Receptor editing constrains development of phosphatidyl choline-specific B cells in V_{H}12-transgenic mice

Alexandra N. Worth\(^1\), Victoria L. Palmer\(^1\), N. Max Schabla\(^1,2\), Greg A. Perry\(^1\), Anna N. Fraser-Philbin\(^1\), Patrick C. Swanson\(^1,3,\ast\)

\(^1\)Department of Medical Microbiology and Immunology, Creighton University, 2500 California Plaza, Omaha, NE 68178, USA

\(^2\)Shoreline Biosciences, San Diego, CA 92121, USA

\(^3\)Lead contact

SUMMARY

B1 B cells reactive to phosphatidyl choline (PtC) exhibit restricted immunoglobulin heavy chain (HC) and light chain (LC) combinations, exemplified by V_{H}12/V_{\kappa}4/5H. Two checkpoints are thought to focus PtC\(^{+}\) B cell maturation in V_{H}12-transgenic mice (VH12 mice): V-J rearrangements encoding a “permissive” LC capable of V_{H}12 HC pairing are selected first, followed by positive selection based on PtC binding, often requiring LC receptor editing to salvage PtC\(^{-}\) B cells and acquire PtC reactivity. However, evidence obtained from breeding VH12 mice to editing-defective dnRAG1 mice and analyzing LC sequences from PtC\(^{+}\) and PtC\(^{-}\) B cell subsets instead suggests that receptor editing functions after initial positive selection to remove PtC\(^{+}\) B cells in VH12 mice. This offers a mechanism to constrain natural, polyreactive B cells to limit their frequency. Sequencing also reveals occasional in-frame hybrid LC genes, reminiscent of type 2 gene replacement, that, testing suggests, arise via a recombination-activating gene (RAG)-independent mechanism.

Graphical abstract

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Correspondence: pswanson@creighton.edu.

AUTHOR CONTRIBUTIONS
Conceptualization, A.N.W. and P.C.S.; methodology, A.N.W., N.M.S., V.L.P., G.A.P., and P.C.S.; software, A.N.W.; investigation, A.N.W., N.M.S., V.L.P., G.A.P., A.N.F.-P., and P.C.S.; writing – original draft, A.N.W. and P.C.S.; writing – review & editing, A.N.W., N.M.S., V.L.P., A.N.F.-P., G.A.P., and P.C.S.; visualization, A.N.W. and P.C.S.; funding acquisition, P.C.S.; supervision, P.C.S.

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.110899.

DECLARATION OF INTERESTS
The authors declare no competing interests.
In brief

Worth et al. show that impairing receptor editing by dnRAG1 transgene expression in HC VH12 transgenic mice enforces BCR reactivity to phosphatidyl choline (PtC). Sequencing suggests that, after positive selection based on HC pairing and PtC binding, receptor editing further reduces PtC+ B cells. Likely RAG-independent type 2 gene replacements are also identified.

INTRODUCTION

A central paradigm in B cell immunobiology is the existence of two major B cell subsets, B-1 and B-2, that are distinguished by their ontological, phenotypical, and functional properties as well as anatomical localization (Hardy, 2006; Kristiansen et al., 2018). B-1 B cells arise early in ontogeny, mainly populate the peritoneal and pleural cavities in mice, and are poorly reconstituted from adult hematopoietic progenitor cells. The largest B-1 B cell subset, B1a, is characterized by a surface immunoglobulin M (IgM)hi and CD5+ immunophenotype, natural antibody production that requires little or no stimulation to induce, innate-like immune functions, and B cell receptor (BCR) specificities that are enriched toward membrane-associated self- and microbial antigens. Antigenic specificities common to this B cell subset include single-stranded DNA (Casali et al., 1987), rheumatoid factor (Hardy et al., 1987), the membrane phospholipid phosphatidyl choline (PtC) (Hayakawa et al., 1984; Mercolino et al., 1988), thymocyte glycoprotein (Hayakawa et al., 2003; Hayakawa et al., 1984), and bacterial lipids and carbohydrates (Forster and Rajewsky, 2003).
1987; Masmoudi et al., 1990). Although evidence suggests that B1 B cells can develop from a dedicated lineage progenitor (Montecino-Rodriguez et al., 2006), other studies show that B cells can be driven to acquire B1 phenotypic and functional properties when BCR specificity is enforced toward antigens commonly recognized by this B cell subset (Graf et al., 2019).

PtC reactivity in murine B cells is conferred by a limited set of heavy chain (HC) and light chain (LC) genes, typically either V\textsubscript{H}11/V\textsubscript{\kappa}9-128 or V\textsubscript{H}12/V\textsubscript{\kappa}4/5H (IGKV14-126 or IGKV4-91, respectively, in the ImMunoGeneTics [IMGT] database) (Mercolino et al., 1989; Pennell et al., 1989). However, PtC reactivity should be viewed as encompassing specificity toward a broader range of phosphorylcholine-containing and oxidation-specific epitopes (OSEs) (Binder et al., 2003; Shaw et al., 2000). Natural antibodies harboring these specificities are thought to play key roles in promoting tissue homeostasis (e.g., clearance of apoptotic cells) and host defense against common microbial pathogens as well as reducing immune responses propagated by OSEs that, when unchecked, can promote chronic inflammation (e.g., atherosclerosis) (Binder et al., 2016). Thus, there is considerable interest in understanding the selection pressures that operate to sustain and restrain PtC\textsuperscript{+} cell populations to support their protective immune functions while limiting the potential risk of autoimmune.

V\textsubscript{H}12-transgenic mice (VH12 mice) (Arnold et al., 1994) have provided key insights into selection pressures occurring at various checkpoints that regulate development of PtC\textsuperscript{+} B cells (reviewed in Wang and Clarke, 2004). One checkpoint occurs as pre-B cells transition to the immature stage, when V\textsubscript{H}12-expressing B cells are selected based on HC pairing with one of a limited number of permissive LCs with which V\textsubscript{H}12 can functionally associate (Tatu et al., 1999). A second checkpoint has been proposed to occur as immature V\textsubscript{H}12-expressing B cells migrate to the periphery and complete maturation. This process is thought to require receptor editing to obtain and select the V\textsubscript{\kappa}4/5H LC that confers PtC binding when paired with the V\textsubscript{H}12 HC (Tatu and Clarke, 2000; Tatu et al., 1999). This conclusion is based on two observations. First, in VH12 mice, most PtC\textsuperscript{+} B cells expressing V\textsubscript{\kappa}4/5H show J\textsubscript{\kappa} gene usage skewed against J\textsubscript{\kappa}1 (Tatu et al., 1999). Second, in VH12 mice expressing a permissive but non-PtC-reactive, targeted V\textsubscript{\kappa}1A LC transgene, most mature B cells have undergone receptor editing to replace V\textsubscript{\kappa}1A with V\textsubscript{\kappa}4/5H (Tatu and Clarke, 2000).

If receptor editing is indeed required for PtC\textsuperscript{+} B cell selection into the mature B cell compartment in VH12 mice, then impairing this process should reduce PtC\textsuperscript{+} B cell frequency. Testing this hypothesis in the context of the endogenous LC repertoire necessitates preferentially blocking V(D)J recombination late in B cell development so primary V(D)J rearrangements can occur normally. We previously developed transgenic mice expressing catalytically inactive RAG1 (recombination-activating gene 1) in the periphery (dnRAG1 mice) that show expansion of B1-like CD5\textsuperscript{+} B cells in the spleen and evidence of impaired receptor editing, including skewing of J\kappa gene usage to J\kappa1 (Hassaballa et al., 2011). Interestingly, dnRAG1 mice also show an increased frequency of PtC\textsuperscript{+} B cells (Palmer et al., 2018). The latter observation is inconsistent with the above prediction based on studies of VH12 mice and led us to question the extent to which receptor editing is required for PtC\textsuperscript{+} B cell development and selection in VH12 mice.
To address this issue, we bred the dnRAG1 transgene onto the VH12 background and directly compared mice for PtC+ B cell frequency and LC gene usage in PtC+ and PtC− B cells. Interestingly, we found that dnRAG1 expression in VH12 mice substantially increased the frequency of splenic PtC+ B cells compared with both of the single-transgenic mice. When the expressed Ig LC gene repertoires were compared using next-generation sequencing (NGS), we found that WT and dnRAG1 mice had LC repertoires distinct from VH12 mice. dnRAG1 expression in VH12 mice had little effect on the LC repertoire in VH12 mice overall but increased the frequency of IGKV4-91 joining to IGKJ5 (Jκ5).

We also detected examples of chimeric IGKV genes in which the 5′ end of one IGKV gene is replaced by another in frame and with perfect joining, reminiscent of type 2 \( V_H \) gene replacement (Darlow and Stott, 2005). Experiments testing the origin of these hybrid sequences suggest that they may arise through a PCR artifact, but other data lend some support to a RAG-mediated process. These data lead us to conclude that receptor editing, rather than promoting positive selection of PtC+ B cells in VH12 mice, as originally proposed, instead reduces their frequency when \( V_H \)12-expressing B cells paired with a permissive, PtC-reactive LC are positively selected. This finding provides a mechanism to restrain excessive development of polyreactive B cells.

RESULTS
dnRAG1 transgene expression in VH12 mice promotes PtC+ B cell development

Hemizygous dnRAG1 and VH12 transgenic mice were interbred to yield cohorts of wild-type (WT), dnRAG1, VH12, and dnRAG1 VH12 mice for analysis. Bone marrow, spleen, and peritoneal lavage cells were characterized to compare cellularity, developmental and mature B cell subsets, and PtC-specific B cells using flow cytometry following previous studies (Allman et al., 2001; Borghesi et al., 2005; Hardy et al., 1991; Palmer et al., 2015, 2018). Total B cells (identified as CD19+) in VH12 mice were significantly lower than in WT and dnRAG1 mice in the bone marrow but were only lower relative to WT mice in the spleen (Figure 1A; Table S1). This observation is consistent with earlier studies showing that VH12 mice have ~50% fewer bone marrow B220+ B cells compared with non-transgenic littermates (Arnold et al., 1994), with significant loss occurring after the B220+CD43+ pro-B cell stage (Figure S1A; Table S1), whereas dnRAG1 mice show no significant abnormalities early in B cell development but lose B cells at the transitional T1 to T2 stage in the spleen (Hassaballa et al., 2011; Figure 1B; Table S1). VH12 mice also had very few IgD+ B cells (Figures S1B and S1C; Table S1), consistent with strong repression of endogenous HC gene expression (Arnold et al., 1994). In line with this finding, the few splenic IgD+ B cells in VH12 mice almost all express the IgMb (endogenous) allotype (Figure S1C; Table S1).

Interestingly, dnRAG1 VH12 mice had significantly higher total B cells in the spleen compared with VH12 mice and were not statistically different from WT or dnRAG1 mice (Figure 1A; Table S1). However, in bone marrow, total B cells, as well as pre-B and immature B cell subsets, remained significantly below their dnRAG1 counterparts (Figures 1A and S1A; Table S1). The elevated number of splenic B cells in dnRAG1 VH12 mice is due to a significant increase in B1-like B cells expressing CD5 (Figure 1B; Table S1). In dnRAG1 and VH12 mice, slightly more than half of the splenic B cells are CD5+, whereas...
in dnRAG1 VH12 mice, it is ~90% (Figure 1B; Table S1). In dnRAG1 VH12 mice, we also observed a marked reduction of transitional T2 B cells with a mild loss of follicular B cells but no change in marginal zone B cells relative to VH12 mice (Figure 1B; Table S1). Slight decreases in IgD+ B cell frequency, the ratio of IgD+ B cells/total B cells, and the fraction of B cells expressing the IgMb allotype in the spleens of dnRAG1 VH12 mice compared with VH12 mice were also noted (Figure S1C; Table S1). These findings lead us to conclude that dnRAG1 expression in VH12 mice enforces allelic exclusion. Peritoneal cell populations were not significantly different between the four mouse strains (Table S1).

To test whether CD5+ B cells in dnRAG1 VH12 mice retain PtC binding activity, we stained cells with PtC-containing fluorescent liposomes (Figure 1C). Consistent with our previous results (Palmer et al., 2018), compared with WT mice, which have few splenic PtC+ or CD5+ B cells, dnRAG1 mice show a small population of PtC+ B cells, most of which are CD5+ (Figure 1C), and a larger population of CD5+PtC− cells. In contrast, splenic B cells in VH12 mice are roughly evenly divided into CD5+PtC+ and CD5+PtC− populations (Figure 1C). Remarkably, almost all splenic B cells in dnRAG1 VH12 mice are CD5+PtC+ (Figure 1C) and are significantly increased compared with dnRAG1 and VH12 mice. Thus, we conclude from these data that dnRAG1 expression in VH12 mice enforces PtC binding and CD5 expression in splenic B cells. In the peritoneal cavity, CD5+PtC+ B cells constitute ~10% of the B cells in WT mice, which expands to ~20% in dnRAG1 mice (Figure 1C; Table S1). In contrast, in VH12 and dnRAG1 VH12 mice, almost all B cells are CD5+PtC+ and are not significantly different in absolute number between these genotypes (Figure 1C; Table S1).

The mean fluorescence intensity (MFI) of PtC and CD5 staining on gated PtC+ B cells varies between the mouse strains (Figures 1C and 1D), with significantly higher PtC staining detected in dnRAG1 mice compared with the other strains (Figure 1D). A similar trend was observed in the peritoneal cavity (Figure 1D). This may in part be due to differences in BCR density because the MFI of IgM staining is about 2-fold higher on splenic CD5+ B cells in dnRAG1 mice compared with VH12 and dnRAG1 VH12 mice (Figure S1D). Staining with PtC-liposomes impairs IgM staining on CD5+ B cells from VH12 mice, presumably because of steric interference, so these values were determined without PtC-liposome co-staining (Figure S1D). The MFI for CD5 in the spleen was significantly increased in dnRAG1 VH12 mice compared with VH12 mice, but in the peritoneal cavity, the opposite was true (Figure 1D).

**Evidence of distinctive and biased LC repertoires and V-J pairing in PtC+ B cells from dnRAG1 and VH12 mice**

To test the idea that receptor editing normally impedes CD5+PtC+ B cell maturation, we first compared the expressed LC repertoire in populations of sorted splenic B cells that were CD5+ or CD5− and either PtC+ or PtC− from WT, dnRAG1, VH12, and dnRAG1 VH12 mice (Figure 2). Subsequently, the LC repertoires in these B cell subsets isolated from the bone marrow and spleen of VH12 and dnRAG1 VH12 mice were directly compared (Figure S2). V_{H12} HC transgene expression and sustained allelic exclusion limiting HC gene expression from the endogenous *Igh* locus in VH12 and dnRAG1 VH12 mice ensures
that all LCs detected in these genotypes are likely to functionally pair with the \( V_H^{12} \) HC (Arnold et al., 1994; Figures S1B and S1C).

The cell staining and sorting strategy and the pipeline for sequencing and analyzing the expressed LC gene repertoire are summarized in Figure S3. The approach relies on IGKV and IGLV family-specific primers and LC constant region primers to prepare double-stranded cDNA from sorted B cell subsets and multiplexing primers to amplify them. These primers also contain sequencing adapter primer sites and index and unique molecular identifier (UMI) sequences to assign sequences to specific mouse B cell populations and genotypes as well as to compress redundant sequences (Table S2). Three biological replicates were performed for each set of experiments. Summaries of total input sequences and sequences passing quality and assembly thresholds, along with average numbers of reads per UMI sequence, are provided in Table S3. Only sequences with 2 or more reads per UMI were analyzed. Graphic representations of V and J gene usage for each B cell population and genotype are shown in Figure 2 and are discussed further below.

Because few B cells are CD5\(^+\) and/or PtC\(^+\) in WT mice, we only analyzed CD5\(^-\)PtC\(^-\) B cells from these animals. The expressed LC repertoire in this population was found to be highly diverse and broadly, but not uniformly, distributed with respect to V gene position within the \( Igk \) and \( Igl \) loci (Figure 2A). Overall, IGKJ gene usage showed similar representation of KJ1, KJ2, and KJ5 sequences, with a slightly lower frequency for KJ4 (Figure 2B). The biased usage of certain IGKV genes in WT B cells is consistent with data reported previously (Aoki-Ota et al., 2012). Specifically, many of the same IGKV genes, particularly at the 5\(^{\prime}\) end of the \( Igk \) locus (e.g., 1–135, 9–120, 10–96, and 19–93) were over represented in both studies; IGKJ4 usage was also underrepresented in both cases (Figure 2; Aoki-Ota et al., 2012).

For dnRAG1 mice, the splenic CD5\(^-\)PtC\(^-\) B cell population showed a pattern of IGKV and IGKJ gene usage similar to WT CD5\(^-\)PtC\(^-\) B cells (Figures 2A and 2B). In contrast, the CD5\(^+\)PtC\(^-\) population showed evidence of repertoire skewing, with IGKV3-5 and IGLV1/2 genes overrepresented relative to the WT LC repertoire (Figure 2A). The increased frequency of Ig\(\lambda^+\) B cells in the CD5\(^+\) subset in dnRAG1 mice was confirmed by flow cytometry (Figure S4A; Table S1). As we noted previously (Hassaballa et al., 2011), this population is biased toward KJ1 and KJ2 usage, with fewer IGKV sequences joined to KJ4 and KJ5. Interestingly, the CD5\(^+\)PtC\(^+\) population showed a strong preference for IGKV14-126 (V\(\kappa\)9-128) usage (Figure 2A), which pairs with the \( V_H^{11} \) HC to confer PtC reactivity (Mercolino et al., 1989; Pennell et al., 1989), as well as IGLV1/2 and several other IGKV genes, including KV3-7, KV6-17, and KV4-91. Notably, KV3-7 (V\(\kappa\)21E) and KV4-91 are known to support PtC binding when paired with \( V_H^{11} \) (Seidl et al., 1997) and \( V_H^{12} \) (Mercolino et al., 1989; Pennell et al., 1989), respectively. KV6-17 has not been reported previously to confer PtC reactivity.

For VH12 mice, CD5\(^+\)PtC\(^+\) B cells showed strong bias toward KV4-91 usage, accounting for ~80% of the total IGKV genes identified in this population, followed by KV3-7 at ~5% (Figure 2A). Consistent with previous studies (Tatu et al., 1999), IGKJ joining by KV4-91 was heavily skewed against KJ1 (Figure 2B). However, this was true regardless
of mouse genotype and population, indicating that KJ1 joining is globally disfavored for functional KV4-91 pairing with HCs. In contrast, KV3-7 joining was biased toward KJ1. As shown previously (Tatu et al., 1999), some CD5−PtC− B cells also use KV4-91, although other genes are used more frequently, particularly KV5-43, KV4-86, KV1-110, and IGLV1. Notably, Vx1A (KV1-110) has been shown previously to support permissive pairing to VH12 but does not confer PtC reactivity (Tatu and Clarke, 2000). There is no consistent pattern of IGKJ gene usage among these IGKV genes; some, like KV4-86, show bias against KJ1 usage, whereas others, like KV1-110 show bias toward KJ1 usage (Figure 2B).

The expression of dnRAG1 in VH12 mice had very little effect on the frequency of KV4-91 gene usage in CD5+PtC+ B cells compared with VH12 mice, and the small difference was not reproduced between the two datasets (Figures 2A and S2A). This was also true for most other KV genes highly represented in CD5+PtC+ B cells as well as for frequently used KV genes in CD5−PtC− B cells. Interestingly, however, the global frequency of KJ5 usage and KV4-91 linkage to KJ5 reproducibly increased in this population relative to VH12 mice (Figures 2B and S2B). Also notable is the observation that usage of IGLV and IGKV14-126 genes prevalent in CD5+ B cells from dnRAG1 mice was reduced substantially in dnRAG1 VH12 mice.

Because previous evidence suggests that LC selection in VH12 mice occurs at the immature and transitional stages of B cell development (Tatu et al., 1999; Wang and Clarke, 2003, 2004; Wang et al., 2007), we directly compared the LC repertoires in bone marrow and spleen populations using AA4.1 staining to differentiate between immature (AA4.1+) and mature (AA4.1−) PtC+ B cells in the bone marrow (Figure S2A). A small population of AA4.1+CD5+PtC+ B cells can be detected in the bone marrow of VH12 mice, whereas splenic CD5+PtC+ B cells in these animals are almost all AA4.1− (Figures S4C and S4D). Interestingly, bone marrow AA4.1+ populations in VH12 and dnRAG1 VH12 mice showed high levels of IGLV1 usage as well as distal KV1-135 and KV1-110 usage in CD5+PtC− and CD5+PtC+ and populations (Figure S2A). Although the AA4.1+CD5+PtC+ B cells showed biased usage of KV4-91 compared with the AA4.1+CD5−PtC− subset, KV4-91 represented only ~10% of the V gene repertoire in this population compared with ~80% in bone marrow AA4.1−CD5+PtC+ and spleen CD5+PtC+ B cell subsets. This result is consistent with previous data suggesting that selection pressures focus usage of KV4-91 during B cell maturation in VH12 mice (Tatu et al., 1999). As with V gene usage, the PtC− and PtC+ subsets among bone marrow AA4.1+ B cells showed much less bias in overall IGKJ usage compared with their AA4.1− counterparts (Figure S2B).

To obtain a more quantitative and unbiased view of the differences in LC repertoires detected by visual inspection, we determined the repertoire dissimilarity index (RDI) (Bolen et al., 2017) for each aggregate B cell population and genotype in the spleen and bone marrow (Figure 3). For individual samples, hierarchical clustering by RDI showed that the LC repertoires for a given B cell population and genotype generally clustered closely together, demonstrating that the replicate experiments are reproducible to justify data aggregation (Figure S5). Hierarchical clustering by RDI and the log fold change between repertoires showed that the LC repertoires in splenic CD5+PtC+ B cells from VH12 and dnRAG1 VH12 mice cluster together (Figures 3A–3D). The LC repertoires in CD5−PtC− B
cells from these animals resembled one another more closely than their counterparts in WT and dnRAG1 mice, which showed a less than 2 log-fold difference. However, CD5<sup>+</sup>PtC<sup>+</sup> B cell LC repertoires in dnRAG1 mice were more closely related to their CD5<sup>+</sup>PtC<sup>−</sup> counterparts than those in CD5<sup>+</sup>PtC<sup>+</sup> B cells from VH12 and dnRAG1 VH12 mice (Figures 3A and 3B). Interestingly, LC repertoires in AA4.1<sup>+</sup> bone marrow CD5<sup>+</sup>PtC<sup>+</sup> B cells in VH12 and dnRAG1 VH12 mice were more similar to their CD5<sup>+</sup>PtC<sup>−</sup> counterparts in bone marrow and spleen (3–5 log-fold difference) than those in CD5<sup>+</sup>PtC<sup>+</sup> B cells from the AA4.1<sup>−</sup> bone marrow fraction or the spleen (7–8 log-fold difference) (Figures 3C and 3D).

Evidence of selection of CDR3 length and composition in PtC-reactive B cells

Previous work has provided evidence that LCs from PtC<sup>+</sup> B cells in VH12 mice exhibit bias in CDR3 length and composition (Tatu et al., 1999), but the conclusions were based on a small number of cloned sequences. Our NGS experiments provided an opportunity to analyze a much larger dataset to determine whether such biases are generalizable. Toward this end, we analyzed three KV genes associated with PtC reactivity (KV4-91, KV3-7, and KV14-126) and compared their CDR3 length distribution and composition with those found in CD5<sup>−</sup>PtC<sup>−</sup> B cells within and among the mouse genotypes (Figures 4A and 4B). We found that, for IGKV4-91, the CDR3 length distribution is significantly more restricted in CD5<sup>+</sup>PtC<sup>+</sup> B cells than in their CD5<sup>−</sup>PtC<sup>−</sup> counterparts in VH12 and dnRAG1 VH12 mice but not dnRAG1 mice. In contrast, CDR3 length distributions for KV3-7 and KV14-126 were not significantly different between PtC<sup>+</sup> and PtC<sup>−</sup> populations within a given genotype. When a KV gene used predominantly in CD5<sup>−</sup>PtC<sup>−</sup> B cells in VH12 mice (KV4-86) was similarly analyzed, CDR3 length was found to be more highly restricted than in its CD5<sup>+</sup>PtC<sup>+</sup> counterpart (Figure 4B). Hence, these data suggest that KV genes used frequently in a given B cell population tend to exhibit restricted CDR3 lengths.

When CDR3 composition was analyzed at the most abundant length (9 amino acids) (Figure 4B), stronger selection biases were generally observed for LCs commonly used in CD5<sup>+</sup>PtC<sup>+</sup> B cells compared with their CD5<sup>−</sup>PtC<sup>−</sup> counterparts. For example, for KV4-91 sequences from VH12 mice, Leu is generally favored over other residues at position 8. A clearer case is seen for KV3-7. In CD5<sup>+</sup>PtC<sup>+</sup> B cells from dnRAG1 mice, His and Trp are clearly favored at positions 2 and 4; these are also favored in VH12 mice, but unlike in dnRAG1 mice, Gly is highly selected at position 8. Interestingly, the CDR3 selection biases noted for KV4-91 and KV3-7 in the splenic B cell populations from VH12 mice were also observed in the AA4.1<sup>−</sup> subsets in the bone marrow but were much less apparent in the AA4.1<sup>+</sup> populations (Figure 4C). For KV14-126 sequences, selection biases, although evident, are not as striking as for KV4-91 and KV3-7 sequences. When KV4-86 sequences were analyzed, amino acid selection appeared to be more stringent in CD5<sup>+</sup>PtC<sup>+</sup> B cells from dnRAG1 mice compared with CD5<sup>−</sup>PtC<sup>−</sup> B cells, but the opposite was true for VH12 and dnRAG1 VH12 mice. However, closer inspection reveals evidence of selection for a sequence motif observed in KV4-91 sequences. As discussed below, this may be attributed to hybrid KV4-86 genes containing CDR3 sequences from the KV4-91 gene.
Evidence of KV gene replacement in VH12 mice

Because the differential use of KV4-91 between CD5+PtC+ and CD5–PtC– B cells in VH12 and dnRAG1 VH12 mice was so striking, we compared V-J junctional sequences between the two populations (Figure 5A). Interestingly, we found that, compared with CD5+PtC+ B cells, KV4-91 sequences from VH12 CD5–PtC– B cells had significantly longer stretches of non-templated or palindromic (NP) nucleotides and an unexpectedly high frequency of long NP additions (NP > 5; 3 of 90 for Experiment 1 [EXP1], and 3 of 208 for EXP2; range, 9–17 nt) associated with a nearly equivalent number of nucleotides deleted at the 3′ end of KV4-91 (see example in Figures 5B and 5D).

The unusual V-J junctions detected in these sequences led us to wonder whether the 3′ end of KV4-91 was replaced by another V gene, by analogy to Ig gene conversion (Lanning and Knight, 2015) or VH gene replacement (Darlow and Stott, 2005). Strikingly, the long sequences identified as insertions showed strong similarity to KV genes 3′ to KV4-91 in the Igk locus, the most frequent of which was KV4-86 (Figure 5A; see Figures 5B and 5C for examples). These events were also detected among KV4-91 sequences from VH12 CD5+PtC+ B cells but at a 3- to 4-fold lower frequency (Figure 5A). Although KV4-91 chimeras with KV4-86 were most commonly observed, KV4-91 sequences fused to IGKV3-7 were almost as frequent (Figure 5A; see Figures 5D and 5E for examples). Because IGKV3-7 confers PtC reactivity when paired with V_H11 (Seidl et al., 1997) and V_H12 (this work), this finding suggested that these hybrid sequences were positively selected. Similar trends were observed in dnRAG1 and dnRAG1 VH12 mice; that is, KV4-91 chimeric sequences were detected in 1%–3% of analyzed sequences in CD5–PtC– B cells and at slightly lower frequencies in CD5+PtC+ B cells. Interestingly, both dnRAG1 VH12 B cell subsets showed similar to lower levels of these events compared with their VH12 counterparts (Figure 5A).

We considered the possibility that this type of event represents an artifact introduced during PCR or sequencing (Brakenhoff et al., 1991; Paabo et al., 1990). Several lines of evidence argued against this possibility. First, reciprocal hybrid sequences containing the 5′ end of KV4-86 and the 3′ end of KV4-91 were not detected at high frequency from CD5+PtC– B cells, as might have been expected from a PCR artifact. Rather, most gene replacement events for KV4-86 involved genes 3′ of KV4-86 in the Igk locus, with KV4-59 being the most frequent hybrid partner (Figure S6A). In contrast, in the CD5+PtC+ B cell subset, the most common KV4-86 hybrid partner was KV4-91, again suggesting that the CDR3 region of KV4-91 may be positively selected for PtC reactivity in these cases. Second, we identified several examples of 2–3 independent IGKV hybrid gene sequences that shared the same donor sequence and were identical or differed by only 1 nt. Importantly, these were identified within and between animals of the same genotype (see Figure S6B for an example). The likelihood that these identical (or nearly so) sequences arose by chance seemed implausible. Third, the hybrid sequences described here are reminiscent of type 2 V_H gene replacement events (Darlow and Stott, 2005) detected in human peripheral and malignant B cells (Bellan et al., 2002; Itoh et al., 2000; Lenze et al., 2003; Wilson et al., 2000), in which the 5′ end of a rearranged V_H gene is replaced in frame without altering the nucleotide sequence at the crossover point. The presence of cryptic heptamer-like motifs as
well as DNA double-strand breaks (Itoh et al., 2000), at or near the hybrid junction provides evidence implicating RAGs in mediating type 2 V<sub>H</sub> gene replacement.

Applying the type 2 V<sub>H</sub> gene replacement mechanism to explain the origin of chimeric KV gene sequences described here (e.g., Figures 5C and 5E) implies that the 5<sup>′</sup> end of the hybrid gene (i.e., KV4-91) was originally not rearranged and that the 3<sup>′</sup> end had undergone prior rearrangement (i.e., KV4-86 or KV3-7). That RAGs might mediate KV gene replacement in this scenario seemed plausible for several reasons. First, the RAG proteins are re-expressed during LC receptor editing, so KV gene replacement events may coincide with this process. Second, the frequency of KV replacement is reduced in dnRAG1 VH12 mice in some cases (e.g., KV4-86; Figure S6A), which is consistent with dnRAG1-mediated interference of endogenous RAG1 activity. This contention is also supported by the evidence described above that dnRAG1 expression in VH12 mice enforces HC allelic exclusion (Table S1). Third, a recent study reported that the heptamer in the 3<sup>′</sup> 12-RSS flanking murine KV genes can serve as an antisense 23-RSS and be cleaved at low frequency by the RAGs (Shinoda et al., 2019), providing precedence for off-target RAG activity in the Igk locus.

Aligning KV genes identified as common partners in KV4-91 chimeric sequences (KV4-86, KV4-59, and KV3-7) as well as two other LC genes identified as commonly expressed in PtC− B cells in VH12 mice but uncommon KV4-91 fusion partners (KV1-110 and KV5-43) revealed that all sequences have conserved FR3 regions harboring conserved tandem, overlapping, cryptic heptamer-like motifs in the antisense orientation (Figure 6A). RSS information content (RIC) scores calculated using RSSsite (Merelli et al., 2010) showed that the heptamers may serve in the context of a cryptic RSS (cRSS; 12-RSS or 23-RSS); however, the RIC scores all fall below a threshold considered “functional.” Although these scores do not reliably reflect relative V(D)J recombination activity (Zhang and Swanson, 2008), it is noteworthy that IGKV genes often fused to KV4-91 have higher 12-RSS RIC scores than those that are uncommon partners and that the RIC scores for these cRSSs are comparable with those determined for other cRSSs known to support cleavage by RAGs, including LMO2 (−42.00, as 12-RSS) and SIL (−73.66, as 23-RSS) (Zhang and Swanson, 2008) and the KV1-117 3′ 23-RSS analyzed by Shinoda et al. (2019) (−56.20). If this model is correct, then RAG-mediated KV3-7 gene replacement by KV4-91 at the FR3 cRSSs in Figure 6A would involve inversional recombination to form a hybrid joint (Figure 6B).

To test this scenario more directly, we cloned the cRSSs identified above into the V(D)J recombination substrate pJH299 (Hesse et al., 1989) according to their presumed genomic orientation prior to KV gene replacement (KV4-91 antisense and either KV3-7 or KV4-86 sense) (Figure 6C). Cell culture-based V(D)J recombination assays were then performed, using PCR to detect inversional rearrangement events. As expected from our previous studies (Kriatchko et al., 2006), pJH299 undergoes V(D)J recombination to produce signal joints detectable by PCR in HEK293T cells expressing WT RAGs but not cells expressing no RAGs or catalytically inactive D600A RAG1 (Figure 6D). In contrast, cells transfected with forms of pJH299 harboring KV4-91 cRSS with the KV3-7 or KV4-86 cRSS showed similar levels of PCR amplification regardless of RAG expression (Figures 6D and S7A). Indeed, transfection was unnecessary to detect amplification from these plasmids (data

Cell Rep. Author manuscript; available in PMC 2022 December 02.
not shown). Immunoblotting confirmed similar RAG protein levels in RAG-expressing cells (Figures 6E and 5B). Sequencing of cloned PCR products confirmed precise signal joint formation in 4 of 4 clones derived from pJH299-transfected cells expressing WT RAGs, whereas PCR products derived from cells transfected with the KV4-91/KV3-7 cRSS derivative of pJH299 and expressing WT or D600A RAG1 had the same KV hybrid gene sequence as shown in Figure 5E in 1 of 3 and 2 of 4 clones, respectively. These data establish that products resembling type 2 V\textsubscript{H} gene replacement events can arise through artifactual PCR amplification.

**DISCUSSION**

Natural polyreactive antibodies like those that bind PtC provide important surveillance against common microbial pathogens and also play critical housekeeping roles (Binder et al., 2016). Developing a B cell repertoire that sustains these protective immune functions while limiting the potential risk of autoimmunity from excessive polyreactive antibody production requires a balance of positive and negative selection pressures. Previous work describing developmental checkpoints for PtC\textsuperscript{+} B cells in VH12 mice (Tatu and Clarke, 2000; Tatu et al., 1999; Wang and Clarke, 2003; Wang et al., 2007) suggests that entry of V\textsubscript{H}12 transgene-expressing B cells into the mature B cell pool requires pairing with one of a limited number of permissive LCs, followed by positive selection, involving receptor editing when necessary, to achieve and focus expression of a specific LC, V\textsubscript{\kappa}4/5H (KV4-91), that confers PtC reactivity (reviewed in Wang and Clarke, 2004). If receptor editing is indeed required for this to occur, then impairing receptor editing in VH12 mice should diminish PtC\textsuperscript{+} B cell frequency.

Alternatively, after initial positive selection based on LC pairing and PtC binding, receptor editing could reduce the frequency of PtC\textsuperscript{+} B cells in VH12 mice. This outcome is consistent with previous studies showing that self-reactive B cell receptor specificities are removed from the human repertoire at the immature B cell stage in the bone marrow and during the transition to mature naive B cells in the periphery (Wardemann et al., 2003). If so, then impairing receptor editing would enforce BCR specificity to PtC by preventing secondary LC gene rearrangements that lead to expression of a non-permissible LC that ultimately results in clonal deletion or expression of a permissible “editor” LC that disrupts PtC binding. Our finding that dnRAG1 expression in VH12 mice rescues spleen cellularity and promotes development of CD5\textsuperscript{+}PtC\textsuperscript{+} B cells leads us to conclude that this alternative scenario is more likely (Figure 7). Our previous and current observations that splenic B cells from dnRAG1 mice are enriched in autoreactive specificities (Palmer et al., 2018), including PtC, suggest that receptor editing may normally operate to remove a range of B1a-restricted auto- and polyreactive antigenic specificities from the repertoire. Identifying the relevant selecting antigens is of interest for future studies.

How can this model be reconciled with the earlier observation that expressing a permissive non-PtC reactive knockin LC transgene (V\textsubscript{\kappa}1A) in VH12 mice (VH12/V\textsubscript{\kappa}1A mice) allows some PtC\textsuperscript{+} B cells to develop to maturity (Tatu and Clarke, 2000)? We suggest that some developing B cells in VH12/V\textsubscript{\kappa}1A mice undergo stochastic antigen-independent RAG-mediated rearrangement and inactivation of the targeted V\textsubscript{\kappa}1A allele, essentially resetting
development for that B cell and necessitating additional rounds of LC gene rearrangement to obtain a V_{H12}-permissive and PtC-reactive LC to achieve positive selection. This scenario does not preclude receptor editing from occurring after positive selection in VH12/V_{\kappa}1A mice. Further studies of VH12/V_{\kappa}1A mice bred onto a RAG1^{-/−} background showed that most B cells failed to mature, being arrested at the transitional T1/T2 stage (Wang and Clarke, 2003). This observation led to the conclusion that V_{H12}-permissive, non-PtC-reactive LCs fail to support B cell maturation. Although this may be true for B cells expressing V_{H12}/V_{\kappa}1A, it may not be a universal feature for all V_{H12}-permissive LCs because KV1-110 (V_{\kappa}1A) usage only accounts for about 5%–10% of the LC repertoire of CD5−PtC− B cells in VH12 mice (Figure 2). Flow cytometry analysis of CD5−PtC− B cells in VH12 mice shows that the majority have a mature B220^{hi}AA4.1−CD21−CD23^{+}IgM_{int} immunophenotype indistinguishable from CD5−PtC− B cells in WT mice (Figures S4C and S4D; data not shown). Thus, we conclude that most V_{H12}-permissive but non-PtC-reactive LCs support normal B cell maturation.

Our conclusion that receptor editing functions to constrain PtC^{+} B cell expansion in VH12 mice contrasts with the findings of Chumley et al. (2000), who found that two indirect readouts of receptor editing, rag2 and Igλ expression, were not spontaneously increased in primary bone marrow from mice expressing V_{H11} and V_{\kappa}9 transgenes that confer PtC reactivity despite expression being detectable in these cells after anti-κ crosslinking in vitro. Some caution is warranted when interpreting those results, however. First, the high-copy BR1 line of V_{H11}/V_{\kappa}9 transgenic mice exhibits suppressed rag expression (Chumley et al., 2000; Wasserman et al., 1998); thus, it may not be surprising that rag2 expression is difficult to detect in freshly isolated bone marrow from these animals. Second, the inability to detect Igλ expression in bone marrow may be secondary to poor rag expression and may be further complicated by negative selection, as seen with VH12 mice (Figure S4A). Hence, the readouts used to assess receptor editing in that model may not accurately reflect the role of receptor editing in VH12 mice, where endogenous LC gene rearrangement must occur to support B cell development.

The observation that dnRAG1 expression in VH12 mice slightly but reproducibly increases the selection frequency of KV4-91 joining to IGKJ5 could argue against the conclusion that dnRAG1 expression impairs receptor editing because skewing of V_{\kappa} rearrangements to J_{\kappa}5 is often viewed as indicative of this process. However, this contention may not be valid for V_{H} genes like V_{H12} with a very limited repertoire of permissive LC partners. In VH12 mice, multiple “primary” V-J rearrangements may be required to obtain a functional KV4-91 LC that supports pairing with V_{H12} and subsequent positive selection before receptor editing acts to alter receptor specificity (Figure 7). Indeed, the large bias against IGKJ1 usage for KV4-91 sequences in all B cells, regardless of phenotype or genotype (Figure 2B), suggests that joining of this IGKJ segment disables KV4-91 for pairing with most IGHV genes and that likely at least one primary V-J rearrangement must be attempted before a second rearrangement could yield a potentially functional KV4-91 LC. We also consider it plausible that LCs encoded by KV4-91 joined to IGKJ5 may induce receptor editing more efficiently than those joined to IGKJ2 or IGKJ4; hence, dnRAG1 expression may block this event from occurring, leading to IGKJ5 over-representation in the resulting repertoire. The slightly higher MFI of PtC staining in splenic B cells from dnRAG1 VH12 mice...
compared with VH12 mice lends some experimental support to this idea (Figure 1D). This observation may also help explain the paucity of PtC⁺ B cells in the WT spleen; based on dnRAG1 mice, developing B cells expressing IGKV14-126, presumably paired with V₃₁₇ (Mercolino et al., 1989; Pennell et al., 1989), which is prevalently expressed in splenic CD5⁺ B cells (Prohaska et al., 2018; Tsuji et al., 2020), are a major source of B cells capable of binding PtC. The strong PtC staining pattern of these B cells (higher than in VH12 or dnRAG1 VH12 mice; Figure 1D) suggests that such reactivity may normally trigger receptor editing to avoid this specificity in the spleen. Otherwise, the observation that the LC repertoires of WT and dnRAG1 PtC⁻ CD5⁻ B cells are very similar (Figures 2A, 3A, and 3B) argues against the possibility that dnRAG1 introduces bias during early B cell development, which is consistent with previous and current findings showing that absolute numbers of bone marrow developmental B cell subsets in WT and dnRAG1 mice are not significantly different (Hassaballa et al., 2011; Table S1).

Consistent with earlier results (Tatu et al., 1999), we found that CDR3 length in KV4-91 sequences from VH12 PtC⁻ B cell subsets are more heterogeneous in length and composition than their PtC⁺ counterparts. Clear selection bias toward sequences encoding Leu at position 96 (position 8 of CDR3) can be observed in PtC⁺ B cells, although Arg, Phe, and Tyr are also used at this position, whereas Phe is favored over Leu at this position in PtC⁻ B cells. Interestingly, Leu96 is similarly favored in PtC⁻ and PtC⁺ Bone marrow AA4.1⁺ subsets in VH12 mice despite KV4-91 usage being strongly enriched in the PtC⁺ population. This observation suggests that positive selection based on permissive LC pairing has occurred by this point but selection based on antigenic specificity has not. Further evidence of this effect can be seen with KV1-110 and KV4-86 usage; these KV genes show enriched and comparable use in both bone marrow AA4.1⁺ B cell populations but are negatively selected in more mature PtC⁺ B cell populations. Given previous evidence that IGLV1 is a non-permissive LC partner for VH12 (Tatu et al., 1999), the prevalent and similar frequency of IGLV1 expression in bone marrow AA4.1⁺ PtC⁻ and PtC⁺ B cell subsets from both VH12 and dnRAG1 VH12 mice was unexpected. This finding suggests that IGLV1 is permissive for VH12 pairing in vivo but not positively selected based on antigenic specificity or that IGLV1 expression occurs via allelic inclusion with Igk expression. We suspect the former is more likely because we observe strong negative selection of IGLV1 in more mature B cell subsets, which is consistent with earlier findings (Tatu et al., 1999).

Chimeric KV genes, such as those shown in Figure 5, have not been identified previously in the murine Igk locus of primary B cells. However, analogous events have been detected previously by PCR in genomic DNA isolated from the human pre-B cell line 697 stimulated by interleukin-1 (IL-1) or LPS in vitro (Ouled-Haddou et al., 2014). In both cases, the composition of these events is reminiscent of type 2 V₃₁ gene replacement (Darlow and Stott, 2005), in which the 5′ end of a rearranged V₃₁ gene is replaced in frame without gain or loss of nucleotides at the junction. In the study by Ouled-Haddou et al. (2014), the IGKV replacements occurred at regions with sequence identity between donor and recipient genes of the same subgroup. The crossover regions lacked obvious cryptic V(D)J recombination signals but were instead flanked by potential hotspots for activation induced cytidine deaminase (AID), although AID expression was not evident. In other reports of type
2 VH gene replacement (Bellan et al., 2002; Itoh et al., 2000; Lenze et al., 2003; Wilson et al., 2000), evidence suggested that these events are RAG mediated at embedded cRSSs near the junction, but the cRSSs were not rigorously tested for functionality. Although we also identified cRSS motifs that, based on composition, could plausibly support V(D)J recombination, functional testing provided no evidence that they undergo RAG-mediated rearrangement. Instead, the PCR-based assay used here showed that these events may arise by template switching because of FR3 homology. However, this finding does not entirely explain the diminished frequency of KV gene replacements in dnRAG1 VH12 mice versus VH12 mice, given the similar KV gene usage patterns between the two genotypes. Thus, it remains possible that RAG-dependent KV gene replacement can occur at low levels.

Regardless of how KV gene replacements originate, such events may inadvertently be masked in outputs from database searches, complicating immune repertoire analyses when not properly identified, with divergent sequences potentially being misconstrued as evidence of somatic hypermutation. As one example, a previous study identified an IGKV gene cloned from lymphocytes collected from an individual with anti-phospholipid syndrome as A20 (IGKV1-27) with mutated sequences in and around CDR3 (Chukwuocha et al., 1999). However, this sequence could alternatively have arisen via type 2 KV gene replacement with IGKV1(D)-39 with no replacement mutations (Figure S7C). Developing algorithms to identify and sequester likely IGKV (and IGHV) gene replacement events in automated sequence identification programs would be helpful in this regard and will be a focus of future efforts.

Limitations of the study

Although the evidence presented here favors the interpretation that receptor editing constrains expansion of PtC+ B cells in VH12 mice, a limitation of the study is that the data were obtained using transgenic mice. When using such models, we acknowledge that position effects related to transgene insertion cannot be entirely excluded from influencing the phenotype of the animals. We also recognize that the importance of receptor editing in shaping the repertoire of PtC+ B cells may not be readily generalized to B cells with other polyreactive specificities because of differences in the degree of restriction in HC-LC pairing, the relative role of LC in defining BCR specificity, and the level of selection pressure to remove a given polyreactive BCR specificity from the host.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to, and will be fulfilled by, the Lead Contact, Patrick C. Swanson (pswanson@creighton.edu).

Materials availability—Newly generated materials will be made available to qualified investigators by the lead contact following completion of an MTA.
Data and code availability

Data availability: LC NGS datasets generated as reported in this publication have been deposited in NCBI’s Gene Expression Omnibus and are accessible through GEO accession number GSE165776. Whole genome sequencing datasets generated for transgene analyses have been deposited in NCBI’s Sequence Read Archive and are accessible through SRA accession number PRJNA697869.

Code availability: All original code used to process NGS and whole genome sequencing datasets has been deposited at GitHub and is publicly available as of the date of publication. DOIs are listed in the Key resources table.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice—Transgenic dnRAG1 mice, described previously by our laboratory (Hassaballa et al., 2011), were bred to VH12-transgenic (6-1) mice (kind gift of Stephen Clarke and Hongsheng Wang) (Arnold et al., 1994). Both strains have been maintained on a C57Bl/6 background and bred as hemizygous transgenic animals. The dnRAG1 and VH12 transgenes were identified by PCR as described (Arnold et al., 1994; Hassaballa et al., 2011), with minor modification to the VH12 genotyping protocol (Hongsheng Wang, personal communication). Briefly, the VH12 transgene was detected by PCR using primers VH12 and JH1E (see Key resources table) under the following thermal cycling conditions: initial denaturation 94°C 3 min; 30 cycles of amplification (94°C 45s; 62°C 45s; 70°C 1min); and final extension (72°C 5 min). Transgene amplification yields an ~320 bp DNA fragment. Mice were euthanized for phenotypic or sequencing analysis at 8-12 weeks of age; littermates of both sexes were randomly assigned to experimental groups.

We further validated the original dnRAG1 and VH12 transgene sequences by analysing whole genome sequencing datasets (generated by Novogene) using aTRAM 2.0 (Allen et al., 2018) and SPAdes (Prjibelski et al., 2020). Briefly, Trimmomatic (Bolger et al., 2014) was used to remove the Illumina flow cell adapter sequences from the 5′ AND-3′ ends of the WGS data. Next, a set of sequence shards and an SQLite database was compiled for the WGS data files for each animal. Then each database was queried using BLAST+ (Camacho et al., 2009) (through aTRAM) with a target FASTA file containing a sequence specific to the transgenic animal (dnRAG1 or VH12; endogenous RAG1 was used as a control sequence for WT animals). The resulting contigs formed by SPAdes and aTRAM were then reviewed and annotated. The WGS datasets have been deposited in NCBI’s Sequence Read Archive and accessible through SRA accession number PRJNA697869. Assembled, annotated full-length transgene sequences and sample codes are available from GitHub (https://github.com/swanson-lab/VH12-Sequencing). These data provided evidence that both the dnRAG1 and VH12 transgenes were inserted in a tandem array.

Mice were housed under specific pathogen-free conditions in microisolator cages in an AAALAC certified animal facility in accordance with university and federal guidelines.
animal protocols were approved by the Creighton University Institutional Animal Care and Use Committee.

**Cell lines**—HEK293T cells were cultured in Dulbecco’s Modified Eagle’s Medium supplemented with 10% fetal bovine serum, penicillin (100 I.U./mL) and streptomycin (100 μg/mL) at 37°C in a humidified incubator under 5% CO₂.

**METHOD DETAILS**

**Cell staining**—Cells were prepared from bone marrow and spleen as previously described (Palmer et al., 2018). Cells were isolated from the peritoneal cavity by slowly injecting 10 mL of PBS into the abdomen of the euthanized mouse using a small gauge (25G) needle and syringe, gently massaging the mouse abdomen to mix the PBS, and draining the PBS back out into a 50 mL conical tube through a large gauge (16G) needle. Cells were stained with cocktails of antibodies and liposomes, and analyzed on a ZE5 flow cytometer (Bio-Rad) as previously described (Palmer et al., 2018).

**Ig LC repertoire amplification and sequencing**—The workflow for the RNA isolation, cDNA synthesis, and sequencing relies on published strategies and primer sequences (Busse et al., 2014; Vollmers et al., 2013). One to four B cell populations were sorted from bone marrow (CD19⁺AA4.1⁺CD5⁻PtC⁻, CD19⁺AA4.1⁺CD5⁺PtC⁺, CD19⁺AA4.1⁻CD5⁻PtC⁻, CD19⁺AA4.1⁻CD5⁺PtC⁺) or spleen (CD19⁺CD5⁻PtC⁻, CD19⁺CD5⁺PtC⁻, and/or CD19⁺CD5⁺PtC⁺) of WT, dnRAG1, VH12, and/or dnRAG1 VH12 mice using a FACSAria flow cytometer (BD Biosciences). Samples were sorted into DPBS with 30% fetal bovine serum, kept at 4°C and immediately pelleted by centrifugation at 300 × g for 5 min. Supernatants were discarded and the pellets were resuspended in 500 μL of cold TRIzol. RNA isolation was performed following protocols reported on the Immunological Genome Project website (http://www.immgen.org/). The isolated RNA was subjected to first-strand cDNA synthesis using SuperScript III Reverse Transcriptase according to manufacturer’s instructions with specific murine Cκ and Cλ primers (Table S2). All primers were purchased from IDT as PAGE purified, 100 nmole scale oligonucleotides. Second-strand cDNA synthesis was performed using Phusion High-Fidelity Polymerase according to manufacturer’s instructions with degenerate Vκ and Vλ primers to amplify specific Vκ and Vλ gene families (Table S2). The Cκ and Cλ primers contained a 15-mer unique molecular identifier (UMI) sequence to differentiate unique cDNAs, and a partial Illumina adapter sequence for amplification and sequencing steps. The Vκ and Vλ primers also contained a partial adapter sequence, but no UMI. The double-stranded cDNA was purified twice using an equal volume of AMPure XP beads according to manufacturer instructions, and then subjected to 30 rounds of amplification using Illumina adapter primers (both of which contained an Illumina-validated barcode to identify the genotype and sorted cell population (phenotype) for each sample using a two-dimensional array) and Platinum Taq DNA Polymerase High Fidelity (initial denaturation 95°C 2 min; 30 cycles of amplification (95°C 30 s; 65°C 2 min; 68°C 2 min); and final extension (68°C 7 min)). Adaptor and index sequences and schemes were obtained from Illumina product literature and related references (Illumina, 2019a, 2019b; Launen, 2017; Wilkie, 2015). Index sequences in Table S2 were based on approved index groupings for a plexity of 3
(PE1) and a plexity of 8 (PE2). The “AA” in front of the Adapted Illumina Index sequence is used to “adapt” the 6-mer sequence of the selected indexes to the 8 cycles of Index Read Sequencing that the Illumina MiSeq will perform when the sample sheet chemistry is set to “Nextera”. The first four bases of the Multiplexing Read 2 Sequencing Primer (“ACAC”) were included in our primers based on published sequences; however, further inspection revealed they are likely dispensable for correct priming. A distinct indexing scheme was assigned for comparing bone marrow to spleen populations to accommodate the increased number of samples per group.

The prepared libraries were purified with 0.7 × volume of AMPure XP beads to remove smaller (<250bp) nucleic acid products from the amplified samples. Library concentration was measured on a Qubit 3 fluorometer using the Qubit dsDNA HS Assay kit. Average product size and purity was calculated and assessed with the Bioanalyzer High Sensitivity DNA Analysis kit for the Agilent 2100 Bioanalyzer Instrument. Samples were pooled to create a 40nM library using the following equation for each sample:

\[
\frac{[\text{Qubit}]\text{ng/µL}}{[\text{Avg} \cdot \text{bp}] + 660 \text{ ng nmol} \text{ bp}} \times 10^6 \text{ µL} = [\text{Sample}]\text{nM} \rightarrow 40\text{nM} + 1\text{µL} = \text{Sample}(\text{µL}) \text{ to pool}
\]

The concentration of the pooled library was again determined using a Qubit fluorimeter and entered into the above equation with the average of all product lengths as the product size to determine concentration of the pooled library in nM.

The library was diluted to 4nM, denatured for 5 min with an equal volume of 0.2N NaOH prepared in DEPC treated water, and then diluted to 20pM in HT1 Buffer from the MiSeq Reagent Nano Kit v2 (500-cycles). In parallel, PhiX DNA was diluted to 4nM with 200mM TrisCl (200mM Tris hydrochloride in nuclease free water, adjusted to pH 7), then denatured for 5 min in 0.2N NaOH and diluted to 20pM in HT1 buffer. The 20pM library (240 µL) was combined with 20pM PhiX DNA (24 µL) to compensate for low sequence diversity due to the small number of targets and HT1 Buffer (336 µL) for a final loading concentration of 8pM library and 0.8pM PhiX DNA, and then added to the sample well on the MiSeq cartridge. Instrument instructions were followed for loading the flow cell, reagents, and cartridge onto the MiSeq for a paired-end 2 x 250bp sequencing run.

The sample sheet was set up to Generate FASTQ only, using Nextera XT as the assay type to allow for amplicon chemistry to be specified. Initial sequencing runs were performed using 16 samples per cartridge; however, coverage was improved when sequencing 8 samples from one complete group of animals at a time (1 sorted population from WT, 3 populations from dnRAG1, and 2 populations each from VH12 and dnRAG1 VH12 animals). For later runs comparing bone marrow and spleen populations from VH12 and dnRAG1 VH12 animals, 12 samples were loaded per cartridge to minimize technical variation among one complete group (4 sorted populations each from bone marrow plus 2 sorted populations each from spleen). Additionally, to ensure adequate population sizes were obtained from the lymphopenic VH12 and dnRAG1 VH12 mice, each experimental sample consisted of cells pooled from at least two animals of these genotypes.
Bioinformatics and computational analysis—The bioinformatics workflow was based on sample workflows available on the Immcantation portal (https://immcantation.readthedocs.io/en/stable/) and relies on published strategies and tools (R Core Team. R, 2019: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/, 2019; Alamyar et al., 2012; Bolen et al., 2017; Busse et al., 2014; Gupta et al., 2015; Li et al., 2013; Madeira et al., 2019; Stern et al., 2014; Vander Heiden et al., 2014; Vollmers et al., 2013; Wagih, 2017; Wickham et al., 2019). Briefly, the bcl2fastq program was used to convert the raw read data into FASTQ files, demultiplex samples based on the indices, and remove the Illumina adapters used to bind the sequence amplicons to the flow cell. The pRESTO package (Vander Heiden et al., 2014) was used to quality filter and combine mate-pair sequences, group sequences based on UMIs, and generate consensus sequences for each UMI group. The resulting FASTA sequence files for each sample were then downloaded from the server and uploaded to IMGT/HighV-QUEST (Alamyar et al., 2012; Li et al., 2013) to identify immunoglobulin V and J gene usage. Change-O (Gupta et al., 2015) was used for sequence annotation, and to parse and reformat the columns of the IMGT data tables.

RStudioServer (R Core Team. R, 2019: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/, 2019) was used to run the R-based ALAKAZAM toolkit (Stern et al., 2014) to subsample the annotated data tables and graph gene usage using ggplot2 (Wickham et al., 2019). The stringr package was then used to pull sequences using specific IGKV genes and further analyze their pairing with IGKJ segments, the inclusion of NP nucleotides, and CDR3 sequence lengths. Ggseqlogo (Wagih, 2017) was used in addition to ggplot2 to create CDR3 sequence logos of the pulled sequences.

The Repertoire Dissimilarity Index (RDI), a non-parametric means of directly comparing the various repertoires through rigorous quantification of V and J gene usage (Bolen et al., 2017), was calculated using the RDI package on the Immcantation website. Briefly, total V and J gene counts were tabulated for each independent sample and were used by the program to calculate repertoire distances that were reported as log fold-change differences between pairs of repertoires.

To analyze IGKV chimeric sequences, individual sequences with NP insertions greater than 5 nucleotides were individually submitted to IMGT/V-QUEST to identify IGKV genes with sequence identity to the 3′ end of the hybrid IGKV genes. For this purpose, we found that removing the first 60 nucleotides of the sequence improved identification. Multiple sequence alignments were performed using Clustal Omega (Madeira et al., 2019).

Plasmid V(D)J recombination assay—Synthetic duplex DNA containing the FR3 cRSSs identified in the text were formed by annealing complementary oligonucleotides (nucleotides in bold font were incorporated to produce sticky ends compatible for cloning into the Sau [KV4-91] or BamH1 [KV3-7 and KV4-86] site of pJH299 (Hesse et al., 1989). These were designed to encompass both 12- and 23-cRSS sequences identified in Figure 6A; putative heptamer and nonamer sequences for the 12-cRSS on the sense (top) strand are shaded.

Cell Rep. Author manuscript; available in PMC 2022 December 02.
IGKV4-91:
5’-
TCGACAGGACATCCAATCTGGCTTCTGGAGTCCCAGCTCGCTTCAGTGGCAGTGG-3’
5’-
TCGACCACTGCCACTGAAGCGAGCTGGGACTCCAGAAGCCAGATGGATGTCCCTG-3’

IGKV3-7:
5’-
GATCCTATGCATCCAACCTAGAATCTGGGGTCCCTGCCAGGTTCAGTGGCAGTGG-3’
5’-
GATCCCACTGCCACTGAACCTGGCAGGGACCCCAGATTCTAGGTTGGATGCA-TAG-3’

IGKV4-86:
5’-
GATCCGAAATATCCAAACTGGCTTCTGGAGTCCCAGCTCGCTTCAGTGGCAGTGG-3’
5’-
GATCCCACTGCCACTGAAGCGAGCTGGGACTCCAGAAGCCAGTTTGGATATTTCG-3’

Plasmids harboring the correct sequence in the desired orientation (KV4-91, antisense; KV3-7 or KV4-86, sense) were identified by sequencing.

pEBB MBP-tagged core RAG1-his-myc and full-length RAG2 expression constructs were obtained from Dr. Michael Lieber (Shimazaki et al., 2009). The MBP coding sequence was removed from the pEBB MBP-core RAG1-his-myc construct by KpnI digestion; sequence encoding a RAG1 active site mutation was inserted by swapping a BsrG1 fragment from pcDNA1 full-length RAG1 D600A (Swanson, unpublished).

The V(D)J recombination assay was performed following procedures described previously (Kriatchko et al., 2006), with minor modifications: each 10 cm dish of HEK293T cells was transfected with 5 μg pJH299 or its derivatives, and 2.5 mg of each pEBB RAG construct with 30 mg polyethylenimine and harvested 48h post-transfection. Plasmid DNA was immediately isolated from half the cells using an Omega BIO-TEK E.Z.N.A. plasmid DNA mini kit I, and adjusted to 5 ng/μL. Recombination products or a template control region were detected from 10 ng of plasmid DNA by PCR using GoTaq Flexi reagents (Promega) and primer sets (2.5 pmol each primer) to detect inversional recombination products (6382 Reverse and 6624 Reverse), or a fragment of the chloramphenicol acetyltransferase gene.
(CAT Forward and CAT Reverse), respectively (see Key resources table). The following thermal cycling conditions were used: initial denaturation (95°C, 5 min); 25 cycles of amplification (95°C, 15 s; 50°C, 30 s; 72°C, 20 s; and final extension (72°C, 5 min). Selected PCR products were gel-isolated, cloned using a TA-TOPO kit (Invitrogen), and plasmids containing inserts of the expected length were sequenced.

Lysates prepared from the remaining half of the cells were probed by immunoblotting as previously described (Schabla et al., 2018) using the following antibodies: anti-Myc to detect RAG1 (clone 9E10 1:5000); anti-MBP to detect RAG2 (clone 8G1; 1:1000); and anti-β-actin (clone AC-74, 1:1000). Blots were developed using HRP-conjugated horse anti-mouse secondary antibody (1:5000) and Pierce ECL 2 substrate (Thermo Fisher Scientific), and imaged using a Typhoon 9410 Variable Mode Imager (GE Healthcare).

QUANTIFICATION AND STATISTICAL ANALYSIS

Data are presented as mean values ± standard error of the means. Collected data were subjected to analysis of variance and post hoc testing using the PASW Statistics 27.0 software package (SPSS Inc., Chicago, IL). Unless otherwise indicated, differences with a p value of <0.05 by Tukey post-hoc testing are considered statistically significant.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank the Creighton University RadLab for configuring the server used for this study. The authors also thank Dr. M. Lieber for providing pEBB RAG expression constructs and Dr. Rob Todd and Dr. Anna Selmecki for valuable assistance and suggestions during LC sequencing protocol development. This work was supported by grants to P.C.S. from the National Institutes of Health (NIH) (R21AI119829 and R21AI159127) and the George F. Haddix President’s Faculty Research Fund and by revenue from Nebraska’s excise tax on cigarettes awarded to Creighton University through the Nebraska Department of Health & Human Services (DHHS). NIH funding to support research laboratory construction (C06RR017417), the Creighton University Animal Resource Facility (G20RR024001), and acquisition of the GE Healthcare Typhoon 9410 Variable Mode Imager (S10RR027352) and Bio-Rad ZE5 Cell Analyzer (3R01GM102487-03S1) used in this study is gratefully acknowledged. This publication’s contents represent the view(s) of the author(s) and do not necessarily represent the official views of the State of Nebraska, DHHS, or the NIH.

REFERENCES

Alamyar E, Duroux P, Lefranc MP, and Giudicelli V (2012). IMGT® tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. Methods Mol. Biol 882, 569–604. 10.1007/978-1-61779-842-9_32. [PubMed: 22665256]

Allen JM, LaFrance R, Folk RA, Johnson KP, and Guralnick RP (2018). aTRAM 2.0: an improved, flexible locus assembler for NGS data. Evol. Bioinf. Online 14, 117693431877454. 10.1177/1176934318774546.

Allman D, Lindley RC, DeMuth W, Rudd K, Shinton SA, and Hardy RR (2001). Resolution of three nonproliferative immature splenic B cell subsets reveals multiple selection points during peripheral B cell maturation. J. Immunol 167, 6834–6840. 10.4049/jimmunol.167.12.6834. [PubMed: 11739500]
Aoki-Ota M, Torkamani A, Ota T, Schork N, and Nemazee D (2012). Skewed primary Igk repertoire and V–J joining in C57BL/6 mice: implications for recombination accessibility and receptor editing. J. Immunol 188, 2305–2315. 10.4049/jimmunol.1103484. [PubMed: 22287713]

Arnold LW, Pennell CA, McCray SK, and Clarke SH (1994). Development of B-1 cells: segregation of phosphatidyl choline-specific B cells to the B-1 population occurs after immunoglobulin gene expression. J. Exp. Med 179, 1585–1595. 10.1084/jem.179.5.1585. [PubMed: 8163938]

Bellan C, Lazzi S, Zazzi M, Lalinga AV, Palummo N, Galieni P, Marafioti T, Tonini T, Cinti C, Leoncini L, et al. (2002). Immunoglobulin gene rearrangement analysis in composite hodgkin disease and large B-cell lymphoma: evidence for receptor revision of immunoglobulin heavy chain variable region genes in Hodgkin-Reed-Sternberg cells? Diagn. Mol. Pathol 11, 2–8. 10.1097/00019606-200203000-00002. [PubMed: 11854595]

Binder CJ, Horkko S, Dewan A, Chang MK, Kieu EP, Goodyear CS, Shaw PX, Palinski W, Witztum JL, and Silverman GJ (2003). Pneumococcal vaccination decreases atherosclerotic lesion formation: molecular mimicry between Streptococcus pneumoniae and oxidized LDL. Nat. Med 9, 736–743. 10.1093/nm876. [PubMed: 12740573]

Binder CJ, Papac-Milicevic N, and Witztum JL (2016). Innate sensing of oxidation-specific epitopes in health and disease. Nat. Rev. Immunol 16, 485–497. 10.1038/nri.2016.63. [PubMed: 27346802]

Bolen CR, Rubelt F, Vander Heiden JA, and Davis MM (2017). The Repertoire Dissimilarity Index as a method to compare lymphocyte receptor repertoires. BMC Bioinformatics 18, 155. 10.1186/s12859-017-1556-5. [PubMed: 28264647]

Bolger AM, Lohse M, and Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. 10.1093/bioinformatics/btu170. [PubMed: 24695404]

Borghesi L, Aites J, Nelson S, Letterov P, James P, and Gerstein R (2005). E47 is required for V(D)J recombinase activity in common lymphoid progenitors. J. Exp. Med 202, 1669–1677. 10.1084/jem.20051190. [PubMed: 16365147]

Brakenhoff RH, Schoenmakers JG, and Lubsen NH (1991). Chimeric cDNA clones: a novel PCR artifact. Nucleic Acids Res. 19, 1949. 10.1093/nar/19.8.1949. [PubMed: 2030976]

Busse CE, Czogiel I, Braun P, Arndt PF, and Wardemann H (2014). Single-cell based high-throughput sequencing of full-length immunoglobulin heavy and light chain genes. Eur. Immunol 44, 597–603. 10.1002/eji.201343917. [PubMed: 24114719]

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, and Madden TL (2009). BLAST+: architecture and applications. BMC Bioinformatics 10, 421. 10.1186/1471-2105-10-421. [PubMed: 20003500]

Casali P, Burastero SE, Nakamura M, Inghirami G, and Notkins AL (1987). Human lymphocytes making rheumatoid factor and antibody to ssDNA belong to Leu-1+ B-cell subset. Science 236, 77–81. 10.1126/science.3105056. [PubMed: 3105056]

Chukwuocha RU, Hsiao ET, Shaw P, Witztum JL, and Chen PP (1999). Isolation, characterization and sequence analysis of five IgG monoclonal anti-beta 2-glycoprotein-1 and anti-prothrombin antigen-binding fragments generated by phage display. J. Immunol 163, 4604–4611. 10.1001/jimmunol.164.2105-10.421. [PubMed: 10779761]

Chumley MJ, Dal Porto JM, Kawaguchi S, Cambier JC, Nemazee D, and Hardy RR (2000). A VH11V kappa 9B cell antigen receptor drives generation of CD5+ B cells both in vivo and in vitro. J. Immunol 164, 4586–4593. 10.4049/jimmunol.164.9.4586. [PubMed: 10779761]

Darlow JM, and Stott DI (2005). V(H) replacement in rearranged immunoglobulin genes. Immunology 114, 155–165. 10.1111/j.1365-2567.2004.02084.x. [PubMed: 15667560]

Forster I, and Rajewsky K (1987). Expansion and functional activity of Ly-1+ B cells upon transfer of peritoneal cells into allogeneic-congenic, newborn mice. Eur. J. Immunol 17, 521–528. 10.1002/eji.1830170414. [PubMed: 2436924]

Graf R, Seagal J, Otipoby KL, Lam KP, Ayoub S, Zhang B, Sander S, Chu VT, and Rajewsky K (2019). BCR-dependent lineage plasticity in mature B cells. Science 363, 748–753. 10.1126/science.aau8475. [PubMed: 30765568]

Gupta NT, Vander Heiden JA, Uduman M, Gadala-Maria D, Yaari G, and Kleinseit SH (2015). Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data: table 1. Bioinformatics 31, 3356–3358. 10.1093/bioinformatics/btv359. [PubMed: 26069265]
Hardy RR (2006). B-1 B cell development. J. Immunol 177, 2749–2754. 10.4049/jimmunol.177.5.2749. [PubMed: 16920907]

Hardy RR, Carmack CE, Shintom SA, Kemp JD, and Hayakawa K (1991). Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. J. Exp. Med 173, 1213–1225. 10.1084/jem.173.5.1213. [PubMed: 1827140]

Hardy RR, Hayakawa K, Shimizu M, Yamasaki K, and Kishimoto T (1987). Rheumatoid factor secretion from human Leu-1+ B cells. Science 236, 81–83. 10.1126/science.3105057. [PubMed: 3105057]

Hassaballa AE, Palmer VL, Anderson DK, Kassmeier MD, Nganga VK, Parks KW, Volkmer DL, Perry GA, and Swanson PC (2011). Accumulation of B1-like B cells in transgenic mice over-expressing catalytically inactive RAG1 in the periphery. Immunology 134, 469–486. 10.1111/j.1365-2567.2011.03509.x. [PubMed: 22044391]

Hayakawa K, Asano M, Shintom SA, Gui M, Wen LJ, Dashoff J, and Hardy RR (2003). Positive selection of anti-thy-1 autoreactive B-1 cells and natural serum autoantibody production independent from bone marrow B cell development. J. Exp. Med 197, 87–99. 10.1084/jem.20021459. [PubMed: 12515816]

Hayakawa K, Hardy RR, Honda M, Herzenberg LA, Steinberg AD, and Herzenberg LA (1984). Ly-1 B cells: functionally distinct lymphocytes that secrete IgM autoantibodies. Proc. Natl. Acad. Sci. USA 81, 2494–2498. 10.1073/pnas.81.8.2494. [PubMed: 6609363]

Hesse JE, Lieber MR, Mizuuchi K, and Gellert M (1989). V(D)J recombination: a functional definition of the joining signals. Genes Dev. 3, 1053–1061. 10.1101/gad.3.7.1053. [PubMed: 2777075]

Illumina. (2019a). Illumina Adapter Sequences. Document # 1000000002694. 06/13/2019. Accessed 09/07/2019. https://support.illumina.com/downloads/illumina-adapter-sequences-document-1000000002694.html.

Illumina. (2019b). Index Adapters Pooling Guide. Document #1000000041074. 06/13/2019. Accessed 09/07/2019. https://support.illumina.com/downloads/index-adapters-pooling-guide-1000000041074.html.

Itoh K, Meffre E, Albessiano E, Farber A, Dines D, Stein P, Asnis SE, Furie RA, Jain RI, and Chiorazzi N (2000). Immunoglobulin heavy chain variable region gene replacement as a mechanism for receptor revision in rheumatoid arthritis synovial tissue B lymphocytes. J. Exp. Med 192, 1151–1164. 10.1084/jem.192.8.1151. [PubMed: 11034605]

Kriatchko AN, Anderson DK, and Swanson PC (2006). Identification and characterization of a gain-of-function RAG-1 mutant. Mol. Cell Biol 26, 4712–4728. 10.1128/Mcb.02487-05. [PubMed: 16738334]

Kristiansen TA, Vanhee S, and Yuan J (2018). The influence of developmental timing on B cell diversity. Curr. Opin. Immunol 51, 7–13. 10.1016/j.coi.2017.12.005. [PubMed: 29272734]

Lanning DK, and Knight KL (2015). Diversification of the primary antibody repertoire by AID-mediated gene conversion. Results Probl. Cell Differ. 57, 279–293. 10.1007/978-3-319-20819-0_12.

Launen L (2017). Illumina Sequencing (for Dummies) - An overview on how our samples are sequenced. In kscbioinformatics (blog). February 13, 2017. https://kscbioinformatics.wordpress.com/2017/02/13/illumina-sequencing-for-dummies-samples-are-sequenced/.

Lenze D, Greiner A, Knorr C, Anagnostopoulos I, Stein H, and Hummel M (2003). Receptor revision of immunoglobulin heavy chain genes in human MALT lymphomas. Mol. Pathol 56, 249–255. 10.1136/mp.56.5.249. [PubMed: 14514917]

Li S, Lefranc MP, Miles JJ, Alamyar E, Giudicelli V, Duroux P, Freeman JD, Corbin VDA, Scheerlinck JP, Frohman MA, et al. (2013). IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. Nat. Commun 4, 2333. 10.1038/ncomms3333. [PubMed: 23995877]

Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, and Lopez R (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47, W636–W641. 10.1093/nar/gkz268. [PubMed: 30976793]

Cell Rep. Author manuscript; available in PMC 2022 December 02.
Masmoudi H, Mota-Santos T, Huetz F, Coutinho A, and Cazenave PA (1990). All T15 Id-positive antibodies (but not the majority of VHT15+ antibodies) are produced by peritoneal CD5+ B lymphocytes. Int. Immunol 2, 515–520. 10.1093/intimm/2.6.515. [PubMed: 1707658]

Mercolino TJ, Arnold LW, Hawkins LA, and Haughton G (1988). Normal mouse peritoneum contains a large population of Ly-1+ (CD5) B cells that recognize phosphatidyl choline. Relationship to cells that secrete hemolytic antibody specific for autologous erythrocytes. J. Exp. Med 168, 687–698. 10.1084/jem.168.2.687. [PubMed: 3045250]

Mercolino TJ, Locke AL, Afshari A, Sasser D, Travis WW, Arnold LW, and Haughton G (1989). Restricted immunoglobulin variable region gene usage by normal Ly-1 (CD5+) B cells that recognize phosphatidyl choline. J. Exp. Med 169, 1869–1877. 10.1084/jem.169.6.1869. [PubMed: 2499651]

Merelli I, Guffanti A, Fabbri M, Cocito A, Furia L, Grazini U, Bonnal RJ, Milanesi L, and McBlane F (2010). RSSsite: a reference database and prediction tool for the identification of cryptic Recombination Signal Sequences in human and murine genomes. Nucleic Acids Res. 38, W262–W267. 10.1093/nar/gkq391. [PubMed: 20478831]

Montecino-Rodriguez E, Leathers H, and Dorshkind K (2006). Identification of a B-1 B cell-specified progenitor. Nat. Immunol 7, 293–301. 10.1038/nim.2006.192. [PubMed: 16429139]

Ouled-Haddou H, Ghamlouch H, Regnier A, Trudel S, Herent D, Lefranc MP, Marolleau JP, and Gubler B (2014). Characterization of a new V gene replacement in the absence of activation-induced cytidine deaminase and its contribution to human B-cell receptor diversity. Immunology 141, 268–275. 10.1111/imn.12192. [PubMed: 24134819]

Paabo S, Irwin DM, and Wilson AC (1990). DNA damage promotes jumping between templates during enzymatic amplification. J. Biol. Chem 265, 4718–4721. 10.1016/s0021-9258(19)39621-8. [PubMed: 2307682]

Palmer VL, Aziz-Seible R, Kassmeier MD, Rothermund M, Perry GA, and Swanson PC (2015). VprBP is required for efficient editing and selection of Igκ+ B cells, but is dispensable for Igλ+ and marginal zone B cell maturation and selection. J. Immunol 195, 1524–1537. 10.4049/jimmunol.1500952. [PubMed: 26150531]

Palmer VL, Worth AN, Scott RL, Perry GA, Yan M, Li QZ, and Swanson PC (2018). IL10 restrains autoreactive B cells in transgenic mice expressing inactive RAG1. Cell Immunol. 331, 110–120. 10.1016/j.cellimm.2018.06.004. [PubMed: 30017086]

Pennell CA, Mercolino TJ, Grdina TA, Arnold LW, Haughton G, and Clarke SH (1989). Biased immunoglobulin variable region gene-expression by Ly-1 B-cells due to clonal selection. Eur. J. Immunol 19, 1289–1295. 10.1002/eji.1830190721. [PubMed: 2503389]

Prijibelski A, Antipov D, Meleshko D, Lapidus A, and Korobeynikov A (2020). Using SPAdes De Novo assembler. Curr. Protoc. Bioinf 70, e102. 10.1002/cpbi.102.

Prohaska TA, Que X, Diehl CJ, Hendriks S, Chang MW, Jepsen K, Glass CK, Benner C, and Witzum JL (2018). Massively parallel sequencing of peritoneal and splenic B cell repertoires highlights unique properties of B-1 cell antibodies. J. Immunol 200, ji1700568–1717. 10.4049/jimmunol.1700568.

R Core Team (2019). R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing). https://www.R-project.org/.

Schabla NM, Perry GA, Palmer VL, and Swanson PC (2018). VprBP (DCAF1) Regulates RAG1 Expression Independently of Dicer by Mediating RAG1 Degradation. J. Immunol 201, 930–939. 10.4049/jimmunol.1800054. [PubMed: 29925675]

Seidl KJ, MacKenzie JD, Wang D, Kantor AB, Kabat EA, Herzenberg LA, and Herzenberg LA (1997). Frequent occurrence of identical heavy and light chain Ig rearrangements. Int. Immunol 9, 689–702. 10.1093/intimm/9.5.689. [PubMed: 9184914]

Shaw PX, Horkko S, Chang MK, Curtiss LK, Palinski W, Silverman GJ, and Witzum JL (2000). Natural antibodies with the T15 idiotype may act in atherosclerosis, apoptotic clearance, and protective immunity. J. Clin. Invest 105, 1731–1740. 10.1172/JCI8472. [PubMed: 10862788]

Shimazaki N, Tsai AG, and Lieber MR (2009). H3K4me3 stimulates the V(D)J RAG complex for both nicking and hairpinning in trans in addition to tethering in cis: implications for translocations. Mol. Cell 34, 535–544. 10.1016/j.molcel.2009.05.011. [PubMed: 19524534]
Shinoda K, Maman Y, Canela A, Schatz DG, Livak F, and Nussenzweig A (2019). Intra-κ cluster recombination shapes the Ig kappa locus repertoire. Cell Rep. 29, 4471–4481.e6. 10.1016/j.celrep.2019.11.088. [PubMed: 31875554]

Stern JNH, Yaari G, Vander Heiden JA, Church G, Donahue WF, Hintzen RQ, Huttner AJ, Laman JD, Nagra RM, Nylander A, et al. (2014). B cells populating the multiple sclerosis brain mature in the draining cervical lymph nodes. Sci. Transl. Med. 6, 248ra107. 10.1126/scitranslmed.3008879.

Tatu C, and Clarke SH (2000). Selective maturation of VH12 B cells in the spleen enriches for anti-phosphatidyl choline B cells: evidence for receptor editing. Curr. Top. Microbiol. Immunol 252, 77–86. 10.1007/978-3-642-57284-5_9. [PubMed: 11125494]

Tatu C, Ye J, Arnold LW, and Clarke SH (1999). Selection at multiple checkpoints focuses V(H)12 B cell differentiation toward a single B-1 cell specificity. J. Exp. Med 190, 903–914. 10.1084/jem.190.7.903. [PubMed: 10510080]

Tsujii N, Rothstein TL, and Holodick NE (2020). Antigen receptor specificity and cell location influence the Diversification and selection of the B-1a cell pool with age. J. Immunol 205, 741–759. 10.4049/jimmunol.1901302. [PubMed: 32561570]

Vander Heiden JA, Yaari G, Uduman M, Stern JN, O’Connor KC, Hafer DA, Vigneault F, and Kleinstein SH (2014). pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. Bioinformatics 30, 1930–1932. 10.1093/bioinformatics/btu138. [PubMed: 24618469]

Vollmers C, Sit RV, Weinstein JA, Dekker CL, and Quake SR (2013). Genetic measurement of memory B-cell recall using antibody repertoire sequencing. Proc. Natl. Acad. Sci. U S A 110, 13463–13468. 10.1073/pnas.1312146110. [PubMed: 23898164]

Wagh O (2017). ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics 33, 3645–3647. 10.1093/bioinformatics/btx469. [PubMed: 29036507]

Wang H, and Clarke SH (2003). Evidence for a ligand-mediated positive selection signal in differentiation to a mature B cell. J. Immunol 171, 6381–6388. 10.4049/jimmunol.171.12.6381. [PubMed: 14662836]

Wang H, and Clarke SH (2004). Positive selection focuses the VH12 B-cell repertoire towards a single B1 specificity with survival function. Immunol. Rev 197, 51–59. 10.1111/j.0105-2896.2004.0098.x. [PubMed: 18025199]

Wardemann H, Yurasov S, Schaefer A, Young JW, Meffre E, and Nussenzweig MC (2003). Predominant autoantibody production by early human B cell precursors. Science 301, 1374–1377. 10.1126/science.1086907. [PubMed: 12920303]

Wasserman R, Li YS, Shinton SA, Carmack CE, Manser T, Wiest DL, Hayakawa K, and Hardy RR (1998). A novel mechanism for B cell repertoire maturation based on response by B cell precursors to pre-B receptor assembly. J. Exp. Med 187, 259–264. 10.1084/jem.187.2.259. [PubMed: 9432984]

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grollemund G, Hayes A, Henry L, Hester J, et al. (2019). Welcome to the tidyverse. J. Open Source Softw 4, 1686. 10.21105/joss.01686.

Wilkie G (2015). Illumina Adapter and Primer Sequences. In CVR Bioinformatics (blog). January 26, 2015. https://bioinformatics.cvr.ac.uk/illumina-adapter-and-primer-sequences/.

Wilson PC, Wilson K, Liu YJ, Banchereau J, Pascual V, and Capra JD (2000). Receptor revision of immunoglobulin heavy chain variable region genes in normal human B lymphocytes. J. Exp. Med 191, 1881–1894. 10.1084/jem.191.11.1881. [PubMed: 10839804]

Zhang M, and Swanson PC (2008). V(D)J recombinase binding and cleavage of cryptic recombination signal sequences identified from lymphoid malignancies. J. Biol. Chem 283, 6717–6727. M710301200 [pii] 10.1074/jbc.M710301200. 10.1074/jbc.m710301200. [PubMed: 18187418]
Highlights

- Impairing receptor editing by dnRAG1 expression in VH12 mice increases PtC+ B cells
- PtC+ B cells in VH12 mice show biased LC KV4-91 usage, selected against joining to KJ1
- dnRAG1 expression in VH12 mice skews KV4-91 joining to KJ5
- Occasional in-frame hybrid LC genes probably arise through a RAG-independent mechanism
Figure 1. dnRAG1 expression in VH12 mice enforces PtC reactivity of splenic CD5+ B cells
(A) Total numbers of cells, lymphocytes (Lymphs), and CD19+ lymphocytes were determined in bone marrow (BM) and spleen (SPL) of WT, dnRAG1, VH12, and dnRAG1 VH12 mice.
(B) SPL lymphocytes from mice in (A) were analyzed by flow cytometry for various surface markers using gating strategies defined under each row. Developmental subsets are shown on the right with corresponding gates.
(C) CD19+ lymphocytes from the SPL and peritoneal cavity (PerC) were analyzed for CD5 expression and PtC staining.
(D) The absolute number of CD5+PtC+ cells and mean fluorescence intensity (MFI) of PtC and CD5 staining are plotted as in (A). The percentage of gated cells is shown for representative animals from 7–8 mice/genotype. Significant differences (*p < 0.05) for group comparisons are indicated. Table S1 summarizes mean values (± SEM) and statistical comparisons for each population and genotype. (A–D) The percentage of gated cells is shown for representative animals from 7–8 mice/genotype. Significant differences (p < 0.05) for group comparisons are indicated by an asterisk. Table S1 summarizes mean values (+− SEM) and statistical comparisons for each population and genotype.
Figure 2. CD5⁺PtC⁺ B cells from dnRAG1 and VH12 mice show restricted but distinctive LC V gene usage and biased IGKJ gene usage patterns based on IGKV gene identity, with evidence of selection occurring in VH12 mice after the AA4.1⁺ stage in the BM.

(A) Usage frequencies for each IGKV and IGLV gene (according to IMGT nomenclature) identified by NGS experiments are plotted for the B cell populations and genotypes identified by color in the legend (inset) as mean ± SEM from three biological replicates. IGKV genes are organized according to genomic position relative to the IGKJ cluster. Four KV pseudogenes (ψ) are also included; all have a functional recombination signal sequence.
(RSS) but harbor an absent initiation codon, nucleotide deletion, or in-frame stop codon. IGLV genes are located next to the distal KV genes.

(B) Frequencies of total IGKJ and LJ gene usage and KJ usage for specific IGKV genes from data in (A). High-resolution images of these graphs are available on GitHub (https://github.com/swanson-lab/VH12-Sequencing).
Figure 3. Repertoire dissimilarity index (RDI) analysis of LC repertoires shows distinct clustering of PtC+ and PtC− B cell subsets

(A–D) Hierarchical clustering of LC repertoires by RDI (A and C) and log fold change differences between repertoires (B and D) based on aggregated data from NGS experiments using all mouse genotypes (A and B) or just VH12 and dnRAG1 VH12 mice (C and D). Variation in log fold change is color coded from green (lowest) to red (highest).
Figure 4. IGKV genes used frequently in PtC⁺ B cells show evidence of restricted CDR3 length and composition

(A) Selected IGKV sequences from CD5⁻PtC⁻ and CD5⁺PtC⁺ B cell populations were aggregated from three biological replicates, and the percentage of sequences at each CDR3 length was calculated. The inset shows the number of sequences analyzed per population and the significance of the observed difference (n.s., not significant).

(B) IGKV sequences in (A) harboring nine amino acids in CDR3 (the most frequent length) were analyzed for composition using ggseqlogo. The number of sequences used to generate each logo is shown below the graph; chemical properties of residues are color coded as

*Cell Rep. Author manuscript; available in PMC 2022 December 02.*
follows: acidic, red; basic, blue; hydrophobic, black; neutral, purple; polar, green. Enlarged logos at position 8 are shown for each cell population on the right.

(C) CDR3 length distribution and composition were analyzed for IGKV4-91 and IGKV3-7 from VH12 BM and SPL PtC+ B cell populations analyzed in Figure S2.
Figure 5. KV4-91 sequences from CD5\textsuperscript{−}PtC\textsuperscript{−} and CD5\textsuperscript{−}PtC\textsuperscript{+} B cell populations show differences in NP nucleotide insertion length and examples of KV gene replacement

(A) KV4-91 sequences were analyzed in two biological replicates (EXP1 and EXP2) for each mouse genotype and B cell population in Figure 2A, listing the number of KV4-91 sequences identified (COUNTS), the average NP length (AVG NP), and the number and frequency of sequences with NP lengths greater than 5 nt (NP > 5NT). Significant differences (p value) between the identified population and CD5\textsuperscript{−}PtC\textsuperscript{−} B cells from the same genotype and experiment were determined by one-way ANOVA. For KV4-91 sequences with NP > 5NT, the likely origin of the sequence fused to the 3′ end was analyzed using IMGT/V-QUEST; the frequencies of common KV fusion partners are listed. Those with limited representation are included in “Other,” and those whose identity was ambiguous are classified as undefined (Undef.).

(B) Example of IMGT/V-QUEST output for a KV4-91 sequence obtained from CD5\textsuperscript{−}PtC\textsuperscript{−} B cells sorted from a VH12 mouse identified as having a large deletion at the 3′ end with a concomitant large non-templated (N) insertion.

(C) IMGT/V-QUEST alignment of the sequence in (B), highlighting sequence identity with KV4-91 at the 5′ end (blue) and KV4-86 at the 3′ end (pink), respectively. Alignment to KJ2 is also shown.
(D and E) Example of IMGT/V-QUEST output (D) and sequence alignment (E) for a hybrid KV4-91 gene with 3′ sequence identity to KV3-7 obtained from CD5+PtC+ B cells of a VH12 mouse.
Figure 6. Framework region three (FR3) of IGKV genes contains conserved antisense cryptic RSS-like motifs, but they do not support detectable levels of RAG-mediated V(D)J recombination

(A) IMGT germline KV genes identified in KV4-91 or KV4-86 hybrid sequences as common (KV4-86, KV4-59, and KV3-7) or uncommon (KV1-110 and KV5-43) partners are aligned. The position of tandem, overlapping, conserved heptamer-like motifs in the antisense orientation are identified in FR3 (shaded or boxed). Putative nonamers 12 bp from the 5′-most heptamer (shaded) or 23 bp from the 3′-most heptamer (boxed) are indicated, with consensus sequences aligned above the cryptic 12-RSS. RIC scores calculated for both
antisense cRSSs are shown on the right; all fail to meet the threshold for functionality. Sequence identities between the KV genes are highlighted (pink).

(B) Model of RAG-mediated gene replacement of rearranged IGKV3-7 with IGKV4-91 at the embedded 3′ cRSS (dashed triangle).

(C) Organization and expected V(D)J rearrangement outcomes of pJH299 or its derivatives containing FR3 cRSSs. Primer sites used to identify rearrangement or a control DNA sequence (red or black half-arrows, respectively) and SalI (S) and BamHI (B) sites are indicated.

(D) Plasmid DNA isolated from cells transfected with pJH299 or its derivative harboring KV4-91 and KV3-7 sequences as in (B) and (C), with or without empty or RAG expression constructs, was subjected to PCR to detect inversion (6382R/6624R) or control DNA (CATFOR/CATREV).

(E) Whole-cell lysates from (D) were probed by immunoblotting (IB) to detect RAG1 (anti-myc tag), RAG2 (anti-MBP tag), or β-actin (loading control).

The results in (D) and (E) are representative of two independent experiments.
During B cell development, the Ig LC loci undergo multiple, sequential “primary” V-J rearrangements to express one of relatively few “permissive” LCs capable of pairing with the V_{H12} HC, followed by antigen-driven positive selection, which focuses KV gene usage to KV4-91 to confer PtC reactivity. After positive selection, PtC-reactive B cells may undergo receptor editing to replace the expressed KV4-91 LC. Those successfully edited to express an alternative permissive LC that disrupts PtC binding develop into CD5^-PtC^- B cells, whereas those failing this process undergo clonal deletion. B cells that do not undergo receptor editing or have their attempts blocked (e.g., by dnRAG1 expression) retain PtC reactivity and develop into CD5^+PtC^- B cells.
| REAGENT or RESOURCE                      | SOURCE                  | IDENTIFIER          |
|------------------------------------------|-------------------------|---------------------|
| **Antibodies**                           |                         |                     |
| PE-CF594 Rat Anti-Mouse CD45R (clone RA3-6B2) | BD Biosciences          | Cat# 562290; RRID: AB_11151901 |
| APC-Cy7 Rat Anti-Mouse CD19 (clone 1D3)   | BD Biosciences          | Cat# 557655; RRID: AB_396770 |
| Biotin Rat Anti-Mouse CD43 (clone S7)     | BD Biosciences          | Cat# 553269; RRID: AB_2255226 |
| BV711 Rat Anti-Mouse Ig, κ, light chain (clone 187.1) | BD Biosciences          | Cat# 742837; RRID: AB_2741088 |
| FITC Rat Anti-Mouse Ig, A1, A2, & A3 light chain (clone R-26-46) | BD Biosciences          | Cat# 553434; RRID: AB_394854 |
| PerCP-Cy 5.5 Rat Anti-Mouse Ly6C (clone AL-21) | BD Biosciences          | Cat# 560525; RRID: AB_1727558 |
| Biotin Mouse Anti-Mouse IgM[a] (clone DS-1) | BD Biosciences          | Cat# 553515; RRID: AB_394896 |
| PE Mouse Anti-Mouse IgM[b] (clone AF6-78) | BD Biosciences          | Cat# 553521; RRID: AB_394902 |
| APC/Cyanine7 anti-mouse CD21/CD35 (CR2/CR1) (clone ZE9) | BioLegend               | Cat# 123417; RRID: AB_1953274 |
| Biotin Rat Anti-Mouse CD23 (clone B3B4)   | BD Biosciences          | Cat# 553137; RRID: AB_394652 |
| Brilliant Violet 421™ anti-mouse IgD (clone 11-26c.2a) | BioLegend               | Cat# 405725; RRID: AB_2562743 |
| CD4 Monoclonal Antibody (clone GK1.5), Alexa Fluor 700, eBioscience™ | Thermo Fisher Scientific (Invitrogen) | Cat# 56-0041-80; RRID: AB_494001 |
| CD5 Monoclonal Antibody (clone 53-7.3), PE, eBioscience | Thermo Fisher Scientific (Invitrogen) | Cat# 12-0051-82; RRID: AB_465523 |
| CD19 Monoclonal Antibody (clone eBio1D3 (1D3)), Alexa Fluor 700, eBioscience | Thermo Fisher Scientific (Invitrogen) | Cat# 56-0193-82; RRID: AB_837083 |
| CD49b (Integrin alpha 2) Monoclonal Antibody (clone DX5), PE-Cyanine7, eBioscience | Thermo Fisher Scientific (Invitrogen) | Cat# 25-5971-81; RRID: AB_469666 |
| CD93 (AA4.1) Monoclonal Antibody (clone AA4.1), PE, eBioscience | Thermo Fisher Scientific (Invitrogen) | Cat# 12-5892-83; RRID: AB_466019 |
| CD93 (AA4.1) Monoclonal Antibody (clone AA4.1), PE-Cyanine7, eBioscience | Thermo Fisher Scientific (Invitrogen) | Cat# 25-5892-81; RRID: AB_469658 |
| IgM Monoclonal Antibody (clone II/41), APC, eBioscience | Thermo Fisher Scientific (Invitrogen) | Cat# 17-5790-82; RRID: AB_469458 |
| IgM Monoclonal Antibody (clone II/41), PE-Cyanine5, eBioscience | Thermo Fisher Scientific (Invitrogen) | Cat# 15-5790-81; RRID: AB_494223 |
| IgD Monoclonal Antibody (clone 11-26c (11-26)), FITC, eBioscience | Thermo Fisher Scientific (Invitrogen) | Cat# 11-5993-85; RRID: AB_465347 |
| APC anti-mouse CD5 Antibody | BioLegend               | Cat# 100625; RRID: AB_2563928 |
| Mouse anti-MBP monoclonal antibody, unconjugated, clone 8G1 | Cell Signaling Technology | Cat#2396; RRID: AB_2140060 |
| Anti-MYC monoclonal antibody, clone 9E10 | Hybridoma | N/A |
| Mouse anti-beta-Actin monoclonal antibody, unconjugated, clone AC-74 | Sigma-Aldrich | Cat#A5316; RRID: AB_476743 |
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Horse anti-mouse IgG-HRP | Cell Signaling Technology | Cat#7076; RRID: AB_330924 |
| Chemicals, peptides, and recombinant proteins | | |
| BV510 Streptavidin | BD Biosciences | Cat# 563261; RRID: AB_2869477 |
| 200nm DOPC/CHOL Liposomes labeled with Fluorescein DHPE | FormuMax | Cat# F60103F2-F |
| 200nm Non-Fluorescent DOPC/CHOL Liposomes | FormuMax | Cat# F60103F2-P |
| TRIzol | Ambion, ThermoFisher Scientific | Cat# 15596018 |
| AMPure XP beads | Beckman Coulter Life Sciences | Cat# A63881 |
| Platinum™ Taq DNA Polymerase High Fidelity | Invitrogen | Cat# 11304011 |
| Precision Plus Protein Dual Color Standards | Bio-Rad Laboratories | Cat#1610374 |
| DEPC-Treated Water | ThermoFisher Scientific | Cat#AM9906 |
| DMEM (Dulbecco’s Modified Eagle’s Medium) | Corning | Cat# 10-017-CV |
| Penicillin-Streptomycin Solution | Corning (ThermoFisher Scientific) | Cat#MT30002CI |
| Fetal Bovine Serum | Gibco (ThermoFisher Scientific) | Cat#16-000-044 |

Critical commercial assays

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Phire Tissue Direct PCR Master Mix | Thermo Fisher Scientific | Cat# F-170S |
| dNTP Mix | Promega | Cat# U1515 |
| GoTag® Flexi DNA Polymerase | Promega | Cat# M8295 |
| SuperScript™ III First-Strand Synthesis System | Invitrogen, ThermoFisher Scientific | Cat#18080051 |
| Qubit dsDNA HS Assay kit | ThermoFisher Scientific | Cat# Q32851 |
| Bioanalyzer High Sensitivity DNA Analysis kit | Agilent | Cat# 5067-4626 |
| MiSeq Reagent Nano Kit v2 (500-cycles) | Illumina | Cat# MS-103-1003 |
| Phusion® High-Fidelity PCR Kit | New England Biolabs Inc. | Cat# E0553L |
| E.Z.N.A.® Plasmid DNA Mini Kit I | Omega Bio-Tek | Cat#D6942-02 |

Deposited data

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Sequence data and analyses related to evaluation of light chain repertoires of WT, dnRAG1, VH12, and dnRAG1 VH12 mice. | This paper | Gene Expression Omnibus (GEO): GSE165776; PRJNA697869 |
| Whole genome data and analyses related to determination of transgene sequence insertion in WT, dnRAG1, and VH12 mice. | This paper | Gene Expression Omnibus (GEO): GSE165776; PRJNA697869 |
| Code and other resources for sequencing and analysis of light chain repertoires of WT, dnRAG1, VH12, and dnRAG1 VH12 mice, and whole genome sequence of WT, dnRAG1, and VH12 mice. | This paper | GitHub https://github.com/swanson-lab/VH12-Sequencing |

Experimental models: Cell lines
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Human: HEK293T Cells | ATCC   | Cat#CRL-3216 |
| Experimental models: Organisms/strains |        |            |
| Mouse: dnRAG1 transgenic; C57Bl/6 background | Hassaballa et al., (2011) | N/A |
| Mouse: VH12 transgenic (line 6-1); C57Bl/6 background B6.Cg-Tg(Igh-V12CH27/IGH-6-1Sb/Ly) | Hongsheng Wang Lab | Jackson Laboratories: Cat#029738; RRID:IMSR_JAX:029738 |
| Oligonucleotides | Integrated DNA Technologies, Hassaballa et al. (2011) | N/A |
| Genotyping Primer: H2Kb Forward: GATCAGAACTCGGAGACGAC | Integrated DNA Technologies, Hongsheng Wang Lab, unpublished | N/A |
| Genotyping Primer: VH12: CTTCCTTACCTGCTCTATTACT | Integrated DNA Technologies, Hongsheng Wang Lab, unpublished | N/A |
| Sequencing primers, see Table S2 | Integrated DNA Technologies, This paper | N/A |
| Recombination assay Primer: 6624 Reverse: CCAAGGGTGGTATATCCAGTG | Integrated DNA Technologies | N/A |
| Recombination assay Primer: 6382 Reverse: CCAAGTGGGATTATTGTGGAG | Integrated DNA Technologies | N/A |
| Recombination assay Primer: PS CAT Reverse: CTTCACCCAGGATGCTGAG | Integrated DNA Technologies | N/A |
| Recombinant DNA |        |            |
| PhiX Control v3 | Illumina | Cat# FC-110-3001 |
| Software and algorithms |        |            |
| FlowJo v10.6.2 | BD | https://www.flowjo.com/ |
| IBM SPSS Statistics 27 | IBM | https://www.ibm.com/analytics/spss-statistics-software |
| Bcl2fastq | Illumina | https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html |
| pRESTO package | Vander Heiden et al. (2014) | https://presto.readthedocs.io/en/stable/ |
| IMGT/HighV-Quest | Alamyar et al. (2012); Li et al. (2013) | http://www.imgt.org/HighV-QUEST/home.action |
| Change-O package | Gupta et al. (2015) | https://changeo.readthedocs.io/en/stable/ |
| RStudioServer | R Core Team. R, 2019 | https://rstudio.com/products/rstudio/ |
| ALAKAZAM package | Stern et al. (2014) | https://alakazam.readthedocs.io/en/stable/ |
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Stringr             | Wickham et al. (2019) | https://stringr.tidyverse.org/ |
| Ggseqlogo           | Wagih (2017) | https://github.com/omarwagih/ggseqlogo |
| Ggplot2             | Wickham et al. (2019) | https://ggplot2.tidyverse.org/ |
| Repertoire Dissimilarity Index | Bolen et al. (2017) | https://rdi.readthedocs.io/en/1.0.0/ |
| Clustal Omega       | Madeira et al. (2019) | https://www.ebi.ac.uk/Tools/msa/clustalo/ |
| Recombination Signal Sequence Site (RSSsite) | Merelli et al. (2010) | https://www.itb.cnr.it/rss/ |
| aTRAM 2.0           | Allen et al. (2018) | https://github.com/juliema/aTRAM |
| SPAdes (3.15.3)     | Pryjelski et al. (2020) | https://cab.spbu.ru/software/spades/ |
| BLAST+ (2.12.0)     | Camacho et al. (2009) | https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastDocs&DOC_TYPE=Download |
| Trimmomatic (0.38)  | Bolger et al. (2014) | http://www.usadellab.org/cms/?page=trimmomatic |