Synonymous Mutations at the Beginning of the Influenza A Virus Hemagglutinin Gene Impact Experimental Fitness

Aneth S. Canale[1], Sergey V. Venev[2], Troy W. Whitfield[2,3], Daniel R. Caffrey[3], Wayne A. Marasco[4], Celia A. Schiffer[1], Timothy F. Kowalik[5], Jeffrey D. Jensen[6], Robert W. Finberg[3], Konstantin B. Zeldovich[2], Jennifer P. Wang[3†] and Daniel N.A. Bolon[1,†]

*1 - **Department of Biochemistry and Molecular Pharmacology,** University of Massachusetts Medical School, Worcester, MA 01655, USA*
*2 - **Program in Bioinformatics and Integrative Biology,** University of Massachusetts Medical School, Worcester, MA 01655, USA*
*3 - **Department of Medicine,** University of Massachusetts Medical School, Worcester, MA 01655, USA*
*4 - **Department of Cancer Immunology & Virology,** Dana-Farber Cancer Institute, Harvard Medical School, 450 Brookline Avenue, Boston, MA 02215, USA*
*5 - **Department of Microbiology and Physiological Systems,** University of Massachusetts Medical School, Worcester, MA 01655, USA*
*6 - **School of Life Sciences,** Center for Evolution & Medicine, Arizona State University, Tempe, AZ. 85281, USA*

***Correspondence to Jennifer P. Wang and Daniel N.A. Bolon:** J.P. Wang is to be contacted at: University of Massachusetts, 364 Plantation Street, LRB 219, Worcester, MA, 01605, USA; D.N.A. Bolon is to be contacted at: University of Massachusetts, 364 Plantation Street, LRB 922, Worcester, MA, 01605, USA. Jennifer.Wang@umassmed.edu; Dan.Bolon@umassmed.edu*
https://doi.org/10.1016/j.jmb.2018.02.009
Edited by Anna Panchenko

Abstract

The fitness effects of synonymous mutations can provide insights into biological and evolutionary mechanisms. We analyzed the experimental fitness effects of all single-nucleotide mutations, including synonymous substitutions, at the beginning of the influenza A virus hemagglutinin (HA) gene. Many synonymous substitutions were deleterious both in bulk competition and for individually isolated clones. Investigating protein and RNA levels of a subset of individually expressed HA variants revealed that multiple biochemical properties contribute to the observed experimental fitness effects. Our results indicate that a structural element in the HA segment viral RNA may influence fitness. Examination of naturally evolved sequences in human hosts indicates a preference for the unfolded state of this structural element compared to that found in swine hosts. Our overall results reveal that synonymous mutations may have greater fitness consequences than indicated by simple models of sequence conservation, and we discuss the implications of this finding for commonly used evolutionary tests and analyses.

Introduction

Synonymous mutations change the nucleotide sequence of a gene without changing the encoded protein sequence. According to a simple interpretation of the central dogma of molecular biology, synonymous mutations do not impact fitness because they do not change protein sequences. In this model, all synonymous mutations are neutral or "silent" and the evolution of synonymous substitutions is thus determined by mutational processes and genetic drift (e.g., the likelihood of different nucleotide substitutions made by replicative polymerases) [1]. The availability of a greater number of gene sequences has made it clear that synonymous mutations are subject to natural selection [2–5].

The signatures of selection acting on synonymous substitutions provide insights into evolutionary and biochemical mechanisms. For example, highly expressed genes evolve at a slow rate and are preferentially composed of codons that have cognate tRNA at high abundance [6–8]. The pattern of synonymous

substitutions in these genes led to the hypothesis that high expression selects for codons that can be translated efficiently and/or accurately in order to minimize fitness costs that scale with expression level [9,10].

Despite sequence conservation patterns indicating that natural selection commonly acts on synonymous mutations [2], identification of synonymous mutations with measureable changes in experimental fitness is rare (e.g., Ref. [11]). Many potential explanations could account for this apparent discrepancy [12]. These include experimental conditions that do not encompass the full breadth of environments encountered in nature [13], non-linear relationships between protein function and fitness [14–16], and dramatically different timescales for natural selection events compared to measuring fitness experimentally [12]. For these reasons, when synonymous mutations are made at sites that exhibit strong nucleotide conservation during natural evolution, they often do not cause measureable changes in experimental fitness [17].

Distinct patterns of sequence conservation of synonymous substitutions have been observed at the beginning of genes [18,19]. Rare or non-preferred codons are disproportionately observed in the first 90–150 nucleotides (nt) of genes. Rare codons at the beginning of genes can impact the efficiency of translation initiation and appear to be important in regulating the spacing of ribosomes on mRNA [20]. Investigation of the impacts of synonymous mutations on GFP expression in bacteria indicated that secondary structure at the beginning of the gene has a large impact on the amount of GFP protein expressed [21]. These observations motivated us to search for potential experimental fitness effects at the beginning of the hemagglutinin (HA) gene of influenza A virus (IAV).

Because IAV has pandemic potential in the human population [22], a strong interest exists in understanding detailed mechanisms of how it can evolve. IAV is a negative-strand RNA virus with two main surface proteins, HA and neuraminidase (NA). HA and NA each bind to sialic acid but have opposing functions: HA mediates membrane fusion that facilitates viral entry [23], while NA cleaves sialic acid on host cells to release newly synthesized virions [24]. HA and NA both utilize host machinery for synthesis and trafficking.

Trafficking of proteins to appropriate cellular compartments is a highly conserved process that viruses utilize for propagation in host cells. Signal sequences, located at the N-terminus of proteins, often serve as molecular addresses for targeting secretory proteins to the endoplasmic reticulum (ER) [25–27]. During synthesis of these proteins, which include HA, the signal recognition particle (SRP) binds to the signal sequence and mediates co-translational trafficking of nascent polypeptide chains to the ER (Fig. 1A) [28,29]. As for many secretory proteins, the signal sequence of HA is cleaved by a host protease and does not form part of the mature protein [30–32]. The biochemical mechanisms of HA synthesis are clearly important for generating infectious IAV. However, the sensitivity of IAV infectivity to quantitative changes in the efficiency of HA synthesis or the level of HA in viral particles has not been well defined. Because the signal sequence is ultimately cleaved off before HA exits the ER, mutations introduced into the signal sequence can in principle alter the amount of HA protein generated without changing the sequence of the mature HA protein.

Signal sequences of different proteins often share common biophysical properties. These generally include three key features: positively charged amino acids, followed by a hydrophobic region, followed by polar amino acids [33,34]. The signal sequence of HA ranges from 14 to 17 amino acids in length and contains all three of these features (Fig. 1B), but the amino acid sequence is generally variable between different HA subtypes (Fig. 1C) [35,36]. Previous studies have shown that large deletions in the signal sequence of HA affect protein maturation and cellular localization, underscoring this region's crucial role [37]. However, a systematic analysis of the impact of signal sequence mutations on HA synthesis, surface expression, and viral function has not been performed.

Several groups have used deep mutational scanning approaches to study effects of mutations on IAV genes, including HA and NA. The Bloom group [38,39] studied mutational tolerance and antigenic evolvability in HA at the amino acid level by calculating the average effect of each mutation in a collection of mutants with closely related genetic backgrounds. Wu *et al.* [40] used high-throughput mutational scanning to measure effects of HA single-nucleotide mutations on viral experimental fitness in order to identify potential targets for vaccine and drug design. The impacts of synonymous substitutions in HA on IAV fitness have not been thoroughly investigated using deep mutational scanning, and all of these papers report the effects of amino acid changes that average across synonymous substitutions. Both groups used error-prone PCR to perform mutagenesis across the full-length HA protein. It is important to note here that experimental evolution studies investigate the effects of mutations in a limited genetic background in a controlled environmental condition. Such studies, including ours, do not perfectly mimic natural evolution where more complex genotypes compete in varying environmental conditions and infect different hosts. For these reasons, the conclusions of experimental evolution studies should be carefully considered relative to potential distinctions with natural evolution.

In the present study, we used the EMPIRIC (Exceedingly Meticulous and Parallel Investigation of Randomized Individual Codons) approach that

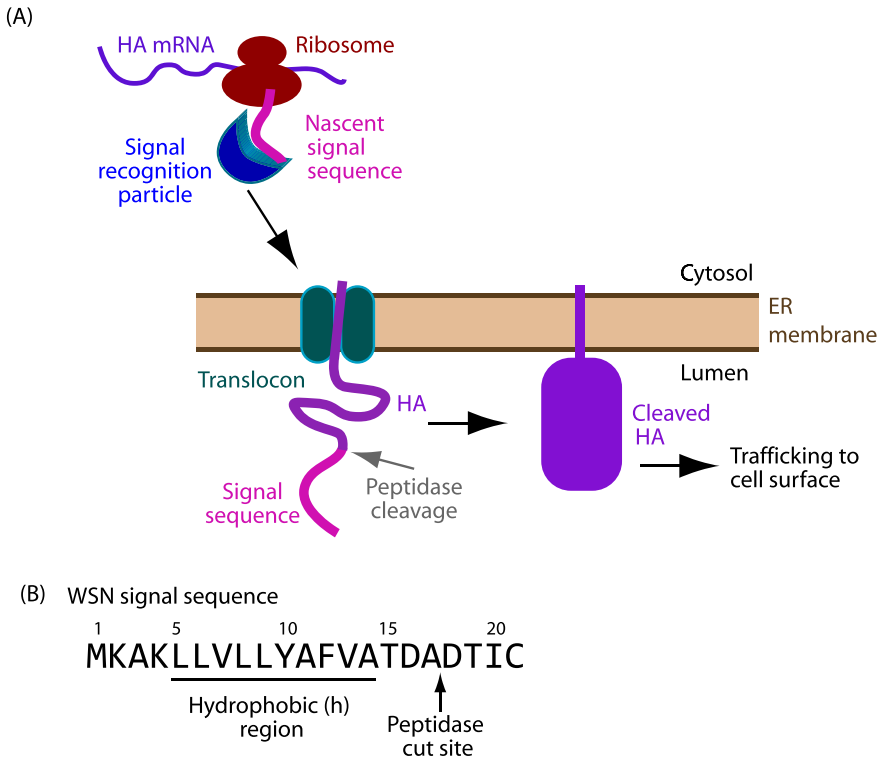

Fig. 1. Function and properties of the signal sequence of HA. (A) Functional illustration. The SRP recognizes the signal sequence of nascent HA, initiating the targeting of HA to the ER lumen. An ER peptidase cleaves the signal sequence, excluding it from the mature protein that is eventually trafficked to the cell surface. (B) Amino acid sequence of the HA signal sequence for A/WSN/33 IAV, highlighting the hydrophobic region important for SRP binding and the peptidase cut site. (C) Amino acid sequence of the HA signal peptide in different HA subtypes. Consensus sequences were generated from 1243 pandemic H1, 1315 non-pandemic H1, 598 H2, 1506 H3, 1585 H5, 1770 H7, 1168 H9, and 925 H10 sequences retrieved from the Influenza Research Database.

we previously developed to quantify the experimental fitness landscapes of genes [14,41–44] to systematically investigate the effects of all possible single-nucleotide mutations, including all possible synonymous mutations, in the signal sequence of HA (Fig. 2A). We measured experimental fitness which, in our study, represents the replication capacity of each mutant calculated as the change in relative mutant frequency before and after bulk competition. The results of our high-throughput screen indicated that many synonymous mutations cause large experimental fitness effects. Using the screen as a guide, we independently analyzed a set of synonymous mutations and again observed many large experimental fitness effects. To probe the physical basis of the experimental fitness effects, we performed biochemical analyses, including estimates of viral RNA (vRNA) and mRNA abundance, and efficiency of surface protein expression on a panel of point mutants.

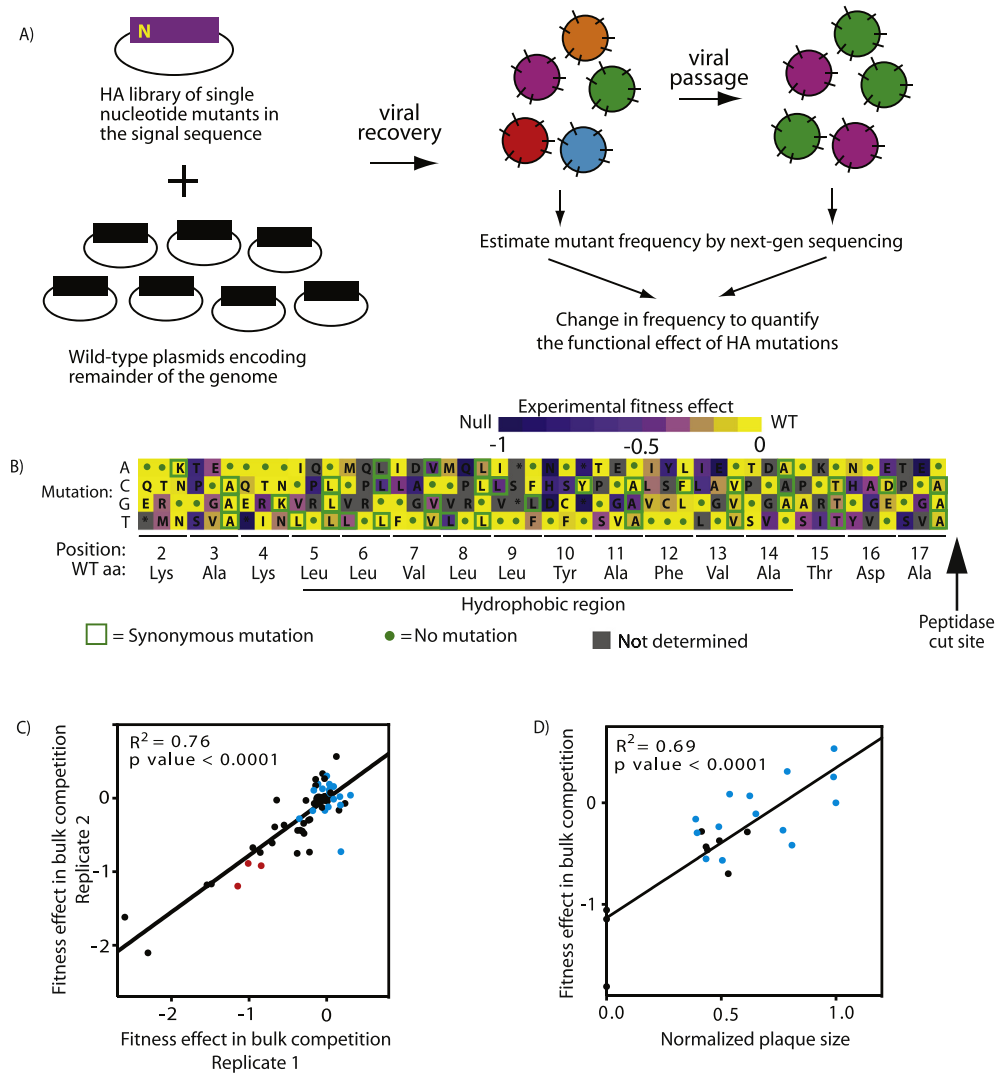

Fig. 2. Experimental fitness effects of mutations in the signal sequence of HA. (A) Systematic approach to quantify the impact of all single base mutations in the signal sequence of HA on experimental fitness. The HA library of single-nucleotide mutants was generated on plasmids that were utilized to recover viruses, recovered viruses were passaged in MDCK cells, and next-generation sequencing was used to estimate the impacts of each mutation on experimental fitness. (B) Heat map representation of the observed impacts of mutations in the HA signal sequence. Synonymous mutations are outlined in green and the wild-type nucleotide at each position is shown by a green dot. (C) Reproducibility of experimental fitness estimates from repeats of the bulk viral expansion for positions 12–21. Synonymous mutations are shown as blue dots, non-synonymous mutations as black dots, and nonsense mutations as red dots (R^2 = 0.76, $P < 0.0001$, from linear regression analysis). (D) Correlation between plaque size of a panel of mutants measured in isolation and average experimental fitness estimates in bulk competition (n = 22, R^2 = 0.69, $P < 0.0001$, from linear regression analysis). Synonymous mutations are shown as blue dots and non-synonymous mutations as black dots.

Results

Quantification of experimental fitness effects of mutations in the HA signal sequence

We systematically generated and measured the relative effects of each single-nucleotide mutation to the HA signal sequence in the A/WSN/33 H1N1 strain (Fig. 2). A number of mutations were depleted during virus recovery from plasmid and/or during subsequent virus infection in the bulk competition assay to the point where the frequency change due to selection could not be accurately determined. While these mutations are likely deleterious, they were marked as “not determined” in the heatmap (Fig. 2B, Supplementary Fig. 1) as the severity of defects could not be quantified. Extra stop codons were included at elevated frequency in our mutant libraries to facilitate quantification of the depletion of

null alleles. Using this approach, the experimental fitness effects of mutations were quantified on a scale from null ($s = -1$) to wild-type ($s = 0$).

Experimental fitness measurements during viral infection were generally reproducible with an $R^2 = 0.76$ in replicates of the bulk competitions (Fig. 2C, Supplementary Fig. 2). As detailed in the Materials and Methods section, experimental repeats were performed for amino acid positions 12–21 in order to provide an estimation of the reproducibility using our methodology. The experimental reproducibility that we observe is in the range seen in similar studies ($R^2 = 0.54$ to 0.78) using deep mutational scanning to study IAV HA [38–40]. For positions 2–11, we relied on changes in mutant frequency between plasmid and P0 to estimate fitness effects. As in our previous work [44], changes in mutant frequency between plasmid and P0 correlate with fitness effects observed between P0 and P1 (Supplementary Fig. 2). For further validation, experimental fitness effects measured in bulk competition were compared to the experimental fitness effects of a panel of 22 HA variants, including 14 synonymous substitutions that were randomly chosen and individually cloned (Supplementary Table 1). For this panel of mutants, plaque size was measured as a surrogate of fitness, larger plaques being indicative of higher experimental fitness. A correlation ($R^2 = 0.69$) was observed between experimental fitness estimates from bulk competitions and plaque size measurements (Fig. 2D) indicating that our high-throughput experiments provide a generally reproducible readout of mutant effects.

Properties underlying selection on the signal sequence of HA

We examined potential similarities and distinctions in selection acting on the signal sequence in our experiments compared with circulating IAV. The experimental fitness effects measured in A/WSN/33 were compared to mutant frequency in 4374 sequenced, non-pandemic H1N1 isolates from the Influenza Research Database [45,46] (see Materials and Methods for details of sequence alignment and filtering). Experimental fitness measurements show a rough trend with mutant frequency in sequenced circulating viral isolates (Fig. 3A and Supplementary Fig. 3). The experimental fitness effects of mutations observed at least twice in circulating isolates show a narrower distribution compared to mutations that were not observed in circulating isolates (Fig. 3A). There were no mutations observed in circulating IAV that exhibited null fitness in our experiments with A/WSN/33, suggesting that similar biophysical properties may mediate selection on the HA signal sequence in our experiments and in circulating IAV. At a finer level of resolution, many mutations observed in circulating IAV exhibited either apparent fitness benefits or defects in our A/WSN/33 experiments (Fig. 3A), indicating that fitness effects of signal sequence mutations may depend on genetic background, or environmental conditions, and/or noise in our experimental fitness measurements.

We investigated how biophysical properties previously associated with signal sequence function correlate with observed experimental fitness effects. The hydrophobic region of the signal sequence serves as the binding site for the SRP and is essential for SRP-mediated targeting of nascent proteins to the ER. The relationship between residue hydrophobicity in the hydrophobic (h-) region of the signal sequence, quantified using the Kyte and Doolittle hydrophobicity score [47], and virus experimental fitness measured in bulk competition was analyzed (Fig. 3B). Hydrophobicity score in the h-region explains roughly 30% ($R^2 = 0.3$) of the variation we observed in experimental fitness. This correlation is lower than what we observed between experimental replicates, indicating that hydrophobicity alone does not fully explain the experimental fitness effects. The most hydrophobic amino acid at positions in the h-region did not always provide the greatest experimental fitness, further suggesting selection on additional features. While mutations to slightly polar amino acids were compatible with robust experimental fitness at some positions, the most polar (least hydrophobic) amino acids were consistently strongly deleterious in our experiments.

Because net hydrophobicity is known to be important for signal sequence function, experimental fitness measurements of mutations that introduced charged residues were analyzed and compared to uncharged mutations. Charged residues were highly depleted in recovered virus despite their high abundance in the plasmid library (Fig. 3C). This observation indicates that charged amino acids in the hydrophobic region are strongly deleterious, consistent with the sparsity of charged amino acids in this region among sequenced isolates [34].

Previous studies have indicated that small amino acids are strongly preferred at the −1 and −3 sites relative to the peptidase cleavage site [31]. We assessed how amino acid size at these positions influenced experimental fitness effects. The molecular weight of residues at the −3 and −1 positions relative to the peptidase cleavage site had a significant correlation with experimental fitness in bulk competition (Fig. 3D, $R^2 = 0.36$, $P = 0.02$). We note that the strength of this correlation is driven by one large amino acid mutation, A17E, that is strongly deleterious. If A17E is removed, the correlation is not statistically significant (Supplementary Fig. 4). Two other large amino acid mutations, T15K and T15R, were severely depleted in recovered virus; although precise quantification of their experimental fitness effects is hindered, the depletion suggests that these substitutions are also strongly deleterious. While our observations

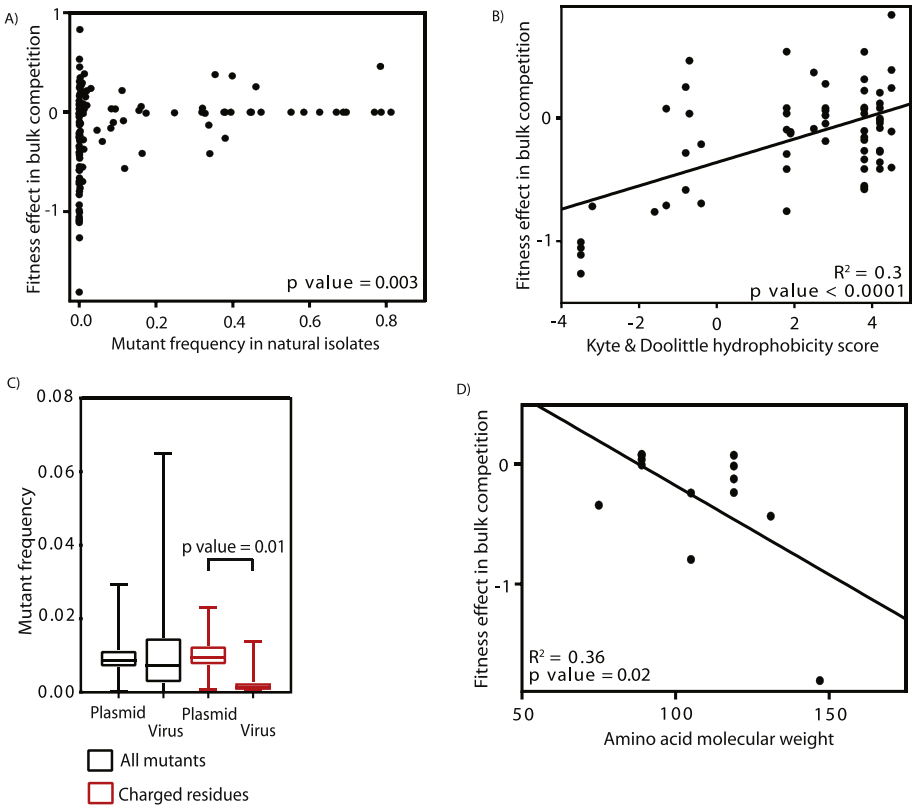

Fig. 3. Relating experimental fitness with conservation patterns and biochemical properties. (A) Comparing experimental fitness measurements to mutant frequency in sequenced circulating viruses. Reported $P = 0.003$ from linear regression analysis. (B) Correlation between experimental fitness measurements in bulk competition and the Kyte–Doolittle hydrophobicity scores for residues in the h-region ($P < 0.0001$ calculated by linear regression). (C) Box plot showing the frequency of 119 individual mutants in the mutant plasmid library before virus recovery and in recovered virus. Mutant frequency for all variants is shown in black and the frequency of charged residues is shown in red. $P = 0.01$ calculated by two-sample *t* test. (D) Correlation between experimental fitness effect and amino acid molecular weight for residues at positions 15 and 17, which correspond to the −3 and −1 positions relative to the peptidase cleavage site ($R^2 = 0.36$, $P = 0.02$ from linear regression).

are consistent with the −1, −3 rule of preference for small amino acids at these sites relative to the peptidase cut site [30,31], the molecular weight of amino acids near the peptidase cut site only explains part of the observed variance in experimental fitness effects. Other factors including structural context may also impact the experimental fitness effects of mutants in this region.

Experimental fitness effects of synonymous substitutions

Previous studies have found evidence of biased codon usage indicating that synonymous substitutions may be under selection for translation efficiency and accuracy [21,48–52]. In our experiments, synonymous substitutions in the signal sequence exhibit a wide distribution of experimental fitness effects (Fig. 2), with fitness ranging from wild-type-like to strongly deleterious. To investigate how selection on synonymous mutations may change with distance from the beginning of the gene, we used identical EMPIRIC procedures to quantify the experimental fitness effects of mutations at amino acids 32–41 of HA1. This 32–41 region is located on the HA stem proximal to the HA2 region involved in membrane fusion and is generally conserved in group 1 and 2 HA (Supplementary Fig. 5). HA is composed of different structural and functional domains that are under varying selection constraints. While the amino acid 32–41 region of HA does not fully represent the rest of HA, we use it here only as a comparison for measuring experimental fitness effects.

In the environmental condition of our experiment, synonymous substitutions had comparable mean fitness effects in the signal sequence and the 32–41 amino acid region, while non-synonymous substitutions were more deleterious in the 32–41 amino acid region (Fig. 4A). While the mean fitness effects of synonymous mutations were similar in the different regions, the variance of fitness effects for synonymous substitutions in the signal sequence compared

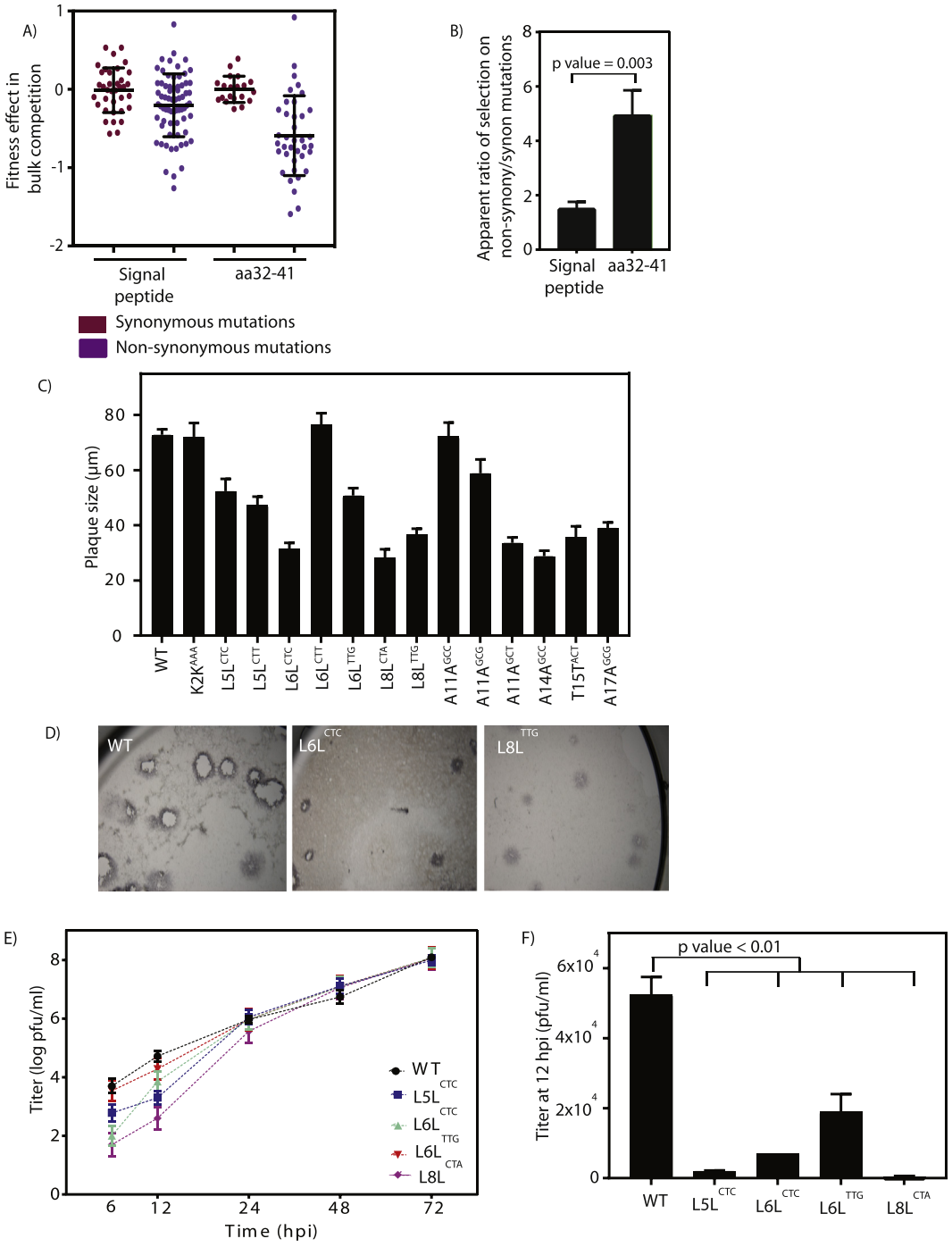

Fig. 4. Synonymous mutations impact experimental fitness in the signal sequence of HA. (A) Beeswarm plot comparing the experimental fitness effects of synonymous and non-synonymous mutations in the signal sequence and in a region of HA encoding amino acids 32–41. Bars represent mean and standard deviation. (B) The apparent ratio of selection acting on non-synonymous/synonymous mutations calculated as the ratio between mean absolute experimental fitness effects of non-synonymous and synonymous mutations. Error bars are based on propagation of standard error of the means and reported *P* value from Student's *t* test. (C) Average plaque size measured for a panel of individually cloned synonymous mutations. Error bars represent standard error of the mean. (D) Images for plaques of wild-type virus and two synonymous substitutions ($L6L^{CTC}$ and $L8L^{TTG}$) that were found to have experimental fitness defects. (E) Infection kinetics of wild-type IAV and four variants carrying synonymous mutations in the signal sequence of HA. Titer was determined based on replicate plaque assays ($n = 2$), and the error bars represent the difference between the two assays. (F) Comparison of viral titer on a linear scale at $t = 12$ h during the infection time course. *P* values from multiple comparison analysis test.

was significantly greater than in the 32–41 amino acid region ($P = 0.02$). Using the ratio in mean absolute experimental fitness effect between non-synonymous and synonymous mutations as an estimate of the apparent ratio of selection on the protein sequence relative to the nucleic acid sequence, the signal sequence's ratio was significantly lower than in the amino acid 32–41 region (Fig. 4B, $P = 0.003$). This difference in experimental fitness ratio indicates the presence of stronger selection on the protein sequence in the amino acid 32–41 region. Our observations are consistent with a combination of stronger selection on the nucleic acid sequence, but weaker selection on the amino acid sequence of the signal sequence relative to a nearby region of HA.

Because of variability in our bulk competitions, we pursued independent experiments on individual clones to further investigate the fitness effects of synonymous mutations in the signal sequence of HA. Consistent with our observations of fitness defects for many synonymous mutations from bulk competition, we also observed that many of the synonymous substitutions in our panel of individual mutants exhibited diminished plaque sizes relative to the parental strain (Fig. 4C and D). Both the bulk competitions and the analyses of individual clones indicate that synonymous mutations in the beginning of the HA gene caused measureable defects in experimental fitness.

To further investigate potential differences in the experimental fitness effects of synonymous mutations, infection kinetics experiments were performed on four synonymous substitutions, L5L^{CTC}, L6L^{CTC}, L6L^{TTG}, and L8L^{CTA}. We chose a small set of synonymous mutations to analyze because of the involved nature of the kinetic experiments. Three of these synonymous mutants exhibited growth defects in bulk competition (L6L^{CTC}, L6L^{TTG}, and L8L^{CTA}), while L5L^{CTC} did not exhibit a growth defect in bulk competition (Fig. 2). As individual mutants, each of these four mutants exhibited right-shifted growth kinetics relative to wild-type virus (Fig. 4E and F).

It is challenging to directly compare the infection kinetics results with the bulk competition data and to reconcile differences in results because of distinctions in the inherent experimental design. In the EMPIRIC experiments, multiple viral mutants are competing against each other, while the infection kinetics experiments follow the expansion of a single viral clone. The host cells available for infection over time likely differ in each of the two types of experiments. To try to account for these distinctions, we compared EMPIRIC selection coefficients with either the viral titer at 12 h or the change in titer between 6 and 24 h from the kinetic experiments and observed poor correlations in both cases (Supplementary Fig. 6A and B). While we were not able to develop a simple and effective comparison between the results of the bulk competitions and the kinetics analyses, we noted distinctions in the kinetic behavior of specific mutants. For example, the L5L^{CTC} mutant that did not show a growth defect in the bulk competition had a low titer at 12 h but exhibited the most rapid expansion between 6 and 24 h in the kinetic experiments.

The discrepancy between the infection kinetics assay and bulk competition (Supplementary Fig. 6B) could also be due to inherent challenges in measuring experimental fitness effects with precision using these techniques. Consistent with our analyses of the reproducibility of experimental fitness measurements, it appears that our experiments can distinguish overall trends but are not precise enough to discern the effect of individual mutations with great confidence. For this reason, we have taken multiple approaches to investigate experimental fitness effects, as each approach has caveats. The slowed infection kinetics of three individual mutations that were deleterious in the bulk competitions lends evidence that synonymous mutations in the signal sequence can cause strong defects in viral expansion.

To further investigate the expansion kinetics results, a phenomenological 5-parameter model of IAV growth kinetics [53] was used to fit the experimental data (see Materials and Methods and Supplementary Information for details). Sensitivity analysis of the model showed that only two parameters (β and τ_E) could be reliably estimated (Supplementary Fig. 7). The β parameter describes the rate of infection of target cells. The τ_E parameter represents the average duration of an eclipse phase, which is the time cells are infected before virus is produced (see Materials and Methods and Supplementary Figs. 7 and 8 for details). Consistent with the infection kinetics data, L5L^{CTC}, L6L^{CTC}, and L8L^{CTA} demonstrated a delayed eclipse phase along with an increased infection rate. The L6L^{TTG} mutant demonstrated WT-like kinetics of viral replication in this model (Table 1).

Signal sequence mutations may affect vRNA structure

It has been reported previously that the folding free energy of a 42-nt-long region beginning four nucleotides before the start codon of bacterial, viral, and eukaryotic mRNA affects translation initiation efficiency [54–57]. In addition, in the influenza virus, packaging signals in HA segment vRNA span the 5′ untranslated region (UTR) and extend for up to 45 nt in the HA ORF [58,59]. Therefore, synonymous mutations in the 5′-end region of the HA segment, which includes part of the signal sequence, can be expected to alter HA expression efficiency as well as vRNA packaging. To investigate the relationship between possible defects in vRNA and experimental fitness, the folding free energy, ΔG, of secondary structures of both mRNA and vRNA in the 5′ region (sense) for all synonymous HA variants was estimated using the ViennaRNA software [60] (Supplementary Fig. 9A). Local secondary structure

Table 1. Viral growth kinetics parameters estimated for wild-type and four synonymous HA mutants, described by rate of infection of target cells β, and duration of eclipse phase τ_E

	$\text{Log}_{10}\ \beta$, [95% BCI]	τ_E, [95% BCI]
WT	−6.28, [−6.50 to 6.05]	6.16, [5.82 to 7.64]
$L5L^{CTC}$	−4.97, [−5.26 to 4.69]	20.8, [19.0 to 22.3]
$L6L^{CTC}$	−4.85, [−5.10 to 4.58]	18.8, [17.1 to 20.0]
$L6L^{TTG}$	−6.16, [−6.40 to 5.87]	8.82, [7.52 to 10.9]
$L8L^{CTA}$	−5.11, [−5.30 to 4.92]	23.1, [22.0 to 24.0]

Median values and 95% bootstrap confidence intervals (BCI) are shown.

of RNA was probed by restricting the sequence lengths for folding to fixed windows of several tested sizes: 24, 36, 42, 54, and 62 nt. Analysis of the correlation between experimental fitness and Δ*G* for different window sizes and positions (Supplementary Fig. 9B) showed a statistically significant correlation between experimental fitness and Δ*G* for a window size of 24 nt beginning at nt 6 in vRNA (permutation test, see Materials and Methods).

No significant correlation between the energy Δ*G* of mRNA structure in the HA signal sequence and experimental fitness was observed. Analysis of predicted vRNA structure at positions 6–30 revealed a small hairpin at positions 7–23 (Fig. 5A). The energy Δ*G* of mutant vRNA in this region has a bimodal distribution, with peaks around −4 kcal/mol (stable hairpin) and −1 kcal/mol (hairpin unfolded) (Supplementary Fig. 10). Mutant viruses predicted to be unstructured in this region exhibited a higher mean experimental fitness compared to viruses predicted to have an intact hairpin (Fig. 5B). While this trend was not statistically significant ($P = 0.29$), it suggests that less structured vRNA in this window may be beneficial for virus replication under our experimental conditions.

To check for similar findings in sequenced HA of circulating IAV, we evaluated HA sequences from the Influenza Research Database [45,46]. Targeted RNA folding for the 30-nt region beginning at position 6 at the 5′ end of the HA segment in the vRNA was performed and the occurrence of different RNA secondary structures in the region was analyzed for strains isolated from either human or swine hosts (Fig. 5C). A two-sample Kolmogorov–Smirnov test indicates that the distributions of RNA folding energy for human-derived and swine-derived IAV are different ($P < 0.01$). While viruses from both human and swine are bimodal

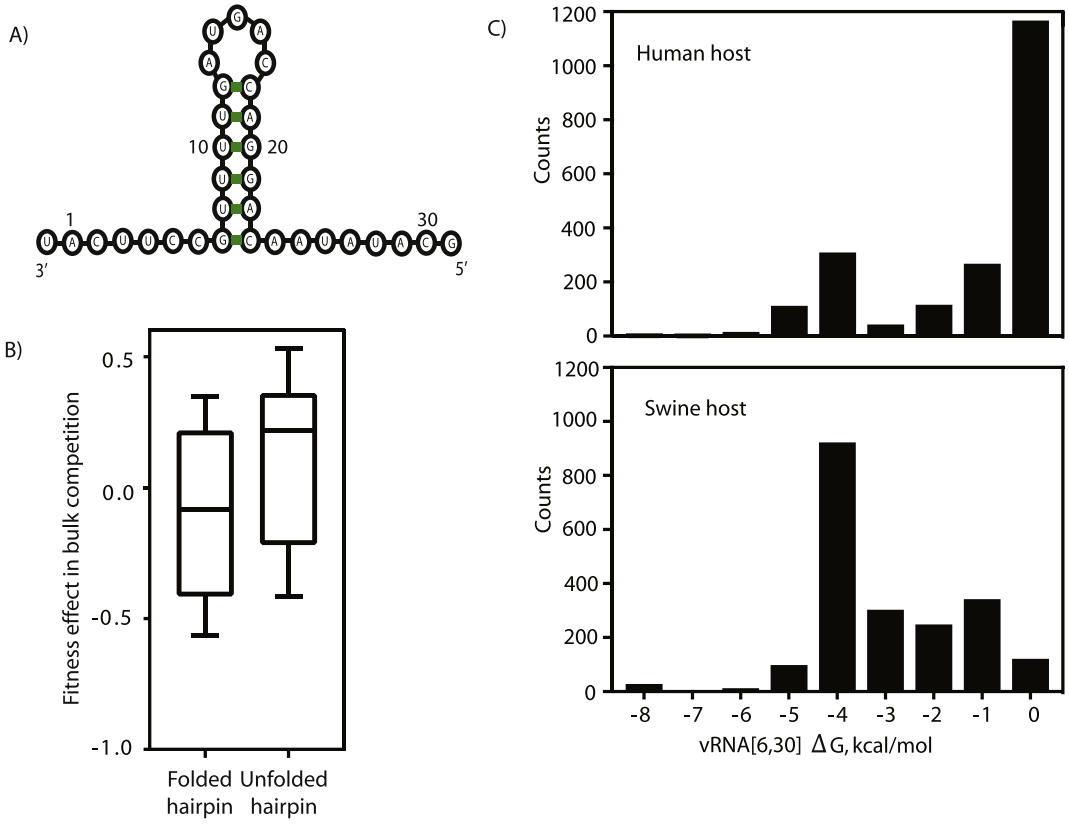

Fig. 5. Probing effects of mutations on vRNA. (A) Illustration of the hairpin at nucleotide position 7–23 in the HA vRNA. (B) Comparing experimental fitness of mutations that disrupt the hairpin at nt 7–23 to mutations that leave the hairpin intact ($P = 0.29$ calculated using two-sample *t* test). (C) Frequency distribution of calculated Δ*G* for sequenced IAV isolated from either human (top panel) [73] or swine (bottom panel) hosts.

with peaks at ≈ –0.5 kcal/mol and ≈ –4 kcal/mol, the human viruses have a higher peak for the unfolded state compared to the swine viruses. These observations are consistent with selection favoring the folded RNA state in swine hosts relative to human hosts.

To investigate relationships between RNA structure and function, we analyzed the structures formed by the stable RNA state. Nearly 90% of the structures with the energy of ≈ – 4 kcal/mol have similar hairpin folds centered around a 5-nt-long loop at nt 13 to 17 (or 3-nt loop at positions 14 to 16), such as shown in Fig. 5A. We extracted the sequences of the 5-nt loop from sequences of circulating strains that have the stable hairpin, and compared sequence occurrence with experimental fitness values in our bulk competition experiment. The sequence variants present in circulating HA all have WT-like experimental fitness, while assayed sequences that are not found in the database all have an experimental fitness defect (Supplementary Table 2). This observation indicates that EMPIRIC fitness measurements for this region are in agreement with trends seen in sequenced natural isolates. In addition, it implies that the hairpin/loop may be important for viral fitness under specific conditions (e.g., when infecting a swine host), similar to the previously reported packaging interactions between vRNA hairpins of IAV segments 2 and 8 that encode the RNA polymerase subunit PB1 and nonstructural protein, respectively [61].

To assess the effects of signal sequence mutations on vRNA packaging, RT-qPCR was used to measure HA vRNA abundance for a panel of individual cell-free mutant viruses. HA vRNA levels were normalized to NA vRNA and used to quantify the packaging efficiency. The mutants analyzed showed reduced levels of HA vRNA relative to wild-type. However, no strong correlation was found between the efficiency of packaging HA vRNA into particles and EMPIRIC fitness (Supplementary Fig. 11). This lack of correlation reveals that viral expansion under our experimental conditions may not be tightly coupled to packaging efficiency, or that we are unable to capture that relationship at our current experimental precision.

Signal sequence mutations influence HA protein abundance in viral particles

Because the HA signal sequence is cleaved and excluded from the mature protein, mutations in this region are likely to affect HA synthesis efficiency and expression rather than the ultimate function of the HA protein. ELISA was used to quantify HA protein expression for a panel of individual mutations to examine how signal sequence mutations impact HA protein abundance in cell-free viral particles (Fig. 6A, see Materials and Methods). All of the studied mutations precede the peptidase cleavage site located at amino acid 17, so mutations would result in identical mature HA protein. Optimal virus experimental fitness, measured in bulk EMPIRIC competition, was supported by a narrow range of HA expression. Either too much or too little HA was associated with experimental fitness defects (Fig. 6B). The relationship between HA abundance and viral experimental fitness was nonlinear with $R^2 = 0.4$ and $P = 0.007$, and appears to be under stabilizing selection. Of note, we excluded L8L^{TTG} from the fit because its expression level is far removed from all other points. Nevertheless, the relationship is still statistically significant even if L8L^{TTG} is included ($P < 0.05$).

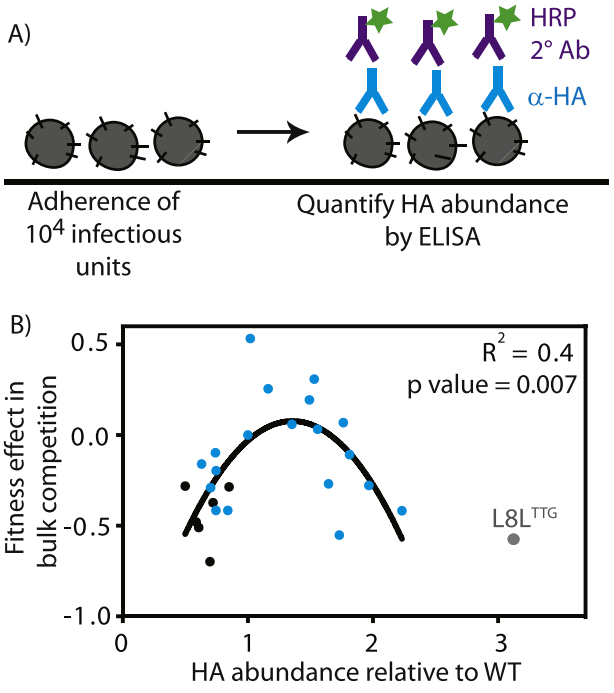

Fig. 6. Relationship between experimental fitness and relative HA protein in cell-free viruses. (A) Experimental setup: viral particles were adhered to wells in plastic plates and HA abundance was quantified by ELISA. (B) Comparison of mutant impact on experimental fitness and HA abundance. Synonymous mutations are shown as blue dots and amino acid changing mutations as black dots. The synonymous mutation of L8L^{TTG}, shown in gray, is excluded from the fit. $R^2 = 0.4$ and P value = 0.007 calculated by quadratic regression.

Changes in HA synthesis efficiency partly explain experimental fitness effects

To measure effects of signal sequence mutations on HA synthesis in the absence of other vRNA segments, HA was transiently expressed in 293T cells and HA synthesis quantified by flow cytometry (Fig. 7A, see Materials and Methods for details). Non-synonymous substitutions had lower HA expression levels compared to the wild-type, while synonymous mutations either reduced or increased

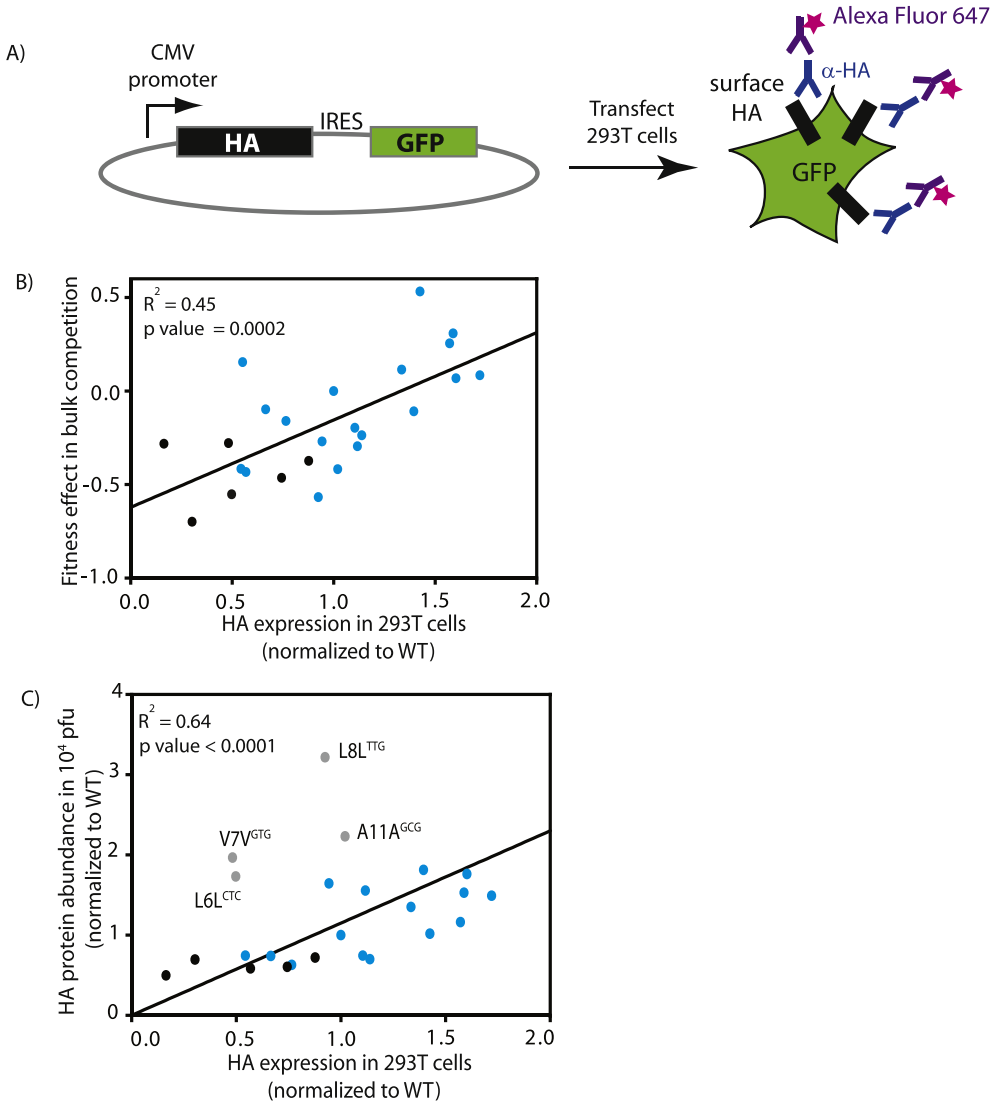

Fig. 7. Relationship between experimental fitness and HA protein expression. (A) HA was expressed in 293T cells and HA protein expression was measured by flow cytometry. (B) Comparing experimental fitness to HA expression in 293T cells. (C) Comparison of HA protein abundance in cell free viruses to HA expression in 293T cells. The points in gray exhibit the largest discrepancy between observed abundance in viruses and 293T cells and are omitted from the fit. In panels B and C, synonymous mutations are shown as blue dots and amino acid changing mutations as black dots. Reported *P* values were calculated by linear regression analysis.

HA expression. In addition, a statistically significant correlation was found between HA abundance and experimental fitness effects measured in bulk (Fig. 7B, $P = 0.0002$).

To further discern whether experimental fitness effects arise from inherent defects in HA synthesis or from HA expression on virus particles, HA expression in 293T cells and HA abundance in viruses were compared. A significant correlation was found between HA synthesis and HA expression in virus particles (Fig. 7C, $P < 0.0001$). However, four synonymous mutations ($L6L^{CTC}$, $V7V^{GTG}$ $L8L^{TTG}$, and $A11A^{GCG}$) deviated far from the trend. The frequency of these variants in sequenced non-pandemic H1N1 isolates was analyzed, and despite being synonymous substitutions, these variants are rare in circulating viruses. This observation suggests the importance of the nucleotide sequence in this region for virus function. When compared to wild-type, these variants had wild-type-like or reduced levels of HA expression in 293T cells, but had up to three times as much HA per plaque-forming unit (PFU). These observations are consistent with an increased ratio of non-infectious to infectious particles, though further confirmatory studies would be required.

Multiple linear regression analysis of experimental fitness effects

To identify measurable factors affecting bulk experimental fitness of IAV mutants, multiple linear regression analysis was performed. As discussed above, the viral experimental fitness determined in the EMPIRIC experiments is correlated with various metrics of HA *in vitro*, using either a linear or quadratic model (see Figs. 6B and 7B correspondingly). To assess whether additional factors explain the observed variance of experimental fitness, a linear regression model was used with HA synthesis in mammalian cells, HA protein abundance in cell-free viral particles, and mRNA folding energy ΔG as explanatory variables. The combination of HA synthesis in mammalian cells and HA protein abundance in cell-free viral particles improved the prediction of experimental fitness (R^2 = 0.68; ANOVA comparison of the models yielding P = 0.014). The best prediction was achieved by a linear model combining all three factors, HA protein synthesis, HA abundance in released virus, and mRNA ΔG, yielding R^2 = 0.752 and P = 0.053 for ANOVA comparison of the two- versus three-factor models. The complete set of P values of nested linear regression models is presented in Supplementary Table 3. The folding energy of mRNA was defined as the free energy ΔG of the viral mRNA segment consisting of nt −4 to 37 relative to the HA start codon, as described in Ref. [21]. Adding a categorical variable representing the DNA library comprising a particular mutant did not improve the fit, suggesting that the measurements were not characterized by batch effect. Likewise, inclusion of additional metrics such as HA vRNA abundance normalized to NA vRNA, vRNA ΔG, and position of the mutation did not improve the predictive power of the multiple linear regression model for experimental fitness (data not shown). However, the effectiveness of additional predictors in the multiple regression was also limited by the overall number of available data points.

Discussion

Our bulk competitions indicate increased selection on synonymous mutations in the signal sequence relative to a nearby region of the HA gene. In addition, many individual synonymous mutations in this region caused defects in viral infectivity when analyzed in isolation. The finding that synonymous mutations in the signal sequence affect viral fitness agrees with previous studies that showed that synonymous mutations in the beginning of genes can have large impacts on translation efficiency [21,48,55]. Biochemical analyses of a panel of synonymous mutations revealed that variants with reduced HA expression had lower experimental fitness, suggesting selection on the efficiency of protein synthesis. Signal sequence mutations also appeared to impact HA protein levels in viruses and during expression in the absence of other IAV genes in 293T cells. Our results are in agreement with selection against reduced HA expression that may in turn impact the ratio of infectious to non-infectious particles leading to apparent stabilizing selection on HA abundance per infectious particle. However, the impacts of synonymous substitutions on translation efficiency appear to only partly explain the experimental fitness effects that we measured.

Changes in vRNA structure also explain some of the experimental fitness effects of synonymous substitutions. We found that change in RNA folding energy in a 24-nt window starting at position 6 is correlated with experimental fitness. Our analyses also indicate that the hairpin loop at position 7–23 may have structural significance as mutations that disrupted this hairpin loop had higher average experimental fitness in our experiments compared to mutations that left the hairpin loop intact. Intriguingly, clustering of HA sequences from circulating IAV based on the existence of the hairpin loop revealed correlations with the host (swine versus human), in line with earlier reports of lineage specificity of RNA structure in IAV segments [62–64]. This variability in structural features across different IAV strains and lineages may be responsible for the lack of conserved secondary structures in IAV vRNA, as proposed by previous genome wide computational analysis [65].

The panel of individual mutations analyzed experimentally caused a reduction in the levels of HA vRNA in virus particles compared to wild-type. This observation further supports the importance of the 5′ UTR in packaging of functional IAV [58,59]. Features beyond hairpin or unfolded conformations in this region may also be important for experimental fitness. While future studies will be required to understand these details further, we note the presence of binding interactions similar to previously reported hairpin interactions between vRNA of segments PB1 and non-structural protein [61]. While the correlation between mRNA folding free energy and experimental fitness was not statistically significant, single-nucleotide synonymous mutations in the signal sequence may introduce subtle structural changes that are poorly resolved by mRNA folding models and also poorly captured by a thermodynamic metric such as ΔG.

The effects of non-synonymous mutations on virus experimental fitness are generally consistent with the previously established biochemical properties of the signal sequence. Our results support the importance of hydrophobicity in the signal sequence [27,34], and the importance of small amino acids near the protease cut site [30,31]. Amino acid changes that introduce charged residues in the hydrophobic region or bulky residues near the protease cut site showed deleterious effects, indicative of structural and functional constraints on the amino acid sequence to ensure binding by the SRP. However, when compared to a region outside the signal sequence (amino acids 32–41), non-

synonymous mutations have higher mean absolute experimental fitness effect, suggesting that selection at the amino acid level is weaker in the signal sequence relative to this nearby region. This pattern agrees with previous reports that the amino acid sequence of the signal peptide of secretory proteins evolves up to five times faster than other regions [66,67].

The biochemical underpinnings of selection on the signal sequence appear complex. Interestingly, we find measurable experimental fitness effects of synonymous mutations in the nucleotides encoding the HA signal peptide—a finding supported by the analysis of sequences from natural isolates as well. Our results demonstrate that selection acting on the signal sequence is due to multiple biochemical pressures, including constraints on RNA secondary structure, RNA packaging efficiency, protein synthesis, and trafficking. This type of selection produces a pattern of sequence evolution that is not readily identified by simple models of conservation. Specifically, our findings indicate that patterns of conservation in natural populations only reveal a fraction, perhaps only the tip of the iceberg, of the impacts that synonymous mutations can have on fitness. This result holds important implications for commonly used *dN*/*dS* (i.e., the ratio of non-synonymous to synonymous divergence) based statistics used for analyzing the selective pressures facing populations. Namely, the fundamental assumption of these approaches is that synonymous mutations are neutral with respect to fitness, and thus, *dS* represents the rate of fixation owing to genetic drift alone (which is simply equal to the mutation rate). *dS* is then compared with *dN*, where a ratio larger than 1 is interpreted as evidence of positive selection on non-synonymous mutations. However, if synonymous mutations are indeed non-neutral as these results suggest, the rate of fixation of synonymous sites is no longer an appropriate proxy for neutrality, potentially resulting in *dN*/*dS* ratios that are greatly misleading. Of note, the use of global *dS* estimates can be utilized to reduce the influence of a limited number of sites under strong synonymous mutation in the identification of sites under positive selection [68]. Thus, future experimental studies aimed at better defining the impacts of synonymous mutations on pathogen fitness and correlations with RNA structure will be of interest, in particular in pathogens with error-prone polymerases.

Materials and Methods

Engineering plasmid libraries

Systematic single-nucleotide mutant libraries of the signal sequence of HA were generated in reverse engineered IAV strain A/WSN/33 as previously described [41,44]. Briefly, nucleotides encoding amino acids 1 to 1650 of HA were transferred to the pRNDM plasmid [41]. PCR was used to introduce inverted BsaI restriction sites. Vector digestion generated directional sticky ends onto which annealed, double-stranded DNA oligonucleotides were ligated. In addition to single-nucleotide mutants, we included additional stop codons at an elevated abundance as an internal control in anticipation of potentially strong selection that might otherwise be difficult to detect. Mutant libraries were transferred to the pHW2000 plasmid [69] using an engineered destination vector and Gibson Assembly (New England BioLabs) as described [70].

Cell culture and viral recovery

MDCK and 293T cells were used for viral recovery and passaging as previously described [44]. Viruses were generated using the pHW2000-based 8-plasmid system using plasmids and protocol developed by Hoffmann and colleagues [69]. Briefly, 1 μg of HA plasmid library and 1 μg of each plasmid encoding the rest of the A/WSN/33 genome were transfected in a co-culture of MDCK and 293T cells using TransIT LT1 reagent (Mirus, Madison, WI). Cell growth media was changed to serum-free media 6 h post-transfection. L-1-tosylamido-2-phenylethyl chloromethyl ketone trypsin (Sigma-Aldrich, St. Louis, MO) was added to a final concentration of 0.5 μg/mL 30 h post-transfection. At 72 h post-transfection, virus-containing supernatant was harvested, centrifuged at 300*g* for 15 min, and stored at −80 °C. Transfections were performed in duplicate and pooled prior to further passaging or analyses. Viral titer of the recovered viruses was quantified by plaque assay as previously described [44]. Briefly, 10-fold serial dilutions of virus-containing supernatant were added to confluent MDCK cells and allowed to bind for 1 h at 37 °C. Unbound viruses were washed off with phosphate-buffered saline (PBS) and infected cells were overlaid with 0.5% agar in DMEM and incubated at 37 °C. Plaques were immunostained with antibody against IAV NP and visually counted 48 h after infection.

Bulk competition experiments

Competition experiments were conducted as previously described [44]. Briefly, MDCK cells were plated at 10^6 cells per well in 6-well plates the day before infection. Cells were then washed with PBS followed by cDMEM Dulbecco's modified Eagle's medium, 100 U/mL penicillin, 100 μg/mL streptomycin, and 7.5% bovine serum albumin. Recovered viruses were diluted in influenza virus growth medium (cDMEM/BSA with 1 μg/mL L-1-tosylamido-2-phenylethyl chloromethyl ketone trypsin) and infections were performed in triplicates at a multiplicity of infection of 0.01. Cells were incubated with virus for 1 h at 37 °C. Unbound virus was washed off twice with PBS. Infected cells were grown in influenza virus growth

media at 37 °C. Virus-containing supernatant was collected 48 h post-infection and stored at −80 °C.

Quantification of mutant abundance

Mutant abundance was quantified as previously described [44]. Briefly, virus-containing supernatant was treated with DNase (New England BioLabs, Ipswich, MA) for 1 h at 37 °C for degradation of plasmid DNA. Influenza vRNA was extracted using the QIAamp Viral RNA minikit (Qiagen, Germantown, MD). cDNA was generated using HA-specific RT primers and Superscript III (Life Technologies, Beverly, MA). Sample processing for sequencing was performed as previously described [41]. Briefly, PCR was used to introduce cut sites for ligation of sample-specific adapters. Adapter-ligated samples were quantified, pooled, and submitted for analysis of mutant abundance by Illumina single read 36-bp sequencing on a Genome Analyzer II. Reads with greater than 99.5% confidence were analyzed. A wild-type sequencing control was included in the deep sequencing sample to discern the sequencing error rate. We observed a per-base sequencing error rate of 0.0004. All mutations that we analyzed were present above this frequency before selection. In addition, we implemented a heuristic cut-off of 100 reads per mutant after selection (about a frequency of 0.0001). Of note, as most sequencing errors generate apparent double mutations that are filtered out of our data, the expected sequencing noise in our data is about 10-fold lower than the sequencing error rate. Fitness effects were calculated as the change in relative mutant abundance before and after selection:

$$\log_2\left(\frac{\text{mutant}}{\text{WT}}\right)_{\text{after selection}} - \log_2\left(\frac{\text{mutant}}{\text{WT}}\right)_{\text{before selection}}$$

Analyzing changes in mutant frequency between the plasmid library and recovered viral population provided an estimate of selection during viral recovery. Due to complications with sequencing of the P1 sample for positions 2–11, we utilized changes between plasmid and P0 to estimate fitness effects in this region.

Growth analyses of individual mutants

Twenty-two individual mutants (Supplementary Table 1) were cloned and recovered as described for mutant libraries. The viral titer for each variant was determined by plaque assay. Images of plaques were acquired using a Nikon SMZ1500 microscope. For each mutant, we used the NIS Elements-BR Analysis program to measure the diameter of 10–20 randomly selected plaques. The average plaque size for each mutant was calculated and utilized as an estimate of the growth rate. Viral infection kinetics for a subset of mutants were determined by infecting MDCK cells at a multiplicity of infection of 0.01, as described earlier, and collecting virus-containing supernatant at 6, 12, 48, and 72 h post-infection. Viral titer analyzed by plaque assay was used to measure viral amplification over time in these experiments.

Investigating the relationship between vRNA/mRNA structural variation and observed experimental fitness effects of synonymous substitutions

Predicted minimum free energy (MFE) of RNA secondary structure was used to test the connection between observed synonymous experimental fitness effects and local structural variation in the mRNA/vRNA of mutated HA segments. MFE was calculated for the synonymous HA mutations in a window of size W, sliding across the first 80 nt (30-nt UTR and 50-nt coding sequence) of the HA mRNA/vRNA sequence. The Pearson correlation coefficient R was used to quantify the relationship between calculated MFE values and measured experimental fitness effects. To assess the statistical significance of this correlation, a random permutation test was used, permuting the energy ΔG and experimental fitness for each window, window size, and position (Supplementary Fig. 9). Regions of mRNA/vRNA demonstrating significant ($P \leq 0.05$) correlation between experimental fitness and MFE were subjected to additional criteria: (1) we set a lower limit of 10 for the number of measured experimental fitness values per folding window, as the number varied with size and starting position of the window, and (2) standard deviation of predicted MFE values per folding window was limited from below ($\sigma \geq 1$), to ensure sufficient structural variation in the window. The biological relevance of regions with significant correlation and both additional criteria satisfied was further investigated. The described analysis was performed independently for mRNA and vRNA sequences, with window sizes W of 24, 36, 42, 54, and 62 nt. As described by Kudla *et al.* [21], we used a zero-based numbering of nucleotides in the mRNA and vRNA. The RNAfold program from the ViennaRNA 2 software package [60] was used for RNA folding with the default settings applied.

Alignment of naturally occurring HA sequences

We downloaded 4452 coding sequences of the HA segment from the NIAID Influenza Research Database [46] using only H1N1 strains and excluding overrepresented pandemic H1N1 from the year 2009. HA sequences of sufficient length (>50 nt) were aligned using TranslatorX [36], which aligns sequences of translated amino acids first and then translates back the amino acids using initial codons. The multiple sequence alignment was further filtered in order to eliminate HA sequences with insertions

and deletions in the first 100 nt near the 5′ end, resulting in 4374 HA sequences.

Analysis of structural variation of mRNA/vRNA in the naturally occurring HA variants

We focused on nt 6 to 30 of the sequences retrieved from the Influenza Research Database, a region that appears biologically relevant according to our experimental and statistical analysis of WSN/33 H1N1 strain. To test if the predicted hairpin was still supported by the naturally evolved sequences, we folded a wider region of vRNA 0–36 nt, while restricting the size of secondary structure to 24 nt.

The RNAfold program from the ViennaRNA 2 software package was used for RNA folding with the maxBPspan parameter set to 24 nt and default settings applied otherwise. All statistical testing and visualization was performed using custom Python scripts, available at https://github.com/sergpolly/EMPIRIC-HA and upon request.

Measuring HA protein abundance in viral particles

Indirect ELISA was used to quantify HA abundance in cell free viral samples. Mutations L9L^{TTG}, L9L^{CTA}, A11A^{GCT}, and V13V^{GTC} were added to the above panel for analysis. Recovered viruses were diluted in coating buffer [3.03 g Na_2CO_3, 6 g $NaHCO_3$, 1 L distilled water (pH 9.6), sterile filtered] and 10^4 PFU per well were incubated in 96-well MaxiSorp plates overnight at 4 °C (Sigma-Aldrich). Unbound virus was removed and wells were blocked with 5% BSA in PBS for 1 h. Primary antibody, IAV H1N1 HA antibody (Genetex, Irvine, CA; Cat. No. GTX 127357), was added and incubated on a rocker at room temperature for 2 h. Wells were washed and incubated with sheep anti-mouse IgG HRP-conjugated secondary antibody (General Electric, UK Limited) for 1 h at room temperature. Wells were washed again and HRP substrate (ThermoFisher Scientific, Waltham, MA) was added to wells and incubated in the dark at room temperature for 30 min. Phosphoric acid of 2 M was added to stop the reaction. Absorbance was measured at 450 nm on a SpectraMax PLUS plate reader (Molecular Devices, Sunnyvale, CA). For each viral stock, three ELISA replicates were performed.

Measuring HA expression in mammalian cells

In order to measure HA expression, the above panel of HA variants was cloned into the JB991 plasmid, kindly provided by J. Bloom (Fred Hutchinson Cancer Research Center, Seattle, WA), that has a GFP reporter, an ampicillin resistance gene, and a CMV promoter. These constructs were transiently expressed in 293T. Cells were washed and re-suspended in isotonic buffer [15 mM MOPS, 145 mM NaCl, 2.7 mM KCl, 4 mM $CaCl_2$ (pH 7.4); 2% heat-inactivated FBS added immediately before use]. HA primary antibody (Genetex; Cat. No. GTX 127357) was added to cells and incubated for 1 h at room temperature. Excess antibody was washed off twice with isotonic buffer. Alexa Fluor 647 goat anti-rabbit IgG secondary antibody (ThermoFisher Scientific) was added and incubated at room temperature for 1 h. Cells were washed, re-suspended in isotonic buffer and analyzed on a BD FACSCalibur. Data analysis was conducted using the FlowJo data analysis software package (Ashland, OR).

Quantification of vRNA levels

To measure effects of signal sequence mutations on vRNA, virus-containing supernatant was treated with DNase for 1 h at 37 °C. Total vRNA was extracted for each mutant variant using the QIAamp Viral RNA minikit (Qiagen). In separate reactions, cDNA was generated using HA- and NA-specific RT primers and SuperScript III (Life Technologies). This cDNA was then used as a template in qPCR performed using SYBR Green and HA- and NA-specific PCR primers (ThermoFisher Scientific) on an Eppendorf MasterCycler RealPlex machine (Eppendorf). Relative HA vRNA abundance was calculated using the efficiency-corrected approach developed by Pfaffl [71] and normalized to NA vRNA.

Modeling of the growth parameters of individual mutants

We used a multi-compartment ordinary differential equation (ODE) model [53,72] to compare viral yield kinetics of silent HA mutants. In the model, initial infection $V_{PFU}(0)$ spreads in a population of N susceptible target cells (T), and infects cells at a rate $\beta V_{PFU}(t)$. Infected cells spend on average τ_E time in an eclipse phase (E) and then, upon switching into infectious mode (I), start producing virus at a rate p for average duration of τ_I. Newly generated virus contributes to the total viral pool $V_{PFU}(t)$ that supports the spread of infection, and also degrades at a rate c. We exclusively considered infectious particles (PFU) in this model, neglecting $V_{RNA}(t)$ (for which we have no data) in the ODE system (Eq. 1 from [53]) and fit the experimental data by minimizing the log-scaled sum of squared residuals (SSR) between experimental measurements and simulated ones for a set of model parameters $\Theta = (\beta, \tau_E, \tau_I, p, c)$:

$$\mathrm{SSR}(\Theta) = \sum_{i=1}^{M_{\mathrm{exp}}} (\log_{10} V_i - \log_{10} V_{\mathrm{PFU}}(t_i, \Theta))^2,$$

where M_{exp} is the number of PFU measurements per mutant and V_i are values of PFU measured at t_i=6, 12, 24, 48, and 72 h. The number of compartments is

fixed in our simulations $n_E = n_I = 20$. The size of initial target cell population was $N = 10^6$, with MOI = 10^{-3}. Supplementary Figure 8 shows an example of the viral growth measurements and the model fit with optimal parameters. Following Paradis *et al.* [53], the odeint function from scipy was used to solve the ODE system and the emcee python module was used to perform Markov Chain Monte Carlo simulations to extract the probability distribution of the model parameters. Initially, we found an optimal set of parameters $\Theta = (\beta, \tau_E, \tau_I, p, c)$ for each mutant and then performed Markov Chain Monte Carlo simulation using 100 walkers for 1000 steps. Upon convergence, data from step 600 and beyond were used to estimate the probability distribution (see Supplementary Fig. 7). In order to assess the difference between estimated distributions, A and B, we used p values derived from a two-sample Z-test,

$$Z = \frac{|\mu_A - \mu_B|}{\sqrt{\sigma_A^2 + \sigma_B^2}},$$

where for a given parameter and mutant, μ is the mean of the distribution and σ is its standard deviation. We used $\log_{10}$ of the β infection rate parameter for placing confidence limits on the estimated model parameters.

Data availability

Custom python scripts are available upon request. The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article.

Acknowledgments

We thank Melanie Trombly and members of the AIRe team for input and comments on the manuscript. This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs, through the Peer Reviewed Medical Research Program (Award No. W81XWH-15-1-0317). Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the Department of Defense.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2018.02.009.

Received 20 October 2017;
Received in revised form 19 January 2018;
Accepted 5 February 2018
Available online 18 February 2018

Keywords:
deep mutational scanning;
selection;
synonymous mutations;
experimental evolution;
influenza A virus

†J.P.W. and D.N.A.B. contributed equally to this work.

Abbreviations used:
HA, hemagglutinin; IAV, influenza A virus; NA, neuraminidase; SRP, signal recognition particle; EMPIRIC, Exceedingly Meticulous and Parallel Investigation of Randomized Individual Codons; UTR, untranslated region; PFU, plaque forming unit; PBS, phosphate-buffered saline; MFE, minimum free energy; ODE, ordinary differential equation.

References

[1] M. Kimura, A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences, J. Mol. Evol. 16 (1980) 111–120.
[2] J.V. Chamary, J.L. Parmley, L.D. Hurst, Hearing silence: non-neutral evolution at synonymous sites in mammals, Nat. Rev. Genet. 7 (2006) 98–108.
[3] T. Zhou, W. Gu, C.O. Wilke, Detecting positive and purifying selection at synonymous sites in yeast and worm, Mol. Biol. Evol. 27 (2010) 1912–1922.
[4] J.B. Plotkin, G. Kudla, Synonymous but not the same: the causes and consequences of codon bias, Nat. Rev. Genet. 12 (2011) 32–42.
[5] S.G. Andersson, C.G. Kurland, Codon preferences in free-living microorganisms, Microbiol. Rev. 54 (1990) 198–210.
[6] L. Duret, tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes, Trends Genet. 16 (2000) 287–289.
[7] T. Ikemura, Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system, J. Mol. Biol. 151 (1981) 389–409.
[8] E.N. Moriyama, J.R. Powell, Codon usage bias and tRNA abundance in *Drosophila*, J. Mol. Evol. 45 (1997) 514–523.
[9] R. Hershberg, D.A. Petrov, Selection on codon bias, Annu. Rev. Genet. 42 (2008) 287–299.
[10] D.A. Drummond, C.O. Wilke, Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution, Cell 134 (2008) 341–352.
[11] A. Acevedo, L. Brodsky, R. Andino, Mutational and fitness landscapes of an RNA virus revealed through population sequencing, Nature 505 (2014) 686–690.
[12] J.I. Boucher, D.N. Bolon, D.S. Tawfik, Quantifying and understanding the fitness effects of protein mutations: laboratory versus nature, Protein Sci. 25 (2016) 1219–1226.
[13] R.T. Hietpas, C. Bank, J.D. Jensen, D.N. Bolon, Shifting fitness landscapes in response to altered environments, Evolution 67 (2013) 3512–3522.
[14] B.P. Roscoe, K.M. Thayer, K.B. Zeldovich, D. Fushman, D.N. Bolon, Analyses of the effects of all ubiquitin point mutants on yeast growth rate, J. Mol. Biol. 425 (2013) 1363–1377.

[15] L. Jiang, P. Mishra, R.T. Hietpas, K.B. Zeldovich, D.N. Bolon, Latent effects of Hsp90 mutants revealed at reduced expression levels, PLoS Genet. 9 (2013), e1003600.
[16] S. Bershtein, J.M. Choi, S. Bhattacharyya, B. Budnik, E. Shakhnovich, Systems-level response to point mutations in a core metabolic enzyme modulates genotype–phenotype relationship, Cell Rep. 11 (2015) 645–656.
[17] I. Mayrose, A. Stern, E.O. Burdelova, Y. Sabo, N. Laham-Karam, R. Zamostiano, et al., Synonymous site conservation in the HIV-1 genome, BMC Evol. Biol. 13 (2013) 164.
[18] H. Qin, W.B. Wu, J.M. Comeron, M. Kreitman, W.H. Li, Intragenic spatial patterns of codon usage bias in prokaryotic and eukaryotic genomes, Genetics 168 (2004) 2245–2260.
[19] E. Firnberg, J.W. Labonte, J.J. Gray, M. Ostermeier, A comprehensive, high-resolution map of a gene's fitness landscape, Mol. Biol. Evol. 31 (2014) 1581–1592.
[20] T. Tuller, A. Carmi, K. Vestsigian, S. Navon, Y. Dorfan, J. Zaborske, et al., An evolutionarily conserved mechanism for controlling the efficiency of protein translation, Cell 141 (2010) 344–354.
[21] G. Kudla, A.W. Murray, D. Tollervey, J.B. Plotkin, Coding-sequence determinants of gene expression in *Escherichia coli*, Science 324 (2009) 255–258.
[22] J.K. Tautenberger, D.M. Morens, 1918 Influenza: the mother of all pandemics, CDC Emerg. Infect. Dis. 12 (2006) 15–22.
[23] I.A. Wilson, J.J. Skehel, D.C. Wiley, Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution, Nature 289 (1981) 366–373.
[24] J.N. Varghese, P.M. Colman, Three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 A resolution, J. Mol. Biol. 221 (1991) 473–486.
[25] G. Blobel, D. Sabatini, Dissociation of mammalian polyribosomes into subunits by puromycin, Proc. Natl. Acad. Sci. U. S. A. 68 (1971) 390–394.
[26] G. Blobel, B. Dobberstein, Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma, J. Cell Biol. 67 (1975) 835–851.
[27] G. von Heijne, Signal sequences: the limits of variation, J. Mol. Biol. 184 (1985) 99–105.
[28] P. Walter, G. Blobel, Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum, Proc. Natl. Acad. Sci. U. S. A. 77 (1980) 7112–7116.
[29] R. Gilmore, P. Walter, G. Blobel, Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor, J. Cell Biol. 95 (1982) 470–477.
[30] R.E. Dalbey, G. Von Heijne, Signal peptidases in prokaryotes and eukaryotes—a new protease family, Trends Biochem. Sci. 17 (1992) 474–478.
[31] G. von Heijne, Patterns of amino acids near signal-sequence cleavage sites, Eur. J. Biochem. 133 (1983) 17–21.
[32] G.M. Air, Nucleotide sequence coding for the "signal peptide" and N terminus of the hemagglutinin from an asian (H2N2) strain of influenza virus, Virology 97 (1979) 468–472.
[33] S. Inouye, X. Soberon, T. Franceschini, K. Nakamura, K. Itakura, M. Inouye, Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane, Proc. Natl. Acad. Sci. U. S. A. 79 (1982) 3438–3441.
[34] J.W. Izard, S.L. Rusch, D.A. Kendall, The amino-terminal charge and core region hydrophobicity interdependently contribute to the function of signal sequences, J. Biol. Chem. 271 (1996) 21579–21582.
[35] J.K. Weltman, G. Skowron, G.B. Loriot, Influenza A H5N1 hemagglutinin cleavable signal sequence substitutions, Biochem. Biophys. Res. Commun. 352 (2007) 177–180.
[36] F. Abascal, R. Zardoya, M.J. Telford, TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations, Nucleic Acids Res. 38 (2010) W7–13.
[37] K. Sekikawa, C. Lai, Defects in functional expression of an influenza virus hemagglutinin lacking the signal peptide sequences, Proc. Natl. Acad. Sci. U. S. A. 80 (1983) 3563–3567.
[38] B. Thyagarajan, J.D. Bloom, The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin, elife 3 (2014).
[39] M.B. Doud, J.D. Bloom, Accurate measurement of the effects of all amino-acid mutations on influenza hemagglutinin, Viruses 8 (2016).
[40] N.C. Wu, A.P. Young, L.Q. Al-Mawsawi, C.A. Olson, J. Feng, H. Qi, et al., High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution, Sci. Rep. 4 (2014) 4942.
[41] R.T. Hietpas, B.P. Roscoe, L. Jiang, D.N.A. Bolon, Fitness analyses of all possible point mutations for regions of genes in yeast, Nat. Protoc. 7 (2012) 1382–1396.
[42] R.T. Hietpas, J.D. Jensen, D.N.A. Bolon, Experimental illumination of a fitness landscape, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 7896–7901.
[43] P. Mishra, D.N. Bolon, Designed Hsp90 heterodimers reveal an asymmetric ATPase-driven mechanism in vivo, Mol. Cell 53 (2014) 344–350.
[44] L. Jiang, P. Liu, C. Bank, N. Renzette, K. Prachanronarong, L.S. Yilmaz, et al., A balance between inhibitor binding and substrate processing confers influenza drug resistance, J. Mol. Biol. 428 (2016) 538–553.
[45] R.B. Squires, J. Noronha, V. Hunt, A. Garcia-Sastre, C. Macken, N. Baumgarth, et al., Influenza research database: an integrated bioinformatics resource for influenza research and surveillance, Influenza Other Respir. Viruses 6 (2012) 404–416.
[46] Y. Zhang, B.D. Aevermann, T.K. Anderson, D.F. Burke, G. Dauphin, Z. Gu, et al., Influenza Research Database: an integrated bioinformatics resource for influenza virus research, Nucleic Acids Res. 45 (2017) D466-D74.
[47] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105–132.
[48] P.M. Sharp, K.M. Devine, Codon usage and gene expression level in *Dictyostelium discoideum*: highly expressed genes do 'prefer' optimal codons, Nucleic Acids Res. 17 (1989) 5029–5039.
[49] N. Stoletzki, A. Eyre-Walker, Synonymous codon usage in *Escherichia coli*: selection for translational accuracy, Mol. Biol. Evol. 24 (2007) 374–381.
[50] J.Y. Choi, C.F. Aquadro, Recent and long-term selection across synonymous sites in *Drosophila ananassae*, J. Mol. Evol. 83 (2016) 50–60.
[51] N.D. Singh, V.L. Bauer DuMont, M.J. Hubisz, R. Nielsen, C.F. Aquadro, Patterns of mutation and selection at synonymous sites in *Drosophila*, Mol. Biol. Evol. 24 (2007) 2687–2697.
[52] D.B. Goodman, G.M. Church, S. Kosuri, Causes and effects of N-terminal codon bias in bacterial genes, Science 342 (2013) 475–479.
[53] E.G. Paradis, L.T. Pinilla, B.P. Holder, Y. Abed, G. Boivin, C.A. Beauchemin, Impact of the H275Y and I223V mutations in the neuraminidase of the 2009 pandemic influenza virus

in vitro and evaluating experimental reproducibility, PLoS One 10 (2015), e0126115.
[54] H. Bedouelle, Mutations in the promoter regions of the malEFG and malK-lamB operons of *Escherichia coli* K12, J. Mol. Biol. 170 (1983) 861–882.
[55] P. Shah, Y. Ding, M. Niemczyk, G. Kudla, J.B. Plotkin, Rate-limiting steps in yeast protein translation, Cell 153 (2013) 1589–1601.
[56] W. Gu, T. Zhou, C.O. Wilke, A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes, PLoS Comput. Biol. 6 (2010), e1000664.
[57] T. Zhou, C.O. Wilke, Reduced stability of mRNA secondary structure near the translation-initiation site in dsDNA viruses, BMC Evol. Biol. 11 (2011) 59.
[58] T. Watanabe, S. Watanabe, T. Noda, Y. Fujii, Y. Kawaoka, Exploitation of nucleic acid packaging signals to generate a novel influenza virus-based vector stably expressing two foreign genes, J. Virol. 77 (2003) 10575–10583.
[59] G.A. Marsh, R. Hatami, P. Palese, Specific residues of the influenza A virus hemagglutinin viral RNA are important for efficient packaging into budding virions, J. Virol. 81 (2007) 9727–9736.
[60] R. Lorenz, S.H. Bernhart, C. Honer Zu Siederdissen, H. Tafer, C. Flamm, P.F. Stadler, et al., ViennaRNA Package 2.0, Algorithms Mol. Biol. 6 (2011) 26.
[61] C. Gavazzi, M. Yver, C. Isel, R.P. Smyth, M. Rosa-Calatrava, B. Lina, et al., A functional sequence-specific interaction between influenza A virus genomic RNA segments, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 16604–16609.
[62] A.V. Vasin, A.V. Petrova, V.V. Egorov, M.A. Plotnikova, S.A. Klotchenko, M.N. Karpenko, et al., The influenza A virus NS genome segment displays lineage-specific patterns in predicted RNA secondary structure, BMC Res. Notes 9 (2016) 279.
[63] A.P. Gultyaev, M.I. Spronken, M. Richard, E.J. Schrauwen, R.C. Olsthoorn, R.A. Fouchier, Subtype-specific structural constraints in the evolution of influenza A virus hemagglutinin genes, Sci. Rep. 6 (2016), 38892.
[64] A.P. Gultyaev, A. Tsyganov-Bodounov, M.I. Spronken, S. van der Kooij, R.A. Fouchier, R.C. Olsthoorn, RNA structural constraints in the evolution of the influenza A virus genome NP segment, RNA Biol. 11 (2014) 942–952.
[65] W.N. Moss, S.F. Priore, D.H. Turner, Identification of potential conserved RNA secondary structure throughout influenza A coding regions, RNA 17 (2011) 991–1011.
[66] E.J. Williams, C. Pal, L.D. Hurst, The molecular evolution of signal peptides, Gene 253 (2000) 313–322.
[67] Y.D. Li, Z.Y. Xie, Y.L. Du, Z. Zhou, X.M. Mao, L.X. Lv, et al., The rapid evolution of signal peptides is mainly caused by relaxed selection on non-synonymous and synonymous sites, Gene 436 (2009) 8–11.
[68] S.J. Spielman, C.O. Wilke, The relationship between dN/dS and scaled selection coefficients, Mol. Biol. Evol. 32 (2015) 1097–1108.
[69] E. Hoffmann, S. Krauss, D. Perez, R. Webby, R.G. Webster, Eight-plasmid system for rapid generation of influenza virus vaccines, Vaccine 20 (2002) 3165–3170.
[70] D.G. Gibson, L. Young, R.Y. Chuang, J.C. Venter, C.A. Hutchison III, H.O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases, Nat. Methods 6 (5) (2009) 343.
[71] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR, Nucleic Acids Res. 29 (2001), e45.
[72] B.P. Holder, C.A. Beauchemin, Exploring the effect of biological delays in kinetic models of influenza within a host or cell culture, BMC Public Health 11 (Suppl. 1) (2011) S10.
[73] A. Nogales, S.F. Baker, E. Ortiz-Riaño, S. Dewhurst, D.J. Topham, L. Martínez-Sobrido, Influenza A virus attenuation by codon deoptimization of the NS gene for vaccine development, J. Virol. 88 (2014) 10525–10540.