

Supplemental Information

Single-cell repertoire tracing identifies rituximab refractory B cells during myasthenia gravis relapses.

Contents:

Table S1. Demographic characteristics of MuSK MG study subjects

Table S2. Counts of reconstructed V(D)J sequences

Figure S1. Distance-to-nearest plots

Figure S2. Persistent B cell clones do not show evidence of significant clonal expansion or accumulation of somatic hypermutations

Table S3. Count of plasmablast derived or tetramer-binding V(D)J sequences

Figure S3. Clonal lineages containing scPCR-derived antibodies from Sanger sequencing

Figure S4. Clonal lineages containing scPCR-derived antibodies from Sanger sequencing

Figure S5. Clonal lineages containing scPCR-derived antibodies from Sanger sequencing

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

Figure S6. Example fluorescence activated cell sorting (FACS) gates

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

Figure S7. Assignment of B cell clusters based on single-cell transcriptome and repertoire

Figure S8. Persistent and non-persistent ASCs are associated with similar V(D)J repertoire

Figure S9. Persistent ASC and memory B cell clones do not show evidence of significant clonal expansion or accumulation of somatic hypermutations

Figure S10. Elevated expression of surface CD20 for B cells with high CD20 mRNA expression.

Figure S11. Cluster 3 has distinct transcriptional and V(D)J repertoire features

Figure S12. Cluster 3 memory B cells have a distinct gene signature

Figure S13. Cluster 3 memory B cells have a MYC and p53/BRCA1 signature

Figure S14. Cluster 3 memory B cells have a MYC and p53/BRCA1 signature

Figure S15. Visual model outlining the likely source of antigen-specific B cells during post-RTX relapse

Patient	Age	Sex	Collection	Event Type	Antibody Titer	MGFA Class
1	37	F	6	Pre-RTX	1:2560	I ¹
			2	Pre-RTX	1:2560	IIa
			7	Pre-RTX	1:5120	IIIb
			10	Post-RTX	5.70 nmol/L ⁴	IIb
2	63	F	3	Pre-RTX	1:2560	IIb
			8	Pre-RTX	1:2560	Asymptomatic ²
			9	Post-RTX	1:5120	IIb ³
3	53	F	1	Pre-RTX	1:160	IIa
			5	Pre-RTX	<1:10	IIa
			12	Pre-RTX	<1:10	IIa
			4	Post-RTX	0.89 nmol/L ⁴	IIa

Table S1. Clinical characteristics of MuSK MG study subjects at each collection time point. The Myasthenia Gravis Foundation of America (MGFA) clinical classification divides MG into five classes (I-V) based on signs and symptoms of disease severity. Reported age is at time of “Post-RTX” relapse.

¹Patient reported mild intermittent fatigable limb weakness prior but none on day of exam/collection, except for mild non-fatigable unilateral ptosis.

²Complete Stable Remission (CSR); patient with clinical relapses 2 months later.

³Patient with clinical relapse/exacerbation

⁴MuSK Antibody Level (reference range: positive > 0.02 nmol/L || Mayo Clinical Laboratory, USA). MuSK Antibody Titer (reference range: positive ≥ 1:20 || Athena Diagnostics, USA) is reported for all others as this was the only or preferred commercial lab available at the time of collection.

Patient	Collection	Relapse	Reads	Unique IgM	Unique IgG	Unique IgA	Clones
1	6	Pre-RTX	321622	19552	25596	26067	20084
	2	Pre-RTX	339025	11995	40274	33127	15015
	7	Pre-RTX	518947	29872	23544	26715	31820
	10	Post-RTX	810822	30492	12343	14239	31902
2	3	Pre-RTX	391364	4274	971	2617	5290
	8	Pre-RTX	108953	9477	1513	4154	10088
	9	Post-RTX	460288	13823	2475	7788	14630
3	1	Pre-RTX	364381	752	5785	2392	2427
	5	Pre-RTX	508120	29	669	1288	239
	12	Pre-RTX	188934	154	2321	4016	860
	4	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 BCR repertoires.

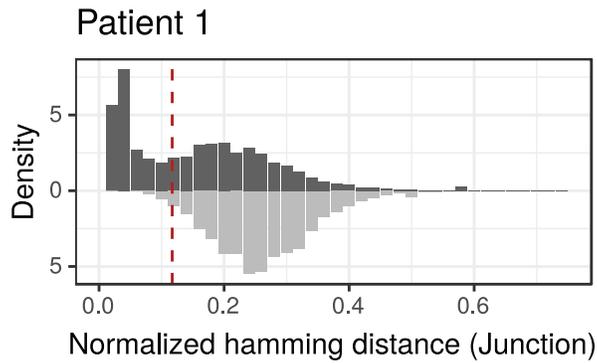
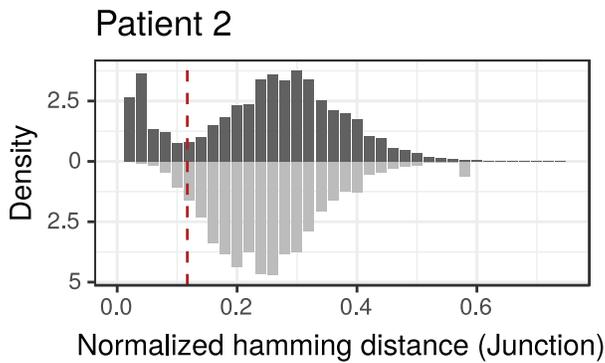
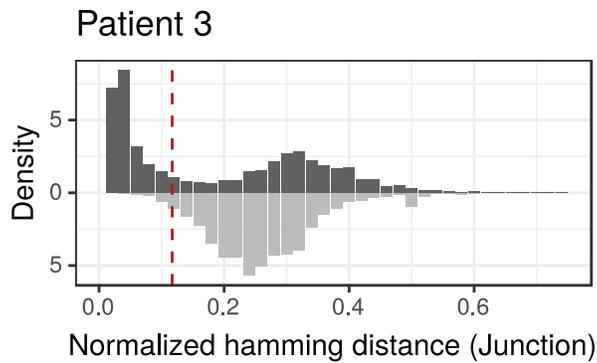
A**B****C**

Figure S1. Distance-to-nearest plots used to identify a common threshold to use for hierarchical clustering based grouping of V(D)J sequences from high throughput sequencing of BCR repertoires. Red dashed lines corresponds to the threshold used for assigning clonal clusters. Dark grey bars represent the distribution of intra-subject distance-to-nearest distances while light grey bars represent the distribution of inter-subject distance-to-nearest distances.

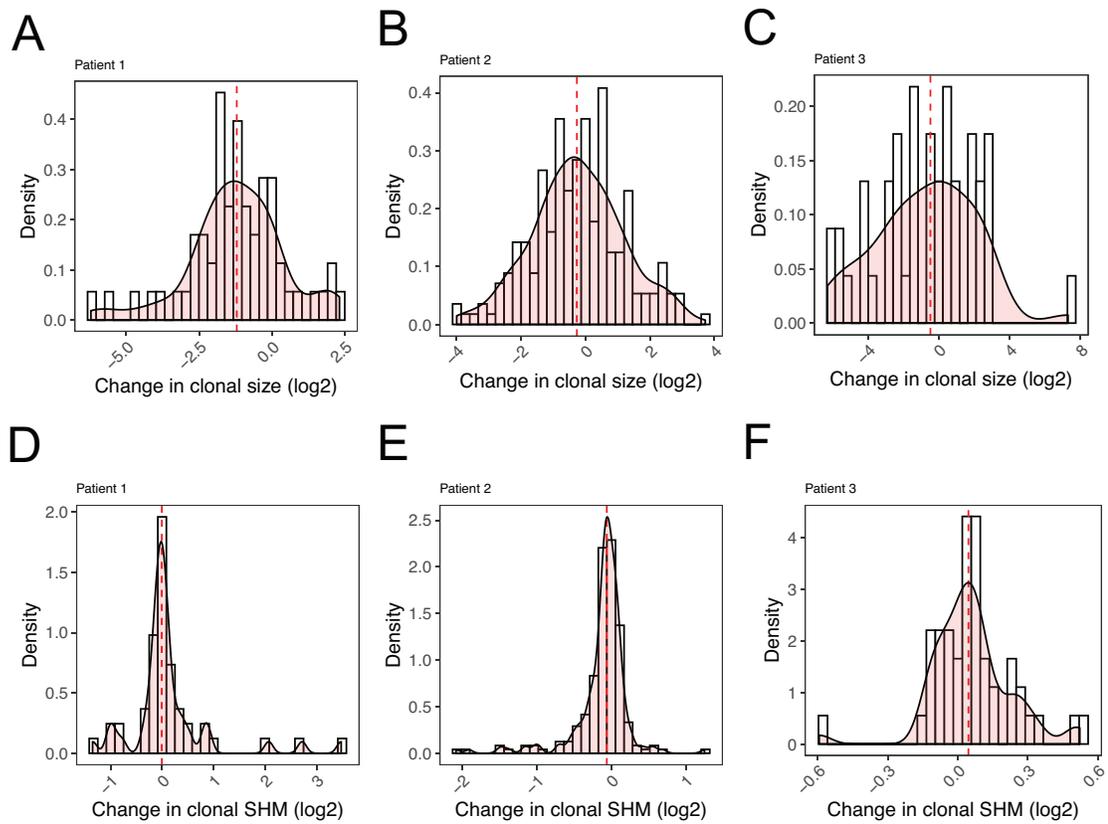


Figure S2. Persistent B cell clones do not show evidence of significant clonal expansion or accumulation of somatic hypermutations when comparing pre-RTX and post-RTX members. The presence of increases in overall somatic hypermutation for individual clones is quantified by comparing the average SHM frequency 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.

Stathopoulos et al. 2017, BCR sequences from scPCR of circulating plasmablasts			
Patient	Clones	Traceable clones	Pre-RTX clones
1	4	3	1
2	0	0	0
3	39	19	5*
Takata et al. 2019, BCR sequences from scPCR of tetramer binding B cells			
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). *One clone was observed to also correspond to a clone previously published to have specificity for MuSK autoantigen. The strongest binding member of this clone was annotated as “3-29” in the previous publication.

Clone ID: 3-3

Clone ID: 3-25

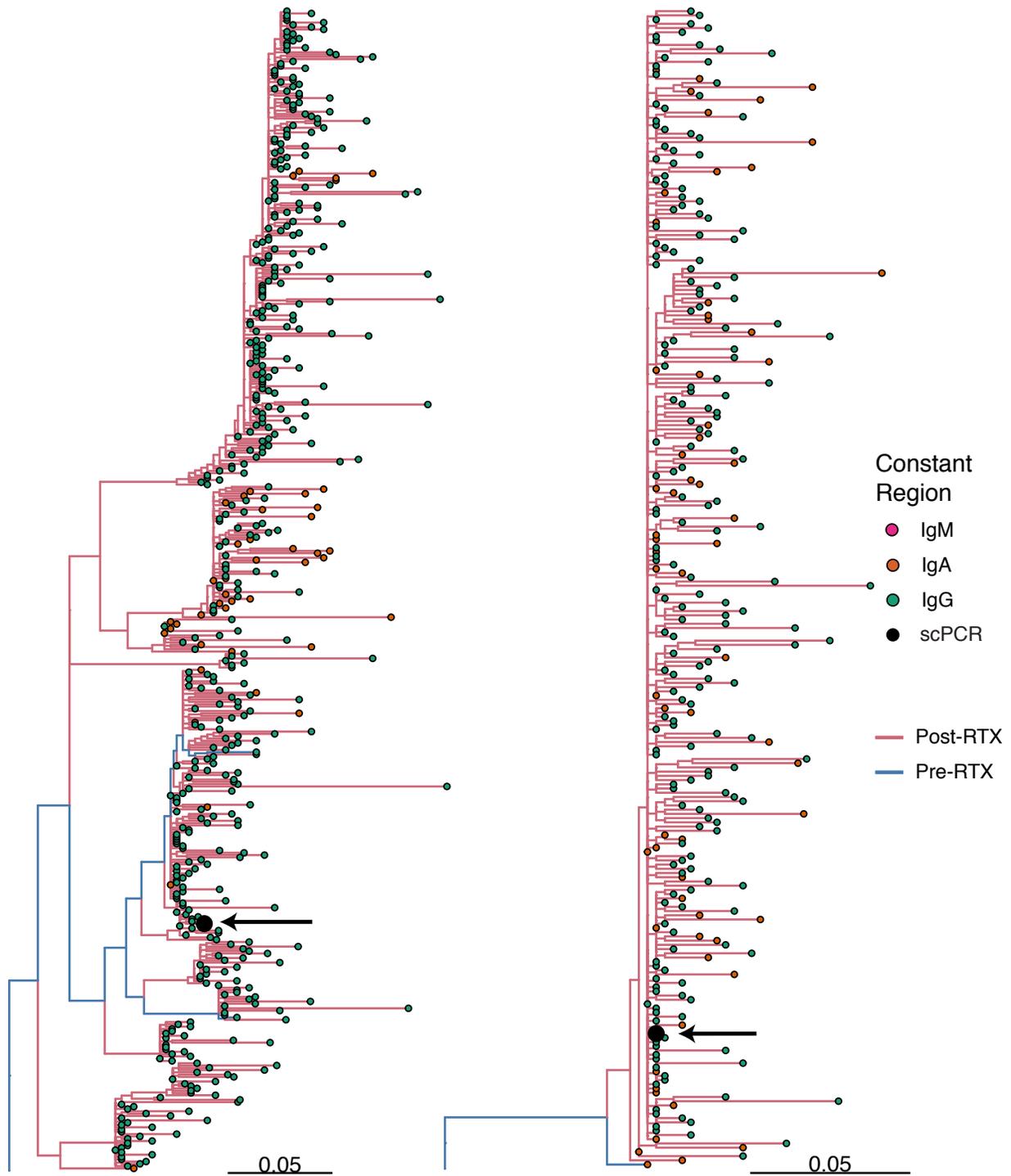


Figure S3. Clonal lineages containing scPCR-derived antibodies from Sanger sequencing identified in post-RTX repertoires. Maximum parsimony trees corresponding to clones that contain scPCR derived V(D)J sequences, and that have clonal variants present pre-RTX and

post-RTX are presented. Sequences isolated using scPCR-based approaches from plasmablasts are denoted as "scPCR" using a large black dot and arrow along with the corresponding identity of the antibody from Stathopoulos et al. 2017 (ref. 15 in main manuscript). Edge lengths are quantified based on number of intervening somatic hypermutations per site 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.

Clone ID: 3-31

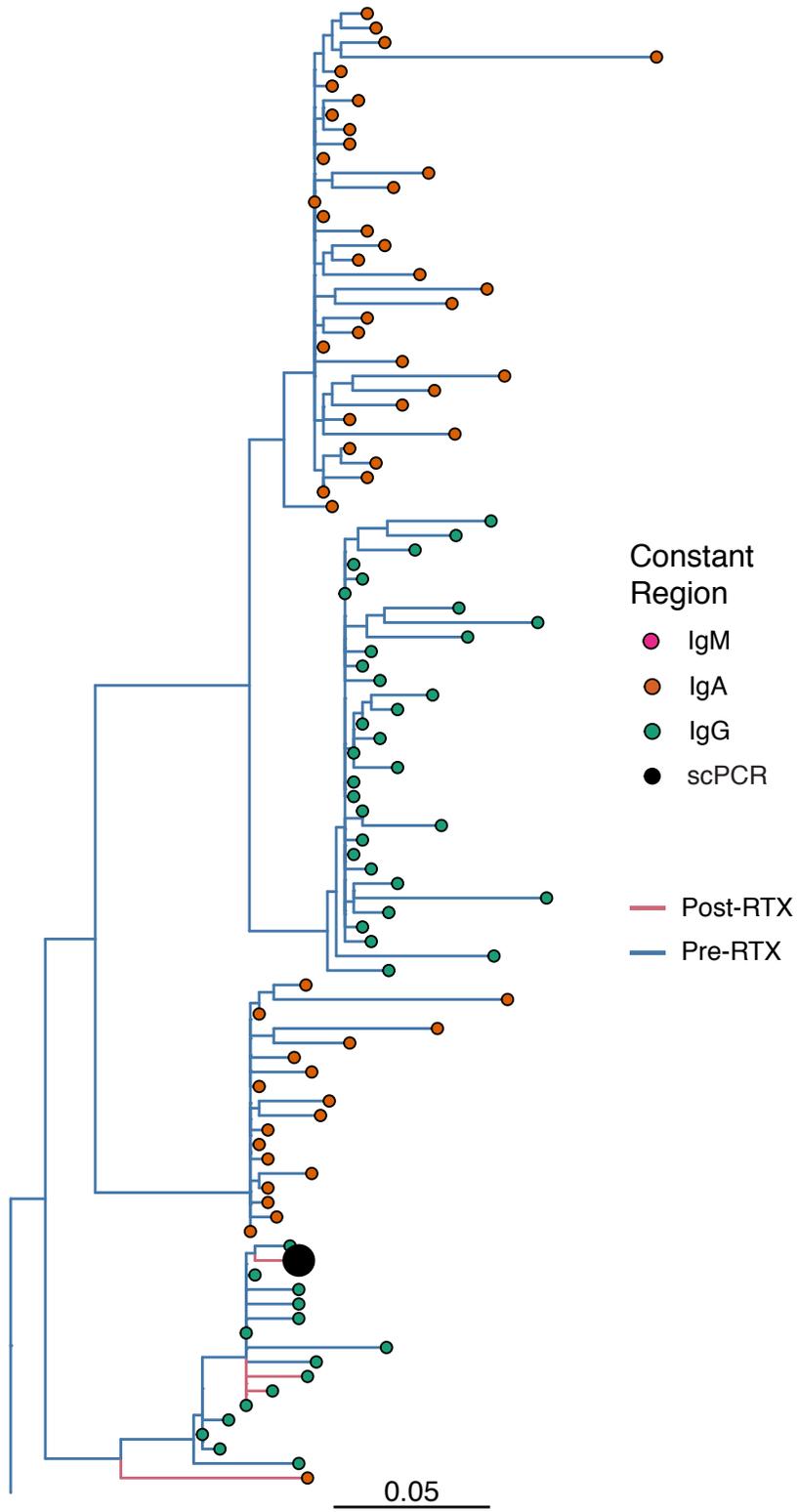


Figure S4. Clonal lineages containing scPCR-derived antibodies from Sanger sequencing identified in post-RTX repertoires. A maximum parsimony tree corresponding to a clone that contains scPCR derived V(D)J sequences, and that have clonal variants present pre-RTX and post-RTX is presented. Sequences isolated using scPCR-based approaches from plasmablasts are denoted as "scPCR" using a large black dot along with the corresponding identity of the antibody from Stathopoulos et al. 2017 (ref. 15 in main manuscript). Edge lengths are quantified based on number of intervening somatic hypermutations per site 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.

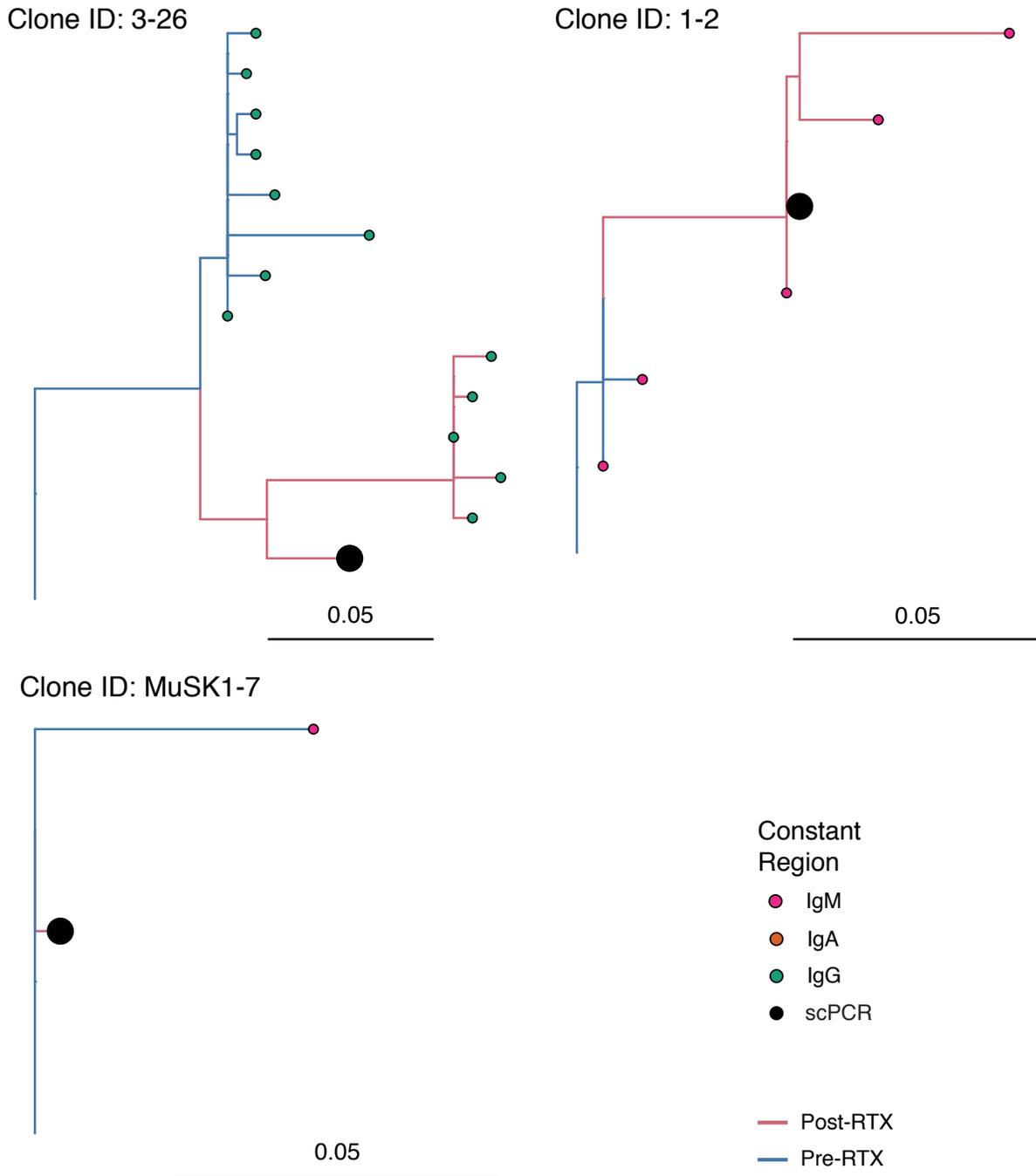


Figure S5. Clonal lineages containing scPCR-derived antibodies from Sanger sequencing identified in post-RTX repertoires. Maximum parsimony trees corresponding to clones that contain scPCR derived V(D)J sequences, and that have clonal variants present pre-RTX and

post-RTX are presented. Sequences isolated using scPCR-based approaches from plasmablasts are denoted as "scPCR" using a large black dot along with the corresponding identity of the antibody from Stathopoulos et al. 2017 (ref. 15 in main manuscript). Of note, the sequence associated with clone "MuSK1-7" was published in Takata et al. 2019 (ref. 16 in main manuscript) and was isolated using an antigen-specific tetramer. Edge lengths are quantified based on number of intervening somatic hypermutations per site 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	A	AChR MG Asymptomatic	708747323	3580	197974	1499	3810
1	10	MuSK MG Post-RTX	303494244	3550	85491	1189	2859
2	9	MuSK MG Post-RTX	130365803	8502	15333	1130	6827
3	4	MuSK MG Post-RTX	112260566	2501	44886	1790	1175

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

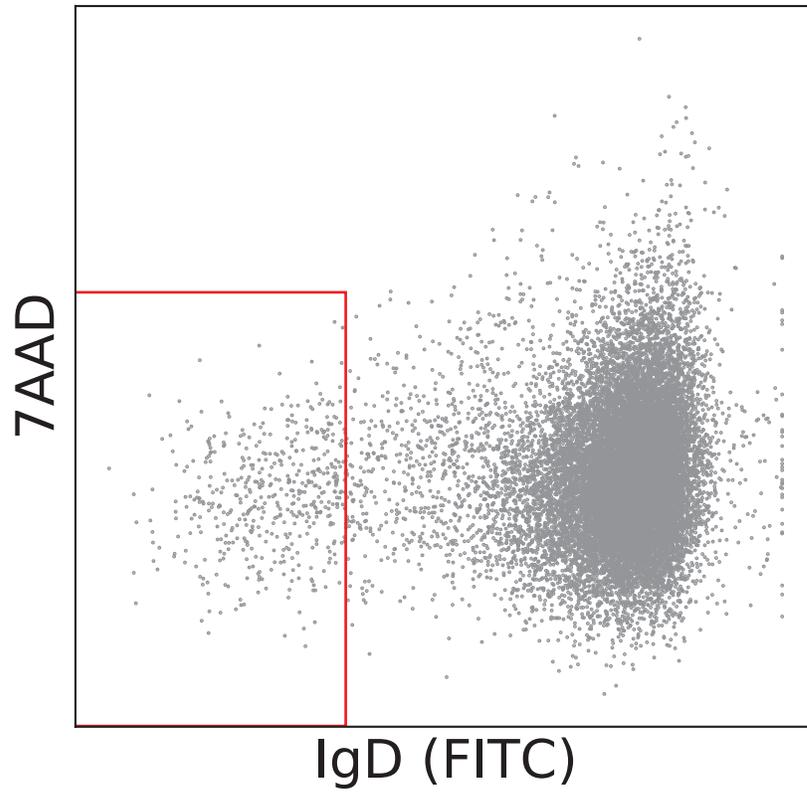


Figure S6. Example fluorescence activated cell sorting (FACS) gates for IgD^{low} B cells used as input for single-cell transcriptomics and BCR repertoire analysis. Example is shown for collection 10.

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. 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.

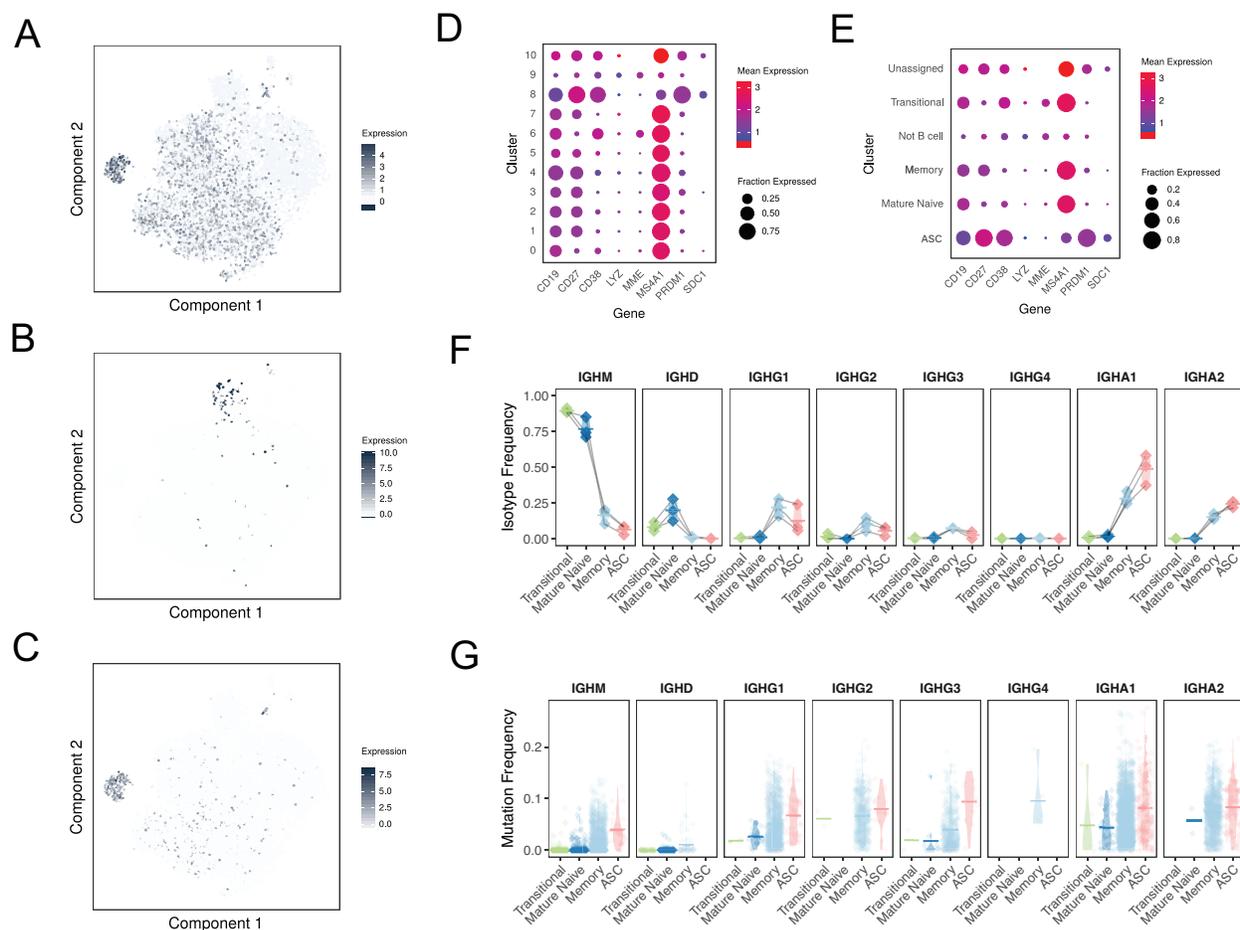


Figure S7. Assignment of B cell clusters based on single-cell transcriptome and repertoire according to known B cell subsets. Plotting of gene expression of key marker genes over t-SNE plot with intensity of shading correlated with amount of scaled log normalized expression for (A) CD27 (B) CD10 (MME) 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.

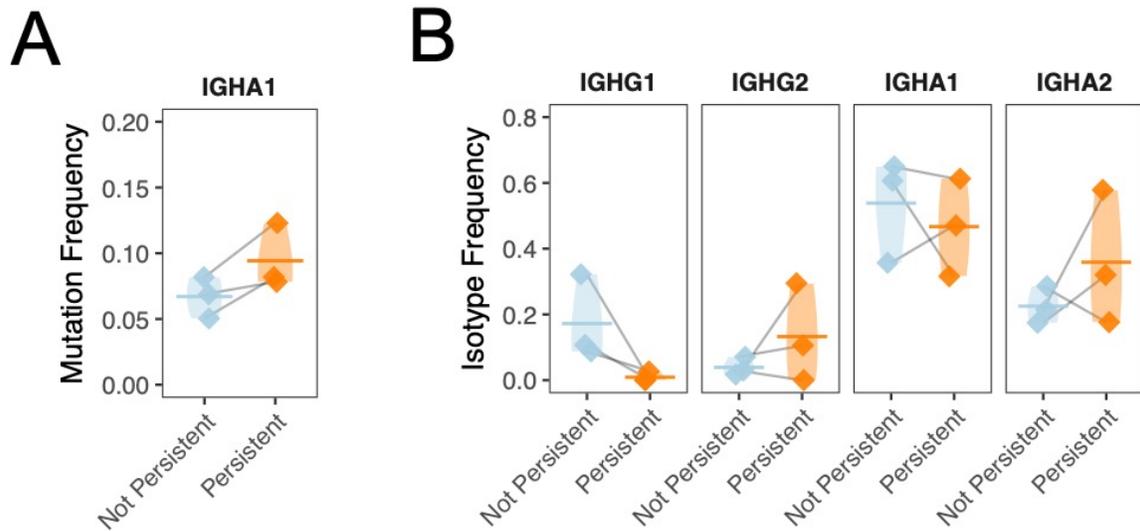


Figure S8. Persistent and non-persistent ASCs are associated with similar V(D)J repertoire features. Overall constant region usage frequencies are quantified for persistent compared to non-persistent cells for the (A) ASC cluster per patient. Horizontal bars show the average frequency of constant region 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. Statistical differences are shown only when significant (**** $P < 0.0001$; *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$).

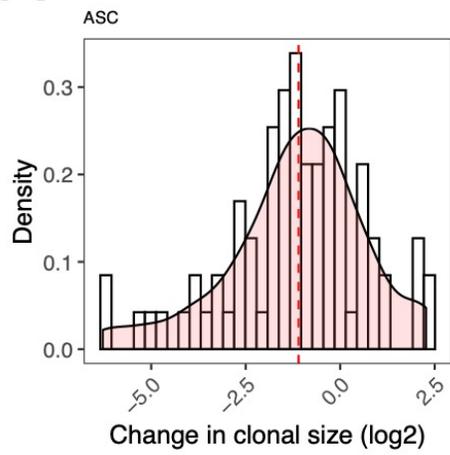
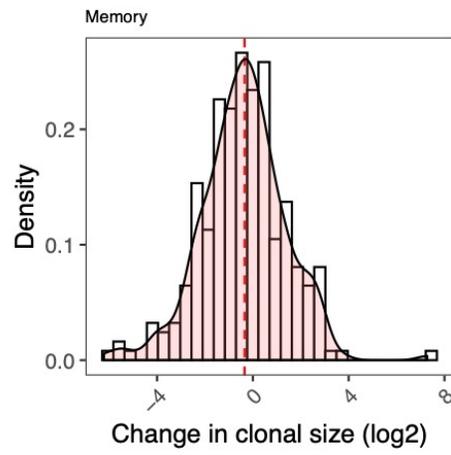
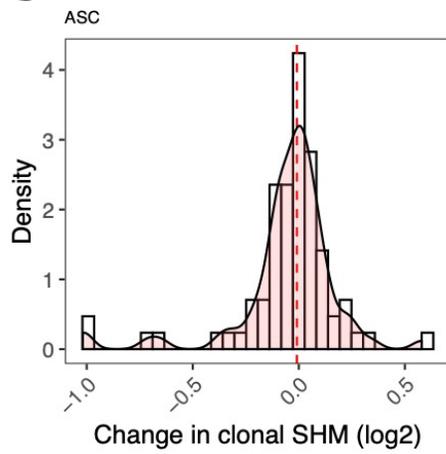
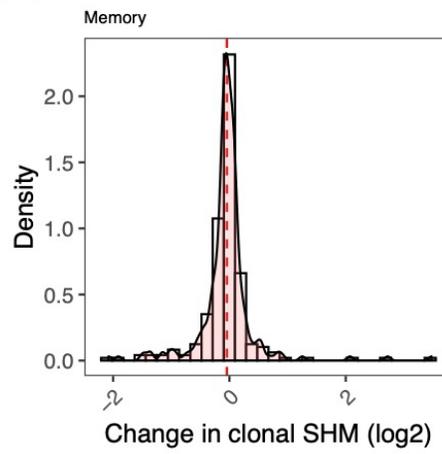
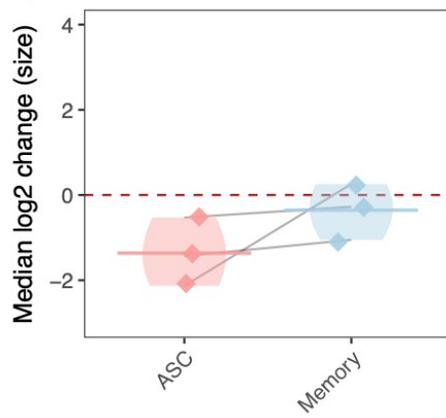
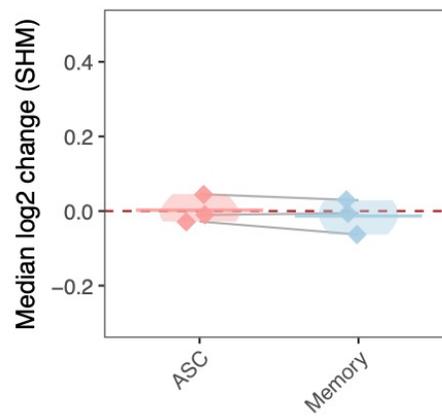
A**B****C****D****E****F**

Figure S9. 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.

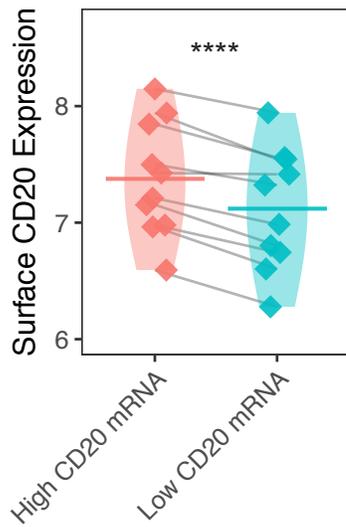


Figure S10. Elevated expression of surface CD20 for B cells with high CD20 mRNA expression. Single cells in the B cell cluster (excluding plasma cells) for each of six subjects in the dataset from Tsang et al. (ref. 30 in main manuscript) were split into high (top 50th percentile) and low (bottom 50th percentile) groups based on CD20 mRNA expression. The mean expression of surface CD20 is presented for each of these groups in each individual sample. Normalized gene expression values are computed from scaled log normalized counts of barcoded transcripts aligned to the CD20 gene. Horizontal bars show the average normalized gene expression across subjects. Expression values belonging to the same subject are paired with a gray line. Statistical differences are shown only when significant (****P < 0.0001; ***P < 0.001; **P < 0.01; *P < 0.05).

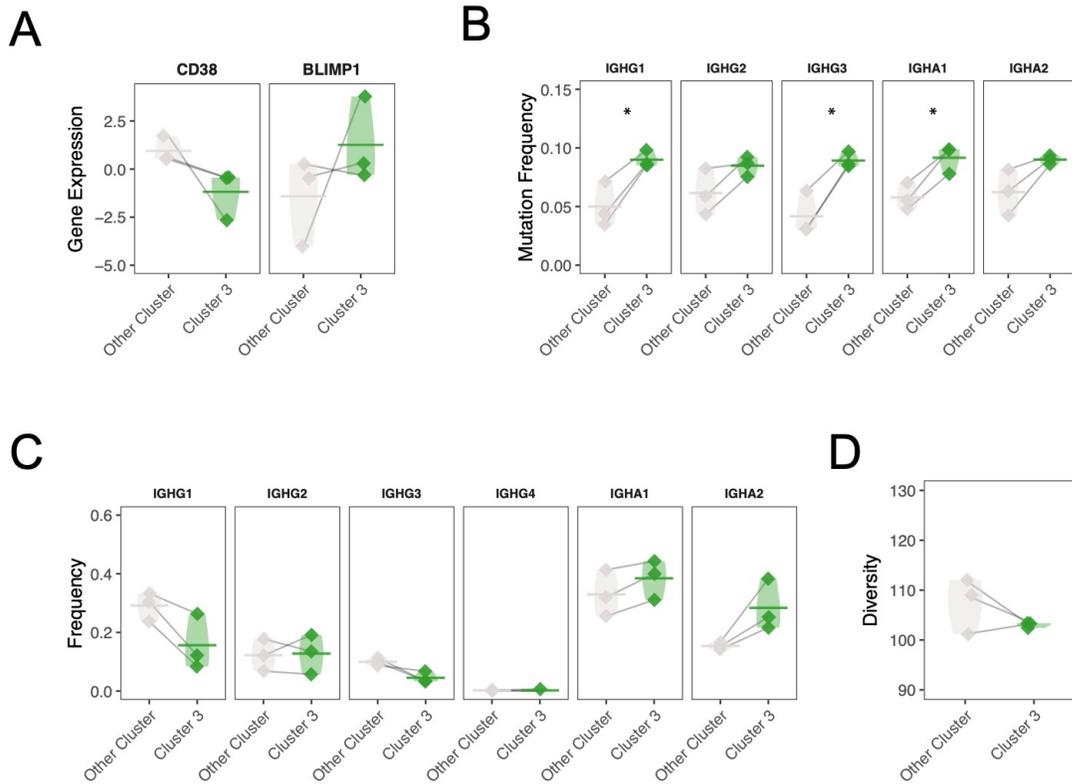


Figure S11. Cluster 3 has distinct transcriptional and V(D)J repertoire features. (A) The expression of plasma cell associated genes (CD38; BLIMP1 or PRDM1) is also presented for members of cluster 3 and members of other memory B cell clusters. Normalized gene expression values are computed from scaled 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. (B) 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. (C) 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. (D) Simpson's diversity of cluster 3 members compared to other memory B cell cluster members. Statistical differences are shown only when significant (****P < 0.0001; ***P < 0.001; **P < 0.01; *P < 0.05).

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.xlsx

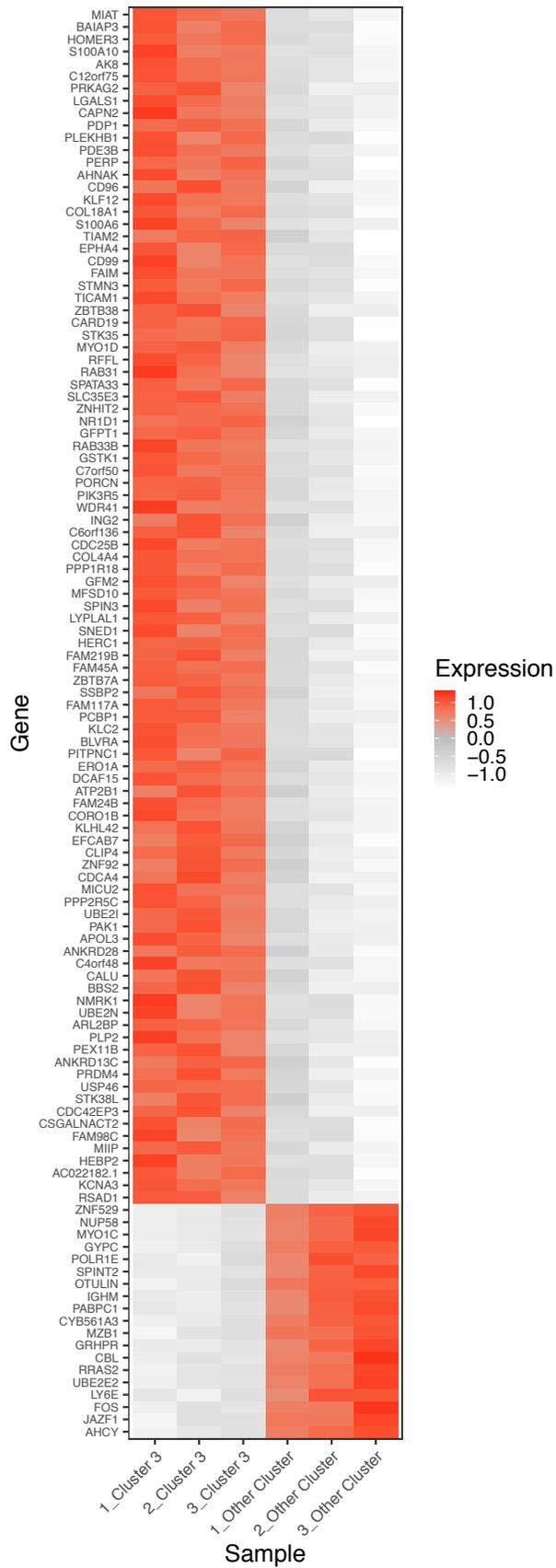


Figure S12. Cluster 3 memory B cells have a distinct gene signature. Heatmap of significantly differentially expressed genes for cluster 3 compared to other memory B cells, with a greater than >0.4 z-score difference between persistent and non-persistent labels.

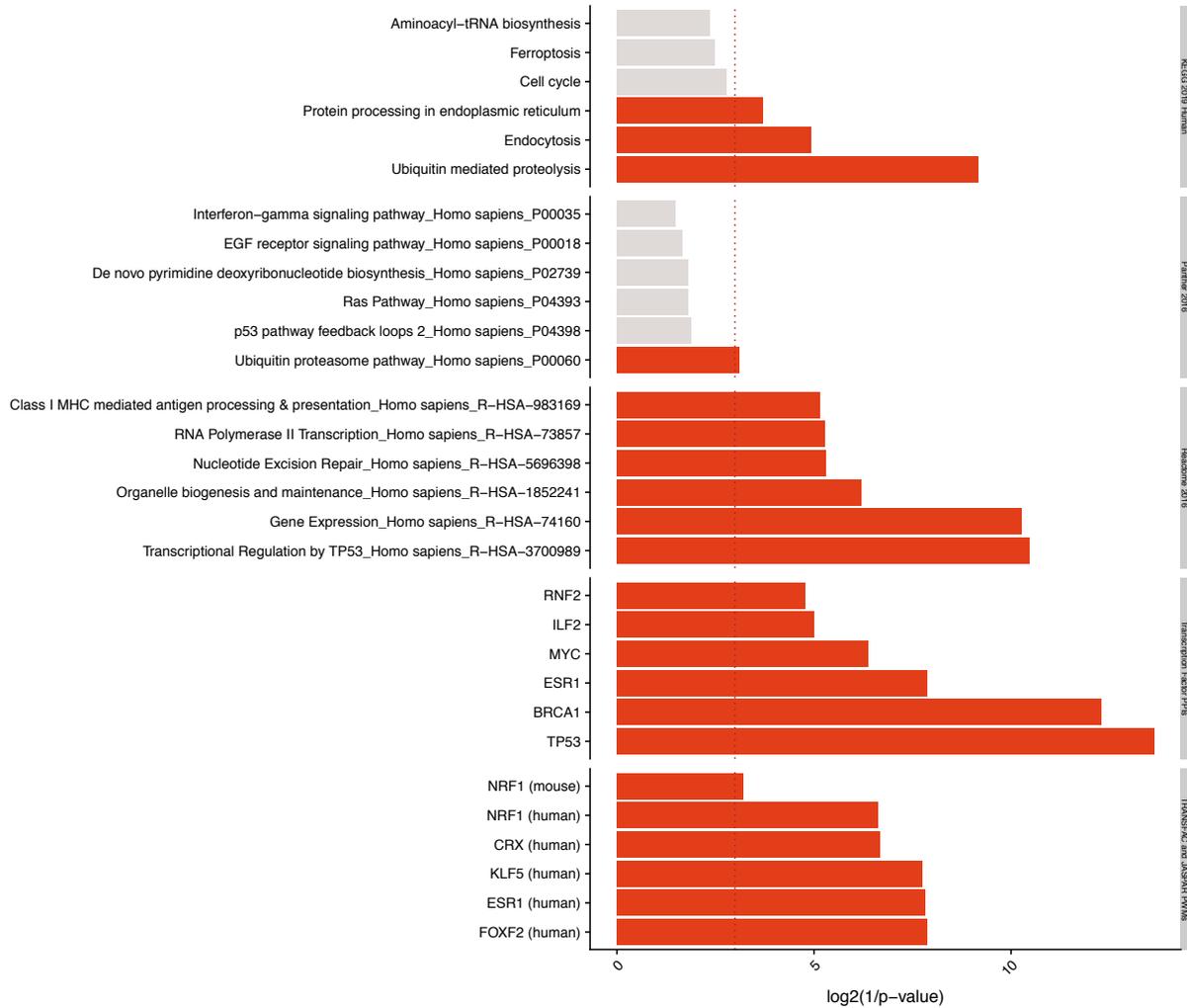


Figure S13. Cluster 3 memory B cells have a MYC and p53/BRCA1 signature. enrichR gene ontology analysis of either genes up-regulated among cluster 3 memory B cells compared to other memory B cells. Only differentially expressed genes with an adjusted p-value less than 0.05 with FDR correction by Storey's method were evaluated. Red bars correspond to significantly associated gene ontology assignments ($p < 0.05$ by Wilcoxon signed rank test).

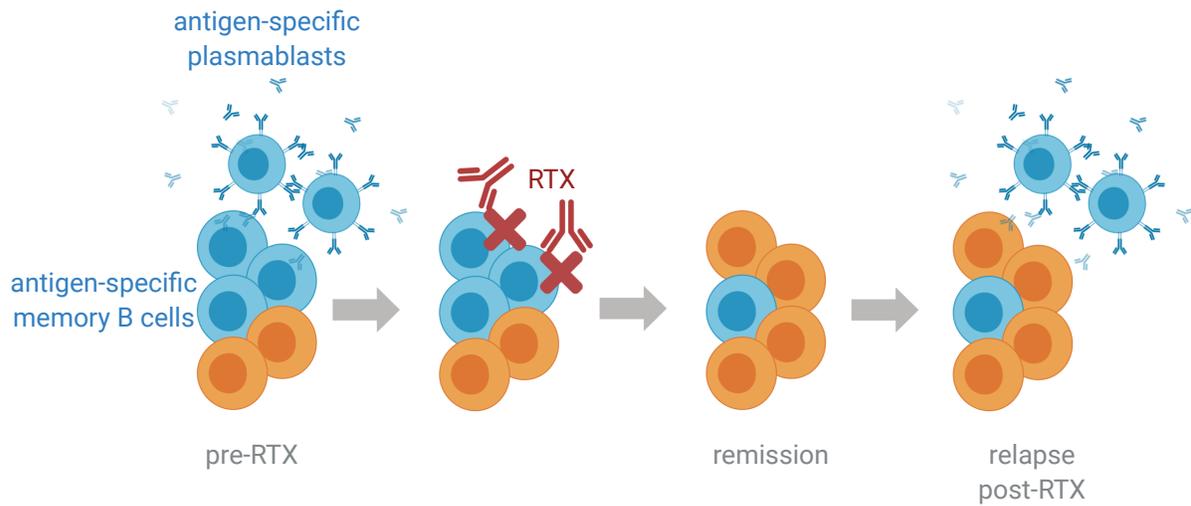


Figure S15. Visual model outlining the likely source of antigen-specific B cells during post-RTX relapse as shown in our study. Blue cells represent antigen-specific B cells whereas orange cells represent non-antigen specific B cells.