



Synchrony of Bird Migration with Global Dispersal of Avian Influenza Reveals Exposed Bird Orders

Received: 29 May 2023

Accepted: 23 January 2024

Published online: 06 February 2024

Check for updates

Qiqi Yang ¹✉, Ben Wang², Phillippe Lemey ³, Lu Dong⁴, Tong Mu ⁵, R. Alex Wiebe¹, Fengyi Guo ¹, Nidia Sequeira Trovão ⁶, Sang Woo Park ¹, Nicola Lewis^{7,8}, Joseph L.-H. Tsui⁹, Sumali Bajaj⁹, Yachang Cheng¹⁰, Luojun Yang¹¹, Yuki Haba¹, Bingying Li², Guogang Zhang^{12,13}, Oliver G. Pybus^{8,9,13}, Huaiyu Tian ^{2,13}✉ & Bryan Grenfell ^{1,5,13}✉

Highly pathogenic avian influenza virus (HPAIV) A H5, particularly clade 2.3.4.4, has caused worldwide outbreaks in domestic poultry, occasional spillover to humans, and increasing deaths of diverse species of wild birds since 2014. Wild bird migration is currently acknowledged as an important ecological process contributing to the global dispersal of HPAIV H5. However, this mechanism has not been quantified using bird movement data from different species, and the timing and location of exposure of different species is unclear. We sought to explore these questions through phylodynamic analyses based on empirical data of bird movement tracking and virus genome sequences of clade 2.3.4.4 and 2.3.2.1. First, we demonstrate that seasonal bird migration can explain salient features of the global dispersal of clade 2.3.4.4. Second, we detect synchrony between the seasonality of bird annual cycle phases and virus lineage movements. We reveal the differing exposed bird orders at geographical origins and destinations of HPAIV H5 clade 2.3.4.4 lineage movements, including relatively under-discussed orders. Our study provides a phylodynamic framework that links the bird movement ecology and genomic epidemiology of avian influenza; it highlights the importance of integrating bird behavior and life history in avian influenza studies.

Since 2014, highly pathogenic avian influenza viruses (HPAIVs) subtype H5, especially clade 2.3.4.4, have caused unprecedentedly large numbers of wild bird deaths worldwide¹. In contrast to previous clades of the A/goose/Guangdong/1996 (Gs/GD) lineage, there have also been more persistent spillovers to local domestic poultry^{2–6}, impacting the poultry farming industry. Despite no onward human-to-human transmission to date, occurrences of zoonotic jumps to humans pose potential threats to public health^{7–12}.

Plausible ecological mechanisms for the global movement of HPAIVs include the live poultry trade and wild bird migration^{13,14}. Endemic in Asia, clade 2.3.2.1 spread throughout the Eurasian and

African continents (Fig. S2). Since its outbreak in wild birds at Qinghai Lake in 2005¹⁵, studies have shown that viral dispersal at long distances is related to wild bird migration^{16,17}. Global dispersal of clade 2.3.4.4 continues to provide virological, epidemiological, and ecological evidence in support of the critical role of migratory wild birds in HPAIV spread and evolution at a global-scale. Compared to previous clades, clade 2.3.4.4a during 2014/15 outbreaks was less pathogenic to some species while being more effectively transmitted^{18–20}, possibly enabling infected birds to migrate between continents. Subsequent phylodynamic work confirmed that the introduction of clade 2.3.4.4a into Europe and North America most likely occurred via long-distance

A full list of affiliations appears at the end of the paper. ✉e-mail: qiqiy@princeton.edu; tianhuaiyu@gmail.com; grenfell@princeton.edu

flights of infected migratory birds²¹. During the 2016/17 outbreaks, the major circulating clade 2.3.4.4b was more transmissible²² and more virulent²³, related to multiple internal genes^{22,23} and potentially their frequent reassortments^{3–5}. Later phylogenetic analysis showed a clear link between the reassortments and migratory birds, as most reassorted gene segments were from migratory wild birds and originated at dates and locations that corresponded to their hosts' migratory cycles²⁴. Integrating host movement in studying HPAIV dispersal is important, though challenging. One challenge is insufficient bird movement data, which limited previous global-scale studies²¹ in accounting for the high variation in bird behaviors across species and locations.

Another challenge of studying HPAIV dispersal in wild birds is the lack of HPAIV prevalence data. Only a few studies document longitudinal HPAIV prevalence in wild bird populations²⁵. Compared to HPAIV, low pathogenic AIV (LPAIV) has better long-term surveillance of infections or seroprevalence and related avian host ecology in disparate bird habitats, e.g., the United States Geological Survey surveillance of birds in Alaska²⁶. While longitudinal records provide insights into the role of life history and ecology of local bird communities in LPAIV circulation²⁷, their conclusions are limited to local dynamics and cannot be easily generalized. To resolve this challenge, ideally, we should have systematic global surveillance for HPAIV. However, this is impossible due to resource constraints.

Instead, we could design effective surveillance strategies by identifying exposed avian species and high-risk geographical regions. Recently, researchers have addressed these questions at a higher taxonomic level to include more diverse species. For example, Hill et al. compared the different roles of species within the Anseriformes and Charadriiformes in the dispersal and spillover of AIVs²⁸. They concluded that wild geese and swans are the main source species of HPAIV H5, while gulls spread the viruses most rapidly. Hicks et al. found that the inter-species transmission of AIVs in North America is positively associated with the overlap of habitats, suggesting the importance of local bird community diversity²⁹. However, they did not use empirical bird movement data. Furthermore, given the heterogeneous biogeographical pattern of bird migrations, identifying geographical hotspots requires linking global and local scales.

To fill this gap, we here focus on two questions related to the contributions of birds, locally and globally, to the spatiotemporal dynamics of HPAIV H5 viruses; specifically, i) how does seasonal bird migration facilitate global virus dispersal and ii) which avian species are exposed to HPAIV H5 and where? To explore these questions, we first illustrate the global circulation history of clade 2.3.4.4 (2010–2017; 2018–2023) and clade 2.3.2.1 (2010–2017) using time-scaled phylogeographic analyses of hemagglutinin (HA) genes of HPAIVs sampled from wild birds and poultry. The split of clade 2.3.4.4 data is based on its history of evolution, epidemiology and sampling intensity, enabling us to maintain the genetic diversity of the dataset while keeping the computational burden manageable. Building upon previous evidence, we propose possible routes of long-distance virus dispersal. We acknowledge two caveats. First, while we only included HA, the neuraminidase (NA) and internal genes contribute to virus evolution, e.g., via reassortment²⁴ and the major circulating subtype shift from H5N8 to H5N1 in clade 2.3.4.4. Second, the geographical bias of virus sampling has a strong impact on the virus lineage movement routes, especially for locations under-sampled. Based on the estimated routes and inferred virus dispersal history, we quantify the contribution of seasonal bird migrations to global virus dispersal and evolution. Second, we model the monthly geographical distribution of bird orders using species distribution models based on environmental factors and bird tracking data. We evaluate the risks of bird orders being exposed to HPAIV H5 at geographical origins and destinations of virus lineage movement by analyzing the statistical association of local bird distributions and virus lineage migration. Our study provides an approach

that integrates bird migration ecology in HPAIV epidemiological studies to disentangle the mechanisms of interaction between HPAIV and wild birds.

Results

Seasonal bird migration associates with global HPAIV H5 dispersal

Is the wide geographical range of HPAIV H5 caused by frequent introductions from one region to another, or a single introduction resulting in subsequent spread within the area? The discrete trait phylogeographical analysis of clade 2.3.4.4 HA genes (before 2018) exhibits scarce virus lineage movements between aggregated regions, most of which are unidirectional (Fig. 1B). It suggests that inter-regional viral introductions over long geographical distance occur at low frequency and in one direction. Furthermore, the sequences are highly clustered by region, implying viral persistence within each region after introduction. Clade 2.3.2.1 shows a similar pattern (Fig. S8). In contrast, clade 2.3.4.4 after 2018 (Fig. 1C) presents more mixed virus lineage movements between regions, suggesting more frequent dispersal between regions. In addition, the trunk composition of the phylogeny shows that most virus diversity is contained in Russia, Africa and Europe. While the two trees of clade 2.3.4.4 (before and after 2018) have shared branches (Fig. 1A), the resulting correlation does not qualitatively affect the subsequent analysis (Fig. 2, Fig. S4).

To test quantitatively whether seasonal bird migration is a key predictor of HPAIV H5 dispersal, we fit a generalized linear model (GLM) parameterization of the discrete phylogeography using a Bayesian model selection procedure^{30,31}. Concurrently, we consider seasonal bird migration, live poultry trade and poultry population size as covariates of the diffusion rates between regions. To incorporate the potential seasonal difference in viral dispersal, we model a time-heterogeneous phylogenetic history³² with three seasons based on bird annual cycle in North Hemisphere: non-migration (mid-November to mid-February, mid-May to mid-September), spring migration (mid-February to mid-May) and fall migration (mid-September to mid-November). Fig. 2 shows the posterior estimates of the inclusion probabilities and conditional effect sizes (on a log scale) of the covariates. It reveals that seasonal bird migration is the dominant driver of the global virus lineage movements of HPAIV H5 clade 2.3.4.4. This is shown by both the log conditional effect size of the seasonal bird migration (for clade 2.3.4.4 (2010–2017), mean log conditional effect size [95% highest probability density interval, HPDI]: -2.35 [0.87, 3.85]; clade 2.3.4.4 (2018–2023): -4.71 [3.60, 6.22]) and the statistical support for its inclusion (for clade 2.3.4.4 (2010–2017), posterior probability (PP) ~ 0.88 and Bayes factor (BF) ~ 56; clade 2.3.4.4 (2018–2023), PP ~ 1 and BF > 16565).

In contrast, poultry population size and the international live poultry trade are not strongly or not consistently associated with the inter-region dispersal of HPAIV H5 (Fig. 2). This is also evident in both the effect size and the statistical support, e.g., the log conditional effect size of the live poultry trade (for clade 2.3.4.4 (2010–2017): -0.62 [0.19, 1.17]; clade 2.3.4.4 (2018–2023): -0.35 [0.08, 0.66]) and the statistical support for its inclusion (for clade 2.3.4.4 (2010–2017), PP ~ 0.60 and BF ~ 12; clade 2.3.4.4 (2018–2023), PP ~ 0.11 and BF ~ 1). For clade 2.3.2.1, neither wild bird migration nor live poultry trade is a major contributing predictor for the inter-regional virus lineage movements.

The viral sample size at origin location might affect the results (clade 2.3.4.4 (2010–2017): -1.22 [-1.75, -0.68], PP ~ 1 and BF > 2647; clade 2.3.4.4 (2018–2023): -0.86 [-1.13, -0.05], PP ~ 0.99 and BF ~ 866; clade 2.3.2.1: -1.52 [-1.85, -1.19], PP ~ 1 and BF > 16,565. Unfortunately, because of the imprecision of host species records, it is challenging to down-sample the data while maintaining genetic diversity. Nevertheless, when controlling for viral sample size at the origin

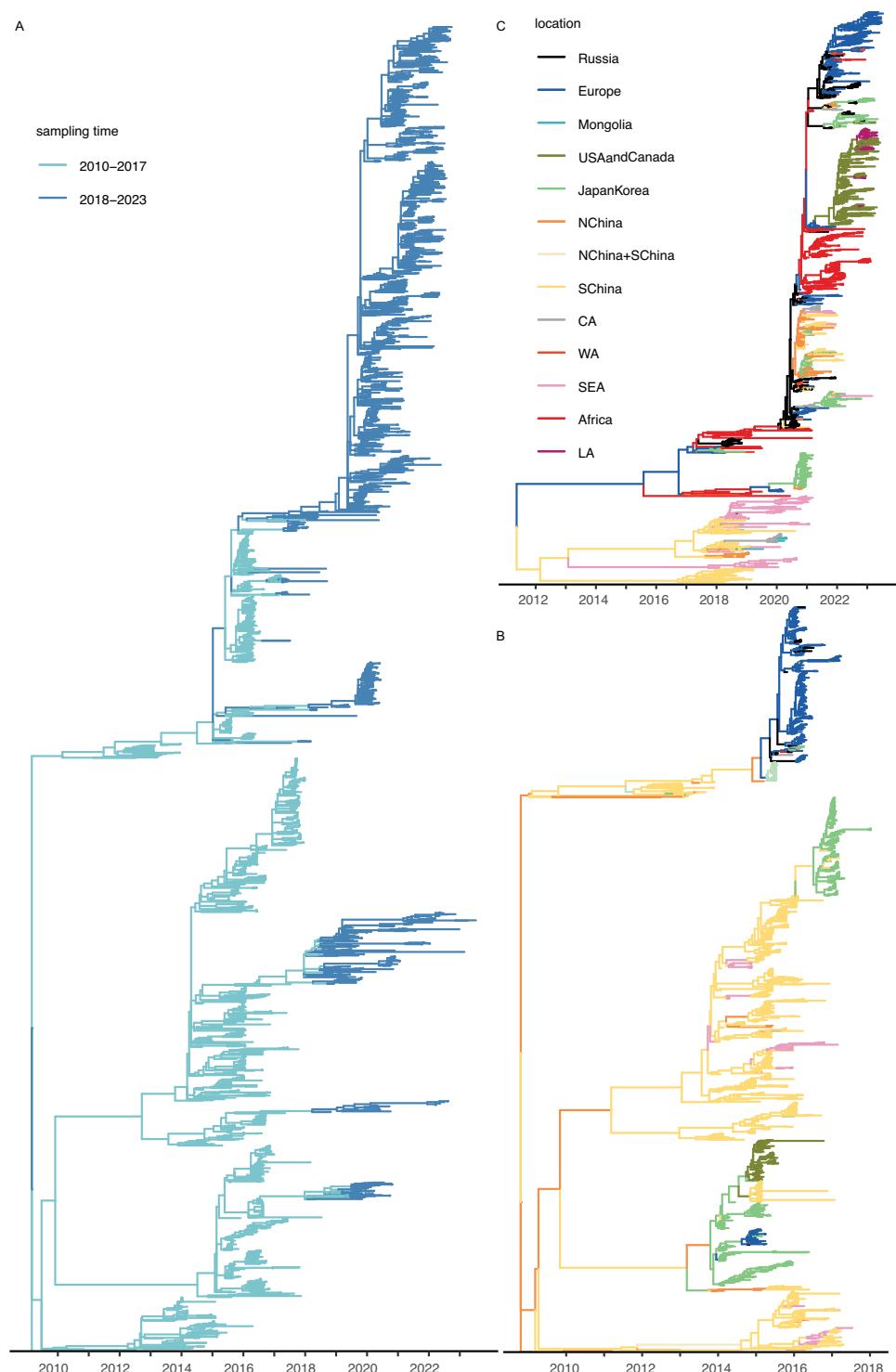


Fig. 1 | The evolutionary history of HPAIV H5 clade 2.3.4.4 and 2.3.2.1. **A** Time-scaled maximum clade credibility (MCC) tree of clade 2.3.4.4 (2010–2023). Sequences sampled from 2010 to 2017 are colored light blue, and those sampled from 2018 to 2023 are darker blue. The phylogeographic reconstructions for the sequences sampled from 2010 to 2017 (**B**) and for those from 2018 to 2023 (**C**),

shown as time-scaled MCC trees with location annotations to summarize the reconstructions. Abbreviations of locations: NChina (North China), SChina (South China), SEA (South-East Asia), CA (Central Asia), WA (Western Asia), LA (Latin America). Supplementary Dataset 3 lists countries in each aggregated region.

location and directly comparing wild bird migration and live poultry trade, the wild bird migration is clearly supported over the live poultry trade (Table S2). Based on these results, we used subsequent analyses to understand the importance of different bird species at the order taxonomy level in the global dispersal and local emergence of HPAIV H5.

Exposed migratory bird orders at origin and destination regions of HPAIV H5 virus lineage movement

We identified virus dispersal routes (Bayes factor > 3) between the aggregated regions in the Northern Hemisphere (Fig. 3A) using the previous phylogeography analyses. Seasonality is reflected in northward and southward virus lineage movements of clade 2.3.4.4.

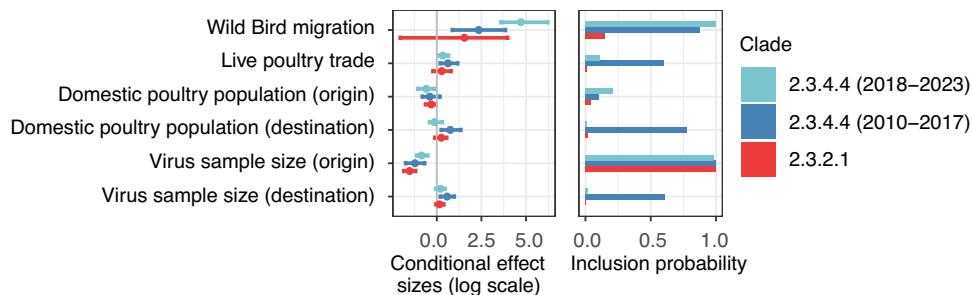


Fig. 2 | Contributions of predictors to worldwide diffusion of H5N1 clade 2.3.2.1 and 2.3.4.4. The virus dispersal is inferred from HA genes by GLM-extended Bayesian phylogeographic inference with heterogeneous evolutionary processes through time. Predictors in the model are shown in Fig. S6. The conditional effect size of each predictor (on a log scale) is presented as mean values with credible

intervals. The phylogeographic GLM estimates were obtained from $n = 1845$, 1844 , and 1163 viral sequences sampled over 12 , 10 , and 11 locations for clade 2.3.4.4 (2018–2023), 2.3.4.4 (2010–2017), and 2.3.2.1, respectively. The light blue, darker blue, and red colors indicate clade 2.3.4.4 (2018–2023), 2.3.4.4 (2010–2017), and 2.3.2.1, respectively.

Furthermore, it overlaps well with the bird migration seasonality. Before 2018, most virus lineage movements (14 of 20) show a single temporal peak (Figs. S3.3–3.4, 3A). The peaks of the northward routes overlap with spring bird migration and/or wintering period (upper rows of Fig. 3A, Fig. S3.3). Only one route (Japan-Korea to USA-Canada) overlaps with the summer breeding period. Most southward virus lineage movements peak during the Fall bird migration period, although some peaks continue in November when birds might still be migrating along some routes (lower half of Fig. 3A, Fig. S3.4). Only one route (Europe to Qinghai) overlaps with the wintering period. In summary, in the Northern Hemisphere, virus lineage movements from south to north occur mainly during the wintering period and spring bird migration, while southward virus lineage movements occur mainly during the fall migration period when birds fly to the south. Since 2018, southward virus lineage movements have expanded the time period to late summer and early winter in addition to fall (Fig. S3.6); northward virus lineage movements shift to mainly spring and summer (Fig. S3.5).

This association of seasonality in bird migration and HPAIV H5 lineage movement of clade 2.3.4.4 implies that breeding grounds might be potential pools of HPAIV H5 genetic diversity for southward virus lineage movements associated with bird migration; wintering grounds might play a similar role for the northward viral lineage movements. Additionally, the results show more virus lineage movements during the fall migration (Southward Markov Jump counts: 11.71 ([8, 16] 95% HPDI, before 2018) and 78.76 ([68, 89], 2018–2023) per month, September–November) than the spring migration (Northward Markov Jump counts: 9.83 ([6, 14], before 2018) and 1.15 ([1, 3], 2018–2023) per month, March–May). Virus lineage movements also have higher relative frequency during the fall migration (shown in the higher peak in Fig. 3). Interestingly, birds also migrate in a larger abundance in the fall than during spring, as the population size becomes larger after breeding.

In contrast, for clade 2.3.2.1, virus lineage movements show no seasonal variation across the year (Fig. S3.1–3.2). This difference might be due to the fact that largely wild birds contribute to the virus genetic diversity of clade 2.3.4.4, while that of 2.3.2.1 is not (Fig. S1); this suggests that clade 2.3.4.4 might be endemic in wild birds in some regions, i.e., wild birds might serve as the reservoir.

Which bird orders might be exposed to HPAIV H5 at the origin or destination regions of virus lineage movements? To explore this question, we examined the synchrony of bird migration and virus lineage movements by testing the correlation between the 2 monthly time series. We then used block bootstrapping to test the statistical significance (See Methods). Along some routes, virus lineage movement frequency is correlated with monthly bird order distribution of the origin/destination location (Table S3). We assume that the bird

population's birth and mortality are stable, and there is no external intervention. Therefore, the change in bird distribution probability is only due to a positive net flux of incoming individuals through migration or a negative net flux of individuals via migration outside a given area. We highlight the following results for clade 2.3.4.4 between 2018 and July 2023 (Fig. S7, Table S3):

- Southern migration of Ciconiiformes from Europe synchronizes with more frequent viral dispersal from Europe to Africa ($r = -0.47$, 97.5% confidence intervals, CI: $[-0.65, -0.29]$, $p = 3.58 \times 10^{-5}$), suggesting that Ciconiiformes might be exposed to clade 2.3.4.4 when they start migrating south from Europe.
- Southern migration of Passeriformes to Africa synchronizes with more frequent viral dispersal from Western Asia to Africa ($r = 0.52$, 97.5% CI: $[0.33, 0.79]$, $p = 5.38 \times 10^{-5}$), suggesting that Passeriformes might be exposed to the viruses when they migrate south to Africa.
- Northern migration of Pelecaniformes from Japan/Korea synchronizes with more frequent viral dispersal from Japan/Korea to Russia ($r = -0.32$, 97.5% CI: $[-0.46, -0.03]$, $p = 8.96 \times 10^{-5}$), suggesting that Pelecaniformes might be exposed to the viruses when they migrate north from Japan/Korea.

For clade 2.3.4.4 before 2018, we detected two virus lineage movement routes, both in Asia, with Anseriformes and Accipitriformes as exposed bird orders, respectively (Fig. 3, Table S3). We do not detect bird orders associated with virus lineage movements of clade 2.3.2.1 (Table S3). Although we cannot conclude any causal relationship from current analyses, the synchrony of virus lineage movement and bird migration suggests the birds might be exposed at the origin/destination location. Future work needs to be conducted to understand the causal relationship.

Discussion

Here, we report a phylodynamic analysis linking the spatial ecology of avian hosts and HPAIV H5 virus lineage movements. Our results support previous findings on the important role of bird migration in disseminating HPAIV H5 clade 2.3.4.4²¹. We found that the seasonal wild bird migration network is associated with the global diffusion and evolutionary dynamics of HPAIV H5. A previous study found that the 2014/2015 outbreaks of HPAIV H5 in Europe and North America were likely introduced by wild bird migration²¹ by comparing the inferred ancestral host-type and location traits of the viral genome sequences²¹. Our study advances this finding by directly integrating the bird migration network into the virus phylogeographic reconstruction. In addition, we found that the inter-regional live poultry trade is not associated with the global HPAIV H5 dispersal, consistent with

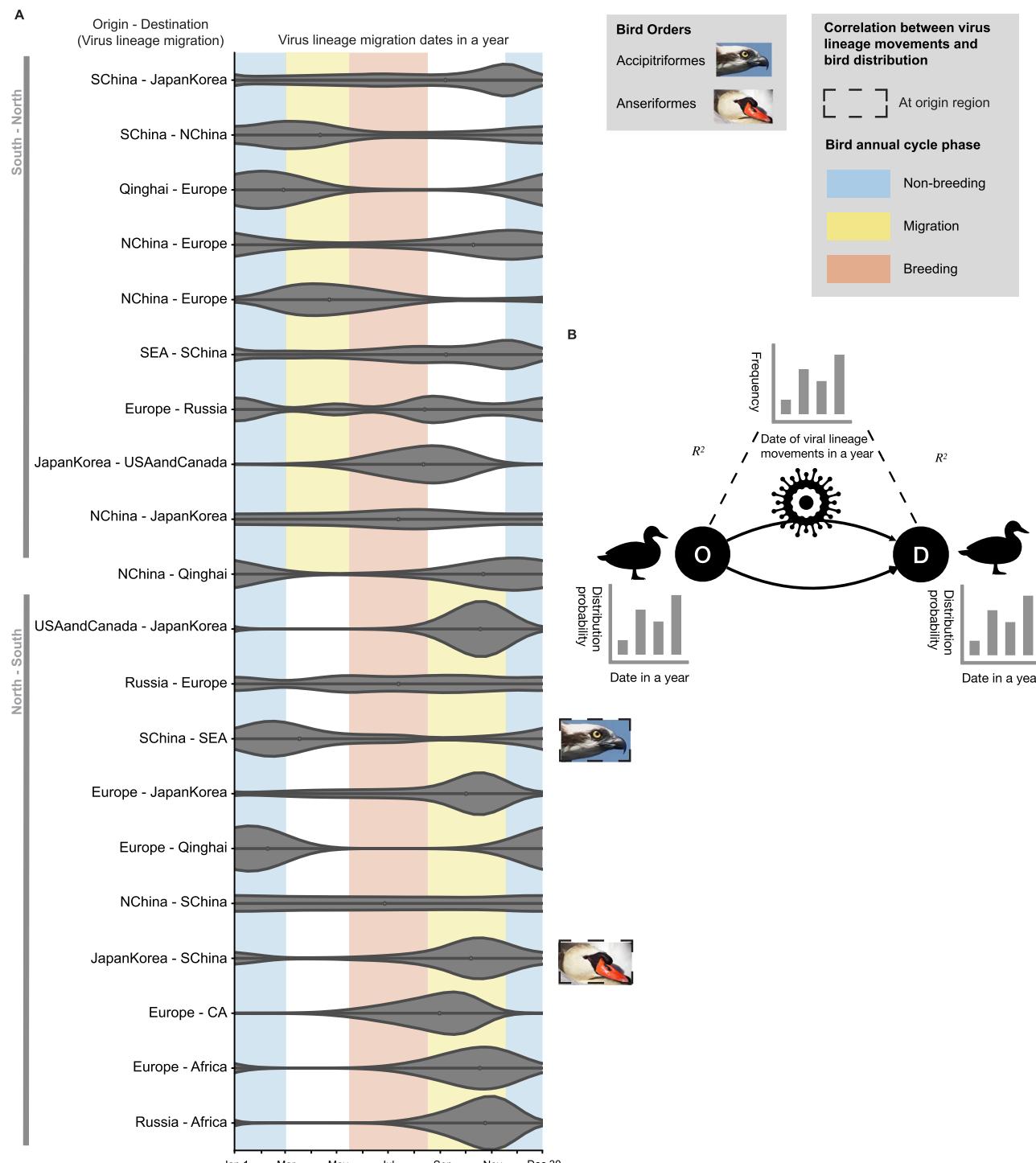


Fig. 3 | The synchrony of virus lineage movements between regions and bird distribution probability at the regions. A Probability density distribution of the virus lineage migration throughout the year, between locations summarized from the discrete trait phylogeography of HPAIV HS clade 2.3.4.4 and the Markov jump counts (Section “Discrete trait phylogeography of HPAIVs and counts of virus lineage migration”). X axis: Virus lineage migration dates in a year; Y axis: origin region - destination region of the virus lineage migration. The width of the violins represents the virus lineage migration probability density. Non-breeding (blue), migration (yellow) and breeding (red) bird annual cycle phases in general are shown in the south-north migration direction and in the north-south migration

direction. Boxes around bird photos show the statistically significant correlation of virus lineage movements and bird order distribution at origin, destination or both regions. We used the following images of bird species from the Macaulay Library at the Cornell Lab of Ornithology: *Pandion haliaetus haliaetus* (MLS2357727) and *Cygnus olor* (ML72775261). Abbreviations of locations are the same as in Fig. 1. **B** Schematic diagram of cross-correlation analyses of virus lineage movement between two locations (O: origin, D: destination) and the bird distribution probability at each location. Block bootstrapping was used to calculate confidence intervals and two-tailed p values. See details in Methods “Discrete trait phylogeography of HPAIVs and counts of virus lineage migration”.

previous studies^{13,21}. The same previous study found that the international poultry trade's direction is opposite to the global spread direction of HPAIV H5 clade 2.3.4.4²¹. Another study demonstrated large-scale HSNI transmission dynamics are structured according to different bird flyways and driven by the Anatidae family, while the Phasianidae family, largely representing poultry, is an evolutionary sink¹³.

Our results provide insights into relatively under-studied bird orders that may be relevant to HPAIV H5 dispersal. Historically, studies of host-pathogen interaction of AIV have focused on Anseriformes as wild bird hosts. However, many other under-studied orders have been affected by clade 2.3.4.4 recently^{33,34}. Interaction of different avian orders might contribute to virus dispersal and local persistence²⁸. A previous study showed that host origins of HPAIV H5 reassorted genes include Anseriformes, other groups of wild birds, and domestic poultry²⁴. Our results highlight the exposure risk of specific under-studied bird orders in addition to Anseriformes, including Accipitriformes, Ciconiiformes, Passeriformes, and Pelecaniformes. Particularly, Passeriformes are rarely considered in avian influenza studies. It is possible that when Passeriformes have migrated south to Africa, they might be exposed to avian influenza from interacting or sharing habitat with the waterfowls in Africa, as Asia-Eastern Africa is a major migratory flyway where waterbirds, Passeriformes and other species migrate from Asia to Eastern Africa for wintering. However, due to the use of light-level geolocators (GLS) in some Passeriformes tracking studies, our results might be geographically inaccurate given the coarse spatial resolution (~200 km) of GLS location estimates³⁵.

Our results also highlight the seasonal pattern of virus dispersal of clade 2.3.4.4. Despite the low relative frequency of virus lineage movements during summer breeding and wintering, they may serve as a gene pool for following virus lineage movement during the migration. A previous study emphasizes the important role of the breeding period in inter-species virus transmission in North America²⁹. Previous surveillance also shows that LPAI prevalence in waterfowls is higher during the wintering period of Eurasian migratory birds in Africa³⁶. Interestingly, clade 2.3.2.1 does not show seasonality in virus lineage movements. The non-seasonal pattern might be because domestic poultry contributes to the virus genetic diversity more than wild birds for clade 2.3.2.1 (Fig. S1), while wild birds contribute much more to the reassortment of the viruses for clade 2.3.4.4 (Fig. S1, ²⁴). The difference also suggests that clade 2.3.2.1 might not be endemic in wild birds, while wild birds might be the reservoir for clade 2.3.4.4.

By comparing clade 2.3.4.4 before 2018 and after (Fig. S4), our results suggest wild bird migration might be the same mechanism for geographical global dispersal of clade 2.3.4.4 throughout the years; differently, the viruses have become endemic in birds in wider geographical regions since 2018, and breeding time might have become more important in virus dispersal.

Caveats. First, our data in *bird tracking* are much limited in species diversity due to low data availability in the tracking community⁽³⁷⁾. Despite variations in migration behaviors within bird orders, with the data, we are not able to account for the variations. Thus, it should be cautioned that not all species within the bird order we found relevant to the virus lineage movements may be exposed to HPAIV H5. Nevertheless, our tracking data are relevant for avian influenza circulation: the families we use in tracking data (Supplementary Dataset 2) included major wild bird species that have been sampled avian influenza viruses (Table S1). In addition, currently available tracking data may be more abundant in individuals that are easier to capture. Despite the use of the species distribution model to account for the empirical data bias, the uneven abundance of tracking data in species makes the model more accurate for species that are represented by better data. We are also not able to include the migration volume of birds in the migration network (Fig. S6). With the increasing availability of bird tracking data, our framework can further study the association of birds at the species level and virus lineage movements; we may find more

associated species that are relatively under-studied; and we can also advance our binary bird migration network by integrating comprehensive bird movement models^{38,39}.

Secondly, the *virus genome* data are biased geographically and in host species. The under-sampling of viruses in some areas hugely impacted the inferred phylogeography. For example, we cannot conclude if the inferred viral lineage movement from Europe to Qinghai or Japan-Korea occurs directly or if geographically proximate areas, e.g., central Eurasia, are middle stops of the movement due to under-sampling in central Eurasia. Despite including sampling size in the phylogeographical analysis, we cannot adjust the geographical sampling biases due to the unknown magnitude of infections at locations. Therefore, our results might be biased in highlighting the regions with better virus sampling. Fortunately, the sampling efforts in some historically under-sampled and no-sampled areas are growing, e.g., in Australia⁴⁰. In the future, given more extensive and evenly sampled spatial data, we may find historically under-studied regions have risks of virus lineage movements. To avoid the effect of geographical sampling bias in phylogeography (although not for no sampling areas), our framework can also easily adapt structured coalescent methods, e.g., MASCOT⁴¹, when being applied to less geographical locations or less sequence data. In addition to the geographical sampling bias in virus genome data, our results may be affected by poultry-sampled viruses. To preserve genetic diversity, we included virus samples from domestic poultry when inferring virus diffusion. Therefore, some patterns in the results could reflect virus transmission between domestic poultry and spillover from wild birds to poultry rather than bird migratory patterns. In the future, with better-recorded host species in virus genome data, we can better understand the interaction and contribution of poultry and wild birds in virus evolution and dispersal, including hosts' contribution to reassortment. Currently, we only considered the HA gene when inferring AIV diffusion and evolution. HA is a key gene in influenza viruses, as it is the receptor-binding and membrane fusion glycoprotein of influenza virus and the target for infectivity-neutralizing antibodies⁴². However, the reassortment events of all internal genes are also important in the dispersal and evolution of HPAIV H5²⁴.

Finally, when analyzing the statistical association between bird distribution and virus lineage movements, we do not include *time lags* between the two time series. Therefore, some bird orders that synchronize with virus lineage movements with time lags may not be detected in our current analyses. Those birds may be exposed to the viruses because of inter-species transmission or environmental transmission if they are possible transmission routes among the birds - it should be noted, however, that currently, little is known about the mechanisms of transmission in wild birds, particularly inter-species transmission among seabirds. With finer-scale bird tracking data on movements other than migratory movements⁴³, future studies could better integrate avian biology in understanding the inter-species transmission mechanism. With longitudinal bird tracking data at locations and virus sampling from the birds, our analysis framework could be advanced to understand the time lags of the synchrony and different roles of bird species in HPAIV H5 dispersal (Table S4). Another limitation in our methods is that bird migration networks in our analyses are binary. Therefore, when we model a non-bird-migration time period in Fig. 2, the binary network does not represent no migration; instead, it represents that all the relative migration rates between locations are the same.

In conclusion, allocating more resources for global surveillance of avian influenza viruses in wild birds would enhance our ability to tackle the challenges of more virulent and transmissible HPAIV H5 spreading in wild birds. To achieve this goal, it is critical to understand "where and in which bird species surveillance is most needed and could have the greatest impact"¹⁴. Given sufficient data in the future, our framework could help conservation and public health policy-making in

designing monitoring and surveillance strategies. More collaboration is needed between ornithologists, movement ecologists, bird conservation experts, avian influenza epidemiologists, disease ecologists and virologists on many aspects, including collaborative data collection/surveillance of AIV and data sharing. For example, if studies were to simultaneously obtain the movement tracking of bird populations and their serology and virology surveillance data, then they could link the bird movement directly with the virus transmission and dispersal. In addition, we need more AIV samples from water bodies to better understand environmental transmission. With such data, we would be able to understand the viral transmission at local scales and therefore develop disease models for bird conservation and potential zoonotic threats.

Methods

Viral sequence data and time-scaled phylogeny of HPAIVs

To infer the phylogeny of avian influenza HPAIV H5 viruses, we accessed sequences of HA genes, NA genes and six internal gene segments from GISAID (Global Initiative on Sharing All Influenza Data^{44–46}). Using the sequences, we estimated a maximum likelihood phylogeny (Fig. S1) for each gene segment, respectively, under a GTR+ γ nucleotide substitution model, with the randomly selected strains as representatives, by FastTree v2.1.4⁴⁷. Genotypes of internal gene segments (Fig. S1) were defined by clustering patterns with background sequences in a previous study⁴⁸. On the phylogeny, the viruses with internal genes from wild birds, e.g., clade 2.3.2.1 and clade 2.3.4.4, showed wider geographical spread^{21,49}, compared to poultry viruses, e.g., clade 2.3.4.1 and clade 2.2, despite the high similarity of their HA genes. This demonstrates the importance of gene reassortment in the evolution and transmission of HPAIVs.

In this study, we focus on clade 2.3.4.4 and clade 2.3.2.1. We inferred their time-scaled phylogenies of HA genes. Before the inference, to test for the presence of phylogenetic temporal structure, we generated a scatterplot of root-to-tip genetic divergence against sampling date by TempEst v1.5⁵⁰. Strong phylogenetic temporal structure was detected in the phylogeny of each clade (Fig. S5). The final datasets were i) 1163 HA sequences of clade 2.3.2.1; ii) 1844 HA sequences of clade 2.3.4.4 (2010–2017); iii) 1845 HA sequences of clade 2.3.4.4 (2018–2023, after down-sampling). For clade 2.3.4.4 (2018–2023), we randomly down-sampled over-represented locations (with sequences more than 250) to 262 for each location to get a similar size of sequences across locations and to get a similar size of total sequences as previous clade 2.3.4.4. The spatial and temporal distribution of the sequences is shown in Fig. S2.

Time-resolved HA phylogenies were estimated using the Markov chain Monte Carlo (MCMC) approach implemented in BEAST v1.10.4⁵¹ with the BEAGLE library⁵². We used an uncorrelated lognormal relaxed molecular clock model⁵³, the SRD06 nucleotide substitution model⁵⁴ and the Gaussian Markov random field Bayesian Skyride coalescent tree prior⁵⁵. For each dataset, MCMC chains were run for 300 million (clade 2.3.2.1) and 400 million (clade 2.3.4.4) generations with burn-in of 10%, sampling every 10,000 steps. Convergence of MCMC chains was checked with Tracer v1.7⁵⁶. A set of 1000 trees for each clade was subsampled from the MCMC chain and used as an empirical tree distribution for the subsequent analysis.

Animal mobility networks and their contribution to HPAIV phylogeography

The bird migration network (Fig. S6) was summarized by searching publicly available migration data on Movebank. An edge between two locations in the network exists if any migration tracking record shows bird migration originating from the start location to the end location. The location-wise live poultry trade values were summed up from country-wise import and export of the live poultry recorded on United Nations Comtrade Database (comtrade.un.org/data/). We accessed the

total net weight and trade value from 1996 to 2016 of live poultry, including fowls of the species *Gallus domesticus*, ducks, geese, turkeys and guinea fowls. Since there are no accessible data of the within-country poultry trade in China, we adapted the inferred poultry trade accessibility between provinces of China from a previous study⁵⁷. Based on the ratio of the inferred accessibility and the empirical trade value between Hong Kong SAR and mainland China, we scaled all the accessibility to the trade value flows among Qinghai, North China and South China.

With the summarized seasonal-varying bird migration network, we statistically quantified the contribution of wild bird migration to avian influenza virus dispersal. We applied the GLM extended Bayesian phylogeography inference³¹ with the 1000 empirical trees as the input. The aggregated locations in the previous discrete trait phylogeography were still used. The epoch model³² was used to account for the time heterogeneity of the contribution. For each clade and each predictor group, MCMC chains were run for 100 million generations with burn-in of 10%, sampling every 10,000 steps. Similarly, we assessed the convergence of the chains in Tracer v1.7⁵⁶.

Wild bird movement tracking and distribution modeling

To assemble the global wild bird observation data, we accessed the worldwide bird tracking data from Movebank in 2021. This dataset amassed from 53 studies across the world^{58–61}. The Movebank study ID, name, principal investigator, and contact person are listed in Supplementary Dataset 2. The dataset is collected by various research groups, and by various sensors, including Global Position System (GPS), Argos, bird ring, radio transmitter, solar geo-locator, and natural mark. It covers over 3542 individual birds (class: Ave), including ten orders and 95 species (Supplementary Dataset 2). For further modeling the migration of the wild birds belonging to different orders, we excluded the observation data on Movebank of Cuculiformes, Caprimulgiformes, Strigiformes, Columbiformes, Phoenicopteriformes, Pici-formes, Sphenisciformes, and Procellariiformes, given their paucity and geographically restricted distribution. Additionally, we accessed GPS tracking data of 193 individuals, including 5 orders and 12 Species between 2006 and 2019 in China from a previous study (ref. 132). Accordingly, we combined the data from China with those on Movebank (Supplementary Dataset 2) and finalized a bird observation dataset consisting of 10 orders and 96 species.

To model the wild bird distribution throughout a year, we developed a model framework based on the species distribution model (SDM). The response variable of the model is bird occurrence (1: presence; 0: pseudo-absence). The independent variables are 20 well-studied environmental predictors^{133–137} (Table S1). We divide the globe into 1-km resolution geographical cells for each month. For each cell, the value of the dependent variable is 1 if there is any observation of an individual in the target order in that month in the bird tracking data, otherwise 0. Furthermore, to infer the probability of bird occurrence between 0 and 1 for each cell, we trained a XGBoost binary classification model¹³⁸ for each bird order, respectively. The method is adapted from a previous bird migration model¹³⁹. We used true-presence and pseudo-absence data (marked as 1 and 0 respectively). We fitted the distribution of birds which manifest as true-presence data and pseudo-absence data. We randomly divided 67% of the data as the training set and the other 33% as the test set. The model finally outputs the probability of the distribution of migratory birds in each month across years. The accuracy was evaluated by the area under the curve in a test set of the ten orders: Pelecaniformes (0.97), Gruiformes (0.97), Passeriformes (0.97), Suliformes (0.98), Ciconiiformes (0.92), Falconiformes (0.98), Charadriiformes (0.94), Anseriformes (0.90), Accipitriformes (0.90). The modeled wild bird distribution was applied in the subsequent analysis to identify key bird orders associated with the global viral dispersal (see Methods) and local virus emergence.

Discrete trait phylogeography of HPAIVs and counts of virus lineage migration

Based on empirical phylogenies, we used a non-reversible discrete-state continuous-time Markov chain model and a Bayesian stochastic search variable selection (BSSVS) approach³⁰ to infer the viral diffusion among locations: i) the most probable locations of the ancestral nodes in the phylogeny and ii) the history and rates of lineage movement among locations³⁰. Sampled countries were divided into 10 locations: Africa, Central Asia, Europe, Japan-Korea, North China, South China, Qinghai, Russia, Southeast Asia and USA-Canada. This regional categorization was done according to the major wild bird breeding areas. Furthermore, to estimate the viral gene flows between locations, we used a robust counting approach^{140,141} to count virus lineage migration events. The basic idea is to count the expected number of lineage movements (Markov jumps) between the locations along the phylogeny branches, as applied in previous studies^{142–146}. The count of virus lineage migration events was used for further analysis below. We also summarized monthly frequency distribution of the virus lineage migration for each pathway (Fig. S3).

To target the key bird orders for each location, we explored the association of wild bird distribution across a year and the virus diffusion. The monthly wild bird distribution probability at each location is generated based on the location's geographical coordinates on the modeled bird distribution probability raster map. We then applied block bootstrapping to generate samples of monthly virus lineage movement counts from all events while keeping the year and the month of the event. The number of samples is decided by $\frac{20}{0.05} \times$ the number of routes \times the number of bird orders. We summarized the counts by month in each sample to get a monthly count sample. We calculated Pearson's correlation (r) for each sample between the virus lineage migration of each route and each bird order distribution at the origin/destination region, respectively. We define the two-tailed p value by doubling the proportion of the bootstrap samples with either $r > 0$ or $r < 0$ (whichever is smaller). To account for multiple comparisons of 9 bird orders, n routes, and 2 locations (origin/destination) on each route, we use p value $< \frac{0.05}{9 \times 2 \times n}$ to define the statistical significance in the correlations. When bird distribution at the origin/destination is correlated with the virus lineage movements positively or negatively, we consider the bird order distribution at the origin/destination correlated with the virus lineage movements.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

We provide Movebank Study ID (unique searchable identifier) and relevant metadata information for bird tracking data available on Movebank database (movebank.org) in Supplementary Dataset 2. We provide accession ID and relevant metadata for virus genomic data available on GISAID database (gisaid.org) in Supplementary Dataset 1. Live poultry trade data are available on the United Nations Comtrade Database (comtrade.un.org/data/). Live poultry trade data within China and the country-level poultry population size data used in phylogeographic analyses are available on the GitHub Repository¹⁴⁷. The environmental variable data used in the species distribution model are publicly available on WorldClim (www.worldclim.com/version2), NASA ARC ECOCAST GIMMS NDVI3g v1p0: Version 1.0. (iridl.ldeo.columbia.edu/SOURCES/.NASA/.ARC/.ECOCAST/.GIMMS/.NDVI3g.v1p0/index.html?Set-Language=en), Global 1-km Consensus Land Cover (www.earthenv.org/landcover) and LP DAVV and Data Products (lpdaac.usgs.gov/about/citing_lp_daac_and_data).

Code availability

All code and scripts are provided on GitHub Repository¹⁴⁷.

References

- Wille, M. & Barr, I. G. Resurgence of avian influenza virus. **376**, 459–460. <https://www.science.org/doi/10.1126/science.abo1232>.
- 2020: OIE - World Organisation for Animal Health. <https://www.oie.int/animal-health-in-the-world/update-on-avian-influenza/2020/>.
- Beerens, N. et al. Multiple reassorted viruses as cause of highly pathogenic avian influenza A(H5N8) virus epidemic, the Netherlands, 2016. *Emer. Infect. Dis.* **23**, 1974–1981 (2017).
- Kwon, J.-H. et al. New reassortant clade 2.3.4.4b avian influenza A(H5N6) virus in wild birds, South Korea, 2017–18. *Emerg. Infect. Dis.* **24**, 1953–1955 (2018).
- Mine, J. et al. Genetics and pathogenicity of H5N6 highly pathogenic avian influenza viruses isolated from wild birds and a chicken in Japan during winter 2017–2018. *Virology* **533**, 1–11 (2019).
- Lee, D.-h., Bertran, K., Kwon, J.-h. & Swayne, D. E. Evolution, global spread, and pathogenicity of highly pathogenic avian influenza H5Nx clade 2.3.4.4. *J. Vet. Sci.* **18**, 269–280 (2017).
- Writing Committee of the Second World Health Organization Consultation on Clinical Aspects of Human Infection with Avian Influenza A (H5N1) Virus; Abdel-Ghafar, A. N., Chotpitayasanunondh, T., Gao, Z., Hayden, F. G., Nguyen, D. H., de Jong, M. D., Naghdaliyev, A., Peiris, J. S., Shindo, N., Soeroso, S., Uyeki T. M. Update on Avian Influenza A (H5N1) Virus Infection in Humans. *N. Engl. J. Med.* **358**, 261–273 (2008).
- Gao, R. et al. Human infection with a novel avian-origin influenza A (H7N9) virus. *N. Engl. J. Med.* **368**, 1888–1897 (2013).
- Yang, Z.-F., Mok, C. K., Peiris, J. S. & Zhong, N.-S. Human infection with a novel avian influenza A(H5N6) virus. *N. Engl. J. Med.* **373**, 487–489 (2015).
- European Centre for Disease Prevention and Control. First identification of human cases of avian influenza A (H5N8) infection. 24 February 2021. (ECDC, Stockholm, 2021).
- Peiris, J. S. M., de Jong, M. D. & Guan, Y. Avian influenza virus (H5N1): a threat to human health. *Clin. Microbiol. Rev.* **20**, 243–267 (2007).
- WHO, FAO & WOAH. Human infection with influenza A(H5N1), Cambodia (2023). <https://www.woah.org/app/uploads/2023/04/prelim-fao-who-woah-joint-rra-h5n1-khm-2023.pdf>.
- Trovão, N. S., Suchard, M. A., Baele, G., Gilbert, M. & Lemey, P. Bayesian inference reveals host-specific contributions to the epidemic expansion of influenza A H5N1. *Mol. Biol. Evol.* **32**, 3264–3275 (2015).
- Russell, C. A. Sick birds don't fly...or do they? *Science* **354**, 174–175 (2016).
- Liu, J. et al. Highly pathogenic H5N1 influenza virus infection in migratory birds. *Science* **309**, 1206–1206 (2005).
- Tian, H. Avian influenza H5N1 viral and bird migration networks in Asia. *Proc Natl Acad Sci USA* **112**, 172–177 (2015).
- Zhou, S. et al. Genetic evidence for avian influenza H5N1 viral transmission along the Black Sea-Mediterranean Flyway. *J. Gen. Virol.* **97**, 2129–2134 (2016).
- Kim, Y.-I. et al. Pathobiological features of a novel, highly pathogenic avian influenza A(H5N8) virus. *Emerg. Microbes. Infect.* **3**, 1–13 (2014).
- Kang, H.-M. et al. Novel reassortant influenza A(H5N8) viruses among inoculated domestic and wild ducks, South Korea, 2014. *Emerg. Infect. Dis.* **21**, 298–304 (2015).
- Sun, H. et al. Characterization of clade 2.3.4.4 highly pathogenic H5 avian influenza viruses in ducks and chickens. *Vet. Microbiol.* **182**, 116–122 (2016). Publisher: Elsevier B.V.
- Lycett, S. J. et al. Role for migratory wild birds in the global spread of avian influenza H5N8. *Science* **354**, 213–217 (2016).
- Blaurock, C. et al. Preferential selection and contribution of non-structural protein 1 (NS1) to the efficient transmission of panzootic

- avian influenza h5n8 virus clades 2.3.4.4a and b in chickens and ducks. **95**, e00445–21. <https://journals.asm.org/doi/10.1128/JVI.00445-21>.
23. Leyson, C. M., Youk, S., Ferreira, H. L., Suarez, D. L. & Pantin-Jackwood, M. Multiple gene segments are associated with enhanced virulence of clade 2.3.4.4 h5n8 highly pathogenic avian influenza virus in mallards. **95**, e00955–21. <https://journals.asm.org/doi/10.1128/JVI.00955-21>.
24. Lycett, S. J. et al. Genesis and spread of multiple reassortants during the 2016/2017 H5 avian influenza epidemic in Eurasia. *Proc. Natl Acad. Sci. USA* **117**, 20814–20825 (2020).
25. Hill, S. C. et al. Comparative micro-epidemiology of pathogenic avian influenza virus outbreaks in a wild bird population. **374**, 20180259. <https://royalsocietypublishing.org/doi/10.1098/rstb.2018.0259>.
26. Reeves, A. B. et al. Influenza a virus recovery, diversity, and intercontinental exchange: a multi-year assessment of wild bird sampling at izembek national wildlife refuge, Alaska. **13**, e0195327. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0195327>.
27. van Dijk, J. G. B. et al. Juveniles and migrants as drivers for seasonal epizootics of avian influenza virus. **83**, 266–275. <https://onlinelibrary.wiley.com/doi/abs/10.1111/1365-2656.12131>.
28. Hill, N. J. et al. Ecological divergence of wild birds drives avian influenza spillover and global spread. **18**, e1010062. <https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1010062>.
29. Hicks, J. T. et al. Host diversity and behavior determine patterns of interspecies transmission and geographic diffusion of avian influenza a subtype among North American wild reservoir species. **18**, e1009973. <https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1009973>.
30. Lemey, P., Rambaut, A., Drummond, A. J. & Suchard, M. A. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* **5**, e1000520 (2009).
31. Lemey, P. et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog.* **10**, 1–17 (2014).
32. Bielejec, F., Lemey, P., Baele, G., Rambaut, A. & Suchard, M. Inferring heterogeneous evolutionary processes through time: from sequence substitution to phylogeography. *Syst. Biol.* **63**, 493–504 (2014).
33. European Food Safety Authority. European Centre for Disease Prevention, Control, European Union Reference Laboratory for Avian Influenza et al. Avian influenza overview December 2021–March 2022. **20**. <https://data.europa.eu/doi/10.2903/j.efsa.2022.7289>.
34. Ramey, A. M. et al. Highly pathogenic avian influenza is an emerging disease threat to wild birds in North America. **86**, e22171. <https://onlinelibrary.wiley.com/doi/abs/10.1002/jwmg.22171>.
35. Lisovski, S. et al. Light-level geolocator analyses: a user's guide. **89**, 221–236. <https://onlinelibrary.wiley.com/doi/abs/10.1111/1365-2656.13036>.
36. Gaidet, N. et al. Understanding the ecological drivers of avian influenza virus infection in wildfowl: a continental-scale study across Africa. **279**, 1131–1141. <https://royalsocietypublishing.org/doi/full/10.1098/rspb.2011.1417>.
37. Bernard, A., Rodrigues, A. S., Cazalís, V. & Grémillet, D. Toward a global strategy for seabird tracking. **14**, e12804. <https://onlinelibrary.wiley.com/doi/abs/10.1111/conl.12804>.
38. van Toor, M. L., Avril, A., Wu, G., Holan, S. H. & Waldenström, J. As the duck flies—estimating the dispersal of low-pathogenic avian influenza viruses by migrating Mallards. *Front. Ecol. Evol.* **6** (2018). <https://www.frontiersin.org/article/10.3389/fevo.2018.00208>.
39. van Toor, M. L. et al. Integrating animal movement with habitat suitability for estimating dynamic migratory connectivity. **33**, 879–893. <https://doi.org/10.1007/s10980-018-0637-9>.
40. Wille, M. et al. Australia as a global sink for the genetic diversity of avian influenza A virus. **18**, e1010150. <https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1010150>.
41. Müller, N. F., Rasmussen, D. & Stadler, T. MASCOT: parameter and state inference under the marginal structured coalescent approximation. *Bioinformatics* (2018). <https://academic.oup.com/bioinformatics/advance-article/doi/10.1093/bioinformatics/bty406/5001387>.
42. Skehel, J. J. & Wiley, D. C. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. **69**, 531–69. <https://www.proquest.com/docview/223690008/abstract/D796D35D31C8400FPQ/1>.
43. Boulinier, T. et al. Migration, Prospecting, Dispersal? What host movement matters for infectious agent circulation? **56**, 330–342. <https://doi.org/10.1093/icb/icw015>.
44. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data—from vision to reality. *Eurosurveillance* **22**, 30494 (2017).
45. Khare, S. et al. GISAID's role in pandemic response. *China CDC Wkly.* **3**, 1049–1051 (2021).
46. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Chall.* **1**, 33–46 (2017).
47. Price, M., Dehal, P. & Arkin, A. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650 (2009).
48. Lam, T.-Y. The genesis and source of the H7N9 influenza viruses causing human infections in China. *Nature* **502**, 241 (2013).
49. Smith, G. J. & Donis, R. O. Nomenclature updates resulting from the evolution of avian influenza A(H5) virus clades 2.1.3.2a, 2.2.1, and 2.3.4 during 2013–2014. *Influenza Other Respir. Viruses* **9**, 271–276 (2015).
50. Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016).
51. Suchard, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018).
52. Ayres, D. L. et al. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Syst. Biol.* **61**, 170–173 (2012).
53. Drummond, A., Ho, S., Phillips, M. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, 88 (2006).
54. Shapiro, B., Rambaut, A. & Drummond, A. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* **23**, 7–9 (2005).
55. Minin, V., Bloomquist, E. & Suchard, M. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* **25**, 1459–1471 (2008).
56. Rambaut, A., Drummond, A., Xie, D., Baele, G. & Suchard, M. Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).
57. Yang, Q. et al. Assessing the role of live poultry trade in community-structured transmission of avian influenza in China. *Proc. Natl Acad. Sci.* **117**, 5949–5954 (2020).
58. Warwick-Evans, V. et al. Data from: Changes in behaviour drive inter-annual variability in the at-sea distribution of northern gannets. *Movebank data repository*. <https://doi.org/10.5441/001/1.30r98t3> (2017).
59. Kölzsch, A., Kruckenberg, H., Glazov, P., Müskens, G. & Wikelski, M. Data from: Towards a new understanding of migration timing: slower spring than autumn migration in geese reflects different

- decision rules for stopover use and departure. *Movebank data repository*. <https://doi.org/10.5441/001/1.31c2v92f> (2016).
60. van Toor, M., Ottosson, U., van der M., T., van Hoorn, S. & Waldenström, J. Data from: As the duck flies: estimating the dispersal of low-pathogenic avian influenza viruses by migrating mallards. *Movebank data repository*. <https://doi.org/10.5441/001/1.3fv21n7m> (2018).
61. Cueto, V. & Bravo, S. Data from: Migratory timing, rate, routes and wintering areas of white-crested elaenia (*Elaenia albiceps chilensis*), a key seed disperser for Patagonian forest regeneration. *Movebank data repository*. <https://doi.org/10.5441/001/1.406327g0> (2017).
62. Ross, J., Bridge, E., Rozmarynowycz, M. & Bingman, V. Data from: Individual variation in migratory path and behavior among Eastern Lark Sparrows. *Movebank data repository*. <https://doi.org/10.5441/001/1.5jd56s8h> (2014).
63. Fleishman, E. et al. Data from: Space use by Swainson's hawk (*Buteo swainsoni*) in the Natomas Basin, California. *Movebank data repository*. <https://doi.org/10.5441/001/1.6p427t2f/1> (2017).
64. Flack, A. et al. Data from: Costs of migratory decisions: a comparison across eight white stork populations. *Movebank data repository*. <https://doi.org/10.5441/001/1.78152p3q> (2015).
65. Lamb, J., Satgé, Y. & Jodice, P. Data from: Influence of density-dependent competition on foraging and migratory behavior of a subtropical colonial seabird. *Movebank data repository*. <https://doi.org/10.5441/001/1.7856r086> (2017).
66. Matthes, D. et al. Data from: Flexibility of continental navigation and migration in European mallards. *Movebank data repository*. <https://doi.org/10.5441/001/1.8dc0v84m> (2013).
67. Friedemann, G., Leshem, Y. & Izhaki, I. Data from: Multi-dimensional differentiation in foraging resource use during breeding of two sympatric raptors: space, habitat type, time and food. *Movebank data repository*. <https://doi.org/10.5441/001/1.c6q8353q> (2016).
68. Hallworth, M. & Marra, P. Data from: Miniaturized GPS tags identify non-breeding territories of a small breeding migratory songbird. *Movebank data repository*. <https://doi.org/10.5441/001/1.h43t9j6b> (2015).
69. Petersen, M. & Douglas, D. Data from: At-sea distribution of spectacled eiders: a 120-year-old mystery resolved. *Movebank data repository*. <https://doi.org/10.5441/001/1.kq7t609j> (2016).
70. Spiegel, O., Getz, W. & Nathan, R. Data from: Factors influencing foraging search efficiency: Why do scarce lappet-faced vultures outperform ubiquitous white-backed vultures? (V2). *Movebank data repository*. <https://doi.org/10.5441/001/1.mf903197> (2014).
71. Arlt, D., Olsson, P., Fox, J., Low, M. & Pärt, T. Data from: Prolonged stopover duration characterises migration strategy and constraints of a long-distant migrant songbird. *Movebank data repository*. <https://doi.org/10.5441/001/1.nn55rh75> (2015).
72. Kleyheeg, E., van Dijk, J., Nolet, B. & Soons, M. Data from: Movement patterns of a keystone waterbird species are highly predictable from landscape configuration. *Movebank data repository*. <https://doi.org/10.5441/001/1.p9c4nc97> (2017).
73. van der Jeugd, H., Oosterbeek, K., Ens, B. J., Shamoun-Baranes, J. & Exo, K. Data from: Forecasting spring from afar? Timing of migration and predictability of phenology along different migration routes of an avian herbivore [Barents Sea data]. *Movebank data repository*. <https://doi.org/10.5441/001/1.ps244r11/2> (2014).
74. Chudzińska, M. & Madsen, J. Data from: Foraging behaviour and fuel accumulation of capital breeders during spring migration as derived from a combination of satellite- and ground-based observations. *Movebank data repository*. <https://doi.org/10.5441/001/1.s9s86319> (2016).
75. Hernández-Priego, J., Rodríguez, C. & Bustamante, J. Data from: Why do kestrels soar? *Movebank data repository*. <https://doi.org/10.5441/001/1.sj8t3r11> (2015).
76. Arlt, D., Olsson, P., Fox, J. W., Low, M. & Pärt, T. Prolonged stopover duration characterises migration strategy and constraints of a long-distance migrant songbird. *Anim. Migr.* **2**, 47–62 (2015).
77. Bengtsson, D. et al. Does influenza A virus infection affect movement behaviour during stopover in its wild reservoir host? *R. Soc. Open Sci.* **3**, 150633 (2016).
78. Bengtsson, D. et al. Movements, home-range size and habitat selection of mallards during autumn migration. *PLOS ONE* **9**, e100764 (2014).
79. Boyd, W. S., Ward, D. H., Kraege, D. K. & Gerick, A. A. Migration patterns of Western High Arctic (Grey-belly) Brant *Branta bernicla*. *Wildfowl* **3**, 3–25 (2013).
80. Bravo, S. P., Cueto, V. R. & Gorosito, C. A. Migratory timing, rate, routes and wintering areas of White-crested Elaenia (*Elaenia albiceps chilensis*), a key seed disperser for Patagonian forest regeneration. *PLOS ONE* **12**, e0170188 (2017).
81. Chudzińska, M. E., Nabe-Nielsen, J., Nolet, B. A. & Madsen, J. Foraging behaviour and fuel accumulation of capital breeders during spring migration as derived from a combination of satellite- and ground-based observations. *J. Avian Biol.* **47**, 563–574 (2016).
82. Cochran, W. & Wikelski, M. *Birds of Two Worlds: The Ecology and Evolution of Migration* (The Johns Hopkins University Press, Baltimore, 2005).
83. Cochran, W. W., Bowlin, M. S. & Wikelski, M. Wingbeat frequency and flap-pause ratio during natural migratory flight in thrushes. *Integr. Comp. Biol.* **48**, 134–151 (2008).
84. DeLuca, W. V. et al. Transoceanic migration by a 12 g songbird. *Biol. Lett.* **11**, 20141045 (2015).
85. Efrat, R., Harel, R., Alexandrou, O., Catsadorakis, G. & Nathan, R. Seasonal differences in energy expenditure, flight characteristics and spatial utilization of Dalmatian Pelicans *Pelecanus crispus* in Greece. *Ibis* **161**, 415–427 (2019).
86. Eichhorn, G. *Northward Bound: Fat for Flight: How Barnacle Geese Budget Their Spring Travels*, 84–90 (KNNV Pub., Utrecht, 2005).
87. Eichhorn, G. *Travels in a changing world flexibility and constraints in migration and breeding of the barnacle goose*, (The University of Groningen, 2008).
88. Flack, A. et al. Costs of migratory decisions: a comparison across eight white stork populations. *Sci. Adv.* **2**, e1500931 (2016).
89. Flack, A., Nagy, M., Fiedler, W., Couzin, I. D. & Wikelski, M. From local collective behavior to global migratory patterns in white storks. *Science* **360**, 911–914 (2018).
90. Fleishman, E. et al. Space Use by Swainson's Hawk (*Buteo swainsoni*) in the Natomas Basin, California. *Collabro* **2**, 5 (2016).
91. Friedemann, G. et al. Multidimensional differentiation in foraging resource use during breeding of two sympatric top predators. *Scientific Reports* **6**, 35031 (2016).
92. Fuller, M. R., Seagar, W. S. & Schueck, L. S. Routes and Travel Rates of Migrating Peregrine Falcons *Falco peregrinus* and Swainson's Hawks *Buteo swainsoni* in the Western Hemisphere. *J. Avian Biol.* **29**, 433–440 (1998).
93. Hallworth, M. T. & Marra, P. P. Miniaturized GPS tags identify non-breeding territories of a small breeding migratory songbird. *Sci. Rep.* **5**, 11069 (2015).
94. Harrison, A.-L., Woodard, P., Mallory, M. & Rausch, J. Sympatrically-breeding congeneric seabirds (*Stercorarius* spp.) from Arctic Canada migrate to four oceans. 181064 bytes Dryad. <https://doi.org/10.5061/DRYAD.NK98SF7V1> (2021).
95. Harrison, A.-L., Woodard, P. F., Mallory, M. L. & Rausch, J. Sympatrically breeding congeneric seabirds (*Stercorarius* spp.) from Arctic Canada migrate to four oceans. *Ecol. Evol.* **12**, e8451 (2022).
96. Hernández-Priego, J., Rodríguez, C. & Bustamante, J. Why do kestrels soar? *PLOS ONE* **10**, e0145402 (2015).

97. Klein, K. et al. Fly with the flock: immersive solutions for animal movement visualization and analytics. *J. R. Soc. Interface* **16**, 20180794 (2019).
98. Kleyheeg, E. et al. Movement patterns of a keystone waterbird species are highly predictable from landscape configuration. *Mov. Ecol.* **5**, 2 (2017).
99. Kochert, M. N. et al. Migration patterns, use of stopover areas, and austral summer movements of Swainson's Hawks. *Condor* **113**, 89–106 (2011).
100. Kölzsch, A. et al. Forecasting spring from afar? Timing of migration and predictability of phenology along different migration routes of an avian herbivore. *J. Anim. Ecol.* **84**, 272–283 (2015).
101. Kölzsch, A. et al. Towards a new understanding of migration timing: slower spring than autumn migration in geese reflects different decision rules for stopover use and departure. *Oikos* **125**, 1496–1507 (2016).
102. Köppen, U., Yakovlev, A. P., Barth, R., Kaatz, M. & Berthold, P. Seasonal migrations of four individual bar-headed geese *Anser indicus* from Kyrgyzstan followed by satellite telemetry. *J. Ornithol.* **151**, 703–712 (2010).
103. Korner, P., Sauter, A., Fiedler, W. & Jenni, L. Variable allocation of activity to daylight and night in the mallard. *Anim. Behav.* **115**, 69–79 (2016).
104. Lamb, J. S., Satgé, Y. G. & Jodice, P. G. R. Influence of density-dependent competition on foraging and migratory behavior of a subtropical colonial seabird. *Ecol. Evol.* **7**, 6469–6481 (2017).
105. Lyons, D. E., Patterson, A. G. L., Tennyson, J., Lawes, T. J. & Roby, D. D. The Salton sea: critical migratory stopover habitat for Caspian terns (*Hydroprogne caspia*) in the North American Pacific Flyway. *Waterbirds* **41**, 154–165 (2018).
106. Nagy, M., Couzin, I. D., Fiedler, W., Wikelski, M. & Flack, A. Synchronization, coordination and collective sensing during thermalling flight of freely migrating white storks. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, 20170011 (2018).
107. Petersen, M. R., Earned, W. W. & Douglas, D. C. At-sea distribution of spectacled eiders: a 120-year-old mystery resolved. *Auk* **116**, 1009–1020 (1999).
108. Petersen, M. R., Flint, P. L., Mulcahy, D. M. & Douglas, D. C. Tracking Data for Common Eiders (*Somateria mollissima*). U.S. Geological Survey. <https://doi.org/10.5066/P9JH3EXV> (2021).
109. Petersen, M. R., Douglas, D. C. & Mulcahy, D. M. Use of implanted satellite transmitters to locate Spectacled Eiders at-Sea. *Condor* **97**, 276–278 (1995).
110. Petersen, M. R. & Douglas, D. C. Winter ecology of spectacled eiders: environmental characteristics and population change. *Condor* **106**, 79–94 (2004).
111. Poli, C. L., Harrison, A.-L., Vallarino, A., Gerard, P. D. & Jodice, P. G. R. Dynamic oceanography determines fine scale foraging behavior of Masked Boobies in the Gulf of Mexico. *PLOS ONE* **12**, e0178318 (2017).
112. Ross, J. D., Bridge, E. S., Rozmarynowycz, M. J. & Bingman, V. P. Individual variation in migratory path and behavior among Eastern Lark Sparrows. *Anim. Migr.* **2**, 29–33 (2015).
113. Senner, N. R. et al. When Siberia came to the Netherlands: the response of continental black-tailed godwits to a rare spring weather event. *J. Anim. Ecol.* **84**, 1164–1176 (2015).
114. Sexson, M. G., Petersen, M. R., Breed, G. A. & Powell, A. N. Shifts in the distribution of molting Spectacled Eiders (*Somateria fischeri*) indicate ecosystem change in the Arctic. *Condor* **118**, 463–476 (2016).
115. Shariati-Najafabadi, M. et al. Environmental parameters linked to the last migratory stage of barnacle geese en route to their breeding sites. *Anim. Behav.* **118**, 81–95 (2016).
116. Shariatinajafabadi, M. et al. Migratory herbivorous waterfowl track satellite-derived green wave index. *PLOS ONE* **9**, e108331 (2014).
117. Shariati Najafabadi, M. et al. Satellite- versus temperature-derived green wave indices for predicting the timing of spring migration of avian herbivores. *Ecol. Indic.* **58**, 322–331 (2015).
118. Silveira, N. S. D., Niebuhr, B. B. S., Muylaert, R. L., Ribeiro, M. C. & Pizo, M. A. Effects of land cover on the movement of frugivorous birds in a heterogeneous landscape. *PLOS ONE* **11**, e0156688 (2016).
119. Soanes, L. M., Atkinson, P. W., Gauvain, R. D. & Green, J. A. Individual consistency in the foraging behaviour of Northern Gannets: Implications for interactions with offshore renewable energy developments. *Mar. Policy* **38**, 507–514 (2013).
120. Spiegel, O., Getz, W. M. & Nathan, R. Factors influencing foraging search efficiency: why do scarce lappet-faced vultures outperform ubiquitous white-backed vultures? *Am. Nat.* **181**, E102–E115 (2013).
121. van Toor, M. L. et al. Flexibility of continental navigation and migration in European mallards. *PLoS One* **8**, e72629 (2013).
122. Torres, L. G., Orben, R. A., Tolkova, I. & Thompson, D. R. Classification of animal movement behavior through residence in space and time. *PLoS One* **12**, e0168513 (2017).
123. Tucker, M. A. et al. Large birds travel farther in homogeneous environments. *Glob. Ecol. Biogeogr.* **28**, 576–587 (2019).
124. van Toor, M. L., Avril, A., Wu, G., Holan, S. H. & Waldenström, J. As the duck flies—estimating the dispersal of low-pathogenic avian influenza viruses by migrating mallards. *Front. Ecol. Evol.* **6** (2018).
125. Warwick-Evans, V. et al. Changes in behaviour drive inter-annual variability in the at-sea distribution of northern gannets. *Mar. Biol.* **163**, 156 (2016).
126. Warwick-Evans, V., Atkinson, P. W., Walkington, I. & Green, J. A. Predicting the impacts of wind farms on seabirds: an individual-based model. *J. Appl. Ecol.* **55**, 503–515 (2018).
127. Weinzierl, R. et al. Wind estimation based on thermal soaring of birds. *Ecol. Evol.* **6**, 8706–8718 (2016).
128. Wikelski, M. et al. True navigation in migrating gulls requires intact olfactory nerves. *Sci. Rep.* **5**, 17061 (2015).
129. Ens, B. et al. SOVON-onderzoeksrapport 2008/10: Tracking of individual birds. Report on WP 3230 (bird tracking sensor characterization) and WP 4130 (sensor adaptation and calibration for bird tracking system) of the FlySafe basic activities project. (SOVON Vogelonderzoek Nederland, Netherlands, 2008). Reporting year: 2008.
130. Georgopoulou, E., Alexandrou, O., Manolopoulos, A., Xirouchakis, S. & Catsadorakis, G. Home range of the Dalmatian pelican in south-east Europe. *Eur. J. Wildl. Res.* **69**, 41 (2023).
131. Scott, T. *Movements of White-Headed and White-Backed Vultures*, (Boise State University, 2020).
132. Zhang, G. et al. Bidirectional Movement of Emerging H5N8 Avian Influenza Viruses Between Europe and Asia via Migratory Birds Since Early 2020. *Mol. Biol. Evol.* **40**, msad019 (2023).
133. Pinzon, J. E. & Tucker, C. J. A non-stationary 1981–2012 AVHRR NDVI3g time series. *Remote Sens.* **6**, 6929–6960 (2014).
134. Tuanmu, M.-N. & Jetz, W. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Glob. Ecol. Biogeogr.* **23**, 1031–1045 (2014).
135. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *37*, 4302–4315. <https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.5086>.
136. Rosen, P. A. et al. Synthetic aperture radar interferometry. *Proceedings of the IEEE* **88**, 333–382 (2000).
137. Kobrick, M. On the Toes of Giants - How SRTM was Born. *Photogramm. Eng. Remote Sens.* 206–210 (2006).
138. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, 785–794 (Association for Computing Machinery, New York, NY, USA, 2016). <https://doi.org/10.1145/2939672.2939785>.

139. Van Doren, B. M. & Horton, K. G. A continental system for forecasting bird migration. *Science* **361**, 1115–1118 (2018).
140. O'Brien, J. D., Minin, V. N. & Suchard, M. A. Learning to count: robust estimates for labeled distances between molecular sequences. *Mol. Biol. Evol.* **26**, 801–814 (2009).
141. Minin, V. N. & Suchard, M. A. Fast, accurate and simulation-free stochastic mapping. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 3985–3995 (2008).
142. Faria, N. R., Suchard, M. A., Rambaut, A., Streicker, D. G. & Lemey, P. Simultaneously reconstructing viral crossspecies transmission history and identifying the underlying constraints. *Philos. Trans. R. Soc. B Biol. Sci.* **368** (2013).
143. Faria, N. R. et al. Distinct rates and patterns of spread of the major HIV-1 subtypes in Central and East Africa. *PLoS Pathog.* **15**, e1007976 (2019).
144. Faria, N. R. et al. The early spread and epidemic ignition of HIV-1 in human populations. *Science* **346**, 56–61 (2014).
145. Vaslyeva, T. I. et al. Molecular epidemiology reveals the role of war in the spread of HIV in Ukraine. *Proc. Natl Acad. Sci. USA* **115**, 1051–1056 (2018).
146. Thézé, J. et al. Genomic epidemiology reconstructs the introduction and spread of Zika Virus in Central America and Mexico. *Cell Host Microbe* **23**, 855–864.e7 (2018).
147. Yang, Q., Wang, B. & Lemey, P. HPAI_Bird_world v1.0.0 <https://doi.org/10.5281/zenodo.10480799>.

Acknowledgements

We gratefully acknowledge all data contributors of virus genome sequences, i.e., the authors and their originating laboratories responsible for obtaining the specimens, and their submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which this research is based. We also gratefully acknowledge all bird tracking data contributors, i.e., the authors and other researchers in their originating groups collecting the data and metadata and sharing via Movebank, on which this research is based. The principle investigator, contact person, citation, or data repository DOI are listed in Supplementary Dataset 2. We thank Qizhong Wu and Joint Center for Earth System Modeling and High Performance Computing, Beijing Normal University, for providing computing resources for time-scaled phylogenetic analyses. We thank Jing Yu for assisting in assembling international poultry trade data. We thank Bram Vrancken for assisting with phylogenetic analyses. We thank insightful comments from Andrew Rambaut and the Virus Club Group at the University of Edinburgh, Grenfell lab at Princeton University, the Influenza group at Francis Crick Institute, Olga Alexandrou at Society for the Protection of Prespa, Yonghong Liu at Chinese Center for Disease Control and Prevention, Yidan Li (formerly at Beijing Normal University), and Nils Stenseth at the University of Oslo. Q.Y. acknowledges funding from the Center for Health and Wellbeing, Princeton University, via Graduate Funding for Health-Focused Research. P.L. acknowledges support from the European Union's Horizon 2020 research and innovation program (grant agreement no. 725422-ReservoirDOCS), from the Wellcome Trust through project 206298/Z/17/Z and from the European Union's Horizon 2020 project MOOD (grant agreement no. 874850). S.W.P. was supported by a Charlotte Elizabeth Procter Fellowship of Princeton University. J.L.H.T. is supported by a Yeotown Scholarship from New College, University of Oxford. Y.H. is supported by a Masason Foundation Fellowship, a Honjo International Fellowship, and a Centennial

Fellowship. O.G.P. was supported by the UKRI GCRF One Health Poultry Hub (grant no. B/S011269/1). G.Z. is supported by National Natural Science Foundation of China (32070530). H.T. is supported by National Key Research and Development Program of China (2022YFC2303803), National Natural Science Foundation of China (82073616), Fundamental Research Funds for the Central Universities (2233300001), and BNU-FGS Global Environmental Change Program (No.2023-GC-ZYTS-11). B.T.G. is supported by Princeton Catalyst Initiative, Princeton University, and High Meadows Environmental Institute, Princeton University. The opinions expressed in this article are those of the authors and do not reflect the view of the National Institutes of Health, the Department of Health and Human Services, or the United States government.

Author contributions

Q.Y., B.T.G., H.T., O.G.P., and G.Z. contributed to the design of the work. Q.Y., G.Z., H.T., L.D., Y.C., and B.L. contributed to the data acquisition. Q.Y., B.W., and P.L. contributed to the data analysis. Q.Y., P.L., L.D., T.M., R.A.W., F.G., S.W.P., N.L., J.T., S.B., Y.H., G.Z., O.G.P., H.T., and B.T.G. contributed to the data interpretation. Q.Y. and B.T.G. drafted the initial manuscript draft. All authors contributed to the critical revision of the manuscript and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-45462-1>.

Correspondence and requests for materials should be addressed to Qiqi Yang, Huaiyu Tian or Bryan Grenfell.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

¹Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. ²State Key Laboratory of Remote Sensing Science, Center for Global Change and Public Health, Faculty of Geographical Science, Beijing Normal University, Beijing, China. ³Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, Belgium. ⁴College of Life Sciences, Beijing Normal University, Beijing, China. ⁵Princeton School of Public

and International Affairs, Princeton University, Princeton, NJ, USA. ⁶Fogarty International Center, U.S. National Institutes of Health, Bethesda, MD, USA. ⁷Animal and Plant Health Agency-Weybridge, OIE/FAO International Reference Laboratory for Avian Influenza, Swine Influenza and Newcastle Disease Virus, Department of Virology, Addlestone, UK. ⁸Department of Pathobiology and Population Science, Royal Veterinary College, London, UK. ⁹Department of Biology, University of Oxford, Oxford, UK. ¹⁰State Key Laboratory of Biocontrol, School of Ecology, Shenzhen Campus of Sun Yat-sen University, Shenzhen, Guangdong, China. ¹¹Institute for Disease Modeling, Bill and Melinda Gates Foundation, Seattle, WA, USA. ¹²Key Laboratory of Forest Protection of National Forestry and Grassland Administration, Ecology and Nature Conservation Institute, Chinese Academy of Forestry, National Bird Banding Center of China, Beijing, China. ¹³These authors jointly supervised this work: Guogang Zhang, Oliver G. Pybus, Huaiyu Tian, Bryan Grenfell. ✉e-mail: qiqiy@princeton.edu; tianhuaiyu@gmail.com; grenfell@princeton.edu