bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The molecular basis of antigenic variation among A(H9N2) avian influenza viruses

Thomas P. Peacock[1,2#a¶], William T. Harvey[3¶], Jean-Remy Sadeyen[1], Richard Reeve[3*] and Munir Iqbal[1]

[1] Influenza Group, The Pirbright Institute, Pirbright, Woking, UK
[2] Department of Virology, Imperial College London, UK
[3] Boyd Orr Centre for Population and Ecosystem Health, Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK
[#a] Current address: Division of Infection and Immunity, University College London, UK

¶These authors contributed equally to this work.
* Address correspondence to richard.reeve@glasgow.ac.uk

Keywords: avian influenza; A(H9N2); antigenic evolution; haemagglutinin; receptor-binding avidity; glycosylation; vaccine selection.

Abstract

Avian influenza A(H9N2) viruses are a threat to global poultry production as well as human health through zoonotic infection and are therefore considered viruses with pandemic potential. Vaccination of poultry is a key element of disease control in endemic countries and human vaccination would be a major component of the response in a pandemic situation. Vaccine effectiveness is however persistently challenged by the emergence of antigenically variant H9N2 viruses. Here we employed a combination of techniques to provide an enhanced understanding of the genetic basis of H9N2 antigenic variability and evaluate the role of different molecular mechanisms of immune escape. We collated every published H9N2 monoclonal antibody escape mutant and systematically tested their influence on polyclonal chicken antiserum binding, determining that many have no significant effect in this vital context. Amino acid substitutions introducing additional glycosylation sites were a notable exception; however, these are relatively rare among circulating viruses. To identify substitutions responsible for antigenic variation among circulating viruses, we performed an integrated meta-analysis of all published H9 haemagglutinin sequences and antigenic data from serological assays. We validated this statistical analysis experimentally using reverse genetics and allocated several new residues to H9N2 antigenic sites using a panel of previously characterised monoclonal antibodies. These results provide new molecular markers of antigenic change for H9N2 viruses that will help explain vaccine breakdown in the field. Conventionally, changes to epitope structure determine antigenicity. We find evidence for the importance of other mechanisms of immune escape, with substitutions increasing glycosylation or receptor-binding avidity exhibiting the largest impacts on chicken antisera binding. Of these, meta-analysis indicates avidity regulation to be more relevant to the evolution of circulating viruses, suggesting that a specific focus on avidity regulation is required to fully understand the molecular basis of immune escape by influenza, and potentially other viruses.

Introduction

In recent years novel avian influenza virus (AIV) strains have emerged as a major threat to animal and human health. H9N2 AIVs are endemic across much of Asia, the Middle East, North and West Africa, where they cause severe economic losses to the poultry industry through drops in egg production, reduced broiler growth rates, as well as moderate to high morbidity and mortality [1-3]. Additionally, H9N2 AIVs are considered to have pandemic potential due to their repeated isolation from humans and their ability to adapt to and transmit between ferrets, the animal model for human transmission of influenza [4-7]. Certain H9N2 virus internal gene constellations have a unique capacity to grant non-H9N2 viruses an

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

elevated zoonotic potential with recent examples including H7N9, H10N8, and H5N6, all known to have high mortality rates in humans [8-10], indicating high levels of H9N2 endemicity are linked to human infections with both H9N2 and as well as divergent AIV strains.

Due to the threat to both poultry and human health posed by endemic H9N2, many countries use poultry vaccination as a major method of viral control, with conventional inactivated, adjuvanted vaccines being used most commonly in the field [11-14]. However, in a similar manner to human influenza virus, poor vaccine seed strain matching due to antigenic drift often results in vaccine failure in the field [3, 11, 15-18]. To mitigate and prevent this failure, better understanding of the molecular basis of antigenicity and antigenic drift in H9N2 viruses is needed to maximise the efficacy of existing vaccine technologies, as well guide the production of novel vaccine platforms.

Immunity to influenza, both after natural infection and vaccination, is primarily achieved by the generation of neutralising antibodies able to block the receptor binding activity of the major influenza antigen, haemagglutinin (HA) [19]. Antibodies interacting with the globular HA domain sterically block attachment of HA to the terminal sialic acid residues on target cells, neutralising viral infectivity. Consequently, antigenic variability of influenza viruses, which is commonly assessed using the haemagglutination inhibition (HI) assay, can largely be mapped to amino acid substitutions in and around the HA receptor binding site (RBS) [20, 21]. These substitutions have been found to contribute to both actual and apparent antigenic change via a variety of mechanisms, namely changes to epitope structure, acquisition of additional glycosylation sites, and modulation of receptor-binding avidity.

Amino acid substitutions that alter the properties (shape, charge, polarity, etc.) of an epitope have the potential to cause antigenic change by directly affecting antibody binding. This manifests as a straightforward drop in titre to homologous antiserum, and often an increase in titre to antisera generated against viruses that contain the introduced residue. A complication arises when certain amino acids, usually charged or bulky, are present within an epitope and generate a robust antibody response; while alternative amino acids, usually small and/or uncharged, occur at this position in other virus strains that generate a weaker antibody response against the equivalent epitope. Consequently antisera generated against 'dominant-epitope'-containing antigens will show a large change in titre when the amino acid is substituted, while the 'sub-dominant-epitope'-raised antisera will show a minimal, or no change, in titre [22].

The acquisition of additional N-linked glycosylation sites is another mechanism by which influenza viruses may escape antibody neutralisation [23]; bulky oligosaccharides can sterically 'shield' HA epitopes from antibody recognition. Glycan acquisition is also expected to reduce HI titre to homologous antisera if neutralising antibodies are inhibited and, less well described, potentially increase titres to antisera generated against viruses with the additional glycans, if the antibody response is directed to alternative antigenic sites in the presence of the additional glycans [24]. Additional glycosylation has been described to reduce HA receptor-binding avidity, in the absence of compensatory mutations, which has been hypothesised to restrict hyperglycosylation as an immune evasion strategy [23, 25].

Neutralising antibodies act as competitive inhibitors of HA-sialylated receptor interactions. Amino acid substitutions that modulate receptor-binding avidity can therefore contribute to apparent changes in antigenicity detected by HI assays, whereby stronger HA binding to sialylated receptors outcompetes antibody binding to HA. This manifests as a consistent change in titre to all antisera tested (a decrease or increase depending on whether avidity is increased or decreased, respectively), which can appear antigenic when only antisera containing or without the substitution (but not both) are examined. Modulation of receptor avidity has recently been hypothesised to be a true form of immune escape rather than an artefact of the HI assay [22, 26].

In our previous study we took a 'classical' mAb escape mutant approach to attempt to dissect the antigenic architecture of the H9 HA [27]. We identified several antigenically important

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

residues that were assigned to two discrete antigenic sites, 'H9-A' and 'H9-B'. Although these residues were important for the binding of murine mAbs, their significance in the context of circulating field viruses remained unknown. Many of the residues found in our previous study, as well as several similar published studies [28-32], were completely conserved amongst circulating viruses indicating they could not be responsible for the antigenic variability observed in the field.

In this paper, we identify amino acid residues and substitutions responsible for H9N2 virus antigenic variability in the field through a variety of approaches. We reconstitute mutations at every published H9 HA monoclonal escape mutant position [27-31] and test the effect of these substitutions on polyclonal chicken antisera binding. We then adapt an approach previously developed for human influenza viruses [20], performing a meta-analysis of all published H9 HI assay and corresponding genetic data to identify naturally-occurring antigenically-important substitutions. We correlate amino acid substitutions that occur alongside antigenic changes, and validate these substitutions using reverse genetics and HI assays. Several novel antigenic sites are predicted, which we demonstrate are indeed relevant to the antigenic diversity of field H9N2 viruses. Overall, this study provides a comprehensive and systematic analysis of the molecular basis of H9 HA antigenic diversity in chickens which are routinely vaccinated by the poultry industry and hence the critical host species for understanding risks of vaccine escape.

Results

Antigenic impact of escape mutant viruses

Six published H9N2 studies describe a total of 39 unique mAb escape mutations in H9 HA1 across 30 different amino acid positions [27-32] (S1 Table). Viruses were generated with mutations at each published escape residue in the HA of the H9N2 virus A/chicken/Pakistan/UDL-01/2008 (UDL1/08) and tested using HI, with a panel of 8 anti-UDL1/08 chicken antisera, to assess the impact of these mutations on antigenicity (Fig 1, S1 Table). When multiple escape mutants were reported at the same position, the mutation, or mutations, that introduced the amino acid with the largest difference in biophysical properties were generated (for example, T129K and T129A are both known escape mutants, however only T129K was tested in this study). Several published mAb escape mutants were unable to be rescued in the background of UDL1/08 and were therefore dropped from further study (S2 Table, Fig 1). Furthermore, several escape mutant residues had alternative mutations introduced that were more biologically relevant to field viruses (25, or 64%, of the 39 published escape mutations were found to be absent or almost absent among sequenced viruses, S3 Table) as well as likely to have larger effects due to differences in biophysical properties (for example S109R was made instead of S108I).

In total, 44 mutants across 30 amino acid positions described in escape mutant studies were assessed by HI assay (Fig 1). Of these 44 mutations, 25 at 20 positions showed significant reductions in HI titre compared to wild type UDL/108 (G72E, R74G, Q115P, T127N, T129K, K131I, D135G, F137L, T145I, Q146H (+T186K), N148D, G149D, G149K, L150S, T179N, T182R, N183D, N183S, N183T, T188N, D189N, L216Q, I217L, I217M, and I217Q). Several of the introduced substitutions – 19 at 16 amino acid positions – caused no significant change in HI titre (R74K, L98Q, S109R, Q115R, T120K, T127S, K131A, K131S, R139G, R139M, K147T, L150A, L150F, P152L, R162W, D178Y, L212P, L217T, and R234Q). Interestingly, of the seven substitutions (T127N, T145I, L150S, T188N, D189N, L216Q, and I217Q) with the largest impact on chicken antiserum binding (≥1.30 $\log_2$ HI titre), four corresponded to the addition of glycosylation sites (T127N, L150S, T188N, and D189N), while the remaining three corresponded to known receptor binding residues (T145I, L216Q and I217Q) [5, 33-35].

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

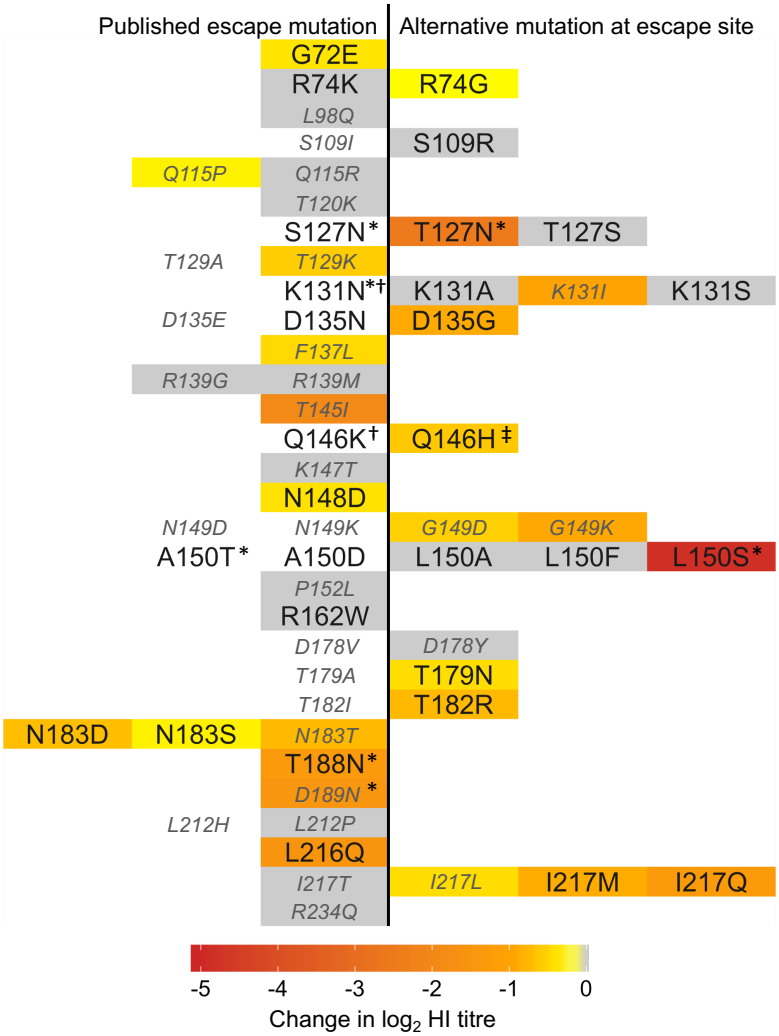

Fig 1. Impact of mutation at published mAb escape sites on haemagglutination inhibition titres performed with polyclonal chicken antisera. Change in $\log_2$ HI titre associated with stated amino acid replacements, based on 8 individual anti-UDL1/08 antisera with a minimum of 5 technical repeats, are indicated by cell colouring according to the colour legend with non-significant differences shown in grey and untested substitutions in white. Published mAb escape mutations are shown to the left of centre while alternative substitutions are shown to the right. Mutations that introduce absent or almost absent amino acid states (<1% of sequenced natural isolates) are labelled in smaller, italicised, grey font. * addition of potential glycosylation site. † would not rescue. ‡ Q146H rescued with additional substitution T186K.

The four successfully rescued escape mutants predicted to add N-linked glycosylation sites (K131N did not rescue) all resulted in large, significant changes in HI titre. Each of these four mutants (T127N, L150S, T188N, and D189N) were tested for the presence of the additional glycans by Western blot and all displayed upward band shifts, consistent with the addition of glycosylation sites (S1 Fig), as described elsewhere [32]. Presence of N-linked glycan motifs at the sites of the four glycosylation site mutants tested are, however, rare among sequenced viruses: 127-129 (6.8%), 148-150 (1.5%), 188-190 (2.6%) and 189-191 (0.1%) (S4 Table). It should be noted that while individual mutations with the potential to add glycosylation sites may be relatively common, the presence of a complete motif across three positions is much rarer (e.g. 150S is present in 11.8% of sequences, however the motif 148N-149X(not P)-150S/T is present in only 1.5% of sequences).

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

These HI results indicate that while several escape mutations identified using murine mAbs do modulate chicken antiserum binding, a sizable group do not. Furthermore, of the tested mutations found to significantly affect chicken antiserum binding, several are unable to contribute significantly to antigenic variability among circulating strains due to their rareness in nature. For example, 25 of 39 published escape mutations are absent or almost absent (frequency <1%) amongst sequenced H9 HAs and the positions 115, 129, 145, and 189 (as well as non-significant positions 98, 137, 147, 152, 212, and 234) are all >98% conserved (S3 Table). Amino acid identity, for a sample of viruses covering all major H9N2 lineages, at each of the 20 escape mutant sites where a significant impact on HI titre was detected is shown in Fig 2 (a similar figure including all 30 escape mutant sites is shown in S2A Fig). Of the 39 published escape mutations, 27 (69%) were either found to have no significant impact on HI titres, to be absent or almost absent among sequenced natural isolates, or both (Fig 1, S3 Table). At 10 of the 30 published escape sites, no evidence of an antigenic impact in HI assays was detected, even when taking alternative substitutions at the site into account. These results suggest that chicken polyclonal antisera do not consistently recognise the same epitopes as mouse mAbs and that further sites are likely to be contributing to antigenic variation of H9N2 viruses seen in the field.

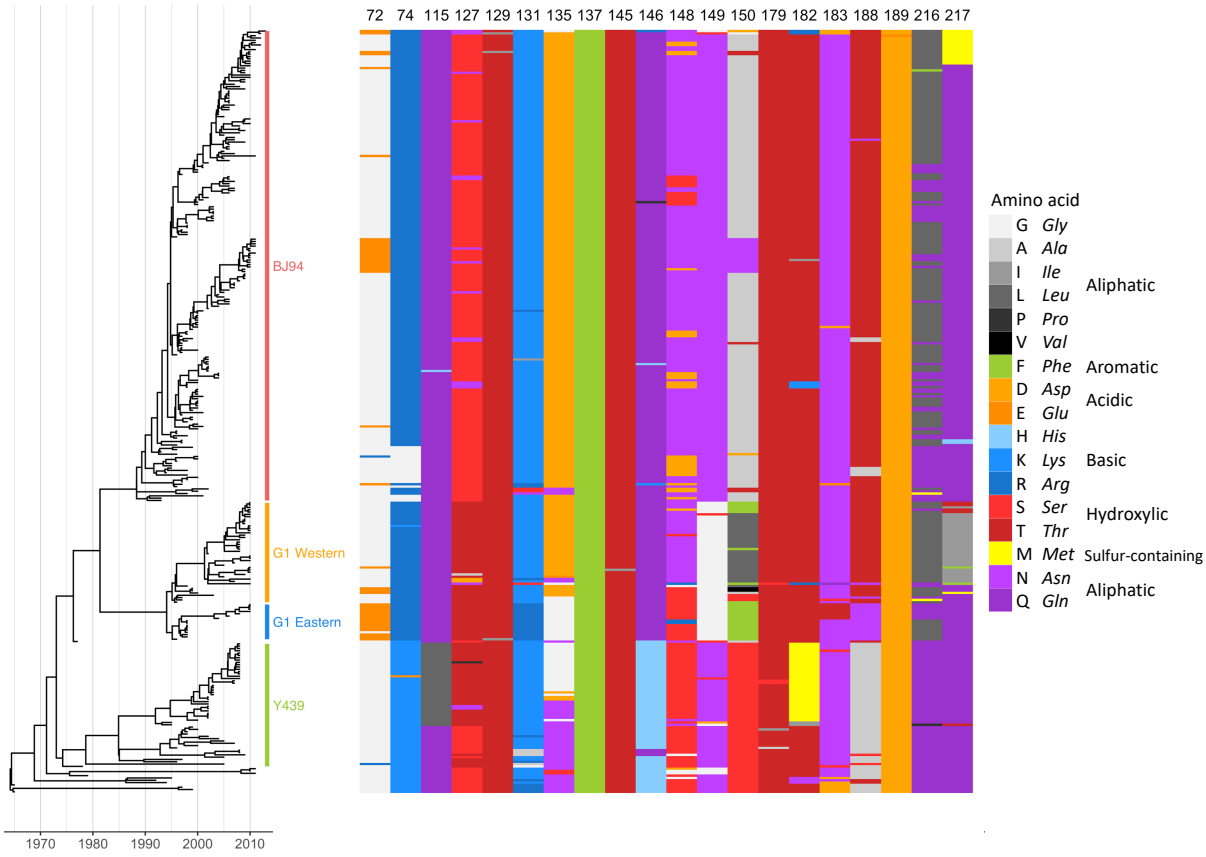

Fig 2. Sequence variation at antigenically significant H9 haemagglutinin escape mutant sites. Time-resolved Bayesian HA1 phylogeny and amino acid identity at each of the 20 mAb escape residues found to significantly affect HI titres with polyclonal chicken antisera. Amino acid identity is shown by colour, grouped by side-chain property, according to the legend. Each virus (n = 330) included in the phylogeny has associated HI data and was included in integrated modelling of genetic and antigenic data.

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Modelling of HI and phylogenetic data identified and prediction of antigenically pertinent substitutions

To better understand the genetic differences responsible for H9N2 field strain antigenic diversity and vaccine escape, we performed an integrated statistical analysis of antigenic and genetic data to predict antigenically pertinent residues and substitutions found in the field. Such an approach has previously been used to identify amino acid substitutions determining the antigenic phenotype of human influenza A and foot-and-mouth disease viruses [20, 36]. All published, as well as some previously unpublished, HI data for H9N2 and the associated HA1 nucleotide sequences were compiled and supplemented with additional HI assays (S5 and S6 Tables) [3, 4, 7, 11, 15, 17, 37-76]. The assembled HI dataset comprised of 330 H9N2 virus sequences, covering all major lineages (Fig 2), including 103 viruses against which antisera were raised with a total of 2,574 individual HI titres measured between 2,131 unique combinations of virus and antiserum.

Table 1. Amino acids substitutions correlating with antigenic change and corresponding antigenic sites.

HA position[a]	Described, or equivalent antigenic site			Antigenically distinct amino acids
	H9	H1	H3	
72	H9-B	Cb	E	E-G-R
74	UC[b]	Cb	E	G-R
121	-[c]			I-T
131	UC		A	K-A/I/R/S, A-R
135	Site II	Ca_2	A	D-G-N, D-E
150	Site I	Sb	B	A/F/N-V/L, F-S-T
180	-	Sb	B	A-E-T, A-V, E-D
183	Site II/H9-A	Sb	B	N-D/E/T
195	-			A-T
198	-	Ca_1	D	D-N
216	Site II	Ca_2	D	L-Q
217	H9-A		D	I-Q/T
249	-			A/V-I
264	-		C	K-N
276	-			K-R
288	-		C	I-V
306	-			I-K

[a]H9 mature numbering used throughout; [b]UC indicates known antigenic residue (by escape mutant residue) in H9 not previously categorised into a specific antigenic site; [c]dash indicates site not previously described as being antigenic.

To identity antigenic relationships and their genetic predictors, we used linear mixed models to account for variation in log_2 HI titres in terms of viral sequence changes. Initial model selection determined variables representing systematic differences in the magnitudes of titres recorded for individual test viruses, for antisera raised against particular reference viruses, for different serum types (i.e. live virus vs inactivated vaccination, or chicken vs ferret), and for different studies. The per-test-virus differences likely represent differences in the receptor-

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

binding avidity of the viruses, while the per-reference-virus changes may represent changes in the immunogenicity of different reference viruses. To control for repeated measures that arise from phylogenetic correlations in the data and prevent false positive support for substitutions, branches of the HA1 phylogeny correlated with drops in HI titre were then identified, and then terms representing substitutions at each variable HA1 position were tested to determine whether their inclusion significantly improved model fit as published previously [20].

This analysis identified antigenically distinct amino acids at 12 HA1 positions (74, 121, 131, 135, 150, 180, 183, 195, 198, 216, 249, and 288). Of these 12 positions, seven had been identified in published H9N2 mouse mAb escape studies (74, 131, 135, 150, 183, 188, and 216), including positions in previously described antigenic sites (Table 1, S1 Table) [27, 29]. Several additional residues were identified when ferret or chicken antisera datasets were analysed independently and were also considered (72, 217, 264, 276, and 306), two of which (72 and 217) had also been described by escape mutant studies [27, 28]. Among the specific substitutions inferred to be antigenically relevant, G72E, K131N, N183T, and L216Q have previously been identified in mAb escape mutant studies; G72E, N183T, and L216Q (K131N could not be rescued) were also successfully confirmed to significantly reduce binding by polyclonal chicken antisera in HI assays above (Fig 1, S1 Table). Amino acid identity at each of the 17 HA positions identified using this modelling approach is shown for each virus included in the analysis in S2B Fig.

Multiple mutants affected virus viability and required additional substitutions to be rescued

To test the effect of the antigenic residues identified in this study on polyclonal antisera binding, mutant libraries were generated in the backgrounds of three antigenically distinct H9N2 viruses (S6 Table), UDL1/08 (as used above), A/Hong Kong/33982/2009 (HK/33982) and A/chicken/Emirates/R66/2002 (Em/R66). Wherever possible mutations were made that were reciprocal between these viruses (i.e. UDL1/08 naturally has 183N and HK/33982 has 183T, substitution between N and T at 183 is predicted to have an antigenic effect therefore UDL1/08 N183T and HK/33982 T183N were generated and tested). When it was possible to introduce reciprocal amino acid substitutions, both mutant viruses were tested using antiserum raised against viruses which naturally had both 'sides' of the substitution. For example, both UDL1/08 N183T and HK/33982 T183N were tested against both UDL1/08 and HK/33982 antisera. HI titres for reciprocal amino acid substitutions tested against antisera possessing and lacking each introduced substitution were analysed to discriminate between the effect of mutation on antigenicity and receptor-binding avidity.

While attempting to rescue the mutant virus libraries using reverse genetics, several mutants were found to produce no infectious virus, some viruses upon sequencing after rescue were shown to have gained compensatory mutations or showed partial, or complete, reversion of the intended substitution, even after three independent rescue attempts (S2 Table). The most marked group of unrescuable mutants were attempted mutations of known and potential RBS residues, for example: the mutant UDL1/08 A180E was unable to be rescued without the addition of either L216Q, I217L, or I217Q; the mutant UDL1/08 A180D partially reverted to D180G; the mutant HK/33982 D180A gained an additional substitution of G215D and HK/33982 Q216L gained the additional substitution of H174E (S2 Table). These observations suggest that the receptor binding site residues of 180, 216, and 217, and to a lesser extent 174 and 215, have important interactions with each other and the context in which these residues are found is extremely important for the viability of the virus (S3 Fig).

Antigenic characterisation of mutant viruses to test predictions

To validate the residues identified by modelling of HI and genetic data, we tested mutant viruses using HI assays and compared titres to those measured using parental viruses. Changes in mean $\log_2$ titre, relative to the parent virus, against both parental and mutant-like

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

antiserum, when available, are detailed in S7 Table. Of the 26 amino acid substitutions from Table 2 introduced, 19 resulted in significantly different titres measured when using parental antiserum (G72E, E72G, R74G, I121T, T121I, K131I, D135G, A150L, A180E, E180A, N183D, N183T, T183N, L216Q, Q216L, I217Q, Q217I, I249V and V249I), while seven did not (K131A, K131S, G135D, L150A, L150F, F150L and D198N). For substitutions introduced into the UDL1/08 background at published escape mutation sites, significant effect sizes are represented in Fig 1 and where reciprocal forward and reverse mutations were introduced in different backgrounds, effect sizes are shown in Fig 3. In Fig 3, significantly lower or higher HI titres recorded for mutant viruses, relative to the parental virus, are shown in shades of blue and red respectively. The first column (left) shows changes in titre against the parental antiserum are generally drops (reduced antigenic recognition) and the second column shows that changes in titre against antiserum possessing the introduced substitution are generally increases (increased antigenic recognition). Of the 22 mutant viruses that could be tested with antisera possessing the introduced mutation, 15 were associated with a significant change in titre to mutant-like antiserum (Fig 3).

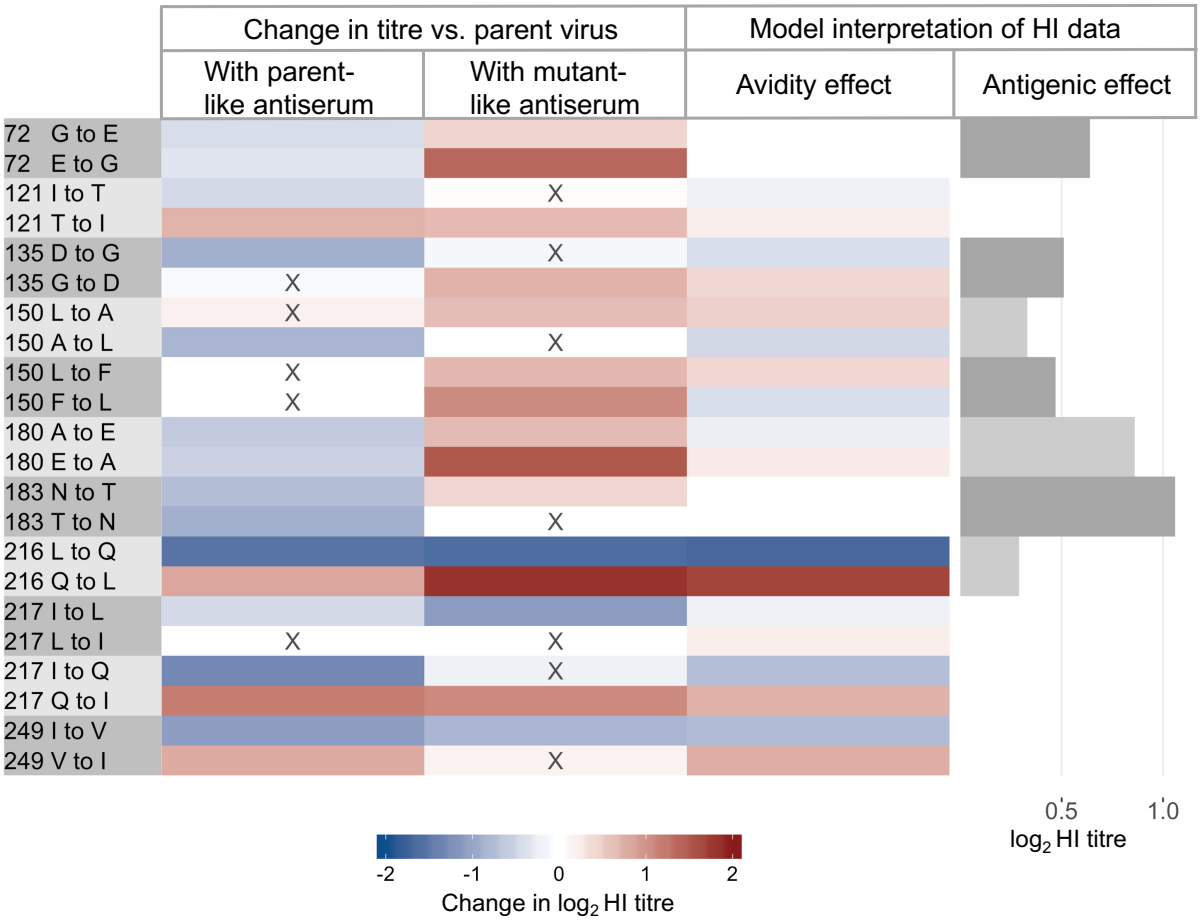

Fig 3. Heat-map showing the impacts of reciprocal amino acid substitutions introduced at sites identified by modelling. Changes in HI titre (log_2) associated with each introduced substitution are shown by heat map, measured using antiserum both lacking the introduced substitution (parent-like) and possessing the introduced mutation (mutant-like) from combinations of the viruses UDL1/08, Em/R66 and HK/33982. Changes in HI titre associated with each pair of substitutions were interpreted as effects of altered avidity and antigenicity using Equation 1, are also shown. Avidity effects cause an increase (red) or drop (blue) in titre depending on the direction of substitution while antigenic effects are the estimated change in titre resulting from antigenic dissimilarity. X indicates not individually significant.

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In addition to antigenic change, individual amino acid substitutions can influence HI titres due to induced changes in other virus properties, notably receptor-binding avidity [26]. By testing mutant viruses using antisera raised against the parent virus and antisera raised against a virus possessing the introduced amino acid state, changes in titres between parent and mutant viruses resulting from antigenic and avidity effects could be distinguished using Equation 1 (Materials and Methods). The impact on HI titre of altered antigenicity and avidity associated with each reciprocal pair of introduced substitutions is shown in Fig 3 and S7 Table. Of the 11 reciprocal pairs considered, two were estimated to influence HI titres as a result of altered antigenicity alone (72 G/E and 183 N/T), four as a result of altered avidity only (121 I/T, 217 L/Q, 217 I/Q and 249 I/V) and five as a combination of these effects (135 D/G, 150 A/L, 150F/L, 180 A/E and 216 L/Q).

The impact of substitutions between A and E at position 180 show a classic antigenic signature. Both A180E and E180A were associated with significant drops in $\log_2$ titre of 0.61 and 0.54 measured using antiserum raised to the parental virus, UDL1/08 (180A) and Em/R66 (180E) respectively (blue cells in column 1, Fig 3). Both A180E and E180A also increased titres, by 0.65 and 1.51 respectively, measured using mutant-like antiserum possessing the introduced amino acid at position 180 (red cells in column 2, Fig 3). These changes in titre indicate that the A180E substitution has decreased and increased the reactivity of antibodies raised against 180A and 180E viruses, respectively; a consistent pattern was observed with the E180A substitution. Consequently, a sizeable antigenic effect (0.86 $\log_2$ titre) is estimated alongside a smaller avidity effect (+/-0.25 $\log_2$ titre). Similar patterns of change in titres indicated sizeable antigenic effects for substitution between G and E at position 72 and between N and T at 183.

In the case of substitution at 183, both N183T and T183N drops in titre, of 0.75 and 0.90 $\log_2$ titre respectively, against antisera raised to viruses lacking the introduced substitution were observed. Increases in titre to antisera raised to viruses possessing the introduced substitution were less dramatic (0.42 and no significant change respectively). This pattern of change indicates that perhaps further antigenic changes to UDL1/08 and HK33982 are required for antibodies raised against one of them to bind to epitopes of the other virus that possess the introduced state at 183, most likely the additional introduction of a change at 180. This is implied by its proximity to position 183 as well as the fact the monoclonal antibody JF7 loses binding to mutant UDL1/08 viruses containing either N183T or A180E (S8 Table), however due to difficulty rescuing UDL1/08 A180D and HK/33982 D180A (S2 Table), this couldn't be conclusively shown.

Exchange of Q and L at position 216 represents the clearest example of substitutions impacting HI titres due to altered receptor-binding avidity. The substitution L216Q reduces titres against antisera both lacking the introduced amino acid state, UDL1/08 (216L) and possessing it, Em/R66 (216Q), by (blue cells in columns 1 and 2, Fig 3). Conversely, Q216L increases titres against antisera both lacking (Em/R66) and possessing (UDL1/08) the introduced amino acid state (red cells in columns 1 and 2, Fig 3). These changes indicate that an H9 HA possessing 216Q has higher avidity than a virus with 216L, resulting in lower HI titres generally regardless of whether antiserum was raised to a 216Q or 216L virus. Similar effects of avidity on HI titres were observed at positions 121, 217 and 249. Interestingly, L216Q, I217L and I217Q, which each result in drops in HI titre due to altered avidity (blue cells, Fig 3) are either published mAb escape mutations or occur at the same position as escape mutations (Fig 1, S1 Table). This indicates that increased avidity facilitated by single amino acid substitutions is a mechanism of immune escape for H9, an observation consistent with previous studies of human H1N1 influenza virus using murine mAbs [26].

Results of HI assays indicated that substitutions at position 150 impacted titres via changes to both antigenicity and avidity fairly evenly (Fig 3). This was also estimated to be the case at position 135, however closer observation of the pattern of change in HI titre suggests an asymmetric antigenic effect of substitutions between aspartic acid and glycine at this position. The average impacts on titre of either D135G or G135D were significant and large when

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

measured using antisera raised against the UDL1/08 (135D) virus (-0.90 and +0.73). In contrast, no significant impact on titres were observed using antisera raised against the HK/33982 (135G) virus, consistent with 135D belonging to an immunodominant epitope. A model that estimated two antigenic effects of substitutions between D and G, depending on which antiserum was used, rather than a single antigenic effect was found to improve model fit as assessed by deviance information criterion [77]. Experimental studies have previously demonstrated an asymmetric effect at the homologous position in human A(H3N2) viruses; viruses possessing 145K were found to have an antigenically dominant epitope including position 145 [78]. The substitution K145N reduced antibody binding with antisera raised against viruses possessing 145K, however the introduction of the N145K substitution did not reduce antibody binding with antisera raised against an H3 virus possessing 145N. The same asymmetric model was also favoured in an analysis of H3 HI titres measured using K145N and N145K mutants producing results consistent with those reported by Li *et al*. (2013) (Supplementary Information) [78]. No further clear evidence of substitutions causing asymmetric antigenic change was found in this study.

Mapping of substitutions into antigenic sites using mAbs

Previously, we described a panel of seven mouse mAbs against UDL1/08 binding to two discrete antigenic sites, 'H9-A' and 'H9-B' [27]. These mAbs were used to perform HI assays against every mutant virus generated in this study to assess whether any of the positions described here could be categorised into these antigenic sites giving a better understanding of the antigenic architecture of H9 HA. If a mutant virus had altered titres to mAbs from either group of mAbs, but not the other, the residue was assigned to the corresponding antigenic site.

Reductions in titres of H9-A binding mAbs were seen in mutants G149K, T179N, A180D and A180E indicated positions 149, 179 and 180 likely form part of antigenic site H9-A alongside residues 145, 146, 183, 212, 217, and 234. Titres recorded for mutant viruses with substitutions G72E and D135G indicated these positions form part of site H9-B, alongside residues 115, 120, 139 and 162 (Fig 4A, Table 2, S8 Table). The four glycosylation mutants (T127N, L150S, T188N, and D189N) reduced titres of both H9-A and H9-B mAbs. It is likely the addition of bulky glycans in the glycosylation mutants may also have the effect of being able to directly block antibody binding to both antigenic sites. The substitutions K131I, D178Y, T182R, L216Q, I217M and I217Q also appeared to slightly reduce binding of mAbs from both groups. These latter residues may either lie between antigenic sites and exert an effect on both, or, these substitutions may overcome antibody binding by increasing HA receptor-binding avidity. The latter hypothesis is supported in the case of L216Q and I217Q as the analysis of reciprocal mutant HI titres indicated that substitutions at these positions exerted an effect on HI titres as a result of altered avidity (Fig 3).

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

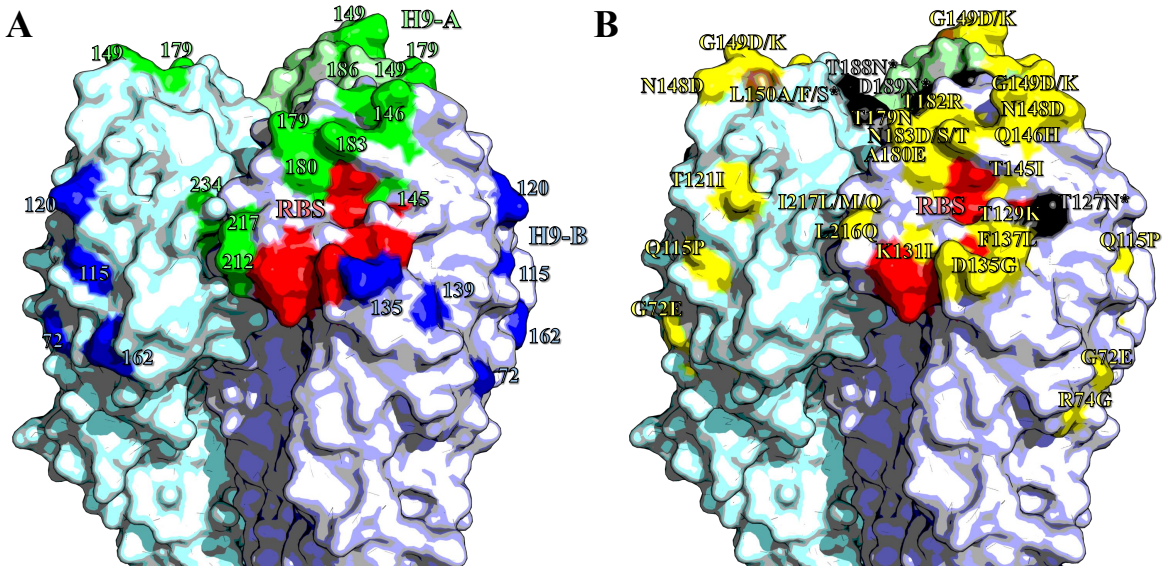

Fig 4. Antigenic structure of H9 HA. Homotrimers of H9 HA. Selected RBS residues shown in red (P92, G128, T129, S130, S131, A132, W142, N173, L184, Y185, N214, G215, G218 and R219). Images made in Pymol [99] (Schrödinger) based on the structure of A/swine/Hong Kong/9/1998 (Protein databank ID:1JSD) [100]. (A) Residues recognised by mouse mAbs, positions with updated site H9-A shown in green, H9-B residues shown in blue. (B) Residues labelled with substitutions that affect the binding of chicken polyclonal antisera. Non-glycosylation altering substitutions shown in yellow; glycosylation site adding mutations shown in black; site 150, which had both glycosylation adding and non-glycosylation adding mutations shown in brown.

Overall these results support the observation that while mouse mAbs and chicken polyclonal antisera often appear to recognize similar epitopes, in many cases different substitutions are important for modulation of chicken antiserum binding and mouse mAb binding. A summary of the results of this paper detailing the impacts of substitutions on mouse mAb and chicken antiserum binding, relevance to circulating viruses, and mapping to revised H9 antigenic sites is shown in Table 2, while the locations of all residues at which substitutions were detected to impact HI assays are shown on the structure of H9 HA in Fig 4B.

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. Summary of residues investigated in this study and their relation to antigenicity and avidity.

Position	Antigenic site	Amino acid diversity[a]	Published mAb escape[b]	Identified by modelling[c]	Impact on HI[d]	Antigenic effect[e]	Avidity effect[e]
72	H9-B	1.92	G-E	E-G-R	G-E	G-E	-
74		1.29	~~R-K~~	R-G	R-G		
98		1.00	~~*L-Q*~~	-	-		
109		2.37	~~S-R~~	-	-		
115	H9-B	1.03	*Q-P*, ~~*Q-R*~~	-	*Q-P*		
120	H9-B	1.26	~~*T-K*~~	-	-		
121		1.06	-	I-T	I-T	-	I-T
127	Overlap	2.72	S-N*	-	T-N*		
129	Site I	1.01	*T-A*, *T-K*	-	*T-K*		
131		1.75	K-N*	~~K-A/S~~, *K-I*, K-R, A-R	*K-I*		
135	H9-B/Site II	2.18	*D-E*, D-N	D-G-N, D-E	D-G	D-G	D-G
137		1.00	*F-L*	-	*F-L*		
139	H9-B	1.12	~~*R-G*~~, ~~*R-M*~~	-	-		
145	H9-A	1.01	*T-I*	-	*T-I*		
146	H9-A/Site I	2.02	Q-K	-	Q-H		
147		1.02	~~*K-T*~~	-	-		
148	Site I	1.65	N-D	-	N-D		
149	H9-A/Site I	2.12	*N-D*, *N-K*	-	*G-D*, *G-K*		
150	Site I	3.93	A-T*, A-D	A/N-V/L, A-F, ~~F-L~~, F-S-T	A-L, L-S*	L-A, A-F	L-A, A-F
152	Site I	1.00	~~*P-L*~~	-	-		
162	H9-B	2.37	~~R-W~~	-	-		
178	Site II	1.13	*D-V*	-	-		
179	H9-A/Overlap	1.06	*T-A*	-	T-N		
180	H9-A	3.45	-	A-E-T, A-V, E-D	A-E	A-E	A-E
182	Site II	1.73	*T-I*	-	T-R		
183	H9-A/Site II	1.77	N-D, N-S, *N-T*	N-D, *N-E/T*	*N-T*	*N-T*	-
188	Overlap	1.29	T-N*	A-T	T-N*		
189	Site II	1.01	*D-N**	-	*D-N**		
195		1.70	-	A-T			
198		2.32	-	~~D-N~~	-		
212	H9-A	1.03	*L-H*, ~~*L-P*~~	-	-		
216	Site II	1.49	L-Q	L-Q	L-Q	L-Q	L-Q
217	H9-A	2.59	~~*I-T*~~	I-Q, ~~*I-T*~~	I-L, I-M, I-Q	-	I-L, I-Q
234	H9-A	1.02	~~*R-Q*~~	-	-		
249		1.42	-	I-V/A	I-V	-	I-V
264		1.85	-	K-N			
276		1.41	-	K-R			
288		1.44	-	I-V			
306		1.02	-	*I-K*			

[a]Inverse Simpson index, 1.00 equivalent to 100% conserved, numbers above 1.00 indicate increasing levels of diversity; [b]Published mAb escape mutant, citations in S1 Table. [c]Identified by integrated modelling of HI titres and gene sequence data (Table 1). [d]Significant impact on HI titre as shown in Figs 1 and/or 3. Estimated impact of antigenicity and avidity change on HI titre if reciprocal mutations were analysed (Fig 2). Substitutions associated with absent or almost absent (frequency <1% among sequenced viruses) are shown in italicised font. Substitutions identified by mAb escape study or

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

modelling but shown not to impact HI titres are indicated by strikethrough. * Potential to add glycosylation site. Shaded cells indicate not tested.

Discussion

Determining the basis of antigenic evolution is of central importance to understanding viral disease dynamics, and vaccine selection and development. We present a systematic analysis of the molecular basis of antigenicity of H9N2 HA using a combination of approaches. Initially we assessed the contribution of residues reported by the literature to be antigenic and their relation to virus antigenic diversity. This showed that previous studies based on murine monoclonal antibody escape mutants are only moderately reliable at providing relevant antigenic information; several false positives found to have no impact on inhibition of chicken antisera were identified, as were residues that are very highly conserved in the field. We also carried out a meta-analysis of all available matching genetic and antigenic data for H9N2 viruses, modelling the relationship between genotype and antigenic phenotype to identify novel residues directly responsible for observed antigenic diversity in the field, and we categorized these residues into recently defined antigenic sites. Analysis of reciprocal mutations introduced by reverse genetics allowed us to differentiate multiple mechanisms facilitating immune escape including increased glycosylation and receptor-binding avidity as well as conventional changes to epitope structure that directly affect antibody binding.

Of all the mutant viruses assessed in their reactivity to chicken antiserum in this study, the substitutions responsible for the eight largest drops in HI titres were associated with two distinct mechanisms of immune evasion. The first group of four (T127N, L150S, T188N and D189N) possessed additional glycosylation sites and represented every such substitution considered in this study. The second group (T145I, L216Q, I217Q, and I249V) possessed substitutions either known to influence receptor binding avidity in H9N2 viruses or determined to do so in this study through analysis of reciprocal mutants, or both [5, 33-35, 79, 80]. These observations demonstrate the potential of alternative mechanisms, other than substitutions that directly reduce antibody recognition by altering epitope biophysics, in facilitating immune escape and determining antigenic diversity of H9N2 viruses. In the analysis of antigenic and genetic data collected from H9N2 field strains, substitutions altering avidity were well represented among those identified suggesting that avidity regulation may be an important mechanism of immune escape in nature. In contrast, substitutions associated with acquisition of additional glycans were largely absent; indeed, the rareness of additional glycosylation sites in nature is quite stark suggesting this to be a less important mechanism, possibly due to an associated fitness cost, as reported for human influenza [23]. Although it is worth noting addition of glycans has also been shown to be able to restore optimal receptor avidity after avidity altering mutations suggesting a possible synergy between these two escape mechanisms [25].

Implementing a systematic approach allows for the comprehensive evaluation of the phenotypic impact of amino acid substitutions. For example, we identify substitutions at the novel antigenic residue 180 (including A-E) that explain significant variation in HI titres carried out using a range of viruses sampled across the H9N2 phylogeny. The analysis of reciprocal substitutions introduced by reverse genetics indicated that the substitution A180E influenced HI titres primarily as a result of an actual antigenic change and not by affecting receptor-binding avidity, while epitope mapping using mAbs indicated that residue 180 is a constituent of the H9-A antigenic site. Furthermore, we find that the amino acid present at position 180 may mask the antigenic effect of substitution at 183, potentially indicating the combinations of amino acids required for antibody binding can be determined from the analysis of HI data.

We demonstrate that, in addition to detecting substitutions that directly affect antibody binding by altering the biophysical structure of epitopes, we can also detect other effects from HI data, such as avidity and immunodominance. Specifically, we found a clear example of an asymmetric antigenic effect, whereby substitutions between aspartic acid and glycine at position 135 (D135G and G135D) were observed to impact titres measured using antisera

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

raised to a D135 virus but not using antisera raised to a G135 virus. These results suggest that an epitope involving position 135 is immunodominant when a virus possesses D135 (aspartate has a prominent acidic side chain, CH_2COO^-, likely to interact with other proteins), but that antibody responses generated against viruses possessing G135 are directed towards regions other than amino acid 135 (which is unsurprising, given that glycine is the smallest amino acid). Substantial change in immunodominance resulting from a single amino acid substitution has previously been reported in human influenza [78]. While we did not detect any other cases of asymmetric antigenic change, we speculate that the phenomenon is perhaps widespread though hard to detect, except in the most dramatic cases. Better understanding of this phenomenon may facilitate the development of vaccines designed to guide the immune response towards more conserved epitopes providing better cross-protection.

In contrast to previously validated applications of the described modelling technique to human influenza A(H1N1) and foot-and-mouth disease virus (FMDV) where very large antigenic datasets (> 10,000 titres) were collected by a single laboratory under relatively consistent methods [20, 81], the data analysed here were drawn from multiple sources and encompass varying experimental designs, serum types, and production methods. This is expected to limit the accuracy of parameter coefficient estimates and potentially introduce false positives in terms of identified residues (e.g. at position 140). Nonetheless, we identify substitutions within previously defined H9 antigenic sites at new amino acid positions and confirm them experimentally as antigenically important. This indicates that such an approach can be extended to influenza subtypes where multiple, smaller databases of antigenic data exist, which may be particularly useful for emerging H5 and H7 viruses.

In conclusion, we show that while mAb escape mutant studies are undoubtedly helpful for understanding virus antigenicity, escape mutations often have little or no effect on inhibition by polyclonal antiserum drawn from chickens, the target species of vaccination. Additionally, many escape mutants are never found to emerge in nature, perhaps due to negative fitness consequences, and are therefore not relevant to the antigenic diversity of circulating viruses. Conversely, integrated modelling of sequence and HI data directly provides information on the molecular basis of antigenic variation specifically for circulating viruses, and its conclusions can be easily validated by reverse genetics. We anticipate that these results will contribute to future vaccine development and seed strain assessment by providing quantitative molecular markers that can help explain vaccine breakdown in the field and predict levels of antigenic drift during virus surveillance.

Materials and methods

Ethics statement

All described animal studies and procedures were carried out in strict accordance with European and United Kingdom Home Office regulations and the Animals (Scientific Procedures) Act 1986 Amendment Regulations, 2012. These studies were carried out under the United Kingdom Home Office approved project license numbers 30/2683 and 30/2952. Additionally, the work had undergone ethical scrutiny before approval by The Pirbright Institute's Animal Welfare and Ethical Review Board (AWERB) under the request numbers AR0000509 and AR000610. Work in 10-day-old embryonated eggs did not require ethical approval as eggs were schedule 1 euthanised before day 14 as stated in the Animals (Scientific Procedures) Act 1986 Amendment Regulations, 2012.

Cells and eggs

MDCKs and HEK 293Ts (obtained from the European Collection of Authenticated Cell Cultures) were grown in Dulbecco modified Eagle media (DMEM) supplemented with 10% foetal bovine serum (FBS) at 37°C, 5% CO_2. Virus was propagated in 10-day-old embryonated chicken eggs (ValoBioMedia GmbH), allantoic and amniotic fluid was harvested and pooled at 48 hours post-inoculation.

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Cloning and rescue of recombinant viruses

All viruses used in this study were generated through reverse genetics as previously described [27, 82]. Recombinant, reassortant viruses were used for generation of chicken antisera, whereby the HA of the named virus was rescued with the remaining segments from A/chicken/Pakistan/UDL-01/2008 (UDL1/08). For all other assays, high growth reassortant 1:1:6 viruses were used where the HA was taken from the named virus, the NA from UDL1/08, and the remaining 6 internal genes from A/Puerto Rico/8/1934 (PR8). Viruses with amino acid substitutions were generated through site directed mutagenesis. After propagation all viruses were Sanger sequenced in the HA1 region to determine if any reversions/compensatory mutations had arisen, as previously described [27].

Antisera production

Post-infection chicken antisera were generated as described previously [83]. Briefly, 3-week-old specific pathogen free chickens were infected intranasally with 10^6 pfu of recombinant virus. Birds were monitored twice daily for clinical signs. At 21 days post-infection birds were euthanised and bled for antisera. All other antisera were either purchased commercially or received as gifts from collaborators.

Haemagglutination and Haemagglutination inhibition assay

Haemagglutination and HI assays were performed as described elsewhere [84], using a solution of 1% chicken red blood cells. All HI assays with mutant viruses were performed with a minimum of three technical repeats with six individual chicken antisera. Linear mixed-effects models were used to determine whether HI titres recorded for mutant viruses were significantly different from those recorded for the parental virus. Mutant viruses were considered in a combined model with fixed effects for each combination of parental virus and for each introduced substitution. Random effects were used to account for any variation in titres attributable to differences in antiserum drawn from different chickens (biological repeats) and in differences between experiments (technical repeats).

Variation in HI titre associated with introduced substitutions resulting from changes in antigenicity and avidity was estimated using a similar methodology to that previously described [20]. When reciprocal substitutions were made (arbitrarily termed forward and reverse for the purpose of explanation), both mutant viruses were tested by HI against antiserum raised against the parental virus lacking the forward substitution (X) and against antiserum raised against the parental virus lacking the reverse substitution (Y). The impact of the introduced substitution on antigenicity (d) and avidity (v) was estimated from differences in titres associated with introduced substitutions:

$$\begin{aligned} H_{XX^+} &\sim N(H_{XX} - d + v, \sigma^2) \\ H_{YX^+} &\sim N(H_{YX} + kd + v, \sigma^2) \\ H_{YY^-} &\sim N(H_{YY} - d - v, \sigma^2) \\ H_{XY^-} &\sim N(H_{XY} + kd - v, \sigma^2) \end{aligned} \quad (1)$$

where titres recorded for each mutant virus are assumed to be normally distributed ($N()$) with a mean depending on titres recorded for the parental virus and the antigenic and avidity effects of the introduced substitution and variance σ^2. H_{AB} is the $\log_2$ HI titre for the virus B tested with the antiserum raised against the virus A, and X^+ and Y^- are the reciprocal mutant viruses possessing forward and reverse substitutions, respectively. Thus, $H_{XY'}$ is the $\log_2$ HI titre for the mutant virus Y^- and antiserum raised against the virus X. The full antigenic effect of an introduced substitution is apparent when measured using parental antiserum, but that effect may be obscured to some degree when measured using an antiserum raised against a distinct virus as a result of other antigenic differences between the parental and distinct viruses. k is

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

estimated as the proportion of the full antigenic impact of an introduced substitution that is observed (*i.e.* not obscured) when measured using antiserum raised against a distinct virus. To investigate a potential asymmetric antigenic effect an alternative model that allowed d on lines one and four of Equation 1 to differ from d on lines two and three. Parameters in the above equations were co-estimated in a Bayesian model with minimally informative priors using JAGS v3.3.0 using the R package *runjags* v2.0.3.2 [85, 86].

Western blotting of glycosylation mutant viruses

Escape mutants with additional potential glycosylation sites were rescued and grown in eggs as described above. Virus-containing allantoic fluid was concentrated by ultracentrifugation at 27,000 rpm for 2 hours. Concentrated viruses were then run on 7.5% SDS-PAGE gels and Western blotted. A cocktail of anti-UDL1/08 mAbs, the previously described JF7, IB3, ID2 and IG10 [27], were used to probe HA1, a band shift was then used to determine presence or absence of glycosylation sites.

Genetic and antigenic data

Haemagglutination inhibition data was collated from a combination of published literature, WHO candidate vaccine virus reports, publicly available theses and previously unpublished data from our laboratory [3, 4, 7, 11, 15, 17, 37-76]. (see S1 Table). Available HA sequences from H9N2 isolates analysed in this study were downloaded from the NCBI influenza database (https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=database) or the GISAID EpiFlu database (http://platform.gisaid.org). Only HI data associated with complete HA1 nucleotide sequences was included for analysis. The HI dataset matched to sequence data comprised 2,547 individual titre measurements between 328 test viruses and antisera raised against 103 different virus strains. In total 2,131 different virus-antisera combinations were represented in the dataset. The antisera used in HI titres analysed was drawn from various species and included both post-infection and post-vaccination (hyperimmune) antisera. 868 titres were measured using chicken hyperimmune antisera, 836 using chicken post-infection, 796 using ferret post-infection, 60 using rabbit hyperimmune, seven sheep hyperimmune, and seven goat hyperimmune.

Phylogenetic analysis

HA1 nucleotide sequences were aligned using MUSCLE alignment software [87]. These genetic data and year of isolation were analysed and time resolved phylogenetic trees were reconstructed using BEAST v1.8.2 [88]. The goodness-of-fit of 88 models of nucleotide substitution were evaluated prior to analysis in BEAST using jModelTest v2.1.7 [89, 90]. The preferred model from this analysis and similar models were tested in BEAST through comparison of Bayes factors [91]. The general time reversible model of nucleotide substitution with a proportion of invariant sires and a gamma distribution describing among-site rate variation with four categories estimated from the data (GTR + I + Γ_4) was identified as the best model of nucleotide substitution. Bayes factor analysis also determined that a codon-position model that allowed rates of nucleotide substitution to vary at the third codon position relative to the first and second positions, a relaxed (uncorrelated) molecular clock with branch rates drawn from a lognormal distribution [92], and a minimally constrained Bayesian skyline demographic model should be used [93]. The maximum clade credibility tree was visualised alongside amino acid alignments using the *ggtree* package [94].

Statistical modelling and model selection

Substitutions responsible for changes in antigenic phenotype during the evolution of the virus were predicted using a modelling approach adapted from that described by Harvey et al. (2016) [20]. The geometric ($\log_2$) HI titre was used as the response variable throughout. Mixed models constructed using the *lme4* package [95] and R v3.3.0 [96] including each of the following variables were assessed by likelihood ratio test: the test virus, the reference virus against which the antiserum was raised, the type of antisera used which encompasses

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

species from which antisera was drawn and method of inoculation (*i.e.* post-infection or post-vaccination), and the study from which data was collected. To prevent false support for substitutions due to repeated measurements that arise due to the evolutionary relationships between viruses, phylogenetic information was included in the model. Branches of the phylogeny associated with antigenicity changing events were identified and added to the model as previously described [20, 36, 81]. Additionally, phylogenetic terms associated with changes in receptor-binding avidity and immunogenicity were identified. These branches led to clades of viruses/antisera that tended to have higher or lower titres, notwithstanding antigenic relationships. Due to the size of the search space, the optimal combination of these various binary variables reflecting phylogenetic structure was determined using random restart hill-climbing as previously described [20, 97].

To investigate the effect of amino acid substitutions at specific positions, the model was then extended with a term indicating the presence, or absence, of substitution between the reference virus and test virus. Amino acid dissimilarity at each variable HA1 position was tested as a predictor of reduced HI titre ($p < 0.05$), both before and after correcting for multiple testing [98]. Positions were tested alongside terms accounting for phylogenetic structure, therefore substitution at identified position must correlate with variation in HI titre in independent branches of the phylogeny: i.e. there are convergent, alternative, or back-substitutions at the same HA1 position associated with reduced cross-reactivity. To identify antigenically distinct amino acids at identified positions, the binary term indicating presence, or absence, of substitution was replaced with a variable with levels for every pair of amino acids observed between two viruses tested together by HI. Estimates of the antigenic impact of specific substitutions were made by examining associated regression coefficients for substitution terms included in models individually alongside phylogenetic terms, and in combination in the absence of phylogenetic terms.

Diversity

Diversity across different amino acid positions was represented using Inverse Simpson's index. Inverse Simpson's index was calculated using all 2523 full length H9 HA sequenced downloaded from the NCBI influenza database (https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=database) as of July 2017.

References

1. Jakhesara SJ, Bhatt VD, Patel NV, Prajapati KS, Joshi CG. Isolation and characterization of H9N2 influenza virus isolates from poultry respiratory disease outbreak. SpringerPlus. 2014;3:196. doi: 10.1186/2193-1801-3-196. PubMed PMID: 24790833; PubMed Central PMCID: PMC4004788.
2. Lee YN, Lee DH, Park JK, Lim TH, Youn HN, Yuk SS, et al. Isolation and characterization of a novel H9N2 influenza virus in Korean native chicken farm. Avian diseases. 2011;55(4):724-7. doi: 10.1637/9774-050911-Case.1. PubMed PMID: 22313001.
3. Pu J, Wang S, Yin Y, Zhang G, Carter RA, Wang J, et al. Evolution of the H9N2 influenza genotype that facilitated the genesis of the novel H7N9 virus. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(2):548-53. doi: 10.1073/pnas.1422456112. PubMed PMID: 25548189; PubMed Central PMCID: PMC4299237.
4. Sorrell EM, Wan H, Araya Y, Song H, Perez DR. Minimal molecular constraints for respiratory droplet transmission of an avian-human H9N2 influenza A virus. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(18):7565-70. doi: 10.1073/pnas.0900877106. PubMed PMID: 19380727; PubMed Central PMCID: PMC2670882.
5. Li X, Shi J, Guo J, Deng G, Zhang Q, Wang J, et al. Genetics, receptor binding property, and transmissibility in mammals of naturally isolated H9N2 Avian Influenza viruses. PLoS

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pathogens. 2014;10(11):e1004508. doi: 10.1371/journal.ppat.1004508. PubMed PMID: 25411973; PubMed Central PMCID: PMC4239090.
6. International Centre for Diarrhoeal Disease Research (ICDDRB). Outbreak of mild respiratory disease caused by H5N1 and H9N2 infections among young children in Dhaka, Bangladesh. Health and Science Bulletin: 2011.
7. Peiris M, Yuen KY, Leung CW, Chan KH, Ip PL, Lai RW, et al. Human infection with influenza H9N2. Lancet. 1999;354(9182):916-7. PubMed PMID: 10489954.
8. Wu S, Wu F, He J. Emerging risk of H7N9 influenza in China. Lancet. 2013;381(9877):1539-40. Epub 2013/04/23. doi: 10.1016/S0140-6736(13)60767-9. PubMed PMID: 23602315.
9. Chen H, Yuan H, Gao R, Zhang J, Wang D, Xiong Y, et al. Clinical and epidemiological characteristics of a fatal case of avian influenza A H10N8 virus infection: a descriptive study. Lancet. 2014;383(9918):714-21. doi: 10.1016/S0140-6736(14)60111-2. PubMed PMID: 24507376.
10. Shen YY, Ke CW, Li Q, Yuan RY, Xiang D, Jia WX, et al. Novel Reassortant Avian Influenza A(H5N6) Viruses in Humans, Guangdong, China, 2015. Emerging infectious diseases. 2016;22(8). doi: 10.3201/eid2208.160146.
11. Zhang P, Tang Y, Liu X, Peng D, Liu W, Liu H, et al. Characterization of H9N2 influenza viruses isolated from vaccinated flocks in an integrated broiler chicken operation in eastern China during a 5 year period (1998-2002). The Journal of general virology. 2008;89(Pt 12):3102-12. doi: 10.1099/vir.0.2008/005652-0. PubMed PMID: 19008399.
12. Choi JG, Lee YJ, Kim YJ, Lee EK, Jeong OM, Sung HW, et al. An inactivated vaccine to control the current H9N2 low pathogenic avian influenza in Korea. Journal of veterinary science. 2008;9(1):67-74. PubMed PMID: 18296890; PubMed Central PMCID: PMC2839114.
13. Banet-Noach C, Perk S, Simanov L, Grebenyuk N, Rozenblut E, Pokamunski S, et al. H9N2 influenza viruses from Israeli poultry: a five-year outbreak. Avian diseases. 2007;51(1 Suppl):290-6. doi: 10.1637/7590-040206R1.1. PubMed PMID: 17494569.
14. El Houadfi M, Fellahi S, Nassik S, Guerin JL, Ducatez MF. First outbreaks and phylogenetic analyses of avian influenza H9N2 viruses isolated from poultry flocks in Morocco. Virol J. 2016;13(1):140. doi: 10.1186/s12985-016-0596-1. PubMed PMID: 27527708; PubMed Central PMCID: PMC4986173.
15. Park KJ, Kwon HI, Song MS, Pascua PN, Baek YH, Lee JH, et al. Rapid evolution of low-pathogenic H9N2 avian influenza viruses following poultry vaccination programmes. The Journal of general virology. 2011;92(Pt 1):36-50. doi: 10.1099/vir.0.024992-0. PubMed PMID: 20861321.
16. Ge F, Li X, Ju H, Yang D, Liu J, Qi X, et al. Genotypic evolution and antigenicity of H9N2 influenza viruses in Shanghai, China. Arch Virol. 2016. doi: 10.1007/s00705-016-2767-1. PubMed PMID: 26935915.
17. Wei Y, Xu G, Zhang G, Wen C, Anwar F, Wang S, et al. Antigenic evolution of H9N2 chicken influenza viruses isolated in China during 2009-2013 and selection of a candidate vaccine strain with broad cross-reactivity. Vet Microbiol. 2016;182:1-7. doi: 10.1016/j.vetmic.2015.10.031. PubMed PMID: 26711021.
18. Lee DH, Fusaro A, Song CS, Suarez DL, Swayne DE. Poultry vaccination directed evolution of H9N2 low pathogenicity avian influenza viruses in Korea. Virology. 2016;488:225-31. doi: 10.1016/j.virol.2015.11.023. PubMed PMID: 26655240.
19. Chambers TM, Kawaoka Y, Webster RG. Protection of chickens from lethal influenza infection by vaccinia-expressed hemagglutinin. Virology. 1988;167(2):414-21. PubMed PMID: 3264429.
20. Harvey WT, Benton DJ, Gregory V, Hall JP, Daniels RS, Bedford T, et al. Identification of Low- and High-Impact Hemagglutinin Amino Acid Substitutions That Drive Antigenic Drift of Influenza A(H1N1) Viruses. PLoS pathogens. 2016;12(4):e1005526. doi: 10.1371/journal.ppat.1005526. PubMed PMID: 27057693; PubMed Central PMCID: PMCPMC4825936.
21. Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GC, Vervaet G, et al. Substitutions near the receptor binding site determine major antigenic change during influenza

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

virus evolution. Science. 2013;342(6161):976-9. doi: 10.1126/science.1244730. PubMed PMID: 24264991.
22. Li Y, Bostick DL, Sullivan CB, Myers JL, Griesemer SB, St George K, et al. Single Hemagglutinin Mutations that Alter Both Antigenicity and Receptor Binding Avidity Influence Influenza Virus Antigenic Clustering. Journal of virology. 2013. doi: 10.1128/JVI.01023-13. PubMed PMID: 23824816.
23. Das SR, Hensley SE, David A, Schmidt L, Gibbs JS, Puigbo P, et al. Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(51):E1417-22. doi: 10.1073/pnas.1108754108. PubMed PMID: 22106257; PubMed Central PMCID: PMC3251056.
24. Linderman SL, Chambers BS, Zost SJ, Parkhouse K, Li Y, Herrmann C, et al. Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013-2014 influenza season. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(44):15798-803. doi: 10.1073/pnas.1409171111. PubMed PMID: 25331901; PubMed Central PMCID: PMCPMC4226110.
25. Kosik I, Ince WL, Gentles LE, Oler AJ, Kosikova M, Angel M, et al. Influenza A virus hemagglutinin glycosylation compensates for antibody escape fitness costs. PLoS pathogens. 2018;14(1):e1006796. Epub 2018/01/19. doi: 10.1371/journal.ppat.1006796. PubMed PMID: 29346435; PubMed Central PMCID: PMCPMC5773227.
26. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, et al. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science. 2009;326(5953):734-6. doi: 10.1126/science.1178258. PubMed PMID: 19900932; PubMed Central PMCID: PMC2784927.
27. Peacock T, Reddy K, James J, Adamiak B, Barclay W, Shelton H, et al. Antigenic mapping of an H9N2 avian influenza virus reveals two discrete antigenic sites and a novel mechanism of immune escape. Scientific reports. 2016;6:18745. doi: 10.1038/srep18745. PubMed PMID: 26738561; PubMed Central PMCID: PMC4704030.
28. Okamatsu M, Sakoda Y, Kishida N, Isoda N, Kida H. Antigenic structure of the hemagglutinin of H9N2 influenza viruses. Arch Virol. 2008;153(12):2189-95. Epub 2008/11/08. doi: 10.1007/s00705-008-0243-2. PubMed PMID: 18989614.
29. Kaverin NV, Rudneva IA, Ilyushina NA, Lipatov AS, Krauss S, Webster RG. Structural differences among hemagglutinins of influenza A virus subtypes are reflected in their antigenic architecture: analysis of H9 escape mutants. Journal of virology. 2004;78(1):240-9. PubMed PMID: 14671105; PubMed Central PMCID: PMC303415.
30. Wan Z, Ye J, Xu L, Shao H, Jin W, Qian K, et al. Antigenic mapping of the hemagglutinin of an H9N2 avian influenza virus reveals novel critical amino acid positions in antigenic sites. Journal of virology. 2014;88(7):3898-901. doi: 10.1128/JVI.03440-13. PubMed PMID: 24429369.
31. Zhu Y, Yang D, Ren Q, Yang Y, Liu X, Xu X, et al. Identification and characterization of a novel antigenic epitope in the hemagglutinin of the escape mutants of H9N2 avian influenza viruses. Vet Microbiol. 2015;178(1-2):144-9. doi: 10.1016/j.vetmic.2015.04.012. PubMed PMID: 25934533.
32. Ping J, Li C, Deng G, Jiang Y, Tian G, Zhang S, et al. Single-amino-acid mutation in the HA alters the recognition of H9N2 influenza virus by a monoclonal antibody. Biochemical and biophysical research communications. 2008;371(1):168-71. doi: 10.1016/j.bbrc.2008.04.045. PubMed PMID: 18424263.
33. Wan H, Perez DR. Amino acid 226 in the hemagglutinin of H9N2 influenza viruses determines cell tropism and replication in human airway epithelial cells. Journal of virology. 2007;81(10):5181-91. doi: 10.1128/JVI.02827-06. PubMed PMID: 17344280; PubMed Central PMCID: PMC1900221.
34. Peacock TP, Benton DJ, Sadeyen JR, Chang P, Sealy JE, Bryant JE, et al. Variability in H9N2 haemagglutinin receptor-binding preference and the pH of fusion. Emerging microbes & infections. 2017;6(3):e11. doi: 10.1038/emi.2016.139. PubMed PMID: 28325922.

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

35. Sang X, Wang A, Ding J, Kong H, Gao X, Li L, et al. Adaptation of H9N2 AIV in guinea pigs enables efficient transmission by direct contact and inefficient transmission by respiratory droplets. Scientific reports. 2015;5:15928. doi: 10.1038/srep15928. PubMed PMID: 26552719.
36. Reeve R, Blignaut B, Esterhuysen JJ, Opperman P, Matthews L, Fry EE, et al. Sequence-based prediction for vaccine strain selection and identification of antigenic variability in foot-and-mouth disease virus. PLoS Comput Biol. 2010;6(12):e1001027. doi: 10.1371/journal.pcbi.1001027. PubMed PMID: 21151576; PubMed Central PMCID: PMCPMC3000348.
37. Gerloff NA, Khan SU, Zanders N, Balish A, Haider N, Islam A, et al. Genetically Diverse Low Pathogenicity Avian Influenza A Virus Subtypes Co-Circulate among Poultry in Bangladesh. PLoS One. 2016;11(3):e0152131. doi: 10.1371/journal.pone.0152131. PubMed PMID: 27010791; PubMed Central PMCID: PMC4806916.
38. Zhang Y, Yin Y, Bi Y, Wang S, Xu S, Wang J, et al. Molecular and antigenic characterization of H9N2 avian influenza virus isolates from chicken flocks between 1998 and 2007 in China. Vet Microbiol. 2012;156(3-4):285-93. doi: 10.1016/j.vetmic.2011.11.014. PubMed PMID: 22177969.
39. Wang Y, Davidson I, Fouchier R, Spackman E. Antigenic Cartography of H9 Avian Influenza Virus and its Application to Vaccine Selection. Avian diseases. 2015. Epub November 10, 2015.
40. Sun Y, Pu J, Jiang Z, Guan T, Xia Y, Xu Q, et al. Genotypic evolution and antigenic drift of H9N2 influenza viruses in China from 1994 to 2008. Vet Microbiol. 2010;146(3-4):215-25. doi: 10.1016/j.vetmic.2010.05.010. PubMed PMID: 20685047.
41. Shanmuganatham K, Feeroz MM, Jones-Engel L, Smith GJ, Fourment M, Walker D, et al. Antigenic and Molecular Characterization of Avian Influenza A(H9N2) Viruses, Bangladesh. Emerging infectious diseases. 2013;19(9). doi: 10.3201/eid1909.130336. PubMed PMID: 23968540.
42. Guan Y. Continuing evolution of H9N2 avian influenza A viruses in poultry in southern China. The HKU Scholars Hub: Hong Kong University; 2011.
43. Nomura N, Sakoda Y, Soda K, Okamatsu M, Kida H. An H9N2 influenza virus vaccine prepared from a non-pathogenic isolate from a migratory duck confers protective immunity in mice against challenge with an H9N2 virus isolated from a girl in Hong Kong. The Journal of veterinary medical science / the Japanese Society of Veterinary Science. 2012;74(4):441-7. PubMed PMID: 22123308.
44. SJCEIRS H9N2 Working Group. Assessing the fitness of distinct clades of influenza A (H9N2) viruses. Emerging microbes & infections [Internet]. 2013 Nov 3924558]; 2(11):[e75 p.]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26038443.
45. Li C, Yu K, Tian G, Yu D, Liu L, Jing B, et al. Evolution of H9N2 influenza viruses from domestic poultry in Mainland China. Virology. 2005;340(1):70-83. doi: 10.1016/j.virol.2005.06.025.
46. Xue Y, Wang JL, Yan ZQ, Li GW, Chen SY, Zhang XB, et al. Sequence and phylogenetic analysis of surface protein genes of emerging H9N2 influenza viruses isolated from poultry in two geographical regions of China. Virus genes. 2014;48(3):479-85. doi: 10.1007/s11262-014-1060-1. PubMed PMID: 24682938.
47. Lin YP, Shaw M, Gregory V, Cameron K, Lim W, Klimov A, et al. Avian-to-human transmission of H9N2 subtype influenza A viruses: relationship between H9N2 and H5N1 human isolates. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(17):9654-8. doi: 10.1073/pnas.160270697. PubMed PMID: 10920197; PubMed Central PMCID: PMC16920.
48. Guan Y, Shortridge KF, Krauss S, Webster RG. Molecular characterization of H9N2 influenza viruses: were they the donors of the "internal" genes of H5N1 viruses in Hong Kong? Proceedings of the National Academy of Sciences of the United States of America. 1999;96(16):9363-7. PubMed PMID: 10430948; PubMed Central PMCID: PMC17788.
49. Shen HQ, Yan ZQ, Zeng FG, Liao CT, Zhou QF, Qin JP, et al. Isolation and phylogenetic analysis of hemagglutinin gene of H9N2 influenza viruses from chickens in South China from

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2012 to 2013. Journal of veterinary science. 2015;16(3):317-24. PubMed PMID: 25643797; PubMed Central PMCID: PMC4588017.
50. Liu J, Okazaki K, Ozaki H, Sakoda Y, Wu Q, Chen F, et al. H9N2 influenza viruses prevalent in poultry in China are phylogenetically distinct from A/quail/Hong Kong/G1/97 presumed to be the donor of the internal protein genes of the H5N1 Hong Kong/97 virus. Avian pathology : journal of the WVPA. 2003;32(5):551-60. doi: 10.1080/0307-9450310001596728. PubMed PMID: 14522712.
51. Zhu Y, Hu S, Bai T, Yang L, Zhao X, Zhu W, et al. Phylogenetic and antigenic characterization of reassortant H9N2 avian influenza viruses isolated from wild waterfowl in the East Dongting Lake wetland in 2011-2012. Virol J. 2014;11:77. doi: 10.1186/1743-422X-11-77. PubMed PMID: 24779444; PubMed Central PMCID: PMC4012091.
52. Nguyen DC, Uyeki TM, Jadhao S, Maines T, Shaw M, Matsuoka Y, et al. Isolation and characterization of avian influenza viruses, including highly pathogenic H5N1, from poultry in live bird markets in Hanoi, Vietnam, in 2001. Journal of virology. 2005;79(7):4201-12. doi: 10.1128/JVI.79.7.4201-4212.2005. PubMed PMID: 15767421; PubMed Central PMCID: PMC1061558.
53. Mase M, Eto M, Imai K, Tsukamoto K, Yamaguchi S. Characterization of H9N2 influenza A viruses isolated from chicken products imported into Japan from China. Epidemiology and infection. 2007;135(3):386-91. doi: 10.1017/S0950268806006728. PubMed PMID: 16824249; PubMed Central PMCID: PMC2870580.
54. Huang Y, Li X, Zhang H, Chen B, Jiang Y, Yang L, et al. Human infection with an avian influenza A (H9N2) virus in the middle region of China. Journal of medical virology. 2015;87(10):1641–8 doi: 10.1002/jmv.24231. PubMed PMID: 25965534.
55. Guan Y, Shortridge KF, Krauss S, Chin PS, Dyrting KC, Ellis TM, et al. H9N2 influenza viruses possessing H5N1-like internal genomes continue to circulate in poultry in southeastern China. Journal of virology. 2000;74(20):9372-80. PubMed PMID: 11000205; PubMed Central PMCID: PMC112365.
56. Nomura N, Sakoda Y, Endo M, Yoshida H, Yamamoto N, Okamatsu M, et al. Characterization of avian influenza viruses isolated from domestic ducks in Vietnam in 2009 and 2010. Arch Virol. 2012;157(2):247-57. doi: 10.1007/s00705-011-1152-3. PubMed PMID: 22068881.
57. Choi YK, Seo SH, Kim JA, Webby RJ, Webster RG. Avian influenza viruses in Korean live poultry markets and their pathogenic potential. Virology. 2005;332(2):529-37. doi: 10.1016/j.virol.2004.12.002. PubMed PMID: 15680418.
58. Cameron KR, Gregory V, Banks J, Brown IH, Alexander DJ, Hay AJ, et al. H9N2 subtype influenza A viruses in poultry in pakistan are closely related to the H9N2 viruses responsible for human infection in Hong Kong. Virology. 2000;278(1):36-41. doi: 10.1006/viro.2000.0585. PubMed PMID: 11112478.
59. Cong YL, Wang CF, Yan CM, Peng JS, Jiang ZL, Liu JH. Swine infection with H9N2 influenza viruses in China in 2004. Virus genes. 2008;36(3):461-9. doi: 10.1007/s11262-008-0227-z. PubMed PMID: 18401696.
60. Negovetich NJ, Feeroz MM, Jones-Engel L, Walker D, Alam SM, Hasan K, et al. Live bird markets of Bangladesh: H9N2 viruses and the near absence of highly pathogenic H5N1 influenza. PLoS One. 2011;6(4):e19311. doi: 10.1371/journal.pone.0019311. PubMed PMID: 21541296; PubMed Central PMCID: PMC3082571.
61. Kandeil A, El-Shesheny R, Maatouq AM, Moatasim Y, Shehata MM, Bagato O, et al. Genetic and antigenic evolution of H9N2 avian influenza viruses circulating in Egypt between 2011 and 2013. Arch Virol. 2014;159(11):2861-76. doi: 10.1007/s00705-014-2118-z. PubMed PMID: 24990416; PubMed Central PMCID: PMC4206084.
62. Wernery U, Shanmuganatham KK, Krylov PS, Joseph S, Friedman K, Krauss S, et al. H9N2 influenza viruses from birds used in falconry. Influenza and other respiratory viruses. 2013. doi: 10.1111/irv.12143. PubMed PMID: 23889772.
63. Saito T, Lim W, Suzuki T, Suzuki Y, Kida H, Nishimura SI, et al. Characterization of a human H9N2 influenza virus isolated in Hong Kong. Vaccine. 2001;20(1-2):125-33. PubMed PMID: 11567756.

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

64. Okamatsu M, Nishi T, Nomura N, Yamamoto N, Sakoda Y, Sakurai K, et al. The genetic and antigenic diversity of avian influenza viruses isolated from domestic ducks, muscovy ducks, and chickens in northern and southern Vietnam, 2010-2012. Virus genes. 2013;47(2):317-29. doi: 10.1007/s11262-013-0954-7. PubMed PMID: 23861018.
65. Zhang H, Xu B, Chen Q, Chen Z. Characterization of H9N2 influenza viruses isolated from Dongting Lake wetland in 2007. Arch Virol. 2011;156(1):95-105. doi: 10.1007/s00705-010-0836-4. PubMed PMID: 21053033.
66. Webster RG, Hulse DJ. Microbial adaptation and change: avian influenza. Revue scientifique et technique. 2004;23(2):453-65. PubMed PMID: 15702713.
67. Kim JA, Cho SH, Kim HS, Seo SH. H9N2 influenza viruses isolated from poultry in Korean live bird markets continuously evolve and cause the severe clinical signs in layers. Vet Microbiol. 2006;118(3-4):169-76. doi: 10.1016/j.vetmic.2006.07.007. PubMed PMID: 16930870.
68. World Health Organisation. Antigenic and genetic characteristics of influenza A(H5N1) and influenza A(H9N2) viruses and candidate vaccine viruses developed for potential use in human vaccines WHO Global Influenza Programme. 2010;February 2010.
69. Peiris JS, Guan Y, Markwell D, Ghose P, Webster RG, Shortridge KF. Cocirculation of avian H9N2 and contemporary "human" H3N2 influenza A viruses in pigs in southeastern China: potential for genetic reassortment? Journal of virology. 2001;75(20):9679-86. doi: 10.1128/JVI.75.20.9679-9686.2001. PubMed PMID: 11559800; PubMed Central PMCID: PMC114539.
70. Peiris M, Yam WC, Chan KH, Ghose P, Shortridge KF. Influenza A H9N2: aspects of laboratory diagnosis. J Clin Microbiol. 1999;37(10):3426-7. PubMed PMID: 10488228; PubMed Central PMCID: PMC85595.
71. Dong G, Xu C, Wang C, Wu B, Luo J, Zhang H, et al. Reassortant H9N2 influenza viruses containing H5N1-like PB1 genes isolated from black-billed magpies in Southern China. PLoS One. 2011;6(9):e25808. doi: 10.1371/journal.pone.0025808. PubMed PMID: 21980538; PubMed Central PMCID: PMC3183077.
72. Aamir UB, Wernery U, Ilyushina N, Webster RG. Characterization of avian H9N2 influenza viruses from United Arab Emirates 2000 to 2003. Virology. 2007;361(1):45-55. doi: 10.1016/j.virol.2006.10.037. PubMed PMID: 17157891; PubMed Central PMCID: PMC2735206.
73. World Health Organisation. Antigenic and genetic characteristics of zoonotic influenza viruses and development of candidate vaccine viruses for pandemic preparedness WHO Global Influenza Programme. 2014;February 2014.
74. Wang B, Liu Z, Chen Q, Gao Z, Fang F, Chang H, et al. Genotype diversity of H9N2 viruses isolated from wild birds and chickens in Hunan Province, China. PLoS One. 2014;9(6):e101287. doi: 10.1371/journal.pone.0101287. PubMed PMID: 24979703; PubMed Central PMCID: PMC4076334.
75. Lee YJ, Shin JY, Song MS, Lee YM, Choi JG, Lee EK, et al. Continuing evolution of H9 influenza viruses in Korean poultry. Virology. 2007;359(2):313-23. doi: 10.1016/j.virol.2006.09.025. PubMed PMID: 17056087.
76. Bahari P, Pourbakhsh SA, Shoushtari H, Bahmaninejad MA. Molecular characterization of H9N2 avian influenza viruses isolated from vaccinated broiler chickens in northeast Iran. Tropical animal health and production. 2015;47(6):1195-201. doi: 10.1007/s11250-015-0848-x. PubMed PMID: 26055889.
77. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series B (Statistical Methodology). 2002;64(4):583-639.
78. Li Y, Bostick DL, Sullivan CB, Myers JL, Griesemer SB, St George K, et al. Single Hemagglutinin Mutations that Alter Both Antigenicity and Receptor Binding Avidity Influence Influenza Virus Antigenic Clustering. Journal of virology. 2013;87(17):9904-10. doi: 10.1128/JVI.01023-13. PubMed PMID: 23824816.
79. Teng Q, Xu D, Shen W, Liu Q, Rong G, Li X, et al. A Single Mutation at Position 190 in Hemagglutinin Enhances Binding Affinity for Human Type Sialic Acid Receptor and

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Replication of H9N2 Avian Influenza Virus in Mice. Journal of virology. 2016. doi: 10.1128/JVI.01141-16. PubMed PMID: 27558420.
80. Gambaryan AS, Tuzikov AB, Pazynina GV, Desheva JA, Bovin NV, Matrosovich MN, et al. 6-sulfo sialyl Lewis X is the common receptor determinant recognized by H5, H6, H7 and H9 influenza viruses of terrestrial poultry. Virol J. 2008;5:85. doi: 10.1186/1743-422X-5-85. PubMed PMID: 18652681; PubMed Central PMCID: PMC2515299.
81. Reeve R, Borley DW, Maree FF, Upadhyaya S, Lukhwareni A, Esterhuysen JJ, et al. Tracking the Antigenic Evolution of Foot-and-Mouth Disease Virus. PLoS One. 2016;11(7):e0159360. doi: 10.1371/journal.pone.0159360. PubMed PMID: 27448206; PubMed Central PMCID: PMCPMC4957747.
82. Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. A DNA transfection system for generation of influenza A virus from eight plasmids. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(11):6108-13. doi: 10.1073/pnas.100133697. PubMed PMID: 10801978; PubMed Central PMCID: PMC18566.
83. Peacock TP, Benton DJ, James J, Sadeyen JR, Chang P, Sealy JE, et al. Immune escape variants of H9N2 influenza viruses containing deletions at the haemagglutinin receptor binding site retain fitness in vivo and display enhanced zoonotic characteristics. Journal of virology. 2017. doi: 10.1128/JVI.00218-17. PubMed PMID: 28468875.
84. Webster R, Cox N, Stöhr K. WHO Manual on Animal Influenza Diagnosis and Surveillance. WHO Global Influenza Programme [Internet]. 2002 19th July 2013.
85. Plummer M. Just Another Gibbs Sampler v3.3.0 (JAGS): A program for analysis of Bayesian graphical models using Gibbs sampling. 2012.
86. Denwood M. runjags: Interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. J Stat Softw. 2016;71(9):1-25.
87. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004;5:113. doi: 10.1186/1471-2105-5-113. PubMed PMID: 15318951; PubMed Central PMCID: PMCPMC517706.
88. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution. 2012;29(8):1969-73. doi: 10.1093/molbev/mss075. PubMed PMID: 22367748; PubMed Central PMCID: PMCPMC3408070.
89. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012;9(8):772. doi: 10.1038/nmeth.2109. PubMed PMID: 22847109; PubMed Central PMCID: PMCPMC4594756.
90. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 2003;52(5):696-704. PubMed PMID: 14530136.
91. Suchard MA, Weiss RE, Sinsheimer JS. Bayesian selection of continuous-time Markov chain evolutionary models. Molecular biology and evolution. 2001;18(6):1001-13. PubMed PMID: 11371589.
92. Drummond AJ, Ho SY, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. PLoS Biol. 2006;4(5):e88. doi: 10.1371/journal.pbio.0040088. PubMed PMID: 16683862; PubMed Central PMCID: PMCPMC1395354.
93. Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular biology and evolution. 2005;22(5):1185-92. doi: 10.1093/molbev/msi103. PubMed PMID: 15703244.
94. Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution. 2017;8(1):28-36. doi: 10.1111/2041-210X.12628.
95. Bates D, Maechler, M, Bolker, B, Walker, S, Christensen, RHB et al. lme4: Linear mixed-effects models using 'Eigen' and S4. 2016.
96. R Core Team. R: A language and environment for statistical computing. . Foundation for Statistical Computing, Vienna, Austria.2016.
97. Russell S, Norvig P. Chapter 4. Artificial intelligence: A modern approach: Pearson Education Limited.; 1995.

bioRxiv preprint doi: https://doi.org/10.1101/312967; this version posted May 2, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

98. Holm S. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. 1979;6:65-70.
99. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010.
100. Ha Y, Stevens DJ, Skehel JJ, Wiley DC. X-ray structures of H5 avian and H9 swine influenza virus hemagglutinins bound to avian and human receptor analogs. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(20):11181-6. doi: 10.1073/pnas.201401198. PubMed PMID: 11562490; PubMed Central PMCID: PMC58807.