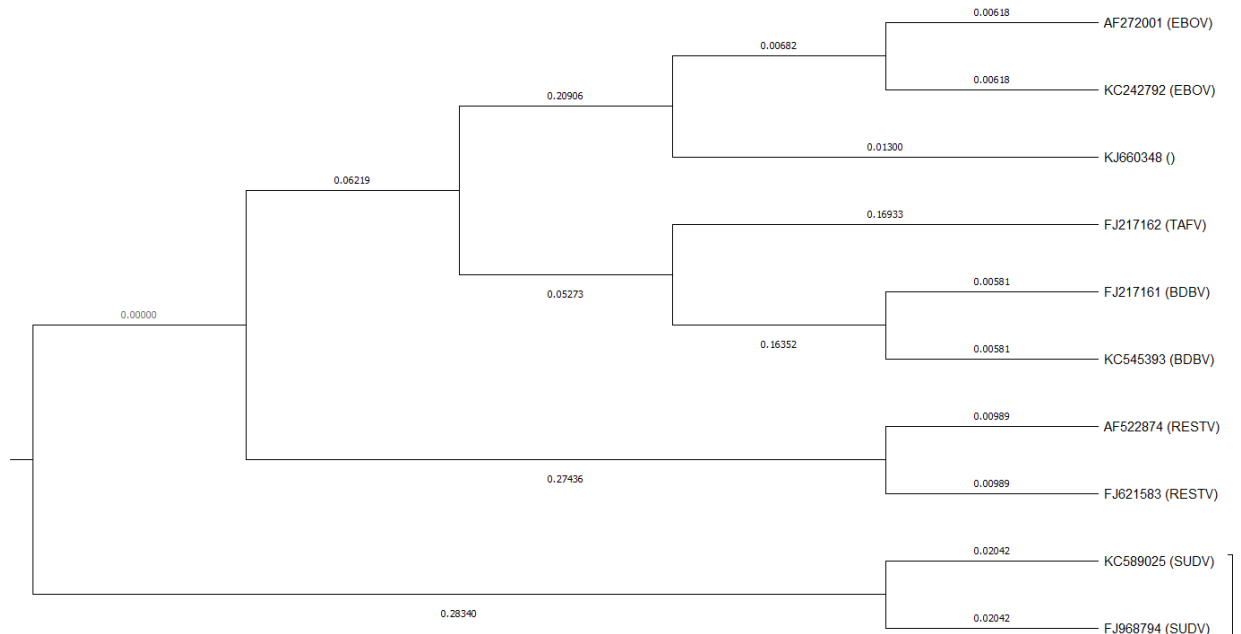**Software Challenge: Constructing Evolutionary Trees**

1. I would expect the species from Guinea be most closely related to TAFV because Tai Forest is closest in physical proximity to Guinea. Therefore, it makes sense to think that the 2 species are most closely related because there is more intermingling with people from the 2 areas than with Guinea and any other area on the list, so it's likely that one evolved directly from the other.

2. The tree is:
   a. 0->4:2.250
   b. 1->5:7.250
   c. 2->4:-0.250
   d. 3->5:7.750
   e. 4->5:3.750
   f. 4->0:2.250
   g. 4->2:-0.250
   h. 5->4:3.750
   i. 5->1:7.250
   j. 5->3:7.750
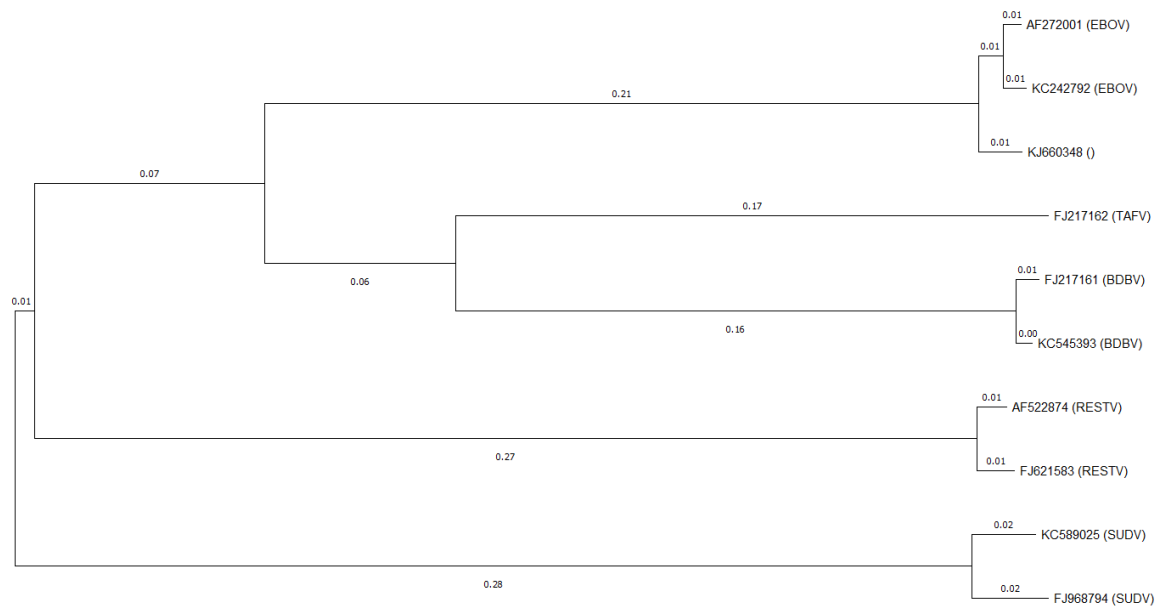
3.
4. The nearest species is EBOV because the 2 strains are by far the closest to the 2014 strain. This is not the same species as in question 1, which was TAFV.

0.00618 — AF272001 (EBOV)

0.00682

0.00618 — KC242792 (EBOV)

0.20906

0.01300 — KJ660348 ()

0.06219

0.16933 — FJ217162 (TAFV)

0.00581 — FJ217161 (BDBV)

0.05273

0.00000

0.16352

0.00581 — KC545393 (BDBV)

0.00989 — AF522874 (RESTV)

0.27436

0.00989 — FJ621583 (RESTV)

0.02042 — KC589025 (SUDV)

0.28340

0.02042 — FJ968794 (SUDV)

5.
6. The root of the tree separates the SUDV viruses' subtree from the subtree of all of the other species.
7. The distance of produced is about 0.30.
8. The distance is 0.28 and to the internal node is 0.27 (approximately).
9. It took approximately 0.02 units. This is simply the sum of the branch lengths of the 2014 virus and common ancestor for other ebola viruses from the internal node (i.e. first common ancestor for other ebola viruses).

10. It was 28.02 years ago.
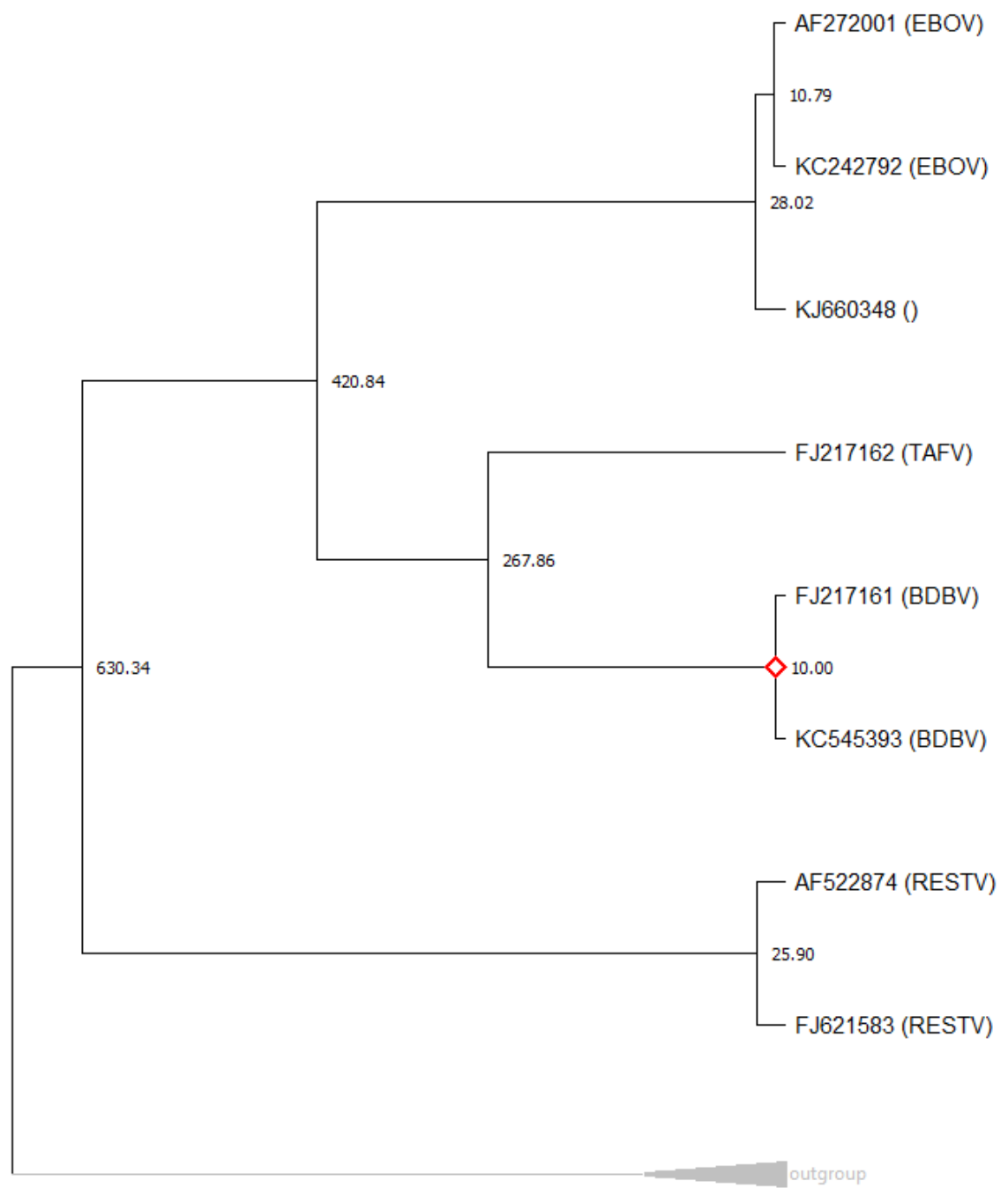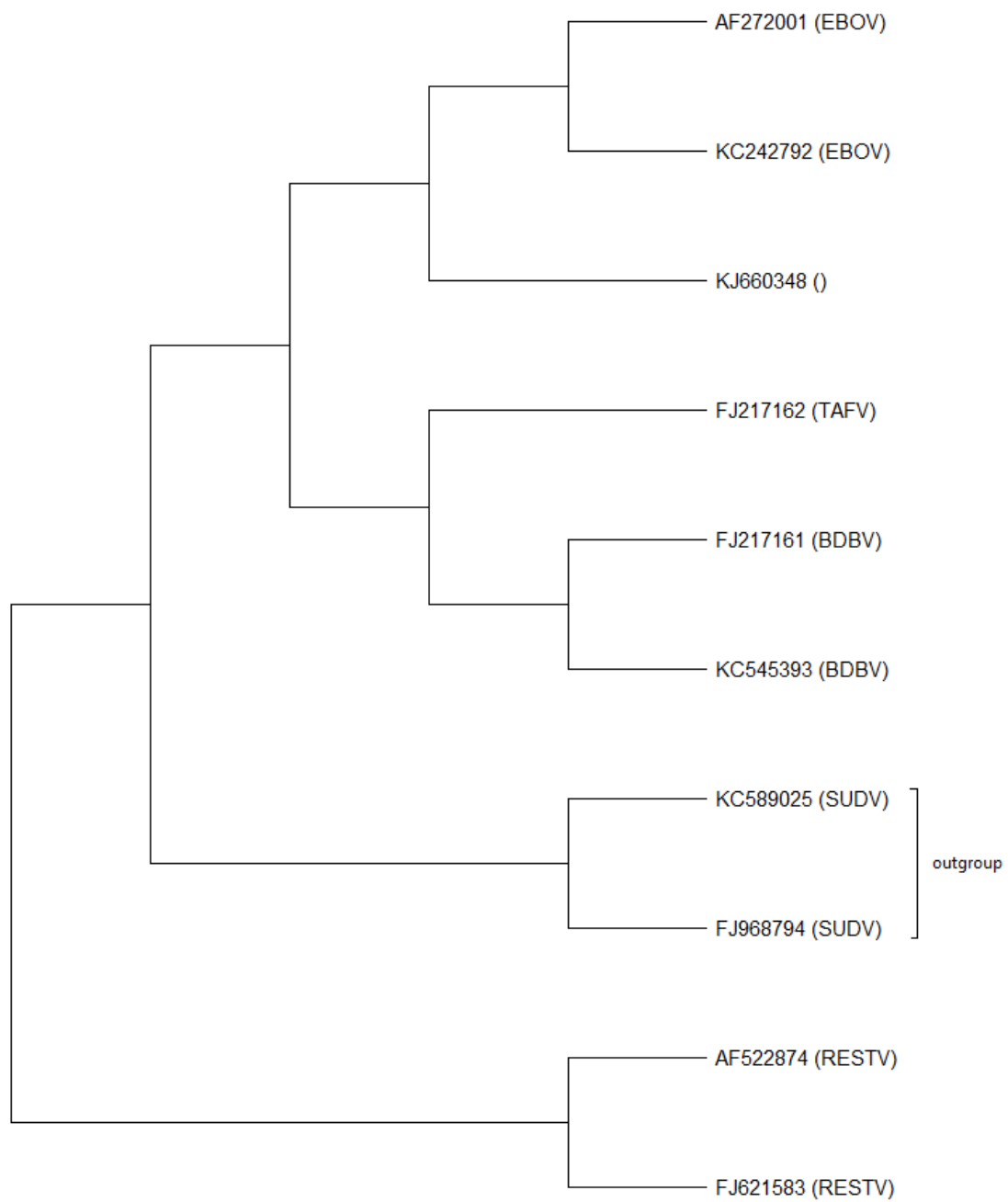
11.
12. It is unrooted.
13. The nucleotides represent the inferred nucleotide at each ancestral node at site 1.
14. There are 409 changes.

15. Both bats connect parts of central Africa with the West African coast. However, the little collard fruit bat has a range that connects Zaire and Guinea more directly.