Emergence of D1.1 reassortant H5N1 avian influenza viruses in North America

Alvin Crespo-Bellido[1§], Nídia S. Trovão[2,3,4], Wendy Puryear[5], Alexander Maksiaev[1], Jonathan E. Pekar[6], Guy Baele[7], Simon Dellicour[7,8,9], Martha I. Nelson[1§]

[1] Division of Intramural Research, National Library of Medicine, Bethesda, MD USA

[2] Department of Pathobiology, College of Veterinary Medicine, University of Illinois Urbana-Champaign, Urbana, IL, USA

[3] Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, IL, USA

[4] National Center for Supercomputing Applications, University of Illinois Urbana-Champaign, Urbana, IL, USA

[5] Infectious Disease and Global Health, Tufts University Cummings School of Veterinary Medicine, North Grafton, MA01536, USA

[6] Institute of Ecology and Evolution, University of Edinburgh, Edinburgh, UK

[7] Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, Belgium

[8] Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, Brussels, Belgium

[9] Interuniversity Institute of Bioinformatics in Brussels, Université Libre de Bruxelles, Vrije Universiteit Brussel, Brussels, Belgium

[§] **Corresponding authors:** Alvin Crespo-Bellido (alvin.crespobellido@nih.gov) and Martha I. Nelson (nelsonma@mail.nih.gov)

Keywords: influenza A virus, highly pathogenic avian influenza, H5N1, evolution, Bayesian phylodynamics, genomic reassortment

Summary

Since 2021, highly pathogenic avian influenza viruses (HPAIVs) belonging to H5N1 clade 2.3.4.4b have circulated widely in North American wild birds and repeatedly spilled over into mammals. In 2025, the first H5N1-associated deaths in humans were recorded in the Western hemisphere, raising questions about how the ongoing evolution of the virus in wild birds impacts spillover risk. Here, our analysis of 21,471 H5N1 genomes identified an evolutionary shift in mid-2024, driven by interhemispheric migration from Asia and reassortment with new antigens. The genotypes that dominated the early years of North America's H5N1 epizootic traced their ancestry back to Europe, but Asia was the source of new "D1.1" genotype viruses that (a) spread faster, (b) have higher reassortment potential, (c) a broader host range, (d) repeatedly spill over to bovines, and (e) cause severe disease in humans, including non-farm workers.

Introduction

Highly pathogenic H5N1 avian influenza viruses from clade 2.3.4.4b have circulated in Eurasia since 2016[1], causing severe mortality events in wild birds[2], poultry[3,4], and farmed mink and foxes[5,6]. In November 2021, H5N1 clade 2.3.4.4b viruses were introduced by wild birds migrating from Europe to Newfoundland, Canada, representing the first detection of clade 2.3.4.4b in the Western hemisphere[7]. The newly introduced H5N1 viruses spread across North America's four migratory flyways during 2021-2022 and underwent genomic reassortment with North American low pathogenicity avian influenza (LPAI) viruses. To categorize the diversity of new H5N1 genotypes that acquired North American LPAI segments, a "GenoFlu" software tool was developed by researchers at the US Department of Agriculture to rapidly categorize new strains (https://github.com/USDA-VS/GenoFLU)[8]. Non-reassorted Eurasian H5N1 viruses are referred to as "A" genotypes. New reassortant genotypes generated by reassortment events between the A genotypes and LPAI are referred to as the "B, C, or D" genotypes. The B genotypes have been the source of most notable outbreaks in the Americas, including mass mortality events in South American coastal wildlife (B3.2)[9–12], spillover to goats in Minnesota in March 2024 (B3.6)[13], and a multistate outbreak in US dairy cattle that began in 2024 (B3.13)[14–16]. In 2024, a new D1.1 genotype emerged with a new N1 segment acquired from LPAI that has been highly costly for US poultry producers, caused the first human deaths by HPAI in the Americas, and spilled over multiple times into US dairy cattle[17].

Whereas B3.13 spillover events primarily occurred in dairy and poultry workers and resulted in mild conjunctivitis symptoms, D1.1 zoonoses have been associated with more severe illness and the animal source has been less clear. The first D1.1 zoonotic infection involved a hospitalized teenager with no known animal exposure in British Columbia, Canada[18] (November 9, 2024, **Figure S1**). In December 2024, an adult with underlying medical conditions died from a D1.1 infection following exposure to backyard poultry in Louisiana, USA, the first H5N1-associated fatality observed in the Americas[19]. In March 2025, a D1.1 infection in Durango, Mexico led to the death of a young child with no underlying medical conditions, no travel history, nor exposure to infected animals[20]. This death was the first reported H5N1 human infection in Mexico.

To determine how the ongoing evolution of H5N1 in North American wild birds impacts spillover risk for livestock, humans, and mammalian wildlife, we performed comparative and phylogenetic analyses using 21,471 H5Nx clade 2.3.4.4b genome sequences collected from all host species in the Americas, as of January 2, 2026. Bayesian phylodynamic approaches were used to infer evolutionary processes that drive H5N1 evolution in North America, including

ongoing intercontinental migration, rapid spatial diffusion across the North American landmass, genomic reassortment with LPAI, and interspecies transmission involving host adaptation. In light of the sudden expansion of the D1.1 genotype in North America, we compared how D1.1 evolution differs from prior H5N1 genotypes across these four domains.

Results

Three introductions of H5Nx clade 2.3.4.4b viruses from Europe into North America's eastern Atlantic flyway. Multiple independent introductions of clade 2.3.4.4b viruses from Eurasia into North America were identified by phylogenetic analyses of individual gene segments from 5,035 H5N1 viruses collected in Europe, Asia and North America (**Figure 1, Figure S2**). The first clade 2.3.4.4b viruses were introduced into North America's Atlantic flyway from Europe in 2021 (genotype "A2"). A2 became the ancestor of the vast majority of H5N1 genetic diversity that evolved in North America during 2022-2025, including the numerous B and C genotypes that evolved via reassortment with North American LPAIs (**Figure 1, Figure S2**) that were dominant from early 2022 to mid-2024 (**Figure 2**). The C genotypes (e.g., C2.1) evolved independently from a genetically drifted descendent of A2 (referred to as "A1," **Figure 1**). Genotypes A5 (H5N1) and A6 (H5N5) were independently introduced from Europe into North America's Atlantic flyway but had no evidence of reassortment with North American LPAI (**Figures S2-S3**).

Seven introductions of H5N1 clade 2.3.4.4b viruses from Asia into North America. Phylogenetic analysis reveals that the "A3" genotype was introduced from Asia into the Pacific flyway of North America (**Figure S4A**) at least six times independently during 2021-2024 (**Figure 3**). Five of the six A3 introductions (introductions 1, 2, and 4-6) did not sustain long-term transmission in North American wild birds or reassort with LPAI. The third A3 introduction, however, circulated widely in wild birds in the Pacific flyway, before undergoing genomic reassortment with North American LPAI to generate genotype D (**Figure 1**). The D1.1 genotype dominated the North American H5N1 epizootic during late 2024 and throughout 2025 and is the only genotype to dominate in North America during two successive fall migration seasons (2024 and 2025) (**Figure 2**).

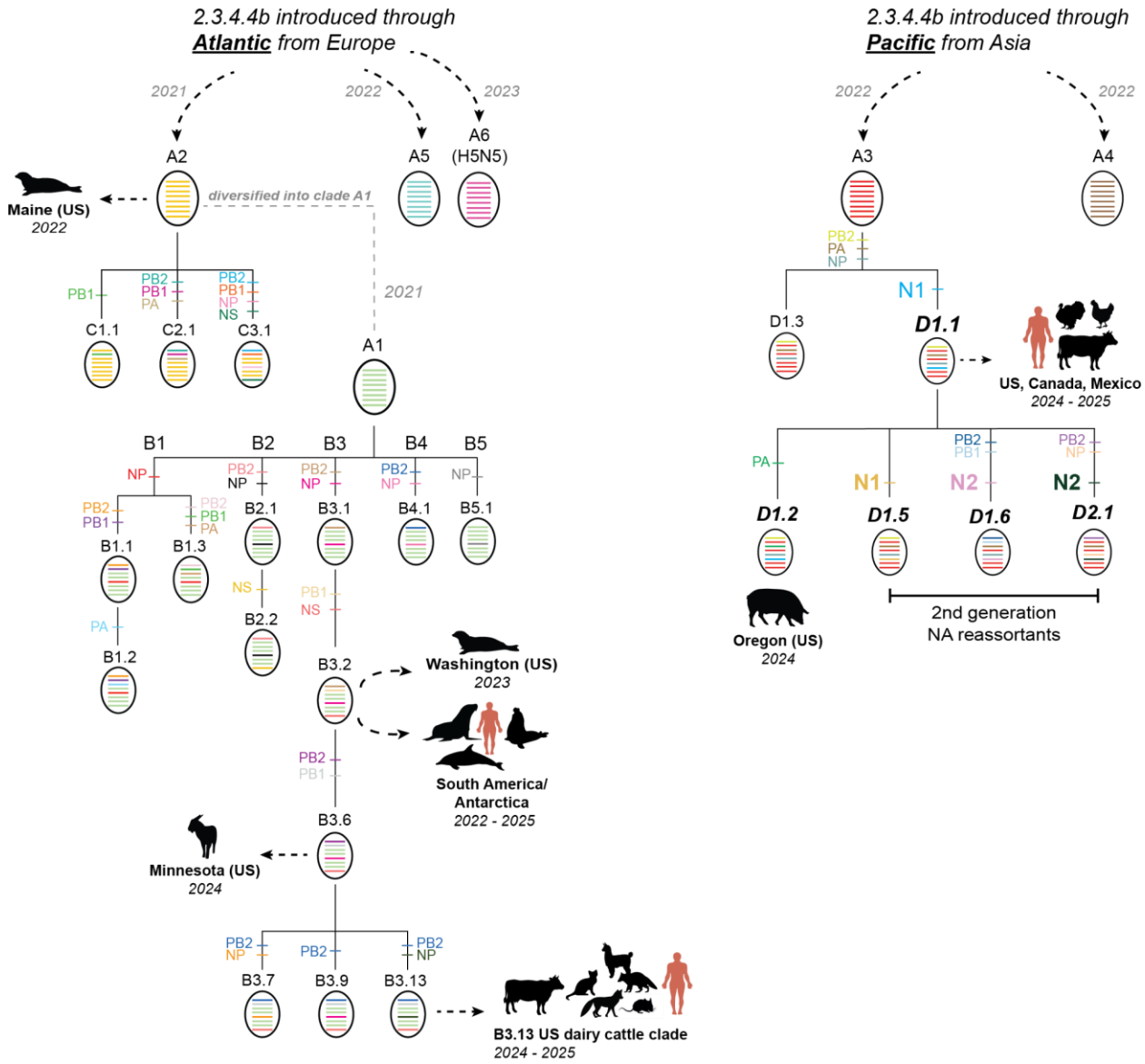


Figure 1. Reassortment history of H5 clade 2.3.4.4b genotypes in the Americas. Each oval represents a genotype (A-D). Each bar inside the oval represents one of the eight segments of the IAV genome, ordered from longest to shortest: PB2, PB1, PA, HA, NP, NA, MP, NS. Bars are shaded by lineage, as defined by GenoFlu (**Supplemental File 1**) and confirmed by phylogenetic analysis (**Figure S2**). For simplicity, all bars from Eurasian-origin "A" genotypes are shaded the same color. Horizontal bars with segment labels indicate where North American LPAI segments were acquired via reassortment, again colored by lineage. Dotted arrows and animal cartoons are provided for select outbreaks associated with spillovers to mammals.

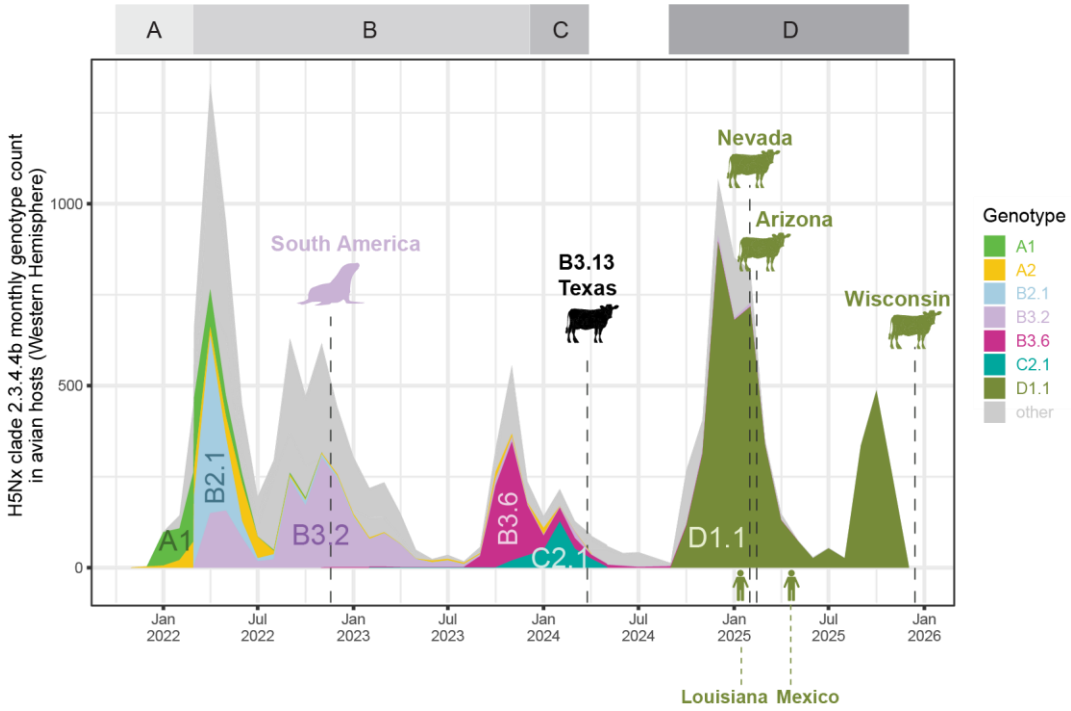


Figure 2. Timing of H5N1 reassortant genotypes in the Americas. The y-axis indicates the number of H5Nx clade 2.3.4.4b sequences detected for each genotype, by month, in avian hosts in the Americas during November 2021 - February 2025, based on 14,390 H5Nx 2.3.4.4b genomes with complete collection dates. The first detected virus is labeled with animal cartoons and dashed lines for select outbreaks/cases: B3.2 in South American marine mammals; B3.13 in Texas dairy cattle; D1.1 in dairy cattle in Nevada, Arizona and Wisconsin; and D1.1-associated deaths in Louisiana and Mexico. Time periods where A-D genotypes are dominant are labeled by boxes above the plot.

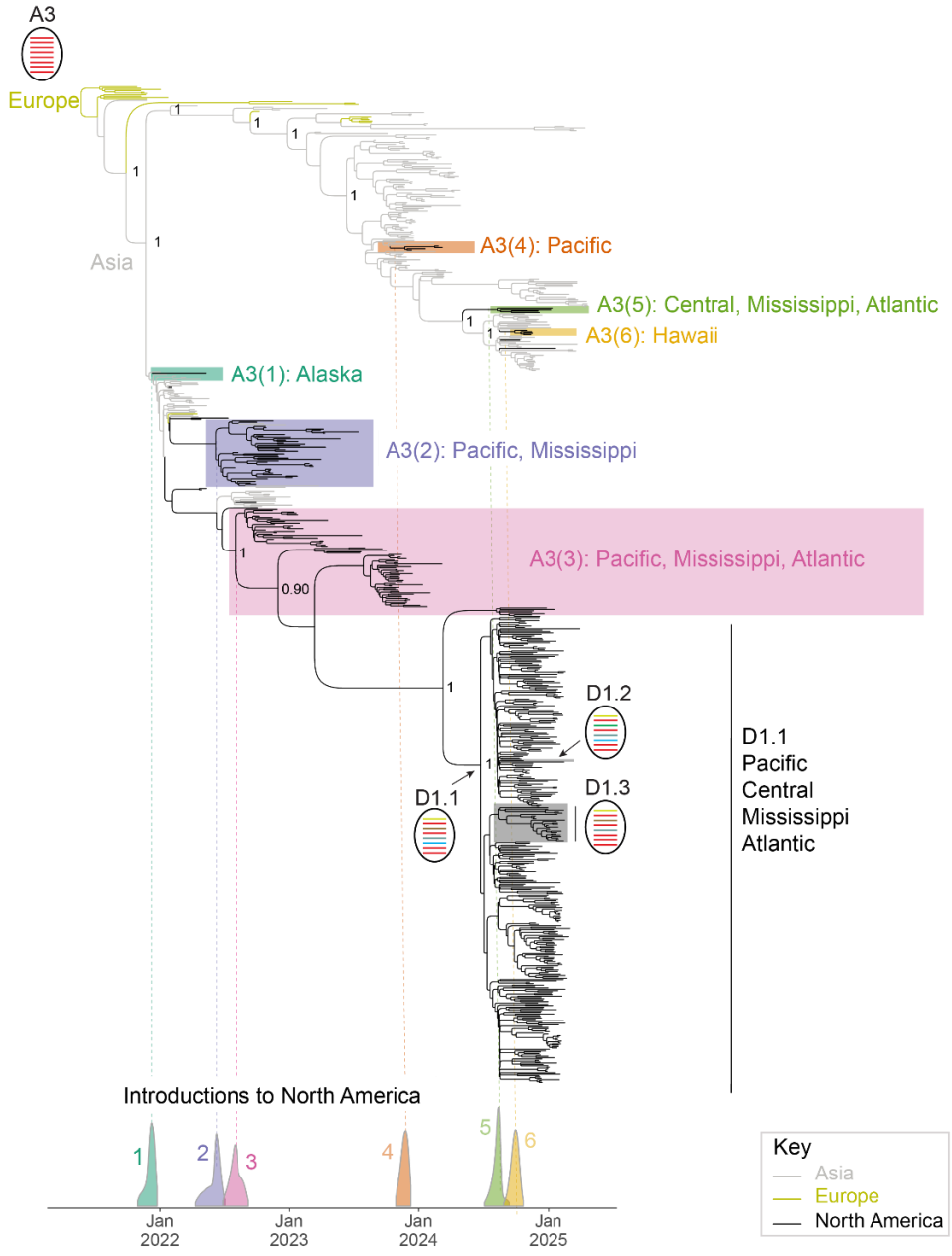


Figure 3. **Introductions of A3 genotype into North American wild birds from Asia.** Time-scaled maximum clade credibility (MCC) tree inferred for four concatenated genome segments (PB1, HA, MP and NS) for which the Eurasian lineage is retained in the D1.1, D1.2, and D1.3 genotypes. Branches are shaded by geographic location. Posterior probabilities are provided for key nodes. The 95% highest posterior density interval for the timing of each of the six A3 North American introductions is provided below the phylogeny. Schematics of reassortant genomes for the genotypes represented in the tree are similar to those presented in Figure 1.

138 **D1.1 acquires N1 from North American LPAI during reassortment.** Prior to the emergence of
139 D1.1 in October 2024, 100% of H5N1 viruses circulating in the Americas still retained the Eurasian
140 N1 (**Figure S5**). Among clade 2.3.4.4b viruses sampled in North American wild birds in 2025 (*n*
141 = 2,576), 92.4% contained an LPAI NA, dominated by D1.1 (**Figure S5**). Most reassortment
142 events between H5N1 clade 2.3.4.4b viruses and North American LPAI during 2022-2025
143 involved the viral ribonucleoprotein (vRNP) complex, which includes the polymerase (PB2, PB1,
144 PA) and NP segments (**Figure S4D-E**). NA and MP segments from LPAI were rarely acquired
145 during reassortment. In a 4:4 reassortment event, the D1.1 genotype acquired the PB2, PA, NP,
146 and N1 segments from LPAI (**Figure S4E**). D1.1 continued to reassort with LPAI to generate
147 "second generation NA reassortants," including D1.6 and D2.1 with N2 segments and D1.5 with
148 yet another distinct N1. (**Figure 1**). The D1.2 virus that infected swine in Oregon in 2024 is similar
149 to D1.2 but underwent another reassortment event in which a different PA was acquired.
150
151
152 **Mutations in D1.1 neuraminidase**. The acquisition of an entire N1 segment from North American
153 LPAI introduced a major genetic change into the H5N1 population. The LPAI N1 acquired by D1.1
154 viruses differs at 22 amino acid positions (**Table S2**) compared to the Eurasian-origin N1 segment
155 found in other H5N1 genotypes in the Americas and globally (mean pairwise genetic percent
156 similarity = 85.6%, **Figure 4A**). The 22 residues are distributed across all major functional
157 domains of the N1 protein (**Table S3**), including 11 in the transmembrane and stalk domains
158 (positions 1-84 in N1 numbering) that anchor the protein in the viral envelope and influence its
159 stability, and 11 on the globular head domain (positions 85+ in N1 numbering) that is a primary
160 target for the host immune system. The mutations on D1.1's globular head domain are arranged
161 in a diagonal pattern on the underside surface of the head, which typically is highly conserved
162 (**Figure 4B**). How these NA mutations affect pandemic risk and immune recognition by antibodies
163 elicited by humans against prior seasonal strains is an important question. The 84.1% genetic
164 similarity observed between the D1.1 N1 and the 2009 H1N1 pandemic (H1N1pdm09) clade is
165 similar to the 84.2% similarity between the B3.13 N1 and H1N1pdm09.

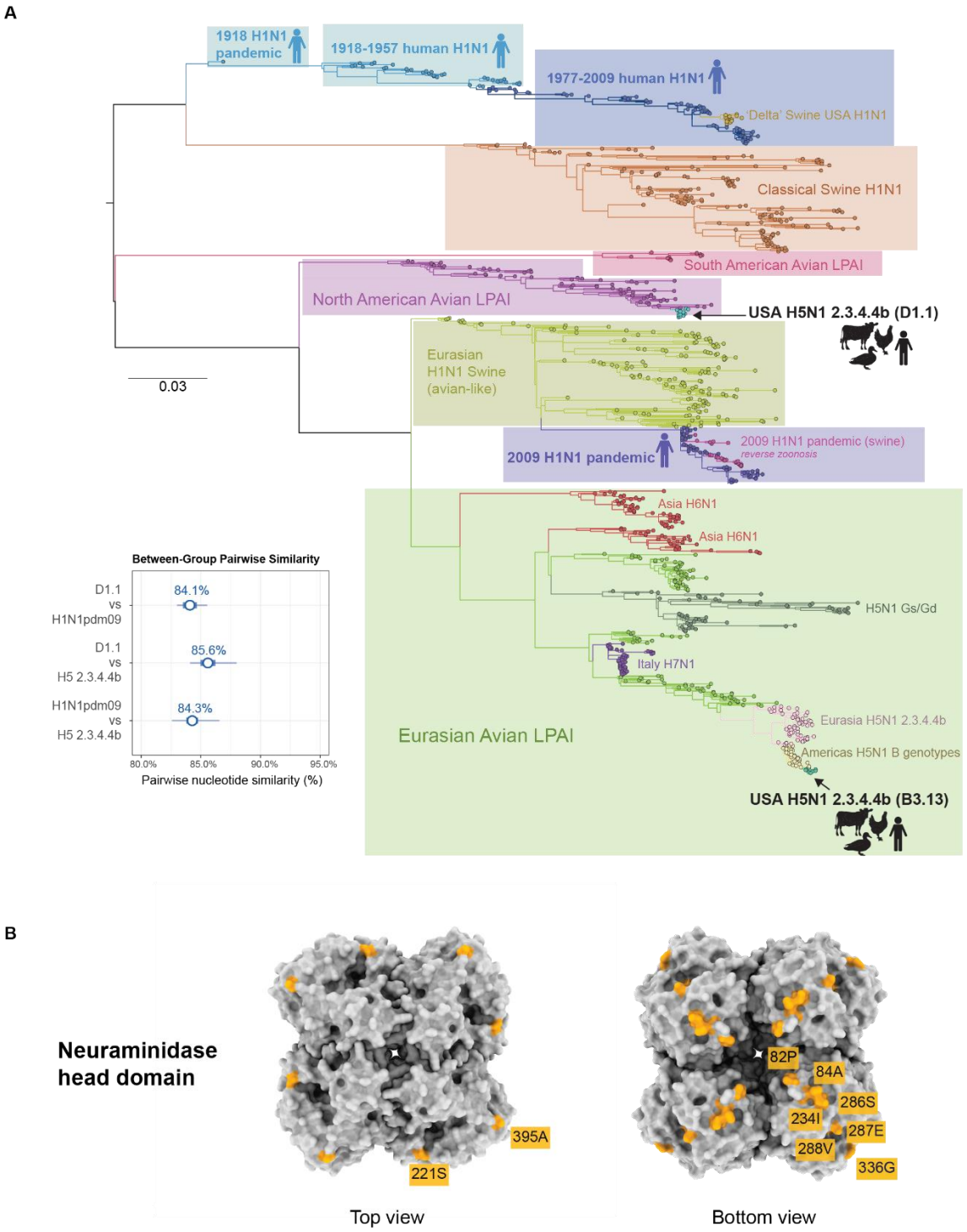


Figure 4. North American N1 segment acquired by D1.1 genotype. (A) Global phylogeny for N1 segment. Major N1 lineages as well as the placement of genotypes D1.1 and B3.13 are highlighted. The inset dot-range plot displays the mean pairwise similarity (dot), the mean ± one standard deviation (thick bar), and the full range of observed pairwise similarities (thin bar) for each between-group comparison. (B) Tetrameric NA head domain protein structure corresponding to H1N1 strain A/California/04/2009 (PDB: 3NSS) annotated with D1.1-specific residues. The surface of the head protein structure is shown at the top while the underside is shown at the bottom. Residues are listed according to N1 (sequential) numbering; the equivalent N2 numbering is listed in Table S3.

Rapid spatial diffusion of D1.1 in North America. Our phylogeographic analysis inferred that the D1.1 genotype emerged in the Central flyway, or possibly Pacific flyway, during the summer of 2024 (95% HPD = 2024-06-17 to 2024-07-24; **Figures S6-S8**). The discrete trait analysis (DTA) provided higher support for the origin of D1.1 in the Central flyway (posterior probability of 0.70, **Figure 5A**), while a post-hoc test based on a continuous phylogeographic analysis provided moderate support for the Pacific flyway (posterior probability = 0.54) and Central flyway (posterior probability = 0.44), with uncertainty reflected in the 95% HPD polygon of the D1.1 MRCA in **Figure 6**. A demographic reconstruction of D1.1 indicates that effective population size increased sharply in the fall of 2024, declined during the spring/summer of 2025, and once again increased during the fall of 2025, consistent with the timing of fall migration (**Figure 5B**). D1.1 lineages tended to remain within a flyway more than expected by chance[21], as evidenced by positive (but not strong) statistical support for a flyway containment effect (Bayes factor = 4.2). D1.1 appeared to persist in the Atlantic flyway during the summer of 2025 (**Figure 5A**). However, viruses sequenced during the second D1.1 wave primarily descended from viruses in the Central flyway positioned deeper in the tree. The continuous phylogeographic reconstruction estimated that D1.1 dispersed across all four flyways during 2024-2025 with the highest weighted diffusion coefficient of North America's four major dominant genotypes during its invasion phase (~6,000 km²/day, **Figure 6**, **Table S4**). The wavefront velocity during D1.1's invasion phase (~16,000 km/year) was substantially higher than the wavefront velocities for B3.2 (~7,000 km/year), B3.6 (~2,500 km/year) and C2.1 (~6,000 km/year, **Figure 6**).

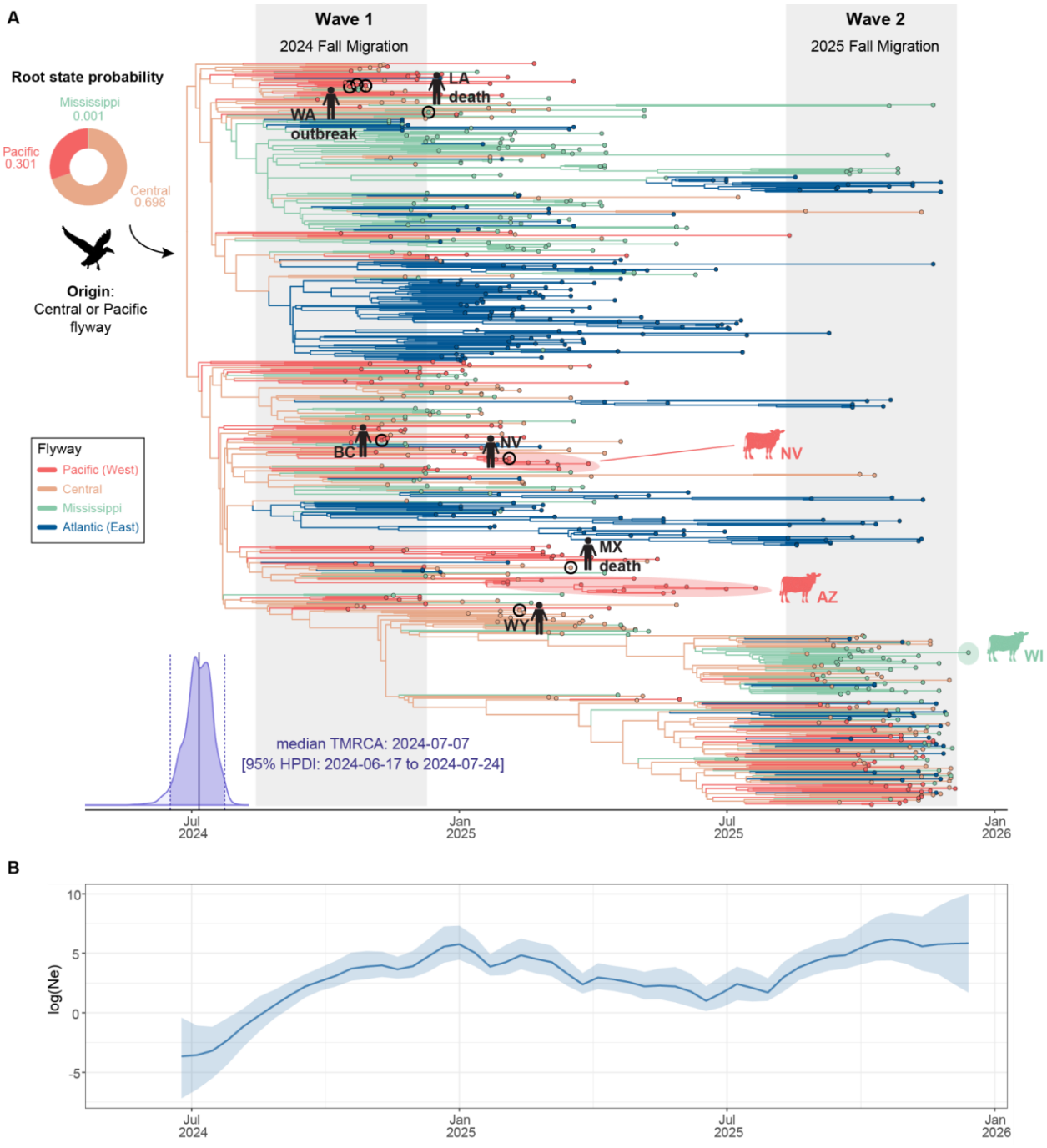


Figure 5. MCC tree of D1.1. (A) Discrete phylogeographic reconstruction of the dispersal history for H5N1 genotype D1.1 across the North American waterfowl migratory flyways. The consensus phylogenetic tree shows the evolutionary relationships of genotype D1.1 viruses, with the x-axis scaled in years. Tips are shaded according to the flyway where they were collected, and branches are shaded according to the flyway inferred at ancestral nodes. The purple curve at the root of the tree represents the posterior density distribution for the TMRCA. (B) Skygrid reconstruction of the effective population size for D1.1. The dark blue line and shaded light blue region represent the median log-transformed viral effective population size *Ne* and its 95% highest posterior density (HPD) interval region, respectively.

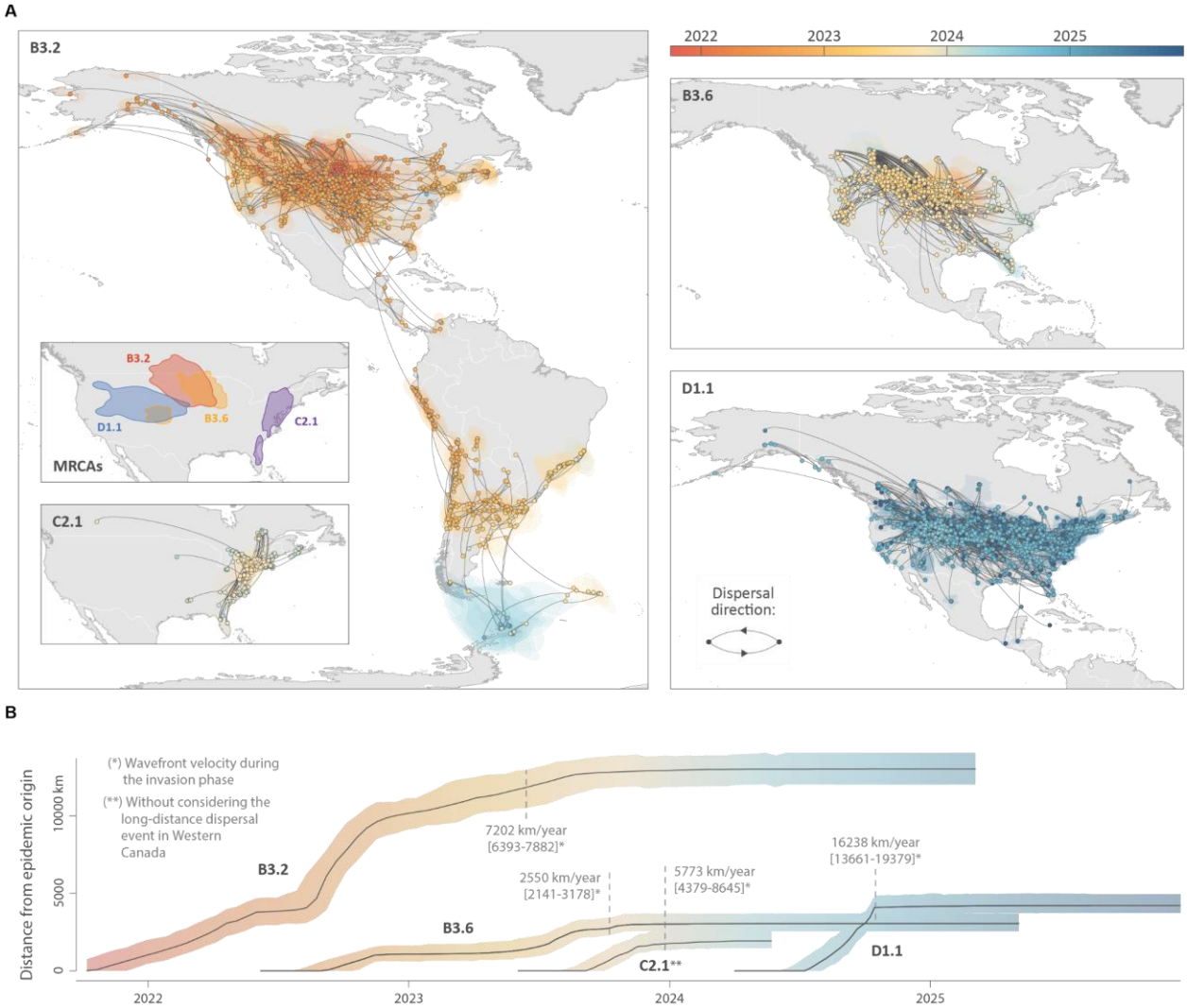


Figure 6. Spatial dispersal of major H5N1 genotypes in the Americas. (A) Phylogeographic reconstruction maps showing maximum clade credibility (MCC) tree branches as curved lines. Nodes are time-colored from red (2022) to blue (2025), with 80% highest posterior density (HPD) polygons indicating uncertainty. Inset maps show the estimated geographic origins as the 95% HPD regions for most recent common ancestors (top) and the C2.1 phylogeographic reconstruction (bottom). (B) Expansion dynamics showing median wavefront distance from geographic origin over time, with 95% HPD intervals. The annotated values indicate median wavefront velocities (km/year) during the initial invasion phase.

D1.1 spillover and adaptation to mammals. D1.1 has frequently transmitted from wild birds to a range of diverse host species (**Figure 7A, Figure S9**). D1.1 has a broad host range compared to prior genotypes and is the only genotype to infect seven host types: wild bird, terrestrial mammal, marine mammal, domestic bird, cattle, domestic mammal (other than cattle), and human (**Figure 7B**). D1.1 also infected a broader range of wild bird orders (*n* = 14, **Figure 7C**) than any other genotype, including the B3.2 genotype that spread widely in South America. Ecologically, most D1.1 spillovers involved direct exposure to wild birds (**Figure 7A**) and these avian-origin

viruses experienced strong selection in mammalian hosts for PB2 mutations that improve virus replication efficiency in mammals. Close to 35% of D1.1 spillovers from wild bird into mammals (e.g., humans, seals, bears, raccoons, skunks) had the E627K mutation in PB2, including the human cases A/Wyoming/01/2025 and A/British Columbia/PHL-2032/2024 (**Figure 8, Figure S10**). The D701N mutation in PB2 seen in South American marine mammals also became fixed in the D1.1 Nevada cattle clade and was found in ~7% of other D1.1 spillovers into mammals (e.g., domestic cat, fox). In contrast, once the PB2 segment became adapted to cattle (D701N for D1.1; M631L for B3.13), there is no evidence of emergence for additional PB2 mutations following spillover events from cattle to other mammalian hosts (**Figure 8, FiguresS10-S11**).

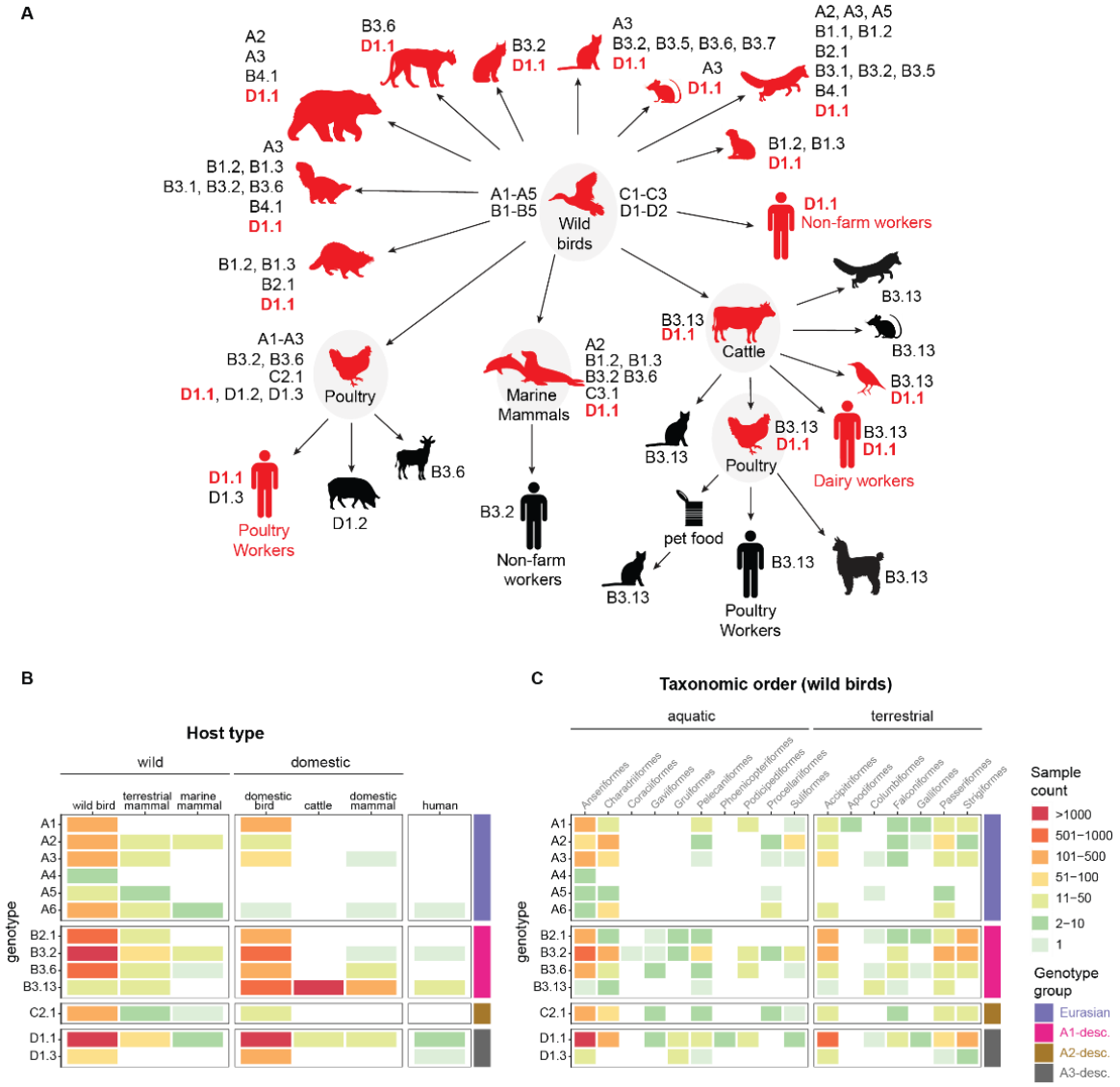


Figure 7. Host range of H5N1 clade 2.3.4.4b genotypes. (A) Spillover of H5N1 clade 2.3.4.4b genotypes between different host species based on phylogenetic evidence. (B) Heatmap shaded by the number of sequences in each wild bird taxonomic order (n = 4,044) and (C) host type (n = 13,488). Select genotypes are shown.

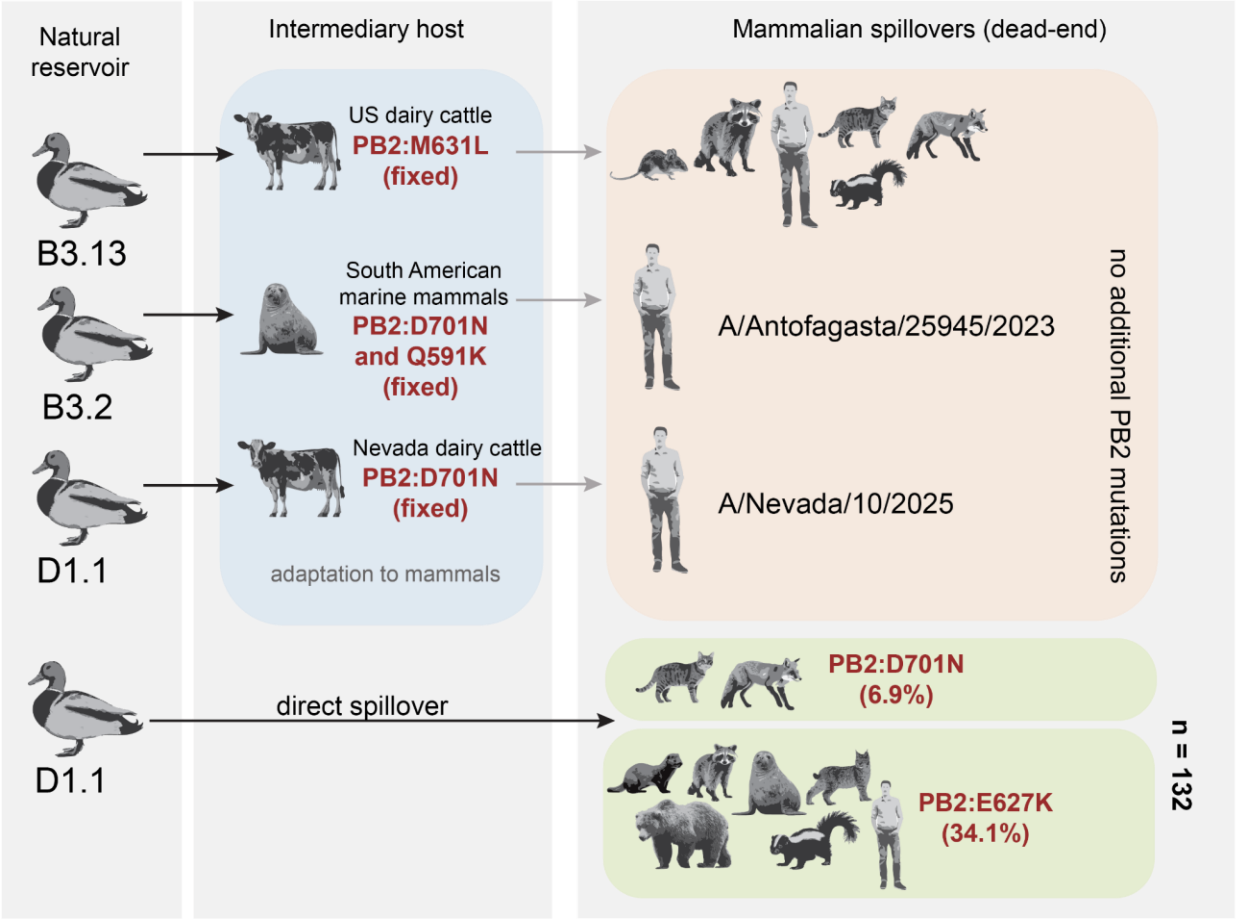


Figure 8. PB2 mutations following spillover events in mammals. Black arrows indicate H5N1 spillover events from the avian natural reservoir to mammalian hosts. Grey arrows = secondary spillover events from intermediary hosts (e.g., bovines, marine mammals) to dead-end mammalian hosts. Blue box = intermediary hosts that acquired PB2 mutations that improve virus replication in mammals. Orange box = secondary hosts that experience little to no selection for additional PB2 mutations because viruses transmitted from intermediaries are already adapted to mammals. Green boxes = mammals infected with avian-origin D1.1 where selection is strong for PB2 adaptations.

Discussion

The emergence of the D1.1 genotype reflects North America's expanding role as a source for novel reassortant H5N1 viruses with pandemic potential and the increasing complexity of the H5N1 epizootic's geography, genetics, and spillover risk. Our findings support a model in which global viral connectivity introduces genetically distinct lineages into new ecological systems, where reassortment with local gene pools generates genotypes with enhanced evolutionary flexibility, enabling rapid expansion and cross-species transmission. Prior to the emergence of D1.1 in 2024, the genetic diversity of the North American 2.3.4.4b epizootic was constrained by

247 the fact that (a) almost all strains traced their ancestry back to the original 2.3.4.4b introduction
248 from Europe and (b) LPAI internal gene segments were almost exclusively acquired during
249 reassortment events, with no swapping of NA antigens[22]. The human population at risk for
250 spillover was likewise constrained, as H5N1 zoonotic cases in 2024 primarily occurred in US farm
251 workers with direct animal contact and generally caused mild illness. It remains unclear whether
252 the H5N1 fatalities caused by D1.1 are an indication of a higher pathogenicity of the virus in
253 humans or a larger denominator of human cases beyond just farm workers. In which case H5N1
254 surveillance in humans has not kept pace with the expanding population at risk.

255 D1.1 is the only genotype to cause two successive epidemic peaks in wild birds, with no
256 sign of being replaced by a different genotype. Prior North American genotypes dominated for
257 one epidemic peak (although B3.2 continues to circulate in South America and the Antarctic
258 region[23]). At the time of this writing, D1.1 is continuing to acquire new N1 and N2 antigens through
259 additional reassortment events with LPAI, suggesting that H5N1 evolution is continuing to
260 accelerate and explore new antigenic pairings.

261 Following years of high conservation of the HA-NA pairing of 2.3.4.4b viruses, it is unclear
262 why a new North American LPAI N1 was acquired in 2024 without any apparent loss in fitness.
263 We found no evidence of permissive mutations[24] in the Asian-origin A3 hemagglutinin in the
264 months leading up to D1.1's emergence that could have explained a greater tolerance for a new
265 NA. The implications of the new N1 for pandemic risk require further research. Human exposure
266 to H1N1 seasonal influenza viruses containing N1 segments from the 2009 H1N1 pandemic,
267 originally sourced from Eurasian swine, is thought to provide a degree of cross-protection for
268 humans against most H5N1 clade 2.3.4.4b viruses circulating in birds globally. The amino acid
269 changes that distinguish the LPAI N1 found in D1.1 from the Eurasian N1 are concentrated on
270 the underside of the globular head domain – or the neuraminidase "dark side" (NDS) – which is
271 positioned away from the enzymatic active site. This region has recently been identified as
272 containing highly conserved epitopes that serve as effective targets for cross-reactive monoclonal
273 antibodies against HxN1 infections[25].

274 As the Americas emerge as a new epicenter for H5N1 evolution, there is an urgent need
275 to expand the region's infrastructure for real-time surveillance and monitoring. The Americas
276 present a different ecological landscape compared to Eurasia, including South American marine
277 mammals that are highly susceptible to H5N1 disease and farming systems with long-range
278 animal movements that sustain virus persistence in dairy cattle[16]. Tracking H5N1 evolution in the
279 Americas is made possible by the collective efforts of government and academic partners,
280 including the U.S. Department of Agriculture[8], Centers for Disease Control and Prevention[26], US

281 Geological Survey[27], the Canadian government[28], South American researchers[10,12], and NIAID-
282 funded Centers of Excellence for Influenza Research and Response (CEIRR) network[29]. The
283 capacity of the USDA's National Milk Testing Strategy (NMTS) to detect three independent D1.1
284 spillover events into US dairy cattle by actively sampling raw bulk milk tanks demonstrates one
285 area where H5N1 active surveillance in North America is advancing. Still, there are severe gaps.
286 A human fatality involving a highly pathogenic H5N5 virus in Washington State (genotype A6)
287 was reported on November 15, 2025 (https://doh.wa.gov/newsroom/grays-harbor-county-
288 resident-dies-complications-avian-influenza), despite no detections of H5N5 at that time in any
289 host outside of the Atlantic flyway on the opposite coast[30], exposing major gaps in our knowledge
290 of which genotypes are circulating in wild birds (and where) caused by a lag in H5N1 data
291 submission to public repositories. Numerous genotypes are listed on GenoFlu for which no or
292 very few genetic sequences are available yet on SRA or GenBank, including some of the latest
293 D genotypes with novel NA segments that could not be included in this study. Tracking how
294 genotypes evolve over time and space is also limited by the large numbers of H5N1 genomes on
295 SRA that are labeled "USA" and have no collection date. It can take months (or more) before a
296 corresponding GenBank entry appears with state and date metadata. H5N1 surveillance is
297 severely underfunded, but it would be sensible to invest more resources into high-throughput data
298 pipelines so that crucial H5N1 data already being generated by government agencies can quickly
299 get into the hands of the global scientific community to track changes in evolution and spillover
300 risk in real time.

301 In recent years, North American avian influenza surveillance has prioritized sequencing
302 HPAI over LPAI, for good reason. However, LPAIs still present an important pandemic risk (e.g.,
303 1918, 1957, 1968, and 2009 influenza pandemics) and contribute key segments to new H5N1
304 reassortants and should not be overlooked. Lower availability of LPAI sequences from countries
305 around the world in public databases decreases phylogenetic resolution and makes it difficult to
306 distinguish reassortment events between HPAI and LPAI. One source of confusion is the A1 and
307 A2 genotypes and whether they represent the first two 2.3.4.4b introductions from Europe. Our
308 updated phylogenetic analysis supports an alternate scenario where A1 and A2 descend from the
309 same introduction from Europe. A2 arrived first, and A1 genetically drifted into a new lineage as
310 mutations accumulated over time (**Figures 1 and S2**). Efforts to develop a new universal global
311 H5N1 lineage classification tool ("ggFlu", https://www.ggflu.com/) represent the latest attempt to
312 bring order to the increasing complexity of global H5N1 evolution (**Table S5**).

Methods

Dataset curation and maximum likelihood phylogenetic analysis for H5 clade 2.3.4.4b. An H5Nx clade 2.3.4.4b dataset for North America, South America and Antarctica was compiled by querying the Global Initiative on Sharing All Influenza Data (GISAID; https://gisaid.org/)[31–33] and NCBI Virus[34] databases for all H5Nx clade 2.3.4.4b samples collected between November 1, 2021 – January 2, 2026 with available genomic sequences for all eight gene segments (accessed on June 2, 2026). Additionally, consensus genomes from raw reads deposited to NCBI's Sequence Read Archive and associated with BioProjects PRJNA1102327, PRJNA1122849, PRJNA1134696, PRJNA1219588, PRJNA1207547, PRJNA980729 were downloaded from the Andersen Lab's avian-influenza github repository (https://github.com/andersen-lab/avian-influenza). The samples in each dataset were assigned a genotype using the GenoFLU genotyping tool[8]. GenoFLU categorizes reassortant clade 2.3.4.4b genotypes as "major" (n = 37, as of April 7, 2026) if they meet the criteria of at least 20 wild bird detections, infection of two or more poultry premises, and/or detection in atypical mammalian host species. The vast majority of reassortant genotypes (123/160) are designated by GenoFlu as 'minor'. Samples with unassigned genotypes were removed from the dataset and duplicate samples across databases were filtered by isolate id. Of the 166 genotypes listed on GenoFLU's website, genetic sequence data was publicly available for 107 of them (~65%). The final Western Hemisphere clade 2.3.4.4b dataset consisted of 21,471 genome sequences. To place the clade 2.3.4.4b genotypes in phylogenetic context and identify reassortment involving American LPAI viruses, a background dataset was compiled by querying the GISAID database for LPAI samples with complete genomes collected in the Americas and Antarctica between January 1, 2015 – January 2, 2026 (accessed on April 4, 2026). Select samples belonging to distinct lineages were removed from the analysis which included H1N1 and H1N2 samples from swine, and H3N8 samples from canines and horses. The search resulted in an LPAI dataset of 3,830 samples. All sequences from the American clade 2.3.4.4b and background LPAI were combined and grouped by segment resulting in eight segment datasets – PB2, PB1, PA, H5, NP, N1, MP and NS.

An additional background dataset was compiled to evaluate the number of North American clade 2.3.4.4b introductions of fully Eurasian genotypes A1-A6 by downloading all European and Asian H5Nx clade 2.3.4.4b and LPAI samples with complete genomes collected between January 1, 2015 and June 13, 2025 (accessed on June 13, 2025). To reduce the size of the datasets while keeping a representative Eurasian background sample set, the sequences were then grouped by segment and used to create local Basic Local Alignment Search Tool (BLAST) databases using

347 BLAST v2.15.0[35]. The resulting Eurasian segment databases were BLASTed against using the
348 American A1-A6 clade 2.3.4.4b datasets as queries, and the top 200 Eurasian BLAST hits with
349 >85% sequence identity across all American clade 2.3.4.4b sequences were included for
350 subsequent phylogenetic analysis. The final American A1-A6 + Eurasian sequence datasets
351 consisted of: H5 (*n* = 4,923), MP (*n* = 5,035), N1 (*n* = 3,935), NP (*n* = 5,035), NS (*n* = 5,035), PA
352 (*n* = 5,035), PB1 (*n* = 5,035), and PB2 (*n* = 5,035).
353 Each gene segment dataset was aligned using MAFFT[36,37] and trimmed to the coding
354 sequence region. Highly diverged sequences that were poorly aligned were removed from each
355 segment dataset. Sequences were also excluded excluded if more than 25% of bases were
356 ambiguous or if their ungapped length fell below 75% of the mean ungapped length for that
357 segment. Maximum-likelihood (ML) phylogenies were inferred for each dataset using IQ-TREE
358 v2.47[38] with a general-time reversible (GTR) model of nucleotide substitution with among-site rate
359 heterogeneity modelled through a discretized gamma distribution (+G) and 1,000 ultra-fast
360 bootstrap (UFBOOT) replicates, using the high-performance computational capabilities of the
361 Biowulf Linux cluster at the National Institutes of Health (http://biowulf.nih.gov).
362 To increase the readability of the American clade 2.3.4.4b and background LPAI
363 phylogenies consisting of >21,000 sequences, ML phylogenetic analyses were also performed
364 on representative, subsampled datasets. Each segment dataset was subsampled using SAMPI
365 (https://github.com/jlcherry/SAMPI)[39]. Subsampling was stratified by collection date, location, host
366 category (avian, cattle, environmental, feline, human, and other mammals), and assigned
367 genotype to maintain representative diversity across these variables. The resulting subsampled
368 datasets consisted of H5 (*n* = 1,160), MP (*n* = 1,340), N1 (*n* =1,120), NP (*n* = 1,354), NS (*n* =
369 1,346), PA (*n* = 1,358), PB1 (*n* = 1,353), and PB2 (*n* = 1,372). Phylogenies were visualized using
370 the "ggtree" R package v3.13.1[40].
371
372 *Bayesian phylodynamic analyses*. Bayesian phylodynamic reconstructions were performed for
373 temporally dominant genotypes B3.2, B3.6, C2.1 and D1.1 using concatenated genome datasets
374 (all eight segments were concatenated using ConcatenatorGUI[41]). For computational efficiency,
375 subsampled sequence datasets were first obtained by conducting subsampling by phylogenetic
376 clustering. This subsampling procedure consists in discarding sequences such that monophyletic
377 clusters of sequences sampled from the same location (i.e. the most precise administrative region
378 of origin of each sample) are only represented by a single sequence[42]. Given that such clusters
379 would largely represent dispersal within a specific administrative region and given that most
380 geographic coordinates are here integrated through a homogeneous prior range delineated by

381 the polygon of this administrative area (see below), conserving more than sequence per cluster
382 would lead to phylogeographic 'noise' and not further inform the overall dispersal history of viral
383 lineages. For genotypes B3.2 and D1.1, we further decreased the size of the resulting datasets
384 by randomly subsampling two sequences per combination of sampling month and region of
385 administrative level 1 (admin-1 level), while preferentially conserving samples associated with the
386 most precise sampling origins. This resulted in the following final datasets – B3.2 (*n* = 594), B3.6
387 (*n* = 526), C2.1 (*n* = 165), D1.1 (*n* = 699). To assess the timing of North American A3 introductions
388 and the emergence of the D1 genotypes, an additional analysis was performed on a concatenated
389 shared segment dataset (i.e., PB1, HA, MP and NS) of all Eurasian (downloaded from GISAID)
390 and North American A3 sequences and a random subset of the D1.1-D1.3 samples (*n* = 725).
391
392 Bayesian phylogenetic and phylogeographic analyses were performed using the software
393 package BEAST X v1.10.5[43], using the BEAGLE high-performance library for computational
394 efficiency[44] and employing default prior settings in BEAUti X[43] for all analyses. We employed a
395 GTR+Γ nucleotide substitution model, an uncorrelated relaxed molecular clock with an underlying
396 lognormal distribution to account for rate variation across lineages[45], and a non-parametric
397 SkyGrid coalescent tree prior[46]. Markov chain Monte Carlo (MCMC) chains were run for at least
398 300 million iterations while sampling every 10^5 iterations. Empirical tree distributions of 1,000
399 trees were collected and subsequently used in the phylogeographic reconstructions[47]. To
400 reconstruct the finer-scale geographic spread of the more recently emerged genotypes B3.2,
401 B3.6, C2.1, and D1.1, we conducted continuous phylogeographic analyses using the relaxed
402 random walk (RRW) diffusion model implemented in the software package BEAST X[43], with a
403 Cauchy distribution to model dispersal velocity heterogeneity[48–50]. For the genomic samples that
404 did not have associated precise sampling coordinates, we used the polygon(s) of their
405 administrative area of origin to define an homogeneous prior range of sampling coordinates and
406 estimate those sampling locations through Bayesian phylogeographic inference[51,52]. We also
407 conducted discrete phylogeographic analyses using the Bayesian stochastic search variable
408 selection (BSSVS) approach[53] implemented in BEAST X and while considering as a distinct
409 discrete location each of the four administrative waterfowl flyways (i.e. the Pacific, Central,
410 Mississippi and Atlantic flyways, as defined by the U.S. Fish and Wildlife Service). For all
411 phylogeographic analyses based on empirical trees, MCMC chains were run for more than 10^7
412 iterations while sampling every 10^5 iterations. MCMC convergence and mixing were evaluated in
413 Tracer v.1.7.3[54], ensuring that all continuous posterior estimates were associated with an effective
414 sample size (ESS) value of at least 200. At least 10% burn-in was removed from each run and

415 maximum clade credibility (MCC) trees were computed using TreeAnnotator X[55]. We used the R
416 package "seraphim"[56] to generate a visualization for the different continuous phylogeographic
417 reconstructions and for estimating dispersal statistics based on these reconstructions. For C2.1,
418 a unique long-distance dispersal event detected towards Western Canada was removed from the
419 dispersal calculations.
420
421 *Global N1 phylogeny and Pairwise nucleotide percent similarity between major N1 lineages*
422 Similarly to Trovão et. al[57], a background of all available neuraminidase sequences for influenza
423 A N1 from GenBank and downloaded specific geographical isolates, not available in GenBank,
424 from GISAID were obtained on 16 July 2019. The sequences were aligned using the MAFFT[36,37]
425 multiple sequence alignment tool. The aligned sequences were manually edited and cleaned in
426 AliView version 1.26 software[58]. The resulting data set was trimmed at the 5′ and 3′ ends to include
427 solely the coding sequence. This data set was subjected to multiple iterations of phylogeny
428 reconstruction using FastTree version 1.0[59] with a general time-reversible (GTR) model and
429 exclusion of outlier sequences whose genetic divergence and sampling date were incongruent
430 using TempEst[60].
431
432 We defined lineages from the same host and geographical region and only considered those
433 circulating for at least 3 years in the region. Each lineage includes only viruses that infect a single
434 host species or closely related host species. Each inferred introduction into a new species creates
435 a new lineage. In some cases, what could be a lineage is subdivided on the basis of geography.
436 Besides the H1N1 pandemic lineage circulating in humans since 2009, we defined two human N1
437 lineages circulating during the periods of 1918-1957 and 1977-2009. The N1 lineage reintroduced
438 in 1977 derives directly from the viruses circulating in 1957, so studying the two lineages together
439 would disrupt the temporal signal and consequently the molecular clock model, due to the lack of
440 accumulation of mutations during the 20-year period where N1 did not circulate[61–64]. For the non-
441 avian lineages, we excluded sequences with mixed subtypes, as marked by "Nx" or "mixed." For
442 some analysis, we combine lineages for the same host. For visualization purposes, we excluded
443 sequences that cover less than 75% of the coding sequence with unambiguous bases and
444 retained only one of identical sequences collected in the same country on the same day or in the
445 same month or year in the case of incomplete date information.
446
447 Next, we applied a subsampling strategy that selected a specific number of sequences per
448 country per time to obtain similarly sized lineages that represented the overall circulating diversity.

449 When sequences had the same country and collection date, we chose, in a decreasing order of
450 priority, those that were longer, those that had at least the month of collection rather than just the
451 year, those with a smaller number of internal gaps with lengths not divisible by 3 (codons), and
452 those with full dates of collection (https://github.com/jlcherry/SAMPI)[39]. We combined all
453 subsampled lineages for each subtype and performed a final iteration of phylogeny reconstruction
454 and exclusion of outlier sequences as described above.
455
456 The abovementioned previously curated dataset was used to place the D1.1 and other clade
457 2.3.4.4b sequences in phylogenetic context. Pairwise nucleotide sequence similarity was
458 calculated between all inter-group sequence pairs across the aligned N1 segment using a custom
459 R script. For each group comparison, similarity was computed as the proportion of identical sites
460 at non-gap, non-ambiguous positions across all pairwise combinations of sequences between the
461 two groups.
462
463 *Identification of genotype-specific amino acid residues and D1.1 phenotypic markers.* American
464 H5Nx clade 2.3.4.4b segment datasets were translated into protein alignments with Nextclade
465 CLI[65] with built-in reference datasets: H5N8 'A/Astrakhan/3212/2020(H5N8)' for HA, H1N1
466 'A/Wisconsin/588/2019' for NA, and H1N1 'A/California/07/2009' for all other segments.
467 Sequences that failed to be aligned to the reference were not considered for this analysis.
468 Genotype-specific amino acid residues were identified among dominant genotypes (B3.2, B3.6,
469 B3.13, C2.1, and D1.1) using a custom R script. We defined genotype-specific residues as those
470 meeting three criteria: (1) present in >90% of sequences within the target genotype, (2) present
471 in <20% of other dominant genotypes, and (3) present in <10% of all other clade 2.3.4.4b
472 sequences combined.
473
474 *Analysis of mammalian adaptation markers.* Known phenotypic markers of mammalian
475 adaptation (PB2:E627K and PB2:D701N) were identified within complete nucleotide alignments
476 for D1.1 (n=5,257) and B3.13 (n=5,280) genotypes using FluMutGUI v3.2.0[66]. Maximum likelihood
477 phylogenies were constructed from the D1.1 dataset using IQTREE as described above to
478 characterize spillover events. Spillover events involving mammalian hosts were categorized by
479 source as "direct spillover" from avian hosts, "AZ cattle clade"-derived, or "NV cattle clade"-
480 derived. For B3.13, groups were defined as "pre-PB2:M631L" or "post-PB2:M631L", given that
481 M631L was fixed early in the cattle outbreak and subsequent spread was driven by spillovers

482 from cattle[14]. We quantified the frequency of PB2 mutations across these spillover categories and
483 different host types.
484
485 **Data and code availability**
486 All data (including supplemental files and GISAID acknowledgement table) and BEAST X XML
487 files with model parametrizations, phylogenetic trees and analysis output files are available at
488 https://github.com/nidiatrovao/H5N1Americas.
489
490 **Competing interests**
491 The authors declare no competing interests.
492
493 **Acknowledgements**
494 This research was supported in part by the Intramural Research Program of the National Institutes
495 of Health (NIH). The contributions of the NIH author(s) are considered Works of the United States
496 Government. The findings and conclusions presented in this paper are those of the author(s) and
497 do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human
498 Services, or the United States government. We gratefully acknowledge all data contributors, i.e.,
499 the Authors and their Originating laboratories responsible for obtaining the specimens, and their
500 Submitting laboratories for generating the genetic sequence and metadata and sharing via the
501 GISAID Initiative, on which this research is based. We thank Dr. Joshua L. Cherry (NIH) for
502 providing a curated global N1 dataset to place the clade 2.3.4.4b viruses in phylogenetic context.
503 We also thank Dr. Andrew Ramey (USGS) and Dr. Christina Ahlstrom (USGS) for providing
504 additional metadata on sequences collected by the USGS. GB and SD acknowledge support from
505 the Research Foundation — Flanders (*Fonds voor Wetenschappelijk Onderzoek — Vlaanderen*,
506 FWO, Belgium; grant n°G098321N), and from the European Union Horizon 2023 RIA project
507 LEAPS (grant agreement no. 101094685). GB also acknowledges support from the Research
508 Foundation — Flanders (*Fonds voor Wetenschappelijk Onderzoek — Vlaanderen*, FWO,
509 Belgium; grant n° G0E1420N), and from the DURABLE EU4Health project 02/2023-01/2027,
510 which is co-funded by the European Union (call EU4H-2021-PJ4) under Grant Agreement No.
511 101102733. SD also acknowledges support from the *Fonds National de la Recherche Scientifique*
512 (F.R.S.-FNRS, Belgium; including grant n°F.4515.22), and from the University of Brussels (ULB)
513 internal fund. This work is supported by the Centers of Excellence for Influenza Research and
514 Response, National Institute of Allergy and Infectious Diseases, National Institutes of Health
515 (NIH), Department of Health and Human Services, under contract 75N93021C00014.

References

1. Xie, R., Edwards, K.M., Wille, M., Wei, X., Wong, S.-S., Zanin, M., El-Shesheny, R., Ducatez, M., Poon, L.L.M., Kayali, G., et al. (2023). The episodic resurgence of highly pathogenic avian influenza H5 virus. Nature *622*, 810–817. https://doi.org/10.1038/s41586-023-06631-2.

2. Rijks, J.M., Leopold, M.F., Kühn, S., In ’t Veld, R., Schenk, F., Brenninkmeijer, A., Lilipaly, S.J., Ballmann, M.Z., Kelder, L., de Jong, J.W., et al. (2022). Mass Mortality Caused by Highly Pathogenic Influenza A(H5N1) Virus in Sandwich Terns, the Netherlands, 2022. Emerg Infect Dis *28*, 2538–2542. https://doi.org/10.3201/eid2812.221292.

3. Fusaro, A., Zecchin, B., Giussani, E., Palumbo, E., Agüero-García, M., Bachofen, C., Bálint, Á., Banihashem, F., Banyard, A.C., Beerens, N., et al. (2024). High pathogenic avian influenza A(H5) viruses of clade 2.3.4.4b in Europe—Why trends of virus evolution are more difficult to predict. Virus Evol *10*, veae027. https://doi.org/10.1093/ve/veae027.

4. European Food Safety Authority, European Centre for Disease Prevention and Control, European Union Reference Laboratory for Avian Influenza, Alexakis, L., Buczkowski, H., Ducatez, M., Fusaro, A., Gonzales, J.L., Kuiken, T., Ståhl, K., et al. (2025). Avian influenza overview December 2024-March 2025. EFSA J *23*, e9352. https://doi.org/10.2903/j.efsa.2025.9352.

5. Agüero, M., Monne, I., Sánchez, A., Zecchin, B., Fusaro, A., Ruano, M.J., Del Valle Arrojo, M., Fernández-Antonio, R., Souto, A.M., Tordable, P., et al. (2023). Highly pathogenic avian influenza A(H5N1) virus infection in farmed minks, Spain, October 2022. Euro Surveill *28*, 2300001. https://doi.org/10.2807/1560-7917.ES.2023.28.3.2300001.

6. Kareinen, L., Tammiranta, N., Kauppinen, A., Zecchin, B., Pastori, A., Monne, I., Terregino, C., Giussani, E., Kaarto, R., Karkamo, V., et al. (2024). Highly pathogenic avian influenza A(H5N1) virus infections on fur farms connected to mass mortalities of black-headed gulls, Finland, July to October 2023. Euro Surveill *29*, 2400063. https://doi.org/10.2807/1560-7917.ES.2024.29.25.2400063.

7. Giacinti, J.A., Signore, A.V., Jones, M.E.B., Bourque, L., Lair, S., Jardine, C., Stevens, B., Bollinger, T., Goldsmith, D., British Columbia Wildlife AIV Surveillance Program (BC WASP), et al. (2024). Avian influenza viruses in wild birds in Canada following incursions of highly pathogenic H5N1 virus from Eurasia in 2021–2022. mBio *15*, e03203-23. https://doi.org/10.1128/mbio.03203-23.

8. Youk, S., Torchetti, M.K., Lantz, K., Lenoch, J.B., Killian, M.L., Leyson, C., Bevins, S.N., Dilione, K., Ip, H.S., Stallknecht, D.E., et al. (2023). H5N1 highly pathogenic avian influenza clade 2.3.4.4b in wild and domestic birds: Introductions into the United States and reassortments, December 2021-April 2022. Virology *587*, 109860. https://doi.org/10.1016/j.virol.2023.109860.

9. Leguia, M., Garcia-Glaessner, A., Muñoz-Saavedra, B., Juarez, D., Barrera, P., Calvo-Mac, C., Jara, J., Silva, W., Ploog, K., Amaro, Lady, et al. (2023). Highly pathogenic avian influenza A (H5N1) in marine mammals and seabirds in Peru. Nat Commun *14*, 5489. https://doi.org/10.1038/s41467-023-41182-0.

10. Uhart, M.M., Vanstreels, R.E.T., Nelson, M.I., Olivera, V., Campagna, J., Zavattieri, V., Lemey, P., Campagna, C., Falabella, V., and Rimondi, A. (2024). Epidemiological data of an influenza A/H5N1 outbreak in elephant seals in Argentina indicates mammal-to-mammal transmission. Nat Commun *15*, 9516. https://doi.org/10.1038/s41467-024-53766-5.

11. Tomás, G., Marandino, A., Panzera, Y., Rodríguez, S., Wallau, G.L., Dezordi, F.Z., Pérez, R., Bassetti, L., Negro, R., Williman, J., et al. (2024). Highly pathogenic avian influenza H5N1 virus infections in pinnipeds and seabirds in Uruguay: Implications for bird-mammal transmission in South America. Virus Evol *10*, veae031. https://doi.org/10.1093/ve/veae031.

12. Pardo-Roa, C., Nelson, M.I., Ariyama, N., Aguayo, C., Almonacid, L.I., Gonzalez-Reiche, A.S., Muñoz, G., Ulloa, M., Ávila, C., Navarro, C., et al. (2025). Cross-species and mammal-to-mammal transmission of clade 2.3.4.4b highly pathogenic avian influenza A/H5N1 with PB2 adaptations. Nat Commun *16*, 2232. https://doi.org/10.1038/s41467-025-57338-z.

13. Detwiler, L. H5N1 Highly Pathogenic Avian Influenza (HPAI) in Livestock - Information for small ruminant (sheep and goat) and camelid stakeholders (Animal and Plant Health Inspection Service).

14. Worobey, M., Gangavarapu, K., Pekar, J.E., Joy, J.B., Moncla, L., Kraemer, M.U.G., Dudas, G., Goldhill, D., Ruis, C., Serrano, L.M., et al. (2024). Preliminary report on genomic epidemiology of the 2024 H5N1 influenza A virus outbreak in U.S. cattle (Part 1 of 2) - Influenza virus H5N1-global. Preprint at Virological.

15. Burrough, E.R., Magstadt, D.R., Petersen, B., Timmermans, S.J., Gauger, P.C., Zhang, J., Siepker, C., Mainenti, M., Li, G., Thompson, A.C., et al. (2024). Highly Pathogenic Avian Influenza A(H5N1) Clade 2.3.4.4b Virus Infection in Domestic Dairy Cattle and Cats, United States, 2024. Emerg Infect Dis *30*, 1335–1343. https://doi.org/10.3201/eid3007.240508.

16. Nguyen, T.-Q., Hutter, C., Markin, A., Thomas, M., Lantz, K., Killian, M.L., Janzen, G.M., Vijendran, S., Wagle, S., Inderski, B., et al. (2024). Emergence and interstate spread of highly pathogenic avian influenza A(H5N1) in dairy cattle. Preprint at Microbiology, https://doi.org/10.1101/2024.05.01.591751 https://doi.org/10.1101/2024.05.01.591751.

17. Pekar, J.E., Gangavarapu, K., Crespo-Bellido, A., Peacock, T.P., Wertheim, J.O., Dudas, G., Joy, J.B., Chand, M., Debarre, F., Gangavarapu, P., et al. (2026). The emergence and molecular evolution of H5N1 influenza viruses in United States dairy cattle. bioRxiv. https://doi.org/10.64898/2026.03.30.713641.

18. Jassem, A.N., Roberts, A., Tyson, J., Zlosnik, J.E.A., Russell, S.L., Caleta, J.M., Eckbo, E.J., Gao, R., Chestley, T., Grant, J., et al. (2025). Critical Illness in an Adolescent with Influenza A(H5N1) Virus Infection. New England Journal of Medicine *392*, 927–929. https://doi.org/10.1056/NEJMc2415890.

19. CDC (2025). Genetic Sequences of Highly Pathogenic Avian Influenza A(H5N1) Viruses Identified in a Person in Louisiana. Avian Influenza (Bird Flu). https://www.cdc.gov/bird-flu/spotlights/h5n1-response-12232024.html.

20. World Health Organization (2025). Avian Influenza A(H5N1) - Mexico. https://www.who.int/emergencies/disease-outbreak-news/item/2025-DON564.

21. Dellicour, S., Lequime, S., Vrancken, B., Gill, M.S., Bastide, P., Gangavarapu, K., Matteson, N.L., Tan, Y., du Plessis, L., Fisher, A.A., et al. (2020). Epidemiological hypothesis testing using a phylogeographic and phylodynamic framework. Nat Commun *11*, 5620. https://doi.org/10.1038/s41467-020-19122-z.

22. Damodaran, L., Jaeger, A.S., and Moncla, L.H. (2025). Ecology and spread of the North American H5N1 epizootic. Nature, 1–10. https://doi.org/10.1038/s41586-025-09737-x.

23. Clessin, A., Brusselmans, M., Hong, S.L., Tornos, J., Lejeune, M., Shao, Y., Briand, F.-X., Abolnik, C., Kaza, B., Suchard, M.A., et al. (2026). Dispersal, adaptation and persistence of H5N1 in the sub-Antarctic and Antarctica. bioRxiv. https://doi.org/10.64898/2026.03.20.713283.

24. Bloom, J.D., Gong, L.I., and Baltimore, D. (2010). Permissive Secondary Mutations Enable the Evolution of Influenza Oseltamivir Resistance. Science *328*, 1272–1275. https://doi.org/10.1126/science.1187816.

25. Hansen, L., McMahon, M., Turner, H.L., Zhu, X., Turner, J.S., Ozorowski, G., Stadlbauer, D., Vahokoski, J., Schmitz, A.J., Rizk, A.A., et al. (2023). Human anti-N1 monoclonal antibodies elicited by pandemic H1N1 virus infection broadly inhibit HxN1 viruses in vitro and in vivo. Immunity *56*, 1927-1938.e8. https://doi.org/10.1016/j.immuni.2023.07.004.

26. CDC (2026). A(H5) Bird Flu: Current Situation. Avian Influenza (Bird Flu). https://www.cdc.gov/bird-flu/situation-summary/index.html.

27. Ramey, A.M., Scott, L.C., Ahlstrom, C.A., Buck, E.J., Williams, A.R., Kim Torchetti, M., Stallknecht, D.E., and Poulson, R.L. (2024). Molecular detection and characterization of highly pathogenic H5N1 clade 2.3.4.4b avian influenza viruses among hunter-harvested wild birds provides evidence for three independent introductions into Alaska. Virology *589*, 109938. https://doi.org/10.1016/j.virol.2023.109938.

28. Signore, A.V., Giacinti, J., Jones, M.E.B., Erdelyan, C.N.G., McLaughlin, A., Alkie, T.N., Cox, S., Lair, S., Jardine, C.M., Stevens, B., et al. (2025). Spatiotemporal reconstruction of the North American A(H5N1) outbreak reveals successive lineage replacements by descendant reassortants. Science Advances *11*, eadu4909. https://doi.org/10.1126/sciadv.adu4909.

29. Lowen, A.C., Baker, A.L., Bowman, A.S., García-Sastre, A., Hensley, S.E., Lakdawala, S.S., Moncla, L.H., Nelson, M.I., Pekosz, A., Poulson, R.L., et al. (2025). Pandemic risk stemming from the bovine H5N1 outbreak: an account of the knowns and unknowns. J Virol *99*, e0005225. https://doi.org/10.1128/jvi.00052-25.

30. Rahman, I., Erdelyan, C.N.G., Signore, A.V., Ayilara, I., Wight, J., Jones, M.E.B., Sullivan, D.S., Lung, O., Hisanaga, T., Wilhelm, S.I., et al. (2024). Recurring incursions and dissemination of novel Eurasian-origin H5Nx avian influenza viruses in Atlantic Canada. Virus Evol *10*, veae111. https://doi.org/10.1093/ve/veae111.

31. Elbe, S., and Buckland-Merrett, G. (2017). Data, disease and diplomacy: GISAID's innovative contribution to global health. Glob Chall *1*, 33–46. https://doi.org/10.1002/gch2.1018.

32. Shu, Y., and McCauley, J. (2017). GISAID: Global initiative on sharing all influenza data – from vision to reality. Euro Surveill *22*, 30494. https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494.

33. Khare, S., Gurry, C., Freitas, L., Schultz, M.B., Bach, G., Diallo, A., Akite, N., Ho, J., Lee, R.T., Yeo, W., et al. (2021). GISAID's Role in Pandemic Response. CCDCW *3*, 1049–1051. https://doi.org/10.46234/ccdcw2021.255.

34. Hatcher, E.L., Zhdanov, S.A., Bao, Y., Blinkova, O., Nawrocki, E.P., Ostapchuck, Y., Schäffer, A.A., and Brister, J.R. (2017). Virus Variation Resource - improved response to emergent viral outbreaks. Nucleic Acids Res *45*, D482–D490. https://doi.org/10.1093/nar/gkw1065.

35. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. Journal of Molecular Biology *215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

36. Katoh, K., and Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform *9*, 286–298. https://doi.org/10.1093/bib/bbn013.

37. Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

38. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution *37*, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

39. Li, X., Trovão, N.S., Wertheim, J.O., Baele, G., and de Bernardi Schneider, A. (2023). Optimizing ancestral trait reconstruction of large HIV Subtype C datasets through multiple-trait subsampling. Virus Evol *9*, vead069. https://doi.org/10.1093/ve/vead069.

40. Xu, S., Li, L., Luo, X., Chen, M., Tang, W., Zhan, L., Dai, Z., Lam, T.T., Guan, Y., and Yu, G. (2022). Ggtree: A serialized data object for visualization of a phylogenetic tree and annotation data. iMeta *1*, e56. https://doi.org/10.1002/imt2.56.

41. Vences, M., Patmanidis, S., Kharchev, V., and Renner, S.S. (2022). Concatenator, a user-friendly program to concatenate DNA sequences, implementing graphical user interfaces for MAFFT and FastTree. Bioinformatics Advances *2*, vbac050. https://doi.org/10.1093/bioadv/vbac050.

42. Dellicour, S., Baele, G., Dudas, G., Faria, N.R., Pybus, O.G., Suchard, M.A., Rambaut, A., and Lemey, P. (2018). Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak. Nat Commun *9*, 2222. https://doi.org/10.1038/s41467-018-03763-2.

43. Baele, G., Ji, X., Hassler, G.W., McCrone, J.T., Shao, Y., Zhang, Z., Holbrook, A.J., Lemey, P., Drummond, A.J., Rambaut, A., et al. (2025). BEAST X for Bayesian phylogenetic, phylogeographic and phylodynamic inference. Nat Methods *22*, 1653–1656. https://doi.org/10.1038/s41592-025-02751-x.

44. Ayres, D.L., Cummings, M.P., Baele, G., Darling, A.E., Lewis, P.O., Swofford, D.L., Huelsenbeck, J.P., Lemey, P., Rambaut, A., and Suchard, M.A. (2019). BEAGLE 3: Improved Performance, Scaling, and Usability for a High-Performance Computing Library for Statistical Phylogenetics. Systematic Biology *68*, 1052–1061. https://doi.org/10.1093/sysbio/syz020.

45. Drummond, A.J., W, H.S.Y., J, P.M., and Andrew, R. (2006). Relaxed phylogenetics and dating with confidence. PLoS Biol *4*, e88. https://doi.org/10.1371/journal.pbio.0040088.

46. Gill, M.S., Lemey, P., Faria, N.R., Rambaut, A., Shapiro, B., and Suchard, M.A. (2013). Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol Biol Evol *30*, 713–724. https://doi.org/10.1093/molbev/mss265.

47. Pagel, M., Meade, A., and Barker, D. (2004). Bayesian estimation of ancestral character states on phylogenies. Syst Biol *53*, 673–684. https://doi.org/10.1080/10635150490522232.

48. Lemey, P., Rambaut, A., Welch, J.J., and Suchard, M.A. (2010). Phylogeography takes a relaxed random walk in continuous space and time. Mol Biol Evol *27*, 1877–1885. https://doi.org/10.1093/molbev/msq067.

49. Pybus, O.G., A, S.M., Philippe, L., J, B.F., Andrew, R., W, C.F., R, G.R., Nimalan, A., L, S.S., P, B.M., et al. (2012). Unifying the spatial epidemiology and molecular evolution of emerging epidemics. Proc Natl Acad Sci U S A *109*, 15066–15071. https://doi.org/10.1073/pnas.1206598109.

50. Dellicour, S., Durkin, K., Hong, S.L., Vanmechelen, B., Martí-Carreras, J., Gill, M.S., Meex, C., Bontems, S., André, E., Gilbert, M., et al. (2021). A Phylodynamic Workflow to Rapidly Gain Insights into the Dispersal History and Dynamics of SARS-CoV-2 Lineages. Mol Biol Evol *38*, 1608–1613. https://doi.org/10.1093/molbev/msaa284.

51. Dellicour, S., Lemey, P., Artois, J., Lam, T.T., Fusaro, A., Monne, I., Cattoli, G., Kuznetsov, D., Xenarios, I., Dauphin, G., et al. (2020). Incorporating heterogeneous sampling probabilities in continuous phylogeographic inference — Application to H5N1 spread in the Mekong region. Bioinformatics *36*, 2098–2104. https://doi.org/10.1093/bioinformatics/btz882.

52. Dellicour, S., Lemey, P., Suchard, M.A., Gilbert, M., and Baele, G. (2022). Accommodating sampling location uncertainty in continuous phylogeography. Virus Evol *8*, veac041. https://doi.org/10.1093/ve/veac041.

53. Lemey, P., Andrew, R., J, D.A., and A, S.M. (2009). Bayesian phylogeography finds its roots. PLoS Comput Biol *5*, e1000520. https://doi.org/10.1371/journal.pcbi.1000520.

54. Rambaut, A., Drummond, A.J., Xie, D., Baele, G., and Suchard, M.A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Syst Biol *67*, 901–904. https://doi.org/10.1093/sysbio/syy032.

55. Baele, G., Carvalho, L.M., Brusselmans, M., Dudas, G., Ji, X., McCrone, J.T., Lemey, P., Suchard, M.A., and Rambaut, A. (2025). HIPSTR: highest independent posterior subtree reconstruction in TreeAnnotator X. Bioinformatics, btaf488. https://doi.org/10.1093/bioinformatics/btaf488.

715 56. Dellicour, S., Rose, R., Faria, N.R., Lemey, P., and Pybus, O.G. (2016). SERAPHIM:
716 studying environmental rasters and phylogenetically informed movements. Bioinformatics *32*,
717 3204–3206. https://doi.org/10.1093/bioinformatics/btw384.

718 57. Trovão, N.S., Khan, S.M., Lemey, P., Nelson, M.I., and Cherry, J.L. (2024). Comparative
719 evolution of influenza A virus H1 and H3 head and stalk domains across host species. mBio
720 *15*, e0264923. https://doi.org/10.1128/mbio.02649-23.

721 58. Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large
722 datasets. Bioinformatics *30*, 3276–3278. https://doi.org/10.1093/bioinformatics/btu531.

723 59. Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum
724 evolution trees with profiles instead of a distance matrix. Mol Biol Evol *26*, 1641–1650.
725 https://doi.org/10.1093/molbev/msp077.

726 60. Rambaut, A., Lam, T.T., Max Carvalho, L., and Pybus, O.G. (2016). Exploring the
727 temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen).
728 Virus Evol *2*, vew007. https://doi.org/10.1093/ve/vew007.

729 61. Nakajima, K., Desselberger, U., and Palese, P. (1978). Recent human influenza A (H1N1)
730 viruses are closely related genetically to strains isolated in 1950. Nature *274*, 334–339.
731 https://doi.org/10.1038/274334a0.

732 62. Kendal, A.P., Noble, G.R., Skehel, J.J., and Dowdle, W.R. (1978). Antigenic similarity of
733 influenza A(H1N1) viruses from epidemics in 1977–1978 to "Scandinavian" strains isolated in
734 epidemics of 1950–1951. Virology *89*, 632–636. https://doi.org/10.1016/0042-
735 6822(78)90207-6.

736 63. Wertheim, J.O. (2010). The Re-Emergence of H1N1 Influenza Virus in 1977: A Cautionary
737 Tale for Estimating Divergence Times Using Biologically Unrealistic Sampling Dates. PLOS
738 ONE *5*, e11184. https://doi.org/10.1371/journal.pone.0011184.

739 64. Carter, R.W., and Sanford, J.C. (2012). A new look at an old virus: patterns of mutation
740 accumulation in the human H1N1 influenza virus since 1918. Theoretical Biology and Medical
741 Modelling *9*, 42. https://doi.org/10.1186/1742-4682-9-42.

742 65. Aksamentov, I., Roemer, C., Hodcroft, E., and Neher, R. (2021). Nextclade: clade
743 assignment, mutation calling and quality control for viral genomes. JOSS *6*, 3773.
744 https://doi.org/10.21105/joss.03773.

745 66. Giussani, E., Sartori, A., Salomoni, A., Cavicchio, L., de Battisti, C., Pastori, A., Varotto,
746 M., Zecchin, B., Hughes, J., Monne, I., et al. (2025). FluMut: a tool for mutation surveillance
747 in highly pathogenic H5N1 genomes. Virus Evol *11*, veaf011.
748 https://doi.org/10.1093/ve/veaf011.

749 **SUPPLEMENTARY MATERIAL**

750 **TABLES**

751 **Table S1. List of H5Nx clade 2.3.4.4b with NA segments acquired through reassortment with North**
752 **American LPAI.** Genotype D1.1 and its major genotype descendants are in bold. GenoFLU-assigned
753 lineages

Genotype	Ancestry	Subtype	PB2	PB1	PA	HA	NP	NA	MP	NS	# of sequences
B3.14	B3.2 descendant	H5N2	am2.1	mx1	ea1	ea1	mx1	mx1N2	ea1	am1.1	1
D1.1	**D1.1 descendant**	**H5N1 (am-NA)**	**am24**	**ea3**	**am4**	**ea3**	**am13**	**am4N1**	**ea3**	**ea3**	**5257**
D1.2	**D1.1 descendant**	**H5N1 (am-NA)**	**am24**	**ea3**	**am5**	**ea3**	**am13**	**am4N1**	**ea3**	**ea3**	**5**
D1.4	**D1.1 descendant**	**H5N1 (am-NA)**	**am3.2**	**ea3**	**am4**	**ea3**	**am13**	**am4N1**	**ea3**	**ea3**	**14**
D1.5	**D1.1 descendant**	**H5N1 (am-NA)**	**am24**	**ea3**	**am4**	**ea3**	**am13**	**am10N1**	**ea3**	**ea3**	**5**
D1.6	**D1.1 descendant**	**H5N2**	**am2.2**	**am22**	**am4**	**ea3**	**am13**	**am11N2**	**ea3**	**ea3**	**5**
D1.7	**D1.1 descendant**	**H5N1 (am-NA)**	**am24**	**ea3**	**am12**	**ea3**	**am16**	**am4N1**	**ea3**	**ea3**	**10**
D1.8	**D1.1 descendant**	**H5N1 (am-NA)**	**am24**	**ea3**	**am4**	**ea3**	**am1.1**	**am4N1**	**ea3**	**ea3**	**5**
D1.9	**D1.1 descendant**	**H5N1 (am-NA)**	**am2.2**	**ea3**	**am14**	**ea3**	**am13**	**am4N1**	**ea3**	**ea3**	**12**
D1.10	**D1.1 descendant**	**H5N1 (am-NA)**	**am2.2**	**ea3**	**am12**	**ea3**	**am1.6**	**am4N1**	**ea3**	**ea3**	**2**
D1.12	**D1.1 descendant**	**H5N1 (am-NA)**	**am18**	**ea3**	**am12**	**ea3**	**am13**	**am4N1**	**ea3**	**ea3**	**3**
D2.1	**D1.1 descendant**	**H5N2**	**am25**	**ea3**	**am4**	**ea3**	**am1.1**	**am5N2**	**ea3**	**ea3**	**2**
Minor101	D1.1 descendant	H5N9	am24	ea3	am4	ea3	am11	am7N9	ea3	ea3	1
Minor102	D1.1 descendant	H5N5	am1.1	ea3	am8	ea3	am13	am8N5	ea3	ea3	1
Minor107	D1.1 descendant	H5N2	am23	ea3	am4	ea3	am13	am9N2	ea3	ea3	1
Minor110	D1.1 descendant	H5N1 (am-NA)	am24	ea3	am8	ea3	am15	am4N1	ea3	ea3	1
Minor111	D1.1 descendant	H5N1 (am-NA)	am2.2	ea3	am9	ea3	am13	am4N1	ea3	ea3	12
Minor112	D1.1 descendant	H5N1 (am-NA)	am24	am9	am4	ea3	am13	am4N1	ea3	ea3	1
Minor114	D1.1 descendant	H5N1 (am-NA)	am24	ea3	am9	ea3	am13	am4N1	ea3	ea3	9
Minor120	D1.1 descendant	H5N1 (am-NA)	am2.2	am25	am4	ea3	am13	am4N1	ea3	am1.2	1
Minor124	D1.1 descendant	H5N1 (am-NA)	am24	ea3	am4	ea3	am17	am4N1	ea3	ea3	1
Minor129	D1.1 descendant	H5N1 (am-NA)	am18	ea3	am8	ea3	am9	am4N1	ea3	ea3	1
Minor131	D1.1 descendant	H5N1 (am-NA)	am24	ea3	am12	ea3	am13	am4N1	ea3	ea3	8
Minor133	D1.1 descendant	H5N1 (am-NA)	am24	ea3	am14	ea3	am13	am4N1	ea3	ea3	1
Minor135	D1.1 descendant	H5N1 (am-NA)	am18	ea3	am4	ea3	am13	am4N1	ea3	ea3	1
Minor137	D1.1 descendant	H5N1 (am-NA)	am24	ea3	am4	ea3	am6	am4N1	ea3	ea3	2
Minor16	A1 descendant	H5N4	am2.2	am3	ea1	ea1	am3	am1N4	ea1	ea1	1
Minor36	A1 descendant	H5N5	am1.3	am1.3	am1.2	ea1	am1.2	am2N6	ea1	ea1	1
Minor97	D1.1 descendant	H5N1 (am-NA)	am24	ea3	am4	ea3	ea3	am4N1	ea3	ea3	4
Minor98	A3 descendant	H5N2	am2.1	ea3	am7	ea3	am14	am6N2	ea3	ea3	6
Minor99	D1.1 descendant	H5N1 (am-NA)	am2.2	ea3	am4	ea3	am13	am4N1	ea3	ea3	1

754 **Table S2. Genotype-defining residues among dominant reassortant genotypes.** Only residues found
755 in >90% of all sequences within the genotype but <20% of other major genotypes in the table and <10%
756 of all other clade 2.3.4.4b sequences combined are listed. All residues follow sequential numbering
757 except HA, which follows an H5 numbering scheme. White cells represent Eurasian segment lineages
758 while colored cells indicate lineages acquired from North American LPAI, with cells shaded with the same
759 color indicating a shared origin between genotypes.

Genotype	PB2	PB1	PA	HA	NP	NA	MP		NS	
							M1	M2	NS1	NEP
B3.2	340K, 464M	215K, 429R	--	--	--	--	--	--	--	--
B3.6	251K, 353R, 663R, 667I	--	--	--	230L	--	--	--	--	--
B3.13	109I, 362G, 441N, 495I, 631L, 649I, 676A	--	219I, 497R	--	482N	--	--	--	7L	7L
C2.1	129A, 470N	386K, 618G	104R, 609R	226T, 503Y	353I, 450S	--	--	--	70G, 111M, 114P, 161T, 176D	--
D1.1	--	16D, 154S, 207R, 614D	--	11I, 52A, 120L, 226V, 341R, 491D, 526I	--	**20I, 23V, 44N, 45H, 48T, 53V, 74L, 75I, 81D, 82P, 84A, 221S, 234I, 241I, 257R, 286S, 287E, 288V, 329S, 336G, 338V, 395A**	200A	61R	75G, 147I, 171N, 193Q	36K, 63E

760 **Table S3. D1.1-specific NA segment residues by protein domain.**

761

NA domain (residue position range)	Sequential (N1) numbering	N2 numbering
Tail (1-6)	--	--
Transmembrane (7-29)	I20	I20
	V23	V23
Stalk (30-84)	N44	N44
	H45	H45
	T48	insertion
	V53	V51
	L74	L72
	I75	L73
	D81	D79
	P82	P80
	A84	A82
Head (85+)	S221	S220
	I234	I233
	I241	I240
	R257	R256
	S286	S285
	E287	E286
	V288	V287
	S329	S329
	G336	G339
	V338	V341
	A395	A399

762 **Table S4. Comparison of dispersal metrics estimated for recent H5N1 genotypes having spread in**
763 **North and South America.** We report the weighted diffusion coefficient (WDC) estimated for each of the
764 four dominant H5N1 reassortant genotypes in the Americas (B3.2, B3.6, C2.1, and D1.1). For each dataset,
765 we report both the posterior median estimate and the 95% highest posterior density (HPD) interval. For
766 comparison, we also report the same dispersal metric previously estimated for the West Nile virus spread
767 in North America.

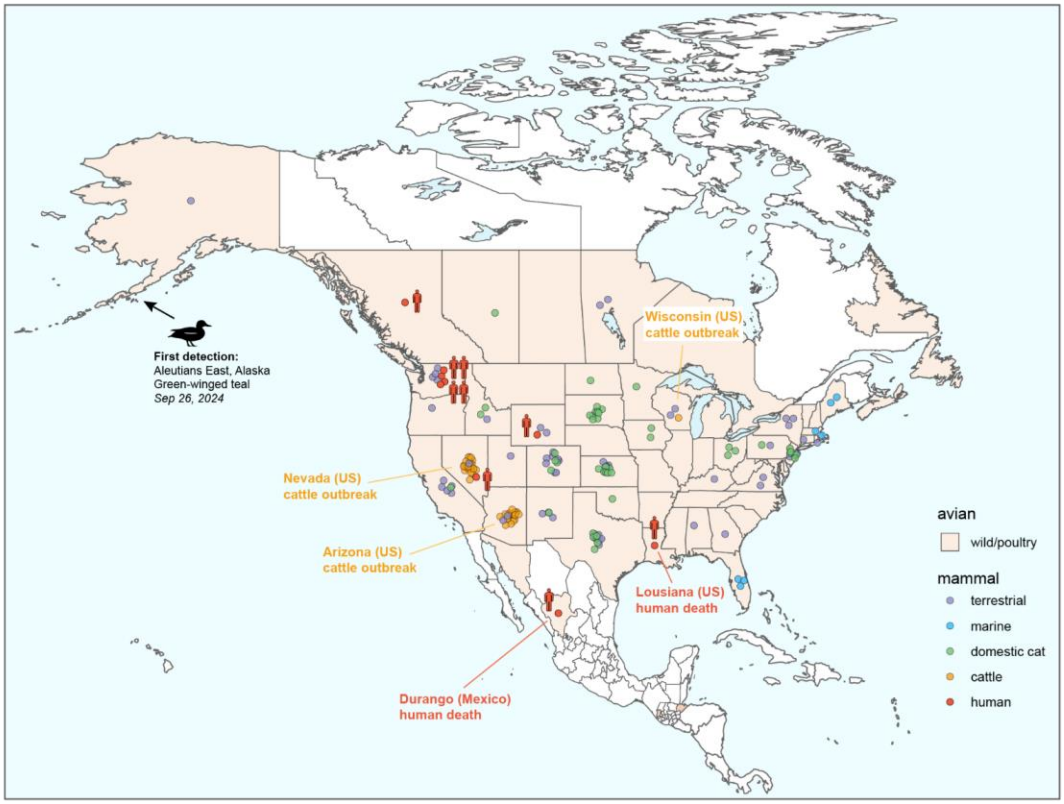
Dataset	Weighted diffusion coefficient (km^2/day)	WDC during invasion (km^2/ day)	WDC after invasion (km^2/ day)
HPAI H5N1 (B3.2)	3030 [2741, 336431]	3418 [3087, 3850]	1908 [1358, 2581]
HPAI H5N1 (B3.6)	2878 [2653, 3126]	3200 [2677, 3884]	2074 [1385, 2794]
HPAI H5N1 (C2.1)	1659 [1376, 1950]	1867 [868, 3475]	536 [296, 815]
HPAI H5N1 (D1.1)	2494 [2268, 2736]	5873 [4248, 7949]	1949 [1711, 2218]
West Nile virus (North America)	161 [151, 170]	312 [280, 373]	119 [107, 135]

768 **Table S5. Equivalent designations between GenoFLU and new global nomenclature system**
769 **ggFLU (https://github.com/YiSong89/ggFlu)**

770

HA clade	Major genotype (ggFLU)	GenoFlu
2.3.4.4b	B.16.1	A1, A2
2.3.4.4b	B.16	A2
2.3.4.4b	G.4	A3
2.3.4.4b	G.6	A3
2.3.4.4b	G.6.2	A3
2.3.4.4b	G.8	A3
2.3.4.4b	B.23.2	A3, A5
2.3.4.4b	B.31	A4
2.3.4.4b	B.24	A6
2.3.4.4b	E.4	B1.1
2.3.4.4b	E.6	B1.2
2.3.4.4b	E.6	B1.3
2.3.4.4b	E.5	B2.1
2.3.4.4b	L.1	B2.2
2.3.4.4b	E.1	B3.1
2.3.4.4b	K.1	B3.10
2.3.4.4b	K.1	B3.11
2.3.4.4b	K.1	B3.12
2.3.4.4b	E.2	B3.13
2.3.4.4b	E.2	B3.2
2.3.4.4b	E.2	B3.3
2.3.4.4b	K.2	B3.4
2.3.4.4b	E.2	B3.5
2.3.4.4b	K.1	B3.6
2.3.4.4b	E.2	B3.7
2.3.4.4b	K.1	B3.8
2.3.4.4b	E.2	B3.9
2.3.4.4b	E.1	B4.1
2.3.4.4b	E.3	B5.1, Minor09
2.3.4.4b	E.12	C1.1
2.3.4.4b	E.13	C2.1
2.3.4.4b	E.15	C3.1
2.3.4.4b	G.6.5	D1.1
2.3.4.4b	P.4	D1.2
2.3.4.4b	P.5	D1.3
2.3.4.4b	P.6	D2.1

FIGURES



B

Virus name	Collection date	Location	Source	Disease severity	Age	Acquired mammalian-adaptive PB2 mutations	Accession (GISAID)
A/Washington/240/2024	10/18/2024	USA-WA	commercial poultry	mild (unspecified)	adult (>18)	--	EPI_ISL_19531299
A/Washington/UW32495/2024	10/18/2024	USA-WA	commercial poultry	mild (unspecified)	adult (>18)	--	EPI_ISL_19541397
A/Washington/254/2024	10/23/2024	USA-WA	commercial poultry	mild (unspecified)	adult (>18)	--	EPI_ISL_19666173
A/Washington/255/2024	10/29/2024	USA-WA	commercial poultry	mild (unspecified)	adult (>18)	--	EPI_ISL_19552697
A/British Columbia/PHL-2032/2024	11/9/2024	Canada-BC	unknown	hospitalization (with conjunctivitis)	teen (13-18)	E627K	EPI_ISL_19548836
A/Louisiana/12/2024	12/11/2024	USA-LA	backyard poultry	death	adult (>18)	--	EPI_ISL_19634828
A/Nevada/10/2025	2/4/2025	USA-NV	dairy cattle	mid (conjunctivitis)	adult (>18)	D701N	EPI_ISL_19726293
A/Wyoming/01/2025	2/11/2025	USA-WY	backyard poultry	hospitalization	adult (>18)	E627K	EPI_ISL_19749443
A/Mexico/Dur_Indre_2292/2025	3/15/2025	Mexico-Durango	unknown	death	child (<13)	--	EPI_ISL_20227034

Figure 1. H5N1 genotype D1.1 sample distribution across North America, 2024-2025. (A) Map of North America showing the geographic distribution of publicly available D1.1 samples collected in 2024-2025 across the US, Canada and Mexico. Circles denote mammalian samples and are colored according to the host type. The coordinates for each data point are not exact and are jittered around the centroid for each US state, Canadian province and Mexican state. Level 1 administrative regions from where wild and/or poultry avian samples have been collected are shaded in beige while regions where D1.1 has not been detected are shaded in white. (B) Description of D1.1 infections in humans that were sampled and sequenced.

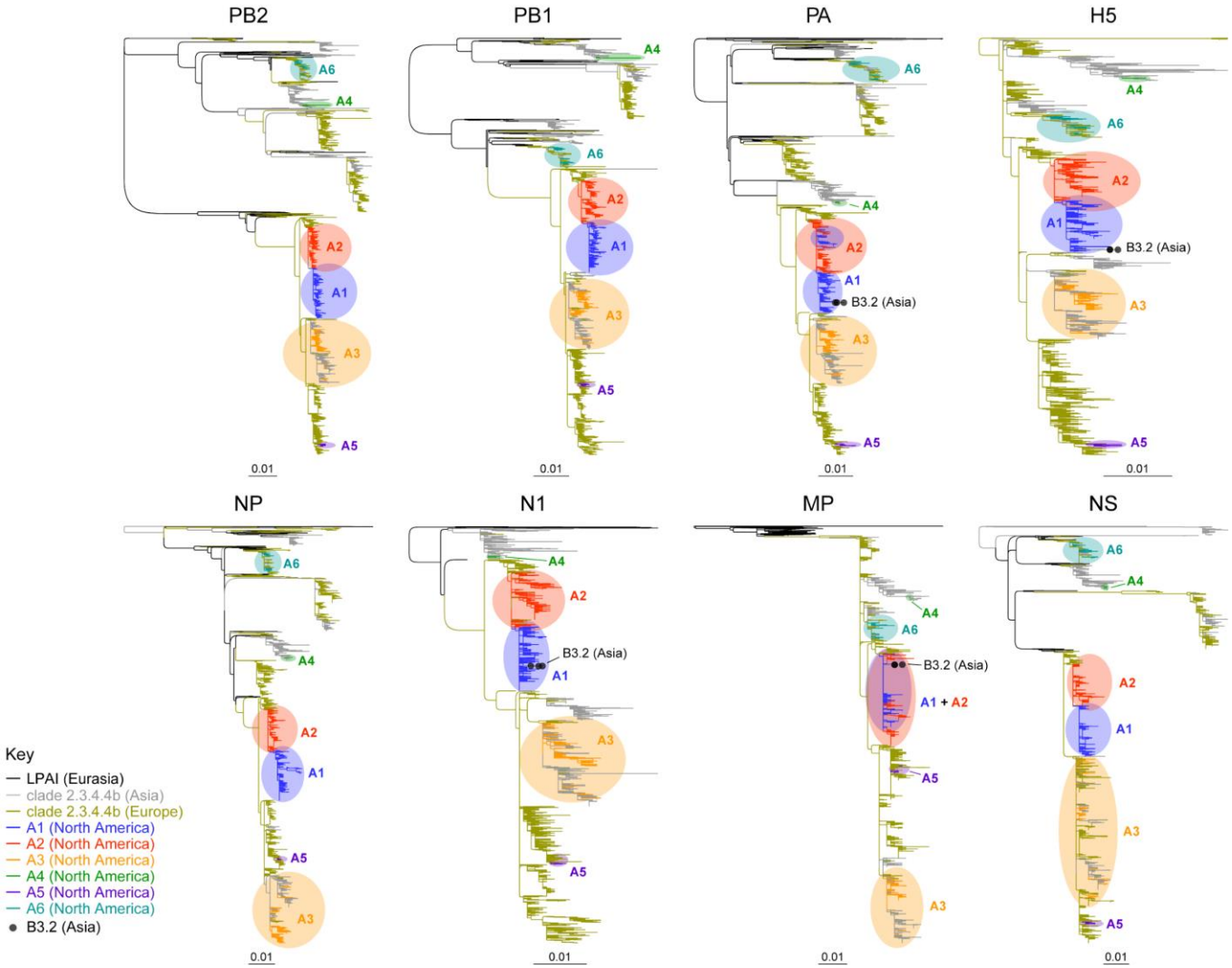


Figure S2. Midpoint-rooted segment phylogenies for EA genotypes (A1-A6) sampled in the Americas, and Eurasian LPAI and clade 2.3.4.4b sequences (collected between January 1, 2015 and June 13, 2025). Tips for North American reassortant genotype B3.2 samples collected in Asia are circled. Full dataset tree files available at https://github.com/nidiatrovao/H5N1Americas.

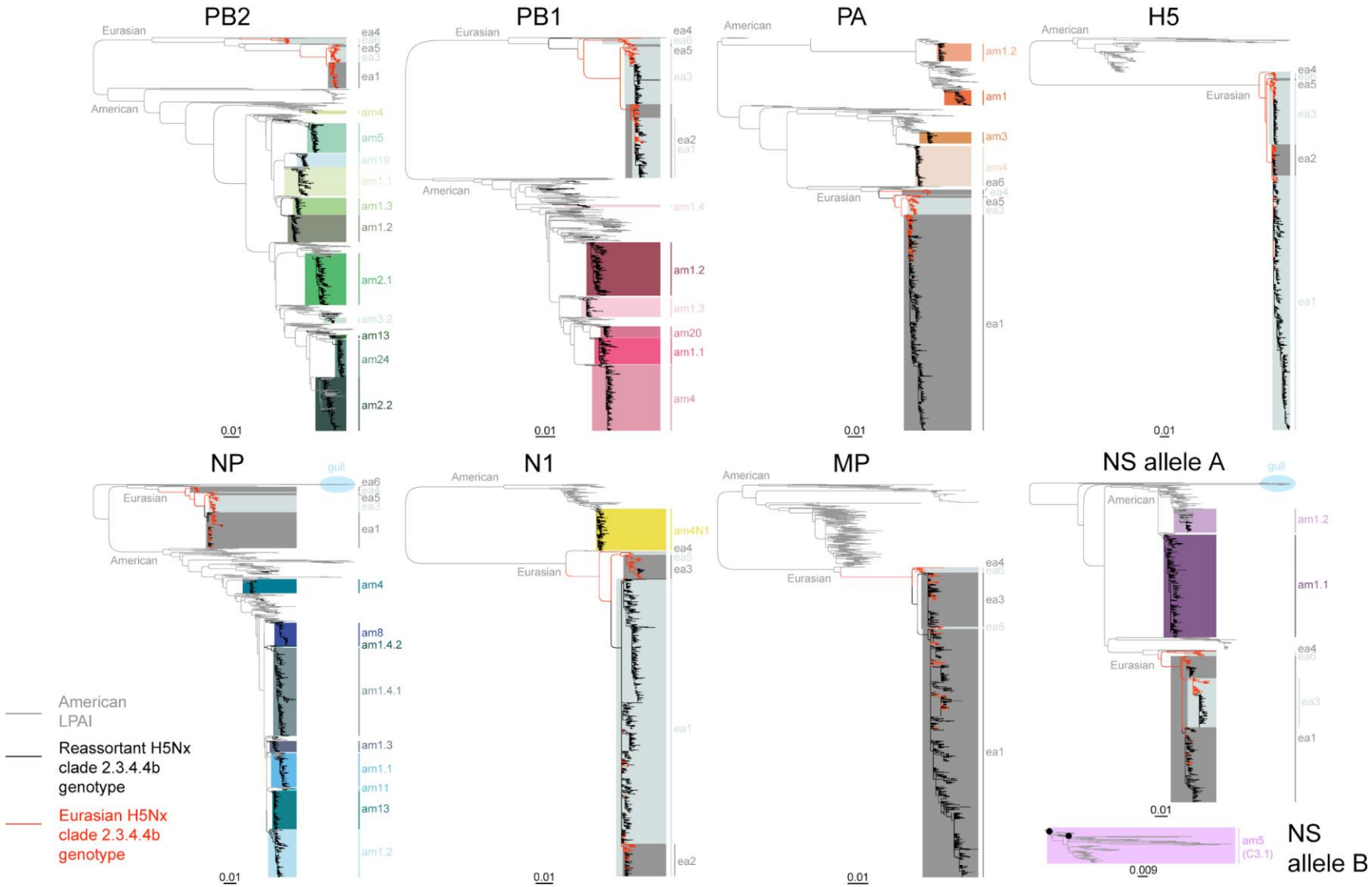


Figure S3. Subsampled, midpoint-rooted segment phylogenies for H5Nx clade 2.3.4.4b and LPAI viruses sampled in the Western hemisphere (collected between January 1, 2015 and June 13, 2025). Full and subsampled dataset tree files are available at https://github.com/nidiatrovao/H5N1Americas

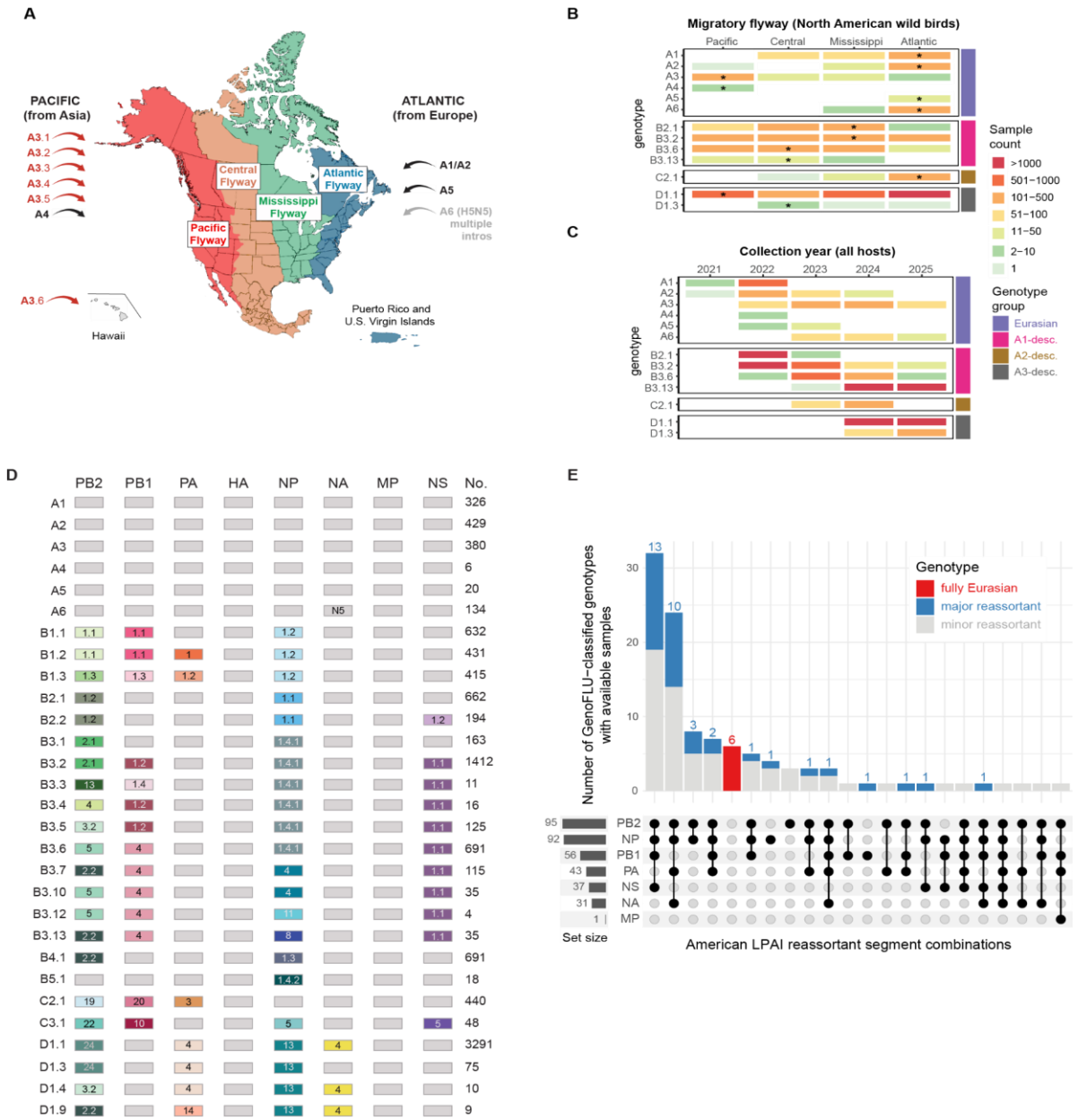


Figure S4. (A) Map of North America's four administrative migratory flyways, defined by the US Fish and Wildlife Service. Arrows indicate independent introductions of Eurasian H5Nx clade 2.3.4.4b genotypes into North American migratory flyways, where black arrows indicate single H5N1 introductions, red arrows indicate repeated introductions of genotype A3, and the grey arrow indicates multiple introductions of H5N5 genotype A6. (B) Heatmap shaded by the number of sequences in each North American migratory flyway for North American wild birds (n = 7,869) and (C) heatmap shaded by collection year across all hosts (n = 17,243). Only select genotypes are shown. Asterisks denote the flyway where the initial detection occurred. (D) Genotype constellation for clade 2.3.4.4b H5N1 major genotypes with >10 sequences available. Grey squares indicate Eurasian lineage; colored squares represent reassorted North American LPAI lineages. The number of genome sequences available from wild birds in the Americas is provided for each genotype in the last column. (E) The y-axis represents the number of H5Nx genotypes with available sequences (defined by GenoFlu; n = 107) with each combination of LPAI segments. Major reassortant genotypes (as defined by GenoFlu, see Methods) are shaded blue; minor reassortant genotypes are shaded grey; non-reassorted fully Eurasian genotypes are shaded red. The count of genotypes with an LPAI-derived segment for each of the eight gene segments is indicated on the left.

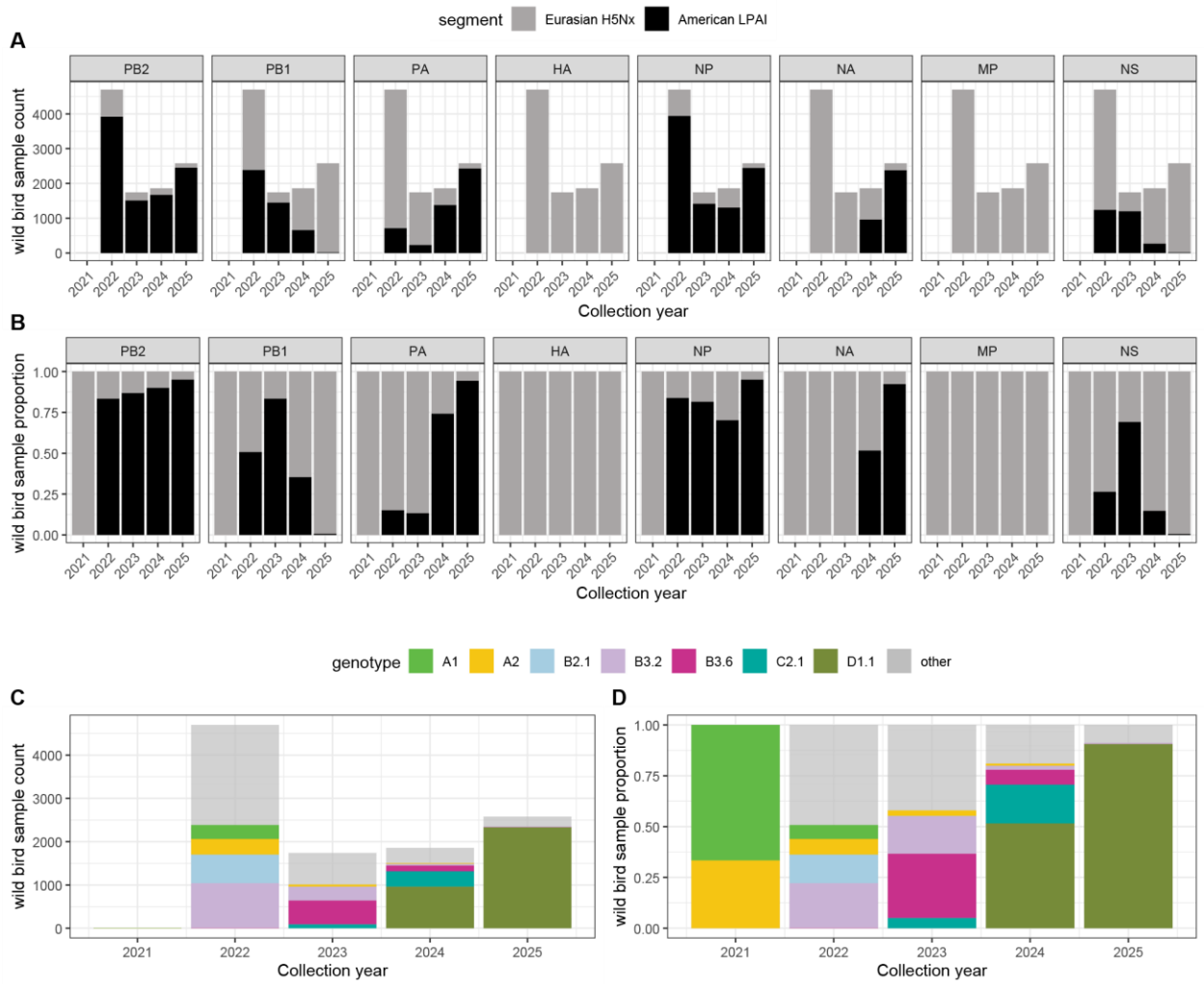


Figure S5. Yearly counts and proportions for H5Nx clade 2.3.4.4b segment sequences during 2021-2025. (A) Yearly counts and (B) proportions for H5Nx clade 2.3.4.4b wild bird segment sequences shaded by Eurasian H5Nx (blue) or American LPAI (black) origin. **(C)** Yearly counts and (D) proportions for H5Nx clade 2.3.4.4b wild bird samples shaded by genotype.

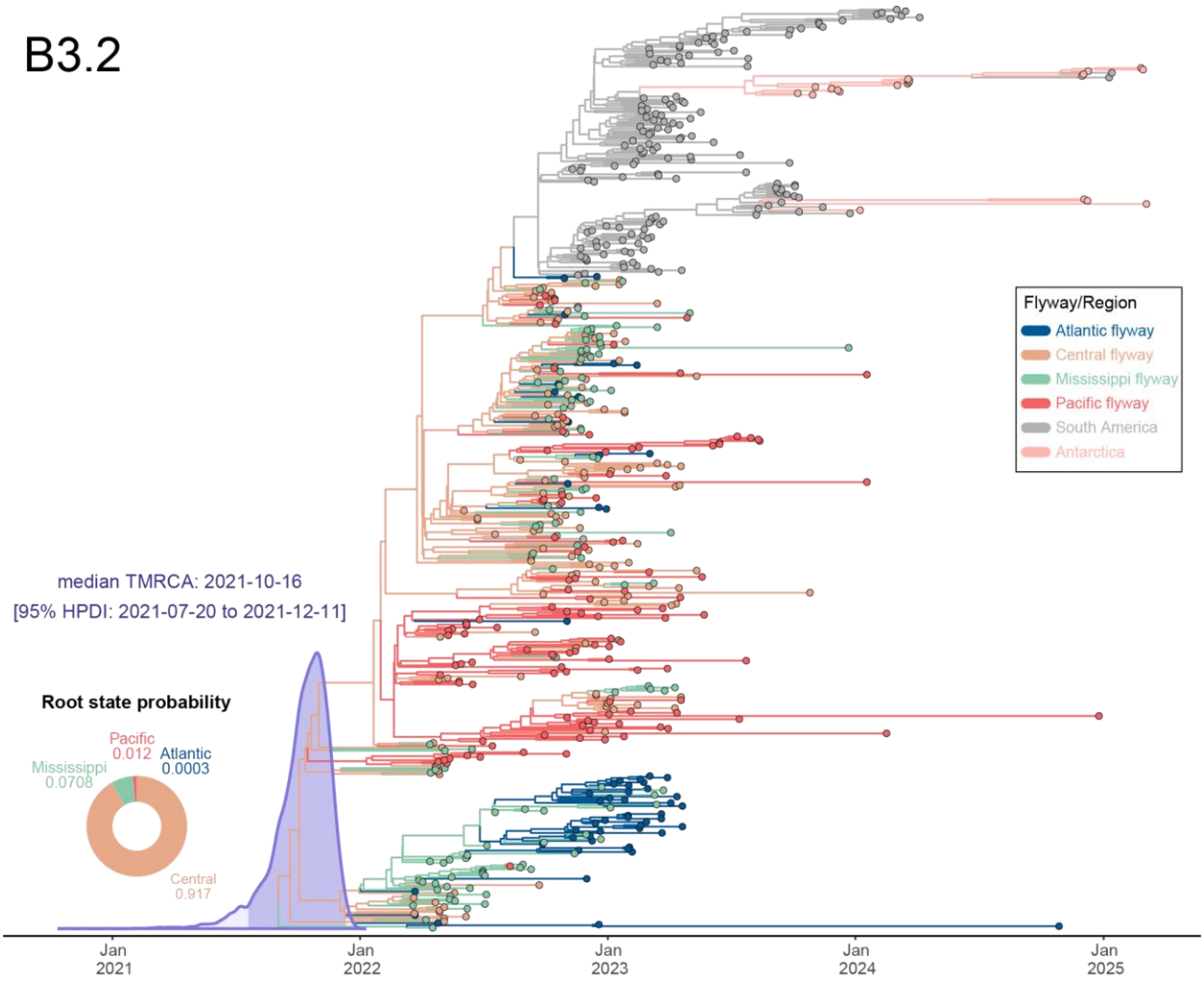


Figure S6. Discrete phylogeographic reconstruction of the dispersal history for H5N1 genotype B3.2 across the North American waterfowl migratory flyways. The consensus phylogenetic tree shows the evolutionary relationships of genotype B3.2 viruses, with the X-axis scaled in years. Tips are shaded according to the region or North American flyway where they were collected, and branches are shaded according to the flyway inferred at ancestral nodes. The purple curve at the root of the tree represents the posterior density distribution for the TMRCA. The donut plot at the base of the phylogeny indicates the posterior flyway root state probabilities.

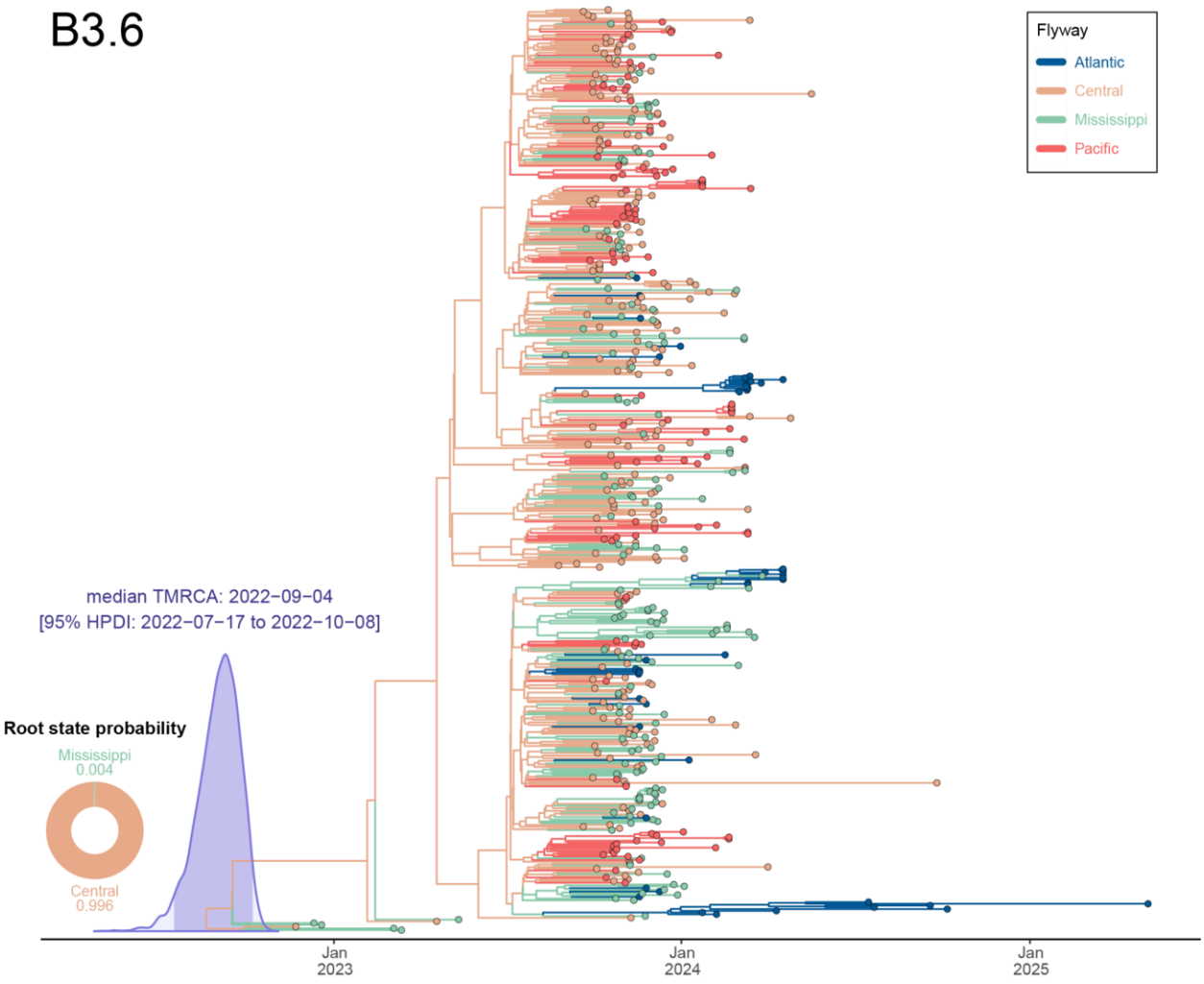


Figure S7. Discrete phylogeographic reconstruction of the dispersal history for H5N1 genotype B3.6 across the North American waterfowl migratory flyways. The consensus phylogenetic tree shows the evolutionary relationships of genotype B3.6 viruses, with the X-axis scaled in years. Tips are shaded according to the flyway where they were collected, and branches are shaded according to the flyway inferred at ancestral nodes. The purple curve at the root of the tree represents the posterior density distribution for the TMRCA. The donut plot at the base of the phylogeny indicates the posterior flyway root state probabilities.

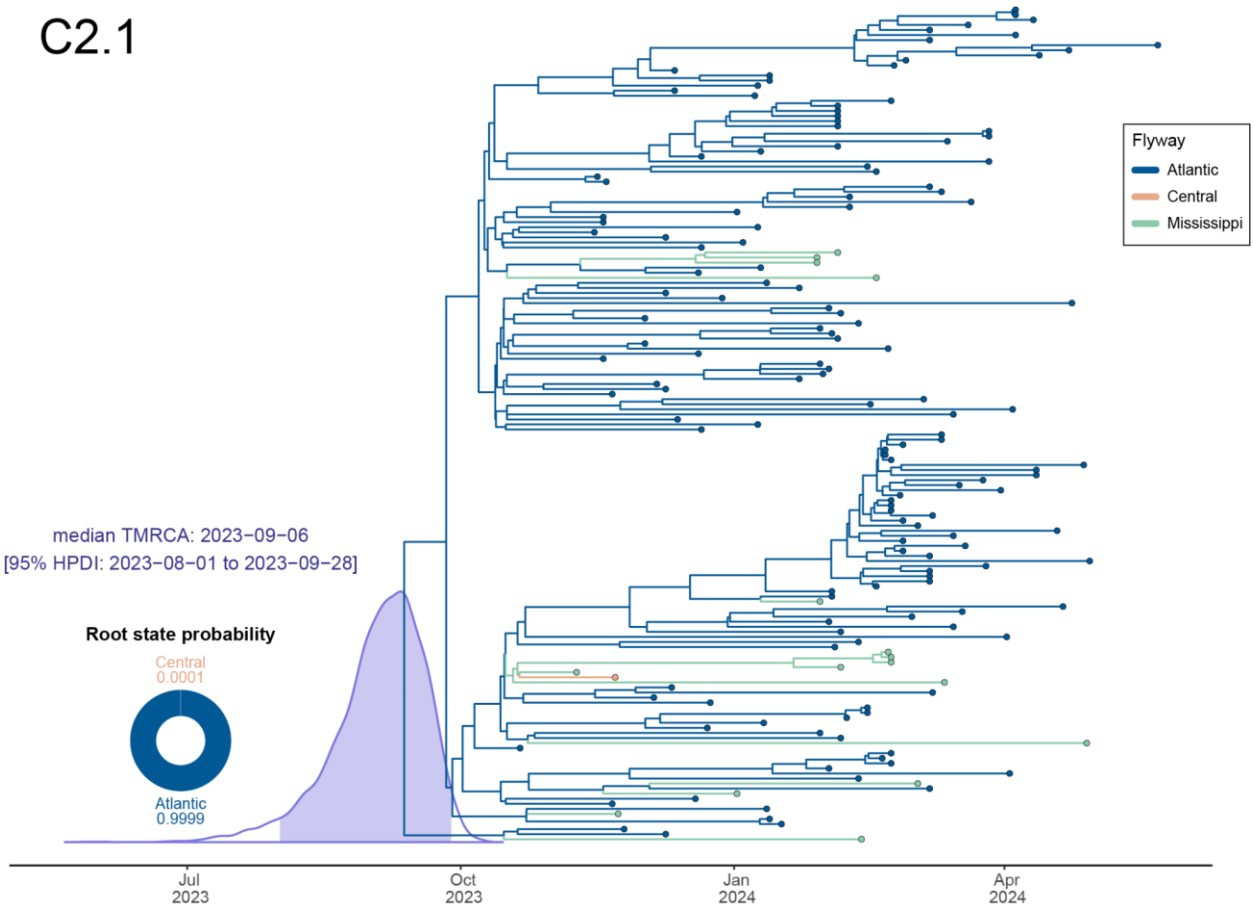
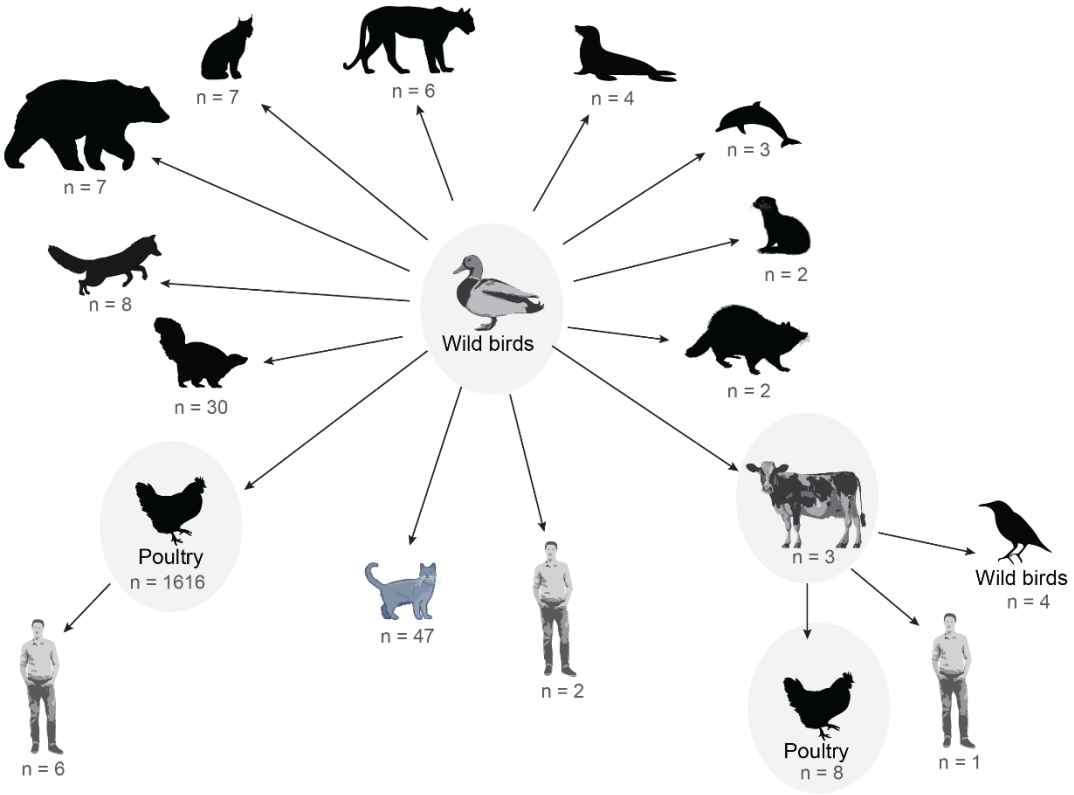


Figure S8. Discrete phylogeographic reconstruction of the dispersal history for H5N1 genotype C2.1 across the North American waterfowl migratory flyways. The consensus phylogenetic tree shows the evolutionary relationships of genotype C2.1 viruses, with the X-axis scaled in years. Tips are shaded according to the flyway where they were collected, and branches are shaded according to the flyway inferred at ancestral nodes. The purple curve at the root of the tree represents the posterior density distribution for the TMRCA.



833

Figure S9. D1.1 transmission routes. Sample sizes represent the number of sequences in each host category.

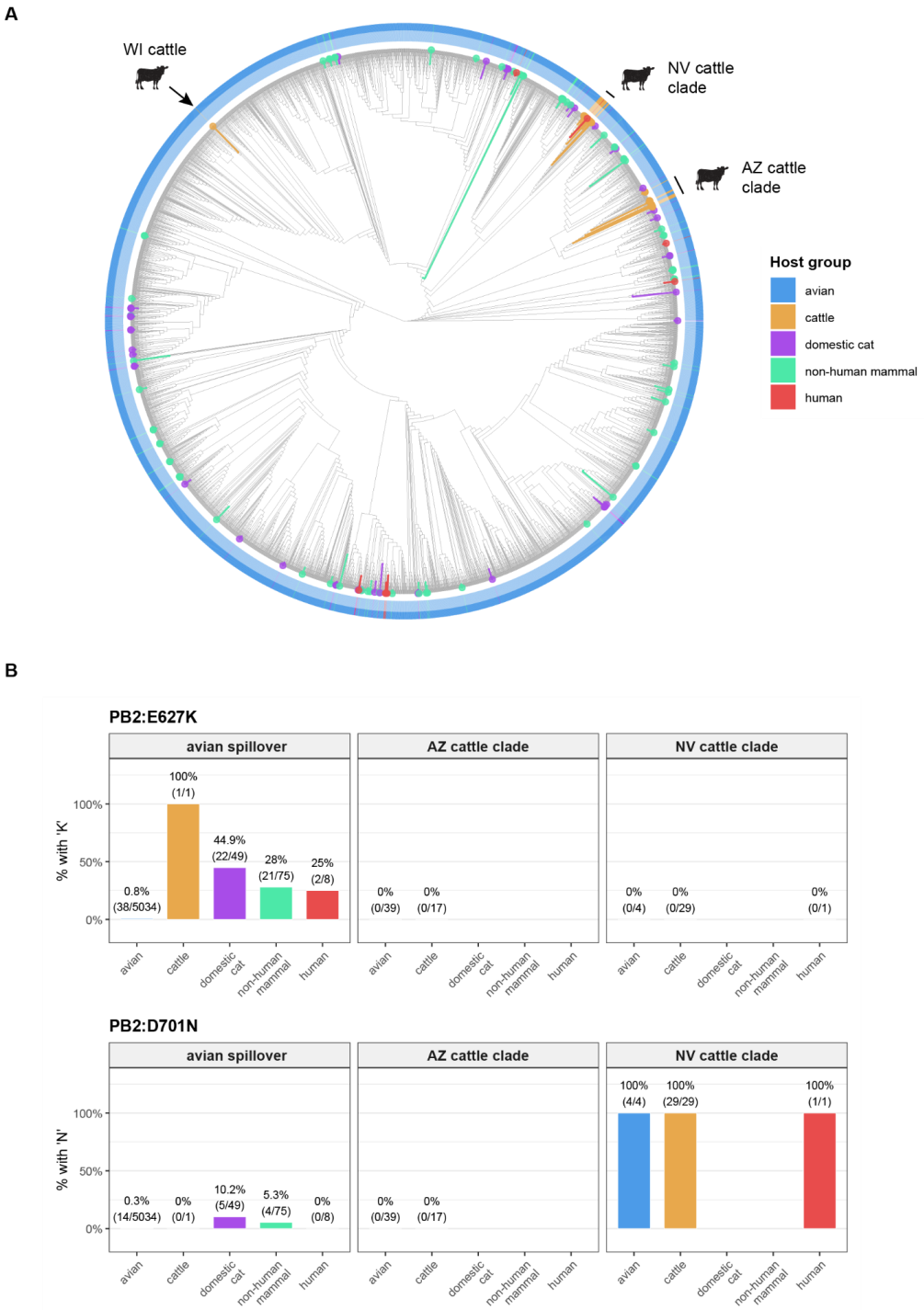


Figure S10. D1.1 spillover. (A) Midpoint-rooted D1.1 ML phylogeny with host group annotations (*n* = 5,257). All three cattle outbreaks are highlighted. (B) Presence of PB2 mammalian-adaptive substitutions E627K and D701N in sequences obtained from different host groups and facetted by transmission route: from cattle (i.e., within "AZ cattle clade" and "NV cattle clade"), and from avian hosts ("avian spillover"). Full PB2 FluMut outputs available at: https://github.com/nidiatrovao/H5N1Americas.

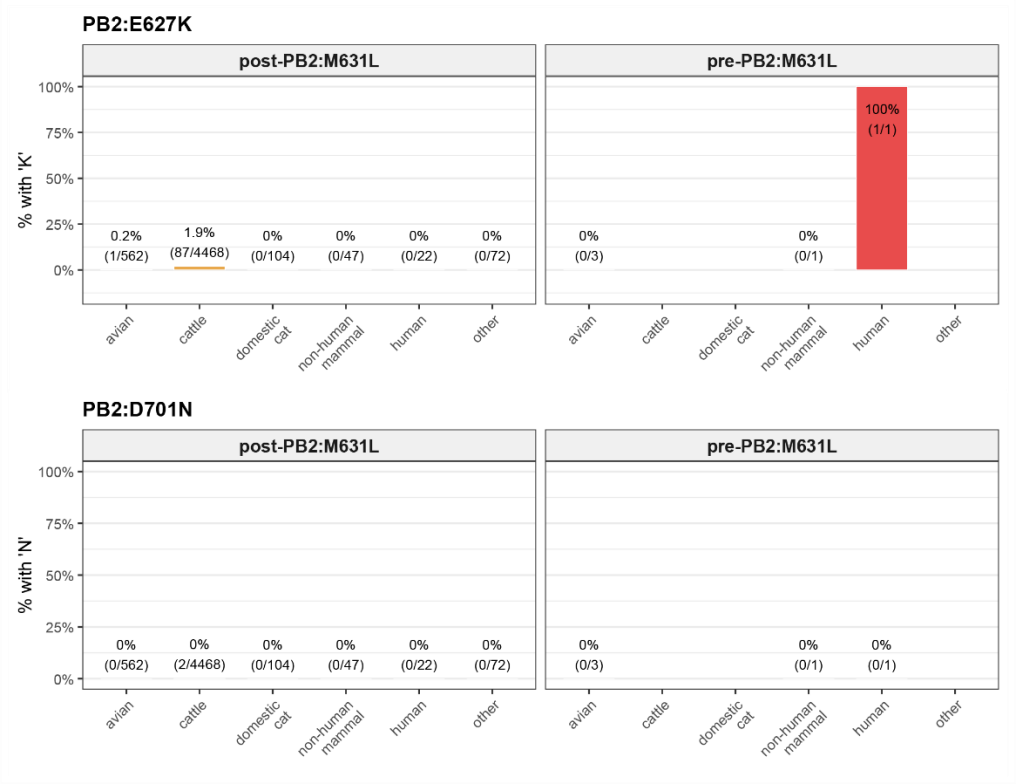
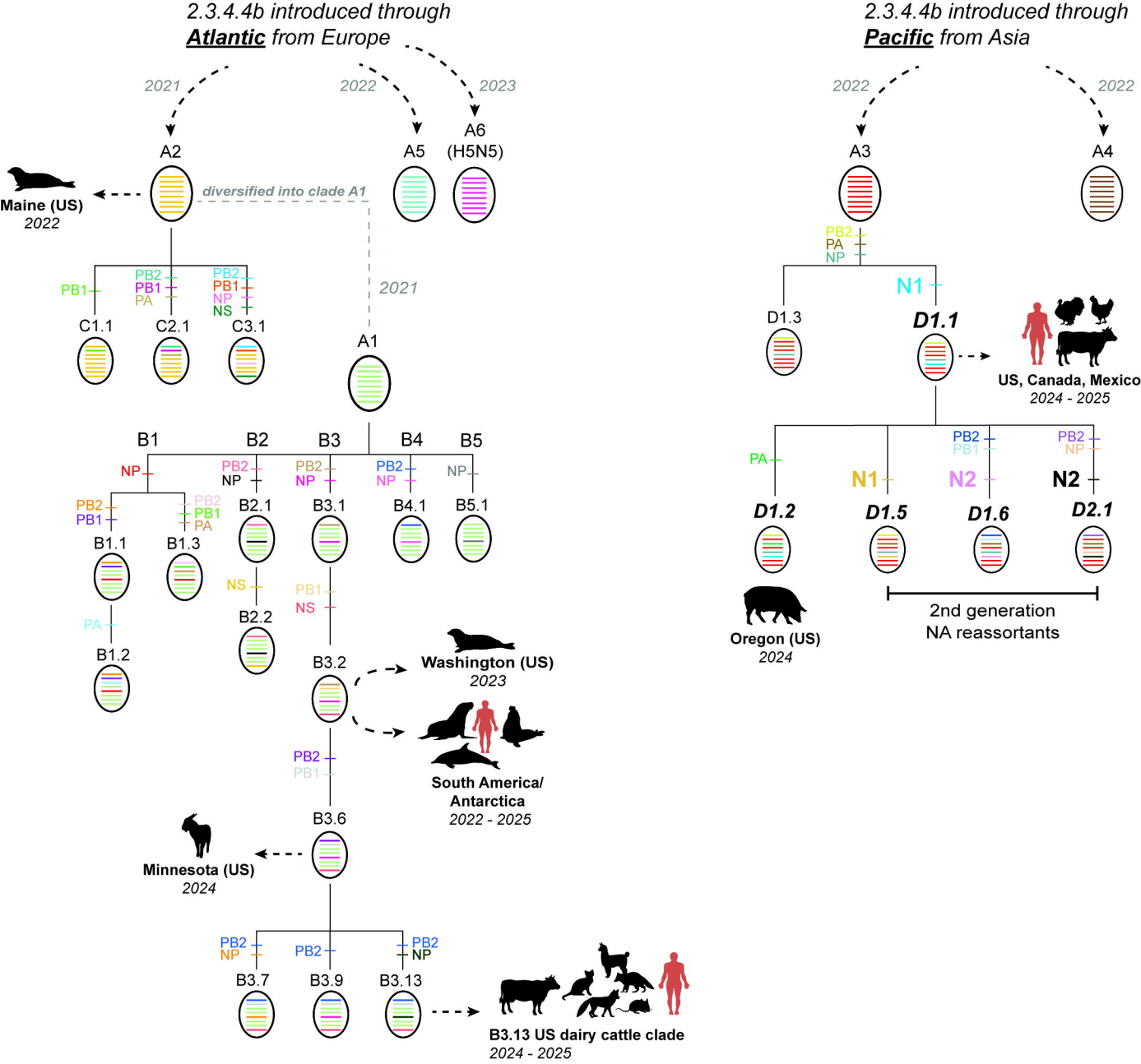
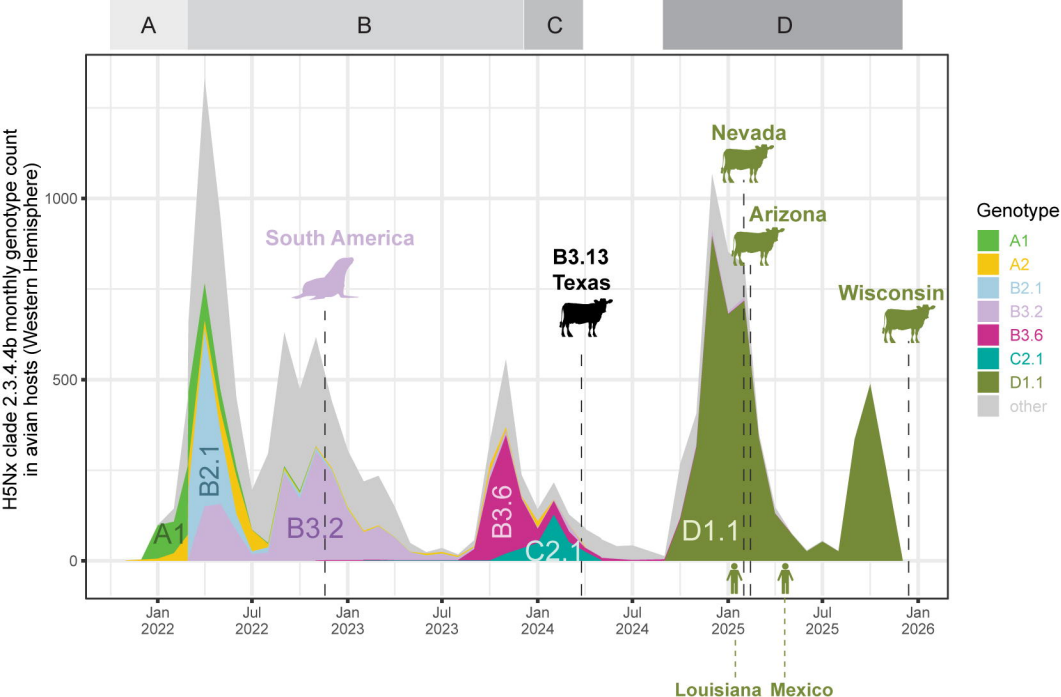
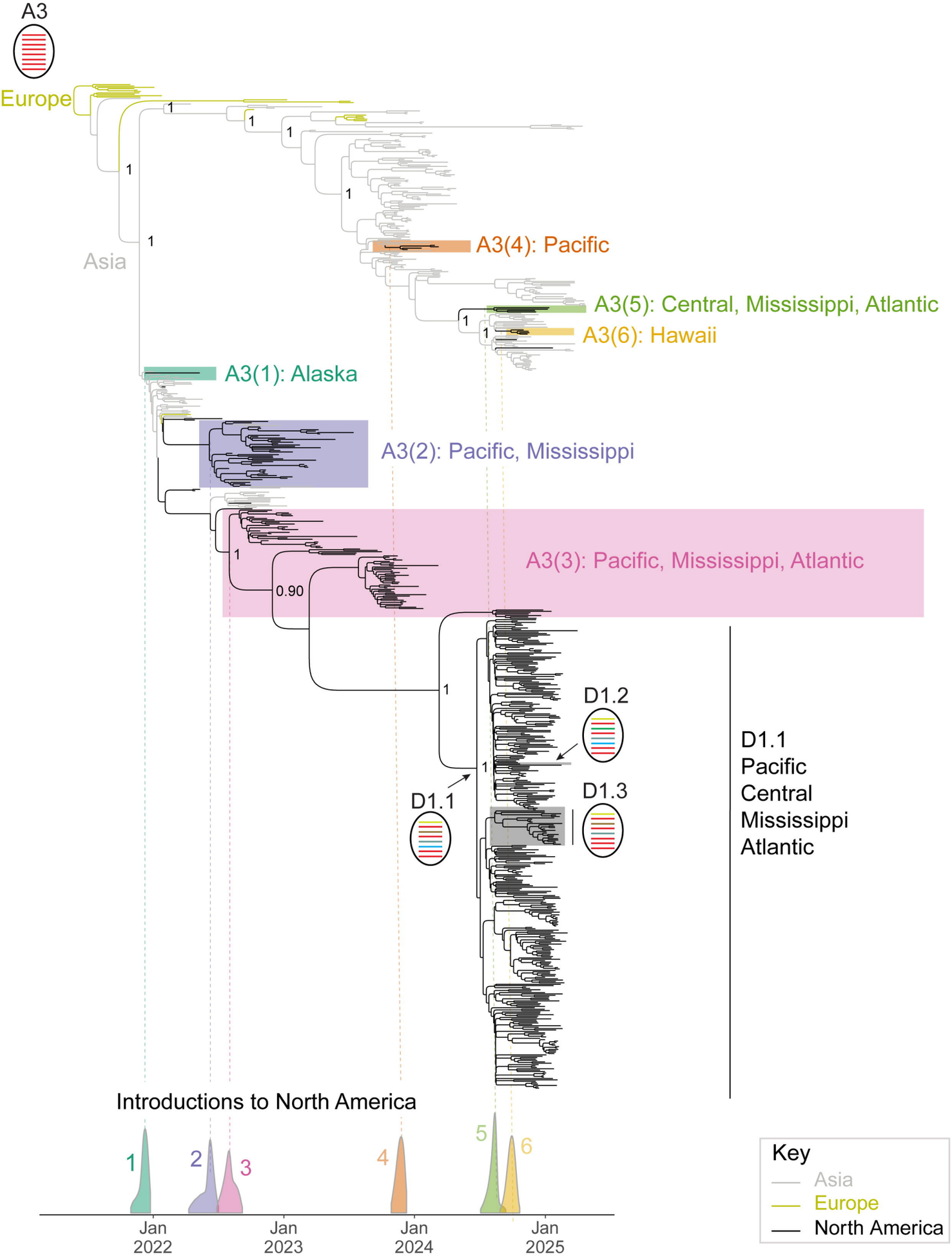


Figure S11. Low frequency of PB2 mammalian-adaptive substitutions E627K and D701N among B3.13 samples. Presence of PB2 mammalian-adaptive substitutions E627K and D701N in sequences obtained from different host groups. Full PB2 FluMut outputs available at: https://github.com/nidiatrovao/H5N1Americas.


2.3.4.4b introduced through
Atlantic from Europe
2021
2022
2023
A2
A5
A6
(H5N5)
Maine (US)
2022
diversified into clade A1
2021
PB1
PB2
PB1
PA
PB2
PB1
NP
NS
C1.1
C2.1
C3.1
A1
B1
B2
B3
B4
B5
NP
PB2
NP
PB2
NP
PB2
NP
NP
PB2
PB1
PB2
PB1
PA
B1.1
B1.3
B2.1
B3.1
B4.1
B5.1
NS
PB1
NS
PA
B2.2
B1.2
B3.2
Washington (US)
2023
South America/
Antarctica
2022 - 2025
PB2
PB1
B3.6
Minnesota (US)
2024
PB2
NP
PB2
PB2
NP
B3.7
B3.9
B3.13
B3.13 US dairy cattle clade
2024 - 2025
2.3.4.4b introduced through
Pacific from Asia
2022
2022
A3
A4
PB2
PA
NP
N1
D1.3
D1.1
US, Canada, Mexico
2024 - 2025
PA
N1
PB2
PB1
N2
PB2
NP
N2
D1.2
D1.5
D1.6
D2.1
Oregon (US)
2024
2nd generation
NA reassortants


A
B
C
D
H5Nx clade 2.3.4.4b monthly genotype count in avian hosts (Western Hemisphere)
1000
500
0
South America
B3.13
Texas
Nevada
Arizona
Wisconsin
A1
B2.1
B3.2
B3.6
C2.1
D1.1
Jan 2022
Jul 2022
Jan 2023
Jul 2023
Jan 2024
Jul 2024
Jan 2025
Jul 2025
Jan 2026
Louisiana
Mexico
Genotype
A1
A2
B2.1
B3.2
B3.6
C2.1
D1.1
other


A3
Europe
Asia
A3(4): Pacific
A3(5): Central, Mississippi, Atlantic
A3(6): Hawaii
A3(1): Alaska
A3(2): Pacific, Mississippi
A3(3): Pacific, Mississippi, Atlantic
0.90
D1.2
D1.1
D1.3
D1.1
Pacific
Central
Mississippi
Atlantic
Introductions to North America
1
2
3
4
5
6
Jan 2022
Jan 2023
Jan 2024
Jan 2025
Key
Asia
Europe
North America

A

1918 H1N1 pandemic

1918-1957 human H1N1

1977-2009 human H1N1

'Delta' Swine USA H1N1

Classical Swine H1N1

South American Avian LPAI

North American Avian LPAI

USA H5N1 2.3.4.4b (D1.1)

0.03

Eurasian H1N1 Swine (avian-like)

2009 H1N1 pandemic

2009 H1N1 pandemic (swine)
reverse zoonosis

Asia H6N1

Asia H6N1

H5N1 Gs/Gd

Italy H7N1

Eurasia H5N1 2.3.4.4b

Americas H5N1 B genotypes

Eurasian Avian LPAI

USA H5N1 2.3.4.4b (B3.13)

Between-Group Pairwise Similarity

D1.1 vs H1N1pdm09 — 84.1%

D1.1 vs H5 2.3.4.4b — 85.6%

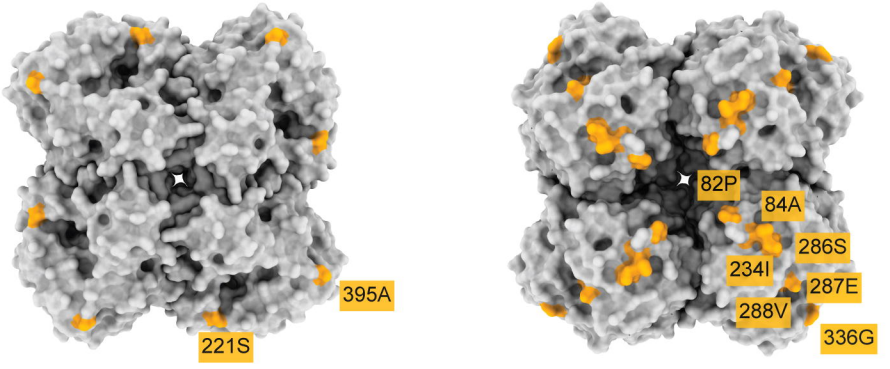
H1N1pdm09 vs H5 2.3.4.4b — 84.3%

80.0% 85.0% 90.0% 95.0%

Pairwise nucleotide similarity (%)

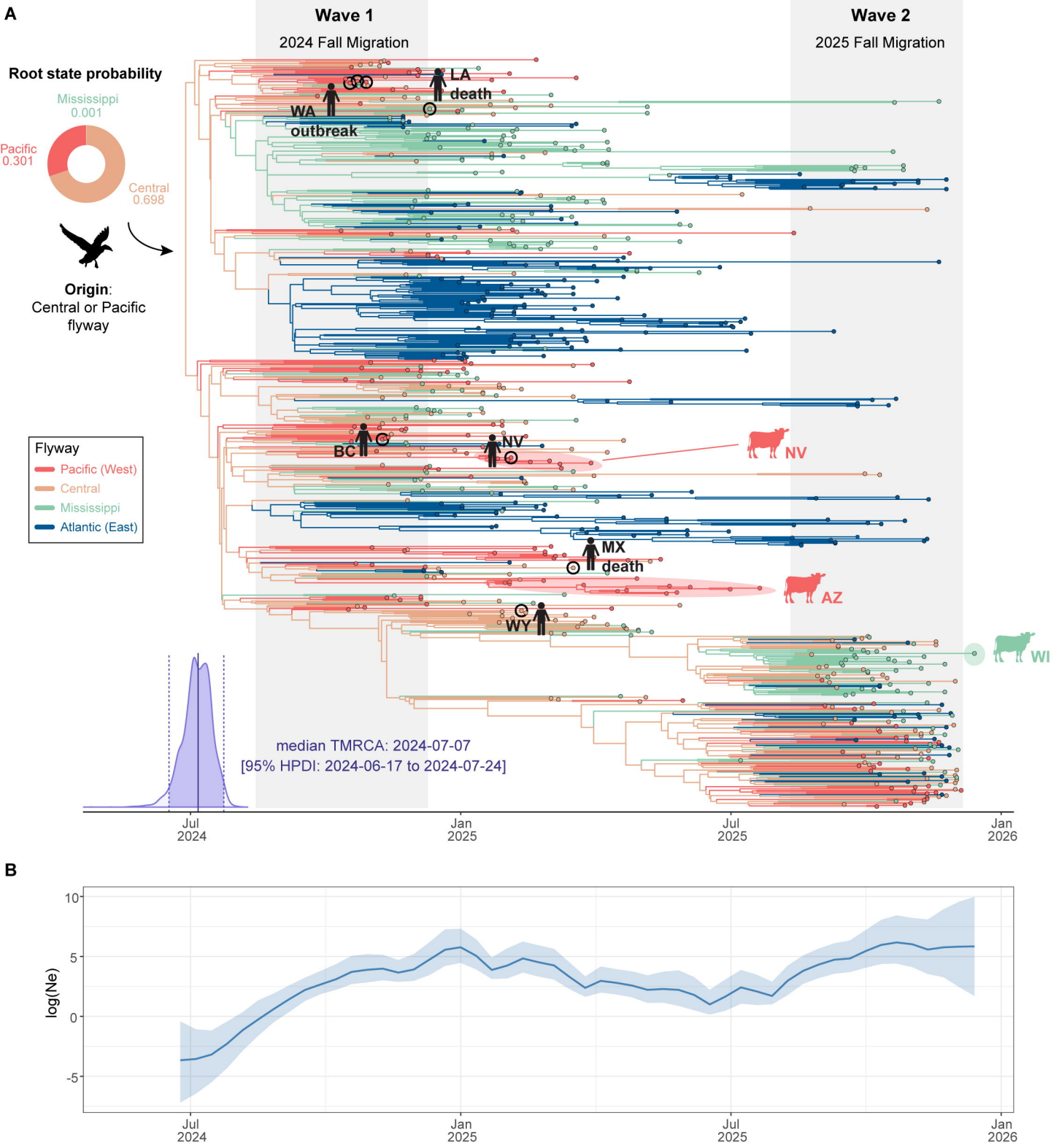
B

Neuraminidase head domain

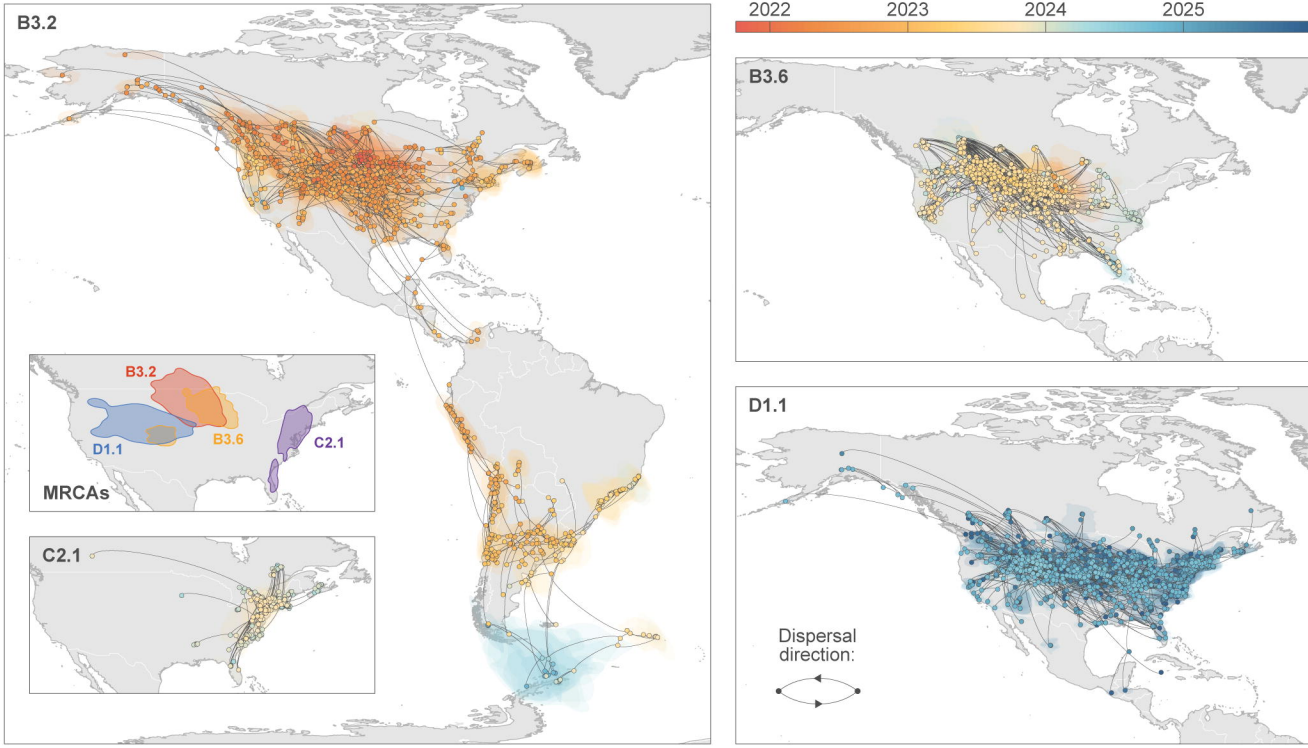


Top view

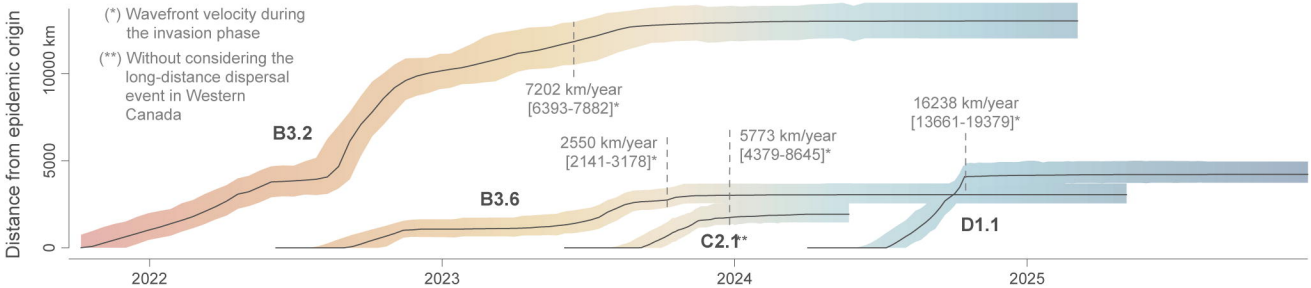
Bottom view

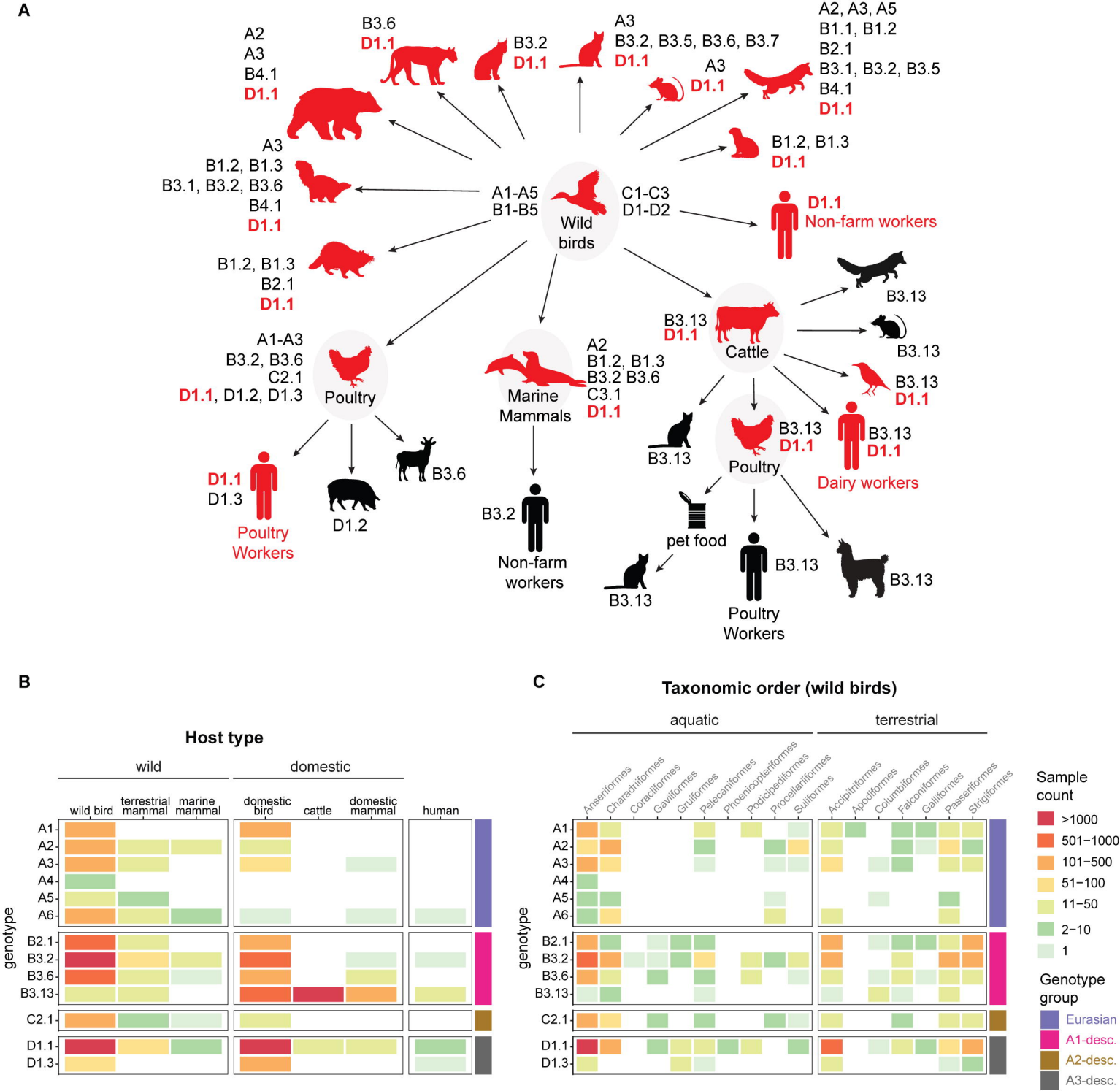

A
Wave 1
2024 Fall Migration
Wave 2
2025 Fall Migration
Root state probability
Mississippi
0.001
Pacific
0.301
Central
0.698
Origin:
Central or Pacific
flyway
Flyway
Pacific (West)
Central
Mississippi
Atlantic (East)
LA
death
WA
outbreak
BC
NV
NV
MX
death
AZ
WY
WI
median TMRCA: 2024-07-07
[95% HPDI: 2024-06-17 to 2024-07-24]
Jul
2024
Jan
2025
Jul
2025
Jan
2026
B
log(Ne)
10
5
0
-5
Jul
2024
Jan
2025
Jul
2025
Jan
2026

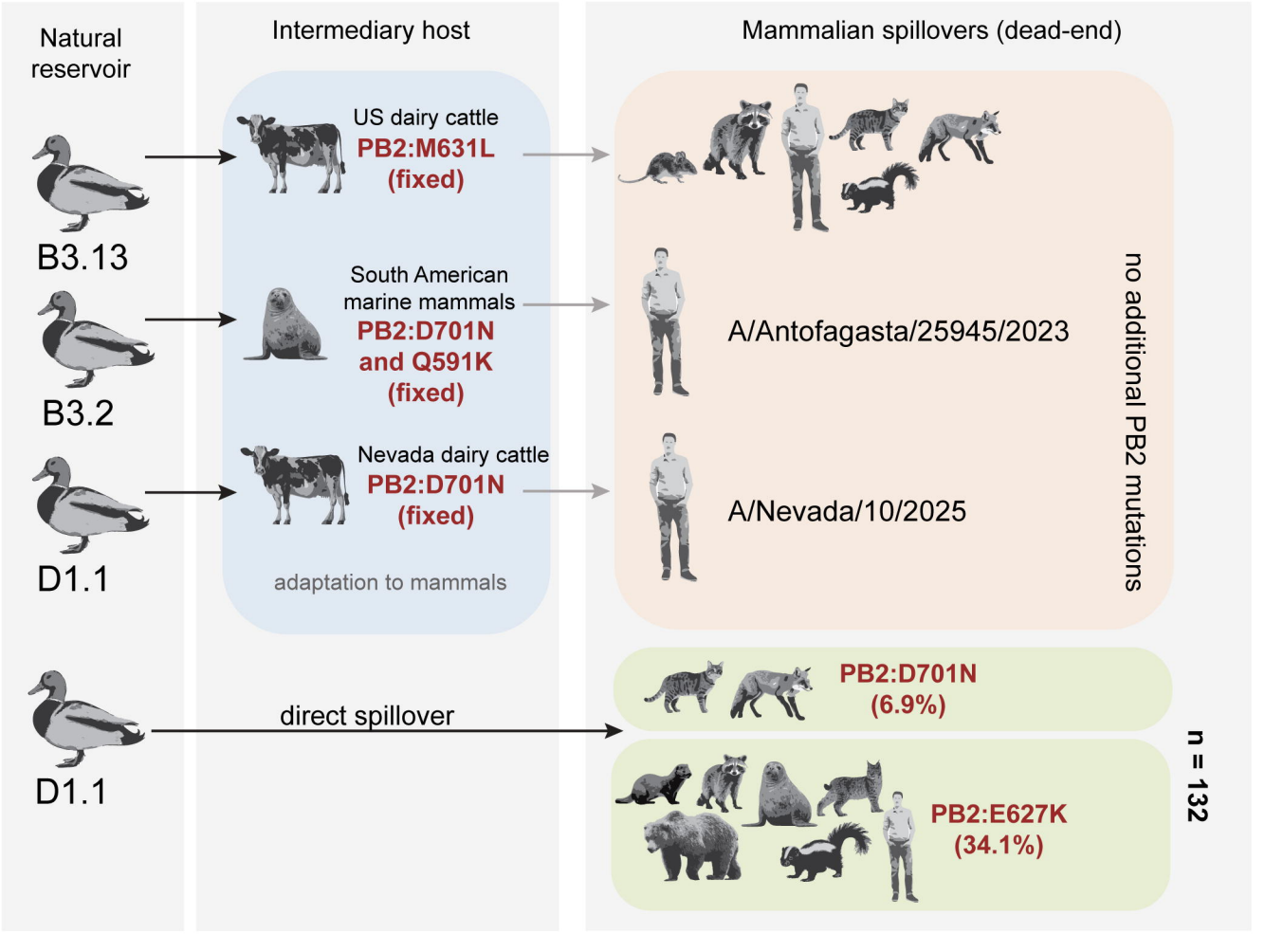
A


2022
2023
2024
2025
B3.2
B3.6
D1.1
C2.1
MRCAs
Dispersal direction:

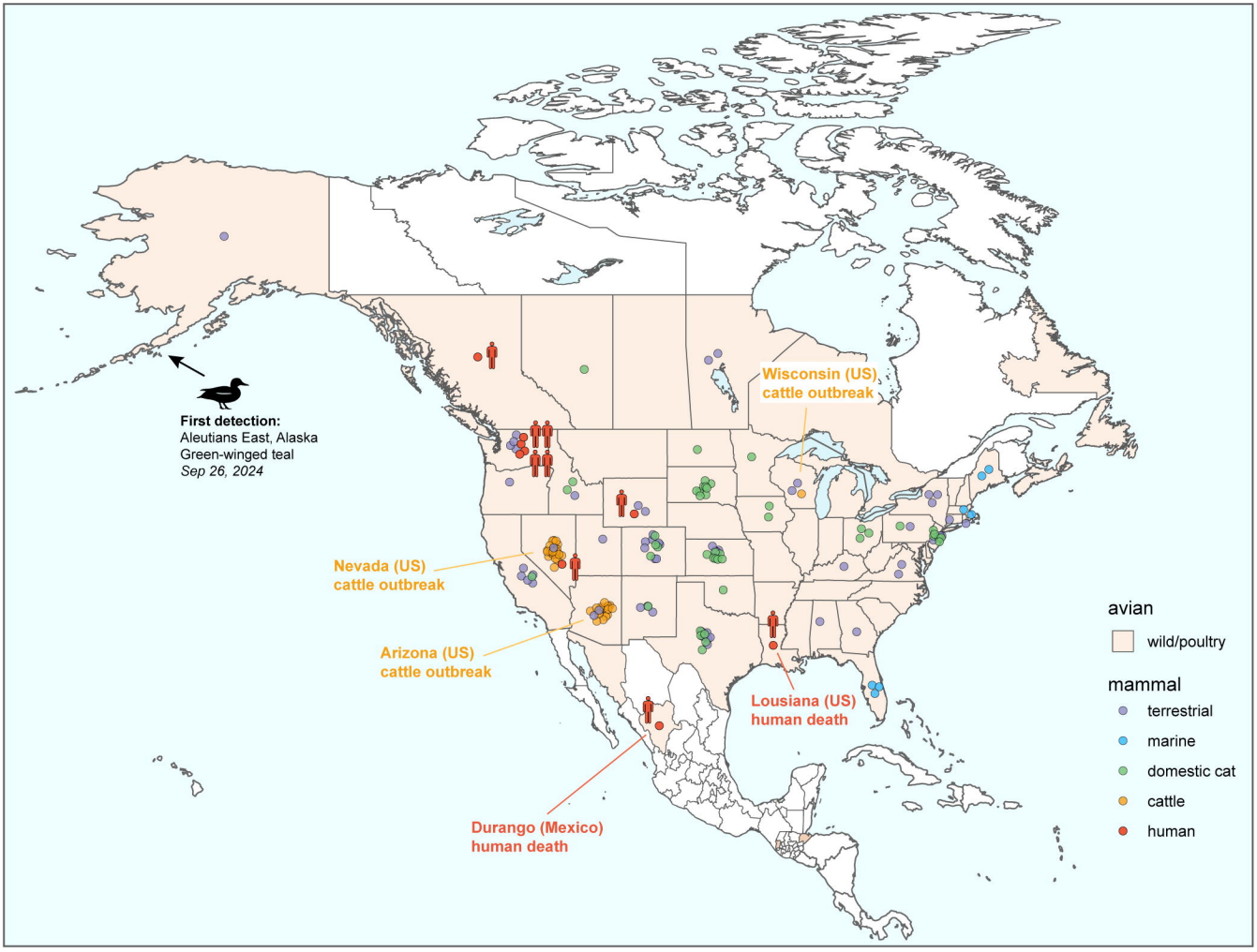
B


(*) Wavefront velocity during the invasion phase
(**) Without considering the long-distance dispersal event in Western Canada
Distance from epidemic origin
0
5000
10000 km
B3.2
7202 km/year [6393-7882]*
B3.6
2550 km/year [2141-3178]*
5773 km/year [4379-8645]*
C2.1**
D1.1
16238 km/year [13661-19379]*
2022
2023
2024
2025


A
A2
A3
B4.1
D1.1
B3.6
D1.1
B3.2
D1.1
A3
B3.2, B3.5, B3.6, B3.7
D1.1
A3
D1.1
A2, A3, A5
B1.1, B1.2
B2.1
B3.1, B3.2, B3.5
B4.1
D1.1
B1.2, B1.3
D1.1
A3
B1.2, B1.3
B3.1, B3.2, B3.6
B4.1
D1.1
A1-A5
B1-B5
Wild
birds
C1-C3
D1-D2
D1.1
Non-farm workers
B1.2, B1.3
B2.1
D1.1
B3.13
B3.13
D1.1
Cattle
B3.13
A1-A3
B3.2, B3.6
C2.1
D1.1, D1.2, D1.3
Poultry
Marine
Mammals
A2
B1.2, B1.3
B3.2 B3.6
C3.1
D1.1
B3.13
D1.1
B3.13
D1.1
B3.13
Poultry
B3.13
D1.1
Dairy workers
D1.1
D1.3
Poultry
Workers
D1.2
B3.6
B3.2
Non-farm
workers
pet food
B3.13
B3.13
Poultry
Workers
B3.13
B
Host type
wild
domestic
wild bird
terrestrial mammal
marine mammal
domestic bird
cattle
domestic mammal
human
genotype
A1
A2
A3
A4
A5
A6
B2.1
B3.2
B3.6
B3.13
C2.1
D1.1
D1.3
C
Taxonomic order (wild birds)
aquatic
terrestrial
Anseriformes
Charadriiformes
Coraciiformes
Gaviiformes
Gruiformes
Pelecaniformes
Phoenicopteriformes
Podicipediformes
Procellariiformes
Suliformes
Accipitriformes
Apodiformes
Columbiformes
Falconiformes
Galliformes
Passeriformes
Strigiformes
Sample
count
>1000
501−1000
101−500
51−100
11−50
2−10
1
Genotype
group
Eurasian
A1-desc.
A2-desc.
A3-desc.

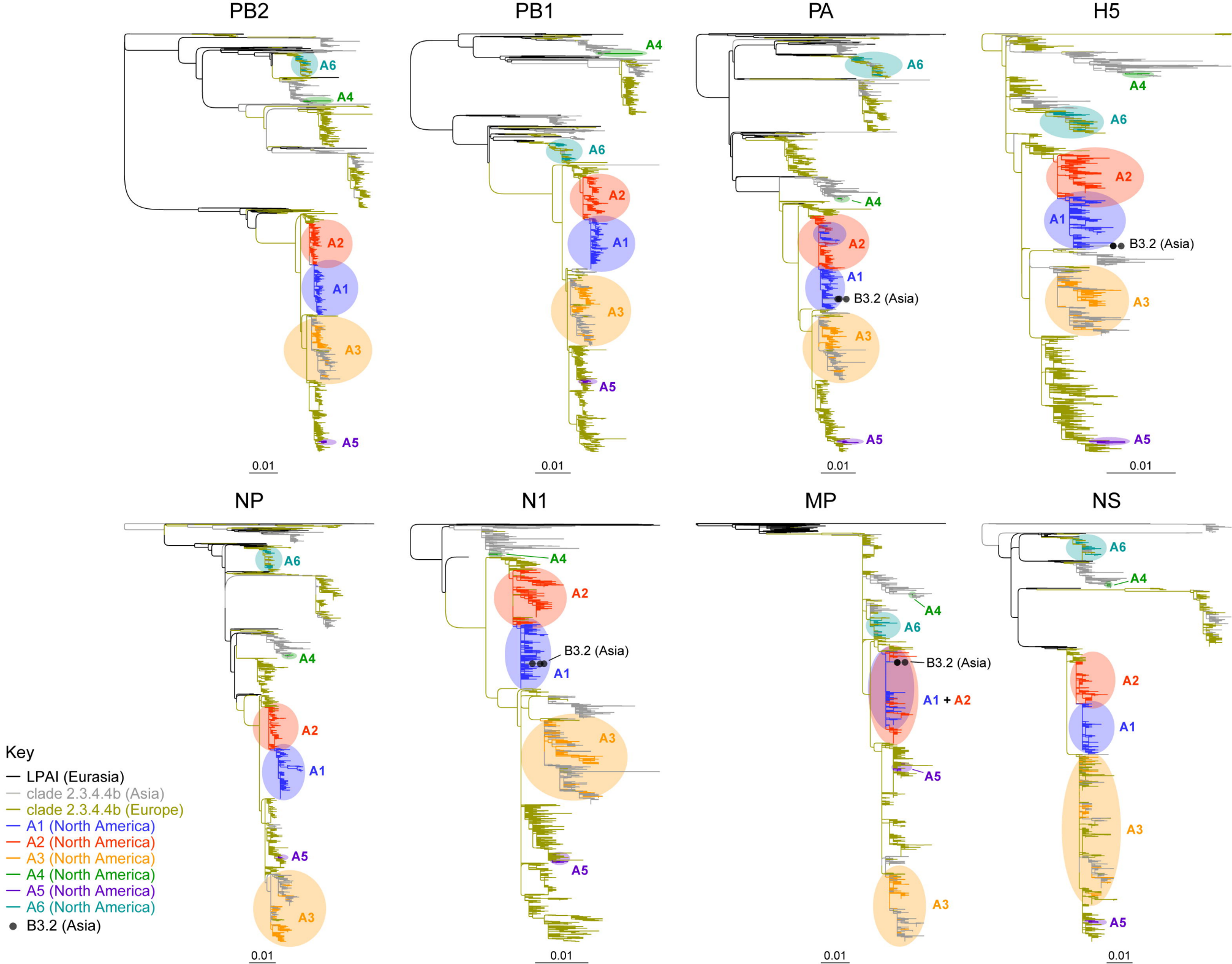

Natural reservoir
Intermediary host
Mammalian spillovers (dead-end)
B3.13
B3.2
D1.1
D1.1
US dairy cattle
PB2:M631L
(fixed)
South American marine mammals
PB2:D701N and Q591K
(fixed)
Nevada dairy cattle
PB2:D701N
(fixed)
adaptation to mammals
A/Antofagasta/25945/2023
A/Nevada/10/2025
no additional PB2 mutations
direct spillover
PB2:D701N
(6.9%)
PB2:E627K
(34.1%)
n = 132

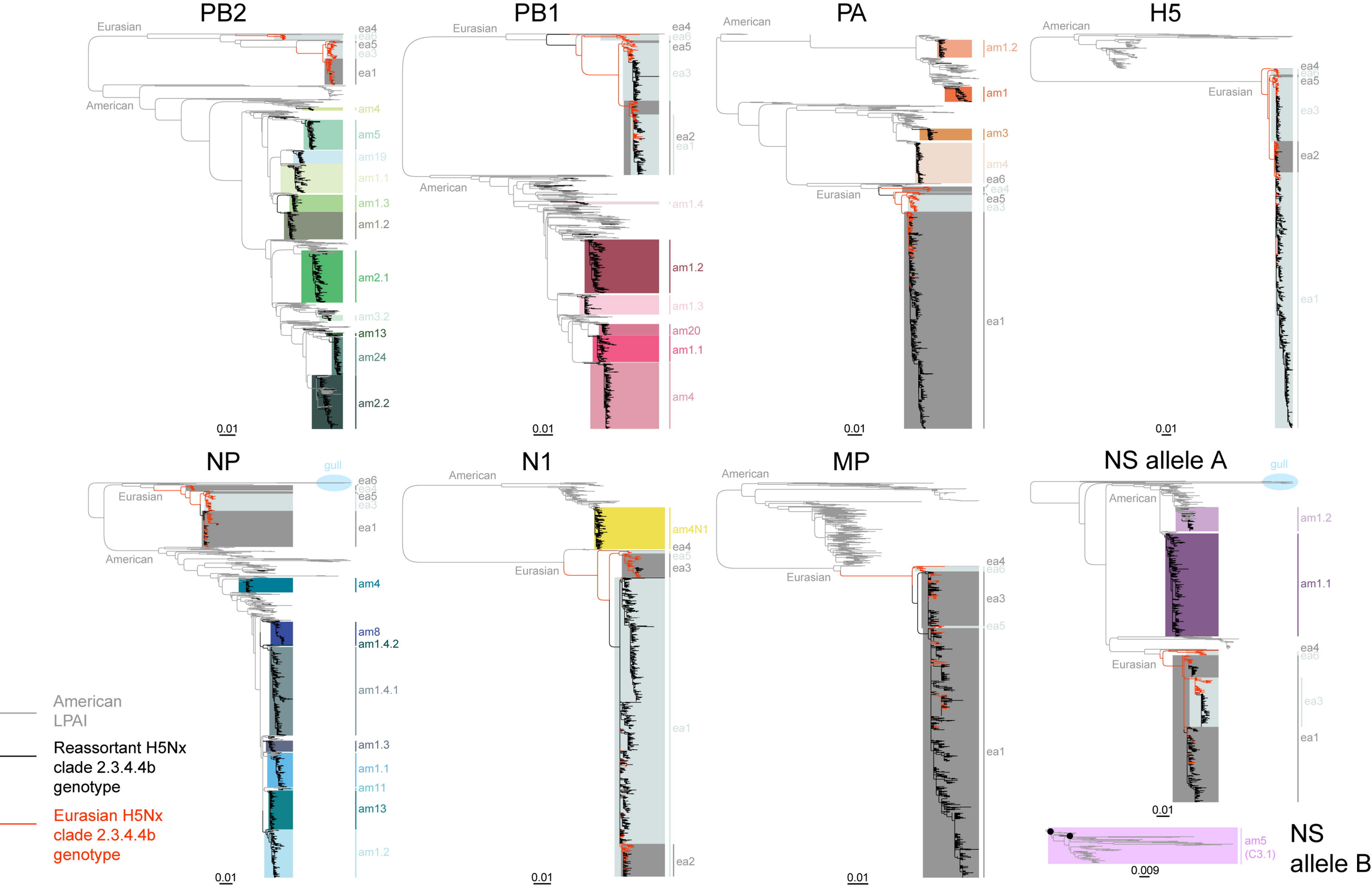
A

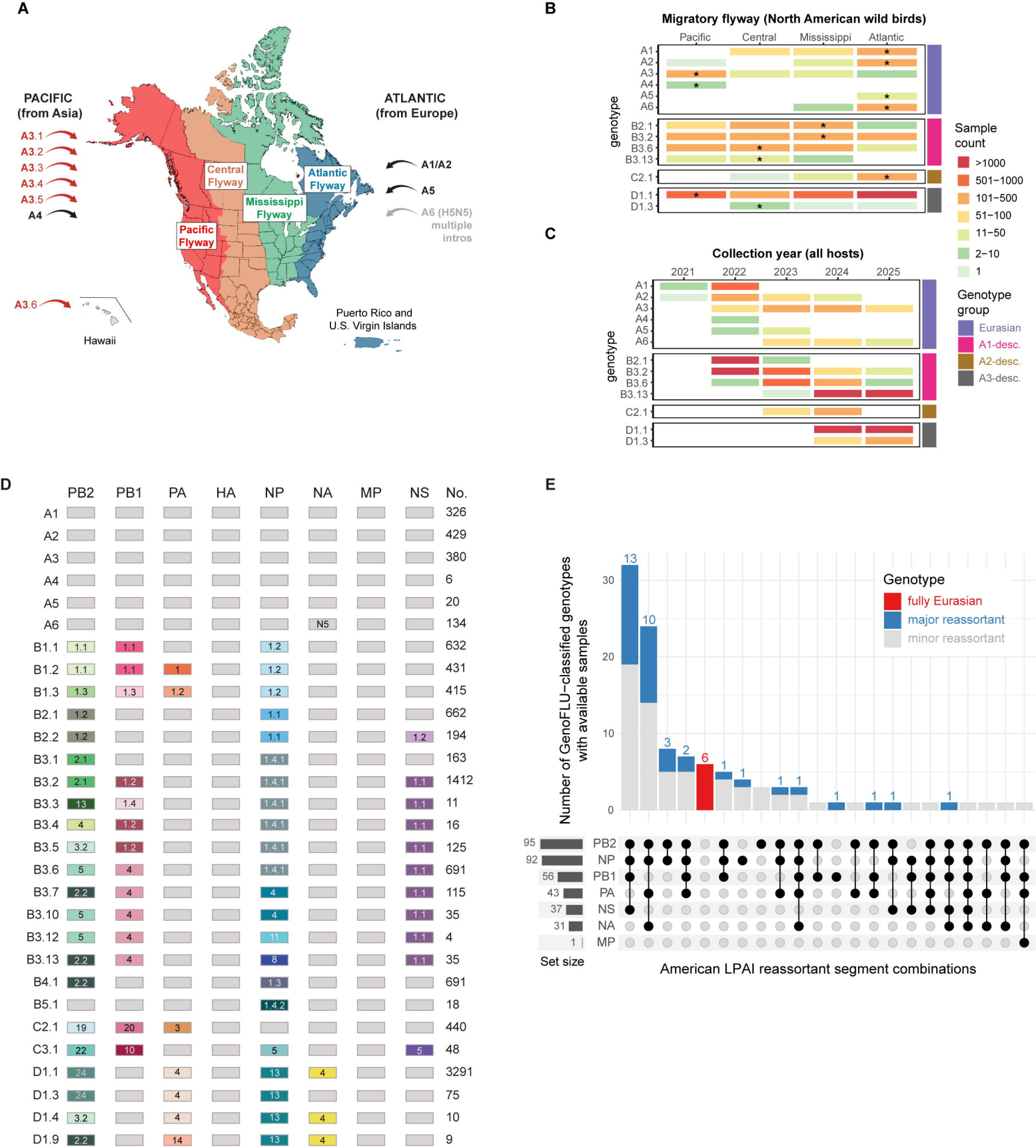


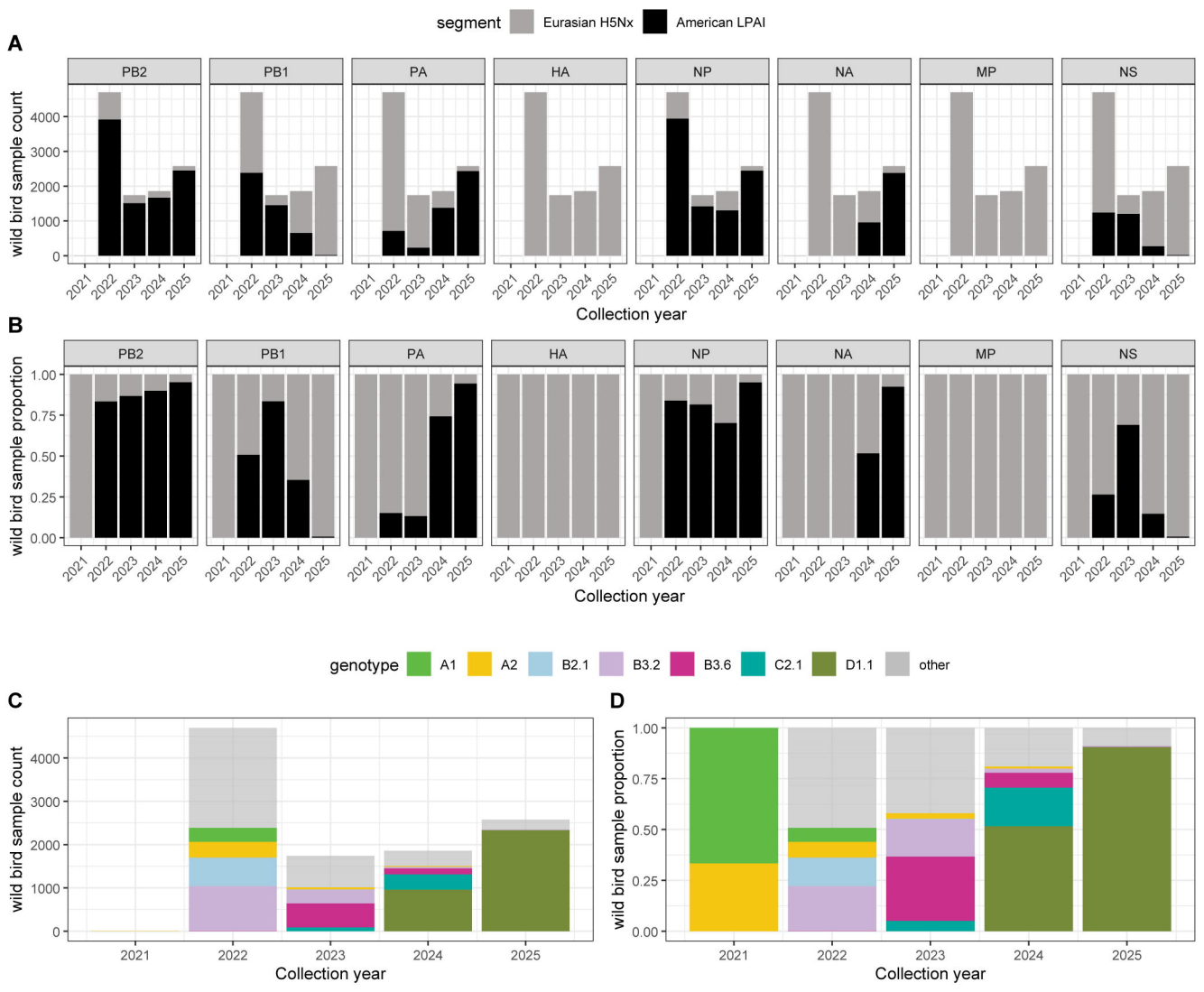
B

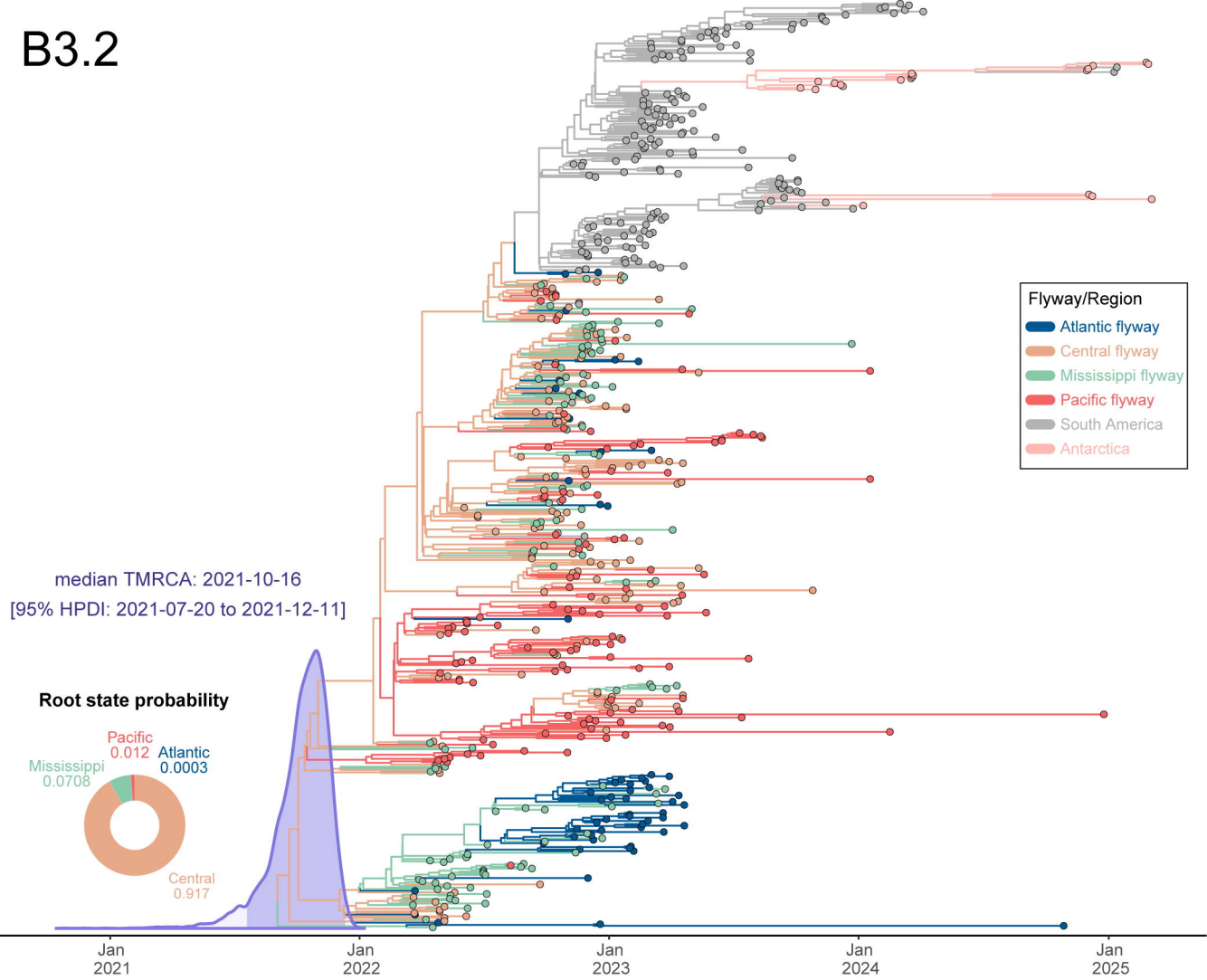
Virus name	Collection date	Location	Source	Disease severity	Age	Acquired mammalian-adaptive PB2 mutations	Accession (GISAID)
A/Washington/240/2024	10/18/2024	USA-WA	commercial poultry	mild (unspecified)	adult (>18)	--	EPI_ISL_19531299
A/Washington/UW32495/2024	10/18/2024	USA-WA	commercial poultry	mild (unspecified)	adult (>18)	--	EPI_ISL_19541397
A/Washington/254/2024	10/23/2024	USA-WA	commercial poultry	mild (unspecified)	adult (>18)	--	EPI_ISL_19666173
A/Washington/255/2024	10/29/2024	USA-WA	commercial poultry	mild (unspecified)	adult (>18)	--	EPI_ISL_19552697
A/British Columbia/PHL-2032/2024	11/9/2024	Canada-BC	unknown	hospitalization (with conjunctivitis)	teen (13-18)	E627K	EPI_ISL_19548836
A/Louisiana/12/2024	12/11/2024	USA-LA	backyard poultry	death	adult (>18)	--	EPI_ISL_19634828
A/Nevada/10/2025	2/4/2025	USA-NV	dairy cattle	mid (conjunctivitis)	adult (>18)	D701N	EPI_ISL_19726293
A/Wyoming/01/2025	2/11/2025	USA-WY	backyard poultry	hospitalization	adult (>18)	E627K	EPI_ISL_19749443
A/Mexico/Dur_Indre_2292/2025	3/15/2025	Mexico-Durango	unknown	death	child (<13)	--	EPI_ISL_20227034

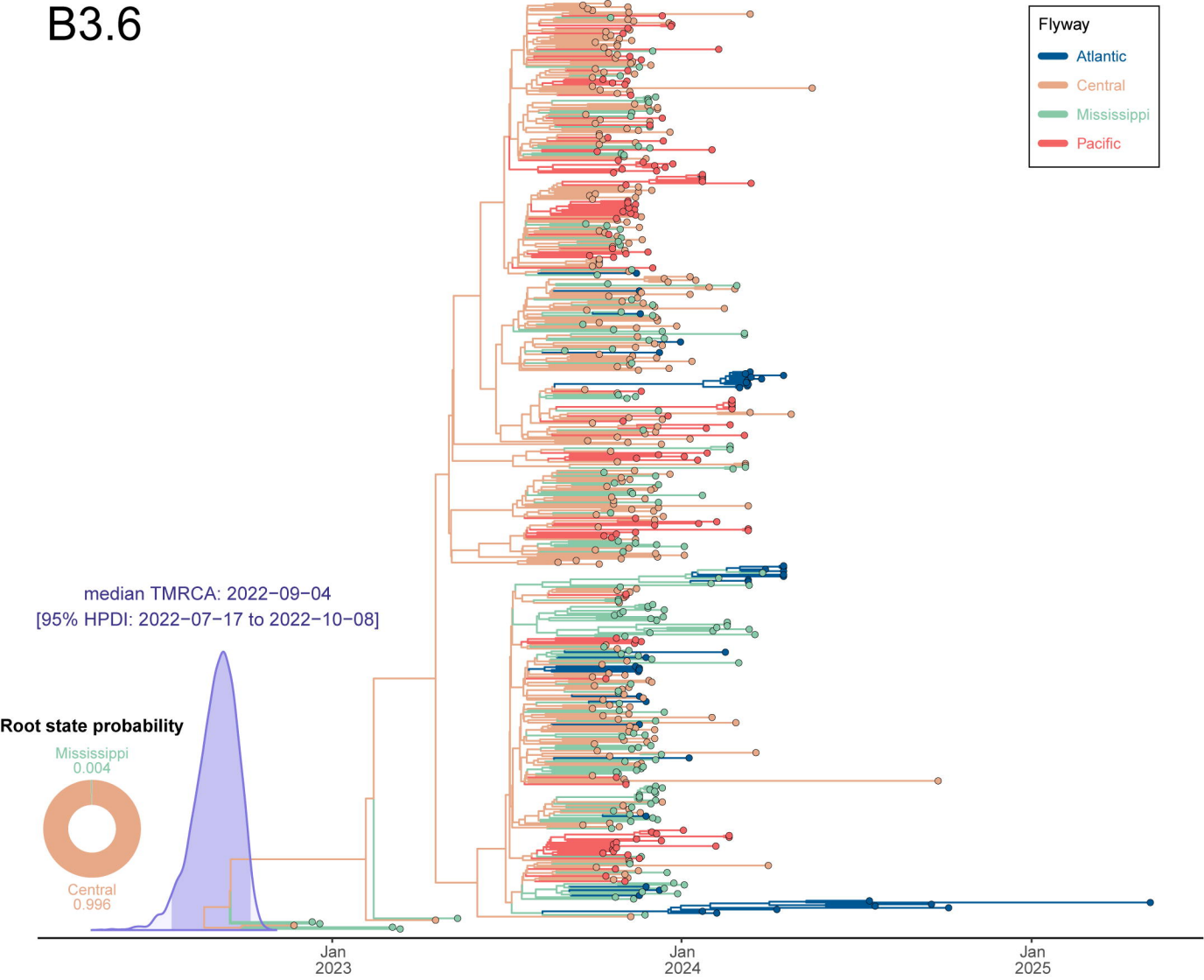

PB2
PB1
PA
H5
NP
N1
MP
NS
A1
A2
A3
A4
A5
A6
B3.2 (Asia)
A1 + A2
0.01
Key
LPAI (Eurasia)
clade 2.3.4.4b (Asia)
clade 2.3.4.4b (Europe)
A1 (North America)
A2 (North America)
A3 (North America)
A4 (North America)
A5 (North America)
A6 (North America)
B3.2 (Asia)

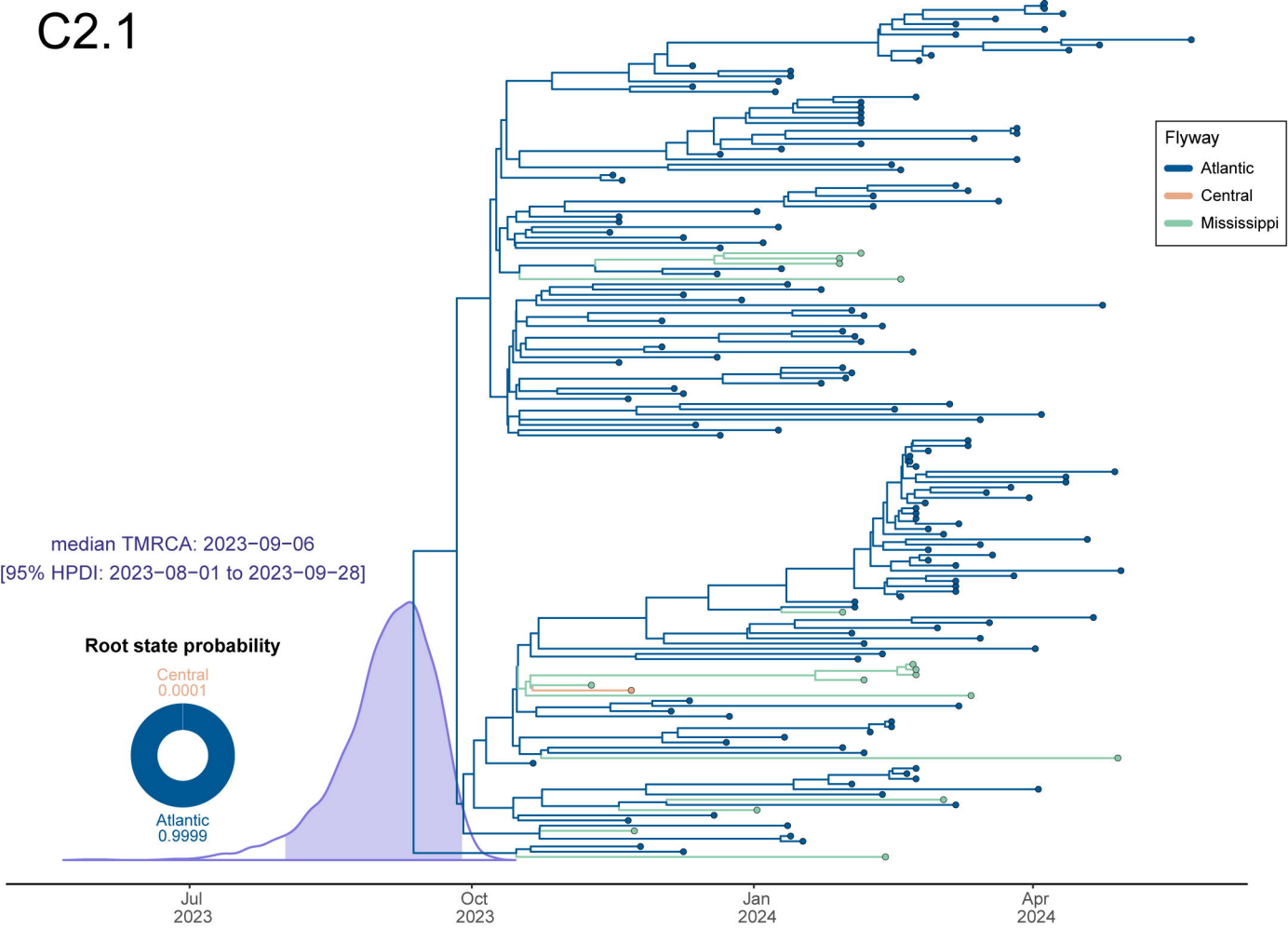

PB2
Eurasian
American
ea4
ea6
ea5
ea3
ea1
am4
am5
am19
am1.1
am1.3
am1.2
am2.1
am3.2
am13
am24
am2.2
0.01
PB1
Eurasian
American
ea4
ea6
ea5
ea3
ea2
ea1
am1.4
am1.2
am1.3
am20
am1.1
am4
0.01
PA
American
Eurasian
am1.2
am1
am3
am4
ea6
ea4
ea5
ea3
ea1
0.01
H5
American
Eurasian
ea4
ea6
ea5
ea3
ea2
ea1
0.01
NP
gull
Eurasian
American
ea6
ea4
ea5
ea3
ea1
am4
am8
am1.4.2
am1.4.1
am1.3
am1.1
am11
am13
am1.2
0.01
N1
American
Eurasian
am4N1
ea4
ea5
ea3
ea1
ea2
0.01
MP
American
Eurasian
ea4
ea6
ea3
ea5
ea1
0.01
NS allele A
gull
American
Eurasian
am1.2
am1.1
ea4
ea6
ea3
ea1
0.01
am5
(C3.1)
0.009
NS
allele B
American
LPAI
Reassortant H5Nx
clade 2.3.4.4b
genotype
Eurasian H5Nx
clade 2.3.4.4b
genotype

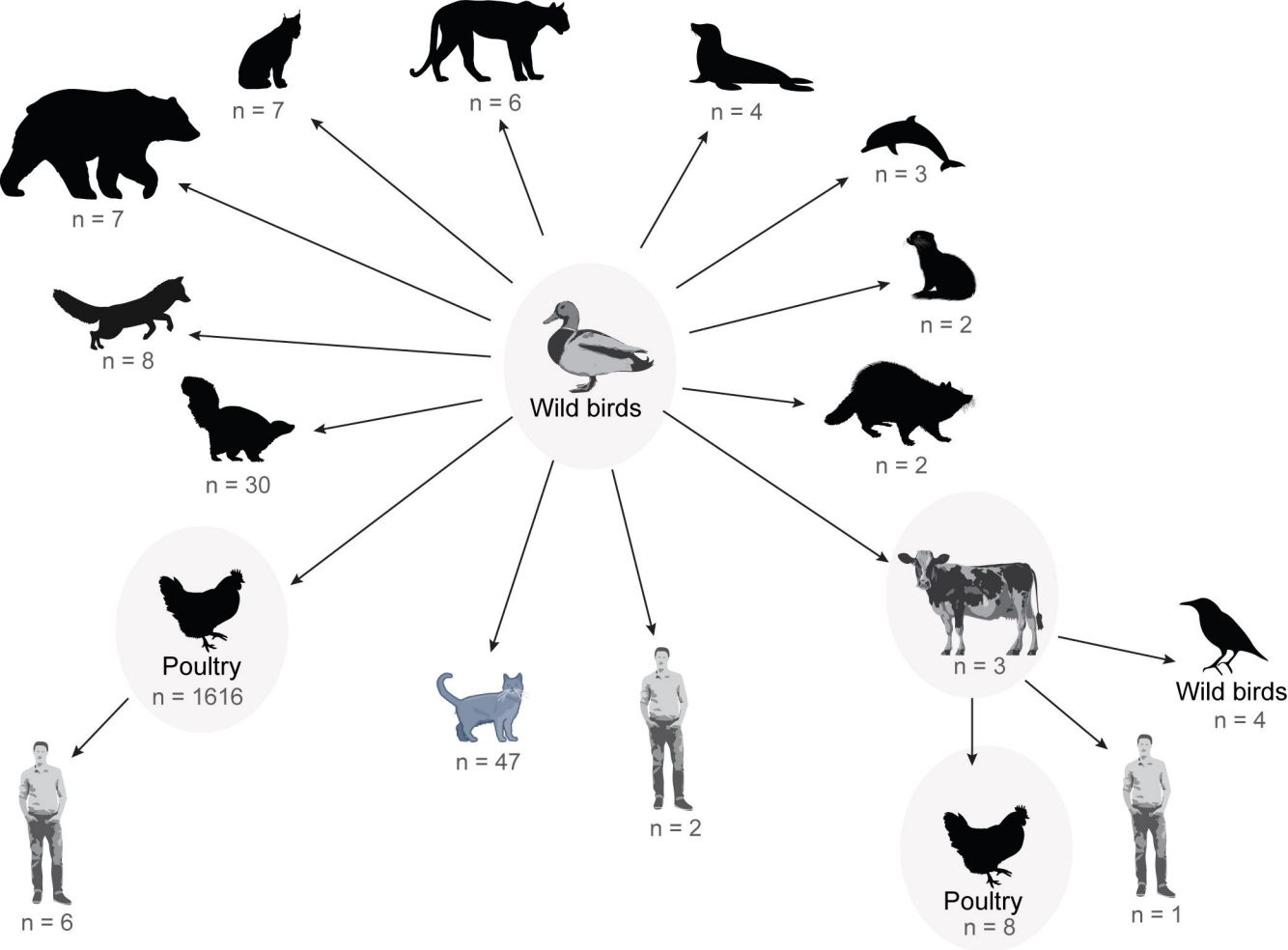

A
PACIFIC
(from Asia)
A3.1
A3.2
A3.3
A3.4
A3.5
A4
A3.6
Hawaii
Central Flyway
Atlantic Flyway
Mississippi Flyway
Pacific Flyway
ATLANTIC
(from Europe)
A1/A2
A5
A6 (H5N5)
multiple
intros
Puerto Rico and
U.S. Virgin Islands
B
Migratory flyway (North American wild birds)
Pacific
Central
Mississippi
Atlantic
genotype
A1
A2
A3
A4
A5
A6
B2.1
B3.2
B3.6
B3.13
C2.1
D1.1
D1.3
Sample
count
>1000
501−1000
101−500
51−100
11−50
2−10
1
C
Collection year (all hosts)
2021
2022
2023
2024
2025
genotype
Genotype
group
Eurasian
A1-desc.
A2-desc.
A3-desc.
D
PB2
PB1
PA
HA
NP
NA
MP
NS
No.
A1 326
A2 429
A3 380
A4 6
A5 20
A6 N5 134
B1.1 1.1 1.1 1.2 632
B1.2 1.1 1.1 1 1.2 431
B1.3 1.3 1.3 1.2 1.2 415
B2.1 1.2 1.1 662
B2.2 1.2 1.1 1.2 194
B3.1 2.1 1.4.1 163
B3.2 2.1 1.2 1.4.1 1.1 1412
B3.3 13 1.4 1.4.1 1.1 11
B3.4 4 1.2 1.4.1 1.1 16
B3.5 3.2 1.2 1.4.1 1.1 125
B3.6 5 4 1.4.1 1.1 691
B3.7 2.2 4 4 1.1 115
B3.10 5 4 4 1.1 35
B3.12 5 4 11 1.1 4
B3.13 2.2 4 8 1.1 35
B4.1 2.2 1.3 691
B5.1 1.4.2 18
C2.1 19 20 3 440
C3.1 22 10 5 5 48
D1.1 24 4 13 4 3291
D1.3 24 4 13 75
D1.4 3.2 4 13 4 10
D1.9 2.2 14 13 4 9
E
Number of GenoFLU−classified genotypes
with available samples
Genotype
fully Eurasian
major reassortant
minor reassortant
Set size
PB2
NP
PB1
PA
NS
NA
MP
American LPAI reassortant segment combinations


segment
Eurasian H5Nx
American LPAI
A
PB2
PB1
PA
HA
NP
NA
MP
NS
wild bird sample count
Collection year
B
wild bird sample proportion
genotype
A1
A2
B2.1
B3.2
B3.6
C2.1
D1.1
other
C
D

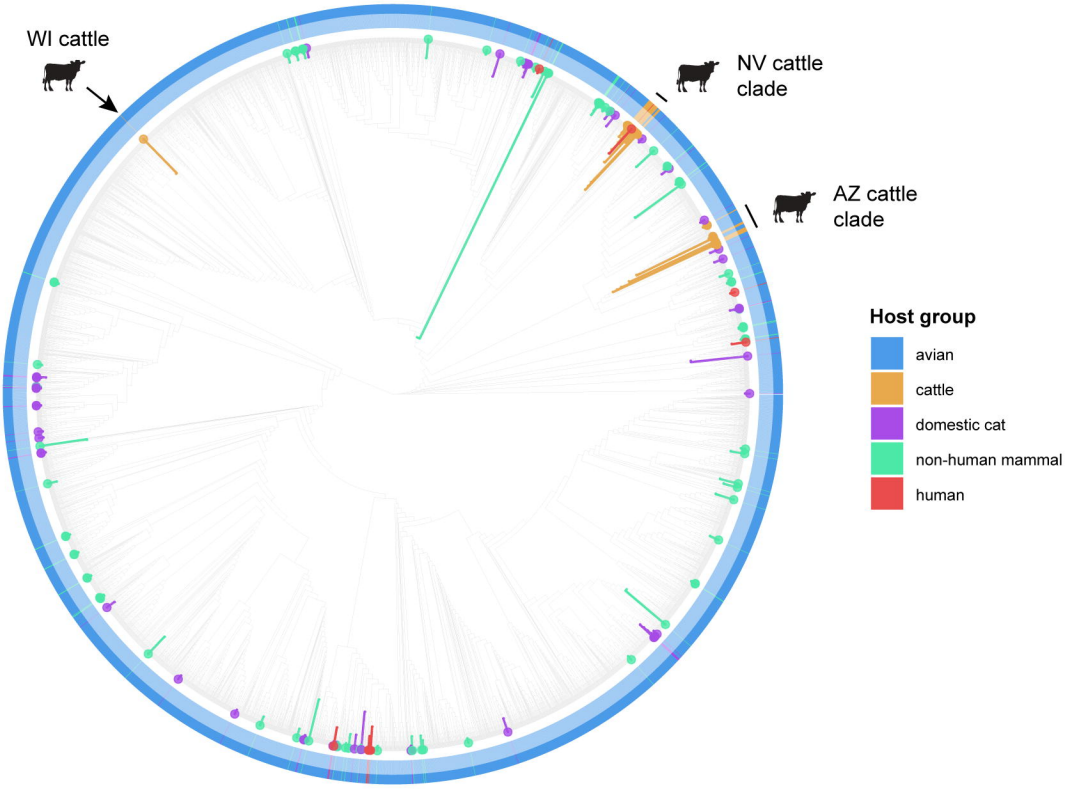

B3.2
Flyway/Region
Atlantic flyway
Central flyway
Mississippi flyway
Pacific flyway
South America
Antarctica
median TMRCA: 2021-10-16
[95% HPDI: 2021-07-20 to 2021-12-11]
Root state probability
Pacific
0.012
Atlantic
0.0003
Mississippi
0.0708
Central
0.917
Jan
2021
Jan
2022
Jan
2023
Jan
2024
Jan
2025


B3.6
Flyway
Atlantic
Central
Mississippi
Pacific
median TMRCA: 2022−09−04
[95% HPDI: 2022−07−17 to 2022−10−08]
Root state probability
Mississippi
0.004
Central
0.996
Jan
2023
Jan
2024
Jan
2025

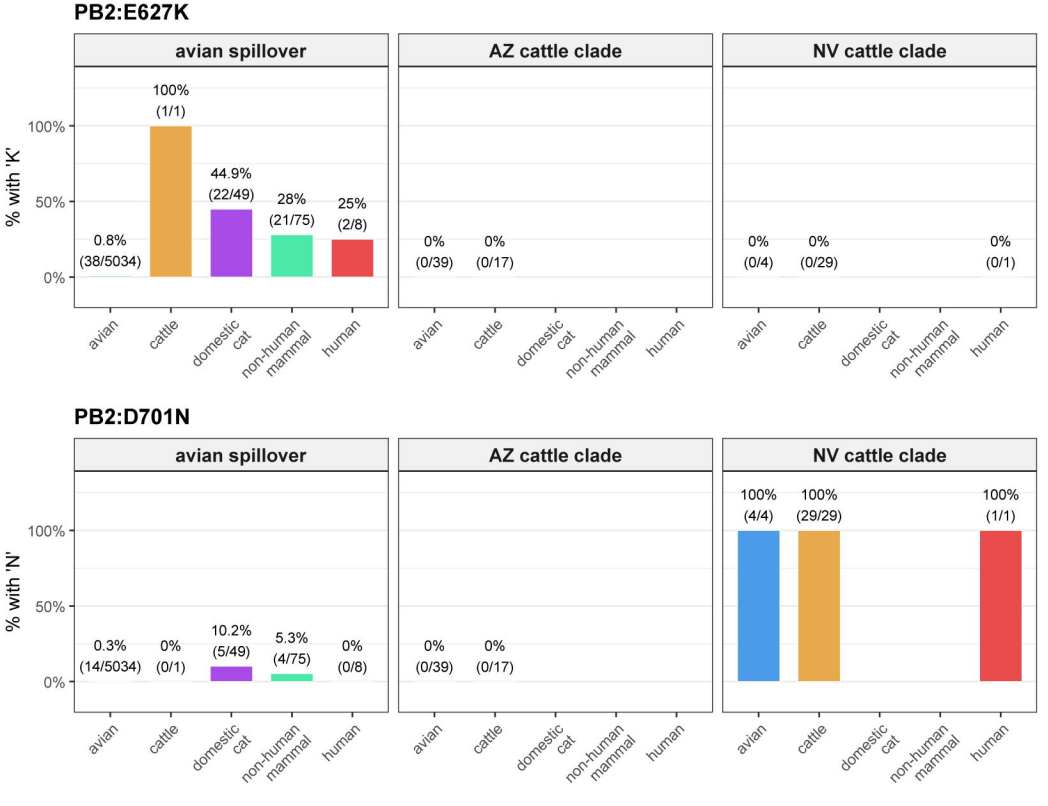

C2.1
Flyway
Atlantic
Central
Mississippi
median TMRCA: 2023−09−06
[95% HPDI: 2023−08−01 to 2023−09−28]
Root state probability
Central
0.0001
Atlantic
0.9999
Jul
2023
Oct
2023
Jan
2024
Apr
2024

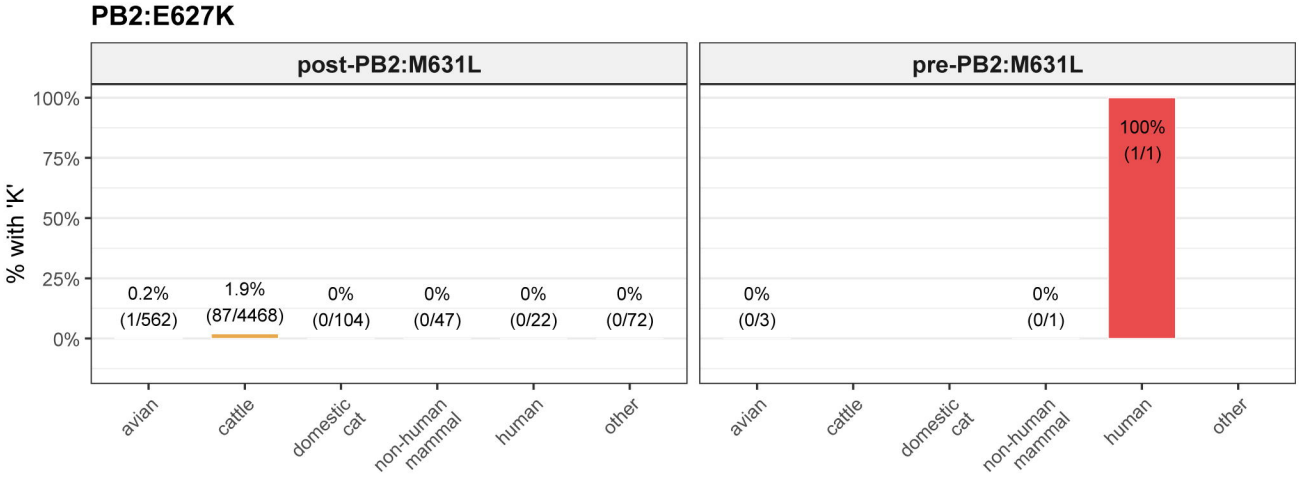

n = 7
n = 6
n = 4
n = 3
n = 7
n = 2
n = 8
Wild birds
n = 2
n = 30
Poultry
n = 1616
n = 3
Wild birds
n = 4
n = 47
n = 2
n = 6
Poultry
n = 8
n = 1

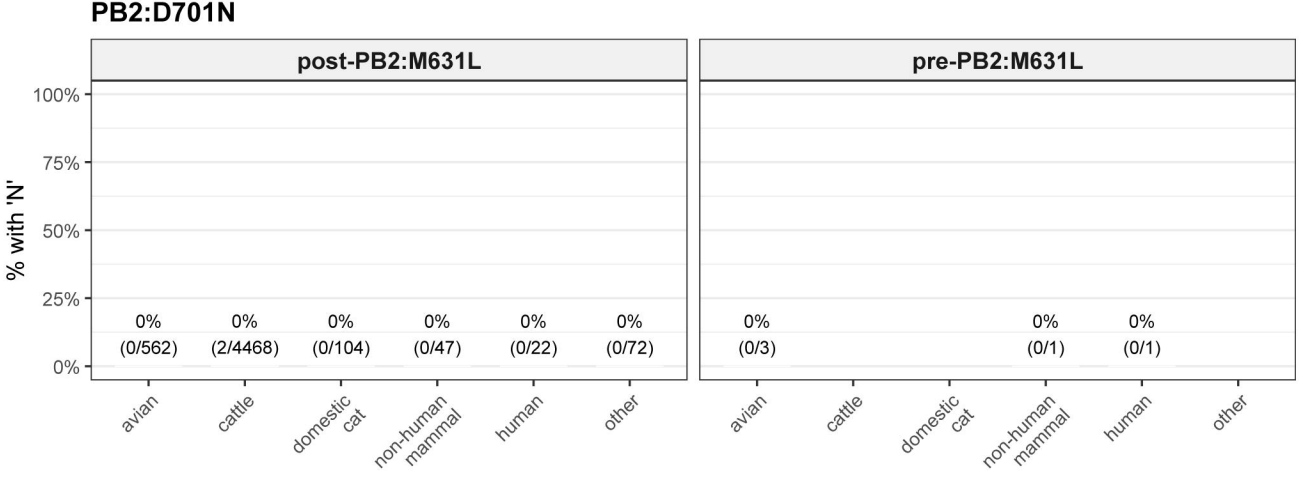
A


WI cattle
NV cattle clade
AZ cattle clade
Host group
avian
cattle
domestic cat
non-human mammal
human

B


PB2:E627K
avian spillover
AZ cattle clade
NV cattle clade
% with 'K'
100%
50%
0%
0.8% (38/5034)
100% (1/1)
44.9% (22/49)
28% (21/75)
25% (2/8)
0% (0/39)
0% (0/17)
0% (0/4)
0% (0/29)
0% (0/1)
avian
cattle
domestic cat
non-human mammal
human
PB2:D701N
% with 'N'
0.3% (14/5034)
0% (0/1)
10.2% (5/49)
5.3% (4/75)
0% (0/8)
0% (0/39)
0% (0/17)
100% (4/4)
100% (29/29)
100% (1/1)


PB2:E627K
post-PB2:M631L
pre-PB2:M631L
% with 'K'
100%
75%
50%
25%
0%
0.2%
(1/562)
1.9%
(87/4468)
0%
(0/104)
0%
(0/47)
0%
(0/22)
0%
(0/72)
0%
(0/3)
0%
(0/1)
100%
(1/1)
avian
cattle
domestic cat
non-human mammal
human
other


PB2:D701N
post-PB2:M631L
pre-PB2:M631L
% with 'N'
100%
75%
50%
25%
0%
0%
(0/562)
0%
(2/4468)
0%
(0/104)
0%
(0/47)
0%
(0/22)
0%
(0/72)
0%
(0/3)
0%
(0/1)
0%
(0/1)
avian
cattle
domestic cat
non-human mammal
human
other