16S Convergence Plots

This document contains plots and statistics for evaluating the convergence of 16S MrBayes runs performed in this paper. Assessment of convergence used the web address http://danwarren.net/plot-comparisons.html written by the authors of the package RWTY.

It seems that both 16S runs have converged. The approximate sampling Estimated Sample Size (ESS) and topology ESS are both above 200 (standard rule of thumb). The tree topology trace shows well-mixed chains and a clear optimum, cumulative split frequencies seem to have become fairly stable, and sliding window split frequencies show large jumps, but a clear search of the tree space. Both runs also seem to have converged on similar posterior probabilities, as shown by the split frequency comparisons. The average standard deviation of split frequencies (ASDSF) is below 0.01 and shows a consistent decrease, as expected. Tree space plots show both runs searching similar areas of tree space and well-mixed chains. The topological autocorrelation plots also appear to have reached a plateau for each run.

Analysis code

```r
library(rwty)

## Loading required package: ape
## Loading required package: ggplot2
# Pull in the trees (only need to indicate a folder)
my.trees <- load.multi("16S", format='mb')

## [1] "16S_con_cds_modified.nex.run1.t"
## [1] "16S_con_cds_modified.nex.run2.t"

# Set burn in
burnin_val <- 2500

# Analysis of trees using **RWTY**
my.trees.rwty <- analyze.rwty(my.trees, burnin=burnin_val, fill.color='LmL')

## [1] "Creating trace for LnL"
## [1] "Creating trace for LnPr"
## [1] "Creating trace for TL"
## [1] "Creating trace for r.A...C."
## [1] "Creating trace for r.A...G."
## [1] "Creating trace for r.A...T."
## [1] "Creating trace for r.C...G."
## [1] "Creating trace for r.C...T."
## [1] "Creating trace for r.G...T."
## [1] "Creating trace for pi.A."
## [1] "Creating trace for pi.C."
```

1
## Creating trace for pi.G.
## Creating trace for pi.T.
## Creating trace for alpha
## Creating trace for m.1.
## Creating trace for tree topologies
## Calculating approximate ESS with sampling intervals from 1 to 100
## Creating topological autocorrelation plot
## Creating sliding window split frequency plot for 20 clades
## Creating sliding window ACSF plot
## Creating cumulative split frequency plot for 20 clades
## Creating cumulative ACSF plot
## Creating treespace plots

## Warning: `panel.margin` is deprecated. Please use `panel.spacing` property instead

## Warning: `panel.margin` is deprecated. Please use `panel.spacing` property instead

## [1] "Creating ASDSF plot"
## [1] "Creating split frequency matrix and ASDSF clustering plots"

# Approximate ESS for topologies

topological.approx.ess(my.trees,burnin=burnin_val)

## [1] "Calculating approximate ESS with sampling intervals from 1 to 100"

## operator approx.ess chain
## 1 = 594.9638 16S_con_cds_modified.nex.run1.t
## 2 = 577.6709 16S_con_cds_modified.nex.run2.t
Plots

Estimated Sample Size

```r
makeplot.pseudo.ess(my.trees, burnin = 2500)
```

## [1] "Creating pseudo ESS plot"
## [1] "Calculating pseudo ESS for 7501 trees and 20 replicates, please be patient"
## [1] "Calculating pseudo ESS for 7501 trees and 20 replicates, please be patient"
## $pseudo.ess.plot
Parameter plot

\[ \text{my.trees.rwty$LnL.trace[[1]]} \]

**LnL trace**

16S_con_cds_modified.nex.run1.t (ESS=1711)

\[ \begin{align*}
\text{LnL} & \quad \text{Generation} \\
-11220 & \quad 4e+06 \quad 6e+06 \quad 8e+06 \quad 1e+07 \\
\end{align*} \]

\[ \text{my.trees.rwty$LnL.trace[[2]]} \]

**LnL trace**

16S_con_cds_modified.nex.run2.t (ESS=1264)

\[ \begin{align*}
\text{LnL} & \quad \text{LnL} \\
-11220 & \quad -11160 \\
\end{align*} \]
Topology trace plots

my.trees.rwty$topology.trace.plot[[1]]

Tree topology trace

16S_con_cds_modified.nex.run1.t (Approximate ESS = 598)

Topological Distance of Tree from Focal Tree

my.trees.rwty$topology.trace.plot[[2]]

Tree topology trace

16S_con_cds_modified.nex.run1.t (Approximate ESS = 598)

Topological Distance of Tree from Focal Tree
Split frequency plots

Cumulative Split Frequencies for 20 clades

Sliding Window Split Frequencies for 20 clades
Split frequency comparisons

con_cds_modified.nex.r

$r = 1.00$

ASDSF = 0.0011

Average Standard Deviation of Split Frequencies

Generation
Tree space plots

my.trees.rwty$treespace.heatmap

Tree space heatmap for 100 trees

my.trees.rwty$treespace.points.plot

Tree space for 100 trees
Autocorrelation plots

my.trees.rwty$autocorr.plot

Topological autocorrelation plot

16S_con_cds_modified.nex.run1.t

Mean Path Difference between Pairs of Trees

16S_con_cds_modified.nex.run2.t

Sampling Interval between Trees