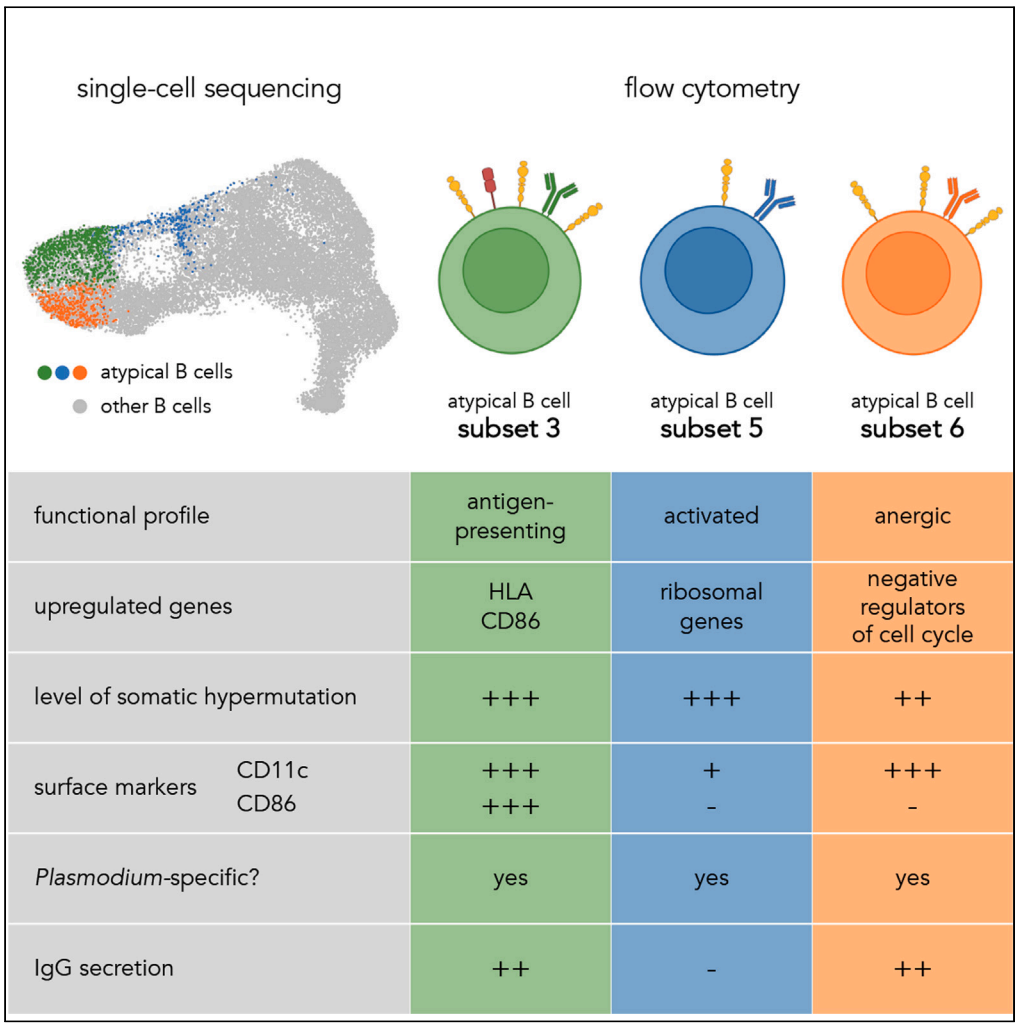
iScience

CellPress
OPEN ACCESS

Article

Atypical B cells consist of subsets with distinct functional profiles



functional profile		antigen-presenting	activated	anergic
upregulated genes		HLA CD86	ribosomal genes	negative regulators of cell cycle
level of somatic hypermutation		+++	+++	++
surface markers	CD11c CD86	+++ +++	+ -	+++ -
Plasmodium-specific?		yes	yes	yes
IgG secretion		++	-	++

Raphael A. Reyes, Gayani Batugedara, Paramita Dutta, ..., Sebastiaan Bol, Ferhat Ay, Evelien M. Bunnik

bunnik@uthscsa.edu

Highlights

Single-cell RNA sequencing reveals three subsets of atypical B cells

Their transcriptomic profiles point to different roles in the immune response

All three subsets are present in *Plasmodium*-exposed children and adults

Subsets differ in their ability to secrete IgG upon T cell help

Reyes et al., iScience *26*, 108496
December 15, 2023
https://doi.org/10.1016/j.isci.2023.108496

iScience

CellPress
OPEN ACCESS

Article

Atypical B cells consist of subsets with distinct functional profiles

Raphael A. Reyes,[1,10] Gayani Batugedara,[1,10] Paramita Dutta,[2] Ashley B. Reers,[1] Rolando Garza,[1] Isaac Ssewanyana,[3,4] Prasanna Jagannathan,[5,6] Margaret E. Feeney,[7,8] Bryan Greenhouse,[7] Sebastiaan Bol,[1] Ferhat Ay,[2,9] and Evelien M. Bunnik[1,11,*]

SUMMARY

Atypical B cells are a population of activated B cells that are commonly enriched in individuals with chronic immune activation but are also part of a normal immune response to infection or vaccination. To better define the role of atypical B cells in the human adaptive immune response, we performed single-cell sequencing of transcriptomes, cell surface markers, and B cell receptors in individuals with chronic exposure to the malaria parasite *Plasmodium falciparum*, a condition known to lead to accumulation of circulating atypical B cells. We identified three previously uncharacterized populations of atypical B cells with distinct transcriptional and functional profiles and observed marked differences among these three subsets in their ability to produce immunoglobulin G upon T-cell-dependent activation. Our findings help explain the conflicting observations in prior studies regarding the function of atypical B cells and highlight their different roles in the adaptive immune response in chronic inflammatory conditions.

INTRODUCTION

Humoral immunity to infection relies on the development of memory B cells that circulate in the blood and through secondary lymphoid organs to patrol for antigen, as well as long-lived plasma cells that home to the bone marrow and function as a source of antibodies that inhibit pathogen replication. In humans, memory B cells are defined based on the expression of the surface proteins CD21 and CD27. Upon acute infection or vaccination, other antigen-experienced B cell populations transiently increase in abundance in the circulation, including $CD21^{-}CD27^{+}$ activated B cells and $CD21^{-}CD27^{-}$ atypical B cells (also called DN2 cells).[1–3] These activated and atypical B cell populations are thought to be part of a normal immune response and peak in abundance 2 to 4 weeks after antigen exposure, followed by a gradual return to baseline over the course of several months.[1,4] However, under conditions of chronic or repetitive immune activation, as seen during HIV or *Plasmodium* infection, and in autoimmune diseases, such as systemic lupus erythematosus (SLE), atypical B cells accumulate and become a more prominent presence in the circulation.[5–8] In healthy adults, atypical B cells make up approximately 2%–4% of all circulating B cells, but in individuals with chronic inflammatory conditions this can increase to 20%–40%.[5,6,9] Whether these atypical B cells contribute to control of infections or negatively affect the host immune response remains incompletely understood. A better understanding of the function of atypical B cells will provide insight into the immune response during chronic infection and may present ways to overcome immune dysfunction or to harness these cells by vaccination.

Studies on the functionality of atypical B cells in the immune response have highlighted various differences between atypical B cells and memory B cells that are consistently observed among various inflammatory conditions. Specifically, atypical B cells display reduced expression of some B cell receptor signaling pathway genes and costimulatory molecules, as well as upregulation of inhibitory receptors.[9–13] In addition, atypical B cells were less responsive to B cell receptor engagement by soluble antigen and showed reduced differentiation into antibody-secreting cells under conditions that efficiently induced differentiation of memory B cells into antibody-secreting cells.[5,8–10,12] These results initially led to the conclusion that atypical B cells are exhausted and non-responsive to stimulation. However, it was recently reported that atypical B cells from *P. falciparum*-exposed individuals respond robustly to high-affinity membrane-associated antigen and showed a markedly

[1]Department of Microbiology, Immunology and Molecular Genetics, Long School of Medicine, The University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA
[2]Centers for Cancer Immunotherapy and Autoimmunity, La Jolla Institute for Immunology, La Jolla, CA 92037, USA
[3]Infectious Disease Research Collaboration, Kampala, Uganda
[4]Department of Infection Biology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
[5]Department of Medicine, Division of Infectious Diseases, Stanford University, Stanford, CA 94305, USA
[6]Department of Microbiology and Immunology, Stanford University, Stanford, CA 94305, USA
[7]Department of Medicine, University of California, San Francisco, San Francisco, CA 94110, USA
[8]Department of Pediatrics, University of California, San Francisco, San Francisco, CA 94110, USA
[9]Department of Pediatrics, University of California San Diego, La Jolla, CA 92093, USA
[10]These authors contributed equally
[11]Lead contact
*Correspondence: bunnik@uthscsa.edu
https://doi.org/10.1016/j.isci.2023.108496

reduced response to low-affinity antigens, a selective mechanism which may serve to limit activation of atypical B cells by low-affinity antigens, such as auto-antigens.[14] Despite these findings, the contribution of atypical B cells to the host immune response during chronic inflammation remains subject of debate and various functions have been proposed for these cells. These include (pre-)antibody-secreting cells, based on the presence of secretory immunoglobulin transcripts and the overlap in B cell receptor sequences between atypical B cells and plasma cells, and antigen-presenting cells, based on the upregulation of surface proteins that promote interactions between B cells and T cells.[9,11,15–19] The multiple and conflicting effector functions previously assigned to atypical B cells may suggest functional heterogeneity among this population.

Here, we used chronic *P. falciparum* parasite exposure as a model to define the heterogeneity of the atypical B cell compartment in humans. Over the course of many years of repetitive *P. falciparum* infections, people living in malaria-endemic regions develop an immune response that protects against disease. As a result, most cases of malaria occur in children under the age of 10, while adults with life-long exposure have asymptomatic infections. To study atypical B cells induced by *P. falciparum* infection, we performed single-cell transcriptomic analysis with cell surface marker and B cell receptor profiling on samples obtained from two children at two time points following malaria. Based on these sequencing data, we defined three subsets of atypical B cells with distinct transcriptional profiles that are likely to play different roles in the immune response. Using flow cytometry, we then analyzed the longevity, *Plasmodium* antigen specificity, and ability to differentiate into antibody-secreting cells of these subsets.

RESULTS

Single-cell transcriptomics identifies various B cell populations

The atypical B cell population is strongly expanded in the first few weeks after malaria and slowly contracts to baseline levels over the course of several months.[4] To study atypical B cells induced as a result of the hyperinflammatory immune response during malaria, we selected two children (5 and 7 years old) from whom peripheral blood mononuclear cells (PBMCs) were collected three weeks after malaria and six months later (Figure 1A; Table S1). Both children were free of symptomatic malaria between the two time points but had detectable *P. falciparum* parasitemia at both the three- and six-month time points (Figure S1).

We first isolated naive and antigen-experienced B cells by fluorescence-activated cell sorting based on the expression of CD21 and CD27 (Figure 1B). Naive B cells typically make up a large fraction of the total B cell pool (43%–60% in the samples used here) but are not of primary interest when studying B cells that have been activated in response to infection. Therefore, we enriched our samples for antigen-experienced B cells but added in some naive B cells to have representation of all B cell populations in the final cell pool. (Table S2). Each sample was then subjected to single-cell RNA sequencing (scRNA-seq) in combination with analysis of select cell surface marker expression and B cell receptor sequencing on the 10× Genomics platform. Cells that passed quality control (n = 21,439; Table S3) were clustered into eleven unique clusters based on gene expression. These clusters were then visualized by Uniform Manifold Approximation and Projection (UMAP) in a 2D plot based on similarity in transcriptional profile (Figure 1C). Cells from all four samples were well represented in all eleven clusters (Figure S2; Table S4). Comparing the expression levels of individual key genes (Figure 1D), as well as gene signatures of various B cell populations generated by others (Figure S3),[11,20,21] allowed us to identify the major B cell populations. This included transitional B cells in cluster 10 (expressing *CD10* and *LEF1*; Figure S4), naive B cells (expressing *IL4R*) in cluster 1, and memory B cells (expressing *CD27*) in cluster 8 (Figure 1D). In total, 83% of all cells in our dataset were found in clusters with a non-naive transcriptional signature, confirming that our dataset was enriched for antigen-experienced cells. Cells in clusters 3, 5, and 6 expressed genes associated with atypical B cells, such as *ZBTB32*, encoding a transcription factor, and *ITGAX*, encoding the surface protein CD11c (Figure 1D), as well as *TBX21*, *FCRL5*, *FCRL3*, and *CXCR3* (Figure S4). Together, these three clusters represented almost 30% of all cells, making this a rich dataset to study heterogeneity among atypical B cells in malaria-experienced individuals.

Phenotypically defined atypical B cells have multiple distinct transcriptomic profiles

We next sought to determine how well the atypical B cell clusters that we identified based on transcriptomic profiles corresponded to $CD21^{-}CD27^{-}$ B cells, which is the phenotypic classification for atypical B cells most commonly used in the malaria field. We therefore used cell surface markers CD21 and CD27 to "gate" cells, similar to how this is commonly done in flow cytometry (Figure S5). Nearly all cells in clusters 3 and 6 were $CD21^{-}CD27^{-}$, in agreement with their transcriptomic profile of atypical B cells (Figures 2A and S6). Additionally, $CD21^{-}CD27^{-}$ B cells made up about 50% of cluster 5. $CD21^{-}CD27^{-}$ B cells were also abundant in other clusters that did not contain atypical B cells. Cluster 10 comprised approximately 50% $CD21^{-}CD27^{-}$ cells, which were transitional B cells that have recently emigrated from the bone marrow and have yet to express CD21 (Figure S3A). Additionally, cluster 2 largely consisted of $CD21^{-}CD27^{-}$ cells that had not undergone class switching and were considered activated naive B cells (see in the following). Thus, using only the absence of CD21 and CD27 surface markers is likely not an ideal way to define atypical B cells.

In a recent study by Sutton et al.,[3] atypical B cells were observed to predominantly express CD11c and CXCR3. We therefore also analyzed the distribution of $CD11c^{+/-}$ $CXCR3^{+/-}$ populations among the eleven B cell clusters (Figure S7). The large majority of cells in clusters 3 and 6 expressed CD11c (75% and 98%, respectively), while expression of CXCR3 was observed in only 40%–45% of all cells in these clusters (Figures 2B and S6). In contrast, 79% of cells in cluster 5 were $CXCR3^{+}$, most of which also expressed CD11c. These percentages were similar when limiting the analysis to $CD21^{-}CD27^{-}$ B cells (Figure S8). These data therefore suggest that CXCR3 marks a subset of atypical B cells but cannot be used as a marker to identify all atypical B cells.

Next, we analyzed the heavy-chain sequences from the B cell receptor-sequencing libraries to determine isotype usage among the different clusters. This analysis showed that, as expected, the transitional cells in cluster 10 and naive B cells in cluster 1 were unswitched cells,

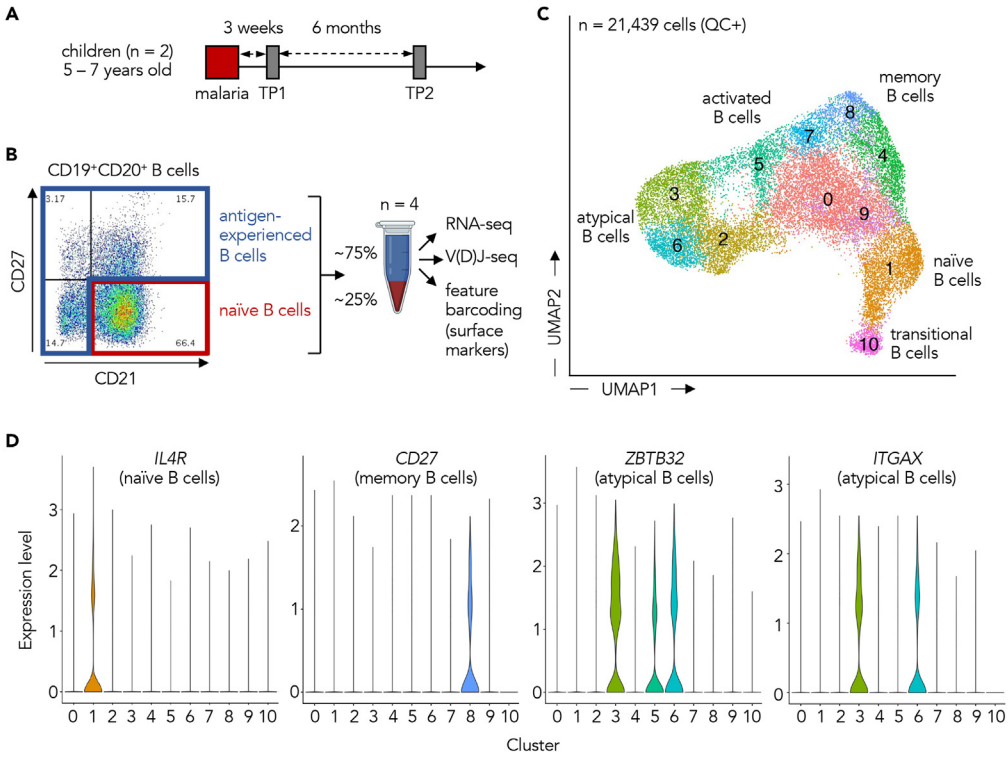


Figure 1. Single-cell sequencing of B cells in malaria-experienced children

(A) Timeline of sample collection at three weeks post-malaria (TP1) and six months later (TP2) in two children. Children did not experience malaria between the first and second sample collection.

(B) Flow cytometry strategy to enrich for antigen-experienced B cells. A majority of antigen-experienced B cells were mixed with naive B cells and used for library preparation using the 10× Genomics platform. This yielded three different libraries for each sample to study gene expression, V(D)J sequences, and cell surface marker expression in the same cells.

(C) UMAP of all B cells that passed quality control (QC+), showing 11 transcriptomic B cell clusters. Major subsets identified based on the expression of key genes and previously published gene signatures are indicated.

(D) Expression levels of genes characteristic for the various B cell populations. TP, time point. See also Figures S1–S4.

expressing only *IGHM* and *IGHD*. Atypical B cell clusters 3, 5, and 6 consisted of 34%–59% class-switched B cells (Figures 2C and S6). Cluster 3 in particular showed a distribution of *IGH* transcripts most similar to that of cluster 8 memory B cells, except for of a lack of *IGHA1* expression, suggesting that cluster 3 atypical B cells have undergone class-switch recombination at similar rates as memory B cells.

To further delineate the distribution of different B cell populations among our transcriptomic clusters, we identified B cells by phenotype based on the surface expression of CD21, CD27, CD11c, immunoglobulin (Ig) M, and IgG following recently published guidelines.[22] Because our cell surface marker analysis did not include IgD, we made two exceptions to these guidelines. First, unswitched B cells were not identified based on the presence of IgD but that of IgM, which is usually co-expressed with IgD. Some B cells express only IgM and lack IgD, but the distinction between IgM^+IgD^+ and IgM^+IgD^- could not be made in this analysis. Second, switched B cells were identified based on the presence of IgG instead of the absence of IgD, isotypes that are mutually exclusive. We first used CD27 and IgM or IgG expression to distinguish naive B cells (NAV, $CD27^-IgM^+$), unswitched memory B cells (unswM, $CD27^+IgM^+$), and switched memory B cells (swM, $CD27^+IgG^+$). These populations were further divided into resting ($CD21^+$) and activated ($CD21^-$). Resting unswitched memory B cells were mainly located in clusters 0, 7, and 9, while the majority of resting class-switched memory B cells were found in clusters 4 and 8. We also identified double-negative (DN) B cells ($CD27^-IgG^+$), which were divided into DN1 ($CD21^+CD11c^-$), DN2 ($CD21^-CD11c^+$), DN3 ($CD21^-CD11c^-$), and DN4 ($CD21^+CD11c^+$) (Figure S9). IgM^+ atypical B cells are also known as activated naive B cells, and IgG^+ atypical B cells are equivalent to DN2 B cells. We recently reported substantial differences in intrinsic properties between IgM^+ and IgG^+ atypical B cells that are likely related to their origin and function.[23] Additionally, CD11c has been identified as the marker that best delineates atypical B cells.[3,24] We therefore decided to use the phenotypic definition of DN2 B cells to identify atypical B cells for the remainder of this study: $CD21^-CD27^-IgG^+CD11c^+$ (Figure 2D). Of all phenotypically defined atypical B cells, 47% were found in cluster 3, 23% in cluster 6, and 17% in cluster 5 (Figure 2E). The majority of cells in cluster 3 (77%) and cluster 6 (75%) were atypical B cells (Figure 2E). Cluster 5 was more diverse with 23% atypical B cells, 19% activated naive B cells, 13% resting switched memory B cells, 16% activated switched memory B cells, and 29% other B cell populations. Of note, clusters 3, 5, and 6 were stable at a clustering resolution ≥ 0.6. At a clustering resolution below 0.6, cluster 6 was merged with cluster 2 that predominantly contained activated naive B cells (Figure S10). To focus our analysis on clusters with a large fraction of atypical B cells, a resolution of 0.6 was chosen for the analysis presented here. Collectively,

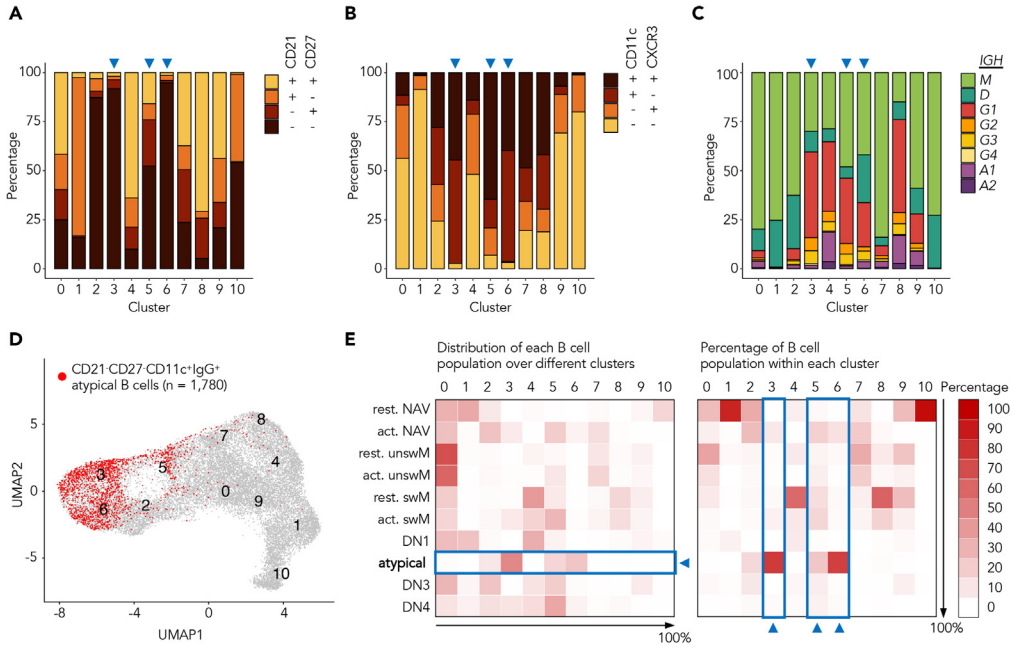


Figure 2. Phenotypically defined atypical B cells (CD21^{-}CD27^{-}IgG^{+}CD11c^{+}) are found in three transcriptionally distinct clusters

(A) Distribution of CD21$^{-/+}$ CD27$^{-/+}$ B cells (based on surface marker expression) among all transcriptional clusters. Clusters 3, 5, and 6 (indicated with blue arrow heads) contain atypical B cells.

(B) Distribution of CD11c$^{-/+}$ CXCR3$^{-/+}$ B cells (based on surface marker expression) among all transcriptional clusters.

(C) Distribution of immunoglobulin heavy-chain (*IGH*) transcripts among all transcriptional clusters.

(D) Projection of phenotypically defined atypical B cells (shown in red) onto the transcriptomics-based UMAP. All non-atypical B cells are shown in gray.

(E) Distribution of each B cell population over the different clusters, shown as a percentage of that population (left), and the percentage of phenotypically defined B cell populations within each cluster (right). The direction in which the values in the rows (left) or columns (right) add up to 100% is indicated with an arrow. Rest., resting; act., activated; NAV, naive, unswM, unswitched memory; swM, switched memory; DN, double-negative. See also Figures S5–S10.

these results show that phenotypically defined class-switched atypical B cells in malaria-experienced children consist of multiple distinct subsets with unique transcriptional profiles.

Differences in gene expression of atypical B cells are associated with different roles in the immune response

To better understand the identity of the three subsets of atypical B cells identified in this study, we analyzed the expression of genes that have previously been described as part of the unique transcriptional profile of atypical B cells.[10–12] The up- or downregulation of various genes in clusters 3, 5, and 6 as compared to resting switched memory B cells in cluster 8 was consistent with previous reports[10–12] (Figure 3A). Two manuscripts recently posted on bioRxiv showed that the transcription factor *ZEB2* acts as an important regulator for the development of atypical B cells.[25,26] Here, we found that expression of ZEB2 was significantly higher in clusters 3, 5, and 6 as compared to cluster 8 (Figure 3A). In contrast, several atypical B cell-associated genes were uniquely expressed in only one of the three clusters (Figure 3B). For example, the integrin-encoding genes *ITGB2* and *ITGB7* were highly expressed in cluster 3. In contrast, *CXCR3* expression was highest in cluster 5, in line with the results of the cell surface marker analysis (Figure 2B). Finally, *CD72*, encoding a surface marker that inhibits B cell responses and prevents antibody production,[27] was more abundantly expressed in cluster 6. These differences in gene expression highlight that bulk atypical B cells are a mixture of various subsets and point to functional differences in activation status and the potential for cell-cell interactions between these subsets.

To gain insight into potential differences in function between cells in clusters 3, 5, and 6, we performed Reactome pathway analysis using genes that were upregulated or downregulated in pairwise comparisons of the three clusters (Figures 3C and 3D; Tables S5 and S6). Atypical B cells in both clusters 3 and 6 upregulated genes in response to stimulation by interferon gamma and cytokines, but the outcome of this response differed between the two clusters. Atypical B cells in cluster 3 were enriched for pathways related to antigen processing and presentation, and immunoregulation (false discovery rate [FDR]-corrected p value $<10^{-4}$). Among the genes upregulated in cluster 3 were genes encoding proteins that promote interactions between B cells and T cells, such as human leukocyte antigen (HLA)-DR/DQ, CD86, and CD74, the latter of which regulates antigen presentation by class II major histocompatibility complex (MHC).[28–30] In addition, we observed upregulation of genes encoding class I MHC (*HLA-A* and *HLA-C*) and several genes that are associated with cytotoxicity, including granzyme M (*GZMM*), cystatin 7 (*CST7*), natural killer cell granule protein 7 (*NKG7*, a target gene of T-bet), and natural cytotoxicity triggering receptor

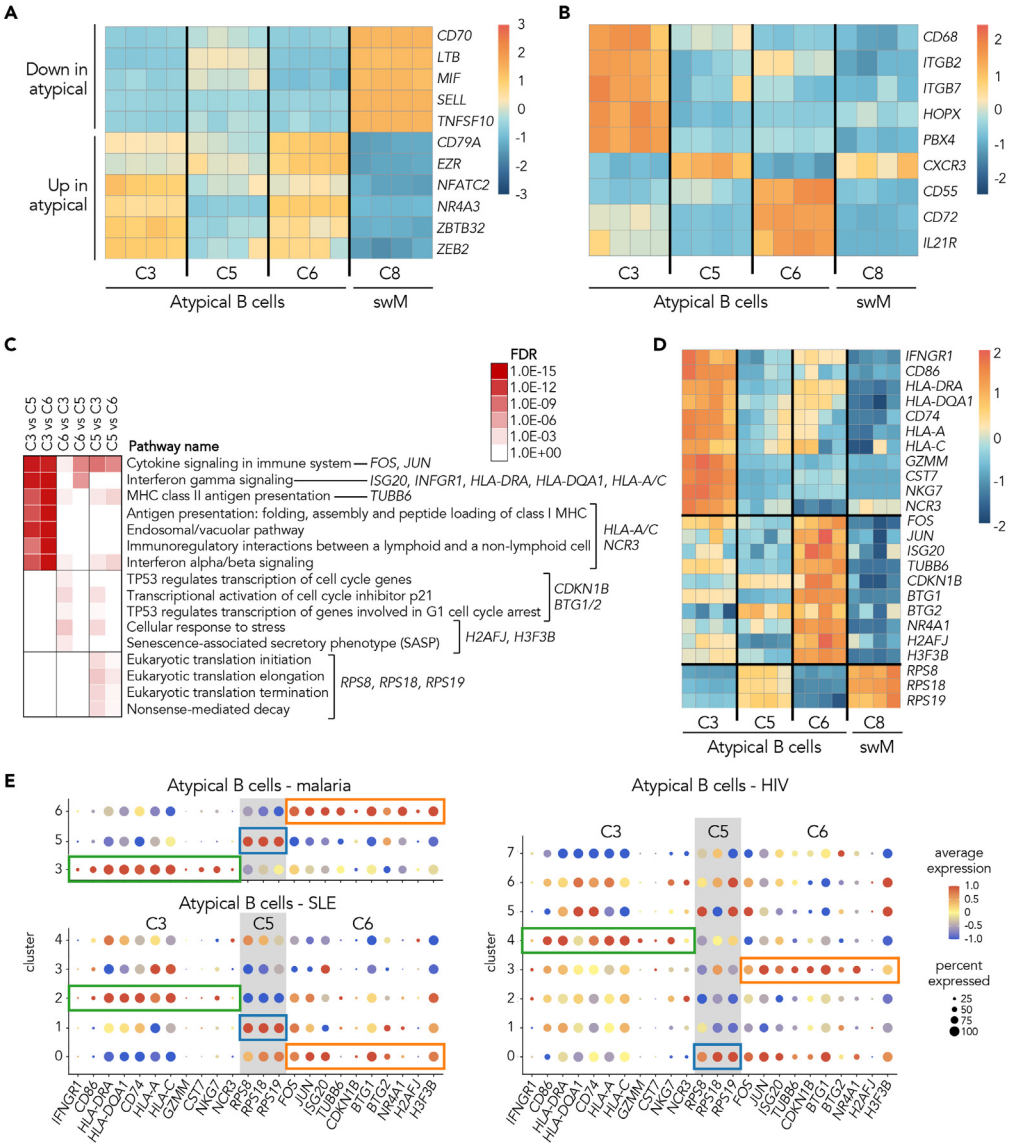


Figure 3. Atypical B cell subsets have different functional profiles

(A) Expression of genes previously identified as part of the unique transcriptional signature of atypical B cells that were consistently expressed 1.4-fold higher or lower in all three clusters of atypical B cells as compared to class-switched resting memory B cells (swM; cluster 8).

(B) Differential expression of atypical B cell-associated genes between the three subsets of atypical B cells.

(C) Selected Reactome pathways that were significantly enriched (FDR-corrected p value $<10^{-1}$) in pairwise comparisons between the three atypical B cell clusters. Key differentially expressed genes in these pathways shown in panel D are listed.

(D) Differentially expressed genes in atypical B cell clusters 3, 5, and 6 that were present in enriched Reactome pathways and are discussed in the text. The four columns in panels A, B, and D are ordered as follows: 3342 – time point (TP) 1, 3342 – TP2, 3119 – TP1, 3119 – TP2. C, cluster.

(E) Expression of genes associated with different functional profiles of atypical B cell clusters identified in this study (top left) in clusters of atypical B cells from SLE patients (bottom left) and HIV-infected individuals (right). See also Figures S11 and S12.

3 (*NCR3*). While it is unlikely that atypical B cells in cluster 3 are cytotoxic, upregulation of these genes could affect cell-cell interactions or antigen presentation.

On the other hand, atypical B cells in cluster 6 upregulated genes encoding components of the Activator Protein 1 (AP1) complex, including *FOS* and *JUN*, as well as interferon-stimulated genes (for example, *ISG20*), and tubulin genes (for example, *TUBB6*). Additionally, cluster 6 atypical B cells were enriched for pathways that signified negative regulation of cell cycle progression. Upregulated genes included the cyclin-dependent kinase inhibitor *CDKN1B*, the anti-proliferative genes *BTG1/2*, and the transcriptional repressor *NR4A1*. Nuclear receptor 4A1 is part of a family of transcription factors that was recently implicated in a negative feedback loop that renders B cells more dependent on T cell help and promotes peripheral B cell tolerance.[31,32] Cluster 6 atypical B cells also showed increased expression of variant histones,

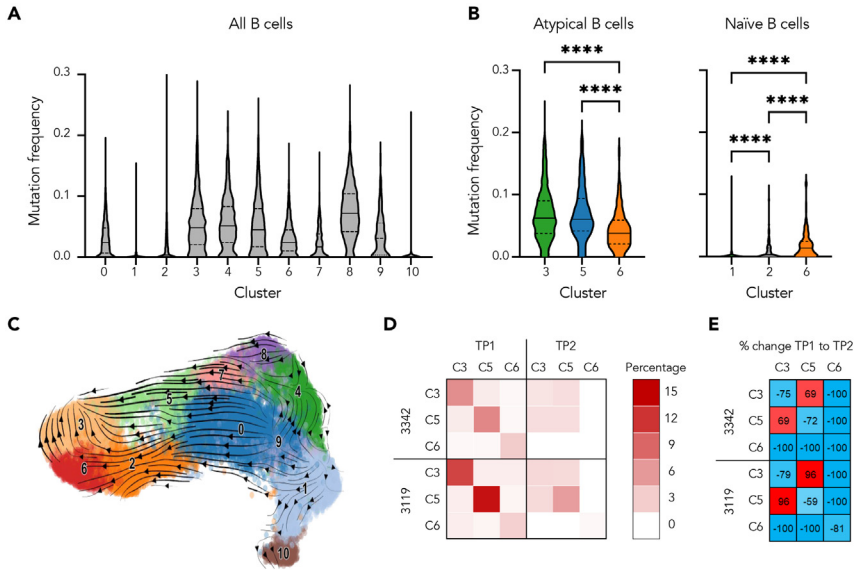


Figure 4. Developmental relationships between atypical B cell subsets

(A) Levels of somatic hypermutation among the transcriptomics-based B cell clusters.

(B) Levels of somatic hypermutation among phenotypically defined atypical B cells in clusters 3, 5, and 6 (left) and among phenotypically defined resting naive B cells in cluster 1 and activated naive B cells in clusters 2 and 6 (right). In panels A and B, the median is shown as a solid horizontal line, while the quartiles are represented by dashed horizontal lines. In panel B, a one-way ANOVA was used to test for statistically significant differences between the groups. p values shown are from Tukey's post hoc test after a one-way ANOVA. ****$p < 0.0001$.

(C) Results of the developmental trajectory analysis of B cells using scVelo.

(D) Clonal expansion within clusters (along the diagonal in each quadrant of the heatmap) and clonal connections between clusters (off-diagonal cells) for the three clusters that contain atypical B cells at 3 weeks post-malaria (TP1) and six months later (TP2).

(E) The percentage change in clonal expansion/connections between the first and the second time point for the three clusters that contain atypical B cells. TP, time point. See also Figures S13 and S14.

associated with a pro-inflammatory stress response resulting in cellular senescence.[33–35] The transcriptional profile of atypical B cells in cluster 6 thus pointed to an interferon-driven response that resulted in senescence or anergy.

Finally, atypical B cells in cluster 5 showed higher expression of 8 out of 99 genes encoding ribosomal proteins (for example, *RPS8*, *RPS18*, and *RPS19*) and were most strongly enriched for pathways related to translation. Ribosomal genes were also highly expressed in cluster 8 switched memory B cells, suggesting that this may be a feature of cells poised to undergo cell differentiation.

To determine whether atypical B cells that arise in other inflammatory conditions also consist of subpopulations with different functional profiles, we reanalyzed scRNA-seq data from SLE patients[36] and HIV-infected individuals.[37] Atypical B cells were extracted from these datasets (n = 3,487 for SLE and n = 55,179 for HIV) and clustered into 5 and 8 clusters, respectively (Figures S11 and S12). Among these were clusters with similar gene expression patterns as observed for the malaria-associated atypical B cells analyzed in this study (Figure 3E). Both datasets also had clusters that did not fit one of the three gene expression profiles identified in this study. This may point to additional heterogeneity among atypical B cells specific to these diseases. Collectively, our results indicate that phenotypically defined IgG^+ atypical B cells can have different transcriptional profiles that are associated with different functions within the immune response.

Atypical B cells form two separate compartments

To study the development of the different subsets of atypical B cells, we turned to an analysis of their B cell receptor sequences. We first determined the level of somatic hypermutation each B cell population has undergone by calculating the B cell receptor mutation frequency in each cluster. This revealed that cells in clusters 3 and 5 on average accumulated similar levels of mutations as class-switched memory B cells in cluster 8 (Figures 4A and S13A). In contrast, the average mutation frequency of cells in cluster 6 was almost 50% lower than that in clusters 3 and 5. This may in part be caused by the larger fraction of unswitched B cells in cluster 6 (Figure 2C). When we included only class-switched atypical B cells in the analysis, we still observed a higher average level of somatic hypermutation in clusters 3 and 5 (0.068) than in cluster 6 (0.043) (Figures 4B and S13B). While the level of somatic hypermutation in unswitched (activated naive) B cells in cluster 6 (0.019) was lower than that of the atypical B cells in cluster 6 (0.043), it was higher than that of their counterparts in cluster 2 (0.010), indicating that the activated naive B cells in cluster 2 could be a precursor population to both unswitched and class-switched B cells in cluster 6 (Figure 4B). Atypical B cells in clusters 3 and 5 had acquired similar levels of somatic hypermutation as memory B cells, which could indicate that they are derived from memory B cells or have a similar developmental trajectory. However, while somatic hypermutation and affinity maturation typically occur

inside germinal centers, these processes can also occur in extrafollicular sites,[38] and our results are therefore not direct evidence for the developmental pathway of atypical B cells.

Next, we sought to better understand the developmental relationship between the atypical B cells in clusters 3, 5, and 6. To this end, we first employed RNA velocity analysis to infer the developmental trajectories of atypical B cells. This analysis uses the ratio of unspliced and spliced transcripts to identify genes that undergo upregulation (more unspliced than spliced transcripts) or downregulation (more spliced than unspliced transcripts).[39] These transcript dynamics are then used to infer the potential transitional relationships among cells together with their directions of development.[40] This analysis revealed two dominant developmental trajectories: (1) the influx for cluster 6 atypical B cells were mainly from activated naive B cells in cluster 2, while (2) for cluster 3 atypical B cells this was predominantly from memory B cells (clusters 0/4/7/8/9), with cluster 5 atypical B cells possibly serving as an intermediate step in this trajectory (Figure 4C). To confirm these developmental connections between the atypical B cell subsets, we calculated the level of expansion of clonal B cell lineages within each cluster at the two time points, as well as the percentage of clonal connections between clusters, as described in more detail previously.[41] B cells were considered part of the same lineage when their heavy-chain variable region contained the same *IGHV* and *IGHJ* genes and shared at least 85% amino acid sequence identity in the heavy-chain (H) CDR3.[42,43] In all three atypical B cell clusters, there was strong B cell expansion at 3 weeks post-malaria, with 3%–14% of B cells within each cluster belonging to the same lineage (Figures 4D and S14A). This was followed by contraction six months later, resulting in a loss of more than 50% of clonal lineages (Figures 4D and 4E). This contraction suggests a return to a resting state and minimal immune activation as a result of the asymptomatic infection at 6-month post-malaria, in comparison to the strong immune response elicited by malaria, as also reported previously.[44] At 3 weeks post-malaria, 0.5%–1.0% of clonal sharing occurred between cluster 6 and clusters 3/5, but these connections were completely lost six months later. In contrast, the connectivity between clusters 3 and 5 was increased at the second time point as compared to the time point shortly after malaria, and this was the only pair of clusters for which this happened consistently (Figures 4E and S14B). These observations suggest that cluster 6 atypical B cells form a compartment separate from the other atypical B cell subsets, while cluster 3 and cluster 5 atypical B cells are more strongly connected, in line with the results of the trajectory analysis. Collectively, these results point to different developmental pathways for atypical B cells that are connected to different functional roles in the immune response.

All three atypical B cell subsets are found in malaria-experienced children and adults irrespective of parasite exposure levels

The identification of atypical B cell clusters by transcriptomics analysis was based on four samples from two children. To validate our findings in a larger number of individuals, we analyzed these cells by flow cytometry. In the single-cell transcriptomics analysis, we observed that the three clusters with atypical B cells could be distinguished by differences in expression level of *ITGAX* (encoding CD11c) and *CD86* (Figure 5A). We therefore sorted three subsets of atypical B cells based on expression of CD11c and CD86: subset 3 ($CD11c^{hi}CD86^{+}$), subset 5 ($CD11c^{int}CD86^{-}$), and subset 6 ($CD11c^{hi}CD86^{-}$). We will refer to these subsets with the number corresponding to the matching transcriptomics cluster, i.e., flow cytometry subset 3 for cells matching transcriptomics cluster 3 (Figures 5B and S15). These three subsets were then analyzed by bulk RNA sequencing and showed similar gene expression profiles as the three transcriptomics clusters identified by scRNA-seq, despite high variability between samples due to low input (100 cells per sample), and differences between the samples used for this analysis (adults, no recent malaria) and those used for the scRNA-seq experiments (children, after malaria) (Figure 5C).

Using spectral flow cytometry, we next studied the abundance of each atypical B cell subset and their longevity in the absence of antigen stimulation and infection-driven inflammation. For this analysis, we included samples from nine children and ten adults collected following known *Plasmodium* exposure and months to years later after a period with low *Plasmodium* exposure due to highly effective insecticide spaying[45,46] (Figure 5D; Table S1). B cells were phenotyped using a panel of 18 surface and intracellular markers that allowed us to identify the three atypical B cell subsets and analyze these for the expression of 12 markers previously shown to be positively (CD19, CD20, CD11c, FcRL5, CD83, CD86, CD95, CXCR3, T-bet) or negatively (CD24, CD38, CXCR5) associated with activated and atypical B cells.[22] To find the atypical B cell subsets identified in our transcriptomic analysis, we first gated manually on $CD19^{+}CD20^{+}CD21^{-}CD27^{-}IgG^{+}CD11c^{+}$ atypical B cells (Figure S16). Atypical B cells from all 38 samples were then plotted in a composite UMAP based on the expression of the 12 cell surface and intracellular markers listed previously (Figure S17). Next, we identified three distinct atypical B cell subsets by clustering these cells based on the expression of CD11c and CD86 using the unsupervised clustering tool FlowSOM (Figures 5E and 5F). We determined the abundance of these three subsets among non-naive B cells. In children, most atypical B cells belonged to subset 5 and subset 6 (Figure 5G). The abundance of the subsets did not change significantly between time points (Figure 5G), which could be the result of the timing of sample collection. At two months post-malaria in a low transmission setting, transient changes to the atypical B cell population induced by malaria may have already reverted to baseline. In adults, the first sample was collected during a period of high *Plasmodium* exposure (Figure 5D). The abundance of subset 3 atypical B cells among total B cells decreased after a period of minimal parasite exposure, while the abundance of subsets 5 and 6 did not change (Figure 5G). Interestingly, the distribution of atypical B cells over the three subsets was similar for adults and children during low parasite exposure at the second time point (Figure 5H). Of note, IgG^{+} B cells (~30%) were on average the most abundant isotype among $CD21^{-}CD27^{-}CD11c^{+}$ atypical B cells, but atypical B cells of other isotypes were also detected (Figure S18). Compared to $CD21^{+}CD27^{+}$ memory B cells, a larger percentage of atypical B cells were IgD^{+} (15% vs. 2%) and a smaller percentage were IgA^{+} (3% vs. 15%) (Figure S19). Collectively, these results demonstrate that we were able to distinguish three subsets of atypical B cells by flow cytometry and that the composition of the atypical B cell population changes in relation to parasite exposure.

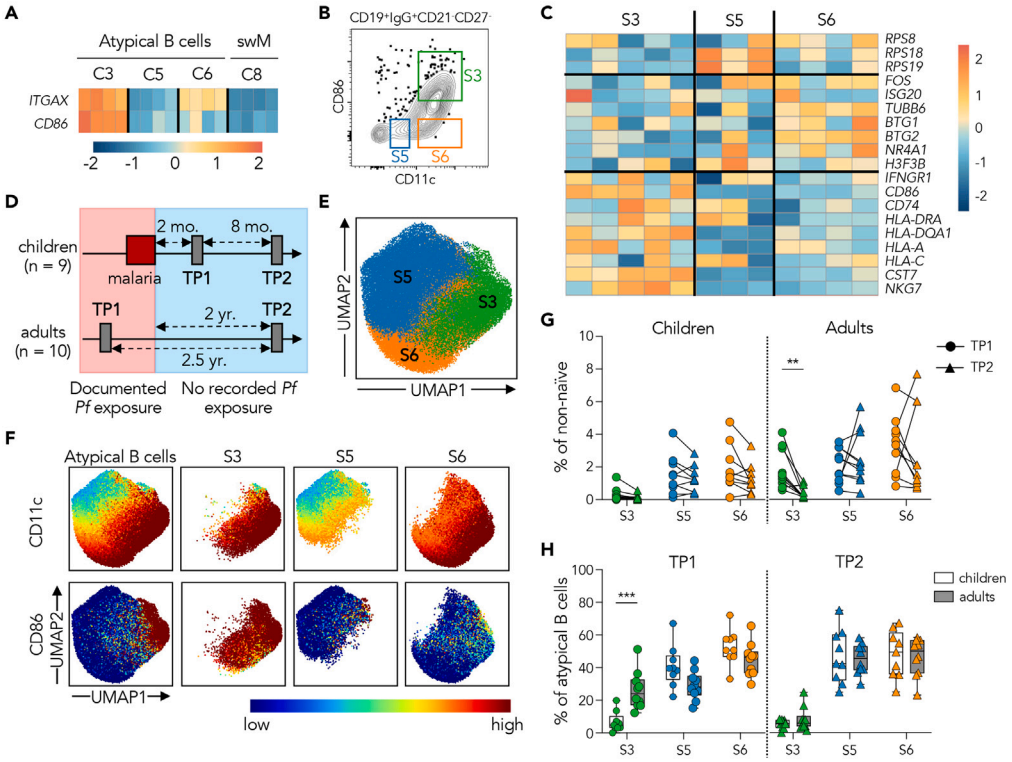


Figure 5. Dynamics of atypical B cell subsets in children and adults

(A) Gene expression profiles of differentially expressed genes encoding cell surface proteins CD11c and CD86 in atypical B cell clusters 3, 5, and 6, as well as resting switched memory B cells (swM, cluster 8) from the four scRNA-seq samples, used to distinguish atypical B cell subsets by flow cytometry.

(B) Gating strategy to sort atypical B cell subsets based on CD11c and CD86 expression.

(C) Expression of genes that underlie the different functional profiles of the three atypical B cell clusters (compare with Figure 3D) in flow-sorted atypical B cells subsets from malaria-experienced adults (n = 3–5 per subset, see STAR Methods).

(D) Schematic overview of sample collection for children and adults during known *P. falciparum* (*Pf*) exposure and in the absence of exposure to *P. falciparum*.

(E) Composite UMAP projection of atypical B cells from all 38 samples. Subsets were defined based on expression of CD11c and CD86 and were named according to the corresponding clusters identified by single-cell RNA-sequencing analysis.

(F) CD11c and CD86 expression projected onto the UMAP of all atypical B cells as well as each of the three subsets individually.

(G) Percentage of each atypical B cell subset among total $CD19^+CD20^+$ B cells in children (n = 9) and adults (n = 10) at the first time point (TP1) and second time point (TP2).

(H) Comparison of relative subset sizes between children and adults at each time point. Statistical analyses were performed using a Wilcoxon matched-pairs signed rank test for paired data (G) and a Mann-Whitney test for unpaired data (H) with a 10% false discovery rate using the two-stage step-up method of Benjamini, Krieger, and Yekutieli. $^{**}p < 0.01$; $^{***}p < 0.001$. See also Figures S15–S19 and S22.

Atypical B cells express higher levels of CXCR3 and CD95 in adults as compared to children

To map additional differences in phenotype between the three subsets of atypical B cells, we plotted the mean fluorescence intensity of CD21, CD27, and the 12 markers that were previously observed to be positively or negatively associated with activated or atypical B cells for all three subsets as well as IgG^+ resting memory B cells ($CD21^-CD27^-$). We observed that subset 3 had the most "atypical" surface marker expression pattern, with the highest expression of CD19, CD20, CD11c, FcRL5, CD95, and T-bet (Figures 6A and S20).

We were also interested in potential differences in phenotype between atypical B cells from *P. falciparum*-exposed children and adults. At the second time point, after months to years of no chronic immune activation, the distribution of atypical B cell subsets was similar between children and adults (Figure 5H). However, we noticed that their atypical B cells had a distinct localization when projected onto the UMAP (Figure 6B). Overlaying the expression of other markers in our flow cytometry panel onto the UMAP revealed that atypical B cells in adults extended into the UMAP area where cells expressed high levels of CXCR3 and CD95 (Figures 6C and 6D). To quantify this, we analyzed the percentage $CXCR3^+$ and $CD95^+$ cells in each subset. Compared to children, all atypical B cell subsets in adults showed more $CXCR3^+$ (Figure 6E) and $CD95^+$ cells (Figure 6F). This difference in CXCR3 and CD95 expression in atypical B cells between children and adults points toward additional heterogeneity among atypical B cells that may be connected to differences in immune status or cumulative parasite exposure levels.

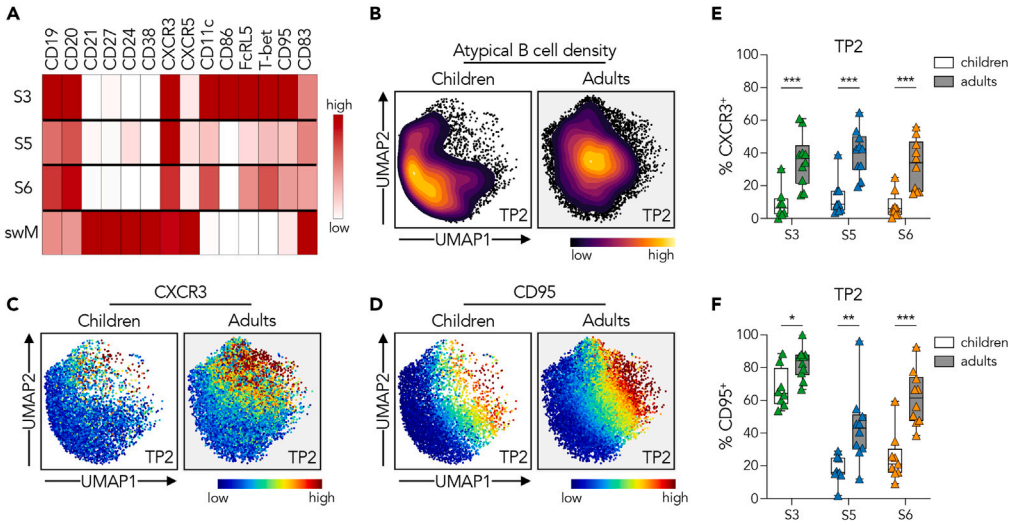


Figure 6. Expression of T-bet and CXCR3 in atypical B cell subsets

(A) Relative median fluorescence intensity of 14 surface and intracellular markers in subsets (S) 3, 5, and 6 and in resting switched memory B cells (swM) calculated using flow cytometry data from all 19 individuals.

(B) Density UMAP plots of atypical B cells detected at the second time point in children and adults.

(C and D) Projection of CXCR3 (C) and CD95 (D) expression in atypical B cells at the second time point projected onto the UMAPs for children and adults.

(E and F) Percentage of $CXCR3^+$ (E) and $CD95^+$ (F) cells within the atypical B cell subsets. For each subset, only individuals with >50 cells were included in the analysis. Statistical analysis was performed using a Mann-Whitney test with a 10% false discovery rate using the two-stage step-up method of Benjamini, Krieger, and Yekutieli. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. See also Figure S20.

All subsets of atypical B cells harbor *P. falciparum*-specific cells

To determine whether the different atypical B cell subsets develop in response to antigen during infection or through bystander activation of non-antigen-specific cells, we used antigen tetramers to detect B cells with specificity to two *P. falciparum* proteins: merozoite surface protein 1 (MSP1) and apical membrane antigen 1 (AMA1). Both antigens are known to elicit strong IgG responses in children and adults.[47] Since both MSP1 and AMA1 tetramers were labeled with the same fluorochromes, MSP1-specific and AMA1-specific B cells are together referred to as MSP1/AMA1-specific B cells (Figure S16). First, we overlayed the MSP1/AMA1-specific B cells onto the previously generated UMAPs of atypical B cell subsets 3, 5, and 6. Antigen-specific cells were identified in all subsets (Figure 7A). We then quantified the number of antigen-specific cells in each of the three atypical B cell subsets and observed that the number of MSP1/AMA1 B cells in each subset was proportional to its total size (Figure 7B). Finally, we analyzed CXCR3 and CD95 expression among MSP1/AMA1-specific atypical B cells in adults. Approximately 25%–40% of all MSP1/AMA1-specific atypical B cells were $CXCR3^+$, with no differences between the three atypical B cell subsets (Figure 7C). In contrast, nearly all MPS1/AMA1-specific atypical B cells in all three subsets expressed CD95, as compared to 60%–70% of all atypical B cells in subsets 5 and 6 (Figure 7D). Increased CD95 expression was also observed in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike-specific B cells in recovered COVID-19 patients[48] and suggests these cells recently underwent antigen-driven activation. Collectively, these results suggest that all three subsets of atypical B cells develop in response to antigen stimulation.

Antibody secretion upon receiving T cell help differs between the three atypical B cell subsets

Recently, Hopp et al. reported that bulk atypical B cells express higher levels of markers associated with T cell interactions than memory B cells.[18] With the help of activated T follicular helper cells, atypical B cells could be stimulated to secrete antibodies, although antibody concentrations in culture were much lower than for memory B cells.[18] Based on our observations that only atypical B cell subsets 3 and 6 expressed markers involved in T cell interactions, we tested which atypical B cell subsets can be induced to secrete antibody under T-cell-dependent activation. Using fluorescence-activated cell sorting (FACS), we sorted the three atypical B cell subset based on CD11c and CD86 expression, as well as switched memory B cells from eight adults with high *P. falciparum* exposure (Figure S15). Each B cell population was cultured with CD40L-expressing 3T3 cells in medium supplemented with interleukin (IL)-2 and IL-21 to mimic T-cell-dependent stimulation. IgG production was measured after 7 and 14 days. Samples from cultures lacking CD40L-expressing 3T3 cells were used as a negative control. In line with the expression of markers involved in T cell interactions, we found that subsets 3 and 6, but not subset 5, secreted IgG in the presence of T cell help. On day 7, the IgG concentration was lower than that produced by switched memory B cells (Figure S21), but this difference disappeared by day 14 (Figure 7E). These results suggest that most but not all atypical B cells can differentiate into antibody-secreting cells upon T-cell-dependent activation.

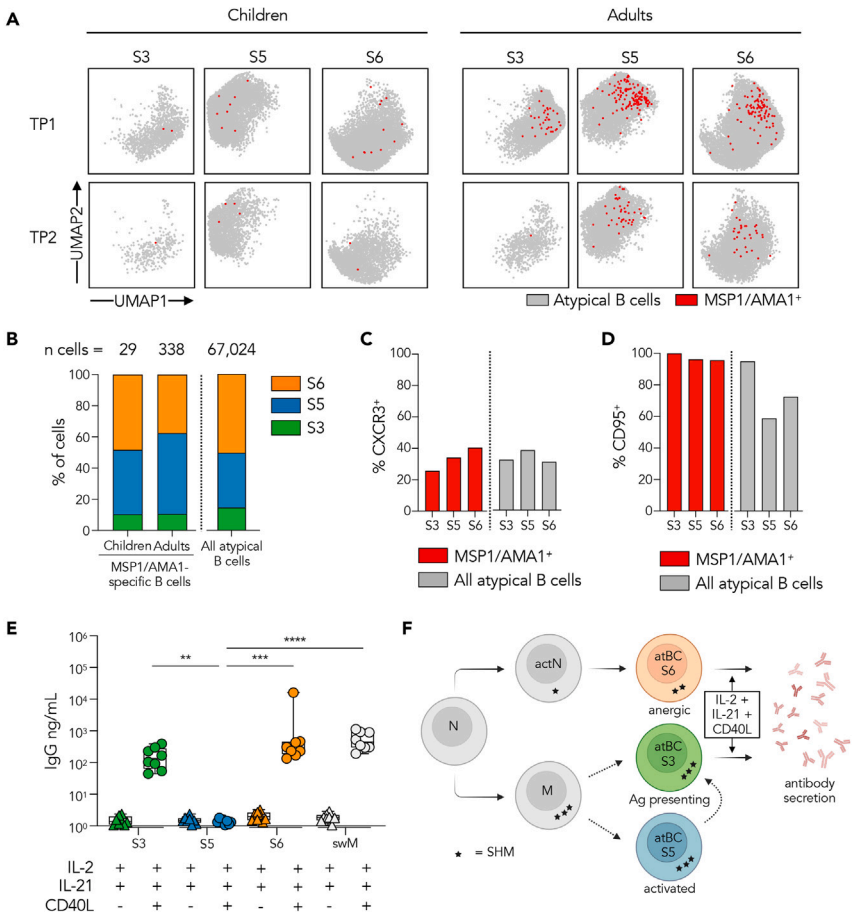


Figure 7. Antigen specificity and antibody secretion among three atypical B cell subsets

(A) Projection of MSP1/AMA1-specific B cells (red) onto the UMAPs of atypical B cell subsets, shown separately for time point 1 (TP1) and time point 2 (TP2) in 9 children and 10 adults.

(B) Distribution of all MSP1/AMA1-specific B cells over the three atypical B cell subsets in children and adults. Due to the small number of antigen-specific B cells, data from both time points were combined.

(C and D) Percentage of all CXCR3$^+$ (C) and CD95$^+$ (D) antigen-specific cells in each atypical B cell subset for samples from 10 adults. Because of the scarcity of antigen-specific B cells, antigen-specific B cells of both time points in adults were combined for this analysis.

(E) IgG concentration in culture supernatant after 14 days of *in vitro* culture. p values shown are from Kruskal-Wallis post hoc test following a one-way ANOVA. **p < 0.01; ***p < 0.001; ****p < 0.0001.

(F) Proposed model of the effector functions and developmental pathways of atypical B cell subsets. N, naive; actN, activated naive; M, memory; atBC, atypical B cell; SHM, somatic hypermutation; Ag, antigen. See also Figures S15, S16, and S21.

DISCUSSION

Atypical B cells accumulate in a range of conditions that share chronic immune activation as the common denominator, including chronic or repetitive infection and autoimmunity. Attempts to understand the function of these cells have resulted in conflicting observations. We hypothesized that these different observations are caused by heterogeneity among atypical B cells and that the underlying subsets are indistinguishable with the markers conventionally used to define atypical B cells. Here, we used multimodal single-cell measurements to better define the heterogeneity among atypical B cells using samples from individuals with chronic exposure to *Plasmodium* parasites. We identified three subsets of atypical B cells, each with a unique transcriptional program and cell surface marker signature. Using spectral flow cytometry, we confirmed the presence of these atypical B cell subsets in a cohort of malaria-experienced children and adults, determined that they can develop in response to antigen stimulation, and assessed their ability to differentiate into antibody-secreting cells.

We generated single-cell sequencing data for almost 21,500 B cells, of which we were able to assign slightly more than 14,000 cells to one of 10 B cell populations defined based on surface marker expression. Among these were just over 4,000 cells that belonged to three clusters containing atypical B cells (here defined as CD21$^-$CD27$^-$CD11c$^+$IgG$^+$). Two recent studies have attempted to assess potential heterogeneity among atypical B cells using single-cell transcriptomics approaches. Holla et al. subclustered 199 B cells from malaria-experienced individuals

with an atypical transcriptomics profile and observed three subsets that were mainly discriminated by isotype usage (i.e., IgD^+IgM^+, IgD^+IgM^{lo}, and IgD^-IgG^+ atypical B cells).[11] Too few cells were present in that dataset to observe subclusters within the IgD^-IgG^+ atypical B cell group, which was the focus of our study. Sutton et al. analyzed a larger number of IgD^- B cells (7,167 cells from malaria-experienced individuals and 5,813 cells from individuals not exposed to *Plasmodium* infection) that were divided into three atypical B cell clusters and several other B cell clusters.[3] However, reanalysis of this dataset did not reveal a clear separation into the same subsets of atypical B cells that we observed in our dataset. One reason for this observation could be that Sutton et al. used samples from immune adults who had not recently had malaria. Likewise, the gene expression profiles of atypical B cell subsets 3 and 6 from *Plasmodium*-exposed adults that were generated in this study were very similar (Figure S22), although differences in expression of genes that underlie the different function profiles of the three atypical B cells subsets were still apparent (Figure 5C). Malaria may drive these functionally distinct gene expression programs, followed by their convergence in the absence of such a strong stimulus.

Using our single-cell transcriptomics dataset of atypical B cells, we confirmed similarities in gene expression between our three atypical B cell clusters and bulk RNA sequencing transcriptional programs that have previously been reported for atypical B cells.[10,12] However, we also discovered that some genes previously linked to atypical B cells were expressed in only one of these three clusters. These genes may be connected to the different functional profiles of these cells. For example, CD72 negatively regulates B cell activation.[27] *CD72* was uniquely expressed in cluster 6 that has a profile of senescence or anergy. Cluster 3, on the other hand, expressed higher levels of the gene encoding CD68, a transmembrane glycoprotein that is mainly associated with the endosomal/lysosomal compartment in monocytes and macrophages but is also expressed by lymphocytes.[49,50] This observation may be related to the pathway of antigen processing that was upregulated in this cluster. The functional significance of CD72 or CD68 expression in the three atypical B cell subsets identified here remains to be determined experimentally.

Using surface markers CD11c and CD86, we were able to identify three corresponding subsets of atypical B cells by flow cytometry. All three subsets harbored *P. falciparum*-specific B cells, suggesting that each of these subsets can develop as part of a specific response to pathogen-derived antigens and not merely through bystander activation. The lower levels of somatic hypermutation set cluster 6 apart from clusters 3 and 5. Indeed, RNA velocity and B cell receptor repertoire analyses suggested that cluster 6 formed a compartment that was mostly separated from clusters 3 and 5 and followed a different developmental trajectory. The RNA velocity analysis and the increase in percentage of B cell clones that were shared between clusters 3 and 5 over time as the B cell response contracted to baseline levels suggest that atypical B cells in subsets 3 and 5 may be developmentally related, either by sharing a common ancestor or through direct differentiation of subset 5 into subset 3.

Based on these data, we propose the following working model (Figure 7F). Chronic immune activation results in the development of three subsets of atypical B cells, named "subset 3," "subset 5," and "subset 6". The latter is derived from activated naive B cells and has relatively low levels of somatic hypermutation. These characteristics are similar to atypical B cells in SLE patients that arise through extrafollicular activation. Under conditions that mimic the interaction with $CD4^+$ T cells, subset 6 atypical B cells can further differentiate into short-lived antibody-secreting cells, although their transcriptional profile suggests that these cells are anergic and may have a high activation threshold. The main developmental trajectory of subset 3 atypical B cells is from memory B cells, possibly via an activated subset 5 state. These cells have undergone high levels of somatic hypermutation, consistent with the normal development of memory B cells. Subset 3 atypical B cells are primed for interactions with T cells and will differentiate into antibody-secreting cells upon stimulation by T cells.[18] Additionally, these cells express proteins that may modulate the function of T cells in return. Subset 5 atypical B cells express lower levels of markers typically associated with atypical B cells, such as FcRL5, T-bet, and CD11c, and do not differentiate into antibody-secreting cells by simulating the interaction with T cells. Other stimuli, such as B cell receptor engagement, additional cytokines, or Toll-like receptor (TLR) engagement may drive the differentiation of this subset into antibody-secreting cells. Indeed, Sullivan et al. reported that a subset of atypical B cells that expressed lower levels of FcRL5 showed more antibody secretion upon B cell receptor stimulation than $FcRL5^+$ atypical B cells.[10] These cells could be functionally immature or exhausted, resulting in the need for additional stimuli to differentiate into antibody-secreting cells. Collectively, the functional profiles of the three subsets of atypical B cells recapitulate many of the observations that have been made about the atypical B cell population as a whole. Our results indicate that specific subsets of atypical B cells are likely responsible for conflicting observations related to anergy/tolerance, antigen presentation, and the ability to differentiate into antibody-secreting cells.

Comparing atypical B cells between *Plasmodium*-exposed children and adults revealed additional heterogeneity. Adults showed a higher percentage of $CXCR3^+$ cells across all three subsets of atypical B cells, which did not decrease after a two-year period of minimal *P. falciparum* exposure. Hopp et al. reported that expression of *CXCR3* is upregulated in atypical B cells in response to malaria in children.[18] The stable percentage of $CXCR3^+$ cells in all atypical B cell subsets in adults suggests that expression of this marker is the result of chronic antigen exposure and may play an important role in the protective immune response against malaria. Sutton et al. observed that *Pf*MSP1-specific B cells in adults expressed higher levels of CXCR3 than the total memory B cell population, although the majority of these cells did not have a $CD21^-CD27^-$ phenotype.[51] We also detected $CXCR3^+$ parasite-specific cells. However, parasite-specific cells were not enriched for CXCR3 expression among atypical B cells. CXCR3 is an interferon-inducible chemokine receptor specific for the chemokines CXCL9, CXCL10, and CXCL11 that are themselves also induced by interferon.[52] Its function has mainly been studied in T cells where it is required for trafficking of T cells to sites of inflammation.[53] The profile of trafficking receptor expression in atypical B cells, including CXCR3, has been linked to the propensity of $T\text{-}bet^{hi}$ B cells to take up residency in the spleen and bone marrow.[51] While no data are available for the tissue distribution of atypical B cells in malaria-experienced individuals, it would make sense for parasite-specific B cells to home to the spleen,

since this is the main anatomical location for the removal of parasite-infected erythrocytes and cellular debris from the circulation. In this regard, it is interesting to note that the elevated expression of T-bet and CD74 in subset 3 atypical B cells fits the profile of memory B cells that reside in the spleen,[51,54] indicating that the loss of these cells in the circulation could be the result of migration into the spleen. However, the observation that the percentage of CXCR3$^+$ atypical B cells does not change over time with minimal parasite exposure suggests that the expression of this marker by itself does not direct homing of B cells to tissues. This notion is strengthened by the recent finding that CXCR3 ablation in murine B cells has no effect on the establishment and maintenance of lung-resident memory B cells.[55] Instead, CXCR3 should probably be considered a marker of activation or prior activation.

In conclusion, we found that atypical B cells can be divided into three functionally distinct subsets. These findings reconcile many of the conflicting observations that have been made about atypical B cells and provide insight into their function in the immune response against infection. Our study also highlights that the composition of the atypical B cell population can alter with changes in pathogen exposure levels. The selection of individuals and timing of sample collection therefore warrant careful consideration when analyzing atypical B cells and their functional profile. Although the transcriptomes of atypical B cells in chronic infections and autoimmunity are remarkably similar,[11] our observations about atypical B cells in malaria-experienced individuals will need to be confirmed in other disease models. Collectively, our results suggest that a large fraction of atypical B cells may contribute to a productive and antigen-specific immune response against infection.

Limitations of the study

One of the complications of studying atypical B cells is that different sets of markers have been used to define atypical B cells. In this study, our definition of atypical B cells is based on how these cells are typically identified in the HIV and malaria fields (CD21$^-$CD27$^-$) and a recent scRNA-seq study that reported CD11c as the most informative cell surface marker for the identification of atypical B cell lineages.[3] Our definition of atypical B cells as IgG$^+$CD21$^-$CD27$^-$CD11c$^+$ also fits the definition of DN2 B cells, which is how these cells are commonly referred to in the autoimmunity field (IgD$^-$CD27$^-$CD21$^-$ CD11c$^+$) and follows recently published recommendations for how these cells should be defined.[22] Not all of the cells that fall within our definition of atypical B cells are T-bet$^+$, which may complicate translating our findings to studies that use T-bet (with or without CD11c) as a defining marker for atypical B cells.[51,56,57] We here show that CD11c and CD86 can be used to subset atypical B cells into three subpopulations. While subset 5 was identified here by the lack of CD86 and expression of intermediate levels of CD11c, a positive marker for this subset would facilitate future studies. The surface marker transmembrane activator and calcium-modulator and CAML interactor (TACI, also known as CD267) was upregulated in cluster 5 as compared to clusters 3 and 6 but will need to be validated as a potential positive marker for this B cell population.

Another limitation of this study is that the single-cell sequencing results were derived from samples from only two children. While the results between these two children were highly consistent, a larger sample size may provide additional insight into the role of atypical B cells in the immune response against malaria parasites. In addition, the phenotypic differences in atypical B cells between children and adults and the lack of a clear separation of atypical B cells into subsets in adults who have not recently had malaria as seen in our bulk RNA sequencing analysis will require additional experimental validation. In this respect, it would be very interesting to compare the transcriptomic profiles of atypical B cells between these two groups that differ in age and cumulative parasite exposure.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - B cell isolations
  - Flow cytometry staining for single-cell analysis
  - Single-cell gene expression, V(D)J, and feature barcoding library preparation
  - Atypical B cell sorting
  - Bulk RNA-Sequencing
  - Spectral flow cytometry
  - B cell cultures
  - Enzyme-linked immunosorbent assays
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Single-cell RNA-seq data analysis
  - Single-cell RNA-sequencing gene signature score analysis
  - Single-cell surface marker analysis
  - Re-analysis of single-cell RNA-seq data sets for HIV and SLE samples

- Single-cell B cell receptor analysis
- RNA velocity analysis
- Bulk RNA-seq analysis
- Spectral flow cytometry analysis
- Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108496.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (R01 AI153425 to E.M.B., R35 GM128938 to F.A., F31 AI169993 to R.A.R., TL1 TR002647 to R.A.R., T32 AI138944 to A.B.R., U19 AI150741 to B.G. and P.J., and U19 AI089674). The authors acknowledge the Cell Analysis Core Facility (Dr. Sandra Cardona) and the Genomics Research Core Facility (Sean Vargas) at the University of Texas at San Antonio for support during this work. Data were generated in the Genome Sequencing Facility, which is supported by UT Health San Antonio, NIH-NCI P30 CA054174 (Mays Cancer Center at UT Health San Antonio), and NIH Shared Instrument grant 1S10OD030311-01 (S10 grant toward NovaSeq6000 sequencing system), and in the Flow Cytometry Shared Resource Facility, which is supported by UT Health San Antonio, NIH-NCI P30 CA054174 (Mays Cancer Center at UT Health), CPRIT (RP210126) and NIH Shared Instrument grant 1S10OD030432-01 (S10 grant toward the purchase of a DB FACSAria Fusion). Plasmids encoding 3D7 MSP1-bio, AMA1-bio, and BirA were a kind gift from Dr. Gavin Wright (Wellcome Sanger Institute).

AUTHOR CONTRIBUTIONS

R.A.R. performed bulk RNA sequencing experiments, spectral flow cytometry experiments, *in vitro* B cell cultures, ELISAs, and data analysis. G.B. performed single-cell sequencing experiments and data analysis. P.D. and F.A. analyzed single-cell sequencing data. A.B.R. analyzed somatic hypermutation data. R.G. provided critical feedback on the manuscript. I.S., P.J., M.E.F., and B.G. provided clinical samples. R.A.R., G.B., S.B., and E.M.B. wrote the manuscript with input from all other co-authors. All authors contributed to the article and approved the submitted version.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: August 30, 2023
Revised: October 30, 2023
Accepted: November 16, 2023
Published: November 20, 2023

REFERENCES

1. Lau, D., Lan, L.Y.L., Andrews, S.F., Henry, C., Rojas, K.T., Neu, K.E., Huang, M., Huang, Y., DeKosky, B., Palm, A.K.E., et al. (2017). Low CD21 expression defines a population of recent germinal center graduates primed for plasma cell differentiation. Sci. Immunol. *2*, eaai8153.
2. Andrews, S.F., Chambers, M.J., Schramm, C.A., Plyler, J., Raab, J.E., Kanekiyo, M., Gillespie, R.A., Ransier, A., Darko, S., Hu, J., et al. (2019). Activation dynamics and immunoglobulin evolution of pre-existing and newly generated human memory B cell responses to Influenza hemagglutinin. Immunity *51*, 398–410.e5.
3. Sutton, H.J., Aye, R., Idris, A.H., Vistein, R., Nduati, E., Kai, O., Mwacharo, J., Li, X., Gao, X., Andrews, T.D., et al. (2021). Atypical B cells are part of an alternative lineage of B cells that participates in responses to vaccination and infection in humans. Cell Rep. *34*, 108684.
4. Sundling, C., Rönnberg, C., Yman, V., Asghar, M., Jahnmatz, P., Lakshmikanth, T., Chen, Y., Mikes, J., Forsell, M.N., Sondén, K., et al. (2019). B cell profiling in malaria reveals expansion and remodelling of CD11c+ B cell subsets. JCI Insight *5*, e126492.
5. Moir, S., Ho, J., Malaspina, A., Wang, W., DiPoto, A.C., O'Shea, M.A., Roby, G., Kottilil, S., Arthos, J., Proschan, M.A., et al. (2008). Evidence for HIV-associated B cell exhaustion in a dysfunctional memory B cell compartment in HIV-infected viremic individuals. J. Exp. Med. *205*, 1797–1805.
6. Weiss, G.E., Crompton, P.D., Li, S., Walsh, L.A., Moir, S., Traore, B., Kayentao, K., Ongoiba, A., Doumbo, O.K., and Pierce, S.K. (2009). Atypical memory B cells are greatly expanded in individuals living in a malaria-endemic area. J. Immunol. *183*, 2176–2182.
7. Wei, C., Anolik, J., Cappione, A., Zheng, B., Pugh-Bernard, A., Brooks, J., Lee, E.H., Milner, E.C.B., and Sanz, I. (2007). A new population of cells lacking expression of CD27 represents a notable component of the B cell memory compartment in systemic lupus erythematosus. J. Immunol. *178*, 6624–6633.
8. Isnardi, I., Ng, Y.S., Menard, L., Meyers, G., Saadoun, D., Srdanovic, I., Samuels, J., Berman, J., Buckner, J.H., Cunningham-Rundles, C., and Meffre, E. (2010). Complement receptor 2/CD21- human naive B cells contain mostly autoreactive unresponsive clones. Blood *115*, 5026–5036.
9. Jenks, S.A., Cashman, K.S., Zumaquero, E., Marigorta, U.M., Patel, A.V., Wang, X., Tomar, D., Woodruff, M.C., Simon, Z., Bugrovsky, R., et al. (2018). Distinct effector B cells induced by unregulated toll-like receptor 7 contribute to pathogenic responses in systemic lupus erythematosus. Immunity *49*, 725–739.e6.

10. Sullivan, R.T., Kim, C.C., Fontana, M.F., Feeney, M.E., Jagannathan, P., Boyle, M.J., Drakeley, C.J., Ssewanyana, I., Nankya, F., Mayanja-Kizza, H., et al. (2015). FCRL5 delineates functionally impaired memory B cells associated with *Plasmodium falciparum* exposure. PLoS Pathog. *11*, e1004894.
11. Holla, P., Dizon, B., Ambegaonkar, A.A., Rogel, N., Goldschmidt, E., Boddapati, A.K., Sohn, H., Sturdevant, D., Austin, J.W., Kardava, L., et al. (2021). Shared transcriptional profiles of atypical B cells suggest common drivers of expansion and function in malaria, HIV, and autoimmunity. Sci. Adv. *7*, eabg8384.
12. Portugal, S., Tipton, C.M., Sohn, H., Kone, Y., Wang, J., Li, S., Skinner, J., Virtaneva, K., Sturdevant, D.E., Porcella, S.F., et al. (2015). Malaria-associated atypical memory B cells exhibit markedly reduced B cell receptor signaling and effector function. Elife *4*, e07218.
13. Austin, J.W., Buckner, C.M., Kardava, L., Wang, W., Zhang, X., Melson, V.A., Swanson, R.G., Martins, A.J., Zhou, J.Q., Hoehn, K.B., et al. (2019). Overexpression of T-bet in HIV infection is associated with accumulation of B cells outside germinal centers and poor affinity maturation. Sci. Transl. Med. *11*, eaax0904.
14. Ambegaonkar, A.A., Kwak, K., Sohn, H., Manzella-Lapeira, J., Brzostowski, J., and Pierce, S.K. (2020). Expression of inhibitory receptors by B cells in chronic human infectious diseases restricts responses to membrane-associated antigens. Sci. Adv. *6*, eaba6493.
15. Wing, E., Sutherland, C., Miles, K., Gray, D., Goodyear, C., Otto, T.D., Breusch, S., Cowan, G.J.M., and Gray, M. (2023). A comprehensive analysis of rheumatoid arthritis B cells reveals the importance of CD11c^{+ve} double-negative-2 B cells as the major synovial plasma cell precursor. Preprint at bioRxiv. https://doi.org/10.1101/2023.02.15.526468.
16. Muellenbeck, M.F., Ueberheide, B., Amulic, B., Epp, A., Fenyo, D., Busse, C.E., Esen, M., Theisen, M., Mordmüller, B., and Wardemann, H. (2013). Atypical and classical memory B cells produce *Plasmodium falciparum* neutralizing antibodies. J. Exp. Med. *210*, 389–399.
17. Holla, P., Ambegaonkar, A., Sohn, H., and Pierce, S.K. (2019). Exhaustion may not be in the human B cell vocabulary, at least not in malaria. Immunol. Rev. *292*, 139–148.
18. Hopp, C.S., Skinner, J., Anzick, S.L., Tipton, C.M., Peterson, M.E., Li, S., Doumbo, S., Kayentao, K., Ongoiba, A., Martens, C., et al. (2022). Atypical B cells up-regulate costimulatory molecules during malaria and secrete antibodies with T follicular helper cell support. Sci. Immunol. *7*, eabn1250.
19. Braddom, A.E., Batugedara, G., Bol, S., and Bunnik, E.M. (2020). Potential functions of atypical memory B cells in *Plasmodium*-exposed individuals. Int. J. Parasitol. *50*, 1033–1042.
20. Stewart, A., Ng, J.C.F., Wallis, G., Tsioligka, V., Fraternali, F., and Dunn-Walters, D.K. (2021). Single-cell transcriptomic analyses define distinct peripheral B cell subsets and discrete development pathways. Front. Immunol. *12*, 602539.
21. 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.
22. Sanz, I., Wei, C., Jenks, S.A., Cashman, K.S., Tipton, C., Woodruff, M.C., Hom, J., and Lee, F.E.H. (2019). Challenges and opportunities for consistent classification of human B cell and plasma cell populations. Front. Immunol. *10*, 2458.
23. Braddom, A.E., Bol, S., Gonzales, S.J., Reyes, R.A., Musinguzi, K., Nankya, F., Ssewanyana, I., Greenhouse, B., and Bunnik, E.M. (2021). B cell receptor repertoire analysis in malaria-naive and malaria-experienced individuals reveals unique characteristics of atypical memory B cells. mSphere *6*, e0072621.
24. Glass, D.R., Tsai, A.G., Oliveria, J.P., Hartmann, F.J., Kimmey, S.C., Calderon, A.A., Borges, L., Glass, M.C., Wagar, L.E., Davis, M.M., and Bendall, S.C. (2020). An integrated multi-omic single-cell atlas of human B cell identity. Immunity *53*, 217–232.e5.
25. Dai, D., Gu, S., Han, X., Ding, H., Jiang, Y., Zhang, X., Yao, C., Hong, S., Zhang, J., Shen, Y., et al. (2023). The transcription factor Zeb2 drives formation of age-associated B cells. Preprint at bioRxiv. https://doi.org/10.1101/2021.07.24.453633.
26. Gao, X., Shen, Q., Roco, J.A., Frith, K., Munier, C.M.L., Nekrasov, M., Dalton, B., He, J.-S., Jaeger, R., Cook, M.C., et al. (2022). ZEB2 regulates the development of CD11c+ atypical B cells. Preprint at bioRxiv. https://doi.org/10.1101/2022.09.01.506173.
27. Akatsu, C., Shinagawa, K., Numoto, N., Liu, Z., Ucar, A.K., Aslam, M., Phoon, S., Adachi, T., Furukawa, K., Ito, N., and Tsubata, T. (2016). CD72 negatively regulates B lymphocyte responses to the lupus-related endogenous toll-like receptor 7 ligand Sm/RNP. J. Exp. Med. *213*, 2691–2706.
28. Giudizi, M.G., Biagiotti, R., Almerigogna, F., Alessi, A., Tiri, A., Del Prete, G.F., Ferrone, S., and Romagnani, S. (1987). Role of HLA class I and class II antigens in activation and differentiation of B cells. Cell. Immunol. *108*, 97–108.
29. Cyster, J.G., and Allen, C.D.C. (2019). B cell responses: cell interaction dynamics and decisions. Cell *177*, 524–540.
30. Stumptner-Cuvelette, P., and Benaroch, P. (2002). Multiple roles of the invariant chain in MHC class II function. Biochim. Biophys. Acta *1542*, 1–13.
31. Tan, C., Hiwa, R., Mueller, J.L., Vykunta, V., Hibiya, K., Noviski, M., Huizar, J., Brooks, J.F., Garcia, J., Heyn, C., et al. (2020). NR4A nuclear receptors restrain B cell responses to antigen when second signals are absent or limiting. Nat. Immunol. *21*, 1267–1279.
32. Tan, C., Mueller, J.L., Noviski, M., Huizar, J., Lau, D., Dubinin, A., Molofsky, A., Wilson, P.C., and Zikherman, J. (2019). Nur77 links chronic antigen stimulation to B cell tolerance by restricting the survival of self-reactive B cells in the periphery. J. Immunol. *202*, 2907–2923.
33. Campisi, J. (2013). Aging, cellular senescence, and cancer. Annu. Rev. Physiol. *75*, 685–705.
34. Contrepois, K., Coudereau, C., Benayoun, B.A., Schuler, N., Roux, P.F., Bischof, O., Courbeyrette, R., Carvalho, C., Thuret, J.Y., Ma, Z., et al. (2017). Histone variant H2A.J accumulates in senescent cells and promotes inflammatory gene expression. Nat. Commun. *8*, 14995.
35. Frasca, D., Romero, M., Diaz, A., Garcia, D., Thaller, S., and Blomberg, B.B. (2021). B cells with a senescent-associated secretory phenotype accumulate in the adipose tissue of individuals with obesity. Int. J. Mol. Sci. *22*, 1839.
36. Nehar-Belaid, D., Hong, S., Marches, R., Chen, G., Bolisetty, M., Baisch, J., Walters, L., Punaro, M., Rossi, R.J., Chung, C.H., et al. (2020). Mapping systemic lupus erythematosus heterogeneity at the single-cell level. Nat. Immunol. *21*, 1094–1106.
37. Joyce, C., Murrell, S., Murrell, B., Omorodion, O., Ver, L.S., Carrico, N., Bastidas, R., Nedellec, R., Bick, M., Woehl, J., et al. (2023). Antigen pressure from two founder viruses induces multiple insertions at a single antibody position to generate broadly neutralizing HIV antibodies. PLoS Pathog. *19*, e1011416.
38. Di Niro, R., Lee, S.J., Vander Heiden, J.A., Elsner, R.A., Trivedi, N., Bannock, J.M., Gupta, N.T., Kleinstein, S.H., Vigneault, F., Gilbert, T.J., et al. (2015). Salmonella infection drives promiscuous B cell activation followed by extrafollicular affinity maturation. Immunity *43*, 120–131.
39. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498.
40. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. *38*, 1408–1414.
41. Gonzales, S.J., Bol, S., Braddom, A.E., Sullivan, R., Reyes, R.A., Ssewanyana, I., Eggers, E., Greenhouse, B., and Bunnik, E.M. (2021). Longitudinal analysis of FcRL5 expression and clonal relationships among classical and atypical memory B cells following malaria. Malar. J. *20*, 435.
42. Hershberg, U., and Luning Prak, E.T. (2015). The analysis of clonal expansions in normal and autoimmune B cell repertoires. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20140239.
43. Meng, W., Zhang, B., Schwartz, G.W., Rosenfeld, A.M., Ren, D., Thome, J.J.C., Carpenter, D.J., Matsuoka, N., Lerner, H., Friedman, A.L., et al. (2017). An atlas of B-cell clonal distribution in the human body. Nat. Biotechnol. *35*, 879–884.
44. Andrade, C.M., Fleckenstein, H., Thomson-Luque, R., Doumbo, S., Lima, N.F., Anderson, C., Hibbert, J., Hopp, C.S., Tran, T.M., Li, S., et al. (2020). Increased circulation time of *Plasmodium falciparum* underlies persistent asymptomatic infection in the dry season. Nat. Med. *26*, 1929–1940.
45. Zinszer, K., Charland, K., Vahey, S., Jahagirdar, D., Rek, J.C., Arinaitwe, E., Nankabirwa, J., Morrison, K., Sadoine, M.L., Tutt-Guérette, M.A., et al. (2020). The impact of multiple rounds of indoor residual spraying on malaria incidence and hemoglobin levels in a high-transmission setting. J. Infect. Dis. *221*, 304–312.
46. Kamya, M.R., Kakuru, A., Muhindo, M., Arinaitwe, E., Nankabirwa, J.I., Rek, J., Bigira, V., Kapisi, J., Wanzira, H., Achan, J., et al. (2020). The impact of control interventions on malaria burden in young children in a historically high-transmission district of Uganda: a pooled analysis of cohort studies from 2007 to 2018. Am. J. Trop. Med. Hyg. *103*, 785–792.
47. Gonzales, S.J., Clarke, K.N., Batugedara, G., Garza, R., Braddom, A.E., Reyes, R.A., Ssewanyana, I., Garrison, K.C., Ippolito, G.C., Greenhouse, B., et al. (2022). A molecular

analysis of memory B cell and antibody responses against *Plasmodium falciparum* merozoite surface protein 1 in children and adults From Uganda. Front. Immunol. *13*, 809264.

48. Reyes, R.A., Clarke, K., Gonzales, S.J., Cantwell, A.M., Garza, R., Catano, G., Tragus, R.E., Patterson, T.F., Bol, S., and Bunnik, E.M. (2021). SARS-CoV-2 spike-specific memory B cells express higher levels of T-bet and FcRL5 after non-severe COVID-19 as compared to severe disease. PLoS One *16*, e0261656.
49. Holness, C.L., and Simmons, D.L. (1993). Molecular cloning of CD68, a human macrophage marker related to lysosomal glycoproteins. Blood *81*, 1607–1613.
50. Gottfried, E., Kunz-Schughart, L.A., Weber, A., Rehli, M., Peuker, A., Müller, A., Kastenberger, M., Brockhoff, G., Andreesen, R., and Kreutz, M. (2008). Expression of CD68 in non-myeloid cell types. Scand. J. Immunol. *67*, 453–463.
51. Johnson, J.L., Rosenthal, R.L., Knox, J.J., Myles, A., Naradikian, M.S., Madej, J., Kostiv, M., Rosenfeld, A.M., Meng, W., Christensen, S.R., et al. (2020). The transcription factor T-bet resolves memory B cell subsets with distinct tissue distributions and antibody specificities in mice and humans. Immunity *52*, 842–855.e6.
52. Groom, J.R., and Luster, A.D. (2011). CXCR3 ligands: redundant, collaborative and antagonistic functions. Immunol. Cell Biol. *89*, 207–215.
53. Groom, J.R., and Luster, A.D. (2011). CXCR3 in T cell function. Exp. Cell Res. *317*, 620–631.
54. Weisel, N.M., Joachim, S.M., Smita, S., Callahan, D., Elsner, R.A., Conter, L.J., Chikina, M., Farber, D.L., Weisel, F.J., and Shlomchik, M.J. (2022). Surface phenotypes of naive and memory B cells in mouse and human tissues. Nat. Immunol. *23*, 135–145.
55. Gregoire, C., Spinelli, L., Villazala-Merino, S., Gil, L., Holgado, M.P., Moussa, M., Dong, C., Zarubica, A., Fallet, M., Navarro, J.M., et al. (2022). Viral infection engenders bona fide and bystander subsets of lung-resident memory B cells through a permissive mechanism. Immunity *55*, 1216–1233.e9.
56. Keller, B., Strohmeier, V., Harder, I., Unger, S., Payne, K.J., Andrieux, G., Boerries, M., Felixberger, P.T., Landry, J.J.M., Nieters, A., et al. (2021). The expansion of human T-bet(high)CD21(low) B cells is T cell dependent. Sci. Immunol. *6*, eabh0891.
57. Wang, S., Wang, J., Kumar, V., Karnell, J.L., Naiman, B., Gross, P.S., Rahman, S., Zerrouki, K., Hanna, R., Morehouse, C., et al. (2018). IL-21 drives expansion and plasma cell differentiation of autoreactive CD11c(hi)T-bet(+) B cells in SLE. Nat. Commun. *9*, 1758.
58. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21.
59. Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., et al. (2020). The reactome pathway knowledgebase. Nucleic Acids Res. *48*, D498–D503.
60. Gupta, N.T., Vander Heiden, J.A., Uduman, M., Gadala-Maria, D., Yaari, G., and Kleinstein, S.H. (2015). Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. Bioinformatics *31*, 3356–3358.
61. Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: computational identification of cell doublets in single-cell transcriptomic data. Cell Syst. *8*, 281–291.e9.
62. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296.
63. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890.
64. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.
65. Kilama, M., Smith, D.L., Hutchinson, R., Kigozi, R., Yeka, A., Lavoy, G., Kamya, M.R., Staedke, S.G., Donnelly, M.J., Drakeley, C., et al. (2014). Estimating the annual entomological inoculation rate for *Plasmodium falciparum* transmitted by *Anopheles gambiae s.l.* using three sampling methods in three sites in Uganda. Malar. J. *13*, 111.
66. Kamya, M.R., Arinaitwe, E., Wanzira, H., Katureebe, A., Barusya, C., Kigozi, S.P., Kilama, M., Tatem, A.J., Rosenthal, P.J., Drakeley, C., et al. (2015). Malaria transmission, infection, and disease at three sites with varied transmission intensity in Uganda: implications for malaria control. Am. J. Trop. Med. Hyg. *92*, 903–912.
67. Van Gassen, S., Callebaut, B., Van Helden, M.J., Lambrecht, B.N., Demeester, P., Dhaene, T., and Saeys, Y. (2015). FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. Cytometry A. *87*, 636–645.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Super Bright 600 anti-human CD10 (clone CB-CALLA)	Thermo	Cat# 63-0106-41; RRID:AB_2688183
Brilliant Violet 421 anti-human CD19 (clone SJ25C1)	BioLegend	Cat# 363017 ; RRID:AB_2564226
Brilliant Violet 785 anti-human CD20 (clone 2H7)	BioLegend	Cat# 302355; RRID:AB_2566315
PerCP-eFluor 710 anti-human CD21 (clone HB5)	Thermo	Cat# 46-0219-41; RRID:AB_2573671
PE/Cyanine7 anti-human CD27 (clone O323)	BioLegend	Cat# 302837; RRID:AB_2561918
TotalSeq-C0181 anti-human CD21 (clone Bu32)	BioLegend	Cat# 354923; RRID:AB_2800953
TotalSeq-C0154 anti-human CD27 (clone O323)	BioLegend	Cat# 302853; RRID:AB_2800747
TotalSeq-C0053 anti-human CD11c (clone S-HCL-3)	BioLegend	Cat# 371521; RRID:AB_2801018
TotalSeq-C0140 anti-human CXCR3 (clone G025H7)	BioLegend	Cat# 353747; RRID:AB_2800949
TotalSeq-C0375 anti-human IgG (clone M1310G05)	BioLegend	Cat# 410727; RRID:AB_2801087
TotalSeq-C0136 anti-human IgM (clone MHM-88)	BioLegend	Cat# 314547; RRID:AB_2800835
Super Bright 645 anti-human CD19 (clone HIB19)	Thermo	Cat# 64-0199-42; RRID:AB_2662583
Spark YG 593 anti-human CD20 (clone 2H7)	BioLegend	Cat# 302367; RRID:AB_2894449
PerCP-eFluor 710 anti-human CD21 (clone HB5)	Thermo	Cat# 46-0219-42; RRID:AB_2573672
Brilliant Violet 605 anti-human CD24 (clone ML5)	BioLegend	Cat# 311123; RRID:AB_2562287
PE/Cyanine7 anti-human CD27 (clone O323)	BioLegend	Cat# 302837; RRID:AB_2561918
APC/Fire 810 anti-human CD38 (clone HB-7)	BioLegend	Cat# 356643; RRID:AB_2860936
PE/Cyanine5 anti-human CD83 (cloneHB15e)	BioLegend	Cat# 305310; RRID:AB_314518
PE-Dazzle 594 anti-human CD86 (clone BU63)	BioLegend	Cat# 374217; RRID:AB_2814336
BUV737 anti-human CD95 (clone DX2)	BD	Cat# 612790; RRID:AB_2870117
Pacific Blue anti-human IgD (clone IA6-2)	BioLegend	Cat# 348223; RRID:AB_2561596
BV570 anti-human IgM (clone MHM-88)	BioLegend	Cat# 314517; RRID:AB_10913816
BV510 anti-human IgG (clone M1310G05)	BioLegend	Cat# 410715; RRID:AB_2728445
FITC anti-human IgA (clone mA-6E1)	Miltenyi	Cat# 130114001; RRID:AB_2726443
PE-Fire 810 anti-human CXCR3 (clone G025H7)	BioLegend	Cat# 353759; RRID:AB_2894484
Brilliant Violet 750 anti-human CXCR5 (clone J252D4)	BioLegend	Cat# 356941; RRID:AB_2832703
APC/Cyanine 7 anti-human CD11c (clone Bu15)	BioLegend	Cat# 337217; RRID:AB_10661724
Brilliant Violet 711 anti-human Tbet (used for intracellular staining step) (clone 4B10)	BioLegend	Cat# 644819; RRID:AB_11218985
BUV805 anti-human FcRL5 (clone 509F6)	BD	Cat# 749599; RRID:AB_2873901
Goat anti-human IgG	Sigma	Cat# I2136; RRID:AB_260147
IgG from human serum	Sigma	Cat# I2511; RRID:AB_1163604
HRP-conjugated donkey anti-human IgG	BioLegend	Cat# 410902; RRID:AB_2686937
Biological samples		
PBMCs from *Plasmodium falciparum*-exposed individuals	This paper (Table S1)	N/A
Chemicals, peptides, and recombinant proteins		
Brilliant Stain Buffer Plus	BD	Cat# 566385
SpectroFlo QC Beads	Cytek	Cat# N7-97355
UltraComp eBeads Plus Compensation Beads	Thermo	Cat# 01-3333-41
APC-conjugated streptavidin	Tonbo	Cat# 20-4317-U100
BUV563-conjugated streptavidin	BD	Cat# 612935

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human TruStain FcX	BioLegend	Cat# 422301
Dextran sulfate sodium salt	Sigma-Aldrich	Cat# 42867-5G
Agencourt AMPure XP beads	Beckman Coulter	Cat# A63880
2x MycoZap Plus-PR	Lonza	Cat# VZA-2021
Human IL-2	GoldBio	Cat# 1110-02-50
Human IL-21	GoldBio	Cat# 1110-21-10
Non-Animal Protein Blocker	G-Biosciences	Cat# 786-190P
Critical commercial assays		
EasySep Human B Cell Isolation Kit	StemCell	Cat#17954
MojoSort Human Pan B Cell Isolation Kit	BioLegend	Cat# 480082
Chromium Single Cell 5′ Library & Gel Bead Kit v1	10x Genomics	Cat# 1000006
Chromium Single Cell A Chip Kit	10x Genomics	Cat# 1000151
Chromium Single Cell 5′ Feature Barcode Library Kit	10x Genomics	Cat# 1000080
Chromium Single Cell 5′ Library Construction Kit	10x Genomics	Cat# 1000020
Chromium Single Cell V(D)J Enrichment Kit	10x Genomics	Cat# 1000016
NEBNext Single Cell/Low Input RNA Library Prep Kit for Illumina	NEB	Cat# E6420S
NEBNext Multiplex Oligos for Illumina	NEB	Cat# E7335S
NEBNext Multiplex Oligos for Illumina	NEB	Cat# E7500S
LIVE/DEAD Fixable Aqua Dead Cell Stain Kit	Thermo	Cat# L34965
Zombie UV Fixable Viability Kit	BioLegend	Cat# 423107
Transcription Factor Staining Buffer Kit	Tonbo	Cat# TNB-0607-KIT
Deposited data		
Raw and analyzed single-cell sequencing data	This paper	GEO: GSE235824
Raw bulk RNA-sequencing data	This paper	SRA: PRJNA1034154
Gene signatures of activated B cells	Ellebedy et al. (2016)[21]	http://www.nature.com/articles/ni.3533
Gene signatures of memory, activated, and atypical B cells	Holla et al. (2021)[11]	https://www.science.org/doi/10.1126/sciadv.abg8384
scRNA-seq data for PBMCs from SLE patients	Nehar-Belaid et al. (2020)[36]	GEO: GSE135779
scRNA-seq data for class-switched memory B cells from HIV-infected patients	Joyce et al. (2023)[37]	GEO: GSE229232
Experimental models: Cell lines		
Human: Expi293F cell line	Thermo	Cat# A14527; RRID:CVCL_D615
Mouse: CD40L-expressing 3T3 cells	Dr. Mark Connors, NIH	RRID:CVCL_1H10
Recombinant DNA		
Plasmid MSP1-bio	Addgene	Cat# 47709; RRID:Addgene_47709
Plasmid AMA1-bio	Addgene	Cat# 47741; RRID:Addgene_47741
Plasmid secretedBirA-8his	Addgene	Cat# 32408; RRID:Addgene_32408
Software and algorithms		
Cell Ranger v3.1.0	10x Genomics	https://www.10xgenomics.com/
Seurat v3.2.3	Stuart et al. (2019)[58]	https://satijalab.org/seurat/
Reactome pathway analysis	Jassal et al. (2020)[59]	https://reactome.org/
Change-O v0.4.4	Gupta et al. (2015)[60]	https://changeo.readthedocs.io/en/stable/
Custom Perl script	Gonzales et al. (2021)[41]	https://github.com/embunnik/clonal_scores
scVelo	Bergen et al. (2020)[40]	https://scvelo.readthedocs.io/en/stable/
Velocyto	La Manno et al. (2018)[39]	http://velocyto.org/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Scrublet	Wolock et al. (2019)[61]	https://pypi.org/project/scrublet/
Harmony	Korsunsky et al. (2019)[62]	https://portals.broadinstitute.org/harmony/index.html
Custom RNA-seq pipeline	N/A	https://github.com/ay-lab/RNA_SEQ_PIPELINE
fastp	Chen et al. (2018)[63]	https://github.com/OpenGene/fastp
STAR v2.7.3a	Dobin et al. (2013)[64]	https://github.com/alexdobin/STAR/
OMIQ	Dotmatics	https://www.omiq.ai/
GraphPad Prism 9	Dotmatics	https://www.graphpad.com/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Evelien M. Bunnik (bunnik@uthscsa.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Raw and processed single-cell sequencing files are available from the NCBI Gene Expression Omnibus under accession no. GSE235824. Bulk RNA-sequencing data are available from the NCBI Sequence Read Archive under accession no. PRJNA1034154.
- This paper does not report original code.
- For any additional information required to reanalyze the data reported in this paper, please contact the lead author.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Individuals included in this study (n = 29) were residents of the Nagongera sub-county in Tororo District, Uganda. This region was historically characterized by extremely high malaria transmission intensity, with an estimated annual entomological inoculation rate of 125 infectious bites per person per year.[65] Since 2015, multiple rounds of insecticide spraying (IRS) have dramatically reduced malaria incidence compared with pre-IRS levels.[45] Individuals were selected for inclusion into this study based on age and *P. falciparum* exposure. In addition, children were selected based on the availability of samples collected at set times relative to malaria (Table S1). All individuals included in this study were enrolled in The Program for Resistance, Immunology, Surveillance, and Modeling of Malaria (PRISM) program[66] and have provided written consent for the use of their samples for research. The PRISM cohort study was approved by the Makerere University School of Medicine Research and Ethics Committee (SOMREC), London School of Hygiene and Tropical Medicine IRB, the University of California, San Francisco Human Research Protection Program, and the Stanford University School of Medicine IRB. The use of cohort samples for this study was approved by the Institutional Review Board of the University of Texas Health Science Center at San Antonio.

METHOD DETAILS

B cell isolations

Cryopreserved PBMCs were thawed in a 37°C water bath and immediately mixed with pre-warmed thawing medium (IMDM/GlutaMAX supplemented with 10% heat-inactivated FBS (USA origin) and 0.01% Universal Nuclease (Thermo, cat. no. 88700). After centrifugation (250 × g, 5 min), the cell pellet was resuspended in thawing medium. A sample of cells was mixed 1:1 (v/v) with 0.2% trypan blue solution in PBS, and viable cells were counted using a Cellometer Mini (Nexcelom) automated cell counter. Next, cells were centrifuged (250 × g, 5 min) and resuspended in isolation buffer (PBS supplemented with 2% heat-inactivated FBS and 1 mM EDTA) at 50 million live cells/mL and filtered through a 35 μm sterile filter cap (Corning, cat. no. 352235) to break apart aggregated cells. B cells were isolated by negative selection using the EasySep Human B Cell Isolation Kit (StemCell, cat. no. 17954) or the MojoSort Human Pan B Cell Isolation Kit (BioLegend, cat. no. 480082) according to the manufacturer's instructions.

Flow cytometry staining for single-cell analysis

Isolated B cells were washed with PBS (250 × g, 5 min), resuspended in 1 ml of PBS containing 1 μl live/dead stain (LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Thermo Fisher Scientific, cat. no. L34965), and incubated for 30 min on ice. Subsequently, cells were washed with cold PBS containing 1% bovine serum albumin (BSA) (250 × g, 5 min) and incubated in labeling buffer (total volume of 100 μl) containing PBS, 1% BSA, 5 μl Human TruStain FcX (Biolegend, cat. no. 422301) and 1% dextran sulfate (Sigma-Aldrich, cat. no. 42867-5G) for 10 min at 4°C before

incubating a further 30 min at 4°C with an antibody cocktail to label B cell surface markers. Details of all antibodies used are given in the key resources table. Before acquisition on a BD FACS AriaII cell sorter, cells were washed with cold PBS containing 1% BSA, centrifuged (250 × g, 5 min, 4°C), diluted to 20-30 million cells/mL in cold PBS containing 1% BSA, and filtered into a FACS tube through a 35 μm sterile filter cap (Corning, cat. no. 352235) to break apart aggregated cells. Lymphocytes were gated using forward and sideward scatter, followed by doublet exclusion. Cells were then gated on live cells, followed by $CD19^{+}CD10^{-}CD20^{+}$ B cells. Naive B cells ($CD21^{+}CD27^{-}$) and antigen-experienced B cells (all other B cells) were separated based on CD21 and CD27 expression and sorted into cold PBS containing 1% BSA. An average of 37,000 antigen-experienced B cells were mixed with naïve B cells up to a total of 50,000 B cells (Table S2). This resulted in an enrichment of antigen-experienced cells from an average of 49% in the original B cell pool to an average of 74% in the final cell suspension. Cells were immediately processed for single-cell sequencing.

Single-cell gene expression, V(D)J, and feature barcoding library preparation

B cells were counted with a hemocytometer and run on a Chromium 10x controller (10x Genomics) in individual lanes with a targeted recovery rate of 10,000 cells per lane, according to the manufacturer's instructions. Library preparation was completed following the recommended protocols for Chromium Single Cell 5′ Gene Expression kit as well as 5′ Feature Barcode and V(D)J Enrichment kits. Details of all kits used are given in the key resources table. Libraries were sequenced using the Illumina NovaSeq6000 (paired-end 150 bp) and HiSeq (paired-end 100 bp) platforms.

Atypical B cell sorting

B cells were isolated by negative selection as described above, washed with PBS (250 × g, 5 min), and resuspended in 1 ml of PBS containing 1 μl live/dead stain (LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Thermo Fisher Scientific, cat. no. L34965) and incubated for 30 min on ice. Cells were subsequently washed with cold PBS containing 1% BSA (250 × g, 5 min, 4°C) and incubated at 4°C for 30 min with a B cell surface marker antibody cocktail, added up to a total volume of 100 μl with PBS containing 1% BSA. Cells were then washed with cold PBS containing 1% BSA, centrifuged (250 × g, 5 min, 4°C), diluted to 20-30 million cells/mL in cold PBS containing 1% BSA, and filtered into a FACS tube through a 35 μm sterile filter cap (Corning, cat. no. 352235) to break apart aggregated cells. Using a BD FACS Aria Fusion cell sorter, atypical B cell subsets and switched memory cells were sorted into 1.5 mL Eppendorf tubes containing IMDM/GlutaMAX supplemented with 10% heat-inactivated FBS (USA origin). Cells were immediately frozen for RNA-seq or plated for *in vitro* B cell activation experiments.

Bulk RNA-Sequencing

RNA-seq libraries were generated following the NEBNext Single Cell/Low Input RNA Library Prep Kit for Illumina (NEB #E6420S) protocol for cells, except that all DNA cleanup steps we done using SparQ magnetic beads (Quantabio). In brief, 200 cells of each B cell subset were FACS sorted directly into 1× NEBNext Cell Lysis Buffer in 1.7 ml low binding tubes (Corning #3207). Cells were then flash frozen in liquid nitrogen for future use. For all samples, cDNA was synthesized and amplified using 21 cycles. cDNA quantity and quality were assessed by Invitrogen Qubit and Agilent 2100 Bioanalyzer. PCR enrichment of adaptor-ligated DNA was performed using NEBNext Multiplex Oligos for Ilumina (NEB #E7335S and #E7500S) following the recommended number of amplification cycles based on the amount of cDNA input used in the fragmentation/end prep reaction. Library quantity and quality were assessed by Invitrogen Qubit and Agilent 2100 Bioanalyzer. Eighteen libraries were multiplexed and sequenced on a NovaSeq600 SP 50PE flow cell. Six libraries were retrospectively removed from data analysis due to low input cDNA and resulting low gene expression quality.

Spectral flow cytometry

C-terminally biotinylated full-length *P. falciparum* 3D7 MSP1 and AMA1 were produced as described previously.[47] Antigen tetramers were synthesized by incubating protein with fluorophore-conjugated streptavidin overnight at 4°C at a molar ratio of 6:1 with rotation. B cells were isolated by negative selection as described above, washed with PBS, centrifuged (250 × g, 5 min), resuspended in 1 ml of PBS containing 1 μl live/dead stain (Zombie UV Fixable Viability kit (Biolegend, cat. no. 423107)) and incubated on ice for 30 min. Cells were subsequently washed with cold PBS containing 1% BSA (250 × g, 5 min, 4°C), resuspended with a cocktail of 25 μM of each antigen tetramer diluted in PBS containing 1% BSA to a volume of 100 μl, and incubated at 4°C for 30 min. Next, the cells were washed twice with cold PBS containing 1% BSA (250 × g, 5 min, 4°C) and incubated at 4°C for 30 min with a B cell surface marker antibody cocktail with 10 μl Brilliant Stain Buffer Plus (BD, cat. No. 566385) diluted in PBS containing 1% BSA up to a volume of 100 μl. The cells were then washed with cold PBS containing 1% BSA (250 × g, 5 min, 4°C), resuspended in 1 ml of Transcription Factor Fix/Perm Concentrate (Tonbo, part of cat. no. TNB-0607-KIT), diluted with 3 parts Transcription Factor Fix/Perm Diluent (Tonbo), and incubated at 4°C for 1 hour. After the incubation, the cells were washed twice with 3 ml of 1× Flow Cytometry Perm Buffer (Tonbo) (300 × g, 8 min, 4°C) and resuspended in 1× Flow Cytometry Perm Buffer with an intracellular marker antibody. After an incubation at 4°C for 30 min, the cells were washed twice with 3 ml cold 1× Flow Cytometry Perm Buffer (300 × g, 8 min, 4°C) and once with 3 ml cold PBS containing 1% BSA, resuspended in cold PBS containing 1% BSA to 20 – 30 million cells/ml and filtered into a FACS tube through a 35 μm sterile filter cap. Cells were analyzed by flow cytometry immediately following intracellular staining.

B cells were analyzed on a Cytek Aurora spectral flow cytometer equipped with five lasers. SpectroFlo QC Beads (Cytek, cat. no. SKU N7-97355) were run prior to each experiment for performance tracking. Quality control and LJ tracking reports were used to ensure machine performance and that settings between different runs were comparable. Paired samples were processed in the same experiment and experiments performed on separate days contained technical replicates. B cells isolated from pooled PBMCs from two malaria-naïve US donors

were used for the compensation of the live/dead stain, for the unstained control, and for technical replicates. Between runs, the expression of cell surface and intracellular markers showed a high correlation between these technical replicates (Spearman r = 0.98, Figure S23). UltraComp eBeads Plus Compensation Beads (Thermo, cat. No. 01-3333-41) were used for compensation of all other fluorophores.

B cell cultures

One day prior to sorting B cells, wells of a 96-well plate were each seeded with 30,000 adherent, CD40L-expressing 3T3 cells (kind gift from Dr. Mark Connors, NIH) in 100 μl IMDM/Glutamax/ 10% FBS containing 2× MycoZap Plus-PR (Lonza #VZA-2021), 100 ng/ml human IL-2 (GoldBio #1110-02-50), and 100 ng/ml human IL-21 (GoldBio #1110-21-10) to promote expansion and differentiation of B cells into antibody-secreting cells. As a control, one plate was prepared with 100 μl IMDM/Glutamax/ 10% FBS containing 2× MycoZap Plus-PR, 100 ng/ml human IL-2, and 100 ng/ml human IL-21, without 3T3 cells. Plates were incubated O/N at 37°C and 8% CO_2. Immediately after B cell sorting, 100 cells of each B cell subset were resuspended in 100 μl IMDM/Glutamax/ 10% FBS, added to the 100 μl culture media with supplements already present in the plates to a total volume of 200 μl, and incubated at 37°C and 8% CO_2. After seven days, 80 μl of supernatant was removed and replenished with 100 μl IMDM/Glutamax/ 10% FBS containing 2× MycoZap Plus-PR, 100 ng/ml human IL-2, 100 ng/ml human IL-21. At day 14, the IgG concentration in the supernatant was determined by enzyme-linked immunosorbent assay (ELISA).

Enzyme-linked immunosorbent assays

To detect IgG, 96-well ELISA plates (Corning #3361) were coated with goat anti-human IgG (Sigma #I2136) antibody at a concentration of 4 μg/ ml diluted in PBS, at a total volume of 100 μl per well. After a one-hour incubation at 37°C or O/N at 4°C, each well was washed once using slowly running (approximately 900 ml/min.) deionized water. All subsequent washes were performed this way. 150 μl blocking buffer (one-third Non-Animal Protein (NAP)- Blocker (G-Biosciences #786-190P) and two-thirds PBS) was added to each well to prevent non-specific binding. After one hour of incubation at 37°C, the wells were washed three times and 5 μl B cell culture supernatant diluted 1:20 in dilution buffer (1% NAP Blocker in PBS; total volume 100 μl) was added per well. Plates were incubated for two hours at 37°C and washed five times. Then, 100 μl 1:2500 diluted (1% NAP Blocker in PBS) HRP-conjugated anti-human IgG antibody (BioLegend #410902) was added to each well. After incubation for one hour at 37°C and three washes, HRP activity was detected using 50 μl TMB (Thermo #PI34024). Plates were incubated in the dark at RT and the oxidation reaction was stopped by adding 50 μl 0.18 M H_2SO_4 (Fisher #FLA300-212) per well when the negative controls (wells that received buffer when test wells received culture supernatant) started to color. Absorbance was measured at 450 nm using a BioTek Synergy H4 microplate reader. A human IgG (Sigma #I2511) standard curve (ten three-fold serial dilutions starting at 20 μg/ml) was used to quantify samples.

QUANTIFICATION AND STATISTICAL ANALYSIS

Single-cell RNA-seq data analysis

Sequencing reads were mapped to the human reference genome (hg38) and assigned to individual cells of origin using cell-specific barcodes in Cell Ranger (10x Genomics) version 3.1.0. The sequencing data were then processed using Seurat version 3.2.3, following their default pipeline as follows. Low-quality cells with high (>2500) or low (<200) UMI counts were filtered from the samples. In addition, to prevent antibody isotype or V(D)J-gene segment usage from influencing the clustering of cells, we removed all IGH, IGK and IGL transcripts from the single cell data before clustering and UMAP visualization, as proposed by Steward et al.[20] Gene expression levels for each cell were normalized by total expression, multiplied by a scale factor (10,000), and log-transformed. Cells with > 5% mitochondrial reads were excluded because these are likely damaged or dying cells that have lost much of their cytoplasmic mRNA. Gene expression values for all genes in all clusters are provided in Table S7. We determined the 2000 most variable genes based on their average expression and dispersion. Principal component analysis was performed on highly variable genes to reduce the dimensionality of the data. We performed clustering using the Louvain algorithm implemented in the *FindClusters* function of Seurat (algorithm = 1, resolution = 0.6, dims. use = 1:20) and displayed the clusters using UMAP. One outlier cluster consisting of 13 cells that were most likely T cells was removed from further analysis. Because cells from all samples were distributed evenly across the UMAP (Figure S2), no batch corrections were performed. We identified the cluster-specific marker genes conserved across different samples using the *FindConservedMarkers* function (default parameters) from Seurat. To identify differentially expressed genes between a pair of clusters, we used the *FindMarkers* function from Seurat (with parameters *min.pct* = 0.20, and log2FC = 0.25), and additionally filtered genes using a false discovery rate of 0.05. Violin plots, dot plots, and UMAP plots overlaid with gene expression level were generated using standard Seurat functions. Reactome pathway analysis was performed using differentially expressed genes that were upregulated in pairwise cluster comparisons. The top 3 - 5 pathways with FDR-adjusted P value < 0.01 were selected for visual representation in Figure 3C. The full set of differentially expressed gene lists and Reactome analysis results are provided in Tables S5 and S6, respectively. For heatmaps in Figure 3, gene expression values for the four samples were first divided by the average value for all clusters for that donor to account for donor-to-donor variation. The data points were then z-scored across all clusters and samples prior to plotting heatmaps using the color palette 'Hiroshige' from the R package MetBrewer.

Single-cell RNA-sequencing gene signature score analysis

We performed gene signature score analysis as described by Holla et al.[11] Briefly, gene expression data from the Seurat object were scaled (z-score across each gene) to remove bias towards highly expressed genes. A cell-specific signature score was computed by first sorting the

normalized scaled gene expression values for each cell followed by summing up the indices (ranks) of the signature genes. These signature scores were then projected onto the transcriptomics-based UMAP.

Single-cell surface marker analysis

From the single-cell RNA sequencing Seurat object, we extracted the antibody capture component (assay = ADT), normalized the antibody matrix using CLR normalization (Seurat method *NormalizeData*), and then scaled the antibody matrix using Seurat routine *ScaleData*. This normalized and scaled antibody matrix was then used to define the following populations of B cells by phenotype: 1) DN1 B cells: $CD27^{-}IgG^{+}CD21^{+}CD11c^{-}$, 2) DN2 B cells: $CD27^{-}IgG^{+}CD21^{-}CD11c^{+}$, 3) DN3 B cells: $CD27^{-}IgG^{+}CD21^{-}CD11c^{-}$, 4) DN4 B cells: $CD27^{-}IgG^{+}CD21^{+}CD11c^{+}$, 5) Resting naïve B cells: $CD27^{-}IgM^{+}CD21^{+}$, 6) Activated naïve B cells: $CD27^{-}IgM^{+}CD21^{-}$, 7) Resting switched memory B cells: $CD27^{+}IgG^{+}CD21^{+}$, 8) Activated switched memory B cells: $CD27^{+}IgG^{+}CD21^{-}$, 9) Resting unswitched memory B cells: $CD27^{+}IgM^{+}CD21^{+}$, and 10) Activated unswitched memory B cells: $CD27^{+}IgM^{+}CD21^{-}$. The UMAP plots highlighting corresponding cell categories were generated using Seurat. Distributions of cell surface expression of the markers CD21 and CD27 (Figure S5), as well as CD11c and CXCR3 (Figure S7) for individual clusters were plotted using the above mentioned normalized and scaled antibody matrix, and employing the 2D kernel density function *kde2d* from the R package *MASS*.

Re-analysis of single-cell RNA-seq data sets for HIV and SLE samples

The HIV scRNA-seq data (CellRanger output feature barcoded matrix) of memory B cells (n = 70 samples) and memory B cell-depleted samples (n = 3 samples) were downloaded from GEO: GSE229232.[37] First, we ran Scrublet[61] for doublet detection. Then we used Seurat[58] to process individual samples and retain only those cells which are not doublets, have nFeature_RNA > 200, percent.mt < 10 and percent.ribo > 4. After filtering the remaining cells from individual samples, we merged all of them using the *merge()* routine of Seurat, performed normalization using the method *LogNormalize* with a scaling factor of 10000, obtained top 2000 variable features using the function *FindVariableFeatures*, scaled data (using the routine *ScaleData*()) by regressing with respect to percent.mt, and executed PCA with a dimension of 30. Batch correction was performed by Harmony,[62] using months post infection (mpi) values as the covariate. The function *FindNeighbors*() with k=20 was used to find the nearest neighbors. Clustering was performed by the Louvain algorithm and using resolution = 0.3. Output clusters were displayed using UMAP. From the clusters, memory B cells (from 70 samples) were extracted and re-clustered using the same parameters (Louvain algorithm, resolution = 0.3). Clusters 0 and 8 were annotated as atypical B cells based on similarities in gene expression profile with DN2 B cells reported by Stewart et al.[20] These two clusters were then extracted and re-clustered using a resolution of 0.5. Dot plots were generated using the default Seurat functions.

For SLE, we downloaded scRNA-seq data from GEO: GSE135779.[36] The data contains 58 samples (33 child SLE, 11 child HD controls, 8 adult SLE, 6 adult HD controls). After running Scrublet for doublet detection, we employed Seurat and retained only those cells from individual samples which are not doublets, have nFeature_RNA > 400, nFeature_RNA < 2500, and percent.mt < 20. Then we followed similar merging procedure as mentioned above. Harmony was used for batch correction, using the condition healthy vs SLE as the covariate. Clustering was performed using a resolution of 0.3. Using the B cell marker *MS4A1*, we extracted B cells and re-clustered these cells using a resolution of 0.4. We then identified atypical B cells in cluster 3 using the gene expression pattern of DN2 B cells from Stewart et al.[20] This cluster was then extracted and re-clustered using a resolution of 0.4. Dot plots were generated using the default Seurat functions.

Single-cell B cell receptor analysis

Single-cell B cell receptor sequencing data was first processed using the command *changeo-10x*. Single-cell heavy chain variable region sequences were then analyzed using the Immcantation pipeline. First, the Immcantation package Change-O (v0.4.4) was used to align lymphocyte receptor sequences to germline sequences for downstream analyses.[60] Isotype-specific data files were converted into standardized tab-delimited database files required for subsequent Change-O modules to operate using IgBLAST (v1.14.0), which is included in Change-O. The frequency of somatic hypermutations was calculated as described previously.[23] Average frequencies for each donor and cluster are reported.

To calculate the percentage of clonal expansion within clusters or clonal connections between clusters, genotyping was performed using the routine *tigger-genotype*, followed by analysis of clonally related B cell receptor sequences using a custom Perl script available for download at https://github.com/embunnik/clonal_scores, as described in detail elsewhere.[41]

RNA velocity analysis

RNA-velocity analysis was performed by scVelo,[40] using the scRNA-seq data. First, gene expression matrices from cell ranger were converted to scVelo compatible loom format using the tool Velocyto.[39] We exported the cluster information and the UMAP coordinates from the Seurat object in .h5ad format, instead of re-computing them in scVelo. RNA-velocity analysis was then performed by the scVelo routine *tl.velocity()* using its default stochastic model. RNA velocity and pseudo-time plots were generated by the default scVelo routines.

Bulk RNA-seq analysis

RNA-seq data was processed by our custom pipeline (https://github.com/ay-lab/RNA_SEQ_PIPELINE). Input fastq files were processed by fastp[63] for QC analysis. Paired-end RNA-seq reads were aligned to the hg38 reference genome using the default parameters of STAR (version 2.7.3a).[64] The GENCODE v28 GTF annotation file was used as a reference. The list of genes that were differentially expressed between the

three atypical B cell clusters in the scRNA-seq data set and were representative of functional differences between these clusters was filtered for an average expression level of >100 transcripts per million (tpm) among the 12 samples that passed post-sequencing QC. Expression values were z-scored across these 12 samples and plotted in a heatmap.

Spectral flow cytometry analysis

The cytometry analysis software OMIQ (Dotmatics) was used for the integration and dimension reduction analysis. In short, atypical B cells were first pre-gated on single/live/$CD19^+$/$CD20^+$/$CD21^-$/$CD27^-$/$CD11c^+$/IgG^+ cells. UMAPs were then created using the expression of CD19, CD20, CD24, CD38, CD83, CD86, CD95, CXCR3, CXCR5, CD11c, FcRL5, and T-bet as features with default parameters (neighbors = 15, minimum distance = 0.4, metric = Euclidean, random seed = 9258) and included all 38 Ugandan donor samples used in this study for initial projection. To define the three atypical B cell subsets, FlowSOM[67] was used to identify three clusters based on CD11c and CD86 expression using the default parameters (metric = Euclidean, random seed = 1226). For the projection of antigen-specific B cells onto the UMAP, gates were manually set to identify populations of interest using two-dimensional displays, which were then overlaid onto the UMAP projection. Mean fluorescence intensities of cell surface and intracellular markers in select B cell subsets were calculated for each sample individually in OMIQ, and then averaged across samples. To visualize mean fluorescence intensities in a heatmap, the average value for each B cell subset was expressed as a percentage of the highest mean fluorescence intensity among all subsets.

Statistical analysis

Statistical analyses of mutation frequency and flow cytometry data were performed in GraphPad Prism 9 with details of statistical tests in the relevant figure legends.