Molecular Evolution, Diversity, and Adaptation of Influenza A(H7N9) Viruses in China

Jing Lu,[1] Jayna Raghwani,[1] Rhys Pryce, Thomas A. Bowden, Julien Thézé, Shanqian Huang, Yingchao Song, Lirong Zou, Lijun Liang, Ru Bai, Yi Jing, Pingping Zhou, Min Kang, Lina Yi, Jie Wu,[2] Oliver G. Pybus,[2] Changwen Ke[2]

The substantial increase in prevalence and emergence of antigenically divergent or highly pathogenic influenza A(H7N9) viruses during 2016–17 raises concerns about the epizootic potential of these viruses. We investigated the evolution and adaptation of H7N9 viruses by analyzing available data and newly generated virus sequences isolated in Guangdong Province, China, during 2015–2017. Phylogenetic analyses showed that circulating H7N9 viruses belong to distinct lineages with differing spatial distributions. Hemagglutination inhibition assays performed on serum samples from patients infected with these viruses identified 3 antigenic clusters for 16 strains of different virus lineages. We used ancestral sequence reconstruction to identify parallel amino acid changes on multiple separate lineages. We inferred that mutations in hemagglutinin occur primarily at sites involved in receptor recognition or antigenicity. Our results indicate that highly pathogenic strains likely emerged from viruses circulating in eastern Guangdong Province during March 2016 and are associated with a high rate of adaptive molecular evolution.

Since its first detection in March 2013, avian influenza A(H7N9) virus has caused 1,534 human infections that, as of November 30, 2017, had resulted in 608 deaths. Recurrent waves of human cases have been reported in 27 provinces in China, indicating sustained transmission of H7N9 viruses (*1*). Moreover, since its emergence, H7N9 virus has reassorted with influenza A(H9N2) viruses that co-circulate in China, resulting in an increasingly diverse array of virus genomes (*2*–*4*). The fifth influenza epidemic wave (2016–17) was marked by a notable increase in the number of human cases (677 during September 2016–May 2017), making it the largest outbreak of influenza A(H7N9) since 2013. Moreover, geographic distribution of human cases suggests that H7N9 virus is now more widespread and that residences of patients have shifted gradually from urban to semiurban and rural areas (*1*,*5*–*7*). These epidemiologic observations have raised public health concerns.

Previous molecular surveillance studies suggested that H7N9 virus lineages originate in 2 densely populated areas, the Yangtze River Delta region in eastern China and the Pearl River Delta region in central Guangdong Province (*8*). Preliminary epidemiologic data suggested that most human infections in the current fifth epidemic wave were caused by viruses from the Yangtze River Delta region (*5*) (previously named lineage C viruses) (*3*). These viruses, in contrast to viruses from the Pearl River Delta region (previously named lineage B viruses) (*3*), appear to exhibit reduced cross-reactivity with existing candidate vaccine virus strains (*9*). Furthermore, a subset of lineage C isolates has also acquired a highly pathogenic (HP) phenotype (*5*,*10*,*11*).

These observations suggest that the increased epidemic activity of H7N9 viruses in China might be driven, at least in part, by ongoing virus evolution and adaptation. Decreased cross-reactivity and increased pathogenicity of some H7N9 viruses was discovered only recently (*9*), and the genetic diversity and evolution of the current fifth epidemic wave of these viruses are not yet well understood. Information necessary to clarify this situation includes geographic distribution of currently circulating H7N9 virus lineages, origin and genetic composition of newly emerged HP H7N9 viruses, and evolutionary and structural characterization of mutations associated with the fifth epidemic wave of these viruses.

Author affiliations: Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China (J. Lu, Y. Song, L. Zou, L. Liang, R. Bai, Y. Jing, P. Zhou, M. Kang, L. Yi, J. Wu, C. Ke); Guangdong Provincial Institution of Public Health, Guangzhou (J. Lu, P. Zhou, L. Yi); University of Oxford, Oxford, UK (J. Raghwani, R. Pryce, T.A. Bowden, J. Thézé, O.G. Pybus); Beijing Normal University, Beijing, China (S. Huang)

DOI: https://doi.org/10.3201/eid2410.171063

[1]These authors contributed equally to this article.

[2]These senior authors jointly supervised this study.

We report 47 hemaglutinnin (HA) and 43 neuraminidase (NA) gene sequences of human-derived and poultry-derived H7N9 viruses that were isolated during 2015–2017 in Guangdong Province, China. We conducted structural and evolutionary analyses of these strains and characterized the evolution and emergence of currently circulating H7N9 viruses in China.

Materials and Methods

Ethics

This study was approved by the institutional ethics committee of the Center for Disease Control and Prevention of Guangdong Province. Written consent was obtained from patients or their guardian(s) when samples were collected. Patients were informed about the study before providing written consent, and data were anonymized for analysis.

Sample Collection

Samples from persons with suspected cases of influenza A(H7N9) were initially tested for avian influenza A virus in provincial clinics in Guangdong Province. Specimens with positive results were subsequently analyzed (*12*,*13*). For poultry-related samples, we obtained samples from locations where poultry were housed and processed (e.g., cages, feeding troughs, defeathering machines) (*12*). Respiratory specimens were collected from persons with suspected cases of influenza A(H7N9) by the Ministry of Health of China.

Sequence Alignment

For phylogenetic studies, we sequenced 47 HA and 41 NA sequences from 20 human samples and 28 poultry-related samples; all belonged to the fourth and fifth epidemic waves of influenza A(H7N9) (GISAID [https://www.gisaid.org/] accession nos. EPI866538–77, 972231–6, 972238–303, 974029, 974523, 974539–42, 997159–60, and 1171786–93). These new H7N9 sequences were combined with all available H7N9 gene sequences whose sampling dates and locations were known. Two gene sequence datasets were generated: H7, HA (n = 737) and N9, NA (n = 610). We constructed multiple sequence alignments by using ClustalW (*14*) and edited these sequences manually by using AliView (*15*).

Molecular Clock Phylogenetic Analysis

We estimated molecular clock phylogenies by using the Bayesian Markov Chain Monte Carlo approach implemented in BEAST version 1.8 (*16*) as described (*4*). We computed 4 independent Markov Chain Monte Carlo runs of 1.5×10^8 steps for each alignment and extracted a subset of 2,000 phylogenies from the posterior tree distribution, subsequently used as an empirical tree distribution for phylogeographic analyses (*17*). We computed maximum clade credibility trees for each dataset by using TreeAnnotator (*16*).

Phylogeographic Analysis of Influenza A(H7N9) Epidemic

We used the discrete phylogeographic method (*18*) implemented in BEAST to investigate spatial dynamics of H7N9 virus lineages from 6 regions in China as classified in a previous study (*4*). The 6 locations were eastern China (Anhui, Shanghai, Zhejiang, Jiangsu, and Shandong); central China (Jiangxi and Hunan); northern China (Beijing, Henan, Hebei, and Xinjiang); southeastern China (Fujian); central Guangdong Province (Guangzhou, Huizhou, Foshan, Dongguan, Zhongshan, Shenzhen, Jiangmen, Zhaoqing Yangjiang, Maoming, and Yunfu); and eastern Guangdong Province (Meizhou, Heyuan, Chaozhou, Jieyang, Shantou, Shanwei, and Shaoguan).

Because sporadic human cases detected in Malaysia and Taiwan were believed to have originated in China, we used available epidemiologic information to assign their location to the most likely source in China. Hong Kong and central Guangdong Province were treated as a single location because of their proximity to each other. We analyzed reported H7N9 virus infections and virus sequences (online Technical Appendix Table 1, https://wwwnc.cdc.gov/EID/article/24/10/17-1063-Techapp1.pdf). To estimate directionality of virus lineage movement, we used asymmetric continuous-time Markov chain phylogeographic model (*19*) and a Bayesian stochastic search variable selection procedure (*18*).

Inferring Phylogenetic Distribution of Amino Acid Changes

We investigated phylogenetic positions of amino acid changes among H7N9 virus isolates by using HA and NA maximum clade credibility trees. We estimated maximum posterior probability amino acid sequences for each internal node by using BEAST with a Jones–Taylor–Thornton amino acid substitution model (*20*), gamma-distributed among-site rate heterogeneity (*21*), and a strict molecular clock model. To infer amino acid substitutions along the trunk branches of the H7N9 phylogeny, we mapped amino acid changes onto internal branches by using a Java script (available on request). Trunk branches corresponded to internal branches that subtended >5 terminal nodes in the fifth influenza epidemic wave.

Structure-Based Mapping Analysis

We used the crystal structure of the HA (Protein Data Bank no. 4BSE) (*22*) and NA (Protein Data Bank no. 2C4L) glycoproteins from an influenza A(H7N9) virus to map amino acid changes identified by evolutionary analysis. We

performed residue mapping onto the H7 and N9 structures by using PyMol (*23*). We calculated solvent accessibility for trimeric hemagglutinin with the ligands removed by using ESPript (*24*) and identified receptor-binding residues by using CONTACT in CCP4 (*25*).

Positive Selection Analyses

To identify sites under positive selection, we used methods implemented in HyPhy (*26*) to estimate the dN/dS ratio of codons in HA. These methods included single-likelihood ancestor counting (*27*), fixed effects likelihood (*27*), mixed effects model of evolution (*28*), and the fast unconstrained Bayesian approximation approach (*29*).

Estimating Rates of Virus Molecular Adaptation

We estimated rates of adaptive substitution in H7N9 virus HA and NA genes by using an established population genetic method related to the McDonald-Kreitman test (*30*,*31*). We used a consensus of H7N9 first-wave sequences as an outgroup to estimate derived and ancestral mutational site frequencies in each subsequent wave. Specifically, we classified polymorphisms into 3 categories according to their frequency in the population (low, 0%–15%; intermediate, 15%–75%; and high, 75%–100%). We calculated the number of adaptive substitutions from the number of synonymous and nonsynonymous sites in each category and assessed statistical uncertainty by using a bootstrap approach (1,000 replicates) (*30,31*).

Serologic Analysis

We obtained serum samples from 4 patients with influenza A(H7N9) 2–3 weeks after clinical symptoms were observed. We performed hemagglutination inhibition assays by using different lineages of H7N9 viruses as antigens (online Technical Appendix). Three lineage C1 strains, 4 lineage C2 strains, 5 lineage B strains, and 4 HP strains were used as antigens (Table). We calculated serum titer for each H7N9 strain as the highest reciprocal serum dilution providing complete hemagglutination inhibition.

Results

Molecular Epidemiology of Viruses Isolated during 2013–2017

During 2013–2017, the influenza A(H7N9) virus epidemic lineage was geographically structured and classified into 3 major lineages, A, B, and C, in accordance with the lineage naming scheme used in a previous study (*3*). H7N9 virus has evolved in a clock-like manner (i.e., there is a strong linear relationship between genetic divergence and sampling time; correlation coefficient 0.95) (Figure 1). The estimated time to the most recent common ancestor (TMRCA) of H7N9 virus HA sequences is November 2012 (95% credible region October–December 2012). The corresponding molecular clock phylogeny for NA (online Technical Appendix Figure 1) also shows A–C lineages and has a similar estimated TMRCA of September 2012 (95% credible region July–October 2012). However, the topology of the NA phylogeny differs from that of HA, suggesting reassortment between HA and NA during emergence of the H7N9 virus epidemic lineage (Figure 2; online Technical Appendix Figure 1).

Different H7N9 virus lineages are associated with different epidemiologic patterns (Figures 2, 3). Specifically, most (86%, 32/37) lineage B viruses that were isolated during the fourth and fifth influenza epidemic waves descended from viruses circulating in central Guangdong Province during earlier epidemic seasons (Figure 2). In addition, lineage B viruses isolated from the fourth and fifth influenza waves

Table. Hemagglutination inhibition titers for serum samples from 4 patients infected with influenza A(H7N9) virus against reference H7N9 strains, China*

H7N9 strain	Date of collection	Clade	Titer by patient and H7N9 strain			
			P1, strain NA	P2, strain NA	P3, strain ZS29	P4, strain ST18
EPI972232/A/CZ009†	2017 Jan	C1	2,048	2,048	512	512
EPI972243/A/MZ011†	2017 Jan	C1	4,096	4,096	1,024	512
EPI1171792/A/ST18†	2017 Jan	C1	1,024	512	512	512
EPI656434/A/ST72	2015 Feb	C2	256	128	1024	128
EPI866569/A/ST021†	2016 Feb	C2	256	128	512	128
EPI656314/A/SW20	2015 Jan	C2	256	128	512	128
EPI1171791/A/SW33†	2015 Feb	C2	256	128	512	128
EPI972259/A/ZS201†	2016 Dec	B	256	128	512	128
EPI972234/A/FS10†	2017 Jan	B	256	256	512	128
EPI656054/A/ZS74	2014 Jan	B	512	256	512	128
EPI656038/A/ZS71	2014 Jan	B	256	256	512	128
EPI656014/A/GZ66	2014 Jan	B	512	256	512	64
EPI1171786/Env/YJ370†	2017 Sep	HP	64	64	512	64
EPI1171788/Env/YJ073†	2017 May	HP	64	64	512	64
EPI1171790/A/ZS29†	2017 Mar	HP	128	64	512	64
EPI919607/A/17SF003	2017 Jan	HP	256	64	512	128

*HP, highly pathogenic; NA, not available; P, patient no.
†Strains isolated and sequenced in this study.

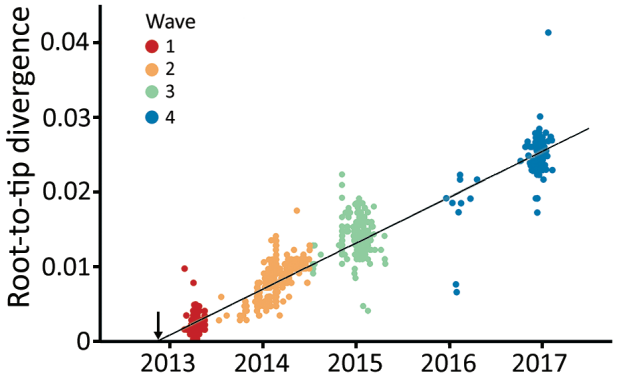


Figure 1. Regression of root-to-tip divergence estimated from hemagglutinin gene of influenza A(H7N9) viruses, China. Arrow indicates the time of the most recent common ancestor of the epidemic lineage.

were almost exclusively restricted to central (rather than eastern) Guangdong Province (Figures 2, 3). In contrast, viruses in eastern China, composed of 2 lineages (A and C) have been exported to and become dominant in multiple regions as the epidemic has progressed (*3*). These findings indicate a comparatively broader geographic dissemination (Figure 3; online Technical Appendix Figure 1).

The new isolates from eastern Guangdong Province, combined with isolates from eastern China (*1*,*5*), suggest that recent H7N9 virus activity is driven primarily by lineage C viruses (Figure 2). The estimated TMRCA of lineage C is December 2013 (95% highest posterior density October 2013–January 2014). For lineage C, we observed 2 clades (Figure 4). The larger of these clades (C1) circulates mainly in central and eastern China, and the smaller clade (C2) is found predominantly in eastern Guangdong Province. Clade C2 also includes recently identified HP viruses (Figures 1, 2, 4).

To investigate these HP viruses, we undertook retrospective screening of poultry-related samples collected in Guangdong Province during January 2016–February 2017 and identified 7 HP influenza virus isolates that belong to the HP cluster within C2 (Figure 2). These HP viruses also form a distinct cluster within lineage C viruses in the NA phylogeny (online Technical Appendix Figure 1). Our analyses indicated that the HP clade likely emerged from clade C2 viruses circulating in eastern Guangdong Province in 2016.

Adaptive Evolution in Virus C Lineage

We then investigated whether the increasing prevalence of lineage C viruses might be associated with virus adaptation. We combined ancestral sequence reconstruction of lineage B and C HA gene sequences (Figures 4, 5) by mapping residues that have undergone changes onto the crystal

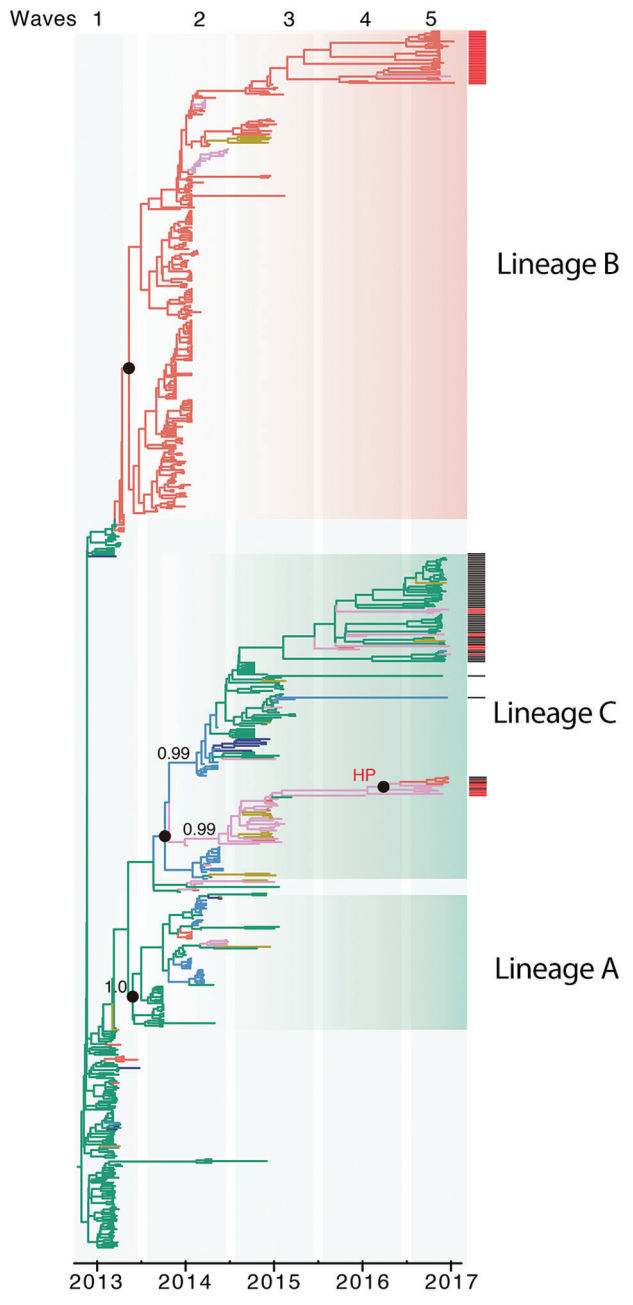


Figure 2. Genetic evolution and spatial spread of epidemic lineage of influenza A(H7N9) viruses, China, 2013–2017. Bayesian maximum clade credibility tree of the hemagglutinin gene is shown. Black bars to the right of the tree indicate sequences (from waves 4 and 5) from other studies (*1*,*5*), and red bars indicate sequences reported in this study from Guangdong Province. Branch colors indicate most probable ancestral locations of each branch. Three major lineages (A, B, and C) of H7N9 viruses were observed. Values along branches indicate bootstrap values. Black circles indicate posterior support >0.95. Location of posterior support is shown for selected clades. An H7N9 strain closely related to the highly pathogenic H7N9 virus cluster is indicated. HP, highly pathogenic.

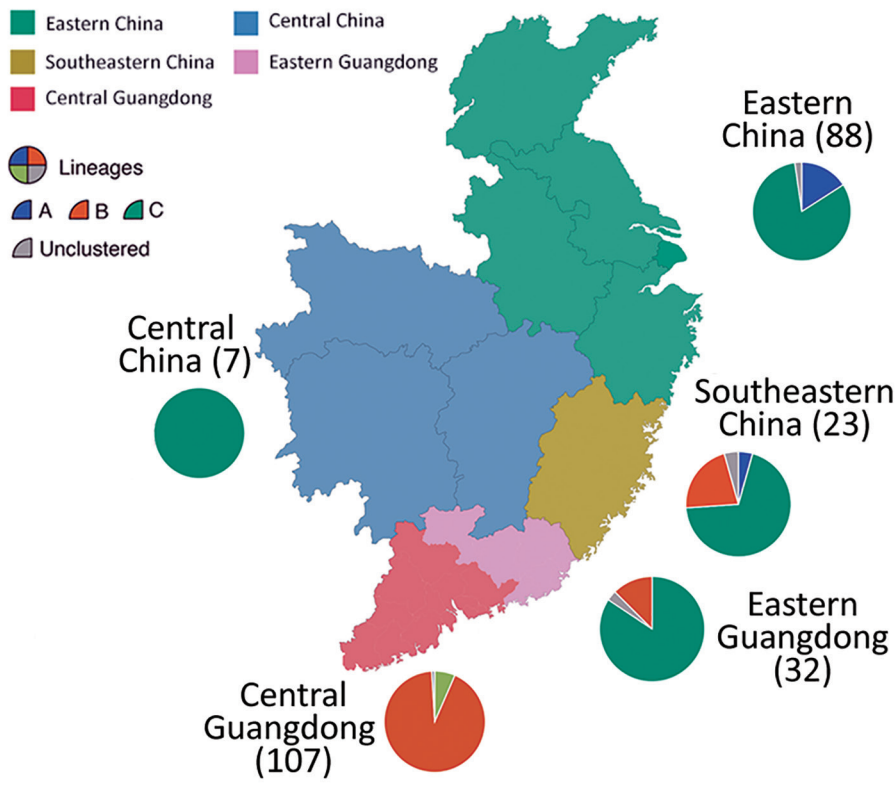


Figure 3. Geographic location and lineage classification of 374 influenza A(H7N9) human viruses, China. Values in parentheses indicate number of sequenced viruses from each region. Pie charts indicate approximate percentages of each virus lineage (A, B, C, or unclustered). Sequences from Xinjiang Province in northern China are not shown.

structure of the trimeric hemagglutinin. Our analysis identified several notable amino acid substitutions that occurred along the internal branches of lineage C viruses (Figure 4).

Around the time of the second influenza epidemic wave, ancestral viruses of lineage C acquired several amino acid changes in HA, specifically L177I, G386A, S489R, and S128N (based on H3 sequence numbering). Three of these mutations (G386A, S489R, and S128N) are located in solvent-accessible regions of HA (Figure 6; online Technical Appendix Table 2). Furthermore, S128N was found within the 130 loop and is proximal to the receptor surface (distance ≈20 Å) (Figure 6).

We found by evolutionary analysis that several HA sites acquired amino acid mutations independently in different phylogenetic clades. First, 4 mutations (A135V, L177I, M236I, and S489N) occurred independently along the trunk lineages that gave rise to the current lineage B and C viruses (Figure 5). Three of these mutations (A135V, M236I, and S489N) were observed only in the fourth and fifth influenza epidemic waves of lineage B (Figure 5). Second, comparison of the C1 and C2 clades also identified parallel amino acid changes within lineage C (A122T/P and M236I) (Figure 4).

The observation of parallel amino acid changes along those particular lineages (online Technical Appendix Tables 2, 3) that have persisted until the fifth influenza epidemic wave (i.e., parallel changes between lineages B and C and between the C1 and C2 clades) is suggestive of convergent, adaptive molecular evolution. The parallel changes in lineage C (A122T/P and M236I) are estimated to be fully or partially solvent accessible and the A135V mutation is located at the receptor-binding site (Figure 6). One subclade of lineage B viruses appears to have acquired mutations A135V and S489N within the last 12 months (Figure 5). Therefore, we suggest that this subclade should be closely monitored in the future.

Within the C2 clade, we found that HA acquired 7 amino acid changes (I48T, A122P, K173E, L226Q, M236I, I326V, and E393K) on the internal branch immediately ancestral to the HP cluster. This internal branch represents a period of approximately 1 year (Figure 4). Although all of these changes appeared in residues with partial or full solvent accessibility, mutations K173E, L226Q, and I326V are particularly noteworthy because they have arisen at or near known antigenic, receptor-binding, and proteolytic cleavage sites, respectively (Figure 6). Furthermore, these mutations in the HP cluster also coincide with appearance of a 4-amino acid insertion (KRTA) near the HA1-HA2 proteolytic cleavage site (Figure 4). A subset of HA substitutions (at sites 57, 114, 140, 226, and 276) that occurred on the trunk branches of lineages B and C viruses was also found to be under positive selection on the basis of dN/dS ratios we estimated by using methods implemented in HyPhy (online Technical Appendix Table 4).

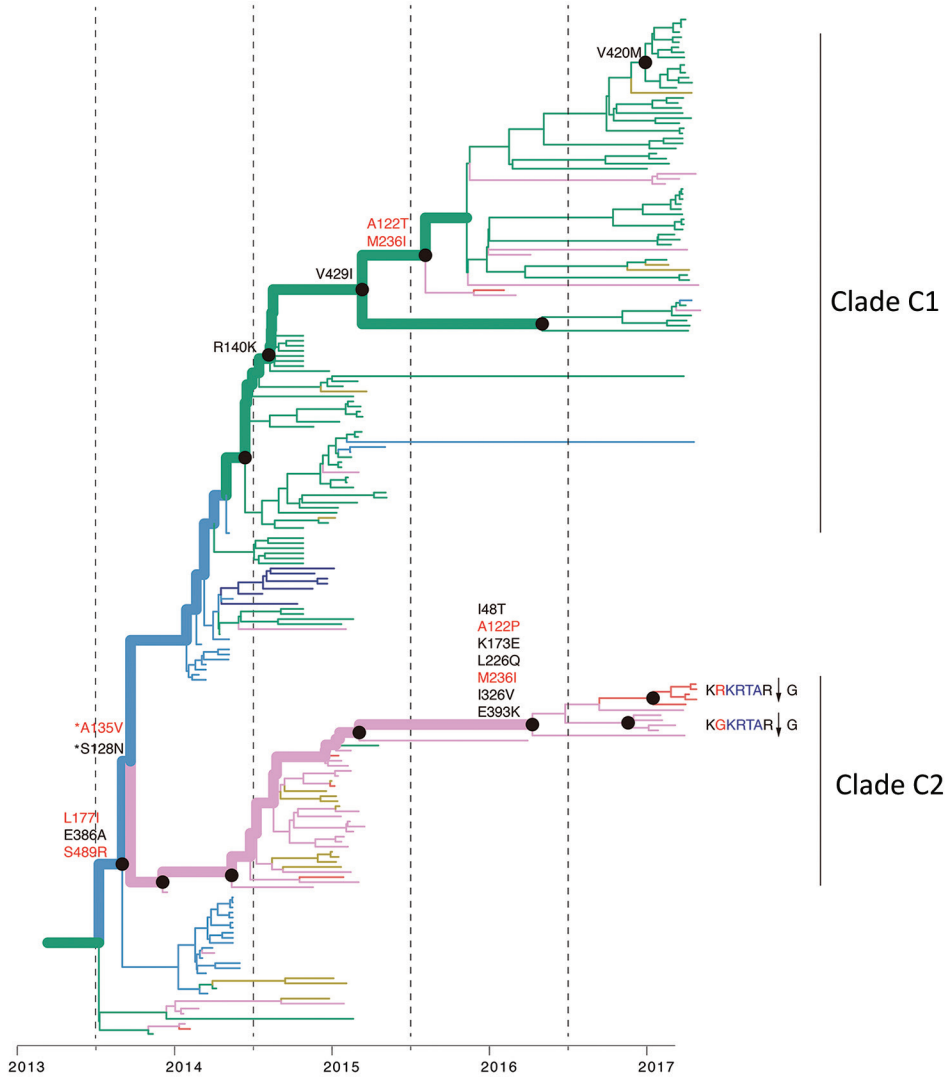


Figure 4. Reconstruction of amino acid changes along trunk of lineage B phylogenies of influenza A(H7N9) viruses, China. Maximum clade credibility tree of hemagglutinin gene sequences from lineage B is shown. Branches are colored according to geographic locations, as in Figure 3. Thicker lines indicate the trunk lineage leading up to the current fifth influenza epidemic wave. Amino acid changes along the trunk are indicated. Red branches indicate sites undergoing parallel amino acid changes across multiple lineages. Mutations correspond to H3 numbering scheme. *Indicates uncertainty about the phylogenetic position of the A135V and S128N mutations because branch posterior support is low.

We also investigated whether amino acid changes in the HA gene during emergence of influenza A(H7N9) virus have been driven by adaptive evolution similar to that observed for seasonal human influenza (*30*). We found evidence for adaptive evolution in HA genes of B and C virus lineages. We estimated that lineage B adapted at a rate of 0.80 (interquartile range [IQR] 0.21–0.95) adaptive substitutions across the whole HA gene per year and lineage A at a rate of 0.60 (IQR 0.10–1.18) adaptive substitutions per year. Within lineage C, the estimated adaptation rate of the C1 clade is ≈1.84 (IQR 1.09–2.14) adaptive substitutions per year and that for the C2 clade (which includes the HP cluster) is 3.12 (IQR 2.45–3.79) adaptive substitutions per year. These results indicate molecular adaptation across the whole H7N9 lineage and suggest that adaptation is faster in the 2 C clades than in the A and B lineages. Previous estimates of rates of adaptive substitution were 1.02 fixations per year in the whole HA gene for seasonal human influenza A(H1N1) virus and 1.52 fixations per year in the whole HA gene for influenza A(H3N2) virus (*30*). In this context, the rate of adaptive evolution observed for lineage C here is notable and raises concern for ongoing evolution of these viruses.

Antigenic Properties

We collected serum samples from 4 patients infected with H7N9 virus during 2015 and 2017 (Table). For patients 3 and 4, the corresponding virus strains were isolated and sequenced. Phylogenetic analysis indicated that patient 3 was infected with clade C1 virus and that patient 4 was infected with HP virus. Hemagglutination inhibition results suggested the presence of 3 antigenic clusters among the 16 H7N9 virus strains selected. Serum samples from patients 1, 2, and 3 showed similar patterns, reacting robustly to clade C1 viruses and moderately to clade C2 and lineage B viruses but poorly to HP viruses. A serum sample from a

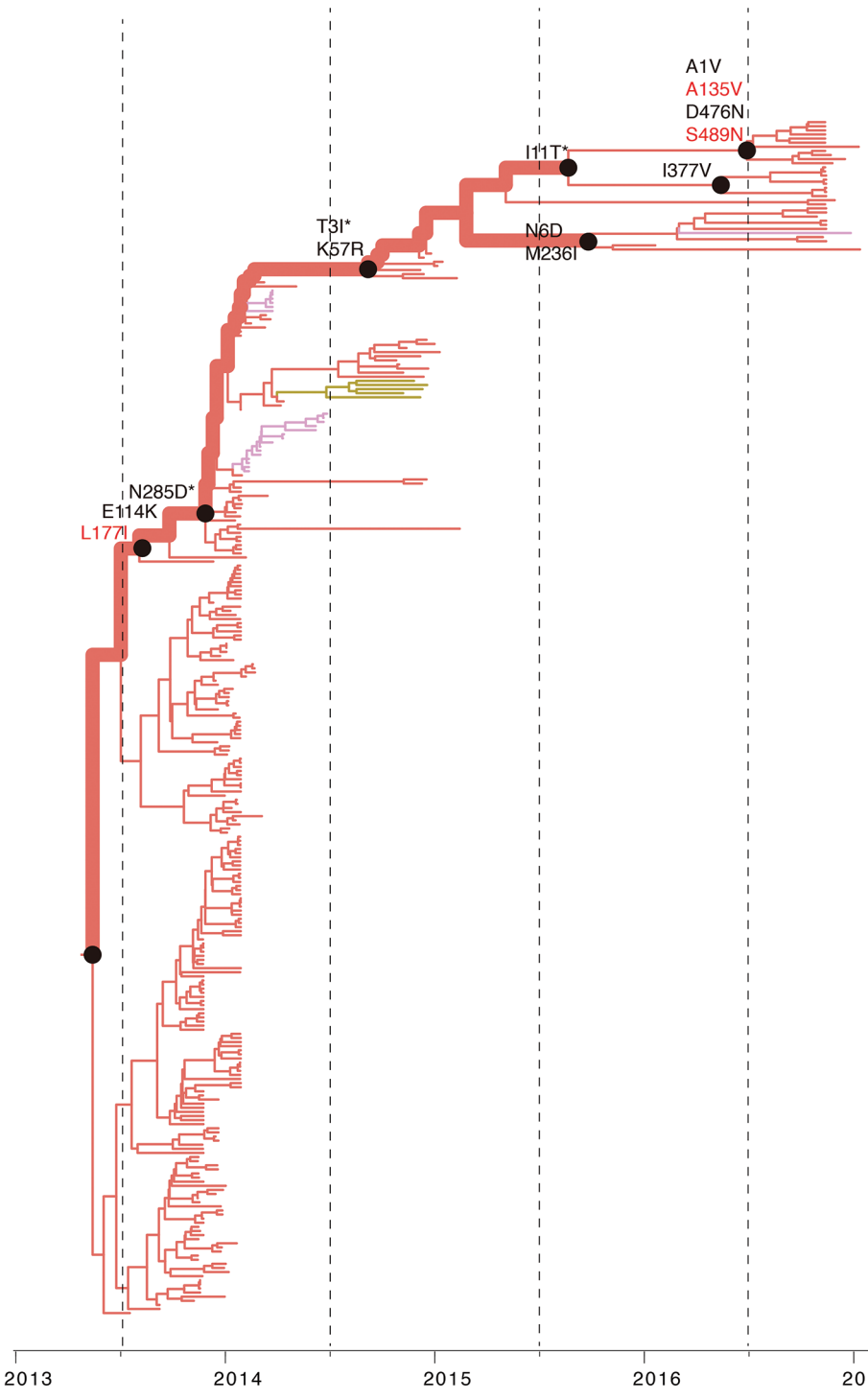


Figure 5. Reconstruction of amino acid changes along trunk of lineage C phylogenies of influenza A(H7N9) viruses, China. Maximum clade credibility tree of hemagglutinin gene sequences from lineage C is shown. Branches are colored according to geographic locations, as in Figure 3. Thicker lines indicate the trunk lineage leading up to the current fifth influenza epidemic wave. Amino acid changes along the trunk are indicated. Red branches indicate sites undergoing parallel amino acid changes across multiple lineages. Mutations correspond to H3 numbering scheme. *Mutation sites not present are numbered according to H7 numbering.

patient infected with an HP H7N9 virus appeared to react strongly to all H7N9 virus strains.

Discussion

Our results show that H7N9 viruses of lineage C, which were prevalent in the recent fifth influenza epidemic wave in China, comprise 2 geographically distinct clades (C1 and C2) that have undergone adaptive evolution. Clade C1 is found primarily in eastern and central China and clade 2 in Guangdong Province, and both clades appear to have circulated in bird populations for ≈3 years. Our ancestral state reconstruction analysis provides crucial evidence that

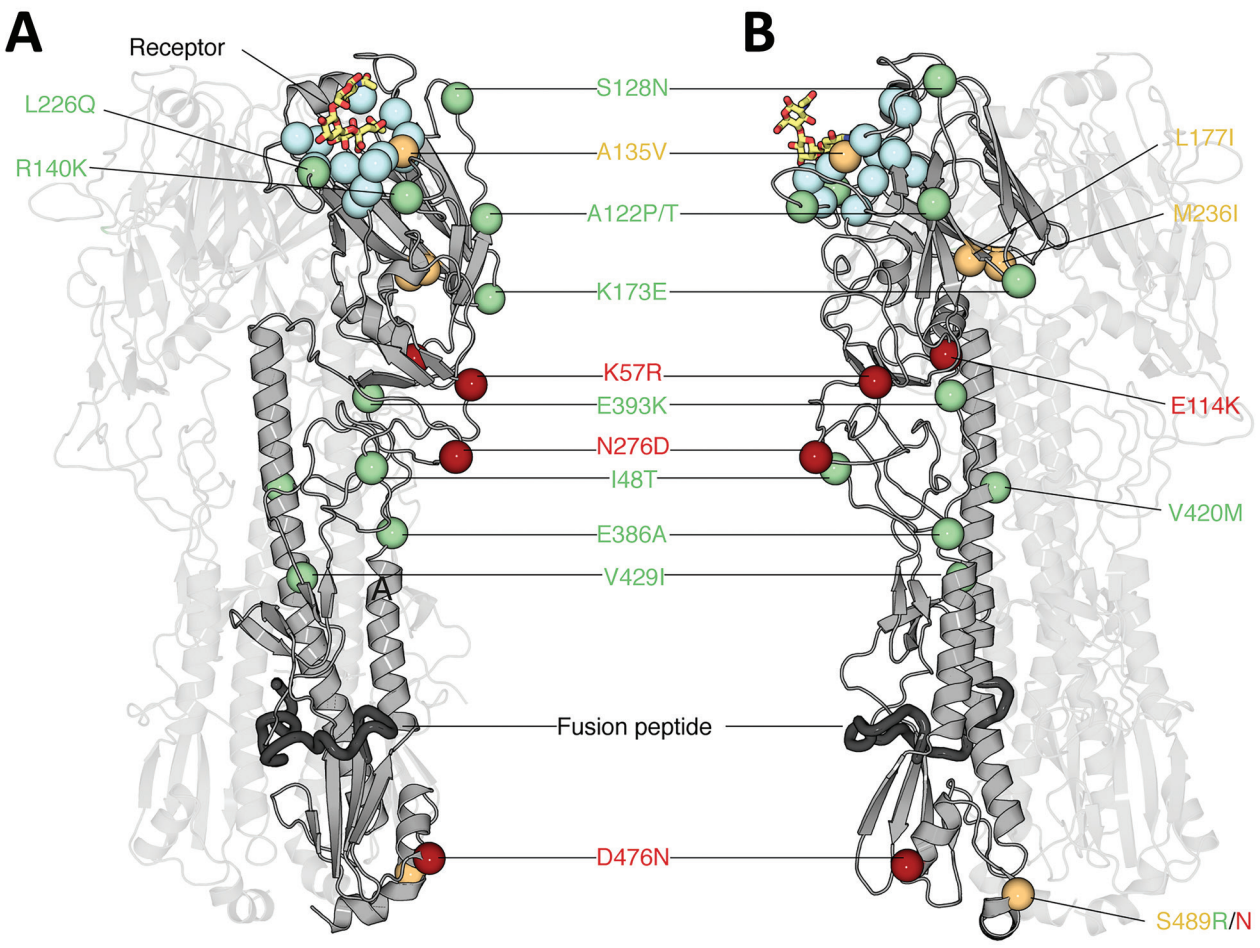


Figure 6. Structural analysis of amino acid changes in hemagglutinin in lineages B and C of influenza A(H7N9) viruses, China. Crystal structure of the homotrimeric H7 hemagglutinin bound to a human receptor analog (Protein Data Bank no. 4BSE) (*27*) (A) and rotated 90° counterclockwise (B) are shown. Two of the 3 protomers are displayed with high transparency to aid visualization. The carbon Cα positions of salient features are shown as spheres. Blue indicates receptor-binding residues, red indicates mutations in lineage B, green indicates mutations in lineage C, and orange indicates mutations in lineages B and C. Human receptor analog α2,6-SLN is shown as sticks colored according to constituent elements: carbon in orange, oxygen in red, and nitrogen in blue. Dark gray indicates the putative fusion peptide (*32*). Residues are numbered according to the H3 numbering system (online Technical Appendix Table 2, https://wwwnc.cdc.gov/EID/article/24/10/17-1063-Techapp1.pdf). A135 and L226 participate in receptor binding and thus are likely to modulate receptor specificity.

2 successful lineages of H7N9 viruses (lineages B and C) have experienced multiple parallel amino acid changes (Figures 4, 5), suggesting the possible action of convergent molecular evolution.

We also observed a higher rate of virus adaptation in eastern Guangdong Province (C2 clade compared with C1). Although clades C1 and C2 are phylogenetically closely related, serum from a clade C1 virus-infected patient has moderate reactivity with C2 strains from 2015–2016 and poor reactivity to the HP virus from 2016–2017. The higher adaptation rate and antigenic changes in clade C2 are of concern from a public health perspective. Introduction of HP avian influenza into domestic poultry might constitute a serious risk, as demonstrated by emergence of goose–Guangdong lineage HP H5N1 viruses, which spilled back into wild birds and caused the longest global outbreak of HP avian influenza to date (*33*).

Parallel amino acid changes in clades C1 and C2 occurred at 2 sites in HA (122 and 236) (Figure 6). Furthermore, we observed 4 mutations that emerged independently in lineages B and C (sites 135, 177, 236, and 489). These results suggest adaptive convergent molecular evolution. Site 135 is located in the receptor-binding region and is near antigenic site A, as defined by Wiley et al. (*34*). Thus, the observed A135V mutation might modulate receptor affinity and contribute to immune escape (Figure 6), as observed in influenza A(H7N1) and A(H7N7) viruses (*35,36*).

Specifically, experimental studies indicate that threonine at position 135 in the HP H7N7 virus A/Netherlands/

219/2003 confers broad-scale resistance to neutralizing monoclonal antibodies against the earliest strain of H7N9 virus (A/Shanghai/02/2013) (*37*). Furthermore, the World Health Organization has reported that recent clade C1 viruses (but not those of lineage B) react less to postinfection ferret antiserum raised against the A/Anhui/1/2013 and A/Shanghai/2/2013-derived candidate vaccine strains (*9*). Consistent with this finding, we found that most clade C1 viruses isolated in 2015 have the A135V mutation. However, this mutation was only detected in a small proportion of recent lineage B viruses (Figure 5).

In this study, we performed a preliminary evaluation of the antigenicity of H7N9 viruses by using patient serum samples collected in 2015 and 2017. Without serum raised in response to early strains from 2013, we cannot discriminate antigenic change between strains from 2013 and those from 2017. However, the limited data we have indicate the presence of 3 antigenic clusters among the 4 phylogenetic clusters circulating during the fifth influenza epidemic wave (Table). HA1 positions 109–301 (H3 numbering) include the A–E antigenic epitopes, which are known to determine antigenicity of influenza A viruses (*34*). The amino acid changes responsible for the antigenic differences between clade C1 and other clades were located in antigenic site A (position 140, H3 numbering; online Technical Appendix Figure 2).

The mutation R140K has been observed in viruses isolated from ferrets experimentally infected with avian influenza A(H7N9) viruses and has been linked to antigenic drift of influenza A(H5N1) viruses (*38–40*). By comparing the sequences of the HP H7N9 virus cluster and other clade C2 viruses, we found a substitution in antigenic site E (position 173, H3 numbering) that could underlie antigenic change in HP H7N9 viruses (online Technical Appendix Figure 2). In future work, we aim to explore the roles of these substitutions in determining viral antigenicity in the context of H7N9 virus genomes by using reverse genetics.

HP H7N9 viruses belonging to lineage C2 were first detected in late 2016, but spread greatly in geographic range in early 2017 (*41*). Several mechanisms for the genesis of an HP virus from a low pathogenicity virus have been proposed, including transcription errors (*42*), stepwise amino acid substitutions (*43*), or recombination (*44*). For H7, emergence of HP viruses is attributed to nonhomologous recombination resulting in simultaneous insertion of several amino acids at the HA cleavage site. These insertions might be derived from host 28S rRNA sequence (*45*) or from other influenza gene segments, such as the matrix (*46*) and nucleoprotein genes (*44*). The 12-nt insert in the HP H7N9 virus strains is 100% identical to a region in the polymerase basic 1 gene in multiple avian influenza A viruses, including subtypes H3N2, H6N2, and H9N2, but is not present in the polymerase basic 1 gene of HP H7N9 virus. H9N2 virus is the most frequently detected avian influenza virus in chickens in China, and the detection rate of this subtype in environmental samples from live poultry markets is ≈20% during the influenza epidemic season (*13*). Therefore, co-infection with H7N9 and other avian influenza viruses, such as influenza A(H9N2) viruses, could, in theory, lead to insertion of a polybasic cleavage site by nonhomologous recombination.

Recent studies have shown that the HP H7N9 virus is more pathogenic in mice, and more thermally stable, than low pathogenicity A/Anhui/1/2013 virus (*47,48*). Current surveillance indicates that HP H7N9 viruses have spread to several provinces in China and are responsible for large influenza outbreaks in poultry in central and northern China that show high mortality rates (http://www.fao.org/ag/againfo/programmes/en/empres/H7N9/situation_update.html). This finding raises the possibility of global dissemination of H7N9 viruses through migration of wild birds, in a manner similar to that observed for HP H5N1 viruses first identified in Guangdong Province (*32*). Although vaccination of poultry against H7N9 virus has now been implemented in some regions of China, virus adaptation and spatial distribution should be more closely monitored.

Acknowledgment

We thank Tommy Lam for performing sequence alignments.

This study was supported by the National Natural Science Foundation of China (grant 81501754) and the National Key Research and Development Program of China (grant 2016YFC1200201). T.A.B. is supported by the Medical Research Council of the United Kingdom (grant MR/L009528/1). The Wellcome Trust Centre for Human Genetics is supported by the Wellcome Trust Centre (grant 203141/Z/16/Z).

C.K., J.W., and J.L. designed the study; J.L., Y.S., L.Z., L.L., R.B., Y.J., P.Z., M.K., and L.Y. collected samples and performed genome sequencing; J.L., J.R., R.P., T.A.B., J.T., S.H., and O.G.P. analyzed data; J.L., J.R., R.P., T.A.B., and O.G.P. interpreted data; J.L., J.R., R.P., T.A.B., and O.G.P. prepared the figures; and J.L., J.R., R.P., T.A.B., and O.G.P. wrote the article

About the Author

Dr. Lu is a virologist at the Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China. His primary research interests are epidemiology, evolution, and transmission of viruses associated with emerging infectious diseases.

References

1. Xiang N, Li X, Ren R, Wang D, Zhou S, Greene CM, et al. Assessing change in avian influenza A(H7N9) virus infections during the fourth epidemic—China, September 2015–August 2016. MMWR Morb Mortal Wkly Rep. 2016;65:1390–4. http://dx.doi.org/10.15585/mmwr.mm6549a2

2. Cui L, Liu D, Shi W, Pan J, Qi X, Li X, et al. Dynamic reassortments and genetic heterogeneity of the human-infecting influenza A (H7N9) virus. Nat Commun. 2014;5:3142. http://dx.doi.org/10.1038/ncomms4142
3. Lam TT, Zhou B, Wang J, Chai Y, Shen Y, Chen X, et al. Dissemination, divergence and establishment of H7N9 influenza viruses in China. Nature. 2015;522:102–5. http://dx.doi.org/10.1038/nature14348
4. Wu J, Lu J, Faria NR, Zeng X, Song Y, Zou L, et al. Effect of live poultry market interventions on influenza A(H7N9) virus, Guangdong, China. Emerg Infect Dis. 2016;22:2104–12. http://dx.doi.org/10.3201/eid2212.160450
5. Iuliano AD, Jang Y, Jones J, Davis CT, Wentworth DE, Uyeki TM, et al. Increase in human infections with avian influenza A(H7N9) virus during the fifth epidemic—China, October 2016–February 2017. MMWR Morb Mortal Wkly Rep. 2017;66:254–5. http://dx.doi.org/10.15585/mmwr.mm6609e2
6. Wang X, Jiang H, Wu P, Uyeki TM, Feng L, Lai S, et al. Epidemiology of avian influenza A H7N9 virus in human beings across five epidemics in mainland China, 2013–17: an epidemiological study of laboratory-confirmed case series. Lancet Infect Dis. 2017;17:822–32. http://dx.doi.org/10.1016/S1473-3099(17)30323-7
7. Artois J, Jiang H, Wang X, Qin Y, Pearcy M, Lai S, et al. Changing geographic patterns and risk factors for avian influenza A(H7N9) infections in humans, China. Emerg Infect Dis. 2018;24:87–94. http://dx.doi.org/10.3201/eid2401.171393
8. Wang D, Yang L, Zhu W, Zhang Y, Zou S, Bo H, et al. Two outbreak sources of influenza A (H7N9) viruses have been established in China. J Virol. 2016;90:5561–73. http://dx.doi.org/ 10.1128/JVI.03173-15
9. Zhu W, Zhou J, Li Z, Yang L, Li X, Huang W, et al. Biological characterisation of the emerged highly pathogenic avian influenza (HPAI) A(H7N9) viruses in humans, in mainland China, 2016 to 2017. Euro Surveill. 2017;22:30533. http://dx.doi.org/10.2807/1560-7917.ES.2017.22.19.30533
10. Ke C, Mok CK, Zhu W, Zhou H, He J, Guan W, et al. Human infection with highly pathogenic avian influenza A(H7N9) virus, China. Emerg Infect Dis. 2017;23:1332–40. http://dx.doi.org/10.3201/eid2308.170600
11. Zhang F, Bi Y, Wang J, Wong G, Shi W, Hu F, et al. Human infections with recently-emerging highly pathogenic H7N9 avian influenza virus in China. J Infect. 2017;75:71–5. http://dx.doi.org/10.1016/j.jinf.2017.04.001
12. Ke C, Lu J, Wu J, Guan D, Zou L, Song T, et al. Circulation of reassortant influenza A(H7N9) viruses in poultry and humans, Guangdong Province, China, 2013. Emerg Infect Dis. 2014;20:2034–40. http://dx.doi.org/10.3201/eid2012.140765
13. Lu J, Wu J, Zeng X, Guan D, Zou L, Yi L, et al. Continuing reassortment leads to the genetic diversity of influenza virus H7N9 in Guangdong, China. J Virol. 2014;88:8297–306. http://dx.doi.org/10.1128/JVI.00630-14
14. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23:2947–8. http://dx.doi.org/10.1093/bioinformatics/btm404
15. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30:3276–8. http://dx.doi.org/10.1093/bioinformatics/btu531
16. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol. 2012;29:1969–73. http://dx.doi.org/10.1093/molbev/mss075
17. Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. PLoS Pathog. 2014;10:e1003932. http://dx.doi.org/10.1371/journal.ppat.1003932
18. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLOS Comput Biol. 2009;5:e1000520. http://dx.doi.org/10.1371/journal.pcbi.1000520
19. Edwards CJ, Suchard MA, Lemey P, Welch JJ, Barnes I, Fulton TL, et al. Ancient hybridization and an Irish origin for the modern polar bear matriline. Curr Biol. 2011;21:1251–8. http://dx.doi.org/10.1016/j.cub.2011.05.058
20. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci. 1992;8:275–82.
21. Yang Z. Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol Evol. 1996;11:367–72. http://dx.doi.org/10.1016/0169-5347(96)10041-0
22. Xiong X, Martin SR, Haire LF, Wharton SA, Daniels RS, Bennett MS, et al. Receptor binding by an H7N9 influenza virus from humans. Nature. 2013;499:496–9. http://dx.doi.org/10.1038/nature12372
23. Schrödinger L. The PyMOL molecular graphics system. Version 1.8. New York: Schrödinger, LLC; 2015.
24. Gouet P, Courcelle E, Stuart DI, Métoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics. 1999;15:305–8. http://dx.doi.org/10.1093/bioinformatics/15.4.305
25. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67:235–42. http://dx.doi.org/10.1107/S0907444910045749
26. Pond SL, Frost SD, Muse SV. HyPhy: hypothesis testing using phylogenies. Bioinformatics. 2005;21:676–9. http://dx.doi.org/10.1093/bioinformatics/bti079
27. Kosakovsky Pond SL, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol. 2005;22:1208–22. http://dx.doi.org/10.1093/molbev/msi105
28. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. PLoS Genet. 2012;8:e1002764. http://dx.doi.org/10.1371/journal.pgen.1002764
29. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol Biol Evol. 2013;30:1196–205. http://dx.doi.org/10.1093/molbev/mst030
30. Bhatt S, Holmes EC, Pybus OG. The genomic rate of molecular adaptation of the human influenza A virus. Mol Biol Evol. 2011;28:2443–51. http://dx.doi.org/10.1093/molbev/msr044
31. Raghwani J, Bhatt S, Pybus OG. Faster adaptation in smaller populations: counterintuitive evolution of HIV during childhood infection. PLOS Comput Biol. 2016;12:e1004694. http://dx.doi.org/10.1371/journal.pcbi.1004694
32. Kapczynski DR, Pantin-Jackwood M, Guzman SG, Ricardez Y, Spackman E, Bertran K, et al. Characterization of the 2012 highly pathogenic avian influenza H7N3 virus isolated from poultry in an outbreak in Mexico: pathobiology and vaccine protection. J Virol. 2013;87:9086–96. http://dx.doi.org/10.1128/JVI.00666-13
33. Subbarao K, Klimov A, Katz J, Regnery H, Lim W, Hall H, et al. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. Science. 1998;279:393–6. http://dx.doi.org/10.1126/science.279.5349.393
34. Wiley DC, Wilson IA, Skehel JJ. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature. 1981;289:373–8. http://dx.doi.org/10.1038/289373a0
35. Monne I, Fusaro A, Nelson MI, Bonfanti L, Mulatti P, Hughes J, et al. Emergence of a highly pathogenic avian influenza virus from a low-pathogenic progenitor. J Virol. 2014;88:4375–88. http://dx.doi.org/10.1128/JVI.03181-13
36. de Wit E, Munster VJ, van Riel D, Beyer WE, Rimmelzwaan GF, Kuiken T, et al. Molecular determinants of adaptation of highly

pathogenic avian influenza H7N7 viruses to efficient replication in the human host. J Virol. 2010;84:1597–606. http://dx.doi.org/10.1128/JVI.01783-09
37. Schmeisser F, Vasudevan A, Verma S, Wang W, Alvarado E, Weiss C, et al. Antibodies to antigenic site A of influenza H7 hemagglutinin provide protection against H7N9 challenge. PLoS One. 2015;10:e0117108. http://dx.doi.org/10.1371/journal.pone.0117108
38. Belser JA, Gustin KM, Pearce MB, Maines TR, Zeng H, Pappas C, et al. Pathogenesis and transmission of avian influenza A (H7N9) virus in ferrets and mice. Nature. 2013;501:556–9. http://dx.doi.org/10.1038/nature12391
39. Cattoli G, Milani A, Temperton N, Zecchin B, Buratin A, Molesti E, et al. Antigenic drift in H5N1 avian influenza virus in poultry is driven by mutations in major antigenic sites of the hemagglutinin molecule analogous to those for human influenza virus. J Virol. 2011;85:8718–24. http://dx.doi.org/10.1128/JVI.02403-10
40. Xu L, Bao L, Deng W, Dong L, Zhu H, Chen T, et al. Novel avian-origin human influenza A(H7N9) can be transmitted between ferrets via respiratory droplets. J Infect Dis. 2014;209:551–6. http://dx.doi.org/10.1093/infdis/jit474
41. Yang L, Zhu W, Li X, Chen M, Wu J, Yu P, et al. Genesis and spread of newly emerged highly pathogenic H7N9 avian viruses in mainland China. J Virol. 2017;91:e01277-17. http://dx.doi.org/10.1128/JVI.01277-17
42. García M, Crawford JM, Latimer JW, Rivera-Cruz E, Perdue ML. Heterogeneity in the haemagglutinin gene and emergence of the highly pathogenic phenotype among recent H5N2 avian influenza viruses from Mexico. J Gen Virol. 1996;77:1493–504. http://dx.doi.org/10.1099/0022-1317-77-7-1493
43. Horimoto T, Rivera E, Pearson J, Senne D, Krauss S, Kawaoka Y, et al. Origin and molecular changes associated with emergence of a highly pathogenic H5N2 influenza virus in Mexico. Virology. 1995;213:223–30. http://dx.doi.org/10.1006/viro.1995.1562
44. Suarez DL, Senne DA, Banks J, Brown IH, Essen SC, Lee CW, et al. Recombination resulting in virulence shift in avian influenza outbreak, Chile. Emerg Infect Dis. 2004;10:693–9. http://dx.doi.org/10.3201/eid1004.030396
45. Kapczynski DR, Pantin-Jackwood M, Guzman SG, Ricardez Y, Spackman E, Bertran K, et al. Characterization of the 2012 highly pathogenic avian influenza H7N3 virus isolated from poultry in an outbreak in Mexico: pathobiology and vaccine protection. J Virol. 2013;87:9086–96. http://dx.doi.org/10.1128/JVI.00666-13
46. Pasick J, Handel K, Robinson J, Copps J, Ridd D, Hills K, et al. Intersegmental recombination between the haemagglutinin and matrix genes was responsible for the emergence of a highly pathogenic H7N3 avian influenza virus in British Columbia. J Gen Virol. 2005;86:727–31. http://dx.doi.org/10.1099/vir.0.80478-0
47. Shi J, Deng G, Kong H, Gu C, Ma S, Yin X, et al. H7N9 virulent mutants detected in chickens in China pose an increased threat to humans. Cell Res. 2017;27:1409–21. http://dx.doi.org/10.1038/cr.2017.129
48. Imai M, Watanabe T, Kiso M, Nakajima N, Yamayoshi S, Iwatsuki-Horimoto K, et al. A highly pathogenic avian H7N9 influenza virus isolated from a human is lethal in some ferrets infected via respiratory droplets. Cell Host Microbe. 2017;22:615–626.e8. http://dx.doi.org/10.1016/j.chom.2017.09.008

Address for correspondence: Changwen Ke or Jing Lu, Guangdong Provincial Center for Disease Control and Prevention, 160 Qunxian Rd, Dashi Town, Panyu District, Guangdong Province, Guangzhou 514300, China; email: kecw1965@aliyun.com or jimlu0331@gmail.com

Article DOI: https://doi.org/10.3201/eid2410.171063

Molecular Evolution, Diversity, and Adaptation of Influenza A(H7N9) Viruses in China

Technical Appendix

Materials and Methods

Collection of Environmental Samples

As described (*1*–*3*), the Guangdong Provincial Center for Disease Control and Prevention launched an environmental surveillance program to monitor avian influenza viruses in live-poultry markets (LPMs) in 21 prefecture cities in April 2013. For each city, ≥1 local LPM was selected on the basis of their poultry sales and market coverage. A total of 625 LPMs were included in the environment surveillance study.

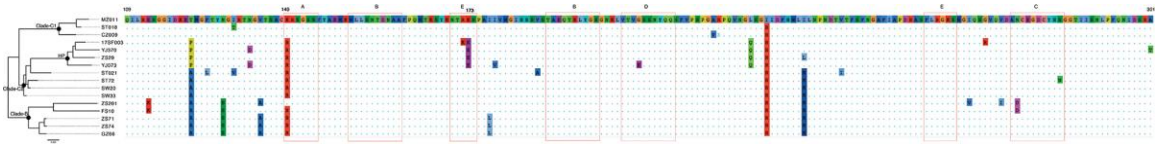
Because of seasonality in the prevalence of avian influenza, surveillance of LPMs was performed during November–May in 11 cities and each month in 10 other cities. At least 20 environmental samples were collected per market each week during the surveillance period. Environmental samples were collected as wet swab specimens from poultry feces, chicken epilator surfaces, chopping board surfaces, case surfaces, and sewage (*3*). If infection of a person with influenza A(H7N9) virus was confirmed and the person had exposure to an LPM, >20 environmental samples were collected from that market. The method of sample collection and detection used in this study has been described (*2*).

Collection of Serum Samples

Serum samples were collected from 4 patients infected with influenza A(H7N9) virus: 2 (patients P1 and P2) in 2015 and 2 (patients P3 and P4) in 2017 (Table). Respiratory specimens were collected for reverse transcription PCR and blindly passaged for 2–3 generations in embryonated chicken eggs for virus isolation. Serum samples were collected 2–3 weeks after clinical symptoms were observed and were aliquoted and stored at −70°C. H7N9 subtype virus strains (EPI1171790/A/ZS29 and EPI1171792/A/ST18) were isolated from respiratory specimens of P3 and P4 (Table).

Hemagglutination Inhibition Assay

Assays were performed according to the laboratory procedure defined by the World Health Organization (



http://www.who.int/influenza/gisrs_laboratory/cnic_serological_diagnosis_hai_a_h7n9_20131220.pdf). For H7N9 subtype virus, horse erythrocytes were used in the assay to achieve improved sensitivity. To exclude nonspecific agglutination, patient serum was treated by using hemadsorption with horse erythrocytes and then receptor-destroying enzyme treatment. In brief, 100 μL of original serum was first heat-inactivated at 56°C for 30 min and then diluted with 400 μL of phosphate-buffered saline. A total of 50 μL of packed horse erythrocytes was then added, and the sample was incubated for 15 min at 4°C for adsorption. After centrifugation, ≈400 μL of diluted serum was recovered and treated with 1.2 mL of receptor-destroying enzyme at 37°C overnight. Treated serum samples were aliquoted and stored at –70°C until used.

Hemagglutinin titers for influenza A(H7N9) virus reference strains from each clade (Table) were first determined by using hemagglutination testing. Samples were then diluted to make a working solution containing 8 hemagglutinating units/50 μL. Two-fold serial dilutions of treated serum samples (25 μL) were subsequently prepared and mixed with 25 μL of standardized virus containing 4 hemagglutinating units. Mixtures were incubated for 30 min at room temperature before hemagglutination testing. Virus back titration was performed in parallel. Each experiment was repeated twice, and HI titers (Table) shows mean values from the 2 experiments.

References

1. Lu J, Wu J, Zeng X, Guan D, Zou L, Yi L, et al. Continuing reassortment leads to the genetic diversity of influenza virus H7N9 in Guangdong, China. J Virol. 2014;88:8297–306. PubMed http://dx.doi.org/10.1128/JVI.00630-14

2. Ke C, Lu J, Wu J, Guan D, Zou L, Song T, et al. Circulation of reassortant influenza A(H7N9) viruses in poultry and humans, Guangdong Province, China, 2013. Emerg Infect Dis. 2014;20:2034–40. PubMed http://dx.doi.org/10.3201/eid2012.140765

3. Kang M, He J, Song T, Rutherford S, Wu J, Lin J, et al. Environmental sampling for avian influenza A(H7N9) in live-poultry markets in Guangdong, China. PLoS One. 2015;10:e0126335. PubMed http://dx.doi.org/10.1371/journal.pone.0126335

4. Gouet P, Courcelle E, Stuart DI, Métoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics. 1999;15:305–8. PubMed http://dx.doi.org/10.1093/bioinformatics/15.4.305

5. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67:235–42. PubMed http://dx.doi.org/10.1107/S0907444910045749

6. Wiley DC, Wilson IA, Skehel JJ. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature. 1981;289:373–8. PubMed http://dx.doi.org/10.1038/289373a0

Technical Appendix Table 1. Geographic distribution of human case of infection with influenza A(H7N9) virus (as of April 8, 2017) and HA and NA gene sequences used in phylogenetic analyses, China*

Location in China	No. clinical cases	No. HA sequences	No. NA sequences
Eastern	709	271	210
Central	134	66	77
Northern	47	12	15
Southeastern	105	26	31
Central Guangdong	211	296	219
Eastern Guangdong	61	66	58

*HA, hemagglutinin; NA, neuraminidase.

Technical Appendix Table 2. Amino acid substitutions (H3 numbering) identified on trunk branches of H7 gene of influenza A(H7N9) viruses, China*

Mutation	Associated lineage	Solvent accessibility†	Salient feature
HA1			
A1V	B	Unknown	NA
N6D	B	Unknown	NA
I48T	C2	Partial	NA
K57R	B	Full	NA
E114K	B	Partial	NA
A122P/T	**C1, C2**	Full	NA
S128N	C	Full	Within 130 loop
A135V	B, C	Full	Receptor binding residue‡
R140K	C1	Full	Antigenic site A§
K173E	C2	Full	Antigenic site E§
L177I	B, C	Inaccessible	Antigenic site E§
L226Q	C2	Partial	Receptor binding residue‡
M236I	**C1, C2**	Partial	NA
N276D	B	Full	Antigenic site C§
I326V	C2	Unknown	Proteolytic cleavage site
HA2			
E386A	C	Full	NA
E393K	C2	Partial	NA
V429I	C1	Inaccessible	NA
D476N	B	Full	NA
S489R/N	B, C	Full	NA

*Independent mutations on two or more lineages are highlighted in bold. HA, hemagglutinin; NA, not applicable.
†As defined by analysis with ESPript (*4*) using the biologic assembly of PDB: 4BSE.
‡Receptor-binding residues were determined by using CONTACT in CCP4 (*5*).
§Antigenic sites as defined by Wiley et al. (*6*).

Technical Appendix Table 3. Amino acid substitutions (N2 numbering) on trunk branches of the N9 gene of avian influenza A(H7N9) viruses, China*

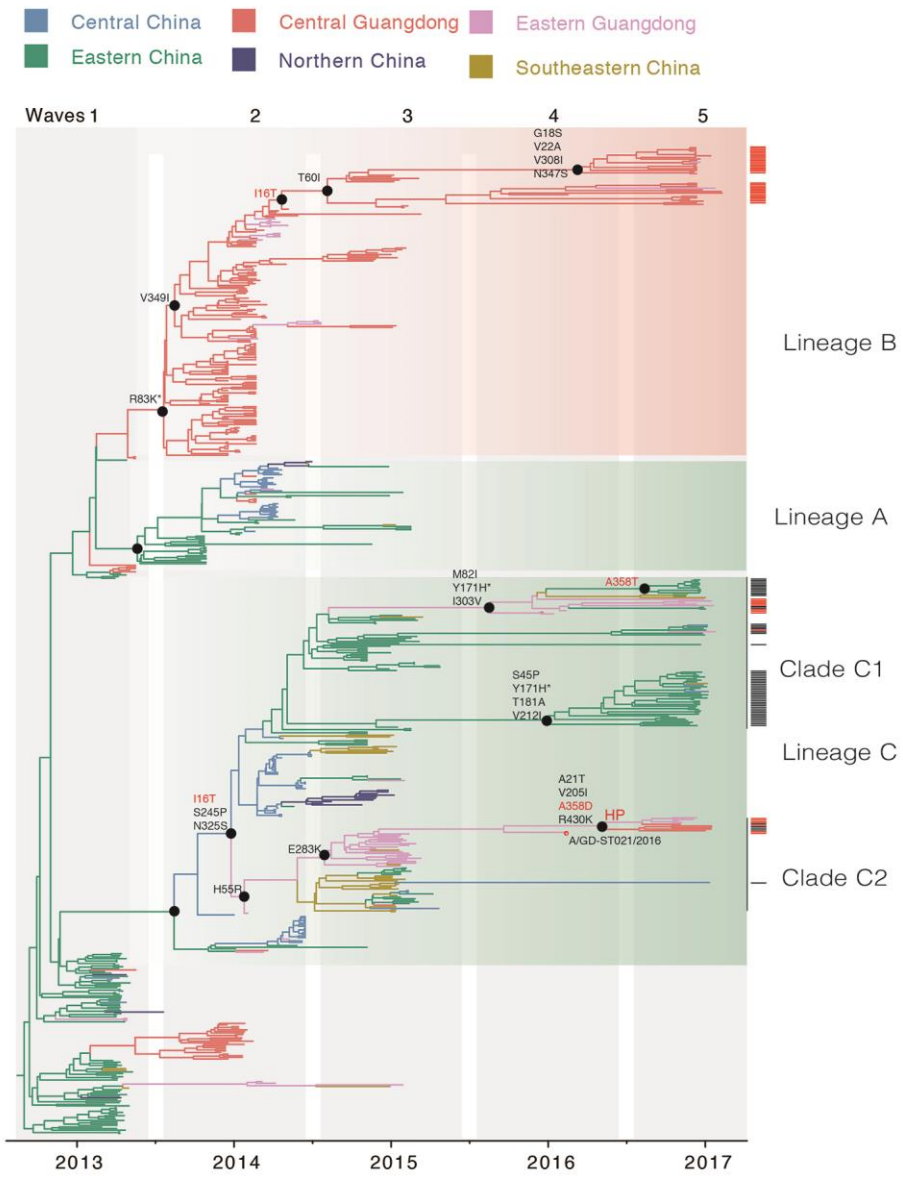
Mutation	Associated Lineage
I16T	**B, C1, C2**
G18S	B
A21T	C2
V22A	B
H55R	C2
T60I	B
M82I	C1
S45P	C1
T181A	C1
V205I	C2
V212I	C1
S245P	**C1, C2**
E283K	C2
I303V	C1
V308I	B
N325S	**C1, C2**
N347S	B
V349I	B
A358T/D	**C1, C2**
R430K	C2

*Independent mutations for ≥2 lineages are indicated in bold.

Technical Appendix Table 4. Putative sites (H3 numbering) that undergo positive selection in influenza A(H7N9) virus lineages B and C (clades C1 and C2), China*

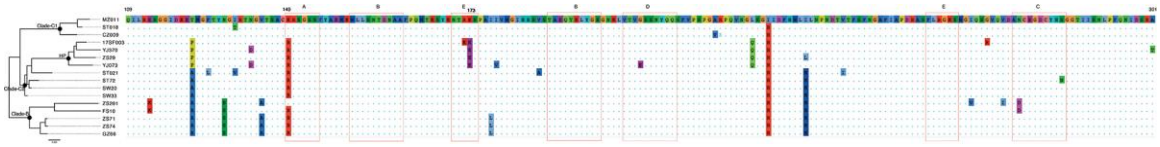
	Method			
Lineage	SLAC	FEL	MEME	FUBAR
B	None	57, 140	57, 140	57, 140, 276, 291, 493
C1	None	57, 424	57, 424	57, 114, 424
C2	226	226	3, 226, 486	226, 332

*FEL, fixed-effects likelihood; FUBAR, fast unconstrained Bayesian approximation; MEME, mixed-effects model of evolution; SLAC, single-likelihood ancestor counting.



Technical Appendix Figure 1. Bayesian maximum clade credibility tree of neuraminidase genes from influenza A(H7N9) viruses, China. Published sequences (from waves 4 and 5) from other studies are indicated by black bars to the right of the tree, and sequences reported in this study from Guangdong are indicated by red bars. Branches are colored according to geographic locations, as in Figure 1. Amino acid changes are mapped on trunk branches of B and C lineages. Parallel changes that occur in both lineages are indicated in red. Mutations are

numbered according to the N2 scheme, with the exception of mutations marked by *, which are not present in the N2 protein and are based on the N9 numbering scheme.



Technical Appendix Figure 2. Sequence alignment of hemagglutinin genes of influenza A(H7N9) virus strains used in the hemagglutination inhibition test in Technical Appendix Table 1, China. The region (109–301, H3 numbering) determining viral antigenicity is shown, and A–E antigenic epitopes are indicated by red boxes. Different colors indicate different amino acids. Dots indicate sequence identity. Scale bar indicates amino acid substitutions per site.