# Title: Single-cell immune repertoire tracing identifies rituximab refractory B cells that emerge during relapse

**Authors:** Ruoyi Jiang [1], Miriam L. Fichtner [1,2], Kenneth B. Hoehn[1,4], Panos Stathopoulos [1,2], Richard J. Nowak [2], Steven H. Kleinstein [1,3,4,†,*], Kevin C. O'Connor [1,2,†,*]

**Affiliations:**

[1]Department of Immunobiology, Yale School of Medicine, New Haven, CT 06520, USA.

[2]Department of Neurology, Yale School of Medicine, New Haven, CT 06520, USA.

[3]Interdepartmental Program in Computational Biology & Bioinformatics, Yale University. New Haven, CT 06520, USA.

[4]Department of Pathology, Yale School of Medicine, New Haven, CT 06520, USA.

*Corresponding authors: Corresponding author. Email: steven.kleinstein@yale.edu (S.H.K); kevin.oconnor@yale.edu (K.C.O)

†Co-senior authors

**One Sentence Summary:** Disease relevant B cells during post-rituximab (RTX) relapse emerge from the unsuccessful depletion of pre-existing B cell clones; single-cell transcriptomic phenotyping combined with lineage tracing using paired B cell receptor repertoires identified both persistent antibody-secreting cell as well as memory B cell subsets.

**Abstract**: Rituximab, an anti-CD20 B cell-depleting therapy, is indicated to treat a growing number of autoimmune disorders. However, disease relapses can occur when the B cell compartment is re-established after treatment. To explore the mechanism of relapse, we studied patients with muscle-specific kinase (MuSK) myasthenia gravis (MG), a paradigm for B cell-mediated autoimmune diseases in which pathology is directly associated with autoantibody production. Whether the failed depletion of specific clones leads to the persistence of autoantibody producing B cell subsets has yet to be elucidated. Here we show that rituximab (RTX) therapy is not fully effective at eliminating disease-associated B cells and provide a mechanistic understanding of relapse. We carried out single-cell transcriptional and B cell receptor (BCR) profiling on longitudinal B cell samples from the peripheral blood of MuSK MG patients who relapsed after rituximab therapy. Computational analysis of these data identified individual B cells that were refractory to rituximab depletion by linking them to clones that were present prior to therapy and continued to persist after therapy. These persistent B cell clones were present at high frequency, and included both antibody-secreting cells and memory B cells. They were disproportionately isotype switched with elevated somatic hypermutation frequencies and a gene expression signature associated with antigen-experience and B cell clonal expansion. We further identified both antibody-secreting cells and memory B cells with specificity for MuSK autoantigen that were clonally expanded during relapse and persistent. Our results provide evidence that post-rituximab relapse in MuSK MG is associated with the failed depletion of autoantigen-specific B cell subsets. This work provides insight into the durability of human B cells during treatment with therapeutic CD20-specific monoclonal antibodies. The transcriptional signatures associated with B cell resistance may represent new therapeutic avenues for treatment and biomarkers of depletion efficacy.

**[Main Text: ]**

## Introduction

Anti-CD20 mediated B cell depletion therapy (BCDT) is indicated for the treatment of multiple autoimmune disorders including rheumatoid arthritis (RA), autoimmune-associated vasculitides (AAV) and multiple sclerosis (MS) *(1–3)*. Acquired Myasthenia gravis (MG) is a chronic autoimmune disorder caused by pathogenic autoantibodies that bind to membrane proteins at the neuromuscular junction leading to muscle weakness. While the majority of MG patients have serum autoantibodies against the acetylcholine receptor (AChR), a subset of patients have autoantibodies against muscle-specific tyrosine kinase (MuSK). MuSK is essential for AChR clustering and synaptic differentiation. Through binding MuSK, serum-derived autoantibodies impair clustering of AChRs, and block neuromuscular transmission causing disease *(4)*. The conspicuous clinical benefit of BCDT in patients with MuSK MG was recently demonstrated *(5–7)*. However, a fraction of patients treated with the anti-CD20 BCDT agent rituximab (RTX) experience post-rituximab (post-RTX) relapses despite symptomatic responses and a fall in anti-MuSK antibody titer *(5, 8–10)*.

Post-RTX relapse has been reported in the context of RA, AAV and systemic lupus erythematosus (SLE), and is typically managed with additional BCDT *(11–13)*. An elevated fraction of antigen-experienced B cell subsets (including memory B cells and plasmablasts) after depletion has been correlated with the occurrence of post-RTX relapse *(14–19)*. Post-RTX relapse in MuSK MG was associated with the presence of circulating plasmablast expansions, which include those secreting MuSK autoantibodies *(20, 21)*.

Previous studies provide some insight into the possible origin of circulating antigen-experienced B cells during post-RTX relapse. Antigen-experienced plasma cells or plasmablasts are known to be resistant to BCDT owing to their low CD20 expression *(22–25)*. However, memory B cells, which express CD20, are abundant immediately after BCDT, and their frequency, as a fraction of total B cells, is elevated in patients with RA who experienced relapses after RTX. *(15, 16)*. Whether or not plasmablasts, plasma cells or memory B cells emerge from inadequately depleted circulating B cells (versus emerging entirely *de novo*) is not known.

The characteristics of antigen-experienced B cells that re-emerge during post-RTX relapse also remain unclear. BAFF (also known as B cell activating factor, B lymphocyte stimulator or BlyS) is a TNF superfamily member produced almost exclusively by myeloid cells *(26, 27)*. BAFF (and another TNF superfamily ligand, APRIL) can bind a family of B cell specific receptors to promote B cell survival. For example, levels of receptors for BAFF/APRIL on B cells such as TACI, BAFF-R and BCMA correlate with poor clinical responses to BCDT but whether these receptors are expressed by persistent B cells or B cells emerging *de novo* has not been investigated *(28–30)*. Clonal expansions have also been correlated with poor responses to RTX, but whether this is because some B cell subsets incompletely depleted by RTX happen to also be clonally expanded remains an open question *(31, 32)*.

Understanding the factors contributing to post-RTX relapse is important for the development of new modalities of monitoring depletion and more effective therapy combinations for treating MG and other autoimmune disorders. Thus, we sought to test the hypothesis that B cell clones resistant to RTX are disease relevant and preferentially clonally expanded post-RTX. We also

sought to test the hypothesis that these B cells are associated with plasmablasts and memory B cells that differentially express CD20 and APRIL/BAFF receptor genes. To this end, we used adaptive immune receptor repertoire sequencing (AIRR-Seq) using next-generation sequencing (NGS) and single-cell gene expression analysis to trace B cell populations temporally and to characterize B cells that emerge during post-RTX relapse. In comparison to the examination of overlap between bulk BCR repertoires alone, this approach allows for the characterization of the expression profile of persistent, disease-relevant B cells (including those specific for antigen) at single-cell resolution. Moreover, this approach allows for the identification of genes representing novel therapeutic targets, is unbiased, and not limited to the set of known B cell surface markers used for characterizing B cell subsets. Finally, it can also be used to remove possible confounders related to B cell subset distribution that could affect the identification of repertoire features associated with persistence. For instance, factors such as clonality, isotype usage and SHM frequency associated with persistence may actually be features associated with persistent subsets such as plasmablasts.

NGS approaches for generating immune repertoires commonly begin with genomic DNA or mRNA and produce high-depth bulk repertoires of unpaired immunoglobulin or T cell receptor V(D)J sequences *(33–35)*. Recent approaches use single-cell isolation techniques with high-throughput sequencing to construct repertoires, sometimes with paired gene expression *(36–39)*. These include studies on tumor-infiltrating lymphocytes from cancer biopsies, infiltrates from the brain of mice and B cells from the spleen after vaccination in mice; no such studies have investigated the circulating B cell repertoire in patients with autoimmune disease thusfar *(37–39)*. Previous immune repertoire studies of BCDT have identified persistent expanded clones and associations between clonality and responder status, and single-cell transcriptome studies have recapitulated flow cytometry results showing increased frequencies of antigen-experienced B cells at the peak of depletion *(32, 40–42)*. In this study, we combine both approaches to leverage depth and single-cell resolution to investigate the basis for differential B cell depletion by RTX in an unbiased, high-throughput manner. This approach, which we refer to as Single-cell Tracing of Adaptive Immune Receptor repertoires (STAIR), allows for unbiased transcriptional characterization of B cells resistant to BCDT.

We demonstrated that circulating memory B cells and post-RTX plasmablasts which produce MuSK-specific antibodies emerge from the incomplete depletion of clonally expanded pre-RTX clones. We also showed that plasmablasts were more persistent than memory B cells. Memory B cells associated with persistence were characterized by low CD20 expression but also low BAFF-R expression and high TACI expression. Clonal expansion was also found to be associated with persistence among memory B cells. These results provide new insights into the mechanism of post-RTX relapse in patients with MuSK MG and potential avenues for improving the treatment of MG and other B cell mediated autoimmune diseases.

## Results

### *A subset of B cell clones is refractory to depletion*

We sought to identify clones that escaped RTX depletion and thus may contribute to relapse that occurs when the B cell compartment repopulates after treatment in patients with MG (fig. 1). To that end, we sequenced the B cell repertoire from three MG patients before and after treatment

(table S1, S2). The number of V(D)J sequences isolated for patient 1 was 639,715 V(D)J; 82,720 from patient 2; and 107,504 from patient 3. Grouping the sequences into clones—those with a common ancestor based on their V, J gene usage and junction sequences *(43)*—generated 187,195; 34,719; and 14,579 clones for patient 1, 2 and 3 respectively. A total of 2,462 clones were found with members that were present in both the pre-rituximab (pre-RTX) and post-rituximab relapse (post-RTX) repertoire (termed persistent clones) in patient 1 (or 1.6% of pre-RTX clones); 902 (4.9%) and 345 (7.6%) were found in patients 2 and 3 respectively (fig. 2A). Given that clonal relationships are inferred computationally, we next sought to substantiate that these persistent clones were not identified by sequence similarity that happened by chance. Because clones cannot be shared across multiple individuals, the average number of clones shared between patients was calculated e.g. the average number of shared clones between patient 1 and 2, and patient 1 and 3 was calculated. Patient 1 shared an average of 0.4% clones in comparisons with other patients, patient 2 shared an average of 1.4%, and patient 3 shared an average of 0.9%. Quantification using a clonal overlap metric demonstrated that the observed overlaps between pre-RTX and post-RTX repertoires were significantly higher relative to this background (t-test P=0.037) (fig. 2B). These collective data demonstrate that circulating B cell clones persist despite RTX-mediated depletion of the repertoire.

*Persistent clones are antigen experienced*

Next we investigated if persistent clones that emerge post-RTX have distinct characteristics compared to non-persistent clones. Given the possible role of persistent B cells in recognizing antigen, we characterized features of the repertoire associated with antigen experience, namely isotype switching or elevated somatic hypermutation (SHM) frequency. We quantified the difference in isotype frequency among V(D)J sequences from persistent and non-persistent post-RTX clones. V(D)J sequences from persistent clones were observed to be mostly isotype switched. While non-persistent sequences were 43.2% IgM, 6.4% IgD, 24.7% IgG and 25.5% IgA of the repertoire on average (fig. 2C), persistent ones were disproportionately switched (15.0% IgM, 2.2% IgD, 27.4% IgG and 55.0% IgA) (ANOVA p=0.027). In addition, the overall SHM frequency of persistent clones was significantly elevated (fig. 2D) (ANOVA P=0.014) independent of isotype (ANOVA p = 0.287). Therefore, persistent clones are preferentially isotype switched with elevated SHM, reflecting a higher frequency of antigen-experienced clones.

*Persistent clones do not consistently expand or gain SHM*

We next sought to determine whether antigen experience was gained during reconstitution of the B cell compartment after RTX treatment. We first tested if persistent clones expand between pre-RTX and post-RTX time points. Clonal expansion was assessed by computing the change in size of persistent clones between pre-RTX and post-RTX time points. Average decreases of -11% (95% CI -93% to +1313%), -85% (95% CI -98% to +439%) and -7% (95% CI -99% to +40.5%) in clonal size respectively were observed across patients 1, 2 and 3 (one-sample t-test P=0.37) (fig. S3A-C). Thus, decreases in clonal size were observed but these decreases were not significant. Given that persistent clones have elevated frequencies of SHM, we tested whether individual clones acquire additional SHM by quantifying the average change in SHM frequency of persistent clones between pre-RTX and post-RTX time points. The median change in SHM frequency was -0.2% (95% CI -89% to +1132%), -2.4% (95% CI -74% to 187%) and +4.5% (95% CI -41% to +67%) for patients 1, 2 and 3 respectively (fig. S3D-F). Thus, we did not detect a consistent change in the SHM frequency of persistent clones between pre-RTX and post-RTX

samples for these patients (one-sample t-test P=0.807). These collective findings suggest that clonal expansion and the acquisition of antigen experience between pre-RTX and post-RTX time points is not a general feature of persistent clones.

*Persistent clones are disease relevant and bind MuSK*

We next sought to investigate if persistent B cell clones were disease-relevant. In two previous studies, using the same three patients described here, we identified B cell clones during post-RTX relapse that had disease relevance (table S3). In the first, 43 B cell clones were used to generate recombinant monoclonal antibodies, all of which were derived from circulating plasmablasts as this cell type was demonstrated to produce MuSK-specific antibodies when cultured *in vitro (20)*. In the first, 43 B cell clones were used to generate recombinant monoclonal antibodies, all of which were derived from circulating plasmablasts as this cell type was demonstrated to produce MuSK-specific antibodies when cultured in vitro *(21)*. In the second, a fluorescent MuSK antigen bait was used to both identify and isolate MuSK-specific B cell clones from which an additional 38 recombinant monoclonal antibodies were produced. Members of these collective 81 clones bound surface expressed MuSK using a cell-based assay across a range of different binding strengths with many of these observed to be weak binders (but some were strong binders) *(20, 21)*. Of these 81 clones, 27 could be found in our post-RTX repertoires. This demonstrates that they were clonally expanded, as they were found in independent samples. When we tested if members of these 81 clones were present in pre-RTX repertoires we observed that seven clones could be found in a pre-RTX repertoire; pre-RTX relatives for the other 74 clones were not observed. Specifically, we observed that one clone from patient 1 could be found in a pre-RTX repertoire, while six clones from patient 3 were also found in a pre-RTX repertoire (fig. 3, fig. S4A-F). One of the six clones from patient 3, which were present in both the pre- and post-RTX repertoire, demonstrated modest MuSK binding activity (fig. 3). Thus, disease-relevant B cells present during relapse can derive from the failed depletion of B cell clones that existed prior to therapy.

*Antibody-secreting cell-like B cells are preferentially persistent but memory B cells also persist*

Given the disease relevance of persistent antigen-experienced B cells during relapse, we next sought to characterize the transcriptional state of persistent B cells in an unbiased manner to investigate how different B cell subsets are represented among persistent populations. To accomplish this, we carried out single-cell transcriptome and paired BCR repertoire profiling of enriched circulating post-RTX B cells. Owing to the high frequency of transitional B cells during post-RTX reconstitution, $IgD^{low}$ B cells were sorted for prior to single-cell analysis (fig. S5). The analysis was limited to the switched IgG and IgA repertoire by computationally filtering out cells with V(D)J's paired with an IgD or IgM constant region. Clones were then identified in single-cell transcriptome and repertoire data sets that were related to clones present in pre-RTX bulk BCR repertoires. This allowed for direct interrogation of the transcriptomes of persistent B cells. That is, we leveraged the resolution offered by high-throughput single-cell transcriptomes and depth offered by bulk BCR analysis to perform single-cell tracing of adaptive immune repertoires (STAIR). From single-cell sequencing, a total of 18,133 cells associated with 14,671 IGH V(D)J sequences were identified (table S4) from the same three post-RTX patients and an additional asymptomatic AChR MG patient used as a control. Gene expression patterns were used to partition the cells into 11 distinct clusters (fig. S6A-E, table S5). Clusters of antibody-secreting cell-like (ASC), memory B cells, mature naive B cells and transitional B cells were assigned by analysis of the gene expression of these clusters (fig. 4A). Because B cells from the

single-cell expression analysis could not be experimentally isolated to assess for antibody secretion or other functional features associated with plasmablasts/plasma cells, we chose to label the cluster most associated with a plasmablast/plasma cell signature as antibody-secreting cell-like (hereafter referred to as ASC). By combining single-cell repertoires with bulk pre-RTX BCR repertoires, a total of 820 persistent clones with members in the single-cell profiling data were identified. Persistent clones could be identified in both the ASC as well as memory B cell clusters. The ASC cluster had a higher frequency (27.9%) of persistent clones in comparison to the memory cluster (14.6%) (fig. 4B, 4C) (t-test P=0.045). The persistence of both ASC and memory cells was not due to the mis-identification of clones, as IGH sequences were also identified from single-cells in the ASC and memory B cell cluster with identical V(D)J sequence as pre-RTX members (fig. 4B, red dots). Thus, while both ASCs and memory B cells are persistent, members of the ASC subset are more likely to be persistent than memory B cells during post-RTX relapse.

*ASCs have persistence-associated features compared to memory B cells*

ASCs are disease-relevant in MG and other autoantibody-mediated diseases. Moreover, we previously demonstrated their direct role in MuSK autoantibody production *(20, 21)*. We next sought to investigate phenotypic characteristics of ASCs compared to memory B cells during post-RTX relapse to further understand their mechanism of persistence and identify possible subsets of persistent ASCs. To correlate the persistence of ASCs with their expression of CD20 (the target of RTX), CD20 expression was examined among B cell subsets and the ASC cluster was observed to express lower CD20 levels when compared to memory B cells (t-test P=0.021) (fig. 5A). However, when we examined the expression of CD20 of persistent compared to non-persistent ASCs, a decrease was observed but this was not significant (t-test P=0.15) (fig. 5B). This finding was also observed when APRIL/BAFF receptor genes (TACI, BAFF-R and BCMA) were examined (P=0.085, 0.555, 0.288 for TACI, BAFF-R, BCMA respectively). Persistent ASCs were also not observed to differ in terms of isotype usage (ANOVA P=0.194) or frequency of SHM (ANOVA P=0.101) compared to their non-persistent counterparts (fig. S7A, B). The clonal diversity (more clonal expansions) of the ASC cluster was lower compared to memory B cells (t-test P=0.011) (fig. 5C). However, when we examined clonal diversity within the ASC cluster, a consistent difference between persistent compared to non-persistent ASCs was not observed (t-test P=0.079) (fig. 5D). The fraction of cells in the ASC cluster was also observed to be 10.0% of the total repertoire in the post-RTX cohort (fig. 5E). This was significantly elevated compared to the 0.5% observed in the control asymptomatic AChR MG sample (one-sample t-test P=0.044). However, when we examined if clonal expansion of persistent B cells between the pre-RTX and post-RTX time points could explain the abundance of ASC B cells that we observed, only some ASC B cell clones were found to be expanded as a fraction of the repertoire between pre-RTX and post-RTX time points; the average change in clonal size was not consistent across patients (t-test P=0.099) (fig. S8A). Collectively these data identify features of the ASC cluster as a whole compared to memory B cells that may explain their higher frequency among persistent clones: ASC cluster members expressed lower CD20 and were more clonally expanded. However, persistent ASC cluster members were not observed to differ from non-persistent members in terms of their expression of CD20 and receptors for BAFF/APRIL, clonal diversity, isotype usage or SHM frequency.

*Persistent memory B cell clones are expanded and express low levels of CD20*

Given the high frequency of persistent memory B cells observed in the data, we next sought to explore if persistent memory B cells had distinct features compared to non-persistent memory B cells. Persistent memory B cells were observed to have significantly lower CD20 expression (t-test P=0.022) and higher expression of TACI (t-test P=0.037), but lower expression of BAFF-R (t-test P=0.037) (fig. 6A). Analysis of SHM and isotype usage revealed that persistent memory B cells also had higher frequencies of SHM across isotypes (ANOVA P=0.037) (fig. 6B); the fraction of persistent memory B cells expressing IgG1 (t-test P=0.022) and IgG3 (t-test P=0.025) (ANOVA P= 0.007) was less than that of their non-persistent counterparts (fig. 6C). We examined if persistent memory B cells were clonally expanded compared to their non-persistent counterparts. They were observed to have lower clonal diversity post-RTX (t-test P=0.030) (fig. 6D). Initial clustering of cells based on single-cell gene expression identified six distinct sub-clusters of memory B cells. To better characterize the gene expression characteristics of persistent memory B cells (fig. S9A, B), we examined the most abundant sub-cluster among persistent memory B cells. This cluster (cluster 3) represented 36.7% of all persistent memory B cells but only 25.0% of all memory B cells (not the most prevalent overall) (fig. S9C). Analysis of gene expression showed that cluster 3 was characterized by lower CD20 expression (t-test P=0.005), higher TACI expression (t-test P=0.008) and lower BAFF-R expression (t-test P=0.014) compared to other memory B cell sub-clusters (fig. S10A). This cluster also displayed a higher frequency of SHM across all isotypes (ANOVA P=1.46e-6) (fig. S10C) and demonstrated increased usage of IgA compared to IgG isotypes (ANOVA P=0.0103) (fig. S10D), but was not more clonally expanded compared to other memory B cells (t-test P=0.169) (fig. S10E).

Unbiased differential gene expression identified 1,282 differentially expressed genes between this sub-cluster and all other memory B cell sub-clusters (data file S1) (FDR < 0.05). The upregulated subset of these genes was enriched for a signature for MYC (or c-Myc, a transcription factor that enhances B cell survival) and BRCA1/p53 (transcription factors that regulate cell senescence and cell cycle) (fig. S11B). The expression of two tissue homing genes that are the targets of the immunotherapy agents, natalizumab and fingolimod (ITGA4 and S1PR1 respectively), were also observed to be upregulated among persistent memory B cells (P=0.049 for ITGA4 and P=0.039 for S1PR1) (data file S1). Thus, persistent memory B cells are clonally expanded and are characterized by a subset of B cells with elevated SHM, high expression of TACI, low expression of CD20 and a gene expression pattern associated with transcription factors that control cell survival, proliferation and senescence.

**Discussion**

RTX is a B cell depleting agent widely used for the treatment of autoimmunity, however relapses after treatment have been reported in the context of diseases such as SLE, RA and AAVs *(12, 13, 44, 45)*. MG is a chronic autoimmune disorder characterized by muscle weakness and caused by pathogenic autoantibodies that bind to membrane proteins at the neuromuscular junction. Pathogenic autoantibodies to muscle-specific tyrosine kinase (MuSK) can be found in patients with MG who do not have detectable antibodies to the acetylcholine receptor. Recent studies have demonstrated the remarkable benefits of RTX-mediated B cell depletion therapy in MuSK MG *(5, 6, 46)*. In addition to diminished autoantibody titers and significant clinical improvement, RTX also allowed for tapering and subsequent discontinuation of other immunotherapies in subsets of MG patients. However, post-RTX relapses do occur *(5, 47)*. We recently demonstrated that relapses after RTX in MuSK MG patients are associated with the

presence of circulating plasmablasts that express MuSK autoantibodies *(20, 21)*. We sought to investigate if these disease-relevant and autoantigen-binding B cells derive either from the failed depletion of pre-existing or *de novo* B cell clones, and we further explored the characteristics of antigen-experienced B cells that emerged following B cell depletion in these patients.

In this study, we show that B cell clones relevant to disease, including autoantibody-secreting B cells present during relapse, can derive from the failed depletion of pre-existing B cell clones. Moreover, by directly tracing individual persistent B cell clones, we also show that persistent B cells use more isotype-switched constant regions, have elevated frequencies of SHM, and are clonally expanded. These are all features of an antigen-driven B cell response. We used a novel approach combining multiple methods of sequencing the repertoire with single-cell gene expression (STAIR) to demonstrate that persistent B cell clones are comprised of both expanded ASCs as well as memory B cell populations expressing low levels of CD20 associated with a unique gene expression profile. Collectively, these data directly show that circulating ASC and memory B cells, including disease-relevant B cells, are incompletely depleted by RTX in a subset of MuSK MG patients who experienced post-RTX relapses. We also characterized the repertoire and transcriptional features of these persistent ASCs and memory B cells in an unbiased fashion.

Our finding that circulating ASCs are persistent is consistent with a large body of evidence showing that ASCs, particularly those expressing lower levels of CD20, resist depletion owing to their CD20 expression and tissue localization *(22, 24, 25, 48)*. ASC frequency post-RTX is also known to be a reliable predictor of treatment response in SLE and RA patients treated with RTX *(11, 49–51)*. By using the repertoire to directly trace and quantify the frequency of these ASCs, we show that many ASCs that reconstitute the B cell compartment after therapy come from pre-existing, circulating B cell populations detectable prior to depletion.

The finding that circulating-memory B cell clones are also persistent is consistent with data showing that the frequency of memory populations after depletion predicts poor responses to RTX in SLE and RA *(15–17)*. While these studies were important for identifying outcome predictions, they were not designed to resolve B-cell subsets due to limitations inherent with flow cytometry. By comparison, our definition of memory using single-cell transcriptomes does not rely on flow cytometry markers, but rather thousands of markers that allow us to be confident these cells share common gene expression patterns associated with memory B cells. Furthermore, we were able to profile persistent memory B cell subsets in detail. For instance, a recent study using bulk AIRR-Seq to investigate the use of RTX for treating AAV, demonstrated that persistent clonotypes are isotype-switched which we also observed in our study *(52)*. However, because this analysis was not combined with single cell gene expression data, it is not clear if this was simply because the associated subsets that use these isotypes are more persistent *(52)*. In this study, persistence is correlated with a specific gene expression signature e.g. lower CD20 expression and differences in APRIL/BAFF receptor expression relative to other memory B cells.

Studies of RTX responses in anti-MAG-neuropathy (Immunoglobulin M(IgM) Anti-Myelin Associated Glycoprotein Peripheral Neuropathy) and pemphigus vulgaris (PV) *(17, 19, 31)* demonstrated that poor responses to RTX are associated with increased repertoire clonality post-

RTX. Our results show that a large fraction of memory B cells is persistent, and that these persistent memory B cells are more clonally expanded than non-persistent memory B cells. These findings suggest that the association between increased repertoire clonality and poor responses to RTX may result from the association between increased clonality and the persistence of memory B cell subsets.

Our work is congruent with studies that have shown that BAFF and the differential expression of receptors for APRIL/BAFF (TACI, BAFF-R, BCMA) on antigen-experienced B cells may lead to poor outcomes after BCDT *(28–30, 53)*. A higher frequency of memory B cells expressing lower BAFF-R can be detected in patients with RA relapsing after RTX *(30)* as well as in patients with GVHD who respond poorly to RTX *(29)*. This is consistent with our observation that persistent memory B cells express lower BAFF-R levels. BAFF levels have been shown to negatively correlate with BAFF-R expression after RTX therapy in primary Sjogren's Syndrome and SLE *(53)*. Moreover, extended exposure of B cells to soluble BAFF decreases surface levels of BAFF-R *in vitro*. By comparison, exposure to anti-BAFF agents has been shown to increase the expression of BAFF-R on B cells *(54)*. By directly characterizing the features of B cell clones that resist RTX depletion we are able to show that memory B cell subsets that persist after RTX express surface markers are associated with poor responses to RTX during relapse. These expression changes may all reflect the underlying effectiveness of B cell depletion.

Our analysis of bulk BCR sequencing showed that the largest fraction of persistent B cells was switched to the IgA isotype. STAIR analysis showed that two subsets observed to be most persistent, ASCs and memory B cells (particularly cluster 3) were also predominantly IgA. We speculate that the abundance of IgA expressing B cells could be related to BAFF expression given the role of BAFF (and receptors like TACI) in inducing class switch recombination, particularly in the generation of gut resident IgA-switched B cells *(55–57)*. The presence of IgA-switched circulating ASCs has been shown to be dependent on BAFF *(58)*. Moreover, IgA secreting plasma cells were shown as beneficial in an experimental autoimmune encephalomyelitis (EAE) mouse model, owing to their production of Il-10. Thus, further characterization of how the induction of IgA producing B cell subsets by RTX may be involved in active disease may be warranted to account for alternative mechanisms of therapeutic benefit from BCDT.

The conclusions of this study offer insights into possible therapeutic and diagnostic avenues for the management of patients with autoimmunity who experience relapses after RTX. Our findings regarding the BAFF/APRIL system suggest that combining RTX with BAFF pathway inhibitors (e.g. belimumab) could deplete populations that persist after RTX including those shown to produce autoantibodies. Combination strategies involving both belimumab and RTX have been successful in a phase 2a randomized clinical trial (RCT) in SLE which was followed with an ongoing phase 3 RCT, BLISS-BELIEVE *(59, 60)*. Similar investigations may be warranted in MG.

We observed that up-regulated genes associated with the subset of memory B cells most resistant to depletion (cluster 3) included those involved in tissue homing (ITGA4, S1PR1). S1P receptor is a gene that is up-regulated in lymphocytes localized to lymph nodes and inflamed tissues and target of the immunotherapy fingolimod. ITGA4 is a component of VLA-4, an integrin dimer

also responsible for the localization of leukocytes to inflamed tissue and target of the biologic agent natalizumab. We speculate that this could reflect a possible tissue-associated origin for this particular B cell subset. This is consistent with evidence showing that memory B cells that resist RTX depletion can be identified in lymph nodes, spleen, bone marrow and tonsils *(14, 15, 18, 61)*. These findings provide a compelling argument for alternative therapies that target the tissue localization of these subsets such as by engineered T cell technology including chimeric-antigen receptor (CAR) or chimeric-autoantibody receptor (CAAR) T cell therapy *(62, 63)*.

Finally, our ability to extract biologically meaningful results from STAIR analysis suggests that tracing individual BCR sequences associated with disease relapse could provide a powerful tool for monitoring patient responses to immunotherapies that rely on the depletion of B cells and T cells. Moreover, in comparison to techniques that use bulk BCR repertoire sequencing to monitor therapy progress, this approach allows for the individual characterization of lymphocytes that resist therapy based on gene expression which may suggest alternative avenues for treatment *(64)*.

The results from this study motivate future investigations. By combining the depth offered by bulk repertoire sequencing with the resolution of single-cell repertoire and gene expression, STAIR analysis allows for the characterization of B cells present at multiple time points in detail. Extending these observations to studies where anti-CD20 therapy has strong evidence for efficacy, such as MS may be particularly worthwhile. These studies could test the hypothesis that the depletion of common B cell subsets may predict efficacy across diseases. Applying STAIR analysis to the simultaneous investigation of other therapies such as anti-CD19 therapy (inebilizumab)—recently shown to be efficacious for the treatment of NMO—could allow for the optimal selection of B cell depleting agents for preventing disease relapse *(65)*. While RTX is a reliable treatment modality for MuSK MG, data for AChR MG has been mixed and patient relapses are common *(8)*. An investigation of factors that contribute to relapse after RTX in the context of AChR MG using similar approaches could lead to the development of improved therapeutics for diagnosing and treating patients with MG regardless of disease subtype.

Several limitations should be discussed in the context of this study. Our study characterized only the single-cell gene expression of B cells after RTX depletion. Determining the phenotype of B cells that resist RTX depletion prior to therapy may be of considerable interest as a way to identify subsets prone to resisting B cell depletion prior to relapse. A recent study demonstrated that not only do B cell clones move between peripheral blood and cerebrospinal fluid over the course of immunotherapy in a subset of patients with MS, but many shift between memory and plasma cell characteristics *(66)*. Similar studies for patients treated with RTX using single-cell transcriptomics could allow for the identification of memory B cell subsets that give rise to autoantibody-producing plasmablasts during relapse. Caveats should be noted given that our results rely on tracing B cell clones by analysis of their IGH V(D)J sequences alone. While we (and others) have demonstrated that clonal relationships can be reliably inferred given the diversity of IGH V(D)J sequences, false positives may occur *(43, 67, 68)*. Additionally, owing to the abundance of transitional B cells that re-populate the repertoire during RTX relapse and our interest in studying antigen-experienced B cell subsets, we specifically excluded IgD+ expressing B cell subsets from our single-cell analysis. Nevertheless, a large fraction of B cell receptors paired with an IGHD constant region (3.2-11.4%) was found in our single-cell data.

This is likely because sorting can flow-based sorting is not absolute in eliminating gated populations. Additionally, the expressed mRNA transcriptome may not always reflect the expression of surface markers used for sorting such as IgD. Given our ability to also capture the associated constant regions of reconstructed V(D)J sequences paired with single-cell gene expression, we were able to computationally filter residual IgD+ B cells (or IgD+ IgM+) by excluding B cells associated with either an IGHM or IGHD constant region.

In this study, disease-relevant B cells emerged from the failed depletion of pre-existing B cell clones. Characterization of post-RTX B cells by single-cell transcriptomics, combined with the identification of clones from paired B cell receptor repertoires identified both persistent ASC as well as memory B cell subsets. By combining the depth offered by bulk repertoire sequencing with the resolution of single-cell repertoire and gene expression, STAIR analysis allows for the characterization of B cells present at multiple times in detail. This could lead to the development and selection of therapies that target ASC and memory B cells that persist across time points, such as those associated with disease relapse.

## Materials and Methods

*Study Design*

The protocol used was approved by the Human Investigation Committee at Yale University School of Medicine. Peripheral blood was collected from MG and HD subjects (Yale Myasthenia Gravis Clinic, New Haven, Connecticut, USA). A MuSK MG cohort (n = 3) was identified for purposes of investigation of post-RTX relapses with these characteristics that we have described previously: (i) a RTX-induced CSR/MM (complete stable remission, minimal manifestation) clinical status (>1 year), (ii) repopulation of the B cell compartment after RTX, and (iii) disease relapse following sustained CSR/MM after rituximab *(20, 21)*. The patients in the study were not selected for only those patients who had never received rituximab prior to inclusion . An additional asymptomatic (by Myasthenia Gravis Foundation of America or MGFA classification score) AChR MG control subject was also included in our study.

*Bulk library Preparation*

RNA was prepared from frozen peripheral blood mononuclear cells using the RNAEasy Mini kit (Qiagen) per manufacturer's instructions. Cells were pelleted and lysed in freezing media without washing to preserve cell populations sensitive to cryopreservation. RNA libraries were made using reagents supplied by New England Biolabs. The RNA was reverse-transcribed into cDNA using a biotinylated oligo dT primer. An adaptor sequence, containing a universal priming site and a 17-nucleotide unique molecular identifier (UMI) was added to the 3' end of all cDNA. Following purification using streptavidin-coated magnetic beads, PCR was performed to enrich for immunoglobulin sequences using a pool of primers targeting the IGHA, IGHD, IGHE, IGHG, and IGHM regions. This immunoglobin-specific primer pool contained tailed sequences with a priming site for a secondary PCR step. The second primer is specific to the adaptor sequence added during the RT step, and contains a sample index for downstream pooling of samples prior to sequencing. Following purification of PCR products using AMPure XP beads (Beckman), the secondary PCR was performed in order to add the full-length Illumina P5 Adaptor sequence to the end of each immunoglobin amplicon. The number of secondary PCR cycles was tailored to each sample to avoid entering plateau phase, as judged by a prior

quantitative PCR analysis. Final products were purified, quantified with a TapeStation (Agilent Genomics) and pooled in equimolar proportions, followed by high-throughput 2x300 base-pair paired-end sequencing with a 20% PhiX spike on the Illumina MiSeq platform according to manufacturer's recommendations, except for performing 325 cycles for read 1 and 275 cycles for read 2.

*Raw read quality control and assembly for bulk libraries*

Bulk repertoire data processing and analysis was carried out using tools in the Immcantation framework (http://immcantation.org). Preprocessing was carried out using pRESTO v0.5.10 *(69)* as follows. 1) Reads with a mean Phred quality score below 20 were removed. 2) Reads were aligned against constant region primer and template switch sequences, with a maximum mismatch rate of 0.2 and 0.5 respectively. Reads failing to match both a constant region primer and template switch sequence were removed. 3) Remaining reads were grouped based on unique molecular identifier (UMI) sequences determined by the 17 nucleotides preceding the template switch site. 4) Separate consensus sequences for the forward and reverse reads within each UMI group were constructed with a maximum error score of 0.1 and minimum constant region primer frequency of 0.6. If multiple constant region primers were associated with a particular UMI group, the majority primer was used. 5) Forward and reverse consensus sequence pairs were aligned to each other with a minimum overlap of 8 nucleotides and a maximum mismatch rate of 0.3, and assembled into full length sequences. Sequence pairs that failed to align to each other were assembled by alignment against the IMGT human germline IGHV reference database (IMGT/GENE-DB v3.1.19; retrieved June 21, 2018; Giudicelli et al. 2005) with a minimum overlap of 0.5 and a E-value threshold of 1x10-5. 6) Isotypes were assigned by local alignment of the 3' end of each consensus sequence to isotype-specific constant region sequences with a maximum mismatch rate of 0.4. 8) Duplicate sequences were removed, except those spanning multiple biological samples and/or with different isotype assignments. Constant region primers, isotype-specific internal constant region sequences, and template switch sequences used are available at: https://bitbucket.org/kleinstein/immcantation/src/default/protocols/AbSeq.

*Single-cell library preparation and gene expression analysis*

B cells were isolated from PBMCs by Human Pan B cell immunomagnetic negative selection kit (StemCell) per manufacturer's instructions. Additional fluorescence-activated cell sorting (FACS) was performed for bead sorted B cells from MuSK MG post-RTX derived samples. B cells were labelled with 7AAD Viability Staining Solution (BioLegend) and 7AAD positive populations were removed. Additionally, B cells were labelled with FITC-labelled anti-IgD antibody (BioLegend clone IA6-2) per manufacturer's instructions and FITC positive populations were removed. Of note, the purpose of this additional filtering was the removal of all naive B cell populations. Sorted B cells were loaded into the Chromium Controller (10x Genomics) to form emulsion droplets. Libraries were prepared using the Chromium Single-cell 5′ Reagent Kit (10x Genomics) for version 3 chemistry per the manufacturer's protocol. Samples were sequenced on the NovaSeq with 100x100 or 150x150 paired end reads for gene expression and BCR libraries respectively. Base calls were converted to fastq sequences and demultiplexed using the cellranger mkfastq function from 10X Genomics 2.2.0 and aligned to the GCRhg38 genome supplied by 10X Genomics. Sparse count matrices, barcode assignments and feature calls were made using the cellranger count function. The resulting output was loaded into Seurat v.2.3.4 for analysis. Cells with less than 50 genes detected, or mitochondrial content above 0.2 of all transcripts were discarded. Individual cells expressing detectable CST3, CD3E, KLRB1,

NKG7 and GATA2 transcripts were excluded to remove non-B cell populations like T cells and myeloid cells. Single-cell expression values were computed from log normalized count values *(70)*. 12,534 cells were identified in the final filtered analysis. Clusters were assigned and tSNE was computed after first scaling the data, removing cells with elevated mitochondrial content, identifying highly variable genes using FindVariableGenes and correcting for batch effects using the first 20 dimensions from canonical correlation analysis as identified by visualizing with MetageneBicorPlot. The first 10 dimensions were used for PCA based on the location of the inflection point from PCElbowPlot. FindClusters was then used to identify clusters in Seurat v.2 which uses shared nearest neighbor clustering after PCA *(70)*.

Cell subsets were identified using a "basis set" of published marker genes called immunoStates *(71)*. The mean expression of each gene was computed for each cluster from the scaled log normalized expression values; each cluster was then assigned to the immunoState subset with the highest Pearson correlation value. The absence of transitional B cells in the immunostates dataset meant this cluster was identified manually using marker genes alone. Small clusters with elevated expression of mitochondrial associated genes were also excluded from further analysis. We confirmed our assignments by visualizing the expression of known gene expression markers for instance, CD27 (for memory B cells), CD20 (for B cells in general except plasma cells), CD10 (for transitional B cells) and BLIMP1 (for plasma cells) (fig. S6A-C). Overall marker expression was also visualized based on the percent of cells in each cluster expressing the gene and the overall mean level of gene expression in each cluster (fig. S6D,E). To further confirm the identities of these single-cell clusters, we assessed their paired IGH repertoire features for mean somatic hypermutation (SHM) frequency and isotype usage which were consistent with expected values (fig. S6F,G).

*Bulk BCR repertoire V(D)J gene annotation, additional sequence filtering, clonal assignment and germline reconstruction for Single-cell tracing of adaptive immune repertoires (STAIR)*

V(D)J germline assignments were made with IgBLAST v1.7.0 *(72)* using the June 21, 2018 IMGT reference gene database from bulk libraries. Reconstructed V(D)J sequences from single cell sequencing were extracted using the cellranger vdj function from fastq reads. V(D)J germline segments were re-assigned using IgBLAST v.1.7.0 also using the June 21, 2018 version of the IMGT gene database. Cells with multiple IGH V(D)J sequences were assigned to the most abundant IGH V(D)J sequence by UMI count. Single-cell V(D)J BCR repertoires as well as previously published Sanger sequencing derived V(D)J sequences were pooled with bulk V(D)J BCR repertoires *(20, 21)*. Following V(D)J annotation, non-functional sequences were removed from further analysis and functional V(D)J sequences were assigned into clonal groups using Change-O v0.3.4 *(43)*. Sequences were first partitioned based on common IGHV gene annotations, IGHJ gene annotations, and junction lengths. Within these groups, sequences differing from one another by a length normalized Hamming distance of 0.126 were defined as clones by single-linkage clustering. This distance threshold was determined based on the average local minima between the two modes of the within-sample bimodal distance-to-nearest histogram for the three patients across multiple time points; thresholds were determined using a kernel density estimate in the same manner described previously *(43)*. Germline sequences were then reconstructed for each clonal cluster (VH) with D segment and N/P regions masked (replaced with Ns) using the CreateGermlines.py function within Change-O v0.3.4.

*Differential gene expression analysis*

Single-cell log normalized expression values were imputed using the ALRA algorithm to account for dropout during differential gene expression analysis *(73)*. The mean value for each gene per sample was subtracted, the mean expression value for cells of a given status and sample was computed and values were z-score normalized for each patient. Student's t-tests without pairing were performed to assign p-values for hypothesis testing. False detection rate (FDR) correction was performed using Storey's method implemented within the qvalue package v3.9 in R from Bioconductor on the 12,066 p-values extracted from our analysis of cluster 3 associated genes *(74)*. Gene ontologies were assigned using the enrichr package v1.0 in R which implements a wilcoxon signed rank test for significance *(75)*. A q-value (corrected p-value) threshold of 0.05 was used for assigning significance for all tests of significance.

*Analysis of SHM, diversity and clonal overlap from BCR repertoires*

Mutations relative to the germline sequence were quantified using SHazaM v0.1.8 in R v3.4.2 *(76)*. Diversity analysis was performed using Alakazam v0.2.11 and assessed using the generalized Hill index, with uniform downsampling (to the number of V(D)J from the sample with the fewest sequences to account for different sequencing depth) and 2,000 replicates of bootstrapping from the inferred complete clonal abundance *(77)*. Clonal overlap was computed using a bray-curtis metric implemented by the function scipy.spatial.distance.braycurtis in scipy v1.1.0 *(78)*.

*Inference of B cell lineage trees and prediction of intermediate timepoints*

B cell lineage tree topologies and branch lengths were estimated using the dnapars program distributed as part of PHYLIP (v3.697) *(79)*. A maximum parsimony algorithm was used to label the internal nodes of each lineage tree as either or pre-RTX, post-RTX, given the date of sampling of each sequence at the tree's tips. This was accomplished using an implementation of the Sankoff parsimony algorithm *(80)* in which switches from post-RTX to pre-RTX are forbidden using a cost matrix that weights post-RTX to pre-RTX switches as 1000 times more costly than switches in the opposite direction. Clusters of internal nodes separated by zero length branches (polytomies), were re-ordered using nearest-neighbor interchange (NNI) moves to minimize the number of label switches along the tree. These analyses were performed using IgPhyML v1.10 *(81)*.

*Statistical analysis*

R v.3.4.2 *(76)* and Python 3.5.4 was used for all statistical analysis. Dataframe handling and plotting was performed using functions from the tidyverse 1.2.1 in R and pandas 0.24.2, scipy 1.1.0 and matplotlib 2.2.2 in python. All parametric statistical testing aside from that used for differential gene expression analysis (described previously) was performed in R using the aov function for two-way ANOVA or t.test functions for paired two-tailed student t-tests (or one-tailed student t-test when appropriate). A significance threshold of $<0.05$ was used and shown on plots with a single asterisk; double asterisks correspond to a $p<0.01$ and triple asterisks correspond to a $p<0.001$.

## Supplementary Materials

Table S1. Demographic characteristics of MuSK MG study subjects at each collection time point.

Table S2. Counts of reconstructed V(D)J sequences by isotype and clones from sequencing of pre-RTX and post-RTX bulk IGH repertoires.

Fig. S1. Timeline of three patients who experienced post-RTX relapses. Symptom score based on MGFA class and MGC scores were recorded at all collection time points.

Fig. S2. Distance-to-nearest plots used to identify a common threshold to use for hierarchical clustering based grouping of V(D)J sequences from high throughout sequencing of BCR repertoires.

Fig. S3. The presence of B cell clonal expansions and somatic hypermutation cannot be detected from analysis of the bulk IGH repertoire.

Table S3. Count of plasmablast derived or tetramer-binding V(D)J sequences clonally related to members of the pre-RTX or post-RTX bulk BCR repertoire (Traceable clones) or to members of the pre-RTX and post-RTX bulk BCR repertoire (Pre-RTX clones).

Fig. S4. Clonal lineage of clonally expanded disease-relevant antibodies from Sanger sequencing identified in post-RTX repertoires.

Table S4. Quality control from paired single-cell transcriptome and repertoire sequencing.

Fig. S5. Example fluorescence activated cell sorting (FACS) gates for IgDlo B cells used as input for single-cell transcriptomics and BCR repertoire analysis.

Table S5. Table of cluster assignments using the immunoStates basis set.

Fig. S6. Assignment of B cell clusters based on single-cell transcriptome and repertoire according to known B cell subsets.

Fig. S7. Persistent and non-persistent ASC-like cells are associated with similar V(D)J repertoire.

Fig S8. The overall accumulation of SHM and presence of clonal expansions between pre-RTX and post-RTX time points cannot be detected for ASCs and memory B cells.

Fig. S9. Identification of B cell clusters associated with B cell clonal persistence through unbiased analysis identifies subset associated with resistance to RTX depletion (cluster 3).

Fig. S10. Cluster 3 has distinct transcriptional and V(D)J repertoire features.

Data file S1. List of differentially expressed genes for cluster 3 compared to all other memory B cell clusters.

Fig. S11. Cluster 3 memory B cells have a MYC and p53/BRCA1 signature.

**References and Notes:**

1. P. Emery, R. Fleischmann, A. Filipowicz-Sosnowska, J. Schechtman, L. Szczepanski, A. Kavanaugh, A. J. Racewicz, R. F. Van Vollenhoven, N. F. Li, S. Agarwal, E. W. Hessey, T. M. Shaw, The efficacy and safety of rituximab in patients with active rheumatoid arthritis despite methotrexate treatment: Results of a phase IIb randomized, double-blind, placebo-controlled, dose-ranging trial, *Arthritis Rheum.* **54**, 1390–1400 (2006).

2. S. L. Hauser, E. Waubant, D. L. Arnold, T. Vollmer, J. Antel, R. J. Fox, A. Bar-Or, M.

Panzara, N. Sarkar, S. Agarwal, A. Langer-Gould, C. H. Smith, B-Cell Depletion with Rituximab in Relapsing–Remitting Multiple Sclerosis, *N. Engl. J. Med.* **358**, 676–688 (2008).

3. R. B. Jones, J. W. C. Tervaert, T. Hauser, R. Luqmani, M. D. Morgan, C. A. Peh, K. Westman, D. R. W. Jayne, E. Vasculitis, S. Group, Rituximab versus Cyclophosphamide in ANCA-Associated Renal Vasculitis, *N. Engl. J. Med.* **363**, 211–220 (2010).

4. J. Mcconville, U. Kishore, R. Metcalfe, J. Newsom-davis, Detection and Characterization of MuSK Antibodies in Seronegative Myasthenia Gravis, *Ann. Neurol.* **55**, 580–584 (2004).

5. R. J. Nowak, D. B. Dicapua, N. Zebardast, J. M. Goldstein, Response of patients with refractory myasthenia gravis to rituximab : a retrospective study, *Ther. Adv. Neurol. Disord.* **4**, 259–266 (2011).

6. I. Illa, J. Diaz-manera, R. Rojas-garcia, J. Pradas, A. Rey, R. Blesa, C. Juarez, E. Gallardo, Sustained response to Rituximab in anti-AChR and anti-MuSK positive, **202**, 90–94 (2008).

7. B. Keung, R. R. Kimberly, D. B. DiCapua, J. B. Rosen, K. C. O'Connor, J. M. Goldstein, R. J. Nowak, Long-term bene fi t of rituximab in MuSK autoantibody myasthenia gravis patients, *J Neurol Neurosurg Psychiatry* **84**, 1407–1410 (2013).

8. K. R. Robeson, A. Kumar, B. Keung, D. B. DiCapua, E. Grodinsky, H. S. Patwa, P. A. Stathopoulos, J. M. Goldstein, K. C. O'Connor, R. J. Nowak, Durability of the rituximab response in acetylcholine receptor autoantibody-positivemyasthenia gravis, *JAMA Neurol.* **74**, 60–66 (2017).

9. S. Blum, D. Gillis, H. Brown, R. Boyle, R. Henderson, D. Heyworth-Smith, P. Hogan, P. Kubler, C. Lander, N. Limberg, P. Pillans, K. Prain, C. Staples, M. Walsh, P. McCombe, R. Wong, Use and monitoring of low dose rituximab in myasthenia gravis, *J. Neurol. Neurosurg. Psychiatry* **82**, 659–663 (2011).

10. J. T. Guptill, D. B. Sanders, A. Evoli, Anti-MuSK antibody Myasthenia gravis: Clinical findings and response to treatment in two large cohorts, *Muscle and Nerve* **44**, 36–40 (2011).

11. M. H. Stradner, C. Dejaco, K. Brickmann, W. B. Graninger, H. P. Brezinschek, A combination of cellular biomarkers predicts failure to respond to rituximab in rheumatoid arthritis: A 24-week observational study, *Arthritis Res. Ther.* **18**, 1–8 (2016).

12. K. G. C. Smith, R. B. Jones, S. M. Burns, D. R. W. Jayne, Long-term comparison of rituximab treatment for refractory systemic lupus erythematosus and vasculitis: Remission, relapse, and re-treatment, *Arthritis Rheum.* **54**, 2970–2982 (2006).

13. B. Terrier, Z. Amoura, P. Ravaud, E. Hachulla, R. Jouenne, B. Combe, C. Bonnet, P. Cacoub, A. Cantagrel, M. De Bandt, O. Fain, B. Fautrel, P. Gaudin, B. Godeau, J. R. Harlé, A. Hot, J. E. Kahn, O. Lambotte, C. Larroche, J. Léone, O. Meyer, B. Pallot-Prades, E. Pertuiset, P. Quartier, T. Schaerverbeke, J. Sibilia, A. Somogyi, M. Soubrier, E. Vignon, B. Bader-Meunier, X. Mariette, J. E. Gottenberg, Safety and efficacy of rituximab in systemic lupus erythematosus: Results from 136 patients from the French autoimmunity and rituximab registry, *Arthritis Rheum.* **62**, 2458–2466 (2010).

14. M. J. Leandro, G. Cambridge, M. R. Ehrenstein, J. C. W. Edwards, Reconstitution of peripheral blood B cells after depletion with rituximab in patients with rheumatoid arthritis, *Arthritis Rheum.* **54**, 613–620 (2006).

15. J. H. Anolik, J. Barnard, T. Owen, B. Zheng, S. Kemshetti, R. J. Looney, I. Sanz, Delayed memory B cell recovery in peripheral blood and lymphoid tissue in systemic lupus erythematosus after B cell depletion therapy, *Arthritis Rheum.* **56**, 3044–3056 (2007).

16. D. G. Adlowitz, J. Barnard, J. N. Biear, C. Cistrone, T. Owen, W. Wang, A. Palanichamy, E. Ezealah, D. Campbell, C. Wei, R. J. Looney, I. Sanz, J. H. Anolik, Expansion of activated peripheral blood memory B cells in rheumatoid arthritis, impact of B cell depletion therapy, and biomarkers of response, *PLoS One* **10** (2015), doi:10.1371/journal.pone.0128269.

17. M. Nakou, G. Katsikas, P. Sidiropoulos, G. Bertsias, E. Papadimitraki, A. Raptopoulou, H. Koutala, H. A. Papadaki, H. Kritikos, D. T. Boumpas, Rituximab therapy reduces activated B cells in both the peripheral blood and bone marrow of patients with rheumatoid arthritis: depletion of memory B cells correlates with clinical response, *Arthritis Res. Ther.* **11**, 1–8 (2009).

18. M. Mamani-matsuda, A. Cosma, S. Weller, A. Faili, C. Staib, O. Beyne-rauzy, C. Fieschi, J. Pers, N. Arakelyan, B. Varet, A. Sauvanet, The human spleen is a major reservoir for long-lived vaccinia virus – specific memory B cells, *Blood* **111**, 4653–4660 (2018).

19. N. Colliou, D. Picard, F. Caillot, S. Calbo, S. Le Corre, A. Lim, B. Lemercier, B. Le Mauff, M. Maho-vaillant, S. Jacquot, C. Bedane, P. Bernard, F. Caux, C. Prost, E. Delaporte, M. Doutre, B. Dreno, N. Franck, S. Ingen-housz-oro, M. Sigal, E. Tancrede-bohin, I. Templier, R. Eming, Long-Term Remissions of Severe Pemphigus After Rituximab Therapy Are Associated with Prolonged Failure of Desmoglein B Cell Response, *Sci. Transl. Med.* **5**, 175ra30 (2013).

20. P. Stathopoulos, A. Kumar, R. J. Nowak, K. C. O. Connor, Autoantibody-producing plasmablasts after B cell depletion identified in muscle-specific kinase myasthenia gravis, *JCI Insight* **2**, e94263 (2017).

21. K. Takata, P. Stathopoulos, M. Cao, M. Mané-Damas, M. L. Fichtner, E. S. Benotti, L. Jacobson, P. Waters, S. R. Irani, P. Martinez-Martinez, D. Beeson, M. Losen, A. Vincent, R. J. Nowak, K. C. O'Connor, Characterization of pathogenic monoclonal autoantibodies derived from muscle-specific kinase myasthenia gravis patients, *JCI Insight* **4** (2019), doi:10.1172/jci.insight.127167.

22. H. Huang, C. Benoist, D. Mathis, Rituximab specifically depletes short-lived autoreactive plasma cells in a mouse model of in fl ammatory arthritis, *Proc. Natl. Acad. Sci.* **107**, 4658–4663 (2010).

23. A. Cho, B. Bradley, R. Kauffman, L. Priyamvada, Y. Kovalenkov, R. Feldman, J. Wrammert, Robust memory responses against influenza vaccination in pemphigus patients previously treated with rituximab, *JCI Insight* **2** (2017), doi:10.1172/jci.insight.93222.

24. F. Medina, C. Segundo, A. Campos-caro, The heterogeneity shown by human plasma cells from tonsil, blood, and bone marrow reveals graded stages of increasing maturity , but local profiles of adhesion molecule expression, *Blood* **99**, 2154–2162 (2018).

25. D. R. Withers, C. Fiorini, R. T. Fischer, R. Ettinger, P. E. Lipsky, A. C. Grammer, T cell – dependent survival of CD20+ and CD20- plasma cells in human secondary lymphoid tissue, *Blood* **109**, 4856–4865 (2018).

26. B. Nardelli, O. Belvedere, V. Roschke, P. A. Moore, H. S. Olsen, T. S. Migone, S. Sosnovtseva, J. A. Carrell, P. Feng, J. G. Giri, D. M. Hilbert, Synthesis and release of B-

lymphocyte stimulator from myeloid cells, *Blood* **97**, 198–204 (2001).

27. P. Schneider, F. Mackay, V. Steiner, K. Hofmann, J. L. Bodmer, N. Holler, C. Ambrose, P. Lawton, S. Bixler, H. Acha-Orbea, D. Valmori, P. Romero, C. Werner-Favre, R. H. Zubler, J. L. Browning, J. Tschopp, BAFF, a novel ligand of the tumor necrosis factor family, stimulates B cell growth, *J. Exp. Med.* **189**, 1747–1756 (1999).

28. J. O. Pers, V. Devauchelle, C. Daridon, B. Bendaoud, R. Le Berre, A. Bordron, P. Hutin, Y. Renaudineau, M. Dueymes, S. Loisel, C. Berthou, A. Saraux, P. Youinou, BAFF-modulated repopulation of B lymphocytes in the blood and salivary glands of rituximab-treated patients with Sjoğren's syndrome, *Arthritis Rheum.* **56**, 1464–1477 (2007).

29. S. Sarantopoulos, K. E. Stevenson, H. T. Kim, W. S. Washel, N. S. Bhuiya, C. S. Cutler, E. P. Alyea, V. T. Ho, R. J. Soiffer, J. H. Antin, J. Ritz, Recovery of B-cell homeostasis after rituximab in chronic graft-versus-host disease, *Blood* **117**, 2275–2283 (2011).

30. E. Becerra, I. De La Torre, M. J. Leandro, G. Cambridge, B cell phenotypes in patients with rheumatoid arthritis relapsing after rituximab: expression of B cell-activating factor-binding receptors on B cell subsets, *Clin. Exp. Immunol.* **190**, 372–383 (2017).

31. M. A. Maurer, F. Tuller, V. Gredler, T. Berger, A. Lutterotti, J. D. Lünemann, M. Reindl, Rituximab induces clonal expansion of IgG memory B-cells in patients with inflammatory central nervous system demyelination, *J. Neuroimmunol.* **290**, 49–53 (2016).

32. M. A. Maurer, G. Rakocevic, C. S. Leung, I. Quast, M. Lukačišin, N. Goebels, C. Münz, H. Wardemann, M. Dalakas, J. D. Lünemann, Rituximab induces sustained reduction of pathogenic B cells in patients with peripheral nervous system autoimmunity, *J. Clin. Invest.* **122**, 1393–1402 (2012).

33. S. D. Boyd, E. L. Marshall, J. D. Merker, J. M. Maniar, L. N. Zhang, B. Sahaf, C. D. Jones, B. B. Simen, B. Hanczaruk, K. D. Nguyen, K. C. Nadeau, M. Egholm, D. B. Miklos, J. L. Zehnder, A. Z. Fire, Measurement and Clinical Monitoring of Human Lymphocyte Clonality by Massively Parallel V-D-J Pyrosequencing, *Sci. Transl. Med.* **1**, 12ra23 (2009).

34. U. Laserson, F. Vigneault, D. Gadala-maria, G. Yaari, M. Uduman, High-resolution antibody dynamics of vaccine-induced immune responses, **111**, 1–6 (2014).

35. U. Hershberg, E. T. Luning Prak, The analysis of clonal expansions in normal and autoimmune B cell repertoires, *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140239 (2015).

36. B. J. DeKosky, O. I. Lungu, D. Park, E. L. Johnson, W. Charab, C. Chrysostomou, D. Kuroda, A. D. Ellington, G. C. Ippolito, J. J. Gray, G. Georgiou, Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires, *Proc. Natl. Acad. Sci.* **113**, E2636–E2645 (2016).

37. M. Singh, G. Al-eryani, S. Carswell, J. M. Ferguson, K. Barton, D. Roden, F. Luciani, T. Phan, A. Swarbrick, High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes, *Nat. Commun.* **10**, 3120 (2018).

38. B. W. Dulken, M. T. Buckley, P. Navarro Negredo, N. Saligrama, R. Cayrol, D. S. Leeman, B. M. George, S. C. Boutet, K. Hebestreit, J. V. Pluvinage, T. Wyss-Coray, I. L. Weissman, H. Vogel, M. M. Davis, A. Brunet, Single-cell analysis reveals T cell infiltration in old neurogenic niches, *Nature* (2019), doi:10.1038/s41586-019-1362-5.

39. J. E. Oh, N. Iijima, E. Song, P. Lu, J. Klein, R. Jiang, S. H. Kleinstein, A. Iwasaki, Migrant memory B cells secrete luminal antibody in the vagina, *Nature* **571**, 122–26 (2019).

40. U. Hershberg, W. Meng, B. Zhang, N. Haff, E. W. S. Clair, P. L. Cohen, P. D. Mcnair, L. Li, M. C. Levesque, E. T. L. Prak, Persistence and selection of an expanded B-cell clone in the setting of rituximab therapy for Sjögren' s syndrome Persistence and selection of an expanded B-cell clone in the setting of rituximab therapy for Sjögren' s syndrome, *Arthritis Res. Ther.* **16**, 1–12 (2014).

41. C. F. A. de Bourcy, C. L. Dekker, M. M. Davis, M. R. Nicolls, S. R. Quake, Dynamics of the human antibody repertoire after B cell depletion in systemic sclerosis., *Sci. Immunol.* **2**, eaan8289 (2017).

42. D. Theil, P. Smith, C. Huck, Y. Gilbart, A. Kakarieka, D. Leppert, C. Rauld, C. Schmid, R. Baumgartner, N. Stuber, F. Cordoba, V. Dubost, K. Darribat, M. Jivkov, W. Frieauff, R. Kneuer, M. Stoeckli, S. Reinker, K. Mansfield, J. M. Carballido, P. Couttet, G. Weckbecker, Imaging Mass Cytometry and Single-Cell Genomics Reveal Differential Depletion and Repletion of B-Cell Populations Following Ofatumumab Treatment in Cynomolgus Monkeys, *Front. Immunol.* **10**, 1–11 (2019).

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

44. A. A. N. Giagounidis, J. Anhuf, P. Schneider, U. Germing, D. Söhngen, K. Quabeck, C. Aul, Treatment of relapsed idiopathic thrombocytopenic purpura with the anti-CD20 monoclonal antibody rituximab: a pilot study., *Eur. J. Haematol.* **69**, 95–100 (2002).

45. T. Turner-Stokes, T. Y. Lu, M. R. Ehrenstein, I. Giles, A. Rahman, D. A. Isenberg, The efficacy of repeated treatment with B-cell depletion therapy in systemic lupus erythematosus: an evaluation., *Rheumatology (Oxford).* **50**, 1401–1408 (2011).

46. J. Díaz-Manera, E. Martínez-Hernández, L. Querol, R. Klooster, R. Rojas-Garciá, X. Suaŕez-Calvet, J. L. Munõz-Blanco, C. Mazia, K. R. Straasheijm, E. Gallardo, C. Juaŕez, J. J. Verschuuren, I. Illa, Long-lasting treatment effect of rituximab in MuSK myasthenia, *Neurology* **78**, 189–193 (2012).

47. J. T. Guptill, J. S. Yi, D. B. Sanders, A. C. Guidon, V. C. Juel, J. M. Massey, J. F. Howard, A. Evoli, K. J. Weinhold, Characterization of B cells in muscle-specific kinase antibody myasthenia gravis, *Neurol. - Neuroimmunol. Neuroinflammation* **2**, 1–9 (2015).

48. E. Hammarlund, A. Thomas, I. J. Amanna, L. A. Holden, O. D. Slayden, B. Park, L. Gao, M. K. Slifka, Plasma cell survival in the absence of B cell memory, *Nat. Commun.* **8** (2017), doi:10.1038/s41467-017-01901-w.

49. M. Y. Md Yusof, D. Shaw, Y. M. El-Sherbiny, E. Dunn, A. C. Rawstron, P. Emery, E. M. Vital, Predicting and managing primary and secondary non-response to rituximab using B-cell biomarkers in systemic lupus erythematosus, *Ann. Rheum. Dis.* **76**, 1829–1836 (2017).

50. R. M. Thurlings, K. Vos, C. A. Wijbrandts, A. H. Zwinderman, D. M. Gerlag, P. P. Tak, Synovial tissue response to rituximab: Mechanism of action and identification of biomarkers of response, *Ann. Rheum. Dis.* **67**, 917–925 (2008).

51. E. M. Vital, S. Dass, M. H. Buch, K. Henshaw, C. T. Pease, M. F. Martin, F. Ponchel, A. C. Rawstron, P. Emery, B cell biomarkers of rituximab responses in systemic lupus erythematosus, *Arthritis Rheum.* **63**, 3038–3047 (2011).

52. R. J. M. Bashford-Rogers, L. Bergamaschi, E. F. McKinney, D. C. Pombal, F. Mescia, J. C. Lee, D. C. Thomas, S. M. Flint, P. Kellam, D. R. W. Jayne, P. A. Lyons, K. G. C. Smith, Analysis of the B cell receptor repertoire in six immune-mediated diseases, *Nature* **574**, 122–126 (2019).

53. J. Sellam, C. Miceli-Richard, J. E. Gottenberg, M. Ittah, F. Lavie, C. Lacabaratz, N. Gestermann, A. Proust, O. Lambotte, X. Mariette, Decreased B cell activating factor receptor expression on peripheral lymphocytes associated with increased disease activity in primary Sjögren's syndrome and systemic lupus erythematosus, *Ann. Rheum. Dis.* **66**, 790–797 (2007).

54. E. Pontarini, M. Fabris, L. Quartuccio, M. Cappeletti, F. Calcaterra, A. Roberto, F. Curcio, D. Mavilio, S. Della Bella, S. De Vita, Treatment with belimumab restores B cell subsets and their expression of B cell activating factor receptor in patients with primary Sjogren's syndrome, *Rheumatol. (United Kingdom)* **54**, 1429–1434 (2015).

55. M. B. Litinskiy, B. Nardelli, D. M. Hilbert, B. He, A. Schaffer, P. Casali, A. Cerutti, DCs induce CD40-independent immunoglobulin class switching through BLyS and APRIL, *Nat. Immunol.* **3**, 822–829 (2002).

56. W. Xu, B. He, A. Chiu, A. Chadburn, M. Shan, M. Buldys, A. Ding, D. M. Knowles, P. A. Santini, A. Cerutti, Epithelial cells trigger frontline immunoglobulin class switching through a pathway regulated by the inhibitor SLPI, *Nat. Immunol.* **8**, 294–303 (2007).

57. E. Castigli, S. A. Wilson, S. Scott, F. Dedeoglu, S. Xu, K. P. Lam, R. J. Bram, H. Jabara, R. S. Geha, TACI and BAFF-R mediate isotype switching in B cells, *J. Exp. Med.* **201**, 35–39 (2005).

58. O. L. Rojas, A. K. Pröbstel, E. A. Porfilio, A. A. Wang, M. Charabati, T. Sun, D. S. W. Lee, G. Galicia, V. Ramaglia, L. A. Ward, L. Y. T. Leung, G. Najafi, K. Khaleghi, B. Garcillán, A. Li, R. Besla, I. Naouar, E. Y. Cao, P. Chiaranunt, K. Burrows, H. G. Robinson, J. R. Allanach, J. Yam, H. Luck, D. J. Campbell, D. Allman, D. G. Brooks, M. Tomura, R. Baumann, S. S. Zamvil, A. Bar-Or, M. S. Horwitz, D. A. Winer, A. Mortha, F. Mackay, A. Prat, L. C. Osborne, C. Robbins, S. E. Baranzini, J. L. Gommerman, Recirculating Intestinal IgA-Producing Cells Regulate Neuroinflammation via IL-10, *Cell* **176**, 610-624.e18 (2019).

59. Y. K. Onno Teng, I. N. Bruce, B. Diamond, R. A. Furie, R. F. Van Vollenhoven, D. Gordon, J. Groark, R. B. Henderson, M. Oldham, P. P. Tak, Phase III, multicentre, randomised, double-blind, placebo-controlled, 104-week study of subcutaneous belimumab administered in combination with rituximab in adults with systemic lupus erythematosus (SLE): BLISS-BELIEVE study protocol, *BMJ Open* **9**, 1–8 (2019).

60. T. Kraaij, S. W. A. Kamerling, E. N. M. de Rooij, P. L. A. van Daele, O. W. Bredewold, J. A. Bakker, I. M. Bajema, H. U. Scherer, R. E. M. Toes, T. J. W. Huizinga, T. J. Rabelink, C. van Kooten, Y. K. O. Teng, The NET-effect of combining rituximab with belimumab in severe systemic lupus erythematosus, *J. Autoimmun.* **91**, 45–54 (2018).

61. T. H. Ramwadhdoebe, L. G. M. van Baarsen, M. J. H. Boumans, S. T. G. Bruijnen, M. Safy, F. H. Berger, J. F. Semmelink, C. J. van der Laken, D. M. Gerlag, R. M. Thurlings, P. P. Tak,

Effect of rituximab treatment on T and B cell subsets in lymph node biopsies of patients with rheumatoid arthritis, *Rheumatology (Oxford).* **58**, 1075–1085 (2019).

62. C. T. Ellebrecht, V. G. Bhoj, A. Nace, E. J. Choi, X. Mao, M. J. Cho, G. Di Zenzo, A. Lanzavecchia, J. T. Seykora, G. Cotsarelis, M. C. Milone, A. S. Payne, Reengineering chimeric antigen receptor T cells for targeted therapy of autoimmune disease, *Science (80-. ).* **353**, 179–184 (2017).

63. R. Kansal, N. Richardson, I. Neeli, S. Khawaja, D. Chamberlain, M. Ghani, Q. ul ain Ghani, L. Balazs, S. Beranova-Giorgianni, F. Giorgianni, J. N. Kochenderfer, T. Marion, L. M. Albritton, M. Radic, Sustained B cell depletion by CD19-targeted CAR T cells is a highly effective treatment for murine lupus, *Sci. Transl. Med.* **11** (2019), doi:10.1126/scitranslmed.aav1648.

64. D. Wu, R. O. Emerson, A. Sherwood, M. L. Loh, A. Angiolillo, B. Howie, J. Vogt, M. Rieder, I. Kirsch, C. Carlson, D. Williamson, B. L. Wood, H. Robins, Detection of minimal residual disease in B lymphoblastic leukemia by high-throughput sequencing of IGH, *Clin. Cancer Res.* **20**, 4540–4548 (2014).

65. B. A. C. Cree, J. L. Bennett, H. J. Kim, B. G. Weinshenker, S. J. Pittock, D. M. Wingerchuk, K. Fujihara, F. Paul, G. R. Cutter, R. Marignier, A. J. Green, O. Aktas, H.-P. Hartung, F. D. Lublin, J. Drappa, G. Barron, S. Madani, J. N. Ratchford, D. She, D. Cimbora, E. Katz, Inebilizumab for the treatment of neuromyelitis optica spectrum disorder (N-MOmentum): a double-blind, randomised placebo-controlled phase 2/3 trial, *Lancet* **6736**, 1–12 (2019).

66. A. L. Greenfield, R. Dandekar, A. Ramesh, E. L. Eggers, H. Wu, S. Laurent, W. Harkin, N. S. Pierson, M. S. Weber, R. G. Henry, A. Bischof, B. A. C. Cree, S. L. Hauser, M. R. Wilson, H.-C. von Büdingen, Longitudinally persistent cerebrospinal fluid B-cells can resist treatment in multiple sclerosis, *JCI Insight* , 1–12 (2019).

67. N. Nouri, S. H. Kleinstein, A spectral clustering-based method for identifying clones from high-throughput B cell repertoire sequencing data, *Bioinformatics* **34**, i341–i349 (2018).

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

69. J. A. Vander Heiden, G. Yaari, M. Uduman, J. N. H. Stern, K. C. O'connor, D. A. Hafler, F. Vigneault, S. H. Kleinstein, PRESTO: A toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires, *Bioinformatics* **30**, 1930–1932 (2014).

70. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species., *Nat. Biotechnol.* **36**, 411–420 (2018).

71. F. Vallania, A. Tam, S. Lofgren, S. Schaffert, T. D. Azad, E. Bongen, W. Haynes, M. Alsup, M. Alonso, M. Davis, E. Engleman, P. Khatri, Leveraging heterogeneity across multiple datasets increases cell-mixture deconvolution accuracy and reduces biological and technical biases, *Nat. Commun.* **9** (2018), doi:10.1038/s41467-018-07242-6.

72. J. Ye, N. Ma, T. L. Madden, J. M. Ostell, IgBLAST: an immunoglobulin variable domain sequence analysis tool., *Nucleic Acids Res.* **41**, 34–40 (2013).

73. G. C. Linderman, J. Zhao, Y. Kluger, Zero-preserving imputation of sc RNA-seq data using

low-rank approximation, *Bioinformatics* (2018).

74. J. D. Storey, R. Tibshirani, Statistical significance for genomewide studies John, **2003**, 6 (2003).

75. M. V Kuleshov, M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, M. G. McDermott, C. D. Monteiro, G. W. Gundersen, A. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update., *Nucleic Acids Res.* **44**, W90-7 (2016).

76. R. C. Team, R: A Language and Environment for Statistical Computing (2013).

77. M. O. Hill, Diversity and Evenness : A Unifying Notation and Its Consequences, *Ecol. Soc. Am.* **54**, 427–432 (2015).

78. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, S. 1. 0 Contributors, SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python, , 1–22 (2019).

79. J. P. Felsenstein, PHYLIP - Phylogeny Inference Package (Version 3.2), *Cladistics* **5**, 164–166 (2004).

80. D. Sankoff, Minimal Mutation Trees, *SIAM J. Appl. Math.* **28**, 35–42 (2019).

81. K. B. Hoehn, J. A. Vander Heiden, J. Q. Zhou, G. Lunter, O. G. Pybus, S. Kleinstein, Repertoire-wide phylogenetic models of B cell molecular evolution reveal evolutionary signatures of aging and vaccination, *PNAS* , 1–9 (2019).

**Acknowledgments:** We thank Jason Vander Heiden for assistance with developing methods for processing V(D)J sequences from 10X Genomics sequencing. We thank Karen Boss and Abeer Obaid for helpful feedback for the manuscript. We thank Eileen Dimalanta and Chen Song for providing reagents used for the preparation of bulk BCR libraries. **Funding:** K.C.O is supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health through grant award numbers AI114780 and AI142198, the National Institute of Neurological Diseases and Stroke under award number NS115054, and a Neuromuscular Disease Research program award from the Muscular Dystrophy Association (MDA) under award number MDA575198. S.H.K. is supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health through grant award numbers AI104739. **Author contributions**: R.J, P.S., S.H.K, and K.C.O. designed the study. R.J. and P.S. performed all the experiments. R.J. and K.B.H. analyzed the data. R.J, M.F., K.B.H, R.J.N, P.S., S.H.K, and K.C.O. wrote the manuscript. R.J.N, S.H.K, and K.C.O. provided resources, reagents, and funding. **Competing interests**: K.C.O. has received research support from Ra Pharmaceuticals and compensation from Momenta Pharmaceuticals for consulting services. S.H.K. receives consulting fees from Northrop Grumman. R.J.N. has received

research support from Alexion Pharmaceuticals, Genentech, Grifols, Ra Pharmaceuticals, and has served as a paid consultant for Alexion Pharmaceuticals, Momenta, Ra Pharmaceuticals, Roivant, Shire and Grifols. All other authors declare that they have no competing interests. **Data and materials availability**: Bulk and single-cell repertoire V(D)J data paired with gene expression will be submitted to NCBI upon acceptance. Code will be made available in a publicly accessible repository (https://bitbucket.org/kleinstein/projects/src/default/).

**Figures**:

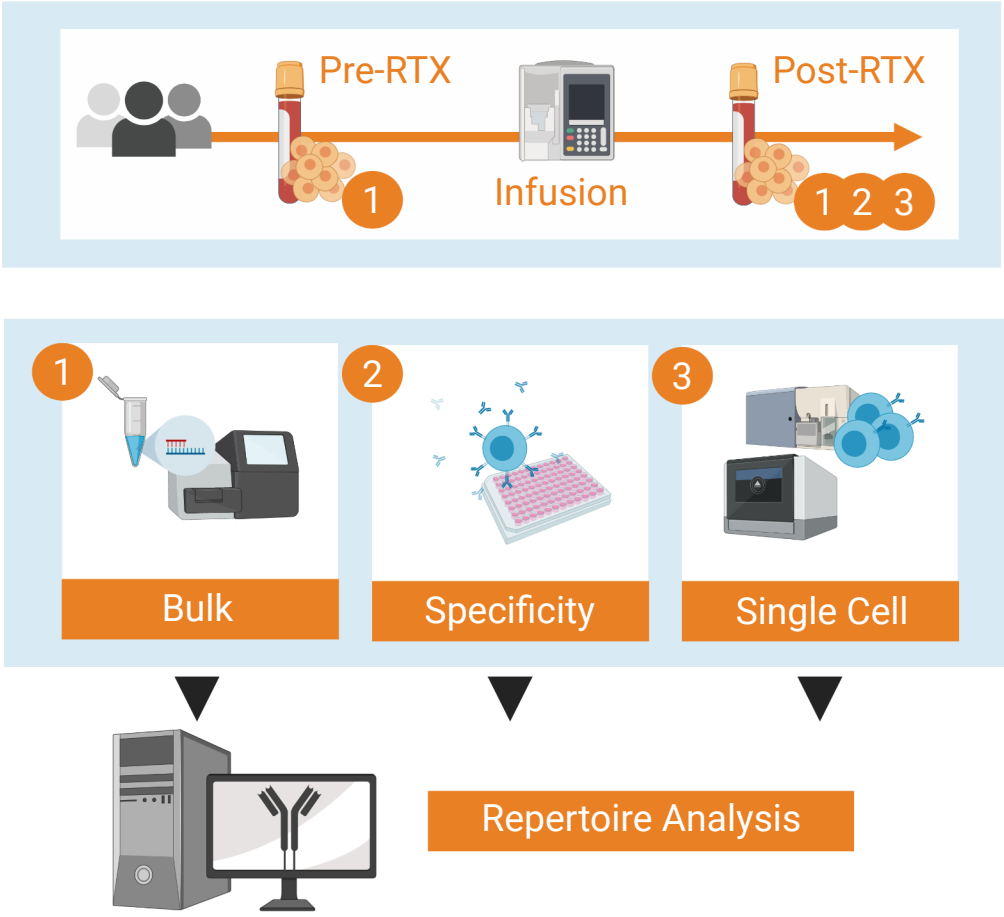


**Fig. 1**. Schematic showing overall workflow from clinical data elements and sample collection to computational analysis of repertoires. In step one, “Bulk” repertoires using next-generation sequencing by isolating unpaired IGH V(D)J sequences directly from RNA were generated. In step two, paired heavy and light chain V(D)J sequences from circulating B cells were expressed as recombinant antibodies and tested for “Specificity” for antigen. In step three, paired heavy and light chain V(D)J sequence repertoires from high-throughput emulsion-based “Single Cell” sequencing were generated. These three types of repertoires were analyzed together in our workflow, "Repertoire Analysis", for each patient across the sampled time points. This workflow involved the identification of clones, quantification of diversity, characterization of differential gene expression, and analysis of SHM frequency and isotype distribution.

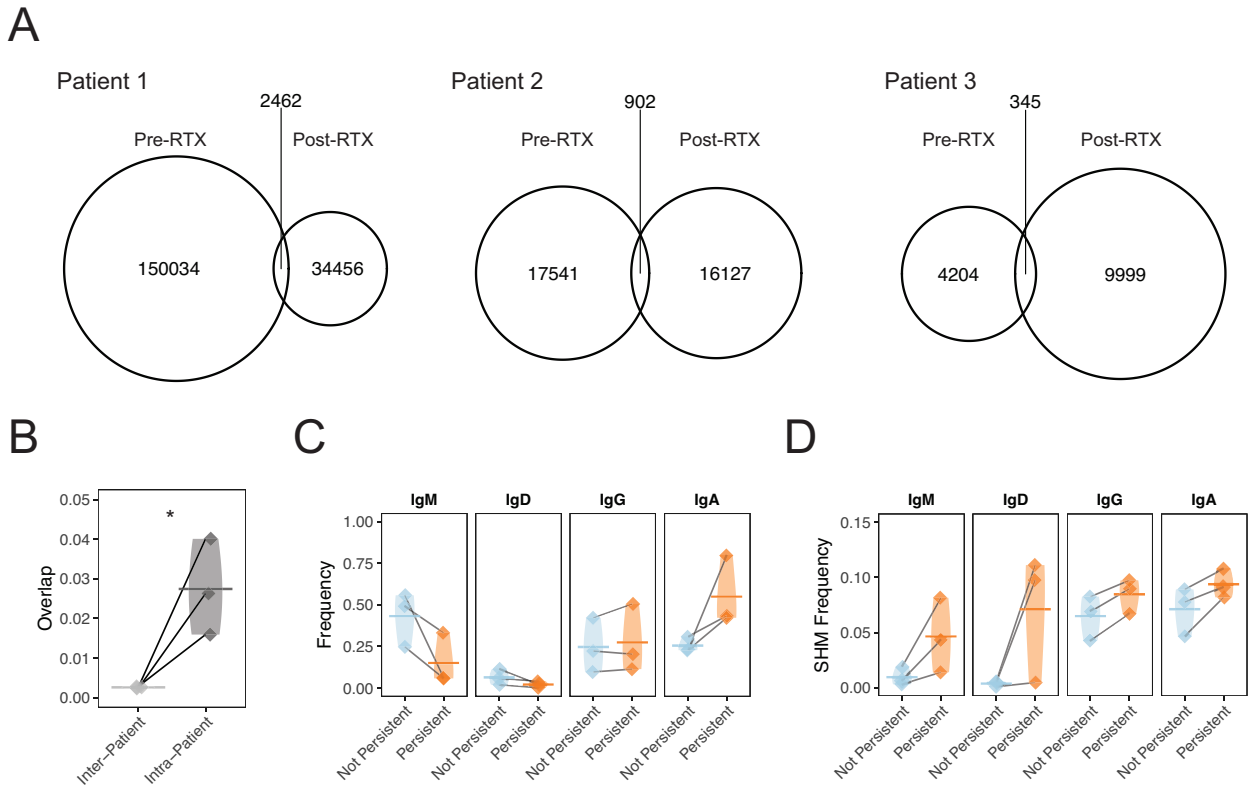


**Fig. 2.** Clones that overlap pre-RTX and post-RTX bulk IGH repertoires (i.e., persistient clones) are associated with switched isoptypes and increased somatic hypermutation frequency. (**A**) Venn diagrams show counts of clones present at pre-RTX and post-RTX time points and those that overlap both time points for all study subjects. (**B**) Overlap between pre-RTX and post-RTX time points quantified as bray-curtis overlap for the same patient (Intra-Patient) or across different patients (Inter-Patient). Horizontal bars show the mean overlap of each comparison. A one-tailed t-test was used to assess the significance of the null hypothesis that Intra-Patient overlap was not higher than Inter-Patient overlap. Overall distribution of frequencies of different isotypes (**C**) and paired average somatic hypermutation frequencies (**D**) among the set of sequences belonging to persistent and non-persistent clones during post-RTX relapse for study subjects. Two-way ANOVA was performed to assess significance for an overall somatic hypermutation difference between non-persistent compared to persistent clones across isotype.

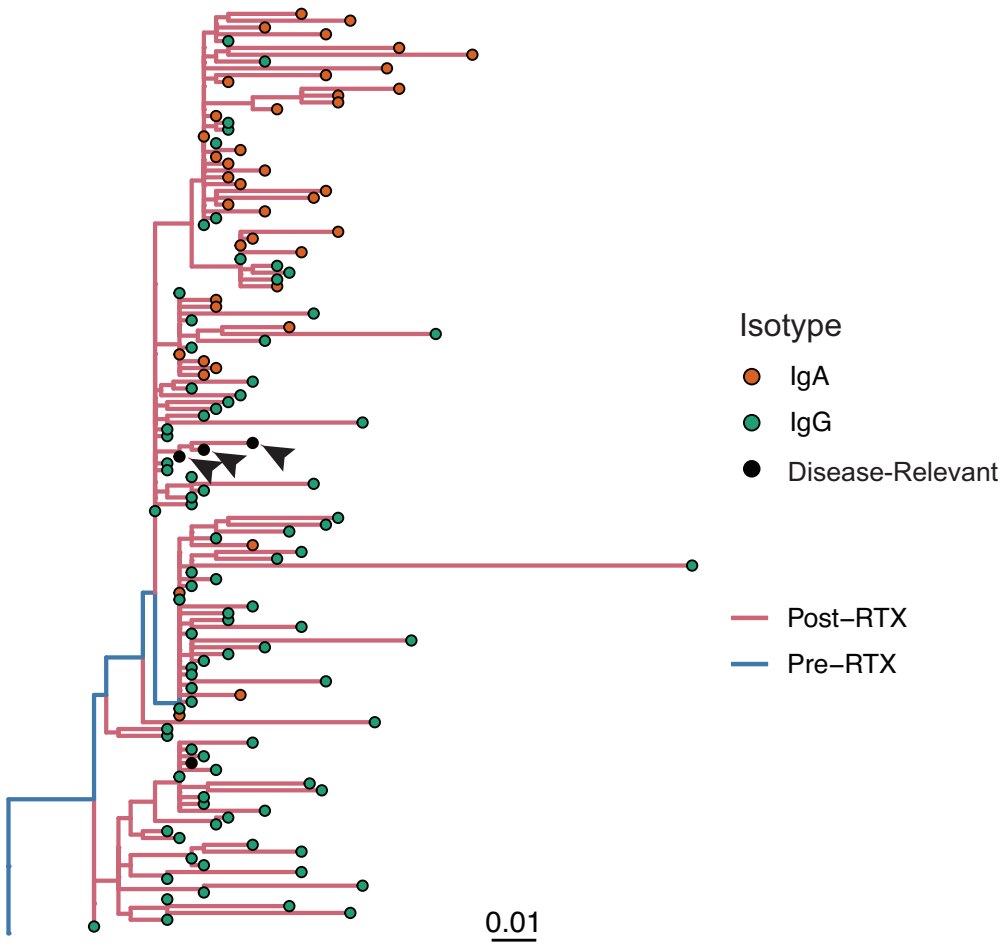


**Fig. 3.** Example lineage tree of a disease-relevant B cell clone spanning both the pre- and post-RTX repertoire. Sequences isolated from plasmablasts using Sanger sequencing-based approaches are denoted as "Disease-Relevant"; some of these bind MuSK weakly including the three shown here (black arrows). Lineage tree topology and branch lengths were estimated using maximum parsimony, with edge lengths representing the expected number of nucleotide substitutions per site (see scale) as estimated using dnapars v3.967 *(79)*. Tips are colored by antibody isotype, and each internal branch is colored by whether its descendant node was predicted to have occurred in the pre- or post-RTX repertoire using a constrained maximum parsimony algorithm (see Methods).

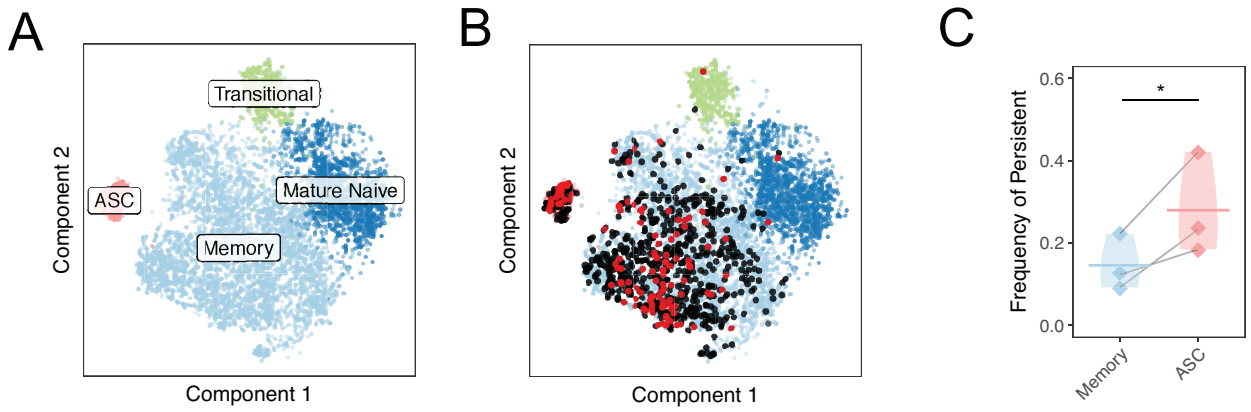


**Fig. 4.** Persistent clones are disproportionately ASC. (**A**) Unbiased clustering of single-cell transcriptomes from all study subjects is visualized by t-SNE and colored based on B cell subset assignment. (**B**) V(D)J sequences associated with persistent B cell clones are visualized over the tSNE plot labelled in (**A**). B cells are colored if they are associated with persistent B cell clones (black) or persistent V(D)J sequences (red, exact same V(D)J sequence a one present in a pre-RTX bulk repertoire). (**C**) The relative fractions of persistent compared to non-persistent B cells for the memory B cell compared to ASC cluster. A one-tailed t-test was used to assess the significance of the null hypothesis that ASCs would not be enriched compared to memory B cells. Horizontal bars show the average frequency of a given B cell subset across patients. Frequencies belonging to the same patient are paired with a gray line.

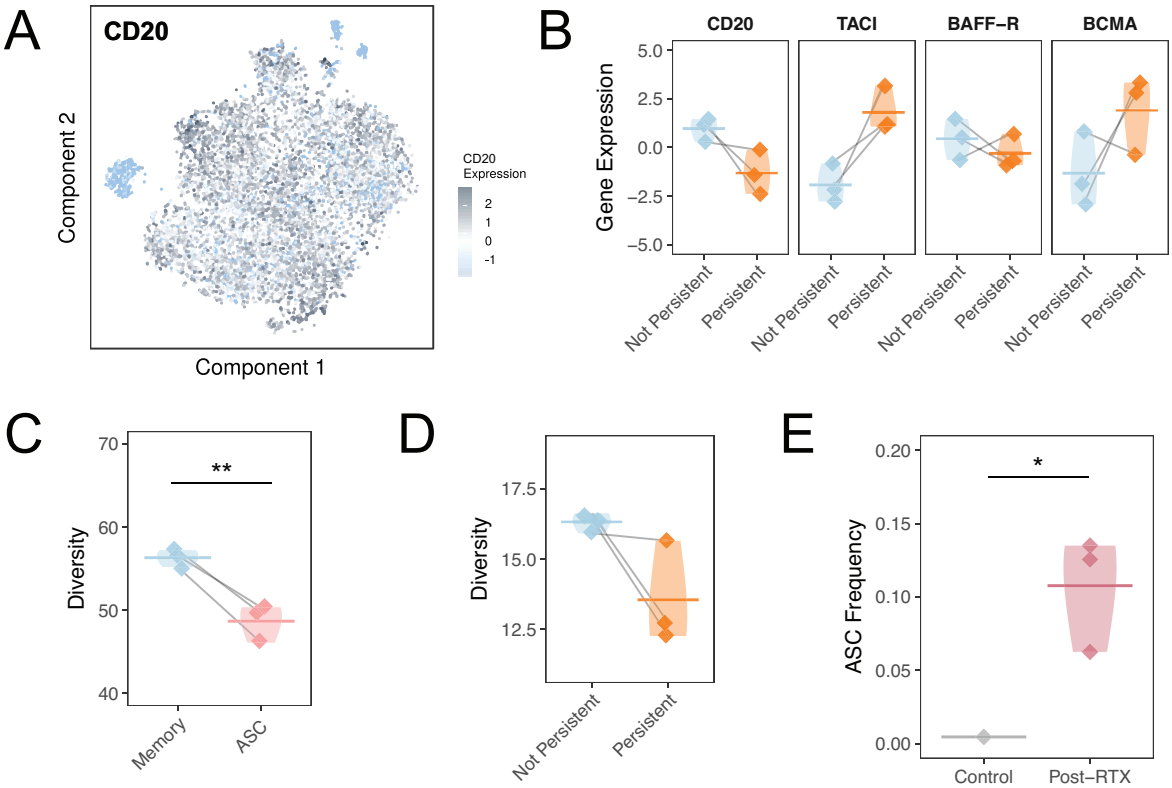


**Fig. 5.** ASCs are expanded but do not have distinct transcriptional and V(D)J repertoire features associated with persistience. (**A**) Expression of log normalized CD20 expression visualized by intensity over t-SNE plot. (**B**) The expression of the receptor target of rituximab CD20 (MS4A1) and BAFF or APRIL receptors (TACI or TNFRSF13B; BAFF-R or TNFRSF13C; BCMA or TNFRSF17) for persistent and non-persistent members of the ASC cluster. (**C**) Simpson's diversity for ASC and memory B cell clusters. (**D**) Simpson's diversity for persistent and non-persistent members of the ASC cluster. A one-tailed t-test was used to assess the significance of the null hypothesis that persistent clones would be less diverse than non-persistent clones. (**E**) The fraction of ASC cluster members is quantified as a ratio of the total fraction of cluster members combined for each patient. A one sample t-test was used to assess significance. Horizontal bars show the average frequency of ASC cluster members for each patient of a given status. Frequencies belonging to the same patient are paired with a gray line.

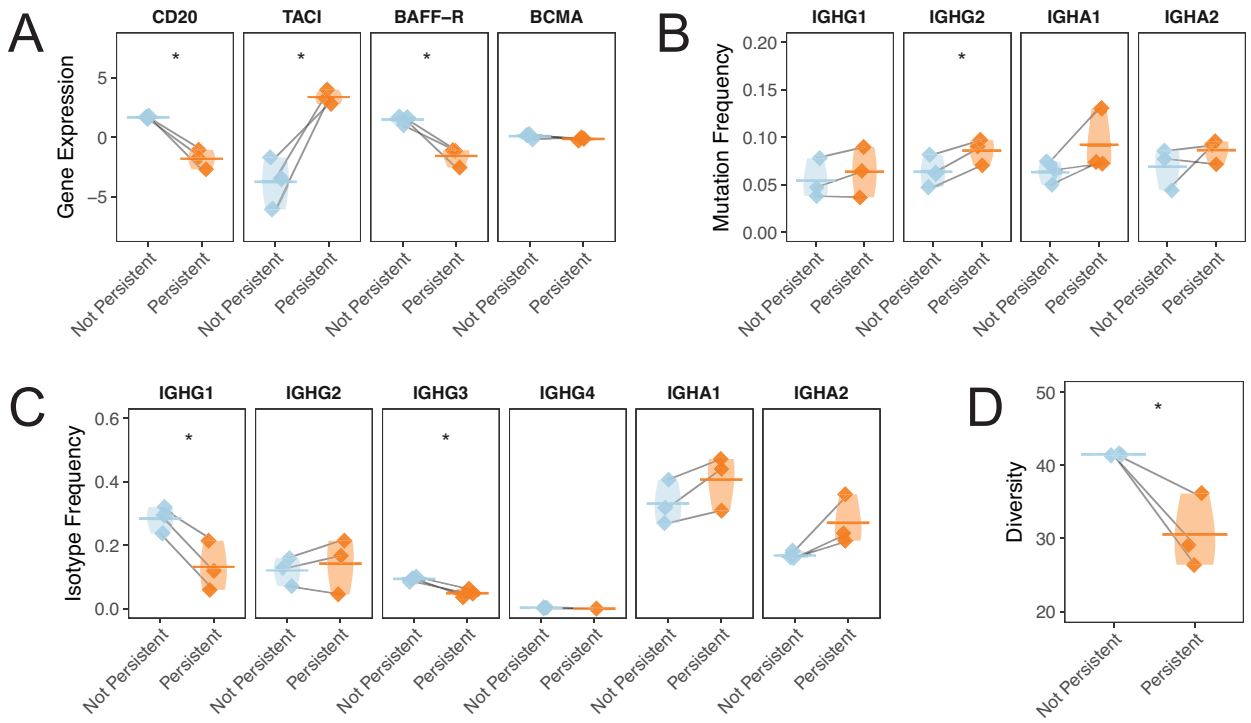


**Fig. 6.** Persistent memory B cells have distinct transcriptional and V(D)J repertoire features associated with persistence. (**A**) The expression of the receptor target of rituximab CD20 (gene symbol, MS4A1) and BAFF or APRIL receptors (TACI or TNFRSF13B; BAFF-R or TNFRSF13C; BCMA or TNFRSF17) is presented for persistent and non-persistent members of the memory B cell cluster. Normalized gene expression values are computed from counts of barcoded gene expression transcripts. (**B**) Overall constant region usage frequencies are quantified for persistent compared to non-persistent memory B cells per patient. Horizontal bars show the average frequency of constant region usage across patients. Frequencies belonging to the same patient are paired with a gray line. Individual SHM frequencies for persistent compared to non-persistent cells are presented for the memory B cell cluster members (**C**). Only mean SHM frequencies are computed for isotypes with more than 3 V(D)J sequences. Horizontal bars show the average somatic hypermutation frequency for a given cluster. Frequencies belonging to the same patient are paired with a gray line. (**D**) Simpson's diversity of persistent compared to non-persistent memory B cell cluster members. A one-tailed t-test was used to assess the significance of the null hypothesis that ersistent clones would be less diverse than non-persistent clones.

**Supplementary Materials:**

| Patient | Age | Sex | Collection | Event Type | Antibody Titer | MGC Score | MGFA Class | MGFA PIS | PI Status Change | Prednisone Dose (mg) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 30-50 | F | MG1 | Pre-RTX | 2560 | | 0 | | Worse (W) | 0 |
| | | | MG12 | Pre-RTX | 2560 | | IIa | | Exacerbation (E) | 0 |
| | | | MG119 | Pre-RTX | | 14 | IIIb | | Worse (W) | 0 |
| | | | MG189 | Post-RTX | 5.7 | 6 | IIb | | | 0 |
| 2 | 30-50 | F | MG37 | Pre-RTX | 2560 | | IIb | | Exacerbation (E) | 0 |
| | | | MG135 | Pre-RTX | 2560 | 0 | 0 | CSR* | Unchanged (U) | 0 |
| | | | MG139 | Post-RTX | 1280 | 7 | IIb | | | 0 |
| 3 | 30-50 | F | MG86 | Pre-RTX | 160 | 3 | IIa | | | 50 |
| | | | MG98 | Pre-RTX | <10 | | IIa | | Improved (I) | 30 |
| | | | MG107 | Pre-RTX | <10 | 6 | IIa | | Improved (I) | 30 |
| | | | MG192 | Post-RTX | 0.89 | 3 | IIa | | | 10 |

**Table S1.** Demographic characteristics of MuSK MG study subjects at each collection time point. MGC Score and MGFA Class were collected for each time point. “MGFA Class” refers to Myasthenia Gravis Foundation of America classification system for MG symptoms. “MGC Score” refers to the Myasthenia Gravis Composite score.

*Complete Stable Remission

| Patient | Collection | Relapse | Reads | Unique IgM | Unique IgG | Unique IgA | Clones |
|---|---|---|---|---|---|---|---|
| 1 | MG1 | Pre-RTX | 321622 | 19552 | 25596 | 26067 | 20084 |
| | MG12 | Pre-RTX | 339025 | 11995 | 40274 | 33127 | 15015 |
| | MG119 | Pre-RTX | 518947 | 29872 | 23544 | 26715 | 31820 |
| | MG189 | Post-RTX | 810822 | 30492 | 12343 | 14239 | 31902 |
| 2 | MG37 | Pre-RTX | 391364 | 4274 | 971 | 2617 | 5290 |
| | MG135 | Pre-RTX | 108953 | 9477 | 1513 | 4154 | 10088 |
| | MG139 | Post-RTX | 460288 | 13823 | 2475 | 7788 | 14630 |
| 3 | MG86 | Pre-RTX | 364381 | 752 | 5785 | 2392 | 2427 |
| | MG98 | Pre-RTX | 508120 | 29 | 669 | 1288 | 239 |
| | MG107 | Pre-RTX | 188934 | 154 | 2321 | 4016 | 860 |
| | MG192 | Post-RTX | 477305 | 7366 | 12952 | 10060 | 7999 |

**Table S2.** Counts of reconstructed V(D)J sequences by isotype and clones from sequencing of pre-RTX and post-RTX bulk IGH repertoires.

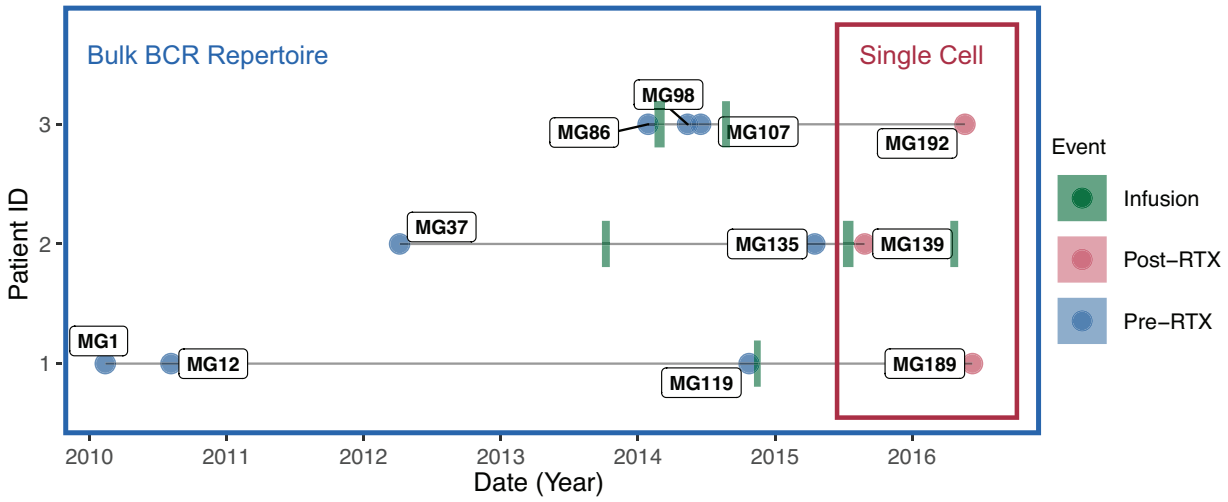


**Fig. S1.** Timeline of three patients who experienced post-RTX relapses. Symptom score based on MGFA class and MGC scores were recorded at all collection time points.

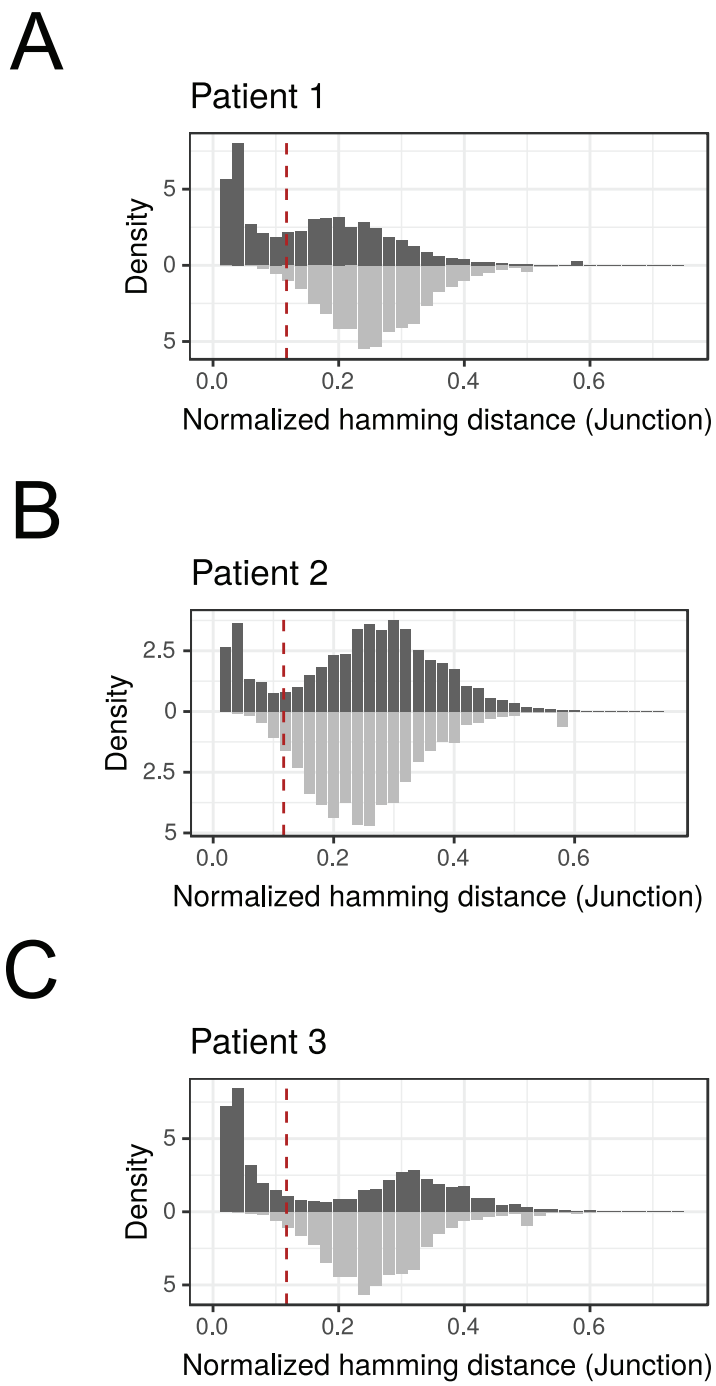


**Fig. S2.** Distance-to-nearest plots used to identify a common threshold to use for hierarchical clustering based grouping of V(D)J sequences from high throughout sequencing of BCR repertoires. Red dashed lines corresponds to the threshold used for assigning clonal clusters.

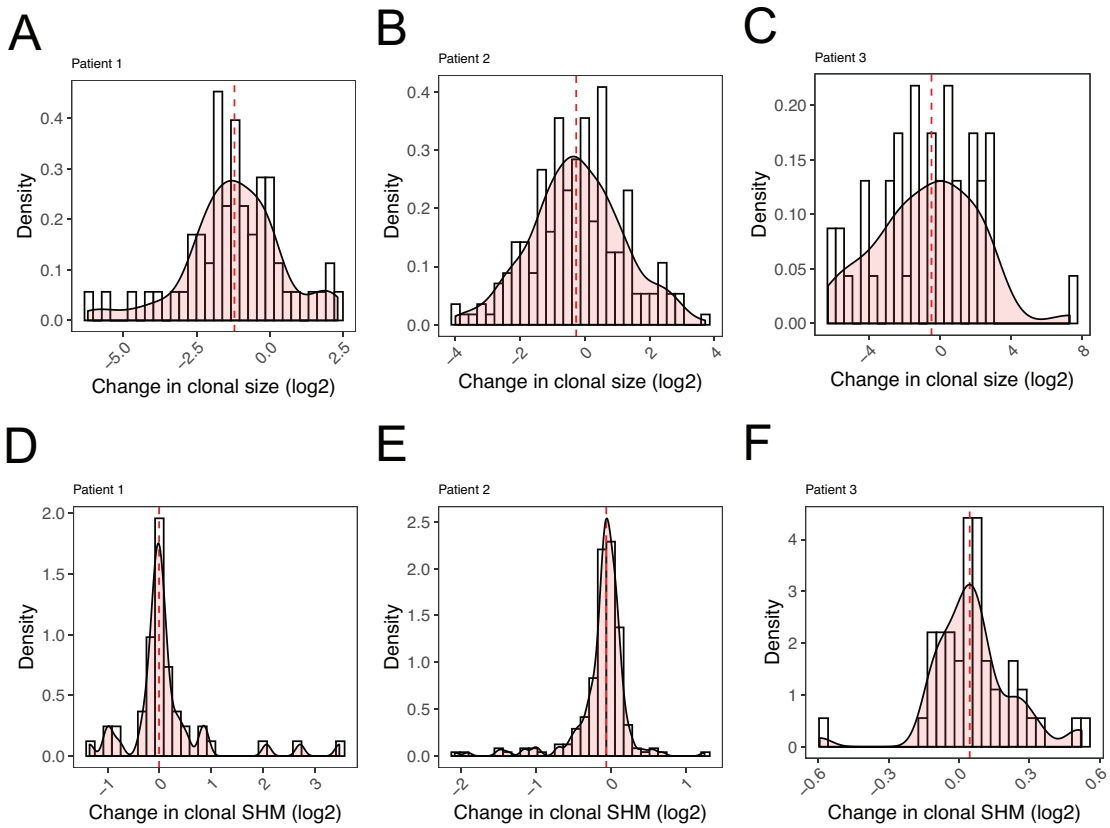


**Fig. S3.** Persistient B cell clones do not show evidence of significant clonal expansion or accumulation of somatic hypermutations when comparing pre-RTX and post-RTX members. The presence of increases in overall somatic hypermutation for individual clones is quantified by comparing the average SHM frequency of each clone in the pre-RTX vs. post-RTX repertoire (A-C). Ratios are visualized as a histogram of the log2 fold changes in somatic hypermutation frequency values for these clones for each patient. Red dashed lines correspond to the median of each distribution. The presence of clonal expansions is evaluated by comparing the overall frequency of individual persistent clones at the post-RTX time point vs. the pre-rituximab time point (D-F). Frequencies are computed as a fraction of total unique V(D)J sequences and visualized in terms of a histogram of the log2 changes in total frequency for these clones. Red dashed lines correspond to the median of each distribution for each patient.

| JCI 2017 | | | |
|---|---|---|---|
| Patient | Clones | Traceable clones | Pre-RTX clones |
| 1 | 4 | 3 | 1 |
| 2 | 0 | 0 | 0 |
| 3 | 39 | 19 | 5 |
| JCI 2019 | | | |
| Patient | Clones | Traceable clones | Pre-RTX clones |
| 1 | 22 | 3 | 1 |
| 2 | 4 | 0 | 0 |
| 3 | 12 | 2 | 0 |

**Table S3.** Count of plasmablast derived or tetramer-binding V(D)J sequences clonally related to members of the pre-RTX *or* post-RTX bulk BCR repertoire (Traceable clones) or to members of the pre-RTX *and* post-RTX bulk BCR repertoire (Pre-RTX clones).

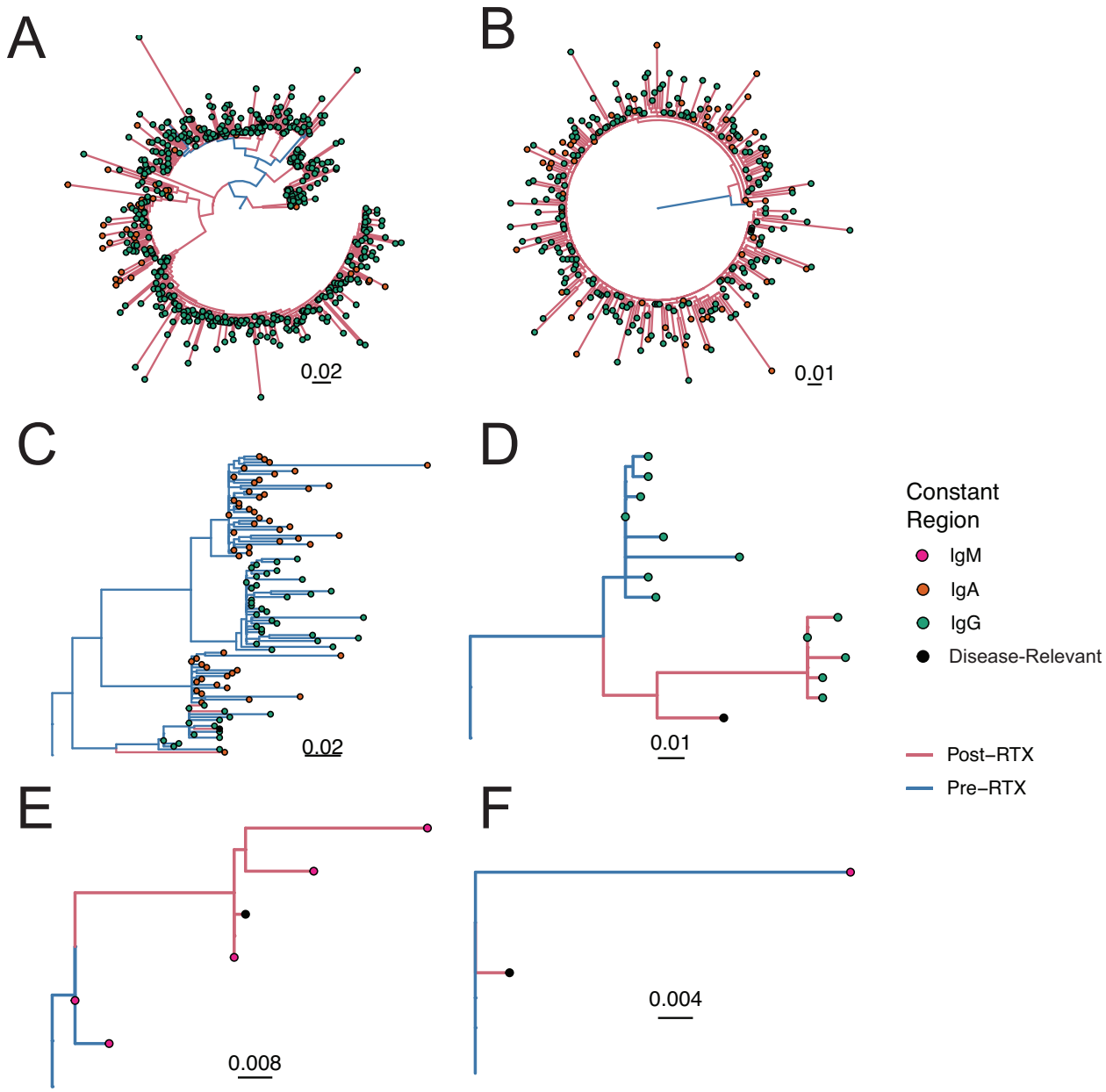


**Fig. S4.** Clonal lineages containing disease-relevant antibodies from Sanger sequencing identified in post-RTX repertoires. Maximum parsimony trees corresponding to clones that contain disease-relevant V(D)J sequences, and that have clonal variants present pre-RTX and post-RTX are presented (A-F). Sequences isolated using Sanger sequencing-based approaches from plasmablasts or by tetramer are denoted as "Disease-Relevant". Edge lengths are quantified based on intervening somatic hypermutations between observed V(D)J sequences per the scale. Colors correspond to whether each V(D)J sequence was collected from a pre-RTX or post-RTX time point (or both) and also the associated constant region.

| Patient | Collection | Status | Read Count | Cell Count | Mean Reads per Cell | Median Genes per Cell | VDJ Count |
|---|---|---|---|---|---|---|---|
| Control | MG295 | AChR MG Asymptomatic | 708747323 | 3580 | 197974 | 1499 | 3810 |
| 1 | MG189 | MuSK MG Post-RTX | 303494244 | 3550 | 85491 | 1189 | 2859 |
| 2 | MG139 | MuSK MG Post-RTX | 130365803 | 8502 | 15333 | 1130 | 6827 |
| 3 | MG192 | MuSK MG Post-RTX | 112260566 | 2501 | 44886 | 1790 | 1175 |

**Table S4.** Quality control from paired single-cell transcriptome and repertoire sequencing.

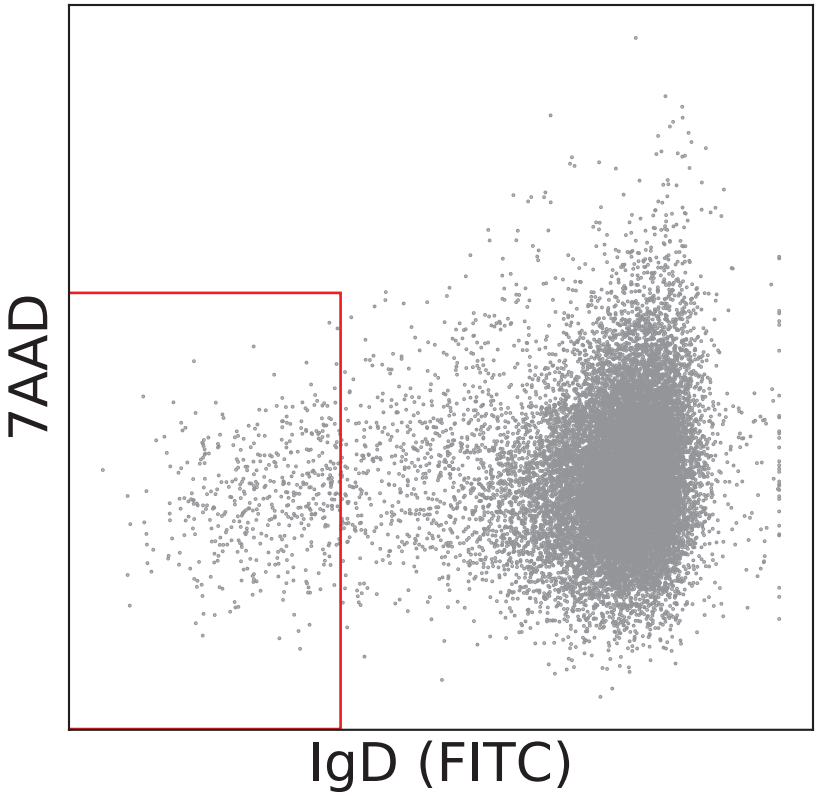


**Fig. S5.** Example fluorescence activated cell sorting (FACS) gates for IgD$^{low}$ B cells used as input for single-cell transcriptomics and BCR repertoire analysis. Example is shown for collection MG189.

| Cluster | Immunostates Assignment | Pearson Correlation | Final Assignment |
|---|---|---|---|
| 0 | naive_B_cell | 0.234 | Mature Naive |
| 1 | memory_B_cell | 0.125 | Memory |
| 2 | memory_B_cell | 0.122 | Memory |
| 3 | memory_B_cell | 0.166 | Memory |
| 4 | memory_B_cell | 0.251 | Memory |
| 5 | memory_B_cell | 0.351 | Memory |
| 6 | naive_B_cell | 0.136 | Transitional |
| 7 | not_assigned | -0.230 | Not Assigned |
| 8 | plasma_cell | 0.644 | ASC |
| 9 | memory_B_cell | 0.212 | Memory |
| 10 | not_b_cell | 0.063 | Not B cell |
| 11 | not_assigned | -0.135 | Not Assigned |

**Table S5.** Table of cluster assignments using the immunoStates basis set. Clusters were assigned to the B cell immunostate with the maximum pearson correlation coefficient when compared with the mean expression value of genes associated with each cluster. Unassigned clusters were manually annotated. One naive B cell cluster was assigned to a transitional B cell subset based on marker expression (cluster 6) and another cluster (cluster 10) was excluded owing to elevated expression of mitochondrial genes.

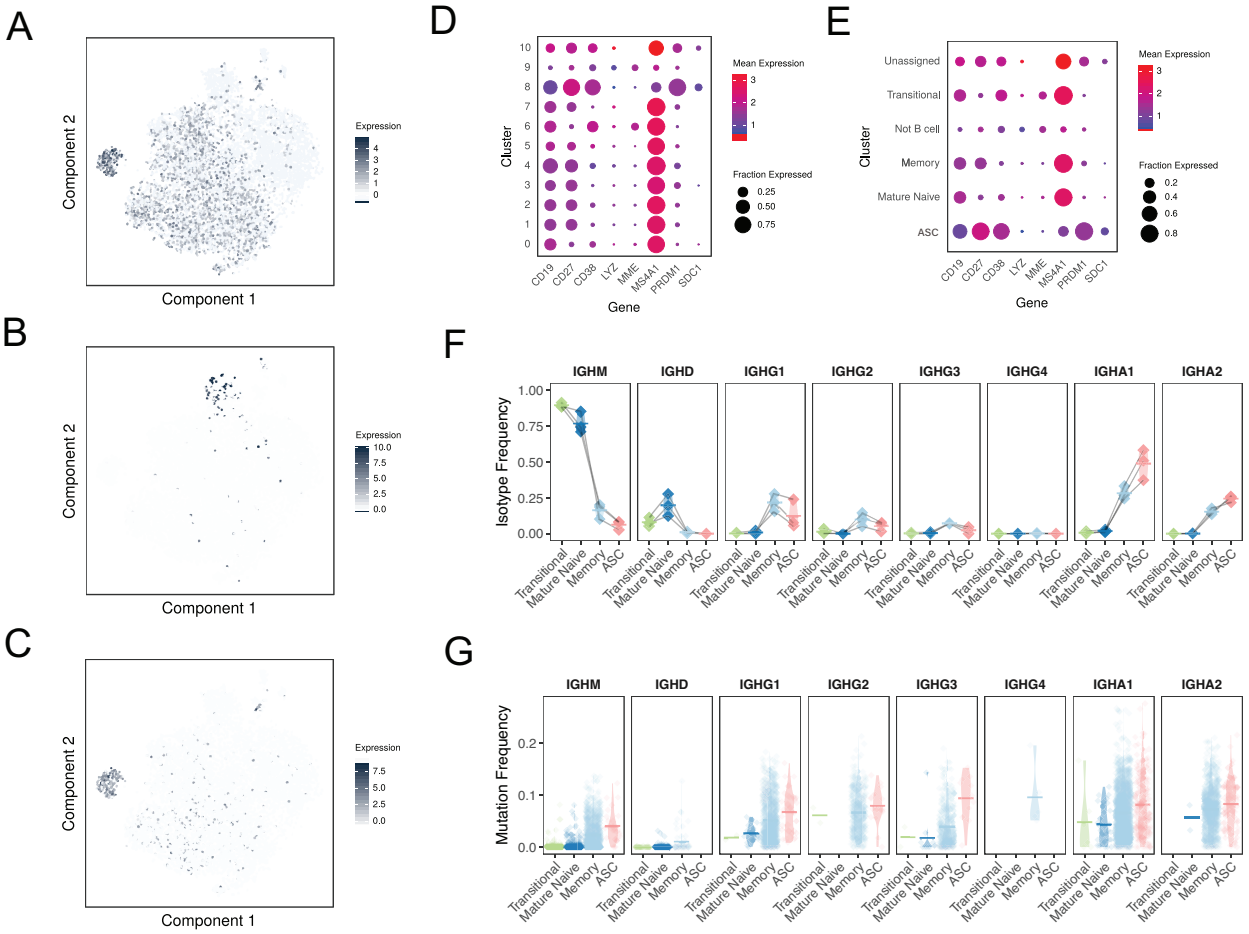


**Fig. S6.** Assignment of B cell clusters based on single-cell transcriptome and repertoire according to known B cell subsets. Plotting of gene expression of key marker genes over t-SNE plot with intensity of shading correlated with amount of scaled log normalized expression for (A) CD27 (B) CD10 (MME) and (C) BLIMP1 (PRDM1). Dot plot of average log normalized expression (color) and fraction of cells expressing a set of marker genes (size) presented for B cell subsets derived from unbiased clustering (D) and when grouped into known B cell subset clusters (E). (F) Overall usage of different constant regions by B cell subset as a fraction of all V(D)J sequences associated with each B cell subset cluster per sample. Horizontal bars show the average frequency of constant region usage across patients. Frequencies belonging to the same patient are paired with a gray line. (G) Distribution of somatic hypermutation frequencies among all V(D)J sequences assigned to each B cell subset cluster. Horizontal bars show the average somatic hypermutation frequency for a given B cell subset cluster.

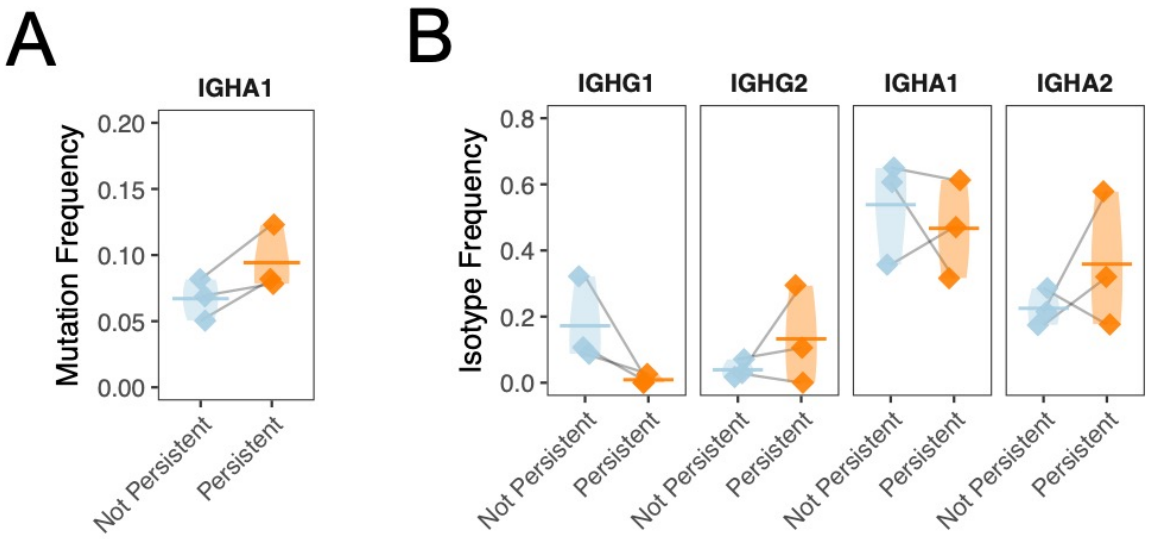


**Fig. S7.** Persistent and non-persistent ASCs are associated with similar V(D)J repertoire. Overall constant region usage frequencies are quantified for persistent compared to non-persistent cells for the (A) ASC cluster per patient. Horizontal bars show the average frequency of constant region usage across patients. Frequencies belonging to the same patient are paired with a gray line. Individual SHM frequencies for persistent compared to non-persistent cells are presented for ASC cluster members (B). Only mean SHM frequencies are computed for isotypes with more than 3 V(D)J sequences. Horizontal bars show the average somatic hypermutation frequency for a given cluster. Frequencies belonging to the same patient are paired with a gray line.

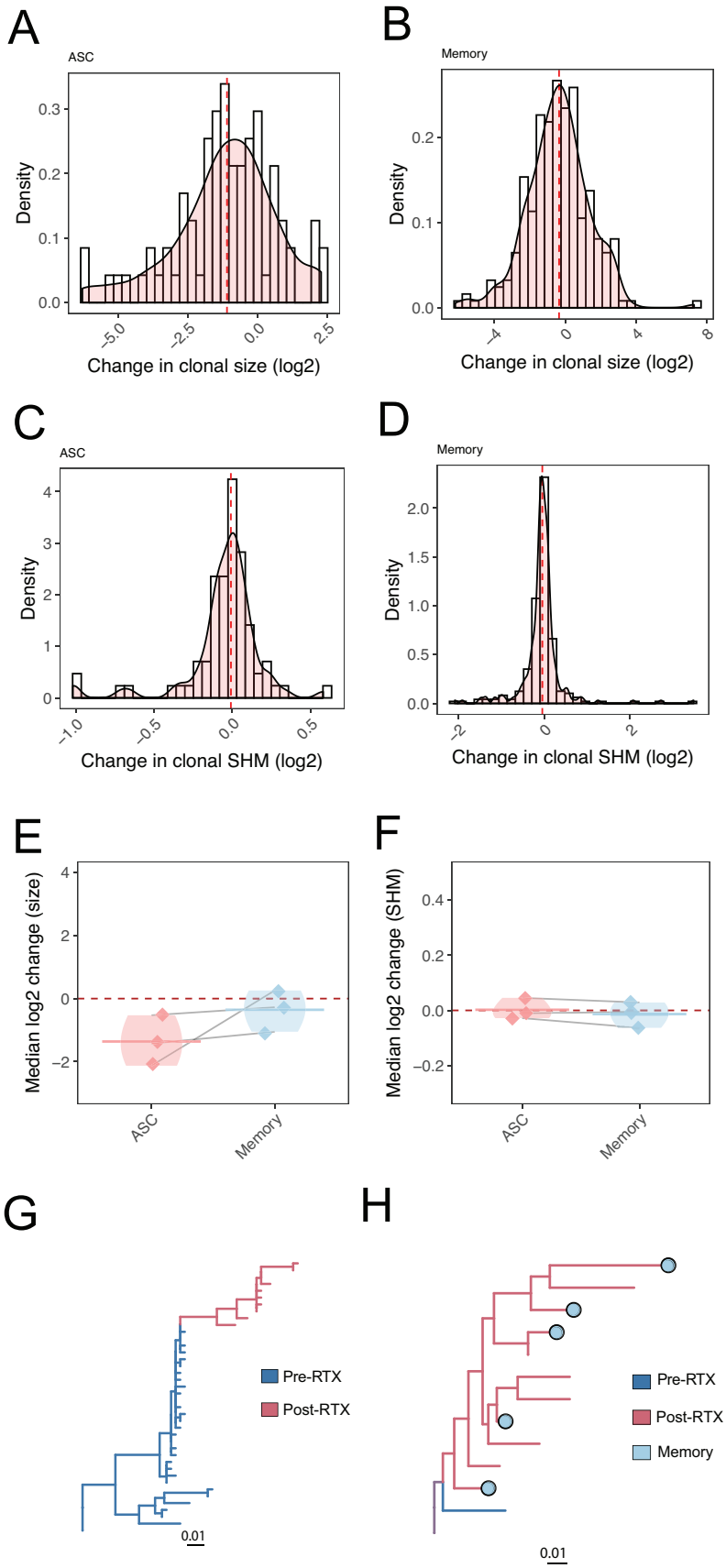
A
ASC
Density
Change in clonal size (log2)
B
Memory
Density
Change in clonal size (log2)
C
ASC
Density
Change in clonal SHM (log2)
D
Memory
Density
Change in clonal SHM (log2)
E
Median log2 change (size)
ASC
Memory
F
Median log2 change (SHM)
ASC
Memory
G
Pre-RTX
Post-RTX
0.01
H
Pre-RTX
Post-RTX
Memory
0.01

**Fig S8.** Persistent ASC and memory B cell clones do not show evidence of significant clonal expansion or accumulation of somatic hypermutations when comparing pre-RTX and post-RTX members. The presence of increases in overall SHM for individual clones is quantified by comparing the SHM frequencies averaged for each clone in the pre-RTX repertoire compared with post-RTX and visualized in terms of a histogram of the natural log change in somatic hypermutation frequency for each clone for each patient. Red dashed lines correspond to the median of each distribution for each patient. The set of clones examined is filtered for clones associated with ASCs (A) or memory B cells (B). The presence of clonal expansions is evaluated by comparing the overall frequency of individual persistent clones at the post-RTX time point compared to the pre-RTX time point from bulk IGH repertoires as a fraction of total unique V(D)J sequences and visualized in terms of a histogram of the natural log change in total frequency for these clones. Red dashed lines correspond to the median of each distribution for each patient. The set of clones examined is filtered for clones associated with ASCs (C) or memory B cells (D). The median of either the change (E) in clonal somatic hypermutation frequencies or clonal size distribution (F) associated with each patient is presented for each filtering strategy ("Total" for no filtering, "ASC" for filtering of ASC associated clones or "Memory" for filtering of Memory B cell associated clones. Horizontal bars show the median natural log change in clonal size or clonal somatic hypermutation frequency across patients. Values belonging to the same patient are paired with a gray line. Example B cell lineage trees consistent with continued somatic evolution after relapse are also shown; tree (G) was obtained from Patient 1, while tree (H) was obtained from Patient 2, and contains multiple memory B cells. Labels of internal nodes were predicted using the maximum parsimony algorithm described in Methods. The root of tree (H) is ambiguously Pre-RTX or post-RTX.

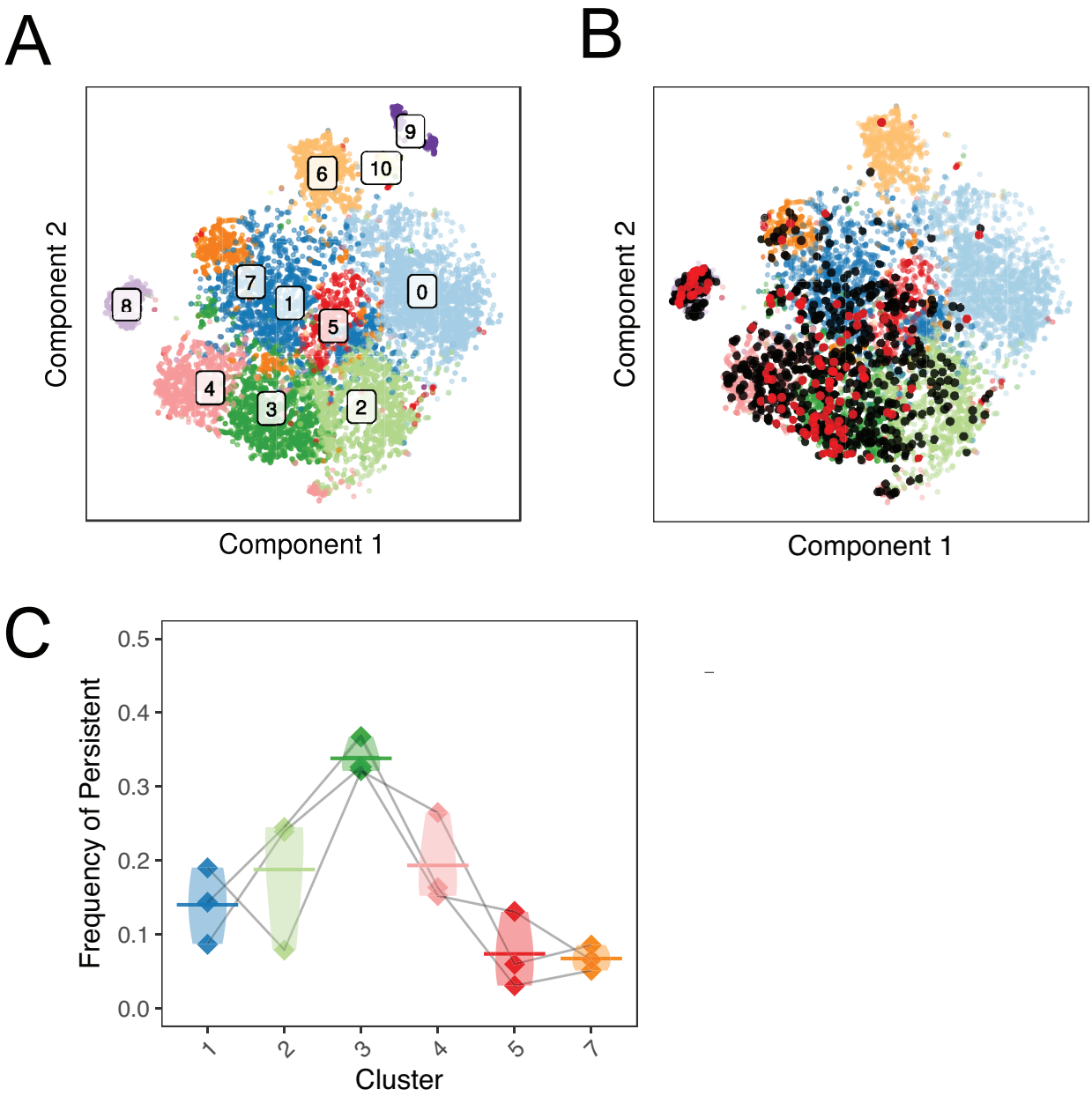


**Fig. S9.** Identification of B cell clusters associated with B cell clonal persistence through unbiased analysis identifies subset associated with resistance to RTX depletion (cluster 3). Visualization of initial shared nearest-neighbor (SNN) clustering of circulating B cells by t-SNE with labels over clusters (A) or paired with B cell clones (black) or exact V(D)J sequence matches (red) (B) that match members of pre-RTX repertoires. (C) The frequency of each cluster among the set of persistent B cell clones is shown.

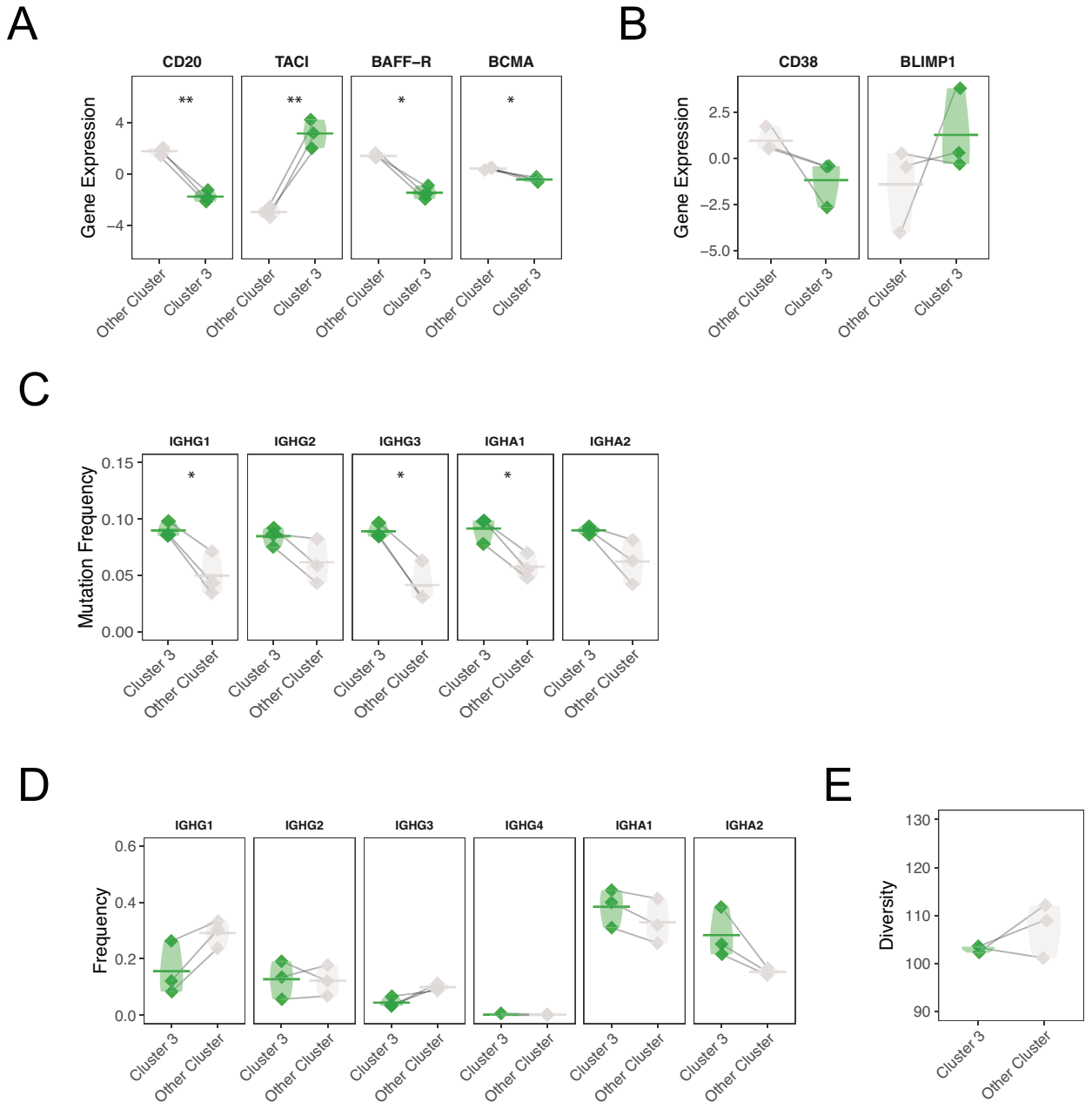


**Fig. S10.** Cluster 3 has distinct transcriptional and V(D)J repertoire features. (A) The expression of the receptor target of rituximab CD20 (MS4A1) is presented for members of cluster 3 and members of other memory B cell clusters. The expression of known receptors for BAFF and APRIL (TACI or TNFRSF13B; BAFF-R or TNFRSF13C; BCMA or TNFRSF17) is also presented. (B) The expression of plasma cell associated genes (CD38; BLIMP1 or PRDM1) is also presented for members of cluster 3 and members of other memory B cell clusters. Normalized gene expression values are computed from scaled log normalized counts of barcoded transcripts aligned to each gene. Horizontal bars show the average normalized gene expression across patients. Expression values belonging to the same patient are paired with a gray line. (C)

Mean somatic hypermutation frequencies for cluster 3 compared to other memory B cells is quantified per patient. Horizontal bars show the average somatic hypermutation frequency for a given set of cells. Frequencies belonging to the same patient are paired with a gray line. (D) Overall constant region usage frequencies are quantified for cluster 3 compared to other memory B cells per patient. Horizontal bars show the average frequency of constant region usage across patients. Frequencies belonging to the same patient are paired with a gray line. (E) Simpson's diversity of cluster 3 members compared to other memory B cell cluster members.

**Data file S1.** List of differentially expressed genes for cluster 3 compared to all other memory B cell clusters. Only differentially expressed genes with an adjusted p-value less than 0.05 with FDR correction by Storey's method are presented.

See attached cluster3.csv

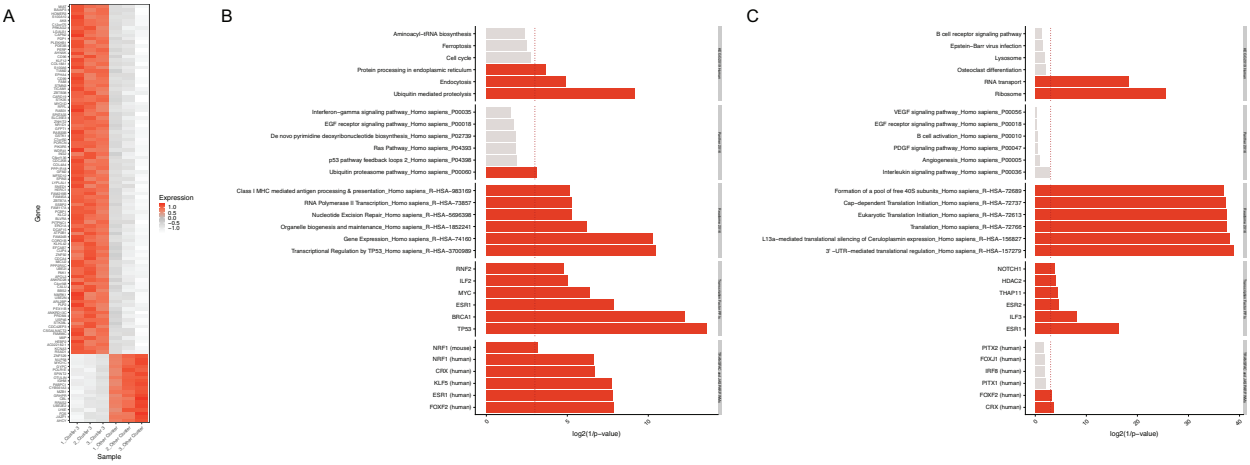


**Fig. S11.** Cluster 3 memory B cells have a MYC and p53/BRCA1 signature. (A) Heatmap of significantly differentially expressed genes for cluster 3 compared to other memory B cells, with a greater than >0.4 z-score difference between persistent and non-persistent labels. enrichR gene ontology analysis of either genes (B) up-regulated or (C) down-regulated among cluster 3 memory B cells compared to other memory B cells. Only differentially expressed genes with an adjusted p-value less than 0.05 with FDR correction by Storey's method were evaluated. Red bars correspond to significantly associated gene ontology assignments ($p < 0.05$ by wilcoxon signed rank test).