

Temporally structured metapopulation dynamics and persistence of influenza A H3N2 virus in humans

Justin Bahl^{a,b,1}, Martha I. Nelson^c, Kwok H. Chan^a, Rubing Chen^d, Dhanasekaran Vijaykrishna^{a,b}, Rebecca A. Halpin^e, Timothy B. Stockwell^e, Xudong Lin^e, David E. Wentworth^e, Elodie Ghedin^{e,f}, Yi Guan^a, J. S. Malik Peiris^{a,g}, Steven Riley^{h,i}, Andrew Rambaut^{c,j}, Edward C. Holmes^{c,k}, and Gavin J. D. Smith^{a,b,1}

^aDepartment of Microbiology, State Key Laboratory of Emerging Infectious Diseases, ^bDepartment of Community Medicine and School of Public Health, Li Ka Shing Faculty of Medicine, and ^gHong Kong University-Pasteur Research Centre, University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, China; ^hLaboratory of Virus Evolution, Program in Emerging Infectious Diseases, Duke-National University of Singapore Graduate Medical School, Republic of Singapore 169857; ^cFogarty International Center, National Institutes of Health, Bethesda, MD 20892; ^dUniversity of Texas Medical Branch, Galveston, TX 77555; ⁱCraig Venter Institute, Rockville, MD 20850; ^fCenter for Vaccine Research, Department of Computational and Systems Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261; ^lMedical Research Council Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College, London W2 1PG, United Kingdom; ^jInstitute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom; and ^kCenter for Infectious Disease Dynamics, Department of Biology, Pennsylvania State University, University Park, PA 16802

Edited by Peter Palese, Mount Sinai School of Medicine, New York, NY, and approved October 12, 2011 (received for review June 9, 2011)

Populations of seasonal influenza virus experience strong annual bottlenecks that pose a considerable extinction risk. It has been suggested that an influenza source population located in tropical Southeast or East Asia seeds annual temperate epidemics. Here we investigate the seasonal dynamics and migration patterns of influenza A H3N2 virus by analysis of virus samples obtained from 2003 to 2006 from Australia, Europe, Japan, New York, New Zealand, Southeast Asia, and newly sequenced viruses from Hong Kong. In contrast to annual temperate epidemics, relatively low levels of relative genetic diversity and no seasonal fluctuations characterized virus populations in tropical Southeast Asia and Hong Kong. Bayesian phylogeographic analysis using discrete temporal and spatial characters reveal high rates of viral migration between urban centers tested. Although the virus population that migrated between Southeast Asia and Hong Kong persisted through time, this was dependent on virus input from temperate regions and these tropical regions did not maintain a source for annual H3N2 influenza epidemics. We further show that multiple lineages may seed annual influenza epidemics, and that each region may function as a potential source population. We therefore propose that the global persistence of H3N2 influenza A virus is the result of a migrating metapopulation in which multiple different localities may seed seasonal epidemics in temperate regions in a given year. Such complex global migration dynamics may confound control efforts and contribute to the emergence and spread of antigenic variants and drug-resistant viruses.

evolution | molecular epidemiology | source-sink | phylogeography

Influenza A is a major respiratory infection of humans, with annual epidemics occurring in Northern and Southern temperate regions (1). On a global scale, viral migration between regions with high seasonal disease incidence is important in determining large-scale epidemiological patterns of influenza (2). Phylogenetic analysis of H3N2 influenza A viruses on a local scale has shown that seasonal increases in relative genetic diversity of the virus population correspond to seasonal increases in viral incidence (3). However, most lineages go extinct soon after their emergence, with few persisting between seasonal epidemics (1–4). Accordingly, local epidemics are annually seeded from other regions (5). In addition, the relative genetic diversity of influenza virus in temperate regions is clearly limited by the population bottlenecks that occur at the end of seasonal epidemics (1–5).

Influenza infections of humans exhibit more variable seasonal patterns in tropical and subtropical zones than in temperate regions (6). Approximately 30% (2.1 billion) of the global population is concentrated in tropical East and Southeast Asia, often in large urban centers (7). Epidemiologically, these densely populated and highly connected regions have year-round in-

fluenza circulation but semiannual increases in influenza activity associated with winter seasonal peaks in both Northern and Southern hemispheres (8–10). This finding, coupled with analyses of phylogenetic and antigenic diversity, led to the idea that East and Southeast Asia may function as reservoirs that maintain high diversity of influenza viruses that seed seasonal influenza outbreaks in temperate regions; that is, influenza viruses circulating in East and Southeast Asia may represent a global source population, but temperate regions represent ecological sinks (3, 11, 12).

The rapid global spread of severe acute respiratory syndrome in 2003, which originated in Southern China, demonstrated how an infectious disease could rapidly spread through a flight-connected network (13, 14). However, equivalent studies to describe the evolutionary behavior and migration dynamics of influenza viruses in tropical regions are lacking and remain a major gap in our understanding of the epidemiological dynamics of influenza on a global scale.

In this study we selected seven geographic regions—Europe (Northern Temperate), New York State in the United States (Northern Temperate), Japan (Northeast Asia), Australia (Southern Temperate), New Zealand (Southern Temperate), Hong Kong (East Asia), and Southeast Asia (collectively Cambodia, Indonesia, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Vietnam, representing a putative reservoir)—to determine migration dynamics and formally assess the suitability of the global source-sink ecological model for seasonal epidemics of H3N2 influenza. Using both HA1 and full genome data, including 105 newly acquired influenza genomes from Hong Kong, we characterize the dynamics of influenza virus evolution during the period 2003 to 2006 and examine how this relates to epidemic patterns in temperate regions, and particularly whether tropical Asia consistently serves as a global source population for influenza outbreaks.

Author contributions: J.B., E.C.H., and G.J.D.S. designed research; J.B., M.I.N., K.H.C., R.C., D.V., and R.A.H. performed research; J.B., R.A.H., T.B.S., X.L., D.E.W., E.G., Y.G., J.S.M.P., and A.R. contributed new reagents/analytic tools; J.B., M.I.N., K.H.C., R.C., D.V., T.B.S., X.L., D.E.W., E.G., Y.G., J.S.M.P., S.R., A.R., E.C.H., and G.J.D.S. analyzed data; and J.B., M.I.N., S.R., A.R., E.C.H., and G.J.D.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (See *SI Appendix* for accession nos.)

¹To whom correspondence may be addressed. E-mail: gavin.smith@duke-nus.edu.sg or justin.bahl@duke-nus.edu.sg.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1109314108/-DCSupplemental.

Results

Virus Surveillance and Dataset Design. We sequenced 105 full genomes of human H3N2 viruses isolated during 2004 to 2005 in Hong Kong (GenBank accession numbers provided in *SI Appendix*). To test the source-sink model of human influenza ecology, we randomly selected 75 sequences from Europe (from a total of 142 sequences available in GenBank), New York (211 sequences), Japan (193 sequences), Australia (170 sequences), New Zealand (247 sequences), Hong Kong (182 sequences), and Southeast Asia (130 sequences) for analysis. The latter consisted

of virus sequence data from Cambodia, Indonesia, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Vietnam. See *SI Appendix*, Figs. S1 and S2 for sampling frequency in each geographic region and for each randomized data set tested.

Global Phylogeny and Geographic Signal. Phylogenetic analysis of the HA1 domain showed viruses isolated from the same year and region tended to cluster together, but with frequent mixing with those from other regions (Fig. 1). To formally test if sequences were more strongly clustered by geographical location than expected by chance alone, and if this effect was supported statistically, given uncertainty in tree topologies, we used a phylogenetic trait-association test (15). This analysis revealed strong and statistically significant geographic structuring for each location ($P < 0.001$), suggesting localized annual epidemics (Table 1).

Region-Specific Evolutionary Behavior. It was evident from the phylogenetic analyses that epidemics in Europe, New York, Japan, Australia, and New Zealand (temperate regions) experienced significant annual population bottlenecks (Fig. 2A). Specifically, the extensive seasonal outbreaks in these regions generated numerous lineages but very few persist locally through time. For example, most lineages went extinct at the end of the New York 2001–2002 seasonal epidemic, except for individual viruses that were detected in New York 2002–2003 (Fig. 2A). Furthermore, no persistent lineages were detected from the New York 2002–2003, 2003–2004, and 2004–2005 epidemics, thereby highlighting the strong seasonal bottlenecks in temperate regions. Moreover, during these temperate epidemics we found that the majority of lineages coalesced 6 to 8 mo before the outbreaks, consistent with the continual introduction of diverse viruses (5) (Fig. 2A).

In contrast, multiple lineages cocirculated in both Hong Kong and Southeast Asia, often with a common ancestor that existed 1 to 2 y before virus sampling, thereby providing evidence of some long-term persistence (Fig. 2A). However, this observation may be, to some extent, an artifact of less rigorous sampling in Hong Kong and Southeast Asia relative to temperate regions, and the long branches, indicative of persistence, could represent unsampled diversity or introductions from other regions (16).

Seasonal Fluctuations in Relative Diversity of Different Geographic Regions.

Bayesian “skyride” reconstructions (17) demonstrated strong seasonal periodicity in relative genetic diversity in temperate zones, with major fluctuations through time (Fig. 2B). In comparison, in Southeast Asia and Hong Kong we observe lower levels of relative genetic diversity of influenza than in temperate regions (Fig. 2B), although in the latter regions there is some biannual fluctuation in diversity that corresponds with both temperate epidemics. We assume that the strength of natural selection, virus mutation rates, and generation times are broadly similar among all human populations, regardless of geographic location, such that any differences observed between regions likely represent differences in viral population size. Hence, these results suggest that influenza transmission in Hong Kong and Southeast Asia was not as extensive as that observed in temperate regions, which are characterized by local epidemics and widespread transmission that generates and sustains greater levels of relative genetic diversity throughout the epidemic.

There are two possible explanations for this pattern: either virus populations are smaller in Hong Kong and Southeast Asia or viruses are repeatedly introduced into Hong Kong and Southeast Asia (where they are not sustained) from other geographic regions experiencing epidemics. Our phylogenetic analysis strongly supports common ancestry of viruses isolated in Hong Kong and Southeast Asia with viruses from other regions. This linkage supports a model of repeated introductions rather than local circulation (Fig. 1). Regardless, the low levels of relative genetic diversity observed make it unlikely that populations in Hong Kong or Southeast Asia are able to support a genetically isolated, locally persistent source population for global influenza epidemics during the period of sampling.

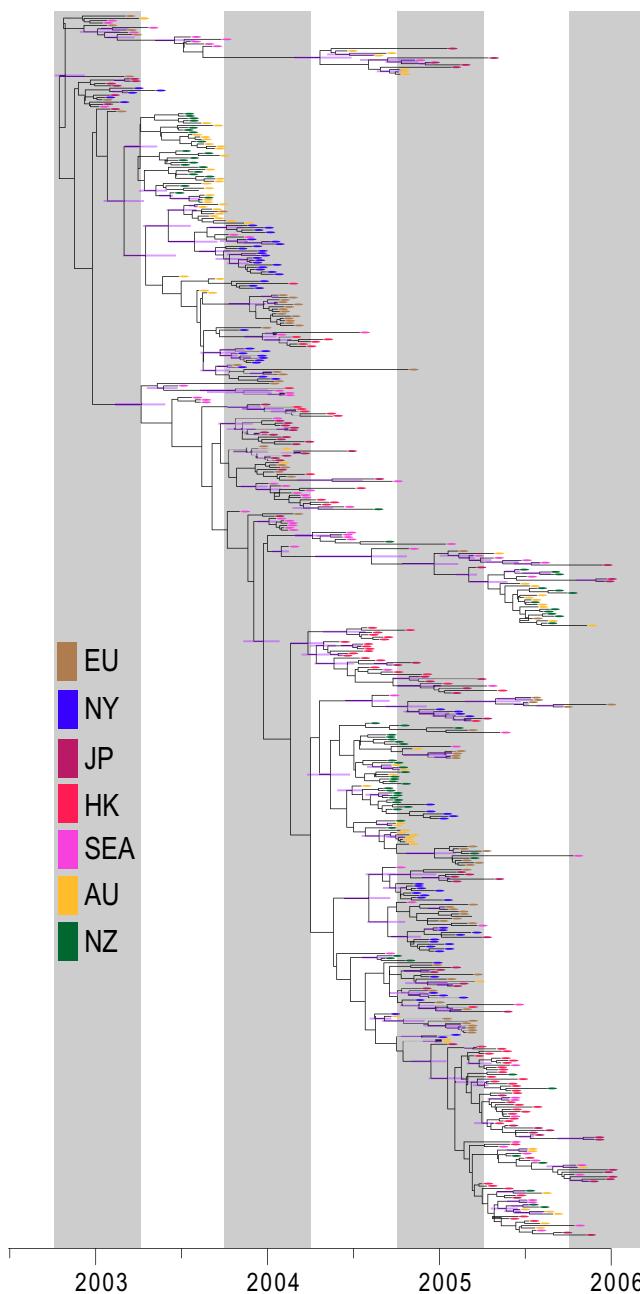


Fig. 1. Temporally structured maximum clade credibility phylogenetic tree showing the mixing of H3N2 influenza A virus global populations. This tree is representative of Bayesian sampled trees used to determine geographic structuring at tree tips for isolates sampled in Australia (AU, yellow), Europe (EU, brown), Hong Kong (HK, red), Japan (JP, purple), New York (NY, blue), New Zealand (NZ, green), and Southeast Asia (SEA, pink). Gray bars indicate the northern temperate epidemic season. See *SI Appendix*, Fig. S3 for virus names.

Table 1. Statistical analysis of geographic structuring for sampled H3N2 virus isolates

| Statistic | Observed mean | Lower 95% CI | Upper 95% CI | Null mean | Lower 95% CI | Upper 95% CI | P value |
|-----------|---------------|--------------|--------------|-----------|--------------|--------------|---------|
| MC (AU) | 5.26 | 4.0 | 7.0 | 2.20 | 1.90 | 2.93 | 0.002 |
| MC (EU) | 14.05 | 14.0 | 14.0 | 2.19 | 1.89 | 2.89 | 0.001 |
| MC (HK) | 7.21 | 7.0 | 8.0 | 2.18 | 1.87 | 2.95 | 0.001 |
| MC (JP) | 10.48 | 6.0 | 11.0 | 2.20 | 1.92 | 2.98 | 0.001 |
| MC (NY) | 13.65 | 13.0 | 17.0 | 2.18 | 1.89 | 2.78 | 0.001 |
| MC (NZ) | 4.99 | 3.0 | 8.0 | 2.19 | 1.89 | 2.98 | 0.001 |
| MC (SEA) | 6.04 | 5.0 | 7.0 | 2.19 | 1.89 | 2.96 | 0.001 |

AU, Australia; EU, Europe; HK, Hong Kong; JP, Japan; NY, New York; NZ, New Zealand; and SEA, Southeast Asia.

Viral Migration and Persistence Through Discrete Geographic Regions.

We tested three different models of viral migration using phylogeographic trait reconstruction (18) to assess the potential persistence of influenza A virus in Southeast Asia and Hong Kong: (i) we assumed there was no seasonality in Southeast Asia and Hong Kong (Fig. 3A); (ii) we assumed biannual epidemics in Southeast Asia and Hong Kong that correspond to temperate seasonal epidemics (Fig. 3B); and (iii) we grouped regions into epidemic zones (Northern, Southern and Tropical), treating Hong Kong plus Southeast Asia as a single discrete character with seasonality corresponding to northern and southern temperate epidemics (Fig. 3C). Although each analysis performed well, the second model possibly suffered from over-parameterization, suggesting that the additional parameters of seasonality in Southeast Asia and Hong Kong are not appropriate, consistent with epidemiological data (8). As a result, it was difficult to recover consistent state transitions between each randomized dataset for the more complicated models, even when

Bayesian sampling exceeded 100 million generations per run (Fig. 3B). Our results showed frequent two-way migration between temperate and tropical Asian regions, as well as direct viral movement between temperate regions. Critically, each model supported populations of viruses migrating between regions with no persistence in Southeast Asia or Hong Kong.

To formally test if one of these regions represents a global source population, we estimated the rate of viral migration between the seven geographic regions, assuming the regions and seasonal patterns as described in Fig. 3A. This analysis revealed that no single location seeded every annual epidemic in other locations; rather, multiple geographic regions occupied sections of the tree backbone (SI Appendix, Fig. S11). Evidence for viral migration among these regions was observed and these phylogeographic linkages were mapped to visualize transmission dynamics (Fig. 4). The highest rates of viral migration with statistical support [Bayes factor (BF) > 10] were observed consistently within a season (Table 2). No evidence of viral

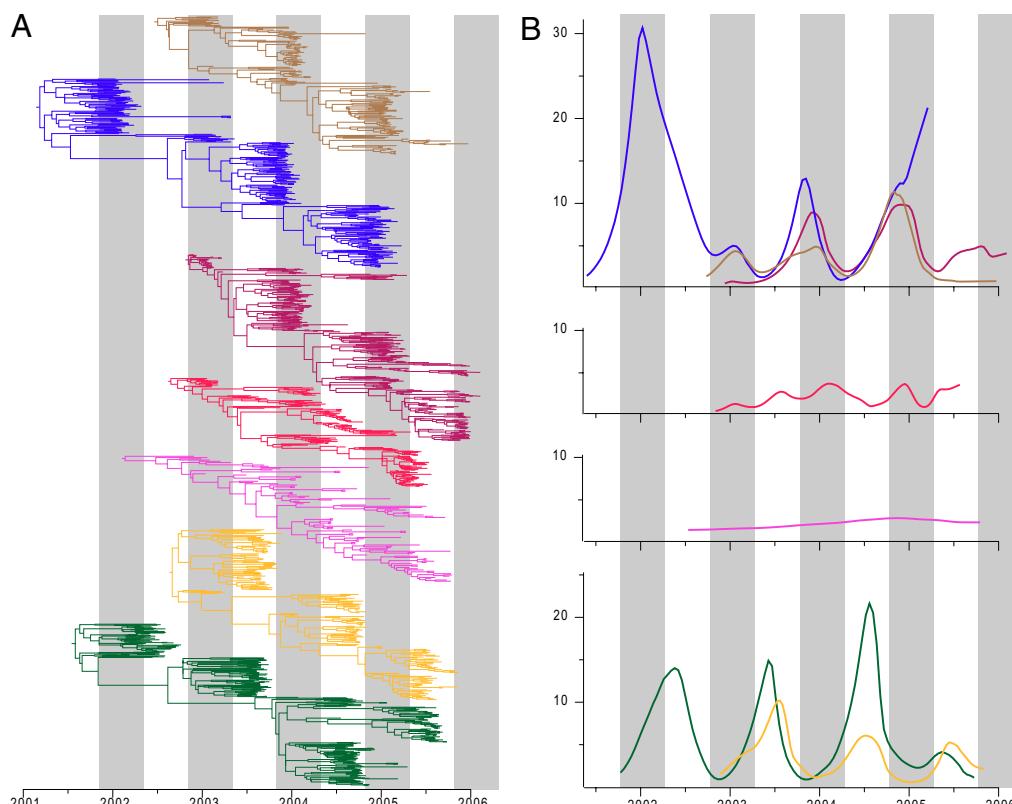


Fig. 2. Temporally structured phylogenetic and coalescent analysis of (A) phylogenies generated from all available sequence data of annual epidemics from viruses isolated in each location studied, and (B) Bayesian skyline analysis depicting fluctuating levels of relative genetic diversity from 2002 to 2006 of each location. The x axis indicates time from youngest sampled sequence to the lower 95% confidence interval (CI) of the tree-root height, and the y axis indicates relative genetic diversity (N_e) as estimated from the skyline coalescent analysis. Location of virus isolation is indicated by color (see legend to Fig. 1). See SI Appendix, Figs. S4, S5, S6, S7, S8, S9, and S10 for virus names and SI Appendix, Table S1 for sequence accession numbers.

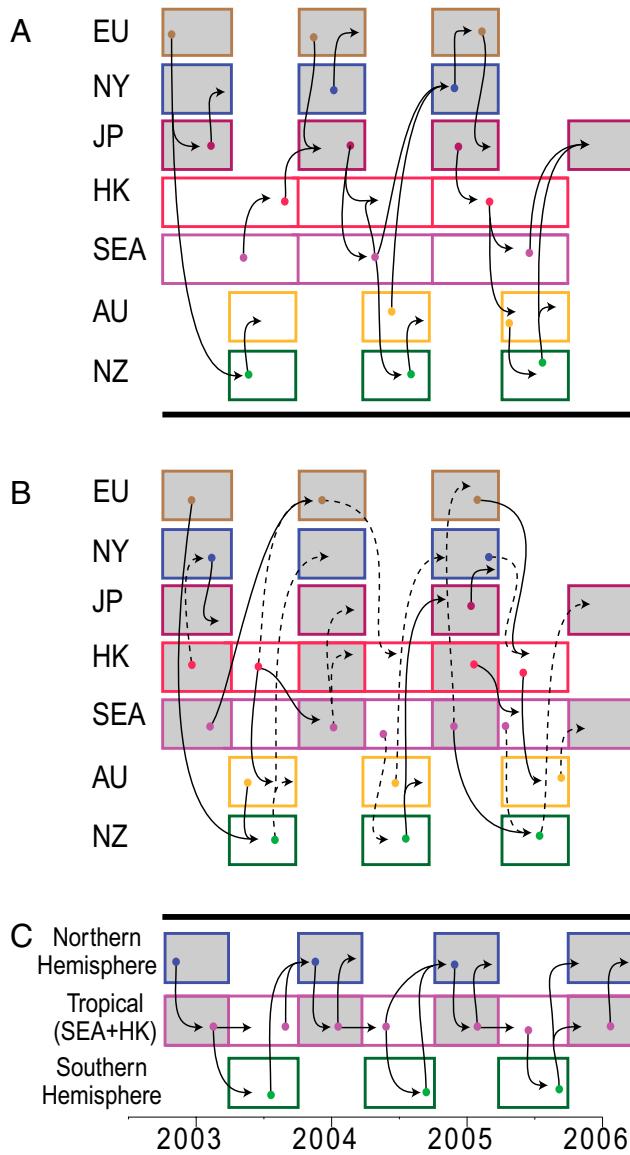


Fig. 3. Supported state transitions recovered from the combined independent Bayesian analyses after removal of burnin ($n = 15,000$ sampling events) for three randomized datasets (solid lines) indicating persistence of a virus metapopulation migrating between geographic regions. Dashed lines indicate state transitions that were supported in two of three randomized datasets. (A) Assumes no seasonality in Southeast Asia or Hong Kong; (B) assumes seasonality that corresponds to epidemics in temperate regions; and (C) groups the metapopulation into three epidemic zones. Location of virus isolation is indicated by color (see legend to Fig. 1).

persistence from one population through to another year in the same location (e.g., Hong Kong 2003 to Hong Kong 2004) was observed, suggesting that no single geographic region supported a separate population through time. However, our results do show strong support for viral migration to multiple regions each year. For example, viruses from Europe (2003), Japan and Southeast Asia (2004), and Hong Kong (2005) each migrated to two or more other regions (Fig. 4 and Table 2).

In addition, our results reveal that outbreaks in temperate regions could be seeded from multiple sources. Specifically, the 2005 Northern hemisphere epidemic was introduced to New York from Australia and Southeast Asia (Fig. 4 and Table 2). This epidemic in New York spread to Europe and then Japan. Direct connections of the European and Japanese epidemic with Southeast Asia or Hong Kong were not supported. The strongly

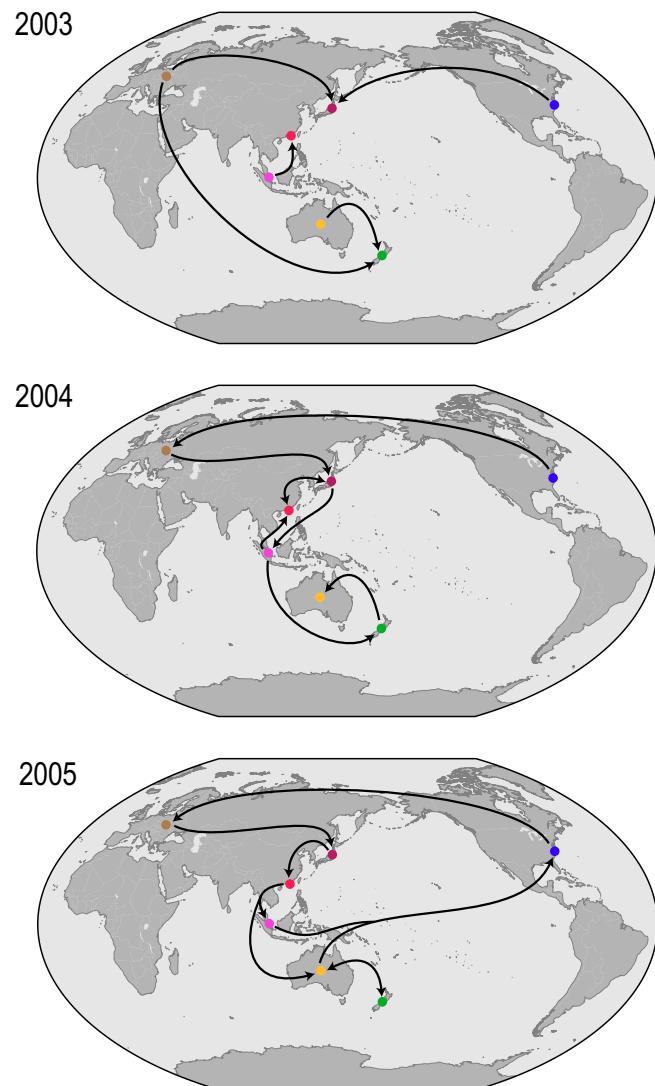


Fig. 4. Supported state transitions indicating migration events and directionality between discrete localities from 2003 to 2005 epidemic seasons visualized on a global map (Model 1, Fig. 3A). Geographic regions are indicated by color (see legend to Fig. 1).

supported migration rates from Europe to New Zealand (2003) and Australia (2005) to New York (2005), also show that the Northern epidemics can be seeded directly from Southern hemispheric regions and vice versa (Table 2).

Taken together, these results suggest that although viral populations migrating between Southeast Asia and Hong Kong may persist, maintenance of this population was dependent on migration from temperate epidemics, and that no single geographical location acted as global source population during the period of sampling, such that it is difficult to predict the global source of any specific outbreak.

Discussion

Our study reveals that, during the period 2003–2006, there was regular migration of populations of H3N2 influenza A virus between regions with coinciding epidemic peaks and connecting annual outbreaks in temporally offset epidemics in temperate zones. Several alternative models have been proposed to explain global dynamics of seasonal influenza, including a hypothetical source population in tropical regions, likely East or Southeast Asia, which seeds annual outbreaks in temperate regions (3, 11). In contrast, we show that migration into East and Southeast Asia

Table 2. Statistically supported migration rates between influenza populations

| 2003 | | | | 2004 | | | | 2005 | | | |
|------|----|------------------|-------|-----------|-----|------------------|-------|------------|-----|------------------|-------|
| From | To | Rate | BF* | From | To | Rate | BF* | From | To | Rate | BF* |
| NZ | AU | 1.72 (0.67–2.95) | >100 | NZ | AU | 1.15 (0.34–2.03) | >100 | JP | HK | 1.89 (0.62–3.34) | >100 |
| SEA | HK | 1.14 (0.21–2.21) | 46.72 | SEA | NZ | 0.53 (0.10–1.05) | 53.00 | NY | EU | 1.17 (0.44–2.00) | >100 |
| EU | JP | 0.67 (0.05–1.59) | 25.96 | NY | EU | 0.71 (0.19–1.31) | 24.08 | HK | SEA | 1.41 (0.32–2.56) | >100 |
| JP | NY | 1.31 (0.02–3.15) | 19.89 | JP | SEA | 1.34 (0.33–2.39) | 22.50 | AU (2004) | NY | 1.18 (0.11–2.33) | >100 |
| EU | NZ | 0.57 (0.02–1.29) | 16.48 | EU | JP | 1.03 (0.27–1.98) | 21.47 | EU | JP | 1.28 (0.38–2.24) | 46.17 |
| | | | | HK (2003) | JP | 1.63 (0.44–3.09) | 20.02 | NZ | AU | 1.83 (0.53–3.24) | 19.65 |
| | | | | SEA | HK | 0.64 (0.20–1.17) | 19.72 | HK | AU | 1.08 (0.27–1.98) | 19.35 |
| | | | | JP | HK | 1.13 (0.35–2.08) | 18.92 | AU | NZ | 1.66 (0.42–3.21) | 19.05 |
| | | | | | | | | SEA (2004) | NY | 1.00 (0.27–1.87) | 17.55 |

AU, Australia; EU, Europe; HK, Hong Kong; JP, Japan; NY, New York; NZ, New Zealand; and SEA, Southeast Asia.

*BF > 100 indicates decisive support, 30 ≤ BF ≤ 100 indicates very strong support, 10 ≤ BF ≤ 30 indicates strong support, and 6 ≤ BF ≤ 10 indicates substantial support for migration between locations. Only statistically supported migrations with indicator values > 0.50 are shown.

occurred throughout the time period studied and that this migration contributed to the local persistence in these regions, and that seasonal epidemics in temperate regions were seeded from a variety of geographical sources, including directly from other temperate outbreaks.

A suggested alternative model is that there are no restrictions for viral migration between temperate or tropical regions (i.e., an equal contacts model, in which China and Southeast Asia and the United States serve as primary hubs for viral migration) (12). However, this study did not account for strong annual population bottlenecks or differing seasonality across large countries (e.g., China), thereby treating the sampling country as the main indicator for estimating migration. Therefore, it is not possible to distinguish the annual contributions to the global influenza virus population of, for example, Hong Kong versus Shanghai or Beijing, which have different climatic conditions and hence might experience very different population dynamics. Although both studies support complex migration dynamics rather than a single source population restricted to one locality, the additional complexity included in our phylogeographic model reveals a continually migrating population.

Even though influenza virus circulates globally in humans, host populations are unevenly distributed within a network of flight-connected, highly dense urban centers (14). The discrete phylogeographic model used here shows that viruses from multiple geographic origins may seed influenza epidemics. The success of a given lineage through time is therefore most likely the result of a combination of competitive exclusion and stochastic factors, rather than its geographic origin alone. Although our study suggests that viral migration between urban centers does occur, the global persistence and re-emergence of annual epidemics in different regions are only likely to occur when viral migration into discrete regions occurs when the environment is suitable, based on factors that are still not entirely understood (i.e., absolute humidity, population density, and behavior) (19–21).

Although the absolute size of the global influenza population is large, the effective population size in any one region may be much smaller. Thus, by examining temperate regions where seasonal influenza epidemics are synchronized, it is possible to capture the full diversity of circulating lineages. However, tropical and subtropical regions do not have synchronized seasons. In source-sink ecological models the source population is expected to maintain high levels of genetic diversity through time such that it can reseed the sink populations that frequently undergo bottlenecks (22). It is possible that the lower levels of genetic diversity observed in Southeast Asia and Hong Kong may represent relatively stronger natural selection, perhaps resulting from a long, slow season in tropical regions, compared with short, fast seasons observed in temperate region. Although we found no evidence for differences in viral behavior in the tropics, the possible effect of geographic structuring with natural selection on phylogenetic structure clearly merits further study.

Here we show that no region included in our study maintains a sufficiently diverse pool of viruses to act as a source population. In contrast, the diversity of the global influenza population is maintained by a series of temporally overlapping epidemics coupled with high rates of migration. Although persistence is higher in regions with less seasonality, epidemics in Southeast Asia and Hong Kong are reseeded from elsewhere. Therefore, our results show it is a globally circulating metapopulation that ensures the long-term survival of influenza virus in humans.

A clear limitation of this analysis is that comprehensive data from many major urban centers is lacking and that we have considered a relatively short time span. In particular, the viral migration to and from major urban centers in continental Europe, Africa, South and Central Asia, and South America remains unknown. Although influenza surveillance is robust in many of these regions, insufficient gene sequencing coupled with associated metadata is available to understand the spatial and temporal evolutionary dynamics in detail. The complex metapopulation dynamics described here suggest that the connectedness of an urban center to other regions may dictate the timing and spatial dissemination of influenza epidemics. Current strategies for the control of influenza by vaccination are based on the biannual selection of vaccine candidates for the Northern and Southern hemispheres and require an understanding of genetic and antigenic variants circulating and the potential for new variants to emerge (23). Although such actions represent an effective vaccine strategy in temperate regions, the complex global dynamics we observe suggest that efforts to control influenza should include region-specific strategies to advance the current global policy.

Materials and Methods

Virus Isolation and Sequencing. The Clinical Virology Unit at Queen Mary Hospital, Hong Kong, is one of the three clinical virology laboratories in Hong Kong and serves the hospitals on Hong Kong Island. The specimens studied here were predominantly those from children hospitalized with acute respiratory illness. The influenza virus isolates were all collected as part of routine clinical work up. In this study we sequenced 105 full genomes of human H3N2 viruses isolated from 2004 to 2005. These sequences were analyzed with additional sequences from GenBank.

Phylogenetic and Coalescent Analysis. We analyzed 1,266 H3-HA sequences ([SI Appendix, Table S1](#)). Shows National Center for Biotechnology Information accession numbers) with at least 900 nucleotides sampled from Europe, New York State in the United States, Japan, Australia, New Zealand, and Southeast Asia (after removing identical sequences). These regions were selected based on the availability of data from well-characterized epidemics (with exact date of isolation) and human migratory networks, including East and Southeast Asia. Therefore, this analysis was limited to a variety of geographical locations from the period 2003–2006, all of which are publicly available via the Influenza Virus Resource at National Center for Biotechnology Information (24). For computational tractability we analyzed three randomly selected datasets with 75 virus HA1 sequences (987 nucleotides) from each region. Our final HA dataset comprised 525 taxa

from seven localities, including Hong Kong, with temporally overlapping influenza epidemics.

We used time-stamped sequence data with a relaxed-clock Bayesian Markov-chain Monte Carlo method as implemented in BEAST v1.6.2 for phylogenetic analysis (25, 26). For all analyses we used the uncorrelated lognormal relaxed molecular clock to accommodate variation in molecular evolutionary rate among lineages, the SRD06 codon position model, with a different rate of nucleotide substitution for the first plus second versus the third codon position, and the HKY85 substitution model then applied to these codon divisions (27), except where noted below. This analysis was conducted with a time-aware linear Bayesian skyride coalescent tree prior over the unknown tree space, with relatively uninformative priors on all model parameters except, a normal prior of 0.0031 substitutions per site per year (SD 0.001) on the mean substitution rate and a normal prior on the mean skyride size (log units) of 11.0 (SD 1.8). We performed at least three independent analyses of 50-million generations. These analyses were combined after the removal of an appropriate burn-in (10–20% of the samples in most cases), with 5,000 generations sampled from each run for a total of 15,000 trees and parameter estimates.

Tests of Geographic Association at Tree Tips. To determine the extent to which the distributed H3N2 influenza populations were geographically structured, based on location the virus isolate was sampled, we performed a phylogenetic-trait association analysis using the posterior distribution of trees produced by BEAST (see above). These geographic regions were coded onto the tips of 5,000 posterior sampled trees and these trees were analyzed using the maximum monophyletic clade size (MC) statistic available within the BaTS program and using 1,000 randomizations (15). This statistic determines the association between sampling location and phylogeny by estimating the size of the largest cluster of sequences sampled from each sampling location. Significance for this test was determined with a critical $P < 0.01$ and $0.01 \leq P \leq 0.05$ marginally significant, and $P > 0.05$ not significant.

Bayesian Skyride Reconstruction of Past Population Dynamics. The annual changing patterns of relative genetic diversity in viral isolates within each locality were compared by Bayesian skyride analyses (17). All available data from each geographical location studied here, was analyzed separately. The Bayesian skyride model uses a piecewise demographic model with a Gaussian Markov random field smoothing prior as described above (17).

Estimates of Viral Migration Through Discrete Geographic Regions. We used a nonreversible continuous-time Markov-chain model to estimate the migration between geographic regions and the general patterns of global circulation of influenza A virus (18). This analysis was restricted to the seven previously identified geographic regions. A constant-population coalescent process prior over the phylogeny and a simplified model of nucleotide

substitution (HKY85) was used in this analysis. Previous reports suggest less restrictive demographic priors have little effect on the phylogeographic inference (18). In this test we assumed each geographic region and influenza season was a discrete character; for example, an isolate collected from October 2002 to February 2003 in New York was assigned the discrete character, "New York 2003," resulting in a total of 22 location/season states. These discrete characters were subsequently mapped onto the internal nodes of the phylogenetic trees sampled during the Bayesian analysis. Given the large number of states, a Bayesian stochastic search variable selection (BSSVS) was used to reduce the number of parameters to those with significantly nonzero transition rates (18). The BSSVS explores and efficiently reduces the state space by using a binary indicator (I) (18). The BSSVS also calculates a BF, allowing the support for individual transitions between location/season states to be assessed. The BF is a function of I and was deemed statistically significant where $I > 0.5$ and the BF > 6 . Supported state transition where the mean I was <0.5 were indicated with a dashed line. In this analysis, statistical supports with $BF > 100$ indicate decisive support, whereas $6 \leq BF \leq 10$, $10 \leq BF \leq 30$, and $30 \leq BF \leq 100$ indicate substantial, strong, and very strong statistical support, respectively.

To further assess if virus populations in any region were persistent, we assumed there was seasonality in Southeast Asia and Hong Kong to discriminate between migration events corresponding to and from the northern and southern epidemic seasons and to assess persistence of viral populations in Southeast Asia and Hong Kong. This analysis assumed there were 29 location/season states. We also grouped regions into epidemic zones (Northern, Southern, and Tropical), again assuming seasonality in the tropical zone (Southeast Asia + Hong Kong), corresponding northern temperate (Europe + New York + Japan), and southern temperate (Australia + New Zealand) annual epidemics (14 location/season states).

We found no evidence for reassortment—a potential confounder—in these analyses (SI Appendix, Figs. S12, S13, and S14).

ACKNOWLEDGMENTS. This study was supported by Contracts HHSN266200700005C and HHSN272200900007 for the Influenza Genome Sequencing Project of the National Institute of Allergy and Infectious Disease, National Institutes of Health, Department of Health and Human Services; the Area of Excellence Scheme of the University Grants Committee (Grant AoE/M-12/06) of the Hong Kong Special Administrative Region Government; the RAPIDD program of the Science and Technology Directorate, US Department of Homeland Security and the Fogarty International Center, National Institutes of Health (S.R.), a career development award under National Institute of Allergy and Infectious Diseases Contract HHSN266200700005C (to G.J.D.S.); and the Duke–National University of Singapore Signature Research Program funded by the Agency for Science, Technology and Research, Singapore, and the Ministry of Health, Singapore (J.B., D.V., and G.J.D.S.).

1. Nelson MI, Holmes EC (2007) The evolution of epidemic influenza. *Nat Rev Genet* 8: 196–205.
2. Nelson MI, Simonsen L, Viboud C, Miller MA, Holmes EC (2007) Phylogenetic analysis reveals the global migration of seasonal influenza A viruses. *PLoS Pathog* 3:1220–1228.
3. Rambaut A, et al. (2008) The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453:615–619.
4. Koelle K, Cobey S, Grenfell B, Pascual M (2006) Epochal evolution shapes the phylogenetics of interpandemic influenza A (H3N2) in humans. *Science* 314:1898–1903.
5. Nelson MI, et al. (2006) Stochastic processes are key determinants of short-term evolution in influenza A virus. *PLoS Pathog* 2(12):e125.
6. Viboud C, Alonso WJ, Simonsen L (2006) Influenza in tropical regions. *PLoS Med* 3(4): e89.
7. Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat. (2009) World Urbanization Prospects: The 2009 Revision Population Database. <http://esa.un.org/wup2009/unup/>. Accessed October 25, 2011.
8. Wong CM, et al. (2006) Influenza-associated hospitalization in a subtropical city. *PLoS Med* 3(4):e121.
9. Chew FT, Doraisingham S, Ling AE, Kumarasinghe G, Lee BW (1998) Seasonal trends of viral respiratory tract infections in the tropics. *Epidemiol Infect* 121:121–128.
10. Lee VJ, et al. (2007) Influenza pandemics in Singapore, a tropical, globally connected city. *Emerg Infect Dis* 13:1052–1057.
11. Russell CA, et al. (2008) The global circulation of seasonal influenza A (H3N2) viruses. *Science* 320:340–346.
12. Bedford T, Cobey S, Beerli P, Pascual M (2010) Global migration dynamics underlie evolution and persistence of human influenza A (H3N2). *PLoS Pathog* 6:e1000918.
13. World Health Organization (2003) Severe acute respiratory syndrome (SARS): Status of the outbreak and lessons for the immediate future. Available at http://www.who.int/csr/media/sars_wha.pdf. Accessed October 25, 2011.
14. Hufnagel L, Brockmann D, Geisel T (2004) Forecast and control of epidemics in a globalized world. *Proc Natl Acad Sci USA* 101:15124–15129.
15. Parker J, Rambaut A, Pybus OG (2008) Correlating viral phenotypes with phylogeny: Accounting for phylogenetic uncertainty. *Infect Genet Evol* 8:239–246.
16. Smith GJD, et al. (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459:1122–1125.
17. Minin VN, Bloomquist EW, Suchard MA (2008) Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol Biol Evol* 25:1459–1471.
18. Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian phylogeography finds its roots. *PLOS Comput Biol* 5(9):e1000520.
19. Shaman J, Kohn M (2009) Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc Natl Acad Sci USA* 106:3243–3248.
20. Tamerius J, et al. (2010) Global influenza seasonality: Reconciling patterns across temperate and tropical regions. *Environ Health Perspect* 119:439–445.
21. Lipsitch M, Viboud C (2009) Influenza seasonality: Lifting the fog. *Proc Natl Acad Sci USA* 106:3645–3646.
22. Pulliam HR (1988) Sources, sinks and population regulation. *Am Nat* 132:652–661.
23. Barr IG, et al. (2010) Epidemiological, antigenic and genetic characteristics of seasonal influenza A(H1N1), A(H3N2) and B influenza viruses: Basis for the WHO recommendation on the composition of influenza vaccines for use in the 2009–2010 Northern Hemisphere season. *Vaccine* 28:1156–1167.
24. Bao Y, et al. (2008) The influenza virus resource at the National Center for Biotechnology Information. *J Virol* 82:596–601.
25. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4(5):e88.
26. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.
27. Shapiro B, Rambaut A, Drummond AJ (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol Biol Evol* 23:7–9.

Temporally structured metapopulation dynamics and persistence of influenza A H3N2 virus in humans – Supporting Information

Appendix

Justin Bahl^{a,b*}, Martha I. Nelson^c, Kwok H. Chan^a, Rubing Chen^d, Dhanasekaran Vijaykrishna^{a,b}, Rebecca Halpin^e, Timothy Stockwell^e, Xudong Lin^e, David E. Wentworth^e, Elodie Ghedin^{e,f}, Yi Guan^a, J. S. Malik Peiris^{a,g}, Steven Riley^{h,i}, Andrew Rambaut^{c,j}, Edward C. Holmes^{c,k} and Gavin J.D. Smith^{a,b*}

^aState Key Laboratory of Emerging Infectious Diseases & Department of Microbiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 21 Sassoon Road, Pokfulam, Hong Kong SAR, China; ^bLaboratory of Virus Evolution, Program in Emerging Infectious Diseases, Duke-NUS Graduate Medical School, 8 College Rd, Singapore, 169857; ^cFogarty International Center, National Institutes of Health, Bethesda, Maryland 20892, United States Of America; ^dUniversity of Texas Medical Branch, Galveston, Texas, United States Of America; ^eJ. Craig Venter Institute, Rockville, Maryland, United States Of America; ^fCenter for Vaccine Research, Department of Computational & Systems Biology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, United States Of America; ^gHKU-Pasteur Research Centre, The University of Hong Kong, Pokfulam, Hong Kong SAR, China; ^hDepartment of Community Medicine & School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong SAR, China; ⁱMedical Research Council Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College, London, U.K. ; ^jInstitute of Evolutionary Biology, University of Edinburgh, Edinburgh, EH9 3JT, U.K.; ^kCenter for

Infectious Disease Dynamics, Department of Biology, The Pennsylvania State University,
University Park, Pennsylvania, 16802, United States Of America

* To whom correspondence should be addressed: email justin.bahl@duke-nus.edu.sg,
gavin.smith@duke.nus.edu.sg

Appendix contents

| | |
|---|----------------|
| 1.0 Detailed Supplementary Materials and Methods | Page 4 |
| 1.1 Virus isolation, sequencing and GenBank Accession numbers | Page 4 |
| 1.2 Phylogenetic and coalescent analysis | Page 5 |
| 1.3 Tests of geographic association at tree tips | Page 6 |
| 1.4 Bayesian skyline reconstruction of past population dynamics | Page 7 |
| 1.5 Estimates of viral migration rates and persistence through discrete geographic regions | Page 8 |
| 1.6 Statistical comparison of genomic phylogenies for reassortment | Page 9 |
| 2.0 References | Page 11 |

Figures and Tables.

| | |
|---|----------------|
| Table S1: Country and GenBank Accession numbers for additional sequences from public databases used in this study | Page 13 |
| Fig. S1. Distribution of virus sampling dates for each location tested. | Page 19 |
| Fig. S2. Distribution of virus sampling dates for each randomly generated dataset | Page 20 |
| Fig. S3. Temporally structured maximum clade credibility phylogenetic tree showing the mixing of H3N2 influenza A virus global populations. This tree is representative of the three randomized datasets. Purple bars on nodes indicated 95% confidence intervals of date estimates. | Page 21 |
| Figs. S4-S10. Temporally structured maximum clade credibility phylogenetic tree showing the phylogenetic trees with taxon names used for the coalescent analysis of each location. Purple bars on nodes indicated 95% confidence intervals of date estimates. | Page 22 |

| | |
|---------------------------------|----------------|
| Fig. S5) Japan | Page 23 |
| Fig. S6) Hong Kong | Page 24 |
| Fig. S7) South East Asia | Page 25 |
| Fig. S8) Australia | Page 26 |
| Fig. S9) New Zealand | Page 27 |
| Fig. S10) Europe | Page 28 |

Fig S11. Phylogenetic tree generated from the discrete phylogeographic analysis showing ancestral state changes at tree nodes recovered from the sampled trees for isolates sampled in each discrete location and influenza season. Taxon names are indicated in Fig S1. The color key indicates the ancestral state location mapped onto the trees. **Page 29**

Figs. S12-S14. Bayesian skyride analysis of influenza genomes depicting fluctuating levels of relative genetic diversity from isolates sampled in New York, New Zealand and Hong Kong. The x-axis indicates time from the youngest sampled sequence to the lower 95% confidence interval of the tree root height, while the y-axis indicates relative genetic diversity (N_{et}) as estimated from the skyride coalescent analysis.

| | |
|-----------------------------|----------------|
| Fig S12) New York | Page 30 |
| Fig S13) New Zealand | Page 31 |
| Fig S14) Hong Kong. | Page 32 |

1.0 Detailed Supplementary Materials and Methods

1.1 Virus isolation and sequencing

The specimens studied here were collected as part of routine clinical work up and predominantly those from children hospitalized with acute respiratory. They were collected at the Clinical Virology Unit at Queen Mary Hospital, Hong Kong, one of the three clinical virology laboratories in Hong Kong. We sequenced 105 full genomes of human H3N2 viruses isolated from 2004-2005 (GenBank Accession numbers CY038567-CY038758, CY038935-CY039078, CY039159-CY039254, CY039487-CY039526, CY040298-CY040361, CY043744-CY043775, CY044397-CY044404 and CY100137-CY100400). These sequences were analyzed with additional sequences from GenBank. Isolate names with GenBank accession numbers are shown in the supplementary phylogenetic trees for each additional sequence used here.

All influenza isolates were grown in MDCK cells and identified as influenza A by immunofluorescence with specific monoclonal antisera and subtyped by the hemagglutination assay (HA) using antisera supplied by the WHO Influenza virus Identification kit (1). RNA extraction was performed as previously described using QIagen RNA extraction kits and the RNA was reverse transcribed and amplified by multisegment RT-PCR DNA concentration was determined using a Nanodrop Spectrophotometer (2). Sequencing reactions were performed using Big Dye Terminator chemistry (Applied Biosystems, Foster City, CA). Additional RT-PCR and sequencing was performed to close gaps and to increase coverage in low coverage or ambiguous regions. Sequencing reactions were analyzed on a 3730xl ABI sequencer and sequences were assembled in a software pipeline developed specifically for this project.

1.2 Phylogenetic and coalescent analysis

To examine whether tropical Asian countries (including Southeast Asia and Hong Kong) were the source of seasonal epidemics in both the northern and southern temperate regions, we conducted an expansive phylogenetic analysis of publically available HA1 gene sequences. We analyzed 1275 H3-HA sequences (with at least 900 nucleotides), sampled from Europe (142 sequences), New York state, USA (211 sequences), Japan (193 sequences), Australia (170 sequences), New Zealand (247 sequences), Hong Kong (182 sequences) and Southeast Asia (130 sequences) (SEA), the latter comprising Cambodia, Indonesia, Vietnam, Malaysia, Myanmar, Philippines, Thailand, and Singapore (See Supplementary Table S1 for accession numbers and Fig. S1 for distribution of virus isolation dates for each geographic location examined). These regions were chosen based on the availability of data from well-characterized epidemics and human migratory networks including East and Southeast Asia. Therefore, this analysis was limited to a variety of geographical locations from the period 2003-2006, all of which are publicly available via the NCBI Influenza Virus Resource at GenBank (3). For computational tractability we randomly selected (3x) 75 virus HA1 sequences (987 nucleotides) from seven discrete regions, including Hong Kong (See Fig. S2 for distribution of sampling dates for each randomized dataset). Our final HA data sets comprised 525 taxa from seven localities with temporally overlapping influenza epidemic seasons including East and Southeast Asia.

To generate temporal phylogenies using time-stamped sequence data we applied a relaxed-clock Bayesian Markov chain Monte Carlo method as implemented in BEAST v1.6.2 (4,5). This method allows variable nucleotide substitution rates among lineages and also incorporates phylogenetic uncertainty by sampling phylogenies and parameter estimates in proportion to their posterior probability (4). The uncorrelated lognormal relaxed molecular clock (ucld) was used for all analyses to accommodate variation in molecular evolutionary

rate amongst lineages. All analyses made use of the SRD06 codon position model, with a different rate of nucleotide substitution for the 1st plus 2nd versus the 3rd codon position, and the HKY85 substitution model then applied to these codon divisions (6). This analysis was also conducted with a time-aware linear Bayesian skyride coalescent tree prior over the unknown tree space and relatively uninformative priors on model parameters except, a normal prior of 0.0031 substitutions/site/year (standard deviation 0.001) on the mean substitution rate and a normal prior on the mean skyride size (log units) of 11.0 (standard deviation 1.8). We performed at least three independent analyses of 50 million generations, each sampled to produce 10,000 trees per data set to ensure adequate sample size of all analysis parameters including the posterior, prior, nucleotide substitution rates, and likelihoods (effective sample size > 200). These analyses were combined after the removal of an appropriate burn-in (10%–20% of the samples in most cases) with 5000 generations sampled from each run for a total of 15,000 trees and parameter estimates. The substitution rates, times to most recent common ancestry (TMRCA), and maximum clade credibility (MCC) phylogenetic trees then were calculated following visual inspection in TRACER version 1.5 (7). The maximum clade credibility phylogenetic tree generated from this analysis showing statistical supports for node age estimates and full taxa names is shown in supporting figure S1.

1.3 Tests of geographic association at tree tips

To determine the extent to which the distributed H3N2 influenza population was geographically structured we performed a phylogenetic trait association analysis using the posterior distribution of trees produced by BEAST (see above). The geographic traits tested were those sequences isolated from each locality (New York, Japan, Australia, New Zealand, Hong Kong, and Southeast Asia). These geographic regions were coded onto the tips of the

final 5000 posterior sampled trees and these trees were analyzed using the maximum monophyletic clade size (MC) statistic available within BaTS program and employing 1000 randomizations (8). This statistic determines the association between sampling location and phylogeny by estimating the size of the largest cluster of sequences sampled from each sampling location. Significance for this test was determined with a critical $p < 0.01$ and $0.01 < p < 0.05$ marginally significant, and $p > 0.05$ not significant.

1.4 Bayesian skyride reconstruction of past population dynamics

The annual changing patterns of relative genetic diversity in viral isolates within each locality were compared by Bayesian skyride analyses (9). Each data set was analyzed separately from each of the six geographical locations. The Bayesian skyride model uses a piecewise demographic model with a Gaussian Markov random field (GMRF) smoothing prior. In contrast to Bayesian skyline methods (10), which require fairly strong prior information regarding the number of change points in a population history (i.e. number of groups), the GMRF prior is relatively uninformative (9). In a population genetic context the skyride coalescent analyses is an estimator of the product of the generation time (t) and the effective size of the population (N_e). For influenza A virus, this product can be interpreted as the effective number of infections averaged over time assuming a neutral evolutionary process. Because both natural selection and seasonality are strong in influenza virus, we interpret the skyride coalescent as relative genetic diversity (11). In our Bayesian coalescent analysis of the HA1 we assume that the virus mutation rates and generation times across populations of humans are similar, regardless of geographic location, such that any differences observed between regions must represent differences in population size. Therefore, it is appropriate to compare each geographic region by coalescent analysis.

Phylogenetic trees associated with the coalescent analysis showing statistical supports for node age estimates and full taxa names are shown in figures S2-S7.

1.5 Estimates of viral migration rates and persistence through discrete geographic regions

We used a non-reversible continuous-time Markov chain model to estimate the migration rates between geographical regions and the general patterns on global circulation of influenza A virus (12, 13). In this test we used a constant-population coalescent process prior over the phylogeny and a simplified model of nucleotide substitution (HKY85) with the UCLD relaxed molecular clock. Here we assumed each geographic region and influenza season was a discrete character. By defining the discrete characters in such a manner we were able to incorporate seasonality into the phylogeographic analysis. For example, an isolate collected from Oct 2002 – Feb 2003 in New York was assigned the discrete character, ‘New York 2003.’ Since reconstructions of Hong Kong and Southeast Asian influenza populations dynamics showed no evidence of seasonal bottlenecks, isolates sampled from these regions were assigned the location and year of isolation as discrete characters. This results in a total of 22 discrete states, which were subsequently mapped onto the internal nodes of phylogenetic trees sampled during the Bayesian analysis. Given the large number of states, a Bayesian stochastic search variable selection (BSSVS) was employed to reduce the number of parameters to those with significantly non-zero transition rates (12). For the BSSVS we assumed a Poisson prior, which assigns a 50% prior probability on the minimal rate configuration. With this scenario we have 21 non-zero rates connecting the 22 location/seasons and a mean Poisson prior of 0.693 used in this analysis. The BSSVS explores and efficiently reduces the state space by employing a binary indicator (I) (12, 13). The BSSVS also calculates a Bayes factor (BF) allowing the support for individual

transitions between location/season states to be assessed. The BF is a function of I and was deemed statistically significant where $I > 0.5$ and the $BF > 6$ in at least 2 out of three randomized datasets. Supported state transition where the mean I was < 0.5 were indicated with a dashed line. Our minimal critical cutoff for statistical supports were $6 \leq BF \leq 10$, indicating substantial support, $10 \leq BF \leq 30$ indicating strong support, $30 \leq BF \leq 100$ indicates very strong support and $BF > 100$ indicating decisive support (12 – 14).

These ancestral traits were then mapped onto the phylogenetic tree (Fig S8). This analysis revealed that no single location seeded every annual epidemic in other locations. Even though Southeast Asia and Hong Kong occupied substantial portions of the tree backbone, multiple geographic regions occupied various sections of the tree indicating extensive migration of the influenza population between these regions. Ancestral trait changes are indicated on the phylogenetic tree (Fig. S8).

1.6 Statistical comparison of genomic phylogenies for reassortment

Reassortment events may potentially confound the results of our analyses. Although the increased co-circulation of lineages provides for frequent reassortment events we found no evidence for reassortment during the study period. In particular, essentially identical skyride plots were observed for each segment from Hong Kong, New York and New Zealand, respectively (Fig. S9-S11). The phylogenetic and coalescent analyses of the full genomes from isolates sampled in New York and New Zealand were qualitatively identical, but this was less clear with Hong Kong isolates (Figures S8-S10). To determine if reassortment had occurred we searched the concatenated full genome sequence alignment of isolates from Hong Kong, New York and New Zealand by single breakpoint analysis as implemented in HyPhy (15). Conflicting phylogenetic signal from various breakpoints was tested by a Kishino-Hasegawa (KH) test against the combined tree as the null hypothesis. One thousand

bootstrap replicates of each breakpoint KH test was conducted in order to determine whether suspected breakpoints were statistically supported and assessed by the Akaike Information Criteria (AIC). No highly supported break-points were identified among concatenated genome segments, suggesting reassortment had not occurred in these regions during the 2003–2006 influenza epidemic seasons ($p < 0.01$).

2.0 References

1. World Health Organization (2007) Antigenic and genetic characteristics of H5N1 viruses and candidate H5N1 vaccine viruses developed for potential use as pre-pandemic vaccines. World Health Organization, Geneva, Switzerland.
2. Zhou B et al (2009) Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses. *J Virol* 83: 10309-10313
3. Bao Y et al (2008) The Influenza Virus Resource at the National Center for Biotechnology Information. *J. Virol* 82:596-601.
4. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4:e88
5. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7:214
6. Shapiro B, Rambaut A, Drummond AJ (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol Biol Evol* 23:7–9.
7. Rambaut A, Drummond AJ (2007) Tracer v1.4, Available from <http://beast.bio.ed.ac.uk/Tracer>
8. Parker J, Rambaut A, Pybus OG (2008) Correlating viral phenotypes with phylogeny: Accounting for phylogenetic uncertainty. *Infect Genet and Evol* 8:239-246
9. Minin VN, Bloomquist EW, Suchard MA (2008) Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol Biol Evol* 25:1459-1471
10. Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) *Mol Biol Evol* 22:1185-1192
11. Rambaut et al (2008) The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453:615-619

12. Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian phylogeography finds its roots. *PLoS Comp Biol* 5:e1000520
13. Edwards CJ et al (2011) Ancient hybridization and an Irish origin for the modern polar bear matriline. *Current Biology* 21:1251-1258
14. Jeffreys H (1961) The Theory of Probability (3e), Oxford University Press
15. Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21: 676-679

Table S1: Country and GenBank Accession numbers for additional sequences from public databases used in this study

| Australia | Europe | Hong Kong | Japan | New York | New Zealand | Southeast Asia |
|-----------|----------|-----------|----------|----------|-------------|----------------|
| DQ865950 | EU106663 | EU501254 | AB262301 | CY000001 | CY002905 | DQ865970 |
| DQ865953 | EU106664 | EU501255 | AB271503 | CY000129 | CY002906 | EF566332 |
| DQ865969 | EU106665 | EU501256 | AB271504 | CY000009 | CY002914 | EU501622 |
| CY015772 | EU106666 | EU501257 | AB271505 | CY000017 | CY002922 | FJ865282 |
| CY015780 | EU106667 | EU501258 | AB271506 | CY000025 | CY002930 | FJ865283 |
| CY015788 | EU106668 | EU501259 | AB271511 | CY000041 | CY002938 | FJ865284 |
| CY015796 | EU106670 | EU501260 | AB271514 | CY000057 | CY002946 | DQ865961 |
| CY015804 | EU106671 | EU501261 | AB271515 | CY000153 | CY002954 | DQ865962 |
| CY015812 | EU106672 | EU501262 | AB271516 | CY000161 | CY002962 | EF566155 |
| CY015820 | EU106674 | EU501264 | AB271517 | CY000169 | CY002904 | EF566164 |
| CY015828 | CY026043 | EU501265 | AB271524 | CY000177 | CY006923 | EF566173 |
| CY015836 | CY026044 | EU501421 | AB271525 | CY000065 | CY006931 | EF566365 |
| CY015844 | CY026046 | EU501422 | AB271526 | CY000081 | CY006939 | EF566051 |
| CY015852 | CY026047 | EU501668 | AB270992 | CY000089 | CY006955 | EU501161 |
| CY015860 | CY026049 | EU501670 | AB270999 | CY000097 | CY006963 | EU501281 |
| CY015876 | CY026052 | EU501671 | EU501222 | CY000105 | CY006971 | EU501283 |
| CY015884 | CY026054 | EU501672 | EU501223 | CY000121 | CY006995 | EU501458 |
| CY015900 | CY026055 | EU501674 | EU501268 | CY000193 | CY007003 | EU501459 |
| CY015908 | CY026058 | EU501675 | EU501269 | CY000145 | CY007011 | EU501460 |
| CY015916 | CY026059 | EU501676 | EU501274 | CY000249 | CY007075 | EU501461 |
| CY015924 | CY026061 | EU501677 | EU501294 | CY000257 | CY007083 | EU501462 |
| CY015932 | CY026062 | EU502208 | EU501300 | CY000265 | CY007131 | EU501463 |
| CY015956 | CY026063 | EU502209 | EU501316 | CY000345 | CY007267 | EU501464 |
| CY015964 | CY026064 | EU856814 | EU501340 | CY000353 | CY007275 | EU501465 |
| CY015980 | CY026067 | EU856822 | EU501363 | CY000361 | CY007291 | EU501466 |
| CY015988 | CY026072 | EU856825 | EU501364 | CY000369 | CY007307 | EU501467 |
| CY015996 | CY026074 | EU856837 | EU501365 | CY000033 | CY007331 | CY091437 |
| CY016004 | CY026075 | EU856841 | EU501367 | CY000377 | CY007347 | CY091445 |
| CY016012 | CY026076 | EU856842 | EU501387 | CY000473 | CY007371 | CY091453 |
| CY016028 | CY026077 | EU856843 | EU501388 | CY000505 | CY007379 | FJ229884 |
| CY016036 | CY026078 | EU856844 | EU501389 | CY000513 | CY007387 | DQ865957 |
| CY016979 | CY026081 | EU856854 | EU501401 | CY000521 | CY007403 | DQ865958 |
| CY017099 | CY026082 | EU856859 | EU501404 | CY000561 | CY007435 | EF566074 |
| CY017107 | CY026084 | EU856860 | EU501405 | CY000753 | CY007451 | EU501168 |
| CY017563 | CY026086 | EU856861 | EU501409 | CY000761 | CY007459 | EU501308 |
| CY017579 | CY026088 | EU856862 | EU501412 | CY000777 | CY007475 | EU501312 |
| CY017587 | CY026090 | EU856864 | EU501413 | CY000785 | CY007483 | EU514653 |
| CY017595 | CY026091 | EU856870 | EU501414 | CY000873 | CY007547 | EU514667 |
| CY017603 | CY026092 | EU856906 | EU501415 | CY000889 | CY007555 | DQ865951 |

| | | | | | | |
|----------|----------|----------|----------|----------|----------|----------|
| CY017611 | CY026093 | EU856909 | EU501416 | CY000881 | CY007563 | DQ865959 |
| CY017619 | CY026094 | EU856928 | EU501434 | CY000769 | CY007579 | EF566141 |
| CY017797 | CY026095 | EU856929 | EU501435 | CY000901 | CY007019 | EF566176 |
| CY017941 | CY026096 | EU856942 | EU501440 | CY000909 | CY007027 | EF566053 |
| CY017949 | CY026098 | EU856944 | EU501442 | CY000957 | CY007067 | EF566068 |
| CY017957 | CY026100 | EU856982 | EU501443 | CY000965 | CY007155 | EU501171 |
| CY017973 | CY026101 | EU856984 | EU501444 | CY000973 | CY007163 | EU501172 |
| CY017981 | CY026102 | EU856986 | EU501445 | CY001013 | CY007171 | EU501174 |
| CY017989 | CY026103 | EU856992 | EU501446 | CY000949 | CY007179 | EU501176 |
| CY018925 | CY026105 | EU856993 | EU501447 | CY000917 | CY007195 | EU501317 |
| CY018973 | CY026108 | EU856994 | EU501448 | CY001037 | CY007203 | EU501318 |
| CY018981 | CY026109 | EU856995 | EU501470 | CY001064 | CY007211 | EU501319 |
| CY018989 | CY026110 | EU856996 | EU501471 | CY001021 | CY007227 | EU501320 |
| CY019005 | CY026114 | EU856997 | EU501475 | CY001088 | CY007235 | EU501321 |
| CY019021 | CY026115 | EU857007 | EU501496 | CY000049 | CY007243 | EU501541 |
| CY019029 | CY026116 | EU857009 | EU501497 | CY001160 | CY007251 | EU501542 |
| CY019747 | CY026117 | EU857010 | EU501498 | CY001053 | CY007283 | EU501544 |
| EF467800 | CY026118 | EU857011 | EU501499 | CY001061 | CY007315 | EU501545 |
| CY020005 | CY026119 | EU857021 | EU501500 | CY001205 | CY007339 | EU501546 |
| CY020029 | CY026121 | EU857023 | EU501513 | CY001221 | CY007395 | EU501547 |
| CY020045 | CY026123 | EU857024 | EU501528 | CY001229 | CY007411 | EU501548 |
| EF566130 | CY026124 | EU857026 | EU501529 | CY001045 | CY007419 | EU501773 |
| EF566132 | CY026125 | EU857027 | EU501532 | CY000137 | CY007427 | EU501774 |
| EF566143 | CY026126 | EU857028 | EU501534 | CY001253 | CY007515 | CY091149 |
| EF566144 | CY026133 | EU857029 | EU501535 | CY001285 | CY007523 | CY091157 |
| EF566148 | CY026134 | EU857030 | EU501536 | CY001405 | CY007539 | CY091165 |
| EF566149 | CY026138 | EU857035 | EU501538 | CY001373 | CY007571 | CY091229 |
| EF566150 | EU501322 | EU857037 | EU501563 | CY001341 | CY007795 | CY091245 |
| EF566157 | EU501776 | EU857038 | EU501587 | CY001293 | CY007955 | CY091253 |
| EF566161 | EU502461 | EU857039 | EU501592 | CY001421 | CY007987 | CY091261 |
| EF566168 | EU502469 | EU857044 | EU501595 | CY001461 | CY008043 | CY091469 |
| EF566170 | EU502492 | EU857045 | EU501597 | CY001469 | CY008051 | CY091485 |
| EF566172 | EU501247 | EU857047 | EU501598 | CY001512 | CY008059 | DQ865945 |
| EF566177 | EU502502 | EU857051 | EU501607 | CY001552 | CY008067 | DQ865949 |
| EF566222 | EU502512 | EU857053 | EU501608 | CY001536 | CY008075 | DQ865955 |
| EF566230 | EF566231 | EU857054 | EU501626 | CY001544 | CY008083 | EF566142 |
| EF566239 | EU501277 | EU857055 | EU501645 | CY001560 | CY008107 | EF566224 |
| EF566241 | EU501453 | EU857058 | EU501646 | CY001624 | CY008212 | EF566227 |
| EF566254 | EU501495 | EU857059 | EU501647 | CY001648 | CY008236 | EF566229 |
| EF566295 | EU501527 | EU857060 | EU501652 | CY001640 | CY008364 | EF566361 |
| EF566305 | EU501710 | EU857061 | EU501653 | CY000073 | CY008372 | EF566362 |
| EF566306 | EU501711 | EU857063 | EU501656 | CY001712 | CY008404 | EF566067 |
| EF566366 | EU502325 | EU857066 | EU501657 | CY002000 | CY008436 | EU501124 |
| EF566367 | EU502490 | EU857067 | EU501658 | CY002008 | CY008396 | EU501125 |
| EF566368 | EU501609 | EU857068 | EU501660 | CY002016 | CY008356 | EU501126 |
| EF566019 | EU502200 | EU857071 | EU501661 | CY002040 | CY008348 | EU501127 |

| | | | | | | |
|----------|----------|----------|----------|----------|----------|----------|
| EF566058 | EU502462 | EU857072 | EU501662 | CY002048 | CY008340 | EU501177 |
| EF566060 | EU502478 | EU857073 | EU501663 | CY002024 | CY008252 | EU501221 |
| EU501208 | EU502480 | EU857079 | EU501664 | CY002056 | CY008228 | EU501279 |
| EU501209 | EU502485 | EU857081 | EU501665 | CY002032 | CY008636 | EU501372 |
| EU501210 | EU502506 | EU857082 | EU501666 | CY002064 | CY008628 | EU501373 |
| EU501211 | GQ161129 | EU857083 | EU501667 | CY002072 | CY008620 | EU501374 |
| EU501212 | GQ161136 | EU857084 | EU501680 | CY002080 | CY008612 | EU501375 |
| EU501214 | GQ161170 | EU857085 | EU501682 | CY002176 | CY008604 | EU501385 |
| EU501215 | EU502463 | EU857086 | EU501683 | CY002184 | CY008596 | EU501441 |
| EU501217 | EU716082 | EU857087 | EU501686 | CY002192 | CY008556 | EU501474 |
| EU501218 | EU716083 | EU857088 | EU501687 | CY002208 | CY008548 | EU501531 |
| EU501219 | EU716084 | EU857090 | EU501688 | CY002216 | CY008540 | EU501558 |
| EU501220 | EU716085 | EU857092 | EU501693 | CY002224 | CY008652 | EU501787 |
| EU501238 | EU502468 | EU857093 | EU501695 | CY002232 | CY009020 | EU501788 |
| EU501301 | EU502504 | EU857094 | EU501696 | CY002200 | CY009036 | EU501789 |
| EU501302 | DQ865971 | | EU501697 | CY002248 | CY009924 | EU501790 |
| EU501303 | EU502489 | | EU501698 | CY002240 | CY011728 | EU501791 |
| EU501306 | EU502497 | | EU501700 | CY002256 | CY011744 | EU501792 |
| EU501324 | EU502222 | | EU501701 | CY002280 | CY011752 | EU501793 |
| EU501326 | EU502470 | | EU501702 | CY002264 | CY011760 | EU501795 |
| EU501327 | EU502488 | | EU501703 | CY002288 | CY011768 | EU501800 |
| EU501347 | CY037351 | | EU501704 | CY000865 | CY012064 | EU502344 |
| EU501348 | EU502465 | | EU501706 | CY002344 | CY012072 | EU514639 |
| EU501351 | EU502467 | | EU501719 | CY002424 | CY012080 | EU514654 |
| EU501377 | AB462528 | | EU501720 | CY002432 | CY012088 | FJ561060 |
| EU501378 | AB462531 | | EU501721 | CY002408 | CY012096 | AB281195 |
| EU501379 | AB462534 | | EU501722 | CY002440 | CY012104 | AB281200 |
| EU501381 | AB462537 | | EU501724 | CY002416 | CY012112 | AB281205 |
| EU501382 | AB462540 | | EU501725 | CY002456 | CY012352 | AB281210 |
| EU501399 | AB462542 | | EU501727 | CY002464 | CY012360 | AB281215 |
| EU501522 | HQ166052 | | EU501728 | CY002472 | CY012368 | AB281232 |
| EU501523 | EF670520 | | EU501730 | CY002488 | CY012376 | AB281235 |
| EU501524 | EU502464 | | EU501732 | CY002480 | CY012384 | AB281238 |
| EU501551 | EU502332 | | EU501736 | CY002504 | CY012392 | AB281241 |
| EU501564 | EU502333 | | EU501737 | CY002520 | CY012400 | AB281244 |
| EU501567 | EU502510 | | EU501738 | CY002584 | CY012408 | AB281247 |
| EU501570 | EU502511 | | EU501739 | CY002592 | CY012416 | AB281250 |
| EU501571 | EU501291 | | EU501744 | CY002600 | CY012424 | |
| EU501572 | EU501292 | | EU501745 | CY002608 | CY012648 | |
| EU501573 | EU501451 | | EU501747 | CY002712 | CY012656 | |
| EU501574 | EU501517 | | EU501748 | CY002720 | CY012664 | |
| EU501575 | EU501591 | | EU501749 | CY002728 | CY012672 | |
| EU501611 | EU501644 | | EU501751 | CY002736 | CY012680 | |
| EU501612 | EU501699 | | EU501758 | CY002784 | CY012688 | |
| EU501613 | EU501726 | | EU501760 | CY002760 | CY012696 | |
| EU501615 | EU501779 | | EU501761 | CY002768 | CY012704 | |

| | | | | |
|----------|----------|----------|----------|----------|
| EU501616 | EU502498 | EU501762 | CY002776 | CY012712 |
| EU501617 | EU502505 | EU501764 | CY003032 | CY012720 |
| EU501618 | EU501246 | EU501765 | CY003056 | CY013120 |
| EU501641 | EU501494 | EU501766 | CY003040 | CY013136 |
| EU501753 | JF710724 | EU501767 | CY003048 | CY013144 |
| EU501755 | JF710725 | EU501768 | CY003072 | CY013152 |
| EU501778 | JF710738 | EU501770 | CY003344 | CY013160 |
| EU501786 | JF710756 | EU501772 | CY003408 | DQ865948 |
| EU501814 | JF710757 | EU501808 | CY003416 | DQ865965 |
| EU501823 | JF972562 | EU501810 | CY003640 | DQ865968 |
| EU501824 | | EU501816 | CY003648 | CY013429 |
| CY091181 | | EU501830 | CY003656 | CY013437 |
| CY091189 | | EU501839 | CY003664 | CY013445 |
| CY091205 | | EU501840 | CY003680 | CY013453 |
| CY091277 | | EU501841 | CY006076 | CY013461 |
| CY091285 | | EU501842 | CY006092 | CY013469 |
| CY091293 | | EU501845 | CY006115 | CY013485 |
| CY091317 | | EU501848 | CY006123 | CY013493 |
| CY091333 | | EU501851 | CY006131 | CY013501 |
| CY091341 | | EU502223 | CY006139 | CY013517 |
| CY091349 | | AB378391 | CY006147 | CY013525 |
| CY091365 | | AB378392 | CY006155 | CY013533 |
| CY091389 | | AB378393 | CY006179 | CY013541 |
| CY091397 | | AB378395 | CY006291 | CY013549 |
| CY091405 | | AB378399 | CY006371 | CY013509 |
| CY091413 | | AB378405 | CY006379 | CY013911 |
| | | AB378406 | CY006435 | CY013919 |
| | | AB378409 | CY006859 | CY013477 |
| | | AB378410 | CY007643 | CY013935 |
| | | AB378411 | CY008164 | CY013943 |
| | | AB378412 | CY003336 | CY013951 |
| | | AB378413 | CY008908 | CY013959 |
| | | AB378414 | CY008900 | CY013967 |
| | | AB378415 | CY008892 | CY013975 |
| | | AB378416 | CY008884 | CY013983 |
| | | AB378417 | CY008868 | CY013991 |
| | | AB378418 | CY008516 | CY014015 |
| | | AB378419 | CY008916 | CY014031 |
| | | AB378426 | CY009260 | CY014039 |
| | | AB378427 | CY009244 | CY014047 |
| | | AB378431 | CY009252 | CY014055 |
| | | AB378432 | CY009268 | CY014063 |
| | | AB378434 | CY019141 | CY014071 |
| | | AB378436 | CY019149 | CY014079 |
| | | AB378439 | CY019165 | CY014087 |
| | | AB378440 | CY019173 | CY014095 |

| | | |
|----------|----------|----------|
| AB378441 | CY019181 | CY014111 |
| AB378445 | CY019189 | CY014119 |
| AB378446 | CY019253 | CY014135 |
| AB378447 | CY019261 | CY015564 |
| AB378448 | CY019269 | CY015596 |
| AB378449 | CY019277 | CY015604 |
| AB378450 | CY019285 | CY015612 |
| AB378452 | CY019293 | CY015620 |
| | CY019301 | CY015628 |
| | CY019309 | CY016212 |
| | CY019317 | CY016443 |
| | CY019811 | CY016451 |
| | EF473613 | CY016467 |
| | EF473618 | CY016475 |
| | EF473619 | CY020341 |
| | EF473625 | CY020893 |
| | CY020533 | EF566129 |
| | EU501484 | EF566131 |
| | EU501485 | EF566134 |
| | EU501486 | EF566139 |
| | EU502299 | EF566154 |
| | EU502300 | EF566165 |
| | EU502301 | EF566166 |
| | EU502302 | EF566171 |
| | EU502303 | EF566182 |
| | EU502305 | EF566285 |
| | EU502306 | EF566363 |
| | EU502307 | EF566035 |
| | EU502309 | EF566045 |
| | EU502310 | EF566073 |
| | EU502313 | CY021965 |
| | EU502315 | CY022573 |
| | EU502316 | CY022597 |
| | | CY023082 |
| | | EU501205 |
| | | EU501206 |
| | | EU501225 |
| | | EU501226 |
| | | EU501227 |
| | | EU501228 |
| | | EU501231 |
| | | EU501234 |
| | | EU501235 |
| | | EU501240 |
| | | EU501241 |
| | | EU501243 |

EU501244
EU501357
EU501369
EU501370
EU501391
EU501392
EU501395
EU501396
EU501397
EU501521
EU501579
EU501637
EU501829
EU501831
EU501832
CY039960
CY039961
CY039962
CY039963
CY039964
CY039965
CY039966
CY096843

Fig.S1. Distribution of virus sampling dates for each location tested.

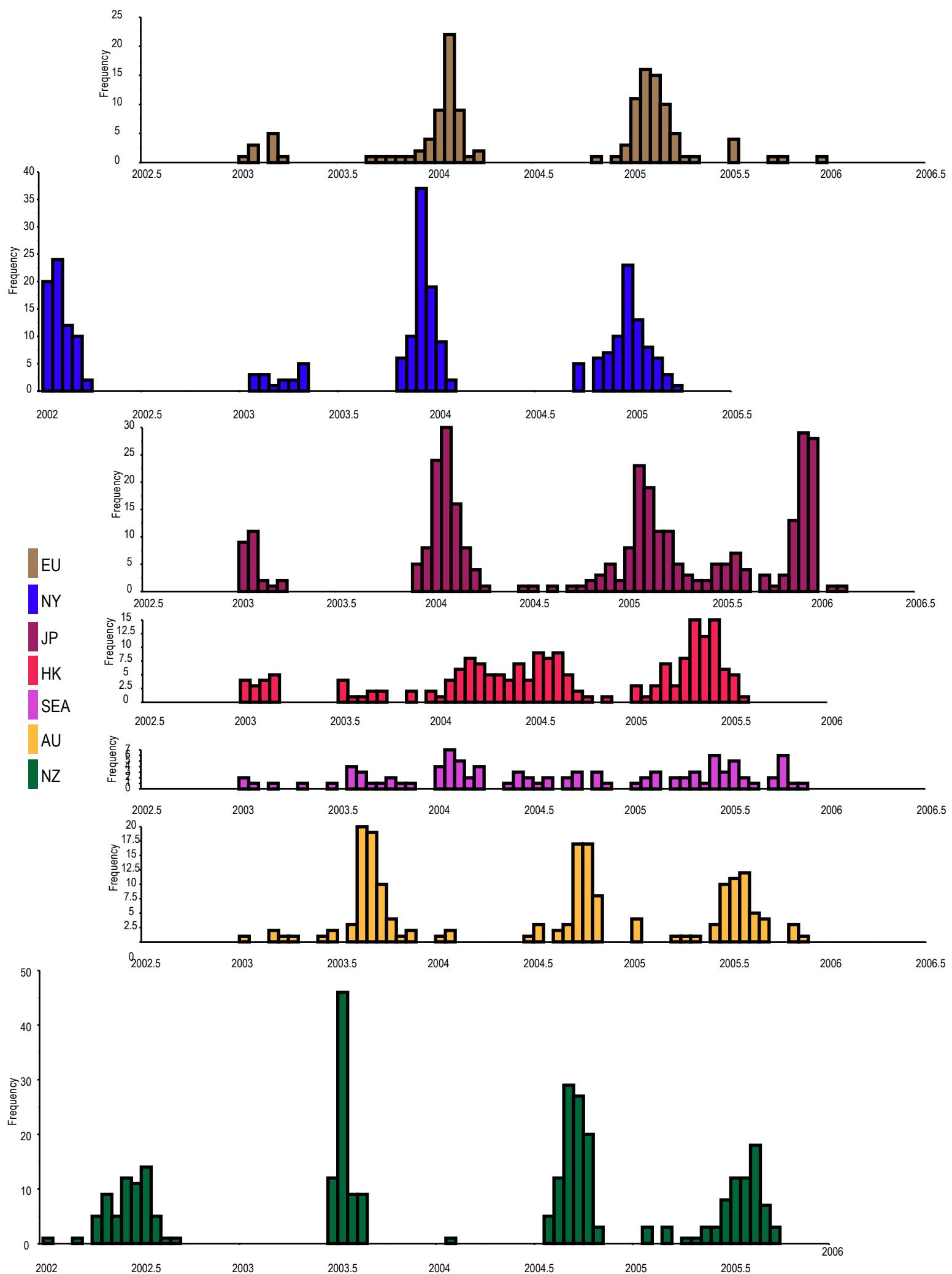


Fig. S2. Distribution of virus sampling dates for each randomly generated dataset

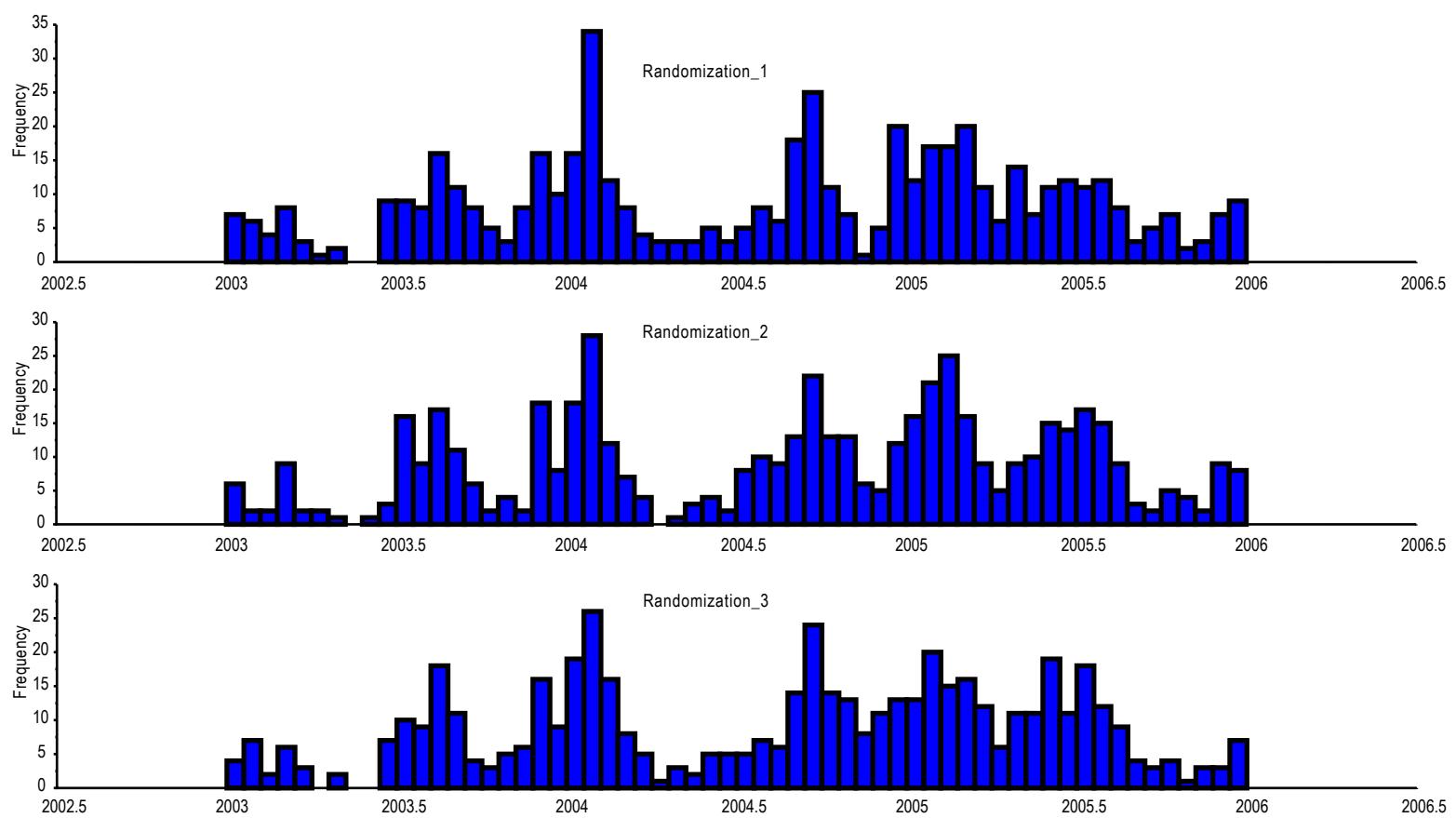
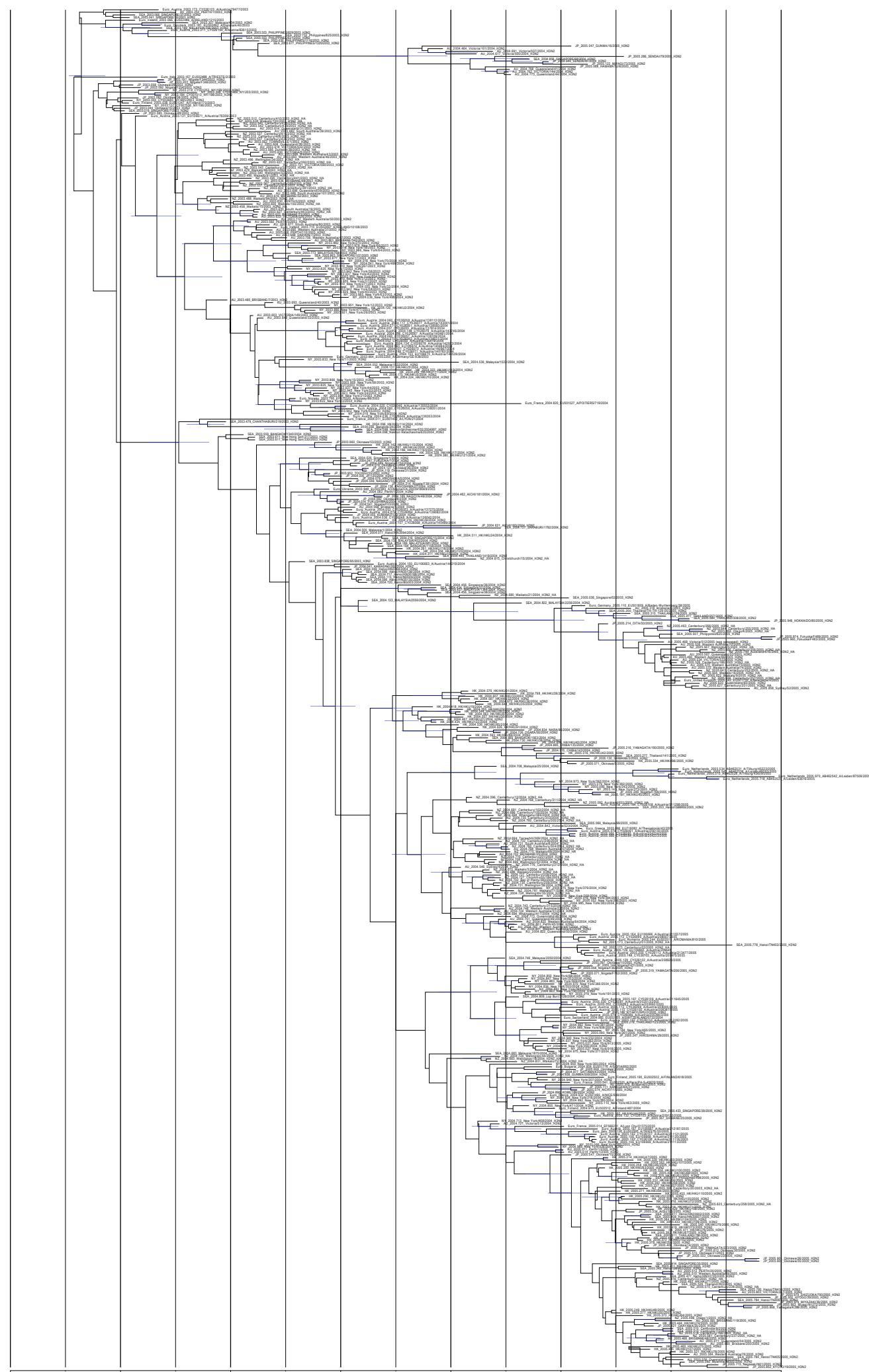
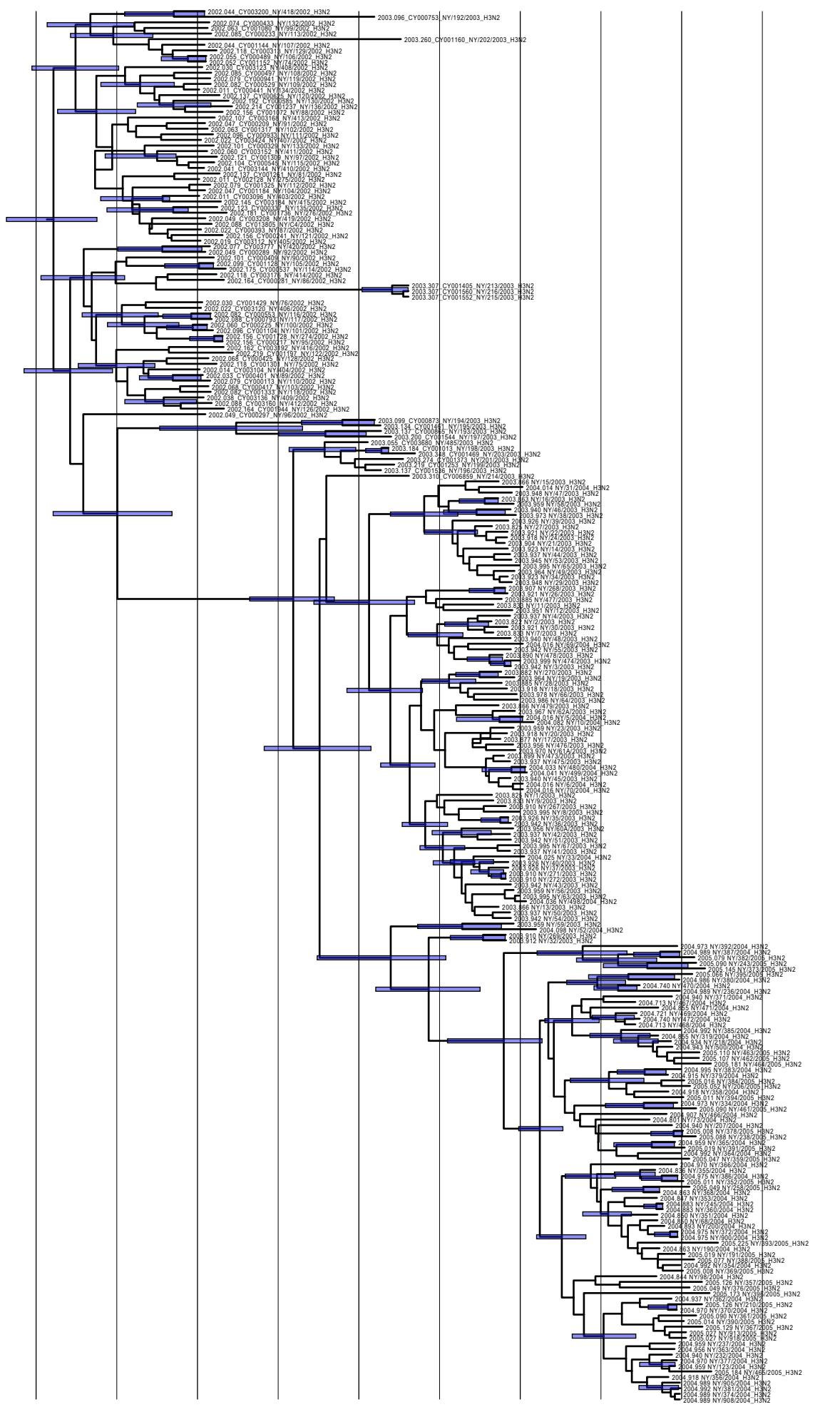


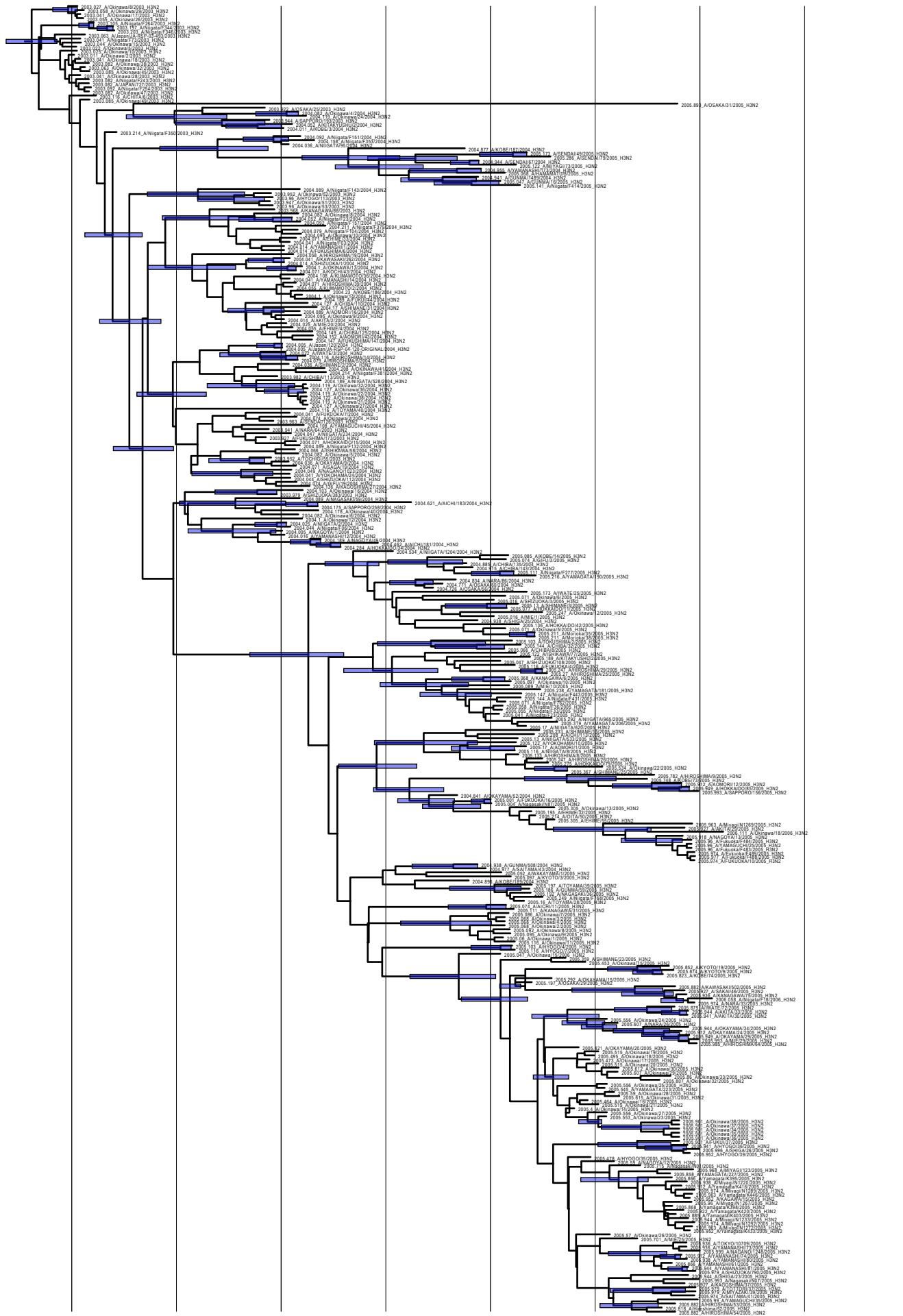
Fig. S3. Temporally structured maximum clade credibility phylogenetic tree showing the mixing of H3N2 influenza A virus global populations. This tree is representative of the three randomized datasets. Purple bars on nodes indicated 95% confidence intervals of date estimates.



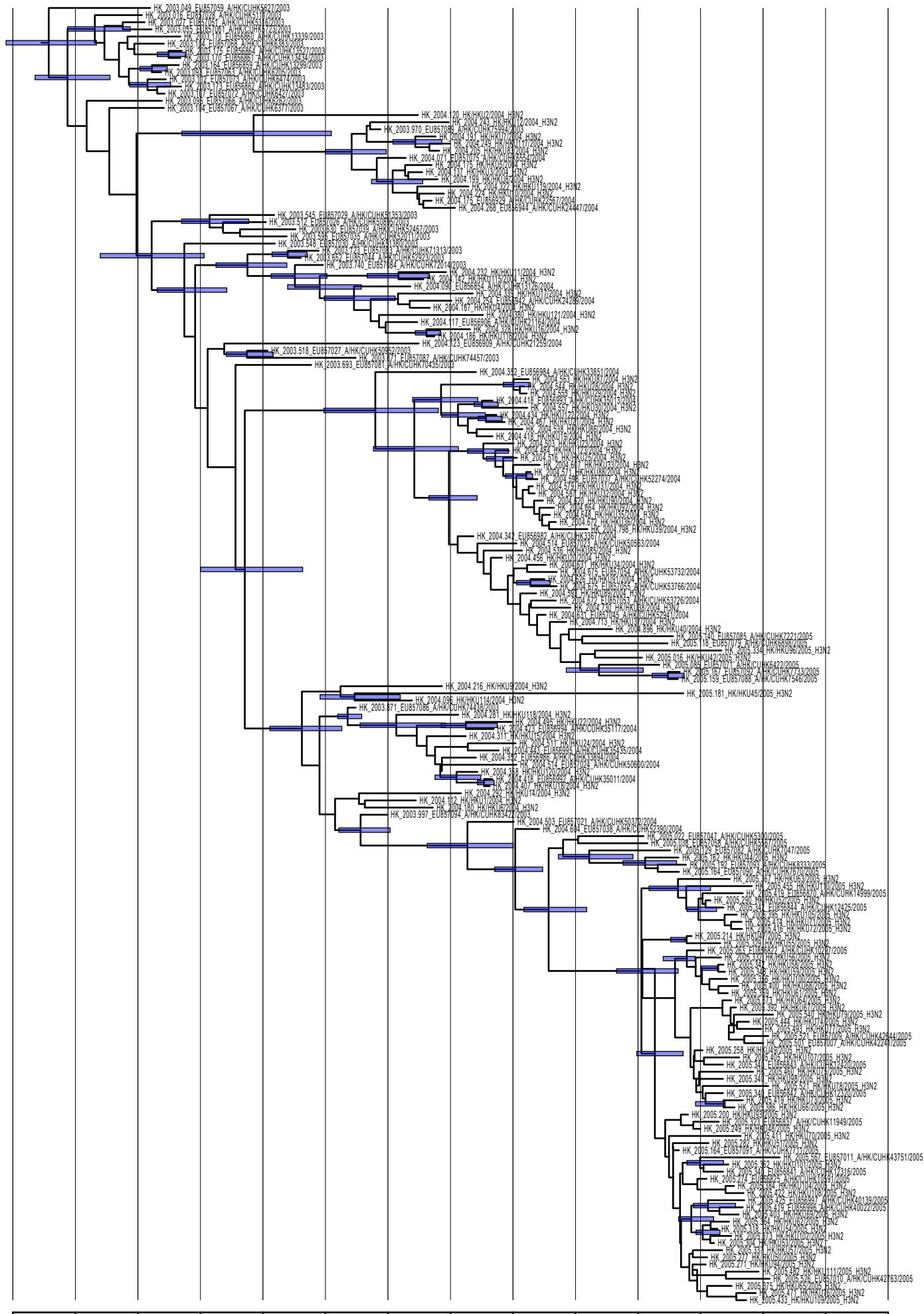
Figs. S4. Temporally structured maximum clade credibility phylogenetic tree showing the phylogenetic trees with taxon names used for the coalescent analysis of New York, Europe. Purple bars on nodes indicated 95% confidence intervals of date estimates.



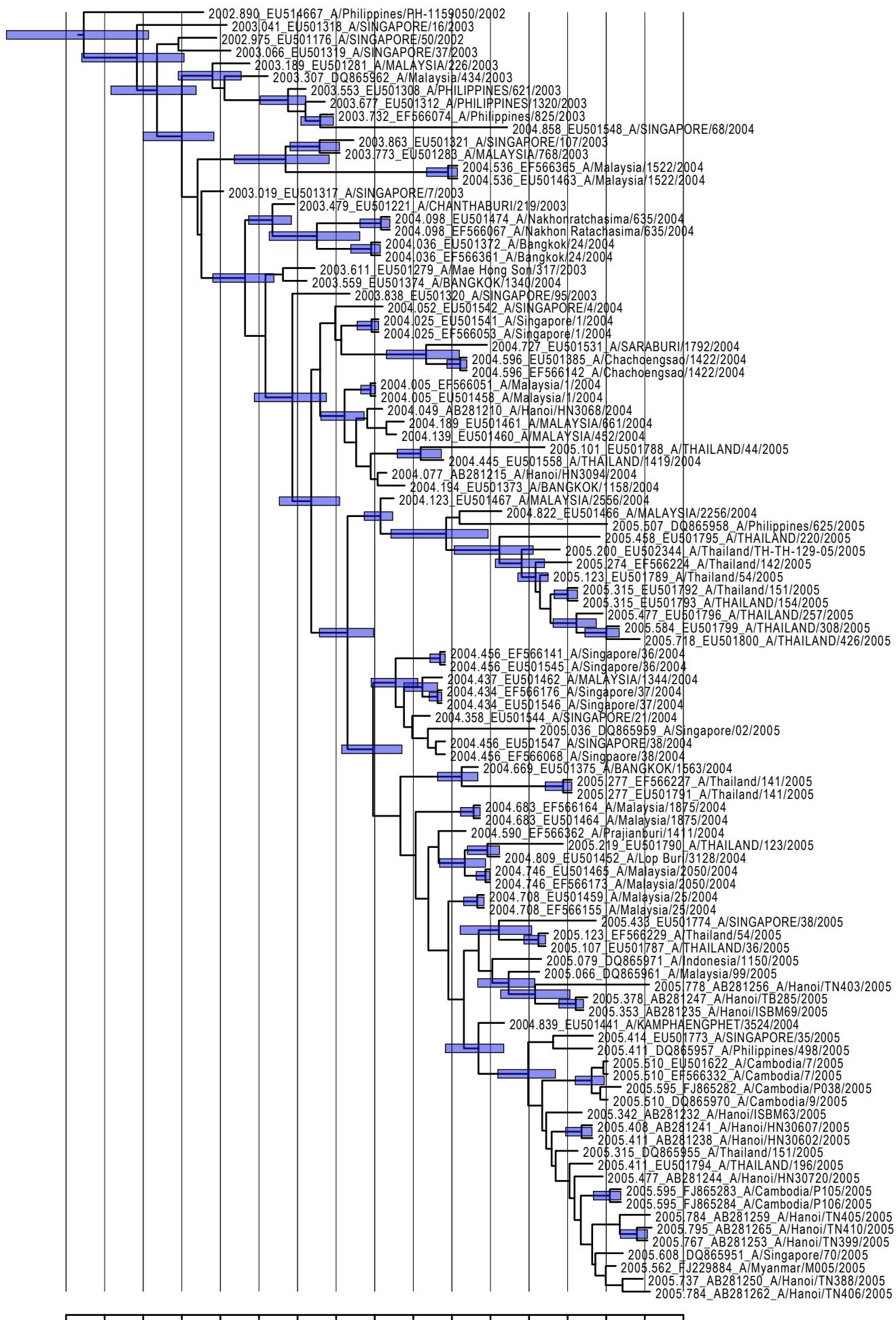
Figs. S5. Temporally structured maximum clade credibility phylogenetic tree showing the phylogenetic trees with taxon names used for the coalescent analysis of Japan. Purple bars on nodes indicated 95% confidence intervals of date estimates.



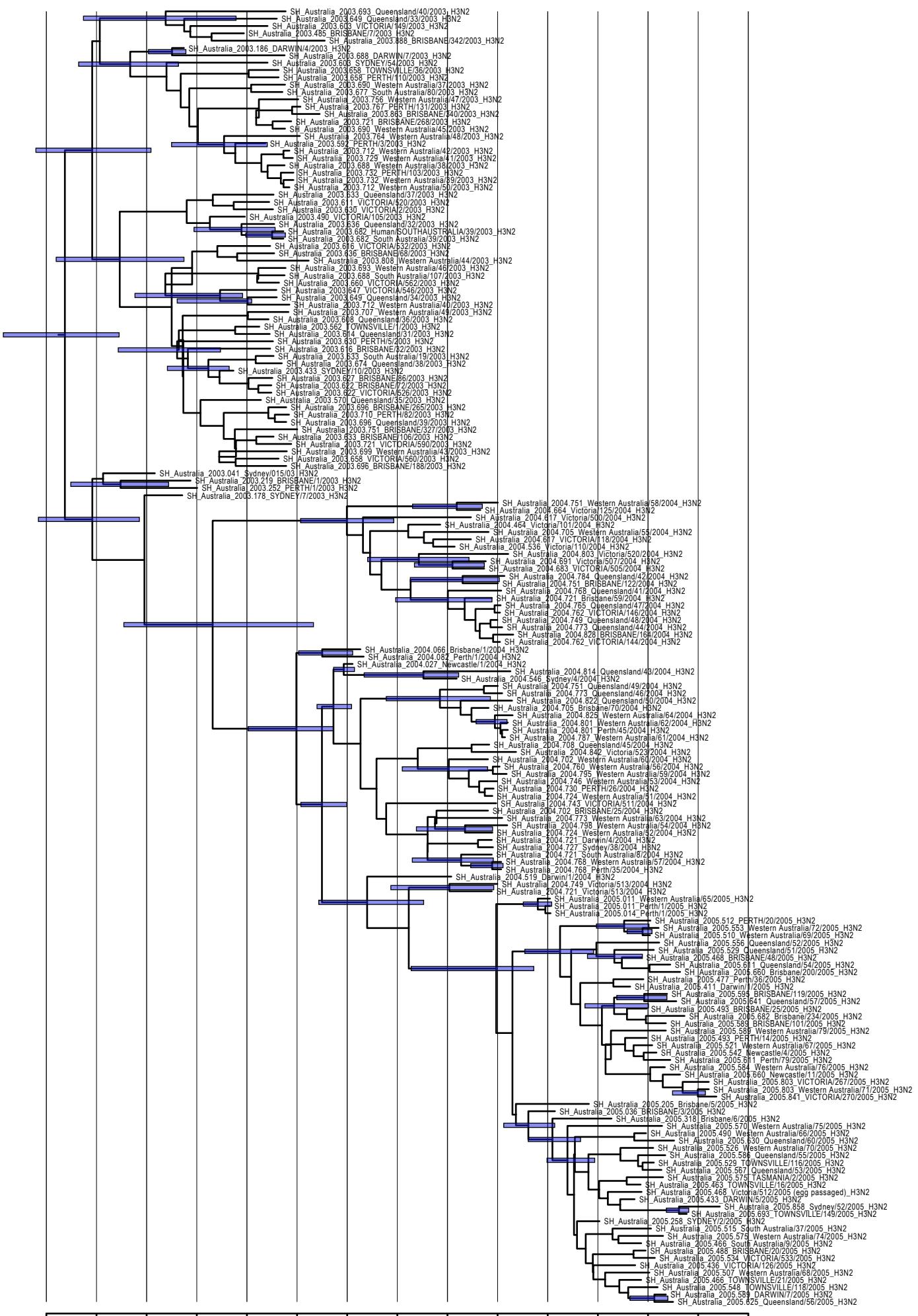
Figs. S6. Temporally structured maximum clade credibility phylogenetic tree showing the phylogenetic trees with taxon names used for the coalescent analysis of Hong Kong. Purple bars on nodes indicated 95% confidence intervals of date estimates.



Figs. S7. Temporally structured maximum clade credibility phylogenetic tree showing the phylogenetic trees with taxon names used for the coalescent analysis of South East Asia. Purple bars on nodes indicated 95% confidence intervals of date estimates.



Figs. S8. Temporally structured maximum clade credibility phylogenetic tree showing the phylogenetic trees with taxon names used for the coalescent analysis of Australia. Purple bars on nodes indicated 95% confidence intervals of date estimates.



Figs. S9. Temporally structured maximum clade credibility phylogenetic tree showing the phylogenetic trees with taxon names used for the coalescent analysis of New Zealand. Purple bars on nodes indicated 95% confidence intervals of date estimates.



Figs. S10. Temporally structured maximum clade credibility phylogenetic tree showing the phylogenetic trees with taxon names used for the coalescent analysis of Europe. Purple bars on nodes indicated 95% confidence intervals of date estimates.

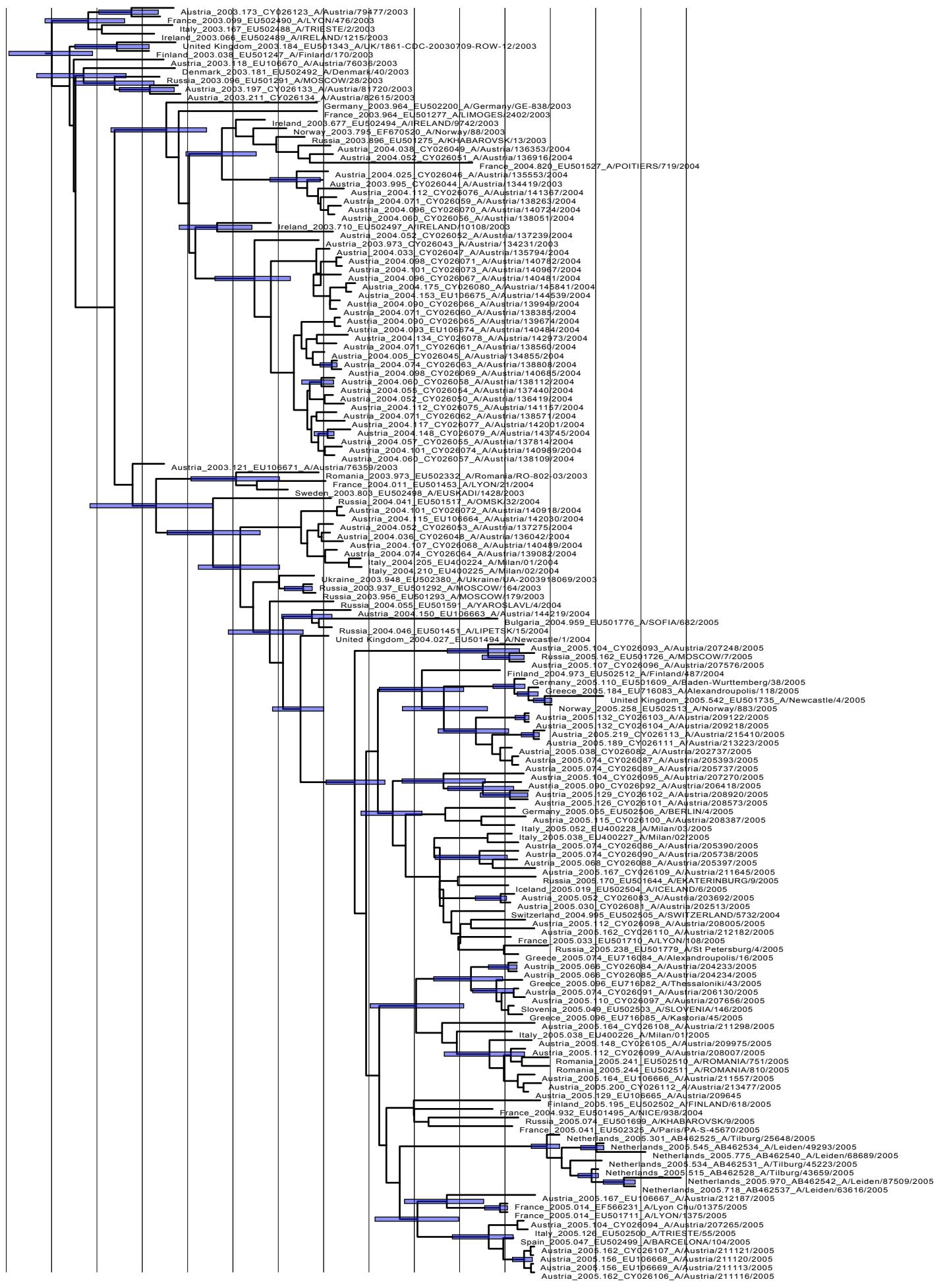
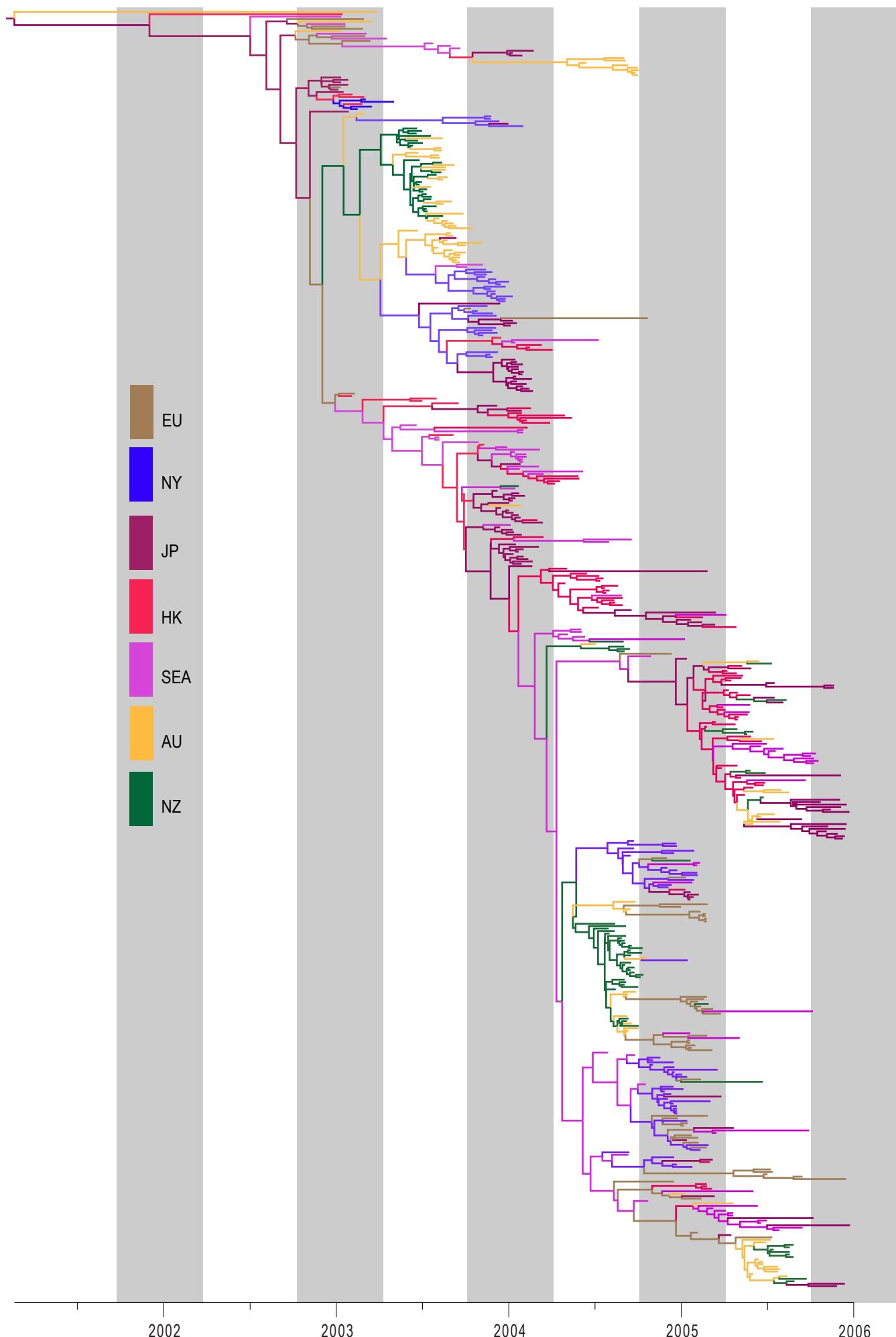
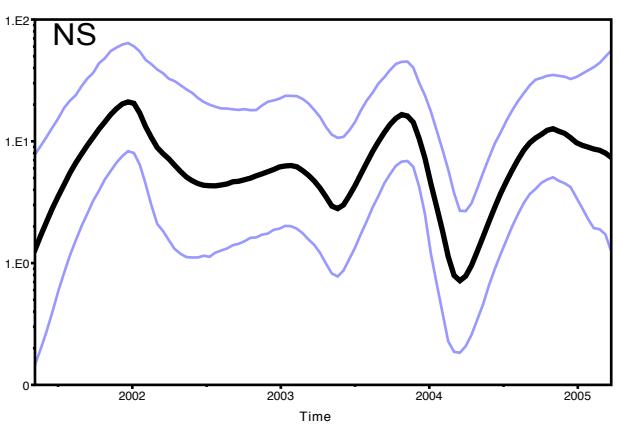
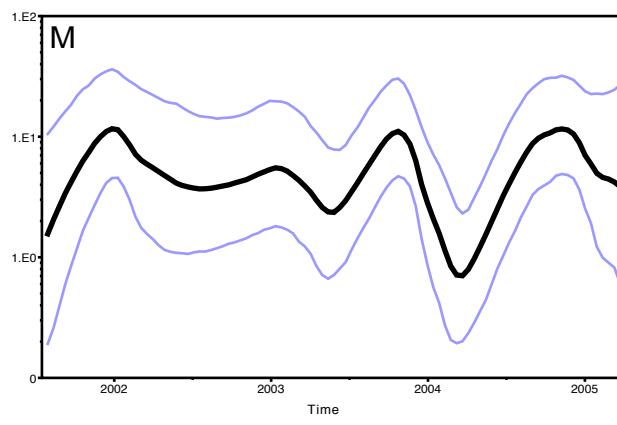
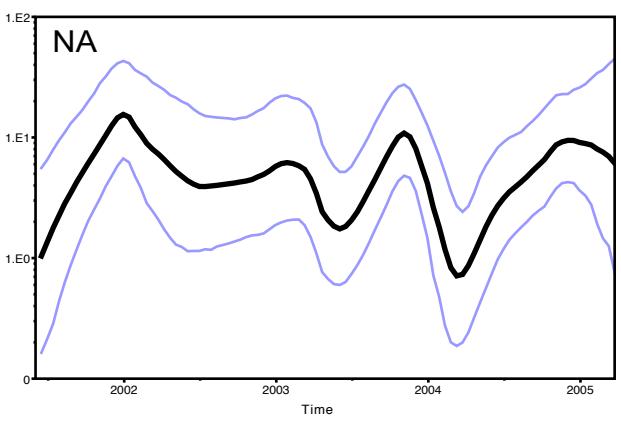
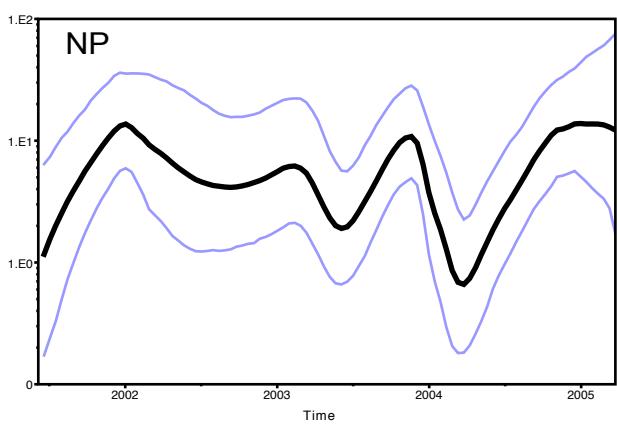
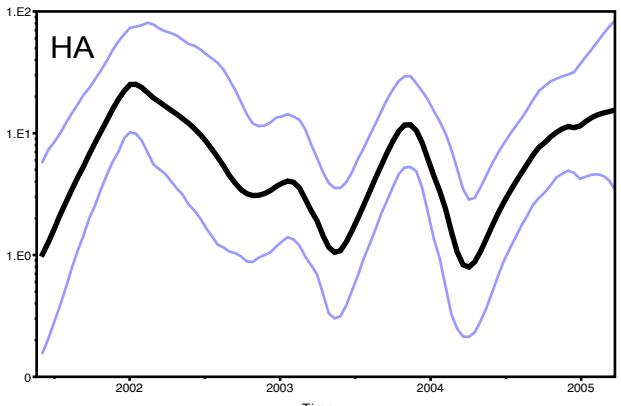
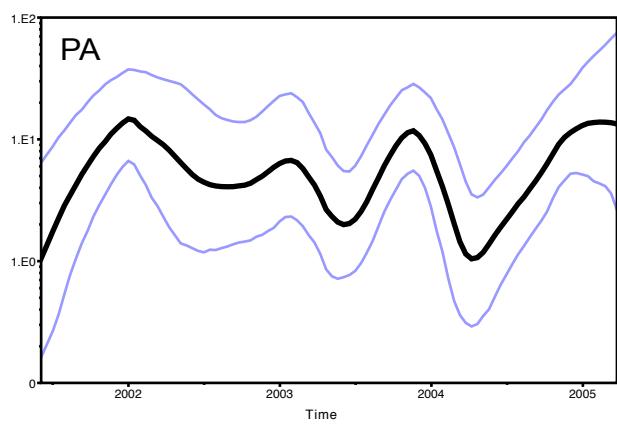
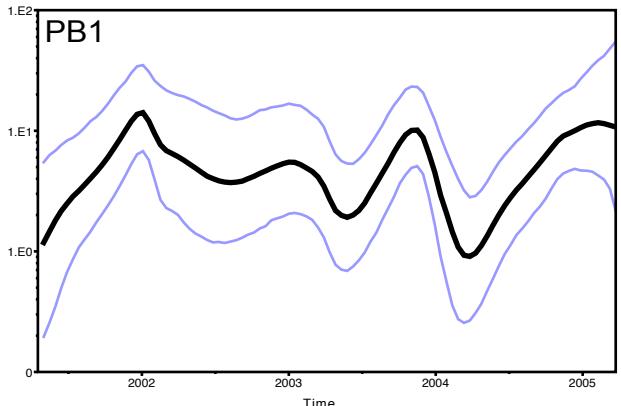
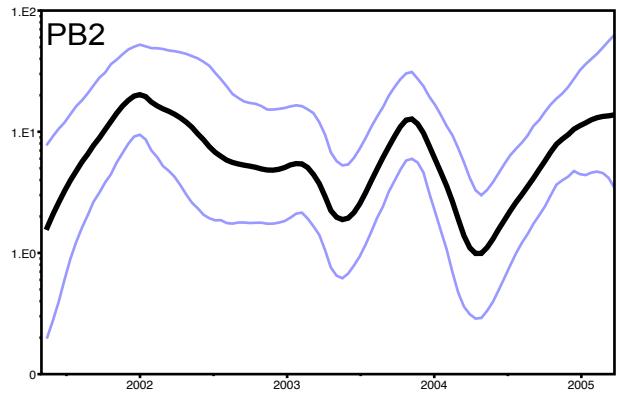


Fig S11. Phylogenetic tree generated from the discrete phylogeographic analysis showing ancestral state changes at tree nodes recovered from the sampled trees for isolates sampled in each discrete location and influenza season. Taxon names are indicated in Fig S1. The color key indicates the ancestral state location mapped onto the trees.



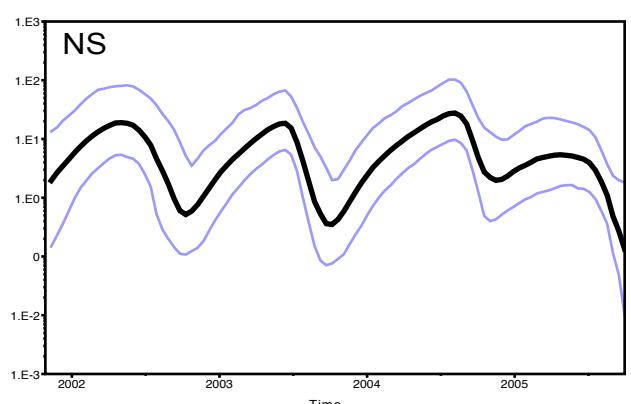
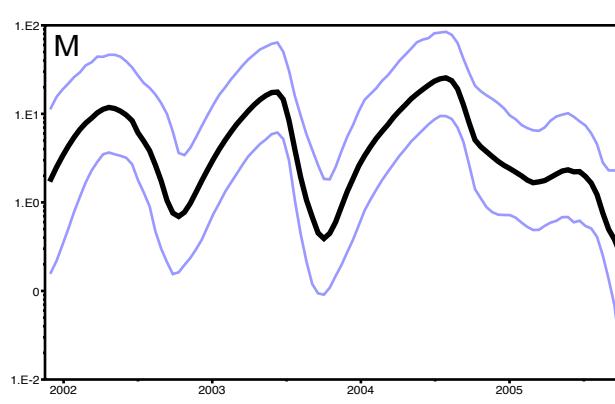
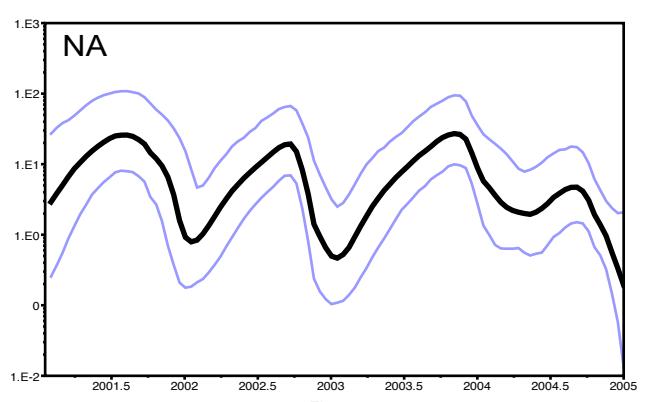
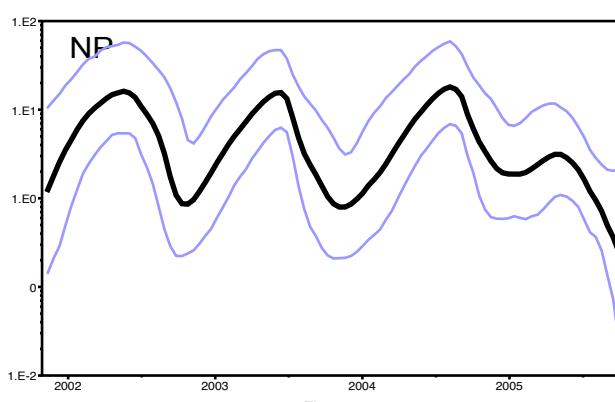
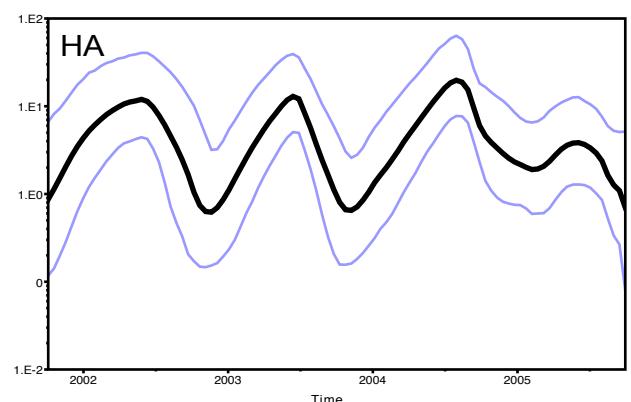
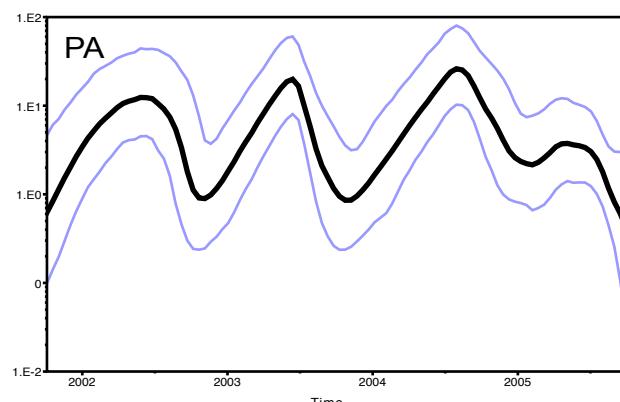
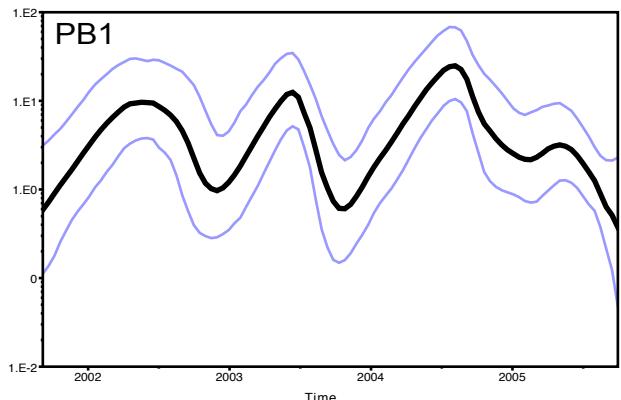
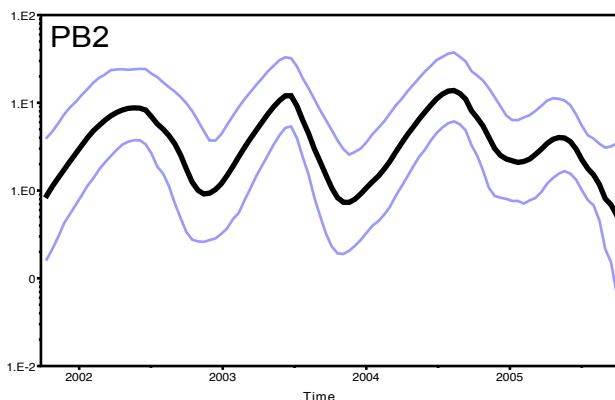
Figs. S12. Bayesian skyride analysis of influenza genomes depicting fluctuating levels of genetic diversity from isolates sampled in New York. The x-axis indicates time from the youngest sampled sequence to the lower 95% confidence interval of the tree root height, while the y-axis indicates relative genetic diversity (Net) as estimated from the skyride coalescent analysis.

New York - GMRF Skyride



Figs. S13. Bayesian skyride analysis of influenza genomes depicting fluctuating levels of genetic diversity from isolates sampled in New Zealand. The x-axis indicates time from the youngest sampled sequence to the lower 95% confidence interval of the tree root height, while the y-axis indicates relative genetic diversity (Net) as estimated from the skyride coalescent analysis.

New Zealand - GMRF Skyride



Figs. S14. Bayesian skyride analysis of influenza genomes depicting fluctuating levels of genetic diversity from isolates sampled in Hong Kong. The x-axis indicates time from the youngest sampled sequence to the lower 95% confidence interval of the tree root height, while the y-axis indicates relative genetic diversity (Net) as estimated from the skyride coalescent analysis.

Hong Kong - GMRF Skyride

