Phylogenetic discordance cannot identify genomic reassortment events: a case study with the avian influenza virus H5Nx

Hugo G. Castelán-Sánchez[1] and Art F. Y. Poon[1,2,3]

[1] Department of Pathology & Laboratory Medicine, Western University, Dental Sciences Building, Rm. 201, London, Ontario N6A 5C1, Canada
[2] Department of Microbiology & Immunology, Western University, 1151 Richmond Street, London, Ontario N6A 3K7, Canada
[3] Department of Computer Science, Western University, Rm. 355, Middlesex College, London N6A 5B7, Canada

Abstract

Recombination plays a central role in the evolution of the RNA virus genome, as it allows the exchange of genetic material between different viral lineages.Reassortment is a form of recombination specific to viruses with segmented genomes that involves the exchange of entire segments between lineages. This process has played a significant role in the adaptation and spread of influenza viruses, where novel combinations of genomic segments are produced from divergent lineages (*i.e.*, antigenic shift). In this study, we investigated reassortment events in a representative sample of $n = 11,765$ complete genomes of H5Nx influenza A viruses from avian hosts. Using tanglegrams and subtree-prune-regraft (SPR) methods to quantify discordance between segment trees, we found evidence of widespread reassortment among these genomes that is consistent with previous studies. However, we obtained the comparable numbers of reassortment events when applying the same methods to sequence data simulated in the complete absence of any reassortment. This outcome indicates a high false positive rate in identifying reassortment events due to uncertainty in phylogenetic reconstruction. There was also no significant difference between SPR events inferred from real data versus simulated false positives with respect to subtree size, or the distance between prune and regraft points in the reference tree. Although there is no doubt that reassortment has played a significant role in the origin of new pandemic strains of influenza viruses, these

results imply that previous reports of extensive reassortment in the avian influenza virus literature should be interpreted with caution.

Introduction

Recombination is an important mechanism of genomic evolution in RNA viruses (Simon-Loriere and Holmes 2011). This process involves the exchange of genetic material between different genomic backgrounds, which is facilitated by the coinfection of a host by two or more distinct virus lineages. Recombination uncouples the evolutionary fate of a mutation from the rest of the genome, and under certain conditions (Otto and Barton 1997), recombination will increase the diversity of genotypes in the population (Pérez-Losada et al. 2015; Wang et al. 2022). Many RNA viruses have segmented genomes, including members of the *Reoviridae* and *Orthomyxoviridae* families (McDonald et al. 2016). In these viruses, recombination can occur by reassortment, in which the genomic segments packaged into a virion are derived from different parental genomes. Additionally, many other viruses have multipartite genomes where segments are packaged into separate virus particles. These viruses do not have to coinfect the same cell for reassortment to occur (Varsani et al. 2018). Reassortment involves the wholesale exchange of entire segments between genomes, which can accelerate the acquisition of combinations of mutations that are required to overcome host barriers more rapidly than the progressive accumulation of individual mutations (Vijaykrishna et al. 2015).

Influenza A virus (IAV; family *Orthomyxoviridae*, genus *Alphainfluenza*) provides several important examples of the possible impacts of genetic reassortment on the health of human and non-human animal populations. This virus has been responsible for numerous and severe epidemics and pandemics (Sutton 2018). IAV has a negative-sense RNA genome that consists of eight genomic segments that encode at least 10 polypeptides (Pinto et al. 2021). This number of segments implies that about 99% virus particles produced by a co-infected host cell will contain a reassortant genome. The co-circulation of multiple IAV subtypes and clades in domesticated animal host

species provides abundant opportunities for reassortment between divergent lineages (Ryt-Hansen et al. 2021; Li et al. 2024). Thus, reassortment can generate new chimeric IAV genomes that potentially increase the capacity of this virus to be transmitted to a new host species and establish a productive infection (Bouvier and Palese 2008; Ma et al. 2016).

Comparative analysis of IAV genome sequences plays an important role in measuring the frequency of reassortment and detecting the emergence of new reassortant viruses (Kong et al. 2015; Taylor et al. 2023). To this end, a number of different bioinformatic methods have been developed to identify reassortment events. In the absence of recombination, the evolutionary history of a sample population of genome sequences may be represented by a single phylogenetic tree. Reassortment causes each segment to follow its own evolutionary history. Consequently, a common technique to detect reassortment is to reconstruct a phylogeny for each segment, and then to identify discrepancies in the topology and/or labels among the resulting trees. Discordant trees are often visualized as tanglegrams (Page 2003), in which two trees with a common set of labels are drawn in opposing orientations, with subtrees rotated to maximize the concordance of tip labels between trees. Tanglegrams can be used to visually assess a small number of reassortments. In addition, the optimization algorithm for untangling pairs of trees provides a rough measure of the extent of discordance that may be due to recombination. However, tanglegrams do not readily scale to resolving large numbers of reassortment events. Evolutionary events that produce discordant trees can be explicitly modeled using an ancestral recombination graph (ARG; Wong et al. 2024). Although there has been a recent resurgence of interest in ARGs (Frost et al. 2015), they remain too computationally complex for large-scale application.

A compromise between these two extremes is to infer the frequency of reassortment events by the number of tree rearrangement operations required to convert one segment tree to another (Svinti et al. 2013), which can be interpreted as an edit distance. Screening for tips that transition between clusters of labels in the respective trees is an equivalent approach Silva et al. 2012; Gong et al. 2021. There are many edit distances available for this purpose, including the Robinson-

Foulds distance (RF; Briand et al. 2020), tree bisection and reconnection (TBR; Kelk and Linz 2020), nearest neighbor interchange (NNI; Collienne and Gavryushkin 2021), and subtree prune and regraft (SPR; Wu 2009). Among these, the SPR distance is frequently used in studies of lateral gene transfer and recombination (Whidden et al. 2014), possibly because the SPR most closely resembles the effect of recombination on the topologies of trees on either side of the breakpoint.

In this work, we performed a phylogenetic analysis of nearly twelve thousand full genome sequences for influenza A H5Nx viruses to identify putative rearrangement events by reconstructing SPR rearrangements. Our main objective was to evaluate the reliability of this approach given the uncertainty in reconstructing accurate phylogenies under realistic conditions. To this end, we reconstructed phylogenies for each IAV H5Nx segment and then selected one of these trees as the template on which we simulated sequence evolution for the other segments in the absence of resorting. We compared the entanglement values and the SPR distances between the reference tree for segment 4 (encoding hemagglutinin, HA) and trees for all other segments for both real and simulated data sets to assess the specificity of these methods for detecting reassortment. In addition, we characterized the SPR operations reconstructed from these data in terms of the size of the subtree and the distance traversed in the tree. In general, we found that the results were indistinguishable between the real IAV H5Nx data and the data simulated in the complete absence of reassortment. These results imply that approaches based on inferring rearrangements in discordant trees suffer from a high false-positive rate in identifying reassortment events in IAV genomes, due to the inherent uncertainty in reconstructing phylogenies.

Methods

Data collection

We retrieved influenza A virus H5Nx genome sequences isolated from avian hosts from the NCBI Genbank and GISAID databases (accessed 5 May 2024). The GISAID query returned 5,994 to 6,000 sequences per segment, and the Genbank query returned 6,157 to 6,170 sequences per seg-

ment. We combined these sequence data sets by genomic segment. There was only a small proportion of sequences that were exact duplicates between databases. For example, out of 12,167 neuraminidase (NA) sequences, there were 8,347 unique sequences, of which 1,417 (17%) had two or more instances — only 171 (12%) of these were represented in both databases, involving a total of 673 sequence records. This low amount of overlap is largely due to a transition away from depositing sequences in Genbank to GISAID around 2020, coinciding with recent outbreaks of avian H5N1 IAV (Supplementary Figure S1). Sequence identity was not sufficient to determine whether the sequences are derived from the same isolate. Consequently, we merged these datasets by harmonizing the sequence labels into the World Health Organization (1980) nomenclature format, including host species, country of sampling, isolate name and year of sampling, and then filtering sequences with duplicate labels. This reduced the combined data to 11,795 sequences per segment. We generated alignments for each segment using MAFFT (v7.490; Katoh and Standley 2013), removed duplicate sequences, and then manually inspected and cleaned each segment alignment with trimAl (Capella-Gutiérrez et al. 2009, p. v1.5.0).

We reconstructed maximum likelihood phylogenies using the standard generalized time-reversible (GTR) model with gamma-distributed rate variation and empirical base frequencies IQ-TREE2 (Minh et al. 2020). To reduce the number of genomes to a representative subset, we used the prune function in Biopython Phylo module to progressively remove terminal branches in the HA tree with a length exceeding a cutoff. After experimenting with different length cutoffs, we selected a cutoff of 0.05 expected substitutions per site to yield a subset of 206 HA sequences. We used the corresponding labels to filter alignments and reconstruct trees for the other segments.

Data simulation

To quantify the discordance among phylogenies reconstructed from sequences that have evolved in the complete absence of reassortment, we simulated sequence alignments along the tree inferred from the real HA alignment (T_{HA}) for every other segment. Since every segment would evolve along exactly the same tree topology, this created a synthetic dataset with a complete absence of

reassortment. In other words, any reassortment event inferred for the simulated trees could be interpreted as a false positive. We proportionally rescaled branch lengths in T_{HA} so that the total tree length corresponded to the length of the tree reconstructed from actual sequences of that segment. Our motivation for adjusting tree lengths in this manner was to accommodate the substantial variation in rates of evolution among segments, which could otherwise confound our assessment of false positive rates. For each rescaled tree, we used the program Pyevolve (Spielman and Wilke 2015) to simulate sequence evolution under the default Goldman-Yang codon substitution model with the dN/dS ratio set to 0.5 and varying segment lengths.

Phylogenetic analysis

We used two approaches to infer reassortment events. First we used a tanglegram, which is a graphical comparison of two trees with a common set of tip labels, where the trees have been rotated and optionally re-rooted to minimize the discordance in the vertical location of labels. Our working assumption is that increasing amounts of reassortment will induce greater amounts of 'entanglement' between trees of different segments. We used the *step2side* method implemented in the Python module *tanglegram* (Nguyen et al. 2022) to disentangle the HA tree and trees from every other segment. The resulting tanglegrams were further optimized and visualized using the nearest-neighbor (NN) tanglegram method implemented in *SplitTree* (Scornavacca et al. 2011).

Second, we used the *Espalier* library (Rasmussen and Guo 2023) in Python to efficiently calculate the approximate number of SPR (subtree prune and regraft) operations required to convert one phylogenetic tree into another. This number provides another means of estimating the number of reassortment events between phylogenetic trees. We obtained SPR distances between the HA tree and the trees for every other segment for the actual IAV genome sequence data, and we made the same comparisons for trees reconstructed from simulated data. These comparisons were visualized as a slopegraph using the R package *ggfree* (https://github.com/ArtPoon/ggfree).

Our hypothesis was that false positive reassortment events due to phylogenetic uncertainty would tend to involve SPR moves separated by shorter distances in the respective trees. To evaluate

this hypothesis, we quantified the total path length separating the inferred graft points for each SPR operation. Let the tree reconstructed from real sequences for the i-th segment be denoted by T_i. Moreover, let the tree reconstructed from simulated sequences that were evolved along the real HA tree T_4 be denoted by T_i^*. For each subtree S_j in the maximum agreement forest between trees T_4 and T_i for $i \neq 4$, we first locate S in T_4 by searching for the internal node with the fewest descendants that match all tip labels in S_j. We call this node the root of subtree j in T_4 and denote it as $r_4^{(j)}$. Next, we located the internal node in T_4 with the same direct common ancestor as $r_4^{(j)}$, which we refer to as its sibling node, $s_4^{(j)}$. Switching to the other tree T_i, we locate the same subtree S_j to identify $r_i^{(j)}$. Finally, we locate the internal node in T_i with the fewest descendants that match all tip labels associated with $s_4^{(j)}$, which we denote as $s_i^{(j)}$. The sibling distance is then defined as the total path length (*i.e.*, sum of branch lengths) from $s_i^{(j)}$ to $r_i^{(j)}$. Put another way, $s_i^{(j)}$ is the approximate location where the SPR operation pruned the subtree from, and $r_i^{(j)}$ is the location where the subtree was grafted to. This distance in T_4 is more convenient because we can consistently compare sibling distances for SPRs involving different segments in the same reference tree. Finally, we calculated the sibling distances for all SPR events between T_4 and the simulated trees T_i^*.

Results

We obtained a total of 11,795 H5Nx influenza A virus genome sequences from the NCBI Genbank and GISAID databases. These sequences represented a variety of avian hosts, with the most frequent annotations being 'chicken' ($n = 2,968$), 'duck' ($n1,774$) and 'turkey' ($n = 927$). There was a total of 527 unique annotations for host organisms at varying taxonomic levels; for instance, mallard duck ($n = 47$) and Muscovy duck ($n = 199$) were both sub-classifications below 'duck'. Similarly, there were 621 unique geolocation annotations with the three most common being 'Netherlands' ($n = 845$), 'Minnesota' ($n = 569$) and 'Vietnam' ($n = 557$). Many records were labeled by state instead of country, obscuring a relative abundance of samples from China and

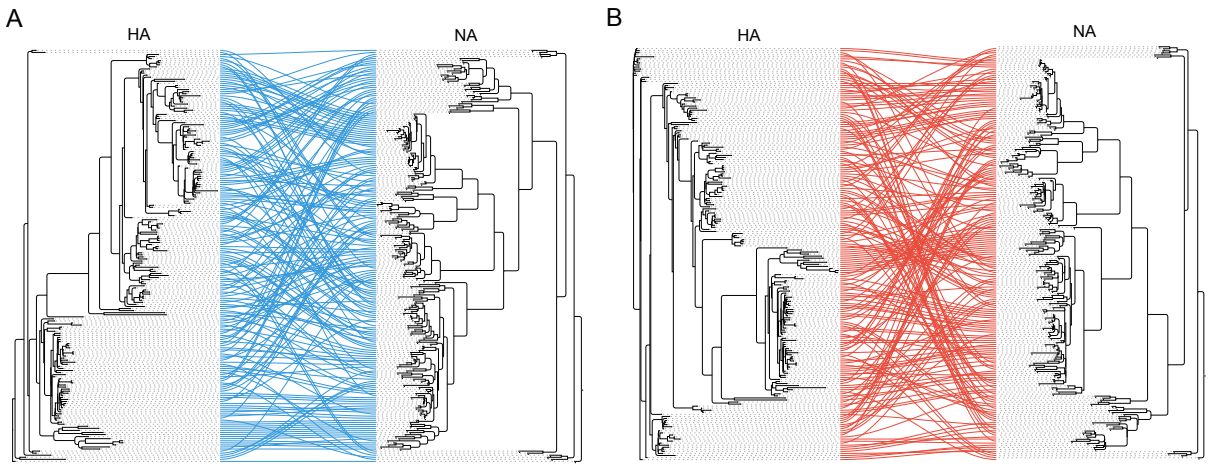

Figure 1: Tanglegrams between $n = 206$ influenza A virus H5Nx sequences for segments 4 (HA) and 6 (NA), with real (A) and simulated (B) data for the latter. Both examples display similar amounts of entanglement (0.27 and 0.22, respectively). These diagrams were generated using SplitTree.

the United States, for example. We reconstructed a maximum likelihood phylogeny relating all HA (segment 4) sequences, and then we generated a representative sub-sample of $n = 206$ (1.7%) sequences by progressively removing the shortest terminal branches. Next, we extracted the corresponding sequences for the other segments and regenerated maximum likelihood phylogenies for each segment.

To get an overview of the potential for reassortment in the evolutionary history of avian H5Nx genomes, we generated tanglegrams for the HA tree against the trees of every other segment. A tanglegram is a graphical method that can re-root and rotate branches in two trees with matching labels to minimize the discordance in the vertical positions of their labels. Thus, it can provide a convenient visual assessment of the extent of reassortment. For example, a tanglegram between the HA and NA trees displays a high level of entanglement (Figure 1A) that could be interpreted as widespread reassortment events in the evolutionary history of these sequences. Entanglement can be quantified by summing the difference in the vertical positions of each label, such that a value of zero indicates that the trees are congruent. The resulting sum can also be normalized such that

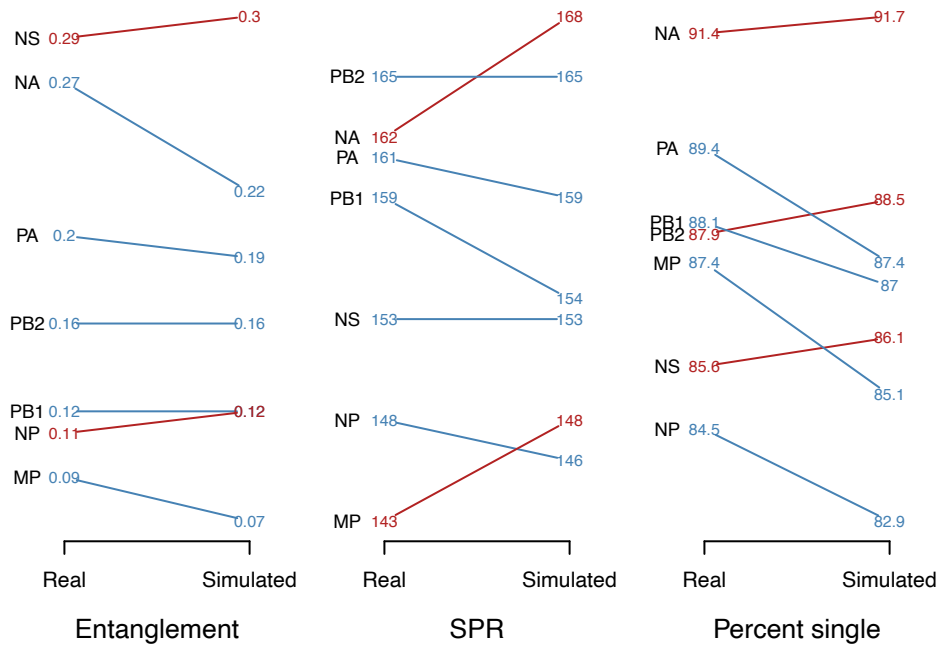

Figure 2: Slopegraphs summarizing the comparisons between phylogenies using entanglement (left), subtree prune and regraft (SPR) distance (middle), or the percentage (%) of inferred SPR rearrangements that involve a single terminal branch ('Percent single'). A reference phylogeny reconstructed from actual HA (segment 4) sequences was compared to phylogenies reconstructed from sequences representing the seven other segments represented either real sequences, or sequences that were simulated along the reference tree.

the value of the worst-case scenario is 1. For the example depicted in Figure 1A, the entanglement value was 0.27.

Although these results suggest that reassortment has occurred between these segments, we do not know how much should be attributed to inaccuracies in reconstructing the respective phylogenies, *i.e.*, false positives. Figure 1B displays a similar tanglegram between HA and NA phylogenies, with the exception that the NA sequences were simulated along the phylogeny of the actual HA sequences. Because the underlying phylogenies are congruent, these simulations should produce a null distribution of entanglement values in the absence of any reassortment. Instead, we observed similar amounts of entanglement compared to the real data across all segments (paired Wilcoxon test, $P = 0.34$; Figure 2).

Another method of quantifying the discordance between phylogenies is the minimum number of subtree-prune-regraft (SPR) rearrangements required to transform one tree into another. A useful characteristic of SPR is that this rearrangement is similar to the effect of reassortment on trees. Using the program Espalier, we calculated the SPR distances between the real HA tree and trees for the other segments with real or simulated data. These results are summarized in Figure 2. Overall, there was no significant difference in the inferred number of SPR rearrangements between real and simulated data sets (paired Wilcoxon test, $P = 0.89$). This implies that there is a high false-positive rate in using phylogenetic discordance to reconstruct reassortment events.

Given that the null distributions of entanglement and SPR distances are not distinguishable between real and simulated trees, we investigated whether one can at least differentiate between SPRs that represent actual reassortments and false positives. First, we evaluated the distributions in the size of subtrees involved in the inferred rearrangements in the MAF. Most SPRs involved the displacement of a single branch in the tree (87.4% of SPRs averaged across segments and datasets). Among segments, the percentages of SPRs involving a single branch were not significantly different between real and simulated datasets (paired Wilcoxon test, $P = 0.22$; Figure 2). Moreover, the full distributions of subtree sizes for SPRs between the real and simulated datasets were not markedly different (Supplementary Figure S2).

Next, we quantified the distance between the prune and regraft points for all inferred SPR rearrangements. This is not straight-forward because these points exist on separate trees, so we approximated the path length by locating the smallest subtree containing the same tip labels as the descendants of the sibling node in a second tree, where the sibling node shares the same immediate ancestor as the subtree affected by the SPR. We refer to this path length as the sibling distance for a given SPR. Figure 3 summarizes the distributions of sibling distances in the real HA tree across segments for real and simulated datasets. The overall mean sibling distance for SPRs was 0.13 expected substitutions per site, with an interquartile range of 0.045 to 0.208. We used a linear mixed effects regression model to evaluate the effects of subtree size and data type (real or

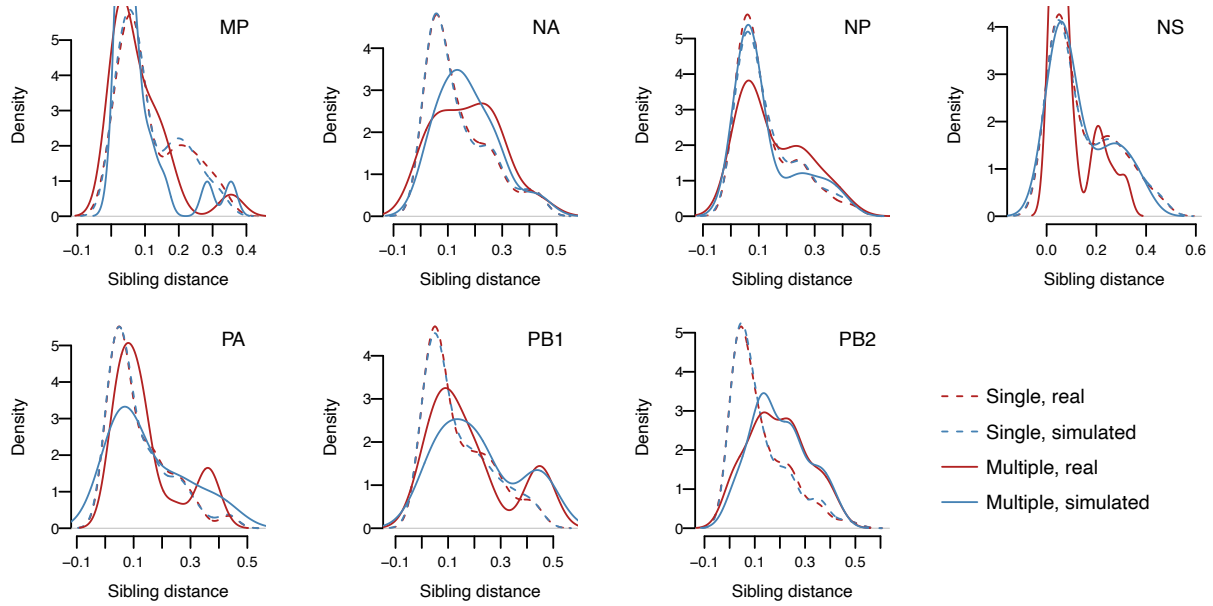

Figure 3: Comparison of the distance between prune and regraft locations in the reference HA tree for real (red) and simulated (blue) data for all other segments. This sibling distance was calculated by locating the sibling node of the subtree root, and then identifying the corresponding node in the other tree. Each line represents a kernel density summarizing the relative distribution of 14 to 154 distance measurements, partitioned by SPRs that involved a single branch ($n = 121 - 154$, dashed line) or multiple branches ($n = 14 - 25$, solid line).

simulated), with segment as a random effect. There was a marginally significant effect of subtree size on sibling distance (95% CI= -0.003 to $+0.014$) in that larger subtrees tended to move further in the tree. However, the sibling distances in real data were not significantly longer or shorter than the distances for SPRs reconstructed from simulated data. Specifically, there was no significant effect of data type on these distances.

Discussion

Reassortment is undoubtedly an important process in the evolution of influenza A viruses in the form of antigenic shift (McHardy and Adams 2009). Several epidemiologically significant strains such as the H1N1 strain associated with the 2009 swine flu pandemic have originated from reassortment combining segments from divergent lineages (Garten et al. 2009; Rambaut et al. 2008). In addition, a wide diversity of influenza A viruses circulate in numerous avian species (Causey and Edwards 2008). This natural host reservoir represents a significant risk for reassortment to generate new genomic combinations with increased pathogenicity and zoonotic potential (Graziosi et al. 2024; Guo et al. 2017; Pulit-Penaloza et al. 2020). For example, Gu et al. (2012) reported a IAV genome isolated from domestic ducks in China in 2012 that appeared to be a reassortment of H11N2, H7N1 and highly pathogenic avian H5Nx lineages.

Consequently there have been many studies using comparative methods to identify reassortment events in the evolutionary history of AIVs. For instance, Lewis et al. (2015) reported extensive reassortment from a tanglegram analysis of 211 complete genomes for a variety of HnNn subtypes of AIVs. Dugan et al. (2008) similarly observed a "remarkably high rate of genome reassortment" among 167 complete genomes of AIVs, using the difference in log-likelihoods of tree topologies reconstructed from different segments as a measure of discordance when fitting those trees to the same data. Although this method can be used to reject a null hypothesis of no reassortment, it does not identify specific reassortment events. Lu et al. (2014) used ancestral reconstruction of HA or NA subtypes as discrete traits associated with the other segments from about

3,000 AIV genomes to quantify a range of reassortment rates among subtypes and host species. Using a larger dataset of over 40,000 IAV genomes from a variety of host species, Gong et al. (2021) partitioned the trees for the genomic segments into 493-663 monophyletic clades, such that each genome could be represented by the clade assignments of its segments. They reported nearly 2,000 putative reassortments, where a genome was assumed to be a reassortant if its clade assignments could be produced from a combination of two or more other genomes.

Although these studies have employed different methods, they share a common premise on utilizing the phylogenetic discordance among genomic segments as evidence of reassortment. It has long been recognized that error in reconstructing phylogenies will also result in discordant trees in the absence of any reassortment (Svinti et al. 2013). Surprisingly, previous studies of reassortment in AIV genomes have seldom checked for these inevitable false positives. Using straight-forward simulation tests, we have found that the estimated prevalence of reassortment events in the absence of any reassortment is not significantly different from the number of events estimated from actual IAV data. Moreover, the subtrees involved in the putative reassortment events were indistinguishable between data simulated in the absence of any reassortment and actual sequence data.

This suggests that results from studies of phylogenetic discordance among influenza virus segments should be interpreted with an abundance of caution. Because the sequences of other segments were simulated along the tree reconstructed from actual HA (segment 4) sequences — rescaling the tree length to match the lengths obtained from actual data for every other segment — the most likely explanation for false positives is the uncertainty in reconstructing trees with limited data. For our analysis, we intentionally selected the $n = 206$ most divergent genome sequences by progressively removing the shortest terminal branches from an initial phylogeny of $n = 11,795$ HA sequences. Assuming that this subset of genomes with the most divergent HA sequences also tended to be highly divergent in the other segments, it is likely that our results underestimate the false positive rate by optimizing the sample for phylogenetic reconstruction. We

note that the rapid nucleotide substitution rates of AIV should provide a relative abundance of phylogenetically-informative sites in genomic data within a certain time scale. Identifying reassortment events for other viruses with segmented genomes and slower molecular clocks may be an even greater challenge (Varsani et al. 2018).

Limitations. In this study, we employed two methods to quantify the extent of reassortment in the evolutionary history of real and simulated data: tanglegrams and the minimum number of subtree-prune-regraft rearrangements. (*i.e.*, the SPR distance). While there are many other approaches to quantifying reassortment that have been described in the literature, we postulate that tangelgrams and SPR distances provide a representative selection of methods. Tanglegrams are pervasively used to visualize phylogenetic discordance (Bansal et al. 2009). An SPR rearrangement is very similar to the effect of reassortment on phylogenies (Svinti et al. 2013; Lin et al. 2024); reconstructing SPRs is arguably the most direct measure of reassortments that is available to us. On the other hand, both methods have their own limitations.

Tanglegrams can be misleading because the number of crossings among line segments connecting the same labels is not actually proportional to topological congruence. Trees with very different topologies can have a relatively low number of crossings, while trees with similar levels of discordance may have different levels of entanglement (Briand et al. 2020). For instance, De Vienne (2019) noted that although larger Robinson-Foulds (RF) distances tend to be associated with higher degrees of entanglement, there is actually a large overlap of degrees of entanglement over a range of topological distances. In our study, pairs of trees with higher entanglement values tended to also have higher SPR distances (Pearson's $r = 0.42$, 95% CI $= -0.14 - 0.78$). Nevertheless, we should be cautious about interpreting differences in entanglement between trees.

A number of programs have been implemented to measure reassortment from phylogenetic discordance (Suzuki 2010). However, only a few of these programs remain available, particularly for inferring the minimum number of SPR rearrangements between trees. SPRSupertrees (Whidden et al. 2014) is a C++ program distributed under the GNU GPLv3 license that calculates the SPR

distance, but its performance decreases for phylogenetic trees with hundreds of branches. In addition, the website from which this program is distributed is not consistently accessible (last attempt February 19, 2025). Espalier (Rasmussen and Guo 2023) uses a heuristic algorithm to compute the SPR distance in a more computationally efficient manner, retaining only strongly supported rearrangements. However, the algorithm employed by Espalier is approximate and can produce distances that are not symmetric, *i.e.*, $d(A,B) \neq d(B,A)$.

Lin et al. (2024) recently described several other issues in estimating the number and distribution of reassortment events from discordant segment trees. First, they noted that reassortment events may alter branch lengths without modifying the topology of the trees. Reassortments that involve an unsampled lineage cannot be reconstructed by comparing trees. In addition, comparing discordant trees can be affected by a 'multiple hits' problem in which a more recent reassortment event obscures other events that precede it. These issues have the opposite effect compared to phylogenetic uncertainty: they would cause the estimated number of SPR rearrangements to underestimate the actual extent of reassortment. Consequently we postulate that measuring the overall rate of reassortment or reconstructing a number of these events by the large-scale comparative analysis of genomes should be considered an open problem for influenza viruses. Fortunately, there has been a resurgence of interest in relevant areas of phylogenetics such as ancestral recombination graphs (Rasmussen and Guo 2023; Barrat-Charlaix et al. 2022; Lewanski et al. 2024).

Acknowledgements

We thank David Rasmussen for providing advice and insight into the algorithms employed by Espalier.

References

Bansal, Mukul S et al. (2009). "Generalized binary tanglegrams: Algorithms and applications". In: *Bioinformatics and Computational Biology: First International Conference, BICoB 2009, New Orleans, LA, USA, April 8-10, 2009. Proceedings*. Springer, pp. 114–125.

Barrat-Charlaix, Pierre, Timothy G Vaughan, and Richard A Neher (2022). "TreeKnit: Inferring ancestral reassortment graphs of influenza viruses". In: *PLoS Computational Biology* 18.8, e1010394.

Bouvier, Nicole M and Peter Palese (2008). "The biology of influenza viruses". In: *Vaccine* 26, pp. D49–D53.

Briand, Samuel et al. (2020). "A generalized Robinson-Foulds distance for labeled trees". In: *BMC genomics* 21, pp. 1–13.

Capella-Gutiérrez, Salvador, José M Silla-Martínez, and Toni Gabaldón (2009). "trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses". In: *Bioinformatics* 25.15, pp. 1972–1973.

Causey, Douglas and Scott V Edwards (2008). "Ecology of avian influenza virus in birds". In: *The Journal of Infectious Diseases* 197.Supplement_1, S29–S33.

Collienne, Lena and Alex Gavryushkin (2021). "Computing nearest neighbour interchange distances between ranked phylogenetic trees". In: *Journal of Mathematical Biology* 82.1, p. 8.

De Vienne, Damien M (2019). "Tanglegrams are misleading for visual evaluation of tree congruence". In: *Molecular biology and evolution* 36.1, pp. 174–176.

Dugan, Vivien G et al. (2008). "The evolutionary genetics and emergence of avian influenza viruses in wild birds". In: *PLoS pathogens* 4.5, e1000076.

Frost, Simon DW et al. (2015). "Eight challenges in phylodynamic inference". In: *Epidemics* 10, pp. 88–92.

Garten, Rebecca J et al. (2009). "Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans". In: *science* 325.5937, pp. 197–201.

Gong, Xingfei et al. (2021). “Reassortment network of influenza A virus”. In: *Frontiers in Microbiology* 12, p. 793500.

Graziosi, Giulia et al. (2024). “Highly pathogenic avian influenza (HPAI) H5 clade 2.3. 4.4 b virus infection in birds and mammals”. In: *Animals* 14.9, p. 1372.

Gu, Min et al. (2012). *Genome sequence of a natural reassortant H5N2 avian influenza virus from domestic mallard ducks in eastern China*.

Guo, Hongbo et al. (2017). “Highly pathogenic influenza A (H5Nx) viruses with altered H5 receptor-binding specificity”. In: *Emerging infectious diseases* 23.2, p. 220.

Katoh, Kazutaka and Daron M Standley (2013). “MAFFT multiple sequence alignment software version 7: improvements in performance and usability”. In: *Molecular biology and evolution* 30.4, pp. 772–780.

Kelk, Steven and Simone Linz (2020). “New reduction rules for the tree bisection and reconnection distance”. In: *Annals of Combinatorics* 24.3, pp. 475–502.

Kong, Weili et al. (2015). “Novel reassortant influenza viruses between pandemic (H1N1) 2009 and other influenza viruses pose a risk to public health”. In: *Microbial pathogenesis* 89, pp. 62–72.

Lewanski, Alexander L, Michael C Grundler, and Gideon S Bradburd (2024). “The era of the ARG: An introduction to ancestral recombination graphs and their significance in empirical evolutionary genomics”. In: *PLoS Genetics* 20.1, e1011110.

Lewis, Nicola S et al. (2015). “Influenza A virus evolution and spatio-temporal dynamics in Eurasian wild birds: a phylogenetic and phylogeographical study of whole-genome sequence data”. In: *Journal of General Virology* 96.8, pp. 2050–2060.

Li, Chong et al. (2024). “Naturally occurring influenza reassortment in pigs facilitates the emergence of intrahost virus subpopulations with distinct genotypes and replicative fitness”. In: *mBio*, e01924–24.

Lin, Qianying et al. (2024). “The number and pattern of viral genomic reassortments are not necessarily identifiable from segment trees”. In: *Molecular Biology and Evolution* 41.6, msae078.

Lu, Lu, Samantha J Lycett, and Andrew J Leigh Brown (2014). “Reassortment patterns of avian influenza virus internal segments among different subtypes”. In: *BMC evolutionary biology* 14, pp. 1–15.

Ma, Eric J et al. (2016). “Reticulate evolution is favored in influenza niche switching”. In: *Proceedings of the National Academy of Sciences* 113.19, pp. 5335–5339.

McDonald, Sarah M et al. (2016). “Reassortment in segmented RNA viruses: mechanisms and outcomes”. In: *Nature Reviews Microbiology* 14.7, pp. 448–460.

McHardy, Alice Carolyn and Ben Adams (2009). “The role of genomics in tracking the evolution of influenza A virus”. In: *PLoS pathogens* 5.10, e1000566.

Minh, Bui Quang et al. (2020). “IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era”. In: *Molecular biology and evolution* 37.5, pp. 1530–1534.

Nguyen, Nghia et al. (2022). “Shuffle & untangle: novel untangle methods for solving the tanglegram layout problem”. In: *Bioinformatics Advances* 2.1, vbac014.

Otto, Sarah Perin and Nick H Barton (1997). “The evolution of recombination: removing the limits to natural selection”. In: *Genetics* 147.2, pp. 879–906.

Page, Roderic DM (2003). *Tangled trees: phylogeny, cospeciation, and coevolution*. University of Chicago press.

Pérez-Losada, Marcos et al. (2015). “Recombination in viruses: mechanisms, methods of study, and evolutionary consequences”. In: *Infection, Genetics and Evolution* 30, pp. 296–307.

Pinto, Rute M et al. (2021). “Accessory gene products of influenza A virus”. In: *Cold Spring Harbor perspectives in medicine* 11.12, a038380.

Pulit-Penaloza, Joanna A et al. (2020). “Characterization of highly pathogenic avian influenza H5Nx viruses in the ferret model”. In: *Scientific reports* 10.1, p. 12700.

Rambaut, Andrew et al. (2008). “The genomic and epidemiological dynamics of human influenza A virus”. In: *Nature* 453.7195, pp. 615–619.

Rasmussen, David A and Fangfang Guo (2023). “Espalier: Efficient Tree Reconciliation and Ancestral Recombination Graphs Reconstruction Using Maximum Agreement Forests”. In: *Systematic Biology* 72.5, pp. 1154–1170.

Ryt-Hansen, Pia et al. (2021). “Co-circulation of multiple influenza A reassortants in swine harboring genes from seasonal human and swine influenza viruses”. In: *Elife* 10, e60940.

Scornavacca, C, F Zickmann, and DH Huson (2011). *Tanglegrams for rooted phylogenetic trees and networks. BioInformatics 27: i248–i256*.

Silva, U Chandimal de et al. (2012). “A comprehensive analysis of reassortment in influenza A virus”. In: *Biology open* 1.4, pp. 385–390.

Simon-Loriere, Etienne and Edward C Holmes (2011). “Why do RNA viruses recombine?” In: *Nature Reviews Microbiology* 9.8, pp. 617–626.

Spielman, Stephanie J and Claus O Wilke (2015). “Pyvolve: a flexible Python module for simulating sequences along phylogenies”. In: *PloS one* 10.9, e0139047.

Sutton, Troy C (2018). “The pandemic threat of emerging H5 and H7 avian influenza viruses”. In: *Viruses* 10.9, p. 461.

Suzuki, Yoshiyuki (2010). “A phylogenetic approach to detecting reassortments in viruses with segmented genomes”. In: *Gene* 464.1-2, pp. 11–16.

Svinti, Victoria, James A Cotton, and James O McInerney (2013). “New approaches for unravelling reassortment pathways”. In: *BMC Evolutionary Biology* 13, pp. 1–14.

Taylor, Kishana Y et al. (2023). “Influenza A virus reassortment is strain dependent”. In: *PLoS pathogens* 19.3, e1011155.

Varsani, Arvind et al. (2018). “Notes on recombination and reassortment in multipartite/segmented viruses”. In: *Current opinion in virology* 33, pp. 156–166.

Vijaykrishna, Dhanasekaran, Reshmi Mukerji, and Gavin JD Smith (2015). “RNA virus reassortment: an evolutionary mechanism for host jumps and immune evasion”. In: *PLoS pathogens* 11.7, e1004902.

Wang, Haiwei et al. (2022). “Recombination in positive-strand RNA viruses”. In: *Frontiers in Microbiology* 13, p. 870759.

Whidden, Christopher, Norbert Zeh, and Robert G Beiko (2014). “Supertrees based on the subtree prune-and-regraft distance”. In: *Systematic biology* 63.4, pp. 566–581.

Wong, Yan et al. (2024). “A general and efficient representation of ancestral recombination graphs”. In: *Genetics* 228.1, iyae100.

World Health Organization (1980). “A revision of the system of nomenclature for influenza viruses: a WHO memorandum”. In: *Bulletin of the World Health Organization* 58.4, pp. 585–591.

Wu, Yufeng (2009). “A practical method for exact computation of subtree prune and regraft distance”. In: *Bioinformatics* 25.2, pp. 190–196.

Supplementary Figures

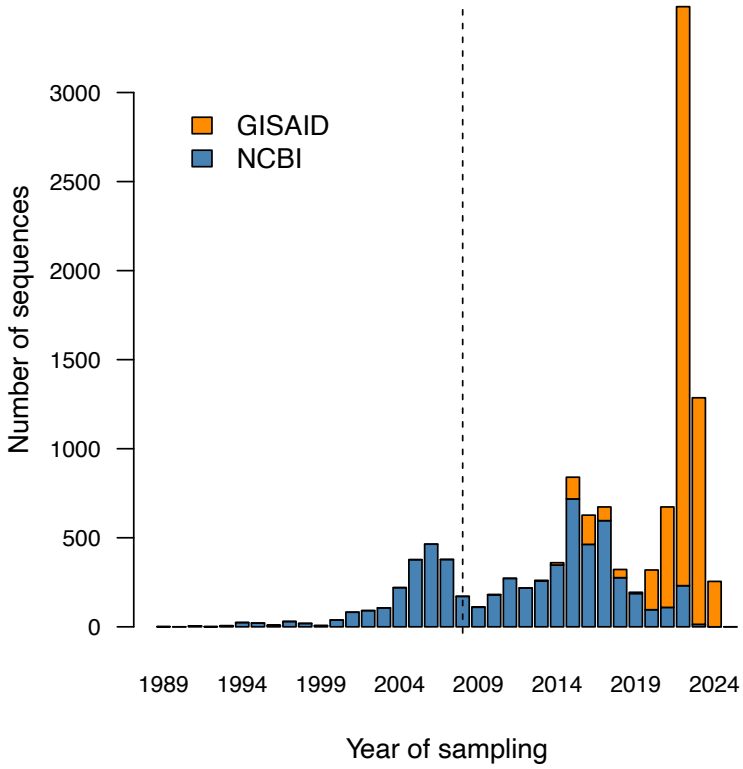

Figure S1: Distributions of sample collection dates for hemagglutinin (HA) sequences from the NCBI Genbank and GISAID databases.

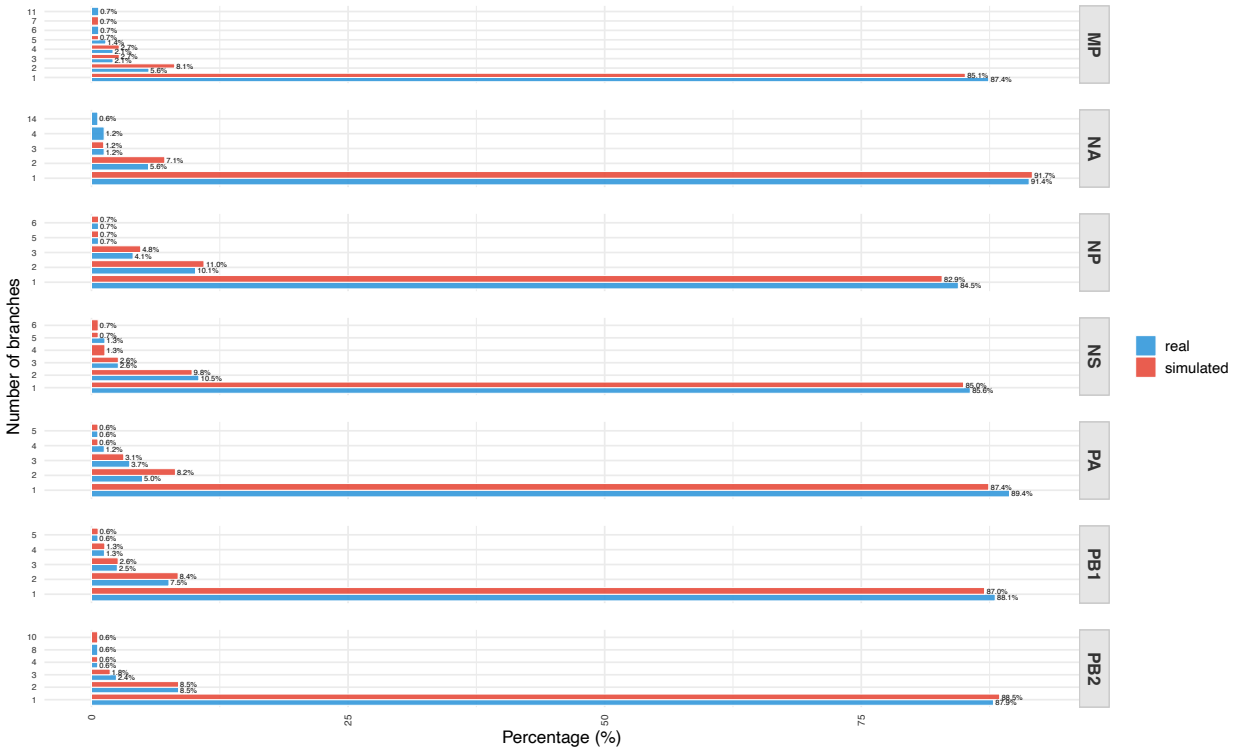

Figure S2: Barplot showing the percentage of frequency of SPR events in phylogenetic trees evaluated from subtrees obtained from MAF. The data show that most SPR events occur in a single branch or in isolated branches, with over 80% of events affecting only one branch for all gene segments. The real data generally show slightly higher percentages compared to the simulated data, suggesting that topological changes are restricted to single branches rather than extensive reassortment.