

# H3N2 HA (trimmed) results

*637 HA sequences sampled between 1993.06 and 2005.25 (1,698bp).*

*Louis du Plessis*

*Last modified: 22 Nov 2019*

## Contents

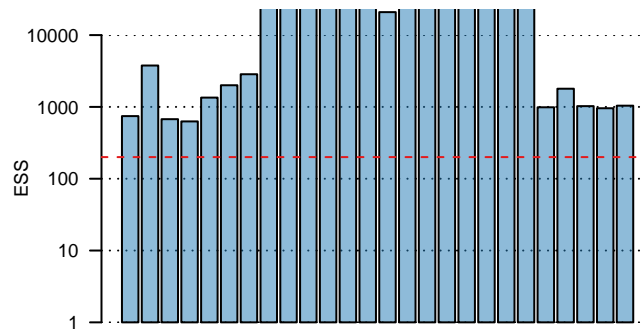
<b>1</b>	<b>Summary</b>	<b>1</b>
<b>2</b>	<b>Combine chains and check convergence</b>	<b>2</b>
2.1	BSP 40 . . . . .	2
2.2	BESP 40/1 . . . . .	3
2.3	BESP 40/12 . . . . .	4
2.4	Convergence summary . . . . .	5
<b>3</b>	<b>Figures</b>	<b>5</b>
<b>4</b>	<b>Session info</b>	<b>14</b>

## 1 Summary

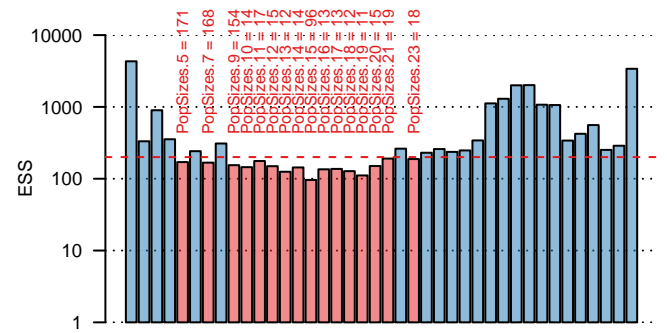
BESP results from analysing 637 HA sequences sampled between 1993.06 and 2005.25 (1,698bp).

## 2 Combine chains and check convergence

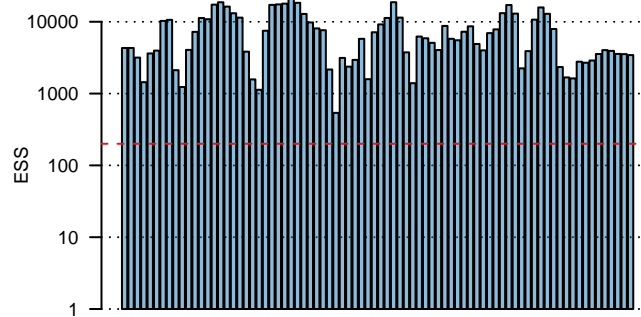
### 2.1 BSP 40



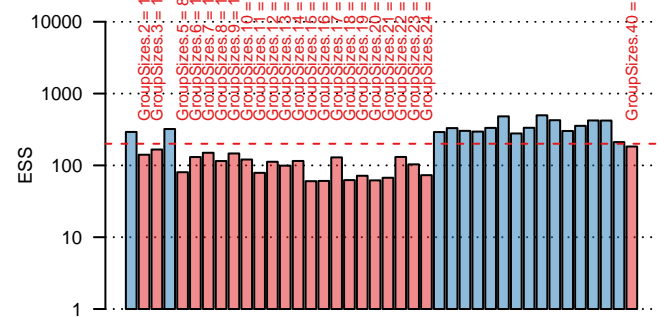
General parameters



PopSize parameters

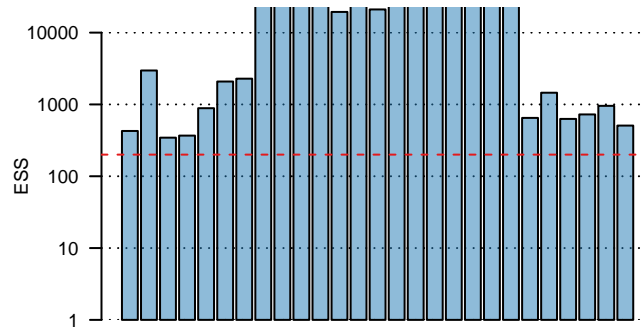


Gridded PopSize parameters

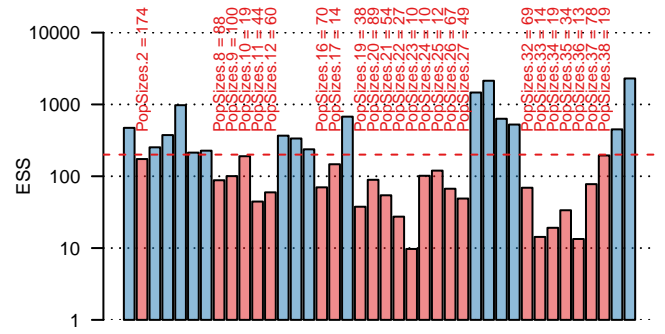


GroupSize parameters

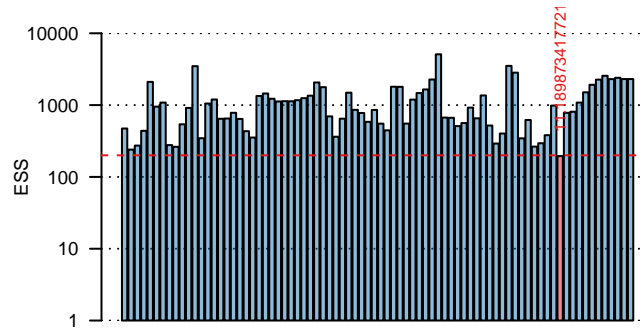
## 2.2 BESP 40/1



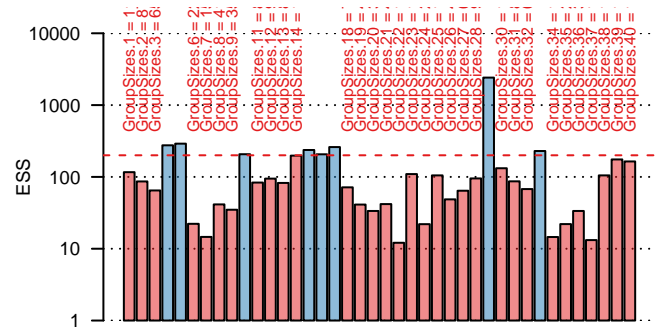
General parameters



PopSize parameters

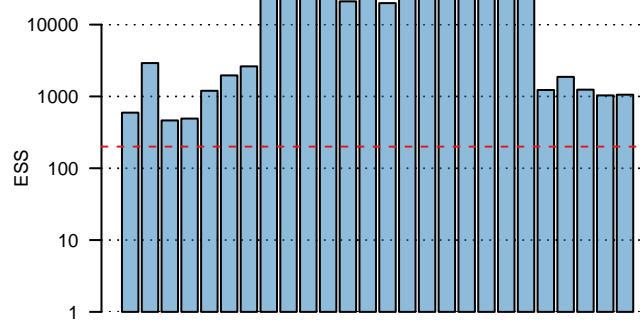


Gridded PopSize parameters

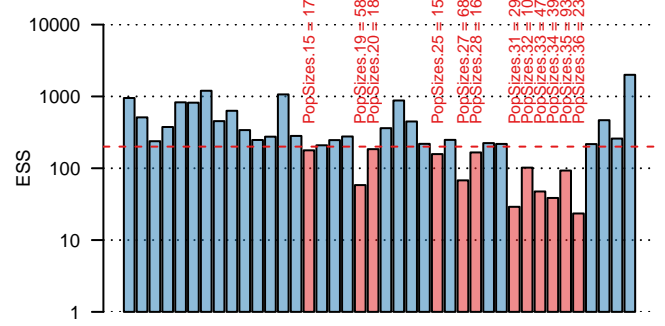


GroupSize parameters

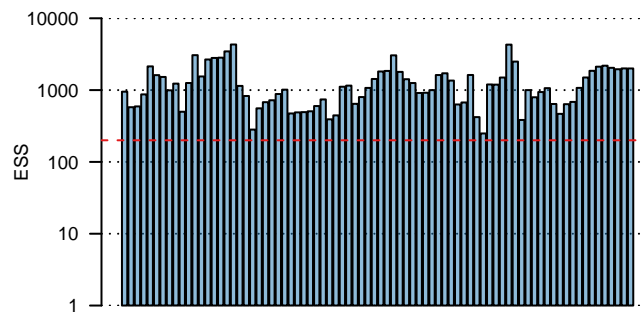
## 2.3 BESP 40/12



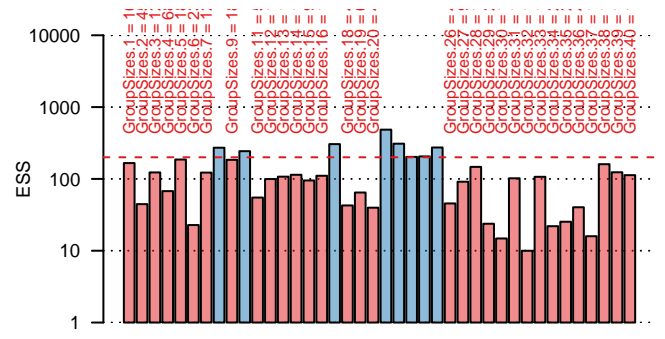
General parameters



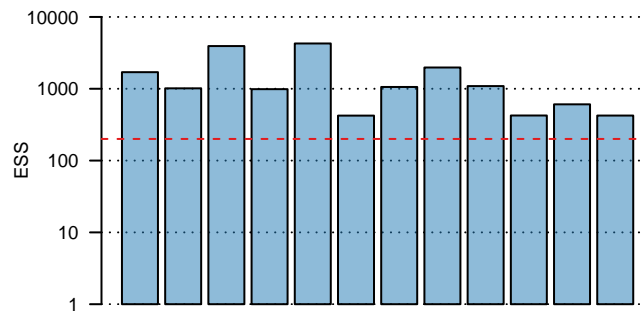
PopSize parameters



Gridded PopSize parameters



GroupSize parameters



Sampling intensity parameters

## 2.4 Convergence summary

Table 1: Number of nonstationary parameters for different models and subsets of parameters.

	BSP 40	BESP 40/1	BESP 40/12
PopSize	16	24	12
Gridded PopSize	0	1	0
Other parameters	0	0	0

Table 2: Lowest ESS values for different models and subsets of parameters.

	BSP 40	BESP 40/1	BESP 40/12
PopSize	96.29	9.73	23.46
Gridded PopSize	-	194.93	-
Other parameters	-	-	-

For checking convergence and combining chains a burn-in of 30% was used, and chains are thinned by a factor of 3, effectively sampling from the MCMC chains every  $3 \times 10^4$  states and resulting in chains with 32676 posterior samples.

For convergence only the general parameters and the gridded population size parameters are important. The logged population size and group size parameters are not always expected to mix well.

## 3 Figures

- BESP 40/12  $\beta$ : 6.03, 14.87, 27.19 (mean of median and 95% HPD)
- BESP 40/1  $\beta$ : 8.38, 11.16, 14.32 (median and 95% HPD)

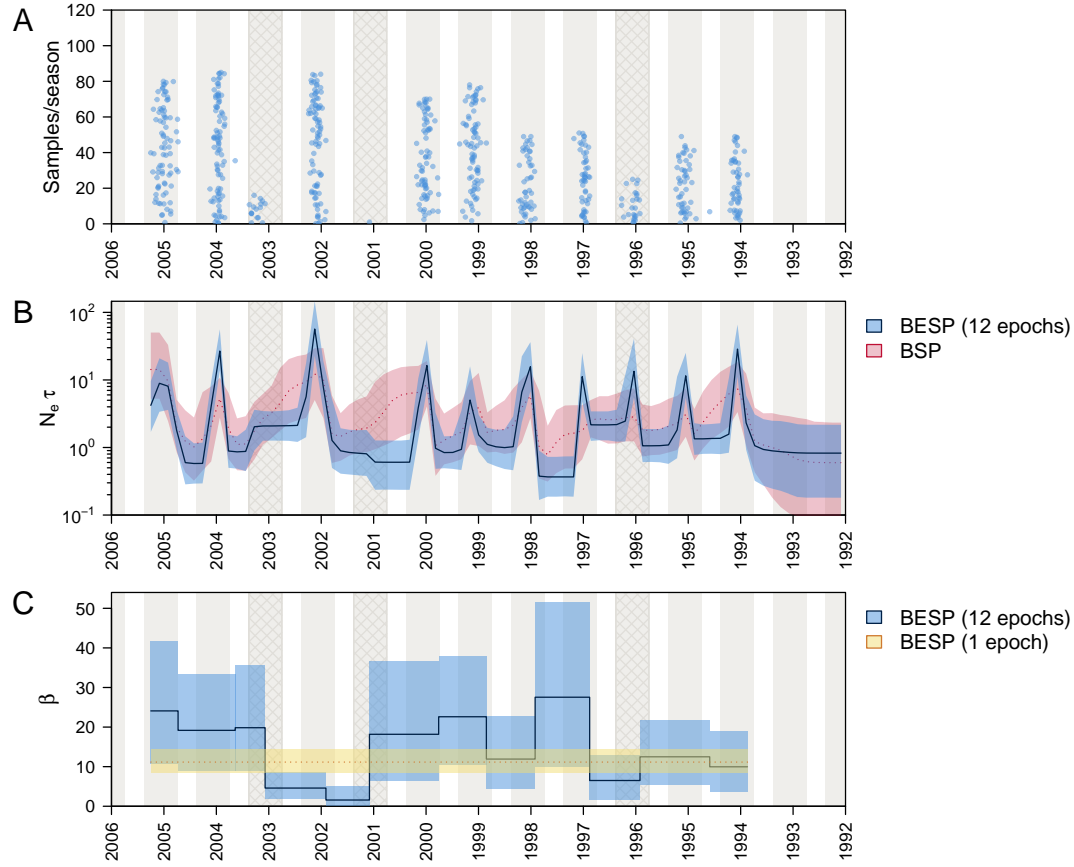


Figure 1: (A) Density of sequence sampling dates through time for the alignment of 637 A/H3N2 HA sequences from NY state that we analysed. Blue dots indicate stripcharts of individual samples for each season. The stripchart heights give the number of samples in each season. Grey shading indicates the approximate period of influenza observation in New York state during each season (epidemiological week 40, to week 20 in the next year). Cross-hatched seasons are those where A/H3N2 was not the dominant influenza virus subtype. (B) Median (solid/dotted line) and 95% highest posterior density (HPD) intervals (shaded areas) for the genetic diversity estimates ( $N_e \tau$ ) through time. The BESP estimate is shown in blue and the BSP estimate is in red. (C) Median (solid line/dotted line) and 95% HPD intervals (shaded areas) of the estimated sampling intensities (beta) for each sampling epoch. The 12-epoch BESP estimates are shown in blue and a single-epoch (density-defined) estimate is in yellow.

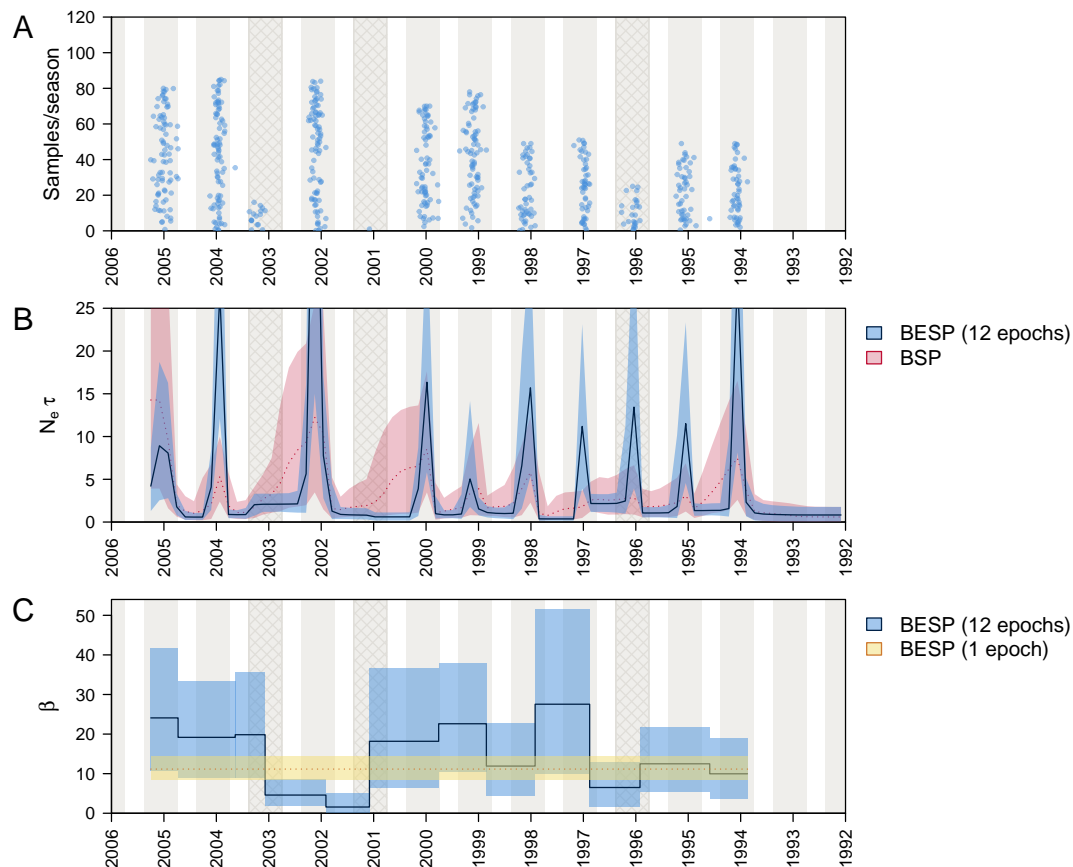


Figure 2: (A) Density of sequence sampling dates through time for the alignment of 637 A/H3N2 HA sequences from NY state that we analysed. Blue dots indicate stripcharts of individual samples for each season. The stripchart heights give the number of samples in each season. Grey shading indicates the approximate period of influenza observation in New York state during each season (epidemiological week 40, to week 20 in the next year). Cross-hatched seasons are those where A/H3N2 was not the dominant influenza virus subtype. (B) Median (solid/dotted line) and 95% highest posterior density (HPD) intervals (shaded areas) for the genetic diversity estimates ( $N_e \tau$ ) through time. The BESP estimate is shown in blue and the BSP estimate is in red. (C) Median (solid line/dotted line) and 95% HPD intervals (shaded areas) of the estimated sampling intensities (beta) for each sampling epoch. The 12-epoch BESP estimates are shown in blue and a single-epoch (density-defined) estimate is in yellow.

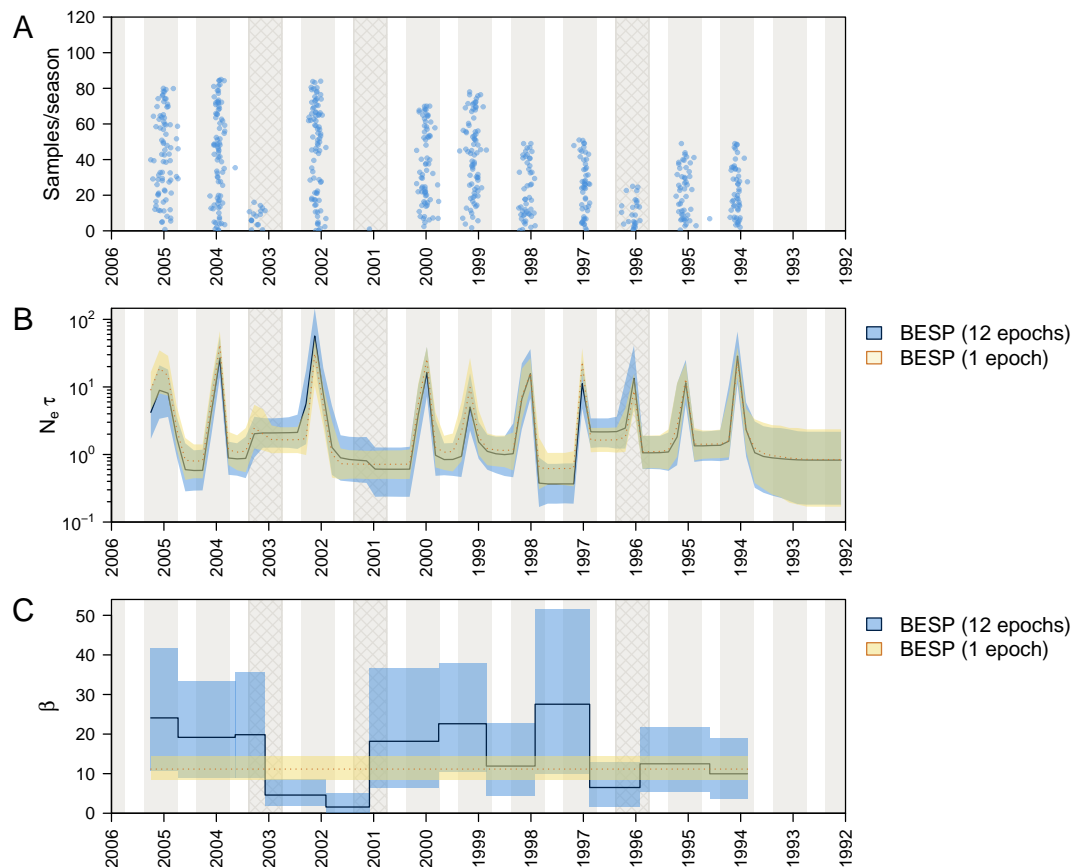


Figure 3: (A) Density of sequence sampling dates through time for the alignment of 637 A/H3N2 HA sequences from NY state that we analysed. Blue dots indicate stripcharts of individual samples for each season. The stripchart heights give the number of samples in each season. Grey shading indicates the approximate period of influenza observation in New York state during each season (epidemiological week 40, to week 20 in the next year). Cross-hatched seasons are those where A/H3N2 was not the dominant influenza virus subtype. (B) Median (solid/dotted line) and 95% highest posterior density (HPD) intervals (shaded areas) for the genetic diversity estimates ( $N_e \tau$ ) through time. The 12-epoch BESP estimate is shown in blue and the single-epoch BESP estimate is in yellow. (C) Median (solid line/dotted line) and 95% HPD intervals (shaded areas) of the estimated sampling intensities ( $\beta$ ) for each sampling epoch. The 12-epoch BESP estimates are shown in blue and a single-epoch (density-defined) estimate is in yellow.



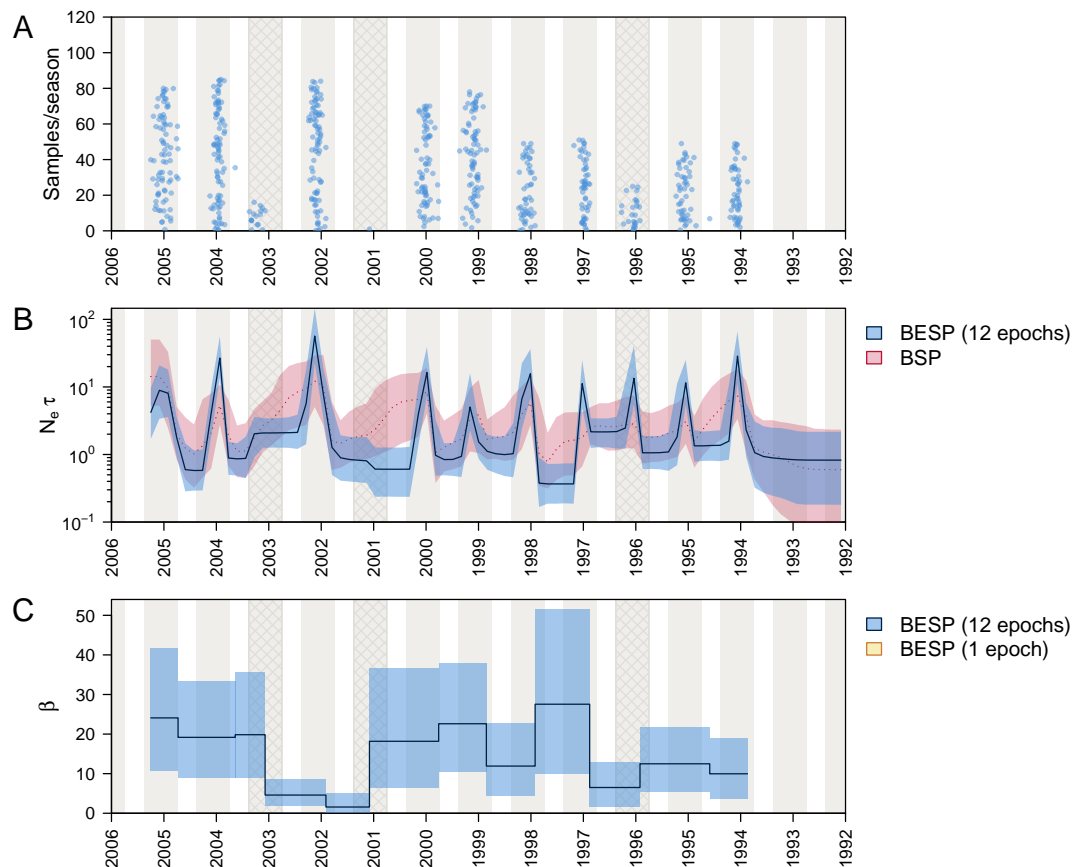


Figure 4: (A) Density of sequence sampling dates through time for the alignment of 637 A/H3N2 HA sequences from NY state that we analysed. Blue dots indicate stripcharts of individual samples for each season. The stripchart heights give the number of samples in each season. Grey shading indicates the approximate period of influenza observation in New York state during each season (epidemiological week 40, to week 20 in the next year). Cross-hatched seasons are those where A/H3N2 was not the dominant influenza virus subtype. (B) Median (solid/dotted line) and 95% highest posterior density (HPD) intervals (shaded areas) for the genetic diversity estimates ( $N_e \tau$ ) through time. The BESP estimate is shown in blue and the BSP estimate is in red. (C) Median (solid line/dotted line) and 95% HPD intervals (shaded areas) of the estimated sampling intensities (beta) for each sampling epoch. The 12-epoch BESP estimates are shown in blue and a single-epoch (density-defined) estimate is in yellow.

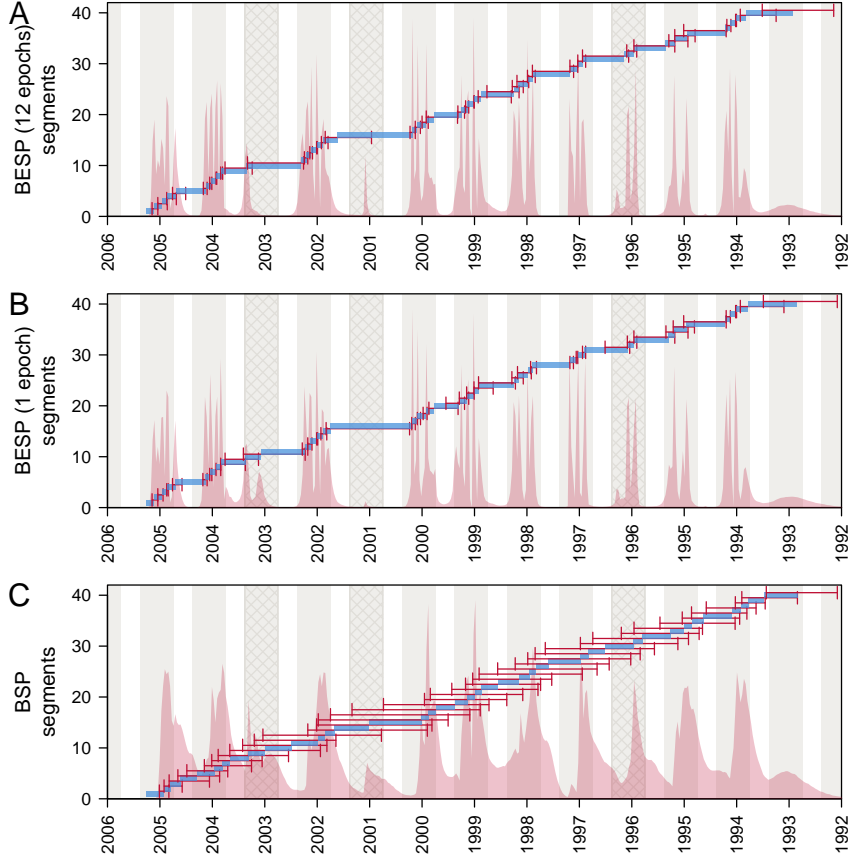


Figure 5: Population size segments for the alignment of 637 A/H3N2 HA sequences from NY state, as estimated under the 12-epoch BESP (A), single-epoch BESP (B) and BSP (C), with  $p = 40$ . Median posterior estimates of segments ( $t_{\{j-1\}}-t_j$ ) are shown in blue. HPD intervals for the segment end-times are indicated by red arrows. Red shading shows the kernel density estimate of the posterior segment times ( $t_j$ ). Grey shading indicates the approximate period of influenza observation in New York state during each season (epidemiological week 40, to week 20 in the next year). Cross-hatched seasons are those where A/H3N2 was not the dominant influenza virus subtype.

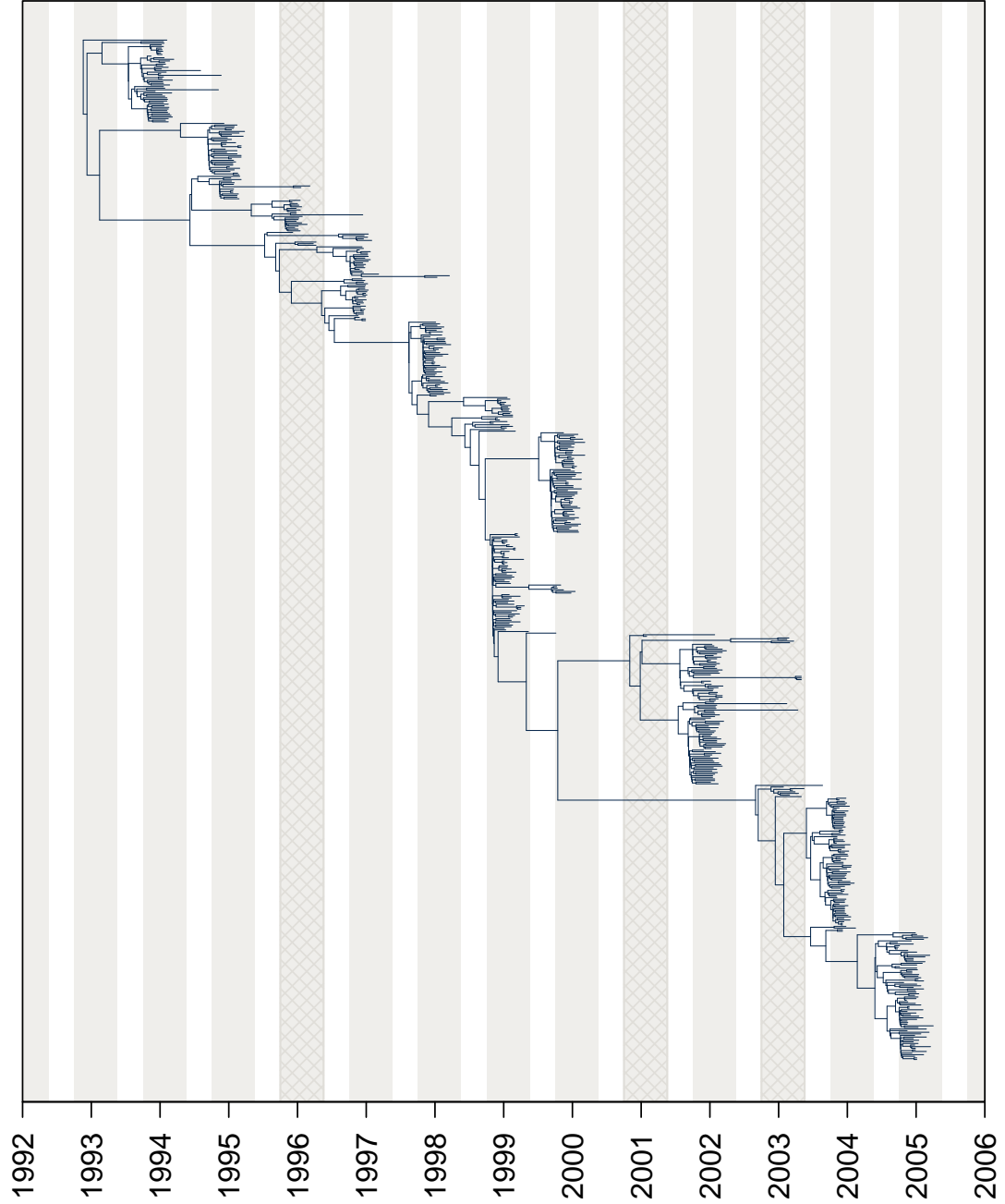


Figure 6: MCC tree of the alignment of 637 A/H3N2 HA sequences estimated under the 12-epoch BESP. Grey shading indicates the approximate period of influenza observation in New York state during each season (epidemiological week 40, to week 20 in the next year). Cross-hatched seasons are those where A/H3N2 was not the dominant influenza virus subtype.

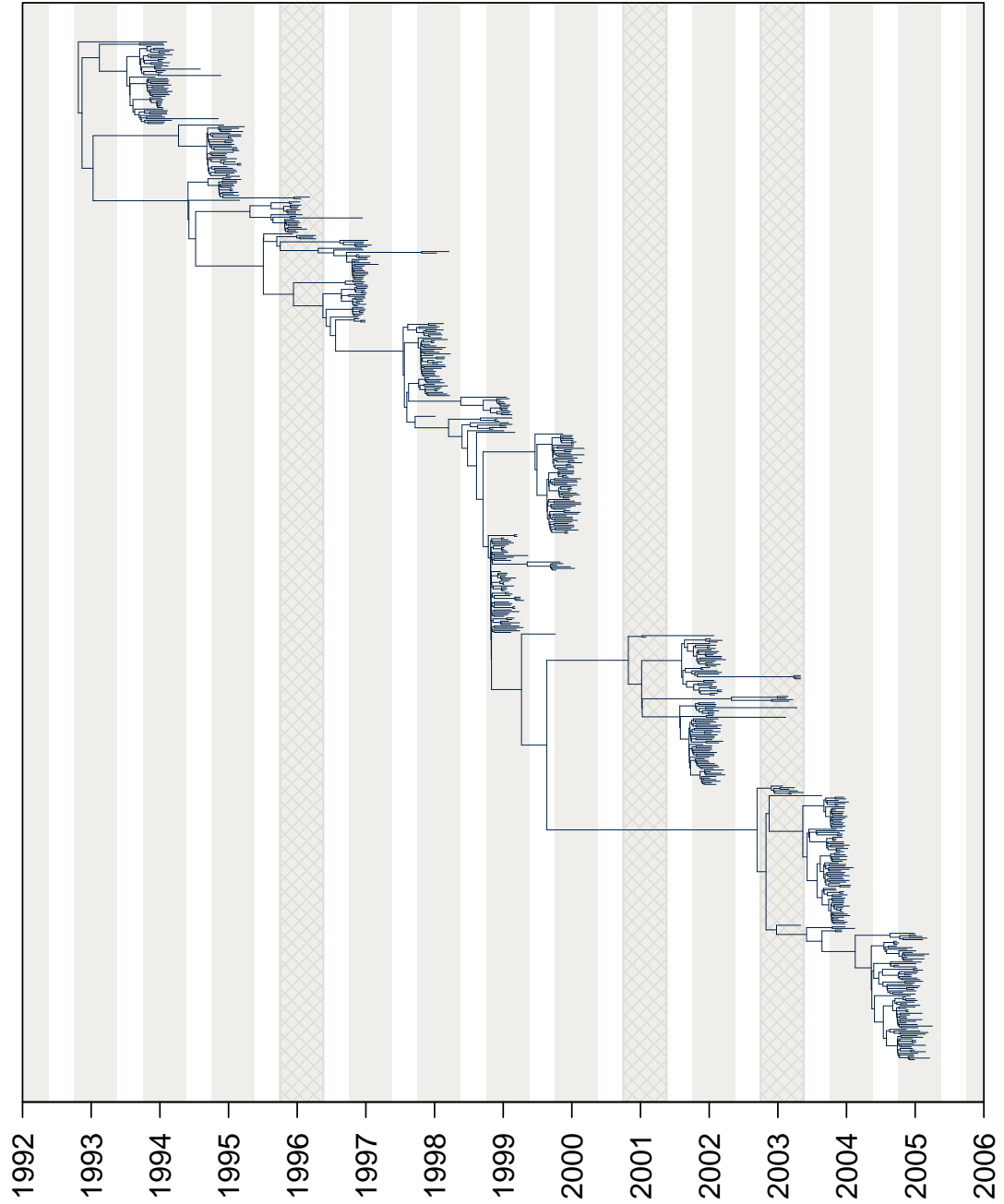


Figure 7: MCC tree of the alignment of 637 A/H3N2 HA sequences estimated under the single-epoch BEP. Grey shading indicates the approximate period of influenza observation in New York state during each season (epidemiological week 40, to week 20 in the next year). Cross-hatched seasons are those where A/H3N2 was not the dominant influenza virus subtype.

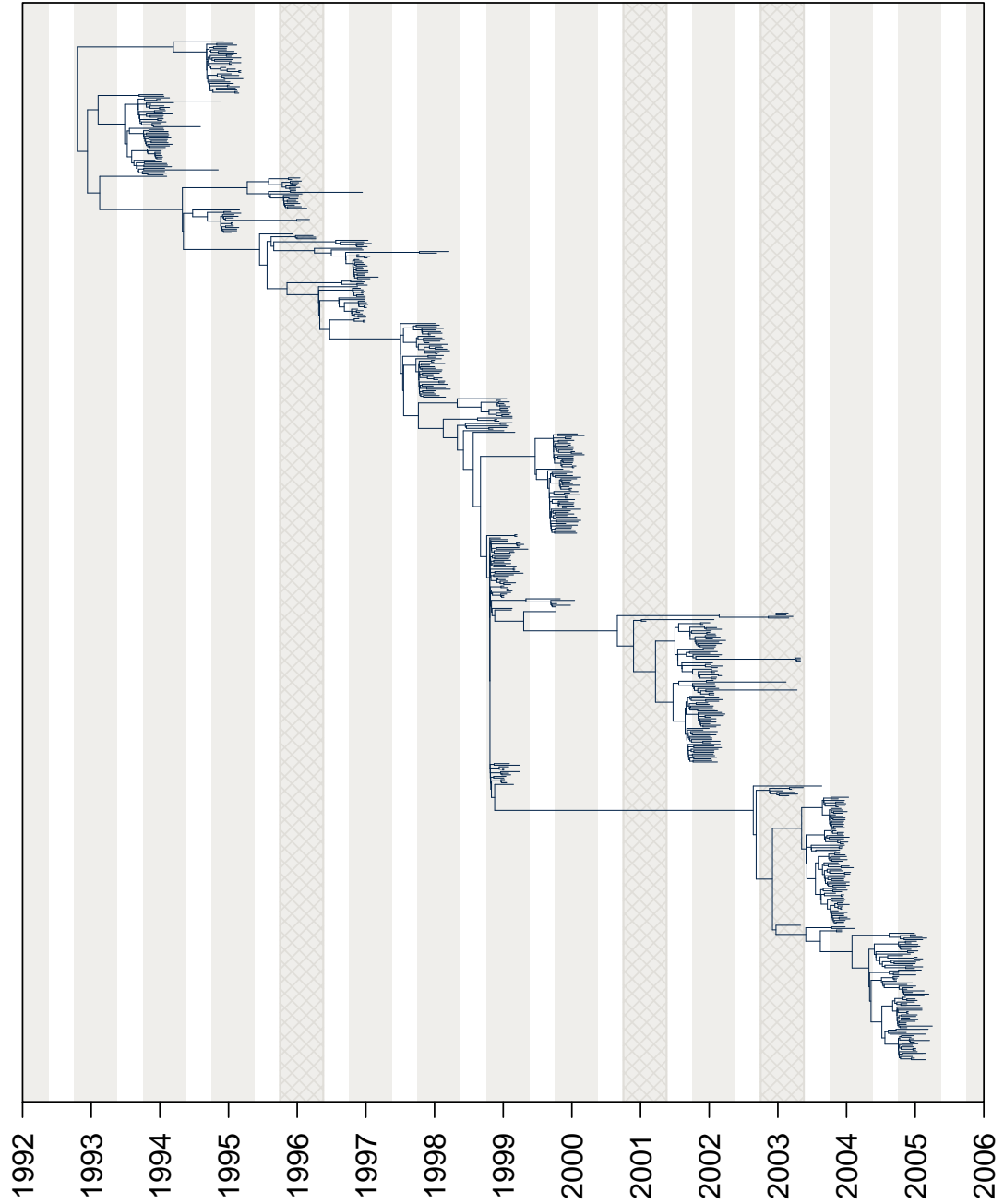


Figure 8: MCC tree of the alignment of 637 A/H3N2 HA sequences estimated under the single-epoch BESP. Grey shading indicates the approximate period of influenza observation in New York state during each season (epidemiological week 40, to week 20 in the next year). Cross-hatched seasons are those where A/H3N2 was not the dominant influenza virus subtype.

## 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] rskylinetools_0.2.1 coda_0.19-3      beastio_0.2.5
## [4] ape_5.3
##
## loaded via a namespace (and not attached):
## [1] Rcpp_1.0.2          lattice_0.20-38    digest_0.6.21
## [4] grid_3.5.1          nlme_3.1-141      magrittr_1.5
## [7] evaluate_0.14       highr_0.8         rlang_0.4.0
## [10] stringi_1.4.3       rmarkdown_1.16    RColorBrewer_1.1-2
## [13] tools_3.5.1         stringr_1.4.0     parallel_3.5.1
## [16] xfun_0.10           yaml_2.2.0        compiler_3.5.1
## [19] htmltools_0.4.0    knitr_1.25
```