

Research Article

Phylogenetics and Molecular Mutations of the Hemagglutinin Affecting Global Transmission and Host Adaptation of H5Nx Viruses

Atanas V. Demirev,¹ Heedo Park,¹ Kyuyoung Lee ,¹ Sejik Park,¹ Joon-Yong Bae,¹ Man-Seong Park ^{1,2,3} and Jin Il Kim ^{1,2,3}

¹Department of Microbiology, Institute for Viral Diseases, College of Medicine, Korea University, Seoul, Republic of Korea

²Biosafety Center, College of Medicine, Korea University, Seoul, Republic of Korea

³Vaccine Innovation Center, College of Medicine, Korea University, Seoul, Republic of Korea

Correspondence should be addressed to Man-Seong Park; ms0392@korea.ac.kr and Jin Il Kim; jinil_kim@korea.ac.kr

Received 7 February 2023; Revised 22 March 2023; Accepted 29 March 2023; Published 14 April 2023

Academic Editor: Karima Al Salihi

Copyright © 2023 Atanas V. Demirev et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Highly pathogenic avian influenza (HPAI) H5 viruses have circulated globally causing incidental human infection with a substantial pandemic threat. The present study investigated the molecular evolution and phylogenetics of hemagglutinin (HA) in avian and human-isolated H5Nx viruses globally circulating since 2000. We investigated the dynamics of amino acid substitution in the HA sequences of avian and human H5Nx viruses and performed a phylogenetic analysis. Our study found that the H5Nx lineages dominantly expanded since 2000 and diverged into multiple sublineages with unique genetic mutations. P185S mutation in HA became a molecular characteristic of dominant H5Nx viruses throughout clades 2.3.4.1 to 2.3.4.4 (2.3.4.1–4). The key mutations, $\Delta E130$ and I155T, and potential N-linked glycosylation at residues 128, 144, and 159 in the HA gene of human-isolated viruses possibly contributed to both the individual and population levels of the H5 evolution and the host adaptation. Our analysis detected heterogeneity in amino acid sites under positive selection in the HA gene of clades 2.3.4.1–4. Accumulated mutations in the HA protein may potentially affect not only the genetic and antigenic diversity of HPAI H5Nx viruses but also increase the functional compatibility with NA subtypes. Given the global spread and incessantly occurring HA mutations of H5Nx viruses, our results emphasize the importance of early identification of HA mutations as well as the need for a comprehensive assessment of H5Nx variants in terms of pandemic preparedness.

1. Introduction

Highly pathogenic avian influenza (HPAI) viruses of the H5 subtype have caused continuous outbreaks in poultry, highlighting the continued risk of transmission in humans [1, 2]. Since the first report of fatal human infection of HPAI H5N1 in 1997 [3], the A/goose/Guangdong/1/96 (Gs/Gd) lineage has diverged into 10 distinct clades (0–9) [4, 5]. Interestingly, the Gs/Gd lineage generated descendant strains which after 2007 showed increasing preference to pair with a variety of neuraminidase (NA) subtypes except for N7 [6–8]. Initially, the HPAI of subtypes H5N2, H5N5,

H5N6, and H5N8 were globally observed and classified in phylogenetic clade 2.3.4.4, one of the major descendant lineages from the Gs/Gd lineage [9]. The clade 2.3.4.4 H5Nx viruses were extensively spread, and in particular, the HPAI H5N8 subtype has caused global outbreaks in poultry [10–13].

Although the molecular determinants for H5Nx viruses to cross interspecies barriers are still not completely understood, the impact of mutation in hemagglutinin (HA) as a key determinant of host adaptation was observed in previous studies [14]. Evolutionary dynamics of the HA protein regarding HA receptor binding specificity play

a central role in the influenza viruses (AIVs) switch from avian to mammalian hosts [15–18]. Specifically, non-synonymous mutations in the proximity of the HA receptor binding site (RBS) appear to enhance host adaptation and aerosol transmission of H5N1 viruses in mammals [19–21]. Naturally occurring mutations within the 130- and 220-loops of H5N1 HAs (such as S137A and K222Q) could increase the affinity for human-like α 2,6-linked sialic acid (SA) receptors [22, 23]. Residues increasing α 2,6 SA binding of the H5N1 HA and lysine at PB2 residue 627 (K627) presented a way to achieve airborne transmission between mammals [16]. It was also reported that human-isolated H5N1 and H5N6 viruses might recognize both avian-like α 2,3 and α 2,6 SA receptors but could not sustain respiratory-droplet transmission in ferrets [24]. In addition, the increasing antigenic diversity in HA and NA can also potentially broaden the host ranges of AIVs because immune pressure in different animal hosts can partially drive the selection and preservation of genetic mutations that arise in both surface glycoproteins [25, 26]. Identifying specific residues within HA and NA by reverse genetics has made it possible to study the effect of certain mutations on the interspecies transmission of H5Nx viruses [16, 17, 24, 27].

Hence, amino acid mutations in HA, especially around the RBS region, of H5Nx variants should be carefully monitored in terms of global public health. This study investigated the molecular and evolutionary dynamics of H5Nx HAs using the complete genomic sequences of the HA gene in recently expanded H5 (sub) clades. The phylodynamics were used to investigate the genealogical relationship of H5Nx HAs, and the selection pressure profile. The genetic diversity in the key antigenic sites and molecular changes in the HA crystal structure were assessed to investigate the evolutionary dynamics of H5Nx HAs on the molecular level and its potential effects on the viral characteristics of H5Nx viruses.

2. Materials and Methods

2.1. Data Collection and Phylogenetic Analyses. A total of 4,069 complete HA sequences of avian and human H5Nx viruses were collected from the database of the National Center for Biotechnology Information (Bethesda, MD, USA) ($n=2,690$) and the Global Initiative on Sharing Avian Influenza Data (GISAID) ($n=1,379$) from 1996 to July 2022. We manually inspected HA sequence data and initially selected 3,890 HA sequences of avian and human H5Nx viruses (H5N1 subtype, $n=1,733$; H5N2, $n=659$; H5N3, $n=90$; H5N4, $n=5$; H5N5, $n=53$; H5N6, $n=379$; H5N8 $n=947$; and H5N9, $n=24$) (Supplemental Table 1 and Supplemental Data S1). The random subsampling was performed only from HAs containing complete genomic information based on the subtype clades from 0 to 9 based on H5 subclades nomenclature [28], year of isolation (1996–2022), isolation region, and host using an in-house command-line script. Finally, our study used 739 HA sequences including 655 avian and human H5Nx HAs and 84 references HAs (Supplemental Table 2 and Supplemental Data S2).

Multiple alignment using fast Fourier transform (MAFFT) (v7.419) was used to align our HA sequences [29]. The phylogeny of 739 HA sequences was reconstructed by the time-framed Bayesian inference method implemented on Bayesian evolutionary analysis by sampling trees (BEAST) (v1.10.4) [30]. The general time-reversible + I + G nucleotide substitution [31] and lognormal uncorrelated relaxed clock [32] models were selected for the phylogenetic estimation. The Markov chain Monte Carlo (MCMC) chain length was 200 million with sampling every 200 thousand steps and the MCMC chains were assessed on Tracer (v1.7.1). The maximum clade credibility (MCC) tree was prepared using TreeAnnotator (v1.10.4) and visualized using FigTree (v1.4.4) (<https://tree.bio.ed.ac.uk>).

2.2. Genetic Diversity Analysis. We built three subsets based on the phylogenetic classification in the MCC tree of 739 HA sequences: the clade 2.3.2.1, the clade 2.3.4.4, and the combined clades 2.3.4.1 to 2.3.4.4 (2.3.4.1–4). The Bayesian Skygrid plots of four datasets (the complete Gs/Gd lineage, the clade 2.3.2.1, the clade 2.3.4.4, and the clades 2.3.4.1–4) were estimated to evaluate temporal dynamics of genetic diversity in the HA gene of H5Nx viruses over time. The coalescent-based Bayesian Skygrid model was used to estimate the lineage through time (LTT) plots reconstructed based on the inferred coalescent trees for the four datasets on Tracer [33]. The lineage-specific evolutionary rate and divergence times of nucleotide substitutions of the HA genes in four datasets were also estimated and summarized based on the posterior probability of sampling phylogenies.

2.3. Selection Pressure Profiles Analysis. Natural selection pressure at individual codon sites was estimated from the ratio of nonsynonymous (dN) vs. synonymous (dS) nucleotide substitutions per site (dN/dS, also known as ω) using online Datamonkey site [34]. The estimated ω (dN/dS) lower than 1 indicated an overall negative or purifying selective pressure and the ω higher than 1 signified that the amino acid position is under positive selection. The individual sites under either positive or negative selection pressure were detected using single-likelihood ancestor counting (SLAC) [35], fast unconstrained Bayesian approximation (FUBAR) [36], and mixed-effects model evolution (MEME) [37] models with the cutoffs of p value <0.1 (SLAC and MEME) and posterior probability >0.9 (FUBAR).

2.4. Frequency Analysis of Amino Acid Mutations. We used 3,890 HA sequences for the frequency analysis of amino acid mutations in H5Nx HAs. The amino acid substitution was investigated at the 13 key antigenic residues around the RBS in the HA gene (118, 130, 131, 133(a), 137, 155, 160, 185, 187, 189, 193, 196, and 222; H3 numbering). Amino acid mutations of H5Nx HAs were then analyzed based on the HA sequences of the A/Anhui/2/2005 strain located at the tip of clade 2.3.4.1 as a reference to all other HAs. We manually calculated the proportion of HA

sequences with each key amino acid substitution in the residue among the total 3,890 sequences by 6 time periods (before 2004, $n = 147$; between 2005 and 2008, $n = 533$; between 2009 and 2012, $n = 425$; between 2013 and 2016, $n = 1,082$; between 2017 and 2020, $n = 842$; and after 2021, $n = 860$). We additionally evaluated the amino acid substitution at the 13 key antigenic residues including two key amino acid substitutions, I155T and T160A, only in the human-isolated H5Nx sequences (before 2004, $n = 61$; between 2005 and 2008, $n = 249$; between 2009 and 2012, $n = 148$; between 2013 and 2016, $n = 67$; between 2017 and 2020, $n = 15$; and after 2021, $n = 9$) (Supplemental Table 3 and Supplemental Data S3).

2.5. Structural Analysis of the HA RBS Region. Our study reconstructed the globular head regions of the HA protein based on H5N1 (A/Vietnam/1194/2004, VN1194, clade 1.1; PDB ID: 4BGX) [18, 38] and H5N6 (A/Sichuan/26221/2014, SC26221, clade 2.3.4.4a; PDB ID: 5HU8) [39] viruses using PyMOL Molecular Graphics System v2.1.0 (Schrödinger, Inc., New York, NY, USA). The conserved amino acids at Y95, W153, H183, and Y195 lining the base of each HA RBS were exposed [40], and molecular interactions within 3.0 Å of specific amino acid residues around the HA RBS were investigated.

3. Results

3.1. Phylogenetic Relationships of H5Nx HAs. Our MCC phylogeny of the HA gene showed that global H5Nx viruses formed a polyphyletic topology with continuous generation and extinction of sublineages. Our study observed independent diversification of two major clades 2.3.2.1 and 2.3.4.4 following the recent H5 nomenclature [28], and each subclade heterogeneously involved multiple genotypes and human infection strains (Figure 1). The clade 2.3.4.4 diverged into multiple subclades from 2.3.4.4a to 2.3.4.4h. The subclade 2.3.4.4b involved the most recent prevalent strains with multiple H5Nx subtypes, globally spread, and formed the largest monophyletic subclade. We noted that subclades 2.3.4.4d, e, f, and h mostly involved H5N2 and H5N6 viruses that have evolved gradually in a single phylogenetic cluster. The subclade 2.3.4.4g involved the H5N1 and H5N6 viruses, although they appeared to share the same common ancestor with the subclades 2.3.4.4d, e, f, and h (tMRCA–2012.47; 10.47–11.49 HPD) (Table 1). Subclade 2.3.4.4a also involved H5N2 and H5N6 viruses, and subclade 2.3.4.4c held H5N1, H5N2, and H5N8 viruses. The H5N6 in the subclades 2.3.4.4d–h were limitedly identified in Southeast Asia, but the H5N8 viruses were globally detected from East Asia to Russia, the Middle East, Africa, and Europe (2.3.4.4b) as well as in North America (2.3.4.4c). The H5N1 subtype mostly belonged to clades 2.3.2.1 and 2.3.4.1–3, which were detected in East Asia, India, the Middle East, Russia, and Central Africa (Figure 1). Some H5N6 and H5N9 HAs were classified in clade 2.3.2.1 and H5N2 HAs in clade 2.3.4.2, but most H5Nx HAs other than the H5N1 HAs were included in clade 2.3.4.4 (Table 1).

3.2. Temporal Dynamics of the Genetic Diversity in HAs. Here, we generated lineage through time (LTT) plots and examined the gradual lineage diversification rate following a coalescent-based Bayesian model separately for clades 2.3.2.1, 2.3.4.4, and 2.3.4.1–4, and for the complete Gs/Gd lineage (Figure 2). The LTT plots of all four datasets generally described the expansion of genetic diversity in the HA gene of H5Nx viruses over time. The LTT plots of each clade showed a difference in the period of the increase in clade genetic diversity. The LTT plot for clade 2.3.2.1 showed two subsequent increases in lineage diversity from 2000 to 2003 and from 2006 to 2013. However, the LTT plot for clades 2.3.4.1–4 exhibited two drastic exponential increases from 2000 to 2006 and from 2012 to 2015 marked by the expansion of subclade 2.3.4.4b and the appearance of subclades 2.3.4.4c, 2.3.4.4d–h (Figure 2). The LTT plots for clade 2.3.4.4 specifically showed a single marked increase from 2012 to 2015. The estimated nucleotide substitution rates were similar around 4.66 to 4.80×10^{-3} substitution/site/year in all four datasets. The mean nucleotide substitution rate of the HA gene in the Gs/Gd lineage was higher than the substitution rates in clades 2.3.2.1, 2.3.4.1–4, and clade 2.3.4.4 (Table 1).

3.3. Key Amino Acid Mutations in H5Nx HAs. Comparing to the HA of A/Anhui/2/2005 as a reference strain, we observed amino acid mutations I118T, D128N, A131T, S133L, S137A, T160A, D187N, A189E, K193N, K222Q, and S227R around the RBS of clade 2.3.4.4 HAs (Table 2). Of these, only T160A, D187N, K193R, K222Q, and S227R were observed in H5N5 HA at the base of clade 2.3.4.4, whereas the mutations S133L, S137A, T160A, and K193R were also identified in clade 2.3.2.1 HAs. The A189E mutation was preserved in the HAs of subclades 2.3.4.4b, 2.3.4.4d, and 2.3.4.4h after 2016 (Table 2, Supplemental Table 3), albeit the 2.3.4.4b evolved independently (Figure 1). Mutations distant from the RBS pocket, such as N98S(T) and Q173R, were found specifically in subclade 2.3.4.4b (and including T199D in 2.3.4.4c). The T160A mutation was persistent in both clades 2.3.2.1 and 2.3.4.4 (Table 2). I155T was detected in subclades 2.3.4.4d–h (except 2.3.4.4g) mostly in H5N6 subtype viruses in East Asia (Figure 1, Table 2).

Among the mutations around the RBS of various H5 subtypes, the substitution ratios of six mutations (A131T, S133L, S137A, D187N, K193N, and R222Q) were with higher proportions in the HAs of the H5N6 and H5N8 subtypes (Figure 3(a), Supplemental Figure 1(a)) and appeared to continuously increase over the past decade (Figure 3(b), Supplemental Figure 1(b)). However, most of these mutations were not detected in the HAs of H5N2, H5N3, H5N5, and H5N9 subtypes before 2020 (Figure 3(a)). Specifically, RBS mutations, such as $\Delta 130$, $\Delta 133$, and I155T mutations, were not found in HAs of the H5N8 (Figure 3(a)) but were detected in 50–80% of the H5N6 and frequently detected in the HAs of human-isolated H5N6 of subclade 2.3.4.4d and 2.3.4.4h (Supplemental Table 3). We noted that the proportion of the A189E mutation increased after 2014 (Figure 3(b)) and was retained in the HAs of two

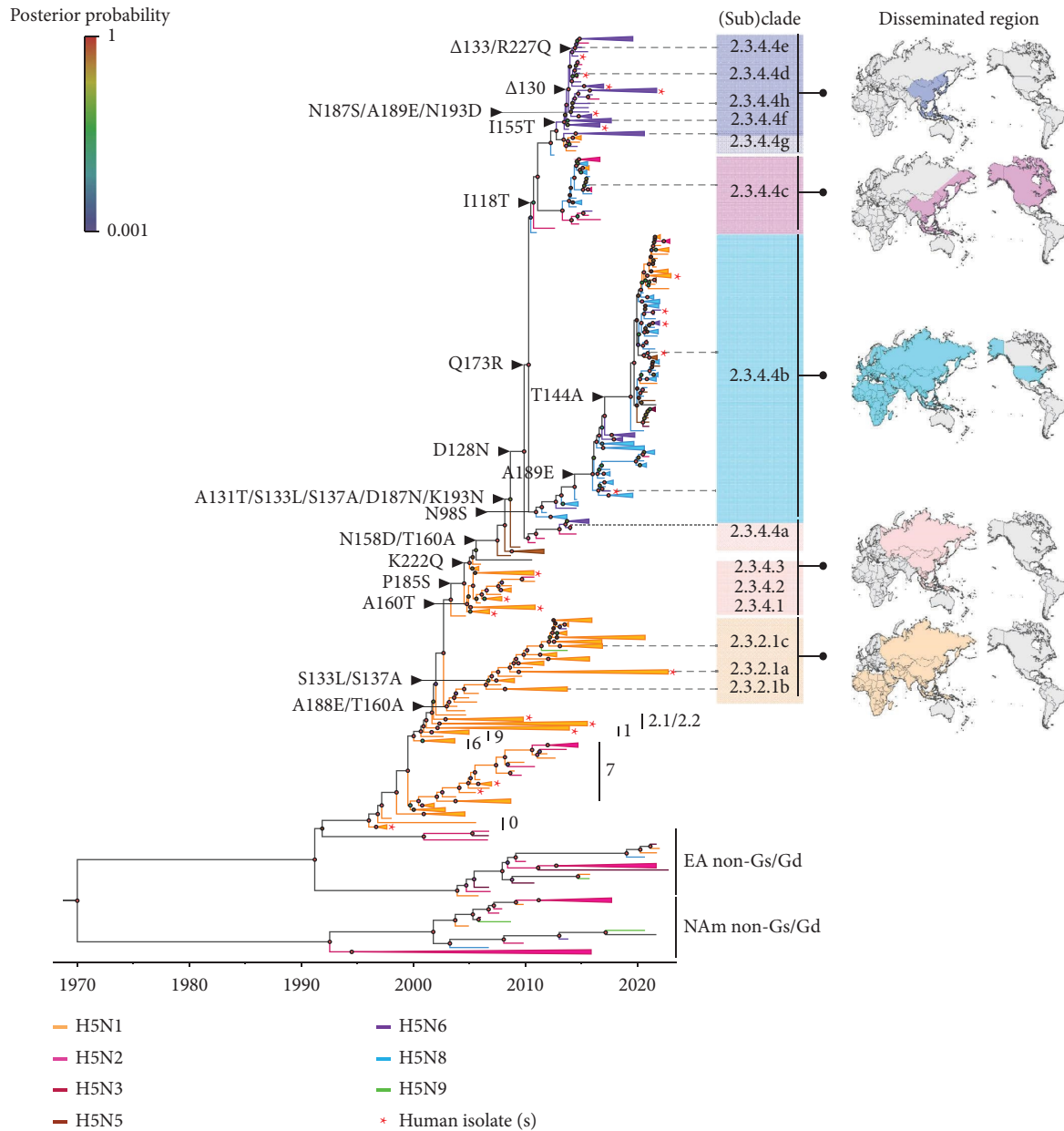


FIGURE 1: Phylogenetic relationships of the H5Nx HAs between 1996 and 2022. The MCC tree of the H5Nx HAs is reconstructed using the Bayesian evolutionary interference method. The phylogenetic branches with specific subtypes are color-labeled in orange (H5N1), strawberry (H5N2), wine red (H5N3), brown (H5N5), purple (H5N6), light blue (H5N8), and green (H5N9). The genetic clusters containing the same H5 subtype are collapsed to simplify the tree representation. The HAs of human-isolated H5Nx viruses are shown with red asterisks. The geographic regions that have confirmed isolated H5 viruses from clades 2.3.2.1, 2.3.4.1-3, and 2.3.4.4 are presented on the right.

independent subclades 2.3.4.4b and 2.3.4.4h after 2016 (Supplemental Table 3). Importantly, the subtle molecular changes appear to be dominant in the H5N6 and H5N8 HAs (Figure 3(a)). Deletions at residues 130 ($\Delta 130$) and 133 ($\Delta 133$) were detected mostly in the H5N6 subtype HAs (Figure 3(a), Table 2), where $\Delta 130$ was in increased proportions between 2013 and 2020, while $\Delta 133$ increased between 2017 and 2019 and then decreased (Figure 3(b)). Mostly observed in the H5N6 HAs, the I155T mutation also

showed an increased proportion between 2013 and 2016. We further noted that the proportions of I155T and T160A mutations in human-isolated H5Nx HAs exhibited a similar pattern from 2006 to 2016 (Supplemental Figure 2). However, the number of human H5Nx HAs with the T160A mutation increased after 2017 (Supplemental Figure 2), while I155T proportions were dramatically reduced, as this mutation was not observed in subclade 2.3.4.4b (Supplemental Table 3).

TABLE 1: Evolutionary rates and time of most recent common ancestor (tMRCA) estimates of H5Nx HAs by (sub)clade.

(Sub)clade	Main subtypes	Evolutionary rate (10^{-3} substitution/site/year)	tMRCA in year
Gs/Gd lineage	H5Nx	4.86 (4.49–5.19) ^a	1995.83 (1994.84–1996.60)
2.3.2.1	H5N1	4.70 (4.08–5.23)	2003.26 (1999.45–2002.99)
2.3.4.1-4	H5Nx	4.67 (4.27–5.18)	2002.58 (1998.25–2004.32)
2.3.4.4	H5Nx	4.66 (4.19–5.18)	2005.02 (2004.10–2005.98)
2.3.4.4a	H5N2 and H5N6	—	2009.77 (2009.49–2010.65)
2.3.4.4b	H5N1, H5N2, H5N3, H5N5, H5N6, and H5N8	—	2010.46 (2010.09–2011.43)
2.3.4.4c	H5N1, H5N2, and H5N8	—	2013.31 (2012.41–2013.57)
2.3.4.4d	H5N2 and H5N6	—	2014.15 (2013.53–2014.24)
2.3.4.4e	H5N2 and H5N6	—	2014.18 (2013.58–2014.32)
2.3.4.4f	H5N6	—	2013.78 (2013.19–2013.86)
2.3.4.4g	H5N1 and H5N6	—	2013.10 (2012.71–2013.70)
2.3.4.4h	H5N6	—	2014.07 (2013.47–2014.20)

^aThe lower and upper limits of 95% highest posterior density (HPD) are presented in parenthesis.

3.4. Genetic Polymorphism and Natural Selection Pressure in H5Nx HAs. Our study found that estimates of selection pressures in all four datasets ranged from 0.172 to 0.207, which were lower than 1 (Table 3, Supplemental Table 4). The Gs/Gd lineage showed the highest estimate of HA genes selective pressure ($\omega = 0.207$) and the clades 2.3.4.1-4 showed a slightly higher ratio ($\omega = 0.19$) than clades 2.3.2.1 ($\omega = 0.18$) and 2.3.4.4 ($\omega = 0.17$). Among the individual sites in the global Gs/Gd lineage analysis (Table 3), we detected four residues around the RBS under positive selection, 133, 159, 160, and 193 (SLAC) that belong to H5 HA antigenic regions A and B [41, 42]. In addition to these residues, the MEME model predicted episodic diversifying pressure at three additional residues (130, 132, and 142), but residue 160 was not involved in the episodic diversification of the global Gs/Gd lineage. FUBAR, on the other hand, detected two residues 144 and 159 within antigenic regions A and B, respectively. Variations of these two residues are specific in clades 2.3.2.1 and 2.3.4.4 (Supplemental Table 3). The FUBAR model further showed that residues 158, 159, and 193 were consistently selected across clade 2.3.2.1, while residues 133, 193, and 199 have continuously been selected in the clade 2.3.4.4 HAs. Among the amino acid residues under episodic selection (MEME), residues 155 and 222 were specifically detected in clade 2.3.4.4 but not in clade 2.3.2.1 (Table 3).

3.5. Molecular Structure and Interaction of Human-Isolated H5Nx HAs. Our study observed structural differences of the HA globular head region in the human A/Changsha/1/2014 (CS1) H5N6 strain in subclade 2.3.4.4h (Figure 3(c)), compared to two other reference HAs in clade 1.1 (human H5N1, VN1194) (Figure 3(d)) [18] and clade 2.3.4.4a (human H5N6, SC26221) (Figure 3(e)) [39]. The simulation of CS1 HA revealed a cluster of polar amino acid residues T131-S(T)132-S133 and T155 were mainly detected in the H5N6 subtype of subclades 2.3.4.4d, 2.3.4.4e, and 2.3.4.4h (Figure 3(c), Supplemental Table 3). The combination of such residues was not observed in the HAs of human H5N6 in subclades 2.3.4.4a (Figure 3(e)) or 2.3.4.4b (Supplemental Table 3), where more hydrophobic side chains were

prevalent (A131/S132/L133 and I155) (Figures 3(d) and 3(e)). We found that the hydrogen bonds (H-bonds) between I155 and E130 in the VN1194 and SC26221 HAs shifted to an analogous interaction between amino acid residues T155 and T131 in CS1 HA. Interestingly, putative physical interaction between residues 155 and 131 was also detected for other simulated HAs of human H5Nx viruses within both clades 2.3.2.1 and 2.3.4.4 isolated after 2017 (Supplemental Figure 3). Furthermore, we found that the P185S mutation might alternatively interact with residues Y95, H183, and E190 residues within the RBS pocket of the simulated CS1 HA (Figure 3(c)). Within the RBS of SC26221 HA, however, in the presence of L133 and I155, the S185 amino acid physically interacted with the Q191 sidechain (Figure 3(e)). CS1 HA simulation (Figure 3(a)) also showed that $\Delta 130$ in CS1 generates a potential N-linked glycosylation (NLG) site at N128 (Table 2, Figure 3(c)), which is in proximity to the previously well-known NLG site at N158 in the HA of VN1194 (Figure 3(d)).

4. Discussion

Our study explored the phylodynamics of influenza H5Nx HAs, detected key determinants on the molecular evolutionary level, and tried to unravel the impact of reassortment with the NA genes with the potential of the emergence of novel zoonotic strains. While fitness difference among AIVs continuously affects the dominance of variant selection through the rapid extinction of variants with low fitness, a variant with beneficial antigenic mutations effectively becomes a dominant strain and generates new sublineages. The HA has been considered a key gene segment that contributes to the fitness of variants due to the impact of its mutation on hosts' partial immunity and transmissibility [43]. Our approach based on the HA gene sequences found that the H5Nx viruses from the Gs/Gd lineage dominantly expanded and diverged into multiple sublineages with unique adaptive genetic mutations in the HAs through their global spread.

Since the HA protein is a constant target for neutralizing antibodies, here, we demonstrated that the HA RBS mutations of H5N6 and H5N8 subtypes would play a key role in

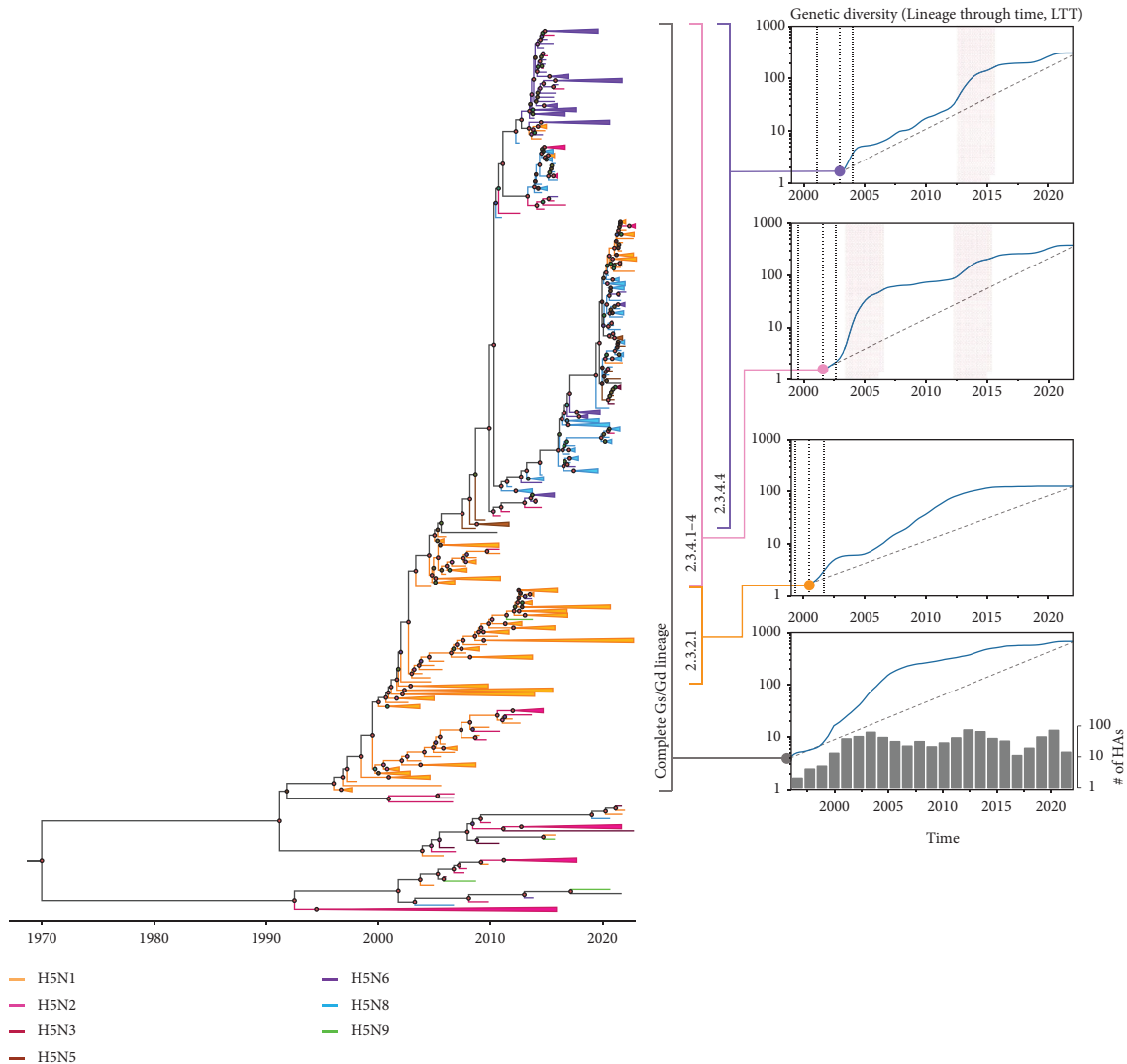


FIGURE 2: Bayesian coalescent LTT plots of H5Nx HAs by clade over time. The LTT plots display the temporal patterns of lineage diversification through time based on the phylogeny inferred by the Bayesian coalescent approach. The solid lines and blue shading indicate the mean and the 95% confidence interval of the effective population size in the log scale at the designated time interval (years). The dashed line under each curve represents the temporal change of the effective population size with an estimated constant rate of diversification without extinction within the lineages. The vertical, bold-dashed lines represent the divergence time of each clade estimated by the molecular clock model. The pink shade explains the period with a dramatic increase in the effective population size. The bar plot under the curve in the complete Gs/Gd lineage explains the number of the HAs used for the phylogenetic analysis.

forming the antigenicity of clade 2.3.4.4. As previous research reported, clades 2.3.2.1 and 2.3.4.4 were distinctly dominant among H5Nx viruses [10, 12, 13, 44]. Our phylogenetic analysis of the HA gene found that clade 2.3.4.4 would genetically diverge into five major groups; a, b, c, g, and others (d, e, f, and h) [28] (Figure 1, Table 1). However, the genetic classification may not follow the antigenicity change of H5Nx viruses [41]. We observed that the amino acid residues in the HA RBS of subclades 2.3.4.4b (H5N8, and more recently variety of H5N1, H5N2, H5N3, H5N5, and H5N6 subtypes) and 2.3.4.4c (H5N2 and H5N8) did not show the significant antigenic difference (Table 2). Despite the low variations, only H5N8 viruses of subclade 2.3.4.4b spread over 51 countries throughout Asia, Europe, the Middle East, Africa, [11] and recently North America by

2019 [45] (Figure 1). We observed the accumulation of mutations in key antigenic sites around the HA RBS in the H5N8 viruses (Figure 3(a)). Prominent molecular substitutions were A131T, S133L, S137A, D187N, A189E, K193N, and R222Q, and especially most H5Nx viruses in the clade 2.3.4.4 identified after 2020 held these mutations (Figure 3(a), Table 2).

The P185S mutation was first detected in clade 2.3.4.1 and became molecular characteristics of the currently dominant HAs of H5Nx viruses throughout clades 2.3.4.1–4 (Figure 1). Yet with largely unknown function in the presence of L133 and I155, the S185 amino acid residue might physically interact with Q191, thus generating a pull toward the 190-helix (Figure 3(e)). The HA RBS structure in subclade 2.3.4.4h H5N6 (CS1), however, revealed that the

TABLE 2: Amino acid mutations in the HA globular head region of H5Nx viruses by (sub)clade.

Representative strains	Subtypes	(Sub)clade	Amino acid signature at a given residue (H3 numbering)																			
			HA head					HA head					190-helix					220-loop				
			98	118	128	130	131	133	137	144	155	160	173	187	189	193	199	222	227			
A/Anhui/2/2005	H5N1	2.3.4.1	N	I	D	E	A	A	S	S	T	I	T	Q	D	A	K	T	K	S		
A/Nepal/19FL1997/2019	H5N1	2.3.2.1a					L	A	N ^a	A		A				R						
A/duck/Bangladesh/17D1821/2022	H5N1	2.3.2.1a			N		L	A	N ^a	A		A				R						
A/duck/Laos/NL-2072062/2020	H5N1	2.3.2.1c					L	A	N ^a	A		A	R			R						
A/goose/Guangdong/k0103/2010	H5N5	2.3.4.4								A						R						
A/duck/Zhejiang/6DK19/2013	H5N2	2.3.4.4a					T	L	A			A			N					R		
A/duck/Vietnam/NCDV15A57/2015	H5N6	2.3.4.4a					T	L	A			A			N					R		
A/duck/Shandong/Q1/2013	H5N8	2.3.4.4b	S				T	L	A			A	R	N		N				R		
A/common teal/Korea/W547/2016	H5N8	2.3.4.4b	S		N		T	L	A			A	R	N	E	N				R		
A/Turkey/Norway/FU496/2020	H5N8	2.3.4.4b	S		N		T	L	A	A		A	R	N	E	N				R		
A/Human/10117/2021	H5N6	2.3.4.4b	S		N		T	L	A	A		A	R	N	E	N				R		
A/England/215201407/2021	H5N1	2.3.4.4b	S		N		T	L	A	A		A	R	N	E	N				R		
A/mallard/Korea/W452/2014	H5N8	2.3.4.4c	T	T	N		T	L	A	A		A	R	N		N	D			R		
A/chicken/Iowa/04-20/2015	H5N2	2.3.4.4c	T	T	N		T	L	A	A		A	R	N		N	D			R		
A/chicken/Guangdong/GD1602/2016	H5N6	2.3.4.4d	T	N ^a	Δ ^b		T		A	M	T	A	R	N	E	N				Q		
A/Hubei/29578/2016	H5N6	2.3.4.4d	T	N ^a	Δ ^b		T		A	M	T	A	R	N	E	N				Q		
A/chicken/Hokkaido/002/2017	H5N6	2.3.4.4e	T	N			T	Δ	A	V	T	A	G	N		N				Q		
A/wild duck/South Korea/1920/2019	H5N6	2.3.4.4e	T	N			T	Δ	A	V	T	A	G	N		N				Q		
A/duck/Vietnam/LBM638/2014	H5N1	2.3.4.4g	T				T	L	A	M		T	R	N	D					R		
A/duck/Vietnam/HU13-65/2019	H5N6	2.3.4.4g	T	N			T	L	A	M		A	S	N		N				R		
A/chicken/Vietnam/BD1113/2017	H5N6	2.3.4.4f	T	N			T		A	V	T	A	G	N		N				Q		
A/Guangdong/18SF020/2018	H5N6	2.3.4.4h	T	N ^a	Δ ^b		T		A	V	T	A	K	S	E	D				R		
A/Chongqing/00013/2021	H5N6	2.3.4.4h	T	N ^a	Δ ^b		T		A	V	T	A	K	S	E	D				R		

^aPotential N-linked glycosylation site. ^bΔ, deletion of amino acid.

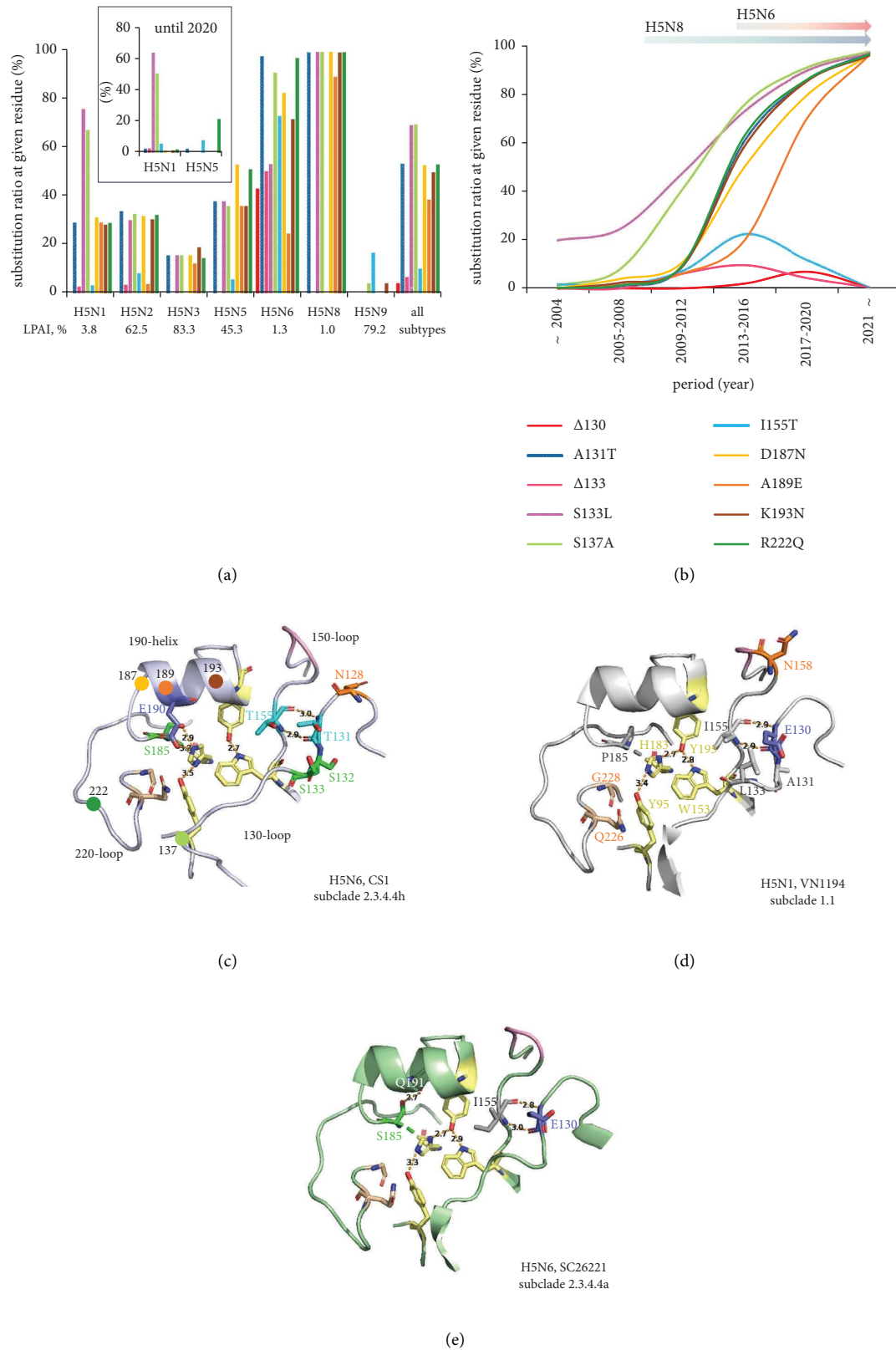


FIGURE 3: The temporal dynamics and molecular interactions of amino acid substitutions around the HA globular head region of H5Nx viruses. The proportion of amino acid substitutions around the HA globular head region of H5Nx viruses is calculated by (a) the number of H5Nx subtypes or by (b) the 6 time periods: (a) ~2004 ($n = 147$), (b) 2005–2008 ($n = 533$), (c) 2009–2012 ($n = 425$), (d) 2013–2016 ($n = 1082$), (e) 2017–2020 ($n = 842$), and (f) 2021~ ($n = 860$). Using the HA structure of VN1194, molecular interactions in the HA globular head region of (c) human-isolated H5N6, A/Changsha/1/2014 (CS1, subclade 2.3.4.4h), (d) human H5N1 HA (VN1194, subclade 1.1), and (e) human H5N6 HA of A/Sichuan/26221/2014 (SC26221, subclade 2.3.4.4a) are analyzed. The amino acid residues of Y95, W153, H183, and Y195 (pale yellow) constitute the conserved floor in the H5 HA RBS. The hydrophobic residues of I155, A131, and L133 are colored gray, the oxygen atoms are red, and the nitrogen atoms are blue. The potential NLG sites at N128 and N158 are colored orange and residues T131 and T155 are colored cyan. S132, S133, and S185 side chains are colored green.

TABLE 3: Selection pressure profiles in H5Nx HAs by clade.

Clades	The number of sequences	dN/ dS	Amino acid residue under positive selection		
			SLAC	FUBAR	MEME
Gs/Gd lineage	697	0.207	40, 133, 159, 160, 173, 192, 193, 199, 481	144, 159	40, 114, 124, 130, 132, 133, 142, 159, 193, 296, 297, 324, 481
2.3.2.1	129	0.182	158, 193	158, 159, 193	130, 132, 140, 145, 158, 159, 173, 193
2.3.4.1-4	358	0.19	119, 133, 173, 192, 199, 326	119, 192, 193, 199, 326, 537	40, 119, 127, 130, 133, 142, 199, 222, 324, 326, 537
2.3.4.4	306	0.172	173, 199	133, 193, 199	40, 130, 133, 142, 155, 173, 199, 222, 273

amino group of H183 also likely form H-bonds with the -OH groups of S185 and E190, causing a conformational shift and averting its amino acid residues from the usual physical connection between Y95, H183, and Y195 side chains (the base of the RBS dictate the receptor-binding mode [40, 46] (Figures 3(d) and 3(e)). This plasticity and physical interactions among Y95, H183, S185, and E190 residues might generate a pull toward the 220-loop in the HA of the H5N6 subtype (Figure 3(c)), therefore, are readily accessible for the SA receptors. An increased polarity in amino acid residues T131, S132, S133, T155, and S185 and a shorter 130-loop (Δ E130) would also be a reason for a more flexible RBS pocket in the human H5N6 HA (Figure 3(c)). Thus, the P185S mutation is a unique and significant determinant in the evolution of H5Nx HAs forming novel RBS plasticity.

H5Nx viruses with deletions in the 130-loop, such as Δ E130 or Δ L/S133, in combination with I155T, might have a potential for zoonotic infections because these were detected in the HAs of H5N1 viruses instigating sporadic human infections between 2007 and 2014 [27, 47]. High proportions of deletion in 130-loop and I155T are rarely observed in subtypes other than H5N6 in the past decade (Figure 3(b)), and Δ E130 and I155T in the HAs were notable in the avian and human-isolated H5N6 viruses of subclades 2.3.4.4d and 2.3.4.4h (Table 2, Supplemental Table 3). The proportions of strains with these two substitutions increased since 2013 and then declined after 2017, which coincided with the period of subclades 2.3.4.4d–h expansion. Another deletion Δ L133 was also reported to enhance the α 2,6 SA receptor recognition [21, 47] and perhaps mimics the shorter 130-loop in H2/H3 HAs [48]. Therefore, the dynamics of H5Nx avian influenza viruses in wild migratory birds with these molecular features should be carefully monitored to prevent the risk of the emergence of zoonotic influenza [49].

Introduction of potential NLG at N128 and/or N158 was detected in the HAs of many human-isolated H5N1 and H5N6 viruses (Supplemental Table 3), whereas NLG at N144 was a characteristic of H5N1 HAs in clade 2.3.2.1 (Table 2). NLG is an important molecular modification due to its impact on the modulation of glycoprotein stability and virus pathogenicity [50, 51]. Since amino acid residues 128 and 158 are structurally close to each other (Figures 3(c) and 3(d)), potential NLGs at N128 and N158 might be mutually exclusive. Sun et al. demonstrated that, although NLG at N128 is positioned away from the RBS, it may improve viral fitness by stabilizing the HA trimeric structure [51]. The impact of the lack of NLG at position N158 around the HA globular head on the pathogenicity of H5 AIVs in mammals remains controversial. Yen et al. [52] and Zhao et al. [53] found that NLG at N158 can increase virulence in mice but does not affect receptor binding preferences. However, Zang et al. [54] and Yin et al. [55] reported that a lack of NLG at N158 enhances virulence in mice. Recently, Gao et al. also showed that NLG at N158 increases H5N6 pathogenicity by inducing higher levels of inflammatory factors in mice [50]. These observations suggest that a functional NLG site with a shorter 130-loop might be an important mechanism for adaptation and increased pathogenicity in mammalian hosts.

Natural selection pressure formed the evolutionary characteristics of H5N1 HAs in clade 2.3.2.1 and H5Nx HAs in clade 2.3.4.4. Both clades shared common residues under selection pressure. Specifically, residues 130, 173, and 193 were consistently diversified in both clades. However, each clade showed its unique positive selection profiles at a molecular level. Amino acid residues 158, 159, and 193 within the proximal antigenic site B [41, 42] were under consistent positive selection in the H5N1 HAs of clade 2.3.2.1. In contrast, the clade 2.3.4.4 HAs uniquely showed positive selection in residues 133, 193, and 199, and an additional genetic variation at both individual host and population levels, including residues 155 and 222 (Table 3). Furthermore, the changes in the selective pressure also did not follow key changes in antigenic properties. The A189E mutation was observed dominantly in subclades 2.3.4.4b as well as 2.3.4.4d and 2.3.4.4h since 2017 (Table 2). Advantageous mutations, such as S133L and S137A, were simultaneously preserved in the clades 2.3.2.1 and 2.3.4.4 HAs. These residues were not recognized in the HA selection profile analysis (Table 3).

The lineage diversity plot for clades 2.3.4.1–4 did not follow the logistic-like growth curve (observed for clades 2.3.2.1 and 2.3.4.4), indicating that genetic shift has led to the continuous replacement of H5N1 with more adapted H5Nx subtypes (Figure 2). Notably, the increased diversity between 2000 and 2006 coincides with the transition from the H5N1 (in clades 2.3.4.1–3) to the H5Nx subtype (clade 2.3.4.4). The second stepwise increase after 2012 marks the global spread and dominance of H5Nx HAs in the LTT plot of clade 2.3.4.4 HAs and the expansion and increasing diversity in subclades 2.3.4.4b, 2.3.4.4c, and 2.3.4.4d–h (Figure 2). The stepwise lineage diversity we observed for clades 2.3.4.1–4 often ensues adaptations that enable a breach in the ecological barrier [56] and provides evidence that the expansion of one genetic group is gained at the expense of another [57]. We also noted a small increase in the genetic diversity after 2019 was characteristic for clades 2.3.4.1–4 and 2.3.4.4 alone, marked by the current expansion and diversity in subclade 2.3.4.4b, which might result in the emergence of novel subclades. Altogether, the lineage diversity in clade 2.3.4.1–4 affected the phylodynamics of the H5Nx HAs in two consequent exponential phases ensuing viral adaptations, one between 2004 and 2006 and another between 2012 and 2014.

Our investigation of the lineage-accumulated genetic changes in H5Nx HAs over time showed modulations in the H5 HA lineage diversity with NA subtype preferences. The present study found the H5N6 HAs belonged to subclades 2.3.4.4a, 2.3.4.4b, and 2.3.4.4d–h, and gene reassortment and consequent antigenic drift by each subclade possibly formed the high heterogeneity in the HAs [10, 13, 58]. Particularly, the I118T and I155T mutations were not observed in the H5N6 HAs of subclades 2.3.4.4a and 2.3.4.4b but persisted in the HAs of subclades 2.3.4.4d, 2.3.4.4e, 2.3.4.4f, and 2.3.4.4h. Furthermore, the persistence and stepwise genetic diversity increase of H5N1 HAs in subclade 2.3.2.1c in the past decade is likely related to the genetic reassortment within their genetic group (H5N6 and H5N9 detected) (Figure 1). It is still unclear how the maintained balance between HA and

NA in the H5N1 subtype suddenly shifted to other stable combinations with various NA subtypes, resulting in the emergence and global spread of H5Nx [59, 60]. An NLG camouflage in major HA epitope regions and a shorter 130-loop might contribute to the adaptation and evasion of H5Nx viruses from the host immune response [51]. The shorter 130-loop might also shift the sialic acid's sitting position and strengthen the epistatic interactions between the host receptor and HA RBS [48]. As shown previously, a combination of shortened 130-loop and polar amino acid residues T131, S132, and T155 (Figure 3(c)) in the H5N1 or H5N6 HAs could enhance α 2,6 SA receptor recognition [47]. Interestingly, these were not observed in the recently reported HAs of human-isolated H5N1 and H5N6 viruses in subclade 2.3.4.4b (Supplemental Table 3), which also suggests an improved genetic and functional balance with the NAs of N1 and N6 subtypes.

The human mortality rate with H5Nx viruses is higher than that with other influenza subtypes, including seasonal influenza viruses [61]. The evolving H5Nx viruses are becoming increasingly virulent in mammalian models [1, 62, 63]. Specifically, the HAs of the H5N1 and H5N6 subtypes isolated after 2020 in the subclades 2.3.4.4b and 2.3.4.4h showed an affinity for both α 2,3 and α 2,6 SA receptors and high replication capacity in mice [1, 64]. Although human-to-human transmission of the H5Nx viruses was very limitedly reported, these H5Nx viruses would have a potential for host adaptation in humans through the accumulation of genetic mutations in the HA key antigenic sites and its functional interaction with the NA as H5N1 and H5N6 viruses showed. Due to the low herd immunity of humans against such novel H5Nx variants, global public health authorities should carefully monitor molecular changes in novel H5Nx viruses. Though our study solely focused on the genetic characteristics of the H5Nx HAs, our findings would be still meaningful because the molecular characteristics of H5Nx HAs in the present study will provide valuable insight for better surveillance to precisely monitor key genetic mutations and their potential for zoonotic influenza H5Nx emergence through both high transmissibility and host adaptation.

5. Conclusions

The present study highlighted unique evolutionary characteristics of the HAs of currently circulating influenza H5Nx viruses. The global spread of the H5Nx viruses promoted the expansion of their genetic diversity, and diverse viral variants continuously gained adaptive advantages by harboring accumulated mutations in HAs and by the genetic shift. Our study suggests key molecular mutations and evolutionary dynamics of the HAs of HPAI H5Nx viruses. We believe this study will facilitate step-changes in our understanding of molecular virology as well as the evolution of H5Nx viruses for the better molecular monitoring of novel influenza variants.

Data Availability

All the data in this study are included within the article and supplementary information files. The HA genomic data of H5Nx viruses are obtained from the database of Influenza

Virus Resource (<https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=database>) and the EpiFlu database of Global Initiative on Sharing Avian Influenza Data (GISAID; <https://gisaid.org>).

Disclosure

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Man-Seong Park and Jin Il Kim conceptualized the study. Atanas V. Demirev, Heedo Park, Man-Seong Park, and Jin Il Kim performed data curation. Atanas V. Demirev, Heedo Park, Kyuyoung Lee, and Sejik Park performed formal analysis. Man-Seong Park and Jin Il Kim performed funding acquisition. Atanas V. Demirev, Heedo Park, Joon-Yong Bae, Kyuyoung Lee, and Sejik Park performed investigation. Man-Seong Park and Jin Il Kim supervised the study. Kyuyoung Lee, Man-Seong Park, and Jin Il Kim performed validation. Atanas V. Demirev and Heedo Park wrote the original draft. Atanas V. Demirev, Heedo Park, Kyuyoung Lee, Man-Seong Park, and Jin Il Kim wrote, reviewed, and edited the manuscript. Atanas V. Demirev and Heedo Park contributed equally to this work.

Acknowledgments

The authors would like to thank all the lab members of the Department of Microbiology and the Institute for Viral Diseases, Korea University, College of Medicine, who supported this study. The authors also thank all the authors of the genomic sequences of influenza viruses registered in the NCBI and GISAID databases that the authors used in this study. This study was supported by the grant of Korea University College of Medicine (K1804341 to J. I. K.), by the grants of the Bio and Medical Technology Development Program of the National Research Foundation (NRF, funded by the Ministry of Science and ICT of Republic of Korea; -NRF-2021M3E5E3081341 to J. I. K.), by the grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI, funded by the Ministry of Health & Welfare, Republic of Korea; No. HV20C0091 to M.-S. P.), and by the grant of Korea Basic Science Institute (National Research Facilities and Equipment Center, funded by the Ministry of Education of Republic of Korea; No. 2021R1A6C101C570 to M.-S. P.).

Supplementary Materials

Supplemental Table 1: the number of the H5Nx HA sequences by subtype collected. Supplemental Table 2: the number of the H5Nx HA sequences by subtype and isolation year used for the phylogenetic analysis. Supplemental

Table 3: amino acid mutations in the HA globular head region of reference human-isolated H5Nx viruses by clade. Supplemental Table 4: summary of natural selection pressure profiles of the H5Nx HAs by clade. Supplemental Data 1: initial sequence set of H5Nx HAs. Supplemental Data 2: selected sequence set of H5Nx HAs. Supplemental Data 3: sequence set of human-isolated H5Nx HAs. Supplemental Figure 1: the proportion of amino acid mutations around the HA globular head region of clade 2.3.4.4 H5Nx viruses by subtype and collection year. (a) The proportion of amino acid mutations around the HA globular head region of clade 2.3.4.4 H5Nx viruses is presented by each subtype (Supplemental Table 1). (b) The H5Nx HA sequences are divided into six periods; (a) 1997–2004 ($n = 147$), (b) 2005–2008 ($n = 533$), (c) 2009–2012 ($n = 425$), (d) 2013–2016 ($n = 1082$), (e) 2017–2020 ($n = 842$) and 2021–2022 ($n = 860$). Supplemental Figure 2: the proportion of the I155T and T160A mutations and molecular interactions of the HA globular head region of H5Nx viruses. (a) The proportion of the I155T (blue) and T160A (magenta) mutations in avian (dashed lines) and human-isolated H5Nx viruses (solid lines) is presented by each period (years). (b) Using the HA structure of VN1194, the receptor-binding pocket in H5 HA contains a conserved floor of residues Y95, W153, H183, Y195, and E190 (pale yellow), and Q226 and G228 (orange) that interact with $\alpha_2,3$ SA receptors. Residues 155 and 160 are colored blue and magenta. Supplemental Figure 3: structural analysis of the HA globular head region of human-isolated H5Nx viruses. Using the HA structure of VN1194, molecular interactions of the HA globular head region residues are estimated; (a) A/Nepal/19FL1997/2019 (H5N1) (NP19FL1197) in subclade 2.3.2.1; (b) A/England/215201407/2021 (H5N1) (END215201407) in subclade 2.3.4.4b; (c) A/Changsha/1/2014 (H5N6) (CS1) in subclades 2.3.4.4e; and (d) A/Guangdong/18SF020/2018 (H5N6) in subclade 2.3.4.4e. (*Supplementary Materials*)

References

- [1] P. Cui, J. Shi, C. Wang et al., “Global dissemination of H5N1 influenza viruses bearing the clade 2.3.4.4b HA gene and biologic analysis of the ones detected in China,” *Emerging Microbes and Infections*, vol. 11, pp. 1693–1704, 2022.
- [2] J. King, C. Schulze, A. Engelhardt et al., “Novel HPAIV H5N8 reassortant (clade 2.3.4.4b) detected in Germany,” *Viruses*, vol. 12, no. 3, p. 281, 2020.
- [3] J. C. de Jong, E. C. Claas, A. D. Osterhaus, R. G. Webster, and W. L. Lim, “A pandemic warning?” *Nature*, vol. 389, no. 6651, p. 554, 1997.
- [4] D. A. Boltz, P. Seiler, E. A. Govorkova et al., R. Mondry, Emergence of H5N1 avian influenza viruses with reduced sensitivity to neuraminidase inhibitors and novel reassortants in Lao People’s Democratic Republic,” *Journal of General Virology*, vol. 91, no. 4, pp. 949–959, 2010.
- [5] Who/Oie/Fao H5N1 Evolution Working Group, “Toward a unified nomenclature system for highly pathogenic avian influenza virus (H5N1),” *Emerging Infectious Diseases*, vol. 14, no. 7, p. e1, 2008.
- [6] Y. J. Lee, H. M. Kang, E. K. Lee et al., “Novel reassortant influenza A(H5N8) viruses, South Korea, 2014,” *Emerging Infectious Diseases*, vol. 20, no. 6, pp. 1086–1089, 2014.
- [7] C. G. Liu, M. Liu, F. Liu et al., “Emerging multiple reassortant H5N5 avian influenza viruses in ducks, China, 2008,” *Veterinary Microbiology*, vol. 167, no. 3–4, pp. 296–306, 2013.
- [8] K. Zhao, M. Gu, L. Zhong et al., “Characterization of three H5N5 and one H5N8 highly pathogenic avian influenza viruses in China,” *Veterinary Microbiology*, vol. 163, no. 3–4, pp. 351–357, 2013.
- [9] E. de Vries, H. Guo, M. Dai, P. J. Rottier, F. J. van Kuppeveld, and C. A. de Haan, “Rapid emergence of highly pathogenic avian influenza subtypes from a subtype H5N1 hemagglutinin variant,” *Emerging Infectious Diseases*, vol. 21, no. 5, pp. 842–846, 2015.
- [10] D. H. Lee, J. Bahl, M. K. Torchetti et al., “Highly pathogenic avian influenza viruses and generation of novel reassortants, United States, 2014–2015,” *Emerging Infectious Diseases*, vol. 22, no. 7, pp. 1283–1285, 2016.
- [11] D. H. Lee, K. Bertran, J. H. Kwon, and D. E. Swayne, “Evolution, global spread, and pathogenicity of highly pathogenic avian influenza H5Nx clade 2.3.4.4,” *Journal of Veterinary Science*, vol. 18, no. S1, pp. 269–280, 2017.
- [12] L. Sims, I. Brown, N. Gaidet et al., *Highly pathogenic H5 avian influenza in 2016 and 2017 – observations and future perspectives FOCUS ON*, vol. 11, 2017.
- [13] L. Yang, W. Zhu, X. Li et al., “Genesis and dissemination of highly pathogenic H5N6 avian influenza viruses,” *Journal of Virology*, vol. 91, no. 5, Article ID e02199-16, 2017.
- [14] K. Subbarao, “The critical interspecies transmission barrier at the Animal(-)Human interface,” *Tropical Medicine and Infectious Disease*, vol. 4, no. 2, p. 72, 2019.
- [15] L. M. Chen, O. Blixt, J. Stevens et al., “In vitro evolution of H5N1 avian influenza virus toward human-type receptor specificity,” *Virology*, vol. 422, no. 1, pp. 105–113, 2012.
- [16] S. Herfst, E. J. Schrauwen, M. Linster et al., “Airborne transmission of influenza A/H5N1 virus between ferrets,” *Science*, vol. 336, no. 6088, pp. 1534–1541, 2012.
- [17] M. Imai, T. Watanabe, M. Hatta et al., “Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets,” *Nature*, vol. 486, no. 7403, pp. 420–428, 2012.
- [18] X. Xiong, P. J. Coombs, S. R. Martin et al., “Receptor binding by a ferret-transmissible H5 avian influenza virus,” *Nature*, vol. 497, no. 7449, pp. 392–396, 2013.
- [19] M. Gu, Q. Li, R. Gao et al., “The T160A hemagglutinin substitution affects not only receptor binding property but also transmissibility of H5N1 clade 2.3.4 avian influenza virus in Guinea pigs,” *Veterinary Research*, vol. 48, no. 1, p. 7, 2017.
- [20] W. Wang, B. Lu, H. Zhou 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,” *Journal of Virology*, vol. 84, no. 13, pp. 6570–6577, 2010.
- [21] S. Yamada, Y. Suzuki, T. Suzuki et al., “Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors,” *Nature*, vol. 444, no. 7117, pp. 378–382, 2006.
- [22] H. Guo, E. de Vries, R. McBride et al., “Highly pathogenic influenza A(H5Nx) viruses with altered H5 receptor-binding specificity,” *Emerging Infectious Diseases*, vol. 23, no. 2, pp. 220–231, 2017.
- [23] Z. Y. Yang, C. J. Wei, W. P. Kong et al., “Immunization by avian H5 influenza hemagglutinin mutants with altered receptor binding specificity,” *Science*, vol. 317, no. 5839, pp. 825–828, 2007.

- [24] S. Herfst, C. K. P. Mok, J. M. A. van den Brand et al., "Human clade 2.3.4.4 A/H5N6 influenza virus lacks mammalian adaptation markers and does not transmit via the airborne route between ferrets," *mSphere*, vol. 3, no. 1, Article ID e00405-17, 2018.
- [25] R. Chen and E. C. Holmes, "Avian influenza virus exhibits rapid evolutionary dynamics," *Molecular Biology and Evolution*, vol. 23, no. 12, pp. 2336–2341, 2006.
- [26] R. G. Webster, W. J. Bean, O. T. Gorman, T. M. Chambers, and Y. Kawaoka, "Evolution and ecology of influenza A viruses," *Microbiological Reviews*, vol. 56, no. 1, pp. 152–179, 1992.
- [27] X. Xiong, H. Xiao, S. R. Martin et al., "Enhanced human receptor binding by H5 haemagglutinins," *Virology*, vol. 456–457, pp. 179–187, 2014.
- [28] Who, "Zoonotic influenza: candidate vaccine viruses and potency testing reagents - development of candidate vaccine viruses for pandemic preparedness," 2022, <https://www.who.int/teams/global-influenza-programme/vaccines/who-recommendations%20-%20last>.
- [29] K. Katoh and D. M. Standley, "MAFFT multiple sequence alignment software version 7: improvements in performance and usability," *Molecular Biology and Evolution*, vol. 30, no. 4, pp. 772–780, 2013.
- [30] M. A. Suchard, P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, and A. Rambaut, "Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10," *Virus Evolution*, vol. 4, no. 1, Article ID vey016, 2018.
- [31] D. Posada and K. A. Crandall, "Selecting the best-fit model of nucleotide substitution," *Systematic Biology*, vol. 50, no. 4, pp. 580–601, 2001.
- [32] A. J. Drummond, S. Y. Ho, M. J. Phillips, and A. Rambaut, "Relaxed phylogenetics and dating with confidence," *PLoS Biology*, vol. 4, no. 5, p. e88, 2006.
- [33] V. Hill and G. Baele, "Bayesian estimation of past population dynamics in BEAST 1.10 using the Skygrid coalescent model," *Molecular Biology and Evolution*, vol. 36, no. 11, pp. 2620–2628, 2019.
- [34] W. Delpont, A. F. Poon, S. D. Frost, and S. L. Kosakovsky Pond, "Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology," *Bioinformatics*, vol. 26, no. 19, pp. 2455–2457, 2010.
- [35] S. L. Pond and S. D. Frost, "Datamonkey: rapid detection of selective pressure on individual sites of codon alignments," *Bioinformatics*, vol. 21, no. 10, pp. 2531–2533, 2005.
- [36] B. Murrell, S. Moola, A. Mabona et al., "FUBAR: a fast, unconstrained bayesian approximation for inferring selection," *Molecular Biology and Evolution*, vol. 30, no. 5, pp. 1196–1205, 2013.
- [37] B. Murrell, J. O. Wertheim, S. Moola, T. Weighill, K. Scheffler, and S. L. Kosakovsky Pond, "Detecting individual sites subject to episodic diversifying selection," *PLoS Genetics*, vol. 8, no. 7, Article ID e1002764, 2012.
- [38] S. A. Bergervoet, C. K. Y. Ho, R. Heutink, A. Bossers, and N. Beerens, "Spread of highly pathogenic avian influenza (HPAI) H5N5 viruses in Europe in 2016–2017 appears related to the timing of reassortment events," *Viruses*, vol. 11, no. 6, p. 501, 2019.
- [39] H. Yang, P. J. Carney, V. P. Mishin et al., "Molecular characterizations of surface proteins hemagglutinin and neuraminidase from recent H5Nx avian influenza viruses," *Journal of Virology*, vol. 90, no. 12, pp. 5770–5784, 2016.
- [40] F. Ni, E. Kondrashkina, and Q. Wang, "Structural and functional studies of influenza virus A/H6 hemagglutinin," *PLoS One*, vol. 10, no. 7, Article ID e0134576, 2015.
- [41] H. Kong, D. F. Burke, T. J. da Silva Lopes et al., "Plasticity of the influenza virus H5 HA protein," *mBio*, vol. 12, no. 1, Article ID e03324-20, 2021.
- [42] J. Li, M. Gu, K. Liu et al., "Amino acid substitutions in antigenic region B of hemagglutinin play a critical role in the antigenic drift of subclade 2.3.4.4 highly pathogenic H5NX influenza viruses," *Transbound Emerg Dis*, vol. 67, no. 1, pp. 263–275, 2020.
- [43] C. J. Russell and R. G. Webster, "The genesis of a pandemic influenza virus," *Cell*, vol. 123, no. 3, pp. 368–371, 2005.
- [44] G. J. Smith and R. O. Donis, "Nomenclature updates resulting from the evolution of avian influenza A(H5) virus clades 2.1.3.2a, 2.2.1, and 2.3.4 during 2013–2014," *Influenza and Other Respiratory Viruses*, vol. 9, no. 5, pp. 271–276, 2015.
- [45] Fao, "H5N8 HPAI Global Situation Update," 2019, https://www.fao.org/ag/againfo/programmes/en/empres/H5N8/situation_update.html.
- [46] M. Lazniewski, W. K. Dawson, T. Szczepinska, and D. Plewczynski, "The structural variability of the influenza A hemagglutinin receptor-binding site," *Brief Funct Genomics*, vol. 17, no. 6, pp. 415–427, 2018.
- [47] Y. Watanabe, M. S. Ibrahim, H. F. Ellakany et al., "Acquisition of human-type receptor binding specificity by new H5N1 influenza virus sublineages during their emergence in birds in Egypt," *PLoS Pathogens*, vol. 7, no. 5, Article ID e1002068, 2011.
- [48] F. Ni, E. Kondrashkina, and Q. Wang, "Determinant of receptor-preference switch in influenza hemagglutinin," *Virology*, vol. 513, pp. 98–107, 2018.
- [49] J. Shin, S. Kang, H. Byeon et al., "Highly pathogenic H5N6 avian influenza virus subtype clade 2.3.4.4 indigenous in South Korea," *Scientific Reports*, vol. 10, no. 1, p. 7241, 2020.
- [50] R. Gao, M. Gu, L. Shi et al., "N-linked glycosylation at site 158 of the HA protein of H5N6 highly pathogenic avian influenza virus is important for viral biological properties and host immune responses," *Veterinary Research*, vol. 52, no. 1, p. 8, 2021.
- [51] H. Sun, G. Deng, H. Sun et al., "N-linked glycosylation enhances hemagglutinin stability in avian H5N6 influenza virus to promote adaptation in mammals," *PNAS Nexus*, vol. 1, no. 3, p. pgac085, 2022.
- [52] H. L. Yen, J. R. Aldridge, A. C. Boon et al., "Changes in H5N1 influenza virus hemagglutinin receptor binding domain affect systemic spread," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 1, pp. 286–291, 2009.
- [53] D. Zhao, L. Liang, S. Wang et al., "Glycosylation of the hemagglutinin protein of H5N1 influenza virus increases its virulence in mice by exacerbating the host immune response," *Journal of Virology*, vol. 91, no. 7, Article ID e02215-16, 2017.
- [54] X. Zhang, S. Chen, Y. Jiang et al., "Hemagglutinin glycosylation modulates the pathogenicity and antigenicity of the H5N1 avian influenza virus," *Veterinary Microbiology*, vol. 175, no. 2–4, pp. 244–256, 2015.
- [55] Y. Yin, S. Yu, Y. Sun et al., "Glycosylation deletion of hemagglutinin head in the H5 subtype avian influenza virus enhances its virulence in mammals by inducing endoplasmic reticulum stress," *Transbound Emerg Dis*, vol. 67, no. 4, pp. 1492–1506, 2020.
- [56] G. G. Simpson, *The Major Features of Evolution*, Columbia University Press, New York, NY, USA, 1953.

- [57] D. N. Reznick and R. E. Ricklefs, "Darwin's bridge between microevolution and macroevolution," *Nature*, vol. 457, no. 7231, pp. 837–842, 2009.
- [58] M. J. Poen, D. Venkatesh, T. M. Bestebroer et al., "Circulation of genetically distinct highly pathogenic avian influenza A clade 2.3.4.4 (H5N6) viruses in wild waterfowl and poultry in Europe and East Asia, 2017-18," *Virus Evolution*, vol. 5, no. 1, p. vez004, 2019.
- [59] K. J. C. Antigua, Y. H. Baek, W.-S. Choi et al., "Multiple HA substitutions in highly pathogenic avian influenza H5Nx viruses contributed to the change in the NA subtype preference," *Virulence*, vol. 13, no. 1, pp. 990–1004, 2022.
- [60] Y.-T. Li, Y. C. F. Su, and G. J. D. Smith, "H5Nx viruses emerged during the suppression of H5N1 virus populations in poultry," *Microbiology Spectrum*, vol. 9, no. 2, pp. 01309211–e201321, 2021.
- [61] S. Wu, Z. Wei, C. M. Greene et al., "Mortality burden from seasonal influenza and 2009 H1N1 pandemic influenza in Beijing, China, 2007-2013," *Influenza and Other Respiratory Viruses*, vol. 12, no. 1, pp. 88–97, 2018.
- [62] L. Ding, J. Li, X. Li, and B. Qu, "Evolutionary and mutational characterization of the first H5N8 subtype influenza A virus in humans," *Pathogens*, vol. 11, no. 6, p. 666, 2022.
- [63] B. T. Duong, D. D. Than, U. Ankhanbaatar et al., "Assessing potential pathogenicity of novel highly pathogenic avian influenza (H5N6) viruses isolated from Mongolian wild duck feces using a mouse model," *Emerging Microbes & Infections*, vol. 11, no. 1, pp. 1425–1434, 2022.
- [64] C. Zhang, H. Cui, L. Chen et al., "Pathogenicity and transmissibility of goose-origin H5N6 avian influenza virus clade 2.3.4.4h in mammals," *Viruses*, vol. 14, no. 11, p. 2454, 2022.