Novel Avian Influenza Virus (H5N1) Clade 2.3.4.4b Reassortants in Migratory Birds, China

Jing Yang,[1] Chunge Zhang,[1] Yue Yuan,[1] Ju Sun,[1] Lu Lu,[1] Honglei Sun, Heting Sun, Dong Chu, Siyuan Qin, Jianjun Chen, Chengbo Zhang, Xiyan Hao, Weifeng Shi, Wenjun Liu, George F. Gao, Paul Digard, Samantha Lycett, Yuhai Bi

Two novel reassortant highly pathogenic avian influenza viruses (H5N1) clade 2.3.4.4b.2 were identified in dead migratory birds in China in November 2021. The viruses probably evolved among wild birds through different flyways connecting Europe and Asia. Their low antigenic reaction to vaccine antiserum indicates high risks to poultry and to public health.

Since the Gs/GD/96-lineage highly pathogenic avian influenza virus (HPAIV) (H5N1) was identified in 1996, H5 HPAIVs have evolved into divergent clades and caused continuous outbreaks in birds (*1*–*11*). Moreover, long-distance transmissions of H5 HPAIVs within a relatively short period indicate a crucial role of migratory birds in global spread of HPAIVs (*7*,*8*). Thus far, H5 viruses have undergone at least 4 waves of intercontinental transmission: H5N1 clade 2.2 during 2005–2006, H5N1 clade 2.3.2.1c during 2009–2010, H5N8 clade 2.3.4.4a and H5N1 clade 2.3.2.1c during 2014–2015, and H5Ny clade 2.3.4.4b during 2016–2017 (*2*–*8*).

Starting during 2020–2021, a new wave of HPAIV H5N1/H5N8 clade 2.3.4.4b outbreaks was reported in wild and domestic birds in Eurasia (*9*–*11*) and Africa (https://wahis.woah.org/#/event-management). Human cases of H5N1/H5N6/H5N8 infection were sporadically documented (https://www.who.int/teams/global-influenza-programme/avian-influenza/monthly-risk-assessment-summary), highlighting the zoonotic risk of H5 HPAIVs. Since 2021, H5 HPAIVs have caused at least 9 outbreaks in wild birds rather than poultry in mainland China (http://www.moa.gov.cn/gk/yjgl_1/yqfb; http://www.xmsyj.moa.gov.cn/yqfb). However, large outbreaks of H5N1 HPAIVs in domestic poultry were reported during 2021–2022 in Europe and the United States (https://www.cdc.gov/flu/avianflu/data-map-commercial.html) (A. Kandeil et al., unpub. data, https://doi.org/10.21203/rs.3.rs-2136604/v1). In this study, we explored the genetic origin, spread patterns, and antigenicity of H5N1 viruses identified from 2 dead migratory birds in China.

The Study

We collected oral swab specimens and lung tissues from a dead whooper swan in northern China (Inner Mongolia) on November 3, 2021, and a deceased black swan in eastern China (Zhejiang) on November 15, 2021. We performed virus isolation in 10-day-old specific pathogen-free chicken embryos (*12*), then confirmed results by quantitative reverse transcription PCR (Mabsky Biotech, http://www.mabsky.com).

We isolated and Sanger sequenced 3 viruses, A/whooper swan/Northern China/11.03 IMEEDSAK1-O/2021 (Ws/NC/AK1-O/2021), A/whooper swan/Northern China/11.03 IMEEDSAK2-O/2021

Author affiliations: Institute of Microbiology, Center for Influenza Research and Early-warning (CASCIRE), Chinese Academy of Sciences–The World Academy of Sciences Center of Excellence for Emerging Infectious Diseases, Chinese Academy of Sciences, Beijing, China (J. Yang, Chunge Zhang, J. Sun, W. Liu, G.F. Gao, Y. Bi); University of Chinese Academy of Sciences, Beijing (J. Yang, Chunge Zhang, W. Liu, G.F. Gao, Y. Bi); Shandong First Medical University, Taian, China (Y. Yuan, W. Shi, Y. Bi); Shanxi Agricultural University, Taigu, China (J. Sun, Y. Bi); University of Edinburgh, Edinburgh, UK (L. Lu, P. Digard, S. Lycett); China Agricultural University, Beijing (Honglei Sun); State Forestry and Grassland Administration, Shenyang, China (Heting Sun, D. Chu, S. Qin); Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, China (J. Chen); Ordos Forestry and Grassland Development Center, Ordos, China (Chengbo Zhang); Hohhot Center for Disease Control and Prevention, Hohhot, China (X. Hao)

DOI: https://doi.org/10.3201/eid2906.221723

[1]These authors contributed equally to this article.

(Ws/NC/AK2-O/2021), and A/black swan/Eastern China/11.15 ZJHZ74-Lg/2021 (Bs/EC/74-Lg/2021). We deposited whole genomes in NMDC (https://nmdc.cn; accession nos. NMDCN0000RD8–NMDCN0000RDV) and GISAID (https://www.gisaid.org; accession nos. EPI195500–EPI195523). We reconstructed phylogenetic trees for each gene of the 3 H5N1 isolates together with reference viruses from GISAID and the National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi), using the maximum-likelihood method with a general time-reversible model plus gamma distribution in RAxML 8.2.12 (https://cme.h-its.org/exelixis/web/software/raxml) (Appendix 1 Table, https://wwwnc.cdc.gov/EID/article/29/6/22-1723-App1.xlsx). We reconstructed Bayesian time-resolved phylogenetic trees in BEAST 1.10.4 (https://beast.community/index.html) using the SRD06 model, the log-normal relaxed clock model, and the Skygrid coalescent model. We mapped spatial coordinates to the post burn-in time-scaled posterior trees using a Brownian motion continuous phylogeographic model. We mapped host type and hemagglutinin (HA) or neuraminidase (NA) subtype on each posterior tree by using a discrete trait phylogeographic model with BSSVS extension to infer the most likely ancestor with statistical support (Appendix 2, https://wwwnc.cdc.gov/EID/article/29/6/22-1723-App2.pdf).

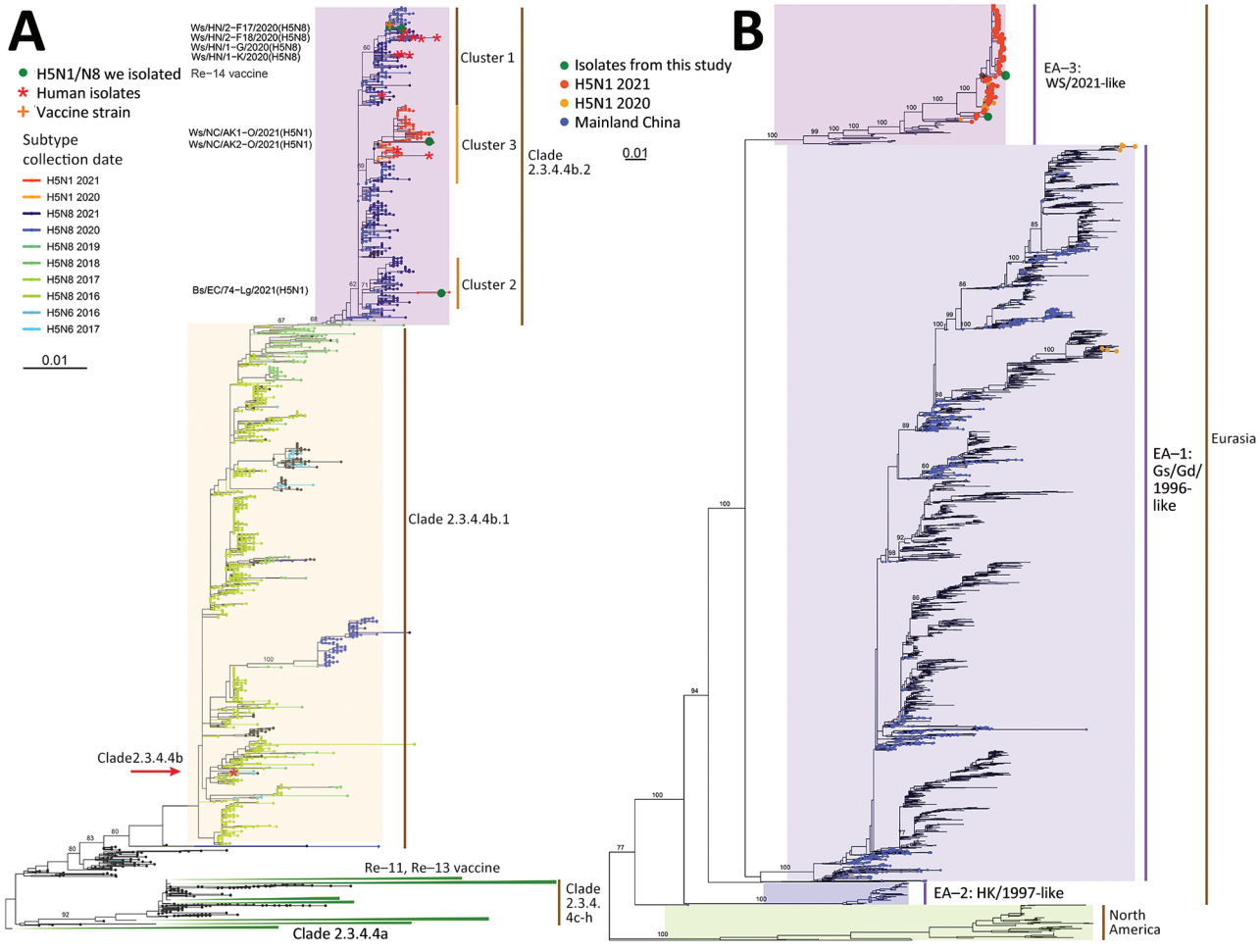


Figure 1. Phylogenetic trees for hemagglutinin genes of clade 2.3.4.4 H5Ny and neuraminidase genes of global H5N1 avian influenza viruses. A) Phylogeny of hemagglutinin genes of global clade 2.3.4.4 H5Ny avian influenza viruses. Solid green circles indicate H5N1 and H5N8 viruses from wild birds isolated in China; sequence names are listed next to corresponding circles. Red asterisks indicate human isolates. H5 vaccine seed strains used in mainland China are listed near the corresponding clades; orange cross indicates Re-14 vaccine. The major H5Ny subtypes within clade 2.3.4.4b are colored by their subtypes and collection dates. Clade 2.3.4.4b was divided into clade 2.3.4.4b.1 and clade 2.3.4.4b.2 because of >2.7% average pairwise nucleotide distance and >60% support for the 2 subclades. B) Phylogeny of global H5N1 neuraminidase genes. Colored circles indicate the novel H5N1 viruses from this study, H5N1 viruses from mainland China, and H5N1 viruses isolated in 2020 and 2021. Numbers on branches represent bootstrap support values for some major clades. Scale bar indicates number of nucleotide substitutions per site. Full phylogenetic trees of hemagglutinin genes of global H5Ny and neuraminidase genes of global H5N1 are provided at https://github.com/judyssister/globalH5N1_2021.

Table 1. HI titers of H5N1 and H5N8 highly pathogenic avian influenza viruses from wild birds in China against antiserum of H5 Re-11, Re-13, and Re-14 vaccines*

Virus	HI titers of chicken antiserum against vaccine strains and H5N1/H5N8 isolates		
	Re-11, clade2.3.4.4h	Re-13, clade2.3.4.4h	Re-14, clade2.3.4.4b
Re-11, Dk/GZ/S4184/2017(H5N6)	256	256–512	128
Re-13, Dk/FJ/S1424/2020(H5N6)	32	256	2–4
Re-14, Ws/SX/4–1/2020(H5N8)	8	16	256
Bs/EC/74-Lg/2021(H5N1)	16	8	128
Ws/NC/AK2-O/2021(H5N1)	2–4	2–4	32
Ws/HN/1-K/2020(H5N8)	8	8	64
Ws/HN/1-G/2020(H5N8)	8	8	64

*H5 Re-11 vaccine was used in poultry in mainland China during December 2018–December 2021. H5 Re-13 and Re-14 vaccines have been deployed since January 2022. HI, hemagglutination inhibition.

We performed hemagglutination inhibitor (HI) assays (https://www.who.int/publications/i/item/manual-for-the-laboratory-diagnosis-and-virological-surveillance-of-influenza) to test the reactivities of antiserum of H5 Re-11/Re-13/Re-14 vaccines against these new H5N1 isolates and H5N8 HPAIVs identified in 2020 (*10*). Re-11 (A/duck/Guizhou/S4184/2017[H5N6], clade 2.3.4.4h) was used in poultry in China during December 2018–December 2021, whereas Re-13 (A/duck/Fujian/S1424/2020[H5N6], clade 2.3.4.4h) and Re-14 (A/whooper swan/Shanxi/4–1/2020[H5N8], clade 2.3.4.4b) have been deployed since January 2022 (http://www.moa.gov.cn/govpublic).

Conclusions

We obtained 3 H5N1 HPAIVs, Ws/NC/AK1-O/2021 and Ws/NC/AK2-O/2021 from a dead whooper swan in northern China and Bs/EC/74-Lg/2021 from a dead black swan in eastern China in November 2021. Consistent with the HPAIV signature of multiple basic amino acids on HA cleavage site, these H5N1 strains caused severe histopathologic changes in the wild birds (Appendix 2 Figure 1).

Phylogenetic analyses showed that all 3 H5N1 HA genes cluster in clade 2.3.4.4b.2 (Figure 1, panel A). Most H5 avian influenza viruses (AIVs) identified during 2020–2021 were in that clade, whereas H5N8 was the dominant subtype during 2019–2021, and H5N1 strains emerged in October 2020 and increased subsequently. In NA phylogeny, most H5N1 viruses identified during 2020–2021 including those 3 H5N1 viruses, were classified into the Eurasian lineage clade EA-3 (Figure 1, panel B). However, almost all H5N1 NA genes from mainland China were identified during 1996–2018 and are clade EA-1.

Given the HA phylogenetic relationships, we defined cluster 1, 2, and 3 in clade 2.3.4.4b.2 (Figure 1, panel A). Cluster 1 includes 4 H5N8 HPAIVs identified from wild birds in 2020 (*10*) and Re-14 vaccine strain. In cluster 2, Bs/EC/74-Lg/2021 was grouped with H5N1 viruses from Japan and South Korea, showing 99.3%–99.6% sequence identity. In cluster 3, Ws/NC/AK1-O/2021 and Ws/NC/AK2-O/2021 are identical (Appendix 2 Table 1) and clustered with H5N1 viruses from Europe, possessing 99.3% nucleotide identity. Most H5N1 viruses identified during 2020–2021 belong to cluster 3. Notably, 8 H5N6 and 1 H5N8 viruses that caused human infections (*13*) are found in cluster 1. Moreover, 2 human infections with cluster 3 H5N1 viruses were reported in the United Kingdom and United States during 2021–2022 (*14*). Therefore, this virus lineage poses a nonnegligible threat to public health, despite these viruses carrying non–mammalian-adapted molecular markers (Appendix 2 Table 2) and avian-type receptor-binding propensity (Appendix 2 Figure 2).

Table 2. Amino acid substitutions on the hemagglutinin antigenic sites between H5N1 2021 and H5N8 2020 highly pathogenic avian influenza viruses and H5 vaccine seed viruses Re-11, Re-13, and Re-14 used in China*

Virus	Position of antigenic sites in hemagglutinin genes (H3 numbering)*															
	63	81	125	131	132	144	145	155	158	159	160	166	188	189	193	202
Re-11, clade2344h	D	R	R	T	S	V	A	T	N	D	A	M	A	E	N	V
Re-13, clade2344h	N	S	E	T	T	V	A	T	N	E	T	K	V	E	D	V
Re-14, clade2344b	D	R	S	E	T	A	P	I	N	D	A	I	A	E	N	I
Bs/EC/74-Lg/2021(H5N1)	D	R	N	E	T	A	P	I	N	D	A	I	A	K	D	I
Ws/NC/AK1-O/2021(H5N1)	D	R	S	E	T	A	P	I	D	D	A	I	A	K	N	I
Ws/NC/AK2-O/2021(H5N1)	D	R	S	E	T	A	P	I	D	D	A	I	A	K	N	I
Ws/HN/1-K/2020(H5N8)	D	R	S	E	T	A	P	I	N	D	A	I	A	E	N	I
Ws/HN/1-G/2020(H5N8)	D	R	S	E	T	A	P	I	N	D	A	I	A	E	N	I
Ws/HN/2-F17/2020(H5N8)	D	R	S	E	T	A	P	I	N	D	A	I	A	E	N	I
Ws/HN/2-F18/2020(H5N8)	D	R	S	E	T	A	P	I	N	D	A	I	A	E	N	I

*Positions of antigenic sites based on the H5 and H3 (sites A–E) antigenic sites.

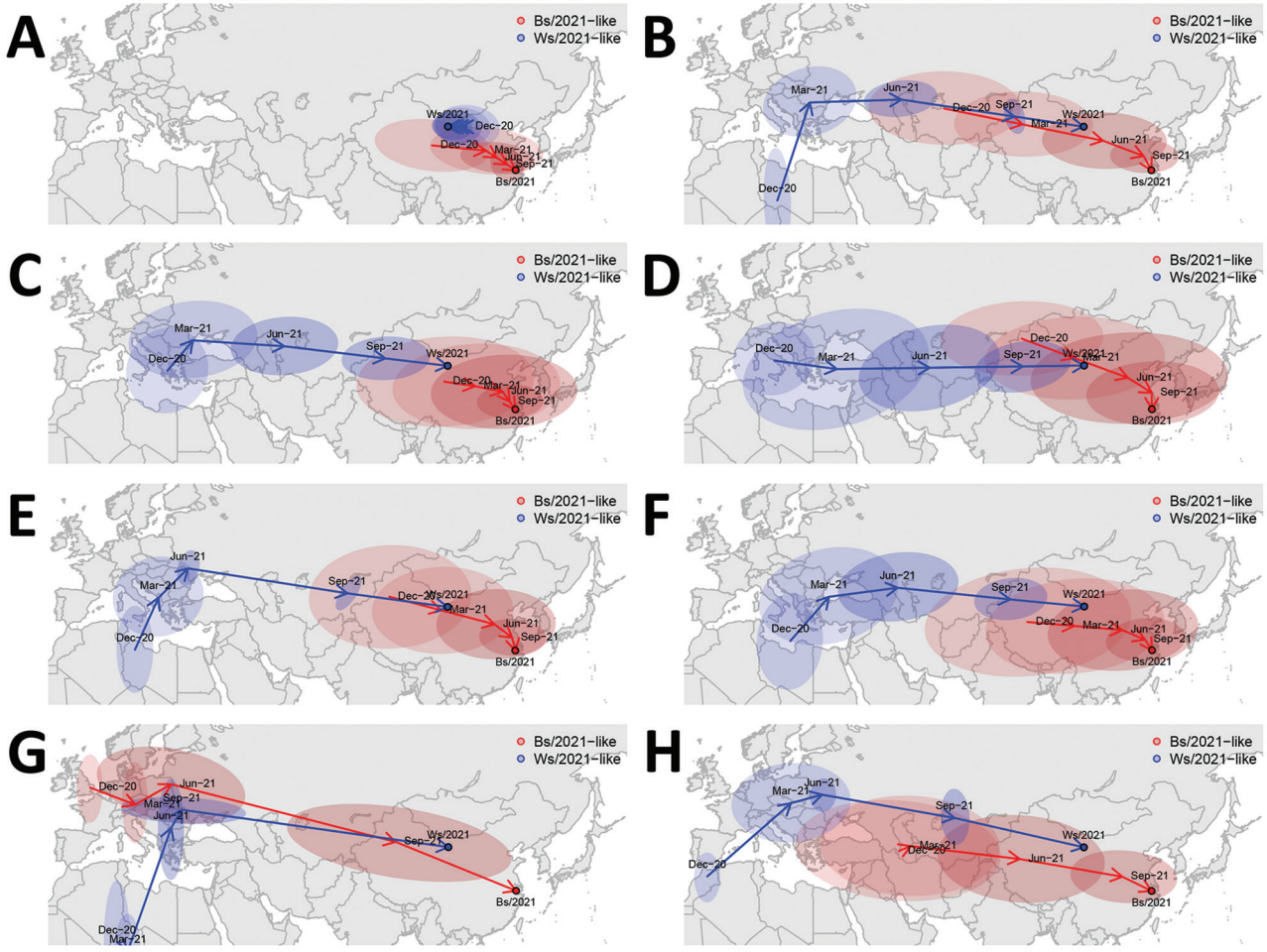


Figure 2. Spread patterns of all 8 gene segments of highly pathogenic avian influenza virus (H5N1), Bs/2021-like and Ws/2021-like reassortants, identified in migratory wild birds in China. Virus spread patterns reconstructed for 8 genes. A) polymerase basic 2 gene. B) polymerase basic 1 gene. C) polymerase acidic gene. D) hemagglutinin gene. E) nucleoprotein gene. F) neuraminidase gene. G) matrix gene. H) nonstructural protein gene. Blue indicates spread patterns of Ws/2021-like and red indicates spread patterns of Bs/2021-like H5N1. The spread patterns were adjusted by interpolating the ancestral space-time points by every 3 months from December 2020 through November 2021. Arrows represent the inferred ancestral locations at corresponding interpolated time (at 3-month intervals going back along their inferred transmission routes), and filled ellipses represent the 95% uncertainty of the inferred ancestral locations.

Current H5N1 viruses have resulted in substantial mortality in domestic and wild birds in Eurasia, Africa, and Americas (https://wahis.woah.org/#/event-management); however, they have only been identified in wild birds in mainland China. Compared with high HI antibody titers (256) between homologous antiserum and antigens of H5 vaccines, the recent H5N1/H5N8 viruses presented low HI titers (2–16) against Re-11/Re-13 antiserum (Table 1). In addition, cluster 1 H5N8 viruses had HI titers of 64 against Re-14 (cluster 1) antiserum, whereas HI titers for the H5N1 viruses were 128 for Bs/EC/74-Lg/2021 (cluster 2) and 32 for Ws/NC/AK1-O/2021 (cluster 3). This finding indicates lower antigenic identities between H5N1/H5N8 viruses circulating in wild birds and vaccines used in domestic poultry, even within the same clade. This antigenic variation may correlate to substitutions at antigenic sites (Table 2; Appendix 2 Tables 3,4, Figure 3).

Phylogenetic analyses uncovered that 3 novel H5N1 viruses could be classified into Ws/2021-like (Ws/NC/AK1-O/2021 and Ws/NC/AK2-O/2021) and Bs/2021-like (Bs/EC/74-Lg/2021) reassortants (Appendix 2 Figure 4). The viruses originated through separate reassortment events between H5N8 HPAIVs (obtaining HA and matrix [M] genes) and low pathogenic avian influenza virus pools (NA, polymerase basic 1, polymerase basic 2, polymerase acidic, nucleoprotein, and nonstructural protein genes) (Appendix 2 Figure 5–12). Phylogeographic

analyses suggested that the H5N1 viruses spread to China by long-distance bird migration through various routes (Figure 2).

We reconstructed the genetic reassortment history for these H5N1 viruses (Appendix 2 Table 5, Figures 13, 14). For the Bs/2021-like reassortant, most gene segments group with viruses from wild Anseriformes in China or its adjacent areas, whereas the M gene likely originated from an HPAIV H5N8 ancestor from Eastern Europe in approximately May 2021 before import into China through bird migration. For the Ws/2021-like reassortant, 7 gene segments originated from Europe and were potentially transmitted by wild birds in February–August 2021, whereas a unique polymerase basic 2 gene originated from an early ancestry in 2017 with unknown origin but most closely related to a Russian H3N6 low pathogenic AIV. The ancestral states of most genes of the 2 reassortants indicate origins in wild Anseriformes and likely transmission through wild Anseriformes over the summer of 2021, whereas a few genes (e.g., Bs/2021-like M gene) potentially originated from domestic poultry. However, sampling bias in sequences might affect ancestral reconstruction by discrete trait phylogeographic models.

In conclusion, we identified 3 H5N1 HPAIVs in wild birds in autumn 2021, China. The antigenic divergence highlights the high-risk introduction of H5N1 circulating in wild birds to incompletely protected vaccinated flocks in China. The H5N1 viruses have experienced complicated reassortment during long-distance spread through various bird migration routes. Therefore, we call for international cooperation on AIV monitoring in migratory birds to help early identification and intervention of the emerging and reemerging AIVs with public health risks.

Acknowledgments

We thank the submitters and originating laboratories for influenza virus genomes in the GISAID and National Center for Biotechnology Information databases.

This work was supported by the National Key R&D Program of China (grant nos. 2021YFC2300900 and 2022YFC2601602), National Natural Science Foundation of China (grant nos. 32261133524 and 32200416), Strategic Priority Research Program of Chinese Academy of Sciences (grant no. XDB29010102), Chinese Academy of Sciences' Southeast Asia Biodiversity Research Institute (grant no. 151C53KYSB20210023), Self-supporting Program of Guangzhou Laboratory (grant no. SRPG22-001), and the National Science and Technology Infrastructure of China (grant no. NPRC-32).

Y.B. is supported by the Youth Innovation Promotion Association of the Chinese Academy of Sciences (grant no. Y2021034) and Innovation Team and Talents Cultivation Program of National Administration of Traditional Chinese Medicine (grant no. ZYYCXTD-D-202208). L.L., S.L., P.D., and W.L. are supported by an Ecology and Evolution of Infectious Diseases collaborative grant with the UK Biotechnology and Biological Sciences Research Council (grant no. BB/V011286/1) and the National Natural Science Foundation of China (grant no. 32061123001). P.D. and S.L. are additionally supported by the Strategic Program grant to Roslin Institute (grant no. BB/P013740/1) and a UK research consortium on avian influenza research gaps (Flu-MAP) (grant no. BB/X006123/1) from the UK Biotechnology and Biological Sciences Research Council and Department for Environment Food and Rural Affairs. P.D. was also supported by an EU Horizon 2020 award (no. 727922 [DELTA-FLU]), and L.L and S.L are supported by an EU Horizon 2020 award (no: 874735 [VEO]).

About the Author

Dr. Yang is an assistant professor at the Institute of Microbiology, Chinese Academy of Sciences. Her research interests are focused on the evolution and spread patterns of emerging and reemerging infectious diseases.

References

1. World Health Organization, World Organization for Animal Health, Food and Agriculture Organization H5N1 Evolution Working Group. Revised and updated nomenclature for highly pathogenic avian influenza A (H5N1) viruses. Influenza Other Respir Viruses. 2014;8:384–8. https://doi.org/10.1111/irv.12230
2. Bi Y, Chen J, Zhang Z, Li M, Cai T, Sharshov K, et al. Highly pathogenic avian influenza H5N1 clade 2.3.2.1c virus in migratory birds, 2014-2015. Virol Sin. 2016;31:300–5. https://doi.org/10.1007/s12250-016-3750-4
3. Shi W, Gao GF. Emerging H5N8 avian influenza viruses. Science. 2021;372:784–6. https://doi.org/10.1126/science.abg6302
4. Liu J, Xiao H, Lei F, Zhu Q, Qin K, Zhang XW, et al. Highly pathogenic H5N1 influenza virus infection in migratory birds. Science. 2005;309:1206. https://doi.org/10.1126/science.1115273
5. Li Y, Liu L, Zhang Y, Duan Z, Tian G, Zeng X, et al. New avian influenza virus (H5N1) in wild birds, Qinghai, China. Emerg Infect Dis. 2011;17:265–7. https://doi.org/10.3201/eid1702.100732
6. Global Consortium for H5N8 and Related Influenza Viruses. Role for migratory wild birds in the global spread of avian influenza H5N8. Science. 2016;354:213–7. https://doi.org/10.1126/science.aaf8852
7. Lee DH, Torchetti MK, Hicks J, Killian ML, Bahl J, Pantin-Jackwood M, et al. Transmission dynamics of highly pathogenic avian influenza virus A(H5Nx) clade 2.3.4.4, North America, 2014–2015. Emerg Infect Dis. 2018;24:1840–8. https://doi.org/10.3201/eid2410.171891

8. Lycett SJ, Pohlmann A, Staubach C, Caliendo V, Woolhouse M, Beer M, et al.; Global Consortium for H5N8 and Related Influenza Viruses. Genesis and spread of multiple reassortants during the 2016/2017 H5 avian influenza epidemic in Eurasia. Proc Natl Acad Sci U S A. 2020;117:20814–25. https://doi.org/10.1073/pnas.2001813117
9. Sobolev I, Sharshov K, Dubovitskiy N, Kurskaya O, Alekseev A, Leonov S, et al. Highly pathogenic avian influenza A(H5N8) virus clade 2.3.4.4b, western Siberia, Russia, 2020. Emerg Infect Dis. 2021;27:2224–7. https://doi.org/10.3201/eid2708.204969
10. Li J, Zhang C, Cao J, Yang Y, Dong H, Cui Y, et al. Re-emergence of H5N8 highly pathogenic avian influenza virus in wild birds, China. Emerg Microbes Infect. 2021; 10:1819–23. https://doi.org/10.1080/22221751.2021.1968317
11. Okuya K, Mine J, Tokorozaki K, Kojima I, Esaki M, Miyazawa K, et al. Genetically diverse highly pathogenic avian influenza A(H5N1/H5N8) viruses among wild waterfowl and domestic poultry, Japan, 2021. Emerg Infect Dis. 2022;28:1451–5. https://doi.org/10.3201/eid2807.212586
12. Bi Y, Li J, Li S, Fu G, Jin T, Zhang C, et al. Dominant subtype switch in avian influenza viruses during 2016-2019 in China. Nat Commun. 2020;11:5909. https://doi.org/10.1038/s41467-020-19671-3
13. Pyankova OG, Susloparov IM, Moiseeva AA, Kolosova NP, Onkhonova GS, Danilenko AV, et al. Isolation of clade 2.3.4.4b A(H5N8), a highly pathogenic avian influenza virus, from a worker during an outbreak on a poultry farm, Russia, December 2020. Euro Surveill. 2021;26:2100439. https://doi.org/10.2807/1560-7917.ES.2021.26.24.2100439
14. Oliver I, Roberts J, Brown CS, Byrne AM, Mellon D, Hansen R, et al. A case of avian influenza A(H5N1) in England, January 2022. Euro Surveill. 2022;27:2200061. https://doi.org/10.2807/1560-7917.ES.2022.27.5.2200061

Address for correspondence: Yuhai Bi, Institute of Microbiology, Chinese Academy of Sciences, No. 1 Beichen West Rd, Chaoyang District, Beijing 100101, China; email: beeyh@im.ac.cn

Article DOI: https://doi.org/10.3201/eid2906.221723

Novel Avian Influenza Virus (H5N1) Clade 2.3.4.4b Reassortants in Migratory Birds, China

Appendix 2

Methods

Phylogenetic analyses

To uncover the potential genetic origin and evolutionary position of the new H5N1 viruses, we constructed phylogenetic trees for each of the eight gene segments given their nucleotide sequences. In dataset 1, we collected all HA nucleotide sequences of H5Ny viruses and all NA sequences of H5N1 viruses from the GISAID EpiFlu database (*1*). In dataset 2, we also collected the first 500 most genetically related sequences to all eight gene segments of each of the three H5N1 viruses identified in this study from the Influenza Virus Resource database at the NCBI (https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi) (*2*) and GISAID EpiFlu database by online BLAST tools. We combined these genetically similar sequences into a local database and ran a local BLAST against the local database to select the first 300 related sequences for the following analyses. We removed replicated sequences in each dataset according to DNA INSDC (international nucleotide sequence database collaboration) number and isolate names. We also removed sequences <75% full-length for each gene segment. We then aligned sequences of each gene segment with default parameters in MAFFT v7.490 (*3*). Maximum likelihood (ML) phylogenetic trees for each gene segment using the GTRGAMMA model with 1000 bootstraps in RAxML v8.2.12 were built (number of sequences 319–602 depending on segment, and 5789 sequences for NA gene of global H5N1), but for phylogeny

construction of global H5 (number of sequences is 13840), 100 bootstraps were used due to the computational complexity (*4*).

Phylogeographic analyses

We assessed the temporal signal of each gene segment in dataset 2 using the ML phylogenetic trees and collection dates of tree tips in Tempest v1.5.3 (http://tree.bio.ed.ac.uk/software/tempest/), resulting in datasets of 197–432 sequences spanning 2–20 years depending on the segment. The Bayesian time-resolved phylogenetic trees were reconstructed in BEAST 1.10.4 (*5*) using the SRD06 substitution model (an HKY substitution model and a site heterogeneity model using gamma distribution with four categories and two partitions for codon position 1+2 versus position 3) (*6*), an uncorrelated relaxed clock with a log-normal distribution, and a Skygrid coalescent model (*7*). A minimum of two independent Markov Chain Monte Carlo (MCMC) chains were run for each segment. Each chain consisted of 100,000,000 steps and was sampled every 10,000 steps, while the first 10% of samples were discarded as burn-in to achieve an effective sample size (ESS) >200. Later, a post burn-in subsample of 1,000 posterior trees for each segment was obtained for the location-annotated MCC tree visualization on the map. Spatial coordinates (latitudes and longitudes) were mapped to the time-scaled trees of each segment using a Brownian motion continuous phylogeographic diffusion model (*8*). In addition, host-type (containing domestic Anseriformes, domestic Galliformes, wild Anseriformes, other wild birds, mammals, and environment), HA subtype and NA subtype were also mapped on each tree using a discrete trait phylogeographic model with BSSVS extension to infer the most likely ancestor with statistical support (*9*). Phylogenetic trees were visualized and annotated in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). The MCC trees generated from the phylogeographic analyses are provided in a GitHub repository (https://github.com/judyssister/globalH5N1_2021).

Receptor binding assay

Receptor binding assays were performed using trisaccharide biotinylated glycan α-2,3-sialic acids (α2–3-SA) and α-2,6-sialic acids (α2–6-SA) receptors, respectively, according to previous studies (*10*). Two human strains, A/California/04/2009(H1N1) and A/Vietnam/1194/2004(H5N1), which have exclusive receptor binding preferences to α2–6-SA (human-type receptor) and α2–3-SA (avian-type receptor), respectively, were used as controls.

Histopathologic analysis of a dead black swan caused by H5N1 AIV

The lung, liver, and small intestine tissues of a freshly dead black swan infected by the H5N1 virus were fixed in 4% paraformaldehyde at room temperature. Serial coronal 4 μm thick sections were obtained and later stained with hematoxylin and eosin (H&E) by the Beijing One-Bio Technology Co., Ltd. The pathological changes were examined by light microscopy and scanned by a KF-PRO-120 Digital Pathology Scanner.

Homology modeling

To understand the positions of amino acid substitutions at antigenic sites, we inferred the structure of the ectodomain of the HA molecule of Ws/NC/AK1-O/2021 by homology modeling using the crystal structure of A/duck/Eastern China/L0321/2010 HA (H5N2; PDB ID: 7DEB) as a template in SWISS-model (*11*). The inferred 3-dimensional structure was annotated and visualized in PyMOL v2.5.2 (http://pymol.org/).

Results

Genetic characteristics of H5N1 isolates from wild birds

The molecular information at known functional sites can provide predictive power for certain phenotypic characteristics (e.g., pathogenicity, transmissivity, and risk of human infections) for novel AIVs (*12*). Typically, the polybasic amino acids at the HA cleavage site are considered as a genetic standard to identify H5 and H7 HPAIVs (*13*). In this study, the Bs/2021-

like H5N1 reassortant possessed PLRERRRKRGL, and the Ws/2021-like H5N1 reassortant and our H5N8 from wild birds in 2020 possessed PLREKRRKRGL at the cleavage sites, which indicated that these viruses belong to HPAIVs (Appendix 2 Table 2).

The combination of Q226L and G228S mutations in the H5 HA can increase the binding ability to the human-like sialic acid receptor (*14*), while a Q226L substitution in combination with N224K can also change the binding preference from avian-like to human-like receptors (*15*). The possession of 224N, 226Q, and 228G in the HA genes of our three H5N1 viruses suggested the viruses would exclusively bind to the avian-type receptor. In addition, as few as five amino acid substitutions (H110Y, T160A, Q226L, and G228S in the HA gene, and E627K in the PB2 gene) for A/Indonesia/5/2005(H5N1), and four substitutions (N158D, N224K, Q226L, and T318I in the HA gene) for A/Vietnam/1203/2004(H5N1) can confer airborne transmission for the H5 viruses among ferrets (*15*,*16*). The H5N1 reassortants we describe here have two mammalian-transmissible signatures in the HA gene, 158D and 160A. This indicates these viruses would most likely not transmit between ferrets.

Amino acid deletions in the NA stalk of AIVs can contribute to better adaption to terrestrial birds for viruses introduced by wild waterfowl and increase virulence in mammals (*17–19*). NA stalk deletions were not present in these H5N1 reassortants, which suggested they have not adapted to terrestrial birds thus far. Additionally, Q591K, E627K, and D701N mutations in the PB2 gene can confer better adaption for AIVs to mammals, such as boosting replication in mammalian cells and increasing pathogenicity in the mouse model (*20–24*). The 591Q, 627E, and 701D sequences in the PB2 gene of our H5N1 viruses indicated that they have not adapted to mammals. H274Y in the NA protein and S31N in M2 protein are NA- and M2-inhibitor resistance mutations, respectively (*25*,*26*). The 274H in NA and 31S in M2 genes of these H5N1 reassortants indicate they are sensitive to NA- and M2-inhibitors. Eight identical N-glycosylation sites are found in Re-14 and H5N1 viruses here, but one additional N-

glycosylation site is found in Bs/EC/74-Lg/2021, which could correlate to antigenic escape. Overall, we conclude that these novel viruses pose a low zoonotic/pandemic threat.

Necropsy and histopathological findings of a dead black swan caused by H5N1 virus

The infected black swan showed severe neurologic signs. At necropsy, serious lesions consisting of hemorrhages and necrosis in the lung, liver, and intestine were observed. Histologically, severe hemorrhage, edema, collapsed alveoli and inflammatory cell infiltration were found in the lung (Appendix 2 Figure 1). The intestine showed mucosal necrosis and hemorrhages. Focal necrosis was seen in the liver. Under high magnification, hepatocytes showed necrotic disintegration and inflammatory cells infiltrated around macrophages within the intervals among sinusoids.

The H5N1 isolates prefer avian-type receptors

The viral affinity to particular isoforms of the sialic acid receptors of host cells is one critical factor for host range of AIVs (*16*). The receptor binding assay showed all three H5N1 isolates have a receptor-binding preference to avian-type (α2–3-SA) receptors (Appendix 2 Figure 2). The avian-like receptor binding preference of the H5N1 is consistent with the molecular signatures observed at receptor binding sites, 224N, 226Q, and 228G (Appendix 2 Table 2).

References

1. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data—from vision to reality. Euro Surveill. 2017;22:30494. PubMed https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494

2. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, et al. The influenza virus resource at the National Center for Biotechnology Information. J Virol. 2008;82:596–601. PubMed https://doi.org/10.1128/JVI.02005-07

3. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80. PubMed https://doi.org/10.1093/molbev/mst010

4. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3. PubMed https://doi.org/10.1093/bioinformatics/btu033

5. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 2018;4:vey016. PubMed https://doi.org/10.1093/ve/vey016

6. Shapiro B, Rambaut A, Drummond AJ. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol Biol Evol. 2006;23:7–9. PubMed https://doi.org/10.1093/molbev/msj021

7. Hill V, Baele G. Bayesian estimation of past population dynamics in BEAST 1.10 using the Skygrid coalescent model. Mol Biol Evol. 2019;36:2620–8. PubMed https://doi.org/10.1093/molbev/msz172

8. Elliot MG, Mooers AØ. Inferring ancestral states without assuming neutrality or gradualism using a stable model of continuous character evolution. BMC Evol Biol. 2014;14:226. PubMed https://doi.org/10.1186/s12862-014-0226-8

9. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLOS Comput Biol. 2009;5:e1000520. PubMed https://doi.org/10.1371/journal.pcbi.1000520

10. Bi Y, Li J, Li S, Fu G, Jin T, Zhang C, et al. Dominant subtype switch in avian influenza viruses during 2016–2019 in China. Nat Commun. 2020;11:5909. PubMed https://doi.org/10.1038/s41467-020-19671-3

11. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296–303. PubMed https://doi.org/10.1093/nar/gky427

12. Bi Y, Zhang Z, Liu W, Yin Y, Hong J, Li X, et al. Highly pathogenic avian influenza A(H5N1) virus struck migratory birds in China in 2015. Sci Rep. 2015;5:12986. PubMed https://doi.org/10.1038/srep12986

13. World Organisation for Animal Health. High pathogenicity avian influenza (HPAI)-situation report 24. December 15, 2021 [cited 2022 Mar 1]. https://www.oie.int/en/document/high-pathogenicity-avian-influenza-hpai-situation-report-24

14. Stevens J, Blixt O, Tumpey TM, Taubenberger JK, Paulson JC, Wilson IA. Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. Science. 2006;312:404–10. PubMed https://doi.org/10.1126/science.1124513

15. Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, et al. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature. 2012;486:420–8. PubMed https://doi.org/10.1038/nature10831

16. Herfst S, Schrauwen EJ, Linster M, Chutinimitkul S, de Wit E, Munster VJ, et al. Airborne transmission of influenza A/H5N1 virus between ferrets. Science. 2012;336:1534–41. PubMed https://doi.org/10.1126/science.1213362

17. Matsuoka Y, Swayne DE, Thomas C, Rameix-Welti MA, Naffakh N, Warnes C, et al. Neuraminidase stalk length and additional glycosylation of the hemagglutinin influence the virulence of influenza H5N1 viruses for mice. J Virol. 2009;83:4704–8. PubMed https://doi.org/10.1128/JVI.01987-08

18. Zhou H, Yu Z, Hu Y, Tu J, Zou W, Peng Y, et al. The special neuraminidase stalk-motif responsible for increased virulence and pathogenesis of H5N1 influenza A virus. PLoS One. 2009;4:e6277. PubMed https://doi.org/10.1371/journal.pone.0006277

19. Cauldwell AV, Long JS, Moncorgé O, Barclay WS. Viral determinants of influenza A virus host range. J Gen Virol. 2014;95:1193–210. PubMed https://doi.org/10.1099/vir.0.062836-0

20. Subbarao K, Klimov A, Katz J, Regnery H, Lim W, Hall H, et al. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. Science. 1998;279:393–6. PubMed https://doi.org/10.1126/science.279.5349.393

21. Hatta M, Gao P, Halfmann P, Kawaoka Y. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. Science. 2001;293:1840–2. PubMed https://doi.org/10.1126/science.1062882

22. Yamada S, Hatta M, Staker BL, Watanabe S, Imai M, Shinya K, et al. Biological and structural characterization of a host-adapting amino acid in influenza virus. PLoS Pathog. 2010;6:e1001034. PubMed https://doi.org/10.1371/journal.ppat.1001034

23. Zhou B, Pearce MB, Li Y, Wang J, Mason RJ, Tumpey TM, et al. Asparagine substitution at PB2 residue 701 enhances the replication, pathogenicity, and transmission of the 2009 pandemic H1N1 influenza A virus. PLoS One. 2013;8:e67616. PubMed https://doi.org/10.1371/journal.pone.0067616

24. Mok CK, Yen HL, Yu MY, Yuen KM, Sia SF, Chan MC, et al. Amino acid residues 253 and 591 of the PB2 protein of avian influenza virus A H9N2 contribute to mammalian pathogenesis. J Virol. 2011;85:9641–5. PubMed https://doi.org/10.1128/JVI.00702-11

25. Hay AJ, Wolstenholme AJ, Skehel JJ, Smith MH. The molecular basis of the specific anti-influenza action of amantadine. EMBO J. 1985;4:3021–4. PubMed https://doi.org/10.1002/j.1460-2075.1985.tb04038.x

26. Pinto LH, Holsinger LJ, Lamb RA. Influenza virus M2 protein has ion channel activity. Cell. 1992;69:517–28. PubMed https://doi.org/10.1016/0092-8674(92)90452-I

Appendix 2 Table 1. Percent identity of nucleotide sequences of H5N1 viruses identified in this study.

Viruses	HA	NA	PB2	PB1	PA	NS	MP	NP
A/whooper swan/Northern China/11.03 IMEEDSAK1-O/2021 A/black swan/Eastern China/11.15 ZJHZ74-Lg/2021	97.9	97.2	94.2	92.9	97.6	95.6	99.0	97.6

The A/whooper swan/Northern China/11.03 IMEEDSAK1-O/2021 and A/whooper swan/Northern China/11.03 IMEEDSAK2-O/2021 were identified from the same whooper swan. They have identical HA, NA, MP, NS, and PB2 genes but one different nucleotide site respectively in PA, NP, and PB1 genes.

Appendix 2 Table 2. Molecular characteristics on key functional sites for HA, PB2, NA, M2 proteins of H5N1 and H5N8 viruses.

	HA (H3 numbering)								PB2		NA (N2 numbering)				M2
Viruses	Cleavage site	110	158	160	224	226	228	318	591	627	701	247	252	274	Stalk deletion	31
A/black swan/Eastern China/11.15 ZJHZ74-Lg/2021(H5N1)	PLRERRRKR/GL	H	N	A	N	Q	G	T	Q	E	D	N	Y	H	No	S
A/whooper swan/Northern China/11.03 IMEEDSAK1-O/2021(H5N1)	PLREKRRKR/GL	H	D	A	N	Q	G	T	Q	E	D	N	Y	H	No	S
A/whooper swan/Northern China/11.03 IMEEDSAK2-O/2021(H5N1)	PLREKRRKR/GL	H	D	A	N	Q	G	T	Q	E	D	N	Y	H	No	S
A/whooper swan/Henan/CAS001-K/2020(H5N8)	PLREKRRKR/GL	H	N	A	N	Q	G	T	Q	E	D	N	Y	H	No	S
A/whooper swan/Henan/CAS001-G/2020(H5N8)	PLREKRRKR/GL	H	N	A	N	Q	G	T	Q	E	D	N	Y	H	No	S
A/whooper swan/Henan/CAS002-F17/2020(H5N8)	PLREKRRKR/GL	H	N	A	N	Q	G	T	Q	E	D	N	Y	H	No	S
A/whooper swan/Henan/CAS002-F18/2020(H5N8)	PLREKRRKR/GL	H	N	A	N	Q	G	T	Q	E	D	N	Y	H	No	S

Appendix 2 Table 3. Amino acid substitutions on the HA antigenic sites of H5N1/H5N8 HPAIVs and H5 Re-11, Re-13 and Re-14 vaccine seed viruses used in China given H5 antigenic sites.

	Position of antigenic sites in HA genes (H3 numbering)																						
Viruses	129	130	131	132	133	140	141	142	143	144	145	155	156	157	158	159	160	161	162	163	164	165	166
Re-11, clade2344h	N	H	**T**	**S**	S	P	Y	Q	G	**V**	**A**	**T**	K	K	**N**	**D**	**A**	Y	P	T	I	K	**M**
Re-13, clade2344h	N	H	**T**	**T**	S	P	Y	Q	G	**V**	**A**	**T**	K	K	**N**	**E**	**T**	Y	P	T	I	K	**K**
Re-14, clade2344b	N	H	**E**	**T**	S	P	Y	Q	G	**A**	**P**	**I**	K	K	**N**	**D**	**A**	Y	P	T	I	K	**I**
Bs/EC/74-Lg/2021(H5N1)	N	H	**E**	**T**	S	P	Y	Q	G	**A**	**P**	**I**	K	K	**N**	**D**	**A**	Y	P	T	I	K	**I**
Ws/NC/AK1-O/2021(H5N1)	N	H	**E**	**T**	S	P	Y	Q	G	**A**	**P**	**I**	K	K	**D**	**D**	**A**	Y	P	T	I	K	**I**
Ws/NC/AK2-O/2021(H5N1)	N	H	**E**	**T**	S	P	Y	Q	G	**A**	**P**	**I**	K	K	**D**	**D**	**A**	Y	P	T	I	K	**I**
Ws/HN/1-K/2020(H5N8)	N	H	**E**	**T**	S	P	Y	Q	G	**A**	**P**	**I**	K	K	**N**	**D**	**A**	Y	P	T	I	K	**I**
Ws/HN/1-G/2020(H5N8)	N	H	**E**	**T**	S	P	Y	Q	G	**A**	**P**	**I**	K	K	**N**	**D**	**A**	Y	P	T	I	K	**I**
Ws/HN/2-F17/2020(H5N8)	N	H	**E**	**T**	S	P	Y	Q	G	**A**	**P**	**I**	K	K	**N**	**D**	**A**	Y	P	T	I	K	**I**
Ws/HN/2-F18/2020(H5N8)	N	H	**E**	**T**	S	P	Y	Q	G	**A**	**P**	**I**	K	K	**N**	**D**	**A**	Y	P	T	I	K	**I**

The substitutions on H5 antigenic sites are shown in bold.

Appendix 2 Table 4. Amino acid substitutions on the HA antigenic sites of H5N1/H5N8 HPAIVs and H5 Re-11, Re-13 and Re-14 vaccine seed viruses used in China according to five antigenic sites A–E for H3 influenza virus.

Viruses	Position of antigenic sites in HA genes (H3 numbering)																						
	Site A																					Site B	
	122	123	124	125	126	127	129	132	133	134	135	136	137	138	140	141	142	143	144	145	146	155	156
Re-11, clade2344h	I	P	K	**R**	S	W	N	**S**	S	G	V	S	A	A	P	Y	Q	G	**V**	**A**	S	**T**	K
Re-13, clade2344h	I	P	K	**E**	S	W	N	**T**	S	G	V	S	A	A	P	Y	Q	G	**V**	**A**	S	**T**	K
Re-14, clade2344b	I	P	K	**S**	S	W	N	**T**	S	G	V	S	A	A	P	Y	Q	G	**A**	**P**	S	**I**	K
Bs/EC/74-Lg/2021(H5N1)	I	P	K	**N**	S	W	N	**T**	S	G	V	S	A	A	P	Y	Q	G	**A**	**P**	S	**I**	K
Ws/NC/AK1-O/2021(H5N1)	I	P	K	**S**	S	W	N	**T**	S	G	V	S	A	A	P	Y	Q	G	**A**	**P**	S	**I**	K
Ws/NC/AK2-O/2021(H5N1)	I	P	K	**S**	S	W	N	**T**	S	G	V	S	A	A	P	Y	Q	G	**A**	**P**	S	**I**	K
Ws/HN/1-K/2020(H5N8)	I	P	K	**S**	S	W	N	**T**	S	G	V	S	A	A	P	Y	Q	G	**A**	**P**	S	**I**	K
Ws/HN/1-G/2020(H5N8)	I	P	K	**S**	S	W	N	**T**	S	G	V	S	A	A	P	Y	Q	G	**A**	**P**	S	**I**	K
Ws/HN/2-F17/2020(H5N8)	I	P	K	**S**	S	W	N	**T**	S	G	V	S	A	A	P	Y	Q	G	**A**	**P**	S	**I**	K
Ws/HN/2-F18/2020(H5N8)	I	P	K	**S**	S	W	N	**T**	S	G	V	S	A	A	P	Y	Q	G	**A**	**P**	S	**I**	K

	Site B												Site C						Site D					
Viruses	157	158	159	160	186	188	189	190	193	194	196	197	50	53	54	275	278	174	201	202	203	204	205
Re-11, clade2344h	K	**N**	**D**	**A**	N	**A**	**E**	E	**N**	L	K	N	K	D	–	G	N	E	Y	**V**	S	V	G
Re-13, clade2344h	K	**N**	**E**	**T**	N	**V**	**E**	E	**D**	L	K	N	K	D	–	G	N	E	Y	**V**	S	V	G
Re-14, clade2344b	K	**N**	**D**	**A**	N	**A**	**E**	E	**N**	L	K	N	K	D	–	G	N	E	Y	**I**	S	V	G
Bs/EC/74-Lg/2021(H5N1)	K	**N**	**D**	**A**	N	**A**	**K**	E	**D**	L	K	N	K	D	–	G	N	E	Y	**I**	S	V	G
Ws/NC/AK1-O/2021(H5N1)	K	**D**	**D**	**A**	N	**A**	**K**	E	**N**	L	K	N	K	D	–	G	N	E	Y	**I**	S	V	G
Ws/NC/AK2-O/2021(H5N1)	K	**D**	**D**	**A**	N	**A**	**K**	E	**N**	L	K	N	K	D	–	G	N	E	Y	**I**	S	V	G
Ws/HN/1-K/2020(H5N8)	K	**N**	**D**	**A**	N	**A**	**E**	E	**N**	L	K	N	K	D	–	G	N	E	Y	**I**	S	V	G
Ws/HN/1-G/2020(H5N8)	K	**N**	**D**	**A**	N	**A**	**E**	E	**N**	L	K	N	K	D	–	G	N	E	Y	**I**	S	V	G
Ws/HN/2-F17/2020(H5N8)	K	**N**	**D**	**A**	N	**A**	**E**	E	**N**	L	K	N	K	D	–	G	N	E	Y	**I**	S	V	G
Ws/HN/2-F18/2020(H5N8)	K	**N**	**D**	**A**	N	**A**	**E**	E	**N**	L	K	N	K	D	–	G	N	E	Y	**I**	S	V	G

Viruses	Site D	Site E

Viruses	Position of antigenic sites in HA genes (H3 numbering)														
	206	207	208	217	218	219	220	62	63	78	79	80	81	82	83
Re-11, clade2344h	T	S	T	I	A	T	R	V	**D**	E	F	I	**R**	P	E
Re-13, clade2344h	T	S	T	I	A	T	R	V	**N**	E	F	I	**S**	P	E
Re-14, clade2344b	T	S	T	I	A	T	R	V	**D**	E	F	I	**R**	P	E
Bs/EC/74-Lg/2021(H5N1)	T	S	T	I	A	T	R	V	**D**	E	F	I	**R**	P	E
Ws/NC/AK1-O/2021(H5N1)	T	S	T	I	A	T	R	V	**D**	E	F	I	**R**	P	E
Ws/NC/AK2-O/2021(H5N1)	T	S	T	I	A	T	R	V	**D**	E	F	I	**R**	P	E
Ws/HN/1-K/2020(H5N8)	T	S	T	I	A	T	R	V	**D**	E	F	I	**R**	P	E
Ws/HN/1-G/2020(H5N8)	T	S	T	I	A	T	R	V	**D**	E	F	I	**R**	P	E
Ws/HN/2-F17/2020(H5N8)	T	S	T	I	A	T	R	V	**D**	E	F	I	**R**	P	E
Ws/HN/2-F18/2020(H5N8)	T	S	T	I	A	T	R	V	**D**	E	F	I	**R**	P	E

The substitutions on H3 antigenic sites are shown in bold.

Appendix 2 Table 5. Origin of the two novel Bs/2021-like and Ws/2021-like reassortants.*

Reassortants	Segment	Ancestral type†	Time of origin		HA	NA	Latitude	Longitude	Ancestral host§	Host type probability¶			
			Median	95% HPD‡						Dom Gal	Wild Ans	Wild other	Dom Ans
Bs/2021-like	PB2	MRCGA	Feb-21	(Oct-20, Jul-21)	H5	N1	35.039	115.893	Wild Ans	0.03	0.97	0	0
		MRCA	Apr-21	(Dec-20, Jul-21)	H5	N1	34.973	111.415	Wild Ans	0.03	0.97	0	0
	PB1	MRCGA	Jun-21	(Apr-21, Aug-21)	H5	N1	36.039	112.431	Wild Ans	0.07	0.93	0	0
		MRCA	Jul-21	(May-21, Sep-21)	H5	N1	35.267	116.604	Wild Ans	0.08	0.92	0	0
	PA	MRCGA	Jun-19	(May-18, Feb-20)	H5	N1	42.291	47.677	Dom Gal	0.66	0.27	0.03	0.03
		MRCA	May-21	(Jan-21, Aug-21)	H5	N1	34.823	115.92	Wild Ans	0.48	0.51	0	0.01
	HA	MRCGA	Jun-21	(Nov-21, Mar-22)	H5	N1	36.54	116.363	Wild Ans	0.39	0.58	0.02	0
		MRCA	Aug-21	(Dec-21, Apr-22)	H5	N1	35.29	119.93	Wild Ans	0.42	0.57	0.01	0
	NP	MRCGA	Jun-21	(Mar-21, Aug-21)	H5	N1	35.982	115.333	Wild Ans	0.28	0.71	0.01	0
		MRCA	Jul-21	(Apr-21, Sep-21)	H5	N1	35.226	118.38	Wild Ans	0.28	0.72	0	0
	NA	MRCGA	Apr-20	(May-19, Mar-21)	H5	N1	38.843	53.758	Wild Ans	0.24	0.47	0.28	0.01
		MRCA	Jul-21	(Mar-21, Aug-21)	H5	N1	35.313	116.577	Wild Ans	0.09	0.87	0.04	0
	MP	MRCGA	May-21	(Apr-21, Jul-21)	H5	N8	37.376	102.744	Dom Gal	0.63	0.35	0	0
		MRCA	Sep-21	(Jul-21, Oct-21)	H5	N1	35.137	112.058	Dom Gal	0.85	0.15	0	0
	NS	MRCGA	Feb-21	(Sep-20, Jun-21)	H5	N1	39.939	55.067	Wild Ans	0.2	0.8	0	0
		MRCA	Jul-21	(Apr-21, Oct-21)	H5	N1	35.261	107.565	Wild Ans	0.27	0.72	0	0
Ws/2021-like	PB2	MRCGA	Jul-17	(Oct-15, Feb-19)	H3	N6	39.802	126.235	Wild Ans	0	0.59	0	0.41
		MRCA	Oct-21	(Aug-21, Nov-21)	H5	N1	39.879	101.088	Wild Ans	0	1	0	0
	PB1	MRCGA	Apr-21	(Feb-21, Jun-21)	H5	N1	46.887	37.112	Wild Ans	0.23	0.77	0	0
		MRCA	Oct-21	(Sep-21, Nov-21)	H5	N1	40.347	96.064	Wild Ans	0	1	0	0
	PA	MRCGA	Mar-21	(Dec-20, May-21)	H5	N1	46.038	31.212	Wild Ans	0.07	0.92	0	0
		MRCA	Oct-21	(Aug-21, Nov-21)	H5	N1	40.294	95.796	Wild Ans	0	1	0	0
	HA	MRCGA	Mar-21	(Jul-21, Nov-21)	H5	N1	39.118	39.471	Dom Gal	0.66	0.33	0.01	0
		MRCA	Oct-21	(Jun-22, May-22)	H5	N1	39.904	94.895	Wild Ans	0	1	0	0
	NP	MRCGA	Jun-21	(Apr-21, Aug-21)	H5	N1	45.988	43.644	Wild Ans	0.03	0.97	0	0

Reassortants	Segment	Ancestral type†	Time of origin		HA	NA	Latitude	Longitude	Ancestral host§	Host type probability¶			
			Median	95% HPD‡						Dom Gal	Wild Ans	Wild other	Dom Ans
		MRCA	Oct-21	(Sep-21, Oct-21)	H5	N1	40.478	109.189	Wild Ans	0	1	0	0
	NA	MRCGA	Feb-21	(Sep-20, Apr-21)	H5	N1	44.859	35.22	Wild Ans	0.07	0.82	0.11	0
		MRCA	Oct-21	(Sep-21, Oct-21)	H5	N1	39.935	99.384	Wild Ans	0	1	0	0
	MP	MRCGA	Aug-21	(Aug-21, Sep-21)	H5	N1	45.415	39.516	Wild Ans	0.47	0.5	0	0.03
		MRCA	Oct-21	(Oct-21, Nov-21)	H5	N1	40.969	93.101	Wild Ans	0.02	0.98	0	0
	NS	MRCGA	May-21	(Mar-21, Jul-21)	H5	N1	49.937	40.477	Dom Gal	0.55	0.45	0	0
		MRCA	Oct-21	(Sep-21, Nov-21)	H5	N1	40.385	97.22	Wild Ans	0	1	0	0

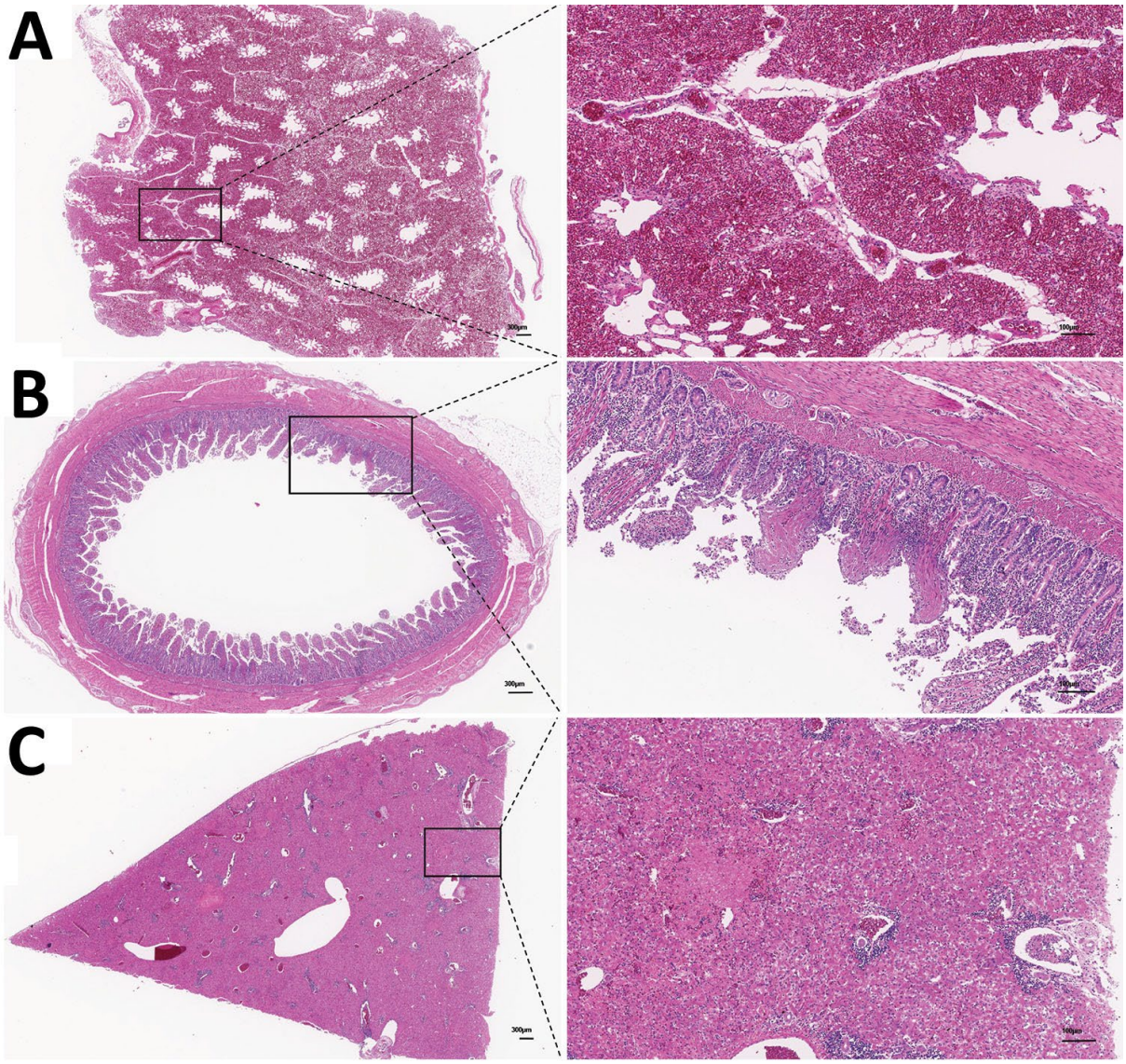
*The estimated origin time, locations, host types and subtypes (HA and NA) of the grand most recent common ancestors (GMRCAs) and the most recent common ancestors (MRCAs) of two reassortants, Bs/2021-like from eastern China and Ws/2021-like from northern China.

†Two types of ancestors: the most recent common ancestor (MRCA) and the most recent common grand ancestor (MRCGA), which is the direct ancestor of the MRCA.

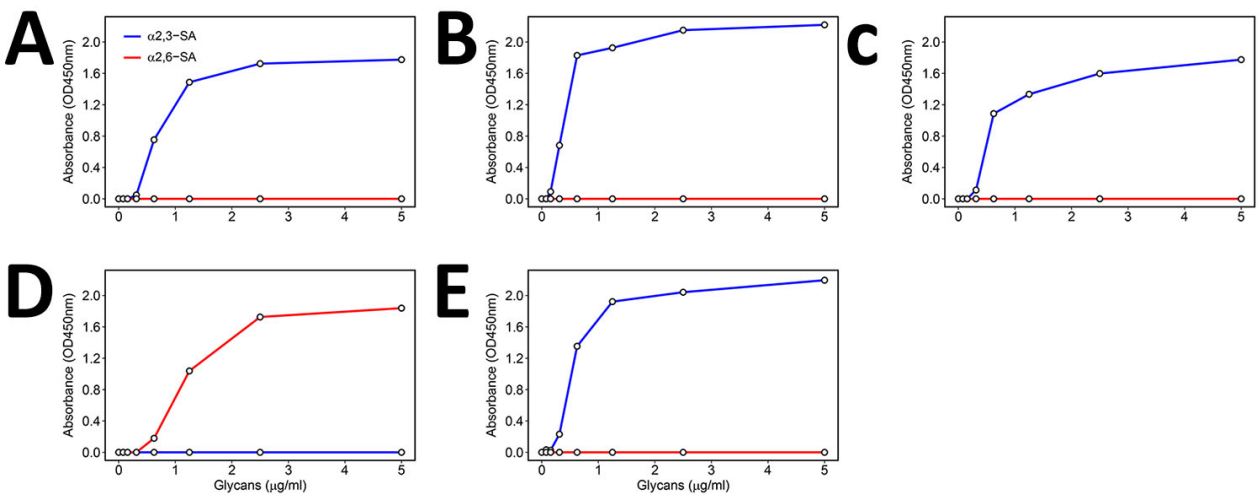
‡HPD is short for the highest posterior density interval.

§Four host types: Domestic Galliformes (Dom Gal), Wild Anseriformes (Wild Ans), Domestic Anseriformes (Dom Ans) and other wild bird types (Wild other).

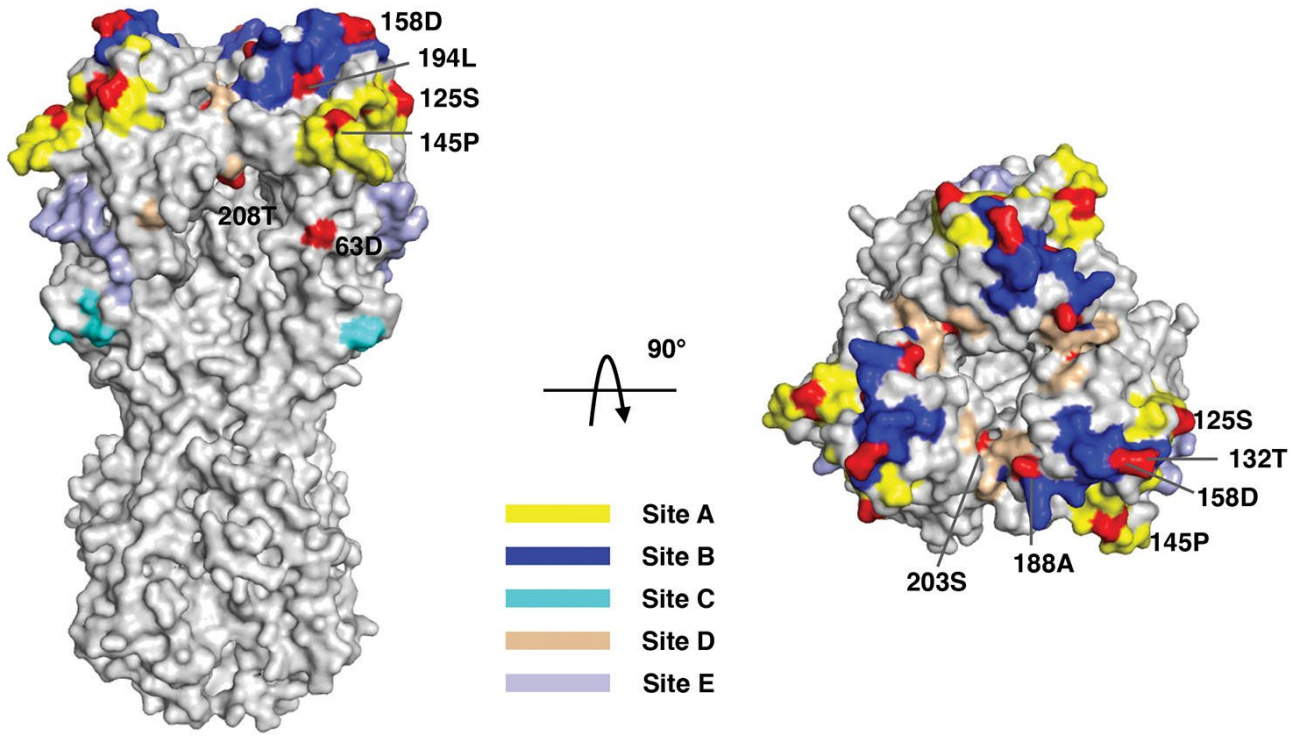
¶The posterior probability of four ancestor host types at the corresponding ancestral nodes.



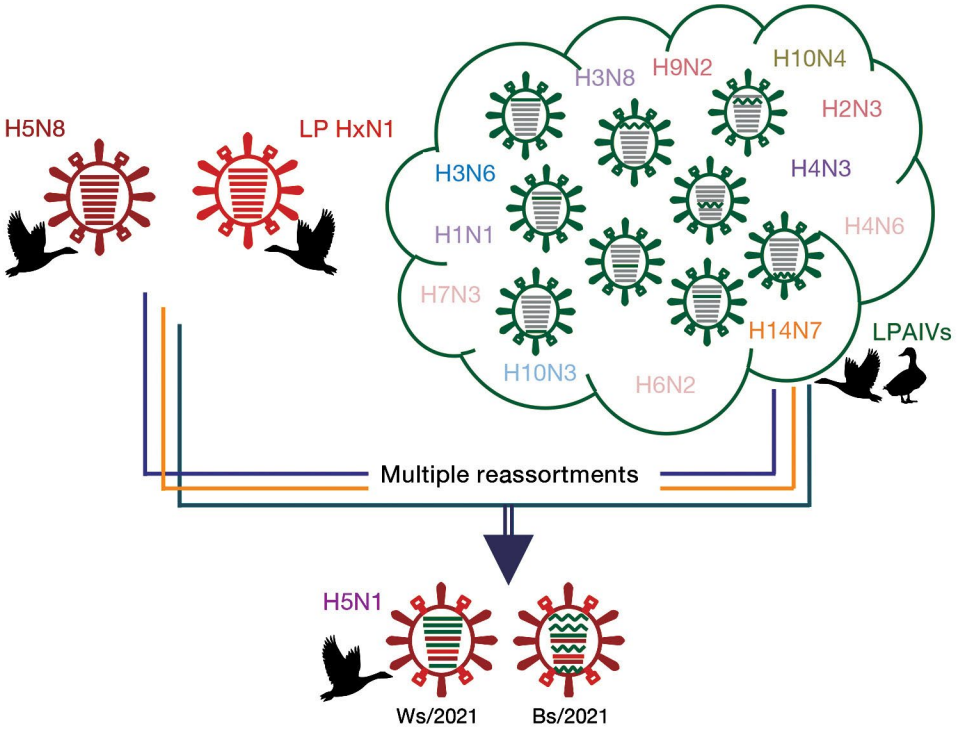
Appendix 2 Figure 1. Histopathologic changes in the tissues of a dead black swan infected by H5N1 virus. Histopathologic changes of (a) hemorrhages in the lung, (b) necrosis in the intestine, and (c) necrosis in the liver. The left panel of each sub-figure shows the panoramic image of corresponding tissues. The right panel shows an enlarged view with obviously histopathologic changes. Scale bars represent 300μm (left panel) and 100μm (right panel), respectively.



Appendix 2 Figure 2. Receptor binding properties for novel H5N1 viruses identified in this study. Receptor binding properties for (a) Ws/NC/AK1-O/2021(H5N1), (b) Ws/NC/AK2-O/2021(H5N1), (c) Bs/EC/74-Lg/2021(H5N1), (d) A/California/04/2009(H1N1), and (e) A/Vietnam/1194/2004(H5N1) virus to α-2,3-linked (blue) and α-2,6-linked (red) sialic acids (SAs), respectively. The A/California/04/2009(H1N1) and A/Vietnam/1194/2004(H5N1) viruses were used as controls with known receptor binding preference to human-type (α-2,6-linked SA) and avian-type receptor (α-2,3-linked SA), respectively. The values represented by points are the mean of three readings on the absorbance value.



Appendix 2 Figure 3. Trimer structure of the ectodomain of HA molecule of the Ws/NC/AK1-O/2021(H5N1) estimated by homology modeling. Five major antigenic sites A–E given antigenicity study on H3 influenza were colored by different colors. The key amino acids on antigenic sites of Ws/NC/AK1-O/2021(H5N1) are colored in red. All positions are shown in H3 numbering. PDB ID of the template used in homology modeling is 7DEB.



Appendix 2 Figure 4. Evolutionary pathways of two H5N1 reassortants identified in this study. The horizontal bars in the virus particle represent the eight gene segments (from top to bottom: PB2, PB1, PA, HA, NP, NA, MP, and NS). Gene segments in descendent viruses are colored according to potential parent viruses. The Ws/2021-like and Bs/2021-like H5N1 reassortants identified in this study resulted from multiple reassortments between the H5N8 HPAIVs from wild waterfowl and the LPAIVs from wild waterfowl or ducks. The reassortants obtained their HA and MP genes from the H5N8 HPAIVs, NA genes from the LPAIV HxN1 virus, and PB2, PB1, PA, NS, and NP genes from multiple LPAIV pools.

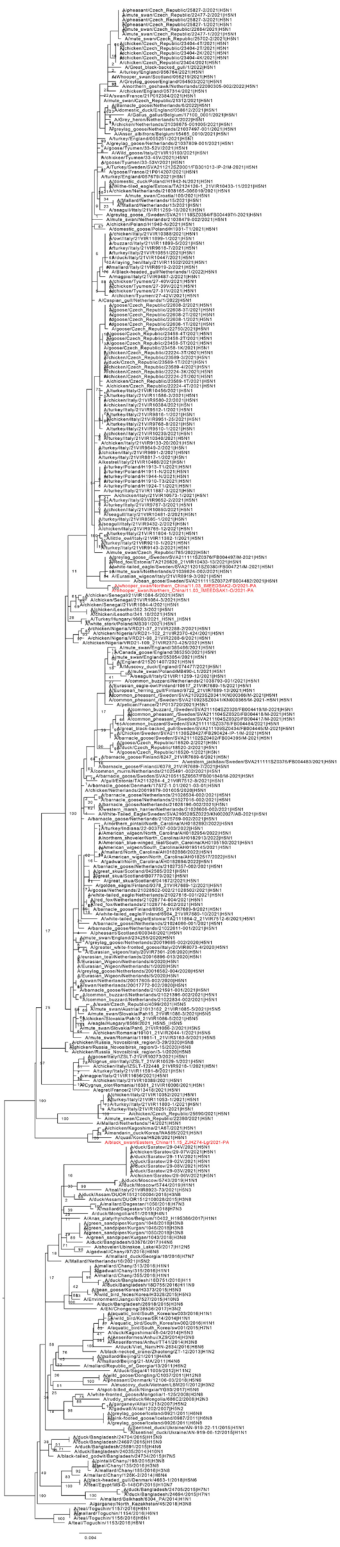
Appendix 2 Figure 5. Phylogenetic tree of HA genes of the novel H5N1 viruses and their reference viruses with the most genetic identities. The viruses identified in this study are colored in red.

Appendix 2 Figure 6. Phylogenetic tree of NA genes of the novel H5N1 viruses and their reference viruses with the most genetic identities. The viruses identified in this study are colored in red.

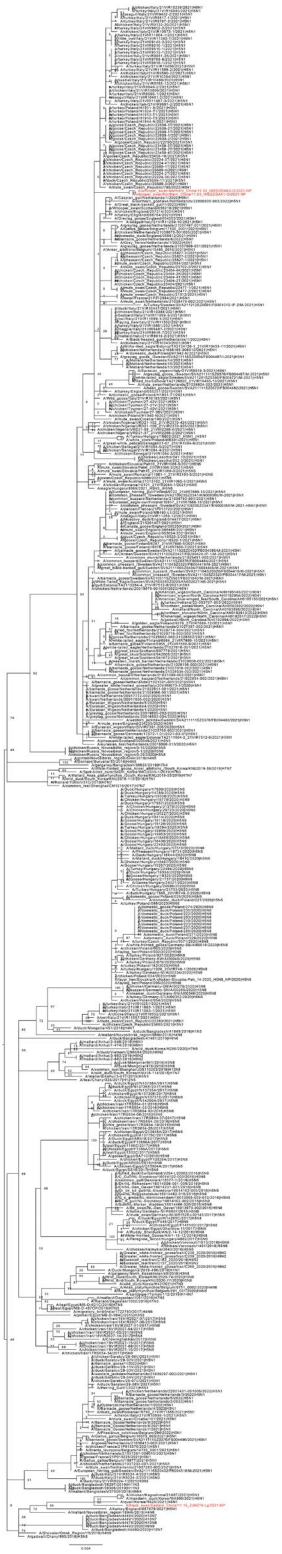
Appendix 2 Figure 7. Phylogenetic tree of MP genes of the novel H5N1 viruses and their reference viruses with the most genetic identities. The viruses identified in this study are colored in red.

Appendix 2 Figure 8. Phylogenetic tree of PB2 genes of the novel H5N1 viruses and their reference viruses with the most genetic identities. The viruses identified in this study are colored in red.

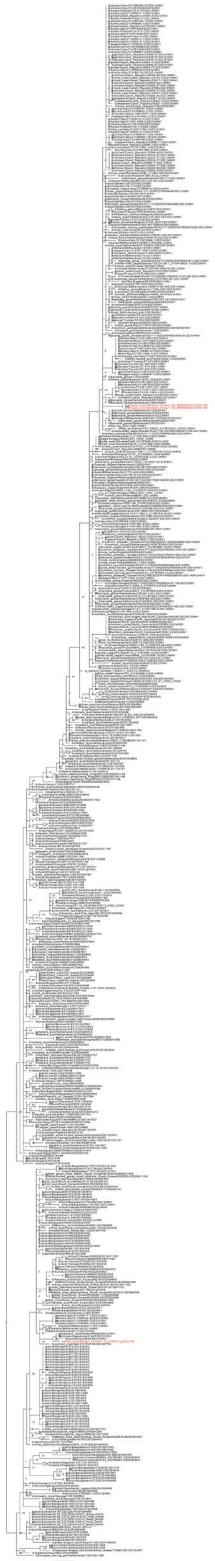
Appendix 2 Figure 9. Phylogenetic tree of PB1 genes of the novel H5N1 viruses and their reference viruses with the most genetic identities. The viruses identified in this study are colored in red.



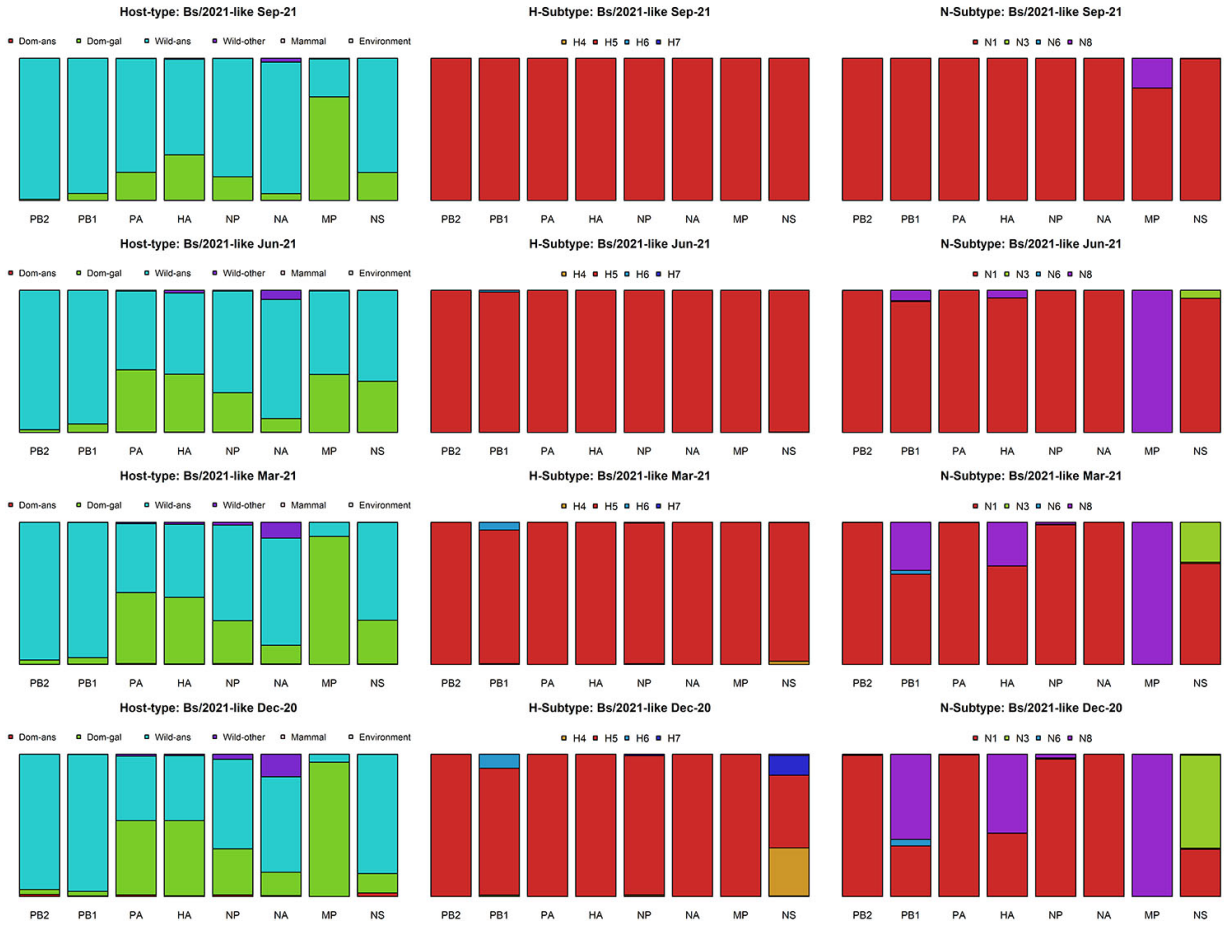
Appendix 2 Figure 10. Phylogenetic tree of PA genes of the novel H5N1 viruses and their reference viruses with the most genetic identities. The viruses identified in this study are colored in red.



Appendix 2 Figure 11. Phylogenetic tree of NP genes of the novel H5N1 viruses and their reference viruses with the most genetic identities. The viruses identified in this study are colored in red.



Appendix 2 Figure 12. Phylogenetic tree of NS genes of the novel H5N1 viruses and their reference viruses with the most genetic identities. The viruses identified in this study are colored in red.



Appendix 2 Figure 13. Inferred ancestral hosts, HA and NA types of Bs/2021-like H5N1 reassortant. The ancestral host and HA/NA type estimations were adjusted by interpolating by every 3 months between December 2020 and November 2021. For each segment, stack bars represent the posterior probabilities of each type of the discrete traits (see legend).

Appendix 2 Figure 14. Inferred ancestral hosts, HA and NA types of Ws/2021-like H5N1 reassortant. The ancestral host and HA/NA type estimations were adjusted by interpolating by every 3 months between December 2020 and November 2021. For each segment, stack bars represent the posterior probabilities of each type of the discrete traits (see legend).