

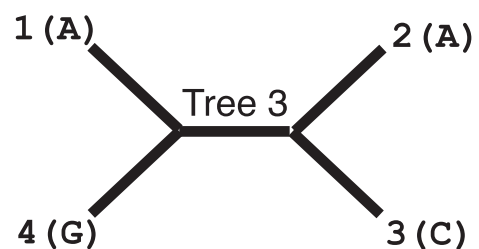
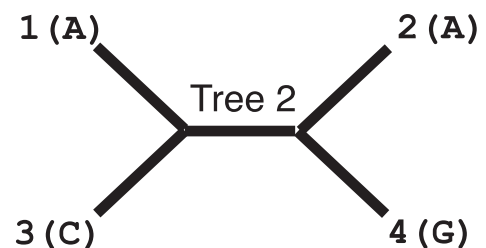
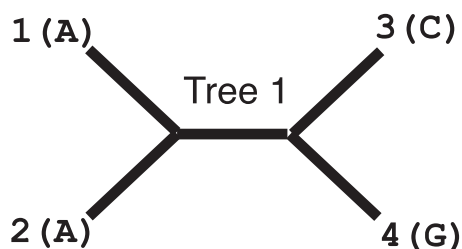
# Computing the likelihood of one column from a DNA alignment

---

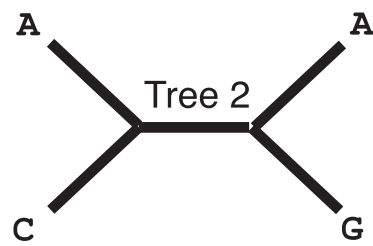
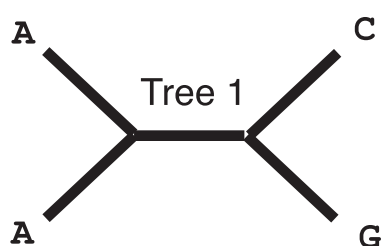
Below is an alignment of four sequences. Only the labeled column shows polymorphism:

```
1:  A A G C T
2:  A A G C T
3:  A C G C T
4:  A G G C T
   *
```

There are three possible trees linking these four species:



You will now use maximum likelihood to decide which tree is the best one. Since tree no. 2 and 3 are identical (except for the labels), we only need to compute the likelihood for tree no. 1 and 2.



# Computing the likelihood of one column from a DNA alignment

---

We will assume that all branches in the tree have the same length and that this branch length corresponds to the transition probability matrix shown below.

**P** = from

	A	C	to	G	T
A	0.67	0.12		0.16	0.04
C	0.08	0.72		0.16	0.04
G	0.08	0.12		0.76	0.04
T	0.08	0.12		0.16	0.65

The four nucleotides have the following equilibrium frequencies:

$$\pi_A = 0.2 \quad \pi_C = 0.3 \quad \pi_G = 0.4 \quad \pi_T = 0.1$$

The teacher will assign a tree and two or three combinations of ancestral states to your group. Compute the probability of the data (the likelihood of your tree) for each combination, write your results on the white board. We will combine all the results and discuss implications when all are done.

