HHS Public Access

Author manuscript

Cell Host Microbe. Author manuscript; available in PMC 2023 October 12.

Published in final edited form as:

Cell Host Microbe. 2022 October 12; 30(10): 1363–1369.e4. doi:10.1016/j.chom.2022.09.003.

The evolutionary potential of Influenza A virus hemagglutinin is highly constrained by epistatic interactions with neuraminidase

Tongyu Liu[1], **Yiquan Wang**[2], **Timothy J C Tan**[3], **Nicholas C Wu**[2,3,4,5,*], **Christopher B Brooke**[1,4,*,^]

[1]Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

[2]Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, Urbana, Illinois, USA

[3]Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

[4]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

[5]Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

Summary

Antigenic evolution of the influenza A virus (IAV) hemagglutinin (HA) gene limits efforts to effectively control the spread of the virus in the population. Efforts to understand the mechanisms governing HA antigenic evolution typically examine the HA gene in isolation. This can ignore the importance of balancing HA receptor-binding activities with the receptor-destroying activities of the viral neuraminidase (NA) to maintain viral fitness. We hypothesize that the need to maintain functional balance with NA significantly constrains the evolutionary potential of the HA. We use deep mutational scanning and show that variation in NA activity significantly reshapes the HA fitness landscape by modulating the overall mutational robustness of HA. Consistent with this, we observe that different NA backgrounds support the emergence of distinct repertoires of HA escape variants under neutralizing antibody pressure. Our results reveal a critical role for intersegment epistasis in influencing the evolutionary potential of the HA gene.

Graphical Abstract

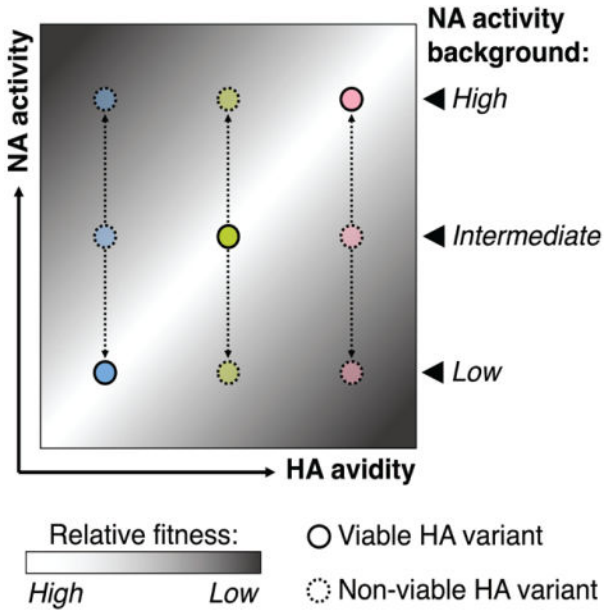
[*]Corresponding authors (nicwu@illinois.edu, cbrooke@illinois.edu).
[^]Lead contact

Author contributions

Conceptualization, C.B.B. and T.L.; Software, Y.W.; Formal Analysis, T.L. and Y.W.; Investigation, T.L. and T.T; Writing – Original Draft, T.L. and C.B.B.; Writing – Review & Editing, C.B.B., N.C.W, Y.W. and T.T.; Supervision, C.B.B. and N.C.W.; Project Administration, C.B.B.; Funding Acquisition, C.B.B. and N.C.W.

Declaration of interests

The authors declare no competing interests.



eTOC Blurb

Antigenic drift of influenza hemagglutinin (HA) poses great challenge to vaccine effectiveness. Predicting the antigenic drift typically examinates HA in isolation. Liu et al. showed the functional interaction between HA and its functional counterpart neuraminidase reshaped the fitness landscape of HA and leaded to distinct antigenic escape profile of HA.

Introduction

Seasonal influenza A viruses (IAV) continue to pose an enormous public health burden (Paget et al., 2019) despite the availability of licensed vaccines. The viral hemagglutinin (HA) glycoprotein mediates cell binding and entry and is the primary target of protective neutralizing antibodies. The persistence of seasonal IAV lineages in the human population depends upon the continual accumulation of antigenically significant substitutions that facilitate evasion of humoral immunity elicited by prior infection or vaccination (Carrat and Flahault, 2007; Petrova and Russell, 2018; Yewdell, 2011).

The antigenic evolution of influenza viruses at the epidemiological scale appears highly constrained, as new antigenic variants typically only emerge every few years despite the high mutation rate of the virus (Brooke, 2017; Pauly et al., 2017; Petrova and Russell, 2018). Further, only a tiny fraction of potential escape variants generally emerge under immune selection, likely because many substitutions capable of reducing antibody binding avidity have deleterious pleotropic effects on HA functions such as receptor binding that offset their fitness benefits (Doud et al., 2017; Koel et al., 2013; Kosik et al., 2018). Defining the specific constraints that govern the emergence of antigenic variants is critical both for the design of next-generation vaccines that elicit more escape-resistant immune responses and for improving our ability to predict future evolutionary trends.

IAV encodes two primary glycoproteins, HA and neuraminidase (NA), that are irregularly distributed across the viral envelope (Harris et al., 2006; Vahey and Fletcher, 2019, 2018; Wasilewski et al., 2012). HA and NA perform opposing functions during infection: HA binds to sialic acid linkages to facilitate cell entry while NA typically cleaves sialic acid linkages to facilitate virion release (Kosik and Yewdell, 2019). Balancing these opposing functions is critical for maintaining viral fitness (Brooke et al., 2014; de Vries et al., 2020; Gaymard et al., 2016; Kosik and Yewdell, 2019; Mitnaul et al., 2000; Wagner et al., 2002; R. Xu et al., 2012). While the importance of HA/NA functional balance for viral fitness is well-established, the implications for the evolutionary potential of HA have not been directly examined.

The NA proteins encoded by different human-origin IAVs exhibit a wide range of relative activity levels (Powell and Pekosz, 2020; R. Xu et al., 2012). We hypothesized that phenotypic variation in NA would alter patterns of mutational tolerance of the HA protein. To test this, we generated two recombinant A/Puerto Rico/8/1934 (rPR8) viruses identical to wild type except for single amino acid substitutions in the NA protein previously reported to reduce NA activity: NA:K253R and NA:H274Y (Bloom et al., 2010; Hensley et al., 2011; Ives et al., 2002). As expected, NA:K253R reduced virion-associated NA activity ~75% compared with WT (Fig S1A), well within the range of natural NA phenotypic variation (Powell and Pekosz, 2020). It also reduced NA protein trafficking to the plasma membrane in infected Madin-Darby canine kidney (MDCK) cells (Fig S1B, C), and decreased the NA protein content of purified virions relative to HA (Fig S1D). We failed to reproduce the previously reported adverse effects of NA:H274Y on NA activity, potentially due to the slightly different virus background used here (Fig S1). As our assays may fail to capture subtle but biologically significant effects of NA:H274Y on NA function, we included it along with NA:K253R in subsequent experiments.

We quantified the effects of NA:H274Y and NA:K253R on the mutational tolerance of the HA1 subunit of HA using deep mutational scanning (DMS) (Doud and Bloom, 2016; Fowler and Fields, 2014; Lee et al., 2018; Wu et al., 2014). We generated a reverse genetics plasmid library in which each codon in the HA1 domain was hyper-mutagenized to ensure sufficient representation of all possible amino acid substitutions as described previously (Doud and Bloom, 2016; Wang et al., 2021; Wu et al., 2014). For each NA genotype (WT, NA:H274Y, and NA:K253R), we rescued three independent recombinant virus populations encoding the mutagenized HA1 domain (HA1dms) and a WT PR8 backbone using reverse genetics. We passaged each population once in MDCK cells at a starting multiplicity of infection (MOI) of 0.05 TCID50/cell to minimize cellular co-infection and thus maintain genotype-phenotype linkages. We then performed barcoded sub-amplicon deep sequencing on each post-passage virus population, along with the mutagenized HA plasmid library used to generate the viruses.

For every possible amino acid substitution in HA1, we calculated an enrichment factor by dividing its post-passage frequency by its frequency in the input plasmid library. We then calculated normalized relative fitness scores for each missense substitution in HA1 by normalizing based on the enrichment factor distributions of nonsense and synonymous substitutions. In brief, we assumed all nonsense substitutions would be lethal, regardless

of genetic background and set their mean relative fitness values at 0 for each experimental replicate. Similarly, we assumed all synonymous substitutions would be neutral and set their mean relative fitness values at 1 for each experimental replicate. The pairwise correlation coefficients of fitness scores for specific substitutions between experimental replicates ranged from 0.665 to 0.812, indicating that our fitness effect measurements were highly reproducible (Fig 1A, Table S1).

A large majority of missense substitutions in the rPR8-WT background had normalized fitness scores of <1, with the overall peak of fitness effects near 0, while only a tiny minority had fitness scores >1, as expected (Fig 1B). This distribution is consistent with previous studies across multiple virus families (including IAV) that indicate the vast majority of mutations have deleterious effects on relative fitness (Sanjuan, 2010; Sanjuan et al., 2004; Visher et al., 2016; Wu et al., 2014). Surprisingly, the normalized distribution of fitness effects (DFE) for rPR8-NA:K253R was shifted significantly compared with WT (Fig 1B, C), with a smaller peak of lethal or near-lethal substitutions with fitness scores of ~0 and a large increase in substitutions with fitness scores of ~1, indicating neutral or nearly-neutral effects on relative fitness. Substitutions at 150 out of 325 HA1 residues in rPR8-NA:K253R exhibited significant shifts ($p<0.01$, t test) in fitness scores compared with rPR8-WT. Finally, the DFE for rPR8-NA:H274Y was also shifted but to a lesser extent, consistent with the minimal effect of this substitution on NA function (Fig 1B). Altogether, these data suggest that phenotypic variation in NA can have widely distributed effects on the mutational tolerance of the HA gene.

To better understand how phenotypic variation in NA can affect mutational tolerance at specific residues, we calculated the differences in normalized fitness scores of individual HA1 substitutions between rPR8-WT and rPR8-NA:K253R (Fig 1D). In this analysis, positive difference values for individual substitutions indicate higher relative fitness in the NA:K253R background compared with NA:WT while negative values indicate a higher relative fitness in the WT background. Negative fitness score differentials were clearly enriched in a subset of residues, many of which are located near the receptor binding site (RBS).

To define how NA influences mutational tolerance across the HA structure, we calculated per-residue mean of absolute difference (MAD) values (quantifies how the combined absolute mutational fitness effects for every possible substitution at each residue change depending on NA background) of every residue in HA1 and plotted these values on the structure (Fig 1E). Key residues involved in receptor specificity, including E190 and D225, had dramatically higher MAD values, indicating that epistatic interactions with NA are enriched in residues involved in receptor binding. Overall, NA-dependent fitness differentials had significantly lower values for RBS-associated residues compared with those elsewhere in HA1 ($p = 0.003739$, unpaired two-sample Wilcoxon test), indicating that RBS substitutions tended to be better tolerated in rPR8-WT than rPR8-NA:K253R (Fig 1F).

From the heatmap in Fig 1D, we observed that the mutational fitness effect patterns of the positively charged amino acids arginine and lysine were distinct from other amino acids. We hypothesized that substitutions that increase net charge may be more tolerated within

the WT NA background as these changes would likely enhance binding to the negatively charged cell surface via electrostatic interactions. As expected, negative charge changes on HA surface determined by GetArea (Fraczkiewicz and Braun, 1998) were better tolerated in rPR8-NA:K253R while positive charge changes had higher fitness in rPR8-WT (Fig 1G). Altogether, these observations indicate that epistatic interactions with NA are most pronounced for HA residues involved in receptor binding and cell adhesion.

Our DMS data clearly demonstrated that variation in NA activity can significantly alter the mutational tolerance of the HA1 domain, particularly at antigenically significant residues surrounding the RBS (Caton et al., 1982; Koel et al., 2013). We next asked whether different NA backgrounds support the emergence of distinct repertoires of escape variants under neutralizing antibody selection. We performed *in vitro* selection experiments using the Sb epitope-specific neutralizing monoclonal antibody (mAb) H36–26 (Yewdell, 2010) and the three recombinant viruses detailed above (rPR8-WT, rPR8-NA:K253R, and rPR8-NA:H274Y). Importantly, all three viruses encoded identical HA sequences. We infected MDCK cells with 10^7 TCID50 of each virus in sextuplicate in the presence of H36–26 at a concentration where neutralization is saturated under these conditions (Fig S2A). We passaged viral populations twice in the presence of H36–26 to reach sufficient titers for sequencing (>10^4 TCID50/mL). We then deep sequenced post-selection viral populations and identified single nucleotide variants (SNVs) that emerged above background using DeepSNV (Gerstung et al., 2014, 2012).

Distinct repertoires of escape substitutions emerged in mutant and WT NA backgrounds. In the NA:WT background, HA:E156K emerged to high frequency (>60%) in 6/6 replicate populations, while K189Q was also observed at frequencies between 2% and 20% in 5/6 populations (Fig 2A). In contrast, neither E156K nor K189Q were observed above background in either rPR8-NA:K253R or rPR8-NA:H274Y. Instead, Q196K/R substitutions emerged to high frequency in 5/5 (one rPR8-NA:K253R replicate failed to grow) or 6/6 replicates for rPR8-NA:K253R and rPR8-NA:H274Y, respectively. Q196R was present above background in 2/6 WT populations while Q196K was not observed. Both E156 and Q189 are located within the canonical Sb epitope (Fig 2B). These results are consistent with the DMS data, where HA:E156K showed higher relative fitness than HA:Q196K in the NA:WT background while HA:Q196R showed higher relative fitness than HA:E156K in the NA:K253R background (Fig S2B).

We hypothesized that viruses with lower relative NA activity would support the emergence of HA escape variants with lower receptor binding avidity compared with viruses with higher NA activity. We compared the receptor binding avidities of HA:E156K and HA:Q196K using bio-layer interferometry (BLI). As expected, HA:E156K exhibited a higher receptor binding on-rate compared with HA:Q169K (Fig S2C).

To test the generality of this observation, we performed a similar selection experiment using the Ca1-specific mAb H17-L2 (Das et al., 2013). Again, we observed distinct repertoires of escape variants in WT versus NA mutant backgrounds (Fig 2C). G173W, S207P/L and G240R were the dominant substitutions that emerged above background for NA:WT viruses but were only sporadically observed in the NA:K253R and NA:H274Y

backgrounds. Instead, I93T, V169M, and G240E were exclusively identified within the NA:K253R and/or NA:H274Y backgrounds. All these residues are within or adjacent to the canonical Ca1 epitope (Fig 2D). Our results clearly demonstrate that different NA genotypes can reproducibly foster the emergence of distinct repertoires of HA escape variants under neutralizing antibody selection.

Our results reveal how phenotypic variation in NA can profoundly reshape the fitness landscape available to the HA gene, thus determining its potential for future adaptation. The importance of balancing the opposing activities of the HA and NA glycoproteins for maximizing viral fitness is well established (Kosik and Yewdell, 2019). The importance of epistatic networks within HA has also been demonstrated to influence antigenic evolution (Kryazhimskiy et al., 2011; Wu et al., 2018). Here, we extend these concepts by demonstrating how intersegment epistasis arising from the intimate functional relationship between HA and NA significantly constrains the evolutionary potential of HA, making the specific mutational pathways taken by HA highly contingent upon the associated NA gene. Our results strongly suggest that the evolution of the HA gene cannot be viewed in isolation and that efforts to predict the evolutionary trajectories of seasonal IAVs must account for the influence of the associated NA segment.

While we observed significant epistatic interactions between NA and the HA RBS as expected, we found that the mutational tolerance of numerous residues distal from the RBS were also significantly influenced by NA background. Epistatic effects of NA on these residues may still be largely driven by HA/NA balance issues, as numerous mechanisms can influence the ability of HA to facilitate receptor binding. Another possibility is that destabilizing substitutions in HA (that potentially decrease proper folding, trafficking, virion incorporation, and/or receptor binding) may be better tolerated in genetic backgrounds with lower relative NA activity.

Related to this, our data also demonstrate how variation in NA activity can modulate the overall mutational robustness of HA. NA:K253R was associated with a positive overall shift in the mean fitness effects distribution for HA1, indicating that the relative fitness costs of a large number of substitutions were decreased in the context of reduced NA activity. We hypothesize that this effect is due to an NA gene with lower relative activity imposing weaker functional constraints on HA and thus better tolerating amino acid substitutions that diminish overall HA receptor binding avidity.

Variation in the overall mutational robustness of the HA gene as a function of NA phenotype has broader implications for IAV evolution. Mutational robustness has been hypothesized to facilitate the adaptive potential of proteins (and by extension, viral populations) under some conditions (Bloom et al., 2006; de Visser et al., 2003; Draghi et al., 2010; Elena, 2012; Lauring et al., 2013; McBride et al., 2008). By buffering mutational fitness effects, increases in robustness can promote the accumulation of genetic variants that may confer enhanced fitness or rescue in changing environmental conditions. In an extreme example, a zoonotic IAV population encoding a more mutationally robust HA gene would be more likely to accumulate substitutions that could facilitate successful cross-species transmission. Alternatively, a recent study demonstrated that increasing the robustness of a bacteriophage

protein actually decreased the potential of the virus to evolve to expand its host range (Strobel et al., 2022). During IAV antigenic evolution, deleterious mutation load has been hypothesized to govern the emergence potential of antigenic escape variants at the host population level, suggesting another mechanism by which HA mutational robustness could influence antigenic drift (Koelle and Rasmussen, 2015). Altogether, our results suggest that variation in NA may have profound, if difficult to predict, effects on the evolvability of the HA gene.

Our results provide proof-of-principle that epistatic interactions between HA and NA play a major role in shaping the fitness landscape of the HA gene (and almost certainly the NA gene as well) and in determining the most likely genetic pathways of antigenic evolution. This study used the PR8 strain, which is mouse- and egg-adapted and thus in many ways not representative of IAV strains circulating in people, thus future studies are needed to define how intersegment epistasis influences the evolution of recent seasonal IAVs. Regardless, our data clearly establish how the need to maintain functional balance between HA and NA activities imposes a significant constraint on the evolutionary potential of influenza viruses and should be considered in efforts to predict future evolutionary trajectories.

STAR Methods

RESOURCE AVALABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Christopher Brooke (cbrooke@illinois.edu)

Materials availability—The PR8 HA1 DMS plasmid library and all reverse genetics plasmids used in this study are available from the lead contact with no restriction.

Data and code availability—Next generation sequencing data generated for this study has been deposited at SRA (Accession: PRJNA842406).

Custom python scripts for analyzing the deep mutational scanning data have been deposited to https://github.com/Wangyiquan95/HA1

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines—Madin-Darby canine kidney (MDCK) and human embryonic kidney HEK293T (293T) cells were passaged and maintained in Minimum Essential Medium (MEM + GlutaMAX, ThermoFisher Scientific) with 8.3% fetal bovine serum (FBS, Avantor Seradigm Premium Grade Fetal Bovine Serum) in 37°C and 5% CO2. MDCK cells and 293T cells were gifts of Dr. Jonathan Yewdell and Dr. Joanna Shisler, respectively. Cell lines were authenticated based on morphology, virus growth, and transfection efficiency phenotypes.

Viruses—A/Puerto Rico/8/1934 (PR8) and the specific PR8 mutants (generated by PCR mutagenesis) used in the study were generated via standard reverse genetics. In brief, ~60% confluent 293T cells in 6-well plate were transfected with 500 ng of each segment cloned in the pDZ reverse genetics vector (jetPRIME, Polyplus Transfection). After 24 hours, the medium was replaced by the infection medium (MEM + 1 μg/mL TPCK-treated trypsin + 1 mM HEPES and 50 μg/mL gentamicin). Supernatants from transfected cells were collected at 48 hours post transfection and used to infect MDCK cells in 6-well plate to generate the seed stock. Seed stocks were collected at 48 hours post infection (hpi) or upon development of cytopathic effect (CPE), which ever came first. Working stocks were generated by infecting MDCK cells in T75 or T175 flask with seed stock at an MOI of 0.0001 TCID50/cell and collecting the supernatant at 48 hpi or when early signs of CPE were observed, which ever came first. Virus stocks were tittered by standard TCID50 assay. The plasmids were generously provided by Dr. Jonathan Yewdell.

The secondary structures forming the receptor binding site are defined as: T132-V135 (130 loop), T155-P162 (150 loop), N187-Y195 (190 helix), A218-D225 (220 loop) (Tzarum et al., 2017; Wu et al., 2013). The residue numbering of HA and NA is based on alignment to structural numbering of H3N2 (strain A/Hong Kong/1/1968 H3N2 (Brown et al., 2001), UniProt: Q91MA7, Q91MA2 for HA and NA) unless specified otherwise.

METHOD DETAILS

MUNANA assay—The substrate 2′-(4-Methylumbelliferyl)-α-D-N-acetylneuraminic acid sodium salt hydrate (MUNANA, Sigma-Aldrich) was dissolve in NA buffer (33 mM MES, 4 mM CaCl2 within 1X PBS, pH = 6.5) and aliquoted. A black 96-well half-well flat bottom plate, the plate reader, NA buffer and the substrate was preheated to 37°C. 25 μL of virus sample (diluted in NA buffer) was mixed with 20 μL of the substrate (200 μM) and taken to the plate reader to measure the fluorescent kinetic (excitation wavelength = 365nm, emission wavelength = 450 nm) for 45 min. V_{max} values were estimated based on data collected after the first 10 min of the assay. The results were normalized based on the genome equivalent of NP segment determined by RT-qPCR.

Cellular surface staining of NA—MDCK cells were infected with MOI = 0.05 based on the TCID50 titer of the viruses. MEM + 8.3%FBS was added on the cells after infection for 1 hour and replaced by NH_4Cl medium (MEM, 50 mM HEPES, 20 mM NH_4Cl, pH = 7.2) to block secondary infection. Cells were collected 16 hpi and stained with NA antibody (NA2-1C1-AF488, 1:1600) without permeabilization and run on a BD FACSAria Flow Cytometer. The NA positive cells were gated based on the no infection controls and the expression level were measured by the mean fluorescence intensity.

Western blotting—The virus stocks used for western blot were purified by ultra-centrifugation. Briefly, 10 mL of 20% sucrose with 30 mL of the virus stock was centrifuged at 27000 rpm, 4°C for 2 hours. The pellet was dissolved in PBS overnight. The product was then added on the top of a cushion of 15% sucrose and 60% sucrose and centrifuged. The virus band was collected after and washed by PBS. The purified viruses were heated at 98°C for 2 min, loaded on Bis-Tris protein gel (Bolt 4–12% Bis-Tris Plus, invitrogen), run

on 150 V for 45 min and then transferred to PVDF membrane (iBlot2 PVDF Mini Stacks, invitrogen). The membrane was then stained with HA antibody (RA5-22, obtained through BEI Resources) and rabbit anti-NA polyclonal antibody (gift of Dr. Jonathan Yewdell).

Deep mutation scanning of HA1—To generate all possible amino acid substitutions within HA1, NNK was introduced into each codon of the interest (D18-S342, H1 numbering from the first amino acid residue) by overlapping PCR (Phusion High-Fidelity DNA Polymerase, ThermoFisher Scientific) to generate full-length HA amplicons (Using primers in Table S2). NNK-mutagenized HA amplicons were then cloned into pDZ vector by T4 ligation (T4 DNA Ligase, New England BioLabs Inc.). The ligation product was then used to transform DH10B competent cells (MegaX DH10B T1, Invitrogen) and yield 9.8×10^5 colonies (150X coverage of all substitutions) in total from two transformation reaction. The colonies were harvested and the plasmids were extracted by Midi prep (HiSpeed Plasmid Midi Kit, QIAGEN).

To rescue the viruses, 1.25×10^7 293T cells and 6×10^6 MDCK cells were mixed and seeded per T175 flask in 25 mL of cell growth medium (MEM + 8.3% FBS). The next day, 7 μg each of the 8 reverse genetics plasmids was mixed with 112 μL of jetPRIME reagent (Polyplus Transfection) and 1.2 mL jetPRIME buffer (Polyplus Transfection) and this mixture was added to the medium. After 24 hours of transfection, the medium was removed, cells were washed with PBS, and 20 mL of infection medium was added. Supernatants were collected 72 hours post transfection. Each virus was passaged once for 24 hours in a T175 flask of confluent MDCK cells at a starting MOI of 0.05 TCID50/cell.

Viral RNA was extracted (QIAamp Viral RNA Kits, QIAGEN) and used to generate the cDNA (SuperScript III, ThermoFisher Scientific). For RT reaction, 11.8 μL of viral RNA was mixed with 1 μL random hexamer (125 ng/μL) and incubated at 65°C for 5 min for primer binding. The mixture was then added to be a 20 μL reaction and incubated at 50°C for 1 hour to generate cDNA (SuperScript III, ThermoFisher Scientific). The HA1 sequence was divided into three fragments for sequencing. Seven Ns were added into the primers as the barcode for barcoded-subamplicon sequencing during the first round PCR using primers described in Table S3 (PrimeSTAR Max DNA Polymerase, Takara Bio, PCR condition was set according to the manufacturer). Equal amount of the products from the three fragments of each sample were purified (PureLink Quick Gel Extraction Kit, ThermoFisher Scientific) and mixed. 1.3×10^6 copies of the PCR product from the mixture above were added into the second round PCR (KOD Hot Start DNA Polymerase, Novagen) to add the adapter for sequencing. The condition of the second round PCR was set according to the manufacture with an annealing temperature of 58°C for 25 cycles. The second round PCR products were then purified and submitted for sequencing on an SP line for 251 cycles from both ends of the fragments on a NovaSeq 6000 (250 nt, paired-end reads).

Monoclonal antibody selection—The saturated neutralization concentration was determined first for the selecting escape variant. Briefly, 10^7 TCID50 of each virus were incubated with the monoclonal antibody with different concentrations in duplicate at 37°C for 30 min to equilibrate binding. Virus-antibody complexes were incubated with MDCK cells for 1 hour at 37°C. The cells were then washed by PBS and infection medium with the

same concentration of antibody was added. The virus supernatants were then collected at 16 hpi and measured the titers by TCID50 assay. The saturated neutralization was decided by where the neutralization curve reached the plateau.

The virus stocks used for selection experiments were pooled from three independent rescues. The selection experiments were carried out as the same setting above with a saturate neutralization concentration escape variants were passaged if required and harvested at 16 hpi to obtain sufficient viral load (>10^4 TCID50/mL) for high quality genome sequencing. No antibody treatment control groups were passaged in parallel. Viral RNA was extracted (QIAamp Viral RNA Kits, QIAGEN) from the supernatants and served as the template for RT-PCR (SuperScript III, ThermoFisher Scientific; Phusion High-Fidelity DNA Polymerase, ThermoFisher Scientific). For RT reaction, 10 μL of viral RNA and 1 μL MBTUni-12 primer (5′-ACG CGT GAT CAG CAA AAG CAG G-3′) was added into a 20 μL reaction. 10 μL cDNA template was amplified with MBTUni-12 and MBTUni-13 (5′-ACG CGT GAT CAG TAG AAA CAA GG-3′) primers and an annealing temperature at 57°C for 25 cycles. The input templates in PCR were normalized by the NP genome equivalents determined by qPCR. The PCR products were purified (PureLink Quick Gel Extraction Kit, ThermoFisher Scientific) and used to generate the shotgun library (KAPA HyperPrep Kits, Roche) and then sequenced one MiSeq flowcell for 251 cycles using a MiSeq 500-cycle sequencing kit version 2. The DeepSNV pipeline (Gerstung et al., 2014, 2012) was used to identify minor sequence variants. 2% minimum frequency threshold was set for variant calling.

Biolayer interferometry (BLI)—The virus stocks used for BLI were purified by ultra-centrifugation as stated above. The protein concentration of the virus stock was determined by Bradford assay (Pierce™ Coomassie Plus (Bradford) Assay Kit, ThermoFisher Scientific). Streptavidin sensors (ForteBio) were coated with 500 nM 3'-SLN-PEG3-biotin (3'-Sialyllactosamine-PEG3-Biotin (Single Arm), Sussex Research). Equal concentrations of each virus were run on the BLI detection system (octet RED96e, ForteBio) for association for 300 seconds in the presence of 10 μM zanamivir to inhibit NA activity. Sensors were then incubated in PBS with zanamivir for 300 seconds for dissociation.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of deep mutational scanning data—Sequencing data were processed as described before (Wang et al., 2021) to generate read counts for each amino acid substitution. The adaptors were trimmed, the FASTQ files were generated and demultiplexed with the bcl2fastq v2.20 Conversion Software (Illumina). Sequencing data was obtained in FASTQ format and analyzed using a custom Python snakemake pipeline (Mölder et al., 2021). First, UMIs were merged using python script with parameters: NNNNNNN 0.8 2, to obtain 5'-end paired UMIs with pattern 'NNNNNNN', at least two sequences with same UMIs, of which at least 80% are consensus sequence. Subsequently, primer sequences were trimmed using cutadapt (Martin, 2011), and then sequencing reads were renamed based on amplicon primer. Before variant calling and the fitness calculations, cleaned paired-end reads were merged by FLASH (Magoč and Salzberg, 2011) using parameters: -m 30 -M 70 -I. Finally, variants and normalized fitness values were calculated by python script as described previously (Wang et al., 2021). Briefly, merged paired-end sequences were firstly

parsed by SeqIO module in BioPython (Cock et al., 2009) and then translated into protein sequences. Reads were filtered and removed if there is no amplicon tag or the sequence length was incorrect. Afterwards, variants of each residue were counted and normalized by amplicon to generate the frequency of each mutation. The enrichment ratio of each amino acid was calculated as follows:

$$enrichment_{aa} = \frac{freq_{aa,\ virus}}{freq_{aa,\ plasmid}}$$

in which $freq_{aa,\ virus}$ represents the frequency of a certain amino acid in the subamplicon in the output virus population after passage; $freq_{aa,\ plasmid}$ represents the frequency of the amino acid in the subamplicon in the input plasmid.

The normalized fitness scores of individual amino acid substitutions were calculated based on the enrichment ratio of the amino acid normalized by the mean enrichment ratios of silent mutations and nonsense mutations (the average of enrichment ratio for all silent mutations and nonsense mutations calculated the same way as above), assuming that nonsense mutations would have fitness scores of 0 (lethal) and synonymous mutations would have fitness scores of 1 (neutral):

$$fitness_{aa} = \frac{log(enrichment_{aa}) - log(enrichment_{nonsense})}{log(enrichment_{silent}) - log(enrichment_{nonsense})}$$

in which $enrichment_{silent}$ and $enrichment_{nonsense}$ are the mean of the enrichment ratio of the silent mutations and the nonsense mutations across HA1 for a given sample.

Statistical tests—The statistical details for each experiment are described in the figure legends. The statistical tests were conducted using R or GraphPad Prism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work has been generously supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under awards K22AI116588 and R01AI139246 to C.B.B., R00AI139445 and R01AI167910 to N.C.W., the Roy J. Carver Charitable Trust under award 17-4905 to C.B.B., and startup funds from the University of Illinois.

We are grateful to Alvaro Hernandez, Chris Wright and the DNA services team in Roy J. Carver Biotechnology center for generating high quality sequencing results and their advice on primer designing. We thank Gloria Rendon in HPCBio at Roy J. Carver Biotechnology center for her assistance with the DeepSNV pipeline.

References

Bloom JD, Gong LI, Baltimore D, 2010. Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. Science 328, 1272–1275. 10.1126/science.1187816 [PubMed: 20522774]

Bloom JD, Labthavikul ST, Otey CR, Arnold FH, 2006. Protein stability promotes evolvability. Proceedings of the National Academy of Sciences of the United States of America 103, 5869–5874. 10.1073/pnas.0510098103 [PubMed: 16581913]

Brooke CB, 2017. Population Diversity and Collective Interactions during Influenza Virus Infection. J. Virol 91. 10.1128/JVI.01164-17

Brooke CB, Ince WL, Wei J, Bennink JR, Yewdell JW, 2014. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. Proc. Natl. Acad. Sci. U.S.A 111, 16854–16859. 10.1073/pnas.1415396111 [PubMed: 25385602]

Brown EG, Liu H, Kit LC, Baird S, Nesrallah M, 2001. Pattern of mutation in the genome of influenza A virus on adaptation to increased virulence in the mouse lung: Identification of functional themes. Proceedings of the National Academy of Sciences 98, 6883–6888. 10.1073/pnas.111165798

Carrat F, Flahault A, 2007. Influenza vaccine: The challenge of antigenic drift. Vaccine. 10.1016/j.vaccine.2007.07.027

Caton AJ, Brownlee GG, Yewdell JW, Gerhard W, 1982. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31, 417–427. [PubMed: 6186384]

Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, de Hoon MJL, 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25, 1422–1423. 10.1093/bioinformatics/btp163 [PubMed: 19304878]

Das SR, Hensley SE, Ince WL, Brooke CB, Subba A, Delboy MG, Russ G, Gibbs JS, Bennink JR, Yewdell JW, 2013. Defining influenza A virus hemagglutinin antigenic drift by sequential monoclonal antibody selection. Cell Host Microbe 13, 314–323. 10.1016/j.chom.2013.02.008 [PubMed: 23498956]

de Visser JAGM, Hermisson J, Wagner GP, Ancel Meyers L, Bagheri-Chaichian H, Blanchard JL, Chao L, Cheverud JM, Elena SF, Fontana W, Gibson G, Hansen TF, Krakauer D, Lewontin RC, Ofria C, Rice SH, von Dassow G, Wagner A, Whitlock MC, 2003. Perspective: Evolution and detection of genetic robustness. Evolution 57, 1959–1972. [PubMed: 14575319]

de Vries E, Du W, Guo H, de Haan CAM, 2020. Influenza A Virus Hemagglutinin–Neuraminidase–Receptor Balance: Preserving Virus Motility. Trends in Microbiology 28, 57–67. 10.1016/j.tim.2019.08.010 [PubMed: 31629602]

Doud MB, Bloom JD, 2016. Accurate measurement of the effects of all amino-acid mutations on influenza hemagglutinin. Viruses 8, 155. 10.3390/v8060155

Doud MB, Hensley SE, Bloom JD, 2017. Complete mapping of viral escape from neutralizing antibodies. PLoS Pathog. 13, e1006271. 10.1371/journal.ppat.1006271 [PubMed: 28288189]

Draghi JA, Parsons TL, Wagner GP, Plotkin JB, 2010. Mutational robustness can facilitate adaptation. Nature 463, 353–355. 10.1038/nature08694 [PubMed: 20090752]

Elena SF, 2012. RNA virus genetic robustness: possible causes and some consequences. Curr Opin Virol 2, 525–530. 10.1016/j.coviro.2012.06.008 [PubMed: 22818515]

Fowler DM, Fields S, 2014. Deep mutational scanning: a new style of protein science. Nat Methods 11, 801–807. 10.1038/nmeth.3027 [PubMed: 25075907]

Fraczkiewicz R, Braun W, 1998. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. J. Comput. Chem 19, 319–333. 10.1002/(SICI)1096-987X(199802)19:3<319::AID-JCC6>3.0.CO;2-W

Gaymard A, Le Briand N, Frobert E, Lina B, Escuret V, 2016. Functional balance between neuraminidase and haemagglutinin in influenza viruses. Clinical Microbiology and Infection 22, 975–983. 10.1016/j.cmi.2016.07.007 [PubMed: 27424943]

Gerstung M, Beisel C, Rechsteiner M, Wild P, Schraml P, Moch H, Beerenwinkel N, 2012. Reliable detection of subclonal single-nucleotide variants in tumour cell populations. Nature Communications 3, 1–8. 10.1038/ncomms1814

Gerstung M, Papaemmanuil E, Campbell PJ, 2014. Subclonal variant calling with multiple samples and prior knowledge. Bioinformatics 30, 1198–1204. 10.1093/bioinformatics/btt750 [PubMed: 24443148]

Harris A, Cardone G, Winkler DC, Heymann JB, Brecher M, White JM, Steven AC, 2006. Influenza virus pleiomorphy characterized by cryoelectron tomography. Proceedings of the National Academy of Sciences 103, 19123–19127. 10.1073/pnas.0607614103

Hensley SE, Das SR, Gibbs JS, Bailey AL, Schmidt LM, 2011. Influenza A Virus Hemagglutinin Antibody Escape Promotes Neuraminidase Antigenic Variation and Drug Resistance. PLoS ONE 6, 15190. 10.1371/journal.pone.0015190

Ives JAL, Carr JA, Mendel DB, Tai CY, Lambkin R, Kelly L, Oxford JS, Hayden FG, Roberts NA, 2002. The H274Y mutation in the influenza A/H1N1 neuraminidase active site following oseltamivir phosphate treatment leave virus severely compromised both in vitro and in vivo. Antiviral Research 55, 307–317. 10.1016/S0166-3542(02)00053-0 [PubMed: 12103431]

Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GCM, Vervaet G, Skepner E, Lewis NS, Spronken MIJ, Russell CA, Eropkin MY, Hurt AC, Barr IG, de Jong JC, Rimmelzwaan GF, Osterhaus ADME, Fouchier RAM, Smith DJ, 2013. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. Science 342, 976–979. 10.1126/science.1244730 [PubMed: 24264991]

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. 10.7554/eLife.07361

Kosik I, Ince WL, Gentles LE, Oler AJ, Kosikova M, Angel M, Magadán JG, Xie H, Brooke CB, Yewdell JW, 2018. Influenza A virus hemagglutinin glycosylation compensates for antibody escape fitness costs. PLoS Pathog. 14, e1006796. 10.1371/journal.ppat.1006796 [PubMed: 29346435]

Kosik I, Yewdell JW, 2019. Influenza Hemagglutinin and Neuraminidase: Yin(−)Yang Proteins Coevolving to Thwart Immunity. Viruses 11. 10.3390/v11040346

Kryazhimskiy S, Dushoff J, Bazykin GA, Plotkin JB, 2011. Prevalence of epistasis in the evolution of influenza A surface proteins. PLoS Genetics 7, e1001301. 10.1371/journal.pgen.1001301 [PubMed: 21390205]

Lauring AS, Frydman J, Andino R, 2013. The role of mutational robustness in RNA virus evolution. Nature Reviews Microbiology 2013 11:5 11, 327–336. 10.1038/nrmicro3003 [PubMed: 23524517]

Lee JM, Huddleston J, Doud MB, Hooper KA, Wu NC, Bedford T, Bloom JD, 2018. Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. Proceedings of the National Academy of Sciences of the United States of America 115, E8276–E8285. 10.1073/pnas.1806133115 [PubMed: 30104379]

Magoč T, Salzberg SL, 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27, 2957–2963. 10.1093/bioinformatics/btr507 [PubMed: 21903629]

Martin M, 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12. 10.14806/ej.17.1.200

Martínez-Sobrido L, García-Sastre A, 2010. Generation of recombinant influenza virus from plasmid DNA. J Vis Exp 2057. 10.3791/2057 [PubMed: 20729804]

McBride RC, Ogbunugafor CB, Turner PE, 2008. Robustness promotes evolvability of thermotolerance in an RNA virus. BMC Evol Biol 8, 231. 10.1186/1471-2148-8-231 [PubMed: 18694497]

Mitnaul LJ, Matrosovich MN, Castrucci MR, Tuzikov AB, Bovin NV, Kobasa D, Kawaoka Y, 2000. Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. J. Virol 74, 6015–6020. [PubMed: 10846083]

Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-Tinch CH, Sochat V, Forster J, Lee S, Twardziok SO, Kanitz A, Wilm A, Holtgrewe M, Rahmann S, Nahnsen S, Köster J, 2021. Sustainable data analysis with Snakemake. 10.12688/f1000research.29032.2

Paget J, Spreeuwenberg P, Charu V, Taylor RJ, Iuliano AD, Bresee J, Simonsen L, Viboud C, 2019. Global mortality associated with seasonal influenza epidemics: New burden estimates and predictors from the GLaMOR Project. Journal of Global Health 9. 10.7189/jogh.09.020421

Pauly MD, Procario MC, Lauring AS, 2017. A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses. eLife 6. 10.7554/eLife.26437

Petrova VN, Russell CA, 2018. The evolution of seasonal influenza viruses. Nat Rev Microbiol 16, 47–60. 10.1038/nrmicro.2017.118 [PubMed: 29081496]

Powell H, Pekosz A, 2020. Neuraminidase antigenic drift of H3N2 clade 3c.2a viruses alters virus replication, enzymatic activity and inhibitory antibody binding. PLoS Pathog 16, e1008411. 10.1371/journal.ppat.1008411 [PubMed: 32598381]

Randall CMH, Biswas S, Selen CV, Shisler JL, 2014. Inhibition of interferon gene activation by death-effector domain-containing proteins from the molluscum contagiosum virus. Proc. Natl. Acad. Sci. U.S.A 111. 10.1073/pnas.1314569111

Sanjuan R, 2010. Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. Philosophical Transactions of the Royal Society B: Biological Sciences 365, 1975–1982. 10.1098/rstb.2010.0063 [PubMed: 20478892]

Sanjuan R, Moya A, Elena SF, 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proceedings of the National Academy of Sciences 101, 8396–8401. 10.1073/pnas.0400146101

Strobel HM, Horwitz EK, Meyer JR, 2022. Viral protein instability enhances hostrange evolvability. PLOS Genetics 18, e1010030. 10.1371/journal.pgen.1010030 [PubMed: 35176040]

Tzarum N, de Vries RP, Peng W, Thompson AJ, Bouwman K, McBride R, Yu W, Zhu X, Verheije MH, Paulson JC, Wilson IA, 2017. The 150-loop restricts the host specificity of human H10N8 influenza virus. Cell Rep 19, 235–245. 10.1016/j.celrep.2017.03.054 [PubMed: 28402848]

Vahey MD, Fletcher DA, 2019. Influenza a virus surface proteins are organized to help penetrate host mucus. eLife 8. 10.7554/eLife.43764

Vahey MD, Fletcher DA, 2018. Low-Fidelity Assembly of Influenza A Virus Promotes Escape from Host Cells. Cell. 10.1016/j.cell.2018.10.056

Visher E, Whitefield SE, McCrone JT, Fitzsimmons W, Lauring AS, 2016. The Mutational Robustness of Influenza A Virus. PLoS Pathog. 12, e1005856. 10.1371/journal.ppat.1005856 [PubMed: 27571422]

Wagner R, Matrosovich M, Klenk H-D, 2002. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. Rev. Med. Virol 12, 159–166. 10.1002/rmv.352 [PubMed: 11987141]

Wang Y, Lei R, Nourmohammad A, Wu NC, 2021. Antigenic evolution of human influenza H3N2 neuraminidase is constrained by charge balancing. eLife 10, e72516. 10.7554/eLife.72516 [PubMed: 34878407]

Wasilewski S, Calder LJ, Grant T, Rosenthal PB, 2012. Distribution of surface glycoproteins on influenza A virus determined by electron cryotomography. Vaccine 30, 7368–7373. 10.1016/j.vaccine.2012.09.082 [PubMed: 23063838]

Wu C, Cheng X, Wang X, Lv X, Yang F, Liu T, Fang S, Zhang R, Jinquan C, 2013. Clinical and molecular characteristics of the 2009 pandemic influenza H1N1 infection with severe or fatal disease from 2009 to 2011 in Shenzhen, China. Journal of Medical Virology 85, 405–412. 10.1002/jmv.23295 [PubMed: 23280524]

Wu NC, Thompson AJ, Xie J, Lin C-W, Nycholat CM, Zhu X, Lerner RA, Paulson JC, Wilson IA, 2018. A complex epistatic network limits the mutational reversibility in the influenza hemagglutinin receptor-binding site. Nat Commun 9, 1264. 10.1038/s41467-018-03663-5 [PubMed: 29593268]

Wu NC, Young AP, Al-Mawsawi LQ, Olson CA, Feng J, Qi H, Chen SH, Lu IH, Lin CY, Chin RG, Luan HH, Nguyen N, Nelson SF, Li X, Wu TT, Sun R, 2014. High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution. Scientific Reports 4. 10.1038/srep04942

Xu Rui, McBride R, Nycholat CM, Paulson JC, Wilson IA, 2012. Structural Characterization of the Hemagglutinin Receptor Specificity from the 2009 H1N1 Influenza Pandemic. Journal of Virology 86, 982–990. 10.1128/JVI.06322-11 [PubMed: 22072785]

Xu R, Zhu X, McBride R, Nycholat CM, Yu W, Paulson JC, Wilson IA, 2012. Functional Balance of the Hemagglutinin and Neuraminidase Activities Accompanies the Emergence of the 2009 H1N1 Influenza Pandemic. Journal of Virology 86, 9221–9232. 10.1128/jvi.00697-12 [PubMed: 22718832]

Yewdell JW, 2011. Viva la revolución: rethinking influenza a virus antigenic drift. Curr Opin Virol 1, 177–183. 10.1016/j.coviro.2011.05.005 [PubMed: 22034587]

Yewdell JW, 2010. Monoclonal antibodies specific for discontinuous epitopes direct refolding of influenza A virus hemagglutinin. Mol. Immunol 47, 1132–1136. 10.1016/j.molimm.2009.10.023 [PubMed: 20047763]

Highlights

- The evolutionary potential of HA is highly constrained by interactions with NA
- Lower NA activity increases the overall mutational robustness of HA
- Specific immune escape pathways taken by HA can be highly influenced by NA background
- NA must be considered when attempting to predict HA evolution

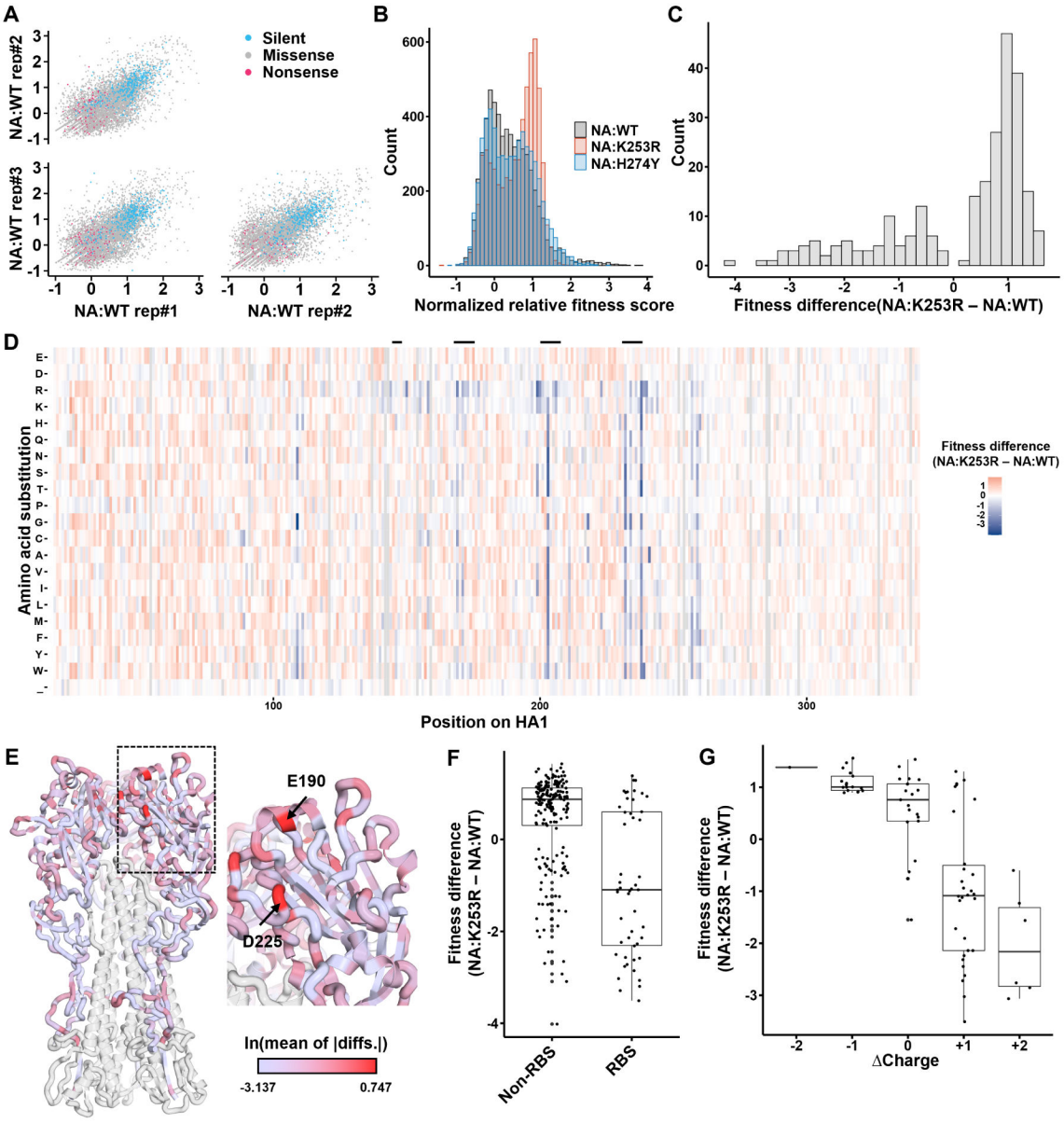


Figure 1. Phenotypic variation in NA reshapes the HA fitness landscape:

(A) Correlations of normalized relative fitness scores across NA:WT replicate populations. Each dot represents the normalized relative fitness score of a specific substitution in the two indicated samples. Silent, missense, and nonsense substitutions are colored as indicated in the legend. **(B)** The distributions of normalized relative fitness scores of all missense substitutions in the indicated genetic backgrounds. Values are averaged across three replicates for each genotype. **(C)** The distribution of differences in missense substitution fitness scores between NA:K253R and NA:WT. Only shows substitutions for which t test on differences between genotypes yielded $p < 0.01$. **(D)** Normalized relative fitness score differences between NA:WT and NA:K253R for each substitution at each residue in HA1 as measured through DMS. Each value was generated by subtracting the mean fitness score of 3 replicates for each genotype. Gray indicates that the mutation had insufficient

coverage in the plasmid library. Residue numbering based on the initiating methionine. Secondary structures forming the receptor binding site (130 loop, 150 loop, 190 helix, 220 loop) are indicated by the black bars above. **(E)** HA structure (PDB:1RU7) showing all HA1 residues colored by the natural log values of per-residue mean of absolute differences (MAD) of all substitutions between NA:WT and NA:K253R. HA2 domain colored in white. **(F)** Normalized relative fitness score differences between residues in HA1 associated with the receptor binding site versus those that are not (only showing substitutions with $p < 0.01$ by t test for comparison between NA:WT and NA:K253R). **(G)** Correlation between normalized relative fitness score differences and charge changes on surface (only showing substitutions with $p < 0.01$ by t test for comparison between NA:WT and NA:K253R).

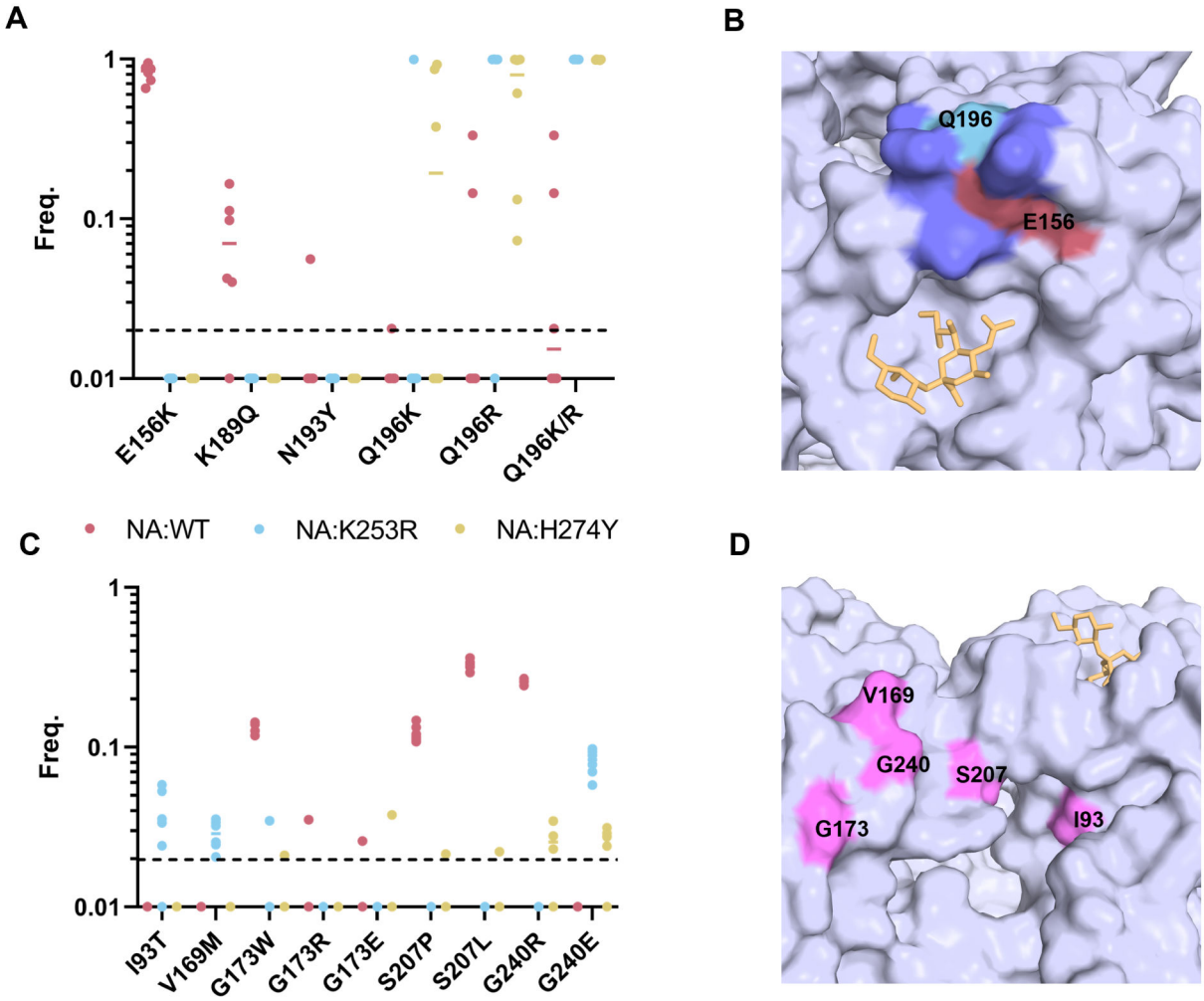


Figure 2: Different NA backgrounds support distinct mutational pathways to escape from neutralizing anti-HA antibodies.

(A) Six independent populations of PR8-WT, PR8-NA:K239R, and PR8-NA:H274Y were passaged twice in MDCK cells in the presence of a neutralizing concentration of the anti-HA mAb H36–26 and deep sequenced. Frequencies of all HA amino acid substitutions detected at frequencies above the 2% frequency threshold (dashed line) across all replicates. Data from no mAb controls not shown but had no variants above 2%. Each dot represents the frequency in a single replicate. **(B)** The position of sialic acid receptor (yellow), E156 (rose), Q196 (cyan), and other Sb epitope residues (purple) on the HA structure (PDB: 3UBQ). **(C)** The frequency of all HA amino acid substitutions detected under H17-L2 selection. Same experimental design as in (A) but using the Ca1-specific mAb H17L2. No mAb controls not shown but had no variants above 2% frequency. **(D)** The position of H17-L2 escape substitutions and the Ca1 epitope in pink and sialic acid receptor (yellow) on HA structure (PDB:3UBQ(Rui Xu et al., 2012)).

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
RA5-22	BEI Resources	Cat# NR-44222
Rabbit anti-NA polyclonal antibody	Jon Yewdell; Brooke et al, 2014	N/A
NA2-1C1	BEI Resources	Cat# NR-50239
H36-26	Jon Yewdell; Brooke et al, 2014	Clone H35-26
H17-L2	Jon Yewdell; Brooke et al, 2014	Clone H17-L2
Bacterial and virus strains		
A/Puerto Rico/8/34/Mount Sinai(H1N1)	Adolfo Garcia Sastre; Martínez-Sobrido and García-Sastre, 2010	N/A
MegaX DH10B T1	Invitrogen	Cat# C640003
Chemicals, peptides, and recombinant proteins		
2′-(4-Methylumbelliferyl)-α-D-N-acetylneuraminic acid sodium salt hydrate	Sigma-Aldrich	Cat# M8639
3'-SLN-PEG3-biotin	Sussex Research	Cat# BT000000
MES hydrate	Sigma	Cat# M5287
jetPRIME	Polyplus Transfection	Cat# 89129
Critical commercial assays		
HiSpeed Plasmid Midi Kit	Qiagen	Cat# 12643
T4 DNA Ligase	New England Biolabs	Cat# M0202
SuperScript III	Invitrogen	Cat# 18080051
Phusion High-Fidelity DNA Polymerase	ThermoFisher	Cat# F530
PrimeSTAR Max DNA Polymerase	Takara Bio	Cat# R045
KOD Hot Start DNA Polymerase	Novagen	Cat# 71086
Bis-Tris protein gel	Invitrogen	Cat# NW04120BOX
PVDF membrane	Invitrogen	Cat# IB24001
Deposited data		
Sequencing data of DMS experiment	This paper	SRA: PRJNA842406
Sequencing data of antibody selection experiment	This paper	SRA: PRJNA842406
Experimental models: Cell lines		
Madin-Darby canine kidney (MDCK) cells	Jon Yewdell; Kosik et al, 2018.	doi:10.1128/JVI.02284-12
Human embryonic kidney HEK293T	Joanna Shisler; Randall et al, 2014	N/A
Oligonucleotides		
HA1 DMS primers, see Table S2	IDT	N/A
HA1 barcode primers for NovaSeq, see Table S3	IDT	N/A
MBTUni-12 primer ACGCGTGATCAGCAAAAGCAGG	IDT	N/A
MBTUni-13 primer ACGCGTGATCAGTAGAAACAAGG	IDT	N/A
Software and algorithms		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HA1 Deep mutation scanning	This paper	https://github.com/Wangyiquan95/HA1 DOI: 10.5281/zenodo.7021429
DeepSNV pipeline	Gerstung et al., 2014, 2012	https://github.com/BROOKELAB/deepSNV_R3.6 DOI: 10.5281/zenodo.7020444
GetArea	Fraczkiewicz and Braun, 1998	http://curie.utmb.edu/getarea.html