Analyses of phylogenetics, natural selection, and protein structure of clade 2.3.4.4b H5N1 Influenza A reveal that recent viral lineages have evolved promiscuity in host range and improved replication in mammals in North America

Daniel Janies[1,2,*], Kary Ocaña[3], Sayal Guirales-Medrano[1,2], Khaled Obeid[1,2], Rachel Alexander[1,2,4], and Colby T. Ford[1,2,4,5]

Affiliations:

1. Center for Computational Intelligence to Predict Health and Environmental Risks (CIPHER), University of North Carolina at Charlotte, Charlotte, NC, USA
2. Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, NC, USA
3. National Laboratory of Scientific Computing, Rio de Janeiro, Brazil
4. School of Data Science, University of North Carolina at Charlotte, Charlotte, NC, USA
5. Tuple LLC, Charlotte, NC, USA

* Corresponding Author: djanies@charlotte.edu

Abstract

H5N1 influenza has been circulating in birds from Eurasia and Africa for more than 146 years, but human infection has been sporadic. H5N1 (clade 2.3.4.4b) has recently infected hundreds of species of wild and domestic birds and mammals in North America. Furthermore, as of February 26, 2025, H5N1 has infected 70 humans in the United States, and one infection proved lethal. Furthermore, in attempts to control H5N1 in the United States, 10s of millions of egg-laying chickens have died or been culled. These efforts have led to very high egg prices in the United States. We have developed an analytical bioinformatics and genomics workflow to understand better how H5N1 is circulating in North America and adapting to new host species. Our workflow consists of: 1) Phylogenetic analyses of large viral sequence datasets to identify subclades of viral lineages causing the current outbreaks in humans and farm animals and closely related viral background lineages. 2) Next, we transfer sequence data from subclades of interest with farm animal and human infection, background data, and vaccine candidate data to analyses of natural selection. 3) Once we identify mutations of interest that underlie recent viral adaptation to animal and human infection, we perform computational structural analyses of binding to host proteins for cell receptors and immune processes. Here, we show that H5N1 (clade 2.3.4.4b) is spreading in North America as two distinct subclades of interest for human and animal health. These viral lineages have achieved a vast host range by efficiently binding the viral surface protein Hemagglutinin (HA) to both mammalian and avian receptors. This novel promiscuity of host range is concomitant with the additional strengthening of the polymerase basic 2 (PB2) viral protein's binding for mammalian and avian immune proteins. Once bound, the immune proteins are disabled, thus allowing for more efficient replication of H5N1 in mammalian and avian cells than seen in the recent past. In conclusion, H5N1 (clade 2.3.4.4b) is causing an animal pandemic through promiscuity of host rage and strengthening ability to evade the innate immune systems of both mammalian and avian cells.

Introduction

Avian influenza (H5N1) lineages have circulated in Eurasia since at least 1878 (Lupiani & Reddy, 2009). In January 2022, H5N1 Influenza A virus (clade 2.3.4.4b) infection was detected in wild birds in the United States [Centers for Disease Control (CDC) 2024]. Since then, this viral lineage has spread to all 50 of the United States and the District of Columbia. This viral lineage has infected over 246 species, including many mammals [World Organization for Animal Health (WOAH) 2025], wild birds (CDC, 2025), cows, cats, and poultry in agricultural settings, and humans coming into contact with wild and farm animals (supplemental_table_1.csv).

As of February 26, 2025, there are 70 confirmed human cases of H5N1 in the United States, including one death in Louisiana (CDC, 2025). Furthermore, tens of millions of egg-laying chickens have died or been culled in the United States due to H5N1, leading to soaring egg prices (Tin, 2025).

In Canada, H5N1 has spread to 132 species of birds and mammals [supplemental table 2; Canadian Food Inspection Agency (CFIA), 2025; Environment and Climate Change Canada, (2024); CFIA, 2025] (supplemental_table_2.csv). For Mexico, the few species reported by The Servicio Nacional de Sanidad Inocuidad y Calidad Agroalimentaria (SENASICA, 2023) are mostly agricultural Galliformes and Anseriformes, with some wild Anseriformes and Pelecaniformes. All avian species in SENASICA reports are redundant with species in reports from the USA and Canada (SENASICA, 2023; WOAH, 2025). The Pan American Health Organization reports no mammal cases in Mexico as of March 4, 2025 (PAHO, 2025).

Here, we have developed a workflow to study the phylogenetics of the host shifts, natural selection, and structural biology of recent lineages of the H5N1 subtype of influenza A virus in North America. We used the workflow to isolate clades of interest for human infection and study those clades for adaptation to mammals. We assess adaptation using computational analyses of natural selection of the nucleotide diversity of HA and PB2 genes. Once mutations under natural selection are identified, we evaluate the nature of the adaptation of H5N1 (clade 2.3.4.4b) viral lineages through computational structural biology. Here, we show that hemagglutinin (HA) mutations have shifted the virus from adaptation to avian cells to promiscuity to many species of wild and domestic mammals and birds. Our analyses of natural selection in Polymerase Basic 2 (PB2) show that mutations have adapted the H5N1 (clade 2.3.4.4b) viral lineage to increase their ability to bind proteins important to the viral invasion and immune response of the avian and mammalian host cells, thus dampening the hosts' innate defenses to infection and setting the stage for increased viral replication.

Methods

We updated a large dataset of 18,289 Hemagglutinin (HA) sequences of H5N1 influenza that was built for Ford et al., (2025) with data from United States National Institutes of Health's NIH's (Genbank; ncbi.nih.nlm.gov) and The Global Initiative to Share all Avian Influenza Data (GISAID; gisaid.org) that was released in the second half of 2024 (e.g., Burrough et al., 2024; Caserta et al., 2024; Pulit-Penaloza et al., 2024).

Multiple sequence alignment

We selected a subset of 9,359 sequences from the aforementioned set of HA sequence data that contained clade 2.3.4.4b and background data (supplemental_table_3.csv) and realigned using MAFFT (Katoh and Standley, 2013). Sequential site numbering was used to count sites after realignment.

For the HA data, we trimmed ragged edges, padded the few remaining leading and trailing gaps with "?", and eliminated sequences that broke the reading frame (supplemental_data_file_4.nexus).

We aligned a dataset of Polymerase basic 2 (PB2) nucleotide sequences of H5N1 influenza containing clade 2.3.4.4b and background data collected from GISAID and Genbank. For the PB2 data, we trimmed ragged edges, padded the few remaining leading and trailing gaps with "?", and eliminated sequences that broke the reading frame (supplemental_data_file_5.nexus of 3415 sequences aligned to 2281 positions). Sequential site numbering was used to count sites after realignment.

Tree searches on full datasets

For tree searches on large datasets mentioned above, we used TNT (Version 1.6; Goloboff and Morales, 2023) with equal weighting for edit costs. For HA, we used GISAID sequence EPI100512 as an outgroup. For PB2, we used the same methods and GISAID sequence GU052525 as an outgroup.

Character evolution

For the HA dataset, we coded 2 independent characters (one for the host and one for the date of isolation) based on metadata associated with the 9359 sequence dataset (supplemental_table_3.csv) and mapped the host and date characters into the tree using Mesquite (Maddison and Maddison, 2023). We examined the host and date character optimization for clades of interest. We focused on two clades of interest for human isolates.

In one subclade dataset of interest, we focused on the influenza HA isolates surrounding the human death in Louisiana (supplemental_data_file_6.phy of 331 Hemagglutinin nucleotide sequences for H5N1 aligned to 1701 positions due to trimming of the stop codon, which was invariant, in preparation for analyses of selection). Trees for this subclade dataset were calculated using RAxML (version 8.2.12; Stamatakis et al., 2014) with Genbank sequence LC718258 as an outgroup.

Another subclade dataset of interest, we focused on the human-bovine-avian transmission events (supplemental_data_file_7.phy of 767 sequences and 1,701 aligned positions due to trimming of the stop codon, which was invariant, in preparation for analyses of selection). Trees for this subclade dataset were calculated using RAxML (Stamatakis et al., 2014) with GISAID sequence EPI3509783 as an outgroup.

For each HA subclade of interest, we calculate StrainHub networks under the criterion of betweenness centrality and colored events of interest (de Bernardi Schneider et al., 2020; Ford et al., 2025) (Figures 1 and 2).

For PB2, we identified a subclade subtended by CY205506 containing the 2022-2024 zoonotic strains of interest corresponding to those strains we discovered in the HA analyses (described above). We removed sequences with stop codons and the trailing stop codons, which were invariant to ready the data for analyses of natural selection. The resulting data contained 1528 PB2 nucleotide sequences and 2277 aligned positions (supplemental_data_file_8.phy).

Analyses of natural selection

All subclade datasets were cleaned of sequences containing stop codons using MACSE (V2; Ranwez et al., 2018) and manual inspection. We added candidate vaccine sequences as baselines (e.g., GISAID sequence EPI1846961 for HA subclade data; and Genbank sequence OQ958041 for PB2 subclade data). Trees were calculated using alignments, as described above, using RAxML (Stamatakis et al., 2014).

To assess natural selection, we used HyPhy (version 2.5.67; Kosakovsky Pond et al., 2020) and the Nielsen and Yang (NY) approach (Nielsen & Yang, 1998) for detecting positive selection. The NY approach estimates the dN/dS ratio at each codon by sampling global dN/dS distributions and calculating Bayesian posterior probabilities to identify sites with dN/dS > 1. We tested several models based on the NY framework, including M0 (Single Rate Model), M1 (Neutral Model), M2 (Selection Model), M3 (Discrete Model), M7 (Beta Model), M8 (Beta & ω Model), and M12 (Two-Class Model). This range of models allowed us to evaluate both purifying and episodic positive selection, accounting for varying selection pressures across sites and ensuring a robust, comprehensive analysis of evolutionary dynamics across the data.

Structural Biology

Hemagglutinin

Our primary objective is to evaluate the binding affinity of HA from different strains to host receptor proteins, providing insights into the expanding host range of H5N1 Influenza A virus (clade 2.3.4.4b).

Four HA sequences of interest were selected based on phylogenetic analyses (i.e. GISAID: EPI3171488, EPI1846961 (a vaccine candidate); Genbank: PQ591824, PQ809550). The amino acid sequences of HA from each strain were manually extracted

and subsequently folded into three-dimensional structures using ColabFold (v1.5.2; Mirdita et al., 2022). To model interactions with host sialic acid (SA) receptors, human and avian SA glycans (chain I) were obtained from existing Protein Data Bank (PDB) structures 4K63 and 4KDO, respectively.

Protein-glycan docking simulations were performed for each HA variant against both human and avian SA glycans. Molecular docking was conducted using HADDOCK3 (v2024.12.0b7), following established protein-glycan docking protocols[1] (Teixeira et al., 2024). Binding interactions were evaluated to assess the potential differences in HA-SA binding affinities among strains.

Polymerase Basic 2

Our primary objective is to evaluate the binding affinity of Polymerase Basic 2 (PB2) from different strains to host immune factors, providing insights into potential changes in immune evasion and replication mechanisms of H5N1 Influenza A virus (clade 2.3.4.4b).

Similar structural and docking analyses were performed for PB2 using sequences from viral strains of interest [Genbank OQ958041 (a vaccine candidate), PQ809562, PQ591825, and PP577947]. We also considered proteins carrying mutations at PB2 protein position 588, including GISAID: EPI3304108 (A588T mutation), EPI3315580 (A588S mutation), and EPI3447228 (A588V mutation)

Known host immune response proteins that interact with viral PB2 protein [e.g., Human mitochondrial antiviral-signaling protein (MAVS; PDB: 2MS8), Importin-α3 (human; PDB: 4UAE) and (avian; Uniprot: A0A2I0M1B5) orthologs, and Human acidic leucine-rich nuclear phosphoprotein 32 family member A (ANP32A; PDB: 6XZQ)] were selected as target proteins for docking studies.

The amino acid sequences of PB2 and its corresponding host protein targets were folded into three-dimensional structures using ColabFold (Mirdita et al., 2022). Protein-protein docking simulations were conducted using HADDOCK3, following the standard protocol for protein-protein interactions[2]. We report parameters and motifs used in the repository: https://github.com/colbyford/Influenza_A_2.3.4.4b_Analyses.

[1] HADDOCK3 Protein-Glycan Configuration Template: https://github.com/haddocking/haddock3/blob/main/examples/docking-protein-glycan/docking-protein-glycan-full.cfg

[2] HADDOCK3 Protein-Protein Configuration Template: https://github.com/haddocking/haddock3/blob/main/examples/docking-protein-protein/docking-protein-protein-full.cfg

Results

HA phylogeny

The results for the curated HA data were an alignment of 9359 sequences and 1704 positions. In two of three independent tree searches in TNT (Goloboff and Morales, 2023) three heuristically parsimonious trees were found at 16950 steps (supplemental_data_file_9.log).

One subclade contains a human isolate that led to death (LDH, 2025; Genbank accession PQ809550) (see the arrow "<----" at line 1516 of supplemental_data_file_10.log for a key point in the tree). Sequences in this subclade are referred to as genotype D1.1.

Another subclade, distant from the subclade containing the isolate from the human death in Louisiana, contains multiple human-bovine-avian transmission events (e.g., see the arrow "<----" at line 8659 of supplemental_data_file_10.log for a key point in the tree). This subclade contains HA sequences from influenza isolates referred to as genotype B3.13.

HA host dynamics

In the StrainHub, two lines (from Anseriformes and another from Galliformes) are incoming to primates. From primates, there are transmissions to Anseriformes, Galliformes, and Accipitriformes (Figure 1).

In the subclade D1.1, the viral isolate (Genbank: PQ809550) from the deceased human patient in Louisiana has host ancestry of Galliformes and is a sister taxon of an isolate from Accipitriformes (Figure 1).

The subclade B3.13, includes a mixture of isolates from many host taxa (i.e. Carnivora, Galliformes, Anseriformes, Gaviiformes, Artiodactyla, Aves, Falconiformes, Passeriformes, Primates, Accipitriformes, and Rodentia). In our StrainHub, we see high-frequency unidirectional transmission from Artiodactlya to Primates (Figure 2).

We also see bidirectional transmission between Galliformes and Primates and a moderate frequency. We see unidirectional transmission from Carnivores to Primates with moderate frequency (Figure 2).

PB2 phylogeny

For the curated dataset of PB2 contained 3415 sequences and 2281 aligned positions (supplemental_data_file_5.nexus), 3 independent TNT searches found 432-465 heuristically parsimonious trees of 12133 steps (supplemental_data_file_11.log).

We chose the first tree to select a subclade of interest based on the position of isolates of interest (marked with "<---" in supplemental_data_file_12.log).

We identified the subclade subtended by the outgroup CY205506 and removed sequences with stop codons to ready the data for analyses of natural selection. The resulting data contained 1528 taxa and 2277 aligned positions (supplemental_data_file_8.phy)

Analyses of natural selection

We tested multiple models to evaluate selection pressures (the full results are in the files supplemental_data_file_13.txt, supplemental_data_file_14.txt, and supplemental_data_file_15.txt)

Below we report values for HyPhy (model 8; Beta & ω; Kosakovsky Pond et al., 2020) which is effective for detecting sites under positive selection.

HA results

Table 1 lists the results for the HA subclade of interest for avian-human exchanges, including the lineage that led to the death of the human patient in Louisiana (Figure 1) and supplemental_data_file_6.phy plus EPI1846961, A/Astrakhan/3212/2020 as a baseline):

Site	Probability	Mutation(s)
140	0.960974714844818	N140D/H
230	0.9685961144425779	A230V
298	0.9697143147489277	V298I/L
490	0.9569366098585621	D490N
509	0.9682055136291944	E509G

Table 1: HA amino acid sites with dN/dS>1 (Posterior cutoff = 0.95) along with mutations of interest for positive selection using sequential site numbering.

Details on the function of these mutations, if known, are provided in the discussion.

Table 2 lists the results for the HA subclade of interest for avian-bovine-human exchanges (Figure 2 and supplemental_data_file_7.phy plus EPI1846961, A/Astrakhan/3212/2020 as a baseline).

Site	Probability	Mutation(s)
11	0.9733589348171219	V11I
147	0.9988807760818452	V147M
172	0.9602969606372659	A172T
250	0.9680511447623218	K250E/N/R
526	0.9894637570633174	I526V

Table 2: HA amino acid sites with dN/dS>1 under the NY model (model 8; Posterior cutoff = 0.95) along with mutations using sequential site numbering.

Details on the function of these mutations, if known, are provided in the discussion.

PB2 results

These are the results for the PB2 subclade of interest (refer to supplemental dataset supplemental_data_file_8.phy plus OQ958041, A/American Wigeon/South Carolina/22-000345-001/2021 as a baseline).

Site	Probability	Mutation(s)
184	0.9670923982786206	T184A/M/K
249	0.9973568275209534	E249G/D/K/V
292	0.9922498777048648	I292L/V/T/M
344	0.9999928310390789	V344M/A/I
463	0.9789001325108995	I463M/V
588	0.9428168996255263*	A588T/S/V
627	0.9999895456621494	E627K/V
676	0.9666441948315898	T676A/I/L/V

Table 3: PB2 amino acid sites with dN/dS>1 under the NY model (model 8; Posterior cutoff = 0.95) along with mutations. *The probability of site 588 is very close to the cutoff and thus was considered in structural analyses.

Details on the function of these mutations, if known, are provided in the discussion.

Analyses of viral-host protein docking

HADDOCK3 provides multiple scoring metrics to assess binding interactions, with Van der Waals energy as this study's primary measure of binding affinity.

Hemagglutinin Structures

Analysis of the HA-SA docking results indicated no significant differences in binding affinity between HA and the human or avian SA glycan variants. Across the selected viral strains (EPI1846961, A/Astrakhan/3212/2020; PP577947, A/Texas/37/2024; PQ591825, A/California/150/2024; and PQ809562, A/Louisiana/12/2024), HA exhibited comparable binding interactions with both human and avian SA, as shown in Figure 3. The Van der Waals energy values, used as a measure of binding strength, showed overlapping distributions, further supporting the observation that these HA variants do not demonstrate higher binding for SA glycan from either avian or human hosts.

Additionally, in Figure 4, we see the spatial orientation and positioning of SA molecules within the docking simulations remained consistent, reinforcing the conclusion that H5N1 clade 2.3.4.4b HA maintains the ability to bind both human and avian SA without structural bias. This suggests that the selected HA variants retain the capacity for a wide host range and adaptation.

Polymerase Basic 2 Structures

In contrast to results for HA binding to SA glycans, docking simulations of polymerase PB2 viral protein motifs with host immune response proteins revealed an increased binding affinity in all recent PB2 strains compared to the vaccine candidate (OQ958041, A/American Wigeon/South Carolina/22-000345-001/2021) as seen in Figures 5-7. These findings indicate an enhanced ability of recent PB2 variants to inhibit host innate immune signaling pathways, contributing to greater immune evasion (Yang et al., 2022).

The results indicate that PB2 exhibits differential binding affinities to its target host proteins. Notably, some recent PB2 strains show stronger binding interactions (lower Van der Waals energy) than the vaccine candidate (OQ958041), for Importin alpha 3 and MAVS. We also see improved binding to human ANP32A as compared to avian ANP32A, which indicates improved replication within human cells (Carrique et al., 2020). This trend indicates strengthening immune evasion capabilities among recent PB2 variants in both human and avian hosts. In summary, binding experiments with viral and host proteins across both North American H5N1 clades of interest show a clear increase in affinity for mitochondrial antiviral signaling proteins in human and mammalian infections, along with enhanced viral replication and nuclear localization in both human and avian hosts.

Binding of viral proteins carrying mutations at PB2 protein site 588, including GISAID: EPI3304108 (A588T mutation), EPI3315580 (A588S mutation) EPI3447228 (A588V mutation)

Improved binding for ANP32A and MAVS remains the same in A588T/S/V variants (i.e. EPI3304108, EPI3315580, and EPI3447228) as isolates representing clades of interest

mentioned above (OQ958041, PQ809562, PQ591825, and PP577947). The 588 mutation occurs in the '627 domain' which we test by binding to Importin alpha 3 (IMA3) (Nilsson et al., 2017). Most isolates with variation at site 588 (A588 have an improved binding affinity in humans than in avian host proteins except for the A588S mutation as seen in Figures 5 and 7. A588S caused a decrease in affinity to human proteins. For A588S, we also see that binding affinity is worse than the vaccine candidate in the human IMA3, but the difference is minimal.

In addition to A588T/S/V, we checked the PB2-F6 constellation of mutations that are pointed out by (Li et al., 2022) to work in synergy with A588T to increase the pathogenicity and transmissibility in chickens and the virulence of mice of H10N8 in the single mutation of PB2-A588V (i.e. PB2: I292V, R389K, T598M, L648V, T676M). We only report wild-type genotypes for this constellation in our H5N1 isolates bearing A588T/S/V at this juncture. Moreover, these A588T/S/V bearing isolates are all currently at wild-type E627. However, many of the H5N1 viruses circulating in North America in 2022-24 carry 627K, including viruses hosted in animals (wild and domestic birds and mammals) and humans and the site is under positive selection in our data.

Residue 676 exhibits mutations T676I, T676L, T676V, T676A. Mutation T676I underlies the expansion of H9N2, H7N7, H7N9 viruses in birds in the 2000's and mammals and subsequently mutated to I676M (Neumann et al., 2014). Caserta et al., (2024) report the mutation T676A in genotype B3.13 H5N1 viruses along with other PB2 mutations (T58A, E362G, D441N, M631L). Despite these reports, the function of variants at site 676 has not been studied.

Discussion

There are several main points of discussion for analyses of HA in this paper: 1) the structure of the clades of interest in North America and the natural selection they exhibit. 2) the changes of binding of HA in terms of host range and by PB2 in host immune suppression.

Structure of clades

The two main clades from the spread of H5N1 in North America in 2022-24 of interest in HA, are 1) a clade surrounding the human death in Louisiana (genotype D1.1 viruses) and 2) clade with human-bovine-avian transmission events (genotype B3.13). Notably, these clades of interest are phylogenetically very far apart in the HA phylogeny.

Moreover, these two HA clades have slightly different inter-host dynamics, as demonstrated by the character evolution and StrainHub analyses. The clade surrounding the human death in Louisiana has largely avian-primate exchanges. However, recent reports have indicated artiodactyl cases (USDA-APHIS, 2025; artiodactyla is represented by Genbank PQ687463 in this subclade).

The clade in which human-bovine-avian interactions are emblematic show high frequency of unidirectional transmission from Artiodactlya to Primates, bidirectional transmission between Galliformes and Primates, and unidirectional transmission from Carnivora to Primates.

Natural selection in HA

The mere fact that two phylogenetically distinct clades are currently infecting North American mammals demonstrates the non-canonical nature of zoonotic evolution of HA. Moreover, these two clades have different selective regimes in the HA. The same motif is under selection but not the same point mutations. The HA has two motifs of interest, the receptor binding domain (RBD; residues 113-265) and the stalk (residues 335-500). While both present antigenic properties, we focused on the RBD and its role for host selection via sialic acid conformation (Ghafoori et al., 2023). The structural evolution of HA is towards viral host promiscuity as the results both show no binding strength differences for this HA motif for avian or mammalian receptors. This result is consistent with the hundreds of wild and domestic avian and mammalian hosts in North America (supplemental_table_1.csv and supplemental_table_2.csv).

For the HA subclade of interest for avian to human exchanges, including the lineage that led to the human death in Louisiana (Figure 1) we see the following sites of functional interest.

Site 230 in the HA protein exhibits mutation A230V in our data. In the literature, this site is described in older H5N1 datasets as M230V. Site 230 is involved in immune escape and stabilization of the HA in experimental ferret infection (Hanson et al., 2015).

For the HA subclade of interest for avian-bovine-human exchanges (Figure 1) we see the following sites of functional interest. Site 172 in the HA protein shows mutation A172T in our data. In the literature, A172T has been favored in vaccine candidate strains as the mutation confers increased replication properties in mammalian cell culture [e.g., Vero cells; (Lee et al., 2017)].

Taken together, the distinct clades, each with wide but slightly different host dynamics and measurements of natural selection in HA, indicate that H5N1 (clade 2.3.4.4b) viral lineages have several non-canonical evolutionary and host pathways to become zoonotic emergent diseases of concern for animal and human health and economic well-being due to large agricultural losses (Tin, 2025). In each of the 100s of hosts now available to the virus, new mutational constellations and reassortment events can evolve leading to a more concerning virus for human and animal health.

Comparison of mutations within the HA RBD to active binding residues reveals that while the mutations are not directly involved in receptor binding, several (140, 147, and 230) lie within 10 amino acids of key binding sites. Their close proximity can introduce steric effects that influence binding affinity, altering host receptor specificity, and increasing host promiscuity.

Dadonaite et al., (2024) points out sets of mutations of interest they discovered in H5N1 HA in combined and computational and pseudo-particle experiments for receptor binding, HA stability, and immune escape mutations. We compared our mutations of interest from tables 1 and 2 with their mutations of interest with adjustments for numbering conventions. We found no exact overlap in sites but some overlap in domains.

Natural selection in PB2

For PB2, we identified from our large dataset a single subclade of interest for zoonotic events. In our analyses of natural selection in this PB2 subclade we find the following sites of functional interest.

Site 249 in the PB2 protein exhibits many mutations in our dataset, including E249K. E249G, E249D, and E249V. Site 249 in the PB2 protein exhibits the E249K mutation in our data. H5N1 isolated from humans in Egypt in 2024 carried E249K (Kayali et al., 2016). E249K was also carried in Human (H1N1) pdm09 cases in India in 2015 (Mukherjee et al., 2016).

A historical study of H5N1 viruses demonstrates that the E249G mutation improves replication in human cells (Yamaji et al., 2015). Most recently (2024) in H5N1 in dairy cows in Kansas the E249G mutation originated in mammals about this time in concert with mutations in NS1 (R21Q) and PB2 (E627K). These mutations form a constellation associated with viral adaptation to mammalian hosts.

Similarly, E249D is among a constellation of mutations that led to increased pathogenicity of Eurasian avian-like swine lineage H1N1 (Cai et al., 2020).

In 2024, a human traveler from India to Australia was infected with a recombinant (clade 2.3.2.1a and 2.3.4.4b) virus with a clade 2.3.4.4b PB2 containing mutation V344M. V344M has been detected in wild birds and poultry in Asia since 2020 (Deng et al., 2025).

Site 463 exhibits mutations I463M and I463V in our data. Mutation I463M in PB2 occurs in low frequency in viruses extracted from pigs experimentally infected with clade 2.3.4.4b H5N1 viruses derived from mink (Kwon et al., 2024).

Results for position 588 are mixed. 1) H5N1 2.3.4.4b viruses carrying variation at 588 have only adapted a fraction of the genotypes for the constellation enabling viral promiscuity in H10N8 discovered by Li et al., 2022. 2) Position 588 is under positive selection that is slightly statistically insignificant. 3) As site 588 has diverse results for binding ANP32A based on which genotype it carries (A588S caused a decrease in affinity to humans proteins studied whereas A588T/V increased binding affinity to human proteins studied). At this juncture, position 588 is important to watch for 588 evolving into constellations already described in future lineages and or 588 evolving its own genomic constellation that may have not been described. The structural data

indicates that some genotypes (e.g. A588T/V) are achieving host promiscuity in the current H5N1 2.3.4.4b genomic backbone (Figures 5-7).

Site 627 exhibits mutations E627K and E627V in our data. E627K is a well-studied mutation underlying the transmission of avian-hosted lineages H5N1 to mammals (Subbarao et al., 1993). E627K also increases the replication and virulence of the H5N1 in mammals (Maines et al., 2005). E627V allows many viruses of avian origin (H9N2, H7N9, and H3N8, which may share PB2 genes with H5N1 via reassortment) to infect and replicate in both chickens and mice by binding ANP32A proteins in both species (Guo et al., 2024). Experimental infection of mice with H5N1 virus bearing E627K leads to increased replication and mortality in mice (Shinya et al., 2004). One human isolate in our data that bears E627K [Genbank: PP577947 A/Texas/37/2024(H5N1)] has been shown in experiments to transmit via respiratory droplets among ferrets, leading to mortality in most of the ferrets (Gu et al., 2024).

Taken together, these mutations (well studied or understudied) point to the adaptation of H5N1 to increased transmission, virulence, and replication in a wide range of avian and mammalian hosts. Thus, our results on PB2 also support the hypothesis for the evolution of host promiscuity and increased viral replication to both avian and mammalian hosts in H5N1 clade 2.3.4.4b in North America in 2022-present.

Misra et al., 2024 (cite) note an array of 12 sites in PB2 that may be of interest. We compared their list to ours and found that only sites 627 (which is well known) and 676 are in common.

Structural biology

The adaptation of influenza A viruses to human hosts is driven by key molecular interactions, particularly in hemagglutinin (HA) receptor binding and polymerase basic 2 (PB2) host factor interactions. HA mediates viral entry by binding to host sialic acid (SA) receptors. PB2 plays a critical role in viral RNA synthesis and replication and often adapts to host-specific nuclear import pathways. Our docking simulations provide insight into how these viral proteins interact with host proteins and the implications for promiscuous transmission and increased replication in a variety of hosts.

Residues that have shown mutations under selection were used to define areas of interest of docking simulations. This resulted in the selection of the HA receptor binding domain and several domains within PB2 such as the amino terminal and the 627 domain (Nilsson et al., 2017). While mutations of interest were not directly involved as active binding residues in the docking simulations, they were present near the active binding residues. These mutations cause conformation changes, changes in charges and side chain sizes, and or effect residue polarity (Gupta et al., 2022).

HA binding to SA remains a central determinant of host specificity. Previous studies have established that avian-adapted influenza strains preferentially bind α2,3-linked SA, whereas human-adapted strains favor α2,6-linked SA (Xiong et al., 2014). Our docking results indicate that HA from the selected H5N1 strains exhibits similar binding affinities

for both avian and human SA. This promiscuity indicates that these strains retain the potential for widening zoonotic transmission and subsequent increases in mutation and reassortment avenues. The structural consistency in HA-SA interactions across different strains indicates that receptor binding alone may not be the primary barrier to human adaptation in these viruses. Instead, additional factors such as polymerase function and host immune evasion likely contribute to host adaptation.

PB2, a key component of the viral RNA polymerase complex, has been shown to influence host adaptation by interacting with cellular factors involved in nuclear trafficking and immune signaling (Ma et al., 2017, Wandzik et al., 2021). Our docking analysis highlights strong PB2 interactions with MAVS and Importin-α3, particularly in recent strains, which bind much stronger to these proteins than the vaccine candidate (OQ958041) does. MAVS is a critical component of the hosts' antiviral response, and increased PB2 binding affinity may facilitate immune evasion by interfering with interferon signaling (Long & Fodor, 2016; Graef et al., 2010). Similarly, Importin-α3 and ANP32A mediates nuclear import of viral ribonucleoproteins. Enhanced PB2-Importin-α3 interactions improve viral replication efficiency in mammalian cells (Yu et al., 2022).

Together, our results indicate that while HA maintains broad receptor-binding capability, PB2 is undergoing adaptive changes that enhance immune evasion and replication in avian and mammalian hosts. This underscores the importance of monitoring both HA and PB2 in emerging influenza strains to assess their potential for human adaptation. Future research integrating structural modeling with functional assays will be crucial for validating these computational predictions and improving our understanding of influenza A H5N1 host adaptation mechanisms.

Limitations and opportunities

All of the work described in this study is computational but sets that stage for empirical confirmation. Despite this limitation, in our experience with SARS-CoV-2, computation is crucial for early insights that can warn of upcoming major events such as seen in our prediction of the increase in cases associated with early ascertainment of the immune invasion of the Omicron variant (Ford et al., 2022).

Sequence data takes time to produce and publish and thus can be behind the activity in nature. Moreover, there is much of the H5N1 genome yet to be analyzed. However, sequence data combined with advanced analytics are the best lenses we have for rapid assessment of the risk of emergent infectious diseases before a crisis worsens. As we have codified a workflow from phylogenetics to structural biology, we can redeploy the steps to new sequence datasets. In doing so our goals are to advance molecular disease surveillance to functional assessment of viral adaptation and risk.

Supplementary Materials

Code, raw sequence data, alignments and results from phylogenetics and analyses of selection, predicted protein structures, and protein-protein binding data are available at https://github.com/colbyford/Influenza_A_2.3.4.4b_Analyses.

Figures

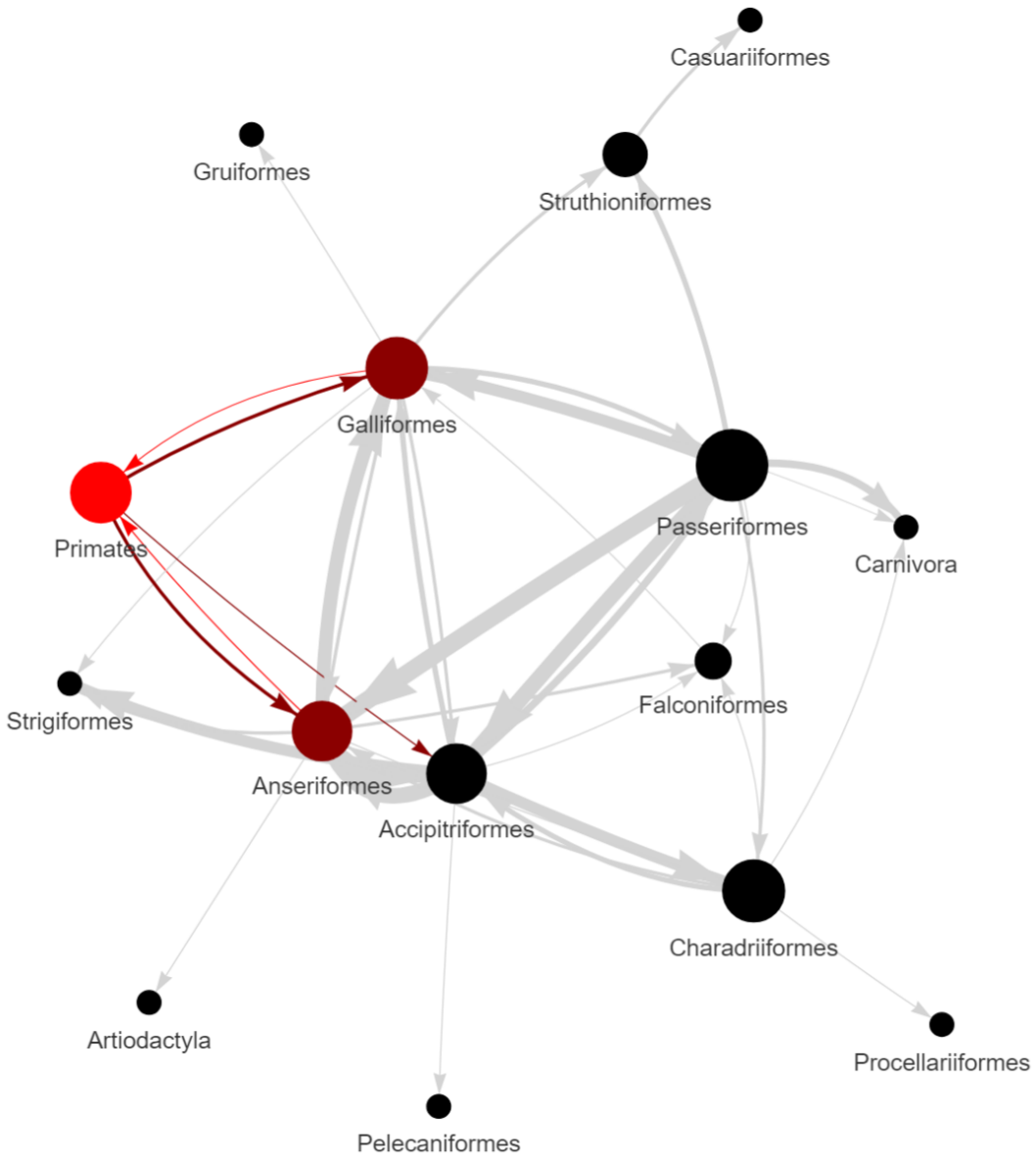

Figure 1: A Strainhub network using a tree based on HA sequence data and host metadata for the clade of interest surrounding the deceased patient infected with H5N1 in Louisiana in 2024 (PQ809550). The metric used to create the network was betweenness centrality. The arrows represent the directionality of viral transmission events recovered. The thickness of the lines represents higher frequency of transmission. The colors represent highlighted zoonotic events such as bidirectional transmission between Primates (humans), Galliformes (e.g., poultry), and Anseriformes (e.g., ducks and geese), and unidirectional transmission from Primates to Accipitriformes (e.g., hawks and eagles).

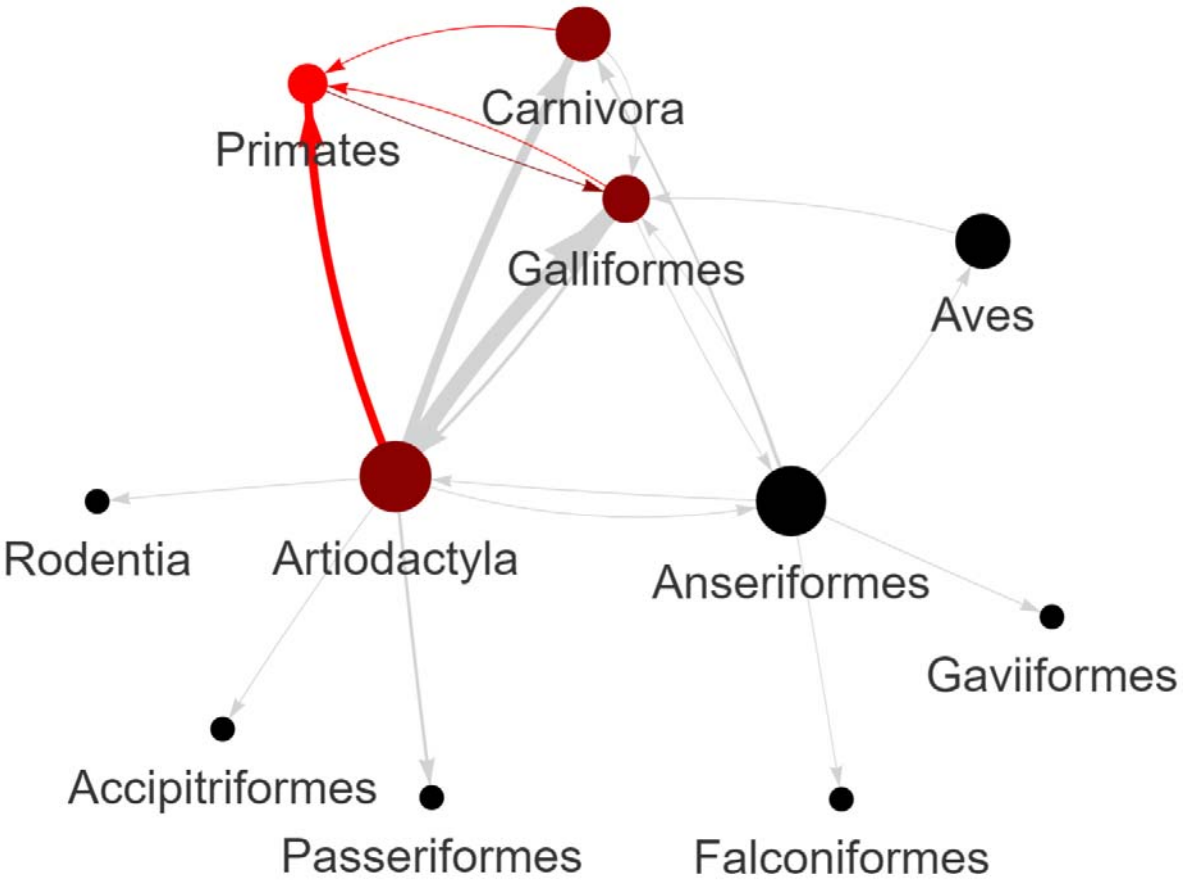

Figure 2: A Strainhub network using a tree based on HA sequence data and host metadata for the clade of interest surrounding the human-bovine-avian transmission events in 2024. The metric used to create the network was betweenness centrality. The arrows represent the directionality of viral transmission events recovered. The thickness of the lines represent higher frequency of transmission. The colors represent highlighted zoonotic events such as unidirectional transmission from Artiodactlya (e.g. cows) to Primates (humans), bidirectional transmission between Primates and Galliformes (e.g., poultry), and unidirectional transmission from Carnivores (e.g. domestic cats) and Primates.

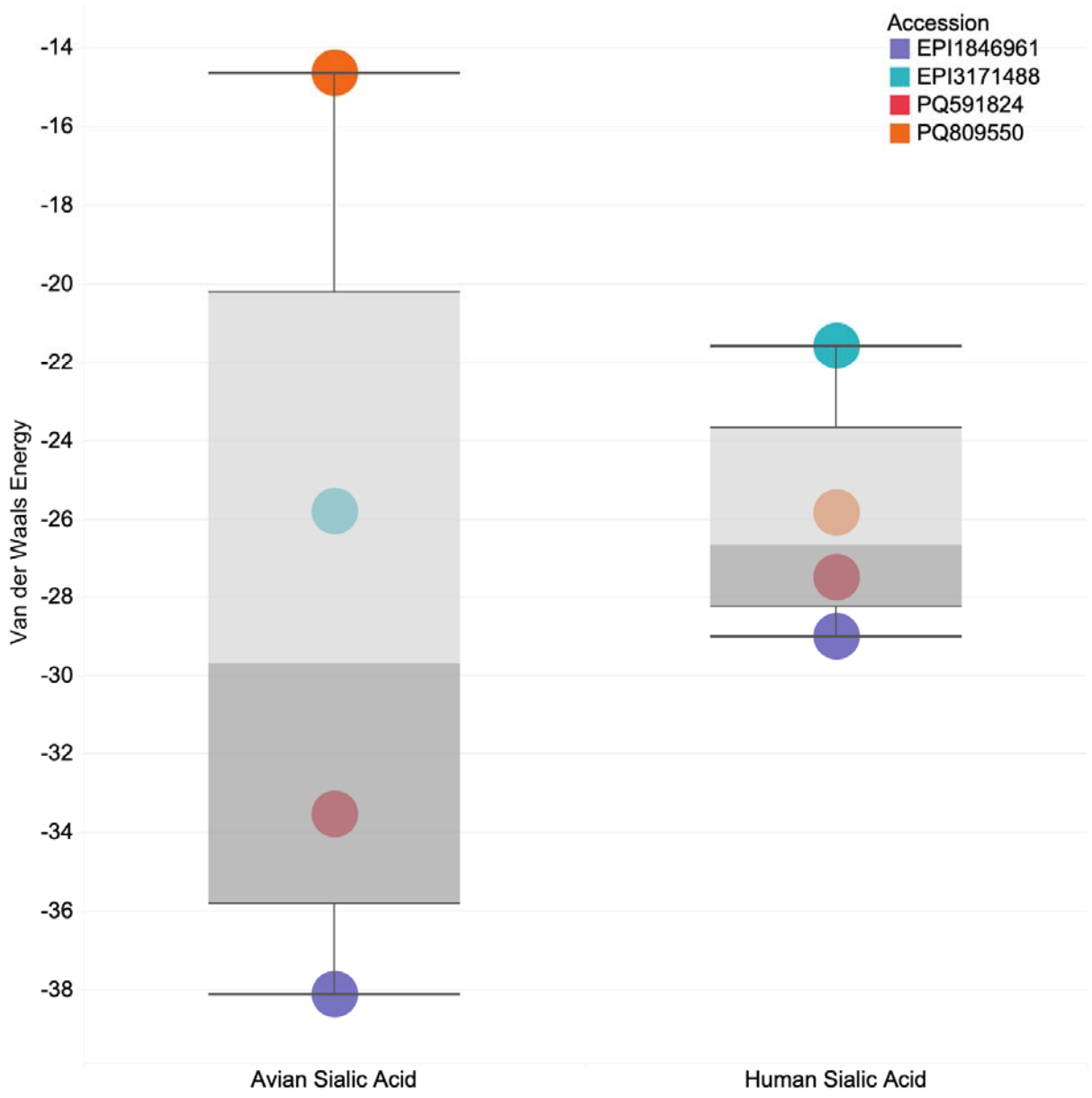

Figure 3: A boxplot graph showing the results of HADDOCK3 docking simulations for hemagglutinin (HA) binding to avian and human sialic acid (SA) glycans. The y-axis represents the Van der Waals energy, which serves as a measure of binding affinity, with lower values indicating stronger binding interactions. The x-axis categorizes the results based on the type of SA (avian or human). Each colored point represents an individual sample [Blue: EPI1846961 (A/Astrakhan/3212/2020, vaccine candidate), Cyan: EPI3171488 (A/Texas/37/2024), Red: PQ591824 (A/California/150/2024), and Orange: PQ809550 (A/Louisiana/12/2024)], and their corresponding energies are plotted for both avian and human SA. The figure suggests that HA binding energies do not show a significant preference between human and avian SA, as the distributions overlap considerably.

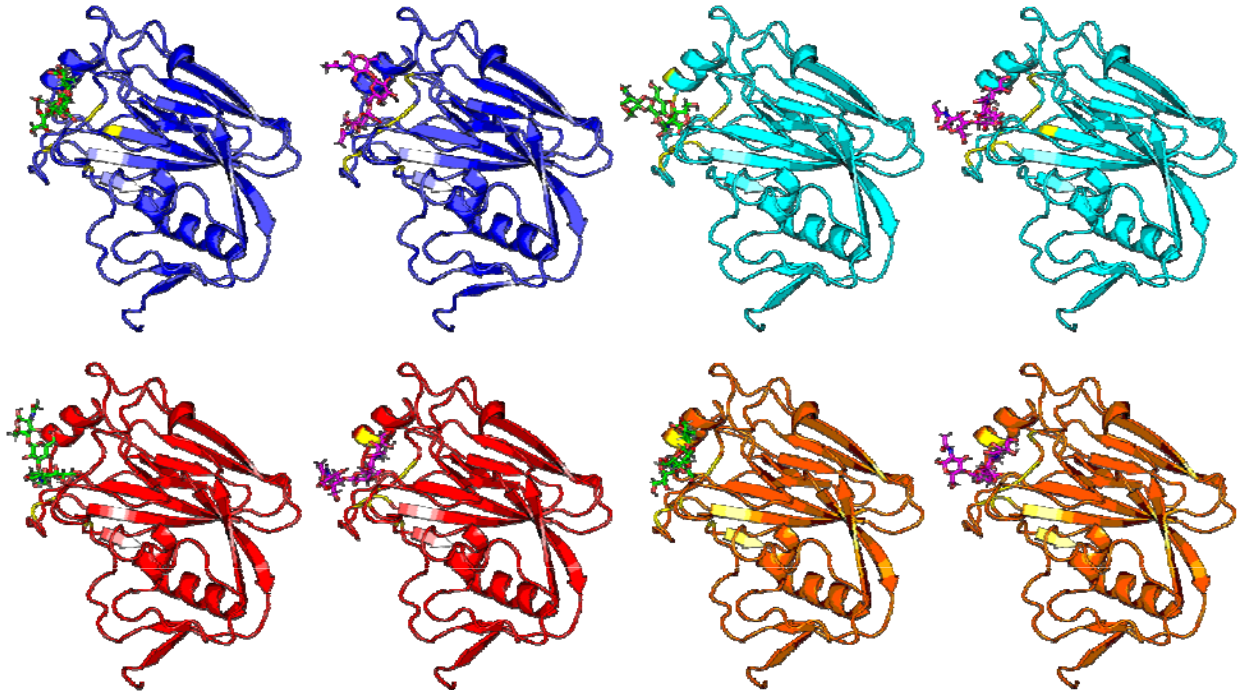

Figure 4: Docking results of avian (green) and human (magenta) sialic acid (SA) with each hemagglutinin (HA) sample. Each colored HA structure represents an individual virus [Blue: EPI1846961 (A/Astrakhan/3212/2020, vaccine candidate), Cyan: EPI3171488 (A/Texas/37/2024), Red: PQ591824 (A/California/150/2024), and Orange: PQ809550 (A/Louisiana/12/2024)]. The analysis reveals minimal variation in the interaction patterns between human and avian SA, as indicated by the similar binding orientations and positions of the SA molecules. Additionally, the binding interactions remain consistent across different HA samples for each type of SA, suggesting a lack of significant structural preference for either human or avian SA.

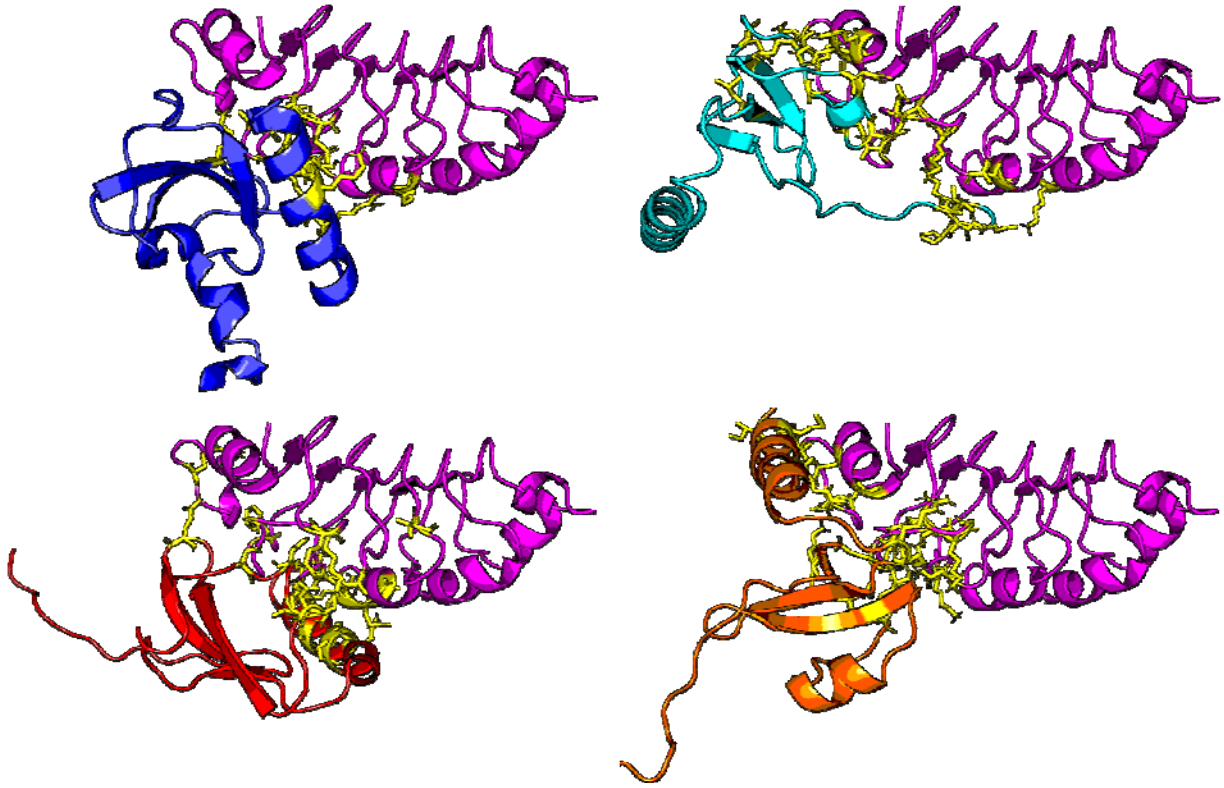

Figure 5: Docking results of ANP32A (in magenta) against PB2 structures. Each colored PB2 structure represents an individual virus [Blue: OQ958041 (A/American Wigeon/South Carolina/22-000345-001/2021, vaccine candidate), Cyan: PP577947 (A/Texas/37/2024), Red: PQ591825 (A/California/150/2024), and Orange: PQ809562 (A/Louisiana/12/2024)].

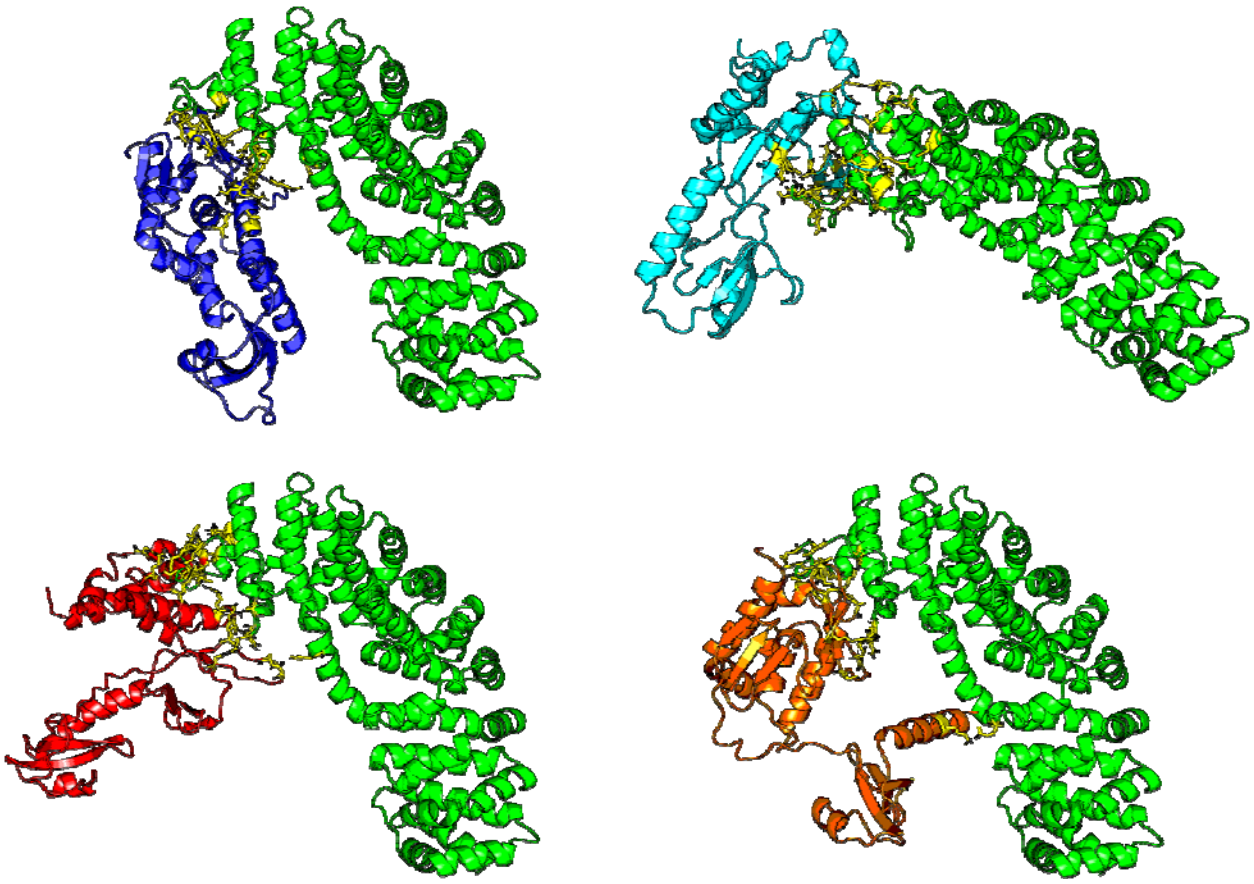

Figure 6: Docking results of IMA3a (in green) against PB2 structures. Each colored PB2 structure represents an individual virus [Blue: OQ958041 (A/American Wigeon/South Carolina/22-000345-001/2021, vaccine candidate), Cyan: PP577947 (A/Texas/37/2024), Red: PQ591825 (A/California/150/2024), and Orange: PQ809562 (A/Louisiana/12/2024)].

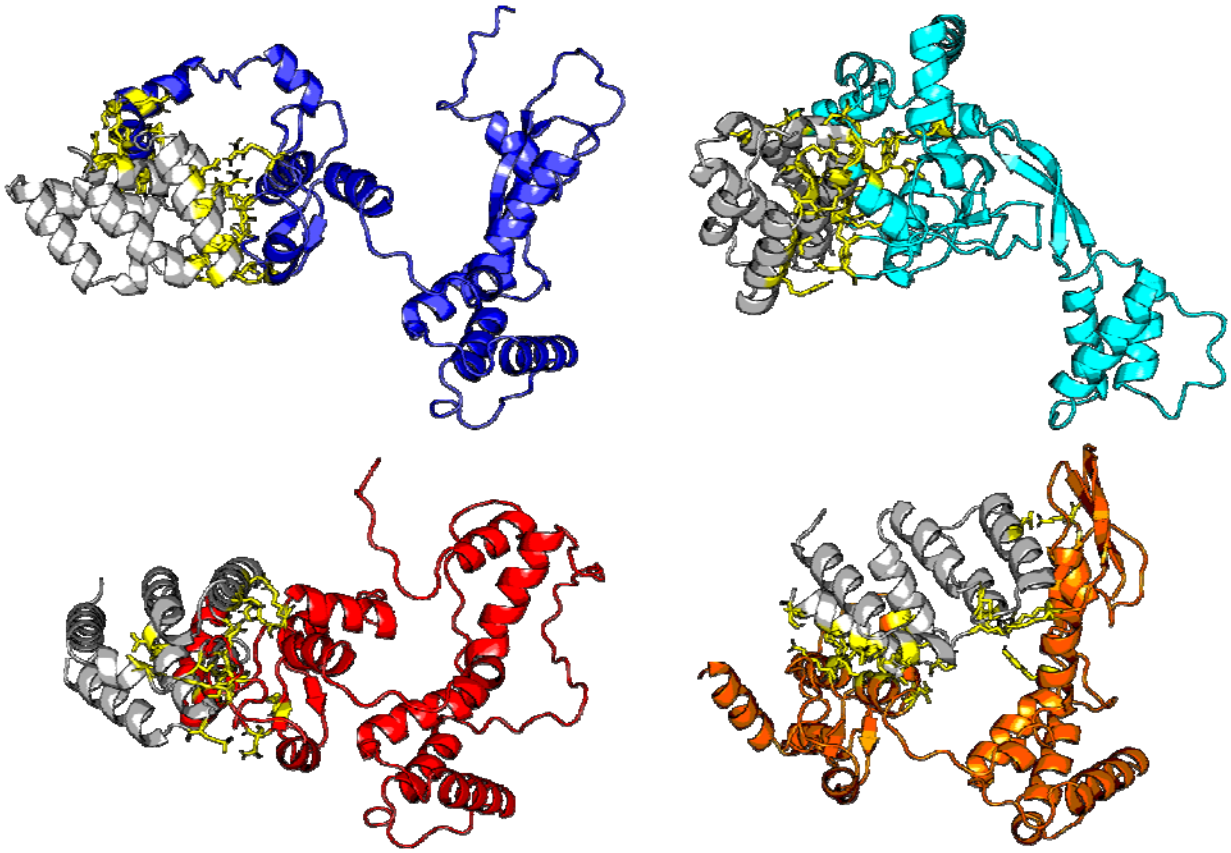

Figure 7: Docking results of MAVS (in grey) against PB2 structures. Each colored PB2 structure represents an individual virus [Blue: OQ958041 (A/American Wigeon/South Carolina/22-000345-001/2021, vaccine candidate), Cyan: PP577947 (A/Texas/37/2024), Red: PQ591825 (A/California/150/2024), and Orange: PQ809562 (A/Louisiana/12/2024)].

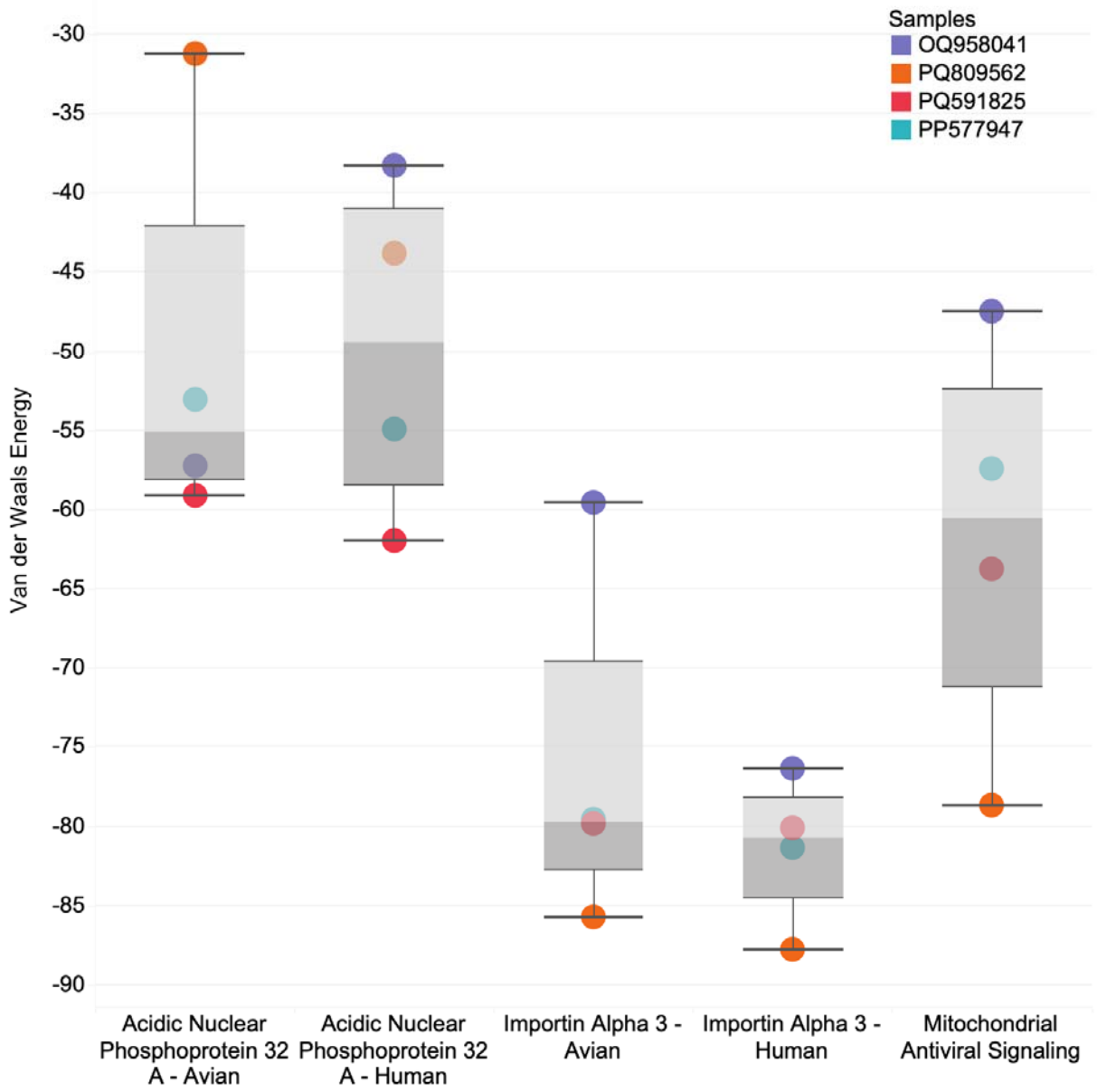

Figure 8: Boxplot graph showing the HADDOCK3 docking results for Polymerase basic 2 (PB2) binding to various host immune response proteins. The y-axis represents the Van der Waals energy. We use this metric to assess binding affinity, with lower values indicating stronger interactions. The x-axis categorizes the results based on the target proteins: ANP32A (avian and human variants), Importin alpha 3 (avian and human variants), and MAVS. Each colored point corresponds to an individual PB2 viral protein [Blue: OQ958041 (A/American Wigeon/South Carolina/22-000345-001/2021, vaccine candidate), Orange: PQ809562 (A/Louisiana/12/2024), Red: PQ591825 (A/California/150/2024), and Cyan: PP577947(A/Texas/37/2024)].

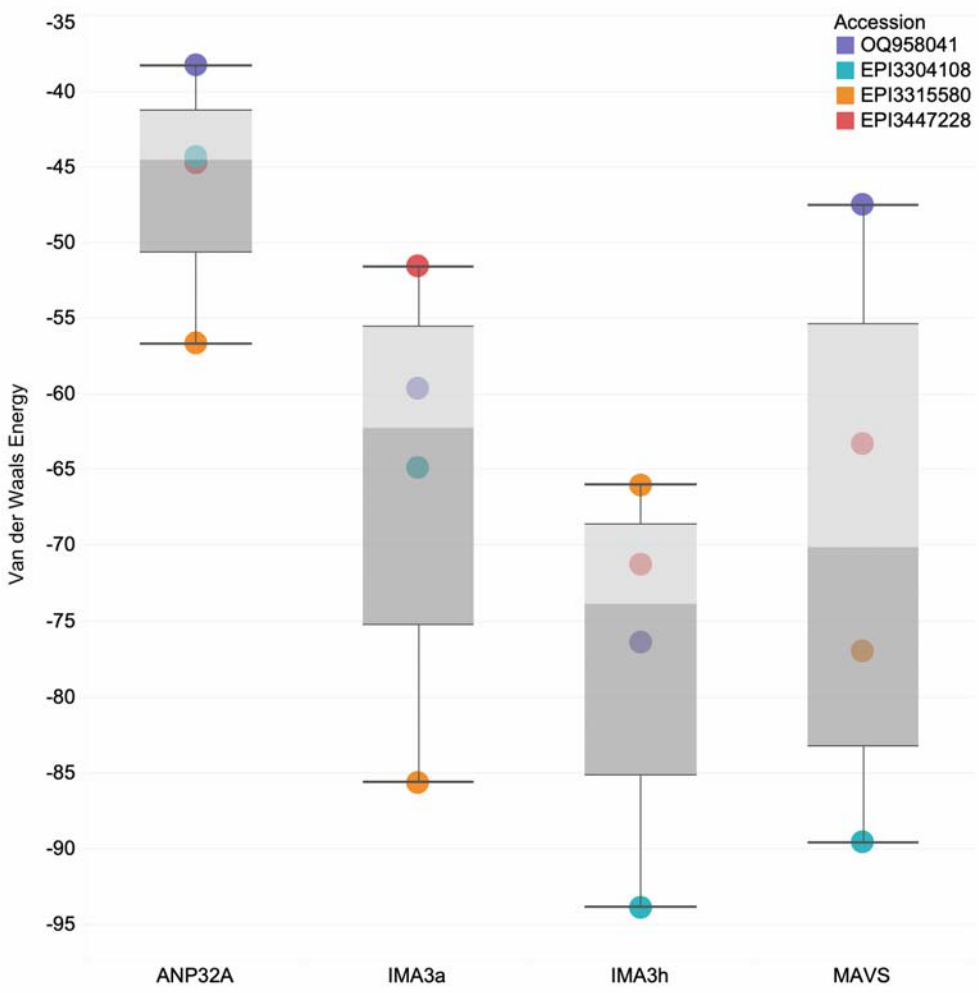

Figure 9: Boxplot graph showing the binding affinity, in Van der Waals energy on the y-axis, from docking simulations of Polymerase Basic 2 (with residue 588 mutations) against target proteins ANP32A, Importin Alpha 3 (avian) (IMA3a), Importin Alpha 3 (human) (IMA3h), and Mitochondrial Antiviral Signaling (MAVS) on the x-axis. Lower Van der Waals energy values indicate stronger binding affinity. The viral sample Blue: OQ958041 (A/American Wigeon/South Carolina/22-000345-001/2021) represents the current vaccine candidate strain carrying the A588 residue. Other samples include Cyan: EPI3304108 (A/Great_Blue_Heron/AB/FAV-0505-25/2022) with the A588T mutation, Orange: EPI3315580 (A/American_Crow/QC/FAV-0050-4/2022) with the A588S mutation, and Red: EPI3447228 (A/chicken/Nebraska/23-040358-001-original/2023) with the A588V mutation. Across all samples, binding affinity to ANP32A and MAVS remains similar to that of the vaccine candidate. Most of the 588 mutants show enhanced binding to human IMA3 compared to avian IMA3, consistent with adaptation to mammalian hosts; however, the A588S mutation (EPI3315580) uniquely shows a decreased affinity to human IMA3, suggesting a potential reduction in nuclear import efficiency. Although all mutant strains exhibit slightly weaker binding to human IMA3 than the vaccine candidate, the differences are minimal.

Acknowledgements

We gratefully acknowledge all GISAID data contributors (i.e., the authors and their originating laboratories) responsible for obtaining the specimens, and their submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which this research is based.

We acknowledge the following entities at the University of North Carolina at Charlotte: Academic Affairs, The Office of Research (Ignite Award), The Center for Computational Intelligence to Predict Health and Environmental Risks (CIPHER), The Department of Bioinformatics and Genomics, The College of Computing and Informatics, and the University Research Computing group. We gratefully acknowledge the support of the Belk Family and the Levine Scholars program.

References

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., Gorden, P. J., Plummer, P. J., & Main, R. (2024). Highly pathogenic avian influenza A(H5N1) clade 2.3.4.4b virus infection in domestic dairy cattle and cats, United States, 2024. Emerging Infectious Diseases, 30(7), 1335–1343. https://stacks.cdc.gov/view/cdc/158121

Cai, M., Zhong, R., Qin, C., Yu, Z., Wen, X., Xian, J., Chen, Y., Cai, Y., Yi, H., Gong, L., & Zhang, G. (2020). The R251K substitution in viral protein PB2 increases viral replication and pathogenicity of Eurasian avian-like H1N1 swine influenza viruses. Viruses, 12(1), 52. https://doi.org/10.3390/v12010052

Canadian Food Inspection Agency. (2025, February 21). Avian influenza (bird flu). https://inspection.canada.ca/en/animal-health/terrestrial-animals/diseases/reportable/avian-influenza

Canadian Food Inspection Agency. (2025). Avian influenza dashboard. https://cfia-ncr.maps.arcgis.com/apps/dashboards/89c779e98cdf492c899df23e1c38fdbc

Carrique, L., Fan, H., Walker, A.P. *et al.* Host ANP32A mediates the assembly of the influenza virus replicase. *Nature* **587**, 638–643 (2020). https://doi.org/10.1038/s41586-020-2927-z

Caserta, L. C., Frye, E. A., Butt, S. L., Laverack, M., Nooruzzaman, M., Covaleda, L. M., Thompson, A. C., Koscielny, M. P., Cronk, B., Johnson, A., Kleinhenz, K., Edwards, E. E., Gomez, G., Hitchener, G., Martins, M., Kapczynski, D. R., Suarez, D. L., Morris, E. R. A., Hensley, T., Beeby, J. S., Lejeune, M., Swinford, A. K., Elvinger, F., Dimitrov, K. M., & Diel, D. G. (2024, July 25). Spillover of highly pathogenic avian influenza H5N1 virus to dairy cattle. Nature. https://doi.org/10.1038/s41586-024-07849-4

CDC. 2020-2024 highlights in the history of avian influenza (Bird flu) timeline. Centers for Disease Control and Prevention. April 30, 2024. Accessed Spring 2025. https://www.cdc.gov/bird-flu/avian-timeline/2020s.html

CDC. (2025, February 26). CDC A(H5N1) Bird Flu Response Update February 26, 2025. https://www.cdc.gov/bird-flu/spotlights/h5n1-response-02262025.html

CDC. (2024). 2020-2024 highlights in the history of avian influenza (bird flu) timeline. https://www.cdc.gov/bird-flu/avian-timeline/2020s.html

CDC. (2025). USDA reported H5N1 bird flu detections in wild birds. https://www.cdc.gov/bird-flu/situation-summary/data-map-wild-birds.html

CFIA NEOC GIS Services. (2025). Avian influenza dashboard. https://cfia-ncr.maps.arcgis.com/apps/dashboards/89c779e98cdf492c899df23e1c38fdbc

Dadonaite B., Ahn, J. J., Ort, J. T., Yu, J., Furey, C., Dosey, A., Hannon, W. W., Vincent, A. L., Webby, R. J., King, N. P., Liu, Y., Hensley, S. E., Peacock, T. P., Moncla, L. H., & Bloom, J. D. (2024). Deep mutational scanning of H5 hemagglutinin to inform influenza virus surveillance. PLoS Biology, 22(11), e3002916–e3002916. https://doi.org/10.1371/journal.pbio.3002916

Deng, Y., Wille, M., Dapat, C., Xie, R., Lay, O., Peck, H., Daley, A. J., Dhanasekaran, V., & Barr, I. G. (2025). Influenza A(H5N1) virus clade 2.3.2.1a in traveler returning to Australia from India, 2024. Emerging Infectious Diseases, 31(1), 135–138. https://doi.org/10.3201/eid3101.241210

Environment and Climate Change Canada. (2024, October 9). Avian influenza in wild birds. https://www.canada.ca/en/environment-climate-change/services/migratory-game-bird-hunting/avian-influenza-wild-birds.html

Focosi, D., & Maggi, F. (2024). Avian Influenza Virus A(H5Nx) and Prepandemic Candidate Vaccines: State of the Art. International Journal of Molecular Sciences, 25(15), 8550–8550. https://doi.org/10.3390/ijms25158550

Ford C. T., Jacob Machado D, Janies DA. 2022. Predictions of the SARS-CoV-2 Omicron Variant (B.1.1.529) Spike Protein Receptor-Binding Domain Structure and Neutralizing Antibody Interactions. *Frontiers in Virology*.;2. doi:10.3389/fviro.2022.830202

Ford, C. T., Yasa, S., Obeid, K., Jaimes, R., III, Tomezsko, P. J., Guirales-Medrano, S., White, R. A., III, & Janies, D. (2025). Large-scale computational modelling of H5 influenza variants against HA1-neutralising antibodies. eBioMedicine, 105632. https://doi.org/10.1016/j.ebiom.2025.105632

Gao, W., Zu, Z., Liu, J., Song, J., Wang, X., Wang, C., Liu, L., Tong, Q., Wang, M., Sun, H., Sun, Y., Liu, J., Chang, K.-C., & Pu, J. (2019). Prevailing I292V PB2 mutation in avian influenza H9N2 virus increases viral polymerase function and attenuates IFN-β induction in human cells. Journal of General Virology, 100(9), 1273–1281. https://doi.org/10.1099/jgv.0.001294

Ghafoori, S. M., Petersen, G. F., Conrady, D. G., et al. (2023). Structural characterisation of hemagglutinin from seven Influenza A H1N1 strains reveal diversity in the C05 antibody recognition site. Scientific Reports, 13(6940). https://doi.org/10.1038/s41598-023-33529-w

Goloboff PA, Morales ME. (2023). TNT version 1.6, with a graphical interface for MacOS and Linux, including new routines in parallel. *Cladistics*. 2023;39(2):144-153. doi:10.1111/cla.12524

Graef, K. M., Vreede, F. T., Lau, Y.-F., McCall, A. W., Carr, S. M., Subbarao, K., & Fodor, E. (2010). The PB2 Subunit of the Influenza Virus RNA Polymerase Affects Virulence by Interacting with the Mitochondrial Antiviral Signaling Protein and Inhibiting Expression of Beta Interferon. Journal of Virology, 84(17), 8433–8445. https://doi.org/10.1128/JVI.00879-10

Gu, C., Maemura, T., Guan, L., Watanabe, T., Kiso, M., Yamayoshi, S., Imai, M., Suzuki, Y., Neumann, G., & Kawaoka, Y. (2024). A human isolate of bovine H5N1 is transmissible and lethal in animal models. Nature, 636, 711–718. https://doi.org/10.1038/s41586-024-08254-7

Guan, L., Eisfeld, A. J., Pattinson, D., Gu, C., Biswas, A., Maemura, T., Trifkovic, S., Babujee, L., Presler, R., Dahn, R., Halfmann, P. J., Barnhardt, T., Neumann, G., Thompson, A., Swinford, A. K., Dimitrov, K. M., Poulsen, K., & Kawaoka, Y. (2024). Cow's Milk Containing Avian Influenza A(H5N1) Virus - Heat Inactivation and Infectivity in Mice. The New England Journal of Medicine. https://doi.org/10.1056/NEJMc2405495

Guo, Y., Shu, S., Zhou, Y., Peng, W., Jiang, Z., Li, Y., Li, T., Du, F., Wang, L., Chen, X., Dong, J., Zhao, C., Wang, M. H., Sun, Y., Sun, H., Lu, L., Digard, P., Chang, K.-C., Yen, H.-L., Liu, J., & Pu, J. (2024). An emerging PB2-627 polymorphism increases the pandemic potential of avian influenza virus by breaking through ANP32 host restriction in mammalian and avian hosts. bioRxiv. https://doi.org/10.1101/2024.07.03.601996

Gupta, D., Kumar, M., Sharma, P., Mohan, T., Prakash, A., Kumari, R., & Kaur, P. (2022). Effect of double mutation (L452R and E484Q) on the binding affinity of monoclonal antibodies (mAbs) against the RBD-A target for vaccine development. Vaccines, 11(1), 23. https://doi.org/10.3390/vaccines11010023

Hanson, A., Imai, M., Hatta, M., McBride, R., Imai, H., Taft, A., Zhong, G., Watanabe, T., Suzuki, Y., Neumann, G., Paulson, J. C., & Kawaoka, Y. (2015). Identification of stabilizing mutations in an H5 hemagglutinin influenza virus protein. Journal of Virology, 90(6), 2981–2992. https://doi.org/10.1128/JVI.02790-15

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution, 30(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kayali, G., Kandeil, A., El-Shesheny, R., Kayed, A. S., Maatouq, A. M., Cai, Z....Ali, M. A. (2016). Avian Influenza A(H5N1) Virus in Egypt. Emerging Infectious Diseases, 22(3), 379-388. https://doi.org/10.3201/eid2203.150593.

Kosakovsky Pond, S. L., Poon, A. F., Velazquez, R., Weaver, S., Hepler, N. L., Murrell, B., Shank, S. D., Magalis, B. R., Bouvier, D., Nekrutenko, A., Wisotsky, S., Spielman, S. J., Frost, S. D. W., & Muse, S. V. (2020). HyPhy 2.5—A customizable platform for evolutionary hypothesis testing using phylogenies. Molecular Biology and Evolution, 37(1), 295–299. https://doi.org/10.1093/molbev/msz197

Kwon, T, Trujillo, JD, Carossino, M, Lyoo, EL, McDowell, CD, Cool, K, Matias-Ferreyra, FS, Jeevan, T, Morozov, I, Gaudreault, NN, Balasuriya, UBR, Webby, RJ, Osterrieder, N, & Richt, JA. (2024). Pigs are highly susceptible to but do not transmit mink-derived highly pathogenic avian influenza virus H5N1 clade 2.3.4.4b. Emerging Microbes & Infections, 13(1), 2353292. https://doi.org/10.1080/22221751.2024.2353292

Li, B., Su, G., Xiao, C., Zhang, J., Li, H., Sun, N., Lao, G., Yu, Y., Ren, X., Qi, W., Wang, X., & Liao, M. (2021). The PB2 co□adaptation of H10N8 avian influenza virus increases the pathogenicity to chickens and mice. Transboundary and Emerging Diseases, 69(4), 1794–1803. https://doi.org/10.1111/tbed.14157

Long, J. C. D., & Fodor, E. (2016). The PB2 subunit of the influenza A virus RNA polymerase is imported into the mitochondrial matrix. Journal of Virology, 90(24), e01384-16. https://doi.org/10.1128/jvi.01384-16

Louisiana Department of Health. (2025). LDH reports first U.S. H5N1-related human death: Current general public health risk remains low. https://www.ldh.la.gov

Lee, I., Il Kim, J., Park, S., Bae, J. Y., Yoo, K., Yun, S. H., Lee, J. Y., Kim, K., Kang, C., & Park, M. S. (2017). Single PA mutation as a high yield determinant of avian influenza vaccines. Scientific Reports, 7, 40675. https://doi.org/10.1038/srep40675

Lupiani, B, & Reddy, SM. (2009). The history of avian influenza. Comparative Immunology, Microbiology and Infectious Diseases, 32(4), 311–323. https://doi.org/10.1016/j.cimid.2008.01.004

Ma, X., Xie, L., Wartchow, C., Warne, R., Xu, Y., Rivkin, A., Sutton, J., Tong, L., & Shih, C. (2017). Structural basis for therapeutic inhibition of influenza A polymerase PB2 subunit. Scientific Reports, 7, 9385. https://doi.org/10.1038/s41598-017-09538-x

Maddison, WP, & Maddison, DR. (2023). Mesquite: A modular system for evolutionary analysis (Version 3.81). http://www.mesquiteproject.org

Maines, T. R., Lu, X. H., Erb, S. M., Edwards, L., Guarner, J., Greer, P. W., Nguyen, D. C., Szretter, K. J., Chen, L., Thawatsupha, P., Chittaganpitch, M., Waicharoen, S., Nguyen, D. T., Nguyen, T., Nguyen, H. H. T., Kim, J., Hoang, L. T., Kang, C., Phuong, L. S., Lim, W., Zaki, S., Donis, R. O., Cox, N. J., Katz, J. M., & Tumpey, T. M. (2005). Avian influenza (H5N1) viruses isolated from humans in Asia in 2004 exhibit increased

virulence in mammals. Journal of Virology, 79(18), 11788–11800. https://doi.org/10.1128/jvi.79.18.11788-11800.2005

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: Making protein folding accessible to all. Nature Methods, 19(6), 679–682. https://doi.org/10.1038/s41592-022-01488-1

Mukherjee, A., Nayak, M. K., Dutta, S., Panda, S., Satpathi, B. R., & Chawla-Sarkar, M. (2016). Genetic characterization of circulating 2015 A(H1N1)pdm09 influenza viruses from Eastern India. PLoS ONE, 11(12), e0168464. https://doi.org/10.1371/journal.pone.0168464

Neumann, G., Macken, C. A., & Kawaoka, Y. (2014). Identification of amino acid changes that may have been critical for the genesis of A(H7N9) influenza viruses. Journal of Virology, 88(9), 4877–4896. https://doi.org/10.1128/jvi.00107-14

Nielsen R, Yang Z. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*. 1998;148(3):929-936. doi:10.1093/genetics/148.3.929

Nilsson, B. E., Te Velthuis, A. J. W., & Fodor, E. (2017). Role of the PB2 627 domain in influenza A virus polymerase function. Journal of Virology, 91(7), e02467-16. https://doi.org/10.1128/JVI.02467-16

PAHO. (2025). Epidemiological Update Avian Influenza A(H5N1) in the Americas Region 4 March 2025 Global Context. https://www.paho.org/sites/default/files/2025-03/2025-mar-4-phe-epidupdate-avianinfluenza-eng-final.pdf

Pulit-Penaloza, J. A., Belser, J. A., Brock, N., et al. (2024). Transmission of a human isolate of clade 2.3.4.4b A(H5N1) virus in ferrets. Nature, 636, 705–710. https://doi.org/10.1038/s41586-024-08246-7

Ranwez, V., Douzery, E. J. P., Cambon, C., Chantret, N., & Delsuc, F. (2018). MACSE v2: Toolkit for the Alignment of Coding Sequences Accounting for Frameshifts and Stop Codons. Molecular Biology and Evolution, 35(10), 2582–2584. https://doi.org/10.1093/molbev/msy159

SENASICA. (2023). Compartimentos libres de influenza aviar. https://datos.gob.mx/busca/dataset/compartimentos-libres-de-influenza-aviar

Singh, G., Trujillo, J. D., McDowell, C. D., Franco Matias-Ferreyra, Sujan Kafle, Kwon, T., Gaudreault, N. N., Fitz, I., Noll, L., Morozov, I., Retallick, J., & Richt, J. A. (2024). Detection and characterization of H5N1 HPAIV in environmental samples from a dairy farm. Virus Genes. https://doi.org/10.1007/s11262-024-02085-4

Shinya, K., Hamm, S., Hatta, M., Ito, H., Ito, T., & Kawaoka, Y. (2004). PB2 amino acid at position 627 affects replicative efficiency, but not cell tropism, of Hong Kong H5N1

influenza A viruses in mice. Virology, 320(2), 258–266. https://doi.org/10.1016/j.virol.2003.11.030

Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies, Bioinformatics, Volume 30, Issue 9, May 2014, Pages 1312–1313, https://doi.org/10.1093/bioinformatics/btu033

Subbarao, E. K., London, W., & Murphy, B. R. (1993). A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. Journal of Virology, 67(4), 1761–1764. https://doi.org/10.1128/jvi.67.4.1761-1764.1993

Teixeira, J. M. C., Honorato, R. V., Giulini, M., Bonvin, A., Alidoost, S., Reys, V., Jimenez, B., Schulte, D., van Noort, C., Verhoeven, S., Vreede, B., Schott, S., & Tsai, R. (2024). haddocking/haddock3: v3.0.0-beta.5 (Version v3.0.0-beta.5) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.10527751

Tin, A. (2025). U.S. egg industry sees record chicken deaths from bird flu outbreak. CBS News. https://www.cbsnews.com/news/egg-industry-chicken-deaths-bird-flu/

USDA-APHIS. (2025). U.S. Department of Agriculture Animal and Plant Health Inspection Service. APHIS confirms D1.1 genotype in dairy cattle in Nevada. https://www.aphis.usda.gov/news/program-update/aphis-confirms-d11-genotype-dairy-cattle-nevada-0

Wandzik, J. M., Kouba, T., & Cusack, S. (2021). Structure and function of influenza polymerase. Cold Spring Harbor Perspectives in Medicine, 11(9), a038372. https://doi.org/10.1101/cshperspect.a038372

World Health Organization. (2024, February 23). Summary of status of development and availability of A(H5N1) candidate vaccine viruses and potency testing reagents. https://cdn.who.int/media/docs/default-source/influenza/cvvs/cvv-zoonotic-northern-hemisphere-2024-2025/h5n1_summary_a_h5n1_cvv_20240223.pdf?sfvrsn=2d559bc4_5

World Organization for Animal Health. (Accessed March 1, 2025). Avian influenza. https://www.woah.org/en/disease/avian-influenza/#ui-id-2

World Organization for Animal Health. (2025). Mexico - Influenza A viruses of high pathogenicity (Inf. with) (non-poultry including wild birds) (2017-). Woah.org.https://wahis.woah.org/#/in-event/5820/dashboard

Xiong, X., McCauley, J. W., & Steinhauer, D. A. (2014). Receptor binding properties of the influenza virus hemagglutinin as a determinant of host range. In R. W. Compans & M. B. A. Oldstone (Eds.), Influenza pathogenesis and control - Volume I (pp. 63–91). Springer International Publishing. https://doi.org/10.1007/82_2014_423

Yamaji, R., Yamada, S., Le, M. Q., Li, C., Chen, H., Qurnianingsih, E., Nidom, C. A., Ito, M., Sakai-Tagawa, Y., & Kawaoka, Y. (2015). Identification of PB2 Mutations

Responsible for the Efficient Replication of H5N1 Influenza Viruses in Human Lung Epithelial Cells. Journal of Virology, 89(7), 3947–3956. https://doi.org/10.1128/jvi.03328-14

Yang, H., Dong, Y., Bian, Y., Xu, X., Liu, X., Zhang, Y., Chen, Y., Wang, X., Li, X., Xiao, H., Li, Z., & Chen, H. (2022). The influenza virus PB2 protein evades antiviral innate immunity by inhibiting JAK1/STAT signalling. Nature Communications, 13, 6288. https://doi.org/10.1038/s41467-022-33909-2

Yu, M., Qu, Y., Zhang, H., Zhang, Y., Li, X., & Chen, H. (2022). Roles of ANP32 proteins in cell biology and viral replication. Animal Diseases, 2, 22. https://doi.org/10.1186/s44149-022-00055-7

Zhou, Y., Wu, X., Yan, D., Chen, C., Liu, X., Huang, C., Fu, X., Tian, G., Ding, C., Wu, J., Xu, J., Li, L., & Yang, S. (2021). V292I mutation in PB2 polymerase induces increased effects of E627K on influenza H7N9 virus replication in cells. Virus research, 291, 198186. https://doi.org/10.1016/j.virusres.2020.198186