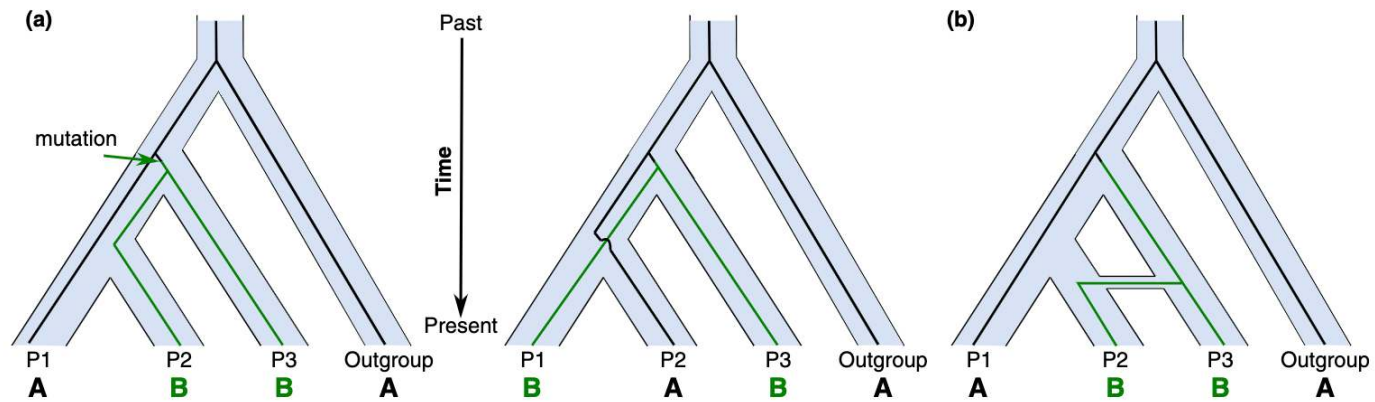
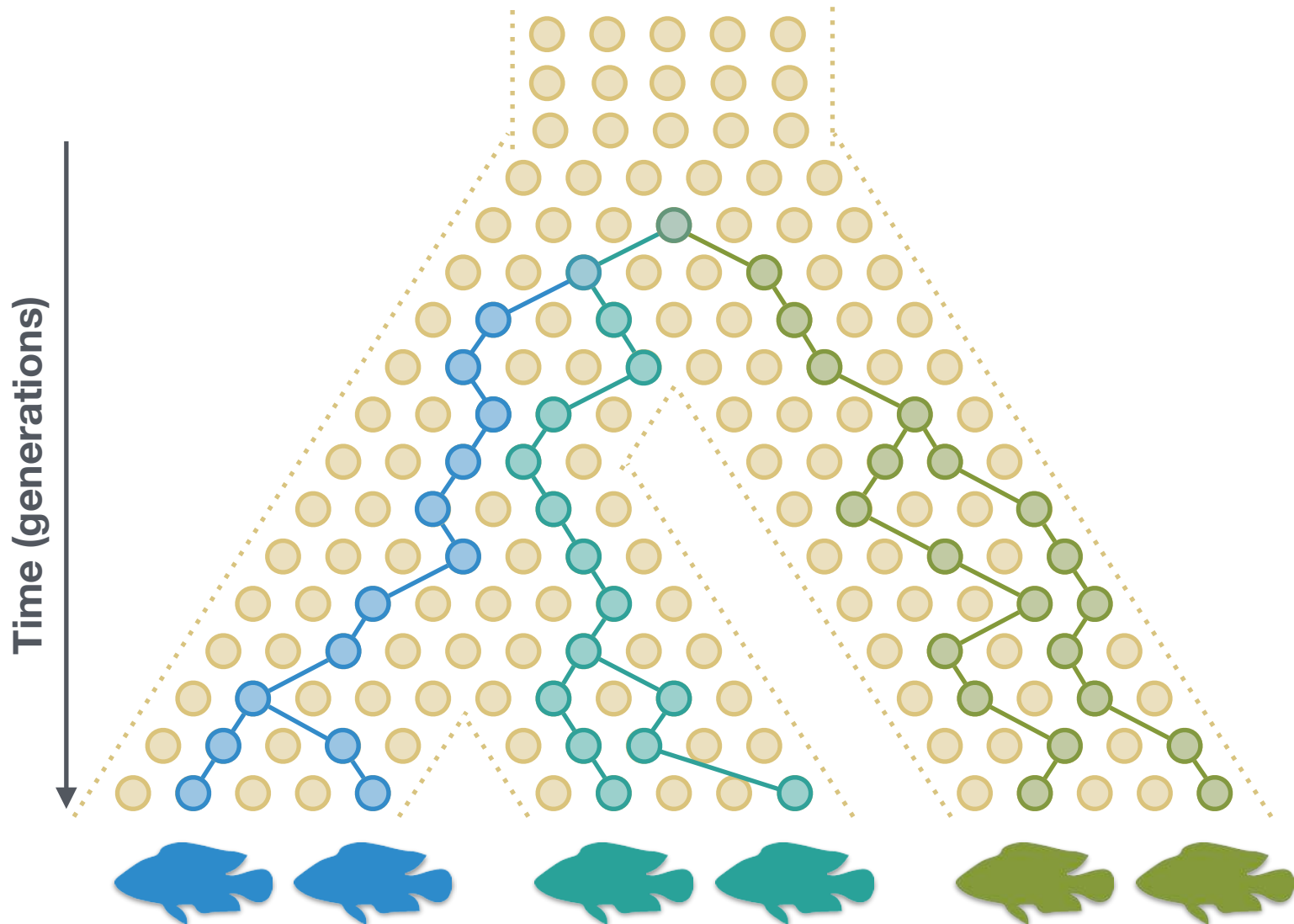


Introgression Activity

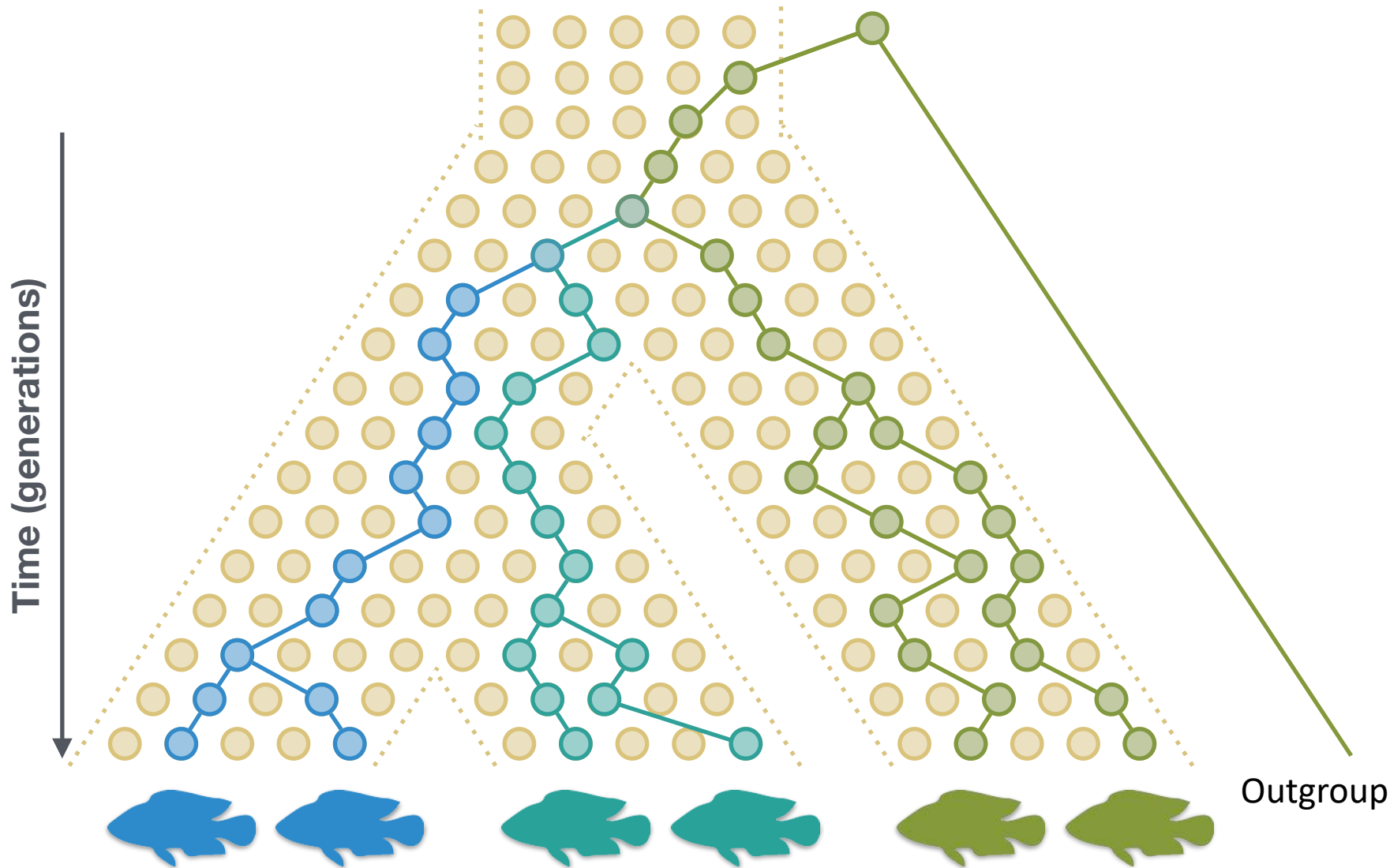
Introgression Activity



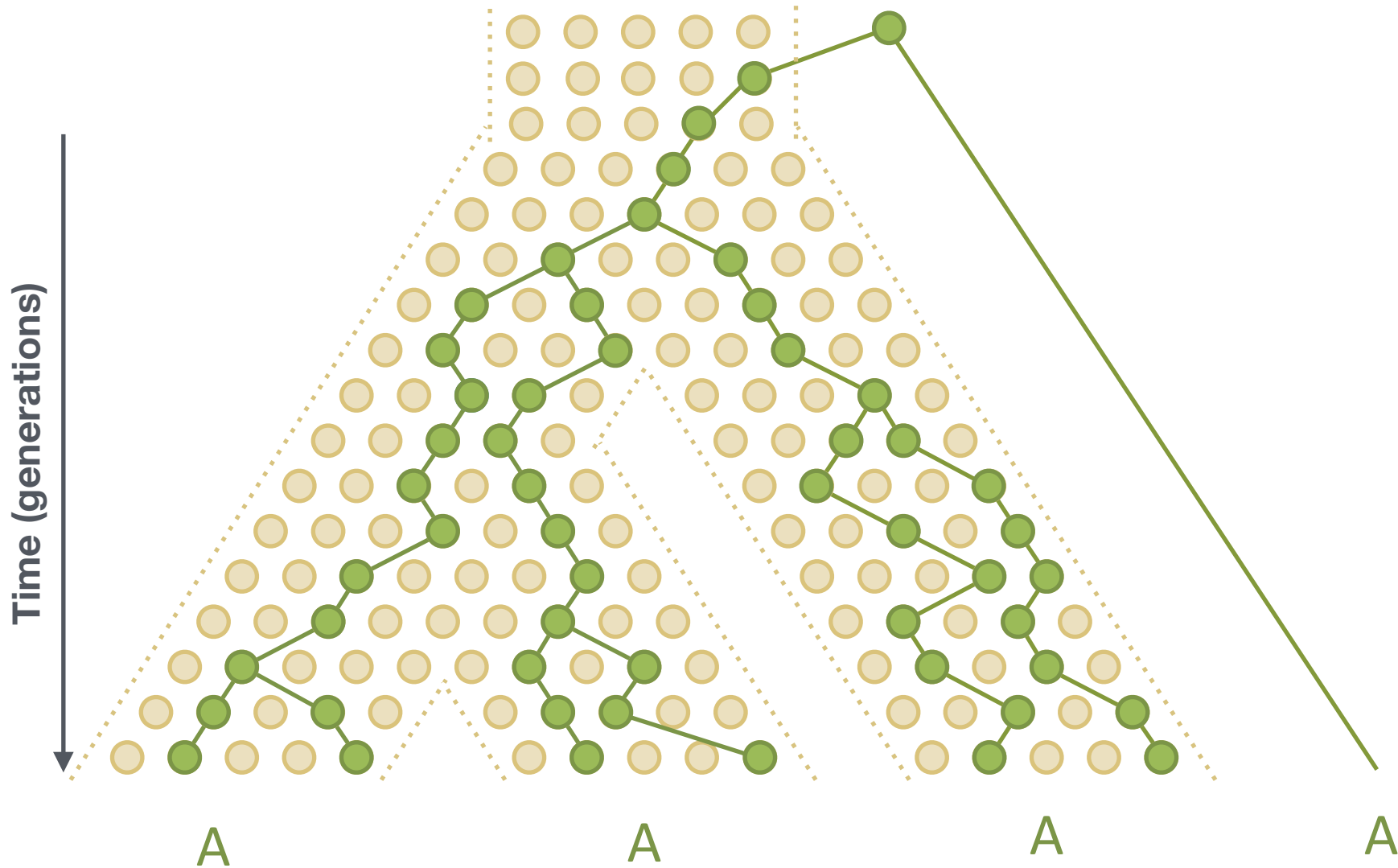
The multi-species coalescent model



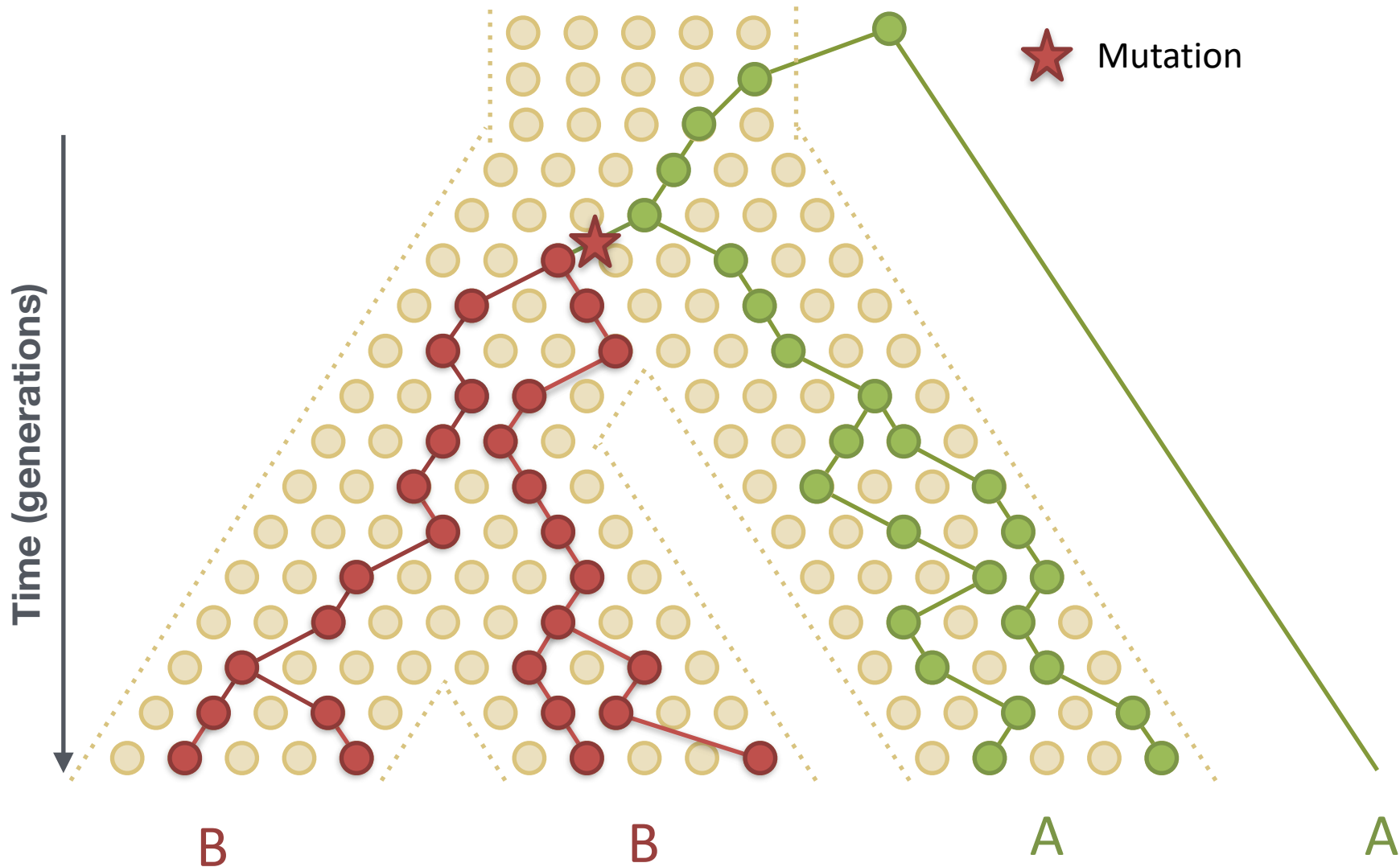
The multi-species coalescent model



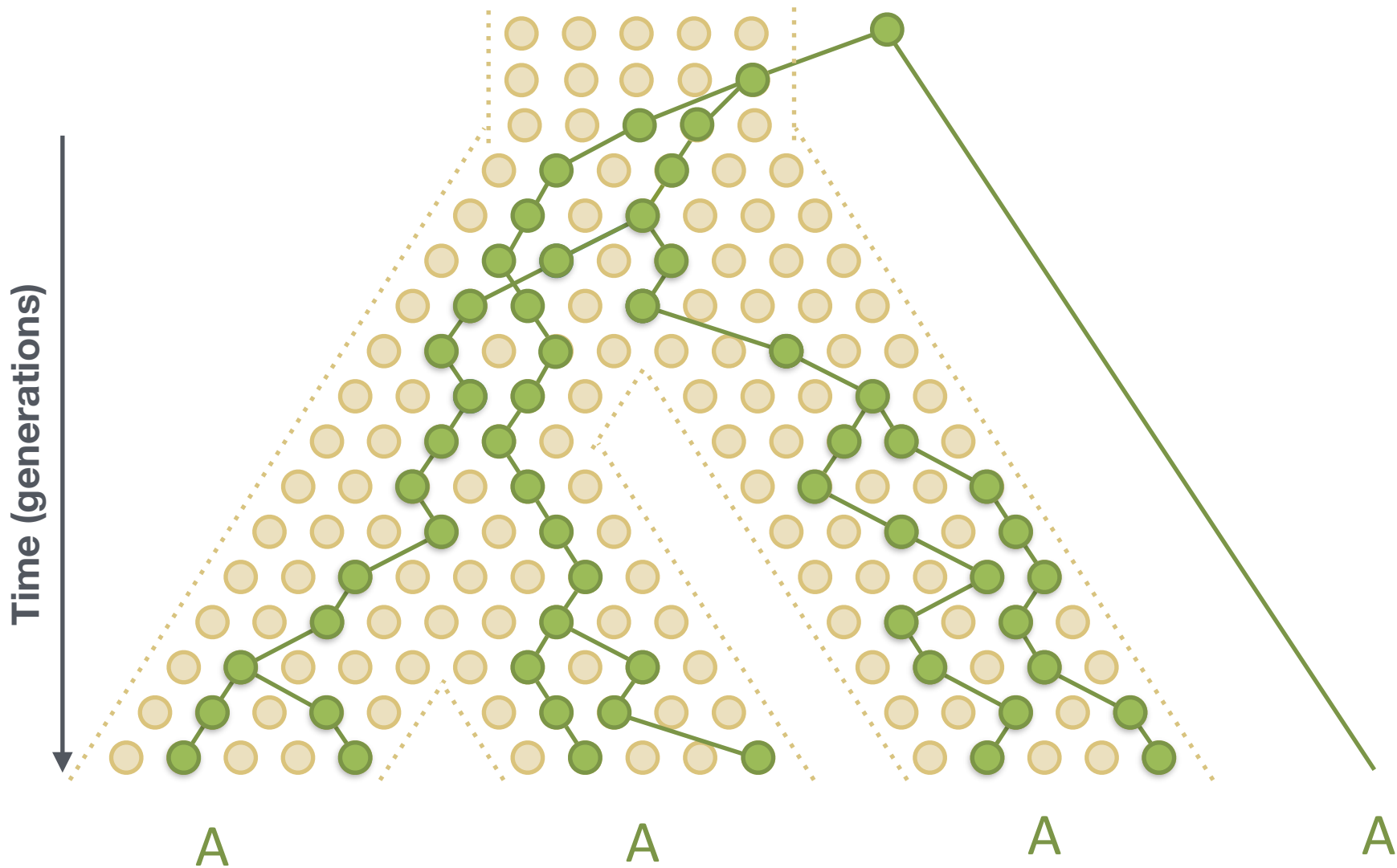
The multi-species coalescent model



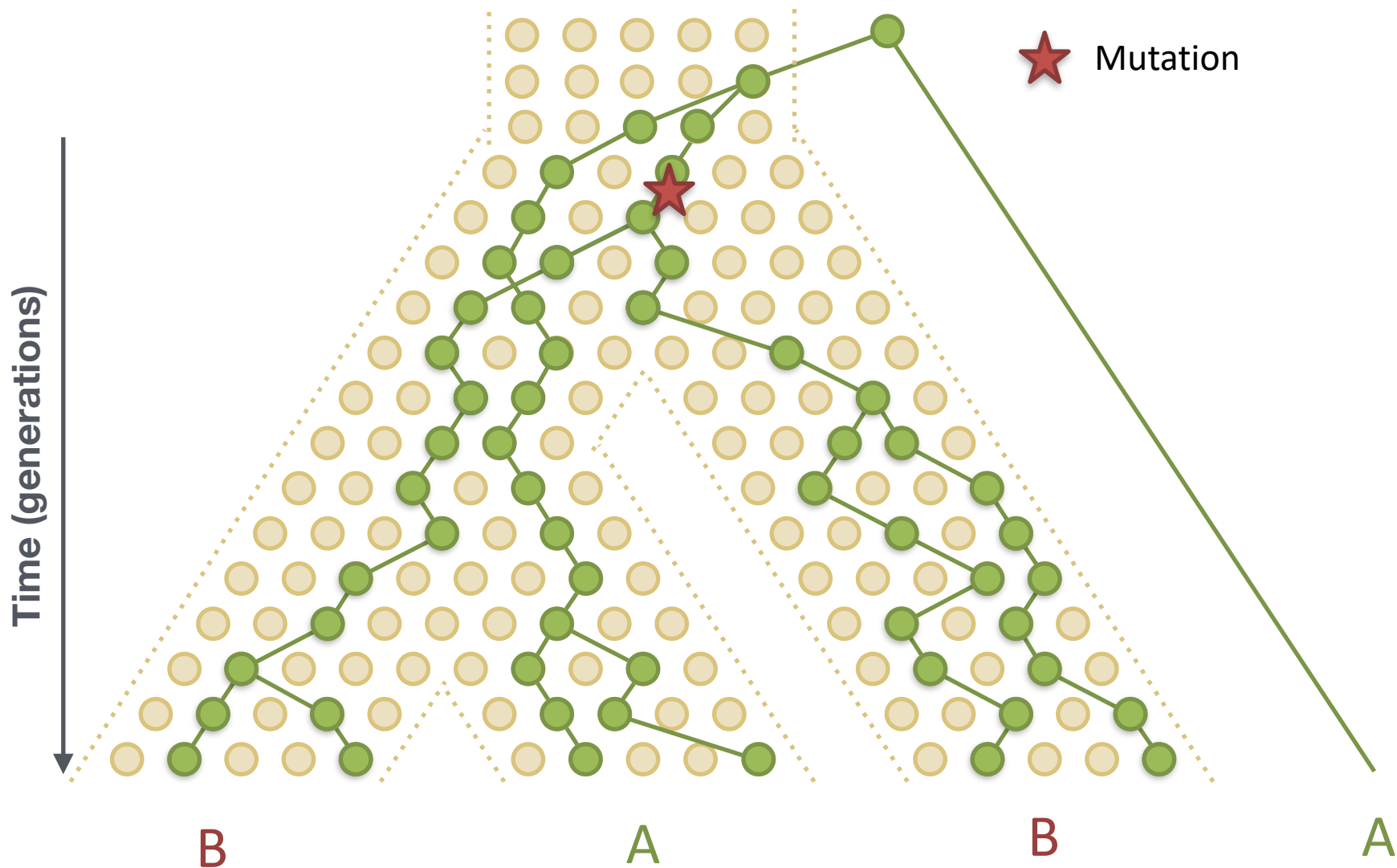
The multi-species coalescent model



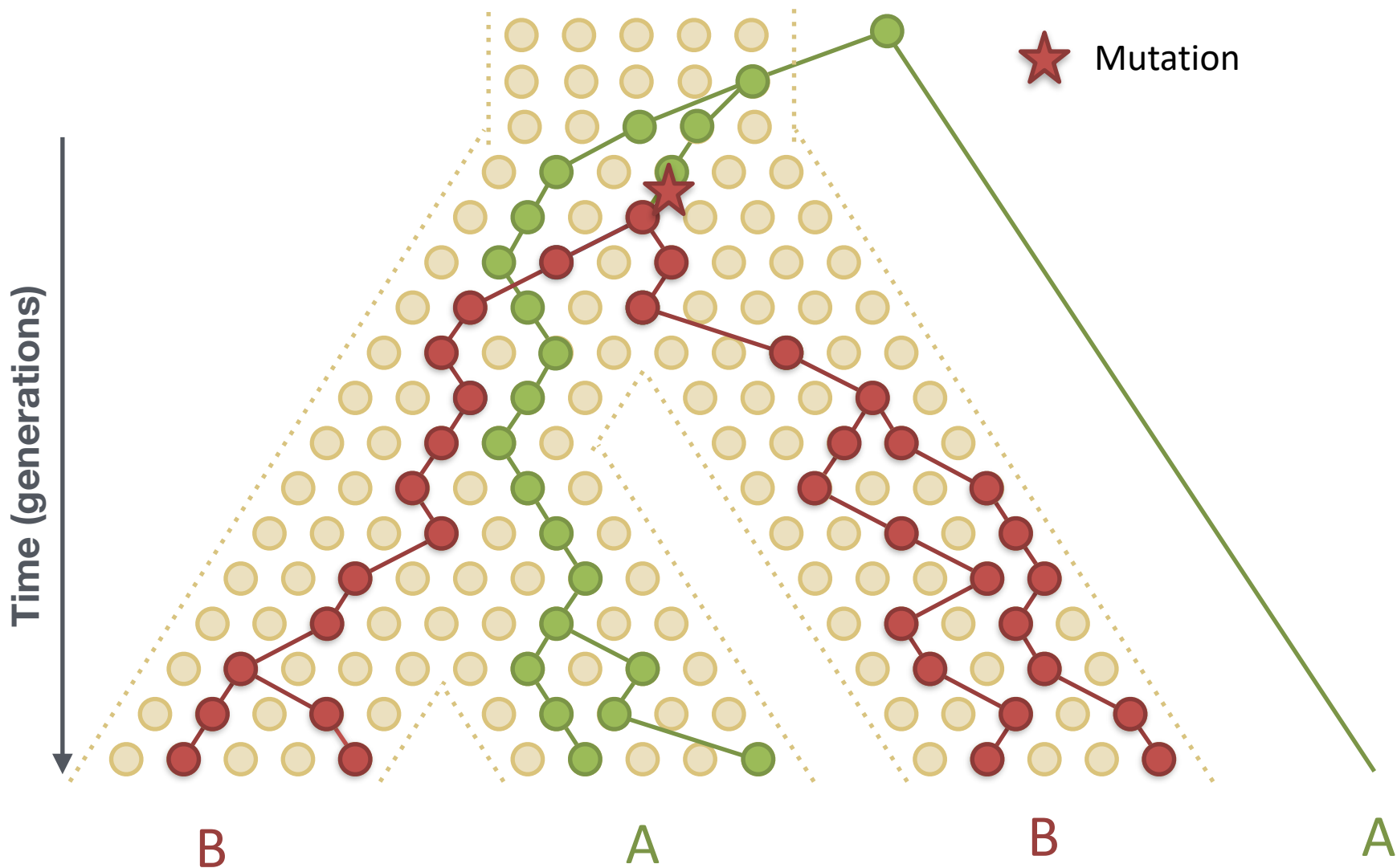
Incomplete lineage sorting



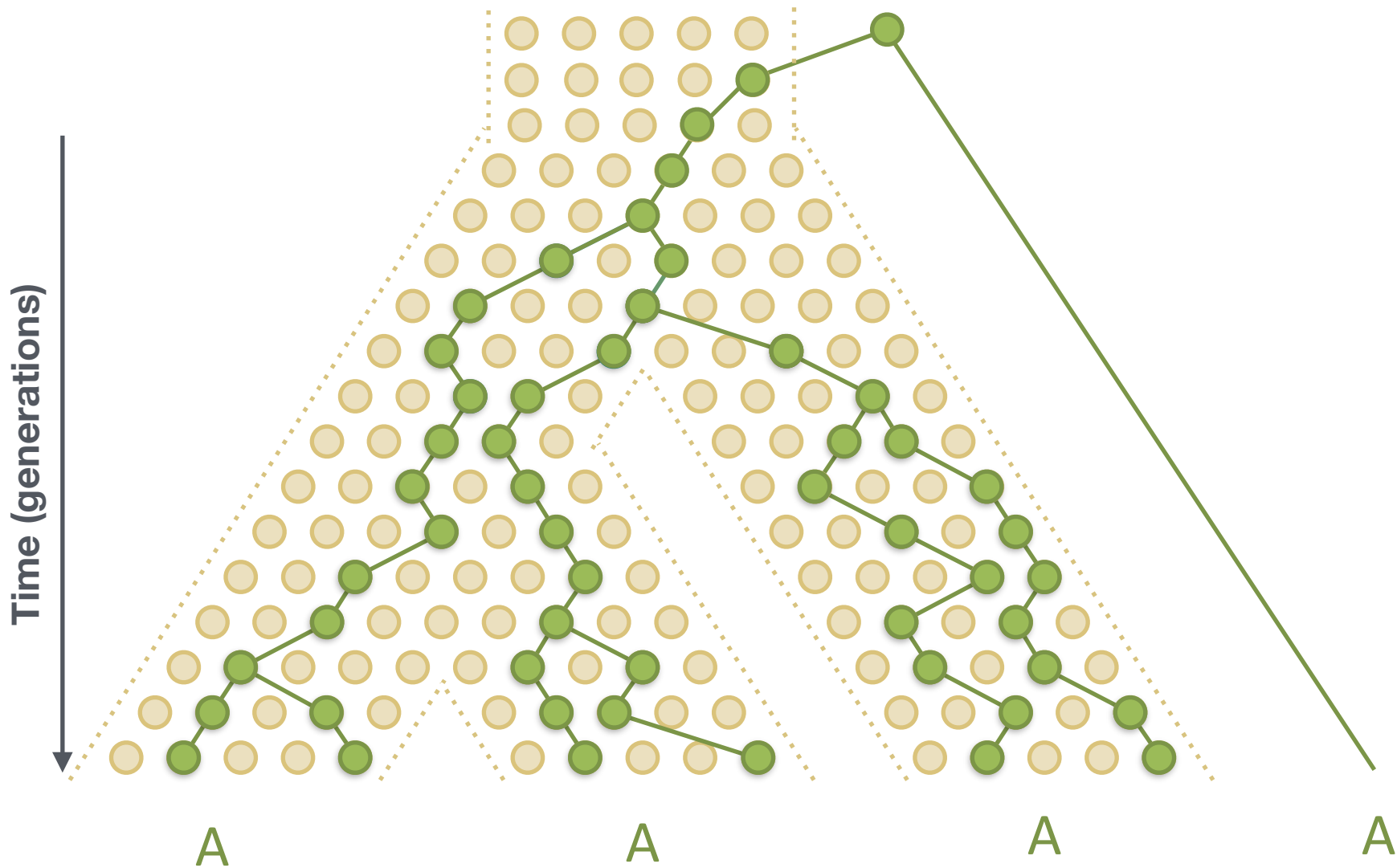
Incomplete lineage sorting



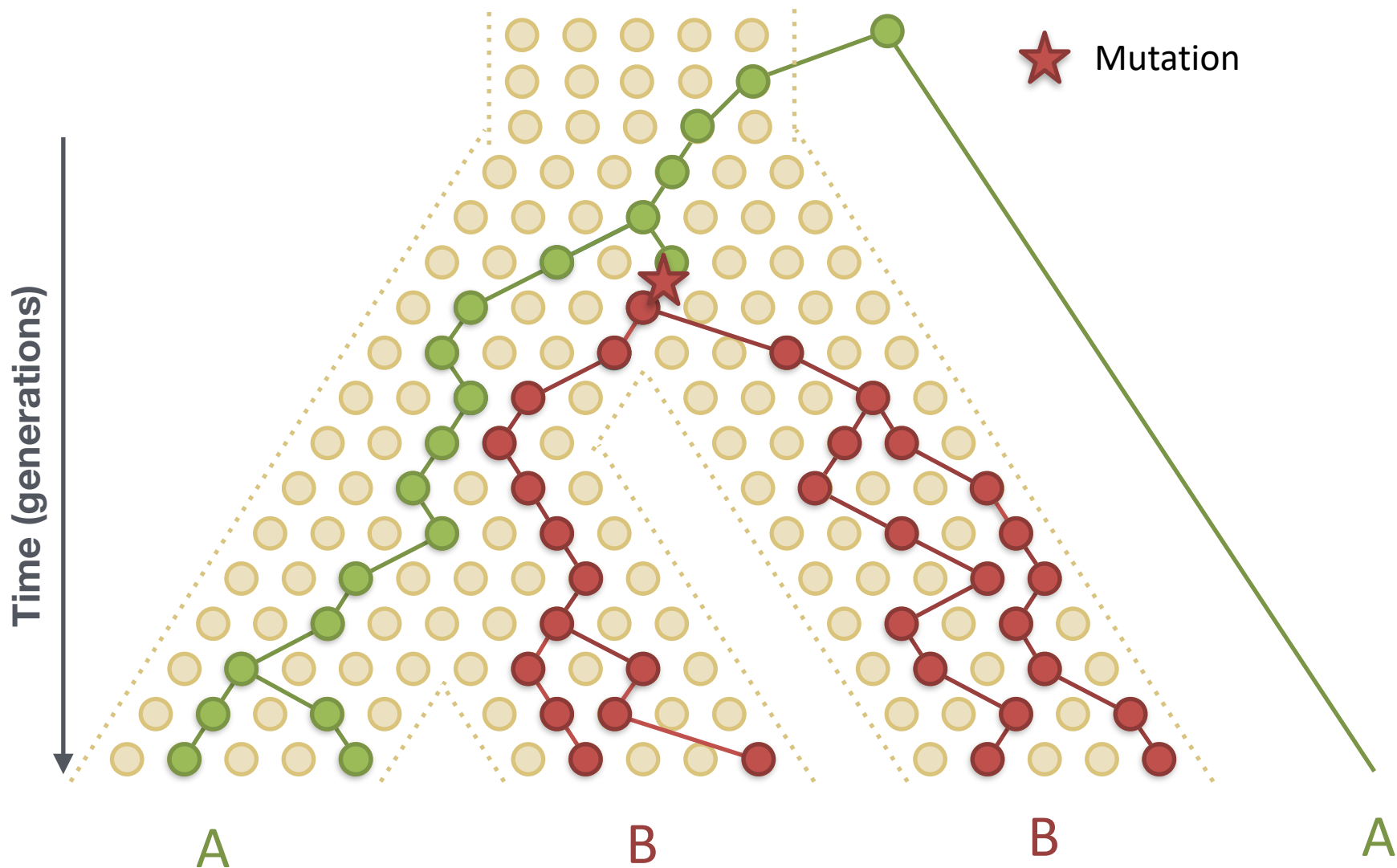
Incomplete lineage sorting



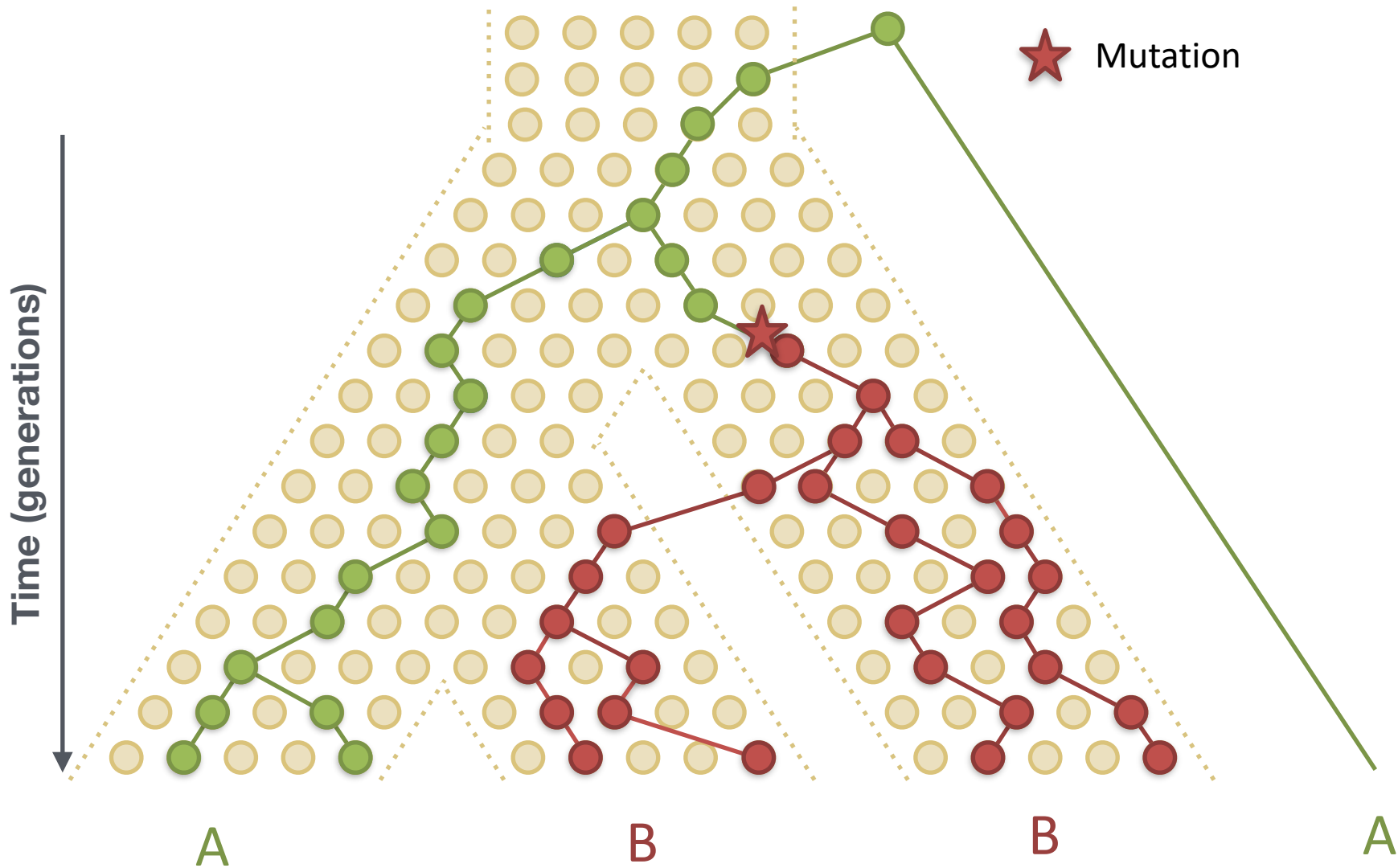
Incomplete lineage sorting



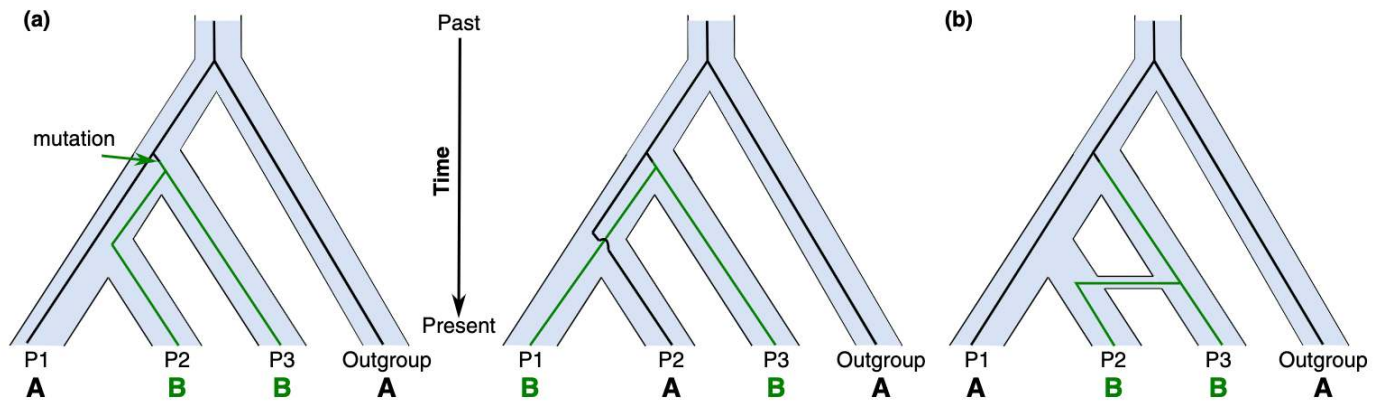
Incomplete lineage sorting



Introgression

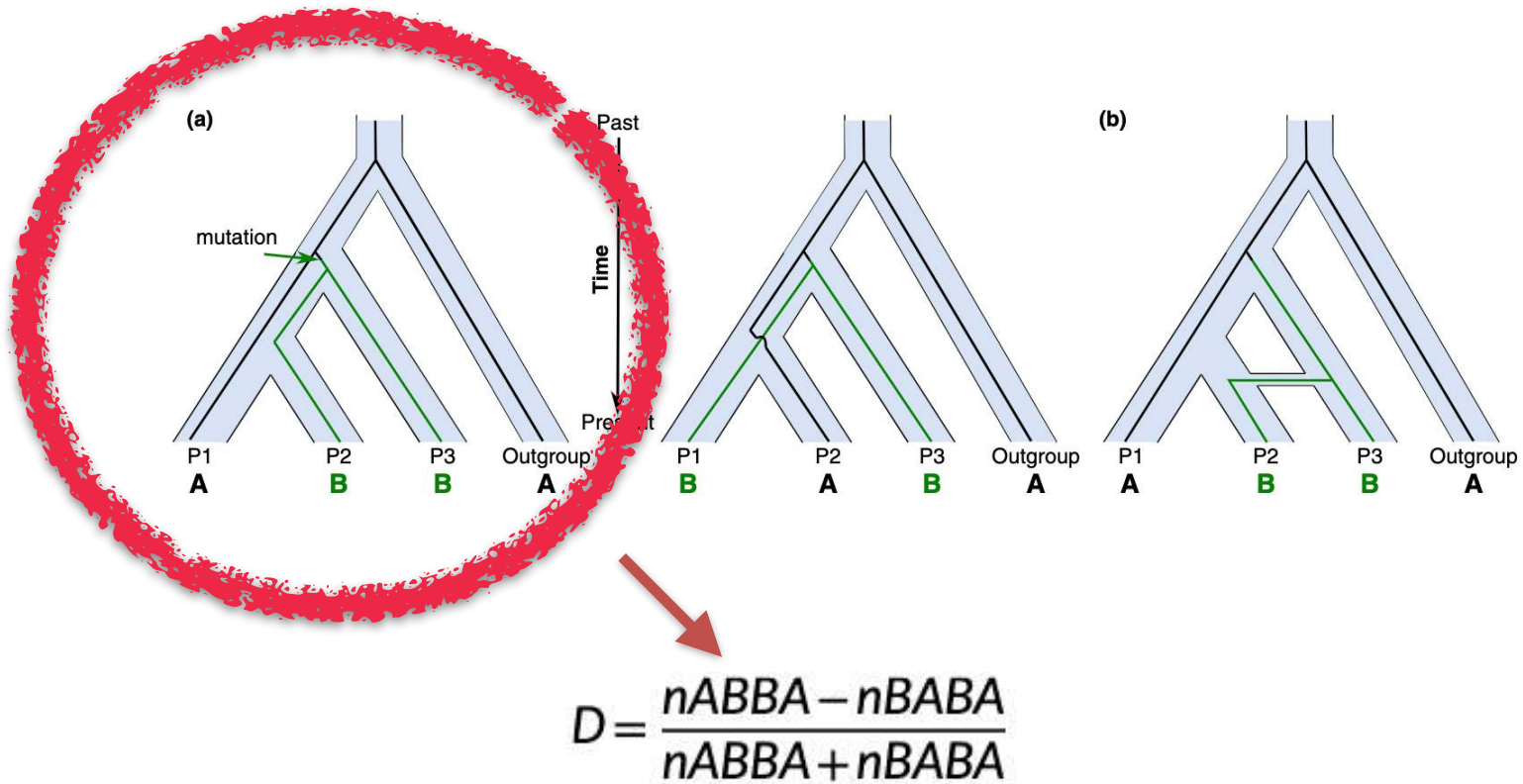


Introgression

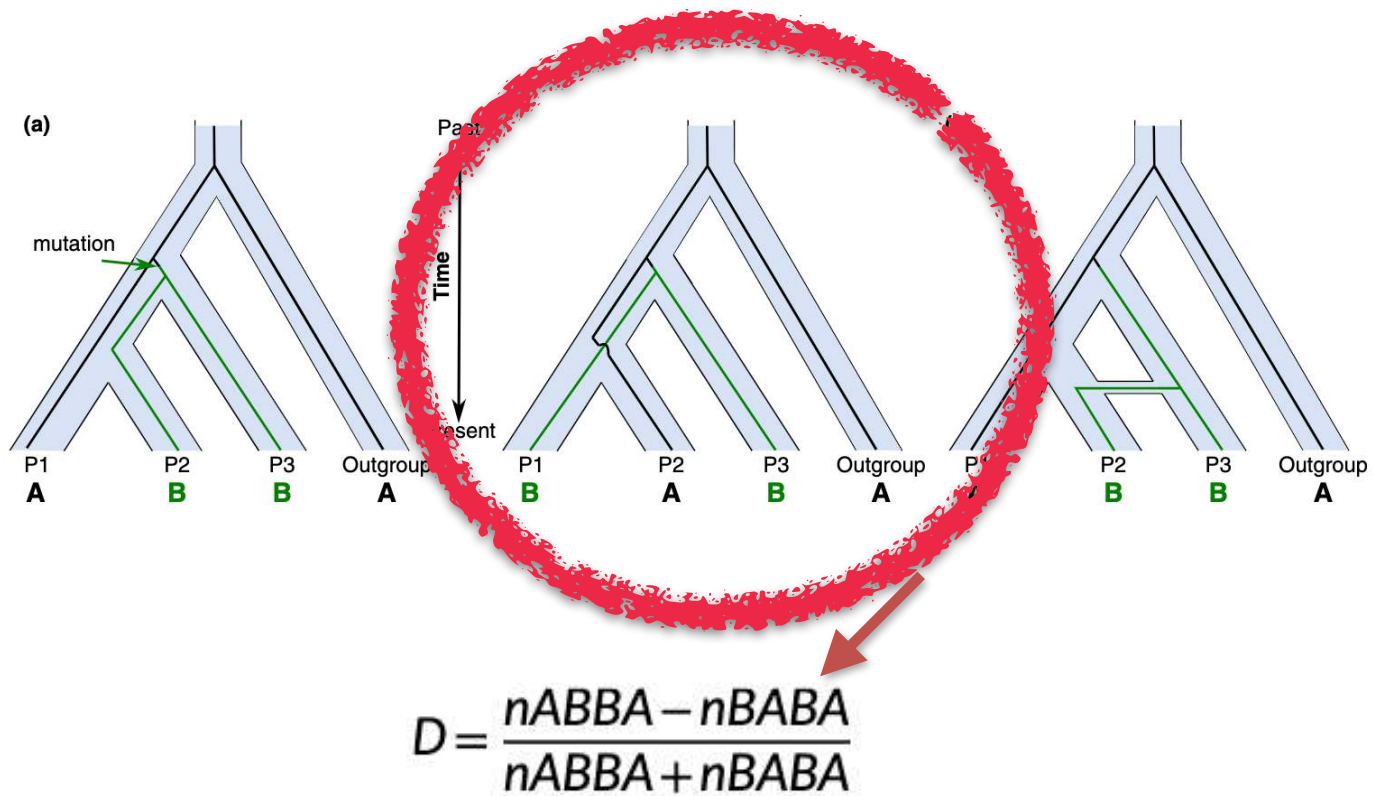


$$D = \frac{n_{ABBA} - n_{BABA}}{n_{ABBA} + n_{BABA}}$$

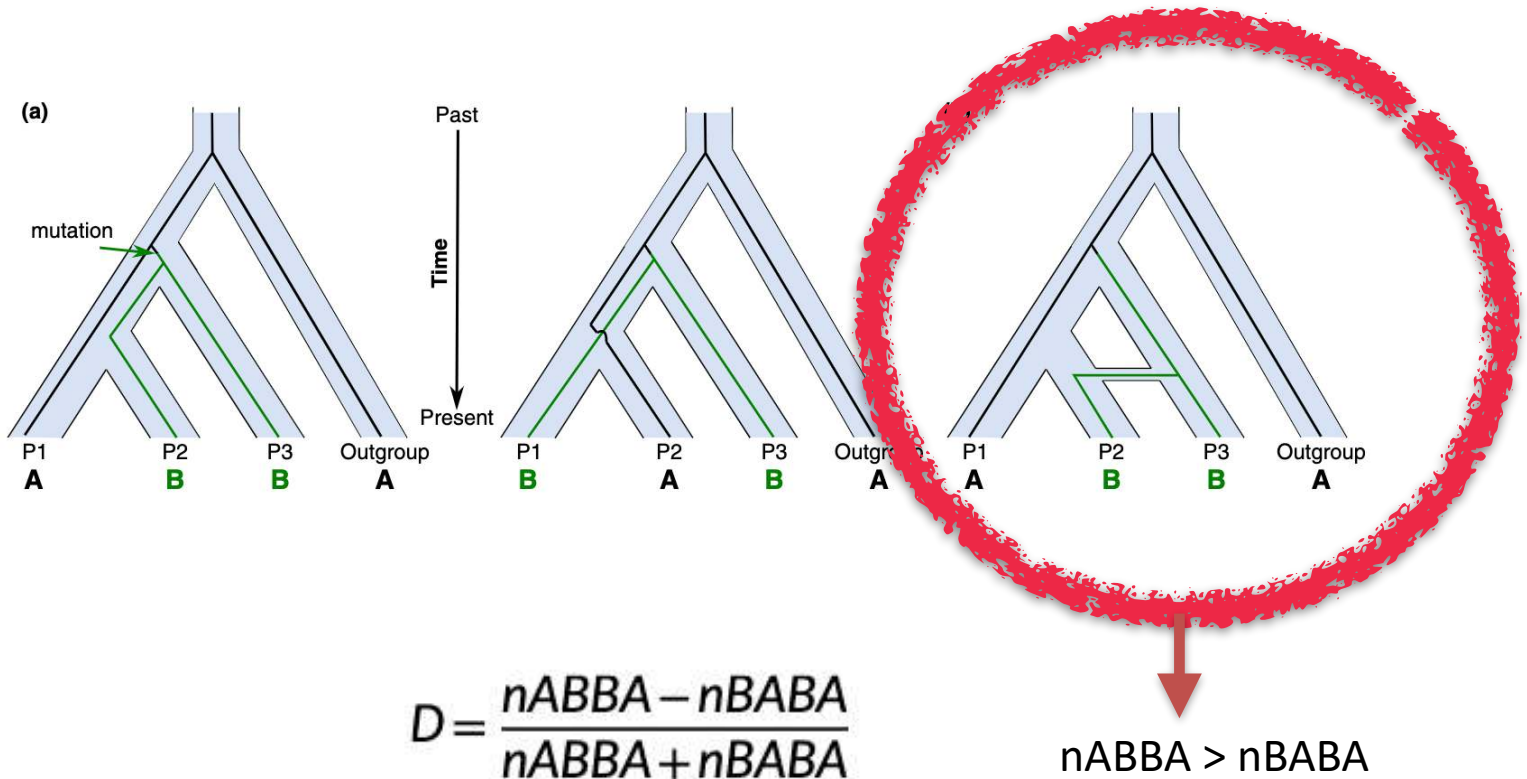
Introgression



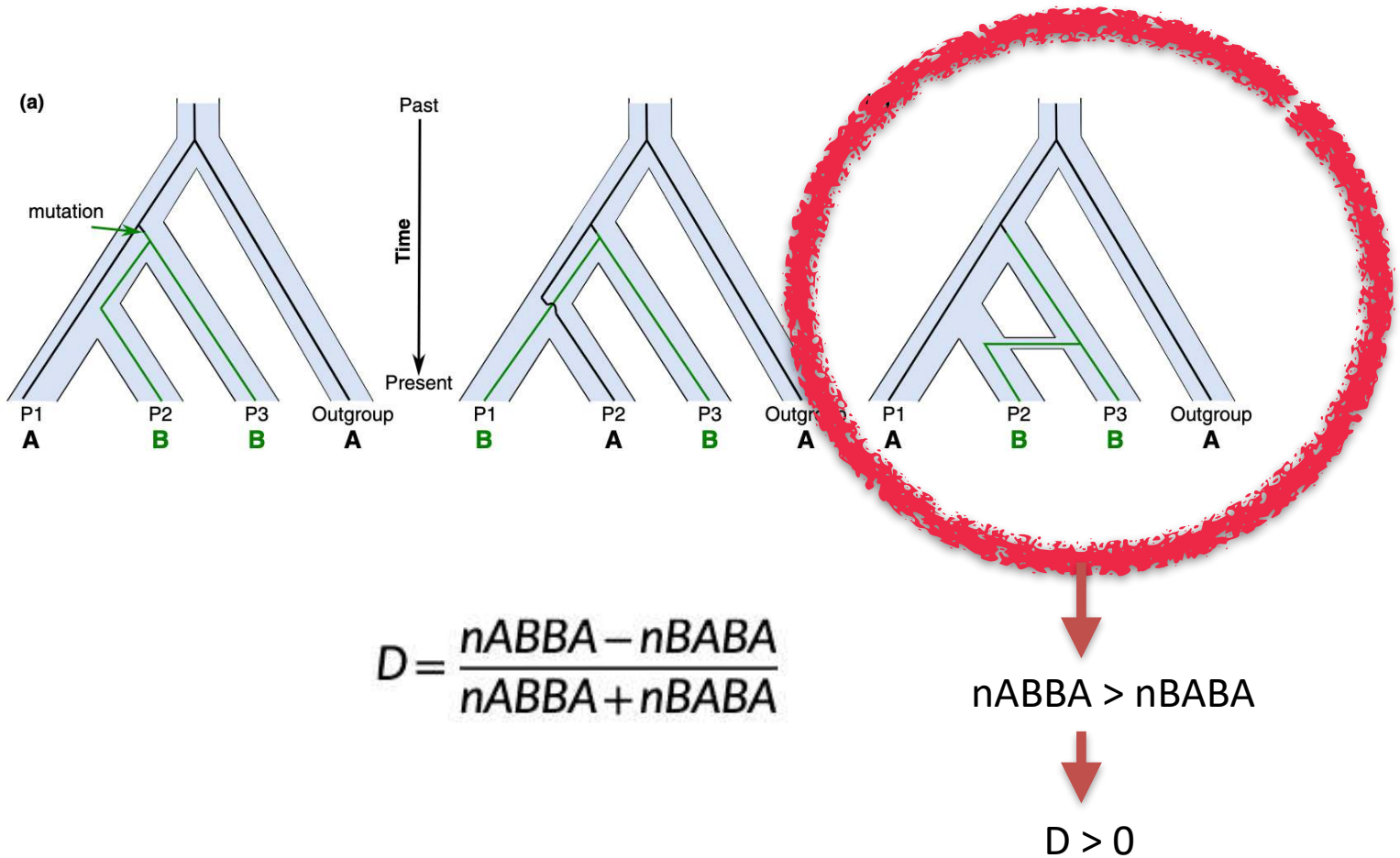
Introgression



Introgression



Introgression



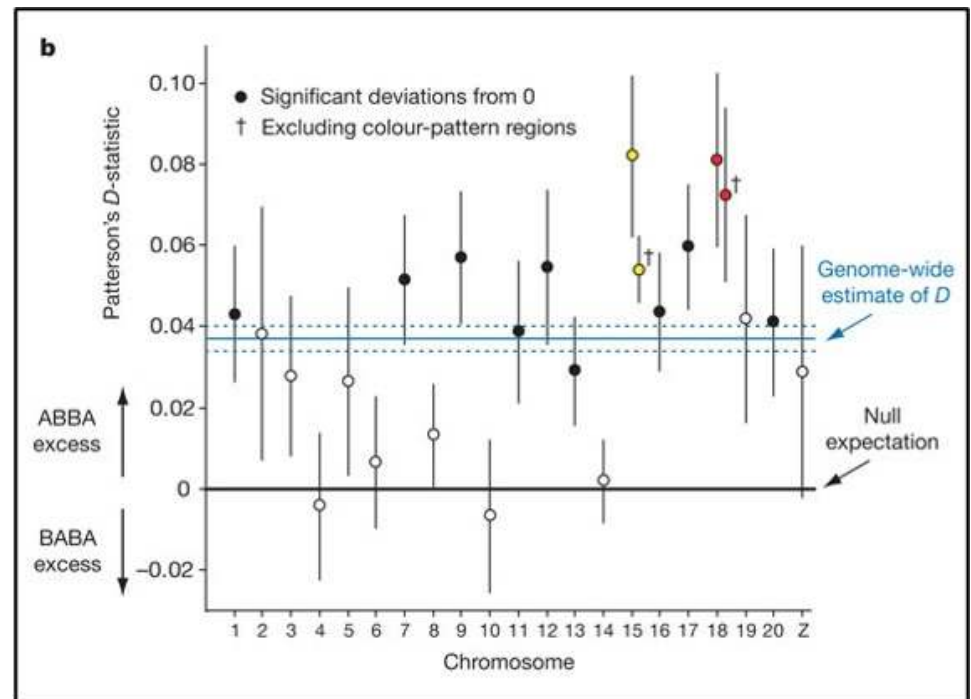
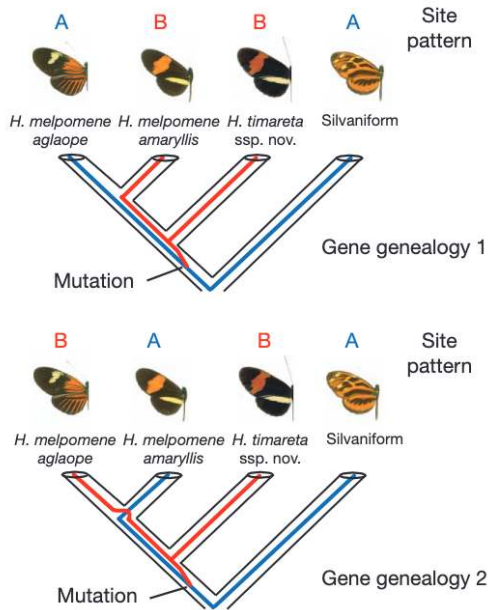
Introgression

LETTER

doi:10.1038/nature11041

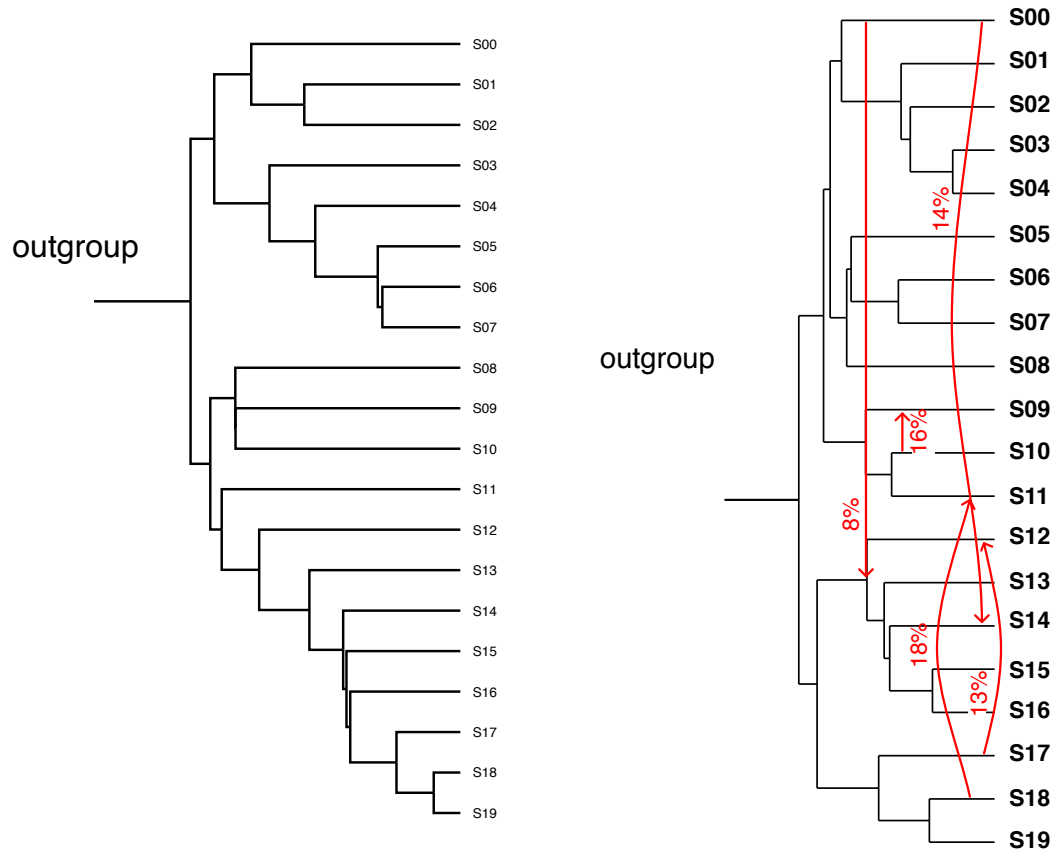
Butterfly genome reveals promiscuous exchange of mimicry adaptations among species

The *Heliconius* Genome Consortium*



Activity:

Inferring gene flow in simulated data



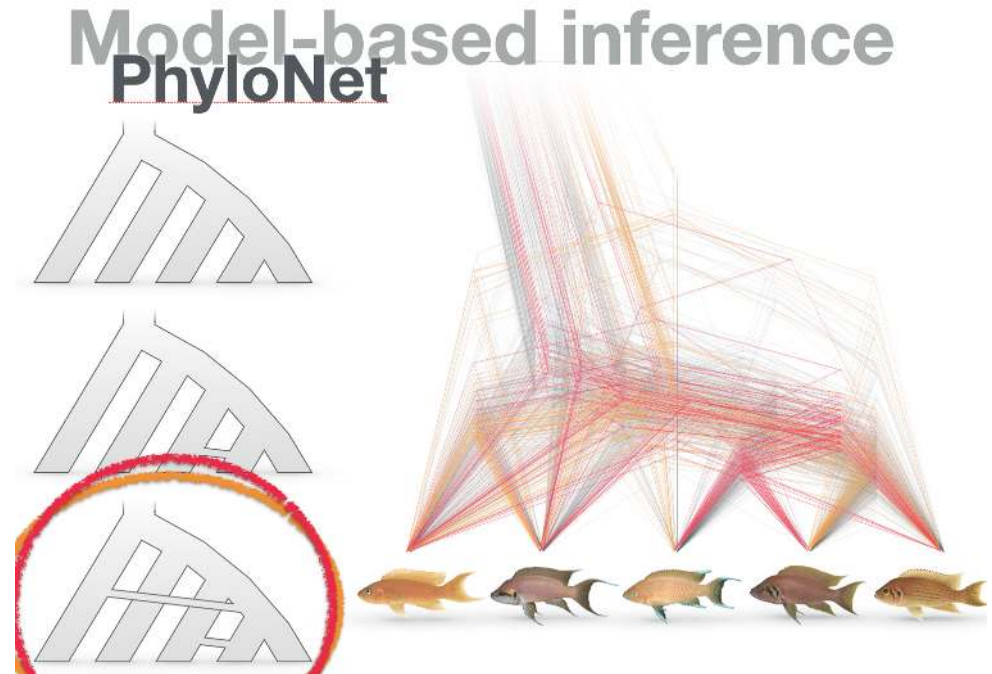
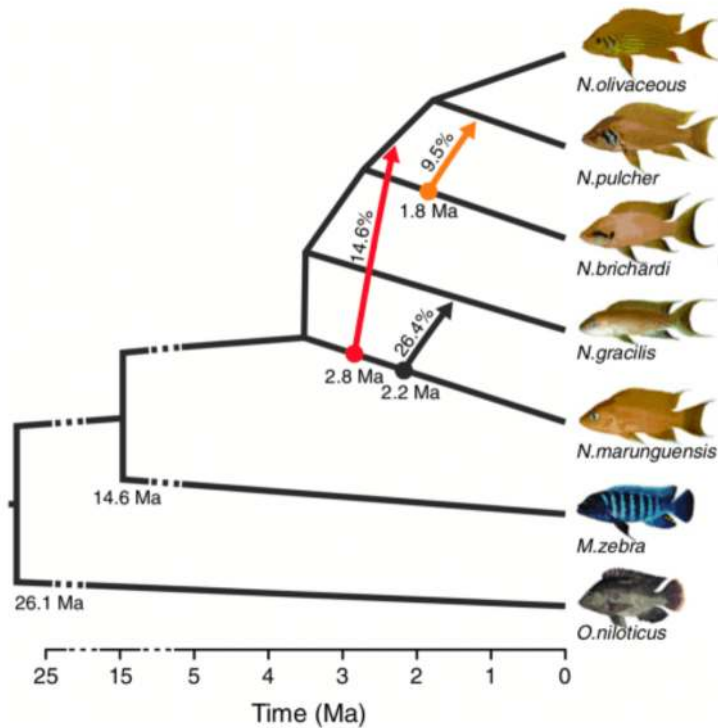
True tree, no introgression

True tree, five introgression
events

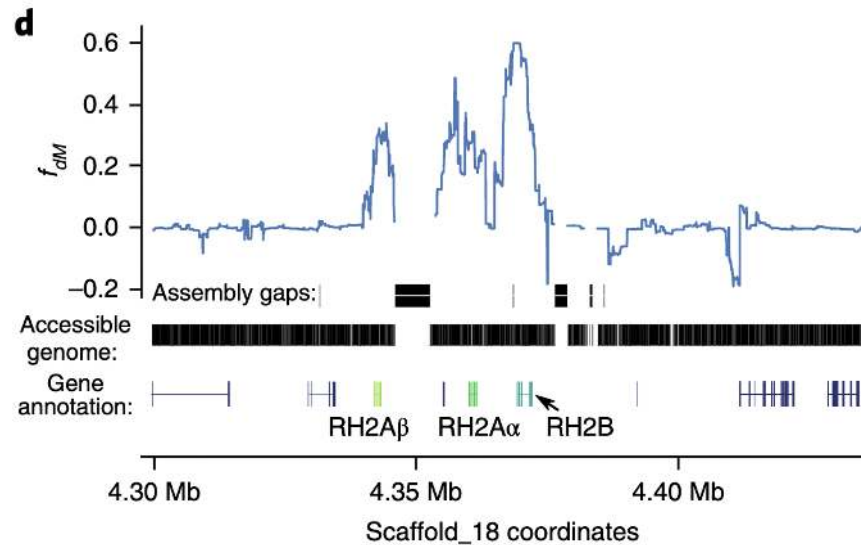
Simulated data (VCF):

- 20 species in 1 million generations
- $N_e = 50,000$
- $\mu, \rho = 1 \times 10^{-8}$
- 20Mb of data

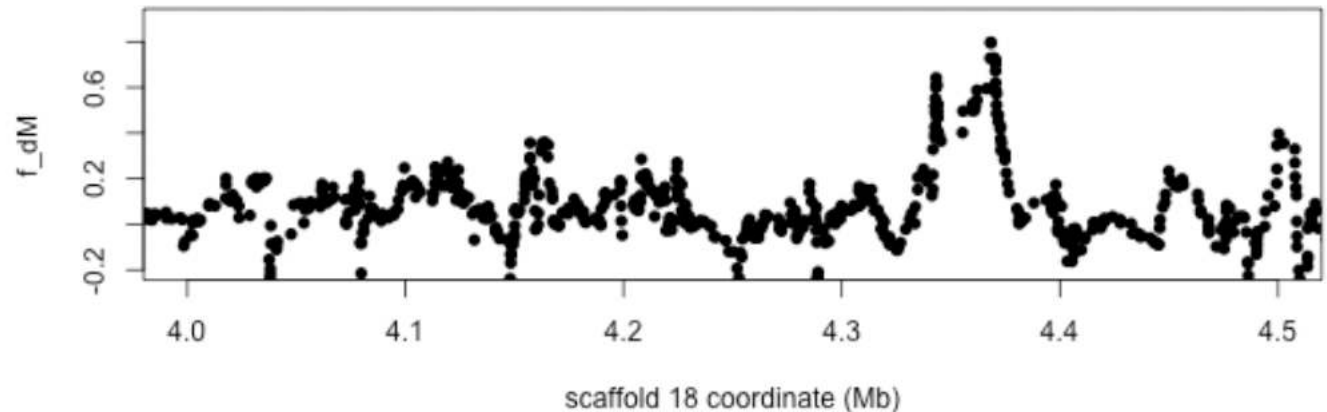
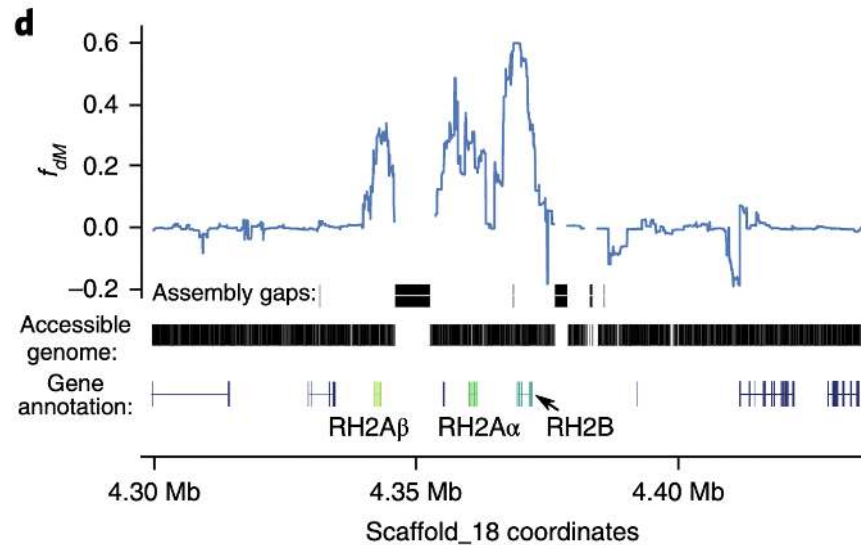
Activity: Finding gene-flow in a real dataset



Activity: Gene flow signal along the genome

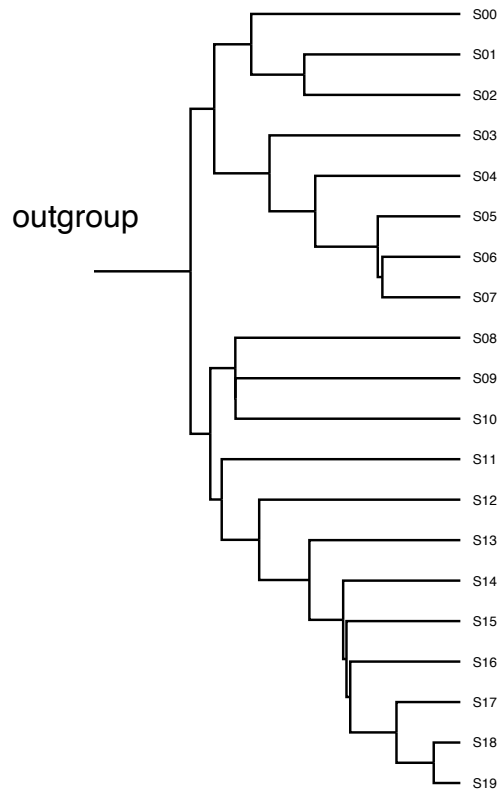


Activity: Gene flow signal along the genome

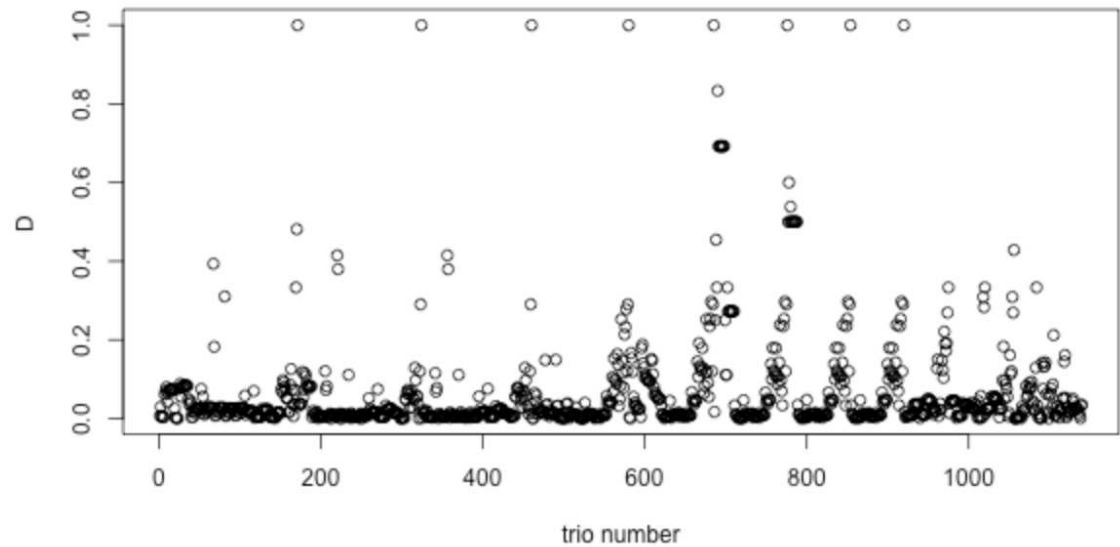


Activity: Results

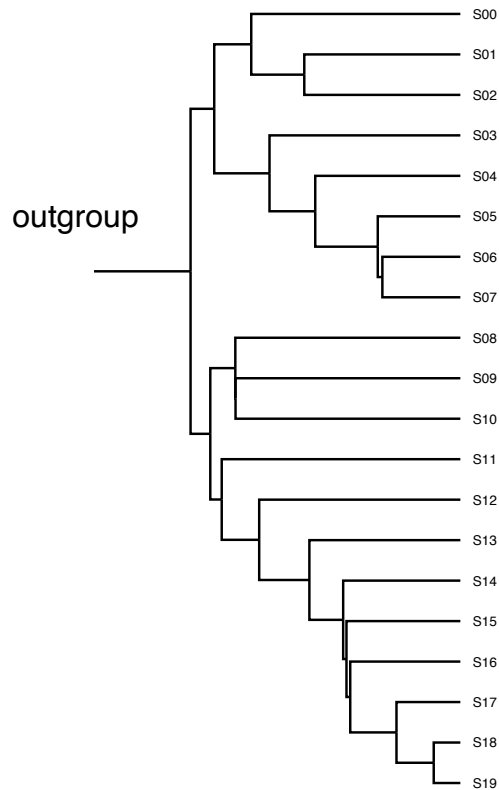
Introgression Activity



True tree, no introgression



Introgression Activity



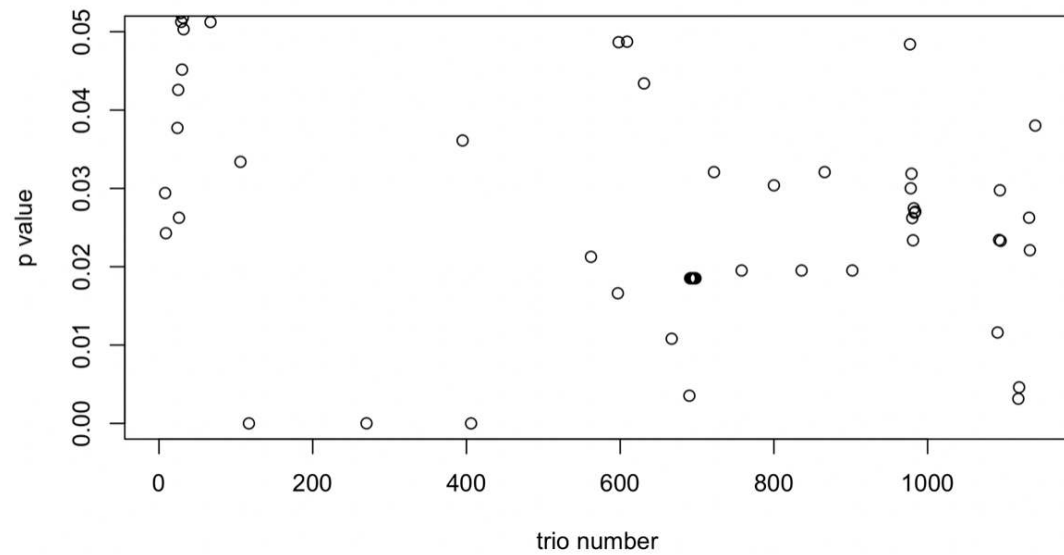
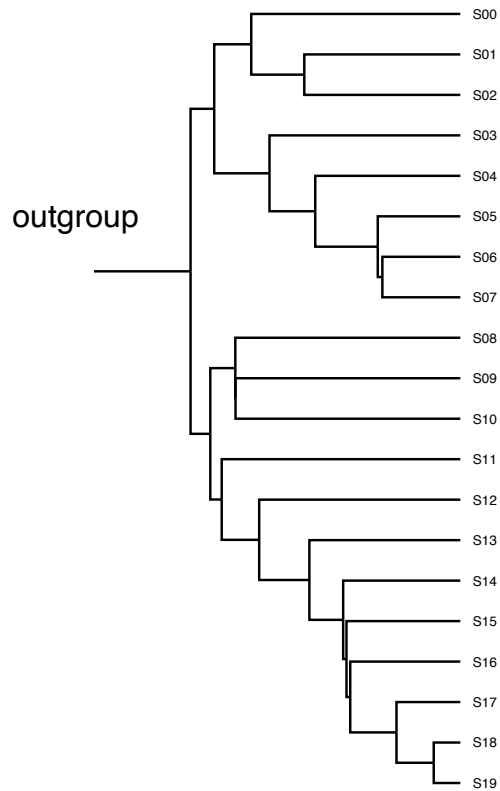
$$D = \frac{nABBA - nBABA}{nABBA + nBABA}$$

D_BBAA_noGF[which(D_BBAA_noGF\$Dstatistic > 0.7),]											
	P1	P2	P3	Dstatistic	Z.score	p.value	f4.ratio	BBAA	ABBA	BABA	
171	S18	S19	S00	1.000000	3.01176e-314	NaN	7.33365e-06	179922	1.5	0	
324	S18	S19	S01	1.000000	3.01176e-314	NaN	7.33100e-06	179814	1.5	0	
460	S18	S19	S02	1.000000	3.01176e-314	NaN	7.31256e-06	179800	1.5	0	
580	S18	S19	S03	1.000000	3.01176e-314	NaN	7.34992e-06	179778	1.5	0	
685	S18	S19	S04	1.000000	3.01176e-314	NaN	7.34833e-06	179765	1.5	0	
690	S06	S05	S11	0.833333	2.91667e+00	0.00353794	4.88206e-05	138120	11.0	1	
776	S18	S19	S05	1.000000	3.01176e-314	NaN	7.34066e-06	179768	1.5	0	
854	S18	S19	S06	1.000000	3.01176e-314	NaN	7.35210e-06	179777	1.5	0	
920	S18	S19	S07	1.000000	3.01176e-314	NaN	7.33584e-06	179771	1.5	0	

True tree, no introgression

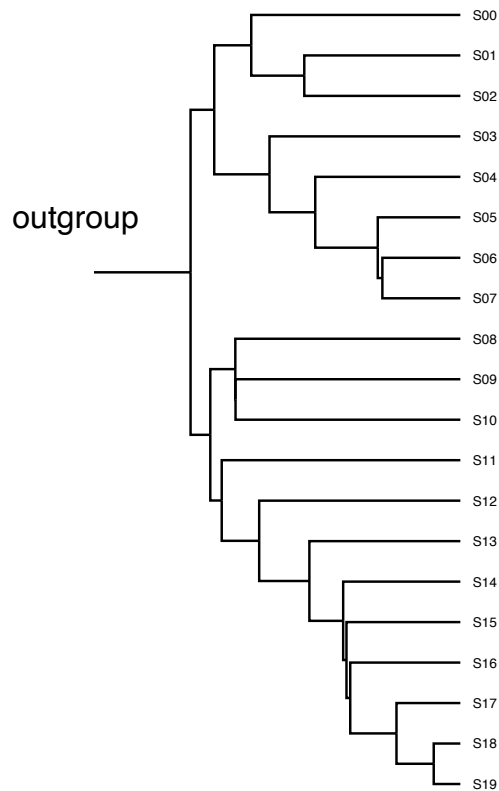
[illegible]

Introgression Activity



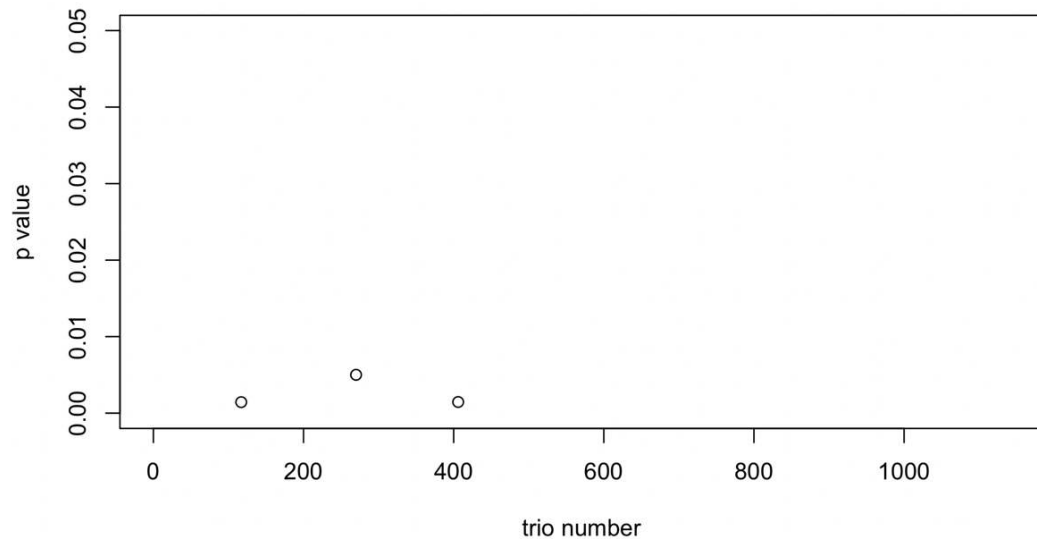
True tree, no introgression

Introgression Activity

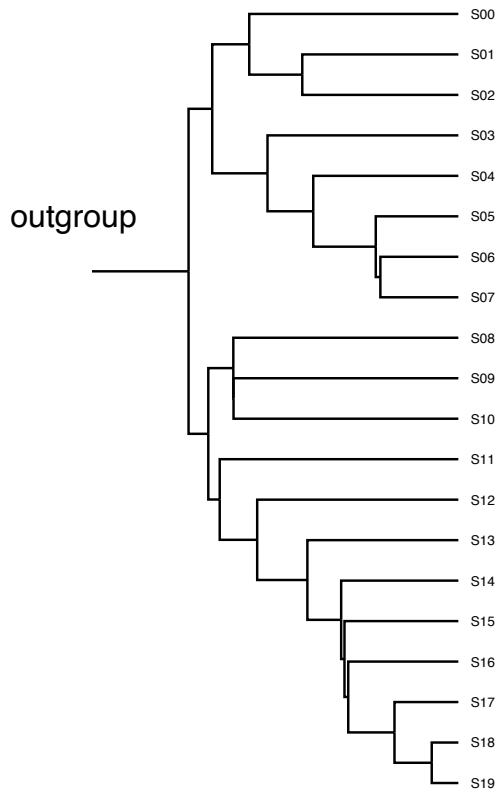


True tree, no introgression

Three significant values after multiple testing correction
FDR=5%

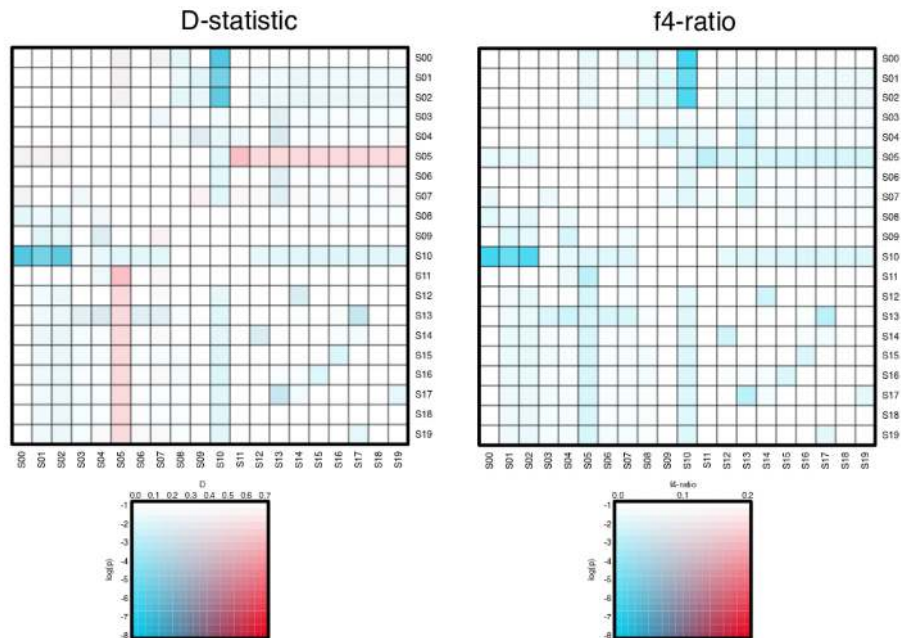


Introgression Activity

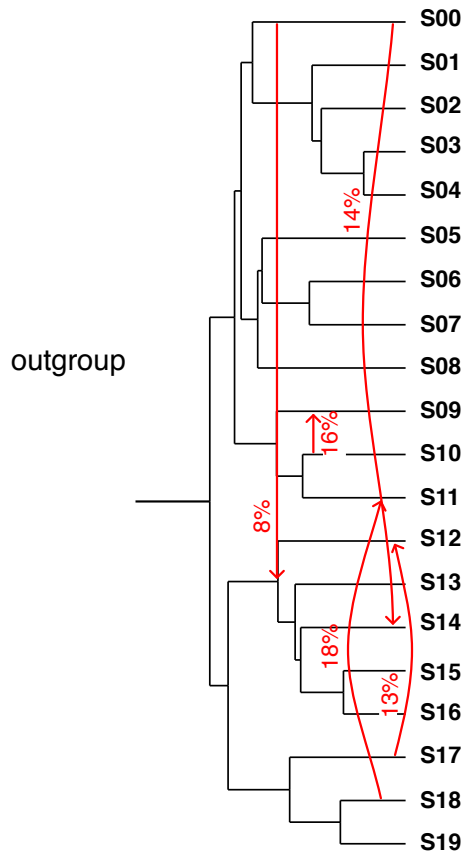


True tree, no introgression

1140 tests reduced to 400 by aggregating over P1

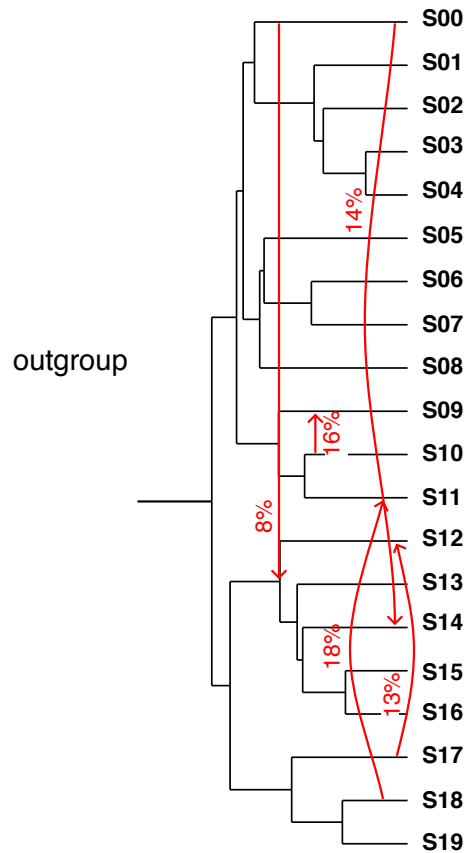


Introgression Activity

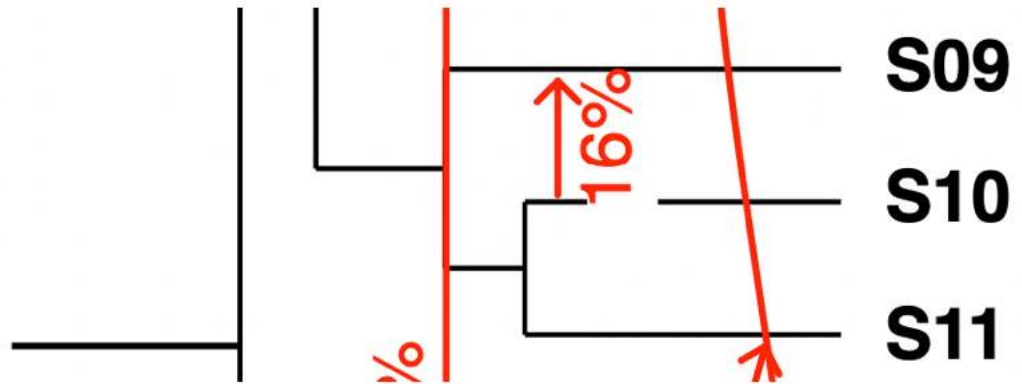


True tree, five introgression events

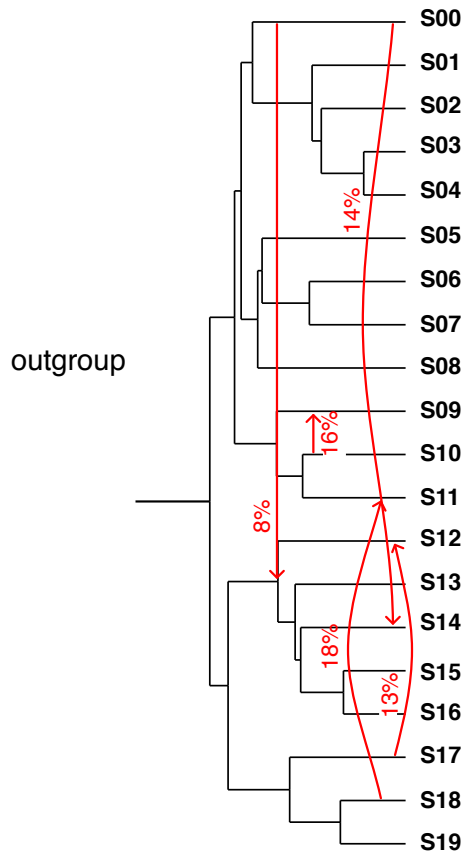
Introgression Activity



True tree, five introgression events



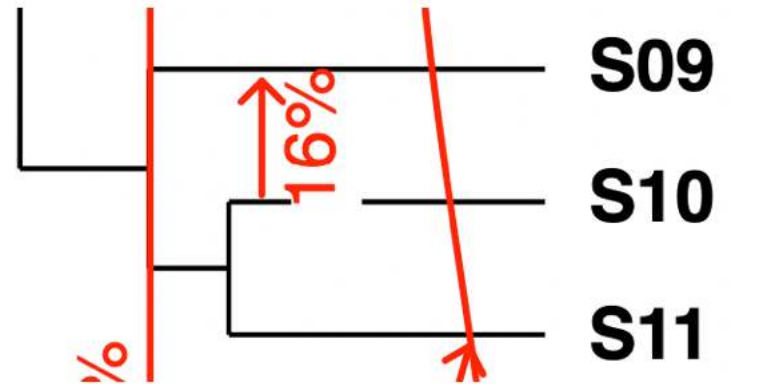
Introgression Activity



True tree, five introgression events

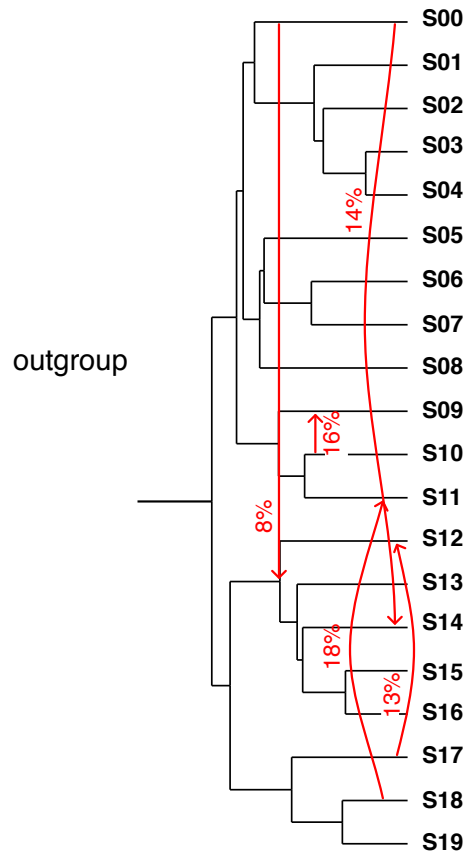


Dstatistic
0.805219

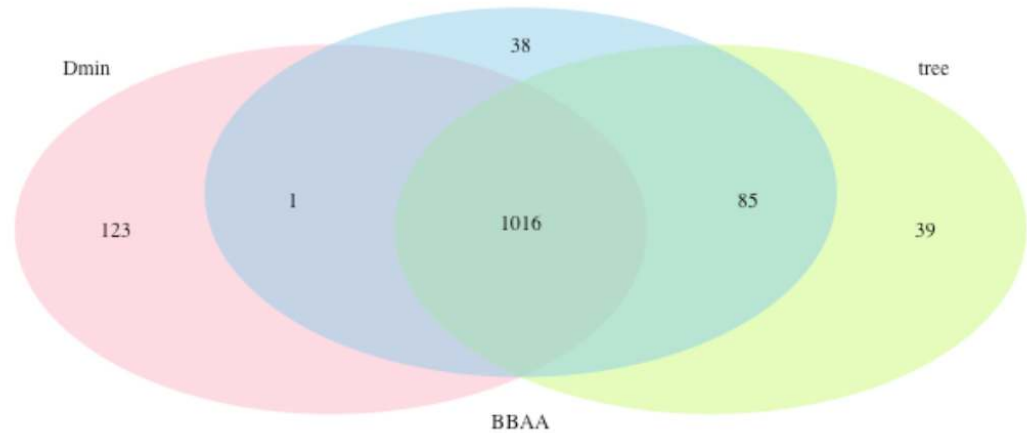


Z-score p-value f4-ratio
72.7811 2.3e-16 0.0950458

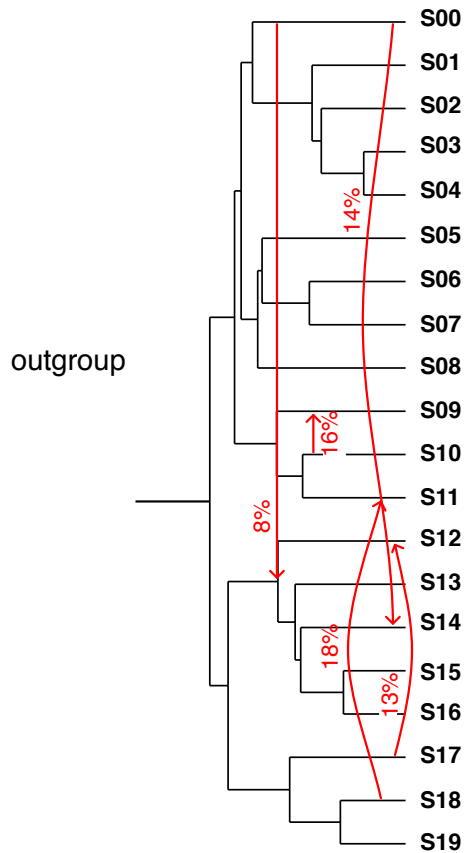
Introgression Activity



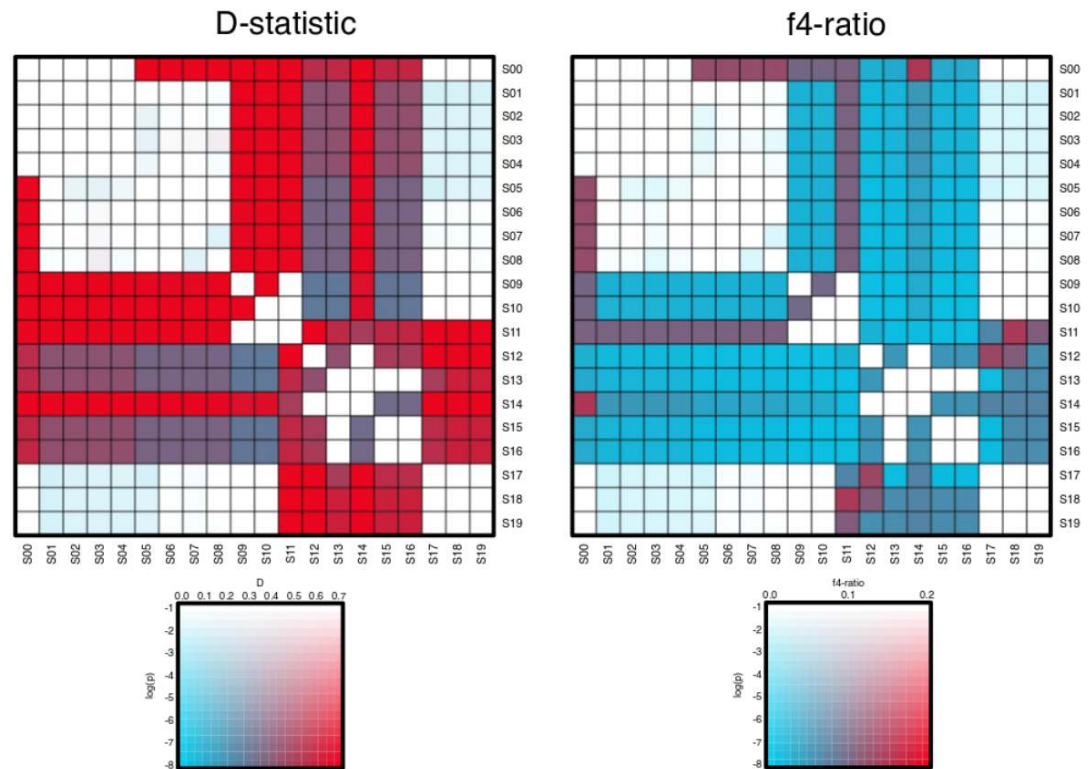
True tree, five introgression events

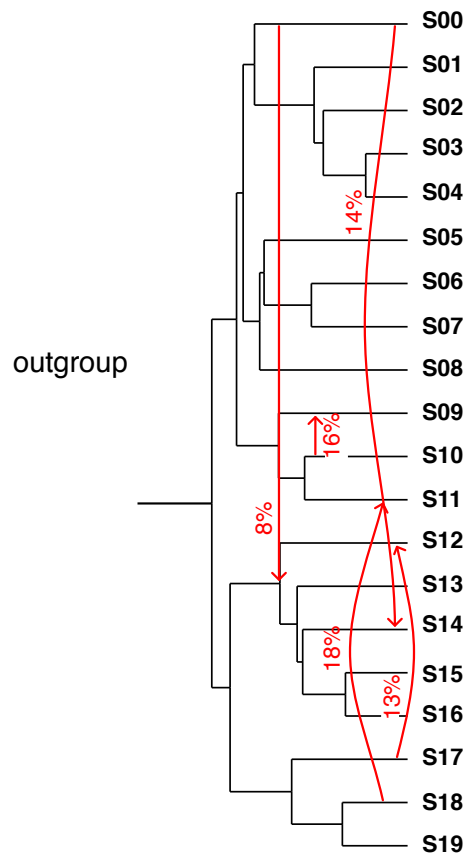


Introgression Activity

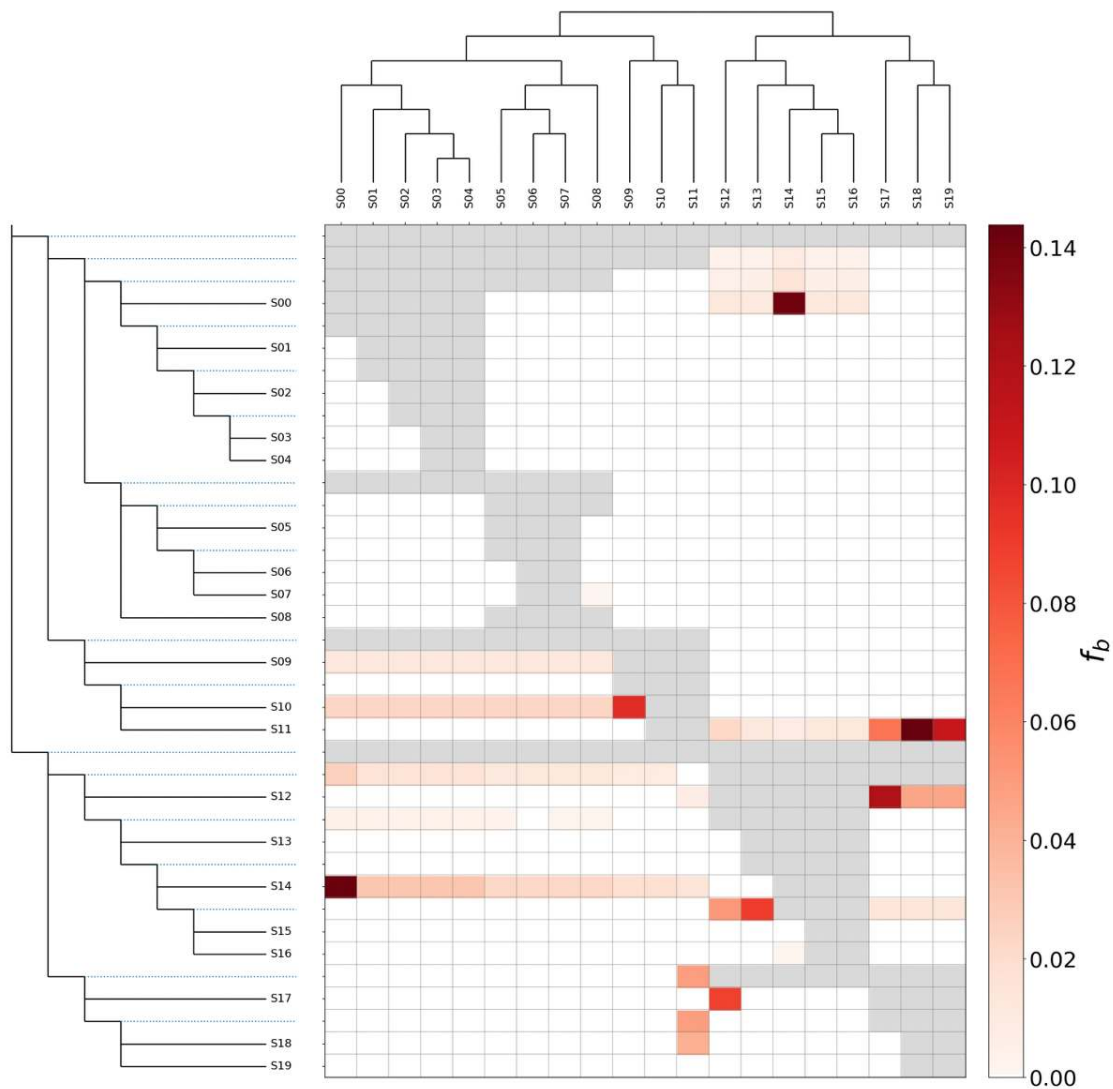


True tree, five introgression events

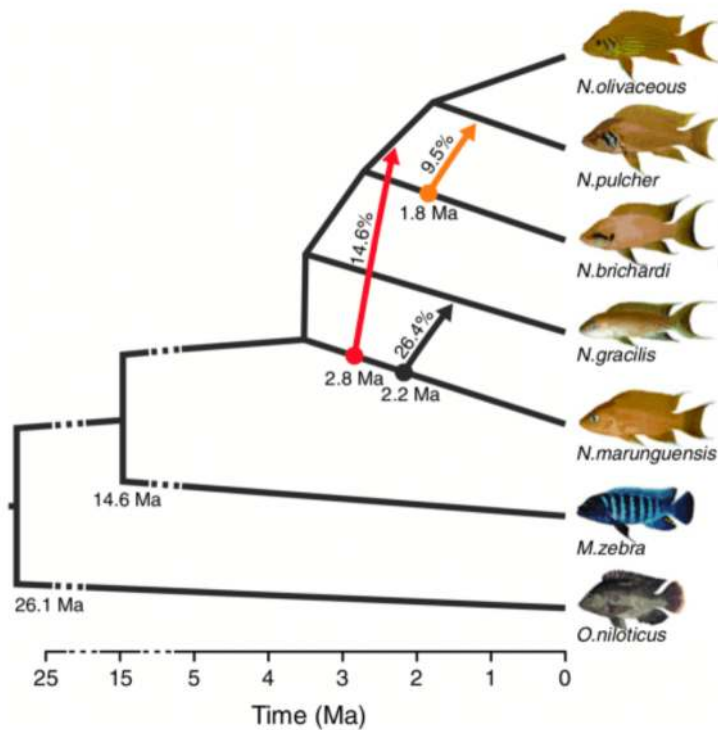




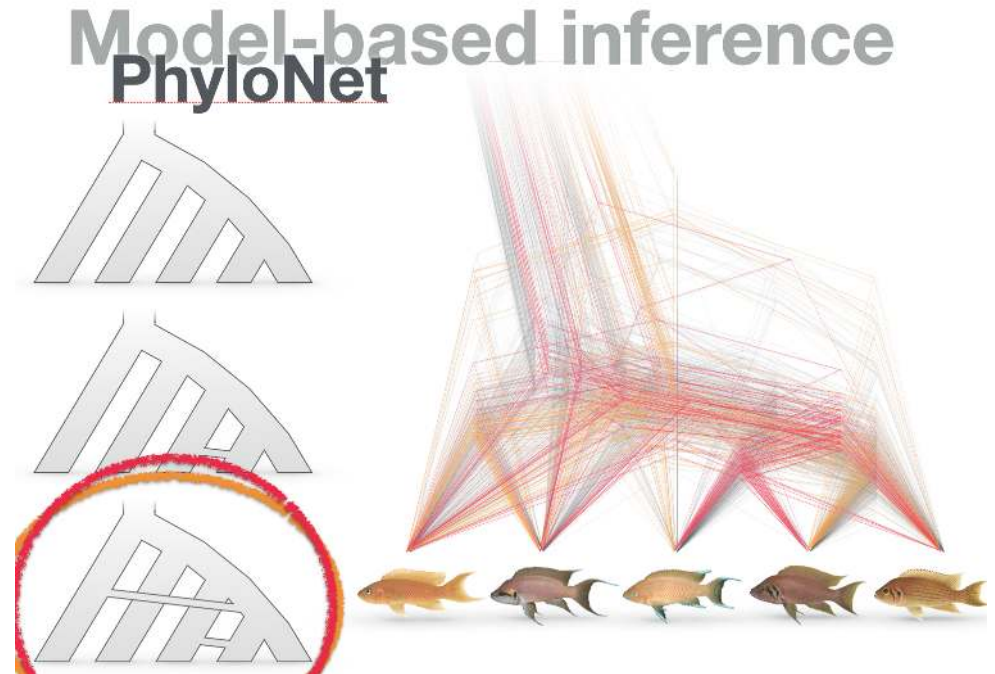
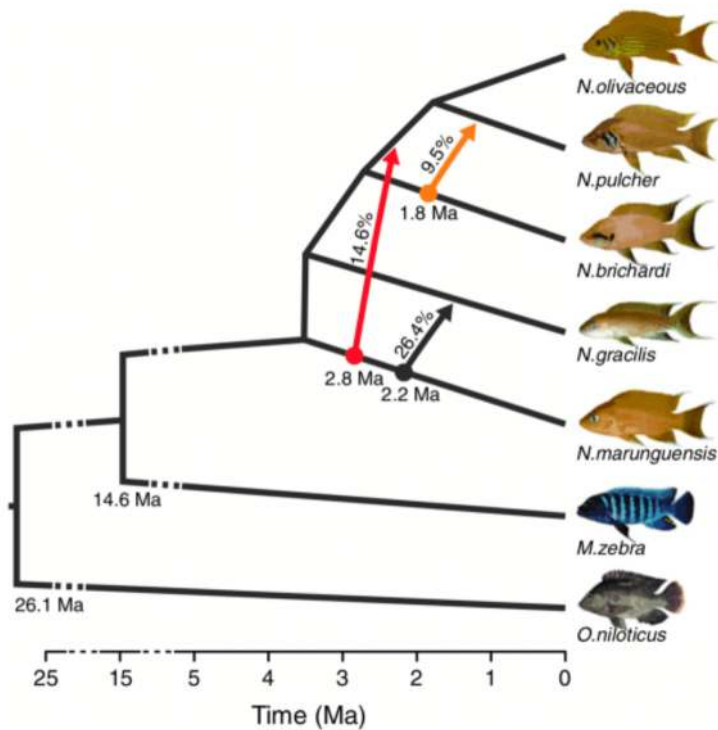
True tree, five introgression events



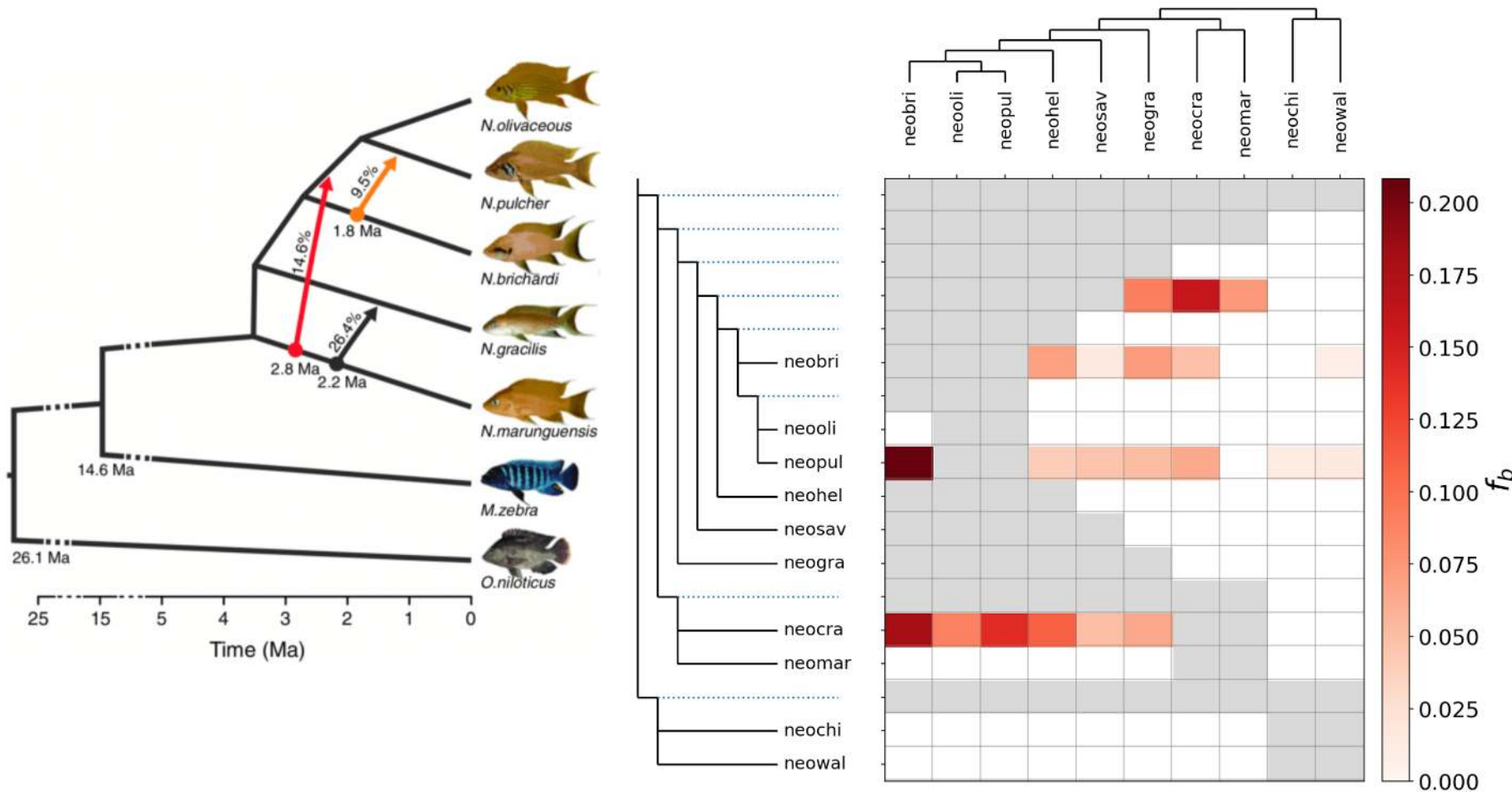
3. Finding geneflow in a real dataset - Tanganyikan cichlids



3. Finding geneflow in a real dataset - Tanganyikan cichlids



3. Finding geneflow in a real dataset - Tanganyikan cichlids



3. Finding geneflow in a real dataset - Tanganyikan cichlids

