Article

Genome-Wide Reassortment Analysis of Influenza A H7N9 Viruses Circulating in China during 2013–2019

Dongchang He [1,†], Xiyue Wang [1,†], Huiguang Wu [1,†], Xiaoquan Wang [1,2,3], Yayao Yan [1], Yang Li [1], Tiansong Zhan [1], Xiaoli Hao [1], Jiao Hu [1,2,3], Shunlin Hu [1,2,3], Xiaowen Liu [1,2,3], Chan Ding [2,4], Shuo Su [5], Min Gu [1,2,3,*] and Xiufan Liu [1,2,3,*]

1   Animal Infectious Disease Laboratory, College of Veterinary Medicine, Yangzhou University, Yangzhou 225009, China; virolog@outlook.com (D.H.); bioyusy@outlook.com (X.W.); hgwu80@163.com (H.W.); wxq@yzu.edu.cn (X.W.); bioyyy@outlook.com (Y.Y.); liyang1631@163.com (Y.L.); biocreater@outlook.com (T.Z.); xlhao@yzu.edu.cn (X.H.); hujiao@yzu.edu.cn (J.H.); slhu@yzu.edu.cn (S.H.); xwliu@yzu.edu.cn (X.L.)
2   Jiangsu Co-innovation Center for Prevention and Control of Important Animal Infectious Diseases and Zoonosis, Yangzhou University, Yangzhou 225009, China; shoveldeen@shvri.ac.cn
3   Jiangsu Key Laboratory of Zoonosis, Yangzhou University, Yangzhou 225009, China
4   Department of Avian Diseases, Shanghai Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Shanghai 200241, China
5   College of Veterinary Medicine, Nanjing Agricultural University, Nanjing 210095, China; shuosu@njau.edu.cn
*   Correspondence: gumin@yzu.edu.cn (M.G.); xfliu@yzu.edu.cn (X.L.)
†   These authors have contributed equally to this work.

Abstract: Reassortment with the H9N2 virus gave rise to the zoonotic H7N9 avian influenza virus (AIV), which caused more than five outbreak waves in humans, with high mortality. The frequent exchange of genomic segments between H7N9 and H9N2 has been well-documented. However, the reassortment patterns have not been described and are not yet fully understood. Here, we used phylogenetic analyses to investigate the patterns of intersubtype and intrasubtype/intralineage reassortment across the eight viral segments. The H7N9 virus and its progeny frequently exchanged internal genes with the H9N2 virus but rarely with the other AIV subtypes. Before beginning the intrasubtype/intralineage reassortment analyses, five Yangtze River Delta (YRD A-E) and two Pearl River Delta (PRD A-B) clusters were divided according to the HA gene phylogeny. The seven reset segment genes were also nomenclatured consistently. As revealed by the tanglegram results, high intralineage reassortment rates were determined in waves 2–3 and 5. Additionally, the clusters of PB2 c05 and M c02 were the most dominant in wave 5, which could have contributed to the onset of the largest H7N9 outbreak in 2016–2017. Meanwhile, a portion of the YRD-C cluster (HP H7N9) inherited their PB2, PA, and M segments from the co-circulating YRD-E (LP H7N9) cluster during wave 5. Untanglegram results revealed that the reassortment rate between HA and NA was lower than HA with any of the other six segments. A multidimensional scaling plot revealed a robust genetic linkage between the PB2 and PA genes, indicating that they may share a co-evolutionary history. Furthermore, we observed relatively more robust positive selection pressure on HA, NA, M2, and NS1 proteins. Our findings demonstrate that frequent reassortment, particular reassorted patterns, and adaptive mutations shaped the H7N9 viral genetic diversity and evolution. Increased surveillance is required immediately to better understand the current state of the HP H7N9 AIV.

Keywords: avian influenza virus; H7N9; highly pathogenic; diversity; reassortment; tanglegram

1. Introduction

Influenza A virus carries enveloped genomes consisting of eight gene segments of single-stranded ribonucleic acid molecules. When two viruses co-infect the same cell, the progeny virions may contain heterogeneous segments from different genomic sources. This process is known as reassortment, which is crucial for viral evolution [1] and adaptation [2]. Reassortment creates epidemiologically significant variants that can influence virus

propagation, illness severity, antiviral resistance, and vaccine efficacy [3–5]. The combination of reassortment and genomic mutations increases the viral diversity while also shaping their short-term evolution [6–8]. Particularly, when a reassortant contains novel antigens against the naive population and is able to achieve cross-host transmission from other species to humans (host jump), it meets the critical requirements for a pandemic with high potential. However, most reassortants are less fit than either parent and then cleaned by the purifying selection [9]. Occasionally, the reassortment leads to the formation of a gene combination with mutations under a particular set of selection pressures, resulting in improved fitness [9] and the potential of a pandemic. Almost all influenza viruses responsible for human pandemics are thought to have evolved through reassortment [2,3,5,10,11].

The reassorted avian influenza H7N9 virus emerged in 2013 with antigenic novelty, resulting in substantial human mortality (30–50%) and posing a significant threat to public health [12]. The six internal genes of the H7N9 virus were derived from at least two different H9N2 virus lineages, and the H7 and N9 genes came from wild birds [13–17]. Nonetheless, the H7N9 virus has continually undergone reassortment with other subtypes of AIVs since its emergence, such as seasonal influenza A virus H1N1, H3N2, and even influenza B virus [18–20], H5N6 [21,22], H6Ny [22,23], H9N2 [23–28], and even H7N9 itself [29]. As a result, multiple novel viruses were generated, including H7N2 [30], H7N3 [31], H7N4 [32], and H7N6 [33]. However, most of these reassortants were transient except for the progeny viruses with the H9N2 (especially the G57 or S genotype H9N2 [34]). The H9N2 virus facilitated the genesis, adaptation, and prevalence of the H7N9 virus [15,35]. Significantly, reassortment with H9N2 has contributed to the onset of the fifth epidemic wave, which is the largest H7N9 outbreak in humans [36]. Due to the ease of notifying the intersubtype reassortment with NA, the actual reassortment events and patterns involving the internal segments may largely remain unknown.

In China, two lineages have been identified and named based on the outbreak sources: the Yangtze River Delta (YRD) and the Pearl River Delta (PRD). The YRD lineage was found to be responsible for the majority of H7N9 outbreaks [17]. Based on this existing nomenclature system, we further divided the H7N9 into eight genotypes. Then, we performed systemic genomic analyses on the eight segments of the H7N9 virus from wave 1 to wave 7. Our study enabled us to estimate the frequency of reassortment and particular patterns among the H7N9 virus. Therefore, we conclude that reassortment and mutation co-shaped the evolution and epidemiology of the H7N9 virus. Monitoring reassortment and mutations in H7N9 virus predominance in chicken flocks is critical for preventing the virus from spreading to humans.

2. Materials and Methods

2.1. Sequences Collection

We downloaded the H5, H6, and H9 subtypes of avian influenza viruses circulating in Asia from 2010 to 2019 from the online databases of the *Influenza Virus Resource*, *NCBI Influenza Research Database*, and *Global Initiative on Sharing All Influenza Data* (accessed on 1 September 2019). We downsampled the collected sequences with a 0.99 threshold by CD-HIT (http://weizhong-lab.ucsd.edu/cdhit-web-server/cgi-bin/index.cgi, accessed on 1 October 2019) to minimize the computation resource consumption. Then, sequences were aligned using MAFFT (v7.453) [37] and trimmed to keep the coding region. As a result, polymerase basic protein 2 (PB2) ($n = 1125$), polymerase basic protein 1 (PB1) ($n = 1118$), polymerase acidic protein (PA) ($n = 1069$), matrix protein (M) ($n = 883$), nucleoprotein (NP) ($n = 1000$), and non-structural protein (NS) ($n = 884$) gene sequences were obtained. For the genomic sequences of H7N9, the cd99 and cd999 datasets were adopted from our previous study [38].

2.2. Phylogenetic Analyses and Diversity Estimates

We used Bayesian evolutionary analysis sampling trees (BEAST, v1.10.4) [39] program to interpret the (cd99 dataset) phylogenetic tree and evolutionary rate (ucld.mean) of eight

segments. The substitution model was determined by ModelFinder in PhyloSuite [40] according to the optimality Bayesian information criterion (BIC). The uncorrelated relaxed molecular clock model was used, and the Bayesian SkyGrid coalescent model was set as the tree prior [41]. Bayesian Markov chain Monte Carlo (MCMC) chains were set to 200 million generations with samples for every 20,000 steps to create the XML files. Each XML file was executed at least three times. Subsequently, the log files were assessed by Tracer (v1.7.1), and the effective sample size (ESS) values greater than 200 were accepted. Lastly, the maximum clade credibility (MCC) trees with median node heights were generated after the first 10% burn-in by TreeAnnotator (v1.10.4). The HA and NA MCC phylogenetic trees were adopted from our previous study [38]. Meanwhile, we performed the diversity test with the H7N9 cd999 dataset by the MEGA X (v10.2) to estimate the nucleotide diversity.

2.3. Phylogenetic Clusters Classification

Phylogenetic clusters were firstly classified based on the MCC trees of eight segments. Phylopart (v2.1) is a methodology for large-scale phylogeny partition [42]. We used it to define monophyletic clades with posterior probabilities of $\geq$90% and a median genetic distance threshold for clusters of 20% [15]. The assignment to these clusters was subsequently manually merged based on cluster tips containing $\leq$ 3 and failed to cluster. Then, each cluster was assigned a unique number based on the increasing tree topology. The remaining 7 MCC trees were assigned to clusters using the same processes. The H7N9 genotypes were classified using the HA clusters result and a previous investigation [17]. Phylogenetic tree and heatmap were plotted and visualized using ggtree [43,44].

2.4. Reassortment Detections

2.4.1. Intersubtype Reassortment

To investigate intersubtype reassortments across the six internal segments (PB2, PB1, PA, NP, M, and NS), we integrated the main subtypes (H5, H6, H7, and H9) of avian influenza viruses that circulated from 2010 to 2019. Multiple sequence alignments were performed with MAFFT [37]. Then, phylogenetic tree construction was performed by Mrbayes (v3.2.7a) [45]. The GTR+G+I model parameters were incorporated into the nexus file. Then, the nexus files were implemented in MrBayes by running 20 million MCMC chains with a sampling frequency of 5000. The phylogenetic trees were visualized and plotted using ggtree and ggtreeExtra [43,44].

2.4.2. Entanglegram and Untanglement

Tanglegram is a visual method to draw an association between two phylogenetic trees with identical tip labels. Tanglegram is also known as the incongruence tree. Theoretically, when the reassortment event is absent, the twines connect the tanglegram in a noncrossing or horizontally way [46]. To uncover the H7N9 AIV intrasubtype reassortment, we used the backronymed adaptable lightweight tree import code (Baltic) to generate incongruence visualization of eight H7N9 AIV MCC trees. The phylogenetic position of each strain was traced and colored according to the HA clusters across all segments. Entanglement was generated from an adapted script (https://github.com/dven42/phylogenetic-incongruence, accessed on 1 June 2021) and modified to visualize the positions of specific isolates for our study.

Untanglegram is a visual method to minimize crossings of hybridization networks between the tanglegram by rotating their branches around ancestral nodes [47]. There is an arrangement of their branches such that the association edges do not intersect if the topologies of the two phylogenies are entirely concordant [47]. The untanglegram phylogeny reveals the extent of reassortment between the HA and seven other segments (PB2, PB1, PA, NA, NP, and M). The identical tips were connected across the trees. Tips and connecting lines were colored according to the HA clusters. Untanglements were plotted with Dendextend using the step1side untangle method [48] by fixing the left tree (HA) and rotating the right tree (the other segments) to generate the minimization hybridization

tanglegram in R (v4.0.5). The bigger the untanglement value is, the worse the reassortment between the paired tree.

2.4.3. Quantified Investigation Reassortment Events

We investigated the reassortment events over the entire genome by combined approaches of Graph Incompatibility based Reassortment Finder (GiRaF, v1.02) [49] and Recombination Detection Program (RDP, v5.05) [50]. In brief, Mrbayes assessed multiple trees per segment for incompatible splits using the GiRaF. After twenty million generations and sampling every 5000 steps, a burn-in of 25% samples was established to obtain trees and subsequently used for input files. Reassortment events with a confidence level greater than 0.9 were deemed accurate. Meanwhile, we also investigated the reassortment events using the RDP from the concatenated segments [51]. The algorithms incorporate the RDP, GENECONV, BOOTSCAN, MAXCHI, CHIMAERA, SISCAN, and 3SEQ techniques. At least four of the seven detection methods with a *p*-value of 10^{-6} were acknowledged as recombination events. Only strains confirmed concurrently by GiRaF and RDP were considered putative reassortment events.

BEAST2/CoalRe (v0.05) [51] estimated intralineage reassortment networks between H7 and N9. Three runs of 200 million MCMC sampling steps were performed following the tutorial (https://taming-the-beast.org/tutorials/Reassortment-Tutorial/, accessed on 1 June 2021). The embed segment tree was drawn using icytree (https://icytree.org/, accessed on 1 October 2021) to depict the reassortment network between HA and NA.

2.5. Multidimensional Scaling

We used multidimensional scaling (MDS) to analyze each viral segment's tree-to-tree branch variation and investigate the overall level of cross-correlation between all segments in a two-dimensional space. Generally, the last 500 phylogenetic trees for each segment were obtained from independently sampled trees in BEAST. The statistical tree-to-tree variation in branch lengths was calculated. Then, the MDS statistics analysis was performed in the R (v4.0.5) to determine the cross-correlation of all segments in two-dimensional space. Only the first two scaling dimensions were plotted using the ggplot2 [52] package for visualization. Theoretically, the viral segments that share similar evolutionary histories occupy similar locations in the two-dimensional Euclidean space where the cloud of points should overlap. Clouds of points in the MDS plot indicate phylogenetic uncertainty based on 500 sampled trees. In contrast, segments are expected to exhibit uncorrelated features due to their unlinked evolutionary histories in response to reassortment. Scripts for the MDS calculation were obtained from Doctor Maude Jacquot et al. [4,53] and modified for this study.

2.6. Selection Pressure Analysis

We quantified the selection pressures acting on the 10 major protein-coding regions (M and NS, respectively, encode at least two proteins) under different models. Before selection analysis, the best-fit model of nucleotide substitution was obtained from Modelfinder [54] based on Bayesian information criterion (BIC) for each segment. The maximum likelihood (ML) phylogenetic tree was interpreted using IQtree in PhyloSuite (v1.2.2). Subsequently, single-likelihood ancestor counting (SLAC) [55], fast unconstrained Bayesian approximation (FUBAR) [56], fixed-effects likelihood (FEL), and mixed-effects model of evolution (MEME) [57] were used to infer sites under episodic or pervasive natural selection on each coding protein. Finally, we recommend the significance levels for FEL ($p < 0.1$), SLAC ($p < 0.05$), MEME ($p < 0.05$), and FUBAR (posterior probability > 0.9). All methods were implemented in the HyPhy (v2.5.2) [58] on a high-performance computing cluster. The selection results of HA and NA were adopted from our previous study [38].

3. Results

3.1. The Surface Glycoproteins Have Faster Evolutionary Rates but Less Genetic Diversity

Nucleotide analyses revealed that the internal segments have higher diversity than the surface glycoproteins (Figure 1). PB2 (3.86×10^{-3}) has the highest genetic diversity among all segments, exhibiting a highly heterogeneous genome. The HA gene (1.83×10^{-3}) has slightly more diversity than the NA gene (1.80×10^{-3}), which is smaller than internal genes.
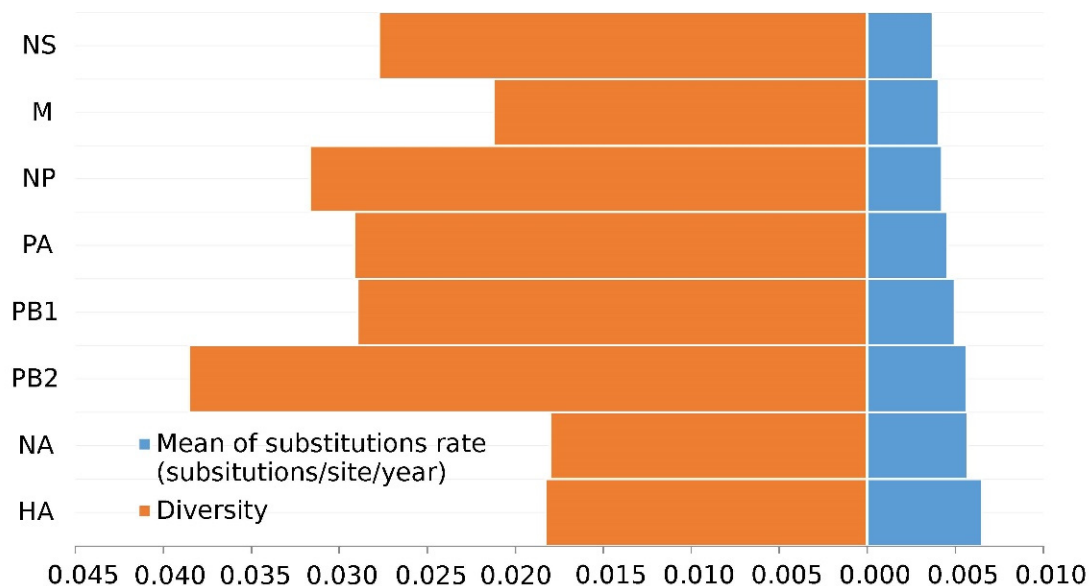

Figure 1. Diversity and evolutionary rates of H7N9 eight gene segments.

3.2. PB2 c05 and M c08 Were the Predominant Clusters in Wave 5

To study H7N9 reassortment, we investigated the clusters of each segment for H7N9 viruses recovered from all hosts in China from 2013 to 2019 (Figure 2). Segment cluster numberings were designated, including HA (c1-c8), NA (c1-c8), PB2 (c1-c5), PB1 (c1-c9), PA (c1-c7), NP (c1-c6), M (c1-c8), and NS (c1-c5) (Figures S1–S7 and Table S1). Cluster 0 (c0) represented the phylogenetic tips that could not be assigned to a cluster. The same number in different segments is not correlative. Eight genotypes were allocated based on HA clusters: the early genotype cluster was designated as W1; the Pearl River Delta (PRD) lineage was divided into two genotypes: PRD-A and PRD-B (Figure 2); and the Yangtze River Delta (YRD) lineage designated as five genotypes: YRD-A, YRD-B, YRD-C, YRD-D, and YRD-E. LP H7N9 viruses of the YRD lineage discovered in wave 5 were mainly assigned within the YRD-E cluster, which was the dominant cluster. In the YRD-E cluster, the dominating segment clusters were c05 PB2 (79.03%) and c08 M (78.23%) (Figure 3). Viruses simultaneously containing PB2 c05 and M c08 were determined in 62.10% (77/124 in the cd99 dataset) of the YRD-E cluster. The circulating HP H7N9 AIVs were assigned to the YRD-C cluster. In the HP H7N9 virus found in waves 6 and 7, the cluster of M c03 was replaced by the c08. Reassortment has a more significant impact on the diversity of viral genotypes.
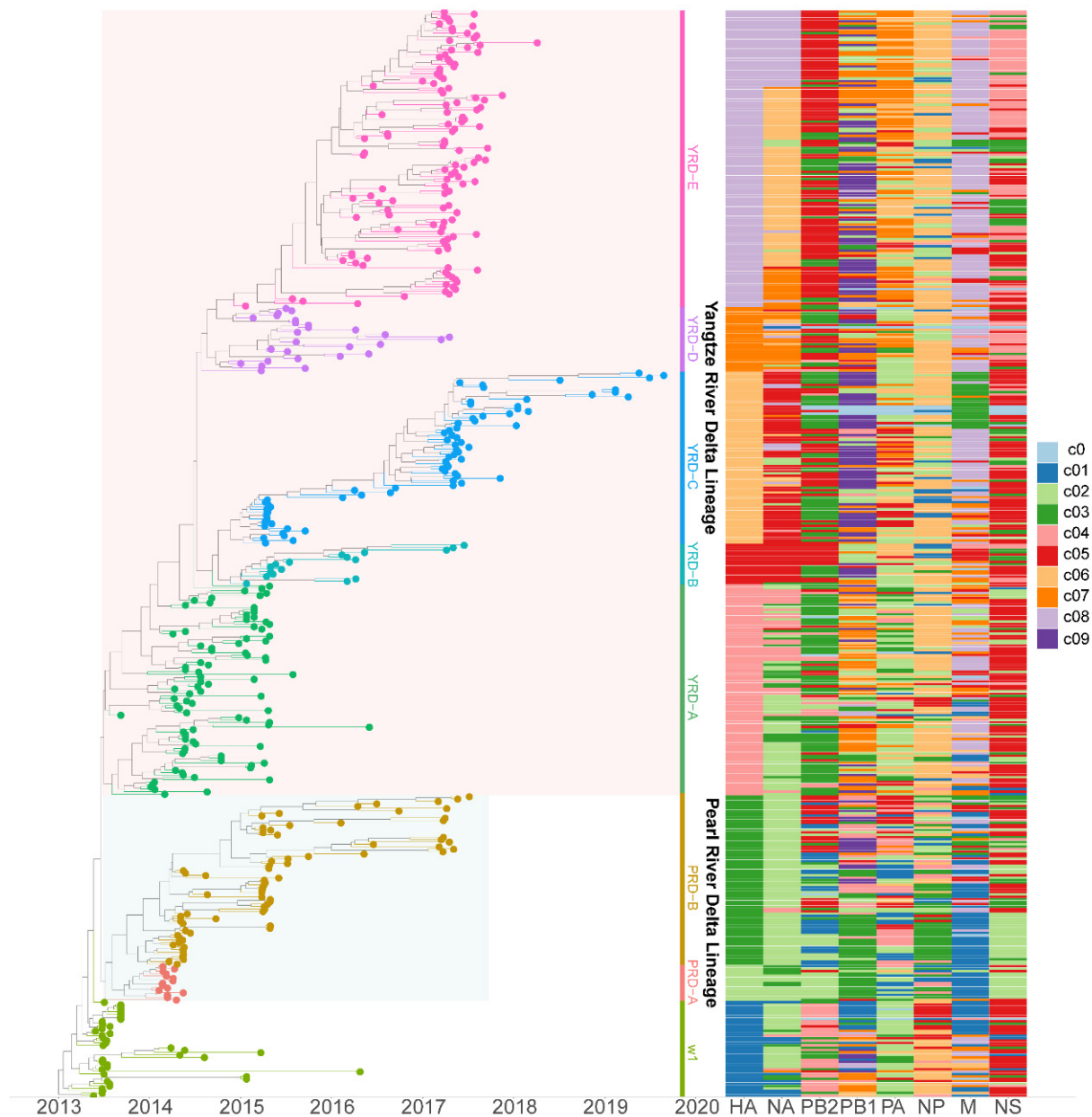
Figure 2. The genotypes heatmap of H7N9 viruses. Each segment's cluster assignment is based on MCC trees with a median branch length distance threshold of 0.20 identified using PhyloPart. The left panel is the time-scaled HA tree with two lineages (Yangtze River Delta and Pearl River Delta), and tips are colored based on clusters. The right panel displays the heatmap of each segment cluster. Shared colors and numbers indicate sequences of the same segment assigned to the same cluster. The same color and number in different segments are not correlative. Detailed clusters are available in the Supplementary Table S1.
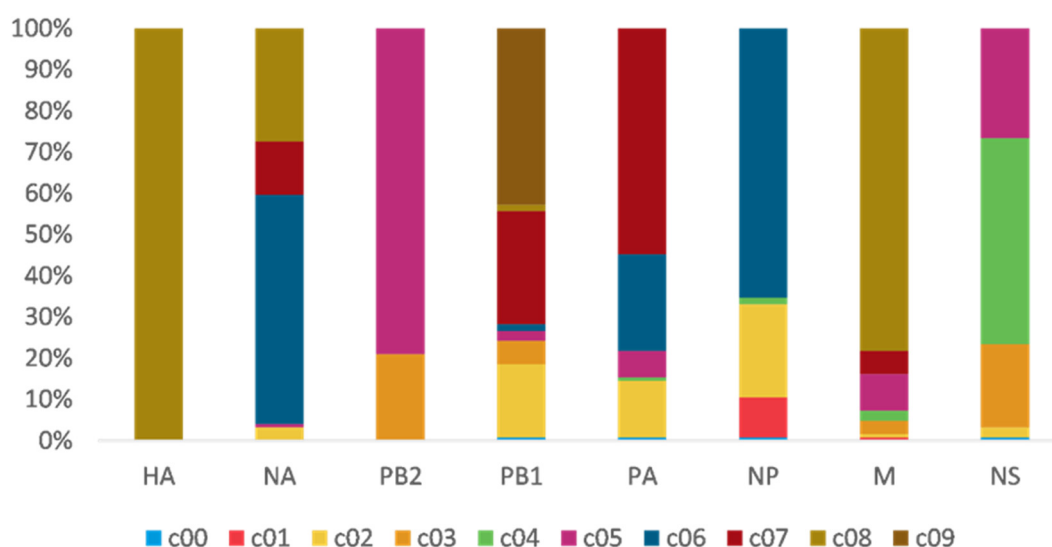
Figure 3. The gene clusters proportion of YRD-E. The same color and number in different segments are not correlative. PB2 c05 (79.03%) and M c08 (78.23%) are dominant clusters in the epidemic wave 5. Viruses have PB2 c05, and M c08 account for 62.10%.

3.3. Intensive Intersubtype Reassortments between H7N9 and H9N2

Phylogenetic tree analyses evaluated intersubtype reassortments between cocirculating subtypes of H5, H6, and H9. Surprisingly, the topology of H7N9 and H9N2 was clustered together in many branches on each tree of internal segments (Figure 4). H7N9 viruses were reassorted in rare cases with other subtypes of AIVs instead of H9N2. Duck-origin viruses in particular have a more complicated genetic background, with internal segments derived primarily from wild bird viruses, such as H5N6, H6N1, H6N2, H6N6, H7N2, H7N3, and H7N7. For example, the M gene of A/chicken/Guangdong/Q1/2016, the early-raised HP H7N9 obtained on 20 June 2016, is closely clustered with the M gene of the H5N6 subtype (A/duck/Yunnan/07.15 DQNPH129/2015, EPI668788). However, all these intersubtype reassortment events were predominantly within the H9N2 subtype (G57 or genotype S).

3.4. Dynamic and Intricate Intrasubtype Reassortment

Baltic was used to visualize inconsistencies between the MCC phylogenetic trees (tanglegram). An abundance of intralineage reassortment events was observed from the tanglegram among the 7 waves (Figure 5). Internal segments in the YRD-E (LP in wave 5) cluster were highly diverse during the wave 5 outbreak. They usually came from 2–3 main clusters on each MCC tree topology (Figure S8). Notably, part of HP H7N9 viruses in the YRD-C inherited their PB2, PA, and M segments from the YRD-E cluster (Figure S9), suggesting the LP H7N9 contributed to the genesis of HP H7N9 and frequent intralineage reassortments. There were also intralineage reassortments between the YRD and PRD lineages. For instance, some YRD-C HP H7N9 viruses obtained the NA gene from the cocirculating PRD-B cluster. Intralineage reassortment events between the YRD and PRD lineages were also commonly and consistently identified throughout all H7N9 epidemic waves by BEAST2/CoalRe (Figure 6). Especially, many intralineage reassortment instances were also determined in waves 2–3 and 5. Although the H7N9 population was severely decreased at the time, the HP H7N9 virus isolated in waves 6 and 7 continued to reassort. For instance, the M gene of A/chicken/Shanxi/SX0256/2019 came from c07 rather than the c03. The M c03 has been the dominant cluster in the previous prevalent viruses since 2018 (Figure 2). Of note, none of the genotypes was predominant in any of the seven epidemic waves. Our results demonstrated that China's H7N9 virus has a dynamic and intricate intrasubtype reassortment pattern.
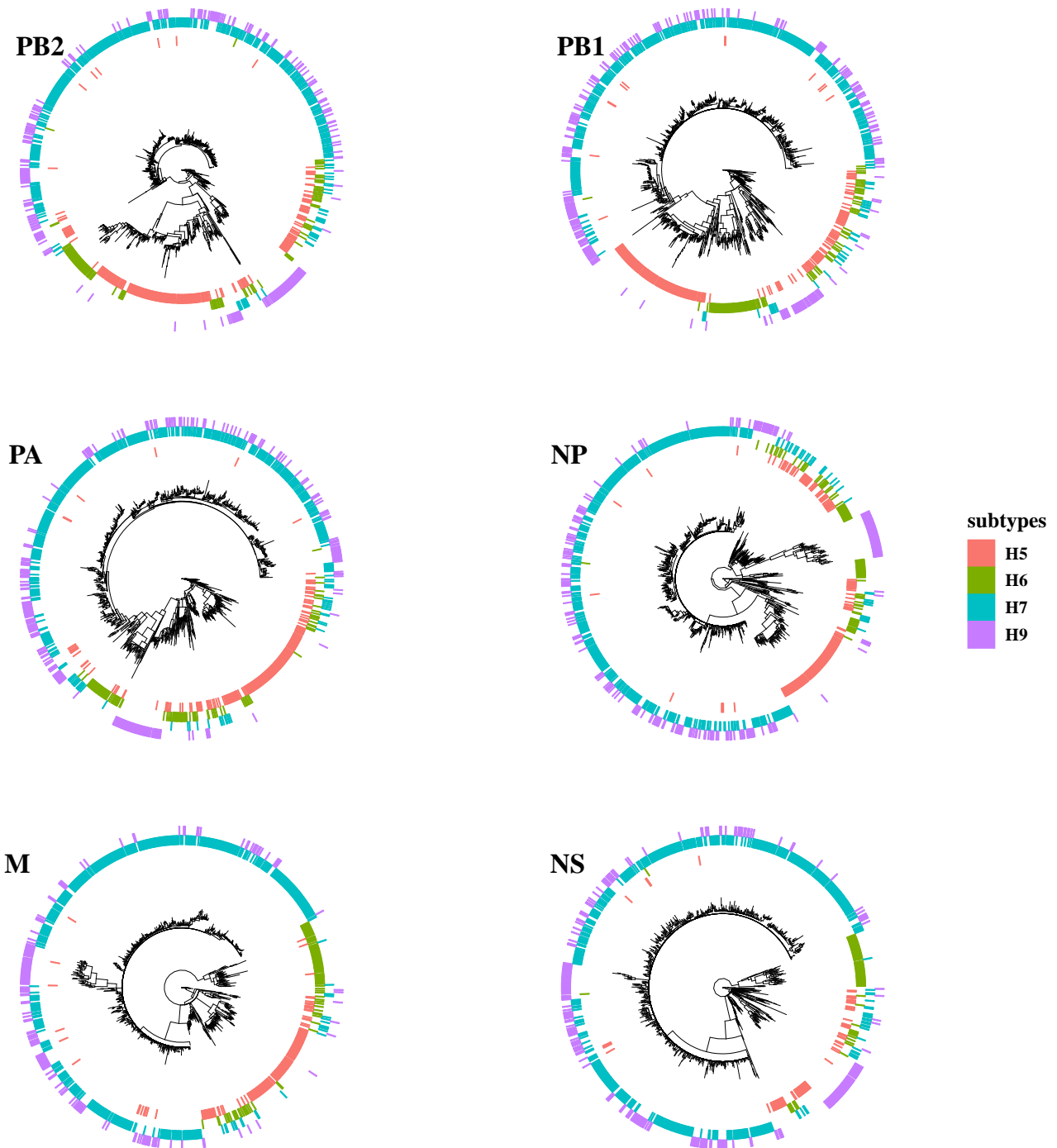
Figure 4. Intersubtype reassortment between H7N9 and H5, H6, and H9 subtypes of avian influenza viruses.
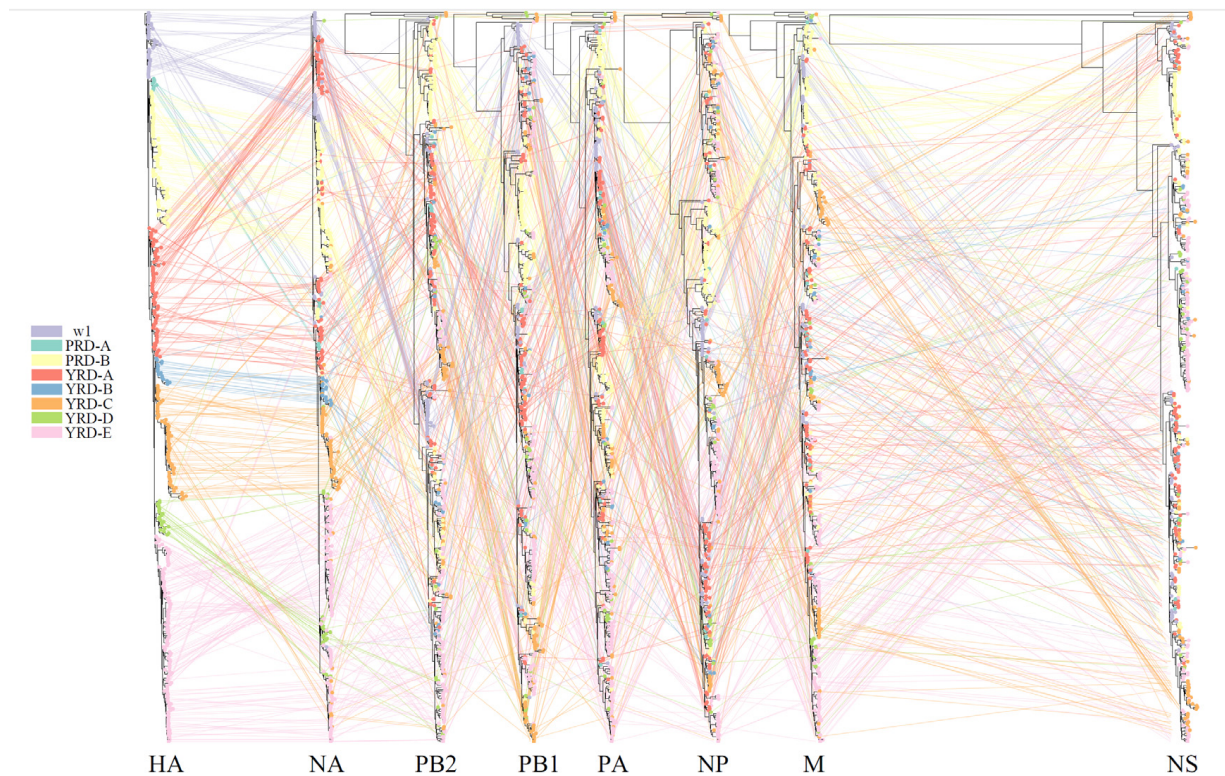
Figure 5. Phylogenetic incongruence analysis. MCC trees for the HA segment and all internal genes NA, PB2, PB1, PA, NP, and M from equivalent strains connect across the trees. Tips and connecting lines are colored according to the HA clusters.
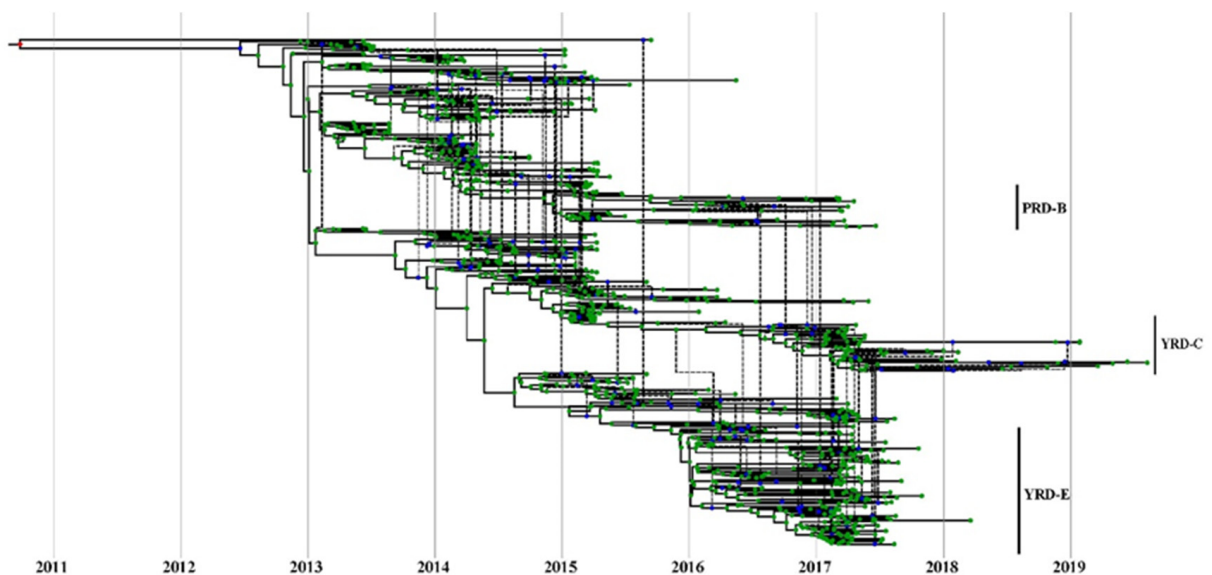
Figure 6. Estimate of MCC reassortment network between HA and NA genes of H7N9 viruses. Icytree visualizes the network as a base tree connected by dotted branches, indicating a reassortment event. The green dots indicate the tips of the tree.

3.5. HA and NA Have the Lowest Reassortment Rate

To measure the severity of reassortment over the entire genome, we untangled the tanglegram between HA and the paired trees shown as untanglement values (the bigger, the worse. Figure 7). The unentanglement values between HA and paired segments (NA,

PB2, PB1, PA, NP, M, and NS) were 0.0381, 0.1997, 0.2340, 0.0591, 0.3234, 0.0672, and 0.4894, respectively. The lowest untanglement value (0.0381) was found between HA and NA, whereas the highest value was found between HA and NS (0.4894).

Figure 7. Evolutionary relationships of each gene segment with HA. Incongruence phylogenetic analysis shows interclade reassortment between the HA segment and seven additional genes (NA, PB2, PB1, PA, NP, and M). Equivalent strains connect across the trees. Tips an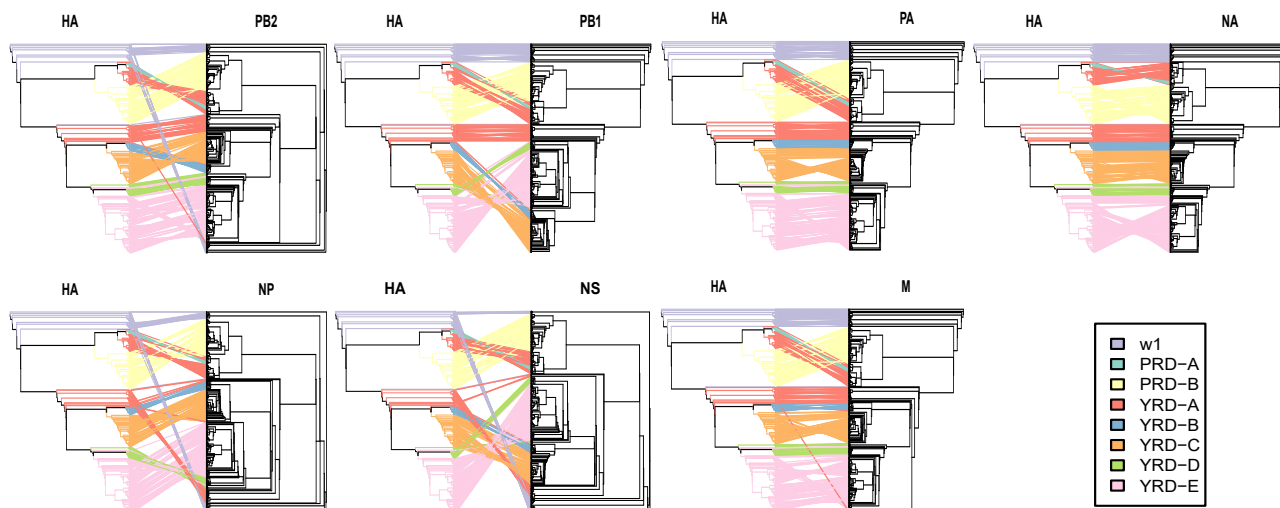d connecting lines are colored according to the HA clusters. Unentanglement is used to minimize crossings of the hybridization network between the paired trees. The values are determined by the degree of inter-segment reassortment.

3.6. High Reassortment Rate in Waves 3 and 5

GiRaF and RDP analyses found that 217 of 454 strains H7N9 (cd99 dataset) were involved with reassortment events (Table S2). Following that, we computed the reassortment rates in different waves. High reassortment rates were discovered in waves 3 (0.76) and 5 (0.79). Wave 1 is notable for having the lowest reassortment rate (0.10).

3.7. PB2 and PA Shared Coevolutionary History

MDS allows for the two-dimensional depiction of the total degree of cross-correlation between all segments, with overlap between observations indicating shared evolutionary history (i.e., linkage) between segments. In comparison, segments that split up due to reassortment are expected to occupy separate plot areas. We found that, except for the PB2 and PA segments, the rest of the H7N9 segments were very distinct (Figure 8). The segments PB1, M, NP, and NS did not show any association, indicating that they did not have any coevolutionary relationship. The PB2 and PA genes almost completely overlapped, indicating they shared a strong evolutionary history (i.e., linkage). In comparison, the capsid proteins of HA and NA only had a weak association.
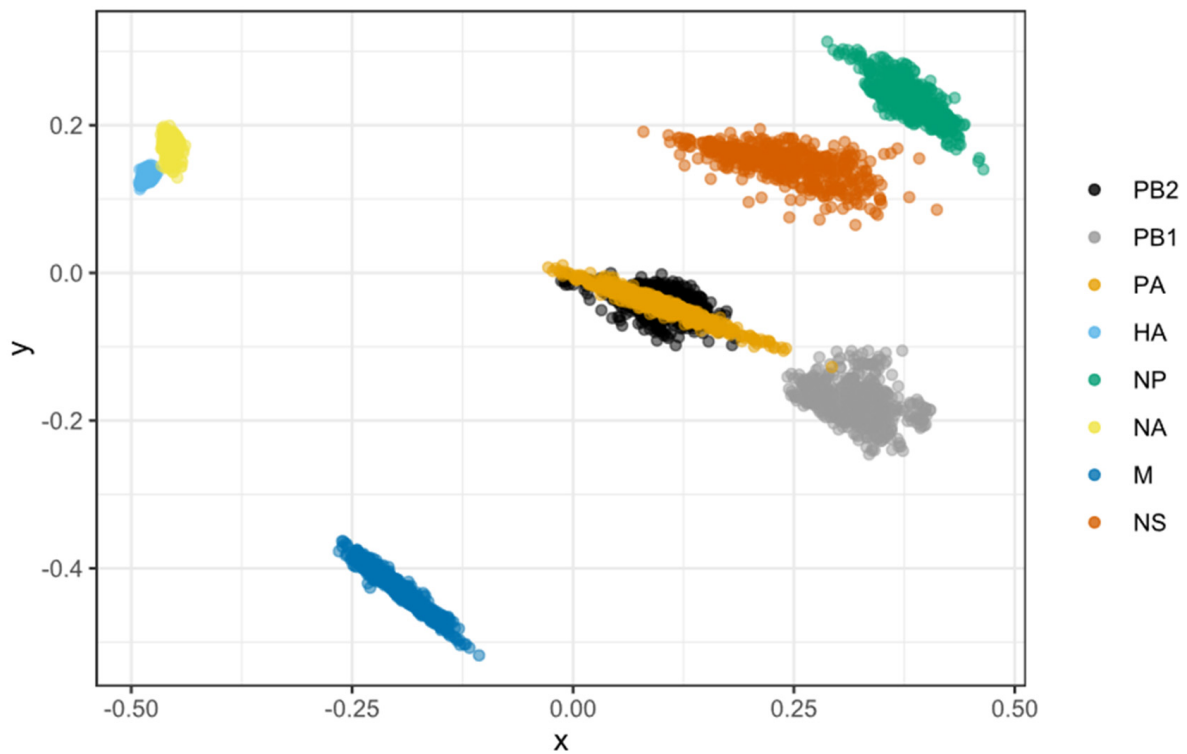
Figure 8. Correlations in time to the most recent common ancestor (tMRCA) among H7N9 viral segments as depicted by a multidimensional scaling (MDS) graphic. MDS enables a two-dimensional representation of the total level of cross-correlation between all segments. Cloud points represent phylogenetic uncertainty based on 500 trees for each segment sampled in the program of BEAST, with pairwise comparisons to other segments limited to viruses sampled in the same year. In the absence of reassortment, segments are likely to have highly correlated TMRCAs due to shared evolutionary history showing as overlapping dots. Reversely, segments split up by reassortment are predicted to inhabit various plot regions. Only the first two scaling dimensions are visible.

3.8. High Selection Pressure Acting on HA, NA, NS1, and M2

The natural selection acting on all coding regions was estimated using the dN/dS ratio. Four site-level detection methods (FEL, SLAC, FUBAR, and MEME) were used to assess positive and negative selection codons. M2 (0.5398), NS1 (0.3087), HA (0.2656), and NA (0.2946) underwent more substantial selection pressure than other segments, whereas M1 protein (0.0752) was the lowest, according to the overall dN/dS values estimated using SLAC (Table 1). Except for M2, NS1, and NEP, the other internal proteins were unaffected by higher purifying selectivity. MEME was used to find more positive selection sites by identifying fixed and sporadic positively chosen codons. As a result, HA (14 codons) and NA (13 codons) had more sites under positive selection pressure than other coding regions. PB2 (10 codons), NS1(10 codons), PB1 (7 codons), and PA (6 codons) also detected many positive selection sites.

Table 1. Selection pressures and positively and negatively selected codons of coding regions of H7N9 viruses circulating between 2013 and 2019.

Coding Region	No. of Codons	dN/dS	No. of Selected Sites (% of Codons)				Negatively Selected (SLAC)
			Positively Selected				
			SLAC [a]	FUBAR [b]	FEL [c]	MEME [d]	
PB2	759	0.1200	4	4	4	10	670 (88.16%)
PB1	757	0.0977	5	6	6	7	657 (75.60%)
PA	716	0.1137	7	4	8	6	573 (79.92%)
HA [38]	564	0.2656	16	13	16	14	374 (66.19%)
NP	498	0.1026	2	1	4	3	413 (82.77%)
NA [38]	465	0.2946	13	8	15	13	280 (60.09%)
M1	252	0.0752	0	1	1	1	190 (75.10%)
M2	97	0.5398	6	7	8	3	26 (26.53%)
NS1	217	0.3087	13	6	14	10	115 (52.75%)
NEP	121	0.2766	0	1	1	2	55 (45.08%)

[a] *p*-value < 0.05; [b] posterior probability of ≥0.9; [c] *p*-value < 0.1; [d] *p*-value < 0.05

4. Discussion

We estimated the inter- and intra- reassortment, reassortment patterns, and adaptative evolution of H7N9 viruses from waves 1–7. Our study found that H7N9 viruses have undergone considerable changes by reassortment, which had a significant impact on the H7N9 genomic composition since its emergence. Notably, the internal genes presented more diversified features than surface genes, and even the surface genes had higher substitution rates. Numerous reassortment events were determined in our study, resulting in a high level of genetic diversity, especially among the internal genes. The dynamic and intricate reassortments may shape the epidemiology and genomic evolution of the H7N9 virus and contribute to its genetic diversity.

The H7N9 haemagglutinin gene evolved faster (6.51×10^{-3} substitutions/site/year) than the seasonal influenza virus, with an average rate of 3.41×10^{-3} [59]. This substitution rate is similar to the H5N1 under vaccination pressure (6.13×10^{-3}–8.87×10^{-3}) but faster than H5N1 without vaccination pressure [60]. According to recent studies, the nucleotide substitution rate of the H7N9 HA gene under vaccination pressure increased to 1.963×10^{-2} during 2018–2019 [61]. Considering that reassortment usually leads to a transient increase in the substitution rate [62], the frequency of reassortments may also contribute to the rapid evolution of H7N9. According to other research, external genes of the influenza virus usually have evolved at faster evolutionary rates than its internal genes [63,64], resulting in more genetic diversity in surface genes. However, the external genes of the H7N9 virus showed less genetic diversity than its internal segments. This contradiction implies that H7N9 internal genes were reassorted frequently rather than in a single event since its emergence.

In the study of intersubtype reassortments, we found that the H7N9 virus was primarily associated with the cocirculating H9N2 virus (genotype S or G57). In contrast, they were rarely reassorted with other AIV subtypes. The H9N2 and H7N9 viruses have been cocirculating in China since 2013. Therefore, numerous opportunities existed for their coinfections. Several investigations in the first five epidemic waves also confirmed the high prevalence of H7N9/H9N2 coinfection in chickens [15,25,27,65]. Influenza virus reassortment occurs with high frequency without segment mismatch [66]. However, most reassortment viruses produced from divergent parental strains are also frequently outcompeted by one or both parental viruses [2]. Specific reassortment genotypes may be inefficient in forming due to incompatibility between the heterologous RNA packaging signals [2,67]. Considering the high genetic similarity between the internal segments of H7N9 and H9N2 AIVs, it may be the biological basis for the high frequency of intersubtype reassortment in the absence of mismatches between their packaging signals. Heterogeneous genome packaging signals combinations other than H7N9/H9N2 viruses may be deleterious or decrease fitness under

natural conditions [68], forcing the reassorted progeny virus particles to be eliminated through purifying selection. Therefore, the high frequency of intersubtype reassortants between H7N9 and H9N2 AIVs could be expected.

The LP H7N9 virus, unlike the other LP AIV, may cause serious illness in humans and other mammalian species, demonstrating its exceptional fitness in mammalian hosts [69]. The fifth wave of 2016–2017 was the biggest epidemic to date, with nearly the same number of human cases (*n* = 758) as the sum of the previous four outbreaks [36]. Nonetheless, most human infections in wave 5 were caused by the LP H7N9 virus, which is primarily associated with YRD-E. Our investigation found that gene clusters PB2 c05 and M c08 have a large proportion of YRD-E cluster strains, which mainly circulated in wave 5 despite a highly diverse genetic background. Our previous studies showed that the reassortments with PB2 and M genes from genotype S H9N2 caused attenuated progeny of H5Nx and H7N9, resulting in optimizing viral fitness viruses in mice [70] and chicken [71]. Therefore, we assume that PB2 c05 and M c08 might optimize the viral fitness and contribute to the wave 5 outbreak. Aside from the PB2 and M, NP c06 and PA c07 have a certain ratio in YRD-E. It is also impossible to overlook their contribution to viral fitness.

In terms of intralineage reassortment, we found that the intralineage reassortment occurs considerably frequently rather than as a single event. Usually, intralineage reassortment is subjected to severe negative selection, which becomes more pronounced as the genetic distance between donor strains increases [72]. Due to the remarkable similarity in their internal genes and genomic packaging signals between the H9N2 and H7N9 viruses, the high compatibility between their internal genes makes the reassorts rise rather naturally. According to our findings, the YRD-E lineage contributed its PB2, PA, and M gene to the YRD-C lineage, dominated by the HP H7N9 viruses. The frequent reassortment leads to such a diversity of genotypes of H7N9. Similarly, Cui et al. documented 27 H7N9 genotypes within 3 months of H7N9 emergence [35]. Nonetheless, it should be highlighted that intralineage reassortment between pretty similar internal gene cassettes is likely to be underestimated. Using tanglegram to estimate the reassortment in phylogenies from high similarity sequences might underestimate or exaggerate reassortment events [46]. Combined with additional pieces of evidence [27–29], we confirmed that the intricate intralineage genetic reassortment of the H7N9 virus had occurred frequently since its introduction.

Usually, wild birds are the primary source of AIV reassortant [73]. However, we found that H7N9 reassortment occurred mainly in the chicken host rather than in wild birds. The most likely reason was that H7N9 and H9N2 were primarily cocirculated in poultry, while H7N9 only had sporadic spillover to wild birds [74]. The other reason is that national surveillance relied predominantly on passive reporting systems and less active and systemic surveillance in wild birds [75]. Therefore, the "sampling strategy" in the H7N9 surveillance may affect the reassortment interpretation.

Many intersubtype reassortments took place between HA and NA in waves 2–3 and 5, coinciding with the periods of extensive intrasubtype and intralineage reassortment. Theoretically, H7N2 and H9N9 should have a high probability of detection since the extensive interlineage reassortment between H7N9 and H9N2. However, the detection of the H7N2 and H9N9 virus subtypes was limited. Even the reassorted H9N9 virus in the laboratory has shown increased fitness features in poultry [76]. Although intersubtype reassortments between H7 and N9 were frequently detected in waves 2–3 and 5, the frequency of reassortment between HA and NA remained low compared to HA with other segments, demonstrating a robust functional balance between H7 and N9. Similar to other avian influenza viruses in poultry and wild bird populations, the NS gene encoding the non-structural protein has a more divergent phylogeny and high reassortment rates [73,77].

Reassortment in the influenza virus is not a random process, which is consistent with observations from other segmented viruses [4,67]. Except for the PB2 and PA segments, the other H7N9 segments did not share their evolutionary history segment. However, the HA-NA capsid proteins and the PB1-PB2/PA RNA polymerase complex proteins only had a weak association. The times to most recent common ancestor (tMRCA) were

consistent in PB2 and PA, indicating they were less amenable to reassortment than the other segments in history. This inherited linkage also implies physical or biochemical interactions between their encoded proteins since epistatic interactions result in tighter evolutionary connections. According to a recent study, reintroducing PB2 and PA with adaptive mutations from cocirculating H9N2 in 2015 resulted in a novel H7N9 genotype, which increased polymerase activity and became dominant in the fifth H7N9 virus epidemic [36]. Because the internal segments primarily arose from genotype S H9N2 [34] (similarly with G57 genotype H9N2 [36]), this linkage PB2-PA pattern most likely originated from H9N2. A previous study found that a coadapted PB1-PB2-HA gene pattern was established during influenza B interlineage reassortment, which has been demonstrated to be critical for whole-genome fitness [78]. As a result, we assume that the particular PB2-PA pattern might confer a specific advantage to the H7N9/H9N2-like gene cassette. Further studies are warranted.

In addition to M2, NS1, and NEP, high dN/dS values were found in the HA and NA proteins (0.5398, 0.3087, 0.2766, 0.2656, and 0.2946, respectively), which were generally in agreement with the previous study in H3N2 [46]. Surface proteins (HA, NA, and M2) are typically more sensitive to positive selection and evolve more rapidly than internal genes, especially when the influenza virus circulates in a naive population [63,64]. The reassortment events lead to a transient increase in the rate of amino acid replacements on the descendant phylogenetic branches [79]. Therefore, the extensive reassortment might increase the H7N9 evolutionary rate and result in rapid adaptive evolution at the molecular level.

In summary, our phylogenetic analyses revealed the comprehensive genetic evolution of the H7N9 viruses. The particular reassortment patterns indicate that the H9N2-original internal gene constellation has superior compatibility to the genesis and evolution of H7N9 over the other AIV subtypes. The H7N9 viruses were further diversified by frequent inter- and intra-lineage reassortment events with adaptative mutation, leading to successful H7N9 genotypes. The evolution and epidemiology of the H7N9 virus in China may be shaped by reassortment and adaptive mutations. Further research is needed to understand the dynamics of reassortment of the circulating H7N9 virus.

Conflicts of Interest: The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Steel, J.; Lowen, A.C. Influenza A virus reassortment. *Curr. Top. Microbiol. Immunol.* **2014**, *385*, 377–401. [CrossRef] [PubMed]
2. Lowen, A.C. Constraints, Drivers, and Implications of Influenza A Virus Reassortment. *Annu. Rev. Virol.* **2017**, *4*, 105–121. [CrossRef] [PubMed]
3. Maljkovic Berry, I.; Melendrez, M.C.; Li, T.; Hawksworth, A.W.; Brice, G.T.; Blair, P.J.; Halsey, E.S.; Williams, M.; Fernandez, S.; Yoon, I.-K.; et al. Frequency of influenza H3N2 intra-subtype reassortment: Attributes and implications of reassortant spread: Attributes and implications of reassortant spread. *BMC Biol.* **2016**, *14*, 117. [CrossRef] [PubMed]
4. Nomikou, K.; Hughes, J.; Wash, R.; Kellam, P.; Breard, E.; Zientara, S.; Palmarini, M.; Biek, R.; Mertens, P. Widespread Reassortment Shapes the Evolution and Epidemiology of Bluetongue Virus following European Invasion. *PLoS Pathog.* **2015**, *11*, e1005056. [CrossRef] [PubMed]
5. White, M.C.; Lowen, A.C. Implications of segment mismatch for influenza A virus evolution. *J. Gen. Virol.* **2018**, *99*, 3–16. [CrossRef] [PubMed]
6. Nelson, M.I.; Simonsen, L.; Viboud, C.; Miller, M.A.; Taylor, J.; George, K.S.; Griesemer, S.B.; Ghedin, E.; Ghedi, E.; Sengamalay, N.A.; et al. Stochastic processes are key determinants of short-term evolution in influenza a virus. *PLoS Pathog.* **2006**, *2*, e125. [CrossRef]
7. Wolf, Y.I.; Viboud, C.; Holmes, E.C.; Koonin, E.V.; Lipman, D.J. Long intervals of stasis punctuated by bursts of positive selection in the seasonal evolution of influenza A virus. *Biol. Direct.* **2006**, *1*, 34. [CrossRef]
8. Rambaut, A.; Pybus, O.G.; Nelson, M.I.; Viboud, C.; Taubenberger, J.K.; Holmes, E.C. The genomic and epidemiological dynamics of human influenza A virus. *Nature* **2008**, *453*, 615–619. [CrossRef]
9. Lowen, A.C. It's in the mix: Reassortment of segmented viral genomes. *PLoS Pathog.* **2018**, *14*, e1007200. [CrossRef]
10. de Silva, U.C.; Tanaka, H.; Nakamura, S.; Goto, N.; Yasunaga, T. A comprehensive analysis of reassortment in influenza A virus. *Biol. Open* **2012**, *1*, 385–390. [CrossRef]
11. Wille, M.; Holmes, E.C. The Ecology and Evolution of Influenza Viruses. *Cold Spring Harb. Perspect. Med.* **2020**, *10*, a038489. [CrossRef] [PubMed]
12. Li, C.; Chen, H. H7N9 Influenza Virus in China. *Cold Spring Harb. Perspect. Med.* **2021**, *11*, a038349. [CrossRef] [PubMed]
13. 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] [PubMed]
14. Lam, T.T.-Y.; Wang, J.; Shen, Y.; Zhou, B.; Duan, L.; Cheung, C.-L.; Ma, C.; Lycett, S.J.; Leung, C.Y.-H.; Chen, X.; et al. The genesis and source of the H7N9 influenza viruses causing human infections in China. *Nature* **2013**, *502*, 241–244. [CrossRef] [PubMed]
15. Pu, J.; Wang, S.; Yin, Y.; Zhang, G.; Carter, R.A.; Wang, J.; Xu, G.; Sun, H.; Wang, M.; Wen, C.; et al. Evolution of the H9N2 influenza genotype that facilitated the genesis of the novel H7N9 virus. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 548–553. [CrossRef] [PubMed]
16. Wu, A.; Su, C.; Wang, D.; Peng, Y.; Liu, M.; Hua, S.; Li, T.; Gao, G.F.; Tang, H.; Chen, J.; et al. Sequential reassortments underlie diverse influenza H7N9 genotypes in China. *Cell Host Microbe* **2013**, *14*, 446–452. [CrossRef]
17. Wang, D.; Yang, L.; Zhu, W.; Zhang, Y.; Zou, S.; Bo, H.; Gao, R.; Dong, J.; Huang, W.; Guo, J.; et al. Two Outbreak Sources of Influenza A (H7N9) Viruses Have Been Established in China. *J. Virol.* **2016**, *90*, 5561–5573. [CrossRef]
18. Zhu, Y.; Qi, X.; Cui, L.; Zhou, M.; Wang, H. Human co-infection with novel avian influenza A H7N9 and influenza A H3N2 viruses in Jiangsu province, China. *Lancet* **2013**, *381*, 2134. [CrossRef]
19. Li, J.; Kou, Y.; Yu, X.; Sun, Y.; Zhou, Y.; Pu, X.; Jin, T.; Pan, J.; Gao, G.F. Human co-infection with avian and seasonal influenza viruses, China. *Emerg. Infect. Dis.* **2014**, *20*, 1953–1955. [CrossRef]
20. Pu, Z.; Yang, J.; Shen, X.; Irwin, D.M.; Shen, Y. Human-isolated H7N9 obtained internal genes from duck and human influenza viruses. *J. Infect.* **2019**, *78*, 491–503. [CrossRef]
21. Yang, L.; Xie, J.; Zhang, Y.; Zhu, W.; Li, X.; Wei, H.; Li, Z.; Zhao, L.; Bo, H.; Liu, J.; et al. Emergence of waterfowl-originated gene cassettes in HPAI H7N9 viruses caused severe human infection in Fujian, China. *Influenza Other Respir. Viruses* **2019**, *13*, 496–503. [CrossRef]
22. Jin, Y.; Ren, H.; Teng, Y.; Hu, M.; Peng, X.; Yue, J.; Liang, L. Novel reassortment of avian influenza A(H7N9) virus with subtype H6N6 and H5N6 viruses circulating in Guangdong Province, China. *J. Infect.* **2017**, *75*, 179–182. [CrossRef] [PubMed]
23. Quan, C.; Shi, W.; Yang, Y.; Yang, Y.; Liu, X.; Xu, W.; Li, H.; Li, J.; Wang, Q.; Tong, Z.; et al. New Threats from H7N9 Influenza Virus: Spread and Evolution of High- and Low-Pathogenicity Variants with High Genomic Diversity in Wave Five: Spread and Evolution of High- and Low-Pathogenicity Variants with High Genomic Diversity in Wave Five. *J. Virol.* **2018**, *92*, e00301-18. [CrossRef] [PubMed]
24. Li, C.; Wang, S.; Bing, G.; Carter, R.A.; Wang, Z.; Wang, J.; Wang, C.; Wang, L.; Wu, G.; Webster, R.G.; et al. Genetic evolution of influenza H9N2 viruses isolated from various hosts in China from 1994 to 2013. *Emerg. Microbes Infect.* **2017**, *6*, e106. [CrossRef] [PubMed]

25. Liu, W.; Fan, H.; Raghwani, J.; Lam, T.T.-Y.; Li, J.; Pybus, O.G.; Yao, H.-W.; Wo, Y.; Liu, K.; An, X.-P.; et al. Occurrence and reassortment of avian influenza A (H7N9) viruses derived from coinfected birds in China. *J. Virol.* **2014**, *88*, 13344–13351. [CrossRef]

26. Ke, C.; Lu, J.; Wu, J.; Guan, D.; Zou, L.; Song, T.; Yi, L.; Zeng, X.; Liang, L.; Ni, H.; et al. Circulation of reassortant influenza A(H7N9) viruses in poultry and humans, Guangdong Province, China, 2013. *Emerg. Infect. Dis.* **2014**, *20*, 2034–2040. [CrossRef]

27. Qi, W.; Shi, W.; Li, W.; Huang, L.; Li, H.; Wu, Y.; Yan, J.; Jiao, P.; Zhu, B.; Ma, J.; et al. Continuous reassortments with local chicken H9N2 virus underlie the human-infecting influenza A (H7N9) virus in the new influenza season, Guangdong, China. *Protein Cell* **2014**, *5*, 878–882. [CrossRef]

28. Bai, R.; Sikkema, R.S.; Munnink, B.B.O.; Li, C.R.; Wu, J.; Zou, L.; Jing, Y.; Lu, J.; Yuan, R.Y.; Liao, M.; et al. Exploring utility of genomic epidemiology to trace origins of highly pathogenic influenza A/H7N9 in Guangdong. *Virus Evol.* **2020**, *6*, veaa097. [CrossRef]

29. Ma, M.-J.; Yang, Y.; Fang, L.-Q. Highly Pathogenic Avian H7N9 Influenza Viruses: Recent Challenges: Recent Challenges. *Trends Microbiol.* **2019**, *27*, 93–95. [CrossRef]

30. Qiu, Y.; Sun, R.; Hou, G.; Yu, X.; Li, Y.; Li, J.; Zhang, Q.; Zou, F.; Liu, H.; Jiang, W. Novel reassortant H7N2 originating from the H7N9 highly pathogenic avian influenza viruses in China, 2019. *J. Infect.* **2019**, *79*, 462–470. [CrossRef]

31. Nakayama, M.; Uchida, Y.; Shibata, A.; Kobayashi, Y.; Mine, J.; Takemae, N.; Tsunekuni, R.; Tanikawa, T.; Harada, R.; Osaka, H.; et al. A novel H7N3 reassortant originating from the zoonotic H7N9 highly pathogenic avian influenza viruses that has adapted to ducks. *Transbound. Emerg. Dis.* **2019**, *66*, 2342–2352. [CrossRef] [PubMed]

32. Qu, B.; Li, X.; Cardona, C.J.; Xing, Z. Reassortment and adaptive mutations of an emerging avian influenza virus H7N4 subtype in China. *PLoS ONE* **2020**, *15*, e0227597. [CrossRef] [PubMed]

33. Zhao, Z.; Liu, L.; Guo, Z.; Zhang, C.; Wang, Z.; Wen, G.; Zhang, W.; Shang, Y.; Zhang, T.; Jiao, Z.; et al. A Novel Reassortant Avian H7N6 Influenza Virus Is Transmissible in Guinea Pigs via Respiratory Droplets. *Front. Microbiol.* **2019**, *10*, 18. [CrossRef] [PubMed]

34. Gu, M.; Chen, H.; Li, Q.; Huang, J.; Zhao, M.; Gu, X.; Jiang, K.; Wang, X.; Peng, D.; Liu, X. Enzootic genotype S of H9N2 avian influenza viruses donates internal genes to emerging zoonotic influenza viruses in China. *Vet. Microbiol.* **2014**, *174*, 309–315. [CrossRef] [PubMed]

35. Cui, L.; Liu, D.; Shi, W.; Pan, J.; Qi, X.; Li, X.; Guo, X.; Zhou, M.; Li, W.; Li, J.; et al. Dynamic reassortments and genetic heterogeneity of the human-infecting influenza A (H7N9) virus. *Nat. Commun.* **2014**, *5*, 3142. [CrossRef]

36. Pu, J.; Yin, Y.; Liu, J.; Wang, X.; Zhou, Y.; Wang, Z.; Sun, Y.; Sun, H.; Li, F.; Song, J.; et al. Reassortment with dominant chicken H9N2 influenza virus contributed to the fifth H7N9 virus human epidemic. *J. Virol.* **2021**, *95*, e01578-20. [CrossRef]

37. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef]

38. He, D.; Gu, J.; Gu, M.; Wu, H.; Li, J.; Zhan, T.; Chen, Y.; Xu, N.; Ge, Z.; Wang, G.; et al. Genetic and antigenic diversity of H7N9 highly pathogenic avian influenza virus in China. *Infect. Genet. Evol.* **2021**, *93*, 104993. [CrossRef]

39. Drummond, A.J.; Suchard, M.A.; Xie, D.; Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **2012**, *29*, 1969–1973. [CrossRef]

40. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [CrossRef]

41. Hill, V.; Baele, G. Bayesian estimation of past population dynamics in BEAST 1.10 using the Skygrid coalescent model. *Mol. Biol. Evol.* **2019**, *36*, 2620–2628. [CrossRef] [PubMed]

42. Prosperi, M.C.F.; Ciccozzi, M.; Fanti, I.; Saladini, F.; Pecorari, M.; Borghi, V.; Di Giambenedetto, S.; Bruzzone, B.; Capetti, A.; Vivarelli, A.; et al. A novel methodology for large-scale phylogeny partition. *Nat. Commun.* **2011**, *2*, 321. [CrossRef] [PubMed]

43. Yu, G. Using ggtree to Visualize Data on Tree-Like Structures. *Curr. Protoc. Bioinform.* **2020**, *69*, e96. [CrossRef]

44. Xu, S.; Dai, Z.; Guo, P.; Fu, X.; Liu, S.; Zhou, L.; Tang, W.; Feng, T.; Chen, M.; Zhan, L.; et al. ggtreeExtra: Compact Visualization of Richly Annotated Phylogenetic Data. *Mol. Biol. Evol.* **2021**, *38*, 4039–4042. [CrossRef]

45. Altekar, G.; Dwarkadas, S.; Huelsenbeck, J.P.; Ronquist, F. Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* **2004**, *20*, 407–415. [CrossRef]

46. Westgeest, K.B.; Russell, C.A.; Lin, X.; Spronken, M.I.J.; Bestebroer, T.M.; Bahl, J.; van Beek, R.; Skepner, E.; Halpin, R.A.; de Jong, J.C.; et al. Genomewide analysis of reassortment and evolution of human influenza A(H3N2) viruses circulating between 1968 and 2011. *J. Virol.* **2014**, *88*, 2844–2857. [CrossRef]

47. Avino, M.; Ng, G.T.; He, Y.; Renaud, M.S.; Jones, B.R.; Poon, A.F.Y. Tree shape-based approaches for the comparative study of cophylogeny. *Ecol. Evol.* **2019**, *9*, 6756–6771. [CrossRef]

48. Galili, T. dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **2015**, *31*, 3718–3720. [CrossRef]

49. Nagarajan, N.; Kingsford, C. GiRaF: Robust, computational identification of influenza reassortments via graph mining. *Nucleic Acids Res.* **2011**, *39*, e34. [CrossRef]

50. Martin, D.P.; Murrell, B.; Golden, M.; Khoosal, A.; Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **2015**, *1*, vev003. [CrossRef]

51. Müller, N.F.; Stolz, U.; Dudas, G.; Stadler, T.; Vaughan, T.G. Bayesian inference of reassortment networks reveals fitness benefits of reassortment in human influenza viruses. *Proc. Natl. Acad. Sci. USA* **2019**, *117*, 17104–17111. [CrossRef] [PubMed]

52. Wickham, H. *Ggplot2: Elegrant graphics for data analysis/Hadley Wickham; with contributions by Carson Sievert*, 2nd ed.; Springer: Cham, Switzerland, 2016; pp. 9–199.

53. Jacquot, M.; Rao, P.P.; Yadav, S.; Nomikou, K.; Maan, S.; Jyothi, Y.K.; Reddy, N.; Putty, K.; Hemadri, D.; Singh, K.P.; et al. Contrasting selective patterns across the segmented genome of bluetongue virus in a global reassortment hotspot. *Virus Evol.* **2019**, *5*, vez027. [CrossRef] [PubMed]

54. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermiin, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589. [CrossRef] [PubMed]

55. 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]

56. 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]

57. 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]

58. Kosakovsky Pond, S.L.; Poon, A.F.Y.; Velazquez, R.; Weaver, S.; Hepler, N.L.; Murrell, B.; Shank, S.D.; Magalis, B.R.; Bouvier, D.; Nekrutenko, A.; et al. HyPhy 2.5-A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol. Biol. Evol.* **2020**, *37*, 295–299. [CrossRef]

59. Chen, R.; Holmes, E.C. Avian influenza virus exhibits rapid evolutionary dynamics. *Mol. Biol. Evol.* **2006**, *23*, 2336–2341. [CrossRef]

60. Cattoli, G.; Fusaro, A.; Monne, I.; Coven, F.; Joannis, T.; El-Hamid, H.S.A.; Hussein, A.A.; Cornelius, C.; Amarin, N.M.; Mancin, M.; et al. Evidence for differing evolutionary dynamics of A/H5N1 viruses among countries applying or not applying avian influenza vaccination in poultry. *Vaccine* **2011**, *29*, 9368–9375. [CrossRef]

61. Wu, Y.; Hu, J.; Jin, X.; Li, X.; Wang, J.; Zhang, M.; Chen, J.; Xie, S.; Qi, W.; Liao, M.; et al. Accelerated Evolution of H7N9 Subtype Influenza Virus under Vaccination Pressure. *Virol. Sin.* **2021**, *36*, 1124–1132. [CrossRef]

62. Zhang, L.; Zhang, Z.; Weng, Z.; Shi, W. Substitution rates of the internal genes in the novel avian H7N9 influenza virus. *Clin. Infect. Dis.* **2013**, *57*, 1213–1215. [CrossRef] [PubMed]

63. Fitch, W.M.; Leiter, J.M.; Li, X.Q.; Palese, P. Positive Darwinian evolution in human influenza A viruses. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 4270–4274. [CrossRef] [PubMed]

64. Suzuki, Y. Natural selection on the influenza virus genome. *Mol. Biol. Evol.* **2006**, *23*, 1902–1911. [CrossRef] [PubMed]

65. Yu, X.; Jin, T.; Cui, Y.; Pu, X.; Li, J.; Xu, J.; Liu, G.; Jia, H.; Liu, D.; Song, S.; et al. Influenza H7N9 and H9N2 viruses: Coexistence in poultry linked to human H7N9 infection and genome characteristics. *J. Virol.* **2014**, *88*, 3423–3431. [CrossRef]

66. Marshall, N.; Priyamvada, L.; Ende, Z.; Steel, J.; Lowen, A.C. Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. *PLoS Pathog.* **2013**, *9*, e1003421. [CrossRef]

67. Gerber, M.; Isel, C.; Moules, V.; Marquet, R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol.* **2014**, *22*, 446–455. [CrossRef]

68. Essere, B.; Yver, M.; Gavazzi, C.; Terrier, O.; Isel, C.; Fournier, E.; Giroux, F.; Textoris, J.; Julien, T.; Socratous, C.; et al. Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E3840-8. [CrossRef]

69. Sun, X.; Belser, J.A.; Pappas, C.; Pulit-Penaloza, J.A.; Brock, N.; Zeng, H.; Creager, H.M.; Le, S.; Wilson, M.; Lewis, A.; et al. Risk Assessment of Fifth-Wave H7N9 Influenza A Viruses in Mammalian Models. *J. Virol.* **2019**, *93*, e01740-18. [CrossRef]

70. Hao, X.; Hu, J.; Wang, X.; Gu, M.; Wang, J.; Liu, D.; Gao, Z.; Chen, Y.; Gao, R.; Li, X.; et al. The PB2 and M genes are critical for the superiority of genotype S H9N2 virus to genotype H in optimizing viral fitness of H5Nx and H7N9 avian influenza viruses in mice. *Transbound. Emerg. Dis.* **2020**, *67*, 758–768. [CrossRef]

71. Hao, X.; Wang, X.; Hu, J.; Gu, M.; Wang, J.; Deng, Y.; Jiang, D.; He, D.; Xu, H.; Yang, Y.; et al. The PB2 and M genes of genotype S H9N2 virus contribute to the enhanced fitness of H5Nx and H7N9 avian influenza viruses in chickens. *Virology* **2019**, *535*, 218–226. [CrossRef]

72. Villa, M.; Lässig, M. Fitness cost of reassortment in human influenza. *PLoS Pathog.* **2017**, *13*, e1006685. [CrossRef] [PubMed]

73. Lu, L.; Lycett, S.J.; Leigh Brown, A.J. Reassortment patterns of avian influenza virus internal segments among different subtypes. *BMC Evol. Biol.* **2014**, *14*, 16. [CrossRef] [PubMed]

74. Yao, Y.; Zhang, T.; Yang, W.; Shao, Z.; He, B.; Chen, X.; Wu, L.; Jin, E.; Liu, H.; Chen, J.; et al. Avian Influenza A (H7N9) Virus in a Wild Land Bird in Central China, Late 2015. *Virol. Sin.* **2018**, *33*, 96–99. [CrossRef] [PubMed]

75. Bui, C.; Bethmont, A.; Chughtai, A.A.; Gardner, L.; Sarkar, S.; Hassan, S.; Seale, H.; MacIntyre, C.R. A Systematic Review of the Comparative Epidemiology of Avian and Human Influenza A H5N1 and H7N9 - Lessons and Unanswered Questions. *Transbound. Emerg. Dis.* **2016**, *63*, 602–620. [CrossRef]

76. Bhat, S.; James, J.; Sadeyen, J.-R.; Mahmood, S.; Everest, H.J.; Chang, P.; Walsh, S.; Byrne, A.M.P.; Mollett, B.; Lean, F.; et al. Co-infection of chickens with H9N2 and H7N9 avian influenza viruses leads to emergence of reassortant H9N9 virus with increased fitness for poultry and enhanced zoonotic potential. *J. Virol.* **2021**, *96*, e0185621. [CrossRef]

77. Dugan, V.G.; Chen, R.; Spiro, D.J.; Sengamalay, N.; Zaborsky, J.; Ghedin, E.; Nolting, J.; Swayne, D.E.; Runstadler, J.A.; Happ, G.M.; et al. The evolutionary genetics and emergence of avian influenza viruses in wild birds. *PLoS Pathog.* **2008**, *4*, e1000076. [CrossRef]
78. Dudas, G.; Bedford, T.; Lycett, S.; Rambaut, A. Reassortment between influenza B lineages and the emergence of a coadapted PB1-PB2-HA gene complex. *Mol. Biol. Evol.* **2015**, *32*, 162–172. [CrossRef]
79. Neverov, A.D.; Lezhnina, K.V.; Kondrashov, A.S.; Bazykin, G.A. Intrasubtype reassortments cause adaptive amino acid replacements in H3N2 influenza genes. *PLoS Genet.* **2014**, *10*, e1004037. [CrossRef]