Genomic epidemiology of influenza B circulation in China during prolonged border closure from 2020 to 2021

Ruopeng Xie
The University of Hong Kong
Dillon Adam
The University of Hong Kong
Kimberly Edwards
The University of Hong Kong https://orcid.org/0000-0003-0944-3118
Shreya Gurung
The University of Hong Kong
Xiaoman Wei
The University of Hong Kong
Benjamin Cowling
University of Hong Kong https://orcid.org/0000-0002-6297-7154
Vijaykrishna Dhanasekaran (veej@hku.hk)
The University of Hong Kong https://orcid.org/0000-0003-3293-6279

Article

Keywords:

Posted Date: March 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1432770/v1

Genomic epidemiology of influenza B circulation in China during prolonged border closure from 2020 to 2021

Ruopeng Xie[1,2], Dillon C. Adam[1], Kimberly M. Edwards[1,2], Shreya Gurung[1,2], Xiaoman Wei[1,2], Benjamin J. Cowling[1], Vijaykrishna Dhanasekaran[1,2*]

[1]School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China

[2]HKU-Pasteur Research Pole, School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China

*Corresponding author:
Vijaykrishna Dhanasekaran
School of Public Health, LKS Faculty of Medicine, The University of Hong Kong,
Hong Kong, China
email: veej@hku.hk

China experienced a resurgence of seasonal influenza activity throughout 2021 under intermittent control measures and prolonged international border closure. We show genomic evidence for multiple A(H3N2), A(H1N1), and B/Victoria transmission lineages circulating for over three years, with the 2021 resurgence driven by two B/Victoria clades. Novel hemagglutinin (HA) gene mutations and altered age profiles of infected individuals were observed, and Jiangxi province was identified as a major source for nationwide outbreaks. An increase in the effective population size and effective reproduction number of B/Victoria outbreaks followed major national holidays, several months prior to their initial detection, underscoring the importance of influenza vaccination prior to holiday periods or travel. Clade 3a2 B/Victoria viruses with the HA1 mutation H122Q and a distinct gene constellation are becoming increasingly dominant in China and warrant close monitoring due to poor antigenicity match with current vaccine strains.

Despite a global decline in seasonal influenza activity during the COVID-19 pandemic, influenza cases were reported to the WHO Global Influenza Surveillance and Response System (GISRS) throughout 2020. Genetic data submitted to Global Initiative on Sharing All Influenza Data (GISAID) indicates that A(H3N2), A(H1N1), and B/Victoria viruses circulated predominantly in South Asia, West Africa, and China respectively, while B/Yamagata lineage viruses appear to have been eliminated [1-4]. As travel restrictions gradually eased in 2021, A(H3N2) viruses became increasingly prevalent across much of the Northern Hemisphere [5], but in China, where international borders remained closed, the incidence of B/Victoria increased since late 2020 and throughout 2021 [3, 6-8]. Here we investigate the evolution and epidemiology of B/Victoria in China amid heightened testing, international border closures, and non-pharmaceutical interventions (NPIs) associated with the COVID-19 pandemic during 2020-2021.

Results

Persistence and resurgence of seasonal influenza activity in China

China maintained an elimination strategy throughout 2020 and 2021 to curb the spread of COVID-19. After the initial lockdown was lifted in April 2020, SARS-CoV-2 outbreaks were locally suppressed through rapid isolation, large-scale testing, and other NPIs [9], enabling the majority of mainland China to maintain a baseline of light social measures and public health

guidelines, such as testing, mask mandates, and periodic inter-provincial border closures while international borders remained closed [10].

According to the Chinese National Influenza Center, all seasonal influenza virus subtypes, including B/Yamagata, were detected in the first week of January 2020 when seasonal influenza activity peaked [11]. The rapid decline thereafter coincided with lockdowns beginning in Wuhan and extending across Hubei and all other provinces [12]. On average, NPIs associated with COVID-19 control reduced influenza activity by 79% across China in comparison to past seasons (2011-2019) [13]. Although little to no influenza activity was reported from April to November 2020 [14], cases in China began to rebound in December 2020 and fluctuated throughout 2021 **(Extended Data Fig. 1)**. Based on sequences available in GISAID (accessed 07-Nov-2021), nine A(H3N2) virus sequences were reported from May to August 2020, all of which were from Yunnan province; one A(H1N1) virus was collected in Fujian in December 2020; and no B/Yamagata viruses were reported after January 2020 [3, 6]. Remarkably, >99% of confirmed influenza cases in China from October 5, 2020 to September 5, 2021 were B/Victoria lineage viruses [15].

V1A was the dominant B/Victoria lineage globally during 2016-2017 from which V1A.3, and subclades 3a1 and 3a2 subsequently emerged. From early-2019 to mid-2020, B/Victoria HA sequences in China belonged predominantly to clade V1A.3 **(Fig. 1)**, whereas genomes sequenced since late 2020 mainly belonged to clades 3a1 and 3a2, with only a few V1A and V1A.3 viruses detected [6]. Clades 3a1 and 3a2 were first detected in China during late 2019, in September and November, respectively **(Fig. 1b)**. Clade 3a1 and 3a2 HA sequences obtained during 2019-2021 from China formed distinct monophyletic clusters with ~99.5% nucleotide identity **(Extended Data Fig. 2)**. Similarly, the few A(H3N2), A(H1N1) and B/Victoria V1A and V1A.3 viruses detected since mid-2020 were most closely related to viruses detected during the 2019-2020 season in China **(Extended Data Fig. 3)**. These results suggest persistent, low-level, domestic circulation of multiple influenza lineages within China despite highly stringent control measures in early 2020.

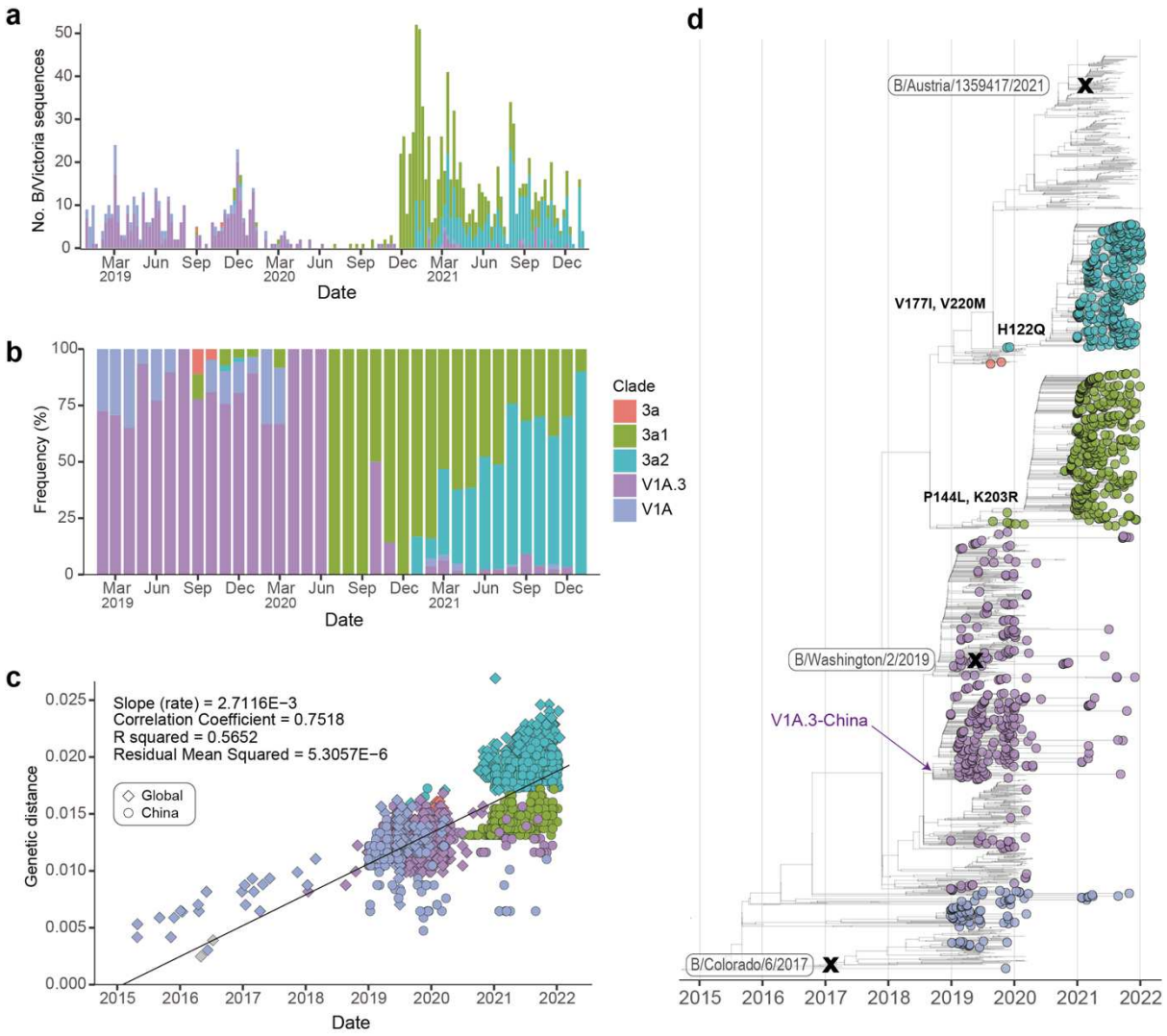

Fig. 1 | Genomic surveillance and evolution of influenza B/Victoria in China since 2019. a, Influenza B/Victoria sequences from China submitted to GISAID [16]. **b**, Proportional clade distribution by month inferred from (**a**). **c**, Root-to-tip regression analysis of the HA gene segments of influenza B/Victoria viruses shown in (**d**). **d**, Time-scaled maximum likelihood phylogenetic tree of the HA gene of influenza B/Victoria viruses, with viruses from China colored by clade. Tips with no color represent global B/Victoria viruses. Amino acid mutations of 3a1 and 3a2 are annotated on ancestral branches. B/Victoria vaccine strains B/Colorado/6/2017 (clade V1A.1), B/Washington/2/2019 (clade V1A.3), and B/Austria/1359417/2021 (clade 3a2) are indicated by "X".

A root-to-tip regression of HA distances and sampling dates suggests an increased evolutionary rate for the ancestral branch leading to clade 3a2 viruses **(Fig. 1c)**. The resurgence of B/Victoria lineages during 2020-2021 in China can likely be attributed to the recent acquisition of HA amino acid mutations in 3a1 and 3a2 that render them fitter than previous variants through partial immune escape. Clades 3a1 and 3a2 feature amino acid substitutions indicative of antigenic drift in the HA1 surface protein, including V177I and V220M in clade 3a1, and P144L and K203R in clade 3a2. Notably, clade 3a2 viruses circulating in China contain an

additional H122Q mutation that is absent from 3a2 viruses outside China, including the vaccine strain B/Austria/1359417/2021, selected in September 2021 **(Fig. 1d)** [17]. While HA1 site 122 lies within the highly variable 120-loop (amino acid positions 116-137) [18, 19], position 122 is highly conserved in both influenza B virus lineages: >97% of B/Victoria viruses globally since 2000 contained Histidine (H) at site 122, while >97% of B/Yamagata viruses contain Glutamine (Q) **(Extended Data Table 1)**.

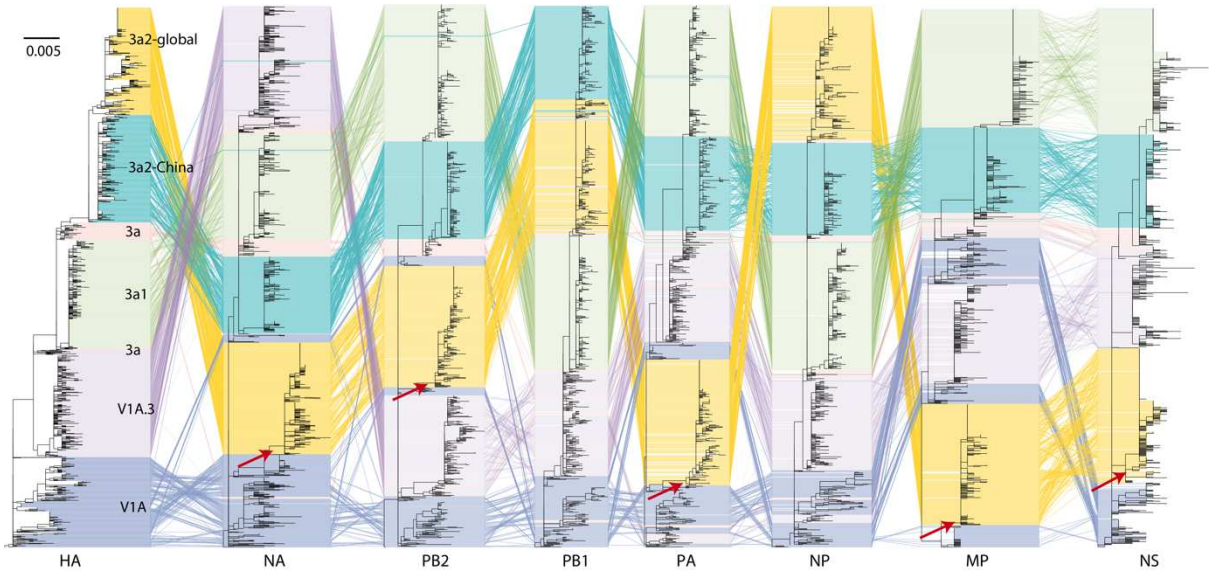

Fig. 2 | Tanglegram of B/Victoria virus reassortment. Colored lines connect each virus across all eight genes, showing incongruence between and within major clades. Red arrows denote the acquisition of NA, PB2, PA, MP, and NS by clade 3a2 viruses in global circulation from clade V1A.

Furthermore, analysis of all eight gene segments showed 3a2 viruses circulating outside China were distinct reassortants, having acquired the neuraminidase (NA) and internal genes PB2, PA, MP, and NS from V1A viruses in 2019 **(Fig. 2)**. V1A continued to be sporadically detected in China from 2018-2021 **(Fig. 1d)**, suggesting the 3a2 HA in global circulation potentially originated in China prior to COVID-19.

Temporal phylodynamics of seasonal influenza in China

To confirm suitability of the dataset for estimation of phylodynamic parameters, we conducted root-to-tip regression analyses of genetic distances and collection dates for circulating clades in China (3a1, 3a2, and V1A.3-China (the largest monophyletic clade of V1A.3 in China during 2019-2021), shown in **Fig. 1d**). While individual gene segments showed weak temporal signal ($r^2 < 0.3$) **(Extended Data Table 2)**, concatenated whole genomes of each clade demonstrated sufficient temporality with no evidence of reassortment (see Methods), enabling subsequent

phylodynamic inference. The mean evolutionary rates of 3a1, 3a2, and V1A.3-China estimated across the genome were 0.5 – 1.8 x 10^{-3} nucleotide substitutions per site (nt/subs/site) **(Extended Data Table 2)**, similar to the long-term evolutionary rates estimated individually for the internal segments (0.14 – 1.7 x 10^{-3} nt/subs/site) but lower than mean estimates for the HA and NA segments (~2 x 10^{-3} nt/subs/site) [20], indicating unremarkable selection pressure on clades circulating in China.

Phylogenetic analysis shows no new introductions occurred for clades 3a1 and 3a2 after international borders closed around March 2020. Estimates of the mean time to most recent common ancestor (tMRCA) indicate that clades 3a1 and 3a2 circulated locally for around six months prior to widespread detection. Clade 3a1, collected since July 2020, had a tMRCA of 25-June-2020 (95% highest posterior density interval (HPD) 12-June-2020 to 7-July-2020), and the tMRCA of clade 3a2 was 18-September-2020 (HPD, 20-August-2020 to 14-October-2020) **(Fig. 3)**. Whole-genome phylogenetic trees suggest 3a1 and 3a2 circulated in eastern China prior to widespread detection in December 2020 and March 2021, respectively. To a lesser extent, clades 3a1 and 3a2 were regionally maintained in the southwest and northwest respectively, for several months.

Temporal phylogenetics of clades 3a1 and 3a2 also point to lineage expansion following three major national holidays in mainland China, National Day (1-7 October 2020), Spring Festival (11-17 February 2021), and Labour Day (1-5 May 2021) (dashed lines, **Fig. 3**). We therefore characterized fluctuations in genetic diversity and transmission rates over time using a non-parametric Bayesian Skyline coalescent model to infer effective population sizes (N_e) [21] and a Bayesian birth-death skyline model to estimate the reproductive rate (R_e) [22]. We found N_e of 3a1 increased substantially after National Day (October 2020), and after a brief reduction in January-February 2021, increased again following the Spring Festival (February 2021) holiday, whereas the N_e of 3a2 steadily increased from October 2020 to August 2021 **(Fig. 3)**.

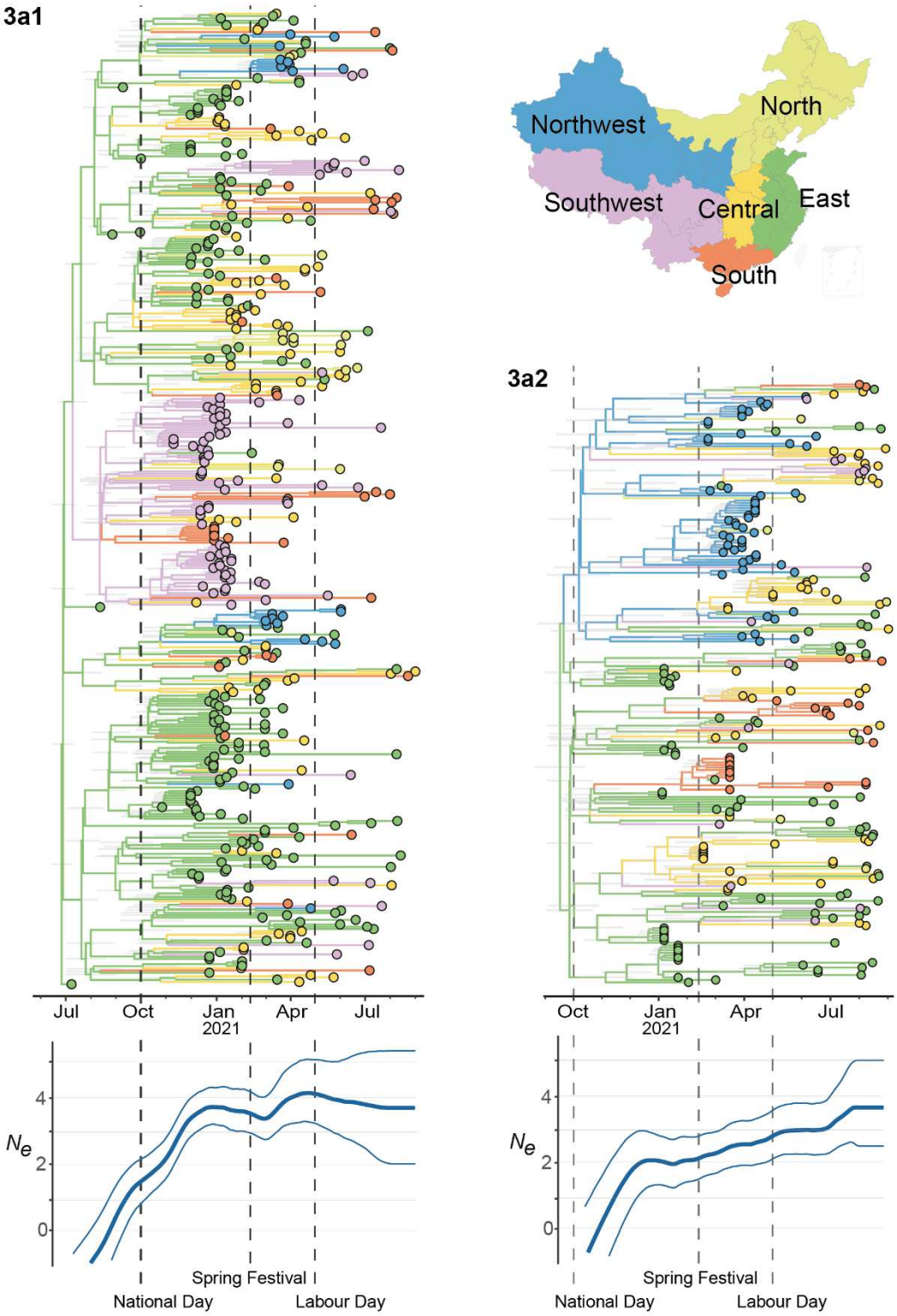

Fig. 3 | Phylodynamic analysis of influenza B/Victoria clades 3a1 and 3a2 in China during 2020-2021. Maximum clade credibility trees colored by geographic regions. Grey bars at nodes denote time to most recent common ancestor (tMRCA). Curves under the trees show change in effective population size (N_e) with 95% confidence intervals. Vertical dashed lines represent major holidays in mainland China (>5 days) including National Day (1-7 October 2020), Spring Festival (11-17 February 2021), and Labour Day (1-5 May 2021).

Transmission rate estimates showed R_e of clade 3a1 increased from R_e = 0.7 (HPD, 0.1-1.4) in April 2020, to R_e = 1.5 (HPD, 0.9-2.3) in June 2020, and to R_e = 2.1 (HPD, 1.7-2.5) from September 2020 until January 2021, when it fell to ~0.9 **(Fig. 4)**. R_e of clade 3a2 fluctuated between ~0.5 and ~2, with peaks following National Day and Spring Festival holidays. Clade 3a2 R_e maintained a high of R_e = 2.1 (HPD, 1.2-3.1) from October to January which rebounded briefly after Spring Festival (during February-March) to R_e = 2.2 (HPD, 1.6-2.9), and peaked again following Labour Day (May-July) to R_e = 1.6 (HPD, 1.1-2.2). For V1A.3-China, N_e and R_e peaked in early 2019, gradually decreased and nearly died out after 2020 **(Fig. 4 and Extended Data Fig. 4)**.

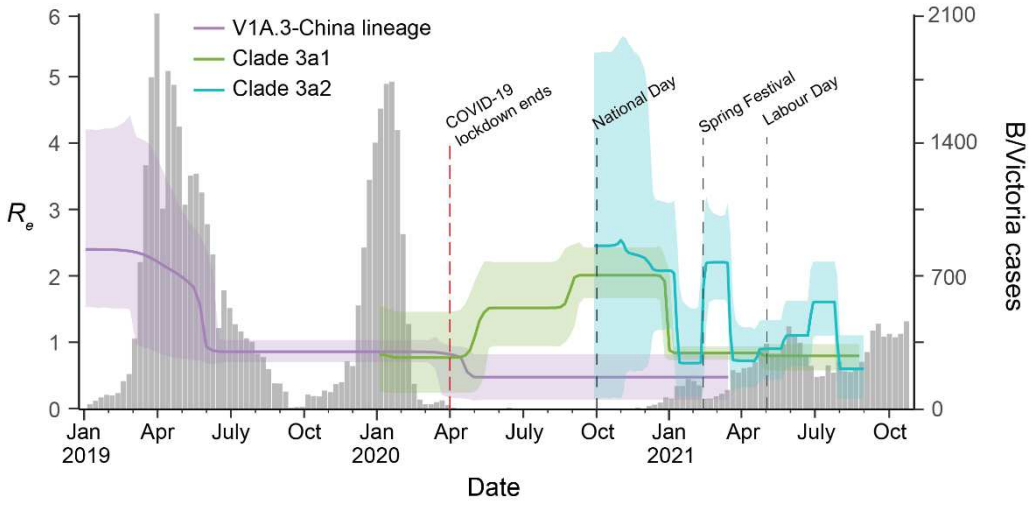

Fig. 4 | Effective reproductive rate (R_e) of B/Victoria 3a1, 3a2, and V1A.3-China lineages. Curves colored by clades (3a1, 3a2, and V1A.3-China) represent the mean effective reproductive rate (R_e) with 95% confidence intervals based on influenza B/Victoria positive cases per week reported from China to Global Influenza Surveillance and Response System (GISRS) [8]. Dashed red line indicates when the nationwide COVID-19 lockdown was lifted in April 2020 [9]. Vertical dashed grey lines represent major holidays in mainland China (>5 days) including National Day (1-7 October 2020), Spring Festival (11-17 February 2021), and Labour Day (1-5 May 2021).

Taken together, our results suggest genetic diversity and transmission rates increased soon after the relaxation of COVID-19 restrictions, earlier for 3a1 than 3a2, and nationwide holidays impacted 3a1 and 3a2 in subtly different ways, with greater effects on N_e of 3a1 and R_e of 3a2.

Phylogeography of B/Victoria in China

To identify seasonal epidemic sources and routes of dissemination we implemented a discrete-trait phylogeographic model utilising Bayesian Stochastic Search Variable Selection (BSSVS) and classified evidence as either 'definitive' (Bayes Factor (BF) > 100) or 'sufficient' (100 > BF > 3). To control for potential sampling bias, our results were confirmed by systematic subsampling (see Methods). Of the 30 possible migration routes between the six geographic regions, we identified four routes with definitive evidence and five with sufficient support for B/Victoria clade 3a1 dissemination **(Fig. 5 and Extended Data Table 3)**. We identified seven definitive and five sufficient migration routes for clade 3a2, and three definitive and three sufficient migration routes for the V1A.3-China lineage **(Fig. 5 and Extended Data Table 3)**. Eastern China acted as a major source for 3a1, 3a2, and V1A.3-China epidemics detected nationwide during 2019-2021. Within eastern China, Jiangxi province was identified as a major epicentre and source for recent influenza B/Victoria virus dissemination. From the east, four migration routes (to central, south, north, and southwest regions) were observed both prior to and during COVID-19 related control measures **(Extended Data Table 3)**. However, new migration routes from northwest to southwest and southwest to central regions were observed for clades 3a1 and 3a2 **(Extended Data Table 3)**. Northern China had the least number of confirmed cases, serving primarily as a regional sink. Notably, a 3a1 lineage from the southwest (Sichuan, Chongqing, Yunnan, and Guizhou provinces) and a 3a2 lineage from the northwest (Gansu and adjacent provinces) acted as important secondary sources for inter-seasonal epidemics throughout 2021 **(Fig. 5)**. Overall, while 3a1 during COVID-19 and V1A.3 prior to COVID-19 were mainly driven by Eastern China, multiple provinces acted as major sources for 3a2 **(Fig. 5)**.

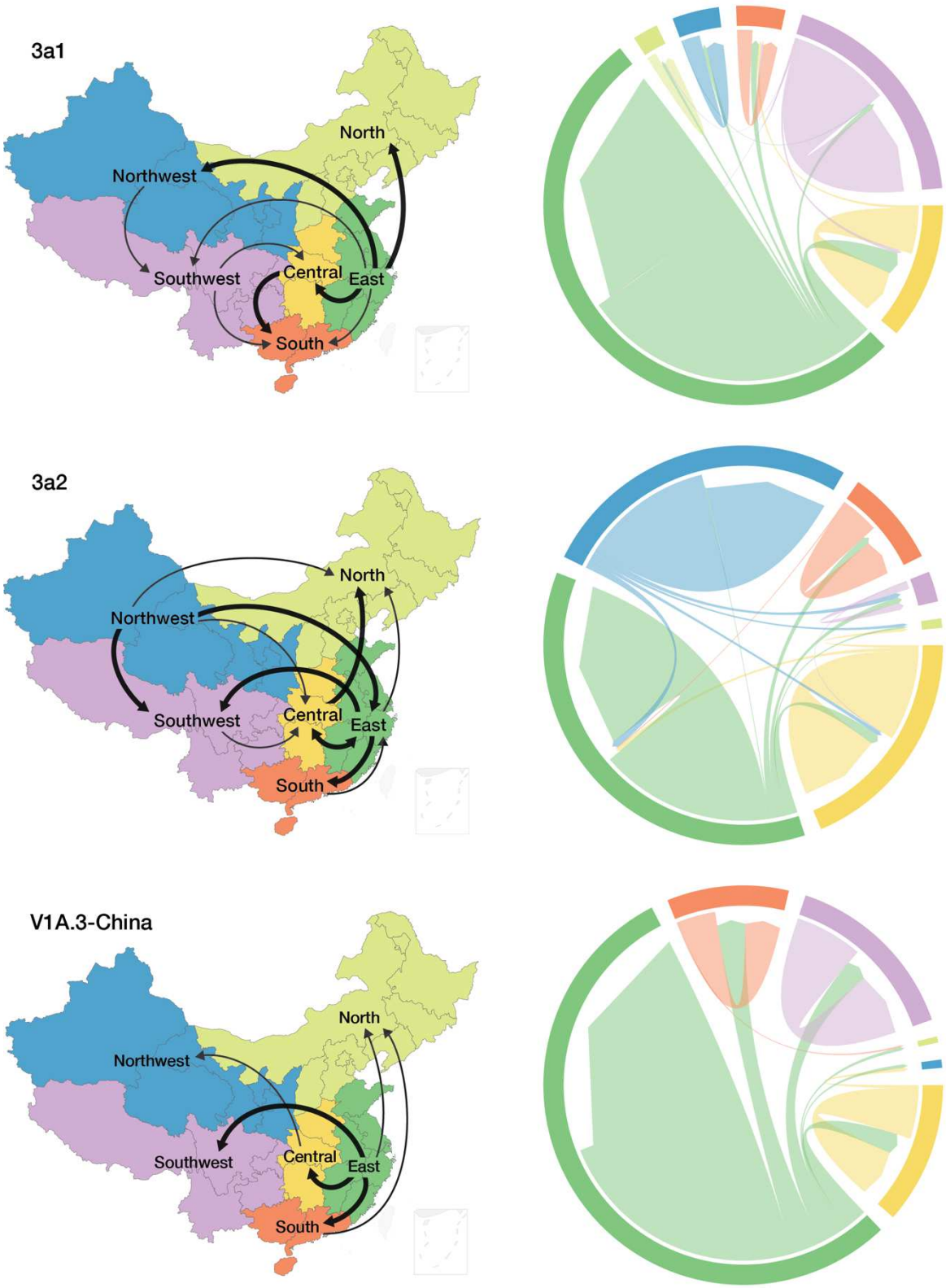

Fig. 5 | Transmission patterns of influenza B/Victoria 3a1, 3a2, and V1A.3-China lineages in China during 2019-2021. Transmission routes with definitive support (Bayes Factor > 100, thick arrows) and sufficient support (Bayes Factor > 3, thin arrows) are indicated on the left, and all migration routes within and between regions are shown by directional arrows in the circular flow diagrams on the right. Regions without reported cases are shown in light grey.

Altered age-distribution of influenza B virus in China

In the absence of selection pressure, demography and behavior of the susceptible host population play a role in viral transmission dynamics, such as social mixing patterns shown previously for global movement of influenza lineages [23], and the early transmission dynamics of SARS-CoV-2 in China [9]. By comparing age distribution of all sequenced human seasonal influenza cases in China before COVID-19 (from 2011 to 2019), we found that while children under 10 years of age accounted for the overwhelming majority of cases regardless of subtype, the age distribution of B/Victoria showed bimodal peaks, with an increased number of cases detected among individuals 20-30 years of age and fewer cases among 10-20 year-olds in comparison to other influenza viruses **(Fig. 6a, Extended Data Fig. 5)**. During 2020-2021, the proportion of infections in individuals over 10 years of age increased from less than 35% to almost 45% **(Fig. 6b)**, which may be attributable to altered mixing patterns between demographics and improvements in testing coverage. A comparison of infection by year of birth showed fewer B/Victoria infections among the early 2000s birth cohort, indicating greater protection against B/Victoria **(Fig. 6a)**.

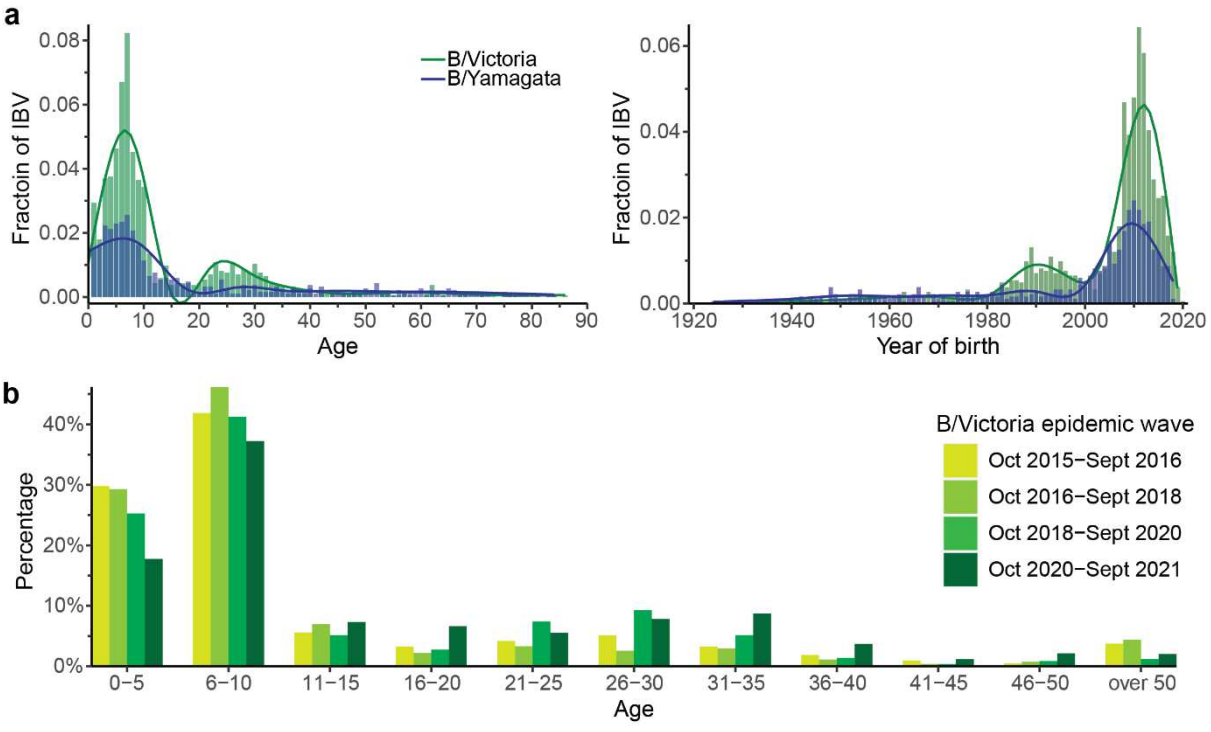

Fig. 6 | Age distribution of sequenced influenza B virus (IBV) cases in China. a, Distribution of cases during 2011-2019 by age at time of infection (in years) and year of birth. The fraction of cases was calculated relative to all sequenced influenza B cases. **b**, B/Victoria age distribution across four seasonal waves of B/Victoria epidemics in China illustrated in **Extended Data Fig. 1**.

Discussion

Since the emergence of SARS-CoV-2 in late 2019, nearly all influenza cases in China have been attributable to B/Victoria lineage viruses. Though A(H3N2), A(H1N1), and B/Yamagata viruses were detected during late 2019 and early 2020, A(H1N1) and A(H3N2) have since been only sporadically detected, and nearby A(H3N2) epidemics in Southeast Asia have not seeded local outbreaks as international borders have remained closed to non-resident travel. Multiple B/Victoria clades survived despite intermittent COVID-19 control measures through persistent regional circulation, however, since late 2020, seasonal influenza epidemics in China have been caused by two B/Victoria clades, 3a1 and 3a2.

The spatiotemporal patterns of seasonal influenza spread have been intensively studied [24-26], yet estimation of phylodynamic patterns in individual countries or continents is often hampered by repeat introductions and cocirculation of multiple transmission lineages. As such, the closure to international travel and the isolated circulation of B/Victoria in mainland China during 2020-2021 offered a unique opportunity to record patterns of B/Victoria transmission in the absence of competition from B/Yamagata, A(H1N1), or A(H3N2) subtype viruses. We identified regions, especially the east, central, northwest, and southwest, capable of sustaining local B/Victoria transmission lineages over several months. Jiangxi province in the southeast acted as a major amplifier and source for seasonal influenza B virus transmission across China, while the northwest and southwest regions became secondary epicentres for sustained 3a2 and 3a1 epidemics, respectively. Prolonged control measures in northern provinces with local SARS-CoV-2 transmission [10], including Beijing, Heilongjiang, Xinjiang, Hebei, and Jilin may have contributed to the absence of influenza in these regions during COVID-19. Temporal phylodynamics suggest increases in N_e and reproductive rate R_e followed major holidays in 2020 and early 2021, yet outbreaks remained largely undetected by surveillance systems before December 2020. Taken alongside evidence of sustained low-level transmission in multiple provinces, as shown for east, northwest, and southwest China, this underscores the importance of seasonal influenza vaccination prior to holiday gatherings or travel to limit epidemic spread. Although some precursor 3a1 and 3a2 viruses were detected in China as early as the second half of 2019, dissemination patterns during early 2020 remain elusive due to a lack of sequence records from early to mid-2020.

Influenza B lineages exhibit contrasting age distribution patterns [20]. Interestingly, the bimodal age distribution of B/Victoria observed in China from 2011-2019 (with the age 0-10 peak and age 20-30 subpeak) was analogously seen in B/Yamagata globally [20, 27]. The higher disease burden of B/Victoria viruses among adults aged 20-30 may have helped the persistence and spread of B/Victoria. Furthermore, B/Victoria viruses have evolved under greater selection pressure and accumulated beneficial substitutions in the HA gene at a faster rate than B/Yamagata [20]. Interestingly, previous clinical trials of influenza B virus vaccination showed B/Yamagata viruses elicit a stronger immune response than B/Victoria [28, 29] and infect a greater proportion of adults [20]. Our longitudinal comparison of age profiles among influenza viruses in China found a shift towards individuals above age 10 during 2020-2021, which may be attributable to changes in social contact patterns, improvements in surveillance coverage, changes in healthcare-seeking behavior, and/or loss of lineage competition. We also found that the cohort born in the early 2000s (15-20 year-olds) in China likely have stronger protection against influenza B/Victoria viruses. A modelling study of influenza age distribution from New Zealand showed additional protection against B/Yamagata in individuals immunologically imprinted with B/Yamagata as their first influenza virus infection [30]. However, definitive reasons for the abnormal distribution and shift of age profiles in China are still unclear and require further analysis and demonstration based on detailed epidemiological records and immunological studies.

Analysis of influenza B viruses collected during the past two decades [20, 31-33] showed B/Victoria and B/Yamagata maintain distinct HA and polymerase genes (PB2, PB1 and PA) [32] and show distinct differences in evolution and epidemiology. We identified a lineage-defining HA1 mutation, H122Q, among B/Victoria clade 3a2 viruses in China which was common among B/Yamagata viruses. A recent study by Huang et *al*. [15, 34] showed that the majority of B/Victoria viruses circulating in China during 2020-2021 are not well recognized by ferret antisera raised against the 2020-2021 vaccine strain B/Washington/02/2019 (clade V1A.3), which does not possess the H122Q mutation. Likewise, the current clade 3a2 vaccine strain, B/Austria/1359417/2021, lacks the H122Q mutation seen uniquely among 3a2 viruses in China. Therefore, we recommend the spread of H122Q B/Victoria viruses be monitored and impacts on age profiles measured upon potential introduction to susceptible populations, particularly amid the potential for reduced vaccine efficacy.

Our study has following limitations. (i) Our phylodynamic inference was based on the concatenated whole genome as there was insufficient temporal signal to analyse gene segments individually. To avoid potential bias introduced due to reassortment, we compared the phylogenetic relationships of each of the segments to confirm that the branching patterns and clades with high support were consistent across the segments. (ii) Studies have shown phylogeographic inference of viral movement can be affected by sampling bias [35, 36]. We therefore repeated our analysis on a systematically down-sampled dataset to confirm our results. (iii) We were unable to investigate the epidemiological dynamics during early to mid 2020 due to a lack of sequence data, and (iv) not all reported cases were sequenced, and lineage information was not available for all reported cases.

Overall, the genomic data analysed here provide important insights into the epidemiology of influenza viruses circulating in humans in China during a unique period of prolonged border closure during 2020-2021. In particular, we find distinct B/Victoria viruses that circulated during the 2019-2020 epidemic maintained low-level circulation in multiple regions within China during 2020, despite COVID-19 associated NPIs. Two of these B/Victoria lineages re-emerged in late 2020 to cause outbreaks of increasing intensity throughout 2021, while local circulation of A(H1N1), A(H3N2) and B/Yamagata virus lineages perished. We find that multiple provinces can maintain influenza transmission lineages for at least six months, Jiangxi province was recognized as a major epicentre during 2020-2021, and showed major holidays promoted lineage expansion and inter-provincial migration, which can inform regional and national control policies. While the observed shift in age-susceptibility profile requires further study, clade 3a2 B/Victoria viruses with the HA1 H122Q mutation should be monitored as travel measures are relaxed.

Online Methods

Dataset curation and phylogenetic analysis

We retrieved all available data on confirmed influenza B/Victoria cases in mainland China from January 2011 to November 2021 from the WHO Global Influenza Surveillance and Response System (GISRS) [8]. From the same study period, we also downloaded all available human seasonal influenza virus sequences from China (n = 2299) and a representative subset of global sequences (n = 955), including all B/Victoria virus vaccine reference strains, from GISAID (accessed 2022-02-17). After removing sequences with the same strain name, multiple

sequence alignments were performed for each gene segment using MAFFT (v.7.22) [37] and trimmed to the coding regions in AliView (v.1.27) [38]. Maximum likelihood phylogenetic analysis was performed using IQ-TREE (v.2) [39] with the best fit nucleotide substitution model identified by ModelFinder [40]. Branch support was assessed by an ultrafast bootstrap algorithm with 1000 replicates [41]. Phylogenetic clusters were labelled according to WHO clade designations. TempEst (v.1.5.3) [42] was used to investigate suitability of the dataset for estimation of phylodynamic parameters, and maximum likelihood dating of the HA gene phylogenetic tree and preliminary molecular clock analyses for each segment were performed using TreeTime [43]. A maximum of 250 sequences were randomly sampled from each major clade for phylogenetic analysis of recombination events.

Phylodynamic analysis

Whole genome sequences were concatenated for each lineage V1A.3-China (n = 140), 3a1 (n = 423), and 3a2 (n = 243) after removing temporal outliers according to the root-to-tip regression analysis using TempEst (v.1.5.3) [42] and potential reassortments within each lineage found by comparing phylogenetic topologies of each segment. The final dataset (Dataset I) was tested for reassortment using RDP5 [44] to apply six methods: RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, and 3Seq. Time-scaled phylogenetic analysis was conducted using BEAST (v.10.1.4) [45] with a GTR + G4 nucleotide substitution model, an uncorrelated, relaxed molecular clock with a lognormal rate distribution prior and a GMRF Bayesian Skyride coalescent model. A Bayesian Skyline coalescent model was additionally applied to estimate changes in the effective population size (N_e) for clades 3a1, 3a2, and V1A.3-China [21]. We also conducted an asymmetric discrete-trait phylogeographic analysis with Bayesian Stochastic Search Variable Selection (BSSVS) in BEAST (v.10.1.4) [45] to reconstruct spatial diffusion. Chinese provinces were grouped into six geographic regions summarized in **Extended Data Table 3**. Finally, a birth-death skyline serial (BDSS) model [22] was implemented in BEAST (v.2.6.3) [46] to estimate changes in the effective reproductive rate (R_e) over time for each lineage. To minimize the impact of regional sampling bias on phylogeographic analysis, we repeated the analysis for 3a1 and 3a2 using a subset containing < 10 randomly sampled sequences per province per month **(Extended Data Table 3 and Extended Data Fig. 6)**. Markov Chain Monte Carlo (MCMC) runs were conducted at least twice each for 50 to 100 million steps, sampling every 1,000 to 5,000 steps, with 10% discarded as burn-in. Sufficient sampling (ESS > 200) from the posterior was inspected by

Tracer (v.1.7.1) [47]. Tree visualizations and circular plots were created with R packages "ggtree" [48] and "circlize" [49]. Map data in this study was downloaded from a public source (https://www.highcharts.com/).

Acknowledgements

The authors are extremely grateful for local, provincial and national CDCs in China for sample collection and sequencing, and uploading of sequences during difficult times of the covid pandemic, and thank all National Influenza Centres and laboratories that supply influenza virus surveillance data to the WHO GISRS, as well as the authors and laboratories who contribute genetic sequence data via GISAID. A complete list of GISAID acknowledgements is provided in Extended Data Table 4.

Author Contributions

V.D. and B.J.C. conceived the study. V.D. designed and supervised the study. R.X. analysed the data. R.X., D.C.A., K.M.E. and V.D. interpreted the results. R.X., K.M.E. and V.D. wrote the paper with input from all authors.

Funding

This study is supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services of the United States, under contract number 75N93021C00016, Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. T11-712/19-N), as well as a postgraduate research scholarship from the University of Hong Kong, China.

Competing Interests statement

The authors declare no conflicts of interest.

References

1. Woolthuis, R.G., J. Wallinga, and M. van Boven, *Variation in loss of immunity shapes influenza epidemics and the impact of vaccination.* BMC Infect Dis, 2017. **17**(1): p. 632.
2. Koutsakos, M., et al., *Influenza lineage extinction during the COVID-19 pandemic?* Nat Rev Microbiol, 2021. **19**(12): p. 741-742.
3. Vijaykrishna, D., et al., *Human seasonal influenza under COVID-19 and the potential consequences of influenza lineage elimination.* Nature Portfolio, 2021.
4. WHO, *Recommended composition of influenza virus vaccines for use in the 2022-2023 northern hemisphere influenza season.* World Health Organization, 25 Feburary 2022.
5. Bolton, M.J., et al., *Antigenic and virological properties of an H3N2 variant that will likely dominate the 2021-2022 Northern Hemisphere influenza season*. 2021, Cold Spring Harbor Laboratory.
6. Neher, R.A. and T. Bedford, *nextflu: real-time tracking of seasonal influenza virus evolution in humans.* Bioinformatics, 2015. **31**(21): p. 3546-8.
7. Center, C.N.I., *Weekly Report, http://www.chinaivdc.cn/cnic/en/Surveillance/WeeklyReport/202112/P020211231823733822502.pdf.* 2021.
8. Hay, A.J. and J.W. McCauley, *The WHO global influenza surveillance and response system (GISRS)-A future perspective.* Influenza Other Respir Viruses, 2018. **12**(5): p. 551-557.
9. Kraemer, M.U.G., et al., *The effect of human mobility and control measures on the COVID-19 epidemic in China.* Science, 2020. **368**(6490): p. 493-497.
10. Yuxi Zhang, T.H., Andrew Wood, Toby Phillips, Anna Petherick, Martina Di Folco, Ke Meng, Tianren Ge, *Chinese provincial government responses to COVID-19.* BSG Working Paper Series, 2021.
11. Center, C.N.I., *Weekly Report, http://www.chinaivdc.cn/cnic/en/Surveillance/WeeklyReport/202002/t20200209_212402.htm.* 2020.
12. Lau, H., et al., *The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China.* J Travel Med, 2020. **27**(3).
13. Feng, L., et al., *Impact of COVID-19 outbreaks and interventions on influenza in China and the United States.* Nat Commun, 2021. **12**(1): p. 3249.
14. Weijuan, H., et al., *Characterization of Influenza Viruses — China, 2019–2020.* 2020. **2**(44): p. 856-861.
15. Weijuan, H., et al., *Epidemiological and Virological Surveillance of Seasonal Influenza Viruses — China, 2020–2021.* 2021. **3**(44): p. 918-922.
16. Shu, Y. and J. McCauley, *GISAID: Global initiative on sharing all influenza data - from vision to reality.* Euro Surveill, 2017. **22**(13).
17. WHO, *Recommended composition of influenza virus vaccines for use in the 2022 southern hemisphere influenza season.* World Health Organization, 24 September 2021: p. 1-13.
18. Shen, J., et al., *Diversifying selective pressure on influenza B virus hemagglutinin.* J Med Virol, 2009. **81**(1): p. 114-24.

19. Wang, Q., et al., *Crystal structure of unliganded influenza B virus hemagglutinin.* J Virol, 2008. **82**(6): p. 3011-20.
20. Vijaykrishna, D., et al., *The contrasting phylodynamics of human influenza B viruses.* Elife, 2015. **4**: p. e05055.
21. Drummond, A.J., et al., *Bayesian coalescent inference of past population dynamics from molecular sequences.* Mol Biol Evol, 2005. **22**(5): p. 1185-92.
22. Stadler, T., et al., *Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV).* Proc Natl Acad Sci U S A, 2013. **110**(1): p. 228-33.
23. Bedford, T., et al., *Global circulation patterns of seasonal influenza viruses vary with antigenic drift.* Nature, 2015. **523**(7559): p. 217-20.
24. Htwe, K.T.Z., et al., *Phylogeographic analysis of human influenza A and B viruses in Myanmar, 2010-2015.* PLoS One, 2019. **14**(1): p. e0210550.
25. Pollett, S., et al., *Phylogeography of Influenza A(H3N2) Virus in Peru, 2010-2012.* Emerg Infect Dis, 2015. **21**(8): p. 1330-8.
26. Cheng, X., et al., *Epidemiological dynamics and phylogeography of influenza virus in southern China.* J Infect Dis, 2013. **207**(1): p. 106-14.
27. Horm, S.V., et al., *Epidemiological and virological characteristics of influenza viruses circulating in Cambodia from 2009 to 2011.* PLoS One, 2014. **9**(10): p. e110713.
28. Skowronski, D.M., et al., *Cross-lineage influenza B and heterologous influenza A antibody responses in vaccinated mice: immunologic interactions and B/Yamagata dominance.* PLoS One, 2012. **7**(6): p. e38929.
29. Skowronski, D.M., et al., *Randomized controlled trial of dose response to influenza vaccine in children aged 6 to 23 months.* Pediatrics, 2011. **128**(2): p. e276-89.
30. Vieira, M.C., et al., *Lineage-specific protection and immune imprinting shape the age distributions of influenza B cases.* Nat Commun, 2021. **12**(1): p. 4313.
31. Virk, R.K., et al., *Divergent evolutionary trajectories of influenza B viruses underlie their contemporaneous epidemic activity.* Proc Natl Acad Sci U S A, 2020. **117**(1): p. 619-628.
32. Dudas, G., et al., *Reassortment between influenza B lineages and the emergence of a coadapted PB1-PB2-HA gene complex.* Mol Biol Evol, 2015. **32**(1): p. 162-72.
33. Langat, P., et al., *Genome-wide evolutionary dynamics of influenza B viruses on a global scale.* PLoS Pathog, 2017. **13**(12): p. e1006749.
34. *Recommended composition of influenza virus vaccines for use in the 2021-2022 northern hemisphere influenza season.* World Health Organization, 2021.
35. Kalkauskas, A., et al., *Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk.* PLoS Comput Biol, 2021. **17**(1): p. e1008561.
36. De Maio, N., et al., *New Routes to Phylogeography: A Bayesian Structured Coalescent Approximation.* PLoS Genet, 2015. **11**(8): p. e1005421.
37. Katoh, K. and D.M. Standley, *MAFFT multiple sequence alignment software version 7: improvements in performance and usability.* Mol Biol Evol, 2013. **30**(4): p. 772-80.
38. Larsson, A., *AliView: a fast and lightweight alignment viewer and editor for large datasets.* Bioinformatics, 2014. **30**(22): p. 3276-8.
39. Nguyen, L.T., et al., *IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies.* Mol Biol Evol, 2015. **32**(1): p. 268-74.
40. Kalyaanamoorthy, S., et al., *ModelFinder: fast model selection for accurate phylogenetic estimates.* Nat Methods, 2017. **14**(6): p. 587-589.

41. Hoang, D.T., et al., *UFBoot2: Improving the Ultrafast Bootstrap Approximation.* Mol Biol Evol, 2018. **35**(2): p. 518-522.
42. Rambaut, A., et al., *Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen).* Virus Evol, 2016. **2**(1): p. vew007.
43. Sagulenko, P., V. Puller, and R.A. Neher, *TreeTime: Maximum-likelihood phylodynamic analysis.* Virus Evol, 2018. **4**(1): p. vex042.
44. Martin, D. and E. Rybicki, *RDP: detection of recombination amongst aligned sequences.* Bioinformatics, 2000. **16**(6): p. 562-3.
45. Drummond, A.J. and A. Rambaut, *BEAST: Bayesian evolutionary analysis by sampling trees.* BMC Evol Biol, 2007. **7**: p. 214.
46. Bouckaert, R., et al., *BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis.* PLoS Comput Biol, 2019. **15**(4): p. e1006650.
47. Rambaut, A., et al., *Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7.* Syst Biol, 2018. **67**(5): p. 901-904.
48. Yu, G., *Using ggtree to Visualize Data on Tree-Like Structures.* Curr Protoc Bioinformatics, 2020. **69**(1): p. e96.
49. Gu, Z., et al., *circlize Implements and enhances circular visualization in R.* Bioinformatics, 2014. **30**(19): p. 2811-2.
50. Wood, S.N., *Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models.* Journal of the Royal Statistical Society Series B-Statistical Methodology, 2011. **73**: p. 3-36.

Extended Data Figures

Extended Data Fig. 1 | Seasonal influenza surveillance in China, 2011-2021. Stacked histogram (left-hand y-axis) represents the number of influenza-positive cases per week colored by subtype in China. Black trend-line (right-hand y-axis) shows the number of specimens tested per week.

Extended Data Fig. 2 | Maximum likelihood phylogenetic tree of the HA gene of influenza B/Victoria viruses, with viruses from China colored by year. Tips without color represent B/Victoria viruses from outside China. Scale bar shows nucleotide substitutions per site, and clade designations are indicated on the right hand side.

Extended Data Fig. 3 | Evolutionary relationship of AH1N1), A(H3N2), B/Victoria V1A and V1A.3 HA clades circulating in China since mid-2020. Samples collected in China since May 2020 are highlighted in red. The presented subtrees of A(H1N1) and A(H3N2) are monophyletic lineages pruned from a large-scale phylogeny of all available HA sequences in GISAID from January 2018 to August 2021. The presented subtrees of B/Victoria are monophyletic lineages pruned from **Extended Data Fig. 2**.

Extended Data Fig. 4 | Phylodynamic analysis of influenza B/Victoria V1A.3-China lineage genomes from 2019-2021. Above, maximum clade credibility tree colored by geographic region of origin. Grey bars denote estimated time to most recent common ancestor (tMRCA). Below, the effective population size (N_e) with 95% confidence intervals.

Extended Data Fig. 5 | Age distribution of human seasonal influenza sequenced cases in China during 2011 to 2019. The regression line was generated by a generalized additive model (GAM) [50].

Extended Data Fig. 6 | Geographic distribution of B/Victoria sequences analysed in this study.

Extended Data Tables

Extended Data Table 1 | Summary of amino acid diversity of B/Victoria and B/Yamagata at HA1 site 122.

Extended Data Table 2 | Molecular evolution estimates for each gene segment and concatenated whole genomes for 3a1, 3a2, and V1A.3-China lineages using TreeTime and Bayesian phylodynamic analysis.

Extended Data Table 3 | The geographic distribution of influenza B/Victoria sequences since 2019 and summary of geographical dissemination routes for 3a1, 3a2, and V1A.3-China lineages.

Extended Data Table 4 | Acknowledgements for sequences obtained from GISAID (accessed on 17-Feb-2022).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ExtendedDataTable3.xlsx
- ExtendedDataFig1.pdf
- ExtendedDataFig2.pdf
- ExtendedDataFig3.pdf
- ExtendedDataFig4.pdf
- ExtendedDataTable4.xlsx
- ExtendedDataTable1.xlsx
- ExtendedDataFig6.pdf
- ExtendedDataTable2.xlsx
- ExtendedDataFig5.pdf