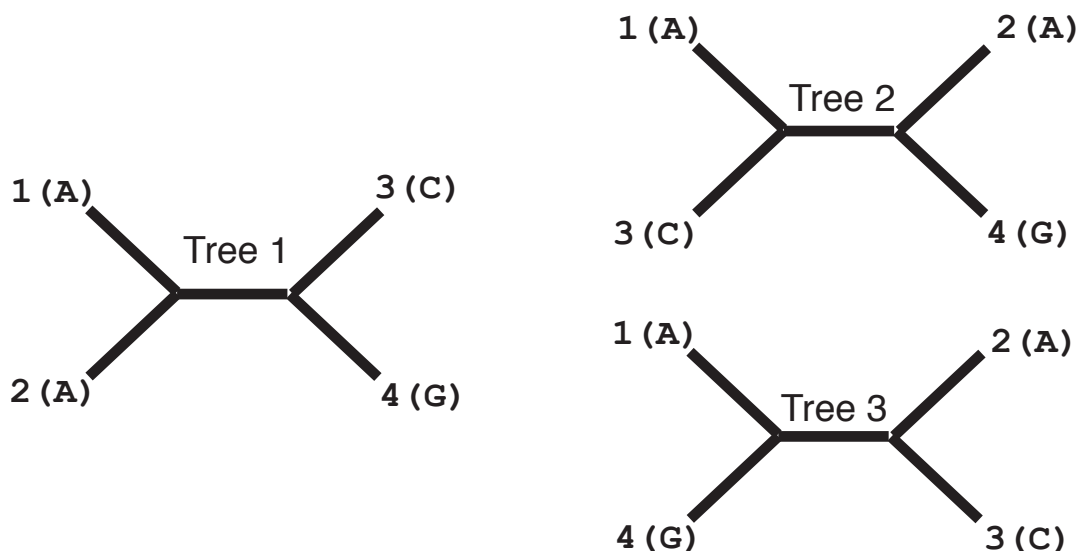


Computing the likelihood of one column from a DNA alignment

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	0.793	0.078	0.104	0.026
C	0.052	0.819	0.104	0.026
G	0.052	0.078	0.844	0.026
T	0.052	0.078	0.104	0.767

The four nucleotides have the following equilibrium frequencies:

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

You should now select one of the two trees and decide on a pair of ancestral nucleotides, which should be placed at the internal nodes labeled "1" and "2" in your selected tree below.

Now, compute the probability of the data (the likelihood of your tree) and write your results on the white board. We will combine all the results and discuss implications when all are done.

