



Deep Sequencing of H7N9 Influenza A Viruses from 16 Infected Patients from 2013 to 2015 in Shanghai Reveals Genetic Diversity and Antigenic Drift

Yong-Li Xiao,^a Lili Ren,^b Xi Zhang,^c Li Qi,^a John C. Kash,^a Yan Xiao,^b Fan Wu,^c Jianwei Wang,^b Jeffery K. Taubenberger^a

^aViral Pathogenesis and Evolution Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA

^bMOH Key Laboratory of Systems Biology of Pathogens and Christophe Mérieux Laboratory, Institute of Pathogen Biology, Chinese Academy of Medical Sciences, Beijing, People's Republic of China

^cShanghai Municipal Center for Disease Control and Prevention, Shanghai, People's Republic of China

ABSTRACT Influenza A virus (IAV) infections are a major public health concern, including annual epidemics, epizootic outbreaks, and pandemics. A significant IAV epizootic outbreak was the H7N9 avian influenza A outbreak in China, which was first detected in 2013 and which has spread over 5 waves from 2013 to 2017, causing human infections in many different Chinese provinces. Here, RNA from primary clinical throat swab samples from 20 H7N9-infected local patients with different clinical outcomes, who were admitted and treated at one hospital in Shanghai, China, from April 2013 to April 2015, was analyzed. Whole-transcriptome amplification, with positive enrichment of IAV RNA, was performed, all 20 samples were subjected to deep sequencing, and data from 16 samples were analyzed in detail. Many single-nucleotide polymorphisms, including ones not previously reported, and many non-synonymous changes that could affect hemagglutinin head and stalk antibody binding epitopes were observed. Minor populations representing viral quasispecies, including nonsynonymous hemagglutinin changes shared by antigenically variant H7N9 clades identified in the most recent wave of H7N9 infections in 2016 to 2017, were also identified.

IMPORTANCE H7N9 subtype avian influenza viruses caused infections in over 1,400 humans from 2013 to 2017 and resulted in almost 600 deaths. It is important to understand how avian influenza viruses infect and cause disease in humans and to assess their potential for efficient person-to-person transmission. In this study, we used deep sequencing of primary clinical material to assess the evolution and potential for human adaptation of H7N9 influenza viruses.

KEYWORDS DNA sequencing, H7N9, avian viruses

Influenza A virus (IAV) infections are a major public health concern, including annual epidemics, epizootic outbreaks, and occasional pandemics (1). Historically, there have been at least 14 influenza pandemics in the last 500 years (2) and 4 in the last 100 years, including the “Spanish” influenza A/H1N1 pandemic in 1918 that resulted in at least 50 million deaths worldwide (3) and the most recent one, in 2009, caused by a novel swine-origin influenza A/H1N1 strain which may have resulted in up to 284,000 deaths globally (4).

Influenza epizootic outbreaks, in which animal-derived IAV infections cause localized outbreaks in the absence of sustained human-to-human transmission, can cause high levels of morbidity and mortality. Such epizootic outbreaks may also presage the emergence of a stably human-adapted IAV strain capable of causing a pandemic. One of the most significant IAV epizootic outbreaks of the past decade is the H7N9 avian

Received 27 August 2018 Accepted 28 August 2018 Published 19 September 2018

Citation Xiao Y-L, Ren L, Zhang X, Qi L, Kash JC, Xiao Y, Wu F, Wang J, Taubenberger JK. 2018. Deep sequencing of H7N9 influenza A viruses from 16 infected patients from 2013 to 2015 in Shanghai reveals genetic diversity and antigenic drift. *mSphere* 3:e00462-18. <https://doi.org/10.1128/mSphereDirect.00462-18>.

Editor Michael J. Imperiale, University of Michigan—Ann Arbor

Copyright © 2018 Xiao et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Jeffery K. Taubenberger, taubenbergerj@niaid.nih.gov. Y.-L.X., L.R., and X.Z. contributed equally to the article.

J.W. and J.K.T. are co-senior authors.

Solicited external reviewers: Elodie Ghedin, New York University; Jonathan Runstadler, Tufts University.

This paper was submitted via the [mSphereDirect™](https://msphere.asm.org/) pathway.

influenza A outbreak, first detected in China in March 2013 (5). H7N9 human spillover infections have spread over 5 waves to many different Chinese provinces, and a number of imported H7N9 cases have been reported in other countries (6, 7). As of 7 December 2017, a total of 1,566 laboratory-confirmed cases of human infection with avian influenza A/H7N9 viruses, including at least 613 deaths, have been reported, with a case fatality rate (CFR) of approximately 39% (8). Previously published studies examined the origin, spread, and continued genetic and antigenic evolution of H7N9 viruses since the emergence of the strain in 2013, including frequent reassortment with other avian IAVs, particularly H9N2 viruses (9, 10); acquisition of mutations associated with mammalian adaptation (7, 8, 11); and reduced susceptibility to viral neuraminidase (NA) inhibitor-based antiviral drugs (6).

During the first 4 waves, or epizootic clusters, of human H7N9 cases (2013 to 2016), almost 90% of diagnosed patients developed pneumonia, approximately two-thirds required intensive care, and 41% had a fatal outcome (6). The H7N9 viral isolates responsible for the most recent (fifth) wave (2016 to 2017) have been shown to have evolved into two distinct genetic lineages, referred to as the Pearl River Delta and Yangtze River Delta lineages, with most recent isolates belonging to the Yangtze River Delta lineage (6, 8). Importantly, viruses in this lineage show reduced antigenic relatedness to candidate prepandemic H7N9 vaccine virus stocks developed in 2013 (6).

Several different methods have been developed to detect IAV infection in patients by the use of different approaches, including methods based on tissue culture, antibody titers in serological assays, nucleic acid amplification, and nucleic acid sequencing (12). Since their commercialization in 2005, next-generation sequencing technologies have evolved dramatically, especially regarding sequencing capacity, which has increased by factors of 100 to 1,000 (13). These sequencing technologies have been widely used in influenza research studies, including detecting IAV and norovirus infections in patients (14), uncovering mixed infection with 2009 pandemic IAV (15), high-throughput sequencing of influenza B viruses (16), evaluating genetic stability of influenza vaccine viruses (17), revealing antigenic variants at low frequencies in IAV-infected patients (18), and high-throughput identification of IAV H3N2 antigenic drift variants (19). However, each of these approaches uses virus-specific primers and virus-specific PCR amplification strategies. Due to the ability of next-generation sequencing to detect low-abundance sequences, avoidance of PCR-introduced errors would be helpful in sequence analyses. It has been shown that most commonly used PCR enzymes, including high-fidelity enzymes, have error rates of 10^{-5} to 10^{-6} point mutations/base pair/duplication (20). Besides well-characterized polymerase base substitution errors, other sources of error have been found to be equally prevalent, including PCR-mediated recombination, template switching, and DNA damage introduced during temperature cycling (21). In addition, lower-quality or partially degraded RNA, particularly that present in clinical or archival samples, can pose a great challenge for PCR-based amplification using influenza virus universal primers (22) or specific primers.

In the current study, we applied a whole-transcriptome amplification (WTA) method, with positive selection of IAV H7N9 viral RNA sequences by target enrichment hybridization, and next-generation sequencing technologies. Using a combined WTA/positive-enrichment target hybridization method, we sought to perform analyses that could be applied to routinely obtained clinical samples that contain limited amounts of RNA, often with degradation. Recently, Briese et al. demonstrated that a similar target enrichment approach with as little as 40% sequence identity of a targeted genome can be used for preparing a comprehensive probe library (23). Using these approaches, we directly sequenced 20 clinical throat swab samples from H7N9-infected patients in Shanghai, China, with different clinical outcomes from 2013 to 2015 by the use of Illumina next-generation sequencing technology. By analyzing sequences from 16 of the 20 samples in detail, we obtained both the consensus viral sequences from each patient with a H7N9 IAV infection and minor populations representing viral quasispecies. Although the samples analyzed in this study were from only one region, Shanghai,

TABLE 1 Summary of all patient data^a

Sample	Date of onset of symptoms	Hospital admission date	Sampling date	Outcome	Antiviral therapy	Antiviral therapy date	Antibiotics
1	30 March 2013	3 April 2013	7 April 2013	Discharge	Tamiflu	3 April 2013	Yes
2	9 April 2013	16 April 2013	17 April 2013	Death	Tamiflu	17 April 2013	Yes
3	4 April 2013	8 April 2013	12 April 2013	Discharge	Tamiflu	12 April 2013	Yes
4	25 March 2013	5 April 2013	7 April 2013	Death	Tamiflu	5 April 2013	Yes
5	15 March 2013	12 April 2013	12 April 2013	Discharge	Tamiflu/Ribavirin	12 April 2013	Yes
6	4 April 2013	9 April 2013	10 April 2013	Discharge	Tamiflu	10 April 2013	Yes
7	30 March 2013	4 April 2013	6 April 2013	Discharge	Tamiflu	4 April 2013	Yes
8	26 December 2013	30 December 2013	3 January 2014	Death	Yes	2 January 2014	Yes
9	1 January 2014	8 January 2014	10 January 2014	Death	Yes	9 January 2014	Yes
10	3 January 2014	9 January 2014	16 January 2014	Death	NA	NA	NA
11	NA	NA	NA	NA	NA	NA	NA
12	6 January 2014	20 January 2014	28 January 2014	Death	Yes	20 January 2014	Yes
13	NA	NA	17 January 2014	NA	NA	NA	NA
14	10 January 2014	17 January 2014	18 January 2014	Death	NA	NA	Yes
15	11 January 2014	17 January 2014	18 January 2014	Death	Yes	17 January 2014	Yes
16	NA	NA	30 January 2014	NA	NA	NA	NA
17	3 January 2015	17 January 2015	22 January 2015	Death	Yes	21 January 2015	Yes
18	4 February 2015	20 February 2015	11 February 2015	Death	Yes	10 February 2015	Yes
19	26 March 2015	1 April 2015	3 April 2015	Discharge	Yes	1 April 2015	Yes
20	27 March 2015	3 April 2015	6 April 2015	Death	Yes	5 April 2015	Yes

^aNA, not available.

China, examination of the viral quasispecies revealed several hemagglutinin nonsynonymous mutations shared with the antigenically variant H7N9 clades identified in the most recent waves of H7N9 infections in 2016 to 2017 (6, 8). In addition, the deep-sequencing technology revealed a number of single-nucleotide polymorphisms (SNPs) in our samples that have not been previously reported in humans.

RESULTS

Patient information. In this study, throat swabs from 20 reverse transcription-PCR (RT-PCR)-confirmed H7N9-infected patients who were hospitalized in Shanghai, China, from 2013 to 2015 were processed. Available clinical information from the patients in the study, including symptom onset, hospitalization and sampling dates, antiviral/antibiotics usage, and patient outcome, is shown in Table 1. Among the cases, 7 were infected during wave 1 (6–8), with hospitalization admissions from 3 April to 16 April 2013. There were 2 deaths in this cohort, with a case fatality rate (CFR) of 28.6%. There were 9 cases during wave 2, with hospitalization admissions from 30 December 2013 to 20 January 2014. There were 3 cases for which outcome information was not available, while the remaining 6 patients all died; thus, for this cohort, the CFR was $\geq 66.7\%$. There were 4 cases in wave 3, with hospitalization admissions from 17 January to 3 April 2015. There were 3 deaths in this cohort, with a CFR of 75.0%. Overall, the CFR in these hospitalized H7N9-infected patients was $\geq 55\%$.

Sequence analysis of positive- and negative-control experiments. Prior to experimental sequence analysis of the patient samples, positive- and negative-control experiments were designed and performed to evaluate the WTA/enrichment method. The sequencing library for the positive control was made from RNA that had been reverse transcribed from the hemagglutinin (HA) open reading frame (KC853766.1) from A/Hangzhou/1/2013 (H7N9) cloned into pBluescript II KS(-). The transcribed HA RNA was processed by the use of the same pipeline, including linear amplification, H7N9 probe enrichment, sequencing, and analysis. Because of the high sensitivity of high-throughput sequencing technology, we also performed a negative-control experiment to determine a baseline of influenza virus contamination in the construction of the libraries. For the negative control, a sequencing library was made from HeLa cell total RNA following the positive-control experiment using the same process pipeline.

From the positive-control HA RNA library, we obtained 355,521,326 reads; 342,112,470 (96.23%) of them aligned to the H7N9 HA reference segment. Similarly,

TABLE 2 Base changes from positive control^a

Reference	nt position	Reference base	Coverage (no. of reads)	No. of insertions	No. of deletions	No. (%) of substitutions					Avg <i>q</i> value
						A	T	G	C	N	
gi_475662453_HA	79	G	39,651	1	0	A = 25 (0.1)	T = 487 (1.2)	ref (G)	C = 17 (0.0)	N = 0 (0.0)	32.16
gi_475662453_HA	1144	T	47,766	1	2	A = 1,114 (2.3)	ref = T	G = 5 (0.0)	C = 24 (0.1)	N = 0 (0.0)	33.958
gi_475662453_HA	1146	G	47,155	0	0	A = 1,003 (2.1)	T = 52 (0.1)	ref (G)	C = 7 (0.0)	N = 0 (0.0)	33.858

^ant, nucleotide; *q*, false-discovery rate; ref, reference nucleotide.

after SNP analysis performed without any specific requirements, the highest SNP value in this run was 2.3% at HA nucleotide position 1144 (46622 reads as T; 1114 reads as A), with coverage of 47,766 reads (Table 2). This provided a baseline of the noise in the library construction and sequencing pipeline. It has been reported that the error rate of the Illumina platform generally increases toward the end of the read, reaching as high as 5% (24). Therefore, to reduce the potential false-positive SNP calling rate, we decided to use 10% as the SNP cutoff level for this study. We obtained 292,901,554 total reads from the negative-control experiment, and there were 3,815 (0.0013%) reads aligned to H7N9 consensus sequences. Therefore, we used this number of sequence reads as an approximate baseline for the sequence contamination level for this study (see below).

H7N9 viral cDNA library sequence analysis. Illumina deep sequencing of enriched cDNA libraries was performed on samples from 20 patients that were H7N9 RT-PCR positive. The summary of the sequencing results is shown in Table 3, and data corresponding to the average sequencing depth of each segment of the genome from all the samples are shown in Table 4. Between 42 million and 324 million reads (reflecting the sum of reads produced in both directions) were generated from the 20 patient sample libraries. After mapping to the reference human H7N9 consensus sequence from an isolate collected in China in 2013, the total number of mapped reads ranged from 3,245 reads to over 178 million reads among the patient sample libraries. Because there were 3,815 mapped reads in our negative-control experiment, samples 10 (number of mapped reads, 10,693), 11 (number of mapped reads, 3,534), 12 (number of mapped reads, 3,830), and 18 (number of mapped reads, 3,245) were eliminated from further SNP analysis to avoid potential false-positive results. Therefore, compared to the human 2013 consensus H7N9 sequence, a total of 509 SNPs were detected from these 16 remaining patient sample libraries, with 454 unique SNP positions identified (Fig. 1A). Synonymous and nonsynonymous SNPs that passed the filter criteria utilized (described in Materials and Methods) from each patient are shown in Table 3 (see also Table S1 in the supplemental material).

Comparison of SNPs detected in this study and SNPs detected from previously sequenced H7N9 human isolates. The precomputed SNP data from human H7N9 IAVs were downloaded on 6 March 2017 from the Influenza Research Database (IRD) (25) with different SNP scores, which were generated from the modified formula as described by Crooks et al. (26). Briefly, the scores represent the normalized entropy of the observed allele distribution at each position. The least polymorphic site would have a score of 0 (no polymorphism), and the most polymorphic site would have a score of 200 (highest polymorphism). Compared to the SNPs from the IRD with a score of >1, among the 454 SNPs identified in the current study, there were 276 previously reported in the IRD. At the level of an SNP score of >10, only 195 SNPs were reported by IRD (Table S1). In addition, at the level of an IRD SNP score of >1, among the same 276 SNP positions, 255 SNP calls were the same as those observed in the current study. There were 178 SNP positions with 106 nonsynonymous changes in human H7N9 that have not been reported at the IRD.

Comparison of H7N9 viral genome SNPs between different patients. Among the 16 remaining sequenced and analyzed patient samples, a total of 509 SNPs that passed

TABLE 3 Summary of all sample data

Sample	Read type	Read length (bp)	Total no. of reads produced/ no. of left reads produced	No. of right reads produced	Total no. of reads mapped to H7N9/ no. of left reads mapped	No. of right reads mapped	Coverage (%)										Total no. of nonsynonymous SNPs	Total no. of mapped reads
							2013H7N9_ NA	2013H7N9_ HA	2013H7N9_ M2_M1	2013H7N9_ consensus_PB2	2013H7N9_ consensus_PB1	2013H7N9_ PA	2013H7N9_ NP	2013H7N9_ consensus_NEP_NS1	Avg coverage	Total no. of SNPs		
1	Single end	80	1.1E+08	1.6E+08	3,637,318	8.5E+07	0.9979	1	0.9996	0.9213	1	0.9947	0.9869	0.9875	75	48	3,637,318	
2	Paired end	100	1.6E+08	6.2E+07	9.4E+07	2E+07	0.9785	1	0.9807	1	0.9967	1	1	0.9945	43	22	1.8E+08	
3	Paired end	100	6.2E+07	6.2E+07	2.6E+07	2E+07	0.4256	0.5264	0.9165	0.1033	0.4816	0.9993	0.6313	0.5767	11	8	4.6E+07	
4	Paired end	100	3.4E+07	3.4E+07	4,519,318	4,315,814	0.6009	1	0.9662	0.6306	0.9805	0.8971	0.9976	0.8841	18	6	8,835,132	
5	Paired end	100	3.5E+07	3.5E+07	209,949	187,212	0.6481	0.9721	0.999	0.8813	0.9921	0.8938	0.9499	0.9168	12	6	397,161	
6	Paired end	100	3.9E+07	3.9E+07	7,884,584	7,067,362	0.6059	0.9447	1	0.9763	0.9991	0.8838	0.673	0.857	38	8	1.5E+07	
7	Paired end	100	6E+07	6E+07	2,784,497	2,463,756	0.4428	0.502	0.9562	0.5456	0.1579	0.537	0.4236	0.5766	11	9	5,248,253	
8	Paired end	100	2.7E+07	2.7E+07	369,432	355,017	0.3569	0.5258	0.9552	0.5228	0.2498	0.9865	0.7673	0.6705	29	9	724,449	
9	Paired end	100	2.2E+07	2.2E+07	126,599	117,241	0.3534	0.4724	0.7994	0.3702	0.1205	0.953	0.6981	0.5815	3	1	243,840	
10	Paired end	100	2.1E+07	2.1E+07	5,554	5,139	0.5322	0.6144	0.6568	0.6075	0.3914	0.9963	0.6516	0.6658			10,693	
11	Paired end	125	7.6E+07	7.6E+07	1,821	1,713	0.877	0.9412	0.8248	0.4654	0.7744	0.9893	0.5668	0.7822			3,534	
12	Paired end	125	8.6E+07	8.6E+07	2,055	1,775	0.9049	0.959	0.6263	0.5061	0.9666	0.9893	0.6265	0.8207			3,830	
13	Paired end	125	9.7E+07	9.7E+07	3,6E+07	3.5E+07	1	1	1	0.9268	1	1	1	0.9909	77	17	7.1E+07	
14	Paired end	125	8.9E+07	8.9E+07	653,106	647,614	0.5622	0.8259	0.9735	0.3412	0.2933	0.6815	0.8675	0.6741	33	14	1,300,720	
15	Paired end	125	8.4E+07	8.4E+07	2,981,673	2,871,180	1	1	0.9674	0.7728	0.9996	1	1	0.9674	75	20	5,852,853	
16	Paired end	125	1E+08	1E+08	8,765,927	8,345,327	0.9964	1	0.9776	0.7018	1	1	1	0.9595	107	35	1.7E+07	
17	Paired end	125	8.5E+07	8.5E+07	9,650,732	9,228,316	1	1	0.9776	0.6557	1	1	1	0.9542	121	23	1.9E+07	
18	Paired end	125	9.9E+07	9.9E+07	2,032	1,213	0.9614	0.9899	0.831	0.5456	0.9974	0.9893	0.9572	0.905			3,245	
19	Paired end	125	8.8E+07	8.8E+07	128,978	123,197	0.6295	0.899	0.834	0.4158	0.8034	0.9777	0.7709	0.7817	29	7	252,175	
20	Paired end	125	8.2E+07	8.2E+07	53,551	51,599	0.4664	0.9721	0.6151	0.4535	0.9494	0.8904	0.753	0.7381	11	6	105,150	

TABLE 4 Average sequencing depth of all samples

Consensus segment	Avg sequencing depth (no. of reads) for sample:																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
PB2_2013	43,366.63	3,176,994.52	2,111,699.39	94,110.75	6,312.50	76,066.89	41,121.49	1,849.69	5.13	17.37	20.78	21.16	318,682.62	751.59	22,865.95	137,128.67	123,741.12	11.90	1,562.34	8.39
PB1_2013	4,312.23	956,734.17	4,550.66	52,287.47	1,749.74	51,385.22	7.78	8.09	1,020.84	8.80	7.64	17.54	443,424.91	0.33	25,870.11	92,391.27	122,526.39	16.38	64.04	9.62
PA_2013	6,286.60	1,409,587.03	79.50	3,373.22	4,201.35	29,885.45	97,808.97	1,942.12	5.02	27.80	22.78	15.57	747,723.04	20,806.25	102,977.92	74,182.53	117,628.43	33.59	2,052.82	160.04
NP_2013	8,352.22	828,919.44	1,109,193.05	74,219.91	5,392.38	71,960.99	96,481.43	7,431.13	17.97	175.48	105.13	31.73	947,654.98	14,806.53	44,093.52	220,890.95	175,262.27	37.16	10,432.36	18.79
NS_2013	4,221.86	364,374.34	1,279.36	1,733.17	365.75	11.34	0.84	5,633.65	16.83	12.65	2.68	2.54	269,508.95	10,758.95	39,148.42	17,559.82	55,246.42	12.08	1.79	1.49
MP_2013	5,747.54	1,279,071.59	2,493,897.05	23,211.51	1,385.23	838,214.05	1,970.88	3,199.19	8.04	52.85	24.44	14.89	393,944.12	4,970.60	34,398.93	63,874.87	114,203.92	12.60	1,910.13	80.45
HA_2013	24,865.55	985,219.43	8,491.03	238,632.09	1,470.89	57,589.48	35,977.54	25,961.20	13,059.48	369.46	62.83	124.78	1,809,859.77	35,499.85	131,745.77	537,585.12	541,936.27	75.83	2,794.49	7,102.20
NA_2013	43,580.60	817,996.07	4,326.90	3,945.42	63.52	85,819.60	11,151.14	1,638.34	4.36	4.08	2.59	5.72	187,019.31	4,614.81	19,543.10	24,959.23	45,620.17	7.15	21.91	2.24

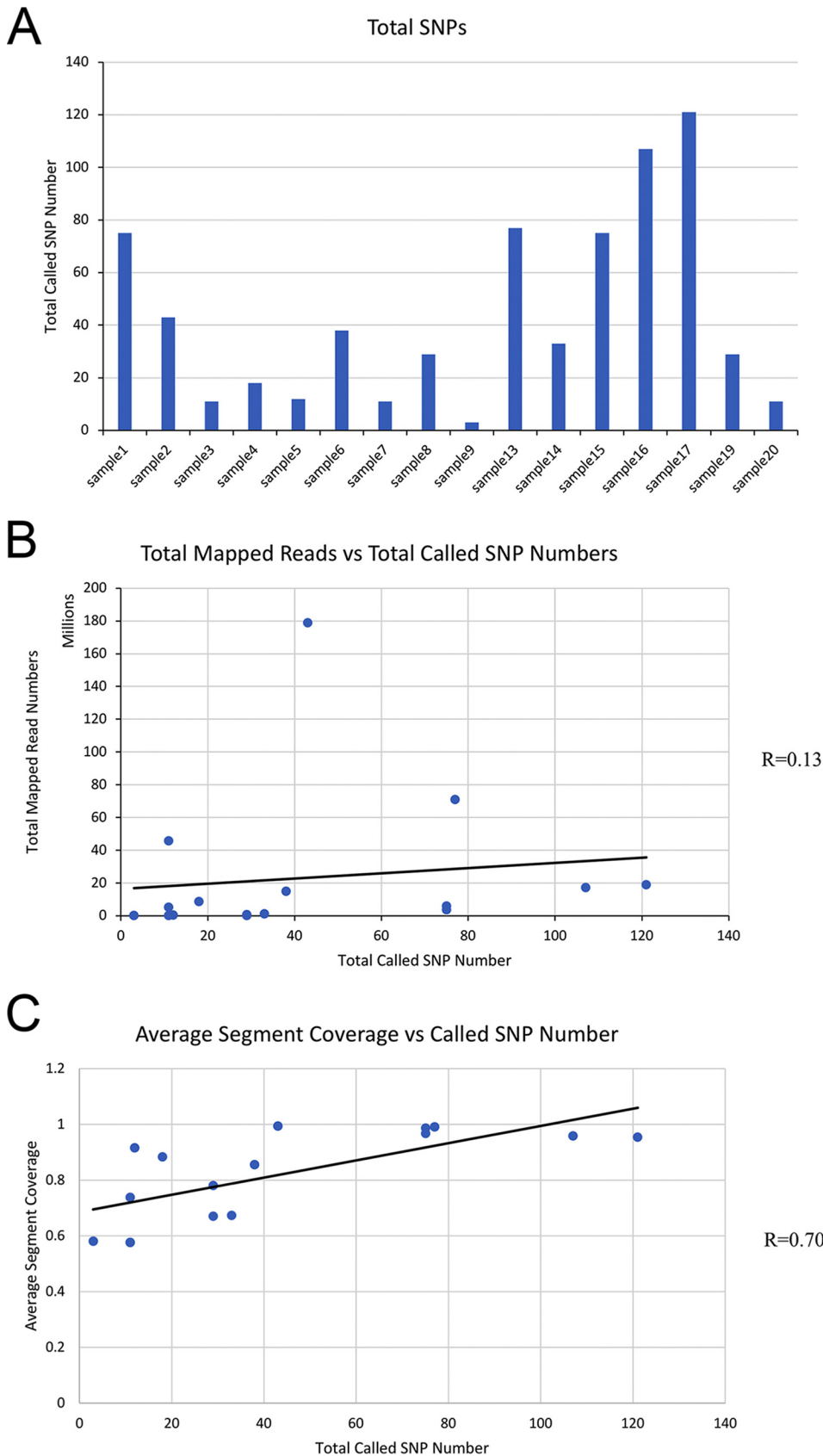


FIG 1 Mapped reads and called SNP numbers. (A) Numbers of called SNPs for all samples. (B) Correlation of the number of mapped reads and the number of called SNPs. (C) Correlation of the average coverage and the number of called SNPs.

filter criteria were detected. The number of SNP locations observed in each sample was quite variable, ranging from 3 SNPs in sample 9 to 121 in sample 17 (Table 3). The correlation (R) between the number of mapped reads and the called SNP number was low ($R = 0.13$) (Fig. 1A and B). For example, sample 3 had more than 45 million mapped reads but only 11 SNPs. On the other hand, the actual numbers of reads containing the SNPs at the SNP positions were better correlated to the depth at those positions (average $R = 0.91$) (Table S2). When the coverage of the H7N9 genome was examined, the correlation (R) of numbers of identified SNP locations in samples and average genome coverages was 0.70 (Fig. 1C). Table 3 suggests that the differences between samples in the numbers of called SNPs were not primarily due to differences in read depth or coverage between samples. Comparing H7N9 sequences from the 16 patients to the reference genome of the consensus sequence of all human H7N9 strains in China in 2013, 509 SNPs, representing 454 unique nucleotide positions, were found. Among them, there were 120 SNP positions shared in two or more samples. Selected amino acid substitutions with known or suspected functional significance with respect to the IAV H7N9 strains observed during the epizootic waves in China (7, 8, 11) from the 16 patient samples analyzed here are listed in Table 5.

Hemagglutinin gene sequences and SNPs from the H7N9 patient samples. As shown in Table 5, a number of nonsynonymous SNPs in the hemagglutinin (HA) gene were observed in the patient samples compared to consensus sequences created from all available epizootic H7N9 virus sequences isolated from birds in China in 2013, 2014, and 2015 or compared to all available H7N9 sequences isolated from spillover infections in humans in China in 2013, 2014, and 2015. The 2013 avian H7N9 and human H7N9 consensus HA sequences differed only at codon 235 (aligning to H3 HA codon 226), and this codon forms a part of the HA receptor binding domain (27, 28), with the human consensus sequences from 2013, 2014, and 2015 containing a Q235L mutation. In addition, the avian consensus sequences from 2014 and 2015 also contain a Q235L mutation, which indicates the evolution of this site in avian hosts as well. Among the patient samples, 7 cases (cases 1, 2, 4, 13, 15, 16, and 17) had a Q235L mutation (Table 5; see also Table S1) and 4 cases (cases 5, 6, 19, and 20) had coverage of fewer than 100 reads at this codon, but all of them included Q235L, and the remaining samples had no coverage at HA codon 235. Structural and glycan array analyses have demonstrated that 235L enhances H7 HA binding to human-type α 2-6 sialic acids (27, 28), while *in silico* modeling has suggested enhanced binding with Q235I (29). A recent study reported that 3 mutations, H7 HA V195K (equivalent to H3 HA V186K), K202T (H3 K193T), and G237S (H3 G228S), in the receptor binding domain of H7N9 viruses switched specificity from avian-type α 2-3 to human-type α 2-6 sialic acids (30). In the patient samples, all matched the consensus, V195, K202, and G237.

Many additional HA SNPs were also seen in this patient cohort (Table 5). To evaluate the relationship between our samples and other H7N9 sequences in China, we constructed approximately maximum-likelihood phylogenetic trees based on 417 unique human HA sequences from isolates collected in China from 2013 to 2017 together with the consensus sequences (with SNP calls at bases over 50%) from the current study (Fig. 2). Previous studies demonstrated evolutionary diversification of H7N9 HA sequences (7), and with the emergence of the fourth and fifth waves (6, 8), two major lineages have been observed and defined as Yangtze River Delta and Pearl River Delta lineages. Yangtze River Delta lineage isolates have demonstrated reduced cross-reactivity with previously produced candidate vaccine strain H7N9 viruses (6). From the HA trees, it is noted that the sample sequences diversified into different clades, including the Yangtze River Delta lineage (sample 17), despite the fact that the patients were all local residents, treated at only one hospital in Shanghai, with the latest sample from April 2015.

Notable here is the HA sequence from sample 17, representing a wave 3 case from January 2015, which is phylogenetically very similar to the antigenically variant Yangtze River Delta H7N9 sequences isolated in 2016 to 2017 in waves 4 and 5 (8). SNPs

TABLE 5 Selected amino acid substitutions observed in H7N9 case samples^a

H7N9 protein ^b	Codon substitution ^c	Amino acid substitution						Function/region	Case no.
		Avian H7N9 consensus 2013	Avian H7N9 consensus 2014	Avian H7N9 consensus 2015	Human H7N9 consensus 2013	Human H7N9 consensus 2014	Human H7N9 consensus 2015		
HA	V104I	V	V	V	V	V	V	HA head	16
	S136N	S	S	N	S	S	N	HA head	17
	A143V	A	A	V	A	A	V	HA head	17
	R148K	R	R	R	R	R	R	HA head	17
	L186I	L	L	I	L	L	I	HA receptor binding	17
	Q235L	Q	L	L	L	L	L	HA receptor binding	1, 2, 4, 13, 15, 16, 17
	R270K	R	R	R	R	R	R	HA head	15, 16
	L394I	L	L	L	L	L	L	HA stalk	8, 13, 14, 15, 16
	E396A	E	E	A	E	E	A	HA stalk	17, 19, 20
	S499R	S	S	R	S	S	R	HA stalk	17, 19, 20
	N551S	N	N	N	N	N	N	HA cytoplasmic tail	1, 2, 3, 6, 8, 13, 14, 15, 16, 20
	G552R	G	G	G	G	G	G	HA cytoplasmic tail	1, 2, 3, 6, 8, 13, 14, 15, 16, 20
	NA	I16T	I	I	I	I	I	T	NA transmembrane domain
T65A		T	T	T	T	T	T	NA stalk	17
R148K		R	K	R	R	R	R	NA catalytic site	1
S242P		S	S	P	S	S	P	NA head	17
E263D		E	E	E	E	E	E	NA head	15, 16
R289K		R	R	R	R	R	R	NA catalytic site	16
N322S		N	N	S	N	N	S	NA head	17
PB2		M473V	M	M	M	M	M	M	Nuclear localization signal
	M535L	M	L	L	M	M	L	T cell epitope	6, 13, 16, 17
	N559T	T	T	T	N	T	T	T cell epitope	17
	E627K	E	E	E	K	K	K	C-terminal region; pathogenicity determinant	1
	N759K	N	N	N	N	N	N	C-terminal region	1, 2
PB1	K237R	K	K	K	K	K	K	T cell epitope	9
PB1-F2	No changes								
PA	D67Y	D	D	D	D	D	D	T cell epitope	13
	I94T, I94V	I	I	I	I	I	I	N-terminal domain	8: 94T, 7: 94V
	Q556R	Q	Q	Q	Q	Q	Q	PB1-binding region, T cell epitope	6
PA-X	D67Y	D	D	D	D	D	D	N-terminal domain	13
	I94T, I94V	I	I	I	I	I	I	N-terminal domain	8: 94T, 7: 94V
	L194P	L	P	P	L	P	P	C-terminal domain	13, 15, 16, 17
	R248K	R	K	R	R	R	R	C-terminal domain	16, 17
NP	R348K	R	R	R	R	R	R	NP oligomerization domain	19
	L466F	L	L	L	L	L	L	PB2 binding domain	14, 16
M1	N82T	N	N	N	N	N	N	Membrane binding domain	14, 16
	F109L	F	F	F	F	F	F	Membrane binding domain	4, 5, 6
	D156E	D	D	D	D	D	D	Membrane binding domain	3, 7
M2	P10L	P	L	L	P	P	L	Ectodomain	2, 13, 15
NS1	R67Q	R	R	R	R	R	R	RNA binding domain	14, 16
	S114P	S	S	S	S	S	S	Effector domain	8, 13, 14, 16
	P213S	P	P	P	P	P	P	C-terminal protein interaction domain	14, 16
NEP/NS2	No changes								

^aConsensus sequences were constructed from all full-length open reading frames available from avian and human H7N9 sequences from 2013, 2014, and 2015 downloaded from the NCBI Influenza Virus Database (<https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=database>).

^bH7N9 proteins, including hemagglutinin, are numbered from the initial methionine as codon 1.

^cNonsynonymous mutations observed according to the the H7N9 case criteria stated in Materials and Methods compared to the reference sequence (the 2013 human H7N9 consensus sequence), except for HA codon 235, where changes from the avian consensus are noted.

identified in sample 17 were notable for the following amino acid changes on the HA head (Table 5): S136N, A143V, R148K, and L186I. These amino acids are highlighted in the head region of the crystal structure of A/Anhui/1/2013 (H7N9) (4R8W) (Fig. 3). While antigenic epitopes on H7 subtype HAs have not been mapped, a computational

Pearl River Delta

Yangtze River Delta

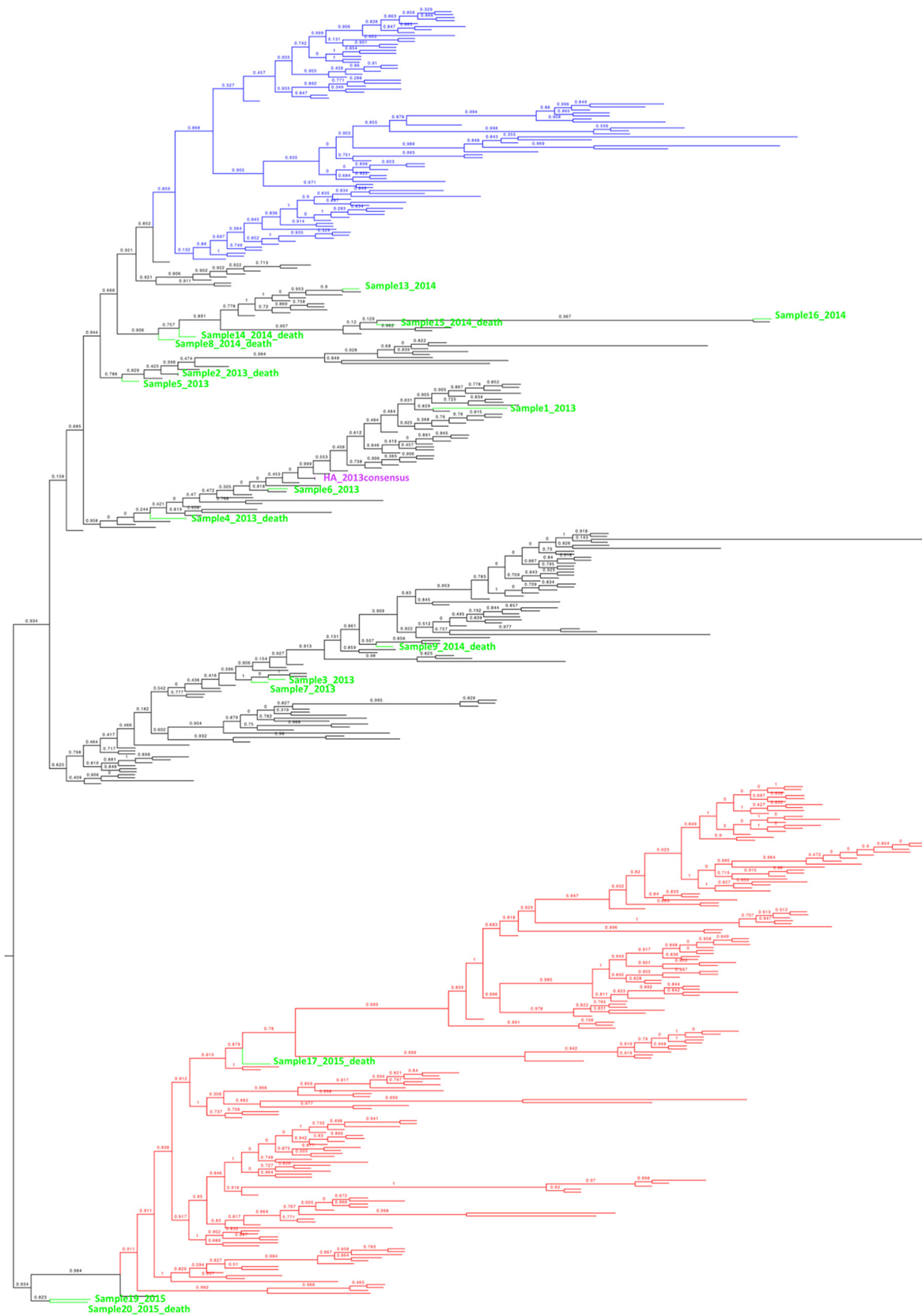


FIG 2 Phylogenetic trees of unique HA segments of human H7N9 in China from 2013 to 2017 and consensus HA sequences (SNP frequency, >50%) of our samples. Approximately maximum-likelihood midpoint-rooted phylogenetic trees with Shimodaira-Hasegawa (SH) values are presented. Purple, reference genome; green, our samples in “sample number_year_outcome” format; blue, viruses in the Pearl River Delta HA lineage; red, viruses in the Yangtze River Delta HA lineage. See Fig. S1 for an expanded form of this tree, in which individual viral taxa are labeled by strain name.

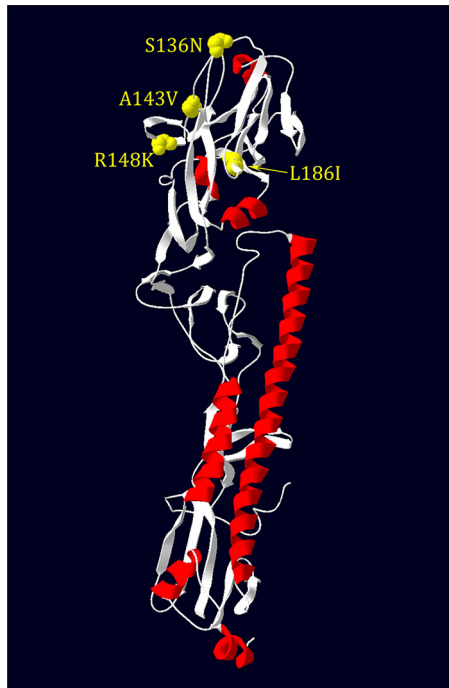


FIG 3 The structure of HA with mutations observed on head regions of sample 17 shown on an H7 HA monomer (A/Anhui/1/2013; see Materials and Methods).

analysis of H7 HA identified several putative epitopes on the HA head. The first three mutations, S136N, A143V, R148K, are near the T130A (H7 numbering) identified in epitope A (31), while the fourth site, L186I, is between the D183S and I188V changes identified in epitopes C and D, respectively. Computational modeling also identified HA receptor binding mutation Q235L/I as an antigenic epitope (epitope D). Interestingly, these amino acid changes are seen in 2016 to 2017 Yangtze River Delta lineage sequences and are likely to contribute to the antigenic evolution observed in these H7N9 strains compared to earlier human 2013 isolates (6). Changes at H7 HA codons 394 and 396 have been detected in 8 patient samples (samples 8, 13, 14, 15, 16, 17, 19, 20), which were located in the first alpha helix of the HA2 domain. The E396A mutation, in particular, was also observed in more recent Yangtze River Delta lineage isolates, and modeling done here suggests that this change may affect binding of broadly neutralizing stalk antibodies (Fig. 4).

The remaining viral gene sequences and SNPs from the H7N9 patient samples.

A number of nonsynonymous SNPs were seen in the H7N9 sample neuraminidase (NA) sequences (Table 5), including a R148K change in the catalytic site in sample 1 (equivalent to R152K in N2 subtype NA) (32). At least 15 of the 20 patients were treated with antiviral drugs, primarily oseltamivir (Table 1); however, only one sample (sample 16) showed a change associated with neuraminidase inhibitor (NAI) resistance in these samples (33), which was the most commonly observed resistance mutation, R289K (R292K in N2 subtype NA), with this SNP present at 41% (Table 5). The conserved N9 amino acids associated with the NA sialic acid-binding site (hemabsorption domain) did not differ in the samples studied (34).

SNPs were also observed in the H7N9 polymerase proteins. In PB2, 21 nonsynonymous SNPs were observed in the samples that differed from the reference sequence, including sample 1, which had a K627E change. It has been reported that 627K is associated with human adaptation of avian-origin PB2 (35) and that it was observed in the 1918 pandemic (36) and in some highly pathogenic H5N1 human isolates (37). The human H7N9 PB2 consensus sequence differs from the avian H7N9 consensus sequence at this site as well (Table 5). None of the samples differed at sites 701 and 702,

which have been reported to be associated with mammalian adaptation (38). One change was observed in PB1 in sample 9, and 3 changes were observed in samples 6, 7, 8, and 13 in polymerase acidic protein (PA). PA codon 100 is typically a valine in avian IAV sequences, and a V100A change is associated with human adaptation (36). Of the 2013 avian and human H7N9 consensus sequences constructed here, both correspond to V100, and no nonsynonymous SNPs were observed at this codon in all the samples sequenced here. SNPs associated with nonstructural protein PA-X (39) were observed in the samples, but the biological significance of these changes remains unclear. Several other samples had nonsynonymous SNPs in M1, M2, and NS1 (Table 5).

To determine which of the nonsynonymous SNPs could be associated with human adaptation and which were observed during two or more H7N9 epidemic waves, Venn analysis was performed on the translated proteins from nonsynonymous SNPs identified in Table 5 (see also Table S1) compared to the 2013 avian H7N9 protein consensus sequences. As shown in Fig. 5, 15% of H7N9 amino acid changes were identified only in wave 1, 36% were identified only in wave 2, and 28% were unique to wave 3. The populations of amino acid changes identified in only a single wave were dominated by sequences identified at a frequency of <10% in previous human H7N9 cases (black text in the figure), novel human mutations (green text), and one sequence, PB2 K627E (blue text), that had been previously identified in >10% of sequences in the Influenza Research Database as of 19 June 2017. Interestingly, only a single SNP (M2 P10L) was identified in both waves 1 and 2, while 3 SNPs (8% of the total) were present in waves 2 and 3, two of which represented changes in the C-terminal domain of host response regulatory protein PA-X. Finally, 4 SNPs (10% of the total) were identified in all three waves and included changes in HA receptor binding (Q235L), the HA cytoplasmic tail (N551S and G552R), and viral polymerase subunit PB2 (M535L). Of these changes, HA N551S and HA G552R were previously identified in H7N9 human sequences at low (<10%) frequencies, while PB2 M535L was previously identified in $\geq 10\%$ of human H7N9 sequences. Significantly, the mutation Q235L in the HA receptor binding domain was present in all three waves and represented a change toward the human consensus sequence (denoted with red text in Fig. 5).

Evidence for reassortment in the H7N9 patient samples. Phylogenetic analyses were performed by constructing gene segment trees using the approximately maximum-likelihood method (40). Trees were inferred using the compiled H7N9 viral gene segment consensus sequences (proportion of observed SNPs, more than 50%) from the 16 patient samples along with 417 H7N9 viral genomic sequences downloaded from GISAID (<http://platform.gisaid.org/epi3/frontend#5277b0>). The eight trees are shown in Fig. S1 in the supplemental material. Most trees showed two major clades, often with two subclades in each major clade. Taxa were colored red in the figure for viruses possessing a Yangtze River Delta clade hemagglutinin and blue for viruses possessing a Pearl River Delta clade hemagglutinin. Evidence of gene segment reassortment is apparent among the 8 trees; for example, several viruses with Pearl River Delta clade HA genes (colored in blue) are contained in the Yangtze River Delta clade (colored in red) of the neuraminidase tree. Following the taxa within the trees, it is possible to observe frequent reassortment within each gene segment, as previously reported (41).

DISCUSSION

Human infections with avian IAV H7 subtype viruses are rare but have been reported before and are usually associated with exposure to poultry (42–44). Most of the previously reported cases of H7 IAV infection in humans have caused only mild or moderate illness (43–45), except for one fatal case resulting from a highly pathogenic avian H7N7 IAV infection in the Netherlands (46). In contrast, the current avian IAV H7N9 epizootic outbreak in China has been much more severe, with more than 90% of the confirmed H7N9 patients being hospitalized with pneumonia or respiratory failure and with a CFR of 39%. This suggests that the 2013 lineage avian H7N9 virus may be more pathogenic in humans than the H7 strains responsible for past IAV epizootic

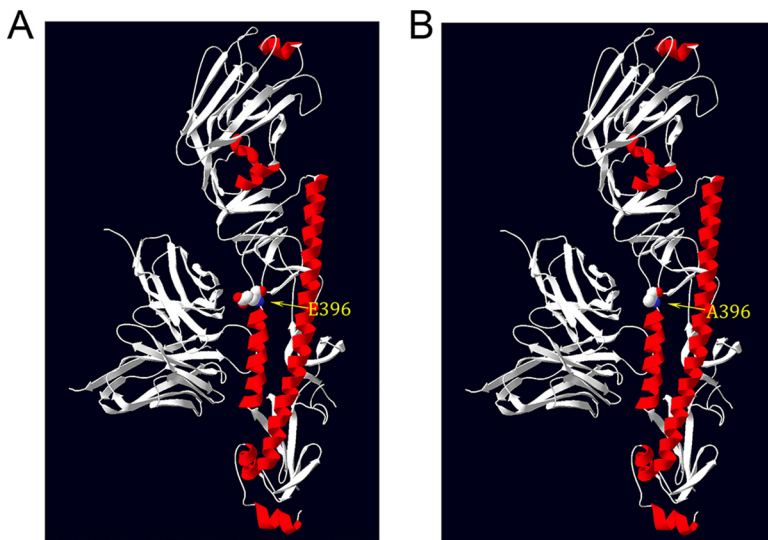


FIG 4 The structure of the HA A396E mutation at the HA stalk related to stalk neutralizing antibody CT149. (A) HA396E. (B) HA396A shown on an H7 HA monomer (A/Anhui/1/2013; see Materials and Methods).

outbreaks (47). Human H7N9 virus infections have also exhibited age-specific and sex-specific morbidity and mortality patterns. The median age of the confirmed H7N9 patients was 58 years. A very few H7N9 cases have occurred in children, teenagers, and young adults. Further, most of the severe and fatal cases occurred in middle-aged and older men rather than in women (48). However, females in their reproductive years have displayed an increased tendency to die of H7N9 influenza in comparison to males (49). In this patient cohort, the observed CFR was high, with an overall fatality rate of at least 55% (some patient outcomes were not available); however, no correlation could be made between outcome and the SNPs identified.

In this study, all 20 patient samples were sequenced and mapped to the H7N9 reference genome. On the basis of the negative-control results with respect to reads mapped to H7N9 genome, 4 samples were eliminated from further analysis. Among the remaining 16 samples, the total numbers of mapped reads ranged from 105,000 to over 178 million and the range of H7N9 genome coverage was 57.7% to 99.5% (Table 3). The differences in the mapped read numbers and coverage percentages depended on the viral RNA concentration and the quality of viral RNA in the different samples. Viral RNA degradation during sample preservation and/or library preparation would also likely decrease coverage.

Previous studies have demonstrated that human host factors are likely important in determining outcome during influenza virus infections. For example, more than 200 host factors are required for influenza virus replication as demonstrated by genome-wide RNA interference screening (50, 51). Other studies showed that interferon-inducible transmembrane 3 (IFITM3) may be critical for defending the host against IAV *in vivo* (52, 53). Even administration of influenza vaccinations at different times of day has been shown to affect the development of host antibody responses (54). A recent study of whole-genome exon capture and sequencing of 18 H7N9 patients identified 21 genes that were highly associated with H7N9 influenza virus infection, some of which may be associated with H7N9 influenza susceptibility (55). Thus, different hosts, different health conditions, and even different times of infection could affect the evolution of IAV.

In addition, when the SNPs identified in the current study were compared to the precomputed SNPs called from human H7N9 sequences in the Influenza Research Database (25), we found that 39.2% (178/454) of the SNPs from the samples had not been previously reported. One reason for this could be that deep-sequencing technol-

Novel human
Human identified (<10%)
Human identified (≥10%)
Human consensus

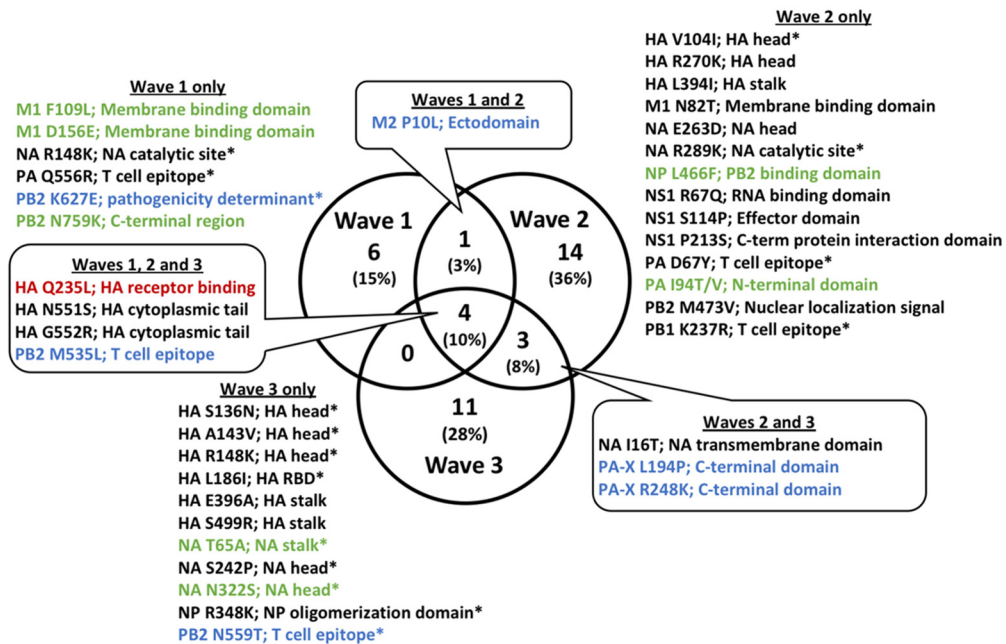


FIG 5 Venn diagram of protein sequences identified in three H7N9 waves between 2013 to 2015 compared to the avian consensus protein sequence. Set analysis was performed on translated nonsynonymous SNPs identified in waves 1, 2, and 3. Novel mutations identified in this study are shown in green; SNPs previously identified in <10% or ≥10% of sequences in the Influenza Research Database are indicated by black text or blue text, respectively. Changes from the avian to the human consensus sequences are shown in red. Percentages of protein sequence changes were rounded to nearest whole-number value. SNPs identified in only a single patient are marked with an asterisk (*). The Venn diagram was created using Venny 2.1 (<http://bioinfogp.cnb.csic.es/tools/venny/>).

ogy was applied to direct sequencing of virus from clinical samples here, which allowed investigation of infected viral populations, including identification of quasispecies. These methods are now becoming more common. Historically, the majority of viral sequences in IRD or GenBank were likely initially cultured and then subjected to PCR amplification and Sanger sequencing to generate consensus genome sequences, some of which may have also acquired culture adaptations (56).

As seen in other human H7N9 isolates, the HA sequences determined in this study showed the change in the receptor binding domain, Q235L/I (H3 Q226L/I) (28, 29, 57, 58), that has been reported to enhance binding to human-type α 2-6 sialic acids (59). Among all the 16 samples examined here, 11 samples with read coverages possessed Q235L, whereas, among the 417 unique HA sequences downloaded from GISAID, only 9 strains possessed Q235I. It will be interesting to see if this mutation is more commonly observed in future human H7N9 isolates. Of the 3 mutations recently modeled *in vitro* to switch HA receptor specificity to α 2-6 sialic acids (30), none were observed in this study.

Sample 17 possessed several nonsynonymous SNPs on the HA head domain that are shared with antigenically variant Yangtze River Delta lineage viruses isolated in waves 4 and 5 in 2016 to 2017 (Fig. 3) and likely reflect some changes related to the antigenic evolution in these H7N9 viruses compared to earlier (2013) isolates (6). Thus, compared to conventional sequencing methods, deep-sequencing strategies may aid in the identification of variant viruses with potentially significant changes associated with human adaptation or with antigenicity acquired more readily and earlier. Since there is no evidence of efficient human-to-human transmission or significant population immunity against H7N9 in humans, antigenic drift, as observed in sample 17 and related

Yangtze River Delta lineage viruses, is likely occurring in the avian reservoir and may reflect widespread use of IAV vaccines in poultry in China (60).

Another nonsynonymous HA mutation identified in the samples is E396A in the HA stalk region. This mutation was observed at a rate of over 98% in samples 17, 19, and 20. Structural modeling based on the cocrystal structure of the H7 HA with the stalk neutralizing antibody CT149 (61) supports the idea that the E396A mutation could affect HA stalk antibody binding in this region (Fig. 4). Among the 417 unique HA sequences downloaded from GISAID, 171 possessed an A396; 60 of those strains were from the year 2017, 43 from 2016, 50 from 2015, and 18 from 2014. This might indicate an antigenic drift trend, but additional experiments are needed to evaluate whether this change would affect broadly neutralizing stalk antibodies.

In this study, 15 of our 20 patients were treated with oseltamivir (Table 1). It has been reported that neuraminidase mutations E119V, I222V, D197N/E, and R292K (N2 numbering) have been commonly detected in H3N2 and type B viruses (62–66) after neuraminidase inhibitor (NAI) treatment. Moreover, corresponding mutations E119V, I222K, and R292K in H7N9 reduced the inhibitory activity of oseltamivir in ferrets (67). In our samples, only R289K (N2 R292K) was detected in patient 16 (59% R and 41% K) (Table 5), but the antiviral treatment status of this patient is unknown. The samples from 3 patients (patients 14, 19, and 20) did not have adequate read coverage at this position, but the remaining samples were all R289 despite most of the patients (except patients 10, 11, 13, and 14, who had incomplete medical records) having received oseltamivir treatment (Table 1). Even lowering the SNP call stringency to a 5% difference compared to the reference, no additional oseltamivir resistance mutations were found in the patient samples. The explanation could be the sampling took place in each case on or right after the dates of the antiviral treatments in all our patients, except patient 12, whose sampling date was 8 days after the antiviral treatment date.

It has been shown that some amino acid changes in the viral polymerase complex of IAV are very important for mammalian adaptation. For example, PB2 E627K and D701N changes are critical for mammalian adaptation (35–37, 68, 69) and the changes could increase polymerase activity and viral replication in mammalian cells and pathogenicity of H7N9 viruses in mice (70–72). It was also found that E627K and D701N mutations in H7N9 could readily occur during transmission of the virus among ferrets via direct physical contact (73). Among our samples, only sample 1 displayed a mixture of E627K mutations (68% E, 32% K). For the PB2 701 site, the samples from all patients with adequate coverage at this site possessed D701.

Interestingly, among the samples, 4 of the key SNPs were present in patients from all 3 waves (Fig. 5), including changes in the viral polymerase subunit PB2 (M535L). Other than being mapped to T cell epitopes, it is not known whether these observed PB2 changes or the change in NP would enhance replication or transmissibility in mammals. Further studies are needed to understand the significance of these changes.

Critical for pre-pandemic surveillance efforts is timely recognition of the emergence of mutations that indicate the appearance of a viral genotype with an increased capacity for forward transmission. In addition to the polymerase mutations that facilitate mammalian viral replication, such as PB2 E627K (35, 74), as well as the changes in the HA receptor binding domain that increase the affinity for α 2-6 sialic acids of the human upper respiratory tract (75, 76), our analysis showed that the PB2 M535L change was identified in multiple patients and was present during all three H7N9 waves between 2013 and 2015. Additionally, changes in the HA cytoplasmic tail (N551S and G552R) were also identified in these three successive waves. The cytoplasmic tail of HA is a site for cysteine acylation (77, 78) and plays an important role during infection by modulating membrane fusion activity during viral assembly and budding (79–81). While it is unclear what role N551S and G552R play in infection, it is tempting to speculate that these changes could be associated with increased viral budding in humans either through acylation-dependent or -independent processes.

To our knowledge, this is the first report of deep-sequencing analysis of H7N9 IAV directly from clinical samples of patients from three consecutive years without initial

viral culture or influenza virus-specific RT-PCR analyses. Applying whole-transcriptome amplification methods, specific enrichment for H7N9 influenza reads, and deep-sequencing methods, we were able to both investigate the consensus sequences of the viral genome and identify low-abundance SNPs. A possible advantage of the methods described in this report, including enrichment of viral RNA by specific probe hybridization, is the ability to obtain sequences from samples with low amounts of or degraded viral RNA. Although the samples analyzed in the present study were all from patients treated at a single hospital in Shanghai, SNP analyses showed that the viruses contained signatures of different evolving H7N9 clusters (6–8), including a virus isolate from a patient infected in 2015 with likely antigenic changes to the HA that are shared with more recently isolated Yangtze River Delta lineage isolates. In addition, we identified 178 SNPs with 106 nonsynonymous changes in H7N9 that have not been reported in the Influenza Research Database. However, further investigation is still needed to determine the function of these newly identified SNPs with respect to virus replication, host adaptation, transmission, or pathogenesis.

MATERIALS AND METHODS

Patient information and sample collection. The samples from 20 patients infected by avian influenza virus H7N9 studied here were collected in Shanghai, China, between 2013 and 2015. All 20 patients were diagnosed with severe community-acquired pneumonia (CAP). Throat swabs were collected from each patient during the inpatient stay and stored in viral transport medium (VTM) (Yocan, China). RNA was extracted from the swab samples using a nucleic acid extraction kit (QIAamp viral RNA minikit; Qiagen, USA). The presence of influenza A/H7N9 virus RNA was tested using the RT-PCR method recommended by the World Health Organization (82). The samples were collected and processed for viral sequence analysis following elimination of personally identifiable information. The sequence analyses of H7N9 virus genomes performed at the National Institutes of Health (NIH) were determined to represent an excluded study, i.e., one not requiring Institutional Review Board (IRB) approval by the NIH Office of Human Subjects Protections (exemption number 13268).

Target enrichment probe design. Enrichment probes were designed by Agilent Technologies (Santa Clara, CA) using A/Hangzhou/1/2013 (H7N9) as the reference sequence. The biotin-labeled RNA probes were 120 bases in length with 5-base spacing density, resulting in a total of 2,034 overlapping probes spanning the H7N9 genome.

Library construction and sequencing. For negative-control experiments, HeLa cell total RNA was obtained from TaKaRa (TaKaRa Bio USA, Inc., Mountain View, CA). For positive-control experiments, the hemagglutinin (HA) gene segment (KC853766.1) from A/Hangzhou/1/2013 (H7N9) was synthesized and cloned into the pBluescript II KS(-) SacI site by GenScript (Piscataway, NJ). Plasmid was linearized by NotI (NEB, Ipswich, MA) digestion, and RNA was transcribed using a HiScribe T7 high-yield RNA synthesis kit (NEB, Ipswich, MA) following the manufacturer's instructions. RNA from the HeLa cells, from the T7 promoter-transcribed HA plasmid, and from the isolated total RNA from patient sample throat swabs was amplified using Ovation RNA-Seq system V2 from NuGEN (NuGEN, San Carlos, CA) following the kit specifications. Each sample, containing 5 μ l of total RNA, was used as input for Ovation RNA-Seq system V2. The amplified total cDNAs were analyzed using an Agilent 2100 Bioanalyzer with an Agilent high-sensitivity DNA kit (Agilent Technologies, Santa Clara, CA) and sheared to 150 bp using a Covaris S2 instrument (Covaris, Woburn, MA). Following this, approximately 300 ng amplified cDNA was used to make each Illumina sequencing library using an Agilent SureSelect^{XT} target enrichment kit (Agilent, Santa Clara, CA) for Illumina multiplex sequencing and the protocol for 200-ng DNA samples. Enriched Illumina sequencing libraries were analyzed on an Agilent 2100 Bioanalyzer using an Agilent high-sensitivity DNA kit. Libraries were then clustered on an Illumina cBot instrument and sequenced on Illumina GAIIX or HiSeq machines according to the instructions of the manufacturer (Illumina, San Diego, CA). In this study, more than 2.8 billion reads and a total of more than 323 Gb of sequences were generated.

Data analysis. A consensus sequence derived from all available H7N9 human strains obtained in China in 2013 was used as the reference genome and was generated at the Influenza Virus Resource site (<https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=databse>) (83) on 16 November 2017. This reference sequence is called the 2013 human H7N9 consensus in this study (the sequences are listed in Text S1 in the supplemental material). Amino acid consensus sequences from avian and human H7N9 isolates from 2013, 2014, and 2015 were generated similarly (the sequences are listed in Text S1). Reads were mapped to Bowtie2 (version 2.2.5; <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) indexed to the consensus sequence of all H7N9 human strains in China in the 2013 genome using Tophat2 (release 2.0.13; <http://ccb.jhu.edu/software/tophat/index.shtml>) downloaded from the Center for Computational Biology, Johns Hopkins University (<http://ccb.jhu.edu/>) (84). SAMtools mpileup (version 2.1.0) was used to make SNP calls, and the minimum base Phred quality score was 25 (85). The reported SNP calls needed to satisfy the indicated criteria at the SNP position, similarly to those described in a previous study (86), were as follows: (i) more than 100 reads at that position; (ii) reads with SNP calls from both read directions; (iii) a SNP call not at the first or last base of all the SNP-containing reads; and (iv) different bases representing more than 10% of the aligned reads.

Multiple-sequence alignment was performed using MUSCLE (87) version 3.8.31, and midpoint-rooted phylogenetic trees using the approximately maximum-likelihood method were constructed with Fast-Tree version 2.1.8 with the Jukes-Cantor + CAT model (40) and displayed in FigTree version 1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>) and included Shimodaira-Hasegawa test values. Other human H7N9 sequences from the outbreak samples collected in China in 2013 to 2017 and used in this study were downloaded from the Global Initiative on Sharing All Influenza Data (GISAID; <http://platform.gisaid.org/epi3/frontend#355fbd>) database on 6 April 2017. Identical sequences from the HA segments were eliminated before alignment. The strain names used in the HA tree were used to identify the sequences to make phylogenetic trees for other IAV H7N9 gene segments. In the sample sequences, at nucleotide positions with read coverage but with fewer than 100 reads, bases from the precomputed human H7N9 consensus sequence from each segment were used to assemble contigs.

Protein structure. The crystal structure of the H7 hemagglutinin from A/Anhui/1/2013 in complex with neutralizing antibody CT149 (4R8W) (61) was downloaded from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) (<http://www.rcsb.org/pdb/home/home.do>) (88), and HA amino acid changes were made and viewed with Swiss-PdbViewer (89) or with the protein structure viewer in the Influenza Research Database (90).

Accession number(s). All sequences have been deposited as a series with accession no. [SRR7444529](https://www.ncbi.nlm.nih.gov/nuccore/SRR7444529) to [SRR7444548](https://www.ncbi.nlm.nih.gov/nuccore/SRR7444548) at the GenBank SRA database.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mSphereDirect.00462-18>.

TEXT S1, DOCX file, 0.03 MB.

FIG S1, PDF file, 0.4 MB.

TABLE S1, XLSX file, 0.03 MB.

TABLE S2, XLSX file, 0.1 MB.

ACKNOWLEDGMENTS

This work was supported in part by the Intramural Research Programs of the National Institutes of Health and National Institute of Allergy and Infectious Diseases and by the CAMS Innovation Fund for Medical Sciences (2016-I2M-1-014).

REFERENCES

- Taubenberger JK, Kash JC. 2010. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* 7:440–451. <https://doi.org/10.1016/j.chom.2010.05.009>.
- Morens DM, Taubenberger JK, Folkers GK, Fauci AS. 2010. Pandemic influenza's 500th anniversary. *Clin Infect Dis* 51:1442–1444. <https://doi.org/10.1086/657429>.
- Johnson NP, Mueller J. 2002. Updating the accounts: global mortality of the 1918–1920 “Spanish” influenza pandemic. *Bull Hist Med* 76:105–115. <https://doi.org/10.1353/bhm.2002.0022>.
- Dawood FS, Iuliano AD, Reed C, Meltzer MI, Shay DK, Cheng PY, Bandaranayake D, Breiman RF, Brooks WA, Buchy P, Feikin DR, Fowler KB, Gordon A, Hien NT, Horby P, Huang QS, Katz MA, Krishnan A, Lal R, Montgomery JM, Molbak K, Pebody R, Presanis AM, Razuri H, Steens A, Tinoco YO, Wallinga J, Yu H, Vong S, Bresee J, Widdowson MA. 2012. Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. *Lancet Infect Dis* 12:687–695. [https://doi.org/10.1016/S1473-3099\(12\)70121-4](https://doi.org/10.1016/S1473-3099(12)70121-4).
- Gao R, Cao B, Hu Y, Feng Z, Wang D, Hu W, Chen J, Jie Z, Qiu H, Xu K, Xu X, Lu H, Zhu W, Gao Z, Xiang N, Shen Y, He Z, Gu Y, Zhang Z, Yang Y, Zhao X, Zhou L, Li X, Zou S, Zhang Y, Li X, Yang L, Guo J, Dong J, Li Q, Dong L, Zhu Y, Bai T, Wang S, Hao P, Yang W, Zhang Y, Han J, Yu H, Li D, Gao GF, Wu G, Wang Y, Yuan Z, Shu Y. 2013. Human infection with a novel avian-origin influenza A (H7N9) virus. *N Engl J Med* 368:1888–1897. <https://doi.org/10.1056/NEJMoa1304459>.
- Iuliano AD, Jang Y, Jones J, Davis CT, Wentworth DE, Uyeki TM, Roguski K, Thompson MG, Gubareva L, Fry AM, Burns E, Trock S, Zhou S, Katz JM, Jernigan DB. 2017. Increase in human infections with avian influenza A(h7n9) virus during the fifth epidemic - China, October 2016–February 2017. *MMWR Morb Mortal Wkly Rep* 66:254–255. <https://doi.org/10.15585/mmwr.mm6609e2>.
- Wang D, Yang L, Zhu W, Zhang Y, Zou S, Bo H, Gao R, Dong J, Huang W, Guo J, Li Z, Zhao X, Li X, Xin L, Zhou J, Chen T, Dong L, Wei H, Li X, Liu L, Tang J, Lan Y, Yang J, Shu Y. 2016. Two outbreak sources of influenza A (H7N9) viruses have been established in China. *J Virol* 90:5561–5573. <https://doi.org/10.1128/JVI.03173-15>.
- World Health Organization. 2018. Influenza at the human animal interface: summary and assessment, 8 December 2017 to 25 January 2018. www.who.int/influenza/human_animal_interface/Influenza_Summary_IRA_HA_interface_25_01_2018_FINAL.pdf?ua=1. Accessed 1 Mar 2018.
- Wu A, Su C, Wang D, Peng Y, Liu M, Hua S, Li T, Gao GF, Tang H, Chen J, Liu X, Shu Y, Peng D, Jiang T. 2013. Sequential reassortments underlie diverse influenza H7N9 genotypes in China. *Cell Host Microbe* 14:446–452. <https://doi.org/10.1016/j.chom.2013.09.001>.
- Cui L, Liu D, Shi W, Pan J, Qi X, Li X, Guo X, Zhou M, Li W, Li J, Haywood J, Xiao H, Yu X, Pu X, Wu Y, Yu H, Zhao K, Zhu Y, Wu B, Jin T, Shi Z, Tang F, Zhu F, Sun Q, Wu L, Yang R, Yan J, Lei F, Zhu B, Liu W, Ma J, Wang H, Gao GF. 2014. Dynamic reassortments and genetic heterogeneity of the human-infecting influenza A (H7N9) virus. *Nat Commun* 5:3142. <https://doi.org/10.1038/ncomms4142>.
- Chen GW, Kuo SM, Yang SL, Gong YN, Hsiao MR, Liu YC, Shih SR, Tsao KC. 2016. Genomic signatures for avian H7N9 viruses adapting to humans. *PLoS One* 11:e0148432. <https://doi.org/10.1371/journal.pone.0148432>.
- Vemula SV, Zhao J, Liu J, Wang X, Biswas S, Hewlett I. 2016. Current approaches for diagnosis of influenza virus infections in humans. *Viruses* 8:96. <https://doi.org/10.3390/v8040096>.
- Goodwin S, McPherson JD, McCombie WR. 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 17:333–351. <https://doi.org/10.1038/nrg.2016.49>.
- Nakamura S, Yang CS, Sakon N, Ueda M, Tougan T, Yamashita A, Goto N, Takahashi K, Yasunaga T, Ikuta K, Mizutani T, Okamoto Y, Tagami M, Morita R, Maeda N, Kawai J, Hayashizaki Y, Nagai Y, Horii T, Iida T, Nakaya T. 2009. Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. *PLoS One* 4:e4219. <https://doi.org/10.1371/journal.pone.0004219>.
- Ghedini E, Laplante J, DePasse J, Wentworth DE, Santos RP, Lepow ML, Porter J, Stellrecht K, Lin X, Operario D, Griesemer S, Fitch A, Halpin RA,

- Stockwell TB, Spiro DJ, Holmes EC, St George K. 2011. Deep sequencing reveals mixed infection with 2009 pandemic influenza A (H1N1) virus strains and the emergence of oseltamivir resistance. *J Infect Dis* 203: 168–174. <https://doi.org/10.1093/infdis/jiq040>.
16. Zhou B, Lin X, Wang W, Halpin RA, Bera J, Stockwell TB, Barr IG, Wentworth DE. 2014. Universal influenza B virus genomic amplification facilitates sequencing, diagnostics, and reverse genetics. *J Clin Microbiol* 52:1330–1337. <https://doi.org/10.1128/JCM.03265-13>.
 17. Laassri M, Majid L, Zagorodnyaya T, Plant EP, Petrovskaya S, Bidzhieva B, Ye Z, Simonyan V, Chumakov K. 2015. Deep sequencing for evaluation of genetic stability of influenza A/California/07/2009 (H1N1) vaccine viruses. *PLoS One* 10:e0138650. <https://doi.org/10.1371/journal.pone.0138650>.
 18. Dinis JM, Florek NW, Fatola OO, Moncla LH, Mutschler JP, Charlier OK, Meece JK, Belongia EA, Friedrich TC. 2016. Deep sequencing reveals potential antigenic variants at low frequencies in influenza A virus-infected humans. *J Virol* 90:3355–3365. <https://doi.org/10.1128/JVI.03248-15>.
 19. Mishin VP, Baranovich T, Garten R, Chesnokov A, Abd Elal AI, Adamczyk M, LaPlante J, St George K, Fry AM, Barnes J, Chester SC, Xu X, Katz JM, Wentworth DE, Gubareva LV. 2017. A pyrosequencing-based approach to high-throughput identification of influenza A(H3N2) virus clades harboring antigenic drift variants. *J Clin Microbiol* 55:145–154. <https://doi.org/10.1128/JCM.01840-16>.
 20. McInerney P, Adams P, Hadi MZ. 2014. Error rate comparison during polymerase chain reaction by DNA polymerase. *Mol Biol Int* 2014: 287430. <https://doi.org/10.1155/2014/287430>.
 21. Potapov V, Ong JL. 2017. Examining sources of error in PCR by single-molecule sequencing. *PLoS One* 12:e0169774. <https://doi.org/10.1371/journal.pone.0169774>.
 22. Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR. 2001. Universal primer set for the full-length amplification of all influenza A viruses. *Arch Virol* 146:2275–2289. <https://doi.org/10.1007/s007050170002>.
 23. Briese T, Kapoor A, Mishra N, Jain K, Kumar A, Jabado OJ, Lipkin WI. 2015. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *mBio* 6:e01491-15. <https://doi.org/10.1128/mBio.01491-15>.
 24. Kao WC, Chan AH, Song YS. 2011. ECHO: a reference-free short-read error correction algorithm. *Genome Res* 21:1181–1192. <https://doi.org/10.1101/gr.111351.110>.
 25. Squires RB, Noronha J, Hunt V, Garcia-Sastre A, Macken C, Baumgarth N, Suarez D, Pickett BE, Zhang Y, Larsen CN, Ramsey A, Zhou L, Zaremba S, Kumar S, Deitrich J, Klem E, Scheuermann RH. 2012. Influenza research database: an integrated bioinformatics resource for influenza research and surveillance. *Influenza Other Respir Viruses* 6:404–416. <https://doi.org/10.1111/j.1750-2659.2011.00331.x>.
 26. Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190. <https://doi.org/10.1101/gr.849004>.
 27. Xu R, de Vries RP, Zhu X, Nycholat CM, McBride R, Yu W, Paulson JC, Wilson IA. 2013. Preferential recognition of avian-like receptors in human influenza A H7N9 viruses. *Science* 342:1230–1235. <https://doi.org/10.1126/science.1243761>.
 28. Shi Y, Zhang W, Wang F, Qi J, Wu Y, Song H, Gao F, Bi Y, Zhang Y, Fan Z, Qin C, Sun H, Liu J, Haywood J, Liu W, Gong W, Wang D, Shu Y, Wang Y, Yan J, Gao GF. 2013. Structures and receptor binding of hemagglutinins from human-infecting H7N9 influenza viruses. *Science* 342:243–247. <https://doi.org/10.1126/science.1242917>.
 29. Rajapaksha H, Petrovsky N. 2014. In silico structural homology modelling and docking for assessment of pandemic potential of a novel H7N9 influenza virus and its ability to be neutralized by existing anti-hemagglutinin antibodies. *PLoS One* 9:e102618. <https://doi.org/10.1371/journal.pone.0102618>.
 30. de Vries RP, Peng W, Grant OC, Thompson AJ, Zhu X, Bouwman KM, de la Pena ATT, van Breemen MJ, Ambepitiya Wickramasinghe IN, de Haan CAM, Yu W, McBride R, Sanders RW, Woods RJ, Verheije MH, Wilson IA, Paulson JC. 2017. Three mutations switch H7N9 influenza to human-type receptor specificity. *PLoS Pathog* 13:e1006390. <https://doi.org/10.1371/journal.ppat.1006390>.
 31. Liu M, Song T, Hua S, Wu A, Jiang T. 2015. Computational analysis of antigenic epitopes of avian influenza A (H7N9) viruses. *Sci China Life Sci* 58:687–693. <https://doi.org/10.1007/s11427-015-4886-4>.
 32. Eweas AF, Abdel-Moneim AS. 2015. In-silico structural analysis of the influenza A subtype H7N9 neuraminidase and molecular docking with different neuraminidase inhibitors. *Virusdissease* 26:27–32. <https://doi.org/10.1007/s13337-014-0245-5>.
 33. Qi Y, Fan H, Qi X, Zhu Z, Guo X, Chen Y, Ge Y, Zhao K, Wu T, Li Y, Shan Y, Zhou M, Shi Z, Wang H, Cui L. 2014. A novel pyrosequencing assay for the detection of neuraminidase inhibitor resistance-conferring mutations among clinical isolates of avian H7N9 influenza virus. *Virus Res* 179:119–124. <https://doi.org/10.1016/j.virusres.2013.10.026>.
 34. Kobasa D, Rodgers ME, Wells K, Kawaoka Y. 1997. Neuraminidase hemadsorption activity, conserved in avian influenza A viruses, does not influence viral replication in ducks. *J Virol* 71:6706–6713.
 35. Subbarao EK, London W, Murphy BR. 1993. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J Virol* 67:1761–1764.
 36. Taubenberger JK, Reid AH, Lourens RM, Wang R, Jin G, Fanning TG. 2005. Characterization of the 1918 influenza virus polymerase genes. *Nature* 437:889–893. <https://doi.org/10.1038/nature04230>.
 37. Hatta M, Gao P, Halfmann P, Kawaoka Y. 2001. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. *Science* 293: 1840–1842. <https://doi.org/10.1126/science.1062882>.
 38. Gabriel G, Herwig A, Klenk HD. 2008. Interaction of polymerase subunit PB2 and NP with importin alpha1 is a determinant of host range of influenza A virus. *PLoS Pathog* 4:e11. <https://doi.org/10.1371/journal.ppat.0040011>.
 39. Jagger BW, Wise HM, Kash JC, Walters KA, Wills NM, Xiao YL, Dunfee RL, Schwartzman LM, Ozinsky A, Bell GL, Dalton RM, Lo A, Efstathiou S, Atkins JF, Firth AE, Taubenberger JK, Digard P. 2012. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* 337:199–204. <https://doi.org/10.1126/science.1222213>.
 40. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
 41. Lu J, Wu J, Zeng X, Guan D, Zou L, Yi L, Liang L, Ni H, Kang M, Zhang X, Zhong H, He X, Monagin C, Lin J, Ke C. 2014. Continuing reassortment leads to the genetic diversity of influenza virus H7N9 in Guangdong, China. *J Virol* 88:8297–8306. <https://doi.org/10.1128/JVI.00630-14>.
 42. Belser JA, Bridges CB, Katz JM, Tumpey TM. 2009. Past, present, and possible future human infection with influenza virus A subtype H7. *Emerg Infect Dis* 15:859–865. <https://doi.org/10.3201/eid1506.090072>.
 43. Koopmans M, Wilbrink B, Conyn M, Natrop G, van der Nat H, Vennema H, Meijer A, van Steenbergen J, Fouchier R, Osterhaus A, Bosman A. 2004. Transmission of H7N7 avian influenza A virus to human beings during a large outbreak in commercial poultry farms in The Netherlands. *Lancet* 363:587–593. [https://doi.org/10.1016/S0140-6736\(04\)15589-X](https://doi.org/10.1016/S0140-6736(04)15589-X).
 44. Tweed SA, Skowronski DM, David ST, Larder A, Petric M, Lees W, Li Y, Katz J, Kraiden M, Tellier R, Halpert C, Hirst M, Astell C, Lawrence D, Mak A. 2004. Human illness from avian influenza H7N3, British Columbia. *Emerg Infect Dis* 10:2196–2199. <https://doi.org/10.3201/eid1012.040961>.
 45. Ostrowsky B, Huang A, Terry W, Anton D, Brunagel B, Traynor L, Abid S, Johnson G, Kacica M, Katz J, Edwards L, Lindstrom S, Klimov A, Uyekei TM. 2012. Low pathogenic avian influenza A (H7N2) virus infection in immunocompromised adult, New York, USA, 2003. *Emerg Infect Dis* 18: 1128–1131. <https://doi.org/10.3201/eid1807.111913>.
 46. Fouchier RA, Schneeberger PM, Rozendaal FW, Broekman JM, Kemink SA, Munster V, Kuiken T, Rimmelzwaan GF, Schutten M, Van Doornum GJ, Koch G, Bosman A, Koopmans M, Osterhaus AD. 2004. Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. *Proc Natl Acad Sci U S A* 101:1356–1361. <https://doi.org/10.1073/pnas.0308352100>.
 47. Li Q, Zhou L, Zhou M, Chen Z, Li F, Wu H, Xiang N, Chen E, Tang F, Wang D, Meng L, Hong Z, Tu W, Cao Y, Li L, Ding F, Liu B, Wang M, Xie R, Gao R, Li X, Bai T, Zou S, He J, Hu J, Xu Y, Chai C, Wang S, Gao Y, Jin L, Zhang Y, Luo H, Yu H, He J, Li Q, Wang X, Gao L, Pang X, Liu G, Yan Y, Yuan H, Shu Y, Yang W, Wang Y, Wu F, Uyekei TM, Feng Z. 2014. Epidemiology of human infections with avian influenza A(H7N9) virus in China. *N Engl J Med* 370:520–532. <https://doi.org/10.1056/NEJMoa1304617>.
 48. Husain M. 2014. Avian influenza A (H7N9) virus infection in humans: epidemiology, evolution, and pathogenesis. *Infect Genet Evol* 28: 304–312. <https://doi.org/10.1016/j.meegid.2014.10.016>.
 49. Hoffmann J, Otte A, Thiele S, Lotter H, Shu Y, Gabriel G. 2015. Sex differences in H7N9 influenza A virus pathogenesis. *Vaccine* 33: 6949–6954. <https://doi.org/10.1016/j.vaccine.2015.08.044>.
 50. Konig R, Stertz S, Zhou Y, Inoue A, Hoffmann HH, Bhattacharyya S, Alamares JG, Tscherne DM, Ortigoza MB, Liang Y, Gao Q, Andrews SE, Bandyopadhyay S, De Jesus P, Tu BP, Pache L, Shih C, Orth A, Bonamy G,

- Miraglia L, Ideker T, Garcia-Sastre A, Young JA, Palese P, Shaw ML, Chanda SK. 2010. Human host factors required for influenza virus replication. *Nature* 463:813–817. <https://doi.org/10.1038/nature08699>.
51. Karlas A, Machuy N, Shin Y, Pleissner KP, Artarini A, Heuer D, Becker D, Khalil H, Ogilvie LA, Hess S, Maurer AP, Muller E, Wolff T, Rudel T, Meyer TF. 2010. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* 463:818–822. <https://doi.org/10.1038/nature08760>.
 52. Everitt AR, Clare S, Pertel T, John SP, Wash RS, Smith SE, Chin CR, Feeley EM, Sims JS, Adams DJ, Wise HM, Kane L, Goulding D, Digard P, Anttila V, Baillie JK, Walsh TS, Hume DA, Palotie A, Xue Y, Colonna V, Tyler-Smith C, Dunning J, Gordon SB, Everingham K, Dawson H, Hope D, Ramsay P, Walsh TS, Campbell A, Kerr S, Harrison D, Rowan K, Addison J, Donald N, Galt S, Noble D, Taylor J, Webster N, Taylor I, Aldridge J, Dorman R, Richard C, Gilmour D, Simmons R, White R, Jardine C, Williams D, Booth M, Quasim T, et al. 2012. IFITM3 restricts the morbidity and mortality associated with influenza. *Nature* 484:519–523. <https://doi.org/10.1038/nature10921>.
 53. Wang S, Chi X, Wei H, Chen Y, Chen Z, Huang S, Chen JL. 2014. Influenza A virus-induced degradation of eukaryotic translation initiation factor 4B contributes to viral replication by suppressing IFITM3 protein expression. *J Virol* 88:8375–8385. <https://doi.org/10.1128/JVI.00126-14>.
 54. Long JE, Drayson MT, Taylor AE, Toellner KM, Lord JM, Phillips AC. 2016. Morning vaccination enhances antibody response over afternoon vaccination: a cluster-randomised trial. *Vaccine* 34:2679–2685. <https://doi.org/10.1016/j.vaccine.2016.04.032>.
 55. Chen C, Wang M, Zhu Z, Qu J, Xi X, Tang X, Lao X, Seeley E, Li T, Fan X, Du C, Wang Q, Yang L, Hu Y, Bai C, Zhang Z, Lu S, Song Y, Zhou W. 2016. Multiple gene mutations identified in patients infected with influenza A (H7N9) virus. *Sci Rep* 6:25614. <https://doi.org/10.1038/srep25614>.
 56. Lee HK, Tang JW, Kong DH, Loh TP, Chiang DK, Lam TT, Koay ES. 2013. Comparison of mutation patterns in full-genome A/H3N2 influenza sequences obtained directly from clinical samples and the same samples after a single MDCK passage. *PLoS One* 8:e79252. <https://doi.org/10.1371/journal.pone.0079252>.
 57. Tharakaraman K, Jayaraman A, Raman R, Viswanathan K, Stebbins NW, Johnson D, Shriver V, Sasisekharan V, Sasisekharan R. 2013. Glycan receptor binding of the influenza A virus H7N9 hemagglutinin. *Cell* 153:1486–1493. <https://doi.org/10.1016/j.cell.2013.05.034>.
 58. Xiong X, Martin SR, Haire LF, Wharton SA, Daniels RS, Bennett MS, McCauley JW, Collins PJ, Walker PA, Skehel JJ, Gamblin SJ. 2013. Receptor binding by an H7N9 influenza virus from humans. *Nature* 499:496–499. <https://doi.org/10.1038/nature12372>.
 59. Zaraket H, Baranovich T, Kaplan BS, Carter R, Song MS, Paulson JC, Rehg JE, Bahl J, Crumpton JC, Seiler J, Edmonson M, Wu G, Karlsson E, Fabrizio T, Zhu H, Guan Y, Hsain M, Schultz-Cherry S, Krauss S, McBride R, Webster RG, Govorkova EA, Zhang J, Russell CJ, Webby RJ. 2015. Mammalian adaptation of influenza A(H7N9) virus is limited by a narrow genetic bottleneck. *Nat Commun* 6:6553. <https://doi.org/10.1038/ncomms7553>.
 60. Swayne DE. 2012. Impact of vaccines and vaccination on global control of avian influenza. *Avian Dis* 56:818–828. <https://doi.org/10.1637/10183-041012-Review.1>.
 61. Wu Y, Cho M, Shore D, Song M, Choi J, Jiang T, Deng YQ, Bourgeois M, Almlil L, Yang H, Chen LM, Shi Y, Qi J, Li A, Yi KS, Chang M, Bae JS, Lee H, Shin J, Stevens J, Hong S, Qin CF, Gao GF, Chang SJ, Donis RO. 2015. A potent broad-spectrum protective human monoclonal antibody cross-linking two haemagglutinin monomers of influenza A virus. *Nat Commun* 6:7708. <https://doi.org/10.1038/ncomms8708>.
 62. Ferraris O, Lina B. 2008. Mutations of neuraminidase implicated in neuraminidase inhibitors resistance. *J Clin Virol* 41:13–19. <https://doi.org/10.1016/j.jcv.2007.10.020>.
 63. Hurt AC, Holien JK, Parker MW, Barr IG. 2009. Oseltamivir resistance and the H274Y neuraminidase mutation in seasonal, pandemic and highly pathogenic influenza viruses. *Drugs* 69:2523–2531. <https://doi.org/10.2165/11531450-000000000-00000>.
 64. Sheu TG, Deyde VM, Okomo-Adhiambo M, Garten RJ, Xu X, Bright RA, Butler EN, Wallis TR, Klimov AI, Gubareva LV. 2008. Surveillance for neuraminidase inhibitor resistance among human influenza A and B viruses circulating worldwide from 2004 to 2008. *Antimicrob Agents Chemother* 52:3284–3292. <https://doi.org/10.1128/AAC.00555-08>.
 65. Okomo-Adhiambo M, Demmler-Harrison GJ, Deyde VM, Sheu TG, Xu X, Klimov AI, Gubareva LV. 2010. Detection of E119V and E119I mutations in influenza A (H3N2) viruses isolated from an immunocompromised patient: challenges in diagnosis of oseltamivir resistance. *Antimicrob Agents Chemother* 54:1834–1841. <https://doi.org/10.1128/AAC.01608-09>.
 66. Simon P, Holder BP, Bouhy X, Abed Y, Beauchemin CA, Boivin G. 2011. The I222V neuraminidase mutation has a compensatory role in replication of an oseltamivir-resistant influenza virus A/H3N2 E119V mutant. *J Clin Microbiol* 49:715–717. <https://doi.org/10.1128/JCM.01732-10>.
 67. Marjuki H, Mishin VP, Chesnokov AP, De La Cruz JA, Davis CT, Villanueva JM, Fry AM, Gubareva LV. 2015. Neuraminidase mutations conferring resistance to oseltamivir in influenza A(H7N9) viruses. *J Virol* 89:5419–5426. <https://doi.org/10.1128/JVI.03513-14>.
 68. Li Z, Chen H, Jiao P, Deng G, Tian G, Li Y, Hoffmann E, Webster RG, Matsuoka Y, Yu K. 2005. Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *J Virol* 79:12058–12064. <https://doi.org/10.1128/JVI.79.18.12058-12064.2005>.
 69. Gabriel G, Dauber B, Wolff T, Planz O, Klenk HD, Stech J. 2005. The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc Natl Acad Sci U S A* 102:18590–18595. <https://doi.org/10.1073/pnas.0507415102>.
 70. Mok CK, Lee HH, Lestra M, Nicholls JM, Chan MC, Sia SF, Zhu H, Poon LL, Guan Y, Peiris JS. 2014. Amino acid substitutions in polymerase basic protein 2 gene contribute to the pathogenicity of the novel A/H7N9 influenza virus in mammalian hosts. *J Virol* 88:3568–3576. <https://doi.org/10.1128/JVI.02740-13>.
 71. Zhang H, Li X, Guo J, Li L, Chang C, Li Y, Bian C, Xu K, Chen H, Sun B. 2014. The PB2 E627K mutation contributes to the high polymerase activity and enhanced replication of H7N9 influenza virus. *J Gen Virol* 95:779–786. <https://doi.org/10.1099/vir.0.061721-0>.
 72. Yamayoshi S, Fukuyama S, Yamada S, Zhao D, Murakami S, Uraki R, Watanabe T, Tomita Y, Neumann G, Kawaoaka Y. 2015. Amino acid substitutions in the PB2 protein of H7N9 influenza A viruses are important for virulence in mammalian hosts. *Sci Rep* 5:8039. <https://doi.org/10.1038/srep08039>.
 73. Zhu W, Li L, Yan Z, Gan T, Li L, Chen R, Chen R, Zheng Z, Hong W, Wang J, Smith DK, Guan Y, Zhu H, Shu Y. 2015. Dual E627K and D701N mutations in the PB2 protein of A(H7N9) influenza virus increased its virulence in mammalian models. *Sci Rep* 5:14170. <https://doi.org/10.1038/srep14170>.
 74. Steel J, Lowen AC, Mubareka S, Palese P. 2009. Transmission of influenza virus in a mammalian host is increased by PB2 amino acids 627K or 627E/701N. *PLoS Pathog* 5:e1000252. <https://doi.org/10.1371/journal.ppat.1000252>.
 75. Van Hoven N, Pappas C, Belser JA, Maines TR, Zeng H, Garcia-Sastre A, Sasisekharan R, Katz JM, Tumpey TM. 2009. Human HA and polymerase subunit PB2 proteins confer transmission of an avian influenza virus through the air. *Proc Natl Acad Sci U S A* 106:3366–3371. <https://doi.org/10.1073/pnas.0813172106>.
 76. de Graaf M, Fouchier RA. 2014. Role of receptor binding specificity in influenza A virus transmission and pathogenesis. *EMBO J* 33:823–841. <https://doi.org/10.1002/embj.201387422>.
 77. Veit M, Kretzschmar E, Kuroda K, Garten W, Schmidt MF, Klenk HD, Rott R. 1991. Site-specific mutagenesis identifies three cysteine residues in the cytoplasmic tail as acylation sites of influenza virus hemagglutinin. *J Virol* 65:2491–2500.
 78. Zurcher T, Luo G, Palese P. 1994. Mutations at palmitoylation sites of the influenza virus hemagglutinin affect virus formation. *J Virol* 68:5748–5754.
 79. Fischer C, Schroth-Diez B, Herrmann A, Garten W, Klenk HD. 1998. Acylation of the influenza hemagglutinin modulates fusion activity. *Virology* 248:284–294. <https://doi.org/10.1006/viro.1998.9286>.
 80. Melikyan GB, Lin S, Roth MG, Cohen FS. 1999. Amino acid sequence requirements of the transmembrane and cytoplasmic domains of influenza virus hemagglutinin for viable membrane fusion. *Mol Biol Cell* 10:1821–1836. <https://doi.org/10.1091/mbc.10.6.1821>.
 81. Wagner R, Herwig A, Azzouz N, Klenk HD. 2005. Acylation-mediated membrane anchoring of avian influenza virus hemagglutinin is essential for fusion pore formation and virus infectivity. *J Virol* 79:6449–6458. <https://doi.org/10.1128/JVI.79.10.6449-6458.2005>.
 82. World Health Organization. 2013. Real-time RT-PCR protocol for the detection of avian influenza A (H7N9) virus. World Health Organization, Geneva, Switzerland.
 83. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, Ostell J, Lipman D. 2008. The influenza virus resource at the national center for biotechnology information. *J Virol* 82:596–601. <https://doi.org/10.1128/JVI.02005-07>.

84. Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25:1105–1111. <https://doi.org/10.1093/bioinformatics/btp120>.
85. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
86. Xiao YL, Kash JC, Beres SB, Sheng ZM, Musser JM, Taubenberger JK. 2013. High-throughput RNA sequencing of a formalin-fixed, paraffin-embedded autopsy lung tissue sample from the 1918 influenza pandemic. *J Pathol* 229:535–545. <https://doi.org/10.1002/path.4145>.
87. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
88. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res* 28:235–242. <https://doi.org/10.1093/nar/28.1.235>.
89. Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18: 2714–2723. <https://doi.org/10.1002/elps.1150181505>.
90. Zhang Y, Aevermann BD, Anderson TK, Burke DF, Dauphin G, Gu Z, He S, Kumar S, Larsen CN, Lee AJ, Li X, Macken C, Mahaffey C, Pickett BE, Reardon B, Smith T, Stewart L, Suloway C, Sun G, Tong L, Vincent AL, Walters B, Zaremba S, Zhao H, Zhou L, Zmasek C, Klem EB, Scheuermann RH. 2017. Influenza research database: an integrated bioinformatics resource for influenza virus research. *Nucleic Acids Res* 45:D466–D474. <https://doi.org/10.1093/nar/gkw857>.