Integrating influenza antigenic dynamics with molecular evolution

Trevor Bedford[1], Marc A. Suchard[2,3,4], Philippe Lemey[5], Gytis Dudas[1], Victoria Gregory[6], Alan J. Hay[6], John W. McCauley[6], Colin A. Russell[7,8], Derek J. Smith[7,8,9] & Andrew Rambaut[1,10]

[1]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK, [2]Department of Biomathematics, David Geffen School of Medicine at UCLA, University of California, Los Angeles CA, USA, [3]Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles CA, USA, [4]Department of Biostatistics, UCLA Fielding School of Public Health, University of California, Los Angeles CA, USA, [5]Department of Microbiology and Immunology, Katholieke Universiteit Leuven, Leuven, Belgium, [6]Division of Virology, MRC National Institute for Medical Research, Mill Hill, London, UK, [7]Centre for Pathogen Evolution, Department of Zoology, University of Cambridge, Cambridge, UK, [8]WHO Collaborating Center for Modeling, Evolution, and Control of Emerging Infectious Diseases, University of Cambridge, Cambridge, UK, [9]Department of Virology, Erasmus Medical Centre, Rotterdam, Netherlands, [10]Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

Abstract

Influenza viruses undergo continual antigenic evolution allowing mutant viruses to evade host immunity acquired to previous virus strains. Antigenic phenotype is often assessed through pairwise measurement of cross-reactivity between influenza strains using the hemagglutination inhibition (HI) assay. Here, we extend previous approaches to antigenic cartography, and simultaneously characterize antigenic and genetic evolution by modeling the diffusion of antigenic phenotype over a shared virus phylogeny. Using HI data from influenza A subtypes H3N2 and H1N1 and influenza B lineages Victoria and Yamagata, we determine patterns of antigenic drift across lineages and show that antigenic drift drives incidence rates and mediates interference between influenza lineages. This work makes possible substantial future advances in investigating the dynamics of influenza and other antigenically-variable pathogens by providing a model that intimately combines molecular and antigenic evolution.

Introduction

Seasonal influenza infects between 10% and 20% of the human population every year, causing an estimated 250,000 to 500,000 deaths annually [1]. While individuals develop long-lasting immunity to particular influenza strains after infection, antigenic mutations to the influenza virus genome result in proteins that are recognized to a lesser degree by the human immune system, leaving individuals susceptible to future infection. The influenza virus population continually evolves in antigenic phenotype in a process known as antigenic drift. A large proportion of the disease burden of influenza stems from antigenic drift, which allows individuals to be infected multiple times throughout their lives. Although influenza vaccines may lack efficacy for a variety of reasons [2], antigenic drift causes efficacy against a fixed vaccine formulation to decline over time. A thorough understanding of the process of antigenic drift is essential to public health efforts to control mortality and morbidity through the use of a seasonal influenza vaccine.

Before 2009, there were four major clades or lineages of influenza circulating within the human population: the H3N2 and H1N1 subtypes of influenza A, and the Victoria and Yamagata lineages of influenza B. In the case of influenza A, subtypes A/H3N2 and A/H1N1 refer to the genes, hemagglutinin (H or HA) and neuraminidase (N or NA), that are primarily responsible for the antigenic character of a strain. In the case of influenza B, Victoria (B/Vic) and Yamagata (B/Yam) refer to antigenically distinct lineages which diverged from a single lineage prior to 1980 [3]. Mutations to the HA1 region of the hemagglutinin protein are thought to drive the majority of antigenic drift in the influenza virus [4,5]. Experimental characterization of antigenic phenotype is possible through the hemagglutination inhibition (HI) assay [6], which measures the cross-reactivity of one virus strain to serum raised against another strain through challenge or vaccination. Sera from older strains react poorly with more recent viruses resulting in new strains having a transmission advantage over previously established strains.

The results of many HI assays across a multitude of virus strains of a single subtype can be combined to yield a two-dimensional map, quantifying antigenic similarity and distance [7]. The antigenic map of influenza A/H3N2 has shown substantial evolution of the influenza virus population since its emergence in 1968. Evolution of antigenic phenotype appears punctuated with episodes of more rapid innovation interspersed by periods of stasis, while genetic evolution appears more continuous [7], suggesting that a relatively small number of genetic changes or combinations of genetic changes may drive changes in antigenic phenotype. The process of antigenic drift results in the rapid turnover of the virus population. Although mutation occurs rapidly, genetic diversity among contemporaneous viruses is low and phylogenetic analysis shows a characteristically 'spindly' tree with a single predominant trunk lineage and transitory side branches that persist for only 1–5 years [8].

Previously, the antigenic and genetic patterns of influenza evolution have been analyzed essentially in isolation. An antigenic map is constructed from a panel of HI measurements, and a phylogenetic tree is constructed from sequence data. However, the opportunity for a combined approach exists as both the antigenic map and the phylogenetic tree often contain many of the same isolates. Here, we implement a flexible Bayesian approach to

characterize jointly the antigenic and genetic evolution of the influenza virus population. We apply this approach to investigate the dynamics of A/H3N2, A/H1N1, B/Vic and B/Yam viruses, and, for the first time, present detailed reconstructions of the antigenic dynamics of all four circulating influenza clades.

Results and discussion

Antigenic and evolutionary cartography

In order to assess patterns of antigenic evolution among influenza strains, we implemented a Bayesian probabilistic analog of multidimensional scaling (MDS), referred to here as BMDS (see Materials and methods). In this model, viruses and sera are given N-dimensional locations, thus specifying an 'antigenic map', such that distances between viruses and sera in this space are inversely proportional to cross-reactivity. In the BMDS model, a map distance of one antigenic unit translates to an expectation of a 2-fold drop in HI titer between virus and sera. Maps that produce pairwise distances most congruent with the observed titers will have a high likelihood and will be favored by the BMDS model. We integrate over sources of uncertainty, such as antigenic locations, in a flexible Bayesian fashion. We apply this model to HI measurements of virus isolates against post-infection ferret antisera for influenza A/H3N2, A/H1N1, B/Vic and B/Yam (see Materials and methods).

We begin with Bayesian analogs of the models used by Smith et al. [7], in which viruses and sera are represented as N-dimensional locations as described in the 'Antigenic cartography' section of Materials and methods. In this case, 'serum effects' are fixed to the maximum titers exhibited by particular ferret sera and give the baseline expectation for titer when virus and serum are antigenically identical. Virus and serum locations follow an uninformative diffuse normal prior. After comparing models of differing dimensions, Smith et al. [7] arrive at a 2D model as the preferred model for their data. Smith et al. [7] implement a form of MDS, seeking to optimize virus and serum locations such that the sum of squared errors between expected and observed titers is minimized (eq. 4 of Materials and methods). Here, in implementing BMDS, we provide a likelihood function for the probability of observing HI data given virus and serum locations (eq. 9 of Materials and methods) and seek to estimate model parameters through Bayesian inference using Markov chain Monte Carlo (MCMC). However, the basic antigenic model describing drop in HI titer as proportional to Euclidean distance between virus and serum locations is identical between these methods.

We test model performance by constructing training datasets representing 90% of the HI measurements for each of the four influenza lineages and test datasets representing the remaining 10% of the measurements for each lineage. By fitting the BMDS model to the training dataset, we are able to predict HI titers in the test dataset and compare these predicted titers to observed titers. We find that a two dimensional model has better predictive power than models of lower of higher dimension in all four influenza lineages (models 1–5; Table 1). We find that this 2D model performs well, yielding an average

absolute predictive error of between 0.78 and 0.91 $\log_2$ HI titers across influenza lineages (model 2; Table 1), in line with the results of Smith et al. [7]. Consequently, we specify a two dimensional model in all subsequent analyses. The finding of a low-dimensional map across influenza lineages extends previous studies in A/H3N2 [7] and remains an interesting and fundamental empirical observation.

Table 1. Average absolute prediction error of $\log_2$ HI titer for test data across models and datasets.

Model	Data	Dimen	Location prior	Serum effects	Virus effects	Test error A/H3N2	A/H1N1	B/Vic	B/Yam
1	HI	1D	Uninformed	Fixed	None	1.35	0.94	0.90	1.08
2	HI	2D	Uninformed	Fixed	None	0.91	0.78	0.82	0.90
3	HI	3D	Uninformed	Fixed	None	0.93	0.80	0.85	0.92
4	HI	4D	Uninformed	Fixed	None	0.98	0.84	0.90	0.97
5	HI	5D	Uninformed	Fixed	None	1.04	0.89	0.98	1.04
6	HI/year	2D	Drift	Fixed	None	0.91	0.75	0.77	0.83
7	HI/year/seq	2D	Diffusion/Drift	Fixed	None	0.89	0.74	0.74	0.83
8	HI/year/seq	2D	Diffusion/Drift	Estimated	None	0.77	0.73	0.66	0.75
9	HI/year/seq	2D	Diffusion/Drift	Fixed	Estimated	0.80	0.72	0.69	0.75
10	HI/year/seq	2D	Diffusion/Drift	Estimated	Estimated	0.76	0.71	0.64	0.72

Previous work on influenza antigenic and genetic evolution has shown that antigenic distance accumulates with increasing genetic distance [7, 9, 10]. Here, we examine pairwise relationships between viruses and observe a correlation between amino acid mutations and antigenic distance (Figure 1) and a similar correlation between phylogenetic distance and antigenic distance (Figure 1). Thus, genetic relationships between viruses provide some predictive power to estimate antigenic distances in the absence of HI data. However, the magnitudes of the coefficients of determination R^2 are low (Figure 1), suggesting that genetic relationships alone will not completely resolve antigenic distances.

Consequently, we seek to flexibly incorporate genetic data by simultaneously modeling antigenic and genetic evolution through an evolutionary diffusion process [11], wherein a virus's antigenic character state evolves along branches of the phylogenetic tree according to a Brownian motion process (see Materials and methods). The phylogenetic diffusion process acts as a prior on virus locations, so that genetically similar viruses are expected to share similar antigenic locations. The antigenic diffusion process includes both systematic drift with time and covariance induced by phylogenetic proximity. We examine the effects of including only systematic drift (model 6; Table 1) and systematic drift plus phylogenetic diffusion (model 7; Table 1), finding a small increase in predictive accuracy of between 0.02 and 0.08 $\log_2$ HI titers when both processes are included. The systematic drift process informs virus and serum locations by dates of isolation and the phylogenetic diffusion process informs virus locations by genetic sequences. Thus, in these models, antigenic locations are inferred using both genetic data and HI data and will differ from locations inferred from HI data alone. If HI data is rich, then we expect minor differences in antigenic locations with the inclusion of genetic data (as may be the case for A/H3N2), while if HI data is spare, then we expect genetic data to play a larger role in determining antigenic locations (as may be the case for B/Vic and B/Yam).
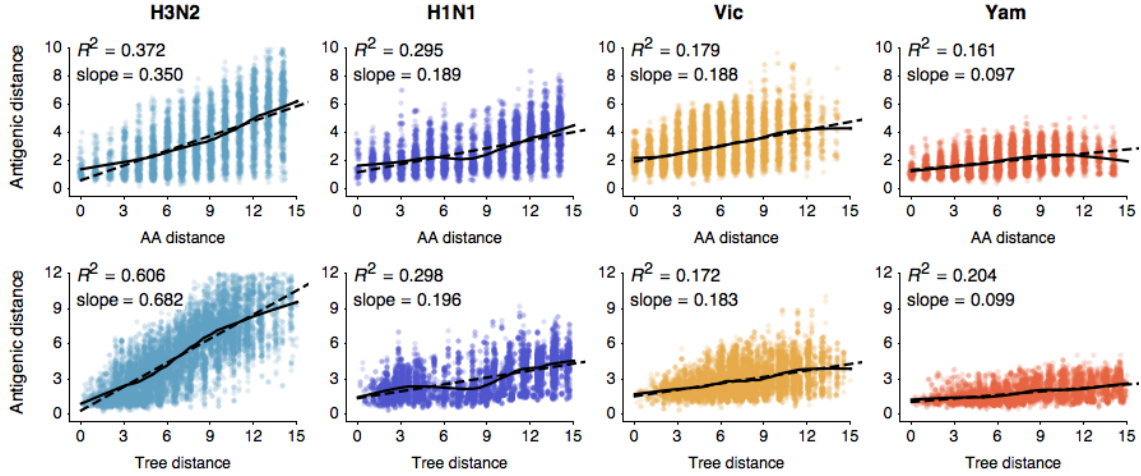
Figure 1. Pairwise correlations between genetic distance, measured as amino acid mutations or as phylogenetic distance, and antigenic distance for influenza A/H3N2, A/H1N1, B/Vic and B/Yam. The top row shows correlations between number of amino acid mutations in HA1 and average antigenic distance between 10,000 random pairs of viruses. The bottom row shows correlations between average phylogenetic distance, measured in terms of years, and average antigenic distance between 10,000 random pairs of viruses. Dashed lines show linear model fits, with R^2 and slope noted, while solid lines show LOESS fits. Antigenic distances derive from model 2 of Table 1.

We further extend the model by estimating 'serum effects' to quantify the overall strength reactivity of each serum, rather than fixing serum effects at maximum titers. Serum effects likely derive from experimental variation in serum extraction and processing or from variation in immune responses between ferrets. Serum effects determine the baseline expectation of titer when virus and serum have identical antigenic locations. However, if serum effect is fixed to the serum's maximum titer, this will often not be the case, as the virus giving the maximum titer is likely to be antigenically distinct from the serum. Thus, fixing serum effects will tend to under-estimate effect size; we observe a mean effect of 10.42 $\log_2$ HI titers for A/H3N2 when fixing serum effects and a mean of 10.94 when estimating serum effects. We find that estimating serum effects improves test error further (model 8, Table 1), with improvements of between 0.01 and 0.12 $\log_2$ HI titers.

Additionally, we include and estimate 'virus effects' in an analogous fashion, which are intended to represent differences in overall reactivity between viruses. With virus effects included, baseline titer derives from both the virus and the serum used in the HI reaction. We find that including virus effects further improves test error, either alone with fixed serum effects (model 9, Table 1) or combined with estimating serum effects (model 10, Table 1). With fixed serum effects, the inclusion of virus effects results in improvements of between 0.02 and 0.09 $\log_2$ HI titers and with estimated serum effects, the inclusion of virus effects results in improvements of between 0.01 and 0.05 $\log_2$ HI titers.

We find that the average absolute error in predicted $\log_2$ HI titer is nearly constant with antigenic distance (Pearson correlation, $r = 0.098$), thus supporting our model assumption that the drop in $\log_2$ titer is proportional to the Euclidean distance separating viruses and

sera on the antigenic map. Additionally, we find that the absolute error in predicted titer is nearly constant with time (Pearson correlation, $r = -0.007$). Antigenic locations inferred by the model are well resolved; estimates of antigenic distance between pairs of viruses show relatively little variation across the posterior. We estimate that virus distances have, on average, a 50% credible interval of ± 0.45 antigenic units for A/H3N2, ± 0.57 units for A/H1N1, ± 0.76 units for B/Vic, and ± 0.65 units for B/Yam.

We find strong correspondence between our results and previous results by Smith et al. [7], with equivalent models producing globally consistent antigenic maps and other models producing locally consistent maps with a small degree of global inconsistency (Figures 9–12). When implementing the same underlying model, differences in the MDS and BMDS approaches reflect greater philosophical differences between maximum-likelihood and Bayesian statistical approaches, with the former seeking the single most likely explanation for the data, and the latter seeking to fully characterize model uncertainty. Additionally, the BMDS method improves flexibility, allowing extensions to the basic cartographic model, such as the incorporation of serum effects and evolutionary priors, that improve fit and add biological interpretability.

Antigenic evolution across influenza lineages

Through our analysis, we reveal the antigenic, as well as evolutionary, relationships among viruses in influenza A/H3N2, A/H1N1, B/Vic and B/Yam, quantifying both antigenic and evolutionary distances between strains (Figure 2). Over the time period of 1968 to 2011, influenza A/H3N2 shows substantially more antigenic evolution than is exhibited by A/H1N1 over the course of 1977 to 2009 or B/Vic and B/Yam over the course of 1986 to 2011. We observe prominent antigenic clusters in A/H3N2 and A/H1N1, but less prominent, though still apparent, clustering in B/Vic and B/Yam. Antigenic clusters show high genetic similarity, so that we observe very few mutation events leading to each cluster, rather than the repeated emergence of clusters. This analysis makes the fate of antigenic clusters obvious, with two clusters in A/H3N2 (Victoria/75 and Beijing/89) appearing to be evolutionary dead-ends. Labeling of prominent antigenic clusters in figures 2 and 3 is intended as a rough guide for orientation and not as exhaustive catalog of antigenic variation.

HI assays lack sensitivity beyond a certain point, so that for A/H3N2, cross-reactive measurements only exist between strains sampled at most 14 years apart, leaving only threshold titers, e.g. '<40', in more temporally distant comparisons. Because of the threshold of sensitivity of the HI assay, it's difficult to distinguish a linear trajectory in 2D antigenic space from a slightly curved trajectory. To solve this problem of identifiability, we assumed a weak prior that favors linear movement in the 2D antigenic space (present in models 6 through 9; Table 1), with the slope of the linear relationship and the precision of the relationship incorporated into the Bayesian model (see Materials and methods). Because of this, we interpret map locations locally rather than globally, and assess rates of antigenic movement without making strong statements about the larger configuration under which the movement occurs.

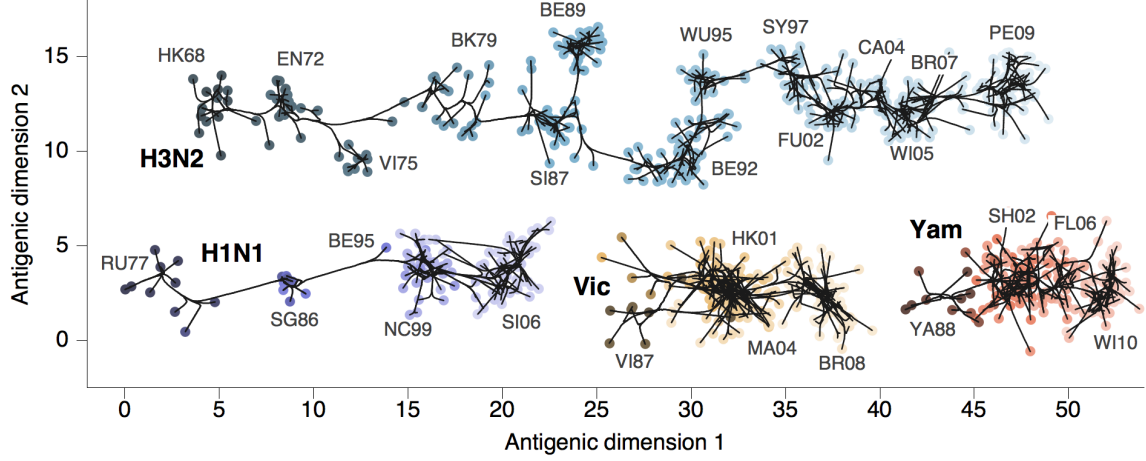We find that influenza A/H3N2 evolved along antigenic dimension 1 at an estimated rate

Figure 2. Antigenic locations of A/H3N2, A/H1N1, B/Vic and B/Yam viruses showing evolutionary relationships between virus samples. Circles represent a posterior sample of virus locations and have been shaded based on year of isolation. Antigenic units represent two-fold dilutions of the HI assay. Absolute positioning of lineages, e.g. A/H3N2 and A/H1N1, is arbitrary. Lines represent mean posterior diffusion paths when virus locations are fixed. Prominent antigenic clusters are labeled after vaccine strains present within clusters, and are abbreviated from Hong Kong/68, England/72, Victoria/75, Bangkok/79, Sichuan/87, Beijing/89, Beijing/92, Wuhan/95, Sydney/97, Fujian/02, California/04, Wisconsin/05, Brisbane/07, Perth/09 (A/H3N2), USSR/77, Singapore/86, Beijing/95, New Caledonia/99, Solomon Islands/06 (H1N1), Victoria/87, Hong Kong/01, Malaysia/04, Brisbane/08 (Vic), Yamagata/88, Shanghai/02, Florida/06, Wisconsin/10 (Yam).

of 1.01 antigenic units per year (Figure 3, Table **??**). However, we observe occasional large jumps in antigenic phenotype (Figure 3), corresponding to cluster transitions identified by Smith et al. [7]. Most variation is contained within the first antigenic dimension, but dimension 2 occasionally shows variation when two antigenically distinct lineages emerge and transiently coexist (Figure 2), as is the case with the previously identified Beijing/89 and Beijing/92 clusters.

We find that other clades of influenza evolved in antigenic phenotype substantially slower than A/H3N2 (Figure 3, Table 2). Influenza A/H1N1 evolved at a rate of 0.62 units per year, but showed a similar pattern of punctuated antigenic evolution with occasional larger jumps in phenotype, such as the emergence of the Solomon Islands/06 cluster. Influenza B/Victoria and B/Yamagata evolved slower still, with mean estimated rates 0.42 units per year and 0.32 units per year, respectively. Punctuated evolution is less obvious in B/Yam and B/Vic compared to A/H3N2 and A/H1N1, but antigenic clusters are still apparent, with recent transitions to the Brisbane/08 cluster in B/Vic [12] and to the Wisconsin/10 cluster in B/Yam [13]. Interestingly, a minor lineage of B/Vic, denoted B/Hubei-Songzi/51/2008 [12], has persisted through 2011, while remaining antigenically distinct from B/Brisbane/60/2008 viruses (Figure 3). Interestingly, although we observe significantly different drift rates between lineages, we observe less variation in diffusion volatility (Table 2). This is reflected in Figure 3, where all four lineages exhibit similar levels of standing antigenic variation, despite A/H3N2 drifting more quickly in antigenic phenotype.
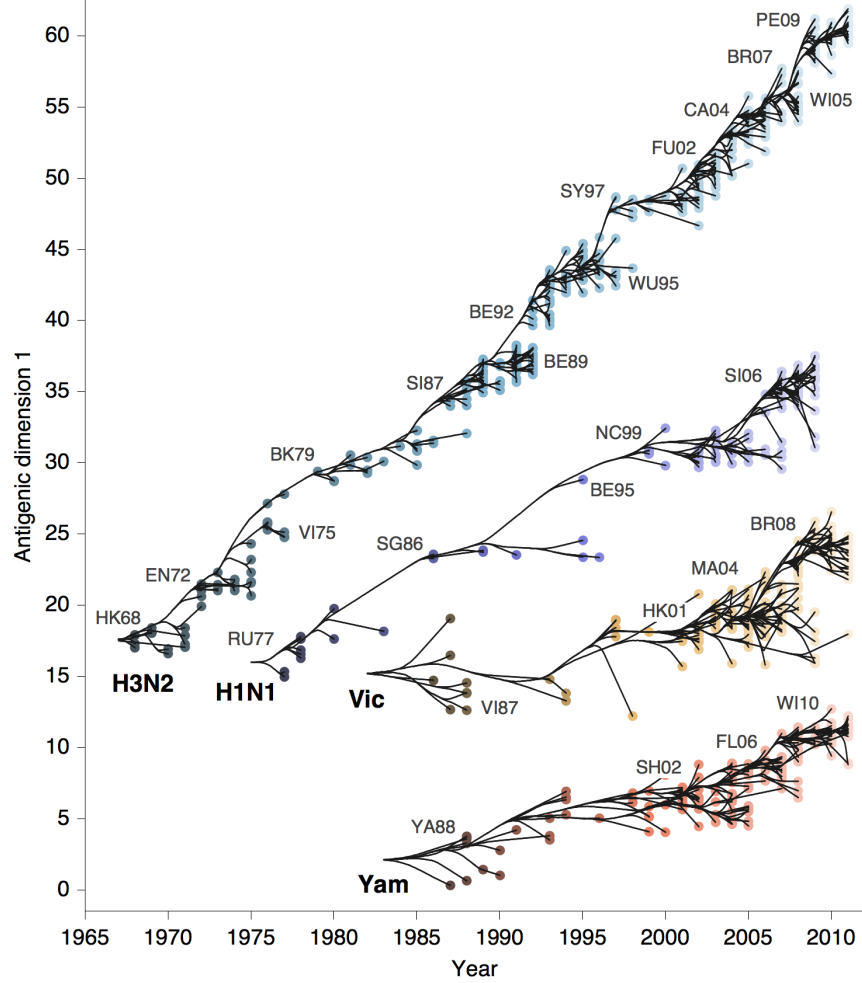
7

Figure 3. Antigenic drift of A/H3N2, A/H1N1, B/Vic and B/Yam viruses showing evolutionary relationships between virus samples. Antigenic drift is shown in terms of change of location in the first antigenic dimension through time. Circles represent a posterior sample of virus locations and have been shaded based on year of isolation. Antigenic units represent two-fold dilutions of the HI assay. Relative positioning of lineages, e.g. A/H3N2 and A/H1N1, in the vertical axis is arbitrary. Lines represent mean posterior diffusion paths when virus locations are fixed. Prominent antigenic clusters are labeled after vaccine strains present within clusters, and are abbreviated from Hong Kong/68, England/72, Victoria/75, Bangkok/79, Sichuan/87, Beijing/89, Beijing/92, Wuhan/95, Sydney/97, Fujian/02, California/04, Wisconsin/05, Brisbane/07, Perth/09 (A/H3N2), USSR/77, Singapore/86, Beijing/95, New Caledonia/99, Solomon Islands/06 (H1N1), Victoria/87, Hong Kong/01, Malaysia/04, Brisbane/08 (Vic), Yamagata/88, Shanghai/02, Florida/06, Wisconsin/10 (Yam).

These patterns of antigenic drift are reflected in the virus phylogenies (Figure 4). Influenza A/H3N2 has characteristically spindly tree showing rapid turnover of the virus population, while A/H1N1 and B have trees that show greater degrees of viral coexistance (Figure 4). The scaled effective population size $N_e\tau$ measures the timescale of coalescence for a phylogeny and quantifies the visual distinction between a 'spindly' tree and a 'bushy' tree [14]. Here, we see that $N_e\tau$ broadly correlates with the rate of antigenic drift (Table 2),

Table 2. Estimates of drift rate μ (in units per year), diffusion volatility σ_x^2 (in units^2 per year) and scaled effective population size $N_e\tau$ (in years) for influenza A/H3N2, A/H1N1, B/Vic and B/Yam including posterior means and 95% highest posterior density intervals.

Lineage	Drift μ	Volatility σ_x^2	Effective pop size $N_e\tau$
A/H3N2	1.01 (0.98–1.04)	1.25 (0.98–2.35)	5.03 (4.42–5.73)
A/H1N1	0.62 (0.56–0.67)	0.92 (0.65–1.56)	6.38 (4.99–8.12)
B/Vic	0.42 (0.32–0.51)	1.22 (0.85–2.25)	10.40 (8.42–12.80)
B/Yam	0.32 (0.25–0.39)	0.71 (0.46–1.36)	9.48 (7.76–11.50)

with A/H3N2 showing fast drift and reduced effective population size as expected from basic epidemiological models [15]. Antigenic drift results in the replacement of antigenically primitive lineages by antigenically advanced lineages, thereby reducing genealogical diversity.
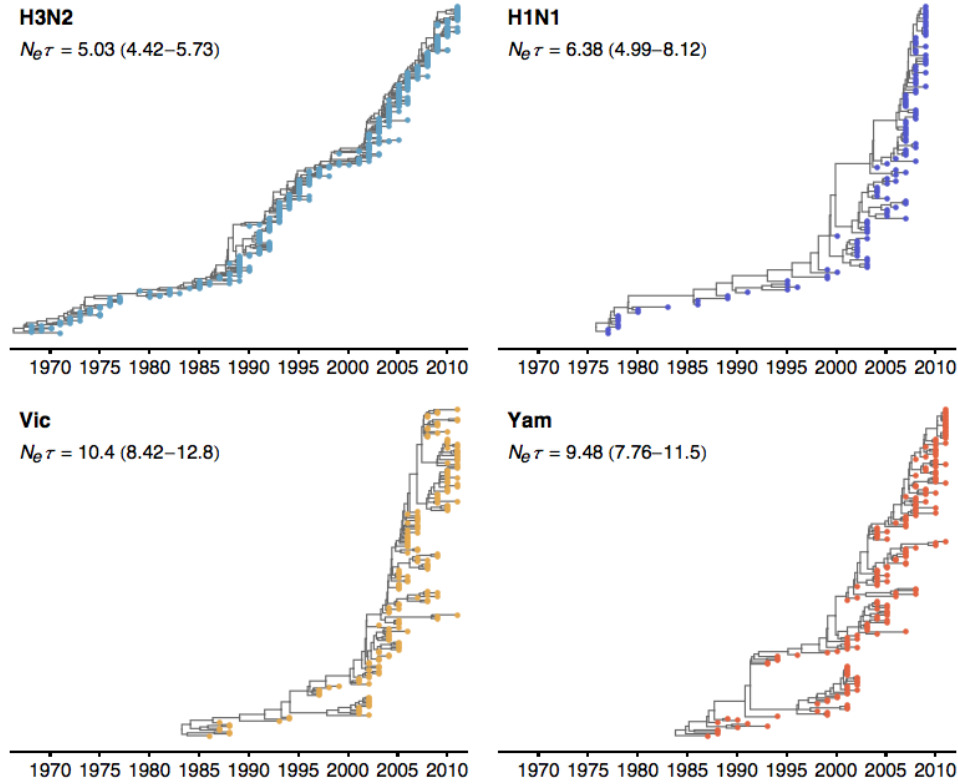

Figure 4. Time-resolved phylogenetic trees of A/H3N2, A/H1N1, B/Vic and B/Yam viruses. The maximum-clade credibility (MCC) tree is shown for each virus. These trees show genealogical relationships, so that branches are measured in terms of years rather than substitutions. The mean posterior scaled effective population size $N_e\tau$, measured in years, and its 95% credible interval is shown for each virus.

Punctuated evolution and its epidemiological consequences

We sought to summarize year-to-year patterns of antigenic drift by calculating the difference in mean location between consecutive years (Figure 5). We estimate year-to-year antigenic drift for years 1992 to 2011 by calculating the average location along dimension 1 of phylogenetic lineages present in the tree at year i and comparing this location to the average location of phylogenetic lineages present in the tree at year $i-1$. There may often be large discontinuities in virus locations across the population; our use of difference in mean location is meant to capture both the distance between antigenic clusters and also the change in cluster frequency over consecutive years. We observe greater heterogeneity in year-to-year antigenic drift in type A than in type B lineages (Figure 5), with standard deviation of year-to-year antigenic drift equal to 0.97 units in A/H3N2, 0.66 units in A/H1N1, 0.46 units in B/Vic and 0.26 units in B/Yam. This analysis classifies drift only to the level of consecutive years; some coarse-graining of the timings of transition events will necessarily occur.

We investigate the relationship between rates of antigenic drift and USA incidence in A/H3N2, A/H1N1, B/Vic and B/Yam. We take estimates of the rate of antigenic drift from the preceding section and compare these rates to measurements of relative incidence for each influenza clade calculated from the proportion of influenza viruses attributable to each clade (see Materials and methods). We analyze incidence from the 1998/1999 to the 2008/2009 seasons to avoid possible complications from the 2009 pandemic, finding that A/H3N2 accounted for 56%, A/H1N1 for 20%, B/Vic for 10% and B/Yam for 14% of viruses detected in the USA during this time period. We find a significant correlation between rate of antigenic drift and relative incidence across the four clades (Pearson correlation, $r = 0.96$, $p = 0.021$).

We follow-up this analysis with a more detailed analysis of year-to-year variation in antigenic drift and clade-specific incidence in the USA. We measure year-to-year antigenic drift within each clade as described in the previous section and shown in Figure 5. We measure seasonal incidence of each clade by taking average influenza-like illness (ILI) percentage and multiplying this by the proportion of viruses attributable to a clade for each season. This measure of incidence has previously been shown to have have predictive power in the analysis of seasonal influenza trends [16]. In correlating antigenic drift to seasonal incidence we standardize this measure of incidence across seasons and across clades to measure effects in terms of standard deviations of incidence. We find that years with high levels of antigenic drift coming into an influenza season, e.g. drift in A/H3N2 from 2000 to 2001, show high levels of incidence in that season, e.g. incidence of A/H3N2 in the 2001/2002 season (Figure 5, Figure 6A). Here, within-clade drift explains 35% of the variance in standardized incidence and represents a significant correlation (Pearson correlation, $r = 0.60$, $p = 1.6 \times 10^{-5}$).

We further show that incidence within a clade decreases with antigenic drift in other clades, e.g. drift in A/H1N1, B/Vic and B/Yam from 2000 to 2001 decrease incidence of A/H3N2 in the 2001/2002 season (Figure 6B). This also represents a significant relationship (Pearson correlation, $r = -0.20$, $p = 0.024$). We fit a linear model that predicts standardized incidence based on drift within each of the four lineages, according to the
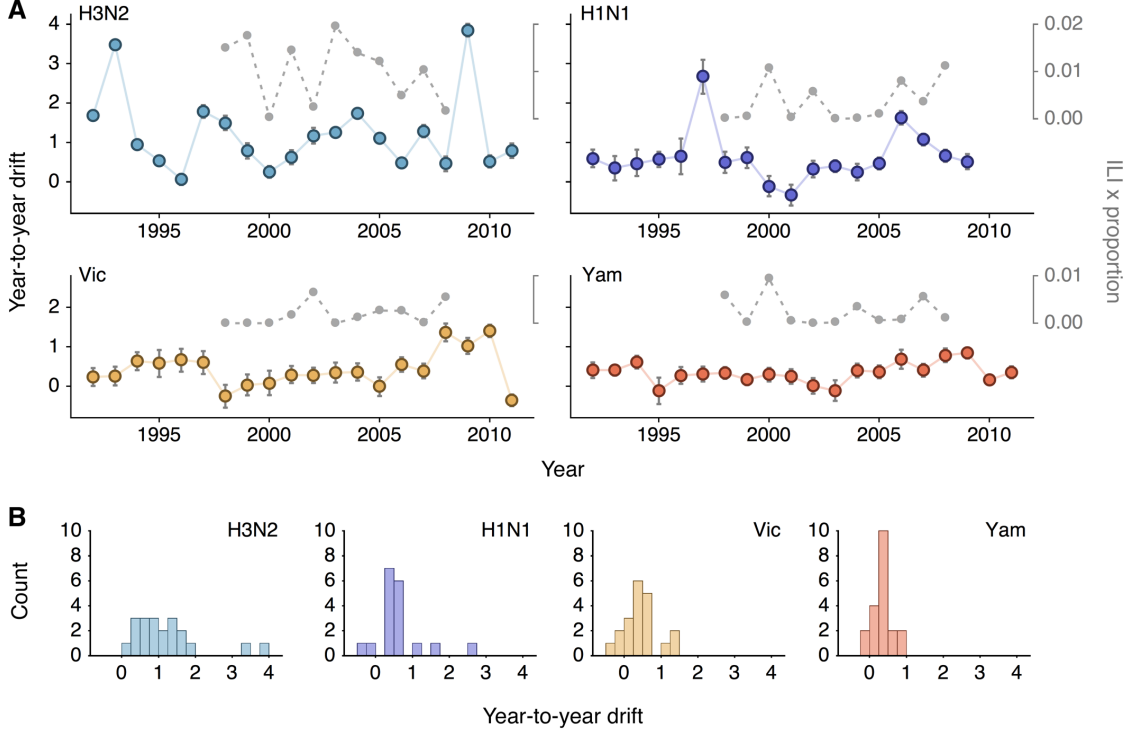
Figure 5. Year-to-year antigenic drift between 1992 and 2009 in A/H3N2, A/H1N1, B/Vic and B/Yam viruses. (A) Time series of year-to-year antigenic drift across influenza clades from 1992 to 2011 and seasonal incidence from 1998/1999 to 2008/2009. Colored lines represent year-to-year antigenic drift, and drift for a virus clade for year i is measured as the mean of antigenic dimension 1 of phylogenetic lineages in year i compared to the mean of antigenic dimension 1 of phylogenetic lineages from the previous year $i-1$. For example, 2000 represents difference in antigenic dimension 1 between 1999 and 2000. Error bars represent 50% Bayesian credible intervals of year-to-year drift. Gray dotted lines represent clade-specific USA incidence taken as average influenza-like illness (ILI) multiplied by proportion of viruses attributable to a clade for each season. Here, 2000 represents the 2000/2001 influenza season. (B) Histograms of year-to-year antigenic drift across influenza clades summarizing observations from 1992 to 2011.

following equations

$$y_{\text{H3}} = \beta_w\, x_{\text{H3}} + \beta_s\, x_{\text{H1}} + \beta_t\, x_{\text{Vic}} + \beta_t\, x_{\text{Yam}}$$
$$y_{\text{H1}} = \beta_s\, x_{\text{H3}} + \beta_w\, x_{\text{H1}} + \beta_t\, x_{\text{Vic}} + \beta_t\, x_{\text{Yam}}$$
$$y_{\text{Vic}} = \beta_t\, x_{\text{H3}} + \beta_t\, x_{\text{H1}} + \beta_w\, x_{\text{Vic}} + \beta_s\, x_{\text{Yam}}$$
$$y_{\text{Yam}} = \beta_t\, x_{\text{H3}} + \beta_t\, x_{\text{H1}} + \beta_s\, x_{\text{Vic}} + \beta_w\, x_{\text{Yam}}, \tag{1}$$

where y represents standardized incidence, x represents drift in antigenic units, β_w scales within-clade drift, β_s scales drift across sister viruses of the same type, and β_t scales drift across viruses of differing type (Table 3). For example, standardized incidence in A/H3N2 would be explained by drift within A/H3N2 scaled by β_w, drift in A/H1N1 scaled by β_s and by drift in B/Vic and B/Yam scaled by β_t. We compared adjusted R^2 across models with different regression coefficients included, and arrived at a best-fitting model that explains 46% of the variance in standardized incidence and includes the terms β_w, β_s and
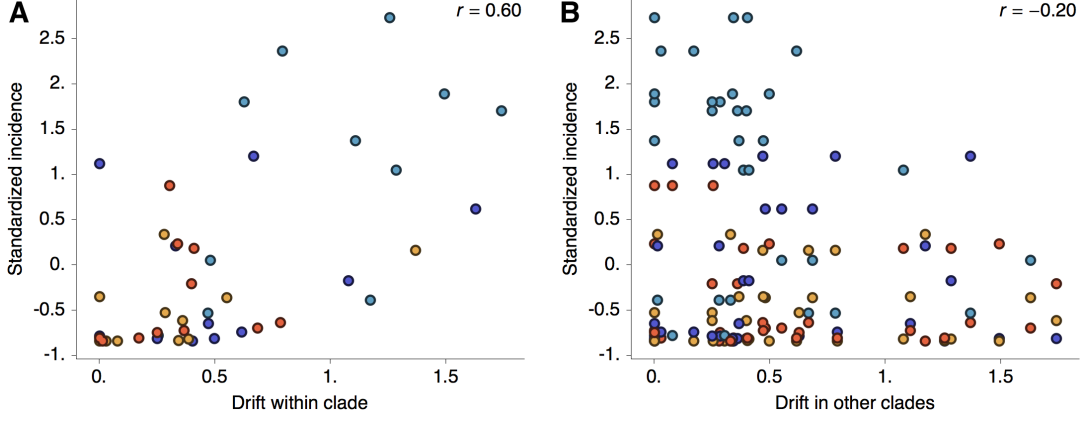
Figure 6. Relationship between antigenic drift within and between clades and seasonal USA incidence for years 1998 to 2009. (A) Standardized incidence within a clade compared to antigenic drift within the same clade. For example, incidence of A/H3N2 in season 1998/1999 is measured against year-to-year antigenic drift of A/H3N2 from 1997 to 1998. (B) Standardized incidence within a clade compared to antigenic drift in other clades. For example, incidence of A/H3N2 in season 1998/1999 is measured against year-to-year antigenic drift of A/H1N1 from 1997 to 1998, as well as antigenic drift of B/Vic from 1997 to 1998. Standardized incidence is measured as yearly ILI percentage multiplied by proportion of viruses identified in a particular clade, and then standardized so that mean and variance across seasons and clades is equal to 0 and 1 respectively. Colors represent the clade in which incidence is measured; light blue for A/H3N2, purple for A/H1N1, orange for B/Vic and red for B/Yam.

β_t, but does not include an intercept term. We find that 1 unit of antigenic drift within a clade increases incidence by 1.26 standard deviations, while 1 unit of drift in a clade of the same type decreases incidence by 0.61 standard deviations and 1 unit of drift in clades of differing type decreases incidence by 0.32 standard deviations. The term for across lineage interference β_t is small in magnitude, but remains statistically significant (Table 3).

Table 3. Estimates of regression coefficients for the relationship of clade-specific USA incidence to antigenic drift within and across clades, where standardized incidence in A/H3N2 is predicted by $\beta_w\, x_{\mathrm{H3}} + \beta_s\, x_{\mathrm{H1}} + \beta_t\, x_{\mathrm{Vic}} + \beta_t\, x_{\mathrm{Yam}}$ and standardized incidence in other clades follows an analogous relationship.

Parameter	Estimate	Standard error	p-value
β_w	1.26	0.22	1.1×10^{-6}
β_s	-0.61	0.22	0.008
β_t	-0.32	0.11	0.005

These findings clearly demonstrate the epidemiological importance of antigenic drift and show that incidence across A/H3N2, A/H1N1, B/Vic and B/Yam is strongly influenced by levels of antigenic drift within each clade, consistent with antigenic drift increasing the susceptible pool available to a clade. The finding of a negative correlation between incidence within one clade and antigenic drift in other clades supports the hypothesis that epidemiological interference exists between clades [16, 17]. The immunological basis of between-clade interference are not entirely clear, however cytotoxic T cells (CTLs) [18],

12

antibodies that broadly neutralize HA [19] and antibodies to internal proteins [20] have been hypothesized as possible mechanisms. Interestingly, we find that interference between sister clades, e.g. B/Vic and B/Yam, appears stronger than interference between more divergent clades, e.g. B/Vic and A/H3N2 (Table 3, $\beta_s = -0.61$ vs $\beta_t = -0.32$). If interference is mediated by CTLs or by antibodies to evolutionary conserved epitopes, then we would expect interference to decrease with evolutionary distance [18].

Furthermore, these findings suggest that the predominant clade of a particular influenza season could be estimated ahead of time by examining the degree of variation in antigenic drift between influenza clades. However, this prediction would not be perfect, and 56% of the variance in incidence remains unexplained by these estimates of antigenic drift. Some of this variance may be explained by incorporating geographic variation in year-to-year antigenic drift; although there exists significant geographic variation in clade-specific incidence [21], we calculated antigenic drift from global samples, rather than USA-specific samples.

Conclusions

Understanding antigenic evolution in seasonal influenza is crucial to our efforts of surveillance and control. Cartographic methods allow complex HI datasets to be compressed to more approachable location-based summaries that quantify antigenic relationships between strains, including relationships not directly assessed via HI. In this study we provide a foundation for evolutionary antigenic cartography, which seeks to simultaneously assess antigenic phenotype and antigenic evolution. We use this approach to characterize competitive dynamics across influenza clades A/H3N2, A/H1N1, B/Vic and B/Yam and show that antigenic evolution drives strain replacement in each clade and interference between clades. We find that varying levels of mutational input and varying efficiencies of the conversion of antigenic polymorphism into fixed differences results in different rates of antigenic drift across influenza clades. Influenza A/H3N2 benefits from a higher rate of new antigenic mutation coupled with strong adaptive evolution and evolves in antigenic phenotype faster than A/H1N1, B/Vic and B/Yam. Correspondingly, we observe substantially greater levels of incidence in A/H3N2 than in other influenza clades. We suggest that antigenic evolution strongly influences competitive dynamics both within and between influenza clades.

The statistical framework presented here represents a baseline to which further advancements in modeling antigenic phenotype and evolution may be made. For example, our likelihood-based model facilitates the inclusion of possible covariates affecting immunological titer, which could include experimental factors such as red blood cell type used in the HI assay [22] and whether oseltamivir is included in the HI reaction [23]. Additionally, this framework should be ideally suited to uncovering genetic determinants of antigenic change, as both the sequence state and antigenic location of internal nodes in the phylogeny may be estimated. In this fashion, it should be possible to correlate sequence substitutions directly to antigenic diffusion.

Identifying viruses that will come to predominate in the global virus population while they

are still at low frequency remains an enormous challenge. However, combining evolutionary and antigenic information may eventually prove useful in identifying low-frequency, but expanding, lineages of antigenically novel viruses that represent ideal targets for vaccine strain selection.

Materials and methods

Antigenic cartography

Antigenic characteristics of viral strains are often assessed through immunological assays such as the hemagglutination inhibition (HI) assay [6]. At heart, these assays compare the reactivity of one virus strain to antibodies raised against another virus strain via challenge or vaccination. In the case of HI, the measurement of cross-reactivity takes the form of a titer representing the dilution factor at which serum raised against a particular virus ceases to be effective at inhibiting the binding of another virus to red blood cells. These factors are commonly assessed by serial dilution, so that HI titers will form a log series, 40, 80, 160, etc .... Because experimental HI titers typically differ by factors of two, we find it convenient to work in $\log_2$ space and represent the titer of virus i against serum j as $H_{ij} = \log_2(\text{HI titer})$, i.e. a titer of 160 has $H_{ij} = 7.32$. Due to experimental constraints, most comparisons cannot be made, leading to a sparse observation matrix $\mathbf{H} = \{H_{ij}\}$. Further, measurements are usually interval and truncated, e.g. inhibition may cease somewhere between the serial titers of 160 and 320, or inhibition may be absent at all titers assayed, suggesting a threshold somewhere between 0 and 40.

Previous work [7, 24] has used multidimensional scaling to place viruses and sera on an 'antigenic map'. These methods heuristically optimize locations of viruses and sera by seeking to minimize the sum of squared errors between titers predicted by map locations and observed titers. Antigenic maps produced by these methods have proved useful in categorizing virus phenotypes [7], but the extension of these methods to integrate genetic data remains notably lacking.

Here, we follow previous models in representing antigenic locations as points in a low P-dimensional antigenic map. One of our initial goals is to find an optimal projection of the high-dimensional distance matrix $\mathbf{H}$ into this lower dimensional space. We conduct this projection using Bayesian multidimensional scaling (BMDS) [25] in which we construct a probabilistic model to quantify the fit of a particular configuration of cartographic locations to the observed matrix of serological measurements. Typically, $P = 2$, but higher or lower dimensions may better reflect the data.

Let $\mathbf{x}_i \in \mathbb{R}^P$ represent the cartographic location of virus i for $i = 1, \ldots, n$, so that $\mathbf{x}_i = (x_{i1}, x_{i2})'$ for $P = 2$. Similarly, let $\mathbf{y}_j$ represent the cartographic location of serum j for $j = 1, \ldots, k$, so that $\mathbf{y}_j = (y_{j1}, y_{j2})'$ for $P = 2$. For notational compactness, we collect together all virus coordinates into an $n \times P$ matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)'$ and all serum coordinates into an $k \times P$ matrix $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_k)'$. Virus and serum may be isolated from / raised against the same strain and have different cartographic locations, and separate serum isolates raised against the same strain may also have different cartographic locations.

This gives a set of distances between virus and serum cartographic locations

$$\delta_{ij} = ||\mathbf{x}_i - \mathbf{y}_j||_2, \tag{2}$$

where $|| \cdot ||_2$ is an L_2 norm.

Traditional approaches to antigenic cartography [7] begin by defining immunological distance as

$$d_{ij} = s_j - H_{ij}, \tag{3}$$

where H_{ij} is the $\log_2$ titer of virus i against serum j and serum effect $s_j = \max(H_{1j}, \ldots, H_{nj})$ is fixed. In following multidimensional scaling (MDS), these approaches attempt to optimize over unknown $\mathbf{X}$ and $\mathbf{Y}$ such that

$$\sum_{(i,j)\in\mathcal{I}} (\delta_{ij} - d_{ij})^2 \tag{4}$$

is minimized, where $\mathcal{I} = \{(i, j) : H_{ij} \text{ is measured}\}$. In the case of threshold measurements, this error function is modified slightly; see [7] for further details.

Here, we instead assume a probabilistic interpretation in which an observed titer is normally distributed around its cartographic expectation with variance φ^2,

$$H_{ij} \sim \mathcal{N}(s_j - \delta_{ij}, \varphi^2). \tag{5}$$

Consequently, the likelihood of observing an exact titer given the placement of antigenic locations is

$$f_|(H_{ij}) = \phi\left(\frac{H_{ij} + \delta_{ij} - s_j}{\varphi}\right), \tag{6}$$

where $\phi(\cdot)$ represents the standard normal probability density function (PDF). Previous BMDS has employed a sampling density truncated to strictly positive quantities since d_{ij} are directly observed, non-negative quantities. In the antigenic setting, these remain random and can be negative since neither s_j is known nor is H_{ij} observed with much precision.

HI assays sometimes show no inhibition at all measured titrations, e.g. a measurement can be reported as '<40'. In this case, the likelihood of observing the threshold measurement follows the cumulative density of the lower tail of the normal distribution

$$f_{\lrcorner}(H_{ij}) = \Phi\left(\frac{H_{ij} + \delta_{ij} - s_j}{\varphi}\right), \tag{7}$$

where $\Phi(\cdot)$ represents the standard normal cumulative distribution function (CDF). Although it is simplest to assume that immunological measurements represent point estimates, it seems more natural to assume that the threshold for inhibition occurs between two titers, e.g. we observe inhibition at 1:160 dilution and no inhibition at 1:320 dilution. Rather than taking the HI titer as 160, we can instead treat this as an interval measurement, assuming that the exact titer for inhibition would occur somewhere between 160
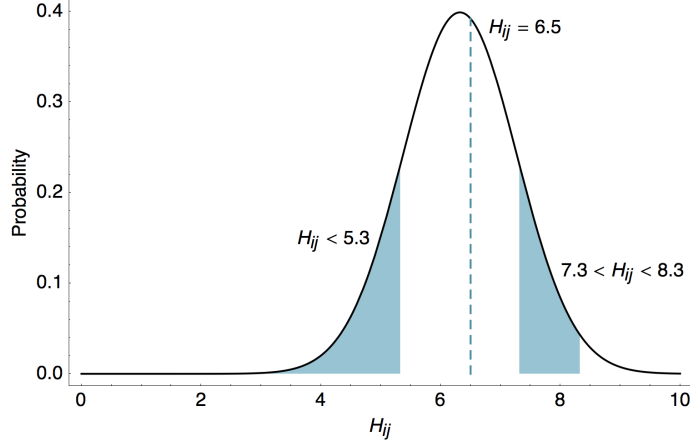
Figure 7. Likelihood of HI titers in the BMDS model. Here we show the likelihoods of observing three different outcomes given $\delta_{ij} = 4$, $\varphi = 0.95$ and $s_j = \log_2(1280) = 10.32$. The likelihood of observing a threshold titer of '<40' is equal to the lower tail of the probability density function $f_\lrcorner(5.32) = 0.146$. The likelihood of observing a point measurement with an exact inhibiting titer of '90.5' is equal to the density function $f_|(6.5) = 0.413$. The likelihood of observing an interval measurement with an inhibiting titer somewhere between '160' and '320' is equal to $f_\sqcup(7.32) = 0.129$.

and 320. HI titers are usually reported as the highest titer that successfully inhibits virus binding, so that in this case, we calculate the likelihood of an interval measurement as

$$f_\sqcup(H_{ij}) = \Phi\left(\frac{H_{ij} + \delta_{ij} - s_j + 1}{\varphi}\right) - \Phi\left(\frac{H_{ij} + \delta_{ij} - s_j}{\varphi}\right). \tag{8}$$

These likelihoods are illustrated in Figure 7. Throughout our analyses, we use interval likelihoods $f_\sqcup$ rather than point likelihoods $f_|$ unless otherwise noted.

We calculate the overall likelihood by multiplying probabilities of individual measurements

$$L(\mathbf{X}, \mathbf{Y}) = \prod_{(i,j) \in \mathcal{I}} f(H_{ij}), \tag{9}$$

using probability functions $f_|$, $f_\lrcorner$ and $f_\sqcup$ as appropriate. We begin by assuming independent, diffuse normal priors on virus and serum locations

$$\mathbf{x}_i \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$$
$$\mathbf{y}_j \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma}), \tag{10}$$

where $\mathbf{m} = (0, \ldots, 0)'$ and $\mathbf{\Sigma}$ is a diagonal matrix with diagonal elements all equal to 10000.

Virus and serum effects

The preceding model represents immunological distance as a drop in titer against the most reactive comparison for a particular serum. However, this model may be biased in some

circumstances. In one example, if a particular serum j is only measured against distant viruses, its maximum titer will be artificially low and the likelihoods concerning this serum will appear poor. To address this issue, we relax the assumption of fixed s_j values and treat the expected $\log_2$ titer when $\delta_{ij} = 0$ as a random variable. In this case, H_{ij} still follows equation 5 with expectation $s_j - \delta_{ij}$, but the vector of 'serum effects' $\mathbf{s} = (s_1, \ldots, s_k)$ is random and estimated rather than fixed. We assume that s_j values are hierarchically distributed according to a normal distribution. We take an Empirical Bayesian approach in specifying the mean and variance of this distribution, set to the empirical mean and empirical variance of the set of maximum titers across sera $\{\max(H_{1j}, \ldots, H_{nj}) : j = 1, \ldots, k\}$. This formulation assumes that particular sera are more reactive in general than other sera.

Additionally, we follow the same logic and assume that some virus isolates are more reactive than other virus isolates and include a 'virus effect' v_i representing the general level of reactivity across HI assays. With virus reactivity included, observed titers follow

$$H_{ij} \sim \mathcal{N}\left(\frac{v_i + s_j}{2} - \delta_{ij}, \, \varphi^2\right), \tag{11}$$

and the vector of virus effects v_i for $i = 1, \ldots, n$ is estimated in an analogous hierarchical fashion, with $\mathbf{v}$ normally distributed with mean and variance equal to the empirical mean and variance of the set of maximum titers across viruses $\{\max(H_{i1}, \ldots, H_{ik}) : i = 1, \ldots, n\}$.

Drift model of antigenic evolution

As presented, multiple configurations of virus and serum locations $\mathbf{X}$ and $\mathbf{Y}$ will give the same likelihood of an observed data matrix $\mathbf{H}$. An example of this phenomenon is shown in Figure 8. In this case, it is impossible to determine from the HI data at hand whether the blue and yellow viruses are antigenically similar (Figure 8A) or antigenically divergent (Figure 8B). This presents an issue of model identifiability, where absolute, as opposed to relative, antigenic locations cannot be determined from observing the serological data alone. Thus, in order to achieve more a interpretable model we impose a weak prior on global locations. In influenza, it's clear that antigenic distance between strains increases with time [7, 24]. To capture this, we replace our previous diffuse prior with an informed prior in which the expected location of viruses and sera increases with date of sampling along dimension one, and each virus and serum location follows an independent normal distribution centered around this temporal expectation, so that

$$\begin{aligned} x_{i1} &\sim \mu\, t_i + \mathcal{N}(0, \sigma_x^2) \\ y_{j1} &\sim \mu\, t_j + \mathcal{N}(0, \sigma_y^2), \end{aligned} \tag{12}$$

where t is the difference between the date of the indexed virus or serum and the date of the earliest sampled virus or serum, and other dimensions follow $x_{im} \sim \mathcal{N}(0, \sigma_x^2)$ and $y_{jm} \sim \mathcal{N}(0, \sigma_y^2)$ for $m \geq 2$. Thus, this model assumes that virus and serum locations drift in a line across the antigenic map at rate μ. The parameter σ_x determines the breadth of the cloud of virus locations at each point in time, while σ_y determines the breadth of the cloud of serum locations.
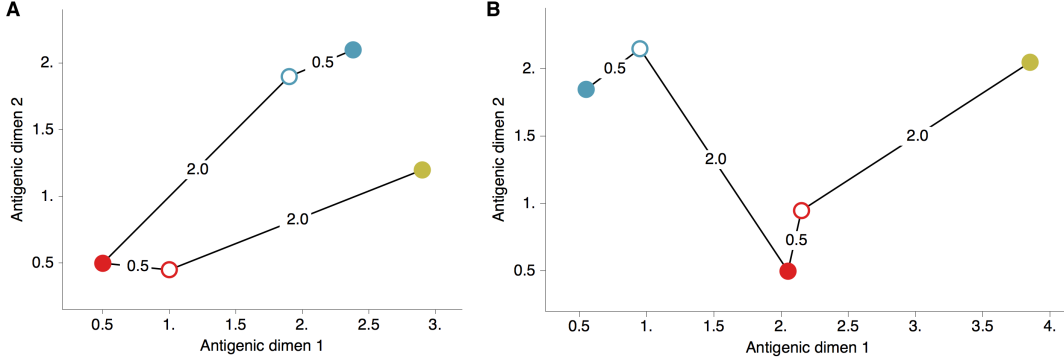
Figure 8. Schematic antigenic map with three viruses and two sera. (A) Map with virus 1 and virus 3 antigenically similar. (B) Map with virus 1 and virus 3 antigenically divergent. Virus 1 is shown in blue, virus 2 is shown in red and virus 3 is shown in yellow. Virus isolates are represented by filled circles, sera raised against viruses are shown as open circles and map distances δ_{ij} are shown as solid lines connecting viruses and sera. Sera from virus 1 is compared against viruses 1 and 2, while sera from virus 2 is compared against viruses 2 and 3. Configurations (A) and (B) represent cartographic models that would give equal likelihoods to a set of serological data $\{H_{11}, H_{21}, H_{22}, H_{32}\}$.

Phylogenetic diffusion model of antigenic evolution

We simultaneously model antigenic locations and genetic relatedness by assuming that virus locations are influenced by evolution following a Brownian motion process [11]. To do this, we replace the previous prior specifying independent virus locations with a prior that incorporates covariance based on shared evolutionary history

$$\mathbf{X} \sim \begin{pmatrix} \mu\,t_1 & 0 \\ \vdots & \vdots \\ \mu\,t_n & 0 \end{pmatrix} + \text{Evolutionary Brownian Process}(\sigma_x, \tau) \tag{13}$$

for $P = 2$, where σ_x is the volatility parameter of the Brownian motion over virus locations and τ is a phylogeny specifying tree topology and branch lengths. Thus, viruses which are genetically similar are induced to have prior locations close to one another on the antigenic map. In the evolutionary Brownian process, the tips of the phylogeny τ correspond to the set of virus locations $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, and the probability of observing tip locations depends on the locations of internal nodes $(\mathbf{x}_{n+1}, \ldots, \mathbf{x}_{2n-2})$ and on the location of the root node $\mathbf{x}_{2n-1}$. This process assumes that a virus location $\mathbf{x}_i$ follows from the location of its parent virus $\mathbf{x}_{f(i)}$, and with the addition of drift along dimension 1, is distributed as

$$\mathbf{x}_i \sim (\mu\,d_i,\, 0)' + \mathcal{N}(\mathbf{x}_{f(i)}, d_i\,\mathbf{\Sigma}) \tag{14}$$

for $P = 2$, where $f(i)$ is a function that maps nodes to parental nodes, d_i is the length of the branch connecting virus i to parent virus $f(i)$, and $\mathbf{\Sigma}$ is a diagonal matrix with diagonal elements all equal to σ_x^2. The root virus location $\mathbf{x}_{2n-1}$ is assumed to follow a normal distribution with expectation $(\mu\,t_{2n-1}, 0)'$ for $P = 2$ and variance determined by the diffusion volatility σ_x [11]. The probability of virus locations $p(\mathbf{X}|\mu, \sigma_x, \tau)$ is determined

18

through analytical integration across internal states following the methods introduced in [11]. This formulation corresponds to a Wiener process with drift, in which the drift term μ only influences the expected states of nodes along the phylogeny, but does not influence the covariance structure among these nodes, which remains the same as it does in a standard Wiener process [26]. This allows the separation in equation 13 between drift terms affecting only expectations and the evolutionary Brownian process that includes covariance among virus locations $\mathbf{x}_1, \ldots, \mathbf{x}_n$.

Here, the phylogenetic tree τ is estimated using sequence data for viruses $1, \ldots, n$ according to well established methods implemented in the software package BEAST [27].

Posterior inference

Top-level priors for $1/\varphi^2$, μ, $1/\sigma_x^2$, and $1/\sigma_y^2$ are assumed to follow diffuse $\mathrm{Gamma}(a, b)$ distributions with $a = 0.001$ and $b = 0.001$. These diffuse priors were chosen to be non-informative and provide little-to-no weight on the resulting posterior distributions. Under the full model, the posterior probability of observing virus and serum locations given immunological data is factored

$$p(\mathbf{X}, \mathbf{Y}|\mathbf{H}) \propto p(\mathbf{H}|\mathbf{X}, \mathbf{Y}, \mathbf{s}, \mathbf{v}, \varphi) \, p(\mathbf{X}|\mu, \sigma_x, \tau) \, p(\mathbf{Y}|\mu, \sigma_y) \, p(\mathbf{s}, \mathbf{v}, \varphi, \mu, \sigma_x, \sigma_y, \tau). \tag{15}$$

We sample from this posterior distribution using the MCMC procedures implemented in the software package BEAST [27]. Metropolis-Hastings proposals include transition kernels that translate individual virus and serum locations $\mathbf{x}_i$ and $\mathbf{y}_j$ and individual virus effects v_i and serum effects s_j, and other transition kernels that scale the entire set of virus and serum locations $\mathbf{X}$ and $\mathbf{Y}$ and that scale parameters φ, μ, σ_x and σ_y. For the present analysis, a two-step approach was taken to sample phylogenies, where a posterior sample of phylogenies was gathered using sequence data and then, in the cartographic analysis, trees from this set were randomly proposed and accepted following the Metropolis-Hastings algorithm [28].

Genetic, antigenic and surveillance data

We compiled an antigenic dataset of hemagglutination inhibition (HI) measurements of virus isolates against post-infection ferret sera for influenza A/H3N2 by collecting data from previous publications [7, 9, 10, 12], NIMR vaccine strain selection reports for 2002 and 2008–2012 [29–35] and the Feb 2011 VRBPAC report [36]. We queried the Influenza Research Database [37] and the EpiFlu Database [38] for HA nucleotide sequences by matching strain names, e.g. A/HongKong/1/1968, and only strains for which sequence was present was retained. If a strain had multiple sequences in the databases we preferentially kept the IRD sequence and preferentially kept the longest sequence in IRD. Many strains had full length HA sequences, while other strains only possessed HA1 sequences. Sequences were aligned using MUSCLE v3.7 under default parameters [39]. This dataset had 2051 influenza isolates (present as either virus or serum in HI comparisons) dating from 1968 to 2011. However, the majority of isolates were present from 2002 to 2007. Because we are interested in longer-term antigenic evolution, we subsampled the data to have at

most 20 virus isolates per year, preferentially keeping those isolates with more antigenic comparisons. We then kept only those serum isolates that are relatively informative to the antigenic placement of viruses, dropping serum isolates that are compared to 4 or fewer different virus isolates. This censoring left 402 virus isolates, 519 serum isolates and 10,059 HI measurements. Each virus isolate was compared to an average of 21.9 serum isolates, and each serum isolate was compared to an average of 18.0 virus isolates.

Antigenic data for influenza A/H1N1 was collected from previous publications [9, 12, 40–53] and NIMR vaccine strain selection reports for 2002–2010 [29–32, 54–62]. The same procedure was followed as was followed for H3N2 to match sequence data and to subsample antigenic comparisons. This procedure yielded 115 virus isolates, 77 serum isolates and 1882 HI measurements over the course of 1977 to 2009. Each virus isolate was compared to an average of 10.0 serum isolates, and each serum isolate was compared to an average of 16.2 virus isolates.

Antigenic comparisons for influenza B/Victoria were collated from previous publications [3, 9, 63–70] and vaccine strain selection reports for 2002–2012 [29–35, 54–61, 71–73]. Here, the sequence matching and subsampling procedure yielded 179 virus isolates, 70 serum isolates and 2003 HI measurements over the course of 1986 to 2011. Each virus isolate was compared to an average of 6.5 serum isolates, and each serum isolate was compared to an average of 16.7 virus isolates.

Antigenic comparisons for influenza B/Yamagata were collected from previous publications [3, 9, 63–70, 74–80] and vaccine strain selection reports for 2002–2012 [29–35, 54–61, 71–73]. For B/Yamagata, the matching and subsampling procedure resulted in 174 virus isolates, 69 serum isolates and 1962 HI measurements over the course of 1987 to 2011. Each virus isolate was compared to an average of 6.9 serum isolates, and each serum isolate was compared to an average of 17.3 virus isolates.

Surveillance data was obtained from the Centers of Disease Control and Prevention Flu-View Influenza Reports from the yearly summaries of influenza seasons 1997–1998 to 2010–2011 [81]. As an example, one report states "collaborating laboratories in the United States tested 195,744 respiratory specimens for influenza viruses, 27,682 (14%) of which were positive. Of these, 18,175 (66%) were positive for influenza A viruses, and 9,507 (34%) were positive for influenza B viruses. Of the 18,175 specimens positive for influenza A viruses, 7,631 (42%) were subtyped; 6,762 (87%) of these were seasonal influenza A (H1N1) viruses, and 869 (13%) were influenza A (H3N2) viruses." In this case, we estimate the relative proportion of A/H3N2 of the four clades as $0.66 \times 0.13 = 0.09$. Similar calculations were performed for A/H1N1, B/Vic and B/Yam.

Implementation

Phylogenetic trees were estimated for A/H3N2, A/H1N1, B/Vic and B/Yam using BEAST [27] and incorporated the SRD06 nucleotide substitution model [82], a coalescent demographic model with constant effective population size and a strict molecular clock across branches. MCMC was run for 60 million steps and trees were sampled every 50,000 steps after allowing a burn-in of 10 million steps, yielding a total sample of 2000 trees. These

trees were treated as a discrete set of possibilities when subsequently sampled in the BMDS analysis [28]. However, it would be possible to jointly sample from sequence data and serological data using these methods.

MCMC was used to sample virus locations $\mathbf{X}$, serum locations $\mathbf{Y}$, virus effects $\mathbf{v}$, serum effects $\mathbf{s}$, MDS precision $1/\varphi^2$, antigenic drift rate μ, virus location precision $1/\sigma_x^2$, serum location precision $1/\sigma_y^2$ and phylogenetic tree τ. MCMC chains were run for 500 million steps and parameter values sampled every 200,000 steps after a burn-in of 100 million steps, yielding a total of 2000 MCMC samples. In all cases when drift parameter μ was included the MCMC chain mixed well and arrived at the same estimated posterior distribution from different starting points. However, without drift parameter μ, maps for A/H3N2 showed some degree of metastability, where some chains would converge on one solution and other chains would converge on a different solution. We favor models that include μ, because its inclusion, in addition to correcting most identifiability issues, yields much improved mixing of antigenic locations.

There is some difficulty summarizing posterior cartographic samples, as sampled virus and serum locations represent only relative quantities, and because of this, over the course of the MCMC, virus locations may shift. Our prior on virus and serum locations removes much of this issue, orienting the antigenic map along dimension 1 and fixing it to begin at the origin. However, local isometries are often still a problem. For example, in A/H3N2 the HK/68, EN/72 and VI/75 clusters may rotate in relation to other clusters. Consequently, it may be difficult to fully align MCMC samples using Procrustes analysis. For the present study, we take a simple approach and sample a single MCMC step and visualize the antigenic locations at this state (Figures 2, Figure 3). Then, for specific quantities of interest, like rate of antigenic drift and rate of diffusion at different points along the phylogeny, we calculate the quantity across MCMC samples to yield an expectation and a credible interval. This approach accurately characterizes uncertainty that may be hidden in an analysis of a single antigenic map.

We summarize diffusion paths of viral lineages (Figure 2, Figure 3) by taking each virus and reconstructing x and y locations along antigenic dimensions 1 and 2 backward through time. We use MCMC to sample tip locations, but when outputting trees sample internal node locations using a peeling algorithm as described in [83]. Thus, after the MCMC is finished we have a posterior sample of 2000 trees each tagged with estimated tip locations and internal node locations. We post-processed each posterior tree by conducting a linear interpolation between parent-child node locations to arrive at x and y values at intervals of 0.05 years for each virus. Then, for each interval, x and y values are averaged across the sample of posterior trees. We draw lines between these locations to approximate mean posterior diffusion paths. As virus lineages coalesce backwards through time down the phylogeny these diffusion paths will also coalesce.

Comparison with previous results

Here, we attempt to compare antigenic locations inferred by our BMDS model to antigenic locations previously inferred by the error minimization methods of Smith et al. [7], referred to here as antigenic cartography by MDS. For this comparison, we use exactly the same HI

data used to produce the results in [7], consisting of 273 virus isolates, 79 serum isolates and a total of 4252 HI measurements taken between 1968 and 2003. We begin with a BMDS analog of the antigenic model used in [7], where serum effects are taken as the maximum titer of a particular ferret serum and the expected $\log_2$ drop in HI titer is proportional to Euclidean distance between virus and serum locations. To bring models into further alignment, we use a Uniform$(-100, 100)$ distribution over virus locations and serum locations. Unsurprisingly, we find that this BMDS model produces results that are strongly congruent with MDS results (Figure 9). Antigenic cluster locations are consistent between methods (Figure 9A–B) and antigenic distances between pairs of viruses are consistent between temporally similar and temporally divergent viruses (Figure 9C–E), suggesting that the resulting maps are consistent at both local and global scales. Credible intervals of antigenic distances for the BMDS model remain narrow across the temporal spectrum (Figure 9C–E), implying a fair degree of rigidity to the map.
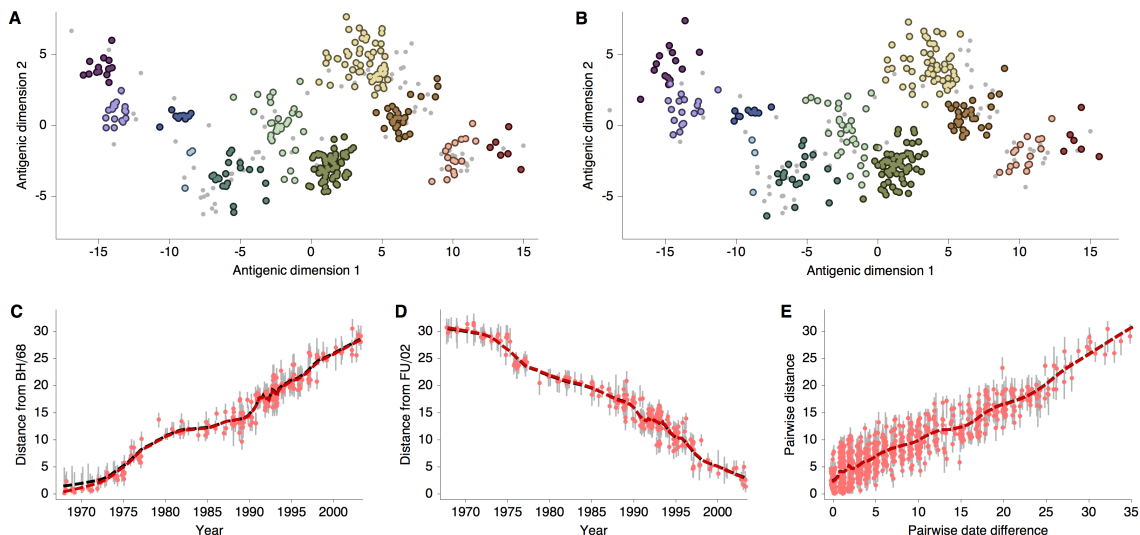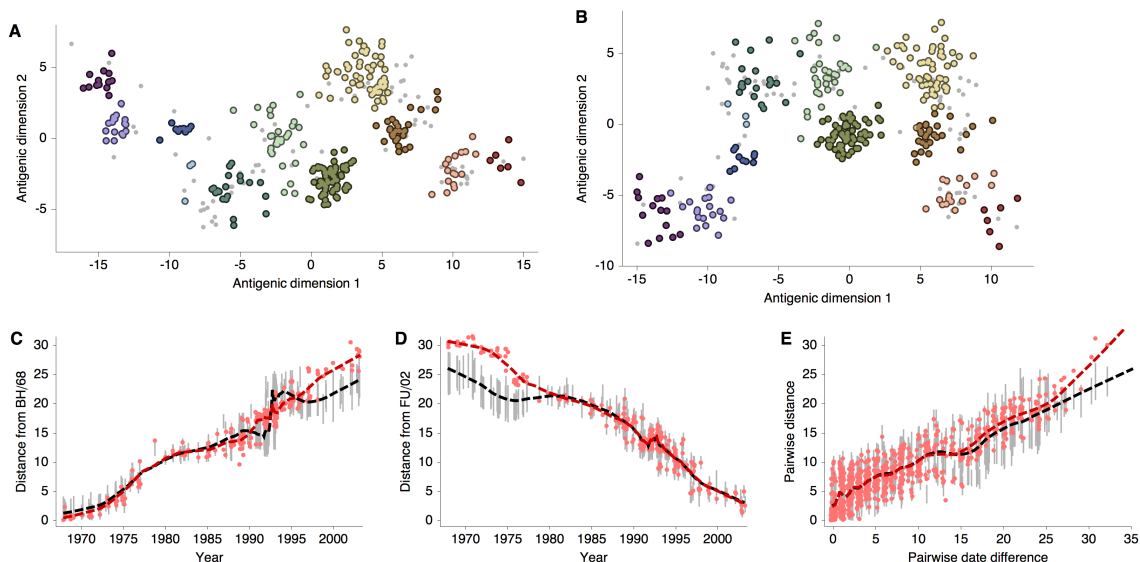

Figure 9. Comparison of A/H3N2 antigenic locations estimated by Smith et al. [7] using MDS and an equivalent BMDS model. (A) MDS antigenic locations, reoriented so that the primary dimension lies on the x-axis rather than on the y-axis as in Figure 1 of [7]. (B) A posterior sample of antigenic locations from an equivalent BMDS model. In (A) and (B), viruses are shown as colored circles, with color denoting antigenic cluster inferred by [7], and sera are shown as gray points. (C) Antigenic distances between A/Bilthoven/16190/1968 and all other viruses determined for both methods. (D) Antigenic distances between A/Fujian/411/2002 and all other viruses determined for both methods. (E) Antigenic distances between 750 random pairs of viruses determined for both methods. In (C), (D) and (E) red points show distances for the MDS model and gray bars show the 95% credible interval of distances for the BMDS model, while the red dashed line shows a LOESS regression to MDS distances, and the black dashed line shows a LOESS regression to the BMDS distances. The BMDS model has a Uniform$(-100, 100)$ prior on antigenic locations and serum effects fixed at maximum titer values.

Smith et al. [7] show that there exist at least two solutions in their assignment of antigenic locations, involving the rotation of clusters HK/68, EN/72 and VI/75 (shown in Figure S2 of [7]). We observe the same metastable behavior in our analysis; some MCMC chains converge on the solution shown in Figure 9B, while other MCMC chains converge on the

alternative solution shown in Figure 10B. The distribution of likelihood values appear highly similar between these two solutions, suggesting that they represent global optima. The rotation of the HK/68, EN/72 and VI/75 clusters creates a map that bends slightly, so that temporally distant viruses appear closer in the rotated solution than in the original solution (Figure 10C–E). In this case, it is clear that the solutions are locally consistent between viruses up to ~15 years divergent, even if there some degree of global flexibility.

Figure 10. Comparison of A/H3N2 antigenic locations estimated by Smith et al. [7] using MDS and an equivalent BMDS model under an alternative solution. (A) MDS antigenic locations, reoriented so that the primary dimension lies on the x-axis rather than on the y-axis as in Figure 1 of [7]. (B) A posterior sample of antigenic locations from an equivalent BMDS model that has converged on the alternative solution. In (A) and (B), viruses are shown as colored circles, with color denoting antigenic cluster inferred by [7], and sera are shown as gray points. (C) Antigenic distances between A/Bilthoven/16190/1968 and all other viruses determined for both methods. (D) Antigenic distances between A/Fujian/411/2002 and all other viruses determined for both methods. (E) Antigenic distances between 750 random pairs of viruses determined for both methods. In (C), (D) and (E) red points show distances for the MDS model and gray bars show the 95% credible interval of distances for the BMDS model, while the red dashed line shows a LOESS regression to MDS distances, and the black dashed line shows a LOESS regression to the BMDS distances. The BMDS model has a Uniform$(-100, 100)$ prior on antigenic locations and serum effects fixed at maximum titer values.

As discussed in the main text, the presence of multiple optima with different degrees of 2D curvature implies an issue of identifiability; the HI likelihood model alone cannot distinguish between these possibilities. Because of this issue, and to more easily estimate rates of antigenic drift, we include a model of systematic drift in antigenic location that favors linear movement in the antigenic map. We find that including this drift prior on antigenic locations removes the problem of identifiability. Antigenic locations produced by this model remain locally consistent with MDS results between viruses ~15 years divergent, but global comparisons show that this BMDS model has partitioned more variance to the first antigenic dimension (Figure 11). We additionally find that including the drift prior

on antigenic locations often results in greater predictive power, with a slight improvement of test error for the A/H1N1, B/Vic and B/Yam datasets (Table 1).
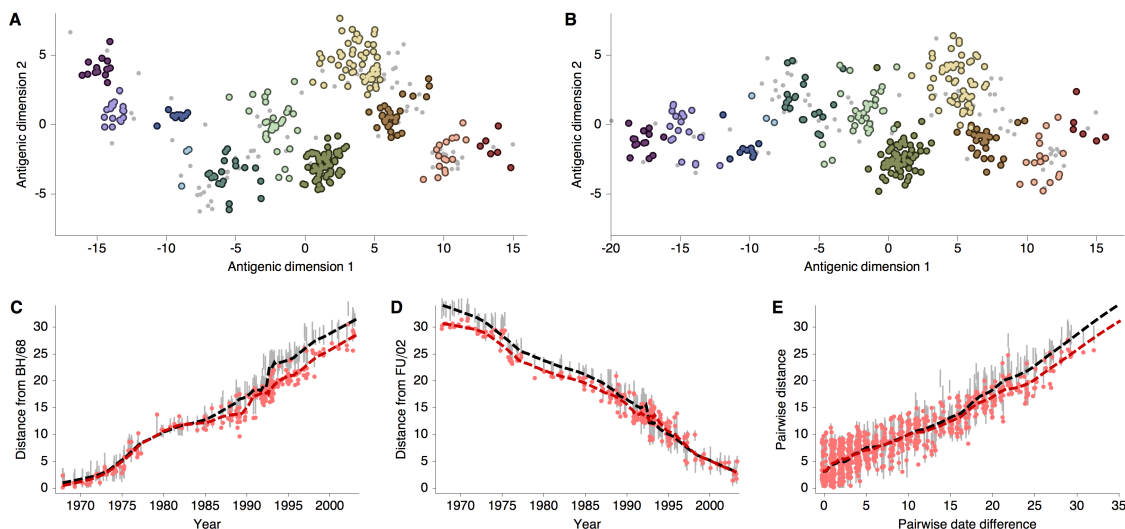

Figure 11. Comparison of A/H3N2 antigenic locations estimated by Smith et al. [7] using MDS and an extended BMDS model that includes date-informed priors on antigenic locations. (A) MDS antigenic locations, reoriented so that the primary dimension lies on the x-axis rather than on the y-axis as in Figure 1 of [7]. (B) A posterior sample of antigenic locations from a BMDS model that includes date-informed priors on antigenic locations. In (A) and (B), viruses are shown as colored circles, with color denoting antigenic cluster inferred by [7], and sera are shown as gray points. (C) Antigenic distances between A/Bilthoven/16190/1968 and all other viruses determined for both methods. (D) Antigenic distances between A/Fujian/411/2002 and all other viruses determined for both methods. (E) Antigenic distances between 750 random pairs of viruses determined for both methods. In (C), (D) and (E) red points show distances for the MDS model and gray bars show the 95% credible interval of distances for the BMDS model, while the red dashed line shows a LOESS regression to MDS distances, and the black dashed line shows a LOESS regression to the BMDS distances. The BMDS model has a date-informed prior on antigenic locations and serum effects fixed at maximum titer values.

Our final BMDS model (model 9, Table 1) differs from antigenic model used by Smith et al. [7] in including temporally- and phylogenetically-informed priors on antigenic locations and also in estimating serum and virus effects. Here, we investigate the impact on antigenic locations of estimating virus and serum effects in the BMDS model. To isolate this difference, we use a Uniform$(-100, 100)$ prior on antigenic locations. Surprisingly, estimating virus and serum effects results in a more linear antigenic map (Figure 12), resembling the appearance of the map incorporating the antigenic drift prior, while preserving local consistency. We generally observe congruence between MDS and BMDS antigenic locations for viruses less than $\sim$10 years divergent (Figure 12E). However, specific viruses may be affected, for instance A/Bilthoven/16190/1968 (Figure 12C), which appears more distant from all other viruses when serum and virus effects are included.

In this dataset, viruses 15 or more years divergent always yield threshold titers, and hence, their relative locations must be indirectly inferred rather than through direct comparison. This may explain why we observe local consistency between models at scales less than
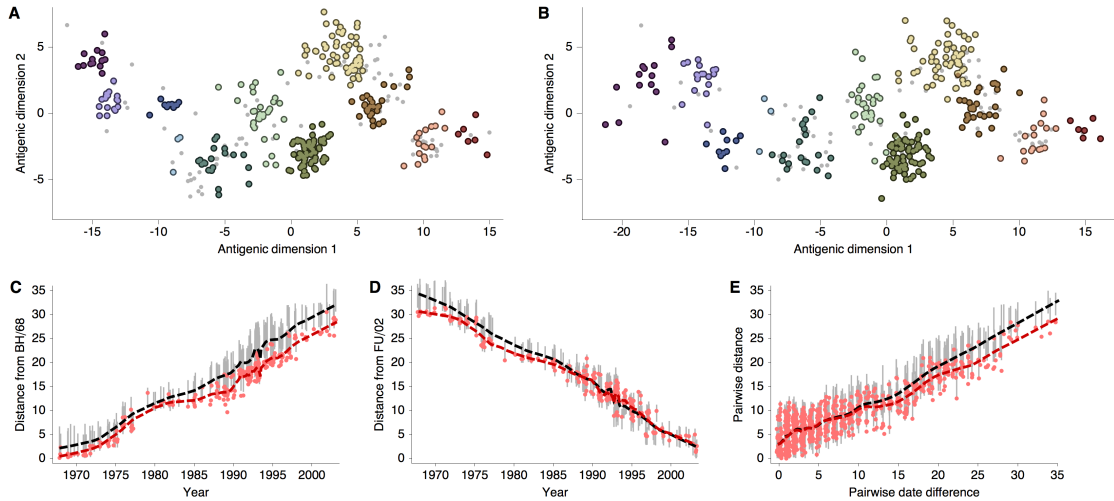
Figure 12. Comparison of A/H3N2 antigenic locations estimated by Smith et al. [7] using MDS and an extended BMDS model that estimates serum and virus effects. (A) MDS antigenic locations, reoriented so that the primary dimension lies on the x-axis rather than on the y-axis as in Figure 1 of [7]. (B) A posterior sample of antigenic locations from a BMDS model that estimates virus and serum effects. In (A) and (B), viruses are shown as colored circles, with color denoting antigenic cluster inferred by [7], and sera are shown as gray points. (C) Antigenic distances between A/Bilthoven/16190/1968 and all other viruses determined for both methods. (D) Antigenic distances between A/Fujian/411/2002 and all other viruses determined for both methods. (E) Antigenic distances between 750 random pairs of viruses determined for both methods. In (C), (D) and (E) red points show distances for the MDS model and gray bars show the 95% credible interval of distances for the BMDS model, while the red dashed line shows a LOESS regression to MDS distances, and the black dashed line shows a LOESS regression to the BMDS distances. The BMDS model has a Uniform$(-100, 100)$ prior on antigenic locations and virus and serum effects estimated in a hierarchical Bayesian fashion.

~15 years, but some degree of global inconsistency. Still, these results suggest that, when making local comparisons, such as those used to calculate year-to-year antigenic drift (Figure 3), outcomes are expected to be robust to many model particulars.

Availability

Source code implementing the cartographic models has been made fully available as part of the software package BEAST [27], and can be downloaded from its Google code repository (http://code.google.com/p/beast-mcmc/). Incidence data used in this analysis is archived with Dryad (doi:10.5061/dryad.rc515).

Acknowledgments

and Control, Hospital Clinic of Barcelona, National Institute of Hygiene of Morocco, National Institute of Infectious Diseases (Japan), National Institute for Medical Research (UK), Norwegian Institute of Public Health, Swedish Institute for Infectious Disease Control, Victorian Infectious Diseases Reference Laboratory (Australia).

Author contributions

AR, MAS, TB conceived the study. AR, MAS, TB, DJS, PL designed the antigenic model. TB, GD, VG, JWM, AJH, CAR, DJS, AR gathered antigenic and genetic data. AR, MAS, TB, PL implemented statistical phylogenetic procedures. TB performed the analysis. TB, AR, DJS, JWM, PL, CAR, AJH, GD, MAS interpreted the results. TB, MAS, AR, JWM, AJH wrote the paper.

Funding

References

1. World Health Organization (2009) Influenza Fact sheet. Available at http://www.who.int/mediacentre/factsheets/fs211/en/.

2. Osterholm M, Kelley N, Manske J, Ballering K, Leighton T, et al. (2012) The compelling need for game-changing influenza vaccines: An analysis of the influenza vaccine enterprise and recommendations for the future. Technical report, Center for Infectious Disease Research & Policy.

3. Rota PA, Wallis TR, Harmon MW, Rota JS, Kendal AP, et al. (1990) Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. Virology 175: 59–68.

4. Wiley D, Wilson I, Skehel J (1981) Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature 289: 373–378.

5. Nelson MI, Holmes EC (2007) The evolution of epidemic influenza. Nat Rev Genet 8: 196–205.

6. Hirst G (1943) Studies of antigenic differences among strains of influenza A by means of red cell agglutination. J Exp Med 78: 407–423.

7. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, et al. (2004) Mapping the antigenic and genetic evolution of influenza virus. Science 305: 371–376.

8. Fitch WM, Bush RM, Bender CA, Cox NJ (1997) Long term trends in the evolution of H(3) HA1 human influenza type A. Proc Natl Acad Sci USA 94: 7712-8.

9. Hay, AJ and Gregory, V and Douglas, AR and Lin, YP (2001) The evolution of human influenza viruses. Phil Trans R Soc Lond B 356: 1861–1870.

10. Russell CA, Jones TC, Barr IG, Cox NJ, Garten RJ, et al. (2008) The global circulation of seasonal influenza A (H3N2) viruses. Science 320: 340–346.

11. Lemey P, Rambaut A, Welch J, Suchard M (2010) Phylogeography takes a relaxed random walk in continuous space and time. Mol Biol Evol 27: 1877–1885.

12. Barr I, McCauley J, Cox N, Daniels R, Engelhardt O, 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.

13. Klimov AI, Garten R, Russell C, Barr IG, Besselaar TG, et al. (2012) WHO recommendations for the viruses to be used in the 2012 Southern Hemisphere Influenza Vaccine: Epidemiology, antigenic and genetic characteristics of influenza A (H1N1) pdm09, A (H3N2) and B influenza viruses collected from February to September 2011. Vaccine 30: 6461–6471.

14. Bedford T, Cobey S, Pascual M (2011) Strength and tempo of selection revealed in viral gene genealogies. BMC Evol Biol 11: 220.

15. Bedford T, Rambaut A, Pascual M (2012) Canalization of the evolutionary trajectory of the human influenza virus. BMC Biol 10: 38.

16. Goldstein E, Cobey S, Takahashi S, Miller J, Lipsitch M (2011) Predicting the epidemic sizes of influenza A/H1N1, A/H3N2, and B: a statistical method. PLoS Med 8: e1001051.

17. Ferguson NM, Galvani AP, Bush RM (2003) Ecological and immunological determinants of influenza evolution. Nature 422: 428–433.

18. Voeten J, Bestebroer T, Nieuwkoop N, Fouchier R, Osterhaus A, et al. (2000) Antigenic drift in the influenza A virus (H3N2) nucleoprotein and escape from recognition by cytotoxic T lymphocytes. J Virol 74: 6800–6807.

19. Corti D, Suguitan Jr AL, Pinna D, Silacci C, Fernandez-Rodriguez BM, et al. (2010) Heterosubtypic neutralizing antibodies are produced by individuals immunized with a seasonal influenza vaccine. J Clin Invest 120: 1663–1673.

20. Laidlaw BJ, Decman V, Ali MAA, Abt MC, Wolf AI, et al. (2013) Cooperativity Between CD8+ T Cells, Non-Neutralizing Antibodies, and Alveolar Macrophages Is Important for Heterosubtypic Influenza Virus Immunity. PLoS Pathog 9: e1003207.

21. Finkelman B, Viboud C, Koelle K, Ferrari M, Bharti N, et al. (2007) Global patterns in seasonal activity of influenza A/H3N2, A/H1N1, and B from 1997 to 2005: viral coexistence and latitudinal gradients. PLoS One 2: e1296.

22. Lin Y, Xiong X, Wharton S, Martin S, Coombs P, et al. (2012) Evolution of the receptor binding properties of the influenza A (H3N2) hemagglutinin. Proc Natl Acad Sci USA 109: 21474–21479.

23. Lin Y, Gregory V, Collins P, Kloess J, Wharton S, et al. (2010) Neuraminidase receptor binding variants of human influenza A (H3N2) viruses resulting from substitution of aspartic acid 151 in the catalytic site: a role in virus attachment? J Virol 84: 6769–6781.

24. Cai Z, Zhang T, Wan XF (2010) A computational framework for influenza antigenic cartography. PLoS Comput Biol 6: e1000949.

25. Oh M, Raftery A (2001) Bayesian multidimensional scaling and choice of dimension. J Amer Statist Assoc 96: 1031–1044.

26. Borodin AN, Salminen P (2002) Handbook of Brownian Motion – Facts and Formulae. Birkhäuser Verlag.

27. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29: 1969–1973.

28. Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. Syst Biol 53: 673–684.

29. Hay AJ, Lin YP, Gregory V, Bennet M (2002) Annual Report. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

30. Hay A, Daniels RS, Lin YP, Zheng X, Gregory V, et al. (March 2008) Characteristics of human influenza AH1N1, AH3N2, and B viruses isolated September 2007 to February 2008. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

31. Hay A, Daniels RS, Lin YP, Zheng X, Hou T, et al. (Feb 2009) Antigenic and genetic characteristics of human influenza A(H1N1), A(H3N2) and B viruses isolated during

October 2008 to February 2009. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

32. McCauley J, Daniels R, Lin YP, Zheng X, Hou T, et al. (Feb 2010) Report prepared for the WHO annual consultation on the composition of influenza vaccine for the Northern Hemisphere. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

33. McCauley J, Daniels R, Lin YP, Zheng X, Hou T, et al. (Sep 2010) Report prepared for the WHO annual consultation on the composition of influenza vaccine for the Southern Hemisphere. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

34. McCauley J, Daniels R, Lin YP, Zheng X, Gregory V, et al. (Sep 2011) Report prepared for the WHO annual consultation on the composition of influenza vaccine for the Southern Hemisphere. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

35. McCauley J, Daniels RS, Lin YP, Zheng X, Gregory V, et al. (Feb 2012) Report prepared for the WHO annual consultation on the composition of influenza vaccine for the Northern Hemisphere. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

36. Cox NJ (Feb 2011) Information for the Vaccines and Related Biological Products Advisory Committee. Seasonal Influenza Vaccines. Technical report, WHO Collaborating Centre for Surveillance, Epidemiology and Control of Influenza, Centers for Disease Control and Prevention, USA.

37. Squires R, Noronha J, Hunt V, García-Sastre A, Macken C, et al. (2012) Influenza research database: an integrated bioinformatics resource for influenza research and surveillance. Influenza Other Respi Viruses 6: 404–416.

38. Bogner P, Capua I, Lipman D, Cox N, et al. (2006) A global initiative on sharing avian flu data. Nature 442: 981–981.

39. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792–1797.

40. Kendal AP, Noble GR, Skehel JJ, Dowdle WR (1978) Antigenic similarity of influenza A(H1N1) viruses from epidemics in 1977–1978 to Scandinavian strains isolated in epidemics of 1950–1951. Virology 89: 632–636.

41. Webster R, Kendal A, Gerhard W (1979) Analysis of antigenic drift in recently isolated influenza A (H1N1) viruses using monoclonal antibody preparations. Virology 96: 258–264.

42. Nakajima K, Nakajima S, Nerome K, Takeuchi Y, Sugiura A, et al. (1979) Genetic relatedness of some 1978–1979 influenza H1N1 strains to 1953 H1N1 strain. Virology 99: 423–426.

43. Nakajima S, Cox NJ, Kendal AP (1981) Antigenic and genomic analyses of influenza A(H1N1) viruses from different regions of the world, February 1978 to March 1980. Infection and Immunity 32: 287–294.

44. Chakraverty P, Cunningham P, Pereira MS (1982) The return of the historic influenza A H1N1 virus and its impact on the population of the United Kingdom. The Journal of Hygiene 89: 89–100.

45. Pereira MS, Chakraverty P (1982) Influenza in the United Kingdom 1977–1981. Epidemiology & Infection 88: 501–512.

46. Chakraverty P, Cunningham P, Shen GZ, Pereira MS (1986) Influenza in the United Kingdom 1982–85. The Journal of Hygiene 97: 347–358.

47. Cox NJ, Bai ZS, Kendal AP (1983) Laboratory-based surveillance of influenza A(H1N1) and A(H3N2) viruses in 1980–81: antigenic and genomic analyses. Bulletin of the World Health Organization 61: 143–152.

48. Daniels RS, Douglas AR, Skehel JJ, Wiley DC (1985) Antigenic and amino acid sequence analyses of influenza viruses of the H1N1 subtype isolated between 1982 and 1984. Bulletin of the World Health Organization 63: 273–277.

49. Raymond F, Caton A, Cox N, Kendal A, Brownlee G (1986) The antigenicity and evolution of influenza H1 haemagglutinin, from 1950–1957 and 1977–1983: Two pathways from one gene. Virology 148: 275–287.

50. Stevens DJ, Douglas AR, Skehel JJ, Wiley DC (1987) Antigenic and amino acid sequence analysis of the variants of H1N1 influenza virus in 1986. Bulletin of the World Health Organization 65: 177–180.

51. Donatelli I, Campitelli L, Ruggieri A, Castrucci MR, Calzoletti L, et al. (1993) Concurrent antigenic analysis of recent epidemic influenza A and B viruses and quantitation of antibodies in population serosurveys in Italy. European Journal of Epidemiology 9: 241–250.

52. Daum LT, Canas LC, Smith CB, Klimov A, Huff W, et al. (2002) Genetic and Antigenic Analysis of the First A/New Caledonia/20/99-like H1N1 Influenza Isolates Reported in the Americas. Emerging Infectious Diseases 8: 408–412.

53. McDonald NJ, Smith CB, Cox NJ (2007) Antigenic drift in the evolution of H1N1 influenza A viruses resulting from deletion of a single amino acid in the haemagglutinin gene. Journal of General Virology 88: 3209–3213.

54. Hay AJ, Lin YP, Gregory V, Bennet M (2003) Annual Report. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

55. Hay AJ, Lin YP, Gregory V, Bennet M (2004) Annual Report. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

56. Hay AJ, Lin YP, Gregory V, Bennet M (Feb 2005) Characteristics of human influenza A H1N1, A H3N2 and B viruses isolated October 2004 to January 2005. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

57. Hay AJ, Lin YP, Gregory V, Bennet M (Sep 2005) Characteristics of human influenza AH1N1, AH3N2 and B viruses isolated February to July 2005. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

58. Hay AJ, Lin YP, Gregory V, Bennet M (March 2006) Characteristics of human influenza AH1N1, AH3N2 and B viruses isolated October 2005 to February 2006. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

59. Hay AJ, Lin YP, Gregory V, Bennet M (Sep 2006) Characteristics of human influenza AH1N1, AH3N2 and B viruses isolated January to September 2006. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

60. Hay AJ, Daniels R, Lin YP, Zheng X, Gregory V, et al. (March 2007) Characteristics of human influenza AH1N1, AH3N2 and B viruses isolated September 2006 to February 2007. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

61. Hay AJ, Daniels R, Lin YP, Zheng X, Gregory V, et al. (Sep 2007) Characteristics of human influenza AH1N1, AH3N2 and B viruses isolated February to August 2007. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

62. Hay AJ, Daniels R, Lin YP, Zheng X, Hou T, et al. (Sep 2008) Characteristics of human influenza AH1N1, AH3N2, and B viruses isolated February to August 2008. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

63. Muyanga J, Matsuzaki Y, Sugawara K, Kimura K, Mizuta K, et al. (2001) Antigenic and genetic analyses of influenza B viruses isolated in Lusaka, Zambia in 1999. Archives of Virology 146: 1667–1679.

64. Shaw MW, Xu X, Li Y, Normand S, Ueki RT, et al. (2002) Reappearance and global spread of variants of influenza B/Victoria/2/87 lineage viruses in the 2000–2001 and 2001–2002 seasons. Virology 303: 1–8.

65. Ansaldi F, Bacilieri S, Amicizia D, Valle L, Banfi F, et al. (2004) Antigenic characterisation of influenza B virus with a new microneutralisation assay: Comparison to haemagglutination and sequence analysis. J Med Virol 74: 141–146.

66. Puzelli S, Frezza F, Fabiani C, Ansaldi F, Campitelli L, et al. (2004) Changes in the hemagglutinins and neuraminidases of human influenza B viruses isolated in Italy during the 2001–02, 2002–03, and 2003–04 seasons. J Med Virol 74: 629–640.

67. Xu X, Lindstrom SE, Shaw MW, Smith CB, Hall HE, et al. (2004) Reassortment and evolution of current human influenza A and B viruses. Virus Res 103: 55–60.

68. Barr I, Komadina N, Durrant C, Sjogren H, Hurt A, et al. (2006) Circulation and antigenic drift in human influenza B viruses in SE Asia and Oceania since 2000. Communicable diseases intelligence 30: 350.

69. Daum LT, Canas LC, Klimov AI, Shaw MW, Gibbons RV, et al. (2006) Molecular analysis of isolates from influenza B outbreaks in the U.S. and Nepal, 2005. Archives of Virology 151: 1863–1874.

70. Lin JH, Chiu SC, Shaw MW, Lin YC, Lee CH, et al. (2007) Characterization of the epidemic influenza B viruses isolated during 2004–2005 season in Taiwan. Virus Res 124: 204–211.

71. Gust I, Barr I, O'Bryan K, Laurie K, Shaw R, et al. (2006) Annual Report. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, Victorian Infectious Disease Research Laboratory, Australia.

72. Hay A, Daniels R, Lin YP, Zheng X, Hou T, et al. (Sep 2009) Antigenic and genetic characteristics of pandemic A(H1N1) viruses and seasonal A(H1N1), A(H3N2) and B viruses isolated during February to August 2009. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

73. McCauley J, Daniels R, Lin YP, Zheng X, Gregory V, et al. (Feb 2011) Report prepared for the WHO annual consultation on the composition of influenza vaccine for the Northern Hemisphere. Technical report, WHO Collaborating Centre for Reference and Research on Influenza, National Institute for Medical Research, UK.

74. Kanegae Y, Sugita S, Endo A, Ishida M, Senya S, et al. (1990) Evolutionary pattern of the hemagglutinin gene of influenza B viruses isolated in Japan: cocirculating lineages in the same epidemic season. J Virol 64: 2860–2865.

75. Nakajima S, Nishikawa F, Nakamura K, Nakajima K (1992) Comparison of the HA genes of type B influenza viruses in herald waves and later epidemic seasons. Epidemiology & Infection 109: 559–568.

76. Nerome R, Hiromoto Y, Sugita S, Tanabe N, Ishida M, et al. (1998) Evolutionary characteristics of influenza B virus since its first isolation in 1940: dynamic circulation of deletion and insertion mechanism. Archives of Virology 143: 1569–1583.

77. Nakagawa N, Nukuzuma S, Haratome S, Go S, Nakagawa T, et al. (2002) Emergence of an influenza B virus with antigenic change. Journal of Clinical Microbiology 40: 3068–3070.

78. Abed Y, Coulthart MB, Li Y, Boivin G (2003) Evolution of surface and nonstructural-1 genes of influenza B viruses isolated in the province of Quebec, Canada, during the 1998–2001 period. Virus Genes 27: 125–135.

79. Ansaldi F, D'Agaro P, de Florentiis D, Puzelli S, Lin YP, et al. (2003) Molecular characterization of influenza B viruses circulating in northern Italy during the 2001–2002 epidemic season. J Med Virol 70: 463–469.

80. Matsuzaki Y, Sugawara K, Takashita E, Muraki Y, Hongo S, et al. (2004) Genetic diversity of influenza B virus: The frequent reassortment and cocirculation of the genetically distinct reassortant viruses in a community. J Med Virol 74: 132–140.

81. Centers for Disease Control and Prevention (2012) FluView Weekly Influenza Reports. Available at http://www.cdc.gov/flu/weekly/.

82. Shapiro B, Rambaut A, Drummond A (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol Biol Evol 23: 7–9.

83. Pybus O, Suchard M, Lemey P, Bernardin F, Rambaut A, et al. (2012) Unifying the spatial epidemiology and molecular evolution of emerging epidemics. Proc Natl Acad Sci USA 109: 15066–15071.