Competitive evolution of H1N1 and H3N2 influenza viruses in the United States

Chaiwat Wilasang[1], Pikkanet Suttirat[1], Anuwat Wiratsudakul[2], Sudarat Chadsuthi[3,4], Charin Modchang[1,5,6,*]

[1]Biophysics Group, Department of Physics, Faculty of Science, Mahidol University, Bangkok 10400, Thailand.

[2]Department of Clinical Sciences and Public Health, and the Monitoring and Surveillance Center for Zoonotic Diseases in Wildlife and Exotic Animals, Faculty of Veterinary Science, Mahidol University, Nakhon Pathom 73170, Thailand.

[3]Department of Physics, Faculty of Science, Naresuan University, Phitsanulok 65000, Thailand.

[4]Research Center for Academic Excellence in Applied Physics, Faculty of Science, Naresuan University, Phitsanulok 65000, Thailand

[5]Centre of Excellence in Mathematics, CHE, Bangkok 10400, Thailand.

[6]Thailand Center of Excellence in Physics, Commission on Higher Education, Bangkok 10400, Thailand.

* Corresponding author

Email: charin.mod@mahidol.edu

Abstract

Seasonal influenza causes vast public health and economic impact globally. The prevention and control of the annual epidemics remain a challenge due to the antigenic evolution of the viruses. Here, we presented a novel modeling framework based on changes in amino acid sequences and relevant epidemiological data to retrospectively investigate the competitive evolution and transmission of H1N1 and H3N2 influenza viruses in the United States during October 2002 and April 2019. To do so, we estimated the time-varying disease transmission rate from the reported influenza cases and the time-varying evolutionary rate of the viruses from the changes in amino acid sequences. By incorporating these time-varying rates into the transmission models, we found that the models could accurately capture the evolutionary transmission dynamics of influenza viruses in the United States. Our modeling results also showed that models incorporating evolutionary change of the virus could provide better modeling performance suggesting the critical role of the evolutionary change of virus on the disease transmission.

Keywords: H1N1, H3N2, Influenza virus, Evolutionary dynamics, Transmission dynamics.

Introduction

The spread of the influenza virus is a major public health threat that leads to mortality, hospitalization, and an economic impact. Each year, there are one billion influenza cases globally, including 3 to 5 million cases of severe illness[1,2] and at least 200,000 to 600,000 recorded deaths[3-6]. In the US, seasonal influenza accounts for more than tens of thousands of deaths during annual epidemics[7]. Moreover, the economic burden of seasonal influenza in the US has been estimated at more than ten billion dollars in healthcare and social costs[7]. The influenza epidemics recur annually and are very challenging to control. These recurrent epidemics are usually governed by a complex interaction of evolutionary and transmission dynamics of influenza viruses[1,6]. A better understanding of the evolutionary transmission dynamics of the virus could provide insights to mitigate the burden of seasonal influenza.

Influenza is caused by a rapidly mutating virus, which is an enveloped negative-strand RNA virus belonging to the family of *Orthomyxoviridae*[6,8,9]. The influenza viruses are classified into types A–D, where influenza type A is the most pathogenic[8]. Among influenza type A, different viral subtypes have been continuously reported. However, the subtype A(H3N2) and A(H1N1) are commonly found during epidemics each year[1]. In the US, these two influenza subtypes are reported almost every winter, usually from October to April. Seasonality is known to affect the transmission dynamics of the influenza virus[10-14]. A change in temperature and humidity throughout the year could profoundly affect the survival of these viruses. For example, in winter, the cold and dry weather can accelerate the growth of the virus and facilitate viral transmisison[15].

Vaccination is one of the most effective interventions used to mitigate the seasonal influenza epidemics[16,17]. Nonetheless, the annual vaccination may fail to fully protect vaccinated individuals due to antigenic changes of the viruses. Influenza viruses exhibit high genetic variability in which the viruses continually undergo an antigenic change through natural selection. For instance, the changes in the surface protein, called hemagglutinin (HA), help the viruses to escape from immune attacks[6,18,19]. The primary target of immunity against influenza is the HA proteins[7,20]. By changing the properties of antigenic sites in the HA proteins, the immune system may not recognize the virus and allows it to infect host cells[6,18,20,21]. Moreover, a high mutation rate has been observed in several studies due to the short generation times and large population

sizes of the viruses [22-24]. These evolutionary processes are believed to be a major driver of seasonal influenza transmission[20].

In the US, influenza subtype A(H3N2) emerges and spreads more frequently than A(H1N1). Nevertheless, the two subtypes seem to be competitive for the susceptible hosts. Each year, the subtype virus that gets the highest fitness from mutation may spread predominantly in the population. In general, individuals infected with a single influenza strain will acquire immunity against that certain strain[16,22]. Partial or cross-immunity to other subtypes is also observed[15,25-29]. However, the level of cross-immunity is weakening with an increasing antigenic change of the influenza virus[2]. Therefore, co-circulation of the two influenza subtypes may result in cross-immunity.

Mathematical models, specifically the compartmental frameworks, have been widely used to investigate the transmission dynamics of influenza viruses. Most of the previous studies focused on the impact of seasonality on influenza transmission dynamics[12,13,30-33]. Subsequently, the evolution of these viruses has increasingly been studied as an important source of disease burden[6]. For instance, in 2014, J. B. Axelsen and co-workers developed a model to examine the role of climatic drivers and viral evolution on influenza dynamics[30]. The model also integrated an antigenic evolution process of the influenza virus by adding a new strain pulse to the year that drastic antigenic change occurs. However, the model did not consider the evolutionary dynamics of the influenza virus. Later, in 2018, X. Du et al. developed advances in mathematical modeling to incorporate such evolutionary changes. Their proposed model was designed specifically for seasonal H3N2 influenza to predict influenza incidence in the US[1]. Incorporating information on evolutionary change into transmission models can accurately improve epidemiological forecasts. Although such models may fit data accurately, their model does not consider the competition between different subtypes of the virus. Indeed, the co-circulation and co-transmission of influenza subtypes may result in different transmission characteristics[15,34].

In this study, we proposed epidemiological models for seasonal influenza A(H3N2) and A(H1N1) viruses that incorporate the evolutionary change of the viruses. The models integrated the data on the changes of amino acid sequences of HA proteins in epitope sites to the transmission models. To do so, we firstly estimated the time-varying evolutionary rate of each influenza subtype using the sequence data. We then incorporated this evolutionary rate into the transmission models

to investigate the transmission dynamics of seasonal influenza in the US. We utilized long-term influenza surveillance data[35] from October 2002 to April 2019 in the US to fit the models. Finally, we demonstrated that evolutionary changes of the virus could play an important role in changing seasonal influenza transmission dynamics at the population level.

Materials and methods

Influenza surveillance data

The influenza surveillance data used in this study were obtained from the FluView of the US Centers for Disease Control and Prevention (CDC)[35]. The US CDC collects seasonal influenza activity data from both the US World Health Organization (WHO) Collaborating Laboratories and National Respiratory and Enteric Virus Surveillance System (NREVSS) laboratories. The US WHO and NREVSS laboratories report the total number of weekly respiratory specimens tested and the number of weekly influenza-like illness (ILI) cases that tested positive for influenza types A and B to the US CDC. Besides, the US WHO collaborating laboratories also report influenza subtype A(H1N2) and A(H3N2). An ILI case was defined as a patient with a fever greater than 38 C° and a cough or sore throat[36,37]. Following X. Du et al.,[1] the weekly influenza cases were computed as the product of the ILI incidence rate, the proportion of ILI samples that tested positive for influenza, the subtype proportion, and the US population size. The population data were downloaded from the US Census Bureau[38]. The monthly influenza cases from October 2002 to April 2019 were obtained by aggregating the weekly data.

Influenza amino acid sequences data

Amino acid sequences of influenza A(H3N2) and A(H1N1) were downloaded from the Global Initiative on Sharing Avian Influenza Data (GISAID)[39,40]. To analyze the HA protein sequences, whole sequences were aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE)[41,42]. The MUSCLE program uses FASTA files containing amino acid sequences as an input and provides the aligned amino acid sequences. Unknown amino acids were replaced by gaps. Aligned sequences were then edited using ClustalW, and the incomplete

sequences were manually removed[43,44]. Since the major target of immunity against influenza is located in the epitope sites of the HA proteins[6,7,20], we used only the amino acid sequences encoding epitope sites of these proteins to analyze the antigenic change of the virus.

Evolutionary transmission model with no competition

To investigate the evolutionary transmission dynamics of influenza, we analyzed two transmission models that explicitly integrate the virus evolutionary dynamics (**Figure 1**). **Figure 1(A)** illustrates the evolutionary transmission model with no competition. In this model, the population was divided into three epidemiological classes, namely, susceptible, infectious, and recovered. The susceptible individuals (S) can be infected if they contact an infectious individual. After being infected, susceptible individuals move to the infectious class at rate $\beta(t)I/N$, where $\beta(t)$ is a time-dependent transmission rate, and N is the total population size. The infectious individuals (I) recover from the disease after the infectious period ($1/\gamma$) and then are moved to the recovered class. Individuals in the recovered class are immune to the virus and cannot be infected. The model also integrates an effect of the virus antigenic change represented by an evolutionary rate ($\varepsilon(t)$). We assumed that recovered individuals (R) could become susceptible again at rate $\varepsilon(t)R$. The evolutionary transmission dynamics of influenza are described by the following ordinary differential equations (ODEs):

$$\frac{dS}{dt} = \mu N - \beta(t)S\left(\frac{I}{N}\right) + \varepsilon(t)R - \mu S, \tag{1}$$

$$\frac{dI}{dt} = \beta(t)S\left(\frac{I}{N}\right) - \gamma I - \mu I, \tag{2}$$

$$\frac{dR}{dt} = \gamma I - \varepsilon(t)R - \mu R, \tag{3}$$

$$\frac{dC}{dt} = \beta(t)S\left(\frac{I}{N}\right), \tag{4}$$

where S, I, R, and C represent the number of susceptible, infectious, recovered, and cumulative infected individuals, respectively. In this model, we assumed that the newborns are susceptible, and the number of natural births and deaths are balanced at a rate μ so that the total population size

is constant. The parameters descriptions and the values used in the model are summarized in **Table 1**.

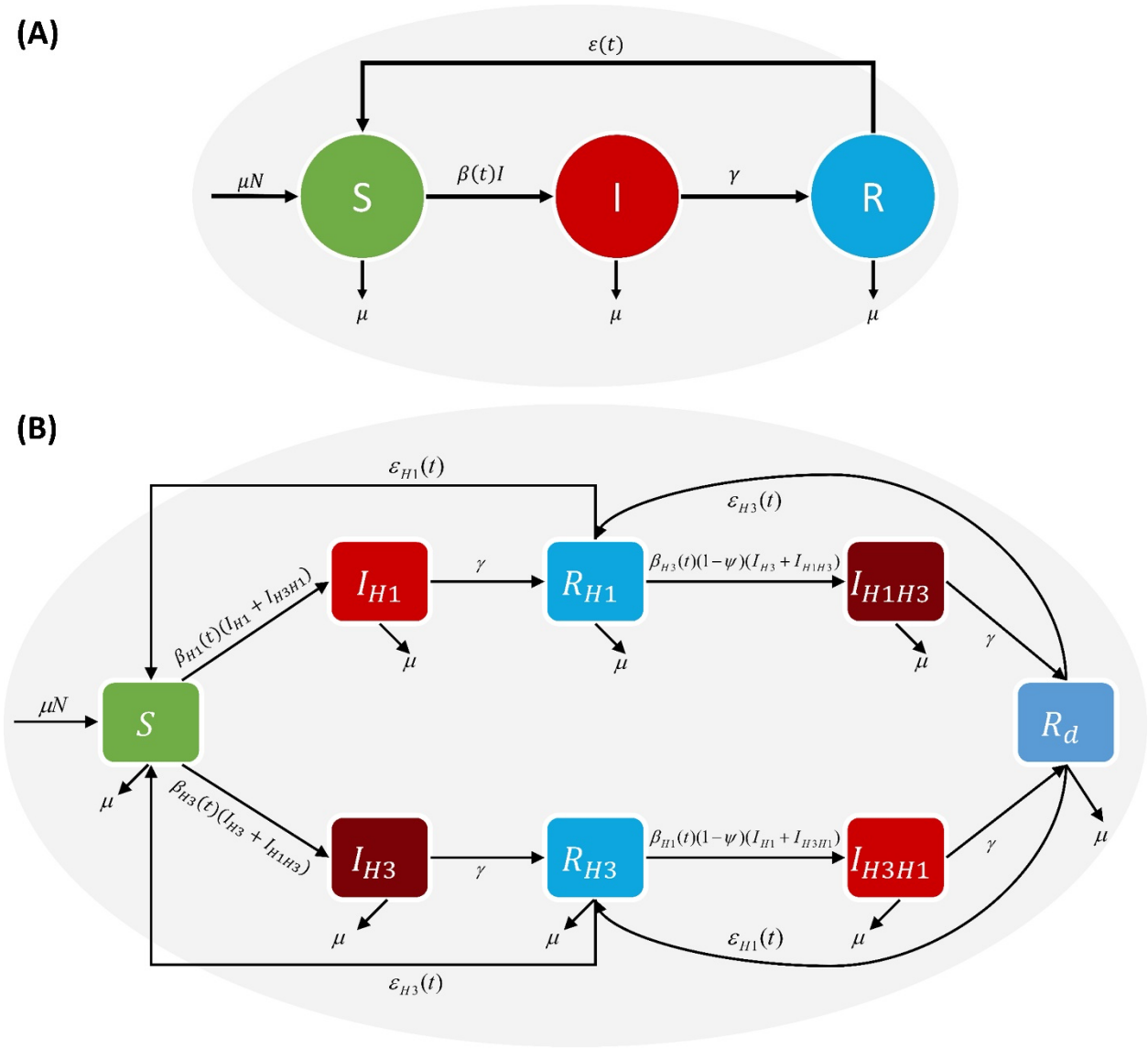


Figure 1. Schematic of the evolutionary transmission model. (A) The flow diagram of the evolutionary transmission model of the A(H1N1) and A(H3N2) influenza viruses with no competition and **(B)** with the selective competition.

Table 1. Model parameters and their values.

Parameter	Definition	Value	References
$\boldsymbol{\beta_i(t)}$	Monthly transmission rate of influenza subtype *i*	[0.001-10]	Estimated
$\boldsymbol{\gamma}$	Recovery rate	1/5 per day	[45-47]
$\boldsymbol{\mu}$	Natural birth and death rate	1/67 per year	[3]
$\boldsymbol{\varepsilon_0}$	Waning rate of immunity	1/20 per year	Estimated
$\boldsymbol{\Psi}$	Cross-immunity induced by primary infection	0.7	[15,48,49]
$\boldsymbol{\alpha}$	The steepness of antigenic change	2	Assumed
$\boldsymbol{N}$	Population size	3.25×10^8	[38]

Evolutionary transmission model with the selective competition

Several influenza virus strains can co-circulate in a population leading to selective competition in susceptible hosts[7,22,25,50,51]. The co-circulation of influenza viruses could result in cross-immunity and mediate the ecological competition between different strains. Consequently, for the modeling pursued here, we constructed an evolutionary transmission model considering cross-immunity and selective competition of the A(H3N2) and A(H1N1) influenza viruses. The schematic for the model with selective competition is shown in **Figure 1(B)**. Similar to the model with no competition, the population was categorized into three different groups, namely. However, each group of individuals here can be infected with either A(H1N1) or A(H3N2), indicated by the subscripts H1 and H3, respectively. Individuals recently recovered from the infection of one stain develop partial protection and hence are less susceptible to another strain. In our equations, we use the order of subscripts to indicate the order of infection. For example, I_{H3H1} represents a group of individuals who had recently recovered from the infection of A(H3N2) but is currently infected with A(H1N1). R_d represents a group of individuals that are immune to the two virus strains. The evolutionary transmission model with selective competition can be described by the following system of ODEs:

$$\frac{dS}{dt} = \mu N - \beta_{H1}(t)S(t)(I_{H1} + I_{H3H1})/N - \beta_{H3}(t)S(I_{H3} + I_{H1H3}) + \varepsilon_{H1}(t)R_{H1} + \varepsilon_{H3}(t)R_{H3} - \mu S, \quad (5)$$

$$\frac{dI_{H1}}{dt} = \beta_{H1}(t)S(I_{H1} + I_{H3H1})/N - \gamma I_{H1} - \mu I_{H1}, \quad (6)$$

$$\frac{dI_{H3}}{dt} = \beta_{H3}(t)S(I_{H3} + I_{H1H3})/N - \gamma I_{H3} - \mu I_{H3}, \quad (7)$$

$$\frac{dR_{H1}}{dt} = \gamma I_{H1} - \varepsilon_{H1}(t)R_{H1} - \beta_{H3}(t)(1-\Psi)R_{H1}(I_{H3} + I_{H1H3})/N + \varepsilon_{H3}(t)R_d - \mu R_{H1}, \quad (8)$$

$$\frac{dR_{H3}}{dt} = \gamma I_{H3} - \varepsilon_{H3}(t)R_{H3} - \beta_{H1}(t)(1-\Psi)R_{H3}(I_{H1} + I_{H3H1})/N + \varepsilon_{H1}(t)R_d - \mu R_{H3}, \quad (9)$$

$$\frac{dI_{H1H3}}{dt} = \beta_{H3}(t)(1-\Psi)R_{H1}(I_{H3} + I_{H1H3})/N - \gamma I_{H1H3} - \mu I_{H1H3}, \quad (10)$$

$$\frac{dI_{H3H1}}{dt} = \beta_{H1}(t)(1-\Psi)R_{H3}(I_{H1} + I_{H3H1})/N - \gamma I_{H3H1} - \mu I_{H3H1}, \quad (11)$$

$$\frac{dR_d}{dt} = \gamma I_{H1H3} + \gamma I_{H3H1} - \varepsilon_{H3}(t)R_d - \varepsilon_{H1}(t)R_d - \mu R_d, \quad (12)$$

where $\beta_{H1}(t)$ and $\beta_{H3}(t)$ is the time-varying transmission rate of infectious individuals who are currently infected with A(H1N1) and A(H3N2), respectively. Ψ is the cross-immunity between the two influenza subtypes. ε_{H1} and ε_{H3} are the evolutionary rate of the A(H1N1) and A(H3N2) strain, respectively.

To generate the long-term seasonal influenza dynamics, the infection was initiated by randomly selected seed infections and run for a period of 100 years to allow the system to reach equilibrium. We then used the averaged population in each compartment as the initial population of the models. For the simulation process, we used the Euler method to solve the sets of ODEs implemented in MATLAB R2019b. The model parameters are summarized in **Table 1**.

Estimating the time-varying evolutionary rate

Seasonal influenza viruses can better survive and spread in a specific environment, especially in winter[30,52]. The recurrence of seasonal influenza epidemics thus frequently occurs between October and March. To consider the effects of seasonality, the period from October 1 to March 31 was assigned as the high transmission season (HTS), while the remaining period was the low transmission season (LTS)[1]. We compared the amino acid sequences of both influenza subtypes between two consecutive seasons to estimate the evolutionary rate.

To investigate the evolutionary dynamics of the seasonal influenza virus, we used amino acid changes in the epitope as a measure of the antigenic distance between the two strains. Once the virus infects the host cells, the antibody attacks the antigen by binding to HA proteins located on the viral surface[18]. This part of an antigen memorized by the immune system is called epitope[3,53]. The mutation within epitope regions could, thus, weaken the immune system[19]. Consequently, the virus can enter and infect the host cell. Epitope regions are categorized differently according to influenza virus subtypes[6,54-57]. The amino acid residues of epitope regions are summarized in **Table S1** and **Table S2** in the Supplementary Information.

To estimate the antigenic change among influenza viruses, we considered aligned amino acid sequences as character strings and computed the number of amino acid differences within all five epitopes, known as the Hamming distance[58]. The Hamming distances were calculated by comparing the viral sequences in the current and previous seasons. However, as the numbers of viral sequences recorded in each season are different, we first sampled 10^6 random pairs of sequences between seasons to avoid sampling biases. Subsequently, we calculated the Hamming distance for each pair of amino acid sequences. We then computed $P_{all-epitopes}$, which is a fraction of the number of amino acid differences within all five epitopes and the total number of amino acids presented[59].

We incorporated the evolutionary rate into the transmission model to investigate how antigenic change affects transmission dynamics. The time-varying evolutionary rate, $\varepsilon(t)$, is given by:

$$\varepsilon(t) = \varepsilon_0 + c\left[\frac{P_{all-epitope}}{1-P_{all-epitope}}\right]^{\alpha}, \tag{13}$$

where ε_0 represents the naturally waning rate of immunity. Parameters α and c were estimated such that the resulting cumulative cases derived from the simulations match the influenza surveillance data. We varied the evolutionary rate and simulated the H1N1 influenza model before the pandemic season of 2009–2010. The values of $\varepsilon(t)$ that the model can search for are uniformly distributed in the interval [0, 1] with 0.001 resolution. We recorded only the values of the evolutionary rate that yield the best simulation results. Subsequently, we calculated the constant parameter c according to Eq. (13). Simultaneously, we varied α; the value of steepness changes.

The estimated parameters are $\alpha = 2$ and $c = 5.9911\times10^{-2}$ as they minimize the difference between simulated and reported data.

Estimating the time-varying transmission rate

In this study, we employed the parameter estimation method developed by P. Suparit et al.[60] to estimate the time series of monthly transmission rates. The monthly influenza cases in the US starting from October 2002 to April 2019 were used to estimate the time series of monthly transmission rates. Briefly, we divided the model simulation into several consecutive monthly time window simulations. We varied $\beta(t)$ ranging from 0.001 to 10 with the resolution of 0.01 and searched for the transmission rate that provides the best fit between the number of monthly simulated cases (C_{sim}) and the number of monthly reported cases (C_{rep}). The values of the monthly transmission rates that minimizes $|C_{sim} - C_{rep}|$ were recorded in the time series of monthly transmission rates. We also recorded the number of individuals in each epidemiological class for the current monthly time window simulation. Subsequently, we used the number of individuals in the previous monthly time window as an initial condition for the next monthly time window simulation. The process was repeated until the complete time series of the monthly transmission rates were obtained.

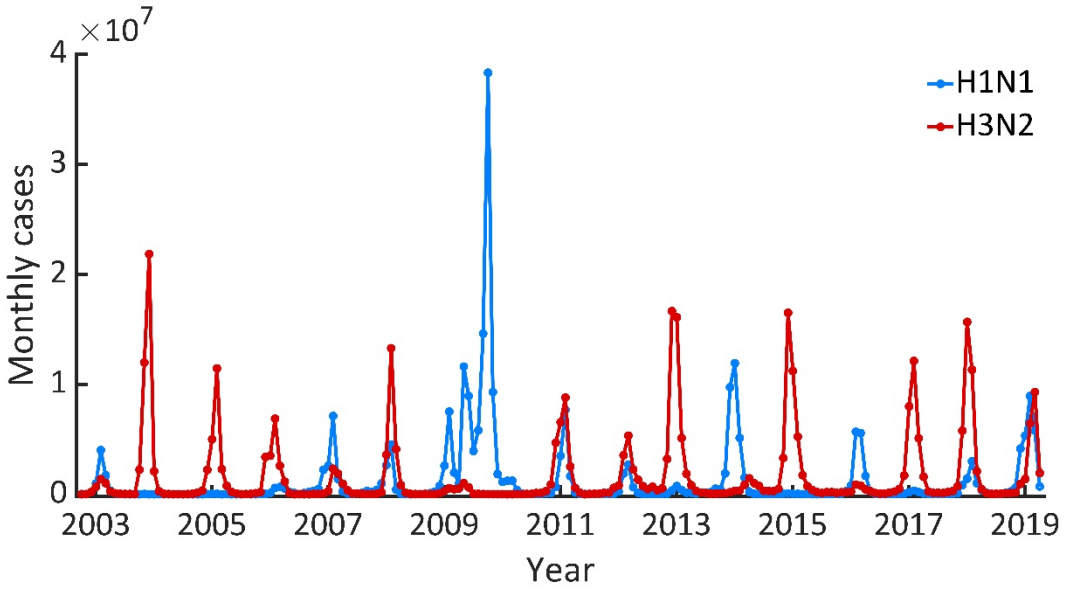
Results

Seasonal patterns of influenza epidemics in the US

The monthly influenza cases from October 2002 to April 2019 (~17 years) reported in the United States are shown in **Figure 2** (see also **Figure S1**). The data shows the incidence of monthly influenza of two subtypes: subtype A(H3N2) and subtype A(H1N1). The outbreaks of influenza A (H3N2) and A (H1N1) usually appear between October and March. The numbers of cases are usually lower in the summer seasons (April to September). We, therefore, classified the period from 1st October to 31st March as a high transmission season (HTS) and the period from 1st April to 30th September as a low transmission season (LTS). Thus, there were 17 HTSs and 16 LTSs in the time series. As illustrated in **Figure 2**, both subtypes of influenza viruses have co-circulated

261 throughout the study period. Aside from the apparent annual cycles, epidemics of A(H3N2) and
262 A(H1N1) usually occur with different magnitudes, especially during the year 2009 when the novel
263 2009 A(H1N1) influenza emerged [47,61,62].

264



265

266 **Figure 2. Influenza epidemics in the US.** The monthly cases of influenza A(H1N1) and A(H3N2)
267 in the US between October 2002 and April 2019.

268

269 Evolutionary rates of influenza virus

270 The recurrent influenza epidemics are a combined outcome of viral antigenic changes,
271 interactions among co-circulating influenza viruses, transmission, and host immune recognition
272 against the virus[6]. When influenza viruses undergo evolution, the immune system may not be able
273 to recognize the viruses, allowing the viruses to infect host cells. Thus, previously infected
274 individuals may eventually become susceptible to the viruses again[19]. To quantify the evolutionary
275 rate, we calculated $P_{all-epitopes}$, which is a measure of the antigenic change between two influenza
276 virus strains[4,59] (see Material and Methods for more details). The average $P_{all-epitopes}$ for both
277 influenza A(H3N2) and A(H1N1) are shown in **Figure 3 (A)** and **(B)**, respectively. $P_{all-epitopes}$
278 were also compared with the monthly influenza cases. According to the results shown in **Figure**

3(A) and **(B)**, we found that the large epidemic curves usually occur after a significant evolutionary change of influenza viruses. Specifically, the antigenic change of influenza A(H1N1) was around 80% in the 2008-2009 HTS, which resulted in the subsequent surge of the 2009 influenza pandemic (**Figure 3(B)**). This result suggested that $P_{all-epitopes}$ might be a potential early warning for a pandemic. We also found that the average $P_{all-epitopes}$ of A(H3N2) is less fluctuate than that of A(H1N1). The corresponding evolutionary rates of influenza A(H1N1) and A(H3N2) are illustrated in **Figure 3(C)** and **(D)**.

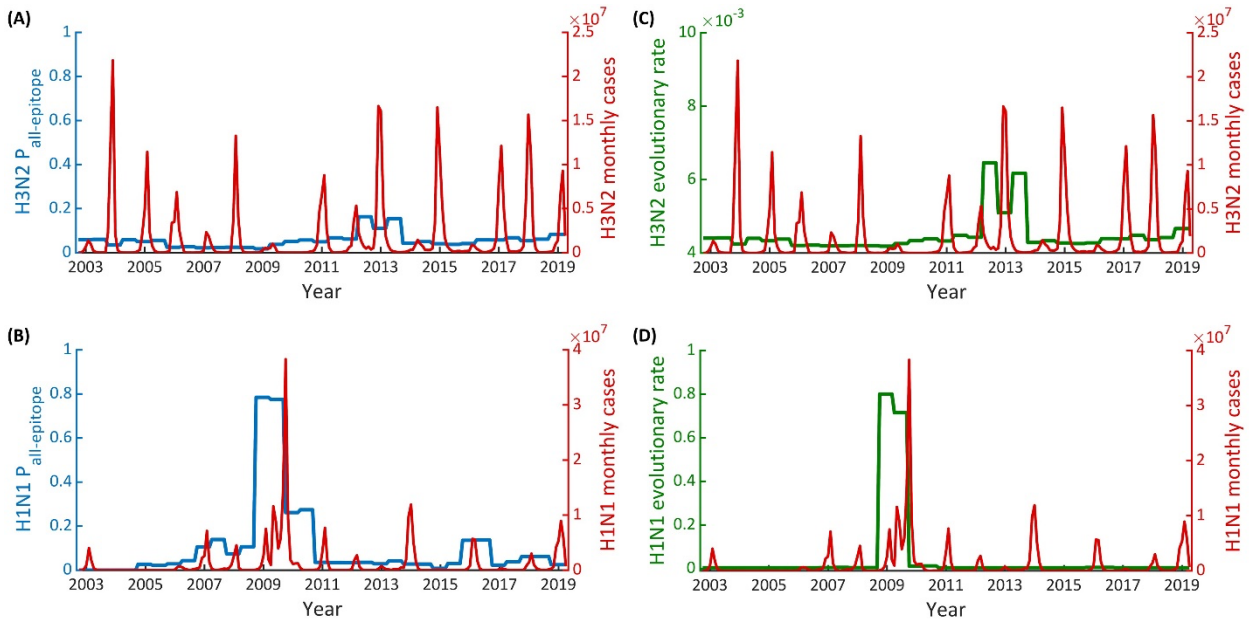


Figure 3. Average monthly $P_{all-epitopes}$ and evolutionary rates. $P_{all-epitopes}$ of influenza **(A)** A(H3N2) virus, and **(B)** influenza A(H1N1) virus, **(C)** the evolutionary rate of influenza A(H3N2) virus, and **(D)** influenza A(H1N1) virus.

Seasonal influenza transmission dynamics in the US

We found that the trends of monthly influenza cases generated from the model agree well with the surveillance data with R-square = 0.9989 (**Figure 4**). During the pandemic season 2008-2009, for example, the total number of influenza cases was predicted at 39.03 million, which is close to the number of reported cases at 38.3 million. We also employed the model to reconstruct

the transmission dynamics of influenza A(H3N2) during the same period in the US. We found that the model also provided a good fit to the observed data with R-square = 0.9989 (**Figure S2**).

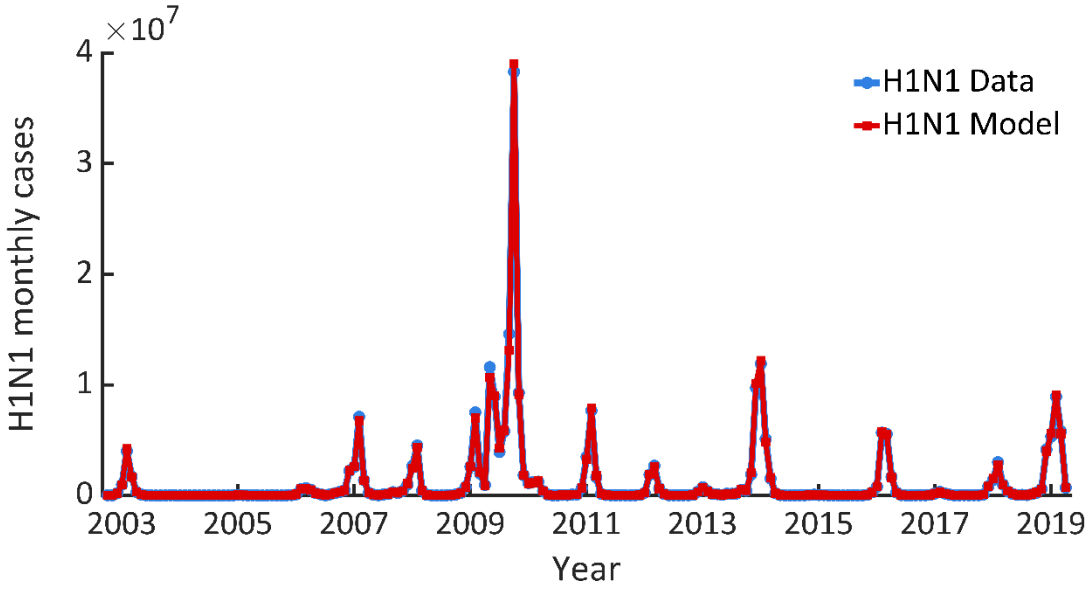


Figure 4. The influenza A(H1N1) transmission dynamics. The model provided a good fit to the 16-year observed data with R-square = 0.9989.

Effect of the evolutionary change of the influenza virus

We found that the estimated transmission rates from the two models are different. The estimated monthly transmission rates that provide the best fit of the two models are shown in **Figure 5**. The results illustrate that the transmission rates estimated from the model without the evolutionary dynamics increased after the major changes of influenza viruses in 2009. In contrast, there is no change in the post-2009 trend of the transmission rates estimated from the model incorporating the evolutionary dynamics. The transmission rate fluctuates between 0.11 and 0.31 per month during the years 2009 and 2019. There is also an approximately 2-fold increase in oscillation amplitude of transmission rates when excluding the evolutionary dynamics in the model. These suggested that the evolutionary change of the virus could affect their transmission dynamics. This effect was also observed in the transmission dynamics of the influenza A(H3N2) virus (see **Figure S3** in the supplementary results).

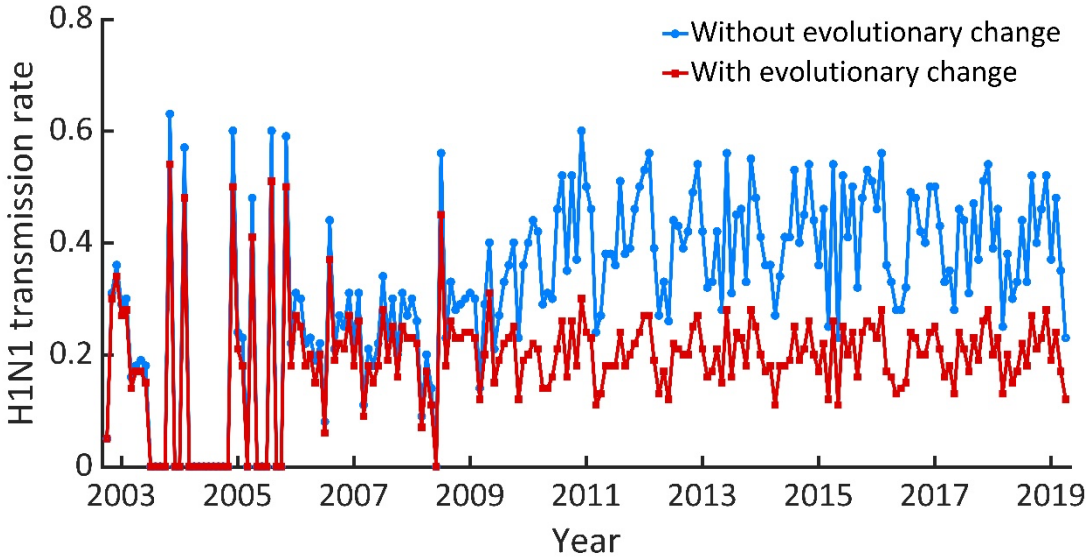


Figure 5. Comparison of H1N1 monthly transmission rates. The time series of monthly transmission rates during October 2002 and April 2019. The blue line shows the monthly transmission rate of the A(H1N1) influenza virus estimated using the model without evolutionary change. Redline denotes the estimated monthly transmission rate of the A(H1N1) influenza virus from the model with the evolutionary change of the virus.

Competitive evolutionary transmission dynamics of the A(H3N2) and A(H1N1) influenza viruses

The infection of a certain influenza subtype in humans results in partial immunity to other subtypes of influenza viruses[22,63]. The empirical evidence suggests that the co-circulation of influenza viruses not only results in cross-immunity but also appears to mediate the ecological competition between the subtypes[7,22,25,50,51]. A competitive model was then constructed to investigate the combined co-evolution and co-transmission dynamics of the A(H3N2) and A(H1N1) influenza viruses. We also incorporated the cross-immunity between the two subtypes into our competitive model. The results illustrated that the models could faithfully reconstruct the influenza epidemics (**Figure S4** and **Figure S5** in the supplementary results). Nevertheless, the competitive model that incorporates the cross-immunity is slightly more accurate in our model validation. For the competitive model incorporating the cross-immunity, the model fit correlated to the data with R-square = 0.9849 for H3N2 influenza cases, and it is 0.9468 in the case of H1N1

influenza cases. For the competitive model without the cross-immunity, we observed a correlation of R-square = 0.9723 and R-square = 0.8905 for H3N2 and H1N1 influenza viruses, respectively. The correlations were driven down partly due to the pandemic season of 2009-2010, in which antigenic evolution of the A(H1N1) influenza virus occurred.

We then further investigated the effects of cross-immunity between the two subtypes. The strength of cross-immunity between the two influenza subtypes is represented by the parameter Ψ. If $\Psi = 0$, the influenza subtypes are completely different antigenically. In contrast, if $\Psi = 1$, the two subtypes are antigenically indistinguishable[25]. **Figure 6** illustrates the effect of the cross-immunity on the transmission rates estimated by the three models. The transmission rate estimated from the competitive model with the cross-immunity $\Psi = 0.7$ was found in the range of 0.01–1.60 per month with an average value of 0.59. In contrast, the transmission rates estimated from the competitive model without cross-immunity ($\Psi = 0.0$) fluctuates in the range of 0.01-0.60 with the average value of 0.20 per month, which is close to the transmission rates estimated from the single strain model (H1N1 model).

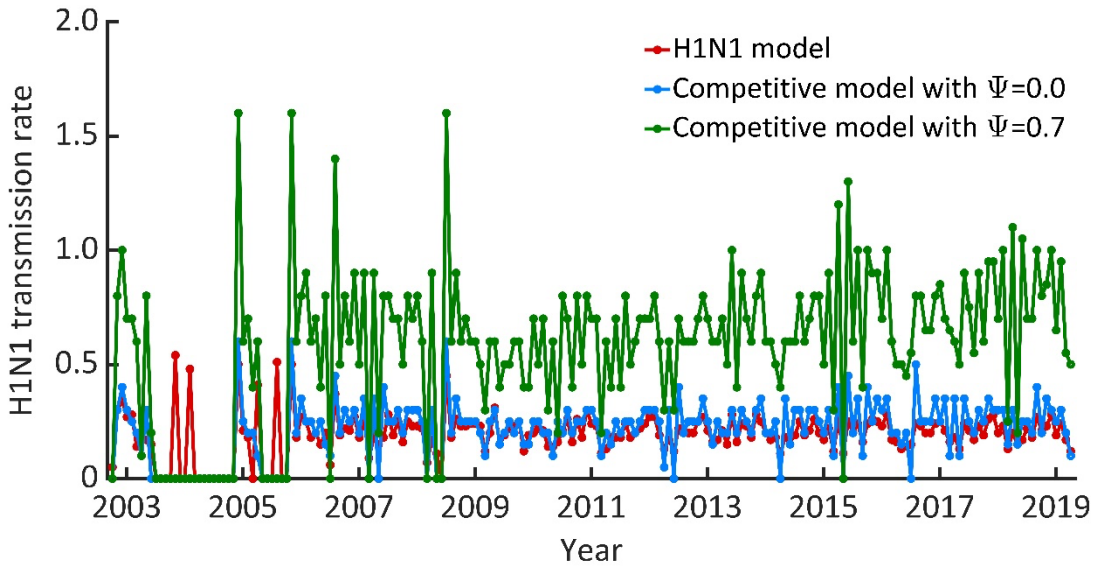


Figure 6. The time series of transmission rate estimated from three different models. The red line shows the monthly transmission rate of the H1N1 influenza virus estimated from the model without evolutionary change of the virus. The blue and green lines display the monthly

transmission rate of the H1N1 influenza virus estimated from the competitive model without cross-immunity ($\Psi = 0.0$) and with the cross-immunity $\Psi = 0.7$, respectively.

Comparing the proportion of immune individuals predicted by each model, we found the fraction of the immune population obtained from the non-competitive model is lower than that of the competitive one (**Figure 7** and **Figure S6(B)** in the supplementary results). Noticeably, at the end of 2009, the immunity levels of influenza A(H1N1) from the three modeling scenarios were predicted at 7.94%, 12.82%, and 20.10%, respectively. Regarding the seroprevalence of influenza-specific antibodies, we found that the competitive model with the cross-immunity provides the proportion of immune population closest to the observed data at 21.0%[64,65].

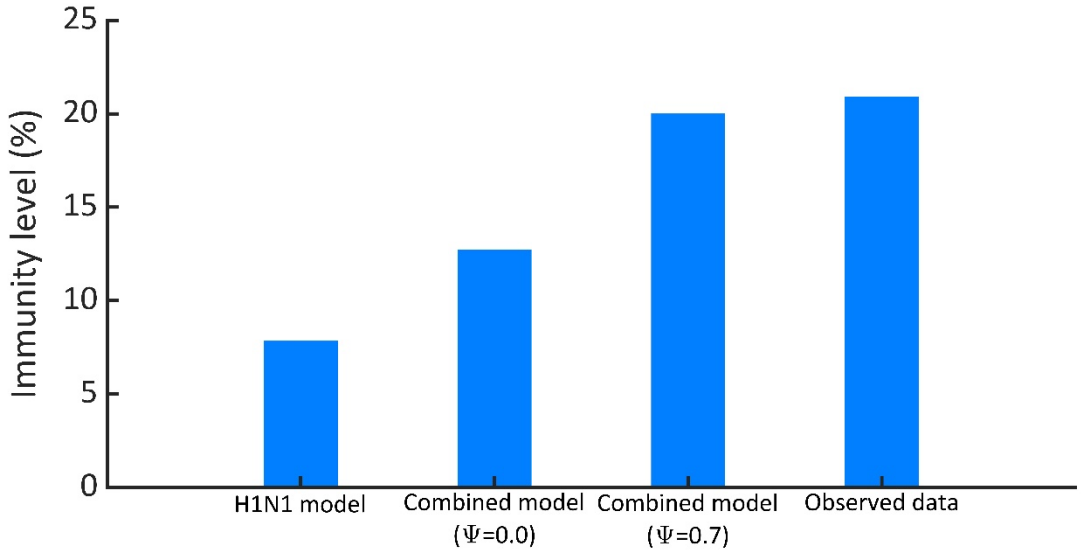


Figure 7. Immunity level of influenza A(H1N1). The immunity levels of influenza A(H1N1) estimated from the H1N1 model, the competitive model without cross-immunity, and the competitive model with Ψ=0.7 are 7.94%, 12.82%, and 20.10%, respectively. The seroprevalence of influenza observed by S. M. Zimmer et al.[64] was approximately 21.0%.

Discussion

Influenza epidemics have been observed annually, resulting in significant morbidity and mortality rates. The recurrence of the influenza epidemic stems from several possibilities[6]. Seasonal factors are known to affect the transmission dynamics of the influenza virus[10-14]. However, one of the most significant factors that drive the annual influenza epidemic is the evolution of the virus[1,19,58]. Thus, the integration of the evolutionary change of the virus into the epidemic model is necessary. To capture the transmission dynamics of the seasonal influenza virus in the US, we proposed mathematical models incorporating the evolutionary change of the virus. The models were employed to reconstruct the long-term influenza epidemics in the US from October 2002 to April 2019. We also compared competitive evolutionary transmission models of influenza A(H1N1) and A(H3N2) viruses with and without cross-immunity.

To investigate the interplay between the epidemic dynamics and the evolutionary dynamics of the influenza virus, we compared the virus evolutionary rate with the monthly influenza incidence (**Figure 3**). The evolutionary rate of influenza A(H3N2) increased a little during the season 2013 (October 2012 to April 2013), which might contribute to the large outbreak in the US in that season, resulting in approximately 30 million influenza illnesses[66]. For influenza A(H1N1), the evolutionary rate drastically increased during October 2008 and April 2009, consequently causing the 2009 A(H1N1) influenza pandemic[67]. Over this period, we observed that the evolutionary rate of influenza A(H1N1) considerably increased from 0.01 to 0.80. Thus, the emerging strain could replace the existing stains rapidly[68].

Several studies have examined the impacts of seasonal factors (e.g., temperature and humidity) on the transmission of seasonal influenza[11-14,25]. Few studies have attempted to bridge the gap between the evolutionary change of the virus and the transmission model[1]. Previously, this was modeled by adding a constant rate to the year that drastic antigenic change occurs to represent the effect of the influenza virus evolution[30]. However, in this work, we intended to link antigenic changes of the influenza virus to disease transmission models. Specifically, we investigated the effect of evolutionary change on seasonal influenza epidemics by incorporating evolutionary rate into the transmission model. We also validated the models by comparing the modeling results with the reported data. The model-fitting results show that the proposed influenza transmission model incorporating evolutionary change provides the best fit between the simulation results and the

403 reported data (R-square=0.9989) (**Figure 5**). Also, our models could reconstruct the observed
404 epidemic dynamics for both high transmission and low transmission seasons for the two subtypes.

405 Furthermore, the evolutionary rate included in the model can describe the transmission
406 dynamics of seasonal influenza. To interpret these findings, we compared the monthly
407 transmission rates estimated from the model with and without evolutionary change of the influenza
408 virus. We then observed significant differences in the transmission rate. The transmission rate
409 estimating from the model incorporating evolutionary rate fluctuates around a certain constant
410 even though there is the drastic change in the genetic sequence of the influenza virus in 2009. In
411 contrast, in the model without evolutionary rate, the fluctuation of the estimated transmission rate
412 is shifted up after the year of pandemics. The increase in transmission rate may result from the
413 large outbreaks in 2009, which drastically depleted the susceptible pool. Consequently, the model
414 compensates by increasing the transmission rate to generate the same number of reported cases as
415 a model incorporating evolutionary change. This means that either human behavior or the viral
416 characteristic has abruptly changed. However, there is no obvious evidence that those factors
417 occurred, unlike the current COVID-19 pandemic[69,70]. Since the transmission rate depends on
418 human behavior, it is unlikely to change after the 2009 pandemic. These results highlight the
419 importance of incorporating the evolutionary rate into the model. The models incorporating
420 evolutionary information of seasonal influenza virus are significantly better than the model that
421 does not. Thus, to accurately reconstruct the dynamics of seasonal influenza transmission, the
422 evolutionary change of the virus should be considered in transmission models.

423 Indeed, people can be infected by multiple strains of the influenza virus. We, therefore,
424 constructed competitive transmission models to investigate the combined co-evolution and co-
425 transmission dynamics of the A(H3N2) and A(H1N1) influenza viruses. Our findings indicate that
426 population susceptibility is associated with the evolutionary change of the virus and the strength
427 of cross-immunity. The evolutionary change of the virus replenishes the pool of susceptible
428 individuals, which increases population susceptibility, whereas cross-immunity can reduce
429 population susceptibility[71]. The antigenic variation of the influenza virus may allow the virus to
430 escape the existing protective immunity and cause new epidemics. This evolutionary change can
431 accommodate the system dynamics, e.g., increased population susceptibility, and allows abrupt
432 increases in population susceptibility in response to epidemic surges. In addition, we also found

433 that there was a competition between influenza subtypes via cross-immunity. Different values of
434 estimated transmission in **Figure 6** illustrate the effect of cross-immunity. These indicate that
435 cross-immunity could modulate the ecological interactions between co-circulating influenza
436 viruses by increasing the transmission rate of the virus. If the cross-immunity level is high, the
437 competition between the two subtypes becomes extensively strict[25]. This finding agrees with the
438 conclusions from other studies[2,25,50,51], which found that cross-immunity among influenza viruses
439 would naturally lead to competition between subtypes. Besides, we also validated our competitive
440 transmission models by comparing the level of immunity from model predictions with
441 seroprevalence data. The competitive transmission model with cross-immunity yields the best
442 simulation result. The model predicted the immunity level at the end of 2009 to be 20.10%, which
443 is similar to other studies[64,65,72-74].

444 Our models made several simplifying assumptions: firstly, we assumed the population is
445 homogeneous mixing, in which all individuals have an equal probability of transmitting the
446 disease. Indeed, as observed in other works[75], models with a homogeneously mixed population
447 usually present a higher attack rate than heterogeneous models, even for the same R_0. Second, we
448 ignore heterogeneity in disease transmissions such as sex, age structure, and geographical region.
449 The individual heterogeneity in infectiousness, e.g., superspreading events, is known to affect the
450 dynamics of disease transmission[76,77]. Moreover, individual heterogeneity in disease transmission
451 can also result in a higher probability of stochastic extinction[77]. Finally, we did not consider human
452 mobility in our model; incorporating these data into the model could improve a more realistic
453 model[78,79]. It is worth stressing that the knowledge of the evolutionary change of the influenza
454 virus is necessary to be investigated. Furthermore, it could also inform vaccine strain selection to
455 prevent yearly epidemics[80,81].

456 In summary, we constructed evolutionary transmission models for investigating seasonal
457 influenza transmission in the US. The models integrate the changes of amino acid sequences of
458 HA proteins in epitope sites and the time-varying disease transmission. We found that the models
459 with information on evolutionary change outperformed the models without it. We also constructed
460 the competitive model to investigate the competitive evolutionary transmission dynamics of the
461 A(H3N2) and A(H1N1) influenza viruses. Our result suggested that the competitive model with
462 cross-immunity is better at capturing the transmission dynamics of seasonal influenza.

Acknowledgments

Chaiwat Wilasang is supported by the Science Achievement Scholarship of Thailand (SAST).

References

1 Du, X., King, A. A., Woods, R. J. & Pascual, M. Evolution-informed forecasting of seasonal influenza A (H3N2). *Science translational medicine* **9** (2017).

2 Yang, W., Lau, E. H. & Cowling, B. J. Dynamic interactions of influenza viruses in Hong Kong during 1998-2018. *PLoS computational biology* **16**, e1007989 (2020).

3 Xu, R. *et al.* Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science* **328**, 357-360 (2010).

4 Gupta, V., Earl, D. J. & Deem, M. W. Quantifying influenza vaccine efficacy and antigenic distance. *Vaccine* **24**, 3881-3888 (2006).

5 Hill, E. M., Petrou, S., De Lusignan, S., Yonova, I. & Keeling, M. J. Seasonal influenza: Modelling approaches to capture immunity propagation. *PLoS computational biology* **15**, e1007096 (2019).

6 Petrova, V. N. & Russell, C. A. The evolution of seasonal influenza viruses. *Nature Reviews Microbiology* **16**, 47 (2018).

7 Zinder, D., Bedford, T., Gupta, S. & Pascual, M. The roles of competition and mutation in shaping antigenic and genetic diversity in influenza. *PLoS Pathog* **9**, e1003104 (2013).

8 Long, J. S., Mistry, B., Haslam, S. M. & Barclay, W. S. Host and viral determinants of influenza A virus species specificity. *Nature Reviews Microbiology* **17**, 67-81 (2019).

9 Maclachlan, N. J. & Dubovi, E. J. *Fenner's veterinary virology*. (Academic press, 2010).

10 Tamerius, J. D. *et al.* Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLoS pathogens* **9**, e1003194 (2013).

11 Steel, J., Palese, P. & Lowen, A. C. Transmission of a 2009 pandemic influenza virus shows a sensitivity to temperature and humidity similar to that of an H3N2 seasonal strain. *Journal of virology* **85**, 1400-1402 (2011).

491 12 Lowen, A. C., Mubareka, S., Steel, J. & Palese, P. Influenza virus transmission is
492 dependent on relative humidity and temperature. *PLoS Pathog* **3**, e151 (2007).
493 13 Lowen, A. C. & Steel, J. Roles of humidity and temperature in shaping influenza
494 seasonality. *Journal of virology* **88**, 7692-7695 (2014).
495 14 Chadsuthi, S., Iamsirithaworn, S., Triampo, W. & Modchang, C. Modeling seasonal
496 influenza transmission and its association with climate factors in Thailand using time-series
497 and ARIMAX analyses. *Computational and mathematical methods in medicine* **2015**
498 (2015).
499 15 Zhang, X.-S. Strain interactions as a mechanism for dominant strain alternation and
500 incidence oscillation in infectious diseases: seasonal influenza as a case study. *PloS one*
501 **10**, e0142170 (2015).
502 16 Ahn, K. W., Kosoy, M. & Chan, K.-S. An approach for modeling cross-immunity of two
503 strains, with application to variants of Bartonella in terms of genetic similarity. *Epidemics*
504 **7**, 7-12 (2014).
505 17 Weir, J. P. & Gruber, M. F. An overview of the regulation of influenza vaccines in the
506 United States. *Influenza and other respiratory viruses* **10**, 354-360 (2016).
507 18 Wiley, D., Wilson, I. & Skehel, J. Structural identification of the antibody-binding sites of
508 Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature*
509 **289**, 373 (1981).
510 19 Ferguson, N. M., Galvani, A. P. & Bush, R. M. Ecological and immunological
511 determinants of influenza evolution. *Nature* **422**, 428 (2003).
512 20 Huang, J.-W. & Yang, J.-M. Changed epitopes drive the antigenic drift for influenza A
513 (H3N2) viruses. *BMC bioinformatics* **12**, 1-10 (2011).
514 21 van de Sandt, C. E. *et al.* Differential recognition of influenza a viruses by M158–66
515 epitope-specific CD8+ T cells is determined by extraepitopic amino acid residues. *Journal*
516 *of virology* **90**, 1009-1022 (2016).
517 22 Zhang, X.-S. *et al.* Co-circulation of influenza A virus strains and emergence of pandemic
518 via reassortment: the role of cross-immunity. *Epidemics* **5**, 20-33 (2013).
519 23 Stern, A. & Andino, R. in *Viral Pathogenesis* 233-240 (Elsevier, 2016).
520 24 Peck, K. M. & Lauring, A. S. Complexities of viral mutation rates. *Journal of virology* **92**
521 (2018).

522 25 Kamo, M. & Sasaki, A. The effect of cross-immunity and seasonal forcing in a multi-strain
523 epidemic model. *Physica D: Nonlinear Phenomena* **165**, 228-241 (2002).
524 26 Laurie, K. L. *et al.* Multiple infections with seasonal influenza A virus induce cross-
525 protective immunity against A (H1N1) pandemic influenza virus in a ferret model. *The*
526 *Journal of infectious diseases* **202**, 1011-1020 (2010).
527 27 Barry, J. M., Viboud, C. & Simonsen, L. Cross-protection between successive waves of
528 the 1918–1919 influenza pandemic: epidemiological evidence from US Army camps and
529 from Britain. *The Journal of infectious diseases* **198**, 1427-1434 (2008).
530 28 Epstein, S. L. & Price, G. E. Cross-protective immunity to influenza A viruses. *Expert*
531 *review of vaccines* **9**, 1325-1341 (2010).
532 29 Bodewes, R. *et al.* Infection of the upper respiratory tract with seasonal influenza A (H3N2)
533 virus induces protective immunity in ferrets against infection with A (H1N1) pdm09 virus
534 after intranasal, but not intratracheal, inoculation. *Journal of virology* **87**, 4293-4301
535 (2013).
536 30 Axelsen, J. B., Yaari, R., Grenfell, B. T. & Stone, L. Multiannual forecasting of seasonal
537 influenza dynamics reveals climatic and evolutionary drivers. *Proceedings of the National*
538 *Academy of Sciences* **111**, 9538-9542 (2014).
539 31 Yaari, R., Katriel, G., Huppert, A., Axelsen, J. & Stone, L. Modelling seasonal influenza:
540 the role of weather and punctuated antigenic drift. *Journal of The Royal Society Interface*
541 **10**, 20130298 (2013).
542 32 Chowell, G., Miller, M. & Viboud, C. Seasonal influenza in the United States, France, and
543 Australia: transmission and prospects for control. *Epidemiology & Infection* **136**, 852-864
544 (2008).
545 33 Shaman, J. & Karspeck, A. Forecasting seasonal outbreaks of influenza. *Proceedings of*
546 *the National Academy of Sciences* **109**, 20425-20430 (2012).
547 34 Zhang, X.-S. & De Angelis, D. Construction of the influenza A virus transmission tree in
548 a college-based population: co-transmission and interactions between influenza A viruses.
549 *BMC infectious diseases* **16**, 1-8 (2015).
550 35 (CDC), C. f. D. C. a. P. *FluView*,
551 <https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html> (2019).

552 36 Fitzner, J. *et al.* Revision of clinical case definitions: influenza-like illness and severe acute
553 respiratory infection. *Bulletin of the World Health Organization* **96**, 122 (2018).
554 37 Kasper, M. R., Wierzba, T. F., Sovann, L., Blair, P. J. & Putnam, S. D. Evaluation of an
555 influenza-like illness case definition in the diagnosis of influenza among patients with
556 acute febrile illness in Cambodia. *BMC infectious diseases* **10**, 1-5 (2010).
557 38 Bureau, U. S. C. *Population and housing unit estimates*,
558 <https://www.census.gov/programs-surveys/popest.html> (2020).
559 39 Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data–from
560 vision to reality. *Eurosurveillance* **22**, 30494 (2017).
561 40 Bogner, P., Capua, I., Lipman, D. J. & Cox, N. J. A global initiative on sharing avian flu
562 data. *Nature* **442**, 981-981 (2006).
563 41 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high
564 throughput. *Nucleic acids research* **32**, 1792-1797 (2004).
565 42 Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and
566 space complexity. *BMC bioinformatics* **5**, 1-19 (2004).
567 43 Thompson, J. D., Gibson, T. J. & Higgins, D. G. Multiple sequence alignment using
568 ClustalW and ClustalX. *Current protocols in bioinformatics*, 2.3. 1-2.3. 22 (2003).
569 44 Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *bioinformatics* **23**, 2947-2948
570 (2007).
571 45 Carrat, F. *et al.* Time lines of infection and disease in human influenza: a review of
572 volunteer challenge studies. *American journal of epidemiology* **167**, 775-785 (2008).
573 46 Nsoesie, E. O., Beckman, R. J. & Marathe, M. V. Sensitivity analysis of an individual-
574 based model for simulation of influenza epidemics. *PloS one* **7**, e45414 (2012).
575 47 Modchang, C., Iamsirithaworn, S., Auewarakul, P. & Triampo, W. A modeling study of
576 school closure to reduce influenza transmission: A case study of an influenza A (H1N1)
577 outbreak in a private Thai school. *Mathematical and Computer Modelling* **55**, 1021-1033
578 (2012).
579 48 Truscott, J. *et al.* Essential epidemiological mechanisms underpinning the transmission
580 dynamics of seasonal influenza. *Journal of The Royal Society Interface* **9**, 304-312 (2012).

581 49 Furuse, Y. & Oshitani, H. Mechanisms of replacement of circulating viruses by seasonal
582 and pandemic influenza A viruses. *International Journal of Infectious Diseases* **51**, 6-14
583 (2016).
584 50 Latorre-Margalef, N. *et al.* Competition between influenza A virus subtypes through
585 heterosubtypic immunity modulates re-infection and antibody dynamics in the mallard
586 duck. *PLoS pathogens* **13**, e1006419 (2017).
587 51 Andreasen, V., Lin, J. & Levin, S. A. The dynamics of cocirculating influenza strains
588 conferring partial cross-immunity. *Journal of mathematical biology* **35**, 825-842 (1997).
589 52 Tang, J. W. The effect of environmental parameters on the survival of airborne infectious
590 agents. *Journal of the Royal Society Interface* **6**, S737-S746 (2009).
591 53 Holmes, E. C. *et al.* Whole-genome analysis of human influenza A virus reveals multiple
592 persistent lineages and reassortment among recent H3N2 viruses. *PLoS biology* **3**, e300
593 (2005).
594 54 Skowronski, D. M. *et al.* Low 2012–13 influenza vaccine effectiveness associated with
595 mutation in the egg-adapted H3N2 vaccine strain not antigenic drift in circulating viruses.
596 *PloS one* **9**, e92153 (2014).
597 55 Munoz, E. T. & Deem, M. W. Epitope analysis for influenza vaccine design. *Vaccine* **23**,
598 1144-1148 (2005).
599 56 Skowronski, D. M. *et al.* Influenza A/subtype and B/lineage effectiveness estimates for the
600 2011–2012 trivalent vaccine: cross-season and cross-lineage protection with unchanged
601 vaccine. *The Journal of infectious diseases* **210**, 126-137 (2014).
602 57 Skowronski, D. M. *et al.* A sentinel platform to evaluate influenza vaccine effectiveness
603 and new variant circulation, Canada 2010–2011 season. *Clinical infectious diseases* **55**,
604 332-342 (2012).
605 58 Klein, E. Y., Serohijos, A. W., Choi, J.-M., Shakhnovich, E. I. & Pekosz, A. Influenza A
606 H1N1 pandemic strain evolution–divergence and the potential for antigenic drift variants.
607 *PloS one* **9**, e93632 (2014).
608 59 Pan, K., Subieta, K. C. & Deem, M. W. A novel sequence-based antigenic distance measure
609 for H1N1, with application to vaccine effectiveness and the selection of vaccine strains.
610 *Protein Engineering, Design & Selection* **24**, 291-299 (2011).

611 60 Suparit, P., Wiratsudakul, A. & Modchang, C. A mathematical model for Zika virus
612 transmission dynamics with a time-dependent mosquito biting rate. *Theoretical Biology*
613 *and Medical Modelling* **15**, 1-11 (2018).
614 61 Shrestha, S. S. *et al.* Estimating the burden of 2009 pandemic influenza A (H1N1) in the
615 United States (April 2009–April 2010). *Clinical Infectious Diseases* **52**, S75-S82 (2011).
616 62 Fraser, C. *et al.* Pandemic potential of a strain of influenza A (H1N1): early findings.
617 *science* **324**, 1557-1561 (2009).
618 63 Chung, K. & Lui, R. Dynamics of two-strain influenza model with cross-immunity and no
619 quarantine class. *Journal of mathematical biology* **73**, 1467-1489 (2016).
620 64 Zimmer, S. M. *et al.* Seroprevalence following the second wave of Pandemic 2009 H1N1
621 influenza in Pittsburgh, PA, USA. *PLoS one* **5**, e11601 (2010).
622 65 Reed, C., Katz, J. M., Hancock, K., Balish, A. & Fry, A. M. Prevalence of seropositivity
623 to pandemic influenza A/H1N1 virus in the United States following the 2009 pandemic.
624 *PloS one* **7**, e48187 (2012).
625 66 Control, C. f. D. & Prevention. Update: influenza activity-United States, 2011-12 season
626 and composition of the 2012-13 influenza vaccine. *MMWR. Morbidity and mortality*
627 *weekly report* **61**, 414-420 (2012).
628 67 Control, C. f. D. & Prevention. Update: influenza activity-United States, 2009-10 season.
629 *MMWR. Morbidity and mortality weekly report* **59**, 901-908 (2010).
630 68 Felsenstein, J. The evolutionary advantage of recombination. *Genetics* **78**, 737-756 (1974).
631 69 Jarvis, C. I. *et al.* Quantifying the impact of physical distance measures on the transmission
632 of COVID-19 in the UK. *BMC medicine* **18**, 1-10 (2020).
633 70 Schlosser, F. *et al.* COVID-19 lockdown induces disease-mitigating structural changes in
634 mobility networks. *Proceedings of the National Academy of Sciences* **117**, 32883-32890
635 (2020).
636 71 Sonoguchi, T., Naito, H., Hara, M., Takeuchi, Y. & Fukumi, H. Cross-subtype protection
637 in humans during sequential, overlapping, and/or concurrent epidemics caused by H3N2
638 and H1N1 influenza viruses. *Journal of Infectious Diseases* **151**, 81-88 (1985).
639 72 Achonu, C. *et al.* Seroprevalence of pandemic influenza H1N1 in Ontario from January
640 2009–May 2010. *PLoS one* **6**, e26427 (2011).

73 Broberg, E., Nicoll, A. & Amato-Gauci, A. Seroprevalence to influenza A (H1N1) 2009 virus—where are we? *Clinical and vaccine immunology* **18**, 1205-1212 (2011).

74 Tsai, T. *et al.* Regional and age-specific patterns of pandemic H1N1 influenza virus seroprevalence inferred from vaccine clinical trials, August-October 2009. *Eurosurveillance* **15**, 19624 (2010).

75 Fox, S. J., Miller, J. C. & Meyers, L. A. Seasonality in risk of pandemic influenza emergence. *PLoS computational biology* **13**, e1005749 (2017).

76 Opatowski, L. *et al.* Transmission characteristics of the 2009 H1N1 influenza pandemic: comparison of 8 Southern hemisphere countries. *PLoS Pathog* **7**, e1002225 (2011).

77 Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355-359 (2005).

78 Barbosa, H. *et al.* Human mobility: Models and applications. *Physics Reports* **734**, 1-74 (2018).

79 Changruenngam, S., Bicout, D. J. & Modchang, C. How the individual human mobility spatio-temporally shapes the disease transmission dynamics. *Scientific Reports* **10**, 1-13 (2020).

80 Łuksza, M. & Lässig, M. A predictive fitness model for influenza. *Nature* **507**, 57 (2014).

81 Agor, J. K. & Özaltın, O. Y. Models for predicting the evolution of influenza to inform vaccine strain selection. *Human vaccines & immunotherapeutics* **14**, 678-683 (2018).

Supporting information captions

S1 Text. Supplementary Information