eLIFE
elifesciences.org

RESEARCH ARTICLE

Figures and figure supplements

Distinct mechanisms define murine B cell lineage immunoglobulin heavy chain (IgH) repertoires

Yang Yang *et al*

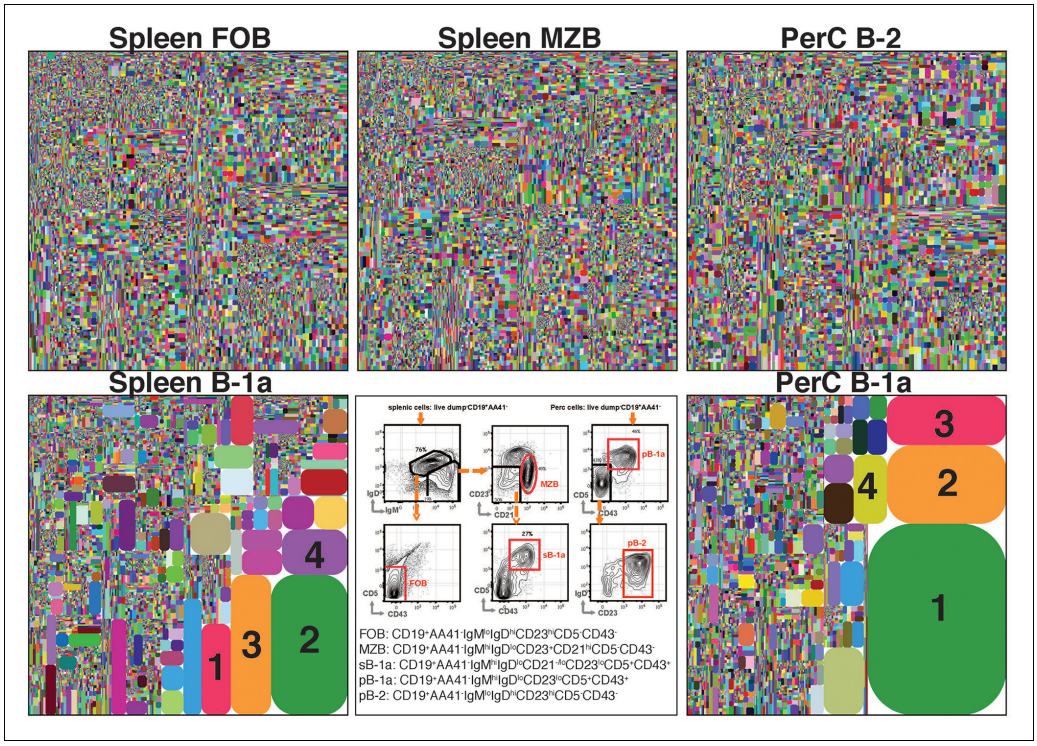

Figure 1. The B-1a IgH CDR3 sequences are much less diverse and recur more frequently than the CDR3 sequences expressed by FOB and MZB B subsets. IgH CDR3 tree-map plots illustrating the IgH CDR3 nucleotide sequences expressed by indicated B cell subsets sorted from one 2-month old C57Bl/6 mouse. Each rectangle in a given tree-map represents a unique CDR3 nucleotide sequence and the size of each rectangle denotes the relative frequency of an individual sequence. The colors for the individual CDR3 sequences in each tree-map plot are chosen randomly thus do not match between plots. The numbers shown in the CDR3 tree-map plots highlight the highly reoccurring CDR3 sequences including PtC-binding CDR3 sequences. 1, ARFYYYGSSYAMDY, V1-55D1-1J4; 2, MRYGNYWYFDV, V11-2D2-8J1; 3, MRYSNYWYFDV, V11-2D2-6J1; 4, MRYGSSYWYFDV, V11-2D1-1J1. *Lower middle panel*: FACS plots showing the gating strategy used to sort the phenotypically defined each B cell subset from spleen (s) or peritoneal cavity (p). Note: peritoneal B-1a cells are well known to express CD11b, a marker expressed on many myeloid cells including macrophage and neutrophils. The level of CD11b expressed on peritoneal B-1a cells, however, is roughly 100 fold lower than the level of CD11b expressed on the myeloid cells. This drastic difference is sufficient to separate the $CD11b^{+}$ B-1a cells from the myeloid cells if monoclonal anti-CD11b reagent is included in the dump channel (***Figure 1—figure supplement 3***).

DOI: 10.7554/eLife.09083.003

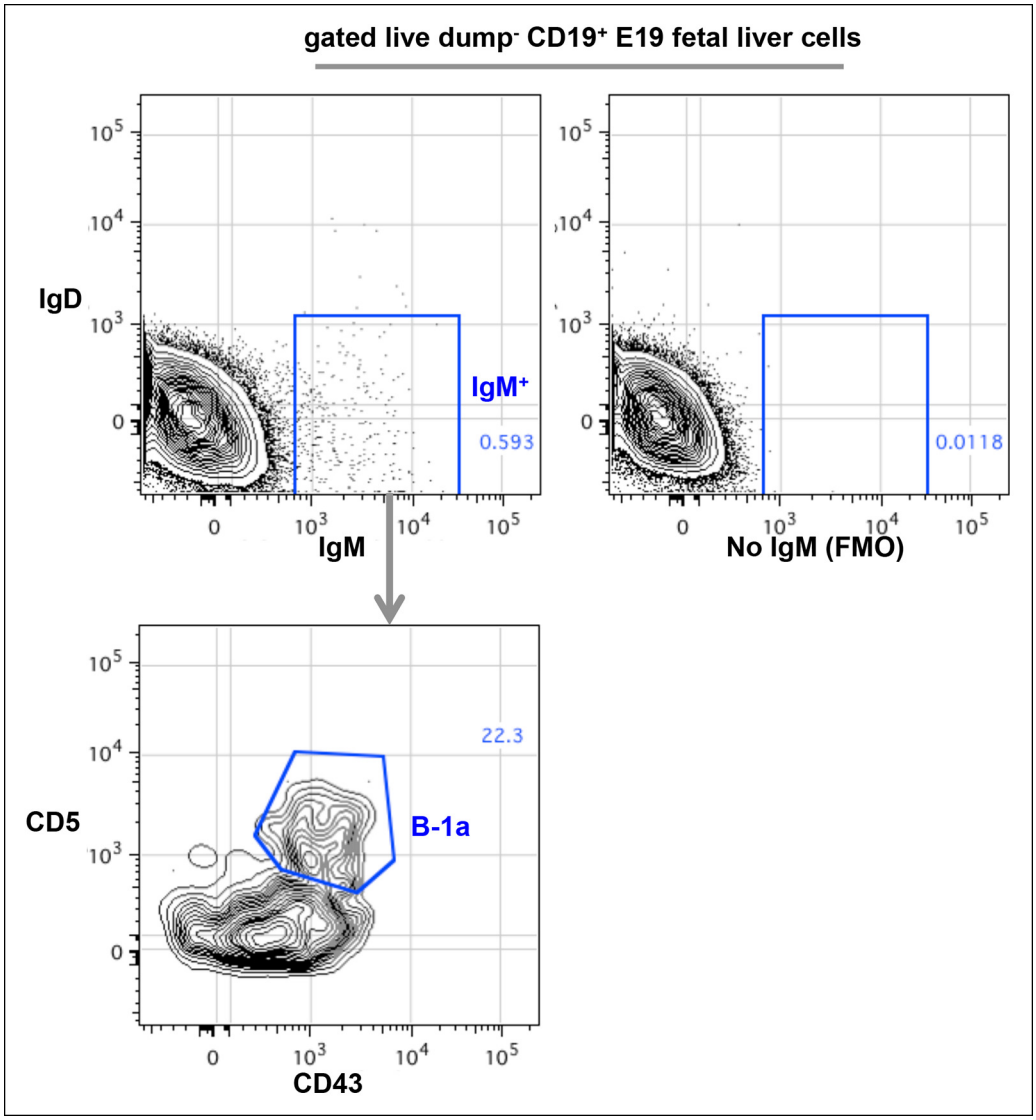

Figure 1—figure supplement 1. FACS plots showing $CD43^+$ $CD5^+$ IgM^+ B-1a cells in E19 fetal liver. Live $dump^-$ ($CD11b^-$ $CD11c^-$ $Gr\text{-}1^-$ $F4/80^-$ $CD3^-$ $TCR\alpha\beta^-$) $CD45^+$ $CD19^+$ cells from E19 fetal liver of C57Bl/6 mouse were gated to show IgM and IgD expression. The boundary for IgM expression was determined from fluorescence-minus-one (FMO) control in which fluorescently labeled anti-mouse IgM antibodies are omitted from the staining sets (right plot). IgM^+ IgD^- cells were further gated to reveal $CD43^+$ $CD5^+$ B-1a cells.

DOI: 10.7554/eLife.09083.004

A: V11-2 D2-6 J1

```
>M01518:53:000000000-A7MVN:1:1101:12630:6931 COPY:2116 mIGHV11-2*01[0] mIGHD2-6*01 mIGHJ1*03[0] mIGHM*03 MRYSNYWYFDV
 T  P  G  K  T  L  E  W  I  G  D  I  N  S  D  G  S  A  I  N  Y  A  P  S  I  K  D  R  F  T  I  F  R  D  N  D  K  S  T  L
          .     .     .     .     .     †     .     .     .     .     .     ‡     .     .
acacctgggaagaccctggagtggaTTGGAGACATTAATTCTGATGGCAGTGCAATAAACTACGCACCATCCATAAAGGATCGATTCACTATCTTCAGAGACAATGACAAGAGCACCCTG
ACACCTGGGAAGACCCTGGAGTGGATTGGAGACATTAATTCTGATGGCAGTGCAATAAACTACGCACCATCCATAAAGGATCGATTCACTATCTTCAGAGACAATGACAAGAGCACCCTG

 Y  L  Q  M  S  N  V  R  S  E  D  T  A  T  Y  F  C  M  R  Y  S  N  Y  W  Y  F  D  V  W  G  T  G  T  T  V  T  V  S  S  E
          .     .     .     †     .     .     .     .     .     .     ‡     .     .     .     .     .     .     .     .
TACCTGCAGATGAGCAATGTGCGATCGGAGGACACAGCCACGTATTTCTGTATGAGATATAGTAACTACTGGTACTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTCTCCTCAGAG
TACCTGCAGATGAGCAATGTGCGATCGGAGGACACAGCCACGTATTTCTGTATGAGATATAGTAACTACTGGTACTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTCTCCTCAGAG
```

B: V11-2 D2-8 J1

```
>M01518:53:000000000-A7MVN:1:1103:15890:4689 COPY:138 mIGHV11-2*01[0] mIGHD2-8*01 mIGHJ1*03[0] mIGHM*03 MRYGNYWYFDV
 T  P  G  K  T  L  E  W  I  G  D  I  N  S  D  G  S  A  I  N  Y  A  P  S  I  K  D  R  F  T  I  F  R  D  N  D  K  S  T  L
          .     .     .     .     .     †     .     .     .     .     .     ‡     .     .
acacctgggaagaccctggagtggaTTGGAGACATTAATTCTGATGGCAGTGCAATAAACTACGCACCATCCATAAAGGATCGATTCACTATCTTCAGAGACAATGACAAGAGCACCCTG
ACACCTGGGAAGACCCTGGAGTGGATTGGAGACATTAATTCTGATGGCAGTGCAATAAACTACGCACCATCCATAAAGGATCGATTCACTATCTTCAGAGACAATGACAAGAGCACCCTG

 Y  L  Q  M  S  N  V  R  S  E  D  T  A  T  Y  F  C  M  R  Y  G  N  Y  W  Y  F  D  V  W  G  T  G  T  T  V  T  V  S  S  E
          .     .     .     †     .     .     .     .     .     .     ‡     .     .     .     .     .     .     .     .
TACCTGCAGATGAGCAATGTGCGATCGGAGGACACAGCCACGTATTTCTGTATGAGATATGGTAACTACTGGTACTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTCTCCTCAGAG
TACCTGCAGATGAGCAATGTGCGATCGGAGGACACAGCCACGTATTTCTGTATGAGATATGGTAACTACTGGTACTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTCTCCTCAGAG
```

C: V11-2 D1-1 J1

```
>M01518:53:000000000-A7MVN:1:1101:10142:6844 COPY:1809 mIGHV11-2*01[0] mIGHD1-1*01 mIGHJ1*03[0] mIGHM*03 MRYGSSYWYFDV
 T  P  G  K  T  L  E  W  I  G  D  I  N  S  D  G  S  A  I  N  Y  A  P  S  I  K  D  R  F  T  I  F  R  D  N  D  K  S  T  L
          .     .     .     .     .     †     .     .     .     .     .     ‡     .     .
acacctgggaagaccctggagtggattgGAGACATTAATTCTGATGGCAGTGCAATAAACTACGCACCATCCATAAAGGATCGATTCACTATCTTCAGAGACAATGACAAGAGCACCCTG
ACACCTGGGAAGACCCTGGAGTGGATTGGAGACATTAATTCTGATGGCAGTGCAATAAACTACGCACCATCCATAAAGGATCGATTCACTATCTTCAGAGACAATGACAAGAGCACCCTG

 Y  L  Q  M  S  N  V  R  S  E  D  T  A  T  Y  F  C  M  R  Y  G  S  S  Y  W  Y  F  D  V  W  G  T  G  T  T  V  T  V  S  S
          .     .     .     †     .     .     .     .     .     .     ‡     .     .     .     .     .     .     .     .
TACCTGCAGATGAGCAATGTGCGATCGGAGGACACAGCCACGTATTTCTGTATGAGATACGGTAGTAGCTACTGGTACTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTCTCCTCA
TACCTGCAGATGAGCAATGTGCGATCGGAGGACACAGCCACGTATTTCTGTATGAGATACGGTAGTAGCTACTGGTACTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTCTCCTCA
```

Figure 1—figure supplement 2. Recurrent V_H11-encoded PtC-binding V(D)J sequences. (**A-C**) lists three V_H11-encoded PtC-binding V(D)J sequences. In each plot, the first line of nucleotides is the obtained sequence read while the second line refers the germline reference sequence. The underlined nucleotides are CDR2 and CDR3.

DOI: 10.7554/eLife.09083.005

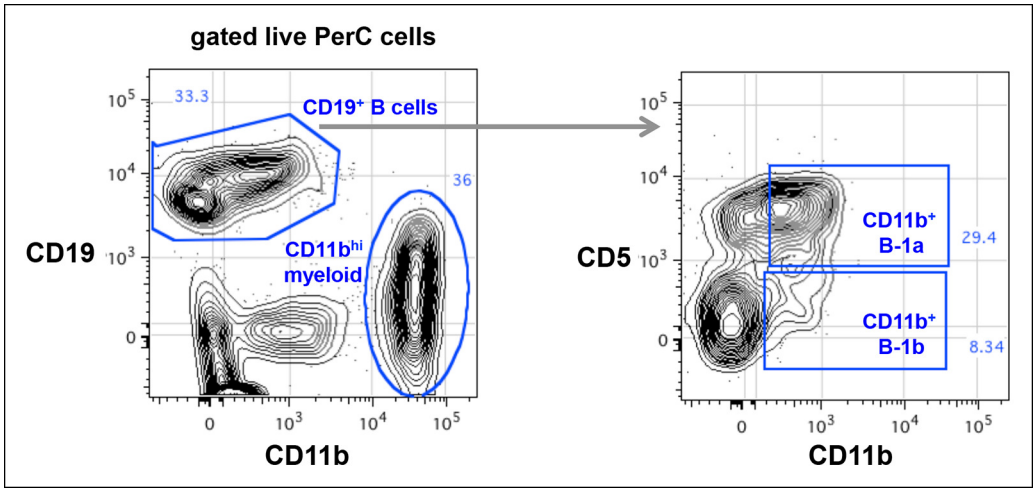

Figure 1—figure supplement 3. CD11b expression on peritoneal B-1a ($CD5^+$) and B-1b ($CD5^-$) is roughly 100-fold lower than the CD11b expression on myeloid cells. Live cells from C57Bl/6 peritoneal cavity were gated to show CD19 and CD11b expression. The $CD19^+$ B cells and $CD11b^{hi}$ myeloid cells were shown. The $CD19^+$ B cells were gated to reveal CD5 and CD11b expression. $CD11b^+$ B-1a and $CD11b^+$ B-1b cells were gated based on FMO control staining where anti-CD11b antibody was omitted in the staining.

DOI: 10.7554/eLife.09083.006

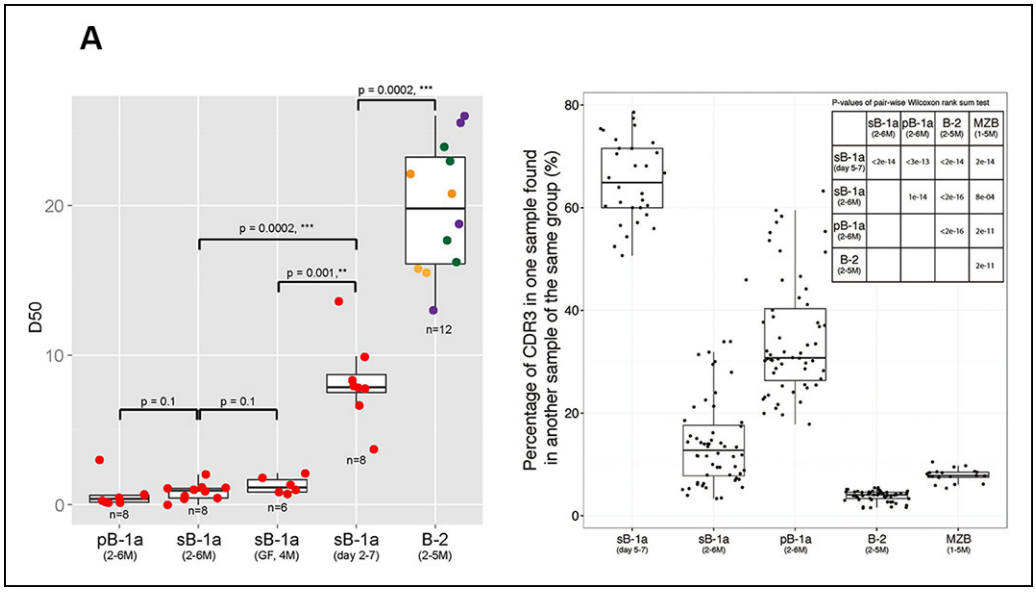

Figure 2. The B-1a pre-immune IgH repertoire is far more restricted than the pre-immune IgH repertoires expressed by splenic FOB, MZB and peritoneal B-2 cells. (**A**) D50 metric analysis quantifying the IgH CDR3 diversity for B cell subsets from mice at the indicated age. Low D50 values are associated with less diversity. Each dot represents the data for a B cell sample from an individual mouse except for the 2 day splenic B-1a data, which are derived from sorted cells pooled from 8 mice. B-1a samples are labeled with red; B-2 samples include FOB (green, n = 4), pB-2 (purple, n = 4) and MZB (yellow, n = 4). The data for germ-free (GF) animals is discussed at the end of the Result section. (**B**) CDR3 peptide pair-wise sharing analysis of IgH repertoire similarity among multiple samples for each B cell group (n = 5-9). Each dot represents the percentage of common CDR3 peptides in one sample that are also found in another sample within a given group. For example, to compute the similarity between sample A and B, the percentage of CDR3 peptides in sample A that are also found in sample B ($p_{A \longrightarrow B}$), together with the percentage of CDR3s in sample B that are also in sample A ($p_{B \longrightarrow A}$) are used as an indicator. For comparison of 6 splenic B-1a samples in 5-7 day group, there are 30 comparisons. *Right upper:* p values showing the statistical significance between two groups. Box plots represent the 10th, 25th, 50th, 75th and 90th percentiles here and in other figures.

DOI: 10.7554/eLife.09083.008

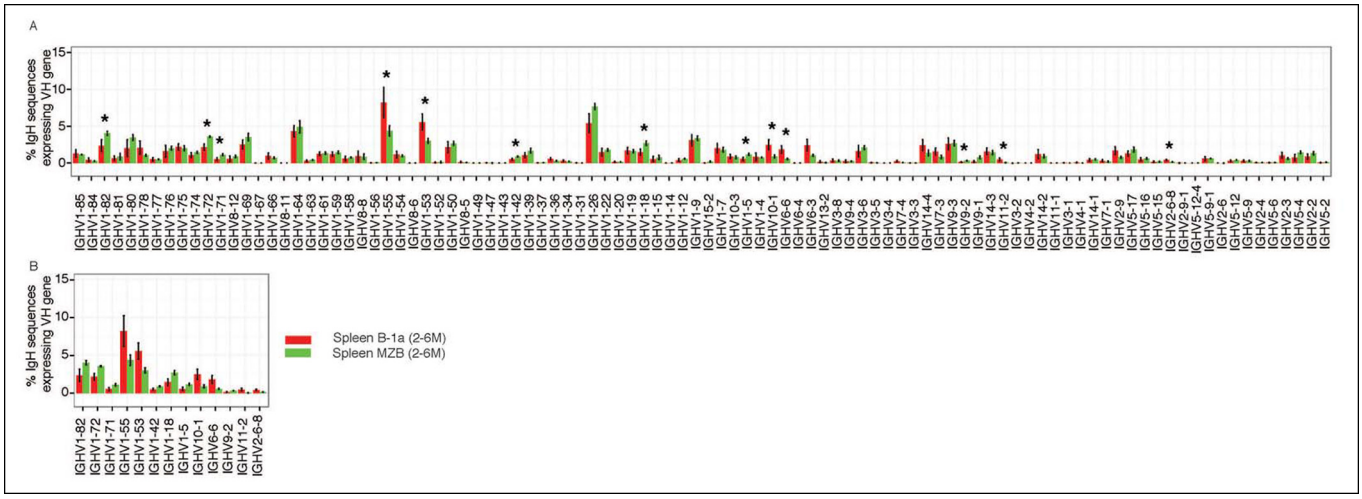

Figure 3. Comparison of V_H gene usage by splenic B-1a vs MZB B cells. (**A**) V_H gene usage profile shown as the percentage of IgH sequences expressing the listed individual V_H genes for individual B cell samples. The profiles are shown for adult splenic B-1a samples (n = 9, red) and for MZB samples (n = 5, green). V_H genes (from left to right) are ordered in 5′- to 3′-direction bases on chromosome location; the IMGT V_H gene nomenclature is used (***Lefranc, 2003***). (**B**) V_H genes showing the statistically significant differences (Welch's t-test $p<0.05$) between two groups are listed and also highlighted with asterisks in the plot. To minimize the impact of the clonal expansion on the V_H gene usage profile, data are presented as the normalized distribution that counts each distinct CDR3 nucleotide sequence expressing a given V_H gene as one, no matter how many times the sequence was detected. Note: V_H12-3 encoded IgH sequences are not detected in this study due to the technical limitations that exclude the V_H12-3 primer from the set of primers designed about three years ago and used for studies presented here. We have since corrected this problem so that V_H12-3 primer is now part of our new set of primers. Comparison of sequence data obtained with old vs. the new set of primers shows that, aside from now detecting V_H12-3 sequences with the new set of primers, the sequences obtained with both primer sets are highly similar (***Figure 3—figure supplement 2***).

DOI: 10.7554/eLife.09083.009

Figure 3—figure supplement 1. V_H gene usage profile pair-wise comparison of B cell groups. The colors shown at the bottom right distinguish the B cell groups (n = 4-9). V_H genes showing the statistically significant differences (Welch's t-test p<0.05) between two groups are listed on the bottom (**A′ to F′**) and also highlighted with asterisks in each plot. The data for germ-free (GF) animals is discussed at the end of the Result section.

DOI: 10.7554/eLife.09083.010

splenic B-1a #1

top 10 CDR3 sequences (old primers)

CDR3 (peptide)	V domain	J domain	Copy
ASYDIYAMDY	mIGHV1-26	mIGHJ4	18787
MRYSNYWYFDV	mIGHV11-2	mIGHJ1	16483
AREGTYGSSPLDY	mIGHV1-76	mIGHJ2	12884
ASYSFAY	mIGHV1-50	mIGHJ3	10907
ARKELGRAHWCFDV	mIGHV1-55	mIGHJ1	5908
ARSEFAY	mIGHV1-50	mIGHJ3	5121
TPGLGAY	mIGHV6-6	mIGHJ3	4717
AHGYLDY	mIGHV1-50	mIGHJ2	4174
ARTYGIRYWYFDV	mIGHV1-76	mIGHJ1	3467
AHATEYYAMDY	mIGHV14-3	mIGHJ4	3399

top 10 CDR3 sequences (new primers)

CDR3 (peptide)	V domain	J domain	Copy
MRYSNYWYFDV	mIGHV11-2	mIGHJ1	10429
AREGTYGSSPLDY	mIGHV1-76	mIGHJ2	10387
ASYDIYAMDY	mIGHV1-26	mIGHJ4	7805
ASYSFAY	mIGHV1-50	mIGHJ3	6820
ARKELGRAHWCFDV	mIGHV1-55	mIGHJ1	4499
TPGLGAY	mIGHV6-6	mIGHJ3	3094
ARTYGIRYWYFDV	mIGHV1-76	mIGHJ1	2942
AHGYLDY	mIGHV1-50	mIGHJ2	2918
ARSEFAY	mIGHV1-50	mIGHJ3	2702
AHATEYYAMDY	mIGHV14-3	mIGHJ4	2267

splenic B-1a #2

top 10 CDR3 sequences (old primers)

CDR3 (peptide)	V domain	J domain	Copy
ASGGNYFSYFDY	mIGHV1-26	mIGHJ2	19808
ARRYYGSSYWYFDV	mIGHV1-55	mIGHJ1	15912
TRSNYYGSRSFDY	mIGHV1-15	mIGHJ2	13684
ASGDYLYYFDY	mIGHV1-84	mIGHJ2	12377
TDPPSDY	mIGHV6-3	mIGHJ2	11233
ARTSSGSFDY	mIGHV9-3	mIGHJ2	10498
ARRYYGSSYAMDY	mIGHV1-50	mIGHJ4	10411
AVRRSSGSLDY	mIGHV1-53	mIGHJ2	9494
ARATLYSNYAMDY	mIGHV1-53	mIGHJ4	9194
ARLETGY	mIGHV1-50	mIGHJ2	7352

top 10 CDR3 sequences (new primers)

CDR3 (peptide)	V domain	J domain	Copy
ASGGNYFSYFDY	mIGHV1-26	mIGHJ2	7548
ARRYYGSSYWYFDV	mIGHV1-55	mIGHJ1	6421
ASGDYLYYFDY	mIGHV1-84	mIGHJ2	4943
TRSNYYGSRSFDY	mIGHV1-15	mIGHJ2	4127
ARRYYGSSYAMDY	mIGHV1-50	mIGHJ4	3964
AVRRSSGSLDY	mIGHV1-53	mIGHJ2	3456
ARATLYSNYAMDY	mIGHV1-53	mIGHJ4	3000
ARLETGY	mIGHV1-50	mIGHJ2	2944
AYGYYHY	mIGHV1-75	mIGHJ2	2824
TDPPSDY	mIGHV6-3	mIGHJ2	2808

Figure 3—figure supplement 2. Almost identical top 10 highly recurring CDR3 sequences are detected for splenic B-1a IgH libraries obtained either with the old or new primer set. We sorted two splenic B-1a populations individually from two 4 month old C57BL/6J mice. We extracted RNA from each population and divided each RNA into two parts. For one part, we prepared an amplified library using the old primer set; and for the other, we prepared an amplified library using the new primer set. We then sequenced these amplified IgH libraries. Analysis of the resultant sequences showed that the sequences obtained from the IgH libraries are highly similar, regardless of the primers used (old or new). In essence, the top 10 highly recurring CDR3 sequences (both peptide and V(D)J recombination) are almost identical and show similar representation order between each pair of libraries. As expected, we detected V_H12-3 encoded sequences from the splenic B-1a IgH libraries prepared with the new primer set, and these V_H12-3 encoded sequences included several published PtC-binding V_H12-3 encode sequences, i.e., AGDYDGYWYFDV (V_H12-3D2-4J1), AGDRDGYWYFDV (V_H12-3D3-2J1), AGDRYGYWYFDV (V_H12-3 D2-9 J1).

DOI: 10.7554/eLife.09083.011

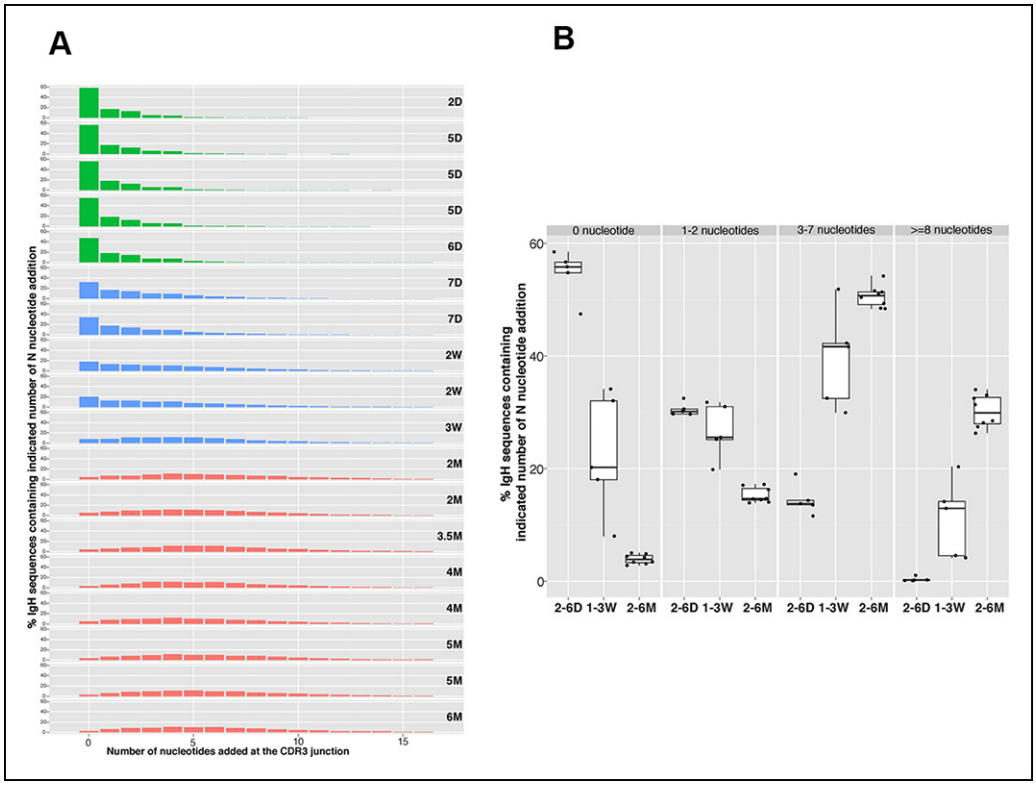

Figure 4. N nucleotide insertion distribution patterns for the B-1a pre-immune IgH repertoires during ontogeny. (**A**) Percentage of IgH sequences containing the indicated number of N nucleotide insertions at the IgH CDR3 junctions (V-DJ + D-J) is shown for each spleen B-1a sample from mice at indicated ages (shown at the right). To minimize the impact of self-renewal on the N-addition profile, normalized data are presented. Thus, each distinct IgH sequence containing indicated N nucleotide insertions is counted as one regardless how many times this sequence was detected. Note that the N insertion pattern changes as animals age. Colors distinguish three age-related patterns: green, D2 to D6; blue, D7 to 3W; red, 2M to 6M. (**B**) Percentages of IgH sequences containing the indicated N-nucleotide insertions (shown at the top) for splenic B-1a samples at the indicated ages are shown. Each dot represents data from an individual mouse, except for day 2 sample, n = 5-7.

DOI: 10.7554/eLife.09083.012

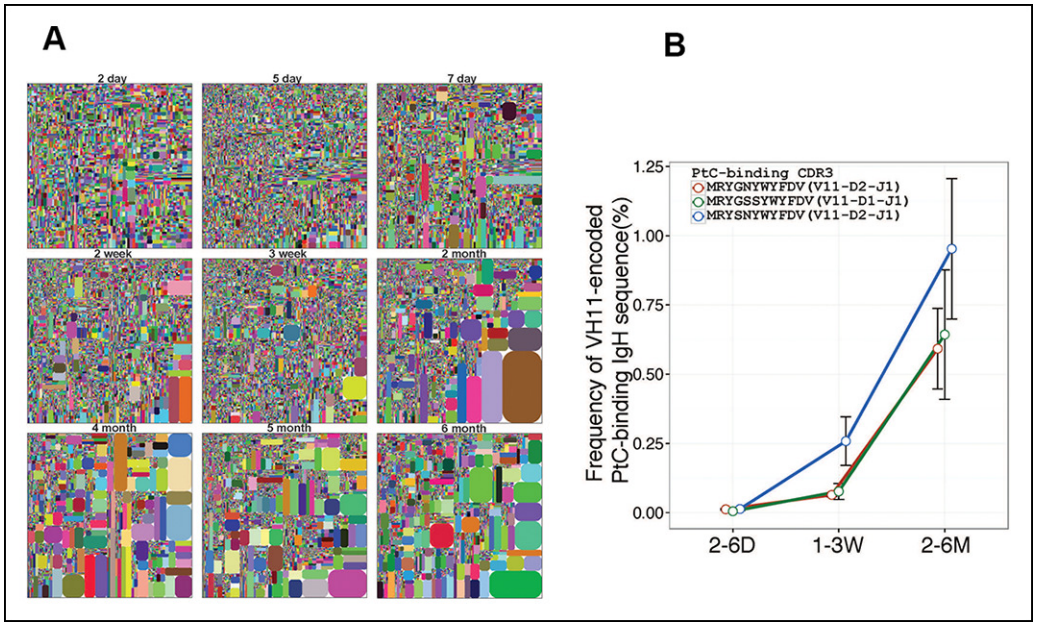

Figure 5. Certain V(D)J sequences increase progressively with age in the B-1a pre-immune IgH repertoire. (**A**) IgH CDR3 tree map plots for splenic B-1a samples from mice at different ages are shown. Each plot represents data for an individual mouse, except for the day 2 sample. Recurrent sequences are visualized as larger contiguously-colored rectangles in each plot. (**B**) Relative frequencies of three PtC-binding IgH CDR3 sequences in indicated splenic B-1a sample groups (n = 5–8 for each group) are plotted with mouse age. Sequence information (peptide and V(D)J recombination) is shown at the top.

DOI: 10.7554/eLife.09083.013

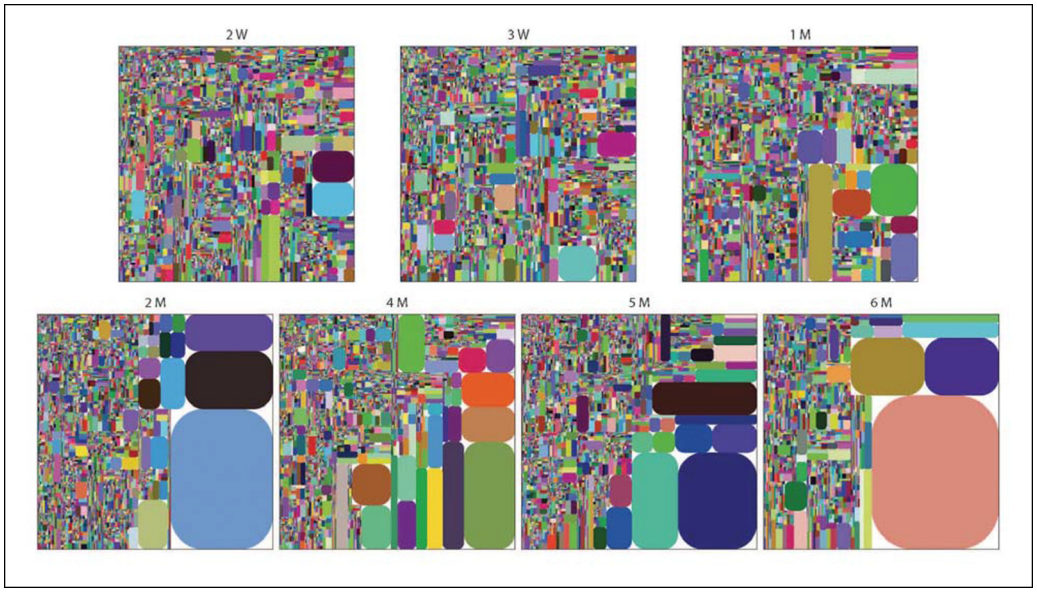

Figure 5—figure supplement 1. The peritoneal B-1a IgH repertoire is increasingly restricted during ontogeny. IgH CDR3 tree map plots for peritoneal B-1a samples from different ontogenic stages. Each plot represents the data for a sample from an age-defined individual mouse, except for the 2 week, 3 week and 1 month samples, which are obtained from cells pooled from several mice. Recurrent sequences are visualized as larger contiguously-colored rectangles in each plot.

DOI: 10.7554/eLife.09083.014

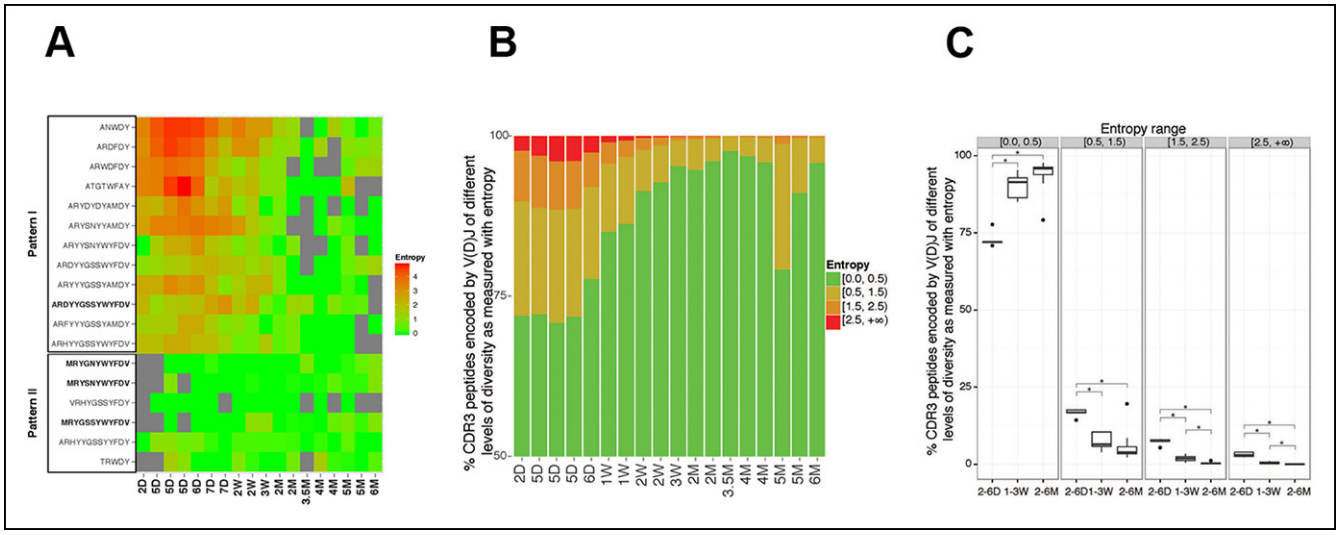

Figure 6. The level of convergent recombination in the B-1a IgH repertoire declines with age. (**A**) Entropy heat map showing the diversity of V(D)J recombination events for each indicated CDR3 peptide (shown at the left) in splenic B-1a samples at different ages (shown at the bottom). The higher the entropy value, the more diverse the V(D)J recombinations for a given CDR3 peptide. CDR3 peptide sequences for T15 Id^{+} anti-PC (pattern I) and anti-PtC (pattern II) antibody are shown in bold. (**B**) The diversities of the V(D)J recombination for each CDR3 peptide for the indicated splenic B-1a samples (shown at the bottom) are quantified as entropy values (see Methods and materials), which are ranked into 4 ranges (shown at the right). For each sample, the frequencies of CDR3 peptide sequences belonging to each entropy range are shown as stacks. (**C**) Splenic B-1a samples are grouped based on age. For each group (n = 5–7), the frequencies of CDR3 peptide sequences belonging to each of four entropy ranges are shown. *p<0.05, Welch's t-test.

DOI: 10.7554/eLife.09083.018

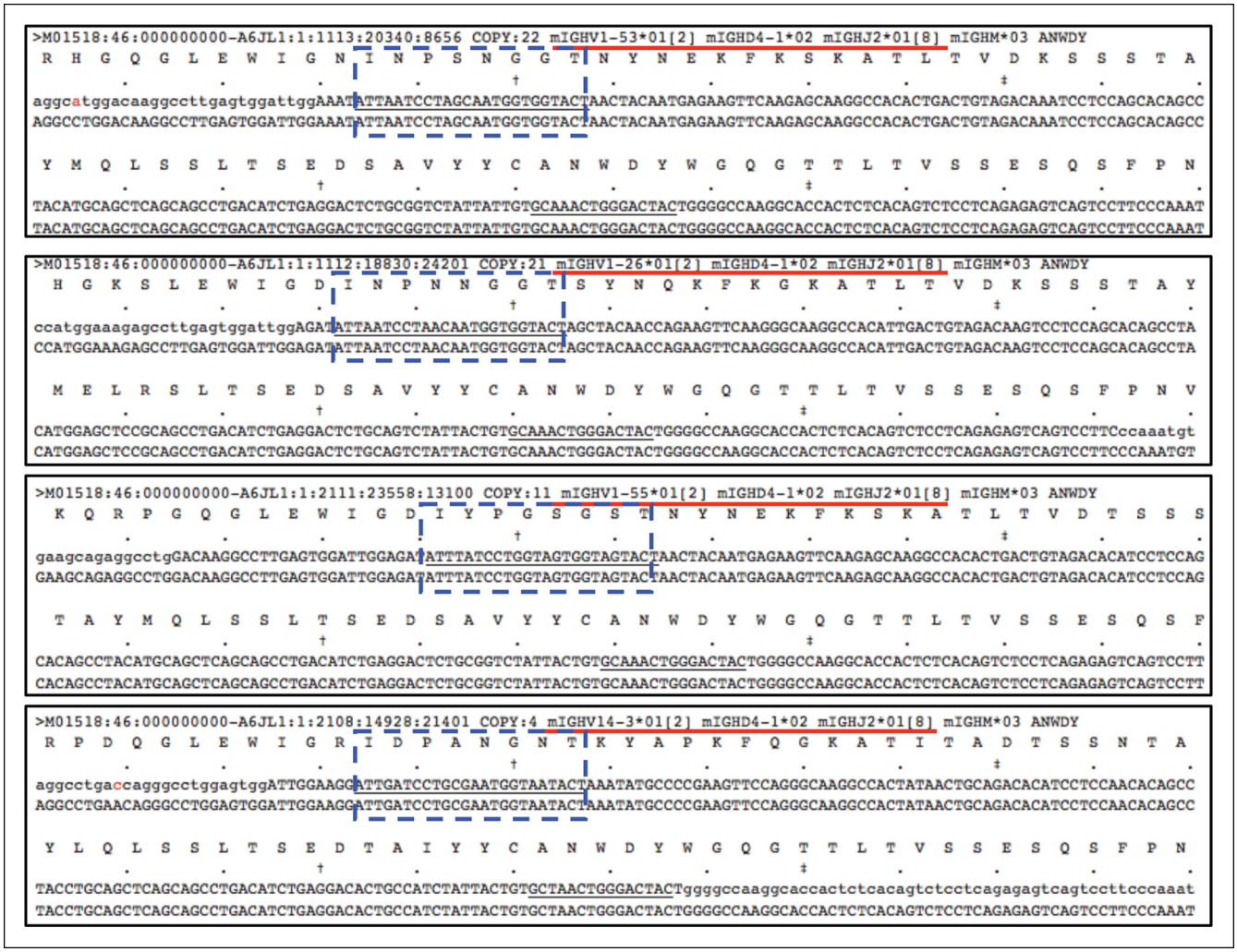

Figure 6—figure supplement 1. Distinct V(D)J sequences encoding the same CDR3 peptide differ in V_H usage. Plots showing an example, in which four different V(D)J sequences expressed by the 5 day splenic B-1a sample all encode the same CDR3 (ANWDY). Red line denotes the V(D)J recombination. CDR2 sequences are highlighted with the blue doted lined box. The V(D)J recombination (V14-3 D4-1 J2) shown in the bottom plot is the predominant V(D)J for ANWDY identified in adult splenic B-1a IgH repertoire.

DOI: 10.7554/eLife.09083.019

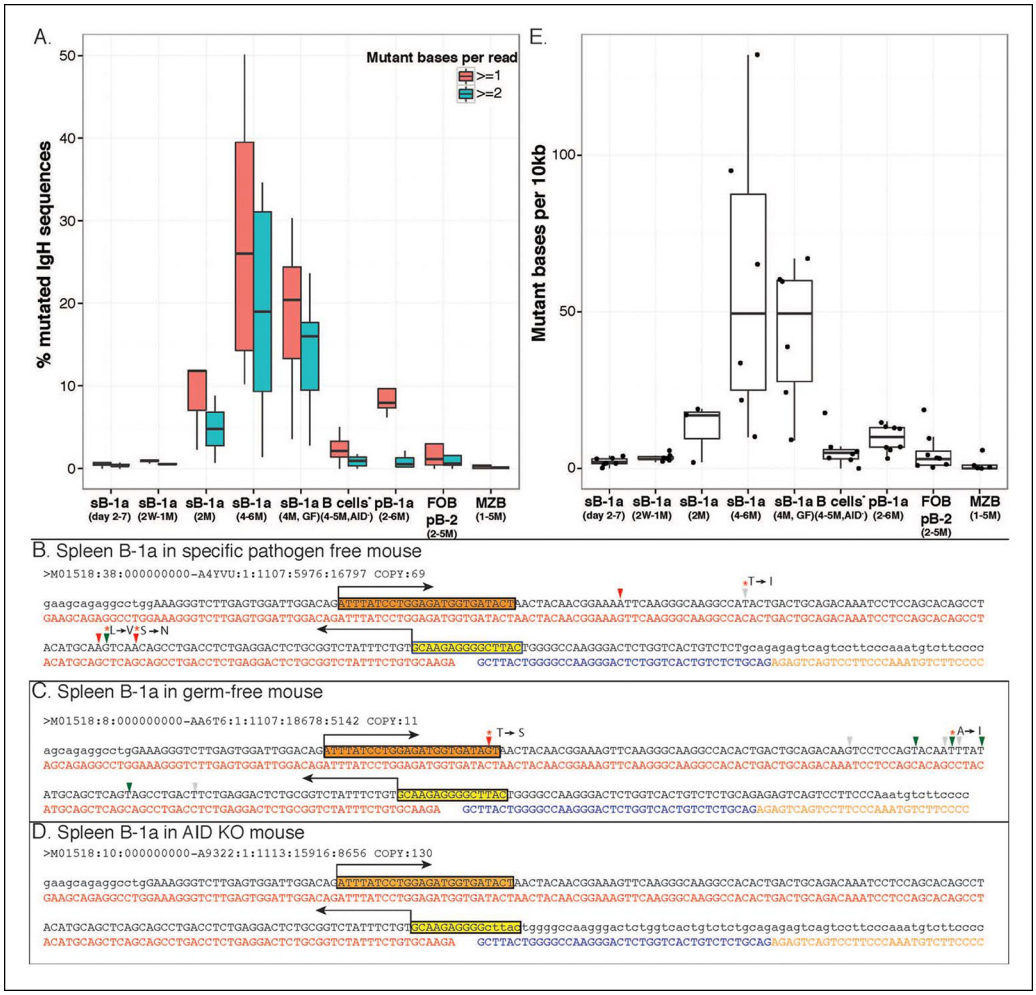

Figure 7. AID-mediated SHM accumulates on splenic B-1a IgV_H with age. (**A**) Percentages of sequences containing > = 1 (red) or > = 2 (green) nucleotide changes for B cell samples from mice at the indicated ages are shown (n = 3-8). Seven B cell samples from 4-5 month old AID knockout mice include sB-1a (n = 4), pB-1a (n = 1), FOB (n = 1) and pB-2 (n = 1). Sequences with the identical V(D)J recombination encoding ARGAY CDR3 peptide obtained from splenic B-1a sample from 4 month old specific pathogen free mouse, (**B**) germ-free mouse (**C**) and AID knockout mouse (**D**) are listed. The nucleotide substitution is analyzed at the V_H region stretching from the start of CDR2 (red box) to the beginning of CDR3 (yellow box). Obtained sequence (upper line) is aligned with the reference (lower line) for V1-80 (red), J3 (blue) and constant region of IgM isotype (orange). Mutations are highlighted with triangles; asterisks indicate mutations resulting in an amino acid change; red and blue triangles denote mutations in DGYW and WRCH motifs, respectively. (**E**) Numbers of mutations per 10^4 base pairs for indicated B cell group are shown. Each dot represents data from an individual sample (n = 3–8). The data for germ-free (GF) animals is discussed at the end of the Result section. Note: The mutation profiles for the splenic B-1a IgH libraries prepared by using either old (V_H12-3 deficient) or new primer set (V_H12-3 included) are highly similar (***Figure 7—figure supplement 3***).

DOI: 10.7554/eLife.09083.020

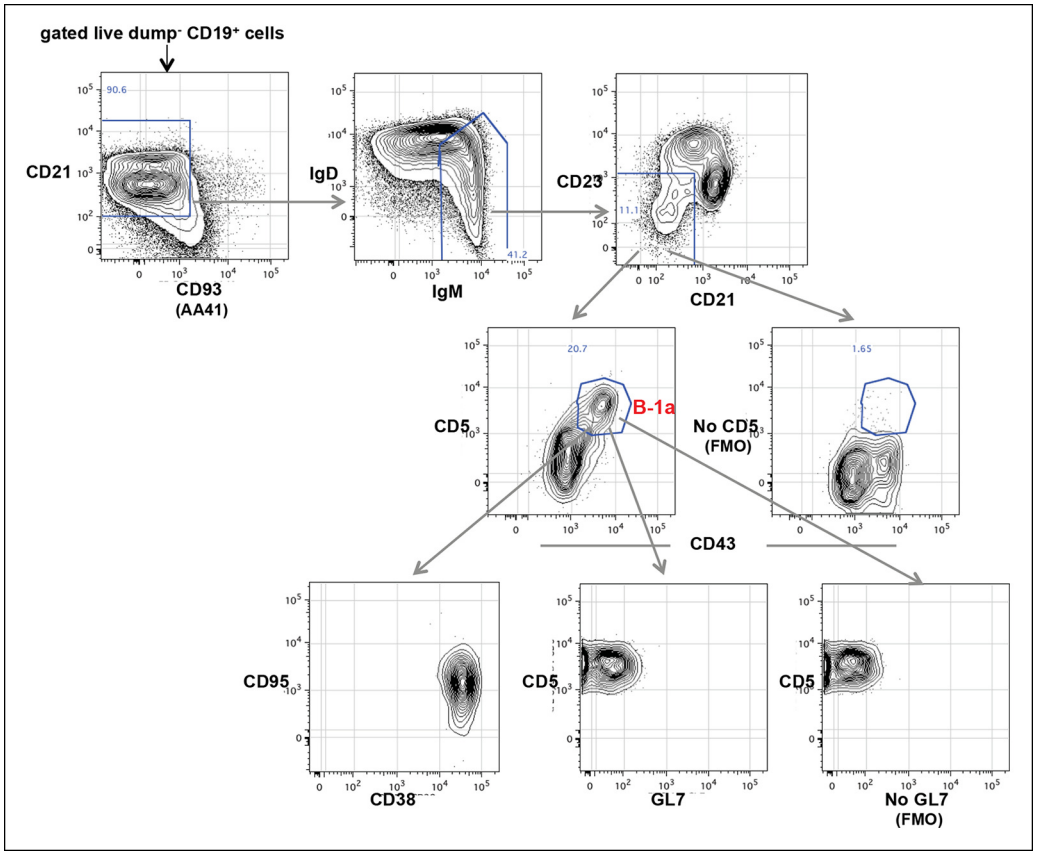

Figure 7—figure supplement 1. Splenic B-1a cells do not contain cells expressing GC phenotype. FACS analysis showing of live dump^{-} CD19^{+} CD93 (AA41)$^{-}$ IgM^{Hi} IgD$^{-/lo}$ CD23^{-} CD21$^{-/lo}$ B cells from spleen of 5 month old C57BL6/J mouse were gated to reveal CD43^{+} CD5^{+} B-1a cells, which were further gated to reveal GL7, CD38 and CD95 expression. GC B cells are GL7^{+} CD38$^{-/lo}$ CD95^{hi}. The boundary for CD5 (*rightmost middle plot*) and GL7 (*rightmost bottom plot*) expression were determined from FMO controls in which fluorescently labeled anti-mouse CD5 or anti-mouse GL7 antibodies are omitted from the staining sets.

DOI: 10.7554/eLife.09083.021

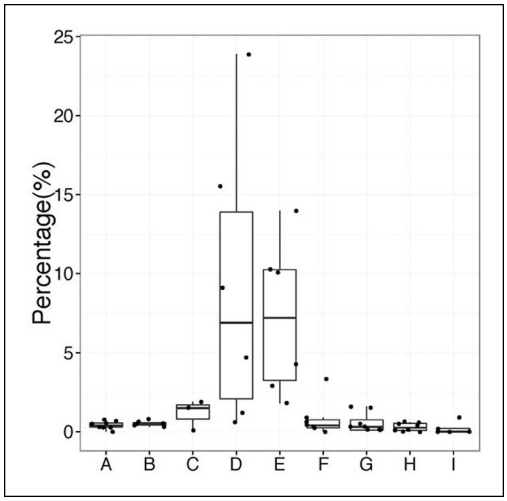

Figure 7—figure supplement 2. Percentage of sequences containing > = 4 nucleotides changes for each B cell group. A, sB-1a (2-7d); B, sB-1a (2W-1M); C, sB-1a (2M); D, sB-1a (4-6M); E, sB-1a (GF, 4M); F, B cells (AIDKO, 4-5M); G, pB-1a (2-6M); H, FOB, pB-2 (2-5M); I, MZB (1-5M). Each dot represents the data for an individual B cell sample, n = 3-8.

DOI: 10.7554/eLife.09083.022

Data for IgH library using old primers

```
>M01518:44:000000000-AGWM4:1:2106:7311:19715 COPY:242 mIGHV1-55*01[0] mIGHD4-1*02 mIGHJ1*03[2] mIGHM*03 ARKELGRAHWCFDV
 R  H  G  Q  G  P  E  W  I  G  D  I  Y  P  G  S  G  T  T  N  Y  N  E  K  F  K  S  K  A  T  L  T  V  D  T  S  S  R  T  V
          .        .        .        .        †        .        .        .        .        ‡        .        .
aggcatggacaaggccctgagtggattggagatatTTATCCTGGTAGTGGTACTACTAACTACAATGAGAAGTTCAAGAGCAAGGCCACACTGACTGTAGACACATCCTCCAGAACAGTC
AGGCCTGGACAAGGCCTTGAGTGGATTGGAGATATTTATCCTGGTAGTGGTAGTACTAACTACAATGAGAAGTTCAAGAGCAAGGCCACACTGACTGTAGACACATCCTCCAGCACAGCC

 Y  M  Q  L  S  S  L  T  S  E  D  S  A  V  Y  Y  C  A  R  K  E  L  G  R  A  H  W  C  F  D  V  W  G  T  G  T  T  V  T  V
          .        .        †        .        .        .        .        ‡        .        .        .        .        .
TACATGCAGCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTGCAAGAAAGGAACTGGGGCGGGCTCACTGGTGCTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTC
TACATGCAGCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTGCAAGA    AACTGGG        ACTGGTACTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTC
```

Data for IgH library using new primers

```
>M01518:44:000000000-AGWM4:1:1114:8333:15960 COPY:162 mIGHV1-55*01[0] mIGHD4-1*02 mIGHJ1*03[2] mIGHM*03 ARKELGRAHWCFDV
 R  H  G  Q  G  P  E  W  I  G  D  I  Y  P  G  S  G  T  T  N  Y  N  E  K  F  K  S  K  A  T  L  T  V  D  T  S  S  R  T  V
          .        .        .        .        †        .        .        .        .        ‡        .        .
aggcatggacaaggccctgagtggattggagataTTTATCCTGGTAGTGGTACTACTAACTACAATGAGAAGTTCAAGAGCAAGGCCACACTGACTGTAGACACATCCTCCAGAACAGTC
AGGCCTGGACAAGGCCTTGAGTGGATTGGAGATATTTATCCTGGTAGTGGTAGTACTAACTACAATGAGAAGTTCAAGAGCAAGGCCACACTGACTGTAGACACATCCTCCAGCACAGCC

 Y  M  Q  L  S  S  L  T  S  E  D  S  A  V  Y  Y  C  A  R  K  E  L  G  R  A  H  W  C  F  D  V  W  G  T  G  T  T  V  T  V
          .        .        †        .        .        .        .        ‡        .        .        .        .        .
TACATGCAGCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTGCAAGAAAGGAACTGGGGCGGGCTCACTGGTGCTTCGATGTCTGgggcacagggaccacggtcaccgtc
TACATGCAGCTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTGCAAGA    AACTGGG        ACTGGTACTTCGATGTCTGGGGCACAGGGACCACGGTCACCGTC
```

Figure 7—figure supplement 3. Identical V(D)J recombination sequences containing identical mutated nucleotides are detected in sequence data sets for IgH libraries obtained by using either old or new primer set. We sorted two splenic B-1a populations individually from two 4 month old C57BL/6J mice. We extracted RNA from each population and divided each RNA sample into two parts. For one part, we prepared an amplified library using the old primer set; and for the other, we prepared an amplified library using the new primer set. We then sequenced two pair of amplified IgH libraries. In two separate comparisons, we detected identical IgH sequences containing identical nucleotides substitutions in each library. One example is shown from comparing one pair of sequence data sets. Red nucleotides are the mutated bases. Upper line of sequence is the obtained sequence reads and the lower line of sequences is the V, D and J reference sequences.

DOI: 10.7554/eLife.09083.023

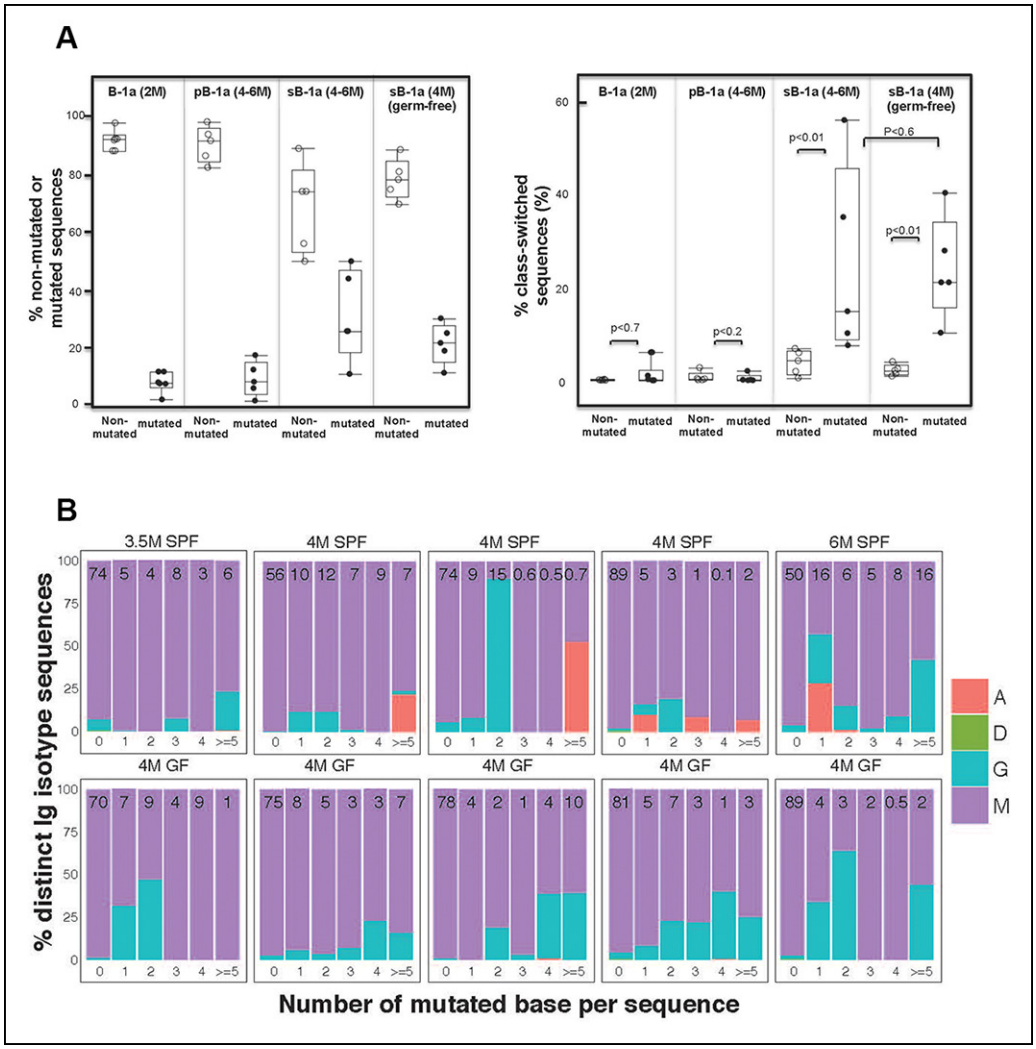

Figure 8. Progressive increase in the splenic B-1a IgV_H mutation frequency with age is accompanied by increased class-switching. (**A**) *Left panel:* The frequencies of non-mutated or mutated (> = 1 nucleotide substitution) IgH sequences obtained from indicated B cell samples are shown; *Right panel:* The frequencies of sequences expressing class-switched isotypes (neither IgM nor IgD) among non-mutated or mutated sequences are shown. Each dot represents data from an individual sample (n = 5–6). p values are calculated based on the Nonparametric Wilcoxon test. (**B**) In each plot, the IgH sequences obtained from each splenic B-1a sample from 3.5–6 month old mice are divided into five categories, based on the number of mutated nucleotides (0, 1, 2, 3, 4, > = 5) per read. In each plot, the values shown at the top are the frequencies of sequences in each category. For each category of sequences, frequencies of the distinct isotype sequences are shown as stacks. A = IgA; D = IgD; G = IgG1 + IgG3 + IgG2c + IgG2b. Each plot represents the data for a splenic B-1a sample from an individual mouse reared under either specific pathogen free (SPF) (*upper plots*) or germ-free (GF) (*lower plots*) conditions. The data for germ-free (GF) animals is discussed at the end of the Result section.

DOI: 10.7554/eLife.09083.024

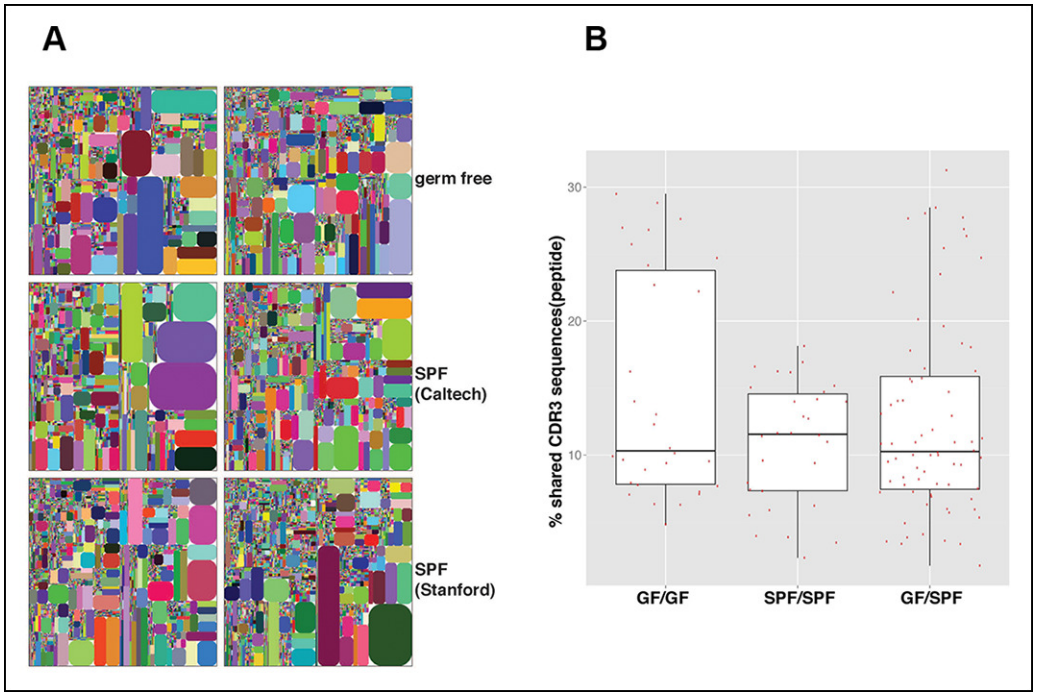

Figure 9. The B-1a IgH repertoires from mice raised in specific pathogen free condition are comparable to the B-1a IgH repertoire from age-matched germ-free mice. (**A**) IgH CDR3 tree map plots for splenic B-1a cells from GF mice (*upper panel*), or SPF mice in Caltech animal facility (*middle panel*), or SPF mice in Stanford animal facility (*bottom panel*). Each plot represents the data for a sample from a 4 month old mouse. Recurrent CDR3 (nucleotide) sequences are visualized as larger contiguously-colored rectangles in each plot. (**B**) CDR3 peptide pair-wise sharing analysis of IgH repertoire similarity between multiple splenic B-1a samples from age-matched GF and SPF mice. GF mice (n = 6); SPF mice (n = 6). CDR3 peptide pair-wise analysis was conducted between GF mice (GF/GF), SPF mice (SPF/SPF) and GF vs. SPF mice (GF/SPF). Each dot represents the percentage of shared CDR3 peptide sequences between two mice. There was no statistical difference between each comparison.

DOI: 10.7554/eLife.09083.027

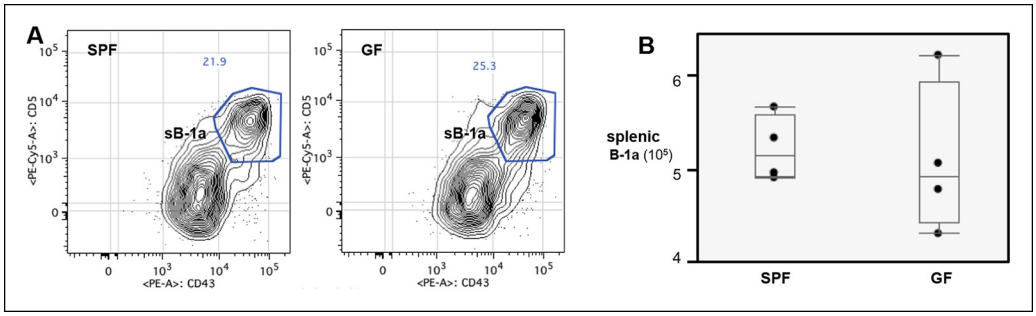

Figure 9—figure supplement 1. Normal splenic B-1a compartment in GF mice. (**A**) FACS plot showing the B-1a population in spleen from SPF or GF mouse. Live dump^{-} CD19^{+} CD93^{-} IgM^{hi} IgD^{lo} CD23$^{lo/-}$ CD21^{-} cells were gated to reveal CD5^{+} CD43^{+} B-1a cells. (**B**) Absolute number of splenic B-1a cells in GF and SPF mice. Each dot represents the data from an individual mouse. There is no significant difference shown between two groups.
DOI: 10.7554/eLife.09083.028