Peer Review File

Manuscript Title: Human germinal centres engage memory and naïve B cells after influenza vaccination

Editorial Notes:

Redactions – unpublished data

Parts of this Peer Review File have been redacted as indicated to maintain the confidentiality of unpublished data.

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

In this well designed and detailed study, the Authors investigate the human B cell response to the seasonal influenza vaccine. They analyse the response in peripheral blood and in fine needle biopsied of draining axillary lymph nodes, a procedure that allows access to germinal centres. As previously reported, they observe a transient wave of circulating antibody-secreting plasmablasts peaking one week after vaccination and producing antibodies specific for previously encountered hemagglutinins, consistent with the "original antigenic sin". Importantly, in 3 out of 8 donors tested, they document a robust GC response that persisted for up to nine weeks in draining lymph nodes. Interestingly, the GC response comprises both highly mutated B cell clones with antibody sequences that overlap with the plasmablast response and with broadly reactive specificities, as well as B cell clones unique to the GC that display significantly lower SHM frequency and are enriched for strain-specific mAbs. The Authors also use electron microscopy-based epitope mapping of polyclonal Fab fragments to show that the strain-specific mAbs generated in the GC response selectively recognize protective epitopes that were not targeted by the early plasmablast response. Furthermore, they show in some cases that the footprint of recombinant Fabs from GC B cells overlaps with that found in plasma, whereas in other cases the Fabs targeted unique epitopes.

Collectively, these results clearly demonstrate that seasonal influenza virus vaccination can elicit a GC response in humans and that a heterogenous population of memory and naïve B cells is recruited to the GC reaction. The work, which is is of the highest technical quality, addresses an important biological and medical question. The results are convincing and are discussed in a concise and objective way.

For these reasons I recommend acceptance of this paper with the highest level of enthusiasm.

Referee #2 (Remarks to the Author):

Turner and colleagues present data on a cohort of volunteers which they immunized with influenza vaccine and then performed fine needle aspirates on to sample GC B cells from the ongoing vaccine response. Although the cohort is very small, the repeated sampling of responding LN makes this a very unique set of samples, which is a major strength of this study. The authors' first major point is that at least some of their subjects clearly formed GCs in response to flu immunization, which answers qualitatively (if not quantitatively given the small number of patients) a question that is still strongly debated in the field. The second major point that the authors make is that the GCs they sample contain a mixture of both memory and naïve B cell clones, and original antigenic sin cannot completely exclude the generation of new antibody responses in GCs. While this is an important point given the history of

work on original antigenic sin in flu infection, I would like to see a more thorough analysis of several aspects of the data, as detailed below. But, overall, this is a unique study that raises many important points that could have an impact on flu vaccine development.

Major points:

1. One issue that could confound this analysis is that this study focuses only on B cells that are "antigen-specific" as defined by HA-baited FACS or mAb ELISA. Multiple reports from mice (e.g. PMID 26948373, 32066977, 22932267) show that most B cells recruited to GCs bind to antigen too weakly to pass these FACS/ELISA thresholds, even though they are almost certainly antigen-specific. This suggests that the present study is likely investigating only a subset of the full HA-induced GC response.

Indeed, even though HA is reportedly the major immunogen of inactivated influenza vaccines, it appears that antigen-binding cells represent only a minor fraction of all GC B cells in most samples (comparing Fig 2i and Ext Fig 2h; as an aside, it would be good if the % of GC B cells positive for antigen were plotted explicitly somewhere). This seems to be the case also according to the numbers reported for mAb reactivity, although it is not clear what exactly those numbers represent. (When the authors say 41/135 mAbs from patient 4 were "clonally distinct and QIV-specific" does it mean that there were 41 different clones? What about the remaining antibodies, were they not specific or not clonally distinct? How were the 135 B cells chosen, were they sorted based on HA binding by FACS or randomly from total GC B cells?) Without this information it is impossible to tell what fraction of the secondary GC response the authors are actually looking at. This is important because memory cells should have higher affinity for antigen, which would make them more likely to fall in the "antigen-specific" group. So restricting clonal analysis to antigen-binding B cells could substantially inflate the overlap between PB and GC, overestimating the fraction of GC B cells coming from memory.

A simple fix would be to recalculate clonal overlaps using all GC B cells not only antigen-specific ones, if the authors have this data available. If not, the authors' analysis is still valid, but it should be made clear in their conclusions that their percentages of GC B cells coming from naïve or memory B cells are calculated for higher-affinity antigen-binding cells only and may not quantitatively represent the situation of the full GC.

2. One of the major advantages of the FNA approach is that the same LN (though probably not the same GC) can be sampled over time, allowing for kinetic analysis of clonal diversity and somatic mutations, among others. In this sense it is disappointing that the authors do not take full advantage of this serial sampling. For example, analysis of the numbers of mutations in figs. 3c and 3e are done on pools of cells from week 1 to week 9, which are unlikely to be homogeneous given that mutations accumulate over time. The cleanest analysis for naive-derived cells would be done at the first point when a clear GC population is present (week 1 or 2, depending on the patient), before these cells have chance to accumulate multiple mutations. One example: line 199 states that "SHM frequency was comparable between total GC and PB populations (Fig. 3c, e)" but GC B cells appear to be pooled from all samples from 1 to 9 weeks, when it is expected that mutations will accumulate over time, whereas PB are a single sample from week 1. There is also no analysis of how the clonal overlap between GC samples and with PB progresses over time. Most antigen-binding mAbs seem to come from early time points (weeks 1 and 2), before the GCs have fully developed, and when GC B cells and activated B cells are difficult to separate. Are the clones in early GCs the same as the ones in late GCs? Is the overlap with PB still as evident when GCs are fully formed?

Minor points:

1. In bulk Ig sequencing, was there a threshold number of reads or UMIs used for defining a true sequence? How was contamination between samples and from the environment dealt with?

2. Very little attention is paid to SHM patterns, which is a pity given the wealth of data available. Building dendrograms from sequences obtained over time would be a simple way to extract further

information from these samples.

3. Lines 64-66: This is a bit misleading. My understanding is that high mutational burdens are not a requirement for broad influenza neutralizing like they are for HIV, and that the extent of mutations in HIV bnabs is much higher.

4. The kinetics of the GC in subject #5 show delayed GC formation and a consistent increase in GC frequency over 9 weeks, which is unusual but not discussed. Do the authors have any insight into this?

5. Line 146, "Given the role of GCs in diversifying B cell responses…" The logic here is a bit flawed. The diversification induced by GCs is unlikely to change the B cell epitopes in a manner that would be distinguishable by EM. What EMPEM seems to be doing is revealing the antibodies that emerge from new incoming clones less or not represented in the serum prior to immunization.

Referee #3 (Remarks to the Author):

Summary:
Turner et al. have conducted an impressive study of germinal center B cell responses to influenza vaccination in humans. They are able to identify vaccine-specific GC B cells from FNAs of multiple individuals over several timepoints. This analysis is key to answering the central question, does vaccination in a previously exposed adult induce a germinal center response incorporating only previously amplified and mutated memory B cells or also include some novel clones? In 3 of there 8 individuals, they find vaccine-induced GC responses that persist for up to 9 weeks. They therefore explore clonal relationships between PBs in PBMCs and GC B cells in FNA in one of the individuals with GCs present across all timepoints. They can find clones that are shared between the two populations as well as unique clones and suggest that novel strain-specific GC clones with lower SHM participate in these GCs. This is a key finding in the context of much work in mice finding that the majority of secondary GCs are seeded by less mutated and lower frequency clones and the work in humans (Andrews et al) that the majority of flu vaccine responses use memory clones amplified in previous exposures. While this work is critical and interesting the data needs to be presented more clearly and the discussion needs to more extensively consider the implications on the general principles of secondary GCs.

Major:
1) Much more context is needed for this work. This study has an exciting opportunity to expand upon the findings in mice that 'naïve' cells/unmutated IgD memory dominate secondary GCs (work from Taylor, Pape, Shlomchik, Victora etc) and studies of human PBMCs after flu vaccination that memory B cells dominate the PB response (Andrews, Wilson etc). This data suggests evidence for both being true and this should be emphasized and discussed. The authors should also include a discussion of whether previous studies of circulating influenza specific memory cells have demonstrated the appearance of new, strain specific clones at a memory timepoint after immunization as this would be hypothesized by the current study and would help to extend its implications.
2) The authors clone QIV-specific BCRs from PBMCs and GCs at several time points to address whether GCs contain unique BCR clones. This essential to the paper, but the presentation in tiny TSNE plots in Fig. 3 is hard to interpret. It might be easier to see the phenotype of each clone in a tree or tracking phenotype of each clone over time and then quantifying the trajectory (ie PBMC memory at d0, PB at week 1, GC at week 2, memory at week 9). This point would also be strengthened by a discussion or demonstration of how the authors control for sampling bias across timepoints (ie not finding a clone in the sample doesn't mean it is not there). It is impressive and very useful that the authors can find the same clones across timepoints and this is a technical advance that should be emphasized and explained. This has critical implications for the future of this work following this experiment into a memory timepoint to see if the new clones in the GCs subsequently contribute to circulating memory.
3) Results are challenging to follow and need more clarity on key aspects of methodology throughout the paper. For a specific example, in reference to Fig.3b the authors say this data represents mAb and bulk

BCR seq, but it is unclear how this information is integrated and contributing to the data in the figure. Also, I am unaware of a precedent for clonal BCR sequences being pulled out of bulk B cells, and if these are just Vh genes, how are you collating the data with the mAb? This needs to be explained or referenced in the text.
4) Much of the analyses and interpretation depends upon GC B cells uniquely expressing BCL6. In mice, studies by Okada and Haberman have shown that BCL6 can also be expressed during T-B clustering before the GC. What does the total BCL6 by KI67 expression look like after vaccination? Is it possible that activated memory cells that are BCL6+ have been included in the GC pool in these analyses? Is it possible to look prior to formation of a GC as these studies and those in mice suggest that secondary/memory GCs do not form with faster kinetics than a primary response? How quickly do HA-specific cells express BCL6 after vaccination? It would help support the authors claims if there was more detailed flow analyses of HA-specific cells instead of pre-gated histograms shown in extended Figure 2I.

Minor:
1) HA+ GCB only detected in 3/8 people, would be valuable to see data for total #GCB collected per timepoint/person as well as antigen specific. Detection error or biology? Extended 2h suggests pt 7 and 8 also had detectable GC, but not HA and all people had responses by ab. Emphasize that got GCB in 6/7 by flow. In Fig2, why not show distribution of B cell types as analyzed in extended fig 2 (how many GCB, MBC etc over time? MFI is good characterization, now pair with quantification from Extended Table 4.
2) Fig 4: What were the circulating strains in the last 3 years that would suggest that only cross-reactive antibody could be memory?
3) Need to report the standard error of the mean to measure variance.
4) Fig 4 d: Were all unique binding antibodies tested in vivo independent of SHM? If not hard to say there is a trend in these ones being protective.
5) Isotype data in Ext. Fig 4 not mentioned in text anywhere, this is useful information.
6) Summary: Precision of language i.e. "proven", "critical" are not supported by the data. Instead use "supports", "contribute".

Author Rebuttals to Initial Comments:

We would like to thank the editor and reviewers for their time and careful consideration of our manuscript. The comments provided were very helpful and we feel that incorporating the recommended analyses and modifications has clarified our conclusions and substantially improved the manuscript. Additionally, the new analyses have allowed us to highlight previously unreported details about the maturation of clones participating in germinal center (GC) responses. Specifically, the revised manuscript includes these critical new analyses:

1. Kinetic analysis of clonal overlap between GC and PB compartments, and for participant 05, SHM analysis over time.

2. A phylogenetic analysis based on somatic hypermutation (SHM) showing the evolution of selected GC B cell clones over time after vaccination.

Below we have provided responses to reviewers' questions detailing the ways in which our revised manuscript addresses their concerns.

Reviewer #1:

In this well designed and detailed study, the Authors investigate the human B cell response to the seasonal influenza vaccine. They analyze the response in peripheral blood and in fine needle biopsied of draining axillary lymph nodes, a procedure that allows access to germinal centres. As previously reported, they observe a transient wave of circulating antibody-secreting plasmablasts peaking one week after vaccination and producing antibodies specific for previously encountered hemagglutinins, consistent with the "original antigenic sin". Importantly, in 3 out of 8 donors tested, they document a robust GC response that persisted for up to nine weeks in draining lymph nodes. Interestingly, the GC response comprises both highly mutated B cell clones with antibody sequences that overlap with the plasmablast response and with broadly reactive specificities, as

well as B cell clones unique to the GC that display significantly lower SHM frequency and are enriched for strain-specific mAbs. The Authors also use electron microscopy-based epitope mapping of polyclonal Fab fragments to show that the strain-specific mAbs generated in the GC response selectively recognize protective epitopes that were not targeted by the early plasmablast response. Furthermore, they show in some cases that the footprint of recombinant Fabs from GC B cells overlaps with that found in plasma, whereas in other cases the Fabs targeted unique epitopes.

Collectively, these results clearly demonstrate that seasonal influenza virus vaccination can elicit a GC response in humans and that a heterogenous population of memory and naïve B cells is recruited to the GC reaction. The work, which is of the highest technical quality, addresses an important biological and medical question. The results are convincing and are discussed in a concise and objective way.

For these reasons I recommend acceptance of this paper with the highest level of enthusiasm.

We would like to thank Reviewer 1 for nicely summarizing our findings and for the kind remarks.

Reviewer #2:

Turner and colleagues present data on a cohort of volunteers which they immunized with influenza vaccine and then performed fine needle aspirates on to sample GC B cells from the ongoing vaccine response. Although the cohort is very small, the repeated sampling of responding LN makes this a very unique set of samples, which is a major strength of this study. The authors' first major point is that at least some of their subjects clearly formed GCs in response to flu immunization, which answers qualitatively (if not quantitatively given the small number of patients)

a question that is still strongly debated in the field. The second major point that the authors make is that the GCs they sample contain a mixture of both memory and naïve B cell clones, and original antigenic sin cannot completely exclude the generation of new antibody responses in GCs. While this is an important point given the history of work on original antigenic sin in flu infection, I would like to see a more thorough analysis of several aspects of the data, as detailed below. But, overall, this is a unique study that raises many important points that could have an impact on flu vaccine development.

Reviewer 2 has made many important comments that we feel have greatly improved the quality of our work. Specific details are outlined below.

Major points:

1. One issue that could confound this analysis is that this study focuses only on B cells that are "antigen-specific" as defined by HA-baited FACS or mAb ELISA. Multiple reports from mice (e.g. PMID 26948373, 32066977, 22932267) show that most B cells recruited to GCs bind to antigen too weakly to pass these FACS/ELISA thresholds, even though they are almost certainly antigen-specific. This suggests that the present study is likely investigating only a subset of the full HA-induced GC response.

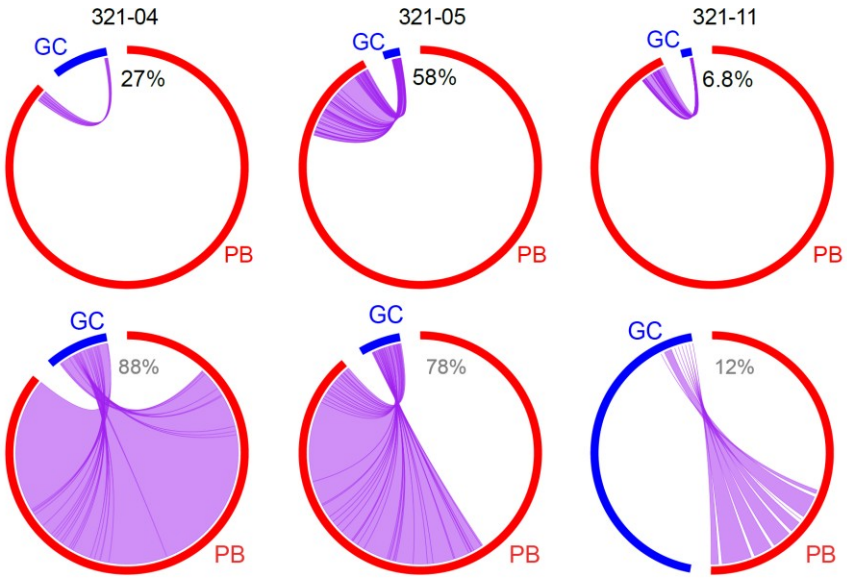
Indeed, even though HA is reportedly the major immunogen of inactivated influenza vaccines, it appears that antigen-binding cells represent only a minor fraction of all GC B cells in most samples (comparing Fig 2i and Ext Fig 2h; as an aside, it would be good if the % of GC B cells positive for antigen were plotted explicitly somewhere). This seems to be the case also according to the numbers reported for mAb reactivity, although it is not clear what exactly those numbers represent. (When the authors say 41/135 mAbs from patient 4 were "clonally distinct and QIV-specific" does it mean that there were 41 different clones? What about the remaining antibodies, were they not specific or not clonally distinct? How were the 135 B cells chosen, were they sorted based on HA binding by FACS or randomly from total GC B cells?) Without this information it is impossible to tell what fraction of the secondary GC response the authors are actually looking at. This is important because memory cells should have higher affinity for antigen, which would make them more likely to fall in the "antigen-specific" group. So restricting clonal analysis to antigen-binding B cells could substantially inflate the overlap between PB and GC, overestimating the fraction of GC B cells coming from memory.

A simple fix would be to recalculate clonal overlaps using all GC B cells not only antigen-specific ones, if the authors have this data available. If not, the authors' analysis is still valid, but it should be made clear in their conclusions that their percentages of GC B cells coming from naïve or memory B cells are calculated for higher-affinity antigen-binding cells only and may not quantitatively represent the situation of the full GC.

Selection process of the responding B cells to be sorted used for mAb generation: We sorted total GC B cells ($CD19^{+}$ $CD4^{-}$ IgD^{lo} $CD20^{hi}$ $CD38^{int}$) from the FNA specimens at multiple timepoints and total PBs ($CD19^{+}$ $CD4^{-}$ IgD^{lo} $CD20^{lo}$ $CD38^{+}$ $CD71^{+}$) from PBMC at the peak of the response, one week after vaccination. We did not sort GC B cells or PBs based on their binding to labeled HA. This point has been emphasized in the updated text (Line 167). After single cell sorting into 96-well plates, the IGHV and IGLV genes were PCR amplified and sequenced. All sequences were then annotated and categorized into different clonal pools based on the IGHV and IGHJ gene usage, the HCDR3 length and the percent homology among HCDR3 (Methods section "Clonal lineage inference", Line 641). We then selected a representative clone from each clonal pool for high-quality amplification and cloning of its IGHV and IGLV genes into expression vectors, and the corresponding mAb was expressed. The specificity of the expressed mAbs was then established by binding to the quadrivalent inactivated influenza vaccine (QIV) that the participants received.

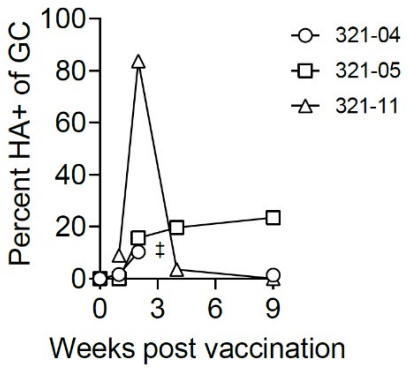
Affinity threshold and designation of vaccine-specificity: The affinity threshold is an important point and we thank the reviewer for drawing our attention to it. As discussed, the specificity of the mAbs we generated was verified by ELISA against QIV. While it is possible this threshold caused us to exclude some of the B cell clones encoding mAbs that weakly bind QIV, we have no objective and practical way of differentiating such clones from those specific for irrelevant antigens. This is important because we detected GC B cells in baseline FNA specimens in participants 04, 05, and 07 (Extended Data Fig. 2h). Given that we recruited participants who had not been vaccinated for at least 3 years, these GC B cells were certainly not QIV-specific, indicating that responses to irrelevant antigens were ongoing in the sampled lymph nodes and casting doubt on the validity of the assumption that all recovered GC B cells are QIV-specific regardless of whether they pass the QIV ELISA binding threshold. Nevertheless, we agree with the reviewer's concern that lower-affinity clones may be excluded from our analysis and have now explained this caveat in the updated manuscript (Line 311).

PB/GC clonal overlap: To illustrate the difference in GC/PB clonal overlap if we include total GC B cell (sorted on surface phenotype $CD19^{+}$ $CD4^{-}$ IgD^{lo} $CD20^{hi}$ $CD38^{int}$) clones versus only those verified by ELISA to be QIV-specific, we performed the analyses shown below (Reviewer Fig. 1). While the frequency of total GC sequences clonally related to PB sequences varied among participants 04, 05, and 11 (top panel), in all cases the percentage increased when restricted to clones that pass the QIV ELISA threshold (bottom panel), although to a much lesser extent for participants 05 and 11 compared to 04. One interpretation of these results is that the QIV ELISA threshold causes us to exclude some of the B cell clones encoding mAbs that weakly bind QIV. However, because we cannot practically distinguish between sub-threshold but still QIV-specific clones and those specific for irrelevant antigens, we report only those we can verify to be QIV+.



Reviewer Figure 1. Total and QIV+ GC and PB clonal overlap. Clonal overlap from total (top) or QIV-specific (bottom) mAbs and bulk repertoire analysis between PBs sorted from PBMC 1-week post-vaccination and GC B cells. Chord width corresponds to clonal population size. Percentages are of GC sequences pooled from all timepoints overlapping with PBs.

Percent of GC B cells that are HA+ by flow cytometry: We thank the reviewer for the suggestion. These data are shown below and have now been added to the updated manuscript as Extended Data Fig. 2i.



Reviewer Figure 2. Percent HA⁺ GC B cells.

2. One of the major advantages of the FNA approach is that the same LN (though probably not the same GC) can be sampled over time, allowing for kinetic analysis of clonal diversity and somatic mutations, among others. In this sense it is disappointing that the authors do not take full advantage of this serial sampling. For example, analysis of the numbers of mutations in figs. 3c and 3e are done on pools of cells from week 1 to week 9, which are unlikely to be homogeneous given that mutations accumulate over time.

The cleanest analysis for naive-derived cells would be done at the first point when a clear GC population is present (week 1 or 2, depending on the patient), before these cells have chance to accumulate multiple mutations. One example: line 199 states that "SHM frequency was comparable between total GC and PB populations (Fig. 3c, e)" but GC B cells appear to be pooled from all samples from 1 to 9 weeks, when it is expected that mutations will accumulate over time, whereas PB are a single sample from week 1.

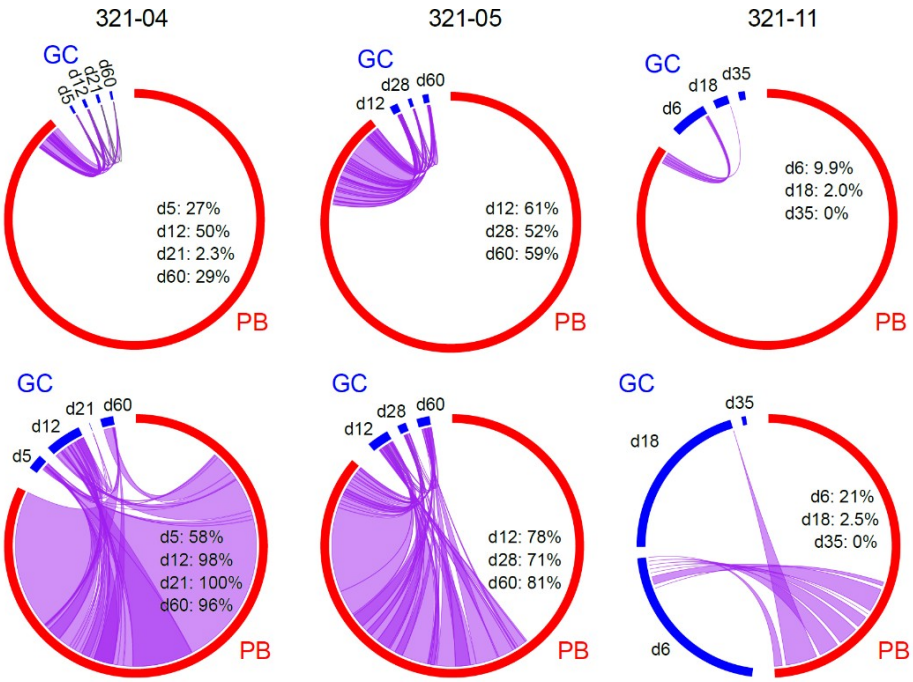
There is also no analysis of how the clonal overlap between GC samples and with PB progresses over time. Most antigen-binding mAbs seem to come from early time points (weeks 1 and 2), before the GCs have fully developed, and when GC B cells and activated B cells are difficult to separate. Are the clones in early GCs the same as the ones in late GCs? Is the overlap with PB still as evident when GCs are fully formed?

Kinetic breakdown of GC B cell SHM: This is a very important point that we have now addressed by including a kinetic breakdown of SHM using the BCR sequences retrieved from participant 05 in the updated Fig. 3e (also shown below in Reviewer Fig. 3a). Notably, the SHM frequency among clones recruited to the GC but not early PB response (blue) remains lower at all timepoints than that of clones found in both compartments (purple). This result is consistent with the conclusion drawn from our initial analysis (in which GC SHM data were derived from cells pooled from weeks 1 to 9) that clones recruited to the GC but not early PB response are significantly less mutated. We agree with the reviewer that it is more informative to include the kinetic breakdown of the data and have now included it in the main figure. Unfortunately, we cannot perform similar kinetic analyses for participants 04 and 11 (for whom scRNAseq analysis was not performed) using sequence data from mAb cloning, since only one representative of each clonal group is chosen for generation of sequences of sufficient quality for SHM analysis, and so we leave Fig. 3c unchanged.

SHM analysis of total vs. QIV specific clones: We have clarified the point that the accumulated SHM frequencies between total GC B cells pooled from all timepoints and total early PB populations (both regardless of QIV specificity) are similar. This is in contrast to the significantly reduced mutational load we observed among the QIV+ clones recruited to the GC but which did not participate in the early PB response, supporting the conclusion that this subset of QIV-specific clones derive from naïve or less mutated MBCs. We have clarified this point in the text (Lines 220-223) and updated Fig. 3e, f (also shown below in Reviewer Fig. 3a, b).

[Redacted]

Kinetic breakdown of PB/GC clonal overlap: The points raised by the reviewer are important, and we have now added a kinetic breakdown of the clonal overlap between the PB and GC compartments for participants 04, 05, and 11 as Extended Data Fig. 3d in the updated manuscript to address them (also shown below as Reviewer Fig. 4). Clonal overlap between the GC and PB compartments was observed across multiple timepoints for the three subjects analyzed, with no clear trends over time (overlap in participant 04 increased, 05 stayed fairly constant, and 11 decreased), although clonal coverage at some timepoints was sparse. Notably, very few clones were found in GCs at multiple timepoints, which may reflect artifacts of limited sampling, sampling different GCs within the same lymph node, or suggest that clonal turnover of GCs in this context is rapid. We have added these points to the text (Line 187).



Reviewer Figure 4. Kinetics of clonal overlap. Clonal overlap from total (top row) and QIV-specific (bottom row) mAbs and bulk repertoire analysis between PBs sorted from PBMC 1-week post-vaccination and GC B cells at the indicated timepoint. Chord width corresponds to clonal population size. Percentages are of GC sequences overlapping with PBs.

Minor points:

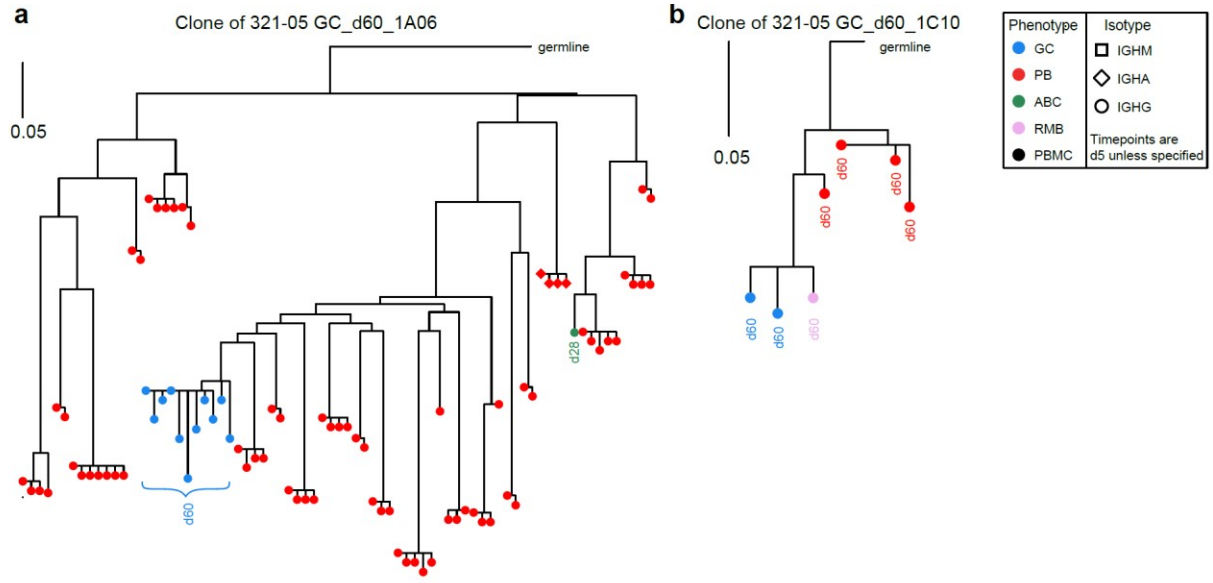
1. In bulk Ig sequencing, was there a threshold number of reads or UMIs used for defining a true sequence? How was contamination between samples and from the environment dealt with?

We clarified in the methods section (Line 587) that only sequences with at least two reads contributing to the UMI consensus sequence were used. An overlap matrix (Reviewer Fig. 5, below) of sequences among samples was constructed and found no indication of cross-sample contamination.

[Redacted]

2. Very little attention is paid to SHM patterns, which is a pity given the wealth of data available. Building dendrograms from sequences obtained over time would be a simple way to extract further information from these samples.

This is an excellent point. We have now included SHM-based dendrograms for multiple QIV-specific B cell clones with clone members found at multiple timepoints in the PBMC and LN compartments. In updated Fig. 4e (also shown below as Reviewer Fig. 6a), we show a dendrogram of the clonal family of mAb 1A06. 1A06 was derived from a GC B cell isolated from participant 5, day 60. As we show in Fig. 4a, it recognizes not only H1 HA from the vaccine strain (A/Michigan), but also representatives from previously circulating H1 HAs (dating back to 1934) and essentially all group 1 HAs (except for H2 HA). A large number of 1A06 clonal siblings were also detected among day 5 PBs with some exhibiting lower, equivalent or higher mutational load than 1A06. Overall, these data indicate that this clonal expansion was seeded by pre-existing memory B cells that recognize a highly conserved epitope in the H1 HA, with some rapidly differentiating into circulating PBs and others (e.g., 1A06) participating in the GC reaction. In contrast, in updated Fig. 4f (shown below as Reviewer Fig. 6b), we show a dendrogram of the clonal family of 1C10, a mAb that was also derived from a GC B cell isolated from participant 5 at day 60. 1C10 only recognized H1 HA from the vaccine strain (A/Michigan), indicating that it binds to a novel epitope. No clonal siblings of 1C10 were detected before day 60 outside the LN, suggesting a delayed engagement in the immune reaction and potentially a naïve B cell origin. Similar patterns were also observed for a pair of influenza B/HA-specific clones (Extended Data Fig. 5h, i). Taken together, these dendrograms support the overarching conclusion of our manuscript; influenza vaccination in humans can induce a GC reaction that engages pre-existing memory B cell as well as naïve clones.



Reviewer Figure 6. Clonal Dendrograms. Dendrograms of clonal families of d60 GC mAbs 1A06 (a) and 1C10 (b). The vertical branch length represents the expected number of substitutions per codon in V region genes, corresponding to the key in the upper left of each panel. Colored symbols represent sequences from cells isolated at day 5 unless otherwise specified, corresponding to the indicated B cell type and Ig isotype.

3. Lines 64-66: This is a bit misleading. My understanding is that high mutational burdens are not a requirement for broad influenza neutralizing like they are for HIV, and that the extent of mutations in HIV bnabs is much higher.

We agree with the reviewer; the text has been revised to reflect the important role of GC maturation with respect to breadth, but deemphasize the role of extensive SHM for broadly cross-reactive influenza mAbs (Lines 69-71).

4. The kinetics of the GC in subject #5 show delayed GC formation and a consistent increase in GC frequency over 9 weeks, which is unusual but not discussed. Do the authors have any insight into this?

We agree that this is an important observation and were surprised to see such a persistent antigen specific GC response after influenza vaccination. Studies in the lab have been initiated to examine the durability of that GC response and the functional implications for its persistence. We have added this point to the discussion in the updated manuscript (Lines 295-303).

5. Line 146, “Given the role of GCs in diversifying B cell responses…” The logic here is a bit flawed. The diversification induced by GCs is unlikely to change the B cell epitopes in a manner that would be distinguishable by EM. What EMPEM seems to be doing is revealing the antibodies that emerge from new incoming clones less or not represented in the serum prior to immunization.

We have clarified that diversification detectable by EMPEM would not be via SHM, but through recruitment of clones specific for novel epitopes to the GC and their eventual plasma cell differentiation (Line 164).

Reviewer #3:

Turner et al. have conducted an impressive study of germinal center B cell responses to influenza vaccination in humans. They are able to identify vaccine-specific GC B cells from FNAs of multiple individuals over several timepoints. This analysis is key to answering the central question, does vaccination in a previously exposed adult induce a germinal center response incorporating only previously amplified and mutated memory B cells or also include some novel clones? In 3 of there 8 individuals, they find vaccine-induced GC responses that persist for up to 9 weeks. They therefore explore clonal relationships between PBs in PBMCs and GC B cells in FNA in one of the individuals with GCs present across all timepoints. They can find clones that are shared between the two populations as well as unique clones and suggest that novel strain-specific GC clones with lower SHM participate in these GCs. This is a key finding in the context of much work in mice finding that the majority of secondary GCs are seeded by less mutated and lower frequency clones and the work in humans (Andrews et al) that the majority of flu vaccine responses use memory clones amplified in previous exposures. While this work is critical and interesting the data needs to be presented more clearly and the discussion needs to more extensively consider the implications on the general principles of secondary GCs.

Reviewer 3 has made excellent comments that have helped us present our findings more clearly, for which we are grateful. Specific details are outlined below.

Major:

1) Much more context is needed for this work. This study has an exciting opportunity to expand upon the findings in mice that 'naïve' cells/unmutated IgD memory dominate secondary GCs (work from Taylor, Pape, Shlomchik, Victora etc) and studies of human PBMCs after flu vaccination that memory B cells dominate the PB response (Andrews, Wilson etc). This data suggests evidence for both being true and this should be emphasized and discussed. The authors should also include a discussion of whether previous studies of circulating influenza specific memory cells have demonstrated the appearance of new, strain specific clones at a memory timepoint after immunization as this would be hypothesized by the current study and would help to extend its implications.

We thank the reviewer for this important suggestion. The introductory and concluding sections have been revised to more fully contextualize our results in the broader field.

2) The authors clone QIV-specific BCRs from PBMCs and GCs at several time points to address whether GCs contain unique BCR clones. This essential to the paper, but the presentation in tiny TSNE plots in Fig. 3 is hard to interpret. It might be easier to see the phenotype of each clone in a tree or tracking phenotype of each clone over time and then quantifying the trajectory (ie PBMC memory at d0, PB at week 1, GC at week 2, memory at week 9). This point would also be strengthened by a discussion or demonstration of how the authors control for sampling bias across timepoints (ie not finding a clone in the sample doesn't mean it is not there). It is impressive and very useful that the authors can find the same clones across timepoints and this is a technical advance that should be emphasized and explained. This has critical implications for the future of this work following this experiment into a memory timepoint to see if the new clones in the GCs subsequently contribute to circulating memory.

We agree with the reviewer's suggestion. This was also a point raised by Reviewer 2 that we addressed by including SHM-based dendrograms of multiple clonal families, each of which encompass phenotypically distinct members (See response to Reviewer 2, minor point #2, above and Extended Data Fig. 5h, i in the revised manuscript). Importantly, for the majority of the clonal families we were not able to detect a representative clone among baseline memory B cells in blood or lymph node samples (data not shown), making it difficult to outline the trajectory of their activation and differentiation. This is not unexpected given the very low frequency of such cells at baseline and the sampling limitation. Nevertheless, the dendrograms support the overall conclusion of our work that influenza virus vaccination in humans can induce a GC reaction that engages pre-existing memory B cell as well as naïve clones. We also agree with the reviewer that limited sampling could bias our results and have emphasized this point throughout the text.

3) Results are challenging to follow and need more clarity on key aspects of methodology throughout the paper. For a specific example, in reference to Fig.3b the authors say this data represents mAb and bulk BCR seq, but it is unclear how this information is integrated and contributing to the data in the figure. Also, I am unaware of a precedent for clonal BCR sequences being pulled out of bulk B cells, and if these are just Vh genes, how are you collating the data with the mAb? This needs to be explained or referenced in the text.

We have attempted to re-write for clarity throughout. To the point specified, we clarified (Line 180) that clonal overlap was calculated using only Vh gene sequencing data from mAbs combined with the bulk Vh gene sequencing from PBs, and added a citation for Vh-based clonal analysis from bulk B cells (Gupta et al., 2017). We would also like to note that the approach of using only Vh

sequences combined from bulk heavy chain repertoire sequencing and single-cell heavy:light paired repertoire sequencing for B cell clonal analysis has also been used by others such as (Horns et al., 2020). It has been previously shown that using only Vh sequences is sufficient to identify B cell clonal relationships with confidence in humans (Zhou and Kleinstein, 2019).

4) Much of the analyses and interpretation depends upon GC B cells uniquely expressing BCL6. In mice, studies by Okada and Haberman have shown that BCL6 can also be expressed during T-B clustering before the GC. What does the total BCL6 by KI67 expression look like after vaccination? Is it possible that activated memory cells that are BCL6+ have been included in the GC pool in these analyses? Is it possible to look prior to formation of a GC as these studies and those in mice suggest that secondary/memory GCs do not form with faster kinetics than a primary response? How quickly do HA-specific cells express BCL6 after vaccination? It would help support the authors claims if there was more detailed flow analyses of HA-specific cells instead of pre-gated histograms shown in extended Figure 2I.

This is an important point. We cannot formally exclude the possibility that a fraction of the cells we sorted as GC B cells include responding activated memory B cells that are in transition and we have now added this point to the updated manuscript (Line 147). However, we are confident in the definition of GC B cells based on the scRNA-seq analysis of participant 05, which are transcriptionally distinguished from activated B cells by their elevated expression of many genes, including RGS13, MEF2B, STMN1, ELL3, and SERPINA9 in addition to BCL6 (Extended Data Fig. 4f, g).

Our temporal resolution of FNA sampling is insufficient to determine the precise dynamics of Bcl6 upregulation in responding B cells. However, of the 3 participants who had HA^+ GCs, only participant 05 did not have HA^+ GC B cells by the first timepoint (day 5) following vaccination. In this sample, HA^+ $Ki\text{-}67^+$ $Bcl6^-$ cells are detected (Reviewer Fig. 7). HA^+ activated/memory B cell and PB gating have been added to Extended Data Fig. 1 and percentages of CD19 in FNAs have been added to Fig. 2i to more fully characterize the HA^+ response in the LN.

[Redacted]

Minor:

1) HA+ GCB only detected in 3/8 people, would be valuable to see data for total #GCB collected per timepoint/person as well as antigen specific. Detection error or biology? Extended 2h suggests pt 7 and 8 also had detectable GC, but not HA and all people had responses by ab. Emphasize that got GCB in 6/7 by flow. In Fig2, why not show distribution of B cell types as

analyzed in extended fig 2 (how many GCB, MBC etc over time? MFI is good characterization, now pair with quantification from Extended Table 4.

We report percent of CD19 for GC, memory B cells, etc. because the numbers of each population per FNA sample are confounded by variability in the total numbers of cells recovered per FNA (Extended Data Fig. 2a). The numbers of cells collected in each FNA correspond poorly to the volume of the LN (as well as cortical height and cortical volume), so we think that plotting numbers of given populations per FNA could be a misleading representation of the response. However, we have reported them in the revised Extended Data Table 1.

We have added emphasis in the updated manuscript that GC B cells were detected in 6 of the participants, with only 3 having HA+ GC B cells. Several reasons could explain this result; we may be not sampling the primary draining LN; it is possible that vaccination in these participants did not elicit a robust GC response. Also, we cannot formally exclude the possibility that these participants had a transient GC response that developed and waned between timepoints. However, we believe that the latter scenario is unlikely given that HA^+ GC B cells were detected for multiple weeks in participants 04, 05, and 11. Another possible technical failure we also believe to be unlikely is that our HA probes failed to identify QIV-responsive GC B cells in these participants. It is unlikely because we were able to detect HA-specific B cells in peripheral blood in the same participants (and in FNA samples of participants 07 and 09) using the same HA probes.

2) Fig 4: What were the circulating strains in the last 3 years that would suggest that only cross-reactive antibody could be memory?

Over the three years prior to our study, three (H1N1, H3N2, B/Vic) out of the four seasonal influenza vaccine strains have changed at least once. However, it is important to note that our projection for responding B cell clones' origin from a memory B cell or a naïve B cell is based on two main criteria. The first is the kinetics of appearance of the clone. It has been established that the early wave of PBs detected in peripheral blood within a week after influenza vaccination of humans is seeded by pre-existing memory B cells (Wrammert et al., 2008). The second criterion is the SHM frequency. In this manuscript, we augmented these criteria with the rationale that clones recognizing novel epitopes are likely to have a naive cell origin, as by definition novel epitopes would have not been previously encountered. While not absolute, specificity for only the vaccine strain within a subtype indicates a higher probability of a mAb recognizing a novel epitope and therefore originating from a naïve cell.

3) Need to report the standard error of the mean to measure variance.

We are not sure what this point exactly refers to. In Fig. 3 we had used Tukey's multiple comparisons test for the SHM frequency comparison between GC, shared, and PB clones. We showed every point, thus representing the full distribution of the data. Upon review and consultation with our bioinformatics colleagues, we have now replaced this test with Dunns' multiple comparisons test which tests for differences between the medians as the measure of centrality, which are now indicated by the horizontal lines in updated Fig. 3c and e, and which, compared to the means, are less prone to being skewed by outliers.

4) Fig 4 d: Were all unique binding antibodies tested in vivo independent of SHM? If not hard to say there is a trend in these ones being protective.

We agree with the reviewer that this was an over-generalization on our part. We picked the two described clones and showed their *in vivo* protective potential as an example that strain-specific mAbs recognizing epitopes not targeted by the early PB response could be protective *in vivo*, but we did not similarly examine all of the strain-specific mAbs. We have modified the text in the updated manuscript to reflect this point.

5) Isotype data in Ext. Fig 4 not mentioned in text anywhere, this is useful information.

We thank the reviewer for pointing out this oversight. These data have now been integrated into the text of the updated manuscript (Line 208).

6) Summary: Precision of language i.e. “proven”, “critical” are not supported by the data. Instead use “supports”, “contribute”.

This language has been modified accordingly in the updated manuscript.

References

Gupta, N.T., Adams, K.D., Briggs, A.W., Timberlake, S.C., Vigneault, F., and Kleinstein, S.H. (2017). Hierarchical Clustering Can Identify B Cell Clones with High Confidence in Ig Repertoire Sequencing Data. J. Immunol. *198*, 2489–2499.

Horns, F., Dekker, C.L., and Quake, S.R. (2020). Memory B Cell Activation, Broad Anti-influenza Antibodies, and Bystander Activation Revealed by Single-Cell Transcriptomics. Cell Rep. *30*, 905-913.e6.

Wrammert, J., Smith, K., Miller, J., Langley, W.A., Kokko, K., Larsen, C., Zheng, N.Y., Mays, I., Garman, L., Helms, C., et al. (2008). Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. Nature *453*, 667–671.

Zhou, J.Q., and Kleinstein, S.H. (2019). Cutting Edge: Ig H Chains Are Sufficient to Determine Most B Cell Clonal Relationships. J. Immunol. *203*, 1687–1692.

Reviewer Reports on the First Revision:

Referee #2 (Remarks to the Author):

In general this is a slight improvement over the previous version, though a few points remain.

Major points:

1. The authors show interesting data to the reviewers comparing clonal overlap between PB and GC in total cells or only when QIV-binding clones are analyzed. This information is key to interpreting this study’s findings, and it is critical that these two datasets be presented side-by-side in figure 2 of the paper. Presenting both datasets while explaining the associated caveats of each in the text provides an unbiased view of the authors’ findings, and I don’t see any downside to doing this. A note: in patient 04 in reviewer fig 1, the width of the overlap link in the circular plot does not seem to correspond to 27% of the blue arc.

2. The kinetic and mutational analyses are very informative, a clear improvement over the last submission. Once again, it is important that data for QIV-binding and non-QIV-binding clones over time

be presented side-by-side in the manuscript, and not only as a reviewer figure. Links between GCs from different time points in the circular graphs are confusing and should be removed, especially if they count towards the % clonal overlap.

Other points:

Line 42: "vaccine-induced GC responses" this should be QIV-binding B cells in 3 of 8 individuals, since the authors can't exclude that the other patients that responded with GCs with good kinetics but didn't contain QIV-binding cells were not vaccine-induced.

Line 43: the authors state that "Up to 88% of the vaccine-induced GC B cell clones overlapped…", which is a bit biased. Better would be to say something like "Between 12% and 88% of QIV-binding (and 6.8% to 58% of total) GC B cell clones overlapped…" Given that the number of subjects is so small it is safer to emphasize the qualitative findings and the variability rather than any specific quantitative data point.

Lines 94-96, 99, 173, 177, 182, etc: consider rephrasing "QIV/HA-specific" as "QIV/HA-binding." Again, it is very likely that many cells with low/undetectable affinity are still recruited specifically by the vaccine.

Regarding time points:

Where does the data from d21 for 321-04 come from, given that this time point was excluded from Fig 2i, extended data Figs. 2h and 2i and extended data table 1? Also, in extended data Fig. 3d, for d21 a 100% overlap is reported which includes a link to a neighboring GC. Is this included in the 100% or is that a single clone?

In Fig. 3A, patient 321-11 has antibodies cloned from week 1, 2 and 5; in extended data Fig. 2i, the reported time points are week 1, 2, 4, and 9 (same as in extended data Fig. 2i,2h); for the kinetic breakdown of the clonal overlap, the days plotted are d6, d18, and d35. Can this be stated more precisely?

Minor:

Line 154: "Given the role of GCs in diversifying B cell responses, we asked whether the epitopes targeted in participants 04, 05, and 11 diversified over time." See my previous comment on how GC diversification is unlikely to change EMPEM patterns.

Line 504: "expoxysilane-coated glass slides (Schott)" is repeated.

Referee #3 (Remarks to the Author):

The authors have made great improvements to the manuscript by putting their work in context and the impact comes across more clearly. The added figures also add valuable evidence to their story and the data support their conclusion that pre-exposed adults can form GC responses to flu immunization and that both memory and less mature memory/naïve cells participate in the GC response to the vaccine. The results section and figure legends could still benefit from more detail and clarity. In particular, the variability between individuals in the cohort is striking and needs additional analysis so that the important findings are not undermined.

1) Context of the work:
a. When discussing the literature related to this work, it is important to distinguish between homologous vs heterologous challenge (lines 61-67, 284-285). While mouse studies of homologous challenge have shown that class-switched and IgM MBCs rarely enter GCs (though IgD antigen-experienced cells

can)(Pape et al. 2011; Zuccarino-Catania et al. 2014; Murugan et al. 2018; Mesin et al. 2020). Studies with heterologous challenge (Purtha et al 2011, Leach 2019) suggest MBCs are key to protection, but do not address participation in secondary GCs. Since the vaccine strains are somewhat distinct from the strains that had circulated in the last 3+ years, this work is examining a heterologous challenge. Thus, this work does not conflict with previous findings, but expands them.
i. line 285: Should also cite all the other mouse work that this repeated.
b. Line 340. The authors conclude that changes to vaccine formulation could be better at inducing GCs, but since seasonal strains are so similar how would a new vaccine avoid high levels of circulating antibodies that the authors suggest are interfering with GC formation? Emphasis on inclusion of novel strains? Do the authors speculate that the only reason they were able to observe GC responses in these individuals is because they hadn't been vaccinated for 3 years?

2) Sample variability
a. Authors claimed all participants had similar responses among PBMCs, but some had more enrichment for HA+ responses in PBMCs (Fig. 1 and Ext Fig. 1). Similarly, Ext. Fig. 2 looks like GCB cells were detected in 5/7 samples, but HA+ GCBs only in 3. It would be clearer to add this sort of summary statement to the results in addition to the more precise listing of individual numbers. Could the authors plot the % GCB and % HA+ GCB of CD19+ both in Figure 2 to make it clear that detection of HA+ GCB cells in only 3 individuals was at least not because the FNAs didn't sample GCs. Can the authors correlate the HA+ PBMC response to the HA+ GC response to also address possible sampling issues? The authors mention in the discussion that high levels of circulating antibody could inhibit GC responses. Is this correlation seen among the cohort or is it too low powered?
b. Line 159-166 (Figure 2jk): The authors claims of diversification over time would be better supported with a statistical measurement of uncertainty, which captures the number of antibodies mapped and how likely it is to capture the same antibody in two subsequent draws if the frequency in the plasma remained the same. Similarly, the abstract statement (line 49-50) about GC-only mAbs "selectively recognized novel epitopes that were not targeted by the early plasmablast response" is based on epitope mapping. This would be more powerful with a statistical measure that enough mAbs from each group were tested to say that there were not plasmablast-derived antibodies that could bind these sites. The IVPM data suggests this could be true.
c. Would be valuable to have the summary measure of % strain-specific vs cross-reactive antibodies for individuals 4 and 11 in Ext. Fig 5 as in Fig 4b, especially because there seem to be fewer cross-reactive clones than among those from individual 5. This variability is not mentioned in the text, were there just too few antibodies? In both Figure 4b and Ext. Figure 5 it would be valuable to see the summary data with PB, shared and GCs clones plotted separately because it is the difference between shared clones and GC-only clones that is key to what types of cells are entering GCs.

3) Data interpretation
a. Line 239. The data suggest that naïve or less mutated MBCs do participate in the GC, but what is the evidence that it is preferential? According to Figure 3b, the majority of GCB cell clones in 2 out of 3 people were shared with PB and were not restricted to the GC.

4) Clear methodology/figure clarity
a. Figure 3b-d legend are the GC B cells QIV-specific? Does the purple in Figure 3b represent QIV-specific clones and the full circle is the bulk repertoire? The text lists that participant 4 only had 4 QIV-specific PB, but there seem to be many more in Figure 3b. This is a key piece of data and could benefit from additional clarity.
b. Line 134 (Figure 2) and Line 202 (Figure 3). The results and methods say that whole PBMCs and enriched IgDlo PBMCs were used for scRNAseq. How is the data from the IgDlo enrichment incorporated into the figures?
c. Line 220-221 (Figure 3f): Based on the figure legend, it seems that Figure 3F is presenting the SHM of total B cells in these populations and not a summary of Figure 3e with only QIV-specific clones. This needs clarification in the legend and text. The point that the GC-only clones are less mutated among all GC clones would be better made by Figure 3c. The total non-QIV specific SHM demonstrates that PB and GCB have more SHM than naïve B cells, but it would be more important to show how the SHM of GC-only QIV clones compares to that of QIV-specific or total naïve B cells.

d. Line 232-234 (Figure 4a): Evidence of cross-reactivity would be made clearer if the authors highlighted the strains contained in the QIV, strains the participants could have been exposed to in past vaccines or natural infections and the strains no longer circulating in humans in Figure 4a and Ext. Figure 5a-c. Did any of the individuals get the flu in the last 3 years (since they weren't vaccinated)?
e. Figure 3b. What was the clonal diversity captured among QIV-specific GCB and QIV-specific PB? For example, what was the proportion of clonal burst and how many cells were participant in the most outgrown clones? This would be valuable for others to build on this work, but not required for the thesis of this paper.

Minor:
1) Would the authors consider a title that is more reflective of their findings; like "Human flu vaccination induces recalled and de novo BCR clones to participate in GC responses"
2) Line 42: Should 'vaccine-induced GC responses' be 'vaccine-specific GC responses'?
3) Line 316: The implication should be addressed. This would suggest that more naïve B cells contribute to the GC than implied by the data.
4) Consistent nomenclature:
a. Label samples as PBMCs not blood if that is accurate (Fig 1).
b. Label samples with individual code e.g. 04 instead of the full 321-04 in both text and figures (eg line 128).
5) What timepoint was used for samples in Ext Fig 2j?
6) Axis and annotation text in figures often unreadably small (eg Fig. 2g).
7) As Figure 2 includes both flow cytometry and scRNAseq data, the authors should clarify the methodology of each panel in the figure legend.
8) Please include the reference for Vh sequencing (Zhou and Kleinstein 2019) in the text.
9) How does Ext Fig 2c differ from Ext Fig 4a? Figure legends need clarification.

Author Rebuttals to First Revision:

We would again like to thank the editor and reviewers for their time and continued consideration of our manuscript. The comments provided were very helpful and we feel that incorporating the recommended modifications has helped further clarify our conclusions and present them more comprehensively. Below we have provided responses to reviewers' questions detailing the ways in which our revised manuscript addresses their concerns.

Reviewer #2:
In general this is a slight improvement over the previous version, though a few points remain.

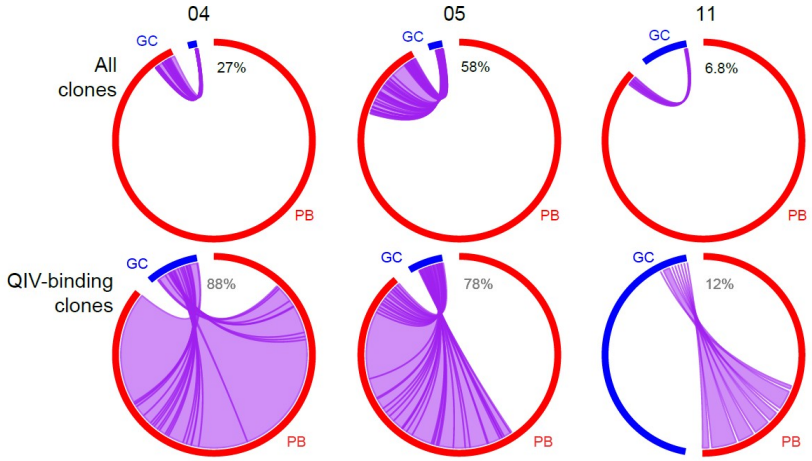
We thank the reviewer for carefully reading our revised manuscript and for the recommendations that helped us more comprehensively and clearly present our findings.

Major points:

1. The authors show interesting data to the reviewers comparing clonal overlap between PB and GC in total cells or only when QIV-binding clones are analyzed. This information is key to interpreting this study's findings, and it is critical that these two datasets be presented side-by-side in figure 2 of the paper. Presenting both datasets while explaining the associated caveats of each in the text provides an unbiased view of the authors' findings, and I don't see any downside to doing this. A note: in patient 04 in reviewer fig 1, the width of the overlap link in the circular plot does not seem to correspond to 27% of the blue arc.

We agree with the reviewer that adding the data from non-QIV binding clones presents a more comprehensive picture of the responses. We have added these data to Fig 3b, also shown as reviewer Fig. 1, below.

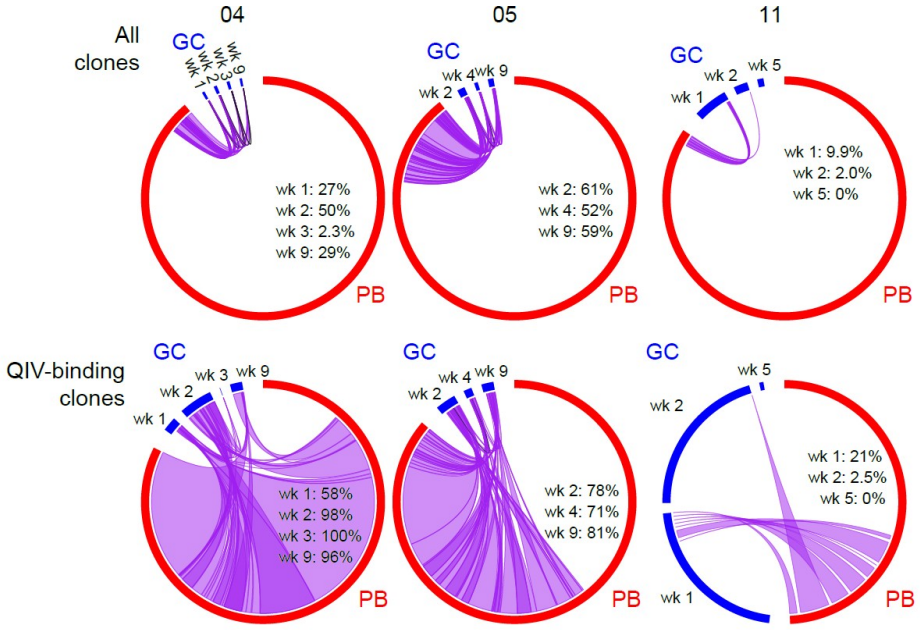
The chord plots for participants 04 and 11 had been inadvertently switched in the previous reviewer figure, which is why the purple chords did not cover 27% of the blue arc; this has been rectified.



Reviewer Figure 1. GC and PB clonal overlap among total and QIV-binding clones. Clonal overlap of sequences from mAb cloning and bulk repertoire analysis between PBs sorted from PBMCs 1-week post-vaccination and GC B cells from all timepoints among total (top) and only QIV-binding (bottom) sequences. Chord width corresponds to clonal population size. Percentages are of GC sequences overlapping with PBs.

2. The kinetic and mutational analyses are very informative, a clear improvement over the last submission. Once again, it is important that data for QIV-binding and non-QIV-binding clones over time be presented side-by-side in the manuscript, and not only as a reviewer figure. Links between GCs from different time points in the circular graphs are confusing and should be removed, especially if they count towards the % clonal overlap.

The clonal overlap chord plots for all clones have been added to Extended Data Fig 3, also shown as reviewer Fig. 2 below. We feel it is important to include the chords between timepoints for non-shared GC clones that were found at multiple timepoints, as their omission could imply that only shared clones were found to persist in GCs. To try to distinguish them more clearly from the shared chords, we have changed their color from grey to black (between all four timepoints for participant 04 (total clones), and between week 2 and 4 for participant 05 (total and QIV-binding clones)). The percent overlaps reported are only those between the GC and PB repertoires and are independent of whether GC clones are observed at multiple timepoints. We have clarified this in the text and legends (lines 180-185, 529-533, and 1039-1044).



Reviewer Figure 2. Kinetics of GC and PB clonal overlap among total and QIV-binding clones. Clonal overlap of sequences from mAb cloning and bulk repertoire analysis between PBs sorted from PBMCs 1-week post-vaccination and GC B cells from the indicated timepoint among total (top) and only QIV-binding (bottom) sequences. Purple chords link overlapping GC and PB clones; black chords link GC clones found at multiple timepoints that did not participate in the early PB response. Chord width corresponds to clonal population size. Percentages are of GC sequences overlapping with PBs.

Other points:

Line 42: “vaccine-induced GC responses” this should be QIV-binding B cells in 3 of 8 individuals, since the authors can’t exclude that the other patients that responded with GCs with

good kinetics but didn't contain QIV-binding cells were not vaccine-induced.

We agree with the reviewer; the text has been revised (line 43).

Line 43: the authors state that "Up to 88% of the vaccine-induced GC B cell clones overlapped…", which is a bit biased. Better would be to say something like "Between 12% and 88% of QIV-binding (and 6.8% to 58% of total) GC B cell clones overlapped…" Given that the number of subjects is so small it is safer to emphasize the qualitative findings and the variability rather than any specific quantitative data point.

We agree with the reviewer; the text has been revised (line 44).

Lines 94-96, 99, 173, 177, 182, etc: consider rephrasing "QIV/HA-specific" as "QIV/HA-binding." Again, it is very likely that many cells with low/undetectable affinity are still recruited specifically by the vaccine.

The text has been revised throughout.

Regarding time points:

Where does the data from d21 for 321-04 come from, given that this time point was excluded from Fig 2i, extended data Figs. 2h and 2i and extended data table 1? Also, in extended data Fig. 3d, for d21 a 100% overlap is reported which includes a link to a neighboring GC. Is this included in the 100% or is that a single clone?

This FNA sample had been excluded from frequency analyses due to poor sampling of GCs (we speculate that interfollicular regions were somehow oversampled given the decreased frequency in GC B cells and increased frequency of $CD38^{hi}$ ASCs, reviewer Fig. 3 below). In an effort to maximize the number of GC B cells cloned, we still sorted GC B cells from it. We identified one QIV-binding clone from this timepoint, which was also found in the early PB response (hence the 100% overlap) and in d12 GC. To better reflect sampling variability, it has been restored to the frequency analyses. Similar patterns of decreased GC sampling (still with CD4/CD14/CD19 profiles consistent with lymphoid tissue) were also observed at week 2 for participant 07 and week 3 for participant 08.

[Redacted]

In Fig. 3A, patient 321-11 has antibodies cloned from week 1, 2 and 5; in extended data Fig. 2i, the reported time points are week 1, 2, 4, and 9 (same as in extended data Fig. 2i,2h); for the kinetic breakdown of the clonal overlap, the days plotted are d6, d18, and d35. Can this be stated more precisely?

We thank the reviewer for bringing our attention to this discrepancy. The timepoints are now consistently labeled throughout.

Minor:

Line 154: “Given the role of GCs in diversifying B cell responses, we asked whether the epitopes targeted in participants 04, 05, and 11 diversified over time.” See my previous comment on how GC diversification is unlikely to change EMPEM patterns.

We have reworded this and clarified our previous response that we are referring to recruitment of new clones with different specificities to the response, not SHM (line 152).

Line 504: “expoxysilane-coated glass slides (Schott)” is repeated.

This has been deduplicated.

Reviewer #3:
The authors have made great improvements to the manuscript by putting their work in context and the impact comes across more clearly. The added figures also add valuable evidence to their story and the data support their conclusion that pre-exposed adults can form GC responses to flu immunization and that both memory and less mature memory/naïve cells participate in the GC response to the vaccine. The results section and figure legends could still benefit from more detail and clarity. In particular, the variability between individuals in the cohort is striking and needs additional analysis so that the important findings are not undermined.

Reviewer 3 has made excellent comments that have helped us present our findings more clearly, for which we are grateful. Specific details are outlined below.

1) Context of the work:
a. When discussing the literature related to this work, it is important to distinguish between homologous vs heterologous challenge (lines 61-67, 284-285). While mouse studies of homologous challenge have shown that class-switched and IgM MBCs rarely enter GCs (though IgD antigen-experienced cells can)(Pape et al. 2011; Zuccarino-Catania et al. 2014; Murugan et al. 2018; Mesin et al. 2020). Studies with heterologous challenge (Purtha et al 2011, Leach 2019) suggest MBCs are key to protection, but do not address participation in secondary GCs. Since the vaccine strains are somewhat distinct from the strains that had circulated in the last 3+ years, this work is examining a heterologous challenge. Thus, this work does not conflict with previous findings, but expands them.

We thank the reviewer for this important suggestion. The discussion has been revised to better contextualize the work and account for homologous and heterologous boost (lines 283-291).

i. line 285: Should also cite all the other mouse work that this repeated.

We have revised this section of the discussion and added the citations.

b. Line 340. The authors conclude that changes to vaccine formulation could be better at inducing GCs, but since seasonal strains are so similar how would a new vaccine avoid high levels of circulating antibodies that the authors suggest are interfering with GC formation? Emphasis on inclusion of novel strains?

While pre-existing antibodies may impede B cell responses targeting novel epitopes, we believe formulations that emphasize presentation of novel epitopes, adjuvants, and/or novel modes of antigen delivery (e.g. mRNA vaccines) which can deliver higher antigen doses may help to overcome this barrier. We have added these points to the discussion (lines 328-333).

Do the authors speculate that the only reason they were able to observe GC responses in these individuals is because they hadn't been vaccinated for 3 years?

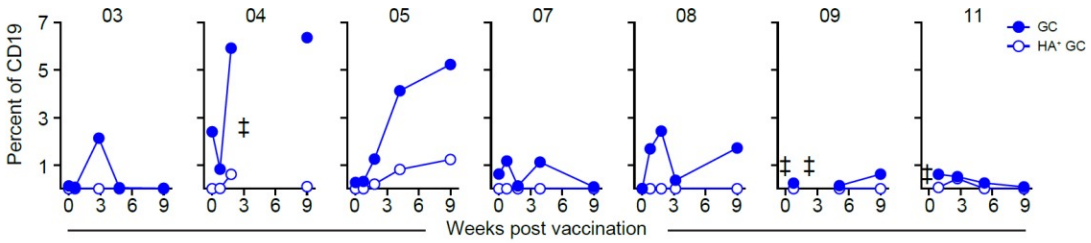
This is an important point. Although we do not yet have strong evidence, we would speculate that the level of pre-existing antibodies can modulate the magnitude and persistence of GC responses.

2) Sample variability
a. Authors claimed all participants had similar responses among PBMCs, but some had more enrichment for HA+ responses in PBMCs (Fig. 1 and Ext Fig. 1). Similarly, Ext. Fig. 2 looks like GCB cells were detected in 5/7 samples, but HA+ GCBs only in 3. It would be clearer to add this sort of summary statement to the results in addition to the more precise listing of individual numbers.

Statements regarding variability among PBMC responses (lines 94-95, 97-98) and a summary statement for total GC responses among the participants (lines 136-139) have been added to the results.

Could the authors plot the % GCB and % HA+ GCB of CD19+ both in Figure 2 to make it clear that detection of HA+ GCB cells in only 3 individuals was at least not because the FNAs didn't sample GCs.

These have been combined in the updated Fig. 2i, also shown below as reviewer Fig. 4.



Reviewer Figure 4. Total and HA⁺ GC B cells. Kinetics of total (closed circles) and HA+ (open circles) GC B cells in FNA, as defined by flow cytometry gates in f and h. Daggers denote samples excluded from analysis due to low cell recovery or blood contamination.

Can the authors correlate the HA+ PBMC response to the HA+ GC response to also address possible sampling issues? The authors mention in the discussion that high levels of circulating antibody could inhibit GC responses. Is this correlation seen among the cohort or is it too low powered?

We agree that such correlation would have been informative, but our study is too small for robust correlative analysis. Participants 04 and 05 were among the most robust responders in terms of peripheral responses and their HA+ GC responses were larger and more prolonged than that of 11, in whom we detected a much smaller peripheral HA+ response, suggesting the two responses may be correlated. However, participant 11 had a less robust peripheral response than most of the participants in whom no HA+ GC response was found at all, though we cannot exclude that these participants had undetected GC responses. A much larger study will be required to determine whether vaccine-induced GC responses correlate with peripheral responses or pre-existing antibody titers.

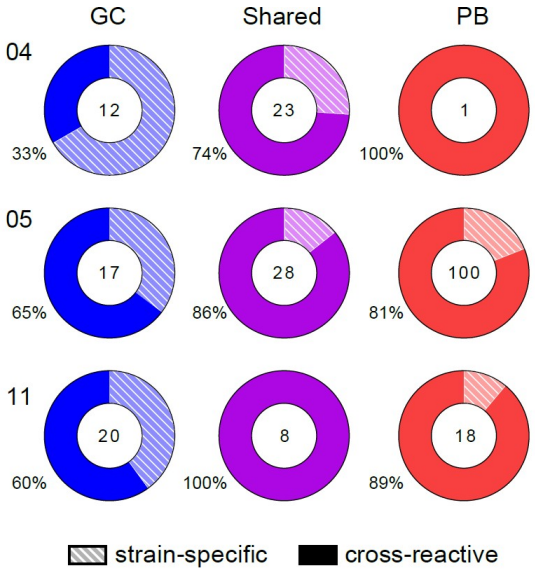
b. Line 159-166 (Figure 2jk): The authors claims of diversification over time would be better supported with a statistical measurement of uncertainty, which captures the number of antibodies mapped and how likely it is to capture the same antibody in two subsequent draws if the frequency in the plasma remained the same. Similarly, the abstract statement (line 49-50) about GC-only mAbs “selectively recognized novel epitopes that were not targeted by the early plasmablast response” is based on epitope mapping. This would be more powerful with a statistical measure that enough mAbs from each group were tested to say that there were not plasmablast-derived antibodies that could bind these sites. The IVPM data suggests this could be true.

We agree with the reviewer that these conclusions would be more strongly supported with quantitative measures. At this stage, we are unable to establish the quantitative parameters of EMPEM. However, EMPEM has been demonstrated as a sensitive measure of circulating antigen-specific pAbs, identifying specificities that were not detected by cloning mAbs from PBs (Bianchi et al., 2018; Han et al., 2020). We sought to use it to sample the epitope landscape of the PB response as thoroughly as possible.
As we cannot exclude the possibility that we failed to detect low-affinity and/or -frequency pAbs from the PB response that targeted novel epitopes by EMPEM, we have moderated our conclusions in this regard. Additionally, we have altered the abstract to remove the reference to GC mAbs’ selective recognition of epitopes not targeted by the PB response, as we have not

comprehensively mapped the epitope landscape for this population. Of note, we have been unsuccessful generating sufficiently stable constructs of some strain-specific GC Fabs for epitope mapping, suggesting this may also occur for some plasma pAbs. We have added these caveats to the discussion (lines 306-310).

c. Would be valuable to have the summary measure of % strain-specific vs cross-reactive antibodies for individuals 4 and 11 in Ext. Fig 5 as in Fig 4b, especially because there seem to be fewer cross-reactive clones than among those from individual 5. This variability is not mentioned in the text, were there just too few antibodies? In both Figure 4b and Ext. Figure 5 it would be valuable to see the summary data with PB, shared and GCs clones plotted separately because it is the difference between shared clones and GC-only clones that is key to what types of cells are entering GCs.

We have further clarified that Fig 4b contains data pooled from participants 04, 05, and 11 (lines 231-235). We run out of power for statistical comparisons if the sequences are divided amongst participants, although the trends are consistent. We have added separated plots as Extended Data Fig. 5d, shown below as reviewer Figure 5.



Reviewer Figure 5. Cross-reactivity of GC, shared, and PB clones. Percentages of mAbs that bound two or more HA strains from participants 04, 05, and 11 from GC clones that did not participate in the early PB response (blue), clones participating in both GC and early PB responses (purple), and from PB clones not found in GCs (red). The number in the middle of the pie charts is the total number of mAbs.

3) Data interpretation

a. Line 239. The data suggest that naïve or less mutated MBCs do participate in the GC, but what is the evidence that it is preferential? According to Figure 3b, the majority of GCB cell clones in 2 out of 3 people were shared with PB and were not restricted to the GC.

We agree with the reviewer that this point needed clarification. We are not saying the GC is primarily comprised of naïve/less mutated MBCs, but rather when naïve/less mutated MBCs respond, they are preferentially recruited to the GC rather than the early PB response (in line with previous reports of the early PB response being primarily comprised of highly mutated cells (Ellebedy et al., 2016; Wrammert et al., 2008)). The evidence for this is the significantly lower SHM level among clones that participate in the GC but not early PB response compared to those that participate in both the GC and early PB responses, as well as to those that only participate in the PB response (Fig. 3c). We have sought to clarify this point (lines 192-197).

4) Clear methodology/figure clarity

a. Figure 3b-d legend are the GC B cells QIV-specific? Does the purple in Figure 3b represent QIV-specific clones and the full circle is the bulk repertoire? The text lists that participant 4 only had 4 QIV-specific PB, but there seem to be many more in Figure 3b. This is a key piece of data and could benefit from additional clarity.

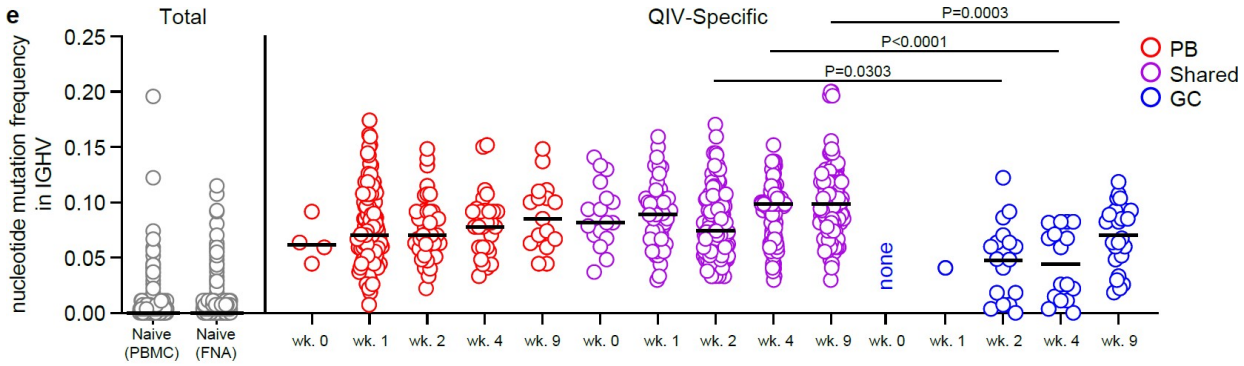
We have clarified that all the clones in the previous Fig. 3b-d are QIV-binding. We are classifying cells/sequences "QIV-binding" by virtue of being clonally related to mAb(s) that tested positive for QIV-binding. We have also added chord diagrams for total clones to the top row of Fig. 3b and sought to clarify what they represent (lines 180-185 and 529-533), see also reviewer Fig. 1 in the response to Reviewer 2's comments above. The full circle in the upper panels of the updated Fig. 3b represents all sequences from mAb cloning of GC B cells (blue arc) and PBs as well as bulk-sorted PBs (red arc). The lower panels (shown previously as Fig. 3b) represent only those clones that encode QIV-binding mAbs. The purple chords are drawn between clonally related sequences that appear in both GC and PB responses.
The four PB mAbs from participant 04 referenced on line 177 in the text were the only QIV-binding PB mAbs that were not clonally related to cells also participating in the GC response (i.e. non-shared). In the lower panel, the portion of the red arc not covered by the purple chords represents all of the sequences (from mAb cloning and bulk PB sorting) clonally related to these four PB mAbs within the PB response. In comparison, the proportion of the red arc that is covered by the purple chords represents all of the PB sequences that were considered QIV-specific by virtue of being clonally related to QIV-binding GC mAbs.

b. Line 134 (Figure 2) and Line 202 (Figure 3). The results and methods say that whole PBMCs and enriched IgDlo PBMCs were used for scRNAseq. How is the data from the IgDlo enrichment incorporated into the figures?

This was a mistake; the libraries from IgD^{lo} enriched PBMC were not included in Fig. 2, and the text and legend have been updated accordingly. The t-SNE plot for IgD^{lo} enriched PBMC is shown in Extended Data Fig. 2g. In Fig. 3, the top row of plots shows sequencing data from whole PBMC and IgD^{lo} libraries combined for each timepoint.

c. Line 220-221 (Figure 3f): Based on the figure legend, it seems that Figure 3F is presenting the SHM of total B cells in these populations and not a summary of Figure 3e with only QIV-specific clones. This needs clarification in the legend and text. The point that the GC-only clones are less mutated among all GC clones would be better made by Figure 3c. The total non-QIV specific SHM demonstrates that PB and GCB have more SHM than naïve B cells, but it would be more important to show how the SHM of GC-only QIV clones compares to that of QIV-specific or total naïve B cells.

We have altered the figure to show only naïve B cells compared to QIV-binding clones, with total GC and total PB removed, also shown below as reviewer Fig. 6. These SHM frequencies come only from scRNAseq data, whereas the mutation data in Fig. 3e comes from mAb cloning and bulk PB sequencing.



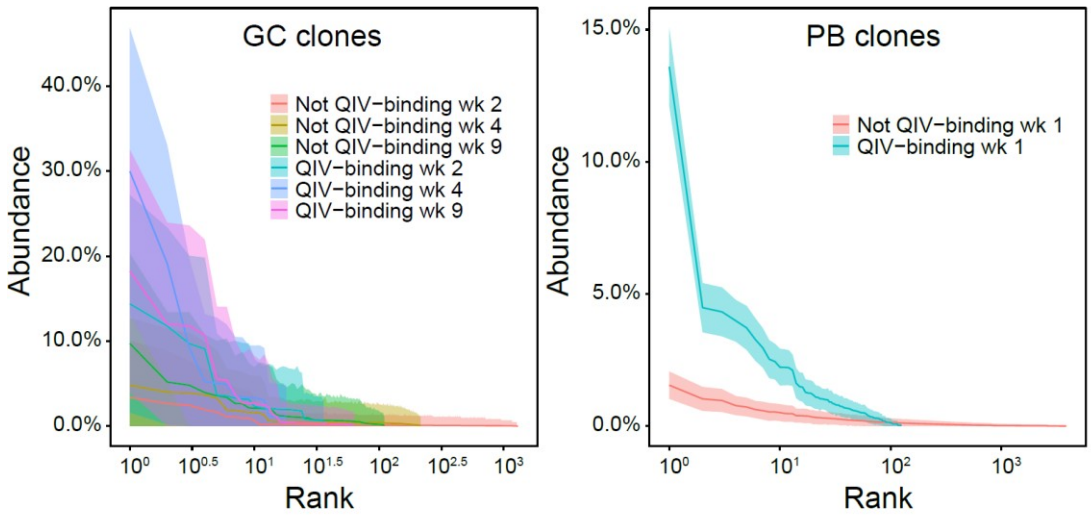
Reviewer Figure 6. *IGHV* mutation frequency of naïve B cells pooled from all time-points (left) and the indicated populations at the indicated timepoint (right) from scRNA-seq of whole and memory B cell-enriched PBMC and FNA samples from participant 05.

d. Line 232-234 (Figure 4a): Evidence of cross-reactivity would be made clearer if the authors highlighted the strains contained in the QIV, strains the participants could have been exposed to in past vaccines or natural infections and the strains no longer circulating in humans in Figure 4a and Ext. Figure 5a-c. Did any of the individuals get the flu in the last 3 years (since they weren't vaccinated)?

Relevant QIV-included and recently circulating strains have been highlighted. We unfortunately do not have a documented disease history for any of the participants.

e. Figure 3b. What was the clonal diversity captured among QIV-specific GCB and QIV-specific PB? For example, what was the proportion of clonal burst and how many cells were participant in the most outgrown clones? This would be valuable for others to build on this work, but not required for the thesis of this paper.

We have added clonal abundance plots from scRNAseq data from participant 05 GC and PB responses showing the clone size and rank as Extended Data Fig. 3e, also shown as reviewer Fig. 7, below.



Reviewer Figure 7. Clonal abundance plots. Clonal rank-abundance distributions of GC B cells from indicated timepoints (left) and of early blood PBs (right). The number of GC B cells or early blood PBs in a clone as a percentage of the total GC or early blood PB repertoire (y-axis) is plotted against the abundance rank of that clone (x-axis). Solid lines represent the estimated clonal abundance curves, with shaded bands representing the 95% confidence intervals from 200 bootstraps.

Minor:

1) Would the authors consider a title that is more reflective of their findings; like “Human flu vaccination induces recalled and de novo BCR clones to participate in GC responses”

We thank the reviewer for the suggestion. We propose the following title, “Human germinal centres engage memory and naïve B cells after influenza vaccination”.

2) Line 42: Should ‘vaccine-induced GC responses’ be ‘vaccine-specific GC responses’?

Yes; this has been changed.

3) Line 316: The implication should be addressed. This would suggest that more naïve B cells contribute to the GC than implied by the data.

We agree; this point has been further discussed (lines 299-306).

4) Consistent nomenclature:

a. Label samples as PBMCs not blood if that is accurate (Fig 1).

We collected blood for both PBMC and plasma.

b. Label samples with individual code e.g. 04 instead of the full 321-04 in both text and figures (eg line 128).

This has been revised throughout.

5) What timepoint was used for samples in Ext Fig 2j?

Histographs were from week 4; MFIs from week 2 or 4, depending on participant. The figure legend has been updated to include this information (lines 1014-1015).

6) Axis and annotation text in figures often unreadably small (eg Fig. 2g).

Where feasible, we have increased the font sizes.

7) As Figure 2 includes both flow cytometry and scRNAseq data, the authors should clarify the methodology of each panel in the figure legend.

The legend has been updated to specify methodology for each panel (line 503).

8) Please include the reference for Vh sequencing (Zhou and Kleinstein 2019) in the text.

This reference has been added (line 181).

9) How does Ext Fig 2c differ from Ext Fig 4a? Figure legends need clarification.

The legend for Extended Data Fig. 4a was incorrect. Extended Data Fig. 2c is overall level clustering including all cells; Extended Data Fig. 4a is secondary clustering of just the cells from the “B cell” cluster of Extended Data Fig. 2e. The legend has been corrected (lines 1055-1057).

References:

Bianchi, M., Turner, H.L., Nogal, B., Cottrell, C.A., Oyen, D., Pauthner, M., Bastidas, R., Nedellec, R., McCoy, L.E., Wilson, I.A., et al. (2018). Electron-Microscopy-Based Epitope Mapping Defines Specificities of Polyclonal Antibodies Elicited during HIV-1 BG505 Envelope Trimer Immunization. Immunity *49*, 288-300.e8.

Ellebedy, A.H., Jackson, K.J.L., Kissick, H.T., Nakaya, H.I., Davis, C.W., Roskin, K.M., McElroy, A.K., Oshansky, C.M., Elbein, R., Thomas, S., et al. (2016). Defining antigen-specific plasmablast and memory B cell subsets in human blood after viral infection or vaccination. Nat. Immunol. *17*, 1226–1234.

Han, J., Schmitz, A.J., Richey, S.T., Dai, Y.-N., Turner, H.L., Mohammed, B.M., Fremont, D.H., Ellebedy, A.H., and Ward, A.B. (2020). Polyclonal epitope cartography reveals the temporal dynamics and diversity of human antibody responses to H5N1 vaccination. BioRxiv 2020.06.16.155754.

Wrammert, J., Smith, K., Miller, J., Langley, W.A., Kokko, K., Larsen, C., Zheng, N.-Y., Mays, I., Garman, L., Helms, C., et al. (2008). Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. Nature *453*, 667–671.

Reviewer Reports on the Second Revision:

Referee #2 (Remarks to the Author):

This is a much improved version of the manuscript with the addition of the analysis of the non-QIV binders and the more balanced discussion. I recommend acceptance.

The authors should ensure that the figure with the non-binders is not removed from the main article when it is edited to fit Nature size limits. The discussion should also be included in an extended text file if allowed.

Referee #3 (Remarks to the Author):

The authors have addressed all of our questions and concerns adequately. This is a very nice contribution to the field.