Response to reviews of "Quantifying the effects of single mutations on viral escape from broad and narrow antibodies to an H1 influenza hemagglutinin" from *Nature Microbiology*

Michael B. Doud, Juhye M. Lee, and Jesse D. Bloom

January 22, 2018

Overview of reviewer response and changes to the manuscript.

We received four reviews for the original submission to *Nature Microbiology*. All four reviewers found the work to be scientifically sound. Two reviewers (#1 and #4) were also enthusiastic about impact. The other two (#2 and #3) concurred about scientific soundness, but were less enthusiastic about impact. The most critical was Reviewer #3, who still wrote that "[t]he study is performed to an extremely high technical standard"—a positive comment to receive from the most negative reviewer!

Because all reviewers found the work to be scientifically sound and at least half found it to be high impact, we request that the paper be considered for *Nature Communications* with the current reviews. Impact is subjective, and the fact that half the reviewers found the paper of sufficient impact for *Nature Microbiology* indicates that it will interest many in the field. We emphasize the following impactful aspects of our study:

- It introduces a rigorous method to compare the potential for immune-escape by single amino-acid mutations across antibodies.
- It is the first paper to completely map how all single amino-acid mutations affect neutralization by broadly neutralizing anti-influenza antibodies.
- It is the first direct head-to-head comparison of immune-escape from anti-head versus anti-stalk antibodies to influenza hemagglutinin.

Below we provide a detailed response to the reviews. We have addressed all the questions / concerns. The only major request was that we show all triplicate neutralization assays and perform statistical analyses, which we now do. The rest of the response is rather long because the reviews were so thorough, but you will see that all points are addressed.

In conclusion, our paper has been evaluated by four reviewers who all found it to be sound, raising only minor points that we have clarified in our revision. The only question was whether our work was of sufficient impact for *Nature Microbiology*, with the reviewers splitting on this question. The *Nature Communications* policies read: "While *Nature* and Nature research journals strive to publish the highest impact research within these disciplines, *Nature Communications* is committed to publishing important advances of significance to specialists within each field." Our study clearly meets these criteria according to the evaluations of all reviewers.

Response to reviews

Below find our responses to the reviews received from *Nature Microbiology*. The original comments are in blue, and our responses are in black.

Thanks for selecting reviewers who cover the full spectrum of relevant expertise.

Reviewer #1 (Remarks to the Author):

In their manuscript Doud and colleagues quantify the effects of single mutations on viral escape from both narrow and broadly binding anti-hemagglutinin antibodies. The methodology is based on a mutant library of an H1N1 strain which are analyzed by deep sequencing after a single infection cycle in the presence or absence of mAb. The data presented in the manuscript is of high significance for the field and the work is technically sound. There are some issues that the authors need to address but no major weaknesses were identified.

We appreciate the fact that the reviewer find our work to be both technically sound and of high impact.

Specific points

1) Multicylce replication might have changed the data somewhat because it would have allowed the virus to acquire additional mutations.

This is true. Our approach systematically quantifies how all *single* mutations affect viral neutralization. This contrasts with an alternative approach of using multiple rounds of growth to select for multi-mutants. There are strengths and weaknesses to both approaches. Examining all single mutations makes the results complete and quantitative—but we can't make claims about the effects of multiple mutations. In contrast, multi-cycle growth sometimes finds combinations of mutations, but is not complete and quantitative, since it is random which mutations arise in any given multi-cycle growth experiment. Therefore, the strength of our approach (examining all single mutations in a way that is comparable across antibodies) comes with the limitation that we do not find multiple mutants.

We have taken the following pains to emphasize this limitation. First, the title of the paper reads "Quantifying the effects of *single mutations* on viral escape from broad and narrow antibodies to an H1 influenza hemagglutinin" (emphasis added).

In addition, we have included the following paragraph in the Discussion:

> *Another important caveat is that our experiments examine* **single** *amino-acid mutations to the HA from one influenza virus strain. The protein evolution literature*

is full of examples of epistatic interactions that enable multiple mutations to access phenotypes not accessible by single mutations. Such epistasis is relevant to HA's evolution. For instance, work by Das et al suggests that the sequential accumulation of mutations can shift the spectrum of available antibody-escape mutations. Wu et al have used deep mutational scanning to directly demonstrate that rampant epistasis enables HA's receptor-binding pocket to accommodate combinations of individually deleterious mutations, some of which affect sensitivity to antibodies. Therefore, our work does not imply any absolute limits on the possibilities for antibody escape when evolution is given sufficient time to explore combinations of mutations. However, single mutations are the most accessible form of genetic variation, and much of influenza virus's natural antigenic drift involves individual mutations that reduce sensitivity to immunodominant antibody specificities. Quantifying the antigenic effects of all such mutations therefore provides a relevant measure of ease of viral antibody escape.

2) The virus used, WSN, is highly artificial and does not reflect actual human influenza viruses. It is trypsin independent (human viruses are not) and it is actually neurotropic (which is highly unusual). Also, it is a mouse adapted virus. While this is not necessarily a problem in the context of the presented data this caveat should be extensively discussed and avoided in the future.

The reviewer is correct that A/WSN/1933 is a lab-adapted virus. As the reviewer also notes, this is not actually a problem for our work since we are examining antibody neutralization rather than pathogenesis, fitness, or transmission. Nonetheless, the reviewer is correct that it is important to clearly mention this caveat. We have added the following text:

We performed our experiments using the lab-adapted A/WSN/1933 (H1N1) strain of influenza. This strain is derived from an early seasonal H1N1 that was extensively passaged in the lab, where it adapted to become neurotropic and trypsin independent. But despite these unusual properties, the virus is neutralized by most broad antibodies that target other H1 viruses, including those used in this study (Figure 3).

3) Page 2, beginning of second last paragraph: It is unclear if such a vaccine would really enhance the pressure dramatically. This would depend on vaccine coverage and potency. Currently, the drift of influenza viruses is driven by immunity induced by natural infection, not by vaccines (global vaccination rates are extremely low, even in high income countries in Europe the coverage is low).

This is a good point. We have changed "will be under strong antigenic selection" to "could come under stronger antigenic selection."

4) Data deposition at GitHub: How long will data/software stored on this platform be available? If versions are updated, is this tracked/recorded?

GitHub is the leading platform for archiving and tracking computer code. It is used widely in computational biology, and also by Google and Microsoft, among others. For instance, see: `https://www.wired.com/2015/03/github-conquered-google-microsoft-everyone-else/`

GitHub uses version tracking to record all updates to computer code, and make it possible to access any prior version of the software. For instance, while `https://github.com/jbloomlab/dms_tools2` provides the latest version of the code, the version used in this study (version 2.2.1)

will always be available `https://github.com/jbloomlab/dms_tools2/tree/2.2.1`, and it is easy to examine modifications made to this version in any new updates.

While it is impossible to guarantee that GitHub will exist forever, the same could be said for *Nature Press* (which could go out of business) or PubMed (which could be de-funded by the NIH or the Trump administration). The best that we can say is that GitHub is considered the most stable and reliable way to archive and track software, as evidenced by its use not only in computational biology but also by major corporations such as Google and Microsoft.

In addition, all the Jupyter notebooks and major results are provided in supplementary files, which will be available on the *Nature Press* website.

5) Figure 6: It might be a good idea to quantify the IC50 shifts and - since the neutralization assay was hopefully done at least in triplicates - check what is significant and what not. E.g. the effect of the -8 mutation for FI6 doesn't look real.

This is a good comment that was also made by several other reviewers. We therefore now show full data for triplicate measurements with significance testing.

Specifically, all neutralization assays with FI6v3 were performed in triplicate on the same day to eliminate batch effects, with each replicate involving independent serial dilution of the antibody in a separate column of a 96-well plate (this is now explained in the Methods). In the original manuscript, we only showed the mean and standard deviation in Figure 6A. We have added Figure S9, which shows all the replicates. In addition, as explained in Figure S9, we fit IC50 values to each replicate, and used the three IC50's to test the null hypothesis that each mutant has an IC50 indistinguishable from wildtype.

The statistical testing in Figure S9B supports the conclusions in the text. The four mutations that are *not* expected to differ from wildtype with respect to FI6v3 neutralization (K280A, V135T, P80D, and M17L(HA2)) do *not* have IC50s that are significantly different than wildtype. The four mutations expected to have the largest increases in neutralization resistance (K280S, G47R(HA2), K280T, and N291S) are all significantly more resistant to FI6v3 than wildtype, with corrected P-values ranging from 0.00072 to 0.023. The mutation found by our mutational antigenic profiling that elicited the most skepticism both from us and the reviewers is K(-8)T. However, the statistical testing suggests that it really is slightly but significantly more resistant to FI6v3, with a corrected P-value of 0.04. However the effect is very small, and in the revised text we clearly emphasize this fact and advise caution in over-interpreting the result:

> *The most unexpected mutations identified in our mutational antigenic profiling were at site -8 in the signal peptide. We tested one of these mutations, K(-8)T, and it did lead to a very slight increase in neutralization resistance (Figure 6A, Figure S9)— although despite the significance testing in Figure S9, we remain circumspect about the magnitude of this effect relative to the noise in our neutralization assays.*

These revisions address this important comment made by several reviewers.

6) It would be important to quantify differences in mAb binding to the different escape mutants shown in Figure 6, e.g. by SPR/BLI. Right now it is unclear which mutant produces a complete loss of binding, if affinity changes result in the loss of neutralization etc. This needs to be explored.

The reviewer is correct that we do not include biophysical measurements of why the mutants have increased neutralization resistance. However, biophysically characterizeing antibody binding is beyond the scope of our study. We do not make any claims about the mechanisms of the mutations—we simply say that they reduce neutralization. This is in line with many other studies in the field. While biophysically oriented studies certainly include such measurements, there are numerous papers on viral evolution and antibody breadth that characterize neutralization as the phenotype of interest. Our study falls into this latter class.

7) Page 15/first page of discussion: The authors did not find a single-mutant virus with full loss of neutralization by anti-stalk mAbs. It has been noted that it is relatively difficult to make escape mutants against anti-stalk mAbs and the process usually involves serial passaging of virus in the presence of sub-neutralizing mAb concentrations - but usually full escape mutants are then recovered. In most cases these escape mutants have a) more than one mutation and b) have no completely loss of binding (but complete loss of neutralization). But there are also exceptions to this rule. Two papers have so far described escape from stalk-based mAbs in detail (`https://www.ncbi.nlm.nih.gov/pubmed/27351973` and `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4362269/`). It is strongly suggested that the authors consult these papers and include their findings in the discussion.

Thank you for pointing out these important references, which we now discuss. Two more references were pointed out by Reviewer #3. Interestingly, all these studies selected their escape mutants in group 2 HAs rather than a group 1 HA like the one in our study. Anecdotally, it seems that there are more described broad anti-stalk antibodies against group 1 than group 2 HAs. We obviously have no way to quantify this rigorously, but in the added text that mentions the studies suggested by the reviewer, we also point out the difference in HA groups:

> Interestingly, most previous studies (Chai, 2016; Ekiert, 2011; Friesen, 2014; Dunand, 2016) that have reported selecting single mutations with large effects ($\gg$10-fold) on neutralization by anti-stalk antibodies have used group 2 (e.g., H3 or H7) HAs rather than group 1 HAs like the one used in our work.

8) It would be good to have a negative control experiment with a non-binding human or mouse mAb. This is missing but should actually be the "real" negative control.

We see the rationale behind this suggestion, but do we not believe it is essential for the current work. Our claim is that we specifically identify mutations that increase resistance to each of the individual antibodies. With the three broad antibodies tested in this paper (FI6v3, C179, S139/1) and the four narrow antibodies tested by Doud et al (2017), we have examined seven different antibodies. The only overlap in the identified resistance mutations comes when the antibodies target similar epitopes (e.g., FI6v3 and C179)—there is no overlap in the mutations identified by antibodies targeting different epitopes. This fact indicates that we are identifying bona fide resistance mutations, rather than mutations with some general effect on all IgGs.

Figure 6 provides further data indicating that we find antibody-specific resistance mutations rather than mutations generally selected by any IgG. For instance, the mutations that increase resistance to H17-L19 (V135T) or H17-L7 (P80D) do *not* affect neutralization by FI6v3, showing they are specific to the antibodies that selected them. Finally, note that there are numerous papers where a monoclonal antibody is tested against a range of viral variants to compare IC50s without including data showing that all viral variants are unaffected by an irrelevant antibody

(e.g., Table 1 of *Nature*, 489:526-532; Figure 3 of *Nature Microbiology*,; 2:16199).

Reviewer #2 (Remarks to the Author):

The authors use a previously described method of mutational antigenic profiling, a very clever and elegant approach exploited in previous papers in PLOS Pathogens and CHM. This approach involves generating viral libraries containing all mutations in the protein of interest, selecting these viruses with or without antibody, and deep-sequencing to determine the relative frequencies of each mutation. The full methodology (except for summary statistics described in Equations 10 and 11), as well as all the results for the narrow antibodies, were described in their previous paper (Doud et al., 2017, PLOS Pathogens). Figure 7 seems to be imported from another paper (Doud et al., 2016, Viruses).

All the results for the broadly neutralizing antibodies are new to this study. The data for the narrow antibodies in Figure 6B,C is from Doud et al (2017), and the mutational tolerance data in Figure 7 is from Doud and Bloom (2016). This fact is clearly stated; for instance, the figure legends say "Shown are neutralization curves for representative escape mutants from H17-L19 and H17-L7 *taken from Doud et al (2017)*" and "The logo plots at bottom show the preference of each HA site of each amino acid under selection for viral replication *as measured by Doud and Bloom (2016)*." There is nothing wrong with including previous data for comparative purposes. For instance, papers on crystal structures often compare the results to prior structures on related molecules. In fact, a major strength of our approach is that it puts the results into units that can be meaningfully compared across studies.

The experimental methodology is the same as in Doud et al (2017), although the entire analysis framework (all equations, not just Equations 10 and 11) are new. The strength of this framework is that it enables the results to be compared across antibodies, which was not possible with the approach in Doud et al (2017).

In this work, they compare three broad and three narrow antibodies. The broad: FI6v3, binds stalk, neutralizes both group 1 and 2 HAs; C179 binds stalk, neutralizes only some group 1 HAs; S139/1, binds receptor-binding pocket, neutralizes both group 1 and group 2 HAs. The authors compare these in a re-analysis of data from narrow, strain-specific antibodies: H17-L19, H17-L10, and H17-L7, bind the Ca2, Ca1, and Cb antigenic regions on HA's globular head.

Using these three broad antibodies and WSN/33 Flu, they selected for escape mutations, and compared them to previous narrow antibody data. They find that escape from narrow or RBS-targeting Ab is relatively easy by single point mutation; while those targeting the stalk are difficult to escape by single point mutation. So breadth of antibody activity per se, is not a predictor of likelihood of escape.

They go on to show that the selected mutations are nearly all in or close to the antibody-binding footprint. While this may not be surprising, it is a nice set of data. Another nice aspect, is that the authors show that numerous different amino acid mutations at each site confer neutralization resistance - most studies tend to identify just one or two, rather than all of the possible aa.

They further show that the mutational antigenic profiling is highly predictive of the results of the neutralization assays, even for small-effect mutations, as was previously shown for large-effect mutations from narrow antibody escape.

The foregoing paragraphs are an accurate summary. We think it is valuable to know that the mutational antigenic profiling also works well for small-effect mutations.

They compare the sites where mutations arise from broad antibody, to previous mutational scanning data that reveals whether these regions are tolerant or not to mutation. They find that this is only partially true, but cannot account for all lack of escape mutations; however, they do not uncover what the other reasons are.

Some questions or comments that come to mind:

1. If the rationale is to ask whether current circulating strains could escape immunity based on broad neutralising antibody elicited by vaccination, could the authors not use circulating strains for their experiments, or at least more lab strains that only WSN?

We agree that it would be interesting to examine additional strains. However, prior to our work, no study had completely mapped the effect of all single mutations to any viral strain with respect to neutralization by any broad antibodies—so doing this with one strain is a substantial step forward. As mentioned in the response to question 2 by Reviewer #1, we now more clearly explain that WSN is a lab-adapted strain.

2. What difference in terms of viral fitness is there between a mutation present in large proportion (for the narrow antibodies) vs present at low frequency (for the stalk-targeting antibodies). Figure 7 partly answers this question; however, it would be interesting to confirm the fitness of those mutants using more conventional growth curves, binding assays, etc.

The neutralization assays in Figure 6 show that mutations for the stalk-targeting antibodies only have weak effects on neutralization, while Figure 7 shows that many of them don't have much effect on viral growth. In addition, we were able to generate all the stalk-antibody escape mutants for the neutralization assays, so they are not too badly attenuated in cell culture. However, we avoid making strong statements about the effects of mutations on viral growth beyond the data in Figure 7 because it is beyond the scope of our study to do growth curves for such a large number of mutations.

3. Are all changes in amino-acid obtained after one nucleotide change? Are the authors able to detect multiple nucleotide or amino acid changes?

Our study tested all amino-acid changes, including those that involved multiple-nucleotide changes. The reason is that our libraries contained all codon mutations (e.g., the 63 non-wildtype codons at each site) rather than just all single nucleotide mutations. We have added text that clarifies this point:

> Importantly, as described previously (Doud and Bloom 2016), these virus libraries were generated by mutagenizing HA at the codon level rather than at the nucleotide level. Performing codon mutagenesis is important, because single-nucleotide mutations access only about a third of the possible amino-acid mutations from a given codon, whereas codon mutations access all possible amino-acid mutations.

4. The major caveat of this study is that it only takes into account one mutation. Authors could check a combination of the mutations they obtain with their experiments (especially the ones escaping the stalk-targeting antibodies).

We agree that this is an important caveat. See our response to question 1 by Reviewer #1.

5. Figure 6 confirms some of the findings using neutralisation assays. As the authors state, it is unclear whether some of the observations actually exceed the noise of the assay. In particular, K(-8)T looks as far from WT as does K280A. It could be that (1) all these differences are due to noise or that (2) there is actually a biological difference between WT and K280A. Whether (1) occurs is unclear: "three biological replicates" should perhaps be more detailed in the Materials and Methods section as to what this means: are they three rows of the same plate performed on the same day, on different days, etc.? According to the authors' previous paper (Doud et al., 2017, PLOS Pathogens), they performed "three replicate dilution columns". The neutralisation curves seem therefore to have been performed at different days (besides, the WT curves are the same or similar to that of Figure 3, and the neutralisation curves from the narrow antibodies were imported from the authors' previous paper). This can induce a batch effect, which may not be linear, thus the normalisation performed (with no antibody and background noise) may not regress this out. From the standpoint of experimental design, these experiments should be repeated, with all viruses for a given antibody performed in the same batch; the triplicates will then give a more accurate measurement of noise. In general, all neutralisation experiments should be repeated, so as to make sure the findings are reproducible. If this has been done, it was not clear from the figure legend, and could be added for clarity. To address (2), authors should provide details of $F_{r,a}$ for K280A. If I understood correctly, because $F_{r,a}^{\text{excess}}$ (defined in Equation 2) hides mutations with lower fitness than average, it is not possible for the reader to assess whether this mutation is "neutral" or it has lower fitness with the antibody under study.

A similar point was made by several other reviewers. We have provided a detailed response in answer to question 5 by Reviewer #1.

Our revisions showithe neutralization curves for all replicates and performing statistical testing. All neutralization curves for antibody FI6v3 were performed on the same day. This is important because as the reviewer notes, there can be batch effects. If you look at Figure 6, you will see that the wildtype curves in panels A, B, and C are slightly different. But each panel is for a single antibody, and all curves with with each antibody were performed on the same day to avoid any confounding batch effects. Therefore, we have already done what the reviewer suggests, and the new Figure S9 now clearly explains this.

As described in the response to question 5 by Reviewer #1, the new statistical analysis finds that the effect for K(-8)T is significant, whereas that for K280A is not. The reviewer also suggests that it might be interesting to look at the $F_{r,a}$ value for K280A, since it might have reduced neutralization resistance. Since the statistical analysis in the new Figure S9 shows that the difference between K280A and wildtype is *not* statistically significant, we do not comment on possible increased neutralization sensitivity or the $F_{r,a}$ for this mutation in the paper. However, we have added a new supplementary file (File S6) that gives all the $F_{r,a}$ values. The value of $F_{r,a}$ for this mutation is 0.00382, which is slightly below the library average of 0.00955.

6. Could the authors detail, in the legend, how the curves of Figure 6 were fitted?

For the quantitative analyses of IC50's, we fit four-parameter logistic curves, although we fixed the top value to one. This is explained in the legend to the new Figure S9, which reports the IC50's and shows the curves for individual replicates. The computer code used to fit the curves is at `https://jbloomlab.github.io/dms_tools2/dms_tools2.neutcurve.html`.

Reviewer #3 (Remarks to the Author):

This paper identifies escape mutations against anti-influenza HA antibodies and quantifies their effects. The authors employed a high-throughput approach known as mutational antigenic profiling, which was developed in their previous study (1). The authors made a minor technical improvement in this study to enable a fairer comparison across profiling results from different antibodies. Two broadly neutralizing antibodies that target the HA stem (mouse C179 and human FI6v3), and one that targets the HA receptor-binding site (S139/1) were examined in this study. Additionally, the authors reanalyzed the results from their previous study (1) on three strain-specific antibodies (H17-L7, H17-L10, and H17-L19). The authors observed strong escape mutations against the strain-specific antibodies and broadly neutralizing antibody S139/1. In contrast, only weak escape mutations were identified against the two stem antibodies. The authors conclude that "breadth is an imperfect indicator of the potential for viral escape". The authors also conclude that broadly neutralizing antibodies targeting the H1 hemagglutinin stalk are "quantifiably harder to escape than the other antibodies tested here". The last two conclusions are not unexpected and are not particularly novel insights.

The study is, however, performed to an extremely high technical standard. Nevertheless, the impact is limited. The improvement in methodology over their previous study (1) is minor. Also, it is known that strong escape mutations against S139/1, one of the broadly neutralizing antibodies tested in this study (although not nearly as broad as FI6v3), can be generated rapidly through passaging the virus in the presence of the antibody (2). In fact, strong escape mutations against some of the stalk antibodies that were not tested in this study, such as CR8020 (3), CR8043 (4), CR6261 (5) and 39.29 (6), have also been generated through single point mutations. Overall, the impact and novelty of the study is not sufficient for publication in Nature Microbiology.

We are disappointed that the reviewer does not view the work of sufficient impact for *Nature Microbiology*, but appreciate the statement that the study is "performed to an extremely high technical standard." We feel that direct quantitative comparison of the escape from narrow and broad antibodies is valuable, since other comparisons have been more anecdotal. We also note that at least two of the reviewers found the work of high impact.

Thank you for the references. Interestingly, (3), (4), and (6) all describe selecting escape from *group 2* HAs (e.g., H3 or H7), not a group 1 HA like the one we used. We now cite these references and mention this point as described in the response to question 7 by Reviewer #1. It is true that (5) does mention an escape mutation from a group 1 HA from CR6261, but it never actually shows any neutralization data for that mutation, so we are unable to comment on it.

Specific comments:

1. Strong escape mutations to C179 have been identified in the original paper that discovered the antibody (7). But it seems like those mutations were not identified here. Can the authors provide some explanations for such a discrepancy?

The original paper on C179 selected a strong escape mutant in an H1 background. That virus had three mutations, and although the paper argues that only one is important, it never shows that the others are not required as secondary changes. We cite this paper, and use it to inform

the discussion of multiple mutants described in response to question 1 from Reviewer #1.

2. How is the curve fitting performed for the neutralization assay results? I do not seem to see any description for the curve-fitting method.

This point was also mentioned by other reviewers. See our response to question 5 of Reviewer #1 and questions 5 and 6 of Reviewer #2.

References

1. Doud MB, Hensley SE, Bloom JD. Complete mapping of viral escape from neutralizing antibodies. PLoS Pathog. 2017;13(3):e1006271.

2. Yoshida R, Igarashi M, Ozaki H, Kishida N, Tomabechi D, Kida H, et al. Cross-protective potential of a novel monoclonal antibody directed against antigenic site B of the hemagglutinin of influenza A viruses. PLoS Pathog. 2009;5(3):e1000350.

3. Ekiert DC, Friesen RH, Bhabha G, Kwaks T, Jongeneelen M, Yu W, et al. A highly conserved neutralizing epitope on group 2 influenza A viruses. Science. 2011;333(6044):843-50.

4. Friesen RH, Lee PS, Stoop EJ, Hoffman RM, Ekiert DC, Bhabha G, et al. A common solution to group 2 influenza virus neutralization. Proc Natl Acad Sci USA. 2014;111(1):445-50.

5. Throsby M, van den Brink E, Jongeneelen M, Poon LL, Alard P, Cornelissen L, et al. Heterosubtypic neutralizing monoclonal antibodies cross-protective against H5N1 and H1N1 recovered from human IgM+ memory B cells. PLoS One. 2008;3(12):e3942.

6. Chai N, Swem LR, Reichelt M, Chen-Harris H, Luis E, Park S, et al. Two escape mechanisms of influenza A virus to a broadly neutralizing stalk-binding antibody. PLoS Pathog. 2016;12(6):e1005702.

7. Okuno Y, Isegawa Y, Sasao F, Ueda S. A common neutralizing epitope conserved between the hemagglutinins of influenza A virus H1 and H2 strains. J Virol. 1993;67(5):2552-8.

Thanks for these references, they are now all cited.

Reviewer #4 (Remarks to the Author):

This paper tackles a really nice question: Are some antibodies broad because they target sites that face little immune pressure, or are they broad because the epitopes they target cannot easily mutate to escape them? Put differently, are conserved epitopes that are targeted by some antibodies conserved because they've not had to change, or are they conserved because no beneficial changes are accessible? This question is obviously relevant for "universal" vaccine development, but it's also important for understanding strain structure and evolution more generally.

To investigate this question, the authors analyzed the ability of a H1 flu strain to escape six well characterized antibodies via single mutations. They showed that for most antibodies, including a broadly neutralizing antibody, single mutations allowed escape from antibody detection. In the other two antibodies, no mutations really rescued the virus. By comparing the mutations conferring antibody escape to those that are generally tolerated during replication, the authors further

showed that mutational tolerance does not fully explain failure to escape: despite substitutions, some antibodies are harder to escape than others.

These are important observations. My only real criticism of the work is that I found the description of mutational effects confusing as early as the abstract. Does the measured frequency of antibody escape mutants reflect only antibody selection, or does it also include selection for replicative capacity? The distinction between antibody selection and general fitness is not made until p. 13, but I think readers (especially those familiar with other work from the lab) might benefit if these distinctions were made in the introduction and review of the methods. My take is that the fraction surviving antibody escape includes both, in that non-viable mutants were rare in the initial library (again implied on p. 13), but this could be clarified. It seems the contribution of each might be quantifiable, or at the relative mutational tolerances of the different epitopes compared.

Thanks for the kind comments on the importance of the paper and the questions it addresses.

The point about clearly distinguishing between how mutations affect antigenicity and viral growth is well taken. We have added text at several places to clarify this. For instance, as soon as we define our measures of antigenic effects, we say:

> *Importantly, Equations 1 and 2 correct for effects on viral growth due to normalization by the mock-selected control, and so measure only antigenicity and not viral growth provided that the virus at least grows well enough to be present in the library.*

Minor points:

1. How do the concentrations for selecting escape mutants in vitro compare to likely concentrations in vivo? Concentrations may be too high or low in vivo to promote adaptation.

We are unable to answer this question, because the concentrations of any given antibody specificity can vary widely among people. In addition, it is difficult to relate *in vitro* neutralization to *in vivo* activity due to Fc-mediated effector functions. This is a limitation of virtually all work on antibody neutralization. We mention this caveat in the Discussion:

> *When interpreting the magnitude of these effects, it is important to note that our experiments only measure how mutations affect neutralization, and not how they affect Fc-mediated functions that are responsible for much of the in vivo protection afforded by anti-stalk antibodies.*

2. Outside these experiments, how much data is there on the antibodies' neutralization capacities within HA subtypes? i.e., how broad or narrow do we really know them to be? Obviously such data might be complementary.

Unfortunately, most of the antibodies have been characterized against only a small number of the many strains within HA subtypes. We summarize the subtype-level information in Figure 2B, and cite all studies that look at breadth within subtypes (references 24, 25, 29, 45, 46). However, because different studies use different strains, the data in these studies do not enable useful comparisons across antibodies. Our study only looks at one strain—but at least it is complete for the mutations to this HA, and so can be compared across antibodies.

3. Additional suggested references:

"the viruses's rapid antigenic evolution erodes the effectiveness of this immunity to that strain's descendants within four to seven years..." – Kucharski et al. PLOS Biol 2015 (although numbers might need revising)

This is a good reference, and we now cite it. The distribution of infection frequencies they derive from epidemiology is similar to the 4-7 years we had previously cited from a human challenge study. However, we have updated the text to "$\sim$5 years" to better reflect the ambiguity.

on immune status driving the rate of antigenic evolution – Wen et al. Proc B 2016

This is a good reference, and we now cite it at the relevant point.

4. In "...we expect a larger fraction of a library to survive an antibody with many escape mutations..." (and similar phrases), the "many escape mutations" read as though they modify "antibody."

This is a good point, we have clarified the wording in these phrases.