Identification of an epitope of limited variability under strong immune selection in the haemagglutinin head domain of H1N1 influenza

Craig P Thompson[1,2*], José Lourenço[1], Adam A Walters[2], Uri Obolski[1], Matthew Edmans[1,2], Duncan S Palmer[1], Kreepa Kooblall[3], George W Carnell[4], Daniel O'Connor[5], Thomas A Bowden[6], Oliver G Pybus[1], Andrew J Pollard[5], Nigel J Temperton[4], Teresa Lambe[2], Sarah C Gilbert[2], Sunetra Gupta[1]

[1]Department of Zoology, University of Oxford, UK, [2]The Jenner Institute Laboratories, University of Oxford, UK, [3]Oxford Centre for Endocrinology, Metabolism and Diabetes, University of Oxford, UK, [4]Medway School of Pharmacy, University of Kent, UK, [5]Oxford Vaccine Group, Department of Paediatrics, University of Oxford, and the NIHR Oxford Biomedical Research Centre, Oxford, UK, [6]Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, UK.

Antigenic targets of influenza vaccination are currently seen to be polarised between (i) highly immunogenic (and protective) epitopes of high variability, and (ii) conserved epitopes of low immunogenicity. This requires vaccines directed against the variable sites to be continuously updated, with the only other alternative being seen as the artificial boosting of immunity to invariant epitopes of low natural efficacy. However, theoretical models suggest that the antigenic evolution of influenza is best explained by postulating the existence of highly immunogenic epitopes of limited variability. A corollary of this model is that a 'universal' influenza vaccine may be constructed by identifying such protective epitopes of low variability and these vaccines would also have the potential to protect against newly emerging influenza strains. Here we report the identification of such an epitope of limited variability in the H1 head domain of the haemagglutinin protein that could be exploited to produce a universal vaccine for the H1 subtype of influenza A.

The antigenic evolution of influenza A viruses is known to occur through mutations in surface glycoproteins, principally in haemagglutinin (HA), allowing strains to escape preexisting host immunity (*1–3*). Epitopes within HA are commonly assumed to be either highly variable due to strong immune selection (and typically located in the head domain of HA) or of very limited variability due to the absence of immune selection (for example, in the stalk of HA). However, our theoretical studies suggest that the antigenic evolution of influenza may be primarily driven by naturally

protective immune responses against epitopes of limited variability located in the head region of HA (*4*, *5*). Under these circumstances, although new strains may be generated constantly through mutation, most of these cannot expand in the host population due to pre-existing immune responses against their less variable epitopes. This accounts for the limited antigenic and genetic diversity observed within an influenza epidemic and also creates the conditions for the sequential (rather than simultaneous) appearance of antigenic types.

We tested the prediction that epitopes of limited variability exist on HA by performing microneutralisation assays against a panel of historical H1N1 influenza isolates using pseudotyped lentiviruses displaying the associated HA protein (hereafter described as pMN assays (*6*, *7*) using sera obtained in 2006/2007 in the UK from 88 children born between March 1994 and May 2000 (Fig 1A). All individuals possessed neutralising antibodies to the A/Solomon Islands/3/2006 strain (belonging to the H1N1 cluster circulating in 1999-2008), and 98% of individuals were able to neutralise A/New Caledonia/20/1999 (found within the same cluster) including those children born in 2000 who may not have been naturally exposed to this strain. Neutralisation of A/USSR/90/1977 was also extremely common and 60% of the serum samples were able to neutralise A/WSN/1933. By contrast, only 3% of individuals possessed neutralising antibodies against A/California/4/2009, and only 3.4% and 9.1% of the serum samples neutralised A/South Carolina/1/1918 and A/PR/8/1934 respectively. Antigen specific HA1 ELISA data against the same seven isolates was consistent with the pMN data and also potentially identified broadly cross-reactive non-neutralising antibodies that bind the HA1 domain of various H1 influenza strains (Fig S1). A number of other studies have suggested antibody responses show some degree of periodic cross-reactivity in agreement with these results (*8–15*).

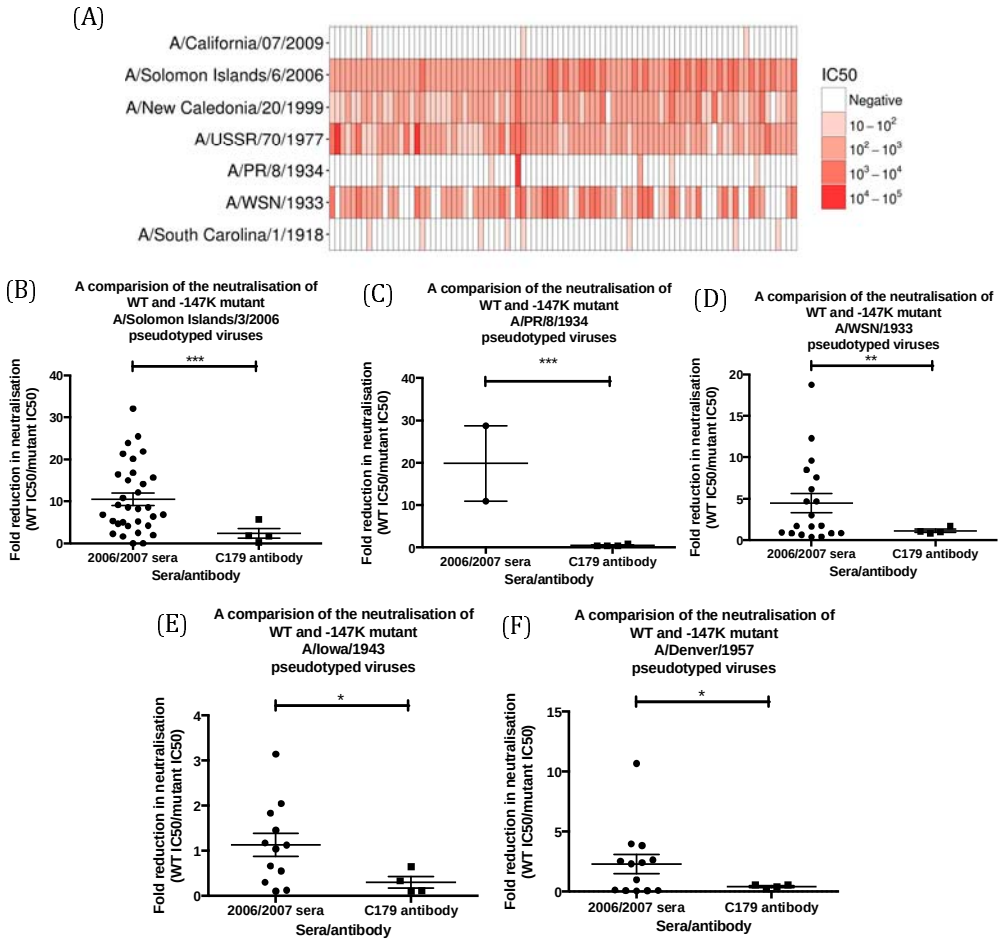


Figure 1: Pseudotype microneutralisation data revealing a cyclic pattern of epitope recognition and the involvement of position 147 in the production of cross-protective antibodies in sera taken from children aged 6 to 12 in 2006/2007. (A) Serum samples from children aged between 6 to 12 years in 2006/2007 (n = 88) were tested for their ability to neutralise a panel of pseudotyped lentiviruses representing a range of historical isolates. (B-F) A lysine residue was inserted at position 147 (linear numbering, where Met = 1) through site-directed mutagenesis in the HAs of pseudotyped lentiviruses A/WSN/1933, A/PR/8/1934 and A/Solomon Islands/3/2006 (included in panel A) as well as A/Iowa/1943 and A/Denver/1957. The ratio of WT IC50 to mutant IC50 was then assessed to determine if there was reduction in neutralisation. Graphs show mean and standard deviation, n=2 (*** $p<0.001$, ** $p<0.01$, * $p<0.05$). A stalk targeting antibody, C179, was used as a control.

We noted that A/Solomon Islands/3/2006, A/New Caledonia/20/1999, A/PR/8/34 and A/WSN/33 all contained a deletion at position 147 (linear numbering, where Met = 1), which otherwise typically contains a positively charged amino acid, as is the case for A/California/4/2009 A/USSR/90/1977 and A/Brevig Mission/1/1918. To determine whether the cross-reactivity observed between these strains could be attributed to this feature, we performed site-directed mutagenesis (SDM) by inserting

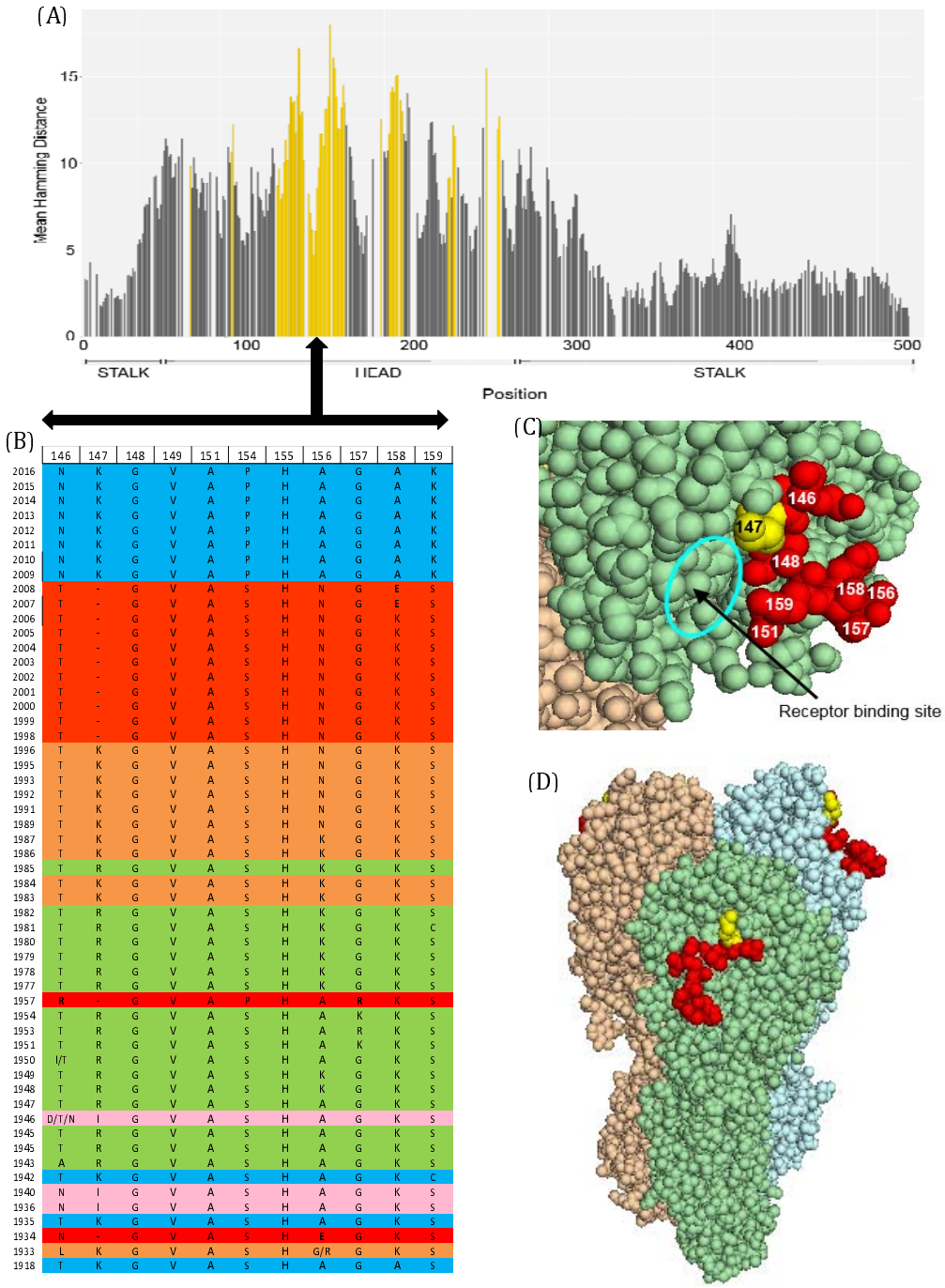
a lysine at position 147 of the A/Solomon Islands/6/2006, A/PR/8/1934 and A/WSN/1933 sequences (Fig 1B-D). Using a microneutralisation assay, up to a 32-fold loss of neutralisation of the A/Solomon Islands/3/2006 pseudotyped lentivirus was observed when a lysine was inserted at position 147 (p-value: 0.0005, Fig 1B). Up to a 18.75-fold loss of neutralisation of the A/WSN/1933 pseudotyped lentivirus was also observed with the insertion of a lysine at position 147 (p-value: 0.0056; Fig. 1D). When a lysine was inserted at position 147 in the HA of A/PR/8/1934, there was a total loss of neutralisation in 4 samples and a reduction in 2 samples indicating that the bulk of cross-reactivity between the A/Solomon Islands/3/2006 and the A/PR/8/1934 strains is mediated through an epitope that contains a deletion at position 147 (p-value: 0.0004, Fig. 1C).

To ascertain whether the absence of an amino acid at position 147 would also be responsible for cross-reactivity with other pseudotyped viruses containing the same deletion, a lysine residue was added at position 147 in the HAs from the A/Iowa/1943 and A/Denver/1957 viruses (Fig 1E & F). Samples that were positive for neutralisation of A/WSN/1933 were used to determine whether a change in neutralisation occurred in these viruses. This resulted in up to a 3-fold loss of neutralisation for A/Iowa/1943 (p-value: 0.012, Fig 1E.). When a lysine was inserted at position 147 in the A/Denver/1957 HA, three samples failed to neutralise the mutant entirely and up to a 12-fold loss of neutralisation was observed for those that did neutralise (p-value: 0.011, Fig 1F). These results imply that at least part of the cross-reactive neutralising immune response within this cohort is mediated through the recognition of an epitope that contains a deletion at position 147 and that the existence of a lysine at position 147 may contribute to the overall lack of neutralisation of A/California/4/2009 and A/South Carolina/1/1918.

Although not included within any of the canonical antigenic sites defined by Caton et al, 1982 (being absent in the A/PR/8/1934 (Mt. Sinai) strain), position 147 has recently been assigned to a new antigenic site denoted ‘Pa’ in Matsuzaki et al, 2014 where it was shown to be responsible for several A/Narita/1/2009 escape mutants (*1*, *16*). Similar 2009 pandemic strain escape mutants have also been identified in Huang et al, 2015 (*14*). Position 147 is also important for the binding of several known

neutralising antibodies: for example, the 5j8 antibody requires a lysine to be present at position 147, whilst the CH65 antibody cannot bind if a lysine is present at position 147 (*17*, *18*). Furthermore, Li et al, 2013 have demonstrated that certain demographics, such as individuals born between 1983 and 1996, possess antibodies that bind to an epitope containing a lysine residue at position 147 (*12*).

We next employed a structural bioinformatic approach in attempt to identify an epitope of limited variability that contained position 147. *In silico* analysis was used to determine how the accessibility and binding site area contributed to the variability of hypothetical antibody binding sites (Fig 2A; Fig S2) for the A/Brevig Mission/1/1918, A/PR/8/1934, A/California/04/2009, A/Washington/5/2011 H1 HA crystal structures (*19–23*). The antibody binding site (ABS) of lowest variability containing position 147 was consistently represented by the disrupted peptide sequence shown in Figure 2B, and could be shown to locate to an exposed loop in the head domain of the H1 HA, not covered by N-linked glycosylation (Figs 2C-D & Fig S3**)**. Analysis of this epitope (hereafter called OREO) suggested that these could be categorised on the basis of variation in positions 147, 156, 157, 158 and 159. Combining these analyses with the SDM results, we arrived at a maximum of 5 functional "allelic" classes or conformations of OREO (Fig 2B).



	146	147	148	149	151	154	155	156	157	158	159
2016	N	K	G	V	A	P	H	A	G	A	K
2015	N	K	G	V	A	P	H	A	G	A	K
2014	N	K	G	V	A	P	H	A	G	A	K
2013	N	K	G	V	A	P	H	A	G	A	K
2012	N	K	G	V	A	P	H	A	G	A	K
2011	N	K	G	V	A	P	H	A	G	A	K
2010	N	K	G	V	A	P	H	A	G	A	K
2009	N	K	G	V	A	P	H	A	G	A	K
2008	T	-	G	V	A	S	H	N	G	E	S
2007	T	-	G	V	A	S	H	N	G	E	S
2006	T	-	G	V	A	S	H	N	G	K	S
2005	T	-	G	V	A	S	H	N	G	K	S
2004	T	-	G	V	A	S	H	N	G	K	S
2003	T	-	G	V	A	S	H	N	G	K	S
2002	T	-	G	V	A	S	H	N	G	K	S
2001	T	-	G	V	A	S	H	N	G	K	S
2000	T	-	G	V	A	S	H	N	G	K	S
1999	T	-	G	V	A	S	H	N	G	K	S
1998	T	-	G	V	A	S	H	N	G	K	S
1996	T	K	G	V	A	S	H	N	G	K	S
1995	T	K	G	V	A	S	H	N	G	K	S
1993	T	K	G	V	A	S	H	N	G	K	S
1992	T	K	G	V	A	S	H	N	G	K	S
1991	T	K	G	V	A	S	H	N	G	K	S
1989	T	K	G	V	A	S	H	N	G	K	S
1987	T	K	G	V	A	S	H	K	G	K	S
1986	T	K	G	V	A	S	H	K	G	K	S
1985	T	R	G	V	A	S	H	K	G	K	S
1984	T	K	G	V	A	S	H	K	G	K	S
1983	T	K	G	V	A	S	H	K	G	K	S
1982	T	R	G	V	A	S	H	K	G	K	S
1981	T	R	G	V	A	S	H	K	G	K	C
1980	T	R	G	V	A	S	H	K	G	K	S
1979	T	R	G	V	A	S	H	K	G	K	S
1978	T	R	G	V	A	S	H	K	G	K	S
1977	T	R	G	V	A	S	H	K	G	K	S
1957	R	-	G	V	A	P	H	A	R	K	S
1954	T	R	G	V	A	S	H	A	K	K	S
1953	T	R	G	V	A	S	H	A	R	K	S
1951	T	R	G	V	A	S	H	A	K	K	S
1950	I/T	R	G	V	A	S	H	A	G	K	S
1949	T	R	G	V	A	S	H	K	G	K	S
1948	T	R	G	V	A	S	H	K	G	K	S
1947	T	R	G	V	A	S	H	A	G	K	S
1946	D/T/N	I	G	V	A	S	H	A	G	K	S
1945	T	R	G	V	A	S	H	A	G	K	S
1945	T	R	G	V	A	S	H	A	G	K	S
1943	A	R	G	V	A	S	H	A	G	K	S
1942	T	K	G	V	A	S	H	A	G	K	C
1940	N	I	G	V	A	S	H	A	G	K	S
1936	N	I	G	V	A	S	H	A	G	K	S
1935	T	K	G	V	A	S	H	A	G	K	S
1934	N	-	G	V	A	S	H	E	G	K	S
1933	L	K	G	V	A	S	H	G/R	G	K	S
1918	T	K	G	V	A	S	H	A	G	A	S

Figure 2: Identification of a site of limited variability in the head domain of the H1 HA through structural bioinformatic analysis. A. Variability of antibody bindings sites (ABS) on the crystal structure of A/California/04/2009 HA; those containing position 147 are shown in yellow. "Position" denotes the amino acid position at the centre of the 1000 $Å^2$ ABS. B. Analysis of the disrupted peptide sequence of the ABS of lowest variability containing position 147 (OREO) C. Location of OREO in the context of the receptor binding site with position 147 shown in yellow and the rest of the site coloured in red. (D) The trimeric A/Brevig Mission/1/1918 H1 HA crystal structure with ABS shown in red and position 147 shown in yellow. Each monomer of the trimer is coloured in blue, hazel or green.

We next substituted the five proposed conformations of OREO (Fig 2B) into H6, H5 and H11 HAs (which have not circulated in the human population) and used these chimeric HAs to vaccinate mice using a DNA-DNA-pseudotyped lentivirus prime-boost-boost regimen (Figure S4). Analysis of sera obtained from the final bleed at 21 weeks prior to influenza challenge demonstrated that vaccinating with these epitopes produces antibodies that are cross-reactive to a number of historical strains. Notably, the 2006-like OREO epitope (red) produce cross-reactive antibodies that mirror the neutralisation profile of sera taken in 2006/2007 from young children aged 6 to 11 (Fig 1A&B). In this instance, both datasets neutralised pseudotype lentiviruses displaying HAs from A/Solomon Islands/3/2006, A/USSR/90/1977, A/PR/8/1934 and A/WSN/1933 via the OREO epitope but not A/California/4/2009 or A/South Carolina/1/1918 (Fig 3A-F). Further cross-reactivity was also induced by other OREO conformations which differ from that produced by the 2006-like (red) OREO epitope.

To test the protection of these epitopes against historical strains and produce proof-of-concept data in support of our vaccine strategy, the red (2006-like) and green (1977-like) groups were challenged with the A/PR/8/1934 (Fig 3G) and the pink (1940-like), blue (2009-like) and orange (1995-like) groups were challenged with the A/California/04/2009 (Fig 3H). In each challenge experiment, an unvaccinated group (n=6) was included as well as a group vaccinated via the DNA-DNA-pseudotyped lentivirus regimen with HAs (H6, H5 and H11) without the substitution of an OREO epitope conformation (n=6). The 2006-like (red) and 1977-like (green) OREO epitopes conferred immunity to challenge with the A/PR/8/1934 virus, which last circulated in 1934. As a control, the 2009-like (blue) OREO epitope conformation conferred immunity to challenge with the A/California/4/2009 virus, which last circulated in 2009.

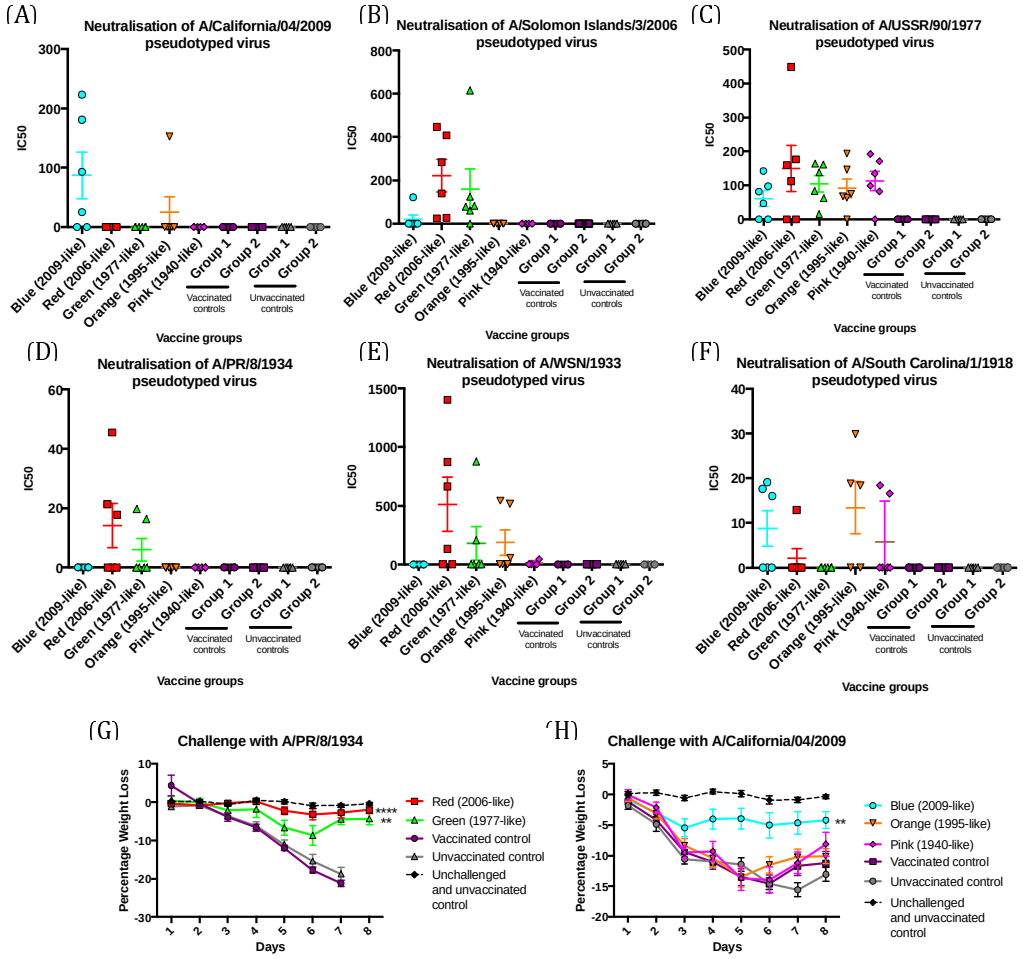


Figure 3. Sequential vaccination using chimeric HA constructs. Five groups of mice were sequentially vaccinated with the sequences outlined in Table S1. substituted into H6, H5 and H11 HAs. Two further groups were sequentially vaccinated with H6, H5 and H11 constructs without any sequence substituted into the HAs. A further two groups were mock vaccinated. The first two vaccinations were administered as a 100 µg *intra muscular* injection of DNA, whilst the final vaccination was administered as an *intra muscular* injection of 8 HI units of lentivirus displaying a chimeric HA with an Alum adjuvant. 2009-like (blue), 2006-like (red), 1995-like (orange), 1977-like (green) and 1940-like (pink) OREO epitope sequences substituted into H6, H5 and H11 (Table S2). (A-E) Pseudotype microneutralisation assays using 0.5 µl of sera from the bleed at 21 weeks. (D-G) Influenza challenge of vaccinated mice with either A/PR/8/1934 or A/California/4/2009. The graphs denote daily weight loss of the mice during the challenge. Mice of the same age, which were not vaccinated or challenged, are shown for reference and denoted 'unchallenged and unvaccinated'. Associated survival curves are shown in Fig S5.

In addition to epitopes of limited variability such as OREO, our analysis identifies a range of highly variable epitopes (Fig 2A); the antigenic trajectory of the latter has been tracked in detail by several previous studies (*24*, *25*), and form the basis of theoretical models of classic "antigenic drift" (*26–30*). Our model (4) requires immune selection to act upon epitopes of both high and limited variability, and here we demonstrate how the immune response can be focussed on an epitope of limited

variability, thereby providing a template for the development of an influenza vaccine conferring broad protection to the H1N1 influenza A subtype. Using the same strategy, we can also produce vaccines against other subtypes of human influenza, as well as swine and poultry influenza viruses, which could have significant economic benefits and also potentially protect us against future influenza pandemics.

Acknowledgements

We would like to thank Dr John S Tregoning, Imperial College, for kindly providing us with the viruses for the influenza challenge. Funding for the study was provide by an ERC grant (DIVERSITY, ID: 268904), an MRC Confidence in Concept grant distributed by the Medical Science Division at the University of Oxford and a Lord Leonard and Lady Estelle Wolfson Foundation Translation Award from The Royal Society. We thank the parents/guardians who gave written informed consent for use of these blood samples for research by the Oxford Vaccine Centre Biobank, with ethical approval by a local research ethics committee (16/SC/0141).

References:

1. A. J. Caton, G. G. Brownlee, J. W. Yewdell, W. Gerhard, The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell*. **31**, 417–427 (1982).
2. D. C. Wiley, I. A. Wilson, J. J. Skehel, Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature*. **289**, 373–378 (1981).
3. J. J. Skehel *et al.*, A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody. *Cell*. **81**, 1779–1783 (1984).
4. M. Recker, O. G. Pybus, S. Nee, S. Gupta, The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proc. Natl. Acad. Sci.* **104**, 7711–7716 (2007).
5. D. Zinder, T. Bedford, S. Gupta, M. Pascual, The Roles of Competition and Mutation in Shaping Antigenic and Genetic Diversity in Influenza. *PLoS Pathog.* **9** (2013), doi:10.1371/journal.ppat.1003104.
6. N. J. Temperton *et al.*, A sensitive retroviral pseudotype assay for influenza H5N1-neutralizing antibodies. *Influenza Other Respi. Viruses*. **1**, 105–112 (2007).
7. G. W. Carnell, F. Ferrara, K. Grehan, C. P. Thompson, N. J. Temperton, Pseudotype-based neutralization assays for influenza: A systematic analysis. *Front. Immunol.* **6** (2015), doi:10.3389/fimmu.2015.00161.
8. D. M. Carter *et al.*, Sequential seasonal H1N1 influenza virus infections protect ferrets against novel 2009 H1N1 influenza virus. *J. Virol.* **87**, 1400–10 (2013).
9. M. S. Miller *et al.*, *Sci. Transl. Med.*, in press, doi:10.1126/scitranslmed.3006637.
10. K. Hancock *et al.*, Cross-reactive antibody responses to the 2009 pandemic

H1N1 influenza virus. *N. Engl. J. Med.* **361**, 1945–52 (2009).
11. S. M. Zimmer *et al.*, Seroprevalence following the second wave of pandemic 2009 H1N1 influenza in Pittsburgh, PA, USA. *PLoS One*. **5** (2010), doi:10.1371/journal.pone.0011601.
12. Y. Li *et al.*, Immune history shapes specificity of pandemic H1N1 influenza antibody responses. *J. Exp. Med.* **210**, 1493–1500 (2013).
13. S. L. Linderman *et al.*, Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013-2014 influenza season. *Proc. Natl. Acad. Sci.* **111**, 15798–15803 (2014).
14. K. Y. A. Huang *et al.*, Focused antibody response to influenza linked to antigenic drift. *J. Clin. Invest.* **125**, 2631–2645 (2015).
15. M. A. Miller, C. Viboud, M. Balinska, L. Simonsen, The Signature Features of Influenza Pandemics — Implications for Policy. *N. Engl. J. Med.* **360**, 2595–2598 (2009).
16. Y. Matsuzaki *et al.*, Epitope Mapping of the Hemagglutinin Molecule of A/(H1N1)pdm09 Influenza Virus by Using Monoclonal Antibody Escape Mutants. *J. Virol.* **88**, 12364–12373 (2014).
17. J. C. Krause *et al.*, A Broadly Neutralizing Human Monoclonal Antibody That Recognizes a Conserved, Novel Epitope on the Globular Head of the Influenza H1N1 Virus Hemagglutinin. *J. Virol.* **85**, 10905–10908 (2011).
18. J. R. R. Whittle *et al.*, Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. *Proc. Natl. Acad. Sci.* **108**, 14216–14221 (2011).
19. S. J. Gamblin, The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science (80-.).* **303**, 1838–1842 (2004).
20. J. Stevens, Structure of the Uncleaved Human H1 Hemagglutinin from the Extinct 1918 Influenza Virus. *Science (80-.).* **303**, 1866–1870 (2004).
21. R. Xu *et al.*, Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science*. **328**, 357–60 (2010).
22. H. Yang *et al.*, Structural Stability of Influenza A(H1N1)pdm09 Virus Hemagglutinins. *J. Virol.* **88**, 4828–4838 (2014).
23. C. Kratsch, T. R. Klingen, L. Mümken, L. Steinbrück, A. C. McHardy, *Virus Evol.*, in press, doi:10.1093/ve/vev025.
24. D. J. Smith, Mapping the Antigenic and Genetic Evolution of Influenza Virus. *Science (80-.).* **305**, 371–376 (2004).
25. B. F. Koel *et al.*, Substitutions Near the Receptor Binding Site Determine Major Antigenic Change During Influenza Virus Evolution. *Science (80-.).* **342**, 976–979 (2013).
26. N. M. Ferguson, A. P. Galvani, R. M. Bush, Ecological and immunological determinants of influenza evolution. *Nature*. **422**, 428–433 (2003).
27. K. Koelle, S. Cobey, B. Grenfell, M. Pascual, Epochal Evolution Shapes the Phylodynamics of Interpandemic Influenza A (H3N2) in Humans. *Science (80-.).* **314**, 1898–1903 (2006).
28. J. R. Gog, B. T. Grenfell, Dynamics and selection of many-strain pathogens. *Proc. Natl. Acad. Sci.* **99**, 17209–17214 (2002).
29. T. Bedford, A. Rambaut, M. Pascual, Canalization of the evolutionary trajectory of the human influenza virus. *BMC Biol.* **10**, 38 (2012).
30. X. Du, A. A. King, R. J. Woods, M. Pascual, Evolution-informed forecasting of seasonal influenza A (H3N2). *Sci. Transl. Med.* **5325**, 198168 (2017).