Vβ Recombination Signal Sequences Mediate Monoallelic And Monogenic Tcrβ Gene Assembly: Implications For The Tcrβ Repertoire And Allelic Exclusion In Health And Disease

Glendon Shou-Shi Wu
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Allergy and Immunology Commons, Immunology and Infectious Disease Commons, Medical Immunology Commons, and the Molecular Biology Commons

Recommended Citation
Wu, Glendon Shou-Shi, "Vβ Recombination Signal Sequences Mediate Monoallelic And Monogenic Tcrβ Gene Assembly: Implications For The Tcrβ Repertoire And Allelic Exclusion In Health And Disease" (2020). Publicly Accessible Penn Dissertations. 4048.
https://repository.upenn.edu/edissertations/4048

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/4048
For more information, please contact repository@pobox.upenn.edu.
Abstract
Monoallelic expression of antigen receptor (AgR) genes is assumed to be critical for the proper
development and function of T and B lymphocytes. AgR loci are composed of variable (V), joining (J), and
sometimes diversity (D) gene segments that must be cleaved and assembled by RAG-endonuclease
mediated V(D)J recombination to form a functional gene. TCRβ, IgH, and Igk allelic exclusion is achieved
by: assembly of an in-frame gene on one allele, transient inhibition of V(D)J recombination signaled by
RAG DNA double stranded breaks (DSBs), and the resulting protein sending feedback signals to
permanently inhibit V-to-(D)J recombination of the non-functional allele. Any definitive mechanisms that
promote asynchronous AgR gene assembly in developing lymphocytes remain unknown. V segment
recombination signal sequence (RSSs) of Tcrb and Igh loci that target RAG have been proposed to
render V recombination inefficient and mediate monoallelic V-to-DJ recombination. To test the role of Vβ
RSSs in promoting Tcrb allelic exclusion, I created and studied mice harboring particular Vβ RSS
replacements with the stronger 3'Dβ1 RSS. I demonstrate a substantial role for weak Vβ RSSs in limiting:
Vβ recombination frequency, biallelic Vβ-to-DJβ recombination before the onset of transient and
permanent feedback mechanisms, biallelic TCRβ expression, and unexpectedly dual TCRβ chain
expression from a single Tcrb allele. These data indicate that weak Vβ RSSs limit Vβ recombination to
promote monogenic Tcrb assembly within the time window before feedback inhibition halts Vβ
rearrangements. I also establish a role for ATM, a key factor in the DNA damage response that promotes
DNA repair and mediates transient feedback inhibition, in guarding against the formation of nonfunctional
deletions for V segments that rearrange by inversion and cooperating with weak Vβ RSSs to impose
TCRβ allelic exclusion. AgR allelic exclusion is most stringently applied to Tcrb and Igh, whose assembly,
expression, and signaling through pre-AgR complexes drive cellular proliferation. As I am able to drive
biallelic recombination of Tcrb genes, I introduce preliminary data that supports the model that
mechanisms directing monoallelic induction of RAG DSBs evolved at Tcrb and Igh loci in part to suppress
DSBs from entering S phase and forming oncogenic AgR translocations. Finally, as the Vβ RSS
replacements result in dramatic shifts in the Vβ repertoire, I present data that suggests the composition
of the TCRβ repertoire may affect antigen-specific immune responses by defining the frequency of naïve
antigen-reactive cells. Together, these studies reveal mechanisms in developing thymocytes that
promote: monoallelic Tcrb assembly and allelic exclusion, repertoire diversity, and perhaps T cell immune
responses and genome stability.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Immunology

First Advisor
Craig H. Bassing
Keywords
Allelic Exclusion, Antigen Receptor, Lymphocyte Development, T Cell Receptor, V(D)J Recombination

Subject Categories
Allergy and Immunology | Immunology and Infectious Disease | Medical Immunology | Molecular Biology

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/4048
Vβ RECOMBINATION SIGNAL SEQUENCES MEDIATE MONOALLELIC AND MONOGENIC
TCRβ GENE ASSEMBLY: IMPLICATIONS FOR THE TCRβ REPERTOIRE AND ALLELIC
EXCLUSION IN HEALTH AND DISEASE

Glendon Shou-Shi Wu

A DISSERTATION

in

Immunology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2021

Supervisor of Dissertation
Craig H. Bassing, Ph.D.

Graduate Group Chairperson
David M. Allman, Ph.D.

Associate Professor of Pathology and
Laboratory Medicine
Taku Kambayashi, M.D., Ph.D., Associate Professor of Pathology and Laboratory Medicine

Dissertation Committee
Laurence Eisenlohr, V.M.D., Ph.D. (Chair), Professor of Pathology and Laboratory Medicine

David M. Allman, Ph.D., Professor of Pathology and Laboratory Medicine

Ranjan Sen, Ph.D., Senior Investigator of the Laboratory of Molecular Biology and Immunology
Vβ RECOMBINATION SIGNAL SEQUENCES MEDIATE MONOALLELIC AND MONOGENIC TCRβ GENE ASSEMBLY: IMPLICATIONS FOR THE TCRβ REPERTOIRE AND ALLELIC EXCLUSION IN HEALTH AND DISEASE

COPYRIGHT
2021
Glendon Shou-Shi Wu

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License

To view a copy of this license, visit

https://creativecommons.org/licenses/by-nc-sa/3.0/us/
DEDICATION

To my parents, George and Teresa, whose immense sacrifices allowed their children to pursue their own dreams and lead prosperous lives.
ACKNOWLEDGMENTS

The phrase “it takes a village” comes to mind when I look back at my journey up to and through graduate school. It may seem gratuitous to thank everyone here, but when else does one get an opportunity to reflect on their academic journey and put into words their appreciation? First, I wish to thank the late Drs. Alan S. Rabson and Ruth Kirschstein, who were frequent clients of my parents’ humble Chinese restaurant in Bethesda, MD. It is because of their generosity that I was connected with and placed into Dr. Lance Liotta’s proteomics lab at the NIH. My first mentors, Dr. Virginia (Ginny) Espina and Dr. Liotta: thank you for taking me on, showing me immense patience, and teaching me my first lessons in biomedical research.

Thank you to Ms. Anika Green, who showed me grace and admitted me into Stanford’s Summer Research Program after having been initially rejected. That summer internship would go on to open several doors of opportunity and changed my professional trajectory. The significance in reversing your decision was never lost on me and I am forever grateful.

To Drs. Jon Zmuda, Kendall Carey, and Xianzhi Zhou from my time at MedImmune: thank you for pushing me to go back to graduate school. Thank you for breaking my cycle of complacency and seeing that I was capable of more.

An immense amount of gratitude must be paid to my mentor Dr. Craig Bassing. Craig is genuinely a very caring and dedicated mentor, enjoys teaching, and is an unrelenting advocate for his students. Thank you for your unwavering support, your positive mental attitude, all of your help in honing my writing skills, and pushing me to be a more meticulous and rigorous scientist.

There are several people in the Bassing Lab, past and present, which I would like to acknowledge. Katherine Yang-lott, Megan Fisher, Rebecca Glynn, Brittney Allyn, and Charline Miot: thank you all for your friendship, mentoring, lunch- and dinner- dates, laugh-sessions… You all have made the lab such a fun and rich environment that coming to work was a joy. Much gratitude to Morgann Klink for performing all of the Southern blots in Chapter 2 and Katharina Hayer for help analyzing the Adaptive datasets. A special shout-out to Kyutae Lee: you were not only a great friend, but your intellect, drive, and optimism pushed me and everyone else in the lab
to be better. I also wish to thank my lab friends and neighbors, Claudia Lanauze and Colleen Harrington, for adding to the laboratory mirth.

Thank you to Adele Harman for injecting, electroporating, culturing, and implanting all of those embryos so that we could have mice to study.

My cohort, the IGG class of 2014, was quite unusual: one, for its size, and another, for it being composed solely of men, earning the moniker “The Brohort.” Thomas Burn (my “RAG-mate”), Tanner Robertson, Brenal Singh, and Omar Khan: thank you for being my second family. You are all amazing scientists and represent some of my most cherished friendships. Thank you as well to the 2015 matriculating year, particularly Jen Wu, as both of our classes became very close. You all have made the last six years some of the best years of my life.

To my friends, Nik Santistevan and Suna Lumeh: thank you for being an endless source of support and laughter during my time in graduate school.

Thank you to my Thesis Committee: David Allman, Ike Eisenlohr, Taku Kambayashi, and Ranjan Sen for your patience and comments on my work.

Thank you to the Penn Immunology Community for being such a welcoming and rigorous training environment. I enjoyed my time in graduate school immensely and am going to miss the atmosphere created by Penn’s faculty and student body. In particular, I will miss Mary Taylor as she was the glue of the IGG; I will miss your smile and your booming laugh.

Finally, this journey would not have been possible without the endless love and support of my family and fiancé. My parents came to this country for the promise of better opportunities and worked tirelessly to support their family of four children. Thank you, Mom and Dad, for all of the life lessons and teaching us the value of hard work, perseverance, and accountability; these teachings were instrumental for my journey throughout graduate school and are central to my approach to life. To my siblings and all of my nieces and nephews: thank you for believing in me, encouraging me to keep pushing, and to not take myself too seriously. And finally, to my fiancé, Shawn O’Reilly: thank you for being my shoulder to lean on in hard times and my person to celebrate the good. I cannot wait to see what our next adventure may have in store.
ABSTRACT

Vβ RECOMBINATION SIGNAL SEQUENCES MEDIATE MONOALLELIC AND MONOGENIC TCRβ GENE ASSEMBLY: IMPLICATIONS FOR THE TCRβ REPERTOIRE AND ALLELIC EXCLUSION IN HEALTH AND DISEASE

Glendon S. Wu
Craig H. Bassing

Monoallelic expression of antigen receptor (AgR) genes is assumed to be critical for the proper development and function of T and B lymphocytes. AgR loci are composed of variable (V), joining (J), and sometimes diversity (D) gene segments that must be cleaved and assembled by RAG-endonuclease mediated V(D)J recombination to form a functional gene. TCRβ, IgH, and Igκ allelic exclusion is achieved by: assembly of an in-frame gene on one allele, transient inhibition of V(D)J recombination signaled by RAG DNA double stranded breaks (DSBs), and the resulting protein sending feedback signals to permanently inhibit V-to-(D)J recombination of the non-functional allele. Any definitive mechanisms that promote asynchronous AgR gene assembly in developing lymphocytes remain unknown. V segment recombination signal sequences (RSSs) of Tcrb and Igh loci that target RAG have been proposed to render V recombination inefficient and mediate monoallelic V-to-DJ recombination. To test the role of Vβ RSSs in promoting Tcrb allelic exclusion, I created and studied mice harboring particular Vβ RSS replacements with the stronger 3'Dβ1 RSS. I demonstrate a substantial role for weak Vβ RSSs in limiting: Vβ recombination frequency, biallelic Vβ-to-DJβ recombination before the onset of transient and permanent feedback mechanisms, biallelic TCRβ expression, and unexpectedly dual TCRβ chain expression from a single Tcrb allele. These data indicate that weak Vβ RSSs limit Vβ recombination to promote monogenic Tcrb assembly within the time window before feedback inhibition halts Vβ rearrangements. I also establish a role for ATM, a key factor in the DNA damage response that
promotes DNA repair and mediates transient feedback inhibition, in guarding against the formation of nonfunctional deletions for V segments that rearrange by inversion and cooperating with weak Vβ RSSs to impose TCRβ allelic exclusion. AgR allelic exclusion is most stringently applied to Tcrb and Igh, whose assembly, expression, and signaling through pre-AgR complexes drive cellular proliferation. As I am able to drive biallelic recombination of Tcrb genes, I introduce preliminary data that supports the model that mechanisms directing monoallelic induction of RAG DSBs evolved at Tcrb and Igh loci in part to suppress DSBs from entering S phase and forming oncogenic AgR translocations. Finally, as the Vβ RSS replacements result in dramatic shifts in the Vβ repertoire, I present data that suggests the composition of the TCRβ repertoire may affect antigen-specific immune responses by defining the frequency of naïve antigen-reactive cells. Together, these studies reveal mechanisms in developing thymocytes that promote: monoallelic Tcrb assembly and allelic exclusion, repertoire diversity, and perhaps T cell immune responses and genome stability.
### TABLE OF CONTENTS

DEDICATION.................................................................................................................................................. III
ACKNOWLEDGMENTS....................................................................................................................................... IV
ABSTRACT........................................................................................................................................................ VI
LIST OF TABLES .................................................................................................................................................. XI
LIST OF FIGURES ............................................................................................................................................. XI

**CHAPTER 1: INTRODUCTION** ...................................................................................................................... 1

- The Adaptive Immune Systems of Jawed-Vertebrates .................................................................................. 1
- Mammalian Antigen Receptors, Their Genes, and V(D)J Recombination ............................................................ 4
- Allelic Exclusion of Antigen Receptor Genes ................................................................................................... 5
- Antigen Receptor Gene Assembly and Lymphocyte Development ................................................................. 6
- Models Contributing to Monallelic Antigen Receptor Assembly and Expression ............................................. 8
- RSSs Orchestrate V(D)J Recombination and Exhibit Varying Levels of Activity .............................................. 10
- Why Antigen Receptor Allelic Exclusion? ......................................................................................................... 11
- Thesis Structure ............................................................................................................................................. 14

**CHAPTER 2: POOR QUALITY Vβ RECOMBINATION SIGNAL SEQUENCES STOCHastically ENFORCE TCRβ**

ALLElic EXCLUSION ..................................................................................................................................... 20

- Abstract .................................................................................................................................................... 20
- Introduction ............................................................................................................................................... 20
- Results ...................................................................................................................................................... 23
- Discussion ................................................................................................................................................ 33

**CHAPTER 3: INEFFICIENT V(D)J RECOMBINATION UNDERLIES MONOGENIC T CELL RECEPTOR β**

EXPRESSION .................................................................................................................................................. 55

- Abstract .................................................................................................................................................... 55
- Introduction ............................................................................................................................................... 55
- Results ...................................................................................................................................................... 57
- Discussion ................................................................................................................................................ 59

**CHAPTER 4: WEAK Vβ RECOMBINATION SIGNAL SEQUENCES COOPERATE WITH THE DNA DAMAGE**

RESPONSE ATM Kinase to Ensure Monoallelic Expression and Broad Repertoire of TCRβ Proteins ............... 64

- Abstract .................................................................................................................................................... 64
- Introduction ............................................................................................................................................... 65
- Results ...................................................................................................................................................... 66
- Discussion ................................................................................................................................................ 68

**CHAPTER 5: A BROAD TCRβ REPertoire DEFINES THE FREQUENCY OF NAÏVE ANTIGEN-SPECIFIC CD8 T**

CELLS ........................................................................................................................................................... 73

- Abstract .................................................................................................................................................... 73
- Introduction ............................................................................................................................................... 73
- Results ...................................................................................................................................................... 76
- Discussion ................................................................................................................................................ 78
CHAPTER 6: DISCUSSION ................................................................. 81

OVERVIEW ...................................................................................... 81

POOR QUALITY Vβ RSSs PROVIDE A STOCHASTIC MECHANISM FOR LIMITING Vβ RECOMBINATION BETWEEN ALLELES BEFORE FEEDBACK INHIBITION ........................................................... 82

POOR QUALITY Vβ RSSs STOCHASTICALLY RESTRAIN Vβ REARRANGEMENTS TO PROMOTE MONOGENIC TCRB ASSEMBLY 84

ROLES FOR ATM IN ALLELIC EXCLUSION AND REPAIR OF INVERSIONAL REARRANGEMENTS .................................................. 86

A WORKING MODEL OF THE MECHANISMS CONTRIBUTING TO MONOALLELIC AGR GENE ASSEMBLY ................................. 87

MODELS CONTRIBUTING TO A DIVERSE Vβ REPertoire .......................................................... 89

LONG-RANGE V RECOMBINATION: RAG SCANNING VERSUS PROXIMITY-BASED DIFFUSION ......................................................... 91

PHYSIOLOGICAL ROLES FOR ALLELIC EXCLUSION IN VERTEBRATE IMMUNE SYSTEMS ............................................................... 92

MONOALLELIC V RECOMBINATION AT IGH AND TCRB LOCI MAY LIMIT ONCOGENIC TRANSLOCATIONS .................................... 95

FINAL THOUGHTS ........................................................................... 98

APPENDIX: MATERIALS AND METHODS .................................................. 109

BIBLIOGRAPHY ............................................................................... 117
LIST OF TABLES

CHAPTER 2:
Table 2.1. Analysis of Tcrb rearrangements in αβ T cell hybridomas .............................................. 51
Table 2.2. Sequence analysis of the V2 and V31 rearrangements on opposite alleles in V2\(^{\beta}\)/V31\(^{\alpha}\)R hybridomas .................................................................................................................. 52
Table 2.3. Analysis of Tcrb rearrangements in Tcrb\(^{\beta}\) T cell hybridomas ............................................. 53
Table 2.4. Recombination information content (RIC) scores of RSSs in this study ............................... 54

APPENDIX:
Appendix Table 1. Primers used in this study .................................................................................... 114
Appendix Table 2. Key Reagents Table ............................................................................................... 116
LIST OF FIGURES

CHAPTER 1:
Figure 1.1. Schematic of V(D)J recombination ................................................................. 15
Figure 1.2. Models of monoallelic versus biallelic antigen receptor gene assembly ............... 16
Figure 1.3. Models contributing to antigen receptor allelic exclusion ................................... 17
Figure 1.4. Potential mechanisms that promote monoallelic initiation of V-to-(D)J recombination between alleles .................................................................................................................. 18
Figure 1.5. Model for Tcrb translocations as a result of biallelic Tcrb gene assembly ........... 19

CHAPTER 2:
Figure 2.1. Normal gross αβ T cell development in V2 and V31 RSS replacement mice .......... 39
Figure 2.2. The 3'Dβ1 RSS replacement increases V2 and V31 recombination in DN3 cells ...... 41
Figure 2.3. Increased utilization of 3'Dβ1 RSS-replaced Vβ segments on αβ T cells ............... 42
Figure 2.4. Peripheral αβ T cells exhibit similar shifts in the Vβ repertoire in RSS replacement mice ................................................................................................................................. 43
Figure 2.5. Vβ RSS replacement alleles compete with normal Tcrb alleles for usage in the TCRβ repertoire ............................................................................................................................... 44
Figure 2.6. 3'Dβ1 RSS-replaced Vβ segments increase biallelic Tcrb gene expression ............ 45
Figure 2.7. αβ T cells exhibiting biallelic Tcrb gene expression seed the periphery ................. 47
Figure 2.8. 3'Dβ1 RSS substitutions increase Vβ usage and TCRβ allelic inclusion independent of c-Fos binding ......................................................................................................................... 48
Figure 2.9. 3'Dβ1 RSS substitutions increase Vβ recombination before enforcement of feedback inhibition ........................................................................................................................ 49
Figure 2.10. 3'Dβ1 RSS substitutions permits direct Vβ-to-Jβ recombination and decreases TCRβ diversity ............................................................................................................................ 50

CHAPTER 3:
Figure 3.1. Increased frequency of dual TCRβ+ cells and altered Vβ repertoire in mice with two Vβ RSS replacements on the same allele ........................................................................................................ 61
Figure 3.2. Expression of two different TCRβ chains from the V2R V31R allele .................... 63

CHAPTER 4:
Figure 4.1. The weak V31 RSS and ATM cooperate to prevent non-functional V31 rearrangements and ATM protects the Vκ repertoire ........................................................................................................ 70
Figure 4.2. The poor V1 and V2 RSSs and ATM cooperate to enforce TCRβ allelic exclusion.... 72
CHAPTER 5:
Figure 5.1. The Vβ repertoire does not alter the frequencies of other immune cell subsets nor the CD8 T cell response during LCMV infection ................................................................. 79
Figure 5.2. The Vβ repertoire alters the frequency of gp33-reactive CD8 T cells during LCMV infection ........................................................................................................................................ 80

CHAPTER 6:
Figure 6.1. Working model of factors enforcing TCRβ allelic exclusion ........................................ 103
Figure 6.2. Model of factors contributing to a diverse Vβ repertoire ............................................ 104
Figure 6.3. Possible physiological functions for AgR allelic exclusion ........................................ 105
Figure 6.4. Weak Vβ RSSs may limit oncogenic Tcrb translocations resulting from biallelic Tcrb recombination .............................................................................................................. 107
Figure 6.5. Theoretical frequencies of in-frame and out-of-frame VDJ joins in cells attempting two or more rearrangements assuming synchronous V recombination ........................................... 108
CHAPTER 1: Introduction

The Adaptive Immune Systems of Jawed-Vertebrates

All living organisms have evolved mechanisms to sense and respond to potentially pathogenic microorganisms and their products. Naturally occurring infections and the injuries that promote them are immensely diverse, can occur at any time, and can be established in any host tissue. Immune systems, therefore, are tasked to manage this immense unpredictability and must identify foreign antigens originating from various entities such as viruses, bacteria, and uni- or multi-cellular parasites (Murphy and Weaver, 2017). Yet, not all microbes are infectious or “mean” to do a host harm, and most organisms form naturally occurring symbiotic relationships with commensal microbes that are mutually beneficial (Haque and Haque, 2017). Thus, not only do immune systems have to discern intrinsic (self) from extrinsic (non-self) molecules, but also have to develop tolerance to antigens that are pervasive or innocuous in the environment (Satitsuksanoa et al., 2018). Various immune processes have evolved to address this potentially enormous array of antigens an organism may encounter over its lifetime. Failure to appropriately identify and respond to potentially infectious materials while maintaining self-tolerance is the basis for various disease states including immunodeficiency, autoimmunity, allergy, and chronic inflammation (Murphy and Weaver, 2017).

The innate immune system is evolutionarily ancient and encompasses general and broad strategies to deal with possible pathogens (Janeway and Medzhitov, 2002; Lemaitre et al., 1996; Murphy and Weaver, 2017). Found in unicellular life and throughout the evolutionary tree, examples of these mechanisms include antimicrobial peptides, multi-component inflammasomes, and germline encoded innate receptors on the cell surface (Broz and Dixit, 2016; Janeway and Medzhitov, 2002; Mahlapuu et al., 2016). Innate immune systems can detect factors that are associated with cellular damage (such as extracellular ATP, histone proteins, and mitochondrial DNA) and infection (Janeway and Medzhitov, 2002; Roh and Sohn, 2018). During an infection, innate immune sensors engage conserved components that are common to certain groups of
microorganisms. A classic example is the detection of bacterial lipopolysaccharide that studs the outer membranes of Gram-negative bacteria by Toll-like receptor (TLR) 4 (Hoshino et al., 1999). Some of these strategies do not necessarily distinguish between a colonizing pathogen from a commensal, but these blanket tactics detect homeostatic perturbations and engage effector programs that slow the initiation of infection. Innate immunity is instrumental for organismal survival, but the breadth of possible antigens the innate immune system can detect is limited due to the germline coding of these sensors and receptors. Over 500 million years ago, vertebrate animals have evolved and rely heavily upon anticipatory adaptive immune systems (Cooper and Alder, 2006; Flajnik, 2018). The basis of these immune systems is the expression of a broadly diverse repertoire of clonotypic antigen receptors (AgRs) by specialized immune cells called lymphocytes. Such a system greatly expands the potential antigenic universe that a host can recognize without a priori knowledge of what may infect it in the future.

Among vertebrates, two different AgR systems evolved that are distinct both in the mechanism that generates receptor diversity and in the structure of the receptors themselves (Cooper and Alder, 2006; Flajnik, 2018). Agnathans are jawless vertebrates that are represented by extant species such as lamprey and hagfish. These animals express variable lymphocyte receptors (VLRs) that are composed of leucine-rich repeat (LRR) domains and are thought to be assembled by a gene conversion mechanism requiring cytidine deaminases (Cooper and Alder, 2006; Flajnik, 2018). In contrast, the jawed vertebrates (gnathostomes) express immunoglobulin superfamily (IgSF) receptors that are assembled by the recombination activating gene (RAG) products RAG1 and RAG2 (Cooper and Alder, 2006; Flajnik, 2018; Hsu, 2009). Gnathostomes do express an immunologically relevant Activation Induced cytidine Deaminase (AID) that works at AgR loci, however in most animals AID participates in somatic hypermutation of the variable region of immunoglobulin genes, and in higher vertebrates promotes antibody class switch recombination (Flajnik, 2018). While both receptor systems are distinct, these highly specific AgRs are expressed by cells derived from hematopoietic precursors, are clonally expressed, have membrane-bound and secreted forms, and the lymphocyte populations expressing them...
have functionally analogous B-like, $\alpha \beta$ T-like, and $\gamma \delta$ T-like populations (Flajnik, 2018). Although separated by >500 million years of evolution, immune pressures experienced by both jawless and jawed vertebrates selected for differing strategies to achieve the same goal: increase AgR genetic diversity to expand the breadth of molecular shapes that can be recognized. These pressures thus established the LRR VLR system for agnathans and IgSF system for gnathostomes (Cooper and Alder, 2006; Flajnik, 2018; Hsu, 2009).

All jawed-vertebrates assemble AgR genes by RAG1/2 (RAG) endonuclease-mediated recombination of AgR loci. However, across vertebrate taxa, AgRs and their genes are immensely diverse in the structure of the receptor, the loci that encode them, the numbers and organization of loci, and the types of AgR genes. For example, immunoglobulins (Igs) are B cell effector molecules that are composed of two Ig heavy (H) chains and two Ig light (L) chains, and they perform a variety of functions, including antigen neutralization and opsonization. Igs are typified by the constant regions that make up part of the IgH chain. While all vertebrate B cells possess genes to express IgM and IgD, taxon-specific variation is also observed: amphibians express IgX and IgY, teleost (bony) fish express IgT, and cartilaginous fish (such as sharks and rays) express IgW and heavy-chain only Ig new antigen receptor (IgNAR) (Flajnik, 2018; Hsu, 2009). Some of these immunoglobulin variants (such as IgX, IgY, and IgT) are functional paralogs to IgA or IgG, but in some instances, no paralogs exist, as is the case for IgNARs (Flajnik, 2018). At the level of gene organization, cartilaginous fish represent the most evolutionarily ancient taxa amongst gnathostomes and their IgH loci are organized in “cluster configurations,” meaning that each IgH constant region possesses their own dedicated V, D, and J gene segments (Pettinello and Dooley, 2014). In contrast, the IgH locus in mammals (as are most other mammalian AgR loci) are in a configuration where all constant region genes share the same pool of V, D, and J segments (Pettinello and Dooley, 2014). Thus, for simplicity and for the purpose of this thesis, mammalian AgR loci, particularly those of the mouse, will be considered as a model for other RAG-based vertebrate immune systems.
Mammalian Antigen Receptors, their Genes, and V(D)J Recombination

Mammalian AgR genes are assembled by the process of V(D)J recombination in developing T and B lymphocytes (Figure 1.1). Four T cell receptor (TCR) loci (Tcra and Tcrb, or Tcrg and Tcrd) encode αβ or γδ TCRs, respectively, and three Ig loci (Igh and Igk or Igl) encode κ+ or λ+ B cell receptors (BCRs). In the germline, TCR and Ig loci are comprised of noncontiguous variable (V), sometimes diversity (D), and joining (J) gene segments that upon assembly form the second variable region exons of AgR genes (Figure 1.1). V(D)J exons are then transcribed with constant (C) region exons, spliced, and translated to generate one chain of an AgR (Figure 1.1). The AgR variable region dictates antigen specificity and is the site of sequence variation. The AgR variable domain is typified by structural framework regions and three “hypervariable” complementarity determining regions (CDRs) that form the antigen-binding site; CDR1 and CDR2 are encoded by the V segment, and CDR3 spans the junctions of the V(D)J joins and is thus the most variable of the CDRs (Dondelinger et al., 2018; Hughes et al., 2003; Kabat and Wu, 1970; Nikolich-Zugich et al., 2004; Rock et al., 1994).

The V(D)J recombination reaction requires the RAG1/2 (RAG) endonuclease, recombination signal sequences (RSSs), and non-homologous end joining (NHEJ) proteins (Schatz and Ji, 2011). RAG targets RSSs that flank all functional V, D, and J segments (Figure 1.1). An RSS is a semiconserved genetic sequence composed of a heptamer and a nonamer sequence separated by 12- or 23- base pair spacer sequences (Schatz and Ji, 2011). The heptamer is directly adjacent to the gene segment coding flank and is the target of RAG-mediated cleavage (Lewis, 1994; Schatz and Ji, 2011). RAG binds to one RSS and captures another RSS of dissimilar length (the 12/23 rule) to form a synaptic complex (Ru et al., 2015; Schatz and Ji, 2011). RAG introduces DNA double stranded breaks (DSBs) between each RSS and its flanking segment, creating blunted signal ends and hairpin-sealed coding ends (Helmink and Sleckman, 2012). Ubiquitous NHEJ proteins stabilize and process these RAG-mediated DNA DSBs to promote their repair (Helmink and Sleckman, 2012). The hairpin-sealed coding ends are enzymatically opened, and nucleotides are removed and/or added by the template-independent
polymerase TdT to introduce additional diversity at V-D, D-J, and V-J junctions. Repair of the DSBs by NHEJ factors form signal and V(D)J coding joins (Helmink and Sleckman, 2012). The genomic orientation of RSSs dictates whether rearrangements proceed via deletional or inversional recombination (Sollbach and Wu, 1995). When rearranging two RSSs that are pointed toward each other in a convergent manner, recombination proceeds via deletion and the resulting signal join and intervening DNA is lost from the genome (Figure 1.1, D). RSSs that are oriented in the same direction recombine by inversion, and the signal join is retained in the chromosome (Figure 1.1, D) (Sollbach and Wu, 1995). Due to the imprecise repair of coding joins, one-third of V(D)J rearrangements assemble an in-frame exon while two-thirds of rearrangements will assemble a nonfunctional out-of-frame exon (Bassing et al., 2002; Schatz and Swanson, 2011). The culmination of all the possible V, D, and J permutations and the imprecise repair of V(D)J joins produces AgR diversity. It is estimated that V(D)J recombination can generate anywhere from $10^{12}$ to $10^{20}$ unique receptors (Lieber, 1991; Miles et al., 2011; Nikolich-Žugich et al., 2004).

**Allelic Exclusion of Antigen Receptor Genes**

Before the structure of AgR genes was solved, Macfarlane Burnet theorized that the cumulative specificities of AgR repertoires were distributed amongst single cells and that each lymphocyte expressed one receptor (Burnet, 1957; Burnet 1976). This tenet of Burnet’s clonal selection theory was largely satisfied when it was discovered that B cells secrete one type of antibody when challenged by two distinct antigens (Nossal and Lederberg, 1958) and express IgH proteins from one allele (Nossal and Lederberg, 1958; Pernis et al., 1965). This phenomenon of allelic exclusion, where one AgR allele is excluded from expression, is also observed at Tcrb and Igk loci (Brady et al., 2010b). The theme of AgR allelic exclusion, and therefore monoallelic gene expression, is observed in several other biological contexts.

In addition to AgR allelic exclusion in lymphocytes, monoallelic gene expression underlies genomic imprinting and X-chromosome activation in many cell types and tissue-specific allelic exclusion of olfactory neuron receptors. Each of these programs involves an initiation and a
maintenance phase and involves epigenetic-based transcriptional silencing (Khamlichi and Feil, 2018). AgR allelic exclusion involves an additional level of regulation due to the obligate assembly of AgR genes through V(D)J recombination. In the absence of any regulation, the frequent assembly of out-of-frame rearrangements and requirement of AgR protein expression for T and B cell development dictates that biallelic expression of any TCR or Ig gene can occur in at most 20% of lymphocytes (Figure 1.2) (Brady et al., 2010b; Mostoslavsky et al., 2004). However, Tcrb, Igh, and Igk loci exhibit more stringent allelic exclusion that is enforced by the assembly of a functional in-frame V(D)J rearrangement on one allele and subsequent feedback inhibition of V rearrangements on the other allele (Brady et al., 2010b; Jung et al., 2006; Levin-Klein and Bergman, 2014; Mostoslavsky et al., 2004; Outters et al., 2015; Vettermann and Schlissel, 2010). Thus, in ~60% of T or B cells only one V-to-(D)J rearrangement is found at each of these loci, while ~40% of T or B cells exhibit V-to-(D)J recombination on both alleles where typically only one rearrangement is in-frame (Figure 1.2) (Jung et al., 2006).

Antigen Receptor Gene Assembly and Lymphocyte Development

Both T and B lymphocyte lineages develop from common lymphoid progenitors in the bone marrow (Murphy and Weaver, 2017). Some of these progenitors emigrate to the thymus to complete their maturation into αβ or γδ T cells, while some remain and differentiate into B cells (Bhandoola et al., 2007; Hardy et al., 1991). AgR gene expression is a requirement for lymphocyte development, and failure to express an AgR on the cell surface results in cell death (Jung et al., 2006; Mostoslavsky et al., 2004). AgR gene assembly is interdependently regulated with T and B cell development as is the tissue- and developmental stage-specific expression of RAG (Nagaoka et al., 2000).

AgR gene assembly initiates in the T lineage-committed CD4^-CD8^- double-negative (DN) thymocytes and B lineage-committed pro-B cells. DN and pro-B cells activate Rag1/Rag2 transcription and induce the transcription, accessibility, and compaction of Tcrb or Igh loci, respectively (Allyn et al., 2020; Brady et al., 2010b; Shih and Krangel, 2013). Histone
modifications and nucleosome repositioning allow RAG to bind at accessible D and J segments, forming a focal recombination center (RC) where D-to-J recombination occurs (Figure 1.1) (Ji et al., 2010). A single V segment then rearranges to a DJ complex one allele at a time (Figures 1.1 and 1.2) (Brady et al., 2010b; Mostoslavsky et al., 2004; Outters et al., 2015; Vettermann and Schlissel, 2010). This V-to-DJ recombination step likely requires V segment accessibility and locus compaction to place V segments in spatial proximity with the RC (Allyn et al., 2020; Brady et al., 2010b; Shih and Krangel, 2013). DSBs induced in DN thymocytes and pro-B cells repress RAG expression (Fisher et al., 2017), which may transiently inhibit further Tcrb and Igh recombination (Steinel et al., 2014) (Figure 1.3). Cells that assemble an out-of-frame VDJ rearrangement on the first allele can attempt V recombination on the other allele (Brady et al., 2010b; Koralov et al., 2006; Lee and Bassing, 2020; Mostoslavsky et al., 2004; Outters et al., 2015; Vettermann and Schlissel, 2010). Following an in-frame VDJ rearrangement, the resultant TCRβ or IgH proteins pair with preTα or surrogate light chains to form pre-TCRs and pre-BCRs, respectively (Clark et al., 2014; von Boehmer and Fehling, 1997). These pre-AgRs signal proliferation, differentiation of CD4⁺CD8⁺ double-positive (DP) thymocytes or pre-B cells, and permanent feedback inhibition of Vβ or VH recombination through epigenetic changes (Figure 1.3) (von Boehmer and Melchers, 2010).

DP thymocytes and pre-B cells re-express RAG to initiate recombination at Tcra or Igk loci but block further V-to-DJ rearrangements at Tcrb and Igh as a result of permanent feedback inhibition. Feedback inhibition signaled by pre-TCRs and pre-BCRs is likely mediated through the transcriptional silencing of unrearranged V segments and locus de-contraction, which places great distance between V segments and DJ complexes (Brady et al., 2010b; Guo et al., 2011; Majumder et al., 2015; Shih and Krangel, 2013; Skok et al., 2007). In DP thymocytes, Vα-to-Jα rearrangements occur on both Tcra alleles until at least one allele yields a TCRα protein that forms a heterodimeric αβ TCR, which becomes selected for further development. DP thymocytes engage thymic epithelial cells to test the interactions between αβ TCRs and peptide-loaded major histocompatibility complex (MHC) proteins. αβ TCRs with sufficient affinity for peptide:MHC are
positively selected and promote the differentiation of CD4+ or CD8+ single-positive (SP) thymocytes and cessation of RAG expression (von Boehmer and Melchers, 2010). Thymocytes with αβ TCRs that cannot derive these signals “die by neglect” and high affinity αβ TCRs are “negatively selected” via apoptosis, resulting in loss of these TCR specificities from the repertoire (von Boehmer and Melchers, 2010). Pre-B cells initiate Vκ-to-Jκ recombination one Igk allele at a time, and the RAG DSBs from one Vκ-to-Jκ rearrangement on one allele signal through the DNA damage response kinase ATM to transiently inhibit recombination of the other allele (Figure 1.3) (Steinel et al., 2013). An autoreactive IgH/Igκ BCR can induce negative selection by apoptosis, anergy, or Igk receptor editing by initiating Vκ recombination on either allele (Casellas et al., 2007). The formation and positive selection of an IgH/Igκ BCR signals permanent feedback inhibition of Vκ recombination and maturation of κ+ B cells (von Boehmer and Melchers, 2010).

As a result of interdependent controls of lymphocyte development and V(D)J recombination between alleles, ~90% of αβ T cells and ~97% of κ+ B cells express one type of AgR (Brady et al., 2010b).

Models Contributing to Monoallelic Antigen Receptor Assembly and Expression

Allelic exclusion of Tcrb, Igh, and Igk loci is thought to be mediated by: 1) assembly of a functional gene through V(D)J recombination on one allele, 2) transient feedback inhibition of V recombination signaled by DNA DSBs, and 3) AgR-mediated permanent feedback inhibition of V rearrangements (Figure 1.3) (Brady et al., 2010b; Steinel et al., 2013). While transient and permanent feedback inhibition mechanisms have been demonstrated experimentally (Steinel et al., 2010; Steinel et al., 2013; Uematsu et al., 1988), the mechanisms by which immature lymphocytes ensure sequential assembly of the two Tcrb, Igh, or Igk alleles prior to feedback inhibition remain unproven (Brady et al., 2010; Levin-Klein and Bergman, 2014; Mostoslavsky et al., 2004; Outters et al., 2015; Vettermann and Schlissel, 2010). In the field there is considerable disagreement about whether these mechanisms are deterministic or stochastic (Brady et al.,
2010; Hewitt et al., 2009; Levin-Klein and Bergman, 2014; Mostoslavsky et al., 2004; Outters et al., 2015; Schlimgen et al., 2008; Vettermann and Schlissel, 2010). Deterministic models entail one allele becoming available for V segment recombination, with activation of the second allele only if the first is assembled out-of-frame. In contrast, stochastic models posit that both alleles are activated within a similar time frame but inefficient V recombination makes it unlikely that a cell will complete assembly of both alleles before feedback inhibition is engaged.

Prevailing models of sequential AgR allele activation invoke epigenetic-based mechanisms, which are known to modulate many aspects of transcription, chromatin accessibility, and chromosome topology (Figure 1.4) (Brady et al., 2010; Shih and Krangel, 2013). Consequently, the field has focused on identifying epigenetic phenomena that correlate with monoallelic recombination of V gene segments at Tcrb, Igh, and Igk loci. In this regard, the two alleles for each of these loci replicate asynchronously in lymphocytes (Mostoslavsky et al., 2001). At least for Igk loci, this process initiates in lymphoid progenitors, is clonally maintained, and correlates with preferential recombination of the early replicating allele (Farago et al., 2012; Mostoslavsky et al., 2001). These data suggest a deterministic mechanism for monoallelic recombination that is associated with DNA replication. For Tcrb, Igh, and Igk, others have shown that one of their respective alleles often reside at transcriptionally repressive nuclear structures such as at the nuclear lamina and pericentromeric heterochromatin (Chan et al., 2013; Chen et al., 2018; Hewitt et al., 2009; Schlimgen et al., 2008; Skok et al., 2007). RAG2 protein is depleted from the nuclear periphery and the association of Tcrb alleles with the nuclear lamina suppresses Vβ recombination (Chan et al., 2013; Chen et al., 2018). These findings have led to models whereby differential positioning of alleles, via deterministic or stochastic mechanisms, governs asynchronous initiation of V recombination (Figure 1.4) (Chan et al., 2013; Chen et al., 2018; Hewitt et al., 2009; Schlimgen et al., 2008; Skok et al., 2007). Critically, although these epigenetic mechanisms might govern the assembly of a functional AgR gene on one allele before silencing of additional V rearrangement by feedback inhibition, causality has not been established for any.
While any of the aforementioned processes can be invoked to direct monoallelic V recombination in a deterministic or stochastic manner, a role for the low strengths of V_{\beta} (and possibly V_{H}) RSSs would argue for a stochastic mechanism (Figure 1.4) (Bassing et al., 2000; Jung et al., 2003; Tillman et al., 2003; Wu et al., 2003; Wu et al., 2007). Sequence features shared by V_{\beta} and V_{H} RSSs, but not in D_{\beta}, J_{H}, V_{\alpha}, or V_{\kappa} RSSs, have been proposed to render V rearrangements inefficient and thus stochastically limit near-simultaneous V-to-DJ recombination on both alleles (Figure 1.4) (Liang et al., 2002). A major aim of this thesis is to investigate whether weak V_{\beta} RSSs contribute in any way to restricting near-synchronous V_{\beta} recombination before the onset of feedback mechanisms.

**RSSs Orchestrate V(D)J Recombination and Exhibit Varying Levels of Activity**

RSSs direct RAG activity during AgR gene assembly. Immediately adjacent to the gene segment coding flank is the RSS heptamer (consensus sequence CACAGTG), a 12 (12RSS) or 23 (23RSS) base pair spacer, and the RSS nonamer (consensus sequence ACAAAAACC). Heptamer and nonamer sequences tend to be more conserved compared to the 12 or 23 spacer sequences, which tend to be divergent across AgR loci (Cowell et al., 2002; Lee, 2003). The most critical and well conserved portion of an RSS is the heptamer CA dinucleotide (Cowell et al., 2002). In fact, CA repeats are dispersed throughout the genome and can support RAG-mediated cleavage as cryptic RSSs (Agard and Lewis, 2000; Sakata et al., 2004; Teng et al., 2015). Variations in heptamer, spacer, and nonamer sequences can have profound effects on recombination frequency both in vitro and in vivo mouse models (Akira et al., 1987; Connor et al., 1995; Gauss and Lieber, 1992; Hesse et al., 1989; Jung et al., 2003; Larijani et al., 1999; Livak et al., 2000; Nadel et al., 1998; Olaru et al., 2004; Ramsden and Wu, 1991; VanDyk et al., 1996; Wei and Lieber, 1993; Wu et al., 2003). These experimental observations are relevant to human physiology as well: an RSS polymorphism in V_{\kappa}A2 decreases its recombination frequency and usage in Navajo Native Americans, correlating with greater susceptibility to *haemophilus*
influenzae type b infection (Feeney et al., 1996), and another RSS polymorphism in Vβ3 correlates with its decreased representation in the TCR repertoire (Posnett et al., 1994a).

At the murine Tcrb locus, Vβ segments are flanked by 23RSSs, Jβ segments by 12RSSs, and Dβ segments are flanked at the 5’ end by 12RSSs and the 3’ end by 23RSSs. Although the 12/23 rule allows direct Vβ rearrangements to Jβ segments, these events occur ~1000 times less often than Vβ rearrangements to DJβ complexes due to the inherent weaknesses of Vβ and Jβ RSSs as compared to 5’Dβ and 3’Dβ RSSs (Bassing et al., 2000; Drejer-Teel et al., 2007; Jung et al., 2003; Tillman et al., 2003; Wu et al., 2003). The shared sequence features of Vβ and VH RSSs is hypothesized to restrict Vβ and VH recombination frequency so as to enforce monoallelic Tcrb and Igh assembly (Liang et al., 2002).

Why Antigen Receptor Allelic Exclusion?

AgR allelic exclusion is a cardinal feature of adaptive immunity and is conserved across >500 million years of evolution (Cooper and Alder, 2006; Flajnik, 2018; Hsu, 2009). Evidence for allelic exclusion is observed at Igh loci in cartilaginous fish, which are the most evolutionarily ancient organisms to assemble TCR and Ig genes by RAG, as well as at VLR loci in lampreys (Boehm et al., 2012; Malecek et al., 2008), indicating that even in completely disparate adaptive immune systems AgR genes are assembled and expressed in a monoallelic fashion. This observation suggests that monoallelic assembly and/or expression of AgR genes is important for organismal survival. Despite the discovery of allelic exclusion in 1965 (Pernis et al., 1965), the role(s) of this process continues to elude immunologists.

The conventional model posits that monoallelic AgR expression is critical to inhibit the development of autoimmune diseases (Heath et al., 1993; Padovan et al., 1993). AgR repertoires are immensely diverse due to mechanisms contributing to combinatorial and junctional diversity (Lieber, 1991; Miles et al., 2011; Nikolich-Žugich et al., 2004). However, given diversity, many AgR specificities will be autoreactive and have the potential to cause harm to the host. Therefore,
to prevent disease lymphocytes bearing autoreactive AgR specificities must be purged from the repertoire, and this is in part achieved by negative selection in the thymus for T cells and the bone marrow for B cells (Marrack et al., 2008; Pelanda and Torres, 2012; Xing and Hogquist, 2012). In a TCR transgenic model, T cells expressing an autoreactive TCR are efficiently deleted in the thymus, but expression of a second non-autoreactive TCR rescues those cells and allows them to seed peripheral tissues (Zal et al., 1996). T cell effector functions can also be activated in these dual TCR T cells through stimulation of the autoreactive receptor (Zal et al., 1996). It is presumed biallelic AgR expression that produces autoreactive and non-autoreactive receptors would permit a developing lymphocyte to evade apoptosis by pro-survival signals derived from the non-autoreactive receptor (Heath et al., 1993; Padovan et al., 1993; Sarukhan et al., 1998; Zal et al., 1996). Consistent with this model, biallelic TCRα expression can potentiate autoimmunity in susceptible mouse models (Sarukhan et al., 1998; Schuldt et al., 2017). However, other models suggest that dual TCR expression is a benefit to the host as expression of a second αβ TCR was found to rescue useful αβ TCR specificities that are not efficiently positively-selected in the thymus (He et al., 2002). Thus, the autoimmune hazard of biallelic AgR gene expression remains unresolved.

Another puzzling observation draws the autoimmunity model under scrutiny. The mechanisms that enforce AgR allelic exclusion are not absolute and are applied to varying levels depending on the AgR locus. Allelic exclusion of Igh loci is the most stringent and estimates posit that 1/10,000 B cells express IgH proteins from both alleles (Barreto and Cumano, 2000; ten Boekel et al., 1998). Biallelic Tcra and Igk expression are estimated at ~3% and 7% of cells, respectively (Aifantis et al., 1997; Balomenos et al., 1995; Casellas et al., 2007; Steinel et al., 2010; Velez et al., 2007). Curiously, in mice and humans Vα recombination is not regulated between alleles and ~30% of αβ T cells possess in-frame VJα rearrangements at both Tcra alleles (Niederberger et al., 2003; von Boehmer and Melchers, 2010). Post-translational mechanisms such as pairing of TCRα chains with TCRβ chains restrict the frequency of dual TCR T cells to ~10%. These observations raise a fundamental question: why is allelic exclusion
applied to greatest effect at Igh and Tcrb loci but not to Tcra loci? Although V(D)J recombination is regulated and coordinated with cellular DNA damage responses and the cell cycle, recombination still poses a hazard for lymphocyte genomes as DNA DSBs are potent translocation substrates (Richardson and Jasin, 2000). Chromosomal translocations that place oncogenes under the influence of strong AgR enhancers are prevalent in some forms of leukemia and lymphoma (Küppers, 2005; Nussenzweig and Nussenzweig, 2010). Unlike all other AgR genes, Igh and Tcrb assembly is tightly linked with cellular proliferation following expression of IgH and TCRβ proteins. Following in-frame rearrangements at Igh and Tcrb loci in G1-phase cells, IgH and TCRβ proteins pair with surrogate light chains and preTα to form pre-BCRs and pre-TCRs, respectively. Pre-AgR expression signals the transcriptional activation of Cyclin D3 to drive cells into S phase where RAG2 is degraded to halt V(D)J recombination and inhibit oncogenic AgR translocations (Jiang et al., 2005; Lee and Desiderio, 1999; Sicinska et al., 2003; Steinel et al., 2014; Zhang et al., 2011). However, if RAG initiates V recombination and cleaves the homologous Igh or Tcrb allele before or as the cell enters S phase, oncogenic translocations between DSBs induced by RAG and by DNA replication errors may occur (Figure 1.5) (Brady et al., 2010b; Helmink and Sleckman, 2012). Therefore, our lab proposed that pressure to suppress oncogenic AgR translocations selected for mechanisms that direct monoallelic induction of RAG DSBs during Igh and Tcrb recombination (Brady et al., 2010b; Hewitt et al., 2009; Steinel et al., 2014). In this model, monoallelic initiation and feedback inhibition of V recombination limits the frequencies of developing lymphocytes with RAG DSBs that evade the G1/S cell cycle checkpoint and enter S phase (Figure 1.5). A distinction worth mentioning is that this translocation model highlights a role for monoallelic AgR gene assembly as compared to the role for monoallelic AgR expression in limiting autoimmunity. Critically, these models are not mutually exclusive and may be relevant to other RAG-based IgSF and/or VLR-based vertebrate adaptive immune systems.
Thesis Structure

This thesis aims to address a long outstanding question in the field AgR allelic exclusion: *How are V rearrangements asynchronously initiated between alleles to ensure monoallelic assembly of AgR genes prior to the engagement of feedback inhibition?*

In Chapter 2, I define a role for weak Vβ RSSs in limiting Vβ recombination frequency to mediate stochastic monoallelic TCRβ assembly.

In Chapter 3, I identify another role for poor quality Vβ RSSs in suppressing the assembly and expression of two distinct TCRβ genes from a single allele.

In Chapter 4, I report a role for the DNA damage kinase ATM in cooperation with weak Vβ RSSs to enforce allelic exclusion by limiting Vβ recombination on both alleles before TCRβ-signaled feedback inhibition silences Vβ recombination. I also highlight a role for ATM in preserving the proper repair of coding joins for V segments that recombine by inversion, thus protecting the composition of V repertoires.

In Chapter 5, I use the mouse models presented in Chapter 3 to test the relationship between diversity of the Vβ repertoire and antigen-specific immune responses.

Finally, in Chapter 6 I discuss the implications and future directions of this work.
Figure 1.1. Schematic of V(D)J Recombination. [A] Tcrb and Igh loci have variable (V), diversity (D), and joining (J) gene segments. To assemble the variable region exons of the TCRβ and IgH proteins, recombination must occur in two steps, first through a D-to-J rearrangement followed by a V-to-DJ rearrangement. The RAG recombinase initiates recombination by cleaving recombination signal sequences (triangles) that flank all gene segments. [B] Tcra and Igk/Igl loci do not possess D segments, thus recombination occurs through a single V-to-J rearrangement step. [C] Transcription with constant (C) region exons, splicing, and translation produce fully assembled αβ T cell receptors (TCRs) and κ+ or λ+ B cell receptors (BCRs). [D] Schematic of deletional and inversion rearrangements.
Intact monoallelic initiation and feedback inhibition of V recombination

A

VDJ\textsuperscript{IF}  
V\textrightarrow DJ  
\rightarrow 3 lymphocytes (60%)

VDJ\textsuperscript{OF}  
V\textrightarrow DJ  
\rightarrow 2 lymphocytes (40%)

VDJ\textsuperscript{IF}  
V\textrightarrow DJ  
\rightarrow Apoptosis

Feedback inhibition

B

No allelic exclusion

VDJ\textsuperscript{IF}  
\rightarrow 1 lymphocyte (20%)

VDJ\textsuperscript{OF}  
\rightarrow 2 lymphocytes (40%)

VDJ\textsuperscript{OF}  
\rightarrow 2 lymphocytes (40%)

VDJ\textsuperscript{OF}  
\rightarrow Apoptosis

Figure 1.2. Models of monoallelic versus biallelic antigen receptor gene assembly. Predicted fractions of in-frame (IF) and out-of-frame (OF) rearrangements. (A) Assuming monoallelic V recombination between alleles and subsequent feedback inhibition, in 60% of cells one allele will possess an IF VDJ rearrangement while the other allele will be unrearranged, and 40% of cells will have VDJ rearrangements on both alleles where one is IF and the other is OF. (B) In the absence of allelic exclusion, all cells will possess VDJ rearrangements on both alleles, but only in 20% of cells will both rearrangements be IF.
Allelic exclusion as \( f(t) \) during lymphocyte development

### Initiation
1. Asynchronous V-to-(D)J recombination
2. ATM-dependent transient inhibition
3. AgR-signaled permanent feedback inhibition

### Maintenance

**Figure 1.3. Models contributing to antigen receptor allelic exclusion.** Antigen receptor (AgR) allelic exclusion is thought to be mediated by: (1) monoallelic initiation of V-to-(D)J rearrangements between alleles, (2) a transient termination of V recombination signaled by RAG-mediated DNA double strand breaks (DSBs), and (3) AgR-signaled permanent feedback inhibition of V recombination. In the *initiation phase* of allelic exclusion, asynchronous V recombination between alleles (1) ensures that AgR assembly occurs one allele at a time. Feedback mechanisms *maintain* allelic exclusion: transient inhibition signaled through the RAG-DSBs and the ATM kinase (2) allows the cell time to test whether the initial rearrangement was in-frame. If not, the temporary inhibition of recombination is lifted and V rearrangements on the other allele can proceed. If the rearrangement is in-frame, the AgR protein is expressed and signals permanent inhibition of V rearrangements (3) to *maintain* monoallelic gene expression.
Figure 1.4. Potential mechanisms that promote monoallelic initiation of V-to-(D)J recombination between alleles. Depicted is the Tcrb locus as an example. Deterministic models propose that one allele is selected to undergo V-to-(D)J recombination and only activates the second allele if the first rearrangement is out-of-frame. Part of the deterministic model is an allelic preference that is imposed early during hematopoiesis and is clonally maintained. In contrast, stochastic models suggest that both alleles can undergo recombination in a similar time frame but inefficient V recombination renders it unlikely that both alleles are assembled before feedback inhibition is implemented. Epigenetic mechanisms that affect transcription, chromatin accessibility, locus topology, and differential positioning of alleles in the nucleus could occur via deterministic or stochastic means. RSS quality affecting the initiation frequencies of V\textsubscript{H}/V\textsubscript{\beta} recombination could be a stochastic mechanism that enforces monoallelic Igh and Tcrb assembly. RSS: recombination signal sequence. TSS: transcription start site.
Figure 1.5. Model for Tcrb translocations as a result of biallelic Tcrb gene assembly. (A) Normally, one Tcrb allele undergoes Vβ-to-DJβ recombination in G1-phase cells, and transient and permanent feedback mechanisms inhibit additional RAG-mediated DNA DSBs as cells move into S phase. (B) As cells progress into S phase, a RAG DSB on the homologous allele could form a chromosomal translocation with DNA DSBs generated by DNA replication. As AgR loci possess strong transcriptional enhancers (Eβ for Tcrb loci), a translocation that places an oncogene under the regulation of these enhancers would drive cellular transformation.
CHAPTER 2: Poor Quality Vβ Recombination Signal Sequences Stochastically Enforce TCRβ Allelic Exclusion

Abstract

The monoallelic expression of AgR genes, called allelic exclusion, is fundamental for highly specific immune responses to pathogens. This cardinal feature of adaptive immunity is achieved by the assembly of a functional AgR gene on one allele, with subsequent feedback inhibition of V(D)J recombination on the other allele. A range of epigenetic mechanisms have been implicated in sequential recombination of AgR alleles; however, we demonstrate that a genetic mechanism controls this process for Tcrb. Replacement of V(D)J recombinase targets at two different mouse Vβ gene segments with a higher quality target elevates Vβ rearrangement frequency prior to feedback inhibition, dramatically increasing the frequency of T cells with TCRβ chains derived from both Tcrb alleles. Thus, TCRβ allelic exclusion is enforced genetically by the low quality of Vβ recombinase targets that stochastically restrict the production of two functional rearrangements before feedback inhibition silences one allele.

Introduction

Monoallelic expression is an essential process that limits the dosage of numerous genes. While genetic imprinting and X-inactivation are vital for normal development and physiology, monoallelic expression of olfactory and antigen receptors is fundamental for highly specific recognition and responses to diverse odors or pathogens. To date, mechanisms enforcing monoallelic gene expression programs have been shown to involve epigenetic-based transcriptional activation of an expressed allele and silencing of the non-expressed allele (Khamlichi and Feil, 2018). Lymphocyte AgR allelic exclusion requires an additional level of regulation due to the obligate assembly of AgR genes by V(D)J recombination.

Developing T and B cells generate AgR diversity through RAG endonuclease-mediated recombination of T cell receptor (TCR) and immunoglobulin (Ig) variable (V), diversity (D), and
joining (J) gene segments. For *Tcrb*, *Igh*, and *Igk* loci, V-to-(D)J recombination between alleles is regulated such that most cells assemble and express a functional gene from only one allele. This allelic exclusion is achieved by initiation of V recombination on only one allele before protein from an in-frame V(D)J rearrangement signals feedback inhibition of V recombination. How developing lymphocytes ensure monoallelic assembly of a functional gene before feedback inhibition remains enigmatic, though both deterministic and stochastic models have been proposed based on correlative observations (Brady et al., 2010b; Levin-Klein and Bergman, 2014; Mostoslavsky et al., 2004; Outters et al., 2015; Vettermann and Schlissel, 2010). Deterministic models invoke that mechanisms initially activate one allele for V-to-(D)J recombination and activate the second only after the first fails to assemble a functional gene. In contrast, stochastic models posit that both alleles can be activated at the same time, but mechanisms lower recombination efficiency to make it unlikely both alleles assemble V(D)J rearrangements before feedback inhibition. Asynchronous replication of *Igk* alleles correlates with preferential Vκ-to-Jκ recombination of the early replicating allele (Mostoslavsky et al., 2001), suggesting a deterministic mechanism linked to DNA replication. In the lymphocyte lineages and developmental stages where *Tcrb*, *Igh*, and *Igk* recombine, one allele of each locus frequently resides at transcriptionally repressive structures, consistent with a role for differential positioning of alleles via deterministic or stochastic mechanisms (Chan et al., 2013; Hewitt et al., 2009; Schlimgen et al., 2008; Skok et al., 2007). In addition, sequence features shared by Vβ and VH RSSs, but not present in Dβ, JH, Vα, or Vκ RSSs, have been proposed to render V rearrangements inefficient and thus stochastically limit near-simultaneous V-to-DJ recombination on both alleles (Liang et al., 2002). Critically, while any of these mechanisms might ensure assembly of a functional gene on only one allele before feedback inhibition, none have been validated by experimentally demonstrating causal relationships.

The mouse *Tcrb* locus offers a powerful physiological platform to elucidate potential contributions of RSSs in monoallelic assembly and expression of functional AgR genes. *Tcrb* has 23 functional Vβs located 250-735 kb upstream of the Dβ1-Jβ1-Cβ1 and Dβ2-Jβ2-Cβ2 clusters,
each of which has one Dβ and six functional Jβs (Figure 2.1, A) (Glusman et al., 2001; Malissen et al., 1986). The locus has another Vβ (V31) located 10 kb downstream of Cβ2 and in opposite transcriptional orientation from other Tcrb coding sequences (Glusman et al., 2001; Malissen et al., 1986). RSSs consist of a semi-conserved heptamer and nonamer separated by a generally non-conserved 12 or 23 nucleotide spacer (Schatz and Swanson, 2011). Upon binding an RSS, RAG adopts an asymmetric tilt conformation that ensures the capture of a second RSS of differing length and bends each RSS by inducing kinks in their spacers (Kim et al., 2018; Ru et al., 2015). In vitro, ~40% of synapses between RSSs with consensus heptamers and nonamers proceed to cleavage (Lovely et al., 2015), and natural variations of heptamers, spacers, and nonamers can have major effects on recombination levels (Akira et al., 1987; Connor et al., 1995; Gauss and Lieber, 1992; Hesse et al., 1989; Larijani et al., 1999; Livak et al., 2000; Nadel et al., 1998; Olaru et al., 2004; Ramsden and Wu, 1991; VanDyk et al., 1996; Wei and Lieber, 1993). The only in vivo confirmation that natural RSS variations influence recombination levels is in the Tcrb locus (Bassing et al., 2000; Horowitz and Bassing, 2014; Jung et al., 2003; Sleckman et al., 2000; Wu et al., 2003; Wu et al., 2007). Vβs are flanked by 23-RSSs, Jβs by 12-RSSs, and Dβs by 5’12-RSSs and 3’23-RSSs (Glusman et al., 2001). Direct Vβ-to-Jβ rearrangements are permitted by the 12/23 rule; however, they rarely occur due to the inherent inefficiency of recombination between Vβ and Jβ RSSs (Bassing et al., 2000; Jung et al., 2003; Tillman et al., 2003; Wu et al., 2003; Wu et al., 2007). The recombination strength of a Tcrb RSS is a property determined at the biochemical level by its interactions with a partner RSS, the RAG endonuclease, and HMGB1 proteins that bend DNA (Banerjee and Schatz, 2014; Drejer-Teel et al., 2007; Jung et al., 2003). In this context, 3’Dβ RSSs are at least 10-fold better than Vβ RSSs at recombining with 5’Dβ RSSs in vitro (Banerjee and Schatz, 2014; Drejer-Teel et al., 2007; Jung et al., 2003). Accordingly, replacement of an endogenous V31 RSS with the 3’Dβ1 RSS increases the percentage of αβ T cells expressing V31+ TCRβ chains due to the elevated
recombination level of V31 relative to other Vβ segments (Horowitz and Bassing, 2014; Wu et al., 2003).

To determine the potential roles of Tcrb RSSs in governing TCRβ allelic exclusion, we made and studied mice carrying replacement(s) of their endogenous V31 and/or Trbv2 (V2) RSSs with a better 3'Dβ1 RSS. All of these mice develop a greater percentage of αβ T cells expressing V2⁺ or V31⁺ TCRβ protein at the expense of cells using another type of TCRβ chain. We demonstrate that each Vβ RSS replacement increases Vβ rearrangement before feedback inhibition, competes with the homologous allele for usage in the TCRβ repertoire, and elevates the percentage of αβ T cells expressing TCRβ proteins from both alleles. We conclude that the poor qualities of Vβ RSSs for recombining with Dβ and Jβ RSSs enforce TCRβ allelic exclusion by stochastically limiting Vβ rearrangements before feedback inhibition from one allele halts further Vβ recombination.

Results

Generation of Vβ RSS Replacement Mice with Grossly Normal αβ T Cell Development

To determine contributions of Tcrb RSSs in allelic exclusion, we established C57BL/6 mice carrying germline replacements of the V2 or V31 RSS with the stronger 3'Dβ1 RSS, referred to as the V2R or V31R modifications (Figure 2.1, A-C). We created mice with each replacement on one allele (V2R/+, V31R/+), both alleles (V2R/R, V31R/R), or opposite alleles (V2R+/V31R/+). The assembly and expression of functional Tcrb genes is essential for αβ T cell development (Bouvier et al., 1996; Mombaerts et al., 1992). In thymocytes, Dβ-to-Jβ rearrangement initiates in ckit⁺CD25⁺ DN1 cells and continues in ckit⁺CD25⁺ DN2 and ckit⁺CD25⁺ DN3 cells, while Vβ-to-DJβ recombination initiates in DN3 cells (Godfrey and Zlotnik, 1993). The expression of a functional Tcrb gene in DN3 cells is necessary and rate-limiting for differentiation of ckit⁺CD25⁺ DN4 cells and then DP thymocytes (Baldwin et al., 2005; Serwold et al., 2007; Shinkai et al., 1992; Yang-Iott et al., 2010). We observed normal numbers and frequencies of splenic αβ T cells and
thymocytes at each developmental stage in every genotype of Vβ RSS replacement mice (Figure 2.1, D-J). In thymocytes, Dβ-to-Jβ rearrangement initiates at the DN1 stage and continues in DN2 and DN3 stages, while Vβ-to-DJβ recombination occurs only at the DN3 stage (Godfrey et al., 1993). V2R and V31R rearrangements initiate in DN3 cells and occur at much greater than normal levels (Figure 2.2, A-C). Thus, replacement of the V2 or V31 RSS with the 3'Dβ1 RSS substantially increases the frequency of V2 or V31 recombination without altering normal development of αβ T cells.

**RSS-replaced Vβ Segments Outcompete Unmodified Vβ Segments in the TCR Repertoire**

In wild-type C57BL/6 mice, the representation of individual Vβ segments within the αβ TCR repertoires of DP thymocytes, SP thymocytes, and naïve splenic αβ T cells is similar and mirrors their relative levels of rearrangement in DN3 thymocytes (Wilson et al., 2001). Thus, we performed flow cytometry on mature naïve αβ T cells (SP thymocytes and splenic αβ T cells) to determine effects of Vβ RSS substitutions on Vβ recombination and resultant usage in the αβ TCR repertoire. We used an antibody for a Cβ epitope contained in all TCRβ proteins in combination with different Vβ-specific antibodies that bind peptides encoded by a single Vβ [V2, Trbv4 (V4), Trbv19 (V19), or V31] or a family of Vβs [Trbv12.1 and Trbv12.2 (V12) or Trbv13.1, Trbv13.2, and Trbv13.3 (V13)]. In WT mice, we observed that 7.0% of SP cells express V2+ or V31+ TCRβ chains on their surface (Figure 2.3, A-C). For mice with V2 or V31 RSS replacement on one or both alleles, we detected a 6-11-fold increased representation of each modified Vβ on SP cells (Figure 2.3, A-C). Specifically, we detected V2+ TCRβ chains on 40.9% of cells from V2R/+ mice and on 61.4% of cells from V2R/R mice, and V31+ TCRβ chains on 50.0% of cells from V31R/+ mice and on 77.1% of cells from V31R/R mice (Figure 2.3, A-C). As all six genotypes exhibit similar numbers of SP cells (Figure 2.1, D), the increased utilization of each RSS-replaced Vβ must be at the expense of other Vβ segments. Indeed, the percentages of V31+ SP cells are reduced in V2R/+ mice compared to WT mice (5.1% versus 7.0%) and in V2R/R mice relative to
V2\textsuperscript{R/+} mice (3.8\% versus 5.1\%, Figure 2.3, A-C). Likewise, the percentages of V2\(^+\) SP cells are reduced in V31\textsuperscript{R/+} mice compared to WT mice (4.3\% versus 7.0\%) and in V31\textsuperscript{R/R} mice relative to V31\textsuperscript{R/+} mice (2.3\% versus 4.3\%, Figure 2.3, A-C). Moreover, the percentage of SP cells expressing V4\(^+\), V12\(^+\), V13\(^+\), or V19\(^+\) TCR\(\beta\) protein is lower than normal in V2\textsuperscript{R/+} and V31\textsuperscript{R/+} mice, and further reduced in V2\textsuperscript{R/R} and V31\textsuperscript{R/R} mice (Figure 2.3, A and D, and data not shown). The V\(\beta\) usage in splenic \(\alpha\beta\) T cells of each V\(\beta\) RSS replacement mouse genotype is altered similarly as on SP thymocytes (Figure 2.4, A-E). These data show that the stronger 3'D\(\beta\)1 RSS empowers V2 and V31 to outcompete normal V\(\beta\) segments for recombination and resultant usage in the \(\alpha\beta\) TCR repertoire.

Notably, each genotype of homozygous V\(\beta\) RSS replacement mice has a ~1.5-fold greater representation of its modified V\(\beta\) compared to the corresponding heterozygous genotype (Figure 2.4, F). This less than additive effect based on allelic copy number suggests that Tcrb alleles compete for rearrangement and resultant usage in the \(\alpha\beta\) TCR repertoire. Our analysis of V2\textsuperscript{R/+}, V31\textsuperscript{R/+}, and V2\textsuperscript{R/+}/V31\textsuperscript{+R} mice yields additional evidence for this competition as each RSS-replaced V\(\beta\) is less represented in V2\textsuperscript{R/+}/V31\textsuperscript{+R} mice relative to V2\textsuperscript{R/+} or V31\textsuperscript{R/+} mice (Figures 2.3, A-C and 2.4, A-C). Specifically, V2 is expressed on 32.2\% of SP cells in V2\textsuperscript{R/+}/V31\textsuperscript{+R} mice compared to 40.9\% in V2\textsuperscript{R/+} mice, and V31 is expressed on 42.0\% of SP cells in V2\textsuperscript{R/+}/V31\textsuperscript{+R} mice relative to 50.0\% in V31\textsuperscript{R/+} mice (Figure 2.3, A-C). We observed similar differences among splenic \(\alpha\beta\) T cells (Figure 2.4, A-C). These differences imply that the overall V\(\beta\) recombination efficiency of each RSS-replaced allele is elevated such that it effectively competes with the other allele for recombination in thymocytes and usage in the \(\alpha\beta\) TCR repertoire.

Competition between Tcrb alleles implies that rearrangement of the unmodified allele in heterozygous V\(\beta\) RSS replacement mice might limit the extent that each RSS-replaced V\(\beta\) outcompetes other V\(\beta\) segments on the modified allele. To test this, we generated mice with the WT, V2\(^R\), or V31\(^R\) allele opposite an allele with deletion of the Tcrb enhancer (E\(\beta\)). As E\(\beta\) deletion blocks all Tcrb recombination events \textit{in cis} (Bories et al., 1996; Bouvier et al., 1996), an E\(\beta\)-
deleted \( (E^\beta) \) allele cannot compete with an active \( Tcrb \) allele. We compared the \( V_\beta \) repertoires of mature \( \alpha\beta \) T cells from \( WT/E^\beta \), \( V2^R/E^\beta \), and \( V31^R/E^\beta \) mice to cells from \( WT \), \( V2^{R/+} \), and \( V31^{R/+} \) mice. The percentages of \( V2^+ \) and \( V31^+ \) SP thymocytes each are equivalent between \( WT/E^\beta \) and \( WT \) mice (Figure 2.5, A-D). In contrast, representation of each RSS-replaced \( V_\beta \) is \( \sim \)1.5-fold greater in \( V2^R/E^\beta \) or \( V31^R/E^\beta \) mice relative to \( V2^{R/+} \) or \( V31^{R/+} \) mice, respectively (Figure 2.5, A-D). Furthermore, the percentages of \( V2^+ \) and \( V31^+ \) cells in \( V2^R/E^\beta \) and \( V31^R/E^\beta \) mice are similar to those of \( V2^{R/R} \) and \( V31^{R/R} \) mice, respectively (compare Figure 2.3, A-C with Figures 2.5, A-D). These comparisons reveal that recombination of a wild-type \( Tcrb \) allele indeed limits the extent to which each RSS-replaced \( V_\beta \) can outcompete other \( V_\beta \) segments on the same allele.

\( V_\beta \) RSS Replacements Increase Biallelic Assembly and Expression of Functional TCR\( \beta \) Genes

We next determined effects of \( V_\beta \) RSS replacements on monoallelic TCR\( \beta \) expression. Due to the absence of allotypic markers that can identify TCR\( \beta \) chains from each allele, the field assays TCR\( \beta \) allelic exclusion by quantifying cells that stain with two different anti-\( V_\beta \) antibodies. This method suggests that 1-3\% of \( \alpha\beta \) T cells exhibits biallelic TCR\( \beta \) expression (Balomenos et al., 1995; Steinel et al., 2014). However, this might be an underestimation as antibodies are not available for all \( V_\beta \) proteins and biallelic \( Tcrb \) expression involving the same \( V_\beta \) segment cannot be discerned. Regardless, we used this approach to determine the percentages of \( \alpha\beta \) T cells expressing two different types of TCR\( \beta \) chains in \( WT \), \( V2^{R/+} \), \( V2^{R/R} \), \( V31^{R/+} \), \( V31^{R/R} \), and \( V2^{R/+}/V31^{R/R} \) mice. We first used an antibody for \( V2 \) or \( V31 \) combined with an antibody for \( V4 \), \( V12 \), \( V13 \), or \( V19 \). For each combination, we observed that 0.05-0.21\% of SP cells stained with both antibodies in \( WT \) mice (Figure 2.6, A-D). In \( V31^{R/+} \) and \( V31^{R/R} \) mice, we detected increased frequencies of SP cells that stained for \( V31 \) and each other \( V_\beta \) tested (Figure 2.6, C and D). Likewise, for \( V2^{R/+} \) and \( V2^{R/R} \) mice, we saw increased frequencies of SP cells that stained for \( V2 \) and each other \( V_\beta \) (Figure 2.6, A and B). We also observed inverse trends where the frequencies of SP cells expressing \( V2 \) and another \( V_\beta \) decreased in \( V31^{R/+} \) and \( V31^{R/R} \) mice, as well as the
frequencies of SP cells expressing V31 and another Vβ decreased in V2R/+ and V2R/R mice (Figure 2.6, A-D). We next quantified V2\(^{-}/V31^{+}\) cells and observed that 0.09% of SP thymocytes stained with both V2 and V31 antibodies in WT mice (Figure 2.6, E and F). In mice carrying V2\(^{R}\) or V31\(^{R}\) on one or both alleles, we detected 0.3-0.68% of SP cells stained with both antibodies (Figure 2.6, E and F). Strikingly, the frequency of V2\(^{-}/V31^{+}\) SP cells is 27-fold higher in V2\(^{R/+}/V31^{+}/R\) mice compared to WT mice (2.47% versus 0.09%, Figure 2.6, E and F). To address any potential background staining from the increased frequencies of V2\(^{+}\) and V31\(^{+}\) cells in V2\(^{R/+}/V31^{+}/R\) mice, we mixed equal numbers of SP cells from V2\(^{R/R}\) and V31\(^{R/R}\) mice. Notably, the frequency of V2\(^{-}/V31^{+}\) cells in V2\(^{R/+}/V31^{+}/R\) mice is 3.5-fold greater than in mixed V2\(^{R/R}\) and V31\(^{R/R}\) cells (Figure 2.6, E and F). This provides firm evidence that V2\(^{-}/V31^{+}\) cells in V2\(^{R/+}/V31^{+}/R\) mice are αβ T cells expressing both V2\(^{+}\) and V31\(^{+}\) TCRβ chains. The sum of the frequencies of double-staining cells for all Vβ combinations tested shows that the total incidence of SP cells expressing two types of TCRβ chains is increased for each Vβ RSS replacement genotype (Figure 2.6, G).

The highest incidence of dual-TCRβ expressing thymocytes is in V2\(^{R/+}/V31^{+}/R\) mice and is 5-fold more than in WT mice (Figure 2.6, G). Similar increased incidences of dual-TCRβ expression was observed in splenic αβ T cells (Figure 2.7, A-G). Collectively, these data provide strong evidence that replacement of a V2 and/or V31 RSS with the 3'Dβ1 RSS elevates the frequencies of mature αβ T cells exhibiting biallelic TCRβ expression.

To determine if Vβ RSS replacements increase biallelic assembly of functional TCRβ genes, we created 102 αβ T cell hybridomas from V2\(^{R/+}/V31^{+}/R\) mice and analyzed Tcrb rearrangements. We compared our data to a prior study of 212 wild-type hybridomas, where 56.6% contained a single Vβ rearrangement on one allele and DJβ rearrangement on the other allele [V(D)J/DJ, Table 2.1], and 43.4% contained an in-frame and an out-of-frame Vβ rearrangement on opposite alleles [V(D)J/V(D)J, Table 2.1] (Khor and Sleckman, 2005). Of our V2\(^{R/+}/V31^{+}/R\) hybridomas, 45.1% were V(D)J/DJ and 31.4% were V(D)J/V(D)J (Table 2.1). Unexpectedly, 23.5% of V2\(^{R/+}/V31^{+}/R\) hybridomas had two Vβ rearrangements (V31 and another
Vβ) on one allele, which has never been observed in WT cells (23.5% versus 0%, Table 2.1, Figure 2.7, H) (Khor and Sleckman, 2005). We also observed V31 recombination directly to Jβ segments in 23.5% of V2R+/V31+/R hybridomas (Table 2.1). Such direct Vβ-to-Jβ rearrangements rarely occur and have only been reported in hybridomas from mice carrying replacement of the V31 RSS with the better 3′Dβ1 RSS or the Jβ1.2 RSS with the stronger 5′Dβ1 RSS (Bassing et al., 2000; Sleckman et al., 2000; Wu et al., 2003; Wu et al., 2007). Finally, we found that eight (7.8%) of the V2R+/V31+/R hybridomas had recombination of both V2R and V31R (Table 2.1). Notably, two of these contained an in-frame V2DJβ rearrangement on one allele and an in-frame V31DJβ rearrangement on the other allele (Tables 2.1 and 2.2), mirroring the 2.47% of V2+V31+ cells detected by flow cytometry. While our hybridoma analysis provides unequivocal evidence that Vβ RSS replacements on opposite alleles increases the overall frequency of Vβ recombination, our sample size precludes concrete evidence for an elevated frequency of biallelic in-frame Vβ rearrangements. To address these limitations of our hybridoma analysis, we turned to an unbiased, comprehensive, and high-throughput sequencing approach.

Congenically marked mouse lgh and lgk constant regions allow for the measurement of global changes in B cell allelic exclusion by flow cytometry (Casellas et al., 2001; Sonada et al., 1997). Unfortunately, these congenic tools are not available for the Tcrb locus, which is why allelic exclusion is quantified by anti-Vβ antibodies. However, this strategy underestimates the degree of biallelic Tcrb assembly as antibodies do not exist for every single Vβ, and staining for the permutations of every Vβ pair is technically cumbersome. To gain more insight into the global changes in Tcrb allelic exclusion due to the RSS replacements, we utilized Adaptive Biotechnologies’ next-generation Immunoseq platform. In brief, the method employs a multi-plex PCR using primers for each Vβ and Jβ segment to amplify all possible V(D)Jβ rearrangements from a genomic DNA sample (Figure 2.6, H) (Carlson et al., 2013). Although next-generation sequencing can be performed on either DNA and RNA, the unequal activity of Vβ promoters and nonsense-mediated degradation of out-of-frame rearrangements can impact the interpretation of
RNA-based sequencing data. V(D)Jβ rearrangements are barcoded, attached to adaptors, and sequenced by Illumina MiSeq (Figure 2.6, H) (Carlson et al., 2013). Importantly, Dβ segments, Jβ segments, and DJβ joins are not amplified by this technique. If allelic exclusion is normal, it would be predicted that ~60% of cells would have one in-frame (IF) VDJβ rearrangement, and ~40% of cells would have two VDJβ rearrangements where one is IF and one is out-of-frame (OF, Figure 2.6, I). Therefore, by sequencing 71.4% of VDJβ rearrangements would be IF and 28.6% of rearrangements would be OF, again because DJβ rearrangements are not sequenced (Figure 2.6, I). A bias of IF VDJβ rearrangements that is lower than the 71.4% theoretical estimate would indicate defects in mechanisms promoting monoallelic TCRβ assembly and expression (Figure 2.6, I). In the complete absence of these mechanisms, up to 20% of αβ T cells could possess two IF VDJβ rearrangements and 80% of cells could possess one IF and one OF VDJβ (Figure 2.6, I). In this case, sequencing would reveal that 60% of VDJβ rearrangements in a population would be IF (Figure 2.6, I). We isolated DNA from sorted WT and V2R+/V31+/R CD4+ SP thymocytes and submitted them for Adaptive Immunosequencing. In WT cells, we observe that 69.4% of unique VDJβ rearrangements are IF, consistent with the theoretical 71.4% of rearrangements if most cells exhibited monoallelic Tcrb assembly and expression (Figure 2.6, J and K). Sequence analysis of V2R+/V31+/R CD4+ SP thymocytes reveals that 60.2% of VDJβ rearrangements are IF, suggesting that most cells are attempting two rearrangements (Figure 2.6, J and K). By considering our hybridoma, flow cytometry, and next-generation sequencing data together, we conclude that replacement of a Vβ RSS with the 3’Dβ1 RSS on opposite alleles increases the overall frequency of Vβ rearrangements, leading to greater incidence of biallelic assembly and expression of functional Tcrb genes.

The Ability of the 3’Dβ1 RSS to Elevate Vβ Recombination Does not Require c-Fos Binding

The increased Vβ recombination and TCRβ allelic inclusion in Vβ RSS replacement mice can be explained by the greater strength of the 3’Dβ1 RSS in recombining with Dβ and Jβ RSSs.
Yet, this interpretation is complicated by the fact that 3'Dβ RSSs, but not Vβ RSSs, contain a c-Fos transcription factor binding site that spans the heptamer and spacer. In vitro, c-Fos interacts with RAG proteins when bound to 3'Dβ RSSs (Figure 2.8A) (Wang et al., 2008), leading to a model where c-Fos deposits RAG on 3'Dβ RSSs to sterically hinder Vβ recombination until Dβ-to-Jβ recombination deletes a 3'Dβ RSS (Wang et al., 2008). To investigate potential effects of c-Fos-mediated RAG deposition and/or transcription-associated accessibility, we made C57BL/6 mice carrying germline V2 or V31 RSS replacements with a 3'Dβ1 variant RSS, referred to as the V2F or V31F modification. The 3'Dβ1 RSS variant contains a two-nucleotide substitution in the spacer that abolishes both c-Fos binding and c-Fos-mediated RAG deposition, but has no obvious effect on the activity of the 3'Dβ1 RSS at recombining to a Vκ RSS in vitro (Figure 2.8, A and B) (Wang et al., 2008). We established mice with each Vβ RSS replacement on one allele (V2F+/+ and V31F+/+) or opposite alleles (V2F+/V31F+/). We performed flow cytometry on WT, V2F+/+, V31F+/+, and V2F+/V31F+/ mice to determine effects of V2F and/or V31F alleles on αβ T cell development, TCRβ repertoire, and TCRβ allelic exclusion. The numbers and frequencies of splenic αβ T cells and thymocytes at each developmental stage are normal in each Vβ RSS replacement genotype (data not shown), indicating no major effects on αβ T cell development. In contrast, replacement of the V2 and/or V31 RSS with the 3'Dβ1 RSS variant increases the frequencies of SP cells expressing V2, V31, or both (Figure 2.8, C-E). Specifically, we found a 4.9-fold greater than normal representation of V2+ cells in V2F+ mice and a 5.4-fold higher than normal frequency of V31+ cells in V31F+ mice (Figure 2.8, C-E). Similar to what we observed in V2F+/V31F+ mice, the V2F and V31F alleles compete for recombination in thymocytes as we detected smaller increases of V2+ and V31+ cells in V2F+/V31F+ mice (Figure 2.8, C-E). Finally, we detected higher than normal frequencies of V2V31+ SP cells in V2F+/, V31F+/, and V2F+/V31F+ mice (Figure 2.8, F and G), revealing that replacement of a V2 and/or V31 RSS with the 3'Dβ1 RSS variant elevates the incidence of biallelic TCRβ expression. These data indicate that neither RAG deposition nor transcription-
associated accessibility from potential c-Fos binding is needed for 3'Dβ1 RSS substitutions to increase Vβ rearrangement and biallelic TCRβ expression.

**Vβ RSS Replacements Increase the Initiation of Vβ Recombination before Feedback Inhibition**

The elevated incidences of biallelic TCRβ expression in Vβ RSS replacement mice could arise from increased initiation of Vβ recombination prior to enforcement of feedback inhibition and/or continued Vβ recombination during feedback inhibition. We found that the rearrangements of RSS-replaced Vβs are elevated in DN3 thymocytes of V2R/Eβ and V31R/Eβ mice (Figure 2.2, C). As these Vβs rearrange independent of competition and feedback inhibition from the other allele, the 3'Dβ1 RSS substitution does increase Vβ recombination before feedback inhibition. To determine if rearrangements of RSS-replaced Vβs are halted by TCRβ-mediated feedback inhibition, we established and analyzed V2R/+ and V31R/+ mice with a pre-assembled functional TCRβ transgene (TcrbTg). Expression of the transgene-encoded V13+ TCRβ chain initiates in DN2 cells and signals feedback inhibition of Vβ rearrangements (Steinel et al., 2010). However, ~3% of TcrbTg αβ T cells also expresses TCRβ protein from VβDβJβ rearrangements that occur in DN3 cells before TcrbTg-mediated feedback inhibition (Steinel et al., 2010). We observed that the TcrbTg more effectively decreases utilization of V2 than V31 when each is flanked by their own RSS or the 3'Dβ1 RSS (Figure 2.9, A-D). Next, we made αβ T cell hybridomas from V2R+, TcrbTgV2R+, V31R+, and TcrbTgV31R+ mice to quantify Vβ rearrangements. We detected V2 rearrangements in 50% of V2R+ clones but not in any TcrbTgV2R+ cells (p = 2.68x10⁻⁵, Pearson’s χ² test with Yates’ correction), and we found V31 rearrangements in 50% of V31R+ clones and in only 15% of TcrbTgV31R+ cells (p = 1.63x10⁻⁵, Pearson’s χ² test with Yates’ correction, Table 2.3). Our previous analysis of 129 TcrbTg αβ T cell hybridomas showed that 2.3% had a V31 rearrangement and an additional 7% carried recombination of a different Vβ (Table 2.3 and data not shown) (Steinel et al., 2010). Collectively, these data indicate that TcrbTg-signaled feedback inhibition suppresses recombination of RSS-replaced Vβs and does so more effectively for V2R...
versus V31R. As feedback inhibition could block accessibility of 5'Dβ RSSs in DN cells (Bassing et al., 2000), RSS-replaced Vβs may continue to recombine with Jβ segments until Vβ recombination is silenced by differentiation of DP thymocytes. To assess this possibility, we quantified V31 targeting to DβJβ complexes or Jβ segments in V31R+ and TcrbTqV31R+ hybridomas where V31 is the only rearranged Vβ segment. These primary V31 rearrangements occurred to Jβ segments in 38% of V31R+ and 14% of TcrbTqV31R+ cells (Table 2.3), revealing that TCRβ-mediated feedback inhibition also suppresses V31R-to-Jβ recombination. Our hybridoma analysis reveals that increased initiation of Vβ recombination prior to TCRβ-mediated feedback inhibition is the predominant, if not sole, mechanistic basis for the elevated frequencies of biallelic TCRβ expression in Vβ RSS replacement mice.

Vβ RSS Replacements Permit Direct Vβ-to-Jβ Recombination that Reduces TCRβ Repertoire Diversity

Although the 12/23 rule allows direct Vβ rearrangements to Jβ segments, these events occur ~1000 times less often than Vβ rearrangements to DJβ complexes (Bassing et al., 2000; Drejer-Teel et al., 2007; Jung et al., 2003; Tillman et al., 2003; Wu et al., 2003). As a result of replacing the V2 or V31 RSS with the 3'Dβ1 RSS, direct joining of Vβ and Jβ segments may occur. To determine if this is the case, we isolated DNA from sorted WT, V2R/R, and V31R/R CD69- TCRβlo DP thymocytes and submitted them for Adaptive Immunosequencing. We measured the lengths of CDR3β sequences since CDR3β length is in part determined by the presence of Dβ segments (Hughes et al., 2003). We then assigned these CDR3β sequences into one of three groups: i) if Dβ1 sequences are identified, ii) if Dβ2 sequences are identified, or iii) “unknown,” if neither Dβ1 or Dβ2 sequences are discernable or if the sequence can come from either Dβ segment. In cases where Dβ1 and Dβ2 sequences are distinguishable, the CDR3β lengths are equivalent among all three genotypes (Figure 2.10, A). However, in instances where Adaptive cannot call Dβ sequences, the CDR3β lengths are shorter in V2R/R and V31R/R pre-selection
thymocytes relative to WT controls, suggesting the exclusion of Dβ segments (Figure 2.10, A). In WT cells, the inability to call Dβ sequences is probably due to considerable chew-back and then addition of nucleotides during the repair of coding ends. This observation of direct Vβ-to-Jβ rearrangements in V2\textsuperscript{R/R} and V31\textsuperscript{R/R} cells was confirmed by conventional and quantitative PCR (data not shown).

Excluding Dβ segments during VDJβ recombination may lead to the loss of TCRβ diversity. RAG introduces DNA DSBs at gene segments to produce hairpin-sealed coding and blunted signal ends that are repaired by ubiquitous NHEJ factors (Helmink and Sleckman, 2012). Coding ends are enzymatically opened by a single-stranded nick, and nucleotides are excised and/or added to introduce additional diversity to V(D)J junctions (Helmink and Sleckman, 2012). CDR3β high-throughput sequencing reveals that hundreds of “public” TCRβ sequences are shared amongst mice presumably due to “convergent recombination” events that generate the same nucleotide TCRβ sequence (Madi et al., 2014; Quigley et al., 2010; Turner et al., 2006). We compared our sequencing data amongst individual mice and observe some low-frequency public sequences in WT pre-selection thymocytes (Figure 2.10, B). We then performed the same comparison among individual mice in V2\textsuperscript{R/R} and V31\textsuperscript{R/R} thymocytes and found an increased abundance and increased frequencies of public TCRβ sequences (Figure 2.10, B). We then determined whether Dβ segments were identifiable amongst these public sequences and found a majority fell into the “unknown” category. This observation makes intuitive sense as a rearrangement containing all three Vβ, Dβ, and Jβ segments possesses two junctions, while a Vβ-to-Jβ join possesses only one junction and lacks the additional diversity from the Dβ segment itself. These data indicate that weak Vβ RSSs promote TCRβ diversity by ensuring Dβ segment incorporation through limiting direct Vβ-to-Jβ rearrangements.

Discussion

Improving the Quality of Vβ RSSs Increases Vβ Rearrangement and Biallelic TCRβ Expression
Here we show that elevating Vβ rearrangement frequency by replacing the endogenous V2 and/or V31 RSS with the 3'Dβ1 RSS increases the incidences of biallelic TCRβ expression before TCRβ-signaled feedback inhibition. Nucleotide sequence differences between the 3'Dβ1 and V2 or V31 RSSs must provide the mechanistic basis for enhanced Vβ rearrangement. Although the 3'Dβ1 RSS can bind c-Fos to recruit RAG in vitro (Wang et al., 2008), we demonstrate that a variant 3'Dβ1 RSS that cannot bind c-Fos maintains elevated levels of Vβ recombination and biallelic TCRβ expression in vivo. The increased rearrangement and preferential targeting of RSS-replaced V2 and V31 to DJβ complexes mirrors the relative in vitro activities of the 3'Dβ1, V2, and V31 RSSs for RAG/HMGB1-mediated synapsis and cleavage (Banerjee and Schatz, 2014; Drejer-Teel et al., 2007; Jung et al., 2003). Context dependent kinking of each RSS spacer must occur for RAG/HMGB1 to pair RSSs for coupled cleavage (Kim et al., 2018). Indeed, the normal and variant 3'Dβ1 RSSs possess spacers of greater bending quality compared to the V2 and V31 RSSs (Kim et al., 2018) and heptamers and nonamers of less overall variation from consensus (Figure 2.1, B). Accordingly, the increased levels of Vβ rearrangement in our RSS replacement mice can be explained mechanistically by the greater quality of the 3'Dβ1 RSS for RAG/HMGB1-mediated pairing and coupled cleavage with 5'Dβ and Jβ RSSs. Although our observations cannot exclude potentially minor contributions of RSS differences in regulating RAG cleavage in Vβ chromatin, our 3'Dβ1 RSS replacements neither introduce a recognized transcription factor binding site nor increase germline transcription of V2 or V31 (data not shown). Some RSSs, albeit not the 3'Dβ1 RSS, position nucleosomes over themselves even within accessible chromatin (Baumann et al., 2003; Kondilis-Mangum et al., 2010). Thus it is possible that Vβ RSSs, but not 3'Dβ1 RSSs, bind nucleosomes to antagonize recombination of otherwise accessible Vβ segments since RSS nucleosome occupancy inhibits RAG access and cleavage (Golding et al., 1999; Kwon et al., 1998). Even so, improving the quality of a Jβ RSS in mice lacking Dβ segments permits robust Vβ-to-Jβ rearrangements where the normal poor quality Vβ RSSs are intact (Bassing et al., 2000). Therefore, we conclude that
the improved Vβ RSS quality for recombining with Dβ and Jβ RSSs is the underlying mechanism for increased Vβ recombination and biallelic TCRβ expression in our Vβ RSS replacement mice.

**The Poor Qualities of Vβ RSSs Provide a Stochastic Mechanism for Limiting Vβ Recombination Before Feedback Inhibition**

Our study offers the first validated mechanism for how only one allele of any AgR locus assembles a functional gene before feedback inhibition. In addition to promoting Vβ recombination and biallelic TCRβ expression, improving the quality of one Vβ RSS reveals that Tcrb alleles compete for Vβ recombination. Our data show that elevating the efficiency of Vβ recombination on one allele competes with Vβ rearrangements on the other. This can only occur if both alleles are activated in a similar time window(s) before the enforcement of feedback inhibition. If deterministic models were correct, recombination of one allele would have no effect on the second allele, and the increases of Vβ recombination in RSS replacement mice would be additive. Consequently, our data demonstrate that the poor qualities of Vβ RSSs for recombining with Dβ and Jβ RSSs provides a stochastic mechanism that serves a major role in limiting the incidence of functional Vβ rearrangements on both alleles before feedback inhibition terminates Vβ recombination.

We propose the following model for how thymocytes enforce TCRβ allelic exclusion. In non-cycling DN3 cells, at least one allele becomes active and undergoes Vβ-to-DJβ recombination. The DNA DSBs trigger transient feedback inhibition at least in part by repression of RAG expression, providing time to test the initial rearrangement (Fisher et al., 2017; Steinel et al., 2014). In cells where both alleles are activated, poor Vβ RSSs limit the likelihood of Vβ recombination on both alleles before loss of RAG. If the rearrangement is out-of-frame, RAG re-expression permits Vβ recombination on the second allele, or on the first allele if a Dβ2Jβ2 complex is available. In the latter case, poor Vβ RSSs again decrease the chance for Vβ recombination on both alleles. When the first rearrangement is in-frame, its TCRβ protein
activates Cyclin D3 to move cells into S phase (Sicinska et al., 2003), where RAG2 is degraded (Lin and Desiderio, 1994). Based on its function in pro-B cells (Powers et al., 2012), Cyclin D3 may repress Vβ accessibility before cells enter S phase. In DN3 cells where RAG is re-expressed between DSB repair and S phase entry, poor Vβ RSSs limit the possibility of further Vβ recombination on the second allele. Additional factors, including inhibition of RC formation, Vβ accessibility, and locus contraction via stochastic interaction of alleles with nuclear lamina (Chan et al., 2013; Chen et al., 2018; Schlimgen et al., 2008), cooperate with poor Vβ RSSs to limit biallelic assembly as DN3 cells attempt Tcrb recombination. Finally, TCRβ signals promote genetic and epigenetic changes that silence Vβ recombination in DP cells where Tcra genes assemble (Jackson and Krangel, 2005; Liang et al., 2002; Majumder et al., 2015; Skok et al., 2007). Notably, all features of this model could apply to IgH allelic exclusion.

The Broader Impacts of Vβ RSSs Being a Major Determinant of Vβ Recombination Frequency

The field has strived to elucidate mechanisms that promote V rearrangements across large chromosomal distances, with emphasis on factors that determine broad usage of V segments or promote allelic exclusion. In vivo experiments have shown that modulation of accessibility and RC contact of a V segment can influence its rearrangement frequency (Fuxa et al., 2004; Jain et al., 2018; Ryu et al., 2004). Correlative computational analyses conclude that V accessibility is the predominant factor for determining relative V utilization at Tcrb and Igh loci, while V RSS quality and RC contact each mainly function as a binary switch to prevent or allow recombination (Bolland et al., 2016; Gopalakrishnan et al., 2013). On the contrary, our data show that the qualities of RSSs flanking V2 and V31 function far beyond reaching a minimal threshold for functional synapsis and cleavage with Dβ RSSs. The increased usage of V2βR and V31βR at the expense of other Vβs on the same allele indicate that most, if not all, Vβs dynamically compete with each other for productive contact with the RC. On a normal allele, RAG bound to Dβ RSSs likely repeatedly capture and release different Vβ RSSs before synapsis proceeds to functional
cleavage (Wu et al., 2003). This sampling of \( V \) segments could occur via diffusional-based synapsis of \( V \) RSSs positioned within a cloud of spatial proximity (Ji et al., 2010) or by RAG chromosomal loop scanning-based synapsis (Hu et al., 2015; Jain et al., 2018). To determine RSS quality (strength), the field typically uses an algorithm that calculates a recombination information content (RIC) score, which is based on statistical modeling of how each nucleotide diverges from an averaged RSS (Cowell et al., 2002). The RIC scores of the RSSs we manipulated predict that the 3'D\( \beta \)1 RSS replacement would decrease \( V2 \) recombination and the variant 3'D\( \beta \)1 RSS substitution would reduce both \( V2 \) and \( V31 \) rearrangements (Table 2.4). These differences between predicted and empirical data could be due to a number of possibilities, including that the RIC algorithm does not address pairwise effects of RSSs on recombination. Regardless, the discrepancies between outcomes predicted by machine-generated associations and our actual \textit{in vivo} results highlight the critical need to experimentally test computationally-based models of V(D)J recombination. Additionally, we identify a role for weak V\( \beta \) RSSs that ensures TCR\( \beta \) repertoire diversity by ensuring the incorporation of D\( \beta \) segments during recombination.

Since the discovery of AgR allelic exclusion (Pernis et al., 1965), the field has worked to identify mechanisms and physiological roles for monoallelic expression of TCR and Ig genes. The predominant, long-standing theory is that the expression of only one type (specificity) of AgR by T and B cells suppresses autoimmunity by ensuring negative selection of cells expressing a self-reactive receptor (Brady et al., 2010a; Heath and Miller, 1993; Padovan et al., 1993). Consistent with this hypothesis, expression of a second endogenous AgR enables cells bearing a transgenic autoreactive AgR to evade negative selection (Auger et al., 2012; Illiev et al., 1994; Sarukhan et al., 1998; Zal et al., 1996), and biallelic TCR\( \alpha \) expression potentiates autoimmune diabetes in the NOD mouse model (Schuldt et al., 2017). Our findings suggest new avenues for investigation. As V\( \beta \) RSSs share sequence features with V\( \text{H} \) RSSs, but not V RSSs of other loci (Liang et al., 2002), experimentally determining if V\( \text{H} \) RSSs are poor for recombining with D\( \text{H} \) RSSs to thus restrain V\( \text{H} \) recombination and enforce IgH allelic exclusion is warranted. In addition, RSS
replacements could test the model that poor activities of Igλ RSSs, relative to Igκ RSSs, help mediate isotypic exclusion (Ramsden and Wu, 1991) where most B cells express Igκ or Igλ, but not both types of light chains (Bernier and Cebra, 1964). V2R+/V31+/R mice and the RSS replacement approach provide unprecedented experimental means to determine the effects of biallelic expression of diverse TCRβ chains in the αβ T cell population. RSS replacement mice also permit testing of possible additional reasons for controlling V rearrangements between alleles, including that monospecificity facilitates robust lymphocyte activation upon antigen encounter (Vettermann and Schlissel, 2010) and our view that asynchronous V-to-(D)J recombination between alleles suppresses RAG-triggered genomic instability and resultant lymphoid cancers.
Figure 2.1. Normal Gross αβ T cell development in V2 and V31 RSS replacement mice. (A) Schematic of the Tcrb locus and relative positions of V, D, and J segments, C exons, and the Eβ enhancer. Not drawn to scale. (B) Sequences of a consensus heptamer and nonamer and the 3′Dβ1, V2, and V31 RSSs. Differences relative to the consensus heptamer and nonamer are indicated in red. Differences of each Vβ RSS relative to the 3′Dβ1 RSS are underlined. (C) Sequence validation of the V2 or V31 RSS replacement with the 3′Dβ1 RSS. The 3′Dβ1 RSSs are highlighted in blue. (D) Total numbers of SP thymocytes and splenic αβ T cells (n ≥ 6 mice per group). (E and F) DN thymocyte development. Representative plots (E) and quantification (F) of DN cells. Gated on Lin−CD4−CD8−TCRβ− thymocytes (n ≥ 5 mice per group). (G and H) Global thymocyte development. Representative plots (G) and quantification (H) of DN, DP, CD4+, and CD8+ thymocytes (n = 5 mice per group). (I and J) Representative plots (I) and quantification (J) of SP αβ T cells in the spleen. Gated on TCRβ+ cells (n = 5 mice per group). [(B) one-way ANOVA followed by Dunnett’s post-tests comparing each RSS mutant to WT. (F, H, and J) two-way ANOVA followed by Dunnett’s post-tests for multiple comparisons, ns=not significant, *p<0.05, **p<0.01]. All quantification plots show mean ± SD. Data are compiled from at least 5 experiments.
Figure 2.2. The 3'Dβ1 RSS replacement increases V2 and V31 recombination in DN3 thymocytes. (A) Quantification of Dβ2-Jβ2.1 rearrangements by TaqMan qPCR in DN1/2 thymocytes (n = 3 mice per group). (B and C) Quantification of indicated Vβ rearrangements by TaqMan qPCR in DN1/2 (B) or DN3 (C) thymocytes (n = 3 mice per group, two-way ANOVA followed by Bonferroni’s post-tests for multiple comparisons). These data are compiled from 3 experiments.
Figure 2.3. Increased Utilization of 3'Dβ1 RSS-replaced Vβ Segments on αβ T Cells. (A) Representative plots of SP thymocytes expressing V2\(^+\), V31\(^+\), or V19\(^+\) TCRβ chains. (B-D) Quantification of V2\(^+\) (B), V31\(^+\) (C), or V19\(^+\) (D) SP thymocytes. (n = 5 mice per group, one-way ANOVA followed by Tukey's post-tests for multiple comparisons. ns=not significant, **p<0.01, ****p<0.0001). All quantification plots show mean ± SD. Data are compiled from 5 experiments.
Figure 2.4. Peripheral αβ T cells exhibit similar shifts in the Vβ repertoire in RSS replacement mice. (A) Representative plots of SP splenocytes expressing V2⁺, V31⁺, or V19⁺ TCRβ chains. (B, C, and D) Quantification of V2⁺ (B), V31⁺ (C), or V19⁺ (D) SP thymocytes (n = 5 mice per group, one-way ANOVA followed by Tukey’s post-tests for multiple comparisons). (E) Quantification of SP splenocytes expressing V4⁺, V12⁺, or V13⁺ TCRβ chains (n = 5 mice per group, two-way ANOVA followed by Tukey’s post-tests for multiple comparisons). (F) Ratio of the V2⁺ and V31⁺ Vβ repertoires. The fold change calculates the frequency of V2⁺ cells from V2⁺/R/R mice divided by V2⁺/R/+ mice. A similar calculation was made for V31⁺ cells from V31⁺/R/R and V31⁺/R/+ mice. All quantification plots show mean ± SD. Multiple post-tests are compared to WT unless indicated by bars, and p-values are corrected for multiple tests. ns=not significant, **p<0.01, ***p<0.001, ****p<0.0001. Data are compiled from 5 experiments.
Figure 2.5. Vβ RSS Replacement Alleles Compete with Normal Tcrb Alleles for Usage in the TCRβ Repertoire. (A and C) Representative plots of SP thymocytes expressing V2⁺ (A) or V31⁺ (C) TCRβ chains. (B and D) Quantification of V2⁺ (B) or V31⁺ (D) SP thymocytes (mean ± SD, n ≥ 3 mice per group, unpaired Student’s t test, ****p<0.0001). Data are compiled from 3 experiments.
**Table:**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>% Unique In-Frame Rearrangements</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>69.4%</td>
<td>0.011</td>
</tr>
<tr>
<td>V2Rho/V31Ro</td>
<td>60.2%</td>
<td>0.002</td>
</tr>
</tbody>
</table>

**Legend:**

- WT
- V2Rho/V31Ro
- Mix
- V2
- V31
- V2/R
- R/R
- R/+
Figure 2.6. 3'Dβ1 RSS-replaced Vβ Segments Increase Biallelic Tcrb Gene Expression. (A, C, and E) Representative plots of SP thymocytes expressing both V2+ and V19+ (A), V31+ and V19+ (C), or V2+ and V31+ (E) TCRβ chains. (B, D, and F) Quantification of SP thymocytes expressing the two indicated TCRβ chains. (F) V2R/ and V31R/ thymocytes were mixed 1:1 and analyzed. (B and D) n = 5 mice per group, two-way ANOVA followed by Dunnett’s post-tests for multiple comparisons. (F) n ≥ 3 mice per group, one-way ANOVA followed by Tukey’s post-tests for multiple comparisons. (G) Quantification of double-staining SP thymocytes for each Vβ combination tested (n = 5 mice per group, one-way ANOVA followed by Dunnett’s post-tests comparing each RSS mutant to WT). (H) Schematic of TCRβ repertoire sequencing by Adaptive Biotechnologies platform. (I) Models of monoallelic versus biallelic Tcrb gene assembly and predicted fractions of in-frame (IF) and out-of-frame (OF) rearrangements. (J) Percent of unique TCRβ rearrangements sequenced from sorted CD4+ SP thymocytes. Frame types indicated: “In” = in-frame, “Out” = out-of-frame, “Stop” = a rearrangement that produces a STOP codon (n = 3 mice per group, unpaired Student’s t test). (K) Averages and standard deviations (SD) of in-frame rearrangements from sorted WT and V2R/V31+ CD4+ thymocytes. All quantification plots show mean ± SD. ns=not significant, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. Data are compiled from 5 experiments.
**Figure 2.7.** αβ T cells exhibiting biallelic Tcrb gene expression seed the periphery. (A, C, and E) Representative plots of SP splenocytes expressing both V2⁺ and V19⁺ (A), V31⁺ and V19⁺ (C), or V2⁺ and V31⁺ (E) TCRβ chains. (B, D, and F) Quantification of SP splenocytes expressing the two indicated TCRβ chains. (n = 5 mice per group, two-way ANOVA followed by Tukey’s post-tests for multiple comparisons [B and D], one-way ANOVA followed by Dunnett’s post-tests comparing each RSS mutant to WT [F]). (G) Quantification of double-staining SP splenocytes for each Vβ combination tested. (H) Depiction of the recombination events that could result in two TCRβ chains expressed from one allele. RSSs indicated as triangles. (n = 5 mice per group, two-way ANOVA). All quantification plots show mean ± SD. Multiple post-tests are compared to WT unless indicated by bars, and p-values are corrected for multiple tests. ns=not significant, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. Data are compiled from 5 experiments.
Figure 2.8. 3’Dβ1 RSS Substitutions Increase Vβ usage and TCRβ Allelic Inclusion Independent of c-Fos binding. (A) Sequences of the normal and inactivated c-Fos binding site in the 3’Dβ1 RSS. The A→C and T→C mutations are indicated in red. (B) Sequence validation of the V2 or V31 RSS replacement with the variant 3’Dβ1 RSS. The inactivated c-Fos binding site is highlighted in blue and mutated nucleotides indicated by red arrows. (C) Representative plots of SP thymocytes expressing V2+ or V31+ TCRβ chains. (D and E) Quantification of V2+ (D) or V31+ (E) SP thymocytes (n ≥ 5 mice per group, one-way ANOVA followed by Tukey’s post-tests for multiple comparisons). (F) Representative plots of SP thymocytes expressing both V2+ and V31+ TCRβ chains. (G) Quantification of SP thymocytes expressing both V2+ and V31+ TCRβ chains (n ≥ 5 mice per group, one-way ANOVA followed by one-way ANOVA followed by Dunnett’s post-tests comparing each RSS mutant to WT). All quantification plots show mean ± SD. ns=not significant, *p<0.05, ****p<0.0001. Data are compiled from 3 experiments.
Figure 2.9. 3’Dβ1 RSS Substitutions Increase Vβ recombination before Enforcement of Feedback Inhibition. (A and C) Representative plots of SP thymocytes expressing V2⁺ (A), V31⁺ (C), or V13⁺ TCRβ chains. (B and D) Quantification of SP thymocytes expressing V2⁺ (B) or V31⁺ (D) TCRβ chains (n = 3 mice per group, unpaired Student’s t test, ****p<0.0001). All quantification plots show mean ± SD. Data are compiled from 3 experiments.
Figure 2.10. 3’Dβ1 RSS Substitutions Permits Direct Vβ-to-Jβ Recombination and Decreases TCRβ Diversity. (A) Total CDR3β lengths in WT, V2R/R, and V31R/R pre-selection thymocytes. CDR3β sequences are binned according to whether Dβ1 or Dβ2 sequences are identifiable. (B) Shared CDR3β sequences between two individual mice, their relative frequencies, and color-coded according to whether Dβ1 or Dβ2 sequences are evident. CD4⁺CD8⁺CD69⁻ TCRβlo pre-selection thymocytes were sorted for sequencing. Representative plots from 3 independent experiments.
Table 2.1. Analysis of Tcrb rearrangements in αβ T cell hybridomas

<table>
<thead>
<tr>
<th>Genotype</th>
<th>( V_{2}^{DR}/V_{31}^{HR} )</th>
<th>WT*</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>% Total</td>
<td>Number</td>
</tr>
<tr>
<td>Clonal Hybridomas Assayed</td>
<td>102</td>
<td>--</td>
<td>212</td>
</tr>
<tr>
<td>Rearrangement Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V(D)J / DJ</td>
<td>46</td>
<td>45.1</td>
<td>120</td>
</tr>
<tr>
<td>V(D)J / V(D)J</td>
<td>32</td>
<td>31.4</td>
<td>92</td>
</tr>
<tr>
<td>V(D)J-V(D)J / DJ</td>
<td>10</td>
<td>9.8</td>
<td>0</td>
</tr>
<tr>
<td>V(D)J-V(D)J / V(D)J</td>
<td>14</td>
<td>13.7</td>
<td>0</td>
</tr>
<tr>
<td>Monoallelic V(D)J</td>
<td>56</td>
<td>54.9</td>
<td>120</td>
</tr>
<tr>
<td>Biallelic V(D)J</td>
<td>46</td>
<td>45.1</td>
<td>92</td>
</tr>
<tr>
<td>1 V(D)J</td>
<td>46</td>
<td>45.1</td>
<td>120</td>
</tr>
<tr>
<td>2 V(D)J</td>
<td>42</td>
<td>41.2</td>
<td>92</td>
</tr>
<tr>
<td>3 V(D)J</td>
<td>14</td>
<td>13.7</td>
<td>0</td>
</tr>
<tr>
<td>V2(D)J</td>
<td>13</td>
<td>12.7</td>
<td></td>
</tr>
<tr>
<td>V31(D)J</td>
<td>58</td>
<td>56.9</td>
<td></td>
</tr>
<tr>
<td>V31-to-DJ</td>
<td>33</td>
<td>32.4</td>
<td></td>
</tr>
<tr>
<td>V31-to-J</td>
<td>24</td>
<td>23.5</td>
<td></td>
</tr>
<tr>
<td>V2(D)J / V31(D)J</td>
<td>8</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td>V2(D)J(^{IF}) / V31(D)J(^{IF})</td>
<td>2</td>
<td>2.0</td>
<td></td>
</tr>
</tbody>
</table>

IF = In-Frame
* = Khor and Sleckman, 2005
\( p \)-values were generated by Pearson’s \( \chi^2 \) test with Yates’ continuity correction.
Table 2.2. Sequence analysis of the V2 and V31 rearrangements on opposite alleles in V2<sup>ab</sup>/V31<sup>ab</sup> T cell hybridomas

<table>
<thead>
<tr>
<th>Clone</th>
<th>V2</th>
<th>N/P</th>
<th>Potential Dβ</th>
<th>N/P</th>
<th>Jβ</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone 1</td>
<td>V2</td>
<td>AGCAGCCAAGA</td>
<td>TA</td>
<td>A</td>
<td>T</td>
<td>CCGGGCAG</td>
</tr>
<tr>
<td>V31</td>
<td>GCCTGGAGTCT</td>
<td>AC</td>
<td>CCAACGAAAGA</td>
<td>TGG</td>
<td>GAACAG</td>
<td>2.7</td>
</tr>
<tr>
<td>Clone 23</td>
<td>V2</td>
<td>AGCAGCCAAGA</td>
<td>T</td>
<td>TGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clone 54</td>
<td>V2</td>
<td>AGCAGCC</td>
<td>CAA</td>
<td>GAC</td>
<td>AAGACGGGCGCAG</td>
<td>2.2</td>
</tr>
<tr>
<td>V31</td>
<td>GCCTGGAGTCT</td>
<td>A</td>
<td>GC</td>
<td>G</td>
<td>TGAACAG</td>
<td>2.7</td>
</tr>
<tr>
<td>Clone 55</td>
<td>V2</td>
<td>AGCAGCC</td>
<td>T</td>
<td>TGG</td>
<td>CTCCTATGAACAG</td>
<td>2.7</td>
</tr>
<tr>
<td>V31</td>
<td>GCCTGGAGTCTC</td>
<td>A</td>
<td>GGGACAGGGGGG</td>
<td>GCGG</td>
<td>GAACAG</td>
<td>2.7</td>
</tr>
<tr>
<td>Clone 59</td>
<td>V2</td>
<td>AGCAGGCC</td>
<td>TC</td>
<td>GG</td>
<td>AGTCAAAAACACCTTG</td>
<td>2.4</td>
</tr>
<tr>
<td>V31</td>
<td>GCCT</td>
<td>GGGC</td>
<td>CACCTTG</td>
<td>2.4</td>
<td></td>
<td>Out-of-Frame</td>
</tr>
<tr>
<td>Clone 62</td>
<td>V2</td>
<td>V31-to-5'D/1' RSS hybrid join, deleting all subsequent Jβ1-Cβ1 and Dβ2/Jβ2-Cβ2 sequences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V31</td>
<td>V31 rearranged to the Vβ coding sequence of an out of frame V29/Dβ/Jβ2 rearrangement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sequences are underlined every 3 nucleotides to orient the reading frame.

Dβ nucleotides in bold are from Dβ1, italicized nucleotides are from Dβ2, and for V2 rearrangements unformatted nucleotides could be from either.
<table>
<thead>
<tr>
<th>Genotype and Source</th>
<th>$Tcrb^{Tg}\times$</th>
<th>$V_{21}^{R\downarrow}$</th>
<th>$Tcrb^{Tg}V_{21}^{R}$</th>
<th>$V_{31}^{R\downarrow}$</th>
<th>$Tcrb^{Tg}V_{31}^{R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>%</td>
<td>Number</td>
<td>%</td>
<td>Number</td>
</tr>
<tr>
<td>Clonal Hybridomas Assayed</td>
<td>129</td>
<td>100.0</td>
<td>18</td>
<td>100.0</td>
<td>56</td>
</tr>
<tr>
<td>V(D)J</td>
<td>12</td>
<td>9.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V2(D)J</td>
<td></td>
<td></td>
<td>9</td>
<td>50.0</td>
<td>0</td>
</tr>
<tr>
<td>V31(D)J</td>
<td>3</td>
<td>2.3</td>
<td></td>
<td></td>
<td>66</td>
</tr>
<tr>
<td>V31-to-DJ</td>
<td>3</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V31-to-J</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V31(D)J on one allele</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V31-to-DJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V31-to-J</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* = Brady et al. 2010
Table 2.4. Recombination information content (RIC) scores of RSSs in this study

<table>
<thead>
<tr>
<th></th>
<th>RIC Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>V2 RSS</td>
<td>-32.1365</td>
</tr>
<tr>
<td>V31 RSS</td>
<td>-37.2563</td>
</tr>
<tr>
<td>3’Dβ1 RSS</td>
<td>-35.7342</td>
</tr>
<tr>
<td>Variant 3’Dβ1 RSS</td>
<td>-38.5529</td>
</tr>
</tbody>
</table>

RIC scores generated by RSSSite (https://www.itb.cnr.it/rss/)

Predicted order of RSS activity:

V2 > 3’Dβ1 > V31 > Variant 3’Dβ1
CHAPTER 3: Inefficient V(D)J Recombination Underlies Monogenic T Cell Receptor β Expression

Abstract

The assembly of TCR and Ig genes by V(D)J recombination generates the AgR diversity that is vital for adaptive immunity. At most AgR loci, V(D)J recombination is regulated so that only one allele assembles a functional gene, ensuring that nearly every T and B cell expresses a single type, or specificity, of AgR. The genomic organizations of some AgR loci permit the assembly and expression of two distinct genes on each allele; however, this is prevented by undetermined mechanisms. We show that the poor qualities of RSSs flanking Vβ gene segments suppress the assembly and expression of two distinct TCRβ genes from a single allele. Our data demonstrate that an intrinsic genetic mechanism that stochastically limits Vβ recombination efficiency governs monogenic TCRβ expression, thereby restraining the expression of multiple AgRs on αβ T cells.

Introduction

The vast diversity of AgRs expressed by T and B lymphocytes is essential for effective adaptive immunity. A TCR is comprised of TCR β and α or γ and δ proteins, whereas a B cell receptor (or antibody) consists of IgH and Igκ or Igλ proteins. Developing T and B cells assemble TCR or Ig genes, respectively, through lymphocyte lineage- and developmental stage-specific recombination of variable (V), diversity (D), and joining (J) gene segments (Bassing et al., 2002; Schatz and Swanson, 2011). Each V(D)J rearrangement and downstream constant (C) region exons comprise a TCR or Ig gene, which encodes protein only if the V and J segments are recombined in-frame. The number of possible joining combinations and imprecise repair of V(D)J coding joins produce enormous AgR gene diversity.

V(D)J recombination poses a challenge for highly specific adaptive immune responses because both alleles of a locus could assemble an in-frame gene, producing a lymphocyte with
two distinct AgRs. To achieve monospecificity, V gene segment rearrangements at TCR and Ig loci are regulated between alleles to enforce monoallelic AgR expression (Brady et al., 2010b; Mostoslavsky et al., 2004). This allelic exclusion is achieved by sequential initiation of V recombination between alleles and subsequent feedback inhibition of V recombination signaled by protein expressed from the in-frame V(D)J rearrangement (Brady et al., 2010b; Mostoslavsky et al., 2004). Notably, the genomic organizations of mammalian TCRβ, TCRγ, TCRδ, and Igλ loci, as well as Ig loci of cartilaginous fish, permit the assembly and expression of multiple genes from a single allele, providing an additional obstacle to achieve monospecificity. Currently, there are no reported mechanisms by which V rearrangements are regulated on individual alleles to achieve monogenic protein expression.

TCRβ loci consist of 23 functional Vβs that recombine by deletion to either one of two downstream Dβ-Jβ-Cβ clusters, and another Vβ (V31) located further downstream that rearranges through inversion to either Dβ-Jβ-Cβ cluster (Figure 3.1, A)(Malissen et al., 1986). Theoretically, each TCRβ allele can assemble and express two distinct genes – one involving an upstream Vβ and another involving V31. However, this has not been observed at any detectable level (Khor and Sleckman, 2005; Lee and Bassing, 2020), indicating that mechanisms control Vβ recombination on each allele to ensure monogenic TCRβ recombination and expression. The semi-conserved RSSs that flank AgR locus gene segments target V(D)J recombinase activity and direct specific V(D)J rearrangements (Bassing et al., 2002). For TCRβ, the poor qualities of Vβ RSSs stochastically restrain the number of Vβ rearrangements before TCRβ protein-signaled feedback inhibition, thereby allowing for monoallelic assembly and expression of functional TCRβ genes (Wu et al., 2020). To determine if poor Vβ RSSs also limit the assembly and expression of two different TCRβ genes from the same allele, we generated C57BL/6 mice carrying replacements of both a V2 RSS and a V31 RSS on the same allele with the stronger 3'Dβ1 RSS (the V2R/V31R allele).
Results

We studied wild-type (WT), heterozygous V2R31R/WT, and homozygous V2R31R/V2R31R mice. The mutant mice had normal numbers and frequencies of mature splenic αβ T cells and thymocytes at each developmental stage (data not shown). Due to the lack of congeneric markers, TCRβ proteins cannot be identified by the allele that encodes them, nor whether they include Cβ1 versus Cβ2 regions. Thus, we performed flow cytometry using anti-V2 and anti-V31 antibodies to quantify cells expressing V2+ and V31+ TCRβ proteins. We assayed CD4+ and CD8+ single-positive (SP) thymocytes as they are mature and naïve αβ T cells. Reflecting published data (Wu et al., 2020; Steinel et al., 2014), we detected a small fraction (0.11%) of cells that stained with both antibodies in WT mice (Figure 3.1, B and C), which is consistent with a small population of V2’V31+ αβ T cells. We observed a 12.4-fold increased fraction of these cells in V2R31R/WT mice, and a 32.8-fold increase in V2R31R/V2R31R mice (Figure 3.1, B and C). These elevated frequencies of dual-TCRβ+ cells corresponded with the greater utilizations of V2 and V31 in expressed TCRβ chains (Figure 3.1, D-F). These data demonstrate that enhancing RSS quality of two Vβs on the same allele increases their rearrangement and consequently the fraction of T cells expressing two distinct types of TCRβ proteins. As the Vβ repertoire of SP thymocytes reflects the relative levels that individual Vβ segments recombine (Wilson et al., 2001), the preferential usage of V31 over V2 reveals that V31R outcompetes V2R for rearrangement. This could be due to greater accessibility of V31 (Gopalakrishnan et al., 2013) or interaction of V31 with Dβ-Jβ segments before TCRβ locus contraction places V2 near Dβ-Jβ segments. Notably, the higher than 2-fold increase of these dual-TCRβ+ cells in V2R31R/V2R31R mice compared to V2R31R/WT mice implies that two distinct V(D)Jβ rearrangements can contribute to TCRβ expression from the same allele.

The possibility that one TCRβ allele could encode two distinct V(D)Jβ rearrangements is further supported by next-generation TCRβ sequencing. Adaptive Biotechnologies’ Immunoseq amplifies all possible V(D)Jβ rearrangements from a genomic DNA sample, but does not amplify
D\(\beta\) segments, J\(\beta\) segments, or DJ\(\beta\) joins (Carlson et al., 2013). When allelic exclusion is normal \(~60\%\) of cells have one in-frame (IF) VDJ\(\beta\) rearrangement and \(~40\%\) of cells have two, where one rearrangement is IF and the other is out-of-frame (OF, Figure 3.1, G). Since only V(D)J\(\beta\) rearrangements are sequenced, 71.4\% of VDJ\(\beta\) rearrangements would be IF and 28.6\% of rearrangements would be OF (Figure 3.1, G). In the complete absence of mechanisms that promote allelic exclusion, 20\% of \(\alpha\beta\) T cells possess two IF VDJ\(\beta\) rearrangements and 80\% of cells will possess one IF and one OF VDJ\(\beta\) (Figure 3.1, G). In this case, sequencing would reveal that 60\% of VDJ\(\beta\) rearrangements in a population would be IF (Figure 3.1, G). We isolated DNA from sorted WT and V2\(R^R\)/V2\(R^R\) CD4\(^+\) thymocytes and submitted them for sequencing. We observe that 69.4\% of unique VDJ\(\beta\) rearrangements are IF in WT cells, which is consistent with a theoretical 71.4\% of IF rearrangements if most cells assembled and expressed TCR\(\beta\) from one allele (Figure 3.1, H and I). Sequence analysis of V2\(R^R\)/V2\(R^R\) CD4\(^+\) thymocytes reveals that 52.2\% of VDJ\(\beta\) rearrangements are IF, which is lower than the theoretical 60\% (Figure 3.1, H and I). A dip below 60\% of IF VDJ\(\beta\) rearrangements is only possible if multiple rearrangements and possible proteins are permitted per allele. Given that the genomic organization of the TCR\(\beta\) locus can permit two TCR\(\beta\) genes to be assembled on one allele, the sequencing data from V2\(R^R\)/V2\(R^R\) CD4\(^+\) thymocytes suggest that some cells are attempting three or possibly four V\(\beta\)-to-DJ\(\beta\) rearrangements.

To determine whether a single TCR\(\beta\) allele can indeed support expression of TCR\(\beta\) proteins from two different V(D)J\(\beta\) rearrangements, by flow cytometry we analyzed mice where one TCR\(\beta\) allele is inactivated by deletion of the TCR\(\beta\) enhancer (E\(\beta\)) (Bories et al., 1996; Bouvier et al., 1996). We assayed mice carrying the E\(\beta\)-deleted allele opposite a WT allele, an allele with an RSS replacement of either V2 (V2\(R^R\)) or V31 (V31\(R^R\)), or both (Wu et al., 2020). We detected a small percentage (0.094\%) of V2\(^-\)V31\(^+\) SP thymocytes in WT/E\(\beta\)\(\Delta\) mice (Figure 3.2, A and B), potentially representing a rare population of cells expressing two different TCR\(\beta\) proteins from the same WT allele. Regardless, we observed V2\(^-\)V31\(^+\) cells at a 1.9-fold greater frequency in
V2\textsuperscript{R}/E\textbetaΔ mice and at a 4.8-fold greater frequency in V31\textsuperscript{R}/E\textbetaΔ mice (Figure 3.2, A and B). Thus, enhancing the quality of either V\textbeta RSS elevates the fraction of cells expressing both V2\textsuperscript{+} and V31\textsuperscript{+} TCR\textbeta proteins. Notably, we detected a 14.4-fold increased frequency of V2\textsuperscript{+}V31\textsuperscript{+} cells in V2\textsuperscript{R}V31\textsuperscript{R}/E\textbetaΔ mice relative to WT/E\textbetaΔ mice (Figure 3.2, A and B), indicating that enhancing quality of two V\textbeta RSSs synergistically increases the percentage of cells expressing both V2\textsuperscript{+} and V31\textsuperscript{+} TCR\textbeta proteins. Indeed, deleting part of the V31 RSS on the V2\textsuperscript{R} allele (the V2\textsuperscript{R}31\textsuperscript{Δ} allele, Figure 3.2, C) dramatically reduces the frequency of V2\textsuperscript{+}V31\textsuperscript{+} cells to levels that are equivalent or less than that in V2\textsuperscript{R}/E\textbetaΔ mice (0.178% versus 0.135%, Figure 3.2, A and B). Collectively, these data confirm that the V2\textsuperscript{R}31\textsuperscript{R} allele promotes expression of two distinct TCR\textbeta proteins from two different V(D)J\textbeta rearrangements on a single allele.

Discussion

Our study demonstrates that an intrinsic genetic mechanism governs monogenic TCR\textbeta assembly and expression. We show that poor quality V\textbeta RSSs cooperate to limit assembly and expression of two distinct TCR\textbeta genes from one allele. We previously showed that poor quality V\textbeta RSSs stochastically restrain V\textbeta recombination frequency before feedback inhibition to decrease biallelic assembly and expression of TCR\textbeta genes (Wu et al., 2020). We now further conclude that low quality V\textbeta RSSs also lower the incidence that both V31 and an upstream V\textbeta recombine on the same allele. These rearrangements could involve either: 1) a deletional V2 rearrangement to the D\textbeta1-J\textbeta1-C\textbeta1 cluster and an inversional V31 rearrangement to the D\textbeta2-J\textbeta2-C\textbeta2 cluster, or 2) an inversional V31 rearrangement to the D\textbeta1-J\textbeta1-C\textbeta1 cluster, which inverts a portion of the locus that contains the D\textbeta2-J\textbeta2-C\textbeta2 cluster, and then an inversional V2 rearrangement to the D\textbeta2-J\textbeta2-C\textbeta2 cluster (Khor and Sleckman, 2005; Lee and Bassing, 2020) (Figure 3.2, D). To achieve monogenic TCR\textbeta assembly and expression, this RSS-based genetic mechanism might function with epigenetic processes that have been implicated to enforce monoallelic V\textbeta recombination. For example, it has been proposed that dynamic interactions of V\textbeta
segments with the nuclear lamina lowers Vβ recombination efficiency by repressing Vβ chromatin accessibility and chromosome looping between Dβ-Jβ clusters and upstream Vβ segments (Chen et al., 2018; Schlimgen et al., 2008). In this context, poor-quality Vβ RSSs could lower the likelihood that two Vβ rearrangements occur on an allele when V3f and an upstream Vβ segment are both accessible and the upstream Vβ is looped in proximity with Dβ-Jβ segments. Thus, the properties of RSSs may underlie monogenic assembly and expression of mammalian TCRγ, TCRδ, and Igλ proteins in mammals, and Ig proteins in cartilaginous fish. Additionally, V RSSs may similarly contribute to Igκ and Igλ isotypic exclusion in B cells.
Figure 3.1. Increased Frequency of Dual TCRβ⁺ Cells and Altered Vβ Repertoire in Mice with Two Vβ RSS Replacements on the Same Allele. (A) Schematic of the TCRβ locus and relative positions of Vβ, Dβ, and Jβ segments, Cβ exons, and the Eβ enhancer. (B and C) Representative plots (B) and quantification (C) of SP thymocytes expressing both V2⁺ and V31⁺ TCRβ chains. (D, E, and F) Representative plots (D) and quantification of SP thymocytes expressing V2⁺ (E) or V31⁺ (F) TCRβ chains. n ≥ 4 mice per group, one-way ANOVA, Dunnett’s multiple post-tests comparing each RSS mutant to WT. (G) Models of monoallelic versus biallelic Tcrb gene assembly and predicted fractions of in-frame (IF) and out-of-frame (OF) rearrangements. (H) Percent of unique TCRβ rearrangements sequenced from sorted CD4⁺ SP thymocytes. Frame types indicated: “In” = in-frame, “Out” = out-of-frame, “Stop” = a rearrangement that produces a STOP codon (n = 3 mice per group, unpaired Student’s t test). (I) Averages and standard deviations (SD) of in-frame rearrangements from sorted WT and V2R31R/V2R31R CD4⁺ thymocytes. ***p<0.001, ****p<0.0001. Data are compiled from 5 experiments.
Figure 3.2. Expression of Two Different TCRβ Chains from the V2\(^R\)V31\(^R\) Allele. (A and B) Representative plots (A) and quantification (B) of SP thymocytes expressing both V2\(^R\) and V31\(^R\) TCRβ chains (n ≥ 6 mice per group, one-way ANOVA, Tukey’s multiple post-tests, ns=not significant, ****p<0.0001). (C) Schematic of the sense strand and truncation of the V31 region of the V2\(^R\)31Δ allele, the V31 RSS indicated in blue. (D) Depiction of the recombination events that could result in two TCRβ chains expressed from one allele. RSSs indicated as triangles. Data are compiled from 4 experiments.
CHAPTER 4: Weak Vβ Recombination Signal Sequences Cooperate with the DNA Damage Response ATM Kinase to Ensure Monoallelic Expression and Broad Repertoire of TCRβ Proteins

Abstract

Monoallelic expression (allelic exclusion) of diverse antigen receptor genes is achieved through incompletely known mechanisms that govern monoallelic initiation and subsequent feedback inhibition of V gene segment rearrangements. At TCRβ loci, Vβ gene segments rearrange to distal Dβ-Jβ segments via deletion of intervening sequences, except for the Dβ-Jβ-proximal V31 segment that rearranges by inversion. All Vβ segments have low-quality RSSs that can impose allelic exclusion by stochastically limiting biallelic initiation of Vβ rearrangements before TCRβ protein expressed from one allele signals to permanently block Vβ recombination on the other allele. The DNA damage response ATM protein kinase that facilitates V(D)J joining and transiently suppresses subsequent V(D)J rearrangements also imposes allelic exclusion. We show here that ATM functions with the poor-quality V31 RSS to preserve V31 representation in the TCRβ repertoire by hindering inversional V31 rearrangements from aberrantly resolving as non-functional deletions. This function for ATM in preserving proper V(D)J joining is extended to the Igκ locus as ~1/3 of Vκ segments rearrange by inversion. We also show that the weak RSSs of the distal V1 and V2 segments cooperate with each other and with ATM to enforce allelic exclusion by limiting initiation of long-range Vβ recombination on both alleles before TCRβ-signaled feedback inhibition perpetually silences Vβ recombination. Our findings demonstrate that the weak V(D)J recombinase targeting elements that render Vβ rearrangements inefficient and the ATM protein that shepherds V(D)J recombinase activity cooperate to create a diverse repertoire of monoallelically expressed TCRβ proteins.
Introduction

The ability of T and B cell populations to express diverse antigen receptors is essential for adaptive immunity. The RAG endonuclease catalyzes assembly of TCR and Ig variable region exons through recombination of variable (V), diversity (D), and joining (J) gene segments (Schatz and Swanson, 2011). Semi-conserved RSSs that flank each segment target RAG endonuclease activity and influence levels of recombination (Schatz and Swanson, 2011). RAG induces DNA DSBs between the RSS and coding sequence of two segments, yielding coding and signal ends (Schatz and Swanson, 2011). RAG functions with the ATM DSB response kinase and other proteins to repair these DNA ends and form coding and signal joins (Bednarski and Sleckman, 2012). TCRβ, IgH, and Igκ recombination is regulated between alleles so that most T and B cells exhibit monoallelic expression (allelic exclusion) of these genes (Brady et al., 2010b). Allelic exclusion is achieved through sequential initiation of V rearrangements between alleles, transient inhibition of further V(D)J recombination triggered by RAG DSBs, and permanent silencing of V recombination signaled by protein expressed from a functional V(D)J rearrangement (Brady et al., 2010b).

The mouse TCRβ locus consists of 23 Vβ segments located upstream of two Dβ-Jβ-Cβ clusters and V31 located downstream in an opposite orientation as other coding sequences (Figure 4.1, A). Distal Vβs rearrange to Dβ-Jβ complexes via deletion of intervening sequences, while V31 rearrangements proceed solely via inversion and sequences are retained in the chromosome (Malissen et al., 1986). Two factors have been documented to enforce TCRβ allelic exclusion by limiting Vβ rearrangements on both alleles before protein from one allele signals feedback inhibition. Vβs have poor-quality RSSs that, at least for V31 and the distal V2 segment, enforce allelic exclusion by limiting Vβ recombination, though mainly of V31, before TCRβ-signaled feedback inhibition (Wu et al., 2020a; Wu and Bassing, 2020b). In response to DSBs, ATM signals repression of RAG expression and V(D)J recombination until DSB repair halts ATM kinase activity (Steinel et al., 2013; Fisher et al., 2017). The inactivation of ATM in mice increases
the frequencies of biallelic V recombination and protein expression at TCRβ, IgH, and Igκ loci (Steinel et al., 2013; Steinel et al., 2014), indicating that ATM is a pervasive factor in enforcing allelic exclusion. To determine whether weak Vβ RSSs generally cooperate with each other and ATM to limit Vβ rearrangements and impose TCRβ allelic exclusion, we generated and analyzed ATM-sufficient or -deficient mice containing replacement of the V31, V2, or most distal V1 RSS on opposite alleles with a higher-quality RSS (Figure 4.1, A).

Results

We started by determining effects of ATM inactivation on Vβ repertoire and TCRβ allelic exclusion in our mice carrying replacement of the V2 and V31 RSS on opposite alleles with the better 3’Dβ1 RSS (Figure 4.1, A) (Wu et al., 2020a). We bred V2R+/V31R+ mice with Atm+/− mice to produce and analyze wild-type (WT), Atm−/−, V2R+/V31R−, and V2R+/V31R−:Atm− mice. The percentages of T cells expressing V2+; V31+, or both V2+ and V31+ TCRβ proteins were similar between WT and Atm−/− mice, but higher in V2R+/V31R− mice (Figure 4.1, B-E). In V2R+/V31R−:Atm− mice as compared to V2R+/V31R+ mice, the frequency of V2+ cells was 1.4-fold greater (46.7% versus 33.4%); however, in marked contrast, the fraction of V31+ cells were 0.66-fold lower (27% versus 41.2%) (Figure 4.1, B and C). Reflecting fewer V31+ cells, the fraction of V2+V31+ cells was lower in V2R+/V31R−:Atm− mice relative to V2R+/V31R+ mice (2% versus 2.7%) (Figure 4.1, D and E). While the elevated fraction of V2+ cells is consistent with loss of ATM-mediated feedback inhibition, the lower frequencies of V31+ and V2+V31+ cells can only be explained by loss of ATM DSB repair functions. In the absence of ATM, some RAG-liberated DNA ends dissociate and fail to form coding or signal joins (Bredemeyer et al., 2006). Escape of 3’Dβ1 RSS ends during V31R− recombination could permit hybrid joins between V31 coding ends and 5’Dβ RSS ends, effectively deleting Dβ-Jβ-Cβ sequences and precluding TCRβ gene assembly (Figure 4.1, F). In contrast, escape of 3’Dβ1 RSS ends during V2R− recombination would not block functional Vβ rearrangements (Figure 4.1, F). To test this, we performed Taqman qPCR
to quantify hybrid joins between V31 and the 5’Dβ1 RSS or 5’Dβ2 RSS in V2R+/V31R/R or V2R+/V31R/R. Atm− mice. Hybrid joins were increased 14-18-fold in the absence of ATM (Figure 4.1, G), indicating that ATM facilitates successful completion of V31R-to-DJβ inversional rearrangements by hindering their resolution as deletions.

A potential extension of the observation that ATM restricts V31 hybrid join formation is that ATM protects the composition of antigen receptor repertoires by ensuring V gene segments that rearrange by inversion are represented. This would have possible implications for the Igκ locus as ~1/3 of Vκ segments rearrange by inversion (Matheson et al., 2017). In germline configuration, Vκ segments 10-96, 19-93, 4-81, and 12-46 are some of the most represented in the Igκ repertoire and are in an opposite orientation dictating that they rearrange by inversion (Matheson et al., 2017). To test the role of ATM in productively resolving inversional rearrangements as VJκ coding joins, we performed Taqman qPCR to quantify hybrid joins between these Vκ segments and Jκ1 in WT and Atm− sorted pre-B cells (Figure 4.1, H). In all cases, Vκ RSS - Jκ1 hybrid joins were difficult to detect in WT pre-B cells, with only a single biological replicate providing signal for Vκ4-81 and Vκ12-46 (Figure 4.1, I). In Atm− sorted pre-B cells these Vκ RSS - Jκ1 hybrid joins were readily detected and, when compared to the single data point for WT cells, are 12.1-fold and 16.9-fold greater for Vκ4-81 and Vκ12-46, respectively (Figure 4.1, I). These data highlight a general function of ATM that ensures inversional V rearrangements are not lost as deletional hybrid joins. However, the anomalous V31R-to-5’Dβ RSS rearrangements in V2R+/V31R/R. Atm− mice preclude conclusion whether weak Vβ RSSs cooperate with ATM to constrain Vβ rearrangements and thereby enforce TCRβ allelic exclusion.

To address this question, we studied mice carrying RSS substitutions of two different distal Vβs. We replaced the most distal Vβ (V1) RSS with the stronger 3’Dβ1 RSS (Figure 4.1, A) (Jung et al., 2003), and bred these V1R+/ mice with V2R+/R and Atm− mice to obtain and analyze WT, Atm−, V1R+/V2R+/R, and V1R+/V2R+/R. Atm− mice. In V1R+/V2R+/R mice compared to WT mice, the frequency of V1+ cells was 2-fold higher (12% versus 6%) and the incidence of V2+ cells was 6.2-
fold greater (38% versus 6.2%) (Figure 4.2, A and B). Moreover, V1\(^{+}\)V2\(^{+}\) cells were 7.5-fold higher in \(V1^{R+/V2^{+R}}\) mice relative to WT mice (0.3% versus 0.04%) (Figure 4.2, A and B). In the only \(Atm^{-}\) mouse that we were able to obtain, we found normal incidences of V1\(^{+}\), V2\(^{+}\), and V1\(^{+}\)V2\(^{+}\) cells (Figure 4.2, C and D). Notably, the frequencies of V1\(^{+}\) and V2\(^{+}\) cells were higher in \(V1^{R+/V2^{+R}:Atm^{-}}\) mice relative to \(V1^{R+/V2^{+R}}\) mice (16.1% versus 12% for V1, and 40.1% versus 38.5% for V2) (Figure 4.2, C and D). Moreover, V1\(^{+}\)V2\(^{+}\) cells were 1.6-fold more prevalent in \(V1^{R+/V2^{+R}:Atm^{-}}\) mice compared to \(V1^{R+/V2^{+R}}\) mice (0.52% versus 0.33%) (Figure 4.2, C and D).

Collectively, these data indicate that the weak RSSs of the distal V1 and V2 segments cooperate with ATM, and with each other, to enforce TCR\(\beta\) allelic exclusion by constraining biallelic V\(\beta\) recombination.

**Discussion**

Our study provides new insights into mechanisms that regulate the monoallelic assembly of diverse antigen receptor genes. We previously showed that the weak V2 and V31 RSSs cooperate to limit cells that express V31\(^{+}\) and V2\(^{+}\) TCR\(\beta\) proteins from opposite alleles, mainly by curbing proximal V31 rearrangements before protein from V2 recombination signals feedback inhibition (Wu et al., 2020a; Wu and Bassing, 2020b). By enhancing RSS quality of two distal V\(\beta\)s, we show that weak V\(\beta\) RSSs serve a general major role in enforcing TCR\(\beta\) allelic exclusion by restraining distal V\(\beta\) rearrangements on both alleles before protein from one allele signals permanent feedback inhibition. We show that this genetic mechanism cooperates with the ATM DSB response protein to further limit biallelic V\(\beta\) recombination before TCR\(\beta\)-signaled feedback inhibition. Although ATM likely cooperates, at least in part, by repressing RAG expression in response to DSBs induced during V\(\beta\) recombination, we cannot rule out an additional role in decreasing the time between RAG DSBs and TCR\(\beta\) protein expression by facilitating coding join formation. V31 is an atypical V\(\beta\) segment as it rearranges solely by inversion, and we show that ATM preserves V31 representation within the TCR\(\beta\) repertoire by preventing non-functional
deletions when a high-quality RSS controls V31 rearrangements. We also show that ATM preserves Vκ usage in the Igκ repertoire for Vκ segments that rearrange by inversion by restricting Vκ RSS - Jκ hybrid joins. We conclude that: weak Vβ RSSs that render Vβ rearrangements inefficient and the ATM protein that shepherds RAG endonuclease activity cooperate to create a diverse repertoire of monoallelically expressed TCRβ proteins, and ATM functions generally to produce diverse antigen receptor repertoires.
Figure 4.1. The weak V31 RSS and ATM cooperate to prevent non-functional V31 rearrangements and ATM protects the Vκ repertoire. (A) The mouse TCRβ locus showing positions of Vβ, Dβ, and Jβ segments and Cβ exons. (B-E) Representative and quantified data of V2⁺ or V31⁺ (B and C) or V2⁻V31⁺ (D and E) thymic T cells from indicated genotypes. Data are from four experiments, each with at least one mouse of each genotype. (F) Diagrams showing resolution of V2 or V31 rearrangements when their cleaved RSS escapes. (G) Quantification of V31-5′Dβ RSS hybrid joins from three experiments with a mouse of each genotype. (H) Diagram of a Vκ rearrangement that occurs through inversion to Jκ1 and its resolution as a non-productive hybrid join. Primers (blue arrows) and probe used for the Taqman qPCR are shown. (I) Quantification of Vκ RSS-Jκ1 hybrid joins from three experiments with a mouse of each genotype. Statistics by two-way (C and G) or one-way (E) ANOVA with Tukey’s (C and E) or Šidák’s (G) multiple post-tests. ns = not significant, ***p<0.001, ****p<0.0001.
Figure 4.2. The poor V1 and V2 RSSs and ATM cooperate to enforce TCRβ allelic exclusion. (A-D) Representative and quantified data of V1⁺ or V2⁺ (A and B) or V1⁺V2⁺ (C and D) thymic T cells from indicated genotypes. Data from four experiments, each with at least one mouse of each genotype, except that only one experiment included Atm⁻/⁻. Statistics by two-way (B) or one-way (D) ANOVA with Tukey's multiple post-tests. **p<0.01, ****p<0.0001.
CHAPTER 5: A Broad TCRβ Repertoire Defines the Frequency of Naïve Antigen-Specific CD8 T cells

Abstract

Broadly diverse AgR repertoires are thought to mediate protective adaptive immunity, yet the extent of AgR diversity that facilitates protection is not clear. αβ TCR diversity is generated during the somatic assembly of TCR genes by V(D)J recombination in developing T cells. CD8 T cells express αβ TCRs and are crucial for the clearance of acute viral infections. Using a mouse model that shifts the composition of the Vβ repertoire without deleting entire segments of the TCRβ locus, we aimed to test the relationship between Vβ repertoire diversity and viral clearance during infection with lymphocytic choriomeningitis virus (LCMV) Armstrong. TCRs specific for the LCMV gp33 peptide utilize a narrow subset of Vβ segments for binding. In mice where the Vβ repertoire is skewed away from these canonical Vβ segments, we observe trended decreases in the frequency of gp33-reactive CD8 T cells, resulting in fewer cells that degranulate and secrete TNFα and IFNγ. Additionally, we find that other Vβ segments can compensate for gp33 binding. These data suggest that the TCRβ repertoire may affect antigen-specific immune responses.

Introduction

Mammalian adaptive immune systems are required to discern a vast array of potentially harmful microorganisms and molecules while maintaining self-tolerance via their lymphocyte AgRs. Hence, a broadly diverse AgR repertoire is thought to ensure proper coverage to establish host immunity. Unique to B and T lymphocytes, V(D)J recombination manufactures AgR genes from a limited number of genetic loci to produce AgR repertoires with near countless specificities (Lieber, 1991; Miles et al., 2011; Nikolich-Žugich et al., 2004). Complete assembly of Ig and TCR genes is initiated by the lymphocyte-specific RAG endonuclease in developing B and T cells. Each locus is a constellation of noncontiguous variable (V), sometimes diversity (D), and joining
(J) gene segments that are assembled into the V(D)J exons encoding the variable regions of AgR proteins (Schatz and Ji, 2011). RAG introduces DNA DSBs at these gene segments to produce hairpin-sealed coding and blunted signal ends that are repaired by ubiquitous NHEJ factors (Helmink and Sleckman, 2012). Coding ends are enzymatically opened by a single-stranded nick, and nucleotides are excised (by Artemis) and/or added (by TdT) to introduce additional diversity to V(D)J junctions (Helmink and Sleckman, 2012). Finally, ligation of the processed coding ends joins V, D, and J segments together, and transcription with constant (C) region exons encode one AgR chain. The diversity of Ig and TCR repertoires is a culmination of the numerous arrangements of V, D, and J combinations (combinatorial diversity) and the imprecise repair of V(D)J coding ends (junctional diversity). Estimates of Ig and TCR diversity generated by V(D)J recombination ranges from $10^{12}$ to $10^{20}$ unique receptors (Lieber, 1991; Miles et al., 2011; Nikolich-Žugich et al., 2004).

αβ T cell responses are critical for the control and clearance of invasive pathogens such as viruses and bacteria. Canonical αβ TCRs recognize protein antigens, but unlike Iggs that can recognize molecular determinants in nearly any configuration, these antigens must be proteolytically processed into linearized peptides and displayed in the context of major histocompatibility complex (MHC) proteins (Davies and Cohen, 1996; Garcia and Adams, 2005; Reinherz et al., 1999). αβ TCR diversity is focused to the hypervariable complementarity determining regions (CDRs) that form the peptide:MHC binding site (Nikolich-Žugich et al., 2004). The variable region of each TCRα and TCRβ chain contains three CDRs: CDR1 and CDR2 are germline encoded by Vα and Vβ segments, and CDR3 is encoded by the junctions spanning VJα and VDJβ joins, respectively. CDR3 is the most variable of the CDRs because it encompasses the effects of both combinatorial and junctional diversity during V(D)J recombination (Hughes et al., 2003; Kabat and Wu, 1970; Miles et al., 2011; Nikolich-Žugich et al., 2004; Rock et al., 1994).

Although V(D)J recombination generates an immense number of unique αβ TCRs, the diversity of the peripheral naïve repertoire contracts considerably due to thymic selection. (Arstila et al., 1999; Attaf et al., 2015; Miles et al., 2011; Tikhonova et al., 2012; Van Laethem et al., 2012;
During αβ T cell development in the thymus, MHC restriction is imposed upon CD4⁺CD8⁺ DP thymocytes expressing mature αβ TCRs (Kisielow et al., 1988). DP thymocytes with αβ TCR specificities that cannot engage peptide:MHC on the surface of cortical thymic epithelial cells (TECs) die by neglect, while those with specificities that signal too strongly are negatively selected (Marrack et al., 2008; Xing and Hogquist, 2012). DP cells that survive this stage of development are "positively selected," differentiate into CD4⁺ or CD8⁺ single positive (SP) thymocytes and migrate to the thymic medulla to test if their αβ TCR engages tissue specific antigens expressed on the surface of medullary TECs (Van Laethem et al., 2012; Xing and Hogquist, 2012). Thus, the αβ TCR repertoire is focused as a result of thymic selection, where ~3% of TCRs (~10¹³ TCRs assuming a theoretical 10¹⁵ maximum) from the pre-selection pool survive to be a part of the naïve repertoire (Attaf et al., 2015; Goldrath and Bevan, 1999).

Yet, not all of these ~10¹³ possible TCRs are made in a single host (Miles et al., 2011), thus raising a basic and fundamental question: what is the required TCR diversity that provides sufficient coverage and protective immune responses? Multiple mouse models that reduce TCR diversity and then measure susceptibility to disease have been generated to try and tackle this question (reviewed in Nikolich-Žugich et al., 2004). However, the preferred methods used to decrease TCR diversity in many of these models rely on either targeted deletions of gene segment clusters at Tcra or Tcrb loci (i.e.: the complete absence of particular Vβ, Dβ, Jβ, Vα, and Jα segments) or use a TCRβ transgene that silences ~98% of endogenous Vβ rearrangements and is expressed by ~98% of cells. The former method necessarily creates holes in the TCR repertoire and specific gene segments may participate in protective immunity to a particular pathogen. This appears to be the case for some Navajo Native Americans who carry an allele that reduces the VκA2 segment from the Ig repertoire and exhibit greater susceptibility to infection by haemophilus influenzae (Feeney et al., 1996). In the latter method, a fully assembled TCRβ transgene is expressed by almost all cells in the host, thus decreasing Vβ diversity, junctional diversity, and also imposes restrictions on the types of TCRα chains that it can pair to.
In contrast, our RSS-replacement model permits dramatic shifts in the TCRβ repertoire without creating holes in the repertoire or imposing a particular type of Vβ chains on nearly all cells.

The LCMV model is useful to dissect virus-specific immune responses and mechanisms of anti-viral immunity. Several tools also exist to track antigen-specific cells and peptides to stimulate T cells ex vivo. CD8 T cells are necessary for LCMV clearance and utilize a focused Vβ repertoire to recognize the immunodominant glycoprotein (gp) 33 epitope presented by H-2Db (Blattman et al., 2000; Kotturi et al., 2007). In WT mice infected with LCMV, V13-2*, V13-3*, and V29+ TCRβ chains account for ~50% of gp33-specific TCRs while V2+ and V31+ TCRβ proteins account for less than ~5% of the gp33-reactive repertoire (Blattman et al., 2000). In naïve WT mice, V2+ and V31+ TCRβ proteins make up ~13% of the thymic and peripheral TCRβ repertoire (Figure 5.1, B and data not shown). In V2R31R/V2R31R mice (Wu and Bassing, 2010b), the Vβ repertoire is shifted towards overutilization of V2 and V31 segments and is biased away from the canonical V13-2, V13-3, and V29 TCRs that bind gp33:H-2Db; the V2 and V31 segments account for 88% of the Vβ repertoire, leaving only ~12% to be represented by other Vβ species (Figure 5.1, B). To test roles for Vβ diversity in mediating protective immunity, we infected WT and V2R31R/V2R31R mice with the acute Armstrong strain of LCMV and analyzed mice at days 7 (acute phase) and 30 (memory phase) post-infection (Figure 5.1, A).

Results

We first characterized the peripheral immune compartments in WT and V2R31R/V2R31R mice at days 7 and 30 post-infection. We observed that splenocyte numbers and their frequencies of B cells, total T cells, NK cells, DCs, neutrophils, and progenitor (p) and inflammatory (i) monocytes are unchanged between WT and V2R31R/V2R31R mice (Figure 5.1, C-E and data not shown). We next analyzed CD8 T cells, which are required for LCMV clearance (Moskophidis et al., 1987), in WT and V2R31R/V2R31R mice and found that their ability to become activated and form short-lived (SL) and memory-precursor (MP) effector cells (ECs) are also
unaffected (Figure 5.1, F and G and data not shown). We conclude that shifting the composition of the Vβ repertoire does not affect other immune cell subsets and the ability of CD8 T cells to respond to LCMV infection.

We next characterized Ag-specific CD8 T cell responses during acute LCMV infection. We quantified the frequency of gp33-reactive CD8 T cells using a gp33:H-2D<sup>b</sup> tetramer and observed a trended decrease of these cells in V<sup>2</sup>R<sup>31</sup>R/V<sup>2</sup>R<sup>31</sup>R mice compared to WT (3.7% versus 5.5%, Figure 5.2, A and B). We stimulated WT and V<sup>2</sup>R<sup>31</sup>R/V<sup>2</sup>R<sup>31</sup>R splenocytes ex vivo with gp33 peptide to measure the ability and frequency of CD8 T cells that degranulate and secrete the effector cytokines TNF<sub>α</sub> and IFN<sub>γ</sub>. Compared to WT, we observed a significant decrease in the frequency of V<sup>2</sup>R<sup>31</sup>R/V<sup>2</sup>R<sup>31</sup>R CD8 T cells that display surface CD107a (a marker of degranulation) and trended decreases in the secretion of TNF<sub>α</sub> and IFN<sub>γ</sub> (Figure 5.2, C-E). These data suggest that the starting frequency of gp33-reactive cells in the naïve CD8 T cell repertoire may be decreased in V<sup>2</sup>R<sup>31</sup>R/V<sup>2</sup>R<sup>31</sup>R mice, leading to fewer cells that respond to antigen stimulation.

Since the V13-2<sup>+</sup>, V13-3<sup>+</sup>, and V29<sup>+</sup> repertoire that binds gp33:H-2D<sup>b</sup> complexes is significantly reduced in V<sup>2</sup>R<sup>31</sup>R/V<sup>2</sup>R<sup>31</sup>R mice, we sought to determine the Vβ segments that are mediating binding to the gp33:H-2D<sup>b</sup> tetramer. Gating on gp33-specific CD8 T cells, we found that in WT mice 6.2% of cells use V<sup>2</sup><sup>+</sup> TCRβ chains and 0.42% of cells use V31<sup>+</sup> TCRβ chains, which is consistent with previously reported figures (Figure 5.2, F-H) (Blattman et al., 2000). In V<sup>2</sup>R<sup>31</sup>R/V<sup>2</sup>R<sup>31</sup>R mice, we find significant increases in the frequency of cells expressing V2<sup>+</sup> or V31<sup>+</sup> TCRβ chains that bind gp33:H-2D<sup>b</sup> (23.2% and 4.4%, respectively, Figure 5.2, F-H). In one V<sup>2</sup>R<sup>31</sup>R/V<sup>2</sup>R<sup>31</sup>R mouse, we also observed a population of cells that were V2<sup>+</sup>V31<sup>+</sup>, indicating a T cell clone (or clones) exhibiting allelic inclusion also can participate in antigen binding (Figure 5.2, F). These data indicate that in the absence of canonical TCRβ proteins that bind a particular antigen, CD8 T cells expressing TCRβ chains composed of other Vβ segments can bind to and recognize the same epitope. By day 30 when LCMV is cleared, the CD8 T cell pool returns to
homeostasis and the frequency of gp33-responsive cells equilibrate between WT and V2\(^R\)31\(^R\)/V2\(^R\)31\(^R\) mice (Figure 5.2, I and data not shown).

**Discussion**

Our study suggests that the TCR\(\beta\) repertoire may affect antigen-specific immune responses. We observe that the gp33-specific V\(\beta\) repertoire shifts from the canonical V13-2\(^+\), V13-3\(^+\), and V29\(^+\) receptors to V2\(^+\) or V31\(^+\) TCR\(\beta\) chains in V2\(^R\)31\(^R\)/V2\(^R\)31\(^R\) mice, suggesting that other V\(\beta\) segments can compensate for the absence of established V\(\beta\)s to detect antigen (Figure 5.2, F). How is this compensation achieved? One explanation is that the V13-2\(^+\), V13-3\(^+\), and V29\(^+\) TCR\(\beta\) proteins are higher affinity and simply outcompete lower affinity receptors for gp33. Another possibility is that the increased frequencies of V2\(^+\) and V31\(^+\) TCR\(\beta\) sequences now contain gp33-reactive TCRs. A clear caveat of this study is that this is a single experiment with three biological replicates for each genotype. Additional experiments will be required to determine if the general trends identified in this study are robust and repeatable. However, the V\(\beta\) RSS-replacement model may be useful to dissect roles for TCR\(\beta\) diversity in infection as the V\(\beta\) repertoire can be altered in young mice without creating repertoire holes from deleting gene segments (Nikolich-Žugich et al., 2004). Additionally, as V2\(^R\)31\(^R\)/V2\(^R\)31\(^R\) T cells also exhibit increased levels of multigenic TCR expression (Wu and Bassing, 2020b), these mice may also be useful to discern the effects of allelic inclusion on host immune responses.
Figure 5.1. The Vβ repertoire does not alter the frequencies of other immune cell subsets nor the CD8 T cell response during LCMV infection. (A) Outline of LCMV infection with indicated analysis timepoints. (B) TCRβ repertoire in WT and V2R31R/V2R31R thymic αβ T cells. (C) Total numbers of splenocytes at D7 and D30 post-infection. (D and E) Gating strategy to identify (D) and enumerate frequencies (E) of B cells, T cells, NK cells, DCs, neutrophils (Neut), and patrolling (p) and inflammatory (i) monocytes (Monos) 7 days post-infection. (F and G) Gating strategy to identify (F) and enumerate frequencies (G) of activated, memory precursor (MP), and short-lived (SL) effector CD8 T cells (ECs) 7 days post-infection. (mean ± SD, n = 3 mice per group, one experiment)
Figure 5.2. The Vβ repertoire alters the frequency of gp33-reactive CD8 T cells during LCMV infection. (A and B) Gating (A) and frequencies (B) of CD8 T cells that bind gp33-H-2D^b tetramer at 7 days post-infection. (C - E) Gating strategy (C), representative flow plots (D), and frequencies (E) of CD8 T cells that degranulate and secrete TNFα and IFNγ in response to stimulation with gp33 peptide 7 days post-infection. (F - H) Representative flow plots (F) and enumeration of V2^+ (G) and V31^+ (H) TCRβ chains that bind gp33-H-2D^b 7 days post-infection. (I) Frequencies of CD8 T cells that degranulate and secrete TNFα and IFNγ in response to stimulation with gp33 peptide 30 days post-infection. (mean ± SD, n = 3 mice per group, one experiment). Statistics by: unpaired Student’s t test (B, G, and H) or two-way ANOVA with Šidák’s multiple post-tests (E). *p<0.05.
CHAPTER 6: Discussion

Overview

A cardinal feature of mammalian adaptive immunity is the monoallelic expression of AgR genes by T and B cells (Flajnik, 2018; Vetterman and Schlissel, 2010). TCRβ, IgH, and Igκ allelic exclusion is achieved by: assembly of an in-frame gene on one allele, transient inhibition of V(D)J recombination signaled by RAG DNA DSBs, and the resulting protein sending feedback signals to permanently inhibit V-to-(D)J recombination of the non-functional allele (Brady et al., 2010b; Steinel et al., 2010; Steinel et al., 2013; Uematsu et al., 1988). However, feedback inhibition and its associated epigenetic silencing takes time, as an in-frame V(D)J rearrangement must be transcribed, translated, and the resulting receptor trafficked to the cell surface to then transmit a signal. This leaves a window of opportunity for secondary rearrangements to occur. If feedback inhibition were the sole enforcer of allelic exclusion, one would expect a large proportion of lymphocytes to express dual antigen receptors. In support of this notion, 10-50% of B cells express two distinct IgH proteins in mice when fully assembled Igh genes are knocked-into both Igh alleles (Sonada et al., 1997). Similar results are also observed in T cells expressing TCRβ chains encoded by two Tcrb transgenes (Sant'Angelo et al., 2001). Therefore, mechanisms must be in place to ensure the monoallelic assembly of Tcrb, Igh, and Igκ genes before receptor-mediated feedback inhibition can enforce permanent silencing. There is much disagreement in the field as to whether deterministic or stochastic mechanisms control asynchronous V rearrangement between alleles. For 20 years a range of diverse epigenetic-based phenomena, such as modulation of chromatin accessibility and locus topology, have been implicated in sequential recombination of AgR genes (Farago et al., 2012; Mostoslavsky et al., 2001). In this thesis, I demonstrate a substantial role for weak Vβ RSSs in limiting biallelic Vβ rearrangements in a stochastic manner to promote monogenic Tcrb assembly within the time window before feedback inhibition halts Vβ recombination.
Poor Quality Vβ RSSs Provide a Stochastic Mechanism for Limiting Vβ Recombination Between Alleles Before Feedback Inhibition

My thesis research has demonstrated the first molecular mechanism dictating initiation of AgR allelic exclusion, revealing that a fundamental genetic mechanism controls monoallelic assembly of a functional TCRβ gene, working even before TCRβ protein-signaled feedback inhibition permanently halts Vβ recombination. Sequence features of Vβ and VH RSSs are thought to render V-to-DJ recombination inefficient and are proposed to mediate monoallelic V-to-DJ recombination (Liang et al., 2002). I demonstrate that replacement of the DJCβ-distal V2 RSS and the DJCβ-proximal V31 RSS with the higher quality 3'Dβ1 RSS in mice: i) elevates recombination of these segments before TCR-signaled feedback inhibition, ii) dramatically increases the frequency of αβ T cells with TCRβ chains derived from both TCRβ alleles, and iii) reveals TCRβ alleles compete for recombination (Chapter 2; Wu et al., 2020). The only way that elevating the efficiency of Vβ recombination on one allele can outcompete Vβ rearrangement on the other is if both alleles are activated in the time window(s) that feedback inhibition requires to block Vβ recombination of the second allele, as is predicted in stochastic models. If deterministic models were correct, the recombination of one allele would have no consequence on the recombination of the other, and one would predict that the increases of Vβ recombination in RSS replacement mice would be additive based on allelic copy number. Importantly, I also show similar findings when I replace the RSSs of the distal V2 and V1 RSSs demonstrating that weak Vβ RSSs serve a broad role in enforcing TCRβ allelic exclusion (Chapter 4). These findings provide unequivocal evidence that weak Vβ RSSs provide a major genetic mechanism that stochastically reduces the incidence of in-frame Vβ rearrangements on both alleles, offering an explanation for the pre-permanent feedback inhibition paradox.

One key observation from my thesis work is that computationally derived correlates of RSS quality may not be predictive for rearrangement frequency in vitro or in vivo. Regression analyses of Tcrb, Igh, and Igk loci conclude that V accessibility, epigenetic modifications, and
transcription factor binding are among the principal factors for determining relative V utilization (Bolland et al., 2016; Gopalakrishnan et al., 2013; Kleiman et al., 2018). Interestingly, V RSSs rank highly as a determinant for V usage but are concluded to act as “binary switches” of V recombination since little to no correlation can be drawn between V usage and RIC score (Bolland et al., 2016; Gopalakrishnan et al., 2013; Kleiman et al., 2018; Matheson et al., 2017). Heptamer, spacer, and nonamer sequences all affect recombination frequency and the RIC score is calculated from a complex probability model of how each nucleotide position diverges from an averaged 12RSS or 23RSS (Cowell et al., 2002; Feeney et al., 2000; Hesslein and Schatz, 2001; Jung et al., 2003). However, several concerns are raised by this approach. First, the RIC score assumes that these averaged 12RSS and 23RSS are the “standards,” and how similar a test RSS is to one of these standards affects the RIC score. Second, the averaged 12RSS and 23RSS are derived from 356 RSSs from both TCR and Ig loci irrespective of whether they bordered V, D, or J segments. Thus, when making comparisons among a group of 12RSS or 23RSS, subtle but conserved nucleotide positions that can alter rearrangement frequency may be lost (Cowell et al., 2002). Third, the RIC score does not factor coding flank sequences that are known to dramatically affect recombination frequency (Boubnov et al., 1995; Ezekiel et al., 1995; Gerstein and Lieber, 1993). Finally, the RIC score does not address pairwise effects of RSSs on recombination. Given these considerations, perhaps it is not surprising that the RIC score was not predictive of V rearrangement frequency at Tcrb, Igh, and Igk loci nor the Vβ repertoire gains from replacing the RSSs of the V1, V2, and V31 segments with the 3'Dβ1 RSS (Bolland et al., 2016; Gopalakrishnan et al., 2013; Kleiman et al., 2018; Matheson et al., 2017; Wu et al., 2020a; Chapter 2; Chapter 4). As context dependent bending of each RSS must occur for RAG/HMGB1 to pair RSSs for coupled cleavage (Kim et al., 2018), perhaps a score that predicts RSS quality that also factors the flexibility of each RSS spacer can be constructed. Although this approach would still lack considerations of coding flank sequences and pairwise interactions, it may provide an improvement to the RIC score.
Poor Quality Vβ RSSs Stochastically Restrain Vβ Rearrangements to Promote Monogenic Tcrb Assembly

Unlike Igh and Igk loci, the Tcrb locus has two DJCβ clusters and the configuration of the locus could permit assembly and expression of two unique TCRβ chains from each allele. Dual Tcrb gene assembly and expression from one allele must involve an upstream Vβ segment and the downstream V31 segment that rearranges solely by inversion. However, this is prevented by undetermined mechanisms. To elucidate whether poor quality Vβ RSSs also restrict the assembly and expression of two distinct TCRβ genes on one allele, I engineered mice that replaced both RSSs to the V2 and V31 segments with the better 3'Dβ1 RSS on one allele. I showed that this increases overall rearrangement frequencies of both Vβs and elevates the fraction of αβ T cells that express two different TCRβ proteins encoded by one allele (Chapter 3; Wu and Bassing, 2010b). To support dual TCRβ chain expression from one allele, the rearrangement events either involve: 1) a deletional V2 rearrangement to the DJCβ1 cluster and an inversional V31 rearrangement to the DJCβ2 cluster, or 2) an inversional V31 rearrangement to the DJCβ1 cluster, which inverts a portion of the locus that contains the DJCβ2 cluster, and then an inversional V2 rearrangement to the DJCβ2 cluster (Lee and Bassing, 2020). In the first scenario, rearrangements need not be ordered so either Vβ can recombine to their respective target DJCβ cluster at any time before the onset of feedback inhibition mechanisms. In the second scenario the order of rearrangements matter because the rearrangement of an upstream Vβ to the DJCβ2 cluster deletes DJCβ1, thus precluding a rearrangement from V31. One would predict that the latter scenario would be less likely than the former given the kinetics, but we identified clones that had either one of these configurations on one allele in our V2R+/V31R/ hybridoma analysis (Chapter 2, data not shown). These data highlight that the low qualities of Vβ RSSs also stochastically restrict Vβ rearrangements occurring along a single allele to restrain the expression of multiple distinct TCRβ chains by αβ T cells. Thematically, the mechanisms that enforce Tcrb allelic exclusion also mediate Tcrb isotypic exclusion (Cβ1* versus Cβ2* TCRβ protein) to effect...
monogenic TCRβ expression. A complete break in Tcrb allelic and isotypic exclusion and Tcra allelic exclusion could produce a small fraction of T cells with four TCRβ chains and two TCRα chains; assuming that there are no restrictions in pairing, a T cell in theory could express eight unique αβ TCRs.

In mammals, Igl and Tcrg loci are also in a configuration that can permit the assembly and expression of multiple AgR chains per allele. Each locus possesses two (Igl) to four (Tcrg) miniclusters that are made up of one or more V segments, one J segment, and a constant region (Carding and Egan, 2002; Vernooij et al., 1993; Vettermann and Schlissel, 2010). Igk recombination precedes rearrangements at Igl in pre-B cells, and the development of Igλ+B cells is facilitated by pro-survival NFκB signaling that promotes pre-B cell survival (Derudder et al., 2009). From the analysis of Igλ secreting hybridomas, 97% of cells contained a single Igl rearrangement, suggesting that Igl recombination is infrequent (Nadel et al., 1990). Consistent with this notion, it is estimated that the frequency of Igl rearrangements is restricted to one per cell and no hybridomas have been observed to contain two Igl rearrangements on one allele, satisfying both allelic and isotypic Igl exclusion (Coleclough et al., 1981; Sanchez et al., 1991). One explanation for this pattern is that similar to Igh, Igk, and Tcra loci, Igl loci also replicate asynchronously, thus possibly indicating a role for epigenetic phenomena that promote monogenic assembly and expression (Mostoslavsky et al., 2001). No two Igl rearrangements have ever been observed on one allele, suggesting either the epigenetic mechanism also identifies and differentially marks miniclusters between alleles or additional levels of regulation must occur for monogenic Igl assembly. Another possibility is that Igl RSSs exhibit lower recombination activity than Igk RSSs in vitro, supporting a stochastic mechanism for monogenic Igl assembly and Igk/Igl isotypic exclusion (Ramsden and Wu, 1991). γδ T cells exhibit phenotypic allelic exclusion because of pairing restrictions between TCRγ and TCRδ chains, but an estimated 10% of cells contain biallelic in-frame Tcrg rearrangements (Boucontet et al., 2005). Yet, not all Tcrg rearrangements are exhausted and whether any cells express two TCRγ proteins.
encoded by one allele is not known. Vγ RSSs share features with Vα and Vδ RSSs that are predicted to be efficient, suggesting that RSS quality may not play a relevant role in promoting monogenic Tcrg assembly (Liang et al., 2002). γδ T cells emerge from the thymus in characteristic “waves” during ontogeny and are typified by their Vγ usage (Carding and Egan, 2002; Sumaria et al., 2019). For example, Vγ5+ cells develop first around embryonic day 12, followed by Vγ6+ cells and then Vγ4+ cells. Given that γδ T cells are developmentally regulated throughout embryonic and fetal development, perhaps epigenetic mechanisms ensure recombination occurs at only one minicluster at a time along a Tcrg allele. Additional investigation of Igl and Tcrg loci is warranted to determine how monogenic gene assembly is achieved.

Roles for ATM in Allelic Exclusion and Repair of Inversional Rearrangements

My thesis work also provides new insights into the role of ATM and weak Vβ RSSs to regulate monoallelic Tcrb assembly. RAG-generated DNA DSBs from V-to-(D)J recombination on one Igk allele activates the ATM kinase to signal temporary feedback inhibition of V recombination. DNA DSBs generated by Igk recombination activate ATM and is correlated with downregulated RAG expression and Vk-to-Jκ recombination of the other Igk allele in pre-B cells (Steinel et al., 2013). Additionally, DSBs decrease RAG expression in pro-B cells and DN thymocytes, and inactivation of ATM increases biallelic V recombination and expression of Igh, Tcrb, and Igk genes (Fisher et al., 2017; Steinel et al., 2014). These data indicate that ATM is an important factor in mediating monoallelic recombination and expression of AgR genes. I show that weak Vβ RSSs cooperate with the ATM kinase to further restrain distal Vβ rearrangements (specifically V1 and V2) on both alleles before the onset of permanent feedback inhibition (Chapter 4). However, ATM also stabilizes DNA ends in RAG post-cleavage complexes; in the absence of ATM these DNA ends can dissociate, hampering coding and signal join formation and leading to the potential loss of the intervening DNA between RAG DSBs (Bredemeyer et al., 2006). I show that ATM deletion does not affect representation of the distal V1 and V2 segments.
that rearrange by deletion as the intervening sequences containing the signal ends are normally lost from the genome. However, the absence of ATM does impact \( V31 \) as it is an atypical \( V\beta \) segment that rearranges solely by inversion, and loss of intervening sequences results in non-functional deletions. I also show that this role for ATM in the proper repair of inversional \( V \) rearrangements is extended to the \( Igk \) locus where \( \approx 1/3 \) of \( V\kappa \) segments recombine by inversion (Matheson et al., 2017; Chapter 4). These data indicate that inefficient \( V\beta \) rearrangements directed by \( V\beta \) RSSs and the ATM kinase cooperate at least in part by repressing RAG expression in response to DSBs to create a diverse repertoire of monoallelically expressed TCR\( \beta \) proteins. However, I cannot rule out an additional role for decreasing the time between RAG DSBs and TCR\( \beta \) protein expression by facilitating coding join formation. Additionally, ATM protects the composition of AgR repertoires by ensuring the representation of \( V \) segments that rearrange by inversion.

**A Working Model of the Mechanisms Contributing to Monoallelic AgR Gene Assembly**

Weak \( V\beta \) RSSs are not the sole regulators that limit biallelic assembly and likely cooperate with additional factors to mediate monoallelic \( Tcrb \) recombination. In \( \approx 95\% \) of DN cells, at least one of the two \( Tcrb \) loci are positioned in a repressive nuclear environment that is refractory to V(D)J recombination such as pericentromeric heterochromatin or the nuclear lamina where RAG2 protein is depleted (Chan et al., 2013; Schlimgen et al., 2008). Disrupting these contacts with the nuclear lamina increases looping to the recombination center and the accessibility and recombination of proximal \( V\beta \) segments (Chen et al., 2018). From these observations, I propose the following model: in non-cycling DN3 cells, at least one allele stochastically becomes active to undergo \( V\beta \)-to-DJ\( \beta \) recombination (Figure 6.1, A-C). When one allele becomes accessible in this manner, weak \( V\beta \) RSSs ensure monogenic \( Tcrb \) assembly (Wu and Bassing, 2020b) (Figure 6.1, B). If both alleles become active in the same time window, weak \( V\beta \) RSSs also ensure monoallelic \( Tcrb \) assembly (Wu et al., 2020a) (Figure 6.1, C). RAG
cleavage of one allele generates DNA DSBs that trigger ATM-mediated transient feedback inhibition at least in part by repression of RAG expression, providing time to test the initial rearrangement (Fisher et al., 2017; Steinel et al., 2014) (Figure 6.1, D). ATM and weak Vβ RSSs cooperate to ensure monoallelic assembly of Tcrb genes, and also ensures proper repair of inversional V31 rearrangements (Figure 6.1, D). If the initial rearrangement is out-of-frame, RAG is re-expressed to permit Vβ recombination on the second allele or on the first allele if a DJβ complex is available. If the first rearrangement is in-frame, the VDJβ join is expressed as a TCRβ protein that pairs with preTα to form a pre-TCR complex (Figure 6.1, E). pre-TCR signaled permanent feedback inhibition silences Vβ recombination, possibly through locus decontraction, nuclear positioning, and epigenetic modifications. Pre-TCR signaling also promotes cell cycle by transcriptional activation of Cyclin D3, which moves G1 phase cells into S phase where RAG2 is ubiquitylated and targeted for degradation (Lin and Desiderio, 1994; Sicinska et al., 2003) (Figure 6.1, F). Cyclin D3 may also directly repress Vβ accessibility before cells enter S phase based on its function in pro-B cells (Powers et al., 2012). However, in thymocytes with two functional VDJβ rearrangements, post-transcriptional silencing of one of the two mRNAs and pairing restrictions with TCRα chains also enforce TCRβ allelic exclusion (Brady et al., 2010b; Steinel et al., 2010). Finally, TCRβ signals promote genetic and epigenetic changes that silence Vβ recombination in DP cells where Tcra genes assemble (Jackson and Krangel, 2005; Liang et al., 2002; Majumder et al., 2015; Skok et al., 2007). Any of these mechanisms and/or the shared sequence features between Vβ and Vλ 23RSSs may contribute to monoallelic Igα assembly and expression.

Yet, Vκ segments are predicted to possess highly efficient 12RSSs (Cowell et al., 2003). Assuming that this premise is true, other mechanisms must then direct monoallelic Igκ assembly. Capitalizing on the genetic differences between the alleles of B6/Castaneous F1-hybrid mice, analysis of single pre-B cell clones by assay for transposase-accessible chromatin sequencing (ATAC-Seq) revealed that approximately 30-40% of Vκ segments are accessible (Levin-Klein et al., 2017). This accessibility also correlates with local transcription and histone acetylation (Levin-
Comparing the profile across several clones, a single Vκ segment adopts one of three states: accessible on both B6 and Castaneous alleles, accessible on one allele, or not accessible on either allele (Levin-Klein et al., 2017). These data support a model where following the expression of the pre-BCR in pro-B cells but prior to the differentiation of pre-B cells, the Igk locus is epigenetically modified in a stochastic manner to activate a subset of Vκ segments to undergo Vκ-to-Jκ recombination. Similar mechanisms of differential V accessibility could occur at Tcrb and Igh loci, but apart from constant region congenic markers, tools to distinguish the two alleles are lacking. Yet, in instances where both Vκ segments are accessible, additional levels of regulation, such as CTCF and cohesin mediated chromosomal loops (Barajas-Mora et al., 2019; Kleiman et al., 2018), must direct monoallelic Igk recombination. Like Tcrb and Igh loci, in pre-B cells one Igk allele is positioned in an active chromatin environment and one allele is located in transcriptionally repressive heterochromatin, suggesting that stochastic and differential positioning of alleles in nuclear structures also enforces Igk allelic exclusion (Goldmit et al., 2005).

**Models Contributing to a Diverse Vβ Repertoire**

Computational models of the various factors that influence V rearrangement frequency at Igk, Igk, and Tcrb loci largely fit into three categories: 1) RSS quality, 2) measures of V accessibility, and 3) proximity/structural elements that bring V segments into contact with the recombination center (Bolland et al., 2016; Gopalakrishnan et al., 2013; Kleiman et al., 2018; Matheson et al., 2017). Replacement of the V1, V2, and V31 RSSs with the same strength 3’Dβ1 RSS reveals differences in the usage of these segments in the Vβ repertoire, thus implicating the role(s) of other factors in directing Vβ rearrangement level. Based upon their usage in the Vβ repertoire, the 3’Dβ1 RSS replacement enhances V31 the most, followed by V2, and then finally V1 (Chapter 2 and Chapter 4). These differences are likely due to their frequencies of rearrangement as I have shown in DN3 cells by Taqman qPCR (Chapter 2; Wu et al., 2020).
can imagine several models that can account for this variation. First, V31 is the most proximal Vβ segment in the Tcrb locus that does not require locus compaction to access to the recombination center (Figure 6.2). Second, establishing CTCF-mediated chromosomal loops that bring all of the other distal Vβs into proximity of the recombination center may take time, perhaps giving V31R a temporal advantage compared to V2R and V1R (Figure 6.2). The V31+ repertoire is comparable between V2R31RWT and V31R mice, while the V2+ repertoire is reduced by 50% in V2R31RWT compared to V2R mice (Chapter 2 and Chapter 3). Therefore, a third explanation is that when both V2R and V31R are on the same allele, perhaps a V31R rearrangement signals feedback inhibition that results in decontraction of the distal end of the Tcrb locus. Finally, greater accessibility of V31 over V2 and V1 may account for the observed recombination patterns (Gopalakrishnan et al., 2013).

Although the RSS replacements at V1, V2, and V31 all result in their increased usage, this is not a general/global phenomenon. During my thesis I have also generated RSS replacements using the same 3'Dβ1 RSS at V4 and V15 segments, and other members of my lab replaced the RSSs to V3, V13-2, V16, and V19. As result of the RSS replacements, V13-2 and V16 utilization increase, V3 and V19 exhibit no discernable change, but contrary to the three Vβ segments that I discussed in depth in this thesis, the utilization of V4 and V15 decrease as a result of the 3'Dβ1 RSS replacement (data not shown). The reason(s) for the decreased rearrangement and utilization of V4 and V15 is not understood and is under active investigation. One possibility is that the V4 and V15 RSSs are of better quality than the 3'Dβ1 RSS. To determine this, the lab is currently generating competitive in vitro plasmid substrates that will compare the rearrangement of the endogenous V4 or V15 RSS and the 3'Dβ1 RSS to a target RSS. This assay allows for a head-to-head comparison of the activity of these RSSs without any other confounding factors. This assay can also be used to clone in and compare the coding flank V4 and V15 sequences to determine if these are a limiting factor during recombination. If these results reveal that the 3'Dβ1 RSS is indeed better than the V4 and V15 RSSs, then some
unknown factor is influencing Vβ rearrangement levels within that chromatin environment. Interestingly, from regression analyses some features that are shared amongst the V1, V2, V13-2, V16, and V31 segments, but not the V4 and V15 segments, are factors and histone marks that are associated with high levels of accessibility (Gopalakrishnan et al., 2013). RSSs can phase nucleosomes over themselves even within accessible chromatin, thus it is possible that the 3′Dβ1 RSS when attached to a particular Vβ segment binds nucleosomes to antagonize recombination (Baumann et al., 2003; Golding et al., 1999; Kondilis-Mangum et al., 2010; Kwon et al., 1998). Perhaps the highly accessible chromatin environment of V1, V2, V13-2, V16, and V31 segments can overcome the antagonism threshold created by nucleosomes positioned over the 3′Dβ1 RSS, but the low accessibility of V4 and V15 segments is further exacerbated by 3′Dβ1 RSS-nucleosome occupancy. Additional studies are required to understand the mechanism(s) that hinders V4 and V15 rearrangements when their RSSs are replaced with the 3′Dβ1 RSS.

Long-Range V Recombination: RAG Scanning versus Proximity-Based Diffusion

How is long-range V-to-(D)J recombination achieved for AgR loci whose V segments can be positioned hundreds of thousands to millions of base pairs away from the recombination center? Locus compaction of Tcra/Tcdr, Tcrb, Igh, and Igk loci presumably brings distal V segments into close proximity to the recombination center and requires the structural zinc-finger protein CTCF, its binding elements (CBE), and cohesin (Allyn et al., 2020; Shih and Krangel, 2013). In general, CBEs are found throughout the V gene clusters and are typically in a convergent orientation with the CBEs positioned around the recombination center. However, this pattern is locus-dependent and the diverse locations and orientations of CBEs may ultimately affect V(D)J recombination. Cohesin is thought to form chromatin loops by bi-directionally extruding the chromatin fiber until the protein complex pauses at CTCF-bound CBEs. Two models have been proposed for how long-range V recombination could occur by CTCF-mediated chromatin looping. The RAG-scanning model posits that the RAG endonuclease binds to a D/J RSS and through the action of cohesin passes chromatin until a compatible V RSS is
encountered. RAG can either form a synaptic complex with the V RSS or it can pass and continue to sample other V RSSs in this manner, in effect “scanning” the V array (Hu et al., 2015; Jain et al., 2018). Evidence for this RAG-scanning model in vivo is the strongest for the lgh locus. All V<sub>H</sub> segments are in a convergent orientation with J<sub>H</sub> segments and thus recombine with DJ<sub>H</sub> complexes by deletion. Inverting the portion of the locus containing the V<sub>H</sub> segments nearly abolishes all V<sub>H</sub> recombination even though V<sub>H</sub> rearrangements could proceed by inversion (Dai et al., 2021). Furthermore, cryptic RSSs within the inverted V<sub>H</sub> cluster are now able to support recombination (Dai et al., 2021). Yet, it is unclear how the RAG-scanning model could accommodate inversional V<sub>k</sub> rearrangements at the Igk locus. Alternatively, the diffusion model proposes that chromatin loops bring V segments into a “cloud of spatial proximity” with the recombination center, thus giving V segments near-equal access to RAG (Ji et al., 2010). If true, a prediction of the RAG-scanning model would be that in mice carrying the V<sup>2<sub>R</sub></sup> allele representation of V<sup>1<sub>+</sub></sup></sup> TCRs would be more diminished compared to the representation of other V<sub>β</sub> segments. In V<sup>2<sub>R<sub>α</sub></sub></sup> and V<sup>2<sub>R<sub>R</sub></sub></sup> mice, V<sup>1</sup>’s decreased utilization is proportional to the decreased utilization of all other V<sub>β</sub> segments surveyed (data not shown), suggesting that proximity-based diffusion plays a greater role for V<sub>β</sub> recombination at the Tcrb locus. Additionally, like Igk, the inversional rearrangement of V31 also presents another obstacle for RAG-scanning. More targeted experiments and analyses are warranted to determine if RAG-scanning plays a role for Tcrb locus recombination.

**Physiological Roles for Allelic Exclusion in Vertebrate Immune Systems**

AgR allelic exclusion is conserved over >500 million years of vertebrate evolution and is a defining feature for both VLR- and IgSF-based adaptive immune systems (Cooper and Alder, 2006; Flajnik, 2018; Hsu, 2009). The lymphocytes from lamprey and hagfish appear to exhibit allelic exclusion even though their VLR genes are organized in a completely different manner from mammals and use a different set of enzymes for assembly (Boehm et al., 2012). Allelic exclusion is also thought to occur in sharks, rays, and skates, which represent the oldest jawed
vertebrate animals that, depending on the species, possess anywhere from a dozen to ~200+ \( Igh \) miniloci (Malecek et al., 2008). In sharks, each minilocus is \(~2\) kilobases in size and typically contains one V segment, several D segments, and one J segment. Since these loci are small relative to most mammalian AgR genes, locus contraction may not be a regulating factor that directs monoallelic recombination. Experiments in nurse sharks reveal that several miniloci recombine at a time and occur in an unregulated manner; V-to-D, D-to-D, and D-to-J rearrangement intermediates are identifiable in nurse shark B cells (Malecek et al., 2008; Zhu et al., 2001). Even though several miniloci appear to rearrange in any B cell, how monoallelic and monogenic \( Igh \) assembly is achieved in these animals remains unknown. Given the number of potential shark \( Igh \) genes that can rearrange and contribute to expression, perhaps some similarities can be drawn with the mammalian olfactory receptor (OR) system. There are approximately 1500-2000 OR genes in mice that are scattered throughout the genome, and OR expression is limited to a single gene from a single allele (Monahan and Lomvardas, 2015). OR genes, like AgR genes, are subject to an initiation phase that identifies one gene to be expressed followed by subsequent feedback inhibition. Perhaps factors such as nuclear localization, chromatin accessibility, transcription factor binding, and post-transcriptional silencing of all but one mRNA species could contribute to monoallelic and monogenic expression of OR and shark \( Igh \) genes (Monahan and Lomvardas, 2015).

**So why allelic exclusion?** Not all B and T lymphocytes exhibit monoallelic expression of AgR genes, and the extent that allelic exclusion is applied varies depending on the AgR locus; allelic exclusion of \( Igh \) loci is the most stringent and 0.01% of B cells express two different IgH chains, while allelic exclusion at \( Tcra \) loci is the least stringent and 10% of T cells express two TCR\(\alpha\) proteins (Barreto and Cumano, 2000; Brady et al., 2010b; Niederberger et al., 2003; ten Boekel et al., 1998; von Boehmer and Melchers, 2010). Although allelic exclusion was discovered 55 years ago, its role(s) in host physiology continues to puzzle immunologists (Pernis et al., 1965). To date, at least four non-necessarily mutually exclusive models have been proposed that ascribe a physiological functional to AgR allelic exclusion, one of which has been disproven. It
has been thought that allelic exclusion prevents: i) the “heavy chain toxicity” effect, ii) immunodeficiency by diluting AgR density for a given antigen, iii) autoimmunity by impaired selection, and iv) oncogenic translocations. The first model, the “heavy chain toxicity” hypothesis, assumes that a B cell displaying biallelic expression of Ig proteins would die from the burden caused to the unfolded protein response in the endoplasmic reticulum (Wabl and Steinberg, 1982). This was summarily disproved when it was shown that B cells in mice with two fully assembled $Igh$ genes can express two distinct IgH proteins and have equivalent sized B cell compartments compared to control mice (Sonada et al., 1997). In the second model, a complete break in allelic and isotypic exclusion mechanisms in B cells could generate a cell expressing two IgH chains, two Igκ chains, and (because of the locus structure) four Igλ chains (Vettermann and Schlissel, 2010) (Figure 6.3, A). Assuming no restrictions in $V_L/V_H$ pairing, these polyspecific B cells could generate twelve unique antigen-binding sites and 144 potential BCRs from the various combinations of heavy and light chains (Figure 6.3, A). Thus, a single cell expressing multiple BCRs may be sub-optimally activated by antigen since the receptor density for a given antigen is “diluted” relative to a cell expressing a single kind of BCR. In support of this theory, Igκ-included B cells develop normally but when stimulated through one IgH/Igκ chain in vitro are partially refractory to BCR signaling (Velez et al., 2007).

The third model is perhaps the most popular and pervasive as it is frequently taught in Immunology 101 classrooms, and it states that allelic exclusion is important to limit autoimmunity from cells expressing both a non-autoreactive receptor (permitting positive selection during development) and an autoreactive receptor (that can detect and attack tissues in the periphery). In a TCR transgenic model, T cells expressing an autoreactive TCR are efficiently deleted in the thymus, but upon expression of a second non-autoreactive TCR, those cells that would normally be deleted are rescued and seed peripheral tissues (Zal et al., 1996). Additionally, T cell effector functions can be activated in these dual TCR T cells through stimulation of the autoreactive receptor (Zal et al., 1996). Formal demonstrations that dual TCR T cells (as a result of TCRα allelic inclusion) can cause autoimmunity in vivo were shown in mouse models of diabetes.
(Sarukhan et al., 1998; Schuldt et al., 2017). These data supported a model where AgR allelic exclusion evolved to ensure the proper negative selection of B and T lymphocytes expressing autoreactive receptors (Figure 6.3, B). Under the condition of allelic inclusion, cells expressing both autoreactive and non-autoreactive receptors allow lymphocytes to evade apoptosis by pro-survival signals derived from the non-autoreactive receptor during development (Heath et al., 1993; Padovan et al., 1993; Sarukhan et al., 1998; Zal et al., 1996) (Figure 6.3, B). However, although dual receptor lymphocytes can be observed in some autoimmune disease states such as systemic lupus erythematosus (Fraser et al., 2015), human autoimmune diseases that are causal from cells exhibiting allelic inclusion have not been demonstrated. Furthermore, some experiments suggest that a receptor saved from deletion by the expression of a second receptor is not necessarily a harmful autoreactive specificity (He et al., 2002). Some TCR specificities, and the cells that express them, are not efficiently positively selected in the thymus and are lost from the repertoire (He et al., 2002). Expression of a second TCR that permits positive selection rescues those useful TCR specificities that can go on to respond to foreign antigens in the periphery. Thus, allelic inclusion can be a benefit to the host by expanding the TCR repertoire (He et al., 2002). Given that most autoimmune diseases tend to develop at later stages in life, perhaps the host tolerates some level of biallelic expression of Tcra genes to balance the selection of a broad and diverse TCR repertoire even if there is an autoimmune hazard.

Finally, the fourth model, proposed by my thesis lab, posits that mechanisms directing monoallelic recombination, and thus induction of DNA DSBs, evolved to suppress DSBs from entering S phase where they can form oncogenic translocations (Brady et al., 2010b). I will discuss this model in more detail in the following section.

**Monoallelic V Recombination at Igh and Tcrb Loci May Limit Oncogenic Translocations**

Leukemias and lymphomas originating from transformed B- or T- lineage cells will account for ~10% of cancer cases in 2020 (Leukemia and Lymphoma Society, 2020). Chromosomal translocations, some of which involve AgR loci that possess strong transcriptional
enhancers, are common in these hematological malignancies (Belver and Ferrando, 2016; Jankovic et al., 2007; Küppers, 2005; Nussenzweig and Nussenzweig, 2010). Assembly and expression of Tcrb and Igh loci are linked with cellular proliferation; in-frame Igh and Tcrb rearrangements that occur in G1-phase cells produce IgH and TCRβ proteins to form pre-BCRs and pre-TCRs, respectively (Figure 6.4, A). Signaling through pre-BCRs and pre-TCRs induces the transcriptional activation of Cyclin D3 (Cooper et al., 2005; Sicinska et al., 2003), which forms complexes with cyclin dependent kinases (CDK) 4 and 6 to push cells into S phase. DNA damage that occurs early enough in G1 activates the G1/S checkpoint to halt the cell cycle and facilitate DNA repair. Upon the G1/S phase transition, RAG2 protein is degraded to halt V(D)J recombination and the formation of DSBs at Tcrb and Igh loci (Jiang et al., 2005; Lee and Desiderio, 1999; Sicinska et al., 2003; Steinel et al., 2014; Zhang et al., 2011). Failure to degrade RAG2 and terminate V(D)J recombination by the time a cell enters S phase causes lymphoid malignancies with clonal antigen receptor translocations, indicating a crucial role for cell-cycle control and the induction of RAG DNA DSBs (Zhang et al., 2011).

If V\(_\text{H}\) or V\(_\beta\) segments possessed stronger RSSs that increase recombination frequency, RAG could cleave the homologous Igh or Tcrb allele too late in G1 to activate the G1/S checkpoint. As DNA DSB intermediates are powerful translocation substrates (Richardson and Jasin, 2000), oncogenic translocations could occur between RAG-induced DSBs and DSBs generated during DNA replication (Figure 6.4, B) (Helmink and Sleckman, 2012). Thus, my thesis lab proposed that the pressure to suppress oncogenic AgR translocations selected mechanisms that direct monoallelic induction of RAG DSBs during Igh and Tcrb recombination (Brady et al., 2010b; Hewitt et al., 2009; Steinel et al., 2014). In G1 phase cells, DSBs trigger the ATM/p53-dependent G1/S cell cycle checkpoint to block S phase entry and to give the cell time to repair DSBs (Helmink and Sleckman, 2012). Loss of p53 or ATM results in more RAG DSBs entering S phase on non-selected Tcrb alleles of TCRβ-selected thymocytes (Dujka et al., 2010; Pedraza-Alva et al., 2006), and ATM loss also causes Tcrb translocations on non-selected alleles (Steinel et al., 2014). Thus, both the ATM/p53-mediated G1/S checkpoint and a potential ATM-mediated
feedback inhibition of recombination may cooperate with weak Vβ RSSs to restrict TCRβ-selected cells from entering S phase with RAG DSBs (Figure 6.4, B).

To test if monoallelic initiation of Vβ-to-DJβ recombination limits oncogenic Tcρ translocations, I generated aged cohorts of p53−/− and p53−/−:V2β31R/Vzβ31R mice and measured their median survival. The median survival age for the 22 mice in the p53−/− cohort is 158.5 days, whereas the 34 mice in the p53−/−:V2β31R/Vzβ31R cohort is 124.5 days (Figure 6.2, C and D). p53 promotes cellular homeostasis in all cell types and its deletion can cause numerous types of cancer including soft-tissue sarcomas (Taubert et al., 1998; Meek, 2009). We observed two instances of these sarcomas in our p53−/− cohort, but all mice in the p53−/−:V2β31R/Vzβ31R cohort possessed clear thymic and/or splenic lymphomas, suggesting that the RSS replacements in p53−/−:V2β31R/Vzβ31R mice increases the penetrance of lymphoid malignancies over other cancer types. This preliminary data demonstrates that by raising the recombination frequency of both V2 and V31 on both alleles decreases survival in p53−/− mice by 34 days. Although suggestive, I did not address whether biallelic Tcρ recombination results in an accumulation of RAG DSBs in S phase cells or if the accelerated mortality in p53−/−:V2β31R/Vzβ31R mice is due oncogenic Tcρ translocations.

I propose the following experiments to test the role of allelic exclusion in suppressing oncogenic translocations in the hope of gaining more mechanistic insight. To determine if infrequent Vβ recombination limits RAG DSBs at Tcρ loci from entering S phase, DNA DSBs can be quantified in sorted S/G2/M phase DN cells of 4-6 week-old p53−/− and p53−/−:V2β31R/Vzβ31R mice. One could analyze young mice at an age before most thymic lymphomas have been established, thus allowing for the analysis of developing thymocytes. Ligation-mediated PCR (LM-PCR) would be a sufficient strategy to quantify DSBs at pre-defined genomic locations such as the coding and signal ends generated by RAG cleavage at V2 and V31 RSSs. p53−/− mice die from thymic lymphomas that lack TCR translocations but exhibit aneuploidy (Liao et al., 1998). If the model is correct, one would expect to detect DSBs at V2 and V31 segments in p53−/−:V2β31R/Vzβ31R S/G2/M phase DN cells but not in p53−/− S/G2/M phase DN cells. Translocations
involving Tcrb loci and their target genes can also be measured genome-wide by translocation capture sequencing (TC-Seq) that detects one translocation in ~10 million cells (Oliveira et al., 2012). TC-Seq can be performed on sorted DP thymocytes from 4-6 week-old p53<sup>+/−</sup> and p53<sup>−/−</sup>:V2<sup>R</sup>31<sup>R</sup>/V2<sup>R</sup>31<sup>R</sup> mice. One would expect that more translocations involving the Tcrb locus will be observed in p53<sup>+/−</sup>:V2<sup>R</sup>31<sup>R</sup>/V2<sup>R</sup>31<sup>R</sup> DP cells. Finally, cytogenetic methods such as spectral karyotyping (SKY) and DNA fluorescence in situ hybridization (FISH) can be performed on thymic lymphomas to identify and map Tcrb translocations. If such a role for monoallelic V recombination in suppressing Igh and Tcrb translocations is established, this would highlight a role for allelic exclusion at the level of AgR gene assembly as compared monoallelic AgR expression. Whether these mechanisms that suppress oncogenic AgR translocations through allelic exclusion are observed in evolutionarily more ancient organisms is unknown. Contrary to popular belief, sharks do spontaneously develop cancers including those of lymphoid origin (Manire et al., 2013; Ostrander et al., 2004). It may be possible that naturally occurring lymphoid cancers in sharks could harbor clonal AgR translocations that were drivers of oncogenesis.

**Final Thoughts**

Measuring the extent of allelic exclusion at any AgR locus under normal or perturbed conditions is hindered by the fundamental lack of tools or by low-throughput and laborious methods. For some loci, such as Igh, Igk, and Tcra, congenically marked constant regions or fluorescently tagged receptors permits the measurement of global changes in allelic exclusion by flow cytometry (Casellas et al., 2001; Sonada et al., 1997; Yang et al., 2020). However, these tools do not exist for all AgR loci, posing a challenge for studying phenotypic allelic exclusion for other AgR genes. One way around this limitation is the use of anti-V antibodies that exist for Tcra, Tcrb, and Tcrg loci, but as I have mentioned, not all V species have a dedicated reagent for analysis resulting in underestimates of allelic exclusion. Even with the ease of genetic manipulations afforded by the CRISPR/Cas9 gene-editing platform, the strategy to congenically and/or fluorescently tag an AgR becomes more complicated for Igλ, Tcrb, and Tcrg loci since they
contain multiple constant regions and each allele could produce more than one unique protein. Another consideration for measuring population level changes in allelic exclusion is that phenotypic exclusion does not necessarily reflect the level of biallelic V rearrangements as post-transcriptional and post-translational mechanisms limit surface expression of multiple AgR proteins (Brady et al., 2010b; Steinel et al., 2010). Some of the strategies used to study biallelic recombination in single cells include Southern analysis of hybridoma clones, single cell DNA sequencing, and DNA FISH. However, analyses of hybridoma clones is low-throughput, laborious, and requires the use of radiation; single-cell DNA sequencing is expensive and is performed on a few hundred to a few thousand cells; and DNA FISH is only practical for loci that rearrange by deletion as the recombination status of an allele is indicated by the excision of a fluorescent probe placed between the V and (D)J gene segments (Hosoya et al., 2018; Steinel et al., 2014). As was implemented in Chapter 2 and Chapter 3 of this thesis, a cheaper, higher-throughput, and commercially available alternative could be used where genomic DNA is isolated from sorted lymphocytes and sequenced in bulk, which is offered by companies such as Adaptive Biotechnologies.

Currently, Adaptive Biotechnologies can perform sequencing of all possible VDJ rearrangements for Igh and Tcrb loci. As discussed in Chapter 2, assuming an AgR locus can only attempt one rearrangement per allele and stringent allelic exclusion is observed, 3/9 cells will make an in-frame rearrangement on one allele, and the receptor produced from that rearrangement will signal to feedback inhibit recombination of the homologous allele. Then, 2/9 cells will possess both an in-frame and an out-of-frame rearrangement as a result of the first rearrangement being nonproductive, and 4/9 cells will die from harboring two out-of-frame rearrangements (Figure 2.6, I). Determining the ratio of in-frame to out-of-frame rearrangements by sequencing would reveal that 71.4% of all VDJ rearrangements would be in-frame and 28.6% of rearrangements would be out-of-frame (Figure 2.6, I). Under the condition of biallelic synchronous V recombination, 1/9 cells will have an in-frame rearrangement on each allele, 4/9 cells will have an in-frame rearrangement on one allele and an out-of-frame rearrangement on the
other allele, and 4/9 cells would die due to out-of-frame rearrangements being made on both alleles (Figure 6.5, A). Thus, 60% of VDJ rearrangements in a population would be in-frame when allelic exclusion is completely subverted (Figure 6.5, A). Yet, as was explored in Chapter 3, the Tcrb locus permits each allele to produce two unique TCRβ proteins because of the presence of two DJCβ clusters and the V31 segment that rearranges by inversion. As a result, the frequency of in-frame rearrangements in V2R/V31R/V2R/V31R mice dipped below the 60% theoretical threshold because more than two rearrangements were attempted per cell. As a consequence of these considerations, the theoretical maximum of Tcrb allelic inclusion can be recalculated based upon whether three or four rearrangements are possible per cell. Assuming each Tcrb allele utilizes an upstream Vβ and V31, the frequency of VDJβ rearrangements in cells attempting three or four rearrangements in a synchronous manner would be 47.37% or 41.54%, respectively (Figure 6.5, B and C). As more rearrangements are possible, the maximum frequency of in-frame VDJ rearrangements decreases. At least for Tcrb, a caveat to this population based modeling is that it cannot account for an upstream Vβ rearrangement on one allele that excises the first VDJβ-Cβ1 join. Thus, even if a thymocyte attempted four rearrangements using only upstream Vβ segments, this sequencing approach would classify this cell as one that only made two rearrangements. Yet, so long as all rearrangements are amplifiable with V and J specific primers, this method could be applied to other AgR loci to assess the degree to which allelic exclusion is perturbed in various experimental settings, even for loci that possess multiple miniclusters such as lgl and Tcrg.

In summary, I have demonstrated that monoallelic and monogenic Tcrb assembly is enforced by low quality Vβ RSSs that stochastically restrict multiple Vβ rearrangements before transient and permanent feedback inhibition mechanisms silence Vβ recombination. These observations, and the models that I have generated, may also illuminate the possibility that weak Vβ RSSs that direct monogenic Tcrb assembly have an essential role in suppressing oncogenic chromosomal translocations. Given that Vh RSSs share similar features as Vβ RSSs, potentially
any or all of these findings are also applicable to the \textit{Igh} locus in B cells. Importantly, the mouse models detailed in this thesis may be useful to test: other physiological roles ascribed to allelic exclusion such as restricting autoimmunity and preventing the dilution of antigen-specific receptors on a lymphocyte surface, and the relationship between diversity of the TCR\(\beta\) repertoire and host immunity.
Figure 6.1. Working model of factors enforcing TCRβ allelic exclusion. In non-cycling DN3 thymocytes, (A) both Tcrb alleles are initially positioned in repressive chromatin environments such as the nuclear lamina and pericentric heterochromatin, which are refractory to V(D)J recombination. (B) Stochastic asynchronous accessibility through differential positioning of Tcrb alleles, histone modifications, and locus compaction allows one allele to become active to undergo Vβ-to-DJβ recombination. When one allele is activated in this manner, weak Vβ RSSs ensure monogenic Tcrb assembly. (C) If both alleles are accessible, weak Vβ RSSs ensure monoallelic Tcrb assembly. (D) RAG-mediated DNA DSBs activate ATM to transiently halt Vβ recombination and downregulate RAG expression, providing time to test the initial rearrangement. (E) If the initial rearrangement is out-of-frame, RAG is re-expressed to allow recombination of the other allele. However, if the rearrangement is in-frame, the VDJβ join is expressed as a TCRβ chain in a pre-TCR complex. pre-TCR signaled permanent feedback inhibition silences Vβ recombination, possibly through locus decontraction, nuclear positioning, and epigenetic modifications, and promotes cell cycle by transcriptional activation of Cyclin D3. (F) Cyclin D3 moves cells into S phase where RAG2 is degraded and may also directly repress Vβ accessibility. DN3 cells differentiate into DP thymocytes where recombination at Tcrb ceases and Tcra loci become accessible to recombine.
Figure 6.2. Model of factors contributing to a diverse Vβ repertoire. Replacement of the V1, V2, and V31 segment RSSs with the stronger 3'Dβ1 RSS reveal marked differences of their usage in the Vβ repertoire (V usage: V31 > V2 > V1). Shown are factors that could contribute to these differences.
12 possible \( V_L/V_H \) pairs = 12 antigen binding specificities
One B cell could express 144 possible antibodies

**B**

**Pre-selection repertoire:**

**Clonal Deletion**

**Tolerant post-selection repertoire:**

**Allelic Exclusion Intact**

**Non-autoreactive specificity: provides pro-survival signaling**

**Autoimmune Specificity (expressed at lower levels)**

**No Allelic Exclusion**

**Post-selection repertoire:**

These clones could contribute to autoimmunity
Figure 6.3. Possible physiological functions for AgR allelic exclusion. Each AgR allele is indicated in green and purple. (A) Allelic exclusion ensures proper lymphocyte function by preventing the cell-surface dilution of antigen-specific receptors. Assuming a B cell produces all possible in-frame rearrangements at \textit{Igh}, \textit{Igk}, and \textit{Igl} loci and there are no restrictions in \textit{V}_{L}/\textit{V}_{H} pairing, one B cell could express 144 unique BCRs with up to twelve unique antigen specificities. Thus, the receptor density for a given antigen is “diluted” relative to a cell expressing a single kind of BCR. (B) Allelic exclusion prevents autoimmunity. Assuming each lymphocyte in the repertoire expresses one AgR specificity, autoreactive B and T cells are efficiently purged from the pre-selection repertoire. Under the condition of allelic inclusion, cells expressing both autoreactive and non-autoreactive receptors allow lymphocytes to evade apoptosis by pro-survival signals derived from the non-autoreactive receptor. These cells can then bind self-antigen in the periphery and contribute to autoimmunity.
Figure 6.4. Weak Vβ RSSs may limit oncogenic Tcrb translocations resulting from biallelic Tcrb recombination. (A and B) Models of monoallelic and biallelic recombination. (A) One Tcrb allele undergoes Vβ-to-DJβ recombination in G1-phase cells, and transient and permanent feedback mechanisms inhibit additional RAG-mediated DNA DSBs as cells move into S phase. (B) As cells progress into S phase, a RAG DSB on the homologous allele could form a chromosomal translocation with DNA DSBs generated by DNA replication. Translocations that place an oncogene under regulation of the Tcrb enhancer (Eβ) could drive cellular transformation. (C) Kaplan-Meier Curve for p53−/− and p53−/−:V2R31R/V2R31R mice. P value was calculated by log-rank Mantel-Cox test. (D) Summary table: median survival time and number of mice used in the tumor cohort.
### A. Assuming Two Possible Rearrangements:

<table>
<thead>
<tr>
<th>Allele 1</th>
<th>Allele 2</th>
<th>Probability</th>
<th>Theoretical Maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td>IF</td>
<td>IF</td>
<td>$1/3^2 = 1/9$</td>
<td>$1/5 = 20%$</td>
</tr>
<tr>
<td>IF</td>
<td>OF</td>
<td>$(1/3)(2/3) = 2/9$</td>
<td>$2/5 = 40%$</td>
</tr>
<tr>
<td>OF</td>
<td>IF</td>
<td>$(2/3)(1/3) = 2/9$</td>
<td>$2/5 = 40%$</td>
</tr>
<tr>
<td>OF</td>
<td>OF</td>
<td>$(2/3)^2 = 4/9$</td>
<td>These Cells Die</td>
</tr>
</tbody>
</table>

By Sequencing:
- VDJ$_{IF}$: $(1/3) + 2(1) + 2(1) = 6$
- VDJ$_{OF}$: $(2/3) + 2(1) = 4$
- $6/10 = 60\%$

### B. Assuming Three Possible Rearrangements:

<table>
<thead>
<tr>
<th>DJ 1</th>
<th>DJ 2</th>
<th>DJ 1 or 2</th>
<th>Probability</th>
<th>Theoretical Maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td>IF</td>
<td>IF</td>
<td>IF</td>
<td>$(1/3)^2 = 1/27$</td>
<td>$1/19 = 5.26%$</td>
</tr>
<tr>
<td>IF</td>
<td>IF</td>
<td>OF</td>
<td>$(1/3)(2/3) \times 3$ permutations</td>
<td>$6/19 = 31.58%$</td>
</tr>
<tr>
<td>IF</td>
<td>OF</td>
<td>OF</td>
<td>$(1/3)(2/3)^2 \times 3$ permutations</td>
<td>$12/19 = 63.16%$</td>
</tr>
<tr>
<td>OF</td>
<td>OF</td>
<td>OF</td>
<td>$(2/3)^2 = 8/27$</td>
<td>These Cells Die</td>
</tr>
</tbody>
</table>

By Sequencing:
- VDJ$_{IF}$: $(1/3) + 4(2/3) + 1(1) = 27$
- VDJ$_{OF}$: $6(1) + 12(2) = 30$
- $27/57 = 47.37\%$

### C. Assuming Four Possible Rearrangements:

<table>
<thead>
<tr>
<th>DJ 1</th>
<th>DJ 2</th>
<th>DJ 1</th>
<th>DJ 2</th>
<th>Probability</th>
<th>Theoretical Maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td>IF</td>
<td>IF</td>
<td>IF</td>
<td>IF</td>
<td>$(1/3)^4 = 1/81$</td>
<td>$1/65 = 1.54%$</td>
</tr>
<tr>
<td>IF</td>
<td>IF</td>
<td>IF</td>
<td>OF</td>
<td>$(1/3)(2/3) \times 4$ permutations</td>
<td>$8/65 = 12.3%$</td>
</tr>
<tr>
<td>IF</td>
<td>IF</td>
<td>OF</td>
<td>OF</td>
<td>$(1/3)(2/3)^2 \times 6$ permutations</td>
<td>$24/65 = 36.9%$</td>
</tr>
<tr>
<td>IF</td>
<td>OF</td>
<td>OF</td>
<td>OF</td>
<td>$(1/3)(2/3)^2 \times 4$ permutations</td>
<td>$32/65 = 49.2%$</td>
</tr>
<tr>
<td>OF</td>
<td>OF</td>
<td>OF</td>
<td>OF</td>
<td>$(2/3)^4 = 16/81$</td>
<td>These Cells Die</td>
</tr>
</tbody>
</table>

By Sequencing:
- VDJ$_{IF}$: $(1/3) + 8(2/3) + 24(2/3) + 32(1) = 108$
- VDJ$_{OF}$: $6(1) + 24(2) + 32(3) = 152$
- $108/260 = 41.54\%$

Figure 6.5: Theoretical frequencies of in-frame and out-of-frame VDJ joins in cells attempting two or more rearrangements assuming synchronous V recombination. Shown are possible rearrangement outcomes and their probabilities for cells attempting two (A), three (B), and four (C) rearrangements. Cells that do not achieve one in-frame rearrangement die and are subtracted from the denominator. The Theoretical Maxima column denotes the frequency of cells with the possible rearrangement pattern. IF: in-frame. OF: out-of-frame.
APPENDIX: Materials and Methods

Mice. All experimental mice assayed in this study were 4-to-6 weeks old, of mixed sex, and housed under specific pathogen-free conditions at the Children’s Hospital of Philadelphia (CHOP). Animal husbandry, breeding, and experiments were performed in accordance with national guidelines and regulations and approved by the CHOP Institutional Animal Care and Use Committee. We used CRISPR/Cas9-mediated genomic editing in C57BL/6 zygotes to create mice carrying replacement of the V2 RSS with the normal 3’Dβ1 RSS (V2R allele) or variant 3’Dβ1 RSS (V2F allele), the V3f RSS with the variant 3’Dβ1 RSS (V3f allele), and the V1 RSS with the normal 3’Dβ1 RSS (V1R allele).

To replace the V2 RSS, we identified a suitable target protospacer adjacent motif (PAM) 5’ AGG 3’ located on the antisense strand of the V2 RSS spacer. We subcloned the 20-mer “Trbv2 gRNA target” sequence (Appendix Table 1) into the pSpCas9(BB)-2A-Puro vector and in vitro transcribed a single stranded guide RNA (sgRNA) using described methods (Ran et al., 2013). The CHOP Transgenic Core microinjected zygotes with a mixture of the sgRNA (8 uM), Cas9 protein (8 uM), and a single strand oligonucleotide (ssDNA) repair template (10 uM) (Chen et al., 2016; Wang et al., 2013). To generate the V2R allele, we used the ssDNA: 5’ GGA CTA CTG AAC TGA GTC CCC AGG CTC AGG TAG ACC AGT TAC ATC AAC AGT TTC CTG GAT CCA AGG AGG TTT TTG TAA AGG CTT CCC ATA GAA TTG AAT CAC CGT GTG TTG GCT GCT GGC ACA GAA GTA TGT GGC CGA GTC ATC AGG CTG AAG G 3’; Integrated DNA Technologies. To generate the V2F allele, we used the ssDNA: 5’ GGA CTA CTG AAC TGA GTC CCC AGG CTC AGG TAG ACC AGT TAC ATC AAC AGT TTC CTG GAT CCA CTG AGG AGG TTT TTG TAA AGG CTT CCC ATA GAA TTG GAG CAC CGT GTC TTG GCT GCT GGC ACA GAA GTA TGT GGC CGA GTC ATC AGG CTG AAG G 3’; Integrated DNA Technologies. For both V2R and V2F alleles, founders were identified by PCR on tail DNA using the 5’V2 and 3’Dβ1RSSRev primers and/or 3’V2 and 3’Dβ1RSS primer pairs. Each RSS replacement was then verified in homozygous mice by PCR-
sequencing using the 5’V2 and 3’V2 primers. For subsequent genotyping, primers 3’V2 and 3’Dβ1RSS were used to identify the V2R allele and primers 3’V2 and V2Fos were used to identify the V2F allele (See Appendix Table 1 for list of all primers).

To generate the V31F allele, we identified a suitable target PAM 5’ AGG 3’ located on the sense strand of the V31 RSS spacer. As outlined for the V2 RSS replacements above, we used the 20-mer “Trbv31 gRNA target” (Appendix Table 1) and the ssoDNA: 5’ CAG GCC GAA GGA CGA CCA ATT CAT CCT AAG CAC GGA GAA GCT GCT TCT CTA CCT CTG TGC CTG GAG TCT CAC GGT GCT CCA ATT CTA TGG GAA GCC TTT ACA AAA ACC ACA CCC TCT CTG TAG TCC TTC CTC CCT CAC TAG GAA CCC TCA CTA GGG ATG GGT GGA GGG GGT TTG CCA CTG AAT TT 3’; Integrated DNA Technologies. Founders were identified by PCR on tail DNA using the 5’V31 and 3’Dβ1RSSRev and/or the 3’V31 and 3’Dβ1RSS primer pairs. The RSS replacement was verified in homozygous mice by PCR-sequencing using the 5’V31 and 3’V31 primers. For subsequent genotyping, primers 3’V31 and 3’Dβ1RSS were used to identify the V31R allele and primers 3’V31 and V31Fos were used to identify the V31F allele (See Appendix Table 1 for list of all primers). We bred the V2R, V2F, and V31F alleles of founding mice to C57BL/6 mice for two-to-five generations. We then crossed heterozygous Vβ RSS replacement mice with each other, V31R/+ mice (Horowitz and Bassing, 2014), EβΔ/Δ mice (Leduc et al., 2000), or TcrbTg mice (Shinkai et al., 1993) to establish experimental mice as well as wild-type controls.

To generate the V1R allele, we adopted a more rapid approach (Quadros et al., 2017) and generated a crRNA using the guide sequence 5’ GAC ACA GTG GTA AAC TCT GC with a protospacer adjacent motif (PAM) 5’ AGG 3’ located on the sense strand of the V1 RSS spacer. The CHOP Transgenic Core microinjected C57BL/6 zygotes with a mixture of crRNA, tracrRNA, Cas9 protein, and a single stranded oligonucleotide (ssoDNA) prepared as described (Quadros et al., 2017). The ssoDNA: 5’ GAG GCT GCA AGT GGC CAA CAT GAG CCA GGG CAG AAC TTT ACA AAA ACC ACA CAC AGA CTA CCC TGC CTT CCA AGC CTT GCT CCT CTG CAA GCC CTC
CTG AGC TTT CTT 3’; Integrated DNA Technologies. For genotyping, the common primer 5’
TCG GCC ACA TTA GCT GTC TAC ATC C 3’ is used with primer 5’ CAC GGT GAT TCA ATT
CTA TGG GAA GCC TT 3’ to identify a V1R allele or primer 5’ CAC AGT GGT AAA CTC TGC
AGG CG 3’ to identify a wild-type allele.

**Flow Cytometry.** Single cell suspensions were prepared from the thymuses and spleens of mice,
depleted of red blood cells, and Fc receptors blocked using anti-CD16/CD32. All antibody stains
were performed in PBS containing 3% FCS and 0.1% NaN₃ (See Appendix Table 2 for list of all
antibodies). To determine effects on gross αβ T cell development, thymocytes were stained with:
anti-CD4, anti-CD8α, anti-TCRβ, anti-c-Kit, anti-CD25, and a lineage (Lin) panel composed of
anti-TCRγδ, CD11b, CD11c, CD19, B220, TER119, and NK1.1 antibodies. Differential
expression of c-Kit and CD25 in Lin’CD4’CD8TCRβ’ cells identify DN1-4 thymocytes. Gross
thymocyte development was assessed based on the expression of CD4 and CD8. Peripheral αβ T
cell numbers were determined by staining splenocytes with anti-CD4, anti-CD8α, and anti-
TCRβ antibodies and identifying CD4’TCRβ+ or CD8’TCRβ+ cells. To monitor Tcrb allelic
exclusion, we wished to avoid potential background staining artifacts that can result as a
consequence of utilizing biotinylated primary antibodies and streptavidin secondaries. Thus, we
ordered directly-conjugated anti-Vβ antibodies, most of which are available in only FITC and PE.
We stained cells in PBS containing 3% FCS and 0.1% NaN₃ with the following antibodies: anti-
CD4 APC-eFluor780, anti-CD8α Pacific Blue, and anti-TCRβ APC. In addition to the
aforementioned antibodies, we stained cells with anti-Vβ4 (V2) PE or anti-Vβ14 (V31) FITC, and
a corresponding antibody in either FITC or PE, respectively. These are: anti-Vβ10b (V4) PE, anti-
Vβ5.1, 5.2 (V12) FITC or PE, anti-Vβ6 (V19) FITC or PE, and anti-Vβ8 (V13) FITC or PE. To
measure allelic inclusion between V1 and V2 or V2 and V31, we stained with anti-CD4 APC-
eFluor780, anti-CD8α Pacific Blue, anti-TCRβ APC, anti-V2 PE, and either anti-V1 FITC or anti-
V31 FITC. Surface TCRβ expression was assayed on singlet and single positive (CD4+ and
CD8\(^{+}\)) cells. Data were collected on an LSR Fortessa and analyzed with FlowJo software (Tree Star). Cells were gated on the basis of forward and side scatter, singlets, and live.

To stain for intracellular cytokines, 10\(^6\) splenocytes were cultured with gp33 peptide in the presence of 2 \(\mu\)g/mL Brefeldin A (Sigma) and 2 \(\mu\)M Monensin (eBioscience) for 4 hours at 37°C. Cells were then stained for surface antigens, permeabilized using the Cytofix/Cytoperm kit (BD Biosciences), and stained for cytokines.

**LCMV Infection.** Mice were infected intraperitoneally with 2x10\(^5\) plaque-forming units of LCMV Armstrong. Mice were euthanized at days and 7 and 30 post-infection. LCMV Armstrong was provided by the Behrens Lab and gp33-tetramer was provided by the Wherry Lab.

**Generating and Analyzing \(\alpha\beta\) T cell Hybridomas.** We generated a panel of \(\alpha\beta\) T cell hybridoma clones using two independent splenocyte cultures from two different \(V2^{+/+}/V31^{-/-}\) mice, employing established methods and reagents (Sleckman et al., 1997). We characterized \(Tcrb\) rearrangements of each clone by Southern blot analyses using strategies and probes previously described (Bassing et al., 2000; Bassing et al., 2008; Khor and Sleckman, 2005; Wu et al., 2003). The \(V31^R\) allele contains an additional 101-bp sequence that distinguishes it from an unmodified \(V31\) allele, and using primers 5'V31 and 3'V31 permits us to determine in clones with \(V31\) rearrangements which \(V31\) rearranged (Horowitz and Bassing, 2014; Wu et al., 2003). We used the 5'V2 and 3'D\(\beta\)1RSSRev primers to PCR-identify which \(V2\) rearranged in clones with \(V2\) rearrangements. For clones with recombination of the RSS-replaced \(V2\) and \(V31\) segments, we PCR-sequenced each rearrangement using the 5'V2 or 5'V31 primer in combination with each of the J\(\beta\) reverse primers and PCR conditions previously reported (Wu et al., 2003). (See Appendix Table 2 for list of all primers).

**Cell Sorting.** Red blood cell-depleted single-cell suspensions of total thymocytes were stained and to sort for desired thymocyte populations. To sort DN1/2 and DN3 thymocytes, thymocytes
were stained with: anti-CD4, anti-CD8α, anti-TCRβ, anti-c-Kit, anti-CD25, and the Lin panel. Thymocytes were first gated on Lin−CD4−CD8−TCRβ− cells, and then sorted c-Kit+ cells to isolate DN1/2 cells or c-Kit−CD25+ cells to isolate DN3 cells. To sort CD4+ SP thymocytes, thymocytes were stained with: anti-CD4, anti-CD8α, and anti-TCRβ. Thymocytes were first gated on TCRβhi cells and then CD4+ cells.

**Real-Time Quantitative PCR Analysis.** To detect Vβ1(Dβ1)Jβ1.1, Vβ(Dβ)Jβ2.1, and Dβ2Jβ2.1 rearrangements, Taqman qPCR assays were performed on DNA isolated from sorted DN1/2 and DN3 thymocytes employing previously described reagents and methods (Gopalakrishnan et al., 2013). Total Vβ1(Dβ)Jβ1.1, Vβ(Dβ)Jβ2.1, and Dβ2Jβ2.1 rearrangements were normalized to an unrearranged region of the genome (CD19).

To detect V31-5’Dβ RSS hybrid joins, TaqMan qPCR was performed on DNA isolated from total thymocytes. 5’Dβ1 RSS probe: 5’ TTC CAG CCC TCA AGG GGT AGA C 3’. 5’Dβ1 RSS primer: 5’ GTC ACC TTC CTT ATC TTC AAC TCC CCC 3’. 5’Dβ2 RSS probe: 5’ GGG TAG GCA CCT GTG GGG AA 3’. 5’Dβ2 RSS primer: 5’ TCC CAG CCC CTC TCA GTC AG 3’. Common V31 primer: 5’ AAA TCA AGC CCT AAC CTC TAC 3’. These rearrangements were normalized to an unrearranged region of the genome (CD19).

**Quantification and Statistical Analysis.** Data are reported as mean ± SD. Statistical analyses were done with Prism 8. Unless otherwise indicated within Figure Legends, unpaired Student’s t tests were performed to generate p-values between two different experimental groups. One-way and two-way ANOVA tests were followed by Tukey, Dunnett, or Bonferroni post-tests.
## Appendix Table 1. Primers used in this study

<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Trbv2</em> gRNA target:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ GAGGAGGTTTGTGTACAGGG 3’</td>
<td></td>
</tr>
<tr>
<td><em>Trbv31</em> gRNA target:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ GAGTCTCACACTGAGTAGGG 3’</td>
<td></td>
</tr>
<tr>
<td>CD19 Probe:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’FAM-AGTCCTTACTGGTGAGGTAGAAGGT-3’ Iowa Black</td>
<td></td>
</tr>
<tr>
<td>CD19 qPCR Forward Primer:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ TGTCCTCTTCCTCCTTCTTCT 3’</td>
<td></td>
</tr>
<tr>
<td>CD19 qPCR Reverse Primer:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ CTCAACTCAAAGACTTTT 3’</td>
<td></td>
</tr>
<tr>
<td><em>Trbv2</em> qPCR Forward Primer:</td>
<td>Gopalakrishnan et al., 2013; IDT</td>
</tr>
<tr>
<td>5’ GTGTCAGAAAAACCATTTAG 3’</td>
<td></td>
</tr>
<tr>
<td><em>Trbv31</em> qPCR Forward Primer:</td>
<td>Gopalakrishnan et al., 2013; IDT</td>
</tr>
<tr>
<td>5’ AAATCAAGCCCTAACCTAC 3’</td>
<td></td>
</tr>
<tr>
<td>Jβ1.1 qPCR Reverse Primer:</td>
<td>Gopalakrishnan et al., 2013; IDT</td>
</tr>
<tr>
<td>5’ CTCAATATGGACACGGAGGACATGC 3’</td>
<td></td>
</tr>
<tr>
<td>Jβ2.1 qPCR Reverse Primer:</td>
<td>Gopalakrishnan et al., 2013; IDT</td>
</tr>
<tr>
<td>5’ CCTGATACAGGGCCTTGGATAGTTA 3’</td>
<td></td>
</tr>
<tr>
<td>Jβ1.1 Probe:</td>
<td>Gopalakrishnan et al., 2013; IDT</td>
</tr>
<tr>
<td>5’FAM-TGTAGTCTCCTTACAAAAACCCAGGT-3’ Iowa Black</td>
<td></td>
</tr>
<tr>
<td>Jβ2.1 Probe:</td>
<td>Gopalakrishnan et al., 2013; IDT</td>
</tr>
<tr>
<td>5’HEX-GGACAGACAGTCTCTTCTCAT-3’ Iowa Black</td>
<td></td>
</tr>
<tr>
<td>5’V2:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ CTGGTGGCAGTCACAGGG 3’</td>
<td></td>
</tr>
<tr>
<td>3’V2:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ AAAAAATTATCAAACCTAGAAGCC 3’</td>
<td></td>
</tr>
<tr>
<td>5’V31:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ TATCCATCAATGGCCAGTGTTCCG 3’</td>
<td></td>
</tr>
<tr>
<td>3’V31:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ AAGAGTAGCCTGTTAAGGAGCCGG 3’</td>
<td></td>
</tr>
<tr>
<td>3’Dβ1RSS:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ CACGGTGATTCTTCTATGGGAAGCCTT 3’</td>
<td></td>
</tr>
<tr>
<td>3’Dβ1RSSRev:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ GTAAGGGCTTCCCATTAGAATTCACCGTG 3’</td>
<td></td>
</tr>
<tr>
<td>V2Fos:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ GCCAGCAGCAGACACACGTTGCTC 3’</td>
<td></td>
</tr>
<tr>
<td>V31Fos:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ CCTCTGTGCTTGAGTGTCACACGCTC 3’</td>
<td></td>
</tr>
<tr>
<td>Jβ1.1Rev:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ GCAGTCGCTACCTTGCCCCG 3’</td>
<td></td>
</tr>
<tr>
<td>Jβ1.2Rev:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ ACCCGAAGACCCCTACG 3’</td>
<td></td>
</tr>
<tr>
<td>Jβ1.3Rev:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ ACCCAGGCATTGGGCCCTAAC 3’</td>
<td></td>
</tr>
<tr>
<td>Jβ1.4Rev:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ CAACACCCCTCAGGAGAAG 3’</td>
<td></td>
</tr>
<tr>
<td>Jβ1.5Rev:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>5’ CCCACCAGTTTGGTCACATAG 3’</td>
<td></td>
</tr>
<tr>
<td>Jβ1.6Rev:</td>
<td>This study; IDT</td>
</tr>
<tr>
<td>Primer Name</td>
<td>Sequence</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>Jβ2.1Rev</td>
<td>5' CAGAAATGGAGCCCCCATACC 3'</td>
</tr>
<tr>
<td>Jβ2.2Rev</td>
<td>5' CCTGTATGGCCTCTGCCTTC 3'</td>
</tr>
<tr>
<td>Jβ2.3Rev</td>
<td>5' CCCATGGTACCCCCAGCTG 3'</td>
</tr>
<tr>
<td>Jβ2.4Rev</td>
<td>5' CCTCCCCCTTCATTACTAGCTCC 3'</td>
</tr>
<tr>
<td>Jβ2.5Rev</td>
<td>5' CCCAACAAAACTATACCCCAGCT 3'</td>
</tr>
<tr>
<td>Jβ2.7Rev</td>
<td>5' CTGAGAACGCGCACGTGG 3'</td>
</tr>
<tr>
<td>Jβ2.7Rev</td>
<td>5' TTGGGTGAAGCGAGATGTGAAT 3'</td>
</tr>
</tbody>
</table>
### Appendix Table 2. Key Reagents Table

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Source</th>
<th>Catalog</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-B220 PE (Clone RA3-6B2)</td>
<td>BD Biosciences</td>
<td>Cat# 553090</td>
</tr>
<tr>
<td>Anti-CD4 APC-eFluor780 (Clone RM4-5)</td>
<td>eBioscience</td>
<td>Cat# 47-0042-82</td>
</tr>
<tr>
<td>Anti-CD8α Pacific Blue (Clone 53-6.7)</td>
<td>BD Biosciences</td>
<td>Cat# 558106</td>
</tr>
<tr>
<td>Anti-CD11b PE (Clone M1/70)</td>
<td>BD Biosciences</td>
<td>Cat# 553311</td>
</tr>
<tr>
<td>Anti-CD11c PE (Clone HL3)</td>
<td>BD Biosciences</td>
<td>Cat# 561044</td>
</tr>
<tr>
<td>Anti-CD16/CD32</td>
<td>BD Biosciences</td>
<td>Cat# 553141</td>
</tr>
<tr>
<td>Anti-CD19 PE (Clone 1D3)</td>
<td>BD Biosciences</td>
<td>Cat# 553786</td>
</tr>
<tr>
<td>Anti-CD25 PE-Cy7 (Clone RA3-6B2)</td>
<td>BD Biosciences</td>
<td>Cat# 552880</td>
</tr>
<tr>
<td>Anti-CD117 (c-Kit) APC (Clone 2B8)</td>
<td>BD Biosciences</td>
<td>Cat# 553356</td>
</tr>
<tr>
<td>Anti-NK1.1 PE (Clone PK136)</td>
<td>BD Biosciences</td>
<td>Cat# 553165</td>
</tr>
<tr>
<td>Anti-TCRβ APC (Clone H57-597)</td>
<td>BD Biosciences</td>
<td>Cat# 561080</td>
</tr>
<tr>
<td>Anti-TCRβ PE (Clone H57-597)</td>
<td>BD Biosciences</td>
<td>Cat# 561081</td>
</tr>
<tr>
<td>Anti-TCRγδ Chain PE (Clone GL3)</td>
<td>BD Biosciences</td>
<td>Cat# 553178</td>
</tr>
<tr>
<td>Anti-TER-119 PE (Clone TER-119)</td>
<td>BD Biosciences</td>
<td>Cat# 553673</td>
</tr>
<tr>
<td>Anti-Vβ4 (Trbv2) PE (Clone KT4)</td>
<td>BD Biosciences</td>
<td>Cat# 553366</td>
</tr>
<tr>
<td>Anti-Vβ5.1, 5.2 (Trbv12) FITC (Clone MR9-4)</td>
<td>BD Biosciences</td>
<td>Cat# 553189</td>
</tr>
<tr>
<td>Anti-Vβ5.1, 5.2 (Trbv12) PE (Clone MR9-4)</td>
<td>BD Biosciences</td>
<td>Cat# 553190</td>
</tr>
<tr>
<td>Anti-Vβ6 (Trbv19) FITC (Clone RR4-7)</td>
<td>BD Biosciences</td>
<td>Cat# 553193</td>
</tr>
<tr>
<td>Anti-Vβ6 (Trbv19) PE (Clone RR4-7)</td>
<td>BD Biosciences</td>
<td>Cat# 553194</td>
</tr>
<tr>
<td>Anti-Vβ8.1, 8.2, 8.3 (Trbv13) FITC (Clone F23.1)</td>
<td>BD Biosciences</td>
<td>Cat# 553861</td>
</tr>
<tr>
<td>Anti-Vβ8.1, 8.2, 8.3 (Trbv13) PE (Clone F23.1)</td>
<td>BD Biosciences</td>
<td>Cat# 553862</td>
</tr>
<tr>
<td>Anti-Vβ10 (Trbv4) PE (Clone B21.5)</td>
<td>BD Biosciences</td>
<td>Cat# 553285</td>
</tr>
<tr>
<td>Anti-Vβ14 (Trbv31) FITC (Clone 14-2)</td>
<td>BD Biosciences</td>
<td>Cat# 553258</td>
</tr>
</tbody>
</table>

| **Chemicals, Peptides, and Recombinant Proteins** | | |
| Recombinant Human IL-2 | Peprotech | Cat# 200-02 |
| Concanavalin A (ConA) | Sigma-Aldrich | C5275 |
| HAT Supplement | ThermoFisher | Cat# 21060017 |

| **Cell Lines** | | |
| Mouse: BW5147 | ATCC | TIB-47 |

| **Mouse Strains** | | |
| Mouse: C57BL/6J | The Jackson Laboratory | JAX: 000664 |
| Mouse: V2"""" | This study | N/A |
| Mouse: V35"""" | Horowitz and Bassing, 2014 | N/A |
| Mouse: V2"""" | This study | N/A |
| Mouse: V35"""" | This study | N/A |
| Mouse: DO11.10 Tcrb transgene (Tcrb") | Shinkai et al., 1993 | N/A |
| Mouse: Eβ Knockout (Eβ") | Leduc et al., 2000 | N/A |
| Mouse: ATM Knockout (Atm") | Borghesani et al., 2000 | N/A |

| **Recombinant DNA** | | |
| pSpCas9(BB)-2A-Puro (PX459) | Addgene | Cat# 62988 |
BIBLIOGRAPHY


