Association of poultry vaccination with the interspecies transmission and molecular evolution of H5 subtype avian influenza virus

Bingying Li[1]†, Jayna Raghwani[2]†, Sarah C. Hill[2], Sarah François[3,4], Noémie Lefrancq[5], Yilin Liang[1], Zengmiao Wang[1], Lu Dong[6], Phillipe Lemey[7], Oliver G. Pybus[2,3]*, Huaiyu Tian[1]*

[1] State Key Laboratory of Remote Sensing Science, Center for Global Change and Public Health, Beijing Normal University, Beijing, China

[2] Department of Pathobiology and Population Sciences, The Royal Veterinary College, London, United Kingdom

[3] Department of Biology, University of Oxford, Oxford, United Kingdom

[4] UMR DGIMI, University of Montpellier, INRAE, Montpellier, France

[5] Department of Genetics, University of Cambridge, Cambridge, UK

[6] Ministry of Education Key Laboratory for Biodiversity and Ecological Engineering, College of Life Sciences, Beijing Normal University, Beijing, China

[7] Department of Microbiology, Immunology and Transplantation, Rega Institute, Clinical and Epidemiological Virology, KU Leuven, 3000 Leuven, Belgium

†These authors contributed equally to this work.

*Corresponding authors: tianhuaiyu@gmail.com (H.Y.); opybus@rvc.ac.uk (O.G.P.).

Abstract

The effectiveness of vaccinating poultry in preventing the transmission of highly pathogenic avian influenza viruses (AIVs) has been questioned for years and its impact on wild birds is uncertain [1-3]. Here we reconstruct movements of H5 subtype AIV lineages among vaccinated poultry, unvaccinated poultry, and wild birds, worldwide from 1996 to 2023. We find that lineage transitions among host types are lagged and that movements from wild birds to unvaccinated poultry were more frequent than those from wild birds to vaccinated poultry. However, we also find that the HA gene of the AIV lineage that circulated predominately among Chinese poultry with high vaccination coverage underwent faster evolution and greater nonsynonymous divergence than other lineages. Further, this Chinese poultry lineage contained more codons inferred to be under positive selection, including at known antigenic sites, and its rates of nonsynonymous divergence and adaptative fixation increased after mass poultry vaccination began. Our results indicate that the epidemiological, ecological and evolutionary consequences of widespread AIV vaccination in poultry may be linked in complex ways, and that much work is needed to better understand how such interventions may affect AIV transmission to, within and from wild birds.

Key words: interspecies virus transmission, avian influenza, poultry vaccination, wild birds, H5 subtype avian influenza viruses

Introduction

During the summer of 2022, seabirds in many European, North American, and African countries suffered unprecedented mortality from avian influenza virus (AIV) [4]. The causative virus, a highly pathogenic H5 subtype AIV (HPAIV) belonging to clade 2.3.4.4b, has been detected at unprecedented incidence in wild birds [5]. Wild birds, the natural reservoir of AIVs, can acquire and transmit viruses to poultry or mammals [6,7] and play a major role in the maintenance and global dissemination of AIVs [8]. Intra- and inter-species transmission of genetically diverse AIVs can result in virus genomic reassortment [9,10] and the emergence of novel HPAIV lineages [11-19].

In order to protect poultry from infection with HPAIV, some countries have implemented vaccination programs in poultry for H5 subtype avian influenza, mostly in Asia and Africa (Fig. 1) [20-23]. Based on national vaccination data from 2010, Egypt had the highest vaccination coverage in poultry (82%), followed by China (73% in 2009 and 87% in chickens and < 30% in ducks in 2018 [24]), Vietnam (31%), and Indonesia (12%) [25]. In 2013, vaccination coverage for commercial layer flocks in two districts of Bangladesh were reported to be 32% and 54% [26]. France and the Netherlands have also vaccinated poultry; however, their overall vaccination coverages are very low (<0.1%). Since 2005 China has implemented a nationwide vaccination program [27] and notably accounts for >90% of the global consumption of H5 AIV vaccines. Due to the continuing evolution of H5 AIV, vaccine strains are

63 frequently updated to ensure their effectiveness [28,29]. Several studies have suggested
64 that the extensive vaccination of Chinese poultry against H5 AIV has suppressed
65 outbreaks effectively and substantially decreased the prevalence of H5 AIV in live
66 bird markets [26,29-32].
67
68 However, there are concern that mass vaccination against H5 AIV could impact the
69 molecular evolution of the virus [33]. For example, after mass vaccination was initiated
70 against H5N1 AIV in China, a significant increase in the evolutionary rate of H5N1
71 AIV was observed during 2005-2010 [34]. Similarly, in North America, the evolutionary
72 rate of H5N2 AIV in Mexico was inferred to be significantly higher between 1993 and
73 2002, following a period of mass avian influenza vaccination in this region, compared
74 to the rate of evolution of H5 AIV in the United States, where vaccination was not
75 used [35].
76
77 Furthermore, AIV lineages circulating in Egypt and Indonesia, where vaccination
78 against H5N1 is prevalent, are characterized by higher evolutionary rates and a greater
79 number of positively selected sites in the HA gene compared to countries where
80 vaccination is about, such as Nigeria, Turkey and Thailand [36,37]. However, these
81 studies have concentrated mainly on viruses from poultry and lacked data from wild
82 birds and at the wild bird-poultry interface. This limitation could potentially create
83 confusion in understanding the evolutionary characteristics of the virus in both wild

84 birds and poultry, contributing to the uncertainty of their conclusions. Given the
85 frequent movement of AIV between wild birds and poultry, it is crucial to understand
86 the impact of mass poultry vaccination on AIV in unvaccinated wild birds.
87
88 Here we conduct phylogenetic analyses to investigate the inter-species transmission of
89 H5 AIV between wild birds and poultry from 1996 to 2023, and compare the
90 evolutionary dynamics of H5 AIV in different host populations that vary in
91 vaccination status.
92
93 **Result**
94 **Interspecies transmission of H5 AIV at the interface of poultry and wild birds**
95 We collated a total of 22,606 hemagglutinin (HA) gene sequences belonging to H5
96 AIV, sampled from each continent since 1996 (Fig. 1). The viral sequences were
97 unevenly distributed over time and space across host populations. Therefore, for the
98 downstream analyses, we focused on Europe and Asia, where sufficient viral genetic
99 data was available from long-term sampling of poultry and wild birds to quantify the
100 dynamics of virus transmission and evolution within and among host populations. By
101 taking into account the implementation of vaccination programs and the availability of
102 viral genetic sequences (Supplementary Fig. S1), we categorised sequences into eight
103 groups based on the country of sampling, host species, and vaccination status: wild
104 birds (without vaccination), European poultry (without vaccination), Japanese poultry

(without vaccination), Korean poultry (without vaccination), Indonesian poultry (with vaccination, low vaccination coverage), Bangladeshi poultry (with vaccination, low vaccination coverage), Vietnamese poultry (with vaccination, low vaccination coverage), and Chinese poultry (with vaccination, high vaccination coverage).

To investigate H5 AIV transmission and evolution, we estimated a time-resolved phylogeny and inferred ancestral state (Fig. 2A; Supplementary Table S1). Inter-species movement of H5 AIV lineages from Chinese poultry to wild birds, and from wild birds to European poultry, were frequently observed (Fig. 2B). Bayesian reconstruction of the host status of H5 AIV lineages indicates there were three waves of spillovers from Chinese poultry to wild birds (CPtoWB), and from wild birds to European poultry (WBtoEP) (Fig. 2C). To investigate lineage transitions among these three host types before 2020, we conducted an extended convergent cross mapping (CCM) analysis on the time series of mean annual lineage transitions (Markov jumps) from Chinese poultry to wild birds, and from wild birds to European poultry, between 1996 and 2019. We find that CPtoWB transitions preceded WBtoEP transitions, with a time lag of approximately 2-3 years (Fig. 2C; Supplementary Fig. S2). The observed pattern is robust to diverse genomic sampling strategies (Supplementary Figs. S3-S6).

Overall, more viral lineage transitions were observed from wild birds to unvaccinated poultry populations than to vaccinated poultry populations (Fig. 3, Supplementary

126 Table S2). Frequent lineage transitions were observed from wild birds to unvaccinated
127 European, Japanese and Korean poultry, especially after 2020 (Fig. 3A). Conversely,
128 there were fewer viral lineage transitions from wild birds to Chinese poultry and other
129 vaccinated poultry populations (Fig. 3B). Inter-species transitions from poultry to
130 wild birds were frequent after 2016 (Fig. 3C and Fig. 3D). Virus lineage movements
131 between poultry populations were not common. The observed pattern is robust to
132 diverse genomic sampling strategies (Supplementary Table S3).
133
134 **Evolutionary dynamics of H5 AIV in different host populations**
135 We next investigated the evolutionary dynamics of H5 AIV in wild birds and poultry
136 populations with different vaccination levels. Poultry-dominated viral lineages were
137 identified only for China (vaccinated since 2005), Bangladesh (vaccinated since 2012)
138 and Indonesia (vaccinated since 2004) (Fig. 4A). Our results indicated that the
139 Chinese poultry lineage had a significantly higher substitution rate (mean rate =
140 5.38×10^{-3} sub/site/year; 95% HPD: $5.02\text{-}5.76\times10^{-3}$; $P < 0.05$) than the early-wild bird
141 lineage (3.39×10^{-3} sub/site/year; 95% HPD: $2.97\text{-}3.83\times10^{-3}$) (Fig. 4B). Notably this
142 result is robust to the time-dependence of virus evolutionary rates [38] because both
143 lineages have similarly high variation in sequence sampling dates (Fig. 4C). Besides,
144 the substitution rate of the Chinese poultry lineage during the vaccination period
145 (5.12×10^{-3} sub/site/year; 95% HPD: $4.71\text{-}5.55\times10^{-3}$) was faster than before
146 vaccination (4.79×10^{-3} sub/site/year; 95% HPD: $4.38\text{-}5.18\times10^{-3}$; $P < 0.05$;

Supplementary Fig. S7). The substitution rates of poultry-dominated viral lineages in Bangladesh and Indonesia were also faster than those of the early-wild bird lineage ($P < 0.05$). Furthermore, the substitution rate of the late-wild bird lineage was faster than that of the Chinese poultry lineage from which it emerged (5.91×10^{-3} sub/site/year; 95% HPD: $5.03\text{-}6.56\times10^{-3}$), however this latter result should be interpreted with caution due to the comparatively shorter range of sequence sampling dates in the late-wild bird lineage.

If the higher evolutionary rate of HA gene in vaccinated poultry lineages is simply a result of increased viral transmission facilitated by high poultry densities, then we should also observe higher rates for poultry lineages in other AIV segments. Conversely, PB2 gene evolution was faster in the wild bird lineage (3.62×10^{-3} sub/site/year; 95% HPD: $3.23\text{-}3.92\times10^{-3}$; $P < 0.05$) than in the Chinese poultry lineage (3.44×10^{-3} sub/site/year; 95% HPD: $3.08\text{-}3.78\times10^{-3}$; Supplementary Fig. S8-S9). This result suggests that the faster evolution of the HA gene in Chinese poultry may not be attributable to higher transmission rates alone and instead could result from selection on the HA gene.

Viral adaptive evolution in host populations with different vaccination status

We then investigated whether the higher rate of molecular evolution observed in the Chinese poultry lineage, compared to the early-wild bird lineage, could be explained

168 by greater viral adaptive evolution in poultry. For each lineage, we calculated the
169 nonsynonymous and synonymous divergence of the HA gene through time, from a
170 known reference sequence (NCBI Reference Sequence: AF144305; Figs. 5A, 5B and
171 5C). Our results indicate higher nonsynonymous divergence for the Chinese poultry
172 lineage than for other lineages, such as wild bird lineages and poultry lineages in
173 Bangladesh and Indonesia (Chinese poultry lineage gradient = 0.0027, $P < 0.05$;
174 early-wild bird lineage = 0.0005, $P < 0.05$; late-wild bird lineage = 0.0007, $P < 0.05$;
175 Bangladeshi poultry lineage = 0.0009, $P < 0.05$; Indonesian poultry lineage I =
176 0.0011, $P < 0.05$; Indonesian poultry lineage II = 0.0021, $P < 0.05$). Additionally, the
177 nonsynonymous divergence rate increased in the Chinese poultry lineage following
178 mass vaccination (gradient pre-2005 = 0.0017, 95% CI = 0.0011-0.0023; gradient
179 between 2005 and 2010 = 0.0046, 95% CI = 0.0041-0.0052; Supplementary Table
180 S4). To test for the potential impact of different sampling intensities among countries,
181 the Chinese poultry lineage was resampled and the results remained consistent
182 (Supplementary Fig. S10). This indicates that the Chinese poultry lineage exhibited
183 more nonsynonymous evolution, especially after vaccination, compared to the wild
184 bird lineages and the poultry lineages that experienced lower vaccination coverage.
185
186 Since nonsynonymous divergence can result from either positive selection or random
187 genetic drift, we also used an independent population genetic method to estimate the
188 rate of accumulation of adaptive substitutions in the HA gene of the Chinese poultry

lineage. The estimated rate of adaptative fixation in that lineage was significantly higher after the initiation of vaccination in poultry (gradient pre-2005 = 0.24 adaptive fixations per codon per year; $P<0.05$; gradient after 2005 = 0.78 adaptive fixations per codon per year; $P < 0.05$; Fig. 5D). These estimated adaptation fixation rates are rapid, but lower than those previously inferred for the HA gene of human influenza subtype H3N2 (1.52) and subtype H1N1 (1.02) [39]. We next estimated dN/dS values for each host-associated lineage using the renaissance counting method [40]. Mean dN/dS is higher in the Chinese poultry lineage (0.24) than in the early wild bird lineage (0.16), the late-wild bird lineage (0.18), and the European poultry lineage (0.16) (Supplementary Table S5). The mean dN/dS values of the Bangladeshi (0.17) and Indonesian poultry I and II lineages (0.18 for both) were also lower than that of the Chinese poultry lineage.

Furthermore, dN/dS values were calculated for individual codons in the HA gene, for each host-specific H5 AIV lineage. We employed multiple methods to detect sites under positive selection (Table 1). Our results indicate the number of amino acid sites exhibiting evidence of positive selection in the Chinese poultry lineage was greater than in the two wild bird lineages and in the Bangladeshi and Indonesian poultry lineages (Table 1). More positively selected sites were also observed in resampled-Chinese poultry lineage (Supplementary Table S6), indicating the finding is robust to diverse sampling strategies (Supplementary Tables S7, S8).

The majority of the positively selected sites were associated with established immune-reactive epitopes. In the Chinese poultry lineage, those were mainly located in the receptor binding subdomain (Supplementary Fig. S11). For example, positive selection was observed at sites 156, 157, and 242, which are recognized as CD8+ T-cell epitopes [41], and at site 87, which is associated with antibody epitope 65C6 [42]. Within the late-wild bird lineage, positively selected site 252 is associated with the $H5_{246-260}$ epitope, which induced activation of T cells in chickens immunized against the HA antigen of H5 AIV [43]. The positive selection analysis also revealed mutations (mostly in the receptor binding subdomain) that are associated with changes within the Chinese poultry lineage (D142E, H154Q /L/N, R156V/T/M/K/N/A, S157P/A, S171D/N, Q185R/K/S/G, L285V/M/I) and with the transition of the virus from the Chinese poultry lineage to the late-wild bird lineage (E142, Q154, A/T156, P157, D171, R185, V285). The amino acid states of these sites in the early-wild bird lineage (D142, N154, R156, S157, N171, Q185, L285) are mostly different to those in the late-wild bird lineage, indicating that these changes are not reversion mutations.

Discussion

Since 2020, H5 subtype avian influenza viruses have caused outbreaks in European and Asian countries, posing a significant real threat to the poultry industry and a potential threat to public health. We reconstructed the inter-species transmission

231 history of H5 subtype avian influenza viruses among wild birds and poultry
232 populations with different vaccination levels. Our analysis reveals a shift from a
233 lineage circulating within Chinese poultry to one circulating among wild birds. The
234 wild bird lineage has frequently transmitted to unvaccinated European poultry, while
235 the spillback of this virus from wild birds to vaccinated poultry appears to be
236 impeded. Further, the virus lineage circulating in highly-vaccinated Chinese poultry
237 exhibits evidence of more non-synonymous and adaptive molecular evolution in the
238 HA gene after the date of introduction of mass poultry vaccination. The Chinese
239 poultry lineage may have experienced more vaccine-driven selection than the other
240 lineages. Further research is needed to determine if this selection has had any impact
241 on the HA gene mutations present in the late-wild bird lineage.
242
243 Due to high virus prevalence, Chinese poultry, as well as Southeast Asian poultry,
244 have been regarded as the primary reservoirs of H5 HPAIV [44]. H5 AIV spread to other
245 regions through bird migration [45] and the poultry trade [46], and our results reveal a 2-3-
246 year delay between the peak of lineage dissemination from Chinese poultry and the
247 peak of lineage introduction into European poultry. Although it has been suggested
248 that the intercontinental spread of H5 AIV may occur within a single avian migratory
249 cycle [8,45,47-50], the delay we observe may be attributed in part to pre-existing immunity
250 in wild birds, which may provide partial protection against infection and disease [51-53],
251 leading to reduced circulation and potentially dampening large-scale outbreaks in wild

252 birds. However, the duration of protection conferred by previous AIV infection and its
253 impact on the epidemiology of AIV have yet to be elucidated fully [54]. Furthermore,
254 the relatively short generation length of many wild birds (~ 3 years) may result in
255 higher turnover [55] of serologically-naïve wild birds in nature. As the reservoir of H5
256 AIV shifted from Chinese poultry to wild birds, frequent migration and large-scale
257 spatial distribution of wild birds likely facilitated inter-species transmission between
258 wild birds and poultry populations [45]. It is crucial that countries and regions enhance
259 regular surveillance of avian influenza viruses in wild birds and actively and closely
260 monitor the dynamics of virus transmission.
261
262 Following the mass poultry vaccination strategy implemented in China since 2005,
263 the spread of H5 AIV there has been relatively well-controlled. Inter-species
264 transmission of these viruses from or to Chinese poultry seems to be limited.
265 However, the Chinese poultry lineage may have experienced more antigenic evolution
266 compared to other lineages. Mutations at amino acid positions 136, 142, 157, 172,
267 201, and 205 in H5 AIV HA gene have been shown to reduce reactivity to specific
268 antibodies [56]. Amino acids at positions 142, 172, and 205 also appear to function as
269 immunodominant epitopes in H5 viruses [56]. We detected positive selection at positions
270 142, 157, 172, and 205 in the Chinese poultry lineage but not in the wild bird lineages
271 or the other poultry lineages. The H5N6 virus currently circulating in Chinese poultry
272 exhibits antigenic divergence from the strains included in the commercial vaccine in

China [57,58], potentially leading to reduced vaccine effectiveness. Our study indicates that when vaccination is used, regular monitoring and refinement of vaccines to target emerging escape variants is necessary to respond to the emergence of novel viruses. Whilst contemporary H5 HPAIVs are considered unlikely to acquire the ability to infect and stably circulate among the human population [59], there is still an urgent need to control the spread of the virus among wild birds, not only for the preservation of wildlife but also for ensuring the safety of poultry [1].

Our study has several limitations. First, although we obtained robust results supported by different datasets, the heterogeneous sampling rates of infected wild birds and poultry may bias ancestral reconstructions. Secondly, although there is a high vaccination rate among Chinese poultry [27], extracting viral lineages circulating exclusively in vaccinated poultry was not possible due to the unknown vaccination status of all sampled hosts. Thirdly, the accurate identification of ancestral host states in AIV lineages is challenging for countries with limited AIV surveillance in wild bird populations. Finally, the lack of viral genetic and surveillance data from many vaccinating/non-vaccinating countries may preclude comparative analysis of evolutionary dynamics among poultry lineages with different vaccination levels.

In conclusion, we find that vaccination in Asian poultry likely reduced the inter-species transmission of these viruses. H5 AIV in Chinese poultry, which are highly

vaccinated, show evidence of greater HA gene molecular evolution and adaptation after the introduction of vaccination. Such circumstances may have increased the probability that birds susceptible to AIV belong to wild species at the interface between wild birds and poultry, leading to shifts in selection pressure on the virus. As avian influenza continues to pose significant challenges to wild and domestic animal health, our research can help inform the development of preventive measures against AIV, such as global vaccination policies.

Method

Sequence data

We collected publicly-available hemagglutinin (HA) gene sequences of H5 AIV sampled in Asia and Europe from January 1996 to February 2023 from Global Initiative on Sharing All Influenza Data (GISAID). Only sequences with available information on date and sampling location were retained for further analysis. We aligned the sequences using MAFFT v7.487 [60] and recombinant sequences were detected using RDP4 [61]. We identified and removed sequences with unexpectedly high or low levels of genetic divergence given their sampling time from our datasets by estimating a maximum likelihood tree using FastTree v2.1.11 [62] under the automatically determined best-fit substitution model and performing a root-to-tip regression analysis in TempEst v1.5.3 [63].

Based on the sustained sampling efforts and the appropriate viral sample size (> 500 sequences; Supplementary Fig. S1), six Asian countries (China, Bangladesh, Japan, Korea, Indonesia, and Vietnam) were selected. All European countries were collectively treated as a single group. To focus on the viral lineages circulating in wild birds and poultry, the sequences were categorized into eight groups based on both geographic information and host type. These groups include wild birds, European poultry (without vaccination against H5 AIV) , Japanese poultry (without vaccination against H5 AIV), Korean poultry (without vaccination against H5 AIV), Bangladeshi poultry (with vaccination against H5 AIV; low vaccination coverage), Indonesian poultry (with vaccination against H5 AIV; low vaccination coverage), Vietnamese poultry (with vaccination against H5 AIV; low vaccination coverage), and Chinese poultry (with vaccination against H5 AIV; high vaccination coverage).

We then downsampled these datasets in a stratified manner to create a more equitable distribution of AIV sequences between wild birds and poultry: (1) wild bird dataset: randomly selected at most 2 sequences per month per country outside China and per month per province in China, comprising 1087 gene sequences from January 1999 to January 2023; (2) European poultry dataset: randomly selecting at most 1 sequence per month per country, including 338 gene sequences from January 1997 to January 2023; (3) Japanese poultry dataset: randomly selected at most 2 sequences per month, including 72 HA gene sequences from January 2000 to January 2023; (4) Korean

poultry dataset: randomly selected at most 2 sequences per month, including 72 gene sequences from October 2008 to October 2022; (5) Bangladeshi poultry dataset: randomly selected at most 1 sequence per month, including 106 HA gene sequences from January 2007 to August 2022; (6) Indonesian poultry dataset: randomly selected at most 1 sequence per month, including 121 HA gene sequences from January 2003 to March 2022; (7) Vietnamese poultry dataset: randomly selected at most 1 sequence per month, including 151 HA gene sequences from 2003 to December 2021; (8) Chinese poultry dataset: randomly selecting at most 1 sequence per month per province, including 462 gene sequences from January 1996 to March 2022.

Considering that also substantial mutations have accumulated in the PB2 gene [64,65], we used a similar method to collate PB2 gene sequences of H5 AIV sampled in Asian and Europe for sensitivity analysis (see Supplementary Materials).

Phylogenetic inference

Evolutionary histories were estimated with the Bayesian phylogenetic package BEAST v1.10.4 [66], using the BEAGLE [67] library to improve computational speed. Specifically, we employed a SRD06 substitution model [45], an uncorrelated lognormal relaxed clock [45] and Gaussian Markov random field (GMRF) Bayesian Skygrid coalescent model [68]. We subsequently used an eight-state discrete trait analysis (DTA) implemented in BEAST 1.10.4 to infer ancestral node hosts on empirical distributions

of 500 time-calibrated trees sampled from the posterior tree distributions [69]. An asymmetric model was used for the host discrete trait, which allows different rates of lineage movement between each pair of host states [70]. Three independent Markov chain Monte Carlo (MCMC) runs were performed for 400 million steps and logged every 20,000 steps. The first 10% of each chain was discarded as burn-in. We confirmed the convergence of all chains in Tracer v1.7.1 [71], ensuring the ESS was >200 for all parameters. A maximum clade credibility tree was estimated using TreeAnnotator v1.10.4 and subsequently visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree) along with the R package ggtree v2.4.1 [72]. We used the BaTS 2.0 software [73] to investigate the uncertainty arising from phylogenetic error (grouped by sampling location and host), which was compared to a null hypothesis that there was no association between the phylogenetic structure and traits by performing tip randomization with 1000 replicates (Supplementary Table S9).

Evolutionary analysis within host-specific populations

To infer virus dissemination within host-specific populations, we removed AIV phylogenetic clades that were determined to represent transmissions between different species [46]. We extracted sequences for each host-specific lineage based on the aforementioned discrete trait analysis and independently estimated the temporal phylogenies using the same substitution, clock and tree models as described above.

Strong phylogenetic temporal structure was observed in all host-specific lineages (Supplementary Fig. S12).

Divergence analysis

To estimate site-specific synonymous and non-synonymous substitutions in the different host-specific lineages, we applied a renaissance counting [40] procedure implementing a codon-position specific HKY nucleotide substitution model along with an uncorrelated lognormal relaxed clock [45] and the Gaussian Markov random field (GMRF) Bayesian Skygrid coalescent model [68]. Sites under positive selection were also identified using complementary approaches. Specifically, the Fast Unconstrained Bayesian AppRoximation (FUBAR) [74], Single Likelihood Ancestor Counting (SLAC) methods [75] and Fixed Effects Likelihood (FEL) methods [75] as implemented in Hyphy v2.5 were used [76].

The synonymous and nonsynonymous divergence for each host-specific lineage were calculated as the average Hamming distance between each sequenced isolate in that lineage and a reference sequence (NCBI Reference Sequence: AF144305). The estimated divergence was calculated by dividing the total number of observed differences between isolate and reference nucleotide sequence that resulted in a substitution(nonsynonymous or synonymous) by the number of possible nucleotide mutations that could result in a substitution, weighted by kappa [77], the

399 transition/transversion rate ratio, which was inferred from host-specific analysis using
400 BEAST.
401
402 **Time series analysis**
403 The extended convergent cross mapping was applied to detect the time lags between
404 time series variables [78]. This analysis was performed using the R package rEDM
405 v1.14.0 [78].
406
407 **Estimating rates of adaptive substitution**
408 We employed an established population genetic method related to the McDonald-
409 Kreitman test [39,79] to estimate the number of adaptive substitutions per codon per year
410 in HA H5 gene from the Chinese poultry lineage. We used a consensus of HA
411 sequences from the earliest time point (sampled in 1996) as an outgroup to estimate
412 ancestral and derived site frequencies at subsequent time points. A bootstrap analysis
413 with 1,000 replicates was conducted to assess statistical uncertainty.
414
415 **ACKNOWLEDGMENTS.** We gratefully acknowledge the authors of the originating
416 and submitting laboratories for their crucial contributions to the generation and
417 sharing of genome sequences and associated metadata through GISAID. This study
418 was supported by the National Key Research and Development Program of China
419 (2022YFC2303803); National Natural Science Foundation of China (82073616,

420 82204160); Beijing Advanced Innovation Program for Land Surface Science
421 (110631111); Fundamental Research Funds for the Central Universities
422 (2233300001); BNU-FGS Global Environmental Change Program (No.2023-GC-
423 ZYTS-11); Research on Key Technologies of Plague Prevention and Control in Inner
424 Mongolia Autonomous Region (2021ZD0006). J.R., and O.G.P. were supported by
425 the UKRI GCRF One Health Poultry Hub (grant no. B/S011269/1). S.C.H. is
426 supported by a Sir Henry Wellcome Postdoctoral Fellowship from the Wellcome Trust
427 (220414/Z/20/Z) (https://wellcome.org/). PL acknowledges support from the
428 European Union's Horizon 2020 research and innovation programme (grant
429 agreement no. 725422-ReservoirDOCS) from the European Union's Horizon 2020
430 project MOOD (grant agreement no. 874850) and from the Wellcome Trust through
431 project 206298/Z/17/Z. For the purpose of open access, the author has applied a CC
432 BY public copyright licence to any Author Accepted Manuscript version arising from
433 this submission. The funders had no role in study design, data collection and analysis,
434 the decision to publish, or in preparation of the manuscript.

435

436 **DECLARATION OF INTERESTS.** The authors declare no competing interests.

437

438 **DATA AND CODE AVAILABILITY.** All original code and data have been
439 deposited at Mendeley Data (https://data.mendeley.com/drafts/mmjmhc394f). The
440 gene segment sequences are available in the GenBank

441 (www.ncbi.nlm.nih.gov/genome/viruses/variation/flu/) and GISAID
442 (platform.gisaid.org/) databases (https://data.mendeley.com/drafts/mmjmhc394f). Any
443 additional information required to reanalyze the data reported in this paper is available
444 from the lead contact upon request.
445
446 **Authors' contributions**
447 H.T. designed the study. H.T., O.G.P., and B.L. designed the analysis. B.L., J.R., and
448 P.L. conducted the analyses. B.L. and Y.L. contributed and collected data. B.L. created
449 figures. B.L., H.T., and O.G.P. wrote the initial draft. O.G.P., J.R., S.C.H., S.F., N.L.,
450 Z.W., and L.D. interpreted the data and edited the manuscript. All authors read and
451 approved the manuscript.
452
453

References

1 Kozlov, M. US will vaccinate birds against avian flu for first time. *Nature*. **618**, 220-221 (2023).

2 Kupferschmidt, K. Bird flu spread between mink is a 'warning bell'. *Science*. **379**, 316-317 (2023).

3 Cohen, J. Bird shots. *Science*. **380**, 24-27 (2023).

4 Kuiken, T. & Cromie, R. Protect wildlife from livestock diseases. *Science*. **378**, 5 (2022).

5 Chinese Science BulletinInfection, G. a. E. o. I., European Food Safety, Aznar, I., Baldinelli, F., Stoicescu, A. & Kohnle, L. Annual report on surveillance for avian influenza in poultry and wild birds in Member States of the European Union in 2021. *EFSA J.* **20**, 7554 (2022).

6 Lebarbenchon, C., Feare, C. J., Renaud, F., Thomas, F. & Gauthier-Clerc, M. Persistence of Highly Pathogenic Avian Influenza Viruses in Natural Ecosystems. *Emerg Infect Dis*. **16**, 1057-1062 (2010).

7 Alexander, D. J. An overview of the epidemiology of avian influenza. *Vaccine*. **25**, 5637-5644 (2007).

8 Olsen, B. *et al.* Global patterns of influenza A virus in wild birds. *Science*. **312**, 384-388 (2006).

9 Taubenberger, J. K. & Kash, J. C. Influenza Virus Evolution, Host Adaptation, and Pandemic Formation. *Cell Host Microbe*. **7**, 440-451 (2010).

475 10 Dugan, V. G. *et al.* The evolutionary genetics and emergence of avian
476 influenza viruses in wild birds. *PLoS Pathog*. **4**, e1000076 (2008).
477 11 Peiris, J. S. M., de Jong, M. D. & Guan, Y. Avian influenza virus (H5N1): a
478 threat to human health. *Clin Microbiol Rev*. **20**, 243-267 (2007).
479 12 Pasick, J. *et al.* Reassortant Highly Pathogenic Influenza A H5N2 Virus
480 Containing Gene Segments Related to Eurasian H5N8 in British Columbia,
481 Canada, 2014. *Sci Rep*. **5**, 9484 (2015).
482 13 Ip, H. S. *et al.* Novel Eurasian Highly Pathogenic Avian Influenza A H5
483 Viruses in Wild Birds, Washington, USA, 2014. *Emerg Infect Dis*. **21**, 886-890
484 (2015).
485 14 Bertran, K. *et al.* Lack of chicken adaptation of newly emergent Eurasian
486 H5N8 and reassortant H5N2 high pathogenicity avian influenza viruses in the
487 US is consistent with restricted poultry outbreaks in the Pacific flyway during
488 2014-2015. *Virology*. **494**, 190-197 (2016).
489 15 Li, X. Y. *et al.* Characterization of avian influenza H5N3 reassortants isolated
490 from migratory waterfowl and domestic ducks in China from 2015 to 2018.
491 *Transbound Emerg Dis*. **66**, 2605-2610 (2019).
492 16 Yeo, S. J. *et al.* Emergence of a Novel Reassortant H5N3 Avian Influenza
493 Virus in Korean Mallard Ducks in 2018. *Intervirology*. **65**, 1-16 (2022).
494 17 Gu, M. *et al.* Novel Reassortant Highly Pathogenic Avian Influenza (H5N5)
495 Viruses in Domestic Ducks, China. *Emerg Infect Dis*. **17**, 1060-1063 (2011).

18 Gu, W. *et al.* Novel H5N6 reassortants bearing the clade 2.3.4.4b HA gene of H5N8 virus have been detected in poultry and caused multiple human infections in China. *Emerg Microbes Infect*. **11**, 1174-1185 (2022).

19 Lee, Y.-J. *et al.* Novel Reassortant Influenza A(H5N8) Viruses, South Korea, 2014. *Emerg Infect Dis*. **20**, 1087-1089 (2014).

20 Kang, Y. M. *et al.* Protection of layers and breeders against homologous or heterologous HPAIv by vaccines from Korean national antigen bank. *Sci Rep*. **10** (2020).

21 Villanueva-Cabezas, J. P., Coppo, M. J. C., Durr, P. A. & McVernon, J. Vaccine efficacy against Indonesian Highly Pathogenic Avian Influenza H5N1: systematic review and meta-analysis. *Vaccine*. **35**, 4859-4869 (2017).

22 Hill, E. M. *et al.* The impact of surveillance and control on highly pathogenic avian influenza outbreaks in poultry in Dhaka division, Bangladesh. *PLoS Comput Biol*. **14** (2018).

23 Hoang, H. T. T. *et al.* Immunization with the H5N1 Recombinant Vaccine Candidate Induces High Protection in Chickens against Vietnamese Highly Pathogenic Avian Influenza Virus Strains. *Vaccines (Basel)*. **8** (2020).

24 Zeng, X. *et al.* Vaccination of poultry successfully eliminated human infection with H7N9 virus in China. *Sci China Life Sci*. **61**, 1465-1473 (2018).

25 Swayne, D. E. Impact of Vaccines and Vaccination on Global Control of Avian Influenza. *Avian Dis*. **56**, 818-828 (2012).

517 26 Parvin, R. *et al.* Controlling avian influenza virus in Bangladesh: Challenges
518 and recommendations. *Viruses*. **12**, 751 (2020).
519 27 Liu, S. *et al.* Control of avian influenza in China: Strategies and lessons.
520 *Transbound Emerg Dis*. **67**, 1463-1471 (2020).
521 28 Sun, Z., Wang, J. & Huang, Z. Assessment of China's H5N1 routine
522 vaccination strategy. *Sci Rep*. **7**, 46441 (2017).
523 29 Wu, J. *et al.* Influenza H5/H7 Virus Vaccination in Poultry and Reduction of
524 Zoonotic Infections, Guangdong Province, China, 2017-18. *Emerg Infect Dis*.
525 **25**, 116-118 (2019).
526 30 Liu, L. L. *et al.* Characterization of Clade 7.2 H5 Avian Influenza Viruses That
527 Continue To Circulate in Chickens in China. *J Virol*. **90**, 9797-9805 (2016).
528 31 Song, W. *et al.* Changes of avian influenza virus subtypes before and after
529 vaccination in live poultry in Nanchang, China from 2016 to 2019. *Microbes*
530 *Infect*. **23**, 104848 (2021).
531 32 Guo, J. *et al.* Pathogen change of avian influenza virus in the live poultry
532 market before and after vaccination of poultry in southern China. *Virol J*. **18**,
533 213 (2021).
534 33 Iwami, S., Suzuki, T. & Takeuchi, Y. Paradox of Vaccination: Is Vaccination
535 Really Effective against Avian Flu Epidemics? *PLoS One*. **4** (2009).
536 34 Wang, Z. *et al.* Increased substitution rate in H5N1 avian influenza viruses
537 during mass vaccination of poultry. *Chin Sci Bull*. **57**, 2419-2424 (2012).

538 35 Lee, C. W., Senne, D. A. & Suarez, D. L. Effect of vaccine use in the evolution
539 of Mexican lineage H5N2 avian influenza virus. *J Virol*. **78**, 8372-8381
540 (2004).
541 36 Cattoli, G. *et al.* Evidence for differing evolutionary dynamics of A/H5N1
542 viruses among countries applying or not applying avian influenza vaccination
543 in poultry. *Vaccine*. **29**, 9368-9375 (2011).
544 37 Kwon, J. H. *et al.* Genetic evolution and transmission dynamics of clade
545 2.3.2.1a highly pathogenic avian influenza A/H5N1 viruses in Bangladesh.
546 *Virus Evol*. **6**, veaa046 (2020).
547 38 Ghafari, M. *et al.* Purifying selection determines the short-term time
548 dependency of evolutionary rates in SARS-CoV-2 and pH1N1 influenza. *Mol*
549 *Biol Evol*. **39**, msac009 (2022).
550 39 Bhatt, S., Holmes, E. C. & Pybus, O. G. The genomic rate of molecular
551 adaptation of the human influenza A virus. *Mol Biol Evol*. **28**, 2443-2451
552 (2011).
553 40 Lemey, P., Minin, V. N., Bielejec, F., Kosakovsky Pond, S. L. & Suchard, M.
554 A. A counting renaissance: combining stochastic mapping and empirical Bayes
555 to quickly detect amino acid sites under positive selection. *Bioinformatics*. **28**,
556 3248-3256 (2012).
557 41 Qiu, X. *et al.* Lineage-specific epitope profiles for HPAI H5 pre-pandemic
558 vaccine selection and evaluation. *Influenza Other Respir Viruses*. **11**, 445-456

(2017).

42 Qian, M. *et al.* Unraveling of a neutralization mechanism by two human antibodies against conserved epitopes in the globular head of H5 hemagglutinin. *J Virol*. **87**, 3571-3577 (2013).

43 Haghighi, H. R., Read, L. R., Haeryfar, S. M., Behboudi, S. & Sharif, S. Identification of a dual-specific T cell epitope of the hemagglutinin antigen of an h5 avian influenza virus in chickens. *PLoS One*. **4**, e7772 (2009).

44 King, J., Harder, T., Conraths, F. J., Beer, M. & Pohlmann, A. The genetics of highly pathogenic avian influenza viruses of subtype H5 in Germany, 2006–2020. *Transbound Emerg Dis*. **68**, 1136-1150 (2021).

45 Lycett, S. J. *et al.* Role for migratory wild birds in the global spread of avian influenza H5N8. *Science*. **354**, 213-217 (2016).

46 Yang, Q. Q. *et al.* Assessing the role of live poultry trade in community-structured transmission of avian influenza in China. *Proc Natl Acad Sci U S A*. **117**, 5949-5954 (2020).

47 Harder, T. *et al.* Influenza A(H5N8) Virus Similar to Strain in Korea Causing Highly Pathogenic Avian Influenza in Germany. *Emerg Infect Dis*. **21**, 860-863 (2015).

48 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).

580 49 Beerens, N. *et al.* Novel Highly Pathogenic Avian Influenza A(H5N6) Virus in
581 the Netherlands, December 2017. *Emerg Infect Dis*. **24**, 770-773 (2018).
582 50 Wu, H. B. *et al.* Novel Reassortant Influenza A(H5N8) Viruses in Domestic
583 Ducks, Eastern China. *Emerg Infect Dis*. **20**, 1315-1318 (2014).
584 51 Caliendo, V. *et al.* Long-Term Protective Effect of Serial Infections with H5N8
585 Highly Pathogenic Avian Influenza Virus in Wild Ducks. *J Virol*. **96**, p.
586 e01233-01222 (2022).
587 52 Koethe, S. *et al.* Modulation of lethal HPAIV H5N8 clade 2.3.4.4B infection
588 in AIV pre-exposed mallards. *Emerg Microbes Infect*. **9**, 180-193 (2020).
589 53 Hill, S. C. *et al.* Antibody responses to avian influenza viruses in wild birds
590 broaden with age. *Proc R Soc Lond B Biol Sci*. **283** (2016).
591 54 Verhagen, J. H. *et al.* Long-Term Effect of Serial Infections with H13 and H16
592 Low-Pathogenic Avian Influenza Viruses in Black-Headed Gulls. *J Virol*. **89**,
593 11507-11522 (2015).
594 55 Bird, J. P. *et al.* Generation lengths of the world's birds and their implications
595 for extinction risk. *Conserv Biol*. **34**, 1252-1261 (2020).
596 56 Zhang, Y. *et al.* Key Amino Acid Residues That Determine the Antigenic
597 Properties of Highly Pathogenic H5 Influenza Viruses Bearing the Clade 2.3.
598 4.4 Hemagglutinin Gene. *Viruses*. **15**, 2249 (2023).
599 57 Jiang, W. M. *et al.* Emerging Novel Reassortant Influenza A(H5N6) Viruses in
600 Poultry and Humans, China, 2021. *Emerg Infect Dis*. **28**, 1064-1066 (2022).

601 58 Zhang, J., Ye, H., Liu, Y., Liao, M. & Qi, W. Resurgence of H5N6 avian
602 influenza virus in 2021 poses new threat to public health. *Lancet Microbe*. **3**,
603 e558-e558 (2022).
604 59 Ciminski, K., Chase, G., Schwemmle, M. & Beer, M. Advocating a watch-
605 and-prepare approach with avian influenza. *Nat Microbiol*. **8**, 1603-1605
606 (2023).
607 60 Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software
608 Version 7: Improvements in Performance and Usability. *Mol Biol Evol*. **30**,
609 772-780 (2013).
610 61 Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4:
611 Detection and analysis of recombination patterns in virus genomes. *Virus Evol*.
612 **1**, vev003 (2015).
613 62 Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2-Approximately Maximum-
614 Likelihood Trees for Large Alignments. *PLoS One*. **5** (2010).
615 63 Rambaut, A., Lam, T. T., Carvalho, L. M. & Pybus, O. G. Exploring the
616 temporal structure of heterochronous sequences using TempEst (formerly
617 Path-O-Gen). *Virus Evol*. **2**, vew007 (2016).
618 64 Escalera-Zamudio, M. *et al.* Parallel evolution in the emergence of highly
619 pathogenic avian influenza A viruses. *Nat Commun*. **11**, 5511 (2020).
620 65 He, W.-T. *et al.* Adaption and parallel evolution of human-isolated H5 avian
621 influenza viruses. *J Infect*. **80**, 630-638 (2020).

622 66 Suchard, M. A. *et al.* Bayesian phylogenetic and phylodynamic data
623 integration using BEAST 1.10. *Virus Evol*. **4**, vey016 (2018).
624 67 Ayres, D. L. *et al.* BEAGLE: an application programming interface and high-
625 performance computing library for statistical phylogenetics. *Syst Biol*. **61**, 170-
626 173 (2012).
627 68 Hill, S. C. *et al.* Wild waterfowl migration and domestic duck density shape
628 the epidemiology of highly pathogenic H5N8 influenza in the Republic of
629 Korea. *Infect Genet Evol*. **34**, 267-277 (2015).
630 69 Lemey, P. *et al.* Unifying Viral Genetics and Human Transportation Data to
631 Predict the Global Transmission Dynamics of Human Influenza H3N2. *PLoS*
632 *Pathog*. **10**, e1003932 (2014).
633 70 Avian DiseasesEdwards, C. J. *et al.* Ancient Hybridization and an Irish Origin
634 for the Modern Polar Bear Matriline. *Curr Biol*. **21**, 1251-1258 (2011).
635 71 Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior
636 Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst Biol*. **67**,
637 901-904 (2018).
638 72 Yu, G. C., Smith, D. K., Zhu, H. C., Guan, Y. & Lam, T. T. Y. GGTREE: an R
639 package for visualization and annotation of phylogenetic trees with their
640 covariates and other associated data. *Methods Ecol Evol*. **8**, 28-36 (2017).
641 73 Parker, J., Rambaut, A. & Pybus, O. G. Correlating viral phenotypes with
642 phylogeny: accounting for phylogenetic uncertainty. *Infect Genet Evol*. **8**, 239-

246 (2008).

74 Murrell, B. *et al.* FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol*. **30**, 1196-1205 (2013).

75 Kosakovsky Pond, S. L. & Frost, S. D. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol*. **22**, 1208-1222 (2005).

76 Pond, S. L. K. *et al.* HyPhy 2.5-A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol Biol Evol*. **37**, 295-299 (2020).

77 Kistler, K. E. & Bedford, T. Evidence for adaptive evolution in the receptor-binding domain of seasonal coronaviruses OC43 and 229e. *Elife*. **10**, e64509 (2021).

78 Ye, H., Deyle, E. R., Gilarranz, L. J. & Sugihara, G. Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci Rep*. **5**, 14750 (2015).

79 Raghwani, J., Bhatt, S. & Pybus, O. G. Faster adaptation in smaller populations: counterintuitive evolution of HIV during childhood infection. *PLoS Comput Biol*. **12**, e1004694 (2016).

80 Gilbert, M. *et al.* Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *PLoS Comput Biol*. **5**, 1-11 (2018).

Figures

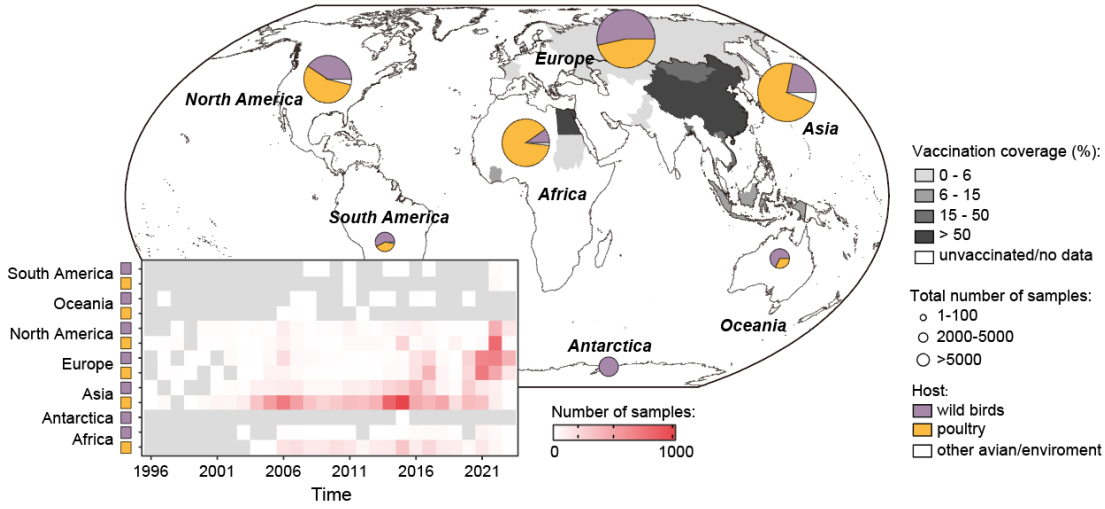


Fig 1. Vaccination coverage and number of H5 AIV HA gene sequences across different continents and countries. Countries are shaded according to the vaccination coverage in poultry in 2010. The pie charts show the total number of H5 AIV HA gene sequences sampled from poultry, wild birds, and environment since 1996. Viral sequences are categorized according to their host: wild birds (purple), poultry (yellow), and other avian/environment (white). Inset: number of H5 AIV HA gene sequences sampled in poultry and wild birds, per year, per continent.

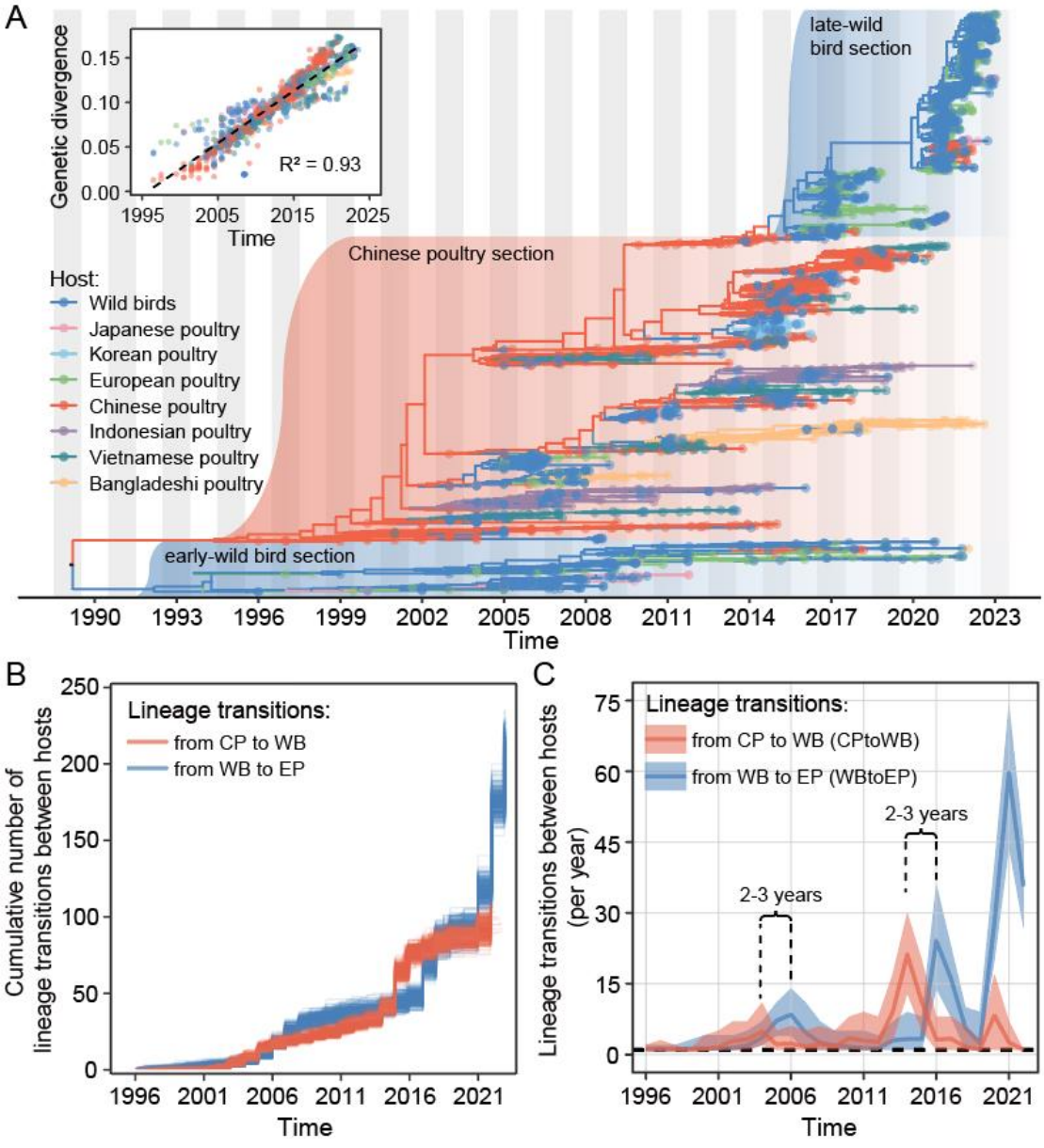
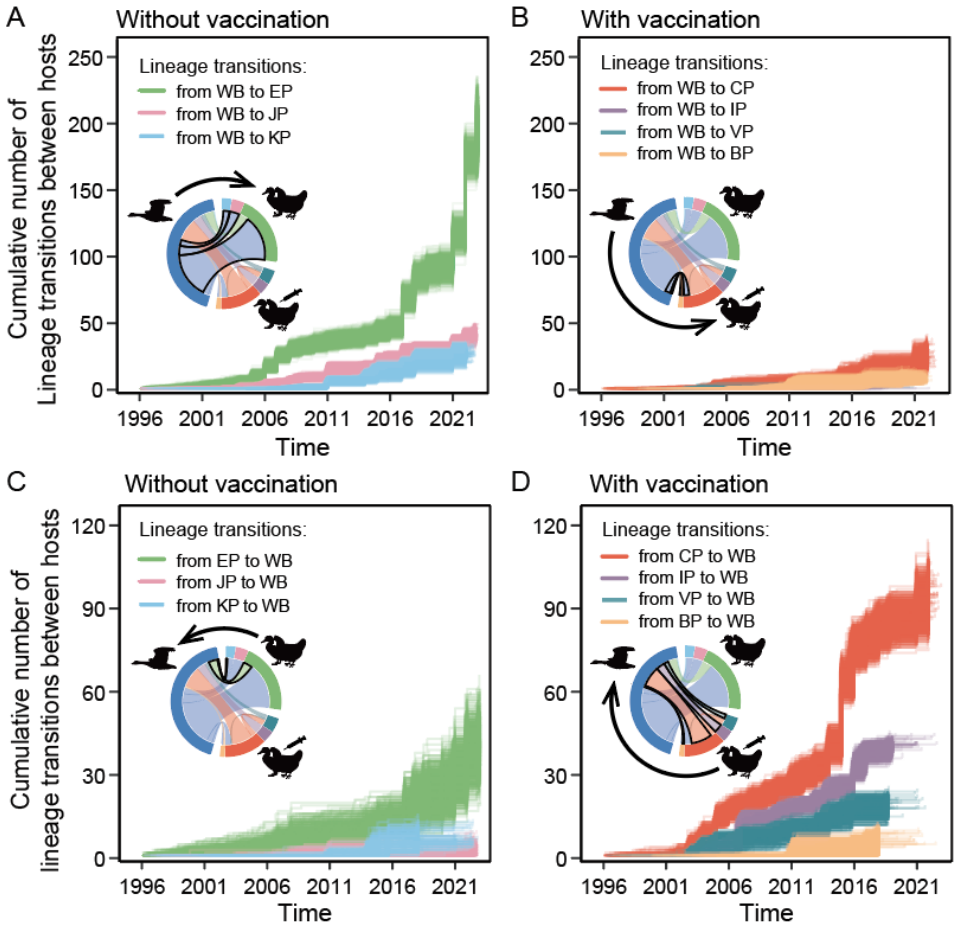


Fig 2. Temporal dynamics of H5 AIV lineage transitions between wild birds and poultry. (**A**) The maximum clade credibility tree of H5 AIV sequences sampled from 1996 to 2023. Tree tips are coloured according to the host from which the sequence was sampled, while internal branches represent ancestral host states inferred using the asymmetric discrete phylogenetic model (dark blue: wild birds; orange: Chinese poultry; light green: European poultry; purple: Indonesian poultry; pink: Japanese poultry; light blue: Korean poultry; dark green: Vietnamese poultry; yellow:

683 Bangladeshi poultry). Inset: a root-to-tip regression of genetic divergence against
684 dates of sample collection. (**B**) Cumulative number of host population changes
685 (Markov jumps) on lineages in the HA gene phylogeny. The lineage transitions
686 between hosts were summarized from a posterior sample of trees from the asymmetric
687 discrete phylogenetic model. (**C**) Time series of the annual mean number of HA gene
688 lineage transitions between wild birds, Chinese poultry and European poultry.
689
690



691

692 **Fig 3. Inter-species lineage transmission between wild birds and poultry**

693 **populations with different vaccination statues.** (**A**) Accumulation of lineage

694 transitions from wild birds to unvaccinated poultry populations. (**B**) Accumulation of

695 lineage transitions from wild birds to vaccinated poultry populations. (**C**)

696 Accumulation of lineage transitions from unvaccinated poultry populations to wild

697 birds. (**D**) Accumulation of lineage transitions from vaccinated poultry populations to

698 wild birds. Chord diagrams show the mean cumulative lineage transitions between

699 different groups of sequences (dark blue: wild birds; orange: Chinese poultry; light

700 green: European poultry, purple: Indonesian poultry; pink: Japanese poultry; light

701 blue: Korean poultry; dark green: Vietnamese poultry; yellow: Bangladeshi poultry).

Animal silhouettes are from PhyloPic.org. The plots were summarized from a posterior sample of trees from the asymmetric discrete phylogenetic model.

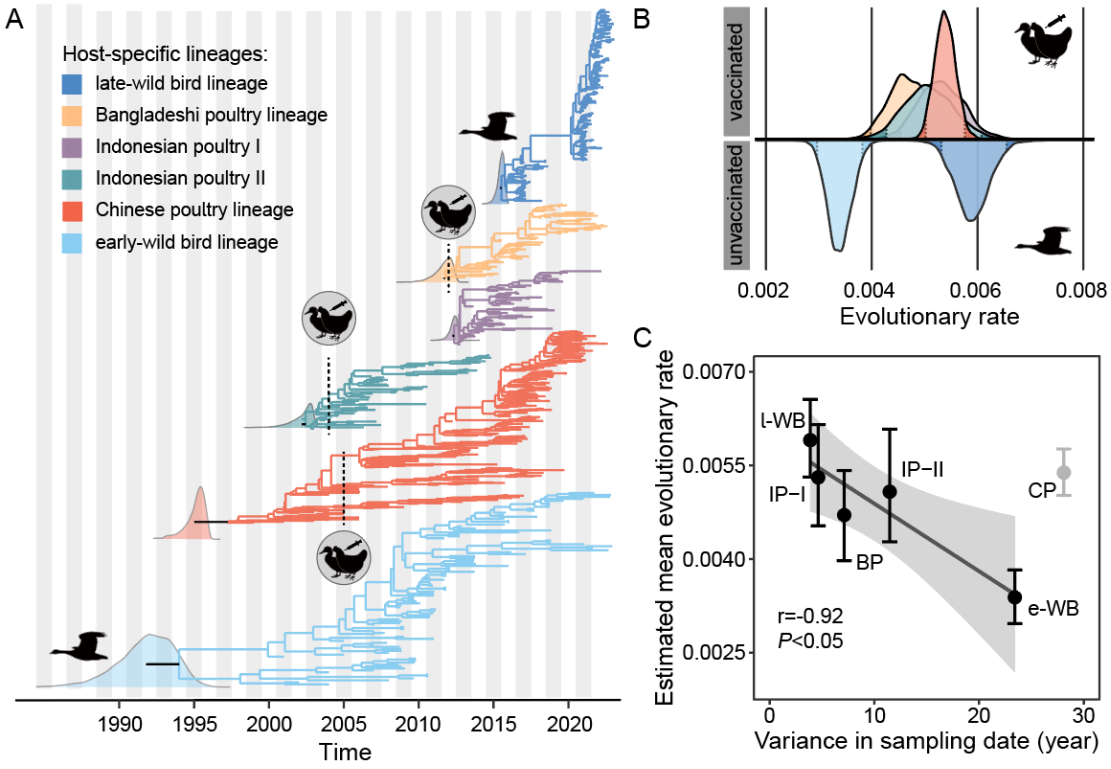


Fig 4. Evolution of the hemagglutinin (HA) gene of H5 AIV in different host-specific lineages. (**A**) Time-resolved phylogenies of H5 AIV lineages in wild birds and vaccinated poultry. The dashed line represents the date when each country started implementing avian influenza vaccination for poultry. The density plot shows the estimated tMRCA for each lineage. (**B**) The estimated substitution rates of the HA gene vary among host-specific H5 AIV lineages. Lineages in unvaccinated host populations (the early- and late-wild bird lineages) are shown below the line and those in vaccinated host populations (the Chinese, Bangladeshi and Indonesian poultry lineages) are shown above the line. Highlighted region shows 95% confidence intervals. (**C**) Scatterplot of the variance in sampling date versus estimated evolutionary rate, for each host-specific lineage. The error bars show the 95% confidence intervals for estimated evolutionary rates. A regression analysis (excluding

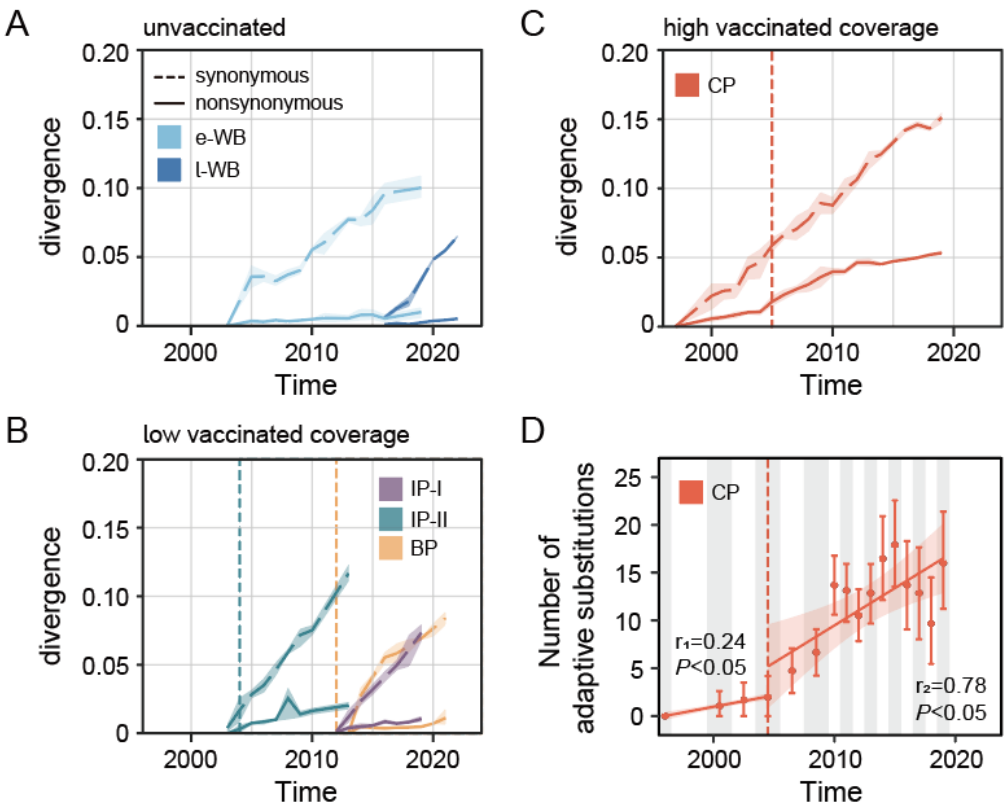
718 the Chinese poultry lineage data point) was conducted to show the negative

719 relationship between expected due to the time-dependency of inferred evolutionary

720 rates ($r = -0.92$; $P < 0.05$), and highlight that the Chinese poultry data point is an

721 outlier.

722



723

724 **Fig 5. Temporal dynamics in divergence and adaptive fixation of the H5 AIV HA**

725 **gene, in different host-specific lineages.** (**A-C**) Nonsynonymous (solid lines) and

726 synonymous (dashed lines) divergence of the HA gene through time. The host-specific

727 lineages were classified into three groups according to the vaccination state:

728 unvaccinated group: (**A**) early-wild bird lineage (e-WB, only the top lineage was

729 retained), late-wild bird linage (l-WB); (**B**) low vaccinated coverage group:

730 Indonesian poultry lineage I (IP-I), Indonesian poultry lineage II (IP-II), Bangladeshi

731 poultry lineage (BP); and (**C**) high vaccinated coverage group: Chinese poultry

732 lineage (CP). Divergences were computed using 1-year sliding windows. Shaded

733 regions show 95% confidence intervals. The dashed line represents the date when

734 each country started implementing avian influenza vaccination for poultry. (**D**) The

735 accumulation of viral adaptative substitutions in the Chinese poultry lineage. Two

736 regression lines were estimated, before (r_1) and after (r_2) 2005. The first sequence,

737 sampled in 1996, was used as the ancestral sequence, those sampled from 1997 to

738 1999 were excluded due to insufficient sample size.

739

740 **Table 1. Positively-selected sites in the HA gene among different host-specific H5**
741 **AIV lineages.**

Host-specific lineages	**Vaccination coverage**	**Livestock density in 2015 (birds/km²) †**		**Methods**			
		chicken	**duck**	**RC**	**FEL (*P* < 0.1)**	**SLAC (*P* < 0.1)**	**FUBAR (PP > 0.9)**
Chinese poultry lineage	73% [25]	921	126	3, 61, 142, 156, 157, 171, 172, 178, 185, 205, 242, 285, 289, 527	87, 142*, 154*, 156, 157*, 172*, 185, 285*, 291*, 325*	3, 142*, 154*, 157*, 172*, 185*, 285*	142*, 154*, 157*, 171, 172*, 285*
Bangladesh poultry lineage	<50% [26]	1448	292	170, 204, 205	204*, 205*	204, 205	170, 204*, 205*
Indonesia poultry lineage I	12% [25]	1262	106	170, 204, 211	205*, 554	205	170*, 205*
Indonesia poultry lineage II				10, 156, 529	3, 11, 156*, 529	156*	5*, 11*, 156*
early-wild bird lineage	-	-		10, 102, 170, 171	102*, 170*, 171	102*, 170	102*, 170*, 171
late-wild bird lineage				10	185, 507	-	252

742 †Mean livestock density was calculated using data from the Gridded Livestock of the
743 World (GLW) [80] website, hosted by Food and Agriculture Organization (FAO;
744 https://dataverse.harvard.edu/dataverse/glw). Regions with <10 birds per square
745 kilometer were excluded from the calculation.
746 *Asterisks mark sites inferred to be under positively selected with posterior
747 probability (PP) >0.95 and <0.05. FEL, Fixed Effects Likelihood. SLAC, Single
748 Likelihood Ancestor Counting. FUBAR, Fast Unconstrained Bayesian
749 AppRoximation. RC, renaissance counting method implemented in BEAST.