Virus Evolution, 2024, **10(1)**, veae075
DOI: https://doi.org/10.1093/ve/veae075
Advance Access Publication 14 September 2024
Research Article

Evolutionary characterization of the establishment of H6 influenza viruses in domestic geese in China: implications for the position of the host in the ecosystem

Liuxia Peng[1], Ziying Jin[2,3], Peiwen Chen[1,2,3], Zengfeng Zhang[4], Xiaohui Fan[4], Wenshan Hong[1], Yongmei Liu[1,2,3,5], David K. Smith[2,5], William Yiu-Man Cheung[1,2,3,5], Jia Wang[1,2], Huachen Zhu[1,2,3,5,*], Tommy Tsan-Yuk Lam[1,2,3,5,*], Yi Guan[1,2,3,5,6,*]

[1]Guangdong-Hong Kong Joint Laboratory of Emerging Infectious Diseases/Joint Laboratory for International Collaboration in Virology and Emerging Infectious Diseases, Joint Institute of Virology (STU/HKU), Shantou University Medical College, Shantou University, 243 Daxue Road, Shantou, Guangdong 515063, China
[2]State Key Laboratory of Emerging Infectious Diseases (SKLEID), School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China
[3]EKIH (Gewuzhikang) Pathogen Research Institute, Shenzhen, Guangdong 518045, China
[4]Department of Microbiology, Guangxi Medical University, 22 Shuangyong Road, Nanning, Guangxi 530021, China
[5]Laboratory of Data Discovery for Health, 19W Hong Kong Science & Technology Parks, Hong Kong SAR, China
[6]Shanghai Institute of Virology, Shanghai Jiao Tong University School of Medicine, 227 South Chongqing Road, Shanghai 200225, China

* Corresponding author. Yi Guan. Shanghai Institute of Virology, Shanghai Jiao Tong University School of Medicine, 227 South Chongqing Road, Shanghai 200225, China. E-mail: yguan@hku.hk; Tommy Tsan-Yuk Lam. Laboratory of Data Discovery for Health, 19W Hong Kong Science & Technology Parks, Hong Kong SAR, China. E-mail: ttylam@hku.hk; Huachen Zhu. School of Public Health, The University of Hong Kong, Hong Kong SAR, China. E-mail: zhuhch@hku.hk

Abstract

Geese, both wild and domestic, are generally considered part of the natural reservoir for influenza A viruses. The highly pathogenic H5 Goose/Guangdong avian influenza virus lineage that is still causing outbreaks worldwide was first detected in domestic geese in 1996. However, while wild geese might have a somewhat restricted role in the influenza ecosystem, the role of domestic geese is little studied. Here, 109 H6 viruses isolated from domestic geese during 2001–2018 in southern China had their phylogeny, evolutionary dynamics, and molecular signatures characterized to examine the role of domestic geese. Our findings demonstrated that all geese H6 viruses were derived from H6 viruses established in ducks and that they subsequently formed three distinct hemagglutinin lineages. Rapid evolution of the hemagglutinin genes was not detected after the duck-to-goose transmissions of H6 viruses that then circulated in geese. Despite long-term persistence in geese, H6 viruses were rarely observed to transmit back to ducks or terrestrial poultry and never exchanged genes with viruses from other subtypes. Most geese H6 viruses maintained the primary molecular signatures of their duck precursors. This study raises the possibility that, rather than being part of the natural reservoir, domestic geese might be more like an aberrant host species for influenza A viruses, and perhaps a "dead-end" host.

Keywords: domestic goose; H6 influenza virus; establishment; evolutionary dynamic.

1. Introduction

Aquatic birds of the orders Anseriformes (ducks, geese, and swans) and Charadriiformes (shorebirds and gulls) are generally regarded as the natural reservoir of influenza A viruses (Webster et al. 1992). Wild waterfowl host a high diversity of lowly pathogenic avian influenza viruses (LPAIVs) in generally asymptomatic infections, with limited viral evolution at the amino acid level, and these fowl are the fundamental source of influenza A viruses infecting domestic poultry and mammals (Webster et al. 1992). Of the LPAIV subtypes, H5 and H7 may sporadically evolve into highly pathogenic avian influenza viruses (HPAIV) by acquiring multiple basic amino acids at the cleavage site of the *HA* protein, with severe consequences, primarily for infected terrestrial poultry (Swayne and Suarez 2000, Banks et al. 2001, Qi et al. 2018). Although domestic waterfowl are also regarded as a major part of the natural avian influenza virus (AIV) reservoir, viruses in these hosts can evolve significantly differently from those in their wild counterparts (Huang et al. 2012, Yoon et al. 2014).

Two major types of domestic waterfowl (ducks and geese) are raised in China, with >70% of the ducks and 93% of the geese farmed in the world currently held there (https://www.fao.org/faostat/en/#data/QCL). The increasing population size and density of poultry farmed in China have impacted the evolution of influenza viruses. The ongoing evolution of H5 and H6

viruses in domestic waterfowl has given rise to distinct virus lineages, as commonly seen in viruses of terrestrial poultry (C. L. Guo et al. 2000, Guan et al. 2002, Cheung et al. 2007, Huang et al. 2010).

Domestic ducks also act as an intermediate host transmitting AIVs to terrestrial poultry (Bahl et al. 2009, Huang et al. 2010, 2012). Interactions between influenza A viruses of wild birds and domestic ducks in China over the last two decades have led to extensive reassortment events between enzootic poultry virus lineages and those from wild birds. This has then contributed to the generation of multiple zoonotic influenza viruses (e.g. H3N8, H7N9, and H10N8) (Lam et al. 2013, Ma et al. 2015, Chen et al. 2023).

Wild and domestic geese are currently considered as part of the natural reservoir of AIVs, although migratory geese might have a more limited role in disseminating AIVs (Jonassen and Handeland 2007, Hoye et al. 2010, Kleijn et al. 2010, Samuel et al. 2015, Yin et al. 2017). Since the detection of the persistent and highly pathogenic Goose/Guangdong (GS/GD) H5Nx lineage, which remains enzootic in domestic geese (Xu et al. 1999), the role of these birds in the genesis of this lineage remains unclear. Domestic geese, even at increased population densities, mostly graze in grassland and sometimes congregate to roost on water bodies (Green and Elmberg 2014), so they could share AIVs with waterbirds and terrestrial poultry. How AIVs evolve in domestic geese and whether these birds might have a similar role to domestic ducks, or act more like a restricted, aberrant, or terrestrial poultry host in the AIV ecosystem has not been systematically explored.

H6 subtype viruses provide an opportunity to explore this as they are widely distributed in wild birds, domestic ducks, geese, and terrestrial poultry (Webby et al. 2003, Huang et al. 2010, Hu et al. 2020). Routine surveillance studies showed that domestic ducks were the predominant host for H6 subtype viruses (Li et al. 2019, Xu et al. 2023). Since 2000, two lineages (group I and group II) of H6 viruses became established in domestic ducks in China (Huang et al. 2010, 2012), with group II subsequently becoming predominant (Huang et al. 2012). After 2011, the rate of virus isolation gradually increased in domestic geese, surpassing that in domestic ducks on some occasions (Luo et al. 2017, Peng et al. 2018). However, the ecological role played by domestic geese in the distribution of H6 viruses is not known.

In this study, our data accumulated from 2001 to 2018 show that viruses of the H6 subtype were frequently detected in domestic geese in southern China. To define the role of domestic geese, we conducted a comprehensive phylogenetic analysis of the genomic evolution of these geese H6 viruses, and evaluated the evolutionary rate of, and the selection pressure acting on, the *HA* gene to interpret their viral evolutionary dynamics, gene reassortment patterns and molecular signatures.

2. Materials and methods

2.1 Surveillance and virus isolates

The record of influenza surveillance in domestic geese in southern China from 2001 to 2018 is summarized in Table 1. Surveillance was conducted in relatively stable, actively trading, live poultry markets (LPMs) on market geese in Guangdong and Guangxi provinces. Samples were collected every 2 weeks, to avoid the possibility of resampling a goose and, while the numbers of geese at market varied due to farming and commercial considerations, efforts were made to sample at least 20% of available birds. Oropharyngeal and cloacal swabs were taken if accessible, otherwise isolated fecal droppings were collected in goose stalls where no other poultry were sold. Samples were immediately preserved in viral transport medium supplemented with antibiotics and placed in 4°C cool boxes before sending to the laboratory. They were then inoculated into 9- to 10-day-old embryonated eggs and incubated for 48–72 h at 37°C for virus isolation. Hemagglutinin-positive allantoic fluids were subtyped by hemagglutinin inhibition (HI) assays using a panel of World Health Organization (WHO) reference antisera. All operations followed WHO guidelines on Animal Influenza Diagnosis and Surveillance.

2.2 Species-origin identification

The species of origin for fecal samples was identified by the method used in previous publications (P. P. Cheung et al. 2009). Host genomic DNA was extracted using a QIAamp Fast DNA Stool Mini Kit (QIAGEN). The avian mitochondrial cytochrome oxidase I gene was amplified by nested PCR using two sets of primers. Bird universal primers were Ext-F(5′-TGTAAAAAGGWCTACAGC CTAACGC-3′) and Ext-R (5′-GTRGCNGAYGTRAARTATGCTCG-3′). Internal primers were Int-F(5′-AACAAACCACAAAGATATCGG-3′) and Int-R (5′-TGGGARATAATTCCRAAGCCTGG-3′). Amplified PCR products were subjected to Sanger sequencing on an ABI3730 genetic analyzer (Applied Biosystems), according to the manufacturer's default protocols. DNA sequences were assembled and compiled using Lasergene version 8.0 (DNASTAR). Their compiled sequences were searched against the NCBI GenBank reference database (Clark et al. 2016) using blastn to collect nearest neighbors based on high coverage and closely related hits in reference databases.

2.3 Full-genome sequencing

Approximately 20% of the goose H6 HI positive isolates were proportionately randomly selected for full-genome sequencing, taking account of the province, sampling date, and isolation rate to achieve as representative as possible a set of sequences across the surveillance period. For fecal droppings, only samples confirmed as from a goose by identifying the species of origin were subjected to sequencing. Viral RNA was extracted using QIAamp Viral RNA Mini Kits (QIAGEN) and treated with DNase (QIAGEN) to remove DNA from embryonated eggs. Complementary DNA was synthesized using PrimeScript II 1st Strand cDNA Synthesis Kits (TaKaRa Bio) with uni12 primers and other influenza-specific primers described previously (Hoffmann et al. 2001). DNA libraries were constructed using the TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme), and were subjected to standard sequencing protocols using MiSeq Reagent Kits v2 (Illumina) on the MiSeq desktop sequencer (Illumina) according to the manufacturer's instructions. Low-quality and short raw reads generated from the standard sequencing process were removed using PRINSEQ. From these preprocessed reads, 3000 reads were randomly selected and assembled into contigs with overlapping regions of 40 nucleotides and >90% identity using the GS De Novo Assembler (version 2.6, Roche). These temporary contigs were searched against the influenza A virus sequences available from GenBank using blastn. The most similar sequences were used as the templates to assemble the rest of the preprocessed reads using MIRA. Samples containing more than one copy of a gene with <97% sequence identity were considered as mixed infections and were removed prior to phylogenetic analyses. The full genomes of 109 unmixed H6 goose viruses obtained in this study have been submitted to GenBank.

Table 1. Isolation of H6 influenza viruses in domestic geese in southern China from 2001 to 2018.

Year	No. of samples[a]			H6[b] (%) |Seq[c]			N_x[d]
	GD	GX	Total	GD	GX	Total	(GD, GX)
2001	363	/	363	18 (4.96) |3	/	18 (4.96) |3	N2 (3, 0)
2002	467	/	467	29 (6.21) |3	/	29 (6.21) |3	N2 (3,0)
2003	/	/	/	/	/	/	
2004	113	481	594	0	0	0	
2005	5374	938	6312	3(0.06) |2	0	3 (0.05) |2	N2 (2, 0)
2006	6435	909	7344	5(0.08) |2	2(0.22) |1	7 (0.10) |3	N2 (0, 1), N6 (2, 0)
2007	/	727	727	/	9(1.24) |2	9 (1.24) |2	N6 (0, 1), N8 (0, 1)
2008	52	576	628	1(1.92) |1	17(2.95) |3	18 (2.87) |4	N2 (1, 1), N6 (0, 2)
2009	418	709	1127	3 (0.72) |2	8 (1.13) |2	11 (0.98) |4	N2 (2, 1), N6 (0, 1)
2010	2286	457	2743	151 (6.61) |8	13 (2.84) |1	164 (5.98) |9	N2 (6, 0), N6 (2, 1)
2011	1240	560	1800	2 (0.16) |1	26 (4.64) |4	28 (1.56) |5	N2(1, 1), N6 (0, 3)
2012	1270	491	1761	32 (2.52) |12	21 (4.28) |6	53 (3.01) |18	N2 (12,4), N6 (0, 2)
2013	1202	561	1763	41 (3.41) |10	35(6.24) |4	76 (4.31) |14	N2 (10, 0), N6 (0, 4)
2014	1252	765	2017	40 (3.19) |5	21 (2.75) |4	61 (3.02) |9	N2 (5, 1), N6 (0, 3)
2015	1200	1183	2383	0	86 (7.27) |8	86 (3.61) |8	N6 (0, 8)
2016	1050	992	2042	0	47 (4.74) |3	47 (2.30) |3	N2 (0, 1), N6 (0, 2)
2017	1206	1215	2421	45 (3.73) |6	147 (12.1) |12	192 (7.93) |18	N2 (3, 9), N6 (3, 3)
2018	1200	1201	2401	0	91 (7.58) |4	91 (3.79) |4	N2 (0, 4)
Total	25,128	11,765	36,893	370 (1.47) |55	523 (4.45) |54	893 (2.42) |109	N2 (48, 23), N6 (7, 30), N8 (0, 1)

[a]Samples were collected in two provinces/regions, GD (Guangdong) and GX (Guangxi).
[b]The isolation rate of H6 virus is shown in parentheses.
[c]Only the numbers of full-length unmixed H6 sequences determined in this study are shown. Initially, 20% of the H6 isolates were randomly selected for sequencing.
[d]The NA subtype of the H6 sequences is shown.
Slash, no sampling.

2.4 Sequence preparation and alignment

Full-genome sequences of the 109 pure goose H6 viruses were combined with 839 full- and 1463 partial-genome sequences of H6 avian influenza A viruses, along with 7980 avian influenza viruses of other subtypes with full-genome sequences retrieved from GISAID (http://www.gisaid.org) in December 2022 for phylogenetic analyses. A total of 770 viruses (e.g. H5, H6, H7, H9, and H10) of avian origin found in mammals were also included for phylogenetic analyses. Only complete gene segments were selected from GISAID. Sequences with ambiguous bases (>0.5%) or of short length (<95% full-length) were removed from the datasets. Multiple sequence alignments were performed for each gene segment (H6 *HA*, N2, and N6 *NA*, *PB2*, *PB1*, *PA*, *NP*, *M* and *NS*) using MAFFT v7.273 with manual adjustment in BioEdit (Katoh and Standley 2013).

2.5 Phylogenetic reconstruction

To preliminarily assess the phylogenetic placement of the domestic geese AIV isolates, an initial global panoramic maximum likelihood (ML) phylogeny for each gene segment was inferred using the fast ML method, Jukes-Cantor + CAT model, and standard SH support values in FastTree v2.1.11 (Price et al. 2010). All the goose H6 isolates obtained from this study clustered with the major Eurasian avian lineage, therefore viruses from the major American lineage were down-sampled and were set as the outgroup in the refined phylogenetic analyses. The datasets were reduced by retaining virus sequences phylogenetically related to the goose H6 isolates, together with selected virus sequences as representatives of each more distant monophyletic clade (e.g. H5 and H9 lineage). ML phylogenies of the gene segments of these viruses were inferred using the GTR-I-Γ4 substitution model implemented in IQ-TREE v1.6.12 (Nguyen et al. 2015), with 1000 bootstrap replicates. Acknowledgment of the sources of the GISAID sequences used in this refined phylogenetic study is given in Supplementary dataset S1. The relative phylogenetic positions of the established duck and goose viruses in the phylogenies of the *HA* and other genes were traced by lines using the *ggtree* package (version 3.4.4) (Guangchuang et al. 2017) in *R*. To avoid misunderstanding of reassortment patterns, caused by uneven locations of viruses in the phylogeny, the branch lengths of ML trees are not displayed in Fig. 2 and Supplementary Fig. S4.

2.6 Bayesian phylogeny and molecular clock inference

For the H6 *HA* gene, the time-scaled maximum clade credibility (MCC) phylogeny and substitution rates were estimated using the Markov Chain Monte Carlo (MCMC) method implemented in BEAST v1.10.4 (Suchard et al. 2018). The GTR model with gamma site heterogeneity with invariant sites and three portions of position 1, 2, 3 was adopted as the substitution model. The clock model and tree prior were defined by the Bayes Factor method and Bayesian model selection through marginal likelihood estimation by path-sampling and stepping-stone sampling. Four clock models of a relaxed molecular clock with (i) uncorrelated lognormal distribution and Bayesian SkyGrid coalescent model, (ii) uncorrelated lognormal distribution and GMRF Bayesian Skyride model, (iii) uncorrelated exponential distribution and Bayesian SkyGrid coalescent model, and (iv) uncorrelated exponential distribution and GMRF Bayesian Skyride model, were tested. Multiple independent MCMC runs of 4×10^8 steps with sampling at every 10,000 steps were performed on the H6 *HA* gene. Convergence of MCMC parameters was evaluated in Tracer v1.7.1 (http://tree.bio.ed.ac.uk/software/tracer/), with effective sample sizes all larger than 200. MCC trees were generated using Tree Annotator v1.10.4, with 10% of MCMC runs removed as burn-in. Evolutionary rates for individual branches (particularly for those leading to host species transfers) were referred to as episodic rates. Visualization was achieved with Figtree v1.4.4 and annotated by the *ggtree* package (version 3.4.4) (Guangchuang et al. 2017) in *R*.

2.7 Estimation of selection pressure

To detect the adaptive molecular evolution, selection pressure (ω = dN/dS) acting on particular branches were inferred using CodeML as implemented in PAML (version 4.9) (Yang 2007, Alvarez-Carretero et al. 2023) and by FitMG94 as implemented in HyPhy (version 2.5.62) (Pond et al. 2007). A branch model with two-ratios was used to test if ω varies across the specific branches of the H6 *HA* gene tree and ω ratios with confidence intervals were inferred separately across branches of the H6 *HA* gene tree and specific clades. These clades were (i) Wild bird-like, (ii) DK-Group I, (iii) DK-Group II, (iv) GS-1, (v) GS-2, and (vi) GS-3 and the specific individual branches ancestral to the DK and GS clades were also examined.

2.8 Ancestral state reconstruction and substitution analysis

Substitutions that occurred along the branches were inferred from the ancestral nucleotide sequences reconstructed by HyPhy 2.5.42 at both ends of the branches in the MCC tree, using the General Reversible Model (GRM), with Global variation and a Gamma Distribution. Sites of sequences at nodes flanking the branches showing interspecies transmission were assessed for episodic positive selection by the mixed effects model of evolution (MEME) (Murrell et al. 2012) with $P \leq .05$. Sites under pervasive positive selection were assessed by Single-Likelihood Ancestor Counting (SLAC) (Kosakovsky Pond and Frost 2005) and Fixed Effects Likelihood (FEL) (Kosakovsky Pond and Frost 2005), using $P \leq .05$ and Fast Unconstrained Bayesian AppRoximation (FUBAR) (Murrell et al. 2013) with posterior probability ≥.9 in the Datamonkey Adaptive Evolution Server (https://datamonkey.org).

3. Results

3.1 H6 viruses were detected persistently in domestic geese

From 2001 to 2018, a total of 36,893 samples were collected from apparently healthy domestic geese at markets (Table 1). From 2001 onwards, H6 viruses were regularly isolated at rates up to 7.93% (Table 1). A total of 893 goose H6 viruses were isolated, giving an overall isolation rate of 2.42% (Table 1). In our earlier surveillance program, from 2000 to 2007, H6 viruses were identified year-round in domestic ducks, with *NA* subtypes of N1, N2, N5, N6, N8, and N9 (Huang et al. 2010, 2012). In geese, almost all the H6 viruses had an N2 or N6 *NA*, with the N8 *NA* gene rarely observed (Table 1). Although H6N2 viruses were generally replaced by H6N6 after 2005 in the duck population (Huang et al. 2012), both *NA* subtypes co-circulated in domestic geese during the surveillance period (Table 1).

3.2 Group I and II H6 viruses co-circulated and became established in domestic geese

To trace the origin and evolution of the H6 viruses in domestic geese, their whole-genome sequences were obtained and phylogenetically analyzed, together with influenza sequences available in GISAID. After pruning of the dataset (see "Materials and Methods" section), 159 H6 avian influenza A viruses were retained for the H6 *HA* analysis and 99 avian influenza A viruses of other subtypes, including those of avian origin found in mammals, were also included in the internal gene datasets. In the *HA* phylogeny, all goose H6 isolates were derived from the group I or group II lineages established in ducks (Huang et al. 2010) (Fig. 1, Supplementary Fig. S1). No H6 viruses from geese clustered with viruses from the natural reservoir or gene pool (Webster et al. 1992) maintained by the diverse populations of wild aquatic birds or with those from terrestrial poultry. Of the 109 full-length goose H6 viruses, 48.6% ($n = 53$) of the H6 *HA* genes were introduced from group I duck viruses, and 51.4% ($n = 56$) from group II. Although spillovers of duck H6 viruses to geese were occasionally observed during 2001–2018, only three introductions from duck viruses were successfully established in geese. These were named as clades GS-1, GS-2, and GS-3 (Fig. 1a, Supplementary Fig. S1).

Clade GS-1 ($n = 46$) contained only H6N2 subtype viruses, with an N2 *NA* gene also introduced from group I viruses (Fig. 1b). These viruses circulated in Guangdong and Guangxi provinces during 2008–2018, even though their source group I viruses could not be detected in the duck population from 2005 (Huang et al. 2012). Within these 11 years (2008–2018) of circulation, transmission of H6 viruses to ducks or chickens was rarely observed in this clade (Figs. 1 and 2, Supplementary Figs. S1 and S2).

The GS-2 and GS-3 clades originated from duck group II viruses. They cocirculated with the GS-1 clade for at least 8 years but had different *NA* subtypes (Fig. 1a, Supplementary Fig. S1). Clade GS-2 included viruses with N2 and N6 subtypes, that both originated from group II viruses (Fig. 1b–c). GS-2 H6N2 viruses were detected in Guangdong during 2009–2014 ($n = 15$), then the N6 *NA* gene replaced the N2 after viruses disseminated to Guangxi. Viruses from clade GS-3 were all H6N6, but the N6 genes originated from separate introductions from duck group II viruses (Fig. 1c). Clade GS-3 circulated locally in Guangxi during 2007–2015, acquiring a separate N6 gene. Over the period that the GS-2 and GS-3 clade viruses circulated, no isolate from wild birds, ducks, or terrestrial poultry phylogenetically clustered with them (Supplementary Fig. S1).

Phylogenetic analyses of the N2 and N6 *NA* genes showed that duck group I/II lineage viruses also reassorted with the N2 gene of H3 and H4 subtype viruses and with the N6 gene of H5 viruses (Fig. 1b–c). However, this did not occur with the goose H6 viruses.

The internal gene phylogenies of goose isolates also formed separate monophyletic clades that originated from duck H6 viruses (Supplementary Fig. S2). After multiple interspecies transmissions from ducks, viruses of clades GS-1, GS-2, and GS-3 were interspersed in the internal gene phylogenetic trees, indicating continuous reassortment and evolution of the viruses in geese. Almost all internal genes of the goose H6 lineages originated from group I/II duck H6 viruses. Only the *PA* gene of the GS-1 clade was predominantly derived from natural reservoir viruses in a monophyletic clade with 100% bootstrap support value (Supplementary Fig. S2-C).

Goose H6 isolates were never found to have contributed genes to viruses of other lineages, including the H5 and H9 subtypes, which are both enzootic in China. Although a group of H5N6 viruses obtained N6 and *PB2* genes from duck H6 viruses (Bi et al. 2016), goose H6 isolates were not involved in the genesis of those H5 viruses for any of the eight gene segments (Fig. 1c and Supplementary Fig. S2-A).

3.3 H6 cross-species transmission did not cause rapid virus evolution

(1) Dating the most recent common ancestor of the goose H6 viruses

The timings of transmission events were evaluated by Bayesian phylogenetic reconstruction, to give maximum clade credibility trees (Supplementary Fig. S3). A relaxed molecular clock with uncorrelated lognormal distribution and Bayesian SkyGrid coalescent model was used as this outperformed other models (Supplementary Table S1). The times to the most recent common

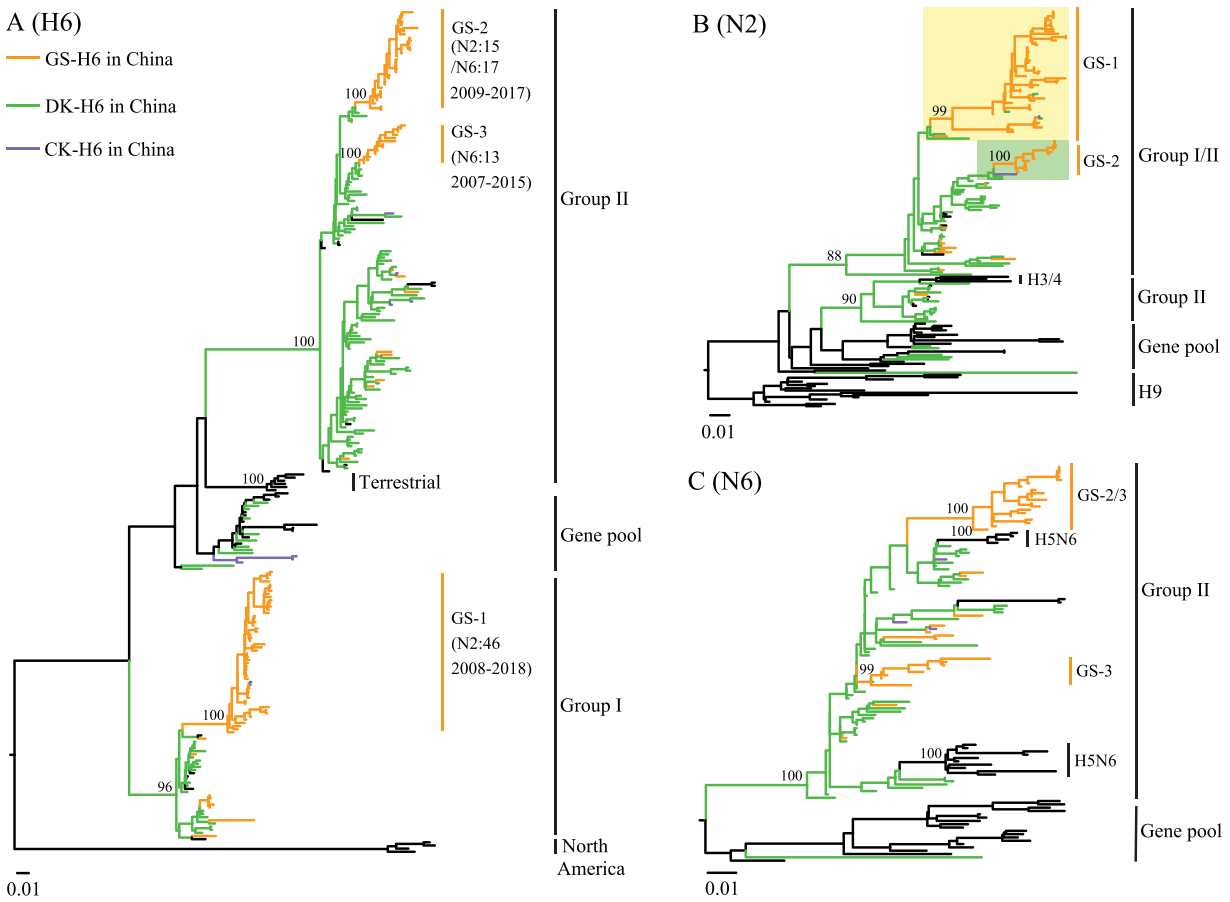

Figure 1. Maximum likelihood phylogenies of the surface protein genes. (a) H6, $n = 268$; (b) N2, $n = 178$; (c) N6, $n = 120$. Numbers above branches indicate bootstrap support values. Viruses isolated from minor poultry were clustered into the clade of terrestrial poultry. All H6 viruses detected in domestic ducks, geese, and chickens in China are colored according to the legend in (a). NA genes of GS H6 lineages introduced from duck group I or II viruses are highlighted in yellow or green boxes. Branch scale represents 0.01 substitutions per site.

ancestor (TMRCA) of the *HA* genes of the clades GS-1, GS-2, and GS-3 that were separately established in geese were estimated as May 2007 (95% HPD, June 2006–March 2008), April 2008 (95% HPD, October 2007–April 2009), and September 2007 (95% HPD, June 2007–October 2007), respectively, which were later than those of the TMRCAs of duck H6 Group I and II viruses (Table 2, Supplementary Fig. S3). The dating of clade GS-1 suggested a 9-year gap (possible range 7–12 years) from the TMRCA of the most closely related duck group I H6 viruses (Supplementary Fig. S3). The direct precursors of clade GS-1 in ducks might not have been detected even though surveillance was conducted continuously in this period (Huang et al. 2010, 2012).

(2) A low evolutionary rate and strong purifying selection were detected in the goose H6 HA gene

H6 *HA* gene sequences from viruses of the natural reservoir, group I and group II lineages, including those from parallel surveillance since 2000 in domestic ducks that were close relatives of the goose H6 lineages (Supplementary Fig. S3), were used to estimate the evolutionary dynamics involving the duck-to-goose transmissions.

The MCMC analysis showed that the mean nucleotide substitution rate of wild bird-like H6 *HA* genes of 3.16×10^{-3} (95% HPD: $2.36 \times 10^{-3} - 4.01 \times 10^{-3}$) rose to 3.76 (Group I, 95% HPD: $2.64 \times 10^{-3} - 4.89 \times 10^{-3}$) and 5.70×10^{-3} (Group II, 95% HPD: $4.98 \times 10^{-3} - 6.43 \times 10^{-3}$) (nucleotide substitutions/site/year) after they became established in ducks (Table 2).

Introductions of H6 viruses from ducks to geese did not increase the nucleotide substitution rate (Table 2, mean rate). An increase in evolutionary rate could be observed on the branches leading to the most recent common ancestor of each goose clade, but the nucleotide substitution rates became lower for clades GS-2 and GS-3 as these H6 viruses persisted in geese (Table 2, Episodic rate, Mean rate). For GS-1 viruses, the evolutionary rate of 4.0×10^{-3} (95% HPD: $3.25 \times 10^{-3} - 4.78 \times 10^{-3}$) nucleotide substitutions/site/year was similar to that of the duck Group I viruses. GS-2 and GS-3 viruses had somewhat lower rates (4.69×10^{-3}, 95% HPD: 3.76×10^{-3}-5.76×10^{-3} and 5.17×10^{-3}, 95% HPD: 3.73×10^{-3}-6.72×10^{-3} nucleotide substitutions/site/year, respectively) than the Group II duck viruses.

To detect variation in selection pressure following viral transmission, different ω (dN/dS) ratios were measured for specific branches or lineages in the *HA* gene phylogeny. Overall, H6 viruses were under purifying selection in ducks and geese (Table 2, Supplementary Table S2, ω), like viruses circulating in natural reservoir hosts (Gorman et al. 1992, Suarez 2000, Spackman et al. 2005). However, the introduction of viruses from the natural reservoir to ducks did increase the dN/dS value for Group-II (Table 2-ω_{Clade}, $\omega_{Wild\ bird\text{-}like} = 0.09$, $\omega_{DK\text{-}Group\ II} = 0.19$). When the duck-to-goose virus transmissions occurred, rises in dN/dS ratios from 0.09 to 0.2 for GS-1 and 0.19 to 0.41/0.97 for GS-2/3 were found in the early

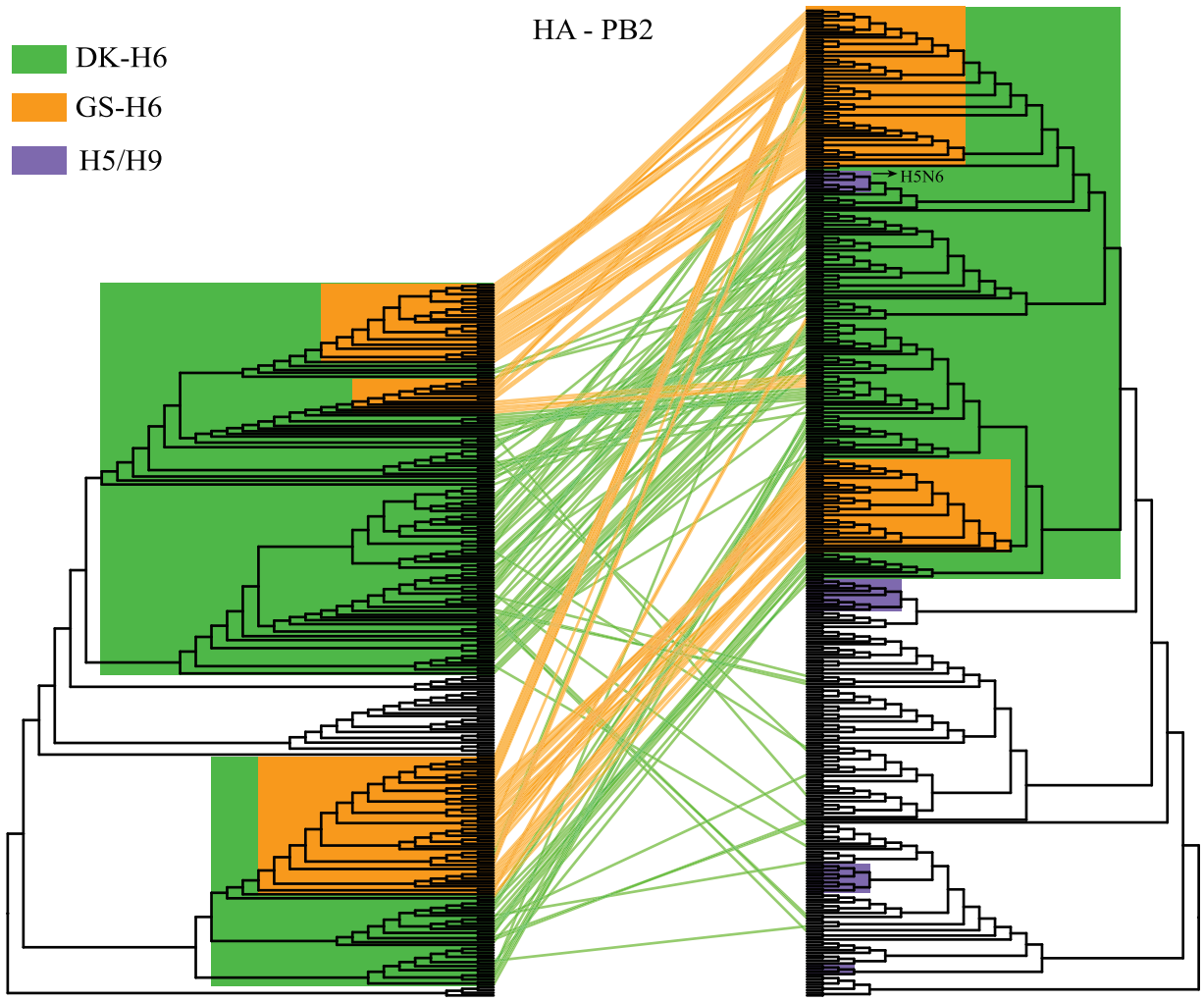

Figure 2. Reassortment patterns of H6 viruses established in ducks and geese. Tanglegrams were built based on ML phylogenies represented without branch-length data. The ML *HA* tree is on the left and the mirrored tree of the *PB2* gene is on the right. Evolutionary positions of the duck H6 and goose viruses are linked by lines in the paired phylogenies. *PB2* genes from H5 and H9 viruses established in terrestrial poultry are indicated by a purple background. Viruses from the natural reservoir were not colored. The exchange between H5 and H6 viruses is highlighted by an arrow. Abbreviations: DK, duck; GS, goose.

Table 2. Evolutionary rates and selection pressures of the *HA* gene.

		Substitution rate (95% HPD) (10^{-3} subs/site/year)		dN/dS (ω)	
Clade	TMRCA[a] (95% HPD)[b]	Episodic[c]	Mean[d]	ω_{branch}[e]	ω_{clade}[f]
Wild bird-like	/	/	3.16 (2.36, 4.01)	/	0.10 (0.06, 0.14)
DK-Group I	Jun. 1995 (Dec. 1992–Jan. 1997)	/	3.76 (2.64, 4.89)	/	0.09 (0.06, 0.12)
GS-1	May 2007 (Jun. 2006–Mar. 2008)	4.50 (3.20,6.00)	4.00 (3.25, 4.78)	0.20 (0.13, 0.30)	0.19 (0.14, 0.24)
DK-Group II	Mar. 2000 (Nov. 1998–Mar. 2001)	/	5.70 (4.98, 6.43)	/	0.19 (0.16, 0.23)
GS-2	Apr. 2008 (Oct. 2007–Apr. 2009)	5.60 (3.00,8.80)	4.69 (3.76, 5.76)	0.41 (0.21, 0.70)	0.11 (0.05, 0.17)
GS-3	Sep. 2007 (Jun. 2007–Oct. 2007)	7.50 (2.60,13.70)	5.17 (3.73, 6.72)	0.97 (0.35, 2.09)	0.16 (0.06, 0.26)

[a]TMRCA, time of the most recent common ancestor of the listed clades.
[b]95% HPD, 95% highest posterior density interval.
[c,e]Nucleotide substitution rates and selection pressures of branch leading to the interspecies clades. (Also indicated by an arrow in Supplementary Figure S3).
[d,f]Mean nucleotide substitution rate and selection pressure of the specific clade.
Slash, not analyzed.

Table 3. Positively selected sites detected in H6 clades.

	Sites detected by selection methods[a]			
Clade	**FEL**	**SLAC**	**FUBAR**	**Positively selected sites[b]**
DK-Group I	None	None	128,173,214,493	0
GS-1	None	None	142,155,158,193	0
DK-Group II	75,**157,160**, **169,186,372**	**157,160**, **186,372**	**157,160,169,186**, 199,367,**372**,487	5
GS-2	**226**	None	92,137,**226**,280	1
GS-3	None	None	172	0

[a]Amino acid sites with statistically significant selection values ($P \leq .05$ for FEL, SLAC, and posterior probability $\geq .9$ for FUBAR) are shown in H3 numbering.
[b]Number of positively selected sites confirmed by more than one method. These sites are highlighted in bold.

stage (Table 2, ω_{branch}), but these ratios dropped to values similar to those in the duck lineages after persistence of the viruses in geese for the Group-II derived viruses (Table 2-ω_{Clade}, $\omega_{GS\text{-}1} = 0.19$, $\omega_{GS\text{-}2} = 0.11$, $\omega_{GS\text{-}3} = 0.16$).

(3) More positively selected sites accumulated in duck viruses before the cross-species transmission of H6 viruses to geese

To identify mutations that might relate to host switching, ancestral sequences at internal nodes of the H6 *HA* phylogeny were inferred. More nonsynonymous mutations were identified on the branches leading to the DK-Group I and II lineages than on the branches leading to the GS-clades (Supplementary Table S3). Most of the mutations that appeared in the domestic duck lineages did not change during the duck-to-goose transmission (Supplementary Table S3). Under the MEME method, two amino acid sites during the formation of the duck lineages were found to be under positive selection, and one site was positively selected in the duck-to-goose (GS-2) host change branch (Supplementary Table S3).

Amino acid sites that might show positive selection during virus circulation in an established lineage or new host were examined by three methods. More positively selected sites were inferred in the duck viruses than in the goose viruses (Table 3), with five positively selected amino acid residues in one duck lineage and one in one goose lineage.

3.4 Gene reassortment was limited in domestic geese

To depict the reassortment pattern of established H6 viruses, tanglegrams, or linked pairs of phylogenies, were constructed by mirroring the internal gene tree against the *HA* phylogeny. Lines that link the same viruses in two phylogenies can visually indicate whether viral reassortment between different clades of viruses occurred.

The association of the *HA* gene from established duck H6 viruses with the internal gene phylogeny for *PB2*, represented by the green lines in Fig. 2, showed that they frequently exchanged genes with multiple subtypes of viruses from the natural reservoir. Reassortment of duck group I/II viruses with H5/H9 viruses of lineages enzootic in terrestrial poultry was also detected in the *PB1*, *M* and *NS* phylogenies (Supplementary Fig. S4, green lines). A duck H6 virus contributed the *PB2* gene to the genesis of highly pathogenic H5N6 viruses of clade 2.3.4.4b.

Goose H6 viruses showed a restricted reassortment pattern (Fig. 2, Supplementary Fig. S4, orange lines), reassorting with other goose viruses or, less commonly, the duck H6 viruses. They did not reassort with natural reservoir viruses, except for rare cases involving the *PA* and *NP* genes (Supplementary Fig. S4). Gene flow between goose H6 and H5/H9 viruses in terrestrial poultry was not detected during our surveillance period.

3.5 Molecular characteristics of goose H6 viruses

Sequence analysis showed that all goose H6 viruses were lowly pathogenic viruses based on having a single basic amino acid in the *HA1/HA2* cleavage site (Fig. 3a). Most goose H6 viruses contained Q226 (79/109) and G228 (107/109) (H3 numbering is used throughout this study) at the receptor binding sites (RBD) that normally indicates avian-type sialic acid-binding preference (Spackman et al. 2005, Liu et al. 2009). Most GS-1 viruses harbored the E190V (43/46) (H3 numbering) mutation, together with Q226 and G228, which indicate higher affinity for avian-type receptors (Bi et al. 2020). The Q226L mutation was found in the *HA* genes of GS-2 viruses (30/32). Whether this mutation would increase the affinity of these viruses for the human-type receptor (α2-6-SA) warrants further investigation.

In the *NA* stalk region, three deletion patterns (Δ66–75, Δ69–78, and Δ63–81) were identified in the goose H6N2 viruses. These had been detected in their duck precursor viruses since 2003 (Supplementary Table S4, Fig. 3b) (Huang et al. 2012) and might be associated with the optimization of gene compatibility (Banks et al. 2001; Cheung et al. 2007). Mammalian adaptation markers like E627K and D701N in the PB2 protein (Zhu et al. 2015) were not detected in any of the goose H6 viruses surveyed in this study.

4. Discussion

As aquatic birds, geese are usually considered part of the natural reservoir of influenza A viruses. Our study systematically investigated the genesis and evolutionary behavior of H6 viruses circulating in domestic geese and raises the possibility that the position of domestic geese in the influenza A ecosystem might be more like that of an aberrant host.

In China, only limited subtypes (H1, H3, H4, H5, H6, H7, H9, and H10) of influenza A viruses have been identified in domestic geese. Our long-term accumulated surveillance data first demonstrated that H6 viruses have been able to persist in this species (Table 1). Phylogenetic analysis showed that only established duck H6 viruses had been transmitted to domestic geese and that these viruses subsequently became established in geese (Fig. 1). Natural reservoir origin H6 viruses, from diverse geographical origins and many species of wild birds, have frequently been detected in domestic ducks (Webster et al. 1992, Olsen et al. 2006), but no natural reservoir virus was found in a domestic goose (Fig. 1). Of the H6 viruses that have circulated in domestic geese for a long time, none are known to have transmitted back to domestic ducks. The ability of domestic geese to transmit influenza viruses to other species might be limited, as might also be the case with wild geese

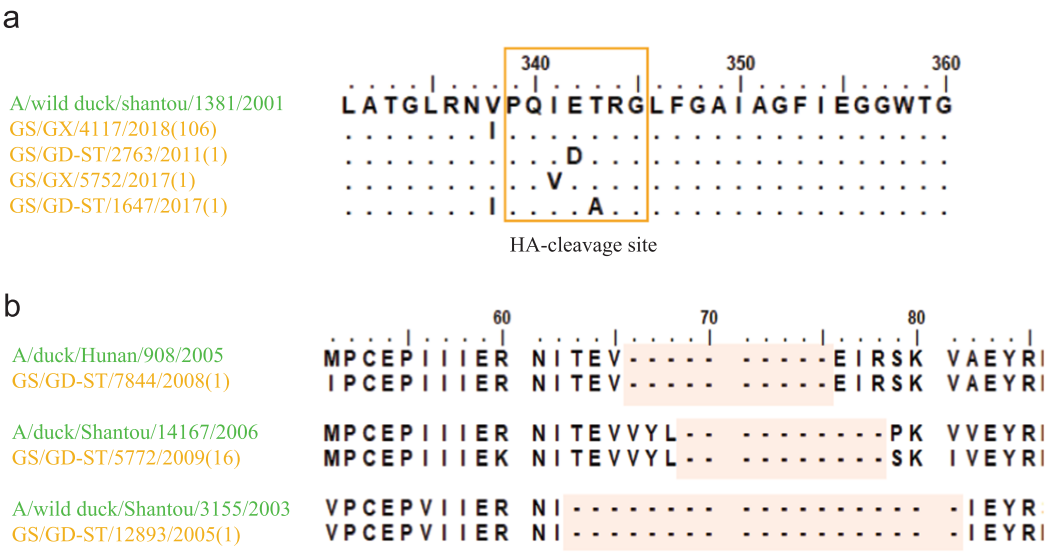

Figure 3. Molecular signatures. (a) The *HA* cleavage sites of the H6 *HA* genes. (b) Amino acid deletion patterns of the N2 *NA* gene in the stalk region. Example sequences of duck and goose H6 viruses are displayed together. The numbers of geese H6 viruses isolated in this study that had these deletion patterns are indicated in parenthesis.

(Jonassen and Handeland 2007, Hoye et al. 2010, Kleijn et al. 2010, Samuel et al. 2015, Yin et al. 2017).

Domestic geese were found to harbor enzootic influenza virus lineages with a stable gene constellation derived from lineages observed in domestic duck H6 viruses (Fig. 1, Supplementary Fig. S2) (Huang et al. 2010, 2012). This viral behavior has also been observed for an H6 virus in minor terrestrial poultry (Cheung et al. 2007). Although the group I domestic duck lineage was replaced by group II lineage viruses in domestic ducks in 2005 (Huang et al. 2012), viruses derived from both lineages have co-circulated in domestic geese up to the end of our surveillance (Fig. 1).

Even when the goose H6 viruses became established in domestic geese, transmission to terrestrial poultry (chickens) was rarely observed (Fig. 1 and Supplementary Fig. S1). As H6N2 viruses from the GS-1 and GS-2 clades do have molecular determinants related to adaption to terrestrial poultry in their surface genes (Cheung et al. 2007, Wang et al. 2015), it is possible that some other factor might restrict the ability of domestic geese to transmit viruses to terrestrial poultry.

Generally, when viruses were introduced to and became established in a poultry host, rapid evolution with an accelerated substitution rate and elevated nonsynonymous to synonymous mutation ratio would be observed (Gorman et al. 1992, Widjaja et al. 2004, Chen and Holmes 2006, Wille and Holmes 2020). This appeared to be the case with the persistence of the H6 *HA* in domestic ducks (Table 2) with host immunoselection or adaptation to the new host possibly driving this higher evolutionary rate. However, transmission and establishment of duck-origin H6 viruses in domestic geese did not show an apparent increase in evolutionary rate (Table 2). Although a rapid accumulation of nonsynonymous mutations was observed in the early phase of interspecies transmission (Table 2, ω_{branch}), most of these mutations were removed during subsequent circulation (Table 2, ω_{clade}). Possibly they were deleterious mutations for the persistence of H6 viruses in the domestic goose host.

More amino acid sites were positively selected during the circulation of H6 viruses in ducks than in geese (Table 3). Circulation of H6 viruses in ducks might have facilitated or "pre-adapted" the virus for transmission and persistence in geese. Our findings could indicate that domestic geese might restrict the evolutionary rate of AIVs, unlike terrestrial poultry (Banks et al. 2001, Shi et al. 2017).

Gene reassortment appeared to be restricted in these domestic goose lineages, which is inconsistent with domestic geese being part of the natural reservoir for influenza A viruses where frequent gene exchange occurs among viruses of different hosts and subtypes. Gene interactions between the established H6 lineage viruses and natural reservoir or enzootic H5/H9 subtype viruses occurred frequently in ducks (Fig. 2, green lines). In geese, this was seldom seen with viruses from wild waterfowl, and reassortment with enzootic H5Nx and H9N2 viruses was not detected in our long-term surveillance (Fig. 2, Supplementary Fig. S4). Domestic geese and ducks are the two major hosts of H5Nx and H6Nx viruses in China (Guan et al. 2002, Huang et al. 2010) and reassortments between H5 and H6 subtype viruses were often observed in domestic ducks, leading to the genesis of the highly pathogenic H5N6 virus (Bi et al. 2016), but this was not seen in domestic goose viruses.

Our study raises the question of whether the GS/GD lineage of viruses, first observed in geese, was indeed generated within geese. Geese acquired only limited types of duck H6 viruses and failed to transmit those viruses to terrestrial poultry in our long-term surveillance. It might be that the domestic goose is effectively a "dead-end" host for influenza A viruses, at least the H6 subtype. We suggest that domestic geese were likely not involved in the genesis of any highly pathogenic H5Nx viruses.

Our extensive long-term surveillance and evolutionary analyses of H6 viruses in southern China indicate that domestic geese received only established duck viruses and maintained the diversity and molecular markers from those viruses. However, domestic geese did not serve as an intermediate host to link viruses from waterfowl to land-based poultry. Domestic geese did not facilitate rapid evolution of influenza viruses and showed restricted gene reassortment with other viruses. It is possible that other reassortant virus forms occurred and were not detected during the surveillance, despite our surveillance being systematic and covering a period of 18 years and that the many other groups also conducting surveillance of AIVs have not reported these forms.

From our findings, it is possible that domestic geese may be an aberrant and "dead-end" host for influenza A viruses, rather than a natural reservoir host. This suggested that the restriction of the role of domestic geese in the influenza ecosystem is more extensive than the apparently restricted dissemination of AIVs by migratory wild geese (Jonassen and Handeland 2007, Hoye et al. 2010, Kleijn et al. 2010, Samuel et al. 2015, Yin et al. 2017). However, potential quarantining effects within the agricultural system need to be more closely evaluated. To confirm whether domestic geese do indeed have a limited role in the influenza ecosystem requires more investigation of other influenza virus subtypes and this should be addressed.

5. Conclusions

Our 18-year surveillance in southern China demonstrated that H6 subtype influenza A viruses were frequently detected in domestic geese. In tracing their origin, and phylogenetic relationships, and examining their evolutionary behavior in domestic geese, we found that domestic geese received a limited range of duck viruses and did not transmit viruses from natural reservoir hosts to terrestrial poultry or generate variants by gene exchange between viruses seen in other species in the influenza ecosystem. These findings raise the possibility that the role of domestic geese in the influenza ecosystem might be more like an aberrant and "dead-end" host. As domestic geese appear to have a limited role in the influenza A ecosystem when compared to other poultry, the option of reducing viral surveillance in this host could be considered.

Acknowledgements

We thank all staff from Joint Laboratory for International Collaboration in Virology and Emerging Infectious Diseases/Joint Institute of Virology (STU/HKU), the Department of Microbiology from Guangxi Medical University, and SKLEID for their technical assistance.

Author contributions

Y.G. and L.P. conceived the study, L.P., X.F., W.H., Z.Z., and Y.L. performed virus isolation and sequencing, X.F., W.H., Z.Z., P.C., and J.W. conducted the surveillance, L.P., Z.J., W.Y.-M.C., T.T.-Y.L., H.Z., and Y.G. contributed to the analysis, and L.P., Z.J., T.T.-Y.L., D.K.S., and Y.G. wrote the manuscript.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Conflict of interest: None declared.

Funding

This study was supported by the National Institutes of Health (National Institute of Allergy and Infectious Diseases contract number 75N93021C00016), Li Ka Shing Foundation, Shenzhen-Hong Kong Science and Technology Cooperation Zone-Shenzhen program (HZQB-KCZYZ-2021014), the AIR@InnoHK fund from the Innovation and Technology Commission of Hong Kong government, and the Guangdong government (for the support programs 190824215544727, 2019B121205009, 190830095586328, and 200109155890863).

Data availability

Genome sequences obtained in this study were submitted to GenBank, under the accession numbers from PP499655 to PP500526.

References

Alvarez-Carretero S, Kapli P, Yang Z. Beginner's guide on the use of PAML to detect positive selection. *Mol Biol Evol* 2023;**40**:msad041.

Bahl J, Vijaykrishna D, Holmes EC *et al*. Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts. *Virology* 2009;**390**:289–97.

Banks J, Speidel ES, Moore E *et al*. Changes in the haemagglutinin and the neuraminidase genes prior to the emergence of highly pathogenic H7N1 avian influenza viruses in Italy. *Arch Virol* 2001;**146**:963–73.

Bi Y, Chen Q, Wang Q *et al*. Genesis, evolution and prevalence of H5N6 avian influenza viruses in China. *Cell Host Microbe* 2016;**20**:810–21.

Bi Y, Li J, Li S *et al*. Dominant subtype switch in avian influenza viruses during 2016-2019 in China. *Nat Commun* 2020;**11**:5909.

Chen P, Jin Z, Peng L *et al*. Characterization of an emergent chicken H3N8 Influenza Virus in Southern China: a Potential Threat to Public Health. *J Virol* 2023;**97**:e0043423.

Chen R, Holmes EC. Avian influenza virus exhibits rapid evolutionary dynamics. *Mol Biol Evol* 2006;**23**:2336–41.

Cheung CL, Vijaykrishna D, Smith GJD *et al*. Establishment of influenza A virus (H6N1) in minor poultry species in southern China. *J Virol* 2007;**81**:10402–12.

Cheung PP, Leung YHC, Chow C-K *et al*. Identifying the species-origin of faecal droppings used for avian influenza virus surveillance in wild-birds. *J Clin Virol* 2009;**46**:90–93.

Clark K, Karsch-Mizrachi I, Lipman DJ *et al*. GenBank. *Nucleic Acids Res* 2016;**44**:D67–72.

Gorman OT, Bean WJ, Webster RG. Evolutionary processes in influenza viruses: divergence, rapid evolution, and stasis. *Curr Top Microbiol Immunol* 1992;**176**:75–97.

Green AJ, Elmberg J. Ecosystem services provided by waterbirds. *Biol Rev Camb Philos Soc* 2014;**89**:105–22.

Guan Y, Peiris M, Kong KF *et al*. H5N1 influenza viruses isolated from geese in Southeastern China: evidence for genetic reassortment and interspecies transmission to ducks. *Virology* 2002;**292**:16–23.

Guangchuang Y *et al*. ggtree: an rpackage for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Meth Ecol Evolut* 2017;**8**:28–36.

Guo YJ, Krauss S, Senne DA *et al*. Characterization of the pathogenicity of members of the newly established H9N2 influenza virus lineages in Asia. *Virology* 2000;**267**:279–88.

Hoffmann E, Stech J, Guan Y *et al*. Universal primer set for the full-length amplification of all influenza A viruses. *Arch Virol* 2001;**146**:2275–89.

Hoye BJ *et al*. Reconstructing an annual cycle of interaction: natural infection and antibody dynamics to avian influenza along a migratory flyway. *Oikos* 2010;**120**:748–55.

Hu C, Li X, Zhu C *et al*. Co-circulation of multiple reassortant H6 subtype avian influenza viruses in wild birds in eastern China, 2016-2017′. *Virol J* 2020;**17**:62.

Huang K, Bahl J, Fan XH *et al*. Establishment of an H6N2 influenza virus lineage in domestic ducks in southern China. *J Virol* 2010;**84**:6978–86.

Huang K, Zhu H, Fan X *et al*. Establishment and lineage replacement of H6 influenza viruses in domestic ducks in southern China. *J Virol* 2012;**86**:6075–83.

Jonassen CM, Handeland K. Avian influenza virus screening in wild waterfowl in Norway, 2005′. *Avian Dis* 2007;**51**:425–28.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;**30**:772–80.

Kleijn D, Munster VJ, Ebbinge BS *et al*. Dynamics and ecological consequences of avian influenza virus infection in greater white-fronted geese in their winter staging areas. *Proc Biol Sci* 2010;**277**:2041–48.

Kosakovsky Pond SL, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* 2005;**22**:1208–22.

Lam TT *et al*. The genesis and source of the H7N9 influenza viruses causing human infections in China. *Nature* 2013;**502**:241–44.

Li J *et al*. Continued reassortment of avian H6 influenza viruses from Southern China, 2014-2016. *Transbound Emerg Dis* 2019;**66**:592–8.

Liu J, Stevens DJ, Haire LF *et al*. Structures of receptor complexes formed by hemagglutinins from the Asian Influenza pandemic of 1957′. *Proc Natl Acad Sci USA* 2009;**106**:17175–80.

Luo S, Xie Z, Xie Z *et al*. Surveillance of live poultry markets for low pathogenic avian Influenza viruses in Guangxi Province, Southern China, from 2012-2015′. *Sci Rep* 2017;**7**:17577.

Ma C *et al*. Emergence and evolution of H10 subtype influenza viruses in poultry in China. *J Virol* 2015;**89**:3534–41.

Murrell B, Moola S, Mabona A *et al*. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* 2013;**30**:1196–205.

Murrell B, Wertheim JO, Moola S *et al*. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* 2012;**8**:e1002764.

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**:268–74.

Olsen B, Munster VJ, Wallensten A *et al*. Global patterns of influenza a virus in wild birds. *Science* 2006;**312**:384–88.

Peng C, Sun H, Li J *et al*. Molecular epidemiological survey and complete genomic phylogenetic analysis of H6 subtype avian influenza viruses in poultry in China from 2011 to 2016′. *Infect Genet Evol* 2018;**65**:91–95.

Pond, Kosakovsky SL, Frost SDW. Estimating selection pressures on alignments of coding sequences Analyses using HyPhy. 2007. https://www.hyphy.org/resources/hyphybook2007.pdf.

Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;**5**:e9490.

Qi W, Jia W, Liu D *et al*. Emergence and adaptation of a novel highly pathogenic H7N9 Influenza virus in birds and humans from a 2013 human-infecting low-pathogenic ancestor. *J Virol* 2018;**92**:10–128.

Samuel MD, Hall JS, Brown JD *et al*. The dynamics of avian influenza in Lesser Snow Geese: implications for annual and migratory infection patterns. *Ecol Appl* 2015;**25**:1851–59.

Shi J *et al*. H7N9 virulent mutants detected in chickens in China pose an increased threat to humans. *Cell Res* 2017;**27**:1409–21.

Spackman E, Stallknecht DE, Slemons RD *et al*. Phylogenetic analyses of type A influenza genes in natural reservoir species in North America reveals genetic variation. *Virus Res* 2005;**114**:89–100.

Suarez DL. Evolution of avian influenza viruses. *Vet Microbiol* 2000;**74**:15–27.

Suchard MA, Lemey P, Baele G *et al*. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10′. *Virus Evol* 2018;**4**:vey016.

Swayne DE, Suarez DL. Highly pathogenic avian influenza. *Rev Sci Tech* 2000;**19**:463–82.

Wang F, Qi J, Bi Y *et al*. Adaptation of avian influenza A (H6N1) virus from avian to human receptor-binding preference. *EMBO J* 2015;**34**:1661–73.

Webby RJ, Woolcock PR, Krauss SL *et al*. Multiple genotypes of non-pathogenic H6N2 influenza viruses isolated from chickens in California. *Avian Dis* 2003;**47**:905–10.

Webster RG, Bean WJ, Gorman OT *et al*. Evolution and ecology of influenza A viruses. *Microbiol Rev* 1992;**56**:152–79.

Widjaja L, Krauss SL, Webby RJ *et al*. Matrix gene of influenza A viruses isolated from wild aquatic birds: ecology and emergence of influenza A viruses. *J Virol* 2004;**78**:8771–79.

Wille M, Holmes EC. The ecology and evolution of influenza viruses. *Cold Spring Harb Perspect Med* 2020;**10**:1–19.

Xu X, Chen Q, Tan M *et al*. Epidemiology, evolution, and biological characteristics of H6 avian influenza viruses in China. *Emerg Microbes Infect* 2023;**12**:2151380.

Xu X, Subbarao K, Cox NJ *et al*. Genetic characterization of the pathogenic influenza A/Goose/Guangdong/1/96 (H5N1) virus: similarity of its hemagglutinin gene to those of H5N1 viruses from the 1997 outbreaks in Hong Kong. *Virology* 1999;**261**:15–19.

Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 2007;**24**:1586–91.

Yin S, Kleijn D, Müskens GJDM *et al*. No evidence that migratory geese disperse avian influenza viruses from breeding to wintering ground. *PLoS One* 2017;**12**:e0177790.

Yoon SW, Webby RJ, Webster RG. Evolution and ecology of influenza A viruses. *Curr Top Microbiol Immunol* 2014;**385**:359–75.

Zhu W, Li L, Yan Z *et al*. Dual E627K and D701N mutations in the PB2 protein of A(H7N9) influenza virus increased its virulence in mammalian models. *Sci Rep* 2015;**5**:14170.

Virus Evolution, 2024, 10(1), veae075 , DOI: https://doi.org/10.1093/ve/veae075, Advance Access Publication 14 September 2024, Research Article