

rstb.royalsocietypublishing.org

Research



Cite this article: Wikramaratna PS,

Sandeman M, Recker M, Gupta S. 2013

The antigenic evolution of influenza: drift or

thrift? *Phil Trans R Soc B* 368: 20120200.

<http://dx.doi.org/10.1098/rstb.2012.0200>

One contribution of 18 to a Discussion Meeting Issue 'Next-generation molecular and evolutionary epidemiology of infectious disease'.

Subject Areas:

health and disease and epidemiology,
theoretical biology

Keywords:

influenza, epitope, epidemiology

Author for correspondence:

Sunetra Gupta

e-mail: sunetra.gupta@zoo.ox.ac.uk

The antigenic evolution of influenza: drift or thrift?

Paul S. Wikramaratna, Michi Sandeman, Mario Recker and Sunetra Gupta

Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

It is commonly assumed that antibody responses against the influenza virus are polarized in the following manner: strong antibody responses are directed at highly variable antigenic epitopes, which consequently undergo 'antigenic drift', while weak antibody responses develop against conserved epitopes. As the highly variable epitopes are in a constant state of flux, current antibody-based vaccine strategies are focused on the conserved epitopes in the expectation that they will provide some level of clinical protection after appropriate boosting. Here, we use a theoretical model to suggest the existence of epitopes of low variability, which elicit a high degree of both clinical and transmission-blocking immunity. We show that several epidemiological features of influenza and its serological and molecular profiles are consistent with this model of 'antigenic thrift', and that identifying the protective epitopes of low variability predicted by this model could offer a more viable alternative to regularly update the influenza vaccine than exploiting responses to weakly immunogenic conserved regions.

1. Introduction

Influenza A viruses are responsible for between three and five million cases of severe disease annually, and up to half a million deaths worldwide. These viruses are classified into subtypes on the basis of variation in their envelope glycoproteins, haemagglutinin (HA) and neuraminidase (NA), and the event of their replacement on a global scale is commonly referred to as an antigenic shift in the virus population. In the last 100 years, we have experienced three such shifts: in 1957, the H1N1 subtype that had been circulating since 1918 was replaced by H2N2; in 1968, H2N2 was replaced by H3N2; H1N1 was reintroduced in 1977 and has been cocirculating since with H3N2, although, in 2009, the current lineage was replaced by one derived from pre-existing swine, avian and human viruses. Each subtype, while in circulation, also undergoes a form of antigenic change culminating in the sequential dominance of antigenically distinguishable strains with very limited cross-sectional genetic diversity. The underlying process is generally visualized as a continuous and incremental transformation principally of the HA glycoprotein, and goes by the name of antigenic drift.

While it has considerable appeal as a verbal explanation for the epidemic behaviour of influenza, a formal link between the process of antigenic drift and patterns of influenza strain replacement is very difficult to make. This is because random mutation is much more likely to lead to a diffuse cloud of antigenic types on a variety of genetic backgrounds than the sequential emergence of discrete strains. A simple, but biologically unsatisfactory solution is to restrict the mode of mutation such that the virus population effectively travels in a pre-ordained straight line or circle [1,2]. The alternative explanation is that most mutants do not succeed, either because they are diffusing through genotypic space along phenotypically neutral networks [3] or because they are out-competed by strains that have achieved greater antigenic distance from the preceding epidemic strains [4] or as a consequence of short-term strain-transcending immunological interference [5]. These additional assumptions can allow the virus population to progress in a linear manner through its available 'antigenic space' by counteracting the diffusive tendencies of antigenic drift. A common assumption here is that the potential for variation of the HA

protein is extremely high and that long-term immune responses to HA are strain-specific, such that immunity to one antigenic type has no effect on any forms other than those that have very recently diverged from it. We have challenged this notion [6] by showing that epidemic behaviour of influenza can be readily explained by assuming that each strain elicits long-term partially cross-protective immune responses in addition to strain-specific immunity. This model—which we will henceforward refer to as the ‘antigenic thrift’ model (as suggested by Eddie Holmes)—departs from the conventional ‘antigenic drift’ hypothesis in a number of important ways: (i) there are a restricted number of unique but inter-connected antigenic states, (ii) the virus population has continuous access to these states, but (iii) most of these are unsuccessful owing to pre-existing partially cross-reactive immune responses in the host population.

We have no shortage of sequence data on influenza, but limited means as yet to use it to discriminate between competing hypotheses concerning the antigenic evolution of the virus population. Phylogenetic trees of influenza A exhibit a spindly structure that has commonly been misinterpreted as evidence of antigenic drift, but, in fact, simply indicates that the populations repeatedly pass through tight bottlenecks either as a result of selection or—as has recently been shown—as a straight-forward consequence of sampling sequences serially through time under neutral evolution [7]. It is clear that several of the models based on a process of antigenic drift [3–5] are capable of generating the observed trees; indeed, a principal focus of these efforts was to reproduce the ladder-like phylogeny of influenza A. It is as yet unknown, however, whether the antigenic thrift model is consistent with the phylogenetics of influenza. The antigenic relationships between different strains of influenza can also be determined using serological methods. An important focus of this paper is how the antigenic thrift model stands up to the scrutiny of sero-epidemiological analyses of the antigenic evolution of influenza.

We first provide a review of the antigenic thrift model; we then show how discriminating between epitopes of high and low variability provides a novel means of reconciling the dynamics of this model with empirical data on the antigenic evolution of influenza. Finally, we discuss how we may use improved serological techniques in conjunction with molecular methods to identify protective epitopes of low variability that may enable us to address the problem of influenza vaccination in a novel and practicable manner.

2. The antigenic thrift model

Figure 1 provides a caricature of the fundamental assumptions of the antigenic thrift model in terms of how the various epitopes might map onto the structure of an HA monomer. In essence, the model combines highly variable, strain-specific epitopes (here visualized as surrounding the binding pocket) as well as epitopes of low to intermediate variability that are shared between strains. The model argues that immune responses against the latter are critical determinants of the protection against disease and onward transmission and drive the population dynamics of influenza in concert with antibodies directed at the uniquely strain-specific epitopes. Those individuals who have been exposed to a particular strain—as defined by a combination of these

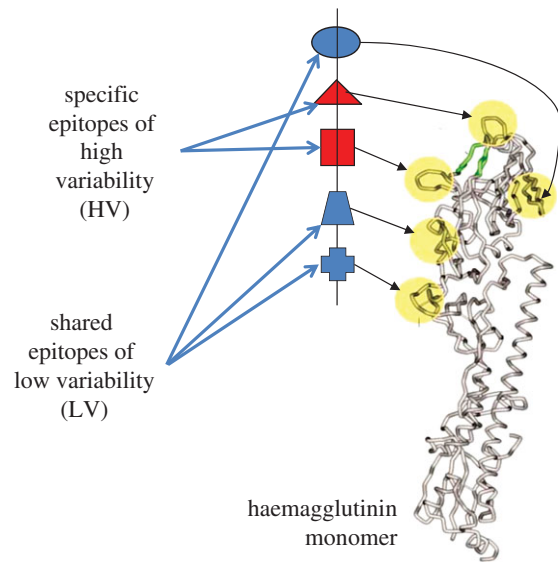
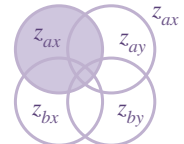


Figure 1. The antigenic thrift model is based on a multi-locus representation of the virus with each locus corresponding to an epitope region. This figure shows how these may locate to the known antigenic sites on a monomer of haemagglutinin (adapted from [71] & [72] with permission from OUP and NEJM, respectively).

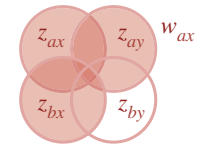
proportion immune to strain ax :

$$\frac{dz_{ax}}{dt} = \beta y_{ax} (1 - z_{ax}) - \mu z_{ax}$$



proportion immune to strains sharing alleles with ax :

$$\frac{dw_{ax}}{dt} = \sum_{ij \sim ax} \beta y_{ij} (1 - w_{ax}) - \mu w_{ax}$$



proportion infectious strain ax :

$$\frac{dy_{ax}}{dt} = \beta y_{ax} ((1 - w_{ax}) + (1 - r)(w_{ax} - z_{ax})) - \sigma y_{ax}$$

Figure 2. The structure of the antigenic thrift model is shown here with reference to a two locus and two allele system. A system of overlapping compartments can be used to indicate the proportions immune to each strain (z) and the proportion immune to antigenically related variants (w), from which the proportion infectious with this strain (y) can be deduced. In the diagram, z_{ax} and w_{ax} are indicated by purple and red shading, respectively. The notation $ij \sim ax$ indicates all strains sharing alleles with ax . The parameters β and $1/\sigma$, respectively, define the transmission coefficient and infectious period of the virus, $1/\mu$ corresponds to the life expectancy of the host population and γ measures the cross-immunity of a host gains from having seen a related but not identical variant.

epitopes—have lifelong immunity to that same strain, but also have partial immunity (also lifelong) to strains related to it by virtue of possessing common epitopes. This network of cross-protection acts to limit the emergence of new antigenic types and can thus reconcile the high mutation rates of influenza with the dominance of a single antigenic type in each season.

Figure 2 summarizes the model structure using a simple schematic based on a system with only two relevant loci and alleles a and b at the first locus and x and y at the second. It is assumed that individuals who have been exposed to a particular strain (say ax , as shown in the diagram) are immune to further infection by the same strain,

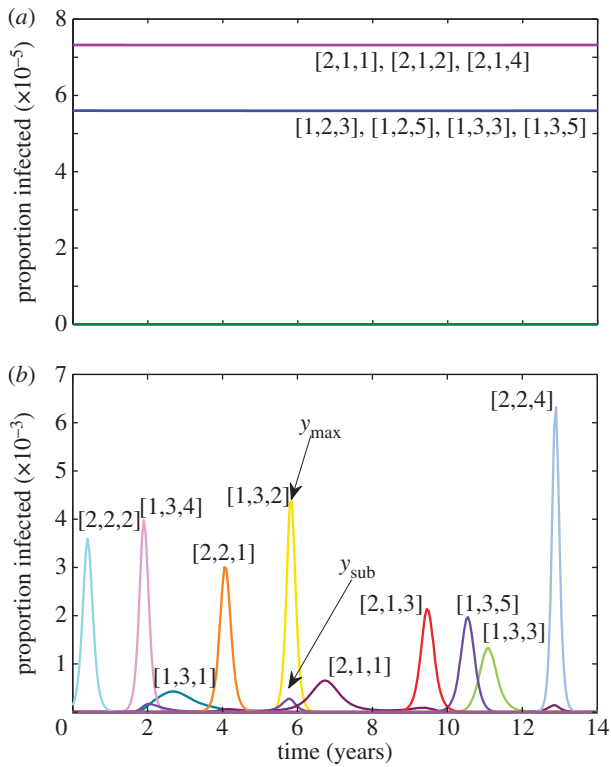


Figure 3. Strain dynamics within a $\{2,3,5\}$ antigenic structure with (a) $\gamma = 0.95$ and (b) $\gamma = 0.8$ ($\beta = 292$; $\sigma = 73$; $\mu = 0.02$).

while those who have been exposed to strains that share alleles with it (in this example, ay and bx) have a reduced probability $(1 - \gamma)$ of transmitting the strain when infected; the parameter γ reflects the strength of allele-specific responses in preventing transmission. This structure can be easily generalized to multiple loci or epitopes with different levels of variability (i.e. number of possible alleles); we will henceforward use the notation $\{2,3,25\}$, for example, to indicate that there are three loci with 2, 3 and 25 alleles, respectively, and $[i,j,k]$ to designate a particular strain or combination of alleles.

Mutation is not explicitly considered in this model. Instead, and since the model is deterministic, each possible antigenic variant is continuously present within the viral population. It is difficult to ascertain how this translates into an explicit mutation rate, but it does mean that at the precise moment a gap emerges in the network of host immunity, this gap can be exploited by any and all appropriate antigenic variants. Thus, it is this network of host immune responses, and not the mutational capability of the virus, that constrains observed antigenic diversity within the premise of antigenic thrift.

Multi-locus systems are capable of exhibiting two kinds of structuring, as shown in figure 3. At high levels of immune selection and provided that there is an equal number of possible alleles at each locus, discrete strain structure emerges with the stable maintenance of a set of strains that do not share alleles [8,9]. This discrete antigenic structure tends to break down in deterministic multi-locus models, when unequal numbers of allelic variants are instead possible at each locus but are recovered by the inclusion of stochastic processes [9]. At intermediate levels of immune selection, cyclical or chaotic strain dynamics (CSS) occurs [10]. We posit that the epidemic behaviour of influenza maps onto an area of CSS that exhibits high single strain dominance [6]. Figure 3b provides an example of this kind of dynamic

for a $\{2,3,5\}$ system; figure 4 traces a section of the antigenic trajectory of the virus population within a three-dimensional space that can be used to represent the relationships between all possible strains.

Single strain dominance can be quantified by the measure of ε by comparing the relative prevalence of the two most common antigenic variants within single epidemics (figure 3), and then averaging across extended periods of time [6]. More formally, averaging across each of P epidemics:

$$\varepsilon = \frac{1}{P} \sum_{i=1}^P \frac{y_{\max}^i - y_{\text{sub}}^i}{y_{\max}^i}$$

High ε , therefore, indicates strong single strain dominance as apparent for the antigenic evolution of influenza. Single strain dominance peaks at certain intermediate levels of γ [6], but also shows a dependence on epitope architecture. Figure 5a shows how different multi-locus systems, all with 32 total variants, differ in the region of γ where they exhibit strong single strain dominance. Interestingly, the combination that most favours strong single strain dominance here is $\{8,2,2\}$, while those that do least well are $\{16,2\}$ and $\{2,2,2,2,2\}$, which respectively minimize and maximize connectivity. The complexity of the relationship between high ε and antigenic architecture is further demonstrated in figure 5b by comparing structures with 400 variants each. It is clear, nonetheless, that structures with high variability (HV) at every locus rarely tend to exhibit $\varepsilon > 0.5$ (even though they are in CSS) and a single-locus system with 400 alleles will not exhibit any oscillations whatsoever. Those combinations that perform well on this measure tend to contain some epitopes of low variability, but can also contain at least one highly variable locus.

3. Serological signatures of antigenic evolution

A common method for recording antibody levels in sera of either naturally or experimentally infected animals is the haemagglutination inhibition (HI) assay. HI assays exploit the ability of the influenza virus to bind to sialic acid receptors, and thus agglutinate avian and mammalian red blood cells [11]; the level of dilution at which a serum sample stops being able to prevent agglutination determines its HI titre. Published HI tables of human influenza A based on antisera raised in ferrets tend to show very high titres to the homologous isolate, but highly variable titres to temporally close isolates and, usually, a complete loss of reactivity against strains isolated more than a few years distant from the considered strain. This has commonly been interpreted as evidence of gradual antigenic drift, with each dilution step corresponding to an increase in antigenic distance. We have previously shown that such empirical data are consistent with the antigenic thrift model under the sampling scheme and multivariate analysis commonly used in their representation in two-dimensional antigenic space [6]. Indeed, the zig-zagging movement through antigenic space revealed by sophisticated cartographic methods applied to the evolution of H3N2 since 1968 [12] may be better explained by the antigenic thrift model. We have suggested that the apparent continuous increase in antigenic distance may be due to the censoring of entries between non-adjacent time-points, and that this signature of drift would

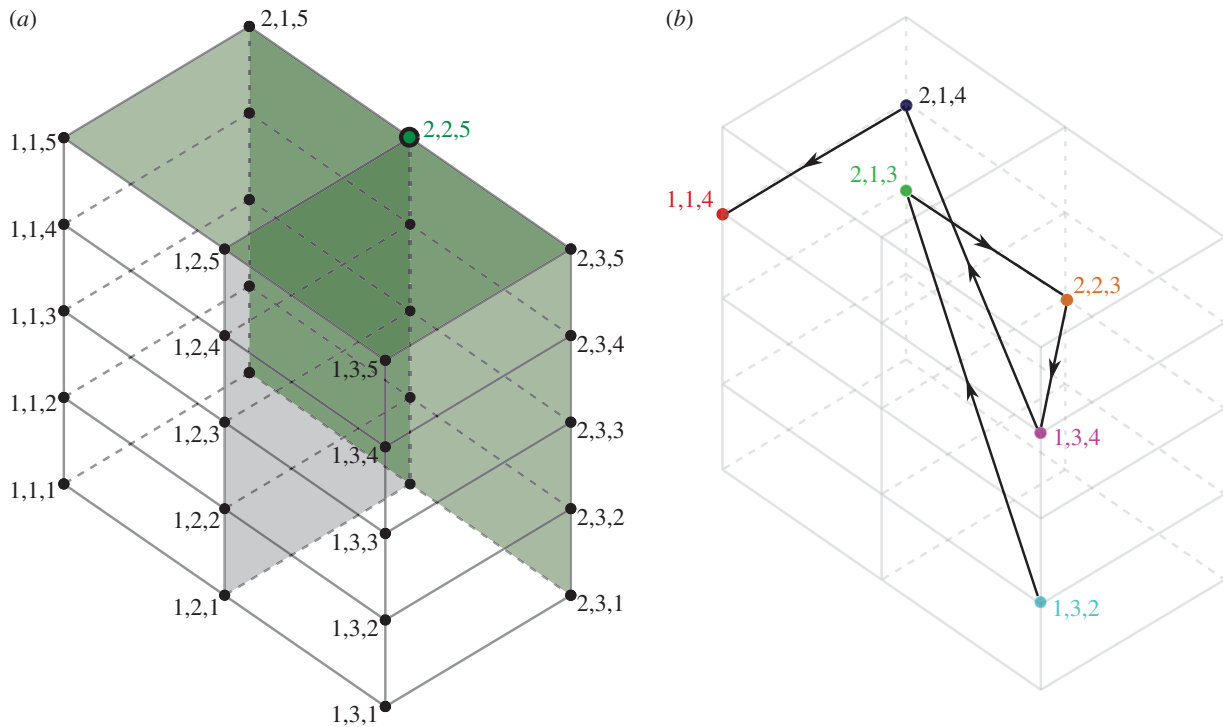


Figure 4. The $\{2,3,5\}$ antigenic structure can be visualized in three-dimensional space with each axis representing a set of alleles at a particular locus, such that each point corresponds to a different antigenic variant. (a) Cross-reactivity against variant $[2,2,5]$ extends within this space along the shaded green planes. (b) An example trajectory of the virus population through this space.

disappear if accurate distance measures were available for all of the elements of the HI data matrix. This may be understood with reference to figure 4b: erasing the links between non-consecutive strains would have the effect of stretching out the trajectory, but nonetheless preserve some of the inherent transverse movement.

There are at least two alternative interpretations that reconcile the absence of reactivity by HI between non-adjacent time-points with the antigenic thrift model. The first is that HI assays selectively provide information on the unique strain-specific epitopes rather than the shared epitopes of limited diversity: in other words that the antibodies that prevent haemagglutination are directed only at the highly variable epitopes putatively clustering around the receptor-binding pocket as shown in figure 1. Figure 6a shows the antigenic relationships between strains within a $\{2,3,25\}$ system as revealed by an assay that focuses solely on the third, highly variable, epitope region. A strong diagonal signature is observed, offering the impression of linear movement through antigenic space when in fact the population is actually zig-zagging within it in a manner analogous to that shown in figure 4b. Assuming instead that the magnitude of the HI titre depends on the precise number of shared epitopes (figure 6b), we can again find a strong diagonal signature, but with more evidence of clustering of similar antigenic variants in time, as seen in human HI data and by genetic analysis of influenza virus sequence evolution [13]. This emerges as a natural property of this model, owing to the time-scales at which epitope-specific immunity declines. In effect, as older hosts die, population-level immunity against the strains that they have specifically experienced wanes, thus creating gaps in the network of herd immunity that may be occupied by similar strains. This can lead to a sequence of antigenically related epidemics, until eventually a completely discordant allele combination is favoured and

a cluster jump occurs. This is evident even in systems of low dimensionality such as the $\{2,3,5\}$ example in figure 3b: the sequence $[2,1,1] \rightarrow [2,1,3] \rightarrow [1,3,5] \rightarrow [1,3,3]$ can be interpreted as a cluster transition between $[2,1,*] \rightarrow [1,3,*]$.

Another consideration in the interpretation of HI is that the laboratory animals used in generating the data may produce a response that is more focused towards the variable regions of HA than humans upon natural infection. HI assays are typically performed on post-infection antisera raised in ferrets and there is evidence suggesting that the cross-reactivities can be different to that of sera to the same isolate raised in mice and rabbits [14,15]. To our knowledge, no comparative study has been performed on post-infection sera taken from ferrets and humans or ducks and chickens, or indeed for many of the pairs of animals, where post-infection sera from the former are used to infer antigenic relationships in the latter.

Antigenic relationships between influenza strains can also be interrogated using microneutralization (MN) assays, where the virus is mixed with varying dilutions of serum and then inoculated into culture, and the presence of virus-specific antibodies in the serum is indicated by impaired or absent viral replication. HI titres correlate well with the results of MN assays [16], suggesting that both are detecting antibodies against highly variable epitopes putatively clustering around the receptor-binding pocket. However, discrepancies between the two assays—such as antisera with high neutralization titre but low or even absent HI titre or vice versa—have been reported for avian [17,18] and swine [19] influenza, and there are monoclonal antibodies that neutralize but do not inhibit haemagglutination and vice versa [20,21], suggesting that they do have different (although possibly overlapping) specificities. In theory, MN assays and other sophisticated serological techniques currently are in development [22], should be able to detect *additional* antibody responses against

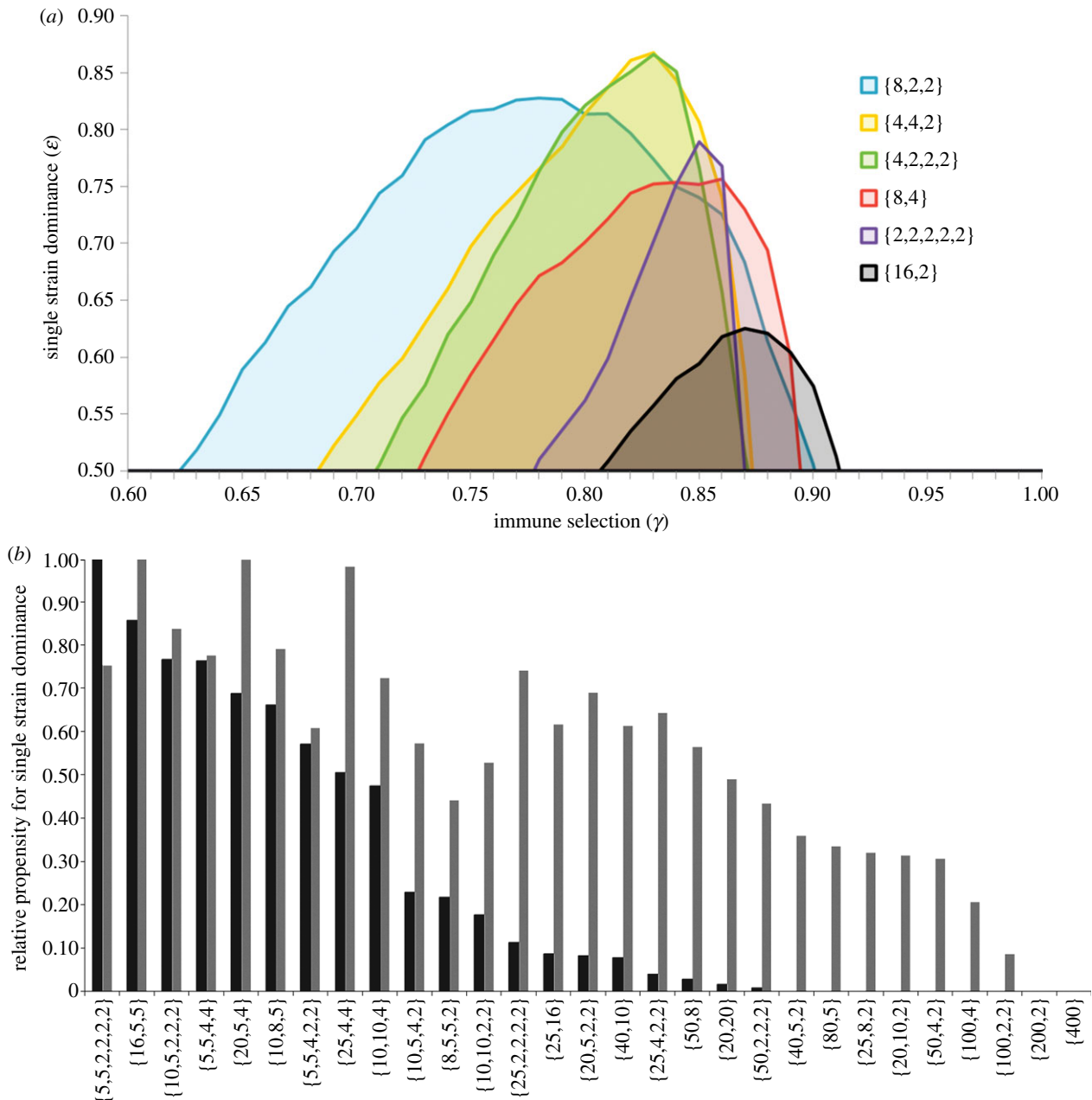


Figure 5. Impact of architecture of antigenic space on strong single strain dominance among different combinations of alleles and loci that produce a total of (a) 32 strains and (b) 400 strains. In (b), this is represented by the area under the curve, where $\epsilon > 0.5$ (black bars) and $\epsilon > 0.1$ (grey bars). Epidemics were identified by local maxima in total prevalence ($\beta = 292$; $\sigma = 73$; $\mu = 0.02$).

the epitopes of limited variability (LV) provided they are performed on the appropriate sera.

There is also growing evidence that antibodies targeting the stem region of HA (HA2) are capable of mediating protection from influenza infection. Such antibodies appear capable of recognizing a variety of different subtypes and strains, resulting also in protection of mice from lethal infection upon challenge (as reviewed by Yewdell [23]). In particular, monoclonal antibodies against HA2 have been shown capable of neutralizing drift variants of H3N2, and also protecting mice from challenge by the same variants [24]. Such responses do not appear to block viral attachment, and therefore do not exhibit cross-reactivity in the HI test, but may instead inhibit viral fusion [24,25]. The epitopes targeted by this class of antibody, may correspond to those of LV posited by the model of antigenic thrift, perhaps not least because escape from them can be precipitated by single point mutations in HA2 [26].

Epitopes of LV may also be located on NA; antibodies against these would act to prevent release of virus from the infected cell. The recent application of antigenic cartography methods to NA inhibition data reveals a high degree of asymmetry that is consistent with the antigenic thrift model. Sandbulte *et al.* [27] further suggest that antigenic evolution of NA may account for unexpected vaccine failures, where there is a good match based on HI data, highlighting the potential importance of more than just HI data in understanding the ways in which the virus can change to evade immune recognition.

4. Re-emergence of antigenic types

Perhaps, the most compelling evidence in support of the antigenic thrift hypothesis comes from the antigenic analysis performed in the wake of the 2009 H1N1 pandemic. Multiple serological studies have found evidence in the elderly of

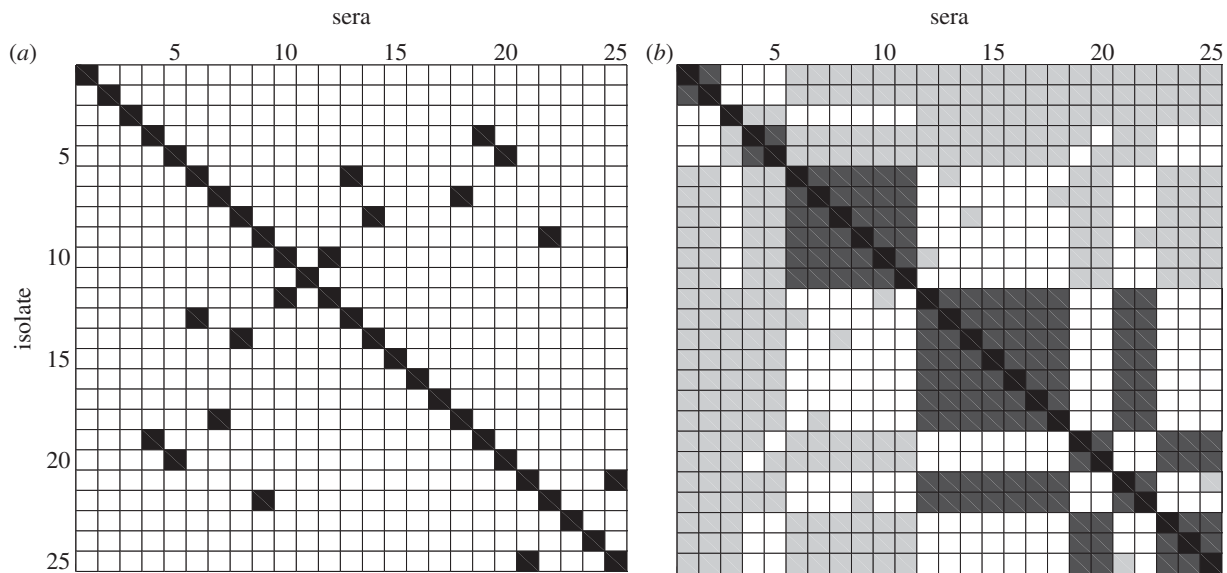


Figure 6. Pseudo HI tables constructed from the strength of cross-reactivity between the dominant variants of the *i*th and *j*th epidemics within a $\{2,3,25\}$ antigenic space under different assumptions about the HI test: (a) HI only picks up the most variable epitope and strains only cross-react in the test (shown in black) if they are identical at this epitope and (b) HI titre depends on the number of shared epitopes with black/dark grey/light grey/white indicating epidemic strains that share 3/2/1/0 alleles, respectively. Epidemics were defined as local maxima in total prevalence more than 1.645 standard deviations from the mean, in line with [73]. ($\gamma = 0.8$, $\beta = 292$, $\sigma = 73$, $\mu = 0.02$).

pre-existing antibodies that recognize the 2009 pandemic strain [28–35], and this appears to explain the low rates of disease in this age group during the pandemic [30,35–37]. Infection of both ferrets and mice with A/New Jersey/76 has been shown to confer strong cross-protection against challenge with 2009 pandemic H1N1, together with high cross-reactivity in the HI test [38–41]. In humans, receipt of the A/New Jersey/76 vaccine has been implicated in superior antibody responses against the 2009 strain [42]. H1N1 isolated prior to 1950 has also been shown to elicit significant, though less complete, protection upon challenge by pandemic H1N1 in both animal models, despite only weak cross-reactivity within HI and MN tests [38–41,43]. An inability of later seasonal isolates to protect against pandemic H1N1 challenge has been linked to glycosylation patterns on the globular head of HA, which may also explain why not all pre-1950 isolates appear to protect equally well [41].

Similar serological and clinical observations of cross-reactivity and clinical protection in the elderly have also been made in each of the previous three pandemics [44]. This could be explained by original antigenic sin (a phenomenon by which individuals continue to produce antibodies against the strain they were first infected with, even when challenged by a different strain/subtype), but this would still require a high degree of similarity between the returning strain and those to which they had been previously exposed.

There are hints of re-emergence of certain epitopes in antigenic analyses performed on H2N2 influenza, with a number of monoclonal antibodies raised against a 1957 strain cross-reacting strongly with a strain isolated in 1964, but not with a 1963 strain [45]. Antigenic analyses of influenza in pigs and birds are also suggestive of the re-emergence of antigenic types. Contractions in antigenic distance can be observed among populations of swine H1 viruses in the USA [46], and HI tables of H5N1 isolates from a variety of avian species demonstrate discordances in cross-reactivity [47] that are more easily explained by recycling of variants than incremental accrual of antigenic distance.

5. Discussion

Many features of the epidemiology of influenza can be explained by assuming that neutralizing antibodies act upon shared epitopes of LV as well as upon epitopes of HV that may be unique to a particular epidemic strain. We argue here that the use of HI assays has focused our attention on HV epitopes, but immune responses against these alone cannot produce a sequential emergence of antigenic types. We propose that the structuring of the virus population is achieved principally through immune responses against the LV epitopes which are not adequately represented in HI tables as they have a limited role in the binding of the virus to RBC. The existence of these additional LV epitopes is consistent with the observation that clinical protection and HI titre are logarithmically related, with limited improvement in protection beyond a certain (fairly low) titre [48], and also explains why high vaccine efficacy is sometimes observed even when HI data indicate that the incoming influenza strain has changed [49]. The results of early studies showing induction of superior HI titres with inactivated vaccines, but inferior protection from disease and shedding when compared with live vaccines [50] can be justified within this framework, if the latter induce a broader response that includes the LV epitopes. It is also tempting to speculate that some of the effects of original antigenic sin may be attributed to the skewing of antibody responses towards LV epitopes in later infections. In other words, a naive individual would produce antibodies to both HV and LV epitopes, but selectively towards LV epitopes in subsequent infections since they are shared between strains. They would thus maintain a strong HI response to the original strain, but may not show high HI titres to more recent infections. A recent Japanese study [51] showing that sera from young children were prone to recognize only the antigenic site B1 of the HA1 region of H3, in contrast to older individuals who had broad recognition, supports this idea.

Our model is not unique in invoking multiple components of immunity, but differs critically from other

frameworks in that the targets of immunity divide between HV and LV epitopes, which both elicit strong specific life-long immunity and do not include epitopes that induce short-term non-specific immunity (as in [5]). The validity of the antigenic thrift model rests on the existence of LV epitopes: identifying these would thus allow us to discriminate between competing hypotheses concerning the antigenic evolution of influenza (although they are by no means mutually exclusive), but, more importantly, could form the basis of a new vaccine that would release us from our dependence on monitoring change in the HV epitopes, and could complement the use of both antibody and T-cell-based vaccines towards fully conserved but weakly immunogenic epitopes that are currently under development [52].

How do we go about looking for LV epitopes, if they are not visible by HI? We anticipate that the dissection of antibody responses in human sera using MN assays, in conjunction with other more advanced techniques in the pipeline [22], will assist in identification, but we can also use molecular sequencing methods? It has long been recognized that the virus possesses a number of discrete antigenic sites [53], and several studies indicate that the number of amino acid differences in these sites is a better predictor of vaccine efficacy than HI data from ferrets [54,55]. We have previously highlighted that there is LV at 18 amino acid positions that have been identified by Bush *et al.* [56] as being under positive selection in H3N2; several of these are represented among the key immunodominant positions identified by pairwise comparison of consecutive epidemic strains [12,57] and by analysis of change in net charge [58]. Amino acid changes in the HA epitopes of H2N2 [45], H1N1 [59] and highly pathogenic H5N1 [60] also seem to be subject to strong restrictions. Epistatic interactions between sites, as documented by Kryazhimskiy *et al.* [61], may act to further reduce the potential nodes within antigenic space that may be occupied by the virus or favour a particular combination even under a very slight increase in transmissibility [62]. Sequence similarities between epitopes of 1918 H1N1 influenza and the 2009 pandemic strain have been used to justify pre-existing immunity to the latter [63,64]; similar arguments have also been made for the presence of neutralizing antibodies

among individuals born before 1957 by comparing the HA sequences of 1957 and 2009 H1N1 strains [65]. Many of these studies rely on HI assays to discriminate between strains: the antigenic thrift model would predict that this method would selectively emphasize the role of mutations near the receptor-binding sites and, indeed, this seems to be the case in a recent study [63]. Rudneva *et al.* [63] also recorded some discordance between HI and ELISA studies, suggesting that the use of HI may not be sufficient to pick-up the LV epitopes. Newer methods such as panning of whole-genome-fragment phage display libraries (GFPDL) with convalescent human sera [66], tend by contrast to direct our attention towards weakly immunogenic conserved epitopes, but may be deployed to pick out LV epitopes with finer resolution of HA and NA gene fragments. It is also important that these analyses include considerations of effects of glycosylation, alteration of biophysical properties [67], complex interactions between residues at antigenic sites both within and between HA and NA [61,68,69] and potential effects of antibody interference [70]. Combining these techniques to elucidate the functional repertoire of human antibodies to influenza will be invaluable in resolving to what extent its epidemiology is determined by epitopes of LV, and whether these may be used to confer broader protection.

Note added in proof

Zinder *et al.* [74] have recently shown that the phylodynamics of influenza can be readily generated within a similar framework with a number of epitope regions of limited diversity, under a somewhat different cross-immunity structure, provided it is also mutation-limited. This suggests that it will be difficult to discriminate between competing hypotheses of antigenic evolution on the basis of their ability to generate realistic phylodynamic patterns.

We thank Maria Zambon, Katja Hoschler, Liz Miller, Adrian Smith, Andy Pollard, Sarah Gilbert and John McCauley for helpful discussions. This work was supported by the European Research Council (ERC Advanced Grant—Diversity). M.R. is a Royal Society University Research Fellow, and S.G. is a Royal Society Wolfson Research Fellow.

References

- Gomes MGM, Medley GF, Nokes DJ. 2002 On the determinants of population structure in antigenically diverse pathogens. *Proc. R. Soc. Lond. B* **269**, 227–233. (doi:10.1098/rspb.2001.1869)
- Gog JR, Grenfell BT. 2002 Dynamics and selection of many-strain pathogens. *Proc. Natl Acad. Sci. USA* **99**, 17 209–17 214. (doi:10.1073/pnas.252512799)
- Koelle K, Cobey S, Grenfell BT, Pascual M. 2006 Epochal evolution shapes the phylodynamics of inter-pandemic influenza A (H3N2) in humans. *Science* **314**, 1898–1903. (doi:10.1126/science.1132745)
- Bedford T, Rambaut A, Pascual M. 2012 Canalization of the evolutionary trajectory of the human influenza virus. *BMC Biol.* **10**, 38. (doi:10.1186/1741-7007-10-38)
- Ferguson NM, Galvani AP, Bush RM. 2003 Ecological and immunological determinants of influenza evolution. *Nature* **422**, 428–433. (doi:10.1038/nature01509)
- Recker M, Pybus OG, Nee S, Gupta S. 2007 The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proc. Natl Acad. Sci. USA* **104**, 7711–7716. (doi:10.1073/pnas.0702154104)
- Gray RR, Pybus OG, Salemi M. 2011 Measuring the temporal structure in serially sampled phylogenies. *Methods Ecol. Evol.* **2**, 437–445. (doi:10.1111/j.2041-210X.2011.00102.x)
- Gupta S, Maiden MC, Feavers IM, Nee S, May RM, Anderson RM. 1996 The maintenance of strain structure in populations of recombining infectious agents. *Nat. Med.* **2**, 437–442. (doi:10.1038/nm0496-437)
- Buckee CO, Recker M, Watkins ER, Gupta S. 2011 Role of stochastic processes in maintaining discrete strain structure in antigenically diverse pathogen populations. *Proc. Natl Acad. Sci. USA* **108**, 15 504–15 509. (doi:10.1073/pnas.1102445108)
- Gupta S, Ferguson N, Anderson R. 1998 Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* **280**, 912–915. (doi:10.1126/science.280.5365.912)
- Hirst GK. 1941 The agglutination of red cells by allantoic fluid of chick embryos infected with influenza virus. *Science* **94**, 22–23. (doi:10.1126/science.94.2427.22)
- Smith DJ, Lapedes AS, De Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, Fouchier RA. 2004

- Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 371–376. (doi:10.1126/science.1097211)
13. Plotkin JB, Dushoff J, Levin SA. 2002 Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *Proc. Natl Acad. Sci. USA* **99**, 6263–6268. (doi:10.1073/pnas.082110799)
 14. Johansson BE, Kilbourne ED. 1992 Influenza vaccine strain selection: equivalence of two antigenically distinct haemagglutinin variants of 1989 H3N2 influenza A virus in protection of mice. *Vaccine* **10**, 603–606. (doi:10.1016/0264-410X(92)90440-U)
 15. Gillim-Ross L, Santos C, Chen Z, Aspelund A, Yang CF, Ye D, Jin H, Kemble G, Subbarao K. 2008 Avian influenza H6 viruses productively infect and cause illness in mice and ferrets. *J. Virol.* **82**, 10 854–10 863. (doi:10.1128/JVI.01206-08)
 16. Wood JM, Major D, Heath A, Newman RW, Hoschler K, Stephenson I, Clark T, Katz JM, Zambon MC. 2012 Reproducibility of serology assays for pandemic influenza H1N1: collaborative study to evaluate a candidate WHO International Standard. *Vaccine* **30**, 210–217. (doi:10.1016/j.vaccine.2011.11.019)
 17. Yoden S, Kida H, Yanagawa R. 1982 An avian influenza virus of which infectivity is neutralized by antisera lacking hemagglutination-inhibition activity: brief report. *Arch. Virol.* **74**, 205–210. (doi:10.1007/BF01314713)
 18. Rowe T, Abernathy RA, Hu-Primmer J, Thompson WW, Lu X, Lim W, Fukuda K, Cox NJ, Katz JM. 1999 Detection of antibody to avian influenza A (H5N1) virus in human serum by using a combination of serologic assays. *J. Clin. Microbiol.* **37**, 937.
 19. Leuwerke B, Kitikoon P, Evans R, Thacker E. 2008 Comparison of three serological assays to determine the cross-reactivity of antibodies from eight genetically diverse U.S. swine influenza viruses. *J. Vet. Diagn. Invest.* **20**, 426–432. (doi:10.1177/104063870802000403)
 20. Gitelman AK, Kaverin NV, Kharitonov IG. 1986 Dissociation of the haemagglutination inhibition and the infectivity neutralization in the reactions of influenza A/USSR/90/77 (H1N1) virus variants with monoclonal antibodies. *J. Gen. Virol.* **67**, 2247–2251. (doi:10.1099/0022-1317-67-10-2247)
 21. Kida H, Brown LE, Webster RG. 1982 Biological activity of monoclonal antibodies to operationally defined antigenic regions on the hemagglutinin molecule of A/Seal/Massachusetts/1/80 (H7N7) influenza virus. *Virology* **122**, 38. (doi:10.1016/0042-6822(82)90375-0)
 22. Katz JM, Hancock K, Xu X. 2011 Serologic assays for influenza surveillance, diagnosis and vaccine evaluation. *Expert Rev. Anti Infect. Ther.* **9**, 669–683. (doi:10.1586/eri.11.51)
 23. Yewdell JW. 2011 Viva la revolution: rethinking influenza A virus antigenic drift. *Curr. Opin. Virol.* **1**, 177–183. (doi:10.1016/j.coviro.2011.05.005)
 24. Wang TT, Tan GS, Hai R, Pica N, Petersen E, Moran TM, Palese P. 2010 Broadly protective monoclonal antibodies against H3 influenza viruses following sequential immunization with different hemagglutinins. *PLoS Pathog.* **6**, e1000796. (doi:10.1371/journal.ppat.1000796)
 25. Sui J *et al.* 2009 Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nat. Struct. Mol. Biol.* **16**, 265–273. (doi:10.1038/nsmb.1566)
 26. Okuno Y, Isegawa Y, Sasao F, Ueda S. 1993 A common neutralizing epitope conserved between the hemagglutinins of influenza A virus H1 and H2 strains. *J. Virol.* **67**, 2552–2558.
 27. Sandbulte MR *et al.* 2011 Discordant antigenic drift of neuraminidase and hemagglutinin in H1N1 and H3N2 influenza viruses. *Proc. Natl Acad. Sci. USA* **108**, 20 748–20 753. (doi:10.1073/pnas.1113801108)
 28. Hancock K *et al.* 2009 Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. *N. Engl. J. Med.* **361**, 1945–1952. (doi:10.1056/NEJMoa0906453)
 29. Itoh Y *et al.* 2009 In vitro and in vivo characterization of new swine-origin H1N1 influenza viruses. *Nature* **460**, 1021–1025.
 30. Bandaranayake D *et al.* 2010 Risk factors and immunity in a nationally representative population following the 2009 influenza A(H1N1) pandemic. *PLoS ONE* **5**, e13211. (doi:10.1371/journal.pone.0013211)
 31. Chi CY, Liu CC, Lin CC, Wang HC, Cheng YT, Chang CM, Wang J-R. 2010 Preexisting antibody response against 2009 pandemic influenza H1N1 viruses in the Taiwanese population. *Clin. Vaccine Immunol.* **17**, 1958–1962. (doi:10.1128/CVI.00212-10)
 32. Gilbert GL, Cretikos MA, Hueston L, Doukas G, O'Toole B, Dwyer DE. 2010 Influenza A (H1N1) 2009 antibodies in residents of New South Wales, Australia, after the first pandemic wave in the 2009 southern hemisphere winter. *PLoS ONE* **5**, e122562. (doi:10.1371/journal.pone.0012562)
 33. Ikonen N, Strengell M, Kinnunen L, Osterlund P, Pirhonen J, Broman M, Davidkin I, Zeigler T, Julkunen I. 2010 High frequency of cross-reacting antibodies against 2009 pandemic influenza A(H1N1) virus among the elderly in Finland. *Euro Surveill.* **15**, 19478.
 34. Mak GC, Choy PW, Lee WY, Wong AH, Ng KC, Lim W. 2010 Sero-immunity and serologic response to pandemic influenza A (H1N1) 2009 virus in Hong Kong. *J. Med. Virol.* **82**, 1809–1815. (doi:10.1002/jmv.21895)
 35. Miller E, Hoschler K, Hardelid P, Stanford E, Andrews N, Zambon M. 2010 Incidence of 2009 pandemic influenza A H1N1 infection in England: a cross-sectional serological study. *Lancet* **375**, 1100–1108. (doi:10.1016/S0140-6736(09)62126-7)
 36. New South Wales Public Health Network. 2009 Progression and impact of the first winter wave of the 2009 pandemic H1N1 influenza in New South Wales, Australia. *Euro Surveill.* **14**, 19365.
 37. Karageorgopoulos DE, Vouloumanou EK, Korbila IP, Kapaskelis A, Falagas ME. 2011 Age distribution of cases of 2009 (H1N1) pandemic influenza in comparison with seasonal influenza. *PLoS ONE* **6**, e21690. (doi:10.1371/journal.pone.0021690)
 38. Kash JC, Qi L, Dugan VG, Jagger BW, Hrabal RJ, Memoli MJ, Morens DM, Taubenberger JK. 2010 Prior infection with classical swine H1N1 influenza viruses is associated with protective immunity to the 2009 pandemic H1N1 virus. *Influenza Other Respir. Viruses* **4**, 121–127. (doi:10.1111/j.1750-2659.2010.00132.x)
 39. Manicassamy B, Medina RA, Hai R, Tsibane T, Stertz S, Nystal-Villán E, Palese P, Basler CF, García-Sastre A. 2010 Protection of mice against lethal challenge with 2009 H1N1 influenza A virus by 1918-like and classical swine H1N1 based vaccines. *PLoS Pathog.* **6**, e1000745. (doi:10.1371/journal.ppat.1000745)
 40. Min JY, Chen GL, Santos C, Lamirande EW, Matsuoka Y, Subbarao K. 2010 Classical swine H1N1 influenza viruses confer cross protection from swine-origin 2009 pandemic H1N1 influenza virus infection in mice and ferrets. *Virology* **408**, 128–133. (doi:10.1016/j.virol.2010.09.009)
 41. O'Donnell CD, Wright A, Vogel L, Wei CJ, Nabel GJ, Subbarao K. 2012 The effect of priming with H1N1 influenza viruses of variable antigenic distance on challenge with 2009 pandemic H1N1 virus. *J. Virol.* **86**, 8625–8633. (doi:10.1128/JVI.00147-12)
 42. McCullers JA, Van De Velde LA, Allison KJ, Branum KC, Webby RJ, Flynn PM. 2010 Recipients of vaccine against the 1976 'swine flu' have enhanced neutralization responses to the 2009 novel H1N1 influenza virus. *Clin. Infect. Dis.* **50**, 1487–1492. (doi:10.1086/652441)
 43. Skountzou I *et al.* 2011 Immunity to Pre-1950 H1N1 influenza viruses confers cross-protection against the pandemic swine-origin 2009 A (H1N1) influenza virus. *J. Immunol.* **185**, 1642–1649. (doi:10.4049/jimmunol.1000091)
 44. Wikramaratna PS, Gupta S. 2009 Influenza outbreaks. *J. Cell. Microbiol.* **11**, 1016–1024. (doi:10.1111/j.1462-5822.2009.01320.x)
 45. Tsuchiya E, Sugawara K, Hongo S, Matsuzaki Y, Muraki Y, Li ZN, Nakamura K. 2001 Antigenic structure of the haemagglutinin of human influenza A/H2N2 virus. *J. Gen. Virol.* **82**, 2475–2484.
 46. Vincent AL, Lager KM, Ma W, Lekcharoensuk P, Gramer MR, Loiacono C, Richt JA. 2006 Evaluation of hemagglutinin subtype 1 swine influenza viruses from the United States. *Vet. Microbiol.* **118**, 212–222. (doi:10.1016/j.vetmic.2006.07.017)
 47. Chen H *et al.* 2006 Establishment of multiple sublineages of H5N1 influenza virus in Asia: implications for pandemic control. *Proc. Natl Acad. Sci. USA* **103**, 2845–2850. (doi:10.1073/pnas.0511120103)
 48. Coudeville L, Bailleux F, Riche B, Megas F, Andre P, Ecochard R. 2010 Relationship between haemagglutination-inhibiting antibody titres and clinical protection against influenza: development and application of a Bayesian random-effects model. *BMC Med. Res. Methodol.* **10**, 18. (doi:10.1186/1471-2288-10-18)
 49. Osterholm MT, Kelley NS, Sommer A, Belongia EA. 2012 Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. *Lancet*

- Infect. Dis.* **12**, 36–44. (doi:10.1016/S1473-3099(11)70295-X)
50. Beare AS, Hobson D, Reed SE, Tyrrell DA. 1968 A comparison of live and killed influenza-virus vaccines. Report to the Medical Research Council's Committee on influenza and other respiratory virus vaccines. *Lancet* **2**, 418–422. (doi:10.1016/S0140-6736(68)90463-7)
 51. Sato K, Morishita T, Nobusawa E, Tonegawa K, Sakae K, Nakajima S, Nakajima K. 2004 Amino-acid change on the antigenic region B1 of H3 haemagglutinin may be a trigger for the emergence of drift strain of influenza A virus. *Epidemiol. Infect.* **132**, 399–406. (doi:10.1017/S0950268803001821)
 52. Lillie PJ *et al.* 2012 Preliminary assessment of the efficacy of a T-cell-based influenza vaccine, MVA-NP + M1, in humans. *Clin. Infect. Dis.* **55**, 19–25. (doi:10.1093/cid/cis327)
 53. Wiley DC, Wilson IA, Skehel JJ. 1981 Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* **289**, 373–378. (doi:10.1038/289373a0)
 54. Gupta V, Earl DJ, Deem MW. 2006 Quantifying influenza vaccine efficacy and antigenic distance. *Vaccine* **24**, 3881–3888. (doi:10.1016/j.vaccine.2006.01.010)
 55. Pan K, Subieta KC, Deem MW. 2011 A novel sequence-based antigenic distance measure for H1N1, with application to vaccine effectiveness and the selection of vaccine strains. *Protein Eng. Des. Sel.* **24**, 291–299. (doi:10.1093/protein/gzq105)
 56. Bush RM, Fitch WM, Bender CA, Cox NJ. 1999 Positive selection on the H3 hemagglutinin gene of human influenza. *Mol. Biol. Evol.* **16**, 1457–1465. (doi:10.1093/oxfordjournals.molbev.a026057)
 57. Lee MS, Chen MC, Liao YC, Hsiung CA. 2007 Identifying potential immunodominant positions and predicting antigenic variants of influenza A/H3N2 viruses. *Vaccine* **25**, 8133–8139. (doi:10.1016/j.vaccine.2007.09.039)
 58. Arinaminpathy N, Grenfell B. 2010 Dynamics of glycoprotein charge in the evolutionary history of human influenza. *PLoS ONE* **5**, e155674. (doi:10.1371/journal.pone.0015674)
 59. Caton AJ, Brownlee GG, Yewdell JW, Gerhard W. 1982 The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell* **31**, 417–427. (doi:10.1016/0092-8674(82)90135-0)
 60. Kaverin NV, Rudneva IA, Govorkova EA, Timofeeva TA, Shilov AA, Kochergin-Nikitsky KS, Krylov PS, Webster RG. 2007 Epitope mapping of the hemagglutinin molecule of a highly pathogenic H5N1 influenza virus by using monoclonal antibodies. *J. Virol.* **81**, 12 911–12 917. (doi:10.1128/JVI.01522-07)
 61. Kryazhimskiy S, Dushoff J, Bazykin GA, Plotkin JB. 2011 Prevalence of epistasis in the evolution of influenza A surface proteins. *PLoS Genet.* **7**, e1001301. (doi:10.1371/journal.pgen.1001301)
 62. Kucharski A, Gog JR. 2012 Influenza emergence in the face of evolutionary constraints. *Proc. Biol. Sci.* **279**, 645–652. (doi:10.1098/rspb.2011.1168)
 63. Rudneva I *et al.* 2012 Escape mutants of pandemic influenza A/H1N1 2009 virus: variations in antigenic specificity and receptor affinity of the hemagglutinin. *Virus Res.* **166**, 61–67. (doi:10.1016/j.virusres.2012.03.003)
 64. Xu R, Ekiert DC, Krause JC, Hai R, Crowe Jr JE, Wilson IA. 2010 Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science* **328**, 357–360. (doi:10.1126/science.1186430)
 65. Greenbaum JA *et al.* 2009 Pre-existing immunity against swine-origin H1N1 influenza viruses in the general human population. *Proc. Natl Acad. Sci. USA* **106**, 20 365–20 370. (doi:10.1073/pnas.0911580106)
 66. Khurana S *et al.* 2009 Antigenic fingerprinting of H5N1 avian influenza using convalescent sera and monoclonal antibodies reveals potential vaccine and diagnostic targets. *PLoS Med.* **6**, e1000049. (doi:10.1371/journal.pmed.1000049)
 67. Stray SJ, Pittman LB. 2012 Subtype- and antigenic site-specific differences in biophysical influences on evolution of influenza virus hemagglutinin. *Viol. J.* **9**, 91. (doi:10.1186/1743-422X-9-91)
 68. Soundararajan V, Zheng S, Patel N, Warnock K, Raman R, Wilson IA, Raguram S, Sasisekharan V, Sasisekharan R. 2011 Networks link antigenic and receptor-binding sites of influenza hemagglutinin: mechanistic insight into fitter strain propagation. *Sci. Rep.* **1**, 200. (doi:10.1038/srep00200)
 69. Baigent SJ, McCauley JW. 2001 Glycosylation of haemagglutinin and stalk-length of neuraminidase combine to regulate the growth of avian influenza viruses in tissue culture. *Virus Res.* **79**, 177–185. (doi:10.1016/S0168-1702(01)00272-6)
 70. Ndifon W, Wingreen NS, Levin SA. 2009 Differential neutralization efficiency of hemagglutinin epitopes, antibody interference, and the design of influenza vaccines. *Proc. Natl Acad. Sci. USA* **106**, 8701–8706. (doi:10.1073/pnas.0903427106)
 71. De Franco A, Locksley R, Robertson M. 2007 *Immunity: The immune response in infectious and inflammatory disease*. Oxford, UK: Oxford University Press.
 72. Treanor J. 2004 Influenza vaccine—outmanoeuvring antigenic shift and drift. *N. Engl. J. Med.* **350**, 218–220. (doi:10.1056/NEJMp038238)
 73. Brammer TL, Izurieta HS, Fukuda K, Schmeltz LM, Regnery HL, Hall HE, Cox NJ. 2000 Surveillance for influenza—United States, 1994–95, 1995–96, and 1996–97 seasons. *MMWR. CDC. Surveill. Summ.* **49**, 13–28.
 74. Zinder D, Bedford T, Gupta S, Pascual M. 2013 The roles of competition and mutation in shaping antigenic and genetic diversity in influenza. *PLoS Pathog.* **9**, e1003104. (doi:10.1371/journal.ppat.1003104)