



## Emergence, prevalence, and evolution of H5N8 avian influenza viruses in central China, 2020

Xiang Li <sup>a</sup>, Xinru Lv<sup>a</sup>, Yi Li<sup>a</sup>, Linhong Xie<sup>b</sup>, Peng Peng<sup>b</sup>, Qing An<sup>a</sup>, Tian Fu<sup>a</sup>, Siyuan Qin<sup>b</sup>, Yuan Cui<sup>c</sup>, Chengbo Zhang<sup>d</sup>, Rongxiu Qin<sup>a</sup>, Fengyi Qu<sup>a</sup>, Zhenliang Zhao<sup>a</sup>, Meixi Wang<sup>a</sup>, Qiuzi Xu<sup>a</sup>, Yong Li<sup>e</sup>, Guoxiang Yang<sup>e</sup>, Guang Chen<sup>e</sup>, Jun Zhang<sup>e</sup>, Hesong Zheng<sup>e</sup>, Enda Ma<sup>f</sup>, Ruifang Zhou<sup>f</sup>, Xiangwei Zeng<sup>a</sup>, Yulong Wang<sup>a</sup>, Zhijun Hou<sup>a</sup>, Yajun Wang<sup>g</sup>, Dong Chu<sup>h</sup>, Yanbing Li<sup>g</sup> and Hongliang Chai <sup>a†</sup>

<sup>a</sup>College of Wildlife and Protected Area, Northeast Forestry University, Harbin, People's Republic of China; <sup>b</sup>National Forestry and Grassland Administration, General Station for Surveillance of Wildlife Disease & Wildlife Borne Diseases, Shenyang, People's Republic of China; <sup>c</sup>Sanmenxia Administration of the National Nature Reserve of the Yellow River Wetland, Sanmenxia, People's Republic of China; <sup>d</sup>Ordos Forestry and Grassland Administration, Ordos, People's Republic of China; <sup>e</sup>Research and Development Center, Hubei Wildlife Rescue, Wuhan, People's Republic of China; <sup>f</sup>Bayannur Forestry and Grassland Administration, Bayannur, People's Republic of China; <sup>g</sup>State Key Laboratory of Veterinary Biotechnology, Chinese Academy of Agricultural Sciences, Harbin Veterinary Research Institute, Harbin, People's Republic of China

### ABSTRACT

Highly pathogenic influenza A(H5N8) viruses have caused several worldwide outbreaks in birds and are able cross the species barrier to infect humans, posing a substantial threat to public health. After the first detection of H5N8 viruses in deceased swans in Inner Mongolia, we performed early warning and active monitoring along swan migration routes in central China. We isolated and sequenced 42 avian influenza viruses, including 40 H5N8 viruses, 1 H5N2 virus, and 1 H9N2 virus, in central China. Our H5N8 viruses isolated in swan stopover sites and wintering grounds showed high nucleotide homologies in the whole genome, revealing a common evolutionary source. Phylogenetic analysis revealed that the H5 viruses of clade 2.3.4.4b prevalent in 2020 have further diverged into two sub-clades: b1 and b2. The phylogeographic analysis also showed that the viruses of sub-clade b2 most likely originated from poultry in Russia. Notably, whooper swans were found to be responsible for the introduction of sub-clade b2 viruses in central China; whooper and tundra swans play a role in viral spread in the Yellow River Basin and the Yangtze River Basin, respectively. Our findings highlight swans as an indicator species for transborder spreading and monitoring of the H5N8 virus.



**ARTICLE HISTORY** Received 6 October 2021; Revised 15 November 2021; Accepted 23 November 2021

**KEYWORDS** Avian influenza viruses; China; H5N8; swan; 2.3.4.4b

### Introduction

Influenza A viruses (IAVs) infect a broad spectrum of hosts, including wild birds, poultry, swine, horses, seals, cats, and humans, occasionally causing major outbreaks or severe disease. Wild birds, especially those of the order *Anseriformes* and *Charadriiformes*, are considered to be the natural reservoirs of IAVs and play a fundamental role in the spread and maintenance of the virus [1]. Moreover, migratory birds contribute to the wide geographic spread and distribution of the circulating highly pathogenic avian influenza (HPAI) H5 virus [2]. In 2014, HPAI H5N8 virus of clade 2.3.4.4 was first identified in South Korea and subsequently spread to Asia, Europe, and North America by migratory waterfowl [3–5]. Another HPAI H5N8 virus of clade 2.3.4.4 began its intercontinental spread

in the autumn in 2016 [6], and it has since been shown to have a sustained prevalence in Europe, Africa, and the Middle East (OIE, <https://www.oie.int/en/animal-health-in-the-world/>). In early 2020, after small H5N8 outbreaks were reported in poultry in Europe, HPAI H5N8 viruses were continuously detected in Iraq, Russia, and Kazakhstan beginning in May [7]. Moreover, seven poultry farm workers infected with a clade 2.3.4.4b H5N8 virus in Russia were first confirmed in December 2020 [8]. Thus, the global spread of the HPAI H5N8 virus has become a major concern to poultry farming, wildlife conservation and security, as well as global public health [9]. The most challenging tasks associated with HPAI disease ecology is to predict and prevent viral spread, transmission pathways, and host-specific contributions to interregional spread.

**CONTACT** Hongliang Chai  [Hongliang\\_Chai@hotmail.com](mailto:Hongliang_Chai@hotmail.com)  College of Wildlife and Protected Area, Northeast Forestry University, Harbin 150040, People's Republic of China

<sup>†</sup>These authors contributed equally to this work.

 Supplemental data for this article can be accessed <https://doi.org/10.1080/22221751.2021.2011622>

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Shanghai Shangyixun Cultural Communication Co., Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Migratory birds moving within several flyways in Eurasia exhibit an overlap in the locations of circum-polar arctic and subarctic breeding regions [10]. Each year, these birds typically perform seasonal movements between breeding and wintering grounds, viral interactions could also occur among different populations and species, during migration or within habitats. Long-distance migration also exposes the domestic or resident bird populations to IAVs at multiple stopover sites. Therefore, understanding the long-distance movement of wild migratory birds among breeding, wintering grounds, and stopover sites is crucial for explaining the interregional spread of HPAI viruses. Among the bird species involved in the HPAI H5 epidemic, swans appear to be the frequently affected wild species [11, 12]. Frequent HPAI H5 virus detection in swans might be related to their unique biological characteristics: (a) white feathers and large size make them easy to find; (b) concentrated habitat distribution; (c) highly susceptible to infection with the HPAI H5 virus; and (d) efficient spreaders of the virus based on duration, route, and concentration of viral excretion [11, 13].

Whooper swans (*Cygnus cygnus*) breed in northern Eurasia and winter in Europe and eastern Asia (i.e. China, South Korea, and Japan). In the Central Asian flyway, they breed in central-northern and western Mongolia, winter in the Sanmenxia Reservoir area in China, and migrate along the Yellow River and its tributaries of Inner Mongolia, Gansu, Ningxia, Shaanxi, and Shanxi [14]. For the tundra swan (*Cygnus columbianus*), the largest flyway population is the Eastern population, which breeds in the Russian arctic and winters in the middle and lower reaches of the Yangtze River in China [15].

It has become increasingly apparent that as HPAI H5 virus-susceptible birds, swans might play a key role in the geographic spread of virus [14, 16]. Here, we report the monitoring and early warning of the clade 2.3.4.4b H5 virus along the swan migration routes in central China during Autumn 2020.

## Materials and methods

### Eggs

Ten-day-old specific pathogen-free chicken embryos were obtained from the National Poultry Laboratory Animal Resource Centre, Harbin Veterinary Research Institute, Chinese Academy of Agriculture Sciences, Harbin 150069, China.

### Sample collection and virus isolation

Sampling was conducted between October and November 2020 along swan flyways in central China. A total of 1,075 samples were collected, of which 872

samples were collected from the Yellow River Valley (Inner Mongolia, n = 189; Shaanxi, n = 108; Shanxi, n = 399; and Henan, n = 176), and 203 were collected from the Yangzi River Valley (Hubei). Swabs or tissue samples were collected from diseased or dead birds found in the field. Fresh and wet faecal samples were collected from the bird's habitat in the morning. All samples were placed into viral transport media, stored in a handheld portable 4°C refrigerator, transported to the laboratory within 24 h, and stored at -80°C immediately for future use. Avian influenza viruses were isolated in 10-day-old specific pathogen-free chicken embryos using viral isolation procedures in accordance with the method described in the WHO manual [17]. The host species of the droppings was determined using the methods as previously described [18, 19].

### Subtyping and sequencing

HA subtypes were initially identified using a hemagglutination inhibition test. Viral RNA was extracted from haemagglutinin-positive incubated allantoic fluid samples using a QIAamp Viral RNA Mini Kit (Qiagen, Germany), reverse transcribed using the influenza A-specific primer, Un12, and subjected to RT-PCR using the method described in the WHO manual [17] to further confirm AIV-positivity. The PCR products of eight fragments of the isolates were sequenced using a set of specific sequencing primers listed in a previous dissertation [20]. The sequencing data were compiled using the SeqMan programme (DNASTAR, Madison, WI, United States).

### Genetic analysis

A Basic Local Alignment Search Tool (BLAST) search was performed against the sequencing data in the Global Initiative on Sharing All Influenza Data databases (GISAID; <https://www.gisaid.org>) to identify and download the closest relatives of our isolates. The top 100 BLAST hits were derived for each gene dataset. All reference sequences were downloaded in April 2021. The sequences were aligned using MAFFT [21] implemented in PhyloSuite 1.21 [22]. The alignment lengths for each dataset were: PB2: 2,277 nucleotides (nt); PB1: 2,271 nt; PA: 2,148 nt; H5: 1,701 nt; H9: 1,680 nt; NP: 1,494 nt; N8: 1,410 nt; N2: 1,407 nt; M: 979 nt; and NS: 835 nt. Repetitive and short sequences were removed for each dataset. The H5 phylogenetic analyses were conducted with an integrated dataset that was comprised of our isolates, sequence data downloaded in the BLAST search, and the 2017–2020 global H5N8 viruses. Additional sequences of the 2020 H5N8 viruses were also downloaded from GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>). A phylogenetic analysis of H9 and N2 was conducted

using our isolate and the sequences derived from the BLAST search. Phylogenetic analyses of the remaining datasets were reconstructed with the inclusion of our isolates, the sequence data downloaded in the BLAST search, and the 2020 global H5N8 viruses. Maximum likelihood (ML) phylogenies were inferred using IQ-TREE [23] under the best-fit substitution model for 10000 ultrafast bootstraps [24]. The best-fit substitution model was selected using the Bayesian information criterion of ModelFinder [25] implemented in PhyloSuite 1.21 [22].

### Phylogenetic analysis

To investigate the transmission patterns of the sub-clade 2.3.4.4b2 H5N8 viruses, we performed a phylogeographic analysis using an asymmetric model, with Bayesian stochastic search variable selection implemented in BEAST version 1.10.4 [26]. We first confirmed that the sequences contained sufficient temporal signal by using TempEst (version 1.5.3, <http://tree.bio.ed.ac.uk/>) as described previously [27]. The best-fit substitution model was selected as described above. We computed the marginal likelihoods using path sampling and stepping-stone sampling [28] to compare two models of among-lineage rate variation (strict clock and uncorrelated lognormal relaxed clock) and three coalescent-based tree priors (constant size, exponential growth, and Bayesian skyline). The Bayes factors revealed support for the Bayesian skyline tree prior and uncorrelated lognormal relaxed clock model. Three independent MCMC analyses were run, each for 100 million iterations and sampling every 10 thousand iterations. All runs were performed via the CIPRES Science Gateway (<https://www.phylo.org>). Tracer version 1.7.1 [29] was used to check for sufficient sampling and convergence among the three sets of samples after discarding the first 10% of the samples as burn-in, and all parameters reached an adequate effective sample size (exceeding 200). The maximum clade credibility tree was computed in TreeAnnotator, a part of the BEAST package. The resulting diffusion rates were used to calculate the Bayes factors using Spread3 [30]. Significant migration pathways were determined based on the combination of a BF of  $\geq 3$  and mean indicator of  $\geq 0.5$ : decisively supported diffusions with BF  $\geq 1000$ , very strongly supported diffusion with  $100 \leq \text{BF} < 1000$ , strongly supported diffusions with  $10 \leq \text{BF} < 100$ , and supported diffusions with  $3 \leq \text{BF} < 10$ .

## Results

### The emergence and monitoring of HPAI H5

On 17 October 2020, two HPAI H5N8 viruses, classified as clade 2.3.4.4b, were detected for the first

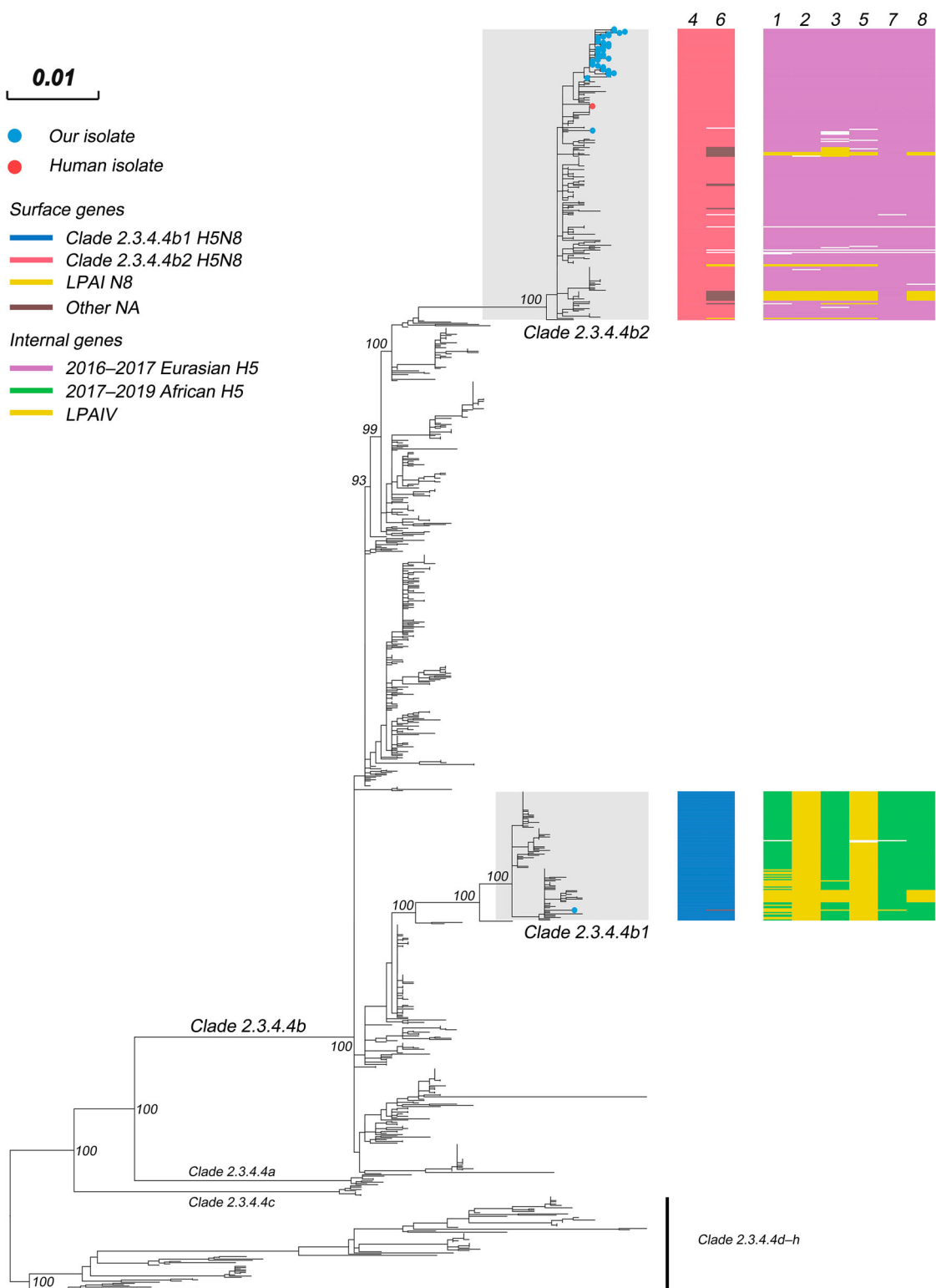
time in two dead swans, whooper swan (*Cygnus cygnus*) and mute swan (*Cygnus olor*) in Inner Mongolia, China [31]. Subsequently, intensified surveillance of HPAI H5 virus was conducted along the reported whooper swan migration routes [14]. Moreover, a total of 42 AIVs were isolated from wild birds in their habitats in central China during autumn of 2020 (Appendix Table 1). In addition to the initial two viruses in Inner Mongolia [31], four H5N8 viruses were detected in the key stopover sites of whooper swan in Shaanxi (Hekou Reservoir and Hongjian Nur) from November 9 to 10; 12 H5N8, 1 H5N2, and 1 H9N2 viruses were detected in Shanxi (Shengtian Lake and Sanwan wetland) on November 10; and 13 H5N8 viruses were detected in Henan (Sanmenxia Reservoir Area) from November 2–11; nine HPAI H5N8 viruses were also isolated in Longan Lake in Hubei, which is primarily the wintering ground of tundra swan, from November 4–16.

### Phylogenetic analysis

The whole genomic sequences of 42 AIVs, including 40 H5N8 viruses (2020–H5N8), 1 H5N2 virus, and 1 H9N2 virus, isolated in central China in autumn of 2020 were obtained (Appendix Table 2). In addition to a single H5N8 virus which was isolated from Eurasian eagle-owl (*Bubo bubo*), 41 AIVs were isolated from migratory waterfowl: whooper swan ( $n = 30$ ), tundra swan ( $n = 7$ ), mute swan ( $n = 1$ ), common teal (*Anas crecca*,  $n = 1$ ), bean goose (*Anser fabalis*,  $n = 1$ ), and whiskered tern (*Chlidonias hybrida*,  $n = 1$ ). The sequences of all isolates were deposited in to the GISAID (Appendix Table 2). All of the isolates were analyzed along with that of the H5N8 viruses isolated in 2020 and other closely related sequences.

The 2020-H5N8 sequences shared high nucleotide identity across all eight gene segments (99.2% – 100%) and were identified as HPAIVs based on the amino acid sequence REKRRKR↓GLFGAI at the haemagglutinin (HA) cleavage site. This finding indicates that they are descendants of a common ancestral virus. The H5N2 isolate, A/whooper swan/Shanxi/SX116/2020 (H5N2) (SX116–H5N2), was found to share the same HA cleavage site with 2020-H5N8, but exhibited a diverse origin as low homology was observed in seven of the genes (91.7% – 96.9%, except for the NA gene). Interestingly, SX116-H5N2 shared quite a high nucleotide identity (99.9% – 100%) with the H9N2 isolate, A/whooper swan/Shanxi/SX126/2020(H9N2) (SX126-H9N2), in six of the genes (PB2, PB1, PA, NP, NA, and M), revealing potential recombination events.

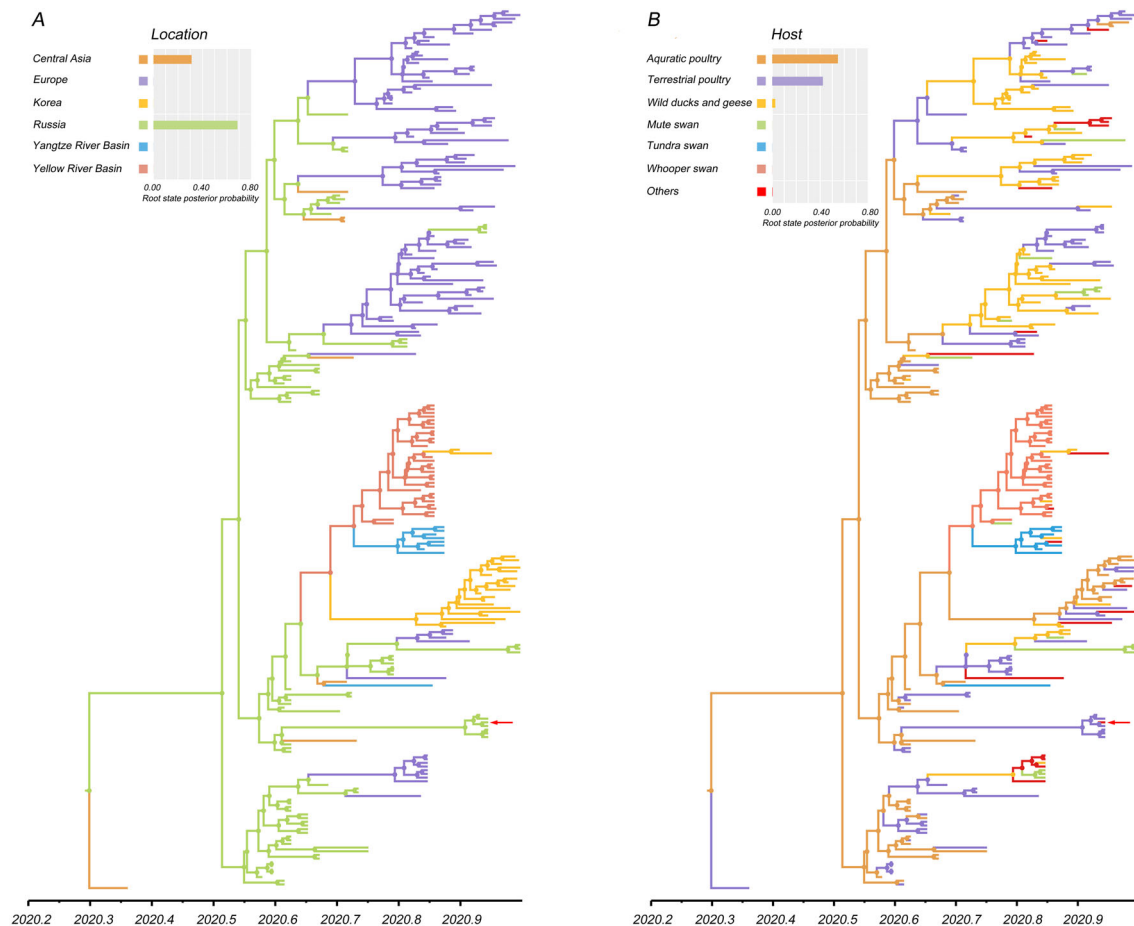
In the Maximum-likelihood (ML) phylogenies of the H5 gene, 2020-H5N8 ( $n = 40$ ) and SX116-H5N2 were both classified in 2.3.4.4b, and were further diverged into two phylogenetically distinct sub-clades: 2.3.4.4b1 and 2.3.4.4b2 (Figure 1) [7, 32, 33]. The high



**Figure 1.** Maximum-likelihood phylogenetic tree of the HA gene segment of HPAI H5 sequences. Our isolates and the Russian human strain are marked by blue and red coloured circles, respectively. For the clade 2.3.4.4b H5 virus isolated in 2020, the clade origins of each gene segment are indicated by different coloured bars. A UFBoot support values of the major branch are indicated. The scale bar represents the nucleotide substitutions per site. Detailed information is available in Supplementary Materials. See also Appendix Figure 1 and Appendix Table 7.

bootstrap value also supported the two novel sub-clades. Isolate SX116-H5N2 was divided into sub-clade 2.3.4.4b1 which was primarily detected in European poultry during early 2020, the immediate outlier virus was A/guinea fowl/Nigeria/

OGGF11T\_19VIR8424-7/2019(H5N8). The viruses of sub-clade 2.3.4.4b1 eventually spread to wild waterfowl in Asia in the autumn of 2020 (i.e. China, Japan [34], and Korea [35]); In May 2020, the sub-clade 2.3.4.4b2 was first emerged in Iraq and followed by



**Figure 2.** Maximum clade credibility (MCC) time-scaled phylogenetic tree of HA sequences of sub-clade 2.3.4.4b2 H5 virus coloured by geographic location (A), and host type (B). The branches are coloured according to the most probable ancestral geographic location, and host type. The Russian human strain is denoted by red arrows. The root state posterior probabilities are shown in each inset panel.

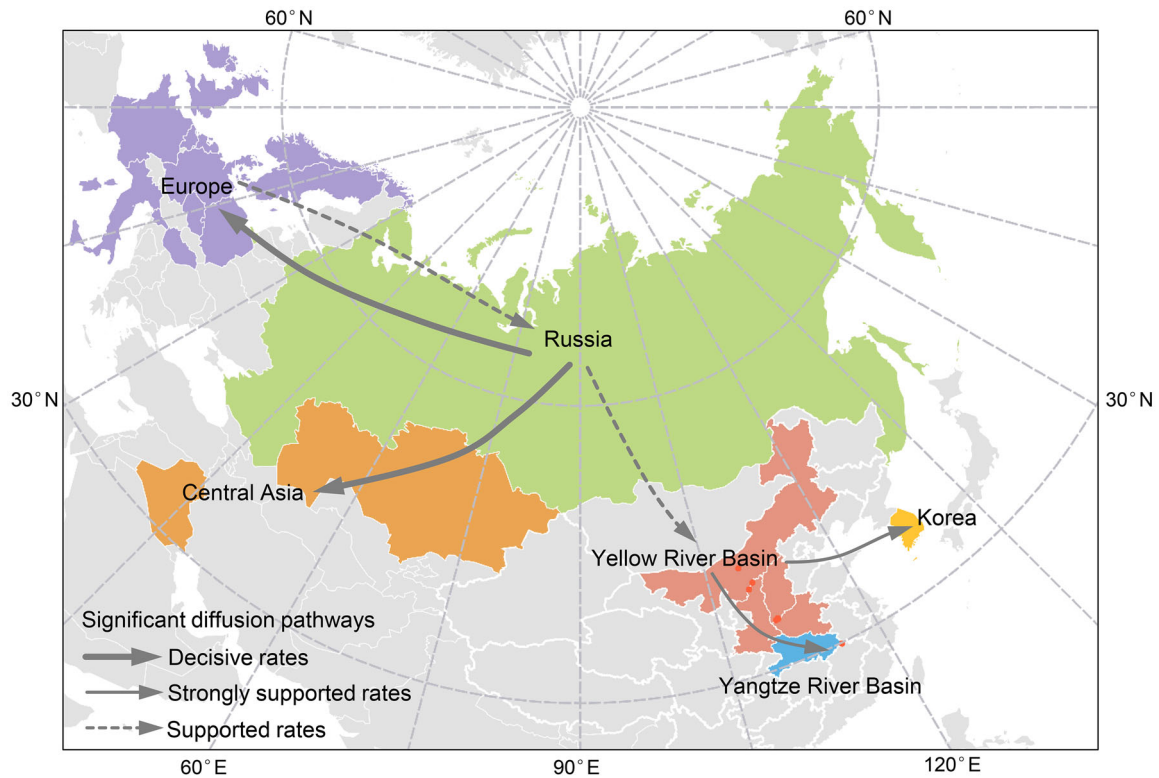
the detection of H5N8 in southern central Russia, Kazakhstan, Netherlands, and other European countries. All of our H5N8 isolates among the H5 viruses isolated in Eurasian in autumn of 2020 were clustered into sub-clade 2.3.4.4b2, of which the ancestral strains were widely circulated in Eurasia in 2017 and 2018.

The phylogenetic analysis also demonstrated that 2020-H5N8 tends to cluster in the complete genome, indicating a common evolutionary source (Appendix Figure 1). In contrast, in the ML trees, SX116-H5N2 further confirmed a process of genetic recombination in which only HA and NS genes were originated from the common ancestor of 2.3.4.4b1. The NA gene of SX116-H5N2 and SX126-H9N2 were closely related to those of low pathogenic avian influenza (LPAI) HxN2 viruses established in wild ducks in East Asia. In the PB2, PB1, PA, NP, and M phylogenies (Appendix Figure 1), SX116-H5N2 and SX126-H9N2 shared most recent common ancestor, which was derived from the Mongolia poultry LPAI virus gene pool. The NS gene of SX126-H9N2 was also derived from the LPAI virus gene pool of Mongolia poultry. The HA gene of SX126-H9N2 was clustered with the

2016/2018 viruses of waterfowl-origin in East Asia and the long branch lengths for SX126-H9N2 suggested that the virus had been circulating undetected for the intervening period.

### Phylogenetic analysis

To gain further insight into the large-scale transmission patterns of the sub-clade 2.3.4.4b2 H5N8 virus in Eurasia during autumn of 2020, and to clarify the contribution of different bird host populations to viral transmission, we performed a phylogeographic analysis of the HA gene sequences. We grouped the isolates into six distinct geographic categories: Central Asia (Iraq and Kazakhstan), Europe, Korea, Russia, Yangtze River Basin (Hubei, China), and the Yellow River Basin (Inner Mongolia, Shaanxi, Shanxi, and Henan, China). The AIV hosts were categorized into seven groups: aquatic poultry, terrestrial poultry, mute swan, tundra swan, whooper swan, wild ducks and geese (except mute swan, tundra swan, and whooper swan), others (Appendix Table 3). Our analyses support the idea that Russia was the geographic origin of sub-clade 2.3.4.4b2, with a posterior probability of



**Figure 3.** Spatial diffusion of the HA gene segment of the sub-clade 2.3.4.4b2 H5 virus. The red dots indicate the sampling site in central China. The bold grey arrow indicates decisively supported diffusions ( $BF \geq 1000$ ); solid grey arrows, strongly supported diffusions ( $10 \leq BF < 100$ ); and dashed grey arrows, supported diffusions ( $3 \leq BF < 10$ ).

0.68 (Figure 2A). Phylogeographic analysis indicates that there are six migration pathways in the spatial diffusion, with three originating from Russia to central Asia (Bayes factor,  $BF > 1,000$ ), Europe ( $BF > 1,000$ ), and the Yellow River Basin ( $BF = 8.64$ ), two from the Yellow River Basin to the Yangtze River Basin ( $BF = 26.05$ ) and Korea ( $BF = 42.46$ ), and the sixth from Europe to Russia ( $BF = 6.28$ ) (Figure 3; Appendix Table 4). As such, our analysis indicates that Russia acted as a primary seeding population while the Yellow River Basin established epidemiological links with other Asian geographical locations in this transmission network.

The most likely ancestral host population for sub-clade 2.3.4.4b2 is poultry, with a posterior probability of 0.97 (aquatic poultry, posterior probability = 0.55; terrestrial poultry, posterior probability = 0.42) (Figure 2B). Poultry acted as a viral source while several statistically significant transitions ( $BF \geq 3$ ) to other host populations were also supported (Appendix Table 5). One diffusion pathway with decisive support of  $BF > 1,000$  was recognized as originating from aquatic poultry to terrestrial poultry. A complex host history with a large amount of viral exchange was also indicated by several very strongly ( $100 \leq BF < 1000$ ) and strongly ( $10 \leq BF < 100$ ) supported diffusion pathways (Appendix Table 5). Our analysis reflected this dynamic with three species of swan playing different roles in spreading HPAI H5 virus: mute swan acted as a receiver, with only one statistically

significant transition from wild ducks and geese ( $BF = 36.08$ ); whooper swan acquired the virus from aquatic poultry ( $BF = 5.41$ ), and subsequently transmitted it to tundra swan ( $BF = 12.26$ ), wild ducks, and geese ( $BF = 76.07$ ), and other hosts ( $BF = 13.28$ ). Therefore, whooper swan acted as the seeding population in viral dissemination.

### Molecular characterization

In the H5 HA protein, the amino acid substitutions Q222L [36] and G224S [36] were not found; however, the S133A, S155N, T156A mutations [37, 38] were identified in all H5 isolates; T188I [37], V210I [39] mutations were found in partial H5 isolates, suggesting that these viruses may be able to potentially bind to human-type receptors ( $\alpha$ -2,6 galactose sialic acids) (Appendix Table 6). Residues Q591 [40], E627 [41], and D701 [40] in the PB2 protein suggest that these viruses have not yet adapted to mammalian hosts. Notably, the D3 V [42] and D622G [43] mutations in the PB1 protein, N383D [44] mutation in the PA protein, and M105 V [45] and A184 K [46] mutations in the NP protein were observed in all isolates, which have been reported to be associated with increased polymerase activity or virulence. Furthermore, pathogenicity-associated mutations (N30D [47], I43M [48], and T215A [47] in the M1 protein), as well as L98F, I101M [49], and T215A [50] in the NS1 protein were observed, which might enhance

viral virulence in mice. In addition, SX116-H5N2 exhibited the N200S mutation and PDZ motif, GSEV, in the NS1 protein and SX126-H9N2 displayed the PDZ motif, ESEV.

## Discussion

The re-emergence of H5N8 in Eurasia, especially its detection in human beings in Russia, poses a serious threat to public health. Recently, whooper swans have drawn increased attention as sentinel species in relation to HPAI H5 dissemination [14, 16]. Since swans in Russia have been reported to be infected with the HPAI H5N8 virus, we proposed an early warning system for HPAI outbreaks among whooper swan migratory routes in central China in accordance with the authenticated migration routes for swans [14, 15].

In the present study, we isolated 42 AIVs from wild birds along the Yellow River Basin (Inner Mongolia, Shaanxi, Shanxi, and Henan) and the Yangtze River Basin (Hubei), including 40 HPAI H5N8 viruses, 1 HPAI H5N2 virus (in Shanxi), and 1 LPAI H9N2 virus (in Shanxi). Our study conclusively demonstrated that the host distribution of the HPAI H5 virus was determined to be primarily in whooper swan in the Yellow River Basin (93.75%, 30/32) and in tundra swan in the Yangtze River Basin (77.78%, 7/9). The Yellow River Basin and the Yangtze River Basin constitute a complex wetland ecosystem in central China where the East Asian/Australasian and central Asian flyways overlap, providing a vital habitat for migratory waterfowl and playing an important intermediate role in connecting AIVs. A similar pattern was evident when compared to the prevalent of H5N8 viruses in 2016, as similar H5N8 viruses were detected in both in whooper swans in Shanxi (the Yellow River Basin) [51] and wild waterfowl in Hubei (the Yangtze River Basin) [52] in late 2016, following the outbreak of HPAI H5N8 in June 2016 in Ubsu-Nur Lake at the Russia-Mongolia border [53]. This finding suggests that HPAI H5N8 outbreaks in central China are closely associated with Russia.

Evolutionary genetic analyses indicate that the clade 2.3.4.4b H5 viruses detected in 2020 clustered into two distinct groups: viruses of sub-clade 2.3.4.4b1, which caused small outbreaks limited to the European poultry sector in early 2020, and spread to Asia by migratory waterfowl during autumn; and sub-clade 2.3.4.4b2, which was first detected in Iraq in May, and contained viruses primarily prevalent during autumn and winter of 2020 in Russia, Kazakhstan, Europe, and China. To further disentangle the host contributions in the spatial expansion of sub-clade 2.3.4.4b2 HPAI H5, we summarized the spatial and host dynamics. While we found that sub-clade 2.3.4.4b2 originated from Russian aquatic poultry,

we could not exclude the possibility that viruses originated from central Asian terrestrial poultry due to the presence of a single isolate in chicken in Iraq.

Outbreaks of sub-clade 2.3.4.4b2 HPAI H5 in southern central Russia were reported in aquatic poultry in late July 2020, after which the virus subsequently circulated in various hosts, including aquatic poultry, terrestrial poultry, migratory waterfowl, and ultimately impacted poultry workers [8]. In autumn, the virus spread from Russia to Kazakhstan, Europe, and China, respectively. Moreover, the virus underwent two historical host shifts before it leaped to humans in Russia. This followed a trajectory of transmission from aquatic poultry to terrestrial poultry, and finally to farm workers (Figure 2B). Therefore, measures should be taken to govern the spill-over risk of HPAI H5 virus in the poultry trade and to protect people who are at risk from possible HPAI pandemics (i.e. via poultry vaccination, strengthened quarantine, and enhanced surveillance in hotspots or sentinel birds).

The analysis of the HA gene segment of sub-clade 2.3.4.4b2 suggests that the strains identified in Autumn of 2020 in China had most likely been introduced from Russia (Russia to the Yellow River Basin, BF = 7.67). Moreover, whooper swans are likely the most responsible for the initial introductions. After the detection of the H5N8 virus in the dead whooper swan in Inner Mongolia (the first port of entry for whooper swans in central China) in October 2020, we strengthened the active monitoring of HPAI H5N8 along the migration routes of swans in central China. Eventually, 40 HPAI H5N8 viruses which exhibited high nucleotide identity were isolated. In the Yellow River Basin (Inner Mongolia, Shaanxi, Shanxi, and Henan), all H5N8 isolates (n = 31) displayed high nucleotide identity at the genomic level and viruses were primarily isolated from whooper swan (28/31). Moreover, several cross-species transmission events from whooper swan to mute swan, common teal, and Eurasian eagle-owl were also observed. Another important cross-species transmission event from whooper swan to tundra swan (BF = 12.26) caused the spread of the virus to the Yangtze River Basin. As such, whooper swan should be considered the sentinel species in central China for monitoring the invasion of HPAI viruses. In contrast to whooper swans, which overwinter in entirely restricted habitats along the Yellow River Basin, tundra swans are associated with a habitat that is widely distributed along the Yangtze River Basin, indicating the risk of further expansion of the epidemic. In addition to the populations of overwintering waterfowl, there is a developed poultry industry in the Yangtze River Basin. Thus, bird host species are brought together in a high-density setting that can facilitate viral persistence, cross-species infection,

and genetic reassortment, indicating the high pandemic potential of HPAI H5 viruses. Therefore, extensive surveillance systems are required to quickly detect HPAI H5 invasions and manage the outbreak in a timely manner.

Whooper swans are also responsible for the transmission of the HPAI H5N8 virus from the Yellow River Basin to Korea; however, there is no evidence of the direct migration of whooper swan from the Yellow River Basin to Korea. The absence of data suggests that it is necessary to strengthen the studies involving satellite tracking of migratory birds.

Although dead whooper swans have been occasionally found in their stopover sites (Inner Mongolia and Shaanxi), corpses have been continuously identified from November to December 2020 in wintering grounds (Henan). This indicates that intensified surveillance in the wintering grounds of whooper swan was necessary after the detection of HPAI H5 virus at their stopover sites. Furthermore, on 26 November 2020, after we detected the HPAI H5 viruses in apparently healthy whooper swan in the Sanwan wetland, Shanxi, dead whooper swans were reported. Early laboratory trials showed a decreased survival time (average of 4 days) following infection with the HPAI H5N1 virus of Asian lineage [13]. The longer duration of a live whooper swan infection would allow for ample time to travel and come into contact with other waterfowl. Thus, it was likely a factor that contributed to the spread of HPAI H5N8 viruses in China.

As early as 2016, the avian-origin HPAI H5N8 virus has been detected in dead grey seals (*Halichoerus grypus*) in Poland [54]. On December 2020, Russian poultry farm workers who participated in a response operation to a poultry H5N8 outbreak tested positive for HPAI H5N8 virus. These are the first reported detections of avian influenza A(H5N8) in humans. The E627 K and D701N substitutions in the PB2 protein have traditionally been thought to be critical for the mammalian adaptation of AIVs; however, no substitutions were observed in the Russia human isolate (A/Astrakhan/3212/2020 [H5N8]) or the HPAI H5 isolates. Multiple mutations (S133A, S154N, and T156A) in the HA protein of Russia human strain and our H5 isolates were identified. Moreover, T188I and V210I mutations were observed in partial isolates, suggesting that these viruses may increase the affinity of the virus to bind to  $\alpha$ 2,6 “human-type” receptors. Taken together, these findings demonstrate that H5N8 viruses pose a considerable threat to public health, despite the limited number of human cases.

## Acknowledgements

This work was also funded by the Surveillance of Wildlife Diseases from the National Forestry and Grassland Administration. We thank the authors and submitting laboratories

of the sequences from the Global Initiative on Sharing All Influenza Data (GISAID, <https://www.gisaid.org>) EpiFlu database. We thank International Science Editing (<http://www.internationalscienceediting.com>) for editing this manuscript.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by National Natural Science Foundation of China: [Grant Number 31970501].

## ORCID

Xiang Li  <http://orcid.org/0000-0003-2879-3957>

Hongliang Chai  <http://orcid.org/0000-0001-7162-2497>

## References

- [1] Webster RG, Bean WJ, Gorman OT, et al. Evolution and ecology of influenza A viruses. *Microbiol Rev.* 1992;56:152–179.
- [2] Lycett SJ, Bodewes R, Pohlmann A, et al. Role for migratory wild birds in the global spread of avian influenza H5N8. *Science.* 2016;354:213–217.
- [3] Lee YJ, Kang HM, Lee EK, et al. Novel reassortant influenza A(H5N8) viruses, South Korea, 2014. *Emerg Infect Dis.* 2014;20:1087–1089.
- [4] Lee DH, Torchetti MK, Winker K, et al. Intercontinental spread of asian-origin H5N8 to North America through beringia by migratory birds. *J Virol.* 2015;89:6521–6524.
- [5] Verhagen JH, Herfst S, Fouchier RAM. How a virus travels the world. *Science.* 2015;347:616–617.
- [6] Li MX, Liu HZ, Bi YH, et al. Highly pathogenic avian influenza A(H5N8) virus in wild migratory birds, qinghai lake, China. *Emerg Infect Dis.* 2017;23:637–641.
- [7] Lewis NS, Banyard AC, Whittard E, et al. Emergence and spread of novel H5N8, H5N5 and H5N1 clade 2.3.4.4 highly pathogenic avian influenza in 2020. *Emerg Microbes Infect.* 2021;10:148–151.
- [8] Pyankova OG, Susloparov IM, Moiseeva AA, 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.
- [9] Shi WF, Gao GF. Emerging H5N8 avian influenza viruses. *Science.* 2021;372:784–786.
- [10] Saito T, Tanikawa T, Uchida Y, et al. Intracontinental and intercontinental dissemination of Asian H5 highly pathogenic avian influenza virus (clade 2.3.4.4) in the winter of 2014–2015. *Rev Med Virol.* 2015;25:388–405.
- [11] Hesterberg U, Harris K, Stroud D, et al. Avian influenza surveillance in wild birds in the European Union in 2006. *Influenza Other Respir Viruses.* 2009;3:1–14.
- [12] Napp S, Majo N, Sanchez-Gonzalez R, et al. Emergence and spread of highly pathogenic avian influenza A(H5N8) in Europe in 2016–2017. *Transbound Emerg Dis.* 2018;65:1217–1226.
- [13] Brown JD, Stallknecht DE, Swayne DE. Experimental infection of swans and geese with highly pathogenic



- avian influenza virus (H5N1) of asian lineage. *Emerg Infect Dis.* 2008;14:136–142.
- [14] Meng W, Yang Q, Vrancken B, et al. New evidence for the east-west spread of the highly pathogenic avian influenza H5N1 virus between Central Asian and east Asian-Australasian flyways in China. *Emerg Microbes Infect.* 2019;8:823–826.
- [15] Fang L, Junjian Z, Zhao Q, et al. Two distinct flyways with different population trends of bewick's swan *Cygnus columbianus bewickii* in east Asia. *Wildfowl.* 2020: 13–42.
- [16] Bi Y, Zhang Z, Liu W, et al. Highly pathogenic avian influenza A(H5N1) virus struck migratory birds in China in 2015. *Sci Rep.* 2015;5:12986–12987.
- [17] Manu U. WHO Manual on Animal Influenza Diagnosis and Surveillance. WHO global influenza programme. (2002). [https://www.who.int/csr/resources/publications/influenza/whocdscsrncs20025\\_rev.pdf](https://www.who.int/csr/resources/publications/influenza/whocdscsrncs20025_rev.pdf).
- [18] Tavares ES, Baker AJ. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. *BMC Evol Biol.* 2008;8:1–14.
- [19] Lee DH, Lee HJ, Lee YN, et al. Application of DNA barcoding technique in avian influenza virus surveillance of wild bird habitats in Korea and Mongolia. *Avian Dis.* 2010;54:677–681.
- [20] Chai H. Molecular epidemiological study on influenza virus in wild birds of Heilongjiang. [dissertation]. Harbin (CN): Northeast Forestry University; 2012.
- [21] Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–780.
- [22] Zhang D, Gao FL, Jakovlic I, et al. Phylosuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol Ecol Resour.* 2020;20:348–355.
- [23] Nguyen LT, Schmidt HA, von Haeseler A, et al. IQ-tree: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 2015;32:268–274.
- [24] Minh BQ, Nguyen M, Haeseler AV. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol.* 2013;30:1188–1195.
- [25] Kalyaanamoorthy S, Bui Quang M, Wong TKF, et al. Modelfinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 2017;14:587–589.
- [26] Suchard MA, Lemey P, Baele G, et al. Bayesian phylogenetic and phylodynamic data integration using beast 1.10. *Virus Evol.* 2018;4:vey016.
- [27] Wei X, Chen M, Cui J. Bayesian evolutionary analysis for emerging infectious disease: an exemplified application for H7N9 avian influenza viruses. *Sci China Life Sci.* 2017;60:1392–1395.
- [28] Baele G, Lemey P, Bedford T, et al. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol Biol Evol.* 2012;29:2157–2167.
- [29] Rambaut A, Drummond AJ, Xie D, et al. Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Syst Biol.* 2018;67:901–904.
- [30] Filip B, Guy B, Bram V, et al. Spread3: interactive visualization of spatiotemporal history and trait evolutionary processes. *Mol Bio Evol.* 2016;33:2167–2169.
- [31] Li X, Lv X, Li Y, et al. Highly pathogenic avian influenza A(H5N8) virus in swans, China, 2020. *Emerg Infect Dis.* 2021;27:1732–1734.
- [32] He G, Ming L, Li X, et al. Genetically divergent highly pathogenic avian influenza A(H5N8) viruses in wild birds, Eastern China. *Emerg Infect Dis.* 2021;27:2940–2943.
- [33] Cui P, Zeng X, Li X, et al. Genetic and biological characteristics of the globally circulating H5N8 avian influenza viruses and the protective efficacy offered by the poultry vaccine currently used in China. *Sci China Life Sci.* 2021;14. DOI: 10.1007/s11427-021-2025-y
- [34] Isoda N, Twabela AT, Bazarragchaa E, et al. Re-invasion of H5N8 high pathogenicity avian influenza virus clade 2.3.4.4b in Hokkaido, Japan, 2020. *Viruses.* 2020;12:1439.
- [35] Jeong S, Lee DH, Kwon JH, et al. Highly pathogenic avian influenza clade 2.3.4.4b subtype H5N8 virus isolated from mandarin duck in South Korea, 2020. *Viruses.* 2020;12:1389.
- [36] Chutinimitkul S, van Riel D, Munster VJ, et al. In vitro assessment of attachment pattern and replication efficiency of H5N1 influenza A viruses with altered receptor specificity. *J Virol.* 2010;84:6825–6833.
- [37] Yang ZY, Wei CJ, Kong WP, et al. Immunization by avian H5 influenza hemagglutinin mutants with altered receptor binding specificity. *Science.* 2007;317:825–828.
- [38] Wang W, Lu B, Zhou H, et al. Glycosylation at 158N of the hemagglutinin protein and receptor binding specificity synergistically affect the antigenicity and immunogenicity of a live attenuated H5N1 A/Vietnam/1203/2004 vaccine virus in ferrets. *J Virol.* 2010;84:6570–6577.
- [39] Watanabe Y, Ibrahim MS, Ellakany HF, et al. Acquisition of human-type receptor binding specificity by new H5N1 influenza virus sublineages during their emergence in birds in Egypt. *PLoS Pathog.* 2011;7:e1002068.
- [40] Taft AS, Ozawa M, Fitch A, et al. Identification of mammalian-adapting mutations in the polymerase complex of an avian H5N1 influenza virus. *Nat Commun.* 2015;6:7491.
- [41] Hatta M, Hatta Y, Kim JH, et al. Growth of H5N1 influenza A viruses in the upper respiratory tracts of mice. *PLoS Pathog.* 2007;3:1374–1379.
- [42] Elgendy EM, Arai Y, Kawashita N, et al. Identification of polymerase gene mutations that affect viral replication in H5N1 influenza viruses isolated from pigeons. *J Gen Virol.* 2017;98:6–17.
- [43] Feng X, Wang Z, Shi J, et al. Glycine at position 622 in PB1 contributes to the virulence of H5N1 avian influenza virus in mice. *J Virol.* 2016;90:1872–1879.
- [44] Song J, Xu J, Shi J, et al. Synergistic effect of S224P and N383D substitutions in the PA of H5N1 avian influenza virus contributes to mammalian adaptation. *Sci Rep.* 2015;5:10510.
- [45] Tada T, Suzuki K, Sakurai Y, et al. NP body domain and PB2 contribute to increased virulence of H5N1 highly pathogenic avian influenza viruses in chickens. *J Virol.* 2011;85:1834–1846.
- [46] Wasilenko JL, Sarmento L, Pantin-Jackwood MJ. A single substitution in amino acid 184 of the NP protein alters the replication and pathogenicity of H5N1 avian influenza viruses in chickens. *Arch Virol.* 2009;154:969–979.
- [47] Fan S, Deng G, Song J, et al. Two amino acid residues in the matrix protein M1 contribute to the virulence difference of H5N1 avian influenza viruses in mice. *Virology.* 2009;384:28–32.

- [48] Nao N, Kajihara M, Manzoor R, et al. A single amino acid in the M1 protein responsible for the different pathogenic potentials of H5N1 highly pathogenic avian influenza virus strains. *Plos One*. 2015;10: e0137989.
- [49] Kuo RL, Krug RM. Influenza a virus polymerase is an integral component of the CPSF30-NS1A protein complex in infected cells. *J Virol*. 2009;83:1611–1616.
- [50] Li Z, Jiang Y, Jiao P, et al. The NS1 gene contributes to the virulence of H5N1 avian influenza viruses. *J Virol*. 2006;80:11115–11123.
- [51] Jiang W, Tian W, Hou G, et al. Analysis of swan H5N8 avian influenza in holy lake of Shanxi province and the current virus lineage of H5 subtype avian influenza. *China Animal Health Inspection*. 2017;34:14–17.
- [52] Ma L, Jin T, Wang H, et al. Two reassortant types of highly pathogenic H5N8 avian influenza virus from wild birds in central China in 2016. *Emerg Microbes Infect*. 2018;7:14.
- [53] Lee DH, Sharshov K, Swayne DE, et al. Novel reassortant clade 2.3.4.4 avian influenza A(H5N8) virus in wild aquatic birds, russia, 2016. *Emerg Infect Dis*. 2017;23:359–360.
- [54] Shin DL, Siebert U, Lakemeyer J, et al. Highly pathogenic avian influenza A(H5N8) virus in gray seals, baltic Sea. *Emerg Infect Dis*. 2019;25:2295–2298.