

Table 1:

|              |   |
|--------------|---|
| Dolphin      | ATG ACC AAC ATC CGA AAA ACA CAC CCT CTA |
| Hippopotamus | ATG ACA AAC ATC CGA AAA TCT CAC CCC TTA |
| Camel        | ATG ACA AAC ATC CGA AAA TCA CAC CCA CTT |
| Cow          | ATG ACT AAC ATT CGA AAG TCC CAC CCA CTA |
| Giraffe      | ATG ATC AAC ATC CGA AAG TCC CAC CCA CTA |
|              |   |
| Sperm whale  | ATG ACC AAC ATC CGA AAA TCA CAC CCA TTA |
| Blue whale   | ATG ACC AAC ATC CGA AAA ACA CAC CCA CTA |
| Pig          | ATG ACC AAC ATC CGA AAA TCA CAC CCA CTA |
| Sheep        | ATG ATC AAC ATC CGA AAA ACC CAC CCA CTA |
| Goat         | ATG ACC AAC ATC CGA AAG ACC CAC CCA TTA |

**Problems.**

1. Table 1 shows the first 30 bp of the mitochondrial cytochrome *b* gene for ten mammals. (The complete gene has 1140 bp for these species.) Find the parsimony score for each of the fifteen possible unrooted bifurcating trees with the first five taxa in the table.
2. A site is said to be *phylogenetically informative* (for parsimony purposes, in any case) if the parsimony score of that site is variable over the set of possible trees. Prove that a site is phylogenetically informative if and only if the site pattern contains at least two different symbols that each appear at least twice.
3. Use the Fitch algorithm to find the parsimony score for each of the phylogenetically informative sites using all the data in Table 1 for the following tree.

((((Dolphin,S. whale),B. whale),((Hippo,Pig),(((Cow,(Sheep,Goat)),Giraffe),Camel))));

4. Assume that transitions ( $A \leftrightarrow G$  and  $C \leftrightarrow T$ ) have a cost of 1 and that transversions (any other change) have a cost of 2. Use Sankoff's algorithm to find the total cost of the tree from the previous problem for each of the phylogenetically informative sites in Table 1.
5. Pick one of the possible unrooted trees for the ten taxa in Table 1 uniformly at random. What is the probability that the tree contains Dolphin, Sperm, Whale, and Blue Whale as a *clade*? (In other words, what is the probability that the tree has an edge that *splits* these three taxa from the other seven?)