Article

The episodic resurgence of highly pathogenic avian influenza H5 virus

https://doi.org/10.1038/s41586-023-06631-2

Received: 2 January 2023

Accepted: 11 September 2023

Published online: 18 October 2023

Check for updates

Ruopeng Xie[1,2], Kimberly M. Edwards[1,2], Michelle Wille[3,4], Xiaoman Wei[1,2], Sook-San Wong[1,2], Mark Zanin[1,5], Rabeh El-Shesheny[6], Mariette Ducatez[7], Leo L. M. Poon[1,2,5], Ghazi Kayali[8], Richard J. Webby[9] & Vijaykrishna Dhanasekaran[1,2]✉

Highly pathogenic avian influenza (HPAI) H5N1 activity has intensified globally since 2021, increasingly causing mass mortality in wild birds and poultry and incidental infections in mammals[1–3]. However, the ecological and virological properties that underscore future mitigation strategies still remain unclear. Using epidemiological, spatial and genomic approaches, we demonstrate changes in the origins of resurgent HPAI H5 and reveal significant shifts in virus ecology and evolution. Outbreak data show key resurgent events in 2016–2017 and 2020–2021, contributing to the emergence and panzootic spread of H5N1 in 2021–2022. Genomic analysis reveals that the 2016–2017 epizootics originated in Asia, where HPAI H5 reservoirs are endemic. In 2020–2021, 2.3.4.4b H5N8 viruses emerged in African poultry, featuring mutations altering *HA* structure and receptor binding. In 2021–2022, a new H5N1 virus evolved through reassortment in wild birds in Europe, undergoing further reassortment with low-pathogenic avian influenza in wild and domestic birds during global dissemination. These results highlight a shift in the HPAI H5 epicentre beyond Asia and indicate that increasing persistence of HPAI H5 in wild birds is facilitating geographic and host range expansion, accelerating dispersion velocity and increasing reassortment potential. As earlier outbreaks of H5N1 and H5N8 were caused by more stable genomic constellations, these recent changes reflect adaptation across the domestic-bird–wild-bird interface. Elimination strategies in domestic birds therefore remain a high priority to limit future epizootics.

Influenza A viruses (genus *Alphainfluenzavirus*, family *Orthomyxoviridae*) are negative-sense, single-stranded, segmented ribonucleic acid viruses categorized into subtypes on the basis of the antigenicity of their two surface proteins. Sixteen out of 18 known hemagglutinin (*HA*) subtypes and 9 out of 11 neuraminidase (NA) subtypes are prevalent as low-pathogenic avian influenza (LPAI) viruses in wild aquatic birds worldwide.

Highly pathogenic avian influenza (HPAI) viruses evolve from LPAI viruses in poultry by acquiring *HA* cleavage site insertions that facilitate systemic infection[4]. Only H5 and H7 subtypes have evolved into HPAI, with most outbreaks contained through culling or die-offs, yielding limited opportunities for spillover. Although an HPAI H5N3 epizootic was reported in wild terns in South Africa as early as 1961[5], the HPAI H5N1 virus that emerged in China in 1996 (goose/Guangdong (gs/Gd) lineage) was the first to establish sustained transmission in domestic poultry[6]. Early evolution of the gs/Gd lineage was characterized by the diversification of the H5 *HA* gene into as many as ten main phylogenetic clades[7], which, through extensive reassortment with LPAI viruses, acquired new combinations of internal genes. Ultimately, clade 2 proved most successful, and repeat spillover to wild aquatic birds since 2005[8] has enabled episodic dissemination across Asia, Europe, Africa and, most recently, North and South America[6,9,10].

The scale of HPAI H5 outbreaks in wild birds has escalated beyond Asia since 2014[11], driven by the emergence of H5 *HA* clade 2.3.4.4 viruses with several *NA* subtypes including H5N2, H5N6 and H5N8 (collectively H5Nx) (Fig. 1)[12]. Before clade 2.3.4.4, HPAI H5 evolution was characterized by a relatively stable internal gene cassette that mainly derived from LPAI viruses in domestic birds (for example, H9N2), with the linkage of H5 *HA* and N1 *NA* surface proteins, despite *HA* diversification (graphical summary in Fig. 6). Experimental animal infections showed that clade 2.3.4.4 viruses were associated with increased adaptation and reduced virulence in wild ducks compared to earlier H5N1 lineages[13]. From 2016, outbreaks in wild birds were repeatedly caused by clade 2.3.4.4b H5N8 viruses that originated in China[11], eliciting a relative increase in virulence for ducks[14]. Most recently, a reassortant HPAI H5N1 virus, which evolved from clade 2.3.4.4b viruses, has almost entirely replaced the formerly dominant clade 2.3.4.4b H5N8 viruses (Fig. 1 and Extended Data Fig. 1)[15]. Since November 2021, this H5N1 virus has

[1]School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China. [2]HKU-Pasteur Research Pole, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China. [3]Sydney Institute for Infectious Diseases, School of Medical Sciences, The University of Sydney, Sydney, New South Wales, Australia. [4]Department of Microbiology and Immunology, at the Peter Doherty Institute for Infection and Immunity, The University of Melbourne, Melbourne, Victoria, Australia. [5]Centre for Immunology & Infection, Hong Kong Science and Technology Park, Hong Kong SAR, China. [6]Center of Scientific Excellence for Influenza Viruses, National Research Centre, Giza, Egypt. [7]IHAP, Université de Toulouse, Institut national de recherche pour l'agriculture, l'alimentation et l'environnement, Ecole Nationale Vétérinaire de Toulouse, Toulouse, France. [8]Human Link, DMCC, Dubai, United Arab Emirates. [9]Department of Infectious Diseases, St Jude Children's Research Hospital, Memphis, TN, USA. ✉e-mail: veej@hku.hk

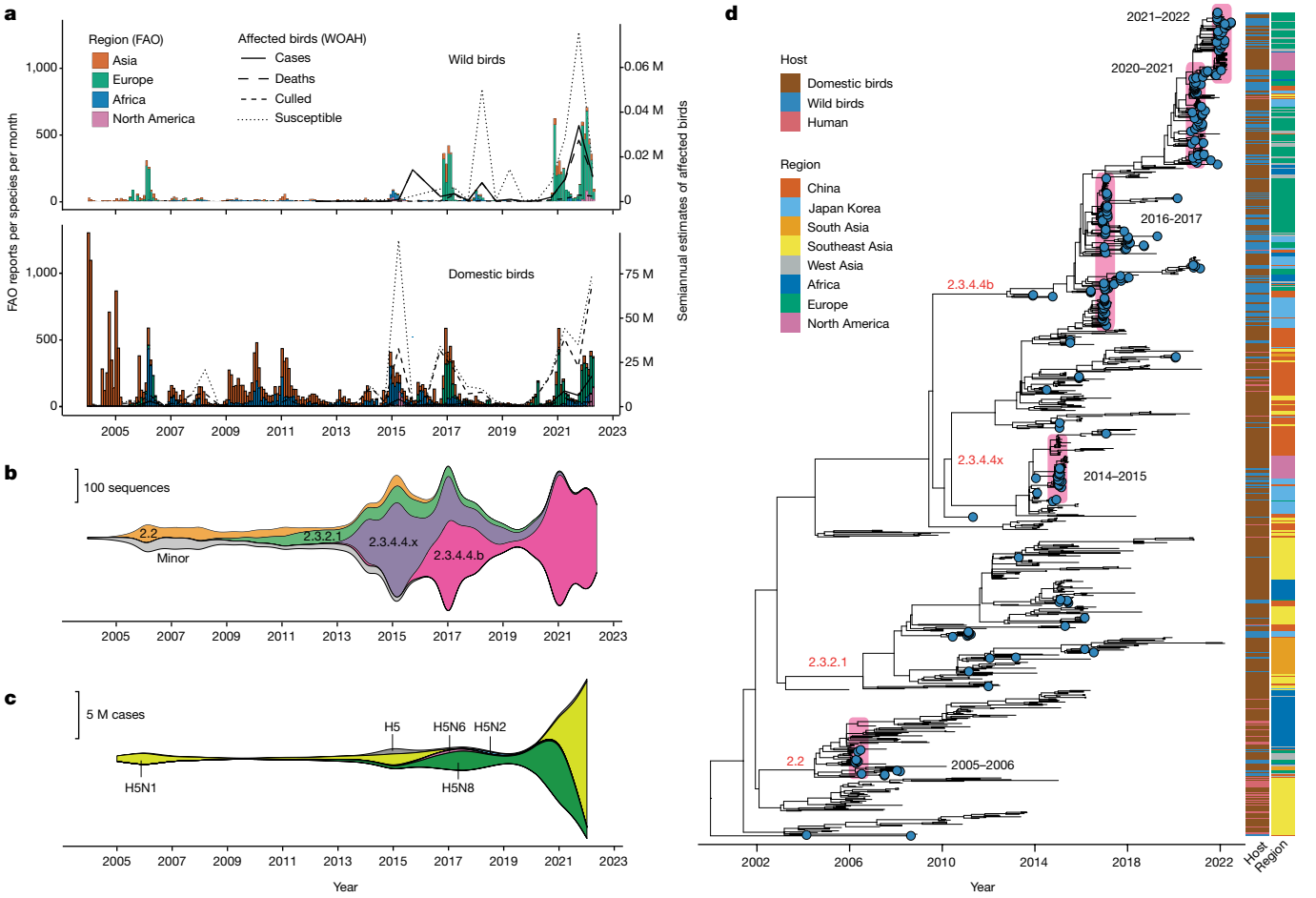


Fig. 1 | HPAI H5 outbreak reports, subtypes and clades. a, Time series comparing HPAI H5 outbreaks in wild birds (top) and poultry (bottom) by geographic region reported to the FAO. Semiannual counts of HPAI H5–affected birds reported to the WOAH are plotted on the right *y* axis. M, million. **b**, Temporal changes in HPAI H5 HA clade prevalence estimated using sample collection dates of sequences submitted to the GISAID and NCBI Influenza Virus Resource databases from January 2004 to June 2022. **c**, Temporal changes in HPAI H5Nx subtype prevalence estimated using observation dates of all reported cases submitted to the WOAH from January 2005 to January 2022. **d**, Time-scaled maximum-likelihood tree of HPAI H5 *HA* on the basis of 1,000 sequences subsampled from all available sequences ($n = 10{,}602$). Four main clades are labelled in red, and significant wild-bird resurgence events are highlighted with pink bars and denoted in black text. Tips representing wild-bird samples are colour-coded in blue. Host and region of isolation are shown as bars.

caused unprecedented outbreaks in diverse wild-bird species across five continents and a significant rise in incidental infections in wild carnivores[3], mink farms[16] and marine mammals[17].

In this Article, we aim to address the uncertain origins of recent HPAI resurgences and the underlying evolution of HPAI H5Nx viruses in wild- and domestic-bird populations. We analysed outbreak data reported to the Food and Agricultural Organization of the United Nations (FAO) and World Organization for Animal Health (WOAH) since 2005 alongside more than 10,000 whole genomes to identify epidemic trends in HPAI ecology and evolution.

Epidemiology of resurgent HPAI H5

According to FAO and WOAH reports, since 2005, both wild and domestic birds have experienced seasonal outbreaks, shifting from Asia and Africa towards Europe and North America (Fig. 1a). Four significant HPAI H5 wild-bird epizootics have occurred since the 2005–2006 resurgence, caused by clade 2.3.4.4 in 2014–2015 and clade 2.3.4.4b in 2016–2017, 2020–2021 and 2021–2022 to the present (Fig. 1b)[9,10]. H5N8 viruses were responsible for the 2014–2015, 2016–2017 and 2020–2021 outbreaks, while H5N6 only played a nominal role in the 2016–2017 wild-bird resurgence (Fig. 1c)[18]. Since 2021–2022, a novel reassortant H5N1 virus has nearly replaced all other HPAI H5 viruses globally (Fig. 1c,d).

Although the 2014–2015 resurgence is notable for spreading across Asia to Europe and North America, resulting in a loss of more than 50 million poultry in the USA[19], poultry outbreaks in Europe and wild-bird detections globally were relatively minor (Fig. 1a). The 2016–2017 epidemic in wild birds lasted five months, with nearly 400 outbreaks per month at its peak (Fig. 1a). The 2017–2018 season saw fewer outbreaks, but a greater number of wild birds were affected across several regions. Following sporadic detections from 2018 to 2020, more than 200 outbreaks per month were reported in 2020–2021, and more than 400 outbreaks per month were recorded during the 2021–2022 season. Substantial outbreaks have continued into 2023. Additionally, the number of infected wild-bird species increased across all affected regions to varying degrees during the 2020–2021 and 2021–2022 seasons (Extended Data Fig. 2). The semiannual estimate of confirmed HPAI H5 cases in wild birds (predominantly dead birds) peaked at 34,000 during the second half of 2021, although in many instances the number of wild-bird cases reported to WOAH/FAO includes only birds tested and positive for HPAI and is therefore a substantial underestimate (Fig. 1a).

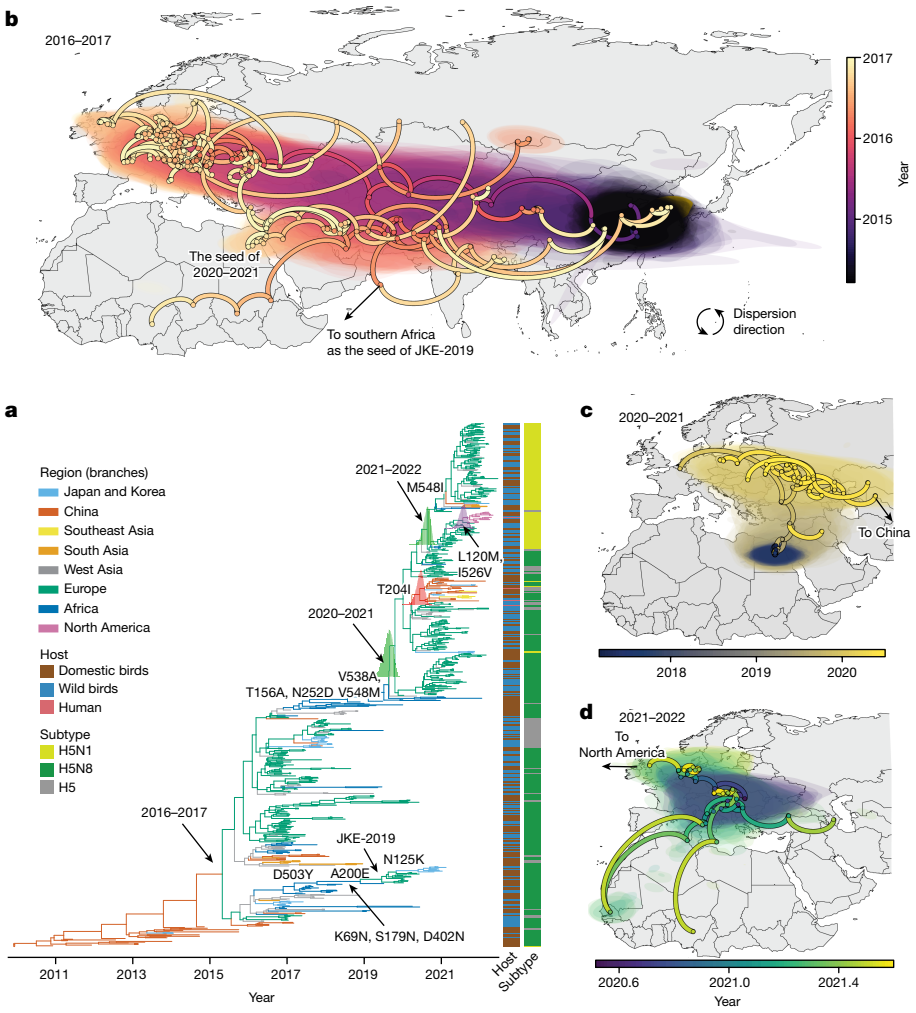


Fig. 2 | Evolution of clade 2.3.4.4b *HA* genes and early migration patterns of resurgent HPAI. a, Maximum clade credibility tree with branches coloured by discrete geographic region. Colour bars indicate host and subtype. The posterior distribution of the tMRCA is shown as bar charts on specific nodes. **b**, Continuous phylogeographic reconstruction of the spread of H5N8 from mid-2014 to 2017. **c**, Early spread of H5N8 from 2017 to mid-2020 before the 2020–2021 resurgence. **d**, Early spread of H5N1 from mid-2020 to 2021 before the 2021–2022 resurgence. Circles represent nodes in the maximum clade credibility phylogeny, coloured by the inferred time of occurrence. Arcs indicate direction of dispersion (counterclockwise) between nodes. An interval of 80% HPD is depicted by shaded areas, illustrating the uncertainty of the phylogeographic estimates.

The number of domestic-bird outbreaks generally corresponded to increases in wild-bird outbreaks. Between January and June 2022, more than 69 million domestic birds were culled. Notably, substantial numbers of poultry outbreaks were also recorded in early 2020, just before wild-bird epizootics spiked during April 2020 (Fig. 1a).

Wild-bird resurgence events exhibit seasonal patterns, with widespread outbreaks beginning in November (Fig. 1a and Extended Data Fig. 3). This seasonality is attributed to the arrival of migratory birds from their arctic breeding grounds[20], coinciding with the 0 °C isotherm[21]. Consistent seasonality has led reporting agencies to describe annual events in waves, starting in September; however, cases in Europe continued through the summer of 2022.

Resurgent HPAI H5 origins

Analysis of HPAI H5 genomes indicates the 2014–2015 and 2016–2017 epidemics originated from independent viral lineages in China (Figs. 1d and 2a). By contrast, all eight genes involved in the 2020–2021 outbreaks evolved from clade 2.3.4.4b H5N8 viruses first detected in Egyptian poultry in 2016–2017 (Fig. 3, Extended Data Figs. 4 and 5 and Supplementary Data 1). Despite limited HPAI H5 surveillance in surrounding poultry networks, continuous detection of ancestral lineages in Egypt strongly implies regional evolution of the 2020–2021 resurgent viruses. The H5N1 viruses responsible for the 2021–2022 epizootic emerged from H5N8 viruses in Europe in 2020 (mean time of most recent common ancestor (tMRCA) of *HA* gene 17 August 2020, 95% higher posterior density (HPD) 13 June 2020, 14 October 2020), with an N1 *NA* and five internal genes (*PB2*, *PB1*, *PA*, *NP* and *NS*) derived through reassortment with LPAI circulating in European wild birds since 2019 (Supplementary Data 1). Estimates of tMRCA showed ancestral genes of all three main resurgent lineages (clade 2.3.4.4b H5N8 in 2016–2017 and 2020–2021 and clade 2.3.4.4b H5N1 in 2021–2022) circulated in wild birds for at least a year before causing widespread outbreaks from November onwards.

A separate clade 2.3.4.4b H5N8 virus lineage that derived from the 2016–2017 resurgence (hereafter termed JKE-2019) was initially reported in domestic poultry in central and southeastern Europe during 2019 and early 2020 and later emerged in wild birds in Japan and Korea in 2020–2021 (Fig. 2a). JKE-2019 retained six of eight genes of clade 2.3.4.4b viruses found in west and southern Africa during 2018–2019, whereas the *PB1* and *NP* genes were obtained from LPAI viruses through reassortment[22] (Extended Data Fig. 5).

We reconstructed the diffusion of *HA* clade 2.3.4.4b using discrete and continuous phylogeographic analyses and found that the initial

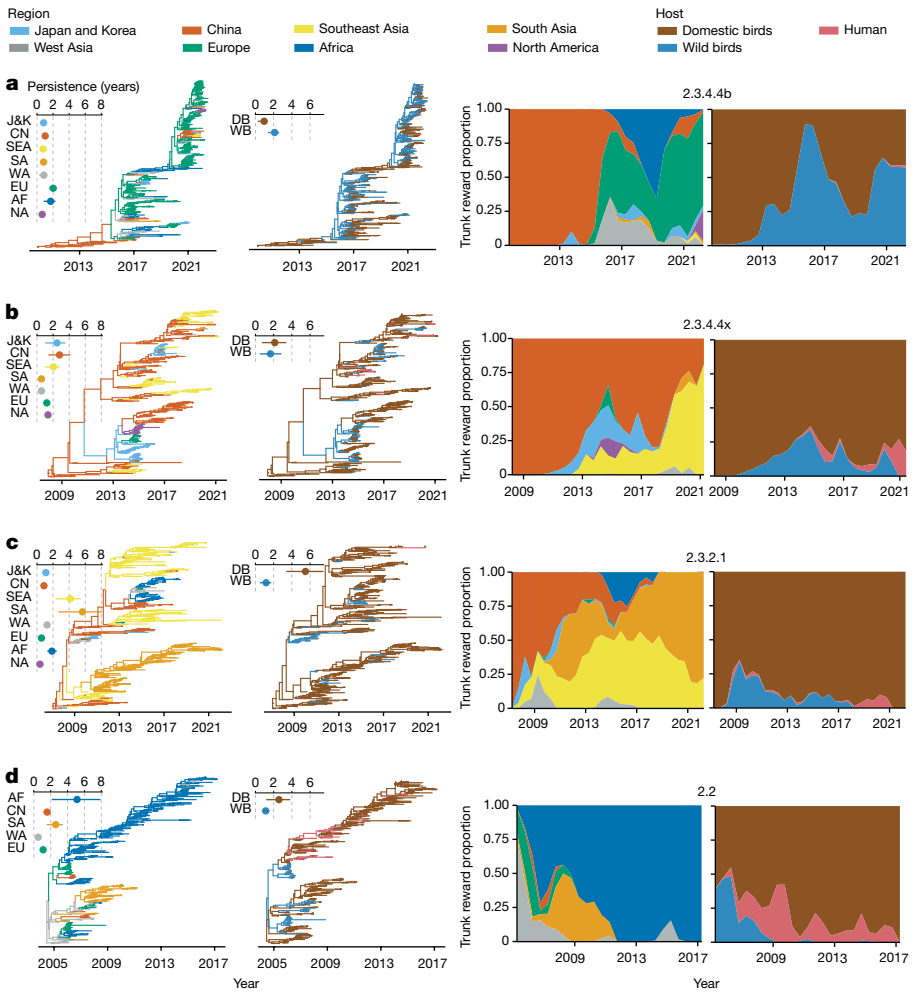


Fig. 3 | Geographic and host transmission patterns of principal HPAI H5 clades. a–**d**, Clades 2.3.4.4b (**a**), 2.3.4.4x (**b**), 2.3.2.1 (**c**) and 2.2 (**d**). First and second columns, maximum clade credibility trees computed from 1,350 posterior trees for each clade coloured by eight geographic regions (Japan and Korea (J&K), China (CN), Southeast Asia (SEA), South Asia (SA), West Asia (WA), Europe (EU), Africa (AF) and North America (NA)) (first column) and hosts (domestic birds (DB), wild birds (WB) and humans (H)) (second column). Insets, duration of region- or host-specific persistence in years, where circles represent the mean and lines represent the interquartile range. Third and fourth columns, proportional ancestral region and host states on the phylogenetic tree trunk over time.

migration of 2.3.4.4b H5N8 from its inferred ancestors in Egyptian domestic poultry (posterior probability, 0.97, Extended Data Fig. 6) was through the Black Sea–Mediterranean flyway towards western Russia and eastern Europe during 2019 (Fig. 2c and Extended Data Fig. 7), following a pattern observed in previous resurgence events[23]. During mid to late 2019, 2.3.4.4b H5N8 was introduced to the northern coastal regions of central Europe (Fig. 2c). Epidemiological reports indicate local circulation before the resurgence in late 2020 (Extended Data Fig. 3). As sampling in non-endemic regions is often limited to outbreak response, early dissemination in Europe is unclear. Nevertheless, HPAI H5N8 spread rapidly across Eurasia and became established in wild birds and poultry during 2020. By October, clade 2.3.4.4b H5N8 viruses with an *HA* T204I mutation migrated east to China, acquiring an N6 *NA* and forming an Asian lineage (tMRCA 9 June 2020, 95% HPD 3 April, 11 August) responsible for 41 of the 65 known human cases of H5N6 infection so far[24].

Continuous phylogeography indicates HPAI H5N1 emerged in eastern Europe during mid-2020, diversifying into two geographically separate lineages (Fig. 2a,d). The lineage that circulated across the northern coastal regions of central Europe in late 2020 (mean tMRCA of *HA* gene 27 November 2020, 95% HPD 30 September 2020, 24 January 2021) seeded North American outbreaks by means of the East Atlantic flyway with two more *HA* mutations (L120M and I526V) in mid-2021 (mean tMRCA of *HA* gene 24 July 2021, 95% HPD 27 May, 19 September). The other European lineage seeded introductions to Africa along the Adriatic flyway around the Mediterranean Sea (Fig. 2d). It acquired a further *HA* substitution (M548I) before causing outbreaks across Eurasia during 2021–2022, indicating the potential importance of this site for adaptation to rapid dispersal among wild birds. Discrete phylogeography confirmed that eastern Europe predominantly seeded HPAI H5N1 outbreaks in other regions (posterior probability, 0.83), with 20.1% Markov rewards (denoting time spent in the region) and 42.2% of Markov jumps (denoting region transitions) (Supplementary Table 1). More than 20% of Markov jumps with definitive support (Bayes factors (BF) > 100) were between western Europe and northern Europe along the East Atlantic flyway (Supplementary Table 1 and Extended Data Fig. 6). From northern Europe to North America, two Markov jumps had sufficient evidence (BF = 75) (Supplementary Table 1), consistent with our continuous phylogeographic inference and previous studies[25,26].

Phylodynamics of clade 2.3.4.4b viruses

The emergence of epizootic lineages in Africa and Europe signifies a shift away from Asia as the HPAI H5 epicentre. To infer changes

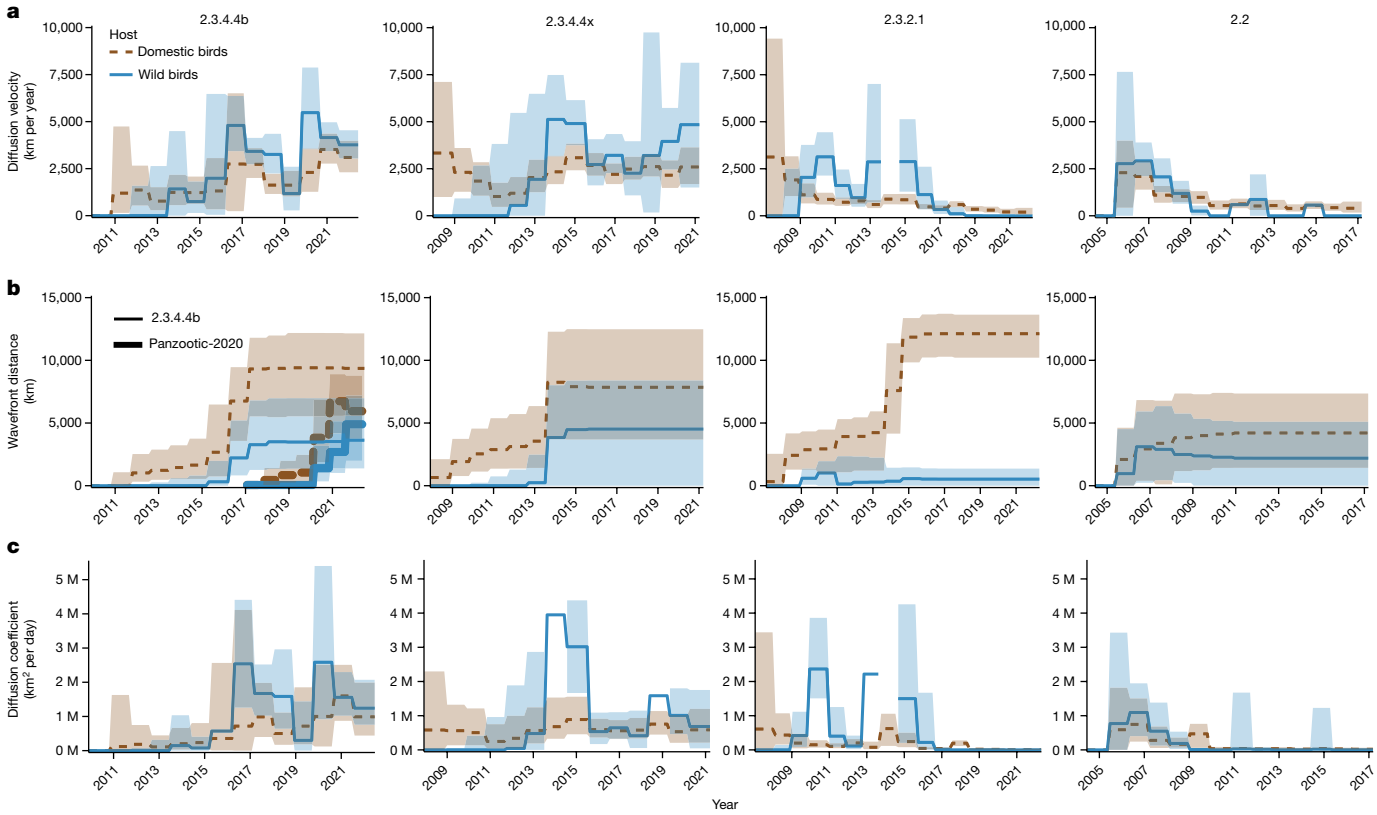


Fig. 4 | The contrasting spatial epidemiology among HPAI H5 clades 2.3.4.4b, 2.3.4.4x, 2.3.2.1 and 2.2. a, Median viral diffusion velocity (km per year) of domestic (brown) and wild birds (blue) over time for each clade. **b**, Median viral wavefront distance (km) of wild and domestic birds over time for each clade. We recalculated the wavefront distance in the panzootic-2020 clade (including 2020–2021 and 2021–2022 resurgences, shown in a thick line), in which Egypt was regarded as the epidemic's origin. **c**, Median viral diffusion coefficient (km^2 per day) over time for each clade. The diffusion velocity of wild birds in clade 2.3.2.1 during 2014 is not shown in the plot owing to abnormal estimation value, possibly caused by insufficient sampling. The shaded areas denote the 95% confidence intervals, in which some extreme values are not shown.

in dispersion and spillover dynamics, we applied an asymmetric discrete-trait model with Bayesian stochastic search variable selection to reconstruct diffusion of principal HPAI H5 clades among eight geographic regions (Africa, China, Europe, Japan and Korea, North America, South Asia, Southeast Asia and West Asia) and different hosts (domestic and wild birds and humans). In contrast to earlier HPAI H5 clades (2.2, 2.3.2.1 and other 2.3.4.4 subclades including a,c–f, collectively 2.3.4.4x), which primarily circulated regionally in domestic birds with brief periods of rapid dispersal by means of wild birds (Fig. 3), clade 2.3.4.4b viruses showed a considerably longer mean duration of persistence in wild birds (51.7% Markov rewards) than domestic birds, with 76.2% of Markov jumps going from wild birds to domestic birds and an epicentre shift from Asia to Europe (Fig. 3, Extended Data Fig. 8 and Supplementary Tables 2 and 3). Significant source populations of clade 2.3.4.4b viruses were identified in Europe (43 Markov jumps, 44.5%), West Asia (22 Markov jumps, 23.3%) and China (11 Markov jumps, 12%) (Extended Data Fig. 8 and Supplementary Table 2). Lineages spent the most time in Europe (51.7% Markov rewards), followed by China (16.7% Markov rewards) and West Asia (10.9% Markov rewards), whereas Africa (13.6% Markov rewards) acted as a sink (Extended Data Fig. 8). However, clade 2.3.4.4b lineages persisted in Europe and Africa for around two years, compared to roughly one year in other regions. Furthermore, whereas domestic birds in China occupied the phylogenetic trunk until 2015 and were the probable source of the 2016–2017 wild-bird outbreaks, Europe represented the primary phylogenetic trunk since 2016, with transmission primarily among wild birds. A notable exception occurred around 2019, when more than 50% of the phylogenetic trunk belonged to lineages circulating in domestic birds in Africa (Fig. 3). Several *HA* sublineages of the 2016–2017 resurgence spread from Europe to Africa and became established in both domestic- and wild-bird populations in South Africa and Nigeria. Iran, Denmark and Bulgaria also detected sustained transmission of 2016–2017-like viruses into 2019, whereas non-2.3.4.4b clades predominated in Asia until 2020. Notably, of the nine lineages detected in 2019–2020, five can be traced to poultry viruses detected in Egypt (Fig. 2a, https://nextstrain.org/community/vjlab/episodic-h5/H5).

In comparison to earlier clades, dissemination routes of 2.3.4.4b viruses were notable. Clade 2.3.4.4b frequently transmitted between Asia and Europe and endured the longest migration route from China to the USA by means of Europe (Extended Data Fig. 7). Clade 2.3.4.4x disseminated frequently from China to Southeast Asia, sporadically to Japan and Korea (over five Markov jumps) and once to North America (Supplementary Table 2). Clade 2.3.2.1 migrated between China and Southeast Asia (over five Markov jumps) and, to a lesser extent, to Japan and Korea and Africa by means of West Asia (Extended Data Figs. 7 and 8). Following the early dispersal of clade 2.2 to Europe and Africa, dispersal from West Asia to Africa by means of Europe was mainly associated with domestic-bird networks, although wild-bird surveillance in Africa was sparse at the time. Wild birds were primarily responsible for transmission between West Asia and Europe. Most of the regional migration of clades 2.2, 2.3.2.1 and 2.3.4.4x probably occurred through domestic poultry networks and associated human activity (Extended Data Fig. 7); however, short-range spread through unsampled wild birds cannot be excluded.

To better understand the spatial epidemiology of resurgent HPAI H5 clades, we estimated the wavefront distance and velocity[27] from a joint

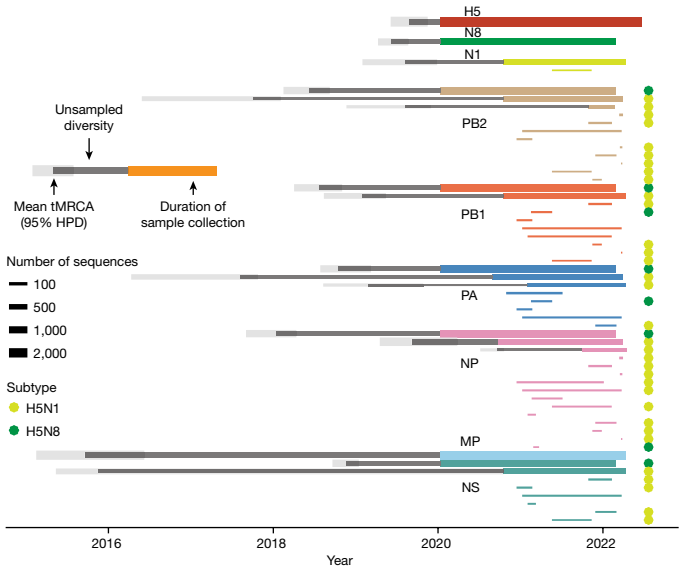


Fig. 5 | Circulation of HPAI gene segments in resurgent 2020–2021 and 2021–2022 viruses. Line thickness indicates the number of sequences in a monophyletic clade (with a minimum of five sequences). Internal genes of H5N1 and H5N8 are labelled with yellow and green snowflakes, respectively. tMRCA is shown for monophyletic clades with more than 100 sequences, estimated with BEAST v.1.10.4 (ref. 48) for the *HA* gene and TreeTime[49] for the remaining genes.

phylodynamic analysis that incorporated continuous spatial diffusion and discrete avian hosts (domestic/wild)[28] (Fig. 4 and Extended Data Fig. 7). The diffusion velocity and coefficient in wild birds was greater than that of domestic birds for all principal gs/Gd clades, meaning that, as expected, wild birds are responsible for faster long-distance dissemination of HPAI H5 viruses. Although previous studies indicate that domestic waterfowl drove the slow but steady regional dissemination of pre-2.3.4.4 HPAI H5 as measured by the wavefront distance[29], and despite a larger bias in wild-bird sampling, we find wild birds disseminated clades 2.3.4.4b (3,630 km, 95% HPD 1,381–6,977) and 2.3.4.4x (4,517 km, 95% HPD 0–8,392) to a greater extent than clades 2.2 (2,199 km, 95% HPD 2–5,097) and 2.3.2.1 (526 km, 95% HPD 0–1,364). Furthermore, wild and domestic birds contributed almost equal wavefront distance (~5,000 km) since the 2.3.4.4b resurgence in wild birds beginning in late 2019 (termed panzootic-2020 in Fig. 4). However, in 2020, the diffusion velocity of 2.3.4.4b from wild birds was nearly three times (5,474 km per year, 95% HPD 3,535–7,877) that of domestic birds, after which the diffusion velocities were similar (3,000–4,000 km per year).

Reassortment and antigenic evolution

Between 1996 and 2004, HPAI H5N1 became established in domestic poultry through extensive reassortment with poultry-adapted LPAI viruses, particularly H9N2 (ref. 30). The recent proliferation of clade 2.3.4.4 also involved extensive reassortment (Extended Data Fig. 9), deriving *NA* subtypes and internal genes from Eurasian wild-bird LPAI viruses[11]. Strikingly, most HPAI H5N8 viruses collected in 2020–2021 maintained a stable genotype with all genes originating from Egyptian poultry, except for one lineage in China that acquired N1 and N6 *NA* genes. Before the 2020–2021 wild-bird resurgence, several *HA* mutations occurred in domestic birds in Egypt between 2017 and 2020, including T156A and N252D, as well as V538A and V548M mutations in the conserved *HA* stalk region (Fig. 2a). Notably, T156A enhances *HA* binding to sialic acid receptors[31]. The 2.3.4.4b H5N1 that subsequently emerged from 2.3.4.4b H5N8 acquired its *NA* and internal genes through reassortment with LPAI viruses (Fig. 5 and Supplementary Data 1). Novel 2.3.4.4b H5N1 reassortants increasingly emerged during 2020–2021 and 2021–2022, with several genotypes enduring into 2021–2022 (Fig. 5). These results indicate that although a predominantly clonal population of HPAI H5N8 caused the 2020–2021 outbreaks, the 2021–2022 H5N1 virus has attained relaxed constraints for reassortment with wild-bird LPAI viruses.

Discussion

Analysis of HPAI H5 episodic resurgence reveals significant shifts in HPAI ecology and evolution, including geographic and host range expansion, increased dispersion velocity and increased reassortment (Fig. 6). The continuing epizootics are being caused by 2.3.4.4b H5 viruses with origins traced to poultry samples collected in Africa, following a significant wild-bird resurgence of HPAI H5N8 in 2016–2017. The dispersion velocity in wild birds increased substantially in 2020, spreading HPAI H5N8 viruses across Eurasia and leading to the subsequent emergence of HPAI H5N1. Central to these findings is the observation that the 2.3.4.4b H5N8 virus that caused the 2020–2021 outbreaks across Eurasia and Africa was of a stable genotype, whereas the novel panzootic 2.3.4.4b H5N1 viruses have attained relaxed constraints for extensive genetic reassortment, indicating potential for enhanced transmission and maintenance across the domestic-bird–wild-bird interface.

Whereas mammalian influenza viruses typically maintain a limited number of stable genome constellations, LPAI viruses in migratory aquatic birds undergo frequent reshuffling, resulting in transient genome constellations[32]. Poultry-adapted LPAI viruses exhibit genetic linkage of the *HA* and *NA* segments (for example, H9N2, H7N9), with limited internal gene-segment exchange[33]. The propensity for HPAI H5 to reassort with other domestic poultry viruses, as was recently observed between clade 2.3.4.4 and H9N2 in Burkina Faso[34], and the recent surge in mammalian HPAI H5 infections are significant causes for concern.

Our findings show clade 2.3.4.4b emergence is shifting from Asia to Africa, underscoring the need for expanded surveillance capacity. Globally, genomic surveillance is limited and biased across time, geography and host species. Despite the substantial number of HPAI H5 outbreaks reported to FAO and WOAH, only 50% of outbreaks and 0.2% of cases were sequenced (Extended Data Fig. 10). Although HPAI H5 sequence availability generally corresponded to regional outbreaks (Fig. 1a and Extended Data Fig. 10), sampling bias was observed at the country level, especially in Africa (Supplementary Data 2). Some countries with endemic HPAI H5 conduct routine poultry surveillance, but many only sample following mass-mortality events. Wild-bird sampling is of lower priority and is often biased towards species that are easily sampled. Despite efforts to subsample equitably (Supplementary Tables 4–8), biased sampling could affect ancestral reconstruction[35], especially for long-distance dispersal events. Additionally, the use of centroid coordinates may skew continuous phylogeography inferences. Nevertheless, our findings are consistent with the timing of reports emerging from new areas.

Despite the sampling bias towards poultry, we found that the dissemination of clade 2.3.4.4b H5 viruses by wild birds was similar to that of domestic birds, which were previously inferred as key drivers of HPAI H5 spread[29]. Domestic birds' dispersal velocity reflects and is constrained by poultry trade and associated human activity, whereas that of migratory birds is more influenced by ecological determinants such as seasonality, migratory flyways and climate. For example, the primary wild-bird resurgence events in Europe are consistently seasonal (Fig. 1a and Extended Data Fig. 3) and correspond with increased wild-bird HPAI H5 dispersion velocity (Fig. 4). Further, several migration routes of HPAI H5 virus were identified along the Black Sea–Mediterranean, East Atlantic and Adriatic flyways, with reports increasing along the Pacific flyway in South America since late 2022[36]. There is also evidence that migratory behaviour and breeding patterns have recently been affected by climatic variation[37,38]. Expansion and persistence of HPAI H5 viruses in wild birds may therefore

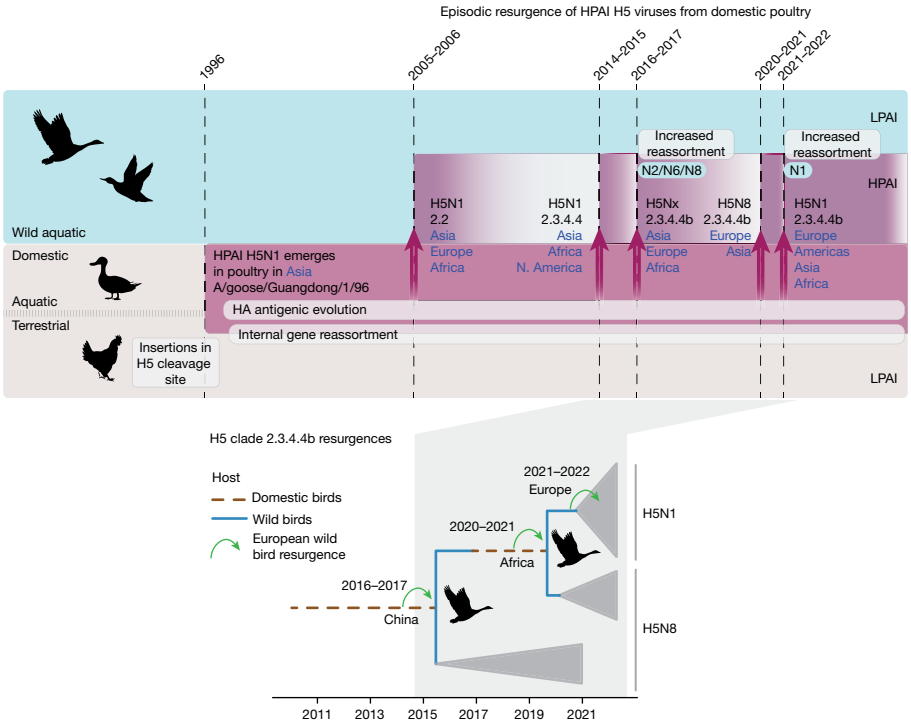


Fig. 6 | Graphical summary of the episodic resurgence of HPAI H5N1 in wild birds. HPAI H5N1 initially emerged in domestic birds (first detected in geese) and then became established in poultry through *HA* diversification and internal gene reassortment with LPAI viruses endemic in poultry. Shaded regions indicate host ecology and avian influenza pathogenicity, with HPAI shown in red and LPAI in wild aquatic birds (blue background) and domestic birds (grey). Dates marked with dashed vertical lines highlight key events in HPAI H5Nx evolution. Blue text highlights the geographic spread of HPAI H5 associated with each primary resurgence event. Arrows represent the key transitions of HPAI H5Nx across the domestic-bird–wild-bird interface. Key evolutionary mechanisms of HPAI H5 are shown in bold. Bottom, phylogenetic tree depicting clade 2.3.4.4b host and region transitions.

continue to be exacerbated by climate change. Future studies will benefit from the integration of host ecology such as species composition, population size and demography and migration patterns (for example, using banding or telemetry data) to inform risk[39,40].

Sustained HPAI H5N1 wild-bird epizootics repeatedly spill over and spill back from domestic birds, increasing zoonotic and pandemic risks. Culling is commonly used to curb HPAI spread in poultry[41,42], and millions of poultry have been culled to stamp out HPAI H5 outbreaks (Fig. 1). However, continuous culling is unsustainable as HPAI H5 is increasingly persistent in wild birds. Mass mortality in wildlife raises concern around a loss of biodiversity and disruption of ecosystem homoeostasis[43]. Subsidized poultry vaccination is therefore increasingly being considered for outbreak prevention in Europe and North America. Several countries in Asia and Africa use vaccination with variable effects[44], whereas Hong Kong and parts of China and Southeast Asia combine vaccination with other control methods[45,46]. A key concern is the role of poultry vaccination in driving endemicity and evolution of HPAI H5 lineages[47]. The proximity of poultry networks to significant flyways in northern Africa, the Middle East and eastern Europe, where vaccination practices are diverse and incomplete, is also concerning. To address these issues, it is necessary to enhance global surveillance and improve multifaceted mitigation strategies for outbreak prevention and response. A better understanding of the ecological properties that enhance and sustain transmission in wild birds and the consequences of poultry vaccination with variable uptake will be crucial to mitigate future HPAI outbreaks, which pose unpredictable epizootic, zoonotic and pandemic threats.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-06631-2.

1. Wille, M. & Barr, I. G. Resurgence of avian influenza virus. *Science* **376**, 459–460 (2022).
2. European Food Safety Authority et al. Avian influenza overview December 2021–March 2022. *EFSA J.* **20**, e07289 (2022).
3. 2022-2023 detections of highly pathogenic avian influenza. *US Department of Agriculture Animal and Plant Health Inspection Service* https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/animal-disease-information/avian/avian-influenza/2022-hpai (2023).
4. Escalera-Zamudio, M. et al. Parallel evolution in the emergence of highly pathogenic avian influenza A viruses. *Nat. Commun.* **11**, 5511 (2020).
5. Becker, W. B. The isolation and classification of Tern virus: influenza A-Tern South Africa-1961. *J. Hyg. (Lond.)* **64**, 309–320 (1966).
6. Chen, H. et al. Establishment of multiple sublineages of H5N1 influenza virus in Asia: implications for pandemic control. *Proc. Natl Acad. Sci. USA* **103**, 2845–2850 (2006).
7. WHO/OIE/FAO H5N1 Evolution Working Group. Continued evolution of highly pathogenic avian influenza A (H5N1): updated nomenclature. *Influenza Other Respir. Viruses* **6**, 1–5 (2012).
8. Chen, H. et al. H5N1 virus outbreak in migratory waterfowl. *Nature* **436**, 191–192 (2005).
9. Ip, H. S. et al. High rates of detection of clade 2.3.4.4 highly pathogenic avian influenza H5 viruses in wild birds in the Pacific Northwest during the winter of 2014–15. *Avian Dis.* **60**, 354–358 (2016).
10. Engelsma, M., Heutink, R., Harders, F., Germeraad, E. A. & Beerens, N. Multiple introductions of reassorted highly pathogenic avian influenza H5Nx viruses clade 2.3.4.4b causing outbreaks in wild birds and poultry in The Netherlands, 2020–2021. *Microbiol. Spectr.* **10**, e0249921 (2022).
11. Global Consortium for H5N8 and Related Influenza Viruses.Role for migratory wild birds in the global spread of avian influenza H5N8. *Science* **354**, 213–217 (2016).
12. Li, Y. T., Su, Y. C. F. & Smith, G. J. D. H5Nx viruses emerged during the suppression of H5N1 virus populations in poultry. *Microbiol. Spectr.* **9**, e0130921 (2021).
13. van den Brand, J. M. A. et al. Wild ducks excrete highly pathogenic avian influenza virus H5N8 (2014–2015) without clinical or pathological evidence of disease. *Emerg. Microbes Infect.* **7**, 67 (2018).
14. 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. *J. Virol.* **95**, e0095521 (2021).
15. Lewis, N. S. et al. Emergence and spread of novel H5N8, H5N5 and H5N1 clade 2.3.4.4 highly pathogenic avian influenza in 2020. *Emerg. Microbes Infect.* **10**, 148–151 (2021).

16. Aguero, M. et al. Highly pathogenic avian influenza A(H5N1) virus infection in farmed minks, Spain, October 2022. *Euro. Surveill.* **28**, 2300001 (2023).
17. Puryear, W. et al. Highly pathogenic avian influenza A(H5N1) virus outbreak in New England seals, United States. *Emerg. Infect. Dis.* **29**, 786–791 (2023).
18. Poen, M. J. et al. Co-circulation of genetically distinct highly pathogenic avian influenza A clade 2.3.4.4 (H5N6) viruses in wild waterfowl and poultry in Europe and East Asia, 2017–18. *Virus Evol.* **5**, vez004 (2019).
19. Ramos, S., MacLachlan, M. & Melton, A. *Impacts of the 2014-2015 Highly Pathogenic Avian Influenza Outbreak on the US Poultry Sector*. Livestock, Dairy, and Poultry Outlook No. (LDPM-282-02) (USDA, 2017).
20. Gass, J. D. Jr et al. Global dissemination of influenza A virus is driven by wild bird migration through arctic and subarctic zones. *Mol. Ecol.* https://doi.org/10.1111/mec.16738 (2022).
21. Reperant, L. A., Fuckar, N. S., Osterhaus, A. D., Dobson, A. P. & Kuiken, T. Spatial and temporal association of outbreaks of H5N1 influenza virus infection in wild birds with the 0 degrees C isotherm. *PLoS Pathog.* **6**, e1000854 (2010).
22. Swieton, E. et al. Sub-Saharan Africa and Eurasia ancestry of reassortant highly pathogenic avian influenza A(H5N8) virus, Europe, December 2019. *Emerg. Infect. Dis.* **26**, 1557–1561 (2020).
23. Napp, S., Majo, N., Sanchez-Gonzalez, R. & Vergara-Alert, J. Emergence and spread of highly pathogenic avian influenza A(H5N8) in Europe in 2016–2017. *Transbound. Emerg. Dis.* **65**, 1217–1226 (2018).
24. Zhu, W. et al. Epidemiologic, clinical, and genetic characteristics of human infections with influenza A(H5N6) viruses, China. *Emerg. Infect. Dis.* **28**, 1332–1344 (2022).
25. Gass, J. D. Jr et al. Global dissemination of influenza A virus is driven by wild bird migration through arctic and subarctic zones. *Mol. Ecol.* **32**, 198–213 (2023).
26. Gunther, A. et al. Iceland as stepping stone for spread of highly pathogenic avian influenza virus between Europe and North America. *Emerg. Infect. Dis.* **28**, 2383–2388 (2022).
27. Pybus, O. G. et al. Unifying the spatial epidemiology and molecular evolution of emerging epidemics. *Proc. Natl Acad. Sci. USA* **109**, 15066–15071 (2012).
28. Trovao, 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).
29. Hill, N. J. et al. Ecological divergence of wild birds drives avian influenza spillover and global spread. *PLoS Pathog.* **18**, e1010062 (2022).
30. Vijaykrishna, D. et al. Evolutionary dynamics and emergence of panzootic H5N1 influenza viruses. *PLoS Pathog.* **4**, e1000161 (2008).
31. Linster, M. et al. Identification, characterization, and natural selection of mutations driving airborne transmission of A/H5N1 virus. *Cell* **157**, 329–339 (2014).
32. Wille, M. et al. Evolutionary features of a prolific subtype of avian influenza A virus in European waterfowl. *Virus Evol.* **8**, veac074 (2022).
33. Pu, J. et al. Reassortment with dominant chicken H9N2 influenza virus contributed to the fifth H7N9 virus human epidemic. *J. Virol.* **95**, e01578-20 (2021).
34. Ouoba, L. B. et al. Emergence of a reassortant 2.3.4.4b highly pathogenic H5N1 avian influenza virus containing H9N2 PA gene in Burkina Faso, West Africa, in 2021. *Viruses* **14**, 1901 (2022).
35. Kalkauskas, A. et al. Sampling bias and model choice in continuous phylogeography: getting lost on a random walk. *PLoS Comput. Biol.* **17**, e1008561 (2021).
36. Jimenez-Bluhm, P. et al. Detection and phylogenetic analysis of highly pathogenic A/H5N1 avian influenza clade 2.3.4.4b virus in Chile, 2022. Preprint at *bioRxiv* https://doi.org/10.1101/2023.02.01.526205 (2023).
37. Rushing, C. S., Royle, J. A., Ziolkowski, D. J. Jr & Pardieck, K. L. Migratory behavior and winter geography drive differential range shifts of eastern birds in response to recent climate change. *Proc. Natl Acad. Sci. USA* **117**, 12897–12903 (2020).
38. McLean, N. et al. Warming temperatures drive at least half of the magnitude of long-term trait changes in European birds. *Proc. Natl Acad. Sci. USA* **119**, e2105416119 (2022).
39. Huang, Z. Y. X. et al. Contrasting effects of host species and phylogenetic diversity on the occurrence of HPAI H5N1 in European wild birds. *J. Anim. Ecol.* **88**, 1044–1053 (2019).
40. 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).
41. Boni, M. F., Galvani, A. P., Wickelgren, A. L. & Malani, A. Economic epidemiology of avian influenza on smallholder poultry farms. *Theor. Popul. Biol.* **90**, 135–144 (2013).
42. Liu, S. et al. Control of avian influenza in China: strategies and lessons. *Transbound. Emerg. Dis.* **67**, 1463–1471 (2020).
43. Lederman, Z. One health and culling as a public health measure. *Public Health Ethics* **9**, 5–23 (2016).
44. Peyre, M. et al. Avian influenza vaccination in Egypt: limitations of the current strategy. *J. Mol. Genet. Med.* **3**, 198–204 (2009).
45. Wu, J. et al. Influenza H5/H7 virus vaccination in poultry and reduction of zoonotic infections, Guangdong Province, China, 2017–18. *Emerg. Infect. Dis.* **25**, 116–118 (2019).
46. Ellis, T. M. et al. Use of avian influenza vaccination in Hong Kong. *Dev. Biol.* **124**, 133–143 (2006).
47. Grund, C. et al. Highly pathogenic avian influenza virus H5N1 from Egypt escapes vaccine-induced immunity but confers clinical protection against a heterologous clade 2.2.1 Egyptian isolate. *Vaccine* **29**, 5567–5573 (2011).
48. Suchard, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018).
49. Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: maximum-likelihood phylodynamic analysis. *Virus Evol.* **4**, vex042 (2018).

Methods

Data source and preparation

HPAI H5 genomes with at least *HA* gene and sample information including collection date, location and host species, were obtained from the Global Initiative on Sharing All Influenza Data (GISAID) (https://platform.epicov.org/) and National Center for Biotechnology Information (NCBI) Influenza Virus Resource (https://www.ncbi.nlm.nih.gov/genomes/FLU/) databases on 11 July 2022. After removing duplicate isolates, laboratory-derived and mixed-subtype isolates, sequences with less than 85% gene coverage and sequences with incomplete collection dates, H5 clades on the basis of the World Health Organization (WHO) gs/Gd H5N1 nomenclature system[50] were determined using LABEL v.0.6.3 (ref. 51). Clade 2.3.4.4 was further assigned into subclades 2.3.4.4a–h on the basis of phylogenetic relationships to WHO H5 candidate vaccine viruses, with known clade assignment estimated using a maximum-likelihood phylogenetic tree generated using the Jukes–Cantor nucleotide substitution model in FastTree v.2.1.1.

Avian hosts were classified as domestic and wild birds using strain names, associated metadata and original publications (https://doi.org/10.5281/zenodo.8251324). Location coordinates were used to analyse diffusion in continuous space and were accurate at least to the province/state level for China, Russia and the United States. For discrete analysis and visualization, locations were classified into countries and regions according to the country–region list in NextStrain[52].

To mitigate sampling bias for phylogeographic analyses, curated datasets were subsampled by either epidemiological information (including country, host and sampling date) or phylogenetic relationship. HPAI H5 *HA* datasets were subsampled randomly with at most two sequences for clade 2.3.4.4b, five sequences for clade 2.3.4.4x, three sequences for clade 2.3.2.1 and six sequences for clade 2.2 per country per host per season of each year (season 1, January–March; season 2, April–June; season 3, July–September; and season 4, October–December). The final subsampled *HA* datasets contained 715, 617, 579 and 563 sequences for the four clades.

The R package 'ggstream' v.0.1 was used to map temporal changes in the sampling of HPAI H5 clades, and 'rworldmap' v.1.3 and 'rnaturalearth' v.0.3 were used to plot world maps using world vector map data from R package 'rnaturalearthdata' v.0.1.

Phylogenetic analyses

Gene sequences were aligned using MAFFT v.7.490 (ref. 53) and trimmed using trimAL v.1.4 (ref. 54) with a 50% gap threshold followed by manual trimming to the open-reading frame. Maximum-likelihood (ML) trees were generated using IQ-TREE v.2.1.4 (ref. 55) with the best-fit nucleotide substitution model. The ML phylogenies were used to check each dataset for molecular clock outliers (sequences that have disproportionately too much or too little root-to-tip divergence for their sampling time) using TempEst v.1.5.3 (ref. 56). The time-scaled ML trees in this study were generated with TreeTime v.0.9.1 (ref. 49). *HA* phylogeny showing sampling location (country/region), host and clade details of all sequences are available as NextStrain[52] builds at https://nextstrain.org/community/vjlab/episodic-h5/H5.

Reassortment analyses

To compare reassortment events that occurred among major clades, including clades 2.3.4.4b, 2.3.4.4x, 2.3.2.1 and 2.2, we first constructed ML trees for each gene using IQ-TREE v.2.1.4. At most 200 isolates with eight segments were subsampled using the Phylogenetic Diversity Analyzer tool v.1.0.3 (http://www.cibiv.at/software/pda)[57] for each major clade and the minor clade, finally resulting in 909 sequences for each *HA, PB2, PB1, PA, NP, MP* and *NS* dataset and 897 sequences for *NA* dataset, which excluded minor subtypes of *NA* genes (N3, N4, N5 and N9). Baltic v.0.1.5 (https://github.com/evogytis/baltic) was used to visualize the incongruence between phylogenetic trees of eight genes. The *HA* tree was rooted using clade 0, and the remaining genes were midpoint rooted. Isolates were coloured according to the *HA* clade. Furthermore, the timeline of reassortment of 2020–2021 and 2021–2022 viruses was inferred using divergence times summarized in Supplementary Data 1 and Fig. 5.

HPAI H5 poultry and wild-bird outbreaks

All reported and confirmed infections of HPAI H5 viruses in humans were obtained from WHO (https://www.who.int/emergencies/disease-outbreak-news). Confirmed detections/outbreaks in domestic and wild birds globally were obtained from World Animal Health Information System, WOAH (https://wahis.woah.org/) and EMPRES-i+ Global Animal Disease Information System, Food and Agriculture Organization (https://empres-i.apps.fao.org/).

Bayesian evolutionary inference

Divergence times and evolutionary rates were estimated using an uncorrelated relaxed clock model under a Bayesian framework using Markov chain Monte Carlo (MCMC) sampling in BEAST v.1.10.4 (ref. 48) and the BEAGLE high-performance library[58]. A flexible Gaussian Markov random field skyride coalescent model and a general time reversible nucleotide substitution model were used with a gamma distribution of substitution rates. At least two independent MCMC chains with 100 million states were performed for each lineage, sampling every 20,000 and discarding 10% as burn-in. Runs were combined to ensure an effective sample size of more than 200 in Tracer v.1.7.1. A subset of 500 trees randomly selected from the posterior distribution was used to generate an empirical tree distribution used in the subsequent phylogeographic analysis, an approach that reduced computational time and burden[28,29]. We used the BaTS[59] to investigate the degree of phylogeographic structure (phylogenetic clustering by sampling location and host), which was compared to a null hypothesis generated by tip randomization with 1,000 replicates.

Discrete phylogeography

To reconstruct spatial diffusion among a set of eight geographic regions (Africa, China, Europe, Japan and Korea, North America, South Asia, Southeast Asia and West Asia) and different hosts (domestic birds, wild birds and humans), we conducted asymmetric discrete-trait phylogeographic analyses with Bayesian stochastic search variable selection in BEAST v.1.10.4 (ref. 48). SpreadD3 v.0.9.6 (ref. 60) was used to estimate BF to determine statistical significance ('definitive' (BF > 100) or 'sufficient' (100 > BF > 3)). To count all the transitions between states and the time spent in states between transitions, we again applied the continuous-time Markov chains model[61] to complete the Markov jump history over time. We combined three independent chains with five million MCMC steps for each lineage and sampled every 10,000 states. The first 10% of each run was discarded as burn-in, resulting in 1,350 posterior trees with estimates of the ancestral region and host for each internal node. The trunk region/host through time and persistence were measured from these posterior phylogenies using PACT v.0.9.5 (https://github.com/trvrb/PACT), where the trunk consists of all branches ancestral to a virus that was sampled within one year of the most recent sample and the persistence is measured by how long a tip takes to leave its sampled location, counting backwards in phylogenetic trees[62]. To further reveal and quantify the transmission patterns of the panzootic-2020 clade (including 2020–2021 and 2021–2022 resurgences) in Europe, we partitioned Europe and repeated the analysis among eight geographic regions (Africa, Asia outside West Asia, West Asia, eastern Europe, western Europe, southern Europe, northern Europe, Africa and North America).

Phylodynamics incorporating geography and host

To complement the discrete spatial diffusion analysis and reconstruct a more detailed geographic history, we estimated the HPAI H5 diffusion

in continuous space (latitude and longitude of country level) using a Cauchy relaxed random walk model[63] with 0.001 jitter window size. Moreover, following the Bayesian method proposed by ref. 28, we incorporated a continuous spatial diffusion process and a discrete host transmission process in a single Bayesian analysis to quantify host-specific diffusion rates and geographic expansion (wavefront distance) for each lineage (https://github.com/vjlab/episodic-h5). At least two independent MCMC chains were performed, sampling every 10,000 steps and discarding 10% as burn-in to ensure effective sample size >200 for each parameter. The continuous phylogeographic analysis was visualized using the R package 'seraphim'[64] and codes from refs. 65,66 using the Python library 'matplotlib'.

Reporting summary

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

Data availability

The avian influenza virus sequences and associated metadata used in this study were downloaded from GISAID and NCBI GenBank. Accession numbers and acknowledgements are provided in Supplementary Data 3. Details of confirmed detections/outbreaks in domestic and wild birds globally are available from World Animal Health Information System, WOAH (https://wahis.woah.org/) and EMPRES-i+ Global Animal Disease Information System, Food and Agriculture Organization (https://empres-i.apps.fao.org/).

Code availability

Data, code and analysis files are available at https://doi.org/10.5281/zenodo.8251324.

50. Smith, G. J. D. et al. 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).
51. Shepard, S. S. et al. LABEL: fast and accurate lineage assignment with assessment of H5N1 and H9N2 influenza A hemagglutinins. *PLoS ONE* **9**, e86921 (2014).
52. Hadfield, J. et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).
53. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
54. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
55. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
56. 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).
57. Chernomor, O. et al. Split diversity in constrained conservation prioritization using integer linear programming. *Methods Ecol. Evol.* **6**, 83–91 (2015).
58. Ayres, D. L. et al. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Syst. Biol.* **61**, 170–173 (2012).
59. Parker, J., Rambaut, A. & Pybus, O. G. Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. *Infect. Genet. Evol.* **8**, 239–246 (2008).
60. Bielejec, F. et al. SpreaD3: interactive visualization of spatiotemporal history and trait evolutionary processes. *Mol. Biol. Evol.* **33**, 2167–2169 (2016).
61. Minin, V. N. & Suchard, M. A. Counting labeled transitions in continuous-time Markov models of evolution. *J. Math. Biol.* **56**, 391–412 (2008).
62. Bedford, T. et al. Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature* **523**, 217–220 (2015).
63. Minin, V. N., Bloomquist, E. W. & Suchard, M. A. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* **25**, 1459–1471 (2008).
64. Dellicour, S., Rose, R., Faria, N. R., Lemey, P. & Pybus, O. G. SERAPHIM: studying environmental rasters and phylogenetically informed movements. *Bioinformatics* **32**, 3204–3206 (2016).
65. McCrone, J. T. et al. Context-specific emergence and growth of the SARS-CoV-2 Delta variant. *Nature* **610**, 154–160 (2022).
66. Dudas, G. et al. Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* **544**, 309–315 (2017).

Acknowledgements The computations were performed using research facilities offered by Information Technology Services, the University of Hong Kong. We gratefully acknowledge the staff from the originating laboratories responsible for obtaining the specimens and the submitting laboratories where the genome data were generated and shared by means of GISAID (Supplementary Data 3). We thank J.L.-H. Tsui for valuable discussions about phylogeographic analysis. The funding bodies had no role in the design of the study and collection, analysis and interpretation of data or the writing of the manuscript. Support came from National Institutes of Health contract number 75N93021C00016 and US National Science Foundation awards 1911955 and 2200310.

Author contributions V.D. conceived and designed the research. R.X. and X.W. curated the genome datasets. K.M.E. curated the outbreak reports. R.X., K.M.E., X.W. and V.D. performed analysis and designed the figures. R.X., K.M.E., M.W. and V.D. wrote the manuscript with input from S-S.W., M.Z., R.E-S., M.D., L.L.M.P., G.K. and R.J.W. All authors discussed and approved the manuscript.

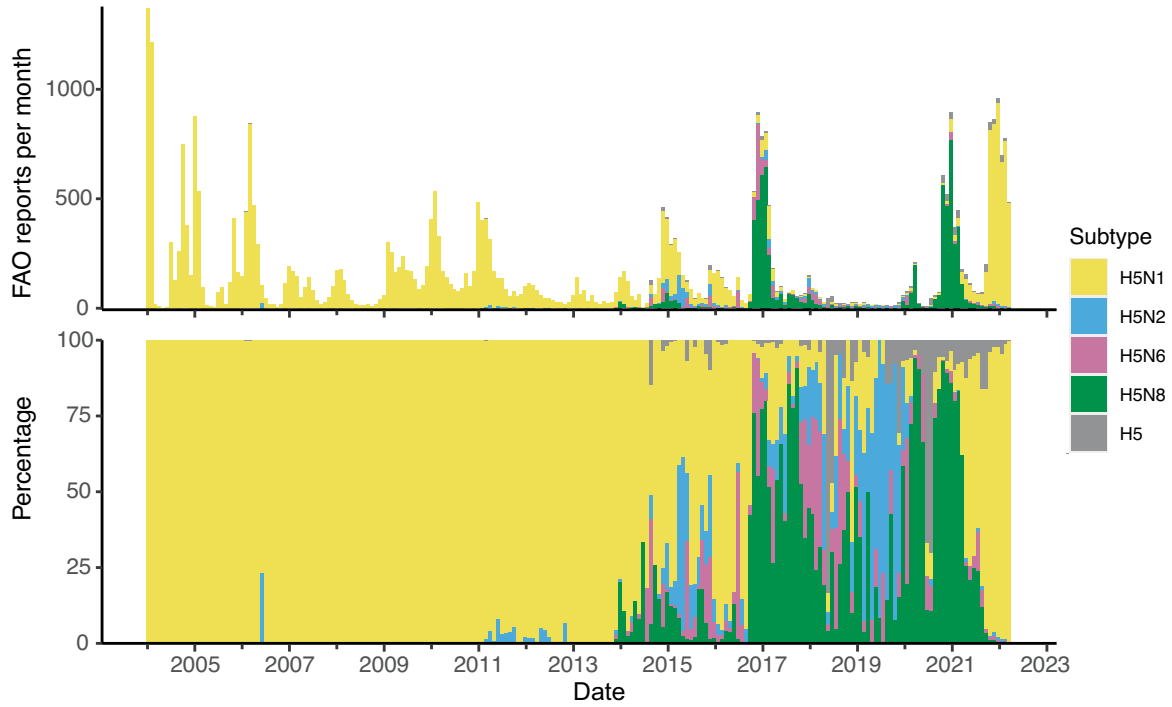
Competing interests The authors declare no competing interests.

Additional information

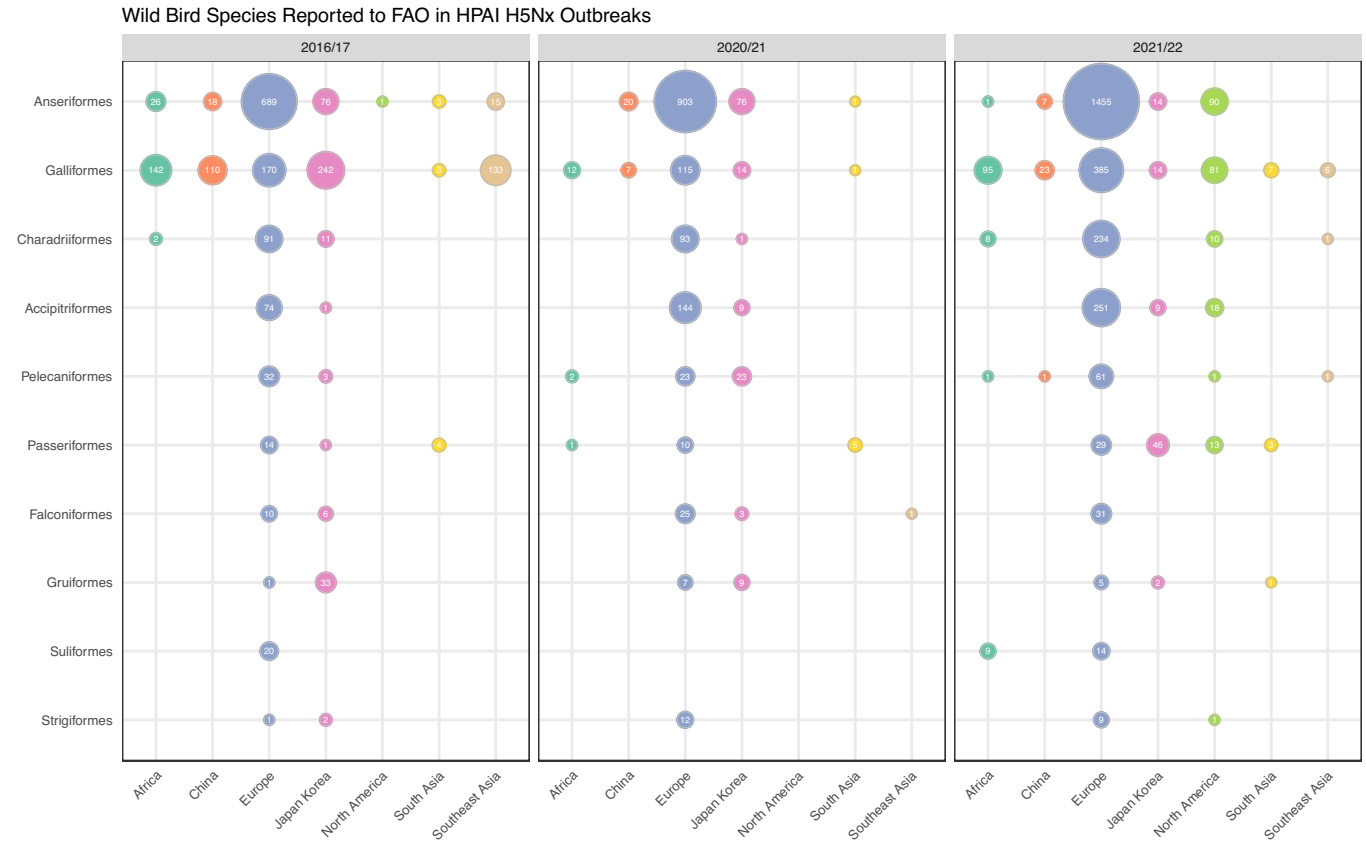
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-06631-2.

Correspondence and requests for materials should be addressed to Vijaykrishna Dhanasekaran.

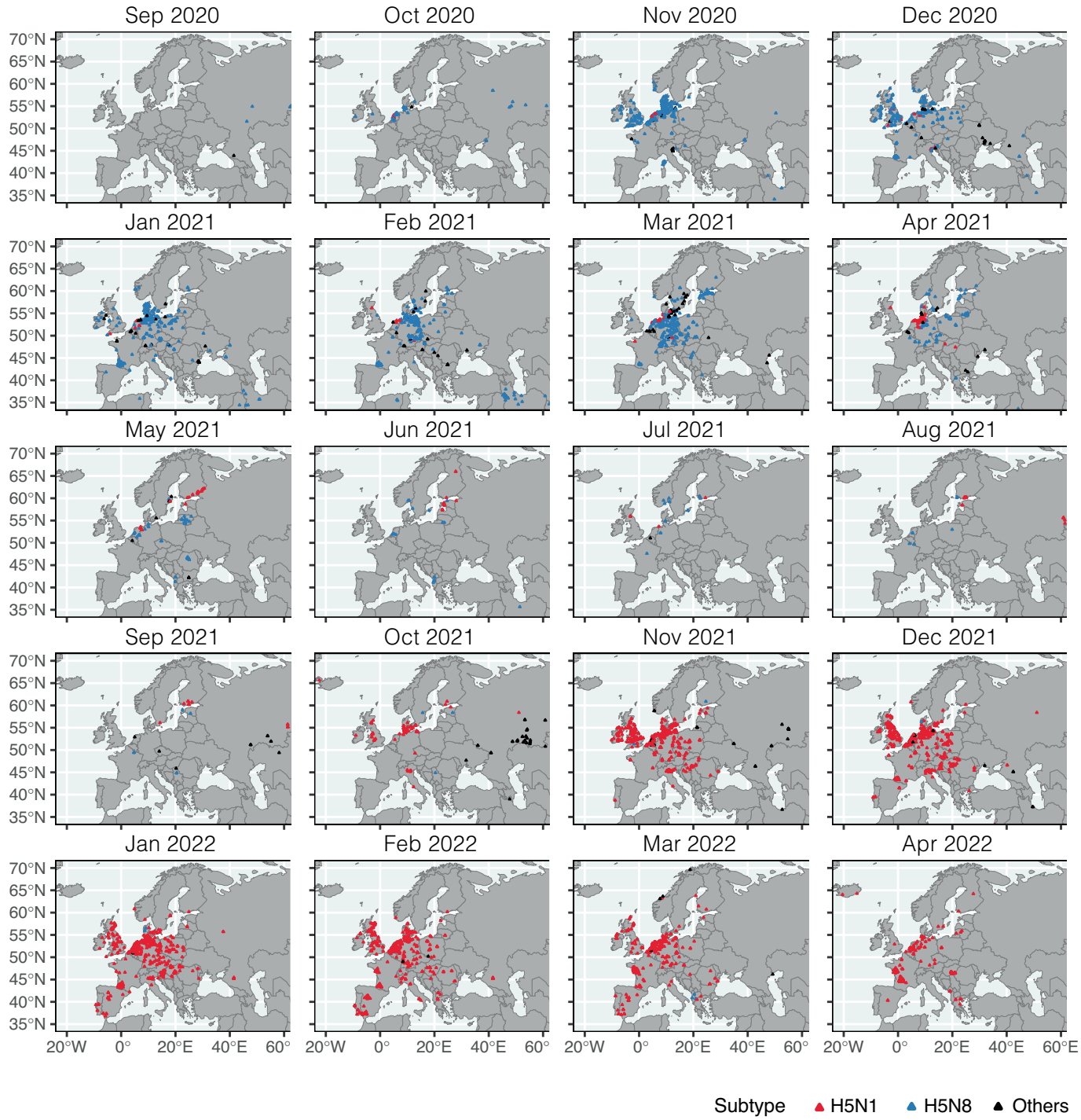
Peer review information *Nature* thanks Matthew Scotch and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



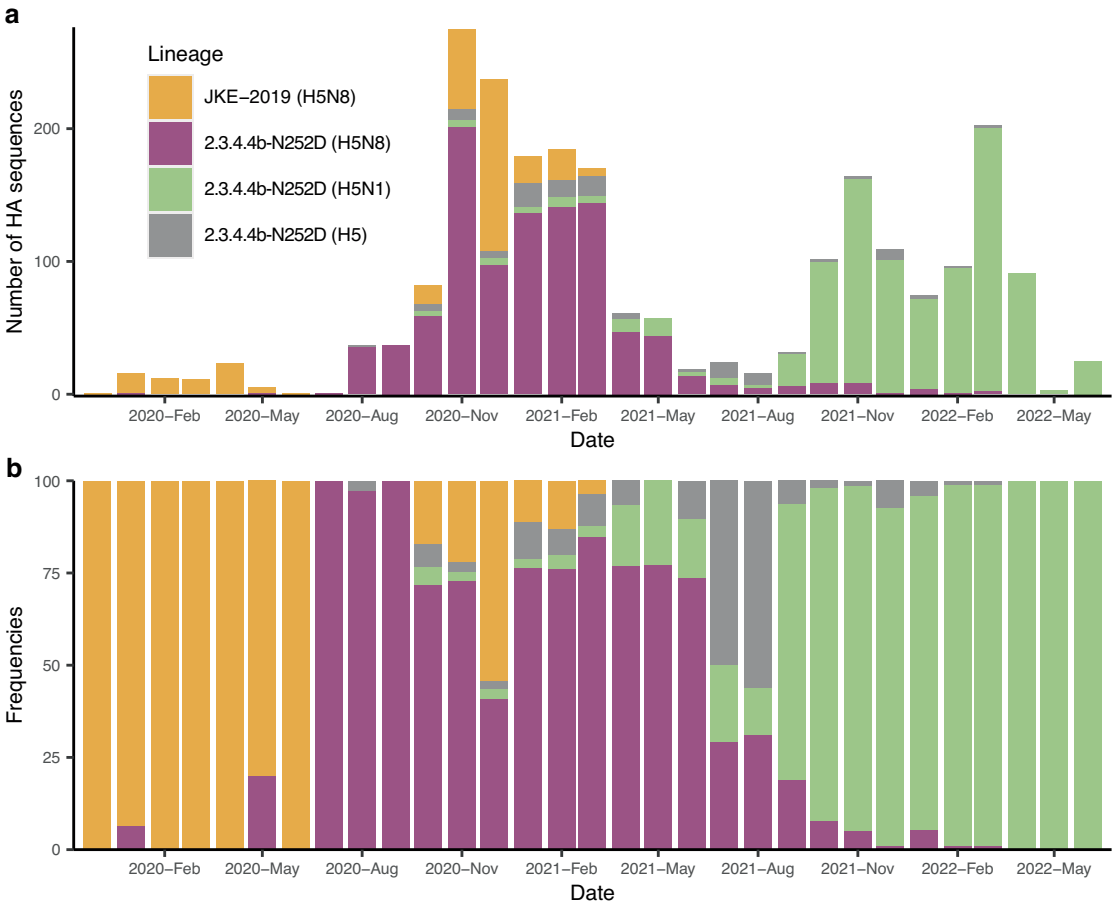
Extended Data Fig. 1 | Number and proportion of HPAI H5 outbreaks. Data according to the EMPRESS-i+ Global Animal Disease Information System, Food and Agriculture Organization of the United Nations (FAO) (https://empres-i.apps.fao.org/) from January 2004 to May 2022 coloured by subtype.



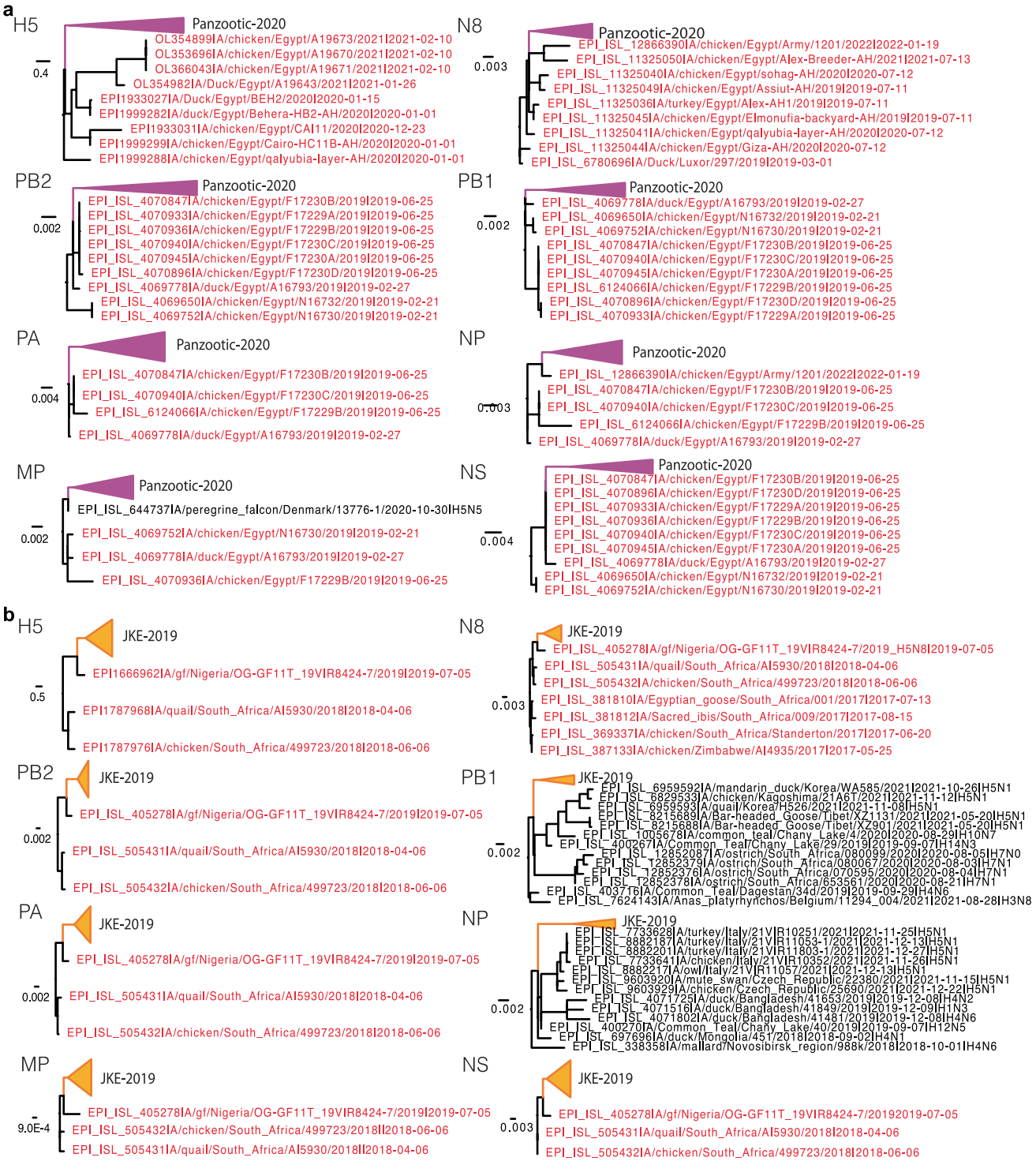
Extended Data Fig. 2 | Comparison of affected wild bird species by taxonomic order. Data reported to the EMPRESS-i+ Global Animal Disease Information System, Food and Agriculture Organization of the United Nations (FAO) (https://empres-i.apps.fao.org/) in HPAI H5Nx outbreaks between 2016/17, 2020/21, and 2021/22. Avian orders with fewer than ten outbreaks are not shown.



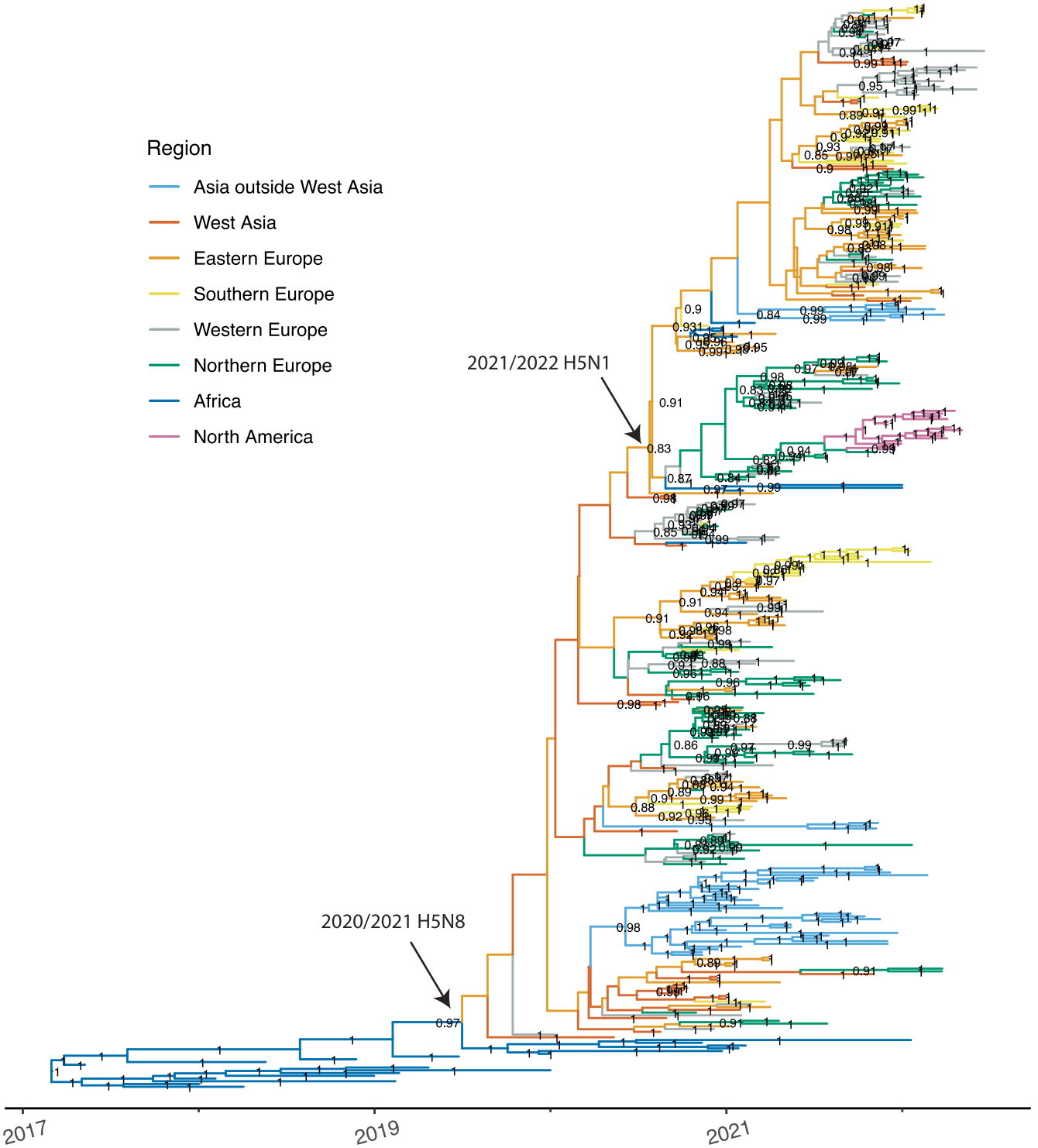
Extended Data Fig. 3 | Spatial distribution of HPAI H5 outbreaks in Europe during 2020–2022. EMPRESS-i+ Global Animal Disease Information System, Food and Agriculture Organization of the United Nations (FAO) (https://empres-i.apps.fao.org/). Outbreaks are coloured by subtype, and maps were generated using the R package "rnaturalearth".



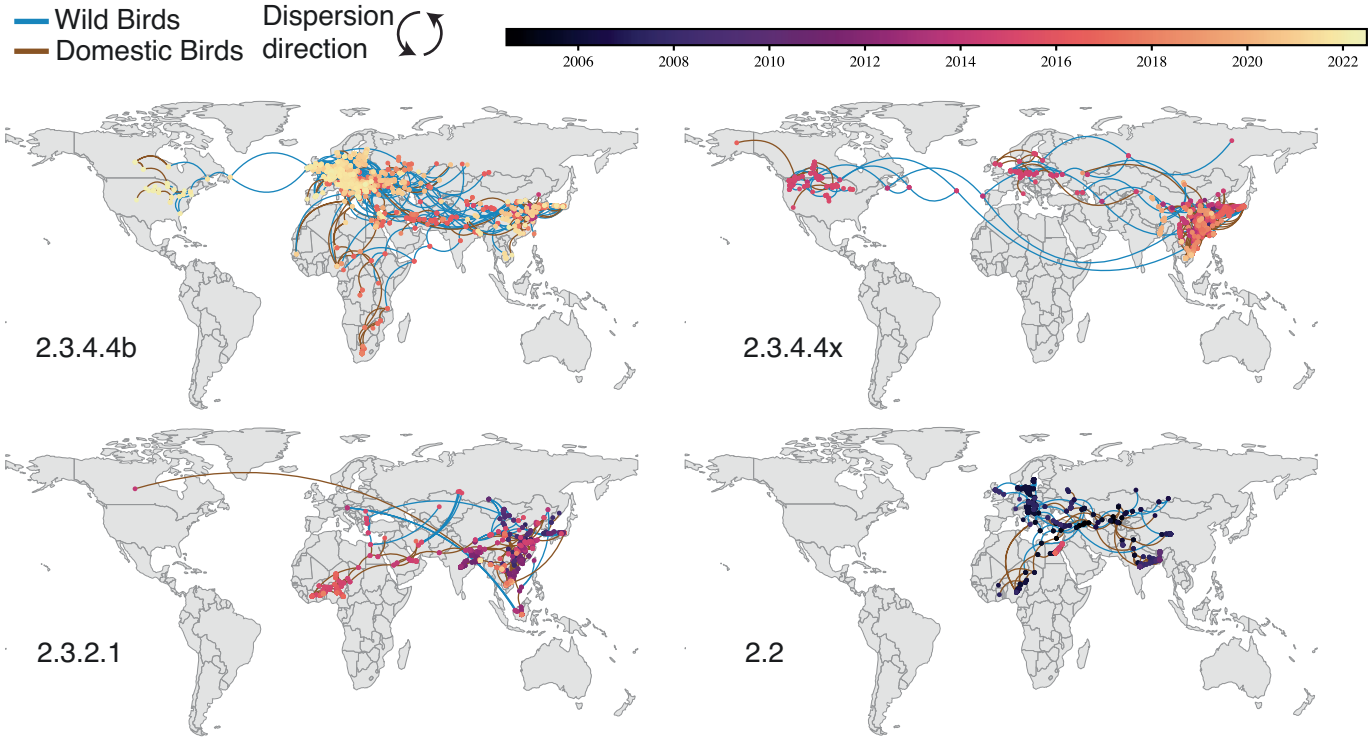
Extended Data Fig. 4 | Temporal changes in HPAI H5 lineage predominance. (**a**) The number of *HA* sequences coloured by lineage since 2020. (**b**) Proportional lineage distribution by month inferred from (**a**).



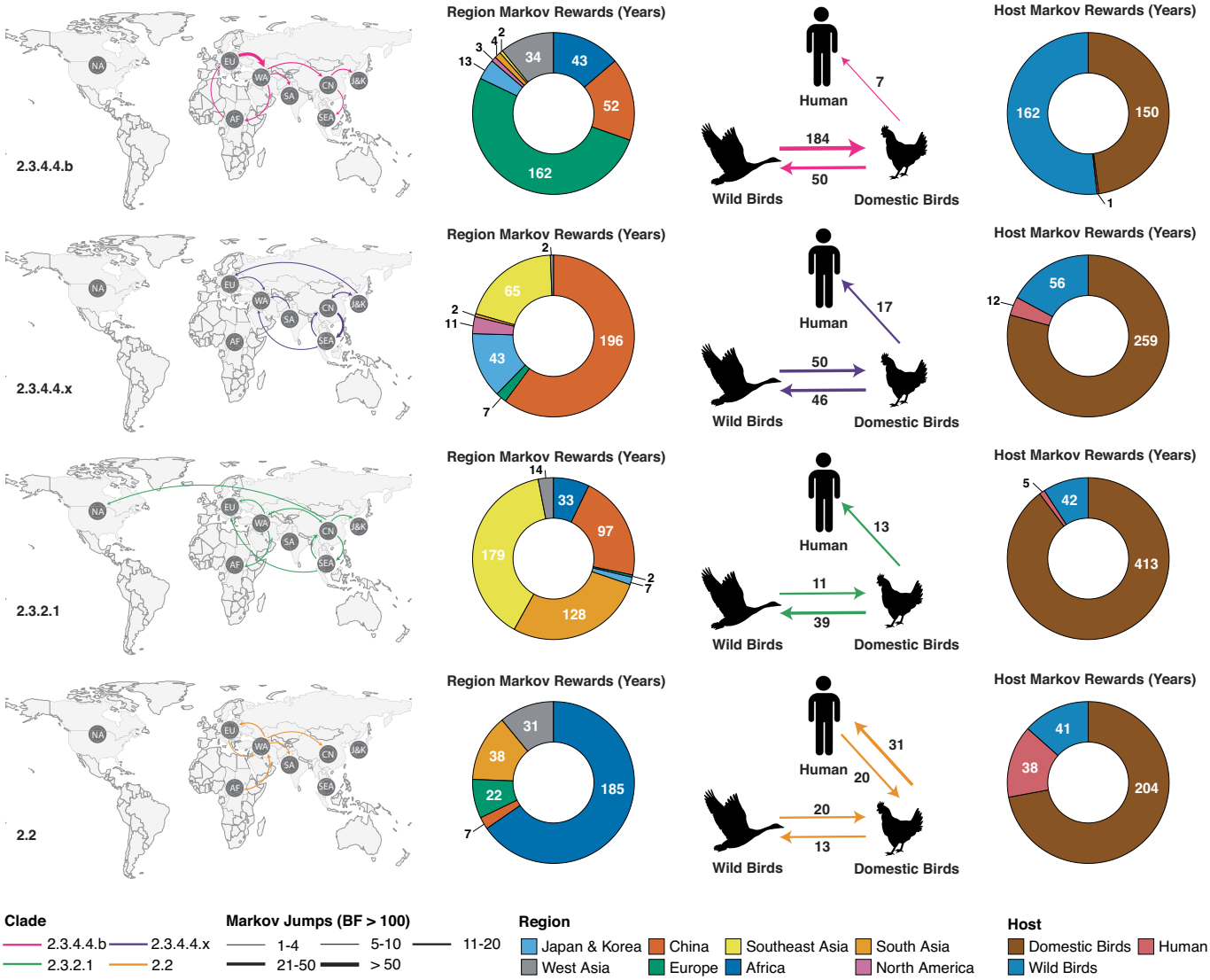
Extended Data Fig. 5 | Evolutionary relationships of Panzootic-2020 (including 2020/21 and 2021/22 resurgence) and JKE-2019 lineage. Maximum-likelihood tree of Panzootic-2020 (a) and JKE-2019 (b) for each of the eight gene segments. Samples collected in Africa are highlighted in red.



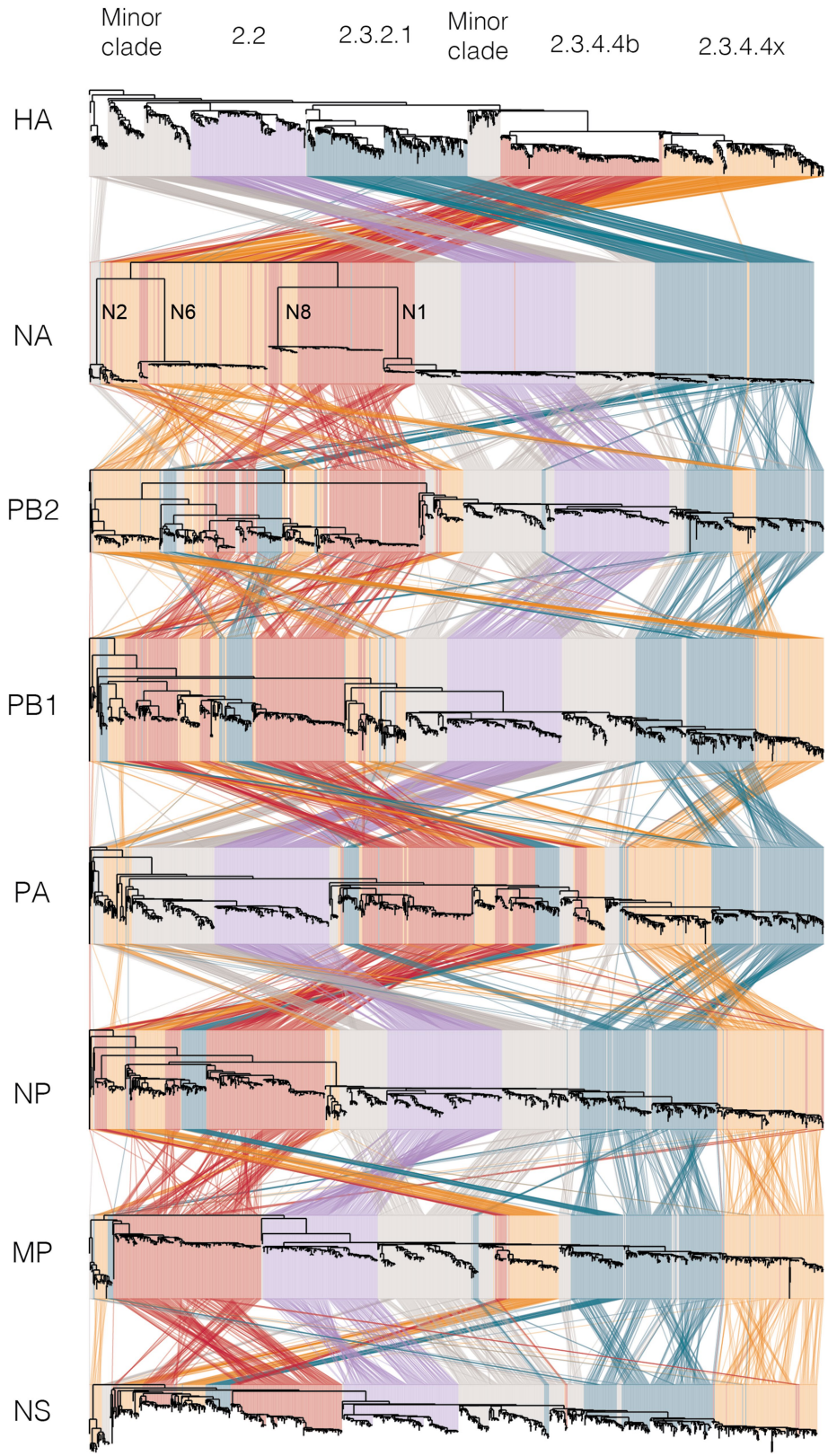
Extended Data Fig. 6 | Maximum clade credibility tree of panzootic-2020 clade with branches coloured by geographic region. The posterior probabilities of regions over 0.8 are annotated in the nodes.



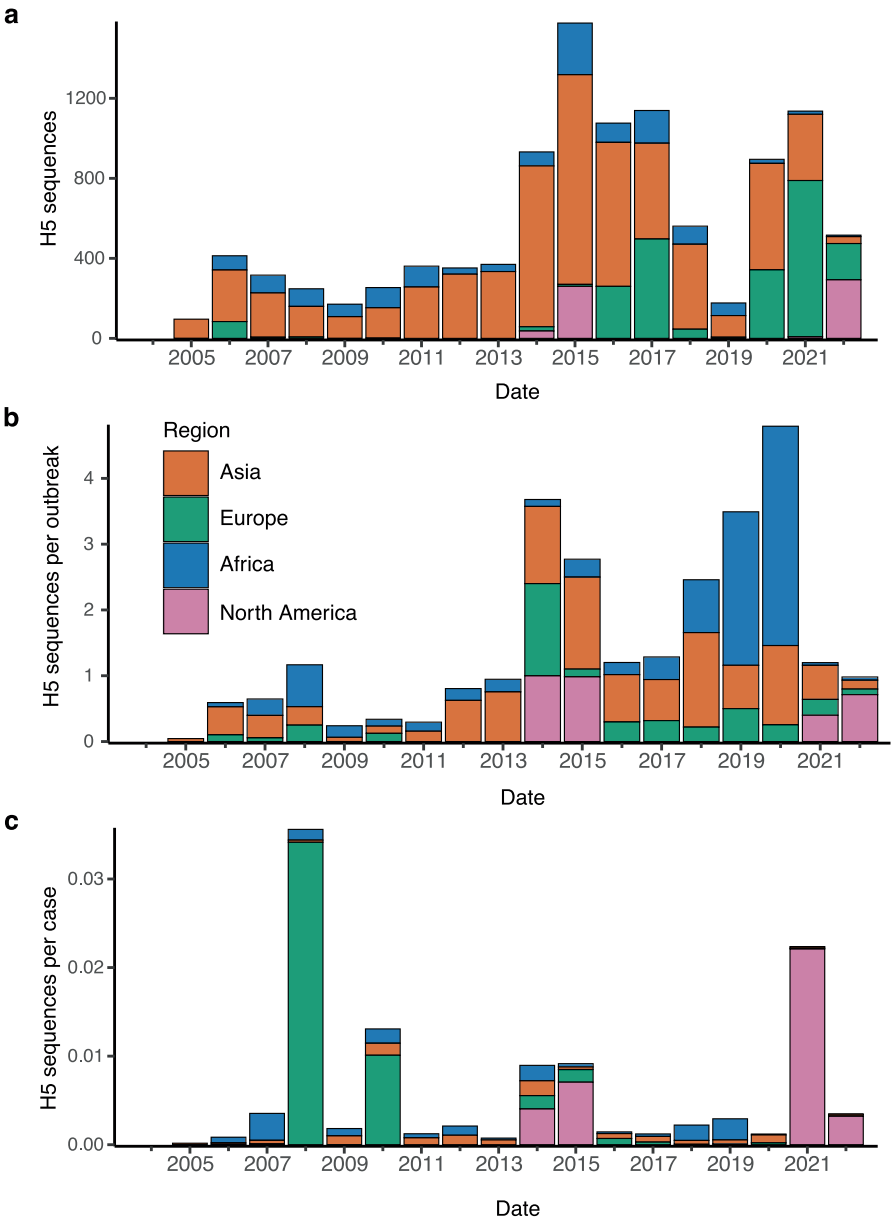
Extended Data Fig. 7 | Dynamics of HPAI H5 transmission lineages in clades 2.3.4.4b, 2.3.4.4x, 2.3.2.1 and 2.2. Virus lineage movements were inferred by continuous phylogeographic analysis for each clade.



Extended Data Fig. 8 | The contrasting geographic and host transmission patterns among HPAI H5 2.3.4.4b, 2.3.4.4x, 2.3.2.1, and 2.2 clades inferred from discrete phylogeography. From left to right, the figures represent regional Markov jumps, regional Markov rewards, host Markov jumps, and host Markov rewards.



Extended Data Fig. 9 | Tanglegram of HPAI H5 virus reassortment. Coloured lines connect each virus across all eight genes, showing incongruence between and within major clades.



Extended Data Fig. 10 | Genomic surveillance of HPAI H5 viruses by geographic region from 2005 to 2022. (**a**) Temporal changes of all HPAI H5 sequences available on the Global Initiative for Sharing All Influenza Data (GISAID) and NCBI Influenza Virus Resource databases from January 2005 to June 2022 using collection dates. (**b**) Temporal changes of HPAI H5 sequences per outbreak reported to the Food and Agriculture Organization of the United Nations (FAO). (**c**) Temporal changes of HPAI H5 sequences per case reported to the World Organization for Animal Health (WOAH).

nature portfolio

Corresponding author(s): Vijaykrishna Dhanasekaran

Last updated by author(s): Aug 17, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection. Code and software (including version numbers) used to process sequence-data generated in this study are provided in the Methods and associated linked repository https://doi.org/10.5281/zenodo.8251324.
Data analysis	All the code developed for the data analysis is available in the linked repository. All software tools used for analysis are open-source unless otherwise stated. H5 clades based on the WHO gs/Gd H5N1 nomenclature system were determined using LABEL v0.6.3. Sequence alignments were performed with MAFFT v7.490 and trimmed using trimAL v1.4. Temporal signal was explored using TempEst v1.5.3. Phylogenetic analyses were performed using using FastTree v.2.1.1, TreeTime v0.9.1, IQ-TREE v.2.1.4 and BEAST v1.10.4. Tracer v1.7.1 was used to visualise and analyse MCMC trace files from Bayesian phylogenetic analysis. SpreadD3 v0.9.693 was used to estimate Bayes factors (BF). BaTS v0.10.1 was used to investigate the degree of phylogeographic structure (phylogenetic clustering by sampling location and host). The trunk region/ host through time and persistence were measured from these posterior phylogenies using PACT v0.9.5. Reassortment analysis was performed using IQ-TREE v.2.1.4 and Phylogenetic Diversity Analyzer (PDA) tool v1.0.3 and baltic v0.1.5 (https://github.com/evogytis/baltic) was used to visualize the incongruence between phylogenetic trees of eight genes. The R package 'ggstream' v.0.1 was used to map temporal changes in the sampling of HPAI H5 clades, and 'rworldmap' v.1.3 was used to plot world maps. All the code and software used for data analysis is available in the data repository associated with the manuscript.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The genome sequences and and associated metadata used in this study are available in Global Initiative on Sharing All Influenza Data (GISAID) (https://platform.epicov.org/) and NCBI Influenza Virus Resource (https://www.ncbi.nlm.nih.gov/genomes/FLU/) databases. Accession numbers are provided in Supplementary Data 2.
Avian hosts were classified as domestic and wild birds using strain names, associated metadata, and original publications and provided in the repository linked to the manuscipt https://doi.org/10.5281/zenodo.8251324

All reported and confirmed infections of HPAI H5 viruses in humans were obtained from WHO (https://www.who.int/emergencies/disease-outbreak-news). Confirmed detections/outbreaks in domestic and wild birds globally were obtained from World Animal Health Information System (WAHIS), World Organisation for Animal Health (https://wahis.woah.org/) and EMPRESS-i+ Global Animal Disease Information System, Food and Agriculture Organisation (https://empres-i.apps.fao.org/). These data are provided in https://doi.org/10.5281/zenodo.8251324

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	In this study, no sample size calculations were performed; however all available HPAI H5 avian influenza sequences in GenBank and GISAID were used.
Data exclusions	For phylogenetic analysis duplicate sequences, laboratory-derived and mixed-subtype isolates, and sequences of low quality or <85% coverage were excluded as specified in the Methods.
Replication	For all phylogenetic anaysis, a minimum of two independent runs were performed, and their convergence and effective sample sizes were confirmed to be adequate prior to inference.
Randomization	Randomization was not required for our study as it contained no experimental groups or comparisons.
Blinding	Blinding was not relevant as publicly available surveillance data, with de-identified patient data, were utilised in the study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging