Simulation Results

Simulation: Epoch linear preferential sampling *(24 epochs – 2 months/epoch)*

Inference: BESP *(24 epochs ~ 2 months/epoch)*

Louis du Plessis

Last modified: 22 Nov 2019

Contents

1 Convergence

1.1 Constant-size

100/100 simulations converged with ESS of all parameters > 200.

1.2 Bottleneck

100/100 simulations converged with ESS of all parameters > 200.

1.3 Boom-bust

100/100 simulations converged with ESS of all parameters > 200.

1.4 Cyclical boom-bust

100/100 simulations converged with ESS of all parameters > 200.

1.5 Logistic growth and decline

100/100 simulations converged with ESS of all parameters > 200.

2 Summary statistics

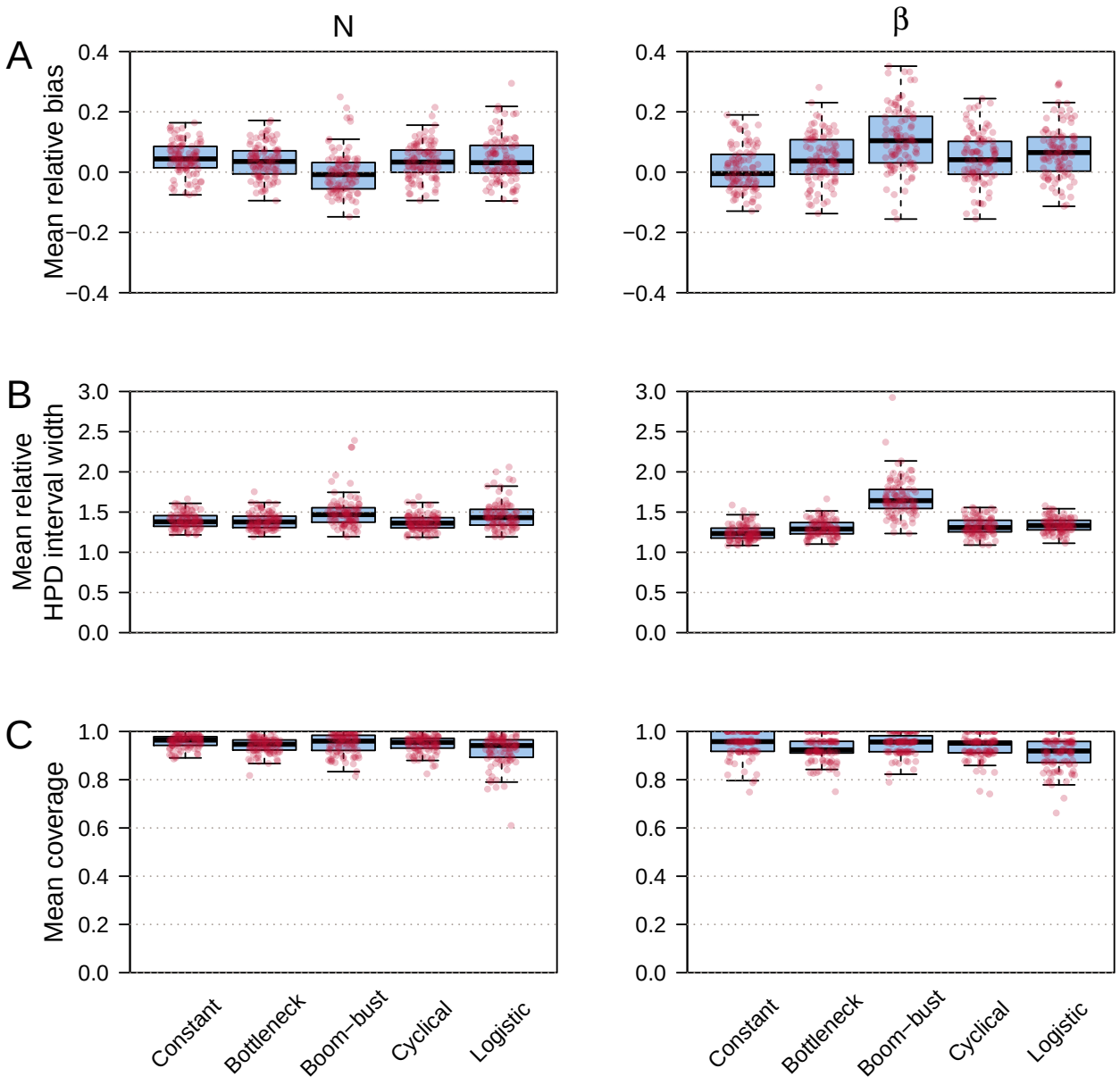

Figure 1: **Sampling period only!** Boxplots and stripcharts displaying measures of the statistical performance of the BESP, evaluated on trees simulated under five different demographic models (constant, bottleneck, boom-bust, cyclical boom-bust, logistic growth and decline). We simulated 100 replicate trees for each scenario. Three measures of estimator performance are shown (A) mean relative bias, (B) mean relative HPD interval size, and (C) mean coverage. The left and right columns illustrate estimation performance for effective population size (N) and sampling intensity (beta), respectively.

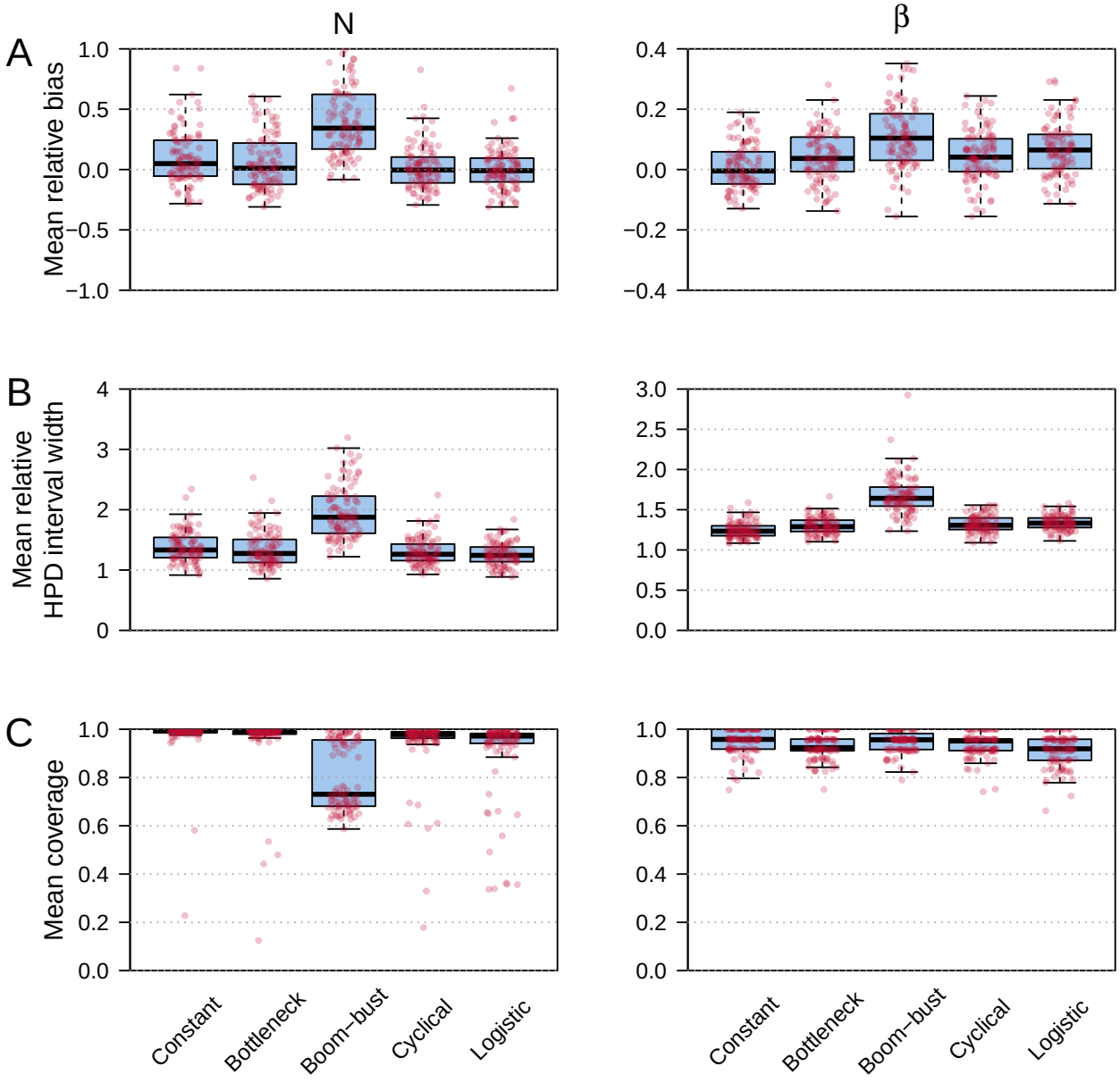

Figure 2: **TMRCA to present!** Boxplots and stripcharts displaying measures of the statistical performance of the BESP, evaluated on trees simulated under five different demographic models (constant, bottleneck, boom-bust, cyclical boom-bust, logistic growth and decline). We simulated 100 replicate trees for each scenario. Three measures of estimator performance are shown (A) mean relative bias, (B) mean relative HPD interval size, and (C) mean coverage. The left and right columns illustrate estimation performance for effective population size (N) and sampling intensity (beta), respectively.

3 Example results

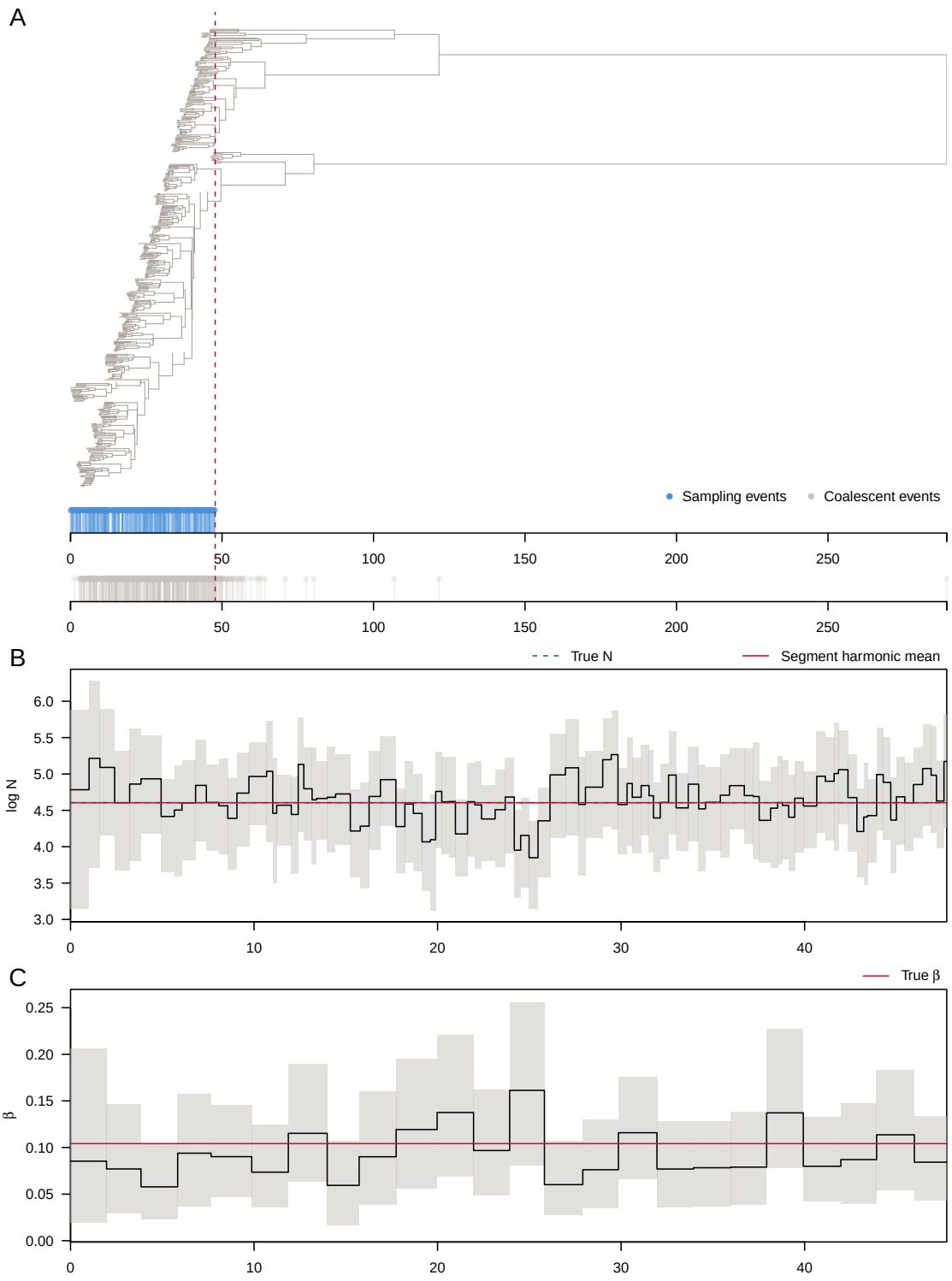

Figure 3: **Constant** *(best ranked simulation, replicate 2)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

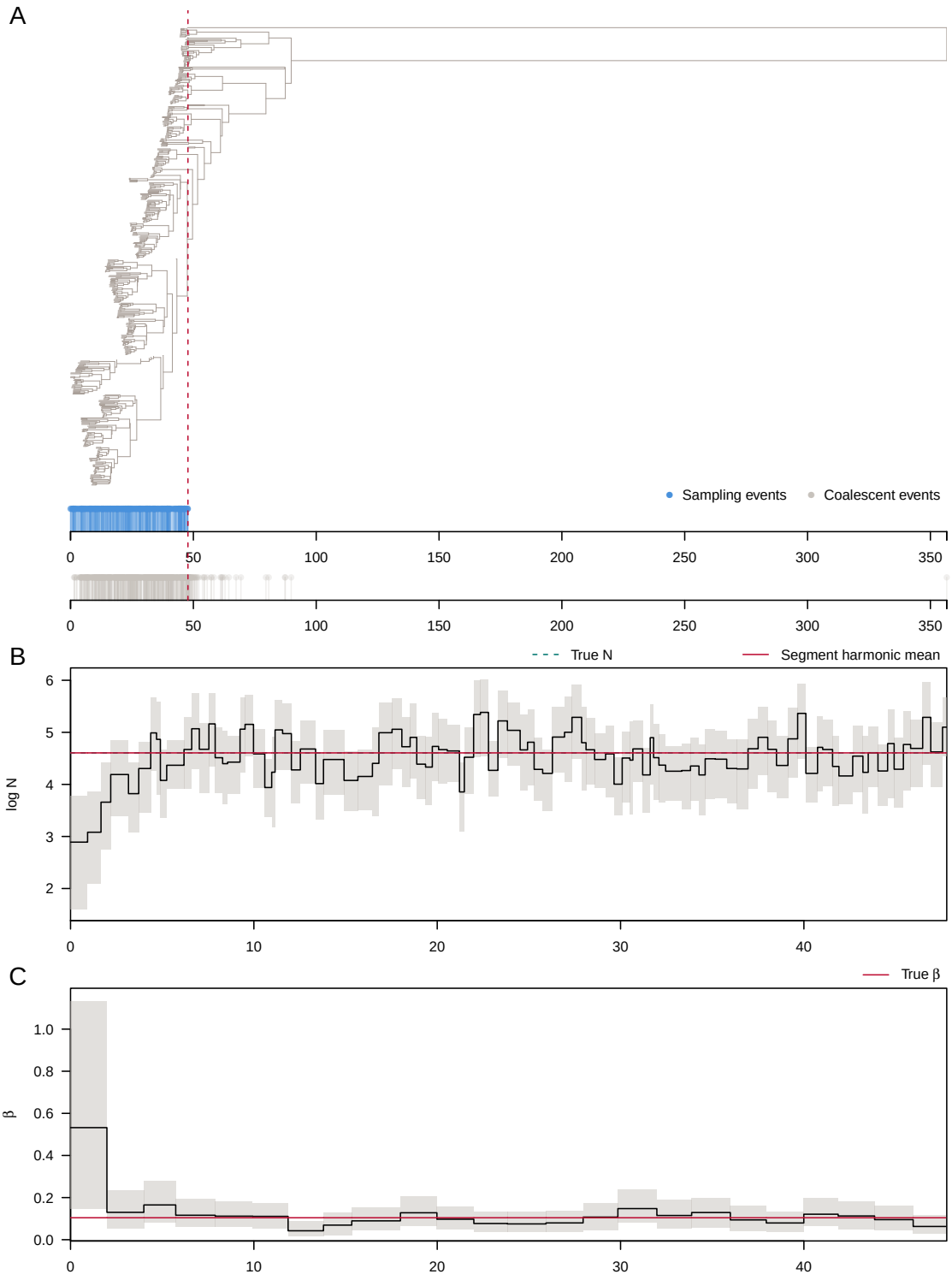

Figure 4: **Constant** *(median ranked simulation, replicate 1)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

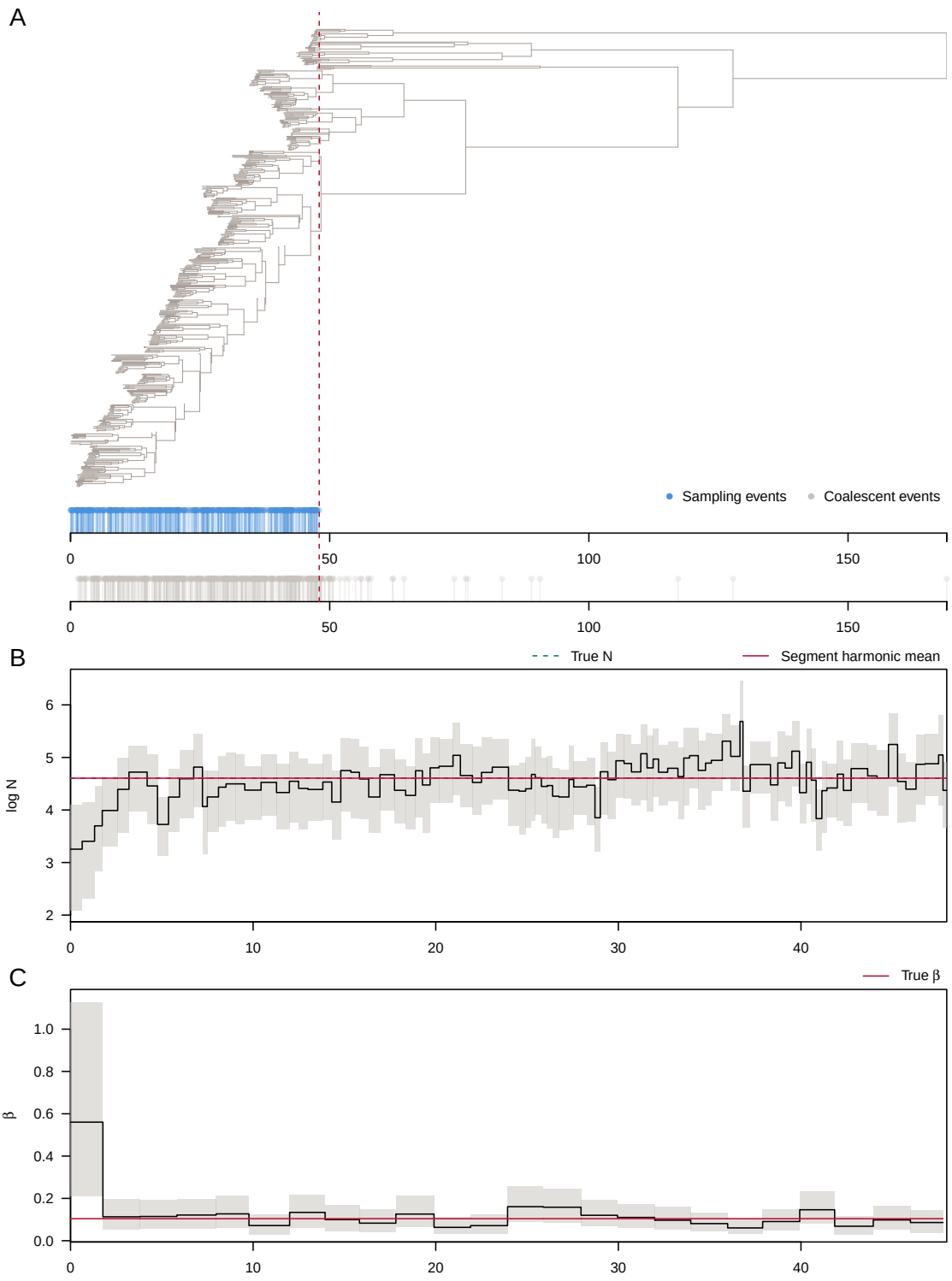

Figure 5: **Constant** *(worst ranked simulation, replicate 68)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

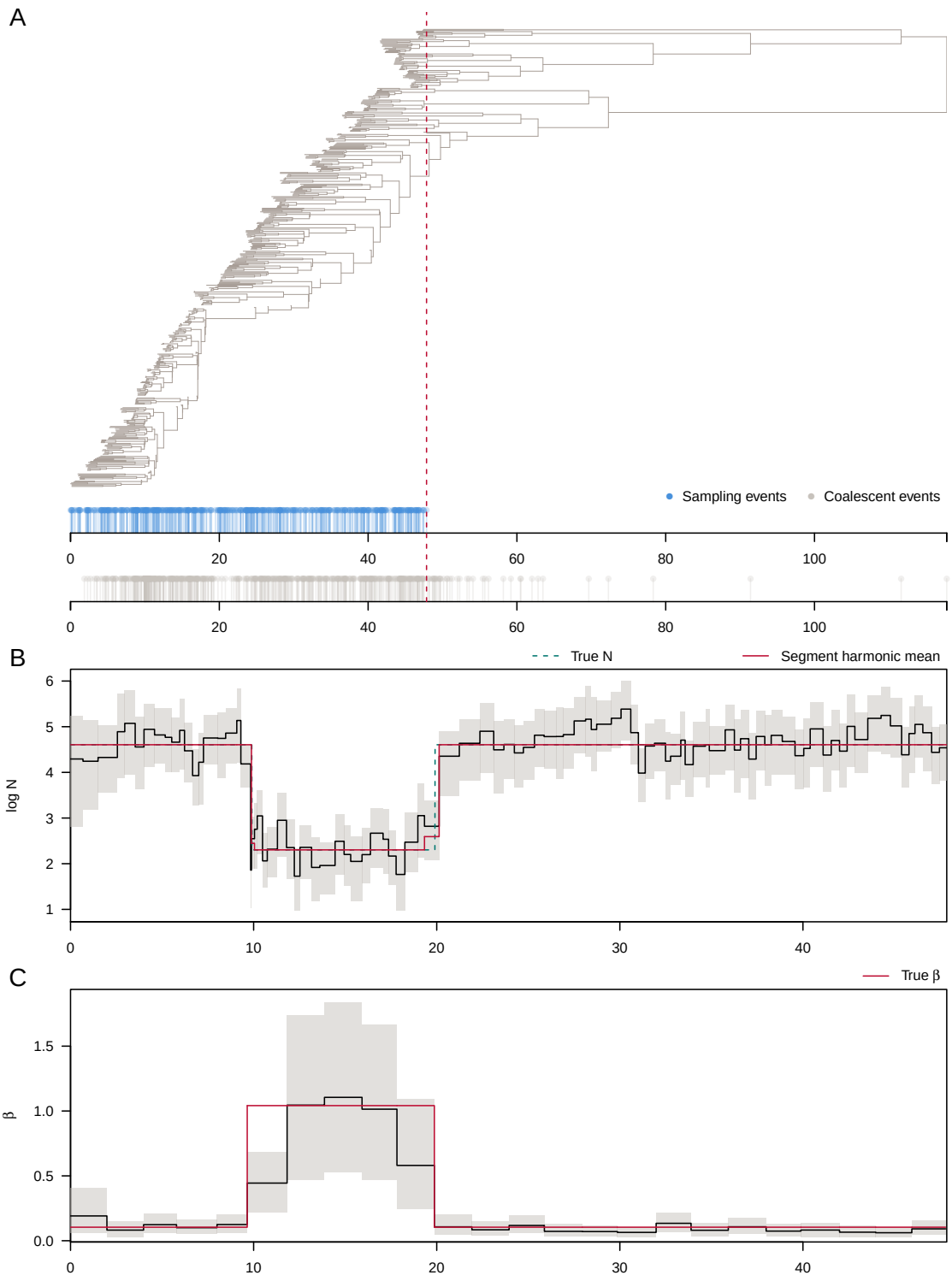

Figure 6: **Bottleneck** *(best ranked simulation, replicate 78)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

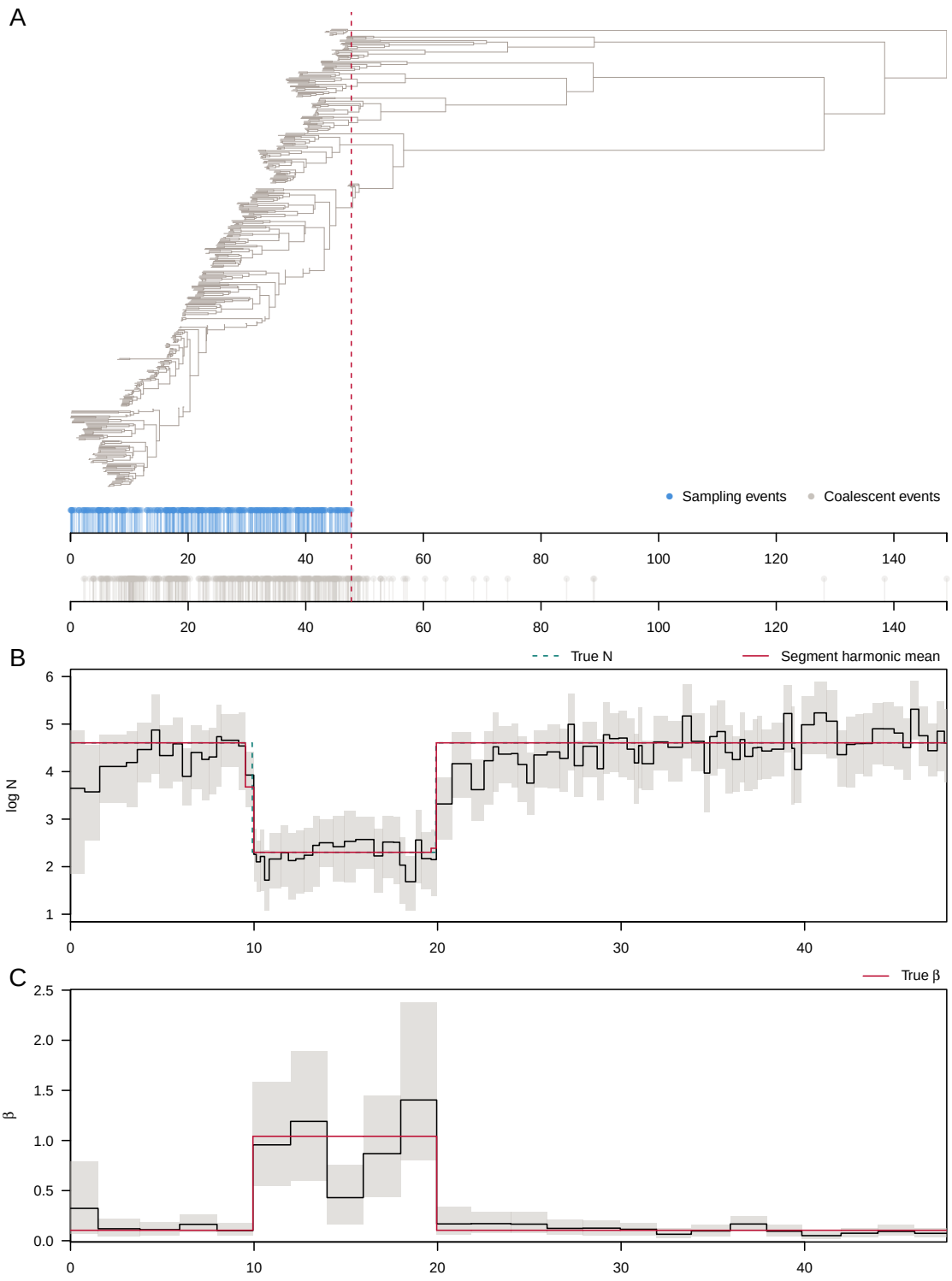

Figure 7: **Bottleneck** *(median ranked simulation, replicate 0)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

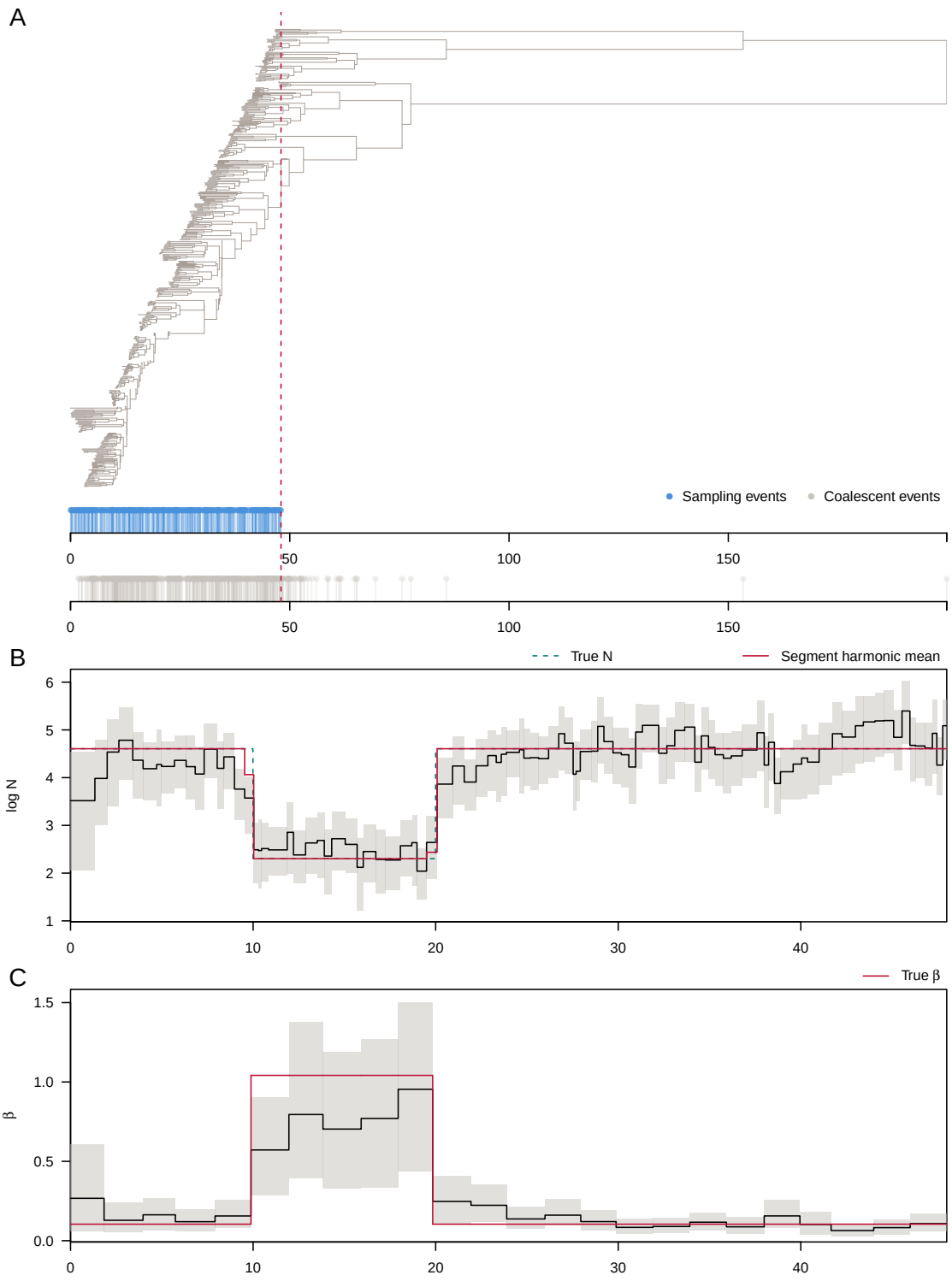

Figure 8: **Bottleneck** *(worst ranked simulation, replicate 73)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

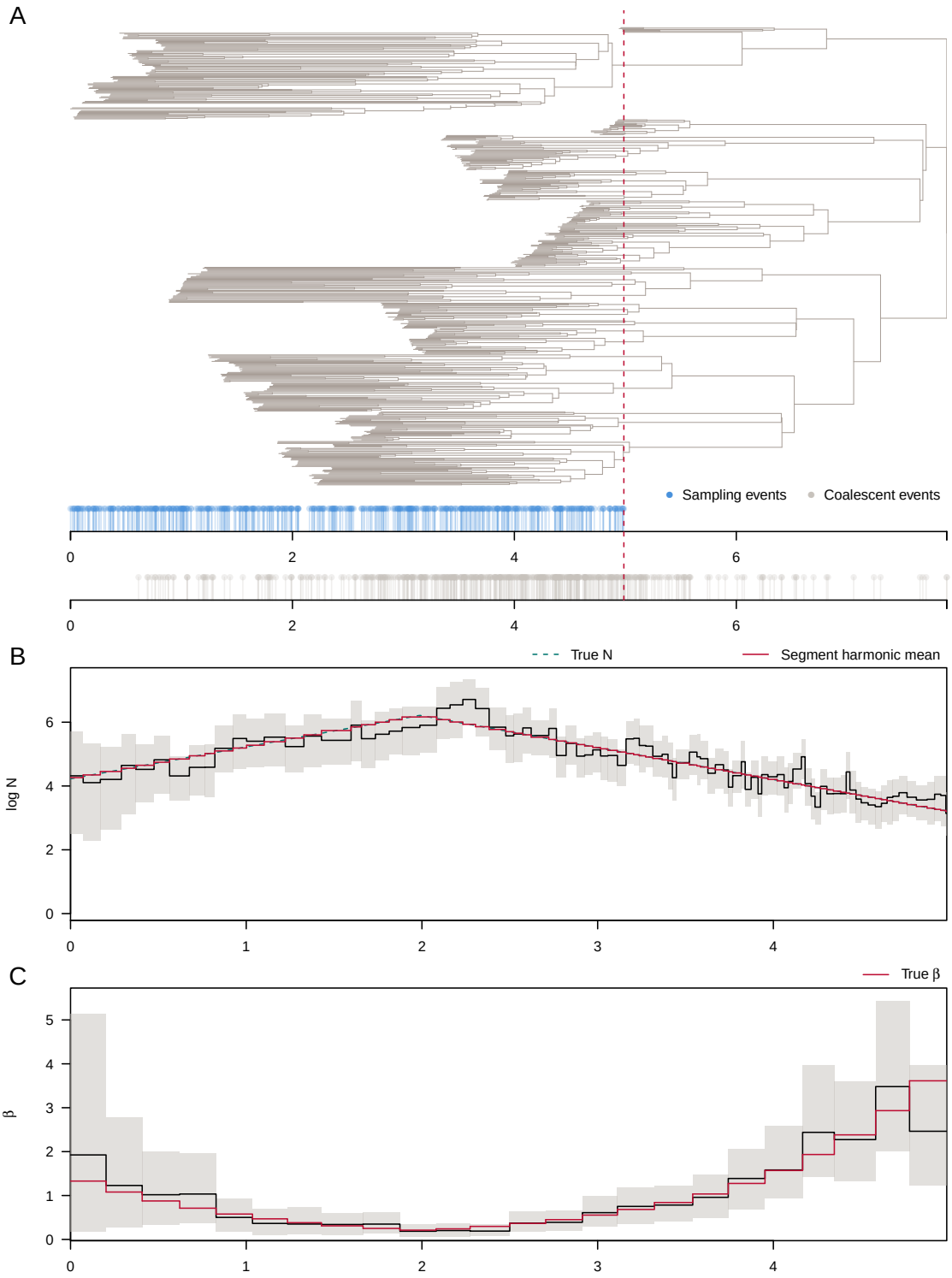

Figure 9: **Boom-bust** *(best ranked simulation, replicate 17)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

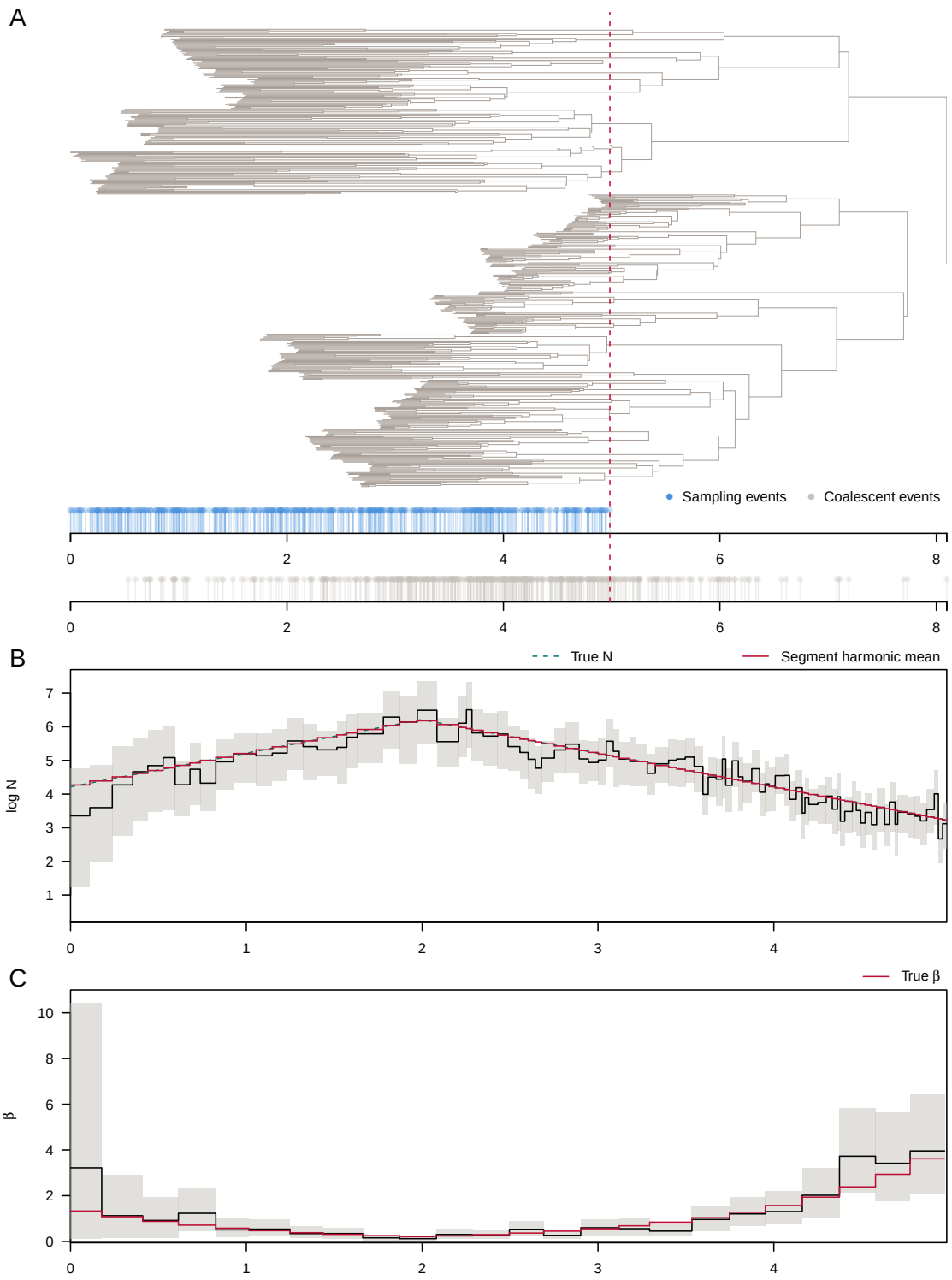

Figure 10: **Boom-bust** *(median ranked simulation, replicate 0)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

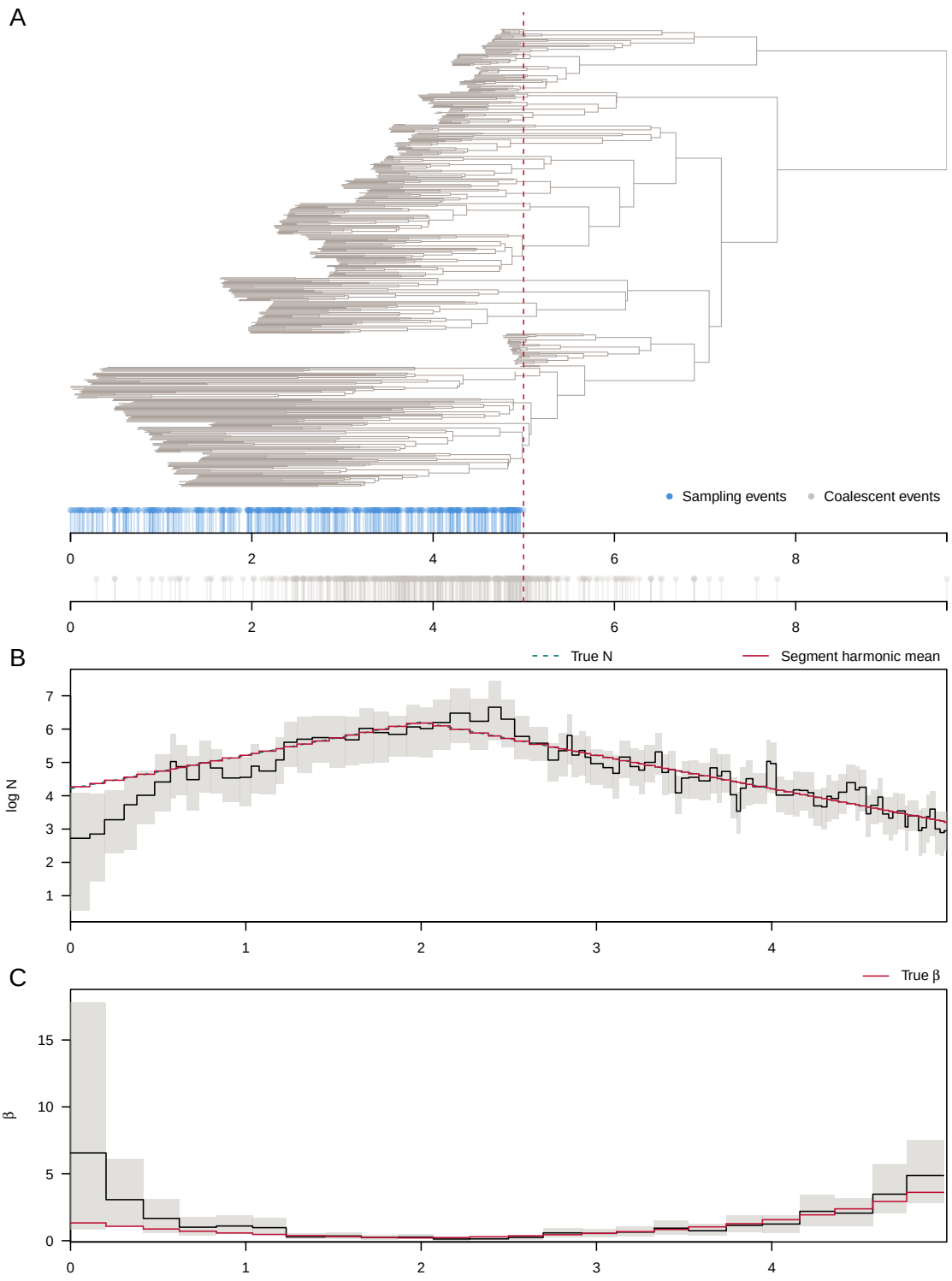

Figure 11: **Boom-bust** *(worst ranked simulation, replicate 40)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

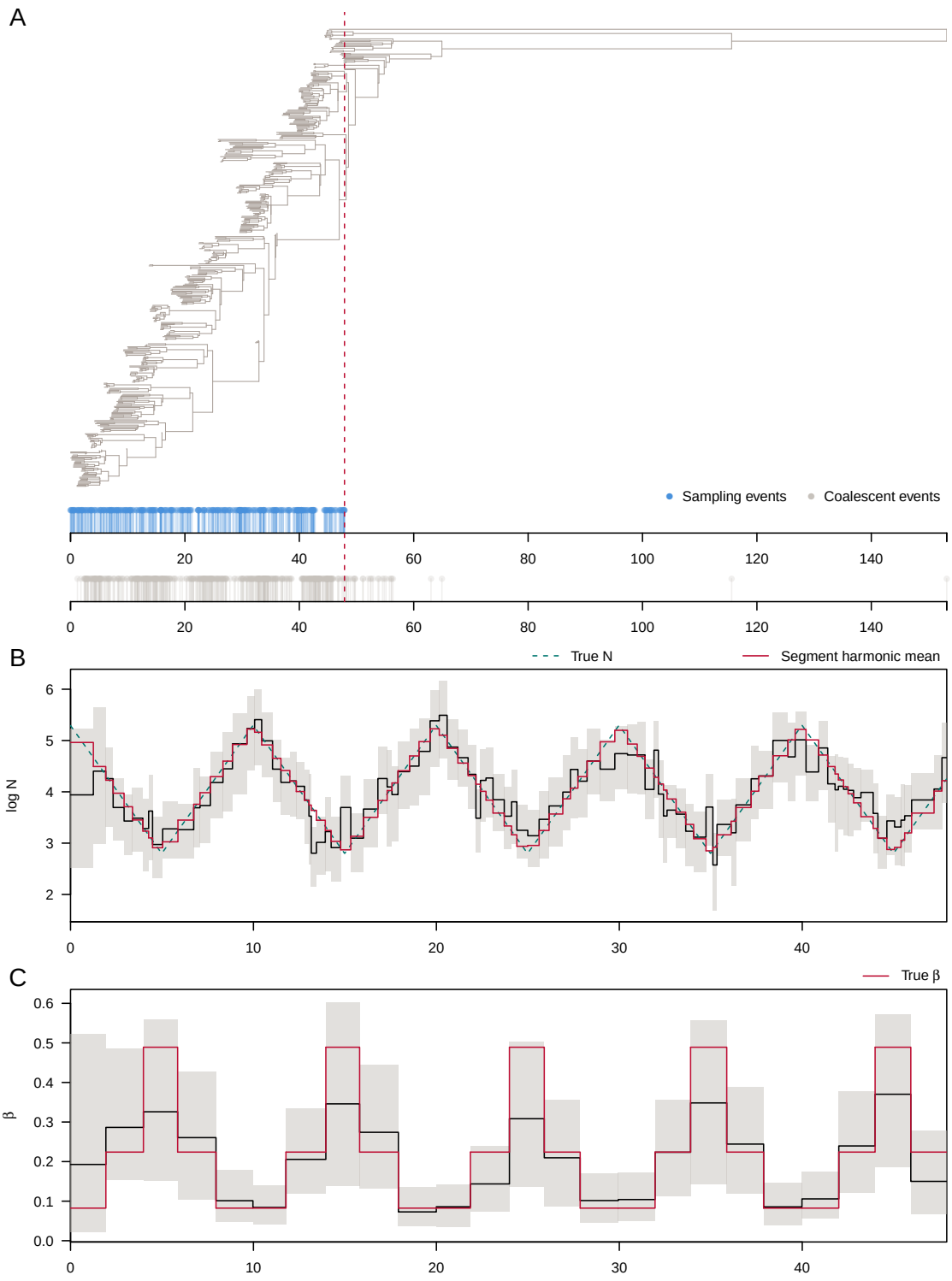

Figure 12: **Cyclical** *(best ranked simulation, replicate 94)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

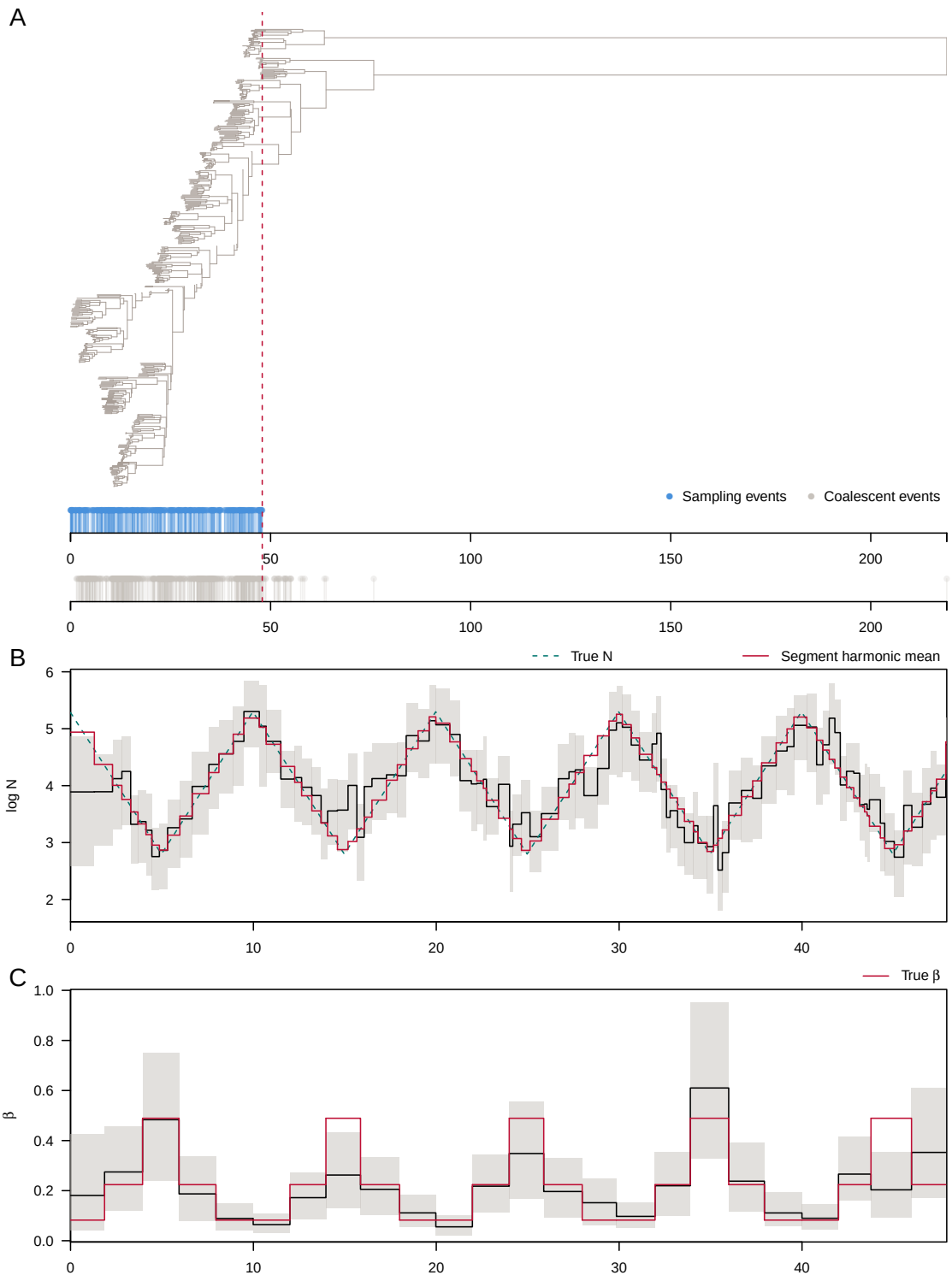

Figure 13: **Cyclical** *(median ranked simulation, replicate 1)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

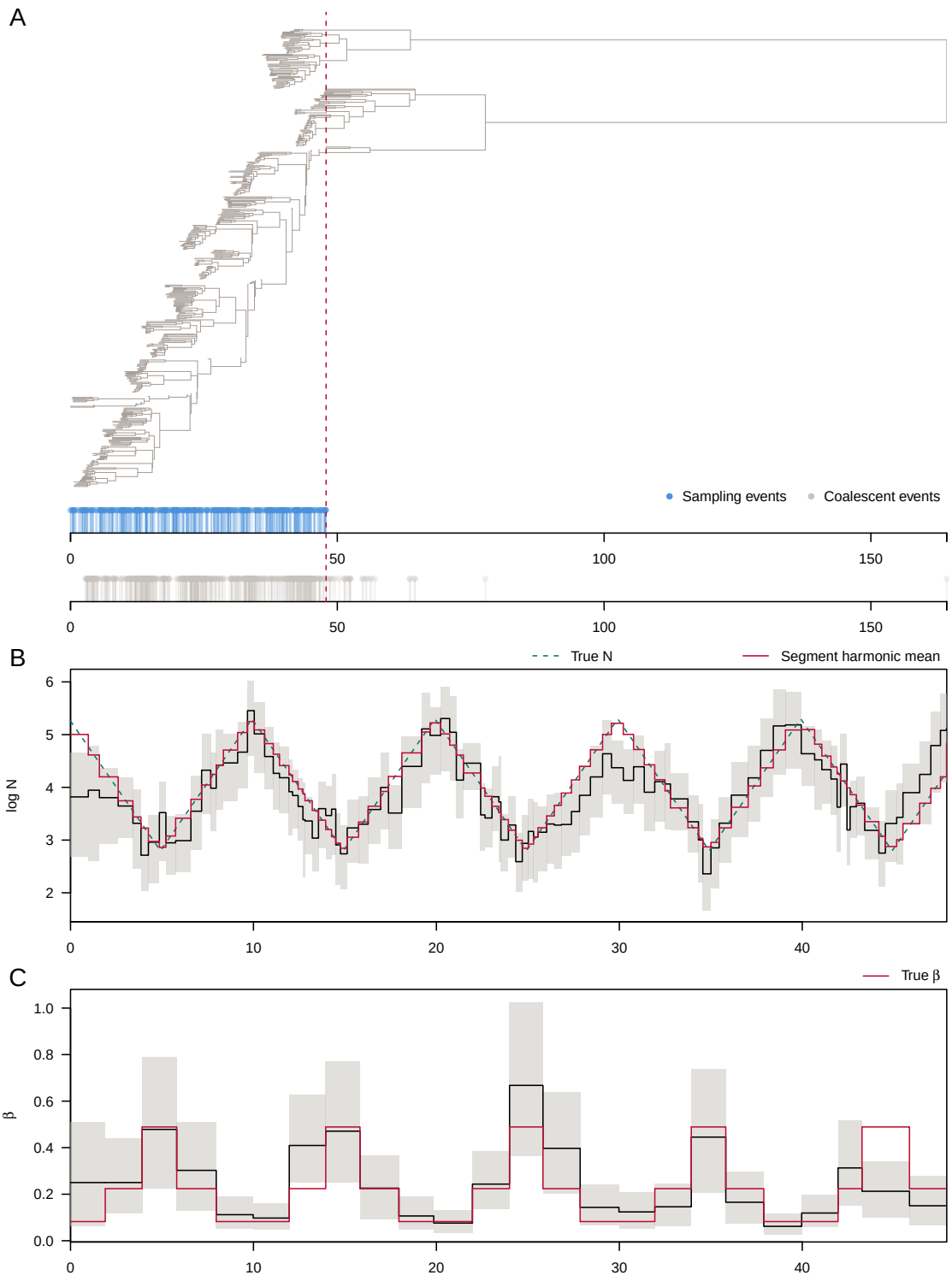

Figure 14: **Cyclical** *(worst ranked simulation, replicate 51)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

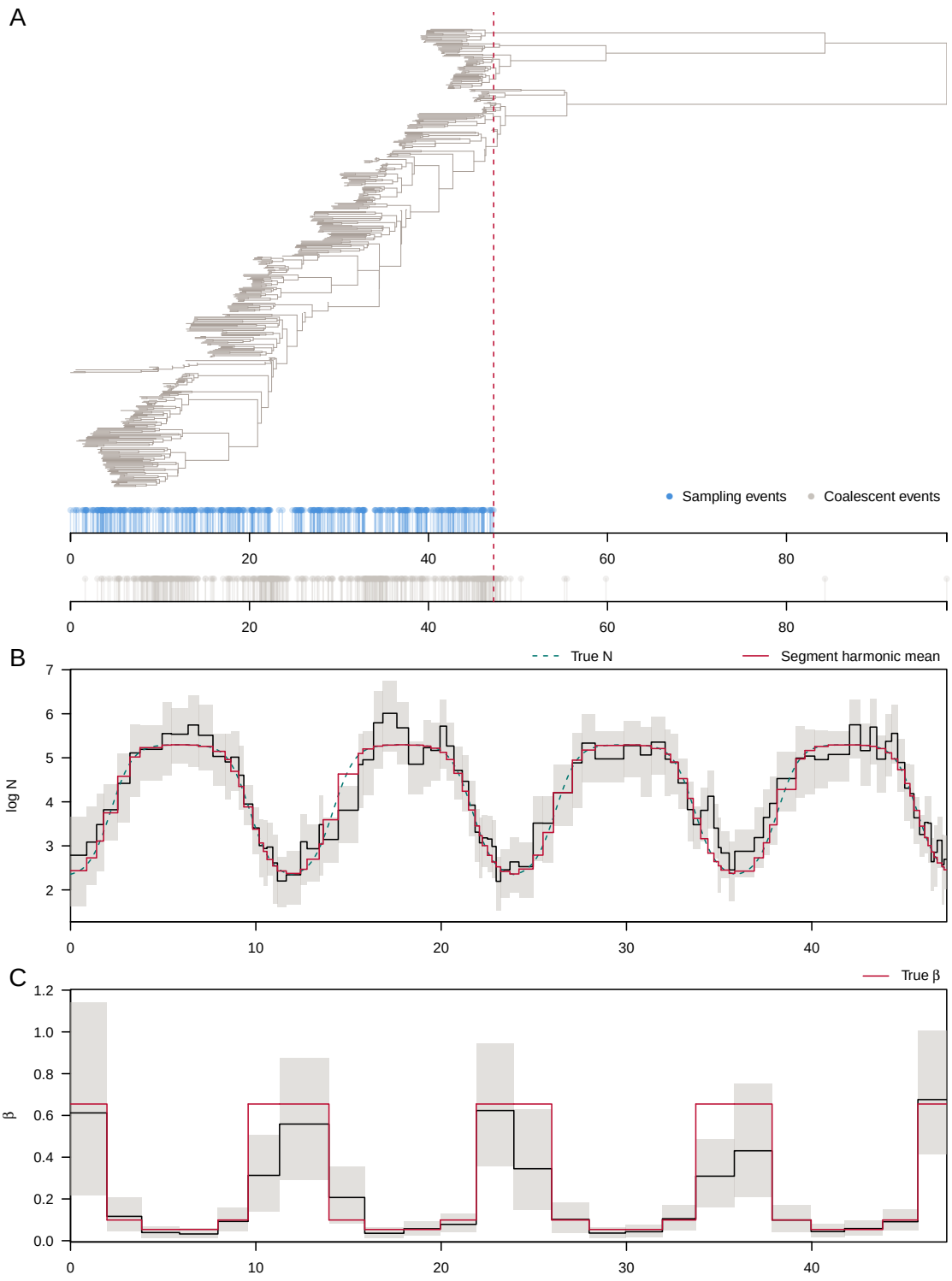

Figure 15: **Logistic** *(best ranked simulation, replicate 38)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

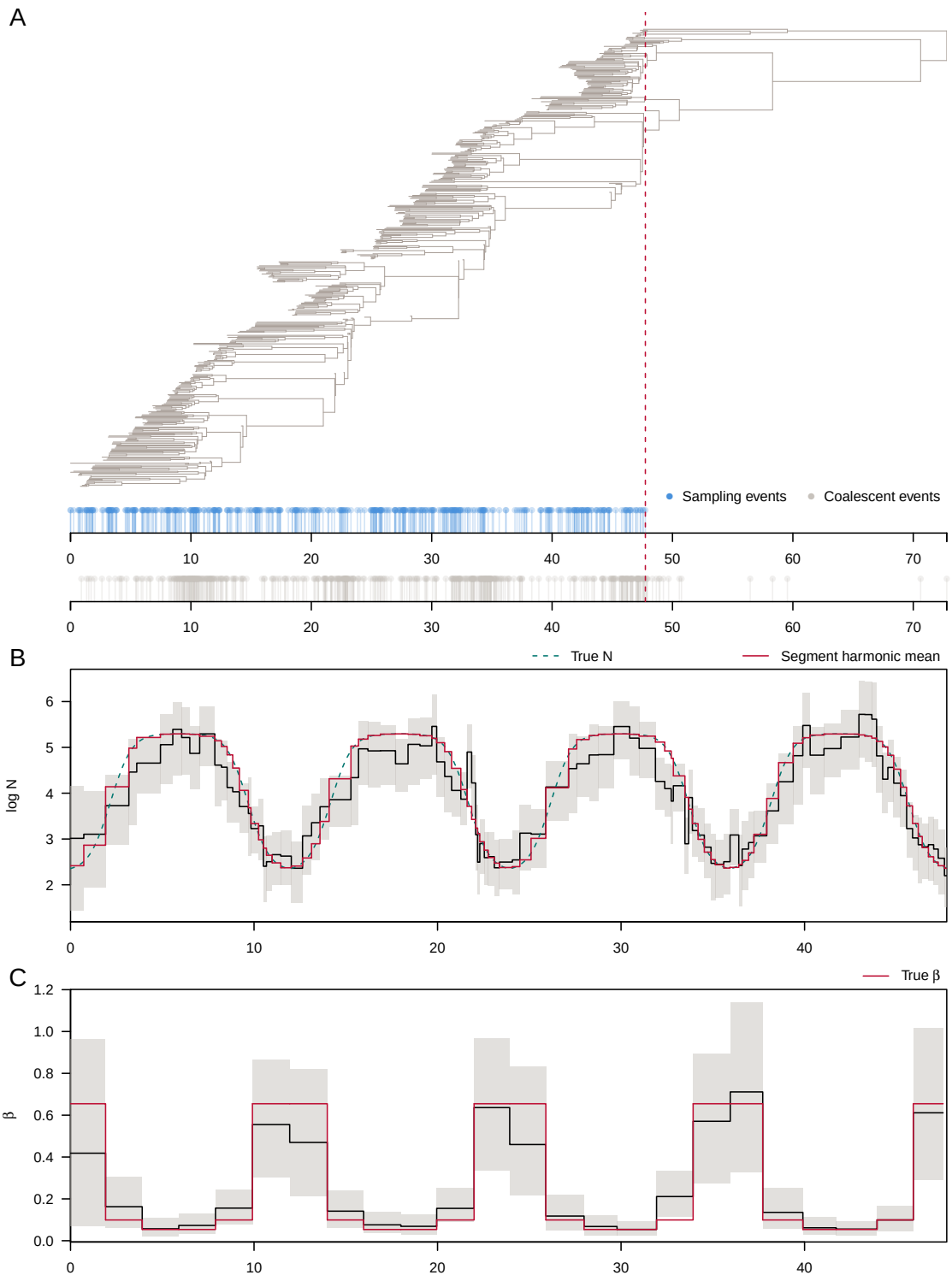

Figure 16: **Logistic** *(median ranked simulation, replicate 3)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

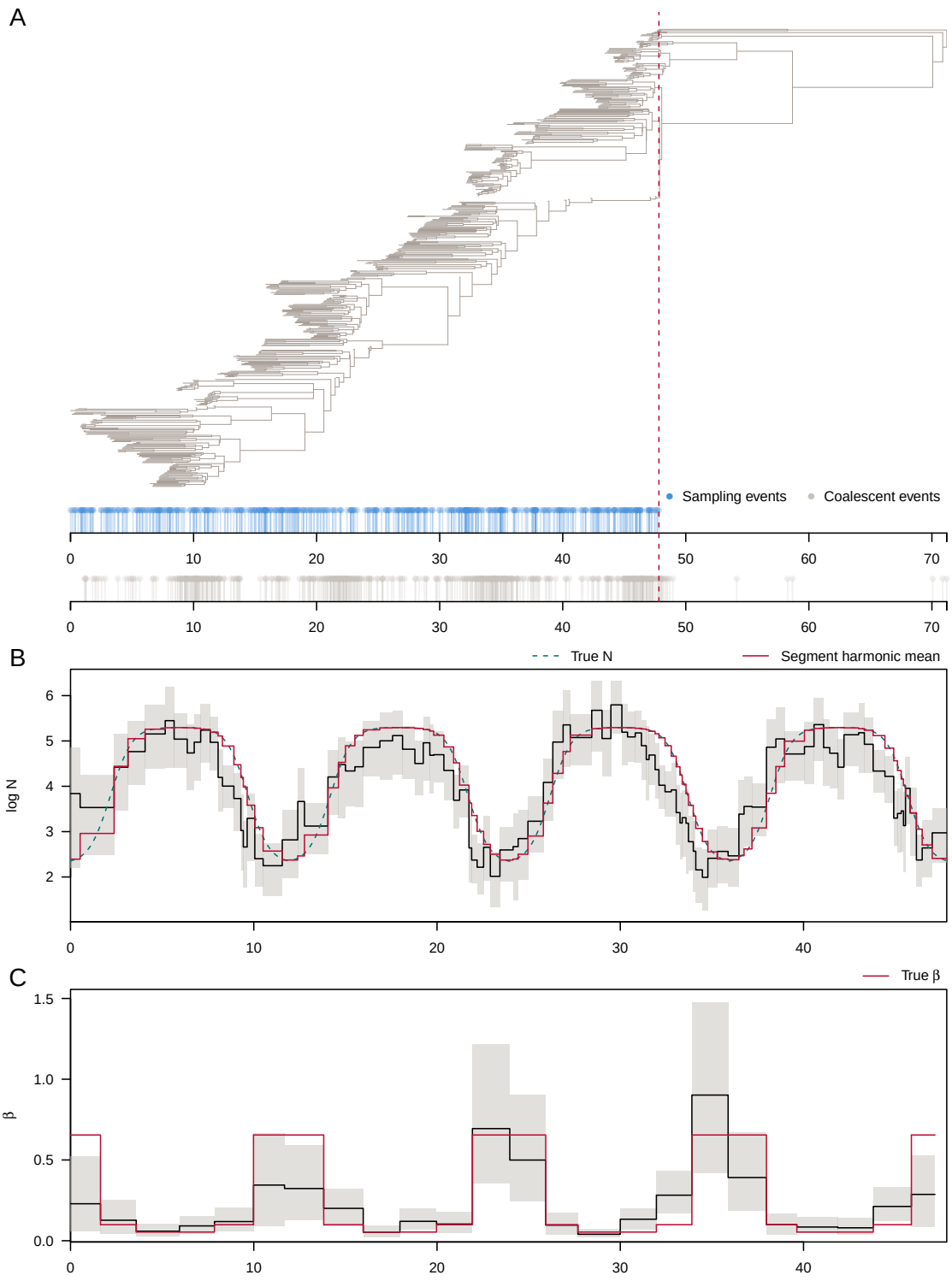

Figure 17: **Logistic** *(worst ranked simulation, replicate 70)*: (A) Simulated tree with sampling (blue) and coalescent (grey) events. (B) Median (solid black line) and HPD intervals (shaded areas) for the N-estimates between the most recent and most ancient samples. The dashed green line shows the true N-trajectory used to simulate the tree and the red line the harmonic mean of the true N during each segment. (C) Median (solid black line) and HPD intervals (shaded areas) for the sampling intensity (beta) estimates for each epoch. The red line shows the true beta used to simulate the tree.

4 Session info

```
## R version 3.5.1 (2018-07-02)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Sierra 10.12.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_GB.UTF-8/en_GB.UTF-8/en_GB.UTF-8/C/en_GB.UTF-8/en_GB.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] ape_5.3        coda_0.19-3    beastio_0.2.5  phylodyn_0.9.0
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.2         codetools_0.2-16   lattice_0.20-38
##  [4] digest_0.6.21      grid_3.5.1         nlme_3.1-141
##  [7] magrittr_1.5       evaluate_0.14      highr_0.8
## [10] rlang_0.4.0        stringi_1.4.3      rmarkdown_1.16
## [13] RColorBrewer_1.1-2 tools_3.5.1        stringr_1.4.0
## [16] parallel_3.5.1     xfun_0.10          yaml_2.2.0
## [19] compiler_3.5.1     htmltools_0.4.0    knitr_1.25
```