

1. **Reversible jump MCMC.**— The following three models are nested: F81, HKY85, and TN93. In each case, we have  $q_{ij} = \mu r_{ij} \pi_j$  for  $i \neq j$ . (We let  $\mu = \left(\sum_i \sum_{j \neq i} \pi_i \pi_j r_{ij}\right)^{-1}$  so that edge lengths are in terms of the expected number of substitutions per site.) Consider a parameterization where  $r_{ij} = 1$  for all transversions in each model. For the F81 model, the transition rates are also one. For the HKY85 model, the transition rates are each  $\kappa$ . For the TN93 model, the two transition rates each have their own parameters,  $\kappa_R$  for the rate between A and G and  $\kappa_Y$  for the rate between C and T. (This differs from the reversible jump MCMC paper I gave you where we assume that the mean  $r_{ij}$  is one for each model.)

Design a reversible jump MCMC algorithm that will move between these three models for a two-taxon tree. Assume that all three models are equally likely a priori and assume an exponential prior on the edge length  $t$  with rate  $\lambda$ . It suffices to choose an update procedure and to calculate the correct acceptance probabilities.

For those with programming experience, it would be better to implement the method as well to find a 95% credible region for the edge length. Use the cox1 data from pig and cow. Estimate  $\pi$  using the observed overall base proportions.

		Cow			
		a	c	g	t
Pig	a	297	38	18	6
	c	32	252	2	42
	g	14	3	131	2
	t	13	51	2	237

2. **Ancestral distributions.**— Consider a four-taxon unrooted tree where all of the edge lengths are one. In the tree topology, taxa  $A$  and  $B$  are each connected to interior node  $X$  and taxa  $C$  and  $D$  are each connected to interior node  $Y$ . The single internal edge connects  $X$  and  $Y$ . Assume the Jukes-Cantor model of nucleotide substitution. At a site where all tips (taxa  $A$ ,  $B$ ,  $C$ , and  $D$ ) are each nucleotide G, what is the joint distribution of the two interior nodes  $X$  and  $Y$ ? At a site where the base at taxa  $A$  and  $B$  is a G and the base at  $C$  and  $D$  is a T, what is the joint distribution of the bases at  $X$  and  $Y$ ?
3. **Sorting by reversal.**— The non-tRNA mitochondrial genome arrangement for the black chiton relative to human is as this.

$$1, 2, 3, -8, -7, -6, -10, 9, -13, -12, -11, 4, 5, 14$$

Draw the breakpoint graph and count the number of breakpoints, cycles, and hurdles. What is the smallest number of reversals to sort this permutation?