

1 **Selection of antigenically advanced variants of seasonal influenza viruses**

2

3 Chengjun Li^{1*}, Masato Hatta^{1*}, David F. Burke^{2,3*}, Jihui Ping^{1*}, Ying Zhang^{1*}, Makoto
4 Ozawa^{1,4}, Andrew S. Taft¹, Subash C. Das¹, Anthony P. Hanson¹, Jiasheng Song¹,
5 Masaki Imai^{1,5}, Peter R. Wilker¹, Tokiko Watanabe⁶, Shinji Watanabe⁶, Mutsumi Ito⁷,
6 Kiyoko Iwatsuki-Horimoto⁷, Colin A. Russell^{3,8,9}, Sarah L. James^{2,3}, Eugene Skepner^{2,3},
7 Eileen A. Maher¹, Gabriele Neumann¹, Alexander I. Klimov^{10†}, Anne Kelso¹¹, John
8 McCauley¹², Dayan Wang¹³, Yuelong Shu¹³, Takato Odagiri¹⁴, Masato Tashiro¹⁴, Xiyan
9 Xu¹⁰, David E. Wentworth¹⁰, Jacqueline M. Katz¹⁰, Nancy J. Cox¹⁰, Derek J. Smith^{2,3,15},
10 and Yoshihiro Kawaoka^{1,4,6,7}

11

12 ¹Influenza Research Institute, Department of Pathobiological Sciences, School of
13 Veterinary Medicine, University of Wisconsin-Madison, Madison, WI 53711, USA.

14 ²Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.

15 ³World Health Organization Collaborating Centre for Modeling, Evolution, and Control
16 of Emerging Infectious Diseases, Cambridge CB2 3EJ, UK.

17 ⁴Department of Special Pathogens, International Research Center for Infectious Diseases,
18 Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan.

19 ⁵Department of Veterinary Medicine, Faculty of Agriculture, Iwate University, Iwate
20 020-8550, Japan.

21 ⁶ERATO Infection-Induced Host Responses Project, Saitama 332-0012, Japan.

22 ⁷Division of Virology, Department of Microbiology and Immunology, Institute of
23 Medical Science, University of Tokyo, Tokyo 108-8639, Japan.

24 ⁸Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA.

25 ⁹Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES,
26 UK.

27 ¹⁰Influenza Division, National Center for Immunization and Respiratory Diseases,
28 Centers for Disease Control and Prevention, Atlanta, Georgia 30033, USA.

29 ¹¹WHO Collaborating Centre for Reference and Research on Influenza (VIDRL) at the
30 Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria 3000, Australia.

1 ¹²Division of Virology, MRC National Institute for Medical Research, The Ridgeway,
2 Mill Hill, London NW7 1AA, United Kingdom.

3 ¹³Chinese National Influenza Center, National Institute for Viral Disease Control and
4 Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China.

5 ¹⁴Influenza Virus Research Center, National Institute of Infectious Diseases, Musashi-
6 Murayama, Tokyo 208-0011, Japan.

7 ¹⁵Department of Virology, Erasmus Medical Center, 3000 CA Rotterdam, Netherlands.

8 *These authors contributed equally to this work.

9 †Deceased.

10

1 **Influenza viruses mutate frequently, necessitating constant updates of**
2 **vaccine viruses. To establish experimental approaches that may complement the**
3 **current vaccine strain selection process, we selected antigenic variants from human**
4 **H1N1 and H3N2 influenza virus libraries possessing random mutations in the**
5 **globular head of the haemagglutinin protein (which includes the antigenic sites) by**
6 **incubating them with human and/or ferret convalescent sera to human H1N1 and**
7 **H3N2 viruses. Further, we selected antigenic escape variants from human viruses**
8 **treated with convalescent sera and from mice that had been previously immunized**
9 **against human influenza viruses. Our pilot studies with past influenza viruses**
10 **identified escape mutants that were antigenically similar to variants that emerged in**
11 **nature, establishing the feasibility of our approach. Our studies with contemporary**
12 **human influenza viruses identified escape mutants before they caused an epidemic**
13 **in 2014–2015. This approach may aid in the prediction of potential antigenic escape**
14 **variants and the selection of future vaccine candidates before they become**
15 **widespread in nature.**

1 The haemagglutinin (HA) protein is the major influenza viral antigen¹ and the
2 primary target of infection-neutralizing antibodies. During influenza virus circulation in
3 humans, the antigenicity of the virus gradually changes due to mutations in the globular
4 head of HA, necessitating frequent updates of the influenza vaccine^{2,3}. Selection of
5 antigenic variants of influenza viruses in humans is believed to be driven by neutralizing
6 antibodies⁴. Currently, seasonal influenza vaccine strains are selected based on the
7 antigenicity of clinical isolates and cross-reactive immunity in human populations in
8 combination with genetic and epidemiologic data⁵. This decision has to be made more
9 than six months prior to the onset of the influenza season to allow sufficient time for
10 vaccine manufacture. At the time of decision making, novel antigenic clusters may not
11 have emerged or may not yet be recognized, occasionally resulting in the selection of
12 vaccine strains that differ antigenically from the viruses circulating during the subsequent
13 influenza season, leading to low vaccine effectiveness^{6,7}. Presently, no experimental or
14 computational methods exist to predict future antigenic variants before they emerge in
15 nature. We, therefore, developed an experimental strategy to predict the antigenic
16 evolution of human influenza A viruses.

17 In our approach (Fig. 1), mutant virus libraries with random mutations in the
18 antigenic region of HA are screened *in vitro* with antisera against circulating viruses or in
19 mice immunized against the circulating viruses. The antigenic properties of the escape
20 mutants are analyzed by using haemagglutination inhibition (HI) assays. Antigenic
21 cartography (an established method to visualize and analyze HI data⁸) is then used to
22 identify novel antigenic clusters that may necessitate the update of an influenza vaccine.

1 This experimental approach could identify novel antigenic clusters before they emerge in
2 nature and complement the current vaccine selection process.

3 **Antigenic evolution of A(H1N1)pdm09 viruses**

4 First, we performed a proof-of-concept study to identify and characterize
5 antigenic escape variants to pandemic (H1N1) 2009 influenza A (A(H1N1)pdm09)
6 viruses. Eight convalescent sera collected from A(H1N1)pdm09-infected individuals in
7 Japan (Supplementary Table 1) were mixed with two A(H1N1)pdm09 clinical isolates
8 and escape viruses were isolated using plaque assays in Madin-Darby canine kidney
9 (MDCK) cells. Escape mutants possessed the K153E/N239K, D127E, or K163E
10 mutations (H1 numbering; Supplementary Table 2). HI assays with ferret antisera against
11 the escape variants confirmed that they were antigenically distinct from their parental
12 viruses (Supplementary Table 3).

13 To further explore the antigenic landscape, we generated a virus library with
14 random mutations in the HA globular head (encompassing amino acids 54–253;
15 Supplementary Fig. 1 and Supplementary Table 4). This region spans the four antigenic
16 domains including antigenic site Sa (positions 124, 125, 153–157, and 159–164)⁹⁻¹² to
17 which two of the selected mutations (K153E and K163E) map. For library generation, we
18 chose the A(H1N1)pdm09 A/Norway/3858/2009 HA whose globular head sequence
19 closely matches the consensus sequence of A(H1N1)pdm09 viruses. The HA mutant
20 library was screened for escape mutants with the 14 human convalescent sera indicated in
21 Supplementary Table 1. We sequenced the HA gene of more than 500 plaque-purified
22 viruses and identified 92 HA mutant clones (Supplementary Table 5), including viruses

1 with the mutations D127E and K163E detected in our preliminary experiments.
2 Importantly, no mutations were found in the HA gene of wild-type virus after three
3 consecutive passages in MDCK cells, indicating that the identified mutations were not
4 selected during virus replication in these cells. HI assays confirmed that single amino
5 acid changes at four positions in the Sa antigenic site (153, 154, 155, 156) caused 4–8-
6 fold reductions in HI titres; Supplementary Table 5). In addition, 2–4-fold reduction in HI
7 titres were caused by mutations at other positions in antigenic site Sa (125, 127, 129 and
8 163) as well as one position adjacent to site Sb (183). Collectively, these data
9 demonstrate the importance of antigenic site Sa for A(H1N1)pdm09 antigenicity.

10 Previously, Koel *et al.* demonstrated that amino acid mutations at seven positions
11 (142, 152, 153, 155, 156, 186, and 190; H1 numbering) of seasonal H3N2 influenza
12 viruses altered antigenicity¹³. Although the primary amino acid sequences of H1 and H3
13 HAs differ by ~57%, their antigenic sites substantially overlap^{9,10,14}. We, therefore,
14 included these positions in our analysis of A(H1N1)pdm09 viruses. On the basis of
15 sequence comparisons of seasonal H1N1 vaccine virus strains from 1977 to 2007, we
16 also selected position 141 (identified in our preliminary escape mutant selection) and
17 position 187 (located close to key antigenic positions 186 and 190) for further analysis.

18 Next, we generated A/Norway/3858/2009 HA libraries that represent all 20 amino
19 acids at each of the 15 selected positions (i.e., HA positions 125, 127, 129, 141, 142, 152,
20 153, 154, 155, 156, 163, 183, 186, 187, and 190) (Supplementary Fig. 1 and
21 Supplementary Table 6). These HA mutant libraries were individually screened with
22 ferret antisera as shown in Supplementary Table 7. Libraries for which we did not have
23 escape mutant-specific antisera (HA positions 141, 142, 152, 186, 187, and 190) were

1 screened with antiserum to wild-type virus; libraries for which we isolated antigenic
2 escape mutants in our earlier experiments (namely, HA positions 125, 127, 129, 153, 154,
3 155, 156, 163, and 183) were screened with antisera to the N125D, D127E, N129D,
4 K153E, K154N, G155E, N156D, K163E, or S183P escape mutants, respectively, to
5 obtain additional mutants at the respective positions. For each screen, at least 11 plaque-
6 purified viruses were characterized for their HA sequence and antigenicity
7 (Supplementary Table 8). Mutations at positions 153–156 reduced HI titres up to 16-fold,
8 suggesting that these HA positions play a major role in A(H1N1)pdm09 HA antigenicity.
9 We did not identify antigenic escape mutants at positions 141, 163, or 187.

10 Our data suggested a key role for amino acid positions 153–156 in
11 A(H1N1)pdm09 antigenicity; therefore, we generated HA mutant libraries with random
12 mutations at two, or all four, of these positions (Supplementary Table 6). Random
13 mutations at positions 153–156 were also combined with the D127E mutation in HA.
14 This mutation is located near the Sa antigenic site, was selected in our screens
15 (Supplementary Tables 2 and 5), and has been detected in several natural
16 A(H1N1)pdm09 isolates. Virus libraries possessing mutations at two sites were screened
17 with 50:50 mixtures of antisera against escape mutants at the respective positions
18 (Supplementary Table 7); libraries possessing mutations at four or five sites were
19 screened with a human serum (A163) obtained from an individual vaccinated with a
20 monovalent, inactivated split-virus vaccine against A/California/07/2009
21 (A(H1N1)pdm09) (Supplementary Table 7). More than 320 escape mutants were isolated
22 that exhibited reduced reactivity with antiserum against the wild-type virus
23 (Supplementary Tables 9–11). Among those, certain amino acid changes were found at

1 higher frequency (e.g., the N156G mutation). Sequence analysis of mutagenized plasmids
2 used for virus library generation did not show sequence biases (Supplementary Tables 12
3 and 13); therefore, more frequently detected variants were most likely selected due to
4 immune pressure and/or structural changes that conferred a fitness advantage.

5 We also carried out antigenic selection studies in which A/California/04/2009
6 virus-infected mice (Supplementary Table 14) were challenged with the mutant HA virus
7 library possessing random amino acid changes at positions 153–156 together with the
8 D127E mutation. We sequenced the HA gene of 53 escape viruses and detected several
9 amino acid combinations that had also been selected in our *in vitro* antigenic selection
10 experiments (Supplementary Table 15). These findings suggest that our *in vitro* selection
11 experiments with human sera in part reflect *in vivo* immune evasion mechanisms.

12 To obtain a comprehensive picture of the antigenic properties of selected escape
13 variants, 469 viruses (including 447 HA escape variants selected in this study) were
14 characterized in HI assays with a panel of ferret antisera against H1N1 influenza viruses
15 (Supplementary Table 16). The HI data were then analyzed using antigenic
16 cartography^{8,13,15,16}. Many of the escape mutants form a single cluster approximately four
17 antigenic units (representing a 16-fold reduction in HI titre) away from wild-type
18 A/Norway/3858/2009 virus (Fig. 2a). Thus, although our *in vivo* selection studies resulted
19 in the isolation of 28 different amino acid combinations, they all localized to the same
20 novel antigenic cluster. Mutations at positions 153–156 had the greatest impact on
21 antigenicity (Fig. 3). Importantly, single amino acid changes are sufficient to produce
22 antigenic variants with >16-fold difference in HI titre (Fig. 2b and Supplementary Fig. 2).

1 To exclude the possibility that the antisera did not adequately resolve the
2 antigenicity of the escape mutants in the antigenic cluster we identified, we raised
3 additional ferret antisera against 10 antigenic variants that lie in the new cluster or were
4 outliers of the original cluster (Supplementary Fig. 3; Supplementary Table 17). HI
5 assays with these antisera did not result in appreciable changes to overall relationships or
6 placement of any of the viruses (Supplementary Fig. 3b). Thus, the identification of only
7 one new antigenic cluster is a valid finding and not an artifact of poor resolution of
8 antigenic differences by the antisera used in this study. Further analysis also revealed that
9 the accumulation of mutations at positions tested here resulted in viruses that are
10 increasingly distinct from the parental virus (Supplementary Fig. 4); however, the effect
11 of individual mutations cannot be summed to predict the effect of combinations of
12 mutations (Supplementary Fig. 5).

13 Since our study identified several potential A(H1N1)pdm09 escape variants, we
14 analyzed the antigenicity of 2,555 recently circulating A(H1N1)pdm09 strains derived
15 from routine WHO surveillance from 2009–2013. Isolates from each season were
16 analyzed separately due to the lack of common antigens and antisera between the datasets
17 (Fig. 2c, Supplementary Fig. 6a–e). Over 6% of the strains were antigenically different
18 (i.e., more than 4-fold difference in HI titre) from the vaccine strain,
19 A/California/07/2009. The antigenic variation seen among these viruses was
20 representative of that seen in A(H1N1)pdm09 viruses currently circulating worldwide¹⁷.
21 The antigenic differences found were identical to those identified with experimentally
22 selected antigenic variants in this study in terms of distance and direction from the main
23 antigenic cluster (typified by the vaccine strain) as well as the formation of a single

1 antigenically distinct cluster. No mutations at any of the other sites identified by Koel *et*
2 *al.* were observed¹³. All but two of the antigenic variants identified among circulating
3 strains contained mutations shown in this study to cause antigenic change [K153E,
4 K154E, G155E or N156D]. The exceptions were two isolates of the same virus: egg-
5 passaged A/Iraq/8529/2009 (E3) and MDCK cell-passaged A/Iraq/8529/2009 (MDCK1).
6 To directly compare the two datasets, we tested 20 field viruses against a subset of ferret
7 antisera used in this study and generated a new antigenic map that merged the datasets
8 from both experimentally generated mutants (Fig. 2b) and field strains (Fig. 2c) (Fig. 2d;
9 a zoomable version is provided in Supplementary Fig. 6f). Isolates that were antigenically
10 advanced, but had no sequence information, were subsequently sequenced. The antigenic
11 properties of these viruses were consistent between the two datasets and demonstrated
12 that a laboratory-based antigenic selection procedure can produce variants seen in
13 surveillance samples.

14 Our experimental data demonstrate that the amino acids at positions 153–156 of
15 HA are critical for antigenicity. This finding is consistent with several studies showing
16 that a small proportion of A(H1N1)pdm09 clinical strains with amino acid changes in HA
17 positions 153–157 showed antigenic drift compared with the A/California/07/2009-like
18 reference viruses¹⁸⁻²⁰. Moreover, mutations at positions 153–157 of HA have been
19 detected after *in vitro* passage of A(H1N1)pdm09 viruses in the presence of monoclonal
20 antibodies^{21,22} or after A(H1N1)pdm09 passage in suboptimally vaccinated ferrets²⁰. The
21 HA positions 153–156 are located at the edge of the receptor-binding pocket; hence, the
22 amino acids at these positions may affect receptor-binding avidity and specificity^{20,23}, in
23 addition to antigenicity.

1 To determine whether the HA mutants selected in our study and the field strains
2 from the new antigenic cluster could indeed evade immunity to current A(H1N1)pdm09
3 strains, mice were immunized by infection or vaccination with an A(H1N1)pdm09 virus
4 (Supplementary Figs 7 and 8). Four weeks later, the mice were challenged with the
5 viruses indicated in Supplementary Figs 9 and 10 (see also Supplementary Tables 18 and
6 19). No virus was detected on day 4 post-challenge in the lungs of immunized mice
7 challenged with A/California/04/2009 or A/Norway/3858/2009 wild-type viruses
8 (Supplementary Figs 9 and 10). By contrast, 12 of the 13 antigenic variants tested
9 replicated in the lungs of immunized mice. Similarly, several viruses representing
10 antigenic variants isolated in our study and found in field strains also replicated in mice
11 immunized against current A(H1N1)pdm09 viruses (Supplementary Fig. 10).

12 Next, we tested whether the selected antigenic variants could evade immunity to
13 the current A(H1N1)pdm09 viruses in ferrets, a commonly used animal model in the
14 evaluation of influenza vaccines. Ferrets were immunized by infection with
15 A/California/04/2009 virus. Pre-challenge serum HI titres were between 1:160 and 1:640
16 (Supplementary Fig. 11). Twelve months later, ferrets were challenged with either
17 A/California/04/2009 control virus, three different mutant viruses that replicated in mice
18 immunized with wild-type virus (Fig. 4), or with an A/Norway/3568/2009 reassortant
19 virus possessing HA changes identified in a field strain (this mutant virus replicated in
20 mice immunized with the wild-type A/California/04/2009 virus or with the
21 A/California/07/2009 split vaccine) (Fig. 4). No virus was detected from the nasal washes
22 collected from the immunized ferrets challenged with A/California/04/2009 virus. By
23 contrast, the three antigenic mutants identified in our *in vivo* screens and the antigenic

1 mutant with HA changes found in the field strain replicated efficiently in immunized
2 ferrets. Thus, viruses possessing these mutations in HA may have the potential to cause
3 epidemics in the future.

4 As influenza viruses evolve and amino acid substitutions become fixed within an
5 antigenic cluster, these changes in genetic background may affect the future evolution of
6 the viruses. If so, antigenic evolution studies should be updated repeatedly with
7 contemporary viruses.

8 **Retrospective antigenic evolution of H3N2 viruses**

9 To further test the potential of our experimental approach to identify future
10 antigenic drift of influenza viruses, we carried out a retrospective study with human
11 H3N2 viruses, whose antigenic evolution is well characterized^{8,13,24,25}. The antigenic
12 transition from the Sydney/1997 (SY97) cluster to the Fujian/2002 (FU02) cluster
13 required two amino acid changes: HA-H155T and HA-Q156H¹³. To assess whether this
14 cluster transition could be recapitulated experimentally, we generated a virus library with
15 random mutations in the HA globular head of A/Hong Kong/CUHK5250/2002
16 (CUHK5250, H3N2), a representative of the SY97 cluster that encodes HA-155H-156Q
17 (the virus library was generated with the NA gene of CUHK5250 virus and the internal
18 genes of A/Puerto Rico/8/34, H1N1). Our sequence analysis of human H3N2 viruses also
19 revealed several natural isolates (for example, A/Kwangju/219/2002; Kwangju/219), that
20 belong to the SY97 cluster, but encode the FU02-like residue HA-155T. This finding
21 suggests that the SY97-to-FU02 transition may have occurred in two steps: from 155H-
22 156Q to 155T-156Q to 155T-156H. Since evolutionary ‘advanced’ viruses (i.e., those

1 with an ‘intermediate’ amino acid sequence) may be the founders of the next antigenic
2 cluster, we also generated a virus library for Kwangju/219 (since the NA proteins of
3 Kwangju/219 and CUHK5250 differ by only two amino acids, the CUHK5250 NA gene
4 was used for Kwangju/219 library generation). *In vitro* screens for escape mutants were
5 carried out individually with 9 human sera collected in Japan in 1999 (for the CUHK5250
6 library; Supplementary Table 20), or with ferret antiserum to CUHK5250 (for the
7 Kwangju/219 library; Supplementary Table 21) to select escape mutants from the SY97
8 cluster; from these screens, we isolated 152 and 155 escape variants, respectively. The
9 antigenic screen of the CUHK5250 mutant library did not result in the isolation of the
10 cluster-transitioning H155T-Q156H mutations. By contrast, nine of the 155 variants
11 isolated from the Kwangju/219 mutant library acquired the cluster-defining HA-Q156H
12 mutation which, together with the HA-155T residue encoded by Kwangju/219, created
13 the FU02 cluster motif of HA-H155T-Q156H.

14 Next, HI assays were performed for 20 reference strains and for 32 and 56 mutant
15 viruses isolated from the CUHK5250 and Kwangju/219 screens, respectively (we focused
16 on viruses with mutations in amino acid positions known to affect antigenicity;
17 Supplementary Tables 20 and 21). All Kwangju/219 mutants that inherently encoded
18 HA-155T and acquired the HA-Q156H mutation localized to the edge of the FU02
19 cluster (Fig. 5, red triangles). Most CUHK5250 mutants did not shift to the FU02 cluster,
20 as expected by the absence of the HA-H155T-Q156H mutations. However, one
21 CUHK5250 mutant possessing HA-Q156H-H183L mutations mapped to the outer edge
22 of the FU02 cluster (Fig. 5, red rhombus). This variant is located in the antigenic map
23 close to the cluster of Kwangju/219 mutants encoding HA-155T-156H (Fig. 5, red

1 triangles). Hence, our *in vitro* antigenic selection identified the same antigenic drift
2 mutants that occurred in nature (i.e., HA-H155T-Q156H) and were critical for the
3 transition from the SY97 cluster to the FU02 cluster.

4 **Antigenic evolution of contemporary H3N2 viruses**

5 In February 2014, the WHO recommended an A/Texas/50/2012(H3N2)-like virus
6 (TX/50) as a vaccine strain for the 2014–2015 influenza season in the Northern
7 hemisphere. From March to September 2014, an increasing percentage of antigenically
8 drifted H3N2 variants belonging to clades 3C.3a or 3C.2a was reported²⁶. The antigenic
9 difference between TX/50 and the novel clade 3C.3a and 3C.2a viruses was most likely
10 correlated with HA-F159S and HA-F159Y mutations, respectively²⁷. The novel antigenic
11 clusters were dominant in most parts of the world by late 2014 and were responsible for
12 the low vaccine effectiveness of the H3N2 component of the 2014–2015 influenza
13 vaccine²⁸. In February 2015, the WHO therefore recommended replacing TX/50 with a
14 clade 3C.3a vaccine virus²⁸. During the first half of 2015, viruses of clade 3C.2a became
15 dominant over those of clade 3C.3a²⁹. In September 2015, the WHO therefore
16 recommended A/Hong Kong/4801/2014 (HK/4801; clade 3C.2a, encoding HA-159Y) as
17 a novel vaccine virus²⁹.

18 To identify antigenically advanced variants of contemporary human H3N2 viruses,
19 we generated a virus library possessing random amino acid changes at HA positions 63–
20 252 (H3 numbering) of TX/50, the WHO-recommended vaccine virus at that time. The
21 library was screened with 10 human antisera collected in December 2013–January 2014,
22 and with three mixtures of these antisera based on the age of the donor (Supplementary

1 Table 22). We identified 139 different potential escape mutants (Supplementary Tables
2 23 and 24); the most frequently mutated amino acid position was HA-159, consistent
3 with the role of this amino acid position in recent cluster transitions. For antigenic
4 analysis, we chose a subset of 26 antigenic mutants that encoded amino acid changes
5 frequently detected in these screens and/or amino acid changes known to affect
6 antigenicity (indicated by HI titres in Supplementary Tables 23 and 24). HI assays were
7 conducted in the presence of 20 nM oseltamivir with 19 reference antisera, and with 19
8 reference strains. The resulting HI data were merged with data from the CDC (Fig. 6).
9 The TX/50 antigenic cluster (encoding HA-159F; blue) and the minor 3C.3b cluster (also
10 encoding HA-159F, yellow) are antigenically closely related, whereas the 3C.3a cluster
11 (characterized by HA-159S; green) and the now dominant 3C.2a cluster (characterized by
12 HA-159Y; red) are antigenically distinguishable from TX/50. Among the experimentally
13 derived escape variants (shown in numbered mid-sized circles), the majority of
14 antigenically ‘advanced’ mutants [i.e., those that shifted from the vaccine virus by more
15 than two grid units (i.e., four HI units)] moved towards the currently dominant 3C.2a
16 cluster, represented by the A/Hong Kong/4801/2014 vaccine virus (#31); some of these
17 possessed the HA-159Y mutation. However, not all of the antigenic escape variants that
18 shifted towards the current cluster encode a mutation at HA-159. In fact, escape mutant
19 #27 (closest to the vaccine strain, #31) encodes mutations HA-K189E and -N225D,
20 which are known to affect antigenic and receptor-binding specificity, respectively^{13,30,31}.
21 The HA-N225D mutation is present in most viruses belonging to clade 3C.2a; however,
22 the HA-K189E mutation has not been detected in human H3N2 influenza viruses
23 collected from January 1, 2014 to February 20, 2016, based on sequence searches in

1 GISAID (www.platform.gisaid.org) and the Influenza Research Database
2 (www.fludb.org). Collectively, these findings indicate that although our approach may
3 not always identify the exact sequence of the next epidemic strains, it can predict the
4 antigenic location of the next antigenic cluster.

5

6 **Discussion**

7 With the current vaccine selection process (based on comparisons of sequence
8 and antigenic data available at the time of the WHO consultations on vaccine strain
9 selection), uncertainty remains as to whether the recommended vaccine virus will provide
10 protective immunity against the next epidemic strain. Here, we provide a methodology
11 that identified H1N1 and H3N2 antigenic clusters similar to those observed in nature.
12 These data can be generated before the variants are detected in nature; in this regard, our
13 approach is conceptually different from methods that predict which circulating variants
14 may become dominant³²⁻³⁵. Our method may therefore improve the current WHO
15 influenza vaccine selection process. These *in vitro* selection studies are highly predictive
16 of the antigenic evolution of H1N1 and H3N2 viruses in human populations. Hence, a
17 limited number of experimental antigenic screens may be sufficient to identify potential
18 future clusters. We found that genetic background can be important for antigenic
19 evolution, as with the SY97-FU02 antigenic cluster transition. To put methods like ours
20 into practice, it will likely be necessary to keep the predictions up to date by reapplying
21 the methodology as new clades emerge. We cannot predict the HA sequence of the next
22 epidemic strain, and (compensatory) mutations in NA may also affect antigenicity;

- 1 however, our methodology may be sufficient to predict the next antigenic cluster for
- 2 proactive vaccine development.

1 **Methods**

2 **Viruses and cells**

3 293T human embryonic kidney cells were maintained in DMEM supplemented
4 with 10% fetal bovine serum. MDCK cells were grown in MEM containing 5% newborn
5 calf serum. After infection with influenza viruses, MDCK cells were maintained in MEM
6 containing 0.3% BSA and 0.5 µg/ml TPCK-trypsin. A/Norway/3206-3/2009 (H1N1),
7 A/Norway/3858/2009 (H1N1), and A/Yokohama/UT-K1205T/2009 (H1N1) were
8 amplified in MDCK cells to generate stock viruses. A/Norway/3568/2009 (H1N1) and
9 A/California/04/2009 (H1N1) were generated using reverse genetics techniques as
10 described³⁶. To generate A/Hong Kong/CUHK5250/2002 (CUHK5250, H3N2), and
11 A/Texas/50/2012 (TX/50, H3N2; cell culture-grown; see #1 in Fig. 6) virus libraries, the
12 HA and NA genes, or just the HA gene, of these viruses were combined with the
13 remaining six or seven genes of A/Puerto Rico/8/34 (H1N1) virus. To generate
14 A/Kwangju/219/2002 (Kwangju/219, H3N2) virus libraries, we combined the
15 Kwangju/219 HA gene, the CUHK5250 NA gene (note that the NA proteins of
16 Kwangju/219 and CUHK5250 differ by only two amino acids), and the remaining genes
17 of A/Puerto Rico/8/34 virus. The cell lines used here were maintained in our laboratory
18 and were negative for mycoplasma contamination; they are not listed in the database of
19 commonly misidentified cell lines.

20

21 **Generation of virus libraries with random mutations in the HA globular head**

22 **region**

1 The cDNAs for the HA gene of A/Norway/3858/2009, CUHK5250,
2 Kwangju/219, and TX/50 viruses were randomly mutagenized using error-prone PCR
3 with the GeneMorph II kit (Stratagene, La Jolla, California, USA) (note that the sequence
4 of our CUHK5250 HA gene differed by two silent nucleotide changes from the sequence
5 reported in the database). The randomized coding regions for amino acid positions 54–
6 253 (for the H1 viruses; H1 numbering) or 63–252 (for the H3 viruses) of HA were
7 replaced with the corresponding region of the respective parental HA cDNA plasmid
8 using a PCR-based cloning technique. Unbiased library composition and the targeted
9 mutation rate (1-2 amino acid substitutions per molecule) were confirmed by sequence
10 analysis. The sizes of our *E. coli* libraries (9×10^6 , 1.3×10^7 , 1.4×10^7 , and 1.5×10^6
11 colony-forming units for A/Norway/3858/2009, CUHK5250, Kwangju/219, and TX/50
12 respectively) should represent all 7.96×10^6 combinations of two-amino acid changes
13 within a stretch of 200 amino acids. The randomized A/Norway/3858/2009 HA plasmid
14 library was then used to generate a virus library in the genetic background of the closely
15 related A/Norway/3568/2009 virus. The CUHK5250 and Kwangju/219 HA plasmid
16 libraries were used to generate virus libraries in the genetic background of the
17 CUHK5250 HA gene and the remaining A/Puerto Rico/8/34 virus genes. The TX/50 HA
18 plasmid library was used to generate a virus library in the genetic background of the 7
19 remaining A/Puerto Rico/8/34 virus genes.

20

21 **Generation of A(H1N1)pdm09 virus libraries representing all amino acids at a**
22 **defined position**

1 For random mutagenesis at a defined amino acid position, the HA plasmid was
2 amplified with degenerate PCR primers representing mixtures of all four nucleotides at
3 the three positions of the respective codon. The PCR products were self-ligated and
4 transformed into *E. coli*, resulting in 10^4 – 10^6 colony-forming units per library. Virus
5 libraries were generated as described above. The titres of the virus libraries were $3.4 \times$
6 10^4 – 7.0×10^7 PFU/ml (Supplementary Table 6).

7

8 **Generation of reassortant viruses possessing HA and NA genes derived from** 9 **A(H1N1)pdm09 field strains**

10 Plasmids for the expression of the HA and NA genes of field strains were
11 generated by site-directed PCR mutagenesis using cDNAs for the HA and NA genes of
12 A/Norway/3858/2009, A/Norway/3568/2009, or A/California/04/2009 as templates (for
13 A/Qingdao/1610/2009, whose NA gene sequence is not publicly available, the NA gene
14 of A/Norway/3568/2009 was used; for A/Miyazaki/40/2011 and A/Okinawa/41/2010,
15 whose HA2 sequences are not publicly available, chimeric HA genes with the HA2
16 sequence of A/Norway/3568/2009 were constructed and used). The resulting plasmids
17 were combined with plasmids encoding the remaining six genes of A/Norway/3568/2009
18 virus; viruses were generated as described above.

19

20 ***In vitro* selection of A(H1N1)pdm09 antigenic escape mutants from clinical isolates**

1 Clinical virus isolates (A/Norway/3206-3/2009 or A/Yokohama/UT-
2 K1205T/2009) were mixed with 2-fold serial dilutions of human convalescent sera and
3 incubated at 4°C overnight. After virus neutralization, virus-serum mixtures were
4 inoculated onto MDCK cells and incubated at 37°C for 1 h. After removal of the
5 inoculum, cells were cultured for 2–4 days. Then, culture supernatants were collected and
6 incubated with the highest concentration of antiserum at which a cytopathic effect was
7 observed. Viruses in the supernatants were plaque-purified in MDCK cells and the HA
8 genes of the purified viruses were sequenced.

9

10 ***In vitro* selection of antigenic escape mutants from A(H1N1)pdm09 or H3N2 virus** 11 **libraries**

12 Virus libraries were mixed with undiluted or 2-fold serial dilutions of the
13 respective ferret or human sera, with the exception of the H3N2 TX/50 virus library,
14 which was mixed with individual or mixed human sera, and incubated at 4°C overnight.
15 The virus-serum mixtures were diluted and subjected to plaque assays in MDCK cells
16 grown in 6-well plates. Antigenic escape variants were isolated from wells that showed
17 inhibition of virus infection (<10 plaques per well), and the HA genes of the isolated
18 viruses were sequenced.

19

20 **Animal experiments**

1 All animal experiments were approved by the Institutional Animal Care and Use
2 Committee (IACUC) at the University of Wisconsin-Madison. The facilities where this
3 research was conducted are fully accredited by the Association for the Assessment and
4 Accreditation of Laboratory Animal Care International (AAALAC). The animal
5 experiments described in this study were not designed to generate data sets for statistical
6 analysis; hence, the sample size was small and randomization and blinding were not
7 carried out.

8

9 ***In vivo* selection of A(H1N1)pdm09 antigenic escape mutants from virus libraries**

10 Six-week-old female BALB/c (Jackson Laboratory, Bar Harbor, Maine, USA) and
11 C57BL6/J (Jackson Laboratory) mice were used in this study. Three mice per group were
12 immunized by intranasal infection with 10 PFU (BALB/c mice) or 10^{1.5} PFU (C57BL6/J
13 mice) (50 µl) of A/California/04/09 virus. Four weeks later, we determined the HI titres
14 of mouse sera against A/California/04/09 virus. Then, animals were challenged by
15 intranasal infection with 10⁶ PFU (50 µl) of either A/California/04/09 virus (control) or
16 the virus library possessing the D127E mutation in HA and random mutations at amino
17 acid position 153–156 of HA. On day 4 post-challenge, virus titres in mouse lungs were
18 determined using plaque assays in MDCK cells. Viruses recovered from the lungs of
19 successfully immunized (HI antibody titres ≥40) and virus library-challenged mice were
20 plaque purified prior to HA sequence analysis.

21

22 **Challenge experiments in mice**

1 Six-week-old female C57BL6/J mice were immunized by intranasal infection
2 with $10^{1.5}$ PFU (50 μ l) of A/California/04/2009 virus or by two vaccinations (with a two-
3 week interval) of a commercial A/California/07/2009 HA split vaccine (CSL, Parkville,
4 Australia, 1.5 μ g in 50 μ l). Four weeks later, serum HI titres were determined. Three-to-
5 four mice per group with HI antibody titres ≥ 40 (for mice immunized by virus infection)
6 or ≥ 160 (for mice immunized by vaccination) were challenged by intranasal infection
7 with 10^6 PFU (50 μ l) of the oseltamivir-sensitive A/California/04/2009 or
8 A/Norway/3858/2009 control viruses, representative antigenic escape mutants, or
9 A/Norway/3568/2009-based reassortant viruses possessing changes in HA detected in
10 field strains. On day 4 post-challenge, virus titres in mouse lungs were determined using
11 plaque assays in MDCK cells

12

13 **Challenge experiments in ferrets**

14 Six- to ten-month-old female ferrets (Triple F Farms, Sayre, Pennsylvania, USA)
15 that were serologically negative by haemagglutination inhibition (HI) assay for currently
16 circulating human influenza viruses were infected by intranasal inoculation with 500 PFU
17 (500 μ l) of A/California/04/2009 virus. Twelve months later, serum HI titres were
18 determined. Three or four ferrets per group with HI antibody titres between 160 and 640
19 were challenged by intranasal infection with 10^7 PFU (500 μ l) of A/California/04/2009,
20 three antigenic escape mutants selected in the *in vivo* screens, or an A/Norway/3568/2009
21 reassortant virus with HA changes identified in a field strain (Fig. 4 and Supplementary
22 Fig. 11). The nasal washes of the challenged ferrets were collected daily from day 1 to

1 day 8 post-challenge. Virus titres in the nasal washes were determined using plaque assay
2 in MDCK cells.

3

4 **Generation of ferret antisera**

5 We used 6–10-month-old female ferrets (Triple F Farms, Sayre, Pennsylvania,
6 USA) that were serologically negative by haemagglutination inhibition (HI) assay for
7 currently circulating human influenza viruses. One or two ferrets per group were
8 anaesthetized intramuscularly with ketamine and xylazine (5–30 mg and 0.2–6 mg/kg of
9 body weight, respectively) and inoculated intranasally with 10^7 PFU (500 μ l) of wild-
10 type or mutant A(H1N1)pdm09, seasonal H1N1, swine H1N1, or human H3N2 virus.
11 Three-to-four weeks later, ferrets were euthanized and sera were collected for subsequent
12 use in HI assays or the selection of escape mutants.

13

14 **Antiserum treatment**

15 Human or ferret serum (100 μ l) was treated with 300 μ l of receptor-destroying
16 enzyme (RDE; Denka Seiken Co., Ltd., Tokyo, Japan) at 37°C for 18–24 h. To denature
17 the RDE solution, the serum-RDE mixture was incubated at 56°C for 30–60 min. After
18 denaturation, 600 μ l of sterile PBS and 100 μ l of pelleted turkey red blood cells (TRBCs)
19 were added and the mixture was incubated at room temperature for 60 min; during that
20 time, the suspension was mixed every 10 min. The mixture was then centrifuged at 3000
21 rpm at 4°C for 5 min. Supernatants were aliquoted and stored at -20°C for use in HI

1 assays. To confirm that all non-specific binding had been removed, 25 μ l of PBS and 25
2 μ l of treated serum were mixed with 50 μ l of 0.5% TRBCs and incubated at room
3 temperature for 45 min. If no hemagglutination was detected, the serum could be used for
4 HI assays; otherwise, the antiserum treatment was repeated.

5

6 **HI assays**

7 HI titrations were performed in 96-well U bottom microtiter plates (Thermo
8 Scientific, Rochester, New York, USA). Sera were serially diluted 2-fold with PBS and
9 then mixed with the amount of virus equivalent to 8 haemagglutination units, followed by
10 incubation at room temperature for 30 min. Finally, 50 μ l of 0.5% TRBCs were added.
11 To screen TX/50 variants, sera were serially diluted 2-fold with PBS containing a final
12 concentration of 20 nM oseltamivir, mixed with the amount of virus equivalent to 4
13 haemagglutination units, and incubated at room temperature for 30 min; finally, 50 μ l of
14 0.75% guinea pig red blood cells with 20 nM oseltamivir was added. Serum, virus, and
15 erythrocytes were gently mixed and incubated at room temperature for 45 min. Titres
16 were recorded as the inverse of the highest antibody dilution that inhibited 8 or 4
17 agglutinating units of virus, respectively.

18

19 **Antigenic cartography of escape mutants**

20 Antigenic maps were constructed using the antigenic cartography method⁸. The
21 HI titres were mathematically transformed to create a table of antigenic distances, using

1 the equation $D_{ij}=b-\log_2(H_{ij})$, where H_{ij} is the titre of antigen i against serum j , and b is the
2 logarithm of the maximum titre against serum j . The error function $(D_{ij}-d_{ij})^2$ was
3 minimized where d_{ij} is the Euclidean distance between two points on the map.

4

5 **Sequence analysis**

6 Viral RNA was extracted from MDCK cell supernatant or purified plaques using
7 the QIAmp viral RNA mini kit (Qiagen, Valencia, California, USA) or the MagMAX™-
8 96 Viral RNA Isolation Kit (Ambion, Austin, Texas, USA). cDNAs were synthesized
9 from vRNAs by reverse transcription with Uni12 primer³⁷, and amplified by PCR with
10 gene-specific primers. Sequencing was performed on an Applied Biosystems DNA
11 analyzer at the Biotechnology Center of the University of Wisconsin-Madison.

12

13 **Accession numbers**

14 The nucleotide sequences of the HA and NA genes of A/Norway/3568/2009
15 (H1N1), A/Norway/3858/2009 (H1N1) and A/Yokohama/UT-K1205T/2009 (H1N1)
16 determined in this study were deposited in GenBank (accession numbers CY189537-
17 CY189542).

18

19

1 **References**

2

- 3 1 Wright, P. F., Neumann, G. & Kawaoka, Y. *Fields Virology Vol. 2*
4 *(Editors David M. Knipe et al.) Ch. 48, 1691-1740 (Wolters Kluwer;*
5 *Lippincott Williams & Wilkins) (2007).*
- 6 2 Hay, A. J., Gregory, V., Douglas, A. R. & Lin, Y. P. The evolution of
7 human influenza viruses. *Philos Trans R Soc Lond B Biol Sci* **356**, 1861-
8 1870 (2001).
- 9 3 Carrat, F. & Flahault, A. Influenza vaccine: the challenge of antigenic drift.
10 *Vaccine* **25**, 6852-6862 (2007).
- 11 4 Wang, T. T. *et al.* Broadly protective monoclonal antibodies against H3
12 influenza viruses following sequential immunization with different
13 hemagglutinins. *PLoS Pathog* **6**, e1000796 (2010).
- 14 5 Russell, C. A. *et al.* Influenza vaccine strain selection and recent studies
15 on the global migration of seasonal influenza viruses. *Vaccine* **26 Suppl 4**,
16 D31-34 (2008).
- 17 6 Belongia, E. A. *et al.* Effectiveness of inactivated influenza vaccines varied
18 substantially with antigenic match from the 2004-2005 season to the 2006-
19 2007 season. *J Infect Dis* **199**, 159-167 (2009).
- 20 7 de Jong, J. C., Beyer, W. E., Palache, A. M., Rimmelzwaan, G. F. &
21 Osterhaus, A. D. Mismatch between the 1997/1998 influenza vaccine and
22 the major epidemic A(H3N2) virus strain as the cause of an inadequate
23 vaccine-induced antibody response to this strain in the elderly. *J Med Virol*
24 **61**, 94-99 (2000).
- 25 8 Smith, D. J. *et al.* Mapping the antigenic and genetic evolution of influenza
26 virus. *Science* **305**, 371-376 (2004).
- 27 9 Gerhard, W., Yewdell, J., Frankel, M. E. & Webster, R. Antigenic structure
28 of influenza virus haemagglutinin defined by hybridoma antibodies. *Nature*
29 **290**, 713-717 (1981).
- 30 10 Caton, A. J., Brownlee, G. G., Yewdell, J. W. & Gerhard, W. The antigenic
31 structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell*
32 **31**, 417-427 (1982).
- 33 11 Xu, R. *et al.* Structural basis of preexisting immunity to the 2009 H1N1
34 pandemic influenza virus. *Science* **328**, 357-360 (2010).
- 35 12 Wei, C. J. *et al.* Cross-neutralization of 1918 and 2009 influenza viruses:
36 role of glycans in viral evolution and vaccine design. *Sci Transl Med* **2**,
37 24ra21 (2010).
- 38 13 Koel, B. F. *et al.* Substitutions near the receptor binding site determine
39 major antigenic change during influenza virus evolution. *Science* **342**, 976-
40 979 (2013).
- 41 14 Webster, R. G. & Laver, W. G. Determination of the number of
42 nonoverlapping antigenic areas on Hong Kong (H3N2) influenza virus
43 hemagglutinin with monoclonal antibodies and the selection of variants
44 with potential epidemiological significance. *Virology* **104**, 139-148 (1980).

1 15 Fouchier, R. A. & Smith, D. J. Use of antigenic cartography in vaccine
2 seed strain selection. *Avian Dis* **54**, 220-223 (2010).

3 16 Sandbulte, M. R. *et al.* Discordant antigenic drift of neuraminidase and
4 hemagglutinin in H1N1 and H3N2 influenza viruses. *Proc Natl Acad Sci U*
5 *S A* **108**, 20748-20753 (2011).

6 17 Klimov, A. I. *et al.* WHO recommendations for the viruses to be used in the
7 2012 Southern Hemisphere Influenza Vaccine: epidemiology, antigenic
8 and genetic characteristics of influenza A(H1N1)pdm09, A(H3N2) and B
9 influenza viruses collected from February to September 2011. *Vaccine* **30**,
10 6461-6471 (2012).

11 18 Strengell, M., Ikonen, N., Ziegler, T. & Julkunen, I. Minor changes in the
12 hemagglutinin of influenza A(H1N1)2009 virus alter its antigenic
13 properties. *PLoS One* **6**, e25848 (2011).

14 19 *WHO Weekly epidemiological record*, 457-468 (2011).

15 20 Guarnaccia, T. *et al.* Antigenic drift of the pandemic 2009 A(H1N1)
16 influenza virus in A ferret model. *PLoS Pathog* **9**, e1003354 (2013).

17 21 O'Donnell, C. D. *et al.* Antibody pressure by a human monoclonal antibody
18 targeting the 2009 pandemic H1N1 virus hemagglutinin drives the
19 emergence of a virus with increased virulence in mice. *MBio* **3** (2012).

20 22 Rudneva, I. *et al.* Escape mutants of pandemic influenza A/H1N1 2009
21 virus: variations in antigenic specificity and receptor affinity of the
22 hemagglutinin. *Virus Res* **166**, 61-67 (2012).

23 23 Hensley, S. E. *et al.* Hemagglutinin receptor binding avidity drives
24 influenza A virus antigenic drift. *Science* **326**, 734-736 (2009).

25 24 Sun, H. *et al.* Using sequence data to infer the antigenicity of influenza
26 virus. *MBio* **4** (2013).

27 25 Bedford, T. *et al.* Integrating influenza antigenic dynamics with molecular
28 evolution. *Elife* **3**, e01914 (2014).

29 26 Recommended composition of influenza virus vaccines for use in the 2015
30 southern hemisphere influenza season. *Wkly Epidemiol Rec* **89**, 441-452
31 (2014).

32 27 Chambers, B. S., Parkhouse, K., Ross, T. M., Alby, K. & Hensley, S. E.
33 Identification of Hemagglutinin Residues Responsible for H3N2 Antigenic
34 Drift during the 2014-2015 Influenza Season. *Cell Rep* **12**, 1-6 (2015).

35 28 Recommended composition of influenza virus vaccines for use in the
36 2015-2016 northern hemisphere influenza season. *Wkly Epidemiol Rec*
37 **90**, 97-108 (2015).

38 29 Recommended composition of influenza virus vaccines for use in the 2016
39 southern hemisphere influenza season. *Wkly Epidemiol Rec* **90**, 545-559
40 (2015).

41 30 Lin, Y. P. *et al.* Evolution of the receptor binding properties of the influenza
42 A(H3N2) hemagglutinin. *Proceedings of the National Academy of*
43 *Sciences of the United States of America* **109**, 21474-21479 (2012).

44 31 Martin, J. *et al.* Studies of the binding properties of influenza
45 hemagglutinin receptor-site mutants. *Virology* **241**, 101-111 (1998).

1 32 Steinbruck, L., Klingen, T. R. & McHardy, A. C. Computational prediction
2 of vaccine strains for human influenza A (H3N2) viruses. *Journal of*
3 *virology* **88**, 12123-12132 (2014).

4 33 Steinbruck, L. & McHardy, A. C. Allele dynamics plots for the study of
5 evolutionary dynamics in viral populations. *Nucleic Acids Res* **39**, e4
6 (2011).

7 34 Luksza, M. & Lassig, M. A predictive fitness model for influenza. *Nature*
8 **507**, 57-61 (2014).

9 35 Neher, R. A. & Bedford, T. nextflu: real-time tracking of seasonal influenza
10 virus evolution in humans. *Bioinformatics* **31**, 3546-3548 (2015).

11 36 Neumann, G. *et al.* Generation of influenza A viruses entirely from cloned
12 cDNAs. *Proc Natl Acad Sci U S A* **96**, 9345-9350 (1999).

13 37 Hoffmann, E., Stech, J., Guan, Y., Webster, R. G. & Perez, D. R.
14 Universal primer set for the full-length amplification of all influenza A
15 viruses. *Arch Virol* **146**, 2275-2289 (2001).

16 38 Xu, R., McBride, R., Nycholat, C. M., Paulson, J. C. & Wilson, I. A.
17 Structural characterization of the hemagglutinin receptor specificity from
18 the 2009 H1N1 influenza pandemic. *J Virol* **86**, 982-990 (2012).

19

1 Correspondence and requests for materials should be addressed to:

2

3 Yoshihiro Kawaoka, DVM, Ph.D.

4 Influenza Research Institute, Department of Pathobiological Sciences, School of
5 Veterinary Medicine, University of Wisconsin-Madison, 575 Science Drive, Madison,
6 WI 53711, USA.

7 Phone: +1-608-265-4925

8 E-mail: kawaokay@svm.vetmed.wisc.edu

9

10 Derek J. Smith, Ph.D.

11 Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ,
12 UK.

13 Phone: +44-0-1223-330933

14 E-mail: dsmith@zoo.cam.ac.uk

15

16 **Acknowledgements**

17 We thank Susan Watson for scientific editing and Peter Jester for technical
18 support. This work was supported by the Bill & Melinda Gates Foundation Global Health
19 Grant OPPGH5383; National Institute of Allergy and Infectious Diseases (NIAID) Public
20 Health Service research grants (USA); ERATO (Japan Science and Technology Agency);
21 the Center for Research on Influenza Pathogenesis (CRIP) funded by the NIAID
22 Contracts HHSN266200700010C and HHSN272201400008C; the Japan Initiative for
23 Global Research Network on Infectious Diseases; Grants-in-Aid for Specially Promoted
24 Research from the Ministry of Education, Culture, Sports, Science, and Technology,
25 Japan; Grants-in-Aid from the Ministry of Health, Labour and Welfare, Japan; grants
26 from the Strategic Basic Research Program of the Japan Science and Technology

1 Agency; and by the Advanced Research & Development Programs for Medical
2 Innovation from the Japan Agency for Medical Research and Development (AMED).
3 C.A.R. was supported by a University Research Fellowship from the Royal Society. The
4 authors acknowledge a Netherlands Organisation for Scientific Research (NWO) VICI
5 grant, European Union (EU) FP7 programs EMPERIE (223498) and ANTIGONE
6 (278976); Human Frontier Science Program (HFSP) program grant P0050/2008;
7 Wellcome 087982AIA; and NIH Director's Pioneer Award DP1-OD000490-01. D.F.B
8 and D.J.S. acknowledge CamGrid, the University of Cambridge distributed computer
9 system. The Melbourne WHO Collaborating Centre for Reference and Research on
10 Influenza is supported by the Australian Government Department of Health. The origins
11 of the HA and NA gene sequences used in this study are recognized in Supplementary
12 Table 25. The content of this report is solely the responsibility of the authors and does not
13 necessarily represent the official views of the National Institutes of Health.

14 NOTE: The methods used in this manuscript could potentially select for variants
15 of A(H1N1)pdm09 viruses that could escape vaccine-based immunity; therefore, our
16 manuscript was reviewed by the CDC's Institutional Biosecurity Board which concluded
17 this study does not constitute Dual Use Research of Concern (DURC). The selection of
18 antigenic escape variants was completed before the US Government issued a Research
19 Funding Pause on October 17, 2014, on selected gain-of-function research on influenza,
20 MERS, and SARS viruses.

21

1 **Author Contributions**

2 C.L., M.H., D.F.B., J.P., Y.Z., M.O., A.S.T., S.C.D., A.P.H., J.S., M.I., P.R.W., T.W.,
3 S.W., M.I., K.I.H., C.A.R., S.L.J., E.S., E.A.M., G.N., A.I.K., A.K., J.M., D.W., Y.S.,
4 M.T., J. K., D.W., N.J.C., D.J.S., and Y.K. designed the experiments; C.L., M.H., D.F.B.,
5 J.P., Y.Z., M.O., A.S.T., S.C.D., A.P.H., J.S., M.I., P.R.W., T.W., S.W., M.I., K.I.H.,
6 C.A.R., S.L.J., T.O., X.X., and E.S. performed the experiments; C.L., M.H., D.F.B., J.P.,
7 Y.Z., M.O., A.S.T., S.C.D., A.P.H., J.S., M.I., P.R.W., T.W., S.W., M.I., K.I.H., C.A.R.,
8 S.L.J., E.S., E.A.M., G.N., A.I.K., A.K., J.M., D.W., Y.S., M.T., N.J.C., D.J.S., and Y.K.
9 analyzed the data; C.L., M.H., D.F.B., M.O., A.S.T., S.C.D., E.A.M., G.N., J.M.K.,
10 N.J.C., D.J.S., and Y.K. wrote the manuscript; C.L., M.H., D.F.B., J.P., and Y.Z.
11 contributed equally to this work.

12

13 **Author Information**

14 The authors declare no competing financial interests.

15

1 **Figure Legends**

2 **Figure 1 | Schematic overview of our experimental approach to predict the antigenic**
3 **evolution of human influenza A viruses.** Virus libraries possessing random or targeted
4 mutations in the antigenic region of HA are screened *in vitro* with antisera against human
5 2009 H1N1 viruses or *in vivo* in mice immunized with a human 2009 H1N1 virus.
6 Mutant viruses are analyzed by using haemagglutination inhibition (HI) assays, and
7 antigenic cartography is then used to identify viruses that differ antigenically from the
8 parent cluster.

9 **Figure 2 | HA antigenic cartography.**

10 An antigenic map, a geometrical representation of the HI assay data. If two viruses have
11 similar HI titres against a panel of sera, then they are close to each other on the map.
12 Conversely, if two viruses differ in their HI titres against a panel of sera, then they will be
13 far apart on the map. Each grid square corresponds to a 2-fold change in HI titre (the x
14 and y axes represent antigenic distances). **(a)** Map for the 447 mutant viruses (blue
15 circles) and 42 ferret antisera (squares) together with A/Norway/3858/2009 (cyan circle),
16 5 wild-type A(H1N1)pdm09 viruses (red circles), 10 A(H1N1)pdm09-based reassortant
17 viruses that possess changes in HA at positions 153–156 detected in field strains (orange
18 circles), 4 seasonal H1N1 viruses (magenta circles), and two H1N1 swine viruses (pink
19 circles). **(b)** Map focused on A(H1N1)pdm09 viruses (light blue circles) and ferret
20 antisera (squares) from this study. Viruses with mutations at positions 153–157 in HA are
21 colored by amino acid type at that position (note: viruses may have more than one
22 mutation; this can result in a change in position [see Supplementary Figs 2 and 5]). **(c)**

1 Antigenic map of 862 A(H1N1)pdm09 circulating strains from WHO surveillance in
2 2009 and 2010 (See key in **b**). For some circulating viruses, the identity of one or more
3 amino acids at positions 153–157 was unclear in sequence data (pale yellow circles).
4 Circulating viruses for which the HA sequence has not been determined are shown as
5 light grey circles. Fig. 2c is available as a labeled, zoomable pdf in the Supplementary
6 figures – see Supplementary Fig. 6a. **(d)** Antigenic map of the combined HI data from
7 Fig. 2b,c. This map includes A(H1N1)pdm09 viruses from the mutant screens in this
8 study (light blue circles); A(H1N1)pdm09-based reassortant viruses that possess changes
9 in HA at positions 153–156 detected in field strains (orange circles); and circulating
10 viruses from WHO surveillance studies with HA sequence data (pale orange circles) or
11 without HA sequence data (light grey circles). Circulating A(H1N1)pdm09 viruses tested
12 with a subset of sera from this study are shown as red circles. Fig. 2d is available as a
13 labeled, zoomable pdf in the Supplementary figures – see Supplementary Fig. 6f.

14 **Figure 3 | Structural basis of antigenic changes.**

15 Structure of the HA of A/California/04/2009; PDB 3UBN³⁸. Amino acid positions at
16 which single mutations result in appreciable antigenic change are shown as increasingly
17 darker shades of green. The structure shown is a co-crystal of HA and a human-type
18 receptor analog, NeuAc α 2-6Gal β 1-4GlcNAc β 1-3Gal β 1-4Glc (LSTc; shown in red).

19 **Figure 4 | Immune evasion in ferrets by HA mutations identified in this study or**
20 **found in a field strain.**

21 Ferrets were infected intranasally by inoculation with 500 PFU of A/California/04/2009
22 virus. Twelve months later, serum HI titres against A/California/04/2009 were

1 determined (Supplementary Fig. 9) and ferrets were challenged with 10^7 PFU of
2 A/California/04/2009 (**a**), with representative antigenic escape mutants selected in the *in*
3 *vivo* screens (**b, c, d**), or with a A/Norway/3568/2009-based reassortant virus possessing
4 changes in HA at positions 155 and 156 detected in a field strain (**e**) (see Supplementary
5 Tables 18 and 19). Shown are nasal wash virus titers collected from day 1 to day 5 post-
6 challenge (all nasal wash samples collected on days 6–8 post-challenge were negative for
7 virus). Dashed lines, virus detection limit ($1 \log_{10}$ PFU/ml).

8 **Figure 5 | Antigenic map of CUHK5250 and Kwangju/219 H3N2 escape mutants.**

9 The antigenic map was generated as described in the legend to Fig. 2. Wild-type strains
10 belonging to the Wuhan 1995, Sydney 1997, or Fujian 2002 clusters are shown in green,
11 magenta, or yellow, respectively. Escape mutants selected from the CUHK5250 and
12 Kwangju/219 libraries are shown as diamonds and triangles, respectively, and colored by
13 the presence of a mutation at position 145 (cyan), 155 (orange), 156 (red), 158 (green),
14 189 (cream), or 193 (wheat), which were recently found to be antigenically important for
15 human H3N2 virus evolution¹³.

16 **Figure 6 | Antigenic map of TX/50 H3N2 escape mutants.**

17 The antigenic map was generated as described in the legend to Fig. 2. Surveillance
18 samples are shown in small circles; experimentally derived antigenic variants are shown
19 in mid-sized circles; vaccine viruses are shown in large circles. Surveillance samples
20 characterized by HI assays, but not sequenced, are indicated in gray. Shown in yellow are
21 viruses belonging to the minor cluster 3C.3b. The TX/50 cluster is indicated in blue.

1 Antigenic escape mutations encoding HA-159F (as found for TX/50) are shown in cyan.
2 The 3C.3a and 3C.2a clusters (characterized by HA-159S and HA-159Y, respectively)
3 are indicated in green and red, respectively. Viruses are numbered as follows: #1, TX/50
4 MK1/MDCK2 (used for virus library generation); #2, TX/50 HA-E172G+F193S; #3;
5 TX/50 HA-F193S+T203A; #4, TX/50 HA-S219F; #5, TX/50 HA-I217T+N225K; #6,
6 TX/50 HA-Q75L+N144I+I192T; #7, TX/50 HA-F174Y; #8, TX/50 HA-Y94F+N225D;
7 #9, TX/50 HA-R220G+N225D; #10, TX/50 HA-N158D; #11, TX/50 HA-Q197R; #12,
8 TX/50 HA-K207N+N225S; #13, TX/50 HA-W127L+N225D; #14, TX/50 HA-
9 Y94H+S107T+N225D; #15, TX/50 HA-V88I+K189N; #16, TX/50 HA-
10 F159Y+K160E+I192V+I242T; #17, TX/50 HA-F159Y; #18, TX/50 HA-L157F+N225D;
11 #19, TX/50 HA-N144I+N158D+F193Y; #20, TX/50 HA-N225D; #21, TX/50 HA-
12 K189E; #22, TX/50 HA-K189E+I242M; #23, TX/50 HA-N122S+F159Y; #24, TX/50
13 HA-E172D+K189E; #25, TX/50 HA-K189E+F193S; #26, TX/50 HA-
14 N128D+F159Y+N225D; #27, TX/50 HA-K189E+N225D; #28, TX/50 HA-
15 H156R+F159S+N246H/WT-NA (propagation and re-generation of viruses possessing
16 HA-F159S+N225D failed; when we attempted to generate a PR8 virus with the NA gene
17 of TX/50 and the TX/50 HA-F159S gene, the virus also acquired HA-H156R and N246H
18 mutations); #29, TX/50 HA-F159S+N246H (propagation and re-generation of viruses
19 possessing HA-F159S+N225D failed; when we attempted to generate a PR8 virus with
20 the TX/50 HA-F159S gene, the virus also acquired a N246H mutation); #30,
21 A/Texas/50/2012 (E5; vaccine strain); #31, A/Hong Kong/4801/2014 (NYMC-263A;
22 vaccine strain); and #32, A/Switzerland/9715293/2013-NIB-88 (E4/E5/E2; vaccine
23 strain).

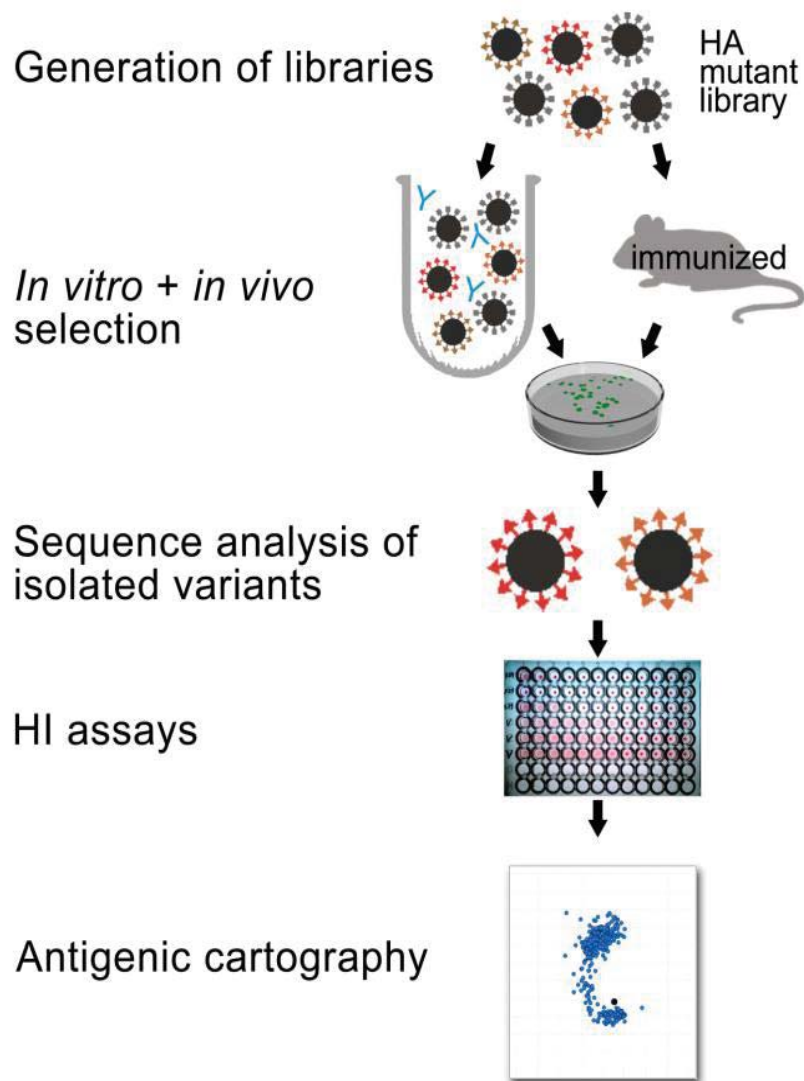
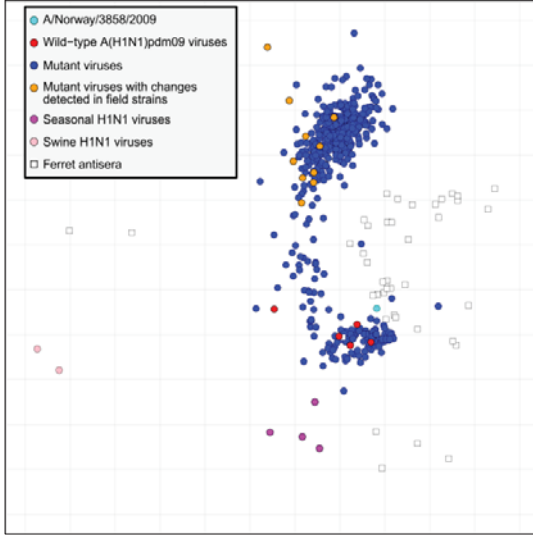
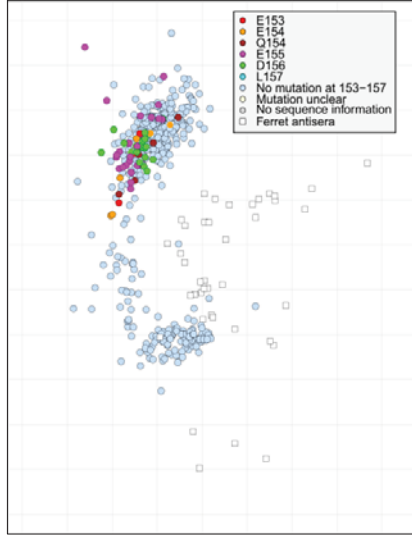


Figure 1 | Schematic overview of our experimental approach to predict the antigenic evolution of human influenza A viruses. Virus libraries possessing random or targeted mutations in the antigenic region of HA are screened *in vitro* with antisera against human 2009 H1N1 viruses or *in vivo* in mice immunized with a human 2009 H1N1 virus. Mutant viruses are analyzed by using haemagglutination inhibition (HI) assays, and antigenic cartography is then used to identify viruses that differ antigenically from the parent cluster.

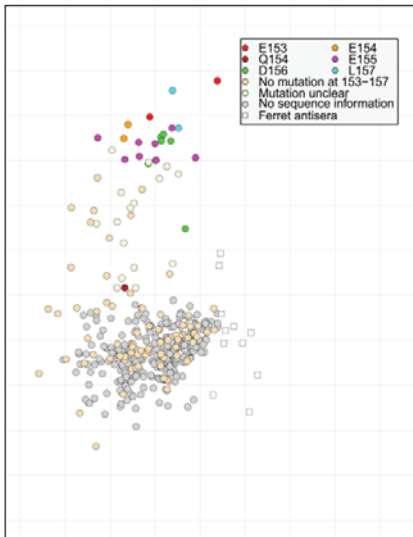
a



b



c



d

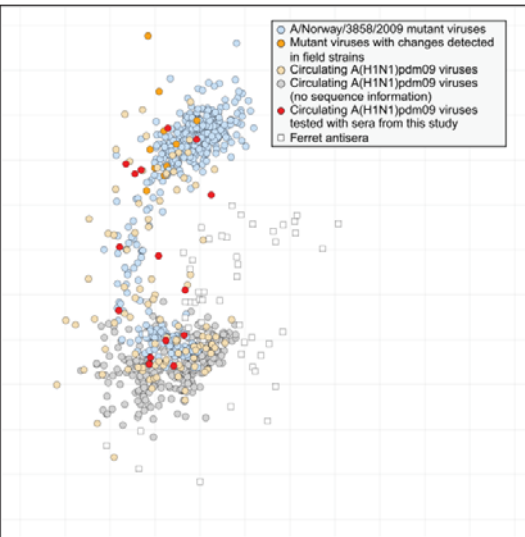


Figure 2 | HA antigenic cartography. An antigenic map, a geometrical representation of the HI assay data. If two viruses have similar HI titres against a panel of sera, then they are close to each other on the map. Conversely, if two viruses differ in their HI titres against a panel of sera, then they will be far apart on the map. Each grid square corresponds to a 2-fold change in HI titre (the x and y axes represent antigenic distances). **(a)** Map for the 447 mutant viruses (blue circles) and 42 ferret antisera (squares) together with A/Norway/3858/2009 (cyan circle), 5 wild-type A(H1N1)pdm09 viruses (red circles), 10 A(H1N1)pdm09-based reassortant viruses that possess changes in HA at positions 153–156 detected in field strains (orange circles), 4 seasonal H1N1 viruses (magenta circles), and two H1N1 swine viruses (pink circles). **(b)** Map focused on A(H1N1)pdm09 viruses (light blue circles) and ferret antisera (squares) from this study. Viruses with mutations at positions 153–157 in HA are colored by amino acid type at that position (note: viruses may have more than one mutation; this can result in a change in position [see Supplementary Figs 2 and 5]). **(c)** Antigenic map of 862 A(H1N1)pdm09 circulating strains from WHO surveillance in 2009 and 2010 (See key in **b**). For some circulating viruses, the identity of one or more amino acids at positions 153–157 was unclear in sequence data (pale yellow circles). Circulating viruses for which the HA sequence has not been determined are shown as light grey circles. Fig. 2c is available as a labeled, zoomable pdf in the Supplementary figures – see Supplementary Fig. 6a. **(d)** Antigenic map of the combined HI data from Fig. 2b,c. This map includes A(H1N1)pdm09 viruses from the mutant screens in this study (light blue circles); A(H1N1)pdm09-based reassortant viruses that possess changes in HA at positions 153–156 detected in field strains (orange circles); and circulating viruses from WHO surveillance studies with HA sequence data (pale orange circles) or without HA sequence data (light grey circles). Circulating A(H1N1)pdm09 viruses tested with a subset of sera from this study are shown as red circles. Fig. 2d is available as a labeled, zoomable pdf in the Supplementary figures – see Supplementary Fig. 6f.

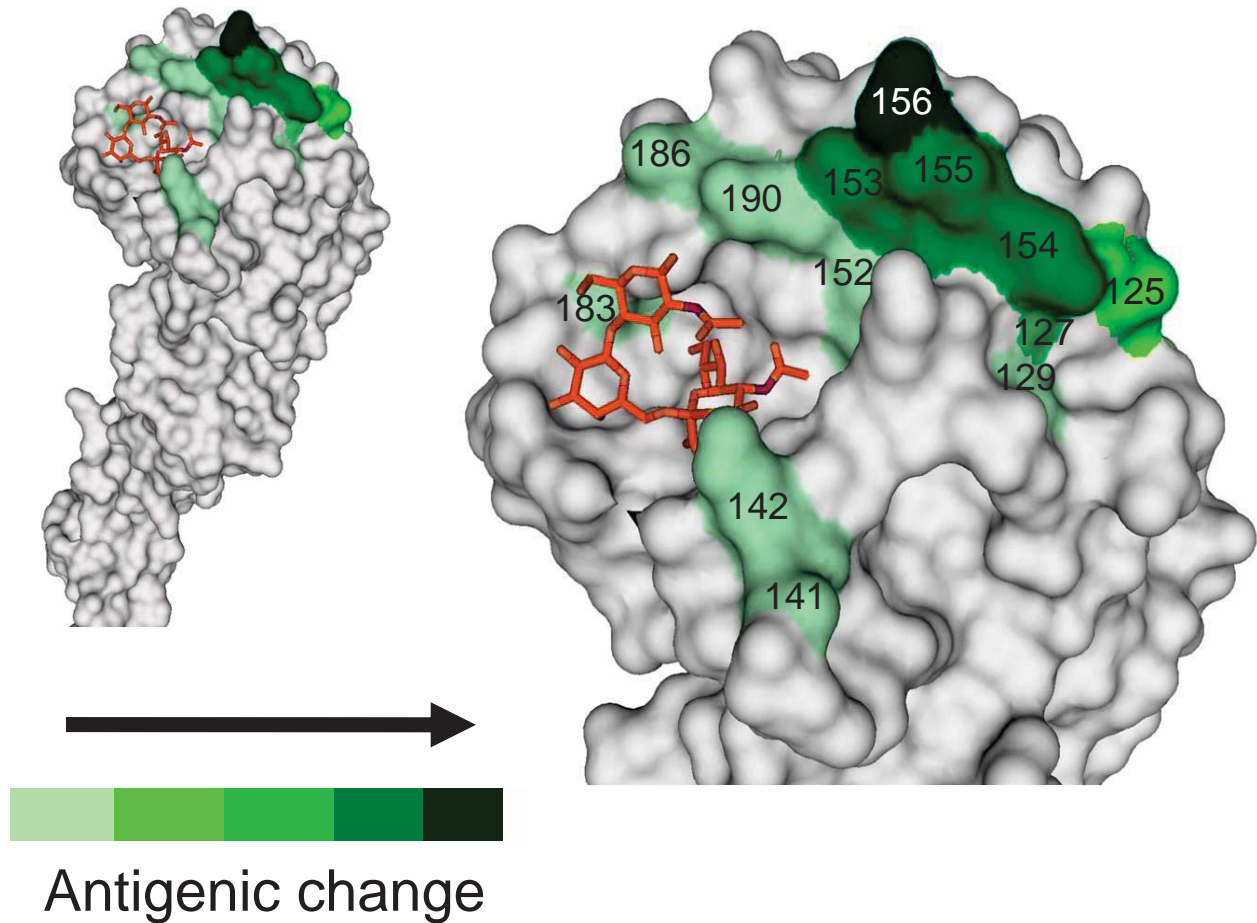


Figure 3 | Structural basis of antigenic changes. Structure of the HA of A/California/04/2009; PDB 3UBN³⁸. Amino acid positions at which single mutations result in appreciable antigenic change are shown as increasingly darker shades of green. The structure shown is a co-crystal of HA and a human-type receptor analog, NeuAc α 2-6Gal β 1-4GlcNAc β 1-3Gal β 1-4Glc (LSTc; shown in red).

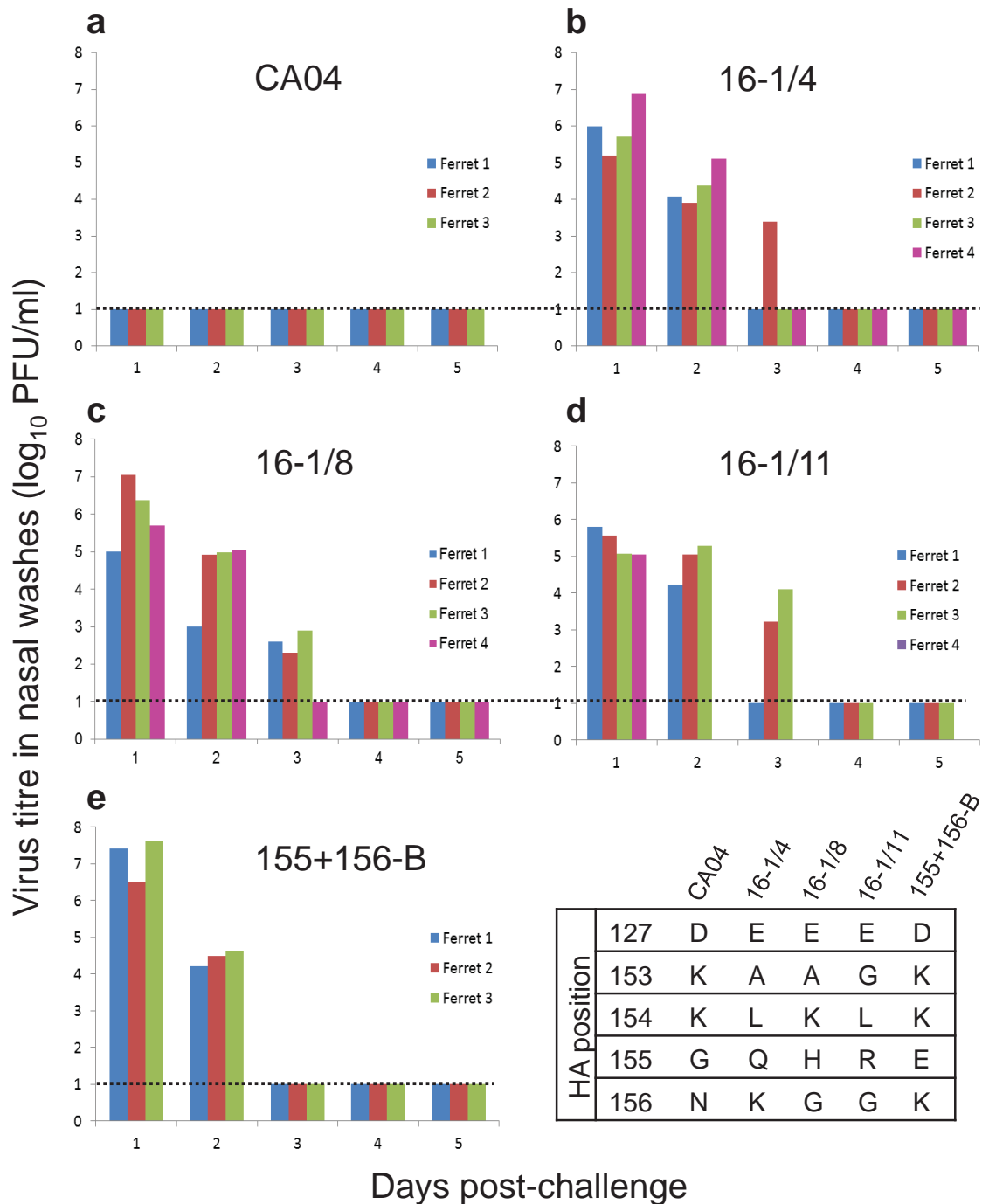


Figure 4 | Immune evasion in ferrets by HA mutations identified in this study or found in a field strain. Ferrets were infected intranasally by inoculation with 500 PFU of A/California/04/2009 virus. Twelve months later, serum HI titres against A/California/04/2009 were determined (Supplementary Fig. 9) and ferrets were challenged with 10^7 PFU of A/California/04/2009 (a), with representative antigenic escape mutants selected in the *in vivo* screens (b, c, d), or with a A/Norway/3568/2009-based reassortant virus possessing changes in HA at positions 155 and 156 detected in a field strain (e) (see Supplementary Tables 18 and 19). Shown are nasal wash virus titers collected from day 1 to day 5 post-challenge (all nasal wash samples collected on days 6–8 post-challenge were negative for virus). Dashed lines, virus detection limit ($1 \log_{10}$ PFU/ml).

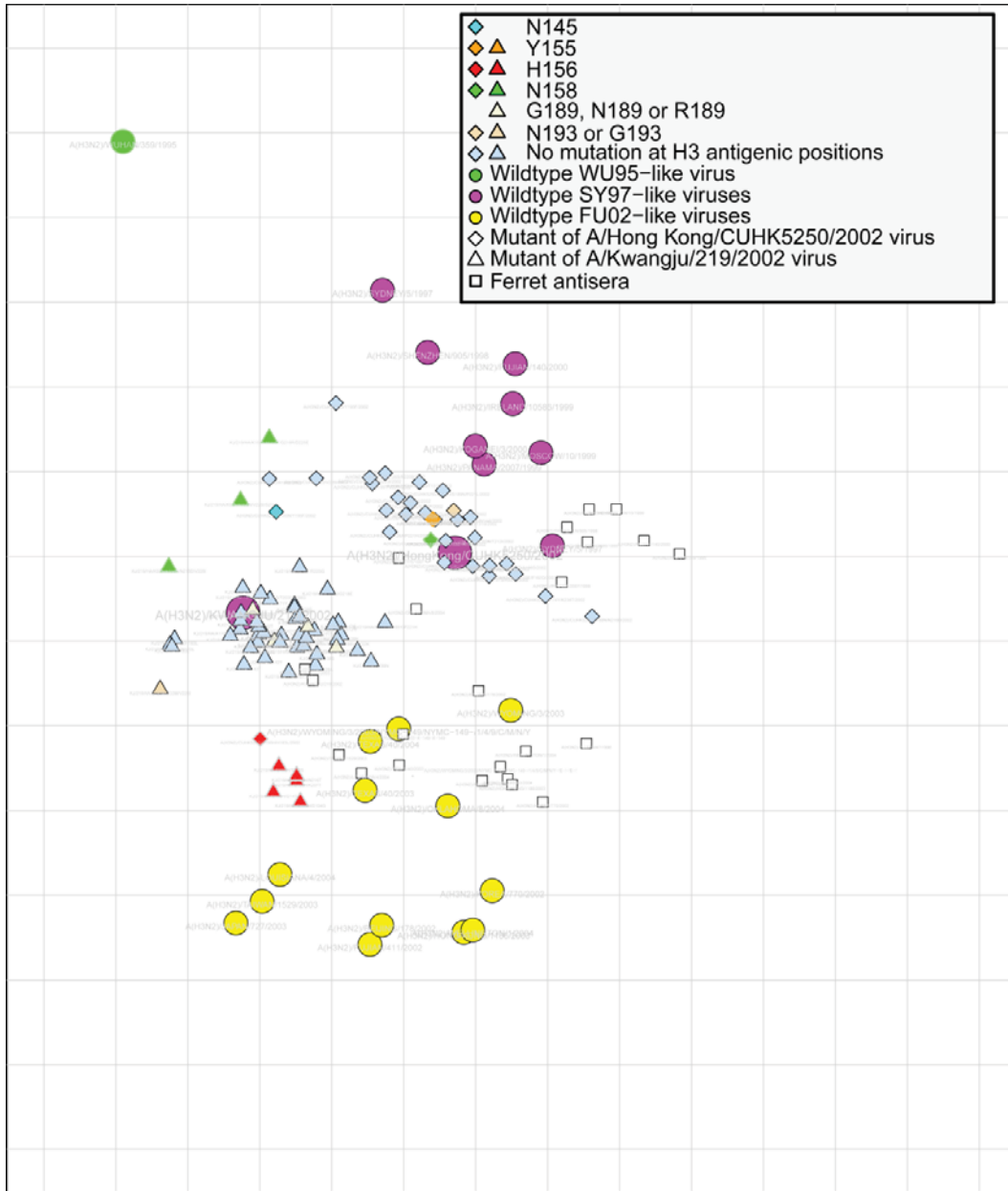


Figure 5 | Antigenic map of CUHK5250 and Kwangju/219 H3N2 escape mutants.
 The antigenic map was generated as described in the legend to Fig. 2. Wild-type strains belonging to the Wuhan 1995, Sydney 1997, or Fujian 2002 clusters are shown in green, magenta, or yellow, respectively. Escape mutants selected from the CUHK5250 and Kwangju/219 libraries are shown as diamonds and triangles, respectively, and colored by the presence of a mutation at position 145 (cyan), 155 (orange), 156 (red), 158 (green), 189 (cream), or 193 (wheat), which were recently found to be antigenically important for human H3N2 virus evolution¹³.

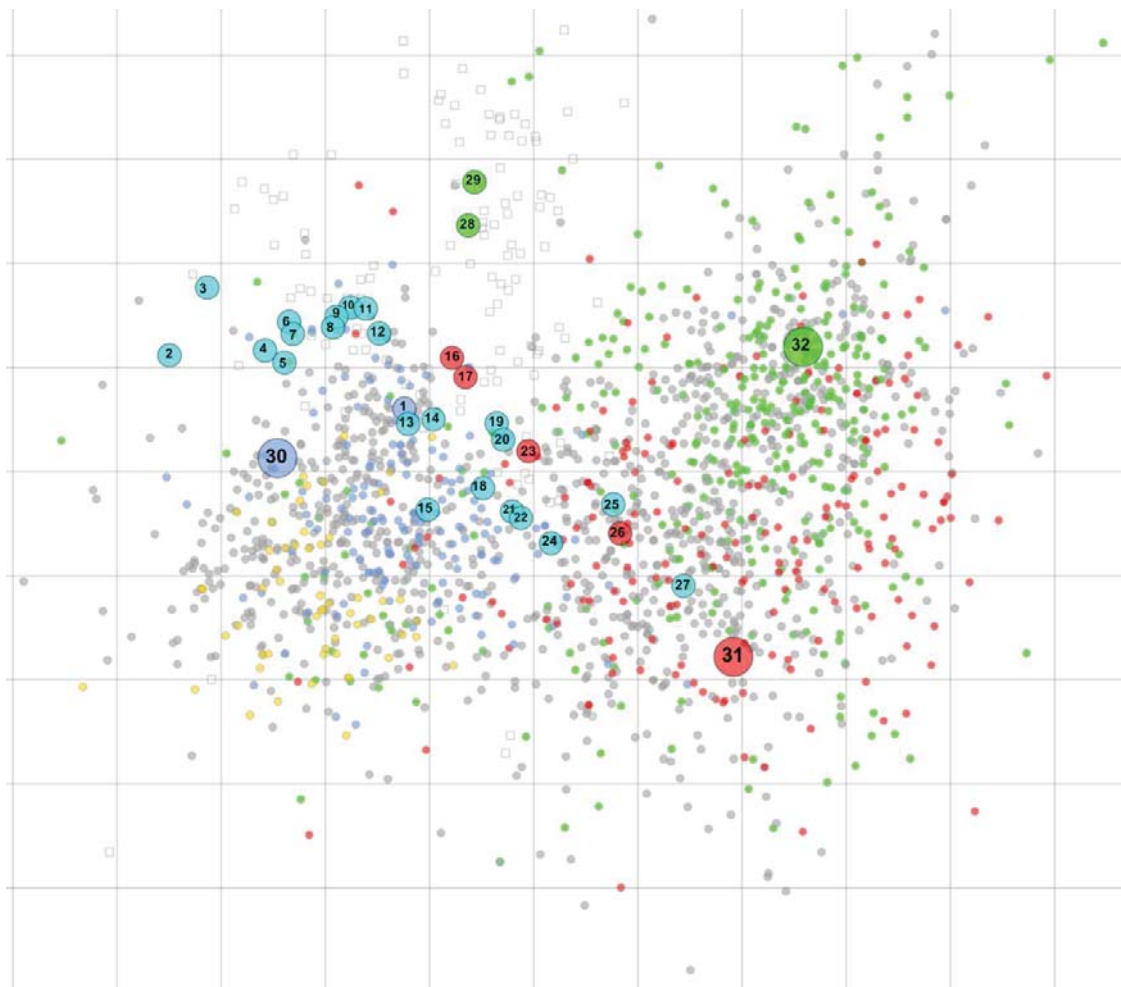


Figure 6 | Antigenic map of TX/50 H3N2 escape mutants. The antigenic map was generated as described in the legend to Fig. 2. Surveillance samples are shown in small circles; experimentally derived antigenic variants are shown in mid-sized circles; vaccine viruses are shown in large circles. Surveillance samples characterized by HI assays, but not sequenced, are indicated in gray. Shown in yellow are viruses belonging to the minor cluster 3C.3b. The TX/50 cluster is indicated in blue. Antigenic escape mutations encoding HA-159F (as found for TX/50) are shown in cyan. The 3C.3a and 3C.2a clusters (characterized by HA-159S and HA-159Y, respectively) are indicated in green and red, respectively. Viruses are numbered as follows: #1, TX/50 MK1/MDCK2 (used for virus library generation); #2, TX/50 HA-E172G+F193S; #3, TX/50 HA-F193S+T203A; #4, TX/50 HA-S219F; #5, TX/50 HA-I217T+N225K; #6, TX/50 HA-Q75L+N144I+I192T; #7, TX/50 HA-F174Y; #8, TX/50 HA-Y94F+N225D; #9, TX/50 HA-R220G+N225D; #10, TX/50 HA-N158D; #11, TX/50 HA-Q197R; #12, TX/50 HA-K207N+N225S; #13, TX/50 HA-W127L+N225D; #14, TX/50 HA-Y94H+S107T+N225D; #15, TX/50 HA-V88I+K189N; #16, TX/50 HA-F159Y+K160E+I192V+I242T; #17, TX/50 HA-F159Y; #18, TX/50 HA-L157F+N225D; #19, TX/50 HA-N144I+N158D+F193Y; #20, TX/50 HA-N225D; #21, TX/50 HA-K189E; #22, TX/50 HA-K189E+I242M; #23, TX/50 HA-N122S+F159Y; #24, TX/50 HA-E172D+K189E; #25, TX/50 HA-K189E+F193S; #26, TX/50 HA-N128D+F159Y+N225D; #27, TX/50 HA-K189E+N225D; #28, TX/50 HA-H156R+F159S+N246H/WT-NA (propagation and re-generation of viruses possessing HA-F159S+N225D failed; when we attempted to generate a PR8 virus with the NA gene of TX/50 and the TX/50 HA-F159S gene, the virus also acquired HA-H156R and N246H mutations); #29, TX/50 HA-F159S+N246H (propagation and re-generation of viruses possessing HA-F159S+N225D failed; when we attempted to generate a PR8 virus with the TX/50 HA-F159S gene, the virus also acquired a N246H mutation); #30, A/Texas/50/2012 (E5; vaccine strain); #31, A/Hong Kong/4801/2014 (NYMC-263A; vaccine strain); and #32, A/Switzerland/9715293/2013-NIB-88 (E4/E5/E2; vaccine strain).