Egg-adaptation pathway of human influenza H3N2 virus is contingent on natural evolution

Weiwen Liang[1], Timothy J.C. Tan[2], Yiquan Wang[3], Huibin Lv[1], Yuanxin Sun[4,5], Roberto Bruzzone[1,6,7], Chris K.P. Mok[4,5,§], Nicholas C. Wu[2,3,8,9,§]

[1] HKU-Pasteur Research Pole, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

[2] Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[3] Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[4] The Jockey Club School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China

[5] Li Ka Shing Institute of Health Sciences, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China

[6] Department of Cell Biology and Infection, Institut Pasteur, Paris Cedex 75015, France

[7] Centre for Immunology and Infection, Hong Kong Science Park, Hong Kong SAR, China

[8] Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[9] Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

§ Correspondence: kapunmok@cuhk.edu.hk (C.K.P.M.) and nicwu@illinois.edu (N.C.W.)

ABSTRACT

Egg-adaptive mutations in influenza hemagglutinin (HA) often emerge during the production of egg-based seasonal influenza vaccines, which contribute to the largest share in the global influenza vaccine market. While some egg-adaptive mutations have minimal impact on the HA antigenicity (e.g. G186V), others can alter it (e.g. L194P). Here, we show that the preference of egg-adaptation pathway in human H3N2 HA is strain-dependent. In particular, Thr160 and Asn190, which are found in many recent H3N2 strains, restrict the emergence of L194P but not G186V. Our results further suggest that natural amino acid variants at other HA residues also play a role in determining the egg-adaptation pathway. Consistently, recent human H3N2 strains from different clades acquire different mutations during egg passaging. Overall, these results demonstrate that natural mutations in human H3N2 HA can influence the egg-adaption pathway, which has important implications in seed strain selection for egg-based influenza vaccine.

Keywords: Influenza, H3N2, egg-based vaccine, mutation, epistasis, evolution

INTRODUCTION

Vaccination is considered the most effective approach to prevent influenza infection. Seasonal influenza vaccine offers protection against two influenza A subtypes, namely H1N1 and H3N2, as well as influenza B viruses. However, the vaccine effectiveness against H3N2 is typically lower than H1N1 and influenza B [1-3]. Besides vaccine mismatch in certain influenza seasons, the occurrence of egg-adaptive mutations during vaccine production can also lead to the reduction of vaccine effectiveness [4-6]. When grown in eggs, human H3N2 viruses often acquire mutations in the receptor-binding site (RBS) of hemagglutinin (HA) to facilitate the binding to α2,3-linked sialic acid, which is an avian-type receptor [7-9]. Since HA RBS partially overlaps with several major antigenic sites [10], some egg-adaptive mutations can alter the antigenicity of HA [4, 5, 9, 11-13].

G186V and L194P are the two most common egg-adaptive mutations in H3N2 HA [14]. Although both mutations are in the RBS, L194P but not G186V significantly alters the antigenicity of HA [15]. Thus, identifying factors that influence the preference of egg-adaptation pathway can help optimize the effectiveness of egg-based influenza vaccine. Our previous study has demonstrated a strong incompatibility between G186V and L194P, in which their co-occurrence is extremely deleterious to the virus [14]. This epistatic interaction suggests that G186V and L194P represent two mutually exclusive egg-adaptation pathways. Given that epistasis is pervasive in the HA RBS and that HA RBS is constantly evolving [13, 16-19], it is possible that different strains have different preferences of egg-adaptation pathways. However, it remains to be explored whether such differential preference exists.

In this study, sequence analysis and mutagenesis experiments are used to identify amino acid variants on human H3N2 HA that can affect the preference of egg-adaptation pathway. We discover that L194P is incompatible with T160 and N190, both of which are prevalent in recent human H3N2 strains. Consistently, recent human H3N2 strains are incompatible with L194P. Our results further indicate that amino acid variants at other residues also contribute to the

L194P incompatibility in recent human H3N2 strains. Along with the observation that human H3N2 strains from different clades prefer different egg-adaptation pathways, we conclude that the preference of egg-adaptation pathway changes as human H3N2 virus evolves.

RESULTS

Differential egg-adaptation pathways in historical vaccine strains

Our previous study has shown that the two major egg-adaptive mutations G186V and L194P are incompatible [14]. To examine whether different human H3N2 strains have different preferences of egg-adaptation pathways, HA sequences from egg-passaged strains were collected from the Global Initiative for Sharing Avian Influenza Data (GISAID) [20] (**Supplementary Table 1**). Specifically, we focused on egg-passaged strains that were derived from WHO-recommended candidate vaccine viruses for influenza seasons between 2008 and 2020. Egg-passaged strains that were derived from different parental viruses have acquired different egg-adaptive mutations (**Figure 1A**). For example, all egg-passaged strains from A/Victoria/361/2011, A/Texas/50/2021, A/Switzerland/9715293/2013, and A/Hong Kong/2671/2019 have acquired G186V, whereas those from A/Hong Kong/4801/2014, A/Singapore/INFIMH-16-0019/2016 (Sing16) and A/Switzerland/8060/2017 (Switz17) have acquired L194P. This observation is consistent with the hypothesis that the preference of egg-adaptation pathway is strain dependent.

To further test our hypothesis, we measured the replication fitness of G186V and L194P in Sing16, Switz17, and A/Kansas/14/2017 (Kansas17) (**Supplementary Table 2**). Our mutagenesis experiments were based on egg-adapted strains. As a result, all the Kansas17 mutants in this experiment were constructed in the presence of the background mutations D190N and S219Y, which were present in 100% (15/15) and 87% (13/15), respectively, of egg-passaged Kansas17 strains in GISAID. Similarly, all the Sing16 and Switz17 mutants were constructed in the presence of the background mutation T160K, which was present in all egg-passaged strains derived from Sing16 and Switz17 in GISAID. The fitness effects of

four different combinations of amino acid variants at residues 186 and 194 were tested, namely 1) G186/L194, which represented the amino acid variants in the unpassaged strains, 2) V186/L194, which represented the single egg-adaptive mutation G186V, 3) G186/P194, which represented the single egg-adaptive mutation L194P, and 4) V186/P194, which represented the co-occurrence of two incompatible egg-adaptive mutations [14].

As expected, all variants with G186/L194 could be rescued, whereas all variants with V186/P194 had no detectable titer (**Figure 1B**). In contrast, the fitness effects of V186/L194 and G186/P194 were different in different strains. Although V186/L194 could be rescued in all three strains **(Figure 1B)**, it only facilitated the egg-adaptation of Kansas17-D190N/S219Y, but not Sing16-T160K and Switz17-T160K (**Figure 1C**). Similarly, G186/P194 could only be rescued in Sing16-T160K and Switz17-T160K, but not Kansas17-D190N/S219Y. Next-generation sequencing further showed that during egg-passaging, V186/L194 were stable in Kansas17-D190N/S219Y, and G186/P194 were stable in both Sing16-T160K and Switz17-T160K (**Figure 1D and Supplementary Table 3**). These results substantiate that there are differential preferences of egg-adaptation pathways among different strains.

Residues 160 and 190 influence the egg-adaptation pathway

The inability of rescuing G186/P194 in Kansas17-D190N/S219Y suggests that G186V is not the only mutation that is incompatible with L194P. We therefore aimed to identify additional mutations that were incompatible with L194P. The HA sequence of Kansas17 X-327, which was an egg-adapted strain with G186V, was compared to that of Switz17 NIB-112 and Sing16 NIB-104, which were two egg-adapted strains with T160K/L194P. Of note, the HA sequences are equivalent between Kansas17 X-327 with V186G/L194P (**Figure 2A**) and the unrescuable Kansas17-G186/P194/N190/Y219 (see above, **Figure 1B**). This sequence analysis led us to identify four candidate amino acid variants in the RBS that might be incompatible with L194P, namely S159, N190, S193 and Y219. Subsequently, S159Y, N190D, S193F and Y219S, which represented the RBS mutations from Kansas17 X-327 to Switz17 NIB-112 and Sing16 NIB-

104, were introduced individually into the HA of Kansas17 X-327 in the presence of the background mutations V186G and L194P. V186G was included in the background since V186 and P194 were known to be incompatible [14]. Our virus rescue experiment showed that the fitness defect of V186G/L194P in Kansas17 X-327 could be restored by N190D (**Figure 2B**). Additionally, Kansas17 X-327 with and without V186G/L194P/N190D had comparable replication fitness in eggs (**Figure 2C**). Next-generation sequencing further confirmed that Kansas17 X-327 with V186G/L194P/N190D was genetically stable during egg-passaging (**Figure 2D and Supplementary Table 3**). To examine if the incompatibility between N190 and P194 could be generalized to other strains, we further introduced D190N into Switz17-K160/P194 and Sing16-K160/P194, which were equivalent to Switz17 NIB-112 and Sing16-K160/P194, respectively. Consistent with the results for Kansas17 X-327, D190N was deleterious in Switz17-K160/P194 and Sing16-K160/P194 (**Figure 2E**). Together, these results demonstrate the incompatibility between N190 and P194.

During our sequence analysis, we noticed that the egg-adaptive mutation T160K [5] often co-occurred with L194P, as exemplified by Switz17 NIB-112 and Sing16 NIB-104. Our rescue experiment showed that T160 was incompatible with P194 in both Sing16 and Switz17 (**Figure 2E**). As a result, in addition to V186 and N190, T160 is another amino acid variant that is incompatible with P194.

Recent human H3N2 strains are incompatible with L194P

Among the amino acid variants that are incompatible with P194, T160 and N190 have reached high occurrence frequency in recent human H3N2 HA. T160, which introduced an N-glycosylation site at N158, was fixed in clade 3C.2a during the 2014-2015 influenza season [5], whereas N190 emerged more recently along with I160 and D186 and reached >90% in 2021 (**Figure 3A**). Not surprisingly, A/Victoria/22/2020 (Vic20) and A/Italy/11871/2020 (Italy20), which were both from clade 3C.2A1b.1a and had T160/N190, could not be rescued in the presence of L194P (**Figure 3B**). In contrast, introducing V186 into Italy20 yielded high titer in

both virus rescue and egg-passaging experiments (**Figure 3C**). These observations illustrate that recent human H3N2 strains have a strong preference against L194P as the egg-adaptation pathway.

While we have shown above that T160 and N190 were incompatible with L194P in Kansas17, Switz17, and Sing16 (**Figure 2**), Italy20 with T160K/N190D was still incompatible with L194P (**Figure 3D**). This result indicates that besides T160 and N190, additional amino acid variants in Italy20 were also incompatible with L194P. We postulated that those additional amino acid variants could be identified by comparing the HA sequences of Italy20 and Sing16, since Sing16 was compatible with L194P in the presence of K160/D190 (**Figure 2E**). Eight mutations in the receptor binding domain (residues 117-265) [21] were identified between Sing16 and Italy20 (**Figure 3E**). Based on spatial proximity between these eight mutations and L194P, we hypothesized that the differential fitness effects of L194P in Sing16 and Italy20 could be due to the presence of mutations G186D, F193S, and S198P in Italy20. Since the compatibility with L194P in Kansas17 was not influenced by the presence of S193 as shown in our rescue experiment above (**Figure 2A-B**), we focused on D186 and P198 here. However, even with mutations D186G and P198S, Italy20-T160K/D186G/N190D/P198S/L194P still could not be rescued (**Figure 3D**). Structural modeling further showed that none of the eight mutations from Sing16 to Italy20 increased the height of the RBS (**Figure 3F**), which was the mechanistic basis of incompatibility between L194P and G186V [14]. Together, these results demonstrate that the L194P incompatibility in recent human H3N2 strains is conferred by multiple amino acid variants including T160 and N190, and likely involves a mechanism that is different from G186V.

Egg-adaptation pathway is clade dependent

The recent emergence of L194P incompatibility suggests that the preference of egg-adaptive pathway has been evolving. To understand the preference of egg-adaptation pathway in recent human H3N2 strains, we analyzed the HA sequences of 72 egg-passaged H3N2 strains

from years 2019 to 2021, which belong to antigenically distinct clades [22-24]. By comparing the HA sequences between the egg-passaged strains and their non-egg-passaged counterparts, we observed that egg-adaptive mutations often emerged at residues 160, 186, 190, 194, 219 and 225 (**Figure 4 and Supplementary Table 4**). In addition, different clades showed different preferences of egg-adaptation pathways. This differential preference could be exemplified by residues 186 and 225. While all non-egg-passaged strains from both clades 3C.2a1b.1a and 3C.2a1b.2a2 had D186, egg-adaptive mutation D186N was common in 3C.2a1b.2a2 but not observed in 3C.2a1b.1a. Similarly, although D225 was present in all non-egg-passaged strains across all clades, egg-adaptive mutation D225G was prevalent in 3C.2a1b.2a2 but not observed in clades 3C.2a1b.2a, 3C.2a1b.2b, and 3C.2a1b.1b. These observations are consistent with the notion that the preference of egg-adaptation pathway in human H3N2 HA is influenced by natural mutations.

DISCUSSION

Egg-adaptation of influenza virus has been central to the production of seasonal influenza vaccine. Our work here shows that the preference of egg-adaptation pathway is strain dependent, due to epistasis between egg-adaptive mutations and natural amino acid variants. Specifically, we identified two natural amino acid variants T160 and N190 that are incompatible with the egg-adaptive mutation L194P. Our results also indicate that additional amino acid variants are responsible for the L194P incompatibility in recent human H3N2 HA. Since L194P is just one of many egg-adaptive mutations in human H3N2 HA, we are likely still far from comprehending the sequence determinants that influence the preference of egg-adaptation pathway.

Our previous studies have shown that epistasis is pervasive between mutations in the HA RBS [16-18]. One interesting observation is that epistasis between two HA RBS mutations can be modulated by a third mutation. For example, epistasis between the engineered mutations L226R and S228E in an H1 HA could only be observed in the presence of G225E [16]. This

type of higher-order epistasis may play a role in determining the egg-adaptation pathway. Despite the incompatibility between T160 and P194 in Sing16 and Switz17 (**Figure 2E**), we noticed the co-occurrence of T160 and P194 in four egg-adapted H3N2 strains (EPI1588456, EPI1584616, EPI1526498 and EPI1440496), which are all belong to clade 3C.2a1b.2b (**Figure 4 and Supplementary Table 4**). This observation implies that the incompatibility between T160 and P194 can be masked by other amino acid variants. As shown by a recent study, epistasis can also alter the antigenic effect of a given mutation in the HA RBS [25]. The complexity of epistasis that involves egg-adaptive mutations as well as the underlying biophysical mechanisms and the antigenic consequences will need to be explored in future studies.

We anticipate that the preference of egg-adaptation pathway will keep changing as human H3N2 HA continues to evolve. As a result, future studies of egg-adaptive mutations in human H3N2 strains are warranted. Given that different egg-adaptive mutations can have different antigenic effects [5, 8, 9, 12, 15, 26-31], characterization of individual egg-adaptive mutations is critical for the selection of vaccine seed strains with minimal antigenic change. Nevertheless, egg-based influenza vaccine production is an 80-year-old technology that needs to be replaced by more advanced alternatives. An important step forward was the commercialization of cell-based and recombinant influenza vaccines, which have shown better effectiveness than egg-based influenza vaccines [32, 33]. In addition, a Phase 1/2 clinical trial of an mRNA seasonal influenza vaccine candidate has initiated recently [34]. Despite the potentially slow transition, these alternatives will provide an ultimate solution to the problems of egg adaptation.

MATERIALS AND METHODS

Sequence analysis

Full-length HA sequences of WHO-recommended H3N2 candidate vaccine virus strains from 2008 to 2020 were downloaded from the Global Initiative for Sharing Avian Influenza Data (GISAID; http://gisaid.org) [20]. Information for these sequences is shown in **Supplementary**

Table 1. Information for the full-length HA sequences of viruses from 2019 to 2021, with and without egg-passaging, is shown in **Supplementary Table 4**. Sequences were aligned by MAFFT version 7 (https://mafft.cbrc.jp/alignment/software/) with default setting [35]. Passaging history was determined by parsing the regular expression in FASTA headers as described [36]. The frequencies of amino acids were plotted using WEBLOGO (https://weblogo.berkeley.edu/logo.cgi) [37]. Phylogenetic tree was built using Maximum Likelihood (ML) method by Geneious Prime (Geneious), with 100 bootstraps.

Cell culture

Humanized Madin-Darby canine kidney (hMDCK) cells with higher and stable expression of human 2,6-sialtransferase were kindly provided by Professor Yoshihiro Kawaoka, the University of Tokyo [38]. Both hMDCK cells and human embryonic kidney (HEK) 293T cells were maintained in minimal essential medium (MEM) supplemented with 10% fetal bovine serum, 25 mM HEPES, and 100 U mL^{-1} penicillin-streptomycin (PS).

Generation of recombinant mutants

All H3N2 viruses generated in this study were based on the influenza eight-plasmid reverse genetics system [39, 40]. The HA and neuraminidase (NA) genes of the strains of interest (**Supplementary Table 2**) were synthesized by Sangon Biotech, and cloned into the pHW2000 vector as previously described [14]. For HA, the ectodomain was from the strains of interest, whereas the noncoding region, N-terminal secretion signal, C-terminal transmembrane domain, and cytoplasmic tail were from H1N1 A/PR/8/34 (PR8). For NA, the entire coding region was from the strains of interest, whereas the noncoding region of NA was from PR8. Mutations were introduced into the constructs by polymerase chain reaction (PCR). Primers used in this study were produced by Integrated DNA Technologies (Integrated DNA Technologies). Recombinant chimeric 6:2 reassortant viruses with the six internal genes (PB2, PB1, PA, NP, M, and NS) from PR8 were rescued by transfecting a co-culture of HEK 293T and hMDCK cells at 6:1 ratio with 70% confluence in a 6-well plate. Transfection was

268 performed using 16 µL of TransIT-LT1 (Mirus Bio) and 1 µg each of the eight plasmids
269 encoding the corresponding influenza virus segments (a total amount of 8 µg of DNA). At 6 h
270 post-transfection, medium was replaced with 1 mL of MEM. At 24 h post-transfection, another
271 1 mL of MEM supplemented with 1 µg mL^{-1} tosylphenylalanyl chloromethyl ketone (TPCK)-
272 trypsin was added. Viruses were harvested at 72 h post-infection and inoculated into hMDCK
273 cells at 90% confluence in a T75 flask. Viruses were harvested when the cytopathic effect
274 (CPE) was observed. Virus titer was determined by titration in hMDCK cells as previously
275 described [41, 42].
276
277 **Virus passaging in embryonic eggs**
278 Embryonic eggs at 10 days post-fertilization were used for the virus passaging. For each
279 passage, 10^4 $TCID_{50}$ of virus was inoculated into one egg. The only exception was for the
280 next-generation sequencing experiment (see below), 10^6 $TCID_{50}$ was used for the first
281 passage of Sing16-K160/V186/L194. The allantoic fluid was harvested at 48 h post-infection.
282 Virus titer was determined by titration in hMDCK cells as previously described [41, 42]. Each
283 experiment consisted of five consecutive passaging and was performed in triplicate.
284
285 **Next-generation sequencing of egg-passaged viruses**
286 The viral RNA of the egg-passaged virus was extracted using QIAamp Viral RNA Mini Kit
287 (Qiagen). The extracted RNA was then reverse transcribed to cDNA using ProtoScript II First
288 Strand cDNA Synthesis Kit (New England Biolabs). Subsequently, PCR was performed using
289 PrimeSTAR Max DNA Polymerase (Takara) with the cDNA as template and virus-specific
290 primers 5'-CAC TCT TTC CCT ACA CGA CGC TCT TCC GAT CT NNN GGT CAC TAG TTG
291 CCT CAT CCG G-3' and 5'- GAC TGG AGT TCA GAC GTG TGC TCT TCC GAT CTG GTG
292 CAT CTG ATC TCA TTA TTG-3'. A second round of PCR was performed to add the rest of
293 adapter sequence and index to the amplicon using primers 5'- AAT GAT ACG GCG ACC ACC
294 GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT-3' and 5'- CAA GCA GAA GAC GGC ATA
295 CGA GAT XXX XXX GTG ACT GGA GTT CAG ACG TGT GCT-3'. Nucleotides in primers

296 annotated by an 'N' or "X" represented the index sequence for distinguishing PCR products
297 derived from different viruses, passages, and biological replicates. The final PCR products
298 were sequenced by Illumina MiSeq PE300.
299
300 **Analysis of next-generation sequencing data**
301 Sequencing data was obtained in FASTQ format and analyzed using in-house Python and
302 shell scripts. Briefly, paired-end reads were merged by PEAR [43]. The merged reads were
303 then parsed by SeqIO module in BioPython and translated into protein sequences [44]. For
304 each sample, the frequency of amino acid variant i was computed as follows:

305 $$frequency_i = \frac{read\ count_i}{sequencing\ depth}$$

306 where $read\ count_i$ represents the number of reads that contain amino acid variant i and
307 $sequencing\ depth$ represents the total number of reads in the sample of interest.
308
309 **Structural modeling**
310 Residues 57 to 261 of a monomer of A/Michigan/15/2014 (H3N2) influenza virus HA (PDB
311 6BKT) [17] was extracted. The PDB file was renumbered using pdb-tools [45]. Point
312 mutagenesis for T128A, V130I, T135K, A138S, G186D, G186V, D190N, F193S, L194P and
313 S198P (all in original numbering) were performed using the fixed backbone (fixbb) backbone
314 design application in Rosetta (RosettaCommons). For each replicate, one hundred poses
315 were generated and the lowest scoring pose was used for downstream analysis. The fixbb
316 application for each mutation was performed in triplicate. A constraint file for each lowest
317 scoring pose was obtained using the minimize_with_cst application in Rosetta and converting
318 the log file to a constraint (cst) file using the convert_to_cst_file application in Rosetta.
319 Subsequently, fast relax was performed on wild type or mutant using the relax application in
320 Rosetta with the corresponding constraint file [46]. For each replicate, 30 poses were
321 generated and the lowest scoring pose was used to measure the height of the RBS, which is
322 the distance between the side chain oxygen atom of Y98 (H3 numbering) and Cα of amino
323 acid 190 (H3 numbering) using PyMOL (Schrödinger).

Data availability

Raw sequencing data have been submitted to the NIH Short Read Archive under accession number: BioProject PRJNA800806. Codes for analyzing next-generation sequencing data have been deposited to https://github.com/Wangyiquan95/HA_egg_passage. Code and source files for structural modeling using Rosetta have been deposited to https://github.com/timothyjtan/HA-RBD_height.

ACKNOWLEDGEMENTS

This work was supported by a fellowship from the Pasteur Foundation Asia (W.L.), a Calmette and Yersin scholarship from the Pasteur International Network Association (H.L.), a startup fund from the University of Illinois at Urbana-Champaign (N.C.W.), the Health and Medical Research Fund (no.19180932) (C.K.P.M) and the National Research Foundation of Korea (NRF) grant funded through the Korea government (NRF-2018M3A9H4055203) (C.K.P.M.). C.K.M.P is a visiting scientist at the Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore.

AUTHOR CONTRIBUTIONS

W.L., H.L., C.K.P.M., and N.C.W. conceived and designed the study. W.L. and Y.S. performed the experiments. W.L., T.J.C.T., and Y.W. analyzed the data. W.L., R.B., C.K.P.M., and N.C.W. wrote the paper, and all authors reviewed and edited the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

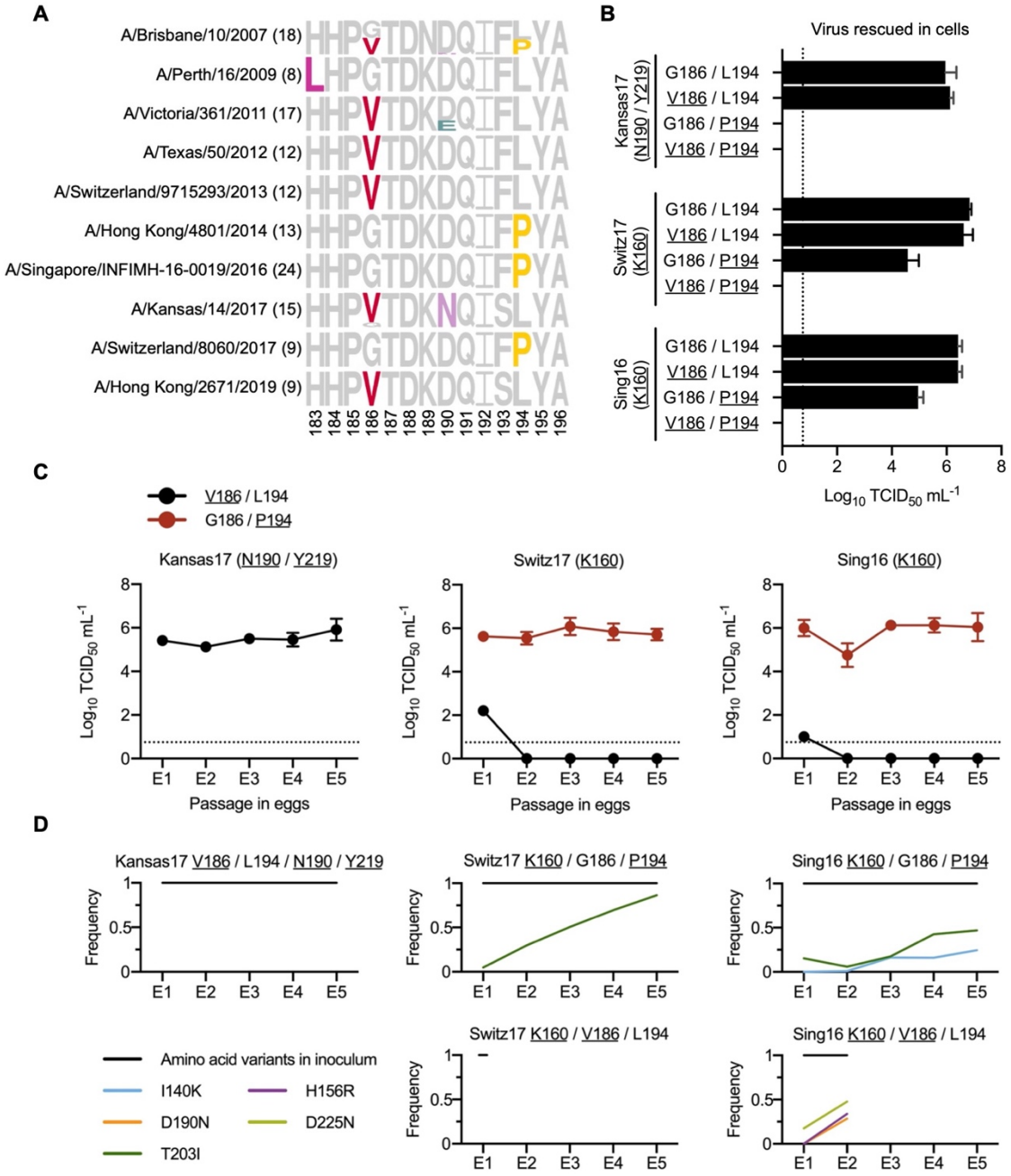

Figure 1. Differential preference of egg-adaptive mutations in human H3N2 vaccine strains. (A) Frequencies of mutations among different egg-passaged seasonal H3N2 vaccine strains are shown as sequence logos. Amino acid variants from residues 183 to 196 are shown (H3 numbering). The number of egg-passaged strains included in this analysis is indicated in the parenthesis. The relative size of each amino acid letter represents its frequency in the sequences. Grey letters represent the amino acid variants that are observed in the corresponding unpassaged parental strains. **(B)** Replication fitness of different mutants of human H3N2 vaccine strains was examined in a virus rescue experiment. Viral titers were

357 measured by $TCID_{50}$. **(C)** The viral titer of each human H3N2 vaccine strain with either
358 V186/L194 or G186/P194 was measured during egg-passaging. **(B-C)** Results are shown as
359 means ± SD of three independent experiments. The dashed line represents the lower
360 detection limit. Amino acid variant representing an egg-adaptive mutation is underlined. **(D)**
361 Frequencies of mutations in the receptor-binding subdomain (residues 117-265) [21] that
362 emerged during serial passaging in eggs. The strain of inoculum for each passaging
363 experiment is indicated above each plot, with those representing egg-adaptive mutations
364 underlined. Frequencies are shown as means of three biological replicates. Only those
365 mutations that reached a minimum average frequency of 10% after the fifth passage are
366 plotted.
367

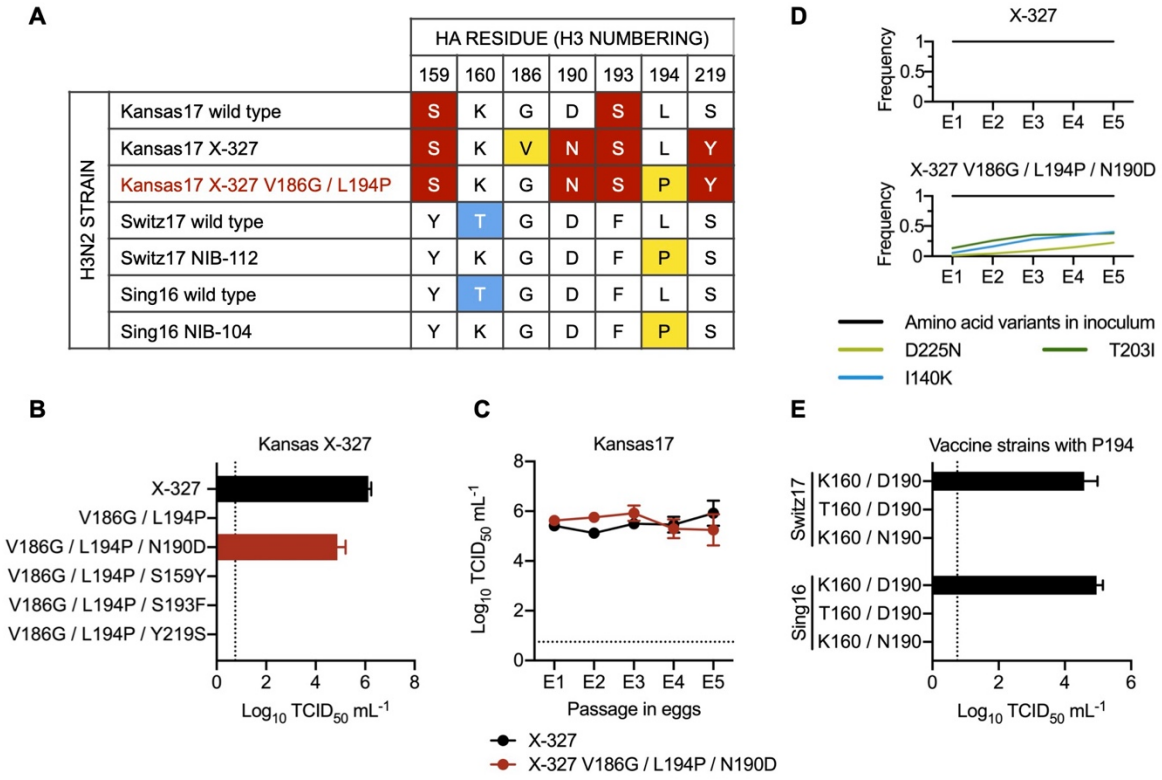

Figure 2. Epistatic interactions that involve egg-adaptive mutations. **(A)** Amino acid variants from residues 150 to 220 among the indicated strains are shown. The rest of the sequences from residues 150 to 220 were completely conserved among these strains. Major egg-adaptive mutations G186V and L194P are in yellow. Amino acid variants that are unique to Kansas17 strains are in red while variants unique to Swiz17 and Sing16 are in blue. Of note, Kansas17 X-327, Switz17 NIB-112, and Sing16 NIB-104 are egg-adapted strains, whereas Kansas17 X-327 V186G/L194P (red) was a mutant generated in this study. **(B)** Replication fitness of Kansas17 X-327 with different mutations was examined in a virus rescue experiment. **(C)** The replication fitness of Kansas17 X-327 and Kansas17 X-327 with triple mutations (V186G/L194P/N190D) in eggs were examined during serial egg-passaging. **(D)** Frequencies of mutations in the receptor-binding subdomain (residues 117-265) [21] that emerged during egg-passaging of the indicated strains. Results are shown as means of three independent biological replicates. Only those mutations that reached a minimum average frequency of 10% after the fifth passage are plotted. **(E)** Replication fitness of Sing16 and Switz17 L194P virus with or without K160T and D190N mutations was assessed in a virus rescue experiment. **(B,**

384 **C, E)** All viral titers were measured by $TCID_{50}$. Results are shown as means ± SD of three

385 independent experiments. The dashed line represents the lower detection limit.

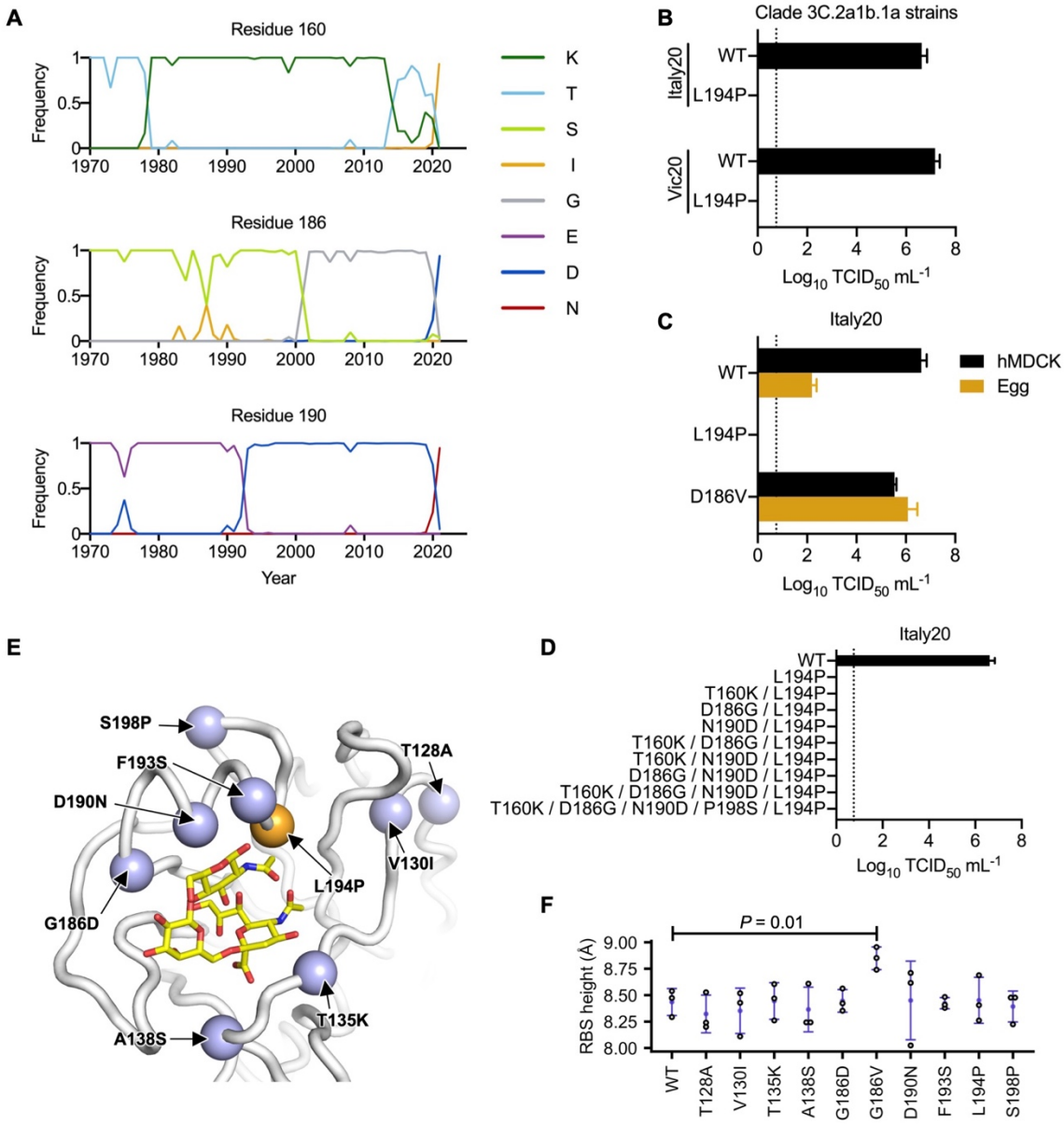

Figure 3. Incompatibility between N190 and P194. **(A)** Frequencies of amino acid variants at residues 160, 186 and 190 in human H3N2 HA over time are shown. Only those variants that reached a maximum annual frequency of 35% are plotted. **(B)** Two strains in clade 3C.2a1b.1a (Vic20 and Italy20) with or without L194P mutation was examined in a virus rescue experiment in hMDCK cells. Of note, most unpassaged strains in clade 3C.2a1b.1a, including both Vic20 and Italy20, naturally contain N190. **(C)** The replication fitness of Italy20 with different mutations was examined in a virus rescue experiment (hMDCK) as well as in an egg-passaging experiment, in which 10^4 $TCID_{50}$ of the rescued viruses were propagated in eggs

for 48 hours. **(D)** The replication fitness of Italy20 with different mutants was measured in a virus rescue experiment. **(B-D)** Viral titers were measured by $TCID_{50}$. Results are shown as means ± SD of three independent experiments. The dashed line represents the lower detection limit. **(E)** The locations of mutations between Sing16 and Italy20 in the receptor-binding subdomain (residues 117-265) [21] are shown in blue (PDB 6BKT) [17]. L194P is in orange. Sialylated glycan receptor is in yellow sticks representation. **(F)** The structural impact of each mutation was modeled by Rosetta [46]. The distance between the phenolic oxygen of Tyr98 (OH_{98}) and the Cα of residue 190 was computed as the height of receptor-binding site (RBS). Three replicates, each with 100 simulations, were performed. Each data point represents the lowest scoring pose in a replicate. Error bars represent means ± SD. *P*-value was computed by two-tailed t-test. Only the difference between WT and G186V is statistically significant ($P < 0.05$).

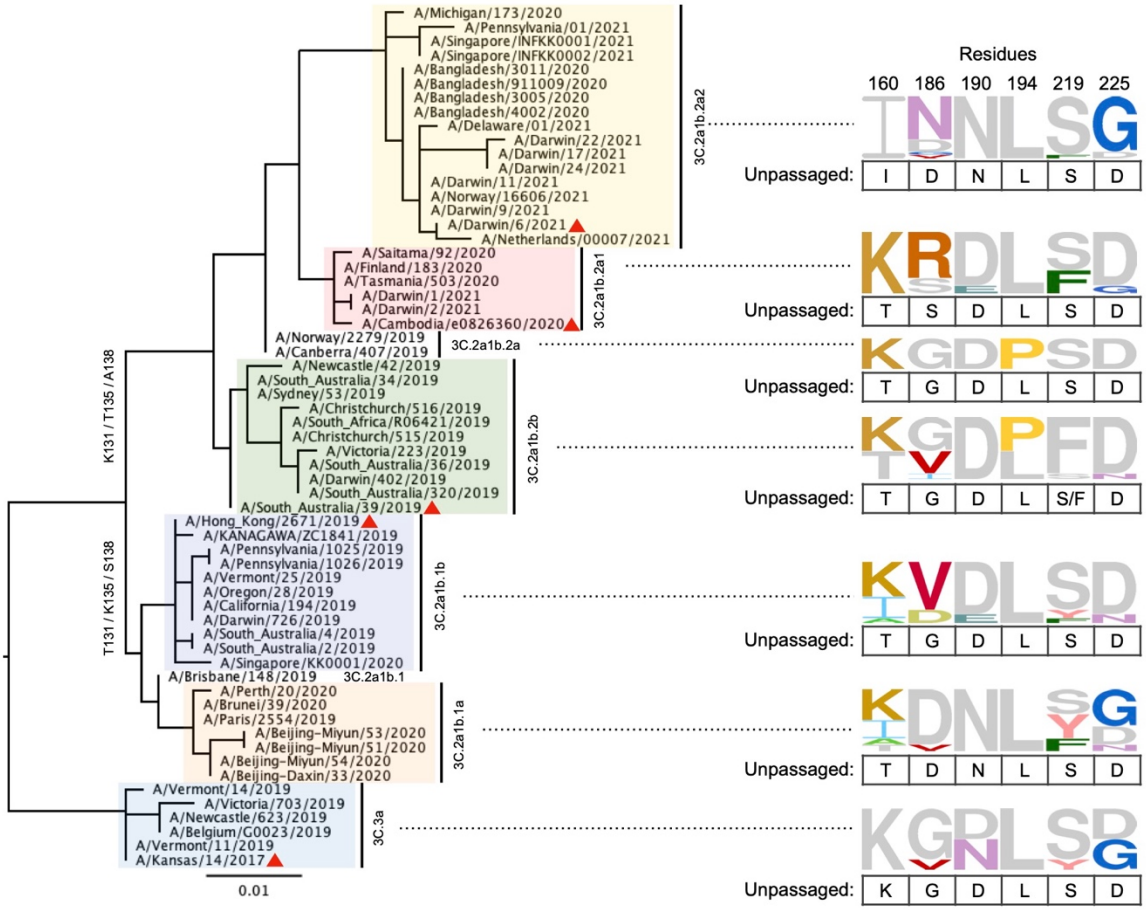

Figure 4. Egg-adaptive mutations in recent human H3N2 HA. A phylogenetic tree was built using the HA sequences of 61 human H3N2 strains without egg-passaging (unpassaged strains). Different clades are highlighted in different colors. WHO-recommended vaccine strains are indicated by a red triangle symbol. These unpassaged human H3N2 strains correspond to the parental strains of 72 egg-passaged strains . Since multiple egg-passaged strains with different HA sequences could be generated from a single unpassaged strain, the number of unpassaged strains is less than that of egg-passaged strains. Frequencies of amino acids at HA residues 160, 186, 190, 194, 219 and 225 among egg-passaged strains in different clades are shown as sequence logos. The relative size of each amino acid letter represents its frequency in the sequences. For each clade, amino acid variants that are also observed in the unpassaged strains are in grey and listed below each sequence logo.

REFERENCES

1. Centers for Disease Control and Prevention. Past seasons vaccine effectiveness estimates [26 August , 2021]. Available from: https://www.cdc.gov/flu/vaccines-work/past-seasons-estimates.html.
2. Belongia EA, Simpson MD, King JP, Sundaram ME, Kelley NS, Osterholm MT, et al. Variable influenza vaccine effectiveness by subtype: a systematic review and meta-analysis of test-negative design studies. Lancet Infect Dis. 2016;16(8):942-51. Epub 20160406. doi: 10.1016/S1473-3099(16)00129-8. PubMed PMID: 27061888.
3. Okoli GN, Racovitan F, Abdulwahid T, Righolt CH, Mahmud SM. Variable seasonal influenza vaccine effectiveness across geographical regions, age groups and levels of vaccine antigenic similarity with circulating virus strains: a systematic review and meta-analysis of the evidence from test-negative design studies after the 2009/10 influenza pandemic. Vaccine. 2021;39(8):1225-40. Epub 20210122. doi: 10.1016/j.vaccine.2021.01.032. PubMed PMID: 33494964.
4. Kodihalli S, Justewicz DM, Gubareva LV, Webster RG. Selection of a single amino acid substitution in the hemagglutinin molecule by chicken eggs can render influenza A virus (H3) candidate vaccine ineffective. J Virol. 1995;69(8):4888-97. doi: 10.1128/JVI.69.8.4888-4897.1995. PubMed PMID: 7609057; PubMed Central PMCID: PMCPMC189303.
5. Zost SJ, Parkhouse K, Gumina ME, Kim K, Diaz Perez S, Wilson PC, et al. Contemporary H3N2 influenza viruses have a glycosylation site that alters binding of antibodies elicited by egg-adapted vaccine strains. Proc Natl Acad Sci U S A. 2017;114(47):12578-83. Epub 2017/11/08. doi: 10.1073/pnas.1712377114. PubMed PMID: 29109276; PubMed Central PMCID: PMCPMC5703309.
6. Rajaram S, Wojcik R, Moore C, Ortiz de Lejarazu R, de Lusignan S, Montomoli E, et al. The impact of candidate influenza virus and egg-based manufacture on vaccine effectiveness: literature review and expert consensus. Vaccine. 2020;38(38):6047-56. Epub 20200626. doi: 10.1016/j.vaccine.2020.06.021. PubMed PMID: 32600916.

7. Ito T, Suzuki Y, Takada A, Kawamoto A, Otsuki K, Masuda H, et al. Differences in sialic acid-galactose linkages in the chicken egg amnion and allantois influence human influenza virus receptor specificity and variant selection. J Virol. 1997;71(4):3357-62. Epub 1997/04/01. doi: 10.1128/JVI.71.4.3357-3362.1997. PubMed PMID: 9060710; PubMed Central PMCID: PMCPMC191479.
8. Barman S, Franks J, Turner JC, Yoon SW, Webster RG, Webby RJ. Egg-adaptive mutations in H3N2v vaccine virus enhance egg-based production without loss of antigenicity or immunogenicity. Vaccine. 2015;33(28):3186-92. Epub 2015/05/23. doi: 10.1016/j.vaccine.2015.05.011. PubMed PMID: 25999284; PubMed Central PMCID: PMCPMC4523127.
9. Parker L, Wharton SA, Martin SR, Cross K, Lin Y, Liu Y, et al. Effects of egg-adaptation on receptor-binding and antigenic properties of recent influenza A (H3N2) vaccine viruses. J Gen Virol. 2016;97(6):1333-44. Epub 2016/03/15. doi: 10.1099/jgv.0.000457. PubMed PMID: 26974849; PubMed Central PMCID: PMCPMC5394856.
10. Wiley DC, Wilson IA, Skehel JJ. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature. 1981;289(5796):373-8. Epub 1981/01/29. doi: 10.1038/289373a0. PubMed PMID: 6162101.
11. Robertson JS, Bootman JS, Newman R, Oxford JS, Daniels RS, Webster RG, et al. Structural changes in the haemagglutinin which accompany egg adaptation of an influenza A(H1N1) virus. Virology. 1987;160(1):31-7. doi: 10.1016/0042-6822(87)90040-7. PubMed PMID: 3629978.
12. Chen Z, Zhou H, Jin H. The impact of key amino acid substitutions in the hemagglutinin of influenza A (H3N2) viruses on vaccine production and antibody response. Vaccine. 2010;28(24):4079-85. Epub 2010/04/20. doi: 10.1016/j.vaccine.2010.03.078. PubMed PMID: 20399830.
13. Wu NC, Wilson IA. A perspective on the structural and functional constraints for immune evasion: insights from influenza virus. J Mol Biol. 2017;429(17):2694-709. Epub

2017/06/27. doi: 10.1016/j.jmb.2017.06.015. PubMed PMID: 28648617; PubMed Central PMCID: PMCPMC5573227.

14. Wu NC, Lv H, Thompson AJ, Wu DC, Ng WWS, Kadam RU, et al. Preventing an antigenically disruptive mutation in egg-based H3N2 seasonal influenza vaccines by mutational incompatibility. Cell Host Microbe. 2019;25(6):836-44 e5. Epub 2019/06/04. doi: 10.1016/j.chom.2019.04.013. PubMed PMID: 31151913; PubMed Central PMCID: PMCPMC6579542.
15. Wu NC, Zost SJ, Thompson AJ, Oyen D, Nycholat CM, McBride R, et al. A structural explanation for the low effectiveness of the seasonal influenza H3N2 vaccine. PLoS Pathog. 2017;13(10):e1006682. Epub 2017/10/24. doi: 10.1371/journal.ppat.1006682. PubMed PMID: 29059230; PubMed Central PMCID: PMCPMC5667890.
16. Wu NC, Xie J, Zheng T, Nycholat CM, Grande G, Paulson JC, et al. Diversity of functionally permissive sequences in the receptor-binding site of influenza hemagglutinin. Cell Host Microbe. 2017;22(2):247-8. doi: 10.1016/j.chom.2017.07.001. PubMed PMID: 28799910; PubMed Central PMCID: PMCPMC5652301.
17. Wu NC, Thompson AJ, Xie J, Lin CW, Nycholat CM, Zhu X, et al. A complex epistatic network limits the mutational reversibility in the influenza hemagglutinin receptor-binding site. Nat Commun. 2018;9(1):1264. Epub 2018/03/30. doi: 10.1038/s41467-018-03663-5. PubMed PMID: 29593268; PubMed Central PMCID: PMCPMC5871881.
18. Wu NC, Otwinowski J, Thompson AJ, Nycholat CM, Nourmohammad A, Wilson IA. Major antigenic site B of human influenza H3N2 viruses has an evolving local fitness landscape. Nat Commun. 2020;11(1):1233. Epub 2020/03/08. doi: 10.1038/s41467-020-15102-5. PubMed PMID: 32144244; PubMed Central PMCID: PMCPMC7060233.
19. Lin YP, Xiong X, Wharton SA, Martin SR, Coombs PJ, Vachieri SG, et al. Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. Proc Natl Acad Sci U S A. 2012;109(52):21474-9. Epub 20121210. doi: 10.1073/pnas.1218841110. PubMed PMID: 23236176; PubMed Central PMCID: PMCPMC3535595.
20. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data - from vision to

505 reality. Euro Surveill. 2017;22(13). doi: 10.2807/1560-7917.ES.2017.22.13.30494.
506 PubMed PMID: 28382917; PubMed Central PMCID: PMCPMC5388101.
507 21. Ha Y, Stevens DJ, Skehel JJ, Wiley DC. H5 avian and H9 swine influenza virus
508 haemagglutinin structures: possible origin of influenza subtypes. EMBO J.
509 2002;21(5):865-75. doi: 10.1093/emboj/21.5.865. PubMed PMID: 11867515; PubMed
510 Central PMCID: PMCPMC125880.
511 22. World Health Organization. Recommended composition of influenza virus vaccines for
512 use in the 2022 southern hemisphere influenza season [24 September, 2021].
513 Available from: https://cdn.who.int/media/docs/default-source/influenza/who-influenza-
514 recommendations/vcm-southern-hemisphere-recommendation-
515 2022/202109_recommendation.pdf.
516 23. Bolton MJ, Ort JT, McBride R, Swanson NJ, Wilson J, Awofolaju M, et al. Antigenic and
517 virological properties of an H3N2 variant that will likely dominate the 2021-2022
518 Northern Hemisphere influenza season. medRxiv [Preprint]. 2021. doi:
519 10.1101/2021.12.15.21267857.
520 24. Siegers JY, Dhanasekaran V, Xie R, Deng YM, Patel S, Ieng V, et al. Genetic and
521 antigenic characterization of an influenza A(H3N2) outbreak in cambodia and the
522 greater mekong subregion during the COVID-19 pandemic, 2020. J Virol.
523 2021:JVI0126721. Epub 2021/09/30. doi: 10.1128/JVI.01267-21. PubMed PMID:
524 34586866.
525 25. Koel BF, Burke DF, van der Vliet S, Bestebroer TM, Rimmelzwaan GF, Osterhaus A, et
526 al. Epistatic interactions can moderate the antigenic effect of substitutions in
527 haemagglutinin of influenza H3N2 virus. J Gen Virol. 2019;100(5):773-7. Epub
528 20190424. doi: 10.1099/jgv.0.001263. PubMed PMID: 31017567; PubMed Central
529 PMCID: PMCPMC7176280.
530 26. Jin H, Zhou H, Liu H, Chan W, Adhikary L, Mahmood K, et al. Two residues in the
531 hemagglutinin of A/Fujian/411/02-like influenza viruses are responsible for antigenic drift
532 from A/Panama/2007/99. Virology. 2005;336(1):113-9. doi: 10.1016/j.virol.2005.03.010.

PubMed PMID: 15866076.

27. Skowronski DM, Janjua NZ, De Serres G, Sabaiduc S, Eshaghi A, Dickinson JA, et al. Low 2012-13 influenza vaccine effectiveness associated with mutation in the egg-adapted H3N2 vaccine strain not antigenic drift in circulating viruses. PLoS One. 2014;9(3):e92153. Epub 20140325. doi: 10.1371/journal.pone.0092153. PubMed PMID: 24667168; PubMed Central PMCID: PMCPMC3965421.
28. Popova L, Smith K, West AH, Wilson PC, James JA, Thompson LF, et al. Immunodominance of antigenic site B over site A of hemagglutinin of recent H3N2 influenza viruses. PLoS One. 2012;7(7):e41895. Epub 20120725. doi: 10.1371/journal.pone.0041895. PubMed PMID: 22848649; PubMed Central PMCID: PMCPMC3405050.
29. Lu B, Zhou H, Chan W, Kemble G, Jin H. Single amino acid substitutions in the hemagglutinin of influenza A/Singapore/21/04 (H3N2) increase virus growth in embryonated chicken eggs. Vaccine. 2006;24(44-46):6691-3. Epub 2006/07/04. doi: 10.1016/j.vaccine.2006.05.062. PubMed PMID: 16814431.
30. Widjaja L, Ilyushina N, Webster RG, Webby RJ. Molecular changes associated with adaptation of human influenza A virus in embryonated chicken eggs. Virology. 2006;350(1):137-45. Epub 20060320. doi: 10.1016/j.virol.2006.02.020. PubMed PMID: 16545416.
31. Lu B, Zhou H, Ye D, Kemble G, Jin H. Improvement of influenza A/Fujian/411/02 (H3N2) virus growth in embryonated chicken eggs by balancing the hemagglutinin and neuraminidase activities, using reverse genetics. J Virol. 2005;79(11):6763-71. doi: 10.1128/JVI.79.11.6763-6771.2005. PubMed PMID: 15890915; PubMed Central PMCID: PMCPMC1112156.
32. Boikos C, Sylvester GC, Sampalis JS, Mansi JA. Relative effectiveness of the cell-cultured quadrivalent influenza vaccine compared to standard, egg-derived quadrivalent influenza vaccines in preventing influenza-like illness in 2017-2018. Clin Infect Dis. 2020;71(10):e665-e71. doi: 10.1093/cid/ciaa371. PubMed PMID: 32253431; PubMed

Central PMCID: PMCPMC7745007.

33. Dawood FS, Naleway AL, Flannery B, Levine MZ, Murthy K, Sambhara S, et al. Comparison of the immunogenicity of cell culture-based and recombinant quadrivalent influenza vaccines to conventional egg-based quadrivalent influenza vaccines among healthcare personnel aged 18-64 years: a randomized open-label trial. Clin Infect Dis. 2021;73(11):1973-81. doi: 10.1093/cid/ciab566. PubMed PMID: 34245243; PubMed Central PMCID: PMCPMC8499731.

34. Abbasi J. Moderna's mRNA vaccine for seasonal flu enters clinical trials. JAMA. 2021;326(14):1365. doi: 10.1001/jama.2021.17499. PubMed PMID: 34636873.

35. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059-66. doi: 10.1093/nar/gkf436. PubMed PMID: 12136088; PubMed Central PMCID: PMCPMC135756.

36. McWhite CD, Meyer AG, Wilke CO. Sequence amplification via cell passaging creates spurious signals of positive adaptation in influenza virus H3N2 hemagglutinin. Virus Evol. 2016;2(2). Epub 2016/10/08. doi: 10.1093/ve/vew026. PubMed PMID: 27713835; PubMed Central PMCID: PMCPMC5049878.

37. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14(6):1188-90. doi: 10.1101/gr.849004. PubMed PMID: 15173120; PubMed Central PMCID: PMCPMC419797.

38. Takada K, Kawakami C, Fan S, Chiba S, Zhong G, Gu C, et al. A humanized MDCK cell line for the efficient isolation and propagation of human influenza viruses. Nat Microbiol. 2019;4(8):1268-73. Epub 2019/05/01. doi: 10.1038/s41564-019-0433-6. PubMed PMID: 31036910.

39. Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. A DNA transfection system for generation of influenza A virus from eight plasmids. Proc Natl Acad Sci U S A. 2000;97(11):6108-13. Epub 2000/05/10. doi: 10.1073/pnas.100133697. PubMed PMID: 10801978; PubMed Central PMCID: PMCPMC18566.

589 40. Neumann G, Watanabe T, Ito H, Watanabe S, Goto H, Gao P, et al. Generation of
590 influenza A viruses entirely from cloned cDNAs. Proc Natl Acad Sci U S A.
591 1999;96(16):9345-50. doi: 10.1073/pnas.96.16.9345. PubMed PMID: 10430945;
592 PubMed Central PMCID: PMCPMC17785.
593 41. Chan MC, Chan RW, Chan LL, Mok CK, Hui KP, Fong JH, et al. Tropism and innate
594 host responses of a novel avian influenza A H7N9 virus: an analysis of ex-vivo and in-
595 vitro cultures of the human respiratory tract. Lancet Respir Med. 2013;1(7):534-42.
596 Epub 20130725. doi: 10.1016/S2213-2600(13)70138-3. PubMed PMID: 24461614;
597 PubMed Central PMCID: PMCPMC7164816.
598 42. Chan MC, Chan RW, Yu WC, Ho CC, Yuen KM, Fong JH, et al. Tropism and innate host
599 responses of the 2009 pandemic H1N1 influenza virus in ex vivo and in vitro cultures of
600 human conjunctiva and respiratory tract. Am J Pathol. 2010;176(4):1828-40. Epub
601 2010/01/30. doi: 10.2353/ajpath.2010.091087. PubMed PMID: 20110407; PubMed
602 Central PMCID: PMCPMC2843473.
603 43. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-
604 End reAd mergeR. Bioinformatics. 2014;30(5):614-20. Epub 20131018. doi:
605 10.1093/bioinformatics/btt593. PubMed PMID: 24142950; PubMed Central PMCID:
606 PMCPMC3933873.
607 44. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely
608 available Python tools for computational molecular biology and bioinformatics.
609 Bioinformatics. 2009;25(11):1422-3. Epub 20090320. doi:
610 10.1093/bioinformatics/btp163. PubMed PMID: 19304878; PubMed Central PMCID:
611 PMCPMC2682512.
612 45. Rodrigues J, Teixeira JMC, Trellet M, Bonvin A. Pdb-tools: a swiss army knife for
613 molecular structures. F1000Res. 2018;7:1961. Epub 20181220. doi:
614 10.12688/f1000research.17456.1. PubMed PMID: 30705752; PubMed Central PMCID:
615 PMCPMC6343223.
616 46. Kellogg EH, Leaver-Fay A, Baker D. Role of conformational sampling in computing

617 mutation-induced changes in protein structure and stability. Proteins. 2011;79(3):830-8.

618 Epub 20101203. doi: 10.1002/prot.22921. PubMed PMID: 21287615; PubMed Central

619 PMCID: PMCPMC3760476.

620