



International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report

IR-10-045

Coexistence conditions for strains of influenza with immune cross-reaction

Ryosuke Omori (omori@bio-math10.biology.kyushu-u.ac.jp)
Ben Adams (b.adams@bath.ac.uk)
Akira Sasaki (sasaki_akira@soken.ac.jp)

Approved by

Ulf Dieckmann
Program Leader, EEP

July 2011

Contents

Abstract.....	1
Introduction	2
Deterministic model	3
Results	8
Individual-based simulation	10
Discussion.....	11
Acknowledgements	13
References	13
Figure legends	16

Coexistence conditions for strains of influenza with immune cross-reaction

Ryosuke Omori^{1*}, Ben Adams^{1,2} and Akira Sasaki^{3,4,5}

¹ Department of Biology, Faculty of Science, Kyushu University,
Fukuoka 812-8581, Japan
phone: +81-92-642-2641
fax: +81-92-642-2645

² Department of Mathematics, University of Bath, Bath, BA2 7AY, UK

³ Department of Evolutionary Studies of Biosystems (Sokendai-Hayama), The Graduate
University for Advanced Studies (Sokendai), Hayama, Kanagawa 240-0193, Japan

⁴ Evolution and Ecology Program, International Institute for Applied Systems Analysis,
A-2361 Laxenburg, Austria

⁵ PRESTO, Japan Science and Technology Agency,
4-1-8 Honcho Kawaguchi, Saitama, Japan

* Corresponding author, email: omori@bio-math10.biology.kyushu-u.ac.jp

Abstract

The accumulation of cross-immunity in the host population is an important factor driving the antigenic evolution of viruses such as influenza A. Mathematical models have shown that the strength of temporary non-specific cross-immunity and the basic reproductive number are both key determinants for evolutionary branching of the antigenic phenotype. Here we develop deterministic and stochastic versions of one such model. We examine how the time of emergence or introduction of a novel strain affects co-existence with existing strains and hence the initial establishment of a new evolutionary branch. We also clarify the roles of cross-immunity and the basic reproductive number in this process. We show that the basic reproductive number is important because it affects the frequency of infection, which influences the long term immune profile of the host population. The time at which a new strain appears relative to the epidemic peak of an existing strain is important because it determines the environment the emergent mutant experiences in terms of the short term immune profile of the host population. Strains are more likely to coexist, and hence to establish a new clade in the viral phylogeny, when there is a significant time overlap between their epidemics. It follows that the majority of antigenic drift in influenza is expected to occur in the earlier part of each transmission season and this is likely to be a key surveillance period for detecting emerging antigenic novelty.

Keywords: partial cross-immunity, antigenic drift, evolution, epidemiology

1. Introduction

One of the most striking characteristics of the influenza A virus is its extraordinarily rapid evolution due to strong selection mediated by the host immune response to viral antigens. If a strain of influenza causes a large epidemic, the majority of hosts acquire immunity against that strain. In order to be successful, subsequent viral strains must therefore escape the residual host immunity and this promotes the fixation of mutants with novel epitopes. A key feature of influenza A evolution is that, despite the appearance and spread of new strains each year, the vast majority become rapidly extinct and the number of persistent branches in the phylogenetic tree remains small (Bush et al. 1999; Smith et al. 2004; Holmes et al., 2005). One of the most important factors responsible for maintaining this slim phylogenetic tree is likely to be cross-immunity (Andreasen et al., 1997; Ferguson et al. 2003; Koelle et al. 2006, Andreasen and Sasaki, 2006), whereby infection with one strain suppresses subsequent infections with antigenically similar strains. Cross-immunity is a major determinant of phylogenetic branching because it can make it difficult for a mutant strain to coexist with the parental or sibling strains.

A key feature of influenza dynamics is the strong seasonality of incidence in temperate regions. To understand the ecology and evolution of influenza, it is vital to understand the way in which host immunity mediates the epidemiological and antigenic interaction of viral strains in the context of seasonal epidemics. An earlier study (Andreasen and Sasaki, 2006) considered a simple model in which discrete annual epidemics of influenza were assumed to be caused by antigenic variants derived from a common ancestor in the previous year. The variants occupied a low-dimensional antigenic space and emerged serially to cause sequential, non-overlapping epidemics. This analysis indicated that strong temporary broad cross-immunity between all viral strains, a high basic reproductive rate and rapid decay in long term strain specific cross-immunity are required to prevent evolutionary branching and maintain the slim antigenic phylogeny of influenza (Andreasen and Sasaki, 2006). Empirical studies, however, show that antigenic variants do not appear in strict sequence each season and there may be extensive co-circulation. In New York State two distinct clades of H3N2 co-circulated in the 2002 – 2003 transmission season and three distinct clades circulated in the 2003 – 2004 season (Holmes et al. 2005). In the 2006 – 2007 season multiple co-circulating clades causing overlapping epidemics were detected (Nelson et al. 2008). In each season each clade is believed to have arisen from a novel introduction (Nelson et al. 2006; 2008).

There are many complex factors governing the epidemiology and evolution of influenza. Here we will focus on the roles that seasonal epidemics and the extent of co-circulation of antigenic variants play in antigenic branching. We show that partially overlapping epidemics can enhance co-existence of antigenic variants, making branching more likely. This effect is strongly dependent on the time lag between the introduction of each variant. It may also be moderated by season to season variation in the sequence in which variants appear. Nevertheless maintenance of the characteristic unbranched phylogeny of influenza may require a higher basic reproductive number or stronger broad cross-immunity than anticipated by previous work.

Models based on linear antigenic spaces are often used to analyze the coupled

dynamics of the host population immune profile and the antigenic properties of the influenza virus (Sasaki, 1994; Andreasen et al., 1996; Haraguchi and Sasaki, 1997; Gog and Grenfell, 2002; Andreasen, 2003; Lin et al., 2003; Boni et al., 2004; Adams and Sasaki, 2007;) because the antigenic drift of influenza A shows approximately linear evolution along a single phylogenetic trunk. However, antigen characteristics are thought to be high dimensional (Smith et al. 2004) and other approaches define antigenic types based on abstractions of the epitope sequence (Sasaki and Haraguchi, 2000; Ferguson et al, 2003; Tria et al, 2005; Adams and Sasaki, 2009). Therefore, we also use an individual-based simulation of the epidemic model to consider the effect of modeling the antigenic evolution through a high dimensional antigenic space. We find that the key characteristics of the evolutionary dynamics in this framework are similar to those of the simpler model, indicating that reducing the antigenic space to one-dimension is indeed a reasonable approximation that produces robust insights.

2. Deterministic Model

We consider an extension of the model proposed by Andreasen and Sasaki (2006), hereafter denoted the AS-model. This framework focuses on the seasonal dynamics of influenza to separate the epidemiological and evolutionary timescales. The epidemiological dynamics are expressed in a continuous time structure within a single transmission season. The evolutionary dynamics are expressed through a discrete time structure. Time is divided into a sequence of consecutive seasons and the host population is classified according to the most recent season in which infection occurred. Within any given season the antigenic distance between viral genotypes in the same clade is assumed to be small relative to the distance between clades or the distance between genotypes in different seasons. The system is then simplified by representing the whole collection of genotypes making up a clade as a single ‘strain’. In temperate regions multiple introductions seed epidemics of antigenically distinct clades in each transmission season. Diversity is replenished each season from an extensive global gene pool but there is little positive selection over the course of the epidemic (Holmes et al. 2005; Nelson et al. 2006, 2008). Therefore the antigenic type of each clade is assumed to evolve at a constant rate between seasons but remain static within any given season.

At the beginning of each season a mutant strain with a fixed antigenic divergence from the ancestral strain in the previous season founds an epidemic in the host population. The epidemiology of the strain follows the standard SIR model, but the infectivity of the virus depends on host cross-immunity due to past infections. Cross-immunity is modeled as an exponentially decaying function of the time since last infection. At some point within the same season a second strain with fixed divergences from the ancestral and sibling strains is introduced. In the AS-model, the second strain is only introduced when the epidemic of the founder strain has finished. Here we also consider the impact of introducing the second strain during the founder epidemic. Persistence of both strains is a necessary condition for the establishment of a new antigenic branch. We are mainly interested in the conditions that prevent such co-existence and may thus lead to a slim phylogenetic tree.

As in the AS-model, let S_k denote the fraction of hosts whose most recent infection occurred k season ago, and have not yet been infected in the current season.

Let I_k denote the fraction of hosts that are currently infected, and whose last previous infection occurred k seasons ago. Similarly, let R_k denote the fraction of hosts that have recovered from infection in the current season and whose last previous infection occurred k seasons ago. For simplicity, the total number of hosts is assumed to be constant ($\sum_j (S_j + I_j + R_j) = 1$) and there is no birth, death, or immigration during the epidemic period. The infectivity of the strain in an infected host is assumed to depend on the number of years since the host was last infected, k , according to $c\tau_k$. Here c is a constant contact rate and τ_k ($0 = \tau_0 < \tau_1 < \dots < 1$) is an increasing function describing the strength of cross-immunity. With these assumptions, the differential equations that describe changes in S_k and I_k in a season are:

$$\begin{aligned} S'_k &= -\tilde{\Lambda}S_k, \\ I'_k &= \tilde{\Lambda}S_k - \gamma I_k, \\ R'_k &= \gamma I_k, \\ \tilde{\Lambda} &= c \sum_k \tau_k I_k, \end{aligned} \tag{1}$$

where the prime ' denotes the time derivative d/dt , $\tilde{\Lambda}$ denotes the force of infection and γ is the recovery rate. As in the AS-model, we assume the following functional form for the cross-immunity:

$$\tau_k = 1 - \alpha^k. \tag{2}$$

Here, α describes how the strength of cross-immunity decays after one season. The cross-immunity decays exponentially with the number k of seasons since the last infection, as shown in Figure 1 for different values of α . We assume that among the past infection events, the most recent infection determines the strength of cross-immunity against the currently circulating strain. Rescaling time in units of the average duration of infection $1/\gamma$, yields

$$\begin{aligned} S'_k &= -\Lambda S_k, \\ I'_k &= \Lambda S_k - I_k, \\ R'_k &= I_k, \\ \Lambda &= \rho \sum_j \tau_j I_j, \end{aligned} \tag{3}$$

where $\rho = c/\gamma$ is the basic reproductive rate, the expected number of secondary infections from a single host infected with a particular strain when the rest of the host population has no immunity to any strain.

Let $S_k^p(\infty)$ be the density of the hosts still uninfected at the end of season p whose most recent infection occurred k seasons ago, and $R_k^p(\infty)$ be the corresponding quantity for the hosts infected and recovered in season p . Then the initial condition of the host population at the start of season $p+1$ is expressed in terms of the state at the end of season p by:

$$\begin{aligned}
S_1^{p+1}(0) &= \sum_{k=1} R_k^p(\infty) \\
S_{k+1}^{p+1}(0) &= S_k^p(\infty) \quad (k \geq 1)
\end{aligned}
\tag{4}$$

The epidemic dynamics in the year $p + 1$ are then described by (3) where the superscripts to designate the year are dropped for notational simplicity.

We now expand the AS-model to evaluate how the extent to which two strains co-circulate within the same epidemic season affects the evolutionary dynamics in terms of the establishment of novel antigenic branches. The immune profile of the host population becomes more complex. To incorporate the immune interaction between co-circulating strains we label the strain present at the start of the season A, the strain emerging at some later point B, fix the strength of partial cross-immunity α and constant contact rate c to be the same for both strains and assume that strain A has been circulating for many seasons but strain B first reaches detectable prevalence in season T . Following the AS-model, we assume that both strains A and B are first present at very low prevalence at the end of season $T - 1$. Therefore, when they reach detectable prevalence, in season T , the antigenic distance between strains A and B is equal to the antigenic distance associated with two seasons in the one strain model. If there are multiple previous infections only the strongest partial cross-immunity is assumed to be effective.

Inter-pandemic influenza subtypes generally circulate for several decades, undergoing antigenic drift throughout this period. The H3N2 subtype currently responsible for the majority of seasonal influenza infections has circulated since 1968 (Kilbourne, 2006) and it is reasonable to assume that the epidemiological dynamics have settled to a stable pattern. The initial immune profile of the population at the start of season T , in terms of the length of time since the most recent infection, is found from this steady state. The population is then challenged with a new strain, A, at the start of the T th season. Later in the same season the population is challenged with another new strain, B. Thereafter, the density of hosts infected with these strains is recorded. The population is now categorized into $9k$ states, susceptible (S), infected (I) and recovered (R) for each strain A and B with the most recent previous infection experienced k seasons ago. For example, the proportion of hosts currently infected with strain A, recovered from infection in the current season with strain B and, besides from the current season, last infected k season ago is given by IR_k . The rate of change of each host state is described by the differential equation system:

$$\begin{aligned}
SS'_k &= -(\Lambda_A + \Lambda_B)SS_k, \\
IS'_k &= \Lambda_A SS_k - IS_k - \nu \Lambda_B IS_k, \\
RS'_k &= IS_k - \nu \Lambda_B IS_k, \\
SI'_k &= \Lambda_B SS_k - SI_k - \nu \Lambda_A SI_k, \\
II'_k &= \nu(\Lambda_A SI_k + \Lambda_B IS_k) - 2II_k, \\
RI'_k &= \nu \Lambda_B RS_k + II_k - RI_k, \\
SR'_k &= SI_k - \nu \Lambda_A SR_k, \\
IR'_k &= \nu \Lambda_A SR_k + II_k - IR_k, \\
RR'_k &= RI_k + IR_k,
\end{aligned} \tag{5}$$

where the parameter ν represents the proportional susceptibility reduction due to temporary non-specific cross-immunity. Ferguson et al. (2003), Tria et al. (2005) and Andreason and Sasaki (2006) all note that it is hard to produce a phylogenetic tree with the shape characteristic of influenza without this factor, unless the basic reproductive ratio is close to 1. Here, $0 \leq \nu \leq 1$ and the effect of temporary non-specific immunity becomes smaller as ν becomes larger. This immunity is only thought to persist for 3-4 months so is not carried over to the next season. Note that co-infection with both strains can occur. However, the II compartment is expected to be very small, and unlikely to play a significant part in the dynamics, because the duration of infection is short and hosts infected with one strain immediately gain non-specific partial immunity to all other strains for the remainder of the season.

The forces of infection of strains A and B, Λ_A and Λ_B , are given by

$$\begin{aligned}
\Lambda_A &= \rho \left(\sum_{k=1} \tau_k (IS_k + II_k + IR_k) + \tau_0 (IS_0 + II_0) + \tau_2 IR_0 \right), \\
\Lambda_B &= \rho \left(\sum_{k=1} \tau_k (SI_k + II_k + RI_k) + \tau_0 (SI_0 + II_0) + \tau_2 RI_0 \right),
\end{aligned} \tag{6}$$

Here $\rho = c / \gamma$ is the basic reproductive rate, the expected number of secondary infections from a single host infected with a particular strain when the rest of the host population has no immunity to any strain. The suffix $k=0$ indicates hosts that have never been infected prior to the current season. The proportional reduction in infectivity due to partial cross-immunity of hosts who are currently infected with strain A, have recovered from an infection with strain B in the current season, but have never been infected prior to the current season (IR_0) is τ_2 .

Following Andreassen and Sasaki (2006), we assume that the host population has experienced annual outbreaks of a single constantly evolving virus line (A) for a long time, and hence the fractions s_k of hosts last infected k seasons ago are at steady state. The equilibrium values of s_k at the onset of season are thus given by

$$\hat{s}_k = (1 - \phi)\phi^{k-1}, \quad (k = 1, 2, \dots), \quad (7)$$

where ϕ is the fraction of hosts that remain susceptible in each year determined from

$$0 = \rho q(1 - \phi) + \log \phi, \quad (8)$$

where $q = \sum_k \tau_k s_k = \sum_k \tau_k (1 - \phi)\phi^{k-1}$. This follows by considering the equilibrium of (4) noting that $S_k^p(0) = S_k^{p+1} = s_k$ and $R_k^p(\infty) = S_k^p(0)\phi = s_k\phi$. With the partial cross-immunity function (2) assumed in this paper, $q = \sum_{k=1}^{\infty} (1 - \alpha^k)(1 - \phi)\phi^{k-1} = 1 - \alpha\phi(1 - \phi) / \{\phi(1 - \alpha\phi)\}$ and hence (8) can be rewritten as

$$0 = \rho(1 - \phi)[1 - \alpha\phi(1 - \phi) / \{\phi(1 - \alpha\phi)\}] + \log \phi. \quad (9)$$

In order to investigate overlapping epidemics of strains A and B, we introduce strain A at the beginning ($t = 0$) of a season. The initial condition is then

$$\begin{aligned} SS_k(0) &= (1 - \varepsilon_A)\hat{s}_k, \\ IS_k(0) &= \varepsilon_A\hat{s}_k \end{aligned} \quad (10)$$

($k = 1, 2, \dots$) where ε_A is a small positive constant representing the initial fraction of strain A infected hosts, and \hat{s}_k is the initial fraction of hosts whose last infection was k seasons ago as given by (7) and (9). All other classes are zero at $t = 0$. Some time $t = T_d$ after the epidemic of strain A starts, strain B is introduced and the initial condition for the integration of (2) for $t \geq T_d$ is:

$$\begin{aligned} SI_k(T_d) &= \varepsilon_B SS_k(T_d - 0), \quad SS_k(T_d) = (1 - \varepsilon_B)SS_k(T_d - 0) \\ II_k(T_d) &= \varepsilon_B IS_k(T_d - 0), \quad IS_k(T_d) = (1 - \varepsilon_B)IS_k(T_d - 0) \\ RI_k(T_d) &= \varepsilon_B RS_k(T_d - 0), \quad RS_k(T_d) = (1 - \varepsilon_B)RS_k(T_d - 0) \end{aligned} \quad (11)$$

($k = 1, 2, \dots$) where ε_B is a small positive constant and the fraction $SS_k(T_d - 0)$, for example, represents the value of $SS_k(t)$ in the limit $t \rightarrow T_d$. The other three classes for each k , $SR_k(t)$, $IR_k(t)$ and $RR_k(t)$ are zero at $t = T_d$.

The invasibility of strain B in any given year is determined by the sign of the rate of change of the strain B infected population at time $t = T_d$, that is $I'_B(T_d)$ where $I_B(t) \equiv \sum_k (SI_k(t) + II_k(t) + RI_k(t))$ is the total fraction of strain B infected hosts. Strain B will invade the population and cause an outbreak in that year if $I'_B(T_d) > 0$. We examine how this invasibility is related to the degree of overlap between the epidemics, which is measured by $(T_{ea} - T_d) / T_{ea}$, where T_{ea} is the time from the beginning of the season to the moment when the fraction of hosts infected with strain A falls below a fixed extinction threshold.

The coexistence of two strains in the same season is a necessary condition for

branching. Influenza phylogenies, however, show distinct branches coexisting for several years until one forms the trunk of the phylogeny and the others become extinct (Bush et al. 1999). Therefore, we now consider the longer term persistence of branches. We assume that strain B successfully increases in the first season it emerges (T) and ask if the progeny of both strains A and B still persist n seasons after season T . To determine the invasibility of the progeny of strains A and B in season $T + n$, we must know the degree of partial cross-immunity they experience i.e., the shortest path connecting the current strain and the strains responsible for any previous infections along the phylogenetic tree as shown in Figure 2. More specifically, we need to know not only the number of seasons k before T that the host has last infected, but also the infection history after season T in terms of the progeny of strains A and B. Let a (b) be the number of seasons since the host was *last* infected by the progeny of strain A (B). Let f_a (f_b) be the number of seasons after T since the host was *first* infected by the progeny of strain A (B) (see Figure 2(a)). The infectivity $c\tau_{k,a,b,f_a,f_b}$ of strain A infected hosts depends on k , a , b , f_a , and f_b . The entries τ_{k,a,b,f_a,f_b} of Table 1 show the infectivity to the progeny strains of A and B in the year $T + n$ for the hosts with state $\{k,a,b,f_a,f_b\} = \{+,0,+,0,+ \}$ i.e. those who were last infected k years ago at the onset of the season T , have not been infected by any progeny of strain A since then ($a = f_a = 0$), and have been infected at least once by progeny of strain B ($b, f_b > 0$). The infectivity $\tau_{k,a,b,f_a,f_b}^{A,n}$ of the progeny of strain A is determined by the most antigenically similar strain i.e. the smaller of $1 - \alpha^{n+k}$ and $1 - \alpha^{2n+2-f_b}$ as shown in Figure 2(b). $\tilde{\tau}_{k,a,b,f_a,f_b}^{A,n}$ is the corresponding quantity when the host has already been infected by the other co-circulating strain in the same year.

Each of the nine host states in the current season is now additionally classified according to infection history k, a, b, f_a, f_b (e.g. $IS_{(k,a,b,f_a,f_b)}$ represents the density of hosts that are currently infected by strain A, have not yet been infected by strain B in the current season and whose infection history with respect to the ancestral strains of A and B is as shown in Fig 2a). The system is described by differential equations extended from (5) in the obvious way, with the forces of infection of strain A and B, Λ_A and Λ_B , given by

$$\begin{aligned} \Lambda_A &= \rho \sum_k \sum_a \sum_b \sum_{f_a} \sum_{f_b} \left\{ \tau_{k,a,b,f_a,f_b}^{A,n} (IS_{(k,a,b,f_a,f_b)} + II_{(k,a,b,f_a,f_b)}) + \tilde{\tau}_{k,a,b,f_a,f_b}^{A,n} IR_{(k,a,b,f_a,f_b)} \right\} \\ \Lambda_B &= \rho \sum_k \sum_a \sum_b \sum_{f_a} \sum_{f_b} \left\{ \tau_{k,a,b,f_a,f_b}^{B,n} (SI_{(k,a,b,f_a,f_b)} + II_{(k,a,b,f_a,f_b)}) + \tilde{\tau}_{k,a,b,f_a,f_b}^{B,n} RI_{(k,a,b,f_a,f_b)} \right\} \end{aligned} \quad (12)$$

Coexistence is determined by the sign of the initial growth rate of the second strain to be introduced. For example, if strain A is initially present and strain B is introduced at $T_d > 0$, then the two strains coexist if $I'_B(T_d) > 0$.

3. Results

In order to understand the coexistence of two strains we the first consider the

model in which only one strain circulates in each season (equation (3)). When the dynamics have stabilized, the distribution of the S_k at the beginning of each season corresponds to the time interval between infections of the same host. As shown in Figure 3, when the basic reproductive number ρ becomes larger, the time interval between infections becomes longer.

With this observation regarding the immune profile in mind, we now analyze the model in which two strains may co-circulate in the same season. Figure 4(a) shows how coexistence in the season that strain B first emerges (T) depends on the basic reproductive rate ρ , the strength of temporary non-specific immunity ν , and the degree of overlap between the epidemics of the two strains. The results can be summarized as follows: (i) If two epidemics do not overlap ($T_d = T_{ea}$) coexistence occurs if ν and ρ are large. (ii) If the epidemics overlap a little ($T_d > (3/4)T_{ea}$) the result remains the same. (iii) If the epidemics overlap a lot, the two strains always coexist except when ν is very small and ρ is very large. The observation that epidemics with different influenza strains often do overlap therefore indicates that, in order to prevent branching, the reproductive number ρ must be larger than previously suggested, or broad cross-immunity ν must be stronger. Since it is known that the basic reproductive number of influenza is relatively small, this suggests that broad temporary cross-immunity is very effective or there are additional factors involved. If the order of emergence is the same in seasons T and $T+1$ the condition for continued coexistence in season $T+1$ is exactly the same. If the order of emergence is reversed between seasons T and $T+1$ the way in which coexistence depends on ρ and ν remains qualitatively similar but is restricted to a smaller region of parameter space (Figure 4 (b)). In particular, if T_d is high and ρ is small coexistence always occurs for any value of ν . Furthermore, the range of values of T_d that result in large changes in the coexistence condition is narrower than when the order of epidemics is the same and the difference between the coexistence regions associated with $T_{ea}/2$ and T_{ea} is almost undetectable.

Figure 5 summarizes the relationship between the sequence of epidemics and the fate of strains when the A and B epidemics in each season overlap ($T_d = (1/4)T_{ea}$) and do not overlap. The solid line shows the boundary for coexistence when the two strains appear in the same order (AB) in both seasons. The broken line shows the corresponding boundary when the two strains appear in the reverse order (BA) in the second season. If there is no overlap between the two epidemics the result agrees with the analysis of Andreasen and Sasaki (2006). If the two epidemics overlap, the region of the $\rho - \nu$ parameter space in which strain B excludes strain A is wider when the order of epidemics is reversed between seasons. The effect of overlap between the two strains can thus be summarized as follows: (i) Coexistence occurs over a more extensive region of the $\rho - \nu$ parameter space. (ii) If coexistence does not occur strain B is more likely to exclude strain A if the order of appearance is reversed between seasons.

We now consider conditions for coexistence in the season in which strains A and B appear together for the first time (T) and the following two seasons. There are four possible combinations for the order of epidemics ($AB \rightarrow AB \rightarrow AB$), ($AB \rightarrow AB \rightarrow BA$), ($AB \rightarrow BA \rightarrow AB$) and ($AB \rightarrow BA \rightarrow BA$). In season $T+2$,

two progeny strains are more likely to coexist than their ancestors due to the increased antigenic distance between them. So if two strains coexist in $T + 1$ and the order of epidemics in season $T + 2$ is the same as $T + 1$, $((AB \rightarrow AB \rightarrow AB)$ or $(AB \rightarrow BA \rightarrow BA))$, coexistence will always continue. If, however, the order of appearance is reversed between seasons $T + 1$ and $T + 2$, coexistence becomes more limited (Figure 6). Thus, reversal of the order of epidemics between seasons reduces coexistence in the same way as was observed when only seasons T and $T + 1$ were considered. Phylogenetic analysis of data from New York State did not find any clear pattern in the order that strains appear from one season to the next (Nelson et al. 2006) suggesting that variation is extensive. Therefore, variation in the order that strains appear is expected to moderate some of the effects of overlapping epidemics and is likely to be a factor in the prevention of branching.

4. Individual-based simulation

In the differential equation based model discussed so far the time scale of the epidemic is separated from that of the drift process. The main conclusion of our analysis was that two strains are more likely to coexist if there is a shorter time lag between their appearance and hence a greater overlap between their epidemics. Branching may be initiated if two strains appear at a similar point of the season. However, it may be subsequently terminated if the order in which these strains appear varies in later seasons. We now re-examine these conclusions using an individual-based model that simulates the epidemic and drift processes on the same time scale. Whereas the differential equation model is based on the approximation of constant antigenic change in each season, we now model mutation explicitly and only assume that the antigenic distance associated with each mutation is constant.

We consider a host population of $N = 10^5$ individuals and record the immune state of each host with respect to each virus strain. The probability that a host susceptible to strain A becomes infected is

$$\Lambda_A = \rho \sum_{x \text{ infected with A}} \tau_{x,A} / N, \quad (13)$$

where the summation is over all hosts, x , infected with strain A and $\tau_{x,A}$ is the reduced infectivity of strain A in host x due to partial cross-immunity, $\rho = c / \gamma$ is the basic reproductive rate, $N = 10^5$ is total host population size, c is a constant contact rate and γ is the recovery rate. In order to determine the strength of partial cross-immunity, we need to determine the antigenic distance between two strains. Here we assume this is equal to the Hamming distance. We consider an epitope consisting of ten amino acid residues (loci) and assume each locus has two variants, 0 or 1. The antigenic distance between two strains is then given by the number of loci with different values. For each infected host, one locus of the infecting strain switches value due to mutation with probability μ . Denoting the infectivity reduction rate associated with one mutation by α , the Hamming distance between strains A and B by $d(A,B)$ and assuming only the strongest partial cross-immunity from all past infections is relevant,

$$\tau_{x,A} = \min_{B|x \text{ recovered from B}} (1 - \alpha^{d(A,B)}) . \quad (14)$$

Infected hosts recover with probability γ and gain complete temporary immunity to all strains, which lasts an average of 1/5 of a year. The probabilities of birth and death are equal, and newborn hosts are susceptible to all strains. Initially ten hosts are infected with the same strain and susceptible to all other strains while the remainder of the population is susceptible to all strains. The system was iterated using a continuous time Markov process.

Figure 7 shows a phylogenetic tree produced by this model. With temporary non-specific immunity the phylogenetic tree shows an approximately linear shape despite rapid turnover of antigenic strains, justifying our assumption that the viral population is nearly monomorphic in each year and escape mutations constantly accumulate. As with the differential equation model, we now focus on the coexistence of a mutant strain M and its immediate progenitor P . We define M and P to be coexistent if they are both present in the population when another mutation occurs in M . This definition rules out situations in which small outbreaks of P occur by chance even though sustained coexistence is not possible.

Using this model, we tested whether the lag between the appearance of strains P and M is correlated with their coexistence. Figure 8 shows the distribution of the number of mutant strains that coexist with their progenitor P when the lag between their appearance is T_d . This sampling distribution is compared with a theoretical distribution for the expected number of coexisting strains when the probability of coexistence is independent of T_d and given by

$$F_0 = \frac{\sum_{T_d=0}^{\infty} M_g(T_d)}{\sum_{T_d=0}^{\infty} M(T_d)} \quad (15)$$

Here $M(T_d)$ denotes the number of strains M for which the appearance lag relative to P is T_d and $M_g(T_d)$ denotes the number of strains that have an appearance lag relative to P of T_d and coexist with P . The Kolmogorov-Smirnov test at the critical level 0.01 rejects the null hypothesis that the sampling distribution corresponds to the theoretical distribution. Hence appearance times are correlated with coexistence. Figure 9 shows the relationship between the lag in the appearance time of P and M (T_d) and the probability of coexistence. Spearman's rank-correlation coefficient between the length of the lag and the coexistence probability is -0.4180821 (p-value < 0.01). We conclude that the key dynamics of the deterministic model based on assumption of continuous, linear antigenic divergence are in good agreement with the stochastic model based on a high dimensional antigenic space. If the lag between the appearance of two strains is shorter then they are more likely to coexist. Overlapping epidemics facilitate branching.

5. Discussion

We have developed a model for influenza evolution over discrete transmission seasons to examine the impact of emergence time and epidemic overlap on the

coexistence of antigenically divergent strains and assess the implications for evolutionary branching. Previous work based on non-overlapping epidemics has shown that maintaining the slim antigenic phylogeny of seasonal influenza is likely to require a relatively high basic reproductive number and strong temporary broad immunity. We have shown that if epidemics overlap an even higher basic reproductive number or stronger temporary immunity are required to prevent branching although variation in the order in which strain appear over several seasons may moderate these constraints. We have also shown that these results arise in both a deterministic model with a basic antigenic space and a stochastic model with a high dimensional antigenic space.

A new branch is established when two distinct strains emerge and coexist. When two strains appear in the same season coexistence is determined by the survival of the second strain and cross-immunity means that hosts infected with, or recovered from, the first strain are difficult for the second strain to re-infect. The key parameters that determine coexistence, therefore, are the decay rate in the strength of partial cross-immunity after one season α , the basic reproductive rate ρ , the proportional reduction in the force of infection due to temporary non-specific immunity ν , and the appearance time of the second strain T_d . The role of α and ν is easy to understand as these parameters determine partial cross-immunity and temporary non-specific immunity directly. The role of T_d and ρ , however, is indirect as these parameters determine the immune profile of the host population. A larger value of the basic reproductive rate ρ reduces the time between infections (Figure 3). Given that partial cross-immunity decays with each passing season, and is almost absent after about five seasons, the fraction of hosts who benefit from partial cross-immunity is larger when ρ is larger. In addition to reducing the time between infections, larger values of the basic reproductive rate increase the magnitude of epidemics as shown by S_1 in Table 2. So ρ affects both the fraction of hosts infected and the time between infections. Increasing the time lag between the appearance of each strain (T_d) means that larger values of ρ , or smaller values of ν , are required for coexistence, as Figures 4 and 6 show. This happens because, if the second strain emerges later, a larger fraction of the host population have immunity due to infection with the first strain.

The sequence in which strains appear is also important for coexistence over two seasons. If the sequence of appearance is the same in seasons T and $T+1$, continued coexistence is more likely than if the sequence is reversed between seasons (Figure 4). In order to understand this phenomenon, we focus on the case in which the sequence is reversed. In season T , strain A emerges first and strain B is suppressed due to host immunity. Hence the number of hosts infected with strain B is small, and so the number of hosts acquiring complete immunity to strain B is also small. This small epidemic may, however, be a herald wave (Glezen et al., 1982), because if strain B appears first in season $T+1$, host immunity is weak and a large strain B epidemic follows. The host immunity arising from this epidemic then suppresses the prevalence of strain A, possibly driving it to extinction. When two epidemics overlap, the exclusion of the preexisting strain by a strain newly established in a herald wave is much more likely to occur (Figure 5). Once two strains succeed in coexisting it becomes more likely in subsequent seasons, although the extent of potential coexistence is always reduced by a reversal in the order of appearance between seasons.

The key result of both the differential equation based model and the

individual based model is that strains are more likely to coexist if the difference between their appearance times is smaller. This conclusion offers important insight into the timing of the emergence of novel influenza strains. For the individual based model, the sampling distribution shown in Figure 8 corresponds to non-neutral evolution whereas the theoretical distribution corresponds to neutral evolution. Here, the degree of overlap between epidemics can also be interpreted as the time at which a mutant emerges from a circulating strain. The peak of the sampling distribution is earlier than that of the theoretical distribution indicating, in agreement with the analysis of Boni et al. (2006), that host immune selection causes excess antigenic drift and most of this drift occurs in the earlier part of each influenza season.

Genetic analysis of H3N2 and H1N1 isolates collected in the United States over the last decade show that there are multiple introduction events each season and several antigenically distinct clades may co-circulate (Nelson et al. 2006, 2007, 2008). Here we have shown that the extent of overlap between the epidemics associated with these clades is an important factor in determining whether or not a new antigenic branch is established, and persists. Clades that appear, by mutation or introduction, at a similar time are more likely to coexist and so result in a new antigenic branch. However, variation in the order in which these clades appear in subsequent seasons limits co-existence and may be an important factor preventing the persistence of new antigenic branches. Analysis of H1N1 epidemiological data has shown that the mutant or clade that emerged earliest in the influenza season caused a major epidemic whilst clades emerging later caused much smaller outbreaks (Nelson et al. 2008). Correspondingly, in our model, strains that emerge when an epidemic of another strain is already well underway are unlikely to lead to significant co-circulation or the establishment of new antigenic branches. This insight suggests that tracking and predicting the antigenic evolution of influenza virus may be improved by focusing attention on the early stages of epidemics.

Acknowledgments

We thank Drs. Yoh Iwasa and Takahiro Irie for their helpful comments. We also thank Yoshihiro Morishita, Jun Nakabayashi, Sayaki Suzuki and Kouichi Saeki for discussion. This work was supported in part by the Research Program of Hayama Center for Advanced Studies of Sokendai, and by the Research Fellowship of Japan Society for the Promotion of Science.

References

- Adams, B. & Sasaki, A. (2007). Cross-immunity, invasion and coexistence of pathogen strains in epidemiological models with one-dimensional antigenic space. *Mathematical Biosciences* 210:680-699.
- Adams, B & Sasaki, A (2009). Antigenic distance and cross-immunity, invasibility and coexistence of pathogen strains in an epidemiological model with discrete antigenic space. *Theor. Popul. Biol.* In press.
- Andreasen, V. (2003). Dynamics of annual influenza A epidemics with immuno-selection. *J. Math. Biol.* 46: 504-536.
- Andreasen, V., Levin, S. and Lin, J. (1996). A model of influenza A drift evolution. *Z. Angew. Math. Mech.* 76: 421-424.

- Andreasen, V., Lin, J., and Levin, S. (1997). The dynamics of cocirculating influenza strains conferring partial cross-immunity. *J. Math. Biol.* 35: 825-842.
- Andreasen, V. & Sasaki, A. (2006). Shaping the phylogenetic tree of influenza by cross-immunity. *Theor. Popul. Biol.* 70: 164-173.
- 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.
- Boni, M.F., Gog, J.R., Andreasen, V. & Feldman, M.W. (2006). Epidemic dynamics and antigenic evolution in a single season of influenza A. *Proc. R. Soc. B.* 273:1307-1316.
- Bush, R.M., Fitch, W.M., Bender, C.A. and Cox, N.J. (1999). Positive Selection on the H3 Hemagglutinin Gene of Human Influenza Virus A. *Mol. Biol. Evol.* 16 1999b 1457-1465.
- Ferguson, N.M., Galvani, A.P. & Bush, R.M. (2003). Ecological and immunological determinants of influenza evolution. *Nature* 422: 428-423.
- Glezen, W.P., Couch, R.B. & Six, H.R. (1982). The influenza herald wave. *Am. J. Epidemiol.* 116: 589-598.
- Gog, J. R. & Grenfell, B. T. (2002). Dynamics and selection of many-strain pathogens. *Proc. Natl. Acad. Sci. USA* 99: 17209-17214.
- Haraguchi, Y. & Sasaki, A. (1997). Evolutionary pattern of intra-host pathogen antigenic drift: Effect of cross-reactivity in immune response. *Phil. Trans. R. Soc. Lond. B* 352: 11-20.
- Holmes, E.C., Ghedin, E., Miller, N., Taylor, J., Bao, Y., St George, K., Grenfell, B. T., Salzberg, S. L., Fraser, C. M., Lipman, D. J., & Taubenberger, J. K. (2005). Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. *PLoS Biol.* 3:e300.
- Kilbourne, E.D. (2006). Influenza pandemics of the 20th century. *Emerg Infect Dis* 12: 9-14.
- Koelle, K., Cobey, S., Grenfell, B. and Pascual, M. (2006). Epochal evolution shapes the phylodynamics of interpandemic Influenza A(H3N2) in humans. *Science.* 314:1898-1903.
- Lin, J., Andreasen, V., Casagrandi, R. and Levin, S. A. (2003). Traveling waves in a model of influenza A drift. *J. Theor. Biol.* 212: 57-69.
- Nelson, M.I., Edelman, L., Spiro, D.J., Boyne, A.R., Bera, J., Halpin, R., Sengamalay, N., Ghedin, E., Miller, M.A., Simonsen, L., Viboud, C. & Holmes, E.C. (2008), Molecular Epidemiology of A/H3N2 and A/H1N1 Influenza Virus during a Single Epidemic Season in the United States. *Plos Pathog.* 4: e1000133.
- Nelson, M. I., Simonsen, L., Viboud, C., Miller, M. A. and Holmes, E. C. (2007). Phylogenetic analysis reveals the global migration of seasonal influenza A viruses. *Plos Pathog.* 3: e131.
- Nelson, M. I., Simonsen, L., Viboud, C., Miller, M. A., Taylor, J., St George, K., Griesemer, S. B., Ghedin, E., Sengamalay, N. A., Spiro, D. J., Volkov, I., Grenfell, B. T., Lipman, D. J., Taubenberger, J. K. and Holmes, E. C. (2006). Stochastic Processes Are Key Determinants of Short-Term Evolution in Influenza A Virus. *Plos Pathog.* 2: e125.
- Sasaki, A. (1994). Evolution of antigen drift and switching: Continuously evading pathogens. *J. Theor. Biol.* 168: 291-308.

- 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.
- Smith, D.J., Lapedes, A.S., de Jong J., Bestebroer, T.M., Rimmelzwaan, G.F., Osterhaus, A.D.M.E. and Fouchier, R.A.M. (2004). Mapping the Antigenic and Genetic Evolution of Influenza Virus. *Science*. 305, 371–375.
- Tria, F., Lassing, M., Peliti, L. & Franz, S. (2005). A minimal stochastic model for influenza evolution. *J. Stat. Mech. Theor. Exp.* P07008.

Figure legends

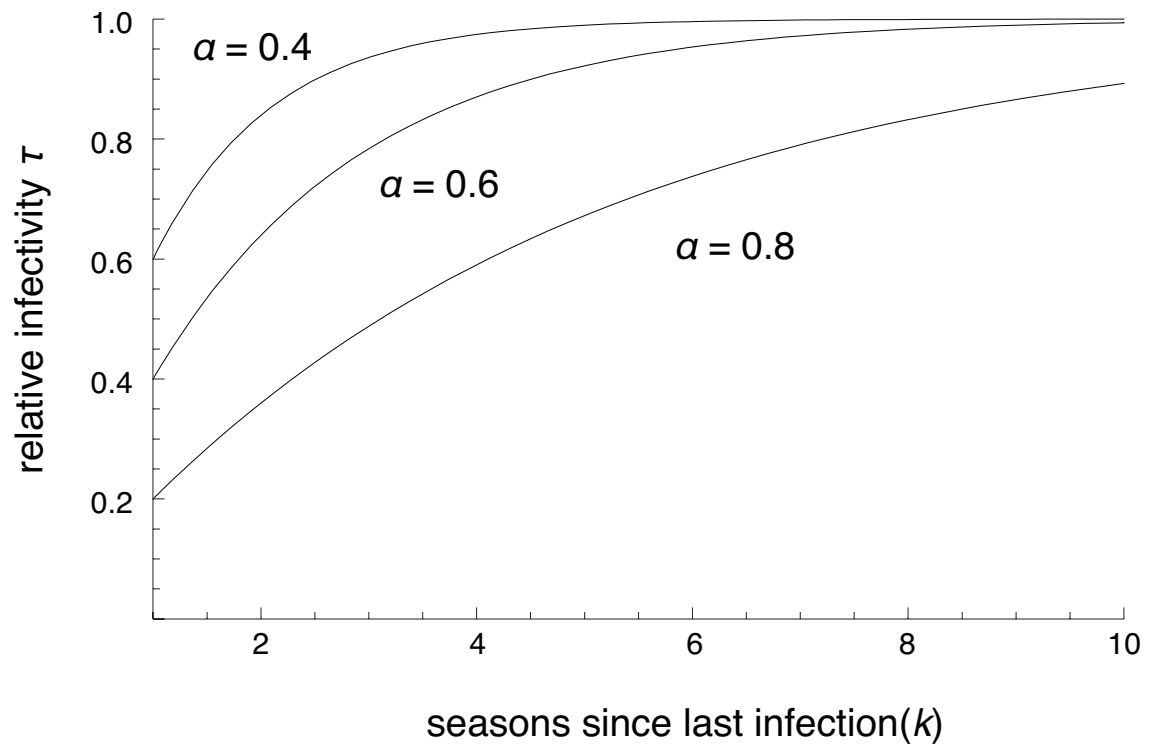


Figure 1: Decay in partial cross-immunity with time. The horizontal axis is the number of seasons k that have elapsed since the most recent infection. The vertical axis is the reduction in the force of infection due partial cross-immunity, $\tau = 1 - \alpha^k$. The parameter α determines the decay in partial cross-immunity after one season has elapsed.

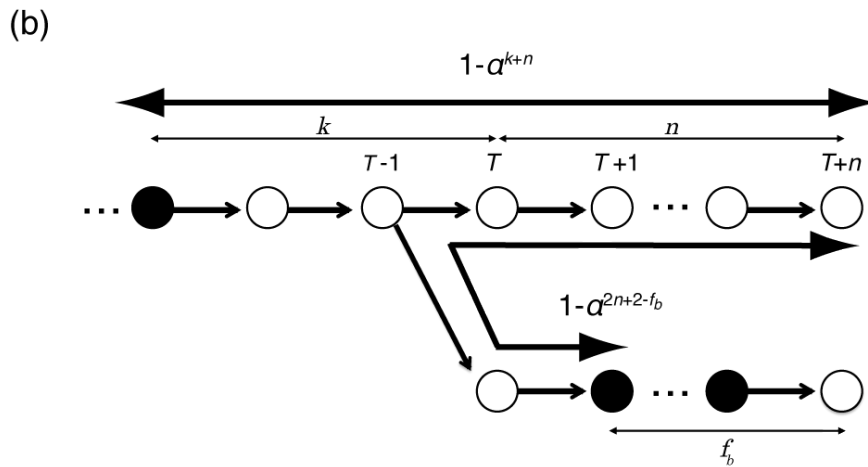
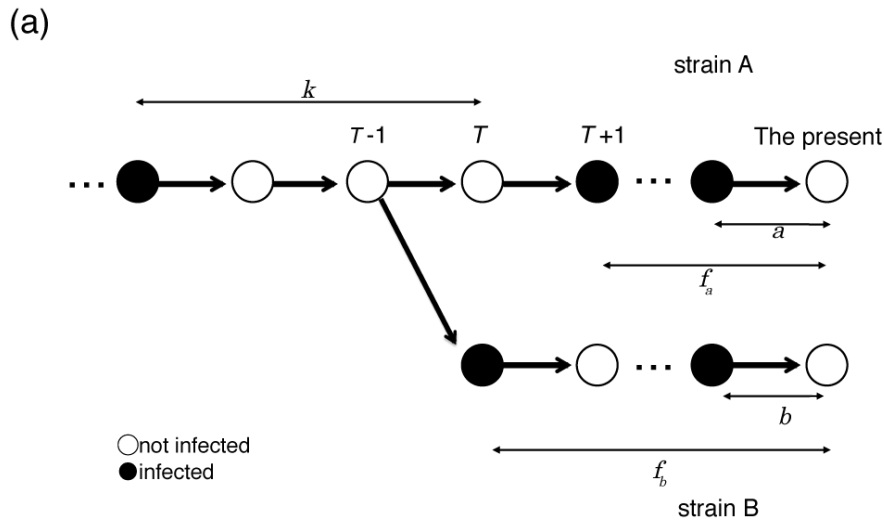


Figure 2: (a) Relationship between the season in which strains circulate relative to the first potential branch point in season T and the antigenic distance between them, indicating the definition of k , a , b , f_a , and f_b . (b) Cross-immunity is determined by the strain most antigenically similar to the one circulating in the current season $T+n$ i.e. the smaller of $1 - \alpha^{n+k}$ and $1 - \alpha^{2n+2-f_b}$.

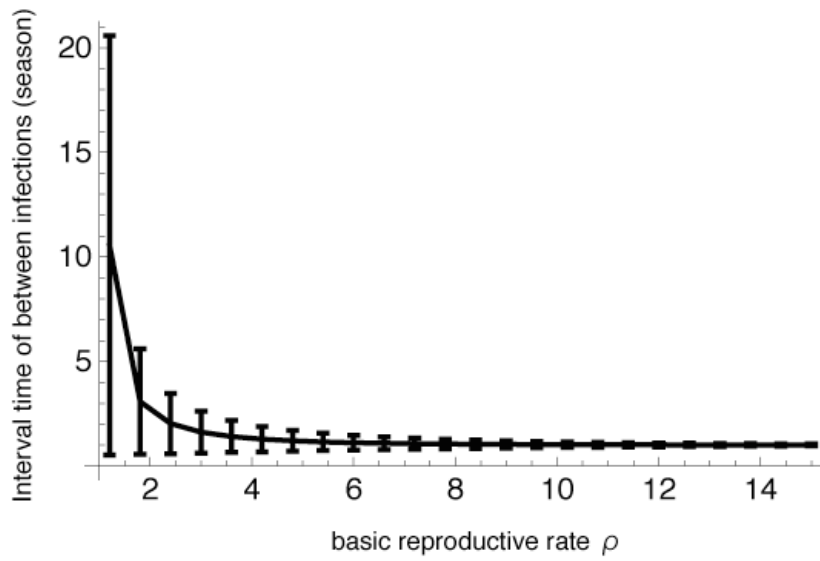


Figure 3: The relationship between the basic reproductive rate ρ and the mean time interval between of infections of same host when there is only one strain and it has reached equilibrium. Error bars denote standard deviation.

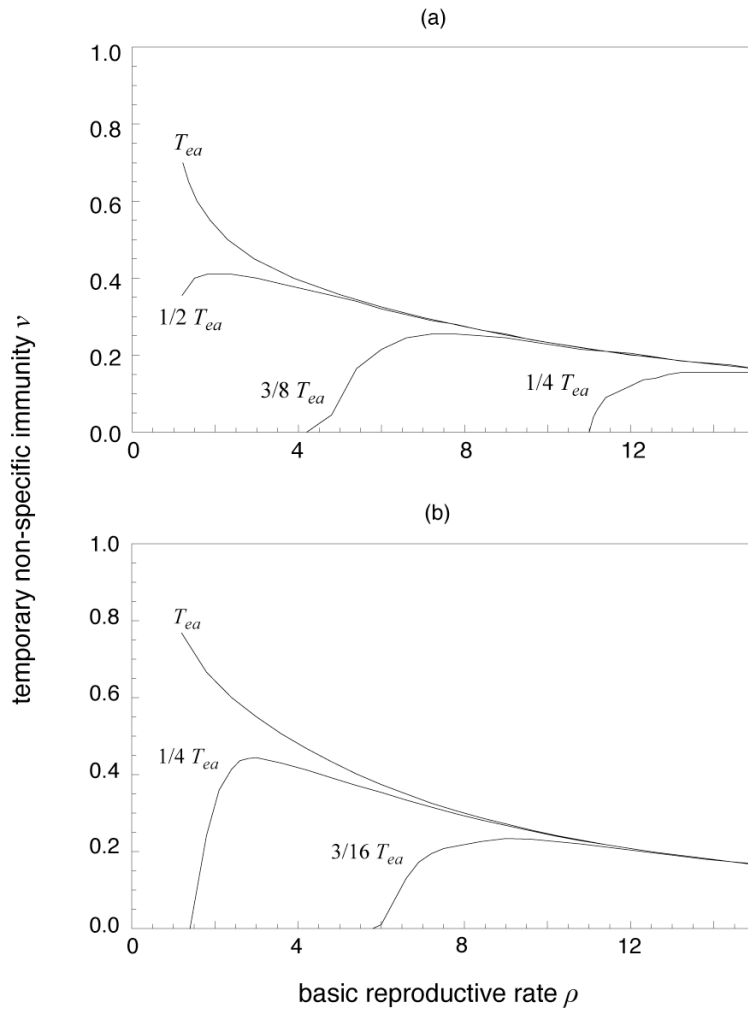


Figure 4: Conditions for coexistence of strains A and B in terms of parameters ρ and v for different lags between the appearance time of the strains, expressed as a proportion of the total epidemic duration T_{ea} (a) The sequence of epidemics is the same in seasons T and $T+1$. (b) The sequence of epidemics is reversed between seasons T and $T+1$. The two strains coexist in the regions above the solid lines, one strain is excluded in the regions below the lines, The initial number of infections when each strain appears is $\varepsilon_A = \varepsilon_B = 10^{-6}$. The extinction threshold for each strain is $\sum_k I_k = 5 \times 10^{-7}$.

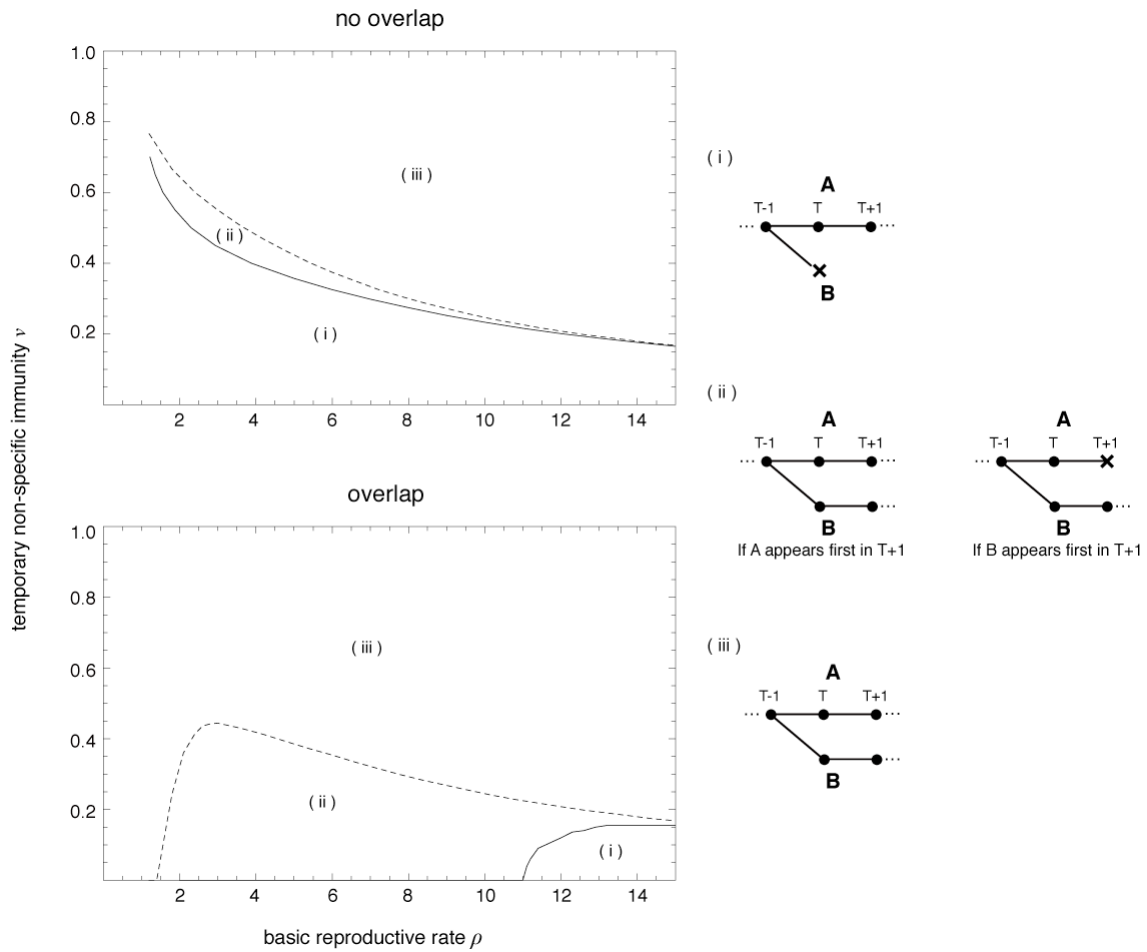


Figure 5: Relationship between the sequence of epidemics and the fate of the progeny of strains A and B when epidemics overlap ($T_d = (1/4)T_{ea}$) and do not overlap. The solid line is the boundary condition for coexistence if the sequence of epidemics is same in seasons T and $T+1$. The broken line is the boundary if the sequence of epidemics is reversed between seasons. The diagrams show which strains persist. ● indicates persistence, × extinction. In region (i) the two strains coexist. In region (iii) one strain is excluded. In the region between the solid and broken lines (ii), the outcome depends on the order of appearance. If strain A appears first in season $T+1$ there is coexistence but if strain B appears first in season $T+1$ strain A will be driven to extinction.

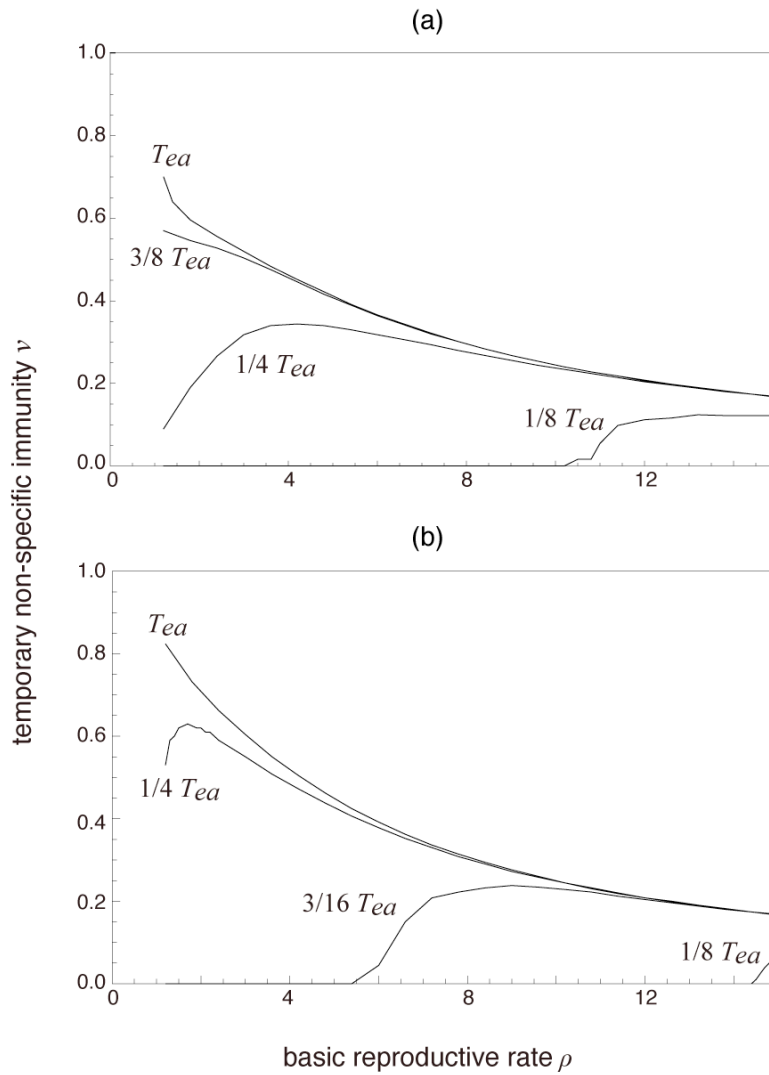


Figure 6: Conditions for coexistence of strains A and B in terms of parameters ρ and ν for different lags between the appearance time of the strains. The two strains coexist in the region above the solid line. One strain is excluded in the region below the line, (a) The sequence of epidemics in season $T+1$ and $T+2$ is $(AB \rightarrow AB \rightarrow BA)$. (b) The sequence of epidemics in season $T+1$ and $T+2$ is $(AB \rightarrow BA \rightarrow AB)$.

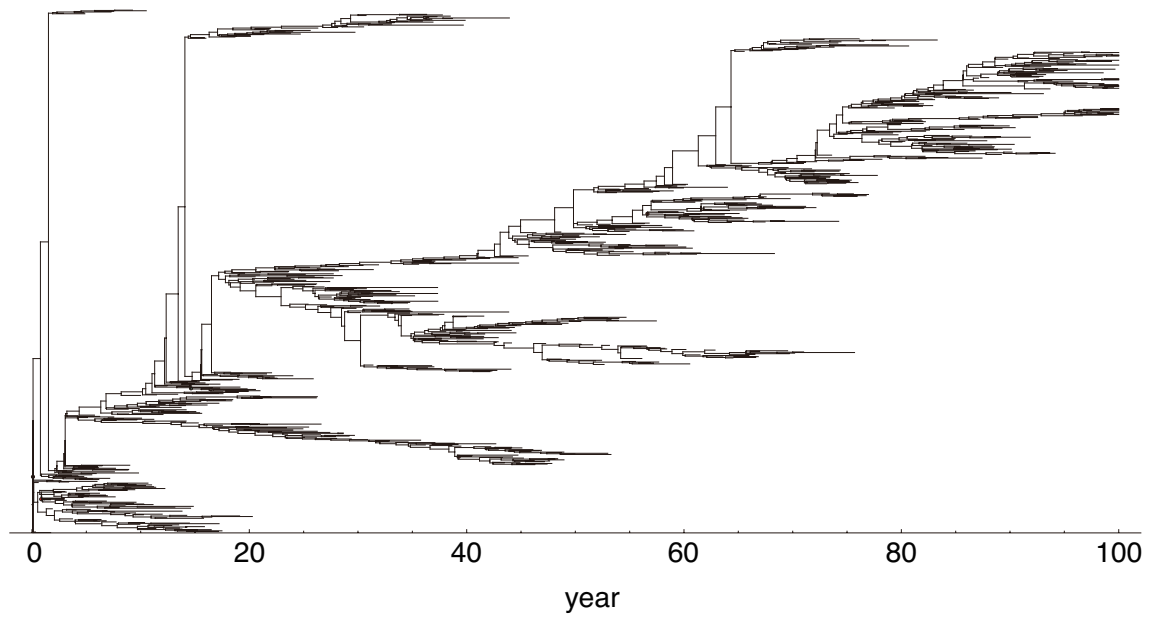


Figure 7: Phylogenetic tree resulting from a 100 year simulation using the individual based model. Strains that caused less than 2000 infections and did not produce mutants have been excluded. Each branch shows a strain created by one site mutation. Basic reproductive ratio ρ is 3.9968, host population size is 10^5 , mutation rate $\mu = 0.0005$, mean duration of protection due to temporary non-specific immunity is $1/5$ year. During temporary immunity hosts cannot be infected by any strain. Initially all of the host population is susceptible to all strains except for 10 hosts infected with the same strain.

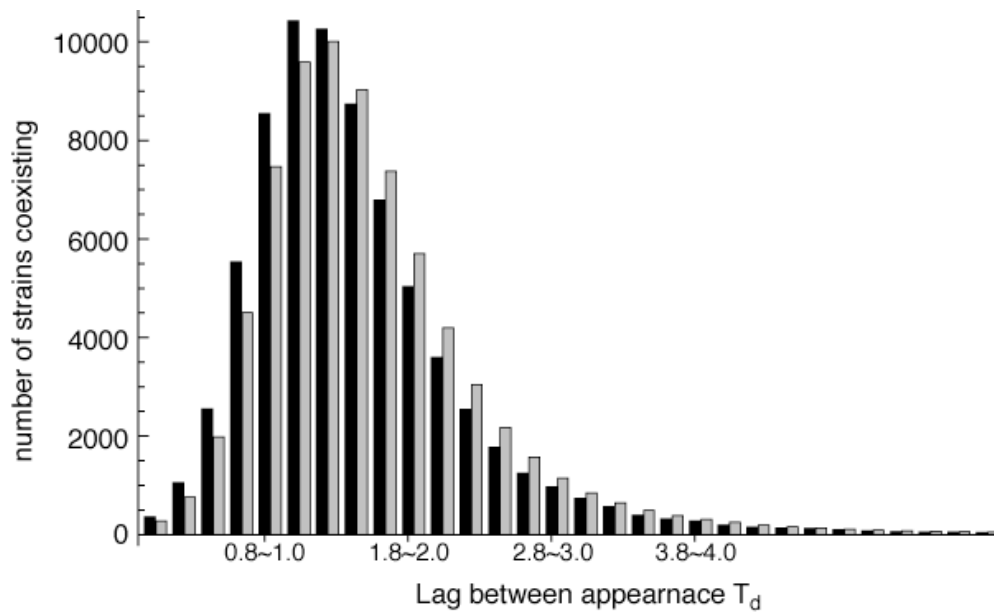


Figure 8: The number of progeny strains that emerged the relative lag T_d after the appearance of their ancestors, and coexisted with the ancestors until they leave the next progeny strain, where the lags are scaled in units of the duration of the ancestral strain epidemic. Black bars: the observed distribution from the individual based model; gray bars: the theoretical distribution when there is no correlation between the probability of coexistence and the appearance time. Earlier emerging progeny strains are more likely to coexist with their ancestral strains than random expectation. See text for the definition of the coexistence between ancestral and progeny strains.

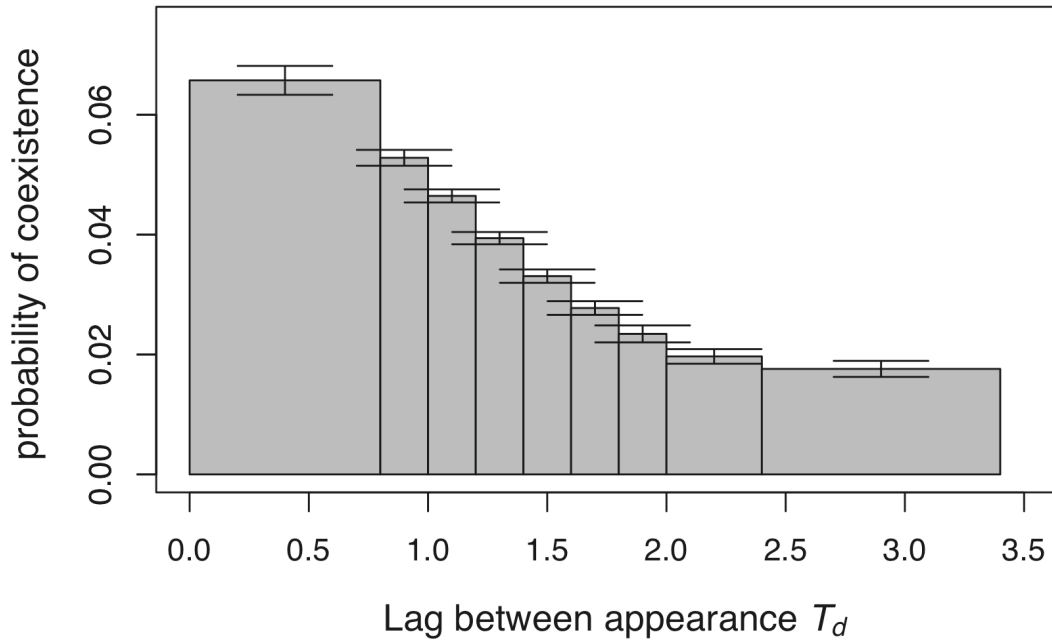


Figure 9: Probability that a mutant strain coexists with its progenitor as a function of the lag between the appearance of the two strains, T_d . Error bars denote standard error. As in Figure 8, this shows that earlier emerging progeny strains are more likely to coexist with their ancestral strains.

(k, a, b, f_a, f_b)	A		B	
	not infected by B $\tau_{k,a,b,f_a,f_b}^{A,n}$	infected by B $\tilde{\tau}_{k,a,b,f_a,f_b}^{A,n}$	not infected by A $\tau_{k,a,b,f_a,f_b}^{B,n}$	infected by A $\tilde{\tau}_{k,a,b,f_a,f_b}^{B,n}$
$(0,0,0,0,0)$	1	$1-\alpha^{2n+2}$	1	$1-\alpha^{2n+2}$
$(0,+,0,+,0)$	$1-\alpha^a$	$1-\alpha^a$	$1-\alpha^{2n+2-f_a}$	$1-\alpha^{2n+2-f_a}$
$(0,0,+,0,+)$	$1-\alpha^{2n+2-f_b}$	$1-\alpha^{2n+2-f_b}$	$1-\alpha^b$	$1-\alpha^b$
$(0,+,+,+,+)$	$1-\alpha^a$	$1-\alpha^a$	$1-\alpha^b$	$1-\alpha^b$
$(+,0,0,0,0)$	$1-\alpha^{k+n}$	$1-\alpha^{k+n} \quad (k-1 < n+1)$ $1-\alpha^{2n+2} \quad (k-1 \geq n+1)$	$1-\alpha^{k+n}$	$1-\alpha^{k+n} \quad (k-1 < n+1)$ $1-\alpha^{2n+2} \quad (k-1 \geq n+1)$
$(+,+,0,+,0)$	$1-\alpha^a$	$1-\alpha^a$	$1-\alpha^{k+n} \quad (k-1 < n+1-f_a)$ $1-\alpha^{2n+2-f_a} \quad (k-1 \geq n+1-f_a)$	$1-\alpha^{k+n} \quad (k-1 < n+1-f_a)$ $1-\alpha^{2n+2-f_a} \quad (k-1 \geq n+1-f_a)$
$(+,0,+,0,+)$	$1-\alpha^{k+n} \quad (k-1 < n+1-f_b)$ $1-\alpha^{2n+2-f_b} \quad (k-1 \geq n+1-f_b)$	$1-\alpha^{k+n} \quad (k-1 < n+1-f_b)$ $1-\alpha^{2n+2-f_b} \quad (k-1 \geq n+1-f_b)$	$1-\alpha^b$	$1-\alpha^b$
$(+,+,+,+,+)$	$1-\alpha^a$	$1-\alpha^a$	$1-\alpha^b$	$1-\alpha^b$

Table 1: Relationship between infection history and strength of partial cross-immunity (τ) in season $T+n$. “(not) infected by A(B)” indicates whether or not the host is infected with strain A(B) in the current season. Suffix k indicates how many seasons before T the most recent infection occurred. Suffix a indicates how many seasons before the present, but after season T , the most recent infection with strain A occurred. Suffix f_a denotes how many seasons before the present, but after T , the first infection with strain A occurred. Suffixes b and f_b are similarly defined for strain B.