

## Lab 7 R Source Code

### Code:

```
1 #
2 # Shawn Schwartz , 2019
3 # EEB C174 UCLA Spring 2019
4 # Lab 7 HW – BAMM
5 #
6
7 rm(list=ls())
8
9 #imports
10 library(BAMMtools)
11 library(coda)
12
13 cwd <- "~/Developer/EEB-C174-Labs/Lab7/"
14 setwd(cwd)
15
16 resources_path <- "resources/"
17 output_path <- "output/"
18
19 tree <- read.tree(paste0(resources_path, "whaleTree.tre"))
20 plot(tree, cex = 0.7)
21 axisPhylo()
22
23 ##### Verify BAMM assumptions #####
24
25 # Tree Must Be Ultrametric
26 is.ultrametric(tree)
27
28 # Tree Must Be Binary (all of the nodes need to lead to 2 branches
29   , and no more than 2)
30 is.binary.tree(tree)
31
32 # Check to make sure all branch lengths are greater than 0
33 min(tree$edge.length)
34
35 ##### Create BAMM Control File #####
36 #estimate priors and create control files
37 priors <- setBAMMpriors(tree, outfile = NULL)
38 priors
```

```
38
39 generateControlFile(file = paste0(resources_path, "
    whaleBAMMcontrolfile.txt"), params = list(
40   treefile = "whaleTree.tre",
41   globalSamplingFraction = "0.8", #This is 1 for complete sampling
    (for homework, put in a fraction like 0.8)
42   seed = sample(1:5000000, 1),
43   overwrite = "0",
44   expectedNumberOfShifts = "1",
45   lambdaInitPrior = as.numeric(priors["lambdaInitPrior"]),
46   lambdaShiftPrior = as.numeric(priors["lambdaShiftPrior"]),
47   muInitPrior = as.numeric(priors["muInitPrior"]),
48   numberOfGenerations = "5000000",
49   mcmcWriteFreq = "1000",
50   eventDataWriteFreq = "1000",
51   printFreq = "1000",
52   acceptanceResetFreq = "1000",
53   outName = "whaleBAMMhomeworkoutput",
54   ##numberOfChains = "2", #set to the number of CPUs
55   numberOfChains = "4", #set to the number of CPUs
56   deltaT = "0.01"
57 ))
58
59
60 labrid_tree <- read.tree(paste0(resources_path, "Labridae.tre"))
61 ##### Verify BAMM assumptions #####
62
63 # Tree Must Be Ultrametric
64 is.ultrametric(labrid_tree)
65
66 library(phytools)
67 labrid_tree <- force.ultrametric(labrid_tree, method=c("nnls",
    "extend"))
68 is.ultrametric(labrid_tree)
69
70 # Tree Must Be Binary (all of the nodes need to lead to 2 branches
    , and no more than 2)
71 is.binary.tree(labrid_tree)
72
73 # Check to make sure all branch lengths are greater than 0
74 min(labrid_tree$edge.length)
75
```

```
76 ##### Create BMM Control File #####
77 #estimate priors and create control files
78 labrid_priors <- setBMMpriors(labrid_tree, outfile = NULL)
79 labrid_priors
80
81 generateControlFile(file = paste0(resources_path, "
  LABRIDBMMcontrolfile.txt"), params = list(
82   treefile = "Labridae.tre",
83   globalSamplingFraction = "0.8", #This is 1 for complete sampling
      (for homework, put in a fraction like 0.8)
84   seed = sample(1:5000000, 1),
85   overwrite = "0",
86   expectedNumberOfShifts = "1",
87   lambdaInitPrior = as.numeric(priors["lambdaInitPrior"]),
88   lambdaShiftPrior = as.numeric(priors["lambdaShiftPrior"]),
89   muInitPrior = as.numeric(priors["muInitPrior"]),
90   numberOfGenerations = "5000000",
91   mcmcWriteFreq = "1000",
92   eventDataWriteFreq = "1000",
93   printFreq = "1000",
94   acceptanceResetFreq = "1000",
95   outName = "LABRIDBMMhomeworkoutput",
96   ##numberOfChains = "2", #set to the number of CPUs
97   numberOfChains = "4", #set to the number of CPUs
98   deltaT = "0.01"
99 ))
100
101 plot(labrid_tree, cex = 0.3)
102
103 whale_data <- getEventData(tree, eventdata = paste0(resources_path
  , "whaleBMMhomeworkoutput_event_data.txt"), burnin = 0.1)
104 summary(whale_data)
105
106 ##### check quality of BMM results #####
107 mcmcout <- read.csv(paste0(resources_path, "whaleBMMhomeworkoutput
  _mcmc_out.txt"))
108 plot(mcmcout$logLik ~ mcmcout$generation)
109
110 # test for convergence of the MCMC chains #
111 burnstart <- floor(0.1 * nrow(mcmcout))
112 postburn <- mcmcout[burnstart:nrow(mcmcout), ]
113 effectiveSize(postburn$N_shifts)
```

```
114 effectiveSize(postburn$logLik)
115
116 computeBayesFactors(paste0(resources_path, "whaleBAMMhomeworkoutput_
    _mcmc_out.txt"), expectedNumberOfShifts = 1, burnin = 0.1)
117 plotPrior(paste0(resources_path, "whaleBAMMhomeworkoutput_mcmc_out.
    txt"), expectedNumberOfShifts = 1)
118
119 s <- plot.bammdata(whale_data, spex = "s", labels = T, font = 3,
    cex = 0.5)
120 title(main = "Mean speciation rate", sub = "time before present")
121 addBAMMlegend(s, location = "topleft", nTicks = 1)
122 axisPhylo()
123
124 css <- credibleShiftSet(whale_data, expectedNumberOfShifts = 1,
    threshold = 5, set.limit = 0.95)
125 css$number.distinct
126 summary(css)
127 sss <- plot.credibleshiftset(css, border = F)
128
129 plot.new()
130 best <- getBestShiftConfiguration(whale_data,
    expectedNumberOfShifts = 1)
131 par(mfrow = c(1,1))
132 par(mar=c(5, 4.5, 2, 0.5))
133 ss <- plot.bammdata(best, labels = T, font = 3, cex = 0.55)
134 title(main = "Best shift configuration", sub = "time before
    present")
135 addBAMMlegend(ss, location = "topleft", nTicks = 1)
136 addBAMMshifts(best, cex = 3, pch = 1)
137 axisPhylo()
138
139 par(font = 1)
140 marg_probs <- marginalShiftProbsTree(whale_data)
141 plot.phylo(marg_probs, cex = 0.55)
142 title(main = "Marginal shift probability")
143 add.scale.bar(x = 0.5, y = 0.5, font = 1)
144
145
146 ##### clade-specific evolutionary rates #####
147 global_rates <- getCladeRates(whale_data)
148 mean(global_rates$lambda)
149 quantile(global_rates$lambda, c(0.05, 0.95))
```

```
150 rateshift_MRCA <- getMRCA(tree, tip = c("Orcinus_orca_AF084061", "  
    Stenella_longirostris_AF084103"))  
151 clade_rates <- getCladeRates(whale_data, node = rateshift_MRCA)  
152 mean(clade_rates$lambda)  
153 quantile(clade_rates$lambda, c(0.05, 0.95))  
154  
155 smaller_clade_MRCA <- getMRCA(tree, tip = c("Delphinapterus_leucas  
    _DLU72037", "Phocoenoides_dalli_PDU09679"))  
156 smaller_clade_rates <- getCladeRates(whale_data, node = smaller_  
    clade_MRCA)  
157 mean(smaller_clade_rates$lambda)  
158 quantile(smaller_clade_rates$lambda, c(0.05, 0.95))  
159  
160 lowest_clade_rates_MRCA <- getMRCA(tree, tip = c("Caperea_  
    marginata_X75586", "Balaenoptera_acutorostrata"))  
161 lowest_clade_rates_MRCA <- getMRCA(tree, tip = c("Ziphius_  
    cavirostris_AF304075", "Mesoplodon_europaeus_X92537"))  
162 lowest_clade_rates <- getCladeRates(whale_data, node = lowest_  
    clade_rates_MRCA)  
163 mean(lowest_clade_rates$lambda)  
164 quantile(lowest_clade_rates$lambda, c(0.05, 0.95))  
165  
166 ##### Rate-through-time analysis #####  
167 par(font = 1)  
168 plotRateThroughTime(whale_data,  
169                      ratetype = "speciation",  
170                      avgCol = "black",  
171                      intervalCol = "gray80",  
172                      intervals = c(0.05, 0.95),  
173                      opacity = 1)  
174  
175  
176  
177 #two planel plot for lab report  
178 par(mfrow = c(1,2))  
179 par(mar=c(5, 1.0, 2, 0.0))  
180 s <- plot.bammdata(whale_data, spex = "s", labels = T, font = 3,  
    cex = 0.5)  
181 title(main = "Mean speciation rate", sub = "time before present (  
    my)")  
182 addBAMMlegend(s, location = "topleft", nTicks = 1)  
183 axisPhylo()
```

```
184
185 par(font = 1)
186 marg_probs <- marginalShiftProbsTree(whale_data)
187 plot.phylo(marg_probs, cex = 0.55)
188 title(main = "Marginal shift probability")
189 add.scale.bar(x = 1.5, y = 0.0, font = 1)
```