

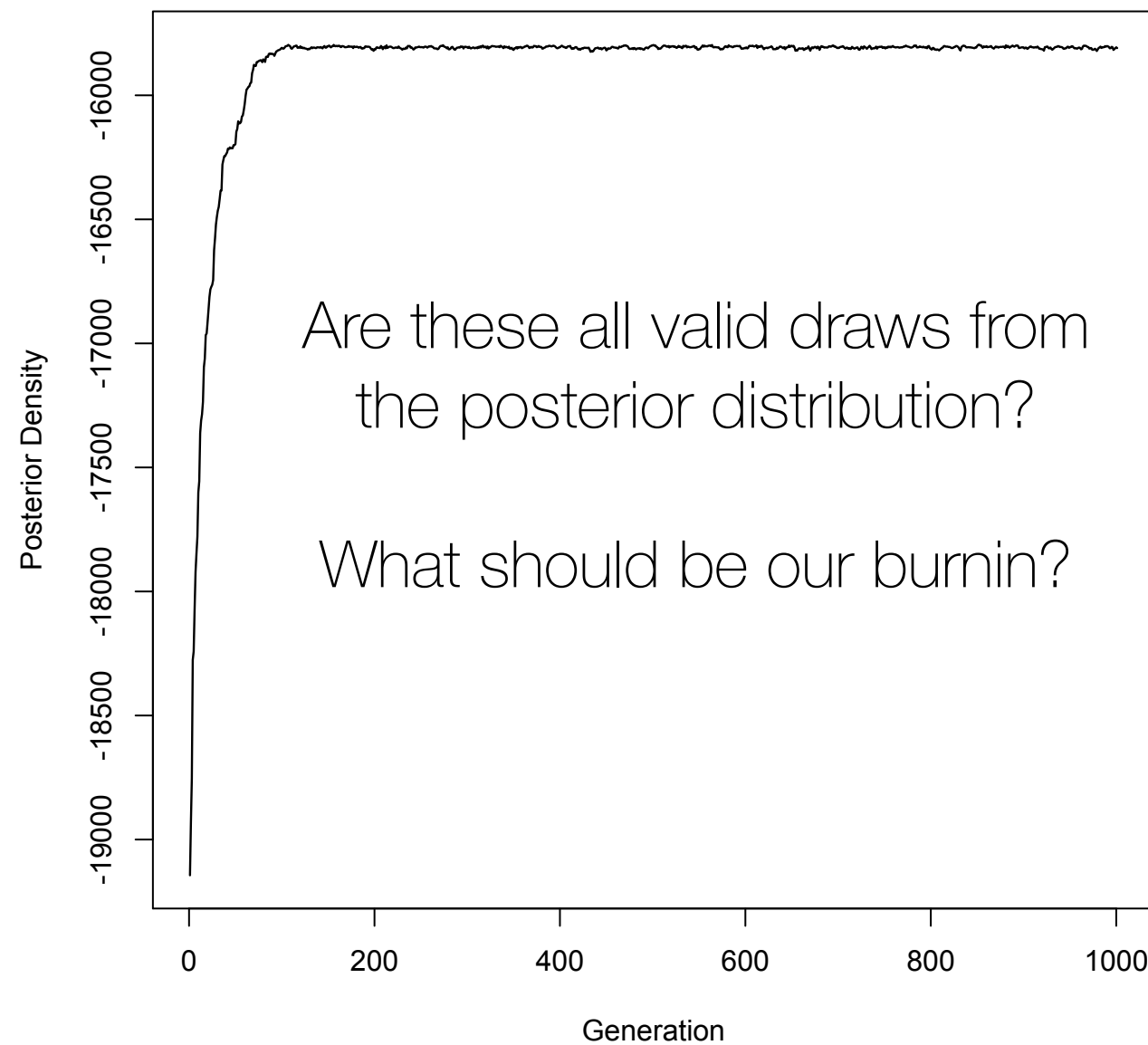
MCMC Convergence and Mixing

Premise

We can never be entirely certain that our MCMC samples give us an accurate view of the posterior, but we can look for indications of problems.

Single Analysis Diagnostics

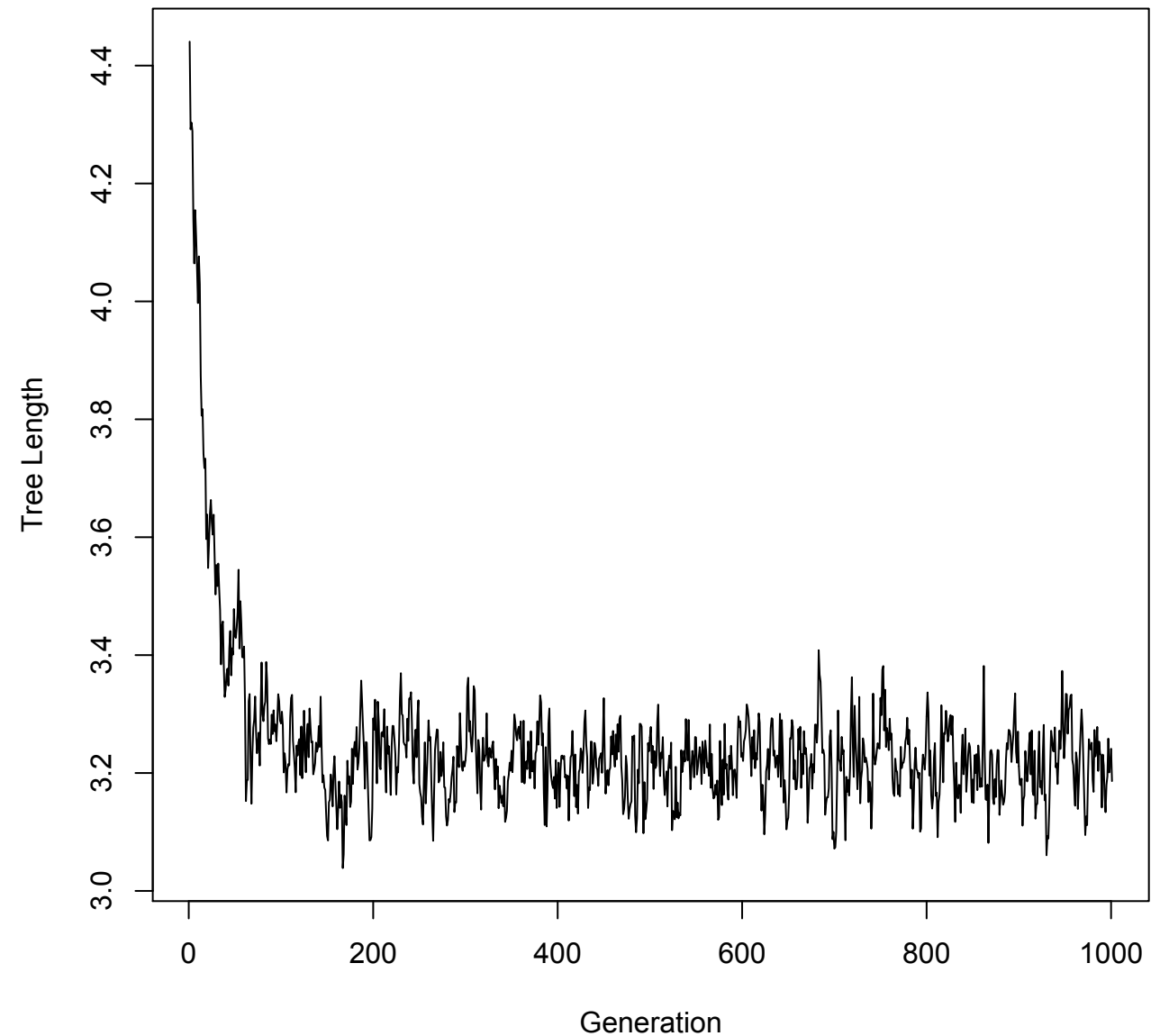
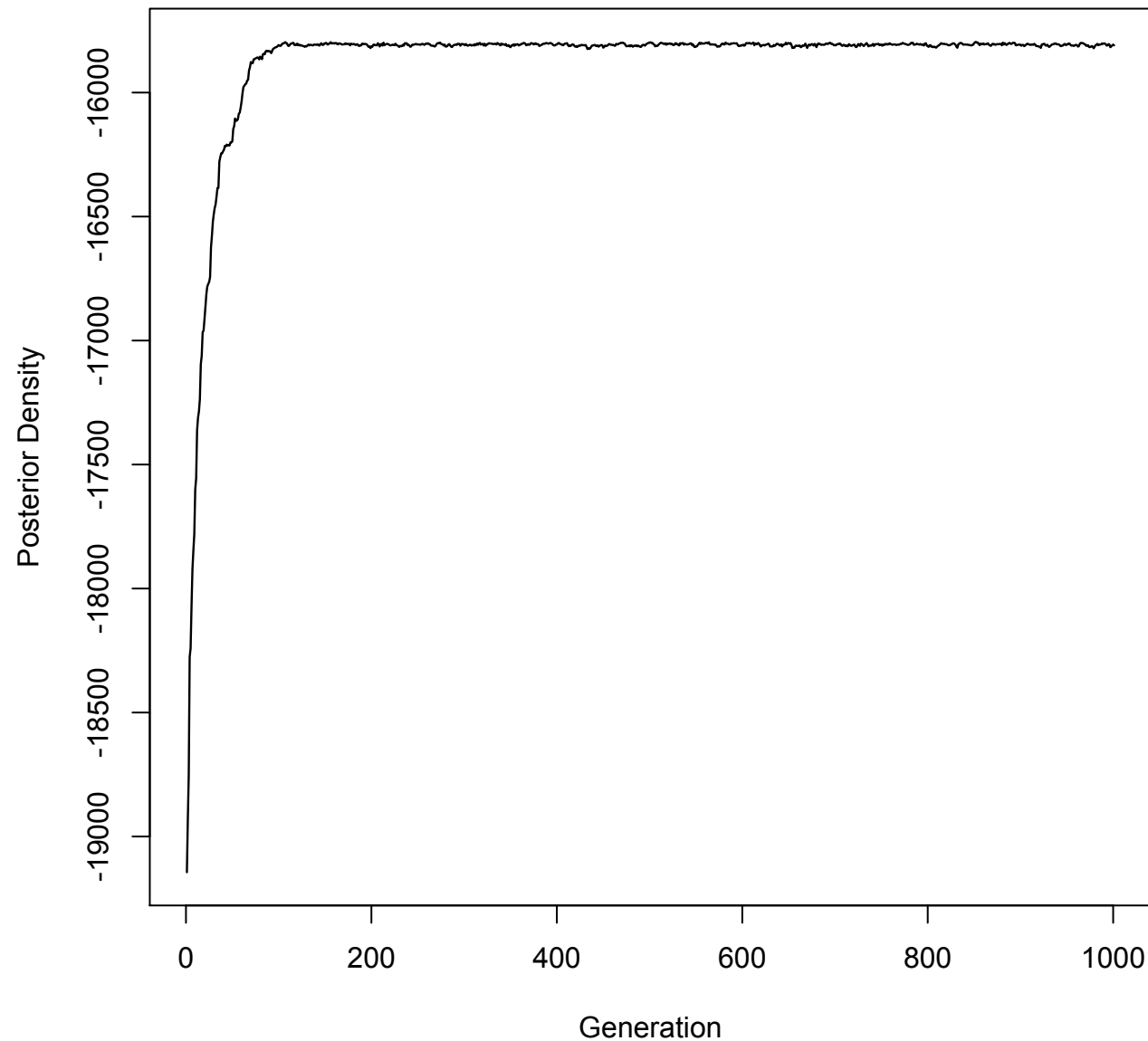
Shape of Traces



Single Analysis Diagnostics

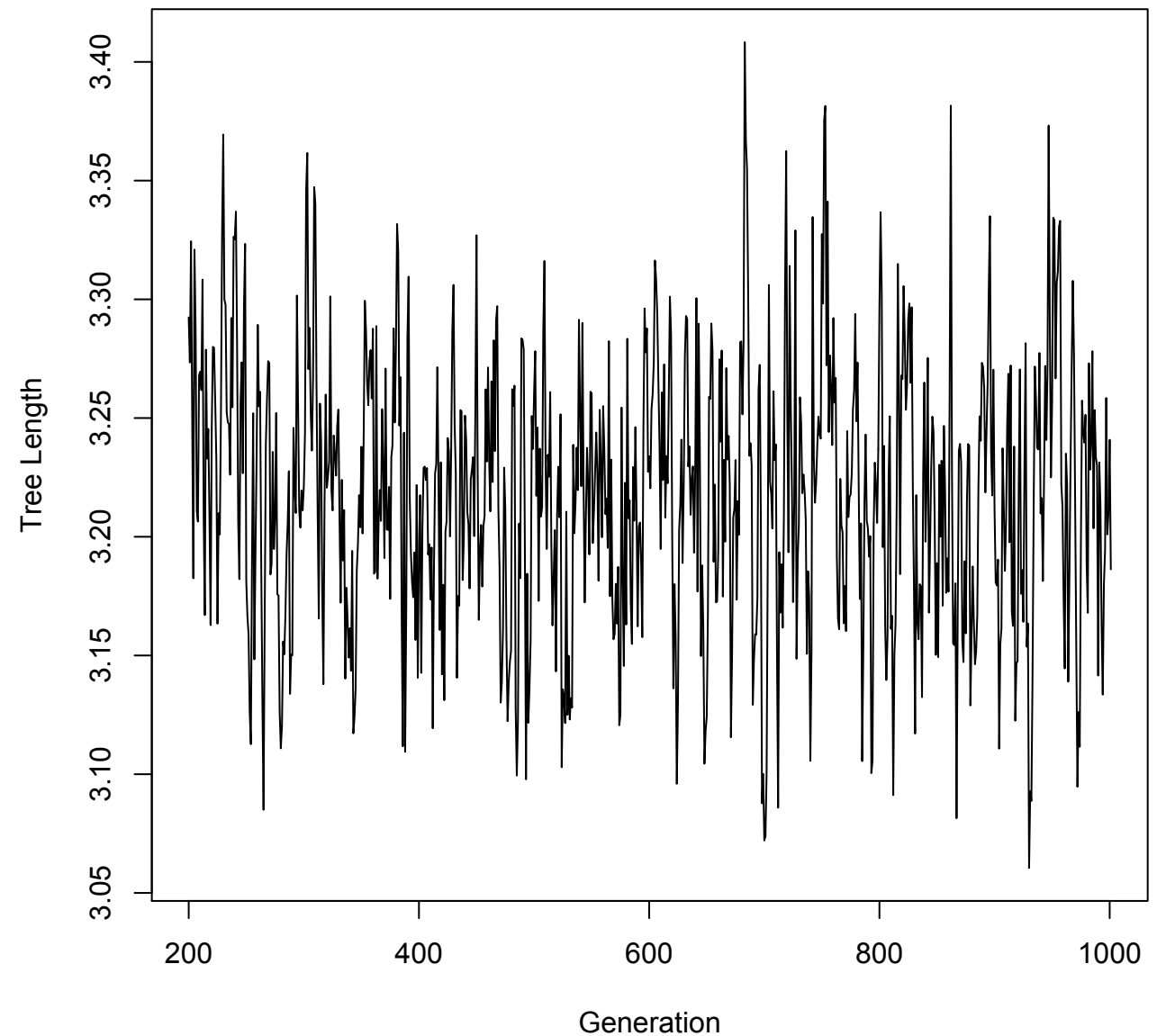
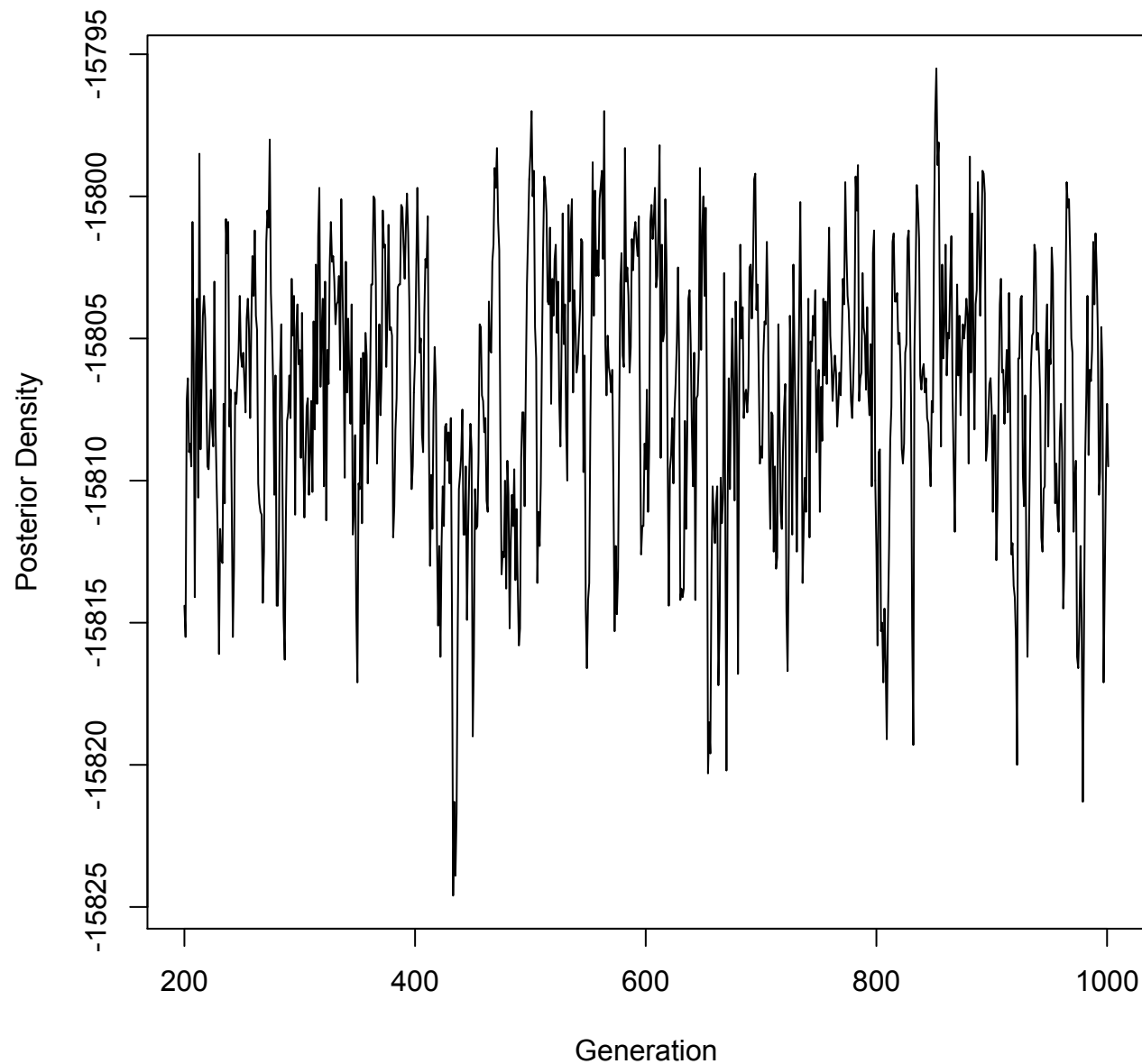
Shape of Traces

What should be our burnin?



Single Analysis Diagnostics

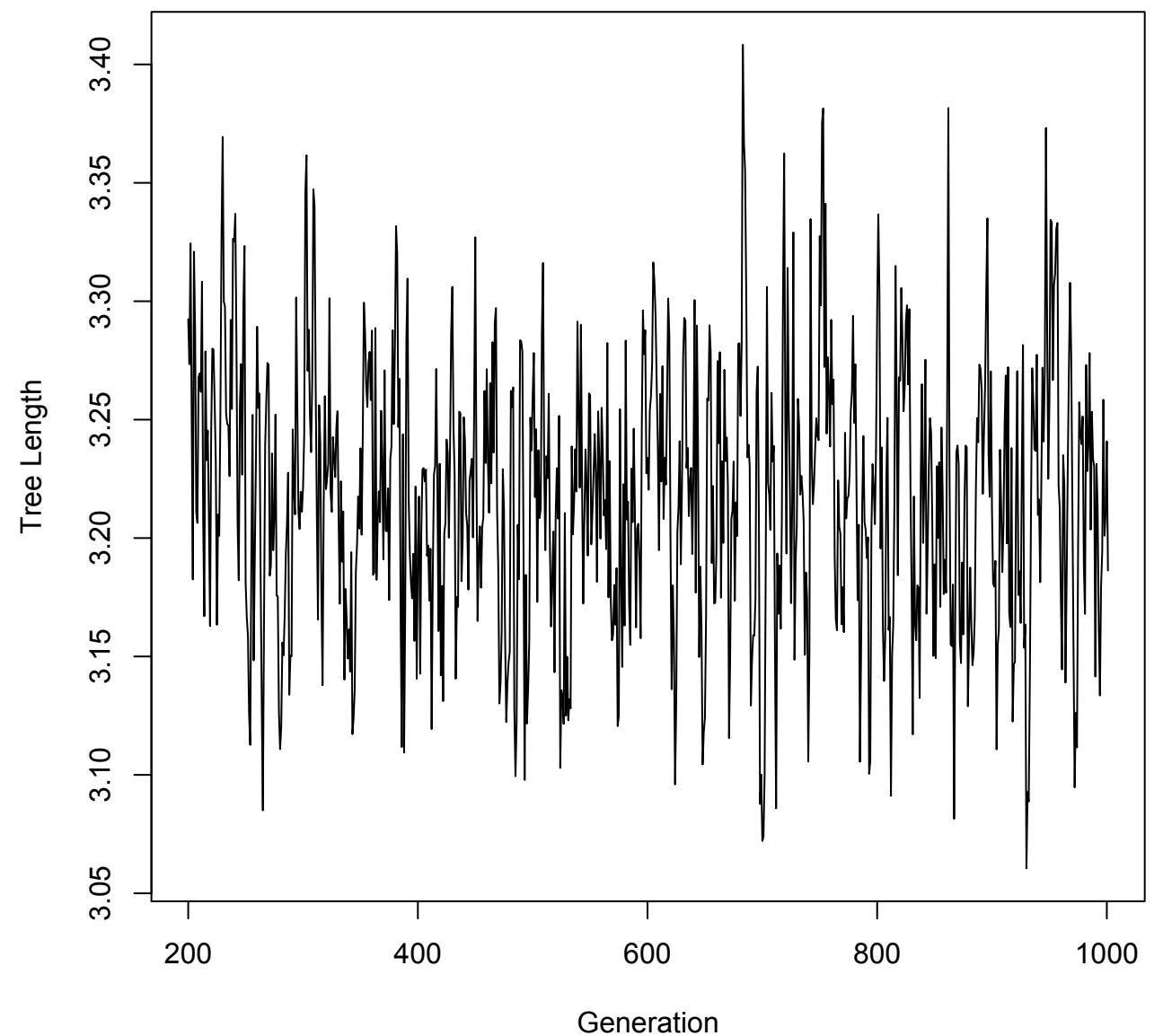
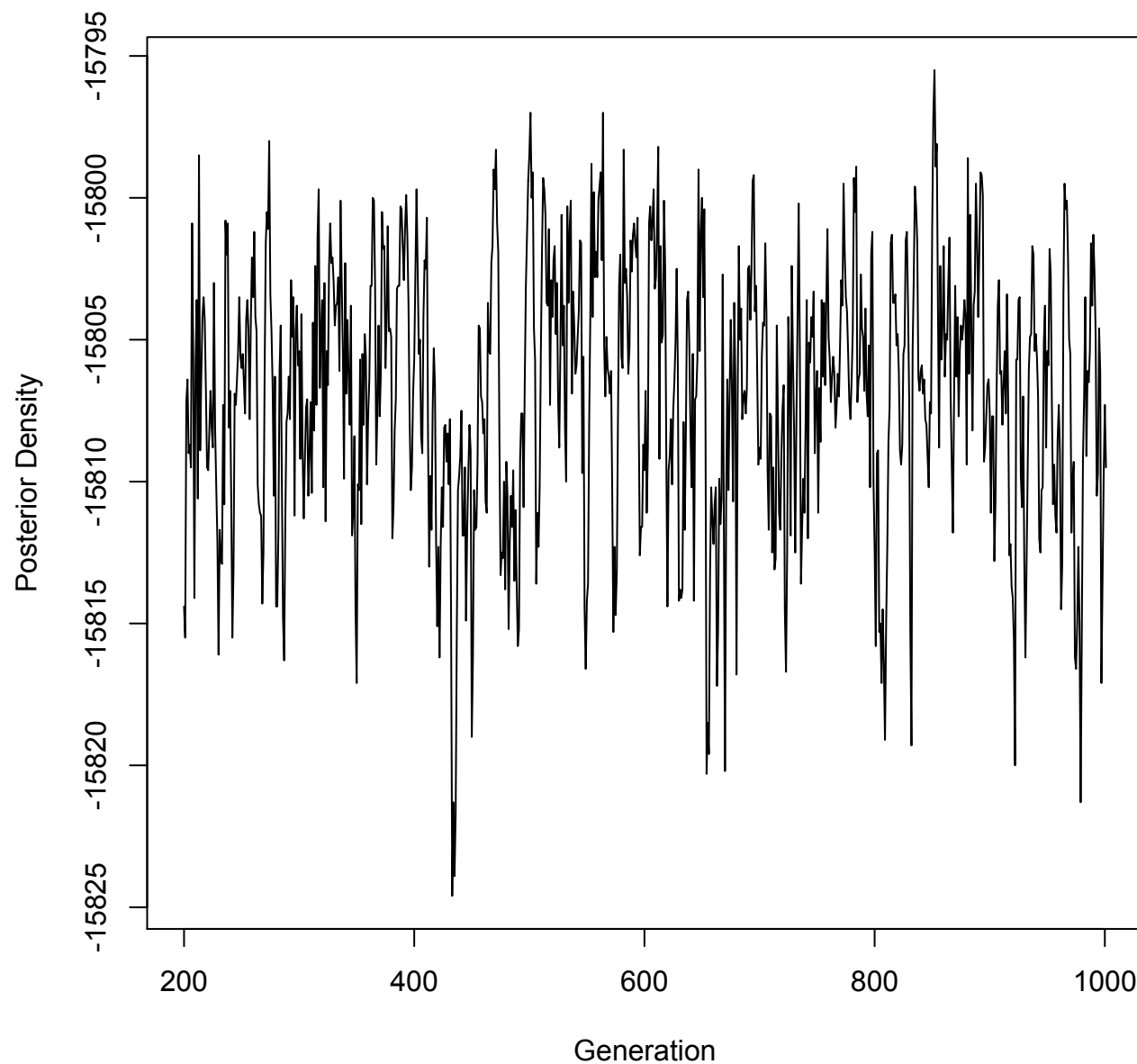
Shape of Traces



Single Analysis Diagnostics

Geweke's Diagnostic

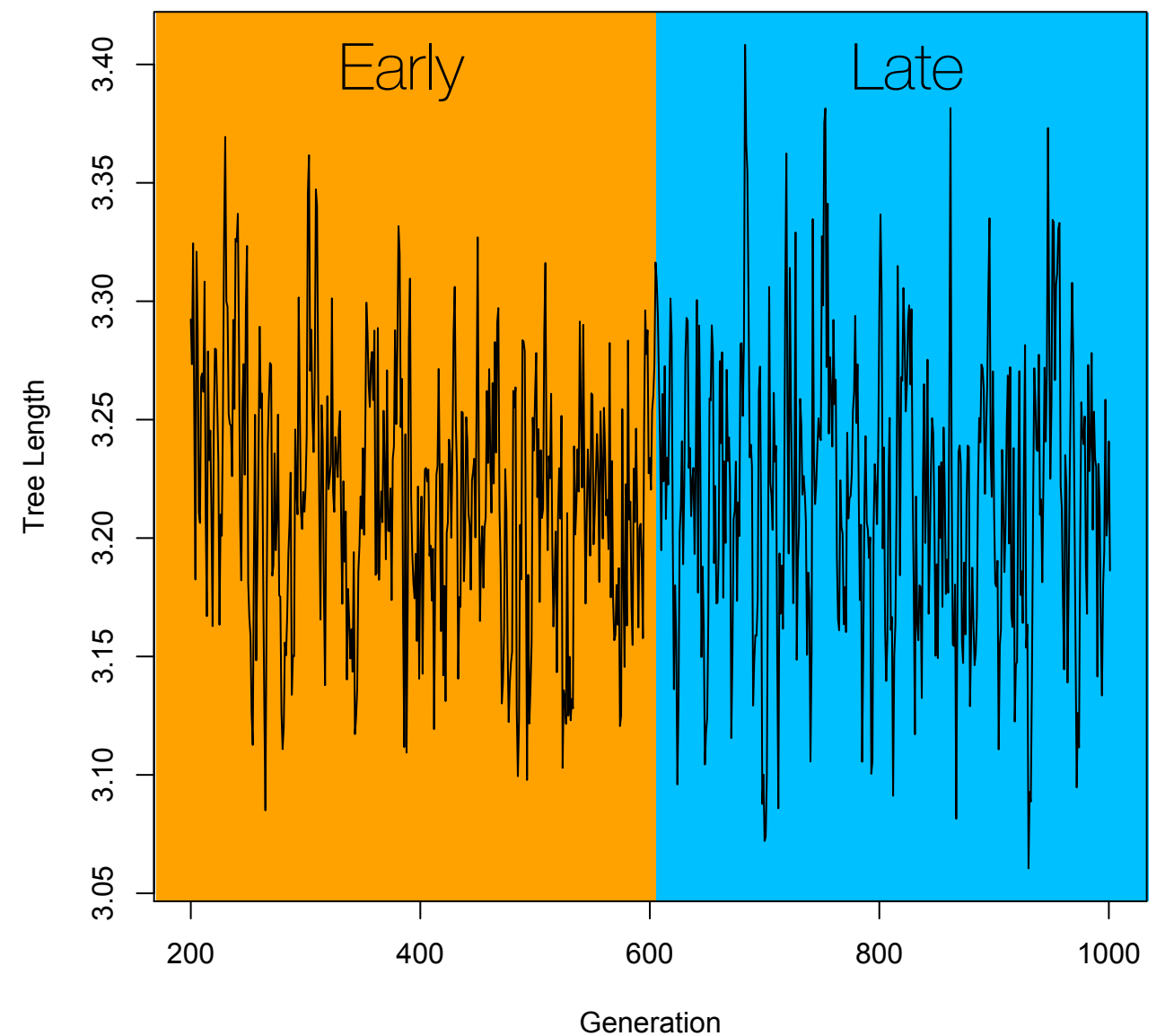
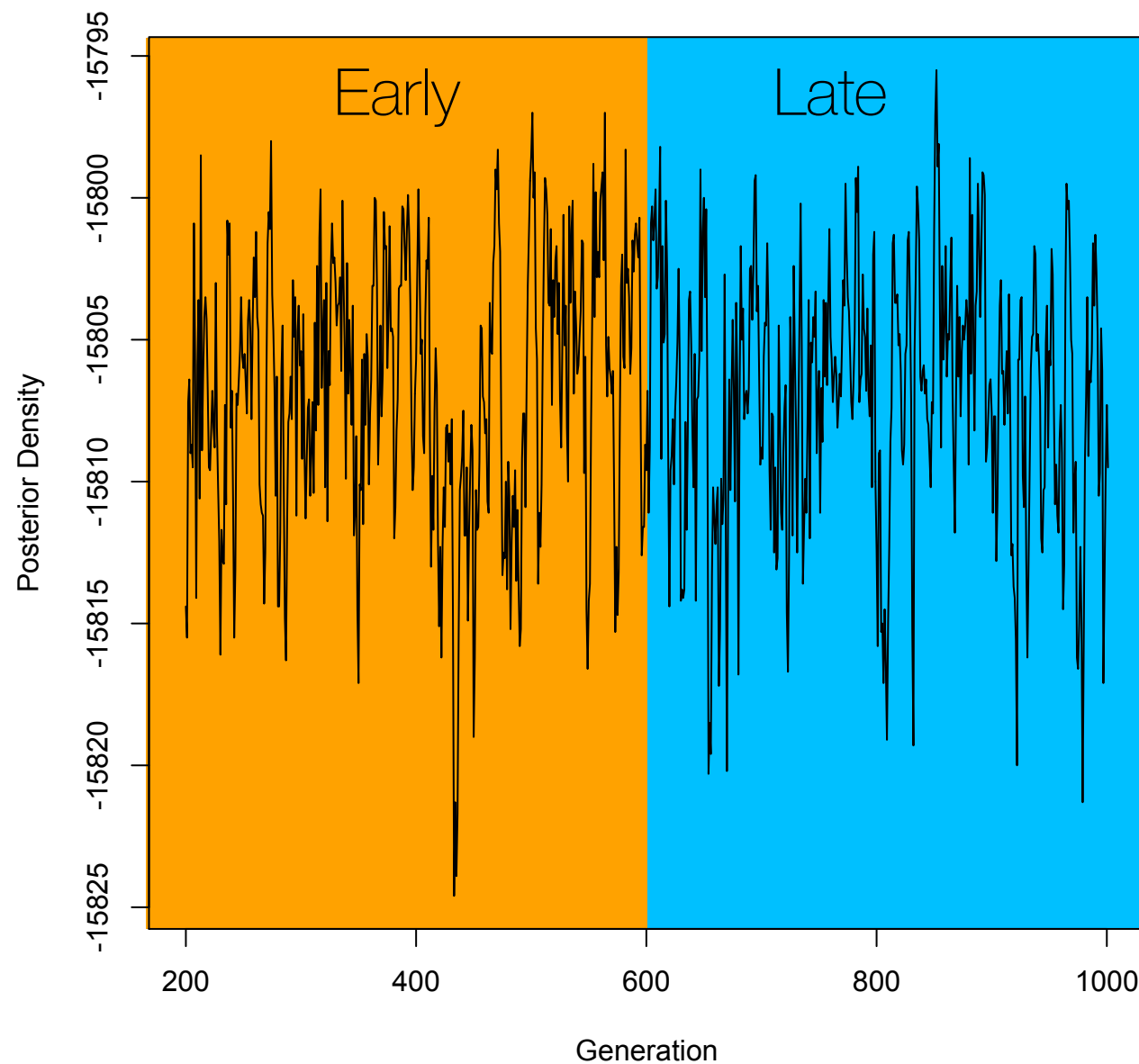
Is there a difference in means between early and late parts of chains?



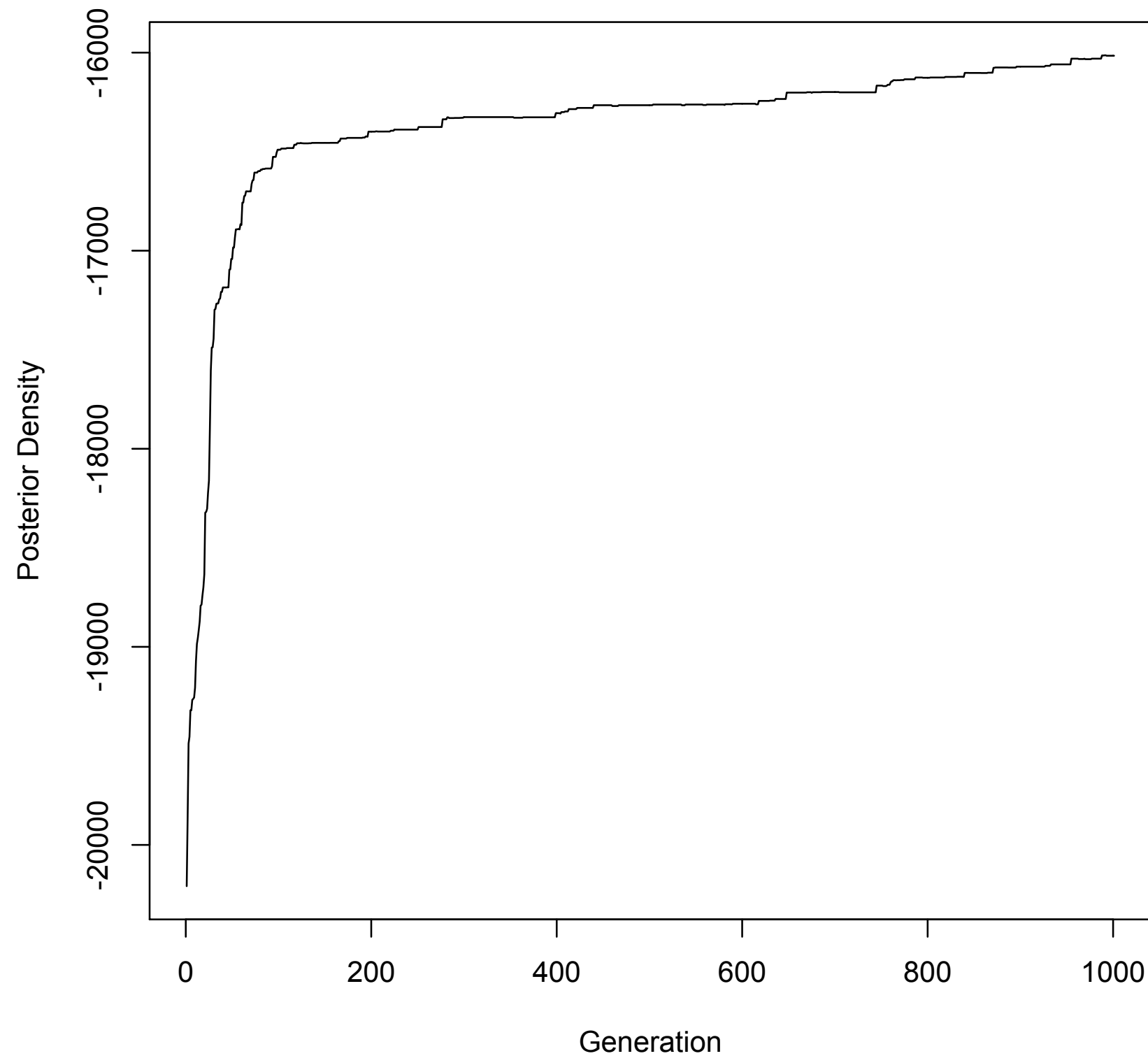
Single Analysis Diagnostics

Geweke's Diagnostic

Is there a difference in means between early and late parts of chains?



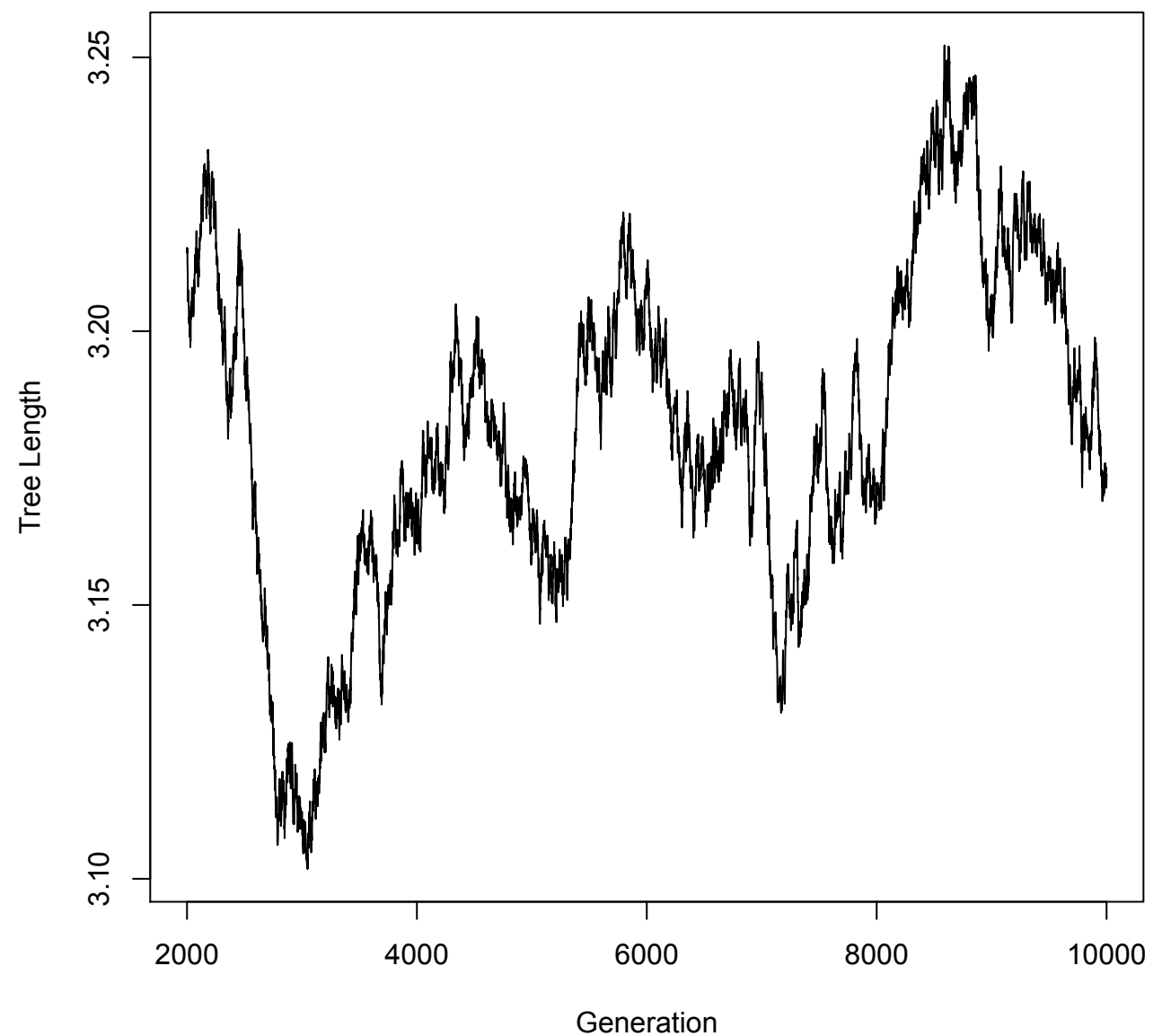
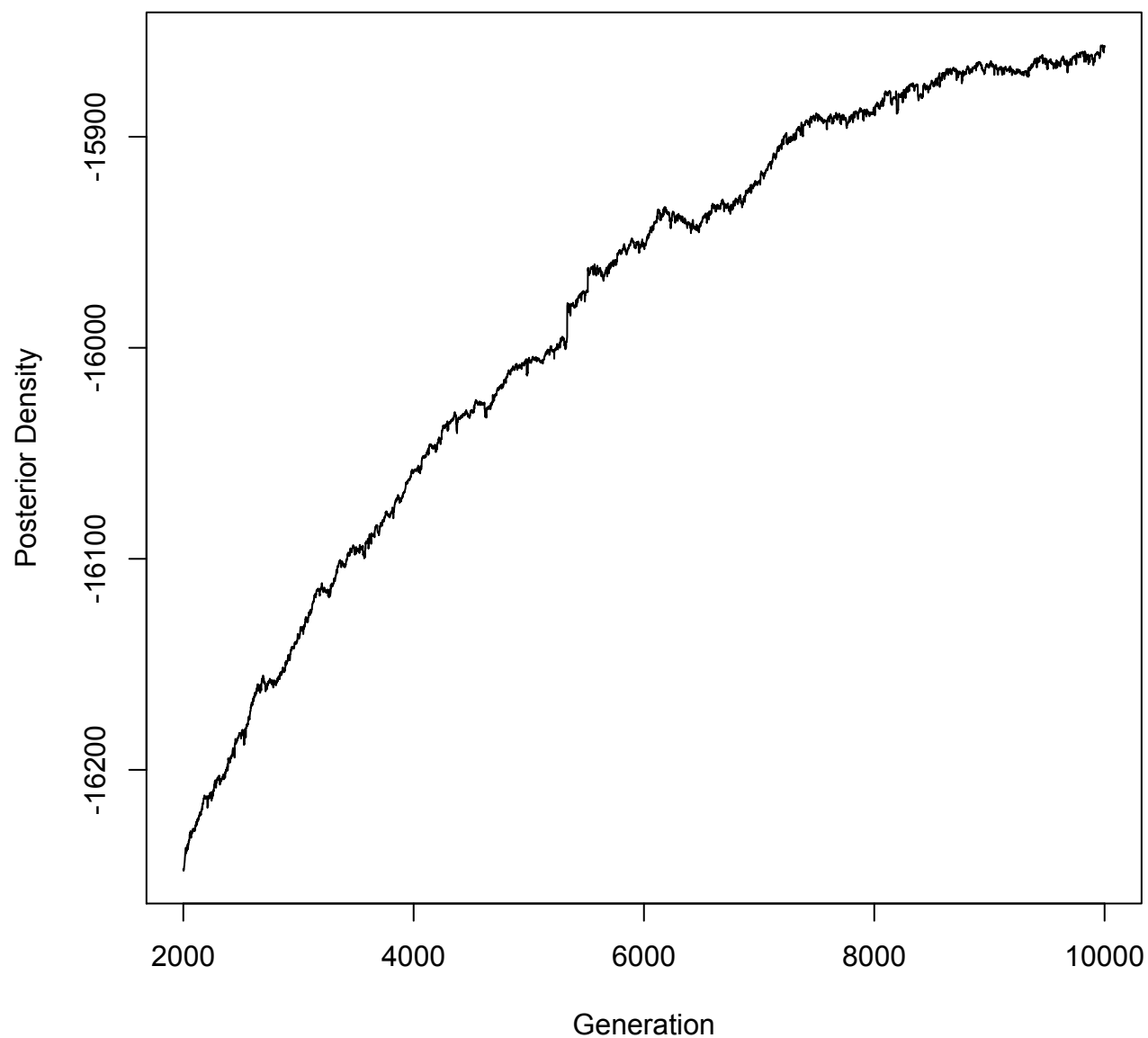
Has this run converged?



Single Analysis Diagnostics

Shape of Traces

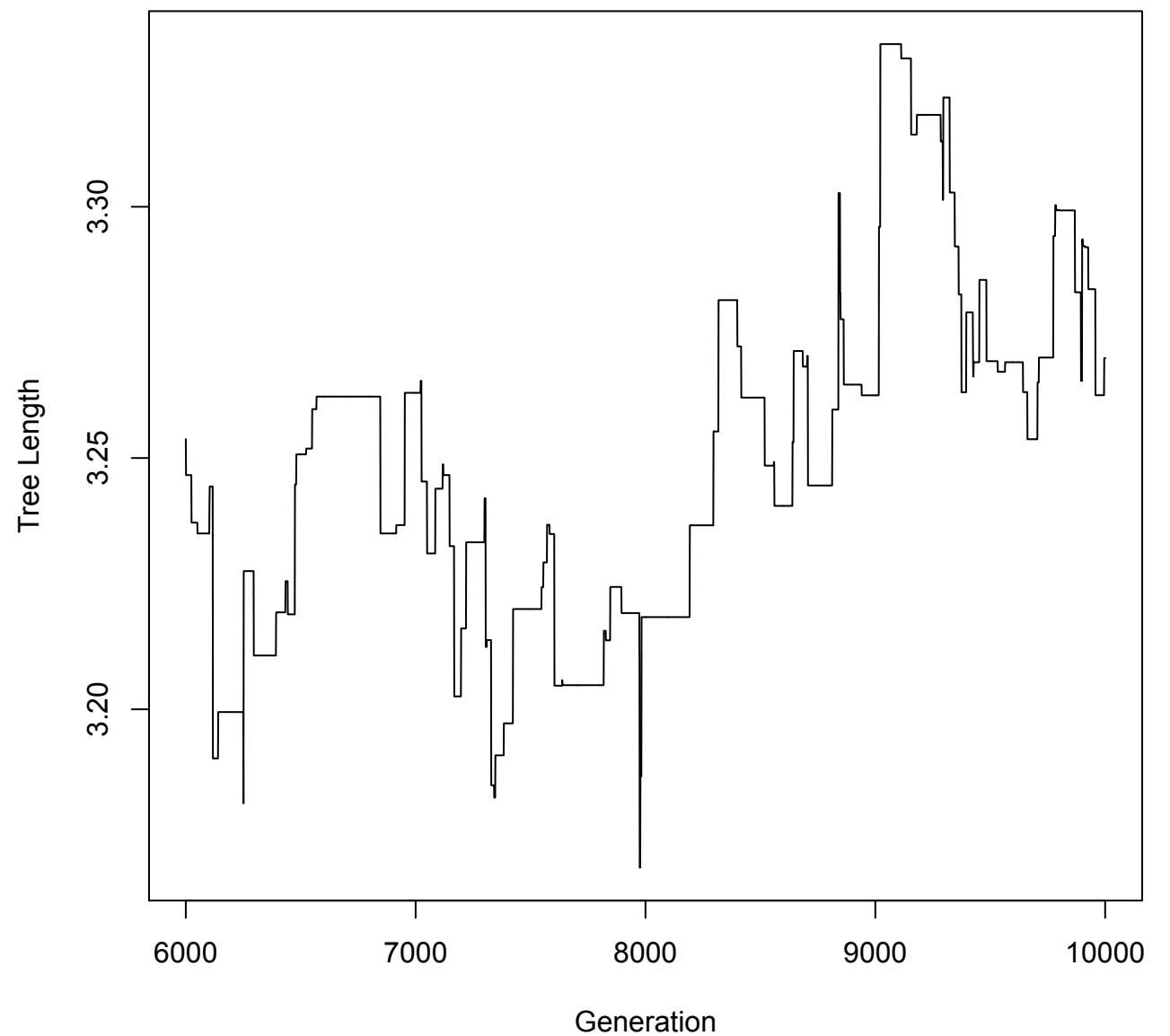
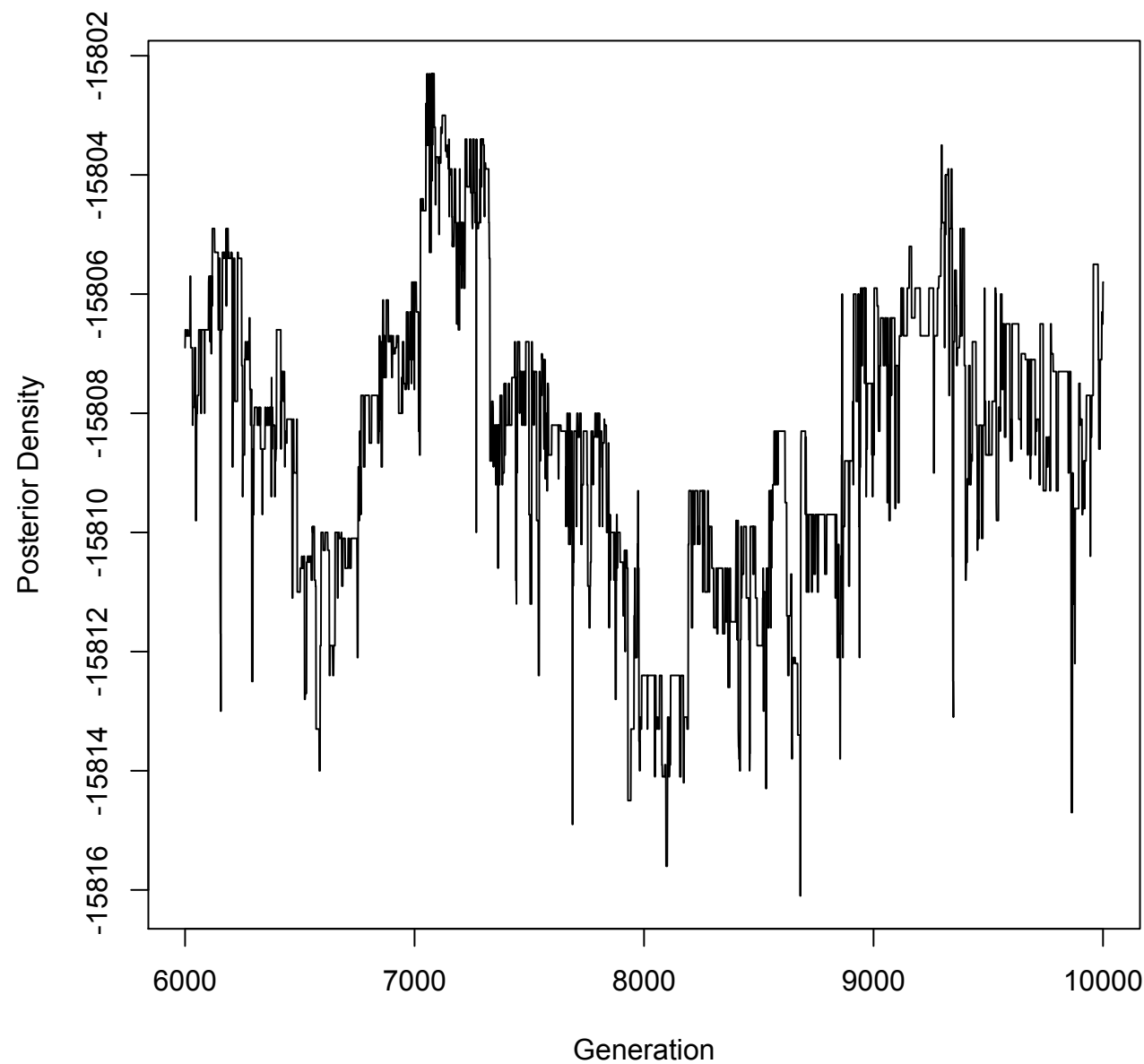
BAD MIXING - Proposals Too **Small**



Single Analysis Diagnostics

Shape of Traces

BAD MIXING - Proposals Too **Big**



Single Analysis Diagnostics

Acceptance Rates

Proposals should ideally be accepted between **20% and 70%** of the time.

> 70% means proposals too small.

< 20% means proposals too big.

By default, RevBayes tunes proposals so that they are in this range.

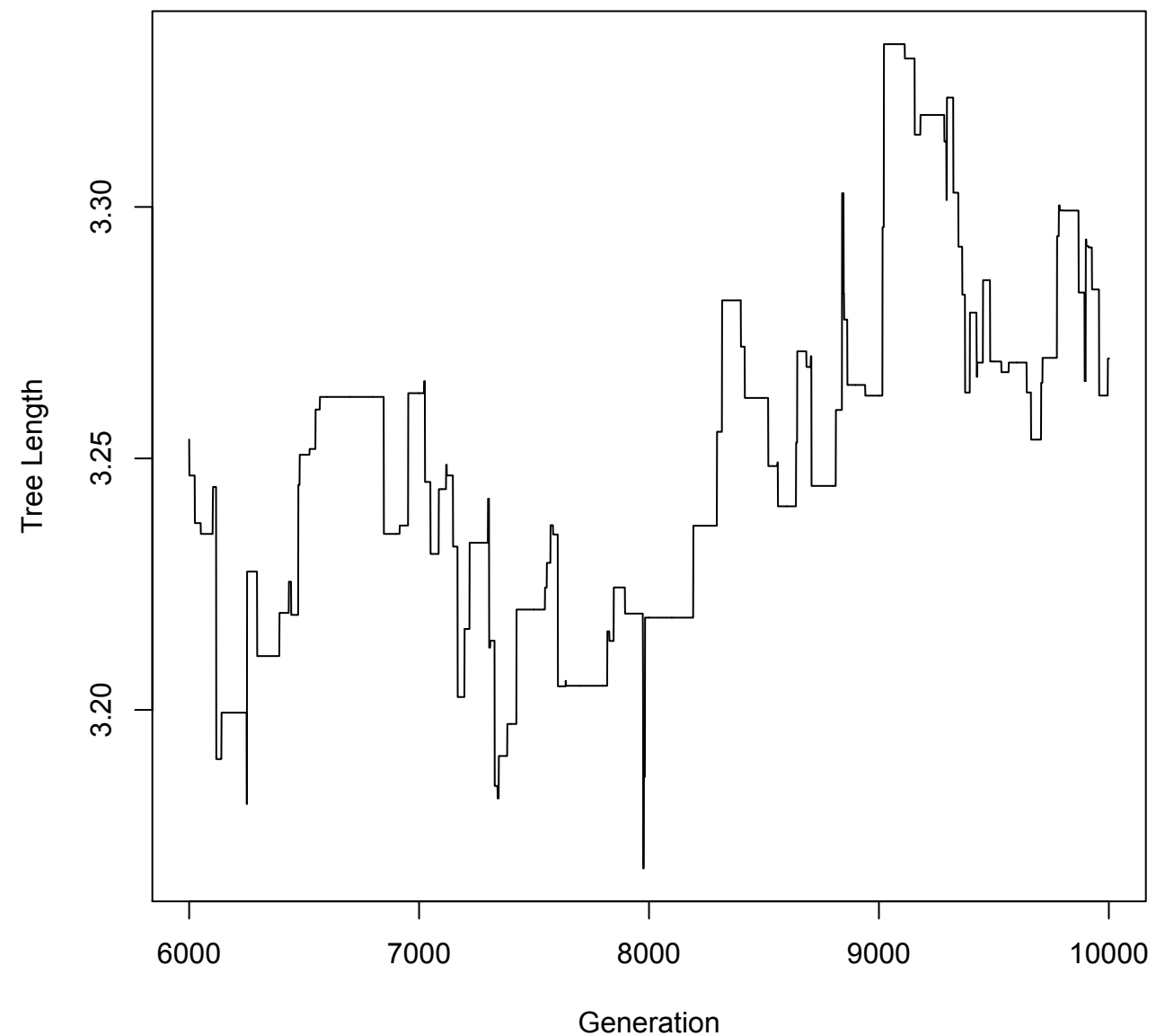
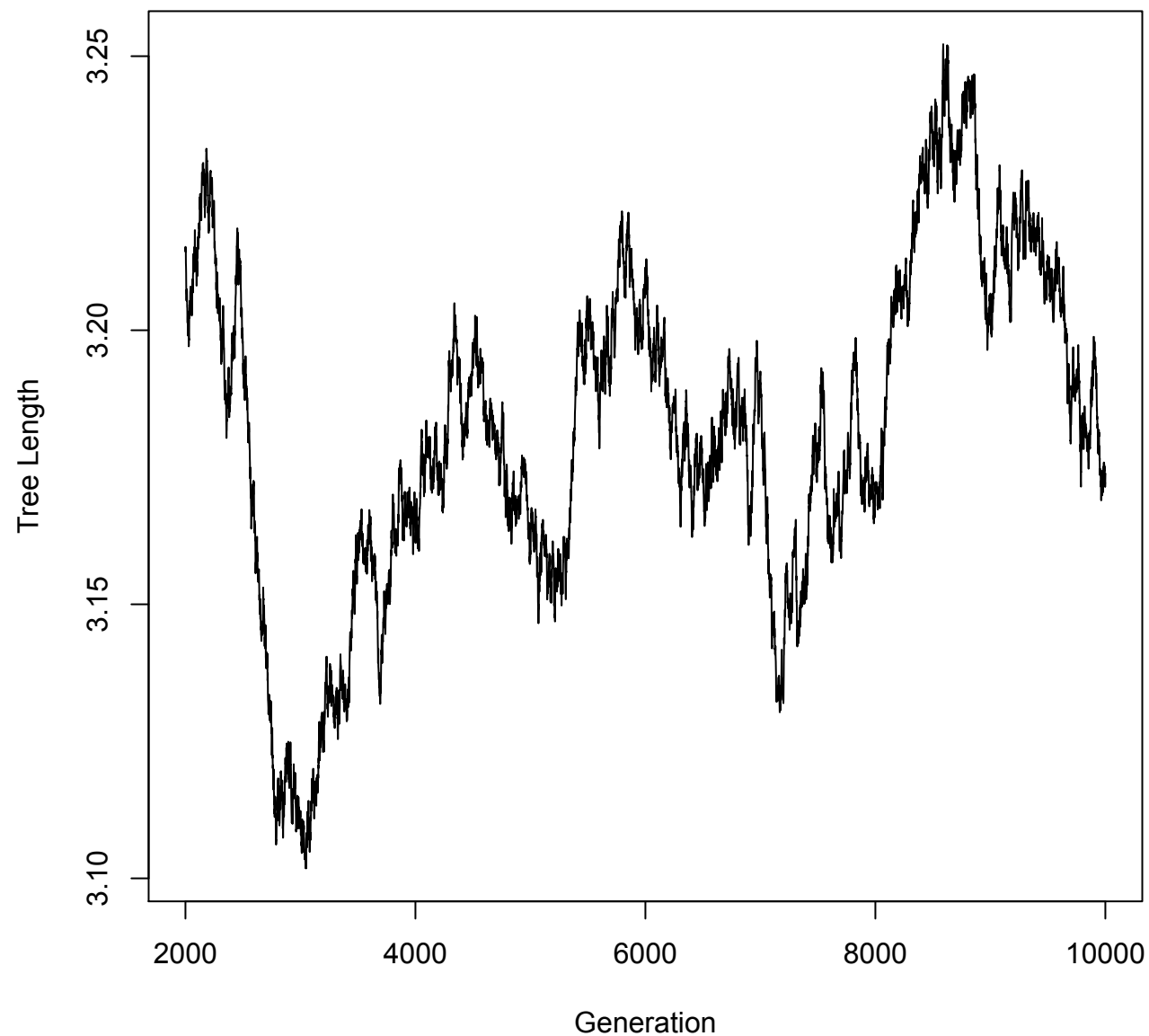
Single Analysis Diagnostics

Autocorrelation

Lag = the interval (number of generations) between samples that are not correlated with one another

Single Analysis Diagnostics

Autocorrelation



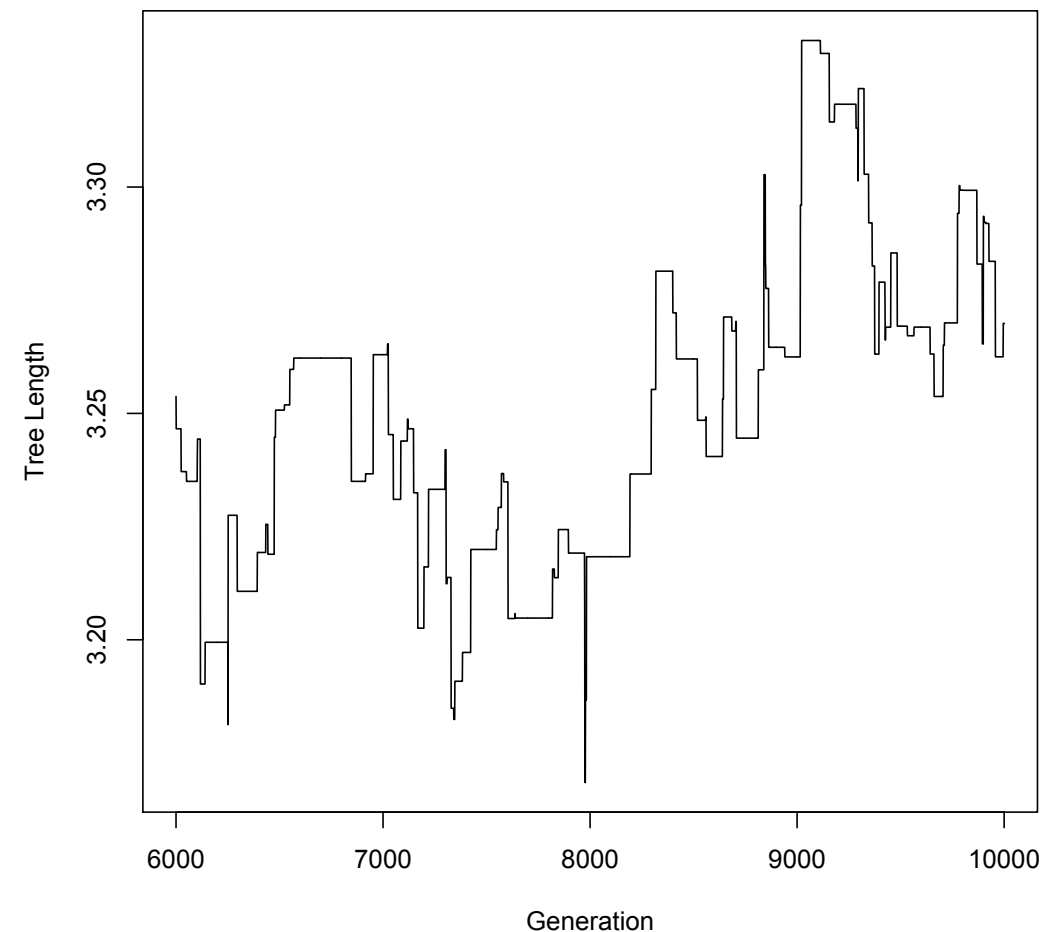
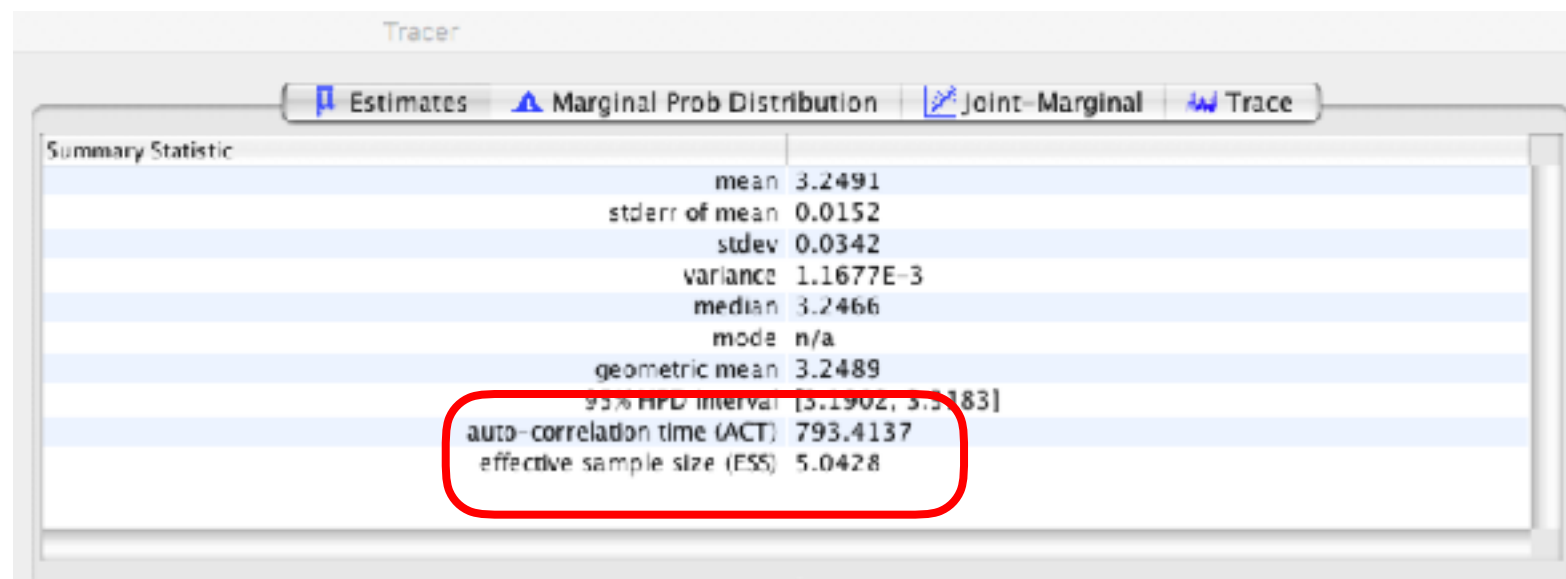
Single Analysis Diagnostics

Effective Sample Size

Given my autocorrelation, how many “independent” samples do I really have?

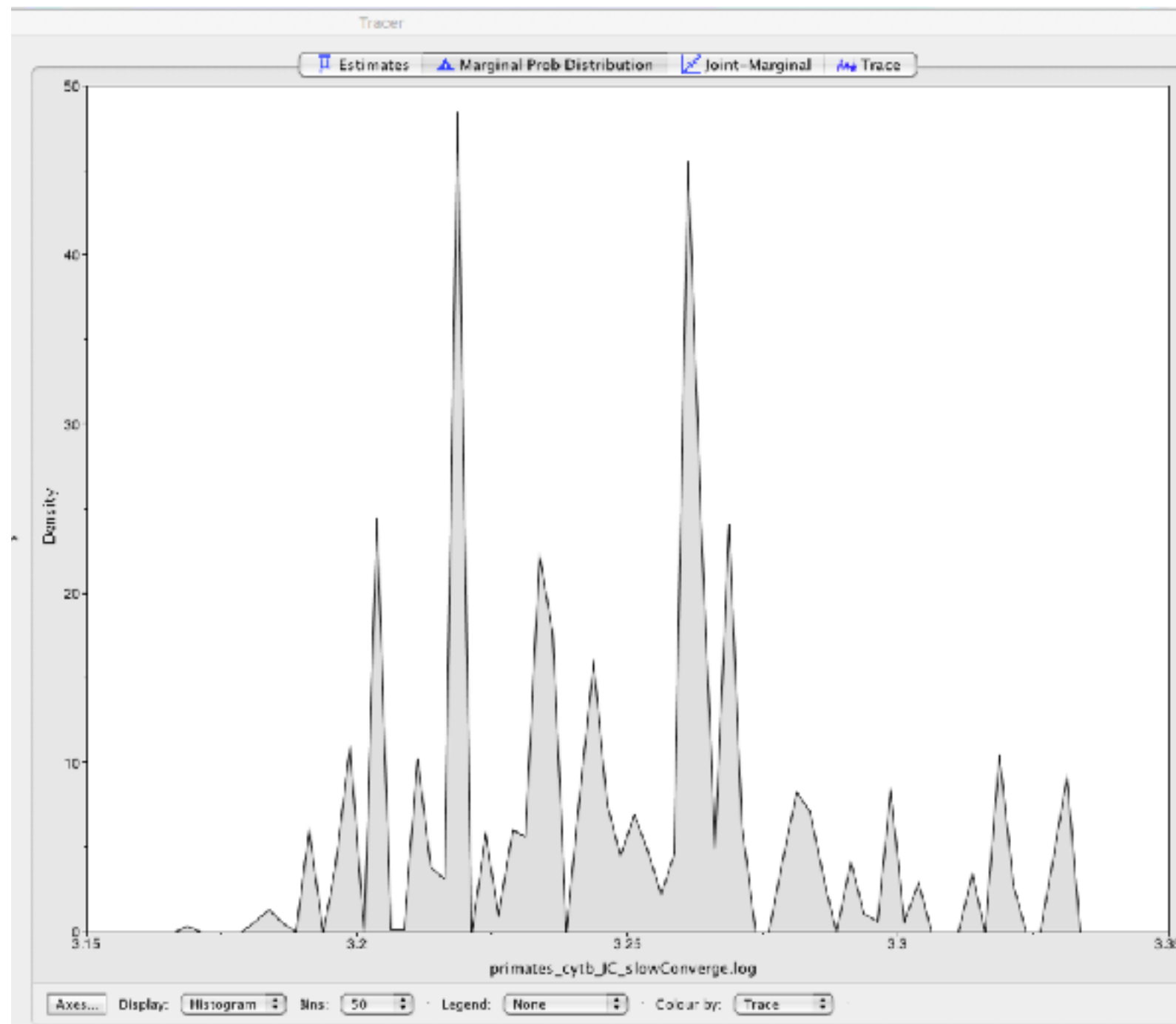
Basically, we’re correcting the total number of samples we’ve gathered for their autocorrelation.

Single Analysis Diagnostics



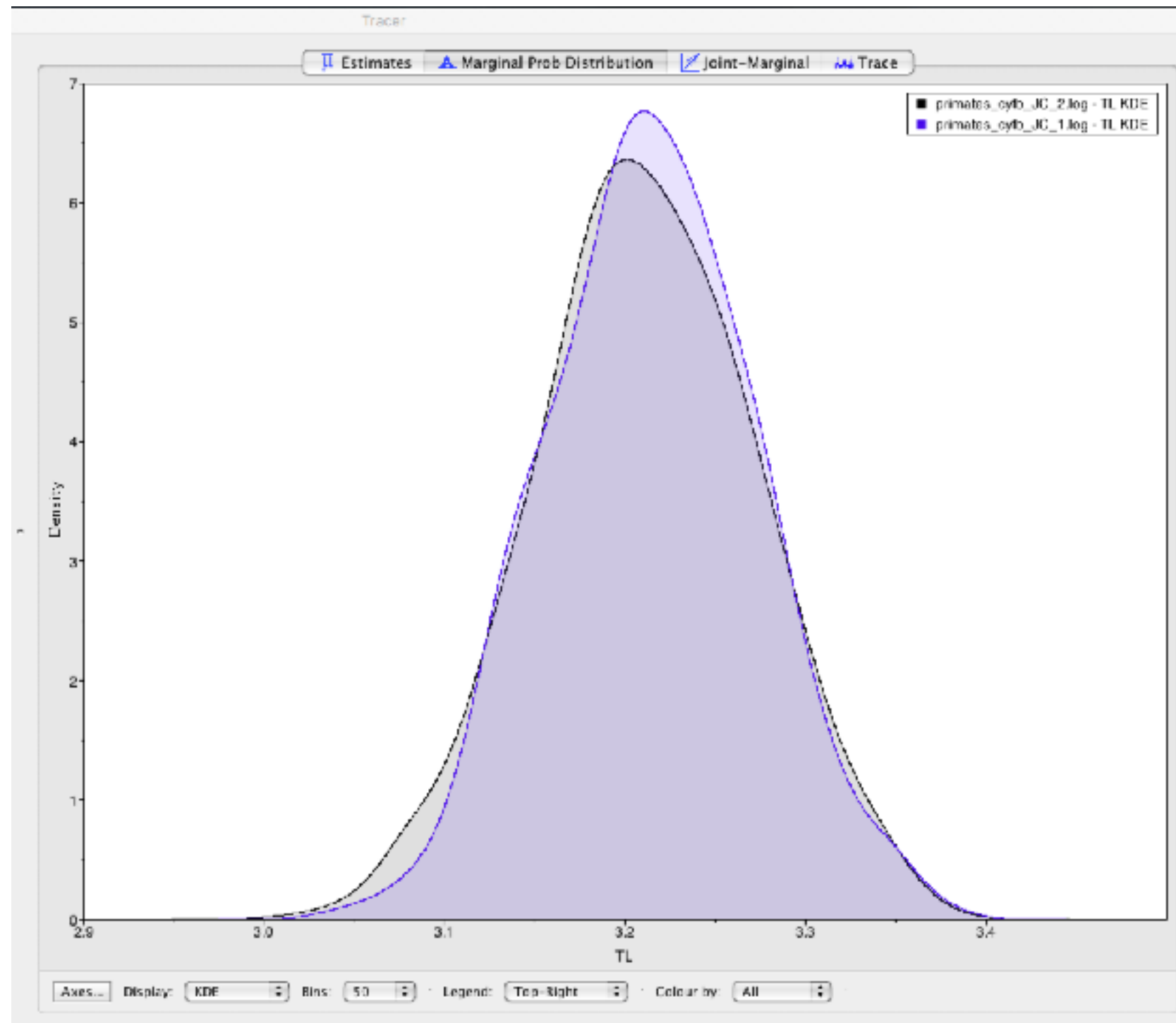
Single Analysis Diagnostics

Jagged Marginal Distributions



Multiple Analysis Diagnostics

Compare Marginal Distributions



Multiple Analysis Diagnostics

Potential Scale Reduction Factor (PSRF)

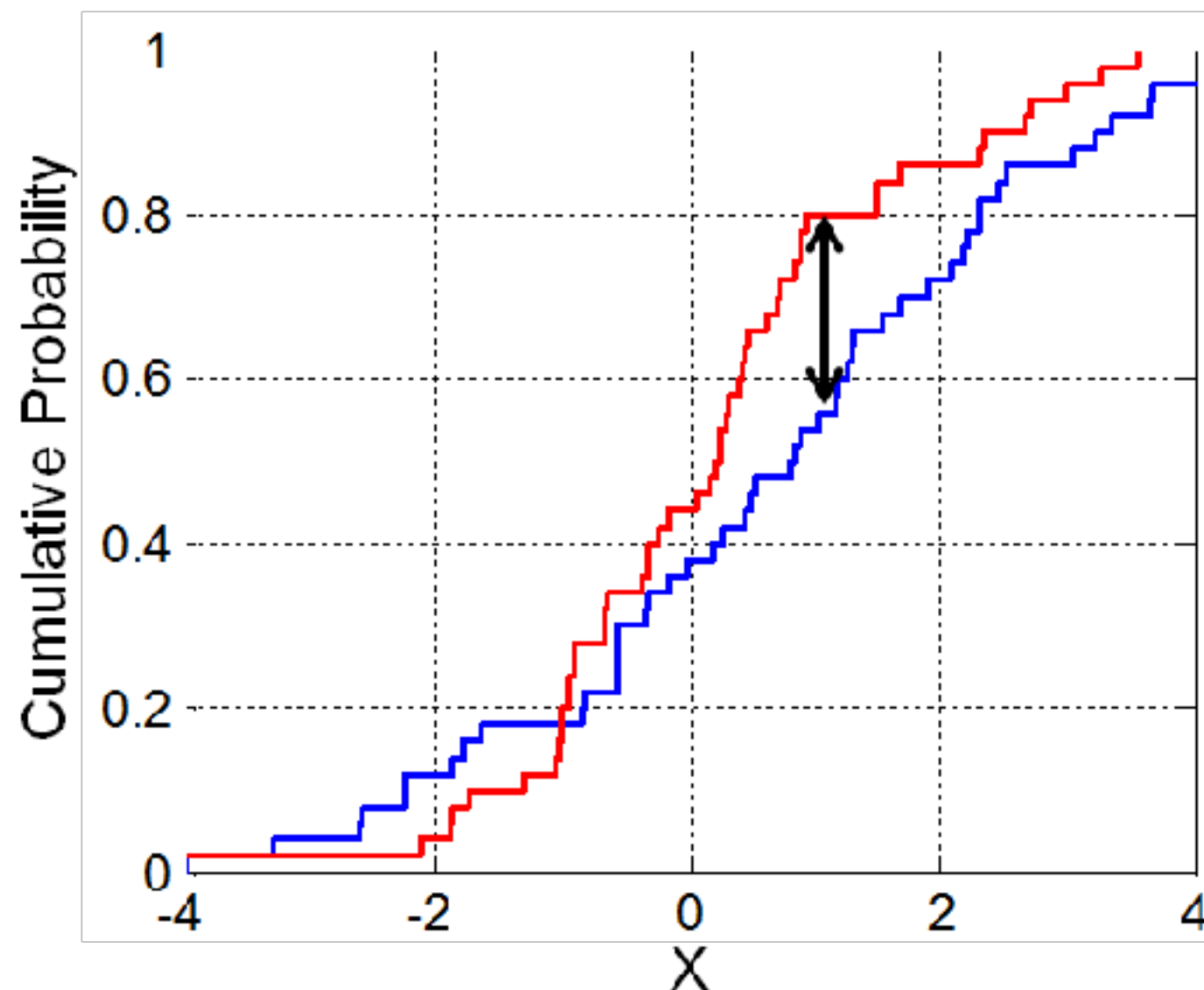
Gelman-Rubin Diagnostic

What is the ratio of the variances within runs to combined runs.

If convergence is good, $PSRF = 1$.

Multiple Analysis Diagnostics

Kolmogorov-Smirnov (KS) Test Compares Two Distributions



Multiple Analysis Diagnostics

Bipartition Probabilities

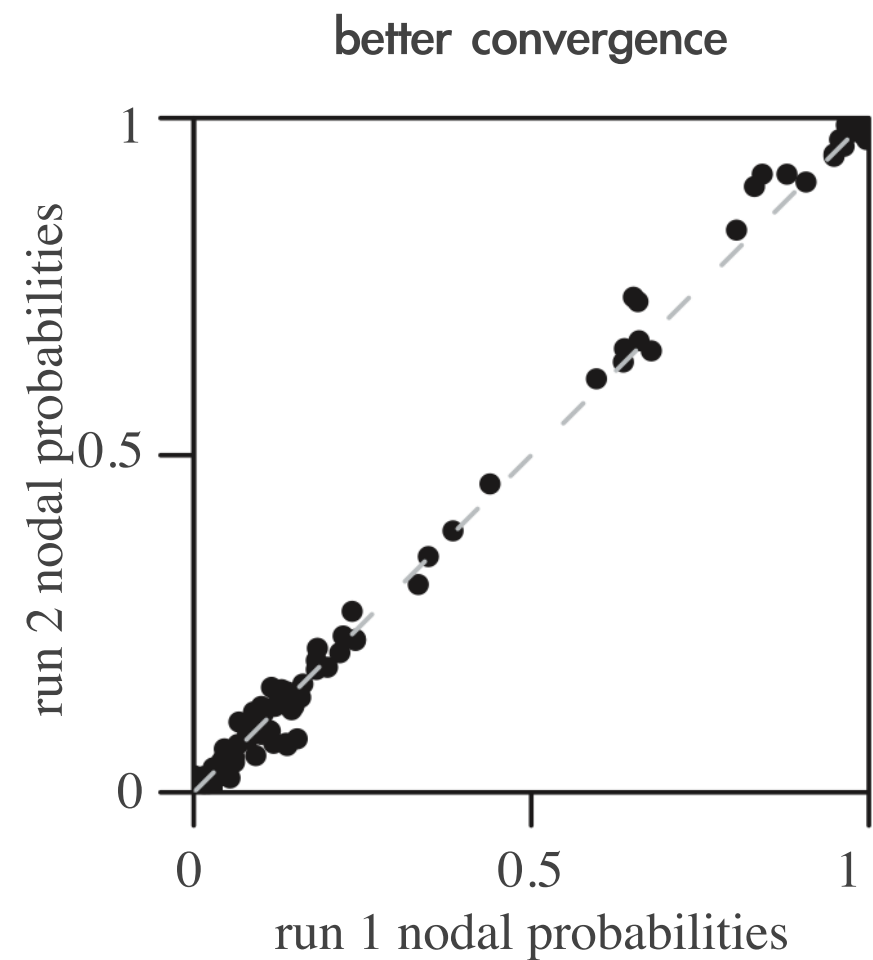
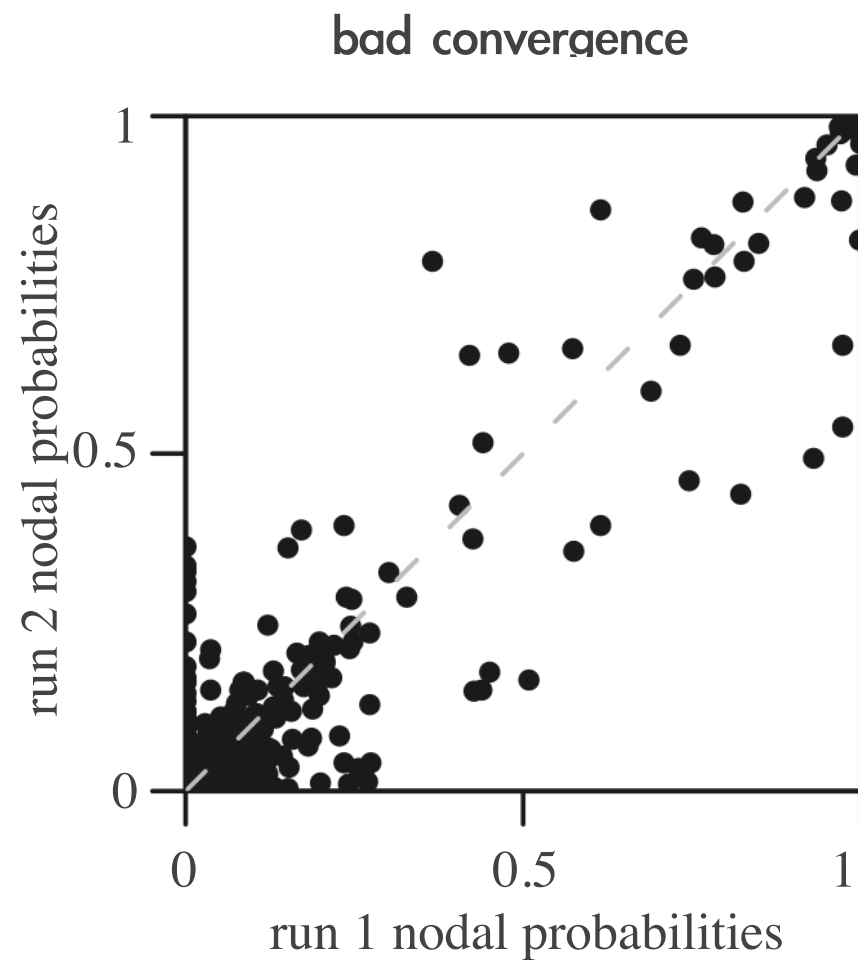
Average Standard Deviation of Split Frequencies
(ASDSF)

Numerical summary of how variable the split (bipartition) frequencies are across runs.

Ideally, close to 0.

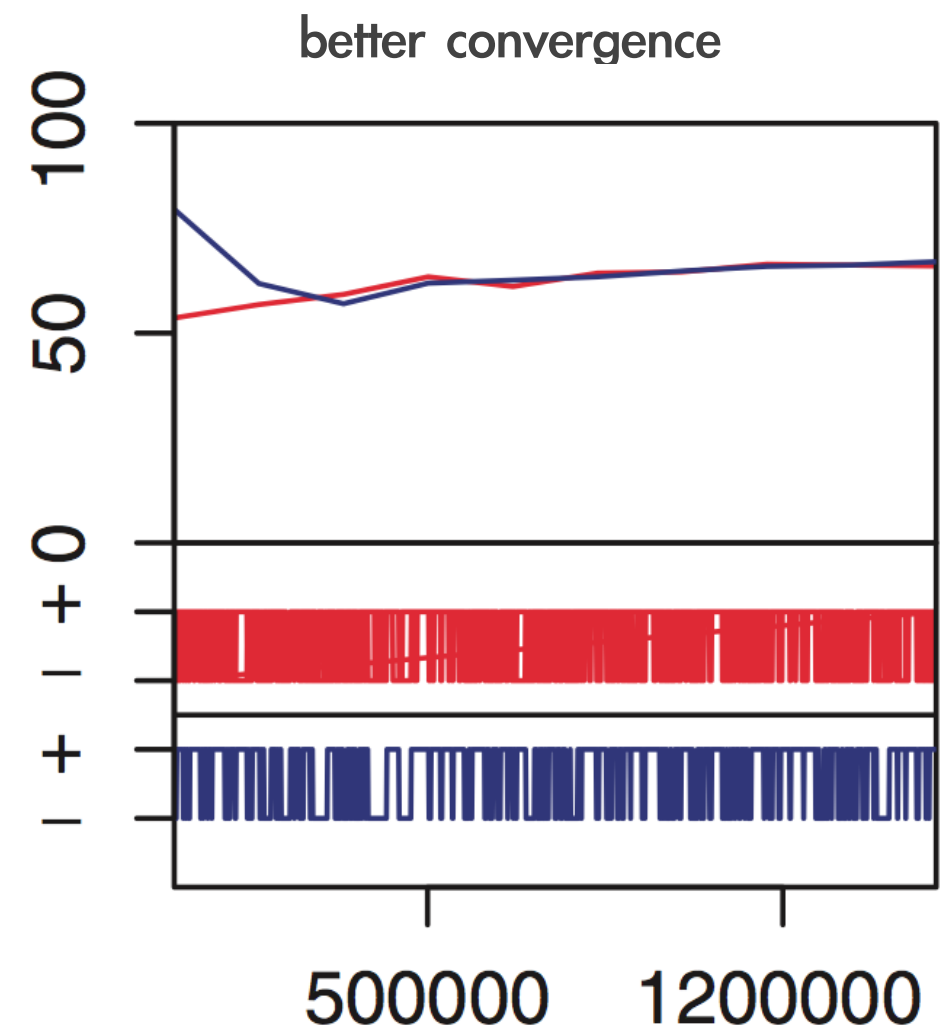
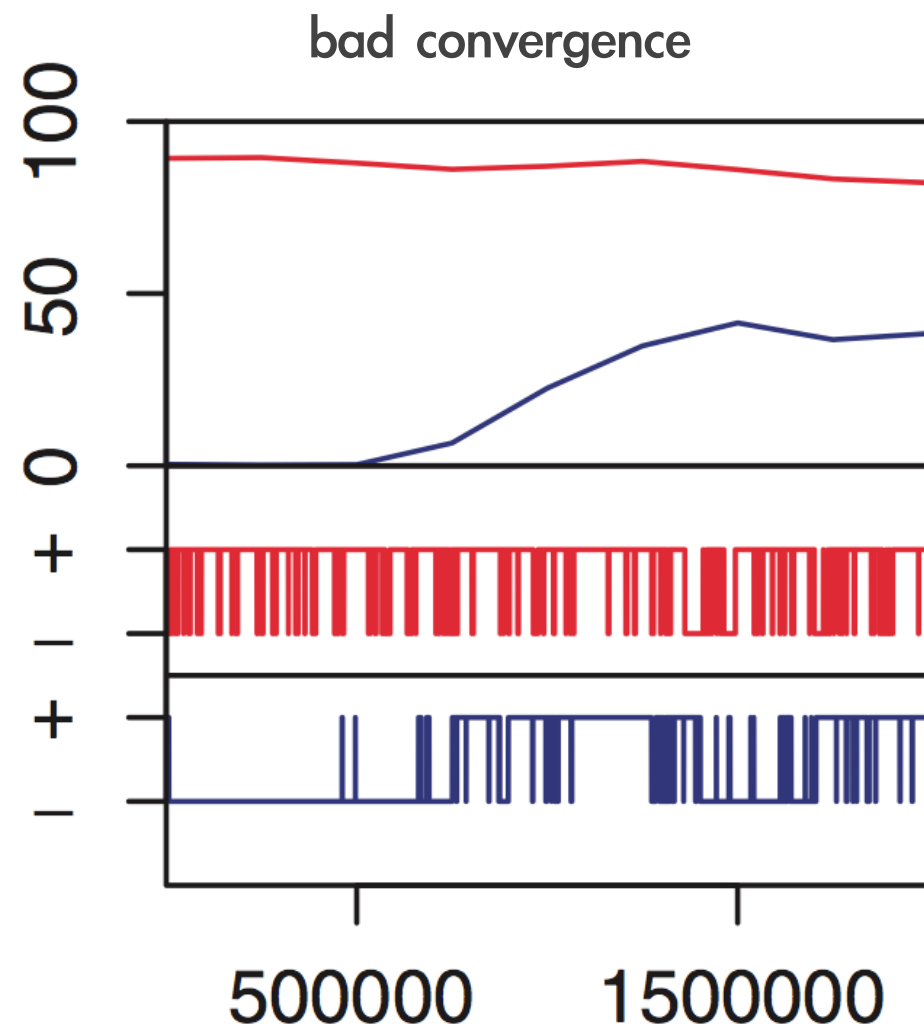
Multiple Analysis Diagnostics

Bipartition Probabilities



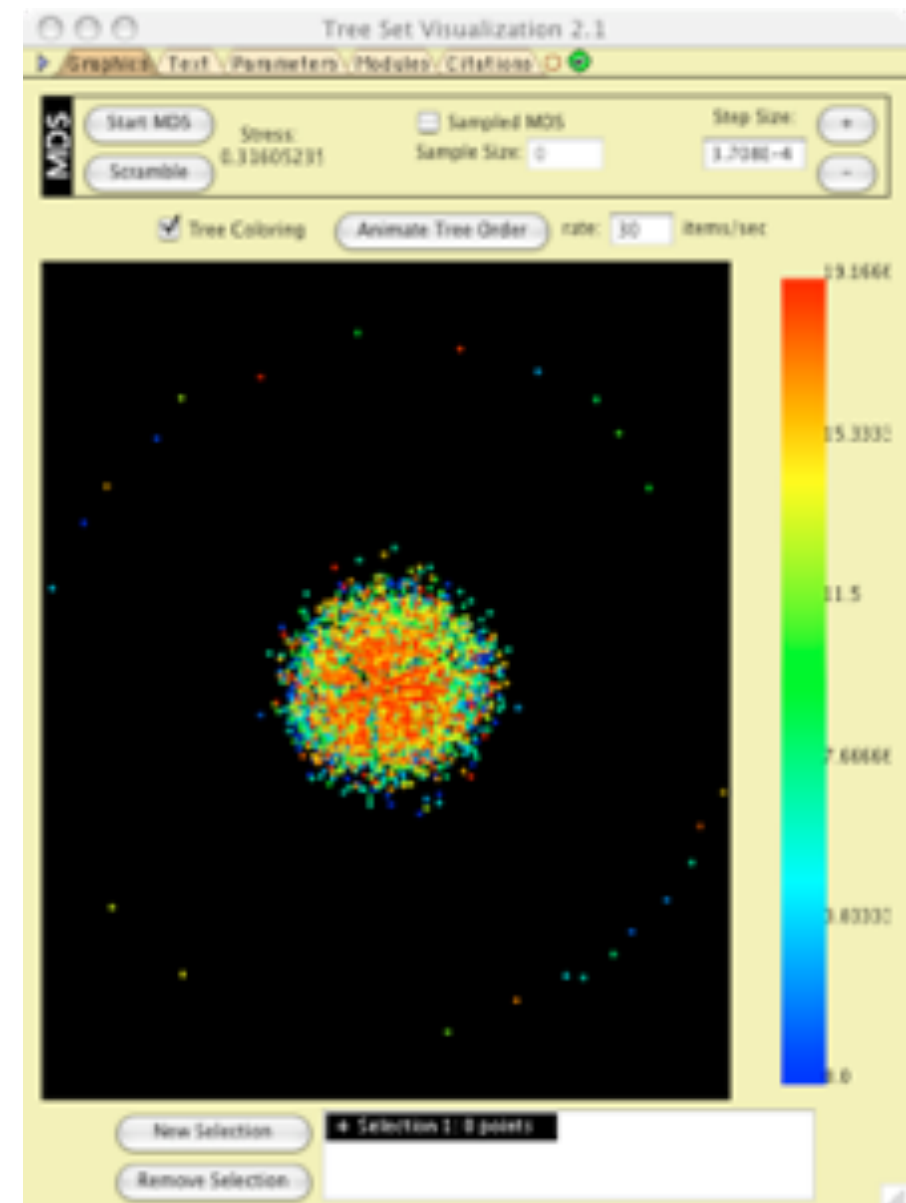
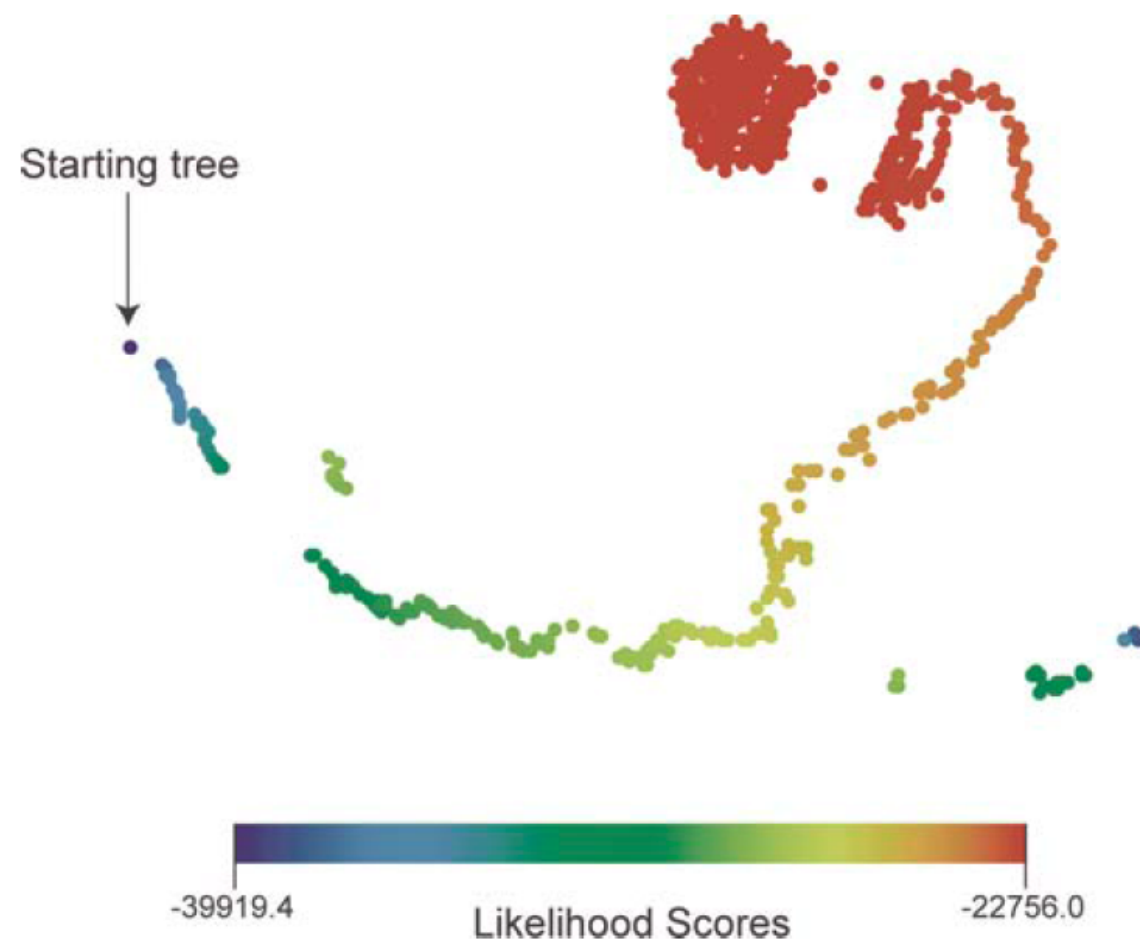
Multiple Analysis Diagnostics

Bipartition Probabilities



Multiple Analysis Diagnostics

Take a look at tree space

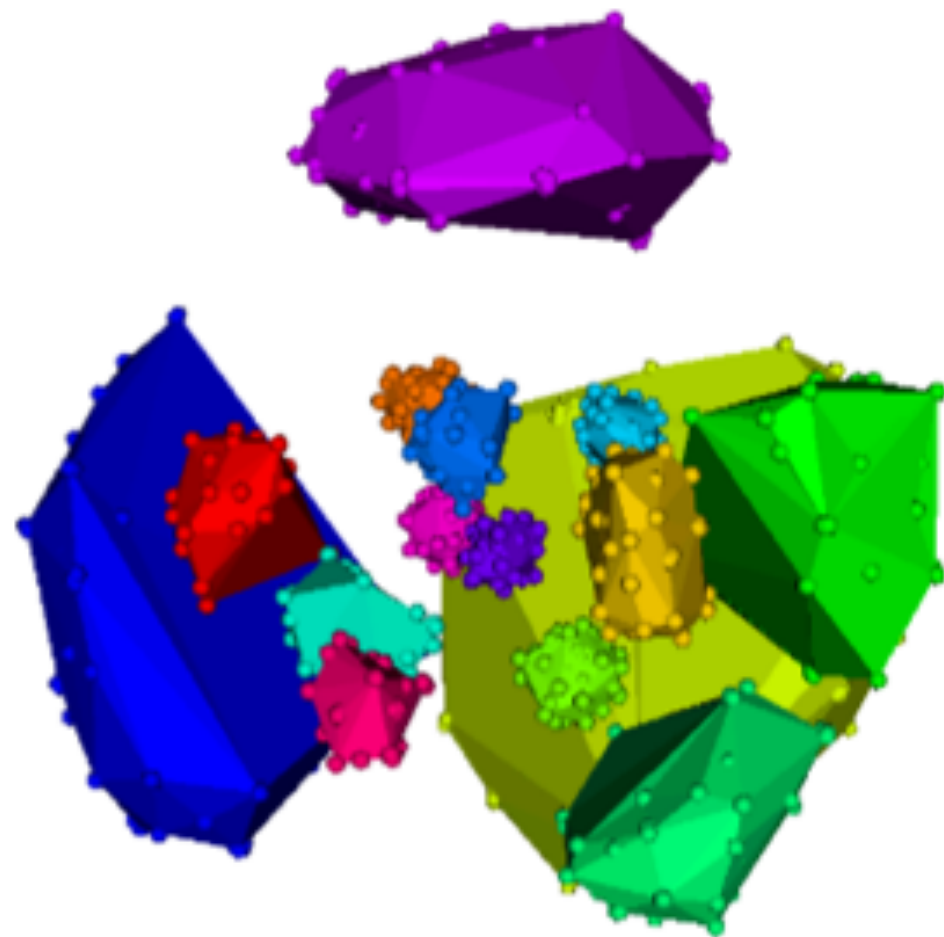


Multiple Analysis Diagnostics

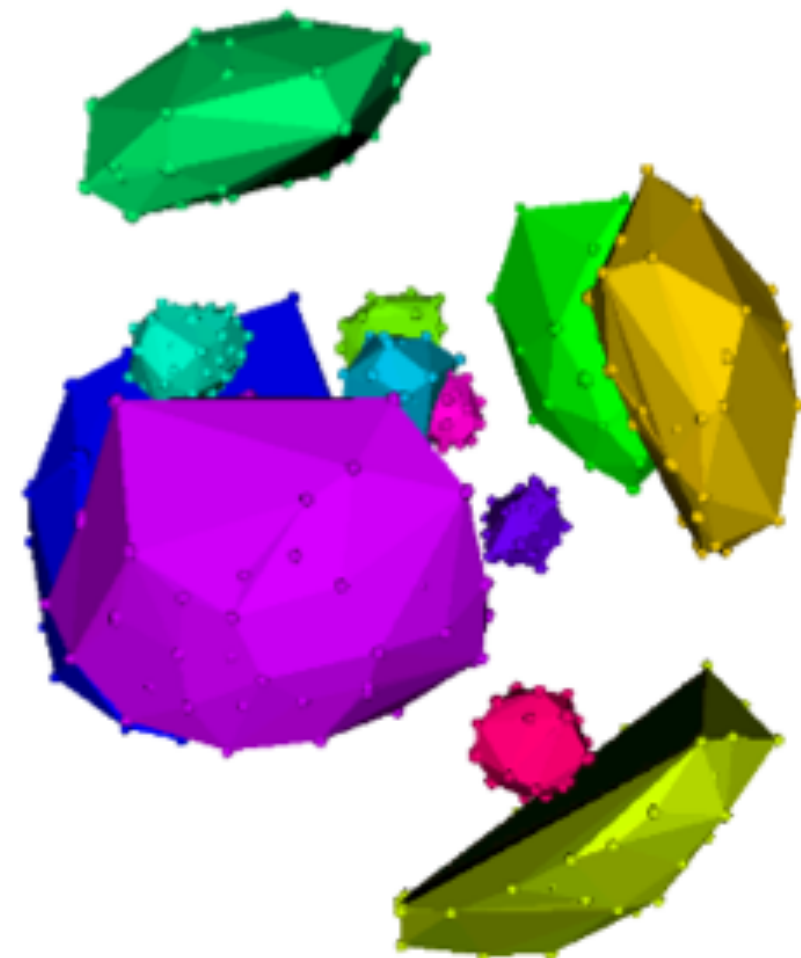
Take a look at tree space

Colors = Trees from Different Genes

Salamanders



Mammals



Software Tools

Convergence Stats for Scalars

Tracer

Bonsai (R - w/ RevBayes)

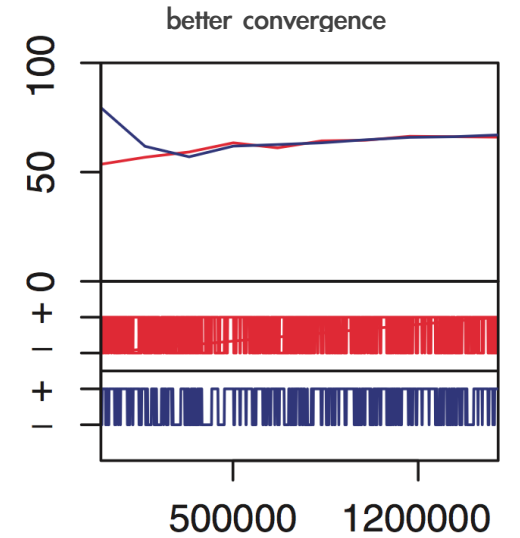
boa, coda (R - generic MCMC)

Clade-Specific Convergence

AWTY

RWTY (R)

Bonsai (R - w/ RevBayes)



Tree Space Projection

TreeScaper

TreeSetViz (Mesquite)

Treespace (R)

