



Published in final edited form as:

Nature. 2013 September 5; 501(7465): 112–115. doi:10.1038/nature12496.

## Microbial colonization influences early B-lineage development in the gut lamina propria

Duane R. Wesemann<sup>a,b,c,d</sup>, Andrew J. Portuguese<sup>a,b,c</sup>, Robin M. Meyers<sup>a,b,c</sup>, Michael P. Gallagher<sup>a,b,c</sup>, Kendra Cluff-Jones<sup>a,b,c</sup>, Jennifer M. Magee<sup>a,b,c</sup>, Rohit A. Panchakshari<sup>a,b,c</sup>, Scott J. Rodig<sup>e</sup>, Thomas B. Kepler<sup>f</sup>, and Frederick W. Alt<sup>a,b,c</sup>

<sup>a</sup>Program in Cellular and Molecular Medicine and Department of Medicine, Children's Hospital Boston, Boston, MA 02115, USA

<sup>b</sup>Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

<sup>c</sup>Howard Hughes Medical Institute, Boston, MA 02115, USA

<sup>d</sup>Division of Rheumatology, Allergy and Immunology, Department of Medicine, Boston, MA 02115, USA

<sup>e</sup>Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA

<sup>f</sup>Department of Microbiology, Boston University School of Medicine, Boston, MA 02215

### Abstract

The RAG1/RAG2 endonuclease ("RAG") initiates the V(D)J recombination reaction that assembles *Ig* heavy (*IgH*) and light (*IgL*) chain variable region exons from germline gene segments to generate primary antibody repertoires<sup>1</sup>. *IgH* V(D)J assembly occurs in progenitor (pro-) B cells followed by that of *IgL* in precursor (pre-) B cells. Expression of *IgH*  $\mu$  and *IgL* (*Ig* $\kappa$  or *Ig* $\lambda$ ) chains generates *IgM*, which is expressed on immature B cells as the B cell antigen-binding receptor ("BCR"). *Rag* expression can continue in immature B cells<sup>2</sup>, allowing continued *Ig* $\kappa$  V(D)J recombination that replaces the initial *V* $\kappa$ *J* $\kappa$  exon with one that generates a new specificity<sup>3–5</sup>. This "receptor editing" process, which also can lead to *Ig* $\lambda$  V(D)J recombination and expression<sup>3,6,7</sup>, provides a mechanism whereby antigen-encounter at the *Rag*-expressing immature B cell stage helps shape pre-immune BCR repertoires. As the major site of post-natal B cell development, the bone marrow is the principal location of primary *Ig* repertoire diversification in mice. Here, we report that early B cell development also occurs within the mouse intestinal

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Correspondence: Frederick W. Alt, [alt@enders.tch.harvard.edu](mailto:alt@enders.tch.harvard.edu), Duane R. Wesemann, [dwesemann@research.bwh.harvard.edu](mailto:dwesemann@research.bwh.harvard.edu), Ph: 617-919-2539; Fax: 617-730-0948.

#### Author Contributions

D.R.W. and F.W.A. designed the study; D.R.W., A.J.P., M.P.G., K. C-J., J.M.M., and R.A.P. performed experiments; R.M.M. and T.B.K performed computational analysis of sequencing data; S.R. performed immunohistochemistry experiments; D.R.W. and F.W.A. wrote the paper.

#### Competing financial interests

The authors declare no competing financial interests.

Microarray data have been deposited in MIAME format into the Gene Expression Omnibus (GEO) database under accession number GSE48870, and repertoire sequencing data have been deposited into the GEO database under accession number GSE48805.

lamina propria (LP), where the associated V(D)J recombination/receptor editing processes modulate primary LP *Ig* repertoires. At weanling age in normally housed mice, the LP contains a population of *Rag*-expressing B lineage cells that harbor intermediates indicative of ongoing V(D)J recombination and which contain cells with pro-B, pre-B, and editing phenotypes. Consistent with LP-specific receptor editing, *Rag*-expressing LP B-lineage cells have similar  $V_H$  repertoires, but significantly different  $V_K$  repertoires, compared to those of *Rag2*-expressing BM counterparts. Moreover, colonization of germ-free mice leads to an increased ratio of *Igλ*-expressing versus *Igκ*-expressing B cells specifically in the LP. We conclude that B cell development occurs in the intestinal mucosa, where it is regulated by extra-cellular signals from commensal microbes that influence gut *Ig* repertoires.

---

Pre-immune *Ig* diversification occurs within the gut or gut associated structures in several vertebrate species, including sheep, rabbits, cattle, pigs and chicken<sup>8–10</sup>. In these species, cells harboring a limited RAG-mediated V(D)J repertoire migrate to gut-associated structures to undergo further diversification through somatic mutation and/or gene conversion to generate a full pre-immune *Ig* repertoire<sup>11</sup>. These examples raise the notion that the gut environment may provide some benefit to the process of primary *Ig* diversification. In this regard, commensal bacteria are required for primary antibody repertoire diversification in pigs and rabbits<sup>8,12</sup>, and may play an important role in stimulating this process in cattle, sheep and chickens shortly after birth<sup>11</sup>. In contrast to the above species, RAG-mediated V(D)J recombination is the major driver of pre-immune diversification in mice and humans. In this context, stimulated by our prior finding of mouse B-lineage tumors that arise in mesenteric lymph nodes from apparent receptor-editing B cells<sup>13,14</sup>, we hypothesized that B cell development might occur in the mouse gut and, thereby, allow *Rag* expression to diversify gut B cell pre-immune repertoires.

To test for RAG expression, we used our *Rag2* reporter mice, which contain a functional *Rag2-Gfp* fusion gene within the endogenous *Rag2* locus that provides *Rag2* expression functionally equivalent to that of the endogenous *Rag2* gene<sup>15,16</sup>. We employed flow cytometry to test for RAG2-GFP in lymphocytes from mesenteric lymph nodes (mLN), small intestinal (SI) lamina propria (LP), and intraepithelial lymphocytes (IEL) of 3 wk-old mice. Cells were gated on the CD19 pan-B lineage marker, and GFP was plotted against the B220 pan-B lineage marker. Staining with dual B cell markers was done to optimize true GFP signal over background auto-fluorescence, which in wild-type controls was approximately 0.1% (Fig. 1a). With this method, we found essentially no *Rag2*-expressing B lineage cells in the IEL and mLN (Fig. 1a). However, we did find a population of *Rag2*-expressing, CD19<sup>+</sup> B220<sup>low</sup> cells within the LP that comprised approximately 3% of total CD19<sup>+</sup> cells (Fig. 1a). Quantitative PCR revealed *Rag1* and *Rag2* expression in wild type small intestinal LP at a level of about 1–10% that of total bone marrow (BM), but little or no *Rag1* or *Rag2* expression in mLN or IEL cells (Supplementary Fig. 1), confirming the flow cytometry results found with the *Rag2-Gfp* reporter mice. Large intestinal LP contained GFP<sup>+</sup> B lineage cells as well, but at a lower level compared to that in the SI LP (Supplementary Fig. 2). We did not find RAG2-GFP in Peyer's patch B cells (Supplementary Fig. 3) or mucosal T cells (Supplementary Fig. 4).

We examined various stages of early post-natal development to determine if levels of RAG2<sup>+</sup> B lineage cells in the gut LP change over time. The proportion of LP RAG2-GFP<sup>+</sup> cells among total CD19<sup>+</sup> cells was low (<0.5%) in the first week of life; however, after that it gradually increased with levels peaking at approximately 4% at age 18–23 days before decreasing to undetectable levels by post-natal day 35 (Fig. 1b). In contrast, the CD19<sup>+</sup> B cell population of peripheral blood contained 20–40% RAG2-GFP<sup>+</sup> cells during the first week of life, which then decreased over time to undetectable levels over the next 4 weeks (Fig. 1c). Similarly, RAG2-GFP<sup>+</sup> cell levels in the spleen appeared highest (10–15%) in the first week of life before decreasing to undetectable levels (Fig. 1d). The finding of low (<0.5%) levels of RAG2-GFP<sup>+</sup> cells in the gut LP in the first week of life, despite the presence of substantial proportions of RAG2-GFP<sup>+</sup> cells in the peripheral blood and spleen suggests that the mechanism responsible for the later emergence of RAG2<sup>+</sup> cells in the gut may not be due to non-specific dissemination driven by high levels of these cells in the blood. As RAG2<sup>+</sup> LP B lineage cells do not express proteins known to promote gut lymphocyte tropism such as the  $\alpha 4\beta 7$  integrin or the CCR9 chemokine receptor (Supplementary Fig. 5), mechanisms underlying their appearance in the gut remain to be determined.

Sixteen days following intraperitoneal (i.p.) alum injection, RAG2-GFP<sup>+</sup> cells accumulate in the peripheral blood and spleens of adult mice due to increased bone marrow output following initial alum-mediated bone marrow suppression<sup>17,18</sup>. To determine if the gut LP in adult mice maintains ability to support RAG2<sup>+</sup> B lineage cells, we injected 4–6 month-old *Rag2-Gfp* mice with i.p. alum and examined gut tissues on day 16. Following alum injection, low levels of RAG2-GFP<sup>+</sup> B lineage cells appeared in IEL, mLN and PP; however, the most striking accumulation was in the LP, where RAG2-GFP<sup>+</sup> cells made up about 2.5% of total CD19<sup>+</sup> cells (Supplementary Fig. 6a, b). Appearance of RAG<sup>+</sup> B lineage cells in the spleen and blood following alum injection is mediated by tumor necrosis factor alpha (TNF $\alpha$ )<sup>19</sup>. To determine whether appearance of RAG<sup>+</sup> B lineage cells in the small intestinal LP at weaning age is also TNF $\alpha$ -dependent, we measured LP *Rag1* and *Rag2* expression by qPCR in 3 wk-old *Tnf $\alpha$*  knock-out and WT control mice and found no differences (Supplementary Fig. 6c, d). Thus, the mechanism of gut LP RAG<sup>+</sup> B lineage cell accumulation at weaning age appears distinct from that of peripheral RAG<sup>+</sup> cell accumulation that occurs after alum immunization.

RAG-expressing B lineage cells in the BM comprise a heterogeneous population of early developmental subsets including pro-B, pre-B, and immature B cells undergoing receptor editing<sup>20</sup>. The expression patterns of I $\mu$  and I $\kappa$  (which accounts for ~95% of mouse IgL) can be used to distinguish these groups<sup>5,21</sup>. In this context, productive assembly of *Igh* V(D)J exons in pro-B cells leads to the cytoplasmic I $\mu$ <sup>+</sup> pre-B cell stage<sup>21</sup>. Assembly of *Ig $\kappa$*  VJ exons in pre-B cells leads to formation of IgM and differentiation to the surface IgM<sup>+</sup> immature B cell stage. *Rag*-expressing cells with cytoplasmic I $\kappa$  and low or absent surface IgM have been defined as immature B cells undergoing receptor editing<sup>2,5,17</sup>. Staining of fixed/permeabilized CD19<sup>+</sup> B220<sup>low</sup> RAG2-GFP<sup>+</sup> LP B cells for cytoplasmic I $\mu$  and I $\kappa$  revealed similar relative levels of pro-B cells (I $\mu$ <sup>-</sup>, I $\kappa$ <sup>-</sup>), pre-B cells (I $\mu$ <sup>+</sup>, I $\kappa$ <sup>-</sup>) and editing B cells (I $\mu$ <sup>+</sup>, I $\kappa$ <sup>+</sup>, RAG2<sup>+</sup>), respectively, to those of the BM (Fig. 2a and

Supplementary Fig. 7). Live CD19<sup>+</sup> B220<sup>low</sup> RAG2-GFP<sup>+</sup> LP B cells that were surface IgM<sup>+</sup> also had similarly low IgM levels to those of this putative editing B-lineage subset in BM (Fig. 2a and Supplementary Fig. 7). In addition, ligation-mediated qPCR further showed that sorted RAG2-GFP<sup>+</sup> LP B cells had a similar level of RAG-dependent DNA double-strand breaks at *J $\kappa$*  as RAG2-GFP<sup>+</sup> BM B lineage cells (Fig. 2b), demonstrating that *Ig $\kappa$*  V(D)J recombination takes place in the RAG2<sup>+</sup> LP B cells at levels similar to that in RAG2<sup>+</sup> BM B lineage cells. Finally, microarray analysis of RAG2<sup>+</sup> B lineage cells in weanling age gut LP revealed no significant differences in general transcript expression profiles in RAG2<sup>+</sup> LP and RAG2<sup>+</sup> BM B lineage cells (Supplementary Fig. 8), demonstrating a strong similarity between RAG2<sup>+</sup> cells in these two sites. Overall, these data indicate that the LP *Rag2*-expressing B lineage cells contain early B lineage developmental subsets representative of those found in the BM, supporting the occurrence of B cell development in the LP.

We performed immunohistochemistry (IHC) to confirm the LP localization of early B lineage cells in mice at post-natal day 18–23. The terminal deoxynucleotidyl transferase (TdT) enzyme is present in pro-B cells and mediates addition of random nucleotides to *IgH* V(D)J DNA ends during V(D)J recombination to increase *IgH* V(D)J junctional diversity<sup>22</sup>. Staining for TdT, which is expressed similarly in BM and LP RAG2<sup>+</sup> B lineage cells (Supplementary Fig. 8), revealed a dense nuclear expression pattern (brown) in SI LP cells similar to TdT<sup>+</sup> cells in the bone marrow (BM) (Fig. 1e). SI sections were also subjected to dual staining for B220 plus TdT, which showed that the TdT<sup>+</sup> cells were also faintly B220-positive, analogous to TdT<sup>+</sup> cells in the BM (Supplementary Fig. 9). There were also cells in the SI lamina propria that stained strongly for B220 but were TdT<sup>-</sup>, representing more mature B cell populations (Supplementary Fig. 9, black arrows). TdT<sup>+</sup> cells were distributed throughout the LP, including within villi; however, they generally appeared to be more proximal to bases of villi, closer to the serosal (antiluminal) intestinal surface compared to B220<sup>high</sup> TdT<sup>-</sup> B cells (Supplementary Fig. 10). These data indicate that gut-resident early B lineage cells inhabit a generally distinct location within the LP compared to more mature B cells. IHC studies of human fetal intestine also identified serosally-positioned pre-B cells in the intestinal LP, suggesting similar early B cell development may occur in the human gut<sup>23</sup>.

Given our finding of primary B cell development in the mouse intestinal LP, we asked whether this process contributes to differential diversification of pre-immune repertoires in developing RAG2<sup>+</sup> LP B lineage cells versus those from BM. To test this, we isolated RNA from sorted RAG2-GFP<sup>+</sup> cells from the BM and LP (Supplementary Fig. 11) of 3 week-old *Rag2-Gfp* mice, and assessed productive *V<sub>H</sub>* and *V $\kappa$*  utilization via 5' rapid amplification of cDNA ends (5'RACE) generated from mature *Ig* gene transcripts, followed by 454 sequencing. To best visualize potential *V* segment utilization differences between RAG2-GFP<sup>+</sup> cells from LP versus BM, we plotted *V* segment utilization from the two sources against individual in-frame *V* segments in the order of highest to lowest utilization in RAG2-GFP<sup>+</sup> BM cells (Fig. 3, Supplementary Fig. 12). These studies showed that *V<sub>H</sub>* usage was very similar in RAG2-GFP<sup>+</sup> LP cells and RAG2-GFP<sup>+</sup> BM cells ( $\chi^2$  test,  $p=0.235$ ) (Fig. 3a, Supplementary Fig. 12a). As a positive control for the method, comparison of *V<sub>H</sub>* usage

between RAG2<sup>+</sup> BM and total non-sorted splenic B cells showed an expected highly significant difference ( $\chi^2$  test,  $P=2.2\times 10^{-16}$ ) (Supplementary Figure 13a). The similar  $V_H$  repertoires of RAG2<sup>+</sup> BM and LP B lineage cells suggests that  $V_H$  utilization during primary IgH V(D)J recombination occurs similarly in the two locations. In contrast, we observed prominent and highly significant differences in  $V_K$  usage in RAG2-GFP<sup>+</sup> cells from LP versus BM ( $\chi^2$  test,  $P=0.00084$ ) (Fig. 3b, Supplementary Fig. 12b); as a negative control, we observed no significant overall differences in  $V_K$  usage between RAG2<sup>+</sup> BM samples compared with other RAG2<sup>+</sup> BM samples; or RAG2<sup>+</sup> LP samples compared with other RAG2<sup>+</sup> LP samples from separate pools of mice ( $\chi^2$  test,  $P=0.560$  and  $0.545$ , respectively) (Supplementary Figure 13b, c). The finding of different  $V_K$  repertoires within RAG2<sup>+</sup> LP and BM B lineage cells indicates that the LP versus BM location of B cell development may influence  $V_K$  usage in developing B cells. In this regard, the very similar  $V_H$  repertoires of LP and BM RAG2<sup>+</sup> B lineage cell suggests that a likely explanation for the marked difference in the LP  $V_K$  repertoires from those of BM is that they are generated in the RAG2<sup>+</sup> receptor-editing LP B cell population.

Intestinal microflora have been shown to influence immune cell development in terms of lymphoid organization and T cell subset accumulation and activity, both locally in the gut as well as systemically<sup>24</sup>. To assess influences of microflora on B cell development, we co-housed 3 week-old WT Swiss-Webster germ-free mice with regular specific pathogen free (SPF) mice for 7 days. Gram staining of SI contents was performed to confirm bacterial colonization of co-housed mice (Supplementary Fig. 14). We employed qPCR to assay colonized mice and germ-free littermates for *Rag1* and *Rag2* expression, which was normalized to *Cd19* expression. In accord with ability of microflora to induce systemic effects on immune cell development<sup>24</sup>, we observed colonization-dependent increases in *Rag1* and *Rag2* expression in the BM and spleen, as well as the gut LP (Supplementary Fig. 15). We also found pro-B cells (identified as CD19<sup>+</sup> B220<sup>low</sup> CD43<sup>+</sup>) to represent an increased percentage of total CD19<sup>+</sup> B cells in colonized mouse BM and LP, and potentially in the colonized spleen (Fig. 4a, Supplementary Fig. 16), correspondingly, the increased *Rag* expression in these tissues likely derives from the increased percentage of pro-B cells. We conclude that gut microflora induce increased levels of pro-B cells systemically, including in the gut LP, in previously un-colonized weaning-age mice.

Given our finding of different  $V_K$  repertoires in LP versus BM RAG2<sup>+</sup> B lineage populations that might be generated via receptor editing, we examined the ratio of Ig $\lambda$ <sup>+</sup> to Ig $\kappa$ <sup>+</sup> B cells in the LP of colonized mice versus their germ-free mice littermates. In this regard, increased Ig $\lambda$  usage in the B cell repertoire is another marker of receptor editing<sup>3,6,7</sup>. Notably, colonization led to a significant and reproducible increase of the ratio of Ig $\lambda$ <sup>+</sup> to Ig $\kappa$ <sup>+</sup> B cells in the LP but not the BM or SPL (Fig. 4b and Supplementary Fig. 17), consistent with a commensal-dependent process leading to increased editing specifically within LP B cells. However, as these analyses were not performed in *Rag2-Gfp* mice, only total B cell populations could be analyzed. Therefore, a non-mutually exclusive possibility would be selection for Ig $\lambda$ <sup>+</sup> B cells in the gut after colonization with commensal microbes, a phenomenon not previously described. In this regard, we do find significant differences in both  $V_H$  and  $V_K$  segment usage in the LP IgM<sup>+</sup> B cell population of colonized mice

compared to germ-free littermates (Supplementary Fig. 18), indicating presence of commensals influences both IgH and IgL mature B cell repertoires.

Consistent with growing evidence demonstrating that the microflora act as a regulators of T lymphocyte subsets<sup>24</sup>, we find that weanlings harbor an intestinal LP B cell developmental process that is influenced in germ-free mice by microbial colonization. As weaning is concurrent with microbial expansion<sup>25</sup>, occurrence of *Rag*-expressing B lineage cell accumulation in weanlings may have evolved to allow B cell primary repertoires to be modulated in response to colonization. In this regard, our findings also suggest that this LP B cell developmental process includes BCR editing, which may contribute to significant differences in the primary *V<sub>k</sub>* repertoire of LP versus BM B lineage populations. Past studies have implicated BCR editing in the BM as a negative selection process, largely based on studies of B cells engineered to make specific self-reactive, high affinity BCRs. Given the natural repertoire in our studies, the degree to which the observed BCR editing process in the gut represents a tolerance mechanism is unclear. However, given the potential special role of the gut in pre-immune diversification in other vertebrates, RAG-dependent editing in the LP may also contribute diversification-related roles. In addition, immature gut-derived B cells also may have specialized roles, such as those suggested for immature splenic B cells<sup>26</sup>. Finally, primary B cell development in the intestine, including mucosal B cell receptor editing, might allow both luminal antigens and peripheral host mucosal components opportunities to shape the pre-immune repertoire. In this regard, the transient nature of LP B cell development implies that there may be windows of opportunity for this influence to occur.

## METHODS

### Mice, Immunizations and Colonization

Mice harboring the *Rag2-Gfp* knock-in fusion gene at the endogenous locus were described previously<sup>15</sup> and were maintained on a 129/SvJ background. Wild type Balb/c mice were purchased from Jackson Laboratories. Swiss-Webster germ-free mice were purchased from Taconic Farms. For each germ-free/colonization experiment, littermate germfree mice were used as controls. Germ-free status and colonization status were confirmed by gram staining of cecal and small intestinal contents as well as microbial culture. Immunization experiments were performed with alum as described<sup>18</sup>. All experiments with mice followed the protocols approved by the Boston Animal Care Facility of the Children's Hospital, Boston, MA 02115.

### Cell Isolation and Flow Cytometry

PP, IEL and LP lymphocytes were isolated essentially as described<sup>27</sup>. PPs were excised from the small intestine, and the remaining tissue was incubated with 1×Hank's balanced salt solution with 1 mM EDTA/10% FBS/PBS for 30 min and room temperature three times for IEL extraction. Residual intestinal tissue was digested in 20% FBS RPMI with 0.05% collagenase from *Clostridium histolyticum* (Sigma) for 1 hr at 37 degrees Celsius three times. IELs and LP cells were centrifuged over Lympholyte (Cedar Lane) per manufacturer's recommendations to minimize mucus contamination. Single-cell suspensions

of mLNs, PPs and spleen were prepared by mashing through a cell strainer (70  $\mu$ m). Cells were stained with fluorophore-conjugated mouse antibodies, and flow cytometry was performed.

### Real-Time qPCR and Microarray Analysis

For quantitative PCR, total RNA was extracted using the TRIzol method (Invitrogen) and reverse transcribed into cDNA using qScript (Quanta Biosciences). *Rag1* and *Rag2* transcripts were then quantified using Taqman qPCR assays Mm01270936\_m1 and Mm00501300\_m1, respectively (Applied Biosystems). The comparative Ct method was used to quantify transcripts that were normalized with respect to *Cd19* expression (Taqman assay Mm00515420\_m1, Applied Biosystems). For comparative transcriptome analysis, B cells were isolated from BM and small intestinal LP of 3 wk-old mice. Cells were sorted (BD FACSAria) into TRIzol on the basis of the following cell surface markers: CD19<sup>+</sup>, B220<sup>low</sup>, GFP<sup>+</sup>. RNA was extracted and then amplified, labeled, and hybridized to Affymetrix GeneChip Mouse Gene 1.0 ST arrays (Expression Analysis, Durham, NC). Raw data were normalized with the RMA algorithm implemented in the Expression File Creator module from the GenePattern suite<sup>28</sup>. Data were visualized with the Multiplot and Hierarchical Clustering Viewer modules. All cell populations analyzed were generated in triplicate from independent experiments consisting of a pool of at least 8 mice for each experiment.

### Immunohistochemistry

Immunohistochemistry was performed using 4-mm-thick formalin fixed paraffin embedded (FFPE) tissue sections. Slides were soaked in xylene, passed through graded alcohols, and put in distilled water. Slides were pretreated with EDTA (pH 8.0) retrieval solution (Zymed, South San Francisco, CA) in a steam pressure cooker (Biocare Decloaking Chamber CD2008US, Biocare Biomedical, Concord, CA) at manufacturers recommended settings. All further steps are performed at room temperature in a hydrate chamber. The slides were blocked for endogenous peroxidase activity with peroxidase block (DAKO), washed 5 minutes in buffer, and followed by 20 minute incubation with serum free protein block (DAKO). For TdT single staining, a polyclonal rabbit antibody (DAKO Cat #A3524) was applied at 1:100 dilution for 1 hour at room temperature followed by washing. The detection of antibody utilized DAKO Rabbit Envision and DAB according to the manufacturer's directions. For TdT/B220 double staining, rabbit anti-TDT was followed with Mach-2 Rabbit AP polymer (Biocare) and developed with Vulcan Fast Red (Biocare). Subsequently, rat anti-B220 (BD Pharmingen, cat # 550286), was applied for 1 hour at 1:200 dilution followed by Goat anti-Rat-HRP (Millipore) and developed with DAB. All slides were counterstained with Harris hematoxylin. Stained slides were scanned at 200 $\times$  magnification using an Aperio ScanScope XT workstation (Aperio Technology, Inc., Vista, CA). Images were visualized, annotated, and microscopic distances quantified using ImageScope software (version 10.0.35.1800, Aperio Technology).

## Ligation-mediated PCR

Sorted CD19<sup>+</sup> B220<sup>low</sup> RAG2-GFP<sup>+</sup> B lineage cells from BM and LP were lysed in SDS lysis buffer (5mM EDTA, 200mM NaCl, 100mM Tris-HCl pH 8.0, 0.2% SDS) with proteinase K (200mg/ml) overnight at room temperature followed by incubation at 37°C for one hour. Following DNA isolation by phenol:chloroform separation and isopropanol precipitation, blunt-end ligation reactions were performed using an oligonucleotide duplex linker consisting of BW-1 (5' GCGGTGACCCGGGAGATCTGAATTC) and BW-2 (5' GAATTCAGATC). DNA was ligated overnight at 16°C in ligation buffer (50mM Tris pH7.5, 10mM MgCl<sub>2</sub>, 10mM DTT, 1mM ATP) and T4 DNA ligase (Promega). Ligase was inactivated by incubation for 10 min at 70°C. Ligation reaction was diluted 1:3 in H<sub>2</sub>O prior to being used for PCR. Nested PCR was used to detect ligation products resulting from both *Jκ1* and *Jκ2* double stranded DNA breaks. The first round of amplification was performed using Qiagen Hot Star Taq (1.25U/rxn) and primers Ko3 (5' AGTGCCACTAACTGCTGAGAAACCT) and BW-1H (5' CCGGGAGATCTGAATTCCAC). The PCR reaction was performed as follows: 95°C for 15min, followed by 26 cycles of 94°C for 45 sec, 57°C for 45 sec and 72°C for 50 sec, followed by 72°C for 5 min. The second round of amplification was performed as quantitative PCR using primers Ko (5' CCACGCATGCTTGGAGAGGGGGTT) and BW-1H primers with an internal Jκ probe (5' 56-FAM/ZEN-3-Iowa Black 5' TGAGGAGGGTTTTTGTACAGCCAGA). Signals were normalized to actin amplified from genomic DNA (Mm00607939\_s1). Each sample was calculated as a percent of a BM standard run on each PCR plate to determine variability within BM and LP samples from three biologic replicates isolated from independent pools of 4–8 mice per experiment.

## Repertoire Sequencing

Total RNA was obtained from purified B cells using TRIzol reagents (Invitrogen). κ-chain and μ chain cDNAs from each sample were synthesized using a SMARTer™-RACE cDNA amplification kit (Clontech), according to the manufacturer's protocol. C<sub>μ</sub>-specific (5'-CAGGTGAAGGAAATGGTGCT) and C<sub>κ</sub>-specific (5'-TTAACTGCTCACTGGATGGTG) primers were used in lieu of oligo dT primers for cDNA synthesis. A total of 0.05 to 0.2 μg total RNA/sample was used. PCR was performed using Phusion DNA Polymerase (Thermo Scientific) and 12.5 ml of first-strand reactions with long and short universal primers (CTAATACGACTCACTATAGGGCAAGCAGTGGTAACAACGCAGAGT and CTAATACGACTCACTATAGGGC) together with either a biotinylated round-1 C<sub>μ</sub> primer (BIO 5'-CTTATCAGACAGGGGGCTCTC) or round-1 C<sub>κ</sub> primer (BIO- 5'-TCACTGGATGGTGGGAAGAT) specific primers. First round PCR reaction conditions were then followed as described elsewhere<sup>29</sup>. PCR round-1 product sizes of 500–700 bp were extracted from agarose gels using a QIAquick gel extraction kit (Qiagen), enriched on streptavidin-coupled Dynabeads (Invitrogen), and purified with a QIAquick PCR purification columns (Qiagen) per manufacturers instructions. Purified round-1 products were then subjected to a second round of PCR with nested primers containing "A" and "B" adapter sequences as well as distinct 10 bp barcode sequences (to distinguish source material) using the nested universal primer (5'-CGTATCGCCTCCCTCGCGCCATCAG[UNIQUE 10BP BARCODE



SEQUENCE]ACGACTCACTATAGGGCAAGCAG) together with either nested C $\mu$  primer (CTATGCGCCTTGCCAGCCCGCTCAG[UNIQUE 10BP BARCODE SEQUENCE]GGGAAGACATTTGGGAAGGA) or C $\kappa$  primer (CTATGCGCCTTGCCAGCCCGCTCAG[10 bp UNIQUE BARCODE SEQUENCE]TGGATGGTGGGAAGATGGAT). PCR products (size 500–700) were extracted from agarose gel, and 100 ng of each amplicon library was combined and used for 454 sequencing analysis. GS FLX Titanium sequencing kit XLR70 (Roche) was used for sample preparation. Data were collected at the sequencing core and at the University of Illinois. Data were analyzed using the empirical Bayes procedure as described<sup>30</sup>. Clonotypes and clone assignments were determined using a recursive set of hypothesis tests on the equality of the V-gene segment mutation rate and that of CDR3. To control for false detection rate, comparisons were made between the same tissues of repeat experiments of BM and LP V $\kappa$  samples (Supplementary Fig. 13). V $H$  segment usage from RAG2<sup>+</sup> BM was also compared to total splenic B cell V $H$  segment usage to ensure that our analysis could identify the expected differences that occur due to selection between these populations.

### Statistics

If not otherwise stated, data were expressed as arithmetic means  $\pm$  s.e.m., and statistical analyses were made by unpaired *t*-test, exact test, or  $\chi^2$  test where appropriate.  $P < 0.05$  was considered statistically significant.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

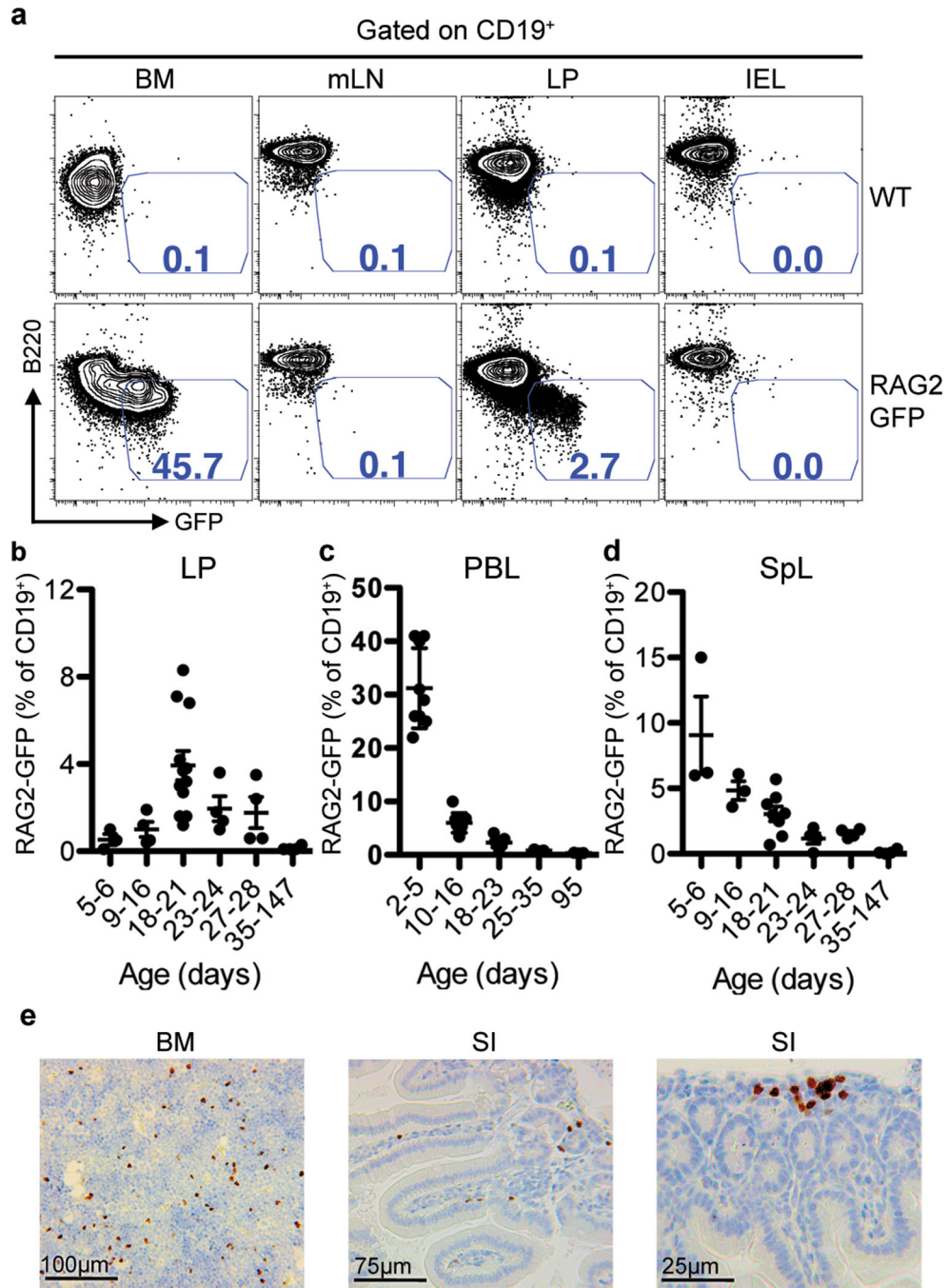
### Acknowledgements

This work was supported by National Institute of Health grants AI020047 (to F.W.A.) and AI89972 (to D.R.W.), Lymphoma and Leukemia SCOR 7009-12 (to F.W.A.), and National Institutes of Health research contract HHSN272201000053C (to T.B.K.). D.R.W. was also supported by an award from the American Academy of Allergy Asthma and Immunology and CSL-Behring and holds a Career Award for Medical Scientists from the Burroughs Wellcome Fund. F.W.A. is an Investigator of the Howard Hughes Medical Institute.

### References

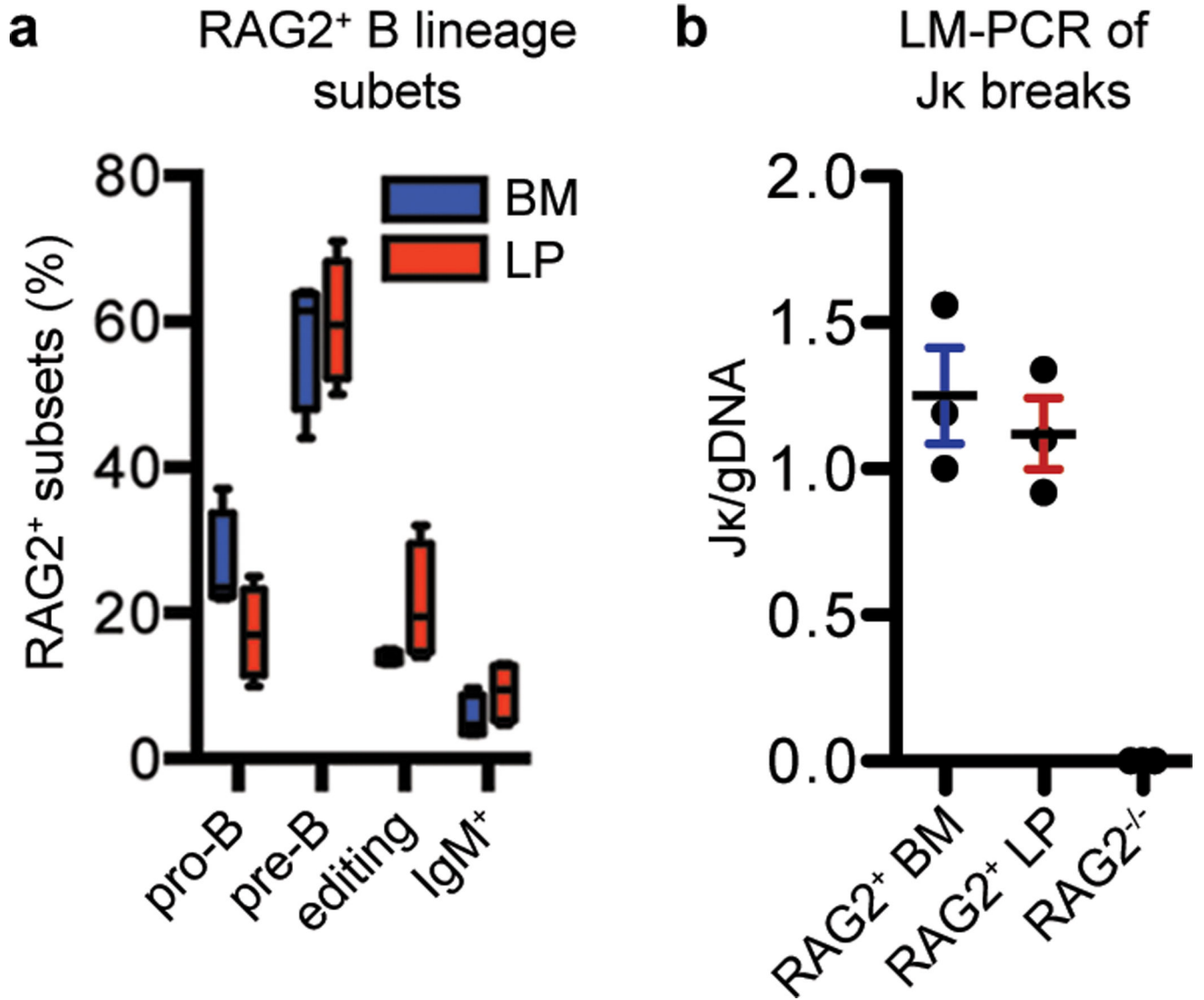
1. Jung D, Giallourakis C, Mostoslavsky R, Alt FW. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol.* 2006; 24:541–570. [PubMed: 16551259]
2. Yu W, et al. Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. *Nature.* 1999; 400:682–687. [PubMed: 10458165]
3. Tiegs SL, Russell DM, Nemazee D. Receptor editing in self-reactive bone marrow B cells. *J Exp Med.* 1993; 177:1009–1020. [PubMed: 8459201]
4. Gay D, Saunders T, Camper S, Weigert M. Receptor editing: an approach by autoreactive B cells to escape tolerance. *J Exp Med.* 1993; 177:999–1008. [PubMed: 8459227]
5. Pelanda R, et al. Receptor editing in a transgenic mouse model: site, efficiency, and role in B cell tolerance and antibody diversification. *Immunity.* 1997; 7:765–775. [PubMed: 9430222]
6. Retter MW, Nemazee D. Receptor editing occurs frequently during normal B cell development. *J Exp Med.* 1998; 188:1231–1238. [PubMed: 9763602]
7. Hertz M, Nemazee D. BCR ligation induces receptor editing in IgM+IgD– bone marrow B cells in vitro. *Immunity.* 1997; 6:429–436. [PubMed: 9133422]

8. Lanning D, Zhu X, Zhai SK, Knight KL. Development of the antibody repertoire in rabbit: gut-associated lymphoid tissue, microbes, and selection. *Immunol Rev.* 2000; 175:214–228. [PubMed: 10933605]
9. Jenne CN, Kennedy LJ, Reynolds JD. Antibody repertoire development in the sheep. *Dev Comp Immunol.* 2006; 30:165–174. [PubMed: 16083958]
10. Ratcliffe MJ. Antibodies, immunoglobulin genes and the bursa of Fabricius in chicken B cell development. *Dev Comp Immunol.* 2006; 30:101–118. [PubMed: 16139886]
11. Lanning DK, Rhee KJ, Knight KL. Intestinal bacteria and development of the B-lymphocyte repertoire. *Trends Immunol.* 2005; 26:419–425. [PubMed: 15967718]
12. Butler JE, et al. Antibody repertoire development in fetal and neonatal piglets. VIII. Colonization is required for newborn piglets to make serum antibodies to T-dependent and type 2 T-independent antigens. *J Immunol.* 2002; 169:6822–6830. [PubMed: 12471114]
13. Wang JH, et al. Mechanisms promoting translocations in editing and switching peripheral B cells. *Nature.* 2009; 460:231–236. [PubMed: 19587764]
14. Wang JH, et al. Oncogenic transformation in the absence of *Xrcc4* targets peripheral B cells that have undergone editing and switching. *J Exp Med.* 2008; 205:3079–3090. [PubMed: 19064702]
15. Monroe RJ, et al. RAG2:GFP knockin mice reveal novel aspects of RAG2 expression in primary and peripheral lymphoid tissues. *Immunity.* 1999; 11:201–212. [PubMed: 10485655]
16. Nagaoka H, Yu W, Nussenzweig MC. Regulation of RAG expression in developing lymphocytes. *Curr Opin Immunol.* 2000; 12:187–190. [PubMed: 10712943]
17. Nagaoka H, Gonzalez-Aseguinolaza G, Tsuji M, Nussenzweig MC. Immunization and infection change the number of recombination activating gene (RAG)-expressing B cells in the periphery by altering immature lymphocyte production. *J Exp Med.* 2000; 191:2113–2120. [PubMed: 10859336]
18. Gartner F, Alt FW, Monroe RJ, Seidl KJ. Antigen-independent appearance of recombination activating gene (RAG)-positive bone marrow B cells in the spleens of immunized mice. *J Exp Med.* 2000; 192:1745–1754. [PubMed: 11120771]
19. Ueda Y, Yang K, Foster SJ, Kondo M, Kelsoe G. Inflammation controls B lymphopoiesis by regulating chemokine CXCL12 expression. *J Exp Med.* 2004; 199:47–58. [PubMed: 14707114]
20. Jankovic M, Casellas R, Yannoutsos N, Wardemann H, Nussenzweig MC. RAGs and regulation of autoantibodies. *Annu Rev Immunol.* 2004; 22:485–501. [PubMed: 15032586]
21. Raff MC, Megson M, Owen JJ, Cooper MD. Early production of intracellular IgM by B-lymphocyte precursors in mouse. *Nature.* 1976; 259:224–226. [PubMed: 814466]
22. Desiderio SV, et al. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. *Nature.* 1984; 311:752–755. [PubMed: 6092963]
23. Golby S, et al. B cell development and proliferation of mature B cells in human fetal intestine. *J Leukoc Biol.* 2002; 72:279–284. [PubMed: 12149418]
24. Hooper LV, Littman DR, Macpherson AJ. Interactions between the microbiota and the immune system. *Science.* 2012; 336:1268–1273. [PubMed: 22674334]
25. Mackie RI, Sghir A, Gaskins HR. Developmental microbial ecology of the neonatal gastrointestinal tract. *Am J Clin Nutr.* 1999; 69:1035S–1045S. [PubMed: 10232646]
26. Ueda Y, Liao D, Yang K, Patel A, Kelsoe G. T-independent activation-induced cytidine deaminase expression, class-switch recombination, and antibody production by immature/transitional 1 B cells. *J Immunol.* 2007; 178:3593–3601. [PubMed: 17339456]
27. Lefrancois L, Lycke N. Isolation of mouse small intestinal intraepithelial lymphocytes, Peyer's patch, and lamina propria cells. *Curr Protoc Immunol.* 2001; Chapter 3(Unit 3):19. [PubMed: 18432783]
28. Reich M, et al. GenePattern 2.0. *Nat Genet.* 2006; 38:500–501. [PubMed: 16642009]
29. Warren RL, et al. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res.* 2011; 21:790–797. [PubMed: 21349924]
30. Kepler TB. Reconstructing a B-cell clonal lineage. I. Statistical inference of unobserved ancestors [v1; ref status: indexed, <http://f1000r.es/z6>]. *F1000 Research.* 2013; 2:1–12. [PubMed: 24358860]



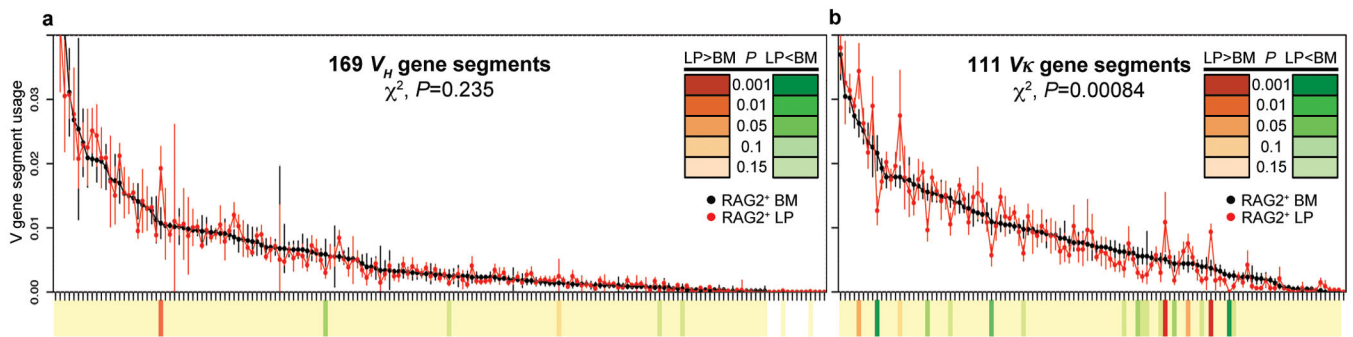
**Figure 1. Gut LP RAG2<sup>+</sup> B Lineage Cells in Weanling Age Mice**

**a**, FACS plots of CD19<sup>+</sup> cells from the indicated tissues taken from wild type (WT) (top) or homozygous *Rag2-Gfp* knock-in (bottom) mice. B220 expression is plotted against GFP fluorescence. Numbers denote percentage of CD19<sup>+</sup> B220<sup>low</sup> RAG2-GFP<sup>+</sup> cells. **b–d**, dot plots showing percentage of RAG2-GFP<sup>+</sup> cells in listed tissues from indicated post-natal ages. Each point represents one mouse. Horizontal bars indicate mean values ± s.e.m. **e**, Immunohistochemistry of paraffin-embedded sections from bone marrow (BM) and small intestine (SI) stained with an anti-TdT antibody. Dark brown indicates TdT-reactivity.

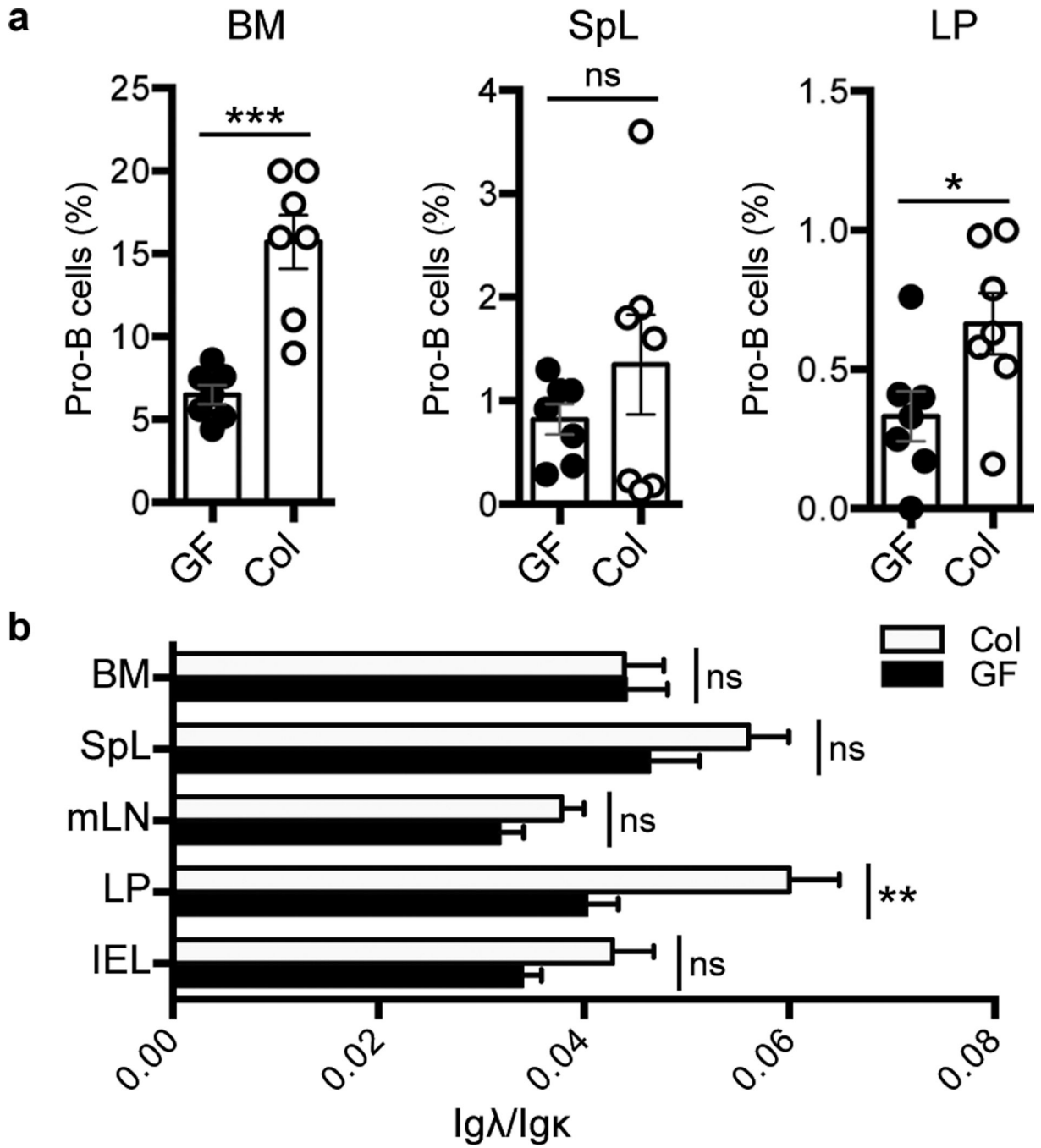


**Figure 2. RAG2-GFP<sup>+</sup> LP B lineage developmental subsets**

**a**, Plots show the relative percentage *Rag2*-expressing pro-B cells, pre-B cells, editing B cells and surface IgM<sup>+</sup> B cells (see text for definition of each) in the bone marrow (blue bars) and lamina propria (red bars). Plotted are mean values  $\pm$  s.e.m. and each are derived from experiments of 4 independent mice at post-natal day 17–24 (see Supplementary Fig. 7 for more details). **b**, Plots show quantitative ligation-mediated PCR of RAG2<sup>+</sup> BM B cells and RAG2<sup>+</sup> LP B cells normalized to genomic DNA. BM cells from RAG2<sup>-/-</sup> mice were a negative control. Values on the y-axis are units relative to the signal obtained from a RAG2<sup>+</sup> BM B cell samples.



**Figure 3. Distinct  $V_K$  segment usage in RAG2<sup>+</sup> cells from BM versus LP**  
**a,b**, Dot plots show contributions (in order of highest to lowest BM utilization) of different  $V_{HS}$  (a) and  $V_{KS}$  (b) to in-frame rearrangements in RAG2-GFP<sup>+</sup> BM (black dots) and LP (red dots) cells. The two most highly utilized  $V_{HS}$  are omitted to increase plot resolution. Each point shows mean  $\pm$  s.e.m. of at least 4 experiments. The  $\chi^2$  calculated  $P$  values for overall differences between BM and LP are indicated. Significant  $V$  segment utilization differences between BM and LP are indicated on heat map ( $P$  values scale indicated in inset). Full data set at increased resolution is in Supplementary Fig. 12.



**Figure 4. Effects of Gut colonization on Development of LP B Lineage Cells**

**a**, Plots of percentage of pro-B cells versus total CD19<sup>+</sup> B lineage cells from bone marrow (BM), spleen (SpL) and lamina propria (LP) of 4 wk-old germfree (GF) mice and littermates colonized (Col) by co-housing with serum pathogen free mice for 7 days. **b**, Bar graphs show ratios of Igλ<sup>+</sup> versus Igκ<sup>+</sup> B cells within mesenteric lymph nodes (mLN), inter-epithelial lymphocytes (IEL) and tissues indicated in (a) from GF and Col mice. Mean values

and s.e.m are shown. The  $P$  values are indicated as: \* $P$  0.05, \*\* $P$  0.01, \*\*\* $P$  0.001. ns = not significant. (Details in Supplementary Fig. 16 and 17)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript