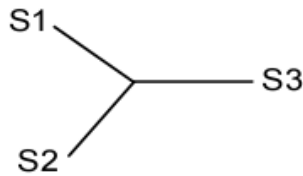Allman Ch 5.1 1,2,3,5,7 and Ch 5 2 1,2,5,13 and Ch5.4 1,2,3,7,9

Ch 5.1
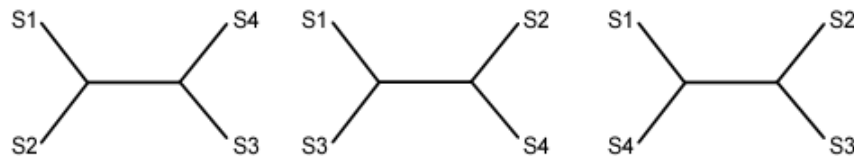
5.1.1. a. $\{T_2, T_3\}$
      b. $\{T_2, T_3, T_5\}$
      c. $\{T_1, T_6\}$, $\{T_2, T_3, T_5\}$
      d. $\{T_1, T_2, T_3, T_4, T_5, T_6\}$
      e. T_4, T_6

5.1.2. a.

      b. In the tree in part (a), the root can be placed along the edge joining the internal node to S1, S2, or S3.

5.1.3. a.

      b. In each of the three trees in part (a), the root can be located on any of the five edges.
Equivalently, for the tree below on the left there are three distinct labelings (from top to bottom) of the leaves: {S1, S2, S3, S4}, {S1, S3, S2, S4}, {S1, S4, S2, S3}, and for the tree on the right there are twelve distinct labelings: {S1, S2, S3, S4}, {S1, S2, S4, S3}, {S1, S3, S2, S4}, {S1, S3, S4, S2}, {S1, S4, S2, S3}, {S1, S4, S3, S2}, {S2, S3, S1, S4}, {S2, S3, S4, S1}, {S2, S4, S1, S3}, {S2, S4, S3, S1}, {S3, S4, S1, S2}, {S3, S4, S2, S1}.

5.1.5.

n	2	3	4	5	6	7	8	9	10
$\frac{(2n-3)!}{2^{(n-2)}(n-2)!}$	1	3	15	105	945	10395	135135	2027025	34459425

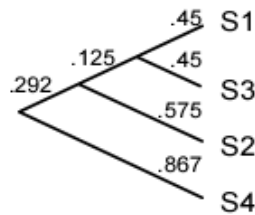**5.1.7.** The most accurate estimate, produced by writing a brief computer program to find the product, is 4.89×10^{296}.
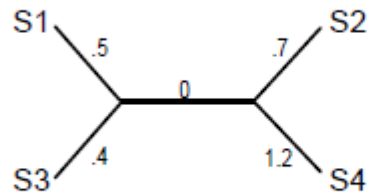
Ch 5.2

5.2.1.

	S1	S2	S3	S4
S1		.425	.27	.55
S2			.425	.55
S3				.55

While the distance between the first two taxa to be joined, $d(S1, S3)$, agrees exactly with the original distance table, the other distances are only close to the original distances. The duplication of some table entries reflects the molecular clock hypothesis, since certain subsets of taxa will be equidistant from a common ancestor.
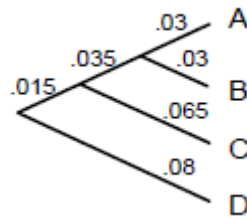
5.2.2.

5.2.5.

Topologically, the trees are the same as unrooted trees. They are not the same metrically. Note, for instance, FM assigns a branch length of 0 to the internal edge, while UPGMA assigns .125.

5.2.13. a. There is no way that all four taxa can be equidistant from a root: Since A and C are not equidistant from the internal node to which they are both joined, if a molecular clock is assumed, the root would have to be on the edge leading to taxon C. If this were the case, then it would be impossible for B and D to be equidistant from the root.

b.

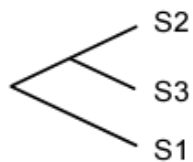|	A	B	C	D
A		.06	.12	.14
B			.14	.12
C				.22

Note: If the labels C and D are exchanged in this tree, then the resulting tree would also be a UPGMA tree, since there were two minima in the collapsed data table.

c. Notice that the UPGMA tree constructed has the wrong topological structure. Since A and B are closest in distance, they are joined first by UPGMA, even though this results in the wrong topology. Neighbor joining, introduced in the next section, will not make this mistake.

d. FM creates the same topological tree as UPGMA, so it too will construct an incorrect tree topology.

Ch 5.4

5.4.1. a. Both trees have parsimony score 3.

b. The most parsimonious trees have score 2.

c. Since there are only four bases, we can always find a tree that requires three substitutions. For example, if we create one joined cluster of taxa with A's, another cluster with G's, another with C's, and a last with T's, then join up these clusters, the resulting tree will have parsimony score 3.

5.4.2. a. The tree on the left has parsimony score 7; the tree on the right has parsimony score 8.

b. The third unrooted tree has parsimony score 10. Therefore, the tree pictured on the left is the most parsimonious unrooted tree.

5.4.3. a. Sites 3, 6, 8, and 11

b. S1 and S4 are neighbors and S2 and S3 are neighbors. The parsimony score for the rooted tree relating them is 5.

c.

5.4.9. a. Informative patterns for four taxa contain two bases, each occurring twice. There are 3 ways to make such patterns without regard to base choices.

b. 25