Epidemic dynamics and antigenic evolution in a single season of influenza A

Maciej F. Boni[1,*], Julia R. Gog[2], Viggo Andreasen[3] and Marcus W. Feldman[1]

[1]*Department of Biological Sciences, Stanford University, 371 Serra Mall, Stanford, CA 94305, USA*
[2]*Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK*
[3]*Department of Mathematics and Physics, University of Roskilde, 4000 Roskilde, Denmark*

We use a mathematical model to study the evolution of influenza A during the epidemic dynamics of a single season. Classifying strains by their distance from the epidemic-originating strain, we show that neutral mutation yields a constant rate of antigenic evolution, even in the presence of epidemic dynamics. We introduce host immunity and viral immune escape to construct a non-neutral model. Our population dynamics can then be framed naturally in the context of population genetics, and we show that departure from neutrality is governed by the covariance between a strain's fitness and its distance from the original epidemic strain. We quantify the amount of antigenic evolution that takes place in excess of what is expected under neutrality and find that this excess amount is largest under strong host immunity and long epidemics.

Keywords: influenza; antigenic drift; population dynamics; population genetics; herd immunity; immune escape

1. INTRODUCTION

Seasonal influenza A epidemics are a significant cause of morbidity and mortality in temperate zones of both hemispheres. In the Northern Hemisphere, annual epidemics occur between November and April and since the early 1900s have caused more cumulative mortality than the three major pandemic events of the twentieth century (Earn *et al.* 2002). The characteristics that make these annual influenza outbreaks unusual among non-childhood diseases is that they are periodic and sustained. Periodicity has often been attributed to seasonal changes in transmissibility or mixing patterns (Schulman & Kilbourne 1962; Davey & Reid 1972; Anderson & May 1991), and more recently to a possible dynamical resonance between low intrinsic seasonality and loss of host immune memory (Dushoff *et al.* 2004). Sustainability of annual epidemics is the result of viral immune escape through antigenic drift (Webster *et al.* 1992; Cox & Subbarao 2000; de Jong *et al.* 2000b; Hay *et al.* 2001; Hampson 2002).

Antigenic drift—the accumulation of point mutations in virus antigens—is easily detectable in sequence data covering almost four decades of influenza activity (Macken *et al.* 2001). The resulting immune escape, as measured by haemagglutinin inhibition (HI) tests, indicates that influenza can escape a significant amount of herd immunity after only 2–3 years (de Jong *et al.* 2000b; Coiras *et al.* 2001; Hay *et al.* 2001). Since hosts lose immunity gradually, the influenza virus population need not mutate to a completely new antigenic form. Rather, influenza benefits from each additional amino acid replacement in its surface proteins by

becoming slightly less recognizable to the hosts on whom it previously conferred immunity. Mutations occur during replication in host epithelial cells, and the virus persists and replicates as long as host contacts sustain a chain of transmission in the host population. These chains of transmission enable influenza to accumulate mutations; the resulting mutated progeny viruses are often called antigenic drift variants or simply drift variants.

Here, we consider the forces that govern antigenic drift in influenza A. While it is well known that significant antigenic drift causes severe influenza outbreaks (Kilbourne 1973; Cox & Subbarao 2000; de Jong *et al.* 2000a; Hay *et al.* 2001), little is known about the effects of influenza outbreaks on antigenic drift. These two processes are of course concurrent and tightly coupled. Epidemic dynamics unroll a series of between-host transmission events, which increase viral population size and offer the influenza virus a means to reproduce, mutate, and escape host immunity. Viral immune escape then lowers the host population's effective immunity and adds momentum to the ongoing epidemic. The benefits of immune escape may be somewhat delayed as hosts maintain some short-term non-specific immunity (Ferguson *et al.* 2003; Xia *et al.* 2005).

To model antigenic drift during a single season of influenza, we use a standard susceptible–infected–recovered (SIR) framework (Kermack & McKendrick 1927; Anderson & May 1991). We classify the various influenza strains by their distance from a reference strain. Our model can be described as an $SI_0I_1\ldots I_n$-model, where the subscript denotes antigenic distance from the reference strain I_0. This antigenic distance reflects a one-dimensional antigenic space as in previous models (Pease 1987; Sasaki 1994; Andreasen *et al.* 1996; Haraguchi & Sasaki 1997; Sasaki & Haraguchi 2000; Gog & Grenfell 2002; Andreasen 2003; Lin *et al.* 2003; Boni *et al.* 2004); however, it is important to remember that a realistic

* Author for correspondence (maciek@charles.stanford.edu).

1307

mapping of antigenic type onto immune escape would have higher dimensionality (Lapedes & Farber 2001).

We first introduce a neutral model with epidemic dynamics and mutation and show that the strain population has a Poisson distribution whose mean moves forward in time according to a molecular clock (Zuckerkandl & Pauling 1965; Kimura 1968, 1969). A second model includes host immunity, where strains that escape host immunity through antigenic drift have higher transmissibility. This non-neutral model has high dimensionality and persistent nonlinearities; we solve it numerically. Fortunately, the model's population dynamics can be naturally expressed in a population-genetic framework, which allows us to extract key viral fitness components and analyse their effects on antigenic drift. Using the neutral model as a baseline, we are able to study the forces that drive influenza antigenic drift in human populations.

2. NEUTRAL MODEL

We first consider a many-strain, single-season influenza epidemic model, where all strains are equally fit. Once recovered, hosts cannot become reinfected, and the epidemic ends when the susceptible pool is depleted. The epidemic begins with a particular strain which we call the epidemic strain or the zero-strain; individuals infected with the zero-strain are said to be in the population class I_0. The zero-strain can mutate, and when it acquires one amino acid change the harbouring individual becomes a member of the population class I_1. The I_1 class represents those hosts that are infected with any strain which is exactly one amino acid different from the original epidemic strain; if a host's infecting virus undergoes another mutation event, that host would move into the I_2 class. In general, hosts in the class I_k are infected with a strain that is k amino acids different from the original epidemic strain.

Real individuals are most likely to be infected with a virus population of great diversity, but we can approximate the genetic distance between a host's infecting influenza virions and the original epidemic-causing strain by considering the mean distance of a host's virus population from the original strain. In this model, we neglect within-host evolution and place individuals in population classes according to the strain they are most likely to transmit at a given moment.

We assume a homogeneous mutation rate across the HA1 segment (987 nt) of influenza's haemagglutinin surface protein. We focus on the HA1 because of its rapid evolution and importance to immune escape (Fitch *et al.* 1997; Bush *et al.* 1999; Plotkin & Dushoff 2003). Since back mutation is highly unlikely and recurrent mutations (double hits) are somewhat unlikely, individuals in the I_k class are assumed to move only to the class I_{k+1} when an amino acid replacement occurs.

Using S to denote susceptible hosts, we write the neutral dynamical model as

$$\left.\begin{aligned}
\dot{S} &= -\beta S \sum_{k=0}^{n} I_k, \\
\dot{I}_0 &= \beta S I_0 - (\nu + \mu) I_0, \\
\dot{I}_k &= \beta S I_k - (\nu + \mu) I_k + \mu I_{k-1} \quad \text{for } 0 < k < n, \\
\dot{I}_n &= \beta S I_n - \nu I_n + \mu I_{n-1},
\end{aligned}\right\} \quad (2.1)$$

where β is the compound parameter describing the transmission rate and the host contact rate, and ν is the hosts' recovery rate from infection. The class I_n denotes individuals infected with a strain at least n amino acids away from the zero-strain. The parameter μ is the non-synonymous mutation rate in the HA1. For the purposes of our model, μ is the RNA polymerase's error rate, times the proportion of possible mutations in the HA1 that cause amino acid changes, times the proportion of new mutations that are not lost due to stochastic fluctuations. We will refer to μ simply as the mutation rate. For low population sizes and slow transmission, stochastic extinction of new mutants would need to be modelled explicitly.

To investigate relative strain frequencies, we write $I(t) = \sum_{k=0}^{n} I_k(t)$ and $i_k(t) = I_k(t)/I(t)$. Using (2.1), the dynamic equations for the strain frequencies are

$$\left.\begin{aligned}
\frac{\mathrm{d}i_0}{\mathrm{d}t} &= -\mu i_0, \\
\frac{\mathrm{d}i_k}{\mathrm{d}t} &= -\mu i_k + \mu i_{k-1} \quad \text{for } 0 < k < n, \\
\frac{\mathrm{d}i_n}{\mathrm{d}t} &= +\mu i_{n-1},
\end{aligned}\right\} \quad (2.2)$$

which is a linear and autonomous system, whose dynamics are governed by a subdiagonal matrix and which is independent of the population dynamic variables S and I. If all the strains at time zero are of type zero, $i_0(0) = 1$ and $i_k(0) = 0$ for $k > 0$, the solution to (2.2) is

$$i_k(t) = \frac{(\mu t)^k}{k!} \mathrm{e}^{-(\mu t)}, \quad (2.3)$$

for all $k < n$ and $t > 0$; the trajectory of i_n is determined by noting that $\sum i_k = 1$. The strain frequencies are Poisson-distributed with mean μt. Thus, even in the presence of host population dynamics, we obtain the standard result that neutral mutation produces a molecular clock (we use the term loosely since our clock follows the mean number of changes in a heterogeneous population, rather than the number of fixation events of new variants). If $\mu = 0.1$ per day and the epidemic lasts 120 days, the strain population at the end of the epidemic will be a mean distance 12 amino acids away from the original epidemic strain. Our neutral system has the same assumptions and behaviour as a standard Poisson process.

3. NON-NEUTRAL MODEL

In this section, we remove the neutrality assumption by allowing for immune structure in the host population and viral fitness differences based on host immunity. In our non-neutral model, influenza strains cause weaker infections in immune hosts and are thus less transmissible by these hosts; a strain's fitness (transmissibility) in a particular host depends on the host's immunity to that particular strain. As the virus population mutates, variants that are distant from this season's epidemic strain will be able to cause increasingly transmissible infections, even in hosts whose immunity to earlier variants may have been quite strong. As in the neutral model, hosts cannot become reinfected and the epidemic ends when it runs out of susceptibles.
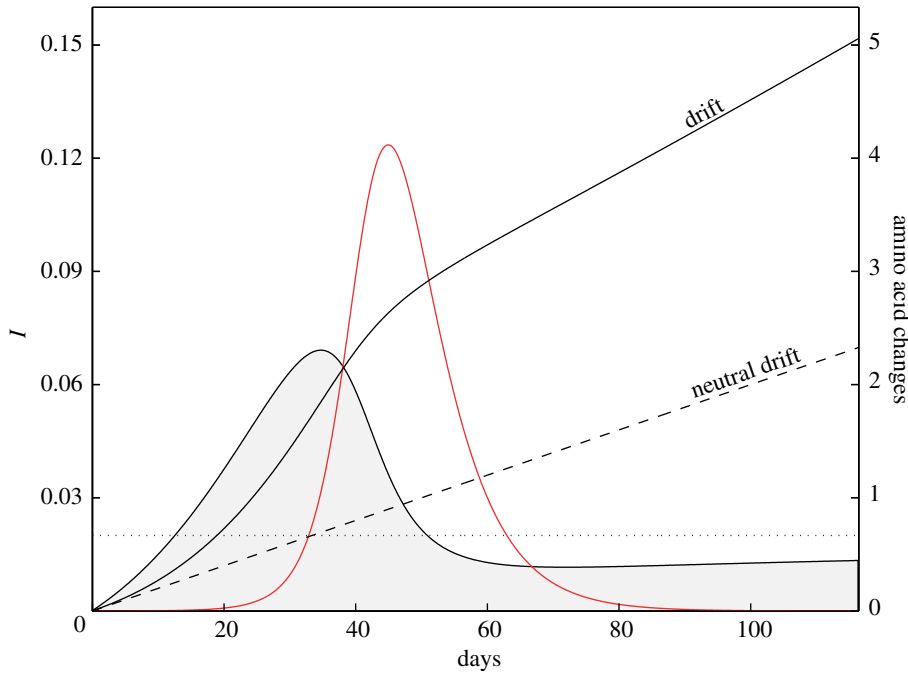
Figure 1. Results from integrating the modified model (3.6)–(3.10). $\beta = 1.0$, $\theta = 0.6$, $a = 0.1$, $N = 10^5$; $\mu = 0.02$ is marked by the dashed line. The red line represents the force of infection I (left-hand scale). The curve bounding the filled area is βQ Cov (left-hand scale) from equation (3.13) and shows here that selection for new variants is more intense in the beginning phases of the epidemic. Comparing the covariance term with the dotted line (μ) shows the relative non-neutral and neutral contributions to total antigenic drift. The two drift lines correspond to values on the right-hand vertical scale. After about 116 days, the epidemic is over and the mean number of amino acid changes is around five; under neutral drift we would expect about two.

We extend the p–q equations from our previous model (Boni *et al.* 2004, p. 180) to include multiple strains; susceptibles are denoted by the variables q_i, where

q_i = frequency of hosts who are susceptible and whose last infection was i amino acids away from this season's zero-strain; $i \in \{0, 1, 2, \dots\}$.

Susceptible hosts $q_0, q_1, q_2, \dots$, have decreasing immunity as the subscript increases. Infected individuals require two subscripts: the current infecting strain and the previous immunizing strain. We define

p_{jk} = frequency of hosts whose last infection was j amino acids away from this season's zero-strain and who are currently infected with strain k; $j \in \{0, 1, 2, \dots\}$, $k \in \{0, 1, \dots, n\}$.

This season's strain k differs by k amino acids from this year's zero-strain, and in a one-dimensional amino acid space, in accordance with our assumptions, individuals in class p_{jk} have a distance of $j + k$ amino acids between their immunizing strain and their current infecting strain. As the distance between challenging strain and immunizing strain increases, immunity decreases (Gill & Murphy 1976; Smith *et al.* 2004). We assume that immunity wanes exponentially with antigenic distance. Individuals in the class p_{jk} have their transmissibility reduced to $1 - \tau_{j+k}$, where $\tau_m = e^{-am}$; the scaling parameter a describes the amount of immune escape conferred by each additional amino acid change. We use the number of amino acid changes as a proxy for immune escape, although their location also plays an important role. An example of this is the 18 strongly selected codons identified by Bush *et al.* (1999), which are known to be associated with antibody-combining sites.

Our dynamical equations for susceptible individuals are:

$$\dot{q}_i = -q_i \left(\beta \sum_{j=0}^{\infty} \sum_{k=0}^{n} (1 - \tau_{j+k}) p_{jk} \right) \tag{3.1}$$

$$\text{for } i = 0, 1, 2, \dots,$$

where the parenthetical term represents the total force of infection in the population. The dynamic equations for the infected individuals are constructed similarly, but we need to take into account the boundary situations $k = 0$ and $k = n$. The equations are:

$$\dot{p}_{j0} = \beta q_j \sum_{i=0}^{\infty} (1 - \tau_i) p_{i0} - (\nu + \mu) p_{j0}, \tag{3.2}$$

$$\dot{p}_{jk} = \beta q_j \sum_{i=0}^{\infty} (1 - \tau_{i+k}) p_{ik} - (\nu + \mu) p_{jk} + \mu c_{jk} p_{j,k-1} \tag{3.3}$$

$$\text{for } 0 < k < n,$$

$$\dot{p}_{jn} = \beta q_j \sum_{i=0}^{\infty} (1 - \tau_{i+n}) p_{in} - \nu p_{jn} + \mu c_{jn} p_{j,n-1}, \tag{3.4}$$

where

$$c_{jk} = \left(\frac{1 - \tau_{j+k-1}}{1 - \tau_{j+k}} \right) \tag{3.5}$$

is a number between 0 and 1. This definition is a fairly close approximation to $c_{jk} = 1$, which would be the natural way to write down the model from first principles. We define c_{jk} as in (3.5) for mathematical convenience; the result of this approximation is a slightly slower rate of antigenic drift. The parameters β, ν and μ are defined as before.

Equations (3.1)–(3.4) now define a complete, infinite-dimensional dynamical system. As in our neutral model,

we collapse some of our variables by defining:

$$S = \sum_{j=0}^{\infty}(1-\tau_j)q_j, \quad Q = \sum_{j=0}^{\infty}q_j, \quad I_k = \sum_{j=0}^{\infty}(1-\tau_{j+k})p_{jk},$$

$$I = \sum_{k=0}^{n}I_k.$$

The variable S denotes the total amount of susceptibility in the population (or the total amount of potential infectivity) and is a number between 0 and 1. Q denotes the total fraction of hosts that are susceptible, irrespective of their immune histories. I_k is the force of infection of strain k, while I is influenza's total force of infection. Q and S obey the dynamical equations:

$$\dot{S} = -\beta SI \quad \text{and} \quad \dot{Q} = -\beta QI,$$

which means that the ratio S/Q does not change with time. We call $\theta = 1 - S/Q$ the immunity in the host population; θ measures the herd immunity to the zero-strain of the susceptible individuals in the host population. Alternatively, θ can be viewed as the expected immunity of any susceptible individual in the population.

As before, $i_k = I_k/I$ is the frequency of strain k, and equations (3.1)–(3.4) reduce to:

$$\frac{\mathrm{d}i_0}{\mathrm{d}t} = \beta Q i_0\left((1-\theta) - \sum_{l=0}^{n}(1-\theta\tau_l)i_l\right) - \mu i_0, \quad (3.6)$$

$$\frac{\mathrm{d}i_k}{\mathrm{d}t} = \beta Q i_k\left((1-\theta\tau_k) - \sum_{l=0}^{n}(1-\theta\tau_l)i_l\right) - \mu i_k + \mu i_{k-1} \quad (3.7)$$

$$\text{for } 0 < k < n,$$

$$\frac{\mathrm{d}i_n}{\mathrm{d}t} = \beta Q i_n\left((1-\theta\tau_n) - \sum_{l=0}^{n}(1-\theta\tau_l)i_l\right) + \mu i_{n-1}. \quad (3.8)$$

With the dynamical equations for Q and I,

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -\beta QI, \quad (3.9)$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta QI \sum_{k=0}^{n}(1-\theta\tau_k)i_k - \nu I. \quad (3.10)$$

Equations (3.6)–(3.10) now describe an $(n+2)$-dimensional dynamical system, which keeps track of the strain frequencies, the total force of infection and the number of susceptibles.

(a) Population genetics

The quantity $w_k = 1 - \theta\tau_k$ has the natural population-genetic interpretation as the fitness of strain k, and

$$W = \sum_{k=0}^{n}(1-\theta\tau_k)i_k \quad \text{where } 0 \le W \le 1 \quad (3.11)$$

is thus the mean fitness of the entire virus population. The dynamical equations (3.6)–(3.10) then resemble standard population-genetic equations where the key determinant of a variant's increase or decrease in frequency is its fitness (w_k) relative to the population's mean fitness (W).

To see how mean fitness behaves as a function of time, we differentiate equation (3.11) and approximating $\tau_n i_n \approx 0$,

$$\dot{W} = \beta Q \,\mathrm{Var}(w_k) + \mu(1-\tau_1)(1-W),$$

which, if we set $\mu = 0$, is the continuous analogue to Fisher's fundamental theorem of natural selection (Fisher 1930).

To investigate the dynamic properties of antigenic drift, we define:

$$D = \sum_{k=0}^{n}ki_k, \quad (3.12)$$

which is the mean antigenic distance from the strain population at time t to the zero-strain. The quantity D follows the dynamics of system (3.6)–(3.10). Approximating $i_n \approx 0$, we have:

$$\dot{D} = \mu + \beta Q\left[\sum kw_k i_k - \sum w_k i_k \sum ki_k\right]$$

$$= \mu + \beta Q \,\mathrm{Cov(distance, fitness)}, \quad (3.13)$$

which is a form of the Price equation (Price 1970, 1972). Distance refers to the number of amino acid replacements a strain is away from the original invading strain, and fitness refers to $w_k = 1 - \theta\tau_k$. Since strain fitness always increases with added distance from the zero-strain, the covariance term—which in the above equation is calculated across the strain frequencies at time t—will always be non-negative. At $t = 0$, the covariance is zero, and its derivative with respect to time is $\beta\mu\theta(1 - \mathrm{e}^{-a})$. At the beginning of the epidemic, when there are strains of high and low fitness, the covariance term will be positive and increasing; then, as Q decreases and as antigenic drift causes most of the strains in the virus population to have a fitness close to one, the covariance term will tend back towards zero. When there are no fitness differences among strains (neutral mutation), the covariance term is always zero and

$$\dot{D} = \mu, \quad (3.14)$$

which can also be derived from (2.3).

Once all the population-genetic structure is extracted from our population-dynamic influenza model, the intensity of selection for antigenically distant strains can be measured by the size of the covariance term in (3.13) relative to the mutation rate μ. The mutation rate μ is responsible for neutral mutation accumulation and sets the baseline pace of the molecular clock, while the covariance term changes throughout the epidemic and accelerates the clock to varying degrees (see figure 1).

(b) Excess antigenic drift

The amount of antigenic drift that occurs during one season is highly dependent on μ^{-1}—the mean number of days it takes for a neutrally mutating flu population to acquire, on the average, one additional amino acid change. Models of within-host flu evolution (Sasaki 1994; Haraguchi & Sasaki 1997) have calculated a drift speed that scales with μ, while some between-host models (Andreasen et al. 1996; Gog & Grenfell 2002; Lin et al. 2003) have found a drift speed that scales approximately with $\sqrt{\mu}$. In this investigation, we focus on the excess antigenic drift, δ, which we define as the difference between the amount of drift occurring under neutral conditions and the amount of drift occurring under non-neutral conditions. We show that δ is relatively insensitive to the mutation rate μ; this means that we can study the factors that affect δ without knowing the true mutation rate.

Table 1. Parameter descriptions and probable ranges for the non-neutral model.

parameter	description	range
a	immune-escape parameter; when a is large, immune escape is rapid (cross-immunity is weak)	$0.01 \leq a \leq 0.50$
β	transmissibility and host contact rate; assumes $\nu = 0.2$, hence $1.2 \leq \mathcal{R}_0 \leq 6.0$	$0.24 \leq \beta \leq 1.20$
μ	non-synonymous mutation rate in the HA1: error rate in the RNA polymerase, times proportion of possible non-silent changes, times proportion of mutants that survive stochastic loss	$0.001 \leq \mu \leq 0.05$
ν	recovery rate from infection; this is set to $\nu = 0.2$, so that an infection lasts 5 days; fixing this parameter has no qualitative effect on model results; realistic range shown at right	$0.1 \leq \nu \leq 0.3$
θ	population-wide immunity or herd immunity; $\theta = 1$ means that the population has full immunity (i.e. everyone is completely immune); $\theta = 0$ means that every host is completely naive	$0 \leq \theta \leq 1$
N^{-1}	inoculum; the initial force of infection that begins the epidemic; the parameter N can be thought of as the host population size	$10^{-9} \leq N^{-1} \leq 10^{-3}$

Let D_S be the amount of antigenic drift that occurs when there is selection for immune escape, and let D_N be the amount of neutral antigenic drift that would be expected to occur. We calculate D_S by numerically integrating equations (3.6)–(3.10) with some initial condition, or inoculum, $I(0)$ that gives the force of infection at time $t = 0$. We set $I(0) = N^{-1}$ and use N as a proxy for population size. Equations (3.6)–(3.10) are numerically integrated from time $t = 0$ until a time t_f such that $I(t_f) = I(0)$; we say that the epidemic ends at time t_f. Using definition (3.12), we let $D_S = D(t_f)$; this is the mean amount of antigenic drift that occurs when selective pressure causes the virus to mutate away from the epidemic strain. To calculate D_N for an epidemic of the same length, we use equation (3.14) and get $D_N = \mu t_f$.

It is then natural to define δ as:

$$\delta = D_S - D_N = \beta \int_0^{t_f} Q(t) \mathrm{Cov}(k, w_k) \mathrm{d}t. \qquad (3.15)$$

The covariance term under the integral is the term from the Price equation (3.13), and t_f is the time at which the epidemic ends. Since the integrand is always non-negative, we see that a longer epidemic results in a larger δ, since it allows selection more time to operate; this phenomenon has been analysed in the context of pathogen emergence by Antia *et al.* (2003). At first glance, it seems that increasing transmissibility and host contact rates (via β) should yield more excess drift; however, higher β-values correspond to shorter epidemics which can in turn yield less excess drift. We will characterize the behaviour of δ as we change the model parameters described in table 1.

(c) Parameter ranges

The high dimensionality of our system forces us to study it numerically. The dynamical system (3.6)–(3.10) has five parameters (a, β, μ, ν and θ), although ν can be scaled out if we wish; we simply set $\nu = 0.2$, fixing the mean infection length at 5 days. The initial condition $I(0) = N^{-1}$ must be set, and we consider it a model parameter. The number of strains n in all simulations is 60. The parameters β, θ and N can be reasonably varied to simulate strains with basic reproduction ratios in the range $1 < \mathcal{R}_0 \leq 6$ (Mills *et al.* 2004), and populations of various sizes with varying levels of herd immunity. The tested ranges for these and other parameters are summarized in table 1.

The parameters a and μ are more difficult to measure and can vary over a wide range of values. The mutation rate in influenza's haemagglutinin has been measured by Fitch *et al.* (1997) and Bush *et al.* (1999) who estimated the observed, rather than neutral, rate of evolution. Moreover, depending on whether one calculates distance from a root strain or mean distance between pairs of strains isolated in consecutive years, estimates of mutation rates can vary by an order of magnitude. Worldwide (Macken *et al.* 2001) and local (Coiras *et al.* 2001; Pyhälä *et al.* 2004) HA1 datasets suggest that the observed mutation rate corresponds to between 1 and 13 amino acid changes per year; the neutral rate can of course be lower. In our numerical simulations, we test the range of μ-values $0.001 \leq \mu \leq 0.050$, which corresponds to between 0.4 and 18 non-synonymous mutations per year. Since δ is not highly sensitive to μ, the choice of range for the neutral mutation rate has little effect on our results.

Finally, a measures immune escape per amino acid change. The range $0.03 \leq a \leq 0.15$ entails that it takes between 5 and 20 amino acid replacements to evade 50% host immunity. This seems reasonable based on published HI tables (de Jong *et al.* 2000b; Coiras *et al.* 2001; Hay *et al.* 2001) and the antigenic map in Smith *et al.* (2004). The tested range for a will be slightly wider: $0.01 \leq a \leq 0.50$. Note that parameter estimates of a and μ are a function of the length of the HA1 molecule (987 nt).

4. RESULTS

According to our model, the keys to generating a large amount of excess antigenic drift are strong herd immunity and long epidemics. Host immunity forces the virus population to mutate to a distant variant so that it can begin spreading efficiently. Figure 2 shows immunity driving antigenic drift within the context of epidemic dynamics, while figure 3 shows excess antigenic drift (δ) increasing as a function of immunity (θ). A slow (and thus long) epidemic allows selection pressure to operate for a longer period of time and allows the virus population to drift further than under a short epidemic. The two key characteristics of a host–parasite system that can lengthen an epidemic are large host population size and low $\mathcal{R}_0$. In our model, if N is large or if our effective $\mathcal{R}_0 = \beta(1 - \theta)/\nu$ is close to 1, the epidemic will be long and excess drift will be large.
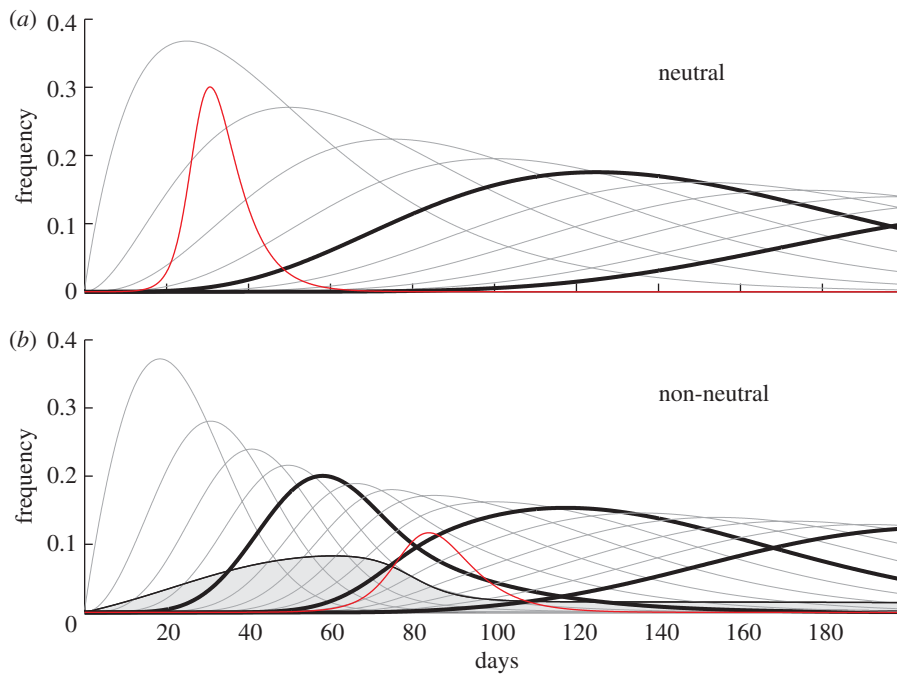
Figure 2. The solid black lines represent the frequencies i_1, i_2, i_3 and so on (i_0 is not shown). Every fifth variant has a bold line so that (a) and (b) can be compared. In (a), i_5 and i_{10} are bold; in (b), i_5, i_{10} and i_{15} are bold. In (a) and (b), $\beta = 0.6$, $\nu = 0.2$, $\mu = 0.04$, $a = 0.1$ and $I(0) = 10^{-5}$. In (a), $\theta = 0.0$, $\mathcal{R}_0 = 3$. In (b), $\theta = 0.6$, and the effective $\mathcal{R}_0 = 1.2$. The red line represents the force of infection I. In (b), the solid curve bounding the filled area corresponds to the term βQ Cov from the Price equation (3.13).
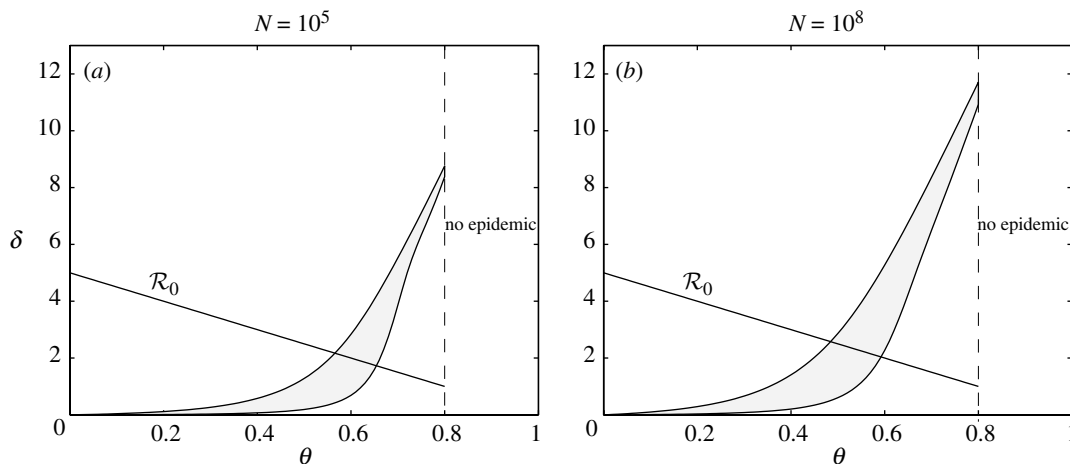

Figure 3. Excess drift as a function of herd immunity. In (a) and (b), the two curves bounding the filled area show the value of δ for $\mu = 0.05$ (upper curve) and $\mu = 0.005$ (lower curve). In (a) and (b), $\beta = 1.0$, $\nu = 0.2$ and $a = 0.05$. The negatively sloped straight line in both panels shows $\mathcal{R}_0$ as a function of θ. When $\theta > 0.8$, the effective basic reproduction ratio is less than 1 and there is no epidemic.

Thus, the parameters β, θ and N have intuitive effects on δ. Decreasing β or increasing θ lowers $\mathcal{R}_0$, lengthens the epidemic, and increases the amount of excess antigenic drift. In addition, increasing θ drives antigenic drift by augmenting the strength of selection for escape mutants. An increase in the population size N decreases the relative size of the inoculum, lengthens the epidemic, and leads to more excess antigenic drift (higher δ).

For the range of a-values that we test, lowering a decreases the amount of excess drift during the course of an epidemic. This happens because when a is small enough, the initial populations of strains that are 1, 2 or 3 amino acids away from the zero-strain are not much more fit than the zero-strain, and natural selection has little fitness variation on which to act. In the case of very small a, antigenic drift is close to neutral. On the other hand, if a

is very large, δ will also be small since fit variants are achieved with few mutations and the epidemics are generally short. Figure 4 shows this behaviour of our model as a function of a; in general, intermediate values of a maximize δ.

Similarly, intermediate values of μ appear to maximize δ. Again, low μ yields little variation, and thus slow natural selection. High μ yields lots of variation and much antigenic drift, but most of this antigenic drift can be explained by the fast mutation rate rather than selective pressure on the virus population—total antigenic drift is high, but excess drift is low.

In general, the mutation rate μ and immune-escape parameter a have relatively little effect on the excess drift δ. Over a set of 83 994 runs using distinct parameter combinations, δ exhibited partial correlations of 0.12
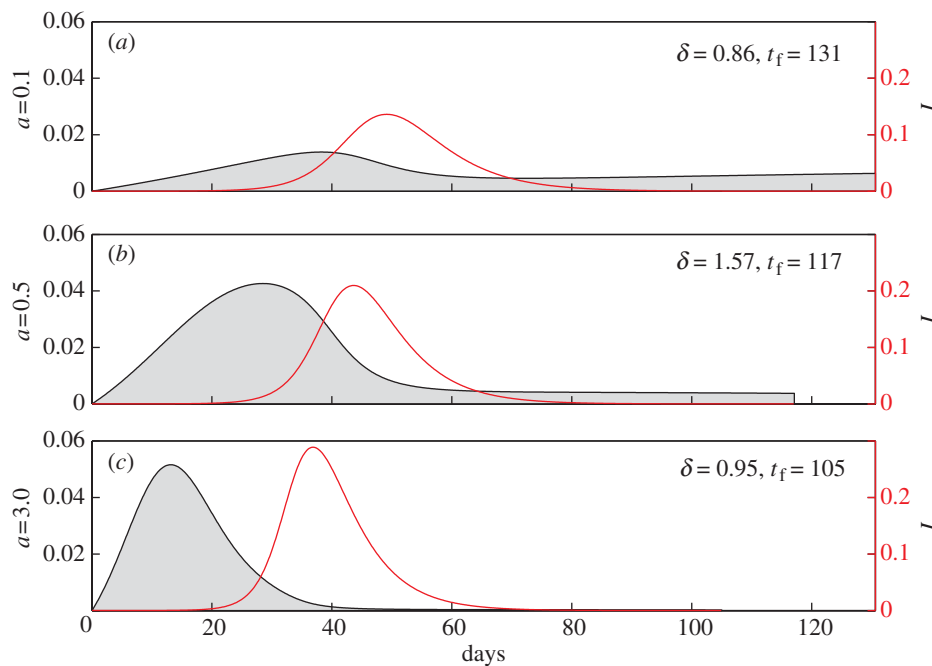
Figure 4. In (a)–(c), $\beta = 0.6$, $\nu = 0.2$, $\theta = 0.3$ (effective $\mathcal{R}_0 = 2.1$ for the zero-strain), $\mu = 0.02$ and $I(0) = 10^{-5}$; $a = 0.1$, 0.5, 3.0 in (a), (b) and (c), respectively. The red line represents the force of infection I and corresponds to the red axis on the right-hand side of the graph. The solid curve bounding the filled area corresponds to the term βQ Cov from the Price equation (3.13). The shaded area under this curve is the value of the integral of βQ Cov from 0 to t_f; this is the excess antigenic drift. Graph (b) maximizes the area under the covariance curve. The value $a = 3.0$ is used for illustrative effect only.

Table 2. Partial correlations between a parameter and a dynamic quantity, when the other four model parameters are held constant. The last column is the partial correlation between δ and a parameter when the other four parameters as well as the length of the epidemic are held constant. For each parameter, the first row uses all simulations (83 994 in all) where the epidemic length was less than 1000 days. The second row, removes some outliers and uses those simulations (71 227) whose epidemics were shorter than 250 days. Note that accounting for the epidemic length does not explain away a correlation between θ and δ. *Abbreviations*: es: epidemic size, $1 - Q(t_f)$; wes: weighted epidemic size, $1 - \theta - S(t_f)$; len: epidemic length, t_f; D: total antigenic drift.

	es	wes	len	D	δ	$\delta^{(\mathrm{len})}$
β	0.76	0.76	-0.54	-0.44	-0.32	-0.06
	0.70	0.76	-0.79	-0.48	-0.34	-0.02
θ	-0.39	-0.94	0.25	0.49	0.59	0.57
	-0.45	-0.98	0.65	0.73	0.78	0.63
a	0.37	0.25	-0.20	-0.12	-0.05	0.09
	0.37	0.31	-0.33	0.04	0.18	0.26
μ	0.15	0.11	-0.09	0.56	0.12	0.21
	0.15	0.13	-0.13	0.84	0.33	0.36
$\log N$	0.17	0.11	0.42	0.35	0.25	-0.02
	0.33	0.29	0.85	0.55	0.39	0.00

with μ and -0.05 with a; this suggests low sensitivity of δ to μ and a. There is no way to test for statistical significance since the correlated quantities are the results of deterministic simulations (see table 2 and electronic supplementary material, Appendix A). Also, δ–μ sensitivity is dependent on the choice of cross-immunity function τ. Alternate functional forms of τ can produce a noticeable sensitivity of δ to μ (electronic supplementary material, Appendix B).

From these explorations of the effects of the model parameters on excess drift, we note two curious behaviours of our single-season model of influenza evolution.

First, the epidemic usually peaks when much of the drift or excess drift has already happened. Therefore, sampling isolates during what we believe to be the beginning of the epidemic may lead us to overestimate the amount of drift that happened between seasons, when in fact, the observed drift may have happened early in the current season. In figure 2a, the neutral epidemic peaks after 30 days, having undergone about 1.2 replacements. In the non-neutral dynamics in figure 2b, the epidemic peaks after 84 days having undergone 7.9 replacements, 4.6 of which are in excess of what can be explained by neutral mutation during that time period. Sampling during the beginning of the non-neutral epidemic (e.g. between days 60 and 70) may lead to an incorrect conclusion about that year's epidemic strain.

This result may help explain a phenomenon described by Schweiger *et al.* (2002), namely, that 'comparable major antigenic differences may result in a severe

outbreak—not necessarily during the first epidemic season [of] their appearance, but during the second.' A slow and mild epidemic can be accompanied by a lot of excess drift in its early phases; in such an epidemic, distant variants may be observed in collected flu isolates. If a distant variant at the end of a mild season starts the epidemic at the beginning of next season, it will benefit by having escaped much of the host population's immunity and may be able to cause a large epidemic. This pattern was observed in Germany during the mild 1997/1998 season and the more severe 1998/1999 season. In general, amino acid changes accumulate within an epidemic season, but short-term non-specific host immunity may prevent their effects from being felt until the following season.

Second, we note that the total size of the epidemic, $1 - Q(t_f)$, measured as the total fraction of hosts infected, as well as the weighted size of the epidemic, $1 - \theta - S(t_f)$, do not always correlate positively with the excess drift δ (partial correlations are -0.15 and -0.06, respectively). Large epidemics do not always result in a lot of antigenic drift in part because larger epidemic sizes correlate negatively (-0.61 and -0.62, respectively) with epidemic length, and epidemic length correlates positively ($+0.60$) with δ. This suggests that a scenario of annual epidemics with runaway antigenic drift (Boni *et al.* 2004) would have to be revisited under the assumption that long epidemics, rather than large epidemics, yield a lot of drift. In such a scenario, the strain distribution at the end of one epidemic and the 'choice' of a particular strain to start next season's epidemic may be critical.

5. DISCUSSION

We analysed a neutral and a non-neutral model of influenza spread and evolution in a single epidemic season in order to investigate the forces that drive antigenic drift in influenza. We solved the neutral model analytically, which provided a basis for comparison of the numerical results of the non-neutral model. In the non-neutral model, we examined the conditions that cause the most excess antigenic drift, which we defined as the drift that occurs beyond that expected under neutral mutation. We found that strong host immunity and long epidemics result in greater excess antigenic drift, that significant amounts of antigenic drift can occur in the early phases of the epidemic when there are still relatively few infected hosts, and that large epidemics tend to be short, generating little excess drift.

We used a standard deterministic SIR formulation with multiple strains; our model had no host births and no immigration so that the epidemic ended when the virus ran out of susceptibles. This restricts our results to closed panmictic populations. Antigenic drift on a global scale would require a meta-population model, which describes human populations exposed to influenza. In particular, the stochastic nature of (i) migration between sub-populations, (ii) summertime transmission dynamics of influenza in temperate zones (Gog *et al.* 2003) and (iii) local extinction of epidemics and new mutants when infected numbers are small (Girvan *et al.* 2002; Park *et al.* 2002) would all need to be better understood.

Our non-neutral model (3.6)–(3.10) fits elegantly into Price's covariance formulation of natural selection. Using Price's population-genetic framework, we can track changes in the mean antigenic drift in the virus population via the covariance between mutations accumulated (k) and immunity escaped (w_k). This covariance term increases initially and then wanes, reflecting the differential selection pressure on the influenza virus population during the course of an epidemic. Price's formulation describes the forces that govern the progression of the virus population's mean antigenic distance from the original epidemic-causing strain, but it tells us nothing about other properties of the strain distribution. An important and open problem is the characterization of the differences between the observed (non-neutral) strain distribution at the end of an epidemic and the neutral Poisson distribution.

We identified the parameters in our system that caused the greatest departures from neutrality. It appears that immunity (θ) has the largest effect on excess antigenic drift (δ). A high level of host immunity puts significant selection pressure on the virus population, and in addition, it slows the epidemic, giving natural selection more time to select for distant antigenic variants. The relationship between θ and δ may have important public health consequences as it indicates that vaccinated populations, as long as they can still sustain epidemics, can cause significant antigenic drift (as suggested by Pease 1987, p. 445). Public health officials may wish to investigate whether the benefits of vaccination during one season conflict with the feasibility of vaccination for the following season. If antigenic drift is indeed greater in more immune populations, preparedness for influenza pandemics (Webby & Webster 2003) may need to include vaccination strategies for the second year after a pandemic with consideration to the effect this will have on the third year after a pandemic.

The δ–θ relationship has further importance due to the discontinuity that appears at $\mathcal{R}_0 = 1$ in figure 3. The stochastic nature of mutation, transmission, vaccination efficacy and population interactions may cause our system to fall on either side of this discontinuity, either yielding an unexpected amount of antigenic drift (to the left of the vertical dashed line) or preventing an epidemic entirely (to the right of the vertical dashed line). The consequences of this particular threshold property will need to be explored with a stochastic model.

With a wealth of sequence data and a high mutation rate, influenza virus ecology and evolution have a broad and important intersection with the growing field of measurably evolving populations (Drummond *et al.* 2003). Techniques for the accurate estimation of mutation rates could be applied to detailed, localized influenza datasets such as the one described by Schweiger *et al.* (2002). A precise estimate of influenza's mutation rate would be a significant step towards accurate predictions of near-term antigenic drift. Similarly, the effects of local population structure during influenza epidemics could be measured with a technique based on allelic mismatch distributions as developed by Fraser *et al.* (2005); this type of study may help determine whether the observed strain distributions result more from host population immunity or host population structure. These methods, along with the techniques presented in this paper, will help quantify the driving forces behind antigenic drift in influenza A.

REFERENCES

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans: dynamics and control.* Oxford, UK: Oxford Science Publications.

Andreasen, V. 2003 Dynamics of annual influenza A epidemics with immuno-selection. *J. Math. Biol.* **46**, 504–536. (doi:10.1007/s00285-002-0186-2)

Andreasen, V., Levin, S. & Lin, J. 1996 A model of influenza A drift evolution. *Z. Angew. Math. Mech.* **76**, 421–424.

Antia, R., Regoes, R. R., Koella, J. C. & Bergstrom, C. T. 2003 The role of evolution in the emergence of infectious diseases. *Nature* **426**, 658–661. (doi:10.1038/nature02104)

Boni, M. F., Gog, J. R., Andreasen, V. & Christiansen, F. B. 2004 Influenza drift and epidemic size: the race between generating and escaping immunity. *Theor. Popul. Biol.* **65**, 179–191. (doi:10.1016/j.tpb.2003.10.002)

Bush, R. M., Fitch, W. M., Bender, C. A. & Cox, N. J. 1999 Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Mol. Biol. Evol.* **16**, 1457–1465.

Coiras, M., Aguilar, J., Galiano, M., Carlos, S., Gregory, V., Lin, Y., Hay, A. & Pérez-Breña, P. 2001 Rapid molecular analysis of the haemagglutinin gene of human influenza A H3N2 viruses isolated in Spain from 1996 to 2000. *Arch. Virol.* **146**, 2133–2147. (doi:10.1007/s007050170025)

Cox, N. J. & Subbarao, K. 2000 Global epidemiology of influenza: past and present. *Annu. Rev. Med.* **51**, 407–421. (doi:10.1146/annurev.med.51.1.407)

Davey, M. L. & Reid, D. 1972 Relationship of air temperature to outbreaks of influenza. *Br. J. Prev. Soc. Med.* **26**, 28–32.

de Jong, J. C., Beyer, W. E. P., Palache, A. M., Rimmelzwaan, G. F. & Osterhaus, A. D. M. E. 2000a Mismatch between the 1997/1998 influenza vaccine and the major epidemic A(H3N2) virus strain as the cause of an inadequate vaccine-induced antibody response to this strain in the elderly. *J. Med. Virol.* **61**, 94–99. (doi:10.1002/(SICI)1096-9071(200005)61:1<94::AID-JMV15>3.0.CO;2-C)

de Jong, J. C., Rimmelzwaan, G. F., Fouchier, R. A. M. & Osterhaus, A. D. M. E. 2000b Influenza virus: a master of metamorphosis. *J. Infect.* **40**, 218–228. (doi:10.1053/jinf.2000.0652)

Drummond, A. J., Pybus, O. G., Rambaut, A., Forsberg, R. & Rodrigo, A. G. 2003 Measurably evolving populations. *Trends Ecol. Evol.* **18**, 481–488. (doi:10.1016/S0169-5347(03)00216-7)

Dushoff, J., Plotkin, J. B., Levin, S. A. & Earn, D. J. D. 2004 Dynamical resonance can account for seasonality of influenza epidemics. *Proc. Natl Acad. Sci. USA* **101**, 16 915–16 916. (doi:10.1073/pnas.0407293101)

Earn, D. J. D., Dushoff, J. & Levin, S. A. 2002 Ecology and evolution of the flu. *Trends Ecol. Evol.* **17**, 334–340. (doi:10.1016/S0169-5347(02)02502-8)

Ferguson, N. M., Galvani, A. P. & Bush, R. M. 2003 Ecological and immunological determinants of influenza evolution. *Nature* **422**, 428–433. (doi:10.1038/nature01509)

Fisher, R. A. 1930 *The genetical theory of natural selection.* Oxford, UK: Clarendon Press.

Fitch, W. M., Bush, R. M., Bender, C. A. & Cox, N. J. 1997 Long term trends in the evolution of H(3) HA1 human influenza type A. *Proc. Natl Acad. Sci. USA* **94**, 7712–7718. (doi:10.1073/pnas.94.15.7712)

Fraser, C., Hanage, W. P. & Spratt, B. G. 2005 Neutral microepidemic evolution of bacterial pathogens. *Proc. Natl Acad. Sci. USA* **102**, 1968–1973. (doi:10.1073/pnas.0406993102)

Gill, P. W. & Murphy, A. M. 1976 Naturally acquired immunity to influenza type A: a clinical and laboratory study. *Med. J. Aust.* **2**, 329–333.

Girvan, M., Callaway, D. S., Newman, M. E. J. & Strogatz, S. H. 2002 A simple model of epidemics with pathogen mutation. *Phys. Rev. E* **65**, 031 915. (doi:10.1103/PhysRevE.65.031915)

Gog, J. R. & Grenfell, B. T. 2002 Dynamics and selection of many-strain pathogens. *Proc. Natl Acad. Sci. USA* **99**, 17 209–17 214. (doi:10.1073/pnas.252512799)

Gog, J. R., Rimmelzwaan, G. F., Osterhaus, A. D. M. E. & Grenfell, B. T. 2003 Population dynamics of rapid fixation in cytotoxic T lymphocye escape mutants of influenza A. *Proc. Natl Acad. Sci. USA* **100**, 11 143–11 147. (doi:10.1073/pnas.1830296100)

Hampson, A. W. 2002 Influenza virus antigens and 'antigenic drift'. In *Influenza* (ed. C. W. Potter), pp. 49–86. Amsterdam, The Netherlands: Elsevier Science B.V.

Haraguchi, Y. & Sasaki, A. 1997 Evolutionary pattern of intra-host pathogen antigenic drift: effect of cross-reactivity in immune response. *Phil. Trans. R. Soc. B* **352**, 11–20. (doi:10.1098/rstb.1997.0002)

Hay, A. J., Gregory, V., Douglas, A. R. & Lin, Y. P. 2001 The evolution of human influenza viruses. *Phil. Trans. R. Soc. B* **356**, 1861–1870. (doi:10.1098/rstb.2001.0999)

Kermack, W. O. & McKendrick, A. G. 1927 Contributions to the mathematical theory of epidemics 1. *Proc. R. Soc. A* **115**, 700–721.

Kilbourne, E. D. 1973 The molecular epidemiology of influenza. *J. Infect. Dis.* **127**, 478–487.

Kimura, M. 1968 Evolutionary rate at the molecular level. *Nature* **217**, 624–626.

Kimura, M. 1969 The rate of molecular evolution considered from the standpoint of population genetics. *Proc. Natl Acad. Sci. USA* **63**, 1181–1188.

Lapedes, A. & Farber, R. 2001 The geometry of shape space: application to influenza. *J. Theor. Biol.* **212**, 57–69. (doi:10.1006/jtbi.2001.2347)

Lin, J., Andreasen, V., Casagrandi, R. & Levin, S. A. 2003 Traveling waves in a model of influenza A drift. *J. Theor. Biol.* **222**, 437–445.

Macken, C., Lu, H., Goodman, J. & Boykin, L. 2001 The value of a database in surveillance and vaccine selection. In *Options for the control of influenza IV* (ed. A. D. M. E. Osterhaus, N. Cox & A. W. Hampson), pp. 103–106. Amsterdam: Elsevier Science.

Mills, C. E., Robins, J. M. & Lipsitch, M. 2004 Transmissibility of the 1918 pandemic influenza. *Nature* **432**, 904–906. (doi:10.1038/nature03063)

Park, A. W., Gubbins, S. & Gilligan, C. A. 2002 Extinction times for closed epidemics: the effects of host spatial structure. *Ecol. Lett.* **5**, 747–755. (doi:10.1046/j.1461-0248.2002.00378.x)

Pease, C. M. 1987 An evolutionary epidemiological mechanism, with applications to type A influenza. *Theor. Popul. Biol.* **31**, 422–452. (doi:10.1016/0040-5809(87)90014-1)

Plotkin, J. B. & Dushoff, J. 2003 Codon bias and frequency-dependent selection on the hemagglutinin epitopes of the influenza A virus. *Proc. Natl Acad. Sci. USA* **100**, 7152–7157. (doi:10.1073/pnas.1132114100)

Price, G. R. 1970 Selection and covariance. *Nature* **227**, 520–521. (doi:10.1038/227520a0)

Price, G. R. 1972 Extension of covariance selection mathematics. *Ann. Hum. Genet. Lond.* **35**, 485–490.

Pyhälä, R., Visakorpi, R., Ikonen, N. & Kleemola, M. 2004 Influence of antigenic drift on the intensity of influenza

outbreaks: upper respiratory tract infections of military conscripts in Finland. *J. Med. Virol.* **72**, 275–280. (doi:10.1002/jmv.10552)

Sasaki, A. 1994 Evolution of antigenic drift/switching: continuously evading pathogens. *J. Theor. Biol.* **168**, 291–308. (doi:10.1006/jtbi.1994.1110)

Sasaki, A. & Haraguchi, Y. 2000 Antigenic drift of viruses within a host: a finite site model with demographic stochasticity. *J. Mol. Evol.* **51**, 245–255.

Schulman, J. L. & Kilbourne, E. D. 1962 Airborne transmission of influenza virus infection in mice. *Nature* **195**, 1129–1130.

Schweiger, B., Zadow, I. & Heckler, R. 2002 Antigenic drift and variability of influenza viruses. *Med. Microbiol. Immunol.* **191**, 133–138. (doi:10.1007/s00430-002-0132-3)

Smith, D. J., Lapedes, A. S., de Jong, J. C., Bestebroer, T. M., Rimmelzwaan, G. F., Osterhaus, A. D. M. E. & Fouchier, R. A. M. 2004 Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 371–376. (doi:10.1126/science.1097211)

Webby, R. J. & Webster, R. G. 2003 Are we ready for pandemic influenza? *Science* **302**, 1519–1522. (doi:10.1126/science.1090350)

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M. & Kawaoka, Y. 1992 Evolution and ecology of influenza A viruses. *Microbiol. Rev.* **56**, 152–179.

Xia, Y., Gog, J. R. & Grenfell, B. T. 2005 Semiparametric estimation of the duration of immunity from infectious disease time series: influenza as a case-study. *J. R. Stat. Soc. C* **54**, 659–672. (doi:10.1111/j.1467-9876.2005.05383.x)

Zuckerkandl, E. & Pauling, L. 1965 Evolutionary divergence and convergence in proteins. In *Evolving genes and proteins* (ed. V. Bryson & H. J. Vogel), pp. 97–166. New York: Academic Press.