**Title: Single-cell sequencing identifies novel atypical B cell subsets with distinct effector functions**

**Authors:** Raphael A. Reyes[1†], Gayani Batugedara[1†], Paramita Dutta[2], Ashley B. Reers[1], Rolando Garza[1], Isaac Ssewanyana[3,4], Margaret E. Feeney[5,6], Bryan Greenhouse[5], Sebastiaan Bol[1], Ferhat Ay[2,7], Evelien M. Bunnik[1*]

**Affiliations:**

[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, USA

[2] Division of Vaccine Discovery, La Jolla Institute for Immunology, La Jolla, CA, USA

[3] Infectious Disease Research Collaboration, Kampala, Uganda

[4] Department of Infection Biology, London School of Hygiene and Tropical Medicine, London, United Kingdom

[5] Department of Medicine, University of California, San Francisco, CA, USA

[6] Department of Pediatrics, University of California, San Francisco, CA, USA

[7] School of Medicine, University of California San Diego, La Jolla, CA, USA

*Correspondencing author. Email: bunnik@uthscsa.edu

[†] These authors contributed equally to this work

**Abstract:** Atypical B cells are a population of activated B cells that are commonly enriched in individuals with chronic immune activation. Prior studies to determine the function of these cells have yielded conflicting results, possibly due to functional heterogeneity among this B cell population. To better define the role(s) of atypical B cells in the host adaptive immune response, we performed single-cell sequencing of transcriptomes, cell surface markers, and B cell receptors in individuals with chronic *Plasmodium falciparum* exposure, a condition known to lead to accumulation of circulating atypical B cells. Our studies identified previously uncharacterized populations of atypical B cells with distinct transcriptional and functional profiles: i) anergy or tolerance, ii) increased ability to interact with T cells, and iii) preparation for high translational activity as seen in pre-antibody secreting cells. We identified a set of cell surface markers to distinguish these distinct atypical B cell subsets and confirmed their presence in malaria-experienced children and adults using flow cytometry. *P. falciparum*-specific cells were present in equal proportions within each of these populations, indicating that all three subsets develop in response to antigen stimulation. Collectively, our findings help explain the conflicting observations in prior studies involving the functions of atypical B cells and indicate the need for reclassification of these cells into multiple distinct subsets to better understand their role in the adaptive immune response in chronic inflammatory conditions.

**One sentence summary:** Atypical B cells consist of distinct subsets that play different roles in the host adaptive immune response

45 **Main Text:**

46 **INTRODUCTION**

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

66 Studies regarding the functionality of atypical B cells in the immune response have highlighted
67 various differences between atypical B cells and memory B cells. Specifically, atypical B cells
68 displayed reduced expression of B cell receptor signaling pathway genes and costimulatory

69 molecules, as well as upregulation of inhibitory receptor expression (*10, 11*). In addition, atypical
70 B cells were less responsive to B cell receptor engagement by soluble antigen and showed
71 reduced differentiation into antibody-secreting cells under conditions that efficiently induced
72 differentiation of memory B cells into antibody-secreting cells (*5, 8-10, 12*). These results initially
73 led to the conclusion that atypical B cells are exhausted and non-responsive to stimulation.
74 However, it was recently reported that atypical B cells responded robustly to high-affinity
75 membrane-associated antigen and showed a markedly reduced response to low-affinity antigens,
76 a selective mechanism which may serve to limit activation of atypical B cells by low-affinity
77 antigens, such as auto-antigens (*11, 13*). Despite these new findings, the contribution of atypical
78 B cells to the host immune response remains subject of debate. Various functions have been
79 proposed for these cells, including (pre-)antibody secreting cells, based on the presence of
80 secretory Ig transcripts, and immune-modulatory cells, based on the upregulation of surface
81 proteins that promote interactions between B cells and T cells (*11, 14-17*). The multiple and
82 conflicting effector functions previously assigned to atypical B cells are suggestive of functional
83 heterogeneity among this population.

84 Here, we used chronic *P. falciparum* exposure as a model to define the heterogeneity of the
85 atypical B cell compartment. Over the course of many years of repetitive *P. falciparum* infections,
86 people living in malaria-endemic regions develop an immune response that protects against
87 disease. As a result, most cases of malaria occur in children under the age of 10, while adults
88 with life-long exposure have asymptomatic infections. To study atypical B cells induced by
89 malaria, we performed single-cell transcriptomics analysis with cell surface marker and B cell
90 receptor profiling on longitudinal samples obtained from two children following malaria. Based on
91 these sequencing data, we defined three subsets of atypical B cells with distinct transcriptional
92 profiles that are likely to play different roles in the immune response. Using flow cytometry, we

93 then analyzed the longevity and *Plasmodium* antigen-specificity among these subsets in samples
94 from malaria-experienced children and adults.

95

96 **RESULTS**

97 **Single-cell transcriptomics identifies B cell populations:**

98 The atypical B cell population is strongly expanded in the first few weeks after a malaria episode
99 and slowly contracts to baseline levels over the course of several months (*4*). To study the
100 identity of atypical B cells induced as a result of the hyperinflammatory immune response during
101 malaria and following return to baseline conditions, we selected two children (5 and 7 years old)
102 for whom peripheral blood mononuclear cell (PBMC) samples were collected three weeks after
103 malaria and six months later (**Figure 1A**, **Table S1**). Both children were free of symptomatic
104 malaria between the two time points, but had detectable *P. falciparum* parasitemia at both the
105 three-month and six-month time points (**Figure S1**). Asymptomatic *P. falciparum* infections
106 result in very low immune activation (*18*), and it can therefore be expected that the atypical B
107 cell population is only minimally affected by these infections, especially in comparison to the
108 strong inflammatory response during malaria.

109 We first isolated naïve and antigen-experienced B cells by fluorescence-activated cell sorting
110 based on the expression of CD21 and CD27 (**Figure 1B**). Naïve B cells typically make up a
111 large fraction of the total B cell pool (43 – 60% in the samples used here), but are not of interest
112 when studying the response of B cells to infection. Therefore, we enriched our samples for
113 antigen-experienced B cells, but added in some naïve B cells to have representation of all B cell
114 populations in the final cell pool. (**Table S2**). Each sample was then subjected to single-cell
115 RNA-sequencing in combination with analysis of select cell surface marker expression and B

116 cell receptor-sequencing on the 10x Genomics platform. Cells that passed quality control (n =
117 21,439; **Table S3**) were clustered into nine populations based on gene expression and
118 visualized by Uniform Manifold Approximation and Projection (UMAP) (**Figure 1C**). UMAP
119 visualizes cell clusters in a 2D plot based on similarity in transcriptional profile. Each of the four
120 samples was well represented in all nine clusters (**Figure S2; Table S4**). Notably,
121 immunoglobulin-related genes were excluded during the clustering to prevent antibody isotype
122 or V(D)J-gene segment usage from influencing these results, as was previously proposed by
123 Stewart *et al*. (*19*). Comparing the expression levels of individual key genes (**Figure 1D**), as well
124 as gene signatures of various B cell populations generated by others (**Figure S3**) (*11, 19, 20*),
125 allowed us to identify the location of major B cell populations in the UMAP. This included
126 transitional B cells in cluster 7, naive B cells (expressing *BACH2*) in cluster 2, and memory B
127 cells (expressing *CD27*) in clusters 0, 3, 5, 6, and 8. In total, 82% of all cells in our data set were
128 found in clusters with a non-naive transcriptional signature, confirming that our data set was
129 enriched for antigen-experienced cells. Clusters 1, 4, and 5 expressed genes associated with
130 atypical B cells, such as *ZBTB32*, encoding a transcription factor, and *ITGAX,* encoding the
131 surface protein CD11c (**Figure 1D**). Together, these two clusters represented almost 30% of all
132 data (6,088 cells), making this a rich dataset to study heterogeneity among atypical B cells in
133 malaria-experienced individuals.

134

135 **Phenotypically defined atypical B cells have multiple distinct transcriptomic profiles:**

136 We next sought to determine how well the atypical B cell clusters that we identified based on
137 transcriptomic profiles corresponded to $CD21^{-}CD27^{-}$ B cells, which is the phenotypic
138 classification for atypical B cells most commonly used in the malaria field. We therefore used
139 cell surface markers CD21 and CD27 to 'gate' cells, similar to how this is commonly done in

140 flow cytometry (**Figure S4**). Nearly all cells in clusters 1 and 4 were $CD21^{-}CD27^{-}$, in agreement
141 with their transcriptomic profile of atypical B cells (**Figure 2A, Figure S5**). However, $CD21^{-}$
142 $CD27^{-}$ B cells also made up about 50% of cluster 5. Cluster 7 was also comprised of
143 approximately 50% $CD21^{-}CD27^{-}$ cells, which were transitional B cells that have recently
144 emigrated from the bone marrow and have yet to upregulate CD21 expression (**Figure S3A**).
145 These cells were therefore excluded from our analysis. In a recent study by Sutton *et al.*,
146 atypical B cells were observed to predominantly express CD11c and CXCR3 (*3*). We therefore
147 also analyzed the distribution of $CD11c^{+/-}$ $CXCR3^{+/-}$ populations among the nine B cell clusters
148 (**Figure S6**). The large majority of cells in clusters 1 and 4 expressed CD11c (75% and 98%,
149 respectively), while expression of CXCR3 was observed in only ~45% of all cells in these
150 clusters (**Figure 2B, Figure S5**). Instead, 79% of cells in cluster 5 were $CXCR3^{+}$, most of which
151 also expressed CD11c. These percentages were similar when limiting the analysis to CD21-
152 CD27- B cells (**Figure S7**). These data suggest that CXCR3 marks a subset of atypical B cells
153 but cannot be used as a marker to identify all atypical B cells. Next, we analyzed the heavy
154 chain sequences from the B cell receptor-sequencing libraries to determine isotype usage
155 among the different clusters. This analysis showed that, as expected, the transitional cells in
156 cluster 7 and naïve B cells in cluster 2 were unswitched cells, expressing only *IGHM* and *IGHD*.
157 Cluster 1 was also mainly made up of unswitched B cells, while clusters 3, 4, and 5 contained
158 larger fractions of class-switched B cells (**Figure 2C, Figure S5**). Cluster 4 in particular showed
159 a similar distribution of *IGH* transcripts to cluster 3 memory B cells, with the exception of a lack
160 of *IGHA1* expression, suggesting that cluster 4 atypical B cells have undergone class-switch
161 recombination at similar rates as memory B cells.

162 To further delineate the distribution of different B cell populations among our transcriptomics
163 clusters, we identified B cells by phenotype based on the surface expression of CD21, CD27,
164 CD11c, IgM, and IgG following recently published guidelines (*21*). Because our cell surface

165 marker analysis did not include IgD, we made two exceptions to these guidelines. First,
166 unswitched B cells were not identified based on the presence of IgD but that of IgM, which is
167 usually co-expressed with IgD. Second, switched B cells were identified based on the presence
168 of IgG instead of the absence of IgD, isotypes that are mutually exclusive. We first used CD27
169 and IgM or IgG expression to distinguish naive B cells (NAV, $CD27^-IgM^+$), unswitched memory B
170 cells (unswM, $CD27^+IgM^+$), and switched memory B cells (swM, $CD27^+IgG^+$). These populations
171 were further divided into resting ($CD21^+$) and activated ($CD21^-$). The majority of resting switched
172 memory B cells was located in cluster 3. We also identified DN B cells ($CD27^-IgG^+$), which were
173 divided into DN1 ($CD21^+CD11c^-$), DN2 ($CD21^-CD11c^+$), DN3 ($CD21^-CD11c^-$), and DN4
174 ($CD21^+CD11c^+$) (**Figure S8**). $IgM^+$ atypical B cells are also known as activated naive B cells,
175 and $IgG^+$ atypical B cells are equivalent to DN2 B cells. We recently reported substantial
176 differences in intrinsic properties between $IgM^+$ and $IgG^+$ atypical B cells that are likely related to
177 their origin and function (*22*). Additionally, CD11c has been identified as the marker that best
178 delineates atypical B cells (*3, 23*). We therefore decided to use the phenotypic definition of DN2
179 B cells to identify atypical B cells for the remainder of this study: $CD21^-CD27^-IgG^+CD11c^+$
180 (**Figure 2D**). Of all phenotypically defined atypical B cells, 40% was found in cluster 1, 40% in
181 cluster 4, and ~20% in cluster 5 (**Figure 2E**). The majority (~80%) of cells in cluster 4 were
182 atypical B cells, whereas cluster 1 was more mixed, with ~40% atypical B cells and ~30%
183 activated naïve B cells (**Figure 2E**). Cluster 5 was more diverse with 23% atypical B cells, 19%
184 activated naïve B cells, 13% resting switched memory B cells, 16% activated switched memory
185 B cells, and 29% other B cell populations. Collectively, these results show that phenotypically
186 defined class-switched atypical B cells in malaria-experienced children consist of multiple
187 distinct subsets with unique transcriptional profiles.

188

189 **Differences in gene expression programs of atypical B cells are associated with different**
190 **roles in the immune response:**

191 To better understand the identity of the three subsets of atypical B cells identified in this study
192 found in clusters 1, 4, and 5, we compared the expression of genes that have previously been
193 described as part of the unique transcriptional profile of atypical B cells (*10-12*). The up or down
194 regulation of various genes in these clusters as compared to resting switched memory B cells in
195 cluster 3 was consistent with previous reports (**Figure 3A**). Two manuscripts recently posted on
196 bioRxiv showed that the transcription factor *ZEB2* acts as an important regulator for the
197 development of atypical B cells (*24, 25*). ZEB2 was also expressed significantly higher in
198 clusters 1, 4, and 5 as compared to cluster 3 (**Figure 3A**). In contrast, several atypical B cell-
199 associated genes were uniquely expressed in only one of the three clusters (**Figure 3B**). For
200 example, *CD72*, encoding a surface marker that inhibits B cell responses and prevents antibody
201 production (*26*), was more abundantly expressed in cluster 1. In contrast, the integrin-encoding
202 genes *ITGB2* and *ITGB7* were predominantly expressed in cluster 4. *CXCR3* expression was
203 highest in cluster 5, in line with the results of the cell surface marker analysis (**Figure**
204 **2B**). These differences in gene expression highlight that bulk atypical B cells are a mixture of
205 various subsets and point to functional differences in activation status and the potential for cell-
206 cell interactions between these subsets.

207 To gain insight into potential differences in function between cells in clusters 1, 4, and 5, we
208 performed Reactome pathway analysis using genes that were upregulated or downregulated in
209 pair-wise comparisons of the three clusters (**Figure 3C-D, Tables S5 – S6**). Atypical B cells in
210 cluster 5 upregulated 17 out of 99 genes encoding ribosomal proteins (for example, *RPS23*,
211 *RPS21*, and *RPS8*) and were most strongly enriched for pathways related to translation,
212 indicating that these cells may be poised to differentiate into antibody-secreting cells. Atypical B
213 cells in both clusters 1 and 4 upregulated genes in response to stimulation by interferon gamma

214 and cytokines, but the outcome of this response differed between the two clusters. Atypical B
215 cells in cluster 1 upregulated genes encoding components of the Activator Protein 1 (AP1)
216 complex, including *FOS* and *JUN*, as well as interferon-stimulated genes (for example, *ISG20*),
217 and tubulin genes (for example, *TUBB6*). Additionally, cluster 1 atypical B cells were enriched
218 for pathways that signified negative regulation of cell cycle progression. Upregulated genes
219 included the cyclin-dependent kinase inhibitors *CDKN1A/B*, the anti-proliferative genes *BTG1/2*,
220 and the transcriptional repressor *NR4A1*. Nuclear receptor 4A1 is part of a family of transcription
221 factors that was recently implicated in a negative feedback loop that rendered B cells more
222 dependent on T cell help and promoted peripheral B cell tolerance (*27, 28*). The transcriptional
223 profile of atypical B cells in cluster 1 thus pointed to an interferon-driven response that resulted
224 in senescence or anergy. On the other hand, atypical B cells in cluster 4 were enriched for
225 pathways related to antigen processing and presentation, and immunoregulation. Among the
226 genes upregulated in cluster 4 were genes encoding proteins that promote interactions between
227 B cells and T cells, such as HLA-DR/DQ, CD86, which is downregulated by NR4A1, and CD74,
228 which regulates antigen presentation by class II MHC (*29-31*). In addition, we observed
229 upregulation of genes encoding class I MHC (*HLA-A* and *HLA-C*) and several genes that are
230 associated with cytotoxicity, including granzyme M (*GZMM*), cathepsin H (*CTSH*), cystatin 7
231 (*CST7*) natural killer cell granule protein 7 (*NKG7*, a target gene of T-bet), and natural
232 cytotoxicity triggering receptor 3 (*NCR3*). This gene signature suggests that these cells have
233 increased exocytosis activity, which could affect cell-cell interactions and have an
234 immunomodulatory function. Of note, clusters 1, 4, and 5 were stable at lower clustering
235 resolution. At higher clustering resolution, cluster 1 was divided into two separate clusters that
236 contained atypical B cells and activated naïve B cells, respectively (**Figure S9**). The results of
237 differential gene expression and pathway analysis for the subcluster of atypical B cells were
238 comparable to those presented above. Collectively, our results indicate that phenotypically

239 defined IgG$^+$ atypical B cells can have one of three different transcriptional profiles that are
240 associated with different functions within the immune response.

241

242 **Atypical B cells form two separate compartments:**

243 To study the development of the different subsets of atypical B cells, we turned to an analysis of
244 their B cell receptor sequences. We assessed the amount of affinity maturation each B cell
245 population has undergone by determining the level of somatic hypermutation in each cluster.
246 This revealed that clusters 4 and 5 on average accumulated similar levels of mutations as class-
247 switched memory B cells in cluster 3, while the average mutation frequency in cluster 1 was
248 more than 50% lower (**Figure 4A, Figure S10A**). However, this may in part be caused by the
249 larger fractions of unswitched B cells in cluster 1 (**Figure 2C**). When including only class-
250 switched atypical B cells, we still observed a higher average level of somatic hypermutation in
251 clusters 4 and 5 than in cluster 1, although the difference was slightly smaller (**Figure 4B,**
252 **Figure S10B**). This difference suggests that cluster 1 atypical B cells have not undergone
253 extensive affinity maturation and seem to be more closely related to naive B cells. Atypical B
254 cells in clusters 4 and 5 have undergone similar levels of affinity maturation as memory B cells,
255 which could indicate that they are derived from memory B cells or have a similar developmental
256 trajectory. However, while somatic hypermutation and affinity maturation typically occur inside
257 germinal centers, these processes can also occur in extrafollicular sites (*32*) and our results are
258 therefore not direct evidence for the developmental pathway of atypical B cells.

259 Next, we wished to better understand the developmental relationship between the three
260 subsets. To this end, we calculated the level of expansion of clonal B cell lineages within each
261 cluster at the two time points, as well as the percentage of clonal connections between clusters,
262 as described in more detail previously (*33*). B cells were considered part of the same lineage if

263 their heavy chain variable region contained the same *IGHV* and *IGHJ* genes, and shared at
264 least 85% amino acid sequence identity in the heavy chain (H) CDR3 (*34, 35*). In all three
265 atypical B cell clusters, there was strong expansion of B cells at 3 weeks post-malaria, with 5 –
266 16% of B cells within each cluster belonging to the same lineage (**Figure 4C, Figure S11A**).
267 This was followed by contraction six months later, resulting in a loss of more than 50% of clonal
268 lineages (**Figure 4C, D**). This contraction points to a return to a resting state and confirms that
269 minimal immune activation occurred as a result of the asymptomatic infection at 6-months post-
270 malaria, as compared to the strong humoral immune response elicited by malaria. At 3 weeks
271 post-malaria, 1.6 – 4.1% of clonal sharing occurred between cluster 1 and clusters 4/5, but
272 these connections were largely lost six months later. In contrast, the connectivity between
273 cluster 4 and cluster 5 increased at the second time point, as compared to the time point shortly
274 after malaria, and this was the only pair of clusters for which this happened (**Figure 4D, Figure**
275 **S11B**). These observations suggest that cluster 1 atypical B cells form a separate compartment
276 from the other atypical B cell subsets, while cluster 4 and cluster 5 atypical B cells are more
277 strongly connected. Collectively, these results point to different developmental pathways for
278 atypical B cells that are connected to different functional roles in the immune response.

279

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

282 The identification of atypical B cell clusters by transcriptomics analysis was based on four
283 samples from two children. To validate our findings in a larger number of malaria-susceptible
284 children and extend our analysis to malaria-protected adults, we turned to an analysis of these
285 cells by spectral flow cytometry. In the single-cell transcriptomics analysis, we observed that
286 differences in expression level of *ITGAX* (encoding CD11c) translated to differences in cell

287 surface expression of CD11c (high expression in cluster 4, intermediate expression in cluster 1,
288 and low / intermediate expression in cluster 5, **Figure 5A**). Based on these data, we reasoned
289 that CD11c in combination with other differentially expressed cell surface markers could be
290 used to discriminate between the three clusters. We selected *CD86* and *FCRL5* that showed
291 consistent expression patterns across all four samples and would allow us to distinguish the
292 three clusters. Cluster 4 expressed high levels of *CD86* and *FCRL5*, cluster 1 expressed only
293 *FCRL5*, and cluster 5 expressed neither of the two (**Figure 5A**).

294 Using flow cytometry, we aimed to study the abundance of each atypical B cell subset and their
295 longevity in the absence of antigen stimulation and infection-driven inflammation. To this end,
296 we included samples from nine children and ten adults collected following known *Plasmodium*
297 exposure and months to years later after a period with low *Plasmodium* exposure due to highly
298 effective insecticide spaying (*36, 37*) (**Figure 5B, Table S1**). B cells were phenotyped using a
299 panel of 18 surface and intracellular markers that would allow us to identify the three atypical B
300 cell subsets and analyze these for the expression of markers previously shown to be positively
301 (CD19, CD20, CD11c, FcRL5, CD83, CD86, CD95, CXCR3, T-bet) or negatively (CD24, CD38,
302 CXCR5) associated with activated and atypical B cells (*21*). To find the atypical B cell subsets
303 identified in our transcriptomic analysis, we first gated manually on $CD19^{+}CD20^{+}CD21^{-}CD27^{-}$
304 $IgG^{+}CD11c^{+}$ atypical B cells (**Figure S12**). The atypical B cells from all samples were then
305 plotted in a composite UMAP based on the expression of the 12 cell surface and intracellular
306 markers previously shown to be associated with activated and atypical B cells (**Figure S13**).
307 Next, we used the unsupervised clustering tool FlowSOM to identify three distinct atypical B cell
308 subsets based on the expression of CD11c, FcRL5, and CD86 (**Figure 5C**). Projection of these
309 subsets onto the UMAP showed that the expression of *ITGAX*, *FcRL5*, and *CD86* matched the
310 cell surface expression signatures of subset 1 ($CD11c^{hi}FcRL5^{+}CD86^{-}$), subset 4
311 ($CD11c^{hi}FcRL5^{+}CD86^{+}$), and subset 5 ($CD11c^{int}FcRL5^{-}CD86^{-}$) (**Figure 5D**). We will refer to

312 these subsets with the number corresponding to the matching transcriptomics cluster, i.e., flow
313 cytometry subset 1 for cells matching transcriptomics cluster 1. Next, we determined the relative
314 abundance of these three subsets among all atypical B cells. In children, most atypical B cells
315 belonged to subset 1 (~50%) and subset 5 (~40%). The relative abundance of the subsets did
316 not change between time points (**Figure 5E**, **Figure S14**). This could be the result of the timing
317 of sample collection. At two months post-malaria in a low transmission setting, transient
318 changes to the atypical B cell population induced by malaria may have already reverted to
319 baseline. In adults, the first sample was collected during a period of high *Plasmodium* exposure.
320 Interestingly, the proportion of subset 4 atypical B cells was larger than in children at the first
321 time point but decreased to the level seen in children after a period of minimal parasite
322 exposure (**Figure 5F**). This was also true when expressing the abundance of subset 4 atypical
323 B cells as a percentage of total B cells (**Figure S14**). Simultaneously, we observed an increase
324 in the relative proportion of subset 5 among all atypical B cells between the two time points in
325 the adults, although the percentage of subset 5 atypical B cells among total B cells did not
326 change (**Figure 5F**, **Figure S14**). Collectively, these results demonstrate that we were able to
327 distinguish three subsets of atypical B cells by flow cytometry and that the composition of the
328 atypical B cell population changes in relation to parasite exposure.

329

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

332 To map additional differences in the phenotype between the three subsets of atypical B cells,
333 we plotted the mean fluorescence intensity of CD21, CD27, and the 12 markers that were
334 previously observed to be positively or negatively associated with activated or atypical B cells
335 for all three subsets as well as $IgG^+$ resting memory B cells ($CD21^-CD27^-$). We observed that

336 subset 4 had the most 'atypical' surface marker expression pattern, with the highest expression
337 of CD19, CD20, CD95, and T-bet (**Figure 6A**, **Figure S15**).

338 We were also interested in potential differences in phenotype between atypical B cells from
339 malaria-experienced children and adults. At the second time point, after months to years of no
340 chronic immune activation, the distribution of atypical B cell subsets was similar between
341 children and adults (**Figure 5F**). However, we noticed that their atypical B cells had a distinct
342 localization when projected onto the UMAP (**Figure 6B**). This observation suggests that atypical
343 B cells from children and adults differ in the expression of one or more other markers.
344 Overlaying the expression of other markers in our flow cytometry panel onto the UMAP revealed
345 that atypical B cells in adults extended into the UMAP area where cells expressed high levels of
346 CXCR3 and CD95 (**Figure 6C**). To quantify this, we analyzed the percentage of $CXCR3^+$ and
347 $CD95^+$ cells in each subset. Adults showed more $CXCR3^+$ cells in all atypical B cell subsets as
348 compared to children, and more $CD95^+$ cells in subsets 1 and 5 (**Figure 6D**). Collectively, this
349 difference in CXCR3 and CD95 expression between children and adults points towards
350 additional heterogeneity among atypical B cells that may be connected to differences in immune
351 status or parasite exposure levels.

352

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

354 To determine whether the different atypical B cell subsets develop in response to antigen during
355 infection or through bystander activation, we used antigen tetramers to detect B cells with
356 specificity to two *P. falciparum* proteins: merozoite surface protein 1 (MSP1) and apical
357 membrane antigen 1 (AMA1). Both antigens are known to elicit strong IgG responses in children
358 and adults (*38*). Since both MSP1 and AMA1 tetramers were labeled with the same
359 fluorochromes, all MSP1-specific and AMA1-specific B cells are referred to as MSP1/AMA1-

360 specific (**Figure S12**). We then determined the percentage of MSP1/AMA1-specific B cells
361 among the three atypical B cell subsets. First, we overlayed the MSP1/AMA1-specific B cells
362 onto the previously generated UMAPs of subset 1, 4, and 5 atypical B cells. Antigen-specific
363 cells were identified in all subsets (**Figure 7A**). We then quantified the number of antigen-
364 specific cells in each of the three atypical B cell subsets and observed that the distribution of
365 MSP1/AMA1 B cells over the three subsets mirrored their total size (**Figure 7B**). Finally, we
366 analyzed CXCR3 and CD95 expression among MSP1/AMA1-specific atypical B cells in adults.
367 Approximately 25 - 40% of all MSP1/AMA1-specific atypical B cells were CXCR3$^{+}$, with no clear
368 pattern related to subset identity (**Figure 7C**). In contrast, nearly all MPS1/AMA1-specific
369 atypical B cells in all three subsets expressed CD95, which was an increase as compared to all
370 atypical B cells. Increased CD95 expression was also observed in Spike-specific B cells in
371 recovered COVID-19 patients (*39*) and suggests that these cells recently underwent antigen-
372 driven activation. Collectively, these results suggest that all three subsets of atypical B cells
373 develop in response to antigen stimulation.

374

375 **DISCUSSION**

376 Atypical B cells accumulate in a range of conditions that share chronic immune activation as the
377 common denominator, including chronic or repetitive infection and autoimmunity. Attempts to
378 understand the function of these cells have resulted in conflicting observations. We
379 hypothesized that these different observations are caused by heterogeneity among atypical B
380 cells and that the underlying subsets are indistinguishable with the markers conventionally used
381 to define atypical B cells. Here, we used multimodal single-cell measurements to better define
382 the heterogeneity among atypical B cells using samples from individuals with chronic
383 *Plasmodium* exposure. We identified novel subsets of atypical B cells, each with a unique

384 transcriptional program and cell surface marker signature, that are associated with different
385 effector functions in the host immune response. Using spectral flow cytometry, we confirmed the
386 presence of these atypical B cell subsets in a cohort of malaria-experienced children and adults
387 and determined that they can develop in response to antigen stimulation.

388

389 Two other recent studies have attempted to assess potential heterogeneity among atypical B
390 cells using single-cell transcriptomics approaches. Holla *et al*. subclustered 199 B cells from
391 malaria-experienced individuals with an atypical transcriptomics profile and observed three
392 subsets that were mainly discriminated by isotype usage (i.e., $IgD^{+}IgM^{+}$, $IgD^{+}IgM^{lo}$, and $IgD^{-}IgG^{+}$
393 atypical B cells) (*11*). Too few cells were present in that data set to observe subclusters within
394 the $IgD^{-}IgG^{+}$ atypical B cell group, which was the focus of our study. Sutton *et al.* analyzed a
395 larger number of $IgD^{-}$ B cells (7,167 cells from malaria-experienced individuals and 5,813 cells
396 from individuals not exposed to *Plasmodium* infection) that were divided into three atypical B
397 cell clusters and several other B cell clusters (*3*). However, no information was provided on
398 differences in gene expression programs between these three clusters that may provide insight
399 into the different identities of these atypical B cell clusters. In our study, we generated data for
400 almost 21,500 B cells, of which we were able to assign slightly more than 14,000 cells to one of
401 10 phenotypically defined B cell populations. Among these were just over 4,000 cells that
402 belonged to three clusters containing phenotypically defined atypical B cells. Using this rich data
403 set, we confirmed similarities in gene expression between our three atypical B cell clusters and
404 bulk RNA sequencing transcriptional programs that have previously been reported for atypical B
405 cells (*10, 12*). However, we also discovered that some genes previously linked to atypical B
406 cells were expressed in only one of these three clusters and may be connected to the different
407 functional profiles of these cells. For example, CD72 negatively regulates B cell activation (*26*).
408 *CD72* was uniquely expressed in cluster 1 that has a profile of senescence or anergy. Cluster 4,

409 on the other hand, expressed higher levels of the gene encoding CD68, a transmembrane
410 glycoprotein that is mainly associated with the endosomal/lysosomal compartment in monocytes
411 and macrophages, but is also expressed by lymphocytes (*40, 41*). This observation may be
412 related to the pathways of antigen processing and exocytosis that were upregulated in this
413 cluster. Hopp *et al.* recently showed that atypical B cells upregulate proteins that mediate B – T
414 cell interactions and can be induced to differentiate into antibody-secreting cells in the presence
415 of follicular helper T cells, which matches our observations for cluster 4 (*16*). Sullivan *et al*.
416 reported that a subset of atypical B cells that expressed lower levels of FcRL5 showed more
417 antibody secretion upon stimulation than FcRL5$^{+}$ atypical B cells, in line with the profile we
418 observed for cluster 5 (*10*). The presence of secretory Ig transcripts in atypical B cells also
419 match the pre-antibody secreting signature of cluster 5 (*14*). Collectively, the functional profiles
420 of the three clusters of atypical B cells recapitulate many of the observations that have been
421 made about the atypical B cell population as a whole. Our results indicate that specific subsets
422 of atypical B cells are likely responsible for conflicting observations related to anergy/tolerance,
423 immune-modulation, and the ability to differentiate into antibody-secreting cells.

424

425 Using differentially expressed surface markers, we were able to identify three corresponding
426 subsets of atypical B cells by flow cytometry. All three subsets harbored *P. falciparum*-specific B
427 cells, suggesting that each of these subsets can develop as part of a specific response to
428 pathogen-derived antigens and not merely through bystander activation. Other stimuli, such as
429 cytokines and CD40 or TLR engagement may influence the development of atypical B cells into
430 one of these subsets. With its lower levels of somatic hypermutation and anergic transcriptional
431 profile, cluster 1 was distinct from clusters 4 and 5 that both seemed to play more functional
432 roles in the immune response. Indeed, B cell receptor repertoire analysis suggested that cluster
433 1 formed a compartment that was mostly separated from clusters 4 and 5, which between

434 themselves showed a higher percentage of clonal connectivity. The percentage of B cell clones
435 that were shared between clusters 4 and 5 increased over time as the B cell response
436 contracted to baseline levels. In addition, we observed that the decrease in subset 4 atypical B
437 cells over time was paralleled by a relative increase in the size of subset 5. Together, these
438 observations suggest that atypical B cells in subsets 4 and 5 are developmentally related.

439

440 Comparing atypical B cells between malaria-experienced children and adults revealed additional
441 heterogeneity. Adults showed a higher percentage of CXCR3$^{+}$ cells across all three subsets of
442 atypical B cells, which did not decrease after a two-year period of minimal *P. falciparum*
443 exposure. Hopp *et al.* reported that expression of *CXCR3* is upregulated in atypical B cells in
444 response to malaria in children (*16*). The stable expression of CXCR3 in adults suggests that
445 this marker is acquired as a result of chronic antigen exposure and may play an important role
446 in the protective immune response against malaria. Sutton *et al.* observed that *Pf*MSP1-specific
447 B cells in adults expressed higher levels of CXCR3 than the total memory B cell population,
448 although the majority of these cells did not have a CD21$^{-}$CD27$^{-}$ phenotype (*42*). In this study, we
449 detected CXCR3$^{+}$ parasite-specific cells but did not observe the same level of enrichment for
450 CXCR3 expression among atypical B cells. CXCR3 is an interferon-inducible chemokine
451 receptor specific for the chemokines CXCL9, CXCL10, and CXCL11 that are themselves also
452 induced by interferon (*43*). Its function has mainly been studied in T cells where it is required for
453 trafficking of T cells to sites of inflammation (*44*). The profile of trafficking receptor expression in
454 atypical B cells, including CXCR3, has been linked to the propensity of T-bet$^{hi}$ B cells to take up
455 residency in the spleen and bone marrow (*45*). While no data is available for the tissue
456 distribution of atypical B cells in malaria-experienced individuals, it would make sense for
457 parasite-specific B cells to home to the spleen, since this is the main anatomical location for the
458 removal of parasite-infected erythrocytes and cellular debris from the circulation. In this regard,

459 it is interesting to note that the elevated expression of T-bet and CD74 in subset 4 atypical B
460 cells fits the profile of memory B cells that reside in the spleen (*45, 46*), indicating that the loss
461 of these cells in the circulation could be the result of migration into the spleen. However, the
462 observation that the percentage of $CXCR3^+$ atypical B cells does not change over time with
463 minimal parasite exposure suggests that the expression of this marker by itself does not direct
464 homing of B cells to tissues. This notion is strengthened by the recent finding that CXCR3
465 ablation in murine B cells has no effect on the establishment and maintenance of lung-resident
466 memory B cells (*47*). Instead, CXCR3 should probably be considered a marker of activation or
467 prior activation.

468

469 One of the complications of studying atypical B cells is that different markers have been used to
470 define atypical B cells. In this study, our definition of atypical B cells is based on how these cells
471 are typically identified in the HIV and malaria fields ($CD21^-CD27^-$) and information from a recent
472 single-cell RNA sequencing study that reported CD11c as the most informative cell surface
473 marker for the identification of atypical B cell lineages (*3*). Our definition of atypical B cells as
474 $IgG^+CD21^-CD27^-CD11c^+$ also fits the definition of DN2 B cells, which is how these cells are
475 commonly referred to in the autoimmunity field ($IgD^-CD27^-CD21^-CD11c^+$) and follows recently
476 published recommendations for how these cells should be defined (*21*). Not all of the cells that
477 fall within our definition of atypical B cells are $T\text{-}bet^+$, which may complicate translating our
478 findings to studies that use T-bet (with or without CD11c) as a defining marker for atypical B
479 cells (*42, 48, 49*). Another caveat of this study is the inclusion of samples obtained ~2 months
480 post-malaria in children for flow cytometric analysis. Because of the timing of sample collection,
481 it is likely that we missed the peak of atypical B cell expansion (*3, 4*). Also, because of low *P.*
482 *falciparum* exposure levels around the time of sample collection, these children had relatively

483 low baseline levels of atypical B cells. As a result, few changes were observed between the two
484 time points and relatively few antigen-specific B cells were detected.

485

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

497

498 **MATERIALS AND METHODS**

499 **Study design and ethics approval**

500 The objective of this study was to assess transcriptional and functional heterogeneity among
501 atypical B cells to determine the function of these cells in the host immune response to infection.
502 Individuals included in this study (n = 21) were residents of the Nagongera sub-county in Tororo
503 District, Uganda. This region was historically characterized by extremely high malaria
504 transmission intensity, with an estimated annual entomological inoculation rate of 125 infectious
505 bites per person per year (*50*). Since 2015, multiple rounds of insecticide spraying (IRS) have
506 dramatically reduced malaria incidence compared with pre-IRS levels (*36*). Individuals were

507 selected for inclusion into this study based on age and *P. falciparum* exposure. In addition,
508 children were selected based on the availability of samples collected at set times relative to a
509 malaria episode (**Table S1**). All individuals included in this study were enrolled in The Program
510 for Resistance, Immunology, Surveillance, and Modeling of Malaria (PRISM) program (*51*) and
511 have provided written consent for the use of their samples for research. The PRISM cohort
512 study was approved by the Makerere University School of Medicine Research and Ethics
513 Committee (SOMREC), London School of Hygiene and Tropical Medicine IRB, and the
514 University of California, San Francisco Human Research Protection Program & IRB. The use of
515 cohort samples for this study was approved by the Institutional Review Board of the University
516 of Texas Health Science Center at San Antonio.

517

518 **B cell isolations**

519 Cryopreserved PBMCs were thawed in a 37°C water bath and immediately mixed with pre-
520 warmed thawing medium (IMDM/GlutaMAX supplemented with 10% heat-inactivated FBS (USA
521 origin) and 0.01% Universal Nuclease (Thermo, cat. no. 88700). After centrifugation (250 × g, 5
522 min), the cell pellet was resuspended in thawing medium. A sample of cells was mixed 1:1 (v/v)
523 with 0.2% trypan blue solution in PBS, and viable cells were counted using a Cellometer Mini
524 (Nexcelom) automated cell counter. Next, cells were centrifuged (250 × g, 5 min) and
525 resuspended in isolation buffer (PBS supplemented with 2% heat-inactivated FBS and 1 mM
526 EDTA) at 50 million live cells/mL and filtered through a 35 µm sterile filter cap (Corning, cat. no.
527 352235) to break apart aggregated cells. B cells were isolated using the EasySep Human B Cell
528 Isolation Kit (StemCell, cat. no. 17954) or the MojoSort Human Pan B Cell Isolation Kit
529 (BioLegend, cat. no. 480082) according to the manufacturer's instructions.

530

531 **Flow cytometry staining for single-cell analysis**

532 Isolated B cells were washed with PBS (250 × g, 5 min), resuspended in 1 ml of PBS containing
533 1 µl live/dead stain (LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Thermo Fisher Scientific, cat.
534 no. L34965), and incubated for 30 min on ice. Subsequently, cells were washed with cold PBS
535 containing 1% bovine serum albumin (BSA) (250 × g, 5 min) and incubated in labeling buffer
536 (total volume of 100 µl) containing PBS, 1% BSA, 5 µl Human TruStain FcX (Biolegend, cat. no.
537 422301) and 1% dextran sulfate (Sigma-Aldrich, cat. no. 42867-5G) for 10 min at 4°C before
538 incubating a further 30 min at 4°C with an antibody cocktail to label B cell surface markers
539 (**Table S7**). Before acquisition on a BD FACS AriaII cell sorter, cells were washed with cold PBS
540 containing 1% BSA, centrifuged (250 × g, 5 min, 4ºC), diluted to 20-30 million cells/mL in cold
541 PBS containing 1% BSA, and filtered into a FACS tube through a 35 µm sterile filter cap
542 (Corning, cat. no. 352235) to break apart aggregated cells. Lymphocytes were gated using
543 forward and sideward scatter, followed by doublet exclusion. Cells were then gated on live cells,
544 followed by $CD19^{+}CD10^{-}CD20^{+}$ B cells. Naive B cells ($CD21^{+}CD27^{-}$) and antigen-experienced B
545 cells (all other B cells) were separated based on CD21 and CD27 expression and sorted into
546 cold PBS containing 1% BSA. An average of 37,000 antigen-experienced B cells were mixed
547 with naïve B cells up to a total of 50,000 B cells (**Table S2**). This resulted in an enrichment of
548 antigen-experienced cells from an average of 49% in the original B cell pool to an average of
549 74% in the final cell suspension. Cells were immediately processed for single-cell sequencing.

550

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

552 B cells were counted with a hemocytometer and run on a Chromium 10x controller (10x
553 Genomics) in individual lanes with a targeted recovery rate of 10,000 cells per lane, according
554 to the manufacturer's instructions. Library preparation was completed following the

555 recommended protocols for Chromium Single Cell 5' Gene Expression kit as well as 5' Feature
556 Barcode and V(D)J Enrichment kits (**Table S8**). Libraries were sequenced using the Illumina
557 NovaSeq6000 (paired-end 150 bp) and HiSeq (paired-end 100 bp) platforms.

558

559 **Single-cell RNA-seq data analysis**

560 Sequencing reads were mapped to the human reference genome (hg38) and assigned to
561 individual cells of origin using cell-specific barcodes in Cell Ranger (10x Genomics) version
562 3.1.0. The sequencing data were then processed using Seurat version 3.2.3, following their
563 default pipeline as follows. Low-quality cells with high (>2500) or low (<200) UMI counts were
564 filtered from the samples. In addition, to prevent antibody isotype or V(D)J-gene segment usage
565 from influencing the clustering of cells, we removed all IGH, IGK and IGL transcripts from the
566 single cell data before clustering and UMAP visualization, as proposed by Steward *et al*. (*19*).
567 Gene expression levels for each cell were normalized by total expression, multiplied by a scale
568 factor (10,000), and log-transformed. Cells with > 5% mitochondrial reads were excluded
569 because these are likely damaged or dying cells that have lost much of their cytoplasmic
570 mRNA. Gene expression values for all genes in all clusters are provided in **Table S9**. We
571 determined the top 2000 most variable genes based on their average expression and
572 dispersion. Principal component analysis was performed on highly variable genes to reduce the
573 dimensionality of the data. We performed clustering using the Louvain algorithm implemented in
574 the *FindClusters* function of Seurat (algorithm = 1, resolution = 0.5, dims. use = 1:20) and
575 displayed the clusters using UMAP. One outlier cluster consisting of 13 cells that were most
576 likely T cells was removed from further analysis. Because cells from all sample were distributed
577 evenly across the UMAP (**Figure S3**), no batch corrections were performed. We identified the
578 cluster-specific marker genes conserved across different samples using the

579 *FindConservedMarkers* function (default parameters) from Seurat. To identify differentially
580 expressed genes between a pair of clusters, we used the *FindMarkers* function from Seurat
581 (with parameters *min.pct* = 0.20, and log2FC = 0.25), and additionally filtered genes using a
582 false discovery rate of 0.05. Violin plots, dot plots, and UMAP plots overlaid with gene
583 expression level were generated using standard Seurat functions. Reactome pathway analysis
584 was performed using differentially expressed genes that were upregulated in pairwise cluster
585 comparisons. The top 3 - 5 pathways with FDR-adjusted P value < 0.01 were selected for visual
586 representation in **Figure 3C**. The full set of Reactome analysis results and differentially
587 expressed gene lists are provided in **Tables S5** and **S6**.

588

589 **Single-cell RNA-sequencing gene signature score analysis**

590 We performed gene signature score analysis as described in (*11*). Briefly, gene expression data
591 from the Seurat object were scaled (z-score across each gene) to remove bias towards highly
592 expressed genes. A cell-specific signature score was computed by first sorting the normalized
593 scaled gene expression values for each cell followed by summing up the indices (ranks) of the
594 signature genes. These signature scores were then projected onto the transcriptomics-based
595 UMAP.

596

597 **Single-cell surface marker analysis**

598 From the single-cell RNA sequencing Seurat object, we extracted the antibody capture
599 component (assay = ADT), normalized the antibody matrix using CLR normalization (Seurat
600 method *NormalizeData*), and then scaled the antibody matrix using Seurat routine *ScaleData*.
601 This normalized and scaled antibody matrix was then used to define the following populations of

602 B cells by phenotype: 1) DN1 B cells: $CD27^-IgG^+CD21^+CD11c^-$ , 2) DN2 B cells: $CD27^-$
603 $IgG^+CD21^-CD11c^+$, 3) DN3 B cells: $CD27^-IgG^+CD21^-CD11c^-$, 4) DN4 B cells: $CD27^-$
604 $IgG^+CD21^+CD11c^+$, 5) Resting naïve B cells: $CD27^-IgM^+CD21^+$, 6) Activated naïve B cells:
605 $CD27^-IgM^+CD21^-$, 7) Resting switched memory B cells: $CD27^+IgG^+CD21^+$, 8) Activated switched
606 memory B cells: $CD27^+IgG^+CD21^-$, 9) Resting unswitched memory B cells: $CD27^+IgM^+CD21^+$,
607 and 10) Activated unswitched memory B cells: $CD27^+IgM^+CD21^-$. The UMAP plots highlighting
608 corresponding cell categories were generated using Seurat. Distributions of cell surface
609 expression of the markers CD21 and CD27 (**Figure S4**), as well as CD11c and CXCR3 (**Figure**
610 **S6**) for individual clusters were plotted using the above mentioned normalized and scaled
611 antibody matrix, and employing the 2D kernel density function *kde2d* from the R package
612 *MASS*.

613

614 **Single-cell B cell receptor analysis**

615 Single-cell B cell receptor sequencing data was first processed using the command *changeo-*
616 *10x*. Single-cell heavy chain variable region sequences were then analyzed using the
617 Immcantation pipeline. First, the Immcantation package Change-O (v0.4.4) was used to align
618 lymphocyte receptor sequences to germline sequences for downstream analyses (*52*). Isotype-
619 specific data files were converted into standardized tab-delimited database files required for
620 subsequent Change-O modules to operate using IgBLAST (v1.14.0), which is included in
621 Change-O. The frequency of somatic hypermutations was calculated as described previously
622 (*22*). Average frequencies for each donor and cluster are reported.

623 To calculate the percentage of clonal expansion within clusters or clonal connections between
624 clusters, genotyping was performed using the routine *tigger-genotype*, followed by analysis of

625 clonally related B cell receptor sequences using a custom Perl script available for download at
626 https://github.com/embunnik/clonal_scores, as described in detail elsewhere (*33*).

627

628 **Flow cytometry staining for spectral flow analysis**

629 C-terminally biotinylated full-length *P. falciparum* 3D7 MSP1 and AMA1 were produced as
630 described previously (*38*). Antigen tetramers were synthesized by incubating protein with
631 fluorophore-conjugated streptavidin overnight at 4˚C at a molar ratio of 6:1 with rotation (**Table**
632 **S10**). Isolated B cells were washed with PBS, centrifuged (250 × g, 5 min), resuspended in 1 ml
633 of PBS containing 1 µl live/dead stain (Zombie UV Fixable Viability kit (Biolegend, cat. no.
634 423107)) and incubated on ice for 30 min. Cells were subsequently washed with cold PBS
635 containing 1% BSA (250 × g, 5 min, 4ºC), resuspended with a cocktail of 25 µM of each antigen
636 tetramer diluted in PBS containing 1% BSA to a volume of 100 µl, and incubated at 4°C for 30
637 min. Next, the cells were washed twice with cold PBS containing 1% BSA (250 × g, 5 min, 4ºC)
638 and incubated at 4°C for 30 min with a B cell surface marker antibody cocktail (**Table S11**) with
639 10 µl Brilliant Stain Buffer Plus (BD, cat. No. 566385) diluted in PBS containing 1% BSA up to a
640 volume of 100 µl. The cells were then washed with cold PBS containing 1% BSA (250 × g, 5
641 min, 4ºC), resuspended in 1 ml of Transcription Factor Fix/Perm Concentrate (Tonbo, part of
642 cat. no. TNB-0607-KIT), diluted with 3 parts Transcription Factor Fix/Perm Diluent (Tonbo), and
643 incubated at 4°C for 1 hour. After the incubation, the cells were washed twice with 3 ml of 1×
644 Flow Cytometry Perm Buffer (Tonbo) (300 × g, 8 min, 4ºC) and resuspended in 1× Flow
645 Cytometry Perm Buffer with an intracellular marker antibody (**Table S11**). After an incubation at
646 4°C for 30 min, the cells were washed twice with 3 ml cold 1× Flow Cytometry Perm Buffer (300
647 × g, 8 min, 4ºC) and once with 3 ml cold PBS containing 1% BSA, resuspended in cold PBS
648 containing 1% BSA to 20 – 30 million cells/ml and filtered into a FACS tube through a 35 µm

649 sterile filter cap. Cells were analyzed by flow cytometry immediately following intracellular
650 staining.
651
652 **Spectral flow cytometry analysis**

653 B cells were analyzed on a Cytek Aurora spectral flow cytometer equipped with five lasers.
654 SpectroFlo QC Beads (Cytek, cat. no. SKU N7-97355) were run prior to each experiment for
655 performance tracking. Quality control and LJ tracking reports were used to ensure machine
656 performance and that settings between different runs were comparable. Paired samples were
657 processed in the same experiment and experiments performed on separate days contained
658 technical replicates. B cells isolated from pooled PBMCs from two malaria-naïve US donors
659 were used for the compensation of the live/dead stain, for the unstained control, and for
660 technical replicates. Between runs, the expression of cell surface and intracellular markers
661 showed a high correlation between these technical replicates (Spearman r = 0.98, **Figure S16**).
662 UltraComp eBeads Plus Compensation Beads (Thermo, cat. No. 01-3333-41) were used for
663 compensation of all other fluorophores. The cytometry analysis software OMIQ (Dotmatics) was
664 used for the integration and dimension reduction analysis. In short, atypical B cells were first
665 pre-gated on single/live/CD19$^{+}$/CD20$^{+}$/CD21$^{-}$/CD27$^{-}$/CD11c$^{+}$/IgG$^{+}$ cells. UMAPs were then
666 created using the expression of CD19, CD20, CD24, CD38, CD83, CD86, CD95, CXCR3,
667 CXCR5, CD11c, FcRL5, and T-bet as features with default parameters (neighbors = 15,
668 minimum distance = 0.4, metric = Euclidean, random seed = 9258) and included all 38 Ugandan
669 donor samples used in this study for initial projection. To define the three atypical B cell subsets,
670 FlowSOM (*53*) was used to identify three clusters based on CD11c, FcRL5, and CD86
671 expression using the default parameters (metric = Euclidean, random seed = 1226). For the
672 projection of antigen-specific B cells onto the UMAP, gates were manually set to identify
673 populations of interest using two-dimensional displays, which were then overlaid onto the UMAP

674 projection. Mean fluorescence intensities of cell surface and intracellular markers in select B cell
675 subsets were calculated for each sample individually in OMIQ, and then averaged across
676 samples. To visualize mean fluorescence intensities in a heatmap, the average value for each B
677 cell subset was expressed as a percentage of the highest mean fluorescence intensity among
678 all subsets.

679

680 **Supplementary materials**

681 Figure S1: Schematic overview of malaria history.
682 Figure S2: Contribution of individual samples to the total data set.
683 Figure S3: Comparisons between B cell clusters identified in this study and previously published
684 gene signatures.
685 Figure S4: Cell surface expression of CD21 and CD27 in each cluster as determined by feature
686 barcoding.
687 Figure S5: Distribution of cell surface marker expression and immunoglobulin heavy chain
688 transcripts per transcriptional cluster in each sample.
689 Figure S6: Cell surface expression of CD11c and CXCR3 in each cluster as determined by
690 feature barcoding.
691 Figure S7: Distribution of CD11c and CXCR3 among all cells and $CD21^{-}CD27^{-}$ cells in the three
692 clusters that contain atypical B cells.
693 Figure S8: Phenotypically defined B cell subsets projected onto the transcriptomics-based
694 UMAP.
695 Figure S9: Clusters obtained at a range of resolutions in Seurat.
696 Figure S10: Levels of somatic hypermutation in each cluster per sample.
697 Figure S11: Clonal expansion within and clonal connections between all transcriptomics-based

698 clusters

699 Figure S12: Gating strategy used to identify atypical B cells and antigen-specific B cells.

700 Figure S13: Projection of the 12 surface and intracellular markers used to generate the

701 composite UMAP of atypical B cells onto this UMAP.

702 Figure S14: Percentage of atypical B cell subsets among total and non-naïve B cells.

703 Figure S15: Mean fluorescence intensity (MFI) of cell surface and intracellular markers used for

704 the flow cytometric analysis of atypical B cell subsets.

705 Figure S16: Spectral flow cytometry technical replicates. The frequency of cells positive for the

706 markers CD11c, FcRL5, CD86 and all other surface and intracellular markers used in the study,

707 with the exception of CD19 and CD20, are shown.

708 Table S1: Information about all individuals included in the study.

709 Table S2: Number of cells sorted and pooled for quality control.

710 Table S3: Summary of sequence stats.

711 Table S4: Numbers of cells per sample in each cluster.

712 Table S5: Differentially expressed genes between clusters 1, 4, and 5.

713 Table S6: Reactome pathway analysis results.

714 Table S7: Antibodies used for flow sorting and single cell sequencing.

715 Table S8: Reagents used for single-cell immune profiling and Feature Barcoding.

716 Table S9: Average gene expression values for all genes in all clusters.

717 Table S10: Streptavidin conjugates used for spectral flow cytometry.

718 Table S11: Antibodies used in spectral flow cytometry analysis.

719 **References and notes**

720 1. D. Lau, L. Y. Lan, S. F. Andrews, C. Henry, K. T. Rojas, K. E. Neu, M. Huang, Y.
721 Huang, B. DeKosky, A. E. Palm, G. C. Ippolito, G. Georgiou, P. C. Wilson, Low CD21
722 expression defines a population of recent germinal center graduates primed for plasma
723 cell differentiation. *Sci Immunol* **2**, (2017).

2. S. F. Andrews, M. J. Chambers, C. A. Schramm, J. Plyler, J. E. Raab, M. Kanekiyo, R. A. Gillespie, A. Ransier, S. Darko, J. Hu, X. Chen, H. M. Yassine, J. C. Boyington, M. C. Crank, G. L. Chen, E. Coates, J. R. Mascola, D. C. Douek, B. S. Graham, J. E. Ledgerwood, A. B. McDermott, Activation Dynamics and Immunoglobulin Evolution of Pre-existing and Newly Generated Human Memory B cell Responses to Influenza Hemagglutinin. *Immunity* **51**, 398-410 e395 (2019).
3. H. J. Sutton, R. Aye, A. H. Idris, R. Vistein, E. Nduati, O. Kai, J. Mwacharo, X. Li, X. Gao, T. D. Andrews, M. Koutsakos, T. H. O. Nguyen, M. Nekrasov, P. Milburn, A. Eltahla, A. A. Berry, N. Kc, S. Chakravarty, B. K. L. Sim, A. K. Wheatley, S. J. Kent, S. L. Hoffman, K. E. Lyke, P. Bejon, F. Luciani, K. Kedzierska, R. A. Seder, F. M. Ndungu, I. A. Cockburn, 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 (2021).
4. C. Sundling, C. Ronnberg, V. Yman, M. Asghar, P. Jahnmatz, T. Lakshmikanth, Y. Chen, J. Mikes, M. N. Forsell, K. Sonden, A. Achour, P. Brodin, K. E. Persson, A. Farnert, B cell profiling in malaria reveals expansion and remodelling of CD11c+ B cell subsets. *JCI Insight* **5**, (2019).
5. S. Moir, J. Ho, A. Malaspina, W. Wang, A. C. DiPoto, M. A. O'Shea, G. Roby, S. Kottilil, J. Arthos, M. A. Proschan, T. W. Chun, A. S. Fauci, 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 (2008).
6. G. E. Weiss, P. D. Crompton, S. Li, L. A. Walsh, S. Moir, B. Traore, K. Kayentao, A. Ongoiba, O. K. Doumbo, S. K. Pierce, Atypical memory B cells are greatly expanded in individuals living in a malaria-endemic area. *J Immunol* **183**, 2176-2182 (2009).
7. C. Wei, J. Anolik, A. Cappione, B. Zheng, A. Pugh-Bernard, J. Brooks, E. H. Lee, E. C. Milner, I. Sanz, 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 (2007).
8. I. Isnardi, Y. S. Ng, L. Menard, G. Meyers, D. Saadoun, I. Srdanovic, J. Samuels, J. Berman, J. H. Buckner, C. Cunningham-Rundles, E. Meffre, Complement receptor 2/CD21- human naive B cells contain mostly autoreactive unresponsive clones. *Blood* **115**, 5026-5036 (2010).
9. S. A. Jenks, K. S. Cashman, E. Zumaquero, U. M. Marigorta, A. V. Patel, X. Wang, D. Tomar, M. C. Woodruff, Z. Simon, R. Bugrovsky, E. L. Blalock, C. D. Scharer, C. M. Tipton, C. Wei, S. S. Lim, M. Petri, T. B. Niewold, J. H. Anolik, G. Gibson, F. E. Lee, J. M. Boss, F. E. Lund, I. Sanz, Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus. *Immunity* **49**, 725-739 e726 (2018).
10. R. T. Sullivan, C. C. Kim, M. F. Fontana, M. E. Feeney, P. Jagannathan, M. J. Boyle, C. J. Drakeley, I. Ssewanyana, F. Nankya, H. Mayanja-Kizza, G. Dorsey, B. Greenhouse, FCRL5 Delineates Functionally Impaired Memory B Cells Associated with Plasmodium falciparum Exposure. *PLoS Pathog* **11**, e1004894 (2015).
11. P. Holla, B. Dizon, A. A. Ambegaonkar, N. Rogel, E. Goldschmidt, A. K. Boddapati, H. Sohn, D. Sturdevant, J. W. Austin, L. Kardava, L. Yuesheng, P. Liu, S. Moir, S. K. Pierce, A. Madi, Shared transcriptional profiles of atypical B cells suggest common drivers of expansion and function in malaria, HIV, and autoimmunity. *Sci Adv* **7**, (2021).

770 12. S. Portugal, C. M. Tipton, H. Sohn, Y. Kone, J. Wang, S. Li, J. Skinner, K. Virtaneva, D.
771 E. Sturdevant, S. F. Porcella, O. K. Doumbo, S. Doumbo, K. Kayentao, A. Ongoiba, B.
772 Traore, I. Sanz, S. K. Pierce, P. D. Crompton, Malaria-associated atypical memory B
773 cells exhibit markedly reduced B cell receptor signaling and effector function. *Elife* **4**,
774 (2015).
775 13. A. A. Ambegaonkar, K. Kwak, H. Sohn, J. Manzella-Lapeira, J. Brzostowski, S. K.
776 Pierce, Expression of inhibitory receptors by B cells in chronic human infectious diseases
777 restricts responses to membrane-associated antigens. *Sci Adv* **6**, eaba6493 (2020).
778 14. M. F. Muellenbeck, B. Ueberheide, B. Amulic, A. Epp, D. Fenyo, C. E. Busse, M. Esen,
779 M. Theisen, B. Mordmuller, H. Wardemann, Atypical and classical memory B cells
780 produce Plasmodium falciparum neutralizing antibodies. *J Exp Med* **210**, 389-399 (2013).
781 15. P. Holla, A. Ambegaonkar, H. Sohn, S. K. Pierce, Exhaustion may not be in the human B
782 cell vocabulary, at least not in malaria. *Immunol Rev* **292**, 139-148 (2019).
783 16. C. S. Hopp, J. Skinner, S. L. Anzick, C. M. Tipton, M. E. Peterson, S. Li, S. Doumbo, K.
784 Kayentao, A. Ongoiba, C. Martens, B. Traore, P. D. Crompton, Atypical B cells up-
785 regulate costimulatory molecules during malaria and secrete antibodies with T follicular
786 helper cell support. *Sci Immunol* **7**, eabn1250 (2022).
787 17. A. E. Braddom, G. Batugedara, S. Bol, E. M. Bunnik, Potential functions of atypical
788 memory B cells in Plasmodium-exposed individuals. *Int J Parasitol* **50**, 1033-1042
789 (2020).
790 18. C. M. Andrade, H. Fleckenstein, R. Thomson-Luque, S. Doumbo, N. F. Lima, C.
791 Anderson, J. Hibbert, C. S. Hopp, T. M. Tran, S. Li, M. Niangaly, H. Cisse, D.
792 Doumtabe, J. Skinner, D. Sturdevant, S. Ricklefs, K. Virtaneva, M. Asghar, M. V.
793 Homann, L. Turner, J. Martins, E. L. Allman, M. E. N'Dri, V. Winkler, M. Llinas, C.
794 Lavazec, C. Martens, A. Farnert, K. Kayentao, A. Ongoiba, T. Lavstsen, N. S. Osorio, T.
795 D. Otto, M. Recker, B. Traore, P. D. Crompton, S. Portugal, Increased circulation time of
796 Plasmodium falciparum underlies persistent asymptomatic infection in the dry season.
797 *Nat Med* **26**, 1929-1940 (2020).
798 19. A. Stewart, J. C. Ng, G. Wallis, V. Tsioligka, F. Fraternali, D. K. Dunn-Walters, Single-
799 Cell Transcriptomic Analyses Define Distinct Peripheral B Cell Subsets and Discrete
800 Development Pathways. *Front Immunol* **12**, 602539 (2021).
801 20. A. H. Ellebedy, K. J. Jackson, H. T. Kissick, H. I. Nakaya, C. W. Davis, K. M. Roskin,
802 A. K. McElroy, C. M. Oshansky, R. Elbein, S. Thomas, G. M. Lyon, C. F. Spiropoulou,
803 A. K. Mehta, P. G. Thomas, S. D. Boyd, R. Ahmed, Defining antigen-specific
804 plasmablast and memory B cell subsets in human blood after viral infection or
805 vaccination. *Nat Immunol* **17**, 1226-1234 (2016).
806 21. I. Sanz, C. Wei, S. A. Jenks, K. S. Cashman, C. Tipton, M. C. Woodruff, J. Hom, F. E.
807 Lee, Challenges and Opportunities for Consistent Classification of Human B Cell and
808 Plasma Cell Populations. *Front Immunol* **10**, 2458 (2019).
809 22. A. E. Braddom, S. Bol, S. J. Gonzales, R. A. Reyes, K. Musinguzi, F. Nankya, I.
810 Ssewanyana, B. Greenhouse, E. M. Bunnik, B Cell Receptor Repertoire Analysis in
811 Malaria-Naive and Malaria-Experienced Individuals Reveals Unique Characteristics of
812 Atypical Memory B Cells. *mSphere* **6**, e0072621 (2021).
813 23. D. R. Glass, A. G. Tsai, J. P. Oliveria, F. J. Hartmann, S. C. Kimmey, A. A. Calderon, L.
814 Borges, M. C. Glass, L. E. Wagar, M. M. Davis, S. C. Bendall, An Integrated Multi-omic
815 Single-Cell Atlas of Human B Cell Identity. *Immunity* **53**, 217-232 e215 (2020).

816 24. S. Gu, X. Han, C. Yao, H. Ding, J. Zhang, G. Hou, B. Qu, H. Zhou, Z. Ying, Z. Ye, J.
817 Qian, Q. Guo, S. Chen, C. G. Vinuesa, D. Dai, N. Shen, The transcription factor Zeb2
818 drives differentiation of age-associated B cells. *bioRxiv*, 2021.2007.2024.453633 (2021).
819 25. X. Gao, Q. Shen, J. A. Roco, K. Frith, C. M. L. Munier, M. Nekrasov, B. Dalton, J.-S.
820 He, R. Jaeger, M. C. Cook, J. J. Zaunders, I. A. Cockburn, ZEB2 regulates the
821 development of CD11c+ atypical B cells. *bioRxiv*, 2022.2009.2001.506173 (2022).
822 26. C. Akatsu, K. Shinagawa, N. Numoto, Z. Liu, A. K. Ucar, M. Aslam, S. Phoon, T.
823 Adachi, K. Furukawa, N. Ito, T. Tsubata, CD72 negatively regulates B lymphocyte
824 responses to the lupus-related endogenous toll-like receptor 7 ligand Sm/RNP. *J Exp Med*
825 **213**, 2691-2706 (2016).
826 27. C. Tan, R. Hiwa, J. L. Mueller, V. Vykunta, K. Hibiya, M. Noviski, J. Huizar, J. F.
827 Brooks, J. Garcia, C. Heyn, Z. Li, A. Marson, J. Zikherman, NR4A nuclear receptors
828 restrain B cell responses to antigen when second signals are absent or limiting. *Nat*
829 *Immunol* **21**, 1267-1279 (2020).
830 28. C. Tan, J. L. Mueller, M. Noviski, J. Huizar, D. Lau, A. Dubinin, A. Molofsky, P. C.
831 Wilson, J. Zikherman, Nur77 Links Chronic Antigen Stimulation to B Cell Tolerance by
832 Restricting the Survival of Self-Reactive B Cells in the Periphery. *J Immunol* **202**, 2907-
833 2923 (2019).
834 29. M. G. Giudizi, R. Biagiotti, F. Almerigogna, A. Alessi, A. Tiri, G. F. Del Prete, S.
835 Ferrone, S. Romagnani, Role of HLA class I and class II antigens in activation and
836 differentiation of B cells. *Cell Immunol* **108**, 97-108 (1987).
837 30. J. G. Cyster, C. D. C. Allen, B Cell Responses: Cell Interaction Dynamics and Decisions.
838 *Cell* **177**, 524-540 (2019).
839 31. P. Stumptner-Cuvelette, P. Benaroch, Multiple roles of the invariant chain in MHC class
840 II function. *Biochim Biophys Acta* **1542**, 1-13 (2002).
841 32. R. Di Niro, S. J. Lee, J. A. Vander Heiden, R. A. Elsner, N. Trivedi, J. M. Bannock, N. T.
842 Gupta, S. H. Kleinstein, F. Vigneault, T. J. Gilbert, E. Meffre, S. J. McSorley, M. J.
843 Shlomchik, Salmonella Infection Drives Promiscuous B Cell Activation Followed by
844 Extrafollicular Affinity Maturation. *Immunity* **43**, 120-131 (2015).
845 33. S. J. Gonzales, S. Bol, A. E. Braddom, R. Sullivan, R. A. Reyes, I. Ssewanyana, E.
846 Eggers, B. Greenhouse, E. M. Bunnik, Longitudinal analysis of FcRL5 expression and
847 clonal relationships among classical and atypical memory B cells following malaria.
848 *Malar J* **20**, 435 (2021).
849 34. U. Hershberg, E. T. Luning Prak, The analysis of clonal expansions in normal and
850 autoimmune B cell repertoires. *Philos Trans R Soc Lond B Biol Sci* **370**, (2015).
851 35. W. Meng, B. Zhang, G. W. Schwartz, A. M. Rosenfeld, D. Ren, J. J. C. Thome, D. J.
852 Carpenter, N. Matsuoka, H. Lerner, A. L. Friedman, T. Granot, D. L. Farber, M. J.
853 Shlomchik, U. Hershberg, E. T. Luning Prak, An atlas of B-cell clonal distribution in the
854 human body. *Nat Biotechnol* **35**, 879-884 (2017).
855 36. K. Zinszer, K. Charland, S. Vahey, D. Jahagirdar, J. C. Rek, E. Arinaitwe, J. Nankabirwa,
856 K. Morrison, M. L. Sadoine, M. A. Tutt-Guerette, S. G. Staedke, M. R. Kamya, B.
857 Greenhouse, I. Rodriguez-Barraquer, G. Dorsey, The Impact of Multiple Rounds of
858 Indoor Residual Spraying on Malaria Incidence and Hemoglobin Levels in a High-
859 Transmission Setting. *J Infect Dis* **221**, 304-312 (2020).
860 37. M. R. Kamya, A. Kakuru, M. Muhindo, E. Arinaitwe, J. I. Nankabirwa, J. Rek, V. Bigira,
861 J. Kapisi, H. Wanzira, J. Achan, P. Natureeba, A. Gasasira, D. Havlir, P. Jagannathan, P.

862 J. Rosenthal, I. Rodriguez-Barraquer, G. Dorsey, The Impact of Control Interventions on
863 Malaria Burden in Young Children in a Historically High-Transmission District of
864 Uganda: A Pooled Analysis of Cohort Studies from 2007 to 2018. *Am J Trop Med Hyg*
865 **103**, 785-792 (2020).
866 38. S. J. Gonzales, K. N. Clarke, G. Batugedara, R. Garza, A. E. Braddom, R. A. Reyes, I.
867 Ssewanyana, K. C. Garrison, G. C. Ippolito, B. Greenhouse, S. Bol, E. M. Bunnik, A
868 Molecular Analysis of Memory B Cell and Antibody Responses Against Plasmodium
869 falciparum Merozoite Surface Protein 1 in Children and Adults From Uganda. *Front*
870 *Immunol* **13**, 809264 (2022).
871 39. R. A. Reyes, K. Clarke, S. J. Gonzales, A. M. Cantwell, R. Garza, G. Catano, R. E.
872 Tragus, T. F. Patterson, S. Bol, E. M. Bunnik, SARS-CoV-2 spike-specific memory B
873 cells express higher levels of T-bet and FcRL5 after non-severe COVID-19 as compared
874 to severe disease. *PLoS One* **16**, e0261656 (2021).
875 40. C. L. Holness, D. L. Simmons, Molecular cloning of CD68, a human macrophage marker
876 related to lysosomal glycoproteins. *Blood* **81**, 1607-1613 (1993).
877 41. E. Gottfried, L. A. Kunz-Schughart, A. Weber, M. Rehli, A. Peuker, A. Muller, M.
878 Kastenberger, G. Brockhoff, R. Andreesen, M. Kreutz, Expression of CD68 in non-
879 myeloid cell types. *Scand J Immunol* **67**, 453-463 (2008).
880 42. J. L. Johnson, R. L. Rosenthal, J. J. Knox, A. Myles, M. S. Naradikian, J. Madej, M.
881 Kostiv, A. M. Rosenfeld, W. Meng, S. R. Christensen, S. E. Hensley, J. Yewdell, D. H.
882 Canaday, J. Zhu, A. B. McDermott, Y. Dori, M. Itkin, E. J. Wherry, N. Pardi, D.
883 Weissman, A. Naji, E. T. L. Prak, M. R. Betts, M. P. Cancro, The Transcription Factor T-
884 bet Resolves Memory B Cell Subsets with Distinct Tissue Distributions and Antibody
885 Specificities in Mice and Humans. *Immunity* **52**, 842-855 e846 (2020).
886 43. J. R. Groom, A. D. Luster, CXCR3 ligands: redundant, collaborative and antagonistic
887 functions. *Immunol Cell Biol* **89**, 207-215 (2011).
888 44. J. R. Groom, A. D. Luster, CXCR3 in T cell function. *Exp Cell Res* **317**, 620-631 (2011).
889 45. G. Fontanesi, P. Costa, R. Rotini, P. Pignedoli, Total P.C.A. knee arthroprosthesis. *Ital J*
890 *Orthop Traumatol* **14**, 41-48 (1988).
891 46. N. M. Weisel, S. M. Joachim, S. Smita, D. Callahan, R. A. Elsner, L. J. Conter, M.
892 Chikina, D. L. Farber, F. J. Weisel, M. J. Shlomchik, Surface phenotypes of naive and
893 memory B cells in mouse and human tissues. *Nat Immunol* **23**, 135-145 (2022).
894 47. C. Gregoire, L. Spinelli, S. Villazala-Merino, L. Gil, M. P. Holgado, M. Moussa, C.
895 Dong, A. Zarubica, M. Fallet, J. M. Navarro, B. Malissen, P. Milpied, M. Gaya, Viral
896 infection engenders bona fide and bystander subsets of lung-resident memory B cells
897 through a permissive mechanism. *Immunity* **55**, 1216-1233 e1219 (2022).
898 48. B. Keller, V. Strohmeier, I. Harder, S. Unger, K. J. Payne, G. Andrieux, M. Boerries, P.
899 T. Felixberger, J. J. M. Landry, A. Nieters, A. Rensing-Ehl, U. Salzer, N. Frede, S.
900 Usadel, R. Elling, C. Speckmann, I. Hainmann, E. Ralph, K. Gilmour, M. W. J. Wentink,
901 M. van der Burg, H. S. Kuehn, S. D. Rosenzweig, U. Kolsch, H. von Bernuth, P. Kaiser-
902 Labusch, F. Gothe, S. Hambleton, A. D. Vlagea, A. Garcia Garcia, L. Alsina, G. Markelj,
903 T. Avcin, J. Vasconcelos, M. Guedes, J. Y. Ding, C. L. Ku, B. Shadur, D. T. Avery, N.
904 Venhoff, J. Thiel, H. Becker, L. Erazo-Borras, C. M. Trujillo-Vargas, J. L. Franco, C.
905 Fieschi, S. Okada, P. E. Gray, G. Uzel, J. L. Casanova, M. Fliegauf, B. Grimbacher, H.
906 Eibel, S. Ehl, R. E. Voll, M. Rizzi, P. Stepensky, V. Benes, C. S. Ma, C. Bossen, S. G.

907 Tangye, K. Warnatz, The expansion of human T-bet(high)CD21(low) B cells is T cell
908 dependent. *Sci Immunol* **6**, eabh0891 (2021).
909 49. S. Wang, J. Wang, V. Kumar, J. L. Karnell, B. Naiman, P. S. Gross, S. Rahman, K.
910 Zerrouki, R. Hanna, C. Morehouse, N. Holoweckyj, H. Liu, T. Autoimmunity Molecular
911 Medicine, Z. Manna, R. Goldbach-Mansky, S. Hasni, R. Siegel, M. Sanjuan, K.
912 Streicher, M. P. Cancro, R. Kolbeck, R. Ettinger, IL-21 drives expansion and plasma cell
913 differentiation of autoreactive CD11c(hi)T-bet(+) B cells in SLE. *Nat Commun* **9**, 1758
914 (2018).
915 50. M. Kilama, D. L. Smith, R. Hutchinson, R. Kigozi, A. Yeka, G. Lavoy, M. R. Kamya, S.
916 G. Staedke, M. J. Donnelly, C. Drakeley, B. Greenhouse, G. Dorsey, S. W. Lindsay,
917 Estimating the annual entomological inoculation rate for Plasmodium falciparum
918 transmitted by Anopheles gambiae s.l. using three sampling methods in three sites in
919 Uganda. *Malar J* **13**, 111 (2014).
920 51. M. R. Kamya, E. Arinaitwe, H. Wanzira, A. Katureebe, C. Barusya, S. P. Kigozi, M.
921 Kilama, A. J. Tatem, P. J. Rosenthal, C. Drakeley, S. W. Lindsay, S. G. Staedke, D. L.
922 Smith, B. Greenhouse, G. Dorsey, Malaria transmission, infection, and disease at three
923 sites with varied transmission intensity in Uganda: implications for malaria control. *Am J*
924 *Trop Med Hyg* **92**, 903-912 (2015).
925 52. N. T. Gupta, J. A. Vander Heiden, M. Uduman, D. Gadala-Maria, G. Yaari, S. H.
926 Kleinstein, Change-O: a toolkit for analyzing large-scale B cell immunoglobulin
927 repertoire sequencing data. *Bioinformatics* **31**, 3356-3358 (2015).
928 53. S. Van Gassen, B. Callebaut, M. J. Van Helden, B. N. Lambrecht, P. Demeester, T.
929 Dhaene, Y. Saeys, FlowSOM: Using self-organizing maps for visualization and
930 interpretation of cytometry data. *Cytometry A* **87**, 636-645 (2015).

931

932 **Acknowledgements:** The authors acknowledge the Cell Analysis Core Facility (Dr. Sandra
933 Cardona) and the Genomics Research Core Facility (Sean Vargas) at the University of Texas at
934 San Antonio for support during this work. Data were generated in the Genome Sequencing
935 Facility, which is supported by UT Health San Antonio, NIH-NCI P30 CA054174 (Mays Cancer
936 Center at UT Health San Antonio), and NIH Shared Instrument grant 1S10OD030311-01 (S10
937 grant towards NovaSeq6000 sequencing system) and in the Flow Cytometry Shared Resource
938 Facility, which is supported by UT Health, NIH-NCI P30 CA054174 (Mays Cancer Center at UT
939 Health), CPRIT (RP210126) and NIH Shared Instrument grant 1S10OD030432-01 (S10 grant
940 towards the purchase of a DB FACSAria Fusion). Plasmids encoding 3D7 MSP1-bio, AMA1-bio,
941 and BirA, were a kind gift from Dr. Gavin Wright (Wellcome Sanger Institute; Addgene plasmids
942 #47709, 47741, and 32408).

943 **Funding:**

944 National Institutes of Health grant R01AI153425 (EMB)

945 National Institutes of Health grant R35GM128938 (FA)

946 National Institutes of Health Translational Science Training award TL1TR002647 (RAR)

947 National Institutes of Health predoctoral fellowship F31AI169993 (RAR)

948 National Institutes of Health Graduate Research in Immunology Program training grant
949 T32AI138944 (ABR)

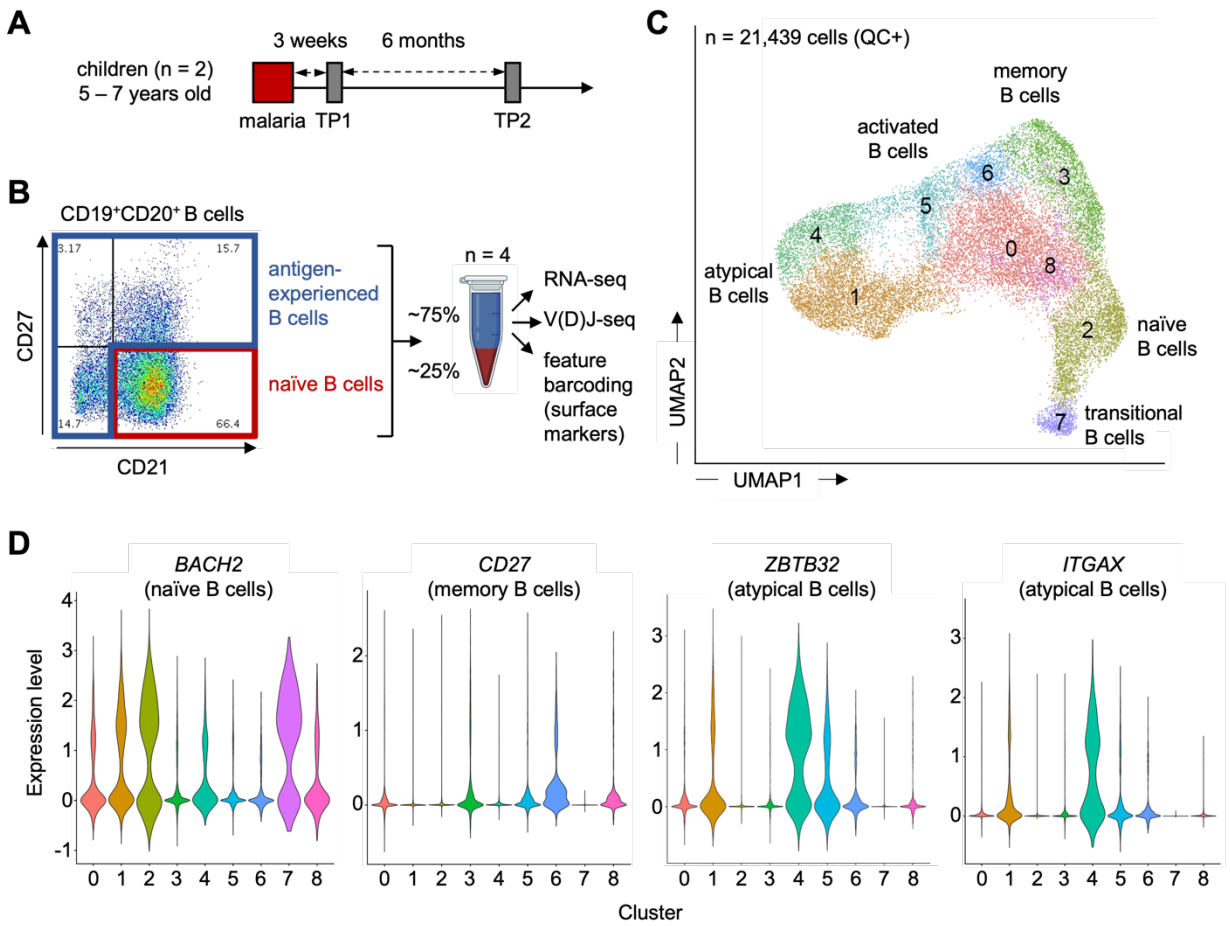
950 National Institutes of Health Research Supplement to Promote Diversity in Health-Related
951 Research R01 AI153425-03W1 (RG)

952 **Author contributions:** GB performed single-cell sequencing experiments and data analysis.
953 RAR performed spectral flow experiments and data analysis. PD and FA analyzed single-cell
954 sequencing data. ABR analyzed somatic hypermutation data. RG provided critical feedback on
955 the manuscript. IS, MEF, and BG provided clinical samples. RAR, GB, SB, and EMB wrote the
956 manuscript with input from all other co-authors. All authors contributed to the article and
957 approved the submitted version.

958 **Competing interests:** The authors declare that they have no competing interests.

959 **Data availability:** All sequencing files are available from the NCBI Sequencing Read Archive
960 under BioProject accession no. TBD.

961 **Figures:**




962

963 **Figure 1: Single-cell sequencing of B cells in malaria-experienced children. A)** Timeline of
964 sample collection at three weeks post-malaria (TP1) and six months later (TP2) in two children.
965 Children did not experience malaria between the first and second sample collection. **B)** Flow
966 cytometry strategy to enrich for antigen-experienced B cells. A majority of antigen-experienced
967 B cells were mixed with naïve B cells and used for library preparation using the 10x Genomics
968 platform. This yielded three different libraries for each sample to study gene expression, V(D)J
969 sequences, and cell surface marker expression in the same cells. **C)** UMAP of all B cells that
970 passed quality control (QC+), showing 9 transcriptomic B cell clusters. Major subsets identified
971 based on the expression of key genes and previously published gene signatures are indicated.
972 **D)** Expression levels of genes characteristic for naïve B cells (*BACH2*), memory B cells (*CD27*),
973 and atypical B cells (*ZBTB32* and *ITGAX*). TP, time point.

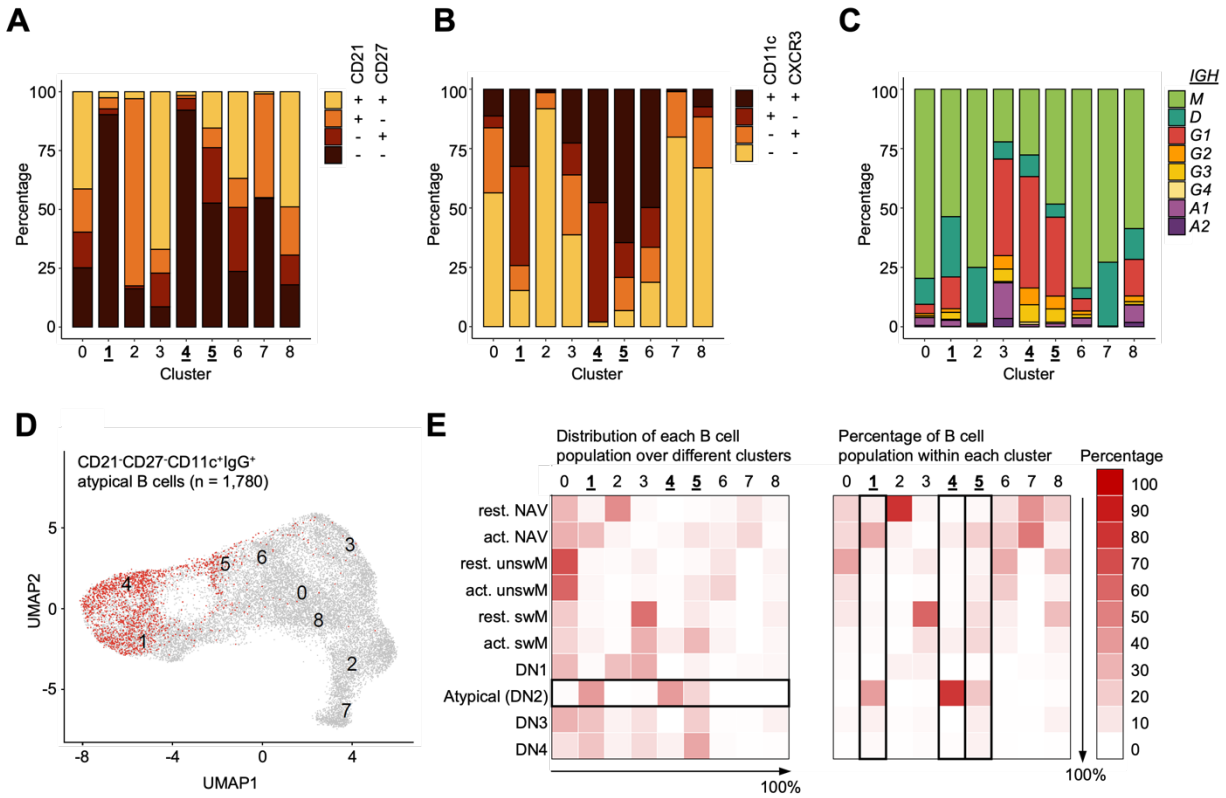



**Figure 2: Phenotypically defined atypical B cells are found in three transcriptionally distinct clusters. A)** Distribution of CD21$^{-/+}$ CD27$^{-/+}$ B cells among all transcriptional clusters. **B)** Distribution of CD11c$^{-/+}$ CXCR3$^{-/+}$ B cells among all transcriptional clusters. **C)** Distribution of immunoglobulin heavy chain 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)** The distribution of each B cell population over the different clusters (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. The atypical B cell population (left heatmap) and clusters 1, 4, and 5 (right heatmap) mentioned in the text are highlighted with borders. Rest., resting; act.,

activated; NAV, naïve, unswM, unswitched memory; swM, switched memory; DN, double negative.

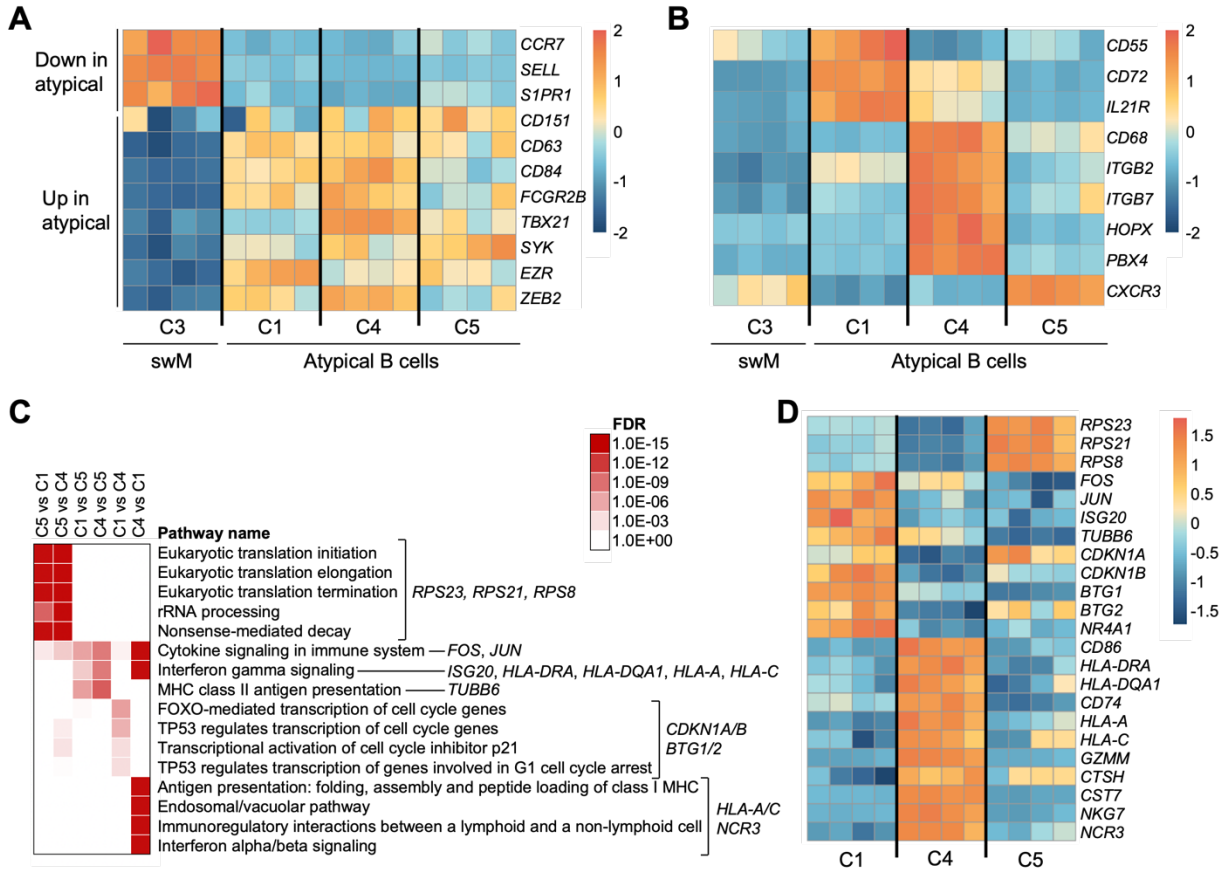



**Figure 3: Phenotypic and functional differences between clusters of atypical B cells. A)** Expression of genes previously identified as part of the unique transcriptional signature of atypical B cells that are consistently expressed higher or lower in all three clusters of atypical B cells as compared to class-switched resting memory B cells (swM; cluster 3). **B)** Differential expression of atypical B cell-associated genes between the three clusters of atypical B cells. **C)** Selected reactome pathways that were significantly enriched (FDR-corrected P value < $10^{-4}$) in pair-wise comparisons between the three clusters that contain atypical B cells. Key differentially

997 expressed genes in these pathways that are shown in panel D are indicated. **D)** Select
998 differentially expressed genes in atypical B cell clusters that were present in enriched Reactome
999 pathways and are discussed in the text.

1000 For heatmaps in panels A, B, and D, gene expression values for the four samples were first
1001 divided by the average value for all clusters for that donor to account for donor-to-donor
1002 variation. The data points were then z-scored across all clusters and samples prior to plotting
1003 heatmaps using the color palette 'Hiroshige' from the R package MetBrewer. The four columns
1004 represent the four samples subjected to single-cell RNA-sequencing analysis. C1, cluster 1; C3,
1005 cluster 3; C4, cluster 4; C5, cluster 5

1006

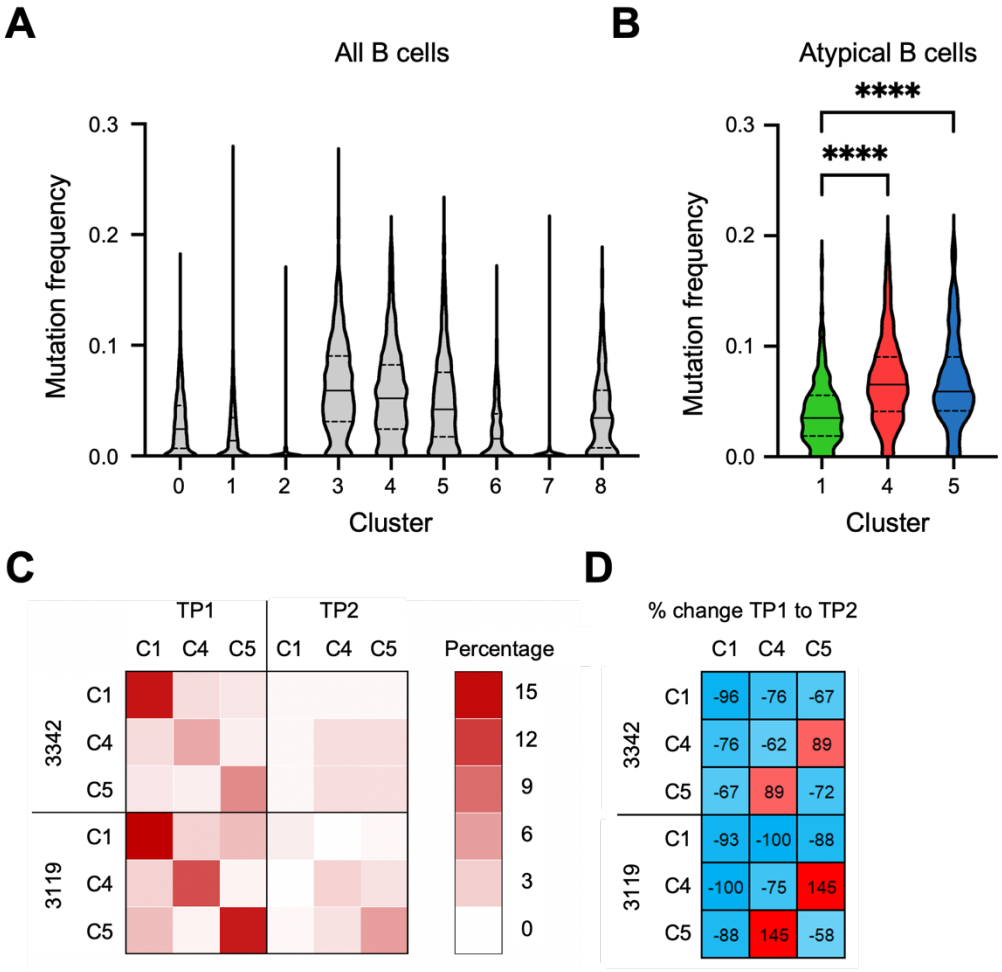
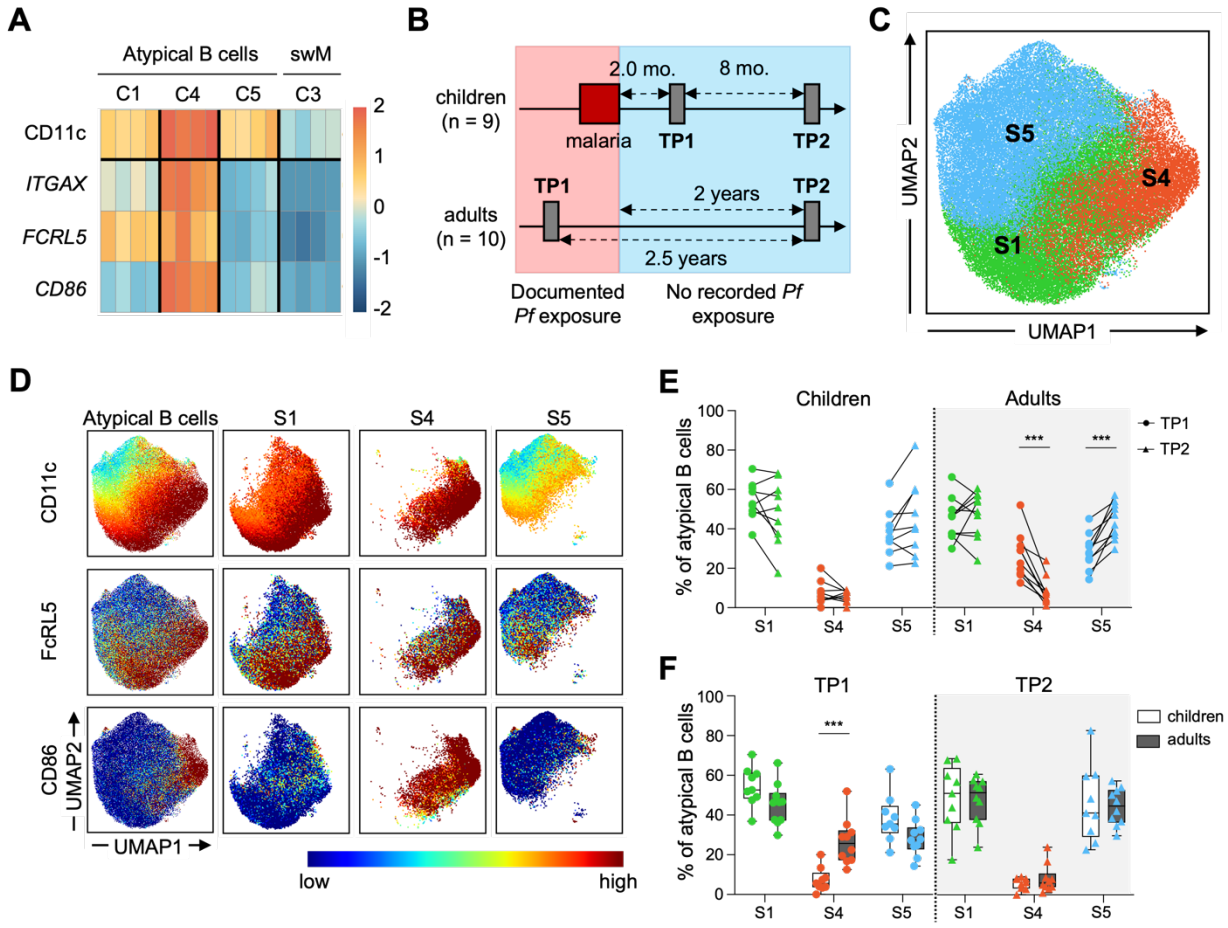



**Figure 4: Differences in somatic hypermutation frequency and clonal connectivity of atypical B cells. A)** Levels of somatic hypermutation among the transcriptomics-based B cell clusters. **B)** Levels of somatic hypermutation among phenotypically defined atypical B cells in clusters 1, 4, and 5. In panels A and B, the median is shown as a solid horizontal lines, while the quartiles are represented by dashed horizontal lines. A one-way ANOVA was used to test for statistically significant differences between the groups. P values are shown from Tukey's post hoc test. **** $P < 0.0001$. **C)** 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

1017 (TP2). **D)** The percentage change in clonal expansion/connections between the first and the

1018 second time point for the three clusters that contain atypical B cells. TP, time point

1019




1020

1021 **Figure 5: Identification of atypical B cell subsets in malaria-experienced children and**

1022 **adults. A)** Gene expression profiles of differentially expressed genes encoding cell surface

1023 proteins CD11c, FcRL5, and CD86 in atypical B cell clusters 1, 4, and 5, as well as resting

1024 switched memory B cells (swM, cluster 3), used to distinguish atypical B cell subsets by flow

1025 cytometry. The top row shows cell surface expression of CD11c to highlight the correlation

1026 between gene expression and protein expression for this marker. **B)** Schematic overview of

1027 sample collection for children and adults during known *P. falciparum* (*Pf*) exposure and in the

absence of exposure to *P. falciparum*. **C)** UMAP projection of atypical B cells from all 19 samples. Subsets were defined based on expression of CD11c, FcRL5, and CD86 and were named according to the corresponding clusters defined by single-cell RNA-sequencing analysis. **D)** CD11c, FcRL5, and CD86 expression projected onto the UMAP of all atypical B cells as well as each of the three subsets individually. **E)** Relative size of each atypical B cell subset in children (n = 9) and adults (n = 10) at the first time point (TP1; circles) and second time point (TP2; triangles). **F)** Comparison of relative subset sizes between children and adults at each time point. Statistical analysis was performed using a Wilcoxon matched-pairs signed rank test for paired data (E) or a Mann-Whitney test for unpaired data (F) 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

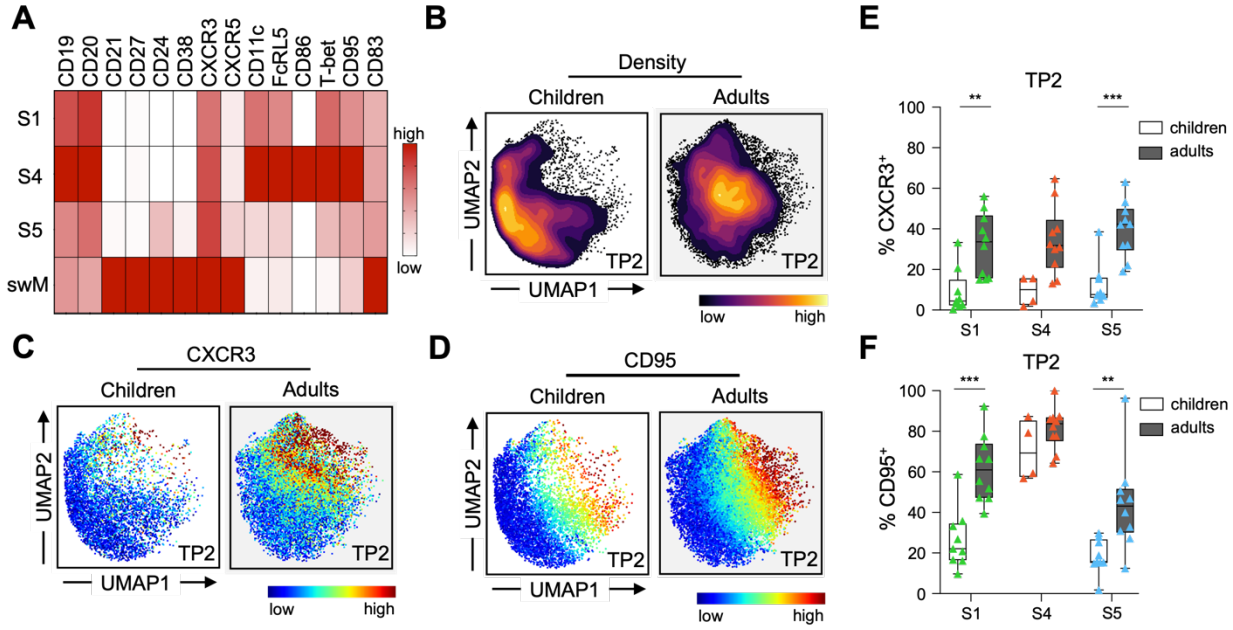
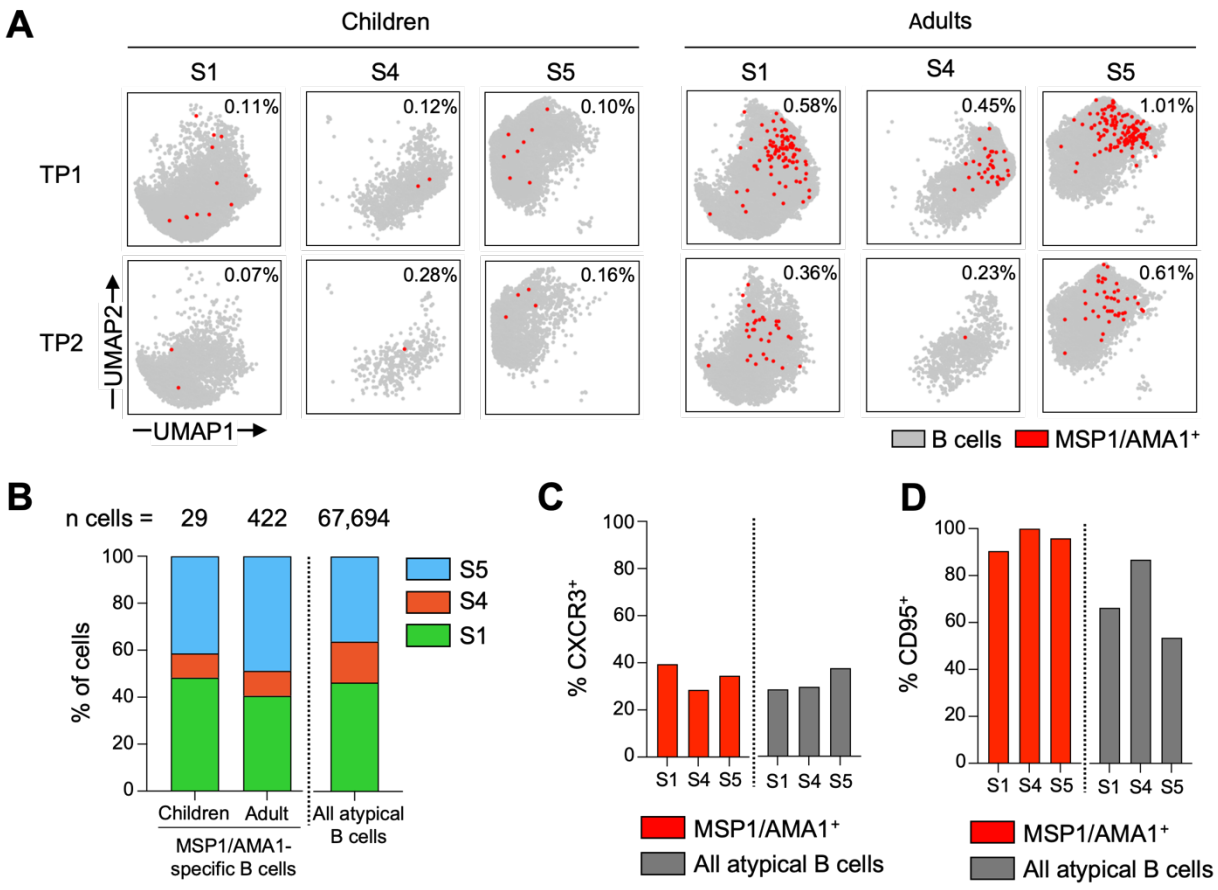



**Figure 6: Detailed phenotyping of atypical B cell subsets. A)** Relative MFI of 14 surface and intracellular markers in subsets (S) 1, 4, and 5 and in resting switched memory B cells (swM).

1043 **B)** Density UMAP plots of atypical B cells detected at the second time point in children and
1044 adults. **C-D)** Projection of CXCR3 (C) and CD95 (D) expression in atypical B cells at the second
1045 time point projected onto the UMAPs for children and adults. **E-F)** Percentage of CXCR3$^+$ (E)
1046 and CD95$^+$ (F) cells within atypical B cell subsets. For each subset, only individuals with >50
1047 cells at both time points were included in the analysis. Statistical analysis was performed using
1048 a Mann-Whitney test with a 10% false discovery rate using the two-stage step-up method of
1049 Benjamini, Krieger, and Yekutieli. * P < 0.05; ** P < 0.01; *** P < 0.001

1050




1051

1052 **Figure 7: Detection of MSP1/AMA1-specific cells among the three atypical B cell subsets.**
1053 **A)** Projection of MSP1/AMA1-specific B cells (red) onto the UMAPs of atypical B cell subsets,

1054 shown separately for time point 1 (TP1) and time point 2 (TP2) in children (n = 9) and adults (n
1055 = 10). **B)** Distribution of MSP1/AMA1-specific B cells over the three atypical B cell subsets in
1056 children and adults. Data from both time points were combined. **C)** Percentage of $CXCR3^+$
1057 antigen-specific cells in each atypical B cell subset for adult samples. **D)** Percentage of $CD95^+$
1058 antigen-specific cells in each atypical B cell subset for adult samples.