Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants

Juhye M. Lee a,b,c,1, John Huddleston a,d,e,1, Michael B. Doud b,b,c, Kathryn A. Hoope r a,e, Nicholas C. Wu f, Trevor Bedford d,g,2, and Jesse D. Bloom a,b,1,2

*Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109; ‡Department of Genome Sciences, University of Washington, Seattle, WA 98195; §Medical Scientist Training Program, University of Washington, Seattle, WA 98195; †Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109; ¶Molecular and Cellular Biology Program, University of Washington, Seattle, WA 98195; ††Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037; and ‡‡Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

Edited by Peter Palese, Icahn School of Medicine at Mount Sinai, New York, NY, and approved July 18, 2018 (received for review April 9, 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 show that mutations that are measured to be more favorable for viral growth are enriched in evolutionarily successful H3N2 viral lineages relative to mutations that are measured to be less favorable for viral growth. Therefore, despite the well-known caveats about cell-culture measurements of viral fitness, such measurements can still be informative for understanding evolution in nature. We also compare our measurements for 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 understand 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 | epistasis | mutational shifts

Seasonal H3N2 influenza virus evolves rapidly, fixing 3 to 4 amino acid mutations per year in its hemagglutinin (HA) surface protein (1, 2). Many of these mutations contribute to the rapid antigenic drift that necessitates frequent updates to the annual influenza vaccine (3). This evolution is further characterized by competition and turnover among groups of strains called clades bearing different complements of mutations (4–8). Clades vary widely in their evolutionary success, with some dying out soon after emergence and others going on to take over the virus population. Several lines of evidence indicate that successful clades have higher fitness than clades that remain at low frequency (4–6, 9). 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 used experimental data to assess the antigenicity of circulating strains (11–16). However, the nonantigenic effects of mutations also play an important role (5, 7, 9, 17). Specifically, due to influenza virus’s high mutation rate (18–20) and lack of intrasegment recombination (21), deleterious mutations become linked to beneficial ones. The resulting accumulation of deleterious mutations can affect nonantigenic properties central to viral fitness (9). However, there are no large-scale quantitative characterizations of how mutations to H3N2 HA affect viral growth.

It is now possible to use deep mutational scanning (22) to measure the functional effects of all single amino acid mutations to viral proteins (10, 23–27). However, the only HA for which such large-scale measurements have previously been made is from the highly laboratory-adapted A/WSN/1933 (H1N1) strain (10, 23, 24). Here, we measure the effects on viral growth in cell culture of all mutations to the HA of a recent human H3N2 strain. We show that these experimental measurements can help discriminate evolutionarily successful mutations from those found in strains that quickly die out. However, the utility of the experiments for understanding natural evolution depends on the similarity between the experimental and natural strains: Measurements made on an H1 HA are less informative for understanding the evolutionary fate of H3 viral strains.

Significance

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 understanding virus evolution in nature.
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 (28, 29). 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 (see SI Appendix, Fig. 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 (see SI Appendix, Fig. 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 (10) (Fig. 1A). These mutant viruses derived all their non-HA genes from the laboratory-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 (Madin–Darby canine kidney cells overexpressing 2,6-sialyltransferase) (30) that we engineered to constitutively express TMPRSS2 (Transmembrane Protease, Serine 2), which cleaves the HA precursor to activate it for membrane fusion (31, 32).

After generating the mutant virus libraries, we passaged them at low multiplicity of infection (MOI) in cell culture to create a genotype–phenotype link and select for functional HA variants (Fig. 1A). All experiments were completed in full biological triplicate (Fig. 1B). We also passaged and deep sequenced library 3 in duplicate (library 3-1 and 3-2) to gauge experimental noise. To further ensure the validity of our measurements, we used the unmutated HA gene to generate and passage viruses carrying wild-type 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% to 45% of their initial frequencies after correcting for error rates estimated by sequencing the wild-type controls (Fig. 1C). The incomplete purging of stop codons is likely because genetic complementation due to co-infection (33, 34) enabled the persistence of some virions with nonfunctional HAs. We also observed selection against many nonsynonymous mutations (Fig. 1C), with their frequencies falling to 30% to 40% of their initial values after error correction.

We next quantified the reproducibility of our deep mutational scanning 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 as described in ref. 39. Because there are 566 residues in HA, there are 566 × 19 = 10,754 distinct measurements [the 20 preferences at each site sum to 1 (39)]. The correlations of the amino acid preferences between pairs of replicates are shown in Fig. 1D. 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 (Fig. 1C), 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 with standard substitution models (35, 40). Table 1 shows that an 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. A substitution model in which the amino acid preferences...
Table 1. Substitution models informed by the experiments describe HA’s evolution better than traditional models

<table>
<thead>
<tr>
<th>Model</th>
<th>ΔAIC</th>
<th>LnL</th>
<th>Stringency</th>
<th>ω</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExpCM</td>
<td>0.0</td>
<td>−8,441</td>
<td>2.47</td>
<td>0.91</td>
</tr>
<tr>
<td>GY94 M5</td>
<td>2.094</td>
<td>−9,482</td>
<td>—</td>
<td>0.36(0.30, 0.84)</td>
</tr>
<tr>
<td>ExpCM, site average</td>
<td>2.501</td>
<td>−9,692</td>
<td>0.67</td>
<td>0.32</td>
</tr>
<tr>
<td>GY94 M0</td>
<td>2.536</td>
<td>−9,704</td>
<td>—</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Shown are the maximum likelihood phylogenetic fits to an alignment of human H3N2 HA’s using ExpCM (35), ExpCM in which the experimental measurements are averaged across sites (site average), and M0 and M5 versions of the Goldman–Yang (GY94) model (36). Models are compared by Akaike information criterion (AIC) (37) 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.

were averaged across all sites (ExpCM, site average) performs no better than conventional substitution models, demonstrating that the reason that our measurements are informative is because they capture site-specific constraints. The relative rate of nonsynonymous to synonymous substitutions (dN/dS or ω) is ≪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 1 (Table 1), indicating that most purifying selection against nonsynonymous substitutions is accounted for in the deep mutational scanning. The ExpCM stringency parameter (35) 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 the amino acid preferences rescaled (35, 40) by this stringency parameter. These rescaled preferences are shown in Fig. 2.

Examination of Fig. 2 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 and 76, 97 and 139, 281 and 305, 14 and 137-HA2, 144-HA2 and 148-HA2) (43) strongly prefer 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 (44–47). A positively charged epitope site is located at position 135-HA2 (48). The top overlay bar indicates whether or not a site is in the set of epitope residues delineated in ref. 38. 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 wild-type amino acid at that site.

Fig. 2. 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 rescaling (35) 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 ref. 38. 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 wild-type amino acid at that site.
charged amino acid at site 329 is important for cleavage of the HA0 precursor into the mature form (48), and this site strongly prefers arginine.

However, there are also some differences between the amino acid preferences measured in our experiments and amino acid frequencies in natural H3 HA sequences (see SI Appendix, Fig. S3). Most surprisingly, the start codon does not show a particularly strong preference for methionine (Fig. 2). We validated that a virus carrying a mutation at this site from methionine to lysine does in fact reach appreciable titers (see SI Appendix, Fig. S4), perhaps because of alternative translation-initiation at a downstream or upstream start site as has been described for other HAs (49). Our measurements also suggest mutational tolerance at some other sites that are relatively conserved among natural HAs, such as the N-linked glycosylation motifs near the beginning of HA1 and the transmembrane domain (Fig. 2). We validated that viruses with mutations to the glycosylation motifs at sites 22 or 38, or a site in the transmembrane domain, do in fact grow to high titers (SI Appendix, Figs. S5 and S6, respectively). The disparity between the relative conservation of these sites in nature and their mutational tolerance in our study could be because cell culture does not fully capture the constraints on HA function in nature or could be because these sites are not under strong immune pressure and so mutations at them are not positively selected in nature.

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 viral growth. 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 rescaled preferences. In prior mutational studies of H1 HAs, the stalk domain was found to be substantially less mutationally tolerant than the globular head (10, 23, 24, 50).

We performed a similar analysis using our new data for the Perth/2009 H3 HA. Surprisingly, the head domain of the H3 HA is not more mutationally tolerant than the stalk domain (Fig. 3). Specifically, whereas solvent-exposed residues in the head domain are substantially more mutationally tolerant than those in the stalk domain for the WSN/1933 H1 HA, the trend is actually reversed for the Perth/2009 H3 HA (Fig. 3B). This difference between the relative mutational tolerances of the H1 and H3 HAs is robust to the cutoff used to define surface residues (see SI Appendix, Fig. S7). For instance, for the H3 HA, the short helix A in the stalk domain is as mutationally tolerant as many surface-exposed residues in the head domain—something that is not the case for the H1 HA. Helix A forms part of the epitope of many broadly neutralizing antistalk antibodies (51–53).

We also see high mutational tolerance in many of the known antigenic regions of H3 HA (54). For instance, antigenic region B is an immunodominant area, and many recent major antigenic drift mutations have occurred in this region (14, 15, 55). 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 (Fig. 3A). Antigenic region C is also notably mutationally tolerant.

Many residues inside HA’s receptor binding pocket are known to be highly functionally constrained (45, 56), and our data indicate that these sites are relatively mutationally intolerant in both H3 and H1 HAs (Fig. 3A). 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 (14, 54).

Our Measurements Can Help Distinguish Between Mutations That Reach Low and High Frequencies in Nature. Mutations occurring in the H3N2 virus population experience widely varying evolutionary fates (Fig. 4). Some mutations appear, spread, and fix in the population, while others briefly circulate before disappearing. We take the maximum frequency reached by a mutation as a coarse indicator of its effect on fitness, since favorable mutations generally reach higher frequencies than unfavorable ones (57). Here, we follow the population genetic definition of “mutation” and track the outcome of each individual mutation event; for example, although R142G occurs multiple times on the
phylogeny, we track each of these mutations occurring on different backgrounds separately. As such, each mutation is shown as a separate circle on a separate branch in Fig. 4. However, because multiple mutations on the same phylogeny branch cannot be entangled, when multiple mutations occurred on a single branch, we assigned a single mutational effect based on the sum of effects of each mutation.

After annotating mutations and their frequencies on the phylogeny in this way (Fig. 4), it is visually obvious that there are relatively few circulating mutations that we measure to be strongly deleterious—and that such deleterious mutations rarely reach high frequency when they do occur.

We next sought to quantify the correlation between a mutation’s experimentally measured effect and the maximum frequency it attained during natural evolution. To calculate a given mutation’s effect, we simply took the logarithm of the ratio of the preferences for the mutant and wild-type amino acids at that site. To minimize effects related to the genetic background of the strain used in the experiment, we excluded mutations closely related to the experimental strain itself and partitioned the remaining mutations into 1,022 mutations predating and 299 mutations postdating the Perth/2009 strain (see SI Appendix, Fig. S8). We additionally excluded mutations from the post-Perth partition that were sampled in 2014 or after, since these mutations have not had enough time for their evolutionary fates to be fully resolved. We used these pre-Perth and post-Perth partitions to test the utility of our measurements for both post hoc and prospective analyses, respectively. We quantified the relationship between mutational effects and maximum mutation frequencies in the H3N2 phylogeny via Spearman rank correlation (Fig. 5A). In both pre-Perth and post-Perth time periods, we found a modest but statistically significant relationship between mutational effect and maximum mutation frequency (pre-Perth \( \rho = 0.17 \), post-Perth \( \rho = 0.15 \)). The similar effect sizes for both the pre- and post-Perth partitions show that our experimental measurements can help explain the evolutionary fates of mutations in strains that postdate the experimentally studied strain as well as to retrospectively analyze mutations that precede the experimental strain.

Many of the HAs in sequence databases are from viral isolates that were passaged in cell culture or eggs, which can cause laboratory-adaptation mutations that confound evolutionary analyses. To check that our results were robust to such laboratory-adaptation mutations, we repeated our analysis using only HA sequences derived from viruses that had not been passaged in the laboratory. Because sequencing of unpassaged primary isolates has only recently become commonplace, we could only perform this analysis for the post-Perth partition of the phylogenetic tree. Fig. 5A shows that the correlation between our measured mutational effects and the maximum frequency was even stronger for mutations from unpassaged viral isolates (\( \rho = 0.24 \)). The trends in Fig. S4 are most strongly driven by the behavior of substantially deleterious mutations. We investigated this further by partitioning mutations into those that reach low, medium, and high frequencies and those that fix in the population (Fig. 5B). The mutations that reach higher frequencies have a more favorable mean effect. Mutations measured to be substantially deleterious almost never reach high frequency. Overall, these results demonstrate that measurements of how mutations affect viral growth in cell culture are informative for understanding the fates of these mutations in nature: In particular, if a mutation is measurably deleterious to viral growth, that mutation is unlikely to prosper in nature.

Measurements Made on an H1 HA Are Less Informative for Understanding the Evolution of H3 Influenza. To determine how broadly experimental measurements can be generalized across HAs, we repeated the foregoing analysis of H3N2 mutation frequencies but using mutational effects measured in our prior deep mutational scanning of the WSN/1933 H1 HA (10) (see SI Appendix, Fig. S9) as which is highly diverged from the Perth/2009 H3 HA (the two HAs only have 42% protein sequence identity). Fig. 6 shows that the correlations between the H1 experimental measurements and the maximum frequency that mutations reach during H3N2 viral evolution are consistently weaker than those using H3 experimental measurements (compare Fig. 6 to Fig. 5A). Therefore, the utility of an experiment for understanding natural evolution degrades as the experimental sequence becomes more diverged from the natural sequences that are being studied.

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 is less useful for understanding the evolution of H3N2 influenza viruses 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 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 (59) and variants of HIV envelope (Env) with 86% amino acid identity (27). However, the H1 and H3 HAs are far more diverged, with only 42% amino acid identity (Fig. 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. Fig. 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 refs. 27 and 59. Fig. 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 (27). However, the typical shift between H3 and H1 is still smaller than that observed when comparing HA to the nonhomologous 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 nonhomologous 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 (Fig. 8A). However, at the domain level, sites in HA’s stalk tend to have smaller shifts than sites in HA’s globular head (Fig. 8B). The HA stalk domain is also more conserved in sequence (60), 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 (Fig. 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 (61, 62). 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 (61, 62). Previous work has defined clades of structurally related HA subtypes (61, 62). One such clade includes H1, H2, H5, and H6, whereas another clade includes H3, H4, and H14 HAs (Fig. 7A). Sites that are conserved at different amino acid identities in these two clades tend to have exceptionally large shifts in amino acid preferences (Fig. 8B). The clade containing H1 has an upward translation of the globular

---

Fig. 5. Experimental measurements are informative about the evolutionary fate of viral mutations. (A) Correlation between the effects of mutations as measured in our deep mutational scanning of the Perth/2009 HA and the maximum frequency reached by these mutations in nature. The plots show Spearman ρ and an empirical P value representing the proportion of 10,000 permutations of the experimental measurements for which the permuted ρ was greater than or equal to the observed ρ. (B) The distribution of mutational effects partitioned by maximum mutation frequency. The vertical black line shows the mean mutation effect for each category. The analysis is performed separately for pre-Perth/2009, post-Perth/2009, and unpassaged isolates from the post-Perth/2009 partitions of the tree (see SI Appendix, Fig. S8).

Fig. 6. Experimental measurements on an H1 HA are less informative about the evolutionary fate of H3N2 mutations. This figure repeats the analysis of the H3N2 mutation frequencies in Fig. 5A but uses the deep mutational scanning data for an H1 HA as measured in ref. 10. SI Appendix, Fig. S10 shows the histograms comparable to those in Fig. 5B. The empirical P value represents the result of 1,000 permutations.
There are large shifts in the effects of mutations between H1 and H3 HAs. \(\text{A}\) indicates the Pearson correlation coefficient. \(\text{B}\) of 10

Lee et al.

which experiments can select for laboratory artifacts or fail to

a vast amount of work in virology has chronicled the ways in

in cell culture do not represent true fitness in nature. Indeed,

gests that deep mutational scanning may be able to provide

“diverse genotypic and phenotypic data” that more realistically

represented the effects of specific mutations. Our work sug-

“predictive viral fitness models that accounted for both anti-

evolutionary forecasting. In their landmark paper introducing

tionarily successful mutations suggests that they might inform

growth. The fact that our experiments can help identify evolu-

nations in nature. Specifically, mutations measured to be more

beneficial for viral growth tend to reach higher frequencies in

nature than mutations measured to be more deleterious for viral

mutations to the Perth/2009 H3 HA on viral growth in cell

We have measured the effects of all possible single amino acid

amino acid preferences can be rationalized in terms of changes

in HA structure.

Discussion

head relative to the clade containing H3. This structural shift has been attributed largely to the interaction between sites 107

and 75(HA2) (61, 62). 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 hydro-

gen bond between Glu-107 and Lys-75(HA2) (Fig. 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 that 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.

We measured the effects of all single amino acid mutations to the Perth/2009 H3 HA on viral growth in cell culture and demonstrated that these measurements have some value for understanding the evolutionary fate of these mutations in nature. Specifically, mutations measured to be more beneficial for viral growth tend to reach higher frequencies in nature than mutations measured to be more deleterious for viral growth. The fact that our experiments can help identify evolutionarily successful mutations suggests that they might inform evolutionary forecasting. In their landmark paper introducing predictive viral fitness models that accounted for both antigenic and nonantigenic mutations, Luksza and Lässig (9) 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 may be able to 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 ways in which experiments can select for laboratory artifacts or fail to capture important pressures that are relevant in nature (63–66). As an example, although we identified G78D as favorable for viral growth in cell culture, this mutation never fixes in nature. Mutations in viral genes other than HA are also important in determining strain success (67, 68). 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. However, before 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 nonantigenic effects of mutations (11–13) or assumed that all nonepitope mutations had equivalent deleterious effects (9). The strength of our measurements are not that they perfectly capture fitness in nature but that they are systematic and quantitative—and so represent an improvement over no information at all. We suspect that performing similar experiments using more realistic and complex selections (e.g., ferrets or primary human airway cultures) might further improve their utility and possibly their generalizability to more divergent strains.

We measured the effects of all single amino acid mutations to a specific HA and then generalized these measurements to other H3N2 HAs from a 50-y timespan. These generalizations will only be valid to the extent that the effects of mutations are conserved during HA’s evolution. Extensive prior work has shown that epistasis can shift the effects of mutations as proteins evolve (69–75). Our work suggests that measurements on a HA from a single human H3N2 viral strain can be usefully generalized to at least some extent across the entire evolutionary history of human H3N2 HA. 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 (27, 59). 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.
The mutant virus libraries were for SI Appendix per microliter for the three library replicates and the Supplementary Text Sites are in H3 numbering, with the signal peptide in negative numbers, HA1 in plain numbers, and HA2 denoted with “(HA2).” When growing influenza virus in cell culture, trypsin is normally added to cleave HA into its mature form.

Creation of MDCK–SIAT1–TMPRSS2 Cells. When growing influenza virus in cell culture, trypsin is normally added to cleave HA into its mature form. To obviate the need for trypsin, we engineered MDCK–SIAT1 cells and MDCK–SIAT1–CMV–PB1 (78) cells to constitutively express the TMPRSS2 protease, which cleaves and activates HA in the human airways (31, 32).

The human TMPRSS2 cDNA ORF was ordered from OriGene (NM_005656) and cloned into a pHAGE2 lentiviral vector under an EF1α-int promoter followed by an IRES driving expression of mCherry to create plasmid pHAGE2-EF1α-int–TMPRSS2–IRES–mCherry-W. We used the lentiviral vector to transduce MDCK–SIAT1 or MDCK–SIAT1–CMV–PB1 cells and sorted an intermediate mCherry-positive population by flow cytometry. We refer to the sorted bulk population as MDCK–SIAT1–TMPRSS2 cells or MDCK–SIAT1–CMV–PB1–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. HA and NA genes for the Perth/2009 viral strain were cloned from recombinant virus obtained from BEI Resources (NR-41803) into the pHW2000 (79) 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 coculture of 293T and MDCK–SIAT1–TMPRSS2 in D10 media (DMEM, supplemented with 10% heat-inactivated FBS, or fetal bovine serum, 2 mM L-glutamine, 100 U of penicillin per milliliter, and 100 μg of streptomycin per milliliter) with equal amounts of pHW-Perth09-HA, pHW-Perth09-NA, the pHW18* series of plasmids (TAG2–EF1α–T1–T3, HDM–Nan95–PA, HDM–Nan95–PB1, HDM–Nan95–PB2, and HDM–Aichi68–PB2), and pHAGE2–EF1α–T1–T3, HDM–Nan95–PA, HDM–Nan95–PB1, HDM–Nan95–PB2, and pHAGE2–EF1α–T1–T3, HDM–Nan95–PA, HDM–Nan95–PB1, HDM–Nan95–PB2, and HDM–Aichi68–PB2, and pHAGE2–EF1α–T1–T3, HDM–Nan95–PA, HDM–Nan95–PB1, HDM–Nan95–PB2, and HDM–Aichi68–PB2, and HDM–Nan95–PA, HDM–Nan95–PB1, HDM–Nan95–PB2, and HDM–Aichi68–PB2, and HDM–Nan95–PA, HDM–Nan95–PB1, HDM–Nan95–PB2, and HDM–Aichi68–PB2 to create pHW-Perth09-HA–G78D-T212I and pHW-Perth09-NA–G78D-T212I and then collected the viral supernatant at 72 h posttransfection. This viral supernatant was 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 enhanced viral growth as shown in SI Appendix, Fig. S1. The HA with these two mutations was cloned into pHW2000 (79) and pICR2 (80) to create pHW-Perth09-HA–G78D-T212I and pICR2-Perth09-NA-G78D-T212I. For all subsequent experiments, we used viruses with the HA containing these two mutations to improve titers and viral genetic stability, and this is the HA that we refer to as Perth/2009. We used all non-HA genes (including NA) from WSN/1933 to help increase titers and reduce biosafety concerns.

The codon-mutant libraries were generated using the approach in ref. 81 with the modifications in ref. 82. See SI Appendix, Supplementary Text for full details.

Generation and Passaging of Mutant Viruses. The mutant virus libraries were generated using the helper-virus approach described in ref. 10 with several modifications, most notably the cell line used. Briefly, we transfected 5 × 10⁸ 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–PB2) and 1,250 ng of one of the three pICR2-mutant-HA libraries (83) (the wild-type control) using Lipofectamine 3000 (ThermoFisher L3000008). We allowed the transfected cells to adhere in six-well plates and 4 h later changed the media to D10 media. Eighteen hours after transfection, we infected the cells with the WSN/1933 HA-deficient helper virus (10) by preparing an inoculum of 500 TCID₅₀ per microliter 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 3 h, we changed the media to fresh IGM. At 24 h after helper-virus infection, we harvested the viral supernatants for each replicate, froze aliquots at −80°C, and titrated them in MDCK–SIAT1–TMPRSS2 cells. There is no selectable marker for the wild-type control, respectively.

We passaged 9 × 10⁸ TCID₅₀ of the transfaction supernatants at an MOI of 0.0035 TCID₅₀ per cell. To do this, we plated 4.6 × 10⁹ MDCK–SIAT1–TMPRSS2 cells per dish in fifteen 15-cm dishes in D10 media and allowed the cells to grow for 24 h, at which time they were at ~1.7 × 10⁸ cells per dish. We replaced the media in each dish with 25 mL of an inoculum of 2.5 TCID₅₀ of virus per microliter in IGM. Three hours postinfection, we replaced the inoculum with fresh IGM for replicates 1, 2, and 3. We did not perform a media change for replicate 3-1. As can be seen in Fig. 1D, the media change does not appear to have a substantial effect, as replicate 3-1 looks comparable to the other replicates. We collected viral supernatant for sequencing 48 h postinfection.

**Materials and Methods**

**Data and Computer Code.** Deep sequencing data are available from the Sequence Read Archive under BioSample accession nos. SAMN08102609 and SAMN08102610. Computer code used to analyze the data are at [https://github.com/jbloomlab/Pernoth2009-DMS-MSuscript](https://github.com/jbloomlab/Pernoth2009-DMS-MSuscript).

**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 to cleave HA into its mature form. To obviate the need for trypsin, we engineered MDCK–SIAT1 cells and MDCK–SIAT1–CMV–PB1 (78) cells to constitutively express the TMPRSS2 protease, which cleaves and activates HA in the human airways (31, 32).
Barcoded Subamplicon Sequencing. To extract viral RNA from the three replicate HA virus libraries and the wild-type HA virus, we clarified the viral supernatant by centrifuging at 2,000 × g for 5 min, then ultracentrifuged 24 h at a clarified 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 0.1% mercaptoethanol, pipetting 30 times, transferring the liquid to a microcentrifuge tube, adding 600 μL 70% ethanol, and proceeding according to the manufacturer’s instructions. The HA gene was reverse-transcribed with AccuScript Reverse Transcriptase (Agilent 2008200) using primers P09-HA-For (5′-AGAGAACGCGGCTGAATCTTTAATC-3′) and P09-HA-Rev (5′-AGATGAAAACCGGTGTTTTTAACTAATAAC-3′). The barcoded-subamplicon library prep and deep sequencing were then performed as in ref. 10. See SI Appendix, Supplementary Text for full details.

Analysis of Deep Sequencing Data. We used the dms.tools2 software package (39) to analyze the deep sequencing data. The amino acid preferences are provided in Dataset S3. Computer code and detailed plots about read depth and other quality control metrics are at https://github.com/bloomlab/dms_tools2 version 2.2.5 to analyze the deep sequencing data. The amino acid preferences are provided in Dataset S3. Computer code and detailed plots about read depth and other quality control metrics are at https://github.com/bloomlab/Perth2009-DMS-Manuscript/blob/master/analysis_code/analysis_notebook.ipynb.

Phylogenetic Model Comparison and Stringency Parameter. Phylogenetic model comparisons and fitting of a stringency parameter were performed using phdms as described in ref. 35. See SI Appendix, Supplementary Text for full details.

Shannon Entropy and Relative Solvent Accessibility. We calculated Shannon entropy, $h_i$, for site $i$ as $h_i = -\sum_{a} p_{ia} \log_2 (p_{ia})$, where $p_{ia}$ is the preference for amino acid $a$ at site $i$. We quantified the absolute solvent accessibility of each site of the H3 HA [PDB ID code 4OSN (41)] or the H1 HA [PDB ID code 1RUX (42)] structure using DSSP (Define Secondary Structure of Proteins) (83). We then normalized to a relative solvent accessibility using the absolute accessibilities in ref. 84.

Quantification of Mutational Effects. The effect of mutating site $r$ from amino acid $a_1$ to $a_2$ was quantified as

$$\log_2 \frac{p_{ia_2}}{p_{ia_1}}$$

where $p_{ia_1}$ and $p_{ia_2}$ are the rescaled preferences for amino acids $a_1$ or $a_2$ at site $r$ as shown in Fig. 2. The WSN/1933 H1 HA amino acid preferences are the replicate-average values reported in ref. 10, rescaled by a stringency parameter of 2.05 (see https://github.com/bloomlab/dms_tools2/blob/master/examples/Doud2016/analysis_notebook.ipynb).

H3N2 Phylogenetic Tree and Maximum Mutation Frequencies. The phylogenetic tree was generated using Nextstrain’s augur pipeline (85), and ancestral state reconstruction and branch length timing were performed with TreeTime (86). Frequency trajectories of mutations were estimated following Nextstrain’s augur pipeline as first implemented in Nextflu (87). See SI Appendix, Supplementary Text for full details.

Analysis of Mutational Shifts. We compared the preferences for the Perth/2009 H3 and WSN/1933 H1 HA viruses using the approach in ref. 27. See SI Appendix, Supplementary Text for full details.

Validation of Individual Point Mutants. To validate the viral growth of Perth/2009 HA point mutants Mi-16K, CS2A, CS2C, T42F, T40V, S287A, and C199(HA2K), we examined the supernatant titers of each of these variants after reverse-genetics generation in the context of PB1flank-GFP viruses (78, 88). See SI Appendix, Supplementary Text for full details.

ACKNOWLEDGMENTS. We thank Sarah Hilton, Hugh Daddox, and Sidney Bell for helpful discussions about data analysis and Richard Neher for sharing analysis code and providing helpful comments on the manuscript. We thank the Reed Hutch Genomics Core for performing the illumina deep sequencing. This work was supported by NIH National Institute of Allergy and Infectious Diseases (NIAID) Grants R01 AI127893 (to J.D.B. and T.B.) and U19 AI117891 (to T.B.). J.M.L. was supported in part by the Center for Inference and Dynamics of Infectious Diseases (CIDID), which is funded by NIH National Institute of General Medical Sciences (NIGMS) Grant U54GM111274. The research of J.D.B. 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. T.B. is a Pew Biomedical Scholar and is supported by NIH Grant R35 GM119774.