RESEARCH ARTICLE SUMMARY

AVIAN INFLUENZA

Emergence and interstate spread of highly pathogenic avian influenza A(H5N1) in dairy cattle in the United States

Thao-Quyen Nguyen†, Carl R. Hutter†, Alexey Markin, Megan Thomas, Kristina Lantz, Mary Lea Killian, Garrett M. Janzen, Sriram Vijendran, Sanket Wagle, Blake Inderski, Drew R. Magstadt, Ganwu Li, Diego G. Diel, Elisha Anna Frye, Kiril M. Dimitrov, Amy K. Swinford, Alexis C. Thompson, Kevin R. Snekvik, David L. Suarez, Steven M. Lakin, Stacey Schwabenlander, Sara C. Ahola, Kammy R. Johnson, Amy L. Baker, Suelee Robbe-Austerman*, Mia Kim Torchetti*, Tavis K. Anderson*

INTRODUCTION: Highly pathogenic avian influenza (HPAI) viruses have critical consequences for animal health and the agricultural economy—and may have pandemic potential. HPAI related to the goose/Guangdong 2.3.4.4 hemagglutinin (HA) H5NX phylogenetic clade has spread to nearly 100 countries, and it is recognized as a panzootic. HPAI virus circulation is ongoing in North America, and after a trans-Atlantic incursion in late 2021, the HPAI H5N1 clade 2.3.4.4b virus caused widespread outbreaks. The outbreaks resulted in extensive mortality events, culling of poultry when detected in agricultural systems, and interspecies transmission events into mammals. It is critical to determine how HPAI clade 2.3.4.4b evolves in wild birds and in nonhuman mammals after spillover to assess the potential for human infection and transmission.

RATIONALE: In late January 2024, veterinarians observed dairy cattle displaying decreased feed intake and changes in milk quality and production. On 25 March 2024, HPAI H5N1 clade 2.3.4.4b was confirmed in dairy cattle in Texas. Shortly thereafter, the virus was identified in cattle in eight other United States (US) states by members of the National Animal Health Laboratory Network. The goal of this study was to analyze genetic sequence data collected after the introduction of HPAI H5N1 in late 2021 into the Atlantic flyway of North America and its onward circulation and reassortment with North American wild bird–origin low-pathogenicity viruses. These data were combined with whole-genome sequence data and epidemiological information from the HPAI H5N1 outbreak among US dairy cattle to help us understand when the interspecies transmission event to cattle occurred and the consequences of animal movement for virus spread.

RESULTS: H5N1 clade 2.3.4.4b genotype B3.13 influenza A virus was confirmed in milk with limited detections in nasal swabs. The initial outbreak included samples from 26 dairy cattle premises across eight states and six poultry premises in three states. The sequences isolated from cattle clustered within a single group in phylogenetic analyses, supporting a single spillover event in late 2023. A reassortment event preceded the spillover, and after introduction, the virus persisted in cattle with evidence for transmission from cattle into poultry and peridomestic animal species. Epidemiological records and phylodynamic modeling documented that the movement of asymptomatic or presymptomatic dairy cattle resulted in the dissemination of the virus across the US. We identified low-frequency within-host sequence variants across the genome that were associated with changes in virulence, host-range specificity, and mammalian adaptation.

CONCLUSION: A single wild bird-to-cattle transmission event of HPAI H5N1 clade 2.3.4.4b occurred in late 2023. The spillover was likely preceded by a reassortment event in wild bird populations followed by the movement of cattle that spread HPAI within the US dairy herd. Molecular markers that may lead to changes in transmission efficiency and phenotype were detected at low frequencies. Continued transmission of H5N1 HPAI within dairy cattle increases the risk for infection and subsequent spread of the virus to humans and other host populations. ∎
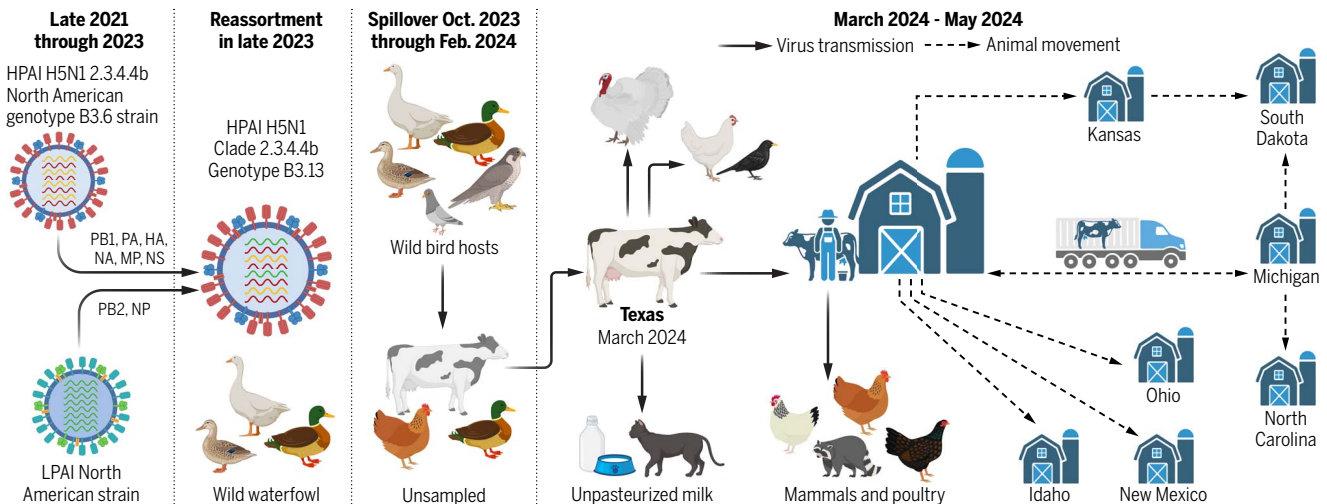
The list of author affiliations is available in the full article online.
*Corresponding author. Email: tavis.anderson@usda.gov (T.K.A.); mia.kim.torchetti@usda.gov (M.K.T.); suelee.robbe-austerman@usda.gov (S.R.-A.)
†These authors contributed equally to this work.
Cite this article as T.-Q. Nguyen *et al.*, *Science* **388**, eadq0900 (2025). DOI: 10.1126/science.adq0900

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adq0900

The emergence and spread of highly pathogenic avian influenza virus H5N1 clade 2.3.4.4b in dairy cattle in the US. Highly pathogenic avian influenza (HPAI) H5N1 was detected in North America in late 2021, and epidemiological information and genomic analysis demonstrated a single spillover into dairy cattle by a reassorted HPAI H5N1 genotype B3.13 virus. The virus disseminated across the US through the movement of asymptomatic or presymptomatic animals and was subsequently reintroduced from dairy cattle back into other host species.

RESEARCH ARTICLE

AVIAN INFLUENZA

Emergence and interstate spread of highly pathogenic avian influenza A(H5N1) in dairy cattle in the United States

Thao-Quyen Nguyen[1,2]†, Carl R. Hutter[1]†, Alexey Markin[1], Megan Thomas[1], Kristina Lantz[3], Mary Lea Killian[3], Garrett M. Janzen[1], Sriram Vijendran[1,2], Sanket Wagle[1,2], Blake Inderski[1], Drew R. Magstadt[4,5], Ganwu Li[4], Diego G. Diel[6], Elisha Anna Frye[6], Kiril M. Dimitrov[7], Amy K. Swinford[7], Alexis C. Thompson[8], Kevin R. Snekvik[9,10], David L. Suarez[11], Steven M. Lakin[3], Stacey Schwabenlander[12], Sara C. Ahola[13], Kammy R. Johnson[13], Amy L. Baker[1], Suelee Robbe-Austerman[3]*, Mia Kim Torchetti[3]*, Tavis K. Anderson[1]*

Highly pathogenic avian influenza (HPAI) viruses cross species barriers and have the potential to cause pandemics. In North America, HPAI A(H5N1) viruses related to the goose/Guangdong 2.3.4.4b hemagglutinin phylogenetic clade have infected wild birds, poultry, and mammals. Our genomic analysis and epidemiological investigation showed that a reassortment event in wild bird populations preceded a single wild bird–to-cattle transmission episode. The movement of asymptomatic or presymptomatic cattle has likely played a role in the spread of HPAI within the United States dairy herd. Some molecular markers that may lead to changes in transmission efficiency and phenotype were detected at low frequencies. Continued transmission of H5N1 HPAI within dairy cattle increases the risk for infection and subsequent spread of the virus to human populations.

Highly pathogenic avian influenza (HPAI) viruses have critical consequences for animal health and the agricultural economy and may have pandemic potential. HPAI related to the goose/Guangdong 2.3.4.4 hemagglutinin (HA) H5NX phylogenetic clade has spread to nearly 100 countries (1), causing infections resulting in mortality events, and is recognized as a panzootic, crossing multiple species barriers. HPAI virus circulation is ongoing in Europe and North America, with recent data indicating a shift in biology and transmission (2). After an initial trans-Atlantic incursion in late 2021 (3, 4), the HPAI H5N1 clade 2.3.4.4b virus caused widespread outbreaks across North America (5). The outbreaks resulted in extensive mortality events in wild bird and mammal species (6, 7), mortality and culling of poultry when detected in agricultural systems, a large number of interspecies transmission events into wild mammals (8), and human infections (9, 10).

The frequent and recent transmission of HPAI clade 2.3.4.4 between avian species in North America has resulted in the emergence of substantial genetic diversity. Experimental studies on some viruses from the HPAI clade 2.3.4.4 have shown that they can bind to both human α2,6-linked and avian α2,3-linked sialic acid receptors (11, 12). Genomic analyses have documented that approximately half of the sequences from mammals globally within the HPAI H5N1 clade 2.3.4.4b have amino acid signatures in the polymerase basic 2 (PB2) protein that have been associated with mammalian

adaptation through enhanced viral replication, host-specific polymerase activity, and temperature sensitivity [E627K (E627→K), D701N, and/or T271A] (5, 13). Additionally, the introduction of HPAI H5N1 into farmed mink in Europe in 2022 provided evidence that transmission to, and within, a population of mammalian hosts could result in mutations in the HA associated with human receptor recognition and in the neuraminidase (NA) protein that affected sialic acid binding in a manner similar to human influenza A viruses (14–16). These results are particularly important because from January 2022 to 1 April 2024, there were 13 reported human cases of H5N1 from the HPAI clade 2.3.4.4b worldwide, with some having severe consequences, including mortality. Consequently, it is critical to determine how evolution of the HPAI clade 2.3.4.4b in wild birds and the associated spillovers and transmission in mammals impacts genomic and phenotype features that alter the potential for human infection and transmission (17, 18).

On 25 March 2024, HPAI H5N1 clade 2.3.4.4b was confirmed in dairy cattle in Texas after reports of decreased milk yields in Texas, Kansas, and New Mexico (19). Shortly thereafter, the virus was identified in cattle in eight other United States (US) states by members of the National Animal Health Laboratory Network (NAHLN) (19–21). The virus was found in mammary tissue and milk; these samples were collected to determine whether bacterial or viral pathogens were driving milk yield changes (19). It was also detected in cats and peridomestic

animals on affected premises. Overall, the detection of influenza A virus (IAV) in cattle has been rarely documented (22, 23), but there is prior evidence for its replication within the mammary gland (22, 23), as well as association with a reduction in milk yield (24), and experimental studies have shown that bovine calves are susceptible to infection and may asymptomatically shed virus (25). The goal of this study was to analyze genetic sequence data collected after the introduction of HPAI H5N1 in late 2021 into the Atlantic flyway of North America and its onward circulation and reassortment with North American wild bird–origin, low-pathogenicity viruses that resulted in more than 100 distinct genotypes in the US (26). These data were combined with newly generated whole-genome sequence data and epidemiological information from the outbreak among US dairy cattle to help us understand when the interspecies transmission event occurred and the consequences of animal movement for the persistence of the virus. We performed phylodynamic analysis of the US HPAI H5N1 viruses detected in dairy cattle alongside epidemiologically linked wild bird, poultry, and peridomestic animal data. We also developed a within-host evolutionary assessment of the virus to determine how transmission in dairy cattle affects genomic diversity and whether this increases the potential for this host to serve as a reservoir for zoonotic IAV.

H5 clade 2.3.4.4b introduction into the US

The H5N1 clade 2.3.4.4b genotype A1 (26) was first identified in American wigeon (*Mareca americana*) and blue-winged teal (*Spatula*

[1]Virus and Prion Research Unit, National Animal Disease Center, Agricultural Research Service, United States Department of Agriculture, Ames, IA, USA. [2]Department of Computer Science, Iowa State University, Ames, IA, USA. [3]National Veterinary Services Laboratories, Animal and Plant Health Inspection Services, United States Department of Agriculture, Ames, IA, USA. [4]Veterinary Diagnostic Laboratory, College of Veterinary Medicine, Iowa State University, Ames, IA, USA. [5]Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University, Ames, IA, USA. [6]Department of Population Medicine and Diagnostic Sciences, Animal Health Diagnostic Center, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA. [7]Texas A&M Veterinary Medical Diagnostic Laboratory, College Station, TX, USA. [8]Texas A&M Veterinary Medical Diagnostic Laboratory, Canyon, TX, USA. [9]Department of Veterinary Microbiology and Pathology, College of Veterinary Medicine, Washington State University, Pullman, WA, USA. [10]The Washington Animal Disease Diagnostic Laboratory, College of Veterinary Medicine, Washington State University, Pullman, WA, USA. [11]Southeast Poultry Research Laboratory, National Poultry Research Center, Agricultural Research Service, United States Department of Agriculture, Athens, GA, USA. [12]Ruminant Health Center, Animal and Plant Health Inspection Services, United States Department of Agriculture, Riverdale, MD, USA. [13]Field Epidemiologic Investigation Services, Animal and Plant Health Inspection Services, United States Department of Agriculture, Ft. Collins, CO, USA.
*Corresponding author. Email: tavis.anderson@usda.gov (T.K.A.); mia.kim.torchetti@usda.gov (M.K.T.); suelee.robbe-austerman@usda.gov (S.R.-A.)
†These authors contributed equally to this work.

discors) collected in December 2021 (*3*). More than 100 genotypes representing different gene constellations have been characterized within the US, but 70% of all viruses fall into only seven genotypes (*7*). Despite genetic differences, spillover events have been associated with the predominance of a genotype rather than any specific link between genotype and host specificity. The introduction of genotype A3 was identified in April 2022 likely via the Pacific flyway and was followed by genotype A4. Another two introductions were identified to have entered the US via the Atlantic flyway, including an A6 Eurasian virus that had a reassorted Eurasian NA (H5N5). Widespread detections in wild bird populations (*27*) continue to result in point source spillovers to poultry (*28*), but there is limited evidence for lateral transmission among poultry flocks. Viruses of multiple genotypes have also been detected in ~20 mammal species in the US, often related to mortality or severe neurologic signs, but these appear to be dead end hosts (*8*, *29*). In early March 2024, HPAI H5N1 was identified in neurologic goat kids on a farm where poultry had recently been depopulated for HPAI control measures (*30*); this event was unrelated to the dairy event and involved a different virus genotype.

Detection of HPAI H5N1 in dairy cattle

In the US, nearly 10 million dairy cattle are raised in farms across all 50 states (*31*, *32*).

Dairy production is concentrated in California, Wisconsin, New York, Idaho, and Pennsylvania (*32*), but these locations are within a network of cow-calf operations that are most frequently located in Texas, Oklahoma, Missouri, Nebraska, South Dakota, and Kansas (*31*). Of the more than 60,000 dairy premises, ~30% house more than 2000 animals (*32*, *33*). Farming is characterized by extensive animal movement within and between farms to facilitate production; for example, nonlactating young stock may move off-farm for rearing with the heifers, returning to the operation by calving or before (*33*). Nearly all dairy operations report animal replacement in the herd, with 40% of replacements raised off-site (*32*).

In late January 2024, production veterinarians observed dairy cattle displaying reductions in milk production, decreased feed intake, and changes in milk quality (*19*). On 20 March 2024, members of the NAHLN identified IAV in milk and in a few nasal swabs from a Texas dairy and forwarded samples to the National Veterinary Services Laboratories (NVSL) for confirmatory testing on 25 March 2024. Testing revealed the presence of H5N1 clade 2.3.4.4b genotype B3.13. Shortly after the identification of HPAI H5N1 genotype B3.13 in Texas, it was confirmed in additional Texas herds and herds in other states (*21*). Samples collected in the initial outbreak between 7 March 2024 and 8 April 2024 have virus characterized as genotype B3.13 that in-

cluded 26 dairy cattle premises across eight states and six poultry premises in three states (data S1). The NVSL conducted whole-genome sequencing and analysis, using a custom software program that identified transmission chains on the basis of genomic similarity (*34*) to provide rapid feedback in support of epidemiologic investigations (data S2 and fig. S1). The phylogenetic and available epidemiologic data derived from questionnaires with dairy producers indicated that the genotype B3.13 virus was moved between dairy cattle premises, as well as poultry premises, via multiple transmission routes (Fig. 1 and fig. S1) (*35*, *36*). Detection of the B3.13 genotype in cattle locations that have no known epidemiologic links to confirmed premises indicates there are affected herds that have not yet been identified.

A single interspecies transmission event

To determine when the HPAI H5 clade 2.3.4.4b virus was introduced into cattle in the US, we conducted a phylogenetic analysis using whole-genome sequence data collected from poultry, wild birds, and mammals (Fig. 2 and fig. S2). From 2022 to the present, clade 2.3.4.4b viruses have been reported in more than 9000 wild birds in at least 163 species across 49 states and Washington, DC, in more than 200 mammals in at least 20 species across 29 states and Washington, DC, and more than 1120 poultry flocks across 48 states. In the time-scaled
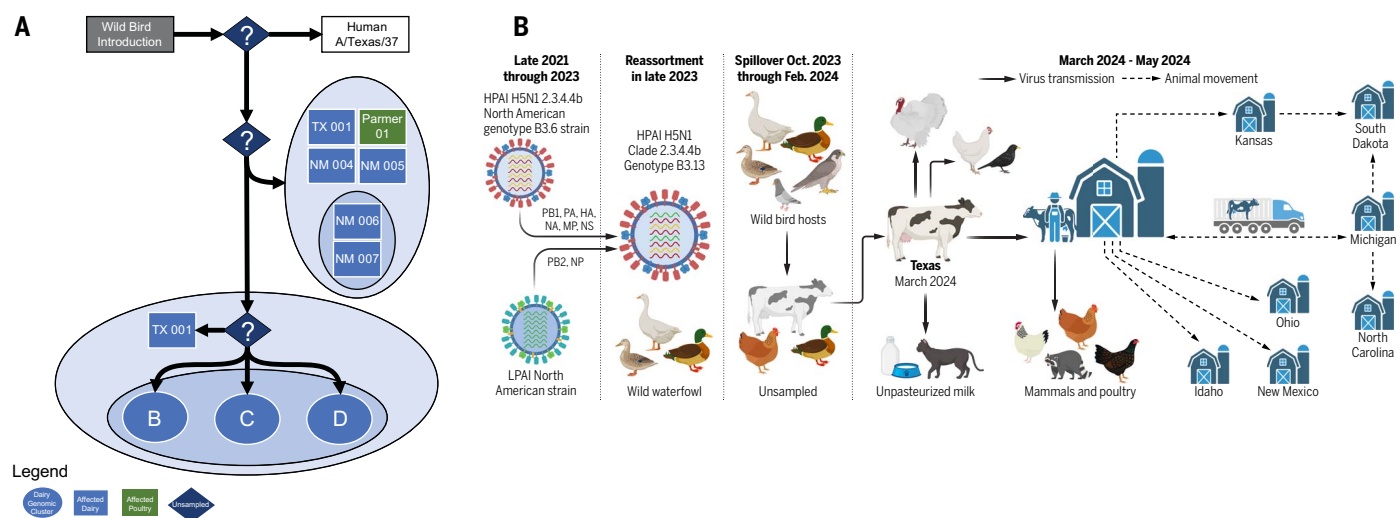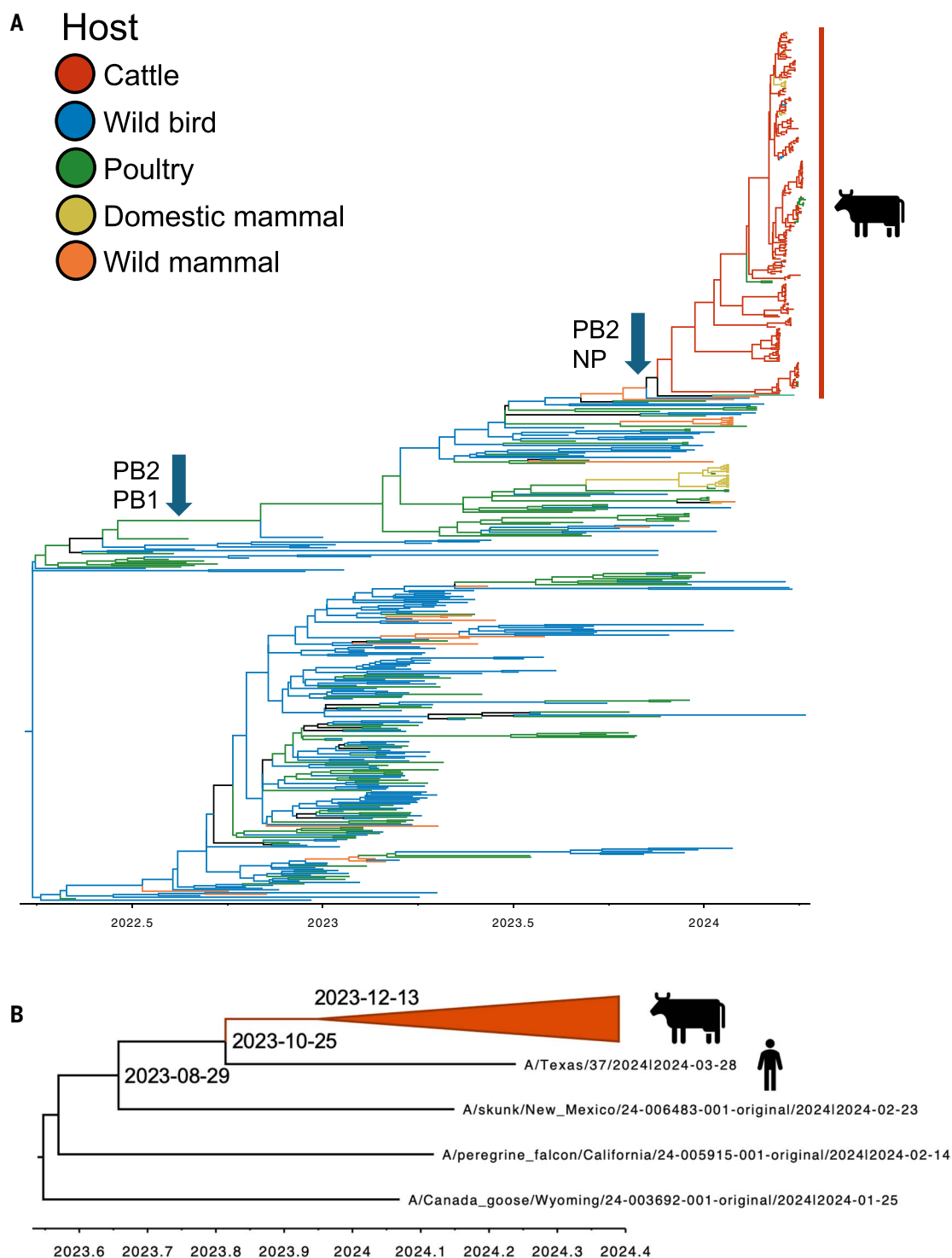

Fig. 1. Putative transmission pathways of HPAI H5N1 clade 2.3.4.4b genotype B3.13 supported by epidemiological links, animal movements, and genomic analysis. Transmission routes were derived from interviews with producers and analyses of genetic sequence data to support a single interspecies transmission event followed by lateral spread within dairy cattle with onward transmission to poultry. (**A**) The gray box represents an unsampled wild bird–origin common ancestor. The white box is the B3.13 genotype human detection distinct from the sampled cattle and poultry clusters. Two major groups emerged from an unsampled common ancestor: one containing dairy premises in Texas and New Mexico along with Parmer 01 poultry, and another containing TX 001 and three clusters labeled B, C, and D, respectively. Links were determined through phylogenetic inference and through epidemiological interviews

to identify animal movements and nonanimal connections between premises. Locations that appear more than once indicate the presence of genetically distinct viruses sequenced from independent samples on those premises. Unsampled common sources are indicated by question marks within dark blue diamonds. A visualization of all epidemiological links between dairy and poultry locations, including the transmission within the B, C, and D groups, is presented in the supplemental figures. (**B**) A conceptual overview of the emergence and interstate spread of the B3.13 genotype; these data demonstrate reassortment to acquire new polymerase basic 2 (PB2) protein and nucleoprotein (NP) within a wild bird host, the subsequent dissemination of the virus across the US through asymptomatic or presymptomatic dairy, and the introduction of HPAI H5N1 from dairy cattle back into other host species.

Fig. 2. Evolutionary history of H5N1 2.3.4.4b in North America before and during the introduction and emergence in US cattle. (**A**) A Bayesian time-scaled phylogeny of the HA gene between December 2021 and April 2024, demonstrating the single introduction of the virus from wild birds to dairy cattle with an estimated date of November 2023 (95% credible interval: October 2023 to January 2024). B3.13 strains inherited PB1, PA, HA, NA, MP, and NS genes from a B3.6 ancestor and acquired different PB2 and NP genes from North American LPAI viruses. The Bayesian phylogeny of the hemagglutinin (HA) gene was paraphyletically subsampled to maintain tree topology and demonstrate the 16 subsequent spillovers from cattle to domestic cats, poultry, and peridomestic animals and is presented in the supplemental figures. (**B**) A simplified visualization of a hypothesis on the evolutionary history of the H5N1 2.3.4.4b B3.13 genotype that emerged in cattle and a range of other mammalian and avian hosts. This time-scaled phylogeny was inferred with maximum likelihood methods from concatenated genomes: Inferred dates are presented at nodes for the most recent common ancestor, and these are congruent with those estimated using Bayesian methods on the HA gene tree presented in (A). The monophyletic clade of B3.13 genotype viruses was collapsed, and the group was presented as a single orange triangle.

Bayesian HA gene phylogeny, the H5N1 2.3.4.4b sequences isolated from cattle clustered within a single monophyletic clade, which supports a single spillover event followed by lateral transmission. Introduction of HPAI from unsampled poultry populations represents an alternate hypothesis, but a single spillover event was also supported by the maximum likelihood phylogeny reconstructed from concatenated whole-genome sequences of B3.13 genotype strains (Fig. 2B), and this pattern was evident in the gene segment maximum likelihood phylogenies: polymerase basic 2 (PB2) protein, polymerase acidic (PA) protein, nucleoprotein (NP), NA, matrix protein (MP), and nonstructural (NS) protein (figs. S3 to S6). For the polymerase basic 1 (PB1) gene segment, >97% of sequences from cattle isolates formed a single mono-phyletic clade; the remaining sequences were placed in a parental clade owing to the high genetic similarity (fig. S6). The single-introduction hypothesis was further supported by an analysis of single-nucleotide polymorphisms (SNPs) (data S2 and fig. S1). A subsequent maximum likelihood phylogenetic analysis on $n = 1156$ genomes of B3.13 genotype strains further supported the proposition of a single origin

from an avian reservoir host (figs. S7 and S10). The inferred evolutionary rate for the HA gene tree was 6.19×10^{-3} substitutions per site/year [95% highest posterior density (HPD), 5.21×10^{-3} to 7.19×10^{-3}]; the estimated evolutionary rate for the cattle clade was not plausible (3.96×10^{-2}, 95% HPD, 2.13×10^{-2} to 6.05×10^{-2}) and likely requires a different dataset or analytical technique. The time to the most recent common ancestor (TMRCA) for the HA segment estimated with Bayesian methods for the cattle clade was estimated as 10 November 2023 (95% HPD, 4 October 2023 to 22 January 2024). A clade 2.3.4.4b genotype B3.13 infection in a dairy worker was diagnosed at the end of March 2024 and was similar to the HPAIV from cattle (10). The TMRCA estimated with Bayesian methods for the cattle group and the human virus was estimated as 27 September 2023 (95% HPD, 18 August 2023 to 9 December 2023). These two groups shared a TMRCA that was estimated as 14 August 2023 (95% HPD, 23 July 2023 to 19 October 2023). These are broad TMRCA estimates and suggest that after the

B3.13 introduction into cattle, there was limited local circulation for >4 months. The Bayesian estimates were congruent with TMRCA inferred by using maximum likelihood methods (Fig. 2B); the TMRCA for each of the gene segments [the TMRCA ranged from 4 October 2023 to 13 February 13 2024 (37)] is feasible given epidemiological information (Fig. 1) and is consistent with independent estimations (38).

Cattle transmission to peridomestic species

The phylogenetic tree topology indicated that after introduction, the virus persisted in cattle populations, with subsequent evidence for transmission from cattle into poultry and peridomestic animal species (Figs. 1 and 2 and figs. S2 and S7). There were as many as nine cattle to poultry, one cattle to raccoon, two cattle to cats, and three cattle to wild bird transmission events. The wild bird transmission events were restricted to the common grackle (*Quiscalus quiscula*), blackbirds, and pigeons. These animals were collected from premises with cattle where genotype B3.13 HPAIV was identified.

Our Bayesian discrete state analysis (Fig. 3) that quantified the movement of HPAIV between six different host categories (poultry, wild bird, cattle, wild mammals, cat, and humans) provided sufficient evidence to verify that HPAI in cattle resulted in infections in other hosts. We cannot exclude the possibility that this genotype is circulating in unsampled locations and unknown hosts, because the existing analysis suggests that data are missing and because incomplete surveillance may obscure transmission inferred with phylogenetic methods (39). The gap in data is highlighted by the human infection with genotype B3.13 HPAIV, in which the HA gene sequence did not nest within the virus HA gene sequences from cattle. This could indicate that HPAIV in unsampled cows was the source of infection or that within-host evolution resulted in divergence sufficient to result in a different phylogenetic grouping. It is most likely, however, that asymptomatic transmission and lack of surveillance in epidemiologically important populations drove this pattern. Our analysis of transmission
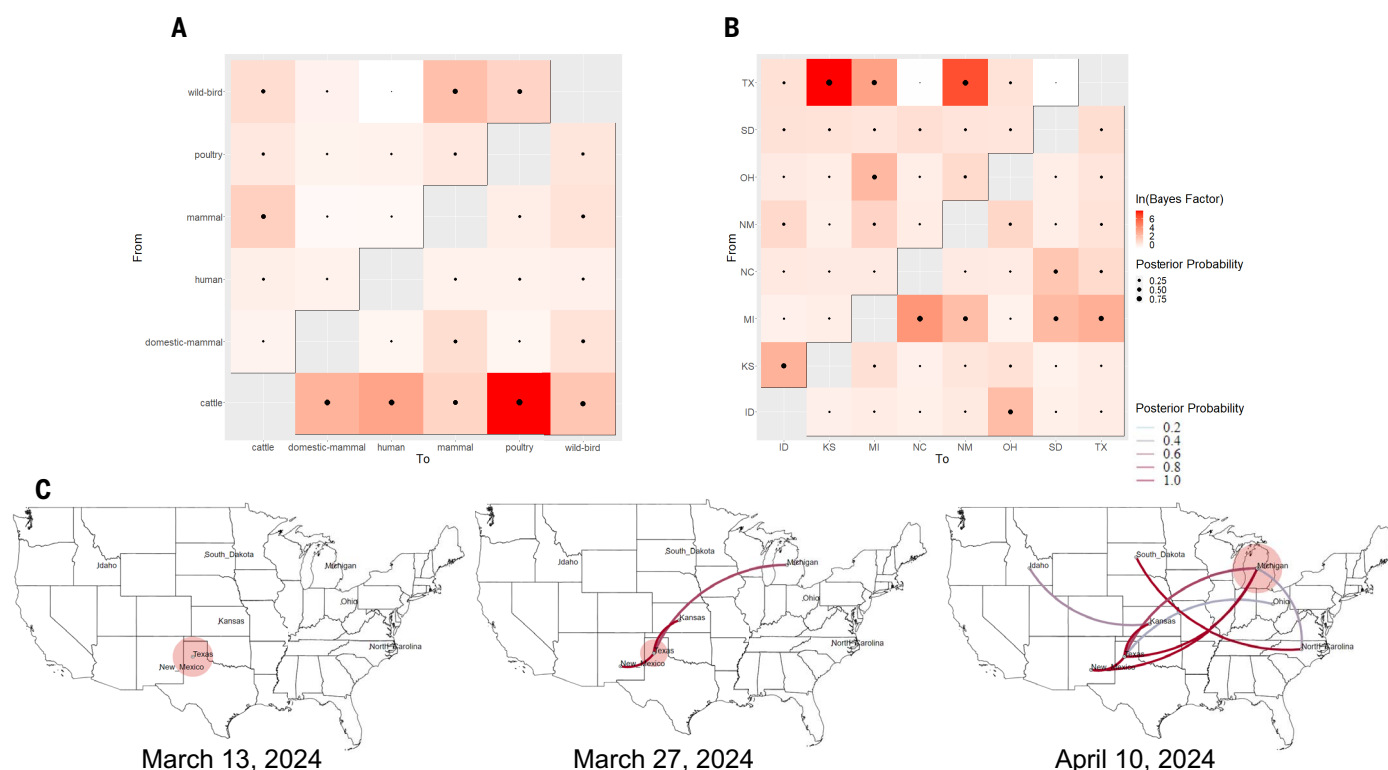

Fig. 3. Bayes factors for inferred movement between different discrete traits of H5N1 clade 2.3.4.4b viruses, demonstrating the frequency of movement. Bayes factors and posterior probabilities were plotted to demonstrate the frequency and support for the movement of the clade 2.3.4.4b B3.13 HA gene between different host species (**A**) or US locations with documented detections of 2.3.4.4b (**B**), with the source host species or US state plotted on the y axis and the recipient host species or US state plotted on the x axis. This analysis was conducted on the subsampled and aligned HA dataset (A) and an extracted subset of the data that covered the monophyletic clade of cattle HA clade 2.3.4.4b genes (B). When imposing a Bayes factor >3 with an associated posterior probability >75%, there were supported transitions from Texas to Kansas, New Mexico, and Michigan, and from Michigan to North Carolina. (**C**) Transitions between locations were inferred with a discrete trait model across time and by using 14-day time points of the spatial spread corresponding to the beginning, middle, and end of our dataset. The centroid of each state was used: For Texas, we located the point in the panhandle where dairy cattle are commonly farmed. The circular polygon reflects the number of lineages within a location, and eastward movements are depicted by lines with a concave curve, whereas westward movements are depicted by lines with a convex curve.

chains within the cattle B3.13 clade using a phylogenomic maximum likelihood approach suggested unsampled transmission in late 2023 and early 2024 (fig. S8), and the TMRCA indicates that there was >4 months of circulation before confirmation. However, given the decline in milk production in monitored dairy herds, it is unlikely that the spillover occurred outside of the described TMRCA ranges.

Reassortment in migratory bird populations

From January 2022 to April 2024, more than 482 commercial and 645 backyard flocks tested positive for HPAI (*40*). A major component in the dissemination of the 2.3.4.4b viruses is movement across four migratory flyways and the infection of nearly 200 wild avian species (*4, 41*). Extensive genetic reassortment with existing North American wild bird LPAIVs is a consequence of the host breadth and the geographic breadth. In our analysis, reassortment generated a spectrum of different genotypes, but the HA, NA, and M genes had preferential pairings in a manner that also occurs in mammalian adapted IAVs (*3, 42*). There was transient detection of reassorted viruses with different combinations of PB2, PB1, PA, NP, and NS genes (*5, 28*). We detected 243 putative reassortment events across the HA phylogeny (fig. S9). These events were associated with PB2 (137/243), PB1 (78/243), PA (36/243), NP (126/243), and NS (52/243) gene segments. With current surveillance data, there was no evidence that most of these events persisted. However, there were 24 reassortment events that resulted in new genotypes with more than 20 downstream detections. The spillover into cattle was preceded by a reassortment event(s) involving different PB2 and NP genes, likely derived from wild bird LPAI in late 2023. This reassortment resulted in the B3.13 genotype, which is maintained across the clade of epidemiologically linked cattle samples and shows no evidence for further reassortment after the spillover. The NP gene acquired by the reassortment event may have resulted in a phenotype change that mediated the emergence of this virus genotype in cattle. There are no comparative phenotypic data to support this assertion, but the NP gene is associated with multiple processes in the virus life cycle (*43*) and was implicated in increasing the transmission efficiency of IAV in the swine host (*44*). An alternate explanation is that the new reassorted virus was a passenger on an overlapping set of ecological parameters (e.g., environmental contamination from infected wild birds) that resulted in the HPAI H5N1 spillover.

HPAI H5N1 dispersal with cattle movements

Epidemiological records documented dairy cattle movement from a Texas herd (at shipment, HPAI status unknown) to North Carolina and Idaho (Fig. 1 and fig. S1). Records also indicated that asymptomatic or presymptomatic dairy cattle—some that were subsequently diagnosed as HPAIV positive—were moved from a Texas herd to Michigan and to Ohio and that a Kansas herd was moved to South Dakota. We also applied an asymmetric discrete-trait model with Bayesian stochastic search variable selection to reconstruct how clade 2.3.4.4b HPAIV H5 isolated in cattle moved among the eight US states (Idaho, Texas, Kansas, Michigan, New Mexico, North Carolina, Ohio, and South Dakota). The interstate HPAIV movement, which is inferred by these phylodynamic techniques (Fig. 3), showed that after the first confirmed case in Texas, HPAIV moved rapidly across the US. The phylogenetic signal within the HA gene of B3.13 genes in cattle was relatively low, and we conservatively identified when a state transition occurred, i.e., a Bayes factor >3 with an associated posterior probability >75%. Using these thresholds, there was phylogenetic evidence for the movement of the B3.13 genotype from Texas to Kansas, Michigan, and New Mexico. Evidence in the phylodynamic analyses indicated that after its introduction into Michigan, B3.13 was moved into North Carolina. However, it was more likely that a confirmed Texas-to-North Carolina animal movement moved the virus, because we were unable to confirm a link between Michigan and North Carolina dairy herds (Fig. 1 and fig. S1). In a phylogenomic analysis (fig. S10), there were cohesive groups of sequences isolated from each US state; this suggested that after mixing and movement of animals in March and April 2024, a spatially structured host and virus population emerged. These data support the proposition that a single genotype was introduced and that direct movement of cattle—based upon production practices—created opportunities for virus dissemination that have since been reduced. The movement of viruses from one region to another will provide opportunities for reassortment and subsequent changes in genomic diversity. Areas that received large shipments of cattle and their viruses will provide additional opportunities for reassortment and the emergence of IAV strains with phenotypes that may have increased zoonotic potential.

Low-frequency genomic variants and within-host evolution

We analyzed 227 cattle viral whole genomes collected during the first phase of the outbreak between 7 March and 3 April 2024, and we include data from an additional 1250 viral genomes collected from cattle between 4 April and 29 October 2024. We identified within-host sequence variants across the genome that were present in >0.5% of whole-genome sequencing reads, matching to the Influenza Research Database Sequence Feature Variant Tool (*45*) and a literature search (table S1). If a sequence variant were to arise that provided a selective advan-

tage, it could increase in frequency through transmission and potentially alter the virus phenotype. Our analyses were restricted to field-collected specimens, with no ability to assess whether single-nucleotide variants (SNVs) affected virus attributes. In addition, most SNVs were present only at low frequencies (data S5). In the cattle sequence data, there were 822 amino acid sites with nonsynonymous amino acid changes; of the variable amino acid sites, 406 nonsynonymous mutations occurred at sites associated with functional changes (mean 44.32 ± 5.06 potential functional changes per sample; range: 24 to 64). An additional 605 synonymous mutations were detected in the dataset. Variants were associated with changes in virulence in HA, MP, NP, and PB2 (e.g., in Table 1: Q134K, Q154R, Q234K, and P337L in HA; S207G in MP; D701N in PB2; see tables S2 to S5 for other animal groups). We also detected SNVs previously associated with host-range specificity (e.g., E229K in NS). In the PB2, variants were associated with mammalian adaptation (an increase in E627K, D701N, and V495I), in which the frequency was 33% in a single animal for E627K. We did not detect the mammalian adaptation PB2 271A in cattle, even at low frequencies; however, one mammalian sample contained the mutation in the consensus gene, and we detected mutations on HA that affected receptor binding affinity (Table 1). Within the PB2, there was the detection of 631L (one strain had 631P), an amino acid position that has been associated with mammalian transmission (*46*). When analyzing an additional 1250 cattle samples collected between 4 April and 29 October 2024, 10 low-frequency SNVs increased to higher frequencies; these included virulence markers in HA (position 336) and NS (position 229) and a mammalian adaptation marker in NA (position 430). We observed 12 other nonsynonymous SNVs with unknown functions that increased in frequency, with 21 SNVs increasing to >5% of the total cattle samples: These may reflect a founder effect or evolution associated with phenotypic change (data S6). Determining the biological relevance of these sequence variants requires additional experimental study and should be monitored because some have been associated with transmission efficiency and mammalian adaptation. We also calculated synonymous (π_S) and nonsynonymous (π_N) site ratio to assess natural selection; π_N/π_S <1 is suggestive of purifying selection, and π_N/π_S >1 is suggestive of positive selection. When we combined the diversity estimates across genes, all cattle strains exhibited π_N/π_S <1; the gene π_N/π_S estimates varied, with none having >1 (HA, π_N/π_S = 0.042; MP, π_N/π_S = 0.566; NA, π_N/π_S = 0.188; NP, π_N/π_S = 0.032; NS, π_N/π_S = 0.355; PA, π_N/π_S = 0.160; PB1, π_N/π_S = 0.048; PB2, π_N/π_S = 0.039), suggesting that the within-host virus populations tended to exhibit

Table 1. Sequence variants detected in sites within H5N1 2.3.4.4b strains isolated in cattle that have been associated with host adaptation and virulence. Raw read data from cattle samples were processed, and high- and low-frequency single-nucleotide variants (SNVs) were identified relative to the most recent common ancestor of the phylogenetic group in this study. The SNVs that induced a coding region change were screened against a database of positions associated with functional change, with a relevant selection shown here. The number of cattle samples with the SNVs were enumerated: The first number represents data from March and April 2024, and the number in parentheses represents data from April to October 2024. The mean allele frequency was calculated, the presence of the mutation within the consensus gene sequence was determined, and the variants detected at low frequencies were counted. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. adapt., adaptation; resist., resistance.

Gene	Coding-region change	Functional type	Cattle with variant (no.)	Mean allele frequency	Consensus sequence	Low-frequency variants
HA	E91K	Mammal adapt.	1 (1)	0.05	0	1
HA	S137F	Mammal adapt.	1 (1)	0.376	0	1
HA	Q154R	Pathogenicity	5 (9)	0.013	0	6
HA	N209T	Mammal adapt.	1 (3)	0.012	0	1
HA	Q234K/R	Virulence	8 (32)	0.039	0	9
HA	G240R	Mammal adapt.	1 (1)	0.025	0	1
HA	S336N	Virulence	18 (245)	0.892	15	2
HA	P337L	Virulence	8 (21)	0.715	5	2
MP	R77K	Virulence	1 (7)	0.006	0	1
NA	T438A/I	Antiviral resist.	3 (9)	0.627	2	1
NA	R430K	Mammal adapt.	1 (82)	0.130	0	1
NS	D125N/G	Virulence	27 (20)	0.873	24	3
NS	E229K	Virulence	21 (85)	0.999	18	0
PB2	R389K/G	Mammal adapt.	2 (5)	0.012	0	2
PB2	E627K	Virulence/adapt.	1 (12)	0.329	0	1
PB2	D701N	Virulence	2 (3)	0.015	0	2

weak purifying selection when compared with an ancestral reference.

Discussion

The HPAI H5N1 genotype B3.13 viruses circulating in cattle represent a pandemic threat, given the evidence of cattle-to-human transmission events and persistence across the US (*10*). On the basis of current information, once infected, a cow may shed virus for 2 to 3 weeks, increasing the window in which zoonotic transmission may occur (*47*, *48*). Additionally, in our field data, we detected some amino acid mutations at sites associated with mammalian adaptation that had already become fixed in the virus population. This result likely reflects the ~4 months of evolution in dairy cattle. Notably, low-frequency sequence variants within cattle were also detected, even within the limited time after the first spillover. The majority of these sequence variants have not increased in detection frequency; however, some changes increased in prevalence over time (data S6). The potential remains for these viruses to acquire other genetic markers that mediate the probability of transmission in mammalian hosts. Experimental studies have demonstrated

limited airborne transmission efficiency in cattle, the B3.13 genotype has rarely been detected in migratory birds, and the majority of cases are restricted to affected dairy or poultry premises (*47*, *48*). These findings suggest that this remains a distinctive spillover with potential for control (*49*). However, turkeys and pigs may be colocated on dairy premises, each species may be infected with IAV, and coinfection could result in reassortment and the emergence of new strains with elevated zoonotic potential (*50*, *51*). Monitoring of cattle and other agricultural animals will inform future risk assessments (*52*) and provide an early warning for whether this interspecies transmission event and dissemination of the viruses represents a future threat to human health.

In countries with endemic HPAI H5N1 in poultry, there may be low levels of immunity against H5N1 viruses (*53*–*55*), but there is no evidence for this in the US. This is concerning because the immunological landscape in humans affects disease severity (*56*), and clade 2.3.4.4b B3.13 virus have demonstrated the capacity to cross the species barrier (*10*). The existing prepandemic candidate vaccine viruses (CVVs) do retain cross-reactivity with

currently circulating clade 2.3.4.4b HPAI H5N1 (*57*). These CVVs are coordinated and shared among the WHO Global Influenza Surveillance and Response Network for use by academic, government, and industry partners for R&D (*53*–*55*, *58*, *59*). However, recent viruses collected in the US had reduced reactivity with the A/Astrakhan/3212/2020 CVV, and on the basis of these data and other genetic and epidemiologic measures, a new clade 2.3.4.4b CVV has been proposed (*57*). An outbreak of HPAI H5N1 B3.13 genotype in humans may be detected—given passive (*60*) and active surveillance at human-animal interfaces (*9*)—and minimized through deployable medical countermeasures (*9*). However, preemptive control of animal IAV in wild and domestic populations before zoonoses is ideal. Testing of cattle before movement between states will reduce HPAI H5N1 dissemination (*61*). In addition, control strategies may be refined through passive IAV surveillance of cattle, swine, and wild birds, alongside a National Milk Testing Strategy to identify IAV transmission hotspots (*62*, *63*). After HPAI detection, poultry locations are depopulated and other farm locations can implement enhanced biosecurity efforts to control wild birds on premises, reduce shared personnel and equipment between locations, apply vaccine interventions to target circulating diversity (*64*), and decrease mixed-species farming (*35*).

In this study, we suggest that the B3.13 genotype had limited local circulation for >4 months. This suggests a lack of surveillance for IAV in nontarget hosts and wildlife species, reducing the potential to detect and subsequently control new viruses before zoonotic transmission. This problem may be rectified through surveillance strategies derived from alternate sample types (e.g., wastewater, bulk milk) alongside better support for passive surveillance of wild mammals and birds. This information will enable detection and identification of outbreak sources and provide the data necessary to monitor virus spread and evolution. Additionally, although genomic surveillance has increased (*65*), there has not been a concomitant increase in the number of phenotypically characterized animal IAV or ecological studies that systematically describe features associated with IAV interspecies transmission. This deficiency is highlighted by our detection of hundreds of reassorted IAVs, yet we are unable to determine whether these viruses possess altered attributes. Ideally, we can share and link genomic surveillance data with phenotypic data to facilitate the identification of IAV for characterization and risk assessment (*59*). More generally, our study demonstrates that IAV is a transboundary pathogen that requires coordination across regulatory agencies and between animal and public health organizations to improve the health of hosts and reduce pandemic risk.

Materials and methods

Sample isolation, whole genome sequencing and assembly

IAV extraction and reverse transcription real-time PCR (RT-rtPCR) were performed at the NAHLN member labs and the US Department of Agriculture (USDA), NVSL, according to the standard operating procedures (3). Influenza A virus RNA from samples was amplified (66) and after amplification was completed, we generated cDNA libraries for iSeq by using the Illumina DNA Sample Preparation Kit, (M) Tagmentation (Illumina, https://www.illumina.com) and the 300-cycle iSeq Reagent Kit v2 (Illumina) according to manufacturer instructions. We performed reference guided assembly of genome sequences using IRMA v1.1.4 (67). We generated 237 whole genomes collected between 7 March 2024 and 3 April 2024 (10 samples did not have sufficient depth for the SNV analysis). We subsequently analyzed public data from 4 April 2024 to 29 October 2024 to confirm initial patterns; this information encompass an additional 1250 whole genomes. All raw read data are provided within the NCBI Sequence Read Archive with consensus genomes assembled and submitted to NCBI GenBank. Data and code used in this manuscript are archived at a GitHub repository (37).

Field epidemiological investigations

Dairy herds were identified following reporting of clinically sick cattle to local animal health officials and confirmatory testing at the USDA, NVSL. Following confirmation of a positive detection of HPAI H5N1, information was collected by the administration of the USDA Dairy Cattle Emerging Health Event: Epidemiological Questionnaire (35, 36). Completion of this document was voluntary and reliant upon the operation and the willingness of the producer. Information included details on animal movements, status and source of animals on the farm, handling of waste milk, manure disposition, the presence of other animal species on the premise, whether locations shared personnel and equipment, the number of visitors to an operation, and whether the location had observed wild bird mortality. The focus of these investigations was to identify potential transmission routes for pathogen spread between affected dairy premises and affected poultry premises (35, 36). We also determined putative transmission routes through an analysis of SNPs in the clade 2.3.4.4b consensus sequences that were collected from each cattle or poultry location. SNPs were annotated and compared using a pipeline called vSNP (68). The pipeline infers phylogenetic trees using RAxML (69) and generates tables of SNPs relative to a reference composed of two North American wild bird origin segments (PB2, NP) and six segments from a H5N1 clade 2.3.4.4b clade virus. Epidemiological data from the

questionnaires were paired with the genomic data to determine viral genome sequence similarity and to identify links between premises. Links were grouped as animal links, non-animal links that included movement distinct from cattle or poultry (e.g., shared farm personnel), and genomic epidemiological links derived from the vSNP analysis (Fig. 1, fig. S1, and data S2).

Data sources and curation for Bayesian phylodynamic analyses

To generate a reference dataset of relevant HPAI H5N1 strains, we queried Global Initiative on Sharing All Influenza Data (GISAID) (67) for all H5 genomes with eight complete gene segments available collected between the dates of 1 January 2020 and 29 March 2024 in the US. This query provided 23,322 sequences from 2915 genomes: we extracted the HA gene sequence and aligned the data using MAFFT v7.526 (70) with default parameters. We subsequently inferred a maximum likelihood phylogenetic tree with IQ-Tree v2.2.2 (71) following automatic model selection (72). Based on this tree, we classified all data using a custom python clade identifier called GenoFLU (https://github.com/USDA-VS/GenoFLU) and filtered the data to only those strains from the HPAI clade 2.3.4.4b. We then removed identical HA sequences, ensuring that we retained the cattle strains that had been published in GISAID and the associated human case (A/Texas/37/2024, GISAID accession EPI_ISL_19027114). This process resulted in a reference gene dataset of 1393 strains. For computational efficiency, these were subsampled using smot v.1.0.0 (73) maintaining representation by sampling 20% of the tips within each monophyletic clade based on an HA gene phylogeny inferred using FastTree v2.1.11 (74). To complement and provide support for Bayesian phylogenetic analyses, we used the resultant dataset of $n = 964$ strains in maximum likelihood phylogenetic analyses. For these, we inferred the ancestral sequences of North American 2.3.4.4b strains for each gene segment by using a maximum likelihood tree inferred using IQ-Tree v.2.2.2 for each gene segment as input for TreeTime v0.9.4 ancestral reconstruction functionality. The ancestral sequences were reconstructed at the deepest nodes on each gene tree that comprised at least 99% of all strains in the subsampled dataset.

Bayesian evolutionary inference and discrete phylogeography

We estimated when HPAI H5N1 clade 2.3.4.4b were introduced into dairy cattle under a Bayesian framework in BEAST v.1.10.4 (75) with the BEAGLE library (76). For this analysis, we began with the representative dataset of $n = 964$ strains, maintained all sequences associated with the dairy cattle clade, and randomly sampled taxa outside of this monophyletic group,

resulting in an aligned HA dataset of $n = 587$ HA genes. We implemented a generalized time reversible (GTR) nucleotide substitution model (77) with gamma-distributed site heterogeneity (78), an uncorrelated relaxed clock with lognormal distribution (79), and a Gaussian Markov random field (GMRF) Bayesian skyride with time aware-smoothing for the coalescent model (80). We conducted 10 independent Markov chain Monte Carlo (MCMC) sampling runs with 50 million iterations with sampling every 5000 iterations. A similar process was conducted on the remaining seven IAV gene segments: we began with the representative dataset of $n = 964$ strains, maintained all sequences associated with the dairy cattle clade, and randomly sampled taxa outside of this monophyletic group with a goal to have aligned datasets of $n \sim 750$ for each gene. For these seven genes, we implemented a GTR nucleotide substitution model (77) with gamma-distributed site heterogeneity (78), a strict molecular clock, and a GMRF Bayesian skyride with time aware-smoothing for the coalescent model. For these seven genes, we ran five MCMC chains for 50 million iterations each, sampling parameters and trees every 5000 iterations. The results were analyzed using the GMRF skyride reconstruction in Tracer v1.7.2 (81), convergence and mixing was assessed, and runs were combined to ensure an effective sample size of more than 200. A time-scaled maximum clade credibility (MCC) tree was generated using TreeAnnotator v1.8.4 using median node heights and ~20% burn-in (75). Dates of time to the most recent common ancestor (TMRCA) were inferred from the nodes between clades, using the 95% HPD as the range of uncertainty. The TMRCA dates inferred from the Bayesian analyses in BEAST were supported through a phylogenomic analysis on concatenated gene segments and gene segment analyses using the TreeTime suite (82).

To reconstruct the spatial diffusion of the H5N1 virus across states within the US and different hosts (poultry, wild bird, cattle, mammal, domestic mammal, and humans), we applied an asymmetric discrete-trait phylogeographic analyses with Bayesian stochastic search variable selection, a strict molecular clock, and an exponential growth model in BEAST v.1.10.4 (75). We conducted this analysis on the subsampled and aligned HA dataset and the extracted monophyletic clade of cattle HA clade 2.3.4.4b genes with any host category (host) or cattle only (location). In this approach, each HA gene had a US state and host group designated as a trait, and transitions from one category to another (e.g. from Texas to New Mexico) were inferred along the internal branches representing the evolutionary history of the virus. These transitions are termed Markov jumps and state change counts for the HA clade trait were reconstructed in BEAST v1.10.4.

We conducted 5 independent MCMC chains with 50 million iterations with sampling every 5000 iterations. Convergence and mixing was assessed in Tracer v1.7.2, as described above, and runs were combined to ensure an effective sampling size of more than 200. We used the resulting posterior trees to provide estimates of the ancestral region and host for each internal node. We then used SpreaD3 v.0.9.6 (*83*) to estimate the Bayes Factor (BF) from the estimated state transition rates and used these as statistical support for spatial movements: we imposed thresholds used in previously published studies, i.e., a definitive spatial movement had a BF > 100, sufficient evidence was assessed as 100 > BF > 3 (*1*, *84*).

Reassortment analyses

We assembled a separate dataset for reassortment analysis that comprehensively included all H5N1 whole genomes collected from January 2020 to 19 April 2024 through downloading n = 27,177 HPAI H5N1 gene sequences from GISAID [accessed 19 April 2024] (*85*). This search included all data from North America, and we maintained only those strains with complete sequences for each of the eight gene segments, resulting in n = 2893 whole genomes. We merged these data with the assembled whole genomes from our study and generated alignments for each gene using MAFFT v7.526 (*70*) with default parameters. We subsequently inferred phylogenetic trees for each gene segment using IQ-Tree v2.2.2 (*86*) under the GTR model of nucleotide substitutions with the stationary probabilities estimated from the empirical base frequencies and five free-rate categories (*87*). To identify reassortment events on the HA phylogeny, we applied TreeSort v.0.1.1 (*42*) with a maximum molecular clock deviation parameter of 2.5. TreeSort uses TreeTime (*82*) to estimate the substitution rates for each gene segment and identifies branches on a tree with a high signal of reassortment. We then removed n = 35 strains that were molecular clock outliers in the HA gene and were not related to the outbreak in dairy cattle and repeated the TreeSort analysis.

Phylogenomic analyses over concatenated B3.13 genomes

We concatenated the coding regions of all gene segments into whole genome sequences for all reassortment-free B3.13 strains from the reassortment analysis above. This resulted in n = 256 aligned whole genome sequences, including n = 220 cattle strains. We inferred a phylogenetic tree from these sequences using an edge-linked partition model in IQ-Tree v.2.2.2 (*88*), where every partition corresponded to an individual gene segment and was associated with a different GTR model with empirical base frequencies and five free-rate categories. IQ-Tree was run with ultrafast bootstrap with 1000 replicates (*89*). We then used the resulting tree topology as input to TreeTime v.0.9.4 (*82*) to infer a time-scaled tree of the B3.13 genomes. TreeTime was executed with an option to account for covariation in tip dates due to shared ancestry. To validate the initial phylogenomic analysis, we repeated this procedure on 29 October 2024 with an updated comprehensive dataset of B3.13 genomes collected up to 29 October 2024. The updated analysis was conducted over n = 1156 whole genomes, n = 1113 of which were cattle.

Within-host genetic diversity analysis

We developed a bioinformatics pipeline for processing, calling SNVs, and analyzing Illumina short read data for influenza A virus called "Flumina" (https://github.com/flu-crew/Flumina). The pipeline uses Python v3.10, R v4.4 (*90*), and SnakeMake (*68*) to organize programs and script execution. A custom Python script organizes raw reads for SnakeMake that subsequently executes other programs for variant calling. The pipeline cleans the raw reads of adapter contamination, low complexity sequences, and other sequencing artifacts using the program FASTP (*91*). IRMA v1.1.4 (*67*) was used on the processed reads to generate consensus contigs that were used for phylogenetic and phylodynamic analyses and other summary statistics and graphs. Following these steps, the pipeline maps reads to an ancestral reference strain (*34*), by indexing using BWA [*bwa index -a bwtsw* function; (*92*)]. Then it uses SAMtools (*71*) to create an fai index with the *faidx* function and generate a sequence dictionary for GATK v4.4 (*93*) using the *CreateSequenceDictionary* function.

For calling of high frequency SNVs, we used GATK v4.4 (*93*) following best practices for discovering and calling SNVs (*94*). The pipeline processes the cleaned reads through the GATK4 functions FastqToSam, RevertSam, and AddorReplaceReadGroups. Next, the BAM of the reads are converted back to Fastq and mapped with BWA using the function *bwa mem -M*. The mapped reads BAM file are merged with the unmapped reads BAM, and duplicate reads are marked after sorting the BAM file by coordinate with the functions *SortSam* and *MarkDuplicates*. Finally, *HaplotypeCaller* is used (parameters: -ERC GVCF -ploidy 1) to call haplotypes, and *GenotypeGVCFs* is used to genotype the sample. Specific variants were selected using the *SelectVariants* function for the SNP type of variant. Using *VariantFiltration*, a filtered set of variants were created by applying the following filters: Qual < 30 (Quality), QD < 2 (Quality by Depth), SOR < 3 (StrandOddsRatio), FS < 60 (FisherStrand), MQ < 40 (MapQuality), MQRankSum < 12.5 (MapQuality RankSumTest), and ReadPosRankSum < 8 (ReadPosRankSumTest). To call low frequency SNVs, the program LoFreq (*95*) was applied, with the similar preprocessing steps where the same BAM was used as input from *HaplotypeCaller*. A database was generated using the Sequence Feature Variant Types tool from the Influenza Research Database (*45*) for all eight genes, and SNVs associated with nonsynonymous changes in each gene were screened to determine whether these had previously been associated with phenotypic change. The numbering in this database uses the full coding region for each gene segment and H5 numbering for the HA gene (*96*). In the NA gene, amino acid changes in the stalk region were excluded due to uncertainty on phenotype impact. To estimate genome-wide estimates of natural selection, we used the program SNPGenie on the VCF files (*97*).

Inference of B3.13 transmission chains in dairy cattle

Inference of transmission chains between hosts, sampled diversity, sampling time, and generation time were estimated from the temporally scaled genome tree using the TransPhylo R package (*98*, *99*). As input, we used the B3.13 tree generated in the phylogenomic analysis. Multifurcations in the tree were suppressed to bifurcations and all edge weights were rounded to at least 1×10^{-9}. The prior generation time and sampling density were gamma-distributed parameters extracted from published studies (*75*, *76*) resulting in a mean of 5 days (*100*, *101*) and a variance of 2 days. We adapted protocol 3 (*99*) ensuring an Effective Sample Size (ESS) >200 at the end of the simulations. The MCMC simulations were run for 1×10^{7} iterations, discarding the first 5% as burn-in. After the simulations converged, the medoid tree from the posterior set of inferred trees was used as the final inferred transmission tree. Transmissions of HPAIV from host to host (e.g., wild bird to cattle) were inferred by TransPhylo (*75*, *76*) and the distribution of incident cases (sampled and unsampled) was generated by adapting protocol 4 (fig. S8) (*99*).

REFERENCES AND NOTES

1. R. Xie *et al.*, The episodic resurgence of highly pathogenic avian influenza H5 virus. *Nature* **622**, 810–817 (2023). doi: 10.1038/s41586-023-06631-2; pmid: 37853121
2. A. Pohlmann *et al.*, Has Epizootic Become Enzootic? Evidence for a Fundamental Change in the Infection Dynamics of Highly Pathogenic Avian Influenza in Europe, 2021. *mBio* **13**, e0060922 (2022). doi: 10.1128/mbio.00609-22; pmid: 35726917
3. S. N. Bevins *et al.*, Intercontinental Movement of Highly Pathogenic Avian Influenza A(H5N1) Clade 2.3.4.4 Virus to the United States, 2021. *Emerg. Infect. Dis.* **28**, 1006–1011 (2022). doi: 10.3201/eid2805.220318; pmid: 35302933
4. V. Caliendo *et al.*, Transatlantic spread of highly pathogenic avian influenza H5N1 by wild birds from Europe to North America in 2021. *Sci. Rep.* **12**, 11729 (2022). doi: 10.1038/s41598-022-13447-z; pmid: 35821511
5. S. Youk *et al.*, 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 (2023). doi: 10.1016/j.virol.2023.109860; pmid: 37572517
6. S. Lair *et al.*, Outbreak of Highly Pathogenic Avian Influenza A (H5N1) Virus in Seals, St. Lawrence Estuary, Quebec, Canada. *Emerg. Infect. Dis.* **30**, 1133–1143 (2024). doi: 10.3201/eid3006.231033; pmid: 38781927

7. W. Puryear et al., Highly Pathogenic Avian Influenza A(H5N1) Virus Outbreak in New England Seals, United States. Emerg. Infect. Dis. 29, 786–791 (2023). doi: 10.3201/eid2904.221538; pmid: 36958010

8. US Department of Agriculture Animal and Plant Health Inspection Service (USDA-APHIS), "Detections of Highly Pathogenic Avian Influenza in Mammals" (2025); https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/mammals.

9. S. Garg et al., Outbreak of Highly Pathogenic Avian Influenza A(H5N1) Viruses in U.S. Dairy Cattle and Detection of Two Human Cases - United States, 2024. MMWR Morb. Mortal. Wkly. Rep. 73, 501–505 (2024). doi: 10.15585/mmwr.mm7321e1; pmid: 38814843

10. T. M. Uyeki et al., Highly Pathogenic Avian Influenza A(H5N1) Virus Infection in a Dairy Farm Worker. N. Engl. J. Med. 390, 2028–2029 (2024). doi: 10.1056/NEJMc2405371; pmid: 38700506

11. J. A. Pulit-Penaloza et al., Characterization of highly pathogenic avian influenza H5Nx viruses in the ferret model. Sci. Rep. 10, 12700 (2020). doi: 10.1038/s41598-020-69535-5; pmid: 32728042

12. R. Yamaji et al., Pandemic potential of highly pathogenic avian influenza clade 2.3.4.4 A(H5) viruses. Rev. Med. Virol. 30, e2099 (2020). doi: 10.1002/rmv.2099; pmid: 32135031

13. C. Adlhoch et al., Avian influenza overview December 2022 - March 2023. EFSA J. 21, e07917 (2023). doi: 10.2903/j.efsa.2023.7917; pmid: 36949860

14. M. Agüero et al., Highly pathogenic avian influenza A(H5N1) virus infection in farmed minks, Spain, October 2022. Euro Surveill. 28, 2300001 (2023). doi: 10.2807/1560-7917.ES.2023.28.3.2300001; pmid: 36695488

15. E. de Vries, C. A. de Haan, Letter to the editor: Highly pathogenic influenza A(H5N1) viruses in farmed mink outbreak contain a disrupted second sialic acid binding site in neuraminidase, similar to human influenza A viruses. Euro Surveill. 28, 2300085 (2023). doi: 10.2807/1560-7917.ES.2023.28.7.2300085; pmid: 36795504

16. M. Agüero et al., Authors' response: Highly pathogenic influenza A(H5N1) viruses in farmed mink outbreak contain a disrupted second sialic acid binding site in neuraminidase, similar to human influenza A viruses. Euro Surveill. 28, 2300109 (2023). doi: 10.2807/1560-7917.ES.2023.28.7.2300109; pmid: 36795502

17. A. Kandeil et al., Rapid evolution of A(H5N1) influenza viruses after intercontinental spread to North America. Nat. Commun. 14, 3082 (2023). doi: 10.1038/s41467-023-38415-7; pmid: 37248261

18. T. Maemura et al., Characterization of highly pathogenic clade 2.3.4.4b H5N1 mink influenza viruses. EBioMedicine 97, 104827 (2023). doi: 10.1016/j.ebiom.2023.104827; pmid: 37812908

19. E. R. Burrough et al., 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 (2024). doi: 10.3201/eid3007.240508; pmid: 38683888

20. X. Hu et al., Highly Pathogenic Avian Influenza A (H5N1) clade 2.3.4.4b Virus detected in dairy cattle. bioRxiv 2024.04.16.588916 [Preprint] (2024); https://doi.org/10.1101/2024.04.16.588916.

21. USDA-APHIS, "HPAI in Livestock: Home" (2024); https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/livestock.

22. C. A. Mitchell, O. Nordland, R. V. Walker, Myxoviruses And Their Propagation In The Mammary Gland Of Ruminants. Can. J. Comp. Med. Vet. Sci. 22, 154–156 (1958). pmid: 17649046

23. C. A. Mitchell, R. V. L. Walker, G. L. Bannister, Studies relating to the formation of neutralizing antibody following the propagation of influenza and Newcastle disease virus in the bovine mammary gland. Can. J. Microbiol. 2, 322–328 (1956). doi: 10.1139/m56-037; pmid: 13316626

24. T. R. Crawshaw, I. H. Brown, S. C. Essen, S. C. Young, Significant rising antibody titres to influenza A are associated with an acute reduction in milk yield in cattle. Vet. J. 178, 98–102 (2008). doi: 10.1016/j.tvjl.2007.07.022; pmid: 17851097

25. D. Kalthoff, B. Hoffmann, T. Harder, M. Durban, M. Beer, Experimental infection of cattle with highly pathogenic avian influenza virus (H5N1). Emerg. Infect. Dis. 14, 1132–1134 (2008). doi: 10.3201/eid1407.071468; pmid: 18598640

26. USDA-APHIS, USDA-VS/GenoFLU (2025); https://github.com/USDA-VS/GenoFLU.

27. USDA-APHIS, "Detections of Highly Pathogenic Avian Influenza in Wild Birds" (2025); https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/wild-birds.

28. USDA-APHIS, "Confirmations of Highly Pathogenic Avian Influenza in Commercial and Backyard Flocks" (2025); https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/commercial-backyard-flocks.

29. E. J. Elsmo et al., Highly Pathogenic Avian Influenza A(H5N1) Virus Clade 2.3.4.4b Infections in Wild Terrestrial Mammals, United States, 2022. Emerg. Infect. Dis. 29, 2451–2460 (2023). doi: 10.3201/eid2912.230464; pmid: 37987580

30. H. Ly, Highly pathogenic avian influenza H5N1 virus infections of dairy cattle and livestock handlers in the United States of America. Virulence 15, 2343931 (2024). doi: 10.1080/21505594.2024.2343931; pmid: 38632687

31. A. H. Cabezas, M. W. Sanderson, C. Y. Lockhart, K. A. Riley, C. J. Hanthorn, Spatial and network analysis of U.S. livestock movements based on Interstate Certificates of Veterinary Inspection. Prev. Vet. Med. 193, 105391 (2021). doi: 10.1016/j.prevetmed.2021.105391; pmid: 34091089

32. USDA-APHIS, "Dairy Industry Manual: FAD PReP, Foreign Animal Disease Preparedness and Response Plan" (2011); https://www.aphis.usda.gov/sites/default/files/dairy_industry_manual.pdf.

33. D. I. Douphrate et al., The dairy industry: A brief description of production practices, trends, and farm characteristics around the world. J. Agromedicine 18, 187–197 (2013). doi: 10.1080/1059924X.2013.796901; pmid: 23844787

34. J. Hicks, T. Stuber, K. Lantz, M. Torchetti, S. Robbe-Austerman, vSNP: A SNP pipeline for the generation of transparent SNP matrices and phylogenetic trees from whole genome sequencing data sets. BMC Genomics 25, 545 (2024). doi: 10.1186/s12864-024-10437-5; pmid: 38822271

35. USDA-APHIS, "Highly Pathogenic Avian Influenza H5N1 Genotype B3.13 in Dairy Cattle: National Epidemiologic Brief" (2024); https://www.aphis.usda.gov/sites/default/files/hpai-dairy-national-epi-brief.pdf.

36. USDA-APHIS, "2024 Highly Pathogenic Avian Influenza (H5N1) - Michigan Dairy Herd and Poultry Flock Summary" (2024); https://www.aphis.usda.gov/sites/default/files/hpai-h5n1-dairy-cattle-mi-epi-invest.pdf.

37. T. Anderson, C. R. Hutter, A. Markin, T. Nguyen, flu-crew/dairy-cattle-hpai-2024: Data and Code from: Emergence and interstate spread of highly pathogenic avian influenza A (H5N1) in dairy cattle in the United States, Version v1.0.4, Zenodo (2025); https://doi.org/10.5281/zenodo.15213504.

38. M. Worobey et al., "Preliminary report on genomic epidemiology of the 2024 H5N1 influenza A virus outbreak in U.S. cattle (Part 1 of 2)" (Virological.org, 2024); https://virological.org/t/preliminary-report-on-genomic-epidemiology-of-the-2024-h5n1-influenza-a-virus-outbreak-in-u-s-cattle-part-1-of-2/970/1.

39. C. Colijn, J. Gardy, Phylogenetic tree shapes resolve disease transmission patterns. Evol. Med. Public Health 2014, 96–108 (2014). doi: 10.1093/emph/eou018; pmid: 24916411

40. N. S. Lewis et al., Emergence and spread of novel H5N8, H5N5 and H5N1 clade 2.3.4.4 highly pathogenic avian influenza in 2020. Emerg. Microbes Infect. 10, 148–151 (2021). doi: 10.1080/22221751.2021.1872355; pmid: 33400615

41. A. Günther et al., Iceland as Stepping Stone for Spread of Highly Pathogenic Avian Influenza Virus between Europe and North America. Emerg. Infect. Dis. 28, 2383–2388 (2022). doi: 10.3201/eid2812.221086; pmid: 36261139

42. A. Markin, C. Macken, A. L. Baker, T. K. Anderson, Revealing reassortment in influenza A viruses with TreeSort. bioRxiv 2024.11.15.623781 [Preprint] (2024); https://doi.org/10.1101/2024.11.15.623781.

43. S. Boulo, H. Akarsu, R. W. Ruigrok, F. Baudin, Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. Virus Res. 124, 12–21 (2007). doi: 10.1016/j.virusres.2006.09.013; pmid: 17081640

44. V. G. Dugan et al., The evolutionary genetics and emergence of avian influenza viruses in wild birds. PLOS Pathog. 4, e1000076 (2008). doi: 10.1371/journal.ppat.1000076; pmid: 18516303

45. Y. Zhang et al., Influenza Research Database: An integrated bioinformatics resource for influenza virus research. Nucleic Acids Res. 45, D466–D474 (2017). doi: 10.1093/nar/gkw857; pmid: 27679478

46. A. Idoko-Akoh et al., Creating resistance to avian influenza infection through genome editing of the ANP32 gene family.

47. Nat. Commun. 14, 6136 (2023). doi: 10.1038/s41467-023-41476-3; pmid: 37816720

47. A. L. Baker et al., Dairy cows inoculated with highly pathogenic avian influenza virus H5N1. Nature 637, 913–920 (2024). doi: 10.1038/s41586-024-08166-6; pmid: 39406346

48. N. J. Halwe et al., H5N1 clade 2.3.4.4b dynamics in experimentally infected calves and cows. Nature 637, 903–912 (2024). doi: 10.1038/s41586-024-08063-y; pmid: 39321846

49. J. O. Lloyd-Smith et al., Epidemic dynamics at the human-animal interface. Science 326, 1362–1367 (2009). doi: 10.1126/science.1177345; pmid: 19965751

50. A. Markin et al., Reverse-zoonoses of 2009 H1N1 pandemic influenza A viruses and evolution in United States swine results in viruses with zoonotic potential. PLOS Pathog. 19, e1011476 (2023). doi: 10.1371/journal.ppat.1011476; pmid: 37488825

51. E. J. Abente et al., A highly pathogenic avian-derived influenza virus H5N1 with 2009 pandemic H1N1 internal genes demonstrates increased replication and transmission in pigs. J. Gen. Virol. 98, 18–30 (2017). doi: 10.1099/jgv.0.000678; pmid: 28206909

52. Centers for Disease Control and Prevention, Influenza Division, "Influenza Risk Assessment Tool (IRAT) Virus Report: Highly pathogenic avian influenza A(H5N1) virus; clade 2.3.4.4b Virus Strain: A/Texas/37/2024" (2024); https://www.cdc.gov/pandemic-flu/media/pdfs/2024/08/IRATATexas.pdf.

53. Y. Qi, H. B. Ni, X. Chen, S. Li, Seroprevalence of highly pathogenic avian influenza (H5N1) virus infection among humans in mainland China: A systematic review and meta-analysis. Transbound. Emerg. Dis. 67, 1861–1871 (2020). doi: 10.1111/tbed.13564; pmid: 32259345

54. E. S. Toner et al., Assessment of serosurveys for H5N1. Clin. Infect. Dis. 56, 1206–1212 (2013). doi: 10.1093/cid/cit047; pmid: 23386633

55. S. Nasreen et al., Seroprevalence of antibodies against highly pathogenic avian influenza A (H5N1) virus among poultry workers in Bangladesh, 2009. PLOS ONE 8, e73200 (2013). doi: 10.1371/journal.pone.0073200; pmid: 24039887

56. J. S. Robertson et al., The development of vaccine viruses against pandemic A(H1N1) influenza. Vaccine 29, 1836–1843 (2011). doi: 10.1016/j.vaccine.2010.12.044; pmid: 21199698

57. World Health Organization, "Genetic and antigenic characteristics of zoonotic influenza A viruses and development of candidate vaccine viruses for pandemic preparedness" (2023); https://cdn.who.int/media/docs/default-source/influenza/who-influenza-recommendations/vcm-northern-hemisphere-recommendation-2023-2024/20230224_zoonotic_recommendations.pdf.

58. B. Khuntirat et al., Absence of neutralizing antibodies against influenza A/H5N1 virus among children in Kamphaeng Phet, Thailand. J. Clin. Virol. 69, 78–80 (2015). doi: 10.1016/j.jcv.2015.05.025; pmid: 26209384

59. R. Yamaji et al., Pandemic risk characterisation of zoonotic influenza A viruses using the Tool for Influenza Pandemic Risk Assessment (TIPRA). Lancet Microbe 6, 100973 (2025). pmid: 39396528

60. S. Louis et al., Wastewater Surveillance for Influenza A Virus and H5 Subtype Concurrent with the Highly Pathogenic Avian Influenza A(H5N1) Virus Outbreak in Cattle and Poultry and Associated Human Cases - United States, May 12-July 13, 2024. MMWR Morb. Mortal. Wkly. Rep. 73, 804–809 (2024). doi: 10.15585/mmwr.mm7337a1; pmid: 39298357

61. USDA-APHIS, "Highly Pathogenic Avian Influenza Emergency Response" (2025); https://www.aphis.usda.gov/animal-emergencies/hpai.

62. USDA-APHIS, "National Milk Testing Strategy" (2025); https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/livestock/nmts.

63. USDA-APHIS, 'National Animal Disease Preparedness and Response Program" (2025); https://www.aphis.usda.gov/funding/nadprp.

64. A. Markin et al., PARNAS: Objectively Selecting the Most Representative Taxa on a Phylogeny. Syst. Biol. 72, 1052–1063 (2023). doi: 10.1093/sysbio/syad028; pmid: 37208300

65. J. T. Ladner, J. W. Sahl, Towards a post-pandemic future for global pathogen genome sequencing. PLOS Biol. 21, e3002225 (2023). doi: 10.1371/journal.pbio.3002225; pmid: 37527248

66. E. Spackman, Avian Influenza Virus Detection and Quantitation by Real-Time RT-PCR. Methods Mol. Biol. 2123, 137–148 (2020). doi: 10.1007/978-1-0716-0346-8_11; pmid: 32170686

67. S. S. Shepard et al., Viral deep sequencing needs an adaptive approach: IRMA, the iterative refinement meta-assembler. BMC Genomics 17, 708 (2016). doi: 10.1186/s12864-016-3030-6; pmid: 27595578

68. F. Mölder et al., Sustainable data analysis with Snakemake. F1000 Res. 10, 33 (2021). doi: 10.12688/f1000research.29032.2; pmid: 34035898

69. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313 (2014). doi: 10.1093/bioinformatics/btu033; pmid: 24451623

70. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780 (2013). doi: 10.1093/molbev/mst010; pmid: 23329690

71. P. Danecek et al., Twelve years of SAMtools and BCFtools. Gigascience 10, giab008 (2021). doi: 10.1093/gigascience/giab008; pmid: 33590861

72. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589 (2017). doi: 10.1038/nmeth.4285; pmid: 28481363

73. Z. W. Arendsee, A. L. Vincent Baker, T. K. Anderson, smot: A python package and CLI tool for contextual phylogenetic subsampling. J. Open Source Softw. 7, 4193 (2022). doi: 10.21105/joss.04193

74. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. PLOS ONE 5, e9490 (2010). doi: 10.1371/journal.pone.0009490; pmid: 20224823

75. A. J. Drummond, M. A. Suchard, D. Xie, A. Rambaut, Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973 (2012). doi: 10.1093/molbev/mss075; pmid: 22367748

76. D. L. Ayres et al., BEAGLE 3: Improved Performance, Scaling, and Usability for a High-Performance Computing Library for Statistical Phylogenetics. Syst. Biol. 68, 1052–1061 (2019). doi: 10.1093/sysbio/syz020; pmid: 31034053

77. S. Tavaré, Some probabilistic and statistical problems in the analysis of DNA sequences. Lect. Math Life Sci. 17, 57–86 (1986).

78. Z. Yang, Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J. Mol. Evol. 39, 306–314 (1994). doi: 10.1007/BF00160154; pmid: 7932792

79. A. J. Drummond, S. Y. Ho, M. J. Phillips, A. Rambaut, Relaxed phylogenetics and dating with confidence. PLOS Biol. 4, e88 (2006). doi: 10.1371/journal.pbio.0040088; pmid: 16683862

80. V. N. Minin, E. W. Bloomquist, M. A. Suchard, Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Mol. Biol. Evol. 25, 1459–1471 (2008). doi: 10.1093/molbev/msn090; pmid: 18408232

81. A. Rambaut, A. J. Drummond, D. Xie, G. Baele, M. A. Suchard, Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Syst. Biol. 67, 901–904 (2018). doi: 10.1093/sysbio/syy032; pmid: 29718447

82. P. Sagulenko, V. Puller, R. A. Neher, TreeTime: Maximum-likelihood phylodynamic analysis. Virus Evol. 4, vex042 (2018). doi: 10.1093/ve/vex042; pmid: 29340210

83. F. Bielejec et al., SpreaD3: Interactive visualization of spatiotemporal history and trait evolutionary processes. Mol. Biol. Evol. 33, 2167–2169 (2016). doi: 10.1093/molbev/msw082; pmid: 27189542

84. A. Alfonso-Morales et al., Evaluation of a Phylogenetic Marker Based on Genomic Segment B of Infectious Bursal Disease Virus: Facilitating a Feasible Incorporation of this Segment to the Molecular Epidemiology Studies for this Viral Agent. PLOS ONE 10, e0125853 (2015). doi: 10.1371/journal.pone.0125853; pmid: 25946336

85. Y. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data - from vision to reality. Euro Surveill. 22, 30494 (2017). doi: 10.2807/1560-7917.ES.2017.22.13.30494; pmid: 28382917

86. L. T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274 (2015). doi: 10.1093/molbev/msu300; pmid: 25371430

87. J. Soubrier et al., The influence of rate heterogeneity among sites on the time dependence of molecular rates. Mol. Biol. Evol. 29, 3345–3358 (2012). doi: 10.1093/molbev/mss140; pmid: 22617951

88. B. Q. Minh et al., IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. 37, 1530–1534 (2020). doi: 10.1093/molbev/msaa015; pmid: 32011700

89. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol. Biol. Evol. 35, 518–522 (2018). doi: 10.1093/molbev/msx281; pmid: 29077904

90. R Core Team, R: A language and environment for statistical computing (R Foundation for Statistical Computing, 2015); https://www.r-project.org/.

91. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890 (2018). doi: 10.1093/bioinformatics/bty560; pmid: 30423086

92. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009). doi: 10.1093/bioinformatics/btp324; pmid: 19451168

93. A. McKenna et al., The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303 (2010). doi: 10.1101/gr.107524.110; pmid: 20644199

94. G. A. Van der Auwera et al., From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. Curr. Protoc. Bioinformatics 43, 10.1, 33 (2013). doi: 10.1002/0471250953.bi1110s43; pmid: 25431634

95. A. Wilm et al., LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. 40, 11189–11201 (2012). doi: 10.1093/nar/gks918; pmid: 23066108

96. D. F. Burke, D. J. Smith, A recommended numbering scheme for influenza A HA subtypes. PLOS ONE 9, e112302 (2014). doi: 10.1371/journal.pone.0112302; pmid: 25391151

97. C. W. Nelson, L. H. Moncla, A. L. Hughes, SNPGenie: Estimating evolutionary parameters to detect natural selection using pooled next-generation sequencing data. Bioinformatics 31, 3709–3711 (2015). doi: 10.1093/bioinformatics/btv449; pmid: 26227143

98. X. Didelot, C. Fraser, J. Gardy, C. Colijn, Genomic Infectious Disease Epidemiology in Partially Sampled and Ongoing Outbreaks. Mol. Biol. Evol. 34, 997–1007 (2017). doi: 10.1093/molbev/msw275; pmid: 28100788

99. X. Didelot, M. Kendall, Y. Xu, P. J. White, N. McCarthy, Genomic Epidemiology Analysis of Infectious Disease Outbreaks Using TransPhylo. Curr. Protoc. 1, e60 (2021). doi: 10.1002/cpz1.60; pmid: 33617114

100. A. Bouma et al., Estimation of transmission parameters of H5N1 avian influenza virus in chickens. PLOS Pathog. 5, e1000281 (2009). doi: 10.1371/journal.ppat.1000281; pmid: 19180190

101. W. H. Kim, S. Cho, Estimation of the Basic Reproduction Numbers of the Subtypes H5N1, H5N8, and H5N6 During the Highly Pathogenic Avian Influenza Epidemic Spread Between Farms. Front. Vet. Sci. 8, 597630 (2021). doi: 10.3389/fvets.2021.597630; pmid: 34250054

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adq0900
Materials and Methods
Figs. S1 to S10
Tables S1 to S5
Data S1 to S7
MDAR Reproducibility Checklist

Submitted 1 May 2024; accepted 14 February 2025
10.1126/science.adq0900