

Coalescents

Joe Felsenstein

GENOME 453, Autumn 2011

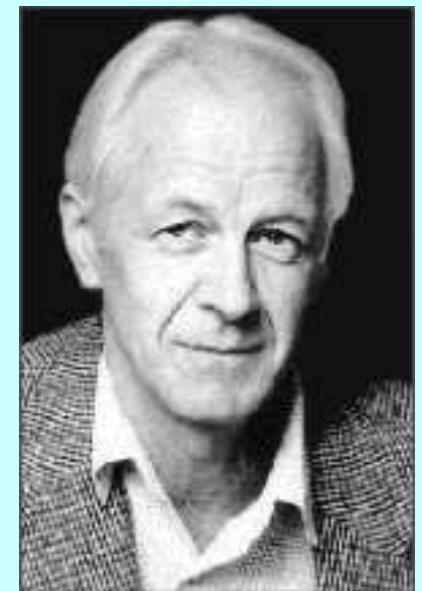
Cann, Stoneking, and Wilson



Becky Cann



Mark Stoneking



the late Allan Wilson

Cann, R. L., M. Stoneking, and A. C. Wilson. 1987. Mitochondrial DNA and human evolution. *Nature* **325:a** 31-36.

Mitochondrial Eve

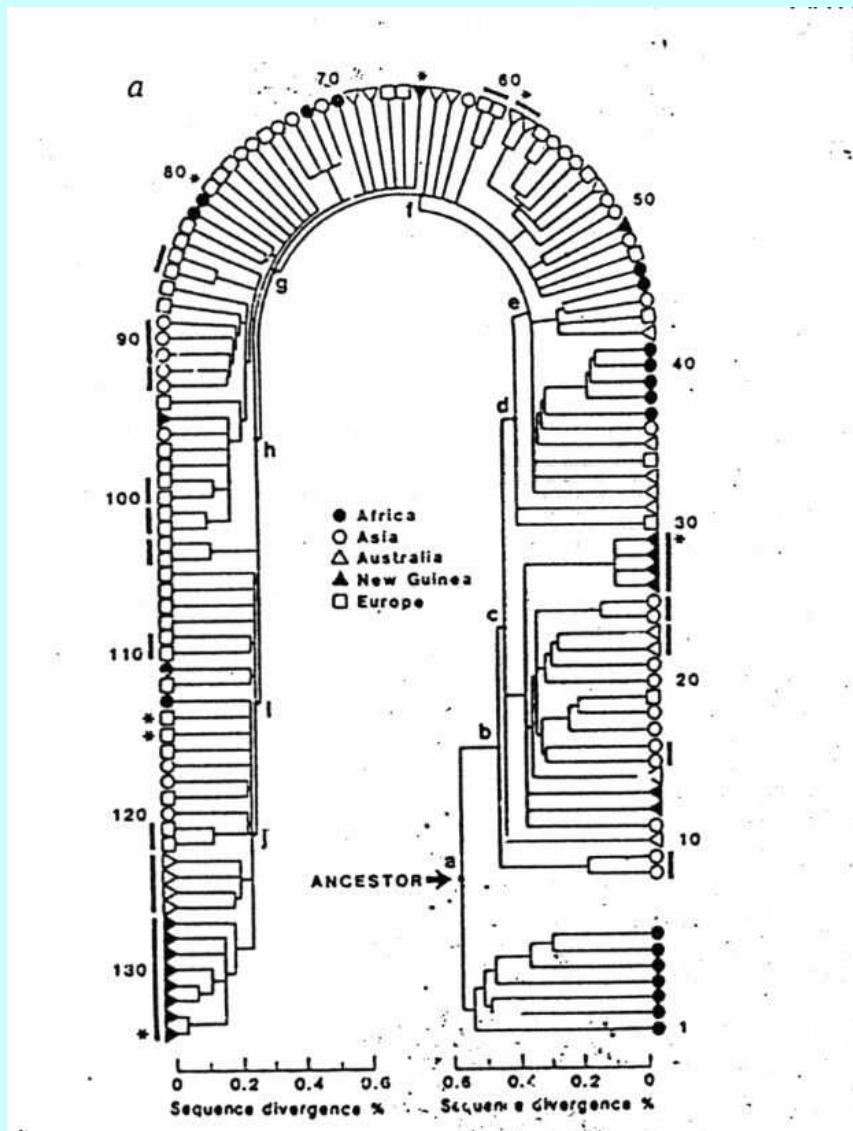
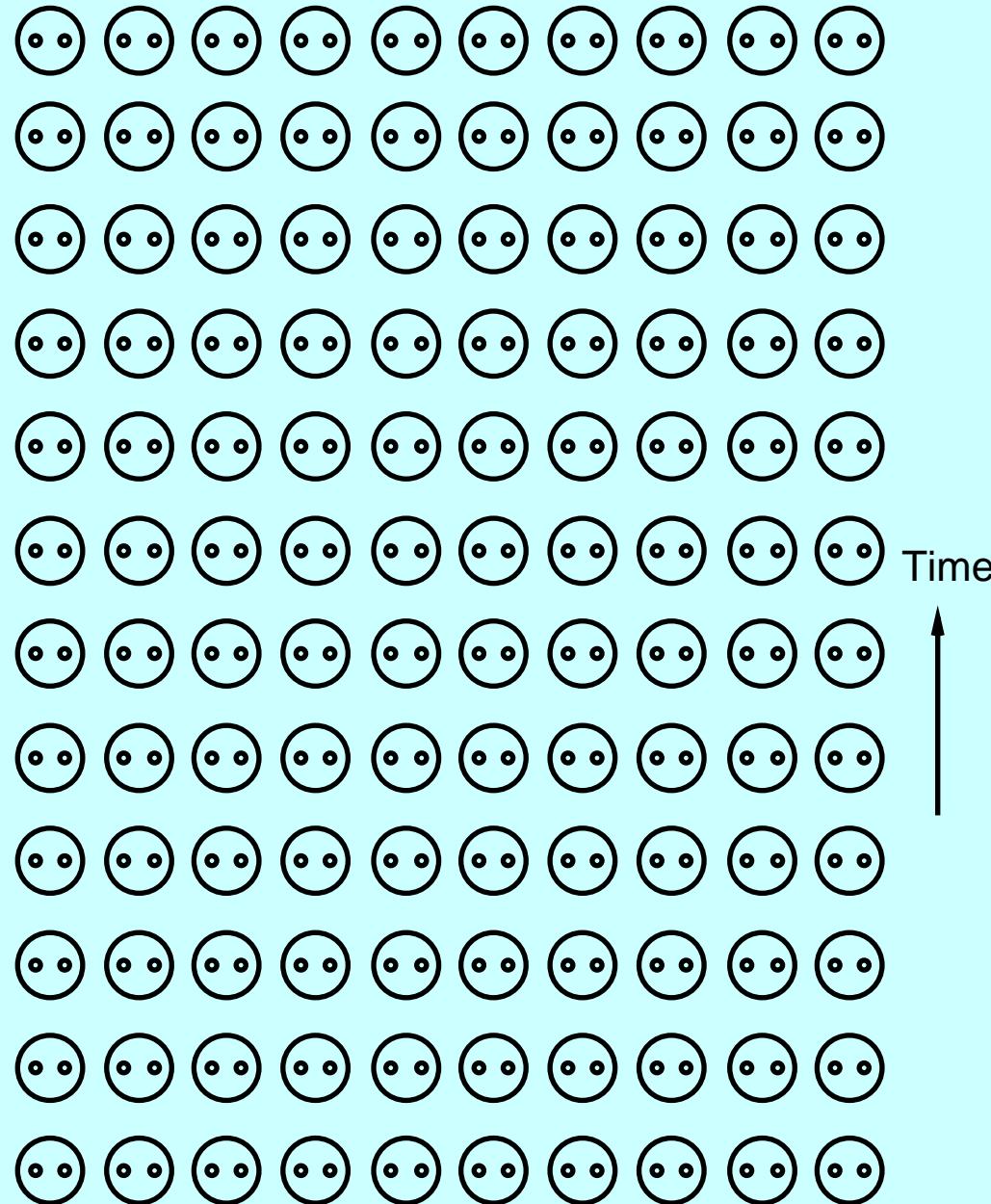
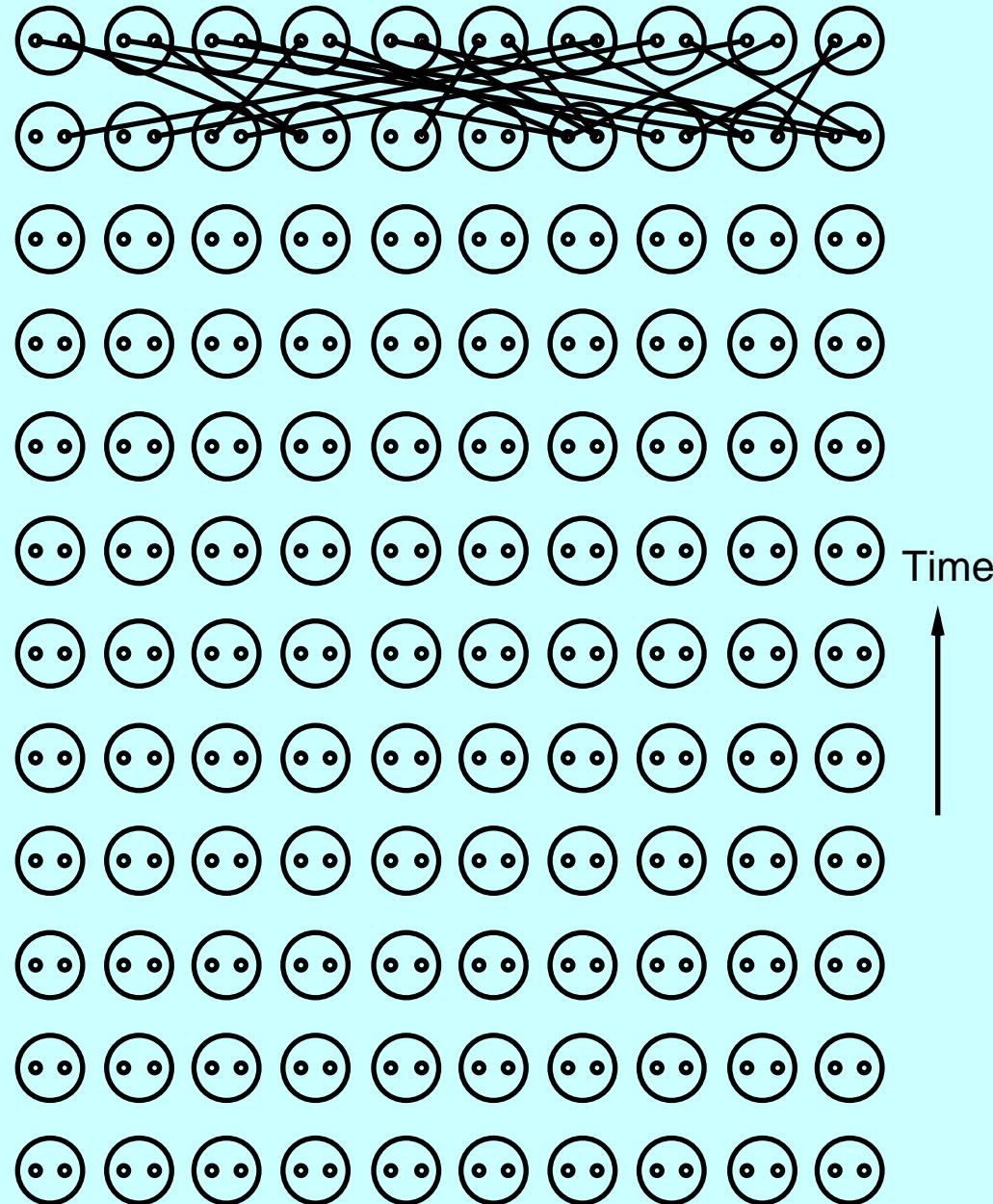


Fig. 3 a, Genealogical tree for 134 types of human mtDNA (133 restriction sites used. The tree accounts for the site differences observed between the various samples. The tree is rooted at the bottom center. The continent of origin of each sample is indicated by a symbol: ● Africa; ○ Asia; ▲ Australia; ▲ New Guinea; □ Europe. The tree is organized into hierarchical clades, labeled with lowercase letters: a, g, h, i, j, l, k, b, c, d, e, f, t. The tree is also marked with numerical values representing sequence divergence percentages, ranging from 1 to 130, along the outer and inner arcs. Two horizontal scales at the bottom indicate 'Sequence divergence %' for both the left and right sides of the tree.

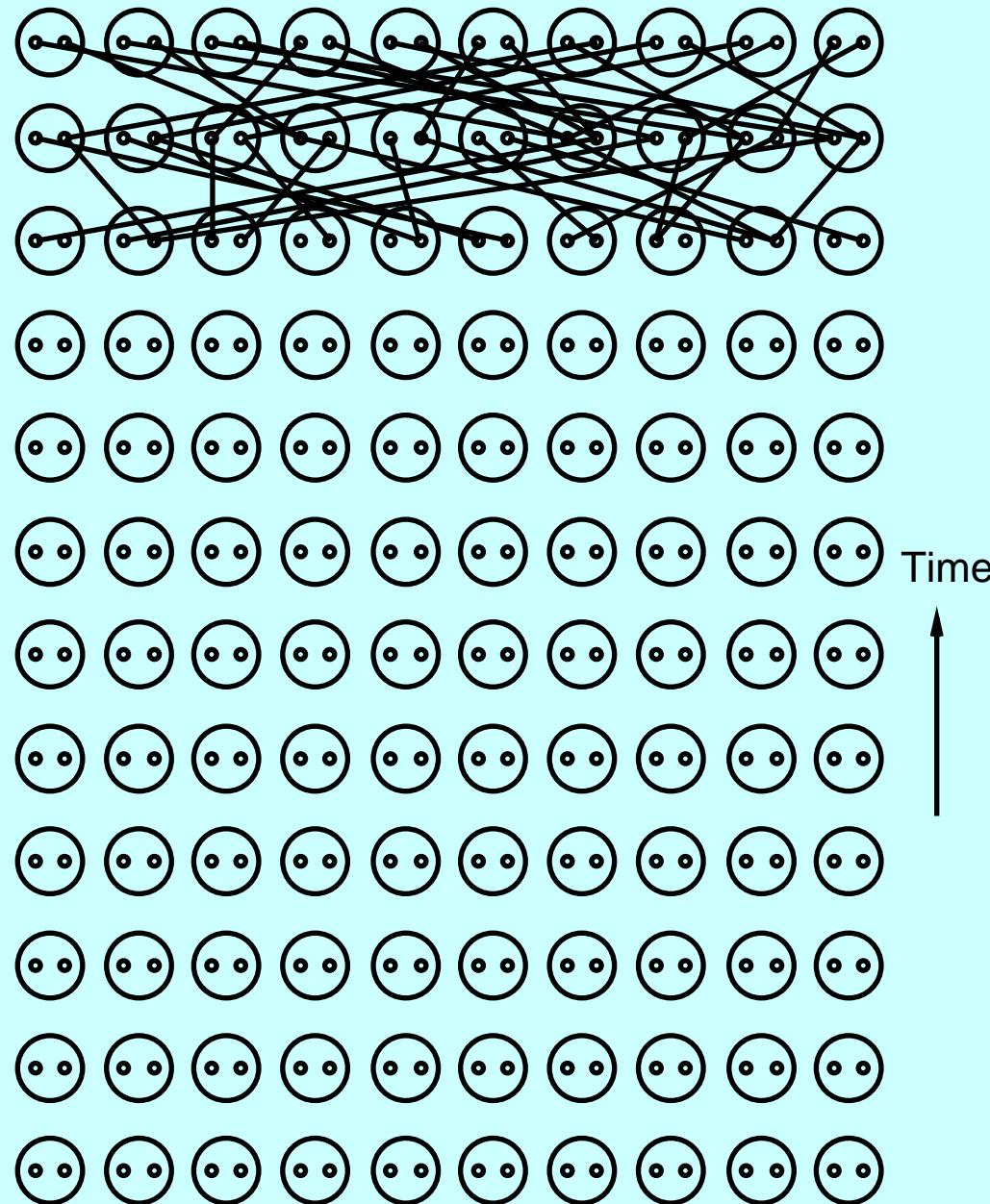
Genealogy of gene copies in a random-mating population



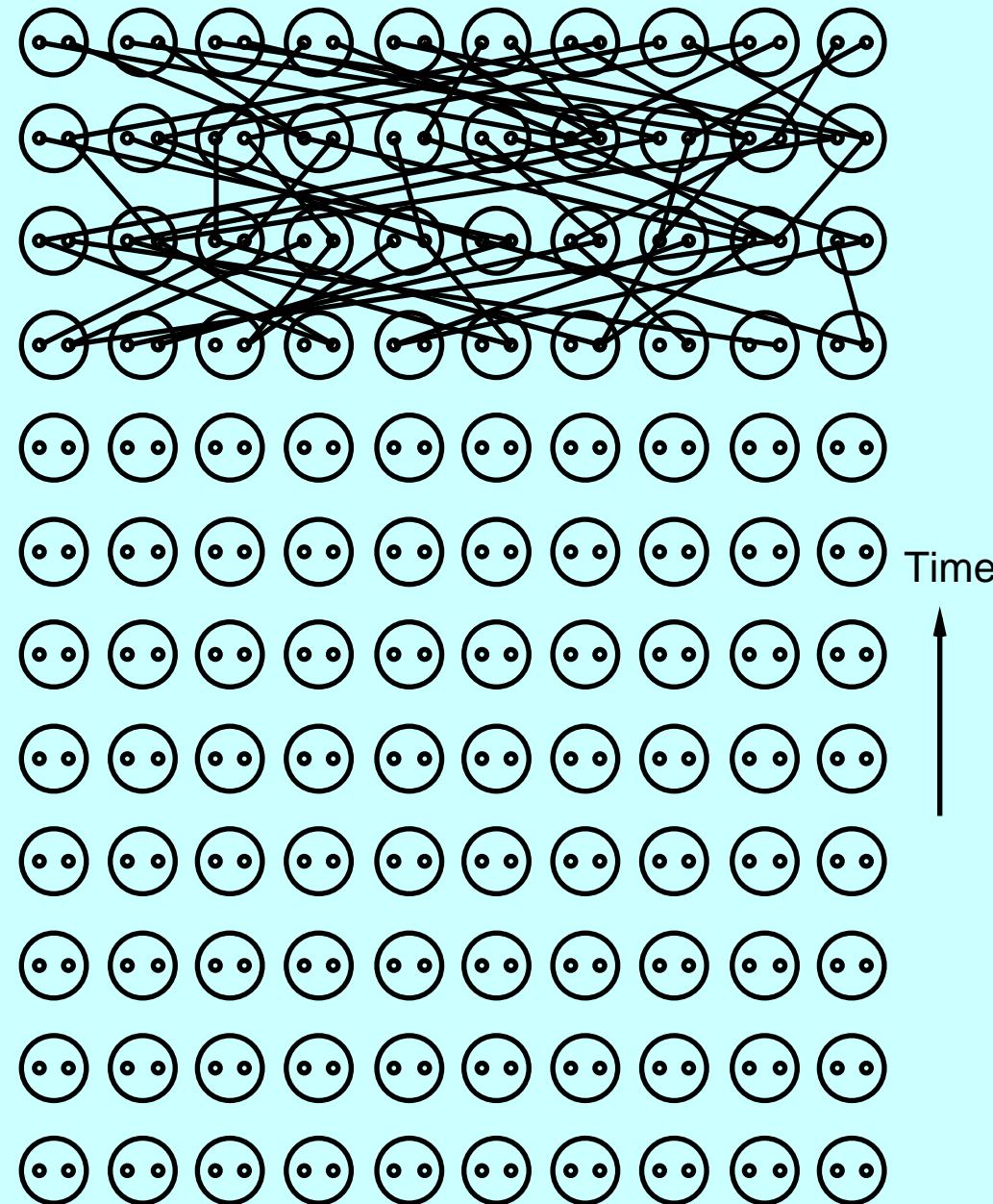
Genealogy of gene copies in a random-mating population



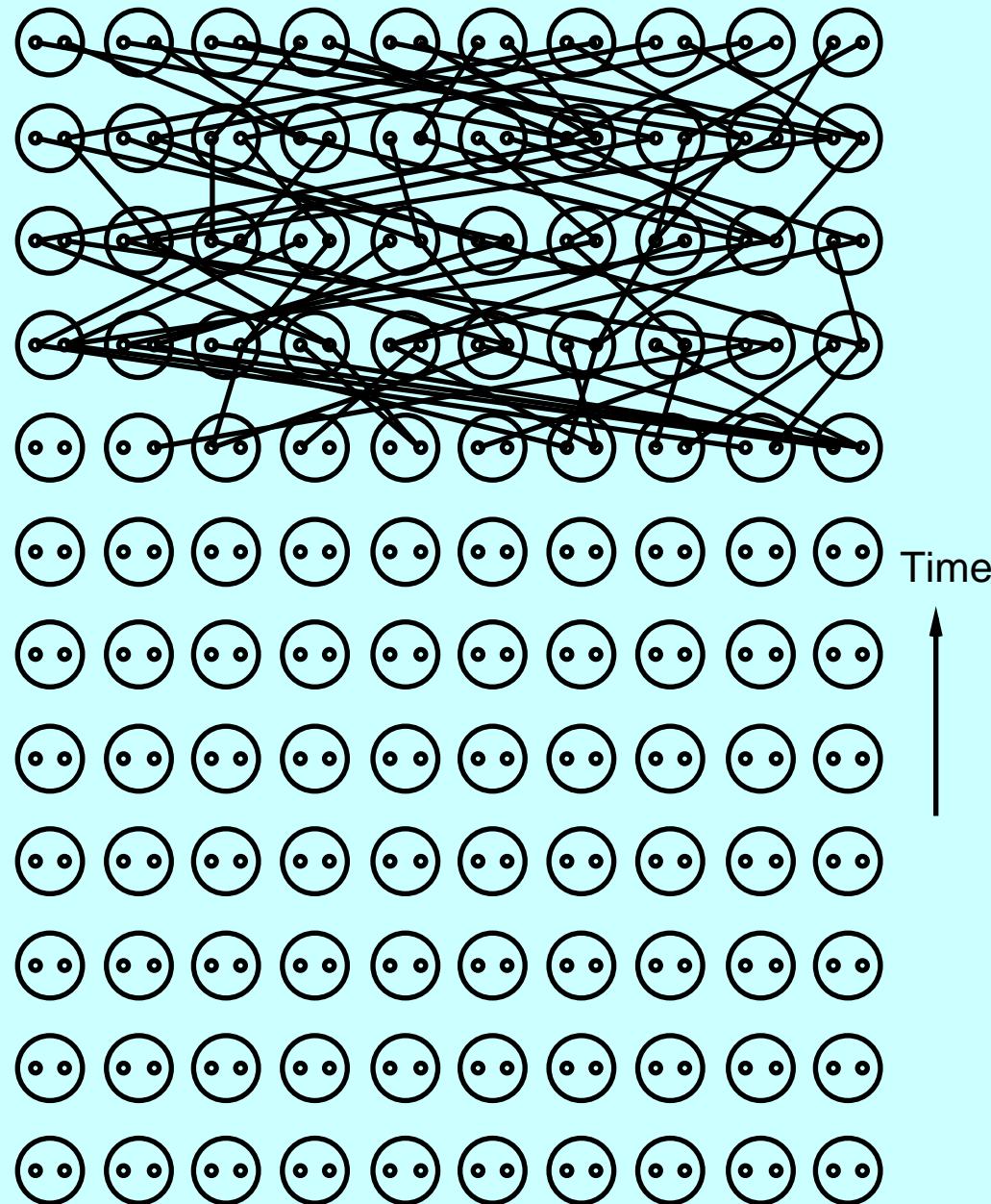
Genealogy of gene copies in a random-mating population



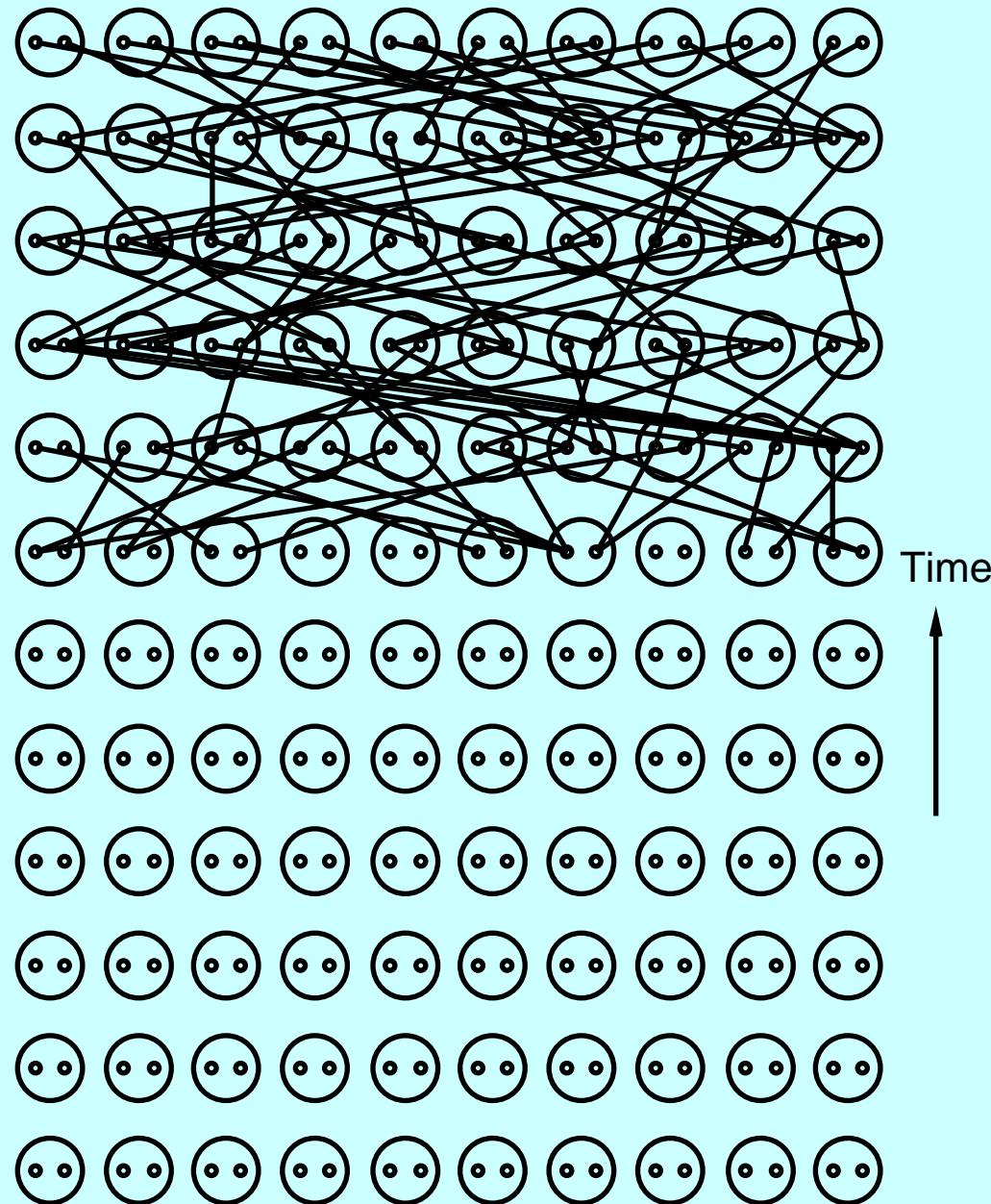
Genealogy of gene copies in a random-mating population



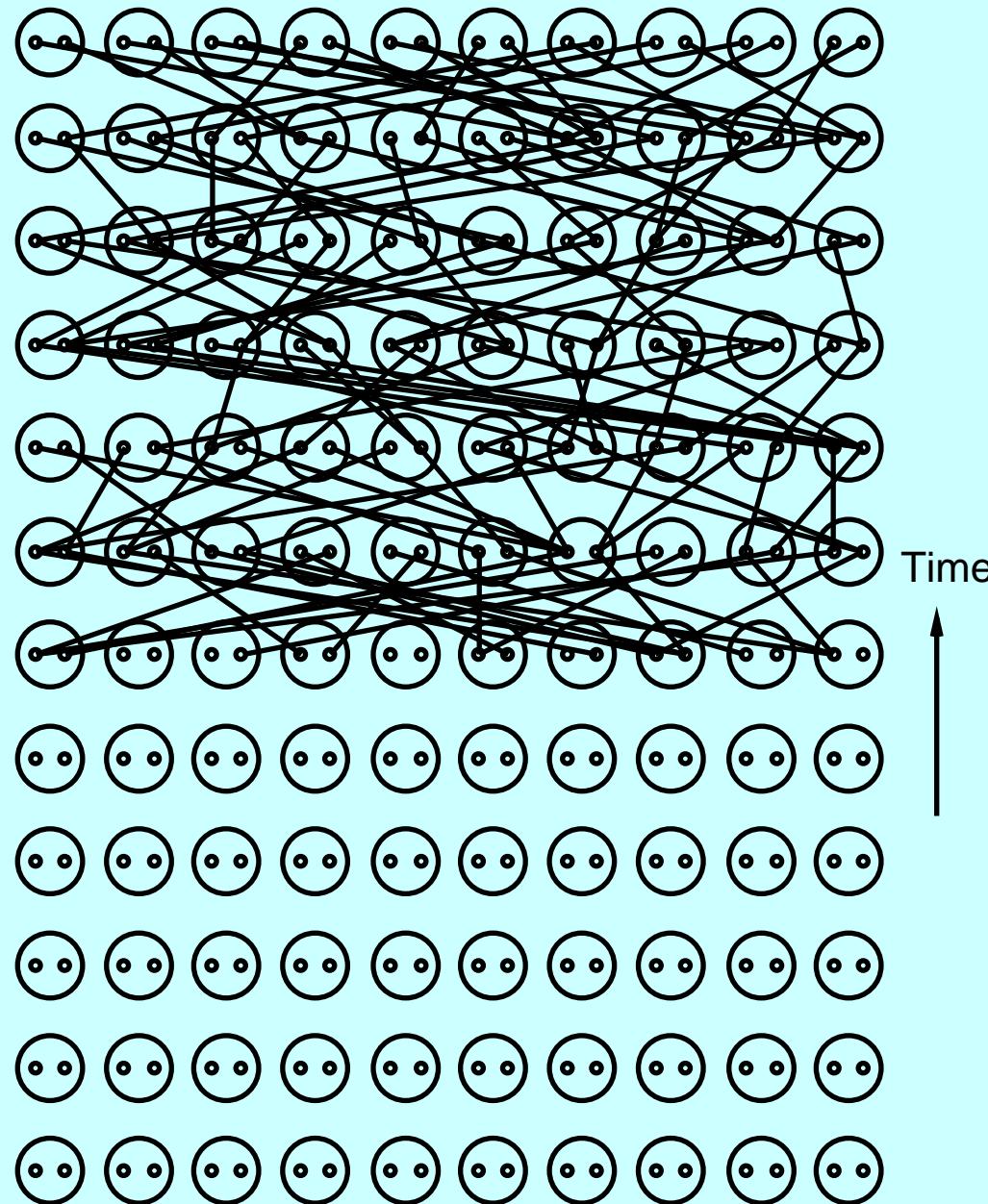
Genealogy of gene copies in a random-mating population



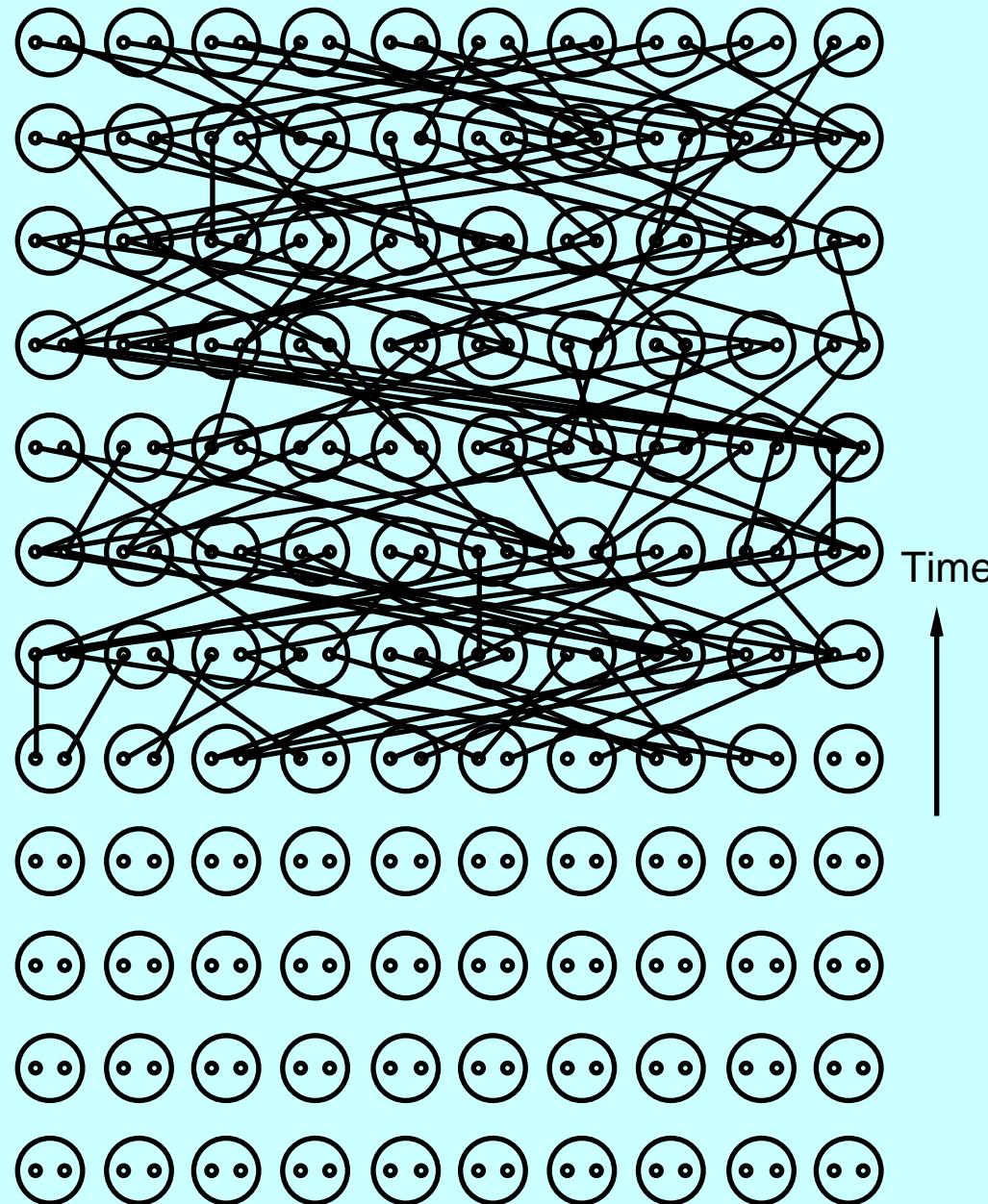
Genealogy of gene copies in a random-mating population



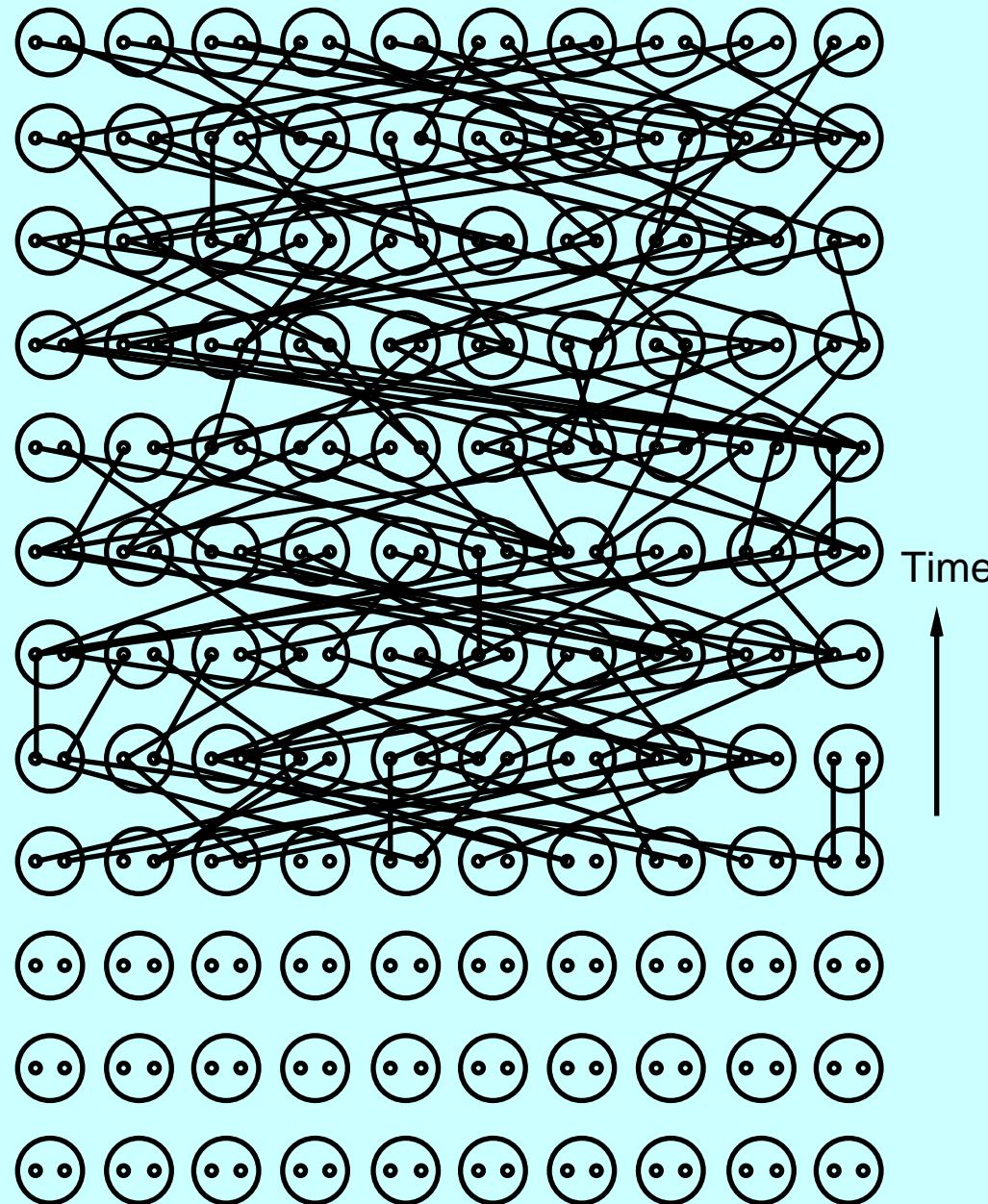
Genealogy of gene copies in a random-mating population



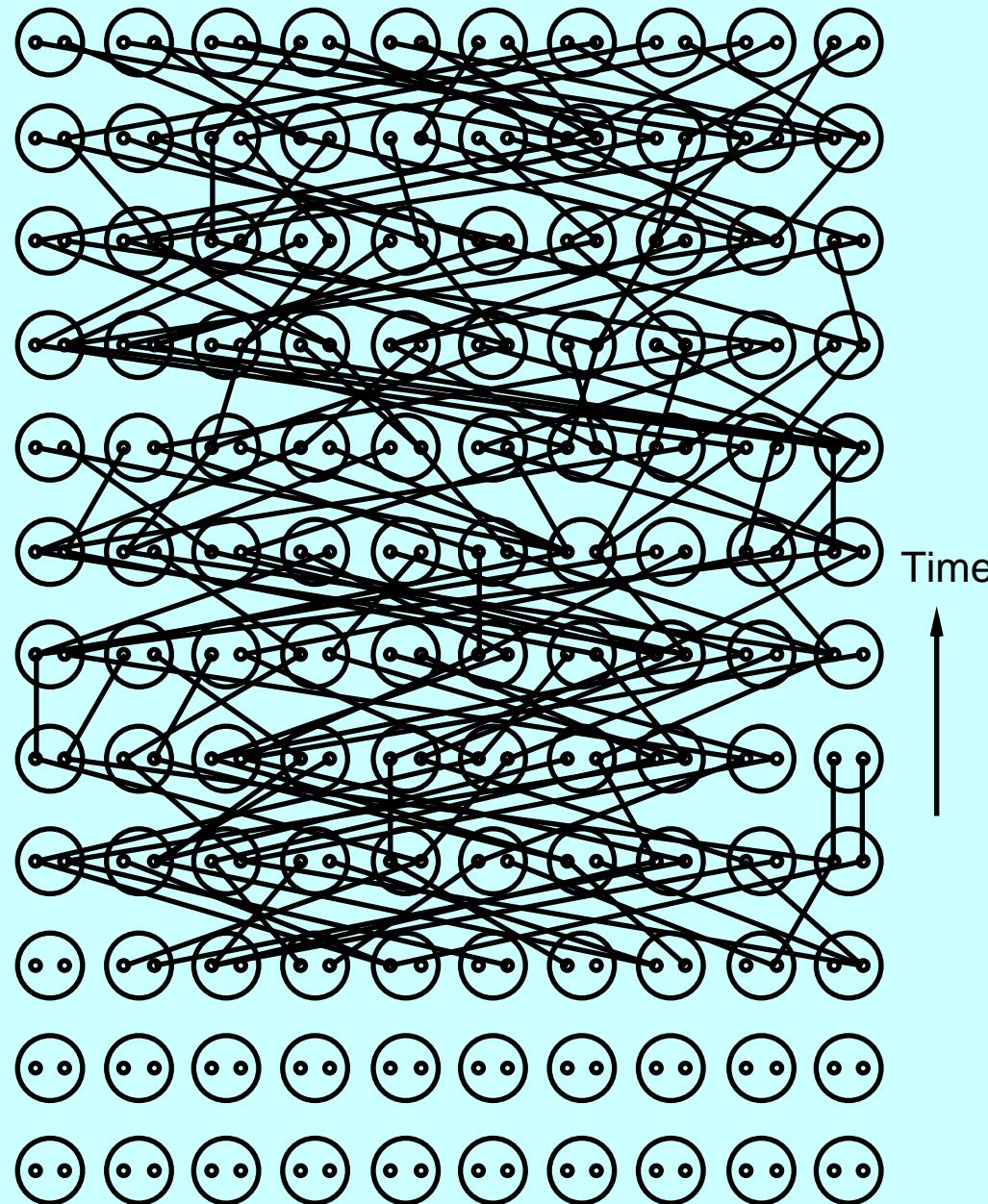
Genealogy of gene copies in a random-mating population



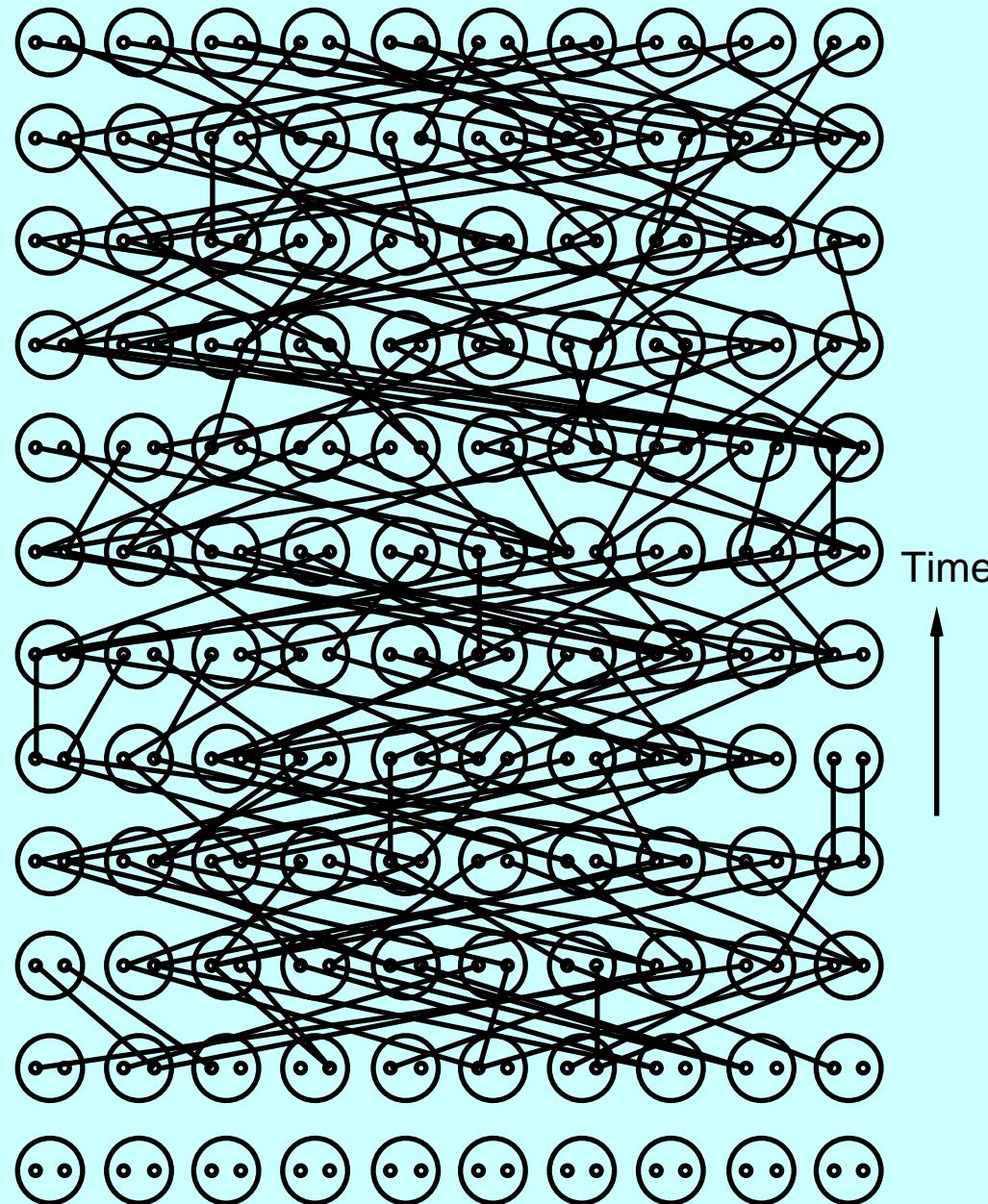
Genealogy of gene copies in a random-mating population



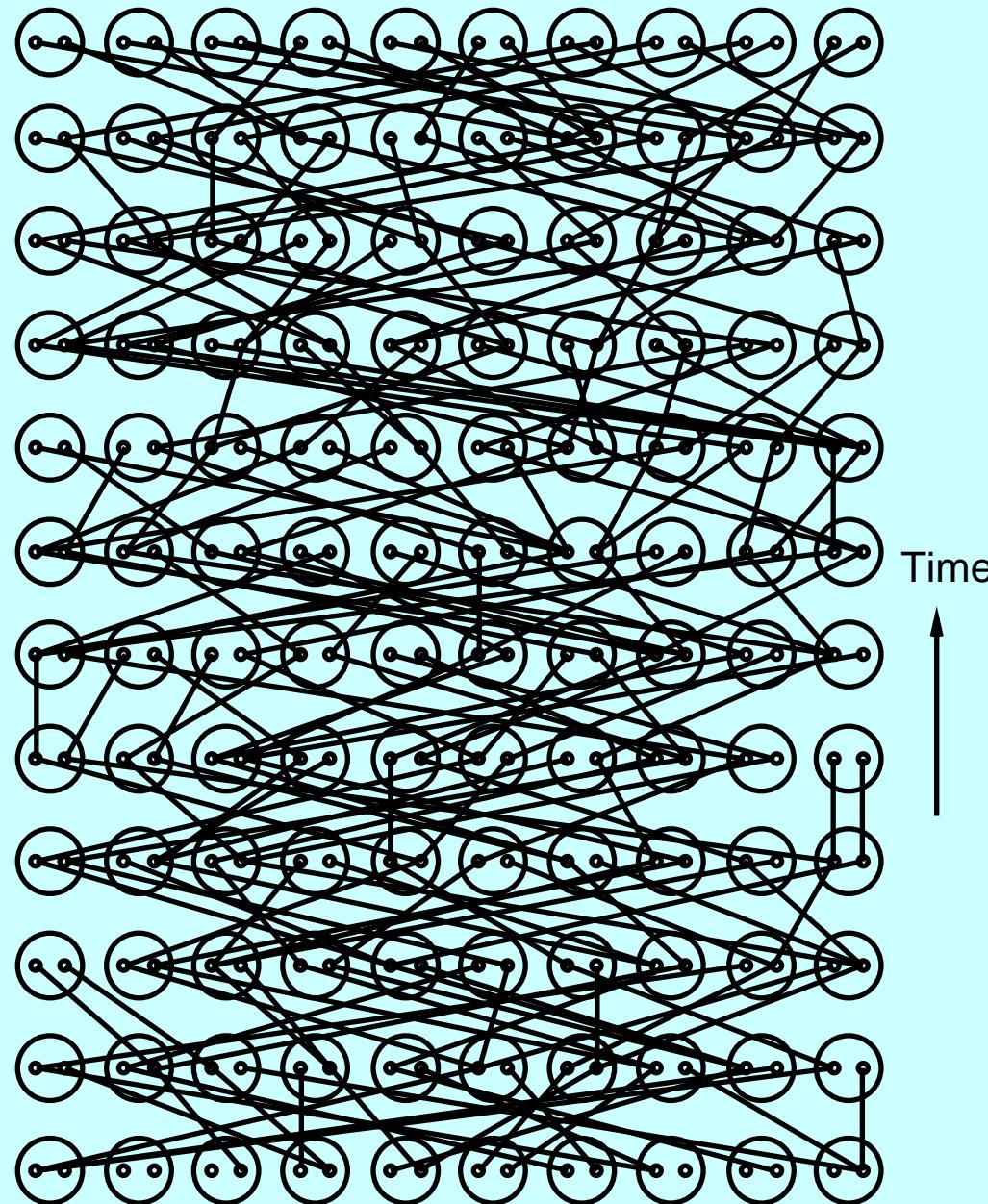
Genealogy of gene copies in a random-mating population



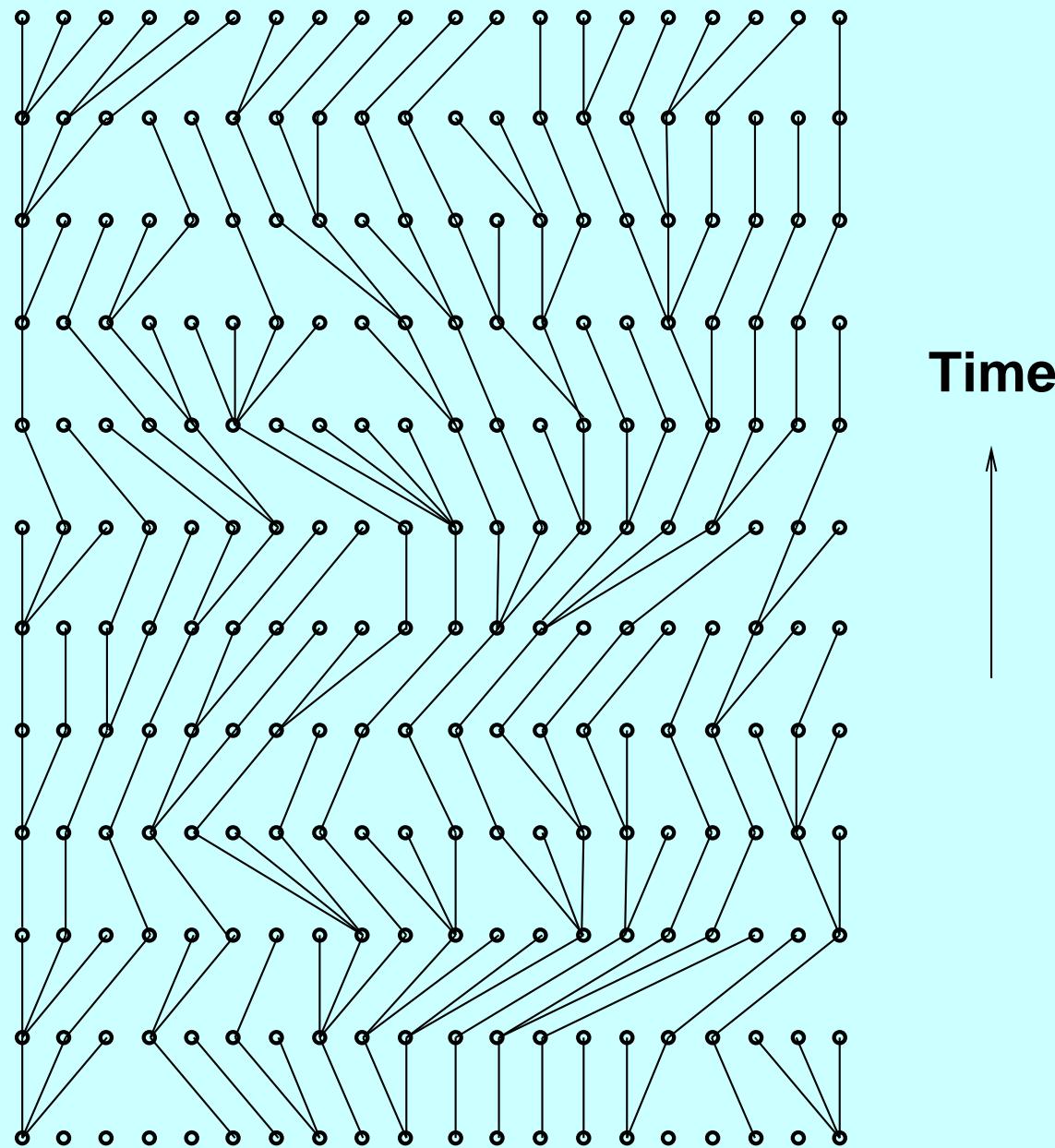
Genealogy of gene copies in a random-mating population



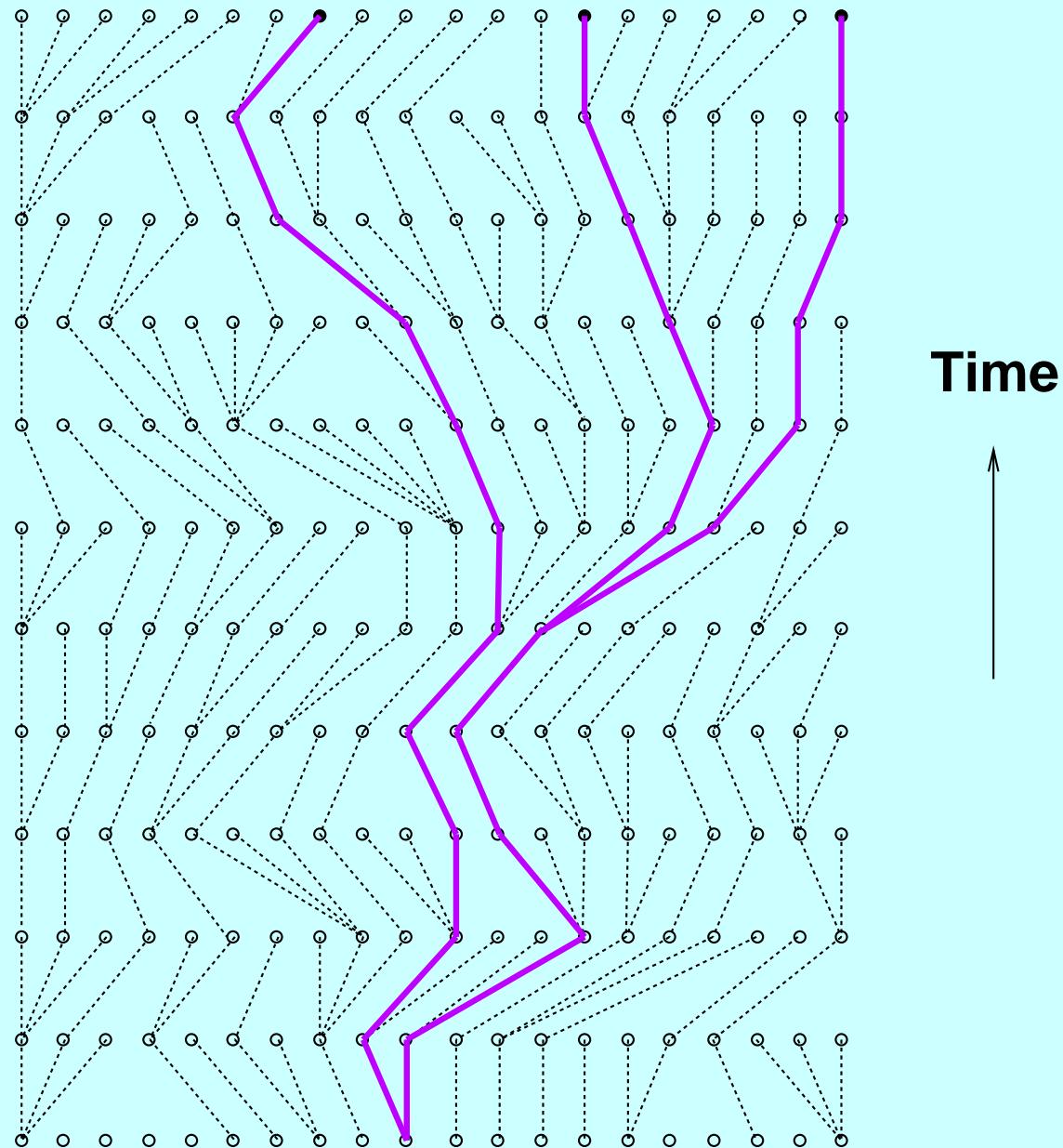
Genealogy of gene copies in a random-mating population



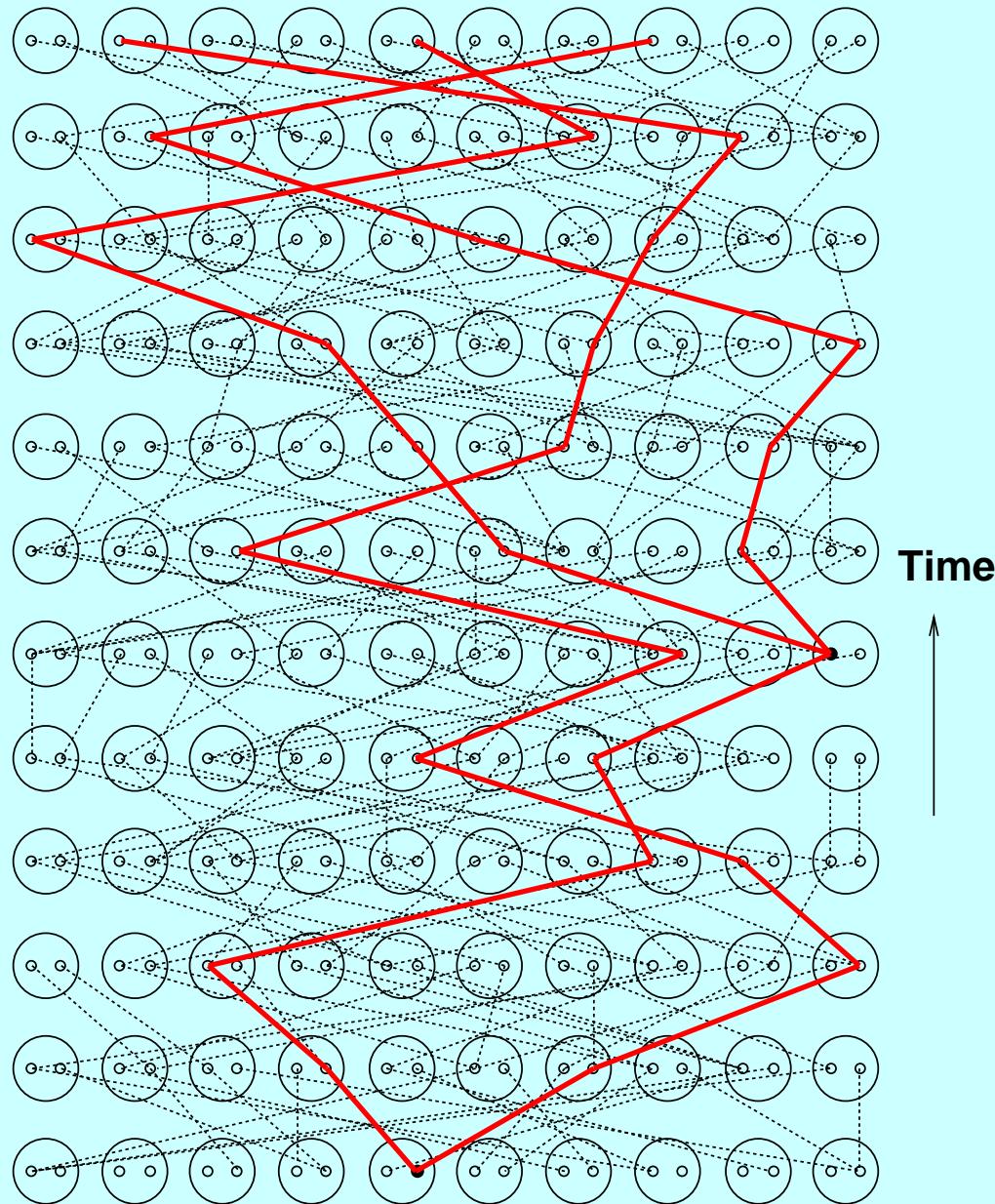
Untangling this genealogy by left-right swappings ...



Genealogy of a sample of gene copies



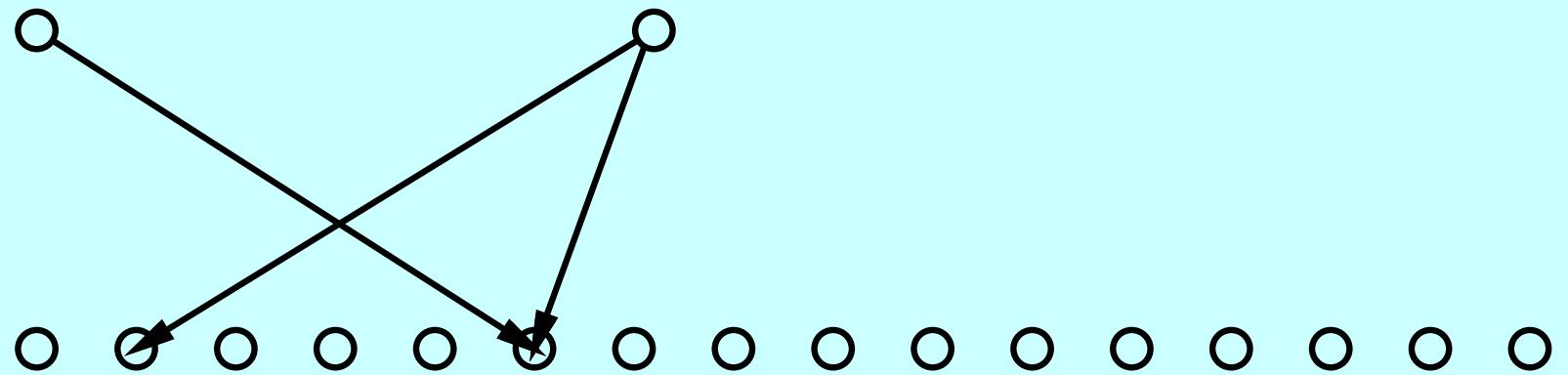
Ancestry of a sample in the population pedigree



Why lineages coalesce

under the Wright–Fisher model

each gene comes from a random copy
in the previous generation



a chance of 1 out of $2N$ that another
one comes from the same copy

hence it takes about $2N$ generations for
two lineages to coalesce

Sir John Kingman

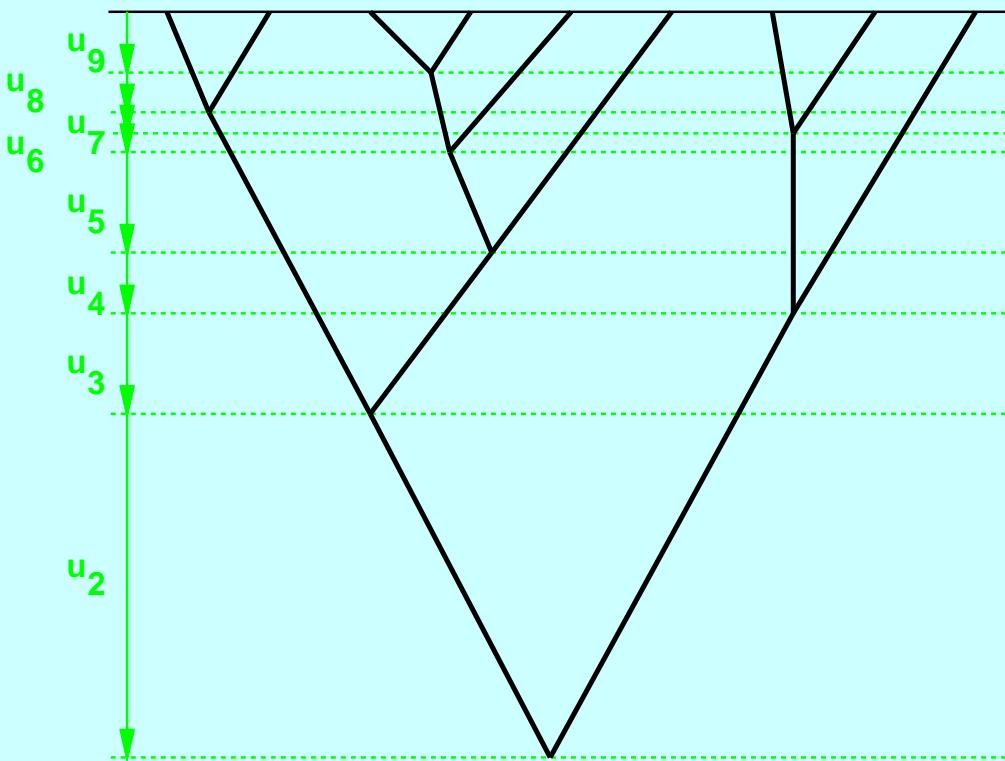


J. F. C. Kingman in about 1983

Currently Emeritus Professor of Mathematics at Cambridge University, U.K., and former head of the Isaac Newton Institute of Mathematical Sciences.

Kingman's coalescent

Random collision of lineages as go back in time (sans recombination)
Collision is faster the smaller the effective population size



In a diploid population of effective population size N_e
Average time for k copies to coalesce to $k-1$ = $\frac{4N_e}{k(k-1)}$

Average time for two copies to coalesce = $2N_e$ generations

Average time for n copies to coalesce = $4N_e \left(1 - \frac{1}{n}\right)$ generations

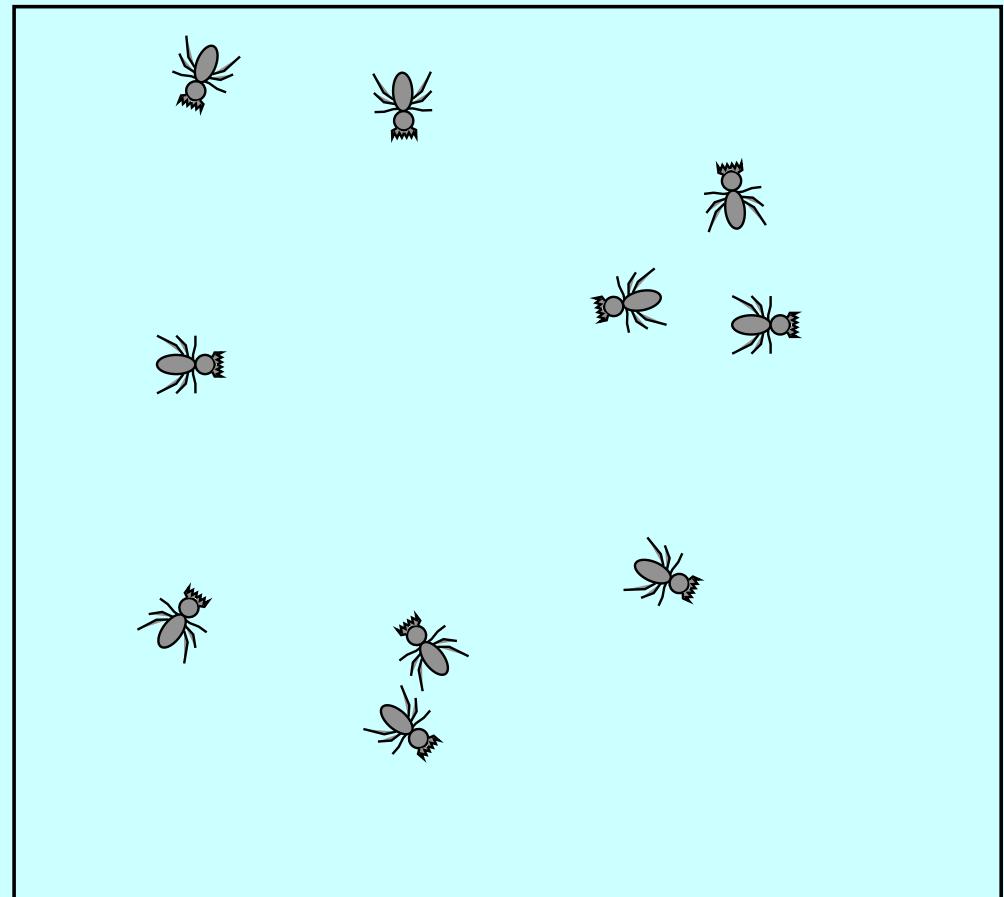
An accurate analogy: Bugs In A Box

There is a box ...



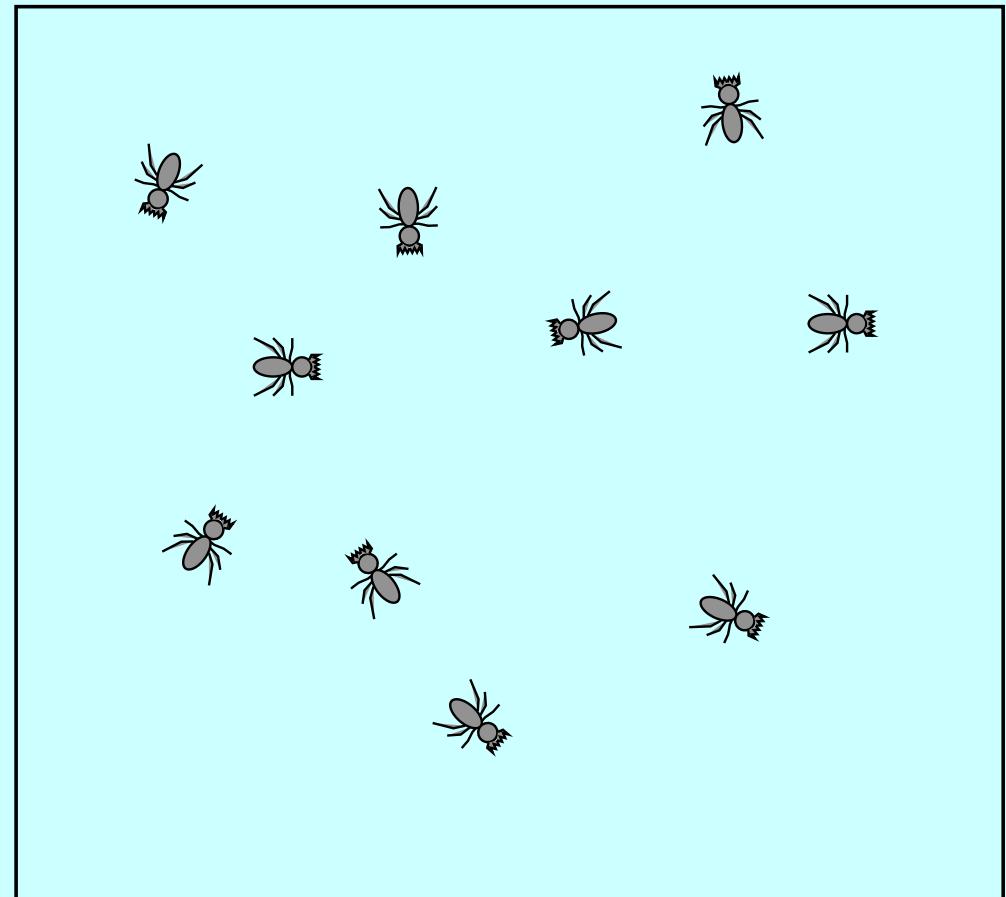
An accurate analogy: Bugs In A Box

with bugs that are ...



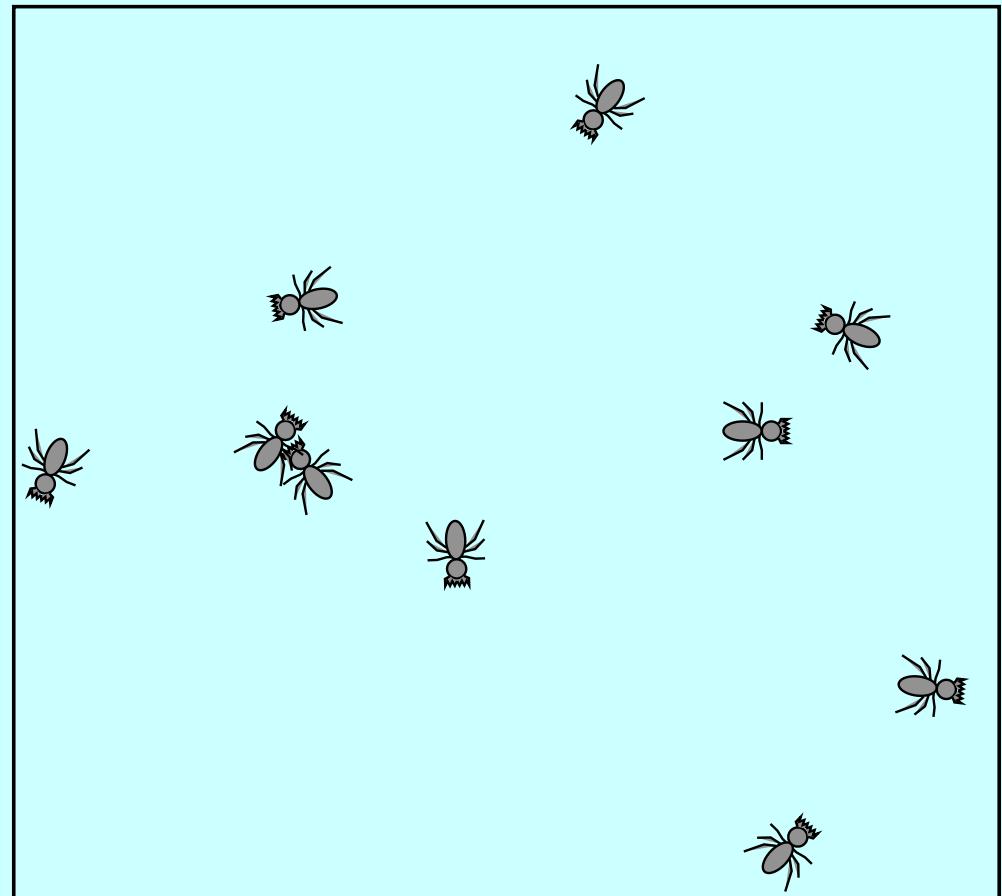
An accurate analogy: Bugs In A Box

... hyperactive ...



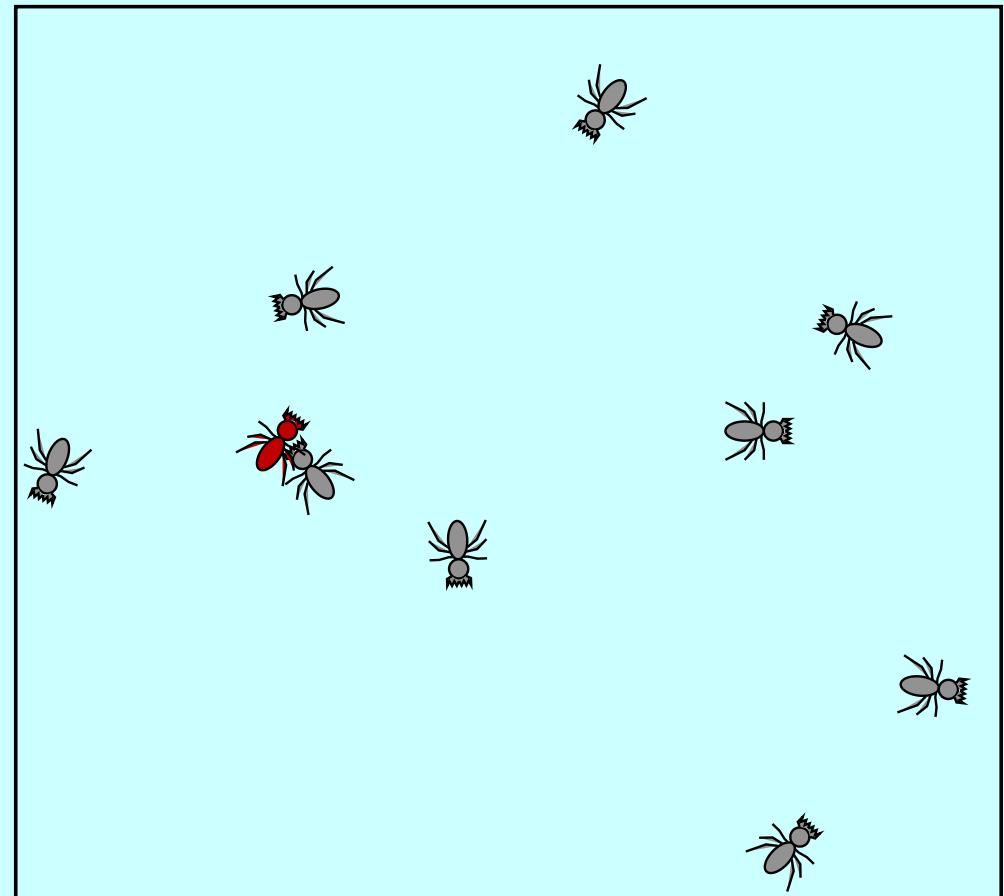
An accurate analogy: Bugs In A Box

... indiscriminate ...



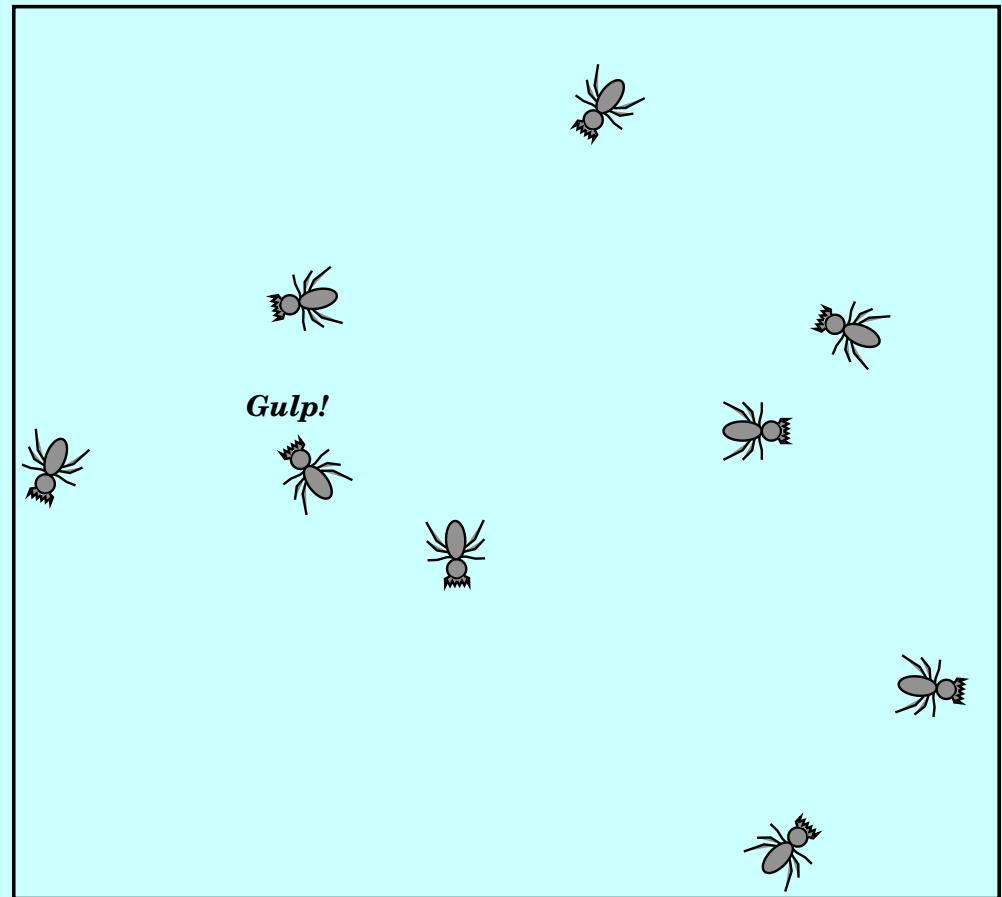
An accurate analogy: Bugs In A Box

... voracious ...



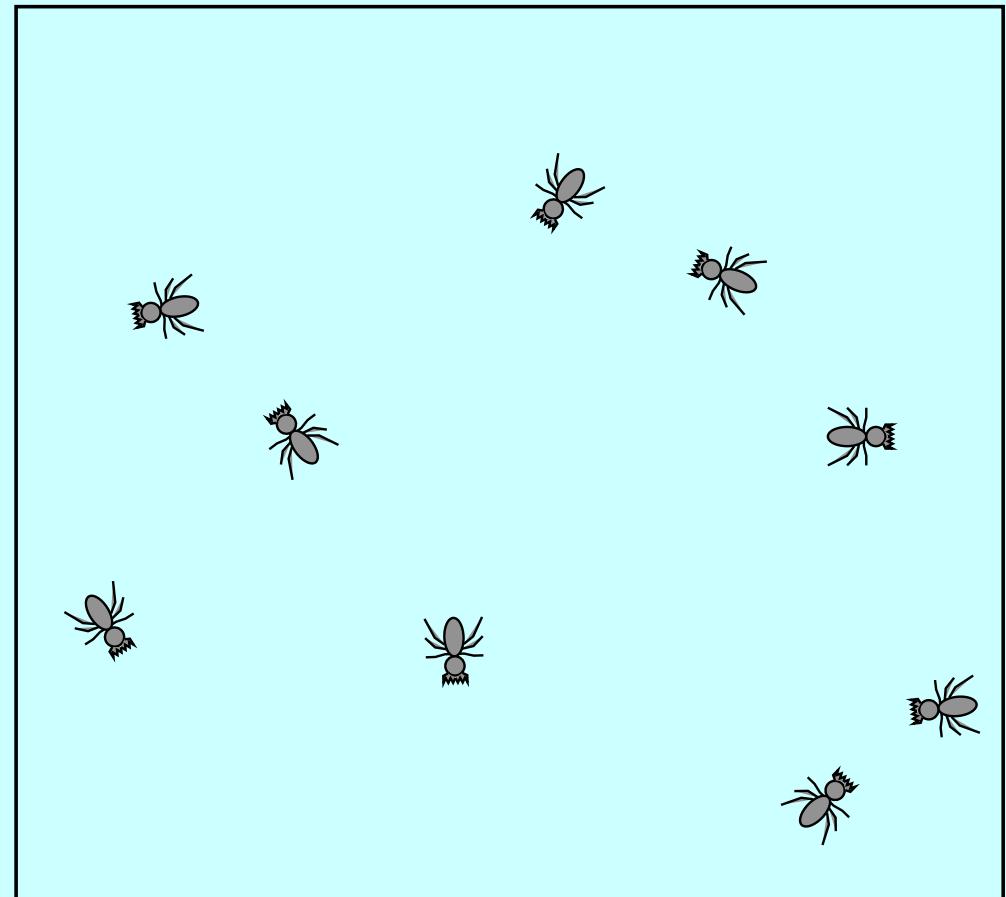
An accurate analogy: Bugs In A Box

... (*eats other bug*) ...

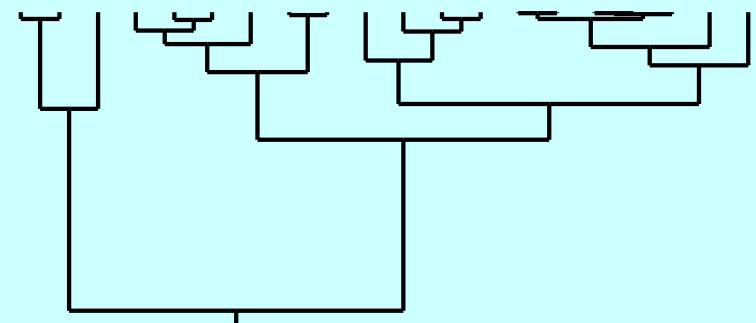
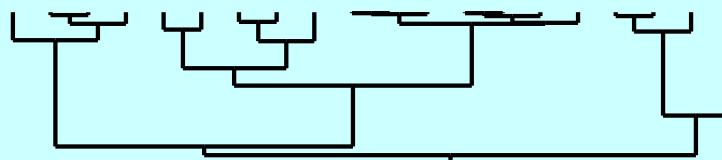
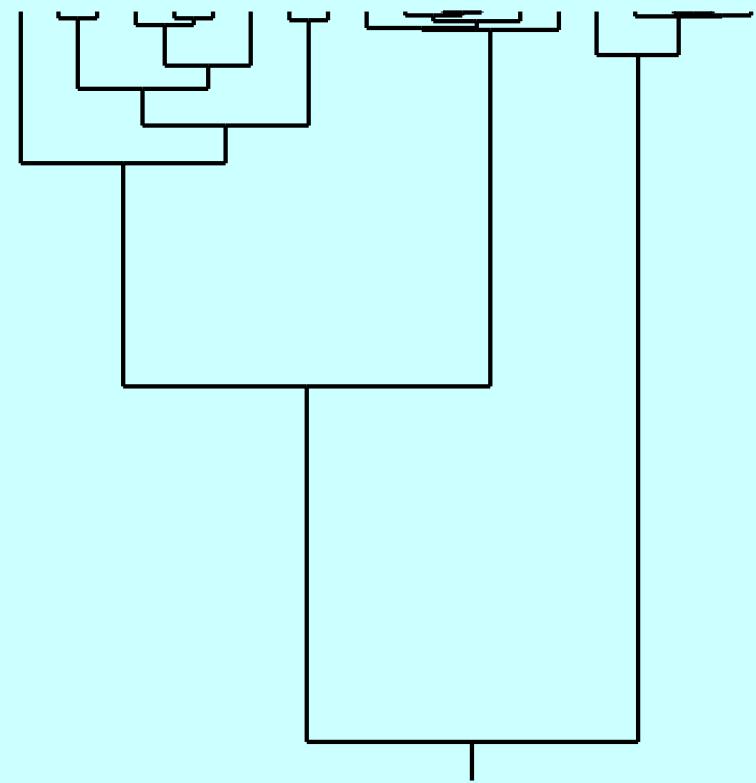
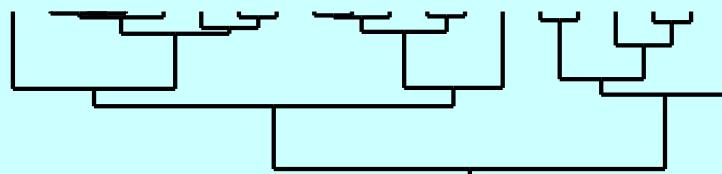


An accurate analogy: Bugs In A Box

... and insatiable.

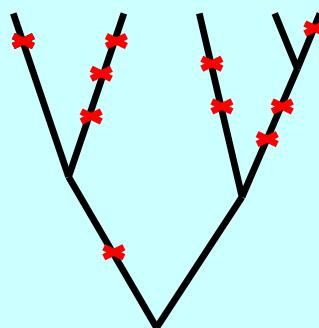


Random coalescents from the same population



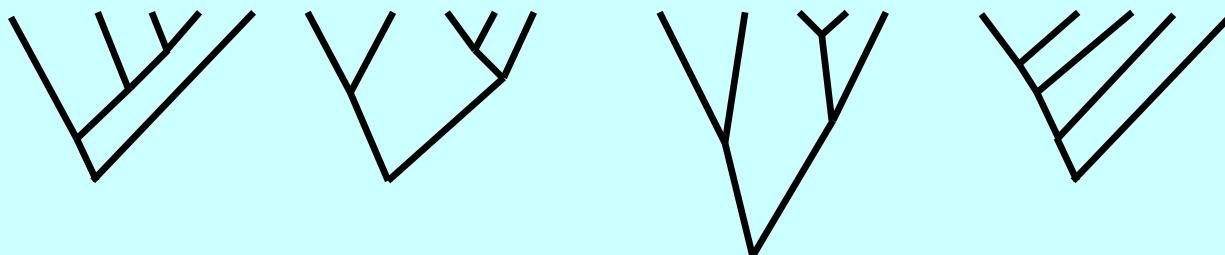
Two sources of variability

(1) Randomness of mutation



affected by the mutation rate u
can reduce variance of
number of mutations per site per
branch by examining more sites

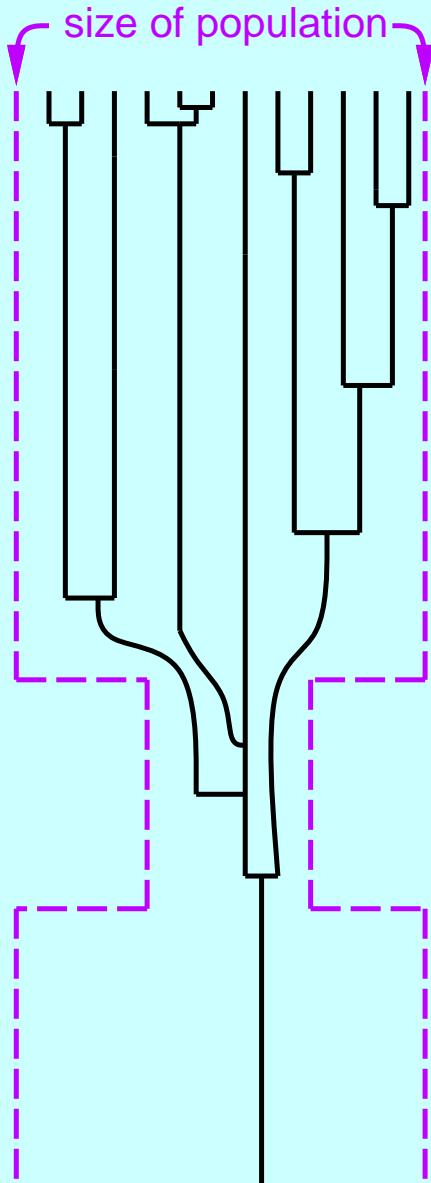
(2) Randomness of coalescence of lineages



affected by effective population size N_e
coalescence times allow estimation of N_e
can reduce variability by looking at

- (i) more gene copies, or
- (ii) more loci

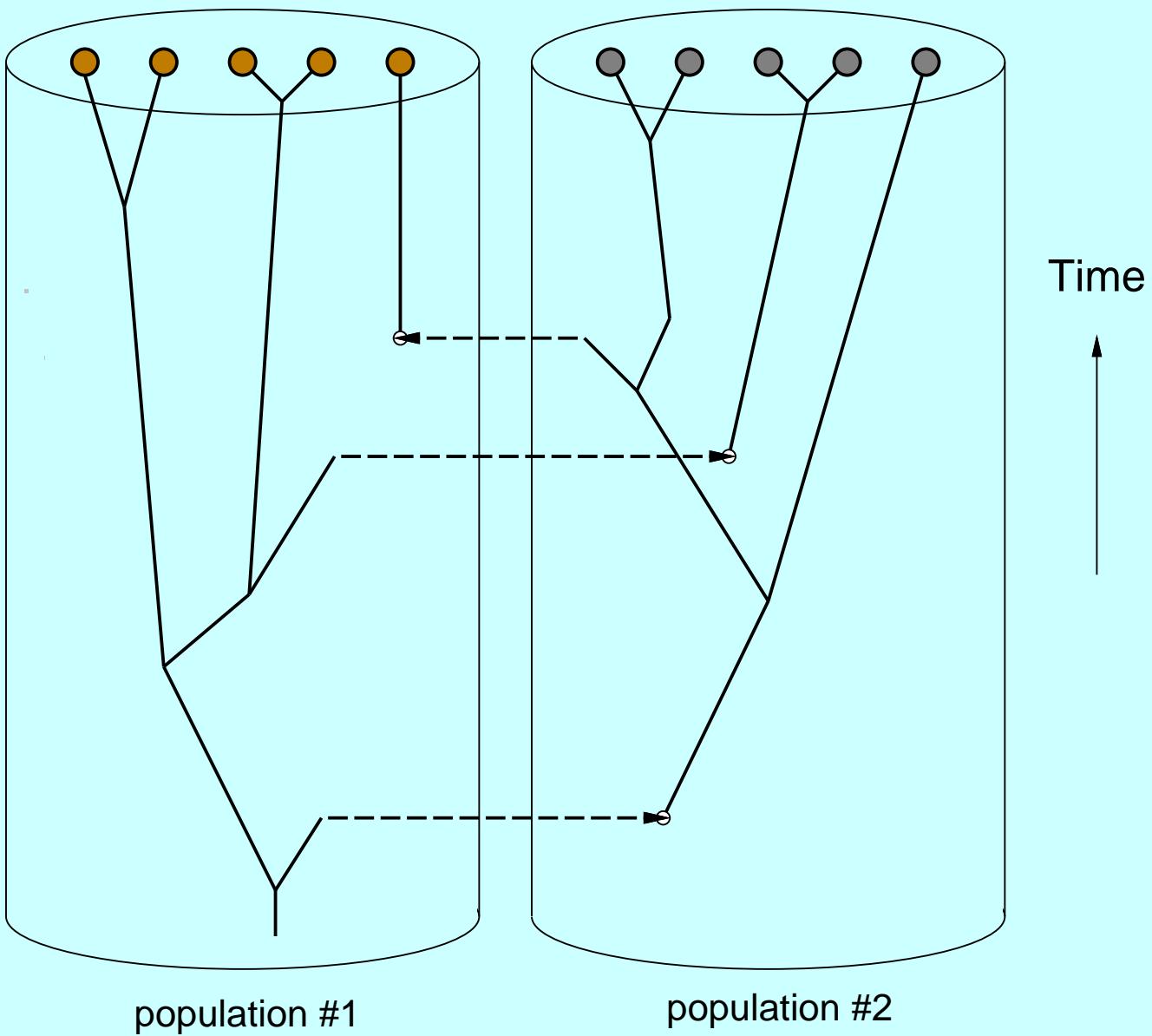
Coalescent with population growth



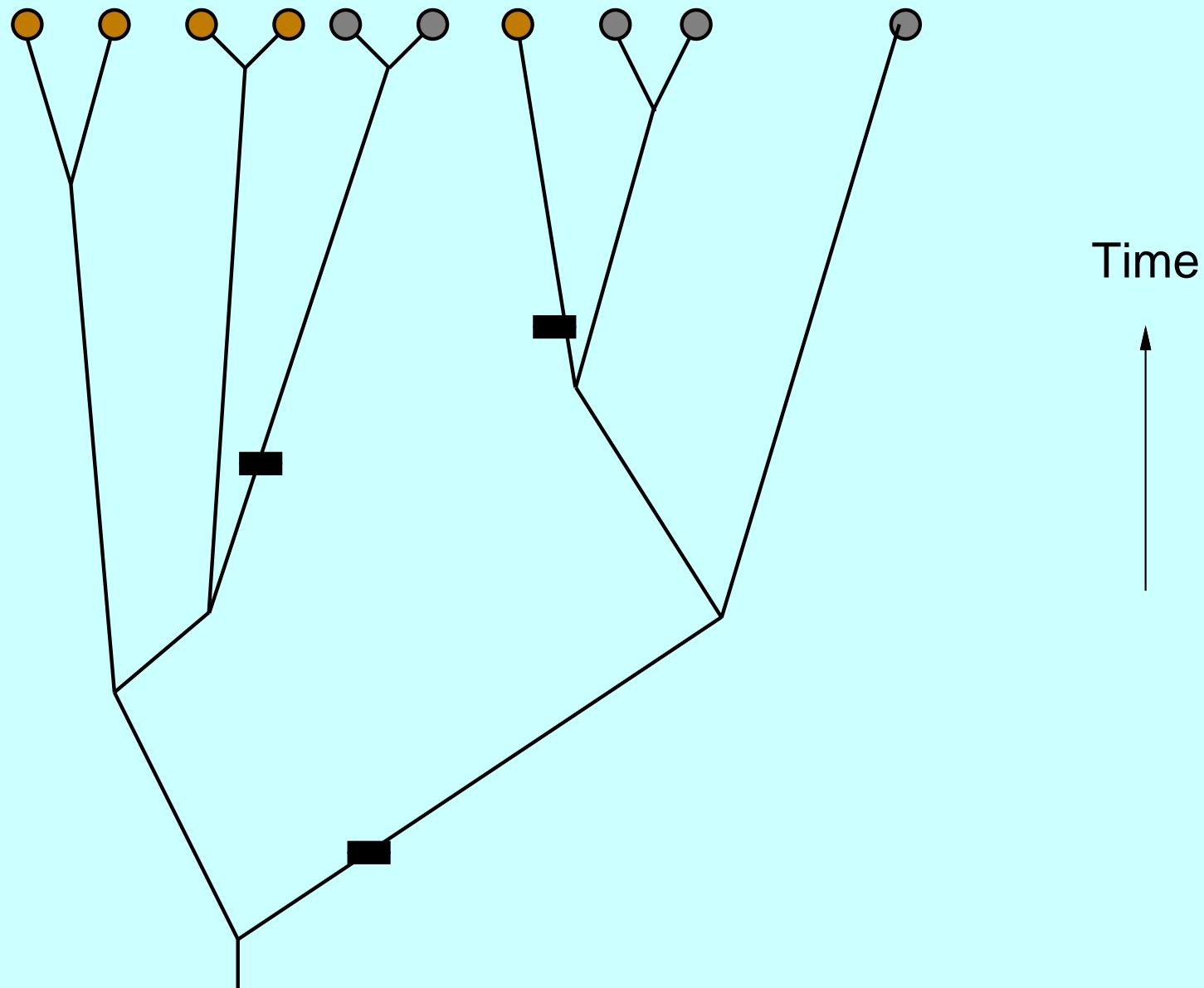
During a population bottleneck
there is expected to be a
burst of coalescence

Down near the root of the tree
effects of population size
become difficult to see

Coalescent with migration



Coalescent with migration, untangled



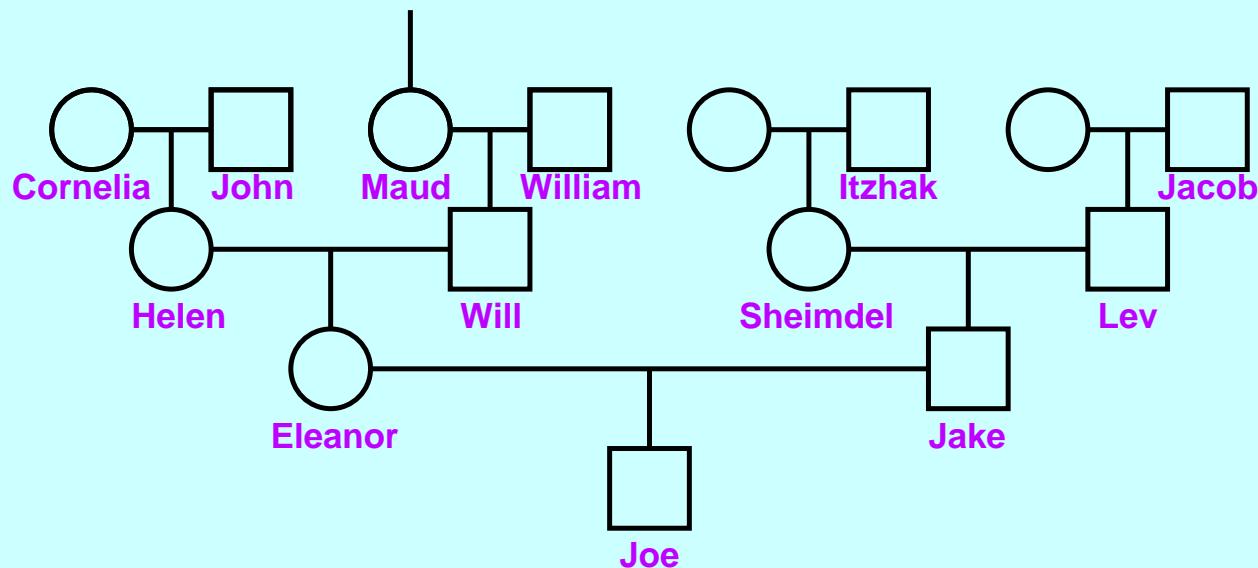
My ancestor?

Charles the Great
(Charlemagne)



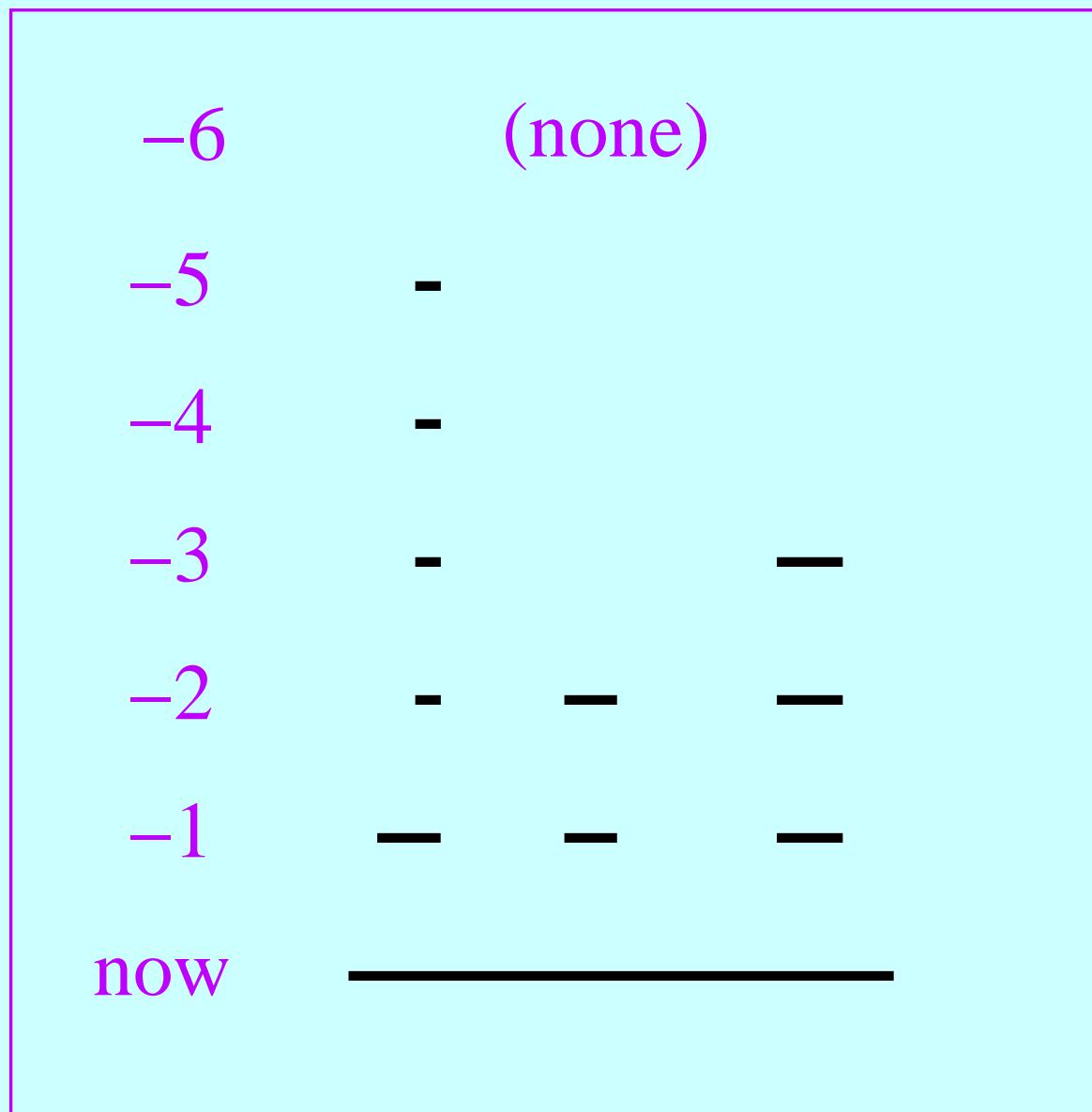
born
747

about 44 more generations



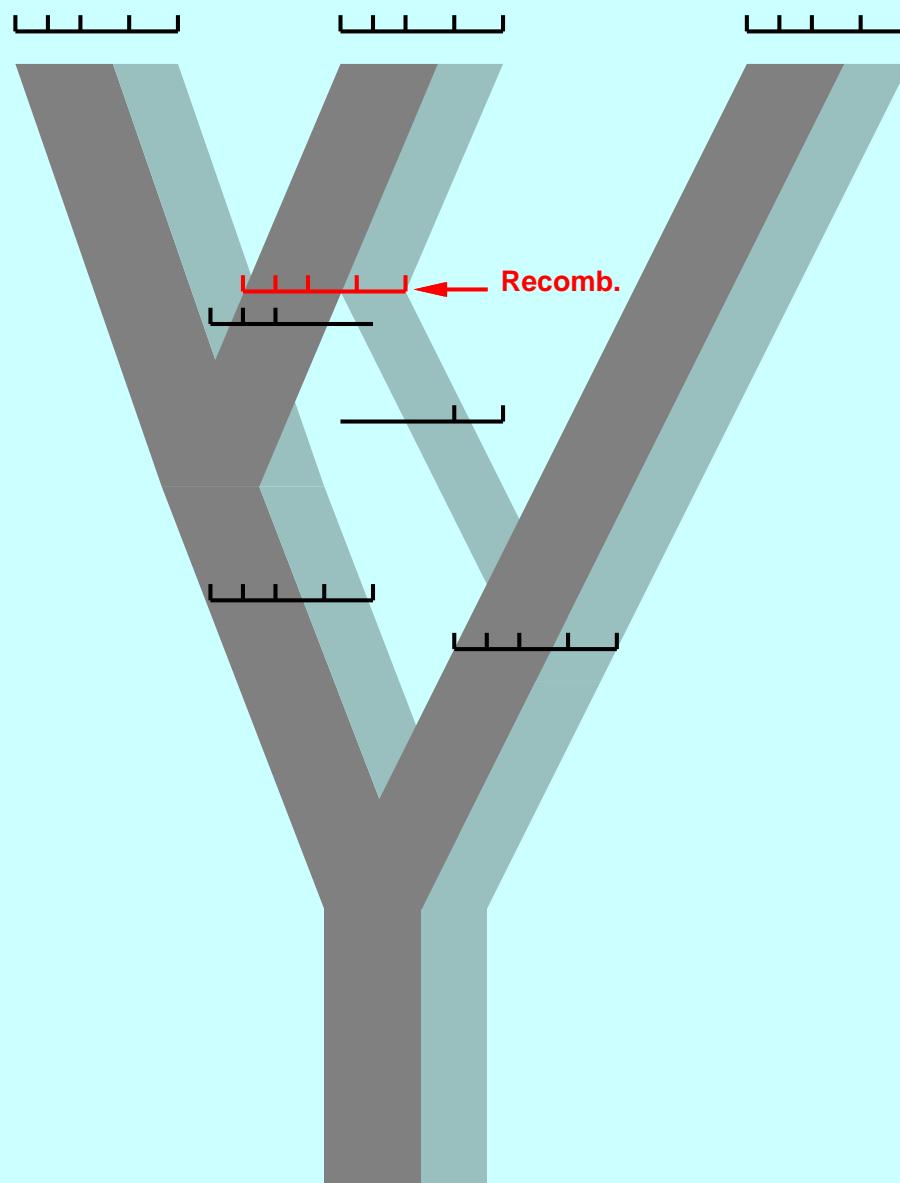
1850s
1880s
1910s
1942

Chromosome 1, back up one lineage

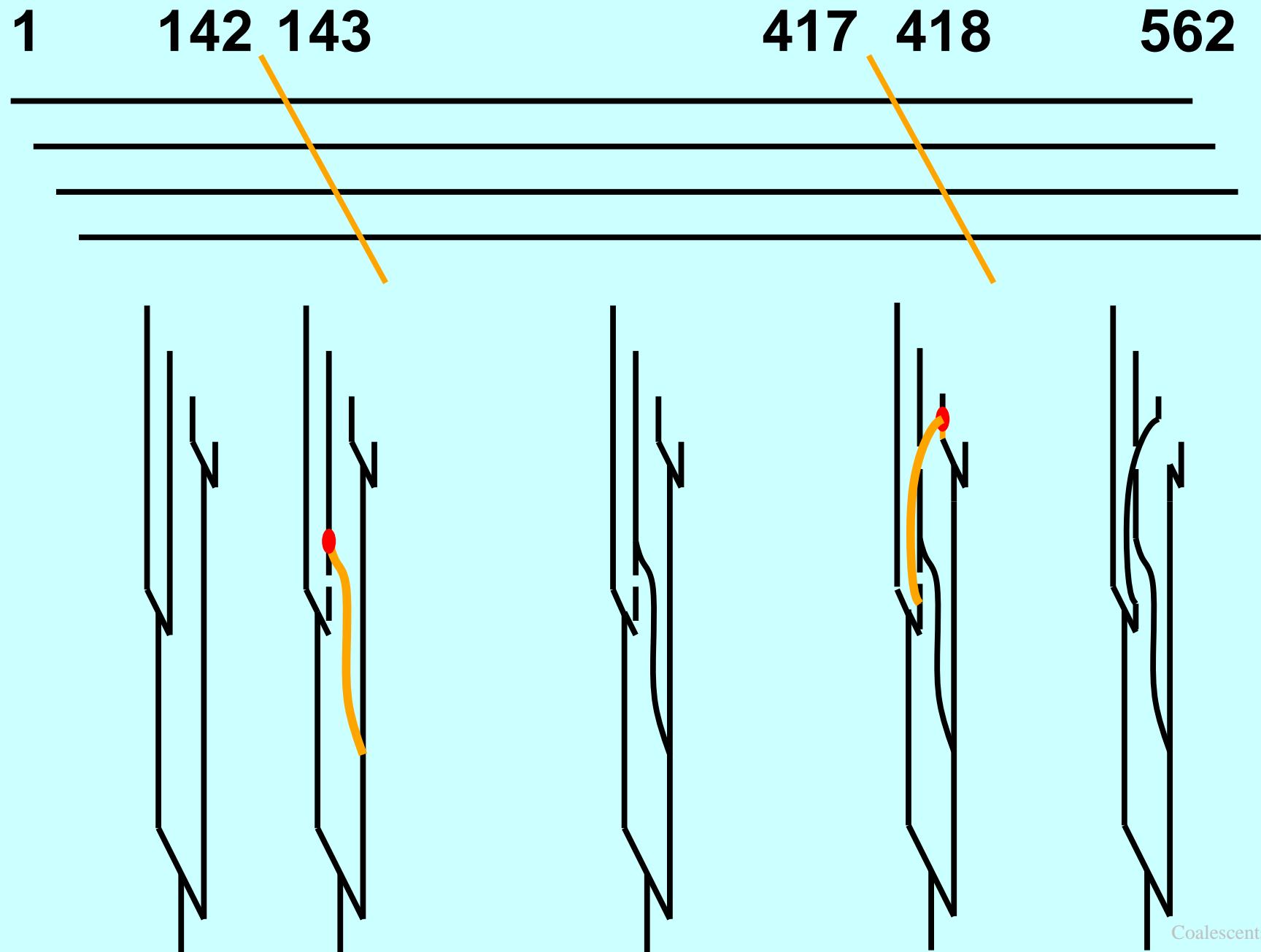


Coalescent with recombination

Different markers have slightly different coalescent trees

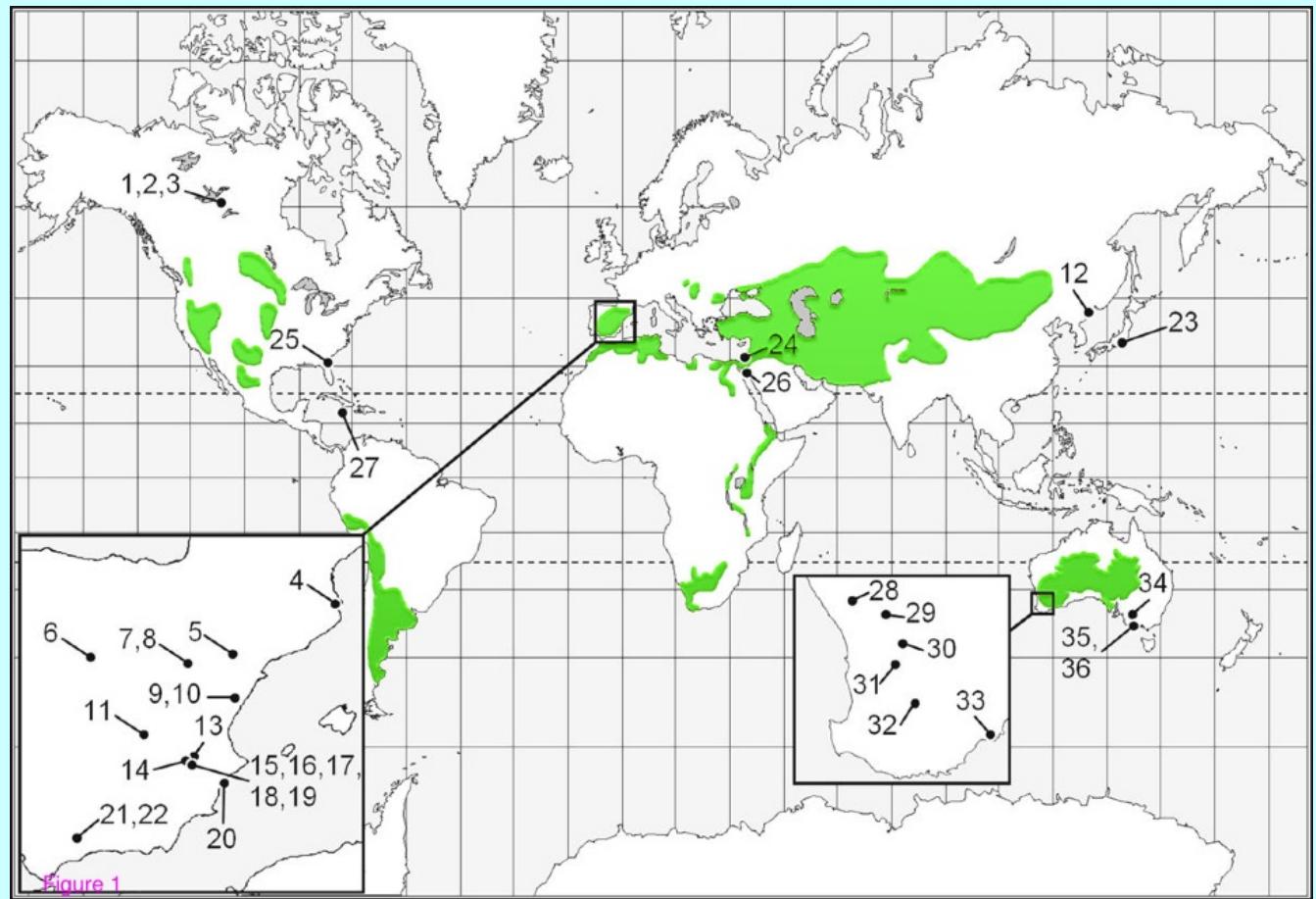
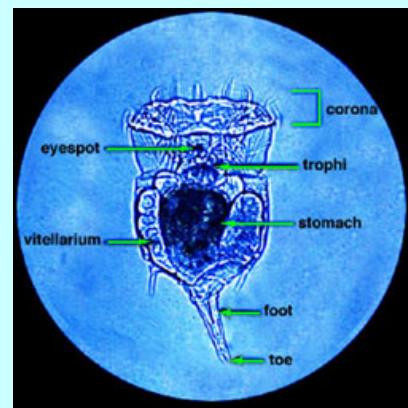


Coalescent with recombination



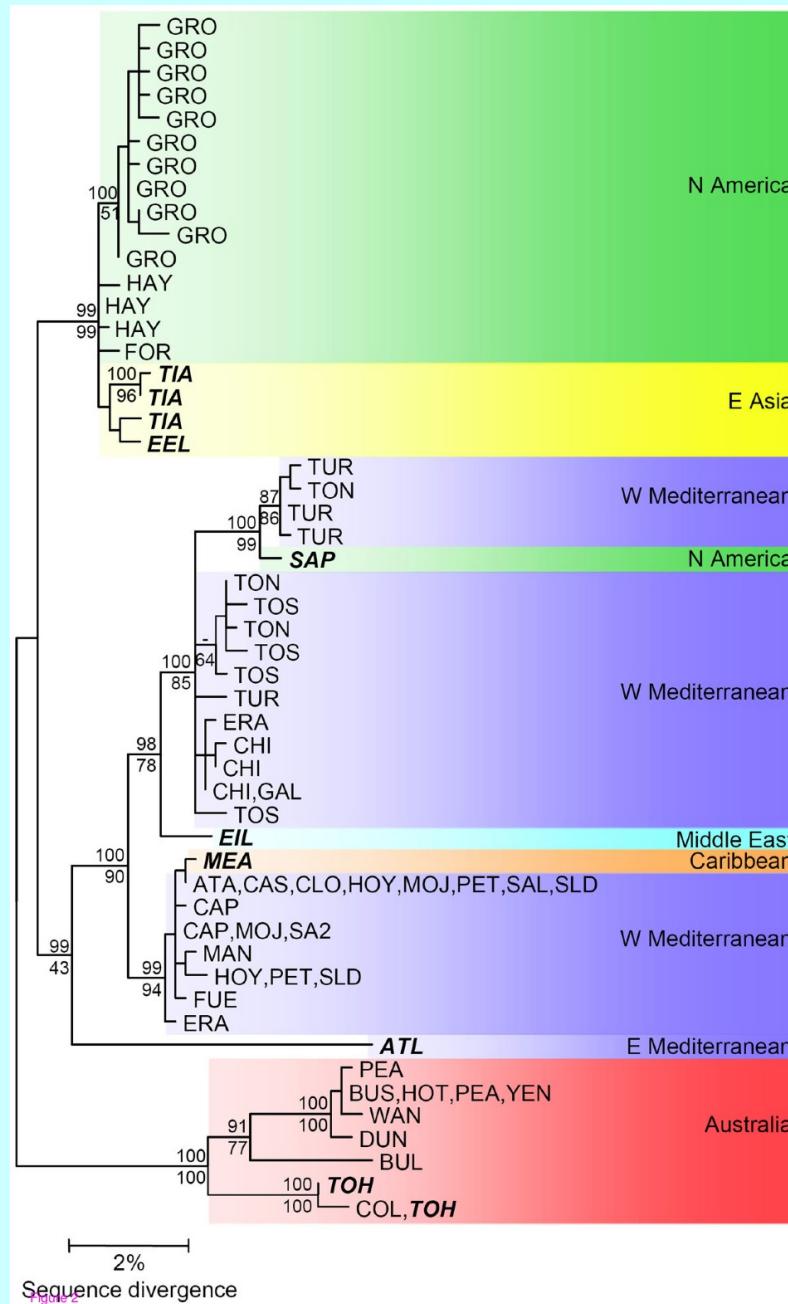
A phylogeographic study

The rotifer *Brachionus plicatilis*

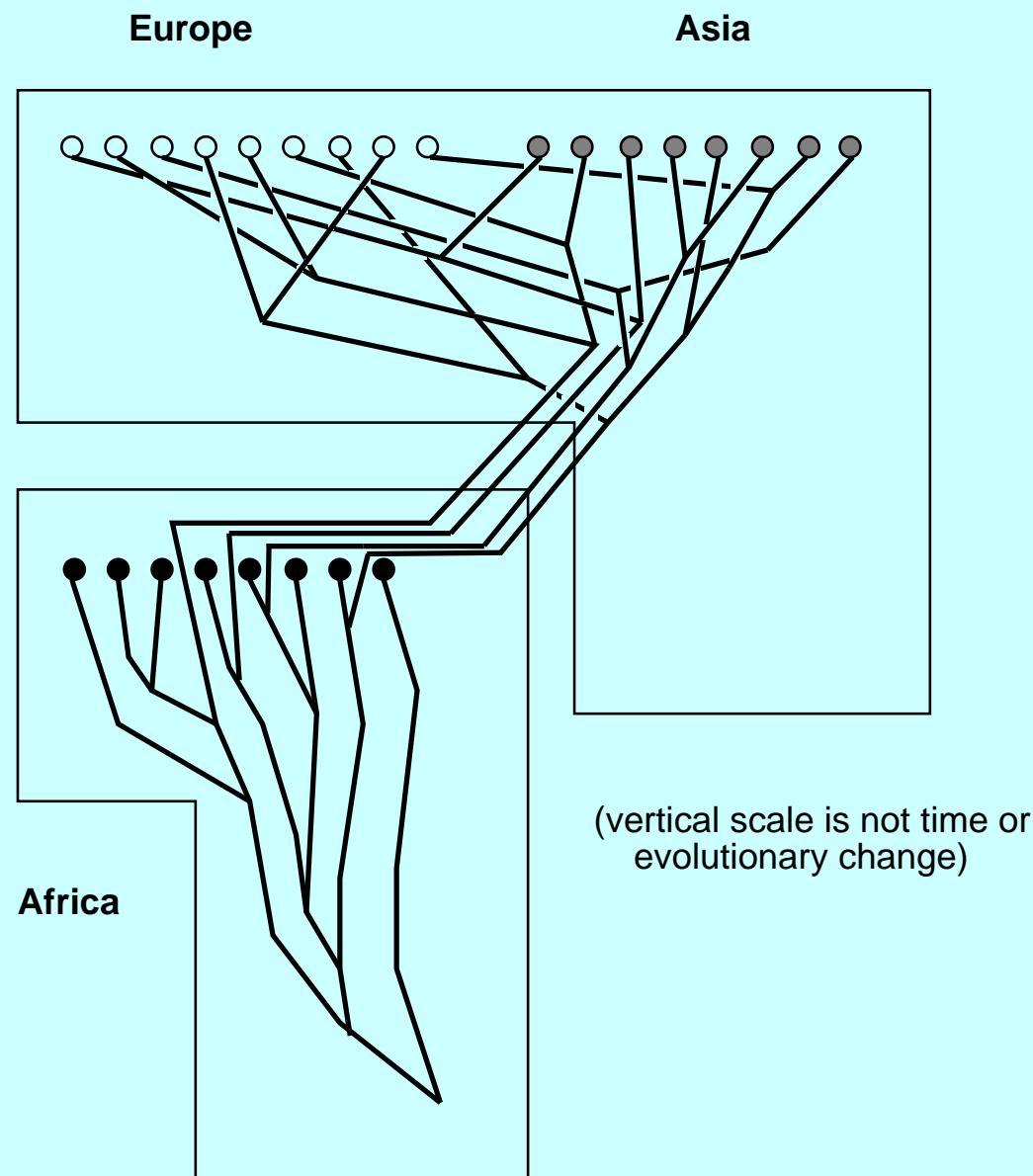


Mills, S., D. H. Lunt, and A. Gomez 2007. Global isolation by distance despite strong regional phylogeography in a small metazoan. *BMC Evolutionary Biology* 7: 225 .

A phylogeographic study



Why mitochondrial eve? The Out-Of-Africa Hypothesis



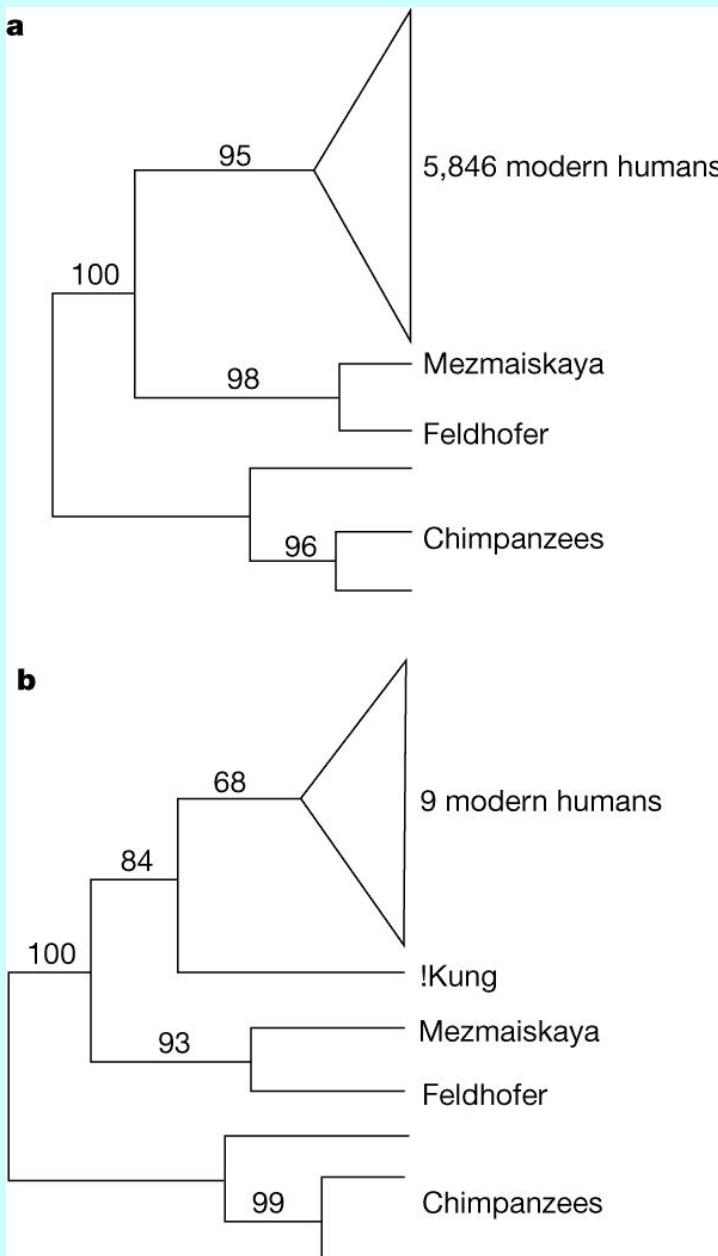
Mitochondria from Neanderthals!



Svante Pääbo

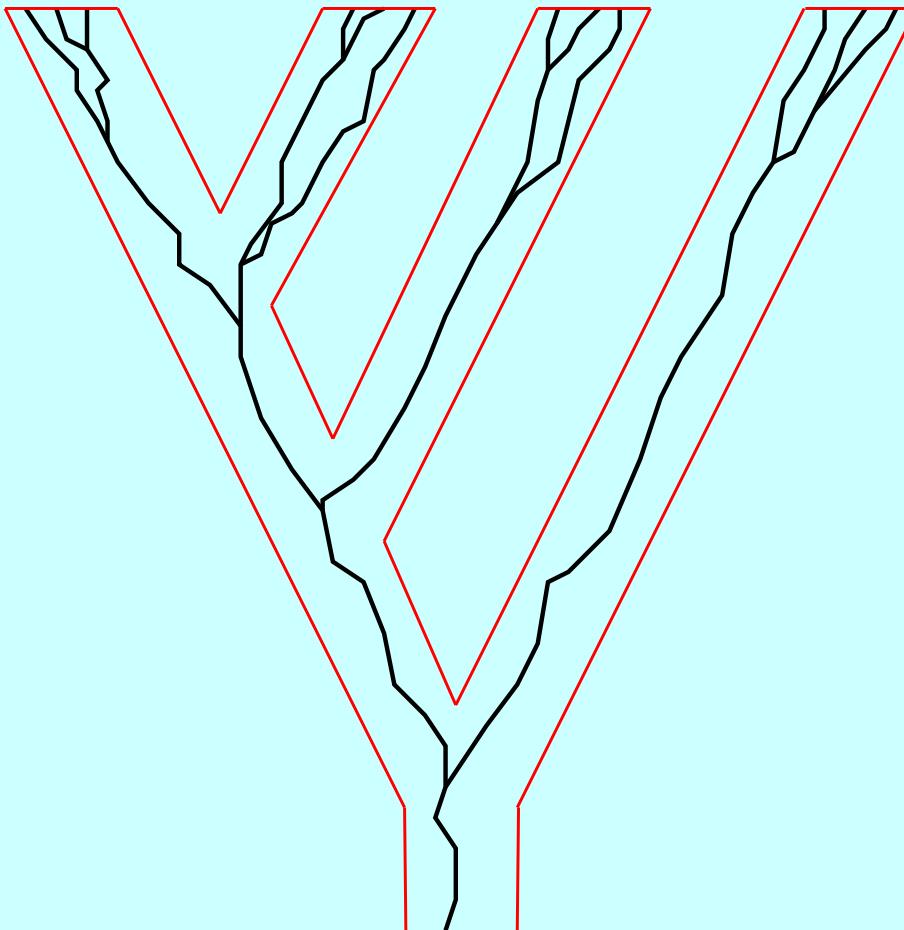
Krings, M., H. Geisert, R. W. Schmitz, H. Krainitzki, and S. Paabo. 1999. DNA sequence of the mitochondrial hypervariable region II from the neandertal type specimen. *Proceedings of the National Academy of Sciences USA* **96**: 5581-5585.
Ovchinnikov, I. V., A. Gotherstrom, G. P. Romanova, V. M. Kharitonov, K. Liden, and W. Goodwin. 2000. Molecular analysis of Neanderthal DNA from the northern Caucasus. *Nature* **404**: 490-493.

The tree with Neanderthals



Coalescents in related species

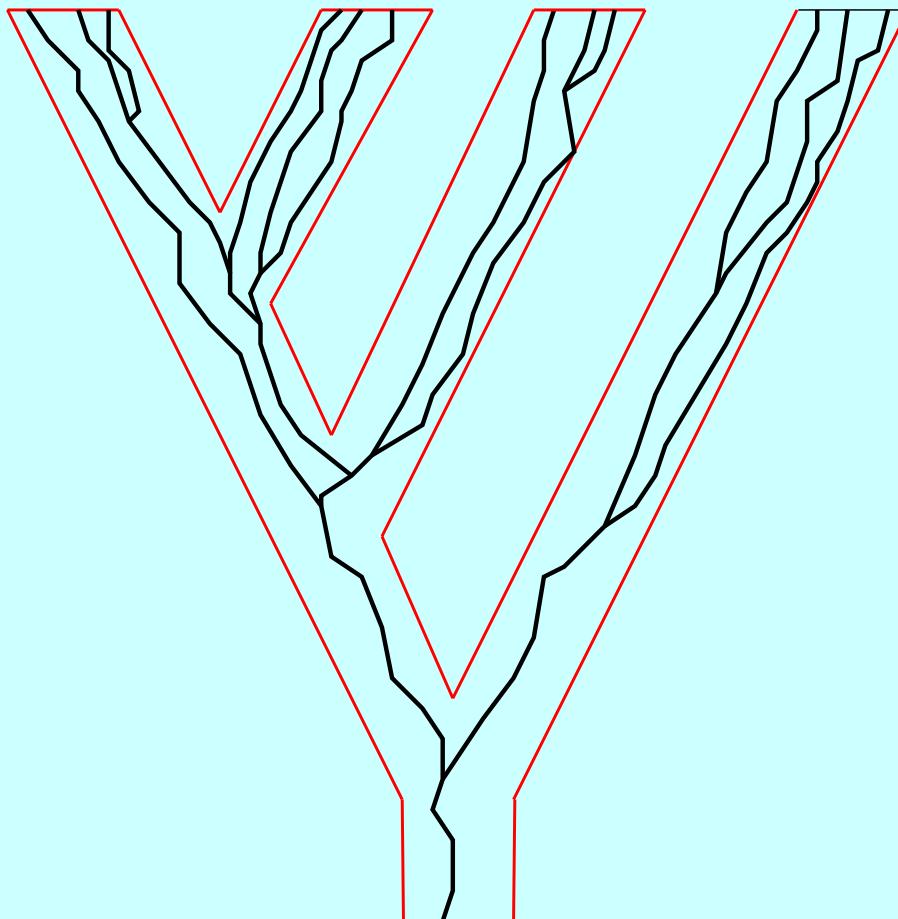
Tree of gene copies, compared with the phylogeny of the species
for the case in which effective population size is small
compared to the number of generations between speciations



The tree of gene copies has topology consistent with the phylogeny

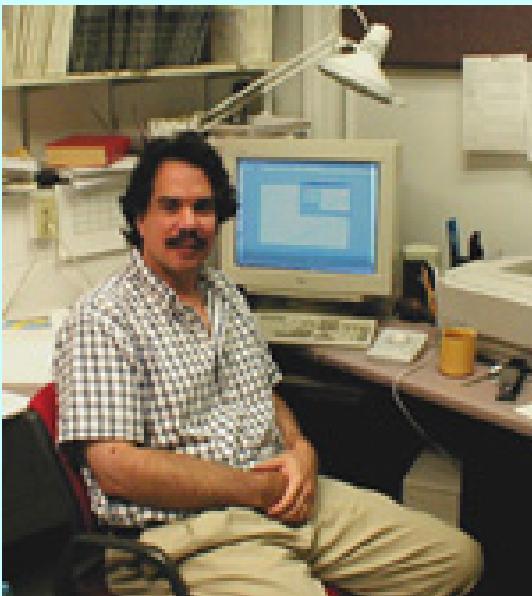
Coalescents in related species

Tree of gene copies, compared with the phylogeny of the species
for the case in which effective population size is large
compared to the number of generations between speciations



The tree of gene copies has topology inconsistent with the phylogeny

Hey and Kliman



Jody Hey

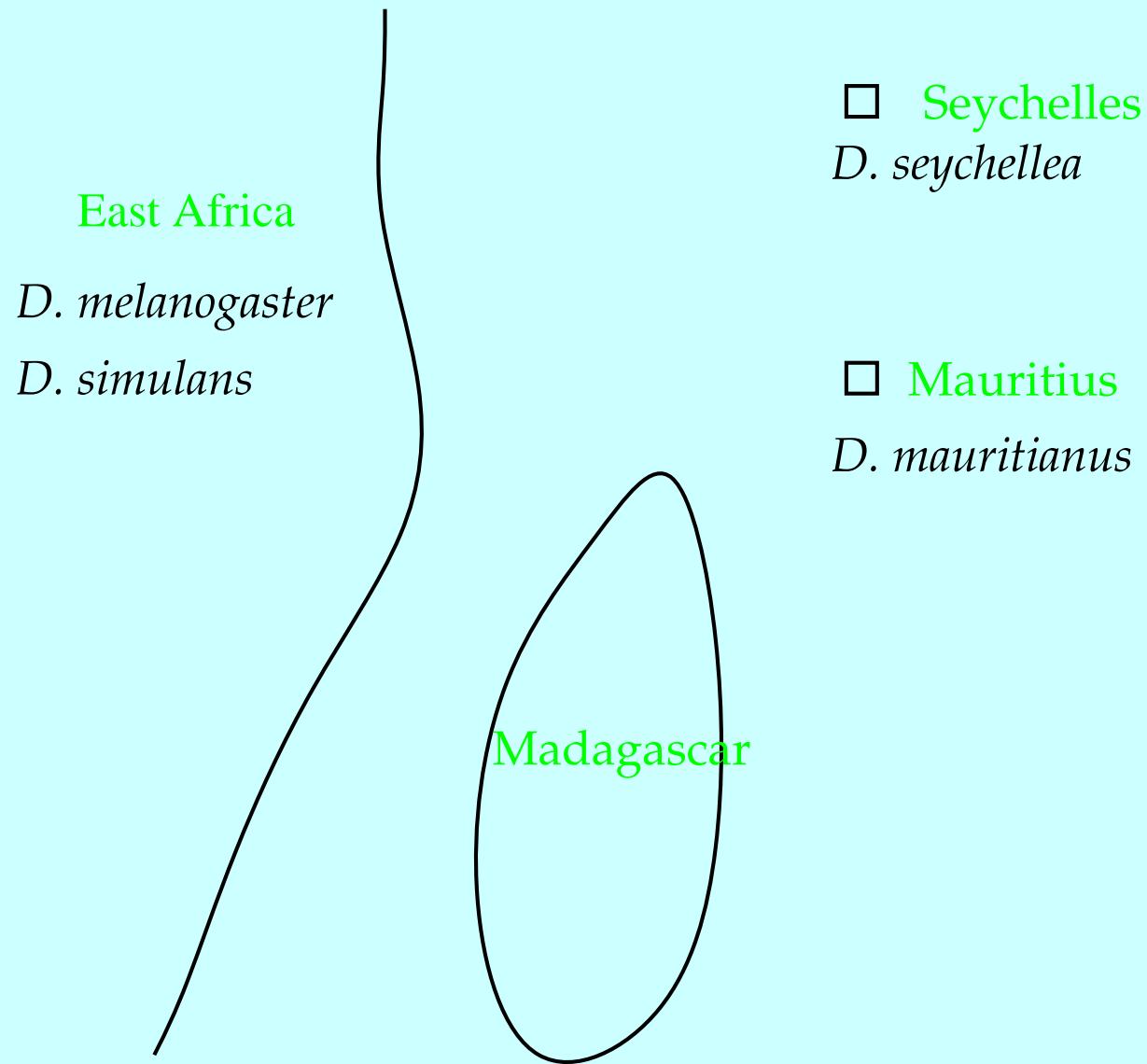
Kliman, R. M., and J. Hey. 1993. DNA sequence variation at the period locus within and among species of the *Drosophila melanogaster* complex. *Genetics* **133**: 375-87



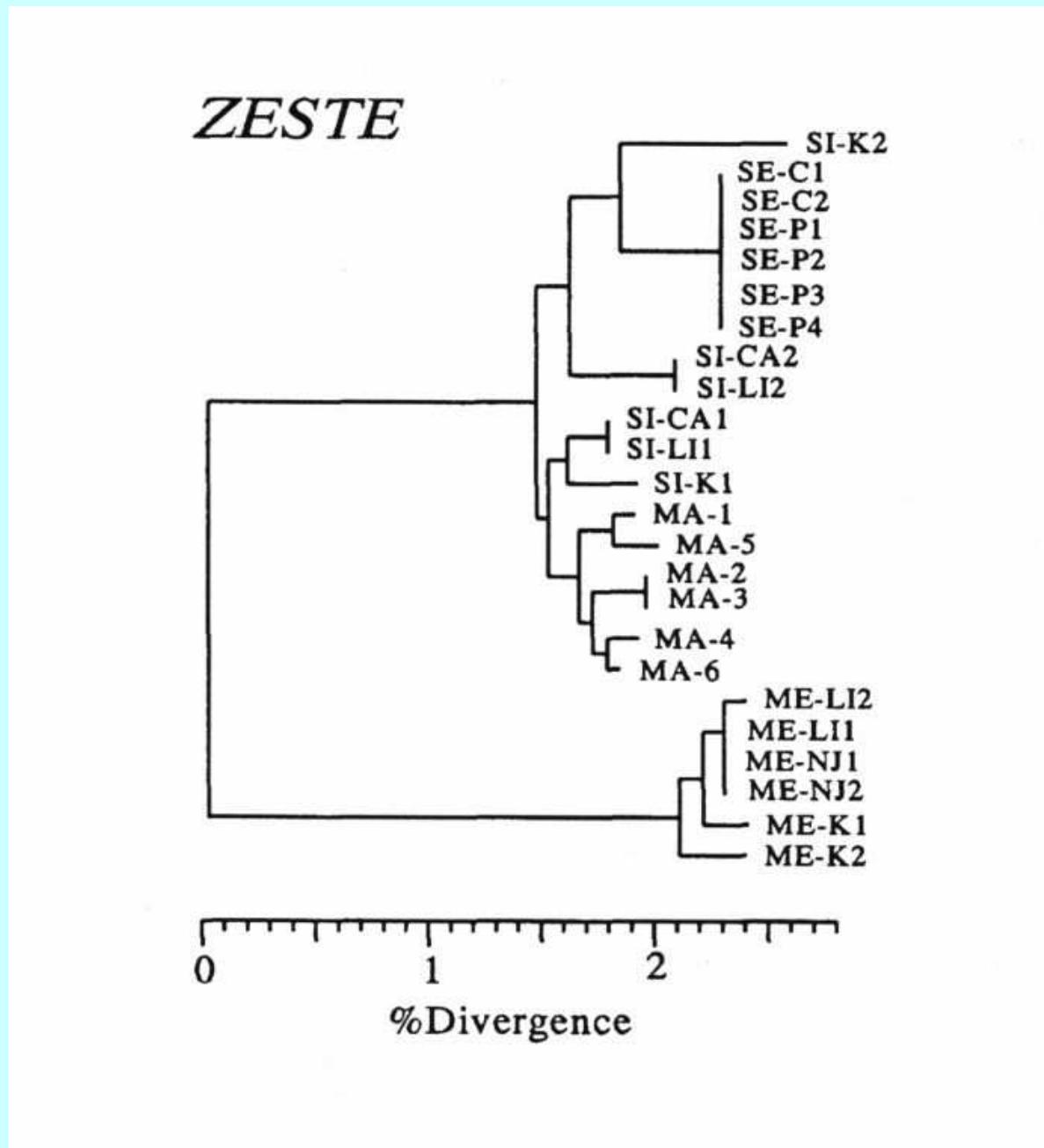
Rich Kliman and friend

Hey, J. and R. M. Kliman. 1993. Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. *Molecular Biology Evolution* **10**: 804-822.

Ranges of East African Drosophila species



Coalescent among *Drosophila* species



Coalescent among *Drosophila* species

