

HW 4: Bayes– Concatenated and Coalescent Analyses

Daisy Fry-Brumit

2022-10-18

Part One: Determine MrBayes Settings Using a Single Gene

using Gonnnet

1: In your MrBayes run, what settings did you use for `nst=`, `rates=`, and `statefreqpr=`, in order to run the K80 +I model? (3 pts)

Model settings for MrBayes run

Parameter	Value
nst	2
rates	propinv
statefreqpr	fixed(equal)

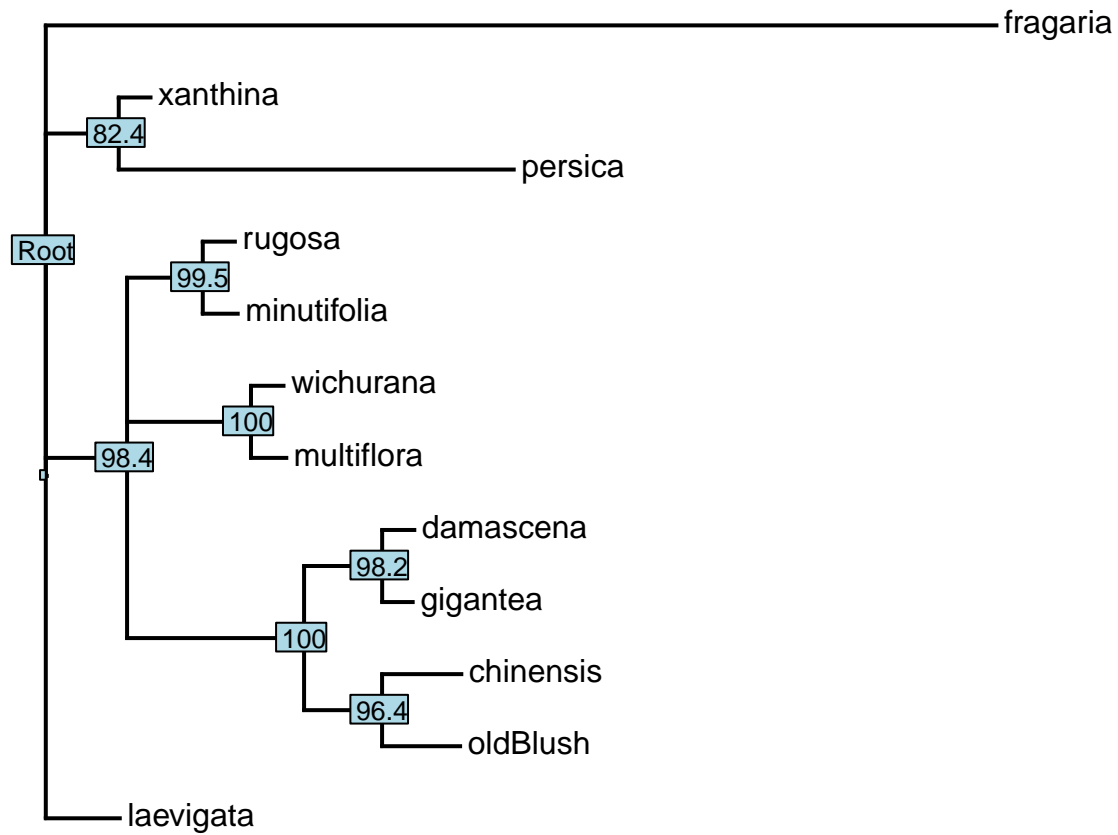
Remaining in step with our class practice, I ran MCMC with 500,000 iterations and showed results at every 10 iterations. (Average standard deviation of split frequencies: 0.002287)

Question 2: Do the values in your .pstat file suggest that the MCMC achieved convergence? Why or why not? Please list at least 2 pieces of evidence that suggest your run converged. If you want, you may use values from other files instead of just .pstat. (3 pts)

.pstat values

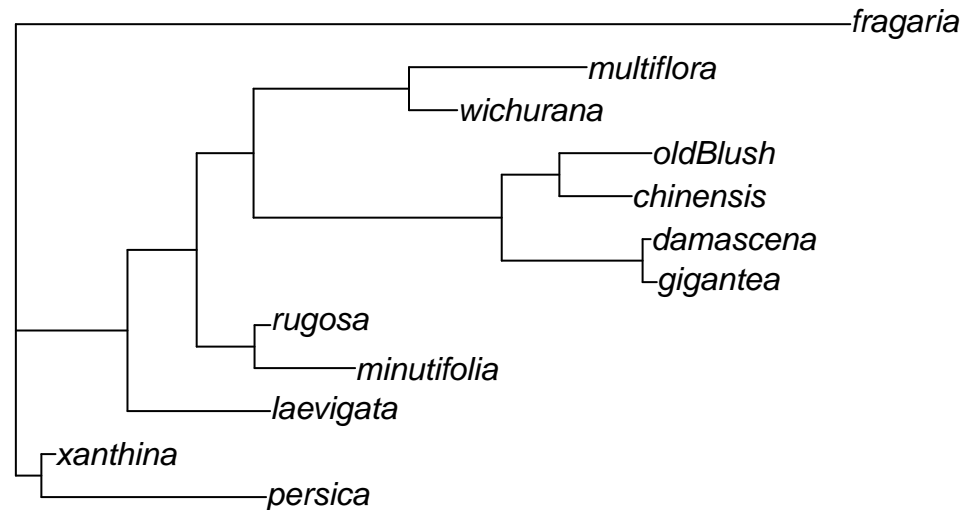
avgESS values are consistently over 100 and PSRF values all round to 1.00. Both of these outcomes suggest convergence!

Question 3: Provide a plot of your rooted consensus tree from MrBayes, with posterior probability values at each node. Make sure your plot is clearly labeled and readable. (2 pts)



Question 4: Provide a plot of your rooted Maximum Clade Credibility tree (support values optional). (1 pt)

Maximum Clade Credibility Tree



Part Two: Concatenated and Partitioned Analysis in MrBayes

using Gonnet

Question 5: Once you have set up your model, please paste in the final block of code/settings from your mb_input file, including the partitions as well as the parameter settings. (5 pts)

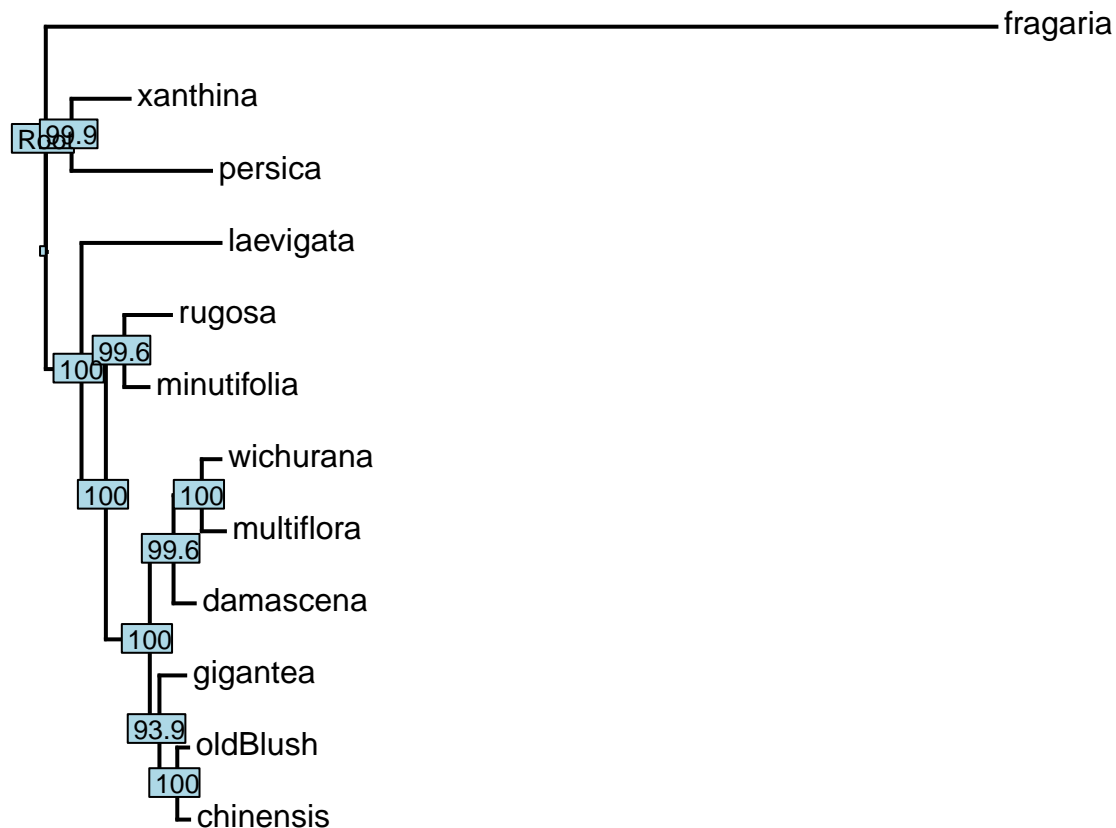
```
begin mrbayes;
  set autoclose=yes nowarn=yes;
  execute RosesConcat.nex;
  charset loc0114=1-314;
  charset loc0180=315-663;
  charset loc0271=664-1135;
  charset loc0273=1136-1517;
  charset loc0535=1518-1817;
  charset loc0692=1818-2122;
  charset loc0716=2123-2526;
  charset loc0751=2527-2858;
  charset loc0778=2859-3237;
  charset loc0895=3238-3591;
```

```

charset loc1018=3592-3907;
charset loc1571=3908-4317;
charset loc1723=4318-4658;
charset loc1762=4659-5108;
charset loc1851=5109-5502;
partition cpGenes = 15:loc0114,loc0180,loc0271,loc0273,loc0535,loc0692,loc0716,loc0751,loc0778,loc0800;
lset nst=2 rates=propinv;
unlink revmat=(all) pinvar=(all) statefreq=(all) shape=(all);
prset applyto=(all) ratepr=variable;
Outgroup fragaria;
mcmc ngen=1000000 samplefreq=500 file=roses;
sump relburnin=yes burninfrac=0.25;
sumt relburnin=yes burninfrac=0.25 conformat=simple;
end;

```

Question 6: Provide a plot of your consensus tree from the concatenated run of MrBayes with posterior probability values at each node. (2 pts)



Question 7: Use treedist() to compare the maximum clade credibility (MCC) tree from your concatenated run with the MCC from just the first locus (0114). How different are the two trees? (1 pt)

##	symmetric.difference	branch.score.difference	path.difference
##	4.000000	4.170159	6.633250

```
## quadratic.path.difference
##                               16.700895
```

The symmetric difference between the two trees is 4, so our 2 trees differ in 2 nodes, which is smaller than I anticipated considering how many loci were introduced in our concatenated run.

Part Three: Run a Coalescent Analysis with ASTRAL

```
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
## using Gonnnet
```

Question 8: Provide the plot of your ASTRAL tree with posterior probability values at each node. Make sure the tree is clearly labeled and readable. (5 pts)

Question 9: Compare your ASTRAL tree with your MCC tree from Part Two (from the concatenated run). Disregarding the branch length scores (since we manipulated the branch lengths) are the two trees identical in terms of topology? What does this result mean in terms of whether or not there was a lot of discordance among the gene trees? (3 pts)