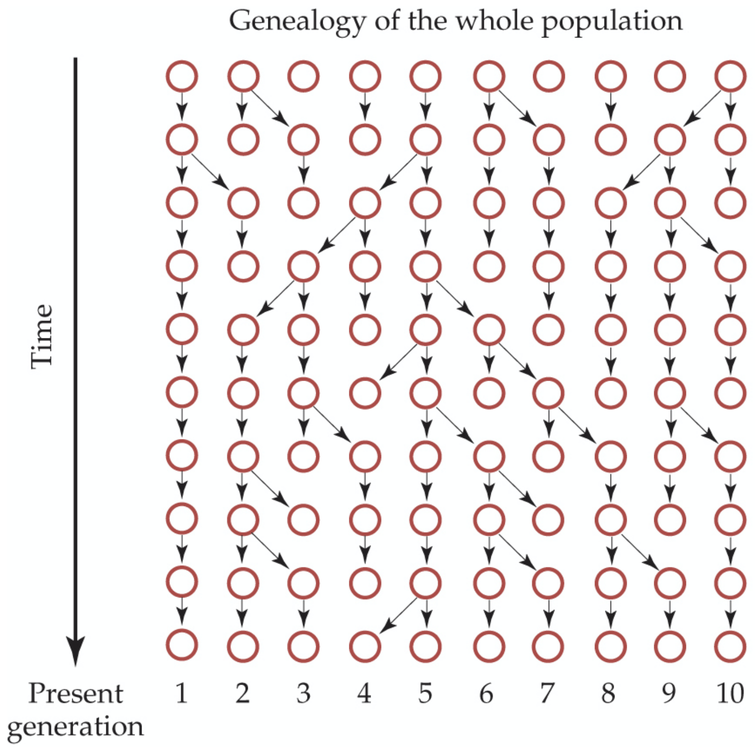
Outline for today

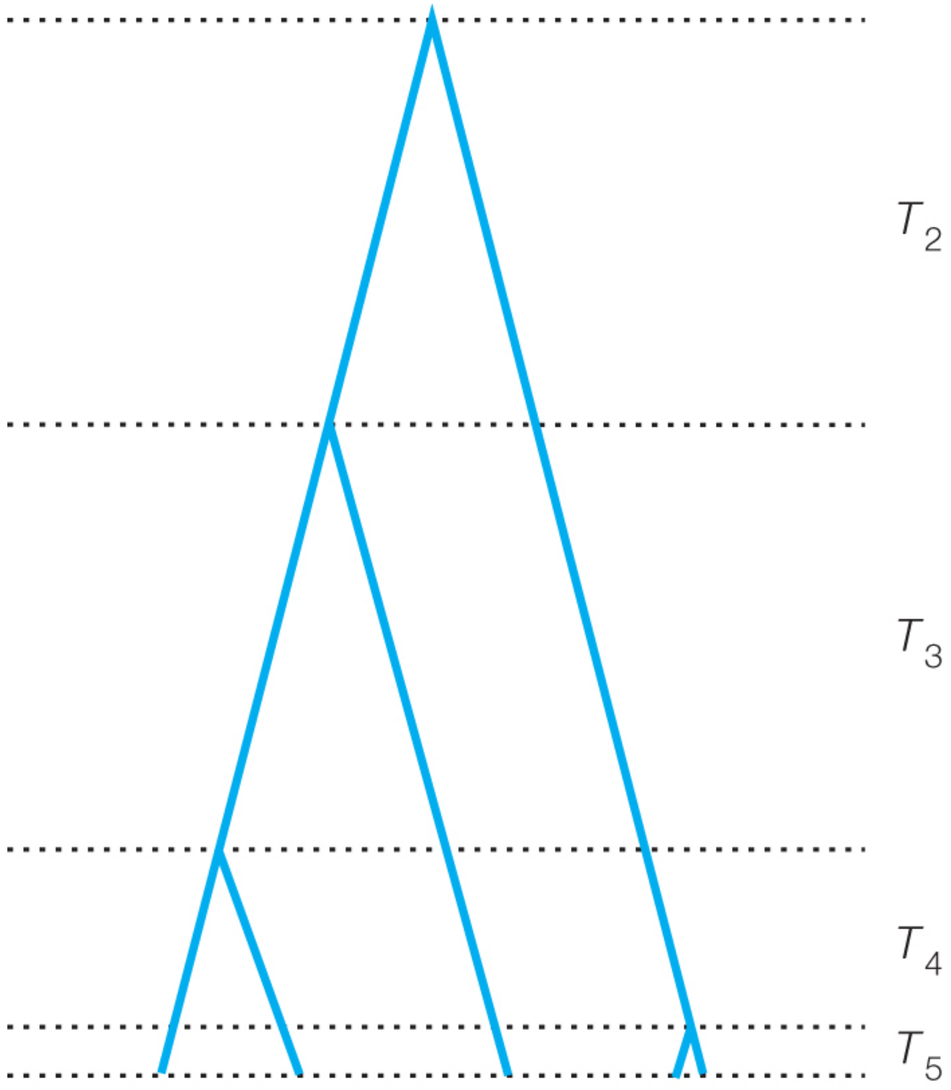
1. Coalescent

2. Multispecies coalescent

Matthew Hahn
mwh@iu.edu
@3rdreviewer

Why do we need the coalescent model?





Algorithm for generating a coalescent tree

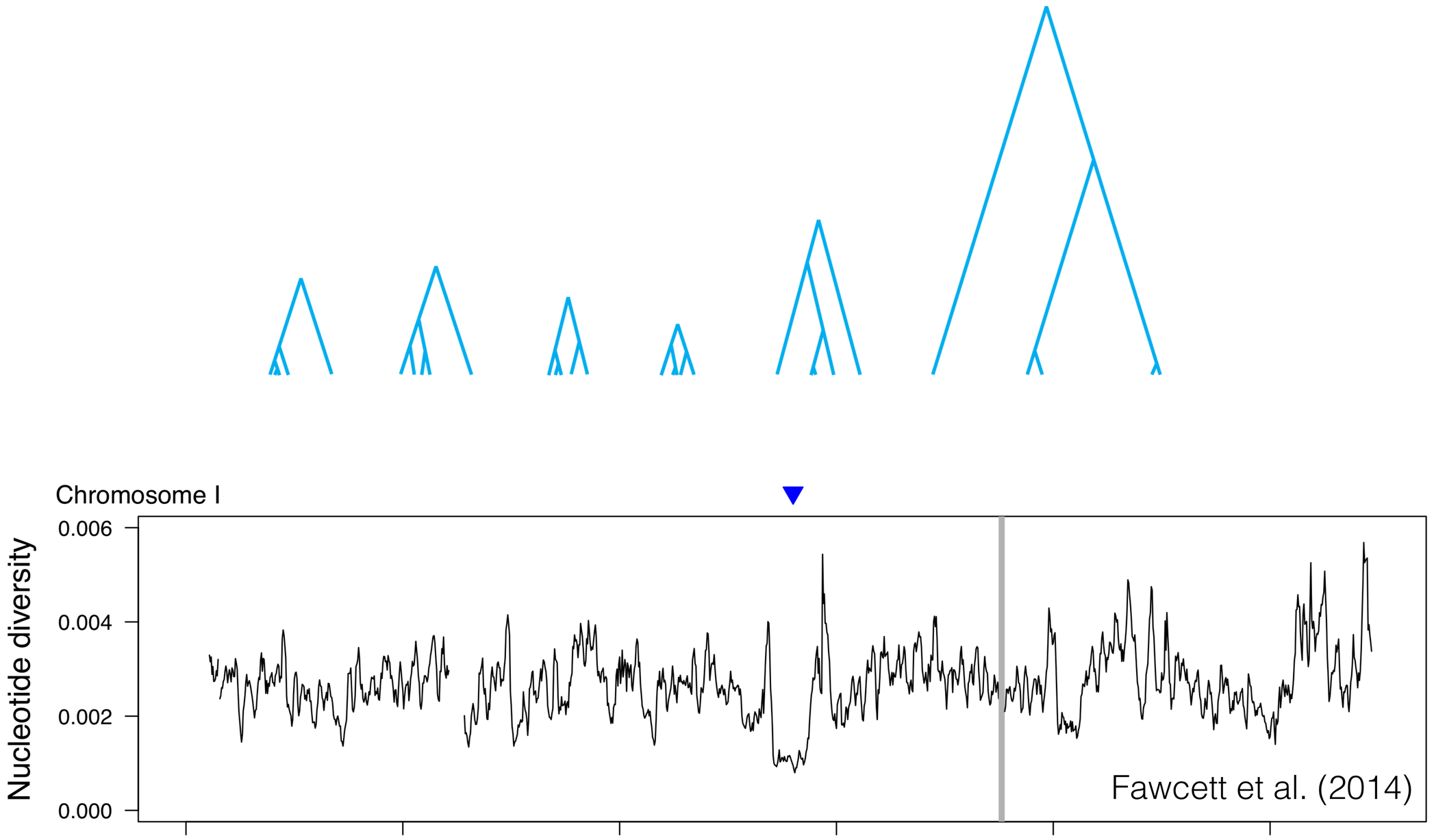
1. Start with $i=n$ chromosomes

2. Choose a time until first coalescence from an exponential distribution with parameter $x= i(i-1)/2$

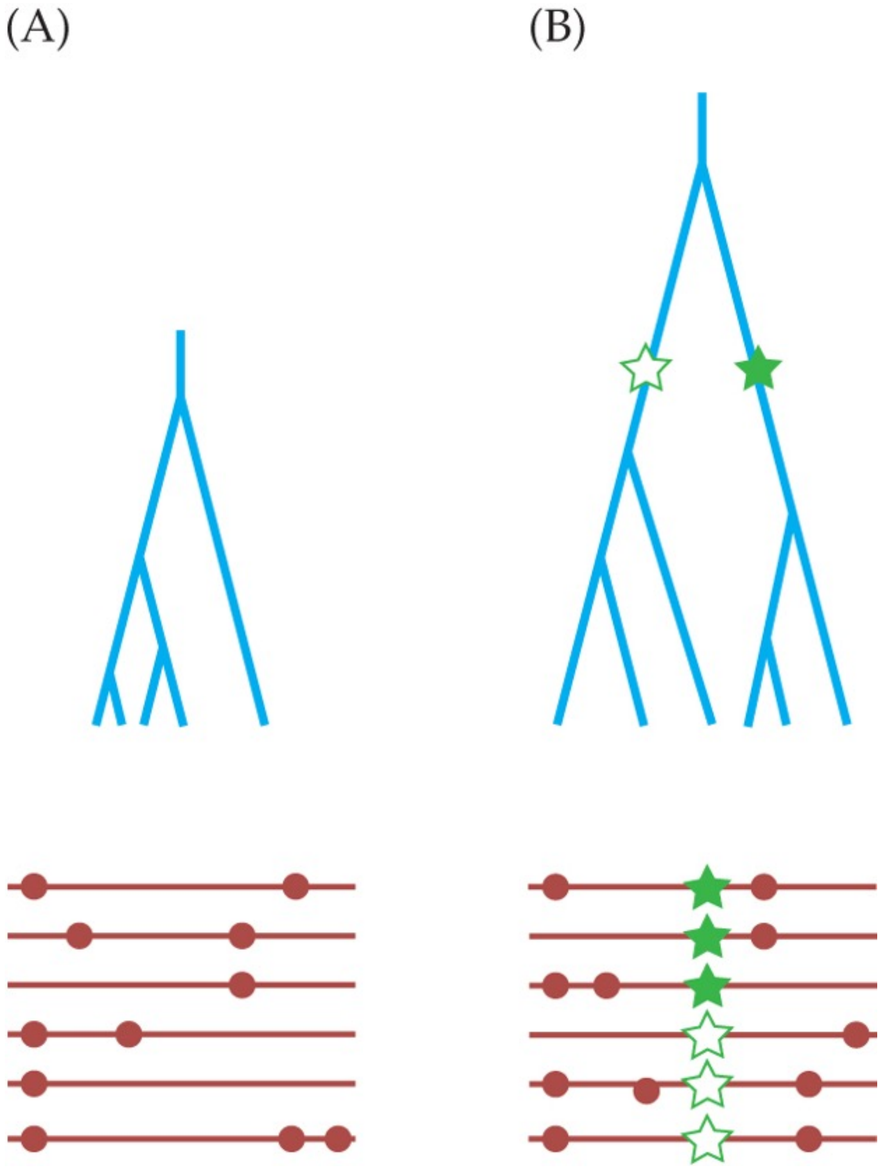
3. Choose two chromosomes at random to coalesce

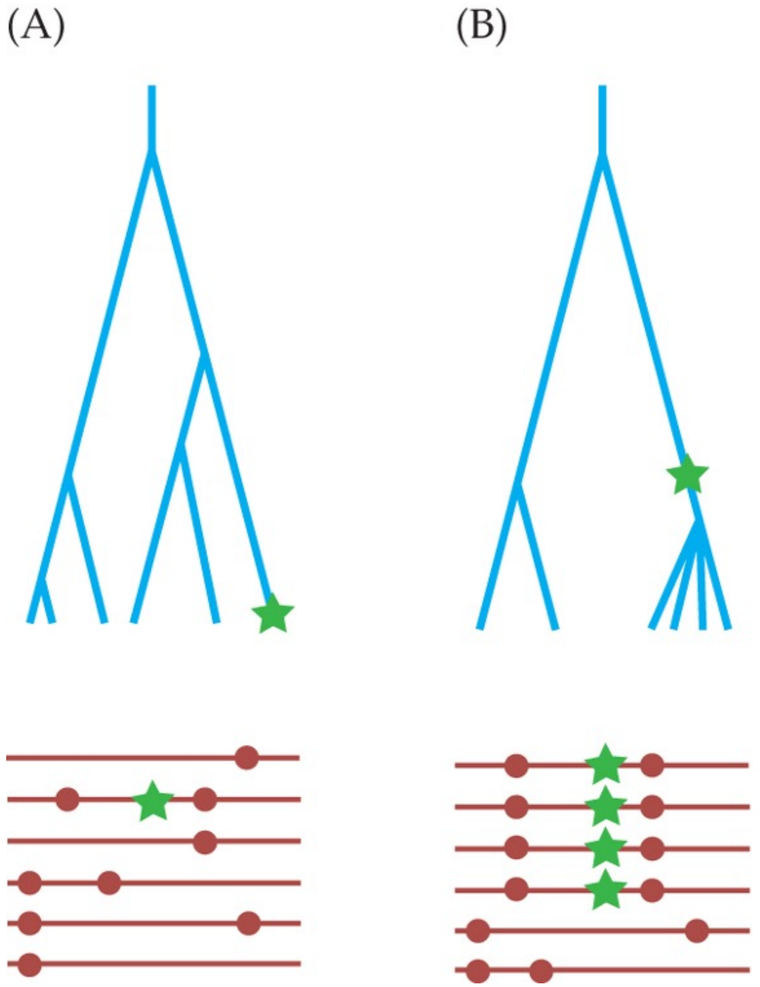
4. Merge the two lineages chosen, and have $i \rightarrow i-1$

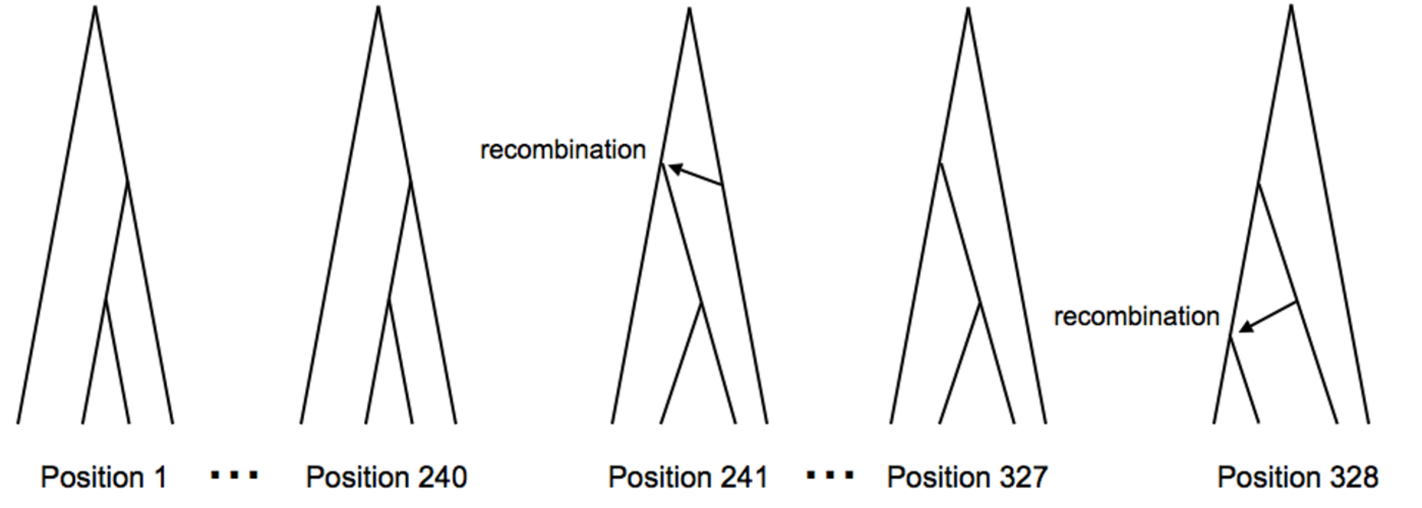
5. If $i>1$ go to 2; if not stop

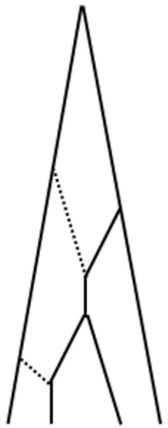

Chromosome I
Nucleotide diversity
0.006
0.004
0.002
0.000
Fawcett et al. (2014)

https://bedford.io/projects/coaltrace/


(A)
(B)

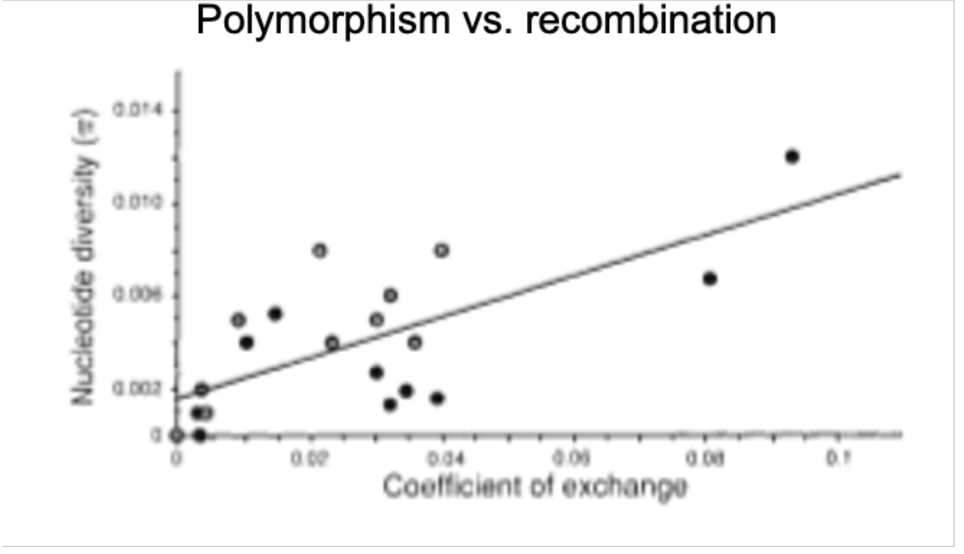

(A)
(B)



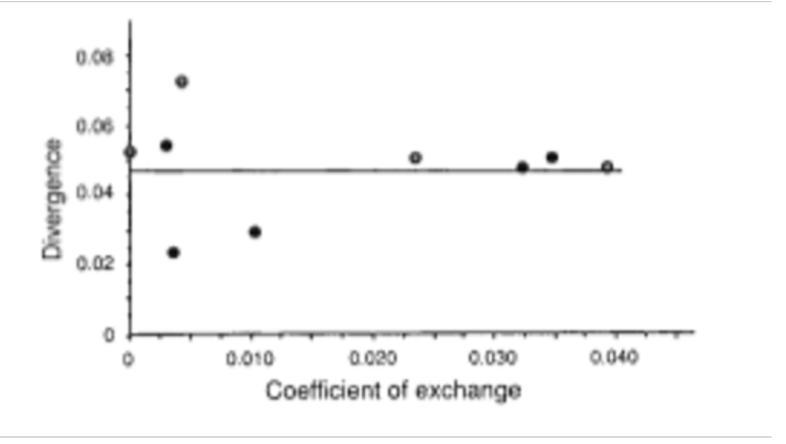


Ancestral recombination graph

Polymorphism vs. recombination



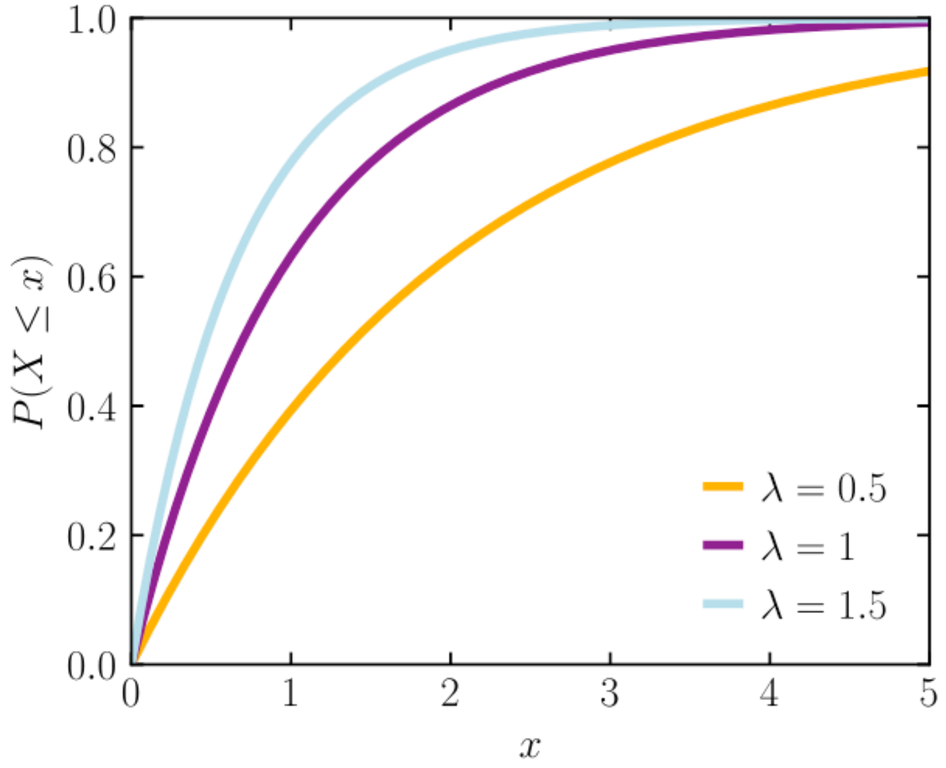
Divergence vs. recombination



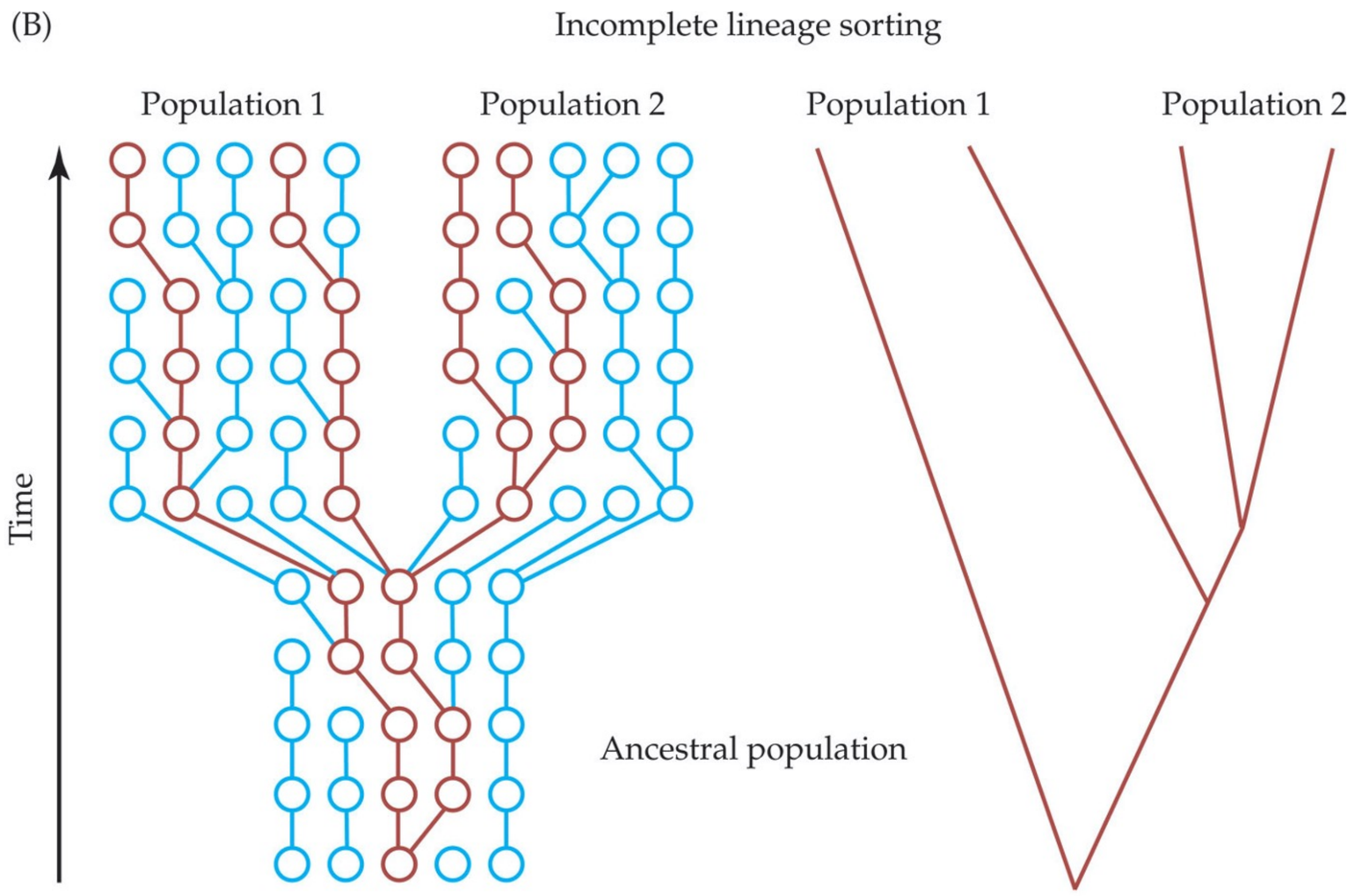
Begun and Aquadro (1992)

Cumulative distribution function of exponential: $1-e^{-\lambda x}$

(For two lineages, lambda=1)

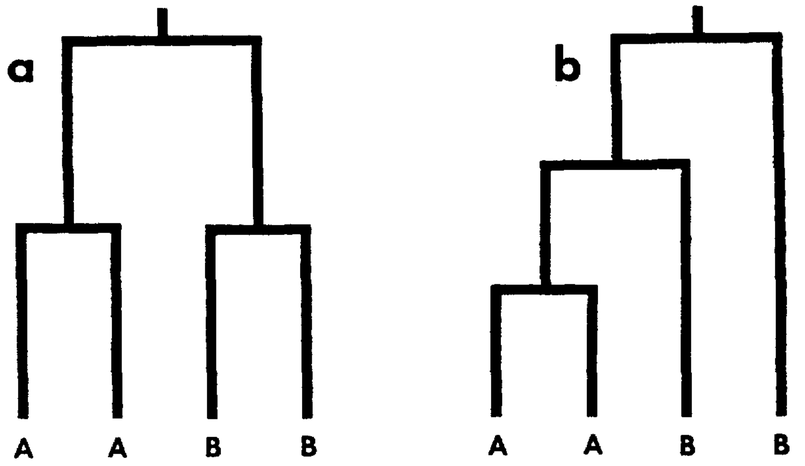


Wikimedia

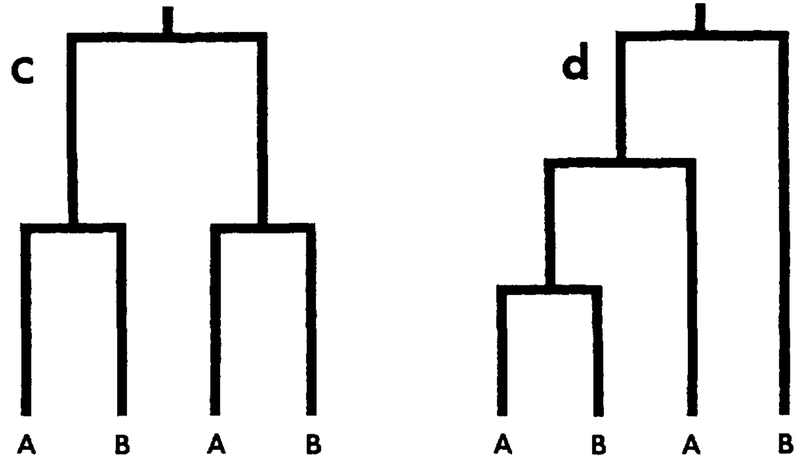


***INTRODUCTION TO POPULATION GENETICS,* Figure 5.5 (Part 2)**

Nielsen and Slatkin (2013)

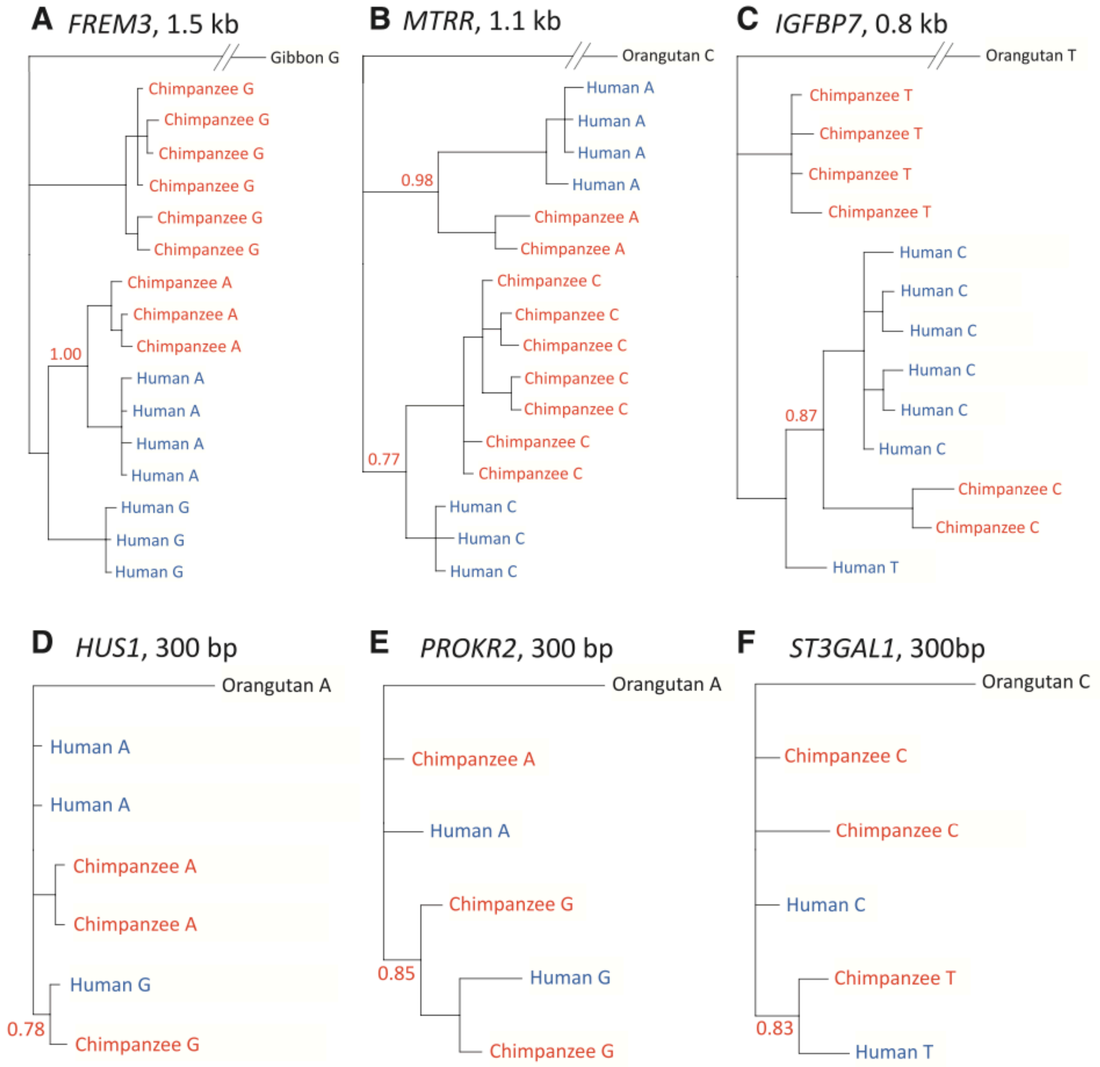


$$P = \left(1 - \frac{2}{3}e^{-\frac{t}{2N}}\right)^2 \qquad P = \frac{4}{3}e^{-\frac{t}{2N}}\left(1 - \frac{5}{6}e^{-\frac{t}{2N}}\right)$$

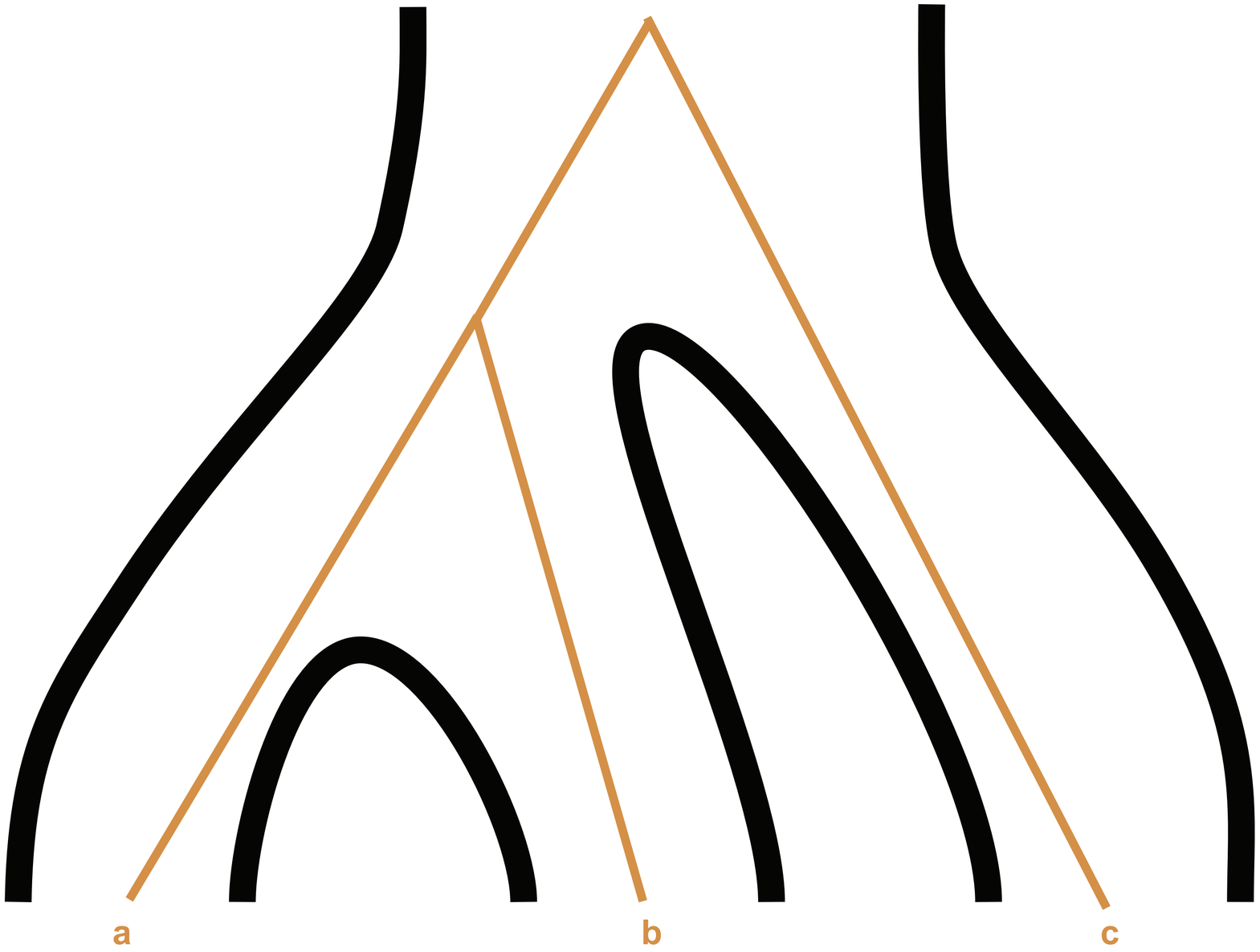


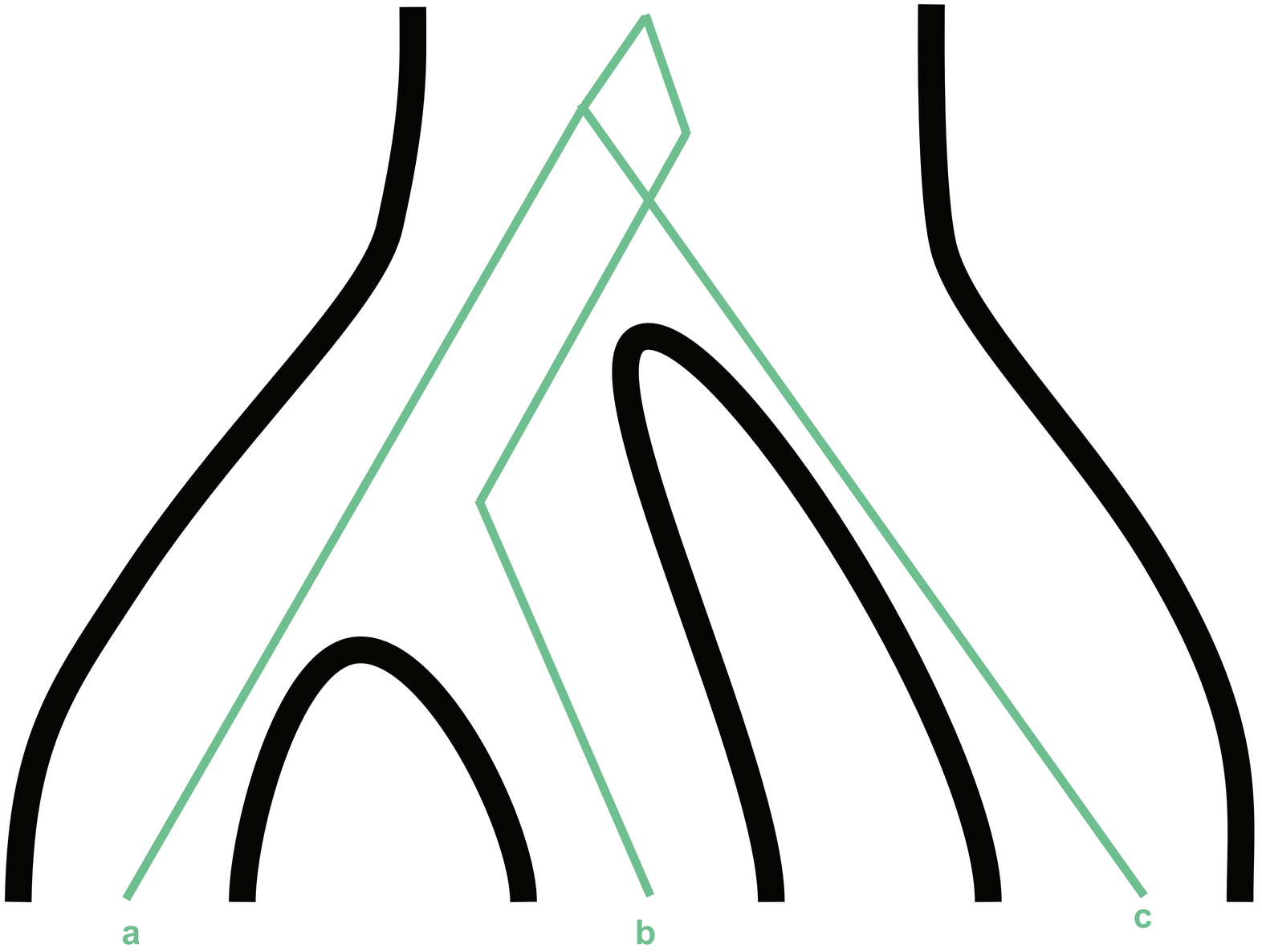
$$P = \frac{2}{9}e^{-\frac{t}{N}} \qquad P = \frac{4}{9}e^{-\frac{t}{N}}$$

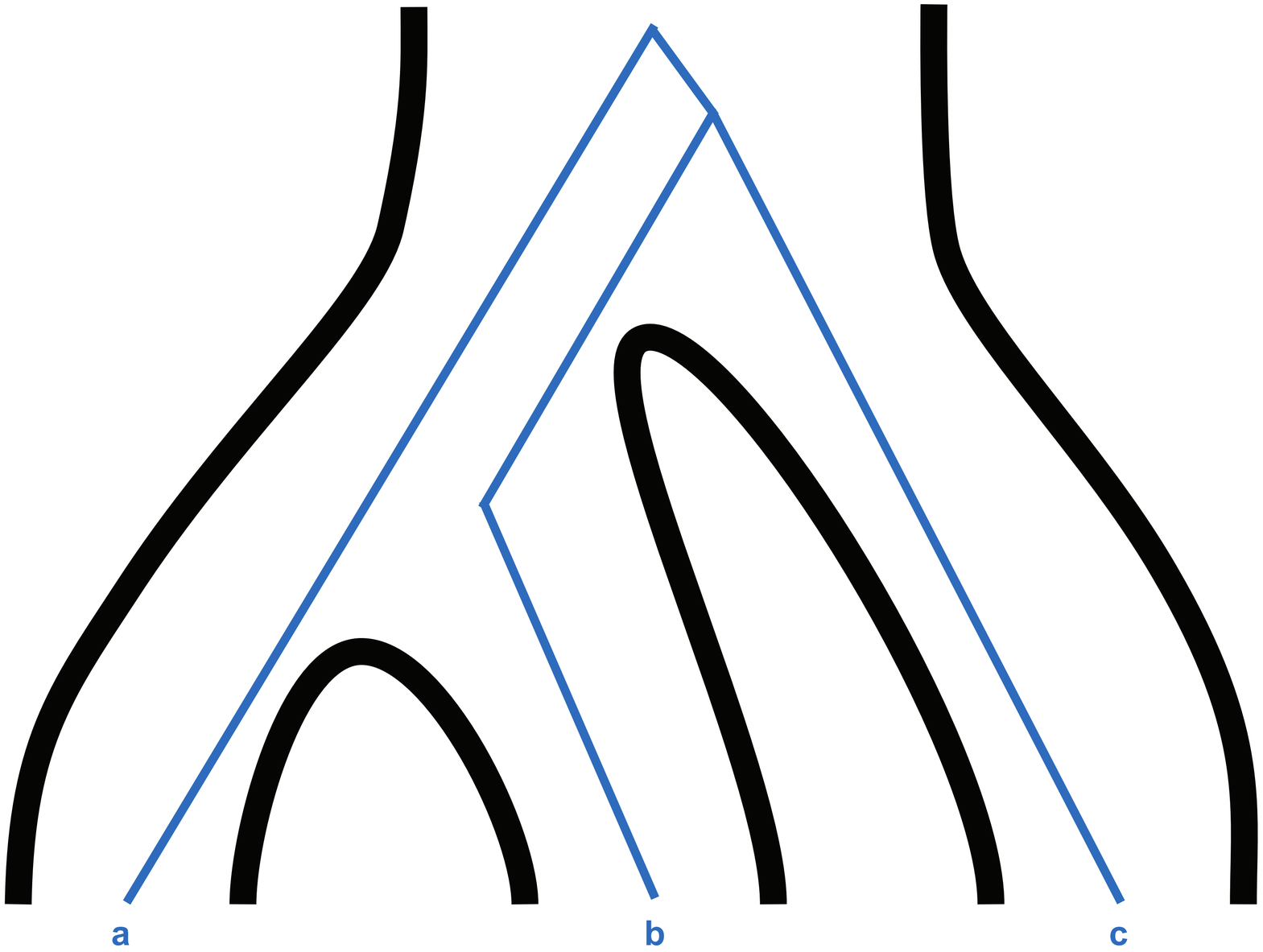
Tajima (1983)



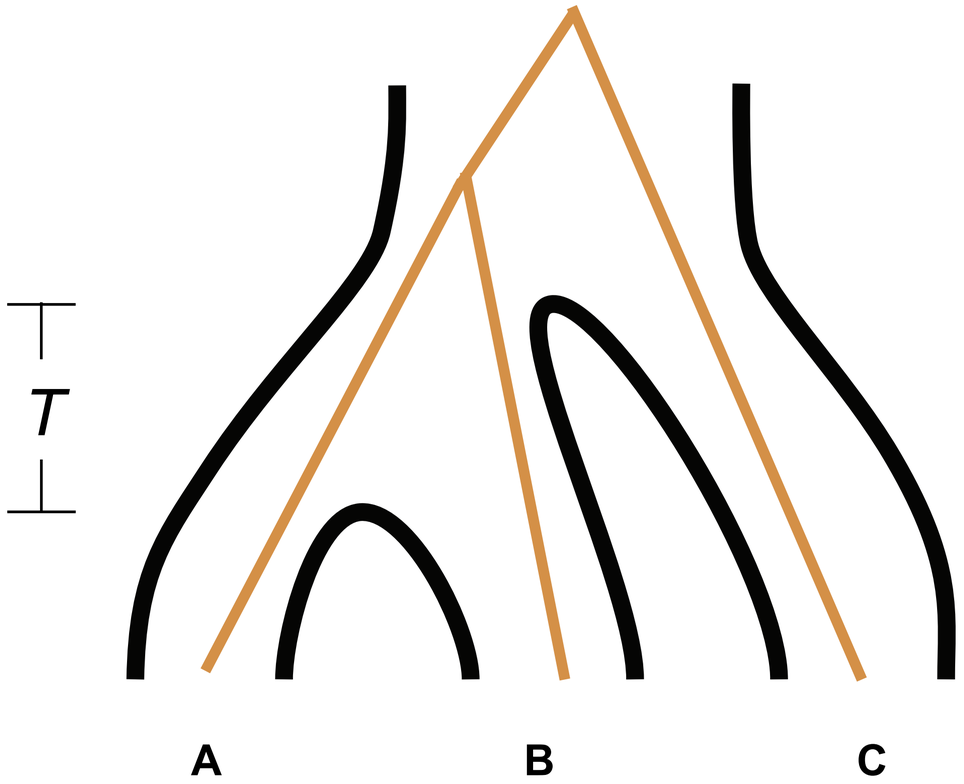
Leffler et al. (2013)


a
b
c


a
b
c


a
b
c

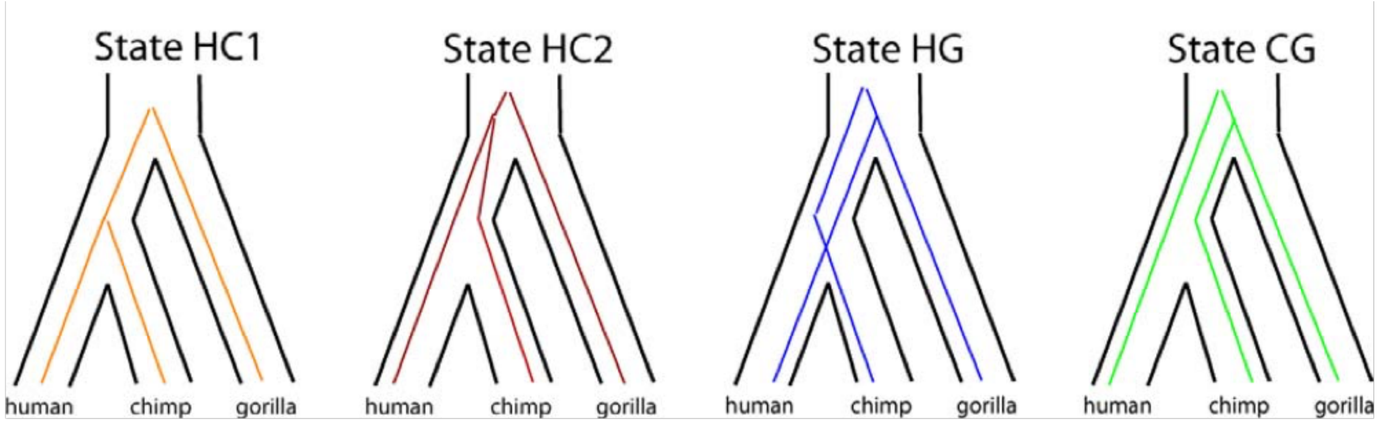
The multispecies coalescent model

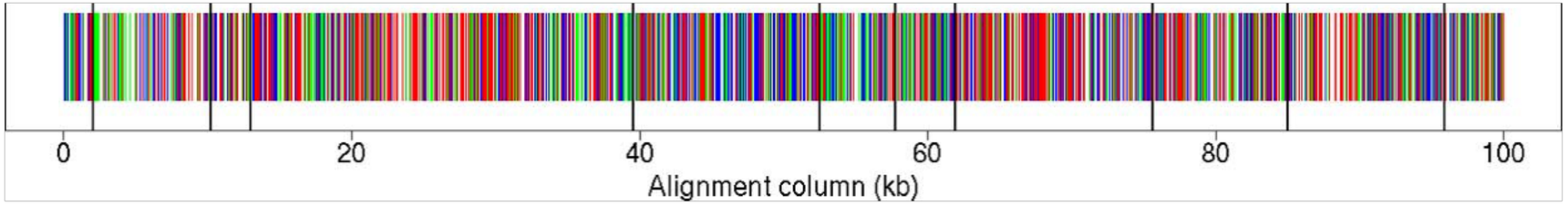


$1-(2/3)e^{-T}$

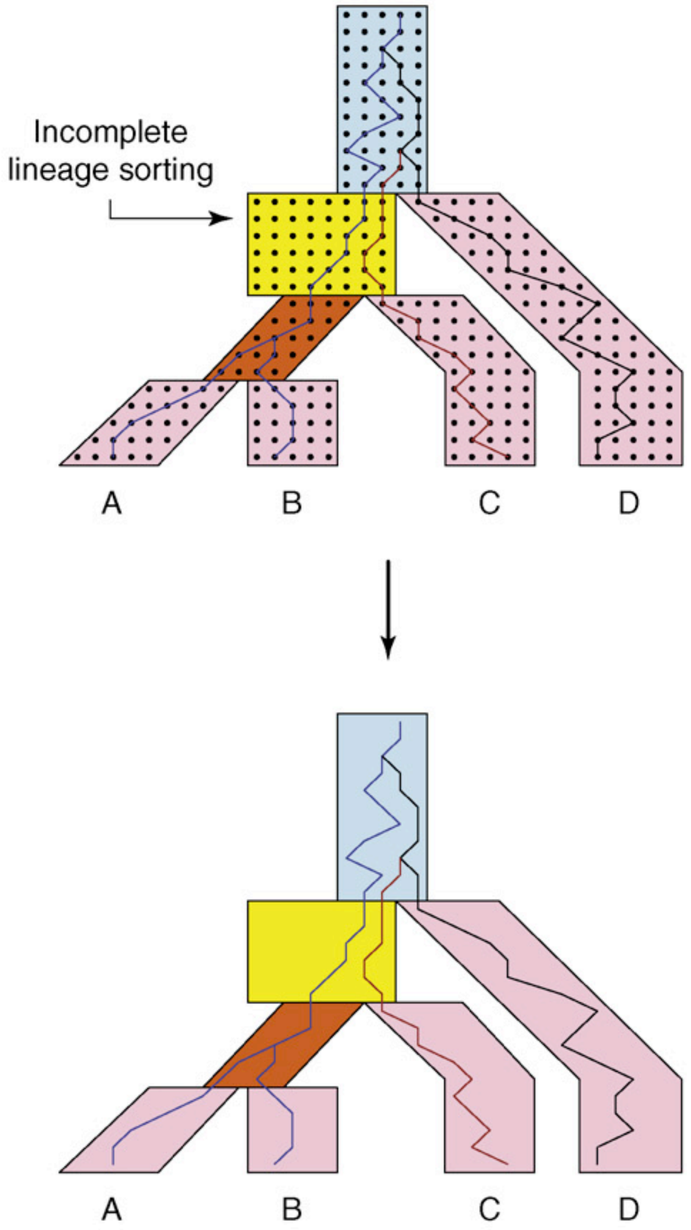
$(1/3)e^{-T}$

$(1/3)e^{-T}$

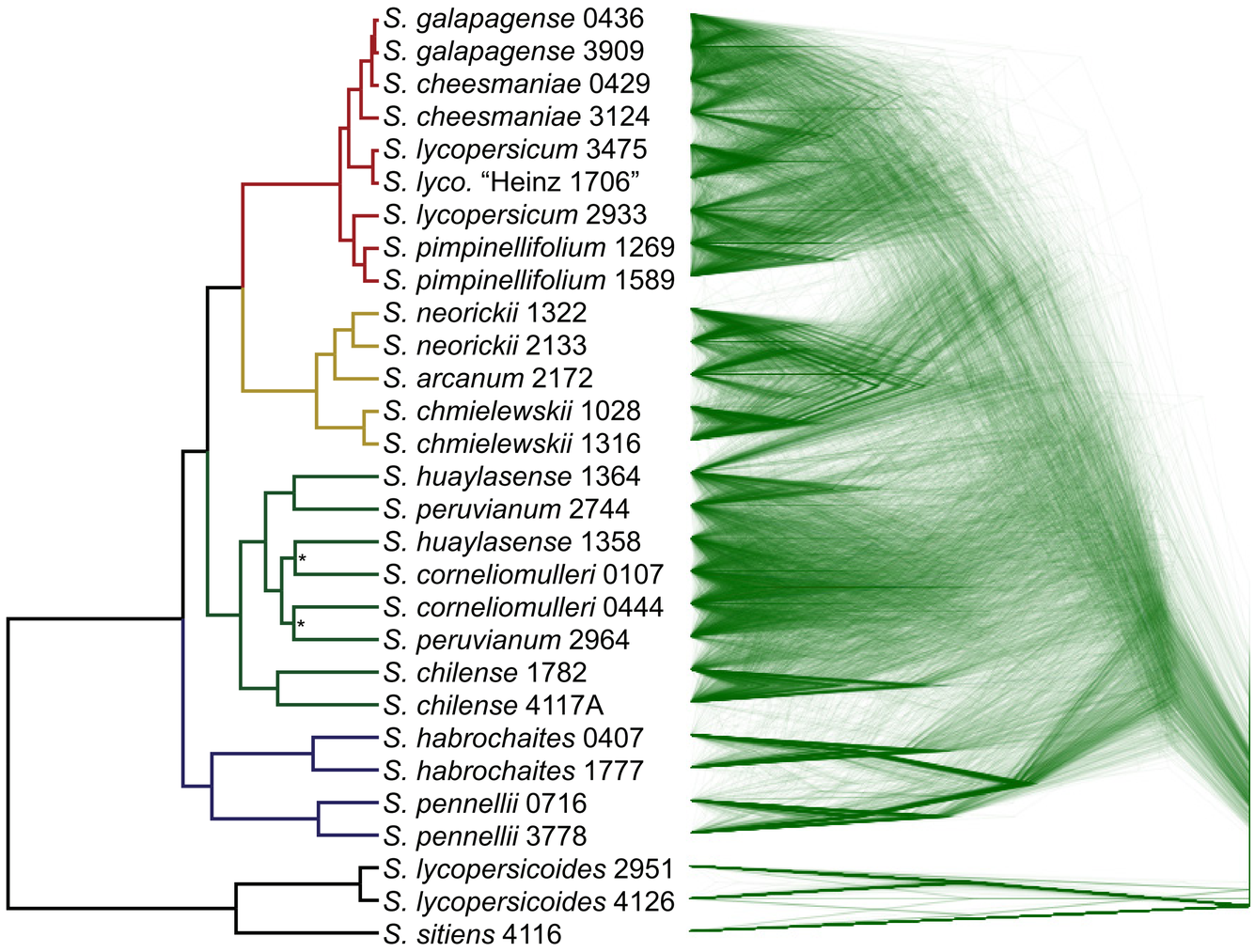




Hobolth et al. (2007)

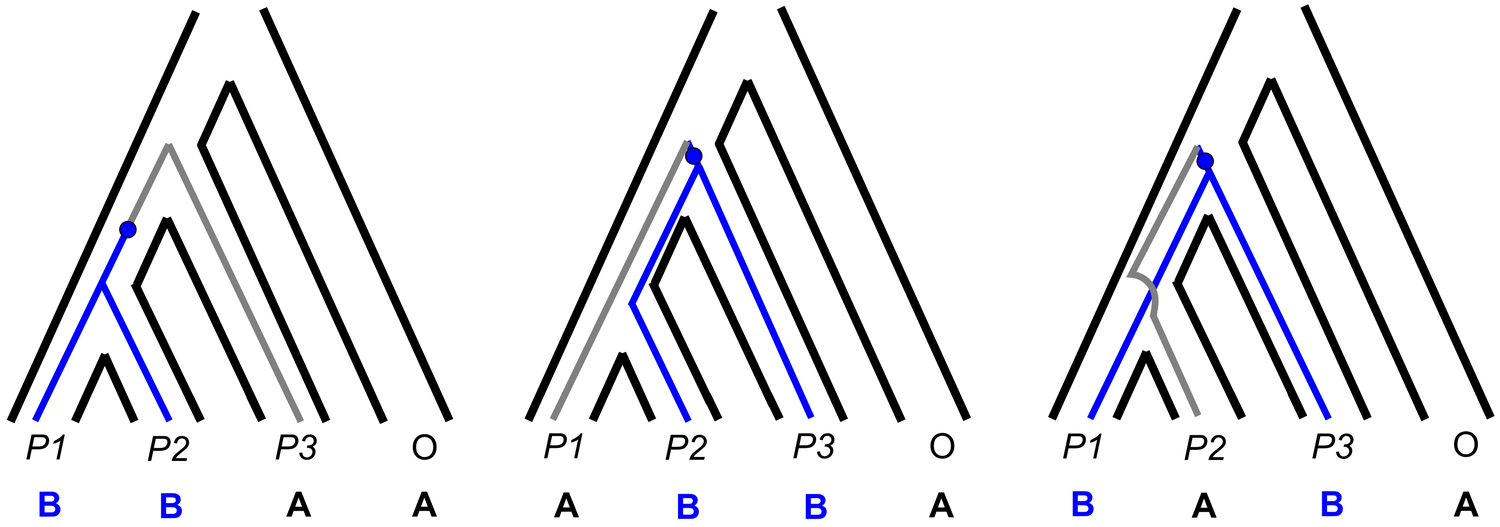


Degnan and Rosenberg (2009)



Pease et al. (2016)

ILS only:



ILS + introgression: