**Supplementary data**

**for**

**Inference of population genetic parameters from the semi-randomly serial-sampled sequences of human seasonal influenza A/H3N2**

by

Myriam Croze and Yuseob Kim

**Table S1.**  Bootstrap tests for the significance of difference in mutation rates, pairwise sequence diversity, and effective population size between segments, for 27-, 10-, and 37-year data sets. Each number is the percentage of resampled data that yield the difference of a given statistic twice larger than the difference in the observed data.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **27-year (1980-2006)** | | | **10-year (2007-2017)** | | | **37 year (1980-2017)** | | |
|  |  |  |  |  |  |  |  |  |  |  |
| **HA vs** | **PB1** | **0.026** | **0.005** | **0.007** | 0.326 | **0.040** | 0.138 | 0.099 | **0** | **0.002** |
| **PB2** | **0.012** | **0** | **0.007** | 0.379 | **0.037** | 0.165 | **0.019** | **0** | **0.002** |
| **PA** | **0.004** | **0.012** | **0.014** | 0.425 | 0.359 | 0.359 | **0.006** | **0.001** | **0.001** |
| **NP** | **0** | **0.002** | **0.018** | 0.100 | 0.359 | 0.134 | **0.007** | **0** | **0.020** |
| **NA** | **0.008** | **0.025** | **0.011** | 0.294 | 0.153 | 0.478 | **0.007** | **0.020** | **0.080** |
| **PB1 vs** | **PB2** | 0.175 | 0.347 | 0.204 | 0.438 | 0.482 | 0.430 | 0.149 | 0.390 | 0.133 |
| **PA** | 0.152 | 0.484 | 0.280 | 0.391 | 0.084 | 0.208 | 0.090 | 0.457 | 0.153 |
| **NP** | **0.017** | 0.225 | 0.068 | 0.178 | 0.106 | 0.356 | 0.059 | 0.406 | 0.140 |
| **NA** | 0.168 | 0.488 | 0.281 | 0.476 | **0.004** | 0.154 | 0.057 | 0.115 | 0.362 |
| **PB2 vs** | **PA** | 0.467 | 0.345 | 0.415 | 0.445 | 0.057 | 0.239 | 0.382 | 0.378 | 0.473 |
| **NP** | 0.114 | 0.329 | 0.155 | 0.138 | 0.092 | 0.314 | 0.294 | 0.324 | 0.442 |
| **NA** | 0.481 | 0.353 | 0.408 | 0.411 | **0.001** | 0.184 | 0.316 | 0.077 | 0.292 |
| **PA vs** | **NP** | 0.133 | 0.190 | 0.128 | 0.119 | 0.482 | 0.169 | 0.390 | 0.433 | 0.446 |
| **NA** | 0.479 | 0.510 | 0.488 | 0.358 | 0.093 | 0.398 | 0.425 | 0.134 | 0.286 |
| **NP vs** | **NA** | 0.167 | 0.241 | 0.132 | 0.241 | 0.086 | 0.154 | 0.469 | 0.158 | 0.247 |

**Table S2.** Sampling-time-corrected sequence diversity and *F*ST in the simulation of metapopulation dynamics under neutrality (*s* = 0) and with τmax = 300 days. The values in bold are the closest to the observation..

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| *Kmax* |  | *m* | | | | |
|  | 0 | 0.00001 | 0.0001 | 0.001 | 0.01 |
| 100 |  | NA | 0.237 | 0.141 | 0.025 | 0.064 |
| *F*ST | NA | 0.952 | 0.916 | 0.459 | 0.0291 |
| 250 |  | 0.255 | 0.243 | 0.128 | **0.043** | 0.064 |
| *F*ST | 0.895 | 0.938 | 0.828 | **0.320** | 0.029 |
| 500 |  | 0.259 | 0.248 | 0.100 | 0.064 | 0.112 |
| *F*ST | 0.803 | 0.924 | 0.675 | 0.225 | 0.014 |
| 1,000 |  | 0.268 | 0.255 | 0.107 | 0.104 | 0.179 |
| *F*ST | 0.647 | 0.887 | 0.542 | 0.132 | 0.005 |
| 2,000 |  | 0.284 | 0.254 | 0.139 | 0.154 | 0.267 |
| *F*ST | 0.387 | 0.818 | 0.423 | 0.074 | 0.005 |
| 5,000 |  | 0.320 | 0.262 | 0.222 | 0.262 | 0.368 |
| *F*ST | -0.105 | 0.643 | 0.272 | 0.033 | 0.003 |
| 10,000 |  | 0.362 | 0.282 | 0.292 | 0.338 | 0.425 |
| *F*ST | -0.544 | 0.502 | 0.176 | 0.015 | 0.000 |

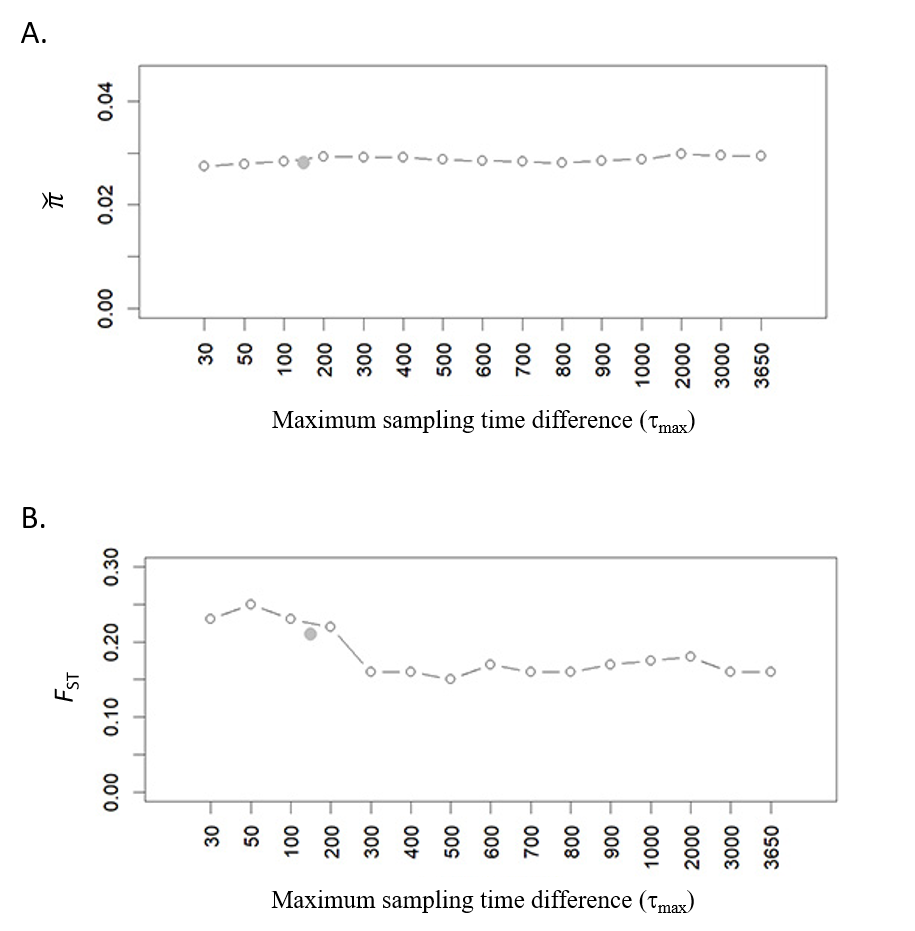
**Table S3.** Sampling-time-corrected sequence diversity and *F*ST in the simulation of metapopulation dynamics with positive selection (*s*) and for various values of migration (*m* ), maximum carrying capacity (*Kmax*) and epitope (ε) (τmax = 300 days).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | *s* = 0.05 | | | | | *s* = 0.1 | | | | | *s* = 0.15 | | | | |
|  |  |  | *m* | | | | | *m* | | | | | *m* | | | | |
| *Kmax* | ε |  | 0 | 0.00001 | 0.0001 | 0.001 | 0.01 | 0 | 0.00001 | 0.0001 | 0.001 | 0.01 | 0 | 0.00001 | 0.0001 | 0.001 | 0.01 |
| 250 | 10 |  | NA | 0.183 | 0.074 | 0.031 | 0.029 | NA | NA | 0.053 | 0.023 | 0.019 | NA | NA | 0.038 | 0.019 | 0.016 |
| *F*ST | NA | 0.903 | 0.789 | 0.312 | 0.027 | NA | NA | 0.760 | 0.297 | 0.022 | NA | NA | 0.738 | 0.280 | 0.007 |
| 20 |  | NA | NA | 0.065 | 0.028 | 0.025 | NA | NA | 0.041 | 0.021 | 0.016 | NA | NA | 0.042 | 0.021 | 0.013 |
| *F*ST | NA | NA | 0.784 | 0.310 | 0.026 | NA | NA | 0.639 | 0.285 | 0.022 | NA | NA | 0.757 | 0.289 | 0026 |
| 30 |  | NA | NA | 0.059 | 0.027 | 0.023 | NA | NA | 0.037 | 0.019 | 0.015 | NA | NA | 0.028 | 0.016 | 0.012 |
| *F*ST | NA | NA | 0.779 | 0.311 | 0.026 | NA | NA | 0.736 | 0.286 | 0.022 | NA | NA | 0.719 | 0.266 | 0.026 |
| 40 |  | NA | NA | 0.057 | 0.026 | 0.021 | NA | NA | 0.035 | 0.018 | 0.015 | NA | NA | 0.025 | 0.015 | 0.012 |
| *F*ST | NA | NA | 0.776 | 0.314 | 0.023 | NA | NA | 0.729 | 0.278 | 0.025 | NA | NA | 0.614 | 0.262 | 0.025 |
| 500 | 10 |  | NA | 0.172 | 0.052 | 0.034 | 0.030 | NA | NA | 0.037 | 0.023 | 0.019 | NA | NA | 0.024 | 0.018 | 0.015 |
| *F*ST | NA | 0.883 | 0.620 | 0.211 | 0.010 | NA | NA | 0.595 | 0.200 | 0.011 | NA | NA | 0.448 | 0.191 | 0.013 |
| 20 |  | NA | NA | 0.048 | 0.030 | 0.027 | NA | NA | 0.032 | 0.021 | 0.017 | NA | NA | 0.025 | 0.021 | 0.014 |
| *F*ST | NA | NA | 0.623 | 0.213 | 0.008 | NA | NA | 0.595 | 0.207 | 0.010 | NA | NA | 0.564 | 0.200 | 0.011 |
| 30 |  | NA | NA | 0.045 | 0.028 | 0.025 | NA | NA | 0.030 | 0.020 | 0.016 | NA | NA | 0.024 | 0.016 | 0.013 |
| *F*ST | NA | NA | 0.626 | 0.213 | 0.010 | NA | NA | 0.589 | 0.202 | 0.013 | NA | NA | 0.572 | 0.198 | 0.014 |
| 40 |  | NA | NA | 0.042 | 0.027 | 0.024 | NA | NA | 0.029 | 0.020 | 0.016 | NA | NA | 0.022 | 0.016 | 0.013 |
| *F*ST | NA | NA | 0.616 | 0.211 | 0.010 | NA | NA | 0.584 | 0.198 | 0.012 | NA | NA | 0.562 | 0.198 | 0.014 |
| 1,000 | 10 |  | 0.255 | 0.141 | 0.048 | 0.035 | 0.034 | NA | 0.088 | 0.034 | 0.024 | 0.020 | NA | NA | 0.025 | 0.019 | 0.016 |
| *F*ST | 0.740 | 0.849 | 0.499 | 0.129 | 0.003 | NA | 0.821 | 0.477 | 0.185 | 0.000 | NA | NA | 0.456 | 0.133 | 0.007 |
| 20 |  | NA | 0.126 | 0.044 | 0.031 | 0.029 | NA | 0.062 | 0.031 | 0.021 | 0.018 | NA | 0.035 | 0.024 | 0.021 | 0.014 |
| *F*ST | NA | 0.838 | 0.496 | 0.139 | 0.002 | NA | 0.773 | 0.482 | 0.136 | 0.003 | NA | 0.683 | 0.455 | 0.140 | 0.008 |
| 30 |  | NA | NA | 0.043 | 0.030 | 0.027 | NA | 0.054 | 0.029 | 0.021 | 0.018 | NA | 0.029 | 0.023 | 0.016 | 0.014 |
| *F*ST | NA | NA | 0.494 | 0.138 | 0.000 | NA | 0.758 | 0.476 | 0.146 | 0.005 | NA | 0.659 | 0.457 | 0.149 | 0.010 |
| 40 |  | NA | 0.103 | 0.041 | 0.029 | 0.026 | NA | 0.046 | 0.029 | 0.020 | 0.017 | NA | NA | 0.023 | 0.016 | 0.013 |
| *F*ST | NA | 0.836 | 0.498 | 0.143 | 0.000 | NA | 0.763 | 0.477 | 0.143 | 0.006 | NA | NA | 0.460 | 0.149 | 0.009 |

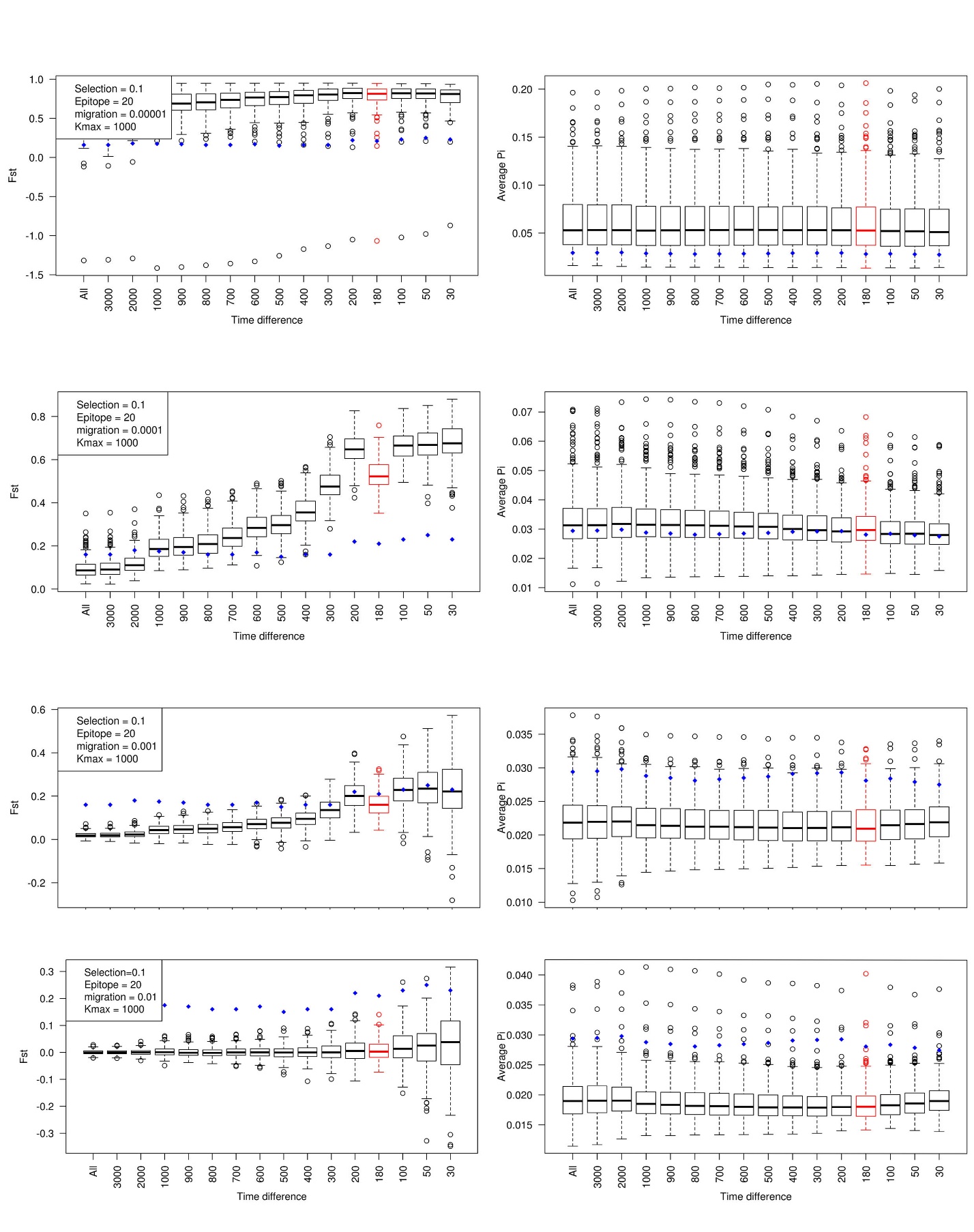
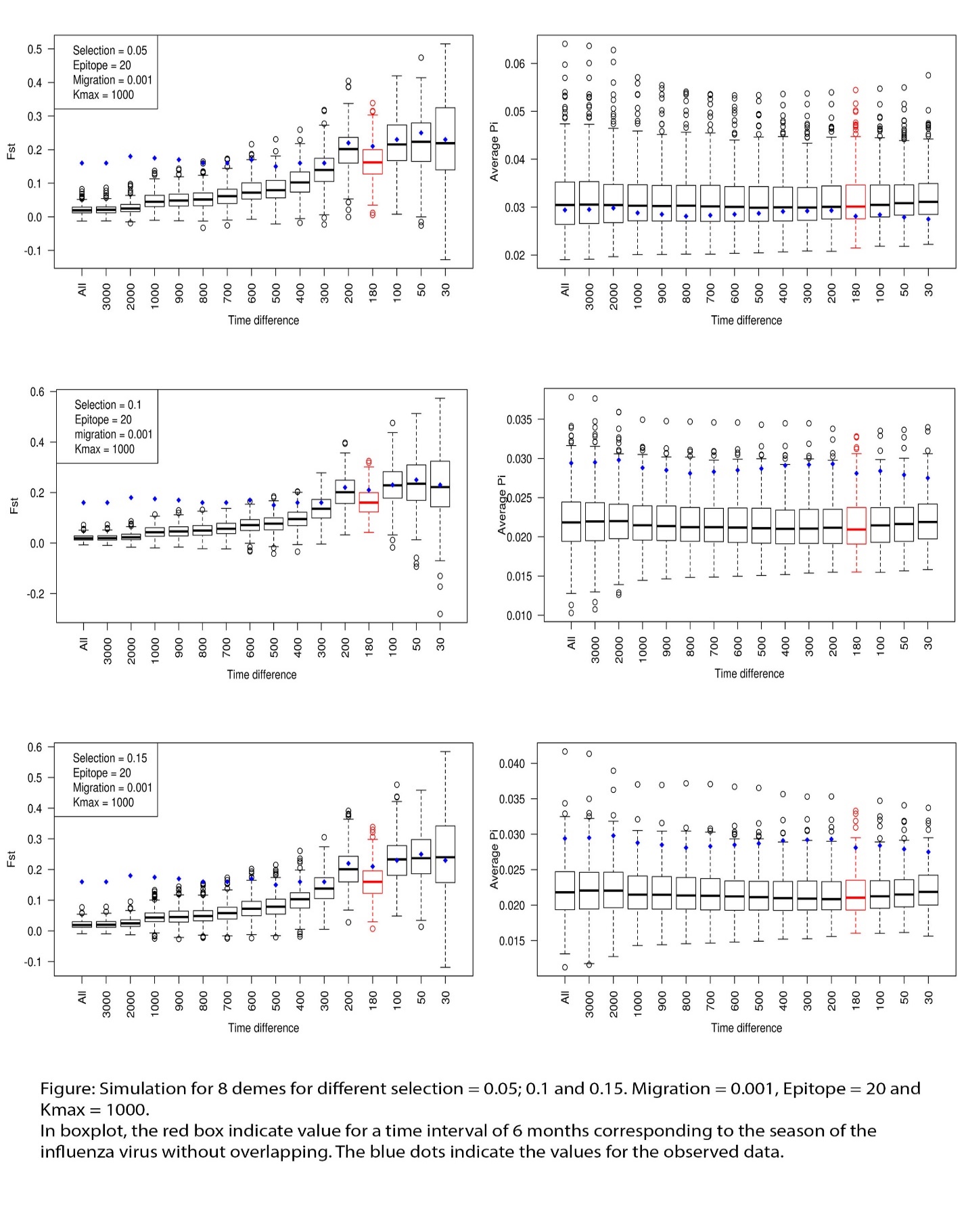
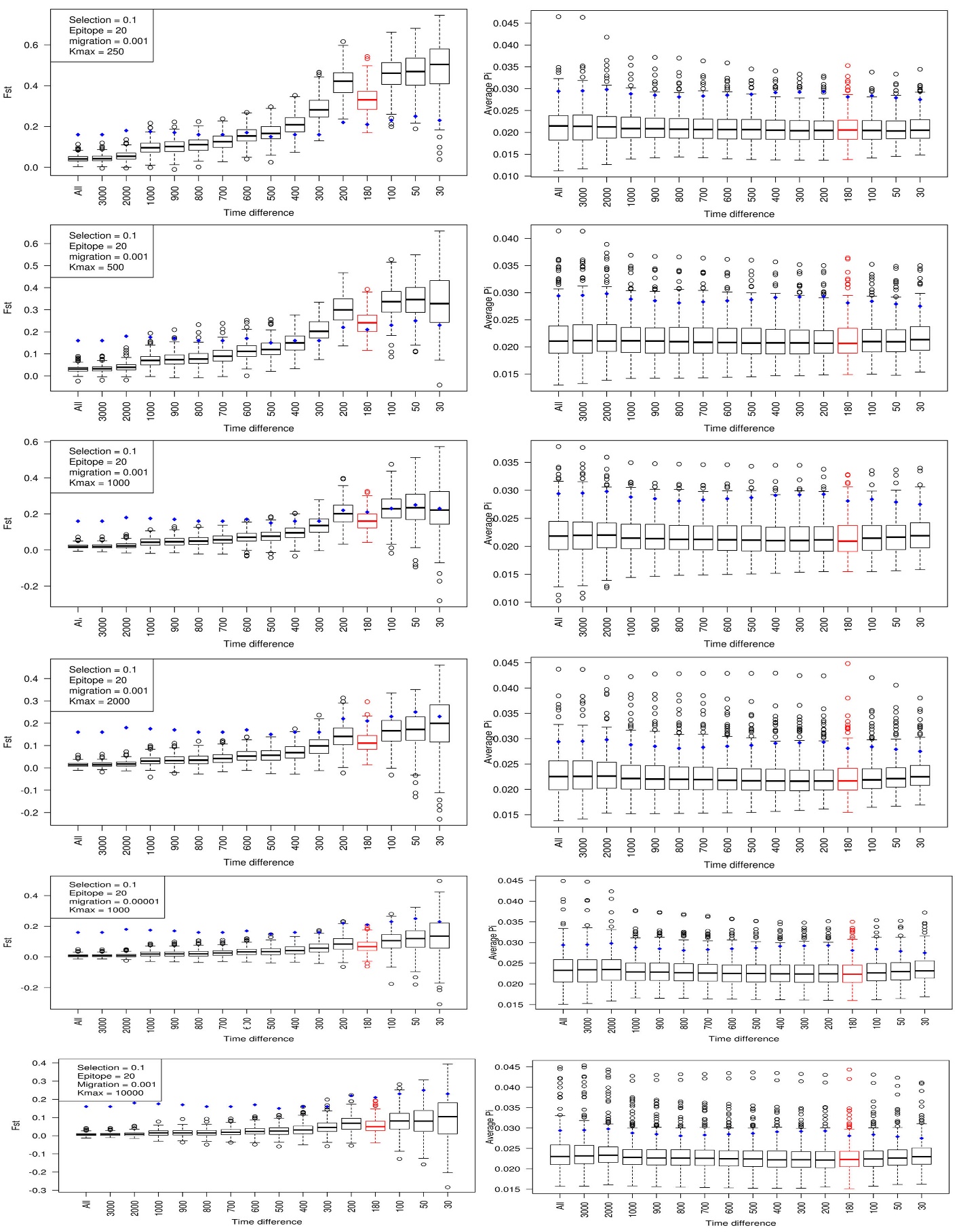
**Table S3 (next).** Sampling-time-corrected sequence diversity and *F*ST in the simulation of metapopulation dynamics with positive selection (*s*) and for various values of migration (*m* ), maximum carrying capacity (*Kmax*) and epitope (ε) (τmax = 300 days).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | *s* = 0.05 | | | | | *s* = 0.1 | | | | | *s* = 0.15 | | | | |
|  |  |  | *m* | | | | | *m* | | | | | *m* | | | | |
| *Kmax* | ε |  | 0 | 0.00001 | 0.0001 | 0.001 | 0.01 | 0 | 0.00001 | 0.0001 | 0.001 | 0.01 | 0 | 0.00001 | 0.0001 | 0.001 | 0.01 |
| 2,000 | 10 |  | 0.260 | 0.108 | 0.049 | 0.036 | 0.036 | NA | 0.070 | 0.033 | 0.025 | 0.022 | NA | 0.044 | 0.026 | 0.019 | 0.017 |
| *F*ST | 0.699 | 0.794 | 0.402 | 0.075 | -0.003 | NA | 0.768 | 0.393 | 0.124 | -0.002 | NA | 0.732 | 0.382 | 0.099 | 0.003 |
| 20 |  | NA | 0.094 | 0.045 | 0.046 | 0.031 | NA | 0.056 | 0.031 | 0.022 | 0.020 | NA | 0.036 | 0.025 | 0.022 | 0.015 |
| *F*ST | NA | 0.792 | 0.407 | 0.030 | -0.002 | NA | 0.763 | 0.399 | 0.097 | 0.000 | NA | 0.732 | 0.378 | 0.101 | 0.003 |
| 30 |  | NA | 0.089 | 0.044 | 0.031 | 0.029 | NA | 0.052 | 0.029 | 0.021 | 0.018 | NA | 0.033 | 0.024 | 0.017 | 0.014 |
| *F*ST | NA | 0.787 | 0.421 | 0.083 | -0.002 | NA | 0.767 | 0.398 | 0.098 | 0.001 | NA | 0.750 | 0.391 | 0.115 | 0.003 |
| 40 |  | NA | 0.086 | 0.041 | 0.030 | 0.027 | NA | 0.049 | 0.029 | 0.020 | 0.018 | NA | 0.031 | 0.023 | 0.017 | 0.014 |
| *F*ST | NA | 0.788 | 0.411 | 0.087 | -0.003 | NA | 0.761 | 0.403 | 0.107 | 0.001 | NA | 0.720 | 0.393 | 0.115 | 0.009 |
| 5,000 | 10 |  | 0.264 | 0.083 | 0.049 | 0.039 | 0.037 | NA | 0.054 | 0.033 | 0.025 | 0.023 | NA | 0.041 | 0.026 | 0.020 | 0.018 |
| *F*ST | 0.674 | 0.653 | 0.290 | 0.036 | -0.006 | NA | 0.637 | 0.294 | 0.079 | -0.003 | NA | 0.624 | 0.382 | 0.071 | 0.000 |
| 20 |  | NA | 0.077 | 0.046 | 0.035 | 0.033 | 0.049 | 0.049 | 0.031 | 0.023 | 0.021 | NA | 0.038 | 0.025 | 0.023 | 0.016 |
| *F*ST | NA | 0.645 | 0.303 | 0.042 | -0.004 | 0.638 | 0.640 | 0.302 | 0.061 | -0.005 | NA | 0.619 | 0.304 | 0.065 | 0.002 |
| 30 |  | NA | 0.071 | 0.044 | 0.033 | 0.031 | NA | 0.045 | 0.031 | 0.022 | 0.020 | NA | 0.035 | 0.024 | 0.018 | 0.015 |
| *F*ST | NA | 0.584 | 0.303 | 0.042 | -0.006 | NA | 0.625 | 0.30 | 0.069 | 0.001 | NA | 0.613 | 0.305 | 0.087 | 0.004 |
| 40 |  | NA | 0.072 | 0.043 | 0.031 | 0.030 | NA | 0.045 | 0.029 | 0.021 | 0.019 | NA | 0.034 | 0.023 | 0.017 | 0.015 |
| *F*ST | NA | 0.658 | 0.308 | 0.046 | -0.002 | NA | 0.630 | 0.306 | 0.071 | 0.005 | NA | 0.630 | 0.311 | 0.089 | 0.006 |
| 10,000 | 10 |  | 0.266 | 0.079 | 0.052 | 0.040 | 0.040 | NA | 0.054 | 0.034 | 0.026 | 0.024 | NA | 0.041 | 0.026 | 0.020 | 0.018 |
| *F*ST | 0.665 | 0.549 | 0.213 | 0.026 | -0.004 | NA | 0.545 | 0.233 | 0.039 | -0.003 | NA | 0.529 | 0.291 | 0.052 | 0.002 |
| 20 |  | 0.261 | 0.073 | 0.045 | 0.035 | 0.035 | 0.048 | 0.048 | 0.031 | 0.023 | 0.021 | NA | 0.035 | 0.025 | 0.023 | 0.07 |
| *F*ST | 0.692 | 0.563 | 0.217 | 0.026 | -0.004 | 0.531 | 0.528 | 0.247 | 0.049 | -0.006 | NA | 0.522 | 0.257 | 0.041 | 0.003 |
| 30 |  | NA | 0.067 | 0.045 | 0.034 | 0.033 | NA | 0.044 | 0.030 | 0.022 | 0.021 | NA | 0.035 | 0.024 | 0.018 | 0.016 |
| *F*ST | NA | 0.539 | 0.233 | 0.026 | -0.005 | NA | 0.530 | 0.253 | 0.048 | -0.002 | NA | 0.524 | 0.265 | 0.066 | 0.001 |
| 40 |  | NA | 0.067 | 0.043 | 0.032 | 0.031 | NA | 0.043 | 0.030 | 0.021 | 0.020 | NA | 0.032 | 0.023 | 0.018 | 0.016 |
| *F*ST | NA | 0.553 | 0.242 | 0.029 | -0.008 | NA | 0.534 | 0.259 | 0.052 | -0.001 | NA | 0.525 | 0.260 | 0.073 | 0.001 |

**Figure S1.** Sampling-time-corrected sequence diversity (A) and *F*ST (B) of segment HA in the 10-year H3N2 data set, with the increasing upper bound of sampling time difference (τmax, in days) between two sequences that are compared. The gray dot indicates the result when each pair of sequences were sampled within the same 6-month flu season.

****

**Figure S2.** *F*ST and from 300 simulated replicates under positive selection (*s* = 0.1) for various τmax, with ε = 20 and *m* = 0.001 (A and B) or 0.0001 (C) and *K*max = 1000 (B and C) or 10000 (A). The red plot indicates the result when each pair of sequences were sampled within the same 6-month flu season. The blue dots indicate the values observed in H3N2 data.



**A.**

**B.**

**C.**

Maximum sampling time difference (τmax)

Maximum sampling time difference (τmax)

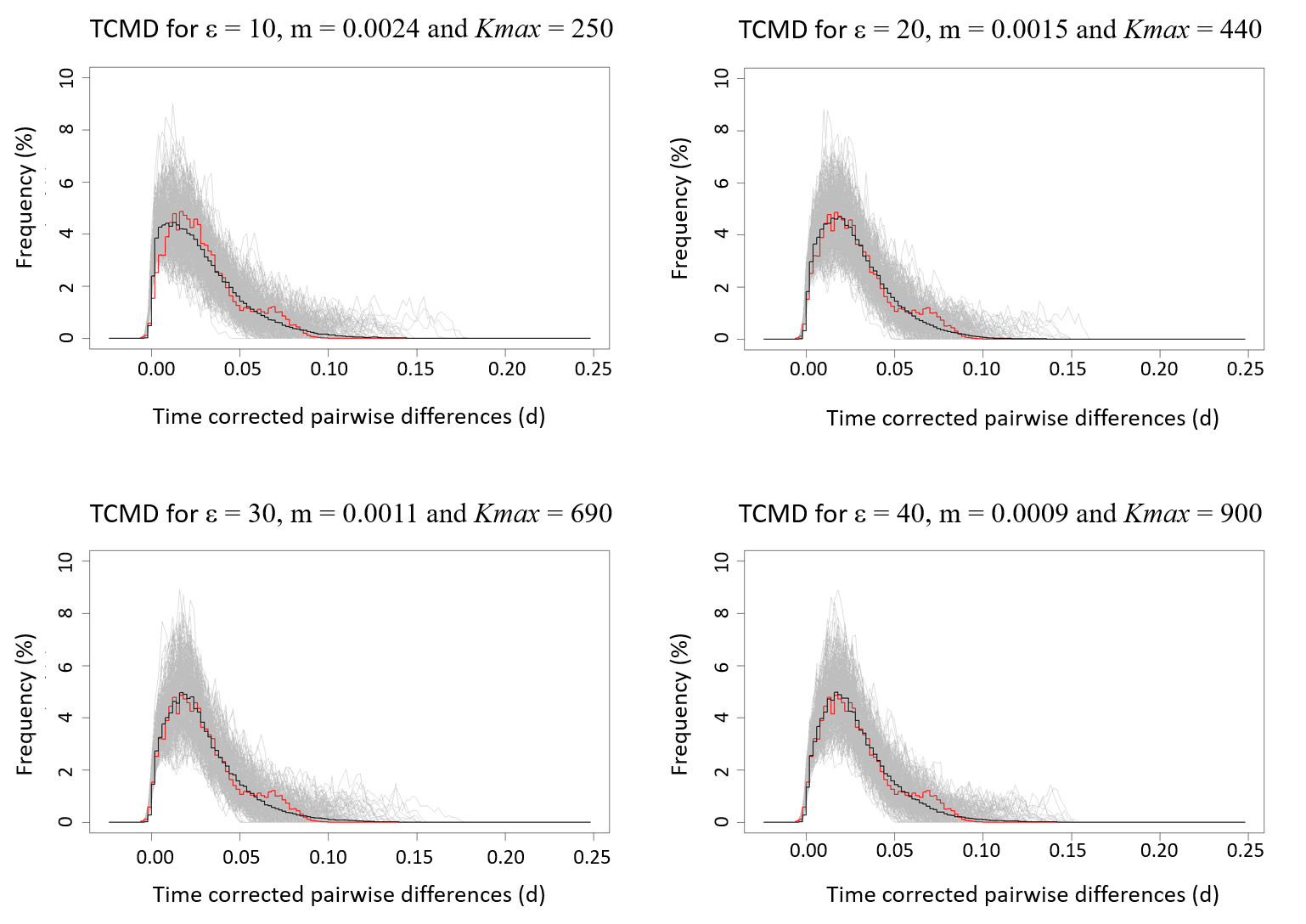
Maximum sampling time difference (τmax)

Maximum sampling time difference (τmax)

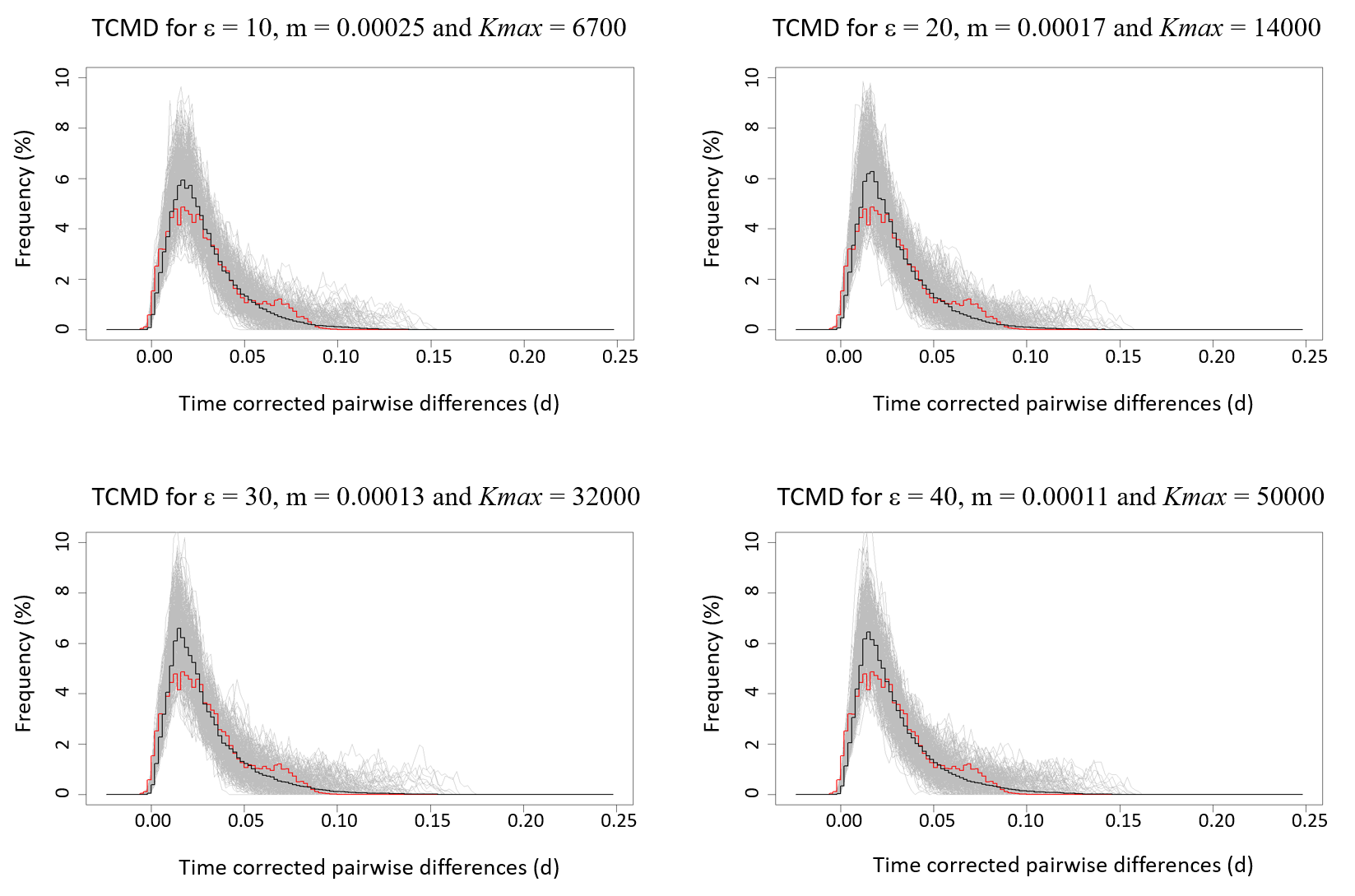
Maximum sampling time difference (τmax)

Maximum sampling time difference (τmax)

**Figure S3:** TCMD for simulated data under positive selection with *s* = 0.05 and ε = 10, 20, 30, and 40. For each set of *s* and *ε*, the best values of *K*max and *m* were found as shown in Table 2. The grey curves represent 300 simulated replicates for the same parameters. The black curve represents the average of those replicates and the red one represents the observed data.

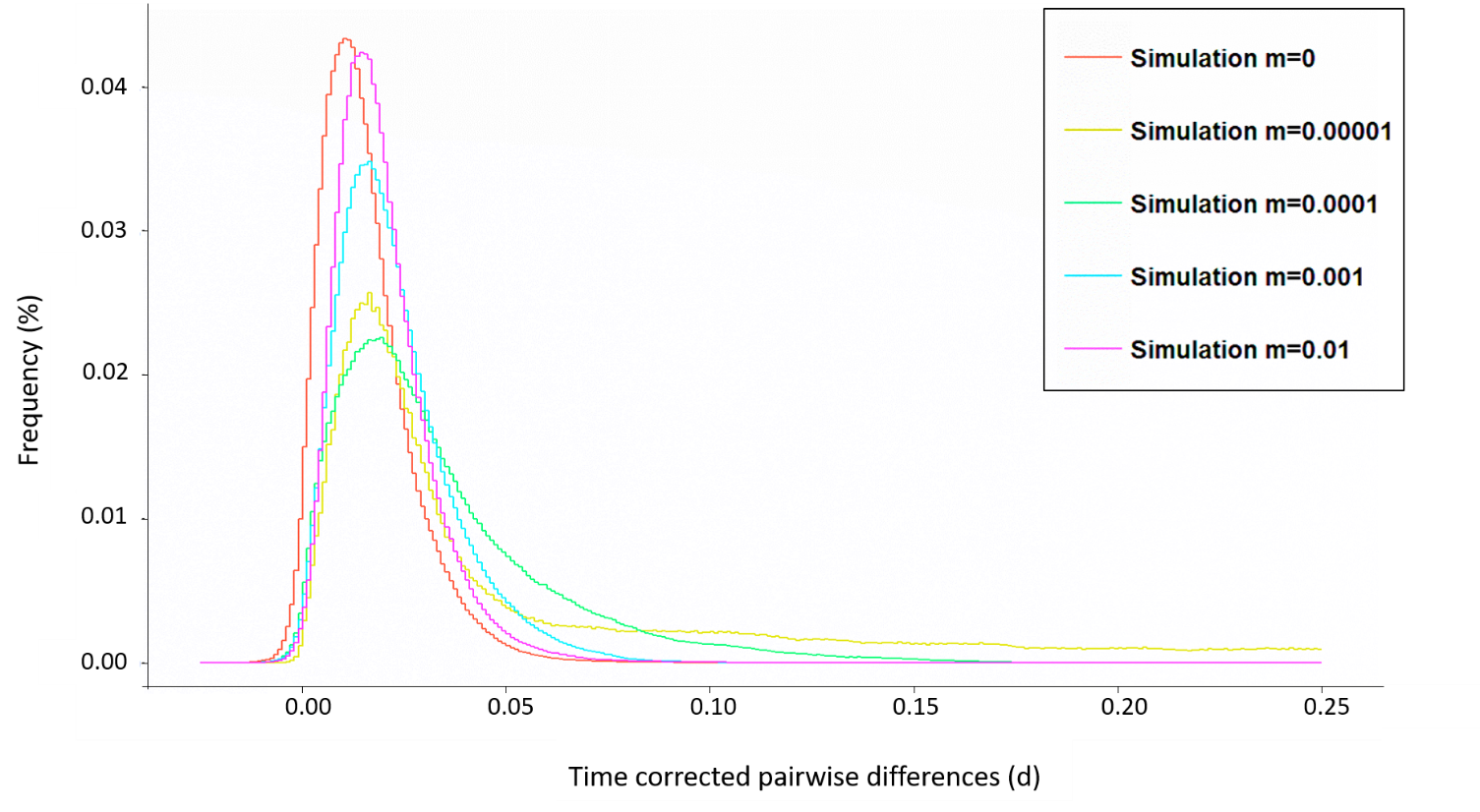


**Figure S4.** TCMD for simulated data under positive selection with *s* = 0.1 and ε = 10, 20, 30, and 40. For each set of *s* and ε, the best values of *K*max and *m* were found as shown in Table 2. The grey curves represent 300 simulated replicates for the same parameters. The black curve represents the average of those replicates and the red one represents the observed data.

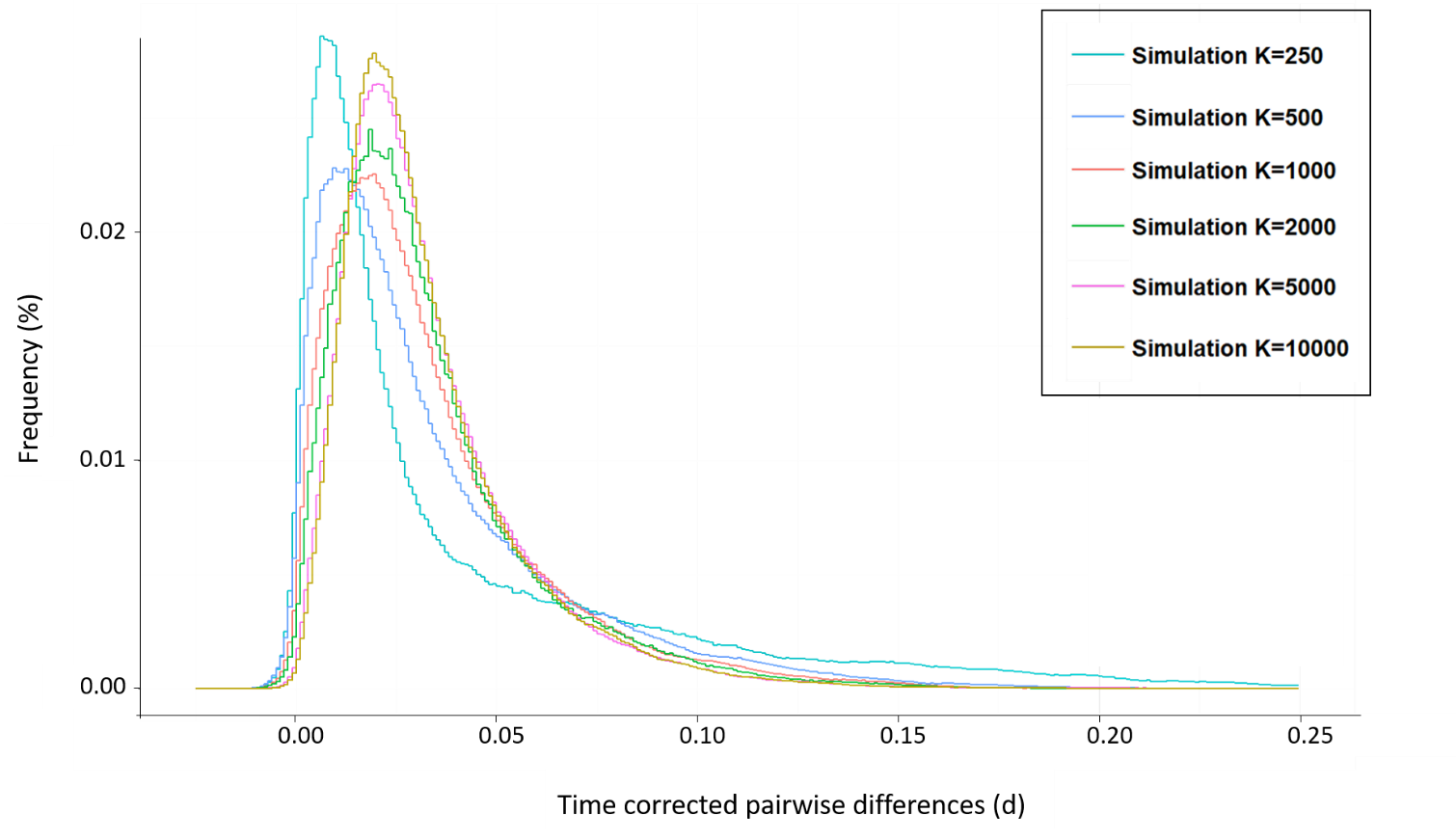


**Figure S5.** TCMD of simulated 10-year HA sequences for different values of *m* (**A**; with *K*max = 1000) and of *K*max (**B**; with *m* = 0.0001). Other parameters: *s* = 0.1 and ε = 10. Each curve is an average over 300 simulated replicates.

**A.**

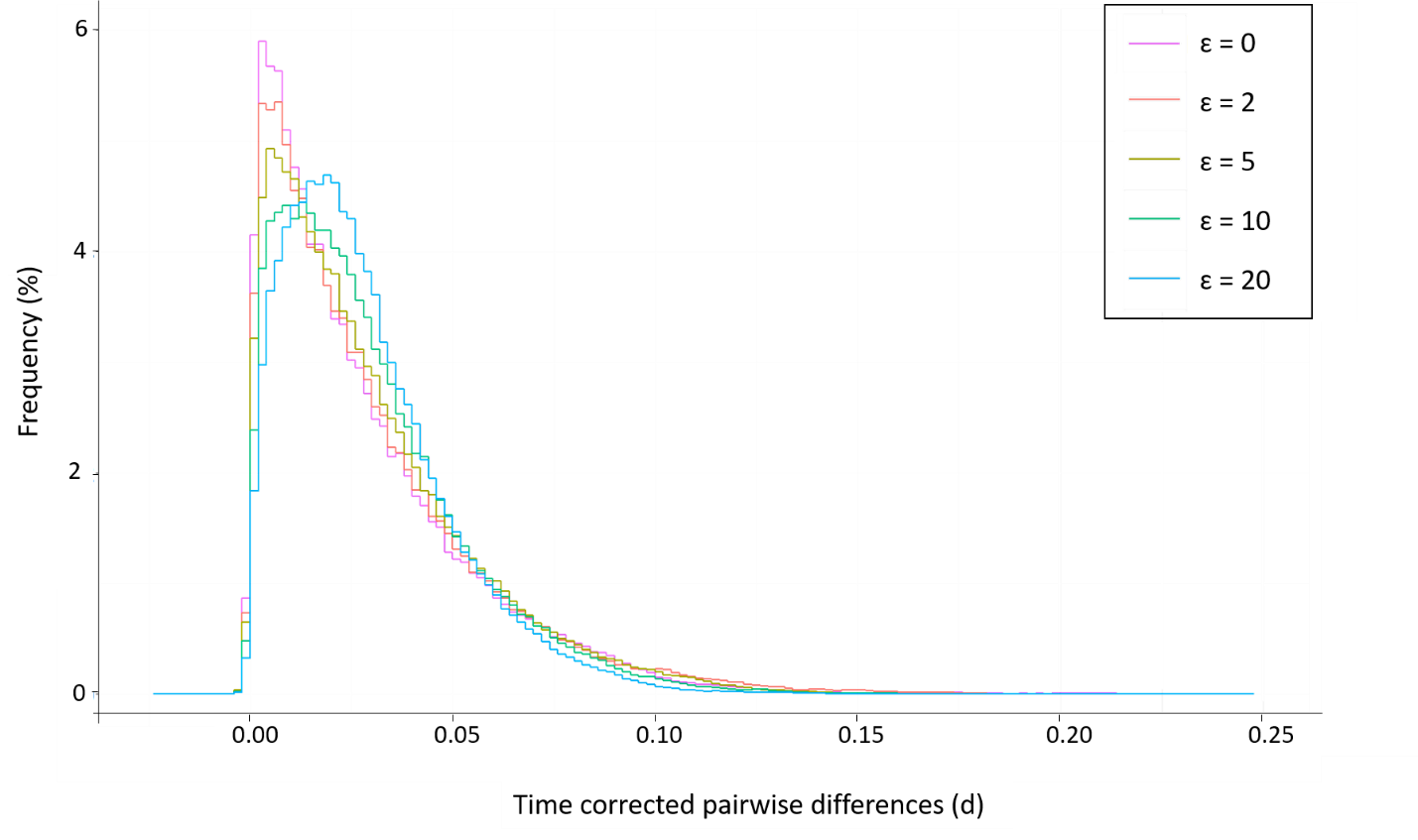
****

**B.**

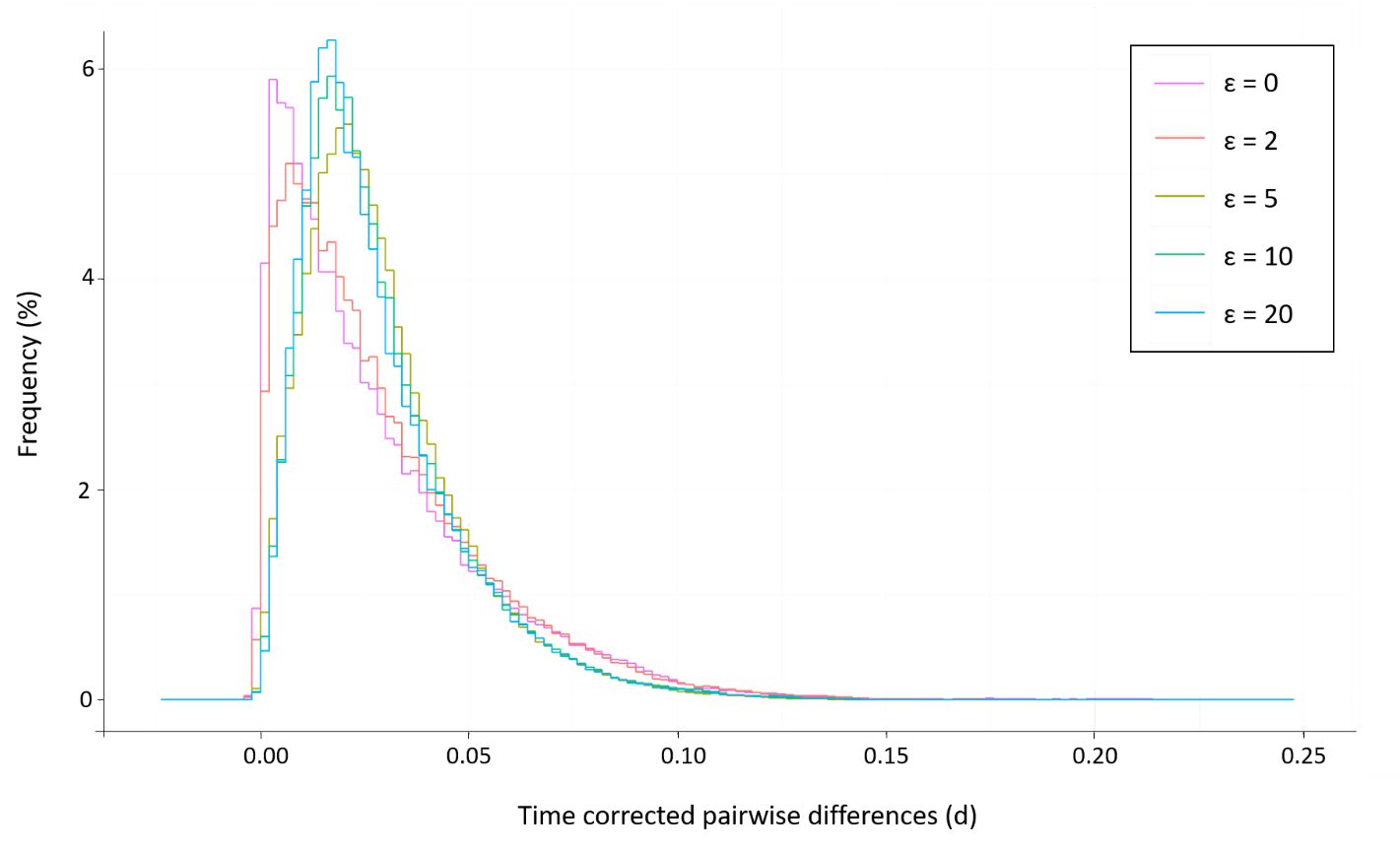
****

**Figure S6.** TCMD for simulated data with increasing frequency of beneficial mutations (ɛ) and other parameters that yield π and *F*ST matching the observed data. Under neutrality (ɛ = 0), *m* = 0.004 and *K*max = 110. Under positive selection with *s* = 0.05 (**A**), we use (ɛ, *m*, *K*max) = (2, 0.0036, 130), (5, 0.003, 170), (10, 0.0024, 250), and (20, 0.0015, 440). With *s* = 0.1 (**B**), we use (ɛ, *m*, *K*max) = (2, 0.003, 180), (5, 0.0004, 2500), (10, 0.00025, 6700), and (20, 0.00017, 1400).

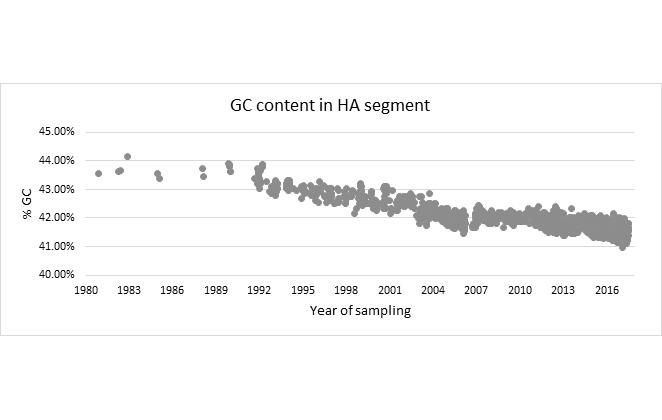
**A.**

****

**B.**

****

**Figure S7.** Plot of the GC content (in %) of the HA segment of H3N2 viruses against the time of sampling.

****