Spec-seq unveils transcriptional subpopulations of antibody-secreting cells following influenza vaccination

Karlynn E. Neu, … , Aly A. Khan, Patrick C. Wilson


Vaccines are among the most effective public health tools for combating certain infectious diseases such as influenza. The role of the humoral immune system in vaccine-induced protection is widely appreciated; however, our understanding of how antibody specificities relate to B cell function remains limited due to the complexity of polyclonal antibody responses. To address this, we developed the Spec-seq framework, which allows for simultaneous monoclonal antibody (mAb) characterization and transcriptional profiling from the same single cell. Here, we present the first application of the Spec-seq framework, which we applied to human plasmablasts after influenza vaccination in order to characterize transcriptional differences governed by B cell receptor (BCR) isotype and vaccine reactivity. Our analysis did not find evidence of long-term transcriptional specialization between plasmablasts of different isotypes. However, we did find enhanced transcriptional similarity between clonally related B cells, as well as distinct transcriptional signatures ascribed by BCR vaccine recognition. These data suggest IgG and IgA vaccine–positive plasmablasts are largely similar, whereas IgA vaccine–negative cells appear to be transcriptionally distinct from conventional, terminally differentiated, antigen-induced peripheral blood plasmablasts.

Find the latest version:
http://jci.me/121341/pdf
Spec-seq unveils transcriptional subpopulations of antibody-secreting cells following influenza vaccination

Karlynn E. Neu,1,2 Jenna J. Guthmiller,2 Min Huang,2 Jennifer La,3 Marcos C. Vieira,4 Kangchon Kim,4 Nai-Ying Zheng,2 Mario Cortese,5 Micah E. Tepora,1 Natalie J. Hamel,1 Karla Thatcher Rojas,2 Carole Henry,2 Dustin Shaw,1,2 Charles L. Dulberger,6 Bali Pulendran,5 Sarah Cobey,4 Aly A. Khan,7 and Patrick C. Wilson1,2

1The Committee on Immunology, 2The Department of Medicine, Section of Rheumatology, 3The Department of Pathology, Molecular Pathogenesis and Molecular Medicine, and 4The Department of Ecology and Evolution, The University of Chicago, Chicago, Illinois, USA. 5Emory Vaccine Center, Emory University, Atlanta, Georgia, USA. 6The Department of Biochemistry and Molecular Biophysics, The University of Chicago, Chicago, Illinois, USA. 7Toyota Technological Institute at Chicago, Chicago, Illinois, USA.

Vaccines are among the most effective public health tools for combating certain infectious diseases such as influenza. The role of the humoral immune system in vaccine-induced protection is widely appreciated; however, our understanding of how antibody specificities relate to B cell function remains limited due to the complexity of polyclonal antibody responses. To address this, we developed the Spec-seq framework, which allows for simultaneous monoclonal antibody (mAb) characterization and transcriptional profiling from the same single cell. Here, we present the first application of the Spec-seq framework, which we applied to human plasmablasts after influenza vaccination in order to characterize transcriptional differences governed by B cell receptor (BCR) isotype and vaccine reactivity. Our analysis did not find evidence of long-term transcriptional specialization between plasmablasts of different isotypes. However, we did find enhanced transcriptional similarity between clonally related B cells, as well as distinct transcriptional signatures ascribed by BCR vaccine recognition. These data suggest IgG and IgA vaccine-positive plasmablasts are largely similar, whereas IgA vaccine-negative cells appear to be transcriptionally distinct from conventional, terminally differentiated, antigen-induced peripheral blood plasmablasts.

Introduction

The medical field has relied on vaccines for improving human health since the days of smallpox inoculation—long before antibodies were identified (1, 2). Vaccines induce a humoral immune response, in which B cells mediate antibody protection against a wide range of bacterial and viral pathogens, including influenza. The influenza vaccine provides annual protection for millions of people worldwide by initiating an immune response against the 4 influenza strains most likely to circulate that year. Viral antigens within the vaccine induce a global immune response that provides immediate protection through antibody-secreting cells (ASCs), while establishing a foundation for future protection within the memory B cell compartment.

The recognition of antigen by surface-bound B cell receptors initiates the humoral immune response. The BCR is generated through a process of recombination and somatic hypermutation of the immunoglobulin (Ig) genes, resulting in a unique receptor sequence that is only shared with the clonal progeny of that cell. The variable domains of the receptor determine the ability to bind certain antigens, while the Fc domain (comprised of the BCR constant domain) is essential for directing downstream immunological responses. After antigen exposure, a B cell can undergo class switch recombination (CSR), where the constant domain of the BCR is swapped. The new constant domain confers a new BCR isotype, which gives the secreted antibody distinct effector functions. These functions are tailored to best combat the current challenge, with different isotypes having different binding affinities for the C1q molecule of the complement cascade and pro- or antiinflammatory Fc receptors. Post-translational modifications to the Fc domain can further alter these binding affinities and give another level of complexity to antibody regulation (3–11).

Prior to the secretion of antibodies (the soluble form of the BCR), B cells must differentiate into ASCs (the effector cells of an ongoing antibody response). Plasmablasts are the main ASCs in the peripheral blood and are thought to be terminally differentiated and short-lived (12, 13). These cells are present at a low level in the peripheral blood of healthy humans and undergo significant expansion 7 days after immunization or infection (14, 15). The vaccine-induced population is predominantly antigen-reactive and has been extensively utilized as a way to characterize the dynamics of an ongoing B cell immune response (14, 16). This population is also a common source for therapeutic and prophylactic human mAb therapies (17). The majority of plasmablasts induced by the inactivated vaccine express the IgG...
receptor isotype (14), although some vaccine-specific cells produce IgA (18). This isotype bias is an inverse of the steady-state plasmablast population, which does not bind influenza and is primarily comprised of cells expressing the IgA receptor isotype (19). These isotype preferences exist due to the functional roles of these populations, with IgG antibodies specializing in antigen clearance through engagement with Fc receptors (20), and IgA antibodies more commonly reducing antigen immunogenicity through neutralization and sequestration (21). Although CSR alters the functional capacity of antibodies, it remains unknown if the process alters the cellular identity of the plasmablast beyond changes to the BCR transcript itself.

The steady-state IgA-secreting plasmablast population is maintained in the peripheral blood during an ongoing vaccine response (19). How these cells are related to and differ from the antigen-induced cells remains unclear. The peripheral blood steady-state plasmablasts are hypothesized to recirculate from mucosal niches (19) where they secrete antibodies that help maintain homeostasis of the microbiota (22). The recent discovery of shared IgA antibody sequences in both the serum and the gut further supports the connection between steady-state circulating IgA-secreting cells and mucosa-localized cells (23). This possible shared origin, combined with the enhanced neutralization capacity of IgA antibodies, makes them an ideal target population for vaccines aimed to induce localized mucosal protection (24–29). This current goal in vaccine development is based on the hypothesis that localized protective antibodies would prevent initial infection of the host, reducing the pathogen burden and the need for neutralizing serum IgG antibodies. Thus, further characterization of IgA plasmablasts, in particular steady-state versus antigen-reactive plasmablasts, is of great interest.

Transcriptional profiling is a powerful way to characterize cellular identity, as it gives insight into all ongoing cellular processes. Here we present Spec-seq, a framework to allow for the simultaneous exploration of the immune transcriptome, receptor repertoire, and receptor functional characteristics at the single-cell level. As a first application of this approach, we address the long-standing questions of how BCR isotype or specificity impacts the B cell transcriptome, and if these features direct functional specialization outside of the BCR itself. We employed Spec-seq to analyze human plasmablasts after influenza vaccination to characterize transcriptional differences dictated by BCR isotype and vaccine reactivity. Specifically, we examine how IgA vaccine-positive (IgAVP) peripheral blood plasmablasts compare with IgG vaccine-positive (IgGVP) and IgA vaccine-negative (IgAVN) populations. Our analysis reveals enhanced transcriptional similarity within clonally related plasmablasts, perhaps due to original B cell activation event imprinting. We also identified transcriptional subpopulations correlated with BCR vaccine binding. Differentially expressed genes between these populations included glycotransferase enzymes previously credited with altering the glycosylation of antibodies, thereby altering their immunogenicity, as well as proteins that regulate cell cycling. These data suggest IgGVP and IgAVP plasmablasts are largely similar, whereas IgAVN cells appear transcriptionally distinct from traditional terminally differentiated, antigen-induced peripheral blood plasmablasts.

Results
The Spec-seq framework (Supplemental Figure 1A; supplemental material available online with this article; https://doi.org/10.1172/JCI121341DS1) was developed to explore plasmablast transcriptional diversity in the context of BCR repertoire and antigen specificity. To address our specific questions of interest, IgA and IgG peripheral blood plasmablasts were single-cell sorted from 6 donors (Supplemental Figure 1D) 7 days after influenza vaccination (Figure 1, A and B). As expected with an intramuscular inactivated vaccine, there was a wide range of plasmablast expansion per donor (1.6%–26.7% of total B cells) (Supplemental Figure 2A), with the majority of plasmablasts expressing IgG (flow cytometry 51%–82%, ELISPOT 62%–85%) (Figure 1C). The total frequency of vaccine-binding plasmablasts per donor ranged from 48%–91% (Figure 1C). Monoclonal antibodies (mAbs) were cloned and generated from each plasmablast and screened by ELISA for binding to the administered vaccine and the viral components of the vaccine (Supplemental Figure 2, B–D). The mAb binding data combined with the receptor sequence information allowed each cell to be binned into 1 of our 3 groups of interest: IgAVP, IgGVP, and IgAVN (Supplemental Table 1). IgGVP mAbs had a slight increase in relative binding affinity, reported as area under the curve, compared with IgGVP mAbs (Supplemental Figure 2E). It is important to note that a minor fraction of the IgAVN cells may be vaccine-induced but were not identified as such because it is well-appreciated that it is impossible to assess all biologically derived vaccine epitopes (30). We believe the majority of these IgAVN cells are derived from the steady-state IgA plasmablast population, which is present in the peripheral blood of humans and is maintained after influenza vaccination (19). Unfortunately, the low frequency of IgG plasmablasts present in the steady-state blood and the huge induction of this population following vaccination prevented us from including a vaccine-inexperienced IgG population for comparison. These samples were then subjected to deep single-cell RNA sequencing (scRNA-seq). We focused our subsequent analyses on 291 plasmablasts that passed quality control filters (Supplemental Figure 3) with 11,895 unique genes detected in at least 5% of the cells.

The BCR isotype classification for each cell was validated on 3 levels: flow cytometry, Sanger sequencing during BCR cloning, and the BCR contig assembled from the scRNA-seq data (31) (Figure 1D). The scRNA-seq data were normalized (Supplemental Figure 3) and visualized using the t-Stochastic Neighbor Embedding algorithm (tSNE), which projects high-dimensional data into a 2-dimensional space and groups cells with similar transcriptomes together (32). As expected, after proper normalization and batch correction, clustering is not associated with experimental batch or donor (Supplemental Figure 3, H–K). Instead, a trend was revealed for the IgAVP and IgAVN cells to cluster together and away from the IgGVP population (Figure 1E), suggesting a plasmablast BCR isotype-dependent signature.

Repertoire analysis found limited differences between the 3 populations of interest, and overall did not suggest distinct cellular compartments of origin (Supplemental Figure 4 and Supplemental Tables 1 and 2). We did observe a significantly increased prevalence of the VH1 gene within the IgGVP population; how-
ever, this is difficult to interpret without repertoire from an IgGVN population. Furthermore, we also observed a substantial increase in the frequency of IgAVN cells versus IgAVP cells utilizing the IgA subtype 2 constant domain (IGHA2), which has been reported as more prevalent within mucosa-localized antibodies (33). Although it is difficult to draw conclusions from repertoire studies of this scale, the enhanced IGH A2 usage further supports a connection between the peripheral blood IgAVN population and the mucosal IgA ASC population.

Further repertoire analysis revealed evidence of clonotype expansion and identified 100 of the total 291 plasmablasts as members of 29 clonal expansions (Figure 2, A and B). The high frequency of clonal expansion is not uncommon after influenza vaccination (14). The clonal families ranged in size from 2–13 detected members and were present in all 3 populations of interest, with 3 unique vaccine-positive clones that span the IgG and IgA compartments (clones 4, 13, and 21). Within the 3 donors where these shared clones were detected, they were found at a 10%–20% frequency, which is similar to what we detected with independent high-throughput repertoire sequencing studies (16.5%–25.4%) (Figure 2, A and C). Clonal expansions containing cells of different isotypes have been previously reported (34, 35), and the tendency for BCR sequences to cluster within them by isotype suggests early CSR divergence before continued affinity maturation (Figure 2B). No difference in the relative binding affinity of IgG or IgA clonal family members was observed in our data, although the significance of this analysis is limited by the number of clonal expansions identified (data not shown). This emphasizes the importance of exploring transcriptional differences between IgGVP and IgAVP plasmablasts, as it increases the possibility of separate transcriptional, or functional, identities beyond the BCR. No BCR overlap was identified between the vaccine binding plasmablasts and the IgAVN population. These data suggest that only the vaccine-binding plasmablasts share cellular ancestors, which furthers our interest in characterizing the unique qualities of the IgAVN plasmablast population.
The identification of clonal expansions within the IgAVN population was surprising due to the expansive combinatorial diversity of the BCR, and we have 2 potential explanations for this clonal expansion. The first is that the soluble activation signals occurring during antigen-specific plasmablast activation were sufficient for BCR-independent B cell activation and subsequent proliferation of these cells. Alternatively, the expansion of IgAVN clones could support a more restricted antigen specificity within the IgAVN population than previously expected, resulting in an overall less diverse repertoire. Both of these hypotheses require further exploration in larger clonal data sets.

This unique data set allowed us to ask if B cells from the same clonal family have transcriptional profiles that are more alike than unrelated B cells. When clonally related B cells were labeled in the tSNE projection, clusters appeared (Supplemental Figure 5A). To quantitatively evaluate this observation, the Pearson correlation coefficient was calculated for each pair of cells in a clonal family as a way to assess overall transcriptional similarity. The average correlation between clonal family members within the vaccine-positive population is significantly higher than that of the unrelated vaccine-positive B cells (Supplemental Figure 5B), indicating increased transcriptional similarity within clonal families. This trend was not detected in the vaccine-negative IgA population (Supplemental Figure 5B). Since B cell clones share the same BCR transcripts, we repeated our analysis in the absence of Ig genes. For clonally related samples in both the vaccine-positive and vaccine-negative compartments, correlations are reduced upon removal of Ig genes (Supplemental Figure 5C), whereas unrelated plasmablast correla-
clones also displayed increased transcriptional similarity, with plasmablasts are biologically significant. As a final measure, we the remaining differences between IgA versus IgG class-switched 6C). However, our sample size was insufficient to determine if all Ig genes still suggests a trend for IgG plasmablasts to cluster together and away from IgA plasmablasts (Supplemental Figure 6C), the observed clustering was appreciably dimin-
ished, identifying the BCR genes as sufficient to drive transcrip-
tional separation. Notably, the tSNE projection after exclusion of Ig constant domain genes (Figure 3C), or all Ig genes (Supplemen-
tal Figure 5D), the latter improvement is not surprising, as diverse BCR transcripts should, in principle, increase the heterogeneity of unrelated B cells. Nevertheless, without Ig genes the average correlation within vaccine-positive clones is still significantly greater than that of unrelated vaccine-positive B cells (Figure 2D), and related B cells still cluster together (Figure 2E). Taken together, although BCR transcripts contribute to transcriptional similarity between clonal plasmablasts, the enhanced transcriptome similarity extends beyond Ig transcripts. We hypothesize this could be due to the fact that plasmablasts within a clone will have similar cellular origins and activation experiences (such as T cell help) that may durably imprint on the transcriptome and be passed on to clonal progeny. It is intriguing to consider that these potential similarities might drive even single pairs of related B cells to have unique transcriptional profiles.

We then hypothesized that beyond just secreting a particular class of antibody, IgG-expressing plasmablasts might be special-
ized for tissue immunity and plasmablasts secreting IgA special-
ized for mucosal immunity. As a means to screen for functional specialization of influenza-induced plasmablasts based on dis-
tinct BCR isotype, the gene expression profiles of IgAVP and IgGVP populations were compared. Visualization of these data indicated separate clustering of the 2 populations (Figure 3A). The most significant differentially expressed genes (DEGs) identified were the genes encoding the BCR constant domains themselves (Figure 3B). In fact, when each single cell was colored according to the expression level of IGHG1 or IGHG1, the clusters aligned perfectly with the assigned isotype annotation (Supplemental Figure 6A and B). When this analysis was repeated after exclusion of the Ig constant domain genes (Figure 3C), or all Ig genes (Supplemen-
tal Figure 6C), the observed clustering was appreciably dimin-
ished, identifying the BCR genes as sufficient to drive transcriptional separation. Notably, the tSNE projection after exclusion of all Ig genes still suggests a trend for IgG plasmablasts to cluster together and away from IgA plasmablasts (Supplemental Figure 6C). However, our sample size was insufficient to determine if the remaining differences between IgA versus IgG class-switched plasmablasts are biologically significant. As a final measure, we determined that clonal expansions consisting of both IgG and IgA clones also displayed increased transcriptional similarity, with and without Ig gene inclusion (Supplemental Figure 5E), indicat-
ing that plasmablasts arising from the same progenitor do not differentiate further after class switching to IgG or IgA. From these analyses we concluded that, besides the BCR genes themselves, peripheral blood IgG or IgA plasmablasts induced by influenza vaccination have highly similar transcriptional profiles. Addition-
ally, the slight difference in IgAVP and IgGVP affinity likely does not account for transcriptional differences.

To determine if differences exist between antigen-induced and vaccine-negative plasmablasts, 2 analyses were performed. The first was to compare the IgAVP and IgAVN samples in isolation. The tSNE projection of the IgA cells suggests transcription-
al segregation between these 2 populations (Figure 4A), which could indicate different functional attributes. A small fraction of the IgAVN cells remains intermingled within the IgAVP cells. The relative binding affinity of these cells does not correlate with their clustering and we hypothesize these are likely to be vaccine induced but specific for epitopes not assessed by our experiments. In total, 62 DEGs with an adjusted P ≤ 0.05 and a fold change ≥ 1.5 were identified between the IgAVP and IgAVN populations (Figure 4B). The significantly increased expression of IGH2A, VH3-23, VH3-7, and IGGC within the IgAVN popu-
lation, and enrichment of IGLJ2 in the IgAVP population, are confirmed by increased total repertoire gene usage frequencies (Figure 4, B and C). To assess the impact of these genes on overall population clustering, the tSNE analysis was repeated without any Ig genes, and a similar level of clustering was identified (Figure 4D). This indicates that transcriptional differences exist beyond repertoire usage. Interestingly, the 2 most significant DEGs identified between IgAVP and IgAVN have been previ-
ously associated with differential BCR glycosylation, FUT8 and B4GALT1 (Figure 4, B and E).

Our second analysis compared the total vaccine-positive samples with the IgAVN population, after exclusion of all Ig genes. The tSNE projection confirms separation based on specificity, showing a preference for the IgGVP and IgAVP cells to cluster together and away from the IgAVN cells (Figure 5A). A larger number of DEGs were identified in this comparison, including FUT8 and B4GALT1 (Figure 5, B and C). Differential glycosylation of antibodies is known to alter immunogenicity, and both FUT8 and B4GALT1 have been

**Figure 3. BCR constant domain transcripts are sufficient to cause clustering by isotype.** (A) tSNE projection of vaccine-positive cells only (n = 190). (B) Volcano plot between IgGVP and IgAVP cells. Genes with fold change less than or equal to 1.5 with a Benjamini-Hochberg adjusted P ≤ 0.05 are colored red. (C) tSNE projection of vaccine-positive cells (n = 190) after Ig constant genes are excluded.
The FUT8 enzyme is required for the addition of fucose moieties to the BCR glycans (3), whereas the B4GALT1 enzyme is necessary for the addition of galactose moieties (8). These modifications are known to shape a humoral immune response through alterations to Fc-receptor binding and complement activation (3–11). The differential expression of these enzymes between vaccine-positive and vaccine-negative cells suggests plasmablast BCR specificity directs transcriptional differences that can ascribe key effector properties to the antibodies secreted by that cell.

**FOXP1** was also differentially expressed between total vaccine-binding and vaccine-negative plasmablasts ($P = 2.06 \times 10^{-15}$ and fold change of 1.49, Figure 5B and Figure 6A), as well as between the IgAVP and the IgAVN cells ($P = 1.96 \times 10^{-5}$ and fold change of $-0.98$, Figure 6B). However, the measured fold change fell below our arbitrary 1.5 fold change requirement to be identified as a DEG of interest. The transcription factor FOXP1 is known to prevent terminal differentiation into antibody-secreting cells through repression of **PRDM1, IFR4**, and **XBPI**, the master transcription factors of the plasmablast fate (36). The increased expression of **FOXP1** within the vaccine-negative plasmablast population may suggest upregulation of pathways designed to prevent terminal differentiation into traditional antigen-induced plasmablasts. To explore this further, we performed transcriptional cell-cycle analysis, as G1 phase arrest is a feature of terminally differentiated antibody-secreting cells (12, 13). This analysis revealed an increased frequency of vaccine-positive cells assigned to the G1 phase, while the IgAVN compartment displayed increased assignment to the S phase (Figure 6C). This trend was also detected between IgAVP and IgAVN cells (Figure 6D). These data suggest that the vaccine-negative plasmablasts may have prevented terminal differentiation through upregulation of **FOXP1** to allow for continued cellular division, perhaps as a mechanism for population renewal and maintenance.

---

**Figure 4. IgA plasmablasts cluster by their ability to bind the vaccine.**

- (A) tSNE projection of the IgA plasmablasts ($n = 176$), population indicated by color and relative binding affinity (area under the curve) indicated by dot size.
- (B) Volcano plot showing the differential expression of all genes between IgA vaccine-negative (IgAVN) and IgA vaccine-positive (IgAVP) plasmablasts. Genes with log fold greater than or equal to 1.5 with a Benjamini-Hochberg adjusted $P \leq 0.05$ (2-sided t test) are colored red. Some of the most significant genes are labeled.
- (C) Pie charts showing repertoire usage frequency between IgAVN and IgAVP plasmablasts. Increased usage of the lambda light chain was found within IgAVP cells ($P = 0.0192, \chi^2$) and increased IGHA2 usage was found within IgAVN cells ($P = 7.45 \times 10^{-5}, \chi^2$).
- (D) tSNE projection of the IgA plasmablasts repeated after the exclusion of Ig genes.
- (E) Violin plots of FUT8 and B4GALT1 between IgAVN and IgAVP plasmablasts, where the location on the y axis shows each cell’s gene expression level and the x axis distribution is randomly assigned to improve spot visualization.
To explore the population dynamics of IgA peripheral blood B cells, longitudinal BCR high-throughput repertoire sequencing was performed for 9 donors across 5 time points spanning 35 days. Analysis of these data identified IgA clones present 28 days prior to vaccination that are maintained within the peripheral blood over time (Supplemental Figure 7A). The average frequency of these clones present at all time points assayed ranged from 10%–20% of the total clonal diversity, and this increased to 20%–55% when considering clones present for at least a month (Supplemental Figure 7B). We think this second value is a more accurate representation, as the possibility of missing a single clone at a given time point is nontrivial. An increase in total clonal diversity is evident 7 days after influenza immunization, due to expansion of many influenza-specific clones that were not detected at the earlier time points. Overall, these data indicate that peripheral blood IgA cells display a level of clonal restriction likely related to antigen selection, and that these cells maintain their presence in the peripheral blood over time. We hypothesize that the capacity for continued cell cycling detected within the IgAVN population could explain this persistence within steady-state plasmablasts.

Altogether, these data have identified transcriptional differences between vaccine-positive plasmablasts and vaccine-negative IgA plasmablasts, which support the possibility of an antigen-induced gene expression profile that may provide a novel means to identify vaccine-specific cells directly from scRNA-seq data without the need for time-exhaustive receptor generation and testing.

Discussion
In this study, we applied Spec-seq to peripheral blood plasmablasts and found that BCR transcripts are a major source of population heterogeneity. However, our data identified unexpected levels of transcriptional specialization within plasmablasts, disproving the long-standing belief that these cells are simply terminally differentiated antibody-secreting cells with limited complexity.
of the single-cell transcriptome of clonally related B cells. Unex-
seq enabled us to perform what we believe is the first-ever analysis
mRNA for the BCR protein will be continuously synthesized and
response. The subsequent accumulation of reactive oxygen spe-
tation of antibodies, which initiates a strong unfolded protein
response. It will be exciting to explore this possibility further in
our understanding of heterogeneity within the humoral immune
activation events, which are passed on to subsequent clonal prog-
scriptional similarity beyond their shared Ig genes. We hypothesize
expectedly, this analysis suggested that clones have increased tran-
scriptional imprinting during original
activation events, which are passed on to subsequent clonal pro-
gen, and eventual cell death (37). Due to this specialization,
mRNA for the BCR protein will be continuously synthesized and
highly overrepresented. This high expression explains why vari-
ble V(D)J-C usage among B cells is responsible for a significant
amount of the plasmablast transcriptional heterogeneity. Our
finding that Ig constant domain gene expression is sufficient for
segregation of the IgG and IgA plasmablasts further emphasizes
this point. Overall, our analysis found no additional significant
changes in the gene expression profiles of cells expressing IgG
versus IgA BCRs, and therefore suggests that prior CSR does not
specifically induce long-term functional specialization within
peripheral blood plasmablasts.

We found that differential expression of $B4GALT1$, $FUT8$, and
FOXP1 is directly related to BCR recognition of the vaccine. Spec-
cifically, an increased frequency of vaccine-positive cells express-
es high levels of the $B4GALT1$ enzyme, which is responsible for the
addition of galactose moieties to a conserved glycan in the BCR
Fc domain (8, 38). This modification increases the ability of IgG
antibodies to facilitate complement-dependent cytotoxicity (10),
while its presence on IgA antibodies enhances antibody-depen-
dent cell-mediated cytotoxicity (ADCC) through improved bind-
ing affinity with the FcγRIIIa (11). This association between antibody
vaccine recognition and galactose modification is supported in
the literature, where galactose moieties have been reported to
increase after influenza vaccination within the antigen-specific
serum antibody compartment when compared with the steady-
state serum profile (9). In contrast, the vaccine-negative plasmab-
lasts displayed enhanced expression of the $FUT8$ and $FOXP1$
gen. The $FUT8$ enzyme mediates the addition of fucose moie-
ties to BCR glycan (3). This Fc modification has been shown to
reduce IgG affinity for the FcγRIIIa, which limits antibody immu-
nogenicity by reducing ADCC (3, 4). Although this modification is
prevalent on steady-state antibodies of both IgG and IgA isotypes
(4, 5), it remains unclear how it specifically alters IgA Fc-receptor
binding (6, 7). Unfortunately, our mAb expression system does
not preserve differential glycosylation patterns, but the increased
expression of $FUT8$ suggests our IgAVN plasmablasts are produc-
ing antibodies with antiinflammatory properties.

The increased expression of $FOXP1$ within IgAVN plasmab-
lasts suggests blockade of terminal ASC differentiation, and this
is supported by the larger frequency of vaccine-negative plasmab-
lasts assigned to the S phase of the cell cycle, indicating DNA
replication and continued cellular division (12, 13). As the IgAVN
cells are hypothesized to be steady-state sentinel cells recirculat-
ing from mucosal sites, it is plausible that proliferation is neces-
sary to maintain peripheral blood population numbers. This pos-
sibility is further supported by our identification of steady-state
IgA B cell clones that are maintained over time in the peripher-
mal blood. Interestingly, we did not detect reliable expression of
gen genes for mucosal homing receptors within this population (data
not shown), which suggests that although peripheral blood ASCs
likely have a shared origin with mucosal ASCs (19, 22, 23), tran-
scriptional profiles primarily reflect the current cellular niche.
Together, these findings suggest that the IgAVN plasmablasts may
represent a new, distinct subset of peripheral blood plasmablasts,
one that performs a homeostatic role at the steady-state through
production of antibodies lacking the inflammatory properties
associated with antigen-induced antibodies and through prolifera-
tive renewal within this compartment. Further characterization

The full-length antibody sequences obtained through Spec-
seq enabled us to perform what we believe is the first-ever analysis
of the single-cell transcriptome of clonally related B cells. Unex-
pectedly, this analysis suggested that clones have increased tran-
scriptional similarity beyond their shared Ig genes. We hypothesize
that this could be due to transcriptional imprinting during original
activation events, which are passed on to subsequent clonal pro-
gen. The capacity for cells to inherit characteristics indicative of
their clonal origin and activation history would significantly alter
our understanding of heterogeneity within the humoral immune
response. It will be exciting to explore this possibility further in
data sets containing larger clonal expansions.

Although novel, our identification of BCR transcripts as the
dominant determinant of plasmablast subpopulations is not a
surprising result. These cells are highly specialized for produc-
tion of antibodies, which initiates a strong unfolded protein
response, the subsequent accumulation of reactive oxygen spe-
cies, and eventual cell death (37). Due to this specialization,
mRNA for the BCR protein will be continuously synthesized and

![Figure 6. FOXP1 expression in IgAVN cells blocks cell-cycle arrest in G1 phase. (A) Violin plot showing FOXP1 expression between total vaccine-positive (VP) and vaccine-negative (VN) plasmablasts. (B) Violin plot for FDG1P expression between IgAVN and IgAVP plasmablasts. (C and D) Box plots showing differences in frequency of each assigned cell-cycle stage between VP and VN plasmablasts (C) and IgAVN and IgAVP plasmablasts (D). Box plots display median cell-cycle score (***P = 0.0016, ****P = 0.00035, and ****P < 0.0001; 2-sided t test).](image-url)
of these antibodies to determine antigen specificity will be very interesting and could solidify a connection to the mucosa-localized, microbiota-regulating IgA ASCs. Our findings have significant implications for the field of vaccine immunology, as our data suggest that vaccine-induced peripheral blood plasmablasts are transcriptionally distinct from the vaccine-negative IgA population. Therefore, although inducing a larger frequency of vaccine-positive IgA plasmablasts in the peripheral blood following vaccination is predicted to result in increased global levels of secretory IgA, the field should continue to focus on other ways of inducing localized mucosal protection (26–29). The direct characterization of mucosa-localized ASCs will likely be instrumental in this process.

Overall, Spec-seq opens new doors for the characterization of transcriptional heterogeneity in the context of antigen specificity. This is particularly important for cells of the adaptive immune system where unique somatically rearranged receptors identify cells of distinct antigen specificities, which equates to distinct effector functions. Many aspects of B cell biology are linked to BCR specificity and this approach will be pivotal for advancing the B cell field.

Methods

**Human peripheral blood isolation.** Fifty healthy subjects were recruited to receive the 2014–2015 quadrivalent influenza vaccine (QIV) Fluarix (GSK) at the University of Chicago with approval by the institutional review board (protocol 09-043A). Peripheral blood (40–80 ml) was drawn 7 days after immunization. B cells were enriched from peripheral blood mononuclear cells (PBMCs) by the RosetteSep Human B cell Enrichment Kit (Stem Cell Technologies, catalog 15024), and treated to receive the 2014–2015 quadrivalent influenza vaccine (QIV) Fluarix (GSK) at the University of Chicago with approval by the institutional review board (protocol 09-043A). Peripheral blood (40–80 ml) was drawn 7 days after immunization. B cells were enriched from peripheral blood mononuclear cells (PBMCs) by the RosetteSep Human B cell Enrichment Kit (Stem Cell Technologies, catalog 15024), and treated and isolated with a Lymphoprep gradient (Corning, catalog 25-072). The 6 subjects with the highest vaccine-positive antibody-secreting cell frequency determined by ELISPOT were selected for inclusion in our study.

**ELISPOT.** B cell–enriched PBMCs (0.25 × 10⁶ to 1.0 × 10⁶ cells) were transferred onto ELISPOT plates (Millipore, catalog MSHAN4B50). Plates were coated with goat anti-human immunoglobulin (whole IgG, IgM, and IgA, KPL, catalog 01-10-07) at 5 μg/ml or with the total administered QIV at 10 μl/ml both in PBS overnight at 4°C and then blocked with complete media for 2 hours at 37°C before the cells were transferred. Plates with cells were incubated overnight at 37°C. The plates were washed and then incubated with biotinylated IgG (Southern Biotech, catalog 2040-08) and biotinylated IgA (Southern Biotech, catalog 2050-08) in separate columns, for each coated antigen type for 2.5 hours at room temperature. After washing, streptavidin–alkaline phosphatase (Southern Biotech, catalog 7100-04) was added for 2.5 hours at room temperature and NBT/BCIP (Thermo Fisher Scientific, catalog 34032) was added to each well for spot reveal.

**Cell sorting.** The remaining enriched peripheral blood B cells were stained for fluorescence-activated cell sorting (FACS) on either a BD FACS Aria II or the BD FACS Aria Fusion at the University of Chicago flow cytometry core. Antibodies used included: CD19 (BioLegend, catalog 30223, clone HIB19), CD27 (BioLegend, catalog 302829, clone O323), CD38 (BioLegend, catalog 303515, clone HIT2), IgG (BD Biosciences, catalog 550931, clone G18-145), IgA (Southern Biotech, catalog 2050-02, clone G1012-YG02B), and IgM (Southern Biotech, catalog 9022-09, clone UHB). After gating on lymphocytes, doublet exclusion was performed. The CD19⁺ B cells were visualized to estimate the total CD27⁺ CD38⁺ plasmablast (PB) response. Samples were required to have more than a 0.5% PB response for inclusion in the study. The IgM CD19⁺ CD27⁺ and CD38⁺ PBs were bulk sorted by IgG and IgA surface expression, prior to single-cell sorting into 96-well plates (BioRad) containing 4 μl of the capturing solution previously described (39). Plates were immediately sealed, placed on dry ice, and frozen at -80°C until use. Each plate contained 41 single cells, 6 unsorted wells, and 1 well with 100 cells.

**Simultaneous single-cell RNA sequencing and monoclonal antibody generation.** This section outlines the modifications required to merge the 2 preexisting protocols as well our optimized modifications to the SmartSeq2 protocol (39).

Here we would also like to draw attention to a few improvements that were made between the processing of the first 2 tagmentation plates, and the later 3 plates. PW2 and PW3 were processed with the SuperScript III reverse transcriptase enzyme. PW4–6 were processed with the Primerscript reverse transcriptase enzyme and 5’ biotin modifications were added to all 3 oligomers: Oligo-dT, TSO, and IS-PCR (for sequences and additional information see ref. 39). We believe all effects of these modifications on the transcriptome data have been removed with experimental batch correction.

**Cell isolation.** The catching solution used for single-cell sorting was identical to that described in the SmartSeq2 protocol (39), except for the oligomer biotin modification mentioned above to reduce undesired primer-to-primer annealing and amplification. FACS was used to isolate the cells of interest, and every sorted plate contained 41 single cells, 6 wells with no cell, and a single 100-cell well. The 100-cell wells were processed at the same time and with the same protocol as that used for the single-cell wells.

**cDNA reaction and amplification.** The plate to be processed was thawed on ice while the reverse transcription (RT) mixture was made in the RNA-only hood. For Primescript reactions, the RT mixture contained 100 U Primerscript (Clonetech, catalog 2680), 10 U RNase inhibitor (Clonetech, catalog 2313), 1M Betaine (Sigma Aldrich, catalog B0300), 6 mM MgCl₂, (Life Technologies, catalog AM9530G), and 1 μM 5’ biotin–TSO oligomer (IDT). The reaction volume was raised to 5.7 μl with nuclease-free water. When the SuperScript III enzyme (Invitrogen, catalog 180800) was used, 5 mM DTT (Invitrogen, catalog 180800) was also included, as suggested in the SmartSeq2 protocol. The Oligo-dT reaction was run prior to the addition of the RT mixture. Then the SmartSeq2 RT-PCR reaction was used for all plates regardless of the reverse transcriptase enzyme.

Once complete, the preamplification PCR mix was assembled on a sterile bench top as previously described (39). We selected 22 preamplification PCR cycles for our PB populations of interest, and amplification was evaluated on an Agilent high-sensitivity DNA chip (catalog 5067-4626) (Supplemental Figure 1B). All PB traces obtained during the course of the experiment contained 2 sharp amplicon peaks, suggestive of extremely high expression for certain transcripts, the smaller of which is likely the BCR itself. The preamplification PCR was as follows: denature at 98°C for 3 minutes, then 22 cycles of 98°C for 20 seconds, 67°C for 15 seconds, and 72°C for 6 minutes, followed by 72°C for 5 minutes before a hold at 4°C.

After bead-based PCR purification (39), 1 μl cDNA from each cell was used in 3 separate PCRs designed to amplify the heavy, kappa, and lambda chains of the BCR. The remaining cDNA was stored at -80°C until next generation sequencing library preparation.
Antibody PCR amplification and cloning. Antibody cloning of samples collected in 2014–2015 (PW2 and PW3) was performed as previously described by Smith et al. (16), whereas the samples collected in 2015–2016 (PW3–5) followed the revised protocol described in Ho et al. (40). No differences in antibody binding were observed with the 2 different cloning approaches (40).

Briefly, both approaches utilized separate sets of PCR to amplify the heavy chain, the kappa chain, and the lambda chain. The first PCR used a cocktail of primers designed to bind to all possible V gene and constant domains utilized by the respective antibody chain. The cocktail of primers used in the second PCR are nested inside the first primer binding sites to allow for additional amplification of the generated PCR product. Sanger sequencing identified the specific V(D)J or VJ combination used in each antibody chain. This information guided the primers for the cloning PCR, which incorporates either the necessary restriction digest sites for vector ligation (16) or the required sequence overhang for Gibson Assembly (40). Regardless of the cloning method used, each antibody heavy chain V(D)J was inserted into the same IgG1 backbone within the AbVec expression vector, whereas each light chain VJ portion was inserted into its respective kappa or lambda AbVec expression vector. All details of this approach are clearly outlined in the referenced papers (16, 40).

Library generation, pooling, and sequencing. The protocol for library generation proceeded according to the SmartSeq2 protocol. For our libraries, 0.5 ng amplified cDNA was used for the tagmentation reaction, and 12 PCR cycles were used for the indexing PCR. After the final round of bead-based PCR purification, the Agilent bioanalyzer high-sensitivity DNA chip was used to measure fragment size and concentration (Supplemental Figure 1C). Equal volumes of all libraries were then pooled into a single tube at 3 nM each and underwent final round of bead-based PCR purification, the Agilent bioanalyzer high-sensitivity DNA chip was used to measure fragment size and concentration for the anti-IgG plates, whereas the supernatant was used at 10 μg/mL. A horseradish peroxidase–conjugated anti-human IgG Fc-specific antibody (Mabtech, catalog 3820-4-250) was used for detection and the plates were developed with Super Aquablu ELISA Substrate (eBiosciences, catalog 00–4203-58). Absorbance was measured at a wavelength of 405 nm on a microplate spectrophotometer (BioRad). If a sample had an OD (minus background) greater than 0.5 on the anti-IgG plate once the positive control reached 1.0, it was considered a successful transfection. The samples on the influenza plates must also have had an OD greater than 0.5 when the positive control OD reached 3.0. The requirement for an antigen binding OD greater than 0.5 was based on previous work in our lab, where 0.5 was identified as 2 standard deviations above the mean absorbance of antibodies from naive B cells (42). Experimental replicates were performed for each antibody to confirm the results. Viral binding of mAbs was confirmed by ELISA. ELISAs were coated at 8 hemagglutination units (HAU) with either A/California/7/2009 H1N1, A/Texas/50/2012 H3N2, B/Massachusetts/2/2012 (B-Yamagata lineage), or B/Brisbane/60/2008 (B-Victoria lineage), and mAbs were added to each plate at 10 μg/mL, diluted 3-fold, and developed as described above. Area under the curve was calculated using prism with the limit of detection of 1.112 × 10^4.

Analyzing the isotype composition of clones from high-throughput sequence data. Clonal expansions containing both IgG and IgA plasmablasts were investigated using high-throughput plasmablast sequences isolated from 3 individuals 7 days after vaccination with the 2010–2011 trivalent influenza vaccine (43). Partis (44) was used to cluster BCR sequences into clonal expansions and infer each clone’s naive ancestral sequence. Briefly, Partis fits hidden Markov models to infer the most likely combination of V, D, and J genes for individual BCR sequences and compares model fit across different partitions of the data into clones to determine the most likely partition. Only productive sequences (in frame, without stop codons) for which Partis was able to infer the germline sequence were used for further analyses. Sequence isotypes were determined by using BLASTN against a custom database of constant region sequences obtained from the National Center for Biotechnology Information (NCBI). Isotypes were assigned based on the best-matching constant region sequence, provided the match had an E-value less than 10^-8. Using an E-value of 10^-5 produced nearly identical results (not shown). Sequences with the same E-value for different constant region sequences were considered to have an undetermined isotype. Matches involving the reverse complement of the query or reference sequences were disregarded. The frequency of clonal expansions with both IgG and IgA was calculated by dividing the number of such clones by the number of clones with more than 1 productive sequence. The maximum-likelihood tree for a representative clone was inferred from an alignment of the V, D, and J gene regions under a general time-reversible model with gamma-distributed across-site rate variation and empirical nucleotide frequencies. The alignment and the phylogenetic inference were performed using MACSE (45) and RaxML (46), respectively.
The Jornal of Clinical Investigation

Step 1: repertoire. Sanger BCR sequences were obtained from the nested PCR and the cloning PCR. Assembled BCR contigs were generated by applying the BASIC (31) algorithm to the raw transcriptome files for all single cells. The BCR sequences from PW2 and PW3 obtained by all 3 approaches were carefully compared previously (31) to confirm the accuracy of BASIC. In general, the assembled sequenc-
es provided by BASIC had less error and often reported a slightly
reduced mutation count. This could be due in part to the high number
of PCR cycles required to amplify the BCR. Therefore, we considered
the assembly by BASIC to be a more accurate report of the original
BCR sequence and used that contig, when available, for repertoire
analysis (Supplemental Tables 1 and 2). If BASIC was unable to assem-
ble a full-length sequence, the Sanger sequence obtained from the
cloning PCR was used in its place.

Clones were identified through the same usage of V and J
genes, as well as highly similar CDR3 sequences. When searching
for any repertoire differences between the 3 populations of interest
(Supplemental Figure 4), all clonal expansions were excluded, as
their inclusion could be confused with population gene preferenc-
es. However, when validating the identified DEGs, clonal expans-
ions were not eliminated, as those cells were present in the tran-
scriptome (Figure 4C).

Step 2: pseudoalignment and transcript quantification. Kallisto (48,
49) was used for pseudoalignment of the raw sequencing data against
the human transcriptome. The -quant command (with the -bias
option) was used to perform pseudoalignment of each sample’s data
against the human transcriptome. Expression quantification is reported as
transcripts per million (TPM), which is adjusted for transcript length
and library size, and facilitates direct comparison of gene expression
profiles across cells.

Step 3: Gene expression data frame, transformation, quality filters,
and gene filters. The data frame of gene expression was summarized
by summing each gene’s transcript variants prior to undergoing log,
(TPM+1) transformation. Next, all values below 0.4, our empirically
identified threshold of detection (Supplemental Figure 3B), were set to 0 and all genes whose expression was never above 0 were
removed from the data set. Our first cell-quality filter excluded cells
for which less than 40% of the total reads mapped to the human tran-
scriptome (Supplemental Figure 3C). Our next cell filter excluded
cells where the V, J, and constant domains did not match between
at least 2 of our 3 approaches: Sanger sequencing of the nested PCR,
cloning PCR, and the assembled BASIC sequence. This filter simulta-
neously confirmed that each sample contained only a single cell.

Step 4: Quantile normalization and batch correction. Quantile nor-
malization using the preprocess core package (50) was performed within
in each experimental plate to reduce differences among the quantiles
of cell-specific TPM counts. Next, the Limma package (51) was used
to remove known batch effects associated with experimental plate and
donor ID, while protecting our variables of interest — BCR isotype and
vaccine recognition (Supplemental Figure 3, H-K).

Step 5: Differential gene identification and visualization. Differen-
tially expressed genes were identified with a Student’s t test, where
samples were either grouped by BCR isotype or BCR vaccine reac-
tivity. We required a fold change greater than or equal to 1.5 and a
Benjamini-Hochberg adjusted P less than 0.05. Visualization of the
scRNA-seq data was obtained with the Rtsne package (52). Default
parameters were used, except theta = 0.001 and initial dims = 10. The
same parameters were used for all data sets.

Cyclone cell-cycle analysis. The R package Cyclone (53) was applied
to the single-cell transcriptome data to assign cell-cycle phases. This
package incorporates previously published gene sets associated with
G1, G2M, or S phases of the eukaryotic cell cycle. The input is the
Limma-adjusted gene expression data frame.

Statistics. A paired Student’s t test was used to assess significant
differences between the cell-cycle phase assignments (Figure 6, C and
D). To identify statistically significant repertoire differences, χ2 tests
were used (Supplemental Figure 4, A-D), in addition to 1-way ANOVA
(Supplemental Figure 4, E-H). To assess global transcriptional similar-
ity within vaccine-positive and -negative cells, we performed 2 quan-
titative steps (Figure 2D and Supplemental Figure 5B). First, we cal-
culated the Pearson correlation between cells within vaccine-positive
clones, and between vaccine-negative cells. Second, since correlation
coefficients are bounded in [-1, +1], we performed a Fisher’s Z tran-
formation [arctanh(p)] on the coefficients and subjected the trans-
formed values to a 1-sided, 2-sample Student’s t test. Specifically, we
tested if the average correlation between clonally related cells is sig-
nificantly greater than nonclonally related cells in various instances.

Data and code availability. Scripts used to generate the data in this
paper can be accessed at http://ttic.uchicago.edu/~aakhan/specseq.
The code for the analyses of clones from high-throughput sequence
data is available at https://github.com/cobeylab/isotype_phylogenetics.
All sequencing data are accessible in the Gene Expression Omnibus
database under accession number GSE11650.

Study approval. Written informed consent was received from all
subjects recruited at the University of Chicago for this study, with
approval of the Institutional Review Board (protocol 09-043A).
Author contributions

KEN carried out the majority of experiments and data analysis. JJG, NYZ, MET, and NJH generated monoclonal antibodies, and JJG, MET, and NJH performed ELISAs and analyzed data. MH performed antibody cloning. JL assisted in computational analysis of scRNAseq data. MCV, KK, and SC performed clonal analysis of the high-throughput repertoire sequencing data for Figure 2. KTR coordinated donor recruitment. NYZ, MC, and BP provided or processed samples for the IgA high-throughput repertoire sequencing studies. KTR, DS, and CLD assisted with sample collection and experiments. CH provided key guidance and experimental assistance. JJG, CH, CLD, and DS made significant contributions to manuscript editing. AAK directed the computational analyses. PCW conceived of the project, directed all experiments, and performed clonal analysis of high-throughput IgA repertoire sequencing studies. KEN, AAK, and PCW wrote the manuscript.

Acknowledgments

The authors thank Yunping Huang and Sarah F. Andrews for assistance with experiments, and the University of Chicago’s Flow Cytometry Facility, Comprehensive Cancer Center DNA Sequencing and Genotyping Facility, and Genomics Facility. Computational analyses were completed in part with resources provided by the University of Chicago Research Computing Center. This project was funded in part with federal funds from the National Institute of Allergy and Infectious Disease, National Institutes of Health grants U19AI082724 (to PCW), P01AI097092 (to PCW), U19AI099946 (to PCW), U19AI057266 (to PCW), HHSN272201400005C (to PCW), P30CA014599 (Genomics Facility), T32GM070281 (to DS), and T32AI007090 (to KEN). MCV and SC were supported by a Complex Systems Scholar Award from the James S. McDonnell Foundation (to SC). KK was supported by Centers of Excellence for Influenza Research and Surveillance grant HHSN272201400005C (to SC).

Address correspondence to: Patrick C. Wilson, Jules F. Knapp Building, 924 East 57th Street, R414, Chicago, Illinois 60637, USA. Phone: 773.702.9009; Email: wilsopn@uchicago.edu. Or to: Aly A. Khan, 6045 South Kenwood Avenue, Chicago, Illinois 60637, USA. Phone: 773.834.2500; Email: aakhan@ttic.edu.

27. Shaffer AL, et al. XBPI, downstream of Blimp-1, expands the secretory apparatus and other


