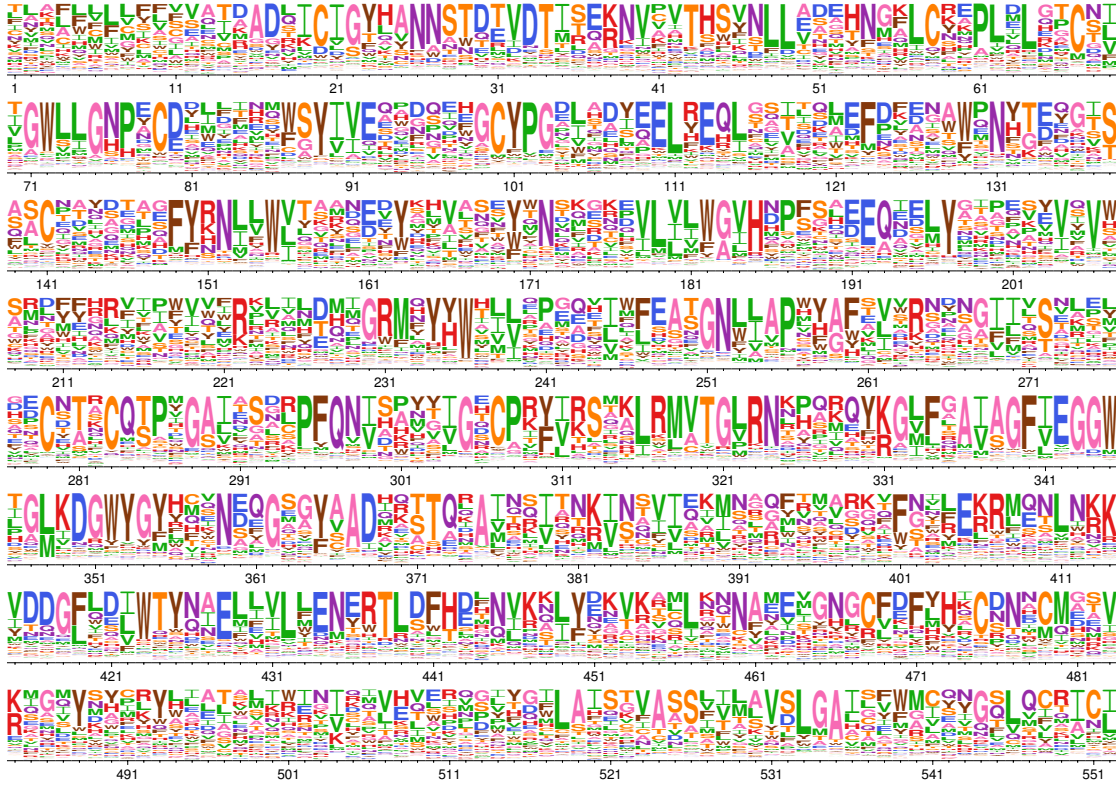
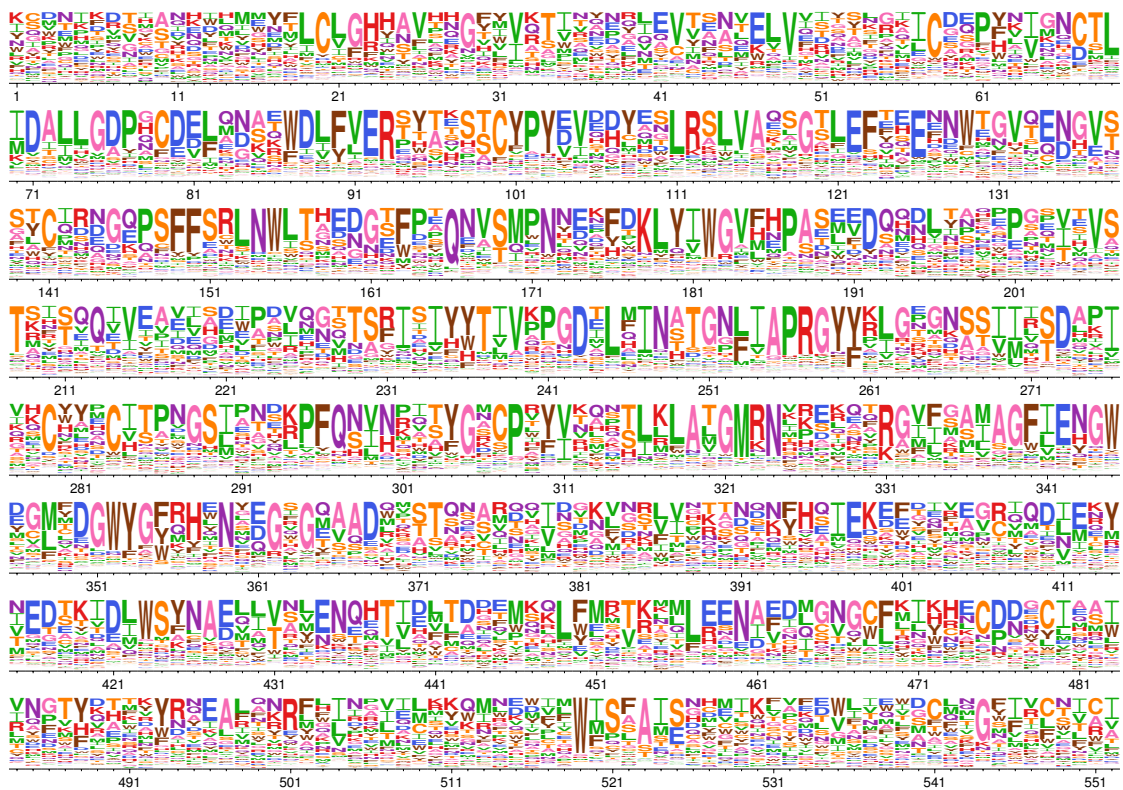


## Supplemental Information



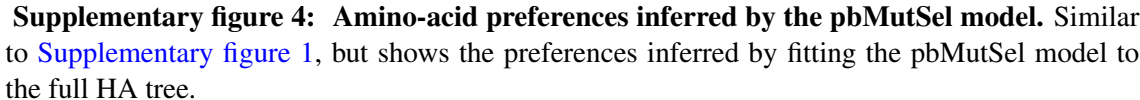
**Supplementary figure 1: H1 HA amino-acid preferences measured by deep mutational scanning.** Each column represents a site in the HA protein, and the height of each letter is proportional to the preference for the amino acid measured by [Doud and Bloom \(2016\)](#) and then re-scaled by the stringency parameter in [Table 1](#). The plot only shows sites that are alignable between the H1 and H3 HAs, and these alignable sites are numbered sequentially starting from 1. The conversion between the numbering scheme in this figure and sequential numbering of the H1 HA reference sequence is in [Supplementary file 1](#).

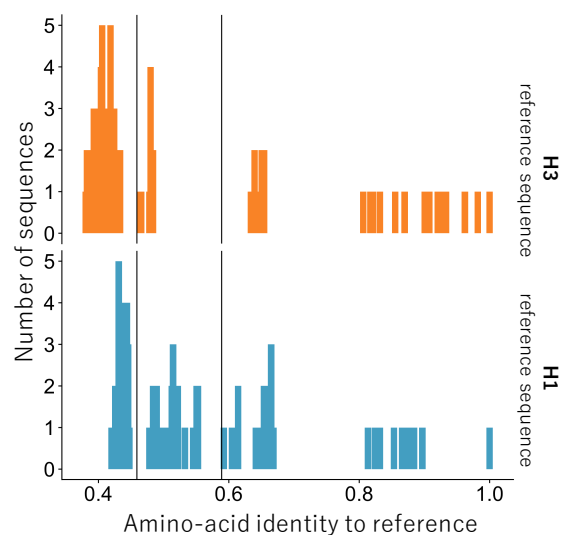


**Supplementary figure 2: H3 HA amino-acid preferences measured by deep mutational scanning.** Similar to [Supplementary figure 1](#) but shows the re-scaled preferences for the H3 HA as measured by [Lee et al. \(2018\)](#).



**Supplementary figure 3: Average of H1 HA and H3 HA amino-acid preferences measured by deep mutational scanning.** Similar to [Supplementary figure 1](#) but shows the re-scaled average of the preferences for the H1 and H3 HAs.





**Supplementary figure 5: Overall divergence for subtrees.** We created two subalignments for each HA used in the deep mutational scanning experiments. The “low divergence” alignments had  $\geq 59\%$  amino-acid identity to either the H1 or H3 reference sequence. The “intermediate divergence” alignments had  $\geq 46\%$  amino-acid identity to the reference sequences.

Model	Tree diameter (average codon substitutions per site)	Percentage of GY94 tree diameter
GY94	12.04	100%
ExpCM(H1)	14.70	122%
ExpCM(H3)	16.28	135%
ExpCM(H1+H3 avg)	19.21	160%
GY94 + $\Gamma\omega$	19.15	159%
ExpCM(H1) + $\Gamma\omega$	24.75	206%
ExpCM(H3) + $\Gamma\omega$	25.03	208%
ExpCM(H1+H3 avg) + $\Gamma\omega$	30.78	256%

**Supplementary table 1: Branch length extension as measured by tree diameter.** We calculated the tree diameter, the distance between the two most divergent tips, for the trees in [Figure 4](#). For each tree, the diameter is reported as a raw value and as a percentage of the GY94 model tree, the smallest of the eight trees.



Model	Parameters
GY94	$\kappa = 3.17, \omega = 0.10,$
	$\phi_{1,A} = 0.32, \phi_{1,C} = 0.14, \phi_{1,G} = 0.28,$
	$\phi_{2,A} = 0.38, \phi_{2,C} = 0.18, \phi_{2,G} = 0.20,$
GY94 + $\Gamma\omega$	$\phi_{3,A} = 0.36, \phi_{3,C} = 0.19, \phi_{3,G} = 0.21$
	$\alpha_\omega = 0.51, \beta_\omega = 3.92, \kappa = 3.49,$
	$\phi_{1,A} = 0.32, \phi_{1,C} = 0.14, \phi_{1,G} = 0.28,$
ExpCM(H1)	$\phi_{2,A} = 0.38, \phi_{2,C} = 0.18, \phi_{2,G} = 0.20,$
	$\phi_{3,A} = 0.36, \phi_{3,C} = 0.19, \phi_{3,G} = 0.21$
	$\beta = 1.56, \kappa = 3.64, \omega = 0.24,$
	$\phi_A = 0.378, \phi_C = 0.17, \phi_G = 0.23$

**Supplementary table 2: Model parameters fit to a low divergence tree.** We fit GY94 models and an ExpCM defined by H1 deep mutational scanning preferences to the “low divergence from H1” tree in [Figure 6](#). We used these model parameters calculate the expected pairwise sequence identity in [Figure 2](#) and simulate the sequences in [Figure 3](#).

**Supplementary file 1:** List of alignable sites between H1 HA and H3 HA. This files provides a conversion between the numbering scheme we use in the paper (sequential numbering of just the alignable sites) to sequential numbering of the H1 HA reference sequence A/Wilson Smith/1933 and the H3 HA reference sequence A/Perth/2009.

**Supplementary file 2:** Amino acid preferences measured by the deep mutational scanning of the H1 HA strain A/WSN/1933 ([Doud and Bloom, 2016](#)). This file only contains measurements for the alignable sites between H1 and H3 HAs. Conversion from this numbering scheme to sequential numbering of A/WSN/1933 is in [Supplementary file 1](#).

**Supplementary file 3:** Amino acid preferences measured by the deep mutational scanning of the H3 HA strain A/Perth/2009 ([Lee et al., 2018](#)). This file only contains measurements for the alignable sites between H1 and H3 HAs. Conversion from this numbering scheme to sequential numbering of A/Perth/2009 is in in [Supplementary file 1](#).

**Supplementary file 4:** The HA sequences for the full HA tree. The sequences in this alignment contain only the alignable sites between H1 and H3 HAs. Conversion from this numbering scheme to sequential numbering of A/Perth/2009 is in in [Supplementary file 1](#).