

Article

Global Prevalence and Hemagglutinin Evolution of H7N9 Avian Influenza Viruses from 2013 to 2022

Qianshuo Liu ^{1,2}, Haowen Zeng ^{1,2}, Xinghui Wu ^{1,2}, Xuelian Yang ² and Guiqin Wang ^{2,*}

¹ College of Life Sciences, Nanjing Normal University, Nanjing 210023, China; 211202116@njnu.edu.cn (Q.L.); zenghw8311@163.com (H.Z.); w63372011@163.com (X.W.)

² Nanjing Advanced Academy of Life and Health, Nanjing 211135, China; xlyang@ipsnanjing.cn

* Correspondence: gqwang@ipsnanjing.cn

Abstract: H7N9 avian influenza viruses have caused severe harm to the global aquaculture industry and human health. For further understanding of the characteristics of prevalence and hemagglutinin evolution of H7N9 avian influenza viruses, we generated the global epidemic map of H7N9 viruses from 2013 to 2022, constructed a phylogenetic tree, predicted the glycosylation sites and compared the selection pressure of the hemagglutinin. The results showed that although H7N9 avian influenza appeared sporadically in other regions worldwide, China had concentrated outbreaks from 2013 to 2017. The hemagglutinin genes were classified into six distinct lineages: A, B, C, D, E and F. After 2019, H7N9 viruses from the lineages B, E and F persisted, with the lineage B being the dominant. The hemagglutinin of highly pathogenic viruses in the B lineage has an additional predicted glycosylation site, which may account for their persistent pandemic, and is under more positive selection pressure. The most recent ancestor of the H7N9 avian influenza viruses originated in September 1991. The continuous evolution of hemagglutinin has led to an increase in virus pathogenicity in both poultry and humans, and sustained human-to-human transmission. This study provides a theoretical basis for better prediction and control of H7N9 avian influenza.

Keywords: H7N9 avian influenza viruses; epidemiology; genetic evolution; selective pressure; glycosylation site



Citation: Liu, Q.; Zeng, H.; Wu, X.; Yang, X.; Wang, G. Global Prevalence and Hemagglutinin Evolution of H7N9 Avian Influenza Viruses from 2013 to 2022. *Viruses* **2023**, *15*, 2214. <https://doi.org/10.3390/v15112214>

Academic Editors: Ayato Takada and Julian Tang

Received: 2 August 2023

Revised: 9 October 2023

Accepted: 16 October 2023

Published: 4 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The novel H7N9 avian influenza viruses [1], which commenced infecting humans in early 2013 [2], were a result of the recombination of the hemagglutinin (HA) of the H7N7 avian influenza viruses, the neuraminidase (NA) of the H7N9 viruses, and the internal gene segments of H9N2 viruses. In 2016, following the introduction of novel H7N9 viruses from the Yangtze River Delta to the Pearl River Delta region, an insertion mutation of four amino acids, RKRT, occurred at the cleavage site of the HA protein. This mutation led to an elevation in the virulence of the viruses in avian species, subsequently resulting in outbreaks across eight provinces in China [3]. The highly pathogenic (HP) H7N9 viruses caused almost complete mortality in infected chickens, leading to significant poultry losses and culling. Moreover, these viruses also inflicted a mortality rate of approximately 50% among human cases [4]. From its initial outbreak until October 2017, the novel H7N9 avian influenza viruses had caused five major epidemics, with the World Health Organization (WHO) reporting a total of 1564 laboratory-confirmed cases of human infection [5]. Thereafter, the Chinese government initiated large-scale mandatory immunization programs for poultry, which proved to be effective in reducing the incidence of avian infection and reducing human infections. During the 2018 flu season, only three cases of human infection were reported, and as of May of that year, there were only five reported incidents of avian outbreaks [6]. However, despite the implementation of mandatory immunization, the novel H7N9 viruses were not eradicated. In 2018 and 2019, these viruses infected poultry again

in chicken farms and zoos in northern China [7]. In late March 2019, after over a year of no reported cases, human infections with the HPAI H7N9 viruses resurfaced [8].

In addition to the new H7N9 virus that emerged in 2013, the earlier H7N9 avian influenza viruses have been reported in wild birds in the USA, Canada, Guatemala, Spain, Sweden, Egypt, Mongolia, and Taiwan [9]. In recent years, the earlier H7N9 influenza viruses have been detected multiple times in poultry and wild birds in various states in the Midwest of the United States [10], including both low-pathogenicity strains and highly pathogenic ones [11,12]. The origins of the earlier H7N9 viruses in Asia can be traced back to the year of 2000 [13]. In 2021, a low-pathogenicity strain of the earlier H7N9 avian influenza viruses was isolated in South Korea [14]. H7N9 avian influenza viruses pose a threat to global poultry farming and human health [15].

HA serves as one of the key membrane proteins of avian influenza A viruses [16]. The HA protein is responsible for binding the viruses to sialic acid receptors on the surface of host cells, and is one of the factors that determines the host species [17,18]. The HA protein is also an important target protein for humoral immunity of influenza virus. The HA plays a critical role in the transmission, evolution, immune response and lifecycle of avian influenza viruses [19,20]. Understanding the genetic evolution characteristics of HA is crucial in preventing and controlling avian influenza pandemics.

To gain insights into the global prevalence and HA evolution characteristics of H7N9 viruses, this study conducted a comprehensive analysis of the global circulation of H7N9 viruses from January 2013 to December 2022. This study involved constructing phylogenetic trees and lineage assignment of the HA protein of these viruses. Our study predicted the N-glycan positions on the HA protein for the virus strains that have continued to circulate since 2019. Furthermore, this study analyzed and compared the selective pressures acting on the HA protein.

2. Materials and Methods

2.1. Data Collection and Compilation

For this study, a total of 2597 H7N9 avian influenza virus strains were collected from the Global Initiative on Sharing All Influenza Data (GISAID) database. The dataset included information such as the date and location of sample collection, as well as the genetic and protein sequences of 2549 virus strains [21]. The study period spanned from 1 January 2013, to 31 December 2022, excluding any virus strains labeled as “reassortant” in their names. Based on the temporal and spatial data, a bar chart was created using the ggplot2 package in the R programming language [22]. Additionally, a global map depicting the prevalence of H7N9 avian influenza was generated using the Datawrapper website (<https://www.datawrapper.de/> accessed on 15 March 2023). For detailed information on virus sample collection dates, locations, and other relevant details, please refer to Supplementary Material S1.

2.2. Construction of the Phylogenetic Tree

In this study, the BioAider V1.527 software [23] was utilized for the alignment of gene sequences using the MAFFT algorithm [24]. The aligned gene sequences were subsequently trimmed in the MEGA 11 software, retaining only the coding region between the start and stop codons. For reducing computational load, the trimmed gene sequences were subjected to 100% deduplication using the BioAider V1.527 software [23]. The neighbor-joining (NJ) method in MEGA 11 software was employed to construct the phylogenetic tree [25,26]. The resulting tree was further refined and annotated using the FigTree 1.4.4 software, which facilitated the labeling of virus lineages, sample collection locations, and dates (<http://tree.bio.ed.ac.uk/software/figtree/>, accessed on 20 January 2023). The non-redundant HA gene sequences can be found in Supplementary Material S2.

We also downloaded all the HA sequences of H7 subtype influenza viruses collected from 2013 to 2022 from the GISAID database. A phylogenetic tree was constructed with

the same method. We used MEGA software to calculate the mean genetic distance of H7, H7N9, and the novel H7N9 avian influenza viruses [25].

2.3. The Estimation of the Time to the Most Recent Common Ancestor (tMRCA)

The HA gene sequences of the viruses, the phylogenetic tree, and the sample collection dates of the viruses were imported into the Treetime V 0.11.1 program. Utilizing the linear relationship between the root-to-tip genetic distance and the sample collection dates of virus strains, we estimated, respectively, the time to the most recent common ancestor (tMRCA) for the HA genes from the H7N9 influenza viruses and the H7N9 viruses in the evolutionary tree [27].

2.4. Prediction of N-Glycan Positions on the HA Protein of H7N9 Avian Influenza Viruses since 2019

The amino acid sequences of the HA protein of the H7N9 viruses after 2019 were input into the online server NetNGlyc 1.0 to predict potential N-glycan positions [28]. The server predicted potential sites based on the presence of the N-X-S/T motif on HA proteins, where X represents any amino acid except proline. A threshold value of >0.5 suggests glycosylation [29]. The amino acid lollipop plot, which described the N-glycosylation sites, was generated using the TrackViewer V 1.36.2 package in R [30].

2.5. Analysis of the Selection Pressure of H7N9 Viruses HA since 2019

The selective pressure analysis and identification of positively selected sites for different lineages of H7N9 viruses were conducted, using the online server Datamonkey [31], after removing stop codons and signal peptides from the aligned HA gene sequences. The Omega value (dn/ds), which represents the ratio of the non-synonymous substitution rate (dn) to the synonymous substitution rate (ds), is used to measure the selection pressure. A dn/ds > 1 indicates positive selection, <1 indicates negative selection, and =1 indicates neutral selection [32].

In order to identify positively selected sites, four different algorithms were utilized: Fixed Likelihood Ancestor Counting (FLAC) [33], Mixed-Effects Model of Evolution (MEME) [34], Fixed Effects Likelihood (FEL) [33], and Fast Unconstrained Bayesian Approximation (FUBAR) [35]. The thresholds for the calculation results of MEME, FLAC, and FEL were set at $p < 0.05$, while the threshold for FUBAR was set at $p > 0.9$. Only when at least two algorithms supported the positive selection at specific sites were those sites considered positively selected.

Furthermore, the dn/ds ratios, reflecting the selection pressure, was calculated separately for the HA genes of different lineages of H7N9 avian influenza viruses, from the different virus hosts, using the MEME and SLAC algorithms. The average values were determined using GraphPad Prism software (version 9.1.0, GraphPad Software, San Diego, CA, USA), and a graph depicting the dn/ds ratios of different lineages was generated.

3. Results

3.1. Global Prevalence of H7N9 Avian Influenza Viruses

Though H7N9 avian influenza also appeared sporadically in other regions worldwide, China has witnessed periodic outbreaks from 2013 to 2017. Since 2017, its outbreaks have transitioned into a more sporadic pattern (Figure 1A), as reported in Asia (Chinese mainland, Taiwan, Hong Kong, South Korea, Japan, Malaysia, and Bangladesh), Europe (the Netherlands and the United Kingdom), North America (Canada and the United States), Africa (Egypt), and Oceania (Australia). Among the regions affected by the viruses, Asia has the highest number of cases, with a significant concentration in the eastern region, particularly Chinese mainland (Figure 1B).

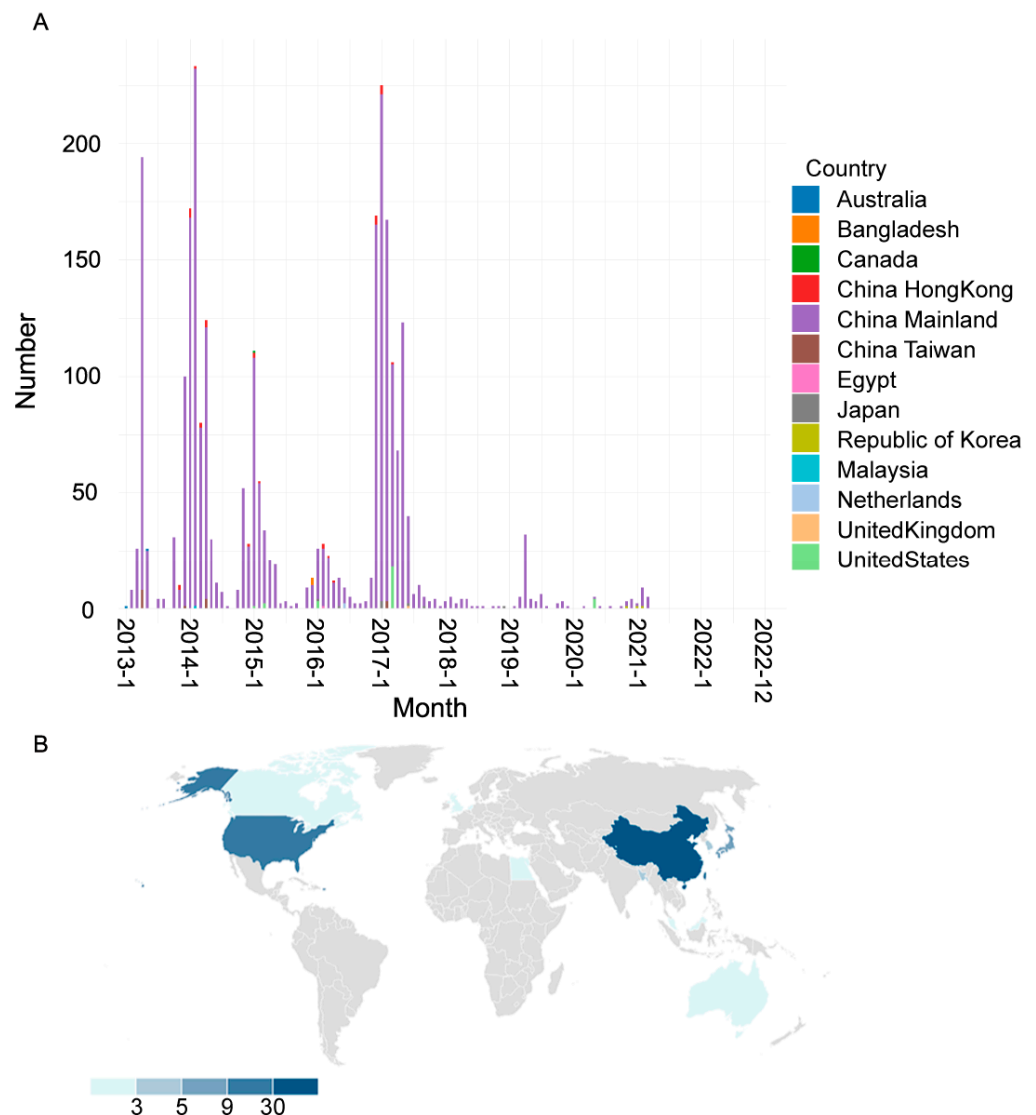


Figure 1. Global prevalence of H7N9 avian influenza from 2013 to 2022. (A) The distribution of H7N9 viruses in different countries and months. (B) The world map shows the total number of H7N9 virus strains isolated in different countries.

By 2018, H7N9 avian influenza had experienced five periodic outbreaks as follows: the first wave from February 2013 to September 2013, the second from October 2013 to September 2014, the third from October 2014 to September 2015, the fourth from October 2015 to September 2016, and the fifth from October 2016 to September 2017. These five major outbreaks primarily occurred in Chinese mainland. Since the middle of 2017, the number of H7N9 virus cases have started to decrease, although sporadic cases still occurred. Since March 2021, H7N9 viruses have not been isolated (Figure 1A).

3.2. Phylogenetic Analysis of HA Genes of H7N9 Avian Influenza Viruses

The phylogenetic tree was constructed using 1540 non-redundant HA gene sequences of the H7N9 avian influenza viruses from 2013 to 2022. Based on the topology of the tree, the location and time of virus isolation, the common ancestor and previous studies, the HA genes were classified into six distinct lineages labeled A, B, C, D, E, and F (Figure 2A).

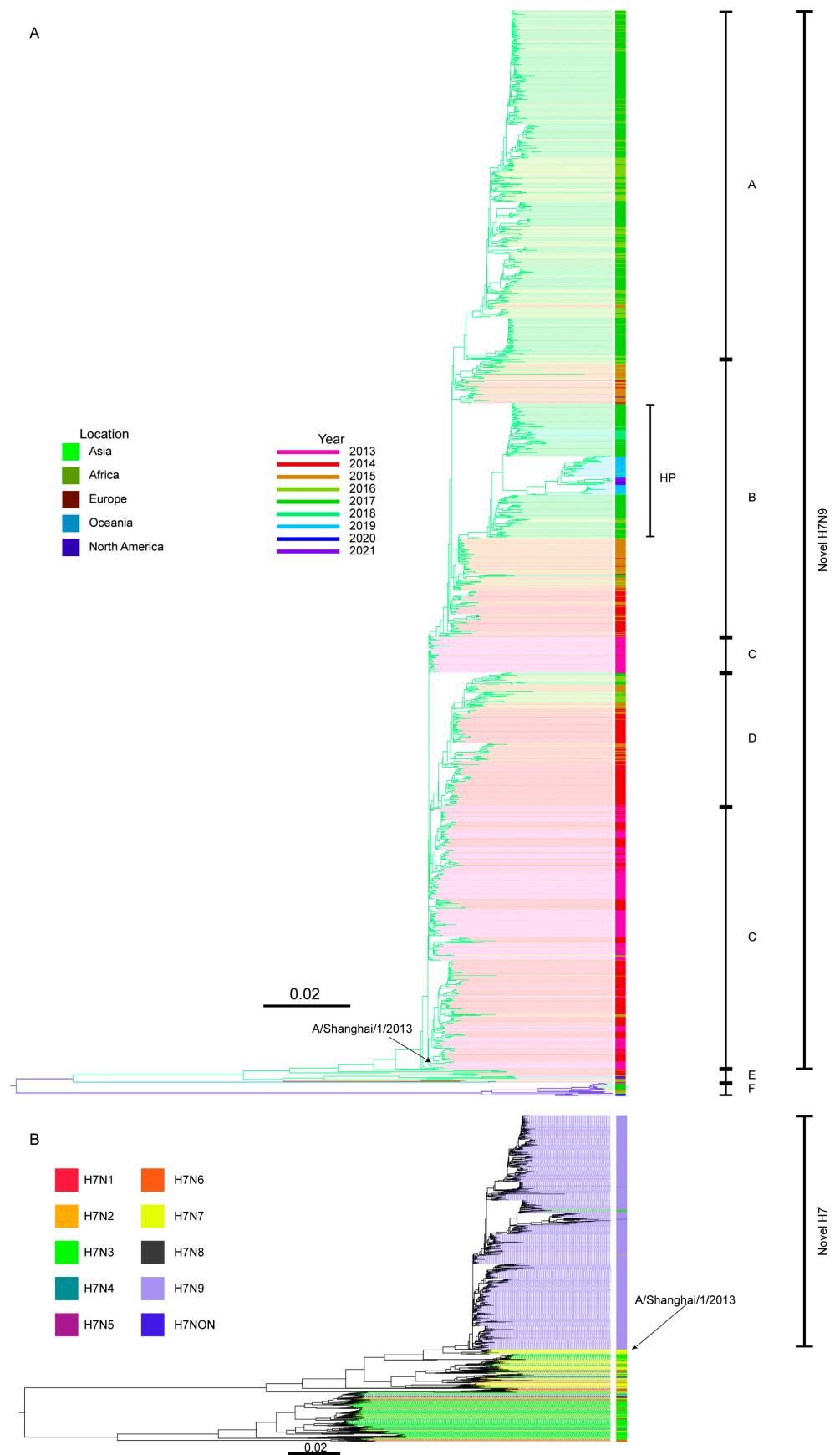


Figure 2. Phylogenetic analysis of the HA genes of H7 influenza viruses from 2013 to 2022. **(A)** Phylogenetic tree of the HA genes of H7N9 influenza viruses. Nucleotide sequences of 1540 full-length

HA genes were analyzed. The letters A, B, C, D, E, and F indicate the lineages. HP: highly pathogenic avian influenza virus. (B) Phylogenetic tree of the HA genes of the H7 influenza viruses. Nucleotide sequences of 2135 full-length HAs were analyzed. Both of the phylogenetic trees are rooted at the midpoint. The trees were generated by the neighbor-joining (NJ) method in MEGA 11 software. The different branch color indicates sample collection location of virus strain. The middle color ribbon indicates the year of virus collection. Scale bar: 0.02 amino acid substitutions per site.

The viruses belonging to the lineages A, B, C, and D were classified as the novel H7N9 avian viruses and were more abundant, while those in lineages E and F, were less numerous (Figure 2A).

In terms of geographic distribution, the H7N9 avian influenza viruses in the different lineages exhibit differences. The viruses in the lineages A, C, and D are primarily prevalent in Asia. Within the lineage B, most of the viruses are isolated in Asia, except for one isolated in North America. Notably, all the high-pathogenicity viruses found in Chinese mainland belong to the lineage B. The viruses in the lineage F are exclusively distributed in North America, while the viruses in the lineage E viruses were be found in Oceania, Africa, Europe, and Asia.

The HA genes from the different lineages exhibit differences in their evolutionary characteristics. Compared with other HA genes, the HA phylogenetic tree branches of the high-pathogenicity viruses in the lineage B and the viruses in the lineage F demonstrates longer branch lengths. That indicates more genetic divergence and low homology of HA genes within these lineages (Figure 2A). The branches of the HA phylogenetic tree of the low-pathogenicity viruses in the lineages A, B, C, and D display a comb-like structure, characterized by short horizontal distances and relatively larger number of virus strains horizontally. This suggests that the high homology of HA genes within these lineages (Figure 2A).

The predominant lineages of the H7N9 avian influenza viruses vary across different years. In 2013, the H7N9 viruses were primarily from the lineages C, D, and E. During the period of 2014 to 2015, the low-pathogenicity viruses was from the lineage B. In 2016, the viruses from the lineages A and B became the main reservoirs, with minor presence in the lineages C and D. From 2017 to 2018, the viruses from the lineages A and B remained predominant, accompanied by a smaller presence in the lineage D. In 2018, all viruses were isolated from the lineages A and B, while after 2019, only the viruses from the B lineages remained. Notably, the viruses from the lineage E exhibited continuous distribution in the evolutionary tree from 2013 to 2021, except for 2017 and 2019. The viruses from the lineage F were observed between 2015 and 2017, and subsequently reemerged in 2020 (Figure 2A).

We constructed the HA phylogenetic tree of the H7 viruses from 2013 to 2022 (Figure 2B). Then, we compared the genetic diversity of the HA sequences from all the H7, H7N9 and novel H7N9 viruses. The genetic diversity was measured by the mean genetic distance. The two have a positive correlation. The mean genetic distance of the HA sequences of all the H7, H7N9 and novel H7N9 viruses from 2013 to 2022 is 0.07, 0.03 and 0.02, respectively. That indicates the degree of HA gene diversity from high to low is the H7, H7N9 and novel H7N9 viruses.

3.3. Time to the Most Recent Common Ancestor (TMRCA)

Using the Treetime V 0.11.1 program, the linear relationship between the root-to-tip genetic distance of HA genes of the H7N9 avian influenza viruses and the date of virus sample collection was calculated. The results revealed that HA genes from the different lineages could be traced back to a common ancestor in September 1991 (Figure 3A). The tMRCA of the HA genes of the novel H7N9 influenza viruses, is January 2013 (Figure 3B).

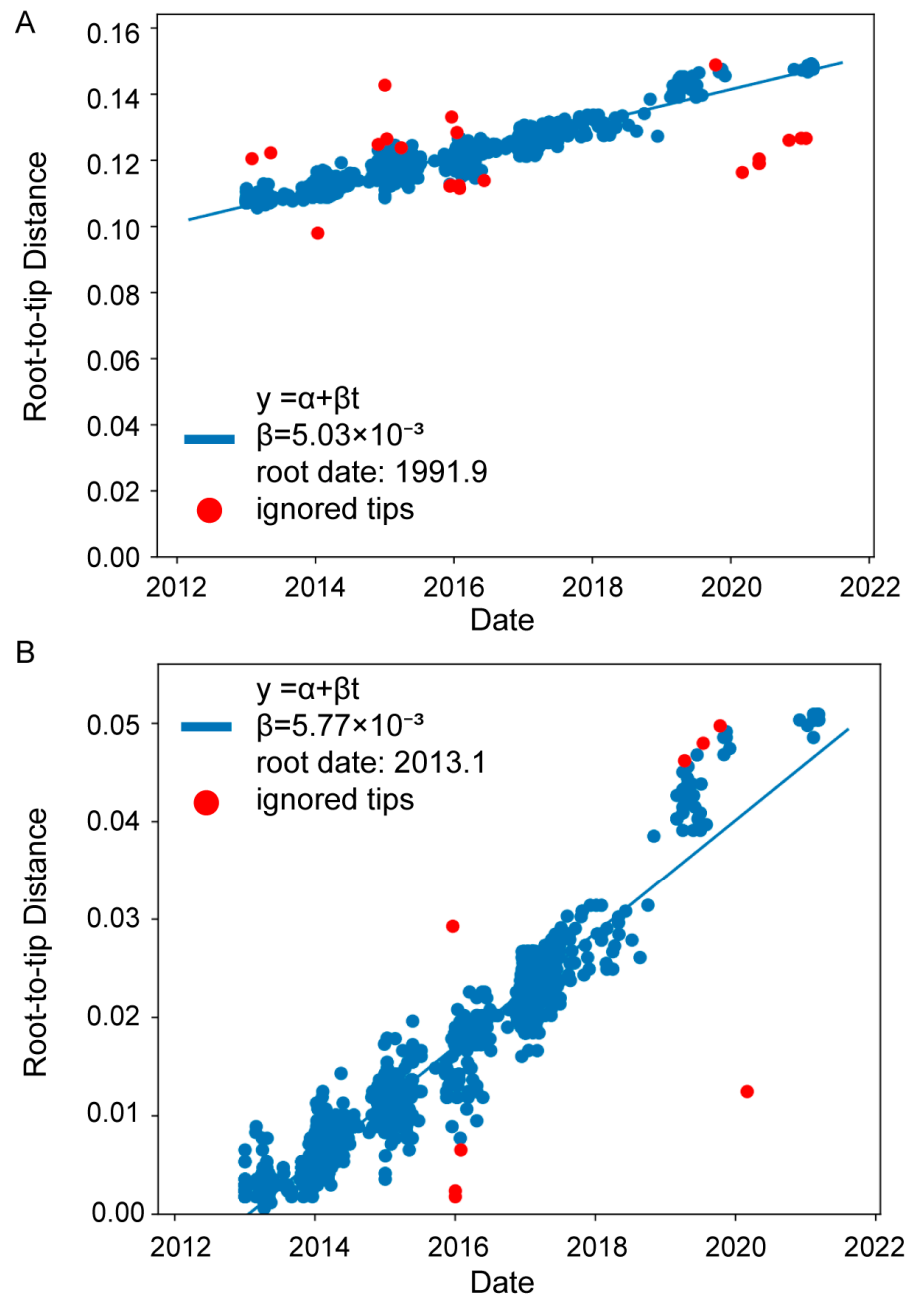


Figure 3. The linear relationship between the root-to-tip genetic distance and the time of virus collection. The tMRCA of the HA sequences from all the H7N9 viruses (A) and the novel H7N9 viruses (B) was estimated, respectively. The linear regression was performed using the Treetime V 0.11.1 software. X-axis represents the time of virus collection and Y-axis represents the root-to-tip genetic distance on the NJ tree. The correlation coefficient (R^2) between the two variables is 0.95. Viruses are represented by circular points, with blue circles indicating conformity to the linear relationship and red circles representing outliers.

3.4. Characteristics of Evolution, Prevalence and N-Glycans of H7N9 Avian Influenza Viruses after 2019

Since 2019, a total of 78 H7N9 viruses have been distributed among the lineages B, E, and F (Figure 2). Among them, the lineage B viruses originated from China, including 71 strains of both high and low pathogenicity. The lineage E viruses were found in South Korea, comprising three strains. The lineage F viruses were identified in the United States, consisting of four strains (Figure 4A,B).

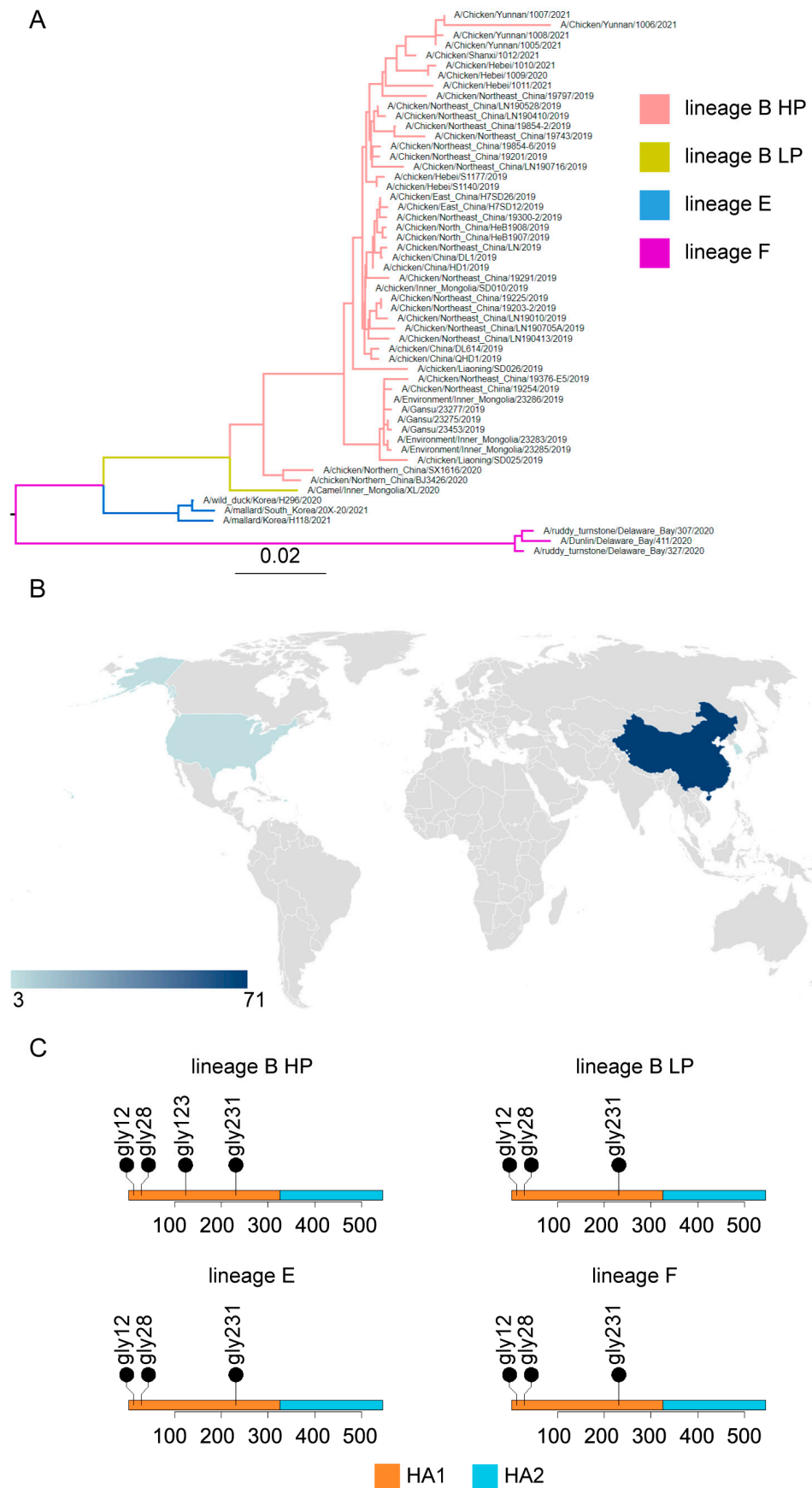


Figure 4. Characteristics of HA evolution, prevalence and N-glycans of H7N9 avian influenza viruses after 2019. (A) Phylogenetic analysis of HA amino acid sequences of the H7N9 avian influenza viruses.

The tree was generated by the neighbor-joining (NJ) method in MEGA 11 software. Amino acid sequences of 54 full-length HAs were analyzed. The different branch color indicates the HA lineages. (B) The world map shows the total number of H7N9 viruses isolated in different countries. (C) Prediction of N-glycan positions on the HA proteins of H7N9 avian influenza viruses. The lollipop plot was constructed using the TrackViewer V 1.36.2 software package in the R. The HA1 and HA2 regions are distinguished using different colors.

The potential N-glycan positions were predicted based on the amino acid sequences of the HA protein of H7N9 viruses after 2019. High-pathogenicity viruses from the lineage B possess four glycosylation sites: 12, 28, 123, and 231, while low-pathogenicity viruses from the lineage B lack the site of 123. Viruses from the lineages E and F have three potential glycosylation sites: 12, 28, and 231. These results indicate that the lineages B, E, and F share the 12, 28, and 231 sites, while the 123 site is unique to the high-pathogenicity viruses in the lineage B (Figure 4C and Table 1). These potential glycosylation sites are located in the HA1 region of the HA protein.

Table 1. Prediction of N-glycan positions on the HA protein.

Lineages	N-Glycan Positions on the HA Protein			
	12	28	123	231
Lineage B China HP	+++	++	++	++
Lineage B China LP	+++	++	Null	++
Lineage E Korea	++	++	Null	++
Lineage F American	++	++	Null	++

++: Potential < 0.5 and Jury agreement (9/9) OR Potential < 0.75; +++: Potential < 0.75 and Jury agreement; Null: non-glycan positions.

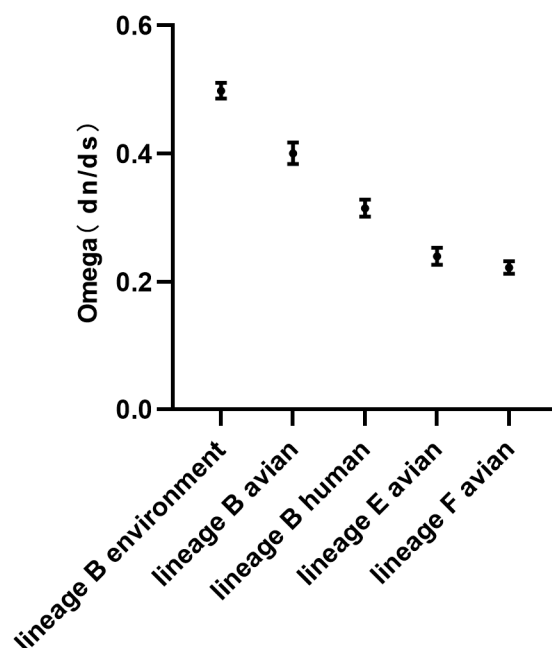
3.5. The Selection Pressure of H7N9 Viruses HA since 2019

Utilizing the Datamonkey online server, the selection pressure of the specific sites of the HA sequences of the H7N9 viruses after 2019 were analyzed. The analyses revealed positive selection signals at amino acid positions 47, 253, 546, and 567 in the HA protein in the lineage B viruses. These variations may play a role in viruses' adaptation to the current environment or other factors. No positive selection sites were detected in the remaining lineages, indicating that the HA protein of the lineage B may exhibit a higher antigenic drift rate.

After 2019, the viruses in the lineages E and F exclusively infected avian, and the lineage B viruses were from avian, humans, mammals and the environment. The Omega values (dn/ds) of the HA genes from the different hosts in the lineages B, E and F were compared. The results showed that the Omega values of HA genes of the B lineage viruses from environmental, avian, and mammalian were 0.499, 0.401, and 0.316, respectively. The Omega values of HA genes in the E and F lineage viruses were 0.401 and 0.316, respectively (Table 2). All the Omega values were less than 1, indicating negative selection pressure on the HA genes. These findings imply that HA genes in the lineages B, E and F tend to retain conserved traits after undergoing limited adaptive evolution. The higher Omega value indicates that the HA genes in the lineage B may experience a relatively moderate negative selection pressure. The dn/ds ratio of the HA genes in the B lineage viruses from human is lower than that from avian, indicating that the negative selection pressure on the HA genes from human is greater than that from avian. The HA genes in the lineages E and F exist at higher synonymous mutation rates and lower non-synonymous mutation rates, indicating that they have stronger negative selection pressure (Figure 5).

Table 2. The dn/ds values calculated by the MEME and SLAC methods.

Lineage	MEME	SLAC	Average
lineage B human	0.306	0.325	0.316
lineage B avian	0.389	0.413	0.401
lineage B environment	0.490	0.507	0.499
lineage E avian	0.231	0.250	0.241
lineage F avian	0.216	0.230	0.223

**Figure 5.** Selection pressure analysis of HA genes of the different lineages after 2019. The dots represent the average value of dn/ds calculated by the MEME and SLAC methods, with the SD values.

4. Discussion

The threat posed by H7N9 avian influenza viruses is escalating. Typically, the viruses require prolonged evolution to cause major epidemics and cross-species transmission [36]. According to this study, the common ancestor of H7N9 avian influenza viruses emerged in September 1991 (Figure 3A). In 2013, the novel H7N9 avian influenza viruses first crossed the species barrier and infected humans. In 2016, an insertion mutation occurred at the HA protein cleavage site, increasing the pathogenicity of the viruses in both poultry and humans [3].

The H7N9 avian influenza viruses have not acquired the ability for sustained human-to-human transmission [37]. It has been confirmed that although viruses from the lineage B were introduced to Canada through human air travel from China, local transmission has not occurred [38]. In 2017, co-circulation of lineages from the Yangtze River Delta and the Pearl River Delta regions in China increased the risk of recombination and mutation among viruses with different pathogenicity [39].

In 2020, a novel low-pathogenic H7N9 virus originating from poultry was transmitted across species, infecting non-human mammals, such as camels [40]. The novel H7N9 viruses, originating in China, have undergone significant changes, with the main circulating lineages shifting the lineages C and D to lineages A and B. Since 2019, the lineage B has been predominant, with high-pathogenicity viruses dominating in this lineage. The H7N9 viruses in the lineages E and F remain relatively low in number, and there have been no major shifts in the predominant lineages (Figure 2A). However, the H7N9 viruses in the

lineages E and F have also been involved in cross-species transmission events between poultry and wild animals [12].

The H7N9 avian influenza viruses in the lineages A, B, C and D, despite their different evolutionary patterns, continue to evolve. It is crucial for governments worldwide to actively implement epidemic prevention and control measures to prepare for the potential occurrence of a large-scale H7N9 avian influenza pandemic.

The accumulation of N-glycosylation sites is a common immune evasion strategy used by various enveloped viruses [41]. By adding glycans in the antibody recognition region, viruses can escape host immunity through steric hindrance [42]. The globular head of HA contains the receptor binding region and major antigenic sites, serving as the primary binding sites for neutralizing antibodies [43]. Glycans have the ability to conceal antigenic epitopes and impede their binding with antibodies, thereby altering the immunological pressure on the population or enhancing viral fitness [29]. This phenomenon has been corroborated in studies on H1N1 and H3N2 avian influenza viruses [44,45]. Upon initial infection with avian influenza viruses or vaccination, the majority of antibody responses are directed toward the HA head domain, which serves as the target of vaccines [46]. Alterations in glycosylation sites can impact the efficacy of vaccines targeting HA. Among the predicted glycosylation sites in this study, positions 123 and 231 are located in the head region of the HA protein, while positions 12 and 28 are in the stalk region [47]. Compared to the HAs in other lineages, the HAs in the lineage B, particularly that of the high-pathogenicity strains, exhibits an additional glycosylation site at position 123, which is associated with a higher number of high pathogenic viruses within the lineage B. Therefore, this glycosylation site may be one of the contributing factors to the sustained circulation of the H7N9 viruses in the lineage B. In contrast to the head region, the stalk region of the HA protein has lower immunogenicity, and glycosylation at these sites primarily affects the structure and oligomerization of HA [48], potentially without a considerable impact on viruses' circulation.

Since 2019, all the H7N9 HAs worldwide have been experiencing purifying selection, indicated by $dn/ds < 1$. This suggests that non-synonymous mutations may contribute to the increased diversity in the natural selection gene pool compared to synonymous mutations [49]. The higher dn/ds values observed in the lineage B indicate a stronger positive selection pressure, further supporting the notion that the increased glycosylation sites in the HA head region enhance the viruses' ability to adapt to new hosts [50]. Amino acid positions 47, 253, 546, and 567 in the lineage B viruses have undergone adaptive changes and are likely involved in the functional evolution of H7N9 viruses.

The novel H7N9 influenza viruses caused five waves of human infections from 2013 to 2017, with a total of 1568 cases, of which 615 were fatal [4]. The novel H7N9 influenza viruses were low pathogenicity in birds. In early 2017, some low-pathogenicity H7N9 strains became highly pathogenic in chickens, causing severe influenza outbreaks in poultry in China [4]. H7N9 viruses posed the severe threat to human and animal health. To control the epidemic, the poultry were mandatorily immunized with H7N9 inactivated vaccine from September 2017 in China [15]. After birds were inoculated with the H7 vaccine, infection cases in humans and birds dramatically reduced. This suggests that the national poultry vaccination strategy has successfully controlled H7N9 infections in poultry and humans. Compared to the earlier novel H7N9 viruses, the HA gene of the H7N9 virus from the latest human case showed a relatively long genetic distance [4]. This indicates that the H7N9 viruses have evolved considerably even though vaccination dramatically limited the H7N9 prevalence. H7N9 viruses will not be eradicated. The viruses lurk in natural hosts such as waterfowls and continue to evolve, occasionally infecting other animals such as poultry. Once the virus has accumulated enough mutations and recombination, it could lead to the next pandemic. Based on World Health Organization (WHO) recommendations, a total of six virus strains were used for the development of influenza A(H7N9) vaccines from 2013 to 2023. The WHO may continue to recommend vaccine strains until all cases of human and poultry infection have disappeared.

This study analyzed the global epidemiology and the HA evolutionary characteristics of the H7N9 viruses from 2013 to 2022. The GISAID data may have a lag, and some viruses may be collected in 2022 but uploaded in 2023. These viruses will not be included in our study. This study also predicted the glycosylation sites of the HA protein in the different lineages of H7N9 strains since 2019. By comparing the selection pressure on the HA protein of H7N9 viruses among the different lineages, this study provided insights into HA evolutionary dynamics. In summary, this study provides a theoretical basis for better prediction and control the spread of H7N9 avian influenza.

5. Conclusions

H7N9 avian influenza appeared sporadically in other regions worldwide, but China witnessed periodic concentrated outbreaks from 2013 to 2017. HA genes of the H7N9 avian influenza viruses originated from a common ancestor in 1991 and have undergone multiple evolutions and mutations. The year 2013 witnessed its first cross-species infection of humans, while 2016 witnessed an insertion of a mutation in the cleavage site of the HA protein, which increased its pathogenicity. There are differences in the HA evolutionary pattern of H7N9 avian influenza viruses of the different lineages, with a higher number of glycosylation sites in highly pathogenic viruses of the B lineage, which may account for their persistent pandemic. Globally, the HA proteins of H7N9 viruses have undergone a natural selection and adaptive changes, and in particular, HA genes of the lineage B have been subjected to a strong positive selection pressure, which further promotes the viruses' ability to adapt to new hosts.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v15112214/s1>, The sample collection time and location information of the viruses is in file S1: Supplementary File1.xlsx. The de-redundant gene sequence is in file S2: Supplementary File2.fas.

Author Contributions: Q.L. processed the data and completed the first draft; X.Y. revised and polished this article; H.Z. and X.W. revised this article; G.W. is responsible for selecting topics, writing, and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the research grants from the Innovation Capacity Building Project of Jiangsu province (BM2020019) and the Science and technology projects of Jiangsu Province (BE2021756) of Jiangsu Provincial Department of Science and Technology.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The viruses' sequence, date and location of sample collection, and other information involved in this study can all be found in GISAID.

Acknowledgments: The research was supported by the technology platform of Nanjing Advanced Academy of Life and Health.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lam, T.T.; Wang, J.; Shen, Y.; Zhou, B.; Duan, L.; Cheung, C.L.; Ma, C.; Lycett, S.J.; Leung, C.Y.; Chen, X.; et al. The Genesis and Source of the H7N9 Influenza Viruses Causing Human Infections in China. *Nature* **2013**, *502*, 241–244. [[CrossRef](#)]
2. Gao, R.; Cao, B.; Hu, Y.; Feng, Z.; Wang, D.; Hu, W.; Chen, J.; Jie, Z.; Qiu, H.; Xu, K.; et al. Human Infection with a Novel Avian-Origin Influenza A (H7N9) Virus. *N. Engl. J. Med.* **2013**, *368*, 1888–1897. [[CrossRef](#)]
3. Qi, W.; Jia, W.; Liu, D.; Li, J.; Bi, Y.; Xie, S.; Li, B.; Hu, T.; Du, Y.; Xing, L.; 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*, e00921-17. [[CrossRef](#)]
4. Li, C.; Chen, H. H7N9 Influenza Virus in China. *Cold Spring Harb. Perspect. Med.* **2021**, *11*, a038349. [[CrossRef](#)]
5. Xiang, D.; Pu, Z.; Luo, T.; Guo, F.; Li, X.; Shen, X.; Irwin, D.M.; Murphy, R.W.; Liao, M.; Shen, Y. Evolutionary Dynamics of Avian Influenza N H7N9 Virus across Five Waves in Mainland China, 2013–2017. *J. Infect.* **2018**, *77*, 205–211. [[CrossRef](#)]

6. Wang, G.L.; Gray, G.C.; Chen, J.M.; Ma, M.J. Will China's H7N9 Control Strategy Continue to Be Effective? *Open Forum Infect. Dis.* **2019**, *6*, ofz258. [[CrossRef](#)]
7. Yin, X.; Deng, G.; Zeng, X.; Cui, P.; Hou, Y.; Liu, Y.; Fang, J.; Pan, S.; Wang, D.; Chen, X.; et al. Genetic and Biological Properties of H7N9 Avian Influenza Viruses Detected after Application of the H7N9 Poultry Vaccine in China. *PLoS Pathog.* **2021**, *17*, e1009561. [[CrossRef](#)]
8. Yu, D.; Xiang, G.; Zhu, W.; Lei, X.; Li, B.; Meng, Y.; Yang, L.; Jiao, H.; Li, X.; Huang, W.; et al. The Re-Emergence of Highly Pathogenic Avian Influenza H7N9 Viruses in Humans in Mainland China, 2019. *Eurosurveillance* **2019**, *24*, 1900273. [[CrossRef](#)]
9. Olson, S.; Gilbert, M.; Cheng, M.C.; Mazet, J.A.K.; Joly, D. Historical Prevalence and Distribution of Avian Influenza Virus A(H7N9) among Wild Birds. *Emerg. Infect. Dis. J.* **2013**, *19*, 2031. [[CrossRef](#)]
10. Lebarbenchon, C.; Pedersen, J.C.; Sreevatsan, S.; Ramey, A.M.; Dugan, V.G.; Halpin, R.A.; Ferro, P.J.; Lupiani, B.; Enomoto, S.; Poulson, R.L.; et al. H7N9 Influenza A Virus in Turkeys in Minnesota. *J. Gen. Virol.* **2015**, *96*, 269–276. [[CrossRef](#)]
11. Lee, D.H.; Torchetti, M.K.; Killian, M.L.; Berhane, Y.; Swayne, D.E. Highly Pathogenic Avian Influenza A(H7N9) Virus, Tennessee, USA, March 2017. *Emerg. Infect. Dis.* **2017**, *23*, 1860–1863. [[CrossRef](#)] [[PubMed](#)]
12. Belser, J.A.; Brock, N.; Sun, X.; Jones, J.; Zanders, N.; Hodges, E.; Pulit-Penalzo, J.A.; Wentworth, D.; Tumpey, T.M.; Davis, T.; et al. Mammalian Pathogenesis and Transmission of Avian Influenza A(H7N9) Viruses, Tennessee, USA, 2017. *Emerg. Infect. Dis.* **2018**, *24*, 149–152. [[CrossRef](#)] [[PubMed](#)]
13. Cheng, M.C.; Lee, M.S.; Ho, Y.H.; Chyi, W.L.; Wang, C.H. Avian Influenza Monitoring in Migrating Birds in Taiwan During 1998–2007. *Avian Dis.* **2010**, *54*, 109–114. [[CrossRef](#)] [[PubMed](#)]
14. Na, E.-J.; Kim, Y.-S.; Kim, Y.-J.; Park, J.-S.; Oem, J.-K. Genetic Characterization and Pathogenicity of H7N7 and H7N9 Avian Influenza Viruses Isolated from South Korea. *Viruses* **2021**, *13*, 2057. [[CrossRef](#)]
15. Shi, J.; Zeng, X.; Cui, P.; Yan, C.; Chen, H. Alarming Situation of Emerging H5 and H7 Avian Influenza and Effective Control Strategies. *Emerg. Microbes Infect.* **2023**, *12*, 2155072. [[CrossRef](#)] [[PubMed](#)]
16. Padilla-Quirarte, H.O.; Lopez-Guerrero, D.V.; Gutierrez-Xicotencatl, L.; Esquivel-Guadarrama, F. Protective Antibodies against Influenza Proteins. *Front. Immunol.* **2019**, *10*, 1677. [[CrossRef](#)]
17. Medina, R.A.; García-Sastre, A. Influenza A Viruses: New Research Developments. *Nat. Rev. Microbiol.* **2011**, *9*, 590–603. [[CrossRef](#)]
18. Jones, S.; Nelson-Sathi, S.; Wang, Y.; Prasad, R.; Rayen, S.; Nandel, V.; Hu, Y.; Zhang, W.; Nair, R.; Dharmaseelan, S.; et al. Evolutionary, Genetic, Structural Characterization and Its Functional Implications for the Influenza A (H1N1) Infection Outbreak in India from 2009 to 2017. *Sci. Rep.* **2019**, *9*, 14690. [[CrossRef](#)]
19. Liu, T.; Wang, Y.; Tan, T.J.C.; Wu, N.C.; Brooke, C.B. The Evolutionary Potential of Influenza a Virus Hemagglutinin Is Highly Constrained by Epistatic Interactions with Neuraminidase. *Cell Host Microbe* **2022**, *30*, 1363–1369.e4. [[CrossRef](#)]
20. Caffrey, M.; Lavie, A. Ph-Dependent Mechanisms of Influenza Infection Mediated by Hemagglutinin. *Front. Mol. Biosci.* **2021**, *8*, 777095. [[CrossRef](#)]
21. Shu, Y.; McCauley, J. GISAID: Global Initiative on Sharing All Influenza Data—From Vision to Reality. *Eurosurveillance* **2017**, *22*, 30494. [[CrossRef](#)] [[PubMed](#)]
22. Villanueva, R.A.; Chen, Z. Ggplot2: Elegant Graphics for Data Analysis (2nd ed.). *Meas. Interdiscip. Res. Perspect.* **2019**, *17*, 160–167. [[CrossRef](#)]
23. Zhou, Z.J.; Qiu, Y.; Pu, Y.; Huang, X.; Ge, X.Y. Bioaider: An Efficient Tool for Viral Genome Analysis and Its Application in Tracing SARS-CoV-2 Transmission. *Sustain. Cities Soc.* **2020**, *63*, 102466. [[CrossRef](#)] [[PubMed](#)]
24. Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)]
25. Tamura, K.; Stecher, G.; Kumar, S. Mega11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. [[CrossRef](#)]
26. Saitou, N.; Nei, M. The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [[CrossRef](#)]
27. Sagulenko, P.; Puller, V.; Neher, R.A. Treetime: Maximum-Likelihood Phylodynamic Analysis. *Virus Evol.* **2018**, *4*, vex042. [[CrossRef](#)]
28. Gupta, R.; Brunak, S. Prediction of Glycosylation across the Human Proteome and the Correlation to Protein Function. *Pac. Symp. Biocomput.* **2002**, *7*, 310–322.
29. Zhang, Z.; Li, S.; Zhu, X.; Hou, J.; Zhang, H.; Zhao, B.; Tian, Z. Increased Genetic Variation of A(H3N2) Virus from Influenza Surveillance at the End of the 2016/2017 Season for Shanghai Port, China. *Sci. Rep.* **2022**, *12*, 17089. [[CrossRef](#)]
30. Ou, J.; Zhu, L.J. Trackviewer: A Bioconductor Package for Interactive and Integrative Visualization of Multi-Omics Data. *Nat. Methods* **2019**, *16*, 453–454. [[CrossRef](#)]
31. Weaver, S.; Shank, S.D.; Spielman, S.J.; Li, M.; Muse, S.V.; Kosakovsky Pond, S.L. Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. *Mol. Biol. Evol.* **2018**, *35*, 773–777. [[CrossRef](#)] [[PubMed](#)]
32. Pond, S.L.K.; Poon, A.F.Y.; Frost, S.D.W. Estimating Selection Pressures on Alignments of Coding Sequences. In *The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing*, 2nd ed.; Vandamme, A.-M., Salemi, M., Lemey, P., Eds.; Cambridge University Press: Cambridge, UK, 2009; pp. 419–490.

33. Kosakovsky Pond, S.L.; Frost, S.D.W. Not So Different after All: A Comparison of Methods for Detecting Amino Acid Sites under Selection. *Mol. Biol. Evol.* **2005**, *22*, 1208–1222. [[CrossRef](#)] [[PubMed](#)]
34. Murrell, B.; Wertheim, J.O.; Moola, S.; Weighill, T.; Scheffler, K.; Kosakovsky Pond, S.L. Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLoS Genet.* **2012**, *8*, e1002764. [[CrossRef](#)] [[PubMed](#)]
35. Murrell, B.; Moola, S.; Mabona, A.; Weighill, T.; Sheward, D.; Kosakovsky Pond, S.L.; Scheffler, K. Fubar: A Fast, Unconstrained Bayesian Approximation for Inferring Selection. *Mol. Biol. Evol.* **2013**, *30*, 1196–1205. [[CrossRef](#)] [[PubMed](#)]
36. Smith, G.J.D.; Bahl, J.; Vijaykrishna, D.; Zhang, J.; Poon, L.L.M.; Chen, H.; Webster, R.G.; Peiris, J.S.M.; Guan, Y. Dating the Emergence of Pandemic Influenza Viruses. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11709–11712. [[CrossRef](#)]
37. Virlogeux, V.; Feng, L.; Tsang, T.K.; Jiang, H.; Fang, V.J.; Qin, Y.; Wu, P.; Wang, X.; Zheng, J.; Lau, E.H.Y.; et al. Evaluation of Animal-to-Human and Human-to-Human Transmission of Influenza A (H7N9) Virus in China, 2013–2015. *Sci. Rep.* **2018**, *8*, 552. [[CrossRef](#)]
38. Skowronski, D.M.; Chambers, C.; Gustafson, R.; Purych, D.B.; Tang, P.; Bastien, N.; Krajden, M.; Li, Y. Avian Influenza A(H7N9) Virus Infection in 2 Travelers Returning from China to Canada, January 2015. *Emerg. Infect. Dis.* **2016**, *22*, 71–74. [[CrossRef](#)]
39. Su, S.; Gu, M.; Liu, D.; Cui, J.; Gao, G.F.; Zhou, J.; Liu, X. Epidemiology, Evolution, and Pathogenesis of H7N9 Influenza Viruses in Five Epidemic Waves since 2013 in China. *Trends Microbiol.* **2017**, *25*, 713–728. [[CrossRef](#)]
40. Yin, Y.; Liu, Y.; Fen, J.; Liu, K.; Qin, T.; Chen, S.; Peng, D.; Liu, X. Characterization of an H7N9 Influenza Virus Isolated from Camels in Inner Mongolia, China. *Microbiol. Spectr.* **2023**, *11*, e01798-22. [[CrossRef](#)]
41. Wu, N.C.; Wilson, I.A. Influenza Hemagglutinin Structures and Antibody Recognition. *Cold Spring Harb. Perspect. Med.* **2020**, *10*, a038778. [[CrossRef](#)]
42. Altman, M.O.; Angel, M.; Košík, I.; Trovão, N.S.; Zost, S.J.; Gibbs, J.S.; Casalino, L.; Amaro, R.E.; Hensley, S.E.; Nelson, M.I.; et al. Human Influenza A Virus Hemagglutinin Glycan Evolution Follows a Temporal Pattern to a Glycan Limit. *mBio* **2019**, *10*, e00204-19. [[CrossRef](#)]
43. Henry Dunand, C.J.; Leon, P.E.; Huang, M.; Choi, A.; Chromikova, V.; Ho, I.Y.; Tan, G.S.; Cruz, J.; Hirsh, A.; Zheng, N.Y.; et al. Both Neutralizing and Non-Neutralizing Human H7N9 Influenza Vaccine-Induced Monoclonal Antibodies Confer Protection. *Cell Host Microbe* **2016**, *19*, 800–813. [[CrossRef](#)]
44. Sun, X.; Jayaraman, A.; ManiPrasad, P.; Raman, R.; Houser, K.V.; Pappas, C.; Zeng, H.; Sasisekharan, R.; Katz, J.M.; Tumpey, T.M. N-Linked Glycosylation of the Hemagglutinin Protein Influences Virulence and Antigenicity of the 1918 Pandemic and Seasonal H1N1 Influenza A Viruses. *J. Virol.* **2013**, *87*, 8756–8766. [[CrossRef](#)] [[PubMed](#)]
45. Wu, N.C.; Zost, S.J.; Thompson, A.J.; Oyen, D.; Nycholat, C.M.; McBride, R.; Paulson, J.C.; Hensley, S.E.; Wilson, I.A. A Structural Explanation for the Low Effectiveness of the Seasonal Influenza H3N2 Vaccine. *PLoS Pathog.* **2017**, *13*, e1006682. [[CrossRef](#)] [[PubMed](#)]
46. Moody, M.A.; Zhang, R.; Walter, E.B.; Woods, C.W.; Ginsburg, G.S.; McClain, M.T.; Denny, T.N.; Chen, X.; Munshaw, S.; Marshall, D.J.; et al. H3n2 Influenza Infection Elicits More Cross-Reactive and Less Clonally Expanded Anti-Hemagglutinin Antibodies Than Influenza Vaccination. *PLoS ONE* **2011**, *6*, e25797. [[CrossRef](#)]
47. Yang, H.; Carney, P.J.; Chang, J.C.; Guo, Z.; Stevens, J. Structural and Molecular Characterization of the Hemagglutinin from the Fifth-Epidemic-Wave A(H7N9) Influenza Viruses. *J. Virol.* **2018**, *92*, e00375-18. [[CrossRef](#)]
48. Daniels, R.; Kurowski, B.; Johnson, A.E.; Hebert, D.N. N-Linked Glycans Direct the Cotranslational Folding Pathway of Influenza Hemagglutinin. *Mol. Cell* **2003**, *11*, 79–90. [[CrossRef](#)] [[PubMed](#)]
49. Makkoch, J.; Suwannakarn, K.; Payungporn, S.; Prachayangprecha, S.; Cheiocharnsin, T.; Linsuwanon, P.; Theamboonlers, A.; Poovorawan, Y. Whole Genome Characterization, Phylogenetic and Genome Signature Analysis of Human Pandemic H1N1 Virus in Thailand, 2009–2012. *PLoS ONE* **2012**, *7*, e51275. [[CrossRef](#)]
50. Bhattacharjee, U.; Chakrabarti, A.K.; Kanungo, S.; Dutta, S. Evolutionary Dynamics of Influenza A/H1N1 Virus Circulating in India from 2011 to 2021. *Infect. Genet. Evol.* **2023**, *110*, 105424. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.