

Deep mutational scanning of hemagglutinin helps identify evolutionarily successful human H3N2 influenza viruses

Juhye M. Lee^{a,d,e}, John Huddleston^{b,f}, Michael B. Doud^{a,d,e}, Kathryn A. Hooper^{a,f}, Nicholas C. Wu^g, Trevor Bedford^{b,c,1}, and Jesse D. Bloom^{a,c,d,1}

^aBasic Sciences Division; ^bVaccine and Infectious Disease Division; ^cand Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA; ^dDepartment of Genome Sciences; ^eMedical Scientist Training Program; ^fMolecular and Cellular Biology Program, University of Washington, Seattle, WA 98195, USA; ^gDepartment of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037

This manuscript was compiled on March 23, 2018

Human influenza virus rapidly accumulates mutations in its major surface protein hemagglutinin (HA). The evolutionary success of influenza virus lineages depends on how these mutations affect HA's functionality and antigenicity. Here we experimentally measure the effects on viral growth in cell culture of all single amino-acid mutations to the HA from a recent human H3N2 influenza virus strain. We then show that mutations in evolutionarily successful H3N2 viral lineages are measured to generally be more favorable than mutations in lineages that die out. Therefore, despite the well-known caveats about cell-culture measurements of viral fitness, such measurements can still be informative for understanding virus evolution in nature. We also compare our measurements for an H3 HA to similar data previously generated for a distantly related H1 HA, and find substantial differences in which amino acids are preferred at many sites. For instance, the H3 HA has less disparity in mutational tolerance between the head and stalk domains than the H1 HA. Overall, our work suggests that experimental measurements of mutational effects can be leveraged to help forecast the evolutionary fates of viral lineages in nature — but only when the measurements are made on a viral strain similar to the ones being studied in nature.

influenza virus | hemagglutinin | deep mutational scanning | antigenic drift | epistasis

Seasonal H3N2 influenza virus evolves rapidly, fixing 3 to 4 amino-acid mutations per year in its hemagglutinin (HA) surface protein. Many of these mutations contribute to the rapid antigenic drift that necessitates frequent updates to the annual influenza vaccine (1, 2). This evolution is further characterized by competition and turnover among groups of strains or clades. As a result, mutations that differentiate clades exhibit dynamic frequency trajectories as these clades appear, circulate, or die out. Several lines of evidence indicate that successful clades, which reach high maximum frequencies or even fix in the global population, have higher fitness than clades that remain at low frequency. A key goal in the study of H3N2 evolution is to identify the features that enable certain clades to succeed as others die out.

Two main characteristics distinguish evolutionarily successful clades from their competitors: greater antigenic change, and efficient viral growth and transmission. In principle, experiments could be informative for identifying how mutations affect these features. Most work on influenza evolution to date has utilized experimental data primarily to assess the antigenicity of circulating strains (3–8). While the antigenic features of a virus certainly contribute to its success, the non-antigenic effects of mutations also play an important role (9–12). Specif-

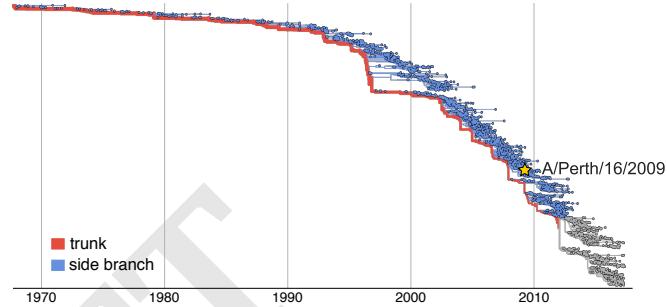


Fig. 1. Human H3N2 HA phylogeny from 1968–2017. The trunk is shown in red, and side branches are shown in blue. The gray branches represent the part of the tree for which we cannot yet distinguish the trunk from side branches. The Perth/2009 strain is indicated with a star.

ically, due to influenza virus's high mutation rate (13–15) and lack of intra-segment recombination (16), deleterious mutations become linked to beneficial ones. The resulting accumulation of deleterious mutations can affect non-antigenic properties that are central to viral fitness (11). However, there are no large-scale quantitative characterizations of how mutations to H3N2 HA affect non-antigenic properties crucial to viral fitness.

It is now possible to use deep mutational scanning (17) to measure the functional effects of large numbers of mutations to viral proteins (18–22). However, the only HA for which such large-scale measurements have previously been made is from

Significance Statement

A key goal in the study of influenza virus evolution is to forecast which viral strains will persist and which ones will die out. Here we experimentally measure the effects of all amino-acid mutations to the hemagglutinin protein from a human H3N2 influenza strain on viral growth in cell culture. We show that these measurements have utility for distinguishing among viral strains that do and do not succeed in nature. Overall, our work suggests that new high-throughput experimental approaches may be useful for evolutionary forecasting.

J.M.L. performed the experiments; J.M.L. and J.H. analyzed data with input from T.B. and J.D.B.; M.B.D. contributed essential reagents and expertise; K.A.H. contributed essential reagents; N.C.W. made key intellectual contributions; J.M.L. and J.D.B. wrote the paper with input from J.H. and T.B.; all authors were involved in project design and revising the paper

The authors declare no conflict of interest.

¹To whom correspondence should be addressed: trevor@bedford.io, jbloom@fredhutch.org

the highly lab-adapted A/WSN/1933 (H1N1) strain (18–20). Here we measure the effects on viral growth in cell culture of all mutations to the HA from a recent human H3N2 strain. We then apply these measurements to examine differences in maximum clade frequencies.

Results

Deep mutational scanning of HA from a recent strain of human H3N2 influenza virus. We performed a deep mutational scan to measure the effects of all amino-acid mutations to HA from the A/Perth/16/2009 (H3N2) strain on viral growth in cell culture. This strain was the H3N2 component of the influenza vaccine from 2010-2012 (23, 24). Relative to the consensus sequence for this HA in Genbank, we used a variant with two mutations that enhanced viral growth in cell culture, G78D and T212I (Figure S1 and Dataset S1). The G78D mutation occurs at low frequency in natural H3N2 sequences, and T212 is a site where a mutation to Ala rose to fixation in human influenza in ~2011.

We mutagenized the entire HA coding sequence at the codon level to create mutant plasmid libraries harboring an average of ~1.4 codon mutations per clone (Figure S2). We then generated mutant virus libraries from the mutant plasmids using a helper-virus system that enables efficient generation of complex influenza virus libraries (20) (Figure 2A). These mutant viruses derived all their non-HA genes from the lab-adapted A/WSN/1933 strain. Using WSN/1933 for the non-HA genes reduces biosafety concerns, and also helped increase viral titers. To further increase viral titers, we used MDCK-SIAT1 cells that we had engineered to constitutively express the TMPRSS2 protease, which cleaves the HA precursor to activate it for membrane fusion (25, 26).

After generating the mutant virus libraries, we passaged them at low MOI in cell culture to create a genotype-phenotype link and select for functional HA variants (Figure 2A). All experiments were completed in full biological triplicate (Figure 2B). We also passaged and deep sequenced library 3 in duplicate (denoted as library 3-1 and 3-2) to gauge the experimental noise occurring *within* a single biological replicate. As a control to measure sequencing and mutational errors, we used the unmutated HA gene to generate and passage viruses carrying wildtype HA.

Deep sequencing of the initial plasmid mutant libraries and the passaged mutant viruses revealed selection for functional HA mutants. Specifically, stop codons were purged to 20–45% of their initial frequencies after correcting for error rates estimated by sequencing the wildtype controls (Figure 2C). The incomplete purging of stop codons is likely because genetic complementation due to co-infection (27) enabled the persistence of some virions with nonfunctional HAs. We also observed selection against many nonsynonymous mutations (Figure 2C), with their frequencies falling to 30–40% of their initial values after error correction.

We next quantified the reproducibility of our deep mutational scanning measurements across biological and technical replicates. We first used the deep sequencing data for each replicate to estimate the preference of each site in HA for all 20 amino acids using the method described in (28). Because there are 566 residues in HA, there are $566 \times 19 = 10,754$ distinct measurements (the 20 preferences at each site are constrained to sum to one (28)). The correlations of the amino-acid pref-

Table 1. Substitution models informed by the experiments describe H3N2's natural evolution better than traditional substitution models.

Model	ΔAIC	LnL	Stringency	ω
ExpCM	0.0	-8441	2.47	0.91
GY94 M5	2094	-9482	—	0.36 (0.30, 0.84)
ExpCM, site avg.	2501	-9692	0.67	0.32
GY94 M0	2536	-9704	—	0.31

Maximum likelihood phylogenetic fit to an alignment of human H3N2 influenza HAs using ExpCM (30), ExpCM in which the experimental measurements are averaged across sites (site avg.), and the M0 and M5 versions of the Goldman-Yang (GY94) model (31). Models are compared by AIC (32) computed from the log likelihood (LnL) and number of model parameters. The ω parameter is dN/dS for the Goldman-Yang models, and the relative dN/dS after accounting for the measurements for the ExpCM. For the M5 model, we give the mean followed by the shape and rate parameters of the gamma distribution over ω .

erences between pairs of replicates are shown in Figure 2D. The biological replicates were well-correlated, with Pearson's R ranging from 0.69 to 0.78. Replicate 1 exhibited the weakest correlation with other replicates; this replicate also showed the weakest selection against stop and nonsynonymous mutations (Figure 2C), perhaps indicating more experimental noise. The two technical replicates 3-1 and 3-2 were only slightly more correlated than pairs of biological replicates, suggesting that bottlenecking of library diversity during viral passage contributes most of the experimental noise.

Our measurements are consistent with existing knowledge about HA's evolution and function. How do the HA amino-acid preferences measured in our experiments relate to the evolution of H3N2 influenza virus in nature? This question can be addressed by evaluating how well an experimentally informed codon substitution model (ExpCM) using our measurements describes H3N2 evolution compared to standard phylogenetic substitution models (29, 30). Table 1 shows that the ExpCM using the across-replicate average of our measurements greatly outperforms conventional substitution models. This result indicates that our experiments authentically capture some of the constraints on HA evolution. The relative rate of nonsynonymous to synonymous substitutions (dN/dS or ω) is $\ll 1$ for conventional substitution models (Table 1). However, the relative rate of nonsynonymous to synonymous substitutions after accounting for the amino-acid preferences measured in our experiments (ω for the ExpCM) is close to one (Table 1), indicating that most purifying selection against nonsynonymous substitutions is accounted for by the constraints measured in the deep mutational scanning. The ExpCM stringency parameter (30) is 2.47 (Table 1), indicating that natural selection favors the same amino acids as the experiments but with greater stringency. Throughout the rest of this paper, we use experimental measurements re-scaled (29, 30) by this stringency parameter. These re-scaled preferences are shown in Figure 3.

Examination of Figure 3 reveals that the experimentally measured amino-acid preferences generally agree with existing knowledge about HA's structure and function. For instance, sites that form structurally important disulfide bridges (sites 52 & 277, 64 & 76, 97 & 139, 281 & 305, 14 & 137-HA2,

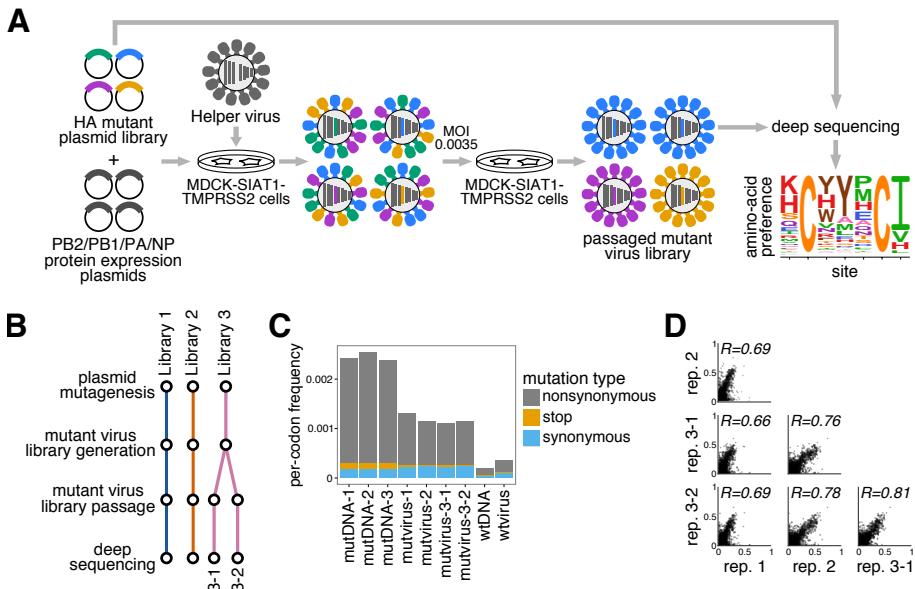


Fig. 2. Deep mutational scanning of the Perth/2009 H3 HA. (A) We generated mutant virus libraries using a helper-virus approach (20), and passaged the libraries at low MOI to establish a genotype-phenotype linkage and to select for functional HA variants. Deep sequencing of the variants before and after selection allowed us to estimate each site's amino-acid preferences. (B) The experiments were performed in full biological triplicate. We also passaged and deep sequenced library 3 in duplicate. (C) Frequencies of nonsynonymous, stop, and synonymous mutations in the mutant plasmid DNA, the passed mutant viruses, and wildtype DNA and virus controls. (D) The Pearson correlations among the amino-acid preferences estimated in each replicate.

144-HA2 & 148-HA2) (33) possess high preference for cysteine. At residues involved in receptor binding, there are strong preferences for the amino acids that are known to be involved in binding sialic acid, such as Y98, D190, W153, and S228 (34–37). A positively charged amino acid at site 329 is important for cleavage of the HA0 precursor into the mature form (38), and this site strongly prefers arginine. However, a notable exception occurs at the start codon at position -16, which does not show a strong preference for methionine. This codon is part of the signal peptide and is cleaved from the mature HA protein. One possible reason that our experiments do not show a strong preference for methionine at this site could be alternative translation-initiation at a downstream or upstream start site, as has been described for other HAs (39).

There is less difference in mutational tolerance between the HA head and stalk domains for H3 than for H1. Our experiments measure which amino acids are tolerated at each HA site under selection for protein function. We can therefore use our experimentally measured amino-acid preferences to calculate the inherent mutational tolerance of each site, which we quantify as the Shannon entropy of the re-scaled preferences. In prior mutational studies of H1 HAs, the stalk domain was found to be substantially less mutationally tolerant than the globular head (18–20, 41).

We performed a similar analysis using the current data for the Perth/2009 H3 HA. Surprisingly, there was much less contrast in mutational tolerance between the stalk and head domains for the H3 HA than for the H1 (Figure 4). Specifically, for solvent-exposed residues, the head domain is more mutationally tolerant than the stalk domain for the WSN/1933 H1 HA, whereas such a difference in tolerances was not seen for the Perth/2009 H3 HA (Figure 4B), and this was true using different relative solvent accessibility cutoffs to define solvent-exposed residues (Figure S5). Furthermore, in the H3 HA, the short helix A in the stalk was found to be very mutationally tolerant. Interestingly, there are more reports of readily selecting escape mutants from broadly neutralizing anti-stalk antibodies in H3 (42–45) than in H1 (46–48) HAs.

We also see high mutational tolerance in many of the known

antigenic regions of H3 HA (49). For instance, antigenic region B is an immunodominant area, and many recent major antigenic drift mutations have occurred in this region (6, 7, 50). We find that the most distal portion of the globular head near the 190-helix, which is part of antigenic region B, is highly tolerant of mutations (Figure 4). Antigenic region C is also notably mutationally tolerant.

Many residues inside HA's receptor binding pocket are known to be highly functionally constrained (35, 51), and our data indicates that these sites are relatively mutationally intolerant in both H3 and H1 HAs. In contrast, the residues surrounding the receptor binding pocket are fairly mutationally tolerant, which may contribute to the rapidity of influenza's antigenic evolution, since mutations at these sites can have large effects on antigenicity (6, 49).

Our measurements reveal differences between clades of low and high maximum frequencies. The maximum frequencies achieved by particular mutations are a reflection of the fitnesses of strains carrying those mutations. Do the experimentally-determined preferences therefore reveal underlying differences in strains that reach various maximum frequencies? To explore this question, we examined the relationship between the experimentally measured effects of all mutations and the maximum frequency achieved by all nodes with at least one tip in a maximum-likelihood H3N2 phylogeny. We estimated the frequency trajectories for every node in the tree by... [I'll let John / Trevor add more here]. Evolutionarily successful nodes that occurred *before* the Perth/2009 strain diverged from the trunk of the phylogenetic tree tend to carry mutations that move towards the Perth/2009 sequence. Because the effects of mutations may vary depending on genetic background, we analyzed nodes that preceded the Perth/2009 strain separately from nodes that occurred after the split of Perth/2009 from the trunk of the phylogenetic tree [need better explanation of pre-and post-Perth/2009 rationale here since it seems like we are going to completely drop the trunk/tip analysis from the paper? Also this is where we would reference the tree supplemental figure]. We found that mutations measured to be very deleterious to viral growth in our experiments were not observed at high

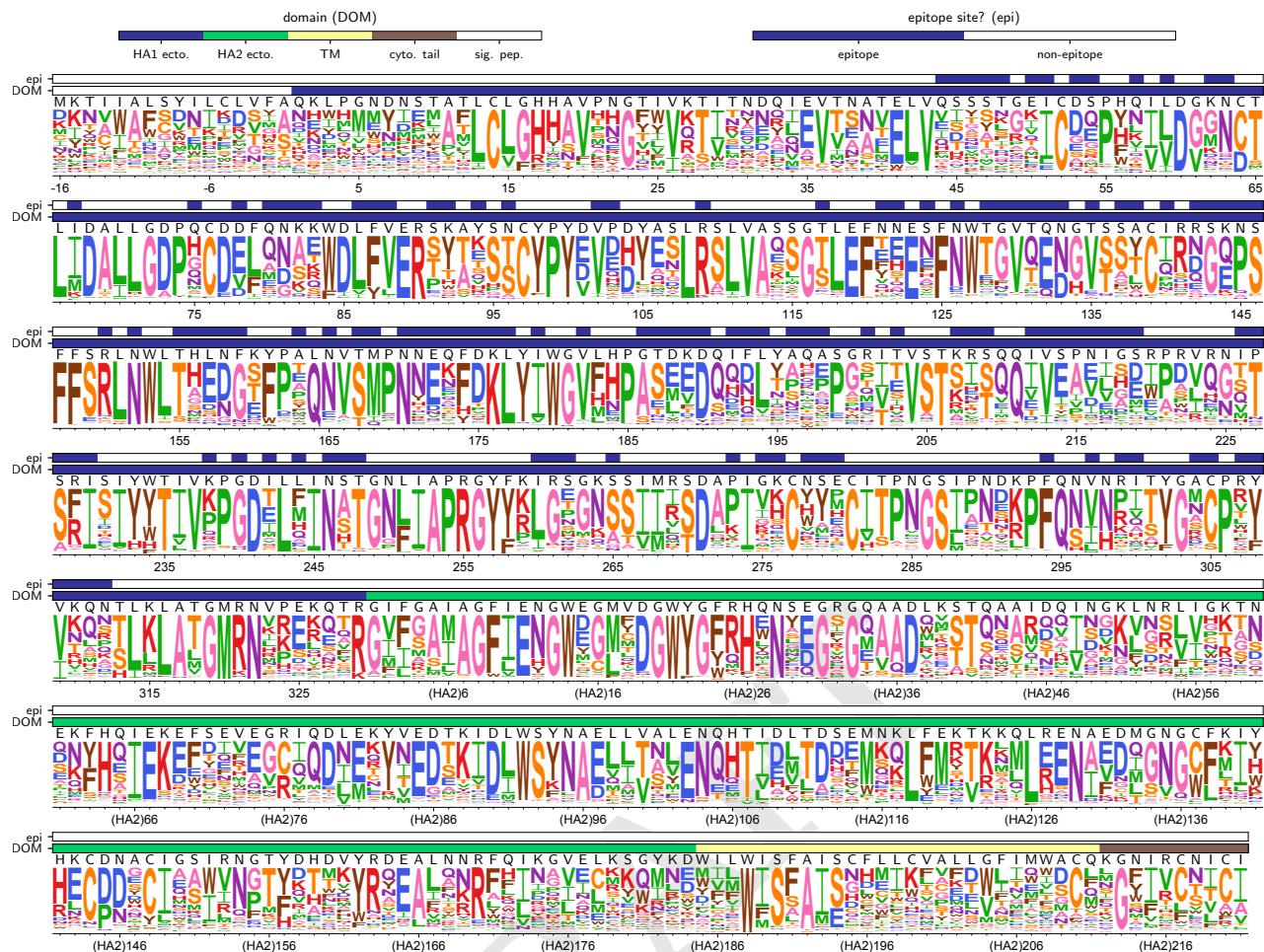


Fig. 3. The site-specific amino-acid preferences of the Perth/2009 HA measured in our experiments. The height of each letter is the preference for that amino acid, after taking the average over experimental replicates and re-scaling (30) by the stringency parameter in Table 1. The sites are in H3 numbering. The top overlay bar indicates whether or not a site is in the set of epitope residues delineated in (40). The bottom overlay bar indicates the HA domain (sig. pep. = signal peptide, HA1 ecto. = HA1 ectodomain, HA2 ecto. = HA2 ectodomain, TM = transmembrane domain, cyto. tail. = cytoplasmic tail). The letters directly above each logo stack indicate the wildtype amino acid at that site.

frequencies in nature. Rather, only mutations with beneficial or small deleterious effects in our experiments appeared to reach high frequencies or even fix in nature, as evidenced by the slight but significant[?] trend towards higher frequencies with more favorable mutations (Figure 5). Furthermore, shifts toward more favorable mutations with increasing maximum frequency were also apparent from plotting the distribution of mutational effects for different frequency bins. Mutations were on average scored as being more detrimental in the lower frequency bins than at higher frequencies. The trend towards more beneficial mutations with increasing maximum frequency was stronger for the pre-Perth/2009 nodes, which may be somewhat driven by high-frequency mutations generally moving towards the Perth/2009 sequence and thereby scored more favorably. Despite this, the post-Perth/2009 nodes exhibited mutations that were on average more favorable for viral growth at higher frequency nodes. Notably, nodes that reached moderate global frequencies ($> 10\%$) had relatively few mutations that were highly deleterious compared to nodes that remained below 2% frequency, suggesting that unfavorable mutations may help determine whether or not a strain is able to surpass low frequencies in the global population. Overall, these results

demonstrate that measurements of mutational effects in the Perth/2009 HA reveal differences in fitness between clades that circulate at low frequencies and those that achieve fixation or near-fixation.

To determine how broadly experimental measurements can be generalized across HAs, we repeated the maximum frequency analyses using amino-acid preferences from our prior deep mutational scanning of the WSN/1933 H1 HA, which is highly diverged from the Perth/2009 H3 HA. Figure 6 shows that the H1 measurements do *not* reveal a trend towards more beneficial mutations for higher-frequency nodes of an H3 phylogeny. This observation indicates that the utility of an experiment for revealing differences between low and high frequency clades degrades over sufficiently long evolutionary distances.

There are large differences between H3 and H1 HAs in the amino-acid preferences of many sites. An obvious hypothesis for why the H1 deep mutational scanning do not reveal differences among H3 strains is that the effect of the same mutation is often different between these two HA subtypes. To determine if this is the case, we examined how much the amino-acid

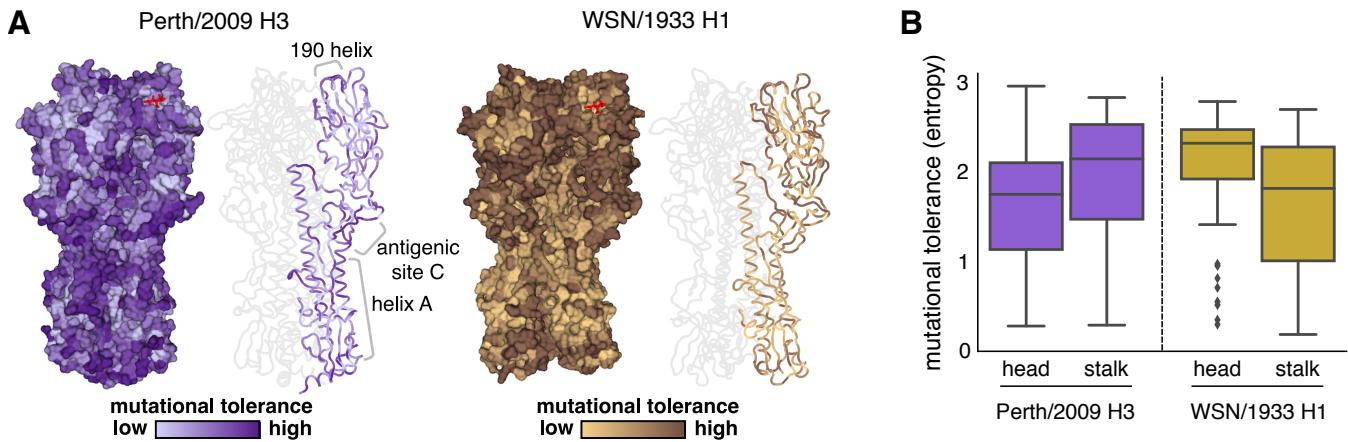


Fig. 4. Mutational tolerance of each site in H3 and H1 HAs. (A) Mutational tolerance as measured in the current study is mapped onto the structure of the H3 trimer (PDB 4O5N; [\(52\)](#)). Mutational tolerance of the WSN/1933 H1 HA as measured in [\(20\)](#) is mapped onto the structure of the H1 trimer (PDB 1RVX; [\(53\)](#)). Different color scales are used because measurements are comparable among sites within the same HA, but not necessarily across HAs. Both trimers are shown in approximately the same orientation. For each HA, the structure at left shows a surface representation of the full trimer, while the structure at right side shows a ribbon representation of just one monomer. The sialic acid receptor is shown in red sticks. (B) The mutational tolerance (Shannon entropy of re-scaled preferences) of solvent-exposed residues with a relative solvent accessibility ≥ 0.2 in the head and stalk domains of the Perth/2009 H3 HA (purple) and WSN/1933 H1 HA (gold). We quantified the absolute solvent accessibility of each site using DSSP [\(54\)](#) and then normalized to a relative solvent accessibility using the absolute accessibilities in [\(55\)](#). Residues falling in between the two cysteines at sites 52 and 277 were defined as belonging to the head domain, while all other residues were defined as the stalk domain. Note that the mutational tolerance values are not comparable between the two HAs.

preferences of homologous sites have shifted between H3 and H1 HAs. Prior experiments have found only modest shifts in amino-acid preferences between two variants of influenza nucleoprotein with 94% amino-acid identity [\(56\)](#) and variants of HIV envelope (Env) with 86% amino-acid identity [\(57\)](#). However, the H1 and H3 HAs are far more diverged, with only 42% amino-acid identity (Figure 7A). One simple way to investigate the extent of shifts in amino-acid preferences is to correlate measurements from independent deep mutational scanning replicates on H1 and H3 HAs. Figure 7B shows that replicate measurements on the same HA variant are more correlated than those on different HA variants.

To more rigorously quantify shifts in amino-acid preferences after correcting for experimental noise, we used the statistical approach in [\(56, 57\)](#). Figure 7C shows the distribution of shifts in amino-acid preferences between H3 and H1 HAs after correcting for experimental noise. Although some sites have small shifts near zero, many sites have large shifts. These shifts between H3 and H1 are much larger than expected from the null distribution that would be observed purely from experimental noise. They are also much larger than the shifts previously observed between two HIV Envs with 86% amino-acid identity [\(57\)](#). However, the typical shift between H3 and H1 is still smaller than that observed when comparing HA to the non-homologous HIV Env protein. Therefore, there are very substantial shifts in mutational effects between highly diverged HA homologs, although the effects of mutations remain more similar than for non-homologous proteins.

Properties associated with the shifts in amino-acid preferences between H3 and H1 HAs. What features distinguish the sites with shifted amino-acid preferences between H3 and H1 HAs? The sites of large shifts do not obviously localize to one specific region of HA's structure (Figure 8A). However, at the domain level, sites in HA's stalk tend to have smaller shifts than sites in HA's globular head (Figure 8B). The HA stalk domain is also more conserved in sequence [\(58\)](#), suggesting

that conservation of amino-acid sequence is correlated with conservation of amino-acid preferences. Consistent with this idea, sites that are absolutely conserved across all 18 HA subtypes are significantly less shifted than sites that are variable across HA subtypes (Figure 8B). Presumably these sites are under consistent functional constraint across all HAs.

Despite their high sequence divergence, H1 and H3 adopt very similar protein folds [\(59, 60\)](#). However, there are differences in the rotation and upward translation of the globular head subdomains relative to the central stalk domain among different HA subtypes [\(59, 60\)](#). Previous work has defined clades of structurally related HA subtypes [\(59, 60\)](#). One such clade includes H1, H2, H5, and H6, whereas another clade includes H3, H4, and H14 HAs (Figure 7A). Sites that are conserved at different amino-acid identities in these two clades tend to have exceptionally large shifts in amino acid preferences (Figure 8B). The clade containing H1 has an upward translation of the globular head relative to the clade containing H3. This structural shift has been attributed largely to the interaction between sites 107 and 75(HA2) [\(59, 60\)](#). Specifically, the clade containing H1 has a taller turn in the interhelical loop connecting helix A and helix B in the stalk domain, and this tall turn is stabilized by a hydrogen bond between Glu-107 and Lys-75(HA2) (Figure 8C). In deep mutational scanning of the H1 HA, site 107 has a high preference for Glu and 75(HA2) strongly prefers positively charged Lys and Arg. In contrast, the interhelical loop in H3 HA makes a sharper and shorter turn which is facilitated by a Gly at 75(HA2). In the deep mutational scanning of the Perth/2009 H3 HA, site 75(HA2) prefers Gly and to a lesser extent Val, while site 107 is fairly tolerant of mutations. Therefore, some of the shifts in HA amino-acid preferences can be rationalized in terms of changes in HA structure.

A/Perth/16/2009 preferences

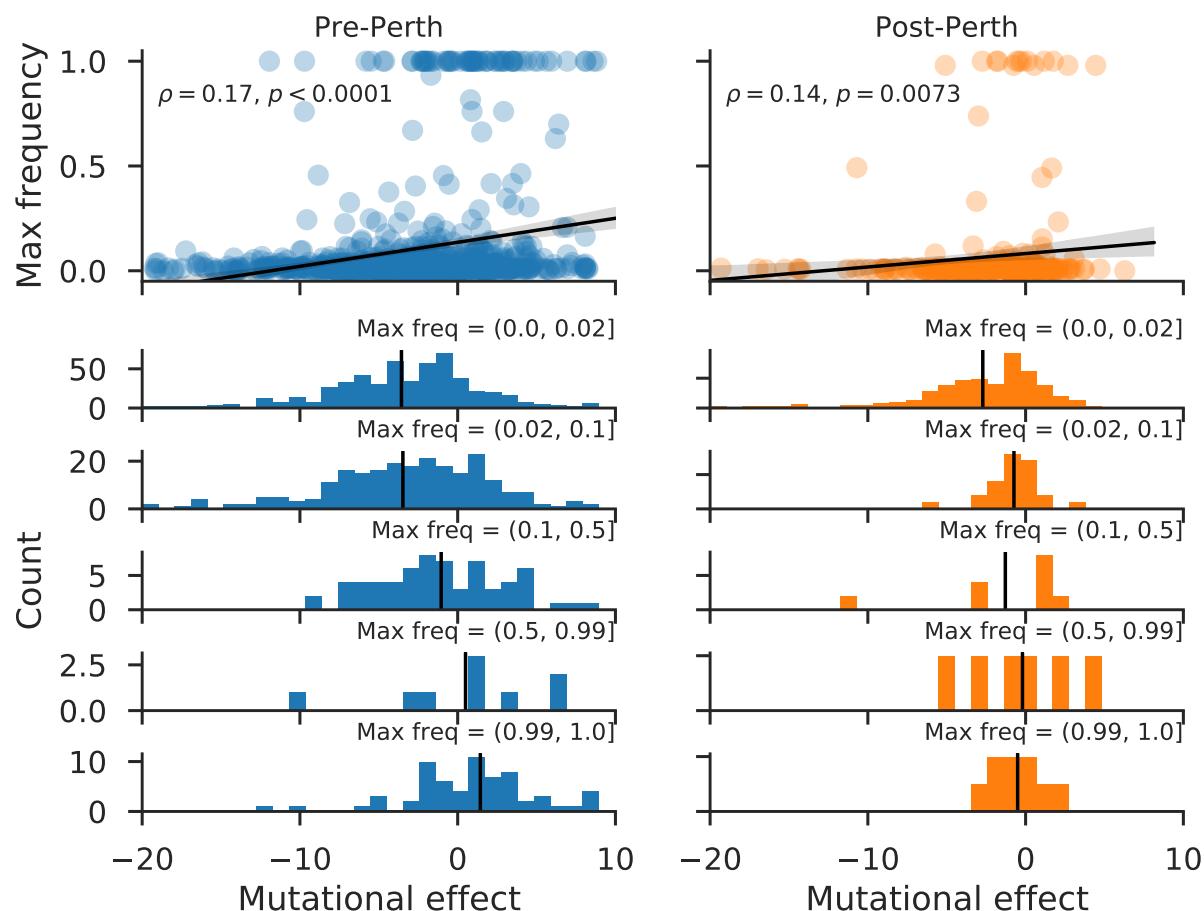


Fig. 5. Highly deleterious mutations are not observed in high-frequency nodes. Figure caption.

Discussion

We have measured the effects of all possible single amino-acid mutations to the Perth/2009 H3 HA on viral growth in cell culture and demonstrated that these measurements are informative for distinguishing the evolutionary fate of human H3N2 viral strains in nature. Specifically, HA sequences on the trunk of the phylogenetic tree tend to have mutations that our experiments measure to be more beneficial for viral growth than mutations that occur in side-branch HA sequences. The fact that our experiments can help distinguish between trunk and side-branch lineages suggests that they might inform evolutionary forecasting. In their landmark paper introducing predictive viral fitness models that accounted for both antigenic and non-antigenic mutations, Łukasza and Lässig (11) noted that the models could in principle be improved by integrating “diverse genotypic and phenotypic data” that more realistically represented the effects of specific mutations. Our work suggests that deep mutational scanning can provide such data.

It is important to emphasize that measurements of viral growth in cell culture do *not* represent true fitness in nature. Indeed, a vast amount of work in virology has chronicled the many ways in which experiments can select for lab artifacts

or fail to capture important pressures that are relevant in nature (61–64). Mutations in viral genes other than HA are also important in determining strain success (65, 66). Given these caveats, it might seem surprising that measuring viral growth in cell culture can be informative about the success of viral strains in nature. Yet, prior to our work, there were no comprehensive studies of the functional effects of mutations to H3 HA on any property that even resembled viral fitness in nature, and modeling work has either omitted the non-antigenic effects of mutations (3–5) or assumed that all non-epitope mutations had equivalent deleterious effects (11). The strength of our measurements are not that they perfectly capture fitness in nature, but that they are systematic and quantitative — and represent a vast improvement over no information at all. That this is true is demonstrated by the fact that these measurements have utility for distinguishing successful and unsuccessful viral lineages over ~50 years of H3N2 evolution. We suspect that performing the experiments using more realistic and complex selections (e.g., ferrets or primary human airway cultures) might further improve their utility.

We measured the effects of all single amino-acid mutations to a specific HA, and then generalized these measurements

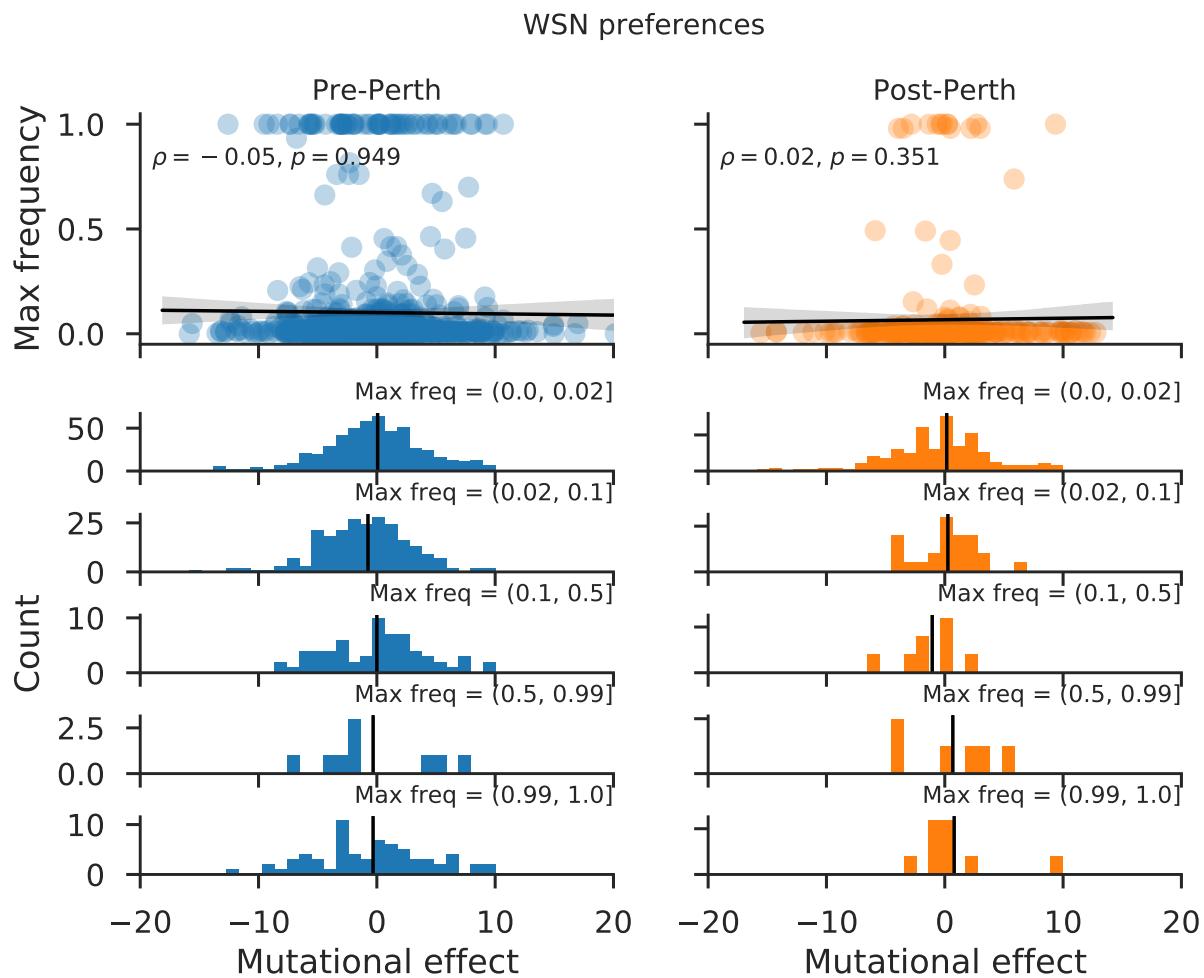


Fig. 6. Experimental measurements on an H1 HA do not reveal differences between low- and high-frequency nodes. Figure caption.

to other H3N2 HAs from a 50-year timespan. These generalizations will only be valid to the extent that the effects of mutations are conserved during HA's evolution. Extensive work on protein evolution has shown that epistasis can shift the effects of mutations (57, 67–71), and the extent to which this occurs is an important question in its own right. The fact that our measurements score trunk mutations more favorably than side branch ones across a 50-year timespan suggests that the mutational effects have not shifted too dramatically among these relatively closely related sequences. On the other hand, when we compared our measurements for an H3 HA to prior measurements on H1 HA, we found substantial shifts at many sites — much greater than those observed in prior protein-wide comparisons of more closely related homologs (56, 57). Further investigation of how mutational effects shift as proteins diverge will be important for determining how broadly any given experiment can be generalized when attempting to make evolutionary forecasts.

Our work did not characterize the antigenic effects of mutations, which also play an important role in determining strain success in nature (5, 6). However, our basic selection and deep-sequencing approach can be harnessed to completely map how mutations affect antibody recognition (47, 72). But

so far, experiments using this approach have not examined antibodies or sera that are relevant to driving the evolution of H3N2 influenza (47, 72), or have used relevant sera but examined a non-comprehensive set of mutations (8). Future experiments that completely map how HA mutations affect recognition by human sera seem likely to be especially fruitful for informing viral forecasting.

Materials and Methods

Data and computer code. Deep sequencing data are available from the Sequence Read Archive under BioSample accessions SAMN08102609 and SAMN08102610. Computer code used to analyze the data and produce the results in the paper are on GitHub at [repository will be made public upon acceptance of paper; if you are a reviewer who wants access please contact the editor].

HA numbering. Sites are in H3 numbering, with the signal peptide in negative numbers, HA1 in plain numbers, and HA2 denoted with "(HA2)". Sequential 1, 2, ... numbering of the Perth/2009 HA can be converted to H3 numbering by subtracting 16 for the HA1 subunit, and subtracting 345 for the HA2 subunit.

Creation of MDCK-SIAT1-TMPRSS2 cells. When growing influenza virus in cell culture, trypsin is normally added in order to cleave

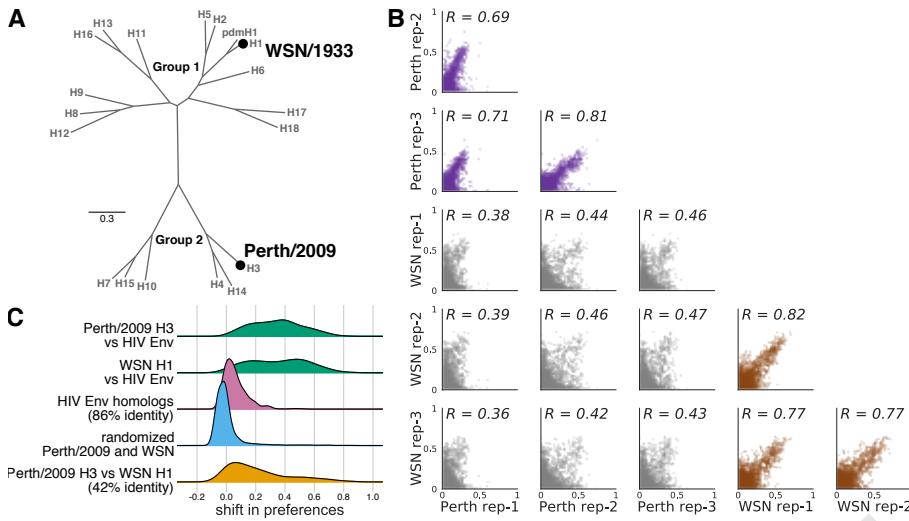


Fig. 7. There are substantial differences in the effects of mutations between H1 and H3 HAs. (A) Phylogenetic tree of HA subtypes, with the WSN/1933 H1 and Perth/2009 H3 HAs labeled. These HAs have 42% amino-acid identity. (B) All pairwise correlations of the amino-acid preferences measured in the three individual deep mutational scanning replicates in the current study and the three replicates in prior deep mutational scanning of an H1 HA (20). Comparisons between H3 replicates are in purple, those between H1 replicates are in brown, and those across H1 and H3 replicates are in gray. R indicates the Pearson correlation coefficient. (C) We calculated the shift in amino-acid preferences at each site between H3 and H1 HAs using the method in (57), and plotted the distribution of shifts for all sites. The shifts between H3 and H1 (yellow) are much larger than the null distribution (blue) expected if all differences are due to experimental noise. The shifts are also much larger than those previously observed between two variants of HIV Env that share 86% amino-acid identity (pink). However, the shifts between H3 and H1 are usually less than the differences between HA and HIV Env (green).

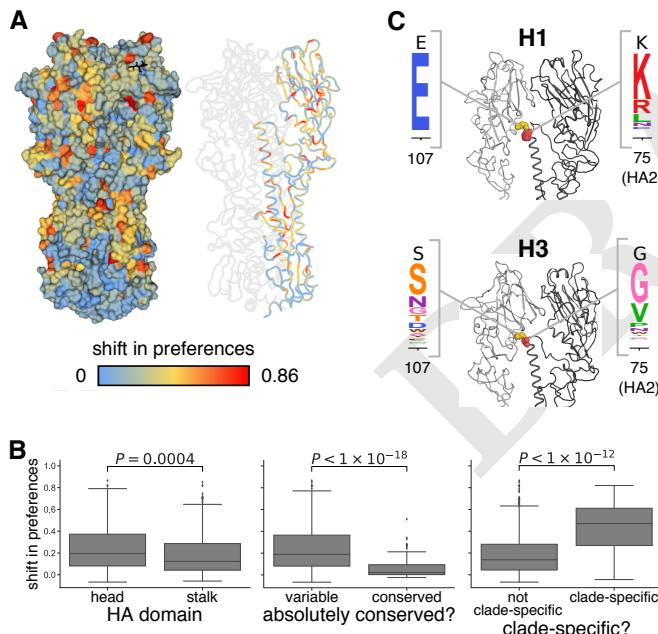


Fig. 8. Sites with strongly shifted amino-acid preferences between H3 and H1 HAs. (A) The shift in amino-acid preferences between the H3 and H1 HA at each site as calculated in Figure 7C is mapped onto the structure of the H3 HA. (B) Amino-acid preferences of sites in the stalk domain are less shifted than those in the head domain. Sites absolutely conserved in all 18 HA subtypes are less shifted than other sites. Sites with one amino-acid identity in the clade containing H1, H2, H5, and H6 and another identity in the clade containing H3, H4, and H14 are more shifted than other sites. (C) Sites 107 and 75(HA2) help determine the different orientation of the globular head domain in H1 versus H3 HAs. These sites are shown in spheres on the structure of H1 and H3 and colored as in panel (A), and the experimentally measured amino-acid preferences in the H1 and H3 HAs are shown. One monomer is in dark gray, while the HA1 domain of the neighboring monomer is in lighter gray.

HA into its mature form. To obviate the need for trypsin, we engineered an MDCK-SIAT1 cell line to constitutively express the TMPRSS2 protease, which cleaves and activates HA in the human airways (25, 26). The human TMPRSS2 cDNA ORF was ordered from OriGene (NM_005656) and cloned into a pHAGE2 lentiviral vector under an EF1 α -Int promoter followed by an IRES driving expression of mCherry to create plasmid pHAGE2-EF1 α Int-TMPRSS2-IRES-mCherry-W. We used the lentiviral vector to transduce MDCK-SIAT1 cells, and sorted an intermediate mCherry-positive population by flow cytometry. We refer to the sorted bulk population as MDCK-SIAT1-TMPRSS2 cells. There is no selectable marker for the TMPRSS2; however, we maintain the cells at low passage number, and have seen no indication that they lose their ability to support the growth of viruses with H3 HAs in the absence of exogenous trypsin.

Generation of HA codon mutant plasmid libraries. The HA and NA genes for the Perth/2009 viral strain were reverse-transcribed and amplified from recombinant virus obtained from BEI Resources (NR-41803), and cloned into the pHW2000 (73) influenza reverse-genetics plasmids to create pHW-Perth09-HA and pHW-Perth09-NA.

We initially created a virus with the HA and NA from Perth/2009 and internal genes from WSN/1933, and passaged it in cell culture to test its genetic stability. To generate this virus, we transfected a co-culture of 293T and MDCK-SIAT1-TMPRSS2 in D10 media (DMEM supplemented with 10% heat-inactivated FBS, 2 mM L-glutamine, 100 U of penicillin/mL, and 100 μ g of streptomycin/mL) with equal amounts of pHW-Perth09-HA, pHW-Perth09-NA, the pHW18* series of plasmids (73) for all non-HA/NA viral genes, and pHAGE2-EF1 α Int-TMPRSS2-IRES-mCherry-W. The next day we changed the media to influenza growth media (IGM, consisting of Opti-MEM supplemented with 0.01% heat-inactivated FBS, 0.3% BSA, 100 U of penicillin/mL, 100 μ g of streptomycin/mL, and 100 μ g of calcium chloride/mL — no trypsin was added since there was TMPRSS2), and then we collected the viral supernatant at 72 hours post-transfection. This viral supernatant was then blind passaged in MDCK-SIAT1-TMPRSS2 a total of six additional times. We isolated viral RNA from these passaged viruses and sequenced the HA gene. The passaged HA had two mutations, G78D and T212I, which we found enhanced viral growth as shown in Figure S1. The HA with these two mutations was cloned into pHW2000 (73) and pICR2 (74) to create pHW-Perth09-HA-G78D-T212I and pICR2-Perth09-HA-G78D-T212I. For all subsequent experiments, we used viruses with the HA containing these two mutations in order to improve titers and viral genetic stability, and this is the HA that we refer to as Perth/2009 in the manuscript text. We also used all non-HA genes (including NA) from WSN/1933 to help increase titers and reduce biosafety concerns.

The codon-mutant libraries were generated in the Perth/2009 HA-G78D-T212I background using the PCR-based approach described in (75) with the primer melting-temperature modifications described in (76), using two rounds of mutagenesis. The script to design the mutagenesis primers is at <https://github.com/jbloomlab/CodonTilingPrimers>. We created three independent libraries, one for each biological replicate. The mutant variants were then cloned at high efficiency into the pICR2 (74) vector using digestion with BsmBI, ligation with T4 DNA ligase, and electroporation into ElectroMAX DH10B competent cells (Invitrogen 18290015). We obtained >6 million transformants for each replicate. We scraped the plates, expanded the cultures in liquid LB + ampicillin at 37°C for 3 h with shaking, and then maxiprepped. We randomly chose 31 clones to Sanger sequence to evaluate the mutation rate (Figure S2).

Generation and passaging of mutant viruses. The mutant virus libraries were generated using the helper-virus approach described in (20) with several modifications, most notably the cell line used. Briefly, we transfected 5×10^5 MDCK-SIAT1-TMPRSS2 cells in suspension with 937.5 ng each of four protein expression plasmids encoding the ribonucleoprotein complex (HDM-Nan95-PA, HDM-Nan95-PB1, HDM-Nan95-PB2, and HDM-Aichi68-NP) (67), and 1250 ng of one of the three pICR2-mutant-HA libraries (or the wild-type control) using Lipofectamine 3000 (ThermoFisher L3000008). We allowed the transfected cells to adhere in 6-well plates and four hours later changed the media to D10 media. Eighteen hours after transfection, we infected the cells with the WSN/1933 HA-deficient helper virus (20) by preparing an inoculum of 500 TCID₅₀ per μL of helper virus (as computed on HA-expressing cells) in IGM, aspirating the D10 media from the cells, and adding 2 mL of the helper-virus inoculum to each well. After three hours, we changed the media to fresh IGM. At 24 hours after helper-virus infection, we harvested the viral supernatants for each replicate, froze aliquots at -80°C, and titered in MDCK-SIAT1-TMPRSS2 cells. The titers were 92, 536, 536, and 734 TCID₅₀ per μL for the three library replicates and the wildtype control, respectively.

We passaged 9×10^5 TCID₅₀ of the transfection supernatants at an MOI of 0.0035 TCID₅₀ per cell. To do this, we plated 4.6×10^6 MDCK-SIAT1-TMPRSS2 cells per dish in 15-cm dishes in D10 media, and allowed the cells to grow for 24 hours, at which time they had reached a density of $\sim 1.7 \times 10^7$ cells per dish. We then replaced the media in each dish with 25 mL of an inoculum of 2.5 TCID₅₀ of virus per μL . We collected viral supernatant for sequencing 48 hours post-infection.

Barcoded subamplicon sequencing. To extract viral RNA from the three replicate HA virus libraries and the wildtype HA virus, we ultracentrifuged 24 mL of the supernatant at 22,000 rpm for 1.5 h at 4°C in a Beckman Coulter SW28 rotor, and extracted RNA using the Qiagen RNeasy Mini Kit by resuspending the viral pellet in 400 μL of buffer RLT supplemented with β -mercaptoethanol, pipetting 30 times, transferring the liquid to a microcentrifuge tube, adding 600 μL 70% ethanol, and proceeding with the extraction according to the manufacturer's instructions. The HA gene was then reverse-transcribed with AccuScript Reverse Transcriptase (Agilent 200820) using the primers P09-HA-For (5'-AGCAAAAGCAGGGGATAATTCTATTAAATC-3') and P09-HA-Rev (5'-AGTAGAACAAAGGGTGTAAATTACTAAATACAC-3').

We generated the HA PCR amplicons for the three plasmid libraries, the three virus libraries, the wildtype plasmid control, and the wildtype virus control using KOD Hot Start Master Mix (EMD Millipore 71842) using the PCR reaction mixture and cycling conditions described in (75) and the P09-HA-For and P09-HA-Rev primers. We prepared the sequencing libraries using a barcoded-subamplicon strategy (19) to increase the accuracy from deep sequencing. The exact details of this approach are described in (20) (also see https://jbloomlab.github.io/dms_tools2/bcsubamp.html). The primers used to generate the subamplicons are in Dataset S2. We performed deep sequencing on a lane of an Illumina HiSeq 2500 using 2×250 bp paired-end reads in rapid-run mode.

Analysis of deep sequencing data. We used the dms_tools2 software package (28) (https://github.com/jbloomlab/dms_tools2, version 2.2.0) to analyze the deep sequencing data. The algorithm used to estimate

the site-specific amino-acid preferences from the deep sequencing counts is described in (28). The amino-acid preferences for each replicate and for the re-scaled, across-replicate average are provided in Dataset S3. Computer code that performs the entirety of this analysis is available on GitHub at [repository will be made public upon acceptance of paper; if you are a reviewer who wants access please contact the editor].

Phylogenetic model comparison and fitting of a stringency parameter. For the analysis in Table 1, we downloaded all full-length H3 HA sequences from the Influenza Virus Resource (77), and randomly subsampled two sequences per year. These sequences were aligned using MAFFT (78) and used to infer a phylogenetic tree using RAxML (79) with a GTRCAT model of nucleotide substitution. We then used phydms (30) (<https://github.com/jbloomlab/phydms>, version 2.2.1) to fit the substitution models listed in Table 1. The amino-acid preferences were re-scaled by the stringency parameter using the approach described in (30).

The phylogenetic tree of HA subtypes in Figure 7A was generated as described in (47).

Inference of human H3N2 phylogenetic tree. We applied Nextstrain's augur pipeline (80) (<https://github.com/nextstrain/augur>; commit 006896d) to publicly available H3N2 HA sequences from GISAID, sampling six viruses per month over the time interval of January 1, 1968 to April 1, 2017. We aligned the resulting 2,139 HA sequences with MAFFT v7.310 (78) and constructed a maximum likelihood phylogeny from this alignment with RAxML 8.2.10 (79). Ancestral state reconstruction and branch length timing were performed with treetime (81). We pruned seven strains with premature stop codons in their amino acid translations that represented likely sequencing errors, producing a final tree with 2,132 strains.

To parse trunk mutations from side-branch mutations, we defined a set of recent nodes sampled on or after January 1, 2017, and traced these nodes back to their most recent common ancestor. We defined all branches ancestral to this most recent common ancestor as the trunk. All nodes that do not have descendants isolated after 2017 were defined as side-branch nodes. Nodes descended from the most recent common ancestor were defined as unresolved, as it is still unclear whether these nodes will ultimately be on the trunk or side branches. The phylogenetic tree is available as a JSON file on GitHub at [repository will be made public upon acceptance of paper; if you are a reviewer who wants access please contact the editor]. The tree was visualized using baltic (<https://github.com/blab/baltic>).

Quantification of mutational effects and sequence preferences from an H3N2 phylogeny. We calculated the effects of mutations from the amino-acid preferences that were estimated from the deep mutational scanning data. The effect of mutating site r from amino-acid a_1 to a_2 was calculated as

$$\log_2 \frac{\pi_{r,a_2}}{\pi_{r,a_1}} \quad [1]$$

where π_{r,a_1} and π_{r,a_2} are the re-scaled preferences for amino acids a_1 or a_2 at site r as shown in Figure 3. The WSN/1933 H1 HA amino-acid preferences are the replicate-average values reported in (20), re-scaled by a stringency parameter of 2.05 (see https://github.com/jbloomlab/dms_tools2/blob/master/examples/Doud2016/analysis_notebook.ipynb). The mutational effects were also calculated for five-year windows from 1968 to 2012.

The average sequence preference per site for every node on the phylogenetic tree shown in Figure 1 was calculated using Equation 1 for all sites, or the set of epitope and non-epitope sites defined in (40) and provided in Dataset S3.

Analysis of mutational shifts. To compare the Perth/2009 H3 and WSN/1933 H1 HA preferences, we first aligned the wildtype HA sequences using MAFFT (78). To quantify the shifts in preference for every alignable site while accounting for experimental noise, we used the approach described in (57) and used the RMSD_{corrected} values as our quantification of the extent of each shift.

For the plots shown in Figure 8B, any residues falling between Cys-52 and Cys-277 were defined as the head domain, and all other residues were defined as the stalk domain. We used the multiple sequence alignment of the HA subtype sequences from (47) to

identify sites that are absolutely conserved across all subtypes, or in the different clades described in Figure 8.

ACKNOWLEDGMENTS. We thank Sarah Hilton, Hugh Haddox, and Sidney Bell for helpful discussions about data analysis and we thank Richard Neher for sharing analysis code. We thank the Fred Hutch Genomics Core for performing the Illumina deep sequencing. This work was supported by grant R01 AI127893 from the NIAID of the NIH to JDB and TB and grant U19 AI117891 to TB. JML was supported in part by the Center for Inference and Dynamics of Infectious Diseases (CIDID), which is funded by grant U54GM111274 from the NIGMS of the NIH. The research of JDB is supported in part by a Faculty Scholar grant from the Howard Hughes Medical Institute and the Simons Foundation, and a Burroughs Wellcome Young Investigator in the Pathogenesis of Infectious Diseases grant. TB is a Pew Biomedical Scholar and is supported by NIH R35 GM119774.

References.

- Smith DJ, et al. (2004) Mapping the antigenic and genetic evolution of influenza virus. *Science* 305(5682):371–376.
- Bhatt S, Holmes EC, Pybus OG (2011) The genomic rate of molecular adaptation of the human influenza A virus. *Molecular Biology and Evolution* 28(9):2443.
- Sun H, et al. (2013) Using sequence data to infer the antigenicity of influenza virus. *MBio* 4(4):e00230–13.
- Harvey WT, et al. (2016) Identification of low-and high-impact hemagglutinin amino acid substitutions that drive antigenic drift of influenza A (H1N1) viruses. *PLoS Pathogens* 12(4):e1005526.
- Neher RA, Bedford T, Daniels RS, Russell CA, Shraiman BI (2016) Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. *Proceedings of the National Academy of Sciences* 113(12):E1701–E1709.
- Koel BF, et al. (2013) Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science* 342(6161):976–979.
- Chambers BS, Parkhouse K, Ross TM, Alby K, Hensley SE (2015) Identification of hemagglutinin residues responsible for H3N2 antigenic drift during the 2014–2015 influenza season. *Cell Reports* 12(1):1–6.
- Li C, et al. (2016) Selection of antigenically advanced variants of seasonal influenza viruses. *Nature Microbiology* 1:16058.
- Pybus OG, et al. (2007) Phylogenetic evidence for deleterious mutation load in RNA viruses and its contribution to viral evolution. *Molecular Biology and Evolution* 24(3):845–852.
- Strelkowa N, Lässig M (2012) Clonal interference in the evolution of influenza. *Genetics* 192(2):671–682.
- Łukasz M, Lässig M (2014) A predictive fitness model for influenza. *Nature* 507(7490):57–61.
- Koelle K, Rasmussen DA (2015) The effects of a deleterious mutation load on patterns of influenza A/H3N2's antigenic evolution in humans. *eLife* 4:e07361.
- Holland J, et al. (1982) Rapid evolution of RNA genomes. *Science* 215(4540):1577–1585.
- Steinhauer D, Holland J (1987) Rapid evolution of RNA viruses. *Annual Reviews in Microbiology* 41(1):409–431.
- Lauring AS, Andino R (2010) Quasispecies theory and the behavior of RNA viruses. *PLoS Pathogens* 6(7):e1001005.
- Boni MF, Zhou Y, Taubenberger JK, Holmes EC (2008) Homologous recombination is very rare or absent in human influenza A virus. *Journal of Virology* 82(10):4807–4811.
- Fowler DM, Fields S (2014) Deep mutational scanning: a new style of protein science. *Nature methods* 11(8):801–807.
- Thyagarajan B, Bloom JD (2014) The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin. *eLife* 3:e03300.
- Wu NC, et al. (2014) High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution. *Scientific Reports* 4:4942.
- Doud MB, Bloom JD (2016) Accurate measurement of the effects of all amino-acid mutations to influenza hemagglutinin. *Viruses* 8:155.
- Haddox HK, Dingens AS, Bloom JD (2016) Experimental estimation of the effects of all amino-acid mutations to HIV's envelope protein on viral replication in cell culture. *PLoS Pathogens* 12(12):e1006114.
- Qi H, Wu NC, Du Y, Wu TT, Sun R (2015) High-resolution genetic profile of viral genomes: why it matters. *Current Opinion in Virology* 14:62–70.
- WHO (2010) Recommended viruses for influenza vaccines for use in the 2010–2011 northern hemisphere influenza season. http://www.who.int/influenza/vaccines/virus/recommendations/2010_02_Recommendation.pdf?ua=1.
- WHO (2011) Recommended composition of influenza virus vaccines for use in the 2011–2012 northern hemisphere influenza season. http://www.who.int/influenza/vaccines/2011_02_recommendation.pdf?ua=1.
- Böttcher E, et al. (2006) Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *Journal of Virology* 80:9896–9898.
- Böttcher-Friebertshäuser, E, et al. (2010) Cleavage of influenza virus hemagglutinin by airway proteases TMPRSS2 and HAT differs in subcellular localization and susceptibility to protease inhibitors. *Journal of Virology* 11:5605–5614.
- Marshall N, Priyamvada L, Ende Z, Steel J, Lowen AC (2013) Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. *PLoS Pathogens* 9(6):e1003421.
- Bloom JD (2015) Software for the analysis and visualization of deep mutational scanning data. *BMC Bioinformatics* 16(1):1.
- Bloom JD (2017) Identification of positive selection in genes is greatly improved by using experimentally informed site-specific models. *Biology Direct* 12(1):1.
- Hilton SK, Doud MB, Bloom JD (2017) phydm: Software for phylogenetic analyses informed by deep mutational scanning. *PeerJ* 5:e3657.
- Yang Z, Nielsen R, Goldman N, Pedersen AMK (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155(1):431–449.
- Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53(5):793–808.
- Waterfield M, Scrase G, Skehel J (1981) Disulphide bonds of haemagglutinin of Asian influenza virus. *Nature* 289:422–424.
- Weis W, et al. (1988) Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature* 333:426–431.
- Martin J, et al. (1998) Studies of the binding properties of influenza hemagglutinin receptor-site mutants. *Virology* 241(1):101–111.
- Nobusawa E, Ishihara H, Morishita T, Sato K, Nakajima K (2000) Change in receptor-binding specificity of recent human influenza A viruses (H3N2): a single amino acid change in hemagglutinin altered its recognition of sialyloligosaccharides. *Virology* 278(2):587–596.
- Yang H, et al. (2015) Structure and receptor binding preferences of recombinant human A (H3N2) virus hemagglutinins. *Virology* 477:18–31.
- Stech J, Garn H, Wegmann M, Wagner R, Klunk H (2005) A new approach to an influenza live vaccine: modification of the cleavage site of hemagglutinin. *Nature Medicine* 11(6):683–689.
- Girard G, Gulyaev A, Olsthoorn R (2011) Upstream start codon in segment 4 of North American H2 avian influenza A viruses. *Infect. Genet. Evol.* 11:489–495.
- Wolf Y, Viboud C, Holmes E, Koonin E, Lipman D (2006) Long intervals of stasis punctuated by bursts of positive selection in the seasonal evolution of influenza A virus. *Biology Direct* 1:34.
- Heaton NS, Sachs D, Chen CJ, Hai R, Palese P (2013) Genome-wide mutagenesis of influenza virus reveals unique plasticity of the hemagglutinin and NS1 proteins. *Proceedings of the National Academy of Sciences* 110(50):20248–20253.
- Ekerti DC, et al. (2011) A highly conserved neutralizing epitope on group 2 influenza A viruses. *Science* 333(6044):843–850.
- Friesen R, et al. (2014) A common solution to group 2 influenza virus neutralization. *Proc. Natl. Acad. Sci. USA* 111:445–450.
- Chan N, et al. (2016) Two escape mechanisms of influenza A virus to a broadly neutralizing stalk-binding antibody. *PLoS Pathogens* 12(6):e1005702.
- Yamayoshi S, et al. (2017) Human protective monoclonal antibodies against the HA stem of group 2 HAs derived from an H3N2 virus-infected human. *Journal of Infection*.
- Okuno Y, Isegawa Y, Sasao F, Ueda S (1993) A common neutralizing epitope conserved between the hemagglutinins of influenza A virus H1 and H2 strains. *Journal of Virology* 67(5):2552–2558.
- Doud MB, Lee JM, Bloom JD (2018) Quantifying the effects of single mutations on viral escape from broad and narrow antibodies to an H1 influenza hemagglutinin. *bioRxiv* p. 210468.
- Anderson CS, et al. (2017) Natural and directed antigenic drift of the H1 influenza virus hemagglutinin stalk domain. *Scientific Reports* 7(1):14614.
- Wiley D, Wilson I, Skehel J, , et al. (1981) Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* 289(5796):373–378.
- Popova L, et al. (2012) Immunodominance of antigenic site B over site A of hemagglutinin of recent H3N2 influenza viruses. *PLoS One* 7(7):e41895.
- Wilson I, Skehel J, Wiley D (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* 289:366–373.
- Lee PS, et al. (2014) Receptor mimicry by antibody F045-092 facilitates universal binding to the H3 subtype of influenza virus. *Nat Commun* 5:3614.
- Gamble B, et al. (2004) The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science* 303(5665):1838–1842.
- Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22(12):2577–2637.
- Tien M, Meyer AG, Spielman SJ, Wilke CO (2013) Maximum allowed solvent accessibilities of residues in proteins. *PLoS One* 8:e80635.
- Doud MB, Ashenberg O, Bloom JD (2015) Site-specific amino acid preferences are mostly conserved in two closely related protein homologs. *Mol. Biol. Evol.* 32:2944–2960.
- Haddox HK, Dingens AS, Hilton SK, Overbaugh J, Bloom JD (2017) Mapping mutational effects along the evolutionary landscape of HIV envelope. *bioRxiv* p. 235630.
- Nobusawa E, et al. (1991) Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. *Virology* 182(2):475–485.
- Ha Y, Stevens DJ, Skehel JJ, Wiley DC (2002) H5 avian and H9 swine influenza virus haemagglutinin structures: possible origin of influenza subtypes. *The EMBO Journal* 21(5):865–875.
- Russell R, et al. (2004) H1 and H7 influenza haemagglutinin structures extend a structural classification of haemagglutinin subtypes. *Virology* 325(2):287–296.
- Daniels R, et al. (1985) Fusion mutants of the influenza virus hemagglutinin glycoprotein. *Cell* 40(2):431–439.
- Sun X, Longping VT, Ferguson AD, Whittaker GR (2010) Modifications to the hemagglutinin cleavage site control the virulence of a neurotropic H1N1 influenza virus. *Journal of virology* 84(17):8683–8690.
- Lee HK, et al. (2013) Comparison of mutation patterns in full-genome A/H3N2 influenza sequences obtained directly from clinical samples and the same samples after a single MDCK passage. *PLoS One* 8(11):e79252.
- Wu N, et al. (2017) A structural explanation for the low effectiveness of the seasonal influenza H3N2 vaccine. *PLoS Pathogens* 13(10):e1006682.
- Memoli MJ, et al. (2009) Recent human influenza A/H3N2 virus evolution driven by novel selection factors in addition to antigenic drift. *The Journal of Infectious Diseases* 200(8):1232–1241.

66. Raghwan J, Thompson RN, Koelle K (2017) Selection on non-antigenic gene segments of seasonal influenza A virus and its impact on adaptive evolution. *Virus Evolution* 3(2).
67. Gong LI, Suchard MA, Bloom JD (2013) Stability-mediated epistasis constrains the evolution of an influenza protein. *eLife* 2:e00631.
68. Natarajan C, et al. (2013) Epistasis among adaptive mutations in deer mouse hemoglobin. *Science* 340(6138):1324–1327.
69. Harms MJ, Thornton JW (2014) Historical contingency and its biophysical basis in glucocorticoid receptor evolution. *Nature* 512(7513):203–207.
70. Starr TN, Thornton JW (2016) Epistasis in protein evolution. *Protein Science* 25(7):1204–1218.
71. Starr TN, Picton LK, Thornton JW (2017) Alternative evolutionary histories in the sequence space of an ancient protein. *Nature* 549(7672):409–413.
72. Doud MB, Hensley SE, Bloom JD (2017) Complete mapping of viral escape from neutralizing antibodies. *PLoS Pathogens* 13(3):e1006271.
73. Höftmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG (2000) A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc. Natl. Acad. Sci. USA* 97:6108–6113.
74. Ashenberg O, Padmakumar J, Doud MB, Bloom JD (2017) Deep mutational scanning identifies sites in influenza nucleoprotein that affect viral inhibition by MxA. *PLoS pathogens* 13(3):e1006288.
75. Bloom JD (2014) An experimentally determined evolutionary model dramatically improves phylogenetic fit. *Molecular Biology and Evolution* 31:1956–1978.
76. Dingens AS, Haddox HK, Overbaugh J, Bloom JD (2017) Comprehensive mapping of HIV-1 escape from a broadly neutralizing antibody. *Cell Host & Microbe* 21:777–787.
77. Bao Y, et al. (2008) The Influenza Virus Resource at the National Center for Biotechnology Information. *J. Virol.* 82:596–601.
78. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30(4):772–780.
79. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688–2690.
80. Hadfield J, et al. (2017) Nextstrain: real-time tracking of pathogen evolution. *bioRxiv*.
81. Sagulenko P, Puller V, Neher RA (2018) TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evolution* 4(1):vex042.

DRAFT

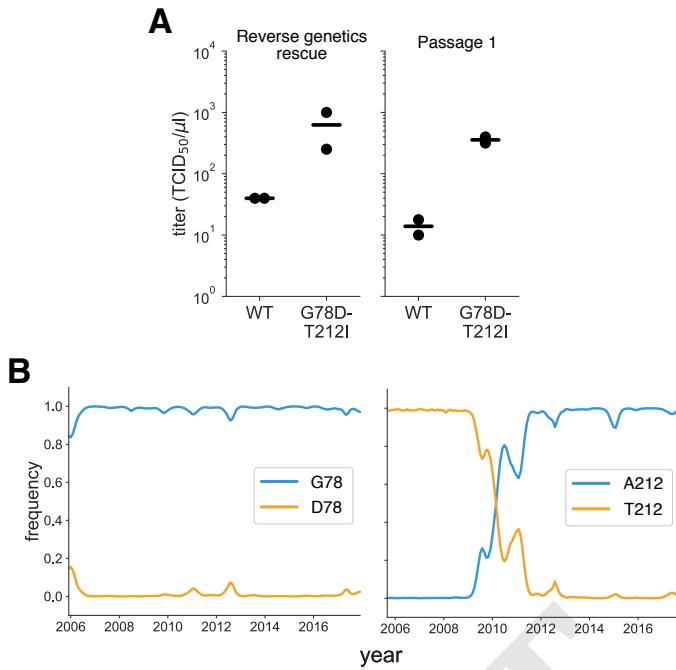


Figure S1. Characterization of the G78D-T212I Perth/2009 HA variant. (A) The G78D-T212I Perth/2009 HA variant supports better viral growth than the wildtype Perth/2009 HA. Viruses were generated in duplicate by reverse genetics with the Perth/2009 NA and WSN internal genes, and passaged once at MOI = 0.01 in MDCK-SIAT1-TMPRSS2 cells. The rescue and passage viral supernatants were collected at 72 hours post-transfection and 44 hours post-infection, respectively, and titrated in MDCK-SIAT1-TMPRSS2 cells. The points mark each duplicate and the bar marks the mean. (B) The D78 variant remained at a low frequency in natural human H3N2 sequences over the past ~10 years. The A212 variant rose to fixation in ~2011, replacing the T212 variant.

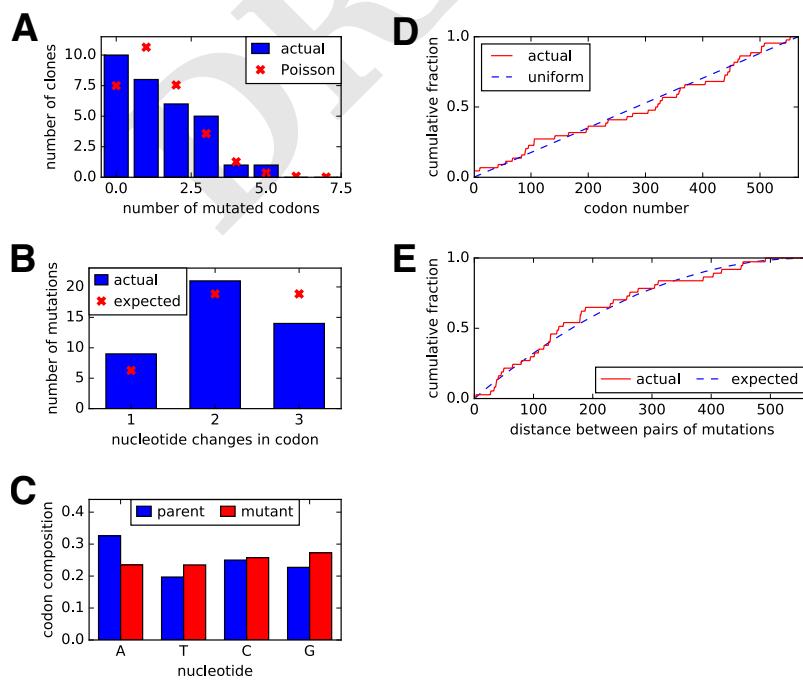


Figure S2. Sanger sequencing of 31 randomly chosen clones from the mutant plasmid libraries. (A) There were an average of ~1.4 codon mutations per clone across the three plasmid mutant libraries. (B) A mixture of one-, two-, and three-nucleotide mutations were present, with slightly fewer triple-nucleotide changes than expected. (C) Nucleotide frequencies were uniform in the codon mutations. (D) The mutations were distributed relatively evenly across the length of the HA coding sequence. (E) We calculated the pairwise distances between mutations for clones carrying more than one mutation. The cumulative distribution of these distances is shown in the red line. The blue line indicates the expected distribution if mutations in multiply mutated genes are randomly dispersed along the sequence.

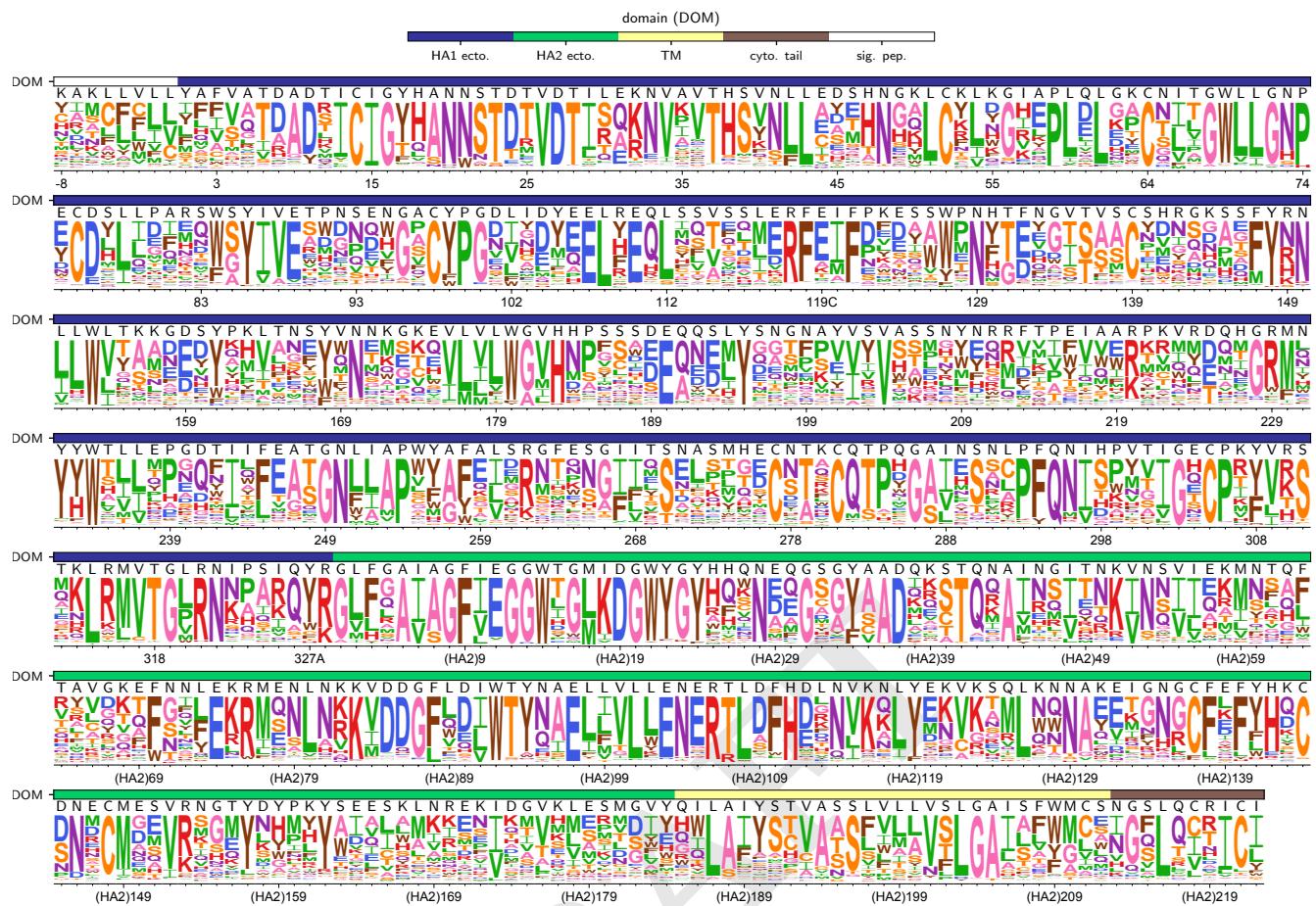


Figure S3. The site-specific amino-acid preferences of the WSN/1933 H1 HA as measured in (20). The amino-acid preferences from (20) after taking the average of the experimental replicates and re-scaling (30) by a stringency parameter of 2.05 (see https://github.com/bloomblab/dms_tools2/blob/master/examples/Doud2016/analysis_notebook.ipynb). The sites are in H3 numbering. The overlays show the same information as in Figure 3 (domain and wildtype amino acid).

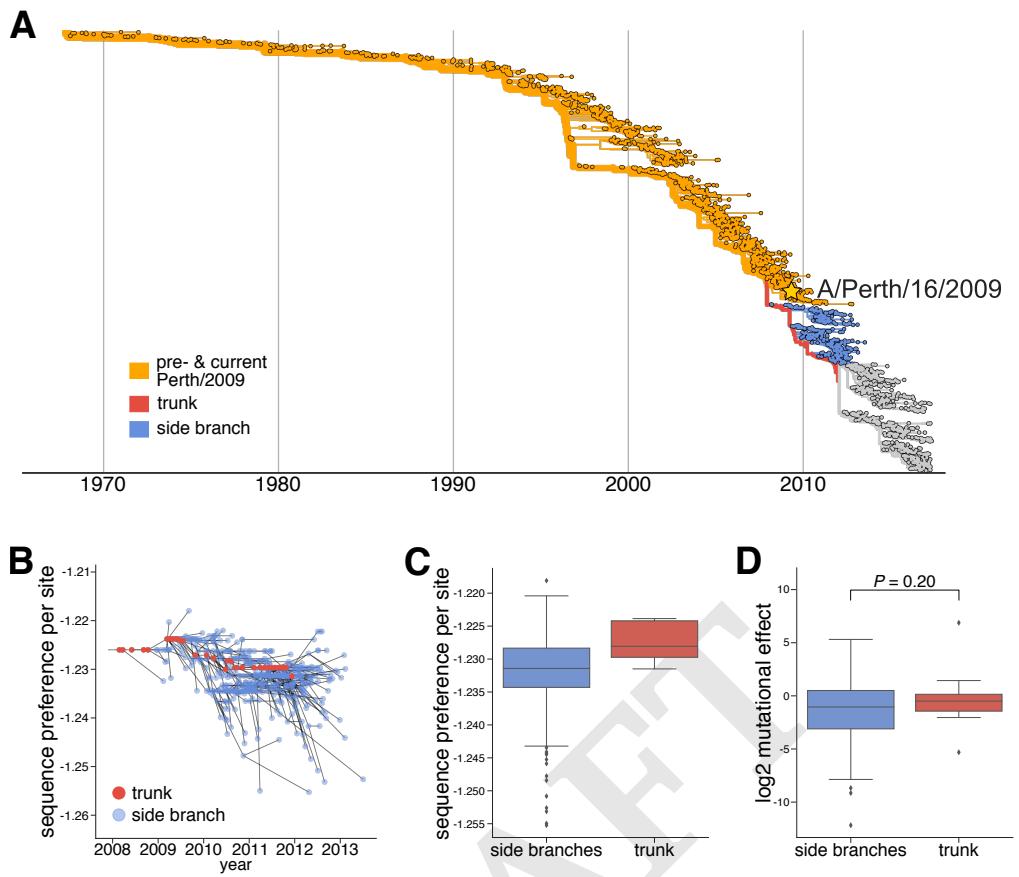


Figure S4. The sequence preference of nodes that occur after the split of Perth/2009 off the trunk. Our experiments show a selective advantage for the trunk even if we limit the analysis to sequences that occur after the split of Perth/2009 off the trunk. (A) Phylogenetic tree showing all nodes that branch from the trunk before or contemporaneously with Perth/2009 (orange), all nodes that occur after the branch to Perth/2009 for which we can resolve the trunk (red) or side branches (blue), and all nodes for which it is not yet clear which sequences will be on the trunk or side branches (gray). In this figure, we restrict the analysis to the nodes in red and blue. (B) Average per-site sequence preference (similar to Figure ??C) for all trunk and side-branch nodes that occur after the branch to Perth/2009. (C) The sequence preferences of nodes on the trunk generally exceed those of nodes on the side branches along the post-Perth/2009 portion of the tree. (D) Post-Perth/2009 trunk mutations on average tend to be more favorable than side-branch mutations. This effect is of similar magnitude to pre-Perth/2009 mutations (see Figure ??) but is not statistically significant because there are relatively few post-Perth/2009 mutations (13 trunk mutations and 324 side-branch mutations).

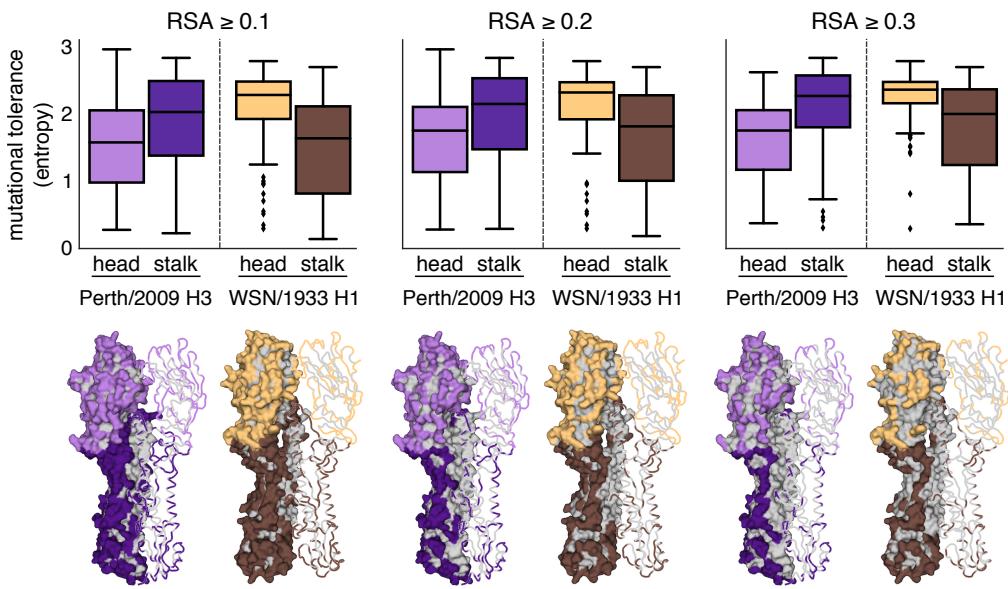


Figure S5. Mutational tolerances of the head and stalk domains at various RSA cutoffs. The mutational tolerances of the head and stalk domains show less disparity for the Perth/2009 H3 HA compared to those for the WSN/1933 H1 HA. We used relative solvent accessibility (RSA) cutoffs of 0.1, 0.2, and 0.3 to define solvent-exposed residues and plotted the mutational tolerances (Shannon entropy of re-scaled preferences) of these residues in the head and stalk domains for the Perth/2009 H3 HA (purple) and the WSN/1933 H1 HA (brown). Residues falling in between the two cysteines at sites 52 and 277 were defined as belonging to the head domain, while all other residues were defined as the stalk domain. The HA structures color the residues that are defined as solvent exposed at a given RSA cutoff. One monomer is shown in surface representation and another monomer shown in ribbon representation. Residues in lighter shades of purple or brown are in the head domain, while residues in darker shades are in the stalk domain. Note that the mutational tolerance values are not comparable between the two HAs.

Dataset S1. Genbank file giving the full sequence of the bidirectional reverse-genetics plasmid pHW-Perth09-HA-G78D-T21I, which encodes the wildtype HA sequence used in this study.

Dataset S2. Excel file providing the primers used to generate the barcoded subamplicons for Perth/2009 HA deep sequencing.

Dataset S3. Excel file giving the amino-acid preferences in sequential 1, 2, ... numbering of the Perth/2009 HA. The unscaled preferences for replicates 1, 2, 3-1, and 3-2 are each in a separate tab of the file. Additional tabs give the across-replicate averaged and re-scaled amino-acid preferences in sequential numbering and in H3 numbering as shown in Figure 3. There are also tabs that give the conversion from sequential to H3 numbering, and the epitope sites in H3 numbering as defined by Wolf et al and used in this study. Each tab can simply be exported to CSV for computational analyses.