



## OPEN ACCESS

## EDITED BY

Jim Procter,  
University of Dundee, United Kingdom

## REVIEWED BY

Raghavendran Partha,  
Independent researcher, Menlo Park,  
United States  
Sam Wilks,  
University of Cambridge, United Kingdom

## \*CORRESPONDENCE

John Huddleston,  
✉ jhuddles@fredhutch.org

<sup>†</sup>These authors have contributed equally  
to this work and share first authorship

## SPECIALTY SECTION

This article was submitted  
to Data Visualization,  
a section of the journal  
Frontiers in Bioinformatics

RECEIVED 13 October 2022

ACCEPTED 08 March 2023

PUBLISHED 22 March 2023

## CITATION

Lee J, Hadfield J, Black A, Sibley TR,  
Neher RA, Bedford T and Huddleston J  
(2023), Joint visualization of seasonal  
influenza serology and phylogeny to  
inform vaccine composition.  
*Front. Bioinform.* 3:1069487.  
doi: 10.3389/fbinf.2023.1069487

## COPYRIGHT

© 2023 Lee, Hadfield, Black, Sibley,  
Neher, Bedford and Huddleston. This is  
an open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Joint visualization of seasonal influenza serology and phylogeny to inform vaccine composition

Jover Lee<sup>1†</sup>, James Hadfield<sup>1†</sup>, Allison Black<sup>2</sup>, Thomas R. Sibley<sup>1</sup>,  
Richard A. Neher<sup>3,4</sup>, Trevor Bedford<sup>1,5</sup> and John Huddleston<sup>1\*</sup>

<sup>1</sup>Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, WA, United States, <sup>2</sup>Chan Zuckerberg Initiative, San Francisco, CA, United States, <sup>3</sup>Biozentrum, Universität Basel, Basel, Switzerland, <sup>4</sup>Swiss Institute of Bioinformatics, Lausanne, Switzerland, <sup>5</sup>Howard Hughes Medical Institute, Seattle, WA, United States

Seasonal influenza vaccines must be updated regularly to account for mutations that allow influenza viruses to escape our existing immunity. A successful vaccine should represent the genetic diversity of recently circulating viruses and induce antibodies that effectively prevent infection by those recent viruses. Thus, linking the genetic composition of circulating viruses and the serological experimental results measuring antibody efficacy is crucial to the vaccine design decision. Historically, genetic and serological data have been presented separately in the form of static visualizations of phylogenetic trees and tabular serological results to identify vaccine candidates. To simplify this decision-making process, we have created an interactive tool for visualizing serological data that has been integrated into Nextstrain's real-time phylogenetic visualization framework, Auspice. We show how the combined interactive visualizations may be used by decision makers to explore the relationships between complex data sets for both prospective vaccine virus selection and retrospectively exploring the performance of vaccine viruses.

## KEYWORDS

influenza, serology, phylogeny, interactive, visualization, pandemic

## 1 Introduction

Seasonal influenza A/H3N2 viruses primarily evolve by acquiring mutations that allow them to escape antibodies from previous infections or vaccinations (Petrova and Russell, 2018). This process, known as antigenic drift, changes the appearance of viral surface proteins hemagglutinin and neuraminidase. Hemagglutinin (HA) is the primary target of our adaptive immunity and the primary component of the seasonal influenza vaccine. Therefore, continued antigenic drift in HA necessitates regular updates to the seasonal influenza vaccine.

The World Health Organization (WHO) Global Influenza Surveillance and Response System (GISRS) tracks antigenic drift throughout the year by sequencing the HA gene of circulating viruses, growing candidate vaccine viruses in cell lines and chicken eggs, and performing serological experiments (Morris et al., 2017). HA gene sequences reveal "clades" or groups of recent influenza viruses that descend from a common ancestor. When the common ancestor of a clade carries amino acid mutations at positions in HA that have been previously targeted by antibodies following infection or vaccination (Wolf et al., 2006; Shih et al., 2007; Koel et al., 2013), viruses in that clade may be able to escape our existing immunity. GISRS researchers select representative viruses from major clades to grow in the

cell line and chicken egg environments used in vaccine production (Katz et al., 2011). Viruses that grow in these conditions become candidate vaccine viruses or vaccine candidates. Serological experiments measure antigenic drift by quantifying how well viruses from one clade can escape detection by antibodies against vaccine candidates from the same clade or other clades. These experimental measurements validate the effects of specific mutations on antigenic drift.

The gold standard of these serological experiments is the hemagglutination inhibition (HI) assay (Hirst, 1943). HI assays measure the minimum concentration or “titer” of antibodies required to prevent a given test virus from binding to and, thereby, agglutinating red blood cells. Typically, antibodies come from previously uninfected ferrets who are then infected with a specific reference virus (e.g., a vaccine candidate). Each HI assay requires a series of two-fold dilutions of ferret antibodies to determine the minimum titer that inhibits agglutination. The higher the antibody titer required to prevent agglutination, the more antigenically drifted the test virus is from the reference virus. To enable comparison between measurements from different reference viruses, researchers normalize titers by the titer required to inhibit the reference virus itself and convert values to a  $\log_2$  scale (Neher et al., 2016). The resulting values represent the magnitude of antigenic distance between a given reference virus and other test viruses. Traditionally, viruses with an antigenic distance greater than  $2\log_2$  units are considered antigenically distinct (Katz et al., 2011).

The WHO convenes influenza vaccine composition meetings (VCMs) twice per year with each meeting occurring approximately 9 months prior to the next influenza season in a given hemisphere (Morris et al., 2017). At these meetings, WHO decision makers must select a single vaccine virus for H3N2 from the pool of available vaccine candidates. This selection strongly depends on the patterns of antigenic drift present in HI titers for clades identified from HA gene sequences. Key questions that decision makers must answer in this process include a) which available vaccine candidates require additional titer measurements against currently circulating clades, b) which vaccine candidates have the lowest antigenic distance to each clade, and c) what is the antigenic diversity of recent clades? The best vaccine candidate will have titer measurements against all current clades and have the lowest antigenic distance across these clades. Such candidates are said to effectively “cover” currently circulating clades. Determining the vaccine candidate that most effectively covers recent clades requires a direct comparison of antigenic distances across all available vaccine candidates.

Historically, WHO decision makers have answered the key questions above using standard phylogenetic visualizations from HA gene sequences (Felsenstein, 2003; Lemey et al., 2012) and separate visualizations of antigenic evolution from HI assays (Smith et al., 2004). Visualization of antigenic evolution in a phylogenetic context is a relatively recent development (Steinbrück and McHardy, 2012; Bedford et al., 2014; Neher et al., 2016). We previously developed two alternate visualizations of antigenic data to inform vaccine selection through reports to the WHO (Bedford and Neher, 2018; Bedford et al., 2019). The first visualization is an interactive phylogenetic view implemented in the genomic epidemiology tool *nextflu* (Neher and Bedford, 2015; 2018). When the user selects a given reference virus from the

phylogeny (by selecting a “gear” icon), the tool plots the pairwise antigenic distances between that reference and the corresponding test viruses in the tree using color to represent the distance values (Figure 1A). This representation reveals clades that the selected reference virus may or may not cover and which clades lack measurements against the reference. While this view provides phylogenetic context of titer measurements and shows the individual pairwise measurements, it does not support direct comparison of antigenic distances from different vaccine candidates. Instead, users must quickly toggle between different vaccine candidates in the tree to get a sense of how well different viruses may perform.

To complement the interactive phylogenetic visualization, we also developed a static heatmap visualization that summarizes the mean antigenic distances between a subset of vaccine candidates and all test viruses within each extant clade. The resulting heatmaps use the  $x$ -axis to encode the names of major clades, the  $y$ -axis to encode the names of reference viruses, and color and text to encode the mean  $\log_2$  distance between a given reference virus and corresponding test viruses in each clade (Figure 1B). These heatmaps allow decision makers to directly compare average antigenic distances for different vaccine candidates and quantify how well these candidates cover extant clades. However, these heatmaps suffer reduced expressiveness by encoding the most valuable quantitative data with color instead of a positional encoding. Additionally, this view shows a summary statistic instead of the underlying distributions of the data, concealing the number and variance of measurements for each reference virus.

To overcome the limitations of existing antigenic visualizations, we applied user-driven design based on the goals of decision makers described above and used standard visual design principles to produce a more expressive and effective visualization for influenza virologists. The result is a new component of Nextstrain’s interactive phylogenetic visualization platform that we call the measurements panel. Below, we describe the measurements panel and provide two case studies that demonstrate the practical value of this tool for vaccine composition decisions with the interactive visualizations available at [nextstrain.org/community/blab/measurements-panel/flu/seasonal/h3n2/ha](https://nextstrain.org/community/blab/measurements-panel/flu/seasonal/h3n2/ha).

## 2 Methods

### 2.1 Visual design

Between the interactive phylogenetic view in *nextflu* (Figure 1A) and the static heatmap visualization used in WHO reports (Figure 1B), the heatmaps addressed the most user goals. The only original goal that heatmaps could not address was the communication of the number of measurements available for each clade that support the summary statistic of mean titer distance. Additionally, both of these prior visualizations either obscured or hid the underlying distributions of the raw data. Since visualization of these distributions can improve confidence during the decision-making process (Correll and Gleicher, 2014; Hullman et al., 2015; Fernandes et al., 2018), we treated the need to view these distributions as an auxiliary user goal.

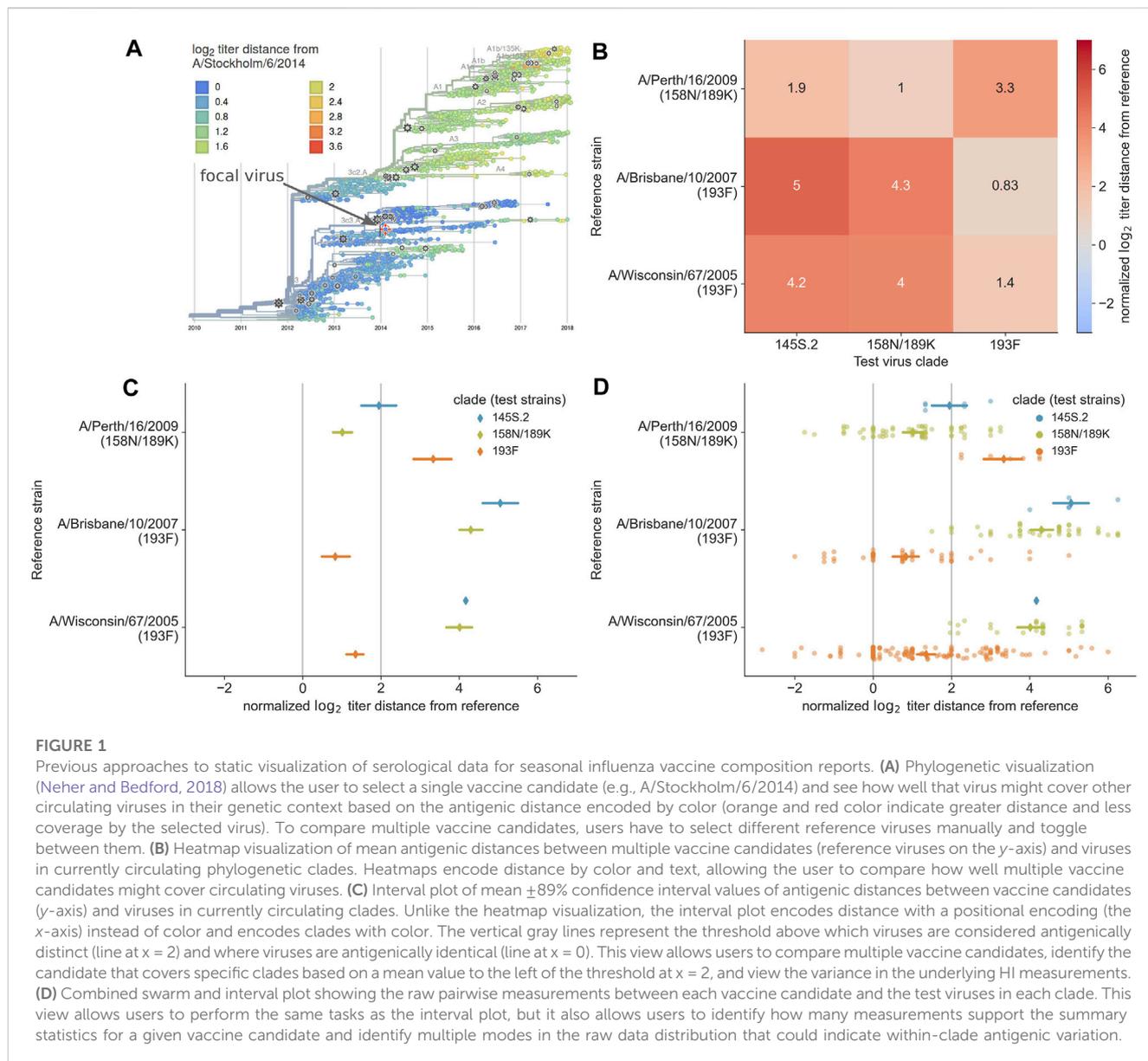


FIGURE 1

Previous approaches to static visualization of serological data for seasonal influenza vaccine composition reports. **(A)** Phylogenetic visualization (Neher and Bedford, 2018) allows the user to select a single vaccine candidate (e.g., A/Stockholm/6/2014) and see how well that virus might cover other circulating viruses in their genetic context based on the antigenic distance between them. To compare multiple vaccine candidates, users have to select different reference viruses manually and toggle between them. **(B)** Heatmap visualization of mean antigenic distances between multiple vaccine candidates (reference viruses on the y-axis) and viruses in currently circulating phylogenetic clades. Heatmaps encode distance by color and text, allowing the user to compare how well multiple vaccine candidates might cover circulating viruses. **(C)** Interval plot of mean  $\pm 89\%$  confidence interval values of antigenic distances between vaccine candidates (y-axis) and viruses in currently circulating clades. Unlike the heatmap visualization, the interval plot encodes distance with a positional encoding (the x-axis) instead of color and encodes clades with color. The vertical gray lines represent the threshold above which viruses are considered antigenically distinct (line at  $x = 2$ ) and where viruses are antigenically identical (line at  $x = 0$ ). This view allows users to compare multiple vaccine candidates, identify the candidate that covers specific clades based on a mean value to the left of the threshold at  $x = 2$ , and view the variance in the underlying HI measurements. **(D)** Combined swarm and interval plot showing the raw pairwise measurements between each vaccine candidate and the test viruses in each clade. This view allows users to perform the same tasks as the interval plot, but it also allows users to identify how many measurements support the summary statistics for a given vaccine candidate and identify multiple modes in the raw data distribution that could indicate within-clade antigenic variation.

In the context of visual design principles, both the phylogenetic and heatmap views encode the most relevant quantitative data of antigenic distance with color. However, quantitative data can be more effectively represented by positional encodings (e.g., x- or y-axis positions) whereas nominal data (e.g., names of phylogenetic clades) can be effectively encoded with color (Mackinlay, 1986). In *nextflu's* phylogenetic view, the two available positional axes represent time and the unitless phylogenetic position of nodes, neither of which are relevant to the user goals described above. In the heatmaps, the two positional axes encode two nominal data types.

We reasoned that we could make a more effective visualization that addressed all user goals by simply changing the encoding of data in the heatmaps. To this end, we swapped the encoding of antigenic distances and test clades, encoding quantitative distances on the x-axis positional scale and encoding nominal test clades with a color scale. The positional encoding of antigenic distances allowed us to

visually encode relevant thresholds for decision-making (e.g.,  $x = 2 \log_2$ ), show all available measurements for each reference virus at once, and display a summary statistic (mean and standard deviation of antigenic distances) for each reference virus. We retained the encoding of nominal reference virus names on the y-axis, since most user goals require comparison of distances between specific vaccine candidates.

## 2.2 Implementation

We implemented this design as a new interactive measurements panel within Nextstrain's visualization tool, Auspice (version 2.43.0), which is freely available on GitHub ([github.com/nextstrain/auspice](https://github.com/nextstrain/auspice)) under the AGPLv3 license. Auspice is a phylogenetic visualization platform inspired by *nextflu* and which maintains the interactive data exploration, with the measurements

panel appearing alongside and in-sync with other views into the data (currently phylogenetic, geographic and genomic diversity views). The visualization requires a minimum of two JavaScript Object Notation (JSON) files that are produced by the Nextstrain bioinformatics toolkit, Augur (Huddleston et al., 2021). The phylogenetic tree is provided *via* a dataset JSON file produced by `augur export v2` and the measurements data is provided *via* a measurements sidecar JSON file produced by `augur measurements`. The measurements file must follow a specific filename format for Auspice to link it to the dataset file, where the dataset filename is `_${name}.json` and the measurements filename must be `_${name}_measurements.json`. The measurements sidecar JSON is expected to have an array of collections, where each collection contains its display configurations for Auspice and an array of measurements. Each measurement must include a numeric value to plot along the *x*-axis and a `strain` name that exactly matches a sequence name in the phylogenetic tree to support interactivity between panels. For complete technical details about the data structure used by the measurements panel, see the [measurements sidecar JSON schema document](#). The application can be cloned and run locally or users can use Auspice through two public websites. Users can drag and drop dataset and measurements files onto [auspice.us/](https://auspice.us/) to visualize the data locally in their own browser. Visualizations can be also shared with others through [nextstrain.org/](https://nextstrain.org/). Full documentation for sharing analyses through Nextstrain can be found at [docs.nextstrain.org/page/guides/share/](https://docs.nextstrain.org/page/guides/share/).

## 2.3 Data curation and analysis

We evaluated the new measurements panel by constructing a Nextstrain analysis (Hadfield et al., 2018) with previously published HA gene sequences and HI titers (Bedford et al., 2014). A full data curation guide is available online at [github.com/nextstrain/measurements-panel/tree/main/data#readme](https://github.com/nextstrain/measurements-panel/tree/main/data#readme) and a full guide to running the bioinformatics analyses is available at [github.com/nextstrain/measurements-panel#readme](https://github.com/nextstrain/measurements-panel#readme). Briefly, we downloaded HI titer data and accessions for H3N2 HA sequences from Bedford et al. (2014)'s GitHub repository. We downloaded and combined HA sequences from the Influenza Virus Resource or GISAID, depending on the original source. We parsed metadata including viral sample name, database accession, collection date, and sequence authors from the sequence headers with `augur parse`. We aligned HA sequences with `mafft v7.508` (Katoh and Standley, 2013), inferred a phylogenetic tree with `IQ-TREE 2.2.0.3` (Minh et al., 2020), and inferred a time tree with `TreeTime 0.9.4` (Sagulenko et al., 2018). We annotated mutations on the phylogenetic tree and constructed the measurements panel data JSON with `Augur 21.0.0` (Huddleston et al., 2021). For improved reproducibility, we automated the execution of these tasks in a Snakemake workflow (Mölder et al., 2021).

In the absence of official WHO clade designations for the time period of this analysis, we algorithmically assigned realistic clade labels to each internal node of the HA phylogeny that a) carried a mutation at one or more of seven previously identified sites associated with antigenic drift (Koel et al., 2013) and b) circulated at or above 10% global frequency at some point during

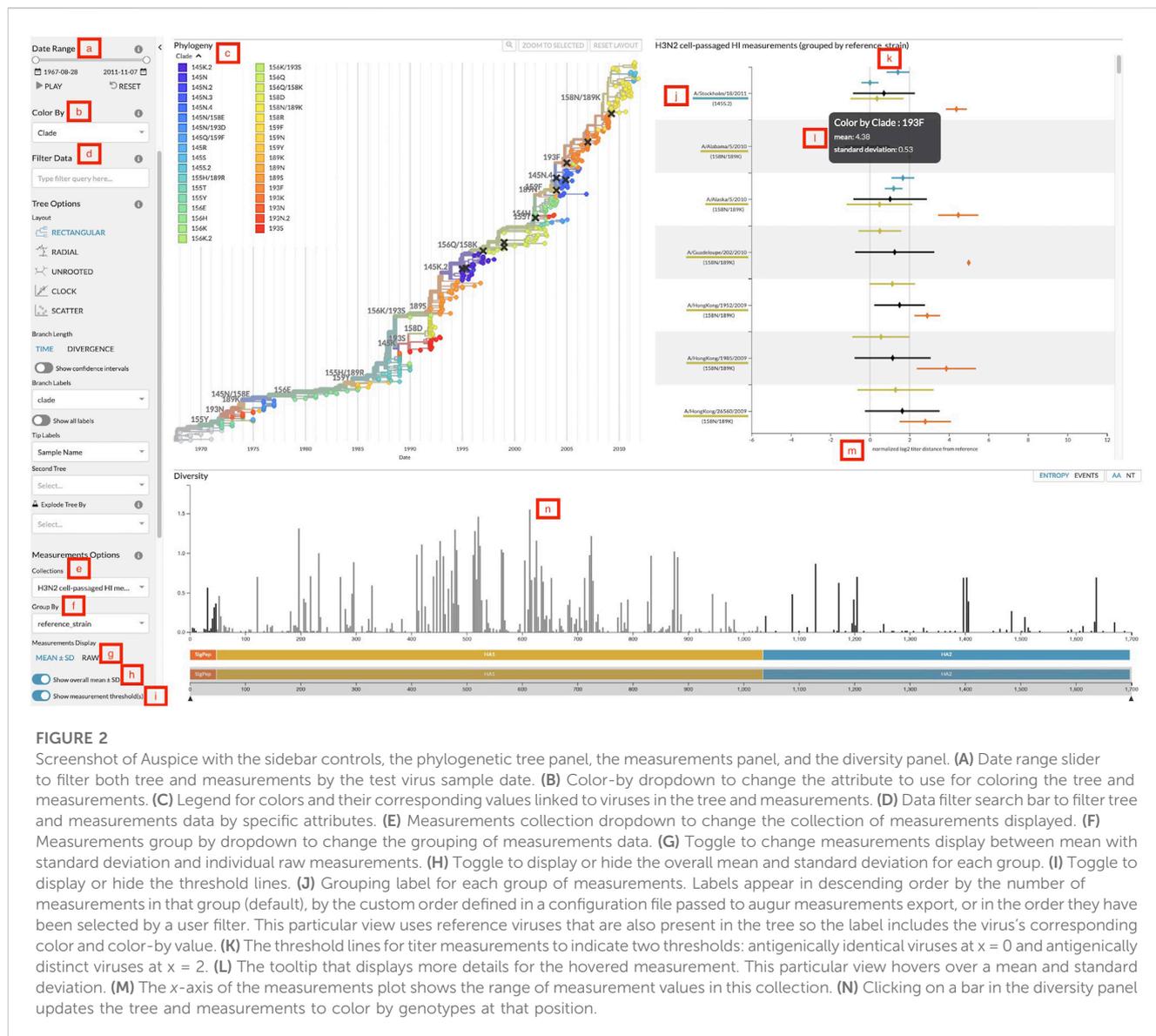
its existence. This algorithm mimics the decision-making approach used to assign clade labels in the last decade based on genetic data and prior knowledge about potentially relevant antigenic sites.

To construct the measurements panel JSON, we first normalized the raw HI titer measurements with Augur's implementation of the titer substitution model (Neher et al., 2016). Normalization  $\log_2$ -transforms raw titer measurements and subtracts the transformed measurement for test virus *a* and reference serum  $\beta$  raised against reference virus *b* from the corresponding measurement for the reference virus *b* and its serum  $\beta$ . This normalization produces an antigenic distance between test and reference viruses that we can compare across HI experiments. Next, we converted these distances to a tab-separated values (TSV) file with a custom Python script (`scripts/get_antigenic_distances_between_strains.py`) and ran the new `augur measurements export` command with this TSV file as input. Using a collection configuration file passed to the `augur measurements export` command, we sorted grouping labels in the measurements JSON file for reference strain, reference clade, and serum by descending order of each clade's minimum *y*-axis position in the phylogeny and then by ascending alphabetical order of reference strain name within each clade. This sorting causes reference strains and other grouping labels to appear in the measurements panel in the same order each reference clade first appears in the phylogeny, keeping closely related clades adjacent to each other in the panel. We sorted the grouping labels for each measurement's "source" by the default which is in descending order by the number of measurements in the grouping. The resulting visualization for this paper has been shared *via* Nextstrain community and can be viewed at [nextstrain.org/community/nextstrain/measurements-panel/flu/seasonal/h3n2/ha](https://nextstrain.org/community/nextstrain/measurements-panel/flu/seasonal/h3n2/ha).

## 3 Results

### 3.1 Interactive visualization of titer measurements

Our goal for incorporating the measurements panel into Auspice is to allow users to explore relationships between the genetic and serological data in one interactive visualization. Test viruses of the titer data are directly linked by name to the viruses displayed in the phylogenetic tree, to ensure that any interactions with the tree also affect the measurements displayed. When users filter the tree by date (Figure 2A), metadata attributes of viruses (Figure 2D), or subtrees selected by clicking on a corresponding tree branch, the measurements panel updates to reflect only the matching test viruses. Users can easily focus on a subset of phylogenetically relevant measurements and examine the measurements of test viruses in recently circulating clades. The measurements are colored by the same coloring attribute (Figure 2B) as the phylogenetic tree, adding the legend values (Figure 2C) as another dimension of nominal data to the titer measurements. This coloring is especially useful for viewing titer data by genotypes of test viruses, allowing users to inspect relationships between specific mutations in HA and antigenic drift quantified by titer measurements. The investigation of genotypes is further facilitated by the diversity panel in Auspice, which shows the diversity of alleles across the genome. Clicking on a bar in the

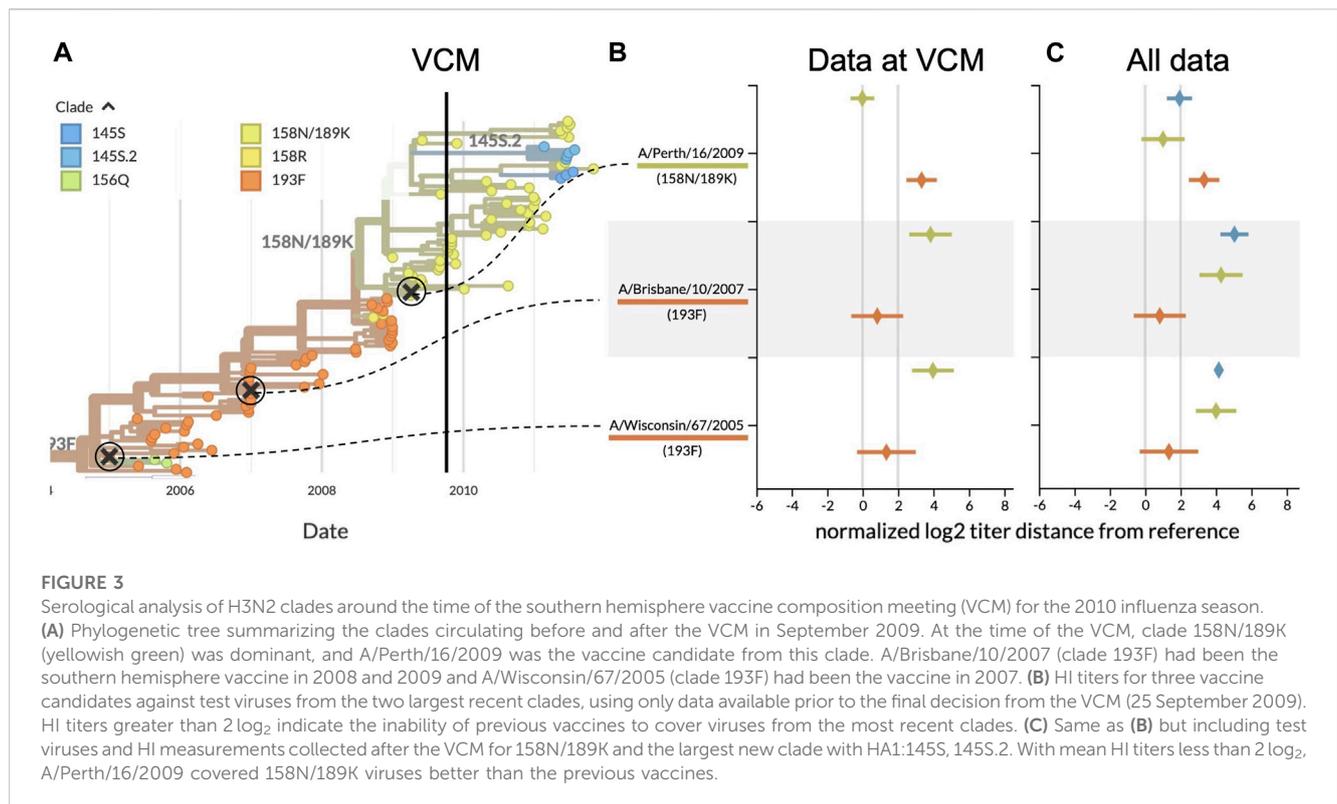


diversity panel will change the coloring of the tree and measurements to the genotypes at that position (Figure 2N).

The measurements JSON can include multiple collections of titers for a single phylogenetic tree and users can change the collection displayed with the collection dropdown (Figure 2E). Users can then review different sets of data for the same phylogenetic tree such as separate measurements of cell- and egg-passaged virus titers. Within each collection, users can compare measurements across different groupings by changing the grouping category with the “group by” dropdown (Figure 2F). For vaccine selection, grouping by the reference virus (Figure 2J) allows decision makers to directly compare titers across multiple vaccine candidates. Other groupings such as data source or ferret serum can be used to explore the variability of the titer measurements. By default, groupings appear in descending order by the number of measurements in each group. Users can specify a custom order for grouping values with a configuration file passed to `augur measurements export`. Users can also manually assign an order for specific groupings in

Auspice by filtering measurements to the corresponding grouping values (Figure 2D). Groupings will appear on the y-axis in the order that the user selects them from the filter field. The overall mean for each grouping can then be toggled (Figure 2H) for whole group comparisons.

The data display can be switched between mean with standard deviation and raw individual measurements (Figure 2G). The means are calculated per color attribute to allow for comparisons across attributes within each group. This view maintains the ability to view the mean antigenic distance for test viruses within each clade that we had implemented in the static heatmaps. The raw measurements view plots each individual measurement to give users a detailed view of the quantity and distribution of titers, which can inform design decisions for future titer experiments. The titers thresholds can be toggled (Figure 2I) to add a clear demarcation of the threshold value for a view of when titer measurements are considered antigenically identical and distinct (Figure 2K). We discuss the application of these features in detail in two case studies in the following sections.



### 3.2 Case study 1: Retrospective vaccine selection in fall 2009

As noted above, the WHO convenes VCMs twice per year (Morris et al., 2017). The northern hemisphere VCM occurs in February or March ahead of a winter season in October through April. The southern hemisphere VCM occurs in September or October ahead of a winter season in April through October. To demonstrate the utility of an interactive visualization of serological measurements for vaccine composition decisions, we performed a retrospective analysis of a H3N2 vaccine update made for the southern hemisphere in the fall of 2009. We used publicly available sequence and titer data (Bedford et al., 2014) to reconstruct a H3N2 HA phylogeny and measurements panel representing information that was available at the time of the VCM (see Methods). We note that the actual selection process used a richer dataset which is not publicly available to include here, and as such these data should be seen as representative of the process only.

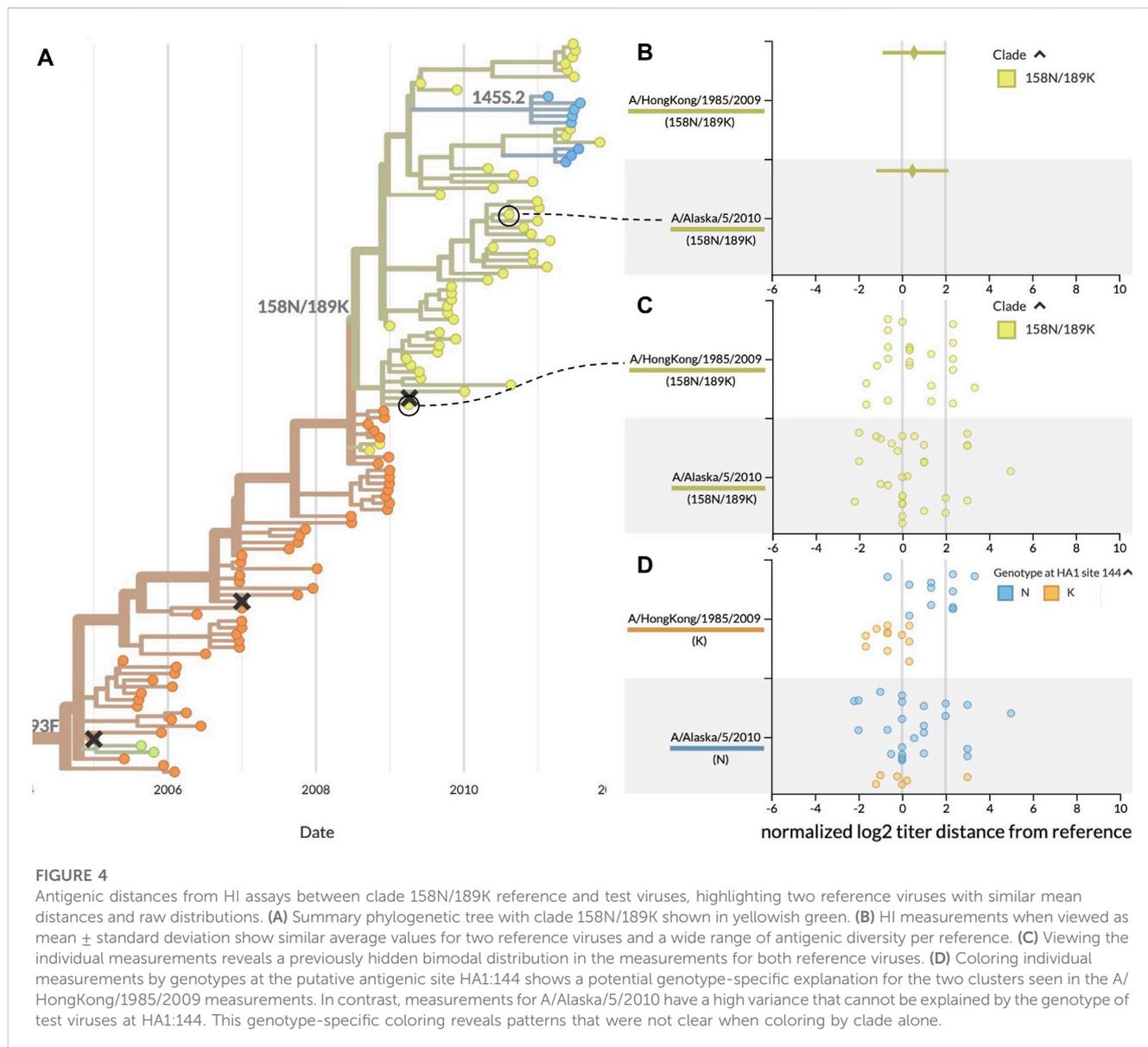
The VCM ahead of the 2010 southern hemisphere season occurred in September 2009. One major clade circulated at that time, 158N/189K, that descended from the previously dominant clade 193F (Figure 3A). The most recent H3N2 vaccine virus, A/Brisbane/10/2007, had been used for the two prior influenza seasons (2008 and 2009) and represented the older 193F clade (WHO, 2010). Prior to the A/Brisbane/10/2007 vaccine, A/Wisconsin/67/2005 had been the vaccine virus in the 2007 southern hemisphere season. While 193F dominated the 2008–2009 season, 158N/189K appeared to be dominant in September 2009. At the time of the VCM for both A/Brisbane/10/2007 and A/Wisconsin/67/2005 against test viruses from 158N/189K exceeded the  $2 \log_2$  threshold (Figure 3B), indicating their inability to cover this recent clade. In contrast, a newer vaccine candidate from

the 158N/189K clade, A/Perth/16/2009, had a titer distance of  $-0.01 \pm 0.67 \log_2$  units (mean  $\pm$  standard deviation) against test viruses from the same clade. These results strongly supported an update from the previous vaccine virus to a vaccine based on A/Perth/16/2009. The WHO announced this decision on 25 September 2009.

Our retrospective analysis allows us to see how the evolution of H3N2 continued after the VCM decision. Data collected after the vaccine selection deadline show that clade 158N/189K dominated for the following H3N2 seasons in both hemispheres, but two smaller clades each with HA1:145S mutations emerged from within this larger clade (Figure 3A). Although later HI measurements show that A/Perth/16/2009 did not cover later viruses from 158N/189K as effectively as it had covered earlier viruses from that clade ( $1.01 \pm 1.22 \log_2$  units), the new vaccine was still a better antigenic match than the previous two vaccines (Figure 3C). A/Perth/16/2009 did not effectively cover most viruses from the larger of the two derived clades, 145S.2 ( $1.94 \pm 0.71 \log_2$  units), indicating that the putative antigenic mutation defining the new clade might be responsible for antigenic drift. Despite this antigenic drift in a derived clade, A/Perth/16/2009 remained the WHO's recommended H3N2 vaccine virus for the 2010–12 southern hemisphere seasons and the 2010–11 and 2011–12 northern hemisphere seasons (WHO, 2022).

### 3.3 Case study 2: Identification of genotype-specific patterns through visualization of raw data

Influenza researchers often define clades of H3N2 viruses based on the presence of mutations that have been previously shown to enable



viruses to escape existing immunity (Wolf et al., 2006; Shih et al., 2007; Koel et al., 2013). The genetic similarity of viruses in the same clade typically corresponds with antigenic similarity of the same viruses as measured by HI assays. However, new mutations may arise within a clade that cause test viruses with those mutations to differ antigenically from earlier viruses in the same clade. Here, we demonstrate how aggregation of antigenic distances by clade can obscure the emergence of antigenically novel test viruses and how visualization of raw measurements can reveal these important patterns.

Using the same data from the previous case study, we inspected the patterns of HI measurements for the clade at the highest global frequency during the fall 2009 vaccine composition meeting, 158N/189K. We identified a reference virus from each of the two largest subclades of 158N/189K where both viruses had similar average HI measurements and distributions of raw measurements (Figure 4B). Serum against A/HongKong/1985/2009 had a mean antigenic distance of  $0.56 \pm 1.45 \log_2$  units to test viruses from 158N/189K, while serum against A/Alaska/5/

2010 had a mean distance of  $0.48 \pm 1.66 \log_2$  units. The distributions of raw HI measurements revealed clusters of values around titer distances of 0 and slightly above 2 for both reference viruses (Figure 4C). We hypothesized that these clusters could be explained by the presence of antigenic mutations in each reference virus's subclade. The two largest subclades of 158N/189K were defined by mutations at HA1 positions 62, 144, and 212. Each of these positions were previously identified as a putative antigenic site in HA where mutations could enable escape from existing immunity (Wolf et al., 2006). Additionally, position 144 is immediately adjacent to HA1:145, a position previously identified to contribute to novel antigenic clusters (Koel et al., 2013). Based on this genetic information, we colored the HI measurements by the genotypes of the test viruses at position 144. The genotype-specific coloring showed two different alleles at position 144 including the ancestral allele 144N and the derived allele 144K (Figure 4D). This view also revealed that test viruses with same 144K genotype as A/HongKong/1985/2009 had lower antigenic distances (mean of  $-0.56 \pm 0.75 \log_2$  units), while viruses with

the ancestral 144N genotype had higher antigenic distances ( $1.58 \pm 1.14 \log_2$  units). Interestingly, measurements against A/Alaska/5/2010 showed similar distances for both test viruses with 144N ( $0.57 \pm 1.72 \log_2$  units) and those with 144K ( $0.13 \pm 1.51 \log_2$  units). We observed the same patterns when grouping measurements by genotypes at the other two subclade-defining positions of HA1:62 and 212.

These results demonstrate how summary statistics can obscure biologically relevant patterns in the raw data. They also show how the ability to interactively color data by different viral attributes like genotype can produce hypotheses to explain the patterns we see in the raw data. For example, the genotype-specific patterns for A/HongKong/1985/2009 indicate the need for additional experiments to verify the antigenic effect of mutations at positions HA1:62, 144, and 212. In contrast, the high, genotype-independent variance of measurements against A/Alaska/5/2010 suggest that this reference virus might not be a stable vaccine candidate. This interactive visualization tool enables decision makers to explore their data and generate new hypotheses in ways that previous tools did not.

## 4 Discussion

Updating the seasonal influenza vaccine composition is a complex process that requires the synthesis of genetic and serological data and the interpretation of these data by a panel of international experts. Effective visualizations facilitate both the synthesis and interpretation by presenting data in a biologically meaningful context. Our interactive visualization tool presents serological data with a phylogenetic context, enabling decision makers to directly compare the antigenic distances between vaccine candidates and investigate patterns in the raw data. This tool regularly informs our discussions of influenza evolution with our collaborators in GISRS.

The move beyond static presentations of analyses towards interactive applications such as this facilitates more widespread usage and analysis of biological data. Specifically, the ability to link static views of the data, such as those found in VCM reports, with URLs that allow an interactive view into the data as presented in the figure is an important bridge between researchers. We hope that the adoption and continued development of biologically-informed visualization tools like this will facilitate a better understanding of pathogen evolution.

The benefits of integrated and interactive visualization of genetic and experimental data extend beyond serological measurements for seasonal influenza. High-throughput experimental measurements of mutational effects and immune escape in both seasonal influenza and SARS-CoV-2 have required custom tools for visualization and interpretation of these high-dimensional data (Hilton et al., 2020; Aksamentov et al., 2021; Garrett et al., 2021; Greaney et al., 2022). The data visualization presented here is amenable to showing similar multi-dimensional data, for instance linking different models to their scores across leaves in the phylogenetic tree in much the same way we have linked reference viruses to their titer measurements. The data structure for the measurements panel is purposefully agnostic to the pathogen or data generation approach. As research continues on the emergence of human pathogens from natural reservoirs in other organisms (Leendertz et al., 2016; Olival et al., 2017) and high-dimensional experimental measurements of

these pathogens accumulate (Soh et al., 2019; Starr et al., 2022), this flexible data structure and the resulting interactive visualizations could impact decision-making related to pandemic preparedness.

## Data availability statement

Details about how to prepare and analyze the data in this study live on GitHub at <https://github.com/blab/measurements-panel> and the dataset therein may be visualized via <https://nextstrain.org/community/blab/measurements-panel/flu/seasonal/h3n2/ha>

## Author contributions

JL, TS, TB, and JAH designed and implemented the measurements panel in Auspice. JOH and AB designed and implemented initial prototypes of interactive measurements visualizations. JOH, AB, and RN designed and implemented static visualizations. JL, JOH, and JAH wrote the manuscript.

## Funding

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number R01AI165821. TB is an Investigator of the Howard Hughes Medical Institute.

## Acknowledgments

This work started as a project for University of Washington (UW) class CSE 512 Data Visualization as a part of the UW eScience Advanced Data Science Option curriculum and we would like to thank Jeffrey Heer, Halden Lin, and Jane Hoffswell for their input on the initial design. We thank the members of GISRS for the high quality sequence and serological data used in this work. We have annotated sequence authors whenever that information was available at [github.com/blab/measurements-panel/blob/main/data/h3n2\\_ha\\_metadata.tsv](https://github.com/blab/measurements-panel/blob/main/data/h3n2_ha_metadata.tsv). Without open data sharing, this work would not be possible.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Aksamentov, I., Roemer, C., Hodcroft, E. B., and Neher, R. A. (2021). Nextclade: Clade assignment, mutation calling and quality control for viral genomes. *J. Open Source Softw.* 6, 3773. doi:10.21105/joss.03773
- Bedford, T., Huddleston, J., Potter, B., and Neher, R. A. (2019). Seasonal influenza circulation patterns and projections for September 2019 to September 2020. bioRxiv. doi:10.1101/780627
- Bedford, T., and Neher, R. (2018). Seasonal influenza circulation patterns and projections for Feb 2018 to Feb 2019. bioRxiv. doi:10.1101/271114
- Bedford, T., Suchard, M. A., Lemey, P., Dudas, G., Gregory, V., Hay, A. J., et al. (2014). Integrating influenza antigenic dynamics with molecular evolution. *Elife* 3, e01914. doi:10.7554/elife.01914
- Correll, M., and Gleicher, M. (2014). Error bars considered harmful: Exploring alternate encodings for mean and error. *IEEE Trans. Vis. Comput. Graph.* 20, 2142–2151. doi:10.1109/tvcg.2014.2346298
- Felsenstein, J. (2003). *Inferring phylogenies*. 2nd Edition. Sunderland, Massachusetts, United States: Sinauer Associates.
- Fernandes, M., Walls, L., Munson, S., Hullman, J., and Kay, M. (2018). “Uncertainty displays using quantile dotplots or cdfs improve transit decision-making,” in ACM Human Factors in Computing Systems (CHI), Seattle, WA, United States, 14 April 2018 (United States: University of Washington).
- Garrett, M. E., Galloway, J., Chu, H. Y., Itell, H. L., Stoddard, C. I., Wolf, C. R., et al. (2021). High-resolution profiling of pathways of escape for SARS-CoV-2 spike-binding antibodies. *Cell* 184, 2927–2938.e11. doi:10.1016/j.cell.2021.04.045
- Greaney, A. J., Starr, T. N., and Bloom, J. D. (2022). An antibody-escape estimator for mutations to the SARS-CoV-2 receptor-binding domain. *Virus Evol.* 8, veac021. doi:10.1093/ve/veac021
- Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., et al. (2018). Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics* 34, 4121–4123. doi:10.1093/bioinformatics/bty407
- Hilton, S. K., Huddleston, J., Black, A., North, K., Dingens, A. S., Bedford, T., et al. (2020). dms-view: Interactive visualization tool for deep mutational scanning data. *J. Open Source Softw.* 5, 2353. doi:10.21105/joss.02353
- Hirst, G. K. (1943). Studies of antigenic differences among strains of influenza A by means of red cell agglutination. *J. Exp. Med.* 78, 407–423. doi:10.1084/jem.78.5.407
- Huddleston, J., Hadfield, J., Sibley, T. R., Lee, J., Fay, K., Ilcisin, M., et al. (2021). Augur: A bioinformatics toolkit for phylogenetic analyses of human pathogens. *J. Open Source Softw.* 6, 2906. doi:10.21105/joss.02906
- Hullman, J., Resnick, P., and Adar, E. (2015). Hypothetical outcome plots outperform error bars and violin plots for inferences about reliability of variable ordering. *PLOS ONE* 10, 0142444–e142525. doi:10.1371/journal.pone.0142444
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi:10.1093/molbev/mst010
- Katz, J. M., Hancock, K., and Xu, X. (2011). Serologic assays for influenza surveillance, diagnosis and vaccine evaluation. *Expert Rev. Anti-infective Ther.* 9, 669–683. doi:10.1586/eri.11.51
- Koel, B. F., Burke, D. F., Bestebroer, T. M., van der Vliet, S., Zondag, G. C. M., Vervaet, G., et al. (2013). Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science* 342, 976–979. doi:10.1126/science.1244730
- Leendertz, S. A., Gogarten, J. F., Düx, A., Calvignac-Spencer, S., and Leendertz, F. H. (2016). Assessing the evidence supporting fruit bats as the primary reservoirs for Ebola viruses. *Ecohealth* 13, 18–25. doi:10.1007/s10393-015-1053-0
- Lemey, P., Salemi, M., and Vandamme, A. (Editors) (2012). *The phylogenetic handbook*. 2nd Edition (Cambridge, United Kingdom: Cambridge University Press). doi:10.1017/CBO9780511819049
- Mackinlay, J. (1986). Automating the design of graphical presentations of relational information. *ACM Trans. Graph.* 5, 110–141. doi:10.1145/22949.22950
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi:10.1093/molbev/msaa015
- Mölder, F., Jablonski, K., Letcher, B., Hall, M., Tomkins-Tinch, C., Sochat, V., et al. (2021). Sustainable data analysis with Snakemake. *F1000Research* 10, 33. doi:10.12688/f1000research.29032.2
- Morris, D. H., Gostic, K. M., Pompei, S., Bedford, T., Luksza, M., Neher, R. A., et al. (2017). Predictive modeling of influenza shows the promise of applied evolutionary biology. *Trends Microbiol.* 26, 102–118. doi:10.1016/j.tim.2017.09.004
- Neher, R. A., Bedford, T., Daniels, R. S., Russell, C. A., and Shraiman, B. I. (2016). Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. *Proc. Natl. Acad. Sci. U.S.A.* 113, E1701–E1709. doi:10.1073/pnas.1525578113
- Neher, R. A., and Bedford, T. (2015). nextflu: real-time tracking of seasonal influenza virus evolution in humans. *Bioinformatics* 31, 3546–3548. doi:10.1093/bioinformatics/btv381
- Neher, R. A., and Bedford, T. (2018). Real-time analysis and visualization of pathogen sequence data. *J. Clin. Microbiol.* 56, 00480–e518. doi:10.1128/jcm.00480-18
- Olivai, K. J., Hosseini, P. R., Zambrana-Torrel, C., Ross, N., Bogich, T. L., and Daszak, P. (2017). Host and viral traits predict zoonotic spillover from mammals. *Nature* 546, 646–650. doi:10.1038/nature22975
- Petrova, V. N., and Russell, C. A. (2018). The evolution of seasonal influenza viruses. *Nat. Rev. Microbiol.* 16, 47–60. doi:10.1038/nrmicro.2017.118
- Sagulenko, P., Puller, V., and Neher, R. A. (2018). TreeTime: Maximum-likelihood phylogenetic analysis. *Virus Evol.* 4, vex042. doi:10.1093/ve/vex042
- Shih, A. C.-C., Hsiao, T.-C., Ho, M.-S., and Li, W.-H. (2007). Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. *Proc. Natl. Acad. Sci.* 104, 6283–6288. doi:10.1073/pnas.0701396104
- Smith, D. J., Lapedes, A. S., de Jong, J. C., Bestebroer, T. M., Rimmelzwaan, G. F., Osterhaus, A. D. M. E., et al. (2004). Mapping the antigenic and genetic evolution of influenza virus. *Science* 305, 371–376. doi:10.1126/science.1097211
- Soh, Y. S., Moncla, L. H., Eguia, R., Bedford, T., and Bloom, J. D. (2019). Comprehensive mapping of adaptation of the avian influenza polymerase protein PB2 to humans. *Elife* 8, e45079. doi:10.7554/elife.45079
- Starr, T. N., Zepeda, S. K., Walls, A. C., Greaney, A. J., Alkhovsky, S., Veessler, D., et al. (2022). ACE2 binding is an ancestral and evolvable trait of sarbecoviruses. *Nature* 603, 913–918. doi:10.1038/s41586-022-04464-z
- Steinbrück, L., and McHardy, A. C. (2012). Inference of genotype-phenotype relationships in the antigenic evolution of human influenza A (H3N2) viruses. *PLoS Comput. Biol.* 8, e1002492. doi:10.1371/journal.pcbi.1002492
- WHO (2010). Recommendations for influenza vaccine composition - archive 1998-2010. Available at: <https://www.who.int/teams/global-influenza-programme/vaccines/who-recommendations/recommendations-for-influenza-vaccine-composition-archive>.
- WHO (2022). Recommendations for influenza vaccine composition. Available at: <https://www.who.int/teams/global-influenza-programme/vaccines/who-recommendations>.
- Wolf, Y. I., Viboud, C., Holmes, E. C., Koonin, E. V., and Lipman, D. J. (2006). Long intervals of stasis punctuated by bursts of positive selection in the seasonal evolution of influenza A virus. *Biol. Direct* 1, 34. doi:10.1186/1745-6150-1-34