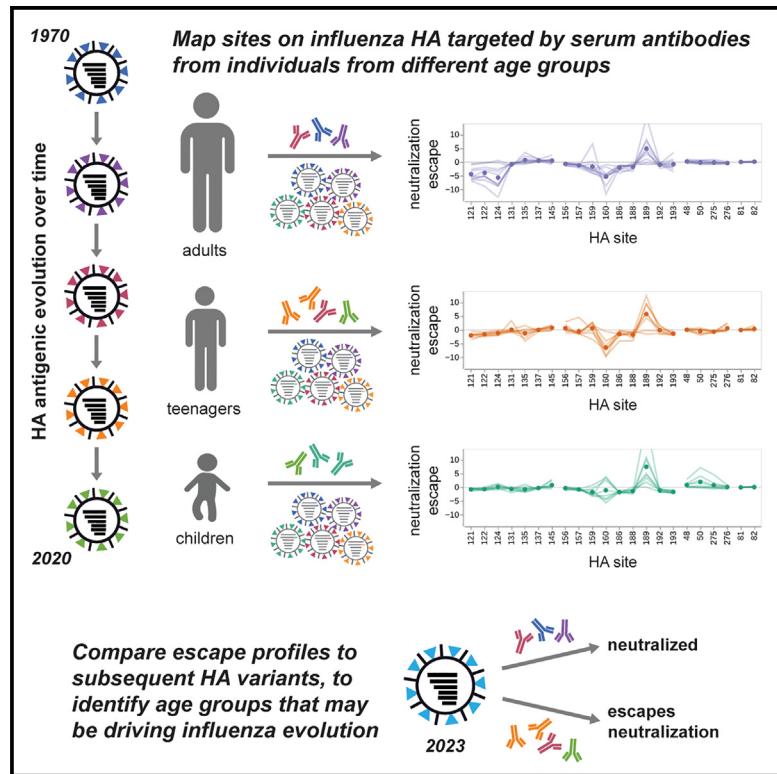


Age-dependent heterogeneity in the antigenic effects of mutations to influenza hemagglutinin

Graphical abstract



Authors

Frances C. Welsh, Rachel T. Eguia,
Juhye M. Lee, ...,
Frederick A. Matsen IV, Maciej F. Boni,
Jesse D. Bloom

Correspondence

jbloom@fredhutch.org

In brief

Welsh et al. measured how mutations to H3N2 influenza hemagglutinin affected neutralization by sera from individuals of different age groups. The mutations that affected neutralization varied among age groups. During its natural evolution, the H3N2 influenza virus especially acquires mutations that escape neutralization by sera from teenagers.

Highlights

- Measure how mutations to influenza hemagglutinin affect neutralization by human sera
- Compare antigenic effects of mutations on sera from individuals of different ages
- Evolution of virus appears to be especially driven by sera from teenagers



Article

Age-dependent heterogeneity in the antigenic effects of mutations to influenza hemagglutinin

Frances C. Welsh,^{1,2} Rachel T. Eguia,^{2,3} Juhye M. Lee,^{2,3} Hugh K. Haddox,² Jared Galloway,² Nguyen Van Vinh Chau,⁴ Andrea N. Loes,^{2,3} John Huddleston,⁵ Timothy C. Yu,^{1,2} Mai Quynh Le,⁶ Nguyen T.D. Nhat,^{4,7} Nguyen Thi Le Thanh,⁴ Alexander L. Greninger,^{8,9} Helen Y. Chu,⁹ Janet A. Englund,¹⁰ Trevor Bedford,^{3,5} Frederick A. Matsen IV,^{2,3} Maciej F. Boni,^{4,7,11} and Jesse D. Bloom^{2,3,12,*}

¹Molecular and Cellular Biology Graduate Program, University of Washington, and Basic Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA 98109, USA

²Basic Sciences Division and Computational Biology Program, Fred Hutchinson Cancer Center, Seattle, WA 98109, USA

³Howard Hughes Medical Institute, Seattle, WA 98109, USA

⁴Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam

⁵Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, WA 98109, USA

⁶National Institutes for Hygiene and Epidemiology, Hanoi, Vietnam

⁷Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK

⁸Department of Laboratory Medicine and Pathology, University of Washington School of Medicine, Seattle, WA 98195, USA

⁹Division of Allergy and Infectious Diseases, University of Washington School of Medicine, Seattle, WA 98195, USA

¹⁰Seattle Children's Research Institute, Seattle, WA 98109, USA

¹¹Center for Infectious Disease Dynamics, Department of Biology, Pennsylvania State University, University Park, PA 16802, USA

¹²Lead contact

*Correspondence: jbloom@fredhutch.org

<https://doi.org/10.1016/j.chom.2024.06.015>

SUMMARY

Human influenza virus evolves to escape neutralization by polyclonal antibodies. However, we have a limited understanding of how the antigenic effects of viral mutations vary across the human population and how this heterogeneity affects virus evolution. Here, we use deep mutational scanning to map how mutations to the hemagglutinin (HA) proteins of two H3N2 strains, A/Hong Kong/45/2019 and A/Perth/16/2009, affect neutralization by serum from individuals of a variety of ages. The effects of HA mutations on serum neutralization differ across age groups in ways that can be partially rationalized in terms of exposure histories. Mutations that were fixed in influenza variants after 2020 cause greater escape from sera from younger individuals compared with adults. Overall, these results demonstrate that influenza faces distinct antigenic selection regimes from different age groups and suggest approaches to understand how this heterogeneous selection shapes viral evolution.

INTRODUCTION

Seasonal human influenza virus undergoes continual antigenic evolution in response to immune selection from the host population. Although human individuals retain potent neutralizing antibodies to historical strains from prior exposures, they are susceptible to re-infection roughly every 5 years due to rapid accumulation of mutations in the hemagglutinin (HA) surface protein that erode antibody neutralization.^{1,2} Antigenically distinct strains of influenza H3N2 emerge every few years,³ resulting in frequent turnover of the virus population.^{4–6} Because of this continual antigenic evolution, individuals in the population have diverse exposure histories, which in turn shape their antibody response to subsequent infection or vaccination.^{7–11} Neutralizing antibody titers are typically highest against strains seen early in life and steadily decline against strains seen later in life,^{1,2,8,12,13} which suggests that immune specificity varies, at least in part, by birth cohort.

Population-level heterogeneity has been proposed as a factor shaping the antigenic evolution of influenza,^{14–16} but we lack a detailed understanding of how the antigenic impacts of viral mutations differ across individuals in the population. Prior work has characterized this heterogeneity by analyzing human serum neutralization or hemagglutination-inhibition of recently circulating strains,^{14,15,17} or analyzing demographics of confirmed influenza cases over many seasons.^{18,19} Kim et al. found that individuals within the same age group have significantly correlated neutralization titers to different HA or NA clades.¹⁷ However, these studies have focused on susceptibility to naturally circulating strains and cannot fully resolve how specific viral mutations affect neutralization by the polyclonal antibodies of different individuals.

Deep mutational scanning can be used to map the mutations in influenza HA that confer escape from neutralization by human sera, and it has shown that single mutations can sometimes lead to large drops in neutralization.²⁰ However, prior deep mutational



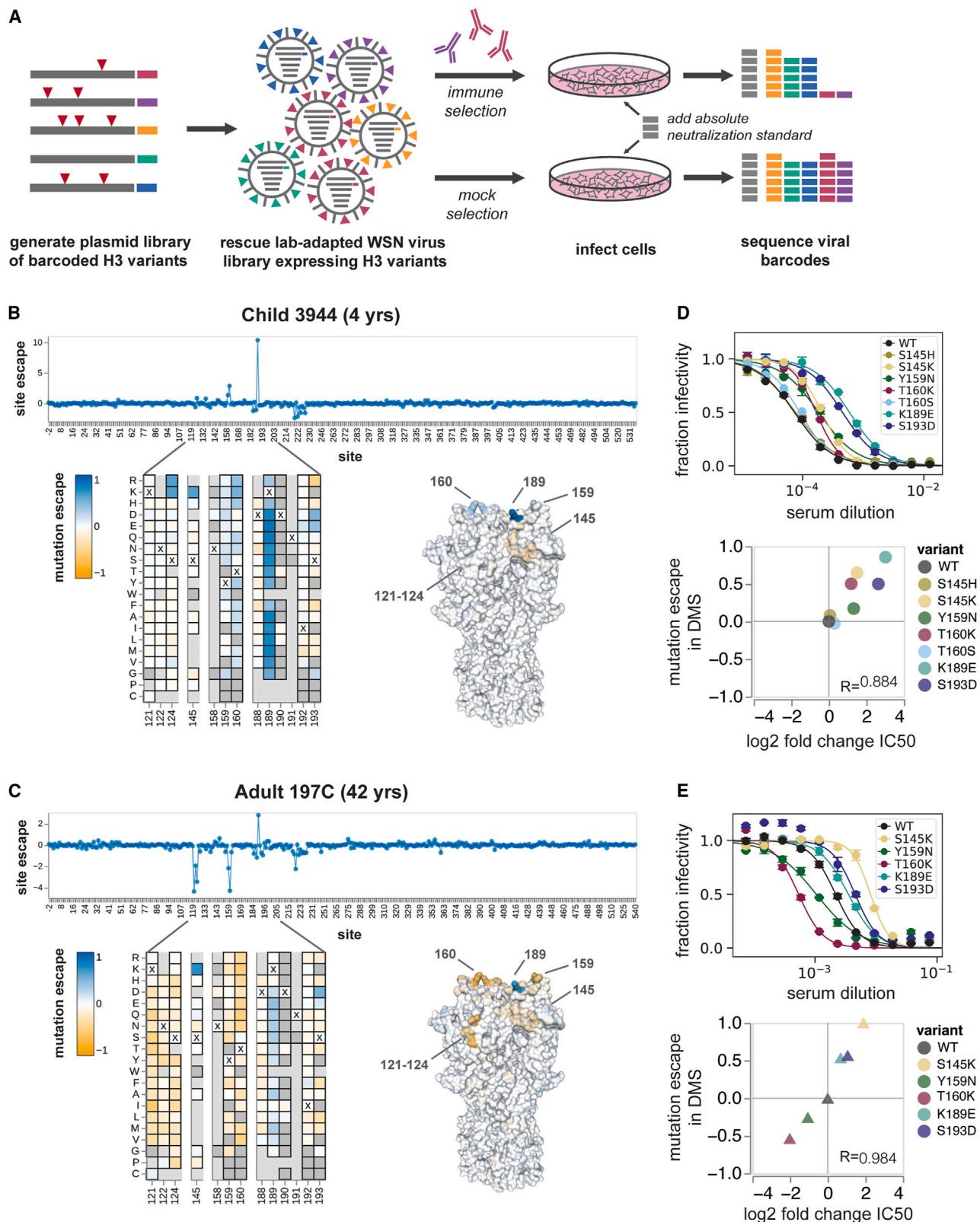


Figure 1. Deep mutational scanning escape maps show how mutations to HA affect viral neutralization

(A) Experimental workflow for mapping HA serum escape. The virus library is incubated with either the serum of interest, or media alone, then added to cells. At 13 h post-infection, RNA is harvested from cells, and viral barcodes are reverse-transcribed and amplified by PCR. Escape scores are calculated for all sampled

(legend continued on next page)

scanning studies have not analyzed the antigenic effects of mutations across many individuals from different age groups. To address this gap, we here use deep mutational scanning to map how mutations to the HAs of two H3N2 vaccine strains (A/Hong Kong/45/2019 and A/Perth/16/2009) affect neutralization by sera collected from individuals of different ages during the time frame that strain was in the vaccine. We find that the antigenic effects of mutations differ across individuals and especially age cohorts, with single mutations often causing large drops in neutralization. Differences among age cohorts are partially explained by inferred exposure history. We also find some evidence that the mutations that fix during natural influenza evolution in subsequent years often confer substantial neutralization escape from child and especially teenage sera. Overall, our work provides a deeper understanding of the antigenic selection exerted by different subsets of the population and suggests ways in which this heterogeneity may shape influenza evolution.

RESULTS

A deep mutational scanning approach to map mutations to HA that affect viral neutralization by human serum

To systematically map how mutations to HA affect neutralization by human sera, we developed deep mutational scanning libraries in the HA of A/Hong Kong/45/2019, which was the H3N2 component of the 2020–2021 influenza vaccine. The HA libraries were incorporated into replication-competent influenza virions, with the other viral genes derived from the lab-adapted strain A/WSN/1933(H1N1) for biosafety reasons (Figure 1A). We improved upon prior influenza deep mutational scanning studies^{20,21} by incorporating barcodes into the viral genome and by using a non-neutralized standard to allow for absolute quantification of neutralization by deep sequencing²² (Figures 1A, S1, and S2). We designed our libraries to include only mutations that are functionally tolerated in HA, as assessed by prior deep mutational scanning and analysis of natural H3 HA sequences. We created two fully independent replicate HA libraries, each containing approximately 30,000 functional HA variants with an average of 3 mutations per gene (Figure S3). The effect of each HA mutation on viral infection in cell culture was calculated by comparing the frequency of the mutation in the plasmid library to the final virus library (Figure S4A). Mutations highly deleterious to HA-mediated infection were excluded from analysis of their effects on serum neutralization (Figure S4B).

mutations by sequencing barcodes from variants that successfully infected cells in the serum-selection condition, and comparing to the mock-selection. A neutralization standard is used to convert relative sequencing counts into absolute neutralization (Figure S2). (B and C) Example serum escape maps from a child (B) and an adult (C). The line plot shows the summed effects of all sampled mutations at each site ("site escape," roughly proportional to fold change in neutralization), with positive values indicating neutralization escape and negative values indicating increased neutralization. Heatmaps show the effects of individual mutations for key sites ("mutation escape"). Note that in some cases, different amino-acid mutations at the same site may have different effects on escape, an effect that can only be fully visualized by looking at the heatmaps rather than the line plots. Due to space, we cannot show the full heatmaps for all sera and sites in the main text figures; see the "Interactive plots of mutation antibody escape" at https://dms-vep.org/flu_h3_hk19_dms/ for full heatmaps showing amino-acid effects for all sera. In the heatmaps, X indicates the wild-type amino acid in A/Hong Kong/45/2019 at each site, dark gray indicates mutations measured to be highly deleterious to HA-mediated infection, and light gray indicates mutations not sampled in the libraries. The structures are colored by the summed site escape scores and show the HA structure for the A/Victoria/361/2011 H3 HA (PDB: 4O5N). (D and E) Validations of escape scores by traditional neutralization assays for selected mutants against the child (D) and adult (E) sera. Correlation plots show the mutation escape scores measured by deep mutational scanning versus the log₂ fold change in IC50 between the wild-type library strain and the mutant of interest. R indicates the Pearson correlation.

Using these deep mutational scanning libraries, we were able to map mutations across HA that increase either resistance or sensitivity of the virus to neutralization. To illustrate the type of data generated by our approach, we first show example escape maps for two sera collected in 2020 in Seattle, one from a 4-year-old child and one from a 42-year-old adult (Figures 1B and 1C). The mutation "escape scores" in these maps are roughly proportional to the log fold change in neutralization titer.²³ Positive escape scores indicate mutations that confer resistance to neutralization. Negative scores indicate mutations that increase neutralization, typically by reverting sites to amino-acid identities present in older viral strains to which the individual was exposed. We validated that the escape scores measured by deep mutational scanning correlated strongly with the results of traditional neutralization assays (Figures 1D and 1E).

The escape map for the child's serum (serum 3944) is dominated by positive escape at just a few sites, most prominently site 189 (H3 numbering), followed by sites 159 and 160 (Figure 1B and interactive plot linked in the figure legend). There is one region (sites 221–228) where mutations modestly increase sensitivity to neutralization. This is likely by decreasing receptor avidity rather than directly affecting antibody binding,²⁴ as mutations at these sites are known to alter receptor binding^{25,26} and have consistent neutralization-sensitizing effects across all mapped sera (Figure S7). At some sites (e.g., site 189), virtually any amino-acid mutation causes escape, while at other sites only specific mutations cause escape (e.g., S145K) (Figure 1B).

By contrast, the escape map for the adult's serum (serum 197C) reveals that some mutations increase neutralization sensitivity while others cause neutralization escape. Specifically, mutations at sites 121–124 and 159–160 increase neutralization by this serum (reflected by negative escape scores), whereas mutations at site 189 cause escape (Figure 1C). The sites with the strongest sensitizing effects all correspond to regions where H3N2 evolution has previously introduced N-linked glycans that mask antibody epitopes. In 1997, H3N2 influenza acquired an N-linked glycan at site 122 that masked a key antigenic site. Mutations at sites 122 and 124 that eliminate this glycan greatly increase sensitivity to neutralization, suggesting that antibodies targeting this now-inaccessible site are still present at high levels in the serum. Similarly, the 158 glycan was introduced in 2014 and masked an immunodominant antigenic region.^{27,28} Increased neutralization sensitivity from mutations that remove glycosylation motifs introduced many years ago is consistent with the well-described concept of antigenic seniority, where

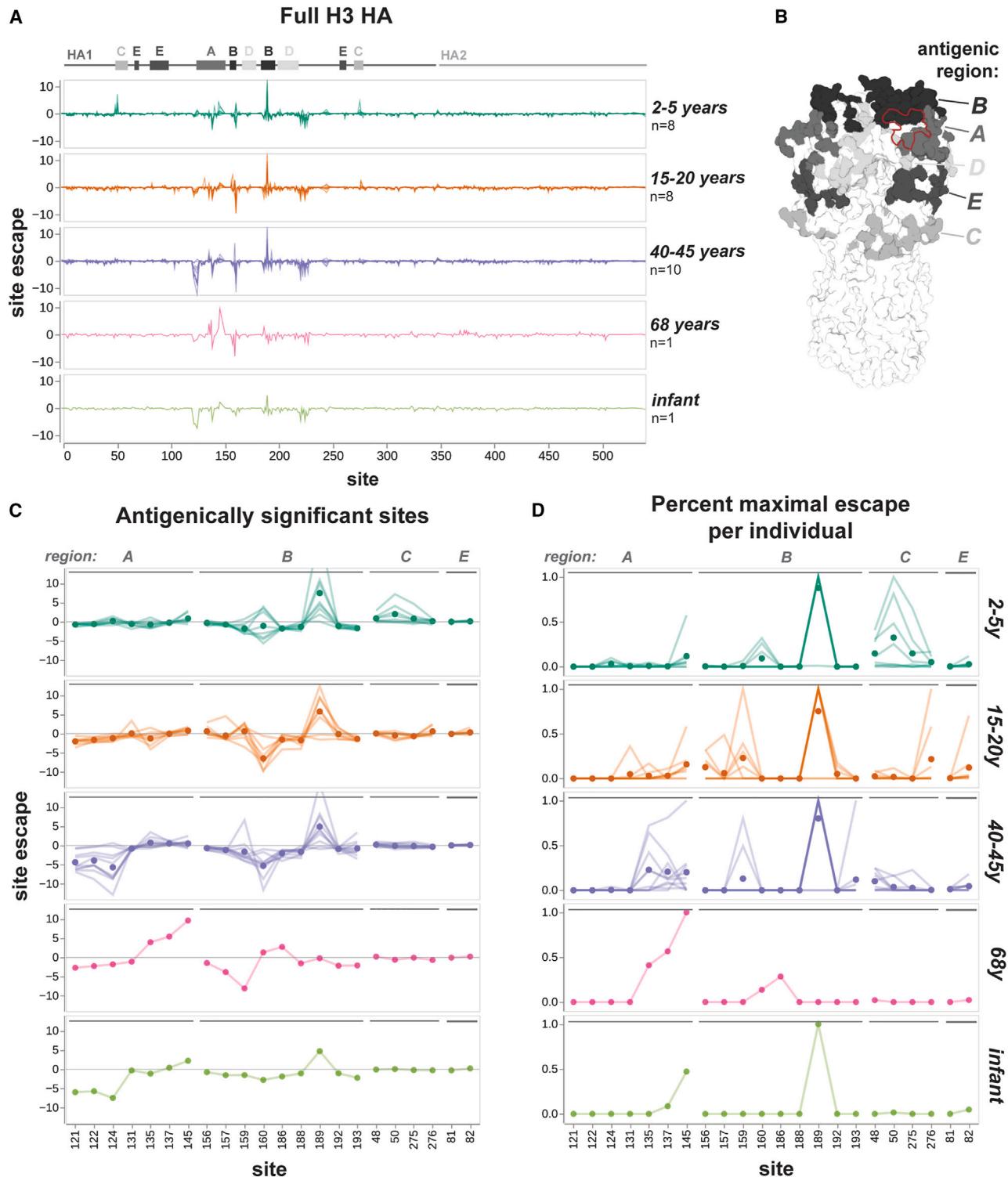


Figure 2. Escape maps from different cohorts show age-dependent trends in serum targeting of A/Hong Kong/45/2019 HA

(A) Serum escape maps for different human individuals in each age cohort. Each line represents an individual, and individuals within the same cohort are overlaid in the same plot. Site escape scores are the summed effects of all sampled mutations at each site. Approximate locations of HA1, HA2, and antigenic regions A through E (as defined by Muñoz and Deem²⁹) are labeled above the plots.

(B) H3 HA trimer colored by antigenic region, with the receptor-binding pocket outlined in red. Structure is from A/Victoria/361/2011 H3 HA (PDB: 4O5N).

(C) Serum escape maps at antigenically significant sites for human individuals in each age cohort. Each line is an individual escape map, and points represent the mean escape score for that age cohort at that site. The y axis is clipped to 12.5 in both (A) and (C), although one child has an escape score of 23.9 at site 189.

(legend continued on next page)

individuals retain neutralizing antibodies against strains encountered early in life.¹²

Antigenic effects of HA mutations differ among individuals and age cohorts

To better understand how antibody escape mutations differ among human individuals, we generated escape maps for sera collected in 2020 from 8 to 10 individuals each in three different age cohorts: children (2–5 years of age), teenagers (15–20 years), and adults (40–45 years) (Figure 2). Most individuals were vaccinated within the past year (Tables S1 and S2). We also mapped sera from an unvaccinated 2-month-old infant and a 68-year-old adult. Sera were chosen from a larger set, based on having good neutralization activity against the parental A/Hong Kong/45/2019 library strain HA (Figure S5). To confirm that our escape maps are replicable, we mapped two sets of sera sampled from the same individual on different days and found that the results were almost identical (Figure S6).

For all sera, the strongest escape mutations are within a small subset of sites in HA1 that correspond to classically defined antigenic regions³⁰ (Figures 2A and 2B). The mutations that cause the greatest escape are in antigenic regions A and B, which are near the sialic-acid receptor-binding pocket and have alternated in immunodominance throughout H3N2 evolution.³¹ Site 189 is a significant escape site across all age cohorts, consistent with prior work identifying this site as a major determinant of antigenic change for H3N2.³² Mutations in antigenic regions C and E, which are lower on the HA head, cause weaker escape from the sera of some individuals. Two outlier sera, child 4584 and teenager 3856, exclusively target regions C and E and have no escape mutations in the immunodominant regions A and B (Figure S7). Notably, out of all mapped sera, these two sera had the lowest neutralization activity against the parental library strain. This observation is consistent with prior work showing that antibodies targeting regions near the receptor-binding pocket (e.g., antigenic regions A and B) tend to be more potent than antibodies targeting sites lower on the HA head (e.g., antigenic regions C and E).^{33–35}

There are clear differences in escape maps among age cohorts, though they share some sites of strong escape. This can best be seen by analyzing both the overall magnitude of escape across sera (Figure 2C) and the sites that confer the strongest escape for each serum, normalizing by their maximum escape score (Figure 2D). For most cohorts, the highest overall escape is at site 189 in the immunodominant antigenic region B. Most adult sera are also escaped by mutations in region A, which was immunodominant in the early 1990s based on analysis of human convalescent sera.³⁶ Teenage sera are more often escaped by mutations at sites 156–159 in region B, which has likely been immunodominant for their entire lifetime^{31,37} (Figure 2D). Some child sera target a different site in region B, site 160, and additional sites in region C. These region C sites are distant from the receptor-binding pocket (Figure 2B) and proximal to the vestigial esterase subdomain, a target of some broadly neutralizing antibodies.³⁸ The escape map for the serum

from an elderly individual is distinct from other cohorts, with the strongest escape at site 145 and no signal at 189. This individual was born 16 years before the introduction of H3N2, and although we cannot draw strong conclusions from a single serum, differences in childhood HA exposure could contribute to their unique escape map.

Sites where mutations strongly increase neutralization sensitivity (indicated by negative values in the escape maps) are more common in older age cohorts (Figure 2C). There are two regions of strongly sensitizing mutations in adults: sites 121–124 and site 160. For teenagers, only site 160 exhibits strongly sensitizing mutations. Some children have weakly sensitizing mutations at site 160. Sites 158–160 and 122–124 are glycosylation motifs that appeared during the lifetimes of certain cohorts, and individuals with sensitizing mutations at these sites were exposed to prior strains without this glycan. The 122–124 glycan appeared in 1997. Adults exposed to H3N2 prior to 1997 could have generated neutralizing antibodies targeting this region, but the epitope has always been masked in strains seen by younger cohorts. The 158–160 glycan appeared in 2014, during the lifetime of both teenagers and adults, so these cohorts have similar sensitizing mutations at site 160. The weak negative signal in some children may be attributed to vaccination against A/Kansas/14/2017 H3N2 in the previous year, which does not express the 158–160 glycan³⁹ (Table S1). There are additional sites of weak negative signal, such as 159 and 193 in adults and 135 in teenagers, that correspond to sites mutated during their lifetimes (Figure S9).

By visual inspection, it appears that teenagers and adults often have more heterogeneous escape maps within their age cohorts than children (Figure 2D). The main escape site for all children is site 189, with the exception of one outlier that exclusively targets antigenic region C. Secondary escape sites are limited and highly consistent among individuals. While 189 is also a major target for most teenagers and adults, the magnitude is comparable to escape at other sites for many individuals, and these other sites are variable between sera. This suggests that a population with limited, likely common prior exposure to influenza develops a narrow and fairly stereotyped antibody response. As exposure history increases, antibody responses diversify, even within an age cohort.

The escape map for the 2-month-old unvaccinated infant is similar to the adults, with neutralization-sensitizing mutations at sites 121–124 (Figure 2C). Influenza antibodies are transferred from mother to fetus during pregnancy and confer passive immunity in newborns for several months.⁴⁰ We hypothesize that this infant gained antibodies targeting the glycan-masked 121–124 region from their mother, alongside antibodies with neutralization activity at sites 145 and 189. However, any conclusions we can draw about infant sera are limited by the availability of data for just one serum sample.

In short, these escape maps demonstrate that sera from different age cohorts are escaped by mutations at different sites in HA. Children with limited prior exposure tend to have serum

(D) Escape maps for key sites with escape floored at zero and normalized to the maximal positive escape score for that individual. Lines represent individuals, and points represent the mean normalized escape score for that age cohort. Note that the site escape scores summarize the effects of all mutations at each site. For more detailed data, see “Interactive plots of mutation antibody escape” at https://dms-vep.org/flu_h3_hk19_dms/ for full heatmaps showing individual mutation effects for all sera.

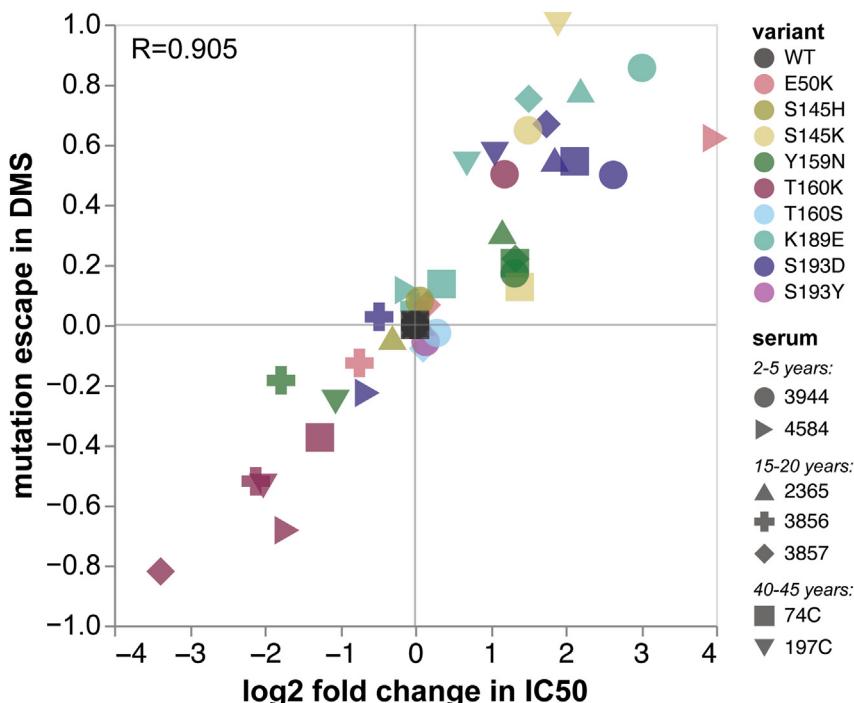


Figure 3. Validation of deep mutational scanning results using neutralization assays

The escape score measured by deep mutational scanning for each mutation is plotted against the log₂ fold change in IC50 between the wild-type library strain and the mutant of interest. R indicates the Pearson correlation. See Figure S8 for full neutralization curves and serum-level correlation plots.

Mutations at site 160 are sensitizing for most sera but confer escape in three children. This is especially notable because 160T is part of a glycosylation motif. It is unclear whether these children have antibodies targeting the glycan itself or a glycan-adjacent region, but we validated that T160K is indeed an escape mutant for child 3944. We also confirmed that T160S, which maintains the glycan, is neutral.

Age-dependent heterogeneity in serum escape is also seen in an unvaccinated cohort measured against a different HA strain

that is escaped by a fairly homogeneous and focused set of mutations, whereas escape mutations become more heterogeneous with age. The prevalence of neutralization-sensitizing mutations in the escape maps for adults indicated that older individuals often retain antibodies that target strains seen earlier in their lives. At a population level, site 189 is the most potent target of neutralizing antibodies, but individual teenagers and adults also target diverse additional sites in several known antigenic regions. This heterogeneity is especially notable given that most individuals were vaccinated against the same strains in the prior season (Tables S1 and S2).

Validation of the deep mutational scanning in traditional neutralization assays

We validated key mutations using traditional neutralization assays for eight representative sera from different age cohorts. The changes in the half-maximal inhibitory concentration (IC50) measured by traditional neutralization assays strongly correlate with the escape scores measured by deep mutational scanning (Figures 3 and S8). Note that these correlations are for the mutation-level escape scores, and different mutations at the same site can have different effects on neutralization.

The validation neutralization assays also confirm the serum-to-serum heterogeneity observed in the deep mutational scanning. For example, Y159N confers escape from four sera but makes the virus more sensitive to neutralization by sera 3856 and 197C (Figure 3). Despite K189E and S193D being strong escape mutations for almost all other sera, they had neutral or slightly sensitizing effects for sera 3856 and 4584. At some sites, only specific mutations confer escape. We confirmed that S145K and S193D were escape mutants, but S145H and S193Y were indeed neutral. These validations emphasize that antigenic effects of mutations are often highly variable across a population.

The aforementioned trends were all based on escape maps for the HA of the A/Hong Kong/45/2019 (H3N2) strain generated using sera collected in 2020 in Seattle, Washington, USA. To assess whether trends in age-dependent heterogeneity in escape mutations are a general phenomenon, we also analyzed sera from nine children (2–4 years of age) and three adults (30–35 years of age) collected between 2010 and 2011 from Ho Chi Minh City, Vietnam, against the HA of the A/Perth/16/2009 (H3N2) strain. In contrast with the high vaccination rates in the Seattle cohort (Tables S1 and S2), vaccination rates in Vietnam are negligible.⁴¹ We also generated escape maps for four ferrets infected with the A/Perth/16/2009 library strain. The A/Perth/16/2009 HA deep mutational scanning used an older approach with previously described libraries^{20,42} that are not barcoded. This older deep mutational scanning approach measures relative escape scores for each serum, but the magnitude of these escape scores is not directly comparable between sera. Furthermore, the older deep mutational scanning approach can only reliably measure escape mutations, not sensitizing mutations.

The ferret and child sera have similar escape mutations within their respective cohorts, while the adult sera target heterogeneous sites (Figures 4A–4D). All four singly infected ferrets have almost identical escape maps and target sites 189 and 193. The children have slightly more variable escape maps but are still quite similar to each other and all target some combination of sites 189, 193, and 159–160. By contrast, each adult serum most strongly targets a different site: 159, 160, 189, or 192. These sites are proximal to one another (Figure 4B), but only confer escape from certain sera, suggesting that neutralizing activity against HA from polyclonal sera is highly focused on specific residues. We validated the antigenic effects of key mutants in neutralization assays and found that they correlated very well with the predicted escape scores (Figures 4E and

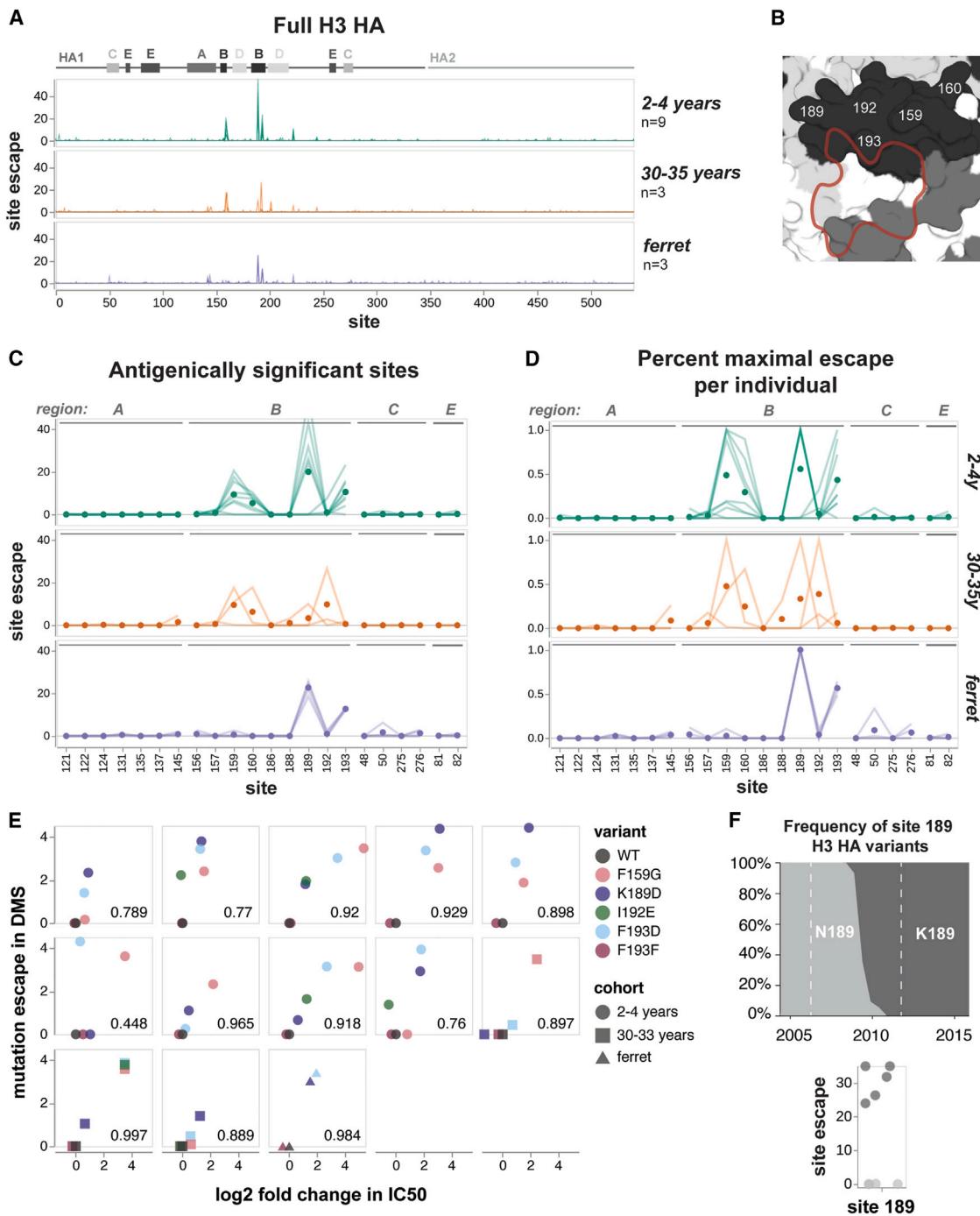


Figure 4. Escape maps from A/Perth/16/2009 H3 HA, measured against sera from an unvaccinated cohort and singly infected ferrets

(A) Serum escape maps for different individuals in each cohort. Each line represents an individual, and individuals within the same cohort are overlaid in the same plot. Site escape scores are the summed effects of all sampled mutations at each site. Approximate locations of HA1, HA2, and antigenic regions A through E are labeled above the plots.

(B) Antigenic region B of H3 HA is shown in black, with the receptor-binding pocket outlined in red. Structure is from A/Victoria/361/2011 H3 HA (PDB: 4O5N). (C and D) Serum escape maps at antigenically significant sites for individuals in each cohort. Each line is an individual escape map, and points represent the mean escape score for the cohort at that site. (C) shows the summed site escape scores, while (D) shows the max normalized escape scores for each individual. Escape scores for singly infected ferrets were originally published in Lee et al.²⁰

(E) Correlation of escape scores measured in deep mutational scanning versus the fold change in IC50 measured in traditional neutralization assays for each serum. F193F is a synonymous mutant used as a control. The Pearson correlation is indicated for each serum. See Figure S11 for full neutralization curves.

(legend continued on next page)

S11). Note that the slope of the correlation line between escape score and measured fold change in IC₅₀ varies among sera in Figure 4E because the magnitude of escape scores for this older deep mutational scanning approach is not directly comparable among sera.

Site 189 has high escape scores for five of the child sera and is neutral for the other four (Figure 4F). The difference between these two groups of children is likely due to the arrival of the N189K mutation in H3N2 HA in 2009, which was halfway between their birth and serum collection. Children with positive escape scores at site 189 have neutralizing antibodies that can bind to K189 in the parental library strain, which suggests that they were exposed to a H3N2 virus with K at site 189. By contrast, we speculate that the children who do not have positive escape scores were likely exposed to H3N2 viruses with N189, and so they do not have antibodies targeting K189 in the library strain. Note that the children with the strongest escape at sites 159–160 are all unaffected by mutations at site 189 (Figure S10).

Ferret sera have escape maps similar to children but not adults (Figures 4A, 4C, and 4D), which suggests that exposure history may be more important than species in shaping specificity. Previous work has shown that singly infected children and ferrets have similar immune specificity, based on hemagglutination inhibition titers against different H3N2 strains.⁴⁵ Our escape maps confirm that children and ferrets generate neutralizing antibodies against similar sites.

Shifts in immune specificity between 2010–2011 and 2020 cohorts reflect H3N2 evolution during this time period

All the 2010–2011 sera mapped against A/Perth/16/2009 are escaped mainly by mutations in antigenic region B, whereas the 2020 sera mapped against A/Hong Kong/45/2019 are escaped by mutations in antigenic regions A, B, C, and E to varying degrees (Figures 5A and 5B). This shift in the location of key escape mutations is likely due to the introduction of the 158 glycan in 2014, which masked nearby sites in antigenic region B²⁸ (Figure 5B). The addition of this 158 glycan may explain why the dominant escape site in A/Hong Kong/45/2019 is 189 for all age cohorts: site 189 is still in the immunodominant antigenic region B but is further from the 158 glycan (Figures 4F and 5B). Out of seven sites characterized as major determinants of H3N2 antigenic change,³² four have gained substitutions between 2010 and 2020, while 189 has not been substituted since 2009. A/Hong Kong/45/2019 has also lost glycans at sites 133 and 144 relative to A/Perth/16/2009, which may contribute to increased escape in antigenic region A in A/Hong Kong/45/2019 by unmasking this region (Figure 5B).

Most escape maps from child sera are dominated by sites 159, 160, 189, and 193 in A/Perth/16/2009. Notably, mutations at three of these four sites rose to at least 90% frequency in the next 5–10 years (Figure 5C). Some children in the 2020 cohort also have sera that are escaped by mutations at site 50. The mutation E50K began to appear in circulating viruses in 2019 and reached 90% frequency in 2023 (Figure 5C). These evolutionary

patterns suggest that the escape sites mapped here, especially from child sera, reflect antigenic pressure on influenza virus.

Site 189 in A/Hong Kong/45/2019 is generally the dominant escape site for the child sera collected in 2020 (Figure 5C). The differences in escape sites in the 2020 versus 2010–2011 child sera can be explained by mutations introduced between 2010 and 2020. While site 159 is a significant escape site for the 2010–2011 cohort, mutations at this site have almost no effect for the 2020 cohort, likely due to the 158–160 glycan hindering antibody binding in this region. Mutations at site 160 in A/Hong Kong/45/2019 removing this glycan moderately increase neutralization for many child sera, which further suggests that the glycan at site 158 is responsible for the reduced neutralizing activity targeting this region.

Similarly, the mutation F193S was introduced in 2019, and children from the 2020 cohort were likely exposed to strains with the prior genotype F193. Reverting to F193 makes A/Hong Kong/45/2019 more sensitive to neutralization by these sera, and 193D remains a strong escape mutation for both cohorts, which indicates that these children have neutralizing antibodies against site 193 in their serum repertoire (Figure 5C). In short, the neutralizing antibody response of children with limited exposure history is largely directed toward sites 160, 189, and 193 in both cohorts.

Recent H3 HA evolution correlates more with immune pressure imposed by children and teenagers than adults

These results show that influenza faces somewhat distinct antigenic selection regimes from different age groups. This raises the possibility that some mutations could be especially beneficial for the virus within subsets of the population (Figure 6A). In order to see if the actual evolution of the virus strongly favored escape mutations from specific age cohorts, we analyzed serum escape for approximately 1,200 representative H3N2 strains circulating from 2012 to 2023. The overall escape score for a given strain was calculated as the sum of the escape scores for all of its HA amino-acid mutations relative to A/Hong Kong/45/2019. For each serum, strains with negative scores are more potently neutralized than A/Hong Kong/45/2019, while strains with positive scores can better escape neutralization.

Analysis of the average strain escape scores from each cohort shows that the evolutionary trajectory of H3N2 during this time frame strongly correlates with escape from teenage sera (slope = 0.14, Pearson $r = 0.92$) (Figure 6B). Escape from child sera also correlates to a lesser degree (slope = 0.06, $r = 0.70$), although this time frame includes years prior to their birth. By contrast, mutations gained after 2020 did not substantially escape neutralization by serum from the adult cohort. We confirmed that the increase over time in average strain escape scores from teenage sera is significantly higher than adult sera ($p = 0.004$) or child sera ($p = 0.022$). When we focus the analysis on strains circulating after 2020, we see a similar trend, where the increase in average strain escape from teenage sera is significantly higher than adult sera (slope = 0.08 for teenagers, -0.14 for adults; $p = 0.002$) (Figure S12).

(F) Global frequency of site 189 H3 HA variants, compared with summed site escape scores at site 189 from 2- to 4-year-old children. Dashed lines on frequency plot indicate maximum exposure period for this cohort, from the earliest birth date to latest serum collection date. Frequency plot adapted from the Nextstrain real-time pathogen evolution website.^{43,44}

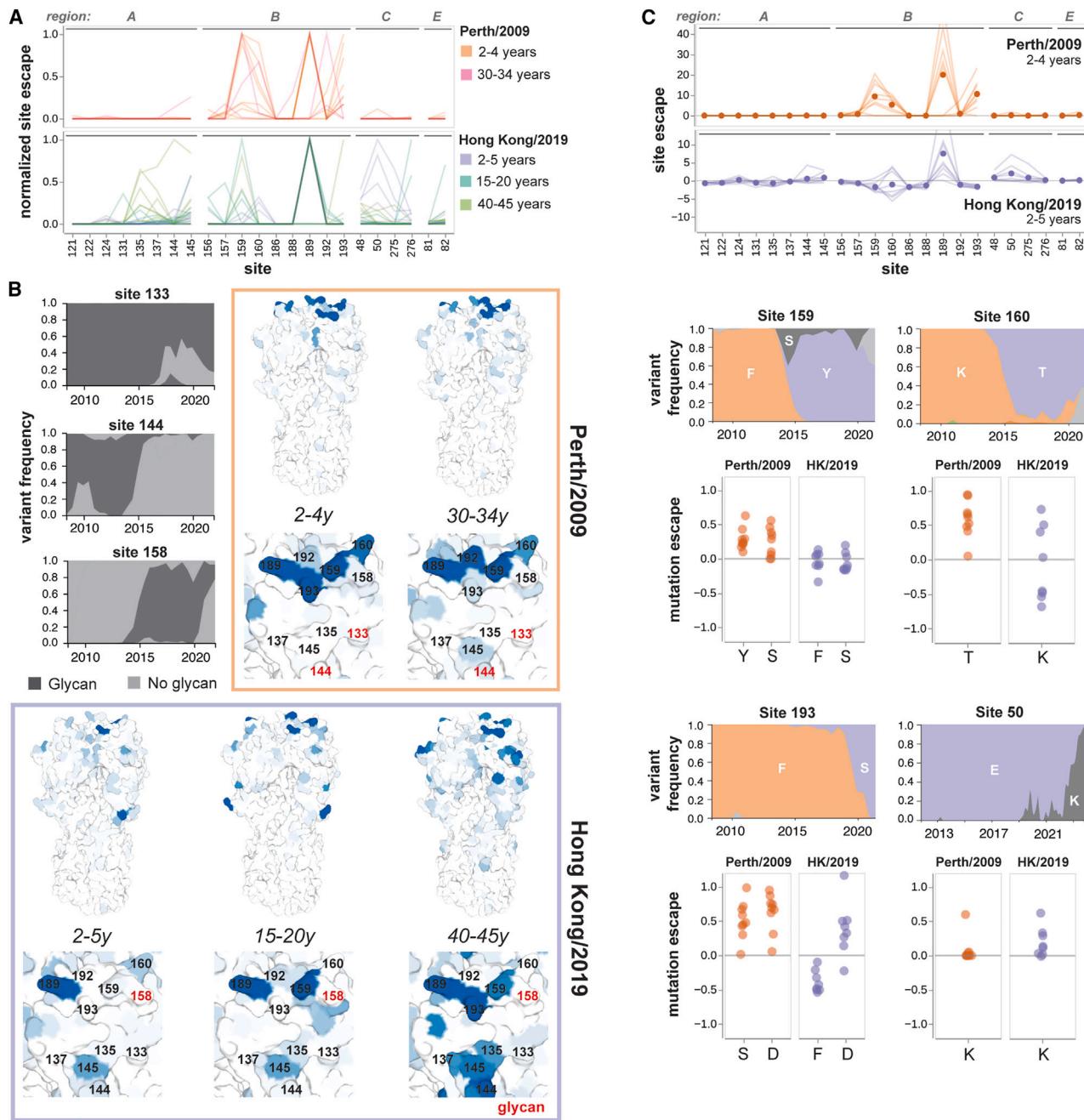


Figure 5. Comparison of serum escape maps between 2010–2011 and 2020 cohorts

(A) Normalized serum escape maps at antigenically significant sites for different individuals in each cohort, colored by age group. Each line is an individual escape map, normalized by the maximum escape for that individual.

(B) Maximal normalized escape score for each age cohort visualized on the H3 HA structure. Zoomed panels show key sites in antigenic regions A and B. Glycosylated sites (sites 133 and 144 for A/Perth/16/2009, and site 158 for A/Hong Kong/45/2019) are indicated with red text, and the global frequency of variants with these glycans over time are plotted to the left. Structures are from A/Victoria/361/2011 H3 HA (PDB: 4O5N).

(C) Changes in escape maps from young children between 2010 and 2020, compared with the emergence of new variants at those sites. Global frequency plots of H3 HA variants at sites 159, 160, and 193 are colored by the wild-type amino acid in A/Perth/16/2009 (F159, K160, F193) or A/Hong Kong/45/2019 (Y159, T160, S193). At site 50, both library strains have the same genotype (E50). Escape scores for mutations to amino acids circulating from 2010 to 2020 are plotted below. Points represent the mutation escape score measured in deep mutational scanning for each individual in the cohort. For the 2010–2011 cohort, escape scores were normalized to one at each site to facilitate comparison to escape scores from the 2020 cohort. Frequency plots in (B) and (C) were adapted from the Nextstrain real-time pathogen evolution website.^{43,44}

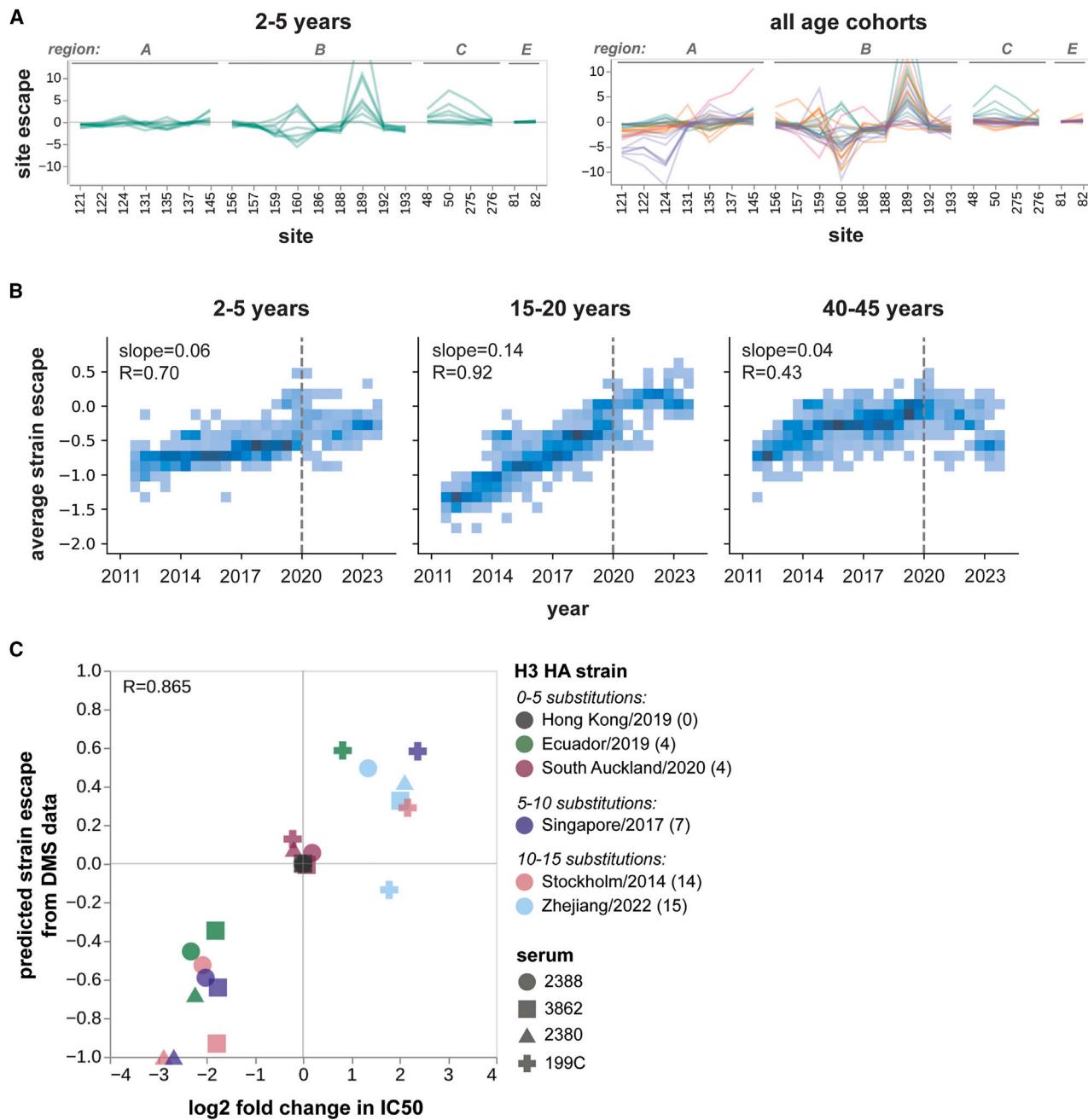


Figure 6. Relationship between escape mapped in deep mutational scanning and actual H3N2 HA evolution in humans

(A) Serum escape maps at key antigenic sites for individuals in either the 2- to 5-year-old cohort, or all age cohorts, tested against the A/Hong Kong/45/2019 HA. Each line is an individual escape map.

(B) Escape scores for H3N2 influenza strains circulating from 2012 to 2023, averaged across sera within each cohort. Strain escape is calculated based on the summed mutation escape scores for all mutations in that strain, relative to the parental library strain A/Hong Kong/45/2019, which has an escape score of zero. Plots show a 2D histogram of escape scores for approximately 1,200 H3N2 HA variants sampled during this time frame. Dashed line indicates the year of serum collection (2020). The slope of the best-fit line relating year to average strain escape (line not shown) is reported for each cohort, along with the Pearson correlation. The slope is significantly higher for the teenage cohort compared with the child ($p = 0.022$) and adult ($p = 0.004$) cohorts.

(C) Validation of predicted strain escape from deep mutational scanning measurements for circulating H3N2 strains. Strain escape is calculated by summing the effects of all mutations in that strain as measured in deep mutational scanning and is plotted against the \log_2 fold change in IC50 between the wild-type library strain and the strain of interest. Legend indicates the number of mutations relative to A/Hong Kong/45/2019 in parentheses after the strain name. R indicates the Pearson correlation. See Figure S13 for full neutralization curves, serum-level correlation plots, and exact strain names.

We also validated that we could predict the neutralization of strains with substantial numbers of HA mutations by simply adding the escape scores for their constituent mutations measured in the deep mutational scanning. Specifically, we selected five strains with between four and fifteen HA amino-acid mutations relative to the A/Hong Kong/45/2019 library strain and measured the actual changes in IC₅₀ for each strain against four sera from different age cohorts. The measured changes in IC₅₀ strongly correlate with the predictions made by summing the mutation effects measured in deep mutational scanning, even for strains with over ten mutations relative to the library strain (Figures 6C and S13).

Although the number of sera in our analysis is modest, these results show how different age cohorts are differentially affected by viral evolution and suggest that teenagers and children could play an especially prominent role in driving viral evolution.

DISCUSSION

We have mapped the mutations that escape neutralization by the polyclonal serum antibodies of 40 human individuals from different age cohorts. Sites of neutralization escape for any given serum are often highly focused, with a few specific mutations often leading to 5- to 10-fold drops in neutralization by polyclonal serum antibodies—a finding consistent with prior work.^{9,20,46,47} We also identify mutations that increase virus sensitivity to neutralization, typically by restoring epitopes present in older strains, such as removal of recently acquired N-linked glycosylation sites. The existence of these sensitizing mutations likely constrains the virus's evolution by disfavoring reversions to epitopes present in older strains.

The effects of mutations on neutralization escape differ between age cohorts in a manner consistent with likely exposure history. Many teenage sera are most escaped by mutations in antigenic region B, whereas adult sera are more strongly affected by mutations in antigenic region A, which reflects the regions that were likely immunodominant during their respective childhoods.^{31,36,37} Heterogeneity between cohorts can be highly specific: although some child and teenage sera are escaped by mutations in antigenic region C, these mutations are at site 50 for children and site 276 for teenagers. The relative number and magnitude of neutralization-sensitizing mutations also increases with age, likely due to older individuals having more antibodies from prior exposures to older strains.

The child sera analyzed in this study are escaped by mutations at a fairly narrow and homogeneous set of sites. This homogeneity suggests a relatively stereotyped initial neutralizing antibody response focused on sites near the receptor-binding pocket of HA, which is consistent with prior characterization of the antibody response in children.^{14,48,49} It is especially notable that two cohorts of children from different years, geographic regions, and vaccination status exhibit similar trends in their neutralizing antibody response. Four out of five sites targeted by children have mutated in subsequent years, suggesting that deep mutational scanning can be used to identify HA sites under strong immune pressure in contemporary human influenza strains.

Exposure history may be more important than species in shaping neutralizing specificity, as ferrets infected with A/Perth/16/2009 have similar escape mutations to children

from the 2010–2011 cohort. The similarities between ferrets and young children, but not adults with complex exposure histories, further illustrates the limitations of using primary ferret antisera as a proxy for adult human sera.^{8,9,20,50} Nevertheless, ferret sera may still be useful for identifying the selective pressures imposed by young children.⁴⁵ Interestingly, recent work has shown that vaccine effectiveness is higher in young children compared with older individuals.⁵¹

The introduction of the 158–160 glycan on the HA head in 2014 resulted in major antigenic change, as it hindered antibody binding to the immunodominant antigenic region B, leaving many individuals highly vulnerable to infection.^{27,28} It was therefore surprising that this glycan was subsequently lost in 2021 with the mutation T160I.⁵² Our analysis of sera from 2020 identifies three children whose serum neutralization is escaped by mutations at site 160, suggesting that the T160I mutation may have been gained in response to antigenic selection from young children.

Our data illustrate how mutations to HA can have heterogeneous effects on virus neutralization by sera from different individuals in the human population. The heterogeneity in escape mutations is substantially clustered by age cohort, raising the question of whether certain age groups (e.g., children, teenagers, or adults) might play an especially important role in driving viral antigenic evolution.^{14,17} Previous work has proposed that school-aged children are important drivers of influenza epidemics in communities, based on transmission modeling,^{53,54} earlier peaks of infection in this age group,¹⁸ and the impact of school closure on transmission.^{55,56} Whether the populations driving epidemics also drive antigenic drift is currently unclear. Although our study does not examine a sufficient number of sera or viruses to draw definitive conclusions, our results suggest that children and teenagers play an important role in shaping viral evolution, since mutations that fix in HA often cause more escape from child and teenage sera than adults. More extensive characterization of the major escape mutations in different age cohorts and viruses could further test this hypothesis, and any insights into which segments of the population drive antigenic evolution could be used to improve evolutionary forecasting for influenza vaccine-strain selection. More speculatively, one might imagine future strategies where vaccines are tailored to different age groups based on heterogeneity between cohorts.

Heterogeneity in the antigenic effects of mutations across the population could also have an effect on the overall rate of evolution of human influenza virus. If a specific mutation confers escape from neutralizing antibodies across all individuals, it will have a greater beneficial effect than if it only confers escape from some individuals, and this heterogeneity could impact evolutionary dynamics.^{57–60} We therefore suggest that continued characterization of population-level heterogeneity in viral escape mutations, and incorporation of these measurements into models of evolution, could shed further light on the evolutionary dynamics of influenza and other viruses.

Limitations of the study

Our experiments analyzed sera from a limited number of individuals from each age cohort, and these cohorts were sampled from a single city for each year (Seattle for the 2020 cohort and Ho Chi Minh City for the 2010–2011 cohort). While this sample size was sufficient to illustrate heterogeneity in serum antibody

targeting between age cohorts, larger sample sizes are needed to draw firm conclusions about how the heterogeneity in the antigenic effects of mutations shapes influenza evolution.

Additionally, our comparison to natural evolution (which uses data from the 2020 Seattle cohort) is based largely on vaccinated individuals. Although there is evidence that natural infection plays a stronger role than vaccination in shaping long-term humoral immune responses,^{61–68} we cannot rule out systematic differences between vaccinated and unvaccinated populations, which may further impact evolutionary dynamics.

A further caveat is that we focused on sera that had high neutralizing titers against the vaccine strains used to build our libraries, in large part to reduce the serum volumes needed to perform the deep mutational scans. Therefore, if the antigenic effects of mutations vary systematically with the overall serum neutralizing titer, then our results might only reflect the subset of the population with high titers.

Our study coincides with the beginning of the SARS-CoV-2 pandemic, which may limit the generalizability of results measured from the 2020 cohort. More specifically, differences in social distancing measures implemented for children and adults may have influenced which age groups were stronger drivers of influenza evolution in subsequent years. Information on sex, gender, ancestry, race, ethnicity, and socioeconomic status for study participants was not available and has not been accounted for in our analysis, which may also limit its generalizability.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
 - Biosafety
 - Human sera
- METHOD DETAILS
 - Data availability statement
 - A/Perth/16/2009 and A/Hong Kong/45/2019 library overview
 - Design of barcoded H3 plasmid with duplicated WSN-HA packaging signals
 - Design of barcoded H3 deep mutational scanning library
 - Generation of plasmid library of barcoded H3 HA variants
 - PacBio sequencing of H3 HA plasmid library
 - Generation of H3 HA variant library in the context of lab-adapted WSN virus
 - Design of external neutralization standards
 - Generation of WSN-H6 HA neutralization standard
 - Generation of RNA spike-in neutralization standard
 - Serum preparation
 - GFP neutralization assays for serum screening and validation
 - Serum selection of library escape variants
 - Barcode amplification from infected cells for Illumina sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Computational pipeline overview
 - PacBio sequencing data analysis
 - Illumina barcode sequencing data analysis
 - Modeling the effects of mutations on H3 HA function
 - Modeling the effects of mutations on antibody and serum escape
 - Estimating escape scores for circulating H3N2 strains

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.chom.2024.06.015>.

ACKNOWLEDGMENTS

We thank Sarah Cobey for scientific advice. We gratefully acknowledge all data contributors, including 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 part of this research is based. This work was supported in part by the NIH/NIAID award R01AI165821 to T.B. and J.D.B., contract 75N93021C00015 to J.D.B. and T.B., and award R01AI146028 to F.A.M. Serum collection in Vietnam was funded by a Wellcome Trust/Royal Society Sir Henry Dale Fellowship (098511/Z/12/Z) to M.F.B. T.C.Y. was supported by the CMB training grant T32 GM007270 and the NSF graduate research fellowship DGE-2140004. J.D.B., T.B., and F.A.M. are Investigators of the Howard Hughes Medical Institute. This research was supported by the Genomics & Bioinformatics Shared Resource (RRID:SCR_022606) of the Fred Hutch/University of Washington/Seattle Children's Cancer Consortium (P30CA015704) and by Fred Hutch Scientific Computing (NIH grants S10-OD-020069 and S10-OD-028685).

AUTHOR CONTRIBUTIONS

Conceptualization, F.C.W., R.T.E., J.M.L., and J.D.B.; methodology, F.C.W., A.N.L., R.T.E., J.M.L., and J.D.B.; software, J.D.B., F.C.W., H.K.H., J.G., J.H., T.C.Y., T.B., and F.A.M.; investigation, F.C.W., R.T.E., and J.M.L.; validation, F.C.W. and R.T.E.; resources, N.V.V.C., M.Q.L., N.T.D.N., N.T.L.T., A.L.G., H.Y.C., J.A.E., and M.F.B.; writing – original draft, F.C.W. and J.D.B.; writing – review and editing, F.C.W., J.D.B., J.M.L., H.K.H., A.N.L., T.C.Y., J.H., A.L.G., H.Y.C., T.B., F.A.M., and M.F.B.; supervision, J.D.B.; funding acquisition, J.D.B.

DECLARATION OF INTERESTS

J.D.B. is on the scientific advisory boards of Apriori Bio, Invivyd, Aerium Therapeutics, and the Vaccine Company. J.D.B., A.N.L., and F.C.W. receive royalty payments as inventors on Fred Hutch licensed patents related to viral deep mutational scanning. H.Y.C. reports consulting with Ellume, Pfizer, and the Bill and Melinda Gates Foundation. She has served on advisory boards for Vir, Merck, and Abbvie. She has conducted CME teaching with Medscape, Vindico, and Clinical Care Options. She has received research funding from Gates Ventures and support and reagents from Ellume and Cepheid outside of the submitted work. J.A.E. receives institutional funding from AstraZeneca, Merck, GlaxoSmithKline, and Pfizer. She is a consultant for Abbvie, AstraZeneca, GlaxoSmithKline, Meissa Vaccines, Moderna, Pfizer, and Sanofi Pasteur. A.L.G. reports contract testing from Abbott, Cepheid, Novavax, Pfizer, Janssen, and Hologic and research support from Gilead outside of the submitted work.

Received: December 12, 2023

Revised: April 19, 2024

Accepted: June 25, 2024

Published: July 19, 2024

REFERENCES

1. Ranjeva, S., Subramanian, R., Fang, V.J., Leung, G.M., Ip, D.K.M., Perera, R.A.P.M., Peiris, J.S.M., Cowling, B.J., and Cobey, S. (2019). Age-specific differences in the dynamics of protective immunity to influenza. *Nat. Commun.* **10**, 1660. <https://doi.org/10.1038/s41467-019-09652-6>.
2. Kucharski, A.J., Lessler, J., Read, J.M., Zhu, H., Jiang, C.Q., Guan, Y., Cummings, D.A.T., and Riley, S. (2015). Estimating the life course of influenza A(H3N2) antibody responses from cross-sectional data. *PLoS Biol.* **13**, e1002082. <https://doi.org/10.1371/journal.pbio.1002082>.

3. Smith, D.J., Lapedes, A.S., de Jong, J.C., Bestebroer, T.M., Rimmelzwaan, G.F., Osterhaus, A.D.M.E., and Fouchier, R.A.M. (2004). Mapping the antigenic and genetic evolution of influenza virus. *Science* 305, 371–376. <https://doi.org/10.1126/science.1097211>.
4. Bedford, T., Suchard, M.A., Lemey, P., Dudas, G., Gregory, V., Hay, A.J., McCauley, J.W., Russell, C.A., Smith, D.J., and Rambaut, A. (2014). Integrating influenza antigenic dynamics with molecular evolution. *eLife* 3, e01914. <https://doi.org/10.7554/eLife.01914>.
5. Fitch, W.M., Bush, R.M., Bender, C.A., and Cox, N.J. (1997). Long term trends in the evolution of H(3) HA1 human influenza type A. *Proc. Natl. Acad. Sci. USA* 94, 7712–7718. <https://doi.org/10.1073/pnas.94.15.7712>.
6. Strelkowa, N., and Lässig, M. (2012). Clonal interference in the evolution of influenza. *Genetics* 192, 671–682. <https://doi.org/10.1534/genetics.112.143396>.
7. Francis, T. (1960). On the doctrine of original antigenic sin. *Proc. Am. Philos. Soc.* 104, 572–578.
8. Fonville, J.M., Wilks, S.H., James, S.L., Fox, A., Ventresca, M., Aban, M., Xue, L., Jones, T.C., Le, N.M.H., Pham, Q.T., et al. (2014). Antibody landscapes after influenza virus infection or vaccination. *Science* 346, 996–1000. <https://doi.org/10.1126/science.1256427>.
9. Linderman, S.L., Chambers, B.S., Zost, S.J., Parkhouse, K., Li, Y., Herrmann, C., Ellebedy, A.H., Carter, D.M., Andrews, S.F., Zheng, N.-Y., et al. (2014). Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013–2014 influenza season. *Proc. Natl. Acad. Sci. USA* 111, 15798–15803. <https://doi.org/10.1073/pnas.1409171111>.
10. Skowronski, D.M., Chambers, C., Sabaiduc, S., De Serres, G., Winter, A.-L., Dickinson, J.A., Gubbay, J.B., Drews, S.J., Martineau, C., Charest, H., et al. (2017). Beyond antigenic match: possible Agent-Host and immuno-epidemiological influences on influenza vaccine effectiveness during the 2015–2016 season in Canada. *J. Infect. Dis.* 216, 1487–1500. <https://doi.org/10.1093/infdis/jix526>.
11. Krammer, F. (2019). The human antibody response to influenza A virus infection and vaccination. *Nat. Rev. Immunol.* 19, 383–397. <https://doi.org/10.1038/s41577-019-0143-6>.
12. Lessler, J., Riley, S., Read, J.M., Wang, S., Zhu, H., Smith, G.J.D., Guan, Y., Jiang, C.Q., and Cummings, D.A.T. (2012). Evidence for antigenic seniority in influenza A (H3N2) antibody responses in southern China. *PLoS Pathog.* 8, e1002802. <https://doi.org/10.1371/journal.ppat.1002802>.
13. Yang, B., Lessler, J., Zhu, H., Jiang, C.Q., Read, J.M., Hay, J.A., Kwok, K.O., Shen, R., Guan, Y., Riley, S., et al. (2020). Life course exposures continually shape antibody profiles and risk of seroconversion to influenza. *PLoS Pathog.* 16, e1008635. <https://doi.org/10.1371/journal.ppat.1008635>.
14. Nakajima, S., Nobusawa, E., and Nakajima, K. (2000). Variation in response among individuals to antigenic sites on the HA protein of human influenza virus may be responsible for the emergence of drift strains in the human population. *Virology* 274, 220–231. <https://doi.org/10.1006/viro.2000.0453>.
15. Sato, K., Morishita, T., Nobusawa, E., Tonegawa, K., Sakae, K., Nakajima, S., and Nakajima, K. (2004). Amino-acid change on the antigenic region B1 of H3 haemagglutinin may be a trigger for the emergence of drift strain of influenza A virus. *Epidemiol. Infect.* 132, 399–406. <https://doi.org/10.1017/s0950268803001821>.
16. Oidtman, R.J., Arevalo, P., Bi, Q., McGough, L., Russo, C.J., Vera Cruz, D., Costa Vieira, M., and Gostic, K.M. (2021). Influenza immune escape under heterogeneous host immune histories. *Trends Microbiol.* 29, 1072–1082. <https://doi.org/10.1016/j.tim.2021.05.009>.
17. Kim, K., Gouma, S., Vieira, M.C., Weirick, M.E., Hensley, S.E., and Cobey, S. (2023). Measures of population immunity can predict the dominant clade of influenza A (H3N2) and reveal age-associated differences in susceptibility and specificity. Preprint at medRxiv. <https://doi.org/10.1101/2023.10.26.23297569>.
18. Worby, C.J., Chaves, S.S., Wallinga, J., Lipsitch, M., Finelli, L., and Goldstein, E. (2015). On the relative role of different age groups in influenza epidemics. *Epidemics* 13, 10–16. <https://doi.org/10.1016/j.epidem.2015.04.003>.
19. Arevalo, P., McLean, H.Q., Belongia, E.A., and Cobey, S. (2020). Earliest infections predict the age distribution of seasonal influenza A cases. *eLife* 9, e50060. <https://doi.org/10.7554/eLife.50060>.
20. Lee, J.M., Eguia, R., Zost, S.J., Choudhary, S., Wilson, P.C., Bedford, T., Stevens-Ayers, T., Boekh, M., Hurt, A.C., Lakdawala, S.S., et al. (2019). Mapping person-to-person variation in viral mutations that escape polyclonal serum targeting influenza hemagglutinin. *eLife* 8, e49324. <https://doi.org/10.7554/eLife.49324>.
21. Doud, M.B., and Bloom, J.D. (2016). Accurate measurement of the effects of all amino-acid mutations on influenza hemagglutinin. *Viruses* 8, 155. <https://doi.org/10.3390/v8060155>.
22. Radford, C.E., Schommers, P., Giesemann, L., Crawford, K.H.D., Dadonaita, B., Yu, T.C., Dingens, A.S., Overbaugh, J., Klein, F., and Bloom, J.D. (2023). Mapping the neutralizing specificity of human anti-HIV serum by deep mutational scanning. *Cell Host Microbe* 31, 1200–1215.e9.
23. Yu, T.C., Thornton, Z.T., Hannon, W.W., DeWitt, W.S., Radford, C.E., Matsen, F.A., 4th, and Bloom, J.D. (2022). A biophysical model of viral escape from polyclonal antibodies. *Virus Evol.* 8, veac110. <https://doi.org/10.1093/ve/veac110>.
24. Hensley, S.E., Das, S.R., Bailey, A.L., Schmidt, L.M., Hickman, H.D., Jayaraman, A., Viswanathan, K., Raman, R., Sasisekharan, R., Bennink, J.R., et al. (2009). Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. *Science* 326, 734–736. <https://doi.org/10.1126/science.1178258>.
25. Matrosovich, M., Tuzikov, A., Bovin, N., Gambaryan, A., Klimov, A., Castrucci, M.R., Donatelli, I., and Kawaoka, Y. (2000). Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. *J. Virol.* 74, 8502–8512. <https://doi.org/10.1128/jvi.74.18.8502-8512.2000>.
26. Lin, Y.P., Xiong, X., Wharton, S.A., Martin, S.R., Coombs, P.J., Vachieri, S.G., Christodoulou, E., Walker, P.A., Liu, J., Skehel, J.J., et al. (2012). Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. *Proc. Natl. Acad. Sci. USA* 109, 21474–21479. <https://doi.org/10.1073/pnas.1218841110>.
27. Zost, S.J., Parkhouse, K., Gumina, M.E., Kim, K., Diaz Perez, S., Wilson, P.C., Treanor, J.J., Sant, A.J., Cobey, S., and Hensley, S.E. (2017). Contemporary H3N2 influenza viruses have a glycosylation site that alters binding of antibodies elicited by egg-adapted vaccine strains. *Proc. Natl. Acad. Sci. USA* 114, 12578–12583. <https://doi.org/10.1073/pnas.1712377114>.
28. Gouma, S., Kim, K., Weirick, M.E., Gumina, M.E., Branche, A., Topham, D.J., Martin, E.T., Monto, A.S., Cobey, S., and Hensley, S.E. (2020). Middle-aged individuals may be in a perpetual state of H3N2 influenza virus susceptibility. *Nat. Commun.* 11, 4566. <https://doi.org/10.1038/s41467-020-18465-x>.
29. Muñoz, E.T., and Deem, M.W. (2005). Epitope analysis for influenza vaccine design. *Vaccine* 23, 1144–1148. <https://doi.org/10.1016/j.vaccine.2004.08.028>.
30. Wiley, D.C., Wilson, I.A., and Skehel, J.J. (1981). Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* 289, 373–378. <https://doi.org/10.1038/289373a0>.
31. Popova, L., Smith, K., West, A.H., Wilson, P.C., James, J.A., Thompson, L.F., and Air, G.M. (2012). Immunodominance of antigenic site B over site A of hemagglutinin of recent H3N2 influenza viruses. *PLoS One* 7, e41895. <https://doi.org/10.1371/journal.pone.0041895>.
32. Koel, B.F., Burke, D.F., Bestebroer, T.M., van der Vliet, S., Zondag, G.C.M., Vervaet, G., Skepner, E., Lewis, N.S., Spronken, M.I.J., Russell, C.A., et al. (2013). Substitutions near the receptor binding site determine

- major antigenic change during influenza virus evolution. *Science* 342, 976–979. <https://doi.org/10.1126/science.1244730>.
33. Angeletti, D., Gibbs, J.S., Angel, M., Kosik, I., Hickman, H.D., Frank, G.M., Das, S.R., Wheatley, A.K., Prabhakaran, M., Leggat, D.J., et al. (2017). Defining B cell immunodominance to viruses. *Nat. Immunol.* 18, 456–463. <https://doi.org/10.1038/ni.3680>.
 34. Zost, S.J., Lee, J., Gumina, M.E., Parkhouse, K., Henry, C., Wu, N.C., Lee, C.D., Wilson, I.A., Wilson, P.C., Bloom, J.D., et al. (2019). Identification of antibodies Targeting the H3N2 Hemagglutinin Receptor Binding Site following Vaccination of Humans. *Cell Rep.* 29, 4460–4470.e8. <https://doi.org/10.1016/j.celrep.2019.11.084>.
 35. Guthmiller, J.J., Han, J., Li, L., Freyn, A.W., Liu, S.T.H., Stovicek, O., Stamper, C.T., Dugan, H.L., Tepora, M.E., Utset, H.A., et al. (2021). First exposure to the pandemic H1N1 virus induced broadly neutralizing antibodies targeting hemagglutinin head epitopes. *Sci. Transl. Med.* 13, eabg4535. <https://doi.org/10.1126/scitranslmed.abg4535>.
 36. Nobusawa, E., Omagari, K., Nakajima, S., and Nakajima, K. (2012). Reactivity of human convalescent sera with influenza virus hemagglutinin protein mutants at antigenic site A. *Microbiol. Immunol.* 56, 99–106. <https://doi.org/10.1111/j.1348-0421.2012.00412.x>.
 37. Chambers, B.S., Parkhouse, K., Ross, T.M., Alby, K., and Hensley, S.E. (2015). Identification of hemagglutinin residues responsible for H3N2 antigenic drift during the 2014–2015 influenza season. *Cell Rep.* 12, 1–6. <https://doi.org/10.1016/j.celrep.2015.06.005>.
 38. Wu, N.C., and Wilson, I.A. (2017). A perspective on the structural and functional constraints for immune evasion: insights from influenza virus. *J. Mol. Biol.* 429, 2694–2709. <https://doi.org/10.1016/j.jmb.2017.06.015>.
 39. Gouma, S., Weirick, M., and Hensley, S.E. (2020). Antigenic assessment of the H3N2 component of the 2019–2020 Northern Hemisphere influenza vaccine. *Nat. Commun.* 11, 2445. <https://doi.org/10.1038/s41467-020-16183-y>.
 40. Nunes, M.C., Cutland, C.L., Dighero, B., Bate, J., Jones, S., Hugo, A., van Niekerk, N., Kuwanda, L., Izu, A., Weinberg, A., et al. (2015). Kinetics of hemagglutination-inhibiting antibodies following maternal influenza vaccination among Mothers With and those without HIV infection and their infants. *J. Infect. Dis.* 212, 1976–1987. <https://doi.org/10.1093/infdis/jiv339>.
 41. Nhat, N.T.D., Todd, S., de Bruin, E., Thao, T.T.N., Vy, N.H.T., Quan, T.M., Vinh, D.N., van Beek, J., Anh, P.H., Lam, H.M., et al. (2017). Structure of general-population antibody titer distributions to influenza A virus. *Sci. Rep.* 7, 6060. <https://doi.org/10.1038/s41598-017-06177-0>.
 42. Lee, J.M., Huddleston, J., Doud, M.B., Hooper, K.A., Wu, N.C., Bedford, T., and Bloom, J.D. (2018). Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. *Proc. Natl. Acad. Sci. USA* 115, E8276–E8285. <https://doi.org/10.1073/pnas.1806133115>.
 43. Hadfield, J., Megill, C., Bell, S.M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T., and Neher, R.A. (2018). Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34, 4121–4123. <https://doi.org/10.1093/bioinformatics/bty407>.
 44. Neher, R.A., and Bedford, T. (2015). nextflu: real-time tracking of seasonal influenza virus evolution in humans. *Bioinformatics* 31, 3546–3548. <https://doi.org/10.1093/bioinformatics/btv381>.
 45. Fonville, J.M., Fraaij, P.L.A., de Mutsert, G., Wilks, S.H., van Beek, R., Fouchier, R.A.M., and Rimmelzwaan, G.F. (2016). Antigenic maps of influenza A(H3N2) Produced With Human Antisera Obtained After Primary Infection. *J. Infect. Dis.* 213, 31–38. <https://doi.org/10.1093/infdis/jiv367>.
 46. Huang, K.-Y.A., Rijal, P., Schimanski, L., Powell, T.J., Lin, T.-Y., McCauley, J.W., Daniels, R.S., and Townsend, A.R. (2015). Focused antibody response to influenza linked to antigenic drift. *J. Clin. Invest.* 125, 2631–2645. <https://doi.org/10.1172/JCI81104>.
 47. Davis, A.K.F., McCormick, K., Gumina, M.E., Petrie, J.G., Martin, E.T., Xue, K.S., Bloom, J.D., Monto, A.S., Bushman, F.D., and Hensley, S.E. (2018). Sera from individuals with narrowly focused influenza virus antibodies rapidly select viral escape mutations in ovo. *J. Virol.* 92, e00859. <https://doi.org/10.1128/JVI.00859-18>.
 48. Meade, P., Kuan, G., Strohmeier, S., Maier, H.E., Amanat, F., Balmaseda, A., Ito, K., Kirkpatrick, E., Javier, A., Gresh, L., et al. (2020). Influenza virus infection induces a narrow antibody response in children but a broad recall response in adults. *mBio* 11, e03243. <https://doi.org/10.1128/mBio.03243-19>.
 49. Islam, S., Mohn, K.G.-I., Krammer, F., Sanne, M., Bredholt, G., Jul-Larsen, Å., Tete, S.M., Zhou, F., Brokstad, K.A., and Cox, R.J. (2017). Influenza A haemagglutinin specific IgG responses in children and adults after seasonal trivalent live attenuated influenza vaccination. *Vaccine* 35, 5666–5673. <https://doi.org/10.1016/j.vaccine.2017.08.044>.
 50. Li, Y., Myers, J.L., Bostick, D.L., Sullivan, C.B., Madara, J., Linderman, S.L., Liu, Q., Carter, D.M., Wrammert, J., Esposito, S., et al. (2013). Immune history shapes specificity of pandemic H1N1 influenza antibody responses. *J. Exp. Med.* 210, 1493–1500. <https://doi.org/10.1084/jem.20130212>.
 51. Hood, N., Flannery, B., Gaglani, M., Beeram, M., Wernli, K., Jackson, M.L., Martin, E.T., Monto, A.S., Zimmerman, R., Raviotta, J., et al. (2023). Influenza vaccine effectiveness among children: 2011–2020. *Pediatrics* 151, e2022059922. <https://doi.org/10.1542/peds.2022-059922>.
 52. Bolton, M.J., Ort, J.T., McBride, R., Swanson, N.J., Wilson, J., Awofolaju, M., Furey, C., Greenplate, A.R., Drapeau, E.M., Pekosz, A., et al. (2022). Antigenic and virological properties of an H3N2 variant that continues to dominate the 2021–22 Northern Hemisphere influenza season. *Cell Rep.* 39, 110897. <https://doi.org/10.1016/j.celrep.2022.110897>.
 53. Basta, N.E., Chao, D.L., Halloran, M.E., Matrajt, L., and Longini, I.M., Jr. (2009). Strategies for pandemic and seasonal influenza vaccination of schoolchildren in the United States. *Am. J. Epidemiol.* 170, 679–686. <https://doi.org/10.1093/aje/kwp237>.
 54. Wallinga, J., Teunis, P., and Kretzschmar, M. (2006). Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. J. Epidemiol.* 164, 936–944. <https://doi.org/10.1093/aje/kwj317>.
 55. Cauchemez, S., Valleron, A.-J., Boëlle, P.-Y., Flahault, A., and Ferguson, N.M. (2008). Estimating the impact of school closure on influenza transmission from Sentinel data. *Nature* 452, 750–754. <https://doi.org/10.1038/nature06732>.
 56. Huang, K.E., Lipsitch, M., Shaman, J., and Goldstein, E. (2014). The US 2009 A(H1N1) influenza epidemic: quantifying the impact of school openings on the reproductive number. *Epidemiology* 25, 203–206. <https://doi.org/10.1097/EDE.0000000000000055>.
 57. Cobey, S., and Pascual, M. (2011). Consequences of host heterogeneity, epitope immunodominance, and immune breadth for strain competition. *J. Theor. Biol.* 270, 80–87. <https://doi.org/10.1016/j.jtbi.2010.11.009>.
 58. Gupta, S., and Galvani, A. (1999). The effects of host heterogeneity on pathogen population structure. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 354, 711–719. <https://doi.org/10.1098/rstb.1999.0424>.
 59. Lukasz, M., and Lässig, M. (2014). A predictive fitness model for influenza. *Nature* 507, 57–61. <https://doi.org/10.1038/nature13087>.
 60. Desai, M.M., and Fisher, D.S. (2007). Beneficial mutation selection balance and the effect of linkage on positive selection. *Genetics* 176, 1759–1798. <https://doi.org/10.1534/genetics.106.067678>.
 61. Davis, C.W., Jackson, K.J.L., McCausland, M.M., Darce, J., Chang, C., Linderman, S.L., Chennareddy, C., Gerkin, R., Brown, S.J., Wrammert, J., et al. (2020). Influenza vaccine-induced human bone marrow plasma cells decline within a year after vaccination. *Science* 370, 237–241. <https://doi.org/10.1126/science.aaz8432>.
 62. Yu, X., Tsibane, T., McGraw, P.A., House, F.S., Keefer, C.J., Hicar, M.D., Tumpey, T.M., Pappas, C., Perrone, L.A., Martinez, O., et al. (2008). Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. *Nature* 455, 532–536. <https://doi.org/10.1038/nature07231>.
 63. Margine, I., Hai, R., Albrecht, R.A., Obermoser, G., Harrod, A.C., Banchereau, J., Palucka, K., García-Sastre, A., Palese, P., Treanor, J.J., et al. (2013). H3N2 influenza virus infection induces broadly reactive

- hemagglutinin stalk antibodies in humans and mice. *J. Virol.* 87, 4728–4737. <https://doi.org/10.1128/JVI.03509-12>.
64. Belongia, E.A., Sundaram, M.E., McClure, D.L., Meece, J.K., Ferdinands, J., and VanWormer, J.J. (2015). Waning vaccine protection against influenza A (H3N2) illness in children and older adults during a single season. *Vaccine* 33, 246–251. <https://doi.org/10.1016/j.vaccine.2014.06.052>.
 65. Fisman, D.N., Savage, R., Gubbay, J., Achonu, C., Akwar, H., Farrell, D.J., Crowcroft, N.S., and Jackson, P. (2009). Older age and a reduced likelihood of 2009 H1N1 virus infection. *N. Engl. J. Med.* 361, 2000–2001. <https://doi.org/10.1056/NEJMc0907256>.
 66. Hancock, K., Veguilla, V., Lu, X., Zhong, W., Butler, E.N., Sun, H., Liu, F., Dong, L., DeVos, J.R., Gargiullo, P.M., et al. (2009). Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. *N. Engl. J. Med.* 361, 1945–1952. <https://doi.org/10.1056/NEJMoa0906453>.
 67. Kissling, E., Valenciano, M., Larrauri, A., Oroszi, B., Cohen, J.M., Nunes, B., Pitigoi, D., Rizzo, C., Rebolledo, J., Paradowska-Stankiewicz, I., et al. (2013). Low and decreasing vaccine effectiveness against influenza A(H3) in 2011/12 among vaccination target groups in Europe: results from the I-MOVE multicentre case-control study. *Euro Surveill.* 18, 20390. <https://doi.org/10.2807/ese.18.05.20390-en>.
 68. Petrie, J.G., Ohmit, S.E., Johnson, E., Truscon, R., and Monto, A.S. (2015). Persistence of antibodies to influenza hemagglutinin and neuraminidase following one or two years of influenza vaccination. *J. Infect. Dis.* 212, 1914–1922. <https://doi.org/10.1093/infdis/jiv313>.
 69. Bloom, J.D., Gong, L.I., and Baltimore, D. (2010). Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. *Science* 328, 1272–1275. <https://doi.org/10.1126/science.1187816>.
 70. Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G., and Webster, R.G. (2000). A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc. Natl. Acad. Sci. USA* 97, 6108–6113. <https://doi.org/10.1073/pnas.100133697>.
 71. Sandbulte, M.R., Gao, J., Straight, T.M., and Eichelberger, M.C. (2009). A miniaturized assay for influenza neuraminidase-inhibiting antibodies utilizing reverse genetics-derived antigens. *Influenza other Respi. Viruses* 3, 233–240.
 72. Crawford, K.H.D., and Bloom, J.D. (2019). alignparse: A Python package for parsing complex features from high-throughput long-read sequencing. *J. Open Source Softw.* 4, 4. <https://doi.org/10.21105/joss.01915>.
 73. Boni, M.F., Chau, N.V.V., Dong, N., Todd, S., Nhat, N.T.D., de Bruin, E., van Beek, J., Hien, N.T., Simmons, C.P., Farrar, J., et al. (2013). Population-level antibody estimates to novel influenza A/H7N9. *J. Infect. Dis.* 208, 554–558. <https://doi.org/10.1093/infdis/jit224>.
 74. Berto, A., Pham, H.A., Thao, T.T.N., Vy, N.H.T., Caddy, S.L., Hiraide, R., Tue, N.T., Goodfellow, I., Carrique-Mas, J.J., Thwaites, G.E., et al. (2018). Hepatitis E in southern Vietnam: seroepidemiology in humans and molecular epidemiology in pigs. *Zoonoses Public Health* 65, 43–50. <https://doi.org/10.1111/zph.12364>.
 75. Quan, T.M., Phuong, H.T., Vy, N.H.T., Thanh, N.T.L., Lien, N.T.N., Hong, T.T.K., Dung, P.N., Chau, N.V.V., Boni, M.F., and Clapham, H.E. (2018). Evidence of previous but not current transmission of Chikungunya virus in southern and central Vietnam: results from a systematic review and a seroprevalence study in four locations. *PLoS Negl. Trop. Dis.* 12, e0006246. <https://doi.org/10.1371/journal.pntd.0006246>.
 76. Lam, H.M., Phuong, H.T., Thao Vy, N.H., Le Thanh, N.T., Dung, P.N., Ngoc Muon, T.T., Van Vinh Chau, N., Rodríguez-Barraquer, I., Cummings, D.A.T., Wills, B.A., et al. (2019). Serological inference of past primary and secondary dengue infection: implications for vaccination. *J. R. Soc. Interface* 16, 20190207. <https://doi.org/10.1098/rsif.2019.0207>.
 77. Vinh, D.N., Nhat, N.T.D., de Bruin, E., Vy, N.H.T., Thao, T.T.N., Phuong, H.T., Anh, P.H., Todd, S., Quan, T.M., Thanh, N.T.L., et al. (2021). Age-seroprevalence curves for the multi-strain structure of influenza A virus. *Nat. Commun.* 12, 6680. <https://doi.org/10.1038/s41467-021-26948-8>.
 78. Thwaites, C.L., Thanh, T.T., Ny, N.T.H., Nguyet, L.A., Nhat, N.T.D., Thuy, C.T., Thanh, N.T.L., Dung, N.T.; OUCRU; Sero-Surveillance Network, and Campbell, J., et al. (2023). Seroprotection against tetanus in southern Vietnam. *Vaccine* 41, 2208–2213. <https://doi.org/10.1016/j.vaccine.2023.02.036>.
 79. Li, X., Gu, M., Zheng, Q., Gao, R., and Liu, X. (2021). Packaging signal of influenza A virus. *Virol. J.* 18, 36. <https://doi.org/10.1186/s12985-021-01504-4>.
 80. Gao, Q., and Palese, P. (2009). Rewiring the RNAs of influenza virus to prevent reassortment. *Proc. Natl. Acad. Sci. USA* 106, 15891–15896. <https://doi.org/10.1073/pnas.0908897106>.
 81. Zhang, J., Leser, G.P., Pekosz, A., and Lamb, R.A. (2000). The cytoplasmic tails of the influenza virus spike glycoproteins are required for normal genome packaging. *Virology* 269, 325–334. <https://doi.org/10.1006/viro.2000.0228>.
 82. Wu, N.C., Xie, J., Zheng, T., Nycholat, C.M., Grande, G., Paulson, J.C., Lerner, R.A., and Wilson, I.A. (2017). Diversity of functionally permissive sequences in the receptor-binding site of influenza hemagglutinin. *Cell Host Microbe* 21, 742–753.e8.
 83. Neumann, G., Watanabe, T., Ito, H., Watanabe, S., Goto, H., Gao, P., Hughes, M., Perez, D.R., Donis, R., Hoffmann, E., et al. (1999). Generation of influenza A viruses entirely from cloned cDNAs. *Proc. Natl. Acad. Sci. USA* 96, 9345–9350. <https://doi.org/10.1073/pnas.96.16.9345>.
 84. Bloom, J.D. (2014). An experimentally determined evolutionary model dramatically improves phylogenetic fit. *Mol. Biol. Evol.* 31, 1956–1978. <https://doi.org/10.1093/molbev/msu173>.
 85. Starr, T.N., Greaney, A.J., Hilton, S.K., Ellis, D., Crawford, K.H.D., Dingens, A.S., Navarro, M.J., Bowen, J.E., Tortorici, M.A., Walls, A.C., et al. (2020). Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* 182, 1295–1310.e20. <https://doi.org/10.1016/j.cell.2020.08.012>.
 86. Hooper, K.A., and Bloom, J.D. (2013). A mutant influenza virus that uses an N1 neuraminidase as the receptor-binding protein. *J. Virol.* 87, 12531–12540. <https://doi.org/10.1128/JVI.01889-13>.
 87. Doud, M.B., Lee, J.M., and Bloom, J.D. (2018). How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. *Nat. Commun.* 9, 1386. <https://doi.org/10.1038/s41467-018-03665-3>.
 88. Bacsik, D.J., Dadonaite, B., Butler, A., Greaney, A.J., Heaton, N.S., and Bloom, J.D. (2023). Influenza virus transcription and progeny production are poorly correlated in single cells. *eLife* 12, RP86852. <https://doi.org/10.7554/eLife.86852>.
 89. Sailer, Z.R., and Harms, M.J. (2017). Detecting high-order epistasis in nonlinear genotype-phenotype maps. *Genetics* 205, 1079–1088. <https://doi.org/10.1534/genetics.116.195214>.
 90. Otwinowski, J., McCandlish, D.M., and Plotkin, J.B. (2018). Inferring the shape of global epistasis. *Proc. Natl. Acad. Sci. USA* 115, E7550–E7558. <https://doi.org/10.1073/pnas.1804015115>.
 91. Shu, Y., and McCauley, J. (2017). GISAID: global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* 22, 30494. <https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
10-beta electrocompetent <i>E. coli</i>	NEB	C3020K
5 α high efficiency <i>E. coli</i>	NEB	C2987
Biological samples		
Residual human sera collected in Seattle from individuals aged 0-20 years	Seattle Children's Hospital	https://www.seattlechildrens.org/research/centers-programs/clinical-and-translational-research/
Human sera collected in Seattle from individuals aged 40-45 years	Hospitalized or Ambulatory Adults with Respiratory Viral Infections (HAARVI) cohort	https://www.chulab.org/haarvi
Residual human sera collected in Seattle from a 68-year-old individual	University of Washington Virology Lab	https://depts.washington.edu/uwviro/
Residual human sera collected in Ho Chi Minh City, Vietnam	Hospital for Tropical Diseases in Ho Chi Minh City, Vietnam	https://doi.org/10.1093/infdis/jit224
Chemicals, peptides, and recombinant proteins		
2x KOD Hot Start Master Mix	Sigma-Aldrich	71842
NotI-HF	NEB	R3189S
BsmBI	NEB	R0739S
XbaI	NEB	R0145S
Receptor-destroying enzyme (RDE) II	VWR	370013
DMEM	Fisher	MT10013CV
Heat-inactivated fetal bovine serum	Gemini	100-106
L-glutamine	Fisher	MT25005CI
Penicillin-streptomycin solution	Fisher	MT30002CI
BioT transfection reagent	Bioland Scientific	B01-02
OptiMEM-I reduced serum media	Life Technologies	31985088
BSA Fraction V Solution	Life Technologies	15260-037
Calcium chloride	Fisher	AC349610250
Medium 199, no phenol red	Fisher	11043-023
HEPES 1M solution	ThermoFisher	15630080
Ampicillin sodium salt	Sigma	A9518-5G
Critical commercial assays		
QIAprep Spin Miniprep Kit	Qiagen	27106X4
AMPure XP Reagent	Beckman Coulter	A63880
Qubit™ 1X dsDNA High Sensitivity (HS) and Broad Range (BR) Assay Kits	Fisher	Q33231
NEBuilder HiFi DNA Assembly kit	NEB	E2621
RNEasy Plus Mini Kit	Qiagen	74134
SuperScript™ III First-Strand Synthesis SuperMix	Invitrogen	18080400
NucleoSpin Gel and PCR Clean-up kit	Takara	740609.5
T7 RiboMAX Express Large Scale RNA Production System	Promega	P1320
Monarch RNA Cleanup Kit	NEB	T2040
Deposited data		
Data files generated by dms-vep-pipeline for A/HongKong/45/2019 H3 HA data	This paper	See https://github.com/dms-vep/flu_h3_hk19_dms

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Data files generated for A/Perth/16/2009 H3 HA data	This paper	See https://github.com/jbloomlab/map_flu_serum_Vietnam_H3_Perth2009
BioProject number for raw sequencing data from all A/HongKong/45/2019 and A/Perth/16/2009 H3 HA experiments	This paper	PRJNA1124381
Experimental models: Cell lines		
Human Embryonic Kidney cells (HEK-293T)	ATCC	CRL-3216
MDCK-SIAT1-TMPRSS2	Lee et al. ⁴²	https://doi.org/10.1073/pnas.1806133115
MDCK-SIAT1	HPA Cultures	05071502
293T-CMV-PB1	Bloom et al. ⁶⁹	https://doi.org/10.1126/science.1187816
MDCK-SIAT1-CMV-PB1	Bloom et al. ⁶⁹	https://doi.org/10.1126/science.1187816
MDCK-SIAT1-CMV-PB1-TMPRSS2	Lee et al. ⁴²	https://doi.org/10.1073/pnas.1806133115
Oligonucleotides		
Primers for A/Hong Kong/45/2019 H3 HA mutagenesis	oPools from Integrated DNA Technologies	See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/targeted_library_primers/results/hk19_primers.csv for general primers, https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/major_epitope_primers/results/hk19_single_epitope_primers.csv for primers targeting major epitopes, and https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/major_epitope_primers/results/hk19_paired_epitope_primers.csv for primers targeting proximal pairs of major epitopes
Additional primers	Integrated DNA Technologies	Detailed in Table S3
Recombinant DNA		
Barcoded WSN-flanked H3 HA gene under control of the pHH21 promoter	This paper	See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/plasmid_maps/3023_pHH_WSNHAflank_HK19_H3_barcode.gb
Recipient plasmid, with GFP in place of the H3 HA ectodomain	This paper	See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/plasmid_maps/2851_pHH_WSNHAflank_GFP_H3-recipient_duppac-stop.gb
pHW18 series of plasmids for influenza reverse genetics	Hoffman et al. ⁷⁰	N/A
pHH-PB1flank-eGFP, which encodes GFP on the PB1 segment	Bloom et al. ⁶⁹	N/A
pHAGE2-EF1aInt-TMPRSS2-IRES-mCherry	Lee et al. ⁴²	N/A
A/Turkey/Mass/1975 H6 HA ectodomain	Sandbulte et al. ⁷¹	GenBank Accession AB296072.1
H6 HA neutralization standard	This paper	See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/plasmid_maps/3238_pHH_WSNHAflank_H6_H3term_mut-R235M_bc1.gb
RNA spike-in neutralization standard	This paper	See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/plasmid_maps/rna-spikein-std_template_DNA.gb
Software and algorithms		
dms-vep pipeline	This paper	See https://github.com/dms-vep/dms-vep-pipeline/tree/28ec231e5ef57d0446f9223374a2b841ba51f8cc
TargetedTilingPrimers	Radford et al. ²²	See https://github.com/jbloomlab/TargetedTilingPrimers/

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
neutcurve	Bloom lab	See https://github.com/jbloomlab/neutcurve
alignparse	Crawford and Bloom ⁷²	See https://github.com/jbloomlab/alignparse
dms_variants	Bloom lab	See https://github.com/jbloomlab/dms_variants
polyclonal	Yu et al. ²³	See https://github.com/jbloomlab/polyclonal
<i>seasonal-flu</i> Nextstrain pipeline, modified to compute escape scores	Hadfield et al. ⁴³	See https://github.com/matsengrp/seasonal-flu-dmsa
<hr/>		
Other		
5-layer tissue culture flask	Corning	353144

RESOURCE AVAILABILITY

Lead contact

Further information and requests for reagents and resources should be directed to and will be fulfilled by the lead contact, Jesse Bloom (jbloom@fredhutch.org).

Materials availability

Influenza H3 HA mutant libraries generated in this study are available from the [lead contact](#) upon request with a completed Materials Transfer Agreement.

Data and code availability

- All data are available via links in the [method details](#) section and [key resources table](#), and are collected at https://github.com/dms-vep/flu_h3_hk19_dms/ along with complete documentation of data processing. The raw sequencing data for this study can be found in the NCBI Sequence Read Archive under BioProject number PRJNA1124381.
- All original code is available at https://github.com/dms-vep/flu_h3_hk19_dms/ or via links in the associated [method details](#) sections.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Biosafety

All mutant viruses in the deep mutational scanning studies derived their HAs from one of two recent human H3N2 vaccine strains (A/Perth/16/2009 or A/Hong Kong/45/2019) and the remaining seven genes from the lab-adapted A/WSN/1933 (H1N1) strain. These viruses are therefore classified as biosafety-level 2, and the work was approved by the Fred Hutchinson Cancer Center biosafety committee.

Human sera

The human sera used for mapping escape from A/Hong Kong/45/2019 were collected in Seattle, Washington in 2020. The sera from the 2-month-old infant, the 2–5-year age cohort, and the 15–20-year age cohort are residual sera collected at Seattle Children's Hospital with approval from the Human Subjects Institutional Review Board. Sera from the 40–45-year age cohort were collected as part of the prospective longitudinal Hospitalized or Ambulatory Adults with Respiratory Viral Infections (HAARVI) cohort study of individuals with SARS-CoV-2 infection. Written informed consent was obtained for each participant. Serum from the 68-year-old individual is deidentified residual serum from the University of Washington Virology Lab that was collected for testing for HBsAb (for which they tested negative), under a study approved by the University of Washington IRB with a consent waiver.

The human sera used for mapping escape from A/Perth/16/2009 were obtained from a serum bank established in 2010 in central and southern Vietnam for the purposes of calculating disease burden, pathogen exposure, and vaccine coverage for a range of pathogens.^{41,73–78} Anonymized samples were collected, as residual serum, from the biochemistry and haematology labs of the Hospital for Tropical Diseases in Ho Chi Minh City, Vietnam, during 2010 and 2011. The study was approved by the Scientific and Ethical Committee of the Hospital for Tropical Diseases and the Oxford Tropical Research Ethics Committee at the University of Oxford.

METHOD DETAILS

Data availability statement

The file https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/results/full_hk19_escape_scores.csv contains the fully processed serum escape data for sera from the 2020 cohort measured against the A/Hong Kong/45/2019 library. This file reports the

mean escape score for each mutation, averaged between the two replicate libraries. Entries are filtered to mutations that are represented in at least three independent variants, are not stop codons, and do not have highly deleterious functional effects. Functional effects represent the effect of a mutation on HA-mediated viral infection in cell culture, calculated based on the ratio of mutation counts in the plasmid library compared to the passaged virus library. This file only includes mutations with functional effects greater than -1.38. Sites are numbered using standard H3 numbering, where the first codon of the ectodomain is site 1, through the HA1 subdomain. Rather than switching to (HA2)1 at the beginning of the HA2 subdomain, as is standard, we then continue with sequential numbering (i.e. site (HA2)1 in standard H3 numbering corresponds to site 330 in this data). See ‘Interactive plots of mutation antibody escape’ at https://dms-vep.org/flu_h3_hk19_dms/ for full heatmaps showing mutation effects on escape for all sera. See https://dms-vep.org/flu_h3_hk19_dms/muteffects_observed_heatmap.html for full heatmap showing mutation functional effects.

The file https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/results/perth2009/merged_escape.csv contains the fully processed serum escape data for sera from the 2010–2011 cohort measured against the A/Perth/16/2009 library. This file reports the median escape score for each mutation from three replicate library measurements. Sites are numbered using standard H3 numbering.

See https://github.com/dms-vep/flu_h3_hk19_dms for the full analysis pipeline for deep mutational scanning of A/Hong Kong/45/2019, and https://dms-vep.org/flu_h3_hk19_dms/ for interactive documentation of the results. See https://github.com/dms-vep/flu_h3_hk19_dms/tree/main/figures for code used to generate all main figures in this paper, with the exception of Figures 6B and S12, which can be found at <https://github.com/matsengrp/seasonal-flu-dmsa> (described in the Computational Methods section ‘Estimating escape scores for circulating H3N2 strains’). See https://github.com/matsengrp/seasonal-flu-dmsa/blob/master/profiles/dmsa-phenotype/notebooks/analysis_code.ipynb for a notebook summarizing the strain escape score analysis shown in Figures 6B and S12.

See https://github.com/jbloomlab/map_flu_serum_Vietnam_H3_Perth2009 for the full analysis of deep mutational scanning of A/Perth/16/2009.

A/Perth/16/2009 and A/Hong Kong/45/2019 library overview

The analysis in this paper uses two distinct libraries: a non-barcoded library in the background of A/Perth/16/2009, and a barcoded library in the background of A/Hong Kong/45/2019. Because the A/Perth/16/2009 library is non-barcoded, these selection experiments use a barcoded-subamplicon sequencing approach, where the HA gene is amplified as a collection of short, tiled fragments for Illumina sequencing (see Doud and Bloom²¹ and https://jbloomlab.github.io/dms_tools2/bcsubamp.html for more detailed descriptions of this approach). There is no linkage information between these fragments, so escape scores are calculated as the enrichment of a given mutation in the serum selection relative to the mock selection. These libraries are therefore designed to be mostly single-mutant variants.

In contrast, the incorporation of barcodes in the A/Hong Kong/45/2019 libraries allows for more straightforward sequencing, and the ability to resolve multiple mutations in a single variant. These libraries were therefore designed to have an average of 2–3 mutations per variant. Escape scores are calculated as the enrichment of a given *variant* in a selection experiment (as measured by sequencing counts for each unique barcode), rather than enrichment of each single mutation.

The A/Perth/16/2009 experiments use the exact library stocks generated in Lee et al.²⁰ Serum selection experiments and data analysis using the A/Perth/16/2009 library were carried out as described in the methods of Lee et al.²⁰ The escape scores from singly-infected ferrets are taken directly from data published in Figure 6 of Lee et al.: see samples ‘ferret Pitt 1 postinf’, ‘ferret Pitt 2 postinf’, ‘ferrett Pitt 3 postinf’, and ‘ferret WHO Perth2009’ (all sera collected 14–23 days post-infection with A/Perth/16/2009).²⁰ The following methods sections all pertain to analyses using the A/Hong Kong/45/2019 barcoded libraries.

Design of barcoded H3 plasmid with duplicated WSN-HA packaging signals

Influenza is multi-segmented, and proper packaging of vRNA into the virion relies on segment-specific RNA sequences called “packaging signals.” These span both coding and noncoding regions at the 3' and 5' ends of the vRNA segment.⁷⁹ To insert a barcode without disrupting vRNA packaging, we designed a chimeric construct with a duplicated 5' packaging signal (negative sense), based on chimeric influenza genes designed by Gao and Palese.⁸⁰ We place the full 150-nucleotide 5' packaging signal (negative sense) from the lab-adapted strain A/WSN/1933 HA after the stop codon of the HA gene. The 90-nucleotide native packaging sequence in the 5' coding region is synonymously recoded to avoid interference between duplicated sequences. We found that maintaining the native H3 HA 5' coding region was essential for virus growth, consistent with prior work.⁸¹ The native 3' packaging signal (defined as the 32-nucleotide noncoding region and first 16 amino acids of HA) is also replaced by the corresponding sequence from A/WSN/1933 HA (32-nucleotide noncoding region and first 19 amino acids) (Figure S1).⁸²

An extra stop codon was added to the end of the coding sequence to minimize translational readthrough. This is followed by a random 16-nucleotide barcode, then the Illumina read1 priming sequence, then the full-length WSN-HA packaging signal. Homology between the native and duplicated packaging signal could lead to barcode loss; we therefore included a stop codon in the duplicated packaging signal that will be in-frame if it replaces the native packaging signal. This ensures that loss of the barcode will generate a truncated, non-functional HA protein (Figure S1).

The barcoded WSN-flanked H3 HA gene is inserted into the pHH21 plasmid, flanked by the human polymerase I promoter and the mouse polymerase I terminator.⁸³ See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/plasmid_maps/3023_pHH_WSNHAflank_HK19_H3_barcoded.gb for an annotated map of the full plasmid.

Design of barcoded H3 deep mutational scanning library

In a library of HA variants with combinations of mutations, the rate of non-functional variants will be much higher than a single-mutant library, as a variant carrying *any* deleterious mutation will likely be non-functional. We therefore maximized the number of functional variants in the library by excluding mutations with highly deleterious functional effects. Using H3 amino acid preferences calculated from previous deep mutational scanning of A/Perth/16/2009 H3 HA,⁴² we retained mutations with functional effects more positive than the 0.75 quantile of stop codon effects (Figure S3A).

We then identified excluded mutations that were tolerated in other sequence contexts by analyzing natural H3 sequences circulating from 1968 to 2021. After basic quality control and exclusion of egg-passaged strains, we aligned 7,597 H3 variants, and identified 44 excluded mutations that were present at a frequency of at least 0.5% in natural sequences. These were added back into the library, for a total of 2,505 targeted mutations (Figure S3C).

To ensure that antigenically significant mutations were highly represented in multiple backgrounds, we biased mutagenesis towards 22 selected epitope sites (Figure S3B). These were defined as sites that had been mutated in at least one major clade from 2015–2021. These sites would be randomly mutated, regardless of functional effects, at a higher rate than the general targeted mutations.

Sites within a primer length of one another (approximately 30bp) will not be mutated in the same variant, because a primer introducing a mutation at position 1 carries the wildtype sequence at position 2. We therefore generated paired epitope primers for epitope sites within a primer length of one another, to ensure that major epitope sites are represented in combination.

Primers were designed to introduce these targeted mutations into the HA of A/Hong Kong/45/2019, and generated as both a forward and reverse primer pool. We generated a separate set of NNK primers to introduce random mutations at the key epitope positions, and mixed them with the corresponding forward or reverse primer pool at a 10:1:1 molar ratio of epitope: paired epitope: general primers.

See https://github.com/dms-vep/flu_h3_hk19_dms/tree/main/library_design for the full analysis described here, including identification of deleterious mutations, analysis of natural H3 sequences, identification of key epitope positions, and generation of the complete set of forward and reverse mutagenic primers.

Generation of plasmid library of barcoded H3 HA variants

See <https://github.com/jbloomlab/CodonTilingPrimers> for a general description of the PCR mutagenesis strategy we use here. Note that we are using primers that generate targeted amino-acid mutations, as described above.

To mutagenize the A/Hong Kong/45/2019 HA sequence, we first amplified the HA coding sequence from the plasmid containing the barcoded WSN-flanked H3 HA gene under control of the pH21 promoter. See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/plasmid_maps/3023_pHH_WSNHAflank_HK19_H3_barcode.gb for the sequence of this plasmid. The PCR was performed with the following conditions: PCR mix: 20 μ L H2O, 1.5 μ L 10 μ M forward linearizing primer (primer_088), 1.5 μ L 10 μ M reverse linearizing primer (primer_089), 2 μ L 5ng/ μ L template plasmid, and 25 μ L 2x KOD Hot Start Master Mix (Sigma-Aldrich, Cat. No. 71842). Cycling conditions: (1) 95C/2min (2) 95C/20sec (3) 70C/1sec (4) 54C/30sec, cooling at 0.5C/sec (5) 70C/40sec (6) Return to Step 2 x19.

The amplified, linearized chimeric H3 HA sequence was gel purified using the NucleoSpin Gel and PCR Clean-up kit (Takara, Cat. No. 740609.5) and then purified using Ampure XP beads (Beckman Coulter, Cat. No. A63881) at 1:1 sample to bead ratio.

The linearized chimeric H3 HA sequence was then mutagenized using a modification of a previously described PCR mutagenesis technique.⁸⁴ Forward and reverse primers for mutagenic PCR were pooled separately at a molar ratio of 10:1:1 between epitope: paired epitope: general mutagenic primers. Mutated fragments of the chimeric H3 HA sequence were then PCR-amplified in two separate reactions, using either the forward primer pool and reverse linearizing primer, or the reverse primer pool and forward linearizing primer. We performed two rounds of mutagenesis, with 7 mutagenic PCR cycles the first round and 11 the second round. The mutagenesis PCRs were performed with the following conditions: PCR mix: 20 μ L H2O, 1.5 μ L 4.5 μ M forward or reverse primer pool, 1.5 μ L 4.5 μ M reverse (primer_089) or forward (primer_088) linearizing primer, 4 μ L 3ng/ μ L linearized template, and 15 μ L KOD. Cycling conditions: (1) 95C/2min (2) 95C/20sec (3) 70C/1sec (4) 50C/30sec, cooling at 0.5C/sec (5) 70C/55sec (6) Return to Step 2 for the number of cycles described above (7) 70C/2min. Note that mutagenesis was performed in duplicate, to generate two independently mutagenized replicate libraries.

Between each round of mutagenic PCR, the mutagenized partial H3 HA fragments from the forward and reverse reactions were joined in a separate PCR reaction to generate full-length H3 HA sequences. The joining PCR was performed with the following conditions: PCR mix: 4 μ L each of 1:4 diluted forward and reverse mutagenesis reactions, 1.5 μ L 4.5 μ M forward primer (primer_088), 1.5 μ L 4.5 μ M reverse primer (primer_089), 3 μ L H2O, and 15 μ L KOD. Cycling conditions: (1) 95C/2min (2) 95C/20sec (3) 70C/1sec (4) 50C/30sec, cooling at 0.5C/sec (5) 70C/40sec (6) Return to Step 2 x19 (7) 70C/2min. Joined PCR mutagenesis products were gel and Ampure bead purified after each joining reaction, as described previously.

We then appended random 16-nucleotide barcodes to the mutagenized H3 HA sequences, followed by a constant overlap sequence for cloning into the backbone plasmid, using a PCR with the following conditions: PCR mix: 11.2 μ L H2O, 0.9 μ L 10 μ M forward linearizing primer (primer_088), 0.9 μ L 10 μ M reverse barcoding primer (primer_090), 2 μ L 15ng/ μ L mutagenized template, and 15 μ L KOD. Cycling conditions: (1) 95C/2min (2) 95C/20sec (3) 70C/1sec (4) 50C/30sec, cooling at 0.5C/sec (5) 70C/40sec (6) Return to Step 2 x9 (7) 70C/2min.

The barcoded mutagenized H3 HA sequences were gel and Ampure bead purified as described previously, using a 1:1 sample to bead ratio, then cloned into a plasmid containing the flanking WSN HA packaging signals, 5' and 3' NCRs, and pH21 promoter and terminator. The plasmid contains GFP in place of the H3 HA ectodomain to ensure that virions cannot be generated from untransformed plasmid. See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/plasmid_maps/2851_pHH_WSNHAflank_GFP_H3-recipient_duppac-stop.gb for detailed map of the GFP recipient plasmid. The plasmid was digested with BsmBI and NotI, then gel and Ampure bead purified. We then used a 2:1 insert to vector ratio in a 1 hour HiFi assembly reaction using NEBuilder HiFi DNA Assembly kit (NEB, Cat. No. E2621). HiFi assembly products were Ampure bead purified and eluted into 20 μ L of H2O for higher electroporation efficiency.

We electroporated 20 μ L of 10-beta electrocompetent *E. coli* cells (NEB, C3020K) with 1 μ L of the purified HiFi product. Electroporated cells were plated on 15cm LB+ampicillin plates at an estimated bottleneck of 80,000 to 100,000 colony forming units to limit library size. After approximately 18 hours of outgrowth, colonies were scraped into liquid LB+ampicillin, and grown at 37°C for 2.5 hours in liquid culture prior to plasmid purification using QIAprep Spin Miniprep Kit (Cat. No. 27104). We generated two independently mutagenized plasmid libraries, referred to as libA and libB. The final structure of the mutagenized H3 HA library construct is illustrated in Figure S1B.

PacBio sequencing of H3 HA plasmid library

We used long-read PacBio sequencing to link variants to their random nucleotide barcodes, and determine the composition of the plasmid library, similar to the approach used in Starr et al.⁸⁵ PacBio sequencing inserts were prepared from the purified plasmid pools via NotI-HF restriction digest (NEB, Cat. No. R3189S), rather than PCR amplification, which eliminates the possibility of PCR strand exchange scrambling barcodes. Digested inserts were then gel and Ampure bead purified, and eluted into 30 μ L of EB. Each library was then barcoded for PacBio sequencing using SMRTbell prep kit 3.0, bound to polymerase using Sequel II Binding Kit 3.2, and then sequenced using a PacBio Sequel IIe sequencer with a 20-hour movie collection time. Analysis is described in the section ‘PacBio sequencing data analysis’.

Generation of H3 HA variant library in the context of lab-adapted WSN virus

Mutant virus libraries were generated by reverse genetics.⁷⁰ We used 6-well plates rather than 15cm plates for transfection, to limit the extent to which jackpotting of certain variants in each well could bottleneck library diversity. We transfected 40 wells of 6-well plates for each library. Each well was plated with 4x10⁵ HEK-293T cells in D10 media (DMEM, supplemented with 10% heat-inactivated FBS, 2 mM L-glutamine, 100 U/mL penicillin, and 100 μ g/mL streptomycin). 21 hours after plating, cells were transfected with 220ng each of H3 HA plasmid library (or wildtype control), the pHW18* series of plasmids⁷⁰ for all non-HA viral genes, and pHAGE2-EF1alnt-TMPRSS2-IRES-mCherry.⁴² The pHW18* plasmids encode genes from the lab-adapted A/WSN/1933 strain, and so all genes in our viruses derive from the lab-adapted WSN strain except for the HAs. We used Bioland BioT (Cat. No. B01-01) for the transfections, and followed the manufacturer’s recommendations for the protocol and DNA / transfection reagent ratios.

At 20 hours post-transfection, we removed the D10 media and added 3x10⁵ MDCK-SIAT1-TMPRSS2 cells per well in low-serum IGM (Influenza Growth Media, consisting of Opti-MEM supplemented with 0.01% heat-inactivated FBS, 0.3% BSA, 100 U/mL penicillin, 100 μ g/mL streptomycin, and 100 μ g/mL calcium chloride). Note that overlay of MDCK-SIAT1-TMPRSS2 cells post-transfection yielded higher virus titers than transfection of co-cultured cells in pilot rescue experiments. At 49 hours post-transfection, the transfection supernatants were harvested, clarified by centrifugation at 2,000xg for 5min, aliquoted, frozen at -80°C, and titered in MDCK-SIAT1-TMPRSS2 cells. The titers were 3,162 TCID50/ μ L for both libA and libB.

To generate a genotype-phenotype link, where the barcoded HA gene carried by each virion matches the HA variant expressed on its surface, we passaged the libraries at a low MOI such that each cell should be infected by one virion at most. For each library, we infected cells in two 5-layer 875 cm² flasks (Corning, Cat. No. 353144) at an MOI of 0.01. For each flask, we gently mixed 7.875x10⁷ MDCK-SIAT1-TMPRSS2 cells in 150mL of IGM (targeting a density of 9x10⁴ cells/cm² when plated) with 7.875x10⁵ TCID50 of library transfection supernatant in a sterile 250mL bottle. The cell-virus mix was then added directly to the 5-layer flask. Virus supernatant was collected at 45 hours post-infection, clarified by centrifugation at 2,000xg for 5min, split into 1mL aliquots, frozen at -80°C, and titered in MDCK-SIAT1-TMPRSS2 cells. The titers were 3.09x10⁴ and 1.43x10⁴ TCID50/ μ L for libA and libB, respectively. Assuming a maximum library size of 100,000 variants, libraries were passaged at over 10x coverage of variant diversity.

For future deep mutational scanning studies, we would recommend rescuing the library at 40 hours post-transfection, and reducing passaging to approximately 36 hours. These conditions yielded similar titers, and minimizing library growth times will reduce bottlenecking from variants carrying tissue-culture-adaptive mutations.

Design of external neutralization standards

In order to generate quantitative escape measurements, we designed barcoded standards that would not be neutralized by human sera (Figure S2A). These would be spiked into the library at a low level (approximately 1%), and used to normalize the barcode frequency between mock selections and serum selections (Figure 1A). This normalization is expected to correct for stochastic variation in virus replication across different conditions and experiments. As more of the library is neutralized at increasingly potent serum concentrations, the proportion of library barcodes to standard barcodes will decrease in a consistent manner (Figure S2B).

We included standards at two stages: infection with the virus library, and harvest of RNA from infected cells. The virus standard is identical to the WSN-flanked H3 HA library construct, but uses the ectodomain and 5' coding region from A/Turkey/Mass/1975 H6

HA, a low-pathogenicity avian HA strain (GenBank Accession AB296072.1) (Figure S2A).⁷¹ Note that although the HA of this viral standard is from a low-pathogenicity avian strain, the virus itself is still only biosafety-level 2, because the seven non-HA gene segments and the H6 HA packaging signals are derived from the lab-adapted WSN strain. This standard will have similar infection dynamics as the library H3 HA variants. We tested neutralization activity of several sera against the WSN-H6 HA standard, and found that any serum used at a dilution of approximately 1:100 or higher begins to neutralize WSN-H6 HA.

Therefore, to allow for analysis of low-potency sera that must be at high concentration for selections, we also included an RNA spike-in standard. This is an RNA fragment that is identical to the WSN-flanked H3 HA sequence from the 5' coding region to the 5' UTR, but replaces the H3 ectodomain with GFP (Figure S2A). This ensures that the RNA spike-in standard is a similar length to the HA RNA. The fragment is generated with a pool of approximately 100 random barcodes. We add a standardized amount of RNA spike-in to the lysis buffer used for harvest of cellular RNA, so that it's equally distributed between wells in each experiment. We tested a range of concentrations, and found that the RNA spike-in behaves similarly to the WSN-H6 HA standard when 0.05ng of RNA spike-in is added per well (Figure S2C).

See https://github.com/dms-vep/flu_h3_hk19_dms/tree/main/library_design/plasmid_maps for annotated plasmid maps of both standards.

Generation of WSN-H6 HA neutralization standard

The WSN-H6 HA standard was generated from the same GFP recipient vector used to clone the WSN-flanked H3 HA library construct (see https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/plasmid_maps/2851_pHH_WSNHAflank_GFP_H3-recipient_duppac-stop.gb for plasmid map). The plasmid was digested with BsmBI (NEB, Cat. No. R0739S) and XbaI (NEB, Cat. No. R0145S), then gel and Ampure bead purified. To avoid sequencing errors, we designed four different barcodes for the WSN-H6 HA standard, such that the nucleotide at a given site was different in all four barcodes. The four barcoded H6 HA ectodomain sequences were synthesized from Twist with a 20bp overlap with the digested recipient vector, and cloned into the recipient vector using a 2:1 insert to vector ratio in a 1 hour HiFi assembly reaction. NEB 5 α High Efficiency Cells (Cat. No. C2987) were transformed with 2 μ L of the HiFi assembly, and plated on LB+ampicillin for 18 hours of outgrowth. Single colonies were grown in liquid LB+ampicillin at 37°C for 18 hours, plasmid was purified using QIAprep Spin Miniprep Kit, and sequence was confirmed by Sanger sequencing. The four barcoded plasmids were then combined at equal concentrations. See plasmids 3238-3241 in https://github.com/dms-vep/flu_h3_hk19_dms/tree/main/library_design/plasmid_maps for detailed map of the WSN-H6 HA plasmid with four unique barcodes.

To generate the virus, we transfected a coculture of 4x10⁵ HEK-293T and 5x10⁴ MDCK-SIAT1 cells with 220ng each of the pooled WSN-H6 HA plasmid, the pHW18* series of plasmids⁷⁰ for all non-HA viral genes, and pHAGE2-EF1alnt-TMPRSS2-IRES-mCherry. Transfections were performed in D10 media using BioT transfection reagent. Media was changed to IGM at 20 hours post-transfection. At 68 hours post-transfection, transfection supernatants were harvested, clarified by centrifugation at 2000xg for 5 min, aliquoted, frozen at -80°C, and titered in MDCK-SIAT1-TMPRSS2 cells. Titers were 3.16x10⁵ TCID50/ μ L.

Generation of RNA spike-in neutralization standard

The RNA spike-in standard was generated from the GFP recipient vector, with the final 46 amino acids from the library strain appended at the 5' end of the GFP sequence (in negative-sense viral orientation), such that the 5' coding, random barcode region, and priming sequence were identical to the library strain. A T7 promoter sequence was added after the 5' NCR for in vitro transcription of the RNA sequence. See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/library_design/plasmid_maps/rna-spikein-std_template_DNA.gb for full sequence of the linear DNA template. The GFP recipient vector was digested with XbaI alone, then gel and Ampure bead purified. The C-terminal insert sequence was PCR-amplified off of the chimeric library plasmid, using a forward primer that appended a 20-bp overlap with the GFP sequence. PCR conditions were as follows: PCR mix: 21 μ L H2O, 1.5 μ L 10 μ M forward GFP overlap primer (primer_115), 1.5 μ L 10 μ M reverse linearizing primer (primer_089), 1 μ L 10ng/ μ L template plasmid, and 25 μ L 2x KOD Hot Start Master Mix. Cycling conditions: (1) 95C/2min (2) 95C/20sec (3) 63C/10sec (4) 70C/10sec (5) Return to Step 2 x19 (6) 70C/2min. Amplified insert was gel purified.

We then appended random 16-nucleotide barcodes to the amplified insert, followed by a constant overlap sequence for cloning into the GFP recipient vector, using a PCR with the following conditions: PCR mix: 10.66 μ L H2O, 0.9 μ L 10 μ M forward linearizing primer (primer_088), 0.9 μ L 10 μ M reverse barcoding primer (primer_090), 2.54 μ L 11.8ng/ μ L insert template, and 15 μ L 2x KOD Hot Start Master Mix. Cycling conditions: (1) 95C/2min (2) 95C/20sec (3) 70C/1sec (4) 50C/30sec, cooling at 0.5C/sec (5) 70C/20sec (6) Return to Step 2 x9 (7) 70C/2min.

The barcoded insert was gel purified, then cloned into the XbaI-digested GFP recipient vector, using a 3:1 insert to vector ratio in a 1 hour HiFi assembly reaction. NEB 5 α High Efficiency Cells were transformed with 2 μ L of the HiFi assembly, and plated on LB+ampicillin for 18 hours of outgrowth. To generate a diverse pool of barcodes, 20 individual colonies were grown in liquid LB+ampicillin at 37°C for 18 hours, and plasmids were purified using QIAprep Spin Miniprep Kit. Sequence was confirmed by Sanger sequencing. The 16 plasmids with no SNPs were pooled at equal concentrations.

From this randomly-barcode pool, we amplified the linear DNA template from the U12 to U13 region, and added a T7 promoter sequence in the reverse orientation after U13 (5' terminus of the linear template, negative-sense viral orientation). PCR conditions were as follows: PCR mix: 11.2 μ L H2O, 0.9 μ L 10 μ M forward U12-annealing primer (primer_113), 0.9 μ L 10 μ M reverse T7-appending primer (primer_116), 2 μ L 10ng/ μ L template plasmid, and 15 μ L 2x KOD Hot Start Master Mix. Cycling conditions: (1) 95C/2min (2)

95C/20sec (3) 62C/10sec (4) 70C/40sec (5) Return to Step 2 x19 (6) 70C/2min. We ran two replicate 30uL PCR reactions, and gel purified both reactions using the same column to maximize concentration of the linear DNA template.

The RNA spike-in standard was then generated via *in vitro* transcription of the linear DNA template with T7 promoter, using the T7 RiboMAX Express Large Scale RNA Production System (Promega, Cat. No. P1320) with 1 μ g DNA template input. Reaction was incubated at 37°C for 45min. To remove DNA template after transcription, we added 1 μ L of RQ1 RNase-Free DNase and incubated for an additional 15 min at 37°C. RNA was then purified using the Monarch RNA Cleanup Kit (NEB, Cat. No. T2040).

Final RNA yield was at a concentration of 2.24 μ g/ μ L. This was diluted to 0.025ng/ μ L, aliquoted, and stored at -80°C. Barcodes were amplified and sequenced on a NextSeq 2000 alongside library selection samples, as described in the section ‘Barcode amplification from infected cells for Illumina sequencing’. The final pool of RNA spike-in neutralization standard contained 125 unique barcodes.

Serum preparation

All sera were heat-inactivated and treated with receptor-destroying enzyme (RDE) before use. RDE-treatment removes residual sialic acids in human serum samples that can bind to the virus libraries. We mixed serum with resuspended, lyophilized RDE II (VWR, Cat. No. 370013) at a ratio of 1:3, and incubated in heatblocks at 37°C for 2.5hr. We then heat-inactivated the RDE and serum by incubating at 55°C for 30min, aliquoted, and stored at -80°C. Note that all sera used in this study are at a starting dilution of 1:4 due to RDE treatment, which is accounted for in reported concentrations.

GFP neutralization assays for serum screening and validation

Neutralization assays were performed using influenza viruses carrying GFP in the PB1 segment. These GFP viruses were generated by reverse genetics using the following plasmids: wildtype WSN-flanked A/Hong Kong/45/2019 HA (the library background strain), pHII-PB1flank-eGFP (which encodes GFP on the PB1 segment),⁶⁹ and the pHW18* series of plasmids⁷⁰ for all viral genes except HA and PB1. The same approach was used to generate GFP virus for validation neutralization assays. For mutation validations, the mutation of interest was introduced into the WSN-flanked A/Hong Kong/45/2019 HA plasmid. For strain validations, the ectodomain of the strain of interest (amino acid sites 1 to (HA2)175 in standard H3 HA numbering) replaces the A/Hong Kong/45/2019 ectodomain in the WSN-flanked H3 HA library construct. See https://github.com/dms-vep/flu_h3_hk19_dms/tree/main/library_design/plasmid_maps for detailed maps of the five WSN-flanked H3 HA strains used for validations in Figure 6B.

We added 4x10⁵ 293T-CMV-PB1⁶⁹ and 5x10⁴ MDCK-SIAT1-CMV-PB1-TMPRSS2⁴² cells in D10 media to each well of a 6-well plate. Cells that constitutively express PB1 are necessary because the PB1 segment is replaced by GFP in these viruses. 18 hours after plating, wells were transfected with 220ng of each reverse genetics plasmid and the pHAGE2-EF1aInt-TMPRSS2-IRES-mCherry plasmid, using BioT transfection reagent. At 20 hours post-transfection, we changed the media in each well to 2mL NAM (Neutralization Assay Media, consisting of Medium-199 supplemented with 0.01% heat-inactivated FBS, 0.3% BSA, 100 U/mL penicillin, 100 μ g/mL streptomycin, 100 μ g/mL calcium chloride, and 25mM HEPES). This media is used because it has low background fluorescence at 488 nm. At approximately 40 hours post-transfection, the transfection supernatants were harvested, clarified by centrifugation at 2,000xg for 5min, aliquoted, and frozen at -80°C.

To titer these viruses, we infected 1x10⁵ MDCK-SIAT1-CMV-PB1-TMPRSS2 in a 12-well plate with different dilutions of virus, chose wells that showed 1-10% GFP positivity at 18 hours post-infection, measured the fraction of GFP-positive cells by flow cytometry, and used this fraction to calculate the titer of infectious particles per μ L.

For accurate measurement of serum neutralization curves, the correlation between number of infectious particles and change in GFP fluorescence should be linear. Adding too much or too little virus per well will result in inaccurate readings. It's therefore important to generate an MOI curve for each GFP virus, where MDCK-SIAT1-CMV-PB1-TMPRSS2 cells in a 96-well plate are infected with 2-fold dilutions of virus, and GFP signal is measured 18 hours post-infection. The virus concentration used for neutralization assays should correspond to the highest MOI in the linear range of MOI:signal.

Neutralization assays were performed as described previously^{86,87} (see also https://github.com/jbloomlab/flu_PB1flank-GFP_neut_assay for detailed protocol). Serum was diluted 2-fold in NAM, for final volumes of 40 μ L per well of a 96-well plate. These dilutions were 2x more concentrated than the target dilutions, to account for the addition of 40 μ L of virus. Note that calculation of neutralization curves assumes that antibody is in excess of virus, and if this is not the case, the measured IC50 will be artificially high. For some highly potent sera (IC50=1x10⁻⁴ dilution or lower), serum dilution volumes were therefore increased to 100 μ L to avoid saturation of serum antibodies. The serum-virus mix was incubated for 1.5 hours at 37°C before adding 4x10⁴ MDCK-SIAT1-CMV-PB1-TMPRSS2 cells per well. When screening for neutralization activity against the wildtype library strain, only one replicate was run per serum. Sera selected for escape mapping were then re-run with two replicates per serum to generate more accurate curves. All validation neutralization assays were run with two replicates per serum and virus, and curves represent the mean and standard error of these replicates.

GFP signal was measured 18 hours post-infection. To calculate fraction infectivity, we subtracted the average background signal from uninfected cells from each condition, then divided the signal from each serum dilution by the average signal from no-serum control infections. Fraction infectivities were then used to fit Hill-like neutralization curves using the neutcurve package (<https://jbloomlab.github.io/neutcurve/>), version 0.5.7. See https://github.com/dms-vep/flu_h3_hk19_dms/tree/main/neut_assays for detailed information on curve fitting for all neutralization assays in this paper.

Serum selection of library escape variants

Selections were performed at three concentrations for each serum, targeting a range of concentrations from serum IC90 to IC99.99. Using multiple potent serum concentrations for selections generates more accurate measurements of enriched escape variants, and improves fitting with our biophysical escape models.²³ Selections were also performed at a theoretical MOI of approximately 0.6 or lower (depending on percentage of neutralized library variants), to avoid competition effects between variants.

For each experiment, we plated MDCK-SIAT1 cells in IGM 4 hours before infection. Because large differences in MOI impact the amount of viral mRNA transcribed by each cell,³⁸ we plated 1.5×10^5 cells per well for each selection condition, and 1×10^6 cells per well for the no-antibody controls, to roughly account for the expected difference in infectious particles. Selections used 12-well plates while no-antibody controls used 6-well plates, to keep cell density relatively consistent. We made a master mix of library virus with WSN-H6 HA neutralization standard added at 0.1% (libA) or 0.3% (libB) of the total virus titer. These ratios were calibrated in pilot selections such that approximately 1% of counts in the no-antibody control conditions would come from the neutralization standard. Using a master mix of virus is important as it ensures comparable results between the no-antibody and serum selection conditions. 400 μ L of virus master mix was mixed with 400 μ L serum (diluted to 2x the targeted concentration, to account for virus volume), or 400 μ L media for the mock selection control. All dilutions were in IGM. The virus-serum or virus-media mixes were incubated at 37°C for one hour, then added dropwise to each well. RNA was harvested from infected cells at 13 hours post-infection (described below).

In initial pilot experiments, we added 6×10^5 TCID50 per infection for each library for approximately 20-fold coverage of variant diversity, and ran selections in minimal volumes to conserve serum. We found that these conditions resulted in serum antibody saturation, where there were more virions than antibodies, and results were skewed by measurement of unbound virions. 800 μ L was the minimal selection volume that avoided saturation effects for potent sera. We also found that the calculated probability of escape for wildtype variants was consistently higher for libB than libA. We therefore reduced the titer of libB added per infection to 3×10^5 TCID50, which led to consistent results between libraries for wildtype variants. This discrepancy is likely due to inconsistency between TCID50, measured as virus spread in cell culture, and selection experiment results, measured based on cell entry. We therefore recommend using the probability escape for wildtype variants from a pilot experiment, which should reflect targeted ICXX values (e.g. 10% probability escape at serum IC90), to calibrate the amount of virus added per selection.

To minimize bottlenecking of variant diversity during infections, we infected 3 wells with no-antibody controls for each library, prepared the RNA and viral barcodes separately (described below), and pooled the Illumina sequencing reads into a single no-antibody control sample.

Barcode amplification from infected cells for Illumina sequencing

RNA was harvested from infected cells 13 hours post-infection using the RNeasy Plus Mini Kit (Qiagen, Cat. No. 74134). We made a master mix of RLT lysis buffer with 1% β-mercaptoethanol to inhibit RNases, and 0.14ng/mL RNA spike-in standard (i.e. 2 μ L of 0.025ng/ μ L standard for every 350 μ L in the master mix). This correlates to 0.05ng of standard in 352 μ L lysis master mix, added to each well. Working with one row of wells at a time, we removed the media, washed with 1mL PBS, and aspirated completely. We then added 352 μ L of the lysis master mix with RNA standard. Cells were washed from each well and transferred to 1.5mL aliquots, then vortexed for 1min at full speed to homogenize. After lysing and homogenizing cells from all conditions in the experiment, we proceeded with RNA extraction, following the extraction kit protocol. RNA was eluted twice from each column in 30 μ L RNase-free water to maximize yields. Extraction was typically completed within 1.5 hours from homogenization.

Viral mRNA was transcribed as cDNA, using SuperScript III First Strand Synthesis SuperMix (Invitrogen, Cat. No. 18080400). We normalized amount of input RNA across all samples in the experiment, and input anywhere from 800ng-1 μ g RNA per sample, depending on the maximum input from the lowest-concentration sample. Reaction was carried out according to manufacturer protocol, using the mRNA-annealing RT primer (primer_110) to amplify off of viral mRNA. We chose to target mRNA as it is more abundant in infected cells than vRNA.

We then used two rounds of PCR to amplify the barcode region and add the necessary Illumina adapters and indices for each sample. The first round of PCR uses a forward primer that anneals to the C-terminus of the HA sequence, and a reverse primer that anneals to the Illumina Truseq Read 1 sequence downstream of the barcode. We use the maximum recommended input of cDNA, 10 μ L, to minimize bottlenecking of barcode diversity during amplification. Conditions were as follows: PCR mix: 14 μ L H2O, 0.5 μ L 5 μ M Illumina round 1 forward primer (primer_098), 0.5 μ L 5 μ M Illumina round 1 reverse primer (primer_099), 10 μ L cDNA template, 25 μ L KOD. Cycling conditions: (1) 95C/2min (2) 95C/20sec (3) 70C/1sec (4) 64C/10sec, cooling at 0.5C/sec (5) 70C/20sec (6) Return to Step 2 x19 (7) 70C/2min.

DNA concentration was quantified using a Qubit Fluorometer (ThermoFisher). A second round of PCR was then performed using custom dual indexing primers, which anneal to the Read 1 or Read 2 sites, and append the P5 Illumina adapter and i5 sample index at the Read 1 terminus, or the P7 Illumina adapter and i7 sample index at the Read 2 terminus. Conditions were as follows: PCR mix: 5ng of round 1 PCR product (variable volumes), H2O to 19 μ L, 3 μ L 2.5 μ M unique i5 indexing primer, 3 μ L 2.5 μ M unique i7 indexing primer, 25 μ L KOD. Cycling conditions: (1) 95C/2min (2) 95C/20sec (3) 70C/1sec (4) 55C/20sec (5) 70C/20sec (6) Return to Step 2 x9 (7) 70C/2min.

DNA concentration of each round 2 PCR product was quantified using Qubit. Serum selection samples were pooled at an even ratio, while no-antibody control samples were added at 3x the concentration of each selection sample, to account for greater barcode diversity in the no-antibody controls. Samples were gel purified and Ampure bead cleaned at a 1:2 sample to beads ratio, then

sequenced using either P2 or P3 reagent kits on a NextSeq 2000. Analysis is described in the section ‘Illumina barcode sequencing data analysis’.

QUANTIFICATION AND STATISTICAL ANALYSIS

Computational pipeline overview

For analysis of A/Hong Kong/45/2019 deep mutational scanning data, we used the modular pipeline <https://github.com/dms-vep/dms-vep-pipeline> (version 2.6.5). The repository https://github.com/dms-vep/flu_h3_hk19_dms contains all analyses performed in this paper, including the main *dms-vep-pipeline* and all scripts, notebooks, and settings necessary to recreate the analysis. Full analysis of A/Perth/16/2009 deep mutational scanning data can be found at https://github.com/jbloomlab/map_fiu_serum_Vietnam_H3_Perth2009 and does not use *dms-vep-pipeline* for analysis. However, we have imported the finalized escape scores and validation neutralization data to the main https://github.com/dms-vep/flu_h3_hk19_dms repository, and integrated these results into the pipeline. See https://dms-vep.org/flu_h3_hk19_dms/ for HTML rendering of key analyses and interactive plots. These pages are the best way to explore the analyses and interactive plots of the results described in this paper.

PacBio sequencing data analysis

PacBio CCSs were analyzed using *alignparse* (see <https://jbloomlab.github.io/alignparse/> for documentation).⁷² We computed the empirical accuracy of the CCSs as the frequency at which CCSs with the same barcode were associated with the same variant sequence (see https://jbloomlab.github.io/alignparse/alignparse.consensus.html#alignparse.consensus.empirical_accuracy for implementation). Empirical accuracies for both libraries were approximately 0.99.

We filtered out any barcodes that had less than two CCSs, to ensure confidence in the consensus sequences. We also filtered out barcodes that had minor fractions of substitutions or indels above 0.2, which would indicate that multiple variants share that barcode. See https://jbloomlab.github.io/alignparse/alignparse.consensus.html#alignparse.consensus.simple_mutconsensus for method for building consensus sequences using these filters. The resulting consensus sequences were saved as barcode / variant lookup tables. This file and the associated analysis notebook, which includes quality control plots, can be found on the main HTML page described in ‘Computational pipeline overview’.

Illumina barcode sequencing data analysis

We processed the Illumina barcode sequencing using the *dms_variants* Illumina barcode parser (see https://jbloomlab.github.io/dms_variants/index.html for documentation) to determine the counts of each variant in each selection condition. We filtered barcoded mutants that had low counts in the no-antibody control conditions, as these were likely to be randomly bottlenecked during infections and generate noisy escape scores. See https://github.com/dms-vep/flu_h3_hk19_dms/blob/main/config.yaml for the configuration file where these thresholds are specified.

Modeling the effects of mutations on H3 HA function

We modeled the functional effects of mutations on A/Hong Kong/45/2019 H3 HA as previously described.²² We computed functional scores based on the frequency of a given mutation in the plasmid library (where there is no functional pressure), compared to the passaged virus library (where non-functional HA variants have been purged). The functional score for a variant v is calculated as $\log_2([n_{post}^v / n_{post}^{wt}] / [n_{pre}^v / n_{pre}^{wt}])$, where n_{post}^v is the variant count in the no-antibody mock selection condition, n_{pre}^v is the variant count in the plasmid library, and n_{post}^{wt} and n_{pre}^{wt} are the counts of all wildtype barcodes in each condition.

Many variants in the library carry multiple mutations. We therefore use global epistasis models^{89,90} to deconvolve this data and estimate the fitness effects of each individual mutation. Under these models, each mutation has an additive effect on a “latent” (unmeasured) phenotype. These additive effects are then transformed by a nonlinear function to generate the “observed” (measured) phenotype. The nonlinear transformation accounts for epistatic interactions between mutations, and thereby generates more accurate estimates for mutation effects.

See https://dms-vep.github.io/flu_h3_hk19_dms/fit_globalepistasis.html for the notebooks used to fit global epistasis models, https://dms-vep.github.io/flu_h3_hk19_dms/muteffects_latent_heatmap.html for the latent effects of mutations averaged across selections, and https://dms-vep.github.io/flu_h3_hk19_dms/muteffects_observed_heatmap.html for the observed effects of mutations averaged across selections.

Modeling the effects of mutations on antibody and serum escape

We modeled the effects of mutations on serum escape as described previously in Dadonaitė et al.²² Briefly, we computed the fraction of non-neutralized variants at each serum selection concentration, normalized by counts from the relevant neutralization standard (either WSN-H6 HA or the RNA spike-in standard). Fraction of non-neutralized variants corresponds to probability of escape, and should fall between 0 (variant is fully neutralized) and 1 (variant fully escapes neutralization). Values greater than 1 are clipped. We then estimated the effects of each individual mutation on escape using the software package *polyclonal*, version 6.2 (see <https://jbloomlab.github.io/polyclonal/> for documentation).²³ Fits were constrained to a single epitope to simplify analysis, as most sera were highly targeted.

For each serum, the heat maps and line plots are generated using the mutation-level escape scores calculated by *polyclonal*. All sera had measurements from two replicate libraries, and the final reported values are the average (median) across those replicates. Files containing the escape scores from individual replicates, the averaged scores, analysis notebooks with relevant quality control plots, and interactive plots summarizing the final estimates are available on the main HTML page described in ‘Computational pipeline overview’.

Figures in this paper only include mutations seen in at least three distinct variants (averaged across libraries), with a functional effect of at least -1.38.

Estimating escape scores for circulating H3N2 strains

The GitHub repository at <https://github.com/matsengrp/seasonal-flu-dmsa> has the code we used to perform the analyses in this section, as well as files with key results. We analyzed a set of 1,478 circulating H3N2 strains obtained from a publicly available Nextstrain⁴³ phylogenetic tree that samples strains from the past 12 years (<https://nextstrain.org/flu/seasonal/h3n2/ha/12y>). This tree is regularly updated. For this paper, we analyzed the strains from the tree that was available on October 19th, 2023. For each strain, we obtained the corresponding HA gene nucleotide sequence and associated metadata from the GISAID database (<https://gisaid.org/>).⁹¹ The above repository provides the GISAID accession number for each strain’s HA sequence and acknowledges contributing labs.

To estimate an escape score for each strain relative to each serum sample, we ran the above sequences and metadata, including the sequence and metadata for the HA sequence of the A/Hong Kong/45/2019 strain used in the DMS experiment, through a version of the core *seasonal-flu* Nextstrain pipeline (<https://github.com/nextstrain/seasonal-flu/>) that we modified to compute escape scores. This pipeline generates a multiple-sequence alignment of translated HA protein sequences. For each strain in the alignment, it then makes a list of all amino-acid mutations that are needed to convert the A/Hong Kong/45/2019 sequence to the sequence of the given strain. Then, to compute that strain’s escape score for a given serum sample, it simply sums the serum-specific DMS-measured escape scores of all mutations in the above list, such that escape scores for the A/Hong Kong/45/2019 DMS strain always equals zero. We only summed escape scores for mutations that were seen in both libraries and in at least three distinct variants (averaged across libraries) in the DMS data for that serum sample. If a strain’s list of mutations included one or more mutations that were not measured by DMS or did not meet the above criteria, we masked that strain in the analysis, unless the mutations in question occurred at sites that we deemed unlikely to affect escape (see the repository for more detail). After the masking step, this process yielded a total of 1,233-1,238 strains with estimated escape scores for each serum sample.

Figure 6B shows the distribution of strain-specific escape scores as a function of strain collection date, with scores averaged across all sera in a given cohort. For each cohort, we fit a linear-regression model that predicts a strain’s average escape score as a function of its collection date and reported the inferred slope of the resulting best-fit line in the figure caption. Next, we used a randomization approach to determine if these slopes were significantly different between cohorts. In each of 1,000 independent trials, we randomized which serum sample was assigned to which cohort, recomputed average strain-specific escape scores using the randomized cohort assignments, and then fit a linear-regression model to the resulting data, fitting one model per cohort as above. We then compared the slopes obtained by randomization to the slopes observed in the non-randomized data. In Figure 6B, the slope associated with teenagers is 0.1 units greater than the slope associated with adults (=0.14-0.04). Out of the 1,000 randomization trials, only four trials resulted in a difference in slope (=teenagers-adults) greater than or equal to 0.1, indicating that the slope for teenagers is significantly larger than the slope for adults in the non-randomized data ($p=0.004$). The randomization test also indicated that the slope for teenagers is significantly larger than the slope for children ($p=0.022$), though the slope for children was not significantly larger than the slope for adults ($p=0.323$).

We also repeated the analysis with only the subset of H3N2 strains circulating after serum collection, from 2020 to 2023 (Figure S12). This analysis included 240 strains, after filtering for mutations not measured in at least three distinct variants (averaged across libraries) in the DMS data for each serum sample. We found that the slope associated with teenagers (0.08, $r=0.41$) is significantly higher than the slope associated with adults (-0.14, $r=-0.60$) based on 1,000 randomization trials as described above ($p=0.002$). The slope associated with children (0.0, $r=0.03$) is also significantly higher than adults ($p=0.045$). There was no significant difference between the children and teenagers ($p=0.226$).