Pset4

1a.

i)

Log likelihood for topology 1 is -2746.80

Log likelihood for topology 2 is -2923.72

Log likelihood for topology 3 is -2923.72

# Topology 1 is the maximum likelihood tree

ii)

Amplify branch length for branch 1-4 by 1000\*

thus, branch length will be:

branch\_len = matrix(c(0.07517\*1000, 0.03059\*1000, 0.03161\*1000, 0.11761\*1000, 0.14289,

0.20843\*1000, 0.03397\*1000, 0.03497\*1000, 0.24952\*1000, 0.00000,

0.20843\*1000, 0.03397\*1000, 0.03497\*1000, 0.24952\*1000, 0.00000),ncol = 5, nrow = 3,byrow = TRUE)

Outcomes of the modified branch length:

Log likelihood for topology 1 is -5140.38

Log likelihood for topology 2 is -5140.38

Log likelihood for topology 3 is -5140.38

The log likelihood of the three trees are exactly the same, making it less likely to distinguish the best tree.

As the branch 5 getting shorter, the difference between the trees are smaller and it is harder to find the maximum likelihood tree.

iii)

Or we can test in some other genes/longer coding region except for APOE. The optimal tree topology should be consistent with the outcomes of APOE.

For the fixed species, we can also test with Sankoff’s algrothm by minimizing the base transitions in each node and get the best tree topology.

1b. output:

root: ATGAAGGTTCTGTGGGCTGCGCTGCTGGTCACATTCCTGGCAGGATGCCAGGCCGAGGTGGAGCTGGAGA

human: ATGAAGGTTCTGTGGGCTGCGTTGCTGGTCACATTCCTGGCAGGATGCCAGGCCAAGGTGGAGCTGGAGA

mouse: ATGAAGGCTCTGTGGGCCGTGCTGTTGGTCACATTGCTGACAGGATGCCTAGCCGAGGGAGAGCAGGTGA

rat: ATGAAGGCTCTGTGGGCCCTGCTGTTGGTCCCATTGCTGACAGGATGCCTGGCCGAGGGAGAGCAGGTGA

dog: ATGAAGGTTCTGTGGGCTGCGCTGGTGGTCACGCTCCTGGCAGGATGCTGGGCCGATGTGCAGCCGGAGC

root: CAGAGCAGACCGAGTGGCAGAGCGGCCAGCCCTGGGAGCTGGCACTGGGCCGCTTCTGGGATTACCTGCG

human: CAGAGCAGACCGAGTGGCAGAGCGGCCAGCGCTGGGAACTGGCACTGGGTCGCTTTTGGGATTACCTGCG

mouse: CAGATCAGCTCGAGTGGCAAAGCAACCAACCCTGGGAGCAGGCCCTGAACCGCTTCTGGGATTACCTGCG

rat: CAGATCAGCTCCCAGGGCAAAGCGACCAACCCTGGGAGCAGGCCCTGAACCGCTTCTGGGATTACCTGCG

dog: CGGAGCAGGCCGGGTGGCAGACTGGCCAGCCCTGGGAGGCGGCGCTGGCCCGCTTCTGGGATTACCTGCG

root: CTGGGTGCAGACGCTGTCTGACCAGGTGCAGGAGGAGCTGCTCAGCTCCCAGGTCACCCAGGAACTGACG

human: CTGGGTGCAGACACTGTCTGAGCAGGTGCAGGAGGAGCTGCTCAGCTCCCAGGTCACCCAGGAACTGAGG

mouse: CTGGGTGCAGACGCTGTCTGACCAGGTCCAGGAAGAGCTGCAGAGCTCCCAAGTCACACAAGAACTGACG

rat: CTGGGTGCAGACGCTTTCTGACCAGGTCCAGGAAGAGCTGCAGAGCTCCCAAGTCACACAGGAACTGACG

dog: CTGGGTGCAGACGCTGTCTGACCAGGTGCAAGAGGGCGTGCTCAACACCCAGGTCACCCAGGAACTGACG

root: GCGCTGATGGACGAGACCATGAAGGAGGTGAAGGCCTACAAATCGGAGCTGGAGGAACAGCTGGGCCCGG

human: GCGCTGATGGACGAGACCATGAAGGAGTTGAAGGCCTACAAATCGGAACTGGAGGAACAACTGACCCCGG

mouse: GCACTGATGGAGGACACTATGACGGAAGTAAAGGCTTACAAAAAGGAGCTGGAGGAACAGCTGGGTCCAG

rat: GTACTGATGGAGGACACTATGACGGAAGTAAAGGCATACAAAAAGGAGCTGGAGGAACAGCTGGGCCCAG

dog: GCGCTGATGGATGAGACCATGAAGGAGGTGAAGGCCTACAAGGCGGAGCTGGACGAGCAGCTGGGCCCCA

root: TGGCGGAGGAGACGCGGGCCCGGCTGGCCAAGGAGCTGCAGGCGGCGCAGGCCCGGCTGGGCGCGGACAT

human: TGGCGGAGGAGACGCGGGCACGGCTGTCCAAGGAGCTGCAGGCGGCGCAGGCCCGGCTGGGCGCGGACAT

mouse: TGGCGGAGGAGACACGGGCCAGGCTGGGCAAAGAGGTGCAGGCGGCACAGGCCCGACTCGGAGCCGACAT

rat: TGGCGGAGGAGACACGGGCCAGGCTGGCTAAAGAGGTGCAGGCGGCACAGGCCCGTCTGGGAGCTGACAT

dog: TGACCTCGGAGACGCAGGCCCGCGTGGCCAAGGAGCTGCAGGCGGCGCAGGCCCGGCTGCGCTCGGACAT

root: GGAGGACGTGCGCAACCGCCTGGTGCAGTACCGCGGCGAGGTGCAGGCCATGCTGGGCCAGAGCACCGAG

human: GGAGGACGTGTGCGGCCGCCTGGTGCAGTACCGCGGCGAGGTGCAGGCCATGCTCGGCCAGAGCACCGAG

mouse: GGAGGATCTACGCAACCGACTCGGGCAGTACCGCAACGAGGTGCACACCATGCTGGGCCAGAGCACAGAG

rat: GGAGGATCTACGCAACCGACTCGGGCAGTACCGCAACGAGGTAAACACCATGCTGGGCCAGAGCACAGAG

dog: GGAGGACGTGCGCAACCGCCTGACGCAGTACCGCGGCGAGCTGCAGGCCATGCTGGGCCAGAGCAGCGAG

root: GAGCTGCGGGCGCGCCTCGCCTCCCACCTGCGCAAGCTGCGCAAGCGGCTGCTGCGGGATGCCGATGACC

human: GAGCTGCGGGTGCGCCTCGCCTCCCACCTGCGCAAGCTGCGTAAGCGGCTCCTCCGCGATGCCGATGACC

mouse: GAGATACGGGCGCGGCTCTCCACACACCTGCGCAAGATGCGCAAGCGCTTGATGCGGGATGCCGAGGATC

rat: GAGCTGCGGTCGCGCCTCTCCACACACCTGCGCAAGATGCGCAAGCGCCTGATGCGGGATGCGGATGATC

dog: GAGCTGCGGGCGCGCTTCGCCTCCCACATGCGCAAGCTGCGCAAGCGGGTGCTGCGGGACGCCGAGGACC

root: TGCAGAAGCGCCTGGCAGTGTACAAGGCCGGGGCCCGCGAGGGCGCCGAGCGCGGCGTGAGCGCCATCCG

human: TGCAGAAGCGCCTGGCAGTGTACCAGGCCGGGGCCCGCGAGGGCGCCGAGCGCGGCCTCAGCGCCATCCG

mouse: TGCAGAAGCGCCTAGCTGTGTACAAGGCAGGGGCACGCGAGGGCGCCGAGCGCGGTGTGAGTGCCATCCG

rat: TGCAGAAGCGCCTGGCGGTGTACAAGGCCGGGGCACAGGAGGGCGCCGAGCGCGGTGTGAGTGCTATCCG

dog: TGCAGAGGCGCCTGGCCGTCTACAAGGCCGGCGTGCGCGAGGGTGCCGAGCGCAGCGTGAGCAGCATCCG

root: CGAGCGCCTGGGGCCCCTGGTGGAGCAGGGCCGCGAGCGCACCGCCAATGTGGGCGCCCTGGCCGCCCAG

human: CGAGCGCCTGGGGCCCCTGGTGGAACAGGGCCGCGTGCGGGCCGCCACTGTGGGCTCCCTGGCCGGCCAG

mouse: TGAGCGCCTGGGGCCTCTGGTGGAGCAAGGTCGCCAGCGCACTGCCAACCTAGGCGCTGGGGCCGCCCAG

rat: TGAGCGCCTGGGGCCACTGGTGGAGCAGGGTCGTCAGCGCACAGCCAACCTAGGCGCTGGCGCCGCCCAG

dog: CGAGCGCCTCTGGCCGCTGCTGGAGCAGGCCCGCGAGCGCAACGCCAAGGTGGGCGCCCTGGCCACGCAG

root: CCGCTGCACGAGCGGGCCCAGGCCTTGGGCGAGCGGCTGCGCGGGCGGCTGGAGGAGATGGGCAGCCGGG

human: CCGCTACAGGAGCGGGCCCAGGCCTGGGGCGAGCGGCTGCGCGCGCGGATGGAGGAGATGGGCAGCCGGA

mouse: CCTCTGCGCGATCGCGCCCAGGCTTTTGGTGACCGCATCCGAGGGCGGCTGGAGGAAGTGGGCAACCAGG

rat: CCCCTGCGCGATCGCGCCCAGGCTTTGAGTGACCGCATCCGAGGGCGGCTGGAGGAAGTGGGCAACCAGG

dog: CCGCTGCTCGAGCGGGCCGACGCCCTGGGCCAGCAGCTGCGCGGGCAGCTGGAGGAGATGAGCAGCCGGG

root: CCCGCGACCGCCTGGAGGAGGTGCGGGAGCAGATGGAGGAGGTGCGCGCCAAGATGGAGGAGCAGGCCCA

human: CCCGCGACCGCCTGGACGAGGTGAAGGAGCAGGTGGCGGAGGTGCGCGCCAAGCTGGAGGAGCAGGCCCA

mouse: CCCGTGACCGCCTAGAGGAGGTGCGTGAGCACATGGAGGAGGTGCGCTCCAAGATGGAGGAACAGACCCA

rat: CCCGAGACCGCCTAGAGGAGGTGCGTGAGCAGATGGAGGAGGTGCGCTCCAAGATGGAGGAGCAGACCCA

dog: CCCGCGGCCACCTGGAGGAGATGCGCGAGCAGATACAGGAGGTGCGGGTGAAGATGGAGGAGCAGGCCGA

root: GCAGATACGCCTGCAGGCCGAGGCCTTCCAGGCCCGCCTCAAGAGCTGGTTCGAGCCCCTGGTGGAAGAC

human: GCAGATACGCCTGCAGGCCGAGGCCTTCCAGGCCCGCCTCAAGAGCTGGTTCGAGCCCCTGGTGGAAGAC

mouse: GCAAATACGCCTGCAGGCGGAGATCTTCCAGGCCCGCCTCAAGGGCTGGTTCGAGCCAATAGTGGAAGAC

rat: GCAGATACGCCTGCAGGCCGAGATCTTCCAGGCCCGCATCAAGGGCTGGTTCGAGCCGCTAGTGGAAGAC

dog: CCAGATACGCCAAAAGGCCGAGGCCTTCCAGGCGCGCCTCAAGAGCTGGTTCGAGCCCCTGCTGGAAGAC

root: ATGCAGCGCCAGTGGGCCGGGCTGGTGGAGAAGGTGCAGGCTGCCGTGGCCACCAGCCCCACCCCTGTGC

human: ATGCAGCGCCAGTGGGCCGGGCTGGTGGAGAAGGTGCAGGCTGCCGTGGGCACCAGCGCCGCCCCTGTGC

mouse: ATGCATCGCCAGTGGGCAAACCTGATGGAGAAGATACAGGCCTCTGTGGCTACCAACCCCATCCCAGTGG

rat: ATGCAGCGCCAGTGGGCAAACCTAATGGAGAAGATACAGGCCTCTGTGGCTACCAACTCCATCACAGTGC

dog: ATGCAGCGCCAGTGGGACGGGCTGGTGGAGAAGGTGCAGGCGGCCGTGGCCACCATCCCCACCTCTAAGC

root: CCATGGAGAATCAATGA

human: CCAGCGACAATCACTGA

mouse: CCCAGGAGAATCAATGA

rat: CCCTGGAGAATCAATGA

dog: CTGTGGAGGAACCATGA

1c.

Note: I highlighted the most of the difference by hand…

root: MKVLWAALLVTFLAGCQAEVELETEQTEWQSGQPWELALGRFWDYLRWVQTLSDQVQEELLSSQVTQELT

human: MKVLWAALLVTFLAGCQAKVELETEQTEWQSGQRWELALGRFWDYLRWVQTLSEQVQEELLSSQVTQELR

root: ALMDETMKEVKAYKSELEEQLGPVAEETRARLAKELQAAQARLGADMEDVRNRLVQYRGEVQAMLGQSTE

human: ALMDETMKELKAYKSELEEQLTPVAEETRARLSKELQAAQARLGADMEDVCGRLVQYRGEVQAMLGQSTE

root: ELRARLASHLRKLRKRLLRDADDLQKRLAVYKAGAREGAERGVSAIRERLGPLVEQGRERTANVGALAAQ

human: ELRVRLASHLRKLRKRLLRDADDLQKRLAVYQAGAREGAERGLSAIRERLGPLVEQGRVRAATVGSLAGQ

root: PLHERAQALGERLRGRLEEMGSRARDRLEEVREQMEEVRAKMEEQAQQIRLQAEAFQARLKSWFEPLVED

human: PLQERAQAWGERLRARMEEMGSRTRDRLDEVKEQVAEVRAKLEEQAQQIRLQAEAFQARLKSWFEPLVED

root: MQRQWAGLVEKVQAAVATSPTPVPMENQ$

human: MQRQWAGLVEKVQAAVGTSAAPVPSDNH$

2a.

[1] "(Bonobo:0.00392,(((((((Rhesus:0.004991,Crab\_eating\_macaque:0.004991):0.003,Baboon:0.008042):0.01961,Green\_monkey:0.027):0.02204,Gibbon:0.02227):0.003471,Orangutan:0.01894):0.009693,Gorilla:0.008964):0.000999999999999997,(Human:0.00655,Chimp:0.00684):0.00122):0.00392)"

2b.

Use UCSC phyloPng tool:

Parameters:

Width:300

Height:512

Decimal places:5

