Anomalous influenza seasonality in the United States and the emergence of novel influenza B viruses

Rebecca K. Borchering[a,b,1], Christian E. Gunning[a,b], Deven V. Gokhale[a,b], K. Bodie Weedop[c], Arash Saeidpour[a,b], Tobias S. Brett[a,b], and Pejman Rohani[a,b,c,d]

[a]Odum School of Ecology, University of Georgia, Athens, GA 30602; [b]Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA 30602; [c]Institute of Bioinformatics, University of Georgia, Athens, GA 30602; and [d]Department of Infectious Diseases, University of Georgia, Athens, GA 30602

Edited by Nils Chr. Stenseth, University of Oslo, Oslo, Norway, and approved December 22, 2020 (received for review August 10, 2020)

The 2019/2020 influenza season in the United States began earlier than any season since the 2009 H1N1 pandemic, with an increase in influenza-like illnesses observed as early as August. Also noteworthy was the numerical domination of influenza B cases early in this influenza season, in contrast to their typically later peak in the past. Here, we dissect the 2019/2020 influenza season not only with regard to its unusually early activity, but also with regard to the relative dynamics of type A and type B cases. We propose that the recent expansion of a novel influenza B/Victoria clade may be associated with this shift in the composition and kinetics of the influenza season in the United States. We use epidemiological transmission models to explore whether changes in the effective reproduction number or short-term cross-immunity between these viruses can explain the dynamics of influenza A and B seasonality. We find support for an increase in the effective reproduction number of influenza B, rather than support for cross-type immunity-driven dynamics. Our findings have clear implications for optimal vaccination strategies.

influenza | viral interference | genetic diversity | epidemiological models | statistical inference

Historically, studies of seasonal influenza epidemics have primarily focused on influenza A viruses (IAVs) rather than influenza B viruses (IBVs) and, in particular, on the scientifically interesting and practically important phenomenon of antigenic evolution of the hemagglutinin (HA) glycoprotein in A/H3N2 viruses (for example, refs. 1–5). The need to understand and predict population immunity to evolving seasonal IAVs has brought attention to the importance of individual-level histories of influenza virus infections (6, 7). The sequence of IAV infections an individual experiences shapes their immune response to future IAV exposures. Thus, much attention has focused on understanding this immunological complexity and predicting the perpetual changes in the antigenic and spatial structuring of circulating IAVs. However, this emphasis on the ecology and evolution of IAVs has come at the expense of that of IBVs, despite awareness that these viruses contribute to the global burden of influenza morbidity and mortality (8, 9).

Consequently, the population biology of IBVs has, until recently (10, 11), received less attention, despite dramatic evolutionary changes over the past decade. Influenza B viruses first diverged into two lineages (Yamagata and Victoria) 30 to 40 y ago (12–14). Recently, both lineages have exhibited higher evolutionary rates and several selective sweeps, but with different mechanisms of evolutionary change (15, 16). Nucleotide deletions in the B/Victoria HA gene segment have been identified and shown to characterize divergent, cocirculating subclades. These nucleotide deletions are accompanied by epistatic mutations in other B/Victoria gene segments and interclade reassortment. In contrast, antigenic drift of the neuraminidase (NA) gene segment has been shown to be a main driver in recent B/Yamagata epidemic activity (16). These evolutionary changes highlight that IBVs are capable of multiple adaptation strategies, yet we lack an understanding of their epidemiological consequences.

Here, we analyze the unusual 2019/2020 influenza season (10, 17, 18), hereafter referred to as the 2019 season, in the United States (independent of the later severe acute respiratory syndrome coronavirus 2 [SARS-CoV-2] introduction). We first characterize the highly atypical dynamics of the 2019 influenza season across the continental United States, relative to those observed over the past decade. We then present phylogenetic analyses based on 5 y of US B/Victoria sequences. We focus here on B/Victoria, as the Centers for Disease Control (CDC) weekly US influenza surveillance system reported that 98.4%of positive influenza B specimens were contained in the B/Victoria lineage (n = 14,077, cumulative from 2019 week 40 to 2020 week 12) (19). Finally, we use a combination of strategic transmission models and likelihood-based statistical inference to challenge alternative hypotheses to explain the observed changes in seasonal influenza dynamics. Ultimately, our modeling leads us to conclude that potential within-host competition between IAVs and IBVs likely operates over too short a time scale to account for the observed shifts in influenza seasonality. Rather, we find that increased effective transmission of novel IBVs provides a parsimonious explanation for the early timing and relative dominance of these viruses in 2019.

Significance

Influenza A viruses are known for persistent change that allows them to reinfect individuals. Recently, influenza B viruses have shown accelerated rates of evolutionary change, but their epidemiological consequences are not understood. We use quantitative methods to explore whether the unusually early start of the US 2019/2020 influenza B season relates to changes in the structure of influenza B/Victoria viruses. A combination of simulation studies and model-fitting exercises demonstrate that atypical 2019 influenza dynamics are better explained by changes in virus transmissibility and population susceptibility than by a competitive interaction between influenza A and B. Anticipating the timing of influenza A and influenza B epidemics could improve vaccination schedules, so that individuals have protection throughout the influenza season.

Author contributions: R.K.B. and P.R. designed research; R.K.B., C.E.G., D.V.G., K.B.W., A.S., T.S.B., and P.R. performed research; R.K.B., C.E.G., D.V.G., K.B.W., A.S., and T.S.B. analyzed data; and R.K.B., C.E.G., D.V.G., K.B.W., A.S., T.S.B., and P.R. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: rebecca.borchering@gmail.com.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2012327118/-/DCSupplemental.

Published January 25, 2021.

Results

In the 2019 season, influenza-like illness (ILI) reports rose in August (Fig. 1*A* and *SI Appendix*, Fig. S1), exhibiting an earlier seasonal influenza takeoff than any season since the 2009 H1N1 pandemic (17). In temperate countries, influenza A epidemics tend to be larger and occur before influenza B epidemics (3). The US 2019 season, however, was characterized by an atypically large and early epidemic of influenza B, with the most severe cases observed in children (18).

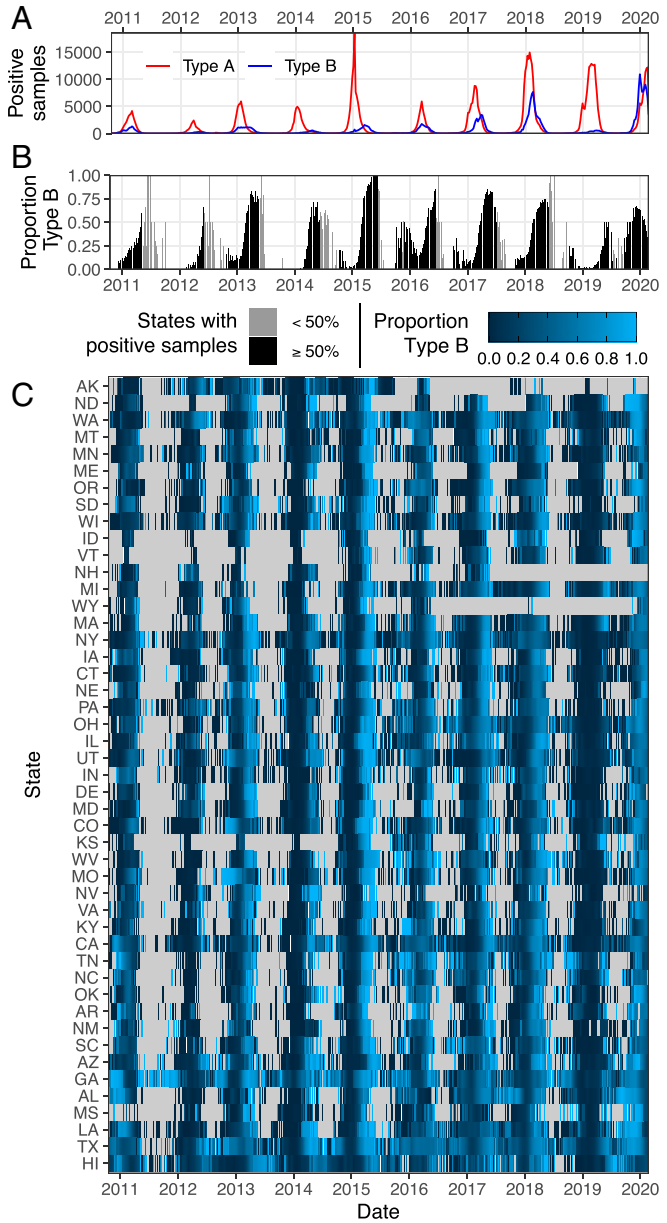

Fig. 1. Timing of influenza type A and type B epidemics in the United States. (*A*) Weekly national total of positive samples by type (see *SI Appendix*, Fig. S1 for total number of samples tested weekly by state). (*B*) Proportion of positive samples of type B (weekly median of all states). Gray shows periods of limited reporting (less than 50% of states reporting positive samples). (*C*) Weekly state-level proportion of positive samples that are type B. Gray indicates weeks lacking positive samples (either because positive A and positive B counts were both reported as zero or because one or more of these counts were not reported). States are organized from north to south (top to bottom). Weekly state-level type A and type B positive samples per 100,000 individuals are presented in *SI Appendix*, Figs. S2 and S3, respectively.

The Relative Timing of Peaks in the Type (A or B) of Confirmed Positive Influenza Samples Exhibited in the 2019 Season Is Unusual. The 2019 influenza season is the only season (of the 10 studied here) in which the week with the largest number of positive type B samples (specimens from patients confirming IBV infection) precedes the corresponding peak week for type A samples (Fig. 1*A* and *SI Appendix*, Figs. S2 and S3). In most seasons, the proportion of positive samples that are type B is lowest at the beginning of the season and then increases in the spring (Fig. 1 *B* and *C* and *SI Appendix*, Fig. S4). Again, the 2019 season is a notable exception, with an early spike in the proportion of positive samples that are type B (see the end of 2019 calendar year in Fig. 1*C*). We observe a high degree of similarity across states within each season (*SI Appendix*, Figs. S5–S8), although variation in timing across seasons is evident.

We next consider the timing of weekly positive samples of each type (by state, per 100,000 residents) relative to type A peak weeks (Fig. 2 *A*–*D*). We first identify, for each state and season, the week with the most positive type A samples and use this peak as a reference week. For each state and season, we then shift weekly positive samples per 100,000 individuals to align the respective peaks (zero on the horizontal axis). Finally, for each type and relative week, we summarize the distribution of positive samples across states and seasons. In Fig. 2 *A* and *B*, we show the 2010 to 2018 seasons, where the majority of type B samples trail the type A reference week. Fig. 2 *C* and *D* highlights the 2019 season, showing most type B samples precede the reference week. We repeat this analysis using type B peaks as reference weeks and observe similar results: Type A cases typically precede type B peaks prior to 2019, yet mostly trail the reference week in the 2019 season (*SI Appendix*, Fig. S9). Within seasons, we observe similar proportions of type B samples attributed to each lineage across age groups (20) (*SI Appendix*, Fig. S10).

To establish whether the early type B peak (Fig. 1*A*) was exceptionally unusual, we employed three separate methods. To start, we constructed two types of generalized additive models (GAMs) that characterize changes in positive samples over time. The first GAM estimates the expected weekly differences between type A and type B samples per 100,000 individuals, highlighting the pattern where type A samples typically exceed type B samples during the height of influenza seasons: approximately weeks 10 to 30 (*Materials and Methods* and Fig. 2*E*). In stark contrast, type B samples peak before those of type A in season 2019, with the expected number of type B samples per 100,000 individuals exceeding that of type A from weeks 7 to 17 (Fig. 2*E* and *SI Appendix*, Fig. S11). Our second GAM estimates the weekly proportion of positive samples that are of type B within each season (see *Materials and Methods* for details). We observe that models fitted to data from the 2010 to 2018 seasons fail to capture the unusually high proportion of type B samples observed in the early weeks of season 2019 (Fig. 2*G* and *SI Appendix*, Figs. S12 and S13).

To provide additional evidence, we use signal processing methods to examine the phase difference between IAV and IBV positive samples (*Materials and Methods*; Fig. 2*F*; and *SI Appendix*, Fig. S14). Of the past 10 influenza seasons, 2019 is the only season in which state-level peaks in IBV samples predominantly precede those of IAV samples (median 2.22 wk; interquartile range [IQR], 0.51 to 4.04 wk).

Finally, we fit a deep autoencoder neural network for unsupervised detection of potential anomalies. At the state level, we characterize an anomaly in phase as a crossing of the 99.5th percentile of the corresponding state's normalized root-mean-square deviation (NRMSD) error distribution (*Materials and Methods* and *SI Appendix*, Figs. S15–S19). We find that the months with the most anomalies occur in early 2019 (22 and 29 states in January and February, respectively). Other months

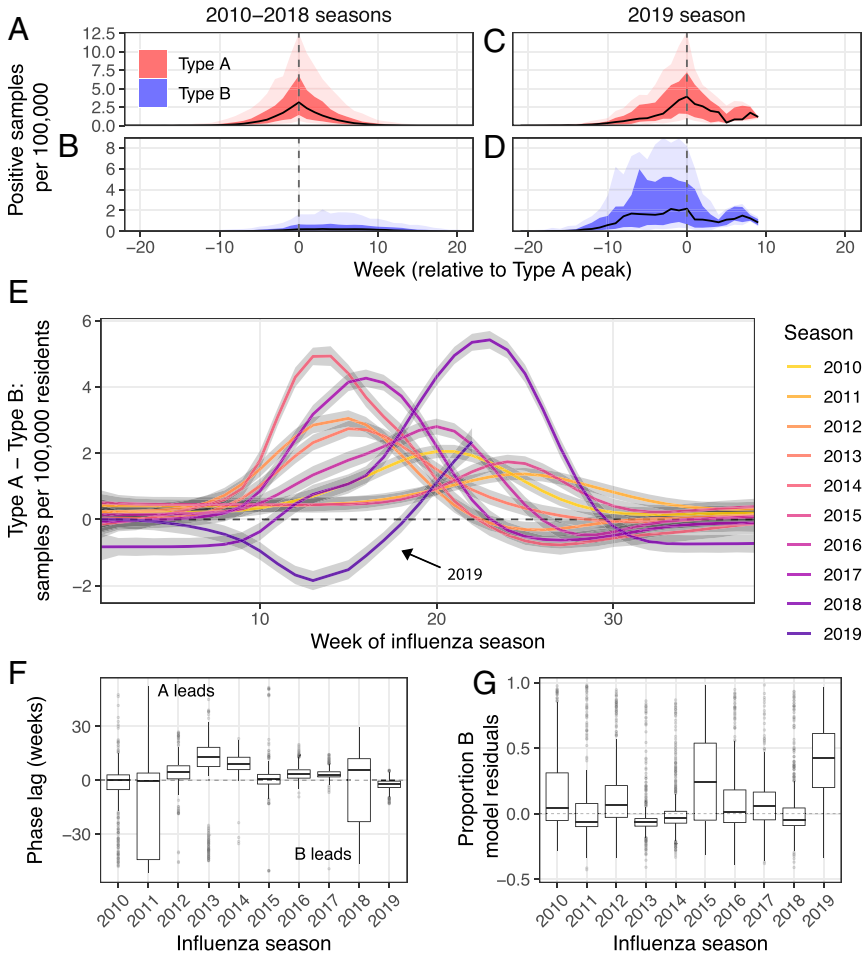

Fig. 2. Relative timing of positive influenza samples of type A and type B. (*A* and *B*) For each state and influenza season (excluding the 2019 season), we identify the week with the largest number of positive type A samples per 100,000 individuals. We then recenter these peak weeks at week zero and consider other weeks of each season relative to their corresponding peak A reference week. We then summarize the distribution of positive samples in each relative week (median in black, IQR and 80% CI in dark and light shading, respectively). In *C* and *D*, we repeat this analysis for the (incomplete) 2019 season (and thus contain fewer observations). *A* and *C* show type A (red) and *B* and *D* show type B (blue). See also *SI Appendix*, Fig. S9. (*E*) Generalized additive model fit showing, for each season (indicated by color), the expected weekly difference (per 100,000 individuals) in positive samples between types A and B (gray shading shows 95% CI). Observed values for each week–state pair are displayed for each season in *SI Appendix*, Fig. S11. (*F*) Phase lag between weekly type A and type B samples for weeks within seasons 2010 to 2019 (median, IQR indicated by boxplot). Dominant periods for each time series were calculated using wavelet transform, with relevant phases extracted from filtered time series using a low-pass filter with cutoff period of 1 y (*Materials and Methods*). (*G*) Performance of season-specific GAMs of weekly proportion of positive samples that are type B, displayed as model residuals (the difference between observed and predicted proportions; *Materials and Methods* and *SI Appendix*, Figs. S12 and S13).

with more than 10 anomalies include August 2013 (13 states) and April 2014 (12 states). Multimodel agreement that the 2019 US influenza season was atypical supports the existence of a mechanistic change in underlying epidemiology.

A Novel Influenza B/Victoria Subclade (V1A.3) Emerged Recently and Dominated Other Subclades during the 2019 Season. Multiple genetically distinct subclades of the B/Victoria lineage (V1A.1 to 4) recently emerged and have cocirculated in recent influenza seasons (2, 21). Subclade V1A.3 dominated the 2019 season, accounting for 94.4% of genotyped B/Victoria samples ($n = 849$) (19). Subclades V1A.1 to 4 arose simultaneously, bearing an amino acid (AA) deletion of either two-AA (162 to 163; V1A.1) or three-AA (162 to 164; V1A.2 to 4) positions in the HA gene segment (16). These AA deletions are coupled with subclade-specific mutations in both the HA and NA gene segments (16). We find that shifts in the seasonal timing and magnitude of US IBV samples coincide with estimated changes in IBV relative genetic diversity observed in the United States (Fig. 3 *A* and *B*). This pattern is consistent with population bottlenecks during the US summer months, when influenza incidence drops sharply. However, the expected seasonal increase in diversity is not observed during the 2018/2019 season. This asynchronous trough in relative genetic diversity, followed by a steep increase during the 2019 season, is consistent with a selective sweep and subsequent expansion of a viral lineage. Indeed, this is exactly the point at which the V1A.3 subclade emerges and rapidly diversifies (Fig. 3 *C* and *D*).

We note that our phylogenetic results are consistent with recent work by Virk et al. (16). Our results show three cocirculating diverged subclades (V1A.1 to 3) within influenza B/Victoria clade V1A. Subclade V1A.3 remains monophyletic in the HA gene segment but segregates into different lineages in the NA phylogeny (Fig. 3 *C* and *D*). A similar pattern was recently described for subclade V1A.1, which indicates some interclade reassortment in the NA gene segment (16). Previous phylodynamic results, however, do not show the 2019 dominance and recent expansion of subclade V1A.3 that we find in our analysis.

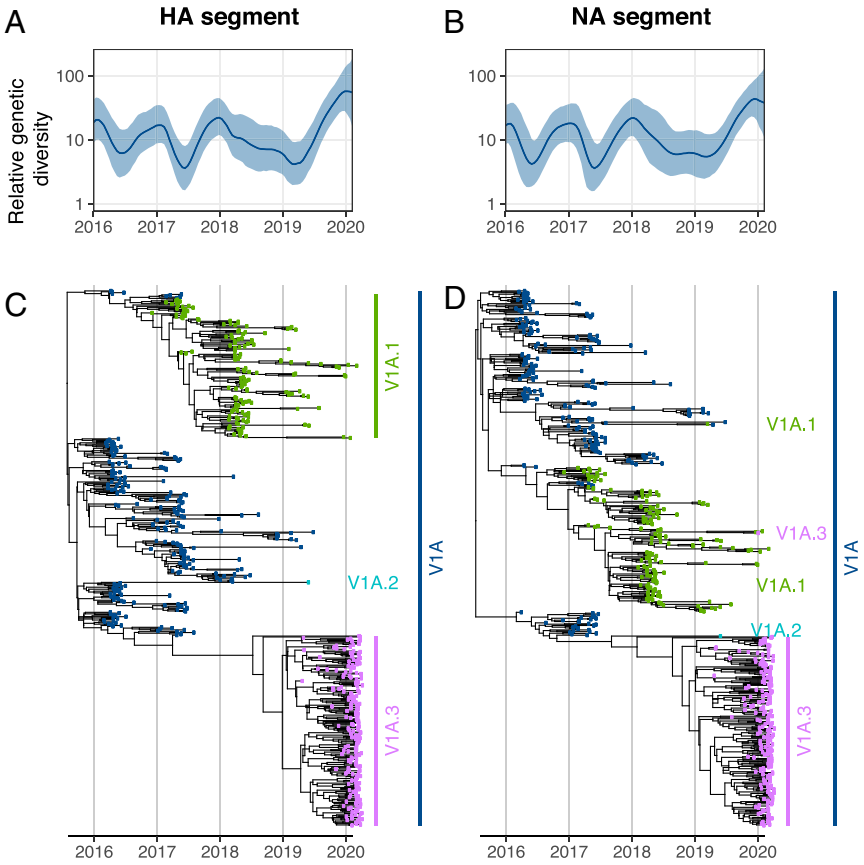

Fig. 3. Phylodynamic analysis of influenza B/Victoria viruses in the United States. (*A* and *B*) Relative genetic diversity of HA and NA gene segments estimated using a Bayesian Skyride model with Gaussian Markov random field (GMRF) smoothing. (*C* and *D*) Reconstructed temporal phylogenies for HA and NA gene segments, respectively. Tip color on phylogenies denotes subclades (V1A.1 to 3) determined by the HA gene segment.

What Mechanisms Can Explain Novel Seasonal Influenza Dynamics? Viral Interference or Changes in Transmissibility and the Susceptible Population? Here we develop transmission models to explore distinct hypotheses for the epidemiological changes we have documented: one related to viral transmissibility and the other determined by within-host competitive mechanisms (22). The recent global spread of novel B/Victoria clades and attendant increases in their diversity have raised the possibility that the HA amino acid deletions have resulted in higher transmission efficiency of these viruses (16). Alternatively, changing epidemiology may be the result of individual-level competitive mechanisms, such as viral interference (22, 23), whereby an individual's innate immune response temporarily hinders the infection ability of other viruses (24). Viral interference has been noted between influenza and noninfluenza viruses (24–26) and has been suggested to play a role in the tendency of type B influenza epidemics to occur later in the northern hemisphere than type A epidemics within a particular influenza season (27). Infection with IBV following an IAV within the same season has been observed in children, with a mean interval between diagnoses of 50 d (28). Challenge experiments in ferrets suggest that, depending on the order of infections, viral interference is possible between IAV subtypes as well as between IAVs and IBVs when the interval between infections is less than 1 wk (29). While evidence indicates that human antibodies can provide cross-protection across influenza B lineages (11), support for antibody-mediated protection across influenza types is lacking. Other immunological mechanisms that might provide a degree of cross-type protection include depletion of susceptible target cells, infection surviving epithelial cells (30), and T cell responses (31–33). Social distancing and isolation of symptomatic individuals during influenza epidemics may also limit the number of potential hosts during an outbreak of a cocirculating influenza type (34–36).

To explore the plausibility of these putative explanations, we develop a strategic, two-type mechanistic compartmental model (*Materials and Methods* and *SI Appendix*, Fig. S20 and Table S1). The model structure permits an investigation of how potential changes in type-specific transmissibility and potential cross-type interference might influence the timing and magnitude of type-specific influenza epidemics. We consider a range of potential cross-protection parameter values (Fig. 4, *Materials and Methods*, and Table 1) and find that changes in type-specific basic reproduction numbers can, in principle, explain the observed changes in the relative timing of their peaks in the absence of cross-type protection. In fact, for a set ratio of type-specific reproduction numbers, changes in average cross-protection do not substantively change the lag between type-specific influenza peaks or differences in peak height (Fig. 4 *A* and *B*, scenarios *b* and *c*). When the type A reproduction number (R_0^A) exceeds the type B reproduction number (R_0^B) (Fig. 4, scenario *a*), we observe dynamics consistent with prior seasons in which type A cases peak first and have a larger peak. By contrast, when the type B reproduction number surpasses that of type A, we observe an early and large peak in B cases (Fig. 4, scenario *d*) as observed in the 2019 US influenza season. We note that such a peak is not observed when the duration of natural immunity for type B is high relative to that for type A (*SI Appendix*, Fig. S21).

Similarly, we find no support for cross-protection as a driving epidemiological mechanism when we use maximum

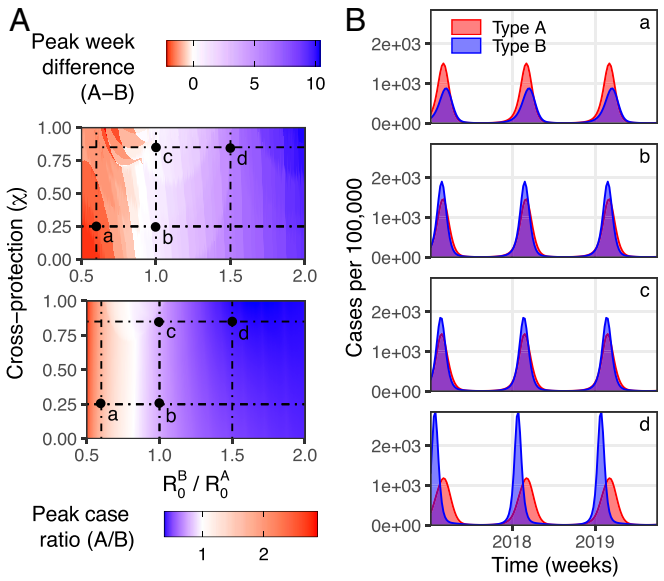

Fig. 4. Simulation study reflecting dynamic effects of cross-protection ($\chi_{AB} = \chi_{BA} = \chi$) and relative changes in type-specific R_0s. (*A*) The difference in the epidemic phases of types A and B (*Top*) and the peaks ratios (*Bottom*) of the two influenza types with relative changes of R_0^B with respect to R_0^A (horizontal axis). (*B*) Epidemic dynamics (cases per 100,000) for the two types resulting from parameter values selected at points *a*, *b*, *c*, and *d* in *A*. R_0^A is fixed at 2, and duration of cross-protection is fixed at 1 m. Type-specific immunity is assumed to last 4 y for both types.

likelihood (see *Materials and Methods* for details) to fit our strategic model to Massachusetts influenza case data for the 2016 to 2018 influenza seasons. In particular, we fit the two-type model (*Materials and Methods* and *SI Appendix*, Table S1 and Fig. S20) and compare its explanatory power of the data to a null model that assumes transmission dynamics of IAV and IBV to be independent (*SI Appendix*, Table S2 and Fig. S22). As we describe in Table 2, the added complexity of the model assuming short-term viral interaction ("cross-protection" model) is not supported by the data (ΔAIC = 90.56).

To examine whether changes in virus transmission potential may underlie the observed epidemiological shifts, we fit the neutral model to incidence data for the 2019 season (*SI Appendix*, Fig. S23 and Table S3). When comparing estimates based on 2016 to 2018 seasons to the 2019 season estimates, we observe an 18.5% increase in estimated R_0^A from 2.59 (95% CI: 2.59, 2.60) to 3.07 (3.06, 3.08) for season 2019 (Table 2 and *SI Appendix*, Tables S2 and S3). Strikingly, we observe a 54.4% increase in the estimate of R_0^B from 1.56 (1.56, 1.58) in seasons 2016 to 2018 to 2.41 (2.40, 2.42) in season 2019. Note that a similar increase is found when R_0^B is estimated from Bayesian Skyride data (*Materials and Methods*), with a rise from a value of 1.15 (95% CI: 1.11, 1.20) in 2016 to 1.82 (1.71, 1.96).

Given that our estimate of R_0^B for the 2019 season is lower than that of R_0^A, an increase in the basic reproductive number alone cannot account for the recent seasonal influenza dynamics. What, then, explains the unusual epidemiological patterns of the 2019 season? To answer this, we observe that simple transmission models describe the growth rate of an epidemic as

$$\frac{dI}{dt} = \gamma I \left(R_0 \frac{S}{N} - 1 \right), \qquad [1]$$

where γ is the recovery rate. Therefore, in the early stages of an influenza season, epidemic takeoff is determined by the combination of pathogen transmissibility, as quantified by R_0, and the fraction of the population that is susceptible, $\frac{S}{N}$. Their product is commonly referred to as the effective reproduction number, R_{eff} (35). We submit, therefore, that the unusually early and rapid rise in IBV cases in the 2019 season resulted from the documented substantial increase in R_0^B which coincided with a large fraction of the population susceptible to influenza B due to historically lower R_0^B values (Fig. 5). In concert, in the late summer months of 2019, these factors led to a higher relative per capita transmission rate for influenza B $\left(\frac{1}{I^B}\frac{dI^B}{dt} = 2.698\right)$ than for influenza A $\left(\frac{1}{I^A}\frac{dI^A}{dt} = 2.584\right)$, using Eq. **1**, with S^B and S^A obtained from the fitted model (*SI Appendix*, Table S3 and Fig. S23).

To illustrate our claim, we present a series of scenario analyses in Fig. 5. We show that an increase in R_{eff}^B in 2019 resulting from the combination of a higher R_0^B and a large susceptible fraction can capture the atypically early and large influenza B outbreak (Fig. 5 *A* and *B*), compared with no change in R_{eff}^B during seasons 2016 to 2019 when R_0^B either remained low (1.56) or high (2.41). Further, as shown in Fig. 1, there were relatively few IBV positive samples in the 2018 influenza season, and therefore we examine whether a buildup of individuals susceptible to influenza B would explain the 2019 season patterns without an accompanying increase in R_0^B. This simulation could also be interpreted as a test of two additional ideas: first, the possibility that amino acid deletions in the novel B/Victoria viruses have resulted in antigenic evolution, by increasing the available susceptible pool via newly susceptible previously infected individuals, and second, the potential increase in the susceptible population due to the dominance of B/Yamagata in the 2016 to 2018 seasons (*SI Appendix*, Fig. S10). We find that an accumulation of susceptible individuals leads to an unusually large outbreak in 2019, but does not affect the lag with the influenza A outbreak (Fig. 5 *C* and *D*).

As a final test of our key idea, we fit the neutral model to the data assuming no change in R_0^B across these four seasons (2016 to 2019). We find strong support for a model with an increase in R_0^B at the outset of the 2019 season (ΔAIC = 530.42; *SI Appendix*, Table S4).

Discussion

We document intriguing differences between the 2019 US influenza season and previous seasons, with an uncharacteristically early surge in influenza type B cases that was accompanied by noteworthy evolutionary transitions within the dominant B/Victoria lineage. Our exploration of transmission models reveals that the relative timing of the epidemics of these viruses is more sensitive to their respective transmission potential, as quantified by R_0, than to any changes to cellular- or

Table 1. Parameters for strategic model

Symbol	Value	Definition
N	6.70×10^6	Total population
$R_0^{(n)}$	1 to 5	Reproductive no. of type n ($n \in \{A, B\}$)
$1/\gamma_A$	2.5 d (57)	Infectious period of type A
$1/\gamma_B$	3.4 d (57)	Infectious period of type B
$1/\phi_n$	1 mo	Cross-protection duration after type n infection
χ_{mn}	0 to 1	Cross-protection against m after infection with n
η_n	1/d	Importation rate of type n
$1/w_A$	4 y (1)	Duration of type A-specific immunity
$1/w_B$	4, 10 y	Duration of type B-specific immunity
b_A	0 to 1	Amplitude of seasonality for type A
b_B	0 to 1	Amplitude of seasonality for type B
t_0^A	day 40	Timing of seasonal peak for type A
t_0^B	day 50	Timing of seasonal peak for type B

Table 2. Parameter estimates and goodness of fit

	2016 to 2018		2019
Estimated parameters	Cross-protection	Neutral	Neutral
R_0^A	2.54	2.59	3.07
R_0^B	2.30	1.56	2.41
χ_{AB}	0.0061		
χ_{BA}	0.95		
b_1^A	0.28	0.44	0.31
b_1^B	0.27	0.38	0.3
ρ_A	0.0038	0.0062	0.0032
ρ_B	0.0014	0.0023	0.0019
t_0^A	0.15	0.13	
t_0^B	0.18	0.15	
logLik	−3,669.89	−3,626.61	−631.54
AIC	7,359.78	7,269.21	1,287.08
ΔAIC	90.56	0	

Cross-protection and neutral models were formally contested to assess relative goodness of fit in the 2016 to 2018 seasons. Goodness of fit was assessed using the Akaike information criterion (AIC), calculated by fitting models to weekly, type-specific, incidence data for influenza for Massachusetts. See *Materials and Methods* for definitions of parameters and *SI Appendix*, Table S3 for associated 95% confidence intervals.

immune-mediated interactions. It is important to emphasize that our results do not argue against such within-host effects; instead, our models indicate that these virus–virus interactions do not leave a strong dynamical footprint in population-level incidence data. This is because of the combination of modest R_0s for both virus types leading to a relatively small fraction of the population affected and the short-term nature of any influenza interference between virus types (potentially lasting only a few days) (29). Our working hypothesis instead is that the early arrival and large epidemic of influenza B cases in the 2019 season can be explained by the combined effects of the emergence of a novel subclade of influenza B/Victoria (V1A.3; Fig. 3) with higher transmissibility than prior IBVs (16) and the availability of an unusually large susceptible pool due to historically lower R_0^B values in past seasons, as illustrated in Fig. 5*A*. This proposed explanation is consistent across our likelihood-based model fitting and our strategic modeling.

An expected consequence of our findings is a selective sweep of B/Victoria viruses by the V1A.3 subclade and the attendant atypically early and intense influenza season. However, as evidenced by contrasting patterns of IAV transmission in North America and Europe in 2017 to 2018 (37), influenza epidemiology and evolution are both complex and context specific, determined in part by local population immunity (6), patterns of mobility (38), and environmental drivers of transmission (39), which may be type specific. Thus, while these V1A.3 viruses have not been associated with a pandemic, sequence data indicate their dominance over other B/Victoria viruses locally (40). Finally, elsewhere an association has been proposed between influenza B epidemiology in Australia and evolutionary changes in influenza B lineages and differences in risk across age (15, 16).

Our analysis is limited by the number of years of data available and seasonal variation reporting (Fig. 1 and *SI Appendix*, Fig. S1). Reporting procedures changed dramatically following the 2009 H1N1 pandemic, and we have restricted our analysis to include seasons after the pandemic. Studies connecting subsequent infections with type A and type B influenza at the individual level, such as ref. 28, are needed to fully determine whether cross-protection between types occurs and, if so, the typical duration of protection.

A key outstanding question is how influenza virus evolution may change the type-specific age distribution of cases (or positive samples) (15). Influenza infection histories are more likely to include one or more prior infections with an influenza B virus for those in older age groups. These infections may allow older individuals to escape IBV infection even when novel IBVs are introduced (*SI Appendix*, Fig. S10). It is possible that dominating lineages play a role in determining the proportion of positive samples that are type B. Important avenues for future research include considering the roles of lineage-determined disease severity, reporting probability, and age-specific susceptibility.

Changes in the timing of type A and type B influenza peaks have implications for optimal vaccination strategies (3),

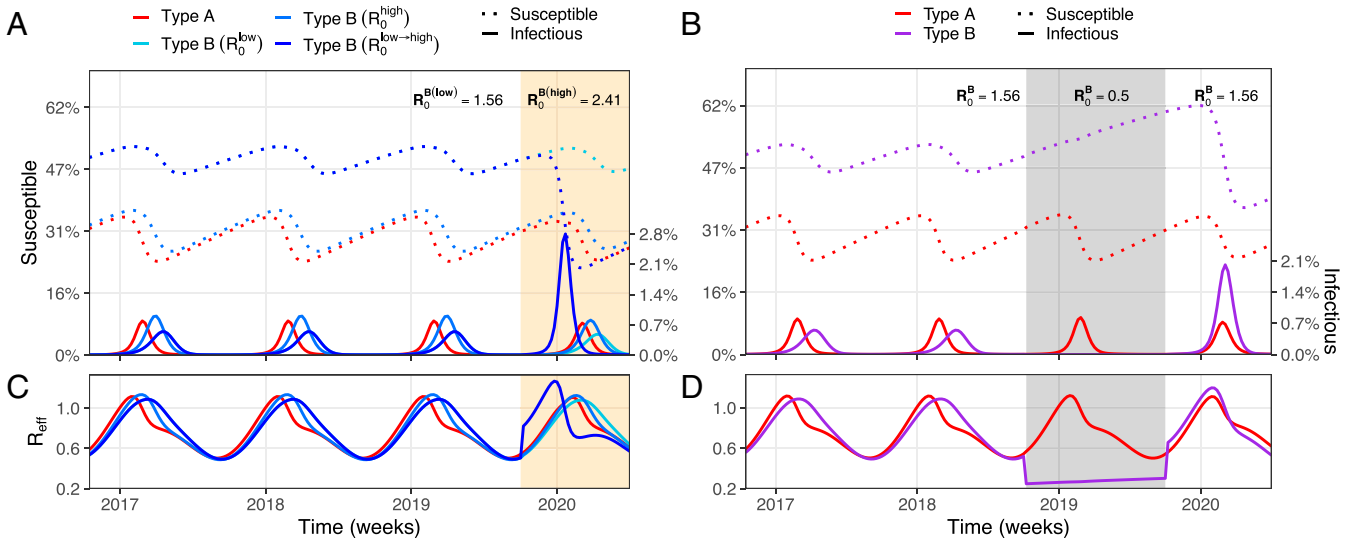

Fig. 5. Illustration of our hypothesis. (*A*) Simulation experiments demonstrating susceptible dynamics (dotted lines) and the corresponding relative timing and amplitude of influenza A (solid red line) and influenza B (solid blue lines). For influenza B, we depict three distinct scenarios: R_0^B is low throughout (=1.56) or high throughout (=2.41), or R_0^B starts low (=1.56), but increases (=2.41) at the start of the 2019 season (highlighted in orange). The associated effective reproductive numbers ($R_{\text{eff}} = R_0 \times \frac{S}{N}$) are presented in *C*. *B* and *D* present similar information to that in *A* and *C* but perform an alternative experiment, testing whether the absence of an influenza B outbreak in the 2018/2019 season highlighted in gray and resulting accumulation of susceptible individuals alone would explain the anomalous dynamics in influenza season 2019. Parameter values are presented in Table 2 and *SI Appendix*, Table S5.

particularly given that the effectiveness of influenza vaccines is thought to wane over a period of approximately 6 mo (41). The apparent connection between evolution of influenza lineages and early timing of influenza B cases suggests that measures of evolutionary change could be a valuable source of information for designating the timing of vaccination campaigns. Type-specific differences between susceptibility to infection and severe disease could also help identify high-risk age groups for targeted vaccination. Thus, it may be beneficial to identify at-risk age groups based on the expected timing of type-specific peaks in infections. Increased viral surveillance and characterization will be needed to effectively determine the optimal timing of vaccination, which may differ between age groups (42). Early awareness of the circulation of a novel influenza clade (such as the B/Victoria subclade V1A.3 identified in the 2019 season) could motivate early initiation of vaccination of younger age groups that may be at a higher risk of infection. This knowledge would be helpful for physicians and public health officials, particularly when other viruses with varying effects across age groups (e.g., SARS-CoV-2) are cocirculating.

Materials and Methods

Epidemiological Data Collation and Sources. We downloaded weekly state-level subtype-specific influenza sample data from CDC FluView Interactive ILI and Viral Surveillance (20). Starting in the 2015/2016 season, data from public health laboratories and clinical laboratories are presented separately (see ref. 43 for additional information). We collated data available (combined) from 2010 to 2015 and clinical laboratory data from 2015 to 2019 by updating field names to be consistent between files. We aggregated positive influenza A sample counts for the cases when subtyping was not performed or it was not possible to subtype the sample, into a single "no subtype" category. Weekly positive type A samples were tabulated into six categories: H1, 2009 pandemic H1, H3, H3N2v, and no subtype. Similarly, weekly positive type B samples were split into three categories: B/Victoria, B/Yamagata, and B-no lineage provided. Finally, we converted epidemiological weeks (Epiweeks) into calendar dates. Unless otherwise specified, our analyses incorporate data from all states that reported sample counts in a particular week for both type A and type B (note that a count of zero samples is considered a reported count). Particularly for summer weeks, it was common for states not to report sample counts (Fig. 1 *B* and *C*).

We used the standard influenza season specification for the United States. The season starts with Epiweek 40 and ends with Epiweek 39 of the following calendar year. We used the convention that the calendar year corresponding to the beginning of the season is designated as the season. We used the centroid latitude to order states from north to south. We extracted population projections from CDC Wonder (44) for each state–year pair. These projections were used to calculate the number of samples per 100,000 individuals (see below). For each state–week pair, we also calculated the proportion of positive samples of type B.

Values for January 3, 2015 were repeated twice for each state: once with the week of year designated as one and once as week 53 of 2014. We removed the week one values from our dataset, so that the date would not be duplicated. Additional detail is provided in the "00-processs-FluView.R" script (45).

Epidemiological Data Exploration. We characterized the relative timing of epidemics of influenza type A and type B within each influenza season. For each type, we identify the season week with the highest count of positive samples and use this as a reference point (week zero). For each type, we considered the median IQRs for positive samples per 100,000 individuals across the states which reported data for the week in consideration. We limited the weeks of consideration to those falling between 25 wk prior to the peak and 25 wk after the peak to avoid boundary effects from adjacent seasons (Fig. 2 *A–D* and *SI Appendix*, Fig. S9 *A–D*).

Next, we characterized spatiotemporal patterns of type A and type B positive samples. To facilitate comparisons between states, we first adjusted weekly positive samples by annual state population projections from CDC Wonder (44) and reported the positive samples per 100,000 people. We inspected the corresponding trajectories for six representative states (Arizona, Georgia, Louisiana, Massachusetts, New York, and Texas) within each season (*SI Appendix*, Fig. S6) and within states across multiple seasons (*SI Appendix*, Fig. S7). We also computed the difference between type A and type B samples per 100,000 individuals by week. We summarized these results using a GAM of this difference across weeks, with a separate model for each season (Fig. 2*E*). The GAM was constructed with the R package "mgcv," with season as random effect, state as fixed effect, and week of season as a cubic regression spline, i.e., gam(..., formula = y s(x, bs = 'cs')) (46). The result is the expected weekly mean difference between type A and type B samples per 100,000 individuals, marginalized across states and conditioned on weeks within a season.

To quantify asynchrony between peaks in type A and type B weekly samples, we performed a phase lag analysis between their state-level time series. We present phase lag results for epidemiological weeks 44 through 8 (Fig. 2*F*). Raw time series were first square rooted to stabilize the variance, then normalized by subtracting the mean and dividing by the standard deviation, and zero padded up to the next power of 2 to mitigate the edge effects. Next, we used a continuous Morlet wavelet transform (47) with a nondimensional frequency of $\omega_0 = 6$ to determine the dominant period of each time series. We then applied a low-pass filter with cutoff period of 1 y to extract the time series around the dominant period. Finally, we applied a Hilbert transform (48) to derive the analytical signal from each filtered time series and used it to find the phase angle. Preprocessing and phase lag calculation steps for the state of Massachusetts are depicted in *SI Appendix*, Fig. S14.

We used a deep autoencoder neural network for unsupervised detection of potential anomalies (49–51) in the time series of influenza B samples across states. The autoencoder (AE) is composed of a long short-term memory (LSTM) encoder layer which encodes a time series of length L into a vector of length m ($m < L$), which is passed to a decoder LSTM to reconstruct the time series $\hat{L}$. The AE is trained on the times series of weekly influenza B samples over the total number of positive samples for each state. We used state-specific error thresholds at 95%, 99%, and 99.5% intervals (*SI Appendix*, Fig. S19).

Regression Model and Analysis. We estimated the probability that a particular positive influenza sample is type B across time and space, again using the mgcv R package to fit a GAM of the proportion of total samples per week that are type B. We used a binomial link function, state as a fixed effect, season as a (random-effect) smoothing term, and week of season as a (fixed-effect) smoothing term, i.e., gam(y $\sim$ state + s(week, bs = 'cs') + s(season, bs = 're')) (46). We evaluated model performance by season using leave-one-out prediction. That is, for each season, we first fit the GAM using all data except that season and the 2019 to 2020 season. We then made out-of-sample predictions for the season under consideration. Finally, we computed model residuals for each state and week: the observed proportion of positive samples of type B minus corresponding model predictions (Fig. 2*G*). As the 2019 season was still in progress, we excluded the 2019 season from all model fitting above and computed only residuals for the available "early" weeks of the 2019 season: 1 to 18 (*SI Appendix*, Fig. S13). We chose not to extend this analysis to include additional weeks to avoid changes in reporting resulting from the COVID-19 pandemic.

Phylogenetic Analysis. A total of 4,302 molecular samples of influenza B/Victoria collected in the United States between 1 January 2016 and 15 March 2020 were retrieved from the GISAID database. These samples were those remaining after filtering for those which had complete sequences for both HA and NA gene segments. The total set of genetic samples was down-sampled for computational efficiency while performing Bayesian phylogenetic reconstruction and population dynamic estimation. Down-sampling was performed by randomly selecting an equal number of samples within each sampled year (625 in total; see *SI Appendix*, section 1 for sequence accession details). Maximum-likelihood (ML) phylogenies of both HA and NA gene segments were reconstructed using RAxML v8.2.11 (52) using the generalized time-reversible nucleotide substitution model and gamma rate heterogeneity. Temporal signal was verified for both gene segments using TempEst v1.5.3 (53) prior to estimation of population dynamics for a final set of 608 sequences using BEAST v1.10.4 (54). An uncorrelated lognormal relaxed clock and a Gaussian Markov random field smoothing of effective population size were used in the estimation of the molecular clock and effective population sizes. A generalized time-reversible nucleotide substitution model with gamma rate heterogeneity plus invariant sites was used through all analyses. Six independent runs of the described analysis were performed using 100 million generations while sampling every 10,000 generations. Tracer v1.7.1 (55) was used to evaluate statistical support and convergence of all independent runs. LogCombiner v1.10.4 was used to combine independent runs after removing the burn-in period (10% or 10 million generations). TreeAnnotator v1.10.4 was used to produce the

maximum clade credibility trees. All phylogenetic tree visualizations were created using the R package "ggtree" v1.16.6 (56).

We estimated the basic reproduction number (R_0) for influenza B/Victoria at the start of each influenza season (2016/2017, 2017/2018, and 2019/2020) using the results of our Gaussian Markov random field Bayesian Skyride analysis and the following equation, derived in ref. 35:

$$\Lambda = \gamma(R_0 - 1), \quad [2]$$

where Λ is the exponential growth rate of the virus population, and $1/\gamma$ is the mean duration of infection (in this case assumed to be 3.4 d) (57). To estimate Λ, we first log-transform the relative genetic diversity data (Fig. 3*A*) and then use linear regression to calculate the slope of the best-fit line. For each season, we choose the number of initial points to be included based on the number of points that maximize the R^2 of the linear regression. We also require that the number of data points exceeds two.

Influenza Transmission Model. To explore the potential role of cross-immunity in determining the timing and magnitude of influenza A vs. B peaks, we developed a two-strain susceptible infectious cross-protected recovered susceptible (SICRS) model (*SI Appendix*, Fig. S20) (35). The population (of size N) is divided into 10 compartments dependent on infection and immune status. Susceptible (S) individuals have no immunity and can be infected with either type at rate λ_n, where λ_n is the force of infection of type n ($n \in \{A, B\}$). Upon infection individuals move to the corresponding I_n compartment. Following the infectious period (with mean duration $1/\gamma_n$) individuals recover and move to the cross-protected compartment, C_n. After recovering, individuals initially have perfect immunity against reinfection with the same type and partial cross-protection against infection with the other type. Cross-protection against type m due to prior infection with n corresponds to a reduction in the infection rate, $(1 - \chi_{mn})\lambda_m$. The cross-protection strength, χ_{mn}, takes values between 0 (no cross-protection) and 1 (perfect cross-protection). Cross-protection wanes with rate ϕ_n. Individuals who have lost cross-protection against type m but retain type-specific immunity move to R_n. Type-specific immunity also wanes, with rate w_n. If individuals experience an infection with the second type, m, before their type-specific immunity wanes they become infectious with type m and move to the I_{mn} compartment. Upon recovery they move to the final model compartment, R, which corresponds to individuals with type-specific immunity to both types. In *SI Appendix*, Tables S1 and S2 summarize the model compartments and parameters, respectively.

The force of infection for type n depends on the number of individuals infectious with type n and the rate cases are imported to the population from external sources (proportional to η_n),

$$\lambda_n(t) = \beta_n(t)(I_n + I_{nm} + \eta_n)/N. \quad [3]$$

Our model allows for seasonality in the transmission rate,

$$\beta_n(t) = \gamma_n R_0^{(n)} \left[1 + b_n \cos\left((2\pi t - t_0^n)/T\right)\right], \quad [4]$$

where $R_0^{(n)}$, b_n, and t_0^n are respectively the basic reproductive number, amplitude of seasonality, and peak day for type n and T is the period seasonality (1 y).

The model dynamics are captured by a system of ordinary differential equations,

$$\frac{dS}{dt} = \sum_n \{-\lambda_n S + w_n R_n\}, \quad [5]$$

$$\frac{dI_n}{dt} = \lambda_n S - \gamma_n I_n, \quad [6]$$

$$\frac{dC_n}{dt} = -(1 - \chi_{mn})\lambda_m C_n + \gamma_n I_n - \phi_n C_n, \quad [7]$$

$$\frac{dR_n}{dt} = -\lambda_m R_n + \phi_n C_n + w_m R - w_n R_n, \quad [8]$$

$$\frac{dI_{mn}}{dt} = (1 - \chi_{mn})\lambda_m C_n + \lambda_m R_n - \gamma_m I_{mn}, \quad [9]$$

$$\frac{dR}{dt} = \sum_n \{\gamma_n I_{nm} - w_n R\}, \quad [10]$$

for $n \in \{A, B\}$. If $n = A$, then $m = B$ and vice versa.

For the likelihood-based inference, we generated the cumulative number of cases according to

$$\frac{dK_A}{dt} = \gamma_A(I_{AB} + I_A), \quad [11]$$

$$\frac{dK_B}{dt} = \gamma_B(I_{BA} + I_B). \quad [12]$$

For formal comparison with data, the weekly counts of new A and B cases were calculated, $Q_n(t) = K_n(t) - K_n(t - \delta)$, where $\delta = 1$ wk. The deterministic model was implemented in R using the package "pomp" (58). Numerical simulations were performed using the resulting pomp object. Simulations were initialized at the endemic equilibrium of the noninteracting two-type model ($\chi_{BA} = \chi_{AB} = 0$) without seasonality ($b_A = b_B = 0$), which possesses an analytical solution (35). The first 100 y were discarded as transient dynamics, before the subsequent 3 y were used in likelihood calculation.

Model Fitting and Comparison. Using type-specific data for Massachusetts (20) influenza seasons 2016 to 2018, we performed maximum-likelihood estimation of the parameters of our two-strain model to investigate whether there was support for a significant strength of cross-protection from viruses belonging to the other type (A vs. B). We fitted two versions of our model, a noninteracting model in which the cross-protection parameters (χ_{AB} and χ_{BA}) were fixed at zero and the full model in which these parameters were allowed to vary. We then performed a formal model comparison of these models (described below).

Maximum-likelihood estimates (MLEs) for the unknown parameters were found using trajectory matching. Weekly case counts were assumed to be subject to Poisson-distributed reporting error. The likelihood function for the data $\{x\}_{t=0}^{T}$ given model parameters Θ was

$$\ell(\Theta) = \prod_{n \in \{A,B\}} \prod_{t=0}^{T} P(x_{n,t} | \rho Q_n(t; \Theta)), \quad [13]$$

where ρ is the reporting probability, $Q_n(t; \Theta)$ is the simulated number of new cases at time t using parameters Θ, and $P(\cdot|\lambda)$ is the Poisson probability mass function with parameter λ. Note that we estimated separate reporting probabilities for influenza types A and B (ρ_A and ρ_B, respectively).

We employed a numerical optimizer routine for two tasks: 1) explore the parameter space and find the MLEs for all unknown parameters and 2) verify whether the inclusion of cross-protection improved model fits to confirmed sample data. Trajectory matching was performed using the R package pomp which allows for the construction and optimization of likelihood-based objective functions (58). The optimization routine was as follows:

1) A population of 500 parameter configurations were used to initialize a stochastic optimizer that implemented a hybrid-genetic algorithm to maximize the objective function. Guesses were generated using a Latin Hypercube (Sobol) design to ensure efficient sampling of the parameter space (59, 60).
2) The genetic algorithm routine was implemented using the R package "GA" (61). This technique evolves the solution until a configuration that maximizes the objective function is found. Probability of random mutations was set to a low value (0.2) while the probability of crossover was set to a high value (0.8). Force of selection was assumed to be linear in fitness of the solution in the objective space.
3) At every significant increase in the objective, an increase in the fitness value by 10%, the solution was pushed through a deterministic optimizer method (limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm [L-BFGS]) (62) to rapidly arrive at intermediate local maxima. We used a hybrid version of the optimizer to decrease convergence times to the global maximum for each of the two models (63).
4) An exhaustive population of evolving solutions provided high initial fitness diversity and a convergence criterion of 500 evolutionary iterations ensured the optimizer's convergence to a global maximum in the objective space (64).

We used the Akaike information criterion (AIC) to assess the relative goodness of fit of the two hypotheses—a noninteracting neutral model and a full two-strain model. At 5% level of significance, an AIC value of 2 was taken to be a significant difference in the comparative agreements of the two models. Parameter estimates as well as their projected type-specific cases were used in this assessment. For the 2016 to 2018 era, models were allowed to reach equilibrium (burn-in period of 100 y) and subsequent 3 y

were used in the fitting procedure. Parameter estimates corresponding to the best-fitting model, neutral, from the initial era were used to infer initial conditions of state variables for the 2019 season. This model was refitted to estimate type-specific reproductive ratios for this season.

Estimation of Uncertainty around Parameter MLE. The method of parametric bootstrap was used to estimate 95% confidence intervals around the parameter MLE obtained in the previous step. This involved the following: 1) For each model in the two eras, MLEs obtained in the preceding section were used to simulate 1,000 synthetic time series. 2) Starting at the MLE, a local search with an L-BFGS optimizer was conducted and the parameter reestimated for every time series. 3) This results in a bootstrapped distribution of parameter estimates around the MLE. Estimate sets corresponding to 2.5 and 97.5 percentiles of this distribution were taken to be as the 95% confidence bounds of the model MLEs.

Data Availability. Source code and data have been deposited in Zenodo (DOI: 10.5281/zenodo.4411959) (45).

ACKNOWLEDGMENTS. We acknowledge Mark Tompkins for valuable discussion. We acknowledge the originating and submitting laboratories for our use of sequences from the GISAID's EpiFlu Database. This work is funded by the NIH through the Models of Infectious Disease Agent Study Project R01GM123007 and by the NSF through the Interdisciplinary Disease Ecology Across Scales Program DGE-1545433.

1. D. J. Smith *et al.*, Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 371–376 (2004).
2. K. Koelle, D. A. Rasmussen, The effects of a deleterious mutation load on patterns of influenza A/H3N2's antigenic evolution in humans. *Elife* **4**, e07361 (2015).
3. J. B. Plotkin, J. Dushoff, S. A. Levin, Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6263–6268 (2002).
4. T. Bedford *et al.*, Integrating influenza antigenic dynamics with molecular evolution. *elife* **3**, e01914 (2014).
5. S. L. Linderman *et al.*, Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013–2014 influenza season. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15798–15803 (2014).
6. K. M. Gostic, M. Ambrose, M. Worobey, J. O. Lloyd-Smith, Potent protection against H5N1 and H7N9 influenza via childhood hemagglutinin imprinting. *Science* **354**, 722–726 (2016).
7. P. Palese, T. T. Wang, Why do influenza virus subtypes die out? A hypothesis. *mBio* **2**, e00150-11 (2011).
8. W. Paul Glezen, J. K. Schmier, C. M. Kuehn, K. J. Ryan, J. Oxford, The burden of influenza B: A structured literature review. *Am. J. Publ. Health* **103**, e43–e51 (2013).
9. M. Tafalla, M. Buijssen, R. Geets, M. V. Noordegraaf-Schouten, A comprehensive review of the epidemiology and disease burden of influenza B in 9 European countries. *Hum. Vaccines Immunother.* **12**, 993–1002 (2016).
10. S. Caini *et al.*, The epidemiological signature of influenza B virus and its B/Victoria and B/Yamagata lineages in the 21st century. *PloS One* **14**, e0222381 (2019).
11. Y. Liu *et al.*, Cross-lineage protection by human antibodies binding the influenza B hemagglutinin. *Nat. Commun.* **10**, 324 (2019).
12. P. A. Rota *et al.*, Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. *Virology* **175**, 59–68 (1990).
13. Y. Kanegae *et al.*, Evolutionary pattern of the hemagglutinin gene of influenza B viruses isolated in Japan: Cocirculating lineages in the same epidemic season. *J. Virol.* **64**, 2860–2865 (1990).
14. G. Dudas, T. Bedford, S. Lycett, A. Rambaut, Reassortment between influenza B lineages and the emergence of a coadapted PB1–PB2–HA gene complex. *Mol. Biol. Evol.* **32**, 162–172 (2015).
15. D. Vijaykrishna *et al.*, The contrasting phylodynamics of human influenza B viruses. *Elife* **4**, e05055 (2015).
16. R. K. Virk *et al.*, Divergent evolutionary trajectories of influenza B viruses underlie their contemporaneous epidemic activity. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 619–628 (2020).
17. F. S. Dawood *et al.*, Interim estimates of 2019–20 seasonal influenza vaccine effectiveness—United States, February 2020. *Morb. Mortal. Wkly. Rep.* **69**, 177–182 (2020).
18. D. Owusu *et al.*, Early season pediatric influenza B/Victoria virus infections associated with a recently emerged virus subclade—Louisiana, 2019. *Morb. Mortal. Wkly. Rep.* **69**, 40–43 (2020).
19. Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases, Weekly U.S. influenza surveillance report. https://www.cdc.gov/flu/weekly/index.htm. Accessed 31 March 2020.
20. Centers for Disease Control and Prevention, Fluview interactive. https://gis.cdc.gov/grasp/fluview/flu_by_age_virus.html. Accessed 5 June 2020.
21. S. Epperson *et al.*, Update: Influenza activity—United States and worldwide, May 19–September 28, 2019, and composition of the 2020 southern hemisphere influenza vaccine. *Morb. Mortal. Wkly. Rep.* **68**, 880–884 (2019).
22. T. DaPalma, B. P. Doonan, N. M. Trager, L. M. Kasman, A systematic approach to virus–virus interactions. *Virus Res.* **149**, 1–9 (2010).
23. L. Elveback, J. Fox, Illness, oral poliovirus vaccine, and interference. *Arch. Environ. Health* **9**, 724–726 (1964).
24. K. F. Chan *et al.*, Investigating viral interference between influenza A virus and human respiratory syncytial virus in a ferret model of infection. *J. Infect. Dis.* **218**, 406–417 (2018).
25. S. Nickbakhsh *et al.*, Virus–virus interactions impact the population dynamics of influenza and the common cold. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 27142–27150 (2019).
26. K. M. Kloepfer, J. E. Gern, Ecological and individual data both indicate that influenza inhibits rhinovirus infection. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 6987 (2020).
27. M. E. Francis, M. L. King, A. A. Kelvin, Back to the future for influenza preimmunity—Looking back at influenza virus history to infer the outcome of future infections. *Viruses* **11**, 122 (2019).
28. J. Möst, G. Weiss, Consecutive infections with influenza A and B virus in children during the 2014–2015 seasonal influenza epidemic. *J. Infect. Dis.* **214**, 1139–1141 (2016).
29. K. L. Laurie *et al.*, Interval between infections and viral hierarchy are determinants of viral interference following influenza virus infection in a ferret model. *J. Infect. Dis.* **212**, 1701–1710 (2015).
30. R. E. Dumm *et al.*, Non-lytic clearance of influenza B virus from infected cells preserves epithelial barrier function. *Nat. Commun.* **10**, 779 (2019).
31. E. B. Clemens, C. Van de Sandt, S. S. Wong, L. M. Wakim, S. A. Valkenburg, Harnessing the power of T cells: The promising hope for a universal influenza vaccine. *Vaccines* **6**, 18 (2018).
32. M. Terajima, J. A. B. Babon, M. D. T. Co, F. A. Ennis, Cross-reactive human B cell and T cell epitopes between influenza A and B viruses. *Virol. J.* **10**, 244 (2013).
33. M. Koutsakos *et al.*, Human CD8+ T cell cross-reactivity across influenza A, B and C viruses. *Nat. Immunol.* **20**, 613–625 (2019).
34. P. Rohani, D. J. Earn, B. Finkenstädt, B. T. Grenfell, Population dynamic interference among childhood diseases. *Proc. R. Soc. Lond. B Biol. Sci.* **265**, 2033–2041 (1998).
35. M. J. Keeling, P. Rohani, *Modelling Infectious Diseases: In Humans and Animals* (Princeton University Press, 2008).
36. P. Rohani, C. Green, N. Mantilla-Beniers, B. T. Grenfell, Ecological interference between fatal diseases. *Nature* **422**, 885–888 (2003).
37. A. Hammond *et al.*, Review of the 2017-2018 influenza season in the northern hemisphere. *Wkly. Epidemiol. Rec.* **93**, 429–445 (2018).
38. C. Viboud *et al.*, Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–451 (2006).
39. E. R. Deyle, M. C. Maher, R. D. Hernandez, S. Basu, G. Sugihara, Global environmental drivers of influenza. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13081–13086 (2016).
40. J. Hadfield *et al.*, Nextstrain: Real-time tracking of pathogen evolution. https://nextstrain.org/flu/seasonal/vic/ha/2y. Accessed 9 December 2020.
41. J. M. Ferdinands *et al.*, Intraseason waning of influenza vaccine protection: Evidence from the US influenza vaccine effectiveness network, 2011–2012 through 2014–2015. *Clin. Infect. Dis.* **64**, 544–550 (2017).
42. J. M. Ferdinands, E. Alyanak, C. Reed, A. M. Fry, Waning of influenza vaccine protection: Exploring the trade-offs of changes in vaccination timing among older adults. *Clin. Infect. Dis.* **70**, 1550–1559 (2020).
43. U.S. influenza surveillance system: Purpose and methods. https://www.cdc.gov/flu/weekly/overview.htm. Accessed 17 April 2020.
44. Centers for Disease Control and Prevention, Population projections, United States, 2004–2030, by state, age, and sex, on CDC WONDER online database. http://wonder.cdc.gov/population-projections.html. Accessed 16 August 2019.
45. R. K. Borchering *et al.*, Anomalous influenza seasonality in the United States and the emergence of novel influenza B/Victoria viruses. Zenodo. http://doi.org/10.5281/zenodo.4411959. Deposited 29 December 2020.
46. S. N. Wood, Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc.* **73**, 3–36 (2011).
47. C. Torrence, G. P. Compo, A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* **79**, 61–78 (1998).
48. M. Le Van Quyen *et al.*, Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. *J. Neurosci. Methods* **111**, 83–98 (2001).
49. J. An, S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability" in *Special Lecture on IE* (Seoul National University Data Mining Center, Seoul, South Korea, 2015), vol. 2.
50. M. Sakurada, T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction" in *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, A. R. Rahman, J. D. Deng, J. L. Li, Eds. (Association for Computing Machinery, New York, NY, 2014), pp. 4–11.
51. C. Zhou, R. C. Paffenroth, "Anomaly detection with robust deep autoencoders" in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, S. M. Matwin, S. U. Yu, Eds. (Association for Computing Machinery, New York, NY, 2017), pp. 665–674.
52. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

53. A. Rambaut, T. T. Lam, L. Max Carvalho, O. G. Pybus, Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016).
54. M. A. Suchard *et al.*, Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018).
55. A. Rambaut, A. J. Drummond, D. Xie, G. Baele, M. A. Suchard, Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).
56. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T. Y. Lam, ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
57. F. Carrat *et al.*, Time lines of infection and disease in human influenza: A review of volunteer challenge studies. *Am. J. Epidemiol.* **167**, 775–785 (2008).
58. A. A. King, D. Nguyen, E. L. Ionides, Statistical inference for partially observed Markov processes via the R package pomp. *J. Stat. Softw.* **69**, 1–43 (2016).
59. I. M. Sobol, On the distribution of points in a cube and the approximate evaluation of integrals. *Zh. Vychislitel noi Mat. Mat. Fiz.* **7**, 784–802 (1967).
60. I. M. Sobol, Uniformly distributed sequences with an additional uniform property. *USSR Comput. Math. Math. Phys.* **16**, 236–242 (1976).
61. L. Scrucca, GA: A package for genetic algorithms in R. *J. Stat. Softw.* **53**, 1–37 (2013).
62. H. Matthies, G. Strang, The solution of nonlinear finite element equations. *Int. J. Numer. Methods Eng.* **14**, 1613–1626 (1979).
63. L. Scrucca, On some extensions to GA package: Hybrid optimisation, parallelisation and islands evolution. *R J.* **9**, 187–206 (2017).
64. A. E. Eiben *et al.*, *Introduction to Evolutionary Computing* (Springer, 2003).