ECOLOGY


The authors note that Scott H. Newman, John Y. Takekawa, Diann J. Prosser, and Xiangming Xiao should be added to the author list. They should appear in the order listed above between Yujun Cui and Yarong Wu. Each of the four new authors should be credited with designing research and performing research. Each new author should also be credited as having contributed equally with H.T., S.Z., L.D., T.P.V.B., and Y.C. to this work.

The authors also note that the following statement should be added to the Acknowledgments: “We thank the international research teams that contributed to the detailed migration movement datasets that are reported elsewhere but synthesized here, including the contribution and leadership of the EMPRES Wildlife Health and Ecology Unit within the Animal Health Service of the Food and Agriculture Organization (FAO) of the United Nations. This part of the research was supported by the National Institutes of Health (NIH R01AI110128-01A1), NIH Fogarty International Center (3R01-TW007869 and 1R56TW009502-01) through the NSF/NIH Ecology of Infectious Diseases program, and National Aeronautics and Space Administration Wild–Domestic Duck Interface and H5NI Transmission (NASA) Public Health Program (NNX11AF66G). The use of trade, product, or firm names in this publication is for descriptive purposes only and does not imply endorsement by the US Government. The views expressed in this information product are those of the authors and do not necessarily reflect the views or policies of FAO.”

The corrected author line, affiliation line, and author footnotes appear below. The online version has been corrected.

Huaiyu Tian\(^a\), Sen Zhou\(^b\), Lu Dong\(^c\), Thomas P. Van Boeckel\(^d\), Yujun Cui\(^e\), Scott H. Newman\(^f\), John Y. Takekawa\(^g\), Diann J. Prosser\(^h\), Xiangming Xiao\(^i\), Yarong Wu\(^j\), Bernard Cazelles\(^k\), Shanqian Huang\(^l\), Ruifu Yang\(^m\), Bryan T. Grenfell\(^n,2\), and Bing Xu\(^o\),\(^b,2\)

\(^a\)State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, People’s Republic of China; \(^b\)Ministry of Education Key Laboratory for Biodiversity and Ecological Engineering, College of Life Sciences, Beijing Normal University, Beijing 100875, People’s Republic of China; \(^c\)Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544; \(^d\)State Key Laboratory of Pathogen and Biosafety, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, People’s Republic of China; \(^e\)Food and Agriculture Organization of the United Nations, Emergency Center for Transboundary Animal Diseases (ECTAD), Hanoi 10000, Vietnam; \(^f\)US Geological Survey, Western Ecological Research Center, Vallejo, CA 94592; \(^g\)National Audubon Society, Science Division, San Francisco, CA 94104; \(^h\)US Geological Survey, Patuxent Wildlife Research Center, Beltsville, MD 20705; \(^i\)Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019; \(^j\)Institute of Biodiversity Science, Fudan University, Shanghai 200433, China; \(^k\)École Normale Supérieure, 75230 Paris Cedex 05, France; \(^l\)Unité de Modélisation Mathématique et Informatique des Systèmes Complexes, Unité Mixte Internationale 209, Institut de Recherche pour le Développement, Centre National de la Recherche Scientifique et Ecole Normale Supérieure, 75230 Paris Cedex 05, France; \(^m\)State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, People’s Republic of China; \(^n\)Department of Geography, University of Utah, Salt Lake City, UT 84112


\(^1\)H.T., S.Z., L.D., T.P.V.B., Y.C., S.H.N., J.Y.T., D.J.P., and X.X. contributed equally to this work.

www.pnas.org/cgi/doi/10.1073/pnas.1505041112
Avian influenza H5N1 viral and bird migration networks in Asia

Huiyu Tian1,2, Sen Zhou1,2, Lu Dong1,2, Thomas P. Van Boeckel1,2, Yujun Cui1,2, Scott H. Newman1,2, John Y. Takekawa3,4, Dianni J. Prosser1,2, Xiangming Xiao5,6,1, Yarong Wu4, Bernard Cazelles1,5,6, Shangqian Huang4, Ruifu Yang4, Bryon T. Grenfell1,7,2, and Bing Xu8,9,10

1State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, People’s Republic of China; 2Ministry of Education Key Laboratory for Earth System Modelling, Center for Earth System Science, and School of Environment, Tsinghua University, Beijing 100084, People’s Republic of China; 3Ministry of Education Key Laboratory for Biodiversity and Ecological Engineering, College of Life Sciences, Beijing Normal University, Beijing 100875, People’s Republic of China; 4Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544; 5State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, People’s Republic of China; 6Food and Agriculture Organization of the United Nations, Rome, Italy; 7Center for Transboundary Animal Diseases (ECTAD), Hanoi 10000, Vietnam; 8US Geological Survey, Western Ecological Research Center, Vallejo, CA 94592; 9National Audubon Society, Science Division, San Francisco, CA 94104; 10US Geological Survey, Patuxent Wildlife Research Center, Beltsville, MD 20705; 11Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019; 12Institute of Biodiversity Science, Fudan University, Shanghai 200433, China; 13Ecologie et Evolution, UMR 7625, Université Pierre et Marie Curie, Centre National de la Recherche Scientifique et Ecole Normale Supérieure, 75230 Paris Cedex 05, France; 14Unite de Modélisation Mathématique et Informatique des Systèmes Complexes, Unité Mixte Internationale 209, Institut de Recherche pour le Développement and Université Pierre et Marie Curie, 93142 Bondy Cedex, France; 15Fogarty International Center, National Institutes of Health, Bethesda, MD 20892; and 16Department of Geography, University of Utah, Salt Lake City, UT 84112

Edited by Burton H. Singer, University of Florida, Gainesville, FL, and approved November 26, 2014 (received for review March 20, 2014)

The spatial spread of the highly pathogenic avian influenza virus H5N1 and its long-term persistence in Asia have resulted in avian influenza panzoootics and enormous economic losses in the poultry sector. However, an understanding of the regional long-distance transmission and seasonal patterns of the virus is still lacking. In this study, we present a phylogeographic approach to reconstruct the viral migration network. We show that within each wild fowl migratory flyway, the timing of H5N1 outbreaks and viral migrations are closely associated, but little viral transmission was observed between the flyways. The bird migration network is shown to better reflect the observed viral gene sequence data than other networks and contributes to seasonal H5N1 epidemics in local regions and its large-scale transmission along flyways. These findings have potentially far-reaching consequences, improving our understanding of how bird migration drives the periodic reemergence of H5N1 in Asia.

Significance

Highly pathogenic avian influenza virus H5N1 first emerged in Asia and subsequently unfolded into the first avian influenza panzootic, causing major economic losses in the poultry sector. However, we still do not understand the regional long-distance transmission and seasonal patterns of H5N1. In this study, we addressed this issue by combining H5N1 outbreak records, whole-genome sequences of viral samples, and satellite tracking data for four species of migratory birds in Asia. We show that timing of H5N1 outbreaks and viral migration are closely associated with known bird migration routes. The flyway is the major viral transmission barrier to the intracontinental spread of H5N1 by migratory birds in Asia, whereas geographic distances within the flyways have little effect on H5N1 transmission.

Migratory birds play important roles in the geographic spread of various zoonotic agents (1). Among these agents, the avian influenza viruses (AIVs) have been shown to be transmitted over long distances during the seasonal migration of birds (2, 3). Wild waterfowl, in particular, are considered the natural reservoir of low-pathogenic avian influenza (LPAI) viruses and have been shown to spread LPAI viruses along migratory flyways in Asia, Africa, and the Americas (4–7). However, one of the fundamental unknowns remaining is the role played by wild birds in the regional spread of AIV (4–7).

Highly pathogenic avian influenza (HPAI) H5N1 first appeared in Asia in 1996 (12), and subsequently spread to Europe, the Middle East, and Africa, causing many human casualties and major economic loss in the booming Asian poultry sector. Despite the low transmissibility of HPAI H5N1 from birds to humans and from humans to humans, the high fatality rate reported in humans after the onset of the epidemic and the potential for H5N1 to become pandemic through migratory bird flyways raised serious concerns (13). The Qinghai lineage of H5N1, in particular, expanded from Qinghai to Eurasia and into the Indian subcontinent and northern and central Africa along migratory flyways. It was also shown experimentally that some species of birds shed the virus before the onset of clinical signs or with no clinical signs (14, 15). This suggests that the large-scale transmission of H5N1 by migratory birds could potentially go undetected. Using satellite telemetry, Gaidet et al. reported that one infected white-faced whistling duck (Dendrocygna viduata) survived HPAI H5N2 infection and was able to migrate for at least 655 km, when tracked with a satellite transmitter for 47 d (16). Other studies have shown that the direction of the geographic spread of HPAI H5N1 is consistent with the major bird migration routes (17, 18). A number of studies have also suggested that long-distance migration may lead to immunosuppression in birds and migratory performance is negatively affected by viral infections (19–21). However, it should be noted that HPAI H5N1 is rarely reported in living and healthy wild birds (22–24).

Recently, HPAI H5N1 clade 2.3.2, the dominant subclade in Asia, was detected in migratory birds during their migration in Mongolia, South Korea, and Japan, and was shown to be associated with wild waterfowl infections (25–27). Furthermore, an isolate of HPAI H5N1 from a common buzzard (Buteo buteo) in Bulgaria showed close genetic proximity to clade 2.3.2.1 isolates from wild birds in the Tyva Republic and Mongolia, suggesting

2To whom correspondence may be addressed. Email: bingxu@tsinghua.edu.cn or grenfell@princeton.edu.
3This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1405216112/-/DCSupplemental.
that the HPAI H5N1 viruses of clade 2.3.2 have spread westward and pose a public health threat (28). These numerous studies have directed our attention to the roles played by migratory birds in the spread of HPAI H5N1 viruses in the last decade.

In this study, we constructed networks of bird and viral gene migrations to evaluate the role of migratory birds in the spread of HPAI H5N1 clades 2.3.2 before 2007, and 2.3.2.1 on and after 2007 (clade 2.3.2 for abbreviation) in Asia. We assembled a unique database of satellite tracking data on wild bird migration patterns, records of HPAI H5N1 outbreaks, and both the viral hemagglutinin (HA) gene and whole-genome nucleotide sequences over the period 2003–2012. The objective of this study was to analyze the association between the networks of bird migration, the networks of viral gene flow, and the timing of HPAI H5N1 outbreaks at different geographic locations.

**Results**

**Bird Migration Routes and Virus Sampling in the Research Regions.** Bird migration routes were acquired from the global positioning system (GPS) tracking data (SI Appendix, Table S1) for four bird species: the bar-headed goose (Anser indicus), swan goose (Anser cygnoides), ruddy shelduck (Tadorna ferruginea), and northern pintail (Anas acuta). The swan goose and northern pintail winter in southern China and Southeast Asia (SEA), breed in Mongolia or the northeast Asian region (NEA, including the Republic of Korea and Japan) along the East Asian–Australasian (EA) flyway. The bar-headed goose and ruddy shelduck winter in South Asia (India, Nepal, Bangladesh, and Myanmar) and breed in Qinghai Province or in Mongolia, along the Central Asian (CA) flyway (Fig. 1A).

HPAI H5N1 outbreaks were geocoded along both the EA and CA flyways (Fig. 1B) and time series, corresponding to the monthly number of HPAI H5N1 outbreaks, were generated in each country or province of China. A wavelet time series analysis between locations revealed that the outbreak lags varied along the flyways (Fig. 2 and SI Appendix, Figs. S1 and S2).

A phylogenetic analysis showed that HPAI H5N1 clades 2.3.2 was structured into two distinct subgroups, corresponding to the CA flyway and EA flyway (Fig. 1C). The viruses isolated from Northeast Asia, southern China, and Southeast Asia showed close genetic proximity. Similarly, viral samples from Qinghai Lake, South Asia, and Mongolia, showed significant genetic proximity. In particular, isolates from Mongolia showed genetic associations with both subgroups, which is attributable to the intersection of these flyways in this region.

**Timing of H5N1 Outbreaks and Bird Migration Patterns.** The outbreaks of HPAI H5N1 in Asia showed a seasonal pattern, but the timing of the outbreaks was highly variable at the country level (or province level in China) across the different flyways. A strong positive correlation was observed between the annual epidemic lag and the average flight time between successive locations along the flyways (Fig. 2A) ($R = 0.70, P < 0.01$ for the CA flyway and $R = 0.78, P < 0.01$ for the EA flyway). Epidemic velocity and bird migration speed (expressed in kilometers per month) were positively associated, indicating that they share similar spatio-temporal patterns along the CA and EA flyways (Fig. 2B and C).

However, a notable exception was observed for Japan and Hong Kong. Faster migration was associated with a faster-spreading epidemic wave. The overall epidemic velocity and migration speed were 1344.79 km/mo (832.04 km/mo when Japan and Hong Kong were excluded) and 789.64 km/mo, respectively, along the EA flyway, and 607.26 and 573.19 km/mo, respectively, along the CA flyway. Epidemic lag and flight duration were not significantly associated with the spatial distance (SI Appendix, Fig. S3), indicating that the relationship between the annual epidemic lag and bird migration along the flyways was not attributable to their underlying correlation with spatial distance.

**Viral Migration Through Regions Along Flyways.** To investigate the association between viral spread and bird migration, a phylogenetic tree of clades 2.3.2 of HPAI H5N1 was constructed based on the sequences of the viral HA gene. The viruses showed genetic proximity with similar isolation time and geographic region, whereas the viruses isolated from geographically distinct regions along one flyway were frequently mixed and no single location seeded every annual epidemic at other locations (Fig. 1C). This observation could result from a scenario in which ancestor viruses were carried by migratory birds and were thus spread to and burst in local regions along the flyway.

To validate whether viral migration in the study regions was associated with bird migrations, we estimated the rates of viral migration across the six regions (Qinghai Lake, South Asia, Mongolia, Southeast Asia, Northeast Asia, and southern China; Fig. 1C) using Bayesian stochastic search variable selection (BSSVS). Evidence of viral migration (i.e., differences in the internal viral genome segments) was observed in these regions (Table 1 and SI Appendix, Fig. S5). When we repeated the BSSVS with larger prior means to determine the number of
nullzero rates, the significant nonzero migration rates remained unchanged (SI Appendix, Table S5).

The viral migration network was jointly estimated from all eight viral gene segments in a single analysis (Fig. 3A). It was insensitive to the sample sizes in the main migration paths (SI Appendix, Table S8). The gene flow structures were similar to the bird migration network (Fig. 3B), for the gene flow between South Asia, Qinghai, and Mongolia (the CA flyway) and for the gene flow through southern China, Mongolia, Japan, and the Republic of Korea (the EA flyway). The EA-1 and EA-2 flyways were defined by the migration routes. EA-1 started at Hong Kong and progressed through southern China, the Republic of Korea, and Japan. EA-2 started at Poyang Lake and progressed through southern China, the Republic of Korea, and Mongolia. The linkage between southern China and Japan–Korea was not strongly supported by the Bayes factor (BF), perhaps because there were gaps in the viral sampling in these two areas. Most of the isolates from southern China were clustered in 2003–2008, whereas the samples from Japan–Korea were clustered in 2009–2011. The migration rates were higher within the flyways than between the flyways (two tailed t tests, P < 0.01; SI Appendix, Fig. S6). The flyways yielded largely congruent results, and the level of gene flow within a flyway was usually higher than that between flyways (3).

Coincidence Between Viral Transmission and Bird Migration Networks. The network of HPAI H5N1 gene flow in Asia was build based on a phylogenetic analysis (Fig. 4). Using randomization tests, we showed that the observed gene flow was best explained using a flyway-based random network (FRN), either within the CA or the EA flyway, rather than by a purely random network (RN) including all geographic locations (Fig. 5A and Table 2). The significance of this result was assessed with a two-tailed t test (P < 0.01). These results reject the hypothesis of random mixing between flyways and suggest that migratory flyways act as a major barrier to the intracontinental spread of HPAI H5N1. Within each FRN, 1,000 random networks were created and compared with the network of bird migrations to reveal potential alternative transmission routes that might better explain the observed gene flow. The results of this analysis showed that the bird migration network best explained the pattern of gene flow in both the EA and CA flyways (Fig. 5B, Table 2, and SI Appendix, Fig. S7). For example, the trajectory Nepal–Qinghai–Mongolia–Bangladesh–India is a likely network for the CA flyway and is consistent with the observed bird migration pattern. The bird migrations were more complex in the EA flyway, and two different flyways were identified: the EA-1 flyway, Hong Kong–Guangdong–Shantou Special Economic Zone–Zhejiang–Republic of Korea–Eastern Siberia–Japan; and the EA-2 flyway, Poyang Lake–Hubei, Hunan–Mongolia–Republic of Korea. Both networks are consistent with the observed migration networks and also showed a good fit with the gene flow in the EA flyway (Fig. 5B, Table 2, and SI Appendix, Fig. S7).

Finally, the inclusion of bird migration histories (Fig. 4B) in the migration networks outperformed the measurement of genetic and spatial divergences when viral migration histories were not considered (Table 2 and SI Appendix, Fig. S8). This was not unexpected because spatial distance measurements that ignore the history of viral migration may fail to reveal the underlying spatial–genetic pattern. In particular, the cycle of bird migrations best described the viral gene flow, and geographic distance had little effect on viral transmission along the flyway. This suggests that spatial distance is not a major ecological barrier to gene transfer within a flyway. Our network analysis consistently demonstrated that migratory birds and HPAI H5N1 share the same migration network, and that the virus could have evolved and spread along the bird migratory pathways in Asia. A randomization analysis of subsets of isolates throughout the FRN analysis revealed weak correlation between the number of isolates per locality and the spatial–genetic coefficient (SI Appendix, Fig. S9).

Discussion

Our study provides unique insight into the association between viral transmission networks, bird migration networks, and the timing of H5N1 outbreaks in Asia. Our results showed that in the period 2003–2012, the regular migration of the HPAI H5N1 virus occurred between migratory avian flyways and that bird migrations coincided with the peaks of HPAI H5N1 epidemics. This suggests that the annual epidemic lags for H5N1 outbreaks are associated with bird migrations along these flyways. The robustness of our findings on the gene flow dynamics of HPAI H5N1 was tested against 1,000 artificially generated random networks. The network analysis performed in this study indicated that the spread of clades 2.3.2 of HPAI H5N1 probably shared the same geographic pattern as the network of bird migrations.

Dozens of studies have discussed the spatiotemporal distribution of H5N1 and the potential effect of covariates on the distribution of its epidemics (29, 30). In this study, we have shown with a wavelet analysis that the temporal lags between epidemics are associated with the flight duration between successive locations along flyways. Our results further support a strong correlation between bird migration speed and epidemic velocity (Fig. 2). We combined satellite tracking data with a phylogenetic analysis to study the spatiotemporal spread of HPAI H5N1 across Asia. We have demonstrated how migration history can be associated with genetic distance and how transmission networks can be inferred from genetic information using a statistical framework. Our analyses demonstrate that the spread of HPAI H5N1 correlates with bird migration networks. Four bird vector species for H5N1 were used to analyze the bird migration...
patterns and represented the major flyways in Asia well. We observed that the contribution of each bird species to the gene flow network was generally consistent with the migration pattern of each species: northern pintail (A. indicus), northern pintail (A. acuta) along the EA-1 flyway, swan goose (A. cygnoides), and ruddy shelduck (T. ferruginea) along the CA flyway. The viral samples used in this study were recovered from most Asian countries in which HPAI H5N1 has been circulating throughout the last decade (31, 32). Because isolates from wild birds constituted a large proportion of the total samples, these viral samples provide a good opportunity to estimate the relationship between the transmission patterns of HPAI H5N1 and bird flyways (SI Appendix, Fig. S10).

Several important limitations of the present analysis should be noted. First, satellite tracking data on wild bird migrations were not available for some areas located along the flyways, in particular in Southeast Asia, where H5N1 is known to be endemic. Second, the birds’ behavior and migration routes might have been affected by the solar-powered GPS trackers. Third, all of the isolates from clades 2.3.2 from Southeast Asia and Yunnan, Guangxi, and Guizhou Provinces in China were taken from poultry samples (SI Appendix, Fig. S10) and showed a close genetic association, so they were excluded from the analysis. Fourth, information on the sampling locations of the isolates from Mongolia was unavailable, and the potential habitats of the birds were scattered across different locations in the country (Fig. 1A). Finally, no gene reassortment was considered in the analysis. Gene reassortment may influence the evolutionary rate of the virus and therefore affect the performance of our models. Furthermore, the rates of sequence evolution were insensitive to coalescent models among the branches of the clade 2.3.2 viruses (SI Appendix, Table S6). Future surveillance efforts should include more frequent and geographically extensive sampling of HPAI H5N1 to improve our knowledge of the association between the spread of the H5N1 virus and bird migration (33).

It has been suggested that the poultry trade, illegal wild trade, and wild bird smuggling may play important roles in the global dispersal of H5N1 (34–36). Therefore, bird migration networks may not be the only route by which H5N1 is transmitted (17). As shown in SI Appendix, Fig. S11, the isolates from wild birds constituted only 20% of the samples isolated along the Asia flyways (SI Appendix, Table S7), and the bird migration network was not the optimal network explaining the patterns of gene flow for all clades. Therefore, we infer that the bird migration network may not be the main route of dispersal for all of the clades, but is one of several underlying transmission networks. Moreover, although some of the H5N1 viruses were isolated from wild birds, their presence may have been the immediate result of a previous outbreak in poultry, and vice versa. This bias is inevitable because of the current sampling strategy. In our analysis, we considered isolates from poultry and from H5N1 outbreaks in poultry based on the most reasonable assumption. However, to address the uncertainties in the model, we require a better surveillance strategy in future studies.

Our results indicate that the transmission of HPAI H5N1 is strongly associated with bird migrations. This calls for increased collaboration between the countries located along the same flyways. Based on the complex dynamics of viral transmission identified in this study, we suggest that future efforts to control HPAI H5N1 should not only include regional surveillance, but
Table 2. Randomization test and bird migration results for the network contribution to patterns of HPAI H5N1 gene flow

<table>
<thead>
<tr>
<th>Model</th>
<th>Migration history</th>
<th>R</th>
<th>P</th>
<th>R mean</th>
<th>R SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>RN†</td>
<td>With</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>FRN†</td>
<td>With</td>
<td>—</td>
<td>—</td>
<td>0.69</td>
<td>0.01</td>
</tr>
<tr>
<td>RCN</td>
<td>With</td>
<td>—</td>
<td>—</td>
<td>0.52</td>
<td>0.05</td>
</tr>
<tr>
<td>EA-1-RN*</td>
<td>With</td>
<td>—</td>
<td>—</td>
<td>0.64</td>
<td>0.01</td>
</tr>
<tr>
<td>EA-1-BN*</td>
<td>With</td>
<td>0.68</td>
<td>&lt;0.01</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>EA-1-BN*</td>
<td>Without</td>
<td>0.64</td>
<td>&lt;0.01</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>EA-2-RN*</td>
<td>With</td>
<td>—</td>
<td>—</td>
<td>0.74</td>
<td>0.01</td>
</tr>
<tr>
<td>EA-2-BN*</td>
<td>With</td>
<td>0.77</td>
<td>&lt;0.01</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>EA-2-BN*</td>
<td>Without</td>
<td>0.76</td>
<td>&lt;0.01</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CA-RN†</td>
<td>With</td>
<td>—</td>
<td>—</td>
<td>0.77</td>
<td>0.02</td>
</tr>
<tr>
<td>CA-BN†</td>
<td>With</td>
<td>0.83</td>
<td>&lt;0.01</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CA-BN‡</td>
<td>Without</td>
<td>0.80</td>
<td>&lt;0.01</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

RN, random network; FRN, flyway random network; EA, East Asian-Australasian flyways; EA-1, the migration started at Hong Kong; EA-2, the migration started at Poyang Lake; CA, Central Asian flyways; RCN, randomly categorized network; BN, bird migration network. Migration history, measuring spatial distance with or without migration history. R is the spatial–genetic correlation coefficient, mean is the average R value, and SD is the SD of 1,000 random networks.

†Randomization tests, virus transmission was assumed to follow a random network.

‡Migration network tests, virus transmission was assumed to follow the bird migration network.

§Migration history tests, virus transmission was assumed without migration history (Fig. 4C), but the cycle length of network was included.

that flyway-based strategies should be implemented, involving all of the countries affected by potential HPAI H5N1 outbreaks. A similar effort is also required to coordinate preventive and interventional strategies, to improve vaccine and prevent nonuniform vaccination policies, which could result in the silent spread of the virus. Vaccination is a key component in the control of HPAI H5N1 in China, Vietnam, and Egypt. However, it seems that such strategies produce unpredictable results (37, 38). Both the evolutionary rates of the virus and the number of positively selected sites were higher in virus populations in countries using an avian influenza vaccine for H5N1 (39). Furthermore, the H5N1 virus can replicate in clinically healthy vaccinated birds (37). Therefore, it is likely that the selection pressure resulting from vaccination influences the viral ecology in countries located along the migratory flyways.

Materials and Methods

Data Collection. The HPAI H5N1 outbreak data were obtained from the Food and Agriculture Organization of the United Nations and the World Organization for Animal Health. The information included the dates and locations of outbreaks, and the host species. The records from both sources were combined and all duplicate reports were removed. Our database contains information on 17,250 independent outbreaks, collected between December 2003 and April 2012 in Asia. Although this dataset contains the most comprehensive outbreak records available, some outbreaks may have gone unreported because a passive surveillance system was used.

The satellite tracking data on wild bird migrations along the EA flyways and CA flyway were collected by the United States Geological Survey. There are four stages in the annual cycles of migratory birds: nonbreeding, spring migration breeding, and autumn migration. The annual cycles are reflected in the time series data. We selected four species of birds as representative of the migrations that occur in the study area: the northern pintail (A. acuta), swan goose (A. clypeata), ruddy shelduck (T. ferruginea), and bar-headed goose (A. indicus) (SI Appendix, Table S1). This choice was based on the population abundances in the study area, the type of clinical symptoms displayed when infected with the HPAI H5N1 virus, and the birds’ virus shedding capacities (14, 40, 41). The arrival dates, departure dates, and flight durations between locations were extracted from the GPS tracking data for each bird species (provided in SI Appendix, Table S3).

The nucleotide sequences of HPAI H5N1 were obtained from the GenBank database, hosted by the National Center for Biotechnology Information, on October 31, 2012 (42). All segments were aligned using the default settings in MUSCLE v3.5 (43). The dataset sizes were as follows: HA, n = 2,379 (alignment length of 1,856 nucleotides); NA, n = 196 (1,410 nt); PB1, n = 194, 2,271 nt; PB2, n = 196, 2,148 nt; NP, n = 196 (1,494 nt); MP, n = 200 (982 nt); and NS, n = 196 (838 nt). The 142 H5N1 sequences were combined with the nucleotide sequences (excluding partial sequences <70% of the full-length sequence) of all influenza A viruses available in GenBank. The final numbers of sequences of the H5N1 viruses for the EA flyways and CA flyway are given in SI Appendix, Table S2.

Each H5N1 sequence was assigned a geographic area according to its province/country of isolation (SI Appendix, Tables S3 and S4). The spatial distance between each pair of countries/provinces was calculated from the latitude and longitude of each state/province center. We collected and analyzed 297 sequences of viruses from clades 2.3.2.2 that were isolated in 2003–2012 (Fig. 1B) in these regions (Fig. 1C), including China (n = 120), Bangladesh (n = 15), India (n = 7), Japan (n = 41), Laos (n = 6), Mongolia (n = 21), Nepal (n = 10), the Republic of Korea (n = 39), and Vietnam (n = 38). The sequences from China were categorized according to locations (Fig. 1B).

Phylogenetic Analysis. Neighbor-joining trees of the H5N1 HA sequences were constructed using a GTR + I + Γ model in PAUP v4.0b10 (44). The best nucleotide substitution model was selected with the Akaike information criterion and a hierarchical likelihood ratio test in ModelTest (45, 46). To assess the robustness of the tree topology, a set of 100 pseudoreplicates was performed and used in our analyses with the neighbor-joining (ML) method implemented in PHYML (47) and the neighbor-joining method implemented in PAUP v4.0b10. These trees were highly congruent to those produced with PAUP above. The sequences associated with clades 2.3.2 were included in our analyses (SI Appendix, Table S2).

Estimates of Viral Migration Through Discrete Geographic Regions. We used a nonreversible continuous-time Markov chain model to estimate the migration rates between regions along flyways and the general patterns of regional circulation of the H5N1 virus. This analysis was restricted to the six geographic regions identified above. In this analysis, the constant size, exponential growth, and Bayesian skyline coalescent prior were used with a strict and uncorrelated lognormal relaxed molecular clock, and a general time-reversible (GTR) nucleotide substitution model in BEAST v2.1 (48). The performance of each combination was compared using BF (SI Appendix, Table S6) (49). We performed three independent analyses of 50–100 million generations. These analyses were combined after the removal of an appropriate burn-in (10–20% of the samples in most cases), with 5,000 generations sampled from each run for a total of 15,000 trees and parameter estimates.

BSSVS was used to infer the phylogeographic diffusion processes under a parsimonious scenario (50). BSSVS searches for the transition rates that are consistent with the data, and efficiently infers the possible bias arising from the uneven sample sizes at each location, we included in our analyses (SI Appendix, Table S2).

The supported state transitions of gene segments were recovered from the independent Bayesian analyses, which indicated the potential persistence of a virus in a metapopulation mixing between geographic regions. The viral migration network was jointly estimated from all eight gene datasets in a single analysis even though the taxon number for each gene dataset was not identical (53). Within-flyway rate estimates were compared with between-flyway rate estimates to measure the viral diffusion patterns. To investigate any possible bias arising from the uneven sample sizes at each location, we performed a sensitivity test using increasingly larger random subsets of isolates (from 50%, 60%, 70%, and 80–90% of the original dataset) throughout the BSSVS procedure, with 50 replicates for each gene (SI Appendix, Table S8).

Network Analyses to Determine Viral Gene Flow and Migration. Under the assumption that viruses circulate across a fixed network with a mutation rate drawn from an underlying normal distribution (estimated with BEAST), then the spatial distance measured in the migration history could reflect the genetic distance between all possible pairs of viral samples in the network. Also, the differences in sampling times among the same or distinct localities can be transformed into the number of viral spread cycles within the network (SI Appendix, SI Materials and Methods, and Eq. 53). The GTR + I + Γ substitution model was used to estimate the pairwise genetic distances between sequences, and the possibly underestimated migration distances of the ancestral virus were identified and not included (SI Appendix, SI Materials and Methods, and Fig. S12).
Phylogeographic models were constructed to determine the extent of the viral gene flow that was influenced by bird migration and to explore the ML transmission networks. In the case that the data arise from an undetected migration process along an unobserved network. Our model can be described as follows: (i) A thousand RN models incorporate all of the sampling locations in a homogeneous network. FRN models incorporate all of the locations within an existing flyway, either CA or EA, according to their sampling locations. All of the scenarios of viral migration between flyways, including the east-west migration paths between Mongolia and northeast Asia, and Qinghai and southern China are considered to display viral migration between the EA and CA flyways, due to the fact that the flyways are connected. Randomly connected networks (RCNs), incorporate localities that were randomly categorized into any of the flyways, regardless of their actual sampling locations, but corresponding to the number of locations along the EA and CA flyways. (ii) Random networks were created to find the best-fitting network for each flyway, which best explained the pattern of gene flow, and then compared with the bird migration network. A total of 1,000 RNs were created in each flyway, as for the FRN and RCN model, but the EA flyway was divided into the EA-1 flyway (starting at Hong Kong) and EA-2 flyway (starting at Poyang Lake). The spatial distances were considered as the routines of bird migration within flyway (Fig. 4A). (iii) Spatial distances, using direct measurements (simple measurements of spatial distances between strains in which the viral migration history was not considered) or measurements with the migration history (the spatial distance traveled by the viral ancestors is summed), were compared between geographic loca-
tions in the FRN model. All networks were one directional and each node was passed only once.

ACKNOWLEDGMENTS. We thank the US Geological Survey, Western Ecological Research Center, Patuxent Wildlife Research Center, and Avian Influenza Program. This research was supported by the Ministry of Science and Technology, China, National Natural Science Foundation of China (2012CB719300 and 2012CB719500), the Ministry of Science and Technology, China, and the National Natural Science Foundation of China (41271099). B.T.G. was supported by the Science and Technology Directorate, US Department of Homeland Security (Contract HSHQDC-C00058 and the Research and Policy Foundation for Avian Influenza Disease Dynamics, the Bill and Melinda Gates Foundation, and the Fogarty International Center, National Institutes of Health. We thank the international research teams that contributed to the detailed migration movement datasets that are reported elsewhere but synthesized here, including the cooperation with GMPRS Wildbird Health and Ecology Unit within the Animal Health Service of the Food and Agriculture Organization of the United Nations. This part of the research was supported by the National Institutes of Health (NIH R01AI0102780-01A1, NIH Fogarty International Center R01WI007865 and 1R01WI009502-01) through the NSF/NIEH Ecology of Infectious Diseases program, and National Aeronautics and Space Administration Wild-Domestic Duck Interface and HSNI Transmission (NASA Public Health Program (NNX11AF66G)). The use of trade, product, or firm names in this publication is for descriptive purposes only and does not necessarily reflect the views or policies of FAO.