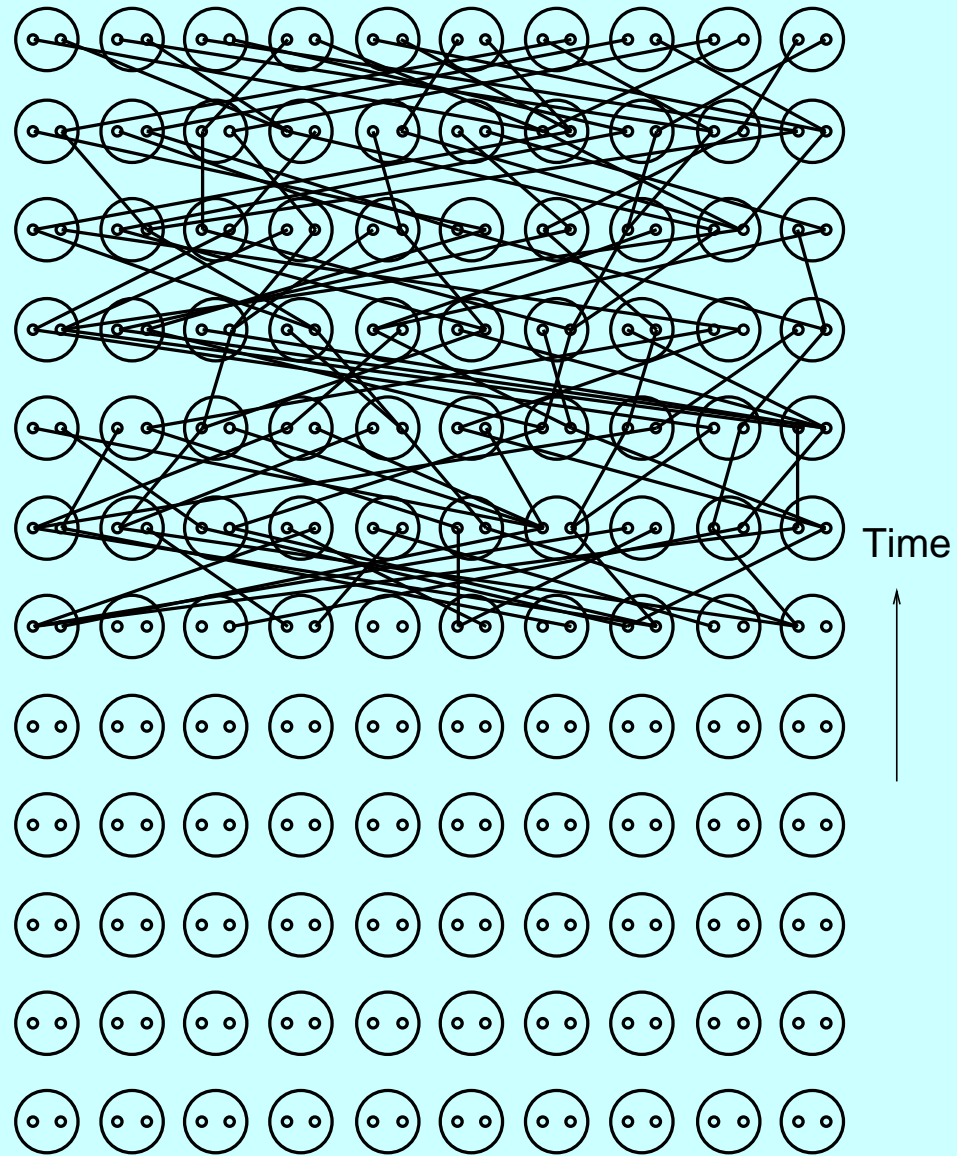




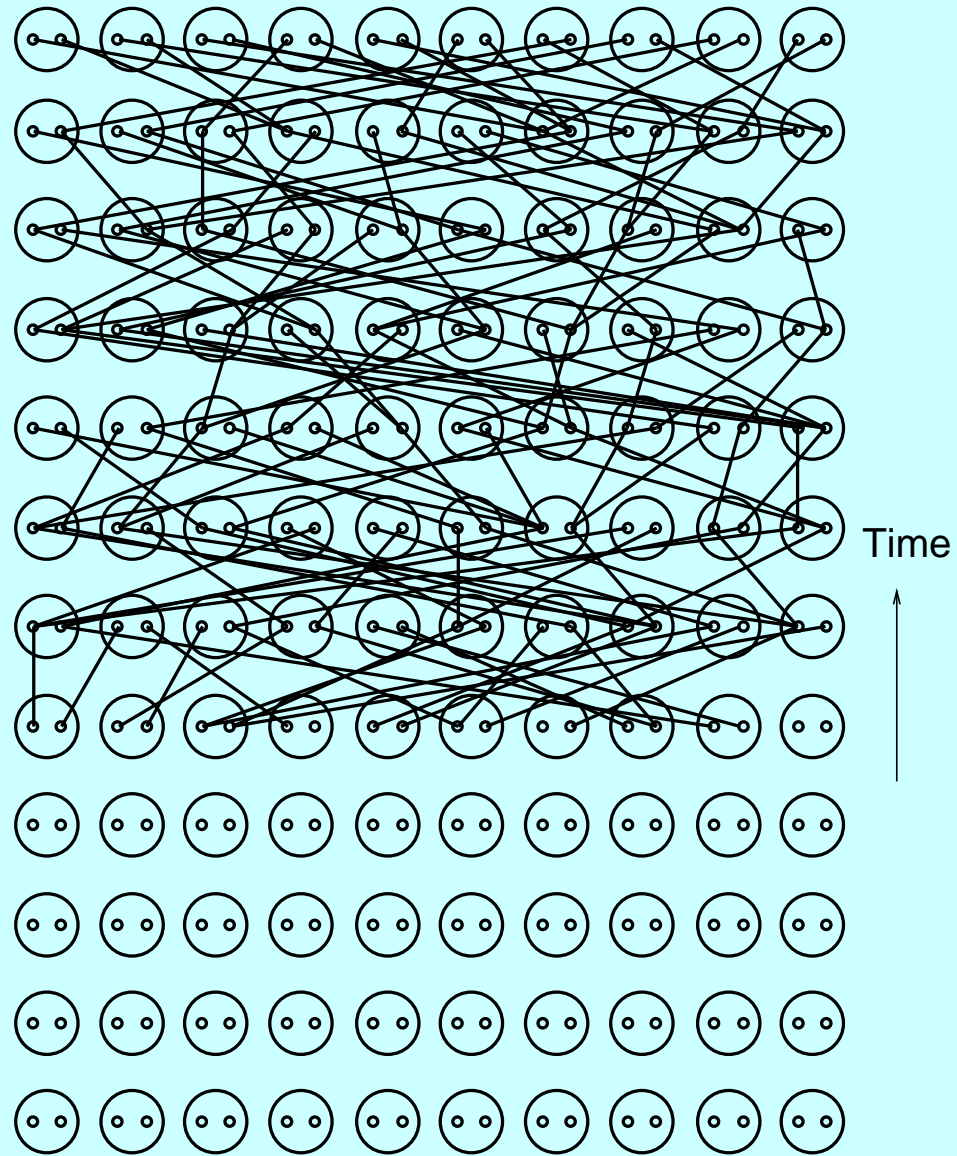
# Coalescent genealogy for one gene



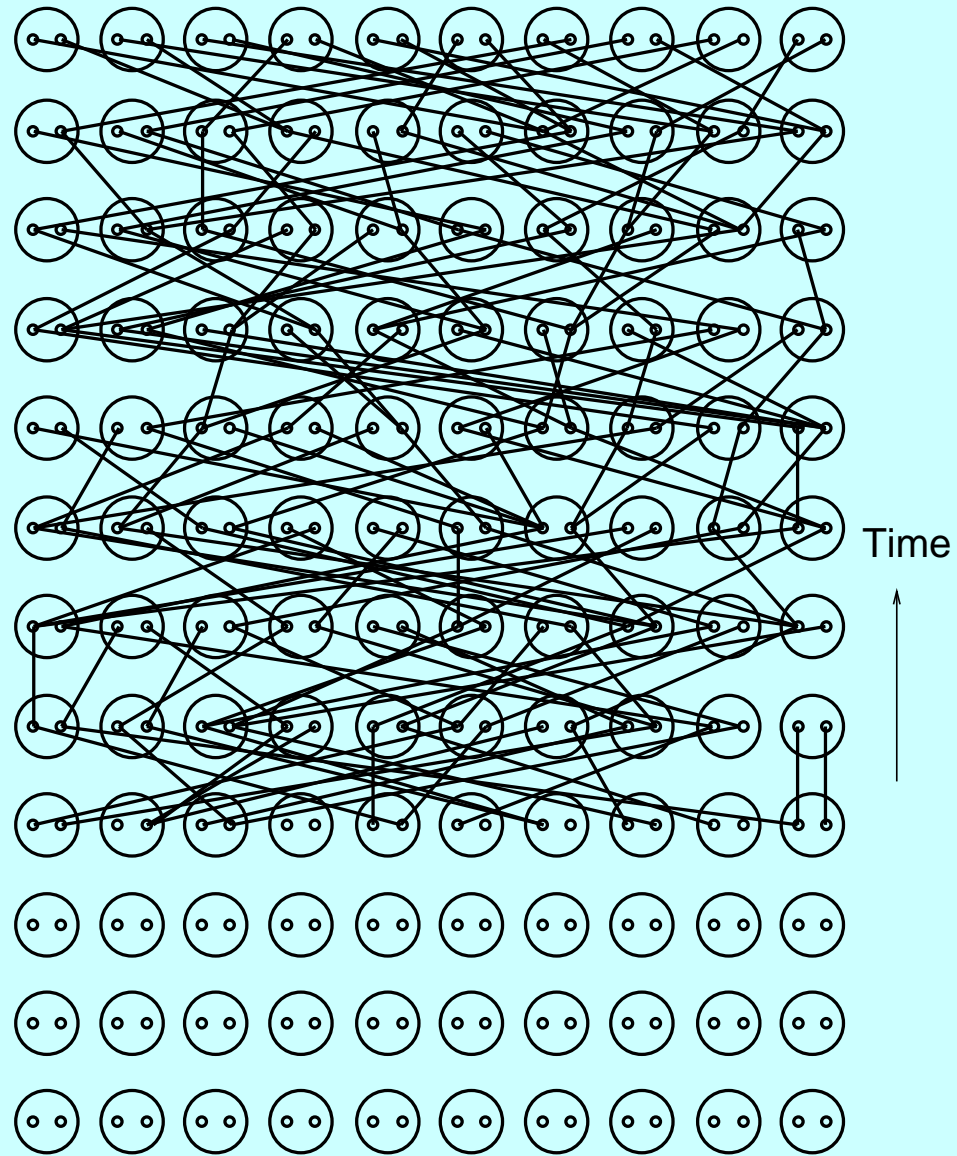
# Coalescent genealogy for one gene



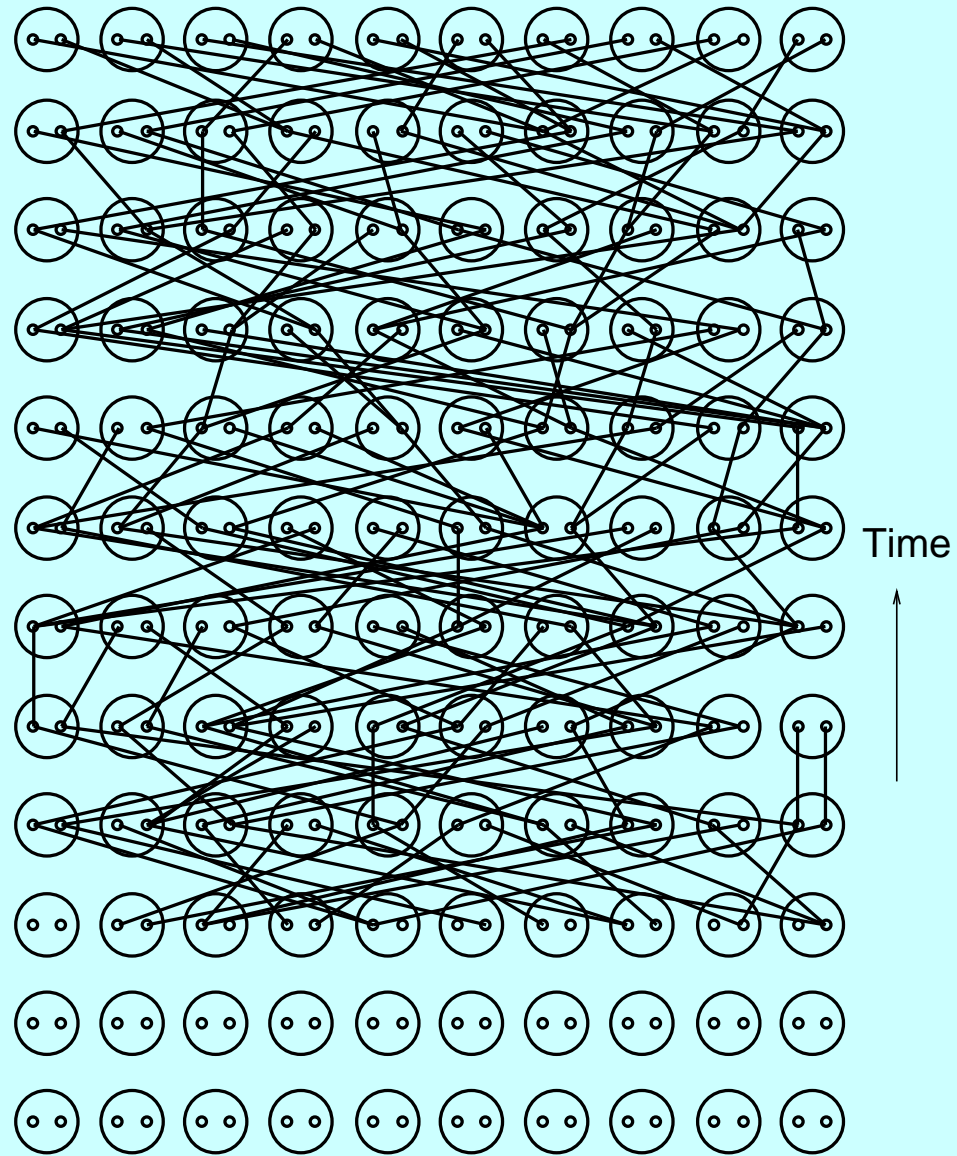
# Coalescent genealogy for one gene



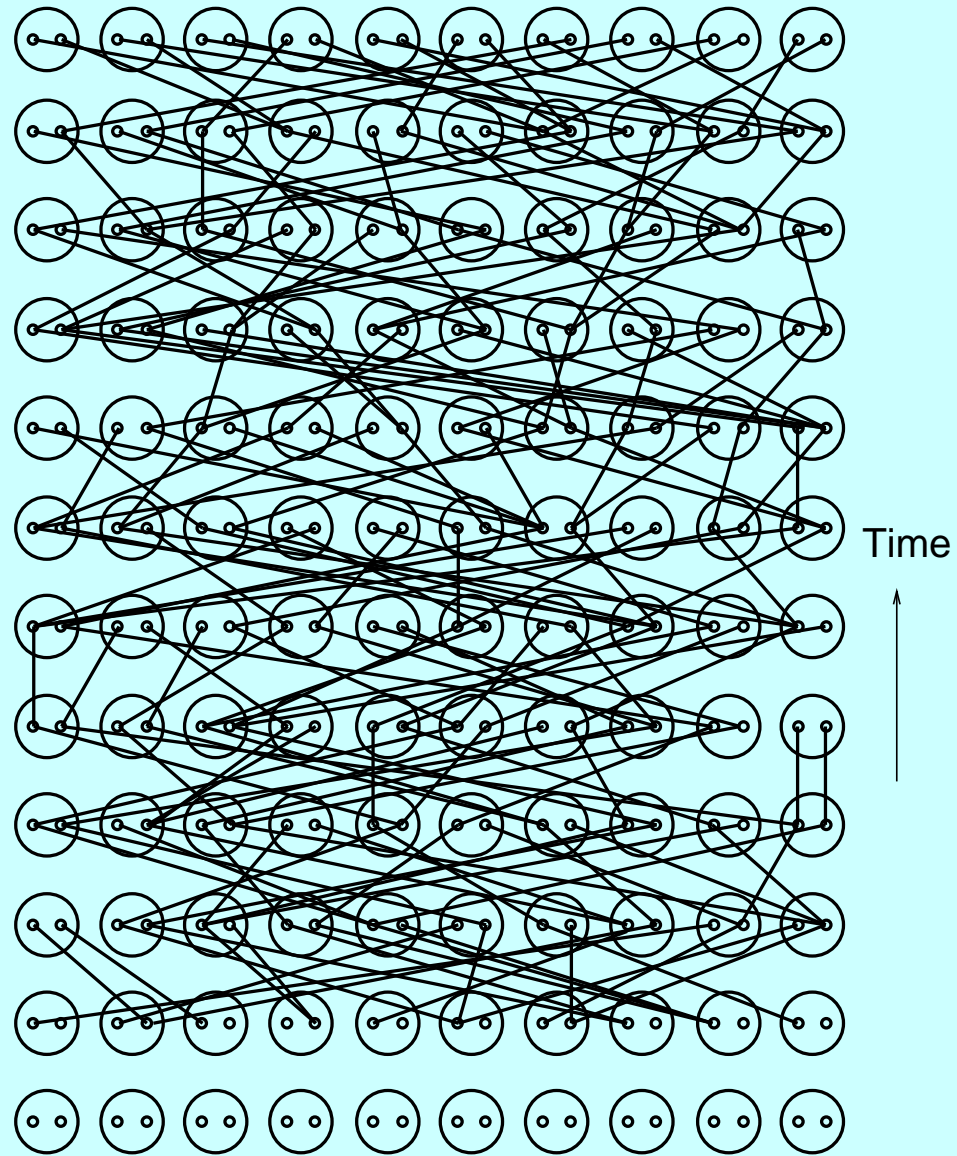
# Coalescent genealogy for one gene



# Coalescent genealogy for one gene

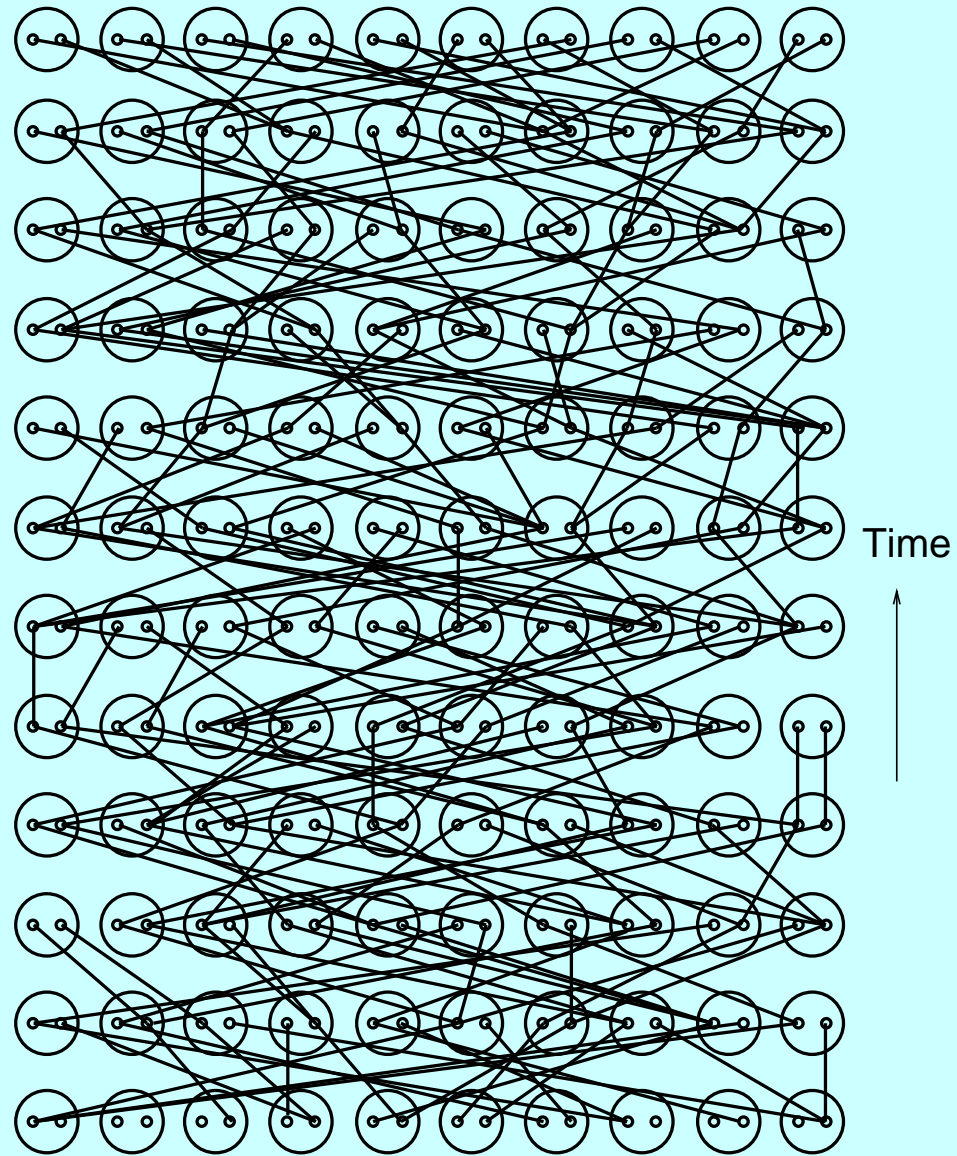


# Coalescent genealogy for one gene

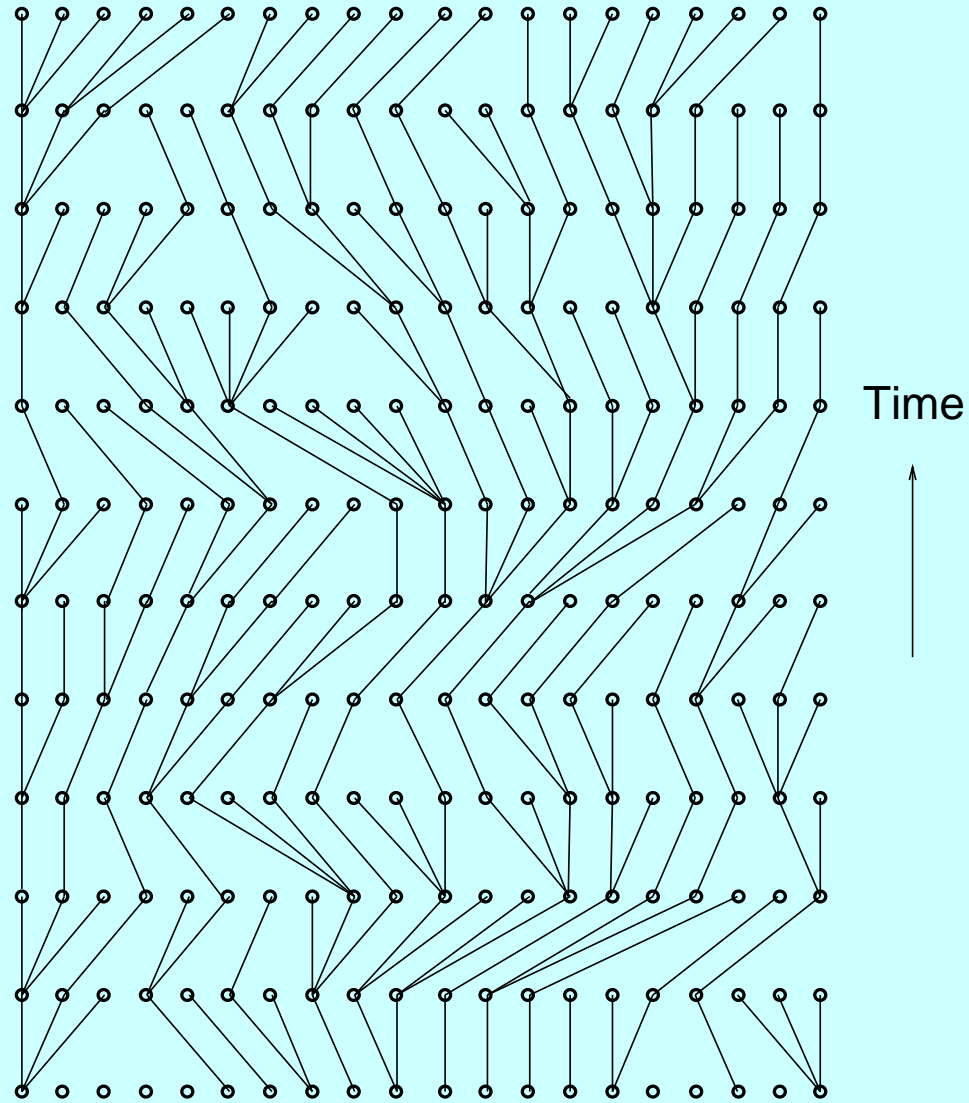




# Coalescent genealogy for one gene

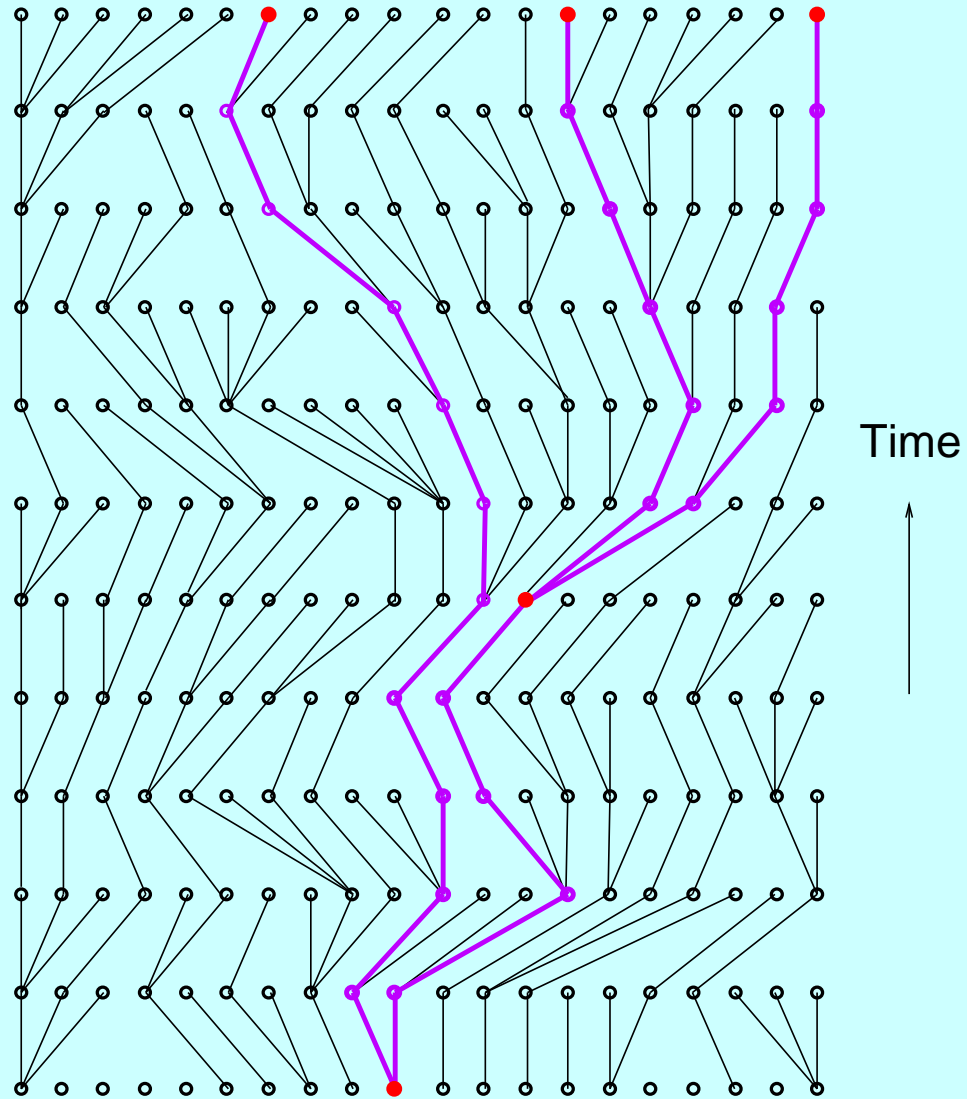


# Untangling the crossed lines ...





# Genealogy of a sample of 3 copies



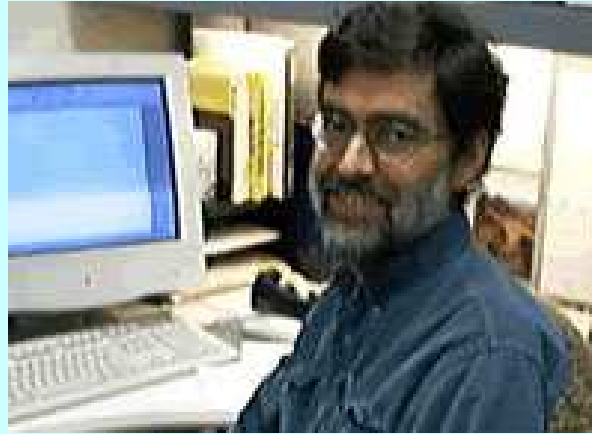
## J. F. C. Kingman's (1982) "coalescent"



1. start with  $n$  tips
2. go back an amount of time  
drawn from Exponential  $\left(\frac{4N}{n(n-1)}\right)$
3. join a random pair of the  $n$
4.  $n \leftarrow n - 1$
5. if  $n = 1$  stop, else go to step 2.

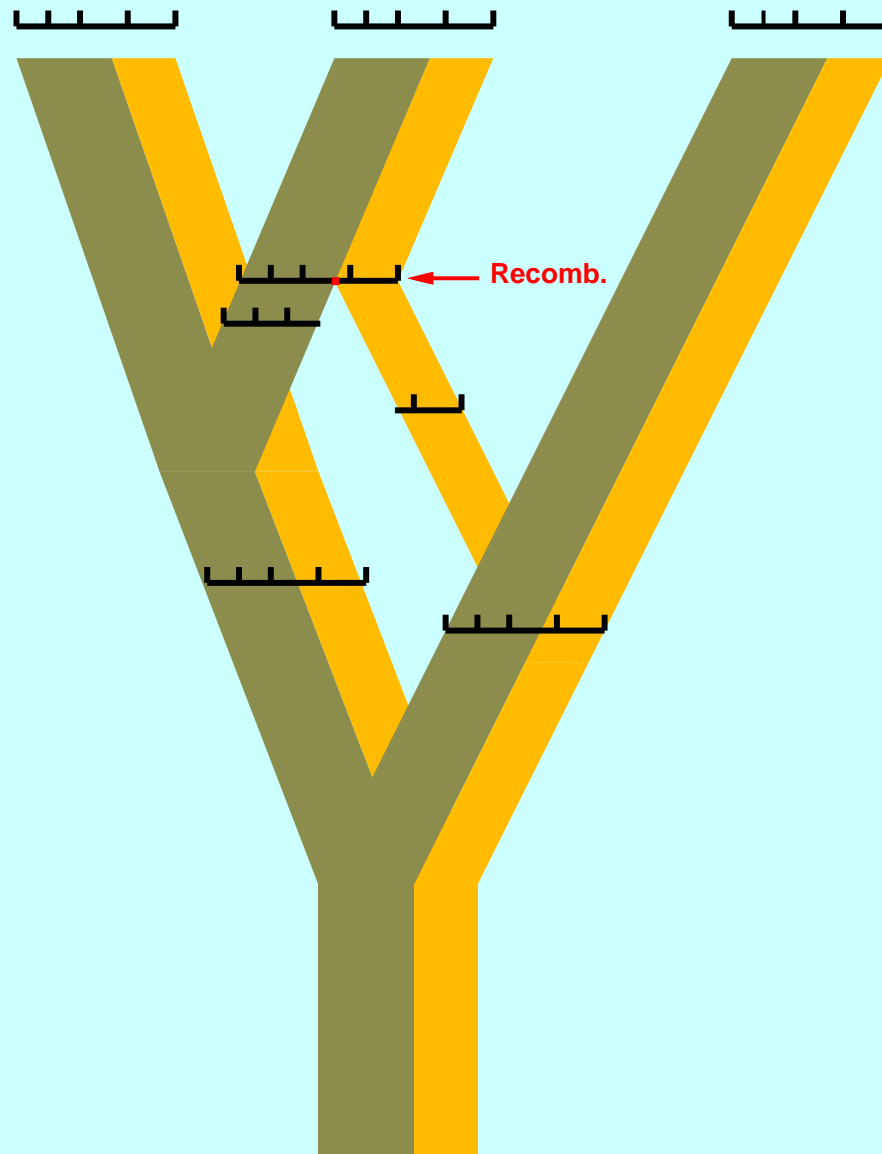
This excellently approximates the distribution of genealogies which arise from samples from a standard (Wright-Fisher) population genetics model with a population size of  $N$ , provided  $n^2 \ll N$

## Pioneer of coalescent theory



Dick Hudson, pioneered understanding of coalescents having recombination or natural selection

# A coalescent with recombination

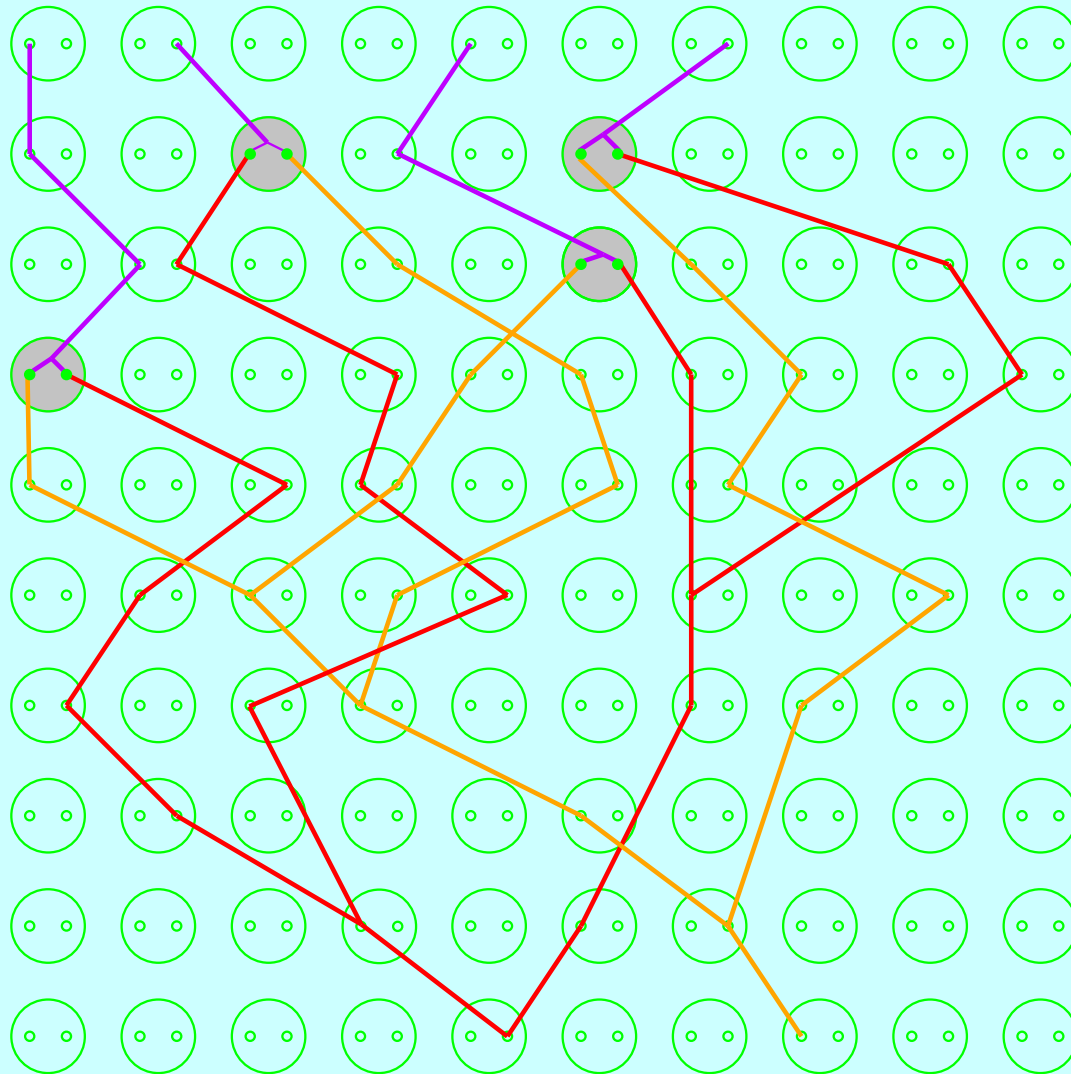


Different markers have slightly different coalescent trees

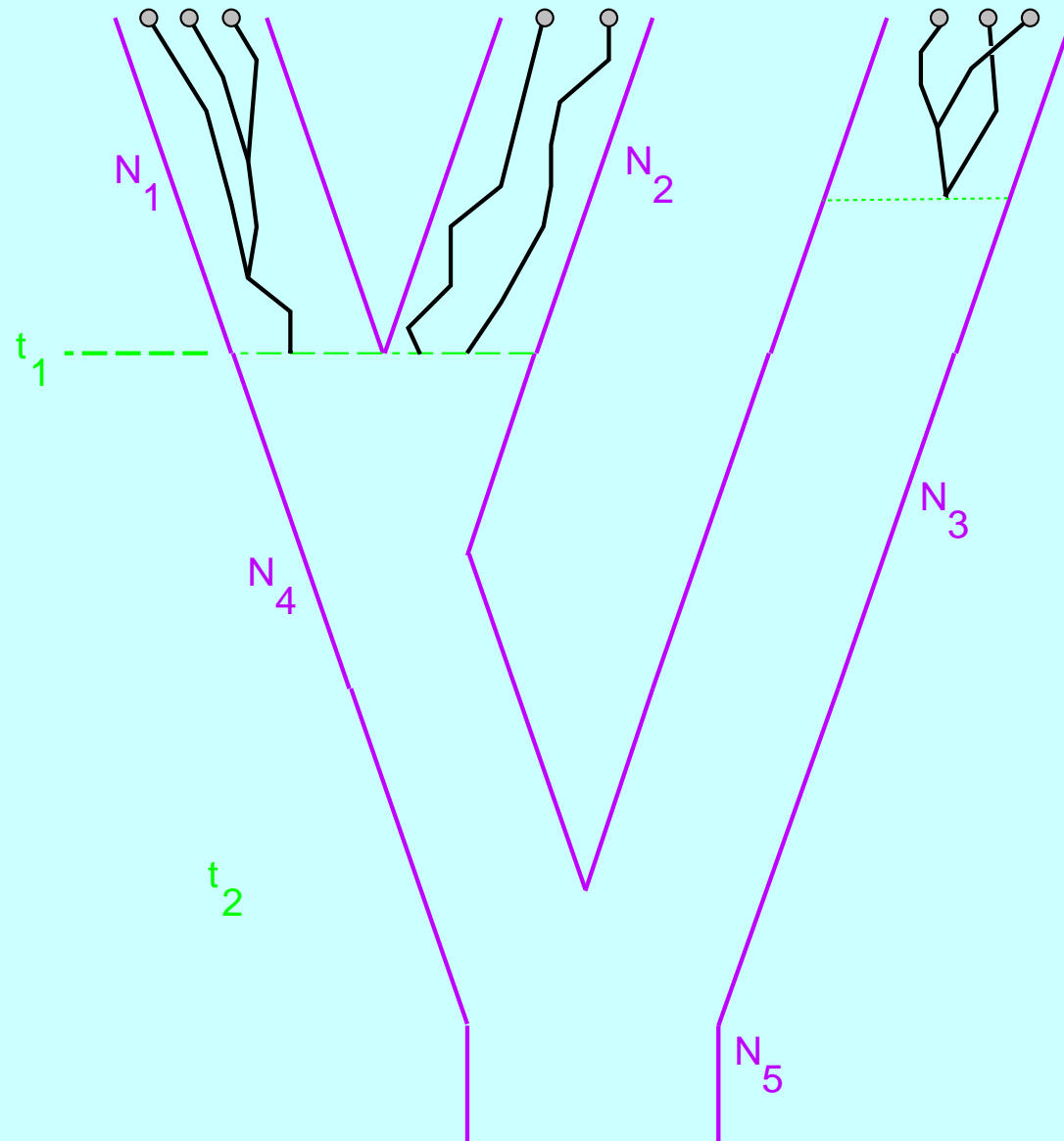
# Coalescents for two loosely-linked genes

— locus A  
— locus B  
— both

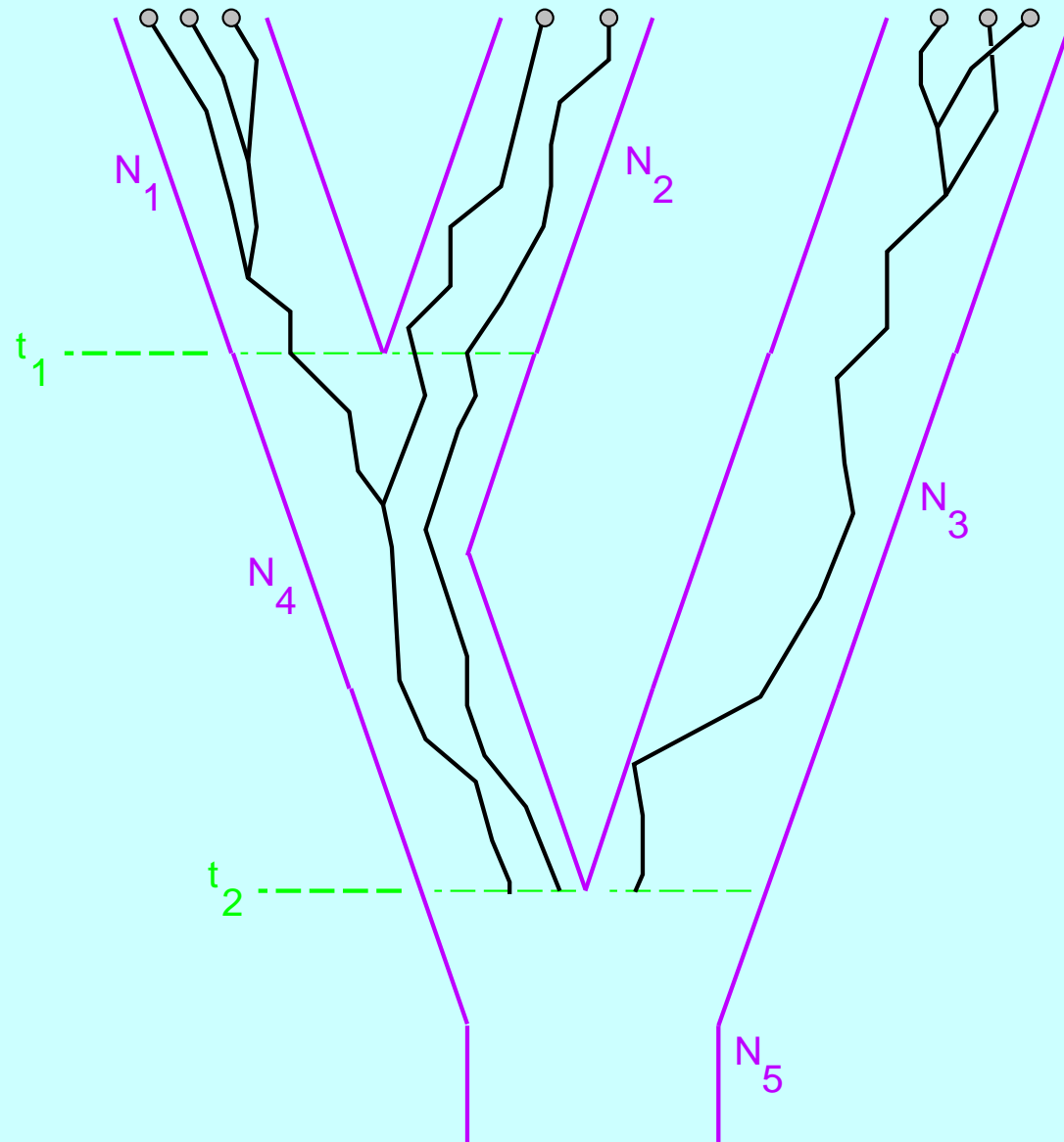
● = recombination



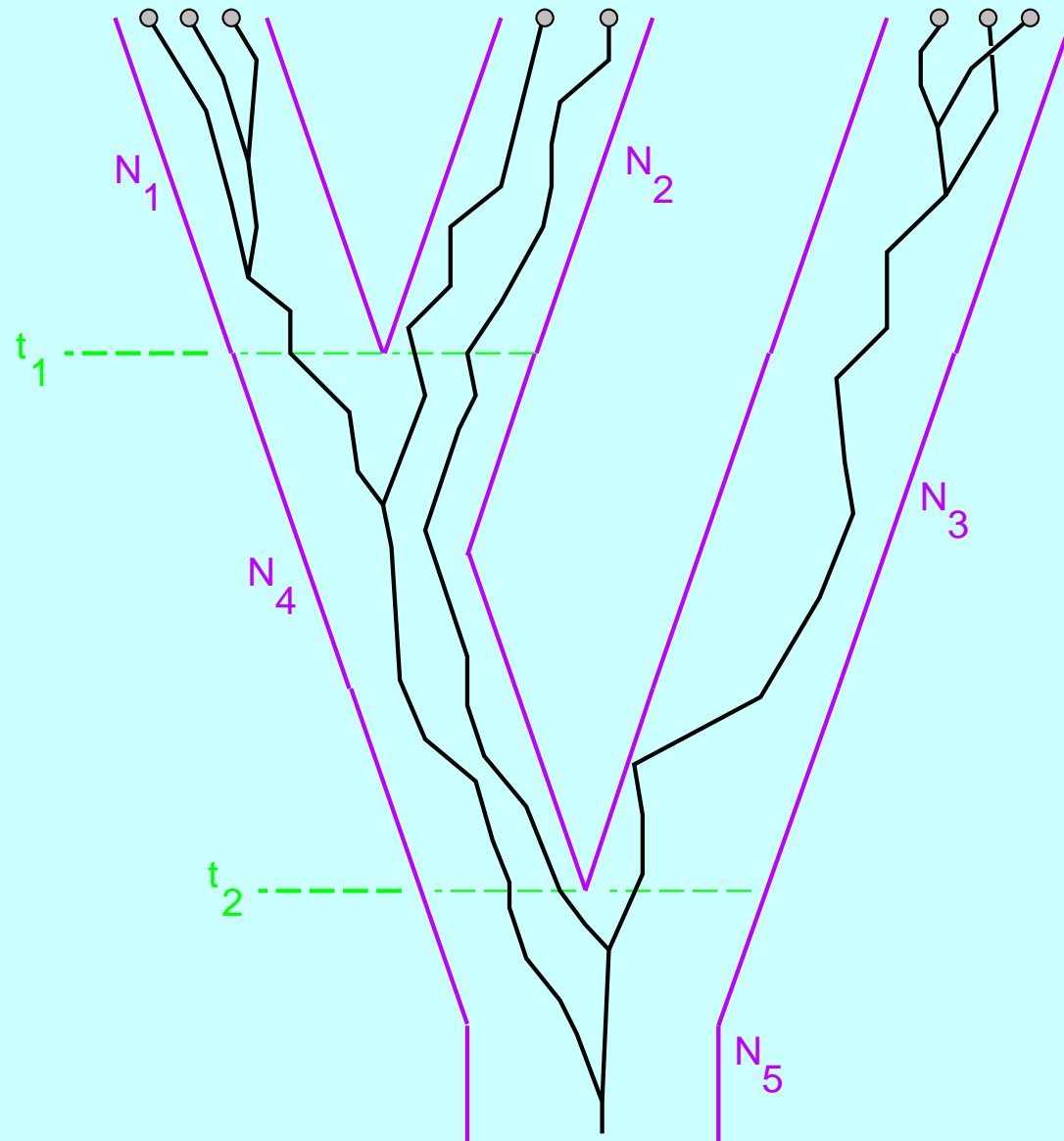
# Species trees and trees of gene copies



# Species trees and trees of gene copies



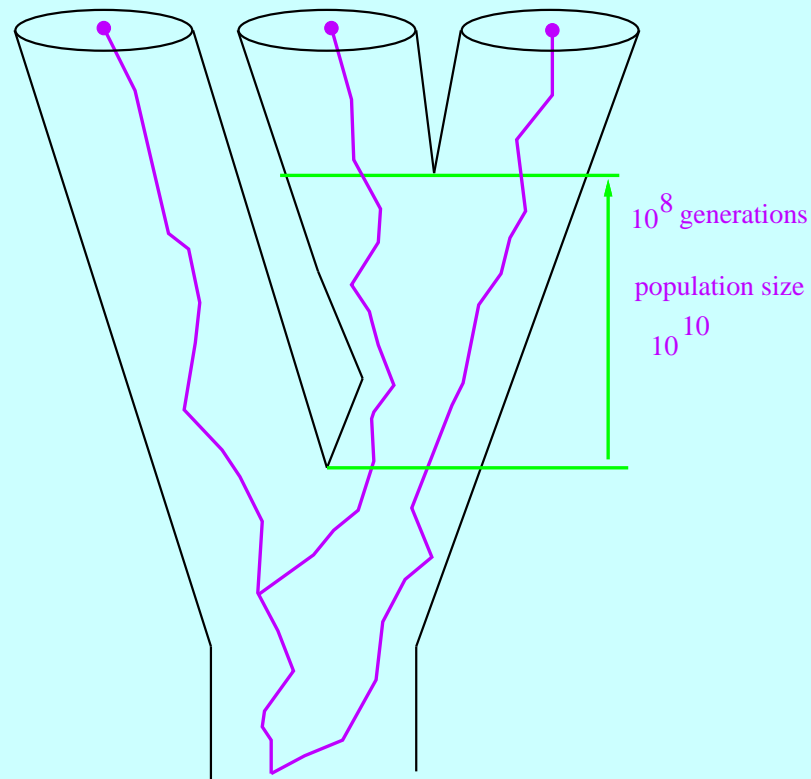
# Species trees and trees of gene copies





# Protists and bacteria – a worry

If protist (or bacterial) populations remain large for long periods of time ...



... is it possible that some apparent horizontal gene transfer events are actually just species-tree / gene-tree discrepancies due to coalescent effects? Has this been examined?

# Approaches to breaching the species barrier



Jerry Coyne



Allen Orr

- Direct assault. Make use of cases where we can cross species or incipient species. Jerry Coyne, Allen Orr, Nick Barton, etc.

# Approaches to breaching the species barrier



Jerry Coyne



Allen Orr

- Direct assault. Make use of cases where we can cross species or incipient species. Jerry Coyne, Allen Orr, Nick Barton, etc.
- QTLs across species. e.g., Toby Bradshaw and Doug Schemske

# Approaches to breaching the species barrier



Jerry Coyne



Allen Orr

- Direct assault. Make use of cases where we can cross species or incipient species. Jerry Coyne, Allen Orr, Nick Barton, etc.
- QTLs across species. e.g., Toby Bradshaw and Doug Schemske
- Going round the other way. Used inadvertently by people studying molecular evolution, then coalescents.

# Approaches to breaching the species barrier



Jerry Coyne



Allen Orr

- Direct assault. Make use of cases where we can cross species or incipient species. Jerry Coyne, Allen Orr, Nick Barton, etc.
- QTLs across species. e.g., Toby Bradshaw and Doug Schemske
- Going round the other way. Used inadvertently by people studying molecular evolution, then coalescents.
  - Study of coalescents at time of speciation. Jody Hey, Rich Kliman.

# Approaches to breaching the species barrier



Jerry Coyne



Allen Orr

- Direct assault. Make use of cases where we can cross species or incipient species. Jerry Coyne, Allen Orr, Nick Barton, etc.
- QTLs across species. e.g., Toby Bradshaw and Doug Schemske
- Going round the other way. Used inadvertently by people studying molecular evolution, then coalescents.
  - Study of coalescents at time of speciation. Jody Hey, Rich Kliman.
  - Synonymous/nonsynonymous comparisons. Masatoshi Nei, Takashi Gojobori, Ziheng Yang, Rasmus Nielsen, John Huelsenbeck

# Between-species and within-species evolutionary work

- They are increasingly coming into contact.

# Between-species and within-species evolutionary work

- They are increasingly coming into contact.
- Signs of this as well in a new wave of work on modelling quantitative genetics in natural populations



# Between-species and within-species evolutionary work

- They are increasingly coming into contact.
- Signs of this as well in a new wave of work on modelling quantitative genetics in natural populations
- Renewed controversy about models of speciation (Doebeli versus Gavrillets)

# Between-species and within-species evolutionary work

- They are increasingly coming into contact.
- Signs of this as well in a new wave of work on modelling quantitative genetics in natural populations
- Renewed controversy about models of speciation (Doebeli versus Gavrillets)
- But ... what do we call the event?

# Between-species and within-species evolutionary work

- They are increasingly coming into contact.
- Signs of this as well in a new wave of work on modelling quantitative genetics in natural populations
- Renewed controversy about models of speciation (Doebeli versus Gavrillets)
- But ... what do we call the event?
  - The Reunion?

# Between-species and within-species evolutionary work

- They are increasingly coming into contact.
- Signs of this as well in a new wave of work on modelling quantitative genetics in natural populations
- Renewed controversy about models of speciation (Doebeli versus Gavrillets)
- But ... what do we call the event?
  - The Reunion?
  - The Final Roundup?

# Between-species and within-species evolutionary work

- They are increasingly coming into contact.
- Signs of this as well in a new wave of work on modelling quantitative genetics in natural populations
- Renewed controversy about models of speciation (Doebeli versus Gavrillets)
- But ... what do we call the event?
  - The Reunion?
  - The Final Roundup?
  - (not The Postneodarwinian Synthesis)

# Between-species and within-species evolutionary work

- They are increasingly coming into contact.
- Signs of this as well in a new wave of work on modelling quantitative genetics in natural populations
- Renewed controversy about models of speciation (Doebeli versus Gavrillets)
- But ... what do we call the event?
  - The Reunion?
  - The Final Roundup?
  - (not The Postneodarwinian Synthesis)
  - (definitely not The Postmodern Synthesis)

## Thousands of SNPs?

- SNPs will help integrate the statistical variation within populations, between populations, and between species.
- They will also allow us to connect QTL statistical genetics with morphological phylogenies
- Still, there will be a lot of statistics to do to correct for false positives.