THE UNIVERSITY OF CHICAGO

CXCR4 REGULATES B CELL RECEPTOR EDITING

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

COMMITTEE ON IMMUNOLOGY

BY

MICHAEL KOFI OKOREEH

CHICAGO, ILLINOIS

DECEMBER 2020

To my parents and countless supporters

Table of Contents

List of Figures

Acknowledgements

I would like to thank my PhD advisor Dr. Marcus Clark for giving me an opportunity to conduct research in his lab. I have had the pleasure of working with Marcus for many years. As undergraduate at the University of California, Riverside, Marcus allowed me to spend a summer in his lab. At the time I had little knowledge of the immune system let alone B cell development. This experience was a turning point for me, as I became very interested in immune tolerance and autoimmunity. Marcus's support did not stop that summer but continued as we applied for NIH funding to allow me to return to the lab and work for a year as a technician. During, that year as I was also applying to medical school and when I was admitted to Pritzker School of Medicine Marcus spoke with Dr. Nancy Schwartz on my behalf to facilitate entry into the MD/PhD program. In graduate school, Marcus showed great support of my interests and ideas as we started this project from ground zero; needing to cryo-recover the κ-macroself mice from Jackson laboratory and receive the Igκ^{del} mice from his colleague Dr. David Nemazee. Through this project, Marcus remained supportive and open-minded as many hypotheses were proven incorrect or reanalysis of data changed our thinking. In all, I am very grateful for Marcus's support not only as a thesis advisor but opening doors and giving me opportunities that have undoubtedly changed the course of my life.

As stated, I have been a member of the Clark lab for years and I have interacted with wonderful people. Keith Hamel, whom I worked with my first summer, was incredibly patient in teaching me B cell biology and walking me through experimental designs and protocols. My year spent as his technician more than prepared me for being successful in graduate school. Malay Mandal, an assistant professor in the lab, was also an immense

supporter of my ideas during graduate school. Malay allowed me to analyze most of the sequencing data for his 2019 Nature Immunology paper on CXCR4 and that was another crucial experience as I have developed a passion for bioinformatics that I will continue to pursue. Domenick Kennedy, a postdoctoral fellow in the lab, was also a great supporter and helping hand during many of my experiments. Especially as the COVID-19 has reduced the number of people that can be in a lab at a time, Dom has been very helpful in helping me push experiments along. Dom also allowed me to be a major contributor to his Nature Immunology paper on germinal center subpopulations. Being a member of the Clark lab has been the best research experience and I am grateful for everyone for fostering such an welcoming environment.

I would like to express my gratitude to my committee members; Dr. Barbara Kee, Dr. Marcelo Nobrega, Dr. Fotini Gounari, and Dr. Mark Maienschein-Cline. This project was very niche, but their insightful questions and critiques were invaluable to the completion of my research. I would like to thank Dr. Mark Maienschein-Cline for his patience with teaching me to analyze the large datasets generated by this research. Mark was very insightful and great at explaining very technical statistical analysis. He pushed me to get better in my bioinformatic skills and that hard work is a major reason why I will continue to pursue bioinformatic analysis in the future.

This journey has been incredibly challenging, but it was made easier because of all the great friends I have made in Chicago. Within the Committee on Immunology, Ceprika (Elaine) Kouame has been a great friend. I admire her advocacy for improving the experience of black students on campus. Former members of the committee on immunology Jasmin Quandt and Leila Haghi were also very kind to me and taught me so much about their respective cultures.

I am also immensely appreciative of Akinola (Junior) Emmanuel. As a graduate student in Fotini's lab, Junior taught me how to use terminal command line and analyze genome datasets. Conversations with him about my project gave me great ideas about how to change my focus when I felt I was running out of ideas and confused about what the data was saying. Although not direct members of the Immunology Committee, I would like to thank Greg Washington for his friendship and engaging discussions, and Gloria Davis for her cheerful and welcoming demeanor. I am also grateful to Paul Cato, a graduate student in the Committee on Social Thought. Paul is like a brother to me and a great supporter during all my years in graduate school. I very much appreciate his friendship.

I am grateful to Euclid Williamson, Jacinta Wallace and Efundunke Hughes from Target H.O.P.E, a non-profit that focuses on providing educational support and help to inner city Black and Latinx high school students applying to college. I worked with Target H.O.P.E for 4 years every Saturday during the school year. I am grateful for this opportunity to give back to my community and inspire the next generation of scientists. Euclid, Jacinta and Efundunke are talented educators and it was an honor to work with them. I would like to thank all the graduate students, physicians, and scientists that gave up their Saturdays to share their experiences with the students.

Lastly, I would like to thank my family for their prayers and unwavering support during these tough years. I always felt rejuvenated when I returned to Chicago after the holidays. I also want to thank my partner Jasmine Arrington for loving me and being the best part of my day.

Abbreviations

ALP	All Lymphoid Progenitor
BAD	BCL2 associated agonist of cell death
BAFF	B-cell activating factor belonging to the TNF family
BCAP	B cell Adaptor for PI3K
BCL2	B-cell lymphoma 2
BIM	Bcl-2-interacting mediator of cell death
BLNK	B cell linker
BLP	B cell Lymphoid Progenitor
BM	Bone Marrow
BP	Base Pair
BRWD1	Bromodomain and WD repeat-containing protein 1
CLPs	Common Lymphoid Progenitors
CXCR4	C-X-C motif chemokine receptor 4
CXCL12	CXC-chemokine ligand 12
DEGs	Differentially expressed genes
dsDNA	Double Stranded DNA
DSBs	Double Stranded Breaks
Eκ3’	3’ Igκ enhancer
Eκi	Intron Igκ Enhancer
ERK	Extracellular- signal-regulated kinase
HEL	Hen Egg Lysozyme
HR	Homologous recombination

HSCs	Hematopoietic Stem Cells
Id	Inhibition of DNA binding
Ig	Immunoglobulin
IL-7R	Interleukin 7 Receptor
IL-7	Interleukin 7
IRS	Intron Recombining Sequence
KDE	Kappa Deleting Element
LC	Light Chain
LMPPs	Lymphoid primed multipotent progenitors
LPS	Lipopolysaccharide
mAbs	Monoclonal antibodies
MCL1	Myeloid Cell Leukemia Sequence 1
MPPs	Multipotent Progenitors
NF-κB	Nuclear Factor κB
NHEJ	Nonhomologous DNA End-Joining
OCRs	Open chromatin regions
OxPhos	Oxidative Phosphorylation
PAX5	Paired-box containing transcription factor
PALS	Periarteriolar lymphoid sheath
PTEN	Phosphatase and Tensin Homologue
PI3K	Phosphoinositide 3-kinases
RAG	Recombination Activating Genes
ROS	Reactive oxygen species

RS	Recombining Segment
RSS	Recombination Signal Sequence
SLC	Surrogate light chain
SLE	Systemic Lupus Erythematosus
STAT5	Signal transducer and activator of transcription 5
T1	Transitional 1 B cells
T1D	Type 1 Diabetes
T2	Transitional 2 B cells
T3	Transitional 3 B cells
TNF	Tumor Necrosis Factor
VCAM1	Vascular cell adhesion protein 1
WT	Wild Type

Abstract

During B lymphopoiesis, B cell progenitors progress through alternating and mutually exclusive stages of clonal expansion and immunoglobulin (Ig) gene rearrangements. Great diversity is generated through the stochastic recombination of Ig gene segments encoding heavy and light chain variable domains. However, this commonly generates autoreactivity. Previous studies have shown that receptor editing is the predominant central tolerance mechanism for self-reactive B cells in the bone marrow. Receptor editing rescues autoreactive B cells from negative selection, by inducing renewed light chain recombination first at *Igκ* then *Igλ* loci. Reports have indicated the λ-chain has distinct physiochemical properties that are particularly effective at quenching autoreactivity to DNA, yet, molecular mechanism of receptor editing that lead to the usage of the λ-chain remain ill defined. Herein, we elucidate transcriptional and epigenetic features of receptor editing cells and show chemokine receptor CXCR4 is essential for maintaining the editing state and orchestrating *Igλ* light recombination. Editing small pre-B cells possess unique open chromatin regions enriched for E2A and NF-κB, key regulators of *Igλ* recombination and receptor editing. These regions showed evidence of having been acquired by partial differentiation into the immature B cell stage. Transcriptionally, editing cells showed upregulation of oxidative phosphorylation, DNA repair mechanisms and CXCR4. CXCR4 deficiency in autoreactive B cells showed a selective defect in *Igλ* recombination and tolerance induction, suggesting CXCR4 is necessary for the differentiation of Igλ+ B cells during receptor editing.

1. Introduction

Together with T cells, B cells participate in a highly ordered, tightly controlled system to respond to immediate and future threats to the host. B cell recognize antigen through their B cell receptor that ultimately leads to antibody secretion to mark infected cells and pathogens for death or internalization of pathogens for presentation of antigens to alert other immune cells. Moreover, during the course of an immune response, B cells possess the unique ability to increase their affinity for antigens, helping to both resolve the current immune response but also to quickly initiate an immune response during re-infection. These qualities make B cells essential, yet potentially harmful due to the significant risk of recognizing self-antigens as foreign. The importance of B cells is demonstrated due to the evolutionary conversed role in lower vertebrates and numerous pathologies that arise through B cell dysfunction. Thus, a precise understanding of B cell development and biology is important for the discovery of therapies and eventually elimination of B cell mediated diseases.

1.1 Specification and Commitment to the B Cell Lineage

Much of our understanding of the immune system comes from studying murine hematopoiesis. Decades of research has elucidated the intermediate cell stages, transcription factors, and environmental cues for many of the different cell lineages. Immune cells develop from hematopoietic stem cells (HSCs), residing in the bone marrow (BM) or fetal liver. These are a self-renewing population of cells with the potential to give rise to all cell types within the blood. This multipotency is achieved through low-level expression of genes affiliated with different cell lineages such as B cell, T cells, macrophages, etc (Mandel and Grosschedl,

2010; Zandi et al., 2010). Integration of environmental ques relayed through surface receptors and upregulation of lineage defining transcription factors initiate the events that lead to HSCs specification; the activation of lineage-specific gene programs and commitment; the repression of alternative lineage-specific gene programs (Kee, 2009). At each developmental stage, precursor cells increasingly lose their ability to differentiate into other cell types.

B cell development begins with asymmetric division of HSCs, into multipotent progenitors (MPPs). Development bifurcates at this point into the myeloid and lymphoid lineages (Boller and Grosschedl, 2014). Progenitors differentiating down the lymphoid lineage upregulate the tyrosine kinase receptor Flt3 and give rise to lymphoid primed multipotent progenitors (LMPPs). Next, upregulation of Interleukin receptor 7 marks the next stage of development represented by common lymphoid progenitors (CLPs) (Rothenberg, 2014). CLPs are not restricted to the B cell lineage but can differentiate into either B, T, or Natural Killer cells (Kondo et al., 1997). CLP multi-lineage potential was demonstrated using single-cell culture systems where a single CLP was plated and was shown to give rise to various other cell types. Recently, use of transmembrane protein Ly6D has further divided CLPs into cells retaining potential for generating all lymphoid progenitors (ALP;$Ly6D^-$) and cells biased toward the B cell lymphoid progenitors (BLP; $Ly6D^+$) (Boller and Grosschedl, 2014; Inlay et al., 2009).

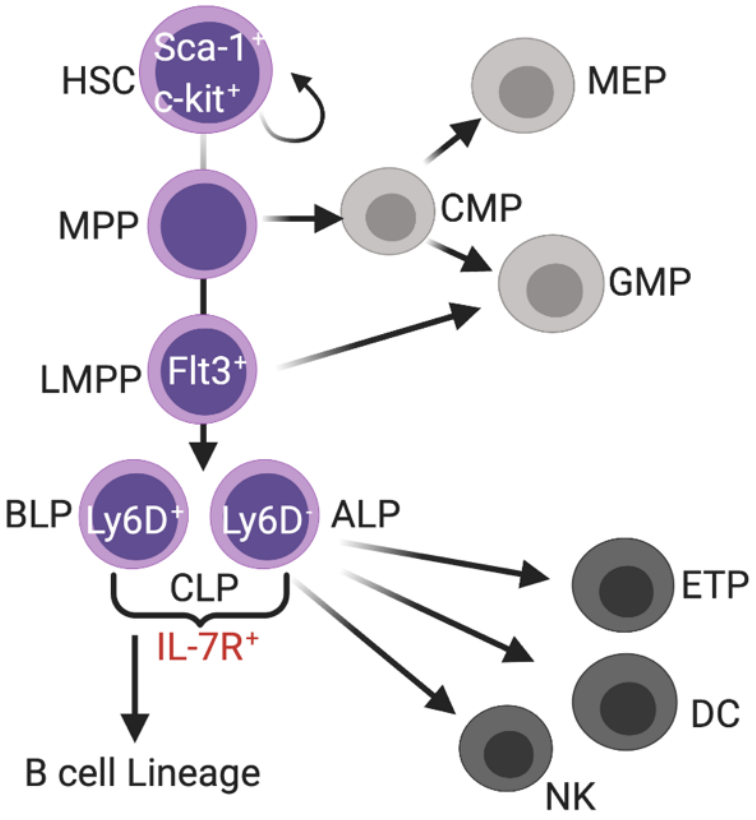


Figure 1.1: Schematic view of B cell lineage commitment
B cell lineage commitment is a stepwise process beginning with Sca-1+ c-kit+ HSCs. Through asymmetric divisions hematopoietic stem cells (HSCs) will yield progenitor cells that lose potential to differentiate into other lineages but gradually gain potential to become B cells. Ly6D has recently become a marker for identifying cells within the common lymphoid progenitor (CLP) population that have a biased potential for B cell formation. Common B lymphoid developmental stages are depicted in purple. Cell types leading to non-B lineages are depicted in gray. Important cell surface markers are outlined in white and red font. MPP, multipotent progenitor; LMPP, lymphoid-primed multipotent progenitor; ALP, all lymphoid progenitor; BLP, B-cell- biased lymphoid progenitor; CMP, common myeloid progenitor; MEP, megakaryocytic/erythrocyte progenitor; GMP, ganulocyte/macrophage progenitor; ETP, early thymic progenitor; DC, dendritic cell; NK, natural killer cell. Adapted from (Boller and Grosschedl, 2014).

CLPs are identified by expression of cytokine receptor Interleukin 7 (IL-7R), which binds Interleukin 7 (IL-7). The binding of IL-7 and IL-7R produces downstream signals that begins upregulation of E protein E2A (Bajoghli et al., 2009). E proteins are transcription factors that contain both a basic DNA-binding domain and helix-loop-helix protein

dimerization domain (Kee, 2009). Mammals express 4 different E proteins: E12 and E47 (alternative splicing products from Tcf3 that bind together forming E2A), HEB and E2-2. E proteins activity is regulated post-translationally by Inhibition of DNA binding proteins (ID1-ID4). ID proteins bind E proteins and prevent their binding to DNA. Genetic deletion of E2A leads to a development arrest at the pre-pro B cell stage that follows the CLP stage (Massari and Murre, 2000; Murre, 2005). E2A deficiency in human seems to halt B cell development at a similar stage as well (Jaleco et al., 1999). Thus, E2A is crucial for the initiation of the B cell lineage program.

In CLPs, a major target of E2A is Early B cell factor 1 (EBF1). E2A together with Forkhead Box O1 (FOXO1) cooperate to induce EBF1 (Rothenberg, 2014). Expression of EBF1 heightens E2A expression due to EBF1 being a strong inhibitor of ID protein expression (Thal et al., 2009). Together EBF1 and E2A upregulate many early B cell lineage genes including Paired-box containing transcription factor (PAX5), which establishes a regulatory network that drives commitment to the B-lineage as pro-B cells (Allman et al., 1999; Hagman et al., 1990). EBF1 and PAX5 co-expression establishes a positive feed-forward loop that amplifies and maintains B cell identity in precursor cells. IL-7R signaling in concert with the aforementioned TFs causes commitment of pre-pro B cells into the B cell lineage. Committed B cells begin to express protein marker CD19 at high levels on the cell surface (Rothenberg, 2014). What follows are B cell progenitor cells that alternate through sequential yet mutually exclusive stages of proliferation and differentiation resulting in an immature B expressing a functional B cell receptor (BCR).

1.2 BM B cell subpopulations

B lymphocytes were first discovered in the mid 60s and 70s by studying animal models and clinical observations of patients with immunodeficiency. In fact, B and T cells were discovered simultaneously by Max Cooper and Robert Good (LeBien and Tedder, 2008). Cooper and Good were studying the bursa (bone marrow) of chickens and observed antibody secreting cells derived from the bursa (hence the name B cell) and cells that need a thymus for cell-mediated hypersensitivity (hence the name T cell) (Cooper et al., 1965; Cooper et al., 1966; Lawton et al., 1972; Raff et al., 1976). Murine models were later used to demonstrate B cell were the source of antibodies and arose from precursor cells that did not express surface immunoglobulin (Ig).

In the 1990s, Randy Hardy and colleagues using newly developed flow cytometry technology and monoclonal antibodies (mAbs) against cell surface proteins (markers) provided resolution of B cell precursor subpopulations (Herzenberg and Herzenberg, 2004; Herzenberg et al., 1976; Van Epps, 2006). Flow cytometry using multicolor lasers to excite fluorescently labeled mAbs bound to cells allowed for sorting distinct cell populations. These sorted cells were cultured in vitro to determine subsequent developmental stages. Hardy divided early B cell development into fractions: A, B, C, C',D, and E (Hardy et al., 1991). Today, these fractions are termed pre-pro B, progenitor B (pro-B; combination of B and C fractions), large pre-B, small pre-B, and immature B cells, respectively.

1.2.1 Pro-B cells

Pro-B cells are the first committed cells to the B cell lineage and express CD19 in a PAX5 dependent manner (Boller and Grosschedl, 2014). Pro-B cells also express high levels

of IL-7R and thus are highly responsive to the cytokine IL-7. IL-7R signaling provides survival signals necessary to maintain pro-B cells (Clark et al., 2014). The primary signal transducer of IL-7R signaling is Signal transducer and activator of transcription 5 (STAT5), which in turn activates pro-survival factors B cell lymphoma 2 (BCL2) and myeloid cell leukemia sequence 1 (MCL1) in both mice and humans (Essafi et al., 2005; Jiang et al., 2004; Malin et al., 2010). Conversely, STAT5 promotes proliferation by upregulating Cyclin D3. In addition to STAT5, IL-7R signaling induces Phosphoinositide 3-kinases (PI3K). Absence of PI3K has a defect in B cell proliferation and B cell specific deletion of regulatory or catalytic PI3K subunits results in impaired B cell development (Fruman et al., 1999; Okkenhaug et al., 2002; Suzuki et al., 1999). PI3K-Akt signaling promotes survival by phosphorylating FOXO1 leading to nuclear export. FOXO1 promotes apoptosis by upregulation pro-apoptotic molecule BCL2 interacting mediator of cell death (BIM) (Powers et al., 2012). Akt directly inactivates BIM through phosphorylation proapoptotic factor BCL2 antagonist of cell death (BAD) (Lam et al., 2000).

It has become clear that the pro-B stage is not homogenous. The pro-B stage seems to possess a continuum of cells that reside in niches with high IL-7 versus low IL-7 (Hamel et al., 2014). Early pro-B cells undergo a few rounds of proliferation under IL-7R signaling and reside in IL-7 rich niches. After a few rounds of division, pro-B cells will cease proliferation and enter a late phase hallmarked by the completion of with V_H to D_HJ_H heavy chain recombination (Clark et al., 2014). Unlike in with D_HJ_H recombination where gene rearrangements are biallelic, the V_H to D_HJ_H events are monoallelic (Alt et al., 1984; Jung et al.).

1.2.2 Large Pre-B cells

Successful completion of heavy chain recombination is a crucial checkpoint for early B cell precursors. The assembled heavy chain gene will pair with surrogate light chain (SLC) molecules VpreB and λ5 along with signaling components Igα/Igβ to form the pre-BCR at the cell surface. Cells expressing a pre-BCR transition to the Pre-B stage and undergo a clonal proliferative burst in the presence of IL-7 (Clark et al., 2014). Indeed, mice deficiency in SLC genes VpreB or λ5 have impaired B development at the late pro-B stage after heavy chain recombination (Gong and Nussenzweig, 1996; Pelanda et al., 2002; Shimizu et al., 2002). Not every heavy chain can successfully pair with the SLC to form a pre-BCR. Thus, the clonal expansion of large pre-B cells is the first checkpoint that shapes the repertoire of heavy chains that comprise the mature B cell pool (Melchers et al., 2000).

The manner in which pre-BCR signaling is initiated is a controversial topic. Pre-BCR activation requires the non-immunoglobulin domain of λ5 for aggregation (Bankovich et al., 2007; Ohnishi and Melchers, 2003; Vettermann et al., 2008). However, it remains unknown if receptor aggregation is an intrinsic property or a consequence of receptor binding to ligand expressed in the BM. Interestingly, heparin sulfate and galectin-1 have been shown to be naturally occurring ligands that can bind the pre-BCR (Bradl et al., 2003; Gauthier et al., 2002; Mourcin et al., 2011; Winkler and Mårtensson, 2018).

Large pre-B cells in addition to pre-BCR expression continue to express IL-7R. Given that the surface expression of the pre-BCR coincides with the large proliferative burst of large pre-B cells, questions arose as to whether pre-BCR signals enhanced proliferative signals of IL-7. PI3K-Akt activity, downstream of IL-7R signaling, is believed to be important in large pre-B cell through upregulation of glucose transporters and glycolysis necessary to

sustain proliferation (McLean and Mandal, 2020). However, it was found that pre-BCR couples poorly with PI3K. Transfection of Rag2 deficient pro-B cells in the presence of IL-7 with a prearranged heavy chain resulting in pre-BCR expression, showed little activation of phosho-Akt (Ochiai et al., 2012). When downstream pre-BCR signaling components B cell Linker (BLNK) or phospholipase Cγ2 were deleted, B cell development was arrested at the proliferating pre-B cell stage. Other experiments over-expressing BLNK demonstrated pre-BCR signaling to be associated with cell cycle exit and not proliferation (Bai et al., 2007; Flemming et al., 2003; Jumaa et al., 1999; Xu et al., 2007).

The definitive experiment demonstrating pre-BCR signaling does not induce proliferation was performed using mice that overexpressed SLC genes VpreB or λ5 (van Loo et al., 2007). Pre-B cells in these mice shown no enhanced proliferation or differentiation. However, immature B cells numbers were greatly reduced and showed evidence of multiple rounds of antibody gene rearrangements. Thus, constitutive pre-BCR signaling is insufficient to induce proliferation. After clonal expansion, large pre-B cells decrease in size and exit cell cycle to become small pre-B cells.

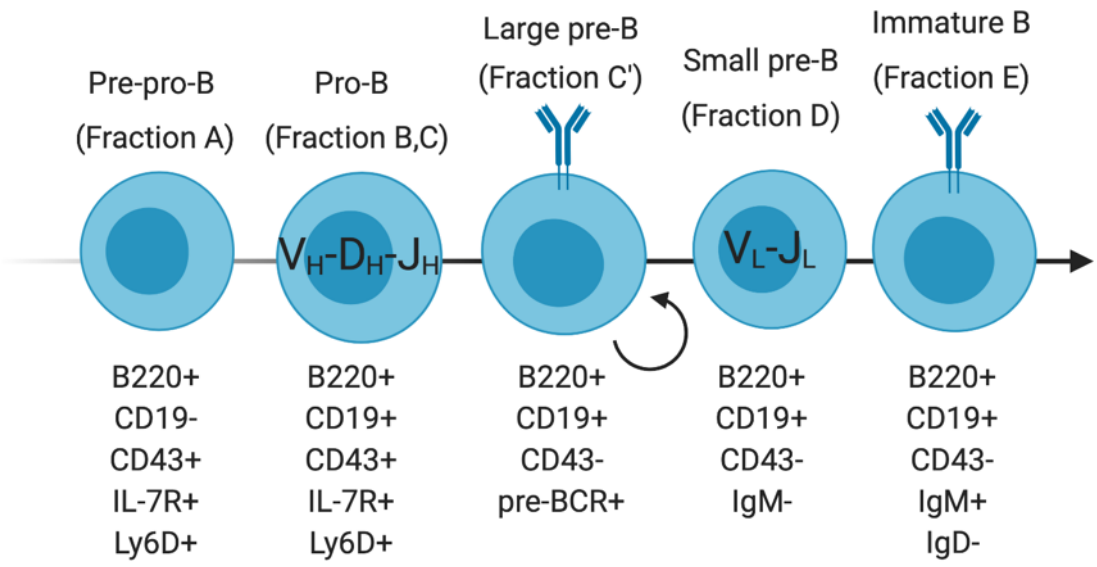


Figure 1.2 Schematic of early B cell development.
B lymphopoiesis occurs in the bone marrow (BM) where progenitor cells progress through alternating and mutually exclusive stages of clonal expansion and quasi-random immunoglobulin (Ig) gene rearrangements. Ig heavy chain Variable (V), Diversity (D) and Joining (J) gene rearrangements occur in pro-B cells and require the activity of recombination activating genes Rag1 and Rag2 together with numerous DNA repair proteins. Pro-B cells with successful Ig heavy chain rearrangements express the encoded protein together with invariant surrogate light chain (SLC) genes VpreB/lambda 5 (λ5) on the cell surface as a pre-B cell receptor (pre-BCR). These, now, large pre-B cells undergo vigorous proliferation highly dependent on the cytokine Interleukin 7 (IL-7) expressed by BM stromal cells binding to IL-7 receptor (IL-7R). After several proliferative rounds, large pre-B cells will cease proliferation, downregulate the pre-BCR to differentiate into small pre-B cells and initiate light chain VJ (kappa, κ, or lambda, λ) recombination. Light chain recombination at *Igκ* or *Igλ* loci occur independently and sequentially with gene rearrangements beginning at *Igκ*. Successful Ig light chain rearrangements in combination with the previously rearranged Ig heavy chain pair together and are expressed as a functional B cell receptor (BCR) of the IgM isotype on immature B cells.

1.2.3 Small Pre-B cells

Light chain recombination occurs at the small pre-B cell stage. This marks another significant stage in B lymphocyte development. Accomplishing light chain recombination involves the integration of IL-7R, pre-BCR and C-X-C motif chemokine receptor 4 (CXCR4) receptor signals to mediate the exit from cell cycle and the derepression of the light locus and upregulation of DNA recombination machinery (Mandal et al., 2019; McLean and

Mandal, 2020). Mammals have two light chains Igκ and Igλ. In mice, Igκ recombination occurs first and represents over 90% of the B cell repertoire. For this reason, much of what is known about light chain regulation concerns Igκ. Yet, the Igλ chain has been shown to be important for rescuing cells that fail Igκ recombination or κ-chain bearing autoreactive BCRs (Nemazee and Weigert, 2000).

Igκ recombination cannot occur in the presence of IL-7. IL-7 signaling represses Igκ accessibility and recombination by STAT5 occupancy at intron enhancer (Eκi) and the Igκ 3′ enhancer (Eκ3) within the Igκ locus (Mandal et al., 2011). Pre-BCR expression both directly inhibits proliferative signals downstream of IL-7R but also upregulates expression of chemokine receptor CXCR4. CXCR4 in response to CXC-chemokine ligand 12 (CXCL12) gradient causes large pre-B cells to migrate away from IL-7 rich niches. Recent evidence has demonstrated that pre-BCR signaling facilitates the transition from IL-7R mediated proliferation to CXCR4 mediated recombination to orchestrate the developmental events at the small pre-B cell stage (Mandal et al., 2019; McLean and Mandal, 2020).

In the absence of IL-7R signaling, pre-BCR signaling dominates and induces upregulation of extracellular- signal-regulated kinase (ERK) and Ras pathway, crucial for *Igκ* recombination (Mandal et al., 2009; McLean and Mandal, 2020). The Ras-ERK pathway induces Rag1/2 and leads to an accumulation of free E2A by inhibiting ID3. E2A then binds to Eκi and Eκ3', which is no longer occupied by STAT5. E2A bind two of the three E boxes within Eκi; κE1 and κE2 (Murre et al., 1989). Genetic targeting of the E-boxes within Eκi impaired *Igκ* recombination, with the greatest reduction in rearrangements occurring when κE1 and κE2 were mutated in combination (Inlay et al., 2004; Sakamoto et al., 2012).

E2A recruitment is crucial for all aspects of *Igκ* gene rearrangements: accessibility, germline transcription and recombination. E2A activation of iEκ recruits histone modifiers to decorate acetyl groups to H3 histones flanking *Cκ* and *Jκ* genes. In particular, E2A acetylation of *Jκ* genes create an epigenetic landscape to recruit epigenetic reader BRWD1 (Bromodomain and WD repeat-containing protein 1) (Mandal et al., 2015; Schatz and Ji, 2011). BRWD1 repositions nucleosomes making *Jκ* accessible to the RAG complex to initiate recombination (Mandal et al., 2018).

For decades, regulation of *Igκ* was thought to be primarily controlled by pre-BCR signaling which dominates in the absence of IL-7R signaling as large pre-B cells migrate away from IL-7^{hi} niches. However, the model for how pre-BCR controlled *Igκ* recombination was confounded by the fact that pre-BCR signaling was associated with concurrent pre-BCR repression. It became unclear whether transient pre-BCR signaling was sufficient to complete the small pre-B cell developmental program or if other receptors where involved. CXCR4 was a likely candidate for involvement in small pre-B cell development as its expression is necessary for limiting IL-7R signaling. Moreover, pre-BCR expression induces IRF4, which positively regulates CXCR4 expression (Mandal et al., 2019).

In vitro cultures of developing pre-B cells in the presence or absence of IL-7 and CXCL12 helped to elucidate the unique contributions of pre-BCR and CXCR4 signaling at the small pre-B stage. Previous experiments used the stromal feeder cell lines OP9 and ST2, which expressed CXCL12, thus obscuring the exact role of CXCR4 signaling (Lagergren et al., 2007). Pre-B cells cultured in stromal cell free conditions, with CXCL12 demonstrated more efficient *Igκ* accessibility and recombination. CXCR4 signaling was demonstrated to repress cell cycle, control chromatin remodeling and promote survival. In fact, many of the

downstream molecules previously ascribed to pre-BCR signaling were induced by CXCR4 signaling. Thus, light chain recombination is accomplished through a coordinated interplay between the IL-7R, pre-BCR and CXCR4. Small pre-B cells that successfully complete light chain recombination, pair the light chain with the preassembled heavy chain on the cell surface as the BCR.

1.2.4 Immature B cells

BCR expression permits entry into the immature B cell stage. In order to progress through the immature B cell stage, the new BCR must (1) be capable of transmitting signals that repress RAG proteins, which prevent further light chain rearrangement and promote allelic exclusion (Tze et al., 2000; Tze et al., 2005), (2) allow egress from the BM to peripheral sites for further maturation (Chung et al., 2003; King and Monroe, 2000), (3) be unable to recognize host derived antigens present in the BM (Norvell et al., 1995; Sandel and Monroe, 1999). Thus, progression through this stage requires the BCR to signal in different ways.

In the unligated form, the BCR delivers intracellular signals termed 'tonic signals' that are important for immature B cell survival and maturation (Monroe, 2006). Tonic signaling largely consists of PI3K activity through tyrosine phosphorylation of adaptor proteins CD19 and B cell adaptor for PI3K (BCAP) by kinases (Fujimoto et al., 2000). PI3K activity negatively regulates RAG via activation of AKT, which phosphorylates FOXO1 and prevents translocation into the nucleus. This pathway ensures cessation of light chain recombination and promotes B cell survival (Tze et al., 2005; Verkoczy et al., 2007). CD19 deficiency can cause defects in immature B cell positive selection. However, this can be overcome by inhibition of PI3K activity inhibitor phosphatase and tensin homologue (PTEN). The RAS-ERK pathway has also been implicated in positive selection. Expression of active RAS

promotes B cell differentiation and can compensate for positive selection defects due to BCR under expression (Rowland et al., 2010). The RAS-ERK pathway can also indirectly promote PI3K activity and downregulate RAG (Llorian et al., 2007). Thus, many components downstream of BCR signaling facilitate positive selection.

Immature B cell stage is also the site for tolerance induction. Tolerance checkpoints during B cell development occur at two sites: BM and secondary lymphoid organs namely, the spleen. Central tolerance describes mechanisms that occur in the bone marrow. These consist of clonal deletion, receptor editing and anergy. Tolerance mechanisms in secondary lymphoid organs include clonal deletion and anergy are collectively called peripheral tolerance. B cells bearing potentially autoreactive BCRs are purged from the immune repertoire. Studies have suggested 50-75% of developing B cells have specificity to nuclear antigens (Giltiay et al., 2012; Meffre and Wardemann, 2008; Wardemann et al., 2003). Some studies suggest up to 50% of the emerging B cell repertoire are targeted for tolerance mechanisms (Casellas et al., 2001; Hippen et al., 2005; Retter and Nemazee, 1998). Thus, self-tolerance mechanisms at the immature B cell stage have significant impact on the immune repertoire.

A model for how BCR signaling can mediate both positive selection and tolerance induction posits that BCR signaling of intermediate strength (such as with tonic signaling) is sufficient to promote positive selection (Nemazee, 2017). However, newly formed autoreactive immature B cells that encounter self-antigen in the BM, will signal strongly before rapid internalization and removal of the BCR from the cell surface. This reduced state of BCR signaling effectively returns immature B cells to the small pre-B cell stage where recombination can resume to replace the existing light chain (LC), editing the original BCR

specificity. B cells can then be positively selected to develop to maturity if the replacement generates a non-self-reactive receptor. These tolerance mechanisms will be discussed in greater detail in later sections.

1.3 Bone Marrow Microenvironment

The bone marrow microenvironment influences the survival, proliferation and responsiveness of developing B cells. Bone marrow stromal cell expression of certain chemokines and cytokines create niches that promote stem cell commitment to the B cell lineage and are necessary for differentiation into antibody secreting cells. Stromal cells must produce the necessary factors and deliver them to B cell precursors at the appropriate cell stages. Fluorescence based immunohistochemistry of BM sections has made it possible to determine the structure, distribution and cell types that comprise BM niches. These studies identified reticular cells, which surround BM sinuses and are the most abundant stromal cell type in the BM, as a probably candidate for establishing these niches (Nagasawa, 2006). While it is clear that distinct stromal cell-lines produce different environmental factors in vitro, it is unclear how these stromal cells function in vivo (Nagasawa, 2006). In vitro studies have implicated IL-7, CXCL12, RANKL and SCF as crucial ligands for developing B cells. However, the importance of IL-7 and CXCL12 niches to B cell development have been best described.

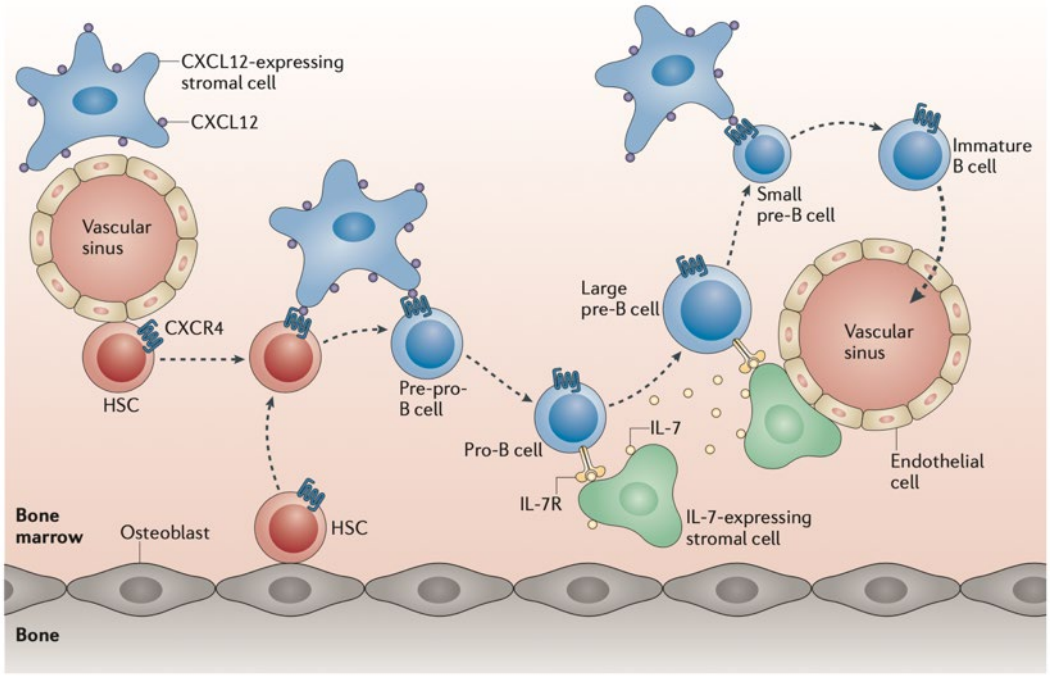


Figure 1.3 Bone marrow microenvironments during early B cell development
The bone marrow (BM) microenvironment consists of stromal and reticular cells that secrete soluble factors necessary for B cell lineage commitment and differentiation. Of particular importance are CXC-chemokine ligand 12 (CXCL12) expressing stromal cells and Interlukin 7 (IL-7) expressing stromal cells. Hematopoietic stem cells (HSCs) express CXCL12 receptor CXC-chemokine receptor 4 (CXCR4) on the cell surface and are drawn into areas enriched in CXCL12. The CXCR4-CXCL12 signaling axis aids in differentiation of HSCs into pre-pro-B cells. Pre-pro-B cells differentiate into pro-B cells that express IL-7 receptor (IL-7) and subsequently reside in niches enriched for IL-7. After successfully rearrangement of the heavy chain pre-B cells express the pre-BCR and clonally divide in the IL-7 expressing niches as large pre-B cells. After proliferation, CXCR4 is upregulated and induces migration away from IL-7 enriched niches toward CXCL12 enriched niches. Attenuation of IL-7R signaling facilitates light chain recombination. After successful light chain recombination, immature B cells express IgM, downregulate CXCR4 and exit the bone marrow. Figure adapted from (Clark et al., 2014).

1.3.1 IL-7 Niche

IL-7 expressing cells in the BM are fibroblast-like cells that express adhesion molecule Vascular cell adhesion protein 1 (VCAM1) (Hirose et al., 2002). Using immunohistochemistry analysis and confocal microscopy studies have demonstrated IL-7 expressing cells scattered throughout the BM (Funk et al., 1995; Mandal et al., 2019). However, the expression IL-7 in the BM is not uniform. Single-cell RNA-sequencing analysis of the BM microenvironment

indicates a subset of cells that express varying amounts of IL-7 (Tikhonova et al., 2019). In agreement, flow cytometry analysis of BM stromal cells has shown stromal cells that express varying amounts of IL-7 (Mandal et al., 2019). A possible explanation for this observation in IL-7 expression pattern may be explained by other studies that suggest IL-7 expression is induced by stimuli including B cell precursors (Sudo et al., 1989). Indeed, the earliest B cell precursors have been observed to reside in IL-7 rich niches and have been suggested to only leave these niches to halt proliferation and undergo recombination.

1.3.2 CXCL12 Niche

CXCL12 expressing cells are also scattered throughout the BM (Mandal et al., 2019). Single cell RNA-sequencing identified leptin-receptor-positive mesenchymal stromal cells to be the highest expressors of CXCL12 (Tikhonova et al., 2019). Early B cell precursors require CXCL12 and thus it has been proposed that CXCL12 attracts precursors immediately after B cell lineage commitment. Furthermore, migration toward CXCL12 has been thought to facilitate escape from IL-7 and allow for DNA recombination of heavy and light chain genes. Large pre-B cells were shown to be highly sensitive to CXCL12 gradients, migrating at a higher frequency compared to small pre-B cells (Mandal et al., 2019). Again, confocal microscopy of the BM of mice with YFP expression tied to $C\kappa$ expression showed small pre-B cells closely associated with CXCL12 expressing cells in the BM (Mandal et al., 2019). Interestingly, terminally differentiated plasma cells have also been observed to reside in areas with high CXCL12 expression (Tokoyoda et al., 2004). Whether plasma cells reside in these CXCL12 niches for retention in the bone or CXCL12 is necessary for maintenance of plasma cells is currently unknown.

1.4 Peripheral B cell Maturation

After leaving the BM, immature B cells primarily seed the spleen and continue differentiation to become mature B cells capable of antibody production. Once in the spleen, B cells migrate through red pulp, penetrate the marginal zone sinus before residing in the periarteriolar lymphoid sheath (PALS) (Chung et al., 2003). Once in the PALS, B cell will undergo transitional stages of development before residing in the B cell follicle as Mature B cells.

1.4.1 Transitional B cells

Transitional B cells are a subpopulation of splenic B cells that have recently arrived from the BM and reside in the PALS (King and Monroe, 2000). Transitional cells can be sub-divided into three population: Transitional 1,2, and 3 B cells based on their surface expression of CD23 and IgM. Transitional 1 (T1) B cells are $CD23^{-}IgM^{+}$, while Transitional 2 (T2) cells are $CD23^{+}IgM^{+}$, and Transitional 3 (T3) cells are $CD23^{+}IgM^{lo}$ (Cambier et al., 2007). Transitional B cells must pass through several developmental checkpoints to be selected into the mature B cell repertoire.

Key differences within the transitional B cell subsets center around the differential outcomes of BCR ligation and responsiveness to T cell help. BCR ligation in T1 cell leads to negative selection as a tolerance mechanism against autoreactive specificities (Allman et al., 2001; Chung et al., 2002). However, while some reports demonstrate T2 cells undergo BCR mediated apoptosis, others have demonstrated T cell help via IL-4 production or CD40 engagement rescues T2 cells from BCR mediated apoptosis compared to T1 cell (Loder et al., 1999; Petro et al., 2002; Su and Rawlings, 2002). As responsiveness to T cell help is a crucial

characteristic of mature B cells, the T1 to T2 transition has been described as a gain of immune competence. On the other hand, the ontogeny of T3 cells is much more obscure. T3 cells are anergic B cells that are maintained within the mature repertoire but have a higher threshold for activation. T3 cells are characterized by high IgD expression but very low IgM expression (Goodnow et al., 1988; Sabouri et al., 2016). These cells are believed to have experienced BCR stimulation at either the T1 or T2 stage but have avoided negative selection by inducing the anergic state.

Apart from differential responses to BCR ligation, transitional B cell development is highly dependent on the sensing survival factors tumor necrosis factor (TNF) and B-cell activating factor belonging to the TNF family (BAFF). T1 cells need to acquire BAFF signaling through the BAFF receptor in order to avoid apoptosis. Indeed, experiments blocking BAFF receptor signaling have prevented the T1 to T2 transitional B cell progression (Schiemann et al., 2001). These mechanisms ensure only the fittest B cells possessing BCRs with appropriate downstream signaling components are selected into the B cell follicle.

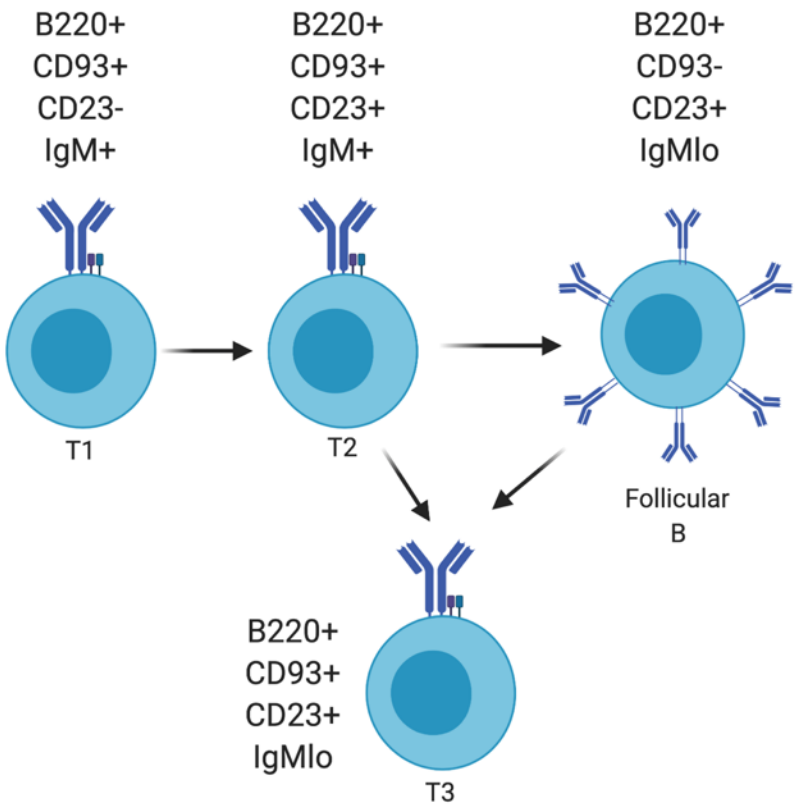


Figure 1.4 Peripheral B cell development.
Immature B cells arrive in the spleen and continue differentiation as transitional B cells. Increased competence of the B cell receptor (BCR) signalosome corresponds to progression of transitional B cells into follicular B cells. As transitional B cells mature, signaling through the BCR results in B cell activation and participation in immune responses as opposed to negative selection or anergy. Studies investigating anergic B cells in mice have identified the T3 population (B220+CD93+CD23+IgMlo) as the anergic population. T1; Transitional 1, T2; Transitional 2, T3; Transitional 3.

1.4.2 Follicular B cells

Follicular B cells are mature B cells capable of participating in an immune response. Binding of antigen by follicular B cells will induce signals that can result in differentiation to many different cell fates. Depending on the context, activated follicular B cells can become short-lived plasma cells that secrete low-affinity antibodies during the early phase of an infection. Other follicular B cells may migrate out of the B cell follicle in order to present antigen to T cells for T cell help (Cyster, 2010). Subsequent interactions with antigen specific T cells will cause formation of a germinal center where somatic hypermutation and affinity

maturation occur to generate long lived plasma cells or memory B cells that have increased affinity for antigen (Boothby et al., 2019).

1.5 Antibody Diversification

During the 1960s-1970s immunologists were fixated on understanding the nature of antibody diversification. Immunologists agreed that vertebrates were capable of secreting millions of diverse antibodies and through the work of Gerald Edelman and Rodney Porter, they understood that an antibody consists of two identical light chains and two identical heavy chains (Edelman, 1973; Porter, 1973). Moreover, they knew these chains consisted of different segments. One segment showed great variability and thus was termed Variable region while the other segment was constant and thus named Constant region. However, immunologists were divided as to whether genetic diversity was acquired during evolution and thus carried in the germline or acquired during development of an organism and thus would be present in only somatic cells and not germline cells. The germline theory proposed that the genome contained a unique gene for every antibody expressed, while the somatic theory maintained that only a limited number of antibody genes were in the germline and that these genes diversified in specialized cells during development. Both theories had significant drawbacks. For the germline theory, it seemed evolutionarily unfavorable for a disproportionate amount of the genome to be used for maintaining genes to account for the vast diversity of antibodies being observed. Moreover, the somatic theory relied on a method of diversification that had not been described in nature. It was not until Susumu Tonegawa in a series of experiments performed between 1974 and 1980, which earned him the Nobel Prize, determined the mechanism of antibody diversification. Using newly developed

restriction enzymes to digest DNA and DNA hybridization techniques, Tonegawa show the immunoglobulin locus was arranged differently in embryo mouse cells versus fully differentiated myeloma cells (Hozumi and Tonegawa, 1976). Probes for the Variable and Constant regions were observed on different fragments of DNA in embryonic mouse cells but were observed on the same DNA fragment in myeloma cells. This suggested a genetic rearrangement had occurred during development between embryonic cells and fully differentiated B cells. Tonegawa concluded that reshuffling of immunoglobulin genes allows for the creation of a wide range of antibodies. In subsequent work, Tonegawa would establish the rules of recombination, proposing proteins necessary to accomplish DNA recombination that would later be discovered by other scientists (Bernard et al., 1978; Brack et al., 1978; Hozumi and Tonegawa, 1976; Tonegawa, 1983).

1.5.1 Mechanism of V(D)J Recombination

Antibody diversity is achieved through permutation of heavy chain V, D and J genes and light chain V and J genes and subsequent pairing. This component of diversity is referred to as combinatorial diversity. V(D)J recombination is a highly ordered process initiated by proteins encoded by recombination activating gene 1 (Rag1) and Rag2. RAG orchestrates recombination between Ig and TCR gene segments by binding highly conserved recombination signal sequences (RSS) that flank these gene segments. RSSs are conserved palindromic heptamer and nonamer sequences separated by 12 or 23 base pair (bp) spacers. RAG binds to RSSs and introduces DNA double-strand breaks (DSBs) between V, D, J segments. RSS ends are joined, forming excision circles, while coding ends are modified with either nucleotide removal or random nucleotide addition. Joining of the coding ends can result in either the inversion or deletion of intervening sequences depending on the

orientation the RSSs. The joining of coding regions is primarily carried out by nonhomologous DNA end-joining (NHEJ). The random addition or removal of nucleotides provides a second layer of diversity in addition to the combinatorial diversity and is referred to as junctional diversity (Bassing et al., 2002).

V(D)J recombination is regulated at 3 levels. First, RAG expression is tightly timed to be upregulated in cells undergoing receptor gene rearrangements. Secondly, gene accessibility facilitated by germline transcription allows for RAG to initiate recombination. Thirdly, three-dimensional structure and chromatin looping of receptor genes are vital in bringing distal gene segments in closer proximity to each other genes, improving the efficiency of V(D)J process (Schatz and Ji, 2011).

1.5.2 Organization of Immunoglobulin Light Chain Loci

Nearly all animals have two light chain isotypes κ or λ (Collins and Watson, 2018). Researchers have speculated that this has been evolutionarily conserved perhaps because each isotype confers unique antigen specificities. However, in some species the B cell repertoire is overwhelmingly represented by either isotype such as κ in mice and λ in horse and chicken (Tallmadge et al., 2014). Moreover, it has been suggested that the predominance of a particular light chain corresponds to the number of available germline V genes in that species (Almagro et al., 1998).

Yet, Igλ+ B cells are important members of a mature B cells repertoire. While WT mice have a B cell repertoire containing a ratio of 20:1 Igκ+ to Igλ+ B cells; the human B cell repertoire has an Igκ+ to Igλ+ ratio of ~2:1, demonstrating an equal importance of the λ-chain to a mature B cell repertoire (Ghia et al., 1995; Gorman et al., 1996; Takeda et al., 1996).

In mice, light chains are on different chromosome κ (Chr2 in human and Chr6 in mice) and λ (Chr22 in humans and Chr16 in mice). Subsequently, many suggestions have been made for why there is a skew towards κ usage in mice. Firstly, the organization of the Igλ locus into VJC clusters is believed to impact recombination. Evidence for this comes by comparative analysis between mice and human Igλ loci. The V genes in Igλ loci in human are clustered together similar to the Igκ locus while there still remains J-C gene cluster downstream. This hypothesis was tested when investigators introduced a knock-in translocus on a yeast artificial chromosome accommodating 60% of λ-chains in humans, including 15 proximal Vλ genes, all Jλ-Cλ segments, and the 3′ enhancer (Popov et al., 1999). They observed similar levels of human λ expression to mouse κ expression. This was suggestive that organization of light chain loci was crucial for representation in the repertoire.

The availability of genomic sequencing allowed for in depth analysis of light chain genes. The human and mouse κ chain loci are organized in similar ways. A cluster of Vκ genes approximately 44 functional genes in human and 101-97 functional genes in mice are 5' to a cluster of 4 functional Jκ genes followed by a single Cκ gene (Karki et al., 2018; Kawasaki et al., 1997; Watson et al., 2015). In contrast, the λ-chain loci are quite different between humans and mice. First, there are more Vλ genes in the human (approximately 38) that are clustered together 5' of the locus similar to the κ locus (Hieter et al., 1981). However, the λ locus does not contain a cluster of Jλ genes but instead has 4 Jλ-Cλ gene cluster 3' of the Vλ clusters.

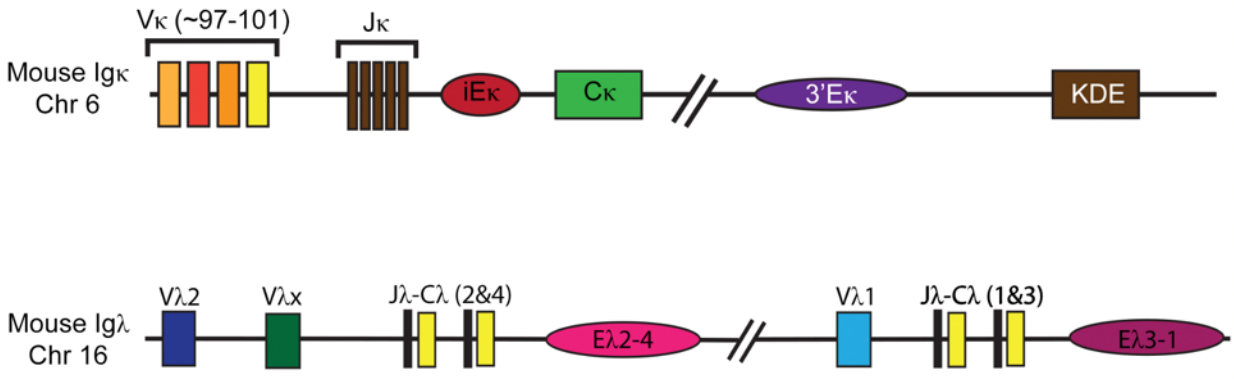


Figure 1.5 Schematic of mouse light chains
In the mouse there are 101-97 *Vκ* genes 5' to a cluster of 4 functional *Jκ* genes followed by a single *Cκ* gene. Two enhancers are present in this locus located 5' of *Cκ* (iEκ) and 3' of *Cκ* (3'Eκ). The *Igλ* locus has 4 *Jλ-Cλ* gene cluster 3' of *Vλ* clusters. The λ locus is greatly reduced, suggesting there may have been gene loss when laboratory mice were inbred. In stark contrast the κ locus, the mouse λ locus includes only three functional V genes.

In mice the λ locus is greatly reduced, suggesting there may have been gene loss when laboratory mice were inbred. In stark contrast the κ locus, the mouse λ locus includes only three functional V genes. These Vλ genes are separated in two topological domains with Vλ2 and Vλ3 associated with one functional Jλ2-Cλ2 pair, while Vλ1 is associated with two functional Jλ-Cλ pairs: Jλ1-Cλ1 and Jλ3-Cλ3 (Degner et al., 2009)(Clark Lab; Unpublished). Furthermore, the functions within these domains are believed to be under the control of enhancers 3-1 and 2-4, that are nearly identical in sequence (Brass et al., 1996; Eisenbeis et al., 1995; Hagman et al., 1990).

Genes of both human and mouse Vλ genes are in the same transcriptional orientation of their corresponding Jλ-Cλ gene clusters (Collins 18,39). Moreover, genes in the heavy-chain loci are all oriented in the same transcriptional direction (Wood and Tonegawa, 1983). Conversely, the κ locus is much more complex in both human and mouse. In mice, the κ locus includes *Vκ* genes that are in the opposite orientation to the *Jκ* cluster and *Cκ* gene. The same

is observed for the human κ locus with more distal *Vκ* genes in the opposite orientation and proximal *Vκ* genes sharing the same orientation as the *Jκ* and *Cκ* genes (Lefranc, 2001).

Early studies of mouse and human light chain rearrangement show that recombination begins at the κ alleles before progressing to the λ alleles (Alt et al., 1980; Hieter et al., 1981). Because of the stochastic nature of V(D)J recombination, failures are common (Karki et al., 2018). The κ locus is unique because its organization permits multiple attempts at recombination if the primary rearrangement was nonproductive or caused autoreactivity specificities of the resulting BCR. The process of additional rounds of secondary rearrangement is called receptor editing. Receptor editing is usually discussed in the context of salvaging autoreactive BCRs during development by allowing selection of a new light chain to pair with the pre-existing heavy chain that leads to loss of self-reactivity (Nemazee, 2006, 2017). However, receptor editing is crucial for shaping the general formation of the B cell repertoire (Casellas et al., 2001; Hippen et al., 2005; Retter and Nemazee, 1998). This will be discussed in greater detail in future sections.

While κ locus rearrangement occurs first, only one allele is targeted for recombination at a time. Secondary rearrangements at κ occur as long as an unrearranged *Vκ* gene remains 5' of a *VκJκ* rearrangement and unrearranged *Jκ* genes remain 3'. Primary rearrangements often occur with either *Jκ1* or *Jκ2* and subsequent secondary rearrangements during receptor editing occur with distal *Jκ4* and *Jκ5* genes. For this reason, many have speculated that κ recombination and by extension receptor editing is a sequential process, where during the first round of recombination a random *Vκ* gene will attempt recombination with *Jκ1*. If that rearrangement is nonproductive or produces self-reactivity

when paired with the heavy chain, subsequent rounds of recombination will utilize the remaining *Jκ* genes in sequential order.

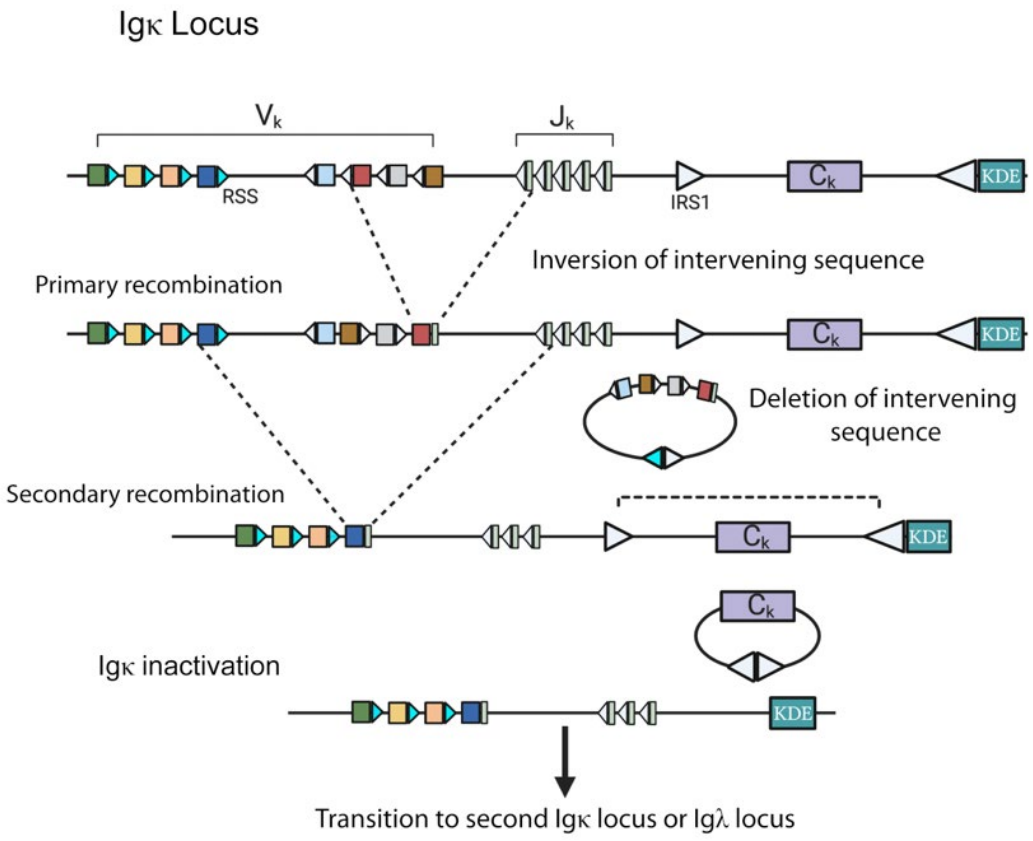


Figure 1.6 Igκ recombination and locus inactivation
Light chain recombination is a sequential process being first at the *Igκ* loci before progressing to the *Igλ* loci. The structure of the *Igκ* locus makes it suitable for repeated attempts at gene rearrangements due to failure of DNA recombination machinery or production of autoreactive specificities. Depicted is the series of events that would lead to inactivation of the *Igκ* locus and progression to the *Igλ* locus. A primary recombination event typically occurs between a 'randomly' chosen *Vκ* gene and *Jκ1*. If that recombination product fails to pair with the heavy chain or produces an autoreactive specificity, then secondary rearrangements can occur at the locus as long as unused *Jκ* genes are present and in frame. The *Igκ* locus can be inactivated by recombination of IRS1 to the KDE, which deletes the *Cκ* gene. Though not depicted, *Vκ* genes can also recombine with the KDE deleting or inverting the intervening gene elements, which would effectively inactivate the locus. The second *Igκ* allele or *Igλ* alleles would be used for further attempts at light chain recombination.

The presence of two light chain alleles poses a significant risk of isotypic inclusion. Single B cell expressing different BCR specificities has been associated with increased risk of

autoimmunity (Pelanda, 2014). Indeed, Systemic Lupus Erythematosis (SLE) patients have enrichment for B cells expressing both Igκ+ and Igλ+ BCRs. However, these 'dual expressing' cells are rare in the normal repertoire (Collins and Watson, 2018). The risk of isotypic inclusion is mitigated by the action of Kappa Deleting Elements (KDE). KDEs, referred to as Recombing Segment (RS) in mice, are highly conserved with the mammalian genome (Das et al., 2009). These sequences are conserved heptamer and nonamer sequences separated by 23 bps. The RS is located within the *Jκ-Cκ* intron and recombines to the RSS at the 3' ends of *Vκ* genes. Recombination of a *Vκ* gene to the RS effectively terminates recombination at that allele because it deletes the intervening *Jκ* sequences. This drives recombination from the first to the second κ allele or from the second κ allele to the λ locus. In this way, λ expressing B cells have lost expression of the κ chain, thus minimizing the potential for dual expression of *Igκ* and *Igλ* (Collins and Watson, 2018).

1.5.3 Tolerance of Self-Reactivity by Light Chains

Lack of the Diversity gene in the light chain limits the diversity of the light chain repertoire. An extensive body of research suggests that an important factor dictating the light chain repertoire is to minimize BCR self-reactivity. This research suggests Igκ+ human antibodies have a greater tendency to be autoreactive, yet through receptor editing at both κ and λ self-reactivity appears to be minimal. Several classical autoimmune disorders: Systemic Lupus Erythematosis (SLE), Type 1 Diabetes (T1D) and myasthenia gravis, have been associated with reduced light chain editing (Panigrahi et al., 2008; Vander Heiden et al., 2017).

Mouse models of autoimmune diseases also support the idea of a specialized role for light chains in tolerance. A classical SLE model uses MRL/*lpr* mice, which have a spontaneous mutation in *Fas* (CD95) a member of the TNF family of receptors that signals for death in the immune system. These mice develop autoimmunity rapidly and recapitulate many aspects of the human disease (Meffre and Wardemann, 2008). BCR transgenic mice cloned from naturally occurring autoantibodies from MRL/*lpr* mice show that receptor editing of the transgene light chain, spares these autoreactive B cells from negative selection (Gay et al., 1993).

1.6 B cell tolerance mechanisms

BCR expression at the early immature B cells stage is a key developmental checkpoint. 'Tonic signals' derived from the newly expressed BCR must be sufficient to suppress further light chain recombination ensuring allelic exclusion and promote survival whereby immature B cells gain competence to emerge from the BM to the spleen for further maturation (Hippen et al., 2005). Lastly, the early immature B stage is the site for tolerance against autoreactive BCR specificities (Nemazee, 2017). Due to the almost stochastic nature of Ig gene rearrangements the generation of BCRs possessing autoreactive specificities is high. Tolerance mechanisms must be equipped to minimize harm to self while not sacrificing the ability to defend against infection.

1.6.1 Clonal Deletion

In the late 1950s, Macfarlane Burnet proposed the clonal selection theory of lymphocyte tolerance to self-antigen (Burnet, 1959). In this theory, B cells clones with specificity to autoantigen are eliminated to prevent autoimmunity. Throughout the 1970s

and 1980s, numerous experiments injecting anti-IgM, anti-IgD, and anti-idiotype antibodies into young mice were performed to test this theory (Cerny et al.; Finkelman et al., 1983; Gause et al., 1987; Grandien et al., 1994; Lawton et al., 1972; Nemazee and Bürki, 1989; Takemori and Rajewsky, 1984). The experiments seemed to corroborate the selection theory as B cells expressing IgM, IgD and idiotype antibodies were eliminated when mice were injected with the corresponding autoantigens. In the 1980s, the availability of transgenic mice made testing selection theory using model antigens in vivo possible. Transgenic mice carrying B cell specificity for Hen egg lysozyme (HEL) or H2-$K^{k,b}$ MHC class I molecules further seemed to demonstrate that when immature B cells encounter autoantigen, these cells underwent apoptosis fairly rapidly after 2-3 days (Nemazee and Bürki, 1989). Negative selection was most prominent with high avidity, abundant membrane bound autoantigen. Interactions with low-avidity autoantigen did not result in deletion. These cells continued to differentiate, however; their lifespan in the periphery was shortened and these cells were eventually eliminated (Goodnow et al., 1988). These early experiments seemed to provide strong evidence for clonal deletion being the primary tolerance mechanism for developing B cells, however future studies would demonstrate clonal deletion was a 'last resort' mechanism after other processes mainly receptor editing, had been exhausted (Pelanda and Torres, 2012).

1.6.2 Receptor Editing

Receptor editing was discovered by David Nemazee and Martin Weigert through the analysis of mice expressing BCRs specific for either dsDNA or mouse MHC class I molecules (Gay et al., 1993; Tiegs et al., 1993). Weigert generated transgenic mice bearing the heavy and light chain for a well characterized antibody specific for dsDNA (3H9) isolated from the

lupus prone MRL/*lpr* mice. Repertoire analysis of mature B cells from adult mice bearing both heavy and light chain transgenes showed the heavy chain transgene paired with endogenous light chains instead of the prearranged transgene. The transgenic B cells transcribed both the light chain transgene as well as the endogenous rearranged light chain. However, only loss of the endogenous light chain resulted in re-expression of the 3H9 double transgene product. Conversely, Nemazee utilized a system where transgenic mice were specific for H2-K^b and H2-K^k MHC class I molecules. B cell development was compared to genetic backgrounds that expressed or lacked the autoantigen: that is, in $H2^b$ and $H2^k$ strains, or in $H2^d$ strains, respectively. In absence of autoantigen, $H2^d$ mice, B cell development was normal with almost all peripheral B cells expressing the transgene-encoded receptor. In the autoreactive context, $H2^b$ or $H2^k$, B cells encountered autoantigen in the BM and were shown to undergo continued recombination at the endogenous, non-transgenic, light chain loci ultimately expressing a new receptor with altered specificity. Moreover, recombination was also observed at the transgenic loci often destroying the transgene encoding the previously expressed Ig LC. Lastly, both of these models showed an increased usage of the λ locus as a consequence of altering receptor specificity (Nemazee and Weigert, 2000). Numerous groups have shown that receptor editing not only occurs frequently during normal B cell development but represents a major force in shaping the antibody repertoire (Casellas et al., 2001; Tze et al., 2005). Taken together, these findings demonstrated receptor editing is a mechanism used by autoreactive immature B cells to escape deletion.

Antigen avidity or the context of antigen encounter were thought to influence whether receptor editing, or clonal deletion took place. Again, using transgenic mice expressing the 3-83 antibody specific for H2-K^k MHC class I molecules or a mutated version

with reduced specificity for H2-K^k MHC class I molecules, Roberta Pelanda and colleagues showed receptor editing occurred equally in both settings (Halverson et al., 2004). Moreover, studies using mice bearing BCRs specific for HEL reported developing B cells cultured with HEL in a soluble or membrane bound form predominately underwent receptor editing to escape deletion (Hippen et al., 2005; Tze et al., 2003). Thus, clonal deletion is a consequence of failed attempts at receptor editing (Pelanda and Torres, 2012).

Receptor editing predominately occurs at light chain loci beginning with *Igκ*. Because the light chain does not contain an intervening Diversity region, *V*κ genes can directly recombine with downstream *J*κ gene elements. Secondary rearrangements can also occur at the heavy chain loci, called heavy chain replacement. However, because each V(D)J_H recombination even deletes all remaining D_H gene elements, secondary rearrangements cannot occur on the same allele. The second allele can then be rearranged and expressed leading to two heavy chains being expression, which is a rare event in WT mice (Barreto and Cumano, 2000). Likewise, secondary rearrangements at *Igλ* is unlikely because the locus contains only three functional V genes separated into distinct recombination clusters with their own constant regions (Carson and Wu, 1989; Pelanda and Torres, 2012; Prak et al., 1994). Thus, alternative methods of tolerance are necessary for B cells expressing auto reactive Igλ+ BCRs.

Receptor editing is not perfect and both insufficient and excessive editing can produce autoreactive BCR specificities (Verkoczy et al., 2004; Yachimovich-Cohen et al., 2003). Ig heavy chain most often dictates antigen specificity and sometimes a light chain is unable to tolerize autoreactivity specificities of the heavy chain (Luning Prak et al., 2011). Interestingly, reports have indicated the λ chain has distinct physiochemical properties and,

in some studies, λx has been shown to be particularly effective at quenching autoreactivity to DNA (Nemazee, 2017; Prak et al., 1994). Nevertheless, defective receptor editing has been associated with autoimmunity in both mice and man. Mouse models of SLE, T1D, and Sjögren's syndrome as well as patients with SLE have reported defects in receptor editing (Clark et al., 2013; Henry-Bonami et al., 2013; Lamoureux et al., 2007; Meng et al., 2012; Yurasov et al., 2005). Conversely, lupus prone NZB/NZW F1 mice utilize secondary rearrangements to generate high-affinity antibodies from low-affinity precursor cells (Yachimovich-Cohen et al., 2003). These findings suggest a narrow window in which receptor editing functions to properly tolerize autoreactivity in the developing B cell repertoire.

A current model of receptor editing postulates a developmental stepwise process whereby if newly developed $IgM^{lo}IgD^{-}$ immature B cells experience crosslinking of the BCR through antigen binding, this results in down-regulation of the BCR (Nemazee, 2006; Sandel and Monroe, 1999). The loss of BCR expression, seems to initiate a 'de-differentiation' program effectively returning immature B cells to the pre-B cell stage. Molecular studies in mice with an inducible deletion of surface Ig via Cre-mediated excision, showed a loss of surface Ig at the immature B cells stage resulted in the 'back differentiation' of these cells to an earlier stage of development characterized by the expression of Pro B genes (Schram et al., 2008; Tze et al., 2005). Immature B cells returning to the pre-B cell stage share a short lifespan of ~3 days for secondary LC recombination to rescue non-selected B cells by generating new BCRs that promote positive selection (Nemazee, 2017).

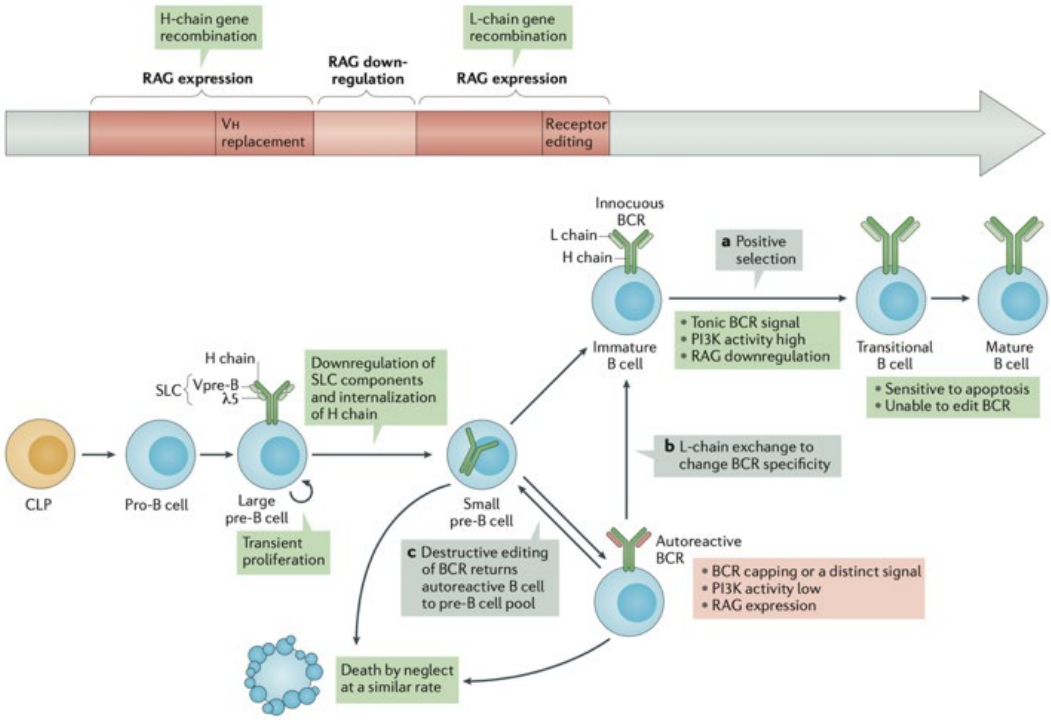


Figure 1.7 B cell central tolerance
The stochastic nature of B cell receptor assembly poses a significant risk of autoreactive BCR formation. Several tolerance mechanisms have been described to minimize this risk. Receptor editing and clonal deletion are prominent mechanisms that have been shown to tolerize autoreactive specificities in the bone marrow. (a) Immature B cells expressing a BCR capable of producing signals that repress light chain recombination and promote survival are selected to progress to the periphery for further differentiation. (b) Newly formed autoreactive immature B cells that encounter self-antigen in the BM, will signal strongly before rapid internalization and removal of the BCR from the cell surface. This reduced state of BCR signaling effectively returns immature B cells to the small pre-B cell stage where recombination can resume to replace the existing light chain (LC), editing the original BCR specificity. B cells can then be positively selected to develop to maturity if the replacement generates a non-self-reactive receptor. Clonal deletion occurs if the resultant light chains are unable to edit autoreactive BCR specificity to non-self. Figure adapted from (Nemazee, 2006)

Efforts to elucidate the molecular mechanisms of receptor editing have proved difficult. While many cells undergo this process, isolating editing cells from non-editing cells in vivo is currently not feasible, because no cell surface markers have been discovered. Nevertheless, BCR transgenic mice models and in vitro assays have proven the most useful in studying receptor editing. From these, a few hallmarks of receptor editing have been established: First, receptor editing involves the maintained expression of RAG1, RAG2 and

other proteins involved in Ig light chain recombination such IRF4 and E2A as a direct result of autoantigen mediated downregulation of the BCR (Beck et al., 2009; Nemazee, 2006). Second, Igλ$^+$ B cells depend on NF-κB signals, which can be substituted by enforced expression of pro-survival molecules such as BCL2 (Ait-Azzouzene et al., 2005; Casellas et al., 2001; Derudder et al., 2009; Sandel and Monroe, 1999). Third, the BM microenvironment and survival signals are necessary for efficient receptor editing to occur (Beck et al., 2014; Nagasawa, 2006; Sandel and Monroe, 1999).

As mentioned, E2A is crucial for early stages of B cell development and *Igκ* LC recombination. E2A has also been implicated in regulating receptor editing and *Igλ* recombination. Mice deficient for E2A gene products E12 or E47 were generated to explore whether they had distinct or redundant roles during early B cell development (Beck et al., 2009). The authors made notable observations that E47 was necessary for developmental progression at the pre-pro B cells stage, while E12 was dispensable. Interestingly, in pre-B and immature B cells E12 and E47 were critical for *Igλ* germline transcription and *VλJλ* gene rearrangements. They showed that increasing levels of E12 and E47 were necessary to promote accessibility at *Igλ* and promote recombination. Lastly, they showed E2A was capable of binding E box motifs present in the λ enhancers: Eλ3-1 and Eλ2-4. Another study demonstrated using the autoreactive 3-83 BCR transgenic model that E2A haplosufficiency impairs secondary LC gene arrangements and RS mediated κ inactivation (Quong et al., 2004). In sum, these data put forward a model that in pre-B and immature B cells, higher levels of E2A activity are required for initiation of receptor editing and *Igλ* recombination.

Another transcription factor implicated in receptor editing is NF-κB. In pre-B cells the induction of *Jκ* germline transcription and subsequent *VκJκ* rearrangements require NF-κB

binding to consensus sites within iEκ. In contrast, *Igλ* enhancers Eλ3-1 and Eλ2-4 contain no NF-κB motifs and thus function independently of NF-κB activity (Hagman et al., 1990). Yet, a number of studies have demonstrated diminished λ-chain usage in the absence of NF-κB activity. Using a transformed pre-B cell line (ts-abl) where inactivation of temperature sensitive v-abl oncogene results in RAG induction and light chain recombination, Bendall et al. demonstrated that cells transfected with the NF-κB inhibitor IκBα had less *Igλ* gene rearrangements (Bendall et al., 2001). This reduction in λ rearrangements was not due to a reduction of RAG expression but instead NF-κB was necessary for *Igλ* locus accessibility as λ germline transcription was lost in the absence of NF-κB. This suggested an indirect role for NF-κB in regulating *Igλ* locus.

In the context of autoreactivity, NF-κB was shown to mark cells engaged in receptor editing (Cadera et al., 2009). In mice where the IκBα gene was replaced with a lacZ β-gal reporter gene and expressed either autoreactive or non-autoreactive BCRs, β-gal was preferentially expressed in cells bearing an autoreactive BCR. A complimentary experiment where a repressor of IκBα was transfected into BM B cells cultures showed κ germline transcription was reduced while λ locus rearrangements were increased, but RAG expression was unaffected.

Apart from modulating locus accessibility, NF-κB has been implicated in promoting survival of pre-B cells undergoing λ gene rearrangements (Derudder et al., 2009). In an elegant set of experiments, Derudder et al generated B specific conditional deletions for a number of NF-κB activating genes including NEMO, Ikk1 and Ikk2. They observed a reduction in the formation of Igλ$^+$ immature B cells while Igκ$^+$ immature B cells were

unaffected. The dependency of Igλ$^+$ immature B cells on NF-κB was overcome by transgenic expression of pro-survival molecule BCL2. This led the authors to propose a model where LC recombination occurs in two phases. An early phase is dominated by Igκ recombination that proceeds independently of NF-κB activity. In the event that *Igκ* recombination fails or produces autoreactivity, cells will enter a later phase dominated by *Igλ* recombination. Cells in this phase, require pro-survival NF-κB signaling to allow for enough time to complete λ gene rearrangements.

The BM microenvironment is crucial for receptor editing to occur. In vivo, transitional cells in the spleen do not undergo receptor editing but instead experience negative selection or anergy tolerance mechanisms (Chung et al., 2003). However, in vitro, when transitional B cells were cocultured with RAG2 deficient BM and stimulated with anti-IgM, transitional B cells were capable of undergoing going receptor editing as evidenced by upregulation RAG2 and *Igλ* recombination (Sandel and Monroe, 1999).Transitional cells stimulated without BM cells underwent rapid negative selection and did not undergo receptor editing.

Early *in vitro* experiments studying B cell development relied on seeding immature B cells on OP9 and S17 BM derived stromal cell lines. Both OP9 and S17 are capable of promoting precursor B cells differentiation into IgM secreting cells (Cumano et al., 1990; Milne et al., 2005; Zetterblad et al., 2010). While OP9 expresses soluble IL-7 and CXCL12, S17 only expresses CXCL12. Experiments studying receptor editing, showed coculturing of S17 stromal cells with HEL specific immature B cells was permissive for receptor editing to occur (Tze et al., 2003). Furthermore, studies investigating B cell egress from the BM to the spleen observed that immature B cells stimulated with autoantigen upregulate CXCR4, presumably to retain autoreactive B cells in the BM to facilitate receptor editing (Beck et al., 2014). This

suggests CXCR4-CXCL12 signaling axis is important for B cell tolerance. Given the newly appreciated role of CXCR4 signaling in establishing small pre-B cell identity, CXCR4 expression on autoreactive immature B cells may extend beyond BM retention and may provide signals that initiation receptor editing.

1.6.3 Anergy

Receptor editing and clonal deletion mechanisms are efficient at minimizing autoreactivity to antigens found in the BM. However, B cells with autoreactive BCRs can escape these mechanisms and reach the periphery. Autoreactive B cells in the periphery lose the ability to receptor edit but instead are tolerized by clonal deletion and anergy (Chung et al., 2003; Goodnow et al., 2005). Anergy is an induced stage of hypo-responsiveness to antigen (Cambier et al., 2007). Typically, anergy is induced by prolonged exposure to low avidity monovalent, or multivalent antigens in absence of activating signals from T cells or Toll-like receptors.

The best studied model of B cell anergy utilizes HEL as an autoantigen. Transgenic mice expressing BCRs (IgM and IgD) for HEL were crossed to mice carrying a transgene that encodes soluble HEL expression. B cells in these double transgenic mice (termed MD4:ML5 mice) developed normally, however, the absolute number of mature B cells was reduced compared to mice lacking soluble HEL expression (Goodnow et al., 1988; Hartley et al., 1993; Hartley et al., 1991; Sabouri et al., 2016). In addition, there was a skew toward T3 cells in the Transitional B cell population, leading many to believe T3 cells are the anergic population during normal B cell development. Importantly, peripheral B cells in these mice failed to mount adaptive responses when immunized with exogenous HEL nor did they generate antibody-secreting cells in response to TLR stimulation with DNA and LPS agonists

(Goodnow et al., 1988). Mouse models with B cell specificity to other self-antigens such as Smith antigen, insulin and aszophenyl-arsonate have all shown development of anergic B cells (Cambier et al., 2007).

Several features have been identified to distinguish anergic B cells. Classically, the anergic B cells phenotype is down regulation of surface IgM but maintenance of surface IgD (Goodnow et al., 1988). It has been suggested that this phenotype helps maintain hypo-responsiveness because IgD signaling attenuates the response to self-antigen mediated by IgM signaling (Sabouri et al., 2016). Anergic B cells also have a reduced lifespan, with an estimated half-life of 3-4 days compared to 4-5 weeks of mature naïve B cells (Fulcher and Basten, 1994). Moreover, anergic B cells have impaired migration and are excluded from entering B cell follicles and instead are arrested at T-B cell boundaries (Cyster, 2010).

Lastly, another characteristic of anergic B cells is recruitment of phosphatases to the BCR signaling complex. Increased expression of SHIP-1 and SHP-1 inhibits signaling from the BCR and increases the threshold for anergic B cell reactivation. B-cell specific deletion of SHP-1 or SHIP-1 is associated with auto-antibody development and lupus-like diseases (O'Neill et al., 2011). Thus, anergy provides another mechanism of tolerance for autoreactive B cells (Oleinika et al., 2020).

1.7 Aims of thesis

As introduced, receptor editing has a major role in shaping the B cell repertoire by salvaging autoreactive B cells. Decades of research have helped identify key transcription factors and genomic events that govern this process. However, many questions remain as to the molecular mechanisms governing receptor editing. This thesis aims to address critical gaps in knowledge that hinder a complete understanding of this main mechanism of central tolerance.

1. A primary criticism of receptor editing research to date, is the obligatory use of B cells expressing mono-specific transgenic BCRs. While these systems have been very useful in providing a homogenous population of cells undergoing the same processes, it does warrant the question of what molecular aspects of receptor editing are observed in a polyclonal setting. For instance, the 'back differentiation' of editing cells upon inhibition of BCR tonic signaling was greatest in mice bearing Ig transgenes than from mice with a polyclonal repertoire. Ig transgenic mice were shown to upregulate VpreB, TdT and λ5 genes associated with the pro-B cell stage while the WT immature B cells upregulated RAG and Ku70 genes (Tze et al., 2005). Thus, an examination of receptor editing in a polyclonal setting is necessary to improve our molecular understanding of receptor editing.
2. The role of chromatin accessibility and epigenetic state in editing cells are also completely unknown. This is easily explained by the lack of technology to easily assess chromatin accessibility during much of the 1990s and early 2000s. The introduction of Assay for Transposase Accessible Chromatin sequencing (ATAC)-seq

technology easily allows for addressing this fundamental gap in knowledge (Buenrostro et al., 2013).

3. Receptor editing promotes the usage of λ-chain; yet, despite reports that estimate 50% of developing B cells are targeted for tolerance, the overall understanding of the *Igλ* locus is poor. Mice have a B cell repertoire containing a ratio of 20:1 Igκ$^+$ to Igλ$^+$ B cells; whereas, the human B cell repertoire has an Igκ$^+$ to Igλ$^+$ ratio of ~1:1, demonstrating an equal importance of the λ-chain to a mature B cell repertoire. Furthermore, reports have indicated the λ-chain has distinct physiochemical properties and, in some studies, λx has been shown to be particularly effective at quenching autoreactivity (Gay et al., 1993; Prak et al., 1994). Therefore, understanding the regulatory mechanisms governing λ-chain recombination in B cell central tolerance will further improve our knowledge of B cell repertoire formation.
4. Our lab recently demonstrated an unprecedented role for CXCR4-CXCL12 signaling axis in establishing small pre-B cell identity and regulating *Igκ* recombination. Combined with the observation that autoreactive immature B cells upregulate CXCR4 when stimulated with auto-antigen, this provides rationale for CXCR4 signaling regulating receptor editing (Beck et al., 2014; Greaves et al., 2019; Schram et al., 2008; Tze et al., 2005).

2 Results

2.1 Molecular mechanisms of B cell receptor editing.

During B lymphopoiesis, B cell progenitors progress through alternating and mutually exclusive stages of clonal expansion and immunoglobulin (Ig) gene rearrangements (Clark et al., 2014; Hamel et al., 2014; Karki et al., 2018). Great diversity is generated through the stochastic, nonhomologus recombination of Ig gene segments encoding heavy and light chain variable domains. However, this also commonly generates autoreactivity. Indeed, up to 75% of immature B cells express autoreactive antigen receptors (Giltiay et al., 2012; Meffre and Wardemann, 2008; Wardemann et al., 2003). Therefore, the initial B cell repertoire must be purged of autoreactivity in the bone marrow (BM), before transit to the periphery.

B cells are primarily tolerized in the bone marrow (BM) by either negative selection (apoptosis) or receptor editing. Previous studies have shown that receptor editing is the predominant central tolerance mechanism for self-reactive B cells (Casellas et al., 2001; Hippen et al., 2005; Retter and Nemazee, 1998). Receptor editing rescues autoreactive B cells from negative selection, by inducing renewed light chain recombination in attempts to edit antigen specificity away from self (Gay et al., 1993; Nemazee, 2006; Radic et al., 1993). Light chain editing occurs first at *Igκ* then proceeds to *Igλ* loci. Renewed light chain gene rearrangements at *Igκ* inactivates the previous *VκJκ* gene rearrangement by genomic inversion or deletion (Collins and Watson, 2018; Vela et al., 2008), while subsequent recombination at the original allele uses downstream *Jκ* segments (Collins and Watson, 2018). Once at *Igλ* the choice of recombination partners is limited with studies showing only 4 major products are produced: λ1, λ2, λ3 and λx. Yet, reports have indicated the λ-chain has

distinct physiochemical properties and, in some studies, λx has been shown to be particularly effective at quenching autoreactivity to DNA (Gay et al., 1993; Nemazee and Weigert, 2000; Prak et al., 1994) In this manner receptor editing promotes λ-chain usage, but due to the low representation of *Igλ* in the mouse repertoire, a complete understanding of the molecular mechanisms leading to *Igλ* usage during receptor editing is poor.

Mouse models utilizing transgenic expression for BCRs specific for model antigens have been useful in elucidating various molecular aspects of receptor editing and *Igλ* recombination (Gay et al., 1993; Halverson et al., 2004; Pelanda et al., 1997; Tiegs et al., 1993; Tze et al., 2005). It is clear that at the immature B cell stage, the loss of BCR expression after self-antigen ligation, initiates a 'de-differentiation' program effectively returning immature B cells to the small pre-B cell stage (Tze et al., 2005). Here, continuous RAG protein expression (Yu et al., 1999) mediates renewed light chain recombination at both *Igκ* and *Igλ* (Gay et al., 1993; Tiegs et al., 1993). Moreover, Igλ+ B cell formation is dependent on NF-κB signals, which are presumed to promote survival and increase the developmental window to allow switching from *Igκ* to *Igλ* (Derudder et al., 2009). Conversely, increasing levels of E2A proteins E12 and E47 promote *Igλ* germline transcription and recombination (Beck et al., 2009). However, it is currently unclear how the *Igλ* locus becomes accessible to recombination during receptor editing.

Failure to successfully rearrange the *Igκ* alleles also leads to *Igλ* recombination. Indeed, recent single cell studies indicate that most initial *Vκ-Jκ* rearrangements are out of frame (Karki et al., 2018). Therefore, there are two very different mechanisms leading to *Igλ* recombination, autoreactivity and failed *Igκ* recombination. These developmental

trajectories leading to *Igλ* recombination for each are radically different and have not been fully described.

Here we demonstrate cells undergoing receptor editing possess unique epigenetic and transcriptional states. Editing small pre-B cells acquired unique open chromatin regions shared with immature B cells and were enriched for NF-κB and E2A transcription factor motifs as well as increased accessibility at the *Igλ* locus. The epigenetic landscape of editing small pre-B cells defined the transcriptional profile of immature B cells. Editing immature B cells were enriched for metabolic and DNA repair pathways as well as upregulation of chemokine receptor CXCR4. CXCR4 deficiency in the setting of autoreactivity resulted in loss of many pathways associated with receptor editing as well as a selective defect in *Igλ* transcription and recombination.

2.2 Two molecular paths to *Igλ* recombination

Igλ expressing B cells represent a small (~10%) proportion of the mouse B cell repertoire, hindering studies of their development in a polyclonal wild-type (WT) setting. Therefore, we utilized two complimentary mouse models that recapitulate the different pathways by which Igλ bearing B cells develop and are selected. The first has a deletion in the J kappa (Jκ) genes and as such models Igλ expression resulting from failed *Igκ* recombination ($Ig\kappa^{del}$) (Chen et al., 1993). The second model utilizes a 'κ-macroself' transgene, which encodes for ubiquitous expression of a single-chain chimeric anti-Igκ antibody (κ-mac) (Ait-Azzouzene et al., 2005). In this model, newly formed Igκ expressing immature B cells are "autoreactive" and forced to edit away from Igκ-chain expression in favor of Igλ expression in order to continue differentiation.

Flow cytometric analysis of WT, $Ig\kappa^{del}$ and κ-mac bone marrow revealed equivalent BM cellularity but different distribution of B cell progenitor subsets (Figure 2.1A, B). Pre-pro B cells (B220+CD19-) were expanded in the κ-mac mice compared to WT and $Ig\kappa^{del}$ mice (Figure 2.1C). However, fully committed pro-B (B220+CD19+CD43+) and subsequent large pre-B (B220+CD19+CD43-IgM-FSC^{lo}) cell numbers were similar between the three genotypes (Figure 2.1D, E). Notably, $Ig\kappa^{del}$ mice had an expanded small pre-B cell (B220+CD19+CD43−IgM−FSC^{lo}) compartment compared to WT and κ-mac mice (Figure 2.1F). Because light chain recombination is sequential, it could be possible that preventing *Igκ* recombination would impair subsequent *Igλ* recombination. However, we observed similar numbers of immature B (B220loCD19+CD43−IgM+) and recirculating mature B cells

between WT and Igκ^{del} mice but a reduced population in κ-mac mice (Figure 2.1G, H). This is presumably a reflection of more cells undergoing negative selection in κ-mac mice as a result

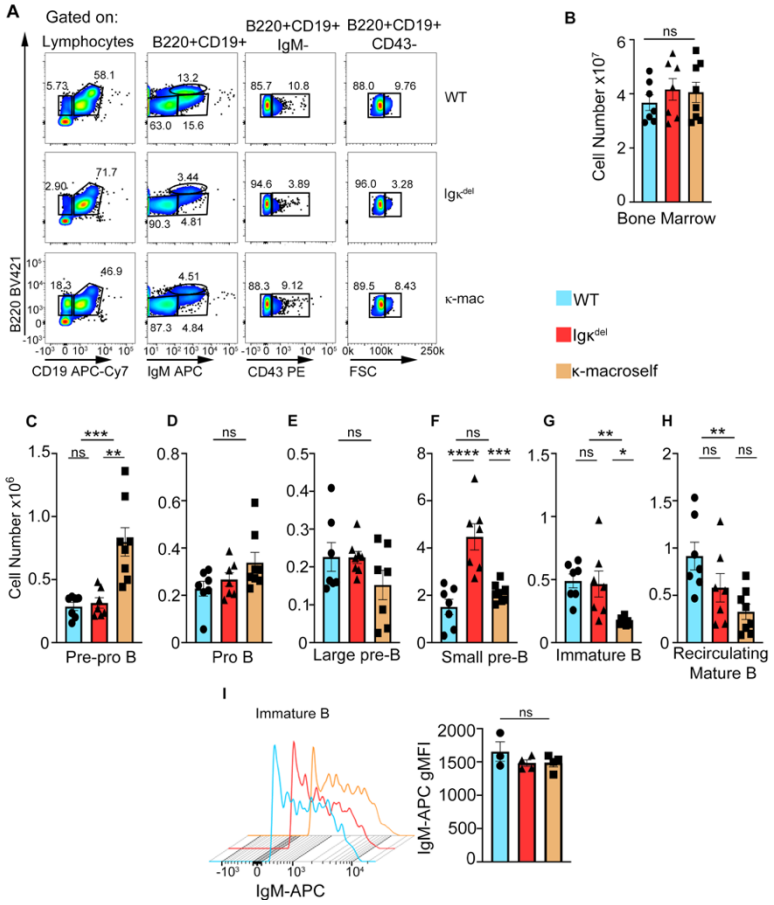


Figure 2.1 Two molecular paths to *Igλ* recombination

A, Representative flow cytometric analysis of different developmental stages of B lymphopoiesis in the BM of WT, Igκ^{del} and κ-mac mice. B cell progenitors subpopulations are defined as follows: Progenitor gate (B220+CD19+), pro-B cells (B220+CD19+CD43+IgM−), large pre-B cells (B220loCD19+CD43−IgM−FSChi) and small pre-B cells (B220loCD19+CD43−IgM−FSClo) and immature B cells (B220loCD19+CD43−IgM+). FSC; Forward Scatter. **B**, Absolute numbers of bone marrow cells from WT (n=7), Igκ^{del} (n=7) and κ-mac mice (n=8). **C**, Absolute number of Pro B cells (B220+CD19+CD43+IgM−) from WT (n=7), Igκ^{del} (n=7) and κ-mac mice (n=8). **D**, Absolute number of large pre-B cells (B220+CD19+CD43−IgM−FSChi) from WT (n=7), Igκ^{del} (n=7) and κ-mac mice (n=8). **E**, Absolute number of small pre-B cells (B220+CD19+CD43−IgM−FSClo) from WT (n=7), Igκ^{del} (n=7) and κ-mac mice (n=8). **F**, Absolute number of immature B cells (B220loCD19+CD43−IgM+) from WT (n=7), Igκ^{del} (n=7) and κ-mac mice (n=8). **G**, Absolute number of mature recirculating B cells (B220hiCD19+CD43−IgM+) from WT (n=7), Igκ^{del} (n=7) and κ-mac mice (n=8). **I**, Flow cytometric analysis of the corresponding cell-surface expression of IgM (top panel) and quantification of mean fluorescence intensity (MFI, bottom panel) on immature B cells from WT (n=7), Igκ^{del} (n=7) and κ-mac mice (n=8). **B-I**, Data are presented as means ± standard error of the mean (S.E.M). P values were determined by ANOVA with Tukey multiple testing (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

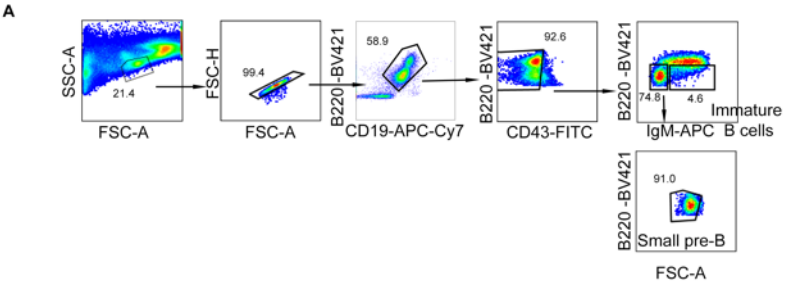
of a majority of cells undergoing receptor editing. Nevertheless, BCR surface expression on immature B cells was similar across all three genotypes (Figure 2.1I). Together, these results demonstrate the neither loss of Igκ recombination capacity nor the κ-macroself severely impair B cell development. Instead, these two models highlight the potential for unique molecular mechanisms governing the two paths to *Igλ* recombination as stages involved in light chain recombination were altered.

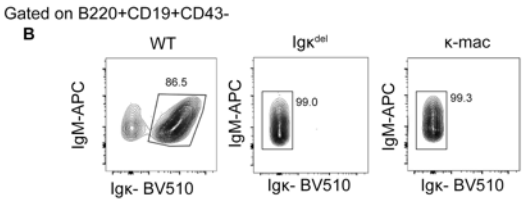
2.3 Gating Strategy

To investigate the gene-regulatory networks that govern receptor editing and *Igλ* recombination, we first used FACS to isolate WT, Igκ^{del} and κ-mac small pre-B cells as well as WT Igκ+, Igκ^{del} and κ-mac Igκ- immature B cells (Figure 2.2A).

Figure 2.2 Gating Strategy

A, Gating strategy for sorting immature B cells and small pre-B cells. **B**, Flow cytometric analysis of Igκ surface expression in WT, Igκ^{del} and κ-mac B cells.





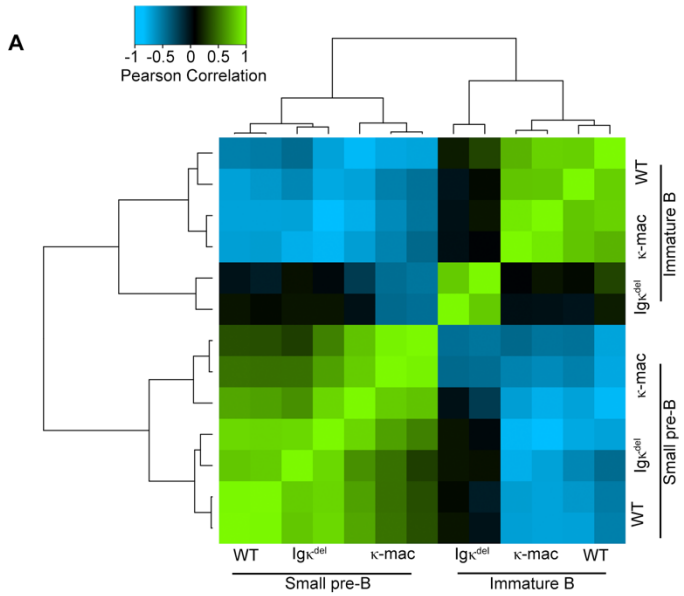
Consistent with published studies, Igκ^{del} and κ-mac mice exclusively produced Igλ+ B cells while WT mice overwhelming produced Igκ+ B cells (Figure 2.2B). For these and subsequent studies, we sorted Igκ- immature B cells from Igκ^{del} and κ-mac mice because commercially available anti-Igλ antibodies are specific for 3 of the 4 known Igλ products: λ1,

λ2, λ3 but not λx. These populations were then subjected to Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) to determine open chromatin regions (OCRs) and RNA sequencing (RNA-seq) to determine transcriptional profiles of WT, Igκ^{del} and κ-mac cells.

2.4 Editing cells have unique chromatin profile

We first analyzed the ATAC-seq data to determine the epigenetic landscape using ChromVAR to assess transcription factor (TF) motif enrichment within accessible chromatin regions of sorted cell populations (Schep et al., 2017) (Figure 2.3A). ChromVAR defines a 'deviation score', which is a measure of the change in accessibility at a particular motif within a sample compared to the average accessibility of that motif in all samples analyzed. We performed clustering on Pearson correlations of normalized deviation scores between sorted samples revealing that large pre-B, small pre-B and immature B cells possessed divergent patterns chromatin accessibility.

Figure 2.3 Pearson correlation of ATAC samples



A, Heatmap of Pearson correlation between samples based on the normalized deviation scores of transcription factor motif accessibility using the ChromVAR R package

Comparisons within each developmental stage revealed clear differences between the WT, Igk^{del} and k-mac genotypes. In particular, in small pre-B cells,

k-mac cells were different from WT and Igk^{del} cells which were relatively similar. In contrast, in immature B cells, WT and κ-mac cells were more similar to each other than Igk^{del} cells. These data suggest that autoreactivity is associated with an early epigenetic reprogramming in small pre-B cells.

2.5 Selective enrichment of E2A and NF-κB motifs in κ-mac B cells

Next, we assessed the TF motifs selectively enriched in between the cell types (Figure 2.4A). Numerous TF motifs necessary for early B cell development such as EBF1 (Rothenberg, 2014), PU.1 (Batista et al., 2017; Christie et al., 2015; Schweitzer and DeKoter, 2004) and MEF2C/MEF2D (Herglotz et al., 2016; Wilker et al., 2008) showed large changes in accessibility across the large pre-B to immature B cell transitions. WT, Igκ^{del} small pre-B cells and Igκ^{del} immature B cells showed increased accessibility at TF motifs involved in chromatin structure (CCCTC-binding factor (CTCF) CTCF and the CTCF paralog BORIS compared to WT and κ-mac small pre-B and immature B cells. This is consistent with findings that during early B cell development small pre-B cells demonstrate a substantial decrease in accessible regions compared to other stages of development (Mandal et al., 2018). Interestingly, κ-mac small pre-B cells and WT large pre-B cells were selectively enriched for accessibility at E-box motifs (HEB and E2A) and Early B cell Factor 1 (EFB1) motifs.

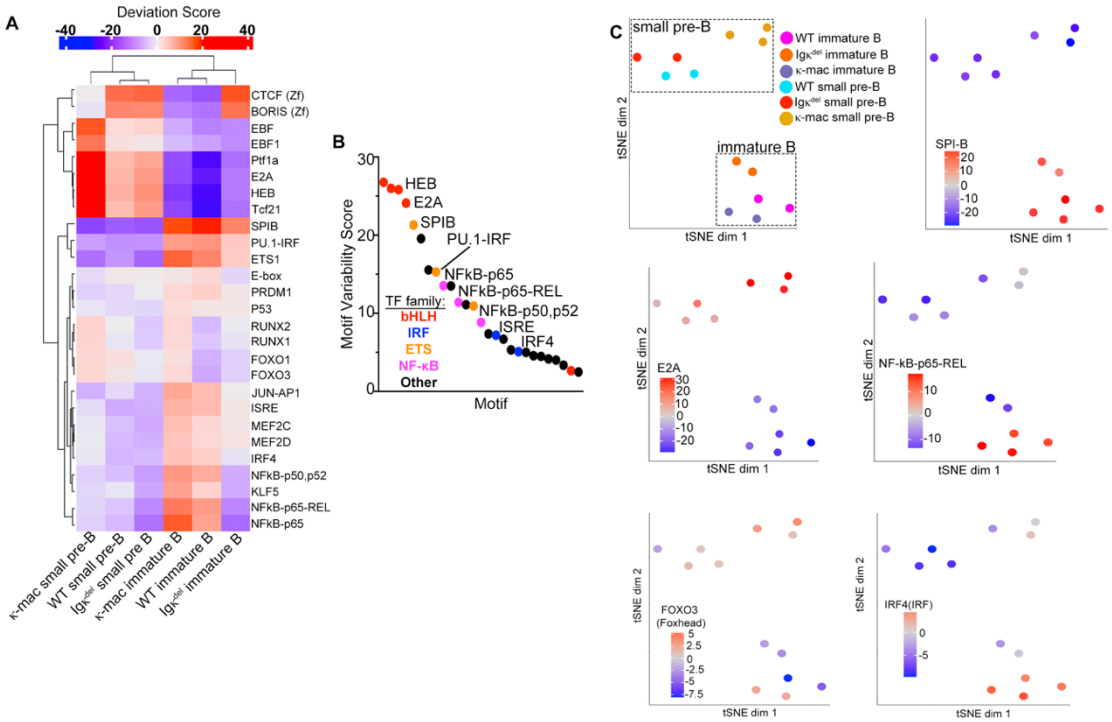


Figure 2.4 Selective enrichment of E2A and NF-κB motifs in κ-mac small pre-B cells
A, An average of top transcription factor motif deviation scores in indicated cell populations. Red indicates an increase in deviation score and blue indicates a decrease in deviation score. Hierarchical clustering performed on Euclidean distances. **B**, Top transcription factor motifs ranked by motif variability score. **C**, T-stochastic neighbor embedding of transcription factor motifs for SPI-B, E2A and composite motifs for NF-κB-p65-REL for indicated cell populations. Scale bar represents z score of deviations calculated using the ChromVAR R package. TSNE was generated using variability threshold = 3.5 and perplexity = 4. Motifs were taken from the "homer_pwms" database.

We next ranked TF with the highest change in accessibility by using ChromVAR's 'variability score', which is the standard deviation of the deviation z-scores across samples (Figure 2.4B). The highest enrichment was for E protein motifs, which are highly represented in large pre-B and κ-mac small pre-B cells. IRF, ETS and NF-κB motifs were also among the highly ranked TF motifs.

T-distributed stochastic neighbor embedding (t-SNE) showed developmental stage dependent clustering with small pre-B cells and immature B cells forming separate clusters (Figure 2.4C). Within these clusters, NF-κB and SPIB motifs were strongly enriched in WT

and κ-mac immature B cells compared to Igκ^{del} immature B cells, while E2A was enriched within WT large pre-B and κ-mac small pre-B cells. These findings are consistent the requirement of both E2A and NF-κB for receptor editing (Beck et al., 2009; Bendall et al., 2001; Cadera et al., 2009; Derudder et al., 2009; Verkoczy et al., 2005). Furthermore, Foxhead motifs (FOXO3) were specifically enriched in κ-mac small pre-B and immature B cells while enrichment for IRF4 motifs were shared between WT and κ-mac immature B cells and WT large pre-B cells. Together, these results demonstrate that in κ-mac small pre-B cells open chromatin regions are enriched for E2A sites while at the immature B cell stage, there is preferential accessibility at SPIB and NF-κB sites.

2.6 Selective opening of *Igλ* locus in editing B cells

Next, we quantified chromatin accessibility at *Igκ* and *Igλ* loci across all encoding genes and regulatory segments. These results were then plotted as a function of relative position across all encoding gene bodies. To capture accessibility at recombination signal sequences (RSS) and gene promoters, flanking regions immediately before and after gene segments were included in the analysis. Remarkably, the *Igλ* locus of κ-mac small pre-B cells (gold) showed increased accessibility compared to both Igκ^{del} (red) and WT (blue) small pre-B cells (Figure 2.5A). Accessibility at *Vλ* and *Jλ* gene segments was significantly increased in κ-mac small pre-B while being moderately increased in immature B cells (Figure 2.5B). In contrast, accessibility at *Igκ* remained largely unchanged in Igκ^{del} and κ-mac small pre-B and immature B cells (Figure 2.5C). As expected, no accessibility was observed at *Jκ* gene

elements in Igκ^{del} small pre-B and immature B cells (Figure 2.5D). Altogether, these data suggest receptor editing is associated with selective opening of *Igλ* locus.

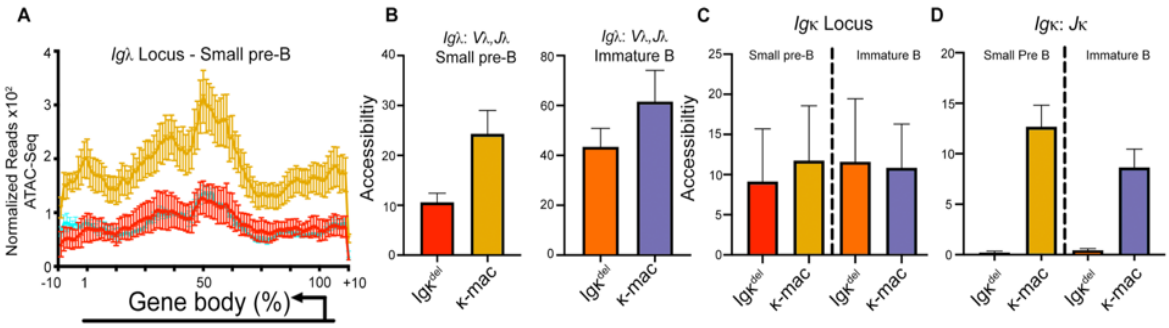


Figure 2.5 Selective opening of *Igλ* locus in editing B cells
A, Meta-analysis of accessibility (ATAC-seq) over the length of each Igλ gene body or regulatory element for WT (Blue), Igκ^{del} (Red) and κ-mac (Gold) small pre-B cells. **B**, Accessibility calculated from - 10% to +10% relative to the transcription start site (TSS) for *Vλ* and *Jλ* small pre-B cells in Igκ^{del} (Red) and κ-mac (Gold) small pre-B cells and Igκ^{del} (Orange) and κ-mac (Purple) immature B cells. **C**, Accessibility calculated from - 10% to +10% relative to the transcription start site (TSS) for *Igκ* genes (*Vκ*, *Jκ*, *Cκ)* small pre-B cells in Igκ^{del} (Red) and κ-mac (Gold) small pre-B cells and Igκ^{del} (Orange) and κ-mac (Purple) immature B cells. **D**, Accessibility calculated from - 10% to +10% relative to the transcription start site (TSS) for *Igκ* genes *Jκ* small pre-B cells in Igκ^{del} (Red) and κ-mac (Gold) small pre-B cells and Igκ^{del} (Orange) and κ-mac (Purple) immature B cells.
A-D) Error bars represent S.E.M.

2.7 Unique chromatin landscape in editing small pre-B cells

Since light-chain recombination occurs in small pre-B cells and it has been shown that autoreactive immature B cells return to the pre-B cell stage upon downregulation of autoreactive receptor (Tze et al., 2005), we performed an in-depth analysis of the open chromatin regions using ATAC-seq data specifically in WT, Igκ^{del} and κ-mac small pre-B cells. As expected, a majority of accessible regions (25,142) were shared between the all three small pre-B genotypes (Figure 2.6A). In contrast, κ-mac small pre-B cells had ~11k unique accessible regions compared to ~1.4k and ~1.7k accessible regions in WT and Igκ^{del} cells respectively. We then annotated the unique accessible regions to promoters, exons, introns

and intergenic regions (Figure 2.6B). Unique κ-mac accessible regions were found at a higher frequency at promoters and exons, compared to WT and Igκ^{del} accessible regions that were somewhat frequent at promoters but not at exons. Igκ^{del} and WT small pre-B accessible regions were mostly found at intergenic regions. Annotation enrichment analysis of observed/expected genomic annotations confirmed enrichment at promoters and exons in κ-mac small pre-B cells compared to Igκ^{del} and WT small pre-B cells (Figure 2.6C).

We then performed gene ontology (GO) pathway analysis of genes in the vicinity of the unique accessible regions (Figure 2.6D). Genes near κ-mac accessible regions were involved in NF-κB and P53 signaling. These pathways are known to be important for light chain recombination and DNA repair respectively (Bendall et al., 2001; Cadera et al., 2009; Derudder et al., 2009; Verkoczy et al., 2007; Williams and Schumacher, 2016). Interestingly, κ-mac open chromatin regions were found near genes involved in cytoprotective response to reactive oxygen species (Keap1-Nrf2 pathway) and chemotaxis. These findings could reflect editing cells establish a unique chromatin landscape to drive transcriptional programs that enable recombination and protection from genotoxic stress.

2.8 Motif enrichment of editing small pre-B cells

Next, we examined the unique accessible regions in WT, Igκ^{del} and κ-mac small pre-B cells for transcription factor binding motifs (Figure 2.7A). Analysis revealed CTCF as a common motif in all accessible regions. Antigen receptor signaling molecule Nuclear Factor of Activated T cells (NFAT) was enriched in WT accessible regions and TWIST1 was enriched in Igκ^{del} accessible regions. In contrast, unique κ-mac accessible regions were enriched for NF-κB subunit REL-B and PU.1 containing composite motifs with IRF and SPI-B transcription

factors. Analysis of motifs within assessible regions shared by all three genotypes termed 'Common - Motifs' revealed enrichment for SPI-B, CTCF, E2A, PRMD1, RUNX, and MEF2C (Figure 2.7B). However, single base-pair resolution mapping of accessibility around these shared TF motifs suggested preferential binding of CTCF, SPI-B and E2A in κ-mac small pre-B cells as we observed significant accumulation of accessibility surrounding these motifs (Figure 2.7C-E). These findings suggest editing small pre-B cells adopt a chromatin landscape poised for *Igλ* recombination with increased accessibility and TF motif enrichment for known mediators of receptor editing and light chain recombination.

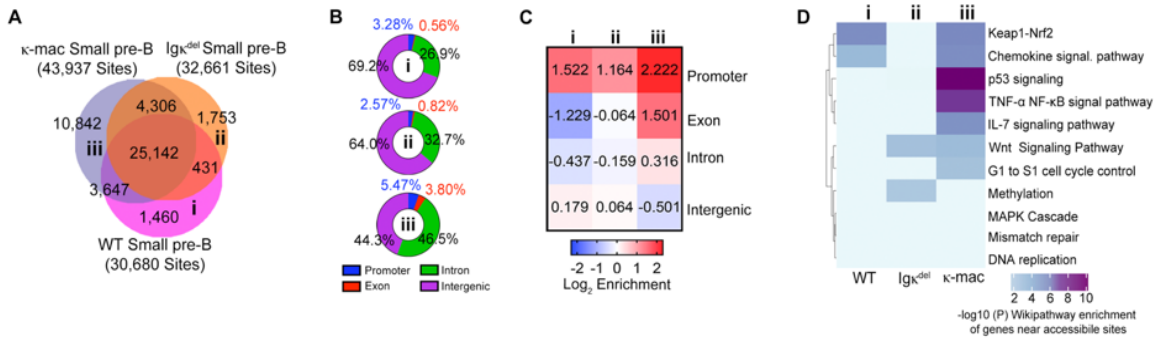


Figure 2.6 Unique chromatin landscape in editing small pre-B cells
A, Total and overlapping open chromatin regions (ATAC-Seq) in flow-purified WT, Igκ^{del} and κ-mac small pre-B cell populations. **B**, Distribution of unique accessible regions identified in A across the genome in promoters (Prom), Exon, Intron, and intergenic regions (Inter.) Roman numerals indicate unique regions: i) WT (n=2), ii) Igκ^{del} (n=2), iii) κ-mac (n=3) small pre-B cells. **C**, Enrichment statistic for peak overlap with different sets of annotation identified in B. **D**, Gene ontology analysis of genes near unique accessible regions identified in A. 'Wikipathway' gene sets were used for this analysis. Scale bar represents log transformed P values obtained using HOMER (see methods).

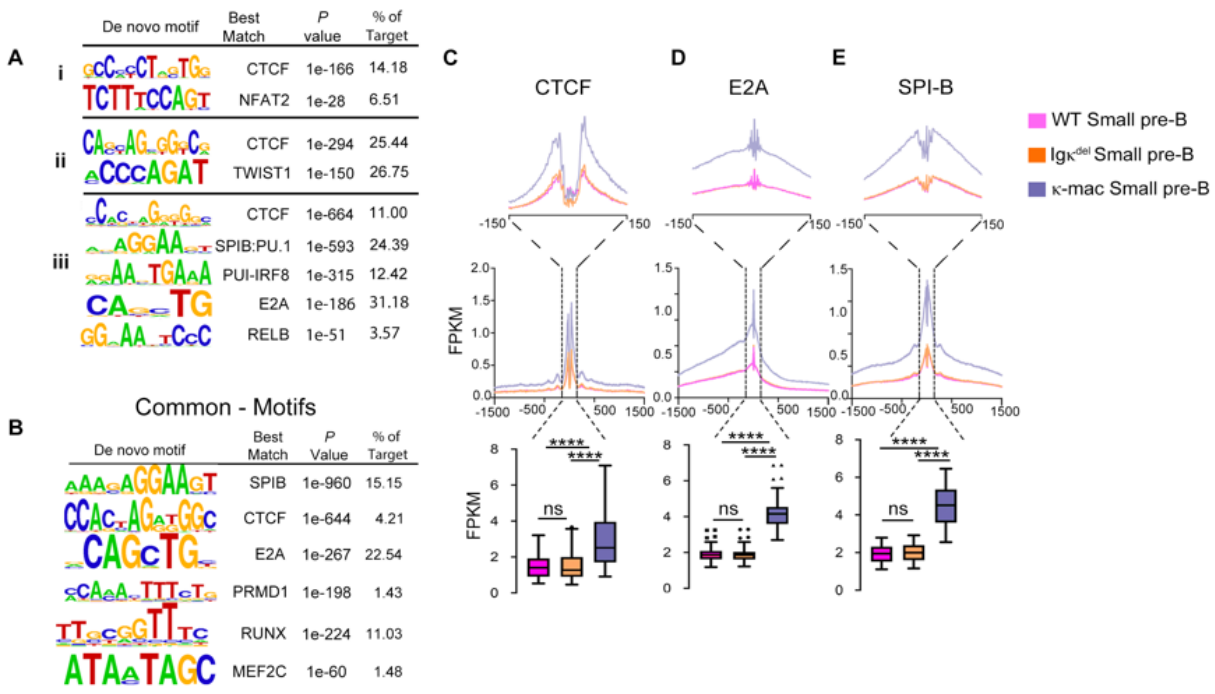


Figure 2.7 Motif enrichment of editing small pre-B cells

A-B, De Novo transcription factor binding motifs associated with unique and common accessible regions. P value and percent of Target (% of Target) determined using HOMER (see methods). **C-E**, ATAC-seq signal single-base-pair resolution at motif-centered peaks containing the *de novo* discovered motifs of CTCF, E2A and SPI-B in WT (Magenta), Igκ^{del} (Orange), or κ-mac (Purple) small pre-B cells. A 300bp (top panel) or 3000bp (middle panel) window around centered motif is shown. Bottom panel shows quantification of accessibility with a 300bp window of centered motif. Boxes represent interquartile ranges (IQRs; Q1–Q3 percentile) and black vertical lines represent median values. Maximum and minimum values (ends of whiskers) are defined as Q3 + 1.5× the IQR and Q1 − 1.5× the IQR, respectively. Outliers are indicated as black dots along the whiskers. Statistical significance determined using ANOVA followed by Kruskal-Wallis multiple comparisons (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

2.9 Unique κ-mac small pre-B open chromatin regions are shared with WT immature B cells

Autoreactive immature B cells return to the small pre-B compartment upon downregulation of the autoreactive BCR. We wondered whether the ~11k unique chromatin regions not shared with WT or Igκ^{del} small pre-B cells were shared with WT immature B cells. Open chromatin regions from WT immature B cells were overlapped with the ~11k unique OCRs

from κ-mac small pre-B cells. We saw a majority of these unique regions were shared with immature B cells (Figure 2.8A).

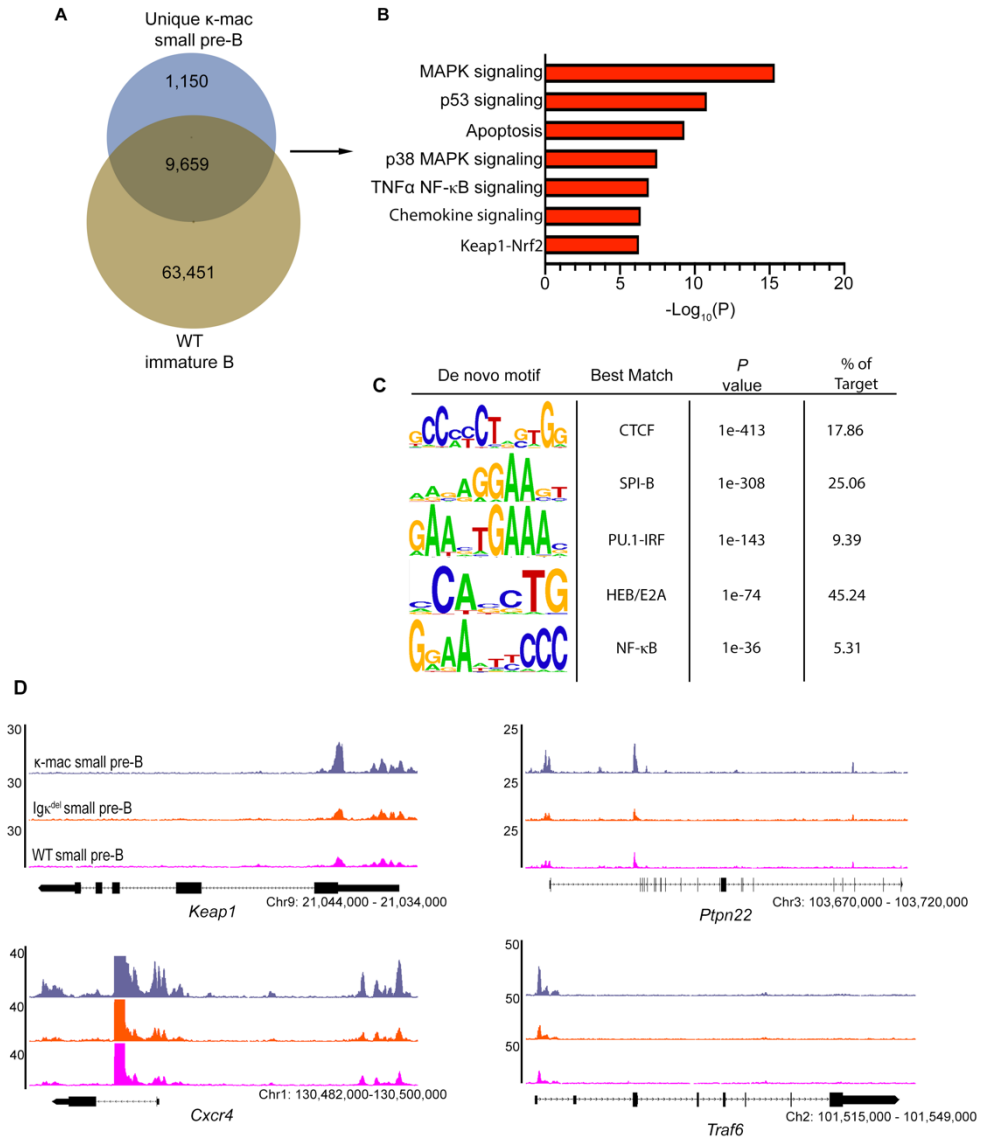


Figure 2.8 Unique κ-mac small pre-B open chromatin regions are shared with WT immature B cells

A, Venn diagram showing the overlaps of open chromatin regions between WT large pre-B and κ-mac small pre and immature B cells. **B**, Gene ontology analysis of genes near DAR depicted in A. 'Wikipathway' gene sets were used for this analysis. Scale bar represents Log_{10} transformed P values obtained from this analysis. **C**, HOMER transcription motifs analysis of shared OCRs. **D,** Genome accessibility at *Keap1, Ptpn22, Cxcr4, and Traf6* in κ-mac, $Igκ^{del}$ and WT small pre-B cells.

Gene pathway enrichment analysis of gene in the vicinity of these shared regions showed the strongest enrichment for MAPK signaling and P53 signaling. Moreover, these OCRs were enriched for key transcription factors such as SPIB, E2A and NF-κB motifs (Figure 2.8C). Indeed, accessibility at genes important for these gene pathways such as *Cxcr4*, *Keap1*, *Ptpn22* and *Traf6* all showed increased accessibility in κ-mac small pre-B cells compared WT and Igκ^{del} small pre-B cells (Figure 2.8D). Thus, the epigenetic landscape of κ-mac small pre-B cells shows evidence of having progressed into the immature B cell stage only to subsequently return to the small pre-B cell stage in order to edit.

2.10 Small pre- and immature κ-mac B cells have unique transcriptional programs

To investigate the transcriptional differences between editing cells and non-editing cells, we performed RNA-sequencing on FACS isolated small pre-B and immature B cells from WT, Igκ^{del} and κ-mac mice. As expected, principal component analysis (PCA) indicated distinct differences between small pre-B and immature B cells, regardless of genotype (Figure 2.9A). Strikingly, small pre-B cells clustered together suggesting remarkable similarity in transcriptional profiles. In contrast, immature B cell showed genotype dependent clustering. Spearman correlation of the 10,408 (FDR < 0.05) differentially expressed genes between WT, Igκ^{del} and κ-mac small pre-B and immature B cells confirmed the PCA clustering (Figure 2.9B). Independent of genotype, small pre-B cells had high spearman correlation values with Igκ^{del} and κ-mac small pre-B cells being virtually identical with a spearman correlation of 0.98. In contrast, spearman correlation values indicated that WT, Igκ^{del} and κ-mac immature B cells were transcriptionally more distinct with spearman correlation values ranging from 0.67 to 0.83.

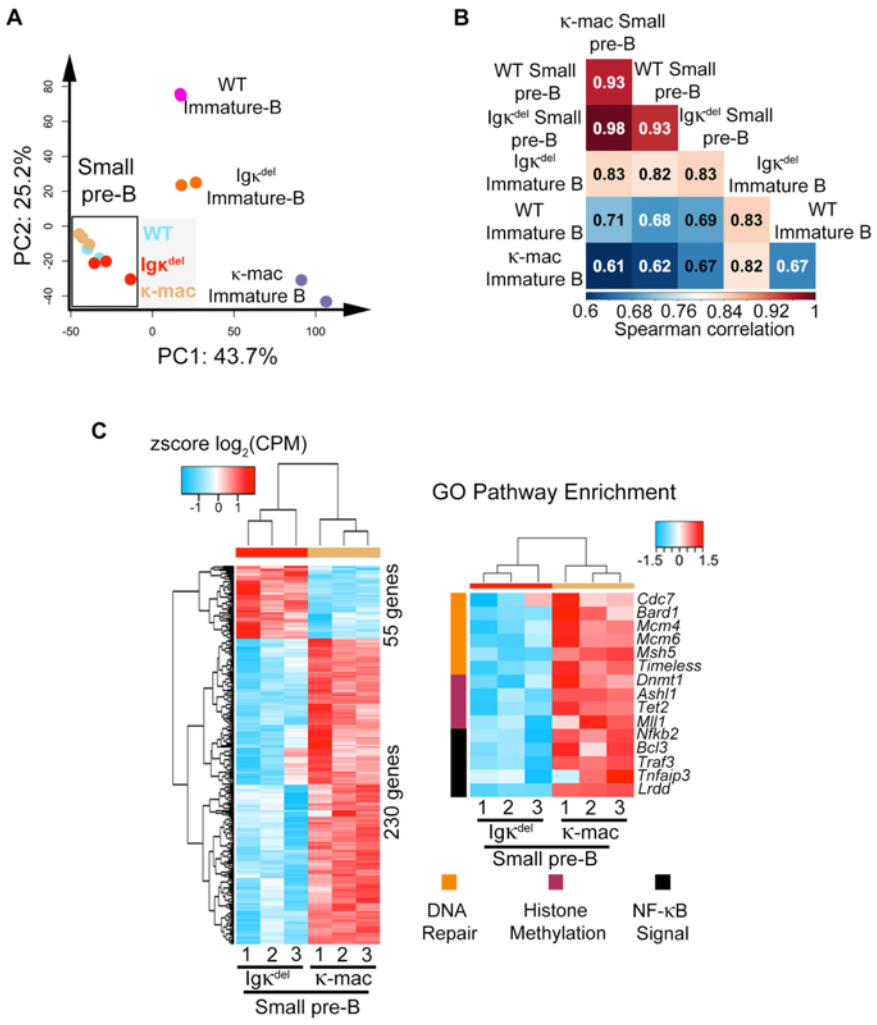


Figure 2.9 Small pre-B cells have unique transcriptional programs
A, Principal component analysis (PCA) of differentially expressed genes (DEGs) RNA-seq data from these cell populations: WT immature B cells (Magenta), Igκ^{del} immature B (Orange), κ-mac immature B (Purple), WT small pre-B (Light Blue), Igκ^{del} small pre-B (Red) and κ-mac small pre-B (Gold). Right panel shows PCA plot of only small pre-B populations. Replicates are shown. **B**, Spearman correlation plot of DEGs from cell populations. Scale indicates spearman correlation coefficients. **C**, Heatmap of DEGs between κ-mac (Gold) and Igκ^{del} (Red) small pre-B cells. Scale bar indicated z score of Log_2 transformed counts per million: z score log2(CPM). Right panel indicates a subset of DEGs upregulated in κ-mac small pre-B cells and representative gene pathways these genes belong to identified using Metascape program. Replicates are shown.

Consistent with their transcriptional similarity, pairwise comparisons between Igκ^{del} and κ-mac small pre-B cell transcriptional profiles revealed only 285 differentially expressed genes (Figure 2.9C). 230 genes were upregulated in κ-mac small pre-B cells and 55 genes

were upregulated in Igκ^{del} cells. However, Metascape gene annotation analysis revealed κ-mac upregulated genes were implicated in DNA repair, histone methylation and NF-κB signaling (Figure 2.9C; right panel) (Zhou et al., 2019). The enrichment of histone modification might explain the unique chromatin landscape seen in κ-mac small pre-B cells. Altogether, these data suggest that the small pre-B cell stage, κ-mac cells begin to establish a distinct transcriptional program predicted to be permissive for receptor editing.

2.11 κ-mac immature B show upregulation of metabolic pathways

Next, we performed differential expression analysis on Igκ^{del} and κ-mac immature B cells. 2,201 genes were significantly (FDR < 0.05) upregulated in Igκ^{del} immature B cells and 2,094 genes were significantly upregulated in κ-mac immature B cells (Figure 2.10A). Gene set enrichment analysis (GSEA) identified differentially enriched gene networks involved in chromatin organization and phosphoinositide 3-kinase (PI3K) signaling in Igκ^{del} immature B cells (Figure 2.10B)(Mootha et al., 2003; Subramanian et al., 2005). BCR tonic signaling primarily consists of PI3K signaling that represses DNA recombination (Clark et al., 2014; Monroe, 2004; Okkenhaug and Vanhaesebroeck, 2003). Among the chromatin organization genes, histone reader *Brwd1* was upregulated in Igκ^{del} immature B cells. We have shown *Brwd1* orchestrates *Igκ* recombination in small pre-B cells by controlling chromatin accessibility and RAG recruitment to *Jκ* genes (Mandal et al., 2015; Mandal et al., 2018). Relatively low expression of *Brwd1* in κ-mac immature cells is consistent with repression of *Igκ* recombination.

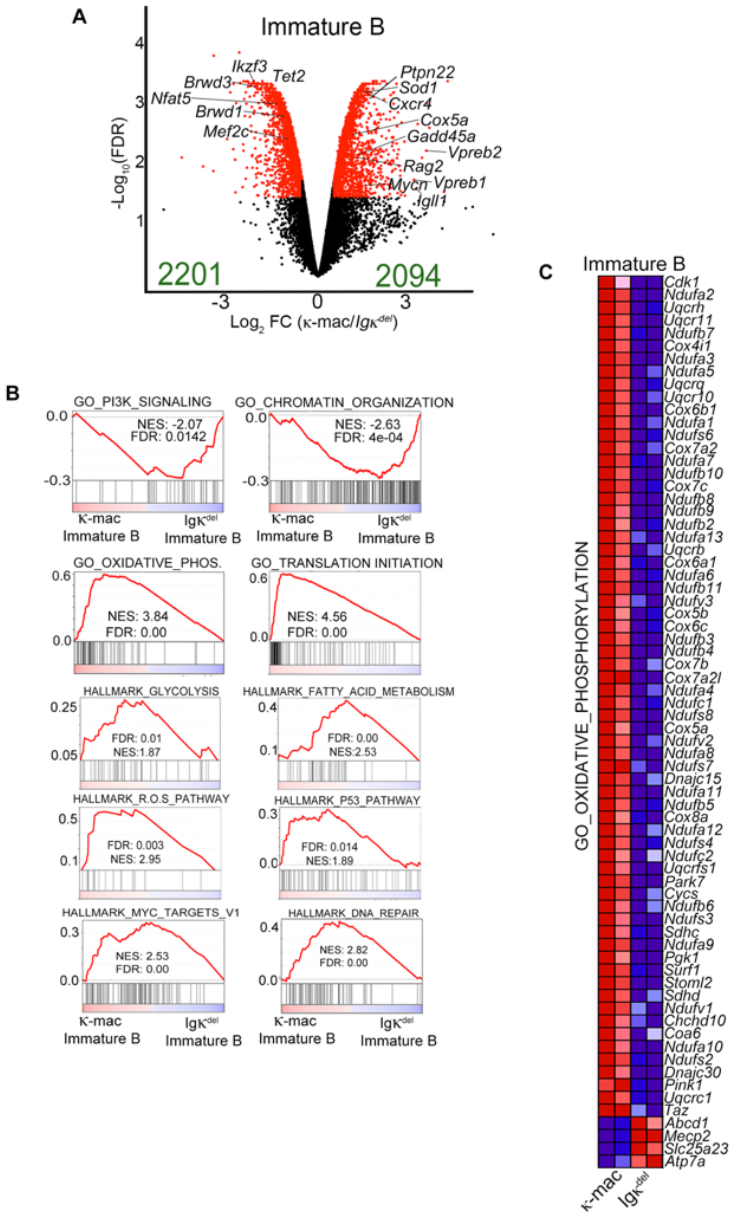


Figure 2.10 κ-mac immature B show upregulation of metabolic pathways

A, Volcano plot of DEGs expressed genes between Igκ^{del} and κ-mac immature B cells. Red dots are genes with significantly increased (right side) or decreased (left side) expression in κ-mac immature B cells. False discovery rate (FDR) was performed on P values using Benjamini-Hochberg correction. **B**, Gene set enrichment analysis (GSEA) for "GO_PHOSPHATIDYLINOSITOL_3_KINASE_SIGNALING", "GO_CHROMATIN_ORGANIZATION", "GO_OXIDATIVE_PHOSPHORYLATION", HALLMARK_DNA_REPAIR", "HALLMARK_GLYCOLYSIS", "HALLMARK_FATTY_ACID_METABOLISM", "HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY", "HALLMARK_P53_PATHWAY", "HALLMARK_MYC_TARGETS_V1" between κ-mac and Igκ^{del} immature B cells. NES: Normalized enrichment score". **C**, Heatmap of GSEA "GO_OXIDATIVE_PHOSPHORYLATION" gene set identified in B.

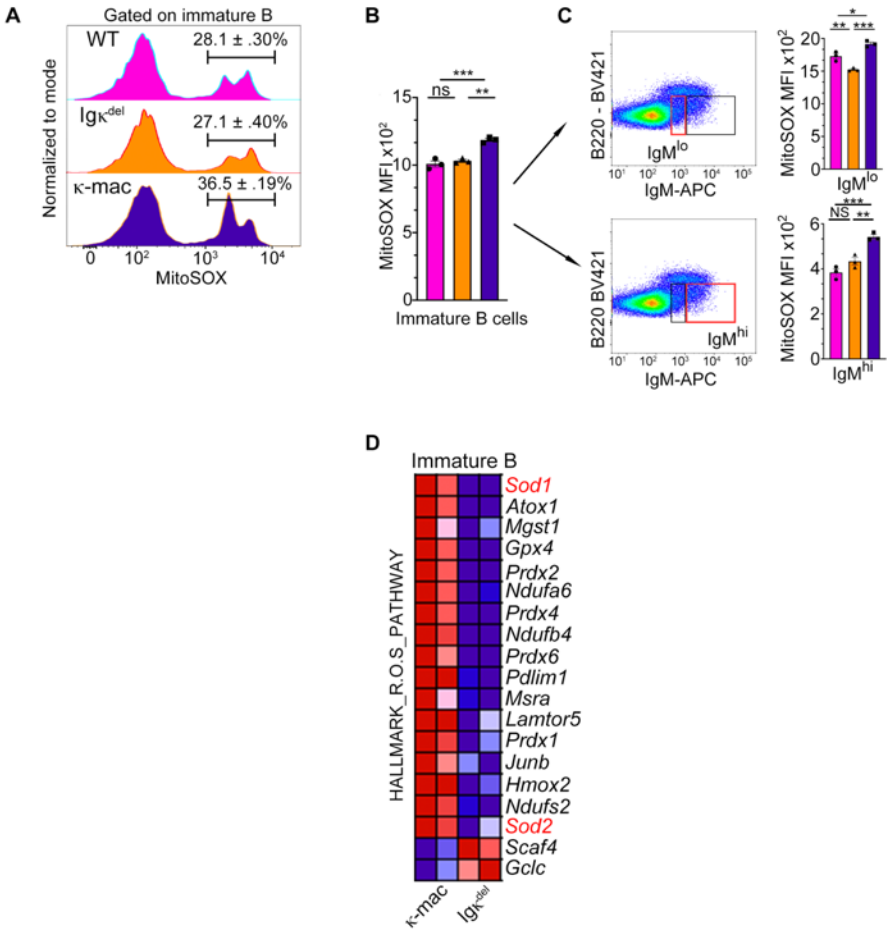
Similar to small pre-B cells, κ-mac immature B cells were enriched for DNA repair/recombination pathways. However, κ-mac immature B cells also upregulated *Rag2* and *Gadd45a. Gadd45a*, a gene induced by DNA damage, which has been shown to regulate *Rag1* and *Rag2* expression during receptor editing (Amin and Schlissel, 2008). Metabolic pathways such as oxidative phosphorylation (OxPhos), glycolysis and fatty acid oxidation, translation initiation, P53 signaling, DNA repair were also enriched in κ-mac immature B cells (Figure 2.10C-G). Although a clear picture of metabolism during B cell differentiation is still emerging, metabolic activity during early B cell development from pro to immature B cells, is characterized by a switch from glycolysis to OxPhos (Urbanczyk et al., 2018). Mature naïve B cells have been shown to rapidly upregulate OxPhos and glycolysis after BCR stimulation and differentiation to plasmablasts in a Myc-dependent manner (Akkaya et al., 2018; Caro-Maldonado et al., 2014; Doughty et al., 2006; Price et al., 2018). These findings suggest immature autoreactive B cells experience strong BCR signaling presumable for self-antigen.

2.12 κ-mac immature B are enriched for oxidative phosphorylation

To validate whether OxPhos was heightened in κ-mac immature B cells, we measured mitochondrial reactive oxygen species production (mitoROS) by flow cytometry. We used the MitoSOX™, which is rapidly oxidized by superoxide produced in mitochondria and becomes highly fluorescent. Consistent with our GSEA results, κ-mac immature B cells showed increased mitoROS production compared to WT and $Ig\kappa^{del}$ immature B cells (Figure 2.11A). Within the immature B cell compartment, cells with lower BCR expression (IgM^{lo}) exhibited higher mitoROS than cells with higher BCR expression (IgM^{hi}) (Figure 2.11B,C).

Moreover, IgM^{lo} κ-mac cells exhibited the highest overall mitoROS levels indicating that cells initiating receptor editing increase OxPhos.

Figure 2.11 κ-mac immature B are enriched for oxidative phosphorylation



A, Flow cytometric analysis of MitoSOX staining in immature B cells from indicated populations. Data presented as mean MitoSOX+ cells and ± 90% confidence interval. N=3 for all genotypes. **B**, MFI of MitoSOX staining in WT (Magenta), Igκ^{del} (Orange) and κ-mac (purple) immature B cells. N=3 for all genotypes. **C**, Representative flow cytometric analysis of immature B cells. IgMlo population is indicated by red box in upper-left panel and IgMhi population indicated in bottom left panel. MitoSOX MFI was quantified in for the IgMlo and IgMhi population in the upper and lower right panels respectively. WT (Magenta), Igκ^{del} (Orange) and κ-mac (purple) immature B cells. N=3 for all genotypes. Data are presented as means ± standard error of the mean (S.E.M). P values were determined by ANOVA followed by Tukey multiple comparisons (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). **D,** Heatmap of GSEA "HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY" gene set identified in A. Genes of interest are highlighted in red.

Increases in mitoROS can cause mitochondrial dysfunction and lead to cell death (Price et al., 2018). We thus searched the genes within the GSEA ROS pathway to identify antioxidant genes that may regulate the levels of mitoROS production. Indeed, superoxide dismutase 1 and 2 (*Sod1*, *Sod2*) were increased in κ-mac immature B cells compared to Igκ^{del}

immature B cells (Figure 2.11D). Altogether, these data show cells undergoing receptor editing upregulate metabolic processes, in particular; OxPhos and ROS production.

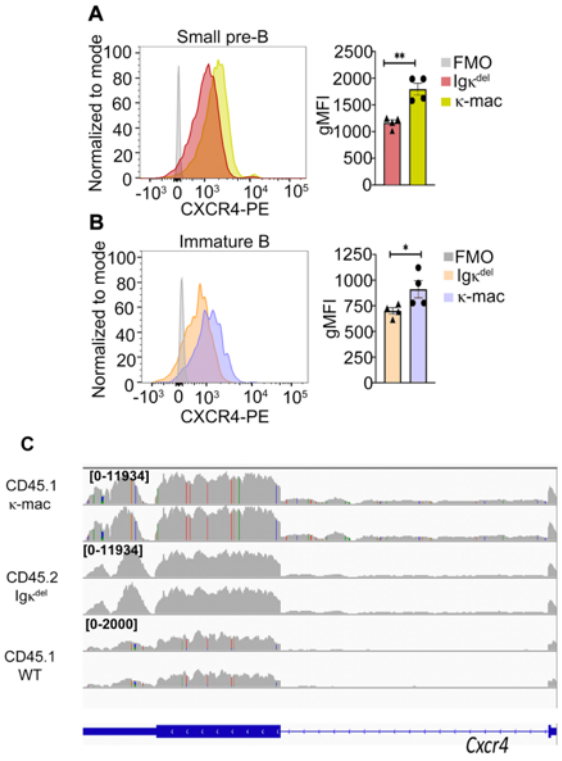


Figure 2.12 CXCR4 is upregulated on κ-mac small pre-B and immature B cells
A, Flow cytometric analysis of CXCR4 surface expression in Igκ^{del} and κ-mac small pre-B cells. Representative image shown to the left and quantification of MFI shown to the right. Height of histogram was normalized to mode. B, Flow cytometric analysis of CXCR4 surface expression in Igκ^{del} and κ-mac immature B. Representative image shown to the left and quantification of MFI shown to the right. Height of histogram was normalized to mode. **A-B**, Data are presented as means ± standard error of the mean (S.E.M). P values were determined by Student's T test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). **C**, IGV browser display of RNA-seq alignment files at *Cxcr4* gene with nucleotide changes from the mm9 genome shown. Colors indicate nucleotide changes: red, T; blue, C; green, A; and orange, G.

2.12 CXCR4 deficiency impairs Igλ+ B cell formation

ATAC-seq analysis highlighted chemokine signaling pathways were enriched in κ-mac cells. C-X-C motif chemokine receptor 4 (*Cxcr4*) a key regulator of B cell chemotaxis was upregulated genes in κ-mac immature B cells was. We further validated increased CXCR4 expression not only on immature B cells but small pre-B cells by flow cytometry (Figure 2.12A,B). Of note, the observed difference in CXCR4 expression was probably an underestimation as κ-mac on the B6-CD45.1 background possess genomic mutations that result in reduced expression of CXCR4 relative to B6-CD45.2 mice, which is on the Igκ^{del} background (Chisolm et al., 2019). RNA-seq analysis of the *Cxcr4* locus confirmed the presence of mutations in κ-mac but not Igκ^{del} mice which are associated with decreased

CXCR4 expression (Figure 2.12C). Yet, despite this genetic predisposition, κ-mac small pre-B and immature B cells highly expressed CXCR4 suggesting an important role in receptor editing.

We recently showed that CXCR4 not only mediates small pre-B chemotaxis but the CXCR4/CXCl12 signaling axis regulates small pre-B cell differentiation and *Igκ* recombination (Mandal et al., 2019). To determine if CXCR4 also has a role in receptor editing, we generated κ-mac *Cxcr4$^{fl/fl}$mb1-cre^{+}* (hereafter κ-mac *Cxcr4KO)* and κ-mac *Cxcr4$^{fl/fl}$* mice. *Cxcr4$^{fl/fl}$ mb1-Cre^{+}* mice have a developmental block at the small-pre B cell stage characterized by an epigenetic landscape reminiscent of large pre-B cells and reduced *Igκ* recombination (Mandal et al., 2019). Consistent with our previous results, pro-B and large pre-B populations were largely unaffected by loss of CXCR4 expression (Figure 2.13A,B). The κ-mac *Cxcr4KO* mice had a similar developmental block at the small pre-B cell stage. However, small pre-B cells were 3-fold reduced and 8-fold and 36-fold reduced in immature and recirculating mature B cells respectively. Cell surface staining for IgM, Igκ and Igλ demonstrated a complete lack of Igκ+ B cells, consistent with these mice expressing the κ-macroself transgene (Figure 2.13C). However, Igλ+ B cells were mostly absent in κ-mac CXCR4KO mice, suggesting CXCR4 is required for Igλ+ B cell formation.

2.14 Igκ recombination is normal in κ-mac *CXCR4KO* mice

Previous work demonstrates CXCR4 is necessary of *Igκ* accessibility. A possible explanation for a lack of Igλ+ B cells in the absence of CXCR4 could be explained by an inability to initiate Igκ recombination. To test whether Igκ recombination was indeed impaired in the κ-mac *Cxcr4KO* mice, we measure intracellular Igκ expression using flow

cytometry. Interestingly, in the context of the κ-macroself Igκ recombination was not significantly impaired as seen by similar frequency of intracellular Igκ in the presence or absence of CXCR4 (Figure 2.14A). As expected, cells expressing surface IgM were negative for intracellular Igκ but positive for intracellular Igλ. Since the antibodies against IgM, Igκ and Igλ recognize assembled immunoglobulins (Li et al., 2002; Yelton et al., 1981), we conclude that the proportion of small pre-B cells that complete primary kappa recombination in the absence of CXCR4 express the completed BCR on the cell surface and transition into the immature B cell stage. However, due to antigen ligation by the κ-macroself, the Igκ+ receptor undergoes rapid down modulation causing the autoreactive cell to return to the small pre-B cell stage where the lack of CXCR4 signaling prevents usage of the Igλ light chain.

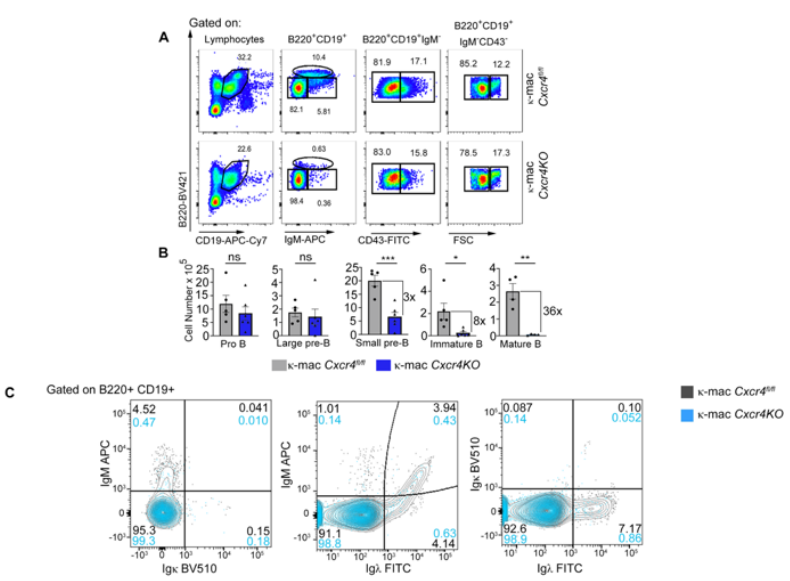


Figure 2.13 CXCR4 deficiency impairs Igλ+ B cell formation

A, Representative flow cytometric analysis of different developmental stages of B lymphopoiesis in the BM of κ-mac *Cxcr4*$^{fl/fl}$ *mb1-Cre-* (κ-mac *Cxcr4*$^{fl/fl}$) and κ-mac *Cxcr4*$^{fl/fl}$ *mb1-cre+* (κ-mac *Cxcr4KO*) mice. B cell progenitors subpopulations are defined as follows: Progenitor gate (B220+CD19+), pro-B cells (B220+CD19+CD43+IgM−), large pre-B cells (B220loCD19+CD43−IgM−FSChi) and small pre-B cells (B220loCD19+CD43−IgM−FSClo) and immature B cells (B220loCD19+CD43−IgM+) and mature B cell (B220hiCD19+CD43−IgM+)). FSC; Forward Scatter. **B**, Absolute cell numbers for populations shown in A. Data are presented as means ± standard error of the mean (S.E.M). P values were determined by Student's T test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). **C**, Representative flow of IgM, Igκ and Igλ expression in κ-mac *Cxcr4*$^{fl/fl}$ *mb1-Cre-* (κ-mac *Cxcr4*$^{fl/fl}$) and κ-mac *Cxcr4*$^{fl/fl}$ *mb1-cre+* (κ-mac *Cxcr4KO*) mice.

2.15 CXCR4 regulates transcriptional programs involved in receptor editing

RNA-seq of FACS isolated small pre-B cells revealed 869 (FDR < 0.05) differentially expressed genes between κ-mac *Cxcr4*$^{fl/fl}$ and κ-mac *Cxcr4KO* mice (Figure 2.15A). Loss of CXCR4 expression was associated with 230 upregulated and 639 repressed genes. Metascape gene annotation analysis revealed CXCR4 was required for endocytosis, regulation of cell activation, cell chemotaxis, MAPK signaling and NF-κB signaling gene programs (Figure 2.15B). As expected, the cell chemotaxis pathway included mediators of cell mobility such as *CXCR4, Sipra* and cell adhesion molecules *Itgam and Itgb2.*

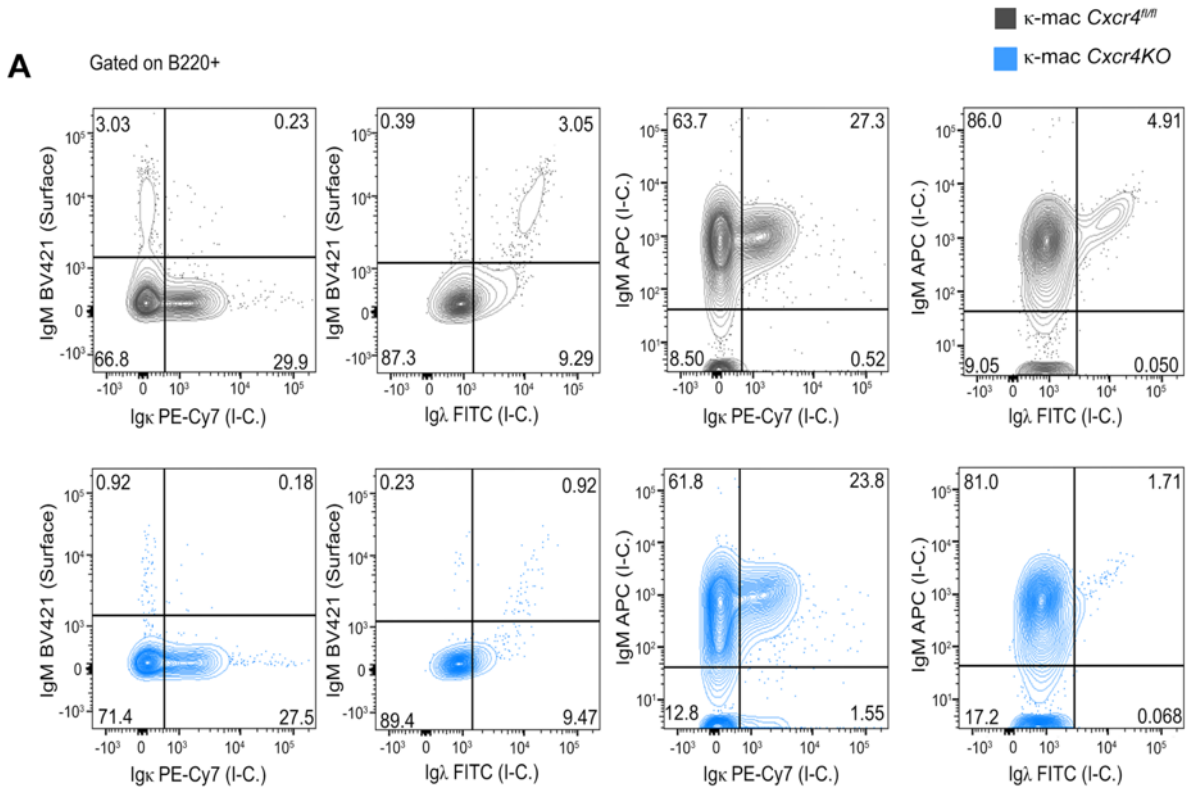


Figure 2.14 Igκ recombination is normal in κ-mac *CXCR4KO* mice
A, Representative flow cytometry of surface IgM and intracellular IgM, Igκ and Igλ expression from κ-mac *CXCR4*$^{fl/fl}$ vs κ-mac *CXCR4KO*. Cell are gated on B220+ cells.

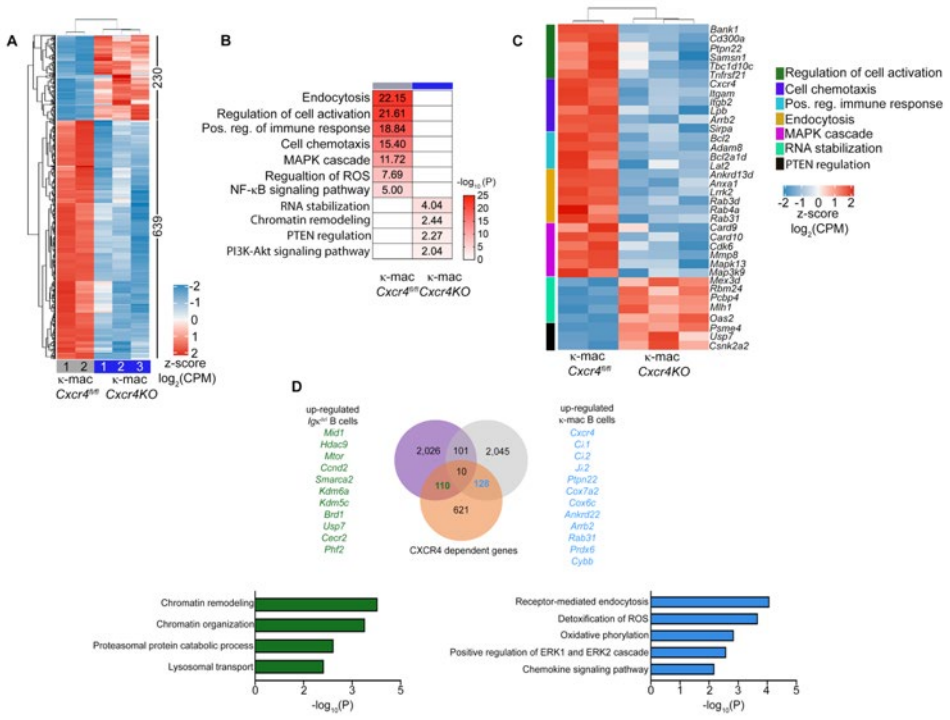


Figure 2.15 CXCR4 regulates transcriptional programs involved in receptor editing
A, Heatmap of DEGs between κ-mac *Cxcr4*$^{fl/fl}$ (n=2) and κ-mac *Cxcr4KO* (n=3) small pre-B cells. Scale indicates z-score of Log_2 transformed normalized counts: z score log_2(CPM). Replicates are shown. **B**, Metascape gene pathway analysis of differentially expressed genes identified in A. Scale bar indicates Log_{10} transformed P values determined by Metascape software. **C**, Heatmap of representative genes in the "Regulation of cell activation" and "Endocytosis", "Mapk cascade", "Nonsense mediated decay", and "Chromatin remodeling." Scale indicates z-score of Log_2 transformed normalized counts: z score log_2(CPM). Replicates are shown. **D**, Venn diagram of overlaps between genes upregulated in the indicated population. Metascape analysis was performed on genes shared by Igκ^{del} and CXCR4 dependent genes and κ-mac and CXCR4 dependent genes.

Enrichment for regulation of cell activation was particularly interesting. 'Tonic' signals from an non-ligated BCR are critical for selection of immature B cells into the periphery (Bannish et al., 2001; Tze et al., 2005). The non-ligated BCR promotes PI3K signaling in immature B cells, which repress RAG genes and halt light chain gene rearrangements (Amin and Schlissel, 2008). Conversely, BCR signaling of autoreactive B cells after antigen ligation must be attenuated to initiate receptor editing and other tolerance

mechanisms (Nemazee, 2017). Thus, regulatory processes involved in modulating BCR signaling are crucial for central tolerance. Genes in the regulation of cell activation pathway such as *Bank1, Tbc1d10c, Tnfrsf21,* Ptpn22 have been demonstrated to inhibit B cell activation (Georg et al., 2019,Stanford, 2014 #363; Schickel et al., 2012; Schmidt et al., 2003) (Figure 2.15C). Of these genes, *Ptpn22* (Protein tyrosine phosphatase, non-receptor type 22) is the best characterized, as *PTPN22* polymorphism in humans is a risk factor for many autoimmune diseases (Stanford and Bottini, 2014). Consistent with this, mice carrying a *Ptpn22* (R619W) mutation have a B cell autonomous defect that promotes autoimmunity and was associated with protection of immature B cells from negative selection (Dai et al., 2013). Altogether, this suggests CXCR4 may be necessary for inhibiting signals from an autoreactive BCR that would prevent tolerance mechanisms.

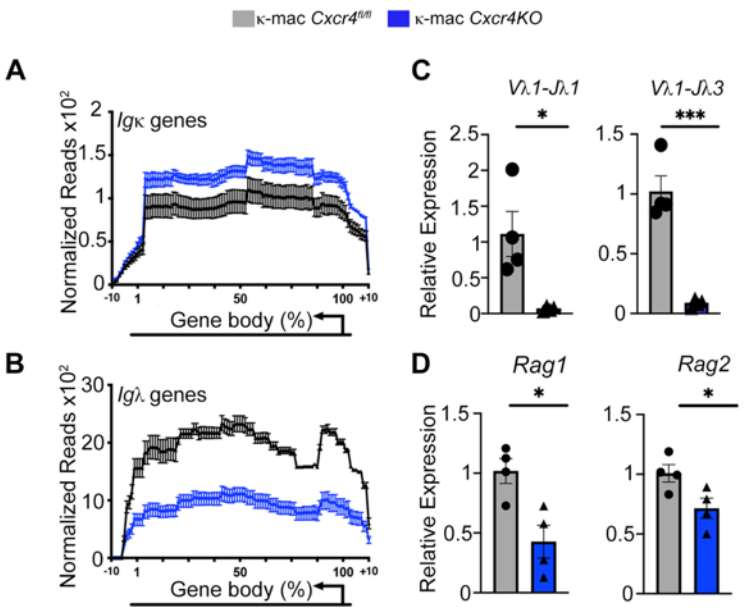


Figure 2.16 CXCR4 expression is necessary for Igλ recombination.
A-B, Quantified and integrated transcription across all Igκ (A) and Igλ (B) gene segments from FACS sorted κ-mac *CXCR4$^{fl/fl}$* (n=2) vs κ-mac *CXCR4KO* (n=3) small pre-B cells. **C-D**, Complementary DNA was synthesized from RNA obtained from sorted κ-mac *CXCR4$^{fl/fl}$* (n=4) vs κ-mac *CXCR4KO* (n=4) small pre-B cells and used to perform qPCR for *Vλ1-Jλ1* and *Vλ1-Jλ3* (C) recombination products and *Rag1* and *Rag2* (D).

Conversely, loss of CXCR4 resulted in enrichment for RNA stabilization, chromatin remodeling, phosphatase and tensin homolog (PTEN) regulation and PI3K-Akt signaling. Genes *USP7 and Csnk2a2* found in the regulation of PTEN pathway were involved in inhibiting PTEN activity through deubiquitylation or phosphorylation respectively (Song et al., 2012). Thus, loss of CXCR4 activity is associated with inhibiting PTEN activity leading to increased PI3K signaling and potentially decreased RAG activity.

Next, we wondered how many of the genes differentially regulated in the Igκ^{del} and κ-mac cells were regulated by CXCR4. Thus, we combined genes that were differentially expressed from the pairwise comparisons of small pre-B Igκ^{del} and κ-mac cells and immature B Igκ^{del} and κ-mac cells. Genes that were upregulated in Igκ^{del} cells represented genes not associated with receptor editing while genes associated with editing were represented by genes upregulated in κ-mac cells. Overlaps of these gene sets with CXCR4 dependent genes revealed a subset of editing and non-editing genes dependent on CXCR4 expression (Figure 2.15D). CXCR4 dependent editing genes included *Cxcr4, Cλ1, Cλ2, Jλ2*, *Cox7a2* and *Cox6c* and phosphatase *Ptpn22*. These genes were associated with pathways necessary for cells undergoing receptor editing such as receptor mediated endocytosis, ROS/OxPhos, and MAPK and chemokine signaling pathways. Conversely, gene pathway analysis of non-editing CXCR4 dependent genes showed processes that would normally occur during the small pre-B to immature B cell transition such as upregulation of chromatin remodeling/organization and protein degradation pathways (Figure 2.15D). Thus, CXCR4 regulates transcriptional programs associated with receptor editing.

2.16 CXCR4 expression is necessary for Igλ recombination.

Since κ-mac mice must express a λ-chain-positive BCR to escape deletion, this gave us the opportunity to assess whether CXCR4 regulates *Igλ* recombination. We examined *Igκ* transcription in κ-mac *Cxcr4KO* and κ-mac *Cxcr4*$^{fl/fl}$ small pre-B cells. Similar to how chromatin accessibility was quantified across the light chain genes, we quantified transcription at *Igκ* and *Igλ* loci across all encoding genes. In agreement with the intracellular flow cytometry data, CXCR4 deficiency did not impair transcription across Igκ genes (Figure 2.16A). In fact, transcription across the gene body was slightly elevated in the absence of CXCR4. However, CXCR4 deficiency resulted in a substantial decrease in transcription along Igλ gene elements and importantly at regions flanking the gene bodies (Figure 2.16B). This suggests a defect in *Igλ* transcription and subsequent recombination.

To validate recombination was impaired in the absence of CXCR4, we measured *Vλ-Jλ* recombination products for λ1 and λ3 using quantitative PCR (qPCR) on cDNA from κ-mac *Cxcr4KO* and κ-mac *Cxcr4*$^{fl/fl}$ small pre-B cells. Recombination of both *Vλ1-Jλ1* and *Vλ2-Jλ3* was severely diminished in κ-mac *Cxcr4KO* sorted small pre-B cells (Figure 2.16C). Moreover, *Rag1* and *Rag2* expression was also diminished in the absence of CXCR4 (Figure 2.16D). Thus, CXCR4 is required not only for preventing BM egress but, in the context of receptor editing, both repression of cellular activation pathways and for re-induction of critical recombination programs including.

2.17 Epigenetic landscape during receptor editing and chromatin accessibility at *Igλ* depends on CXCR4 expression

As demonstrated, κ-mac small pre-B cells possess a unique epigenetic landscape. Thus, we investigated whether CXCR4 regulated the chromatin landscape of editing small pre-B. In the

absence of CXCR4, there were approximately 2,300 regions where chromatin accessibility was altered in κ-mac small pre-B cells as assessed by ATAC-seq (Figure 2.17A). Accessibility at ~1,100 of these regions were significantly decreased in κ-mac *Cxcr4KO,* while ~1,200 regions showed increased accessibility. Interestingly, of the 1,100 regions that lost accessibility due to CXCR4 deficiency, roughly 70% were located at intergenic or intronic regions and the remaining 30% were located at promoter or exonic regions (Figure 2.17B). In contrast, regions that gained accessibility in κ-mac *Cxcr4KO* were almost exclusively found at intergenic or intronic regions (90%) with the remaining regions distributed between promoter or exonic regions. Pathway analysis of genes near differential OCRs identified CXCR4 was necessary for signaling pathways related to metabolism, NF-κB, MAPK signaling among other pathways (Figure 2.17C). Many of these pathways were highly enriched in κ-mac small pre-B cells when compared to WT or $Ig\kappa^{del}$ small pre-B cells. Genes in the vicinity of chromatin sites that gained accessibility in the absence of CXCR4 were involved in P53

signaling and oxidative stress. These findings suggest CXCR4 preferentially opens up regions for signaling pathways involved in receptor editing.

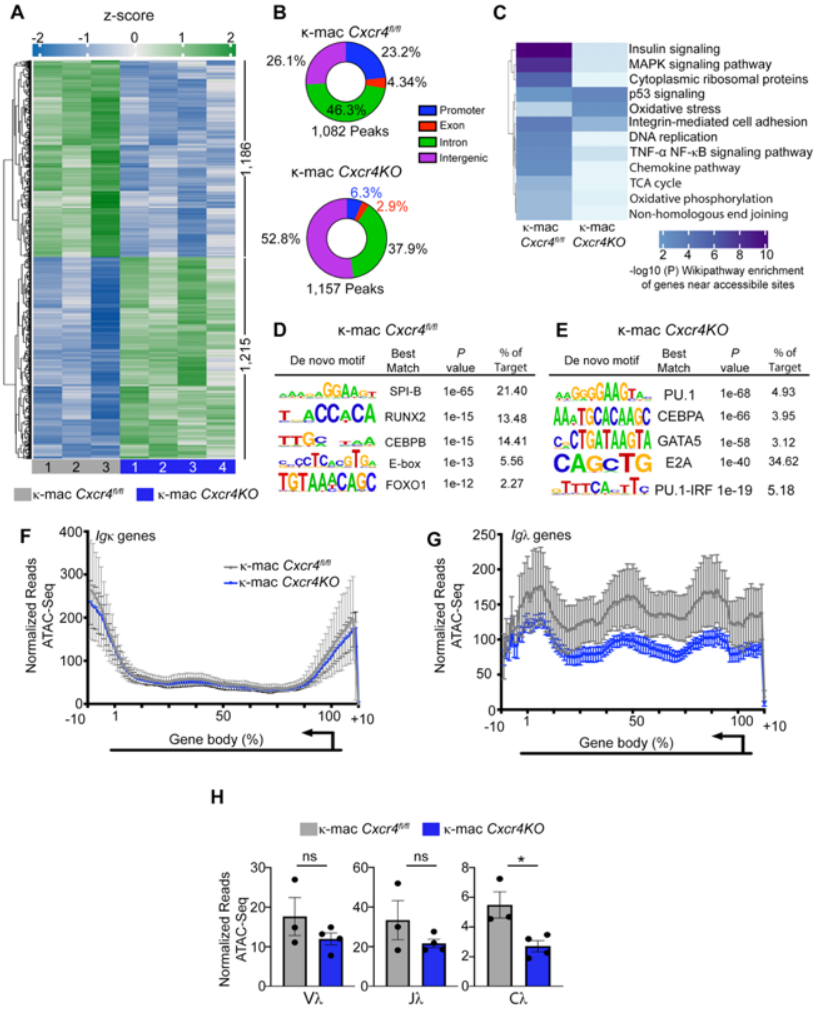


Figure 2.17 Epigenetic landscape during receptor editing and chromatin accessibility at *Igλ* depends on CXCR4 expression

A, Heatmap depicting differentially accessible regions (DARs) determined by ATAC-Seq of FACS sorted κ-mac *CXCR4*$^{fl/fl}$ (n=3) vs κ-mac *CXCR4KO* (n=4) small pre-B cells. The row z score of Log_2 transformed normalized counts were used to construct the heatmap. Biological replicates are indicated. **B**, Frequency of DARs depicted in **A** at annotated genomic regions: Promoter (blue), Exon (red), Intron (green) and Intergenic (purple). **C,** Gene ontology analysis of genes near DAR depicted in A. 'Wikipathway' gene sets were used for this analysis. Scale bar represents Log_{10} transformed P values obtained from this analysis. **D-E**, DARs were analyzed for transcription factor motif enrichment using HOMER's 'findmotifsgenome.pl'. De Novo motifs are depicted with significance determined using hypergeometric distribution. **F-G,** Quantified and integrated accessibility determined by ATAC-seq immediately before and after the length of each Igκ (F) or Igλ (G) gene body from κ-mac *CXCR4*$^{fl/fl}$ (n=3) vs κ-mac *CXCR4KO* (n=4) small pre-B cells. **H**, Histogram depicting ATAC-Seq defined accessibility at Vλ, Jλ and Cλ genes separately.

Analysis of binding motifs present in the differentially accessible chromatin regions showed enrichment for similar TFs (Figure 2.17D,E). SPI-B, PU.1 and E-box binding motifs were present at regions opened and closed by CXCR4, all of which are TFs important light chain recombination. However, CXCR4 specifically opened regions containing FOXO1 binding motifs, a TF known to directly regulate RAG induction. Thus, CXCR4 modulates regions where key TF motifs are found.

Next, we assessed accessibility at the light chain loci. In agreement with transcriptional analysis at *Igκ,* accessibility was nearly identical in the presence or absence of CXCR4 (Figure 2.17F). Chromatin accessibility was highest at the flanking 5' and 3' regions of Igκ genes indicating the locus is open and capable of undergoing recombination. However, unlike the transcriptional analysis, *Igλ* accessibility was only marginally affected in the absence of CXCR4 (Figure 2.17G). Assessing accessibility separately at Vλ, Jλ and Cλ segments revealed that loss of accessibility was most significantly observed at Cλ genes in the absence of CXCR4 (Figure 2.17H). Taken together, CXCR4 is necessary for establishing aspects of the epigenetic landscape during receptor editing and complete opening of the Igλ locus.

2.18 CXCR4 deficient small pre-B cells are found in the periphery

Loss of CXCR4 expression might promote premature egress from the BM. Indeed, splenocytes from κ-mac *Cxcr4KO* mice had an accumulation of B220+CD19+CD93+CD23-IgM- (hereafter CD23-IgM-) cells (Figure 2.18A). Transitional cells were slightly elevated in κ-mac *Cxcr4KO* mice, but mature B cell numbers were comparable to κ-mac *Cxcr4*$^{fl/fl}$ mice (Figure 2.18B-G). As these cells express CD93, they may be IgM^{-} transitional B cells.

Therefore, we assessed IgD expression as it is acquired as Transitional 1 (T1) cells differentiate into Transitional 2 (T2) cells (Boothby et al., 2019; Chung et al., 2003). Surprisingly, we observed reduced IgD expression on CD23-IgM- cells from κ-mac *Cxcr4KO* mice compared to CD23-IgM- κ-mac *Cxcr4*$^{fl/fl}$ mice (Figure 2.18H). These data suggest the CD23-IgM- cells may arise from either premature BM emigrates or defective T1 to T2 transition. Given the normal frequency and number of mature B cells in κ-mac *Cxcr4KO* mice compared to floxed control, we favor the hypothesis that the CD23-IgM- are primarily derived from the BM.

To further explore the progeny of the CD93+CD23-IgM- population in κ-mac *Cxcr4KO*, we FACS isolated small pre-B, immature B and CD93+CD23-IgM- cells and performed RNA-Seq. PCA showed these population to be transcriptionally distinct (Figure 2.19I). We observed 1,182 (FDR < 0.05) differentially expressed genes between the cell populations. Upon further analysis, K-means clustering identified 4 gene expression patterns (Figure 2.18I,J; Clusters1-4). Cluster 1 was the smallest cluster (85 genes) and were unique for κ-mac *Cxcr4KO* immature B cells. Clusters 2 and 4 were comparable in size (382 and 378 genes respectively) and these showed unique expression for the CD23-IgM- and small pre-B populations respectively. Cluster 3 contained 337 genes that shared upregulation in small

pre-B and CD23-IgM- cells. Thus, CD23-IgM- in κ-mac *Cxcr4KO* mice are likely derived from premature egress of small pre- B cells from the BM.

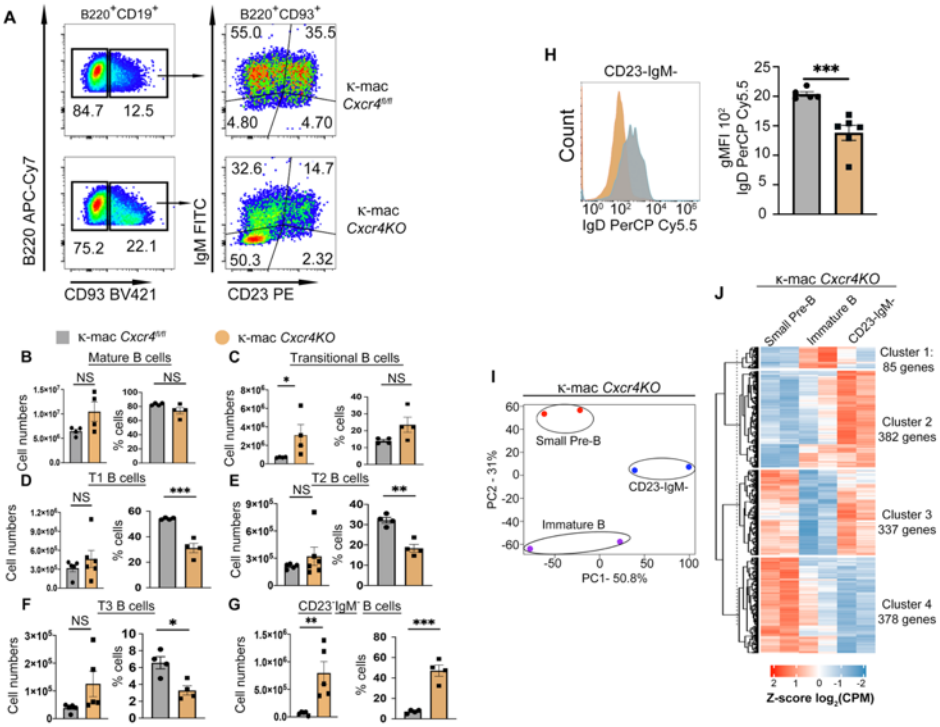


Figure 2.18 CXCR4 deficient small pre-B cells are found in the periphery

A, Flow cytometric results of splenic B cell populations from κ-mac *Cxcr4*$^{fl/fl}$ *mb1-cre-* (κ-mac *Cxcr4*$^{fl/fl}$) and κ-mac *Cxcr4*$^{fl/fl}$ *mb1-cre+*(κ-mac *Cxcr4KO*). Left panel indicates gates for mature B cells (B220+CD19+CD93-) and Transitional B cells (B220+CD19+CD93+). The right panel indicates gates for subpopulations within Transitional B cells. These population are as follows: Transitional 1 (T1) B220+CD19+CD93+IgM+CD23-, Transitional 2 (T2) B220+CD19+CD93+IgM+CD23+, Transitional 3 (T3) B220+CD19+CD93+IgMloCD23+ and a fourth population B220+CD19+CD93+IgM-CD23-. **B**, Absolute cell number and frequency of mature B cells in κ-mac *Cxcr4*$^{fl/fl}$ (Gray) κ-mac *Cxcr4KO* (Gold). **C**, Absolute cell number and frequency of Transitional B cells in κ-mac *Cxcr4*$^{fl/fl}$ (Gray) κ-mac *Cxcr4KO* (Gold). **D**, Absolute cell number and frequency of Transitional 1 B cells in κ-mac *Cxcr4*$^{fl/fl}$ (Gray) κ-mac *Cxcr4KO* (Gold). **E**, Absolute cell number and frequency of Transitional 2 B cells in κ-mac *Cxcr4*$^{fl/fl}$ (Gray) κ-mac *Cxcr4KO* (Gold). **F** Absolute cell number and frequency of Transitional 3 B cells in κ-mac *Cxcr4*$^{fl/fl}$ (Gray) κ-mac *Cxcr4KO* (Gold). **G**, Absolute cell number and frequency of CD23-IgM- cells in κ-mac *Cxcr4*$^{fl/fl}$ (Gray) κ-mac *Cxcr4KO* (Gold). **B-G**, Littermate control were used for these experiments with n=4 for each group. **H,** Flow cytometric analysis of IgD surface expression in Igκ^{del} and κ-mac small pre-B cells. Representative image shown to the left and quantification of MFI shown to the right. Height of histogram was normalized to mode. **B-H**, Data are presented as means ± standard error of the mean (S.E.M). P values were determined by Student's T test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). **I**, Principal component analysis of RNA-seq data from indicated cell populations. **J**, Heat map of differentially expressed genes between the indicated cell populations. Clusters were determined using consensus k-means clustering with 500 iterations. Numbers of genes in each cluster are indicated and replicates are shown. Scale bar indicated z-score of Log_2 transformed counts per million: z score log2(CPM).

3 Results: Addendum

Many other aspects of Igλ regulation was pursued over the course of this study but were tangential to the primary results discussed earlier. We were first interested in understanding the 3-D chromatin architecture of the Igλ locus and how this may influence the resulting Igλ repertoire observed in mice. Next, we sought to understand the regulatory importance of the two Igλ enhancers 2-4 and 3-1 and their role in locus accessibility, recombination, and repertoire formation. Lastly, analysis implicating CXCR4 in promoting DNA recombination is presented.

These works presented as vignettes represent interesting novel findings that will improve our understanding of Igλ regulation and CXCR4 signaling should others pursue follow up experiments.

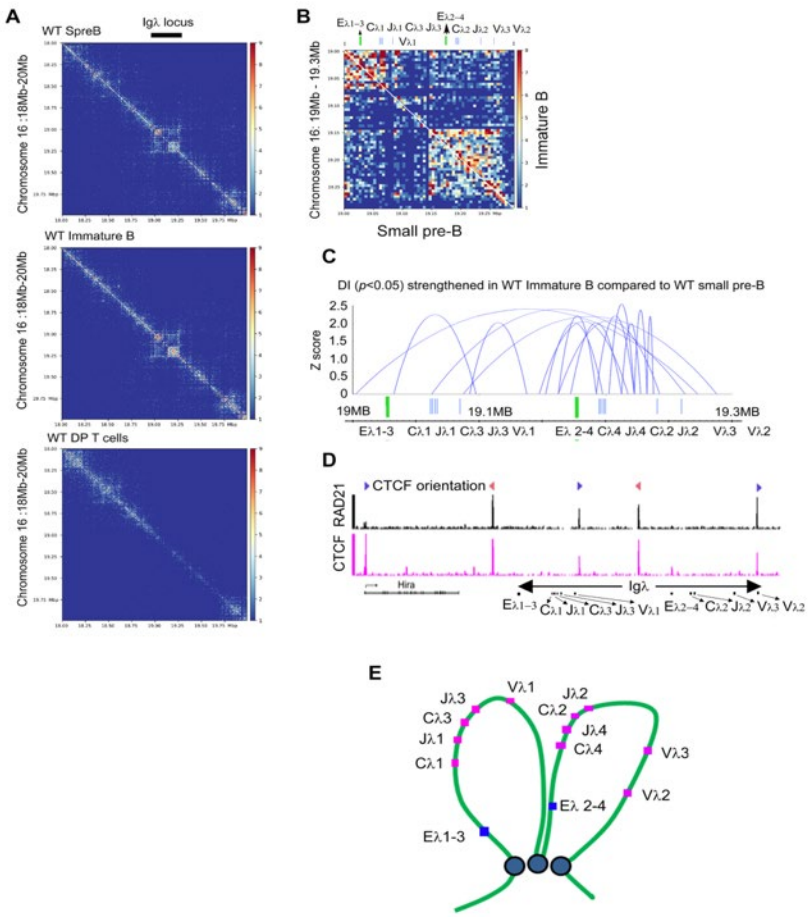
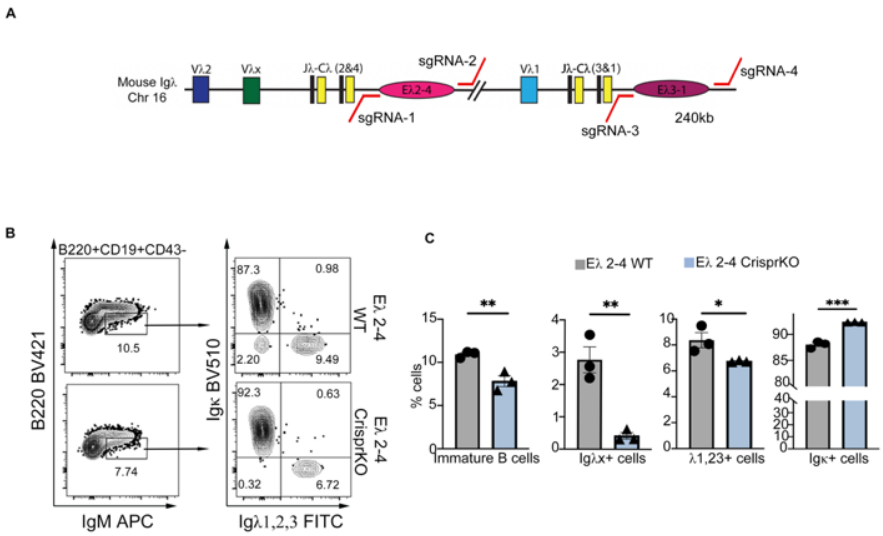


Figure 3.1 HI-C shows Igλ locus is organized into two topologically associated domains
A, Schematic of Igλ locus with gRNAs targeting the Eλ2-4 and Eλ3-1 highlighted in red. **B**, Flow cytometric analysis of BM cells from Eλ2-4 control and knockout (Eλ 2-4 CrisprKO) mice. **C**, Frequency of indicated cells populations in Eλ2-4 control and knockout (Eλ 2-4 CrisprKO) mice. Igλx cells are found in the Igκ- Igλ- gate. Data are presented as means ± standard error of the mean (S.E.M). P values were determined by Student's T test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

3.1 Igλ locus is organized into two topologically associated domains

To understand the chromatin structure of the Igλ locus, Malay Mandal and Aazam Moshin performed in situ Hi-C on sorted WT small pre-B and immature B cells (Rao et al., 2014; Yu and Ren, 2017). In situ Hi-C contact maps of the Igλ locus showed the increased chromosomal contact in immature B cells compared to small pre-B (Figure 3.1A-B). As a control WT double positive T cells showed no chromosomal interactions at the Igλ locus. We assessed the significance of the chromosomal contacts strengthened in WT immature B cells compared to small pre-B cell and observed many statistically significant ($P<0.05$) DNA contacts clustered around Eλ2-4 in immature-B cells.

Figure 3.2 Eλ2-4 act locally to orchestrate recombination.



A, Normalized Hi-C contact maps for Chr16:18Mb-20Mb of WT small pre-B, WT Immature B and WT double positive (DP) T cells indicating the Igλ locus. Bars depict signal intensities. **B**, In situ Hi-C contact matrices of the Igλ locus showed the presence of two topologically associated domains (TADs) separating two functional enhancers of Igλ locus, Eλ2-4 and Eλ1-3 into the two TADs. **C**, Comparison of interaction matrix identifying differential DNA interactions between small pre-B cells and Immature B cells with statistically significant (P<0.05) DNA contacts between small pre-B and immature-B cells. Immature B cells showed increased interactions in the TAD containing Eλ2-4. **D,** CTCF and cohesion protein (RAD21) have been associated with TAD boundary formation of Igλ locus separating enhancers Eλ2-4 and Eλ1-3 into the two TADs. **E**, Schematic of loop formation interpreted from CTCF and RAD21 ChIP-seq showed in D.

TADs are chromosomal domains wherein DNA sequences physically interact at higher frequency than with neighboring chromosomal domains (Beagan and Phillips-Cremins, 2020). CTCF and cohesion proteins have been associated with TAD boundary formation with TADs flanked by convergent CTCF binding sites forming the most stable boundaries (Dixon et al., 2012; Nora et al., 2012; Yu and Ren, 2017). Indeed, ChIP-seq of CTCF and RAD21, a cohesion subunit (Degner et al., 2011; Degner et al., 2009), in sorted WT pro-B cells showed binding of both CTCF and RAD21 across the Igλ locus (GSE40173 and GSE47766) (Figure 3.1B). A schematic representation of predicted loop formation the Igλ locus is composed of two CTCF-mediated loops (Figure 3.1C). Interestingly, these loops each contained an enhancer (Eλ3-1 or Eλ2-4), Jλ and Cλ genes, and either one or two Vλ genes.

Given the high sequence homology of the enhancers and their location in different TADs this raised the possibility that these enhancers orchestrate Vλ-Jλ recombination locally within their respective TADs.

To test this hypothesis, we used CRISPR to delete the Eλ3-1 and Eλ2-4 from WT C57BL/6J mice. We designed 4 gRNA guides, a pair for each enhancer (3.2.A). We were successful in recovering a homozygous deletion of the Eλ2-4. Flow cytometric analysis of BM population revealed a slight reduction of immature B cells in Eλ2-4 CRISPRKO mice compared to control mice derived from embryos inject with Eλ2-4 gRNAs but were negative for deletion (Figure 3.2B). Staining for Igκ and Igλ revealed an increase in the frequency Igκ+ B cells in Eλ2-4 CRISPRKO but a decrease in Igλ+ B cells. Interestingly, we observed a complete loss of the Igκ- Igλ- cells, presumed to be the Igλx product. Thus, Eλ2-4 and presumably Eλ1-3 act locally to orchestrate recombination.

3.2 CXCR4 signaling represses homologous recombination

Immunoglobulin gene recombination is dependent upon non-homologous end-joining (NHEJ) DNA repair and not homologous DNA recombination (HR) (Brandsma and Gent, 2012; Malu et al., 2012).Therefore, we examined if components of DNA repair were transcriptionally regulated by CXCR4. Overall, CXCR4 repressed DNA repair pathways (Figure 3.3A). However, examination of specific mediators of NHEJ by real time RT-PCR, including *Samrca2*, *Smarca4*, *Dclre1c* (encodes Artemis), *Xrcc4*, *Xrcc6* (Ku70), *Xrcc5* (Ku80), *Prkdc* (DNA-PKC) and *Lig4* revealed that they were not repressed by CXCR4 (Figure 3.3B). Rather, *in vitr*o analysis determined that CXCL12, and not IL-7 withdrawal, strongly and broadly repressed almost all the mediators of homologous recombination (HR) (Figure 3.3C,

left panel) including signature HR genes *Brca1, Brca2, Aurka, Aurkb, Ezh2, Rad21, Rad51, Check2, E2fa and Foxm1* (Figure 3.3D**,** left panel). The effect of CXCR4 *in vitro* is seen *in vivo* (Figure 3.3E**,** right panels). Transition from large pre-B cell to small pre-B cell stage is associated with a marked repression of HR gene expression which is compromised in *Cxcr4*$^{-/-}$ small pre-B cells.

To determine if the observed repression of HR genes was functionally significant, we used R*osa26* Direct Repeat-GFP (RaDR-GFP) mice (Sukup-Jackson et al., 2014), in which a repeat HR substrate is inserted in the ubiquitously expressed *Rosa26* locus. HR between two truncated *GFP* expression cassettes gives rise to fluorescence following gene conversion, sister chromatid exchange, and replication fork repair, but not following single strand annealing (SSA)(Figure 3.3F). Harvested BM B220^{+}IgM^{-} cells from RaDR-GFP mice were cultured in the presence or absence of IL7 with or without CXCL12. Three hours prior to assaying, cells were treated with Camptothecin, a topoisomerase inhibitor, which induces the HR, but not NHEJ, pathway. Flow cytometry for the frequency of GFP+ cells revealed similar rates of HR in the presence or absence of IL-7. In contrast, HR was significantly

depressed in the presence of CXCL12 (Figure 3.3F). These data suggest that by repressing HR, CXCR4 favors NHEJ DNA repair in small pre-B cells undergoing recombination.

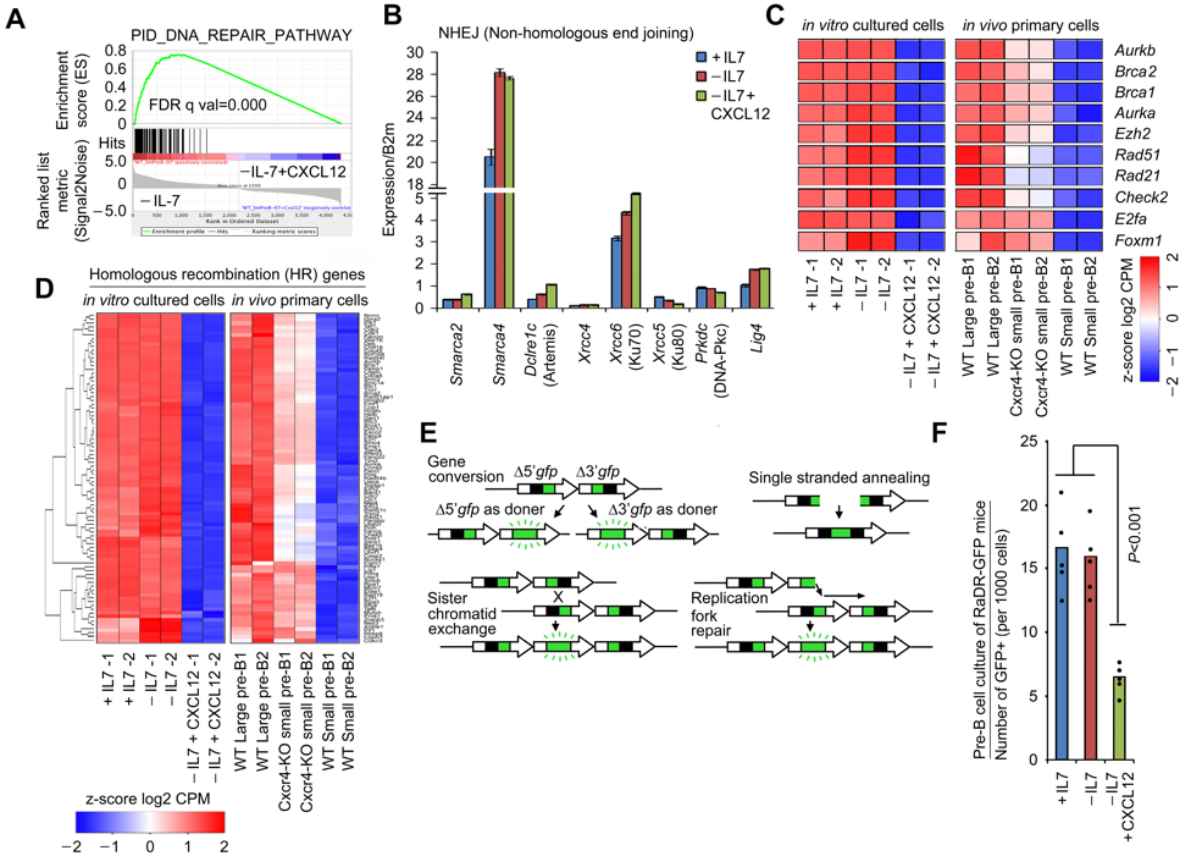


Figure 3.3 CXCR4 signaling represses homologous recombination

A, GSEA of DNA-repair pathway in –IL7+CXCL12 cultured pre-B cells compared to –IL7 cultured pre-B cells. **B**, Quantitative PCR analysis of the expression of classical Non Homologous End Joining (NHEJ) pathway genes *Smarca2, Smarca4, Dclre1c* (encodes Artemis), *Xrcc4*, *Xrcc6* (Ku70), *xrcc5* (Ku80), *Prkdc* (DNA-PKC) *and Lig4* (Ligase 4) in WT pre-B cells cultured as indicated (n=3). Data presented as average ± s.d. **C**, Heat maps showing expression of genes (from *in vitro* and *in vivo* RNA-Seq experiments) involved in the Homologous Recombination (HR) repair pathway in WT pre-B cells cultured in the indicated conditions (left panel) and in WT large pre-B, *Cxcr4*$^{fl/fl}$-*mb1*-cre$^{+/-}$ small pre-B and WT small pre-B cells (right panel). **D**, Heat maps showing expression of HR repair pathway signature genes (from *in vitro* and *in vivo* RNA-Seq) *Aurkb*, *Brca1*, *Brca2*, *Aurka*, *Ezh2*, *Rad51*, *Rad21*, *Check2*, *E2fa* and *Foxm1* in cultured WT pre-B cells cultured at the indicated conditions (left panel) and in WT large pre-B, *Cxcr4*$^{fl/fl}$-*mb1*-cre$^{+/-}$ small pre-B and WT small pre-B (right panel). **E**, Schematic diagram showing how HR at the RaDR-GFP substrate gives rise to fluorescence following gene conversion, sister chromatid exchange, and replication fork repair, but not following SSA. **f,** Number of GFP^{+} B cell progenitors generated after activation of HR repair pathway when pre-B cells from RaDR-GFP mice were cultured as indicated with treatment of Camptothecin (10μM) for final 3 hrs (n=5). Individual experiments plotted.

4 Discussion

Mechanisms of B cell tolerance are a major force in shaping the mature B cell repertoire. Development of autoreactive specificities is a necessary risk for the ability to recognize and defend against countless pathogens. Receptor editing represents an important mechanism that enables once autoreactive B cells from being lost by negative selection. This is made possible by usage of two light chain isotypes (κ and λ) present in almost all mammals. Decades of research have helped identify key events that govern this process. Through those work it has become clear that light chain editing occurs first at *Igκ* then proceeds to *Igλ* loci. However, reports have indicated the λ-chain has distinct physiochemical properties particularly effective at quenching autoreactivity to DNA, Yet, molecular mechanism of receptor editing that lead to the usage of the λ-chain remain ill defined. Herein, we elucidate transcriptional and epigenetic features of editing cells and show chemokine receptor CXCR4 is essential for maintaining the editing state and orchestrating *Igλ* light recombination. Editing small pre-B cells possess unique open chromatin regions enriched for E2A and NF-κB; key regulators of *Igλ* recombination and receptor editing. These regions showed evidence of having been acquired by partial differentiation into the immature B cell stage. Transcriptionally, editing cells showed upregulation of oxidative phosphorylation, DNA repair mechanisms and CXCR4. CXCR4 deficiency in autoreactive immature B cells showed a selective defect in *Igλ* recombination and tolerance induction, suggesting CXCR4 is necessary for the differentiation of Igλ+ B cells during receptor editing. We then suggest mediators downstream of CXCR4 signaling that may be selectively promoting *Igλ* recombination.

4.1 Transcriptional profile of receptor editing

As stated earlier, transcription factors such as NF-κB (Verkoczy et al., 2005), Rag1/Rag2 (Tiegs et al., 1993) and E2A (Beck et al., 2009) have been demonstrated to be essential for receptor editing to occur. However, many of these studies relied on transgenic expression of rearranged BCRs that were monospecific to model antigens such as Hen egg lysozyme. This raised the question of whether these observations reflected events that would occur in a normal polyclonal setting. Our data presented here confirms the validity of these earlier studies.

An initial concern of this work was our ability to capture the initiating events of receptor editing in immature B cells. This is because the immature B cell compartment of κ-mac mice consists of three different populations of cells: 1) Those that have completed κ recombination and are beginning to express an Igκ+ BCR on the cell surface. These cells are immediately rendered autoreactive by ubiquitously expressed κ-macroself transgene and return to the small pre-B cell stage (Pelanda and Torres, 2012). This may be an iterative process as other Igκ+ BCR produced from secondary rearrangements will also return to the small pre-B cells stage to continuing editing. 2) Autoreactive B cells that finally complete λ recombination and begin to express an Igλ+ BCR on the cell surface. 3) Lastly, a minority of cells that only underwent successful light recombination at the *Igλ* locus and never expressed a recombined κ chain and now express an Igλ+ BCR. Separate isolation of these cells would be technically very challenging. We thought to separate the immature compartment into IgM^{lo} vs IgM^{hi} cells as other studies have noted that in WT mice, cells engaged in editing can be captured in the IgM^{lo} gate (Cadera et al., 2009). For instance, WT immature B cells express low levels of RAG (Mandal et al., 2018). In fact, PI3K dependent

signaling from an unligated BCR represses RAG to enforce isotype exclusion (Pelanda, 2014). However, in WT mice when the immature B cell compartment is divided into IgM^{lo} vs IgM^{hi} cells, IgM^{lo} cells still express high levels of RAG (Yu et al., 1999). In our hands, we have made a similar observation. This suggests that examination of the entire immature B cell compartment could hinder capturing the transcriptional profile of editing cells. However, we reasoned that since virtually all cells undergo receptor editing in κ-mac mice, we would be able to robustly capture genes involved with receptor editing (Ait-Azzouzene et al., 2005). Indeed, RNA-seq of $Ig\kappa^{del}$ and κ-mac immature B cells showed over 4,000 genes were differentially expressed between them. Key genes upregulated in κ-mac immature B cells included *Rag2*, *Gadd45a* and SLC genes *Vpreb1*, *Vpreb2*, and *Igll1.* These genes have all been demonstrated to be involved in receptor editing (Nemazee, 2017). Since the entire immature B cell compartment was sorted for these studies, defection of these genes is indicative of our ability to robustly capture genes involved in receptor editing.

GSEA of the differentially expressed genes (DEGs) between $Ig\kappa^{del}$ and κ-mac immature B cell highlighted a surprising number of metabolic pathways highly enriched in κ-mac immature B cells. These pathways included GO_OXIDATIVE_PHOSPHORYLATION, HALLMARK_GLYCOLYSIS, HALLMARK_FATTY_ACID_OXIDIATION_METABOLISM. The OxPhos pathway was the most enriched and we sought to validate this finding. We used MitoSOX™ reagent that selectively binds to mitochondrial ROS, a byproduct of OxPhos. We observed higher levels of ROS in immature κ-mac cells compared to WT and $Ig\kappa^{del}$ cells. When we partitioned the immature B cell gate into IgM^{lo} vs IgM^{hi} cells, we observed higher ROS in the IgM^{lo} gate. This confirmed that editing cells utilized OxPhos. Evidence for OxPhos and ROS production was not just found transcriptionally but many genes involved in OxPhos

pathways had open chromatin regions (OCRs) present only in κ-mac cells. For instance, genes involved in detoxification of ROS were uniquely upregulated or more accessible in κ-mac cells. *Keap1* had more accessibility at the promoter in small pre-B cells while *Sod1* and *Sod2* were upregulated in κ-mac immature pre-B cells.

Upregulation of OxPhos gene programs is likely a consequence of BCR ligation. *In vitro* differentiation of naïve mature B cells stimulated with anti-IgM upregulate OxPhos during differentiation to plasmablasts (Akkaya et al., 2018). In this system, stimulation through the BCR alone was sufficient to induce OxPhos and increase ROS production, which led to mitochondrial dysfunction. However, when naïve B cells were treated with cognate antigen and TLR agonist (CpG), this resulted in upregulation of OxPhos and ROS detoxification genes allowing for efficient plasmablast differentiation. This suggests that OxPhos and more importantly regulation of ROS is necessary for B cell differentiation.

Many studies had noted that ROS is associated with genomic instability and DNA damage (Ito et al., 2007; Sallmyr et al., 2008). High levels of ROS would be dangerous for cells engaged in receptor editing because recombination inherently induces localized genomic instability. Likewise, strong upregulation of antioxidant genes in κ-mac immature B cells is perhaps a compensatory mechanism for higher levels of ROS produced by increased OxPhos. This could be crucial to safeguard renewed light chain recombination as autoreactive immature B cells return to the small pre-B cell stage. We suggest CXCR4 regulates levels of ROS during receptor editing, as loss of CXCR4 expression led to a decrease in ROS detoxification pathways. While the precise role of OxPhos in receptor editing remains to be determined, these data highlight metabolism as another regulated step in this process.

4.2 Receptor editing cells possess a unique epigenetic profile

To date, no studies have assessed the epigenetic profile of editing cells. We suspected given the dynamics involved in halting immature B cell differentiation, down regulation of the autoreactive BCR and renewing light chain recombination would result in profound epigenetic differences between editing and non-editing cells. Indeed, when we compared OCRs within κ-mac and Igκ^{del} small pre-B and immature B cells we discovered they were dissimilar in meaningful ways. Hierarchical clustering of WT, Igκ^{del} and κ-mac small pre-B and immature B cells based on the Pearson correlations of TF motif enrichment within OCRs, showed Igκ^{del} small pre-B cells were more similar to WT small pre-B cells. Conversely, κ-mac immature B cells were more similar to WT immature B cells. Motifs for E2A where highly enriched in κ-mac small pre-B cells compared to Igκ^{del} and WT cells. We interpreted this result as support for a current model of λ recombination requiring elevated amounts of E2A (Beck et al., 2009). Increased accessibility at E2A motifs genome-wide would facilitates E2A binding and activating important genes. NF-κB is another important transcription factor necessary for the development of Igλ+ B cells (Derudder et al., 2009). Likewise, κ-mac immature B cells were uniquely enriched for NF-κB motifs. An in-depth analysis of OCRs within small pre-B cells showed ~11k unique regions in κ-mac small pre-B cells. These regions were selectively enriched at gene promoters and exons. Moreover, gene in the vicinity of these unique OCRs were enriched for chemokine signaling pathways, P53 signaling and NF-κB pathways. Together, this demonstrates editing cells selectively open chromatin sites for key TFs necessary for receptor editing and *Igλ* recombination.

Autoreactive BCR down modulation after antigen ligation is an initiation step for receptor editing. This process has to be quick to attenuate signals that would lead to

overactivation, premature differentiation and cell death (Setz et al., 2018). After removal of surface BCR, immature B cells return to the small pre-B stage in order to initiate secondary rearrangements at *Igκ* before progressing to *Igλ*. We saw evidence of this 'back differentiation' when we overlapped the ~11k unique OCRs of κ-mac small pre-B cells with OCRs of WT immature B cells. Almost 10,000 of these regions overlapped with large pre-B and immature B cells. Studies profiling OCRs during early B cell development showed a significant loss of OCRs at the small pre-B cell stage compared to the large pre-B and immature B cell stage (Mandal et al., 2018). Thus, the significant overlap of the ~11k unique κ-mac small pre-B OCRs with immature B cells is most likely indicative of κ-mac small pre-B cells having differentiated into immature B cells. Together, these data support a model that autoreactive B cells returning to the small pre-B cell stage retain OCRs from the immature B cells stage.

Comparing the transcriptional profile and epigenetic state of Igκ^{del} and κ-mac small pre-B and immature B cells, we found remarkable transcriptional similarity between small pre-B cells yet large transcriptional differences at the immature B cell stage. Conversely, at the small pre-B cell stage these cells were epigenetically very different. Large transcriptional differences at the immature B cell stage is perhaps a reflection of the numerous processes that are occurring during receptor editing initiation. It is reasonable to expect multiple gene regulatory networks such as endocytosis, metabolism, and chemotaxis to be recruited to halt autoreactive immature B cells in development and to initiate a partial reversal in development. Once at the small pre-B cells stage, the light chain recombination machinery would dominate much of the transcriptional profile. Since κ-mac and Igκ^{del} small pre-B cells would require similar mechanisms for light chain recombination, some transcriptional

similarity at the small pre-B cell stage is expected. Furthermore, given that both Igκ^{del} and κ-mac small pre-B cells need to rearrange λ light change genes to fully differentiate into immature B cells, the degree of transcriptional overlap in small pre-B cells is not surprising. Thus, a combination of forward and reverse development imprints an epigenetic and transcriptional landscape leading to a unique cell state permissive for receptor editing.

4.3 *Igλ* locus during receptor editing

We were surprised to find that accessibility at the *Igλ* locus was higher in κ-mac small pre-B cells than Igκ^{del} and WT small pre-B cells. We expected accessibility at the *Igλ* locus to be lowest in WT small pre-B cells because most of those cells engage in κ recombination. However, we observed comparable accessibility levels at *Igλ* in both Igκ^{del} and WT small pre-B cells. This could reflect a delay in Igκ^{del} small pre-B cell development as these cells try and fail to recombine *Igκ* genes. The accumulation of small pre-B cells in Igκ^{del} mice supports this possibility. Furthermore, accessibility at *Igκ* was similar in κ-mac and Igκ^{del} mice both at the small pre-B and immature B cell stage. This suggests that the *Igκ* locus is accessible for recombination in both models. Eventually, *Igλ* becomes accessible for recombination in Igκ^{del} mice and these cells can progress to the immature B cells stage. Conversely, in κ-mac small pre-B cells secondary rearrangements at *Igκ* leads to κ inactivation at both alleles which hastens the use of the *Igλ* locus for recombination.

4.4 CXCR4 is essential for receptor editing

An extensive body of work has shown that autoreactive B cells in the BM increase surface expression of CXCR4 upon antigen ligation (Beck et al., 2014; Greaves et al., 2019; Schram et al., 2008; Tze et al., 2005). For years this was thought to retain autoreactive B cells in the BM

to facilitate tolerance mechanisms. However, recent evidence that CXCR4 signaling directs *Igκ* recombination independent of BM positional effects extends the functions of CXCR4 (Mandal et al., 2019). Herein, we demonstrate CXCR4 is also necessary for receptor editing and *Igλ* recombination. We show that in the context of receptor editing absence of CXCR4 impairs early B cells development beginning at the small pre-B cell stage. Compared to *Cxcr4$^{fl/fl}$ mb1-cre+* mice, κ-mac *Cxcr4KO* mice had a greater loss of immature B cells. Interestingly, while surface expression of Igκ was not detected in κ-mac *Cxcr4KO* immature B cells, intracellular *Igκ* was detected in equal proportions to κ-mac *Cxcr4$^{fl/fl}$* mice. This demonstrated that *Igκ* recombination was likely unaffected in these mice and that *Igλ* recombination was selectively impaired in the absence of CXCR4.

Equal proportions of intracellular Igκ+ was a surprising result given that CXCR4 directs *Igκ* recombination (Mandal et al., 2019). In the *Cxcr4$^{fl/fl}$ mb1-cre+* mice, loss of CXCR4 resulted in diminished *Igκ* rearrangement yet some cells were still able to assemble a BCR and differentiate into immature B cells. In the κ-mac setting, those relatively few cells that successfully rearranged *Igκ* genes were strongly selected into the editing pool were lack of CXCR4 completely blocked *Igλ* rearrangements. This potentially demonstrates *Igλ* is more dependent on CXCR4 for completion of light chain recombination than *Igκ*.

A simple explanation for the lack of recombination in κ-mac *Cxcr4KO* small pre-B cells is because of a reduction in RAG genes. Loss of CXCR4 was associated with increased expression of genes that inhibit PTEN function. With loss of PTEN, PI3K signaling dominates and RAG is inhibited. However, in the absence of *CXCR4* the reduction of *Rag1* and *Rag2* was modest. While the reduction in RAG was sufficient to halt *Igλ* recombination, *Igκ*

recombination remained intact. This may be because mouse κ light chain RSSs are more efficient at mediating recombination than RSSs of the λ light chain (Ramsden and Wu, 1991). Thus, it is tempting to conclude that increasing the levels of RAG might rescue recombination, but that seems unlikely to be a mechanism that would be evolutionarily conserved. RAG has been demonstrated to have off-target effects (Teng et al., 2015); therefore, increasing concentrations of RAG protein places editing cells under risk of genomic instability and aberrant translocation. Thus, other mechanisms necessary for recombination are likely impaired in the absence of CXCR4.

κ-mac *Cxcr4KO* small pre-B cells showed a severe defect in transcription of virtually all λ locus genes compared to control κ-mac *Cxcr4*$^{fl/fl}$. In particular, transcription across the RSSs was also lost in the absence of CXCR4. Assaying for *Vλ1-Jλ1* and *Vλ1-Jλ3* recombination products using qPCR of cDNA from sorted small pre-B cells supported this model because recombination was nearly absent in κ-mac *Cxcr4KO* cells. This suggest that lack of transcription was sufficient to suppress recombination. Indeed, Abarrategui et al, demonstrated that transcriptional blockade through the mouse T cell receptor-α locus suppressed Vα-to-Jα recombination (Abarrategui and Krangel, 2006). Therefore, CXCR4 dependent transcription of λ genes could be necessary for successful recombination.

We next assessed *Igλ* chromatin accessibility in κ-mac *Cxcr4KO* small pre-B cells. Given the lack of transcription, we expected chromatin accessibility to also be reduced at the *Igλ* locus in κ-mac *Cxcr4KO* small pre-B cells. However, ATAC-seq of the *Igλ* locus showed that *Igλ* accessibility was mostly unaffected, with only accessibility at *Cλ* genes significantly impaired. *C* light chain genes do not participate in recombination and are joined to a recombined *VJ* gene segment via mRNA splicing. At the *Igκ* locus *Cκ* is approximate 4kb from

the *Jκ* promoters while *Jλ-Cλ* gene clusters are much closer at 1kb apart. Given the close proximity of *Jλ-Cλ* gene clusters, lack of accessibility at *Cλ* may affect transcription of the recombined *Vλ-Jλ* gene segment. This implies that recombination, in fact, is not impaired in the absence of CXCR4 but transcription post recombination is impaired due to the lack of accessibility at *Cλ* genes.

Therefore, two possibilities remain for how CXCR4 regulates λ recombination: 1) A reduction in RAG expression in the absence of CXCR4. This suggests the defect in λ recombination is because RAG is not expressed at efficient levels to drive *Igλ* recombination. 2) The lack of accessibility at *Cλ* impairs transcription of a recombined VJ-Cλ gene and this was missed because cDNA was used instead of genomic DNA to measure recombination. Thus, to test these two possibilities we propose assessing λ recombination at the genomic level in sorted small pre-B cells from κ-mac *Cxcr4$^{fl/fl}$* and κ-mac *Cxcr4KO* mice. A lack of recombination products would strongly implicate failure to induce recombination machinery in the absence of CXCR4 as a reason for the lack of *Igλ*+ B cells. Although, this would not rule out the possibility that transcription of the recombined *VJ-Cλ* gene is impaired. However, if *Vλ-Jλ* recombination is intact, then this would strongly implicate transcription post recombination as the reason for the lack of *Igλ*+ B cells. This would establish a paradigm that autoreactive B cells upregulate CXCR4 upon antigen stimulation in order to initiation transcription of *Igλ* genes to then promote recombination.

Still, an intriguing question is what molecules downstream of CXCR4 regulates *Igλ* recombination. A likely candidate downstream of CXCR4 signaling is E2A (Mandal et al., 2019). Higher levels of E2A induce λ germline transcription, promote H3K4Me3, a histone

mark for RAG recruitment, and initiate *VJ* recombination (Beck et al., 2009). Furthermore, E2A has been shown to bind both *Igλ* enhancers Eλ2-4 and Eλ3-1. E2A is activated downstream of ERK, a principal signaling mediator of CXCR4 (Lin et al., 2010; Mandal et al., 2019; Mandal et al., 2011). Moreover, OCRs in κ-mac small pre-B cells compared to other small pre-B cells were enriched for E2A binding motifs. Thus, the lack of light chain recombination in κ-mac *Cxcr4KO* at *Igλ* could be a consequence of reduced ERK signaling and E2A activity. Several experiments can be done to test this hypothesis. First, E2A ChIP at *Igλ* enhancers Eλ2-4, Eλ3-1 can be performed in κ-mac *Cxcr4*$^{fl/fl}$ *and* κ-mac *Cxcr4KO* small pre-B cells to assess whether loss of E2A binding the λ enhancers is the cause of diminished λ recombination. Perhaps a more definitive experiment would be to overexpress E2A in κ-mac *Cxcr4KO* small pre-B cells and assess whether λ recombination is restored.

Our data also identified other transcriptional pathways CXCR4 is involved in such as regulation of cell activation, NF-κB signaling, cell chemotaxis. (Derudder et al., 2009; Tze et al., 2005). Regulation of cell activation is important in preventing autoimmunity. Numerous autoimmune diseases have been linked to aberrant signaling and activation in B and T cells (Stanford and Bottini, 2014). 'Tonic' signals from autoreactive BCRs assist cells in positive selection and escape for tolerance mechanism. Many of the genes found in the regulation of cell activation pathway were genes associated with inhibiting B cell activation and dampening signals from antigen receptors. We did not specifically investigate whether κ-mac *Cxcr4KO* developed autoimmunity, but that outcome is unlikely. A majority of the cells were unable to progress to the immature B cell stage and there was an accumulation of IgM- cells in the spleen. We think this would prevent development of autoimmune pathology.

Many studies have shown that prolonged survival can rescue defects in receptor editing (Ait-Azzouzene et al., 2005; Hertz and Nemazee, 1997; Setz et al., 2018; Vela et al., 2008). The classical experiment involves overexpression of Bcl2 in BCR stimulated immature B cells undergoing receptor editing will result in more Igλ+ B cells. Derudder et al. identified NF-κB signals as the source of these survival signals. B cell specific deletion in many components of the canonical and non-canonical NF-κB pathway produced less Igλ+ B cells, while *Igκ* recombination was not impaired (Derudder et al., 2009). They also demonstrated that Bcl2 overexpression could rescue Igλ+ B cell development in the absence of NF-κB. In our data, we observed NF-κB motif enrichment at OCRs in κ-mac immature B cells, as well as upregulation of transcriptional pathways. CXCR4 deficiency also resulted in a loss of NF-κB pathways. We also observed loss of Bcl2 in CXCR4 deficient κ-mac small pre-B cells as well. This implies that CXCR4 may promote survival of editing cells as well as directing *Igλ* recombination.

In conclusion, our results further extend the cellular functions of CXCR4. The BM microenvironment has been known for many years to be necessary for receptor editing (Sandel and Monroe, 1999). The ability of CXCR4 to halt BM egress and upregulate many aspects of receptor editing outlined in this work implicates the CXCL12-CXCR4 signaling axis as vital molecules governing B cell central tolerance.

4.5 Pending Experiments

The data outlined here raise several experiments that will be performed to solidify the conclusions being made. The experiments were discussed above but are summarized in greater detail below:

1) To address the possibility that CXCR4 regulates *Igλ* recombination through transcription, we will sort small pre-B κ-mac *CXCR4KO and* κ-mac *CXCR4$^{fl/fl}$* cells and isolate genomic DNA and synthesize cDNA. We will determine the status of *Igλ* rearrangement from genomic DNA for λ1 λ3 and λx gene products. Using cDNA, we will determine germline transcription of *Vλ1* and *Jλ1*. These results will allow us to determine whether recombination is impaired, or transcription post recombination is impaired in κ-mac *CXCR4KO*.

2). To address whether lack of E2A binding to *Igλ* enhancers prevents *Igλ* recombination, we will perform ChIP for E2A binding in sorted sort small pre-B κ-mac *CXCR4KO and* κ-mac *CXCR4$^{fl/fl}$* cells. Furthermore, retroviral transduction of E2A into the same populations could be an alternative means to address a similar question of E2A levels dedicating *Igλ* recombination efficiency. Lastly, we will culture κ-mac *CXCR4KO* and κ-mac *CXCR4$^{fl/fl}$* cells small pre-B cells in the presence of CXCL12 and in the presence or absence of ERK inhibitor (FR 180204; CAS 865362-74-9; Millipore) and determine *Igλ* recombination is impaired in the absence of ERK signaling.

Model

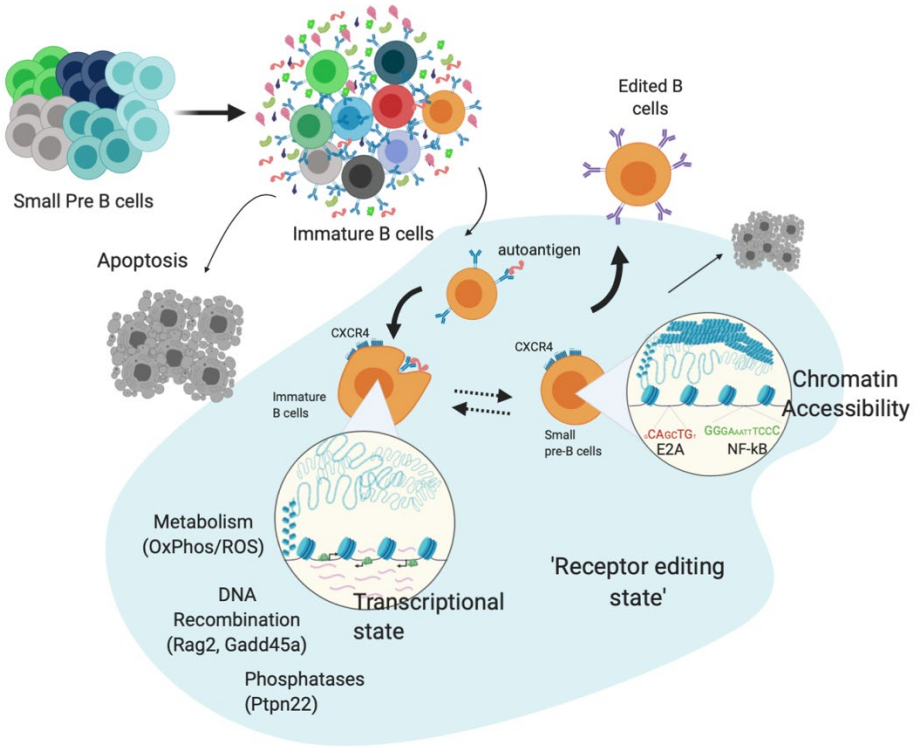


Figure 4.1 Unique transcriptional and epigenetic state
A schematic depicting key features of the 'receptor editing state.' As small pre-B cells differentiate into immature B cells, those immature B cells with autoreactive specificities undergo the receptor editing process. Editing small pre-B cells possess an epigenetic landscape that is a blend of both the small pre-B and immature B cell stage. These cells possess open chromatin regions acquired from having partially differentiated into immature B cells. The open chromatin regions are enriched for key transcription factors such as NF-κB and E2A. This epigenetic landscape establishes the transcriptionally processes occurring at immature B cells stage. Immature B cells engaged in receptor editing are transcriptional and epigenetically distinct from other developing B cells. Editing cells have a unique transcriptional profile marked by upregulation of metabolic processes such as oxidative phosphorylation and reactive oxygen species production. As expected, these cells also express genes involved in recombination. Lastly, editing cells upregulate phosphatases, presumably to dampen strong signaling from autoreactive BCR ligation.

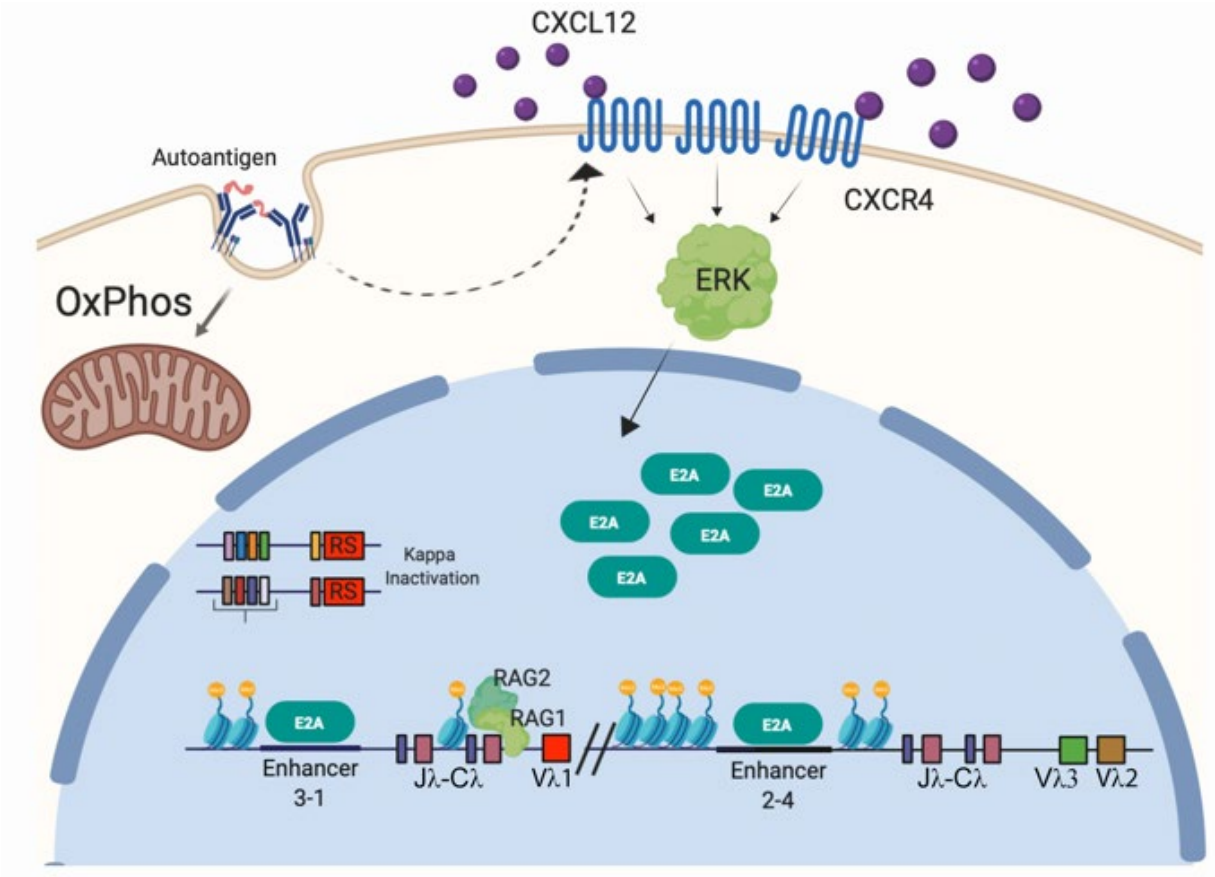


Figure 4.2 CXCR4 regulation of *Igλ* recombination
A schematic depicting the proposed mechanism of how CXCR4 directs *Igλ* recombination. Rapid internalization of the autoreactive BCR leads to upregulation of CXCR4 expression. CXCR4 signaling leads to ERK activation. ERK activation leads to the accumulation of E2A and facilitates *Igλ* recombination by binding the Igλ enhancers.

On the basis of the data presented here, we can summarize that receptor editing is comprised of unique transcriptional and epigenetic states. Autoreactive immature B cells initiating receptor editing at the immature B cell stage upregulate CXCR4 surface expression and are retained in the BM. Editing cells upregulate metabolic pathways namely oxidative phosphorylation which leads to an accumulation of reactive oxygen species. CXCR4 induces upregulation of phosphatases and antioxidants to limit ROS accumulation and BCR signaling. Editing cells also increase accessibility at E2A and NF-κB motifs and transition back to the

small pre-B cell stage to undergo light chain recombination. Here, CXCR4 promotes recombination of *Igλ* genes following inactivation of *Igκ* alleles. ERK signals downstream of CXCR4 increase availability of E2A that binds *Igλ* enhancers and promotes recombination. *Created by BioRender*

5 Future Directions

The data presented here define the transcriptional and epigenetic state governing B cell receptor editing and demonstrates CXCR4 as a crucial mediator of this process. However, this work raises intriguing questions about understanding *Igλ* repertoire formation, the contribution of autoreactive BCR signaling to the epigenetic and transcriptional state of editing cells.

5.1 Contribution of autoreactive BCR signaling to receptor editing

While many important genes upregulated in κ-mac cells are dependent on CXCR4 signaling, the number of these genes was small compared to the total number of genes upregulated in κ-mac cells compared to Igκ^{del} cells. This implicates an alternative receptor cooperates with CXCR4 to establish the transcriptional program of cells undergoing receptor editing. A likely candidate is an autoreactive BCR. The initiating events of receptor editing involve strong BCR signaling prior to receptor down modulation. As outlined earlier in the introduction, loss of BCR expression at the immature B cell stage induces expression of many genes including CXCR4 but the transcriptional consequence of signals produced from an autoreactive BCR and how these signals may cooperate with CXCR4 during receptor editing has not been established. To address this, an *in vitro* system could be developed where immature B cells from BCL2 transgenic mice harboring autoreactive specificities from a knock-in heavy and light chain can be cultured with or without its cognate antigen in the presence or absence of CXCL12. A subset of these conditions could be treated with tyrosine kinase inhibitors such as Herbimycin to block signals from the BCR. RNA-seq of these various conditions would provide greater resolution into what genes are dependent on autoreactive BCR signaling, the

loss of tonic BCR signaling and what genes are dependent on CXCR4 signaling during receptor editing.

5.2 Receptor editing is an instructive process

Mouse κ light chain RSSs are more efficient at mediating recombination than RSSs of the λ light chain (Ramsden and Wu, 1991). This is presumably because the κ locus RSSs more closely resemble canonical RSSs. Within the λ locus, the RSS of the V and J genes contain different mutations from each other leading to speculation that the overrepresentation of λ1 in the mouse repertoire is because the RSSs for the *Vλ1-Jλ1* facilitate Rag1 recognition and binding than other λ gene products. In agreement with this hypothesis, ChIP-seq of catalytically inactive Rag1-D708 transgenic mice showed an increase in Rag1 binding signal at *Jλ1* compared to *Jλ3* and *Jλ2* (Teng et al., 2015). This has the potential to obscure instructive processes during receptor editing that would promote usage of other λ allotypes. For instance, λx has been shown to be particular effective at tolerizing autoreactive B cells specific for dsDNA (Prak et al., 1994). Our ATAC-seq data shows increased accessibility at Eλ2-4 and Vλx in κ-mac small pre-B cells. In addition, Hi-C of WT immature pre-B cells showed more chromosomal interactions within the TAD that contained λx gene components. To formally test the hypothesis that receptor editing is an instructive process, we propose using CRISPR to revert the λ RSSs to canonical RSSs. These mice would be crossed with κ-mac mice containing a recombined heavy and κ light chain knock-in. We predict these mice would develop normally to the immature B cell stage where antigen ligation with the κ-macroself would force usage of the *Igλ* locus and λx would be a major product. In the absence

of receptor editing, the different λ allotypes would be represented evenly in the mature repertoire.

5.2 CXCR4 in splenic B cell development

We showed that in the spleen of κ-mac *Cxcr4KO* mice there was a large accumulation of B220+CD19+CD93+CD23-IgD^{lo}IgM^{-} cells. Transitional cells in the κ-mac *Cxcr4KO* mice were also diminished but mature B cells were comparable with their floxed controls. When we compared RNA-seq of this population with κ-mac *Cxcr4KO* small pre-B and immature B cells and performed K-means clustering, these cells shared the most similarity with κ-mac *Cxcr4KO* small pre-B cells. However, this is not enough evidence to pinpoint the true origin of these cells. Their cellular phenotype is reminiscent of anergic B cells (Goodnow et al., 1988) that are also IgM^{lo}IgD^{hi}; but anergic B cells also express CD23, which these do not. We encourage further investigation of these observations as they would provide a more complete understanding of CXCR4 in B cell biology.

6 Materials and Methods

Mice

κ-macroself (stock #006259) were cryorecovered from Jackson laboratories. All κ-macroself mice used were hemizygous for the κ-macroself gene. Note: breeding should be performed with hemizygous κ-macroself female mice paired with WT CD45.1 male mice. Igκ^{del} mice were a generous gift from David Nemazee (The Scripps Research Institute). Wildtype CD45.1 C57BL/6J (stock #002014) mice were purchased from Jackson laboratories and housed in the University of Chicago animal facilities. *Cxcr4$^{fl/fl}$ mb1-cre+* mice were a gift from Malay Mandal (University of Chicago). Female and male mice were used at 7–12 weeks of age, and studies carried out in accordance with the guidelines of the Institutional Animal Care and Use Committee at the University of Chicago (Protocol No. 71577).

Isolation and flow cytometry of BM B cell progenitors.

BM was collected from WT, Igκ^{del}, κ-mac, κ-mac *Cxcr4*$^{fl/fl}$ *mb1*-cre- or κ-mac *Cxcr4$^{fl/fl}$ mb1*-cre+ mice and cells were resuspended in staining buffer: (3% (v/v) fetal bovine serum (FBS) in 1x phosphate buffered saline. Erythrocytes were lysed with ACK lysis buffer (Lonza cat # 10-548E) and cells were stained with rat anti-CXCR4 (2B11), rat anti-CD43 (S7), rat anti-IgM (R6–60.2), rat anti-IgD (11–36), rat anti-CD19 (1D3) and rat anti-B220 (RA3-6B2l), (all from BD Biosciences) as described previously (Mandal et al., 2011; Mandal et al., 2009) and viability dye eFluor 506 (eBioscience). Pre-pro-B cells (Lin−CD19−B220+IgM−), pro-B cells (Lin−CD19+B220+CD43+IgM−), large pre-B cells (Lin−B220+CD43−IgM−FSChi), small pre-B cells (Lin−B220+CD43− IgM−FSClow) and immature B cells (Lin−B220+CD43−IgM+) were isolated by cell sorting with a FACSAria II (BD Biosciences).

Detection of mitochondrial reactive oxygen species (MitoSOX™)

Flow cytometric detection of ROS production, BM was extracted from mice and after erythrocyte lysis, $2x10^6$ cells were resuspended in fresh 1x HBSS (cat # 14025092; pH 7.2). Cells were stained with viability dye eFluor 506 (eBioscience) to exclude dead cells. Cells were resuspended in 2% FCS and HBSS (pH 7.2) and stained with antibodies. Cells were washed twice 2% FCS and 1% HBSS (pH 7.2) and stain with HBSS supplemented with 5 μM of MitoSOX™ red (Thermo Fisher Scientific cat # M36008) in the dark at 37°C for 30 minutes. Cells were then washed twice with plain HBSS and resuspended with 2% FCS and 1% HBSS (pH 7.2) and rapidly analyzed in flow cytometry.

Assay for Transposase Accessible Chromatin and Sequencing (ATAC-seq)

$1.2x10^5$ small pre-B and immature B cells sorted from WT, $Ig\kappa^{del}$ κ-mac, κ-mac *CXCR4*$^{fl/fl}$ and κ-mac *CXCR4*$^{fl/fl}$ *mb1-cre*+ mice were used for ATAC-seq. Cells were centrifuged at 500g at 4°C for 5 minutes, washed with 1X PBS, and centrifuged again. Cells were resuspended in lysis buffer (10mM Tris-HCl pH7.4, 10mM NaCl, 3mM MgCl2, 0.1% IGEPAL CA-630) and immediately centrifuged at 500g at 4°C for 10 minutes. Pellet was resuspended in transposition reaction buffer (25μl 2x Tagment Buffer (FC-121-1030, Illumina), 2.5μl Tagment DNA Enzyme, 22.5μl nuclease free H_2O) for 30 minutes at 37°C. DNA was purified with a Qiagen MinElute Kit and amplified with Nextera PCR Primers (Illumina Nextera Index Kit) and NEBNext PCR Master Mix (M0541, New England BioLabs) for 11 cycles. Amplified DNA was purified with a Qiagen PCR cleanup kit. Libraries were sequenced on a HiSeq 4000 sequencer at University of Chicago Genomics Facility.

RNA isolation and RNA-Sequencing (RNA-seq)

FACS sorted cells from small pre-B and immature B cells sorted from WT, Igκ^{del} κ-mac, κ-mac *CXCR4$^{fl/fl}$* and κ-mac *CXCR4$^{fl/fl}$ mb1-cre*+ mice placed in total RNA was isolated using Trizol (15596026, Invitrogen) following the protocol described by the Immunological Genome Project (Immgen.org). Libraries were generated (Kit, RS-122-2101 TruSeq Stranded mRNA LT-SetA) and sequenced by the University of Chicago Genomics Facility on the NextSeq500.

ATAC-seq and RNA-seq quantification

For ATAC–seq, read alignments were first adjusted to account for TAC transposon binding: +4 bp for +strand alignments, −5 bp for −strand alignments. The open chromatin enrichment track was generated by first creating a bedGraph from the raw reads using bedtools genomcov57, then converted to bigWig using UCSC tool bedGraphToBigWig; tracks were normalized by the sum of alignment lengths over 1 billion. Open chromatin peaks were called using Macs2 (Zhang et al., 2008) with no model set and no background provided; peaks with a score of >5 were retained. To quantitatively measure changes in epigenetic enrichment, we first identified empirical regulatory elements (RE) based on the peak calls obtained from each sample in the study. Peaks were merged into a uniform set of regulatory elements using bedtools merge (Quinlan and Hall, 2010). Enrichment levels for each RE were then quantified with featureCounts (Liao et al., 2019).

For RNA-seq, raw reads were aligned to reference genome mm9 in a splice aware manner using STAR (Dobin et al., 2013; Dobin and Gingeras, 2015). Gene expression was quantified using featureCounts within the RSubread R package (Liao et al., 2019) against UCSC genes with Ensembl Ig genes from mm10 converted to mm9 coordinates with UCSC liftOver.

Hi-C preparation and Analysis

Flow sorted WT small pre-B and WT Immature B cells (5x10^6 for each stage) were crosslinked with 1% formaldehyde following wash with ice cold PBS. Then the membranes were lysed keeping the nuclei as intact as possible following by restriction digestion with 100U of MboI (NEB R0147). The DNA ends were then marked with biotin, ligated proximally and reversed the crosslinks. DNA shearing and Size selection were then performed for fragments 300-500bp. DNA amount was quantified by Qubit dsDNA High Sensitivity Assay (Life technologies, Q32854) and biotin pull down was done to prepare final in-situ Hi-C library to be quantified and sequenced using an Illumina sequencing platform.

Sequenced paired-end DNA reads were aligned to the mouse mm9 genome with Bowtie2 and reads were filtered further, with the HiC-Pro and HiC-bench workflow (Lazaris et al., 2017; Servant et al., 2015), to verify which read pairs should be used for downstream analysis. HiC-contact files were then generated with Juicer tools. Further analyses were performed using JuiceBox, Homer HiC software, and Hi-C domain caller (Heinz et al., 2010; Lin et al., 2012; Servant et al., 2015).

Differences between the Hi-C profiles of WT small pre-B and WT immature B was calculated using diffHIC and HiC-Compare implemented through R (Lun, A) as well as Homer's "getHiCcorrDiff.pl". The genome was divided into active (A) versus inactive (B) chromatin compartments based on the correlation matrix ("runHiCpca.pl"). Changes in chromatin compartments between B cell subpopulation were determined using "findHiCCompartments.pl". TADs were determined with the Hi-C Domain caller, TADbits and HiCExplorer (Ramírez et al., 2018; Serra et al., 2017; Wolff et al., 2018; Wolff et al., 2020). TAD discovery will be confirmed and visualized with Juicebox (Durand et al., 2016).

BedGraph files of compartments and domains will be visualized with IGV, IGB or UCSC browsers. Significant 3D interactions will be identified with HiCCUPS, DiffHiC and Homer over a resolution range of 5-25 kB, and compared between WT Small pre-B and WT Immature B. Significant interactions associated with nearby genes was done using Homer's "annotateInteractions.pl", over a 10 kb window. Contact matrices, significant genome interactions, was visualized with Juicebox, Circos, Cytoscape, Hic-Explorer and HiCPlotter (Akdemir and Chin, 2015; Krzywinski et al., 2009; Ramírez et al., 2018; Shannon et al., 2003; Wolff et al., 2018; Wolff et al., 2020). Difference detection between multiple HiC datasets was performed using the R package HiCcompare (Stansfield et al., 2018) and differences with an average expression lower than a 10% quantile were filtered out. Finally, only significant differences, with a p-value less than 0.01were kept and graphed on an MD plot. Significant differences were also graphed on individual chromosome contact plots using HiCPlotter (Akdemir and Chin, 2015).

qPCR

QPCR reactions were analyzed in triplicate in a 25 µl reaction containing 10 µM of each primer. Cycling conditions for all qPCRs were: 50°C for 2 min, 95°C for 3 min, followed by 45 cycles of 95°C for 15 s and 61°C for 1 min. C_T values were determined using the Applied Biosystems 7300 Real-Time PCR System and the provided application-specific software. Data were exported and analyzed with Microsoft Excel. Data were analyzed according to the $\Delta\Delta C_T$ method. The C_T mean for each sample was calculated and standard deviations (s) were calculated for each mean C_T value. C_T means were first defined as the difference between sample and corresponding input control and the amount of target gene was normalized to the negative control *HPRT1*.

Differential expression (RNA-seq and ATAC–seq)

Differential expression statistics (fold-change and P value) were computed using edgeR, on raw expression counts obtained from quantification (either genes or ATAC open chromatin regions) (McCarthy et al., 2012; Robinson et al., 2010). Group comparisons were made using the generalized linear modeling capability in edgeR. In all cases, P values were adjusted for multiple testing using the FDR correction of Benjamini and Hochberg. Significant genes were determined based on an FDR threshold of 5% (0.05) in the group comparison.

For differential expression between κ-mac *Cxcr4*$^{fl/fl}$ *mb1*-cre- and κ-mac *Cxcr4*$^{fl/fl}$ *mb1*-cre+ small pre-B cells, batch correction was performed using the Remove Unwanted Variation from RNA-Seq Data (RUV-Seq) package from bioconductor in R. Specifically the RUVs with RUV-Seq program was used.

Principal component analysis (PCA) was performed using the R function prcomp with the parameter ‘‘scale=TRUE’’.

Pathway analysis and motif analysis (ATAC-seq)

ATAC-seq analysis for WT, Igκ^{del} and κ-mac small pre-B cells was performed HOMER (Hypergeometric Optimization of Motif EnRichment)(Duttke et al., 2019). Motif enrichment analysis was performed using ‘findmotifsgenome.pl’ command with -size 300 and other default parameters. On occasion statistical significance was determined using a hypergeometric distribution by setting the -h flag. Open chromatin regions were assigned to their nearest using the ‘annotatepeaks.pl’ command. Gene pathway enrichment of genes near open chromatin regions was performed by setting the -go flag. The results from the ‘wikipathway’ were explored and data was replotted using ‘ComplexHeatmap’ in R. For Figure 2.8 C-E, histograms of reads around transcription factor binding motifs were

generated by centering open chromatin regions containing the investigated motifs using the 'annotatepeaks.pl' command, followed by counting reads from individual experiments at single base pair resolution in a radius of 1500 bp (or 150 bp) around the peak centers using the 'annotatepeaks.pl' function with flags '-hist -fragLength 1.' Overlaps of open chromatin regions from ATAC-seq data was performed using the HOMER ''mergePeaks'' command

Gene Pathway analysis (RNA-seq)

Metascape web portal (www.metascape.org) was used to pathway analysis of differential expressed genes found from analyzing RNA-seq data (Zhou et al., 2019).

Visualization of ATAC-seq

Visualization of ATAC-seq data was done using Integrated Genome Browser (Freese et al., 2016).

GSEA

Gene Set Enrichment Analysis (GSEA) was performed on normalized Log_2 [counts per million] values from RNA-seq data. 'Ratio of Classes' metric was used to determine statistical significance. Gene sets found within the HALLMARK and C5 GO molecular signatures pathways were used. GSEA program can downloaded here: http://software.broadinstitute. Results from GSEA analysis were replotted using the custom R function ReplotGSEA.R (https://github.com/PeeperLab/Rtoolbox/blob/master/R/ReplotGSEA.R).

Clustering and heatmaps

Clustering for heatmaps was performed using complete linkage hierarchical clustering of z-scored normalized values from RNA-seq or ATAC-seq and plotted as a heatmap using the

'hclust' and 'heatmap.3' functions in or the 'ComplexHeatmap' from Bioconductor in R (Gu et al., 2016; Gu et al., 2014).

Primers

qPCR

The following primers were obtained from (Beck et al., 2009)

Rag1-Fwd 5'- ACCCGATGAAATTCAACACCC-3'

Rag1-Rev 5'- CTGGAACTACTGGAGACTGTTCT-3'

Rag2-Fwd 5'- ACACCAAACAATGAGCTTTCCG-3'

Rag2-Rev 5'- CCGTATCTGGGTTCAGGGAC-3'

HPRT1-Fwd 5'- GTTAAGCAGTACAGCCCCAAA-3'

HPRT1-Rev 5'- AGGGCATATCCAACAACAAACTT-3'

The following primers were obtained from (Bai et al., 2007), (Zou et al., 1993) and (Pathak et al., 2008)

Vλ1 5'-AGAAGCTTGTGACTCAGGAATCTGCA-3'

Jλ1 5'-CAGGATCCTAGGACAGTCAGTTTGGT-3'

Jλ2,3 5'-TAGGACAGTGACCTTGGTTCCACCG-3'

Genotyping

κ-mac primers

Transgene-Fwd 5'-TGGTGAGGGGAGGGATAAGTGAGGC-3'

Transgene-Rev 5'-CGCTGAACCTTGATGGGACCCC-3'

Internal Control Fwd 5'-CTAGGCCACAGAATTGAAAGATCT-3'

Internal Control Rev 5'-GTAGGTGGAAATTCTAGCATCATC C-3'

Igκ^{del} primers

Null Kappa Fwd 5'ATGTACTTGTGGATGCAGAGGCTG-3'

Null Kappa Rev 5'-ATGTGGAATGTGTGCGAGGCCAGA-3'

Cxcr4-flox primers

Fwd 5'-CCACCCAGGACAGTGTGACTCTAA-3'

Rev 5'-GATGGGATTTCTGTATGAGGATTAGC-3'

Cre primers

Fwd 5'-ACACACCTGGAAGATGCTCCT-3'

Rev 5'-CAAAGTCAGTGCGTTCAAAGG-3'

Genotyping PCR protocol

For κ-mac mice PCR cycling:

For Igκ^{del} mice PCR cycling: 94°C for 2 mins, then for 30 cycles 94°C for 45 seconds 68C for 1 min and 72°C for 1 min. Final extension was performed at 72°C for 7 mins.

For κ-mac mice PCR cycling: 94°C for 2 mins, touchdown (94°C for 20 secs 65°C for 15 secs and 68°C for 10 secs) for 10 cycles then for 30 cycles 94°C for 15 secs 60°C for 15 sec and 72°C for 10 secs. Final extension was performed at 72°C for 2 mins.

For all other genotypes PCR cycling: 94°C for 5 mins, touchdown (94°C for 45 secs 65°C for 45 secs and 72°C for 45 secs) for 5 cycles then for 30 cycles 94°C for 45 secs 60°C for 15 sec and 72°C for 45 secs. Final extension was performed at 72°C for 10 mins.

Computer software versions

R_3.5.3, edgeR_3.28.1, circlize_0.4.10, ComplexHeatmap_2.2.0, chromVAR_1.4.1, RUVSeq_1.20.0, Homer_v.4.10.3, bedtools_v2.27.1-9-g5f83cacb, macs2_2.2.6, GSEA_4.0.3., IGB_9.1.4.

Statistical analysis

Statistical analyses were performed with GraphPad Prism. For analyses with multiple comparisons, data were analyzed by analysis of variance in combination with Tukey's multiple comparisons test. Bar graphs are displayed as the mean ± S.E.M. Significance as defined by *P* value or FDR (*q* value) are defined in the figures, figure legends, corresponding text or in tables. Additional quantitative methods and statistical criteria are mentioned above based on their respective technology and analysis approaches.

Accession Codes

Gene Expression Omnibus accession codes for our publicly available datasets are as follows: GSE103057 (ATAC-Seq and RNA-Seq for WT pro-B, WT large pre-B, WT small pre-B and WT immature B cells); GSE103057 (IRF4 ChIP-Seq for WT small pre-B cells), ChIP-seq CTCF and RAD21 (GSE40173 and GSE47766)

7 References

Abarrategui, I., and Krangel, M.S. (2006). Regulation of T cell receptor-α gene recombination by transcription. Nature Immunology *7*, 1109-1115.

Ait-Azzouzene, D., Verkoczy, L., Peters, J., Gavin, A., Skog, P., Vela, J.L., and Nemazee, D. (2005). An immunoglobulin C kappa-reactive single chain antibody fusion protein induces tolerance through receptor editing in a normal polyclonal immune system. Journal of Experimental Medicine *201*, 817-828.

Akdemir, K.C., and Chin, L. (2015). HiCPlotter integrates genomic data with interaction matrices. Genome Biol *16*, 198.

Akkaya, M., Traba, J., Roesler, A.S., Miozzo, P., Akkaya, B., Theall, B.P., Sohn, H., Pena, M., Smelkinson, M., Kabat, J.*, et al.* (2018). Second signals rescue B cells from activation-induced mitochondrial dysfunction and death. Nature Publishing Group *19*, 871-884.

Allman, D., Li, J., and Hardy, R.R. (1999). Commitment to the B lymphoid lineage occurs before DH-JH recombination. J Exp Med *189*, 735-740.

Allman, D., Lindsley, R.C., DeMuth, W., Rudd, K., Shinton, S.A., and Hardy, R.R. (2001). Resolution of three nonproliferative immature splenic B cell subsets reveals multiple selection points during peripheral B cell maturation. J Immunol *167*, 6834-6840.

Almagro, J.C., Hernández, I., Ramírez, M.C., and Vargas-Madrazo, E. (1998). Structural differences between the repertoires of mouse and human germline genes and their evolutionary implications. Immunogenetics *47*, 355-363.

Alt, F.W., Enea, V., Bothwell, A.L., and Baltimore, D. (1980). Activity of multiple light chain genes in murine myeloma cells producing a single, functional light chain. Cell *21*, 1-12.

Alt, F.W., Yancopoulos, G.D., Blackwell, T.K., Wood, C., Thomas, E., Boss, M., Coffman, R., Rosenberg, N., Tonegawa, S., and Baltimore, D. (1984). Ordered rearrangement of immunoglobulin heavy chain variable region segments. Embo j *3*, 1209-1219.

Amin, R.H., and Schlissel, M.S. (2008). Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. Nature Publishing Group *9*, 613-622.

Bai, L., Chen, Y., He, Y., Dai, X., Lin, X., Wen, R., and Wang, D. (2007). Phospholipase Cgamma2 contributes to light-chain gene activation and receptor editing. Molecular and Cellular Biology *27*, 5957-5967.

Bajoghli, B., Aghaallaei, N., Hess, I., Rode, I., Netuschil, N., Tay, B.H., Venkatesh, B., Yu, J.K., Kaltenbach, S.L., Holland, N.D*., et al.* (2009). Evolution of genetic networks underlying the emergence of thymopoiesis in vertebrates. Cell *138*, 186-197.

Bankovich, A.J., Raunser, S., Juo, Z.S., Walz, T., Davis, M.M., and Garcia, K.C. (2007). Structural insight into pre-B cell receptor function. Science *316*, 291-294.

Bannish, G., Fuentes-Pananá, E.M., Cambier, J.C., Pear, W.S., and Monroe, J.G. (2001). Ligand-independent signaling functions for the B lymphocyte antigen receptor and their role in positive selection during B lymphopoiesis. J Exp Med *194*, 1583-1596.

Barreto, V., and Cumano, A. (2000). Frequency and characterization of phenotypic Ig heavy chain allelically included IgM-expressing B cells in mice. J Immunol *164*, 893-899.

Bassing, C.H., Swat, W., and Alt, F.W. (2002). The Mechanism and Regulation of Chromosomal V(D)J Recombination. Cell *109*, S45-S55.

Batista, C.R., Li, S.K.H., Xu, L.S., Solomon, L.A., and DeKoter, R.P. (2017). PU.1 Regulates Ig Light Chain Transcription and Rearrangement in Pre-B Cells during B Cell Development. Journal of immunology (Baltimore, Md. : 1950) *198*, 1565-1574.

Beagan, J.A., and Phillips-Cremins, J.E. (2020). On the existence and functionality of topologically associating domains. Nature Genetics *52*, 8-16.

Beck, K., Peak, M.M., Ota, T., Nemazee, D., and Murre, C. (2009). Distinct roles for E12 and E47 in B cell specification and the sequential rearrangement of immunoglobulin light chain loci. The Journal of Experimental Medicine *206*, 2271-2284.

Beck, T.C., Gomes, A.C., Cyster, J.G., and Pereira, J.P. (2014). CXCR4 and a cell-extrinsic mechanism control immature B lymphocyte egress from bone marrow. The Journal of Experimental Medicine *211*, 2567-2581.

Bendall, H.H., Sikes, M.L., and Oltz, E.M. (2001). Transcription factor NF-kappa B regulates Ig lambda light chain gene rearrangement. The Journal of Immunology *167*, 264-269.

Bernard, O., Hozumi, N., and Tonegawa, S. (1978). Sequences of mouse immunoglobulin light chain genes before and after somatic changes. Cell.

Boller, S., and Grosschedl, R. (2014). The regulatory network of B-cell differentiation: a focused view of early B-cell factor 1 function. Immunological reviews *261*, 102-115.

Boothby, M.R., Hodges, E., and Thomas, J.W. (2019). Molecular regulation of peripheral B cells and their progeny in immunity. Genes & Development *33*, 26-48.

Brack, C., Hirama, M., Lenhard-Schuller, R., and Tonegawa, S. (1978). A complete immunoglobulin gene is created by somatic recombination. Cell *15*, 1-14.

Bradl, H., Wittmann, J., Milius, D., Vettermann, C., and Jäck, H.M. (2003). Interaction of murine precursor B cell receptor with stroma cells is controlled by the unique tail of lambda 5 and stroma cell-associated heparan sulfate. J Immunol *171*, 2338-2348.

Brandsma, I., and Gent, D.C. (2012). Pathway choice in DNA double strand break repair: observations of a balancing act. Genome Integr *3*, 9.

Brass, A.L., Kehrli, E., Eisenbeis, C.F., Storb, U., and Singh, H. (1996). Pip, a lymphoid-restricted IRF, contains a regulatory domain that is important for autoinhibition and ternary complex formation with the Ets factor PU.1. Genes & Development *10*, 2335-2347.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods *10*, 1213-1218.

Burnet, F.M. (1959). The Clonal Selection Theory of Acquired Immunity. Cambridge University Press, Cambridge, UK.

Cadera, E.J., Wan, F., Amin, R.H., Nolla, H., Lenardo, M.J., and Schlissel, M.S. (2009). NF-kappaB activity marks cells engaged in receptor editing. The Journal of Experimental Medicine *206*, 1803-1816.

Cambier, J.C., Gauld, S.B., Merrell, K.T., and Vilen, B.J. (2007). B-cell anergy: from transgenic models to naturally occurring anergic B cells? Nature Immunology *7*, 633-643.

Caro-Maldonado, A., Wang, R., Nichols, A.G., Kuraoka, M., Milasta, S., Sun, L.D., Gavin, A.L., Abel, E.D., Kelsoe, G., Green, D.R., and Rathmell, J.C. (2014). Metabolic reprogramming is required for antibody production that is suppressed in anergic but exaggerated in chronically BAFF-exposed B cells. Journal of immunology (Baltimore, Md. : 1950) *192*, 3626-3636.

Carson, S., and Wu, G.E. (1989). A linkage map of the mouse immunoglobulin lambda light chain locus. Immunogenetics *29*, 173-179.

Casellas, R., Shih, T.-A.Y., Kleinewietfeld, M., Rakonjac, J., Nemazee, D., Rajewsky, K., and Nussenzweig, M.C. (2001). Contribution of Receptor Editing to the Antibody Repertoire. Science *291*, 1541-1544.

Cerny, A., Heusser, C.H., Scandinavian, S.S., and 1986 Generation of Agammaglobulinaemic Mice by Prenatal and Postnatal Exposure to Polyclonal or Monoclonal Anti-IgM Antibodies. Wiley Online Library.

Chen, J., Trounstine, M., Kurahara, C., Young, F., Kuo, C.-C., Xu, Y., Loring, J.F., Alt, F.W., and Huszar, D. (1993). B cell development in mice that lack one or both immunoglobulin kappa light chain genes. embopress.org.

Chisolm, D.A., Cheng, W., Colburn, S.A., Silva-Sanchez, A., Meza-Perez, S., Randall, T.D., and Weinmann, A.S. (2019). Defining Genetic Variation in Widely Used Congenic and Backcrossed Mouse Models Reveals Varied Regulation of Genes Important for Immune Responses. Immunity *51*, 155-168.e155.

Christie, D.A., Xu, L.S., Turkistany, S.A., Solomon, L.A., Li, S.K.H., Yim, E., Welch, I., Bell, G.I., Hess, D.A., and DeKoter, R.P. (2015). PU.1 opposes IL-7-dependent proliferation of developing B cells with involvement of the direct target gene bruton tyrosine kinase. Journal of immunology (Baltimore, Md. : 1950) *194*, 595-605.

Chung, J.B., Sater, R.A., Fields, M.L., Erikson, J., and Monroe, J.G. (2002). CD23 defines two distinct subsets of immature B cells which differ in their responses to T cell help signals. Int Immunol *14*, 157-166.

Chung, J.B., Silverman, M., and Monroe, J.G. (2003). Transitional B cells: step by step towards immune competence. TRENDS in Immunology *24*.

Clark, A.G., Fan, Q., Brady, G.F., Mackin, K.M., Coffman, E.D., Weston, M.L., and Foster, M.H. (2013). Regulation of basement membrane-reactive B cells in BXSB, (NZBxNZW)F1, NZB, and MRL/lpr lupus mice. Autoimmunity *46*, 188-204.

Clark, M.R., Mandal, M., Ochiai, K., and Singh, H. (2014). Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. Nature Reviews Immunology *14*, 69-80.

Collins, A.M., and Watson, C.T. (2018). Immunoglobulin Light Chain Gene Rearrangements, Receptor Editing and the Development of a Self-Tolerant Antibody Repertoire. Frontiers in Immunology *9*, 2249.

Cooper, M.D., PETERSON, R.D., and GOOD, R.A. (1965). DELINEATION OF THE THYMIC AND BURSAL LYMPHOID SYSTEMS IN THE CHICKEN. Nature *205*, 143-146.

Cooper, M.D., Peterson, R.D.A., South, M.A., and Good, R.A. (1966). THE FUNCTIONS OF THE THYMUS SYSTEM AND THE BURSA SYSTEM IN THE CHICKEN. Journal of Experimental Medicine *123*, 75-102.

Cumano, A., Dorshkind, K., Gillis, S., and Paige, C.J. (1990). The influence of S17 stromal cells and interleukin 7 on B cell development. Eur J Immunol *20*, 2183-2189.

Cyster, J.G. (2010). B cell follicles and antigen encounters of the third kind. Nature Immunology *11*, 989-996.

Dai, X., James, R.G., Habib, T., Singh, S., Jackson, S., Khim, S., Moon, R.T., Liggitt, D., Wolf-Yadlin, A., Buckner, J.H., and Rawlings, D.J. (2013). A disease-associated PTPN22 variant promotes systemic autoimmunity in murine models. J Clin Invest *123*, 2024-2036.

Das, S., Nikolaidis, N., and Nei, M. (2009). Genomic organization and evolution of immunoglobulin kappa gene enhancers and kappa deleting element in mammals. Mol Immunol *46*, 3171-3177.

Degner, S.C., Verma-Gaur, J., Wong, T.P., Bossen, C., Iverson, G.M., Torkamani, A., Vettermann, C., Lin, Y.C., Ju, Z., Schulz, D*., et al.* (2011). CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. Proceedings of the National Academy of Sciences *108*, 9566-9571.

Degner, S.C., Wong, T.P., Jankevicius, G., and Feeney, A.J. (2009). Cutting Edge: Developmental Stage-Specific Recruitment of Cohesin to CTCF Sites throughout Immunoglobulin Loci during B Lymphocyte Development. The Journal of Immunology *182*, 44-48.

Derudder, E., Cadera, E.J., Vahl, J.C., Wang, J., Fox, C.J., Zha, S., van Loo, G., Pasparakis, M., Schlissel, M.S., Schmidt-Supprian, M., and Rajewsky, K. (2009). Development of immunoglobulin lambda-chain-positive B cells, but not editing of immunoglobulin kappa-chain, depends on NF-kappaB signals. Nature Publishing Group *10*, 647-654.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Dobin, A., and Gingeras, T.R. (2015). Mapping RNA-seq Reads with STAR. Curr Protoc Bioinformatics *51*, 11.14.11-11.14.19.

Doughty, C.A., Bleiman, B.F., Wagner, D.J., Dufort, F.J., Mataraza, J.M., Roberts, M.F., and Chiles, T.C. (2006). Antigen receptor-mediated changes in glucose metabolism in B lymphocytes: role of phosphatidylinositol 3-kinase signaling in the glycolytic control of growth. Blood *107*, 4458-4465.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst *3*, 99-101.

Duttke, S.H., Chang, M.W., Heinz, S., and Benner, C. (2019). Identification and dynamic quantification of regulatory elements using total RNA. Genome research *29*, 1836-1846.

Edelman, G.M. (1973). Antibody structure and molecular immunology. Science *180*, 830-840.

Eisenbeis, C.F., Singh, H., and Storb, U. (1995). Pip, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcriptional activator. Genes & Development *9*, 1377-1387.

Essafi, A., Fernández de Mattos, S., Hassen, Y.A., Soeiro, I., Mufti, G.J., Thomas, N.S., Medema, R.H., and Lam, E.W. (2005). Direct transcriptional regulation of Bim by FoxO3a mediates STI571-induced apoptosis in Bcr-Abl-expressing cells. Oncogene *24*, 2317-2329.

Finkelman, F.D., Mond, J.J., and Metcalf, E.S. (1983). Effects of neonatal anti-delta antibody treatment on the murine immune system. I. Suppression of development of surface IgD+ B cells and expansion of a surface IgM+ IgD- B lymphocyte population. The Journal of Immunology *131*, 593-600.

Flemming, A., Brummer, T., Reth, M., and Jumaa, H. (2003). The adaptor protein SLP-65 acts as a tumor suppressor that limits pre-B cell expansion. Nat Immunol *4*, 38-43.

Freese, N.H., Norris, D.C., and Loraine, A.E. (2016). Integrated genome browser: visual analytics platform for genomics. Bioinformatics *32*, 2089-2095.

Fruman, D.A., Snapper, S.B., Yballe, C.M., Davidson, L., Yu, J.Y., Alt, F.W., and Cantley, L.C. (1999). Impaired B cell development and proliferation in absence of phosphoinositide 3-kinase p85alpha. Science *283*, 393-397.

Fujimoto, M., Fujimoto, Y., Poe, J.C., Jansen, P.J., Lowell, C.A., DeFranco, A.L., and Tedder, T.F. (2000). CD19 regulates Src family protein tyrosine kinase activation in B lymphocytes through processive amplification. Immunity *13*, 47-57.

Fulcher, D.A., and Basten, A. (1994). Reduced life span of anergic self-reactive B cells in a double-transgenic model. J Exp Med *179*, 125-134.

Funk, P.E., Stephan, R.P., and Witte, P.L. (1995). Vascular cell adhesion molecule 1-positive reticular cells express interleukin-7 and stem cell factor in the bone marrow. Blood *86*, 2661-2671.

Gause, A., Yoshida, N., Kappen, C., and Rajewsky, K. (1987). In vivo generation and function of B cells in the presence of a monoclonal anti-IgM antibody: implications for B cell tolerance. European Journal of Immunology *17*, 981-990.

Gauthier, L., Rossi, B., Roux, F., Termine, E., and Schiff, C. (2002). Galectin-1 is a stromal cell ligand of the pre-B cell receptor (BCR) implicated in synapse formation between pre-B and stromal cells and in pre-BCR triggering. Proc Natl Acad Sci U S A *99*, 13014-13019.

Gay, D., Saunders, T., Camper, S., and Weigert, M. (1993). Receptor editing: an approach by autoreactive B cells to escape tolerance. Journal of Experimental Medicine *177*, 999-1008.

Georg, I., Díaz-Barreiro, A., Morell, M., Pey, A.L., and Alarcón-Riquelme, M.E. (2019). BANK1 interacts with TRAF6 and MyD88 in innate immune signaling in B cells. Cellular & Molecular Immunology.

Ghia, P., Gratwohl, A., Signer, E., Winkler, T.H., Melchers, F., and Rolink, A.G. (1995). Immature B cells from human and mouse bone marrow can change their surface light chain expression. Eur J Immunol *25*, 3108-3114.

Giltiay, N.V., Chappell, C.P., and Clark, E.A. (2012). B-cell selection and the development of autoantibodies. Arthritis research & therapy *14 Suppl 4*, S1-13.

Gong, S., and Nussenzweig, M.C. (1996). Regulation of an early developmental checkpoint in the B cell pathway by Ig beta. Science *272*, 411-414.

Goodnow, C.C., Crosbie, J., Adelstein, S., Lavoie, T.B., Smith-Gill, S.J., Brink, R.A., Pritchard-Briscoe, H., Wotherspoon, J.S., Loblay, R.H., Raphael, K*., et al.* (1988). Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. Nature *334*, 676-682.

Goodnow, C.C., Sprent, J., Fazekas de St Groth, B., and Vinuesa, C.G. (2005). Cellular and genetic mechanisms of self tolerance and autoimmunity. Nature *435*, 590-597.

Gorman, J.R., van der Stoep, N., Monroe, R., Cogne, M., Davidson, L., and Alt, F.W. (1996). The Igκ 3′ Enhancer Influences the Ratio of Igκ versus Igλ B Lymphocytes. Immunity *5*, 241-252.

Grandien, A., Fucs, R., Nobrega, A., Andersson, J., and Coutinho, A. (1994). Negative selection of multireactive B cell clones in normal adult mice. European Journal of Immunology *24*, 1345-1352.

Greaves, S.A., Peterson, J.N., Strauch, P., Torres, R.M., and Pelanda, R. (2019). Active PI3K abrogates central tolerance in high-avidity autoreactive B cells. The Journal of Experimental Medicine *216*, 1135-1153.

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847-2849.

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize Implements and enhances circular visualization in R. Bioinformatics *30*, 2811-2812.

Hagman, J., Rudin, C.M., Haasch, D., Chaplin, D., and Storb, U. (1990). A novel enhancer in the immunoglobulin lambda locus is duplicated and functionally independent of NF kappa B. Genes & Development *4*, 978-992.

Halverson, R., Torres, R.M., and Pelanda, R. (2004). Receptor editing is the main mechanism of B cell tolerance toward membrane antigens. Nature Immunology *5*, 645-650.

Hamel, K.M., Mandal, M., Karki, S., and Clark, M.R. (2014). Balancing Proliferation with Igκ Recombination during B-lymphopoiesis. Frontiers in Immunology *5*, 139.

Hardy, R.R., Carmack, C.E., Shinton, S.A., Kemp, J.D., and Hayakawa, K. (1991). Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. Journal of Experimental Medicine *173*, 1213-1225.

Hartley, S.B., Cooke, M.P., Fulcher, D.A., Harris, A.W., Cory, S., Basten, A., and Goodnow, C.C. (1993). Elimination of self-reactive B lymphocytes proceeds in two stages: arrested development and cell death. Cell *72*, 325-335.

Hartley, S.B., Crosbie, J., Brink, R., Kantor, A.B., Basten, A., and Goodnow, C.C. (1991). Elimination from peripheral lymphoid tissues of self-reactive B lymphocytes recognizing membrane-bound antigens. Nature *353*, 765-769.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Henry-Bonami, R.A., Williams, J.M., Rachakonda, A.B., Karamali, M., Kendall, P.L., and Thomas, J.W. (2013). B lymphocyte "original sin" in the bone marrow enhances islet autoreactivity in type 1 diabetes-prone nonobese diabetic mice. J Immunol *190*, 5992-6003.

Herglotz, J., Unrau, L., Hauschildt, F., Fischer, M., Kriebitzsch, N., Alawi, M., Indenbirken, D., Spohn, M., Müller, U., Ziegler, M.*, et al.* (2016). Essential control of early B-cell development by Mef2 transcription factors. Blood *127*, 572-581.

Hertz, M., and Nemazee, D. (1997). BCR Ligation Induces Receptor Editing in IgM+IgD− Bone Marrow B Cells In Vitro. Immunity *6*, 429-436.

Herzenberg, L.A., and Herzenberg, L.A. (2004). Genetics, FACS, Immunology, and Redox: A Tale of Two Lives Intertwined. Annual Review of Immunology *22*, 1-31.

Herzenberg, L.A., Sweet, R.G., and Herzenberg, L.A. (1976). Fluorescence-activated cell sorting. Sci Am *234*, 108-117.

Hieter, P.A., Korsmeyer, S.J., Waldmann, T.A., and Leder, P. (1981). Human immunoglobulin kappa light-chain genes are deleted or rearranged in lambda-producing B cells. Nature *290*, 368-372.

Hippen, K.L., Schram, B.R., Tze, L.E., Pape, K.A., Jenkins, M.K., and Behrens, T.W. (2005). In vivo assessment of the relative contributions of deletion, anergy, and editing to B cell self-tolerance. The Journal of Immunology *175*, 909-916.

Hirose, J., Kouro, T., Igarashi, H., Yokota, T., Sakaguchi, N., and Kincade, P.W. (2002). A developing picture of lymphopoiesis in bone marrow. Immunol Rev *189*, 28-40.

Hozumi, N., and Tonegawa, S. (1976). Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. Proceedings of the National Academy of Sciences.

Inlay, M.A., Bhattacharya, D., Sahoo, D., Serwold, T., Seita, J., Karsunky, H., Plevritis, S.K., Dill, D.L., and Weissman, I.L. (2009). Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. Genes Dev *23*, 2376-2381.

Inlay, M.A., Tian, H., Lin, T., and Xu, Y. (2004). Important roles for E protein binding sites within the immunoglobulin kappa chain intronic enhancer in activating Vkappa Jkappa rearrangement. J Exp Med *200*, 1205-1211.

Ito, K., Takubo, K., Arai, F., Satoh, H., Matsuoka, S., Ohmura, M., Naka, K., Azuma, M., Miyamoto, K., Hosokawa, K*., et al.* (2007). Regulation of reactive oxygen species by Atm is essential for proper response to DNA double-strand breaks in lymphocytes. J Immunol *178*, 103-110.

Jaleco, A.C., Stegmann, A.P., Heemskerk, M.H., Couwenberg, F., Bakker, A.Q., Weijer, K., and Spits, H. (1999). Genetic modification of human B-cell development: B-cell development is inhibited by the dominant negative helix loop helix factor Id3. Blood *94*, 2637-2646.

Jiang, Q., Li, W.Q., Hofmeister, R.R., Young, H.A., Hodge, D.R., Keller, J.R., Khaled, A.R., and Durum, S.K. (2004). Distinct regions of the interleukin-7 receptor regulate different Bcl2 family members. Mol Cell Biol *24*, 6501-6513.

Jumaa, H., Wollscheid, B., Mitterer, M., Wienands, J., Reth, M., and Nielsen, P.J. (1999). Abnormal development and function of B lymphocytes in mice deficient for the signaling adaptor protein SLP-65. Immunity *11*, 547-554.

Jung, D., Giallourakis, C., Rev, R.M.A., and 2006 Mechanism and control of V (D) J recombination at the immunoglobulin heavy chain locus. annualreviews.org.

Karki, S., Kennedy, D.E., Mclean, K., Grzybowski, A.T., Maienschein-Cline, M., Banerjee, S., Xu, H., Davis, E., Mandal, M., Labno, C*., et al.* (2018). Regulated Capture of Vκ Gene Topologically Associating Domains by Transcription Factories. CellReports *24*, 2443-2456.

Kawasaki, K., Minoshima, S., Nakato, E., Shibuya, K., Shintani, A., Schmeits, J.L., Wang, J., and Shimizu, N. (1997). One-megabase sequence analysis of the human immunoglobulin lambda gene locus. Genome Research *7*, 250-261.

Kee, B.L. (2009). E and ID proteins branch out. Nature Reviews Immunology *9*, 175-184.

King, L.B., and Monroe, J.G. (2000). Immunobiology of the immature B cell: plasticity in the B-cell antigen receptor-induced response fine tunes negative selection. Immunological reviews *176*, 86-104.

Kondo, M., Weissman, I.L., and Akashi, K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell *91*, 661-672.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome research *19*, 1639-1645.

Lagergren, A., Månsson, R., Zetterblad, J., Smith, E., Basta, B., Bryder, D., Akerblad, P., and Sigvardsson, M. (2007). The Cxcl12, periostin, and Ccl9 genes are direct targets for early B-cell factor in OP-9 stroma cells. J Biol Chem *282*, 14454-14462.

Lam, E.W., Glassford, J., Banerji, L., Thomas, N.S., Sicinski, P., and Klaus, G.G. (2000). Cyclin D3 compensates for loss of cyclin D2 in mouse B-lymphocytes activated via the antigen receptor and CD40. J Biol Chem *275*, 3479-3484.

Lamoureux, J.L., Watson, L.C., Cherrier, M., Skog, P., Nemazee, D., and Feeney, A.J. (2007). Reduced receptor editing in lupus-prone MRL/lpr mice. J Exp Med *204*, 2853-2864.

Lawton, A.R., Asofsky, R., Hylton, M.B., and Cooper, M.D. (1972). Suppression of immunoglobulin class synthesis in mice. I. Effects of treatment with antibody to mu-chain. Journal of Experimental Medicine *135*, 277-297.

Lazaris, C., Kelly, S., Ntziachristos, P., Aifantis, I., and Tsirigos, A. (2017). HiC-bench: comprehensive and reproducible Hi-C data analysis designed for parameter exploration and benchmarking. BMC Genomics *18*, 22.

LeBien, T.W., and Tedder, T.F. (2008). B lymphocytes: how they develop and function. Blood *112*, 1570-1580.

Lefranc, M.P. (2001). Nomenclature of the human immunoglobulin kappa (IGK) genes. Experimental and clinical immunogenetics *18*, 161-174.

Li, Y., Li, H., and Weigert, M. (2002). Autoreactive B cells in the marginal zone that express dual receptors. J Exp Med *195*, 181-188.

Liao, Y., Smyth, G.K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res *47*, e47-e47.

Lin, Y.C., Benner, C., Mansson, R., Heinz, S., Miyazaki, K., Miyazaki, M., Chandra, V., Bossen, C., Glass, C.K., and Murre, C. (2012). Global changes in the nuclear positioning of genes and

intra- and interdomain genomic interactions that orchestrate B cell fate. Nat Immunol *13*, 1196-1204.

Lin, Y.C., Jhunjhunwala, S., Benner, C., Heinz, S., Welinder, E., Mansson, R., Sigvardsson, M., Hagman, J., Espinoza, C.A., Dutkowski, J.*, et al.* (2010). A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. Nat Immunol *11*, 635-643.

Llorian, M., Stamataki, Z., Hill, S., Turner, M., and Mårtensson, I.L. (2007). The PI3K p110delta is required for down-regulation of RAG expression in immature B cells. J Immunol *178*, 1981-1985.

Loder, F., Mutschler, B., Ray, R.J., Paige, C.J., Sideras, P., Torres, R., Lamers, M.C., and Carsetti, R. (1999). B cell development in the spleen takes place in discrete steps and is determined by the quality of B cell receptor-derived signals. J Exp Med *190*, 75-89.

Luning Prak, E.T., Monestier, M., and Eisenberg, R.A. (2011). B cell receptor editing in tolerance and autoimmunity. Annals of the New York Academy of Sciences *1217*, 96-121.

Malin, S., McManus, S., Cobaleda, C., Novatchkova, M., Delogu, A., Bouillet, P., Strasser, A., and Busslinger, M. (2010). Role of STAT5 in controlling cell survival and immunoglobulin gene recombination during pro-B cell development. Nature Immunology *11*, 171-179.

Malu, S., Malshetty, V., Francis, D., and Cortes, P. (2012). Role of non-homologous end joining in V(D)J recombination. Immunol Res *54*, 233-246.

Mandal, M., Hamel, K.M., Maienschein-Cline, M., Tanaka, A., Teng, G., Tuteja, J.H., Bunker, J.J., Bahroos, N., Eppig, J.J., Schatz, D.G., and Clark, M.R. (2015). Histone reader BRWD1 targets and restricts recombination to the Igk locus. Nature Immunology *16*, 1094-1103.

Mandal, M., Maienschein-Cline, M., Maffucci, P., Veselits, M., Kennedy, D.E., McLean, K.C., Okoreeh, M.K., Karki, S., Cunningham-Rundles, C., and Clark, M.R. (2018). BRWD1 orchestrates epigenetic landscape of late B lymphopoiesis. Nature communications *9*, 3888.

Mandal, M., Okoreeh, M.K., Kennedy, D.E., Maienschein-Cline, M., Ai, J., McLean, K.C., Kaverina, N., Veselits, M., Aifantis, I., Gounari, F., and Clark, M.R. (2019). CXCR4 signaling directs Igk recombination and the molecular mechanisms of late B lymphopoiesis. Nature Immunology *14*, 69.

Mandal, M., Powers, S.E., Maienschein-Cline, M., Bartom, E.T., Hamel, K.M., Kee, B.L., Dinner, A.R., and Clark, M.R. (2011). Epigenetic repression of the Igk locus by STAT5-mediated recruitment of the histone methyltransferase Ezh2. Nature Immunology *12*, 1212-1220.

Mandal, M., Powers, S.E., Ochiai, K., Georgopoulos, K., Kee, B.L., Singh, H., and Clark, M.R. (2009). Ras orchestrates exit from the cell cycle and light-chain recombination during early B cell development. Nature Immunology *10*, 1110-1117.

Mandel, E.M., and Grosschedl, R. (2010). Transcription control of early B cell differentiation. Curr Opin Immunol *22*, 161-167.

Massari, M.E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. Molecular and Cellular Biology *20*, 429-440.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res *40*, 4288-4297.

McLean, K.C., and Mandal, M. (2020). It Takes Three Receptors to Raise a B Cell. Trends in Immunology.

Meffre, E., and Wardemann, H. (2008). B-cell tolerance checkpoints in health and autoimmunity. Current Opinion in Immunology *20*, 632-638.

Melchers, F., ten Boekel, E., Seidl, T., Kong, X.C., Yamagami, T., Onishi, K., Shimizu, T., Rolink, A.G., and Andersson, J. (2000). Repertoire selection by pre-B-cell receptors and B-cell receptors, and genetic control of B-cell development from immature to mature B cells. Immunol Rev *175*, 33-46.

Meng, W., Li, Y., Xue, E., Satoh, M., Peck, A.B., Cohen, P.L., Eisenberg, R.A., and Luning Prak, E.T. (2012). B-cell tolerance defects in the B6.Aec1/2 mouse model of Sjögren's syndrome. J Clin Immunol *32*, 551-564.

Milne, C.D., Zhang, Y., and Paige, C.J. (2005). Stromal cells attract B-cell progenitors to promote B-cell-B-cell contact and maturation., Vol 62 Suppl 1 (John Wiley & Sons, Ltd).

Monroe, J.G. (2004). Ligand-independent tonic signaling in B-cell receptor function. Current Opinion in Immunology *16*, 288-295.

Monroe, J.G. (2006). ITAM-mediated tonic signalling through pre-BCR and BCR complexes. Nat Rev Immunol *6*, 283-294.

Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E.*, et al.* (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nature Genetics *34*, 267-273.

Mourcin, F., Breton, C., Tellier, J., Narang, P., Chasson, L., Jorquera, A., Coles, M., Schiff, C., and Mancini, S.J. (2011). Galectin-1-expressing stromal cells constitute a specific niche for pre-BII cell development in mouse bone marrow. Blood *117*, 6552-6561.

Murre, C. (2005). Helix-loop-helix proteins and lymphocyte development. Nature Immunology *6*, 1079-1086.

Murre, C., McCaw, P.S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell *56*, 777-783.

Nagasawa, T. (2006). Microenvironmental niches in the bone marrow required for B-cell development. Nature Immunology *6*, 107-116.

Nemazee, D. (2006). Receptor editing in lymphocyte development and central tolerance. Nature Reviews Immunology *6*, 728-740.

Nemazee, D. (2017). Mechanisms of central tolerance for B cells. Nature Reviews Immunology *17*, 281-294.

Nemazee, D., and Weigert, M. (2000). Revising B cell receptors. Journal of Experimental Medicine *191*, 1813-1817.

Nemazee, D.A., and Bürki, K. (1989). Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. Nature *337*, 562-566.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Norvell, A., Mandik, L., and Monroe, J.G. (1995). Engagement of the antigen-receptor on immature murine B lymphocytes results in death by apoptosis. The Journal of Immunology *154*, 4404-4413.

O'Neill, S.K., Getahun, A., Gauld, S.B., Merrell, K.T., Tamir, I., Smith, M.J., Dal Porto, J.M., Li, Q.Z., and Cambier, J.C. (2011). Monophosphorylation of CD79a and CD79b ITAM motifs initiates a SHIP-1 phosphatase-mediated inhibitory signaling cascade required for B cell anergy. Immunity *35*, 746-756.

Ochiai, K., Maienschein-Cline, M., Mandal, M., Triggs, J.R., Bertolino, E., Sciammas, R., Dinner, A.R., Clark, M.R., and Singh, H. (2012). A self-reinforcing regulatory network triggered by limiting IL-7 activates pre-BCR signaling and differentiation. Nature Immunology *13*, 300-307.

Ohnishi, K., and Melchers, F. (2003). The nonimmunoglobulin portion of lambda5 mediates cell-autonomous pre-B cell receptor signaling. Nat Immunol *4*, 849-856.

Okkenhaug, K., Bilancio, A., Farjot, G., Priddle, H., Sancho, S., Peskett, E., Pearce, W., Meek, S.E., Salpekar, A., Waterfield, M.D.*, et al.* (2002). Impaired B and T cell antigen receptor signaling in p110delta PI 3-kinase mutant mice. Science *297*, 1031-1034.

Okkenhaug, K., and Vanhaesebroeck, B. (2003). PI3K in lymphocyte development, differentiation and activation. Nature Reviews Immunology *3*, 317-330.

Oleinika, K., Mauri, C., and Blair, P.A. (2020). B cell Activation and B cell Tolerance. In The Autoimmune Diseases, N.R. Rose, and I.R. Mackay, eds. (Academic Press), pp. 171-187.

Panigrahi, A.K., Goodman, N.G., Eisenberg, R.A., Rickels, M.R., Naji, A., and Luning Prak, E.T. (2008). RS rearrangement frequency as a marker of receptor editing in lupus and type 1 diabetes. The Journal of Experimental Medicine *205*, 2985-2994.

Pathak, S., Ma, S., Trinh, L., and Lu, R. (2008). A Role for Interferon Regulatory Factor 4 in Receptor Editing. Molecular and Cellular Biology *28*, 2815-2824.

Pelanda, R. (2014). Dual immunoglobulin light chain B cells: Trojan horses of autoimmunity? Curr Opin Immunol *27*, 53-59.

Pelanda, R., Braun, U., Hobeika, E., Nussenzweig, M.C., and Reth, M. (2002). B cell progenitors are arrested in maturation but have intact VDJ recombination in the absence of Ig-alpha and Ig-beta. J Immunol *169*, 865-872.

Pelanda, R., Schwers, S., Sonoda, E., Torres, R.M., Nemazee, D., and Rajewsky, K. (1997). Receptor Editing in a Transgenic Mouse Model: Site, Efficiency, and Role in B Cell Tolerance and Antibody Diversification. Immunity *7*, 765-775.

Pelanda, R., and Torres, R.M. (2012). Central B-Cell Tolerance: Where Selection Begins. Cold Spring Harbor Perspectives in Biology *4*, a007146.

Petro, J.B., Gerstein, R.M., Lowe, J., Carter, R.S., Shinners, N., and Khan, W.N. (2002). Transitional type 1 and 2 B lymphocyte subsets are differentially responsive to antigen receptor signaling. J Biol Chem *277*, 48009-48019.

Popov, A.V., Zou, X., Xian, J., Nicholson, I.C., and Brüggemann, M. (1999). A human immunoglobulin lambda locus is similarly well expressed in mice and humans. The Journal of experimental medicine *189*, 1611-1620.

Porter, R.R. (1973). Structural studies of immunoglobulins. Science (New York, N.Y.) *180*, 713-716.

Powers, S.E., Mandal, M., Matsuda, S., Miletic, A.V., Cato, M.H., Tanaka, A., Rickert, R.C., Koyasu, S., and Clark, M.R. (2012). Subnuclear cyclin D3 compartments and the coordinated regulation of proliferation and immunoglobulin variable gene repression. J Exp Med *209*, 2199-2213.

Prak, E.L., Trounstine, M., Huszar, D., and Weigert, M. (1994). Light chain editing in kappa-deficient animals: a potential mechanism of B cell tolerance. Journal of Experimental Medicine *180*, 1805-1815.

Price, M.J., Patterson, D.G., Scharer, C.D., and Boss, J.M. (2018). Progressive Upregulation of Oxidative Metabolism Facilitates Plasmablast Differentiation to a T-Independent Antigen. CellReports *23*, 3152-3159.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics (Oxford, England) *26*, 841-842.

Quong , M.W., Martensson , A., Langerak , A.W., Rivera , R.R., Nemazee , D., and Murre , C. (2004). Receptor Editing and Marginal Zone B Cell Development Are Regulated by the Helix-Loop-Helix Protein, E2A. Journal of Experimental Medicine *199*, 1101-1112.

Radic, M.Z., Erikson, J., Litwin, S., and Weigert, M. (1993). B lymphocytes may escape tolerance by revising their antigen receptors. Journal of Experimental Medicine *177*, 1165-1173.

Raff, M.C., Megson, M., Owen, J.J., and Cooper, M.D. (1976). Early production of intracellular IgM by B-lymphocyte precursors in mouse. Nature *259*, 224-226.

Ramírez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., Grüning, B.A., Villaveces, J., Habermann, B., Akhtar, A., and Manke, T. (2018). High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nature Communications *9*, 189.

Ramsden, D.A., and Wu, G.E. (1991). Mouse kappa light-chain recombination signal sequences mediate recombination more frequently than do those of lambda light chain. Proceedings of the National Academy of Sciences of the United States of America *88*, 10721-10725.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

Retter, M.W., and Nemazee, D. (1998). Receptor Editing Occurs Frequently during Normal B Cell Development. Journal of Experimental Medicine *188*, 1231-1238.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Rothenberg, E.V. (2014). Transcriptional control of early T and B cell developmental choices. Annual Review of Immunology *32*, 283-321.

Rowland, S.L., DePersis, C.L., Torres, R.M., and Pelanda, R. (2010). Ras activation of Erk restores impaired tonic BCR signaling and rescues immature B cell differentiation. J Exp Med *207*, 607-621.

Sabouri, Z., Perotti, S., Spierings, E., Humburg, P., Yabas, M., Bergmann, H., Horikawa, K., Roots, C., Lambe, S., Young, C.*, et al.* (2016). IgD attenuates the IgM-induced anergy response in transitional and mature B cells. Nature communications *7*, 13381.

Sakamoto, S., Wakae, K., Anzai, Y., Murai, K., Tamaki, N., Miyazaki, M., Miyazaki, K., Romanow, W.J., Ikawa, T., Kitamura, D.*, et al.* (2012). E2A and CBP/p300 act in synergy to promote chromatin accessibility of the immunoglobulin κ locus. J Immunol *188*, 5547-5560.

Sallmyr, A., Fan, J., and Rassool, F.V. (2008). Genomic instability in myeloid malignancies: increased reactive oxygen species (ROS), DNA double strand breaks (DSBs) and error-prone repair. Cancer Lett *270*, 1-9.

Sandel, P.C., and Monroe, J.G. (1999). Negative Selection of Immature B Cells by Receptor Editing or Deletion Is Determined by Site of Antigen Encounter. Immunity *10*, 289-299.

Schatz, D.G., and Ji, Y. (2011). Recombination centres and the orchestration of V(D)J recombination. Nature Reviews Immunology *11*, 251-263.

Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nature Methods *14*, 975-978.

Schickel, J.N., Pasquali, J.L., Soley, A., Knapp, A.M., Decossas, M., Kern, A., Fauny, J.D., Marcellin, L., Korganow, A.S., Martin, T., and Soulas-Sprauel, P. (2012). Carabin deficiency in B cells increases BCR-TLR9 costimulation-induced autoimmunity. EMBO Mol Med *4*, 1261-1275.

Schiemann, B., Gommerman, J.L., Vora, K., Cachero, T.G., Shulga-Morskaya, S., Dobles, M., Frew, E., and Scott, M.L. (2001). An essential role for BAFF in the normal development of B cells through a BCMA-independent pathway. Science *293*, 2111-2114.

Schmidt, C.S., Liu, J., Zhang, T., Song, H.Y., Sandusky, G., Mintze, K., Benschop, R.J., Glasebrook, A., Yang, D.D., and Na, S. (2003). Enhanced B cell expansion, survival, and humoral responses by targeting death receptor 6. J Exp Med *197*, 51-62.

Schram, B.R., Tze, L.E., Ramsey, L.B., Liu, J., Najera, L., Vegoe, A.L., Hardy, R.R., Hippen, K.L., Farrar, M.A., and Behrens, T.W. (2008). B Cell Receptor Basal Signaling Regulates Antigen-Induced Ig Light Chain Rearrangements. The Journal of Immunology *180*, 4728-4741.

Schweitzer, B.L., and DeKoter, R.P. (2004). Analysis of gene expression and Ig transcription in PU.1/Spi-B-deficient progenitor B cell lines. The Journal of Immunology *172*, 144-154.

Serra, F., Baù, D., Goodstadt, M., Castillo, D., Filion, G.J., and Marti-Renom, M.A. (2017). Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors. PLoS Comput Biol *13*, e1005665.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biology *16*, 259.

Setz, C.S., Hug, E., Khadour, A., Abdelrasoul, H., Bilal, M., Hobeika, E., and Jumaa, H. (2018). PI3K-Mediated Blimp-1 Activation Controls B Cell Selection and Homeostasis. CellReports *24*, 391-405.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.

Shimizu, T., Mundt, C., Licence, S., Melchers, F., and Mårtensson, I.L. (2002). VpreB1/VpreB2/lambda 5 triple-deficient mice show impaired B cell development but functional allelic exclusion of the IgH locus. J Immunol *168*, 6286-6293.

Song, M.S., Salmena, L., and Pandolfi, P.P. (2012). The functions and regulation of the PTEN tumour suppressor. Nature Reviews Molecular Cell Biology *13*, 283-296.

Stanford, S.M., and Bottini, N. (2014). PTPN22: the archetypal non-HLA autoimmunity gene. Nat Rev Rheumatol *10*, 602-611.

Stansfield, J.C., Cresswell, K.G., Vladimirov, V.I., and Dozmorov, M.G. (2018). HiCcompare: an R-package for joint normalization and comparison of HI-C datasets. BMC Bioinformatics *19*, 279.

Su, T.T., and Rawlings, D.J. (2002). Transitional B lymphocyte subsets operate as distinct checkpoints in murine splenic B cell development. J Immunol *168*, 2101-2110.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences *102*, 15545-15550.

Sudo, T., Ito, M., Ogawa, Y., Iizuka, M., Kodama, H., Kunisada, T., Hayashi, S., Ogawa, M., Sakai, K., and Nishikawa, S. (1989). Interleukin 7 production and function in stromal cell-dependent B cell development. J Exp Med *170*, 333-338.

Sukup-Jackson, M.R., Kiraly, O., Kay, J.E., Na, L., Rowland, E.A., Winther, K.E., Chow, D.N., Kimoto, T., Matsuguchi, T., Jonnalagadda, V.S.*, et al.* (2014). Rosa26-GFP direct repeat (RaDR-GFP) mice reveal tissue- and age-dependence of homologous recombination in mammals in vivo. PLoS Genet *10*, e1004299.

Suzuki, H., Terauchi, Y., Fujiwara, M., Aizawa, S., Yazaki, Y., Kadowaki, T., and Koyasu, S. (1999). Xid-like immunodeficiency in mice with disruption of the p85alpha subunit of phosphoinositide 3-kinase. Science *283*, 390-392.

Takeda, S., Sonoda, E., and Arakawa, H. (1996). The kappa:lambda ratio of immature B cells. Immunol Today *17*, 200-201.

Takemori, T., and Rajewsky, K. (1984). Specificity, duration and mechanism of idiotype suppression induced by neonatal injection of monoclonal anti-idiotope antibodies into mice. European Journal of Immunology *14*, 656-667.

Tallmadge, R.L., Tseng, C.T., and Felippe, M.J.B. (2014). Diversity of immunoglobulin lambda light chain gene usage over developmental stages in the horse. Dev Comp Immunol *46*, 171-179.

Teng, G., Maman, Y., Resch, W., Kim, M., Yamane, A., Qian, J., Kieffer-Kwon, K.R., Mandal, M., Ji, Y., Meffre, E.*, et al.* (2015). RAG Represents a Widespread Threat to the Lymphocyte Genome. Cell *162*, 751-765.

Thal, M.A., Carvalho, T.L., He, T., Kim, H.G., Gao, H., Hagman, J., and Klug, C.A. (2009). Ebf1-mediated down-regulation of Id2 and Id3 is essential for specification of the B cell lineage. Proc Natl Acad Sci U S A *106*, 552-557.

Tiegs, S.L., Russell, D.M., and Nemazee, D. (1993). Receptor editing in self-reactive bone marrow B cells. Journal of Experimental Medicine *177*, 1009-1020.

Tikhonova, A.N., Dolgalev, I., Hu, H., Sivaraj, K.K., Hoxha, E., Cuesta-Domínguez, Á., Pinho, S., Akhmetzyanova, I., Gao, J., Witkowski, M.*, et al.* (2019). The bone marrow microenvironment at single-cell resolution. Nature *569*, 222-228.

Tokoyoda, K., Egawa, T., Sugiyama, T., Choi, B.-I., and Nagasawa, T. (2004). Cellular niches controlling B lymphocyte behavior within bone marrow during development. Immunity *20*, 707-718.

Tonegawa, S. (1983). Somatic generation of antibody diversity. Nature *302*, 575-581.

Tze, L.E., Baness, E.A., Hippen, K.L., and Behrens, T.W. (2000). Ig light chain receptor editing in anergic B cells. The Journal of Immunology *165*, 6796-6802.

Tze, L.E., Hippen, K.L., and Behrens, T.W. (2003). Late Immature B Cells (IgMhighIgDneg) Undergo a Light Chain Receptor Editing Response to Soluble Self-Antigen. The Journal of Immunology *171*, 678-682.

Tze, L.E., Schram, B.R., Lam, K.-P., Hogquist, K.A., Hippen, K.L., Liu, J., Shinton, S.A., Otipoby, K.L., Rodine, P.R., Vegoe, A.L.*, et al.* (2005). Basal immunoglobulin signaling actively maintains developmental stage in immature B cells. PLoS biology *3*, e82.

Urbanczyk, S., Stein, M., Schuh, W., Jäck, H.-M., Mougiakakos, D., and Mielenz, D. (2018). Regulation of Energy Metabolism during Early B Lymphocyte Development. International Journal of Molecular Sciences *19*, 2192.

Van Epps, H.L. (2006). Bringing order to early B cell chaos. J Exp Med *203*, 1389.

van Loo, P.F., Dingjan, G.M., Maas, A., and Hendriks, R.W. (2007). Surrogate-light-chain silencing is not critical for the limitation of pre-B cell expansion but is for the termination of constitutive signaling. Immunity *27*, 468-480.

Vander Heiden, J.A., Stathopoulos, P., Zhou, J.Q., Chen, L., Gilbert, T.J., Bolen, C.R., Barohn, R.J., Dimachkie, M.M., Ciafaloni, E., Broering, T.J.*, et al.* (2017). Dysregulation of B Cell Repertoire Formation in Myasthenia Gravis Patients Revealed through Deep Sequencing. The Journal of Immunology *198*, 1460-1473.

Vela, J.L., Ait-Azzouzene, D., Duong, B.H., Ota, T., and Nemazee, D. (2008). Rearrangement of mouse immunoglobulin kappa deleting element recombining sequence promotes immune tolerance and lambda B cell production. Immunity *28*, 161-170.

Verkoczy, L., Ait-Azzouzene, D., Skog, P., Märtensson, A., Lang, J., Duong, B., and Nemazee, D. (2005). A role for nuclear factor kappa B/rel transcription factors in the regulation of the recombinase activator genes. Immunity *22*, 519-531.

Verkoczy, L., Duong, B., Skog, P., Aït-Azzouzene, D., Puri, K., Vela, J.L., and Nemazee, D. (2007). Basal B cell receptor-directed phosphatidylinositol 3-kinase signaling turns off RAGs and promotes B cell-positive selection. J Immunol *178*, 6332-6341.

Verkoczy, L.K., Mårtensson, A.S., and Nemazee, D. (2004). The scope of receptor editing and its association with autoimmunity. Curr Opin Immunol *16*, 808-814.

Vettermann, C., Herrmann, K., Albert, C., Roth, E., Bösl, M.R., and Jäck, H.M. (2008). A unique role for the lambda5 nonimmunoglobulin tail in early B lymphocyte development. J Immunol *181*, 3232-3242.

Wardemann, H., Yurasov, S., Schaefer, A., Young, J.W., Meffre, E., and Nussenzweig, M.C. (2003). Predominant Autoantibody Production by Early Human B Cell Precursors. Science *301*, 1374-1377.

Watson, C.T., Steinberg, K.M., Graves, T.A., Warren, R.L., Malig, M., Schein, J., Wilson, R.K., Holt, R.A., Eichler, E.E., and Breden, F. (2015). Sequencing of the human IG light chain loci from a hydatidiform mole BAC library reveals locus-specific signatures of genetic diversity. Genes Immun *16*, 24-34.

Wilker, P.R., Kohyama, M., Sandau, M.M., Albring, J.C., Nakagawa, O., Schwarz, J.J., and Murphy, K.M. (2008). Transcription factor Mef2c is required for B cell proliferation and survival after antigen receptor stimulation. Nature Publishing Group *9*, 603-612.

Williams, A.B., and Schumacher, B. (2016). p53 in the DNA-Damage-Repair Process. Cold Spring Harb Perspect Med *6*, a026070.

Winkler, T.H., and Mårtensson, I.-L. (2018). The Role of the Pre-B Cell Receptor in B Cell Development, Repertoire Selection, and Tolerance. Frontiers in Immunology *9*, 2423.

Wolff, J., Bhardwaj, V., Nothjunge, S., Richard, G., Renschler, G., Gilsbach, R., Manke, T., Backofen, R., Ramírez, F., and Grüning, B.A. (2018). Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. Nucleic Acids Res *46*, W11-w16.

Wolff, J., Rabbani, L., Gilsbach, R., Richard, G., Manke, T., Backofen, R., and Grüning, B.A. (2020). Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and single-cell Hi-C data analysis, quality control and visualization. Nucleic Acids Res *48*, W177-W184.

Wood, C., and Tonegawa, S. (1983). Diversity and joining segments of mouse immunoglobulin heavy chain genes are closely linked and in the same orientation: implications for the joining mechanism. Proc Natl Acad Sci U S A *80*, 3030-3034.

Xu, S., Lee, K.G., Huo, J., Kurosaki, T., and Lam, K.P. (2007). Combined deficiencies in Bruton tyrosine kinase and phospholipase Cgamma2 arrest B-cell development at a pre-BCR+ stage. Blood *109*, 3377-3384.

Yachimovich-Cohen, N., Fischel, R., Bachar, N., Yarkoni, Y., and Eilat, D. (2003). Autoimmune NZB/NZW F1 mice utilize B cell receptor editing for generating high-affinity anti-dsDNA autoantibodies from low-affinity precursors. Eur J Immunol *33*, 2469-2478.

Yelton, D.E., Desaymard, C., and Scharff, M.D. (1981). Use of monoclonal anti-mouse immunoglobulin to detect mouse antibodies. Hybridoma *1*, 5-11.

Yu, M., and Ren, B. (2017). The Three-Dimensional Organization of Mammalian Genomes. Annual Review of Cell and Developmental Biology *33*, 265-289.

Yu, W., Nagaoka, H., Jankovic, M., Misulovin, Z., Suh, H., Rolink, A., Melchers, F., Meffre, E., and Nussenzweig, M.C. (1999). Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. Nature *400*, 682-687.

Yurasov, S., Wardemann, H., Hammersen, J., Tsuiji, M., Meffre, E., Pascual, V., and Nussenzweig, M.C. (2005). Defective B cell tolerance checkpoints in systemic lupus erythematosus. J Exp Med *201*, 703-711.

Zandi, S., Bryder, D., and Sigvardsson, M. (2010). Load and lock: the molecular mechanisms of B-lymphocyte commitment. Immunol Rev *238*, 47-62.

Zetterblad, J., Qian, H., Zandi, S., Mansson, R., Lagergren, A., Hansson, F., Bryder, D., Paulsson, N., and Sigvardsson, M. (2010). Genomics based analysis of interactions between developing B-lymphocytes and stromal cells reveal complex interactions and two-way communication. BMC genomics *11*, 108-115.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun *10*, 1523.

Zou, Y.R., Takeda, S., journal, K.R.T.E., and 1993 (1993). Gene targeting in the Ig kappa locus: efficient generation of lambda chain-expressing B cells, independent of gene rearrangements in Ig kappa. embopress.org *12*, 811-820.