RNA:DNA Heteroduplex Resolution in B-Lymphocyte Immunoglobulin Diversification and Genomic Maintenance

David Kazadi

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2016

ABSTRACT

RNA:DNA Heteroduplex Resolution in B-Lymphocyte Immunoglobulin Diversification and Genomic Maintenance

David Kazadi

Immunoglobulin (Ig) gene diversification plays an essential role in adaptive immunity. Faced with a continuous yet varied stream of self, non-self, and possibly harmful molecules, many organisms have mechanisms in their arsenal that have evolved to match the diversity of the antigens they encounter. In humans and mice, developing B and T lymphocytes go through a first round of genomic alteration — V(D)J recombination — in the bone marrow and the thymus, respectively. B cells can subsequently undergo two additional Ig gene diversification processes in secondary lymphoid tissues. Through somatic hypermutation (SHM), Ig variable regions of stimulated germinal center (GC)-forming B cells are mutated and further diversified, enabling affinity maturation. During class-switch recombination (CSR), on the other hand, B cells in the GC or prior to entering the GC recombine Ig constant regions, swapping the IgM-encoding locus for another isotype constant regions gene (e.g., IgG_1, IgG_3, IgE, IgA) to allow for different effector functions. Both B cell-specific genomic alterations are initiated when the single-stranded DNA (ssDNA) mutator enzyme activation-induced cytidine deaminase (AID) catalyzes the removal of the amino group off deoxycytidine residues, resulting in deoxyuridines and dU:dG mismatches. Low-fidelity cellular responses to the presence of dU, including the mismatch repair (MMR) and the base-excision repair (BER) pathways, are then thought to introduce mutations in

SHM and CSR, as well as cause double-strand breaks (DSBs) repaired through canonical and alternative non-homologous end-joining in CSR.

Though necessary for proper physiological function, these lymphocyte genome diversification processes are rife with danger for B cells and there is strong selective pressure to carefully orchestrate and target them so as not to threaten the genomic integrity of the cells through breaks or other mutations at non-Ig loci. Yet, these events can still occur, as demonstrated by the implication of AID with translocations found in some cancers (e.g., *c-MYC*:*IGH* in Burkitt's lymphoma). Therefore, the mechanisms underlying AID mutagenic activity targeting to physiological deamination substrates have been the focus of several studies.

Protein kinase A (PKA)-dependent phosphorylation of AID at its serine 38 residue has been shown to enable its interaction with replication protein A (RPA) before binding to ssDNA. Others have reported that SPT5 helps target AID to sites of RNA polymerase II (Pol II) stalling, such as the Ig switch sequences. Another cofactor, the RNA exosome complex, helps target the ssDNA mutator AID to both strands of DNA in vivo. The RNA exosome had hitherto been described in the context of RNA processing and degradation as 3' → 5' exoribonuclease. Sterile transcript-generating transcription at Ig loci was known to be required for proper AID catalytic activity; the newly described link between the RNA exosome and AID activity raised the prospect that RNA processing, and not mere transcription, might be playing a role in shaping the diversification of the immune repertoire in B lymphocytes.

During CSR, transient three-strand structures called R loops are generated. R loops are formed as the nascent transcript invades the DNA duplex, hybridizing to the template strand, and displacing the non-template one. The G-rich nature of the non-template strand is posited to help

stabilize the R loop, which allows the ssDNA mutator AID to use the exposed, non-template strand as a substrate. AID must then access the template strand. Here, we investigate the role that the RNA exosome and a potential cofactor, the putative RNA/DNA helicase senataxin (SETX), play in the sequence of biological events that result in CSR while protecting the cell from R-loop accumulation-associated genomic instability.

Table of Contents

List of Figures

Acknowledgments

It takes a village, the saying goes. For my PhD, it took a Congolese-French-American global village to get me to where I find myself today. Accordingly, the acknowledgments will go back, way back, to where it all began, and yet, I know it will be incomplete…

Academic. It goes without saying that I am where I am, because of the cumulated help, nudging, attention I received along the way: Les Oisillons, René Descartes/L'École Française, Les Loupiots, Collège Saint-Ouen, Pensionnat Jean-Baptiste de la Salle, Hunter College/City University of New York, Europa Kolleg, and, finally, Columbia University in the City of New York. I would like to heartily acknowledge the Hunter College MBRS Program (Barbara Thorsen, Janerie Rodriguez) and David Mootoo's laboratory, which introduced me to academic research, and the Hunter College Thomas Hunter Honors Program. Philip Alcabes, thank you for Angels in America. I would as well as like to acknowledge the Columbia University Summer Research Program (Andrew Marks, Ron Liem) and the Liem lab (Adijat Adebola in particular), which introduced me to biomedical research and to Columbia. Few people have helped me as much as Ronald Liem has. I sincerely do not know how to thank him enough. The MD/PhD program at Columbia (Patrice Spitalnik, Ron Liem, Michael Shelanski, Steve Reiner, Jeffrey Brandt, Stacy Warren) supported me when I needed it most; for this, I am so very grateful. Patrice, thank you. The medical school at Columbia has also been very supportive. Hilda Hutcherson, thank you, I will never forget. Lisa Mellman, thank you. We still have the whole major clinical year ahead of us. I hope it all goes smoothly. Alessandra Pernis-Lowell, thank you, thank you, thank you. Kind, kindred spirit comes to mind whenever I think of you. Emily

Dimango, thank you. The Integrated Program in Cellular, Molecular, and Biomedical Studies (Ron Liem, Lorraine Symington, Zaia Sivo, Lori Sussel): a warm thank you, and a socially awkward scientist hug. The grad school at Columbia (Fred Loweff) and the Department of Microbiology & Immunology (Elizabeth Scott, Carla Horne, Amir Figueroa): the laboratories, particularly those of the 9th floor: Reiner Lab (Will Adams, Julie Chaix, Simone Nish, Nyanza Rothman), Ivaylov Lab (Casandra Panea), Reizis Lab (Vanja Sisirak); let me also thank the Goff Lab on the 13th floor, as well as the Ghosh and Symington Labs on the 12th floor, though I'm convinced I have borrowed things from every lab in the department. Many thanks to the Cancer Center (Murty Vundavalli, your collaboration on the T-FISH analysis has been so very valuable; Kenta Yamamoto in the Zha Lab, thank you). Some additional support came from the National Institutes of Health/National Institute of Allergy and Infectious Diseases in the form of a Ruth Kirchstein pre-doctoral award: thank you for this grant.

I would like to emphatically express my gratitude to my qualifying examination committee (Ulf Klein, Kang Liu, Sankar Ghosh) and my dissertation committee (Lorraine Symington, Sankar Ghosh, Boris Reizis; I should also include Peter Nagy, who has agreed to be on my defense committee, although it's a preëmptive thank you). I feel that my committee has had my best interests at heart and I appreciate the investment in my growth as a scientist. I have truly been fortunate.

I would like to thank my lab, the Basu lab (not the Kazadi lab — yet!): today in the lab we have Jianbo Sun, Gerson Rothschild, Junghyun Lim, Pankaj Giri; but yesterday, we had Evangelos "Angelo" Pefanis. Jaime Chao, Michael Closser, Sabrina Tasnova, Celia Keim….and Veronika Grinstein. I have learned so much in these past few years. Angelo, without your mouse,

none of this would have been possible: thank you. I have greatly valued your thoughts and opinions on science over these years. Jianbo, you have been an intellectual constant in the lab; many thanks for your help all these years. Guoxin, I thank you for the IP in Ramos cells. Gerson, thank you for your help throughout these years; one really doesn't know what hard work is until one sees Gerson in action. Pankaj, thank you for introducing me to Endnote. What I would have done without this tool, I do not know. A particular thank you goes out to Veronika, for her indispensable support, for being my lab sibling, and overall for being a true friend. I wish you all the best. Another particular thank you goes out to Junghyun. Here is another kindred spirit. Your intellectual acumen, your work ethic, and your overall demeanor are quite unique. I have learned and benefited so much from having you in the lab. Thank you!

And finally, Uttiya. Part of me feels like nothing I write will do justice to the amount of gratitude, regard, affection, admiration, and respect I have for you. I have been fortunate to come across you. From the time I stepped into your lab on the 13th floor, you have been an excellent mentor, a true role model. You knew when to be tough, and you have always had my back. I feel like I have grown as a scientist, and as a human being, thanks to you. Your drive, perfectionism, and energy are really something to behold. I look forward to the hopefully many years of fruitful collaboration.

Personal. I will start with my family, my cornerstone, my anchor. I love you guys. We've been through oh-so-much. *But still, I'll rise!* Thanks for your support, thanks for being there, whether I'm up or down. Les mots me manquent. I love you. A special note for my friends, whose company I have so missed. Thanks for always extending an invitation, even when you knew the answer would be "I can't, I'm in/I've got to go to lab!" A few of you who have been of

particular help over the years in matters scientific, or related to my development as a scientist. Gérard Brun, *merci*! Andrew Pantoja, Olivia Kouadio: thank you guys for your support. I love you. Alejandro “Jaño” Ramirez Vallejo, Alexander Lyashchenko, Marco Russo, Catherine Chang, Veronika Grinstein (again), Vanja Sisirak (again), Fred Hitti, Kimberly Robinson Point du Jour, Yiorgios Mountoufaris, Tim Zeiske, Yann Ravussin, James Papizan, Kara Marshall, Aarti Sharma, Kevin Kanning, Brian Stater, Luis Arnes. Guys: we did it!

Education has too long been the province of the powerful,

a privilege for those at the top.

Today still, somewhere in the world,

a little child will be prevented

from learning how to read

because she is a girl;

a little boy will have to quit school

and start working to help his family.

In another life, at another time, that could have been my lot.

This work is dedicated to the memory of those

who fought,

whose faces I will never know,

and who, through acts small and large,

contributed to making this world a place

where liberty and equality

are decreasingly foreign,

and where someone like me

could live freely and fulfill,

if not his life's full potential, at least a substantial part of it.

Chapter I: Introduction

I.1. An Evolutionary Perspective

Conceptual evidence for what we today refer to as the immune system can be found as far back as in some of the earliest forms of terrestrial life. Deoxyribonucleic acid (DNA) restriction and surface exclusion in bacteria and archaea, for instance, evolved alongside horizontal gene transfer as a way to curtail the oftentimes maladaptive consequences of the latter. In eukaryotes, ribonucleic acid (RNA)-interference provides an early and commonly found example of a mechanism to defend against invading foreign genomic material. Similarly, the acquisition of nutrients from the environment is thought to have been accompanied by the emergence of rudimentary mechanisms, which helped distinguish between self and non-self, and which preceded more complex innate immune systems. However, the appearance of the adaptive immune system is relatively recent on the evolutionary time scale (Figure 1.1).

Investigators of the immune system have traditionally distinguished between innate and adaptive immunity. The former allows for a rapid, stereotyped, and largely nonspecific response, while the latter directs a less rapid but more specific response, typically within a multicellular organism, with the added feature of a "memory" of the response-eliciting agent, should it be encountered again. Adaptive immunity is thought to have first appeared in jawed vertebrates around 550 million years ago (Figure 1.1), though new discoveries have suggested its presence in jawless vertebrates as well, perhaps by convergent evolution (Kaufman, 2010, Hirano et al., 2011, Boehm, 2012). It is believed that these jawed animals were likely to have been predatory, and to have evolved to sustain the food-chain homeostasis by maintaining high predator-to-prey ratio and, consequently, having few offspring. According to this "offspring protection hypothesis," the need to protect the few offspring they did produce may have provided the

evolutionary pressure that selected for this very potent form of immune system defense

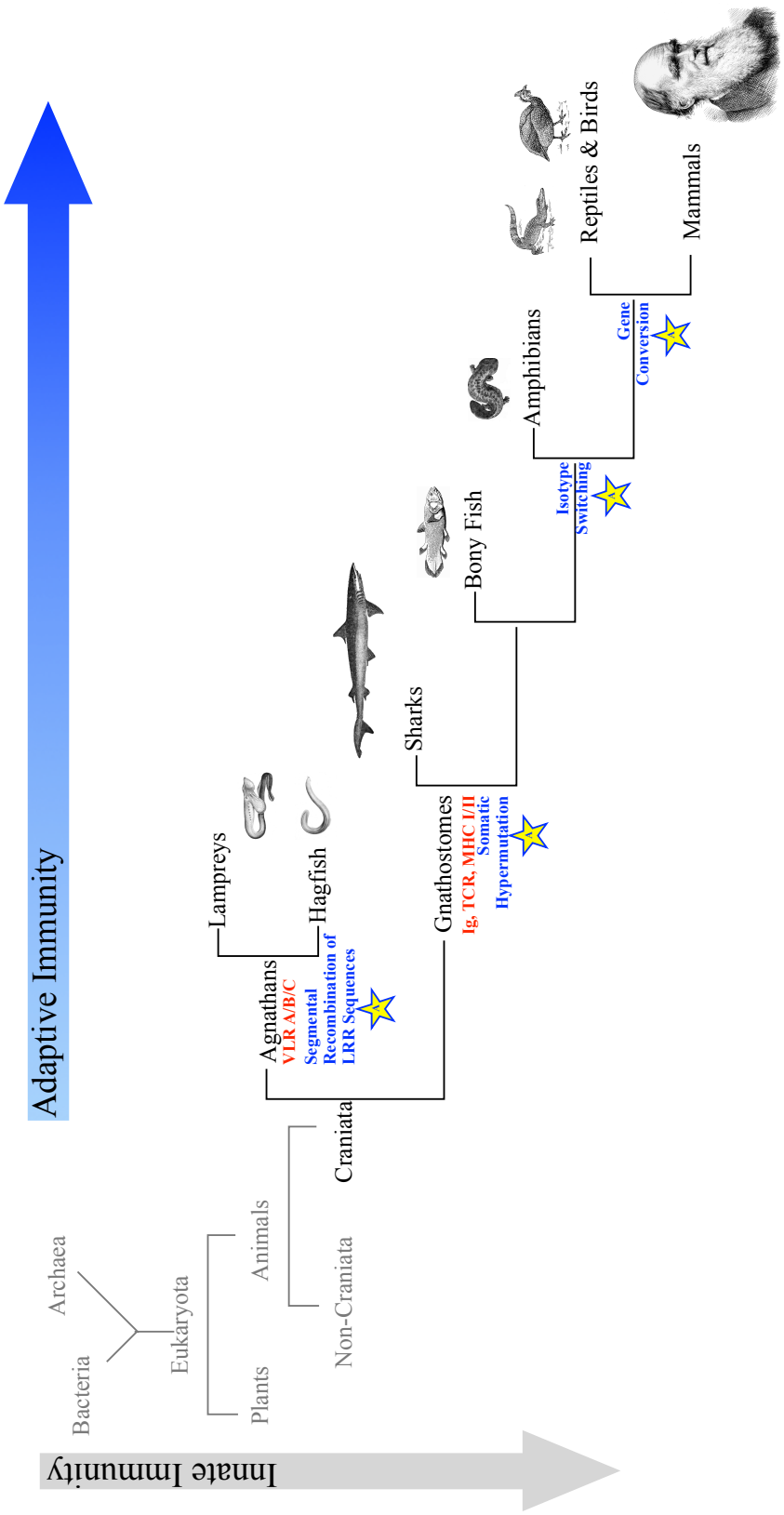


Figure 1.1. Evolution of innate immunity and adaptive immunity. Major classification groups are shown. Members of the different vertebrate groups are illustrated. The appearance of receptors mediating adaptive immunity is indicated in red. In blue are shown the immune receptor diversification mechanisms when they first appear. The yellow star containing an "A" points to the involvement of a cytidine deaminase.

mechanism (Flajnik and Kasahara, 2010). Another hypothesis stipulates that because the early vertebrates were so morphologically and functionally complex, they were more likely to occupy a wider variety of ecological niches, which would, as a consequence, expose them to a wider variety of pathogens. For this "habitat hypothesis," habitat expansion, and the accompanying expanded pathogen exposure, may have been a strong selective drive in the evolution of early vertebrate immune systems (Hedrick, 2004, Boehm, 2012). A third hypothesis suggests that the habitat that the early vertebrates themselves provided to a host of non-pathogenic microbes may have helped shape the evolution of the adaptive immune system. This "mutualism hypothesis" posits that the novel, complex features specific to the immune system of early vertebrates favored the evolution of complex, reciprocally beneficial relationships between autochthonous flora and host organism. The metabolic windfall in nutrients or the protection from pathogenic, allocthonous flora may then have helped drive the evolution of the adaptive immune system in early vertebrates (Lee and Mazmanian, 2010, Boehm, 2012).

Though these three, non-mutually exclusive hypotheses help us understand how an immune system as sophisticated as the one we and most other mammals now possess may have arisen, they say little about the tradeoff that may have taken place. In particular, we know that in humans, a number of pathologies are intrinsically related to the immune system: allergies and hypersensitivities, proliferative and oncological disorders, and immunodeficiencies. The evolution of the adaptive immune system may offer some clues. The different immune diversification methods seem to have appeared in association with the activity of members of the cytidine deaminase family: in agnathans, segmental recombination of the leucine-rich repeats (LRR) of variable lymphocyte receptors (VLR) resulted from the combined activity of DNA

cytosine deaminases 1 and 2 (CDA 1 and CDA 2); in gnathosomes, somatic hypermutation (SHM), class-switch recombination (CSR), and gene conversion, are all initiated by activation-induced cytidine deaminase (AID) (Figure 1.1) (Kaufman, 2010, Hirano et al., 2011, Boehm, 2012). It therefore behooves us to better understand how cytidine deaminase activity and the principal actors of adaptive immune system reactions interact to achieve proper physiological events, as well as how they may contribute to immunopathologies. We shall do so once we briefly review the different cell types and functions of the immune system, discuss the molecular machinery behind adaptive immunity recombination reactions, and, finally, give a presentation of the relevance of our projects.

One final note, before we delve into the above: there is some poetic "justice" of the evolutionary kind (from our vertebrate perspective, at least), in realizing that the invasion of the immunoglobulin loci ancestor by transposons (along with genome-wide duplications) may have set in motion a series of selective events that resulted in the rise of the adaptive immune system of vertebrates, a system much better tailored against the vagaries of infection by epitope-altering invaders like viruses, which, of course, are themselves related to transposons (Flajnik and Kasahara, 2010). In other words, it's as if the evolution of the adaptive immune system is the result of the invaded organism modifying and using the very weapons of the invader to fend off further invasions. An early pressure for the evolution of deaminase-type enzymes might have been to combat viral/transposon presence in the cell, a defense device which might have then become coöpted to diversify the transposon-invaded immunoglobulin loci ancestor. This might perhaps also give us an enhanced appreciation for the different nucleic acid tropisms of deaminases.

I.2. Cell Types and Functions of the Immune System

I.2.1. Development of Myeloid and Lymphoid Cellular Lineages

Over evolutionary time, different cells have arisen, which deal directly with the immune response, by sensing pathogens, secreting cytokines, or neutralizing antigens. The cells implicated in these functions broadly fall into two lineages, which largely correspond to the two branches of immunity (Figure 1.2). Myeloid lineage cells derive from the common myeloid progenitor, and are as varied as erythrocytes, macrophages, and granulocytes; myeloid immune cells are typically involved with innate responses. Lymphoid lineage cells, including B and T lymphocytes, stem from the common lymphoid progenitor, and mediate adaptive immunity. Both myeloid and lymphoid lineages are part of the hematopoietic system and share a common ancestor, the pluripotent hematopoietic stem cell.

Hematopoiesis starts as early during mouse and human embryonic development as at the yolk sac stage, and continues in several locations throughout development and after birth, including the aorta, gonad, mesonephros (or AGM, a region neighboring the dorsal aorta at the level of the embryonic middle kidney), the placenta, the fetal liver, and finally, the thymus and the bone marrow (Müller et al., 1994, Galloway and Zon, 2003, Gekas et al., 2005, Ottersbach and Dzierzak, 2005, Inman and Downs, 2007, Orkin and Zon, 2008). It is at the fetal liver stage that the pluripotent hematopoietic stem cell (HSC) gives rise to both multipotent progenitors: the common myeloid progenitor and the common lymphoid progenitor (Orkin and Zon, 2008). The common myeloid progenitor in turn yields other multipotent progenitors: the megakaryocyte/ erythroid progenitor (which ultimately gives rise to red blood cells and megakaryocytes/

platelets), the granulocyte/macrophage precursor (from which descend monocyte/macrophages and granulocytes, i.e., mast cells, eosinophils, and neutrophils), and the macrophage/DC progenitor (from which monocytes, some macrophages, and the common DC precursors derive) (Orkin and Zon, 2008, Liu et al., 2009, Geissmann et al., 2010). Several transcription factors are important in the various differential development programs. Among them, GATA-1, a member of a family of transcription factors that recognize the DNA consensus motif 5'-WGATAR-3' (where W= A or T, and R = A or G), in particular, is crucial at different time points for several lineages, including: erythrocytes, megakaryocytes, mast cells, and eosinophils (Orkin and Zon, 2008). Another transcription factor, PU.1, which recognizes the purine-rich PU-box 5'-GAGGAA-3', is necessary for the proper development of monocytes/macrophages and granulocytes (Orkin and Zon, 2008).

One important component of the myeloid lineage is the extensively heterogeneous class of cells discovered over 40 years ago by Ralph Steinman and Zanvil Cohn and called dendritic cells (DCs), based on their "branching" morphological appearance (Steinman and Cohn, 1973, Steinman and Cohn, 1974). DCs can be grouped based on their tissue residence status in two sets: lymphoid tissue or conventional DCs and non-lymphoid tissue DCs. Conventional DCs can be divided into two large groups ($CD8\alpha^+$ DCs and $CD4^+CD11b^+$ DCs), and are distinct from migratory DCs and plasmacytoid DCs. Plasmacytoid DCs are involved in type I interferon responses. Migratory DCs typically express either CD103 or CD11b, though they express both in the lamina propria. Moreover, DCs in the outer epidermal layer of stratified epithelia are termed Langerhans cells. Finally, inflammation can cause the appearance of inflammatory DCs derived from monocytes (Ganguly et al., 2013, Schlitzer and Ginhoux, 2014).

The common lymphoid progenitor gives rise to T cells, B cells, and innate lymphoid cells (ILCs). ILCs constitute a class of effector lymphoid cells that do not undergo receptor rearrangement like T cells, and that bear markers typically found on some myeloid cells (Killig et al., 2014). ILCs are comprised of natural killer (NK) cells (type 1 ILC), type 2 ILC, and type 3 ILC. While the description of the ILC2 and ILC3 groups is still in its infancy, NK cells have been investigated more deeply, and their MHC I-mediated surveillance of tumors and virally infected cells is well characterized (Vivier et al., 2008).

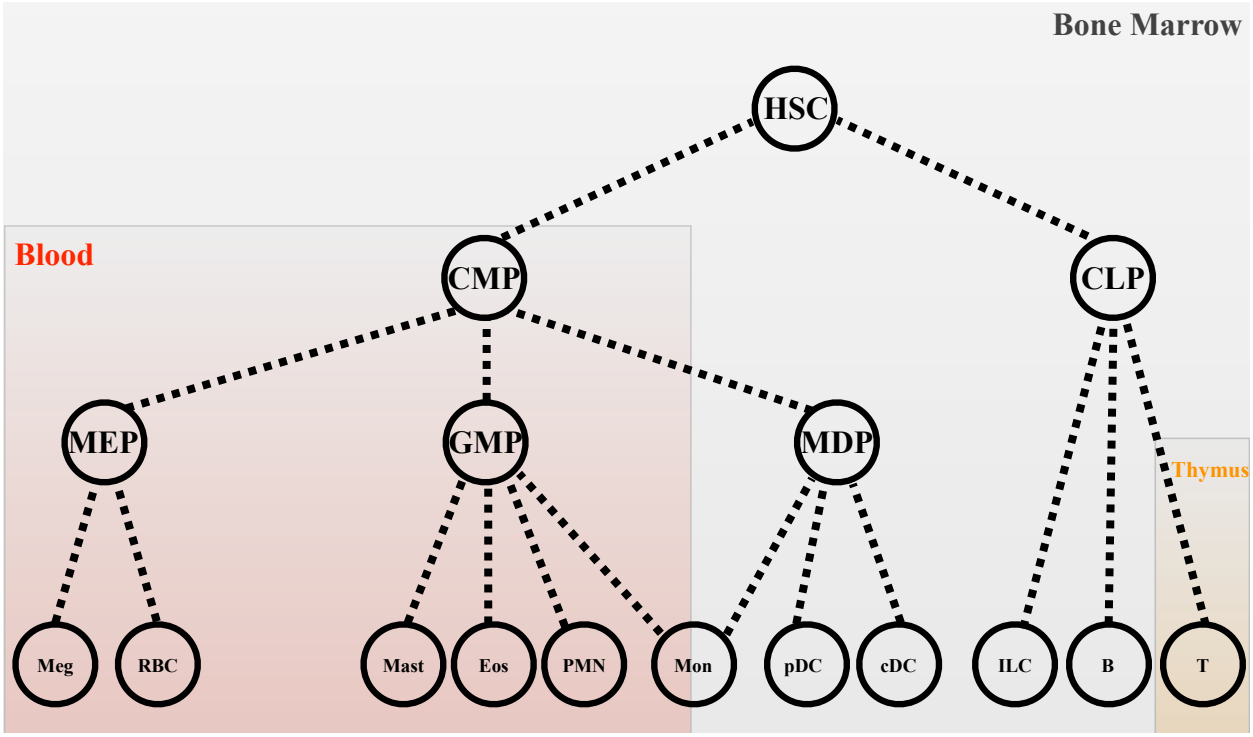


Figure 1.2. Development of myeloid and lymphoid compartments. Cells involved in the immune response derive from the hematopoietic stem cell (HSC). Hematopoiesis starts *in utero*. Shown here are the different immune cells (before further differentiation in the periphery) and their precursors, as well as the tissues in which they develop in the adult. HSC: hematopoietic stem cell; CMP: common myeloid progenitor; CLP: common lymphoid progenitor; MEP: megakaryocyte/erythroid progenitor; GMP: granulocyte/monocyte progenitor; MDP: monocyte/dendritic progenitor; Meg: megakaryocyte; RBC: red blood cell; Eos: eosinophil; PMN: polymorphonuclear cells (neutrophils); Mon: monocyte; pDC: plasmacytoid dendritic cell; cDC: conventional dendritic cell; ILC: innate lymphoid cell.

T cells, so-named because part of their development is done in the thymus, play a central, integrative role in the adaptive immune system reaction, particularly in activating B cells to produce antibodies against protein antigens. Some of the transcription factors that allow for the proper development of T cells (γδ T cells, αβ $CD4^+$ T cells, αβ $CD8^+$ T cells) include Notch and T cell factor-1 (TCF-1) (Orkin and Zon, 2008). αβ $CD4^+$ T cells can further differentiate into T helper 1 (T_H1), T_H2, T_H17, induced regulatory T cells (iT_{REG}), and T follicular helper (T_{FH}), whose role is crucial in aiding B cell activation in the germinal center reaction; these different subsets appear to be more plastic than initially thought, even if they still possess a distinctive transcription factor signature: T-bet for T_H1, GATA-3 for T_H2, RORγt for T_H17, FOXP3 for iT_{REG}, and BCL-6 for T_{FH} (O'Shea and Paul, 2010).

This year marks the 50th anniversary of the recognition by Max Cooper, Raymond Peterson, and Robert Good that B lymphocytes constitute a cell lineage distinct from T cells (Cooper et al., 1965). Naturally, their seminal contribution was made possible in part by preceding discoveries, including findings that antibodies mediate humoral immunity, that antibodies are γ globulins (by the Department's Elvin Kabat, then a postdoctoral fellow in Sweden), that plasma cells make antibodies, and that bursectomy compromises bird antibody production (Behring and Kitasato, 1890, Tiselius and Kabat, 1939, Fagraeus, 1948, Cooper, 2015). Since then, much more has been learned about B cells. These cells owe their name to the Bursa of Fabricius, a lymphoid organ where they develop in birds. In adult mice or humans, there are two broad subsets of B cells, B1 and B2, with B2 accounting for the vast majority of B lymphocytes (Dorshkind and Montecino-Rodriguez, 2007, Monroe and Dorshkind, 2007). B1 cells, which make up about 1 out of every 20 B cells can be found in the spleen, the intestine,

and serous cavities (pleural and peritoneal cavities). Although they can be generated from the bone marrow, they are most efficiently derived from embryogenesis-stage progenitors and are thought of as innate immunity effector cells, since they respond to T cell-independent antigens (e.g. microbial carbohydrates), with B1-a cells spontaneously secreting IgM (for instance against encapsulated bacteria) and B1-b cells producing antibodies only after induction (Kroese et al., 1992, Kantor and Herzenberg, 1993, Haas et al., 2005, Montecino-Rodriguez et al., 2006, Dorshkind and Montecino-Rodriguez, 2007, Monroe and Dorshkind, 2007).

B2 cells develop in the bone marrow. A number of transcription factors are essential for proper B cell development. Chief among them is paired box protein 5 (PAX5), which simultaneously helps activate B cell-lineage specific genes and contributes to the repression of genes that would favor other differentiation programs (Busslinger, 2004, Cobaleda et al., 2007). Additional players that are important include Ikaros, PU.1, E2A, EBF, and BCL-11a (Orkin and Zon, 2008). The AGM-originating HSC-derived precursor CLP first appears in the fetal liver and then seeds the bone marrow (Müller et al., 1994, Medvinsky and Dzierzak, 1996, Pieper et al., 2013). At this stage, the progenitor B cell, or pro-B cell, will start undergoing recombination-activating gene (*RAG*) 1 and 2 product-mediated scrambling of its immunoglobulin heavy and light chain loci. First, a diversity (D) and a joining (J) segments from the heavy chain locus will recombine to form a DJ segment, which will then be recombined to a cis-located variable (V) segment in early pre-B stage (Tonegawa, 1987). Functional, rearranged H chains can then be expressed into intracellular pre-BCR (B cell receptor) if their expression product can pair with a surrogate light chain composed of V-preB and λ-like proteins. *RAG* gene expression, which had transiently gone down is upregulated once again, with the beneficial effect of catalyzing light

chain locus (κ or λ) V and J rearrangement. Once a successful light chain locus VJ rearrangement has taken place, its light chain product replaces the surrogate light chain, resulting in the assembly of an immunoglobulin M (IgM) protein, which can be expressed on the B lymphocyte cell surface. These IgM-expressing, immature B cells leave the bone marrow, reach the secondary lymphoid tissues, and produce cell-surface IgD BCRs by alternative splicing of a large mRNA that can also translate into IgM BCR, making them mature B cells (Flajnik, 2002). Mature B cells can then further differentiate into marginal zone B cells and follicular B cells, notably through the germinal center reaction (Pieper et al., 2013). The IFNβ-promoter transcriptional repressor B lymphocyte-induced maturation protein-1 (BLIMP-1) has been found — thanks in part to contributions from the Department's Kathryn Calame — to play an essential role in the terminal differentiation steps of the B cell lineage into immunoglobulin-secreting cells (Martins and Calame, 2008).

Finally, older and recent findings have shed light on the suppressive capacity of B lymphocytes, particularly with the description of regulatory B cells (B_{REG} cells) (Katz et al., 1974, Neta and Salvin, 1974, Wolf et al., 1996, Fillatreau et al., 2002, Mizoguchi et al., 2002, Mauri et al., 2003, Mauri and Bosma, 2012). B_{REG} cells, by secreting IL-10, TGF-β, and IL-35, mediate the immunosuppression of pro-inflammatory actors (e.g., TNFα-producing monocytes, IL-12-producing dendritic cells, pathogenic T helper cells), the induction of immunosuppressive T cells, and the maintenance of invariant natural killer cells (Rosser and Mauri, 2015). Although a B_{REG} cell precursor has so far not been identified, it is believed that various B_{REG} cell subsets can be induced in different immuno-inflammatory environments (Rosser and Mauri, 2015).

I.2.2. Functions and Interactions

I.2.2.1. Innate Role

More than being a mere simpler, coarser, more ancient alternative to the more sophisticated adaptive response, innate immunity plays a complex and central role in shaping the way the organism counters antigen intrusion. Charles Janeway put this concept forth almost 30 years ago; it stipulates that the recognition of well conserved viral, bacterial, and fungal pathogen-associated molecular patterns (PAMPs) by pattern-recognition receptors (PRRs) belonging to innate immunity cells triggers an inflammatory response and innate response and activates the adaptive response (Janeway, 1989, Iwasaki and Medzhitov, 2015). PAMPs are found on the cell walls of bacteria or fungi, for instance, and in viral nucleic acid. PRRs are principally made up of 3 families: the Toll-like receptors (TLRs), the nucleotide-binding oligomerization domain (NOD)-, leucine-rich repeat–containing receptors (NLRs), and the retinoid acid-inducible gene 1- (RIG-I-) like receptors (RLRs). Additional PRRs include C-type lectin receptors (CLRs), absent in melanoma 2- (AIM 2-) like receptors, and intracellular nucleic acid sensors, such as oligoadenylate synthase (OAS) proteins and their homologue cyclic GMP–AMP synthase (cGAS) (Janeway, 1989, Kawai and Akira, 2010, Rathinam et al., 2012, Wu and Chen, 2014, Iwasaki and Medzhitov, 2015). The different receptor classes have predilections for different PAMPs, help determine the kind of pathogen that the organism is facing, and therefore help steer lymphocytic reaction towards the most apposite response (Figure 1.3).

Pathogen detection can happen through two broad avenues: it can be structure-based or function-based. Structural features allow the organism to sense the presence of antigen in two ways: one is the more intuitive pattern recognition, the other is the less obvious missing self

recognition. Pattern recognition is an old strategy that can be divided into a cell-extrinsic branch, and a cell-intrinsic branch, although most PAMPs can be detected by both cell-extrinsic and cell-intrinsic branches. Cell-extrinsic pattern recognition is mostly facilitated by the PRRs of macrophages and dendritic cells: TLR1, TLR2, TLR4, TLR5, TLR6, CLRs (such as mincle, and dectins 1 and 2) have evolved towards the detection of β-glucans, flagellin, glycolipids, lipopeptides, lipopolysaccharides, and lipoteichoic acids, which are all conserved bacterial and fungal cell wall components. TLR3, TLR7, TLR8, and TLR9, on the other hand have evolved towards the detection of viruses (Iwasaki and Medzhitov, 2015). Unlike the bacterial and fungal wall component detectors, which are expressed on the cell surface, these TLRs are expressed in endosomes.

The cell-intrinsic branch of pattern recognition for viral pathogens is mostly mediated via RLRs (for viral RNAs) and cytosolic DNA sensors (for viral DNA) (Rathinam and Fitzgerald, 2011, Gurtler and Bowie, 2013, Wu and Chen, 2014, Iwasaki and Medzhitov, 2015, Yoneyama et al., 2015). Cell-intrinsic pattern recognition bacterial detection is thought to be more dependent on the invasiveness, viability, and replication of the pathogen (Vance et al., 2009, Sander et al., 2011, Blander and Sander, 2012, Stuart et al., 2013, Iwasaki and Medzhitov, 2015). This branch includes nucleic acid sensors like cGAS, PRRs such as NOD1 and NOD2, and NLRs that detect cytosolic PAMPs; interestingly, the inflammasome complex may be induced by some virus- and bacteria-sensing NLRs, leading to inflammatory cytokine production of the interleukin 1 (IL-1) family (Lamkanfi and Dixit, 2012, Burdette and Vance, 2013, von Moltke et al., 2013, Iwasaki and Medzhitov, 2015).

In addition to cell-extrinsic and cell-intrinsic pattern recognition, the organism can sense pathogen through another structural features-based approach: missing self recognition (Kärre et al., 1986). This is an evolutionary ancient method of distinguishing self and non-self and is used by natural killer (NK) cells, macrophages, and the complement system (Iwasaki and Medzhitov, 2015). NK cells, for instance, rely on the sensing of cell-surface class I major histocompatibility complex (MHC-I) by their immunoreceptor tyrosine-based inhibitory motif- (ITIM-) containing inhibitory receptors to recognize and not kill non-stressed or virally non-infected cells; but after some viral infections, oncogenic changes, or other forms of stress, MHC-I molecules may be downregulated, precipitating their lysis by NK cells (Raulet, 2006).

Beyond pathogen recognition based on structural features, the innate immune system can also sense pathogen presence indirectly through the detection of changes in function that imply the presence of a pathogenic intruder, like cysteine protease excretion or basement membrane disruption (Chovatiya and Medzhitov, 2014, Iwasaki and Medzhitov, 2015). There are a few advantages to recognizing functional features, as opposed to structural ones (Iwasaki and Medzhitov, 2015). One is that it allows the detection of a wide spectrum of pathogens based on the common features of their various infectious processes, instead of necessitating a specific receptor for each. Another advantage is that it can complement the information garnered from structural features-based recognition to discriminate between pathogens and non-pathogens; this would be beneficial, for instance, in the choice of appropriate responses to commensal versus pathogenic bacteria, given that both can be sensed by structural features-based recognition. One third advantage is that functional features-based recognition may provide a fail-safe mechanism, for instance for organisms that lack the stereotyped components (bacterial wall fragments, etc.)

recognized by TLRs, CLRs, and other PRRs. This might be the case for multicellular parasites with no invariant PRR targets (Iwasaki and Medzhitov, 2015). One illustrative mediator of functional features-based pathogen detection is the Nod-like receptor family, pyrin domain containing 3 (NLRP3) inflammasome, which can be activated by both organelle ion channel-forming viroporins and bacterial pore-forming exotoxins (Mariathasan et al., 2006, Martinon et al., 2009, Ichinohe et al., 2010).

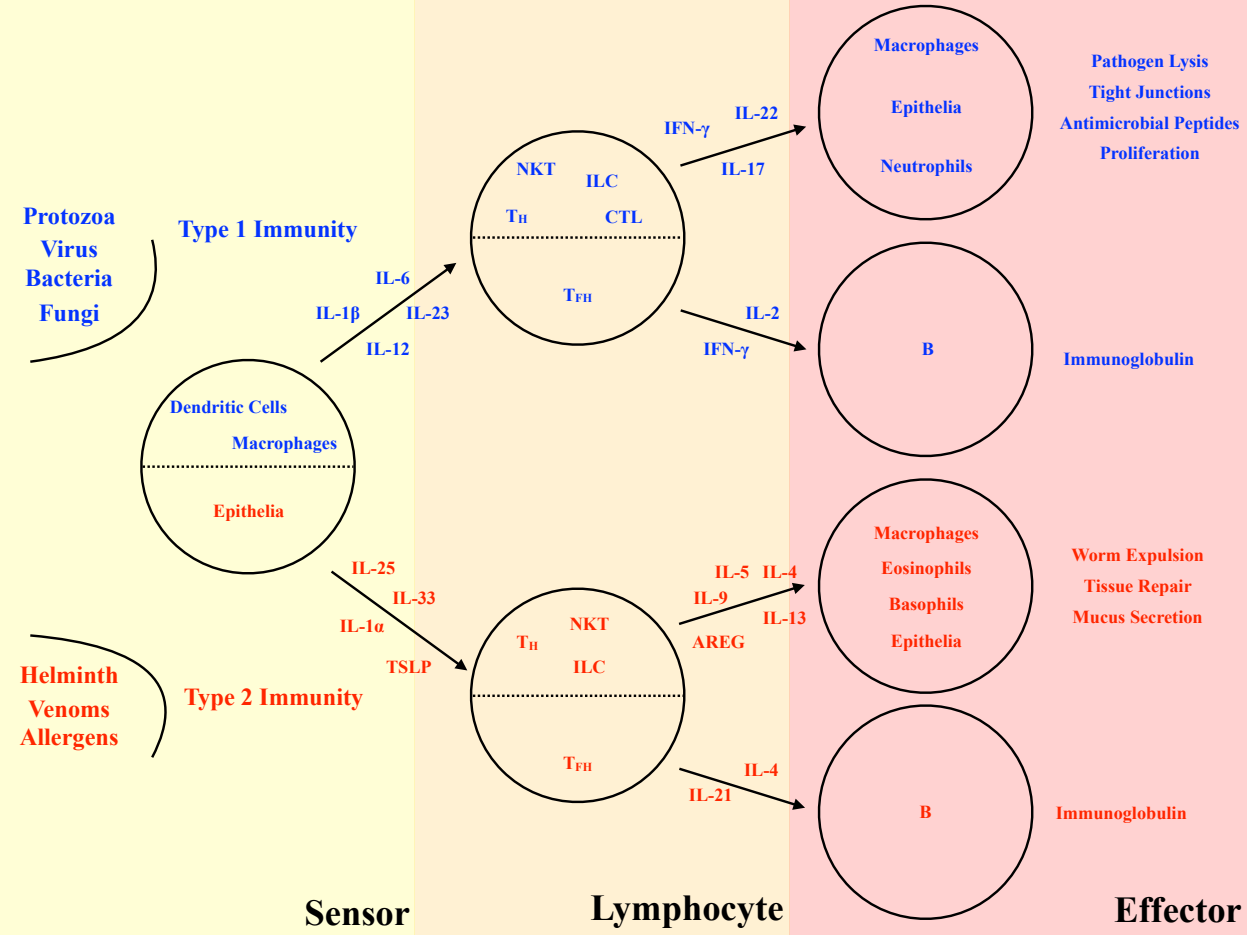


Figure 1.3. Integration of innate & adaptive immune systems. Illustrated here is a two-wave response mounted against a pathogen. In both type 1 immunity and type 2 immunity, lymphocytes bridge the sensing by dendritic cells/macrophages and epithelia at one end, and the effector cells at the other end of the response. High affinity, highly effective immunoglobulin production by B cells relies on interaction with T_{FH} cells in secondary lymphoid tissues.

The cells bearing these sensors of foreign material are strategically located around the organism; their activation will often notify the immune system of the kind and degree of threat that is posed by the pathogen, and will result in the mounting of an appropriate response (Iwasaki and Medzhitov, 2015). One such strategic location is the mucosal epithelium, or more precisely its luminal interface, as is the case for supra-epithelial phagocytes, such as luminal alveolar macrophages in the lungs, and luminal dendritic cell extensions in the lungs and intestines (Rescigno et al., 2001). Another location is the epithelium itself. The epithelial barrier, even when it is not the final pathogenic destination, is often breached by an invading pathogen, and can send valuable information to the immune system and initiate an immune response to ready the host for the possibility of viral dissemination (Saenz et al., 2008, Abreu, 2010, Hermesh et al., 2010, Iwasaki and Medzhitov, 2015). The third strategic location is the lamina propria, where dendritic cells, macrophages, and mast cells can also play a sentinel role, by detecting PAMPs and secreting cytokines to summon other leukocytes: basophils, eosinophils, neutrophils, and monocytes (Uematsu et al., 2006, Iwasaki and Medzhitov, 2015). Finally, pathogenic presence in the blood itself (the fourth location) is promptly detected by monocytes and neutrophils, and results in a full-on response and the stemming of pathogen spread by various ways, including enhanced monocyte phagocytosis, eosinophilic granule release, respiratory burst, bacterial entrapment by DNA extrusion, acute phase response protein production, and plasmacytoid dendritic cell- (pDC-) mediated systemic type 1 interferon response (Asselin-Paturel et al., 2001, Barchet et al., 2002, Brinkmann et al., 2004, Asselin-Paturel et al., 2005, Pecchi et al., 2009, Reizis et al., 2011, Thomas and Schroder, 2013, Iwasaki and Medzhitov, 2015).

Among the different PRR-bearing cells, DCs play an important, coordinating role by activating the adaptive immune system. Different types of tissue-resident DCs have been described: stratified squamous epithelium-resident $CD207^+$ Langerhans cells, BATF3-dependent $CD103^+CD207^+CD11b^-$ and IRF4-dependent $CD301b^+CD207^-CD11b^+$ dermis- and submucosa-resident DCs, and intestine- and lung-resident $CD103^+CD11b^+$ DCs. Tissue-resident DCs migrate to a secondary lymphoid tissue (draining lymph nodes, spleen, and mucosa-associated lymphoid tissues), where they, along with lymphoid organ-resident $CD8\alpha^+$ and $CD11b^+$ DCs, will help prime and shape adaptive immune system responses (Satpathy et al., 2012, Ganguly et al., 2013, Merad et al., 2013, Persson et al., 2013, Schlitzer et al., 2013, Briseno et al., 2014, Iwasaki and Medzhitov, 2015).

Some innate responses will elicit the release of a wave of cytokines with a direct effector result; such is the case for pDCs and type I interferons, and mast cells and histamine, prostaglandins, and proteases (Iwasaki and Medzhitov, 2015). Oftentimes, however, the different innate immunity pathogen sensing modalities are integrated in complex and elegant fashion (Figure 1.3). PRR-mediated responses by DCs and macrophages and PRR-less epithelial cell-mediated responses can trigger the release of a first wave of cytokines, which cause various lymphocyte subsets to release a second wave of cytokines targeting effector cells such as neutrophils and B lymphocytes (Iwasaki and Medzhitov, 2015). It is at the level of this integration of information that the germinal center reaction occurs.

I.2.2.2. The Germinal Center

Innate support for the specific response of the adaptive immune system culminates when antigen is brought from the periphery hematogenously or lymphogenously to secondary

lymphoid tissues to be exposed to B cells residing in primary follicles, in a process that can be mediated by Follicular Dendritic Cells (FDCs) and that can ultimately result in the production of high-affinity antibodies against the offending antigen (Figures 1.4 and 1.5).

The production of specific, high-affinity antibodies has long been enigmatic in immunobiology. The theory of clonal selection and the description of lymphocyte receptor gene rearrangement allowed for a deeper understanding of the diversification of the immune repertoire (Burnet, 1976, Hozumi and Tonegawa, 1976, Bernard et al., 1978, Early et al., 1980). The discovery that immunoglobulin affinity for antigen increases with time and that V(D)J recombination is supplemented by additional somatic alterations to the immunoglobulin genes gave currency to the theory of affinity maturation (Jerne, 1951, Eisen and Siskind, 1964, Weigert

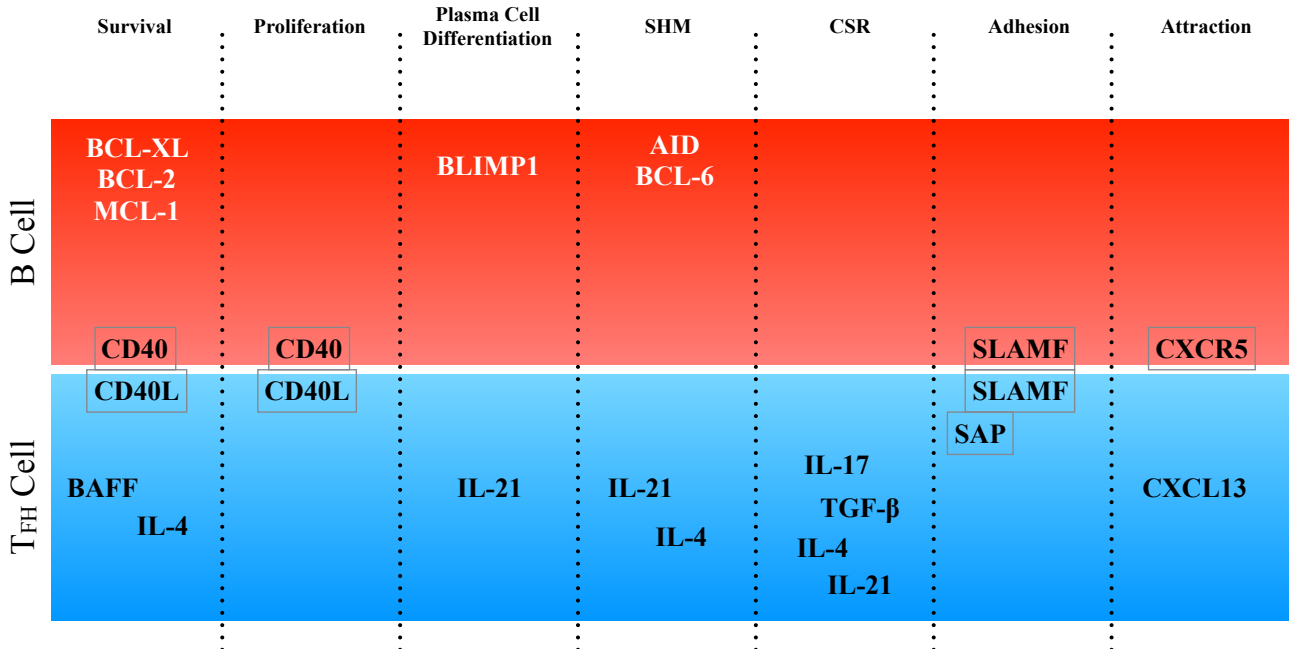


Figure 1.4. Interactions between B and T_{FH} cells. In the course of their encounter during the germinal center reactions, several signals are mediated: survival, proliferation, differentiation, antibody diversification, adhesion, attraction. Shown in white are the genes upregulated in B cells during the interaction. Shown in boxes are the cell-surface proteins involved in different signals. Finally, also shown are different cytokines produced by T_{FH} lymphocytes in their interactions with B cells.

et al., 1970, McKean et al., 1984, Berek and Milstein, 1987, Victora and Nussenzweig, 2012). Affinity maturation is now understood to be the outcome of an iterative process, whereby sequential rounds of competitive and fine-tuning interactions between B and T lymphocytes result in the selection and proliferation of B cells producing immunoglobulins with high antigen specificity (Figure 1.4). These reactions take place in a micro-anatomical and functional structure first described over 130 years ago by Walther Flemming: the germinal center (Nieuwenhuis and Opstelten, 1984, Victora and Nussenzweig, 2012, De Silva and Klein, 2015).

The first 7 days of the germinal center reaction to a potentially dangerous antigen are divided into three stages. The early initiation stage, lasting about two days post immunization, starts when naive B cells see antigen and become activated. They leave the B cell-rich and FDC-containing primary follicle and migrate to the interfollicular region, where they can come into MHC/TCR- and CD40/CD40L-mediated contact with activated T lymphocytes that have similarly left the T cell-populated T cell zone after encountering cognate antigen (Kerfoot et al., 2011, De Silva and Klein, 2015). This first, long-lived encounter allows for the T cell to acquire a T follicular helper cell (T_{FH}) profile and for the B cell to become more fully activated. From day 3 to day 4 after immunization, the immune reaction will be in its late initiation phase. During this period, T_{FH} cells, and later, B cells, will leave the interfollicular region and will penetrate the follicle, where, in the presence of FDCs, B cells will start to proliferate, creating an early germinal center at the center of the follicle. As a result, the non-activated, resident B cells are displaced into forming a mantle zone around the early germinal center. This germinal center-containing follicle is also referred to as the secondary follicle. The third and last stage is marked by the massive proliferation of germinal center B cells and the establishment of the mature

germinal center. This period also sees the appearance of a dark zone and a light zone. The dark zone contains a network of reticular cells similar to the FDCs, within which B cells continue to vigorously proliferate; it is so-named due to its histological features, in contrast to the less densely populated light zone, where B cells, FDCs, T_{FH} cells, and macrophages are encountered (MacLennan, 1994, Victora and Nussenzweig, 2012, Bannard et al., 2013, De Silva and Klein, 2015).

B cell centroblasts residing in the dark zone were once believed to move to the light zone to differentiate into centrocytes, before becoming plasma cells and exiting the germinal center (MacLennan, 1994). Mounting evidence now suggests, however, that the intragerminal movements of B cells between the two zones may be slightly more dynamic, in that B cells can re-enter the dark zone after exiting it once to go into the light zone (Figure 1.5) (Allen et al., 2007, Hauser et al., 2007, Schwickert et al., 2007, Victora et al., 2010). B cells that have exited the dark zone can indeed access the light zone and be selected, if post-V(D)J genomic alterations to the immunoglobulin receptor variable region-encoding loci has resulted in higher binding affinity to the antigen. In the light zone environment comprised notably of T_{FH} cells and FDCs, B cells with higher B cell receptor-binding ability to antigen internalize the latter with higher efficiency and present it more readily on their surface Major Histocompatibility Complexes (MHC). As a consequence, neighboring T_{FH} cells activated against the same antigen interact through their own T cell receptors with the higher-affinity B cells and select these cells for proliferation and differentiation. Some of these selected light zone B cells, however, instead of differentiating into memory or plasma cells, will recirculate into the dark zone, where additional rounds of alterations to the loci that encode the variable regions of their immunoglobulin

receptors can occur (Allen et al., 2007, Hauser et al., 2007, Schwickert et al., 2007, Victora et al., 2010). These recirculating cells will therefore go through an additional round of proliferation and somatic mutation before re-entering the light zone, where they can get further selected. This iterative process ensures the preponderance of plasma and memory B cells producing antibodies with the most enhanced binding to the immunizing antigen. This newer account of secondary follicle dynamics highlights two important, novel aspects of germinal center B cells. One is the

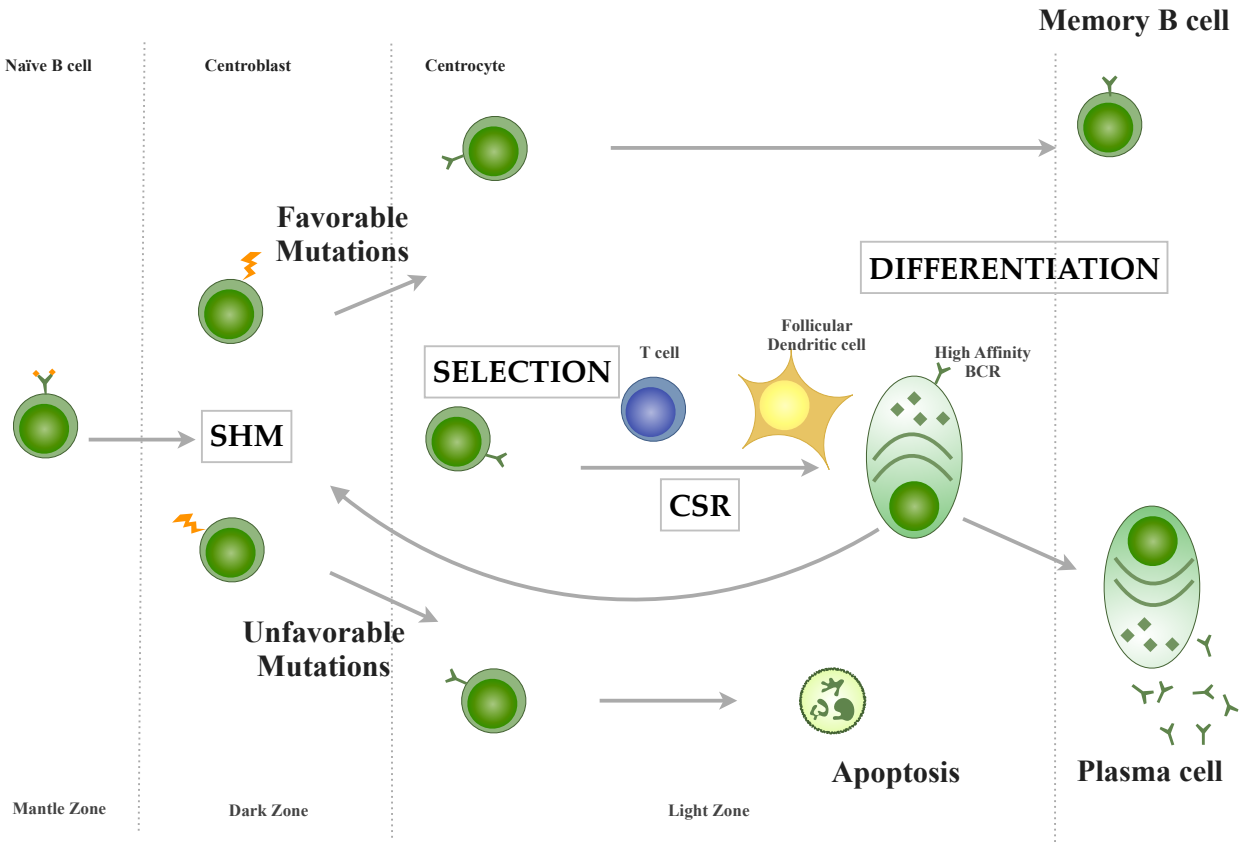


Figure 1.5. The germinal center reaction. Naive B cells migrate through different secondary lymphoid tissue compartments in the course of the germinal center reaction. Illustrated here is the itinerary of a B cell as it is being stimulated in the spleen to undergo somatic hypermutation and class-switch recombination. Several outcomes are possible for a naive B cell going though the germinal center reaction: it can undergo apoptosis, or differentiate into a plasma cell, or a long-lived memory B cell, readying the organism for subsequent encounters with the same antigen.

concept of recirculation, which helps explain the efficiency of the process that leads to highly specific antibody production. The second is the realization that the major force behind the B cell germinal center reaction (i.e., B cell selection and affinity maturation) is competition for T_{FH} cell help before and during the germinal center reaction, as opposed to direct competition for antigen (Victora et al., 2010, Victora and Nussenzweig, 2012).

The interzonal dynamics of the germinal center reaction are accompanied by a variety of changes at the molecular level in the different stages. During the early initiation phase, antigen activation coincides with, or precedes, the activity of several key factors, among which B cell lymphoma 6 (BCL-6), Myocyte-specific enhancer factor 2B (MEF2B), Myeloid cell leukemia 1 (MCL1), Interferon-regulatory factor 4 (IRF4), and Myelocytomatosis (MYC). BCL-6 in particular is crucial at this stage, since in its absence, or that of its repression domain 2 (RD2), pre-germinal center B cells no longer migrate into the follicle, as they fail to upregulate the dark zone B cell-expressed chemokine receptor CXCR4, and fail to downregulate the mediator of follicle exit sphingosine-1-phosphate receptor type 1 (S1PR1) (Allen et al., 2004, Caron et al., 2009, Kitano et al., 2011, Huang et al., 2014, De Silva and Klein, 2015). The cross-talk between B cells and T cells would not be spared in the absence of BCL-6 either, since its levels in precursor B cells correlate with the integrin-mediated interaction of the two cell types (Cannons et al., 2010, Kitano et al., 2011). The role of MEF2B seems to be intimately linked to that of BCL-6, as it seems that MEF2B induces *BCL6* expression (Ying et al., 2013). MCL1 is an anti-apoptotic protein whose deletion impairs germinal center formation (Vikstrom et al., 2010). The role of IRF4 has not cogently been characterized yet. Peculiarly, evidence indicates that IRF4 seems to both activate and repress *BCL6* transcription (Saito et al., 2007, Ochiai et al., 2013).

Finally, MYC function is equally complex and not entirely defined; however, MYC does seem to be important, since its deletion in activated mouse B cells hinders germinal center formation (Calado et al., 2012).

Later in the initiation phase, while IRF4 expression decreases, MYC, MEF2B, BCL-6, and MCL1 are still expressed, and remain so for the rest of the germinal center reaction, except for MYC, which ceases to be expressed during the proliferation and establishment (De Silva and Klein, 2015). As the B cell differentiates into a quiescent, long-lived plasma cell or memory B cell, different factors are again involved and play important roles, namely BLIMP1, IRF4 and XBP1 for plasma cells, and BACH2, PAX5, and PU.1-IRF8 for memory B cells (Nutt et al., 2015).

Much remain to be untangled to gain a fuller appreciation of the regulation of the germinal center reaction dynamics at a molecular level, though a general picture continues to emerge. A considerable amount of knowledge has been accumulated, however, about the genomic changes that take place during the germinal center reaction. The alterations of the heavy and light chain immunoglobulin loci are at the very center of the adaptive immune system, and also shed some light on B cell-origin pathologies.

I.3. B-Cell Specific Lymphoid Immune System Reactions

I.3.1. Somatic Hypermutation

When B cells emerge from the bone marrow as naive B lymphocytes, they have already undergone substantial alterations to the genes that encode the B cell receptor. Indeed, it is estimated that at the conclusion of RAG-mediated V(D)J combinatorial recombination, B cells have so diversified their immunoglobulin repertoire that they can now recognize an astronomical 5×10^{13} different molecules — more than there are stars in the Milky Way (Pieper et al., 2013). Yet this primary repertoire only represents an anlage, the basis for further diversification at the behest of activation-induced cytidine deaminase (AID), the mutator enzyme of B cells, through somatic hypermutation (SHM). The increased Ig diversity will allow for selection of the B cells with the best suited set of alterations for antigen-binding and T-cell interaction. In a sense, SHM can be understood as real-time evolution, the molecular facilitator of affinity maturation.

Among the different genomic alterations described in B lymphocytes, SHM is thought to have appeared with the first jawed vertebrates, before isotype switching and gene conversion (Figure 1.1) (Kaufman, 2010). The concept of SHM was introduced about 45 years ago, when it was argued that the most parsimonious explanation to account for the pattern of variability in mouse light chains was the "generat[ion] by somatic spontaneous mutation and by sequential selection by antigen of single step mutants" (Weigert et al., 1970). Support for this proposal soon followed when it was shown through sequencing that mutations in mouse myeloma immunoglobulin genes were confined to the Ig variable region, though the constant region could be SHM-targeted as well if Ig constant region transcription was induced (Kim et al., 1981, Selsing and Storb, 1981, Peters and Storb, 1996, Storb, 2014). The realization that transcription

plays an essential role in SHM led to a model linking somatic mutation to transcription-coupled DNA repair (Peters and Storb, 1996). In this insightful model, mutations are catalyzed by a mutator factor coupled to the transcription machinery from initiation at the promoter region and causing mutations during polymerase pausing. That mutator factor was later identified as AID (Muramatsu et al., 1999).

The mutations of SHM are mostly single base pair substitutions (transitions more so than transversions) — though insertions and deletions are also observed — with rates of 10^{-5} to 10^{-3} mutations per base pair per generation, or approximately 10^{6}-fold higher than elsewhere in the genome (Rajewsky et al., 1987, Longerich et al., 2006). Analyses in the immunoglobulin heavy chain locus variable region indicate that the mutations start appearing about 150 to 200 bp downstream of the transcription start site and can still be detected for another 1.5 kilobases, clustering at WRC/GYW hotspots (where W = dA or dT, R = dA or dG, Y = dC or dT), but without extending into the constant region (Li et al., 2004).

I.3.2. Immunoglobulin Isotype Switching

Isotype switching is the replacement of one antibody heavy chain isotype by another, along with the attendant changes in effector capabilities of a B cell in a mechanism termed class-switch recombination (CSR). CSR is thought to have first appeared with amphibians and land animals after SHM had already evolved (Kaufman, 2010). Isotype switching from IgM to IgG was first described in birds (Kincade et al., 1970). However, the original version of today's prevailing CSR model was proposed about 8 years later (Honjo and Kataoka, 1978).

CSR is a recombination process that takes place in proliferating B cells, usually during the protein antigen-triggered, T cell- and dendritic cell-assisted germinal center reaction, although T cell-independent recombination also occurs (Figure 1.6) (Hodgkin et al., 1996, Rush et al., 2005, Bergqvist et al., 2006). The heavy chain immunoglobulin (IgH) locus is composed of several constant (C_x) regions, each preceded by 1-12 kb-long, G:C oligomer-rich switch sequences (S_x, except for C_δ, which has a rudimentary S_δ-like , σ_δ recently shown to undergo CSR as well), and a promoter exon I_x (Saintamand et al., 2015). Concomitant with cytokine-inducible sterile transcription from I_x promoters, CSR-initiating AID deamination activity at the donor, upstream S_μ switch sequence, and at a 65-160 kb downstream, acceptor switch sequence may result in the presence of contemporaneous double-stranded DNA breaks (DSBs) at both sites of mutagenic activity that are repaired mostly through canonical non-homologous end joining (NHEJ) of the donor and acceptor switch sequences, as the DNA sequence intervening between the two DSBs is looped out and excised. Rarely, a switch region other than S_μ will recombine with another, further downstream acceptor sequence. NHEJ is favored in part because of the low homologies between the different switch sequences and because it is a predominant repair pathway during the G_1 phase, which is when donor and acceptor DSBs appear (Dunnick et al., 1993, Petersen et al., 2001, Schrader et al., 2007, Sharbeen et al., 2012, Khair et al., 2014). The less-well defined Ku-independent, microhomology-directed, alternative non homologous end-joining has recently been shown to be mediated by the DNA end-processing factor CtBP-interacting protein (CtIP), while Mre11, Rad50, Sae2 have also been implicated in yeast studies (Yan et al., 2007, Mimitou and Symington, 2010, Boboila et al., 2012, Bunting and Nussenzweig, 2013).

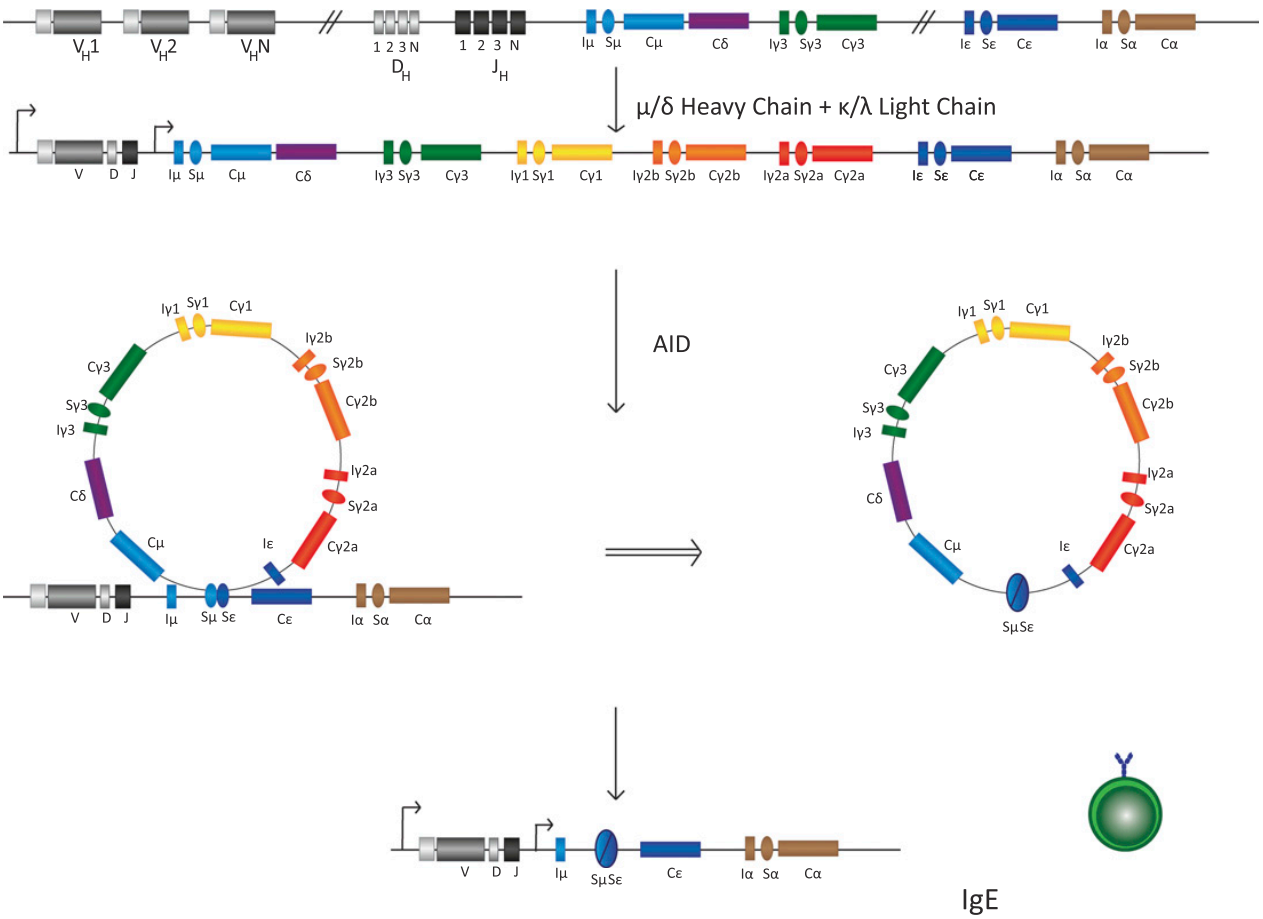


Figure 1.6. Class-switch recombination. The immunoglobulin heavy- and light-chain loci undergo RAG1/2-mediated V(D)J recombination as B and T cells develop in the bone marrow. Mature B cells are capable of an additional deletion and end-joining process that results in their producing differently enabled immunoglobulin proteins. Shown here is isotype switching to IgE, an important effector against helminthic infections. IgA is secreted as a dimer at mucosal sites (e.g., gastro-intestinal tract); IgG, found in the blood, is a potent opsonizer.

The different immunoglobulin isotypes feature well-characterized properties that confer specific effector functions to the B cell that produces them. For instance, IgM can be produced in pentameric form and is a strong complement system activator; IgA, secretable as a dimer across mucosal barriers, is crucial in keeping at bay pathogenic antigen; IgG is found in the blood (in subsets named for their relative abundance 1 through 4) and plays an essential role in opsonization and gestational and perinatal immunity; IgE, finally, is involved in granulocyte sensitization and protection against helminths. Depending on the port of entry, pathogenicity, and nature of the instigating antigenic insult, an assortment of cellular mediators is involved and a cocktail of cytokines is produced by epithelial and innate immune cells, influencing the germinal center milieu and shaping isotype choice in CSR. The B lymphocyte's CD40, Toll-like receptor (TLR), and immunoreceptor tyrosine-based activation motif (ITAM)-containing protein complex CD79-associated B-cell receptor pathways integrate primary stimuli from T cells (e.g., CD40-ligand CD154 on T_{FH} cells) or antigen, resulting in the activation of nuclear factor-κB (NF-κB) through both its canonical and non-canonical modalities (Xu et al., 2012). NF-κB induces the expression of homeobox protein C4 (*HOXC4*); HOXC4 and NF-κB then induce the expression of the master mutator AID-encoding *AICDA*. In parallel, secondary signals carried by cytokines such as interleukin 4 (IL-4), transforming growth factor β (TGF-β), and interferon γ (IFNγ), up-regulate transcription at the appropriate I_x-S_x-C_x transcriptional unit (Xu et al., 2012). Finally, and interestingly, chromosome capture techniques have been employed to show that chromosomal looping, involving notably the IgH 3' regulatory region and the intron enhancer E_μ,

plays a role in isotype selection (Wuerffel et al., 2007, Sellars et al., 2009, Chatterjee et al., 2011, Kenter et al., 2012, Stavnezer and Schrader, 2014).

I.3.3. Gene Conversion

In some animals, most notoriously the chicken but also the rabbit, the product of V(D)J recombination is further altered by a process known as somatic gene conversion (Dildrop et al., 1982, Kleinfield et al., 1986, Reth et al., 1986, Reynaud et al., 1987, Reynaud et al., 1989). Upstream of functional variable light and heavy chain immunoglobulin loci in the chicken lie clusters of nonfunctional families of pseudogenes, termed ΨV_L and ΨV_H (Ratcliffe, 2006). Though lacking the signal sequences that would facilitate their participation in V(D)J recombination, these pseudogene V elements can nonetheless be used by B lymphocytes to increase the post-V(D)J diversity of their immunoglobulin receptors, as stretches from these pseudogenes get copied and pasted into the functional V(D)J exon, in a B cell deaminator AID-dependent diversification reaction, much like SHM and CSR (Arakawa et al., 2002, Fugmann and Schatz, 2002). Gene conversion represents another vertebrate mechanism to facilitate adaptive immunity by increasing B cell receptor diversity. AID requirement in this process might inform current and future investigations of the biology of its involvement in the adaptive immune system diversification schemes of murine and human B lymphocytes.

I.4. Pathophysiology of B-cell Diseases

I.4.1. Immunoglobulin Production-Related Immunodeficiencies

Given the prominent function occupied by B lymphocytes in adaptive immunity, it comes as no surprise that B cell disorders are likely to negatively impact the immune competency of an organism. In addition, processes taking place in other cell types, whose function is intimately related to antibody production by B cells, will have an effect on antibody synthesis. Primary immunoglobulin deficiencies can therefore have a B-cell intrinsic origin or a B-cell extrinsic origin.

B-cell originating immunoglobulin production defects can be classified into 6 broad categories: developmental defects, migration defects, survival defects, activation defects, cytokine responsiveness defects, and CSR defects. Inauspicious mutations in key proteins involved in B lymphocyte development have been associated with diseases such as agammaglobulonemias (many mutated proteins, among which: Bruton tyrosine kinase, BCR-associated proteins CD79a and CD79b, and B cell linker protein BLNK), severe combined immunodeficiency (SCID, including adenosine deaminase (ADA), for which the first successful human gene therapy was conducted, RAG, adenylate kinase 2, and NHEJ components), dyskeratosis congenita, also known as Hoyeraal-Hreidarsson syndrome (which involves sheltering complex proteins), and the newly described dendritic cell, monocyte, B and NK lymphoid (DCML) deficiency or MonoMAC (monocytopenia with Mycobacterium avium complex), associated with GATA-2 mutations (Bigley et al., 2011, Durandy et al., 2013). B cell migration defects may be due to mutations in DOCK8 or MST1 (leading to combined immunodeficiency), WIP or WASP (leading to Wiskott Aldrich syndrome), CXCR4 (leading to

WHIM syndrome, for warts, hypogammaglobulinemia, infections and myelokathexis) (Durandy et al., 2013). B cell survival defects have been associated with mutations in the pro-inflammatory cytokine TWEAK (TNF-like weak inducer of apoptosis) and tumor necrosis factor receptors BAFFR/TACI (B-cell activating factor receptor/Transmembrane activator and CAML interactor), leading to common variable immunodeficiency (Durandy et al., 2013). B cell activation defects are numerous, and run the gamut from ICF syndrome (immunodeficiency, centromeric instability, facial anomalies, associated with mutations in DNA (cytosine-5-)-methyltransferase 3β), to hypogammaglobulonemia related to CD19 or CD20 mutations, to combined immunodeficiency related to CARD11 mutations (caspase recruitment domain family, member 11). Two other defects in this category include selective polysaccharide antibody deficiency (extensively disparate in origin) and PLCγ2-associated antibody deficiency and immune dysregulation. Described cytokine responsiveness defects include hyper-IgE syndrome and severe combined immunodeficiency (both associated with mutations in signal transducer and activator of transcription STAT3 and Janus kinase JAK3), as well as combined immunodeficiency resulting from IL21R mutations (Durandy et al., 2013).

Class-switch recombination defects have been associated with Ataxia-Telangiectasia, Nijmegen-breakage syndrome, and severe combined immunodeficiency; the proteins involved in these defects include important mediators of CSR: AID, CD40, UNG, PMS2, ATM, MRE11, NBS1, RNF168 (for AT and NBS), and NHEJ components for SCID (Durandy et al., 2013). In parallel to these B-cell intrinsic causes are B-cell extrinsic origins of immunoglobulin production deficiencies. Here, too, the defects can be categorized in 4 broad classes: T-cell differentiation defects, T-cell function defects, T_{FH} defects, and innate immunity defects. In the first category are

found DiGeorge syndrome (caused in humans by a deletion in the long arm of chromosome 22) and combined immunodeficiency (in this case due to mutations in CD3 subunits or IL7Rα) (Durandy et al., 2013). Another form of combined immunodeficiency can be understood as deriving from T cell malfunction and is associated with mutations in the regulatory factors of MHC class II transcription, as well as the proteins ZAP70 (ζ-chain-associated protein kinase 70) , ITK (interleukin-2-inducible T-cell kinase), and the protein tyrosine kinase $p56^{Lck}$ (Durandy et al., 2013). Defective T_{FH} can be manifested by combined variable immunodeficiency (with mutations in ICOS, inducible T-cell costimulator), isotype switching defects (involving CD40 ligand CD154 mutations), hyper IgE syndrome (if STAT3 mutations occur), and X lymphoproliferative and related disease (concomitant with mutations in CD27 or SAP, SLAM (signaling lymphocytic activation molecule)-associated protein) (Durandy et al., 2013). Finally, signaling defects in the TLR/IL1R pathways, due to mutations in IRAK4 (interleukin-1 receptor-associated kinase 4), TLR and IL-1 adapter protein MyD88 (myeloid differentiation primary response 88), or TIRAP (toll-interleukin 1 receptor (TIR) domain containing adaptor protein), can induce innate immunity defects leading to B cell extrinsic primary antibody production deficiency (Durandy et al., 2013).

I.4.2. B Lymphocyte-related Immunoproliferative Disorders

Lymphocytes, B cells in particular, have a certain predilection for oncogenic transformation. It is the evolutionary cost paid for having a system capable of massive proliferation (e.g., in the germinal center) and directed mutagenesis (RAG-mediated recombination, but especially AID-initiated alterations). This unfortunate oncogenic trade-off can

therefore typically manifest itself at times that coincide with the different B-cell development stages and immunoglobulin diversification reactions.

During B-cell lymphopoiesis in the bone marrow, a translocation involving the breakpoint cluster region (*BCR*) and the Abelson tyrosine-protein kinase homologue 1 (*ABL1*) and producing the "Philadelphia Chromosome" in precursor B cells can result in B cell acute lymphocytic leukemias (B-ALLs); additional sources for these leukemias involve mutations or translocations targeting runt-related transcription factor 1 (*RUNX1*)-associated mutations, pre-B cell leukemia homeobox 1 (*PBX1*), paired box protein 5 (*PAX5*), mixed-lineage leukemia (*MLL*), early B cell factor 1 (*EBF1*), transcription factor *E2A*, protein tyrosine phosphatase non-receptor type 11 (*PTPN11*), and Rat sarcoma (*RAS*) genes (Rickert, 2013). These genes play physiologically important roles in B cell differentiation and pre-BCR signaling, a crucial component of the proliferation/survival signal received by these cells.

Later on in primary follicles of secondary lymphoid tissues, mature B cells can give rise to mantle cell lymphoma (MCL, associated with mutations in cyclin D1 *CCND1*) and the non-mutated form of B cell chronic lymphocytic leukemia (B-CLL), where the variable region has not undergone SHM (Seifert et al., 2012, Rickert, 2013).

In the marginal zone, mutations in NOTCH2 have been associated with splenic marginal zone lymphoma. Another lymphoma thought to originate from marginal zone B cells is mucosa-associated lymphoid tissue (MALT) lymphoma, and mutations in the protein B cell lymphoma 10 (*BCL10*) locus and mutations/translocations involving the protease mucosa-associated

lymphoid tissue lymphoma translocation protein 1 (*MALT1*) are thought to contribute to the development of this indolent lymphoma (Rickert, 2013).

B cells that undergo SHM and/or CSR in the germinal center secondary follicle can give rise to a number of neoplasms. Follicular lymphoma, diffuse large B-cell lymphoma, and Burkitt's lymphoma derive from germinal center B cells, while memory B cells are thought to be the origin of the immunoglobulin variable-region mutated version of B-CLL, and plasma cells may give rise to multiple myeloma. Multiple myeloma has been associated with mutations in the transcription factor musculoaponeurotic fibrosarcoma (*MAF*), fibroblast growth factor receptor 3 (*FGFR3*), interferon regulatory factor 4 (*IRF4*), and *CCND1* (Rickert, 2013). The high proliferation rate of germinal center B cells, coupled to the genomic alterations occurring during that developmental stage, promote a particularly mutagenic environment, one that is mirrored by the multiplicity of neoplasms that can originate here. Both Burkitt's lymphoma and follicular lymphoma have their origin in the dark zone of germinal centers. Translocations have been identified in variants of both lymphomas. In follicular lymphoma, a translocation thought to result from an error in RAG-mediated V(D)J recombination involves the immunoglobulin heavy chain locus (*IGH*) and the B cell lymphoma 2 (*BCL2*) protein locus. In Burkitt's lymphoma, it is thought that mistakes arising during AID-mediated immunoglobulin heavy and light chain loci remodeling cause translocations, most notably between *IGH* and *c-MYC*. These translocation events are thought to juxtapose promoters or enhancers with an oncogene, causing the dysregulated ectopic expression of the latter (Basso and Dalla-Favera, 2015). In addition, Burkitt's lymphoma is noteworthy for disruptions in phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K) signaling, the regulatory axis composed by basic helix–loop–helix transcription

factor E2A and its negative modulator inhibitor of DNA binding 3 (ID3), and the Gα13 pathway, which negative modulates of GC B cell migration and PI3K signaling (Basso and Dalla-Favera, 2015). In addition to the ectopic expression of BCL-2 resulting from the *IGH-BCL2* translocation of follicular lymphoma, the genetic inactivation of the histone methyltransferase MLL2 is thought to play a fundamental role in the pathogenesis of this lymphoma (Basso and Dalla-Favera, 2015). Diffuse large B cell lymphoma (DLBCL) is often described as coming in two subtypes: a germinal center B cell-like DLBCL (GCB-DLBCL) subtype, thought to originate from light zone B cells and an activated B cell-like DLBCL subtype (ABC-DLBCL), thought to originate from B cells later in the germinal center reaction (plamablasts). While the pathophysiology is not exactly ascertainable, certain unmistakeable features can be highlighted. Gain-of-function mutations in the enhancer of zest homologue 2 (EZH2) seen in GCB-DLBCL are believed to drive epigenetic changes favorable to the lymphoma; in addition, translocations resulting in the ectopic expression of such oncogenes as *BCL2* and *MYC* have also been noted. Moreover, Gα13 pathway inactivation is also observed in GCB-DLBCL (Basso and Dalla-Favera, 2015). For ABC-DLBCL, one hallmark feature is the constitutive activation of the NF-κB pathway; another feature is B lymphocyte-induced maturation protein 1 (BLIMP1)-encoding *PRDM1* inactivation. Both ABC- and GCB-DLBCL display genetic aberrations thought to contribute to their pathogenesis, including *BCL6* constitutive expression, inactivation of CREB-binding protein-encoding *CREBBP* or E1A-binding protein p300-encoding *EP300*, inactivation of *MLL2*, and immune escape as a result of mutations in the genes encoding β2 microglobulin, HLA-A, B, and C. In addition to the above alterations, changes in ploidy can occur and confer neoplastic potential to the different B cell developmental stages (Basso and Dalla-Favera, 2015).

I.5. Molecular Machinery of SHM and CSR

I.5.1. AID and the APOBECs

The B cell-specific genome mutator enzyme activation-induced cytidine deaminase (AID) is a member of the APOBEC (Apolipoprotein B mRNA editing enzyme, catalytic polypeptide) zinc-dependent cytidine deaminase family of proteins (Figure 1.7) (Conticello, 2008). The bona fide origin of the APOBECs is placed at about 500 million years ago, when cartilaginous fish (e.g. sharks) and tetrapods are estimated to have diverged (Saunders and

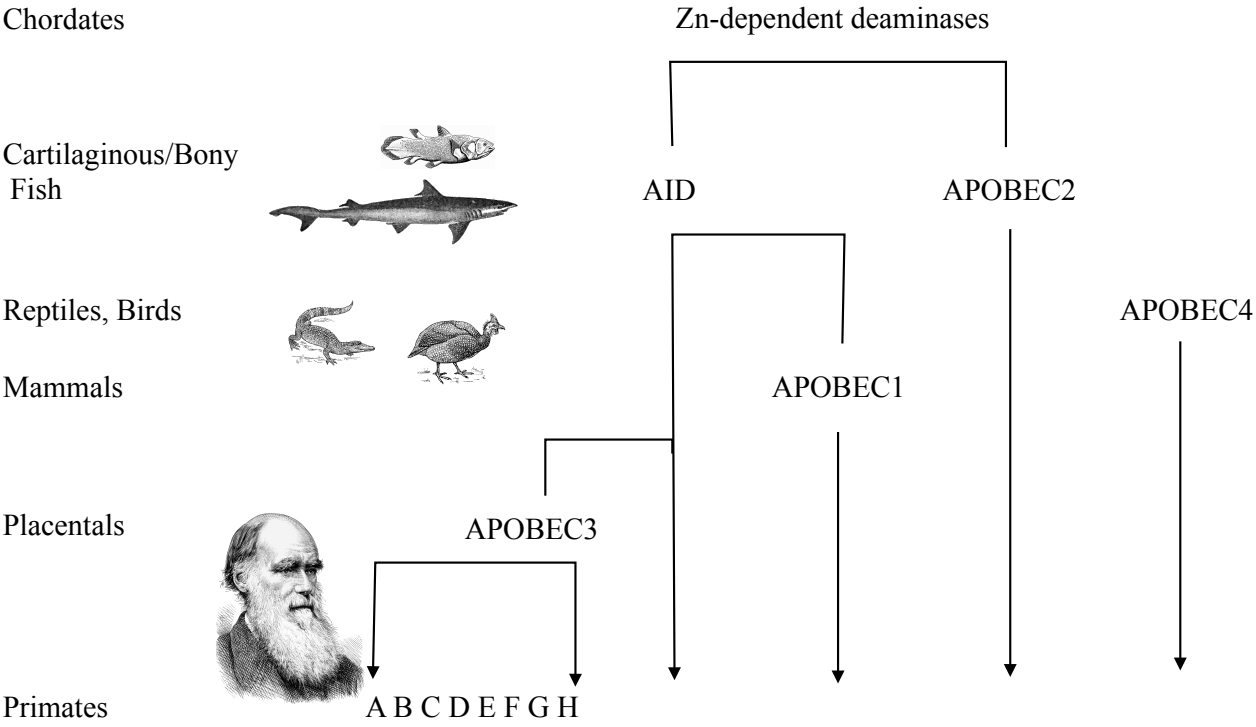


Figure 1.7. Evolution of APOBECs. While the evolution of APOBEC4 is not entirely understood, AID and APOBEC2 are the oldest APOBEC members, and appear as early as the first jawed vertebrates. APOBEC2 is found in muscle cells. APOBEC1, the apolipoprotein mRNA mutator behind the family name, appears to have evolved from AID around the time mammals appeared. The APOBEC3 subgroup is also thought tho have evolved from AID, though at a later time than APOBEC1, and are active in viral restriction.

Magor, 2004, Conticello et al., 2005, Zhao et al., 2005, Conticello et al., 2007)). Some APOBECs target RNA (e.g., APOBEC1), while others use DNA as their substrate (such as AID and the APOBEC3s). APOBEC1 plays an important role in lipid biology, as it edits the pre-mRNA of apolipoprotein B100, deaminating a cytidine into a uridine and introducing a UAA stop codon, thereby allowing for the production of the shorter apolipoprotein B48 (Teng et al., 1993). While APOBEC2 and APOBEC4 have been little characterized, members of the APOBEC3 genes (APOBEC3A to 3G) have been implicated in the control of primate retrovirus propagation (Jarmuz et al., 2002, Sheehy et al., 2002). Based on studies in yeast cytosine deaminase, a model for Zn-dependent deaminase catalytic activity was proposed, whereby the loss of the amino group is the result of a hydrolytic reaction by nucleophilic attack, in which participate the hydroxide of a Zn-activated water molecule and a glutamate found in the enzymatic catalytic pocket (Figure 1.8) (Ko et al., 2003, Conticello et al., 2007).

AID was discovered a little over 15 years ago through a subtractive hybridization screen, when cDNA from an activated B cell line was compared to that of the same cell line after cytokine stimulation to undergo *in-vitro* isotype switching (Muramatsu et al., 1999). It was shortly thereafter shown that post-V(D)J immunoglobulin remodeling through somatic hypermutation and class-switch recombination were abrogated in the absence of the deaminase, as $AID^{-/-}$ mice fail to undergo either genomic alteration event, and that gene conversion in chicken was similarly unable to take place (Muramatsu et al., 2000, Arakawa et al., 2002). The enzyme is relatively small at 198 amino acids (Figure 1.9). Its catalytic domain is contained between amino acids 55 and 94, though other domains, such as the SHM-required N-terminus and the CSR-required C terminus, and co-factors help regulate its function as well (Barreto et al.,

2003, Ta et al., 2003, Chaudhuri and Alt, 2004, Shinkura et al., 2004, Ellyard et al., 2011, Ranjit et al., 2011). Although the only APOBEC family crystal structures available are those of APOBEC2 and APOBEC3G carboxy-terminal domain, a lot has been learned about AID by analogy to the other APOBECs (Prochnow et al., 2007, Chen et al., 2008, Conticello, 2008). However, and for a while, the target substrate of AID proved to be a controversial topic. Based

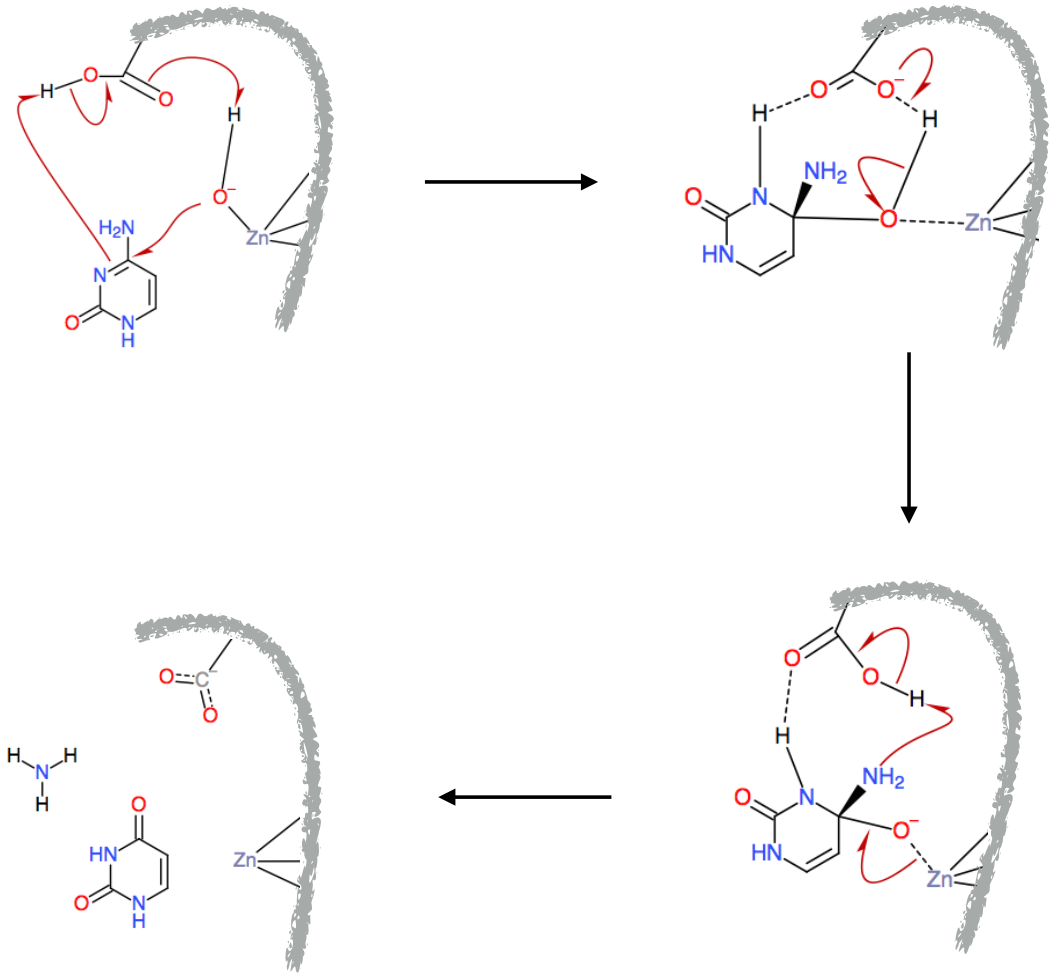


Figure 1.8. The deamination reaction. In the catalytic pocket of AID, a Zn atom is coordinated by 3 amino acids (2 cysteines and 1 histidine), and activates a water molecule that mounts a nucleophilic attack on the amine-bearing carbon of the deoxycytidine ring; this attack is facilitated by the participation of a glutamate residue, the carboxyl group of which is also depicted. At the end of the reaction, the amino group is lost, and the carbon is double-bonded to an oxygen atom.

on the deamination model of apolipoprotein B100 mRNA by APOBEC1, it was postulated that AID, similarly, used RNA substrates (MacGinnitie et al., 1995, Muramatsu et al., 2000, Steele and Blanden, 2001). In this model, AID, perhaps along with a substrate specificity-conferring cofactor, deaminates some mRNA, which can now encode a CSR recombinase or a somatic hypermutator, which in turn can deaminate Ig locus targets to initiate isotype switching or somatic hypermutation; in a variant of this model, AID-mutated Ig RNA is reverse-transcribed into DNA, from which RNA and eventually protein is made (Muramatsu et al., 2000). However, a number of studies have presented convincing evidence that physiological AID activity in B-cell immunoglobulin diversification is not on RNA but on DNA substrates. No reverse transcriptase in germinal center B cells has been described (Storb et al., 1999). AID was shown *in vitro* to catalyze the deamination of deoxycitidine residues on single-stranded DNA (Bransteitter et al., 2003, Dickerson et al., 2003, Shen and Storb, 2004). *Escherichia coli* experiments confirmed this by showing that AID transformation of *E.coli* results in the presence of dC:dG transition mutations on exposed single-stranded DNA in a transcription-enhanced phenomenon (Petersen-Mahrt et al., 2002, Ramiro et al., 2003, Sohail et al., 2003). Finally, *in vivo* reports using glycosylase-deficient mice that replication of DNA undergoing SHM yields transition mutations, and recent comprehensive, high-throughput RNA sequencing observations from mouse models provide direct and conclusive support to the DNA-deamination model (Rada et al., 2002, Fritz et al., 2013).

Given the potentially deleterious effects that uncontrolled deamination activity could have, it is to be expected that AID activity regulation mechanisms have evolved. *AICDA* (activation-induced cytidine deaminase, encoding AID) is expressed specifically in stimulated

secondary lymphoid tissue B cells (Muramatsu et al., 1999). Though AID substrate DNA is nucleus-bound and the enzyme exhibits an N-terminus nuclear localization signal, AID is principally found in the cytoplasm, by virtue of a chromosome region maintenance/exportin 1 (CRM1)-dependent, leucine-rich C-terminal nuclear export signal and a cytoplasmic retention mechanism, and its nuclear import being an active, energy-dependent transport (Ito et al., 2004, McBride et al., 2004, Geisberger et al., 2009, Patenaude et al., 2009). There are several additional ways in which the expression and function of AID are regulated. Nuclear AID can be polyubiquitinated, precipitating its degradation and shortening its half-life compared to cytoplasmic AID (Aoufouchi et al., 2008, Delker et al., 2013). AID can be phosphorylated at different residues, chief among them Ser38, but also Ser3, Thr140 and Tyr184; Ser38 phosphorylation by the holoenzyme protein kinase A is important for CSR and SHM, as it mediates interaction with the heterotrimeric replication protein A (RPA), a single-stranded DNA-binding protein (Figure1.9) (Chaudhuri et al., 2004, Basu et al., 2005, McBride et al., 2006, Pasqualucci et al., 2006, Basu et al., 2008, Cheng et al., 2009). Several proteins have been

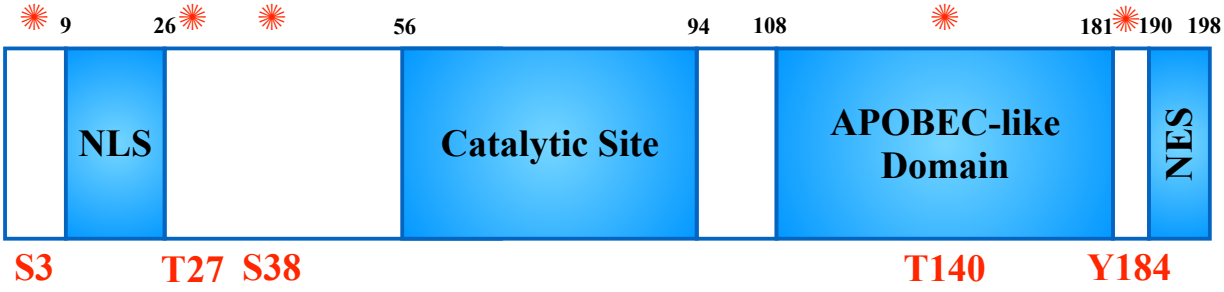


Figure 1.9. Domain structure of AID. The potent single-stranded DNA mutator is tightly regulated. It contains a nuclear localization signal at its amino terminal, and a nuclear export signal at its carboxy- end. Also highlighted are residues that get phosphorylated. The kinase PKA catalyzes the phosphorylation of AID at its serine 38 residue, allowing it to interact with RPA and bind single strands of DNA at its SHM and CSR target loci.

reported to regulate AID levels in B cells: heat shock protein 40 (HSP40) family member, J domain-containing DnaJa1 in its farnesylated form is necessary for physiological folding and levels of AID as well as for CSR (Orthwein et al., 2012); the ubiquitin- and ATP-independent protein degrader REG-γ, deficiency of which increases AID and CSR levels (Uchimura et al., 2011); RNA polymerase-associated factor PAF1, as well as RNA splicing factors CTNNBL1 and SRSF1, defects in which negatively impacts CSR, physiological AID levels in the nucleus, and somatic hypermutation, respectively (Ganesh et al., 2011, Kanehiro et al., 2012, Willmann et al., 2012). Finally, AID levels are also impacted more directly at transcription levels by E protein transcription factors (helix-loop-helix protein Id3 overexpression antagonizes *AICDA* expression), and post-transcriptionally by micro RNAs (miR-181b and miR-155) (Sayegh et al., 2003, De Yébenes et al., 2008, Dorsett et al., 2008, Teng et al., 2008).

Mature B lymphocytes express higher levels of AID protein upon stimulation *in vivo* and *in vitro* (Muramatsu et al., 2000). AID deaminates cytidine residues on transcribed IgH switch regions (in CSR) or transcribed IgH or IgL variable regions (in SHM), a reaction that effectively "turns" the cytidines into uridines (Petersen-Mahrt et al., 2002, Neuberger et al., 2003, Di Noia and Neuberger, 2007, Maul et al., 2011). Post-deamination, several, non-mutually exclusive options, are available (Figure 1.10). First, during replication, the uracil could be read as a thymine by DNA polymerase, leading ultimately to a transition from dG:dC base-pair to a dA:dT base-pair. Also, the dU, which is present as a result of AID activity, can be the substrate of Uracil-DNA Glycosylase isoform 2 (UNG2) (Kavli et al., 2005). UNG2 enables uracil base removal by catalyzing the hydrolysis of the N-glycosidic bond (Ide and Kotera, 2004). This creates an apyrimidinic site. If UNG2 activity, which most often appears in the initial steps of the

base excision repair (BER) pathway, is instead directly followed by replication across the apyrimidinic site, then both transitions and transversions can occur, as each one of the four DNA nucleotides could be incorporated where the uracil base was once found (Chaudhuri et al., 2007). A third option implicates the MSH2/MSH6 complex from the mismatch repair (MMR) pathway. This complex could, in the course of dU:dG mismatch processing, introduce single-stranded gaps, the filling of which, by error-prone polymerases (e.g., polymerase η), could usher in further mutations (Li et al., 2004, Wilson et al., 2005). These options could explain SHM; but for CSR, it is predicted that UNG2 activity is succeeded by that of the apurinic/apyrimidinic endonuclease 1 (APE1, also a BER pathway member), which will break the phosphodiester bond that UNG2,

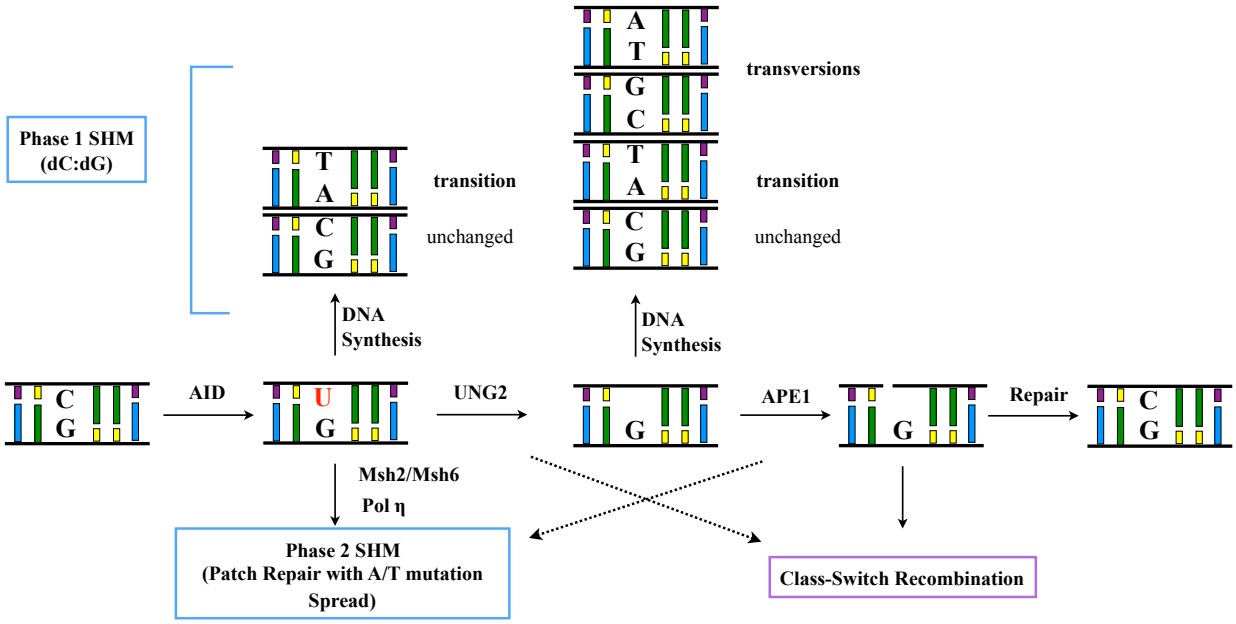


Figure 1.10. Cellular responses to AID deamination. The product of the AID-catalyzed reaction is processed through various low-fidelity pathways in B lymphocytes resulting in the diversification of both the binding and effector abilities of immunoglobulin proteins. Phase 1 SHM results from replication across a mismatch dU:dG or across an abasic site; Phase 2 SHM involves the mismatch repair (MMR) pathway, and induces patch repair. CSR may result from the engagement of the base-excision repair (BER) pathway, including the glycosylase UNG2 and the endonuclease APE1. Low levels of CSR are witnessed in the absence of this pathway, presumably as a result of some amount of MMR redundancy.

being a monofunctional glycosylase, could not disrupt, and create a nick in the switch region (Petersen-Mahrt et al., 2002). AID, though a single-stranded DNA deaminase, deaminates both the transcribed and the non-transcribed DNA strands (Milstein et al., 1998, Shen et al., 2006, Xue et al., 2006). If these nicks and/or gaps occur temporally and spatially closely enough, and on both DNA strands, then they could result in a DSB. This is referred to as the Neuberger model, in honor of the late Cambridge University immunologist and chemist.

CSR is thought to result from the recombination of a donor and an acceptor switch-sequence DSBs principally through non-homologous end-joining (NHEJ) (Stavnezer and Schrader, 2014). NHEJ and homologous recombination (HR) represent the two main pathways cells use to repair DSBs; which pathway is taken seems to hinge on the cell cycle and the availability and the relative abundance of the proteins needed to steer DNA repair along the pathway. HR uses the sister chromatids of replicated DNA strands as homologous templates to engineer repair, while NHEJ allows the synapsis of two DNA strands with little or no homology. While HR operates during the S and G2 cell cycle phases, NHEJ predominantly occurs during G1 (Symington and Gautier, 2011). Following the generation of a DSB, the initiation of resection to reveal homologous sequences suitable for HR signals commitment to this pathway. Resection is more pronounced during S/G2 than in G1, and this would help explain why NHEJ is the B cell-favored repair pathway in the G1-restricted CSR; furthermore, S regions display a lack of sufficient homology to undergo successful homologous recombination (Dunnick et al., 1993, Petersen et al., 2001, Schrader et al., 2007, Stavnezer et al., 2008, Symington and Gautier, 2011, Sharbeen et al., 2012, Khair et al., 2014) .

NHEJ requires the coordinated activities of several important proteins. DNA-dependent protein kinase (DNA-PK) is formed by the Ku heterodimer (Ku70 and Ku80) associated with DNA-PKcs (catalytic subunit) and is thought to bind DSB ends first due to the relative abundance of Ku (Stavnezer et al., 2008). Another complex, formed by MRE11, NBS1, and RAD50 (MRN complex) is also thought to rapidly bind to DSB ends and serve to assist in tethering them together. Additionally, the nuclease ARTEMIS associates with Ku and processes DNA ends, while the kinase ATM (ataxia telangiectasia mutated) is recruited and activated by MRN, and, in turn, phosphorylates several other proteins including NBS1, 53BP1, P53, CHK2, MDC1, and H2AX; this leads to more MRN accumulation, as well as the activation of cell-cycle checkpoints. Finally, Ku facilitates the binding of the ligase complex DNA ligase IV-XRCC4-XLF, whose activity is necessary for the proper completion of the synapsis reaction (Stavnezer et al., 2008).

I.5.2. AID Substrate-Targeting and in the Context of Chromatinized DNA

The faculty for AID to target its substrate sequences has long been the source of investigations. In cell-free assays, AID binds RNA but prefers DNA, even when it does not contain deoxycytidine residues, though the deamination reaction would only be observed on deoxycytidine-containing single-stranded DNA or transcribed double-stranded linear DNA (Chaudhuri et al., 2003, Dickerson et al., 2003, Pham et al., 2005, Ronai et al., 2007). It was also recognized that the enzyme had a proclivity for WRC/GYW motifs (where W = dA or dT, R = dA or dG, Y = dC or dT), since AID deamination was 4.6 times higher at these motifs than at coldspots during *in vitro* experiments, and 7.2 times higher at the same motifs than at coldspots

during *in vivo* experiments (Shapiro et al., 2002, Pham et al., 2003, Li et al., 2004, Larijani et al., 2005, Chelico et al., 2009).

Despite the many regulatory mechanisms, at multiple levels, that help ensure AID activity is not promiscuous, several non-immunoglobulin genes have been reported as targets of somatic hypermutation or AID; an early example of such a gene is *BCL6*, which is found to be mutated in normal memory B cells as well as their diffuse large B cell lymphoma (DLBCL) counterparts (Pasqualucci et al., 1998, Shen et al., 1998, Pasqualucci et al., 2001). In addition, other non-Ig targets of AID have been identified, including *PAX5* and *PIM1*, in DLBCL cells (Pasqualucci et al., 2004, Kotani et al., 2005, Okazaki et al., 2007). Another pathophysiological target of AID is *c-MYC*, the translocation of which with *IGH* is pathognomonic for Burkitt's lymphoma; AID has been found to be required for the necessary double-stranded DNA breaks that lead to that translocation (Ramiro et al., 2004, Robbiani et al., 2008). It appears that non-Ig targets of AID may share similar sequence motifs as those present in physiological target enhancers, but that they are also more the subject of error-free repair than physiological AID targets (Liu et al., 2008, Duke et al., 2013, Storb, 2014).

What has long been clearer is the role played by the chromatin milieu and state in general, and transcription in particular, to physiological AID activity. Transcription was initially implicated in AID-mediated immunoglobulin remodeling with the observation that in transgenic mice, immunoglobulin κ transcription enhancers were necessary for efficient somatic hypermutation, although the identity of the promoter did not matter much, since replacing the immunoglobulin-specific V_κ promoter with a β-globin promoter did not affect the ability of the cells to undergo SHM (Betz et al., 1994). It was thereafter shown that chromatin compaction as a

result of DNA hypermethylation of bacterial sequences inserted into Ig transgenes negatively impacts SHM, presumably because transcription is also negatively affected (Storb et al., 1996). Similarly, transcription through immunoglobulin heavy and light chain V(D)J exons or switch regions is required for SHM and CSR (Chaudhuri et al., 2007, Di Noia and Neuberger, 2007, Yang and Schatz, 2007, Maul and Gearhart, 2010). Observations that the constant region could be SHM-targeted as well if Ig constant region transcription was induced led to a model in which "a mutator factor that is present only in mutating B cells loads on to the transcription initiation complex assembled at the promoter and remains associated with the elongating complex" that will intermittently pause/stall, allowing for mutations to be catalyzed (Peters and Storb, 1996). AID was found to be the predicted mutating factor, and several additional co-factors were proposed, including replication protein A, protein kinase-Ar1α, and CTNNBL1, as well as, more recently, the non-coding immunoglobulin switch RNA generated following Lariat debranching (Chaudhuri et al., 2004, Basu et al., 2005, McBride et al., 2006, Pasqualucci et al., 2006, Conticello, 2008, Zheng et al., 2015). One important factor proposed in the last few years is suppressor of Ty 5 (Spt5 in yeast, SPT5 in human), a polymerase II pausing/stalling factor, which was found to interact with AID and to be required for physiological CSR (Pavri et al., 2010). The proposal of this protein as a cofactor of AID activity tentatively helped explain the mutation distribution profile of AID at its target sequences. Another important factor recently described is the RNA exosome complex (Basu et al., 2011, Pefanis et al., 2014, Pefanis et al., 2015).

I.5.3. The Ig R-Loop Reaction: RNA exosome and Senataxin

The implication of transcription in immunoglobulin remodeling strategies was all the more appreciable, because it helped reconcile in part the observation that AID is a single-strand DNA deaminase lacking activity on double-stranded DNA with the observation that both the template and the non-template DNA strands get mutated (Chaudhuri et al., 2003, Dickerson et al., 2003, Ramiro et al., 2003, Sohail et al., 2003). It was therefore proposed that the transcription reaction, by temporarily melting the DNA duplex and separating the two strands, exposes at least the non-transcribed strand to AID deamination activity. Indeed, mammalian switch region transcription results in the appearance of structures called R loops, that are formed when the RNA product of transcription hybridizes to the transcribed DNA and displaces the non-transcribed strand.

I.5.3.1. R Loops

Short RNA:DNA hybrids can be seen in different physiological processes, as ~10 base pair-long primers for Okazaki fragments during DNA replication for instance (Sakabe and Okazaki, 1966, Okazaki et al., 1968, Egli et al., 1992, Waga and Stillman, 1998). During transcription, an ~8 bp-long RNA:DNA hybrid can be found within the active site of the RNA polymerase (Nudler et al., 1997). The RNA strand and the transcribed DNA strand exit the polymerase through two different channels, which impairs their ability to remain base-paired (Westover et al., 2004, Costantino and Koshland, 2015). However, the negative supercoils that arise in DNA in the wake of transcription by pol II may unwind the two DNA strands to the

extent that the synthesized RNA is allowed to anneal back to the transcribed DNA strand, forming a stable, longer RNA:DNA heteroduplex, and displacing the non-transcribed strand (Liu and Wang, 1987, Aguilera, 2002, Gowrishankar and Harinarayanan, 2004, Li and Manley, 2006, Kim and Jinks-Robertson, 2012, Costantino and Koshland, 2015). The relevance of supercoils was underscored by reports that AID, through the formation of RNA polymerase II transcription-generated supercoils, may be able to be active on both the transcribed and non-transcribed DNA strands of some transcribed plasmids (Shen et al., 2005, Besmer et al., 2006, Basu et al., 2011).

One additional feature favoring the stability of R loops is that thermodynamic investigation of RNA:DNA hybrids suggests that depending on base composition some heteroduplexes can be quite stable compared to their DNA homoduplex counterparts (Thomas et al., 1976, Chien and Davidson, 1978, Conn et al., 1999). What is more, evidence suggests that heteroduplexes in which the RNA moiety is purine-rich (as would be the case during transcription at immunoglobulin switch sequences) are more stable than heteroduplexes with a pyrimidine-rich RNA strand (Hung et al., 1994, Ratmeyer et al., 1994, Lesnik and Freier, 1995, Wang and Kool, 1995). Factors that are antagonistic to R-loop formation include in bacteria the coupling of transcription and translation, which limits the extent to which RNA finds itself naked and able to base-pair with the DNA template strand (Gowrishankar and Harinarayanan, 2004). Though transcription and translation occur in different cellular compartments in eukaryotes, naked nascent RNA is co-transcriptionally coated with proteins that mediate splicing, export into the cytoplasm, or secondary structure formation, which has the effect of not only decreasing relative amounts of available RNA in the nucleus, but also preventing the accumulation of naked, linear RNA, which would more easily anneal to DNA (Kim and Jinks-Robertson, 2012). Finally,

the superhelical stress-relaxing enzyme topoisomerase 1 (TOP1) holds R-loop formation in check (Tuduri et al., 2009, Kim and Jinks-Robertson, 2012).

The presence of regulatory mechanisms that prevent R-loop accumulation suggests that the latter may be problematic for the cell. Early evidence of negative consequences related to the presence of R loops in the form of genomic instability events comes from observations in the yeast *hpr1* hyper-recombinant mutant phenotype, in which the THO-TREX complex of mRNA metabolism and export, of which HPR1 is a subunit, is no longer able to assemble RNA into ribonucleoproteins, promoting R-loop formation, impairment of transcription elongation (with the attendant collision conflicts with the replication machinery), and recombination (Chávez and Aguilera, 1997, Huertas and Aguilera, 2003, Li and Manley, 2006, Rondón et al., 2010, Kim and Jinks-Robertson, 2012). In chicken DT40 cells, too, a genomic instability phenotype was observed when serine/arginine-rich splicing factor (SRSF1) depletion resulted in splicing complex assembly defects, subsequent R-loop accumulation, and high levels of genomic rearrangements (Li and Manley, 2005, Kim and Jinks-Robertson, 2012). In general, it is believed that the stretches of single-stranded DNA that are exposed in R loops are less stable than double-stranded DNA, and are at increased risk of transcription-associated mutagenesis (TAM), recombination (TAR), and double stranded DNA breaks (DSBs) (Beletskii and Bhagwat, 1996, Muers, 2011, Kim and Jinks-Robertson, 2012, Wimberly et al., 2013, Costantino and Koshland, 2015).

A number of human pathologies have been associated more or less convincingly with R-loop accumulation (Groh and Gromak, 2014, Costantino and Koshland, 2015). A well-defined example is Fragile X syndrome, the most common genetic form of mental retardation, in which

R-loop formation has been shown to increase with CGG trinucleotide repeat element expansion within the 5'untranslated region of the *FMR1* (the fragile X mental retardation 1 gene), impairing the physiological expression of the protein and resulting in gene silencing (Colak et al., 2014). Another example concerns the neurodegenerative disorders amyotrophic lateral sclerosis and frontotemporal dementia, where another nucleotide repeat expansion (involving the hexanucleotide GGGGCC) favors the formation of R loops, leading to transcription abortion in the hexanucleotide repeat expansion region, and molecular cascades that lead to pathologies (Haeusler et al., 2014). Finally, R-loop accumulation and the genomic instability that results from it have been proposed as being involved in the pathophysiology of some cancers. The observation that depletion of tumor suppressor BRCA2 (which is mutated and inactivated in some breast cancers) leads to R-loop accumulation and DNA damage checkpoint activation suggests that BRCA2 may work to suppress cancer development in part by preventing R-loop accumulation (Roy et al., 2012, Bhatia et al., 2014, Costantino and Koshland, 2015). Similarly, increased activity levels of the recombinase RAD51 and the myeloproliferative neoplasm-related FIPL1 in some cancers, may contribute to genomic instability, given that both these proteins promote RNA:DNA hybrid formation in yeast (Gotlib and Cools, 2008, Klein, 2008, Wahba et al., 2013, Costantino and Koshland, 2015).

The preceding deliberation notwithstanding, R loops have paradoxically also been shown to be involved in several physiological functions. The expression in the mouse-ear cress *Arabidopsis thaliana* of the long non-coding RNA cold-induced long antisense intragenic RNA (COOLAIR) regulates the flowering locus, and has been reported to be silenced by R-loop formation in the promoter region of COOLAIR (Sun et al., 2013b, Costantino and Koshland,

2015). Additionally, a link has been suggested between R-loop formation, chromatin compaction, and epigenetic modifications, notably H3S10 phosphorylation both in yeast and human cells, hinting at a the possibility of R loops as regulators of chromosomal structure (Skourti-Stathaki et al., 2011, Castellano-Pozo et al., 2013, Skourti-Stathaki et al., 2014, Costantino and Koshland, 2015). Moreover, R-loop formation has been described at telomeric regions that produce the non-coding telomeric repeat–containing RNA (TERRA); telomeric R-loop formation has been proposed to impede senescence in the budding yeast *Saccharomyces cerevisiae* by mediating telomere elongation in homologous recombination-competent telomerase mutant backgrounds (Balk et al., 2013, Costantino and Koshland, 2015, Cusanelli and Chartrand, 2015).

An additional, most important physiological function ascribed to R loops relates to their formation as a result of transcription at immunoglobulin switch sequences, where they can top 1 kb in length (Daniels and Lieber, 1995, Tian and Alt, 2000, Shinkura et al., 2003, Yu et al., 2003). R loop formation at switch sequences is thought to be favored in part by the asymmetric G-richness of the looped-out, non-transcribed DNA strand, which, unlike the transcribed strand, is hypothesized to be able to form stabilizing G-quadruplex DNA (G4-DNA) comprised by stacks of Hoogsteen hydrogen-bonded guanine quartets (Sen and Gilbert, 1988, Dunnick et al., 1993, Yu et al., 2003, Duquette et al., 2004, Maizels, 2006, Kim and Jinks-Robertson, 2012). The looped-out, non-transcribed DNA finds itself in single-stranded conformation, and is a suitable target for AID activity, as evidenced in part by the isotype switching defect seen when the G-rich sequence is inverted or otherwise altered (Shinkura et al., 2003, Yu et al., 2003). This, however, does not fully address the question of how AID is able to access the transcribed DNA

strand, which also gets deaminated at immunoglobulin switch sequences (Milstein et al., 1998, Shen et al., 2006, Xue et al., 2006).

I.5.3.2. The RNA Exosome Complex

These i*n vivo* observations that AID deamination activity was present on both the immunoglobulin switch region transcribed strand and its non-transcribed base-pairing counterpart had stumped investigators for whom cell-free, biochemical assays had only helped explain transcription-mediated AID access to the non-transcribed strand (Chaudhuri et al., 2003). Even at physiological, R loop non-forming somatic hypermutation double-stranded DNA substrates, AID phosphorylated on its serine 38 residue by protein kinase A had been shown to access mainly the non-template strand (Basu et al., 2005, Chaudhuri et al., 2007, Basu et al., 2008). Other experiments, involving the expression of exogenous AID in bacterial and yeast assays, revealed the prevalence of mutations on the non-template strand of transcribed substrates (Ramiro et al., 2003, Gómez-González and Aguilera, 2007). In the looped-out model for AID access to the non-transcribed strand, enzymatic deamination of the transcribed strand is seemingly impaired by the nascent RNA hybridizing to its template DNA and hiding it from mutagenesis. The absence of reports of AID deamination activity on hybrid RNA:DNA strands, or of evidence that the endoribonuclease ribonuclease H (RNase H) exposes switch sequence R-loop template strands seemed to dash any hope of a parsimonious mechanistic explanation of the transcribed strand access problem (Yu and Lieber, 2003, Lieber, 2010). This changed when the RNA exosome complex was shown to target the catalytic activity of AID to both the non-transcribed and the transcribed strands of transcribed duplex DNA substrates, initiating the generation of DNA double-stranded break intermediate substrates for constant region

immunoglobulin remodeling and chromosomal interactions (Wuerffel et al., 1997, Petersen et al., 2001, Ramiro et al., 2004, Schrader et al., 2005, Robbiani et al., 2008, Basu et al., 2011).

The RNA exosome is a conserved, multi-subunit complex whose global architecture is seen under various forms in all domains of life (Figure 1.11). There are three classes of exoribonuclease catalytic activity: bacterial RNase II and RNase R, and eukaryotic ribosomal RNA processing protein 44 (Rrp44) have processive hydrolytic RNA decay activity; bacterial RNase D and eukaryotic Rrp6 possess distributive hydrolytic RNA decay activity; and, finally, bacterial and mitochondrial polynucleotide phosphorylase (PNPase) and the archaeal RNA exosome catalyze processive phosphorolytic exoribonuclease activity (Januszyk and Lima, 2014). In eukaryotes, the RNA exosome is a ubiquitous endoribonuclease and 3′ → 5′ exoribonuclease composed of an RNA-binding, non-catalytic nine-subunit core similar in architecture to PNPase but sans its phosphorolytic activity (Exo9, composed of Csl4 (for Cep1 synthetic lethal), Rrp4, Rrp40, Rrp41, Rrp46, Mtr3 (for mRNA transport 3), Rrp42, Rrp43, and Rrp45), which, depending on the subcellular localization, interacts with and regulates either or both catalytic subunits Rrp44 and Rrp6 (Mitchell et al., 1997, Allmang et al., 1999, Symmons et al., 2000, Anderson et al., 2006, Houseley et al., 2006, Liu et al., 2006, Oddone et al., 2007, Greimann and Lima, 2008, Lebreton et al., 2008, Schaeffer et al., 2009). Yeast studies suggest that in its cytoplasmic form, the RNA exosome complex is composed of Exo9 and Rrp44, while in the nucleus RNA exosome contains Rrp6 added to the Exo9-Rrp44 pairing found in the cytoplasm; finally, in human, a nucleolar complex is posited to be comprised of Exo9 and Rrp6 (Mitchell et al., 1997, Tomecki et al., 2010, Wasmuth et al., 2014). The subunits comprising Exo9 form a pseudo-hexameric ring with RNA-binding activity inside the channel; the ring is

	Protein Subunit	Characteristic Domain	Human Gene
1	CSL4	S1	*EXOSC1*
2	RRP4	S1/KH	*EXOSC2*
3	RRP40	S1/KH	*EXOSC3*
4	RRP41	RNase PH	*EXOSC4*
5	RRP46	RNase PH	*EXOSC5*
6	MTR3	RNase PH	*EXOSC6*
7	RRP42	RNase PH	*EXOSC7*
8	RRP43	RNase PH	*EXOSC8*
9	RRP45	RNase PH	*EXOSC9*
10	RRP6	RNase D	*EXOSC10*
11	RRP44	RNase R	*EXOSC11*

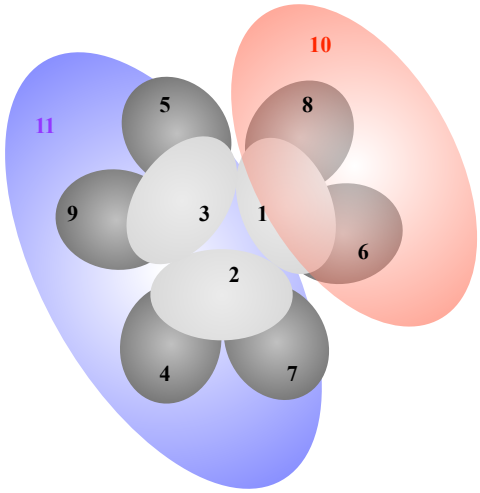


Figure 1.11. The RNA exosome complex. Illustrated here is the complex composed of its nine non-catalytic core subunits: the 3-subunit cap in light grey, the 6-subunit ring in dark grey. The cap can bind RRP6, which has exoribonuclease activity; the bottom ring can associate with RRP44, which has both exo- and endo-ribonuclease active sites.

made up of three heterodimeric pairings: Rrp41-Rrp45, Rrp46-Rrp43, and Mtr3-Rrp42 (Januszyk and Lima, 2014). The ring is then bound by the heterotrimeric ring cap comprised of Csl4, Rrp4, and Rrp40, whose channel is continuous with that formed by the hexameric non-cap ring (Januszyk and Lima, 2014). It has been proposed that Rrp44 associates with the bottom of Exo9 through interactions between its Rrp44 endoribonuclease-containing, N-terminal Pilus-forming N-terminus (PIN) and one of its cold-shock domains (CSD1) with core subunits Rrp41, Rrp45, and Rrp43; the same model also suggests that Rrp6 associates with the top of Exo9 through its Rrp6 C-terminal domain (6CTD) interacting with Csl4, Mtr3, and Rrp43 (Makino et al., 2013, Januszyk and Lima, 2014). When the Exo9 core binds both catalytic subunits, evidence from

current structural models points to a channel-containing, RNA-binding non-catalytic hetero-nonamer, sandwiched by catalytic subunits (Makino et al., 2013, Januszyk and Lima, 2014).

The RNA exosome complex has been shown to interact with a number of factors that help it identify which RNA substrates to process/degrade. In yeast, RNA exosome participates in the $3' \rightarrow 5'$ degradation of mRNA in the cytoplasm by interacting with the Ski complex and the translation machinery (Araki et al., 2001, Wang et al., 2005, Halbach et al., 2013, Januszyk and Lima, 2014). In the nucleus, RNA exosome has been shown to interact with Rrp47 (also known as C1D in human) via the catalytic Rrp6, and with Mpp6 to process RNA species like the 3' extended version of the precursor to 5.8S rRNA (Mitchell et al., 2003, Schilders et al., 2005, Stead et al., 2007, Milligan et al., 2008, Costello et al., 2011, Januszyk and Lima, 2014). The nuclear RNA exosome has also been shown to interact with the TRAMP complex (Trf4/Trf5 poly(A) polymerases, Air1/Air2 Zn-knuckle RNA-binding proteins, and Mtr4 RNA helicase) to carry out aberrant RNA degradation and surveillance (LaCava et al., 2005, Callahan and Butler, 2010, Januszyk and Lima, 2014). One putative interaction is with the nuclear RNA exosome targeting (NEXT) complex (comprised, in its human iteration, of hMTR4, the Zn-knuckle protein ZCCHC8, and the putative RNA binding protein RBM7) to facilitate promoter upstream transcript degradation (Preker et al., 2008, Lubas et al., 2011, Januszyk and Lima, 2014). In addition to its interaction with AID, two other interactions of particular importance have been suggested (Basu et al., 2011). One is with the NNS complex (Nrd1, Nab3, and Sen1 — alternatively called the NRD complex), allowing the degradation and/or processing of several small nuclear and small nucleolar RNA polymerase II transcripts (Vasiljeva and Buratowski,

2006, Rondón et al., 2009, Januszyk and Lima, 2014). The other is with RNA polymerase II transcription elongation cofactors Spt5/6 (Andrulis et al., 2002).

I.5.3.3. RNA in the Ig R-loop Reaction: RNA exosome, Stalling, and Senataxin

The RNA exosome complex proved to be a long-sought cofactor of AID targeting in post V(D)J B-cell genome AID-dependent remodeling reactions through *in vitro* and *in vivo* experiments (Basu et al., 2011, Pefanis et al., 2014). The RNA exosome complex is required for proper SHM substrate AID deamination activity on the non-transcribed and the transcribed DNA strands. The RNA exosome complex accumulates at the switch sequences of B cells activated to undergo isotype switching in an AID-dependent fashion, where it also associates with the mutator enzyme. CSR is severely impaired when the RNA exosome complex cannot function properly.

The successful production of functional antibodies necessitates full-length transcription at immunoglobulin loci and is required to avoid apoptosis (Liu et al., 1989). The presence of AID-initiated alterations in these loci suggests that full-length transcription occurs after the DNA mutation events have taken place. In an updated model for AID activity at its physiological Ig substrates, the deamination events caused by AID occur when the mutator enzyme has successfully and stochastically been loaded onto a transcription machinery complex at the transcription start site, which becomes associated with transcription pausing/stalling factors; this would favor pausing of the complex, permit the degradation of the nascent RNA such that transcription is aborted and polymerase released, and allow AID deamination reactions on

exposed transcribed DNA (Figure 1.12) (Kuehner et al., 2011, Sun et al., 2013a, Storb, 2014). While the RNA exosome complex is thought to mediate the removal of the transcription product, SPT5 has been reported to associate with CSR-involved polymerase II at AID target sequences (Pavri et al., 2010). With SPT4, SPT5 forms a heterodimeric complex called 5,6-dichloro-1-β-d-ribofuranosyl-benzimidazole (DRB) sensitivity inducing factor (DSIF), which, along with negative elongation factor (NELF), has been shown to interact with polymerase II and lead to pausing (Wada et al., 1998, Yamaguchi et al., 1999a, Yamaguchi et al., 1999b).

Elongating polymerase II complexes exist in several states (Adelman and Lis, 2012, Jonkers and Lis, 2015). The poised pol II pre-initiation complex near the transcription start site contains pol II as well as general transcription factors, is bound to the promoter, but has not started synthesizing RNA. The pol II complex can then become stalled, a term signifying that it is associated with pausing-inducing factors, such as NELF, DSIF, and transcript cleavage factor TFIIS, all of which are involved in pausing, arresting, or terminating transcription by the complex. Paused pol II is able to resume elongation, as the RNA is properly aligned in its polymerase active site, although the RNA 3' end may have unraveled to a negligible degree, and RNA synthesis momentarily been stopped (Adelman and Lis, 2012). Arrested pol II has backtracked along the transcribed DNA such that the RNA 3' end is no longer in place in the active site. Arrested pol II complex restart is possible but requires additional factors. The nascent RNA could alternatively become dissociated from the DNA template if the pol II elongation complex is too unstable, terminating transcription and releasing the polymerase, which can recycle at the promoter (Muse et al., 2007, Zeitlinger et al., 2007, Lee et al., 2008, Adelman and Lis, 2012).

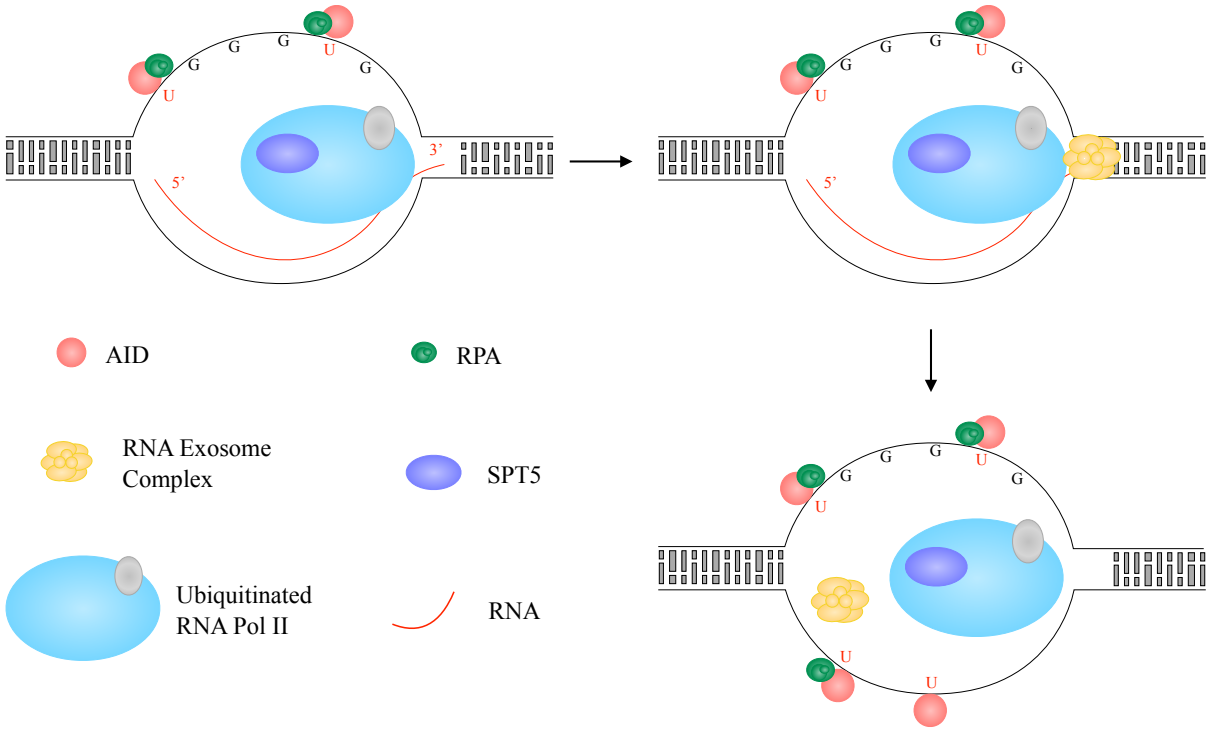


Figure 1.12. The Ig R-loop model. At *IGH* switch sequences, SPT5-mediated RNA polymerase II stalling/pausing and transcription-favored R-loop formation contribute to give AID access to its substrates in single-stranded DNA conformation. The RNA exosome complex, through its ribonuclease activity is a crucial co-factor of AID activity here, in part by recognizing the RNA 3' end exposed by backtracked polymerase, and helping remove/degrade it, to reveal the transcribed strand and make it susceptible to AID activity. Not depicted here is the hypothetical involvement of RNA/DNA helicase senataxin.

Some genomic regions are more prone to stalling, and transcription termination could help prevent the accumulation of paused RNA polymerase II complexes on the DNA (Anindya et al., 2007, Svejstrup, 2010, Sun et al., 2013a). It is thought that RNA pol II release from the template DNA strand can be precipitated by a polyubiquitination event that is initiated by a monoubiquitination reaction catalyzed by the homologous to the E6-AP carboxyl terminus (HECT) domain-containing E3 ligase NEDD4 (Rotin and Kumar, 2009). The NEDD4-mediated

monoubiquitination of the lysine 63 residue of RNA pol II is followed by the polyubiquitination of RNA pol II K48 by a different enzyme (Harreman et al., 2009, Sun et al., 2013a). It was recently shown that pol II transcription complex at both noncanonical and CSR substrates of AID is ubiquitinated by NEDD4 (Sun et al., 2013a). NEDD4 activity at RNA exosome target sequences was shown to be required, linking pol II release from its DNA template to RNA exosome access to the nascent RNA 3' end (Keim et al., 2013).

Yet, transcription termination of sterile transcripts at AID target sequences as a mechanistic step in antibody diversification is still not entirely elucidated. RNA:DNA hybrids that transiently appear during transcription are assembled into heteroduplex helices that vary in conformation depending on nucleotide content between A-form and B-form structures (Rich, 1960, Furth et al., 1961, Egli et al., 1992, Egli et al., 1993, Hung et al., 1994, Ratmeyer et al., 1994, Lesnik and Freier, 1995, Wang and Kool, 1995). To date, only hydrolytic ribonuclease catalytic activities have been described for the RNA exosome complex, leaving open the possibility that its activity is supported by a cofactor with helicase activity (Januszyk and Lima, 2014). Reports that in yeast, Sen1 as a member of the Nrd1-Nab3-Sen1 NNS transcription termination/processing pathway that generates non-coding RNA, may be associated with the RNA exosome complex via Nrd1, or that SETX, the human homologue of Sen1, may be associated with the RNA exosome by interacting with RRP45, one of its nine core subunits, hints at the tantalizing possibility that a helicase might be involved in the transcription termination reaction steps that accompany AID-initiated mutations at Ig switch sequence transcription-generated R loops (Vasiljeva and Buratowski, 2006, Rondón et al., 2009, Richard et al., 2013, Porrua and Libri, 2015).

Sen1 was identified through a tRNA-*s*plicing *en*donuclease activity screen almost 30 years ago, which showed that mutations in *Sen1* caused temperature-sensitive growth, reduced *in vitro* endonuclease activity, and *in vivo* accumulation of unspliced pre-tRNAs (Winey and Culbertson, 1988). Based on the sequence motif-based comparisons with other enzymes and the realization that a wider spectrum of RNA species are affected by defects in Sen1 activity, it was proposed that Sen1 is a RNA helicase member of the helicase superfamily I (SFI) (Koonin, 1992, Ursic et al., 1997). Later, more sophisticated analyses confirmed the SF1 classification of Sen1 as a non-ring forming, DExxQ motif-containing, Upf1-like, ATP-dependent, 5' → 3' RNA and DNA helicase, with putative activity on RNA:DNA heteroduplexes akin to some members of the SF1 superfamily such as Pif1 (Boulé and Zakian, 2007, Fairman-Williams et al., 2010, Jankowsky, 2011). The discovery that the human orthologue of Sen1, senataxin (for Sen1 - ataxia - protein), is mutated in a loss-of-function manner in the neuromotor disorder ataxia with oculomotor apraxia type 2 (AOA2) was further evidence that RNA biogenesis, even of non-coding species, plays an important role in cellular health (Figure 1.13) (Moreira et al., 2004, Fogel et al., 2009). Other disorders, most prominently, a juvenile form of Lou Gehrig's disease, amyotrophic lateral sclerosis type 4 (ALS4), have since been associated with dominant mutations in senataxin (SETX) (Chen et al., 2004).

I.5.4. Summary of Major Projects and Hypothesis

From the very start, investigators of adaptive immunity have wrestled with a **central conundrum**: how cells engineer a large number of varied antigen-binding proteins from a finite number of genes. Over the years, the discovery of immunoglobulin-gene recombination reactions

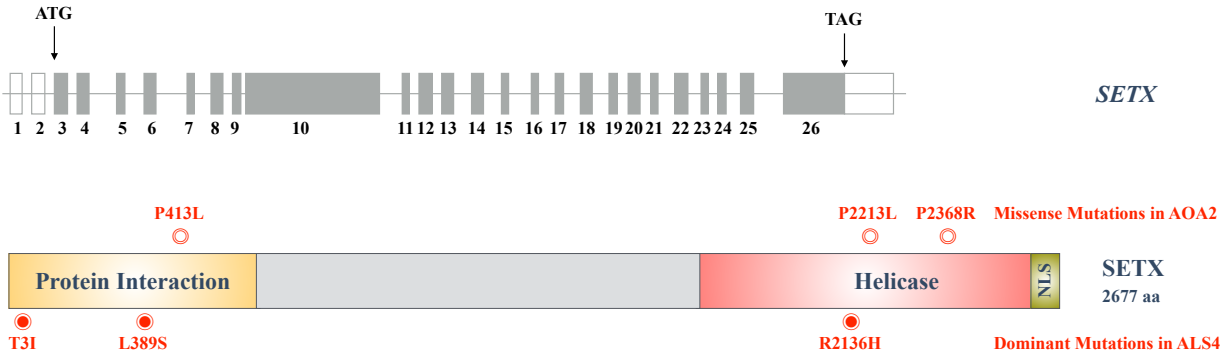


Figure 1.13. Domain structure of SETX. The open reading frame for *SETX* is large (>8000kb) and includes 26 exons, though the first 2 do not code for proteins. Some reports suggest the presence of alternative transcription start sites in exon 10, and additional isoforms from differential splicing involving exon 24 and 25 (Fogel et al., 2009). Regions of homology between SETX and Sen1 are shown (protein interaction N-terminal domain; and carboxy-terminal helicase domain and nuclear localization signal (NLS). Also shown are reported loss of function mutations in AOA2 and gain of function mutations in ALS4.

has considerably improved our understanding of that central question, particularly as it pertains to **B lymphocytes**.

Like T cells, B cells undergo antigen-receptor locus rearrangement through V(D)J recombination. Unlike T cells, they further modify their antigen receptor loci through the processes of Somatic Hypermutation (SHM) and Class-Switch Recombination (CSR), two physiological alterations initiated by the mutator enzyme **Activation-Induced cytidine Deaminase (AID)**, whose expression is most distinctive in the Germinal Center-(GC) B cells of lymph nodes and the spleen. These three types of physiological, controlled damage to the DNA require **transcription** in order to proceed; and the recruitment of repair factors to sites of physiological DNA damage may depend in part on the transcript(s) produced and on the chromatin milieu.

The study of B cells is of great value, given that mutations in AID have been linked to immunodeficiencies, and that some oncogenic translocations in B-cell lymphomas have been associated with AID, highlighting the central place that the **maintenance of the B-cell genomic organization occupies both physiologically and pathophysiologically**.

In this context, we set out to investigate different aspects of the biological processes that are thought to contribute to the maintenance of B-cell genomic organization. In particular, the projects discussed below aim to elucidate the **roles played by the RNA exosome and senataxin — by helping resolve RNA:DNA hybrids — in sustaining, and restricting, needed physiological genomic alteration reactions to suitable substrates and in preventing the generation of catastrophic mutagenic events**.

RNA Exosome and AID-dependent B-cell genomic instability. AID, a single-stranded DNA deaminase, accesses both the template and the non-template strands of DNA *in vivo*. Our group has reported that the RNA processing/degradation complex RNA exosome enables AID access to both strands of DNA in B cells during CSR. Beyond its role in facilitating access to both DNA strands, we want to investigate the role played by the RNA exosome in targeting the mutagenic consequences of AID activity to its physiological substrates, thereby preventing genomic instability. Our hypothesis is that in the absence of the RNA exosome, AID sub-stoichiometrically accesses the RNA-hybridized template DNA strand; as a result, the non-template strand bears the brunt of AID deaminase activity. The product of this dysregulated activity may no longer be suitable for physiological recombination during isotype switching, but may instead prefer undergoing translocation events with non-physiological partners, such as the

proto-oncogenic *c-MYC* locus. In such a setting, the RNA exosome would indeed emerge as a protector of AID-mediated genomic integrity.

Senataxin and AID-dependent B-cell genomic instability. We set out to investigate the role that the RNA/DNA helicase senataxin (SETX) plays in the sequence of biological events that result in CSR. SETX is the human homologue of Sen1, a member of the yeast RNA/DNA helicase superfamily I. In yeast, Sen1 forms a complex with Nrd1 and Nab3 thought to bind RNA polymerase II, unwind the RNA:DNA hybrid, and therefore play an important role in transcription termination. One way this may occur is by the resolution of the R loops that are created at G-rich pause sites downstream of polyadenylation signals. We reasoned that SETX may have a similar role in the G-rich Ig switch sequences found in B cells, and that in its absence, the RNA:DNA hybrids may not be unwound, obstructing RNA exosome-mediated removal or degradation of sterile transcripts, and, ultimately, preventing AID access to the template strand. In this context, senataxin, like RNA exosome, would also play a role in guaranteeing that the genomic integrity of the cell is maintained.

Chapter II: Materials and Methods

II.1. Materials

II.1.1. Conditionally-Deletable RNA exosome Mouse

The recombined mouse designed by, and generated under the guidance of, Evangelos Pefanis, has several particularities that have been published (Pefanis et al., 2014). It contains at its *ROSA26* locus a transgene encoding the Cre recombinase fused to the estrogen receptor T2. At the *EXOSC3* locus, two pairs of mutually exclusive heterologous Cre substrate *loxP* sites flank the portion of the locus bounded by exons 2 and 3. In intron 3 are also inserted the inverted sequences for a splice acceptor, a T2A-GFP cassette, and a polyadenylation transcription termination signal. When Cre-mediated inversion occurs, involving two facing monads of either pair of *lox* sites, exon 1 can be spliced onto the splice acceptor, and transcription termination can cease after a GFP reporter RNA has been synthesized; moreover, a deletion reaction is now

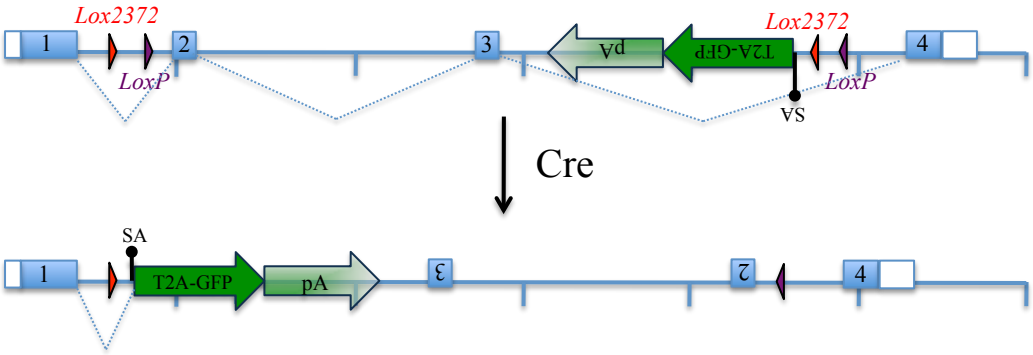


Figure 2.1 The *EXOSC3* allelic inversion-deletion reaction. Cre recombinase activity on same-orientation homologous *lox* sites can result in deletion of the intervening sequence, while Cre recombinase activity on opposite-orientation homologous *lox* sites can result in reversible inversion of the intervening sequence. When Cre-ERT2 is translocated into the nucleus, it can catalyze an inversion reaction between the 2 *loxP* sites or the 2 *lox2372* sites. The pair involved in the inversion reaction will now have one of its sites flanked by the two sites of the pair that did not engage in the inversion reaction; the two *lox* sites that did not engage in the inversion reaction find themselves in the same orientation and can mediate the deletion of the intervening sequence, meaning they can now facilitate the deletion of one of the 2 *lox* sites involved in the preceding inversion reaction, preventing it from being reversible and locking in place the inverted conformation. Shown above is an example using the *loxP* sites as being involved in the inversion reaction, and the *lox2372* sites as the ones mediating the deletion reaction.

possible between two homologous adjacent *loxP* sites and traps in the post-inversion conformation. The strategy is summarized in Figure 2.1.

II.1.2. HB-GFP Construct

We are grateful to have received the construct for a recombined protein encoding the RNA:DNA Hybrid Binding domain of RNaseH1 which has been fused at its amino- end to a GFP tag (Bhatia et al., 2014). The backbone is the mammalian expression vector pEGFP-C1, which encodes kanamycin resistance for growth in bacteria at 30 $\mu g.mL^{-1}$. The construct was introduce into CH12F3 cells by electroporation.

II.1.3. SETX-Knockdown shRNAs and real-time quantitative PCR

We introduced two shRNA constructs in the CH12F3 cell line. The two constructs are:

5’CCGGCCCAGCTAAGTGGGAAAGAAACTCGAGTTTCTTTCCCACTTAGCTGGGTTTT TG3’

and

5’CCGGGCGGTTGATGAACTTATGAAACTCGAGTTTCATAAGTTCATCAACCGCTTTTT G3’

We used the polyethylenimine (PEI) method for transfection. The shRNAs were mixed with the intended DNA/packaging vectors (psPAX2 and pMD2.G) at a ratio of 4:3:1 (shRNA:psPAX2:pMD2.G) and PEI was added so as to have a DNA:PEI ratio of 3:4. Reduced serum medium Opti-MEM was added to the total at this point would be 1 mL. At the same time,

293 cells (passaged the night before so s to be ~75% confluent) are in 4 mL medium, and the 1 mL nucleotide-PEI-Opti-MEM mixture is added. The medium is changed once after 4-6 hrs, and a second time after overnight. The virus-containing supernatant is collected 48 hrs later. In a 1:1 ratio, the virus-containing supernatant is added to a CH12F3 cells plated at 2 $x10^5$ cells.mL^{-1} and containing polybrene at 10-20 μg.mL^{-1}, so the final polybrene concentration is 5-10 μg.mL^{-1}. The cells can then be spun for 90 min at 1230 *g* at 30˚C. Finally, 75% of the supernatant can be removed and replaced with appropriate growth medium. After 48 hrs, the cells can be exposed to the selecting agent (in our case puromycin, at 0.5 μg.μL^{-1}).

II.1.4. SETX-Knockout in CH12F3 Cells

To knockout *SETX* in CH12F3 cells, we made use of the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 technology (Ran et al., 2013b). We used the D10A RuvC/Cas9 nickase mutant, which can nick but will not cause a double-stranded break (DSB). A particular region of interest must therefore be targeted by two different nickases in order to induce DSB formation, with the beneficial effect of increasing DSB specificity over a system using wild-type Cas9 (Ran et al., 2013a, Ran et al., 2013b). We targeted mouse *SETX* and *AICDA*, the genes that encode the putative RNA/DNA helicase senataxin (SETX) and the B cell mutator enzyme activation-induced cytidine deaminase (AID). For each target, we devised two pairs of single guide RNA (sgRNA) sequences, which help the CRIPSR-Cas9 system recognize its target in DNA.

SETX-targeted sgRNA pair 1: 5’-CTGCTACTGCTTAGAGTGTG-3'

and 5'-TGCTAGATGCATAGCGCTTT-3'

SETX-targeted sgRNA pair 2: 5'-AAGCGCTATGCATCTAGCAC-3'

and 5'-TCAATGGTGGAAGACCCACC-3'

AICDA-targeted sgRNA pair 1: 5'-TGAGACCTACCTCTGCTACG-3'

and 5'-CATGCCGTCCCTTGGCCCAG-3'

AICDA-targeted sgRNA pair 2: 5'-ACCTCTGCTACGTGGTGAAG-3'

and 5'-GTAGGTCTCATGCCGTCCCT-3'

The *SETX* sgRNAs target exon 2, while exon 3 is targeted by the *AICDA* sgRNAs.

II.1.5. SETX-Mutant Mouse

We obtained a mouse model to study SETX, wherein exon 4 has been deleted and replaced by a *lox* site (Becherel et al., 2013). To genotype we used the following primers:

Forward Primer: 5′-TTTAAGGAACAGTGCTGC-3′

Reverse Primer: 5′-ATGAAGCAGGTAGGATT-3′

Reverse Primer lox: 5′-CGAAGTTATATTAAGGGT-3′

II.2. Methods

II.2.1. *EXOSC3* Allelic Deletion in Murine Cells

Mus Musculus animals 6-8 weeks in age were euthanized by CO_2 inhalation as recommended by the 2000 Report of the American Veterinary Medical Association Panel on Euthanasia and approved by our university's Institution's IACUC. In addition, cervical dislocation will be used as a secondary physical method of euthanasia. We then dissected out their spleens, dissociated the tissue using a syringe plunger as mortar and a cell strainer, and

isolated B lymphocytes by negative selection using CD43 beads Miltenyi Biotec. To allow for the inversion/deletion RNA exosome abrogation reaction to occur *ROSA*$^{Cre\text{-}ERT2/+}$*; EXOSC3*$^{COIN/COIN}$ ES cells, B lymphocytes, and their respective controls were exposed to 100 nM 4-hydroxytamoxifen (from Sigma) in the appropriate cell culture conditions for 24 hours, after which the 4-OHT was washed away, and CSR medium was added for B cells, or regular tissue culture growth medium is added for ES cells. B cells must receive proliferation signal, lest they die; therefore during the inversion-deletion reaction, LPS is added to the growth medium at 20 $\mu g.mL^{-1}$.

II.2.2. Bacterial Artificial Chromosome (BAC)-Fluorescence In-Situ Hybridization

The BAC-FISH assay starts when cells are harvested (after having been metaphase-arrested by colcemid treatment, as in our case, or not). The metaphases are then fixed dropped on glass sides. At this point, labeled BAC-DNA containing the DNA sequence homologous to the target of hybridization can be added to the glass side. The DNA labeling is done with biotin or digoxigenin by nick translation: 1 µg of template BAC DNA and 4 µL of nick translation mix are briefly centrifuged and incubated at 15˚C; the reaction is then stopped by heating up to 85˚C and adding EDTA. Once the labeled BAC DNA probes are added to the glass side, denaturation from dsDNA to ssDNA is performed, as is the probe DNA-Genome DNA hybridization. Digoxigenin is then detectable by FITC-anti-dig fragment, and biotin by Cy5-streptavidin. We used two BAC-DNA sequences flanking the IgH locus: BAC 207 (RP 22-207i23, coordinates are 12: 116023270 - 116172952) is more telomeric, and slightly 5' of the variable region of *IgH*, and BAC 199 (CT7-199m11, coordinates 12: 113129176 - 113255643) is more centromeric and spans the C_α-3'RR region of *IgH*.

II.2.3. Cell Culture Conditions

All cells were incubated at 37°C in a 5% CO_2 humidified incubator. Primary B cells were grown in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 15% Fetal Bovine Serum (FBS) and 1X Non-Essential Amino Acids, 1mM Sodium Pyruvate, 10mM 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES), Penicillin-Streptomycin (100 $U.mL^{-1}$ Penicillin and 100 $\mu g.mL^{-1}$ Streptomycin), 55 μM β-mercaptoethanol. CH12F3 and Ramos cells were grown in similar medium with the addition of 2mM L-glutamine. Human Embryonic Kidney 293 T cells were cultured in Dubelcco's Modified Eagle Medium (DMEM) supplemented with 1X Non-Essential Amino Acids, 1mM Sodium Pyruvate, 10mM 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES), Penicillin-Streptomycin (100 $U.mL^{-1}$ Penicillin and 100 $\mu g.mL^{-1}$ Streptomycin), 55 μM β-mercaptoethanol. ES cells were grown in Dubelcco's Modified Eagle Medium (DMEM) supplemented with 15% Fetal Bovine Serum (FBS), 4mM L-glutamine, 1X Non-Essential Amino Acids, 1mM Sodium Pyruvate, Penicillin-Streptomycin (100 $U.mL^{-1}$ Penicillin and 100 $\mu g.mL^{-1}$ Streptomycin), 110 μM β-mercaptoethanol, and 120 μL ESGRO (Leukemia Inhibitory Factor, LIF).

II.2.4. Class-Switch Recombination Assay

For primary B cells, CSR to IgG1 was initiated by adding 20 $\mu g.mL^{-1}$ LPS and IL-4 (from R&D Systems) at 20 $ng.mL^{-1}$ for 72 hrs (or 48 hrs if the inversion-deletion reaction was taking place in the presence of LPS for 24 hrs). For CH12F3 cells, CSR to IgA was initiated by adding 20 $\mu g.mL^{-1}$ LPS, 20 $ng.mL^{-1}$ IL-4, and 1 $ng.mL^{-1}$. After culture, cells were placed in 2.5% FBS in 1X PBS, stained with the appropriate conjugated antibodies (FITC-conjugated anti- mouse IgA,

IgG1, CD19, anti-Rat IgG). We acquired data on a FACSAria cell sorter (BD Biosciences) and used FlowJo software for analysis.

II.2.5. Immunoprecipitation and Immunoblotting Assay

Cells are lysed by douncing in buffer A (20 mM Tris, pH7.5, 1mM DTT, 10mM $MgCl_2$) plus 10mM NaCl, 10% glycerol and protease inhibitors; an equal volume of buffer A + 1 M NaCl +10% glycerol and protease inhibitors is added, followed by sonication and douncing. The extract is dialyzed against buffer A + 100 mM NaCl and centrifuged to remove debris. Cell lysates are pre-cleared with IgG-conjugated Pierce protein A/G agarose beads (Life Technologies) for at least 2 hours and then incubated with anti-human SETX antibody (Santa Cruz Biotechnology) plus protein A/G agarose beads for more than 6 hours, followed by several washes with cell lysis buffer plus 10% glycerol. The immunoprecipitated beads are boiled in protein loading buffer for SDS-PAGE. After the gel is run and the transfer to a PVDF membrane is completed, immunoblotting is performed with the appropriate antibodies. We used anti-RRP45 (Novus), anti-AID, anti-human SETX (Santa Cruz).

II.2.6. Electroporation of Cells

We used an Amaxa Nucleofactor II electroporator device from Lonza, and followed the manufacturer's instructions. Depending on the plasmid size, 2-4 x 10^6 cells to be electroporated are suspended in an electroporation buffer solution (100 µL) containing the desired plasmid and placed in a cuvette in which the reaction will take place. After electroporation has occurred the solution 500 µL of 37 ˚C-medium are added to the cuvette, and the mixture is moved to 1 mL of growth medium pre-warmed to 37˚C.

II.2.7. Immunofluorescence Assay

Cells are placed on coverslips pre-treated with poly-L-lysine for 1 hr at room temperature and allowed to incubate for 60 min at room temperature or at 37˚C. The cells are then fixed using 3-4% paraformaldehyde (PFA) at room temperature for 20 min. To quench autofluorescence, the coverslips are placed in 50 mM NH_4Cl. They are then permeabilized with 0.1 % Triton X-100 for no longer than 1 min, before undergoing blocking (with 0.25% fish skin gelatin and 0.01% saponin in PBS) for 30 min. The appropriate primary antibody can then be used. In our case, we used the S9.6 anti-RNA:DNA hybrid antibody at 1:250 (Kerafast) (Boguslawski et al., 1986, Ginno et al., 2012). As secondary antibody, we used Alexa 488-conjugated Goat anti-Mouse at 1:300 (Sigma). In another instance we used anti-γH2AX and anti-H2AX (Novus, Bethyl Laboratories) with their appropriate fluorophore-tagged secondary antibody. Images were captured with an EVOS digital inverted fluorescent microscope, and analyzed with the ImageJ software from the National Institutes of Health. For radiation-dependent experiments, cells were irradiated using an Atomic Energy of Canada Gammacell 40 Cesium irradiator at ~82 $rad.min^{-1}$ = 0.82 $Gy.min^{-1}$ = 0.82 $J.kg^{-1}$.

II.2.8. RNA preparation and Real-Time quantitative PCR

Total RNA was isolated from cells using Trizol reagent (Life Technologies). RNA was resuspended in water and quantified using a Nanovue Plus spectrophotometer (GE Healthcare Life Sciences). RNA samples were treated with DNase I (Turbo DNA-free kit, Life Technologies), eluted in water and re-quantified. 1.5 µg of RNA were then converted to cDNA using random hexamers and the Superscript III First-Strand Synthesis System for RT–PCR (Life

Technologies). We used two *SETX* primer pairs for real-time quantitative PCR reactions. These two pairs are:

Primer Pair 1: Forward: 5'GTGGGTCTTCCACCATTGATG3'

Reverse: 5'TGCTACTGCTTAGAGTGTGTGG3'

Primer Pair 2: Forward: 5'AAAATTAGCGCAGAGAAGTCTGG3'

Reverse: 5'TCATTAAAGGATGGCTCTGTTGG3'

We also used primers for the germline transcripts at switch sequences μ and α, *GAPDH* housekeeping control, and *AICDA*.

For μ: 5′-CTCTGGCCCTGCTTATTGTTG-3′

and 5′-GAAGACATTTGGGAAGGACTGACT-3′

For α: 5′-CCTGGCTGTTCCCCTATGAA-3′

and 5′-GAGCTGGTGGGAGTGTCAGTG-3′

For *GAPDH*: 5′-TGGCCTTCCGTGTTCCTAC-3′

and 5′-GAGTTGCTGTTGAAGTCGCA-3′

For *AICDA*: 5′-GGAACAGCAGAACTTCCAGACTTTG-3′

and 5′-CCTGAAAGTGAGCCTTAGAGGGAA-3′

Real-Time PCR was performed using the Light Cycler 480 II system by Roche Applied Science, using SYBR Green. Analysis was performed using the Δ-ΔCT method.

II.2.9. Spectral Karyotyping

Spectral karyotyping assay was outsourced; a brief description of the steps follows. Cells on glass slides are treated with trypsin and dehydrated with ethanol washes. Their chromosomes are heat-denatured for no longer than 1.5 min and dehydrated again. Spectral karyotyping reagent aliquot is well mixed, denatured, and added to the denatured chromosome preparation; a cover slip is added for overnight incubation at 37˚C, removed the next day, washed, sequentially stained, and placed under coverglass. Spectral imaging software can then be used to analyze the data.

II.2.10. Statistical Analysis

ANOVA and unpaired, two-tailed student t tests were used when possible. We used the software Prism for all statistical analysis calculations. Standard error of the mean was used whenever error bars are displayed.

II.2.11. Telomere Fluorescence In-Situ Hybridization

Metaphase arrest is produced by cell exposure to N-desacetyl-N-methylocolchicine (also known as colcemid, at 200 $ng.mL^{-1}$ for 2-4hrs). The cells are then spun and resuspended in prewarmed hypotonic solution (75 mM KCl) and incubated at 37˚C for 15 min. Fixative solution (3:1 methanol:glacial acetic acid) is added drop-wise. The mixture is spun and the pellet resuspended in fixative solution, and incubated at 4˚C for 15 min. The mixture is spun, the pellet resuspended in fixative solution, and allowed to incubate at 4˚C for 15 min once more. This is done a third time, after which the pellet, resuspended in fixative solution, can be dropped on glass slides that have been “hydrated” by being passed over water vapor emanating from a water at 80-90˚C (with side to be dropped on facing the vapor). Once the metaphases have been

dropped, the slide is passed over the water vapor again, on both sides, for 3-5 seconds at a time, then placed on a plate pre-heated to ~55-75 ˚C so as to dry the slide without burning the sample. Slow drying has tended to be better. The chromosomes can be fully visualized at this stage. The slides are left overnight, and the next day are washed with PBS, then fixed in 4% formaldehyde/ PBS, washed again in PBS, digested with acidified pepsin, washed and fixed again, and dehydrated with serial ethanol washes, before being allowed to air dry. The hybridization step follows and consists in adding the prepared fluorochrome-labelled telomeric peptide nucleic acid (PNA) probe to the slide and denaturing at 80˚C for 3 min. The probe mix contains 10 mM Tris at pH 7.2, buffer MgCl2, 70% deionized formatted, 0.25 % Boehringer Mannheim blocking regent, the hybridization probe (at 0.5 $\mu g.mL^{-1}$) in water. After the denaturation step, the slide is allowed to incubate with the probe for 2 hours at room temperature in a humified chamber, at the conclusion of which formamide washes are performed (70% formamide, 10mM Tris, 0.1% BSA in water), followed by PBS/Tween washes, serial dehydration ethanol washes, and air drying. The slides are covered with Vectashield with DAPI mounting medium (from Vector Laboratories), and sealed with nail polish, before being stored in the dark. Images were captured with an EVOS digital inverted fluorescent microscope, and analyzed with the ImageJ software from the National Institutes of Health. Earliest T-FISH assays were conducted in collaboration with the Cancer Center.

II.2.12. VPD450 Proliferation Assay

Washed cells are labelled by adding the prepared VPD450 dye from BD Bioscience to a concentration of 1 μM and allowing the cells to incubate in a 37˚C water bath for 10-15 minutes. The cells are then thoroughly washed and resuspended in appropriate growth medium, and can

be analyzed by flow cytometry. As with the CSR assay, we used FlowJo software to process our flow cytometry data.

II.2.13. 5' S_μ Sequencing

The 5' subregion of Sμ was amplified by PCR with the following primers:

Forward: 5'-AATGGATACCTCAGTGGTTTTTAATG-3'

Reverse: 5'-GCGGCCCGGCTCATTCCAGTTCATTA-3'

We used Turbo Pfu as follows: 28 cycles of 94°C (30 s), 58°C (30 s), and 72°C (40 s). The PCR products were then incubated with GoTaq DNA polymerase at 72°C for 10min and purified with a gel extraction kit (Qiagen) after electrophoresis in agarose gel. Purified PCR product was ligated with pGEM-T Easy vector and electroporated into XL-1 blue competent cells. White colonies were selected for minipreps and sequenced with T7 primer. Sequences were aligned with ClustalX.

Chapter III: Results

III.1. Genomic Instability Increase in RNA Exosome-Deficient B Cells

III.1.1. CSR Defects in RNA Exosome-Deficient B cells

We wanted to investigate the extent to which impairing RNA exosome complex physiological function during the immunoglobulin R-loop reaction impacts B cell genomic integrity. The presence — and importance — of R loops to isotype switching by class-switch recombination (CSR) has increasingly been suspected and supported, culminating in the identification of the RNA exosome complex as an activity and targeting cofactor of AID deamination of deoxycytidine residues on both the looped-out non-transcribed DNA strand and the RNA-hidden template DNA strand at immunoglobulin switch sequence AID targets (Basu et al., 2011, Sun et al., 2013a, Pefanis et al., 2014, Pefanis et al., 2015). The interest in these RNA:DNA heteroduplex-containing R loop structures is augmented by the implication of RNA:DNA hybrids in pathophysiological processes, namely genomic instability (Wahba et al.,

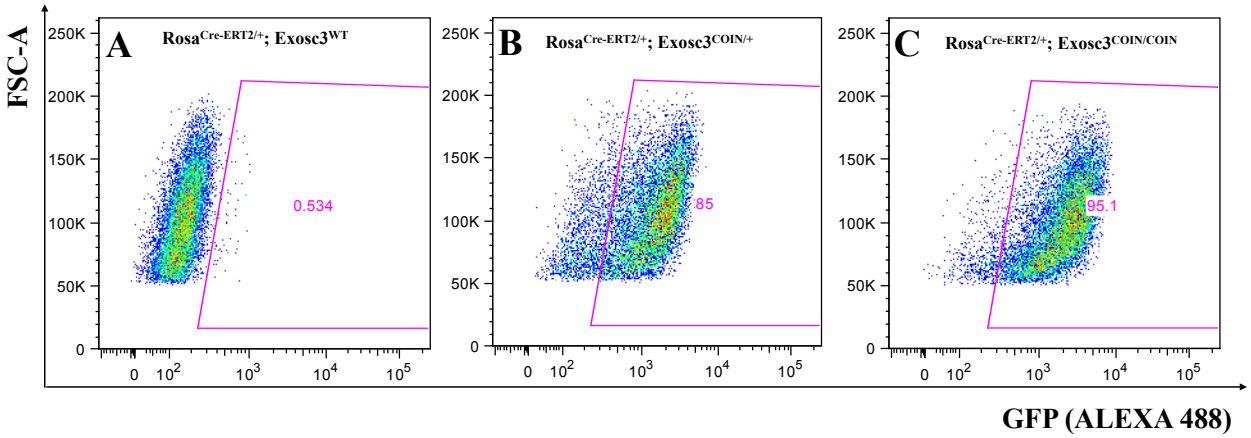


Figure 3.1. Effectiveness of inversion-deletion reaction. Exposure to 4-hydroxytamoxifen results in the nuclear translocation of the ubiquitously expressed recombined Cre-ERT2 protein where it can catalyze the inversion-deletion reaction of the *EXOSC3*$^{COIN/COIN}$ allele and the expression of a reporter *GFP* cassette. When both *EXOSC3* alleles are wild-type, no GFP is detected by FACS (A). In the heterozygous background, GFP is detected in the vast majority of cells (B). In the homozygous *EXOSC3*$^{COIN/COIN}$ background, close to 100% deletion is achieved.

2011). Transcription termination is hypothesized to be a required process at canonical AID targets for the desired, physiological immunoglobulin genomic alterations to take place before productive, full-length transcription of Ig loci (Storb, 2014). We set out to elucidate what the consequences of perturbing this process would be by abrogating RNA exosome activity, which had been described in the processing and/or degradation of many RNA species (Januszyk and Lima, 2014).

We utilized an inducible system to delete the gene *EXOSC3* in mouse, which encodes RRP40, one of the nine core, essential subunits of the RNA exosome complex with a S1 and K-homology RNA-binding domain (Pefanis et al., 2014, Pefanis et al., 2015). We reasoned that in the absence of appropriate levels of the different RNA exosome complex core components, the assembly of the 11 subunit RNA exosome complex would not proceed properly, and the activity of the complex would be abrogated. The deletion system employed allows the nuclear translocation, conditioned upon exposure to 4-hydroxy tamoxifen (4-OHT), of the constitutive and ubiquitous expression product of a *Cre-ERT2* recombinase allele introduced at the *ROSA26* locus. In the nucleus, the recombined recombinase recognizes *lox* sites flanking the *EXOSC3* allele, and proceeds to delete it via an inversion-deletion reaction that results in the expression of a *GFP* cassette. First, to ensure that the expression of the *GFP* cassette in splenic B lymphocytes exposed to 4-OHT for 24 hrs (along with lipopolysaccharide, LPS, to prevent cell death) proceeded as planned, we analyzed these cells by fluorescence-activated cell sorting (FACS, Figure 3.1). We use the detection of GFP by FACS as a substitute for successful *EXOSC3* allelic inversion-deletion. We observe that $ROSA26^{Cre\text{-}ERT2/+}$; $EXOSC3^{WT}$ cells do not have detectable levels of GFP, while the $ROSA26^{Cre\text{-}ERT2/+}$; $EXOSC3^{COIN/+}$ and the $ROSA26^{Cre\text{-}ERT2/+}$;

EXOSC3$^{COIN/COIN}$ do, with the *ROSA26*$^{Cre\text{-}ERT2/+}$*; EXOSC3*$^{COIN/COIN}$ cells having levels that are slightly higher. The high detection levels of GFP in both *EXOSC3*$^{COIN/+}$ and *EXOSC3*$^{COIN/COIN}$ backgrounds indicate that the inversion-deletion reaction proceeded with success.

As previously reported, we observed that the ability of isolated RNA exosome-deficient splenic B lymphocytes to engage in isotype switching from IgM to IgG1 is severely impaired in contrast to RNA exosome-sufficient *ROSA26*$^{Cre\text{-}ERT2/+}$; *EXOSC3*$^{COIN/+}$ cells, indicating that the Ig R loop reaction may be negatively impacted by RNA exosome functional abrogation (Figure 3.2) (Basu et al., 2011, Pefanis et al., 2014). However, a fraction of B lymphocytes thought to have abrogated RNA exosomse activity can still remodel their Ig loci to produce the IgG1 isotype (Figure 3.2).

III.1.2. Structural Catastrophic Events in Stimulated RNA Exosome-Deficient B cells

We then set out to find out whether the abrogation of RNA exosome function affected the genomic integrity of the B cells lacking the complex. Numerous accounts have suggested that the accumulation of RNA:DNA species is detrimental to cellular health by generating genomic instability (Huertas and Aguilera, 2003, Tuduri et al., 2009, Wahba et al., 2011). These RNA:DNA species may accumulate as a result of deficient RNA processing mechanisms that increase the relative amounts of RNA available for base-pairing with negatively supercoiled DNA, either through co-transcriptional defects that may result in increased RNA:DNA hybrids formed *in cis*, or through post-transcriptional defects that may result in increased RNA:DNA hybrids formed perhaps *in trans* at a different DNA locus sharing some homologue (Kim and

Jinks-Robertson, 2012, Wahba et al., 2013, Costantino and Koshland, 2015). Given the implication in yeast of catalytic RNA exosome subunit Rrp6 in the genesis of genomic instability, we decided to investigate the consequences of RNA exosome activity abrogation in

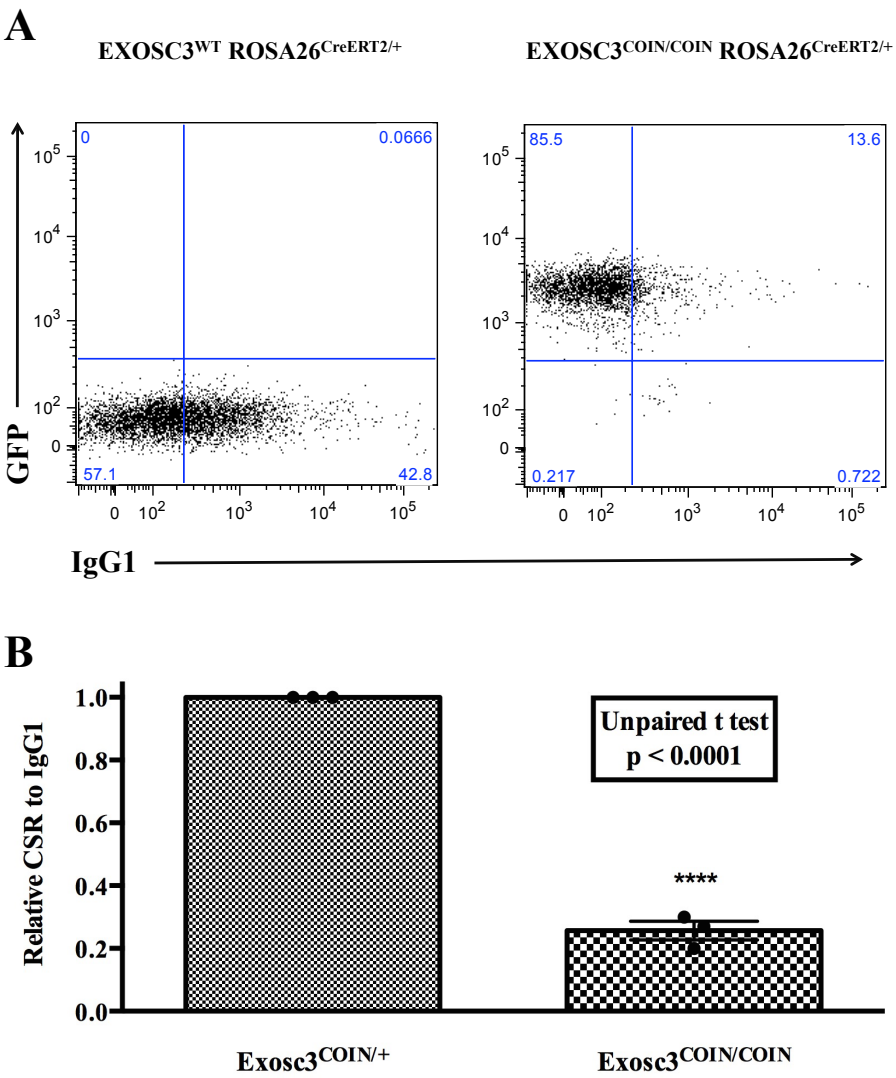


Figure 3.2. CSR in exosome-deficient mouse B cells. *Ex vivo* 4-OHT-exposed RNA exosome-deficient B lymphocytes exhibit a defect in their ability to complete isotype switching to IgG_1 upon stimulation with IL-4 and LPS. A. *ROSA26*$^{Cre\text{-}ERT2/+}$; *EXOSC3*$^{COIN/COIN}$ cells have high levels of GFP and low levels of IgG_1, while a substantial portion of cells wild-type for *EXOSC3* are IgG_1^+. B. Relative RNA Exosome-competent backgrounds, RNA Exosome-deficient cells exhibit stunted ability to switch to IgG_1. 3 animals were used in this analysis (Pefanis et al., 2014).

mouse B cells by performing a telomeric fluorescence in situ hybridization (T-FISH) assay on them (Wahba et al., 2011).

We thought that a T-FISH assay would allow us to detect general structural abnormalities in chromosomal integrity in the absence of the key RNA processing machinery RNA exosome. We isolated splenic B lymphocytes from mice with different susceptibilities to the CreERT2-mediated inversion-deletion reaction, exposed them to 4-OHT, induced them to undergo IgM-to-IgG1 class-switch recombination by stimulating them with an LPS-interleukin 4 (IL-4) cocktail for 3 days, and caused them to metaphase-arrest by providing them with colcemid (a less toxic relative of the microtubule-disruptive compound colchicine) that prevents metaphase-to-anaphase passage (Rieder and Palazzo, 1992). We then performed a number of washes and fixed the cells, before "dropping" these metaphase preparations on glass slides, where a fluorophore-tagged telomeric probe was allowed to hybridize to the genomic material. When compared to RNA exosome-sufficient backgrounds, *ex vivo* CSR-activated RNA exosome-deficient B lymphocytes displayed a marked increase in the proportion of metaphases containing genomic instability. We noticed a number of changes, including the loss of telomeres (both at centromeric and telomeric chromosomal ends), the presence of acentric fragments (which simply denotes the absence of a centromere in this DNA specimen), the fusion of two chromosomes into one dicentric chromosome, the end-to-end fusion of sister chromatids intra-chromosomally, and, rarely, the formation of so-called radials (Figures 3.3A and 3.3B). Radials are chromatid-type errors thought to occur at DNA polymerase stalling sites, where a chromatid break is repaired by the insertion of another chromosome (which also bears a chromatid break in the case of quadriradials) (Savage, 1976, Scully et al., 2000, Satoh et al., 2002). For all subtypes of genomic

instability, more catastrophic events occurred in the RNA exosome-deficient B cells than in RNA

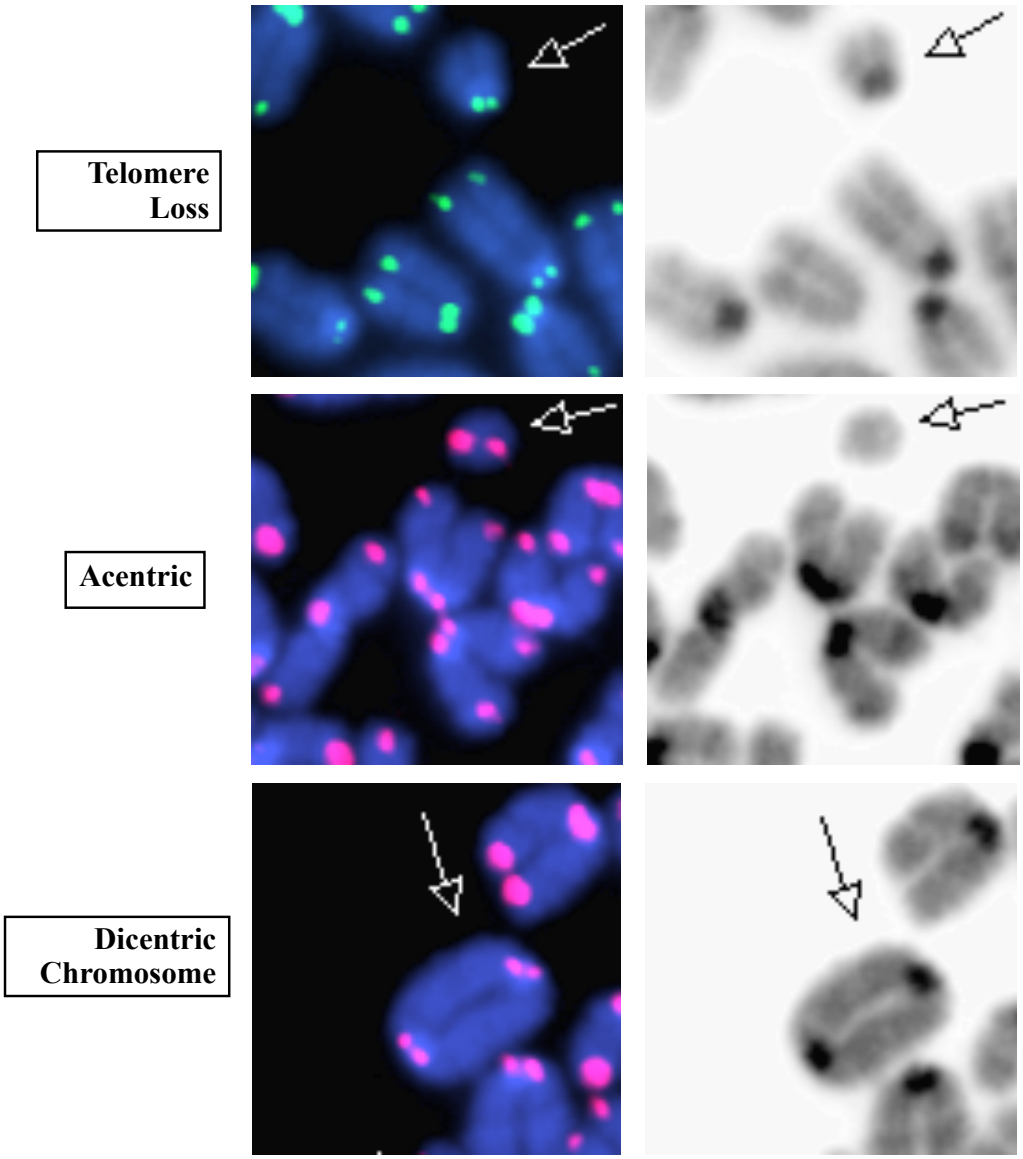


Figure 3.3A. Structural abnormalities in RNA exosome-deficient stimulated B cells. Examples are shown of telomeric telomere loss, acentric chromosomal fragments, and dicentric chromosomes. These abnormal chromosomal structures accrue at a higher degree in activated *ROSA26*$^{Cre\text{-}ERT2/+}$*; EXOSC3*$^{COIN/COIN}$ B lymphocytes than in other their RNA exosome-competent counterparts. Structures are shown with fluorophore and 4',6-diamidino-2-phenylindole (DAPI) staining (on the left) and with transmitted light (on the right). Densities are indicative of centromeres. Chromosomal fragments without centromeres are acentric (middle row), while chromosomal structures with 2 centromeres are dicentric (bottom row). Fluorophore-labeled telomeric probes denote the presence or absence of telomeres (centromeric or telomeric telomeres; seen here is telomeric telomere loss in the top row).

exosome-sufficient B cells (Figure 3.4). Breaks and fragments, which are tallied from

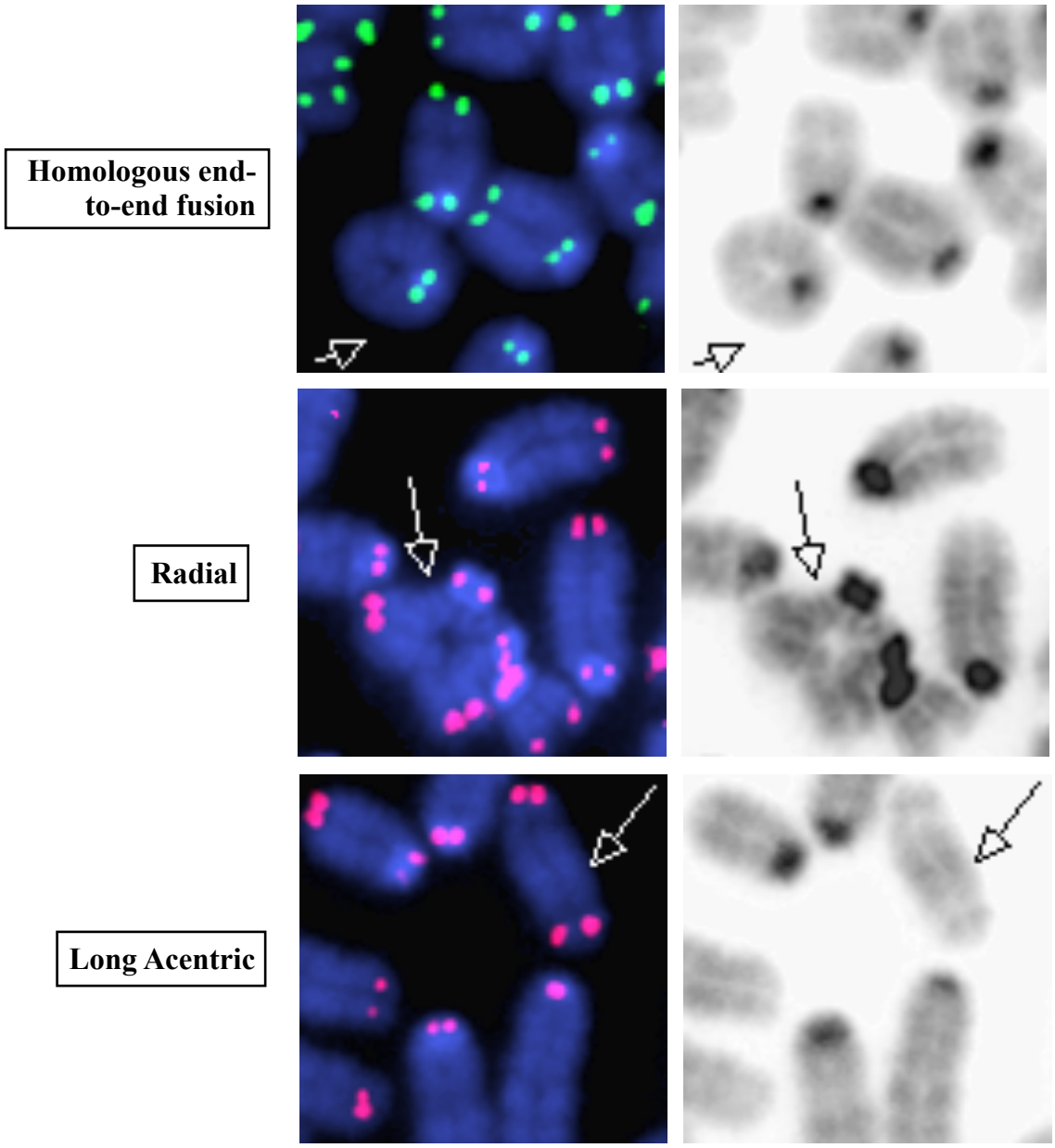


Figure 3.3B. Structural abnormalities in RNA exosome-deficient Stimulated B Cells. Additional representations are shown for homologous end-to-end fusion, radial, and long acentric chromosomes. These abnormal chromosomal structures also accrue at a higher degree in activated *ROSA26*$^{Cre\text{-}ERT2/+}$*; EXOSC3*$^{COIN/COIN}$ B lymphocytes than in other their RNA exosome-competent counterparts. Note the chromatid fusion and absence of telomeric telomeres in the top picture pair; the middle pair shows a quadriradial chromosome; the bottom pair shows a long acentric chromosome: while in 2 mouse chromosomes, lack of a centromere is expected, this chromosome is excessively long to be one of those.

metaphases with telomeric loss (whether telomeric telomeres or centromeric telomeres) and from metaphases with acentric chromosomes, vastly account for the difference observed in the proportions of metaphases with instability between RNA exosome-deficient and RNA exosome-sufficient backgrounds (Figure 3.4).

The increase in catastrophic events in the absence of physiological RNA exosome activity was striking. Still, we thought that our assay restricted our analysis only to those genomic

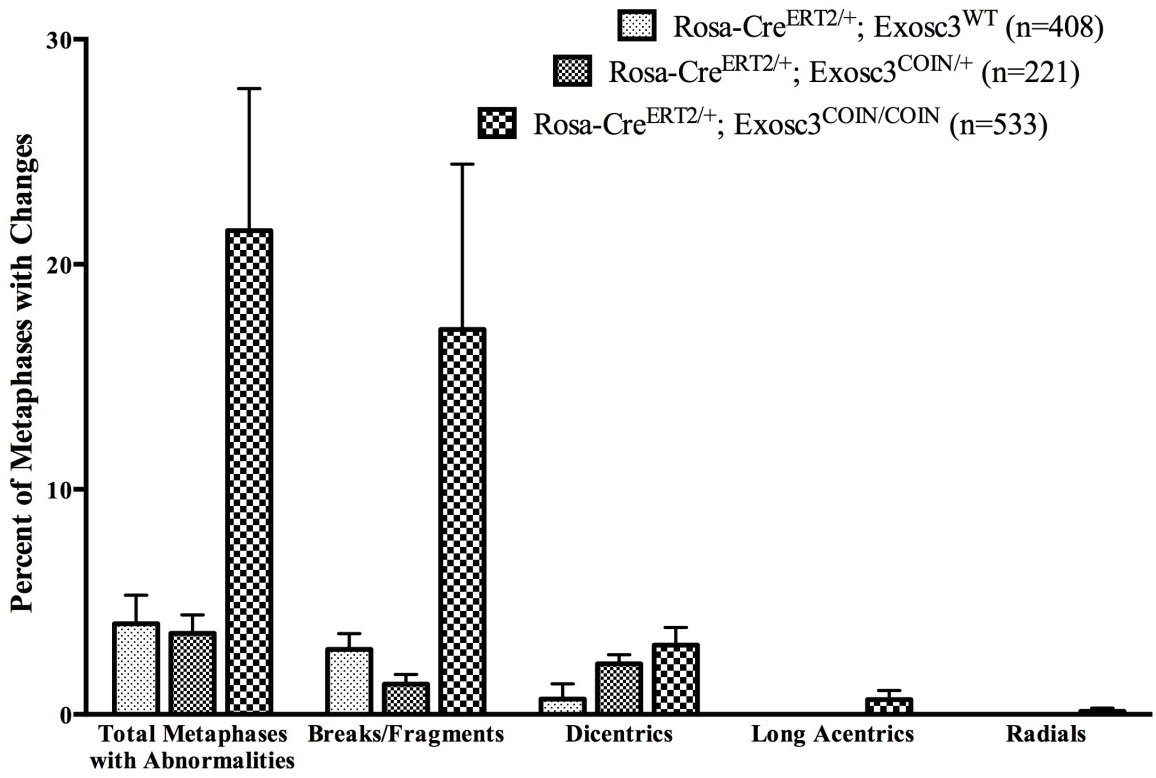


Figure 3.4. Breakdown of chromosomal instability events. A compilation of the different phenotypes is plotted in this bar graph for the three backgrounds. For all abnormalities, we observed a statistically significant difference between the RNA exosome-deficient B cells and the RNA exosome-competent backgrounds: ANOVA, $p = 0.0409$, $R^2 = 0.5989$. 4 animals were used for the WT background, 2 for the heterozygous background, and 4 for the *ROSA26*$^{Cre\text{-}ERT2/+}$; *EXOSC3*$^{COIN/COIN}$ background. The respective number of metaphases analyzed is indicated for each mouse background.

instability instances that resulted in chromosomal structural changes perceptible by T-FISH, namely, improperly repaired chromosomal breaks, dicentric or acentric chromosomes, and intra-chromosomal chromatid end-to-end fusion events. Were a break to occur and be successfully repaired, our assay could not enable us to recognize such an accidental occurrence; similarly, were a balanced, pathophysiological translocation event to take place, our T-FISH assay would not be able to reveal that disease-predisposing genetic material rearrangement. We would not able to detect any changes that may have occurred more proximal to the presumably RNA exosome absence-related insult, which may have been repaired by the *a priori* intact cellular repair mechanism. We therefore set out to increase the stress encountered by these B cells by exposing them to ionizing radiation and assaying for instability.

III.1.3. Genomic Instability in IR-Exposed Stimulated RNA Exosome-Deficient B cells

As we did before, we isolated splenic B lymphocytes and grew them *ex vivo* in the presence of LPS and 4-OHT for the first 24 hours to abrogate RNA exosome activity, followed by an LPS and IL-4 treatment for an added 48 hours to allow for CSR to occur. We then irradiated the cells for different periods: 0 min, 2.5 min, and 24.5 min, meaning that the cells were correspondingly exposed to 0 Gy, 2 Gy, and 20 Gy of irradiation. We had initially set out to perform a T-FISH assay on these cells; but multiple attempts yielded no usable metaphase spreads, perhaps because the exposure to colcemid in addition to irradiation proved excessively toxic for the cells. In parallel, we had wanted to perform an immunofluorescence (IF) assay on these cells using antibodies that recognize proteins involved in the repair of double-stranded DNA breaks. To do this, we first allowed irradiated cells to recover for 45 min, after which we

washed, fixed, and prepared them for IF. We settled on 45 min because after longer recovery periods (5.5 hrs and 10.5 hrs), assayed DSB repair proteins were mostly no longer detectable by IF, and shorter recovery periods were not logistically consistently achievable, given the number of assay steps and samples involved. We performed IF using antibodies against γH2AX, the phosphorylated form of the histone variant protein H2A and analyzed the data using the software ImageJ. It is thought that upon sensing of double-stranded DNA breaks by the complex formed by MRE11 (meiotic recombination homolog 11), RAD50 (radiation sensitive mutant 50), and NBS1 (Nijmegen-breakage syndrome 1), phosphatidylinositol-3-kinase-like family of protein kinases (which include DNA-dependent protein kinase catalytic subunit (DNA-PK_{cs}), Ataxia telangiectasia mutated (ATM), and ATM- and RAD3-related (ATR)) will be activated and in turn will phosphorylate H2AX at its serine 139 residue to produce γH2AX, which will induce the recruitment of additional DNA repair proteins (Kinner et al., 2008). The aggregation of these phosphorylated γH2AX proteins into foci at or near sites initially recognized by the cell can therefore be used as a biomarker proxy for the presence of a double-stranded DNA break (Kuo and Yang, 2008).

We observe that at 0 Gy, stimulated B cells, RNA exosome-sufficient and -deficient alike, have minimal amounts of γH2AX (Figure 3.5). It is not all too surprising that there should be some amount of γH2AX detection, reflecting the DSBs that are bound to be formed in proliferating cells in general, and stimulated B cells in particular, since DSBs are an intermediate for class-switch recombination. When we evaluate the percentage of stimulated, irradiated, fixed cells that accumulate 3 or more γH2AX foci, we see that at baseline 0 Gy levels very few to no RNA exosome-sufficient B cells accumulate these foci, and that B cells with compromised RNA

exosome activity already exhibit a statistically significant difference (by ANOVA), small though it be (Figure 3.5). Exposing the *ex vivo* activated B lymphocytes to 2 Gy does not seem to drastically change the observations made at baseline (Figure 3.6). The RNA exosome-sufficient *ROSA26*$^{Cre\text{-ERT2/+}}$; *EXOSC3*^{WT} and *ROSA26*$^{Cre\text{-ERT2/+}}$; *EXOSC3*$^{COIN/+}$ backgrounds display around the same amount of γH2AX as detected at 0 Gy in our assay, which is again lower, in a statistically significant manner, than the amount of γH2AX detected in the *ROSA26*$^{Cre\text{-ERT2/+}}$; *EXOSC3*$^{COIN/COIN}$ background. Similarly, when we examine the percentage of activated, 2 Gy-

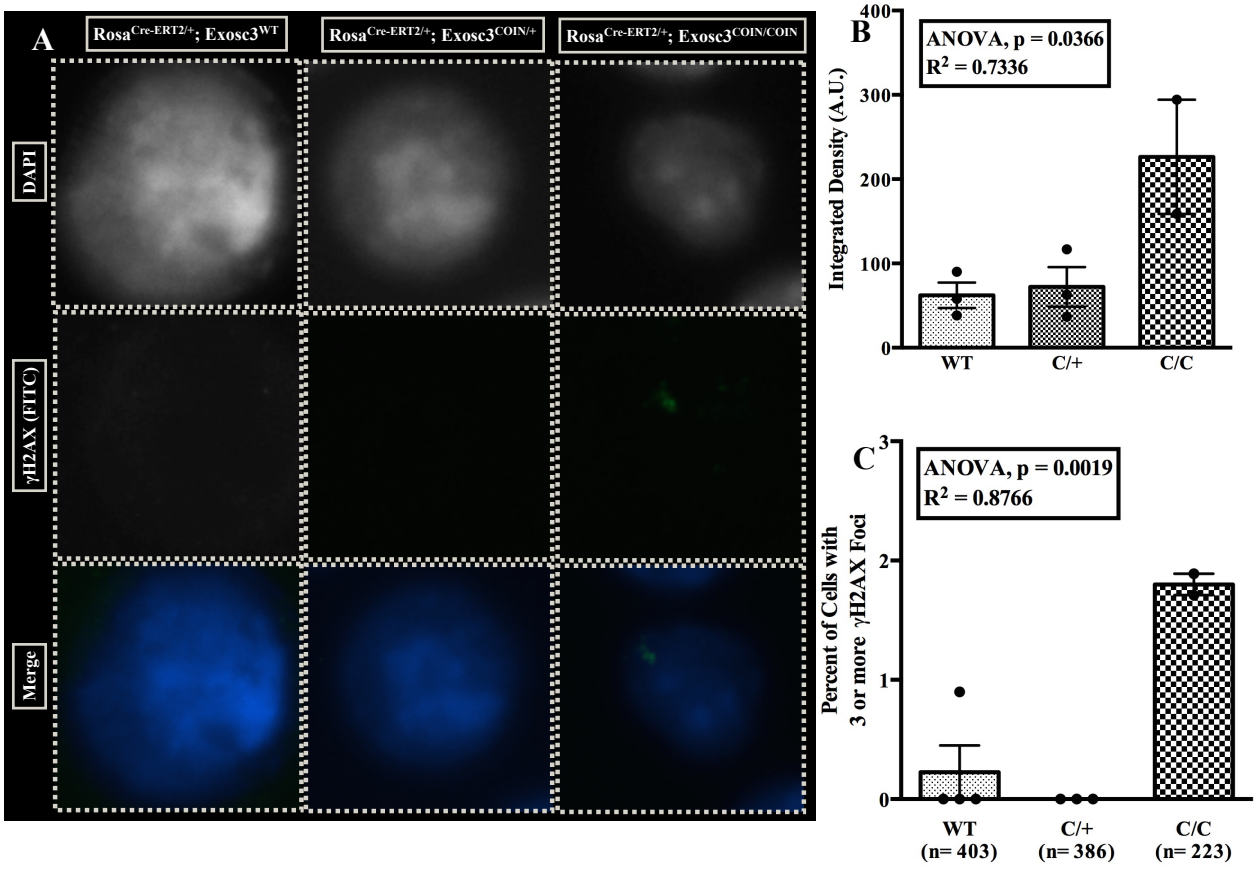


Figure 3.5. γH2AX immunofluorescence assay of 0-Gy irradiated stimulated B cells of different backgrounds. A. IF Images of cells with different backgrounds. B. Low levels of γH2AX are detected by integrated density, though a small, significant difference is apparent among RNA exosome-non-deficient and RNA exosome-deficient backgrounds by ANOVA. By unpaired, two-tailed Student's t-test, p = 0.0561 for WT and C/C, 0.07916 for C/+ and C/C, and 0.738228 for WT and C/+. C. Very few cells accumulate 3 or more γH2AX as detected in our assay. RNA exosome-deficient cells show a statistically significant increase nonetheless.

irradiated, fixed cells that accumulate 3 or more γH2AX foci, RNA exosome-sufficient B cells tend to not accumulate these foci, whereas RNA exosome-deficient B cells exhibit a statistically significant increase, comparable to the baseline difference (Figure 3.6). It appeared that the increase in IR stress did not translate into a higher preponderance of DSBs as measured by γH2AX foci accumulation in our assay.

We then increased the IR stress by exposing the stimulated B cells to 20 Gy, a tenfold change (Figure 3.7). What became apparent at this level of ionizing radiation was that both the

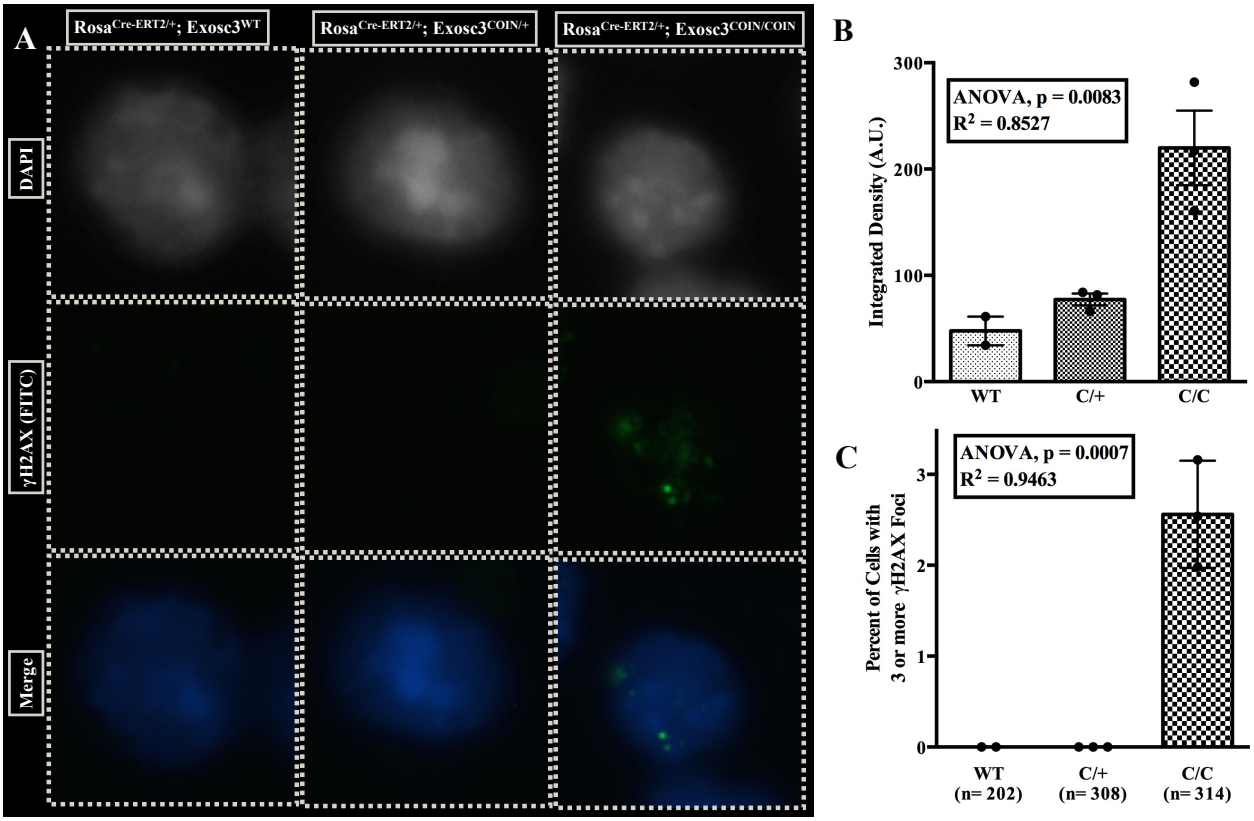


Figure 3.6. γH2AX immunofluorescence assay of 2-Gy irradiated stimulated B cells of different backgrounds. A. IF Images of cells with different backgrounds now start to show γH2AX in the RNA exosome-deficient background. B. The difference between RNA exosome-deprived and RNA exosome-containing backgrounds is slightly more pronounced than at 0 Gy, though still relatively small. Still, a statistically significant difference is perceptible among the different backgrounds. By unpaired, two-tailed Student's t-test, p = 0.0339 for WT and C/C, 0.0159 for C/+ and C/C, and 0.0954 for WT and C/+. C. Cells with 3 γH2AX or more are more frequent in the RNA exosome-deficient primary B cell background.

backgrounds with intact RNA exosome function and B cells with abrogated RNA exosome activity experienced an increase in the amounts of γH2AX detected in our IF assay. Overall γH2AX detection levels increased in all three backgrounds compared to the previous IR exposures, with the RNA exosome-sufficient cells seeing a ~ twofold increase in γH2AX, while RNA exosome-deficient cells had around triple the levels of γH2AX (Figure 3.7). But analyzing the subsets of cells that had 3 or more foci reveals that the increase in the RNA exosome-

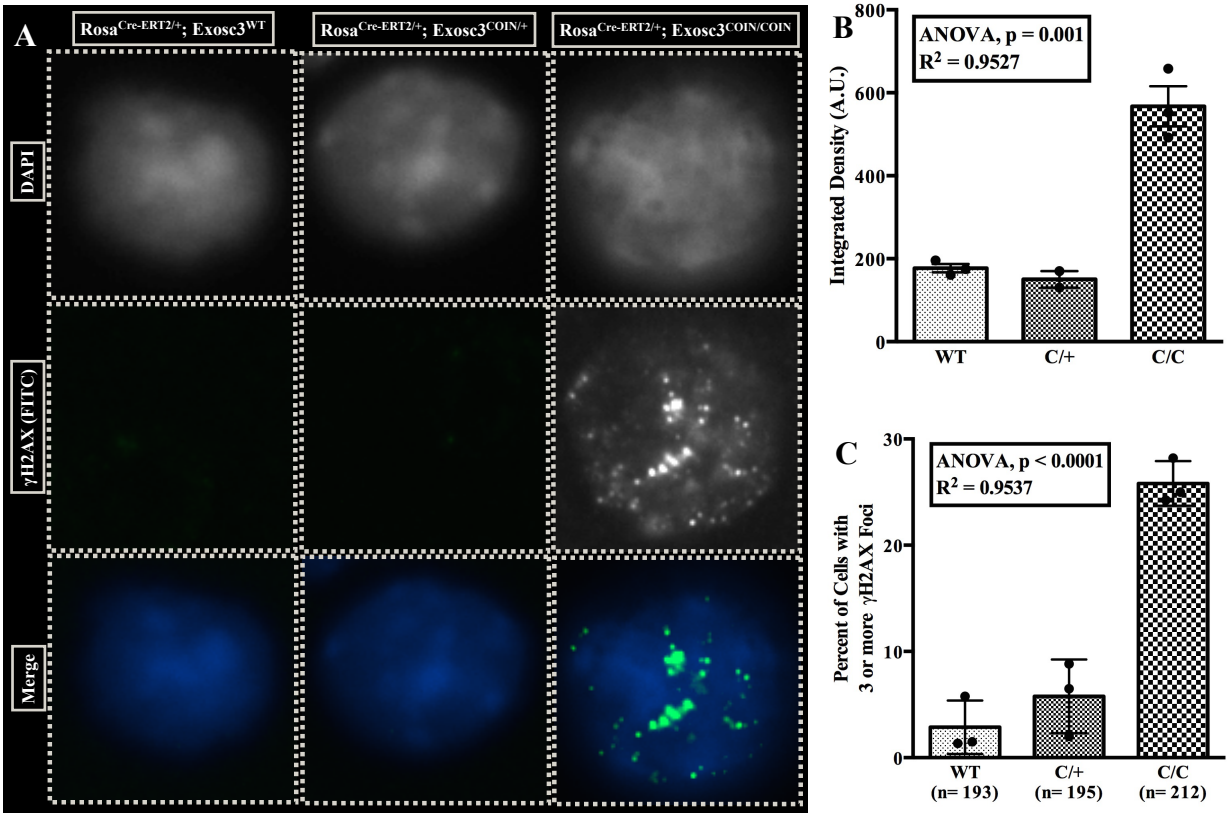


Figure 3.7. γH2AX immunofluorescence assay of 20-Gy irradiated stimulated B cells of different backgrounds. A. IF Images of cells with different backgrounds show massive γH2AX foci accumulation in the RNA exosome-deficient background. B. As IR stress increases to 20 Gy, γH2AX is also detected in cells with proper RNA exosome activity, though at lower levels than in RNA exosome-deficient background. By unpaired, two-tailed Student's t-test, p = 0.0014 for WT and C/C, 0.0013 for C/+ and C/C, and 0.3419 for WT and C/+. C. All cell backgrounds exhibit at least a small percentage of cells with 3 or more γH2AX foci subsequent to irradiation, though that percentage is much larger in the RNA exosome-deficient primary B cells.

deficient background was drastically larger, at about tenfold (Figure 3.7). A small but non-negligible portion of RNA exosome-sufficient backgrounds also displayed an accumulation of γH2AX foci, though much less than the percentage seen in the deficient background. The differences observed at 20 Gy were again statistically significant (Figure 3.7). We ensured that the changes in γH2AX were not due to higher levels of non-phosphorylated H2AX protein in the cells by performing the same IF assay on the various RNA exosome backgrounds using an anti-H2AX antibody after irradiating the cells at 20 Gy (Figure 3.8). We see no statistically significant difference among the three RNA exosome backgrounds, indicating that the changes reported in γH2AX are likely not due to differences in H2AX protein levels. To ease the

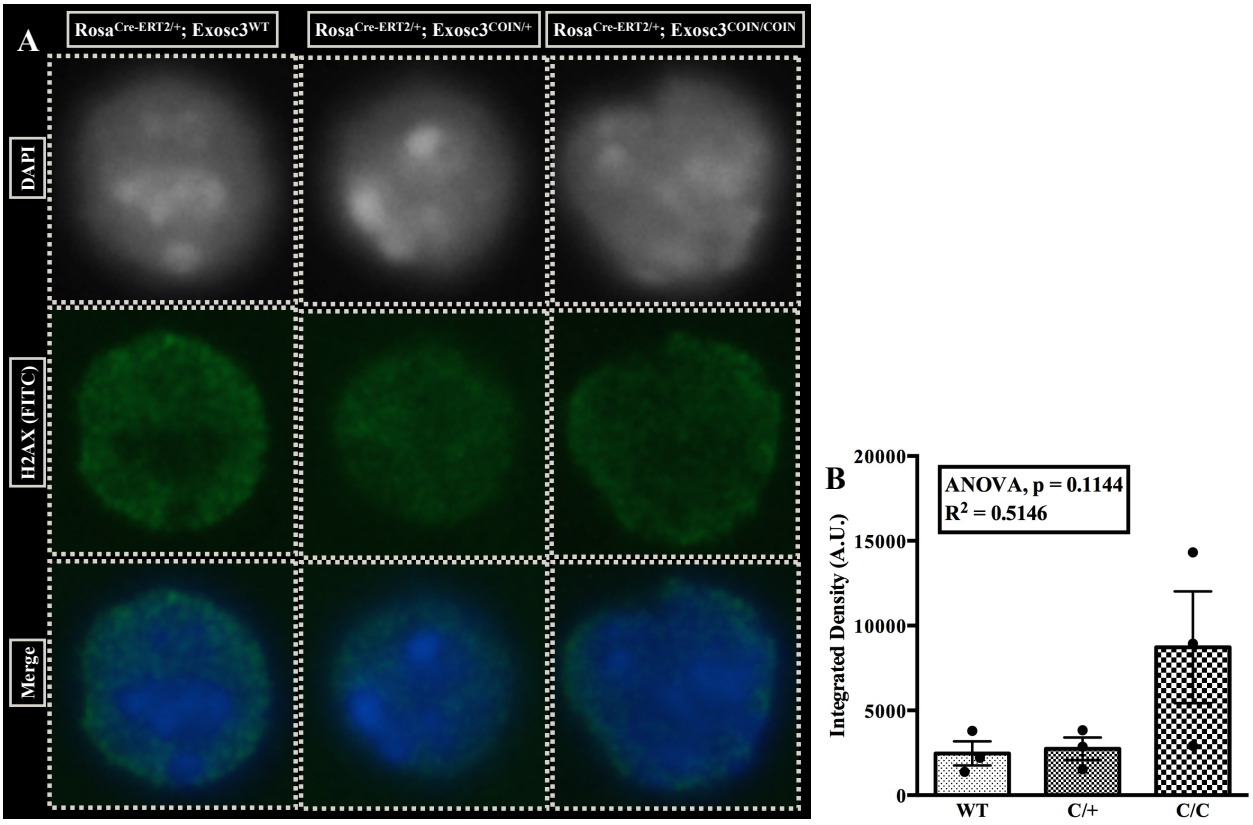


Figure 3.8. Non-phosphorylated H2AX immunofluorescence assay of 20-Gy irradiated stimulated B cells of different backgrounds. IF images of CH12F3 cells with different backgrounds show non-phosphorylated H2AX is readily detected in all three backgrounds (A), and no statistically significant difference among the backgrounds is noted (B) by ANOVA and by unpaired, two-tailed Student's t-test (p = 0.1367 for WT and C/C, 0.1497 for C/+ and C/C, and 0.7888 for WT and C/+)

comparison among levels of detected γH2AX in the different backgrounds, the various measurements were re-plotted on the same graphs (Figure 3.9).

III.1.4. Identity Clues in Chromosome SKY

The T-FISH and IR/IF assays gave us some appreciation of the magnitude of compromise in chromosomal structural integrity that results from the absence of physiological levels of RNA exosome complex in stimulated B lymphocytes. We had wanted to identify the loci involved in these catastrophic events. We set out to perform Giemsa-banding (G-banding), which we thought would provide valuable resolution information in the identification of loci involved in genomic

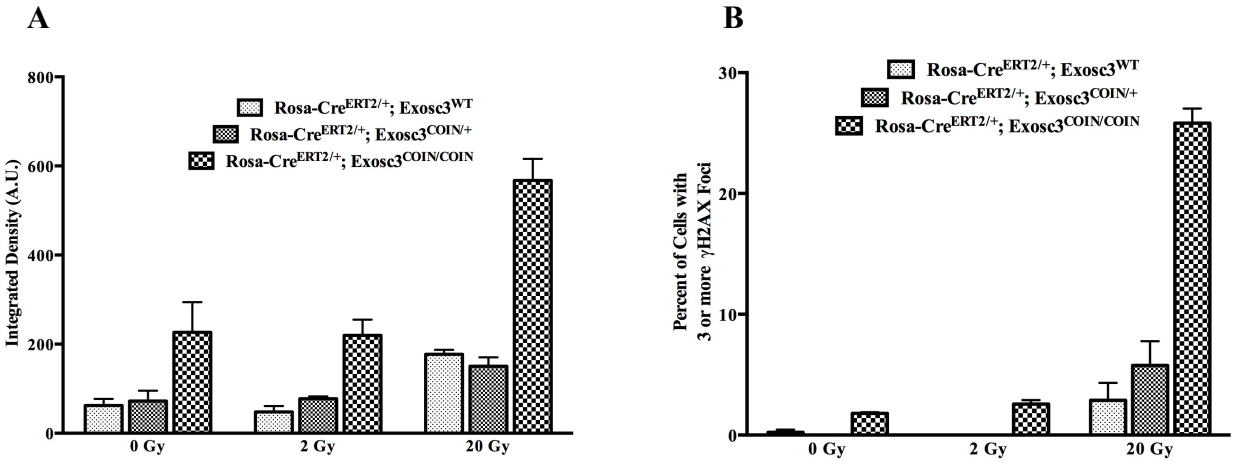


Figure 3.9. Compilation of immunofluorescence/irradiation data. The different RNA exosome backgrounds are plotted for γH2AX accumulation after ionizing radiation at different levels. A. The leftmost plot represents the detected γH2AX intensity in integrated density arbitrary units. B. The rightmost plot shows the percentage of analyzed fixed cells with 3 γH2AX foci or more. Irrespective of the amount of IR, RNA exosome-deficient backgrounds accumulate more double-stranded breaks as assayed by γH2AX IF; at 20 Gy, in conditions of high mutagenic stress, RNA exosome-deficient cells show a dramatic increase in γH2AX accumulation, suggesting they are more sensitive to increases in mutagenic stress.

instability events (Sumner et al., 1971). Ultimately, we opted for a different approach. We thought to perform spectral karyotyping of RNA exosome-deficient activated B cells. Spectral karyotyping (SKY) a ca. 20-year old multicolor FISH-based technology that allows for the hybridization-mediated simultaneous discernment of the different chromosomes and their painting in different colors (Schrock et al., 1996). We obtained SKY results in a CSR-stimulated, RNA exosome-deficient *ROSA26*$^{Cre\text{-}ERT2/+}$; *EXOSC3*$^{COIN/COIN}$ background, compiled in Figure 3.10. What is immediately noticeable is one translocation event involving murine *IGH*-locus bearing chromosome 12. The involvement of chromosome 12 is additional evidence that RNA exosome-mediated instability is possibly partly related to the biology of R loop resolution, notably at Ig loci. We observed that the translocation event resulted in the formation of a dicentric chromosome, with acentric chromosome 12 byproducts also detectable. These observations are reminiscent of some of the results we obtained in our T-FISH screen.

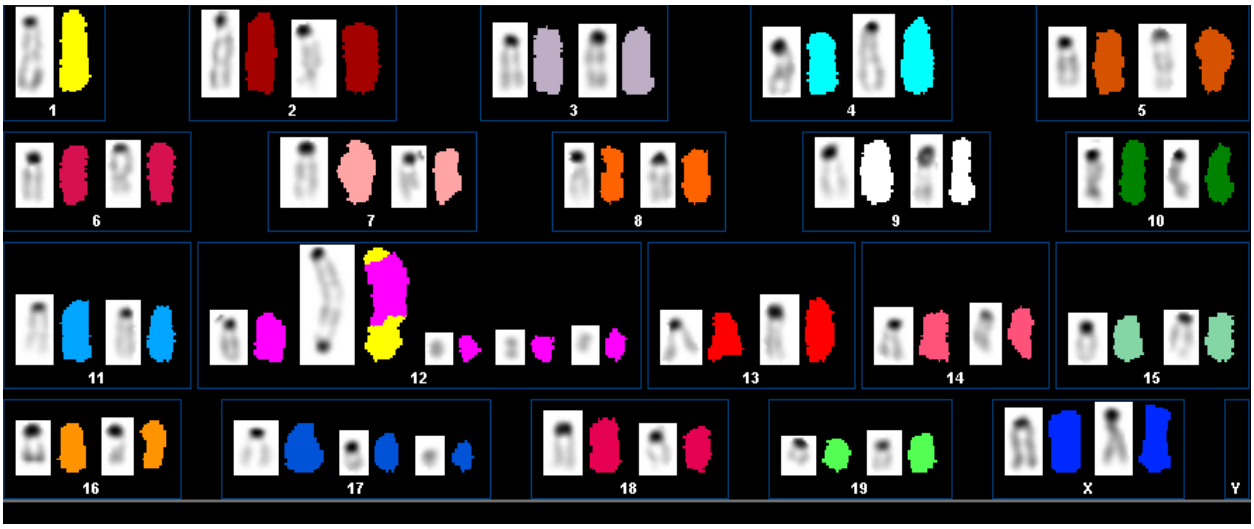


Figure 3.10. Spectral karyotyping. Stimulated RNA exosome-deficient B lymphocytes were metaphase arrested and prepared for spectral karyotyping. Each chromosome is painted in a unique color, allowing for easy chromosomal identification. Of note here is the translocation event involving chromosome 12 and resulting in a dicentric chromosome formed with chromosome 1. Also visible is the fragmentation of chromosome 17.

Interestingly, we noticed that the other chromosome involved in the translocation transaction is chromosome 1, which in mouse bears two genes (*BCL2* and *PTPRC*) associated, respectively, with B-cell cancers or B-cell cancers, severe combined immunodeficiency (SCID), and multiple sclerosis (Strasser et al., 1990, Tchilian et al., 2001, Wu et al., 2002, Carulli et al., 2008, Basso and Dalla-Favera, 2015). Another noteworthy observation is the genomic instability event occurring on chromosome 17, which in mouse is where the AID-mutated proto-oncogene *PIM1* is found, and where we reported finding RNA:DNA hybrid accumulation in the absence of proper RNA exosome activity (Hilkens et al., 1986, Pasqualucci et al., 2001, Pasqualucci et al., 2004, Pefanis et al., 2014). Additional SKY-assayed metaphase-arrested stimulated RNA exosome-deficient B cells ought to be examined before we are able to make conclusive remarks. Still, our results give us clues as to which genomic loci might be involved during the genomic instability events and might be implicated in the pathophysiology of a given B-cell immunoproliferative or an immunodeficiency disorder.

III.1.5. RNA exosome Deficiency-Related Instability in ES Cells

Subsequently, we undertook the question of the AID- and/or class-switch recombination (CSR)-relatedness of these observations. AID is a powerful mutagen, whose activity is highly regulated and mostly channeled towards physiological mutagenesis-initiated immunoglobulin locus remodeling in activated B cells, though it is also sometimes associated with a number of disorders, as was discussed previously. We wanted to find out whether the genomic instability phenotype seen after abrogation of RNA exosome physiological activity was associated to AID

and/or CSR. We hypothesized that to the extent that R loops may be conducive to instability, at least some of this instability would indeed be associated to AID-activity in the context of a B lymphocyte undergoing CSR, while some may be unrelated to AID mutagenic activity or to the double-stranded break intermediates generated during CSR, and may instead reflect other dynamic transcription-associated mutagenesis (TAM) and recombination (TAR) events, particularly in light of the recent description of early replicating fragile sites, some of which seem to be AID-independent while other seem associated to AID (Kim and Jinks-Robertson, 2012, Barlow et al., 2013).

To explore this question, we performed the T-FISH assay on embryonic stem (ES) cells from mice with the *ROSA26*$^{Cre\text{-ERT2/+}}$; *EXOSC3*$^{COIN/COIN}$ background. ES cells provide a convenient tool in which to study the effects of RNA exosome activity abrogation in a high proliferation system where neither the B lymphocyte-specific somatic hypermutation nor class-switch recombination are occurring. Before exposure to 4-OHT, cells bearing the re-engineered *EXOSC3* conditional locus express normal levels of RRP40, which then go down once exposure to tamoxifen has taken place (Pefanis et al., 2014). We assayed the cells before exposing them to 4-OHT as well as after 4-OHT exposure. As with B cells lacking proper RNA exosome activity, we observed that compared to pre-abrogation ES cell metaphases, more RNA exosome-deficient ES fixed metaphases displayed structural chromosomal abnormalities (Figure 3.11). We then wanted to compare the instability observed in the ES-cell and B-cell backgrounds. To do this, we calculated a "genomic instability index," which gives us an estimation of the instability in RNA exosome-deficient cells relative to their RNA exosome-sufficient counterparts in both the ES-cell and B-cell backgrounds (Figure 3.11). We calculated that the genomic instability index derived

from RNA exosome-deficient B cells is 62% higher than the genomic instability derived from RNA exosome-deficient ES cells (4.331 vs 2.672, respectively), suggesting that the genomic

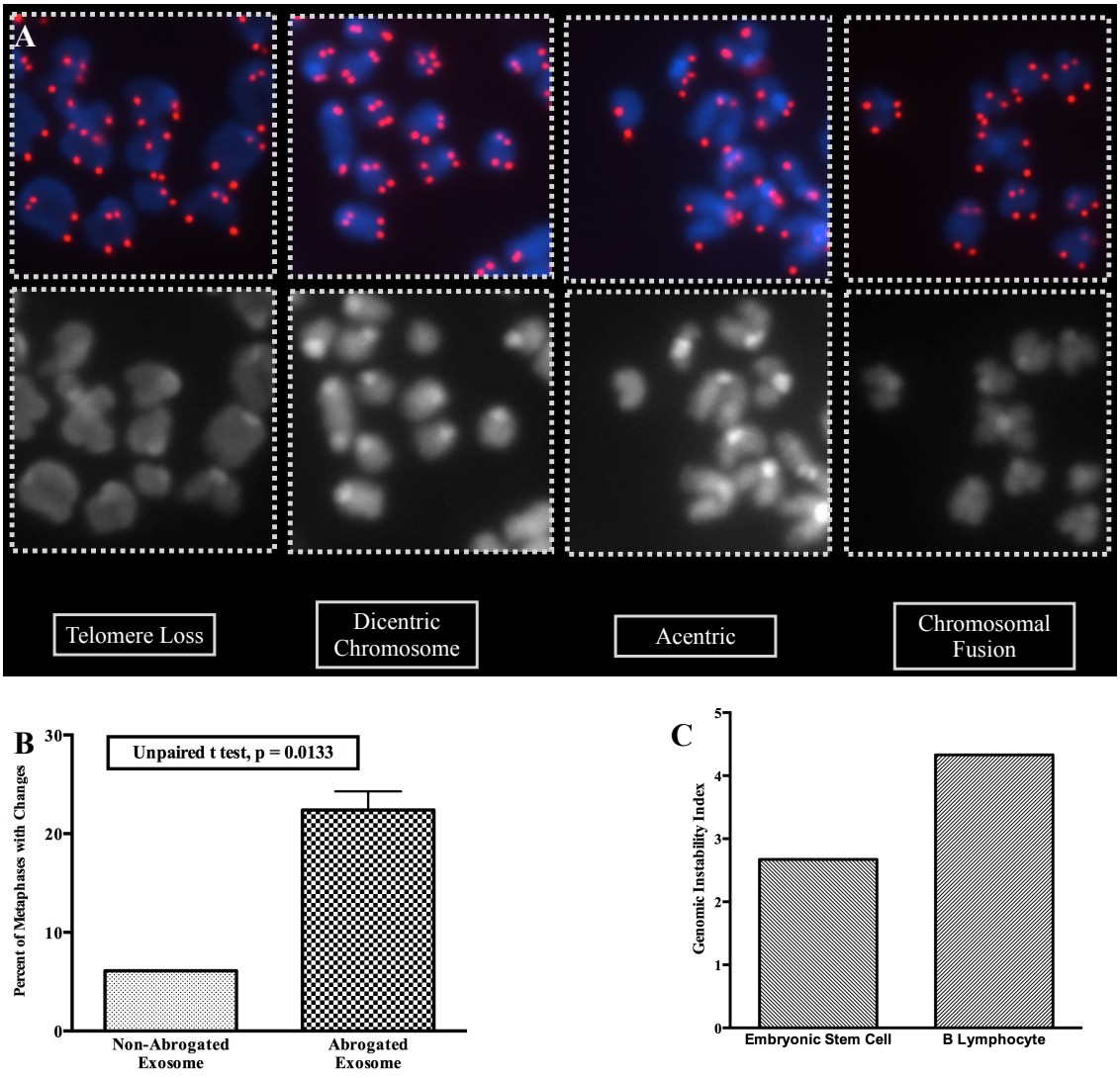


Figure 3.11. Embryonic stem cell T-FISH. *ROSA26*$^{Cre\text{-ERT2/+}}$; *EXOSC3*$^{COIN/COIN}$ ES Cells were subjected to a T-FISH assay before exposure to 4-OHT (therefore before RNA exosome function-abrogation) as well as after exposure to 4-OHT (when RNA exosome assembly is impaired). A. The images showcase examples of different structural abnormalities akin to those seen in stimulated B lymphocytes lacking proper RNA exosome function. B. The left plot compares the percentage of fixed metaphase with chromosomal abnormalities before and after RNA exosome abrogation. C. The right plot shows the genomic instability indices devised to compare the level of genomic instability observed when RNA exosome activity is impaired in both ESCs and B cells.

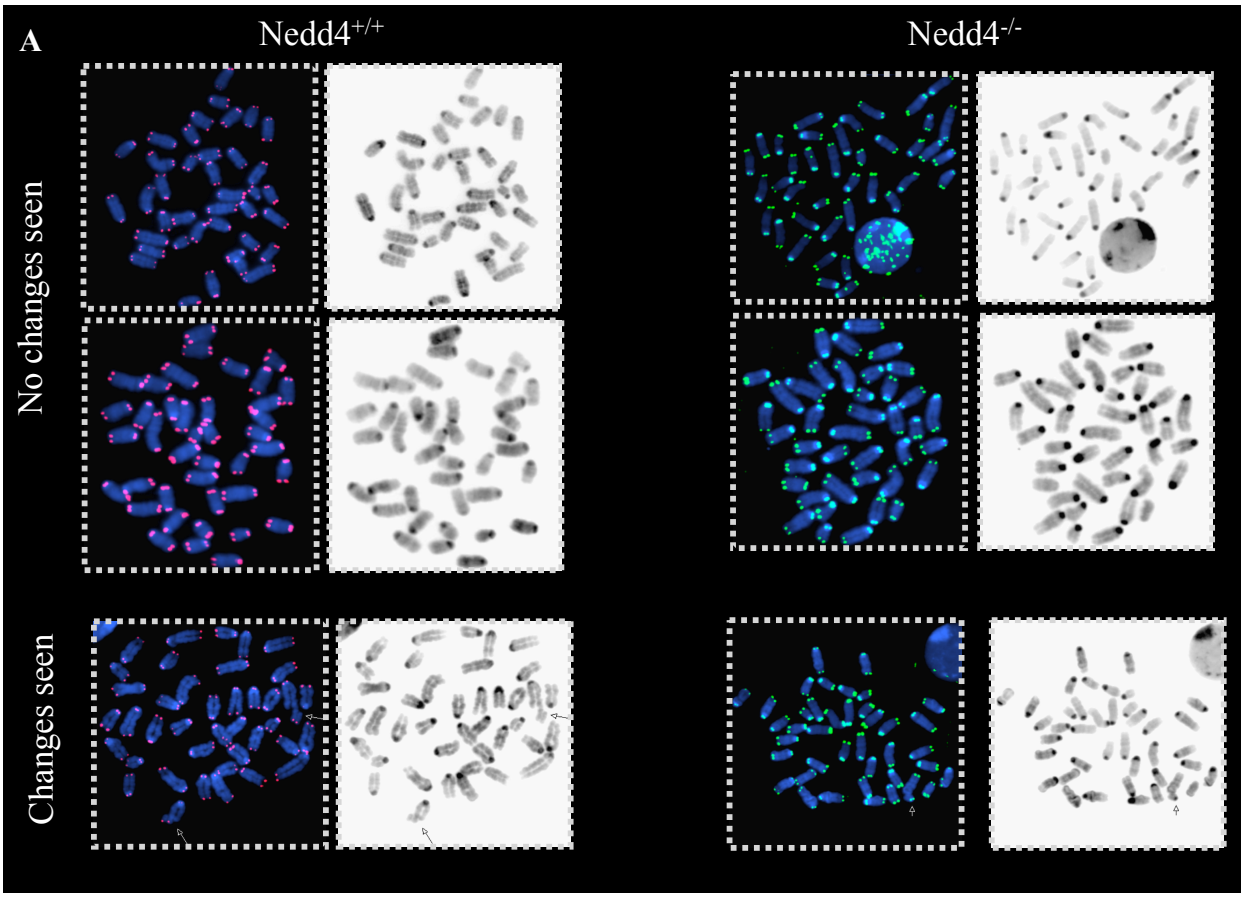
instability burden in activated, RNA exosome-deficient B cells is higher than in proliferating RNA exosome-deficient ES cells, even when normalized to their respective baseline genomic instability levels. These results provide some indication that part of the genomic instability seen in B cells may be non-related to AID activity, while a portion of this instability may in part be due to the genomic alterations that are specific to these cells.

III.1.6. Ubiquitination Dynamics In the Ig R-loop Reaction

There is evidence that RNA polymerase II gets monoubiquitinated by the E3 ligase NEDD4, which can instigate pol II polyubiquitination by another enzyme and precipitate its destabilization from loci where it had stalled, providing access to the template DNA strand to the single-stranded mutator AID to initiate physiological genomic alterations (Anindya et al., 2007, Harreman et al., 2009, Rotin and Kumar, 2009, Cheung and Cramer, 2012, Sun et al., 2013a). NEDD4 would therefore be involved in the putatively stabilizing monoubiquitination of RNA polymerase II and the degradation-mediating polyubiquitination of RNA polymerase II. Different cellular outcomes are possible for each of these NEDD4 roles, and the interactions among the different post-translational modification pathways are still not entirely elucidated.

We isolated splenic B cells from wild-type mice and grew them *ex vivo* with the CSR-inducing LPS and IL-4, along with the E1 heterodimer component protein APPBP1/NAE inhibitor MLN4924, a small molecule structurally related to 5'-monophosphate (AMP), which is produced in the ATP-dependent E1 reaction (Soucy et al., 2009). We wanted to see what the effect of impairing this pathway would be on the genomic integrity of activated B cells. Prior

work in our laboratory had shown that NEDD4 supports CSR and promotes AID interaction with its cofactors in the Ig R loop reaction, Spt5 and RNA exosome (Sun et al., 2013a). The NAE inhibitor MLN4924 was used at different concentrations (0 $mol.L^{-1}$, 62.5 $nmol.L^{-1}$, and 125 $nmol.L^{-1}$) for 72 hours, along with the CSR cocktail, after which colcemid was added to induce metaphase arrest. A series of washes followed to fix the metaphases and perform a T-FISH assay. Although MLN4924-non-treated cells were able to yield metaphases, only few metaphases were obtainable with the drug-treated B lymphocytes, preventing us from performing a proper analysis. It seems that the drug had some toxic effects, as almost no usable metaphase spreads were obtained from the 125 $nmol.L^{-1}$ treatment (Data not shown) (Whitby et al., 1998, Hjerpe et al., 2012, Enchev et al., 2015). We then attempted to investigate NEDD4 involvement in the Ig R loop reaction-mediated genomic instability differently by using B cells from mouse fetal liver chimeras expressing a C-terminally deleted form of NEDD4, i.e., mutants that no longer have an intact ubiquitin ligase-containing HECT domain (Yang et al., 2008). Impairing NEDD4 could in theory affect not only AID neddylation, but RNA polymerase II mono/polyubiquitination as well. We set out to find out if perturbing NEDD4 levels would have a consequence on genomic instability. We isolated splenic B cells from these chimeric mice, grew them *ex vivo* in medium contain LPS and IL-4 for 3 days, arrested them at metaphase by colcemid exposure, and completed the T-FISH assay by fixing and hybridizing the genomic material; control B cells were from fetal chimeric mice expressing non-truncated NEDD4 (Sun et al., 2013a). Unlike with the MLN drug treatment T-FISH assay, metaphases were readily obtainable in the absence of physiological NEDD4 from fetal liver chimeric mouse B cells. We noticed that there was no increase in genomic instability in the NEDD4-deficient background (Figure 3.12).



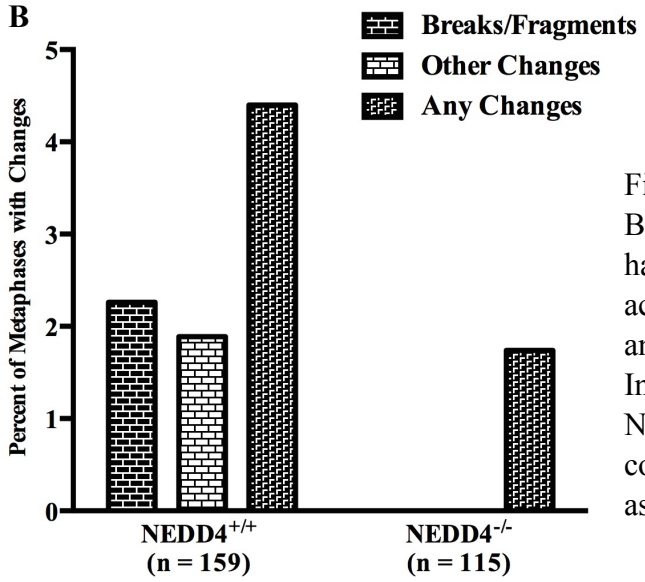


Figure 3.12. NEDD4$^{-/-}$ T-FISH. CSR-stimulated B cells from mouse fetal liver chimeras that have either wild-type or abrogated NEDD4 activity are prepared and assayed for T-FISH, and images from these assays are shown in A. In B, a plot with instability levels of both NEDD4$^{-/-}$ and NEDD4$^{+/+}$ backgrounds shows comparable levels of genomic instability as assayed by T-FISH.

III.2. Senataxin in the Ig R-loop Reaction

III.2.1. SETX Associates with AID and RNA Exosome

In our quest to decipher the intricacies behind the immunoglobulin R-loop reaction that permit the access of the mutator enzyme AID to both the readily accessible non-transcribed DNA strand and the RNA-hidden DNA strand of its target sequences during transcription, we identified the RNA exosome complex as a critical cofactor or AID activity and targeting (Basu et al., 2011, Pefanis et al., 2014). We also contemplated the possibility that other players might be involved in the Ig R-loop reaction. Previous work in *Saccharomyces cerevisiae* had shown that a complex formed by Sen1 (the yeast orthologue of SETX) with Nrd1 (N-arginine dibasic convertase) and Nab3 (nuclear polyadenylated RNA-binding protein 3) physically interacts through Nrd1 with the RNA exosome complex to mediate transcription termination (Vasiljeva and Buratowski, 2006). More recently, a report using a yeast two-hybrid screen, with as prey a human brain cDNA library and as bait the amino terminal region of SETX, suggested a SUMO-dependent interaction between the RNA Exosome subunit RRP45 and SETX (Richard et al., 2013). The RNA exosome is composed of a hetero-nonameric core that can interact with catalytic cofactors to bring about its 3'→5' exo- or endoribonuclease activities to degrade or process various RNA classes, and RRP45 is a core-component subunit of the RNA exosome complex (Januszyk and Lima, 2014). By analogy to the involvement of its orthologue in transcription termination at G-rich sequences downstream of polyA sites, we thought that SETX could, too, be involved in transcription termination during the Ig R-loop reaction. In order to investigate the potential involvement of SETX in antibody diversification mechanisms, we

therefore set out to find out whether SETX interacts with AID, the initiator and *bona fide* protein of antibody diversification reactions, and the RNA exosome complex in a B-cell background.

Because of the dearth of usable antibodies against murine SETX, we performed immunoprecipitation (IP) assays on Ramos cells, a human Burkitt's lymphoma cell line (Klein et al., 1975). We prepared Ramos cell lysates, washed them, incubated them with antibody-coupled agarose A/G protein, and eluted protein complexes before sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblot assaying with the appropriate antibodies. We observe that, similar to corroborating evidence in previous but different organism backgrounds, SETX interacts with RRP45 in Ramos cells (Figure 3.13). Furthermore, we observed that SETX also interacts with AID (Figure 3.13). Taken together, these results place the helicase SETX at the very center of the immunoglobulin R-loop reaction and strongly indicate that SETX may occupy an important role in antibody diversification mechanisms.

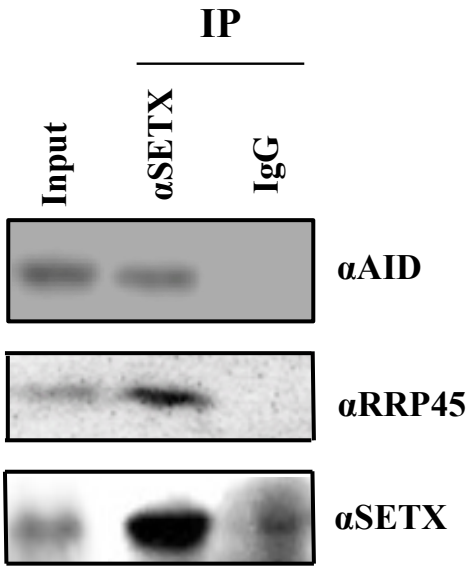


Figure 3.13. SETX immunoprecipitation. Human Burkitt's lymphoma Ramos cells were harvested, lysed, and prepared for SETX immunoprecipitation. Whole cell lysate and IgG controls are displayed., along with Western blotting using antibodies against AID and RNA exosome core subunit RRP45.

III.2.2. SETX-deficient CH12 Cells Exhibit Impaired CSR

We next set out to determine if, in fact, SETX occupies such a role in class-switch recombination. We reasoned that if SETX plays an important role during CSR, defects in switching ability should be perceptible when SETX activity is impaired. We proceeded to knockdown the expression of the *SETX* gene using the CH12F3 cell line, a murine B-cell lymphoma cell line that is able to undergo isotype switching upon stimulation (Arnold et al., 1983, Kunimoto et al., 1988, Nakamura et al., 1996). We introduced lentiviral backbone and shRNA-containing plasmids into 293T cells to produce packaged virus for the spinfection of our CH12F3 cells to cause the down regulation of *SETX*. Our knockdown approach, evaluated by RT-PCR using sets of primers that recognize two distinct regions of the gene, resulted in a statistically significant decrease of about 40 to 60 percentage points in SETX mRNA levels

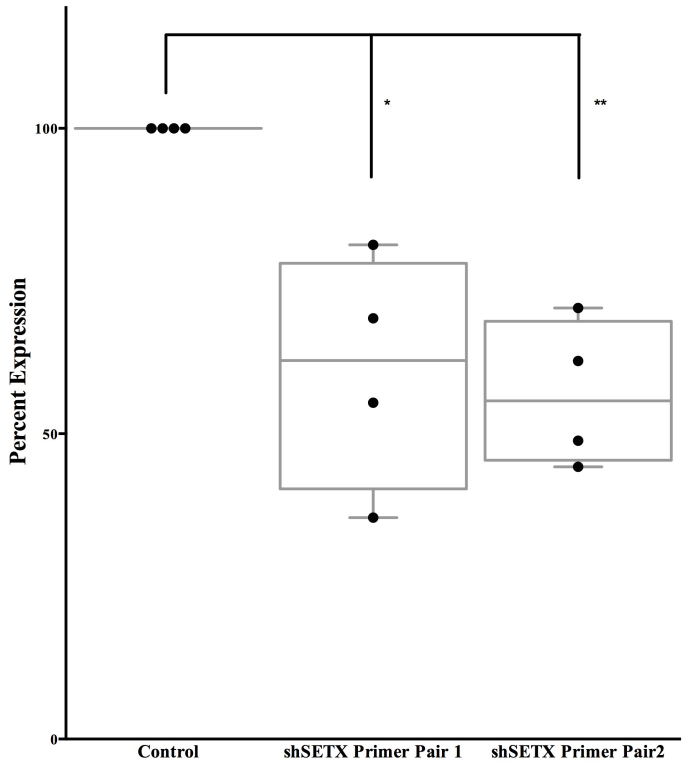


Figure 3.14. *SETX* Transcription levels in knockdown background. quantitative PCR was performed on murine CH12F3 cells containing a control vector or a senataxin knockdown construct. Displayed here are results of an analysis using two primer sets. Results are shown relative to control vector levels. Student t test was used in either case, and primer pair 1 had p = 0.0061, while for primer pair 2, p = 0.0003

(Figure 3.14). Because CH12F3 cells are of murine background, we were not able to assay and compare SETX protein levels in the different knockdown conditions.

Transcription plays a crucial role for proper, AID-initiated physiological CSR (Ramiro et al., 2003). To evaluate whether germline transcription decreased as a result of knocking down SETX, we monitored μ and α switch sequence transcription in the *SETX* knockdown background. We observed that germline transcription at switch sequences was not adversely impacted under these conditions (Figure 3.15). AID is the critical mutagen that initiates both SHM and CSR (Muramatsu et al., 2000). Given this critical role in Ig diversification, we also checked the mRNA levels of *AICDA* in the *SETX* knockdown background. As we had before, we observed that the knockdown conditions did not seem to negatively impact AID mRNA levels (Figure 3.15).

To further minimize confounding factors, we appraised the proliferation proficiency of the CH12F3 cells in our experiment, hoping to get a sense of their general viability in conditions

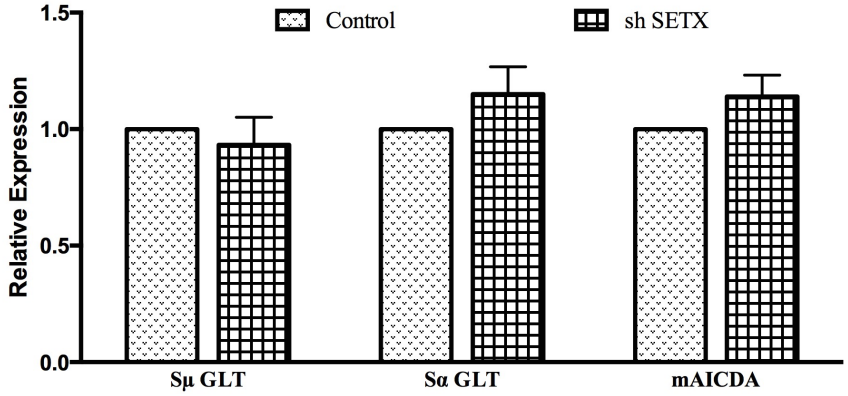


Figure 3.15. Transcription at *IGH* switch sequences and *AICDA* in K/D Background. S_μ and S_α germ line transcripts (GLT) in shSETX knockdown backgrounds are shown, as is the transcription at mouse *AICDA*, relative to control background. We observed no statistically appreciable defect in transcription at any of the three loci.

of *SETX* knockdown. For this, we used the violet laser excitable Violet Proliferation Dye 450 (VPD 450), which passively diffuses into cells and is cleaved by esterase activity within viable cells to become fluorescent; upon dividing, these cells partition their dye into similar amounts between the next-generation cells, causing a dilution in the overall dye intensity. We measured VPD450 at 24-hour intervals over 72 hours. We observed that there was no proliferation deficiency and that the knockdown and control backgrounds behaved similarly (Figure 3.16).

Next, we stimulated the cells to undergo isotype switching to IgA. We exposed CH12F3 cells with different backgrounds to LPS, IL-4, and transforming growth factor β (TGF-β) for 48 hours and evaluated the percentage of IgA^+ cells (Figure 3.17). Knocking down *SETX* in

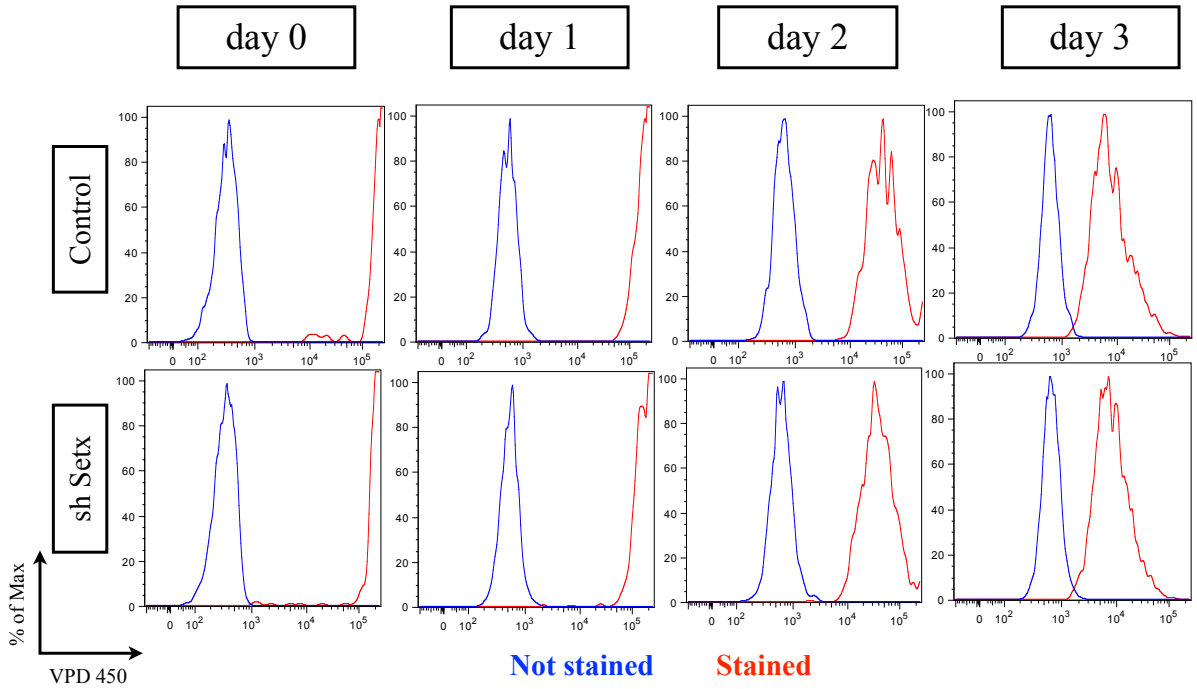


Figure 3.16. Proliferation of *SETX* knockdown CH12F3 cells. Using the dilution dye VPD450, we tracked the proliferation of CH12F3 cells containing control vector or shSETX knockdown construct over 3 days. As the cells proliferate, the intensity of the dye decreases. Shown in blue are cells that were not stained, while in red are stained cells. As the cells proliferate, the VPD450 intensity of stained cells is diluted and nears that of the not stained cells.

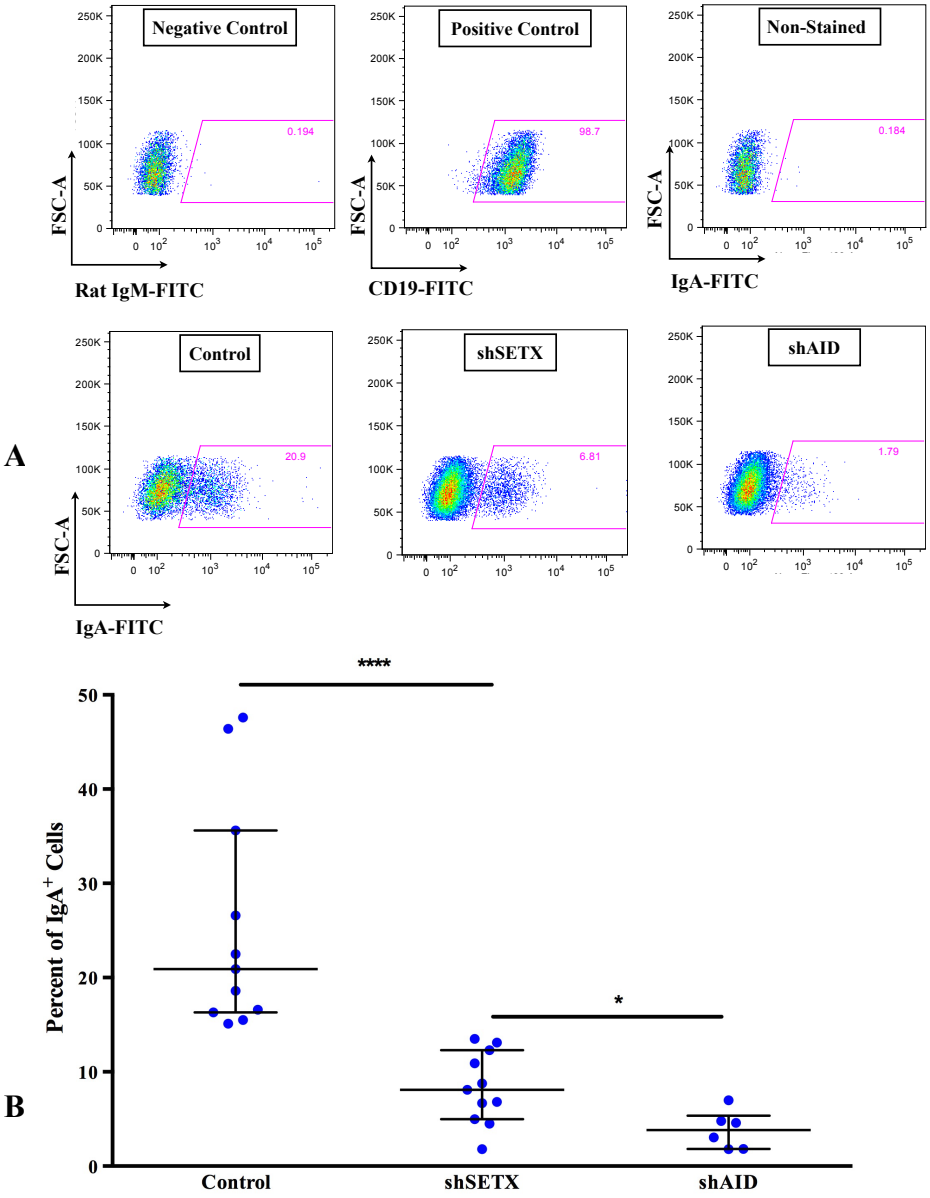


Figure 3.17 Class-switch recombination of *SETX*-deficient CH12F3 cells. A. FACS plots show on the top row staining controls, and on the bottom row experimental results, with control vector CH12F3 cells, SETX-deficient cells, and AID-deficient cells. Stimulated CH12F3 cells deficient in SETX show CSR results that are comparable to the AID control. B. After multiple repeats of the isotype switching assay, the comparative plot was drawn. There is a statistically significant decrease from the control vector background to the SETX-deficient background (unpaired student t test, $p = 0.0002$), as well as between the SETX-deficient and the AID-deficient backgrounds (unpaired student t test, $p = 0.0182$).

CH12F3 cells resulted in a substantial reduction in their ability to undergo CSR, as measured by the proportion of SETX-deficient CH12F3 cells expressing IgA on their surface (Figure 3.17). In this regard, CSR in SETX-deficient CH12F3 cells greatly resembled CSR in AID-deficient CH12F3 cells, which we used as our CSR-abrogation positive control, and are much different from the negative control background, which uses an empty vector backbone (Figure 3.17).

III.2.3. SETX-deficient CH12 Cells Accumulate RNA:DNA Hybrids

Our results suggested that SETX indeed may play a role in the immunoglobulin R-loop reaction. Upon elimination of possible other explanations (such as poor cellular proliferation, lack of switch sequence transcription, or lack of mutator gene *AICDA* transcription), we entertained the hypothesis that the putative RNA/DNA helicase SETX is mechanistically implicated in the Ig R-loop reaction by helping unwind the RNA:DNA hybrids that arise when R loops are formed during CSR. In SETX-deficient backgrounds, the RNA moiety of the still-wound RNA:DNA heteroduplex may not be amenable to degradation or removal by the RNA exosome complex, leading to an accumulation of these RNA:DNA hybrids, and, as a consequence, to the non-resolution of R loops. We therefore attempted to find evidence of these accumulating heteroduplexes.

First, we made use of a mouse monoclonal antibody ("S9.6") that recognizes RNA:DNA hybrids (Boguslawski et al., 1986, Bhatia et al., 2014). We stimulated SETX-proficient and SETX-deficient CH12F3 cells for 48 hrs with LPS, IL-4, and TGF-β, then harvested, washed, fixed, and prepared them for IF. We observed that in the SETX-deficient CH12F3 background, there was an increase in the proportion of prepared cells with 2 foci or more, while the proportion of prepared cells with no foci or 1 focus decreased compared to the SETX-sufficient

background (Figure 3.18). Additionally, although these fixed cells with no foci or 1 focus detected accounted for ~ 82% of all the fixed cells in the SETX-sufficient background, but comprised less than 50% in the SETX-deficient background (Figure 3.18).

We also attempted to corroborate the accumulation of RNA:DNA heteroduplexes in SETX-deficient CH12F3 cells using a recombined protein that recognizes RNA:DNA hybrids and contains a GFP reporter ("HB-GFP"). This recombined protein was engineered from the hybrid-binding domain of RNaseH1 ("HB") fused to a GFP cassette (Bhatia et al., 2014). The HB-GFP construct was introduced into CH12F3 cells by electroporation; the cells are allowed to recover overnight, and are then stimulated for CSR. After 1.5 days, GFP-positive cells are sorted, permeabilized with Triton X-100, allowing the HB-GFP recombined protein to leak out of the cells — unless the protein is bound to RNA:DNA hybrids, impairing its leaking out. This HB-

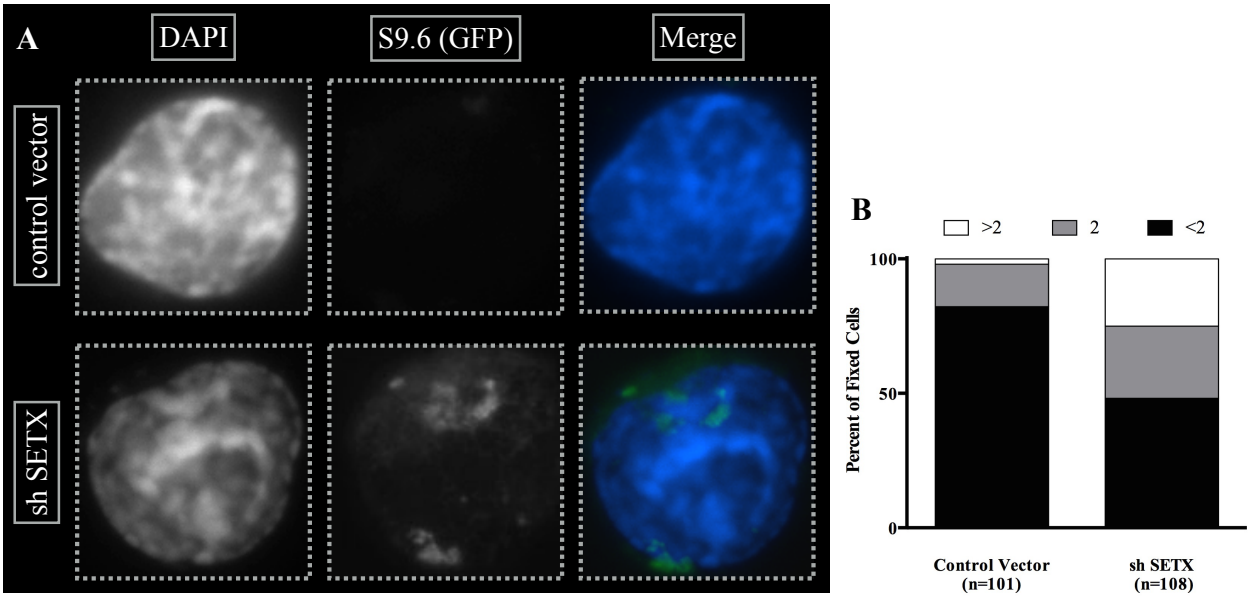


Figure 3.18. S9.6 immunofluorescence assay. A. Images of CH12F3 cells containing control vector or SETX knockdown vector, stimulated for isotype switching, and assayed by immunofluorescence with the anti-RNA:DNA hybrid antibody S9.6. B. A plot of the two backgrounds shows an increase in the proportions of prepared cells that have 2, or more than 2, S9.6 foci as measured by ImageJ in our assay.

GFP retention assay would have allowed us to assess RNA:DNA hybrid accumulation using a different method; however, we weren't satisfied with the quality of the assay in CH12F3 cells, and could not confidently rely on observations obtained from it.

III.2.4. SETX-Knockdown CH12F3 Cells Exhibit an Altered Mutational Profile

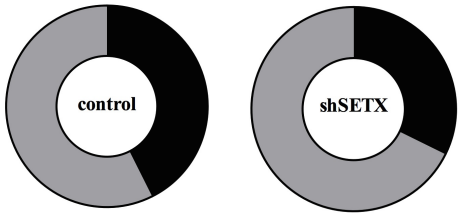
Beyond the prospects of a mechanistic implication of SETX in the immunoglobulin R-loop reaction, we wondered what the consequences of such an implication may be. Based on our hypothesis, the accumulation of RNA:DNA heteroduplexes would impair proper diversifying enzyme AID access to the transcribed DNA strand, and therefore result in a different mutational profile at switch sequences. We decided to probe this prediction. We sequenced the 5'-most region of switch sequence μ in SETX-sufficient and -deficient CH12F3 cells stimulated to undergo isotype switching to IgA by exposure to IL-4, LPS, and TGF-β. We observe that the proportion of mutations that occur at GC residues is higher in SETX-deficient CH12F3 cells than in control CH12F3 cells; in parallel, the proportion of mutations that occur at AT residues is lower in SETX-deficient CH12F3 cells than in SETX-sufficient CH12 cells (Figure 3.19A). We wanted to know whether the proportional mutational increase at GC residues occurred at the expense of mutations at AT residues. Tellingly, our analysis shows that the mutation rate at GC residues in SETX-deficient CH12F3 cells is higher than in SETX-sufficient CH12F3 cells (Figure 3.19B). Our analysis also shows that the mutation rate at AT residues is higher in SETX-deficient CH12F3 cells than in SETX-sufficient CH12F3 cells (Figure 3.19B). However, the mutation rate increase from SETX-sufficient to SETX-deficient is more pronounced for GC residues compared to AT residues (Figure 3.19B).

A

control					
↱	A	C	G	T	
A		1	14	1	16
C	0		1	8	9
G	21	0		1	22
T	0	6	1		7
	21	7	16	18	54

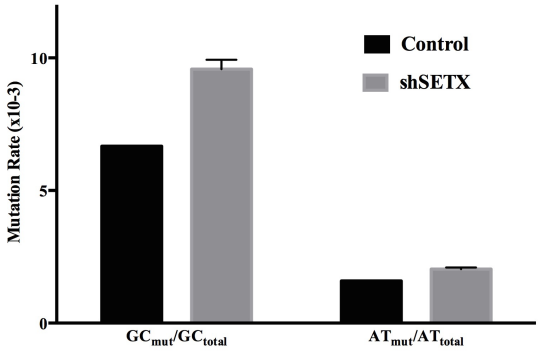
shSETX					
↱	A	C	G	T	
A		0	16	3	19
C	0		3	23	26
G	29	0		0	29
T	3	6	0		9
	32	6	19	26	83

	control	shSETX
AT	42.6%	32.3%
GC	57.4%	67.7%



B

	Control	shSETX
mutations at GC/ total GC	75/11238 = 6.67 x 10^{-3}	105/11394 = 9.22 x 10^{-3}
mutations at GC/ total ACGT	75/25748 = 2.91 x 10^{-3}	105/26130 = 4.02 x 10^{-3}
mutations at AT/ total AT	23/14510 = 1.59 x 10^{-3}	31/14736 = 2.10 x 10^{-3}
mutations at AT/ total ACGT	23/25748 = 0.893 x 10^{-3}	31/26130 =1.19 x 10^{-3}



C

Non-Transcribed Strand		
Control	G	22/7064 = 3.11 x 10^{-3}
	C	9/4174 = 2.16 x 10^{-3}
shSETX	G	30/7158 = 4.19 x 10^{-3}
	C	35/4226 = 8.29 x 10^{-3}

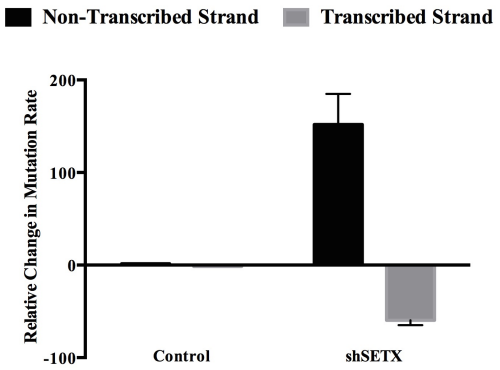


Figure 3.19. Mutation profile of stimulated SETX-deficient CH12F3 cells. A. Distribution of mutations in control and SETX-deficient backgrounds. In the SETX-deficient background, compared to the control background, a decrease in the proportion of mutations at AT residues and an increase in the proportion of mutations at GC residues are noticeable. B. Further analysis of the mutations in the SETX-deficient background shows that overall mutations at AT and GC both increase, but that this increase is more pronounced at GC residues, accounting for the opposite trends in A. C. Mutation Rate at C residues increases on the non-transcribed strand, and decreases on the transcribed strand, relative to the control vector.

We further probed the alteration of the mutational profile in SETX-deficient CH12F3 cells. We compared the mutation rate at deoxycytidine residues on the non-transcribed strand to the transcribed strand deoxycytidine mutation rate. One would predict that dC residues are getting mutated at higher rates on the non-transcribed strand than on the transcribed strand upon impairment of SETX activity. Indeed, we notice that compared to the control background, there is a marked increase in the mutation rate at dC residues found on the non-transcribed strand in SETX-deficient CH12F3 cells (Figure 3.19C). In parallel, there is a decrease in the mutation rate at dC residues found on the transcribed strand of SETX-sufficient CH12F3 cells (Figure 3.19C).

III.2.5. SETX-Mutant Mice Allow for Limited *In Vivo* Studies

Our findings in CH12F3 cells seemed compelling, and we subsequently made an effort to expand our investigation to an *in vivo* model of SETX activity. We obtained four mice (two males and two females) heterozygous for *SETX*, where one allele was wild-type and the other reengineered so its open reading frame (ORF) exon 4 (the second of 23 coding exons, the first two ORF exons being noncoding) is constitutively and ubiquitously deleted (Becherel et al., 2013). The mice were mated (Figure 3.20), and splenic B lymphocytes from 6-8 week-old

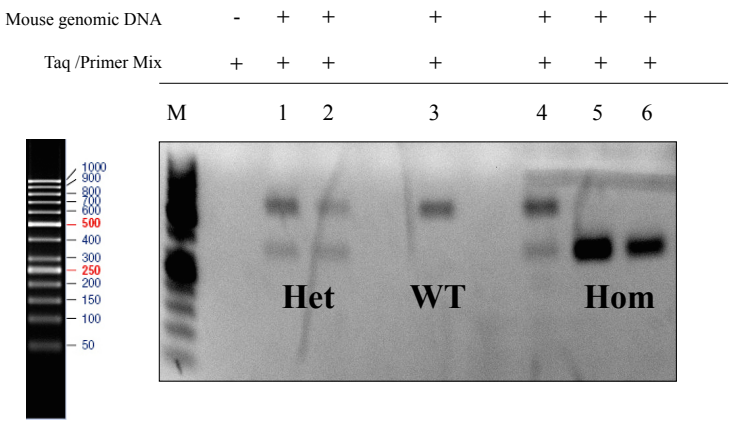


Figure 3.20. *SETX* mutant mouse. Genotyping PCR of 6 animals shows the different *SETX* backgrounds. Het: heterozygous, WT: wild-type, Hom: homozygous.

offspring homozygous for this *SETX* mutation were harvested and grown *ex vivo* for 72 hours in

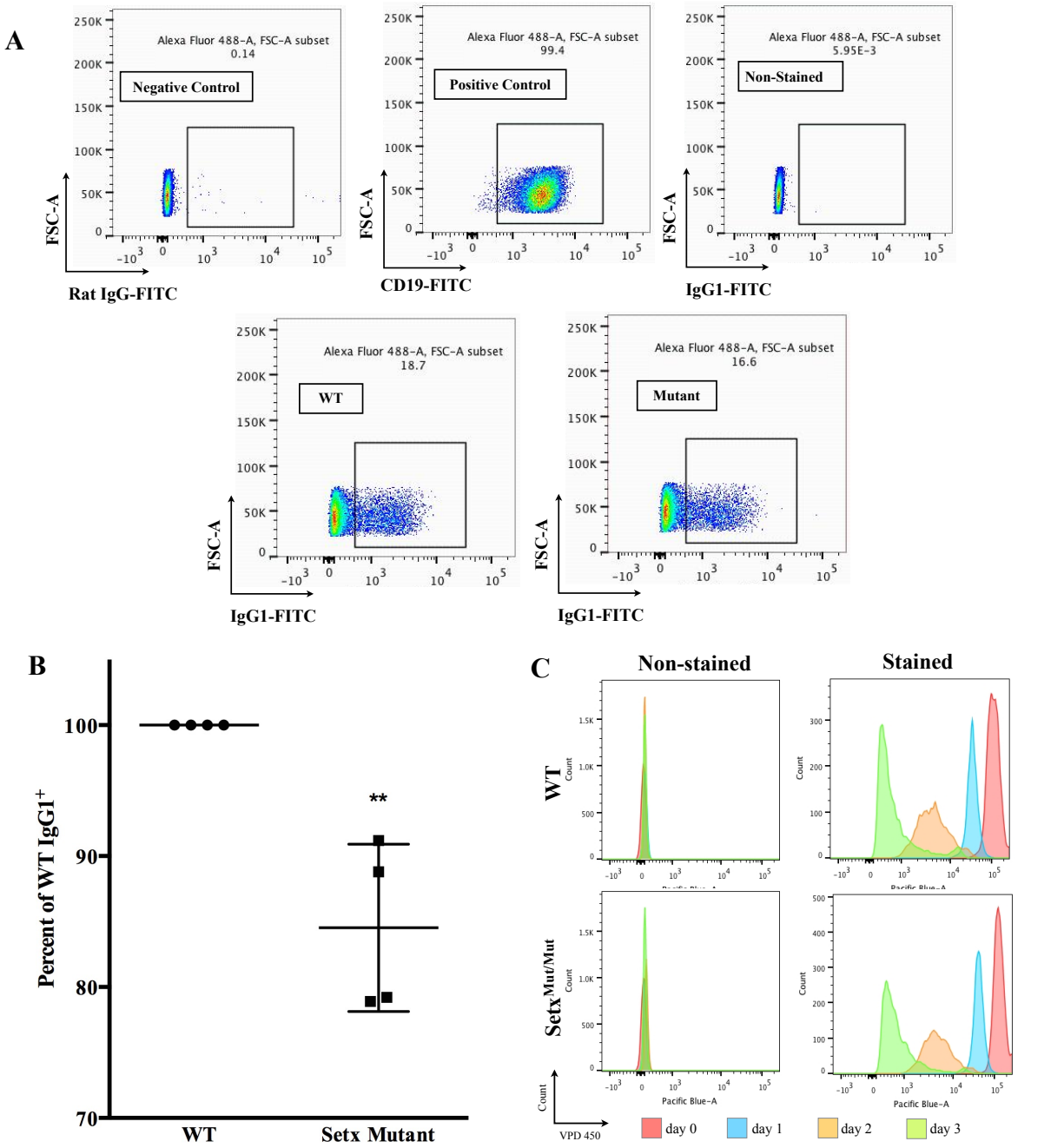


Figure 3.21. *SETX* mutant primary B cell isotype switching and proliferation. A. Staining controls are displayed on the top row, while experimental FACS plots are shown on the bottom row. B. Several assays suggest a small, statistically significant defect in CSR in *SETX* mutant primary B cells compared to their wild-type counterparts (unpaired student t test, p = 0.0029). C. Proliferation assay using VPD450 shows no noticeable defect in proliferation in the mutant background.

the usual LPS and IL-4 cocktail to stimulate CSR to IgG1. We then assayed the cells for CSR (Figure 3.21A). A number of iterations of the assay later, we observed a small though persistent decrease in isotype switching ability in *SETX*-mutant mice (Figure 3.21B). It seemed like B lymphocytes from mutant SETX backgrounds only had a modest statistically significant relative decrease in their ability to undergo isotype switching (Figure 3.21B). As with the CH12F3 cell line, mutant SETX B cells appeared to proliferate with no ostensible defect (Figure 3.21C).

Finally, we also set out to find out if splenic B lymphocytes from mutant SETX mice exhibited structural chromosomal instability akin to what we observed with the RNA exosome-deficient splenic B cells. The relative accumulation of RNA:DNA hybrids, as could be

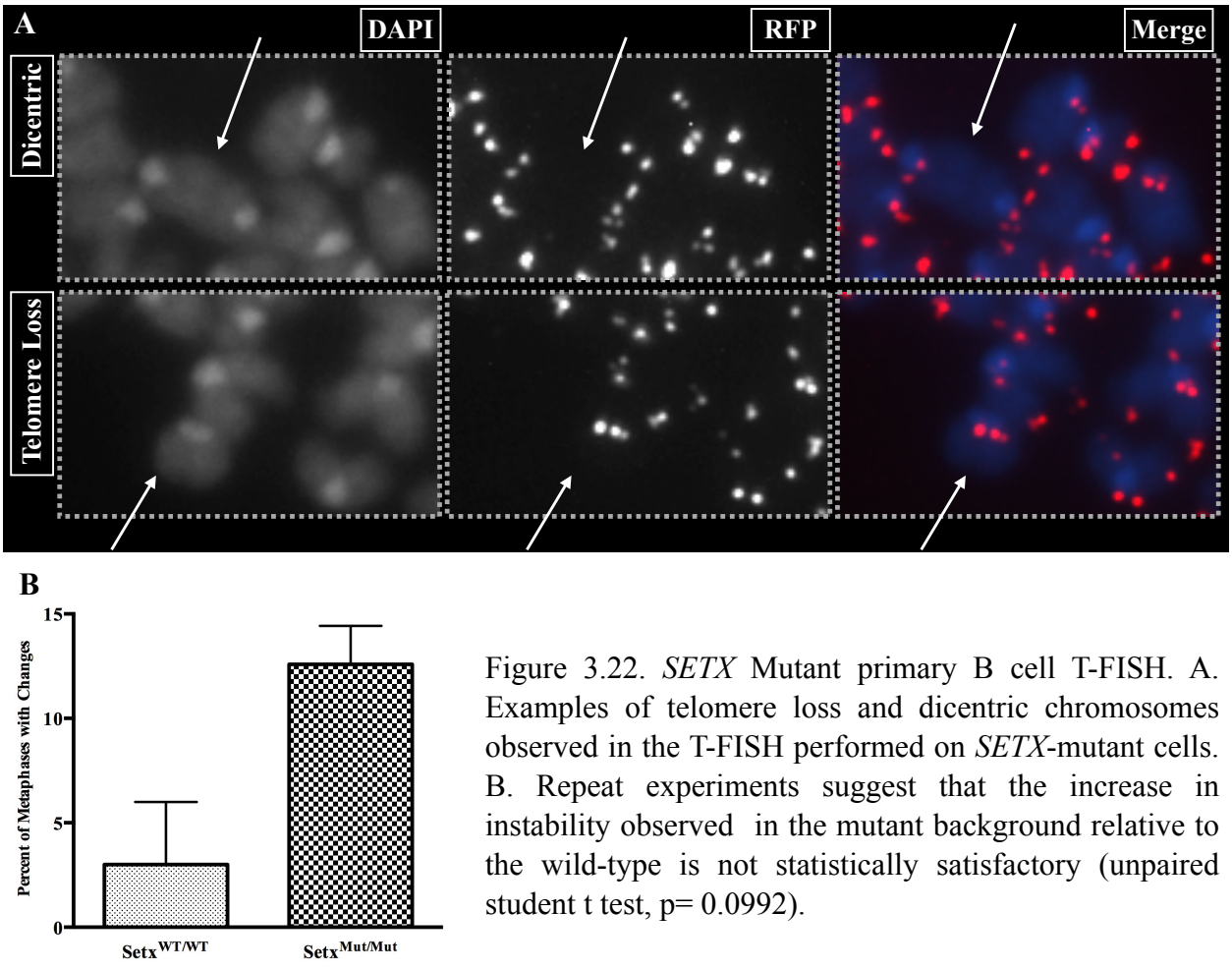


Figure 3.22. *SETX* Mutant primary B cell T-FISH. A. Examples of telomere loss and dicentric chromosomes observed in the T-FISH performed on *SETX*-mutant cells. B. Repeat experiments suggest that the increase in instability observed in the mutant background relative to the wild-type is not statistically satisfactory (unpaired student t test, p= 0.0992).

conjectured to be the case in the absence of properly active SETX, would be conducive to an increase of genomic instability by making R loops more perennial. We therefore performed a T-FISH assay on splenic B lymphocytes isolated from SETX mutant and wild-type mice, activated them for isotype switching for 72 hours, metaphase-arrested them by colcemid exposure, and washed, fixed, and prepared them for telomeric probe hybridization. Although there is some level of increase in genomic instability as measured by T-FISH in the SETX-mutant B lymphocytes compared to wild-type SETX B lymphocytes, the increase does so far not appear to be statistically significant, mirroring our earlier observations that the defects in these mutant mice in isotype switching are not very conspicuous (Figure 3.22).

Chapter IV: Discussion

IV.1. Relevance of Work

As aspects of the adaptive immune system still seem to be under strong selective pressure, critical features of this system have come under scrutiny, particularly as it appears that immunoproliferative disorders may be the evolutionary cost for the emergence of an adaptive immune system as malleable as ours is. This trade-off is by and large advantageous for us. Nonetheless, each time disease strikes, it can be a tragedy; furthermore, our insights into health and disease have so far not proved sufficient to staunch overall cancer prevalence or incidence rates. Non-Hodgkin Lymphoma rates in particular are of interest to us, as many immunoproliferative disorders of mature B-cell origin fall into this category. As is readily

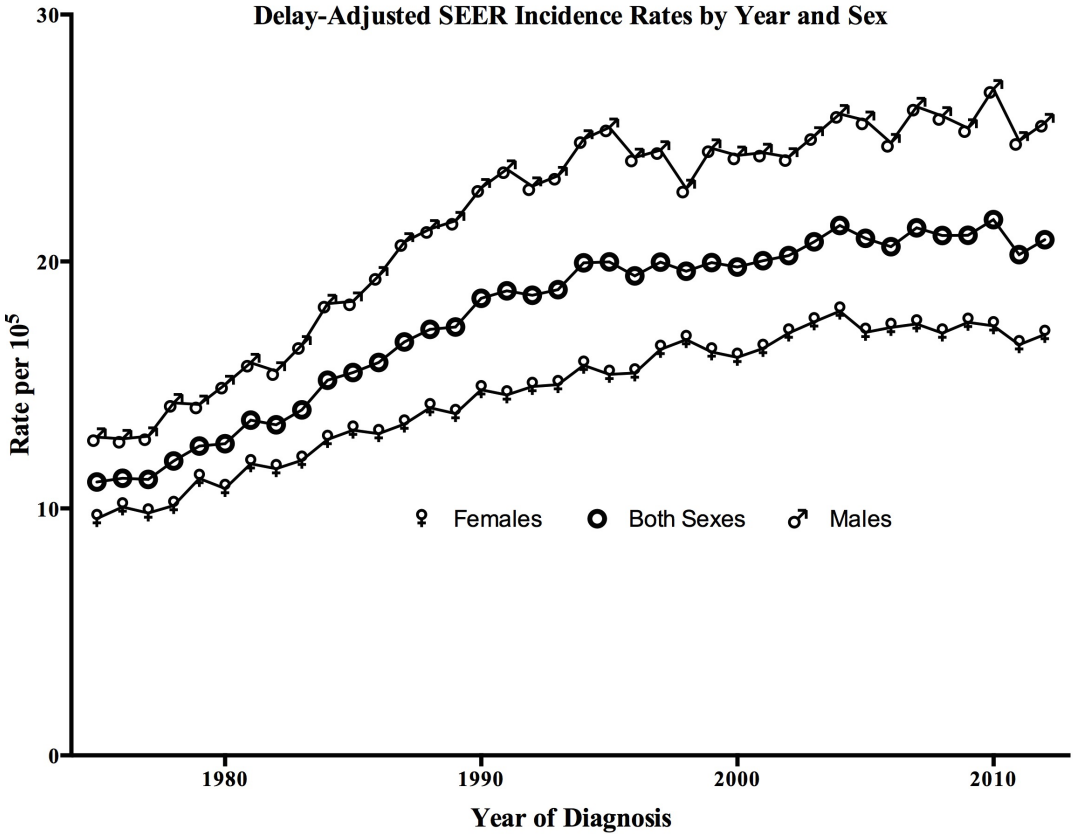


Figure 4.1. Incidence rates of non-Hodgkin lymphoma. Most cancers of B cell origin fall into the non-hodgkin lymphoma category, which has been trending upwards in the past four decades in both males and females. Raw data were obtained from the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute.

appreciable from Figure 4.1, incidence rates of these cancers remain elevated. Similarly, the constant threat of infectious diseases, including those caused by multi-drug resistant microbes and emergent zoonoses with the possibility of wide and rapid spread, compels us to learn more about the adaptive immune system in the hopes of engineering new or more effective ways to combat disease.

The B-cell arm of the immune system is central to a successful adaptive immune response. In particular, its ability to generate and produce high affinity immunoglobulin proteins with varied biological functions is a defining characteristic. Much has been learned in the 50 years since the B cell lineage was identified, including how the first round of lymphocyte immunoglobulin diversification occurs, that another round of diversification also occurs specifically in B lymphocytes, and that this second round of diversification is instigated by the APOBEC deaminase family member AID (Cooper et al., 1965, Kincade et al., 1970, Weigert et al., 1970, Honjo and Kataoka, 1978, Tonegawa, 1987, Muramatsu et al., 1999). Recently, efforts have been invested in untangling the complex machinery by which the immunoglobulin mutator AID is successfully targeted to its physiological substrates while largely sparing other locations, and understanding how it sometimes is mis-targeted to regions where its activity might portend mutations with dire consequences. Several co-factors of AID activity have been identified, including the RNA exosome complex (Basu et al., 2011, Pefanis et al., 2014).

Our work has contributed mechanistic insight into the current understanding by providing evidence that the immunological R-loop reaction is the molecular key that opens the door to the possibility of physiological deamination by AID, and that the regulation of this window of carefully orchestrated, physiological instability is mediated by the resolution of transcription-

initiated R loops. The regulation of this molecular window of instability will ensure that the deamination activity of AID happens and is restricted both in time (when AID activity takes place) and in space (where AID activity takes place). Transcription allows for the transient availability of a looped-out, non-transcribed strand that is rendered more lasting in part by the G-richness of this strand through the formation of an R loop. The access of the single-stranded deaminase AID to its transcribed strand substrate is then facilitated by transcription termination during abortive transcription of *Ig* target loci, notably because these target regions are more prone to RNA polymerase stalling and pausing, an occurrence facilitated by the activity of stalling/pausing cofactors and that precedes the backtracking and/or release of RNA polymerase II, freeing the 3'end of the synthesized RNA that can now be removed or degraded by RNA exosome complex, aided in this respect by the unwinding activity of putative helicase SETX; this not only unveils the transcribed DNA strand to AID activity, but it also allows the DNA duplex to reconstitute, which will close the R loop, and prevent further AID deaminating activity at this locus.

We explored this hypothesis in various ways. We investigated how altering the activity of RNA exosome might impact the structural integrity of the genome. We considered whether or how the post-translational modification experienced by RNA polymerase II is involved in the maintenance of genomic stability. Finally, we inquired about how SETX may be active during the immunoglobulin R-loop reaction.

IV.2. RNA Exosome and Related Instability

We used a conditional deletion model for the abrogation of RNA exosome activity in mouse cells. We knew that RNA exosome activity is essential for robust processes of genomic remodeling in the immunoglobulin locus (Basu et al., 2011, Pefanis et al., 2014). We suspected that the RNA exosome was centrally involved in the regulation of processes like CSR, and more specifically, that it is involved in the resolution of R loops found at *Ig* switch sequences. Given the relevance of these structures to the maintenance of genomic integrity, we hypothesized that RNA exosome activity during the Ig R-loop reaction helps prevent chromosomal instability. We observed that when we prevent RNA exosome complex assembly, and therefore abrogate its function, activated B lymphocytes experience a number of catastrophic events, as evidenced through our T-FISH assay, and that they are more sensitive to additional mutagenic stress, as evidenced by our ionizing radiation immunofluorescence assay.

The catastrophic events we witnessed included breaks and fragments (telomere loss, acentric chromosomal fragments), as well as rearrangements (dicentric chromosomes, homologous end-to-end fusions, radials). One explanation for the loss of telomeres or the presence of acentric fragments is cartooned in Figure 4.2, which shows the process by which these structures may be encountered in metaphase, and reflects the coinciding first appearance of breaks with AID activity in the cell cycle G_1 phase before replication.

The rearrangement phenotypes can appear when breaks and fragments are repaired intrachromosomally through end-to-end homologous fusion or with a different chromosome to form a dicentric (Figure 4.3) (rearranged chromosomes are only perceptible to a certain extent with our T-FISH assay; some balanced translocations or short intrachromosomal deletions, for

instance, would not be detected). These observations were reminiscent of Breakage-Fusion-Bridge cycles seen in some cancers, when chromosome breakage is repaired by a fusion event (between two sister chromatids or between two chromosomes), which is followed by the formation of a "bridge" during mitosis that gets broken again as the spindle pulls on each centromere to re-create broken ends available for another cycle (Sorzano et al., 2013, Feijoo et al., 2014). As for the radials, they are thought of as chromatid-type break (affecting only one chromatid, in contrast to both chromatids being affected in the chromosome-type breaks) that arise at sites of stalled replication forks and are repaired by the presumably non-physiological insertion of another chromosome (Savage, 1976, Scully et al., 2000, Satoh et al., 2002). Our results would then hint at the possibility that chromosome-type breaks (involving both chromatids, typically G_1, pre-replicative in origin) occur more frequently than chromatid-type breaks (involving one chromatid) (Figure 3.4). Common fragile sites, where DNA polymerase

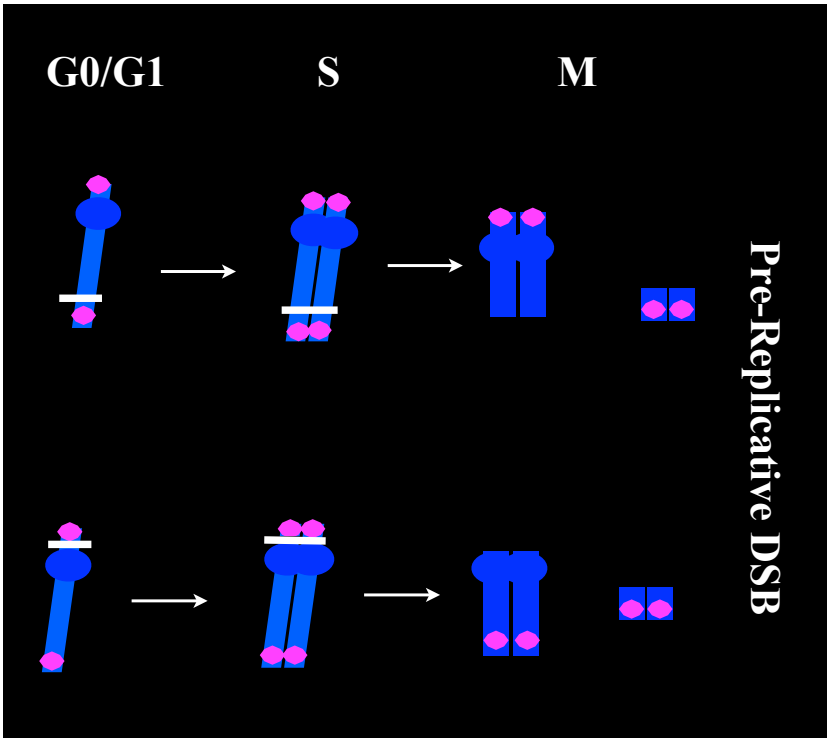


Figure 4.2. Generation of Breaks/Fragments Phenotype. After a break in G_1 phase that fails to be properly repaired, replication in S phase proceeds, and the resulting telomere loss or acentric fragment is detected in the following metaphase.

stalling occurs, and which are more susceptible to replication fork stalling and collapse, are often associated with these chromatid-type breaks that occur later in replication.

These observations also highlighted the possibility that instability events might be occurring at different times in the cell cycle (like the less frequent radials) or at additional sites sensitive to RNA exosome activity abrogation. Indeed, R-loop formation has been reported at several other locations, including the non-coding telomeric repeat–containing RNA (TERRA) transcripts (Cusanelli and Chartrand, 2015). The description of early replicating fragile sites (ERFSs), which similarly to CFSs, are sites prone to breakage, but contrary to them, occur early during replication, poses the question of whether they, too, could be loci where R loops are created (given that their high GC content) and where RNA exosome activity abrogation might be

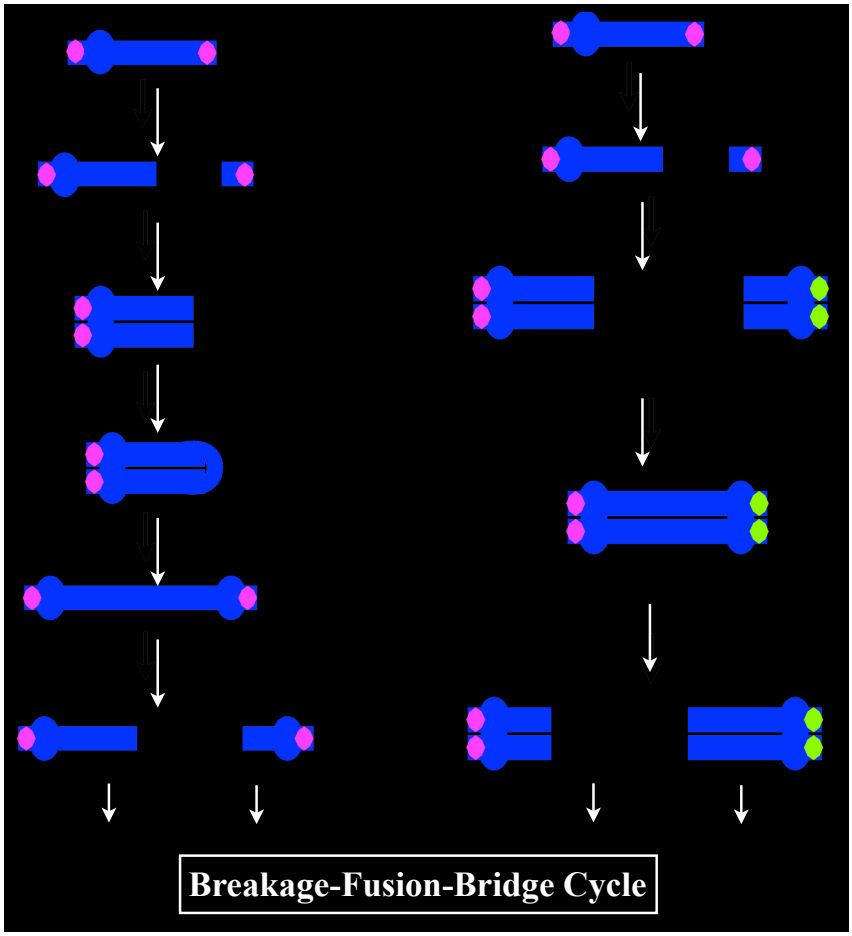


Figure 4.3. Generation of Fusion Phenotypes. Dicentric chromosomes and homologous end-to-end chromosomes may arise when the cell repairs a break in a chromosome by adjoining two sister chromatids or a different chromosome in a translocation event.

particularly consequential (Barlow et al., 2013). Comparing ERFS hotspots and AID targets, we could see that at least two loci belonged in both categories, including one, IRF4, affected in B cell cancer (Figure 4.4). Finally, the recent identification by our group of long non-coding RNAs and R-loop forming divergently transcribed transcription start-site and enhancer RNAs as substrates of the RNA exosome complex is auspicious, particularly given that some of these newly identified RNA exosome substrates, such as *BIRC3* and *NCOA3* enhancers, are involved in translocation events with the *IGH* locus described in translocation capture studies (Klein et al., 2011, Pefanis et al., 2014, Pefanis et al., 2015). This is strong indication that subsequent studies should investigate the identity of the loci involved in instability events with greater resolution than our SKY assay or than G-banding studies. One possible route to follow would be the use of 2-color *IGH* BAC-FISH, which would produce further evidence of *IGH* locus involvement, possibly in a translocation event (Figure 4.5). Another possible, more high-throughput avenue would include translocation capture assays of activated RNA exosome-deficient B cells. These would give us insights into the presence of translocation events, as well as the identity of such events.

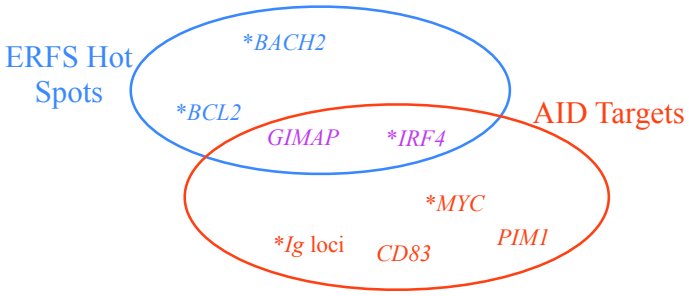


Figure 4.4. ERFS & AID Hotspots. Venn diagram illustrates the overlap existing in some instances between identified early replicating fragile sites and sites of AID activity. Asterisks denotes loci involved in translocation events

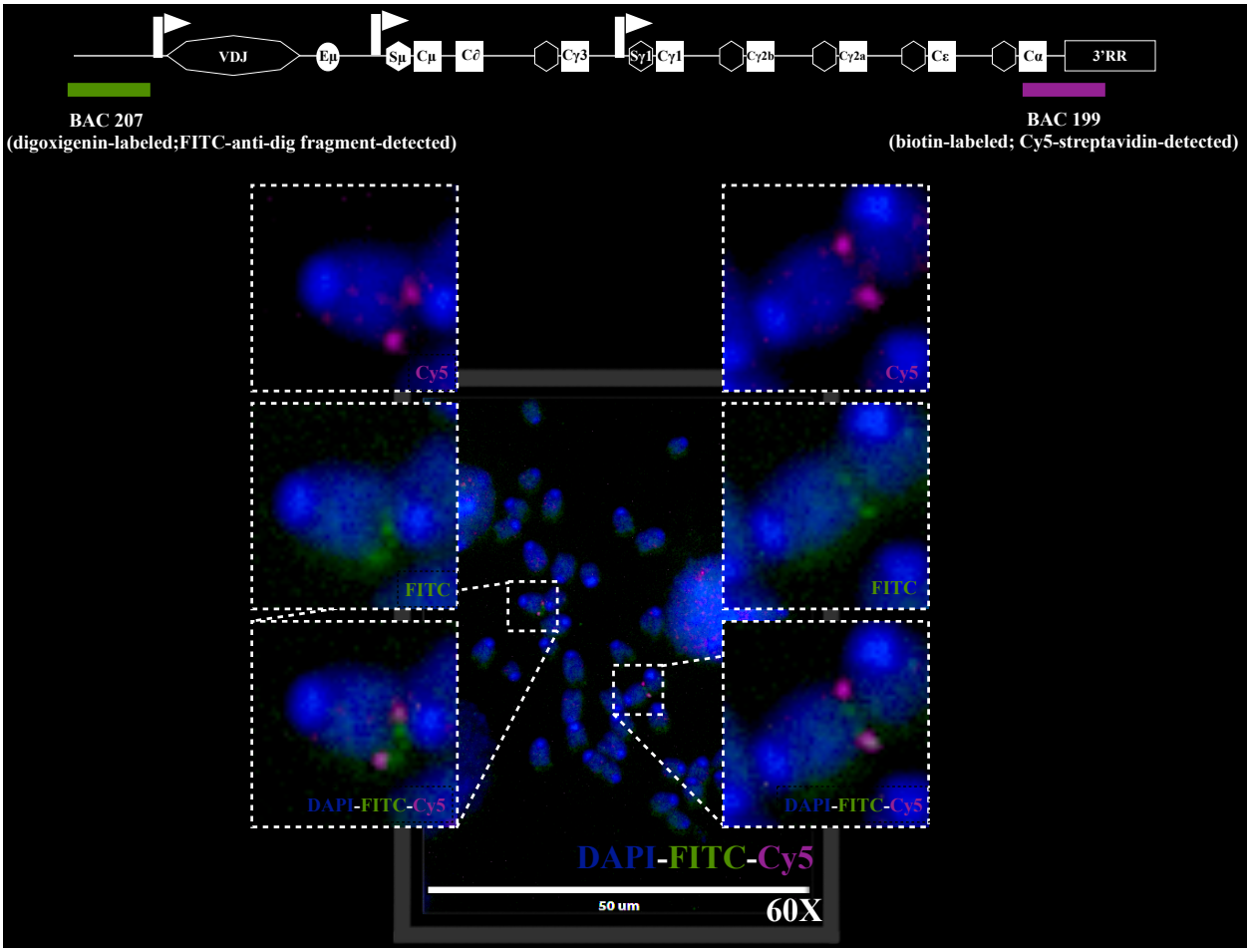


Figure 4.5. 2-Color *IGH* BAC-FISH. Bacterial artifical chromosome-fluorescence in situ hybridization using 2 BAC probes (digoxigenin-labeled upstream of VDJ and more telomeric, and biotin-labeled downstream of Cα and more centromeric) flanking the *IGH* locus on metaphase-arrested B lymphocytes. The probes are then detected using FITC-tagged anti-digoxigenin fragment or Cy5-tagged streptavidin.

IV.3. AID Dependence?

Our experiments in B cells were accompanied by similar experiments in an ES cell background. We wanted to explore the same questions we asked for stimulated B cells in a background deprived of class-switch recombination ability.

We used the same *ROSA26*$^{Cre\text{-ERT2}/+}$; *EXOSC3*$^{COIN/COIN}$ animal background as for our B cell RNA exosome studies. Initially, we had sought to induce RNA exosome activity-deficient B cell proliferation with RP105, which reportedly allows B cells to proliferate without eliciting AID expression (Miyake et al., 1995, Callén et al., 2007). But we turned to ES cells, when these attempts were not satisfactory; moreover, one potent benefit of using B cells is that they might provide a better control system because they negate possible inter-animal differences since all the ES cells we study are from the same embryo, and what is compared is the instability profiles before and after RNA exosome activity abrogation. T-FISH assays on RNA exosome-deficient metaphase-arrested ES cells in ES cells revealed the accumulation of structural chromosomal catastrophic events of the same nature as observed in our B-cell T-FISH assay. In an effort to compare the two different backgrounds, we normalized the levels of instability to their respective cell-specific baseline instability levels (that is, instability level observed without abrogating RNA exosome activity) to create an instability index, and noticed that the latter was more elevated in B cells (4.33) than in ES cells (2.67). These results would therefore suggest that AID might account for the difference between the two instability indices. In this scenario, the instability observed in the non-switching RNA exosome-deficient ES cells may be accounted for by the presence of stress related to proliferation, particularly if transcription-associated and/or replication-associated events at fragile DNA segments prone to forming non-B (including G4)

DNA, TERRA loci, fragile sites (early and/or common), divergently-transcribed transcription start sites, and loci encoding long non-coding RNA or enhancer RNA are occurring (Reddy and Vasquez, 2005, Kim and Jinks-Robertson, 2012, Pefanis et al., 2014, Cusanelli and Chartrand, 2015, Pefanis et al., 2015).

In a different experimental setup, we performed T-FISH assays on activated, then metaphase-arrested, NEDD4$^{-/-}$ B lymphocytes (from fetal liver chimeras) (Sun et al., 2013a). Our results suggest that there is no increase in genomic instability in the absence of NEDD4 in activated B cells. This is interesting, because NEDD4 is thought to promote the activity of AID in B cells, in part by facilitating its interaction with its SPT5 and RNA exosome cofactors. At the same time, although NEDD4 is thought to monoubiquitinate RNA polymerase II (an event that precedes RNA polymerase II polyubiquitination and its degradation), neither RNA polymerase II levels nor immunoglobulin transcript production at switch sequences are decreased in the absence of NEDD4 (Sun et al., 2013a). However, RNA exosome core subunit RRP40 is much less recruited to switch sequence IgS_{μ} in NEDD4 knockdown CH12F3 cells stimulated for isotype switching (Sun et al., 2013a). The observation that no increase in genomic instability was detected in our T-FISH assay of NEDD4$^{-/-}$ primary B cells, even though NEDD4 knockdowns in CH12F3 cells suggest that AID activity is deficient, RNA exosome recruitment is deficient, and transcription is unaffected, indicates that the genomic instability we observed in B cells in the absence of proper RNA exosome function is at least partly AID-dependent. Mating AID knockout mice to our conditional deletion system background would provide us with a useful experimental tool, particularly in light of the hypothesis that AID may mediate epigenetic regulation in pluripotent tissues (Morgan et al., 2004).

IV.4. A Role for Senataxin

As we attempted to understand the workings of the Ig R-loop reaction by investigating the role played by the RNA exosome complex, we decided to study the possible involvement of the putative RNA/DNA 5′ → 3′ helicase senataxin (SETX) and hypothesized it helps unwind RNA:DNA hybrid duplexes to facilitate the RNA exosome-mediated removal/degradation of R-loop RNA:DNA heteroduplex RNA moieties (Figure 4.6).

We saw that in human Burkitt's lymphoma cell line RA.1 (Ramos cells) SETX immunoprecipitates with RNA exosome core subunit RRP45 and with CSR initiator mutagen AID, placing it in the middle of the Ig R-loop reaction. We established a SETX-knockdown system in the murine CH12F3 cell line, where class-switch recombination could be assayed, and found that when *SETX* expression is decreased, CSR becomes impaired. We also found evidence that RNA:DNA hybrids accumulate in SETX-knockdown CH12F3 cells. Our studies in this

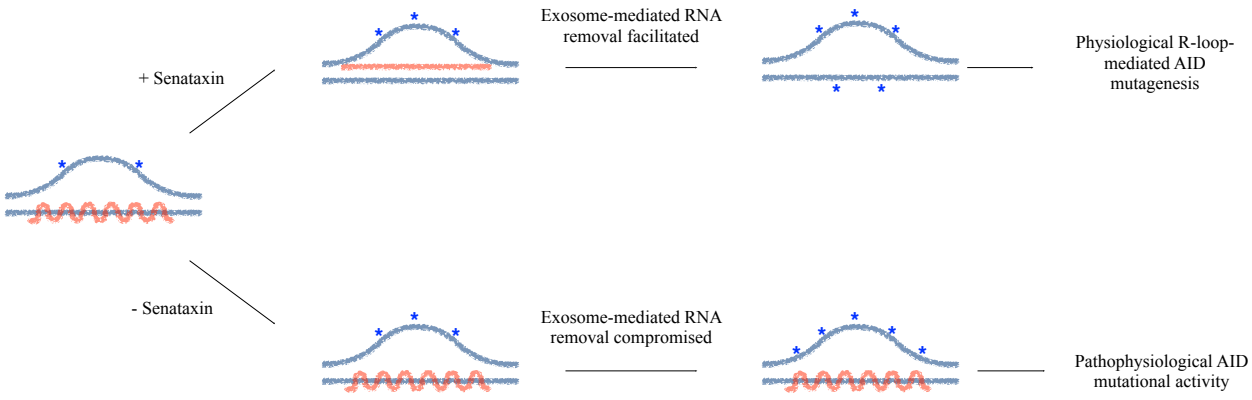


Figure 4.6. SETX Hypothesis. Physiologic SETX levels would permit the unwinding of RNA:DNA hybrids and favor their removal by RNA exosome to resolve R loops and carry out physiological AID-mediated mutagenesis. In the absence of physiological SETX levels, RNA:DNA hybrids would accumulate, not get resolved, and mediate pathophysiological breaks, some of which at least are mediated by AID.

system were limited because no satisfactory commercial antibody against mouse SETX is available. Nevertheless, and importantly, we studied a 5'-most switch sequence S_μ segment for mutational analysis, and observed that in senataxin-deficient backgrounds, the mutational profile at these AID target loci changes, with mutations accruing on the non-transcribed strand relative to the transcribed strand, as would be predicted from our hypothesis (Figure 4.6).

We used a mouse model where *SETX* exon 4 has been deleted, abrogating the function of the protein constitutively (Becherel et al., 2013). Our findings using this model were decidedly more tempered than in the knockdown CH12F3 background. We saw a small, though consistent decrease in the percentage of IgG_1^+ activated B cells from mutant mice compared to wild-type activated B lymphocytes. Moreover, our T-FISH assay on these cells revealed a modest increase in genomic instability. These observations suggest the possible existence of redundant pathways in primary B cells that are somehow not present or less functional in the CH12F3 cells. Several candidates could fall into this category. MTR4 (mRNA transport 4) is a member of the RNA helicase superfamily 2 that forms the TRAMP complex (Trf4/Trf5 poly(A) polymerases, Air1/ Air2 Zn-knuckle RNA-binding proteins, and Mtr4 RNA helicase) known to interact with the RNA exosome complex (Houseley et al., 2006, Houseley and Tollervey, 2009, Schuch et al., 2014). Another potential candidate is aquarius (AQR), which, unlike MTR4, but like SETX, is a putative RNA/DNA helicase member of RNA helicase superfamily 1; AQR activity, like that of SETX, has also been associated to R-loop resolution (Sollier et al., 2014, De et al., 2015). Studies of the Ig R-loop reaction in the absence or impairment of one or a combination of all these helicases would surely be insightful. Given that some of them are reportedly essential, like SETX, a possible way to study them might be the generation of mutants in lieu of outright

constitutive deletion. This evokes another explanation for our mouse phenotype, which is that it may represent a hypomorph of some sort, or a differently spliced gene product, given the absence of pronounced neurological phenotype despite the association of SETX with at least ALS4 and AOA2 (Becherel et al., 2013). This may be worth exploring given reports of alternative splicing and transcription start sites (Fogel et al., 2009).

We initiated the CRISPR-directed generation of *SETX* mutants in CH12F3 cells. Preliminary observations in non-fully characterized lines suggest that the ability of these cells to undergo isotype switching to IgA might be impaired, and, just as in the knockdown experiments, their phenotype is more akin to an AID-mutant control, without discernible growth defects (Data not shown). It remains to be seen what the result of these investigations will be once single clones are isolated. However, if they are successful, they could give us a deeper understanding of the mechanisms behind RNA:DNA unwinding and RNA removal/degradation during the Ig R-loop reaction. In addition, the generation of mice harboring these mutations could provide a system to investigate whether SETX or other helicase activity impairment can precipitate genomic instability in stimulated B lymphocytes, which we were not able to test in our CH12F3 system, given the transformed nature of the cell line.

IV.5. Concluding Remarks

Our work helps demonstrate why disrupting physiological R loop balance may be problematic in activated B lymphocytes (Figure 4.7). While suboptimal class-switch recombination results from the inadequate generation of R loops, the accumulation of the latter without their timely resolution is a source of mutations, double strand breaks, pathophysiological translocations, the occurrence of all of which hinders proper CSR. Failure to adequately resolve

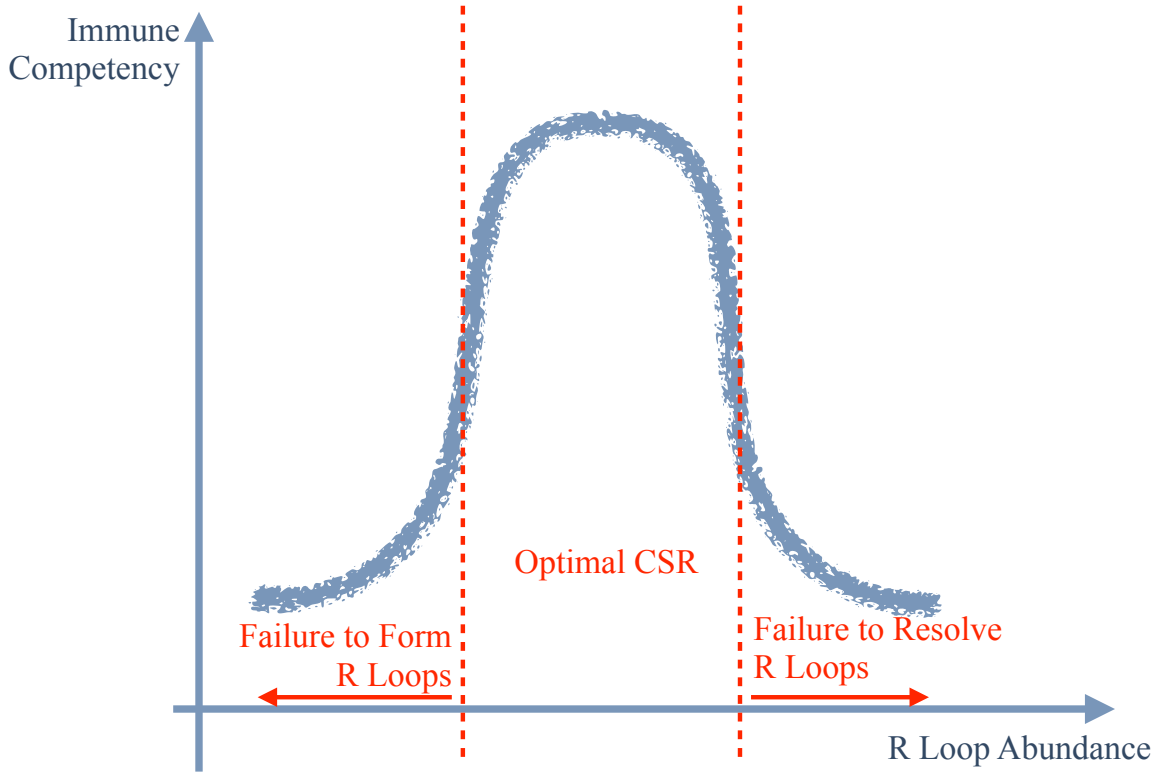


Figure 4.7. R Loop in CSR and Immunocompetency. R loop formation and resolution constitute the mechanism by which a molecular window of instability is opened and closed to allow for carefully orchestrated, time- and space-restricted, physiological mutagenesis by AID. Altering physiological R loop balance is then problematic, and may translate into suboptimal CSR, mutations, double-stranded breaks, pathophysiologic translocations, the combined occurrence of all of which hinders proper CSR. Failure to adequately resolve R loops, as in exosome or senataxin-deficient backgrounds, impairs physiological NHEJ-mediated CSR, as the mutations and breaks that are produced are addressed in ways that are incompatible with CSR.

R loops, as in RNA exosome- or senataxin-deficient backgrounds, impairs physiological NHEJ-mediated CSR, as the mutations and breaks that are produced are addressed in ways that are incompatible with CSR. These catastrophic events may be even compounded by AID-independent breaks in B cells, because of the presence of sensitive sites.

RNA Exosome and B-cell genomic instability. Using a conditional inversion/deletion system that abrogates the synthesis of core RNA exosome subunit RRP40 and prevents the proper assembly of the RNA exosome complex, we observed that B cells stimulated to undergo isotype switching are deficient in the physiological recombination reaction and instead accumulate structural chromosomal instability events. We hypothesized that some of these events are the result of inordinate AID mutational activity at unresolved RNA exosome-related R loops, causing double stranded breaks that fail to be repaired physiologically. One prediction from our R-loop accumulation hypothesis is that these cells are more sensitive and prone to double stranded breaks in conditions of increased mutagenic load. We stressed the cells by exposing them to varying levels of ionizing radiation, and observed that, indeed, RNA exosome-deficient B cells are more likely to accumulate DSBs than RNA exosome-sufficient cells, particularly when confronted with increasing mutagenic stress. Finally, we were also able to identify by spectral karyotyping technology structural changes involving the *IGH*-bearing chromosome 12 as well as chromosomes 1 and 17, which in mouse contain proto-oncogenes.

Senataxin and B-cell genomic instability. In parallel, we made an effort to identify possible RNA exosome co-factors that similarly may be implicated in maintaining the integrity of the B-cell genome. We used both a knockdown, *in-vitro* approach, as well as a mouse model.

More recently, we also initiated a knock-out *in*-vitro system. We report that SETX associates with RNA exosome complex subunit RRP45 as well as AID in a B-lymphocyte cell line. Importantly, we see that SETX-deficient cells exhibit impaired CSR. Moreover, cells lacking physiological SETX activity accumulate RNA:DNA heteroduplexes. Finally, we observe that SETX deficiency leads to an increase in mutations on the template strand, particularly at deoxycitidine residues. These results help us better understand the mechanistic underpinnings of AID targeting to its substrates during isotype switching.

AID dependence. We wanted to test the extent to which our observations were AID-dependent. We used ES cells from our conditional *ROSA26*$^{Cre\text{-}ERT2/+}$; *EXOSC3*$^{COIN/COIN}$ mouse model which proliferate extensively but do not undergo isotype switching and performed T-FISH assays that revealed an increase in instability in the absence of physiological RNA exosome activity. Tellingly, this ES-cell instability does not appear to be as elevated as in B cells. The dysregulation of proper RNA processing mechanisms in B cells may thus lead to oncogenic translocations as a result of the concerted effects of both AID-related mutagenesis and replication- and transcription-related proclivity to DNA damage.

References

ABREU, M. T. 2010. Toll-like receptor signalling in the intestinal epithelium: how bacterial recognition shapes intestinal function. *Nat Rev Immunol,* 10**,** 131-44.

ADELMAN, K. & LIS, J. T. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet,* 13**,** 720-731.

AGUILERA, A. 2002. The connection between transcription and genomic instability. *Embo j,* 21**,** 195-201.

ALLEN, C. D., OKADA, T., TANG, H. L. & CYSTER, J. G. 2007. Imaging of germinal center selection events during affinity maturation. *Science,* 315**,** 528-531.

ALLEN, C. D. C., ANSEL, K. M., LOW, C., LESLEY, R., TAMAMURA, H., FUJII, N. & CYSTER, J. G. 2004. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat Immunol,* 5**,** 943-952.

ALLMANG, C., PETFALSKI, E., PODTELEJNIKOV, A., MANN, M., TOLLERVEY, D. & MITCHELL, P. 1999. The yeast exosome and human PM-Scl are related complexes of 3' → 5' exonucleases. *Genes and Development,* 13**,** 2148-2158.

ANDERSON, J. R., MUKHERJEE, D., MUTHUKUMARASWAMY, K., MORAES, K. C. M., WILUSZ, C. J. & WILUSZ, J. 2006. Sequence-specific RNA binding mediated by the RNase PH domain of components of the exosome. *RNA,* 12**,** 1810-1816.

ANDRULIS, E. D., WERNER, J., NAZARIAN, A., ERDJUMENT-BROMAGE, H., TEMPST, P. & LIS, J. T. 2002. The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. *Nature,* 420**,** 837-841.

ANINDYA, R., AYGUN, O. & SVEJSTRUP, J. Q. 2007. Damage-induced ubiquitylation of human RNA polymerase II by the ubiquitin ligase Nedd4, but not Cockayne syndrome proteins or BRCA1. *Mol Cell,* 28**,** 386-97.

AOUFOUCHI, S., FAILI, A., ZOBER, C., D'ORLANDO, O., WELLER, S., WEILL, J. C. & REYNAUD, C. A. 2008. Proteasomal degradation restricts the nuclear lifespan of AID. *Journal of Experimental Medicine,* 205**,** 1357-1368.

ARAKAWA, H., HAUSCHILD, J. & BUERSTEDDE, J.-M. 2002. Requirement of the Activation-Induced Deaminase (AID) Gene for Immunoglobulin Gene Conversion. *Science,* 295**,** 1301-1306.

ARAKI, Y., TAKAHASHI, S., KOBAYASHI, T., KAJIHO, H., HOSHINO, S. I. & KATADA, T. 2001. Ski7p G protein interacts with the exosome and the ski complex for 3′-to-5′ mRNA decay in yeast. *EMBO Journal,* 20**,** 4684-4693.

ARNOLD, L. W., LOCASCIO, N. J., LUTZ, P. M., PENNELL, C. A., KLAPPER, D. & HAUGHTON, G. 1983. Antigen-induced lymphomagenesis: identification of a murine B cell lymphoma with known antigen specificity. *J Immunol,* 131**,** 2064-8.

ASSELIN-PATUREL, C., BOONSTRA, A., DALOD, M., DURAND, I., YESSAAD, N., DEZUTTER-DAMBUYANT, C., VICARI, A., O'GARRA, A., BIRON, C., BRIERE, F.

& TRINCHIERI, G. 2001. Mouse type I IFN-producing cells are immature APCs with plasmacytoid morphology. *Nat Immunol,* 2**,** 1144-50.

ASSELIN-PATUREL, C., BRIZARD, G., CHEMIN, K., BOONSTRA, A., O'GARRA, A., VICARI, A. & TRINCHIERI, G. 2005. Type I interferon dependence of plasmacytoid dendritic cell activation and migration. *J Exp Med,* 201**,** 1157-67.

BALK, B., MAICHER, A., DEES, M., KLERMUND, J., LUKE-GLASER, S., BENDER, K. & LUKE, B. 2013. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol,* 20**,** 1199-1205.

BANNARD, O., HORTON, R. M., ALLEN, C. D., AN, J., NAGASAWA, T. & CYSTER, J. G. 2013. Germinal center centroblasts transition to a centrocyte phenotype according to a timed program and depend on the dark zone for effective selection. *Immunity,* 39**,** 912-24.

BARCHET, W., CELLA, M., ODERMATT, B., ASSELIN-PATUREL, C., COLONNA, M. & KALINKE, U. 2002. Virus-induced interferon alpha production by a dendritic cell subset in the absence of feedback signaling in vivo. *J Exp Med,* 195**,** 507-16.

BARLOW, J. H., FARYABI, R. B., CALLEN, E., WONG, N., MALHOWSKI, A., CHEN, H. T., GUTIERREZ-CRUZ, G., SUN, H. W., MCKINNON, P., WRIGHT, G., CASELLAS, R., ROBBIANI, D. F., STAUDT, L., FERNANDEZ-CAPETILLO, O. & NUSSENZWEIG, A. 2013. Identification of early replicating fragile sites that contribute to genome instability. *Cell,* 152**,** 620-32.

BARRETO, V., REINA-SAN-MARTIN, B., RAMIRO, A. R., MCBRIDE, K. M. & NUSSENZWEIG, M. C. 2003. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Molecular Cell,* 12**,** 501-508.

BASSO, K. & DALLA-FAVERA, R. 2015. Germinal centres and B cell lymphomagenesis. *Nat Rev Immunol,* 15**,** 172-184.

BASU, U., CHAUDHURI, J., ALPERT, C., DUTT, S., RANGANATH, S., LI, G., SCHRUM, J. P., MANIS, J. P. & ALT, F. W. 2005. The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature,* 438**,** 508-11.

BASU, U., MENG, F. L., KEIM, C., GRINSTEIN, V., PEFANIS, E., ECCLESTON, J., ZHANG, T., MYERS, D., WASSERMAN, C. R., WESEMANN, D. R., JANUSZYK, K., GREGORY, R. I., DENG, H., LIMA, C. D. & ALT, F. W. 2011. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell,* 144**,** 353-63.

BASU, U., WANG, Y. & ALT, F. W. 2008. Evolution of phosphorylation-dependent regulation of activation-induced cytidine deaminase. *Mol Cell,* 32**,** 285-91.

BECHEREL, O. J., YEO, A. J., STELLATI, A., HENG, E. Y., LUFF, J., SURAWEERA, A. M., WOODS, R., FLEMING, J., CARRIE, D., MCKINNEY, K., XU, X., DENG, C. &

LAVIN, M. F. 2013. Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genet,* 9**,** e1003435.

BEHRING & KITASATO 1890. Ueber das Zustandekommen der Diphtherie-Immunität und der Tetanus-Immunität bei Thieren. *Dtsch med Wochenschr,* 16**,** 1113-1114.

BELETSKII, A. & BHAGWAT, A. S. 1996. Transcription-induced mutations: Increase in C to T mutations in the nontranscribed strand during transcription in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America,* 93**,** 13919-13924.

BEREK, C. & MILSTEIN, C. 1987. Mutation drift and repertoire shift in the maturation of the immune response. *Immunol Rev,* 96**,** 23-41.

BERGQVIST, P., GÄRDBY, E., STENSSON, A., BEMARK, M. & LYCKE, N. Y. 2006. Gut IgA Class Switch Recombination in the Absence of CD40 Does Not Occur in the Lamina Propria and Is Independent of Germinal Centers. *The Journal of Immunology,* 177**,** 7772-7783.

BERNARD, O., HOZUMI, N. & TONEGAWA, S. 1978. Sequences of mouse immunoglobulin light chain genes before and after somatic changes. *Cell,* 15**,** 1133-44.

BESMER, E., MARKET, E. & PAPAVASILIOU, F. N. 2006. The transcription elongation complex directs activation-induced cytidine deaminase-mediated DNA deamination. *Molecular and Cellular Biology,* 26**,** 4378-4385.

BETZ, A. G., MILSTEIN, C., GONZÁLEZ-FERNÁNDEZ, A., PANNELL, R., LARSON, T. & NEUBERGER, M. S. 1994. Elements regulating somatic hypermutation of an immunoglobulin κ gene: Critical role for the intron enhancer/matrix attachment region. *Cell,* 77**,** 239-248.

BHATIA, V., BARROSO, S. I., GARCÍA-RUBIO, M. L., TUMINI, E., HERRERA-MOYANO, E. & AGUILERA, A. 2014. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature,* 511**,** 362-365.

BIGLEY, V., HANIFFA, M., DOULATOV, S., WANG, X.-N., DICKINSON, R., MCGOVERN, N., JARDINE, L., PAGAN, S., DIMMICK, I., CHUA, I., WALLIS, J., LORDAN, J., MORGAN, C., KUMARARATNE, D. S., DOFFINGER, R., VAN DER BURG, M., VAN DONGEN, J., CANT, A., DICK, J. E., HAMBLETON, S. & COLLIN, M. 2011. The human syndrome of dendritic cell, monocyte, B and NK lymphoid deficiency. *The Journal of Experimental Medicine,* 208**,** 227-234.

BLANDER, J. M. & SANDER, L. E. 2012. Beyond pattern recognition: five immune checkpoints for scaling the microbial threat. *Nat Rev Immunol,* 12**,** 215-25.

BOBOILA, C., ALT, F. W. & SCHWER, B. 2012. Chapter One - Classical and Alternative End-Joining Pathways for Repair of Lymphocyte-Specific and General DNA Double-Strand Breaks. *In:* FREDERICK, W. A. (ed.) *Advances in Immunology.* Academic Press.

BOEHM, T. 2012. Evolution of vertebrate immunity. *Curr Biol,* 22**,** R722-32.

BOGUSLAWSKI, S. J., SMITH, D. E., MICHALAK, M. A., MICKELSON, K. E., YEHLE, C. O., PATTERSON, W. L. & CARRICO, R. J. 1986. Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. *J Immunol Methods,* 89**,** 123-30.

BOULÉ, J. B. & ZAKIAN, V. A. 2007. The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates. *Nucleic Acids Res,* 35**,** 5809-18.

BRANSTEITTER, R., PHAM, P., SCHARFFT, M. D. & GOODMAN, M. F. 2003. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 4102-4107.

BRINKMANN, V., REICHARD, U., GOOSMANN, C., FAULER, B., UHLEMANN, Y., WEISS, D. S., WEINRAUCH, Y. & ZYCHLINSKY, A. 2004. Neutrophil extracellular traps kill bacteria. *Science,* 303**,** 1532-5.

BRISENO, C. G., MURPHY, T. L. & MURPHY, K. M. 2014. Complementary diversification of dendritic cells and innate lymphoid cells. *Curr Opin Immunol,* 29**,** 69-78.

BUNTING, S. F. & NUSSENZWEIG, A. 2013. End-joining, translocations and cancer. *Nat Rev Cancer,* 7, 443-54

BURDETTE, D. L. & VANCE, R. E. 2013. STING and the innate immune response to nucleic acids in the cytosol. *Nat Immunol,* 14**,** 19-26.

BURNET, F. M. 1976. A modification of Jerne's theory of antibody production using the concept of clonal selection (Reprinted from Australian Journal of Science, vol 20, 1957). *CA Cancer J Clin,* 26**,** 119-21.

BUSSLINGER, M. 2004. Transcriptional control of early B cell development. *Annu Rev Immunol,* 22**,** 55-79.

CALADO, D. P., SASAKI, Y., GODINHO, S. A., PELLERIN, A., KOCHERT, K., SLECKMAN, B. P., DE ALBORAN, I. M., JANZ, M., RODIG, S. & RAJEWSKY, K. 2012. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat Immunol,* 13**,** 1092-1100.

CALLAHAN, K. P. & BUTLER, J. S. 2010. TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. *Journal of Biological Chemistry,* 285**,** 3540-3547.

CALLÉN, E., JANKOVIC, M., DIFILIPPANTONIO, S., DANIEL, J. A., CHEN, H.-T., CELESTE, A., PELLEGRINI, M., MCBRIDE, K., WANGSA, D., BREDEMEYER, A. L., SLECKMAN, B. P., RIED, T., NUSSENZWEIG, M. & NUSSENZWEIG, A. 2007. ATM Prevents the Persistence and Propagation of Chromosome Breaks in Lymphocytes. *Cell,* 130**,** 63-75.

CANNONS, J. L., QI, H., LU, K. T., DUTTA, M., GOMEZ-RODRIGUEZ, J., CHENG, J., WAKELAND, E. K., GERMAIN, R. N. & SCHWARTZBERG, P. L. 2010. Optimal

germinal center responses require a multistage T cell:B cell adhesion process involving integrins, SLAM-associated protein, and CD84. *Immunity,* 32**,** 253-65.

CARON, G., LE GALLOU, S., LAMY, T., TARTE, K. & FEST, T. 2009. CXCR4 expression functionally discriminates centroblasts versus centrocytes within human germinal center B cells. *J. Immunol.,* 182**,** 7595-7602.

CARULLI, G., CANNIZZO, E., ZUCCA, A., BUDA, G., ORCIUOLO, E., MARINI, A. & PETRINI, M. 2008. CD45 expression in low-grade B-cell non-Hodgkin's lymphomas. *Leuk Res,* 32**,** 263-7.

CASTELLANO-POZO, M., SANTOS-PEREIRA, J., RONDÓN, A., BARROSO, S., ANDÚJAR, E., PÉREZ-ALEGRE, M., GARCÍA-MUSE, T. & AGUILERA, A. 2013. R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Molecular Cell,* 52**,** 583-590.

CHATTERJEE, S., JU, Z., HASSAN, R., VOLPI, S. A., EMELYANOV, A. V. & BIRSHTEIN, B. K. 2011. Dynamic Changes in Binding of Immunoglobulin Heavy Chain 3′ Regulatory Region to Protein Factors during Class Switching. *Journal of Biological Chemistry,* 286**,** 29303-29312.

CHAUDHURI, J. & ALT, F. W. 2004. Class-switch recombination: Interplay of transcription, DNA deamination and DNA repair. *Nature Reviews Immunology,* 4**,** 541-552.

CHAUDHURI, J., BASU, U., ZARRIN, A., YAN, C., FRANCO, S., PERLOT, T., VUONG, B., WANG, J., PHAN, R. T., DATTA, A., MANIS, J. & ALT, F. W. 2007. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv Immunol,* 94**,** 157-214.

CHAUDHURI, J., KHUONG, C. & ALT, F. W. 2004. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature,* 430**,** 992-8.

CHAUDHURI, J., TIAN, M., KHUONG, C., CHUA, K., PINAUD, E. & ALT, F. W. 2003. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature,* 422**,** 726-730.

CHÁVEZ, S. & AGUILERA, A. 1997. The yeast HPR1 gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. *Genes Dev,* 11**,** 3459-70.

CHELICO, L., PHAM, P. & GOODMAN, M. F. 2009. Stochastic properties of processive cytidine DNA deaminases AID and APOBEC3G. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 364**,** 583-593.

CHEN, K. M., HARJES, E., GROSS, P. J., FAHMY, A., LU, Y., SHINDO, K., HARRIS, R. S. & MATSUO, H. 2008. Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature,* 452**,** 116-9.

CHEN, Y.-Z., BENNETT, C. L., HUYNH, H. M., BLAIR, I. P., PULS, I., IROBI, J., DIERICK, I., ABEL, A., KENNERSON, M. L., RABIN, B. A., NICHOLSON, G. A., AUER-

GRUMBACH, M., WAGNER, K., DE JONGHE, P., GRIFFIN, J. W., FISCHBECK, K. H., TIMMERMAN, V., CORNBLATH, D. R. & CHANCE, P. F. 2004. DNA/RNA Helicase Gene Mutations in a Form of Juvenile Amyotrophic Lateral Sclerosis (ALS4). *The American Journal of Human Genetics,* 74**,** 1128-1135.

CHENG, H. L., VUONG, B. Q., BASU, U., FRANKLIN, A., SCHWER, B., ASTARITA, J., PHAN, R. T., DATTA, A., MANIS, J., ALT, F. W. & CHAUDHURI, J. 2009. Integrity of the AID serine-38 phosphorylation site is critical for class switch recombination and somatic hypermutation in mice. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 2717-2722.

CHEUNG, A. C. & CRAMER, P. 2012. A movie of RNA polymerase II transcription. *Cell,* 149**,** 1431-7.

CHIEN, Y. H. & DAVIDSON, N. 1978. RNA:DNA hybrids are more stable than DNA:DNA duplexes in concentrated perchlorate and trichloroacetate solutions. *Nucleic Acids Res,* 5**,** 1627-37.

CHOVATIYA, R. & MEDZHITOV, R. 2014. Stress, inflammation, and defense of homeostasis. *Mol Cell,* 54**,** 281-8.

COBALEDA, C., SCHEBESTA, A., DELOGU, A. & BUSSLINGER, M. 2007. Pax5: the guardian of B cell identity and function. *Nat Immunol,* 8**,** 463-470.

COLAK, D., ZANINOVIC, N., COHEN, M. S., ROSENWAKS, Z., YANG, W. Y., GERHARDT, J., DISNEY, M. D. & JAFFREY, S. R. 2014. Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science,* 343**,** 1002-1005.

CONN, G. L., BROWN, T. & LEONARD, G. A. 1999. The crystal structure of the RNA/DNA hybrid r(GAAGAGAAGC)·d(GCTTCTCTTC) shows significant differences to that found in solution. *Nucleic Acids Research,* 27**,** 555-561.

CONTICELLO, S. G. 2008. The AID/APOBEC family of nucleic acid mutators. *Genome Biol,* 9**,** 229.

CONTICELLO, S. G., LANGLOIS, M. A., YANG, Z. & NEUBERGER, M. S. 2007. DNA Deamination in Immunity: AID in the Context of Its APOBEC Relatives. *In:* FREDERICK, W. A. & TASUKU, H. (eds.) *Advances in Immunology.* Academic Press.

CONTICELLO, S. G., THOMAS, C. J. F., PETERSEN-MAHRT, S. K. & NEUBERGER, M. S. 2005. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Molecular Biology and Evolution,* 22**,** 367-377.

COOPER, M. D. 2015. The early history of B cells. *Nat Rev Immunol,* 15**,** 191-197.

COOPER, M. D., PETERSON, R. D. & GOOD, R. A. 1965. Delineation of the thymic and bursal lymphoid systems in the chicken. *Nature,* 205**,** 143-6.

COSTANTINO, L. & KOSHLAND, D. 2015. The Yin and Yang of R-loop biology. *Current Opinion in Cell Biology,* 34**,** 39-45.

COSTELLO, J. L., STEAD, J. A., FEIGENBUTZ, M., JONES, R. M. & MITCHELL, P. 2011. The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D small nucleolar RNA 3′-maturation. *Journal of Biological Chemistry,* 286**,** 4535-4543.

CUSANELLI, E. & CHARTRAND, P. 2015. Telomeric repeat-containing RNA TERRA: a noncoding RNA connecting telomere biology to genome integrity. *Front Genet,* 6**,** 143.

DANIELS, G. A. & LIEBER, M. R. 1995. RNA:DNA complex formation upon transcription of immunoglobulin switch regions: Implications for the mechanism and regulation of class switch recombination. *Nucleic Acids Research,* 23**,** 5006-5011.

DE, I., BESSONOV, S., HOFELE, R., DOS SANTOS, K., WILL, C. L., URLAUB, H., LÜHRMANN, R. & PENA, V. 2015. The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. *Nat Struct Mol Biol,* 22**,** 138-144.

DE SILVA, N. S. & KLEIN, U. 2015. Dynamics of B cells in germinal centres. *Nat Rev Immunol,* 15**,** 137-148.

DE YÉBENES, V. G., BELVER, L., PISANO, D. G., GONZÁLEZ, S., VILLASANTE, A., CROCE, C., HE, L. & RAMIRO, A. R. 2008. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *Journal of Experimental Medicine,* 205**,** 2199-2206.

DELKER, R. K., ZHOU, Y., STRIKOUDIS, A., STEBBINS, C. E. & PAPAVASILIOU, F. N. 2013. Solubility-based genetic screen identifies RING finger protein 126 as an E3 ligase for activation-induced cytidine deaminase. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 1029-1034.

DI NOIA, J. M. & NEUBERGER, M. S. 2007. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.,* 76**,** 1-22.

DICKERSON, S. K., MARKET, E., BESMER, E. & PAPAVASILIOU, F. N. 2003. AID mediates hypermutation by deaminating single stranded DNA. *Journal of Experimental Medicine,* 197**,** 1291-1296.

DILDROP, R., BRÜGGEMANN, M., RADBRUCH, A., RAJEWSKY, K. & BEYREUTHER, K. 1982. Immunoglobulin V region variants in hybridoma cells. II. Recombination between V genes. *EMBO Journal,* 1**,** 635-640.

DORSETT, Y., MCBRIDE, K. M., JANKOVIC, M., GAZUMYAN, A., THAI, T. H., ROBBIANI, D. F., DI VIRGILIO, M., SAN-MARTIN, B. R., HEIDKAMP, G., SCHWICKERT, T. A., EISENREICH, T., RAJEWSKY, K. & NUSSENZWEIG, M. C. 2008. MicroRNA-155 Suppresses Activation-Induced Cytidine Deaminase-Mediated Myc-Igh Translocation. *Immunity,* 28**,** 630-638.

DORSHKIND, K. & MONTECINO-RODRIGUEZ, E. 2007. Fetal B-cell lymphopoiesis and the emergence of B-1-cell potential. *Nat Rev Immunol,* 7**,** 213-219.

DUKE, J. L., LIU, M., YAARI, G., KHALIL, A. M., TOMAYKO, M. M., SHLOMCHIK, M. J., SCHATZ, D. G. & KLEINSTEIN, S. H. 2013. Multiple transcription factor binding sites predict AID targeting in non-Ig genes. *Journal of Immunology,* 190**,** 3878-3888.

DUNNICK, W., HERTZ, G. Z., SCAPPINO, L. & GRITZMACHER, C. 1993. DNA sequences at immunoglobulin switch region recombination sites. *Nucleic Acids Research,* 21**,** 365-372.

DUQUETTE, M. L., HANDA, P., VINCENT, J. A., TAYLOR, A. F. & MAIZELS, N. 2004. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev,* 18**,** 1618-29.

DURANDY, A., KRACKER, S. & FISCHER, A. 2013. Primary antibody deficiencies. *Nat Rev Immunol,* 13**,** 519-533.

EARLY, P., HUANG, H., DAVIS, M., CALAME, K. & HOOD, L. 1980. An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. *Cell,* 19**,** 981-92.

EGLI, M., USMAN, N. & RICH, A. 1993. Conformational influence of the ribose 2'-hydroxyl group: crystal structures of DNA-RNA chimeric duplexes. *Biochemistry,* 32**,** 3221-37.

EGLI, M., USMAN, N., ZHANG, S. G. & RICH, A. 1992. Crystal structure of an Okazaki fragment at 2-A resolution. *Proceedings of the National Academy of Sciences,* 89**,** 534-538.

EISEN, H. N. & SISKIND, G. W. 1964. Variations in affinities of antibodies during the immune response. *Biochemistry,* 3**,** 996-1008.

ELLYARD, J. I., BENK, A. S., TAYLOR, B., RADA, C. & NEUBERGER, M. S. 2011. The dependence of Ig class-switching on the nuclear export sequence of AID likely reflects interaction with factors additional to Crm1 exportin. *European Journal of Immunology,* 41**,** 485-490.

ENCHEV, R. I., SCHULMAN, B. A. & PETER, M. 2015. Protein neddylation: beyond cullin-RING ligases. *Nat Rev Mol Cell Biol,* 16**,** 30-44.

FAGRAEUS, A. 1948. The plasma cellular reaction and its relation to the formation of antibodies in vitro. *J Immunol,* 58**,** 1-13.

FAIRMAN-WILLIAMS, M. E., GUENTHER, U.-P. & JANKOWSKY, E. 2010. SF1 and SF2 helicases: family matters. *Current opinion in structural biology,* 20**,** 313-324.

FEIJOO, P., DOMINGUEZ, D., TUSELL, L. & GENESCA, A. 2014. Telomere-dependent genomic integrity: evolution of the fusion-bridge-breakage cycle concept. *Curr Pharm Des,* 20**,** 6375-85.

FILLATREAU, S., SWEENIE, C. H., MCGEACHY, M. J., GRAY, D. & ANDERTON, S. M. 2002. B cells regulate autoimmunity by provision of IL-10. *Nature Immunology,* 3**,** 944-950.

FLAJNIK, M. F. 2002. Comparative analyses of immunoglobulin genes: surprises and portents. *Nat Rev Immunol,* 2**,** 688-698.

FLAJNIK, M. F. & KASAHARA, M. 2010. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat Rev Genet,* 11**,** 47-59.

FOGEL, B. L., LEE, J. Y. & PERLMAN, S. 2009. Aberrant splicing of the senataxin gene in a patient with ataxia with oculomotor apraxia type 2. *Cerebellum,* 8**,** 448-53.

FRITZ, E. L., ROSENBERG, B. R., LAY, K., MIHAILOVIC, A., TUSCHL, T. & PAPAVASILIOU, F. N. 2013. A comprehensive analysis of the effects of the deaminase AID on the transcriptome and methylome of activated B cells. *Nat Immunol,* 14**,** 749-755.

FUGMANN, S. D. & SCHATZ, D. G. 2002. One AID to Unite Them All. *Science,* 295**,** 1244-1245.

FURTH, J. J., HURWITZ, J. & GOLDMANN, M. 1961. The directing role of DNA in RNA synthesis. *Biochem Biophys Res Commun,* 4**,** 362-7.

GALLOWAY, J. L. & ZON, L. I. 2003. Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo. *Curr Top Dev Biol,* 53**,** 139-58.

GANESH, K., ADAM, S., TAYLOR, B., SIMPSON, P., RADA, C. & NEUBERGER, M. 2011. CTNNBL1 is a novel nuclear localization sequence-binding protein that recognizes RNA-splicing factors CDC5L and Prp31. *Journal of Biological Chemistry,* 286**,** 17091-17102.

GANGULY, D., HAAK, S., SISIRAK, V. & REIZIS, B. 2013. The role of dendritic cells in autoimmunity. *Nat Rev Immunol,* 13**,** 566-577.

GEISBERGER, R., RADA, C. & NEUBERGER, M. S. 2009. The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 6736-6741.

GEISSMANN, F., MANZ, M. G., JUNG, S., SIEWEKE, M. H., MERAD, M. & LEY, K. 2010. Development of Monocytes, Macrophages, and Dendritic Cells. *Science,* 327**,** 656-661.

GEKAS, C., DIETERLEN-LIEVRE, F., ORKIN, S. H. & MIKKOLA, H. K. 2005. The placenta is a niche for hematopoietic stem cells. *Dev Cell,* 8**,** 365-75.

GINNO, P. A., LOTT, P. L., CHRISTENSEN, H. C., KORF, I. & CHEDIN, F. 2012. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell,* 45**,** 814-25.

GÓMEZ-GONZÁLEZ, B. & AGUILERA, A. 2007. Activation-induced cytidine deaminase action is strongly stimulated by mutations of the THO complex. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 8409-8414.

GOTLIB, J. & COOLS, J. 2008. Five years since the discovery of FIP1L1-PDGFRA: What we have learned about the fusion and other molecularly defined eosinophilias. *Leukemia,* 22**,** 1999-2010.

GOWRISHANKAR, J. & HARINARAYANAN, R. 2004. Why is transcription coupled to translation in bacteria? *Mol Microbiol,* 54**,** 598-603.

GREIMANN, J. C. & LIMA, C. D. 2008. Chapter 10 Reconstitution of RNA Exosomes from Human and Saccharomyces cerevisiae. Cloning, Expression, Purification, and Activity Assays. *In:* MAQUAT, L. E. & KILEDJIAN, M. (eds.) *Methods in Enzymology.*

GROH, M. & GROMAK, N. 2014. Out of balance: R-loops in human disease. *PLoS Genet,* 10**,** e1004630.

GURTLER, C. & BOWIE, A. G. 2013. Innate immune detection of microbial nucleic acids. *Trends Microbiol,* 21**,** 413-20.

HAAS, K. M., POE, J. C., STEEBER, D. A. & TEDDER, T. F. 2005. B-1a and B-1b cells exhibit distinct developmental requirements and have unique functional roles in innate and adaptive immunity to S. pneumoniae. *Immunity,* 23**,** 7-18.

HAEUSLER, A. R., DONNELLY, C. J., PERIZ, G., SIMKO, E. A. J., SHAW, P. G., KIM, M. S., MARAGAKIS, N. J., TRONCOSO, J. C., PANDEY, A., SATTLER, R., ROTHSTEIN, J. D. & WANG, J. 2014. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature,* 507**,** 195-200.

HALBACH, F., REICHELT, P., RODE, M. & CONTI, E. 2013. The yeast ski complex: Crystal structure and rna channeling to the exosome complex. *Cell,* 154**,** 814-826.

HARREMAN, M., TASCHNER, M., SIGURDSSON, S., ANINDYA, R., REID, J., SOMESH, B., KONG, S. E., BANKS, C. A. S., CONAWAY, R. C., CONAWAY, J. W. & SVEJSTRUP, J. Q. 2009. Distinct ubiquitin ligases act sequentially for RNA polymerase II polyubiquitylation. *Proceedings of the National Academy of Sciences,* 106**,** 20705-20710.

HAUSER, A. E., JUNT, T., MEMPEL, T. R., SNEDDON, M. W., KLEINSTEIN, S. H., HENRICKSON, S. E., VON ANDRIAN, U. H., SHLOMCHIK, M. J. & HABERMAN, A. M. 2007. Definition of germinal-center B cell migration in vivo reveals predominant intrazonal circulation patterns. *Immunity,* 26**,** 655-67.

HEDRICK, S. M. 2004. The acquired immune system: a vantage from beneath. *Immunity,* 21**,** 607-15.

HERMESH, T., MOLTEDO, B., MORAN, T. M. & LOPEZ, C. B. 2010. Antiviral instruction of bone marrow leukocytes during respiratory viral infections. *Cell Host Microbe,* 7**,** 343-53.

HILKENS, J., CUYPERS, H. T., SELTEN, G., KROEZEN, V., HILGERS, J. & BERNS, A. 1986. Genetic mapping of Pim-1 putative oncogene to mouse chromosome 17. *Somat Cell Mol Genet,* 12**,** 81-8.

HIRANO, M., DAS, S., GUO, P. & COOPER, M. D. 2011. The evolution of adaptive immunity in vertebrates. *Adv Immunol,* 109**,** 125-57.

HJERPE, R., THOMAS, Y., CHEN, J., ZEMLA, A., CURRAN, S., SHPIRO, N., DICK, L. R. & KURZ, T. 2012. Changes in the ratio of free NEDD8 to ubiquitin triggers NEDDylation by ubiquitin enzymes. *Biochem J,* 441**,** 927-36.

HODGKIN, P. D., LEE, J. H. & LYONS, A. B. 1996. B cell differentiation and isotype switching is related to division cycle number. *The Journal of Experimental Medicine,* 184**,** 277-281.

HONJO, T. & KATAOKA, T. 1978. Organization of immunoglobulin heavy chain genes and allelic deletion model. *Proc Natl Acad Sci U S A,* 75**,** 2140-4.

HOUSELEY, J., LACAVA, J. & TOLLERVEY, D. 2006. RNA-quality control by the exosome. *Nature Reviews Molecular Cell Biology,* 7**,** 529-539.

HOUSELEY, J. & TOLLERVEY, D. 2009. The many pathways of RNA degradation. *Cell,* 136**,** 763-76.

HOZUMI, N. & TONEGAWA, S. 1976. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci U S A,* 73**,** 3628-32.

HUANG, C., GONZALEZ, D. G., COTE, C. M., JIANG, Y., HATZI, K., TEATER, M., DAI, K., HLA, T., HABERMAN, A. M. & MELNICK, A. 2014. The BCL6 RD2 domain governs commitment of activated B cells to form germinal centers. *Cell Rep,* 8**,** 1497-508.

HUERTAS, P. & AGUILERA, A. 2003. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol Cell,* 12**,** 711-21.

HUNG, S.-H., YU, Q., GRAY, D. M. & RATLIFF, R. L. 1994. Evidence from CD spectra that d(purine)-r(pyrimidine) and r(purine)-d(pyrimidine) hybrids are in different structural classes. *Nucleic Acids Research,* 22**,** 4326-4334.

ICHINOHE, T., PANG, I. K. & IWASAKI, A. 2010. Influenza virus activates inflammasomes via its intracellular M2 ion channel. *Nat Immunol,* 11**,** 404-10.

IDE, H. & KOTERA, M. 2004. Human DNA glycosylases involved in the repair of oxidatively damaged DNA. *Biol Pharm Bull,* 27**,** 480-5.

INMAN, K. E. & DOWNS, K. M. 2007. The murine allantois: emerging paradigms in development of the mammalian umbilical cord and its relation to the fetus. *Genesis,* 45**,** 237-58.

ITO, S., NAGAOKA, H., SHINKURA, R., BEGUM, N., MURAMATSU, M., NAKATA, M. & HONJO, T. 2004. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 1975-1980.

IWASAKI, A. & MEDZHITOV, R. 2015. Control of adaptive immunity by the innate immune system. *Nat Immunol,* 16**,** 343-353.

JANEWAY, C. A., JR. 1989. Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb Symp Quant Biol,* 54 Pt 1**,** 1-13.

JANKOWSKY, E. 2011. RNA Helicases at work: binding and rearranging. *Trends in biochemical sciences,* 36**,** 19-29.

JANUSZYK, K. & LIMA, C. D. 2014. The eukaryotic RNA exosome. *Current Opinion in Structural Biology,* 24**,** 132-140.

JARMUZ, A., CHESTER, A., BAYLISS, J., GISBOURNE, J., DUNHAM, I., SCOTT, J. & NAVARATNAM, N. 2002. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics,* 79**,** 285-96.

JERNE, N. K. 1951. A study of avidity based on rabbit skin responses to diphtheria toxin-antitoxin mixtures. *Acta Pathol Microbiol Scand Suppl,* 87**,** 1-183.

JONKERS, I. & LIS, J. T. 2015. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol,* 16**,** 167-177.

KANEHIRO, Y., TODO, K., NEGISHI, M., FUKUOKA, J., GAN, W., HIKASA, T., KAGA, Y., TAKEMOTO, M., MAGARI, M., LI, X., MANLEY, J. L., OHMORI, H. & KANAYAMA, N. 2012. Activation-induced cytidine deaminase (AID)-dependent somatic hypermutation requires a splice isoform of the serine/arginine-rich (SR) protein SRSF1. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 1216-1221.

KANTOR, A. B. & HERZENBERG, L. A. 1993. Origin of murine B cell lineages. *Annu Rev Immunol,* 11**,** 501-38.

KÄRRE, K., LJUNGGREN, H. G., PIONTEK, G. & KIESSLING, R. 1986. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature,* 319**,** 675-8.

KATZ, S. I., PARKER, D. & TURK, J. L. 1974. B-cell suppression of delayed hypersensitivity reactions. *Nature,* 251**,** 550-1.

KAUFMAN, J. 2010. Evolution and immunity. *Immunology,* 130**,** 459-62.

KAVLI, B., ANDERSEN, S., OTTERLEI, M., LIABAKK, N. B., IMAI, K., FISCHER, A., DURANDY, A., KROKAN, H. E. & SLUPPHAUG, G. 2005. B cells from hyper-IgM patients carrying UNG mutations lack ability to remove uracil from ssDNA and have elevated genomic uracil. *J Exp Med,* 201**,** 2011-21.

KAWAI, T. & AKIRA, S. 2010. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol,* 11**,** 373-84.

KEIM, C., KAZADI, D., ROTHSCHILD, G. & BASU, U. 2013. Regulation of AID, the B-cell genome mutator. *Genes Dev,* 27**,** 1-17.

KENTER, A. L., FELDMAN, S., WUERFFEL, R., ACHOUR, I., WANG, L. & KUMAR, S. 2012. Three-dimensional architecture of the IgH locus facilitates class switch recombination. *Ann N Y Acad Sci,* 1267**,** 86-94.

KERFOOT, S. M., YAARI, G., PATEL, J. R., JOHNSON, K. L., GONZALEZ, D. G., KLEINSTEIN, S. H. & HABERMAN, A. M. 2011. Germinal center B cell and T follicular helper cell development initiates in the interfollicular zone. *Immunity,* 34**,** 947-60.

KHAIR, L., GUIKEMA, J. E. J., LINEHAN, E. K., UCHER, A. J., LEUS, N. G. J., OGILVIE, C., LOU, Z., SCHRADER, C. E. & STAVNEZER, J. 2014. ATM Increases Activation-Induced Cytidine Deaminase Activity at Downstream S Regions during Class-Switch Recombination. *The Journal of Immunology,* 192**,** 4887-4896.

KILLIG, M., GLATZER, T. & ROMAGNANI, C. 2014. Recognition strategies of group 3 innate lymphoid cells. *Front Immunol,* 5**,** 142.

KIM, N. & JINKS-ROBERTSON, S. 2012. Transcription as a source of genome instability. *Nat Rev Genet,* 13**,** 204-214.

KIM, S., DAVIS, M., SINN, E., PATTEN, P. & HOOD, L. 1981. Antibody diversity: somatic hypermutation of rearranged VH genes. *Cell,* 27**,** 573-81.

KINCADE, P. W., LAWTON, A. R., BOCKMAN, D. E. & COOPER, M. D. 1970. Suppression of Immunoglobulin G Synthesis as a Result of Antibody-Mediated Suppression of Immunoglobulin M Synthesis in Chickens. *Proceedings of the National Academy of Sciences,* 67**,** 1918-1925.

KINNER, A., WU, W., STAUDT, C. & ILIAKIS, G. 2008. γ-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. *Nucleic Acids Research,* 36**,** 5678-5694.

KITANO, M., MORIYAMA, S., ANDO, Y., HIKIDA, M., MORI, Y., KUROSAKI, T. & OKADA, T. 2011. Bcl6 protein expression shapes pre-germinal center B cell dynamics and follicular helper T cell heterogeneity. *Immunity,* 34**,** 961-72.

KLEIN, G., GIOVANELLA, B., WESTMAN, A., STEHLIN, J. S. & MUMFORD, D. 1975. An EBV-genome-negative cell line established from an American Burkitt lymphoma; receptor characteristics. EBV infectibility and permanent conversion into EBV-positive sublines by in vitro infection. *Intervirology,* 5**,** 319-34.

KLEIN, H. L. 2008. The consequences of Rad51 overexpression for normal and tumor cells. *DNA Repair,* 7**,** 686-693.

KLEIN, ISAAC A., RESCH, W., JANKOVIC, M., OLIVEIRA, T., YAMANE, A., NAKAHASHI, H., DI VIRGILIO, M., BOTHMER, A., NUSSENZWEIG, A., ROBBIANI, DAVIDE F., CASELLAS, R. & NUSSENZWEIG, MICHEL C. 2011. Translocation-Capture Sequencing Reveals the Extent and Nature of Chromosomal Rearrangements in B Lymphocytes. *Cell,* 147**,** 95-106.

KLEINFIELD, R., HARDY, R. R., TARLINTON, D., DANGL, J., HERZENBERG, L. A. & WEIGERT, M. 1986. Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a Ly 1+ B-cell lymphoma. *Nature,* 322**,** 843-846.

KO, T. P., LIN, J. J., HU, C. Y., HSU, Y. H., WANG, A. H. J. & LIAW, S. H. 2003. Crystal structure of yeast cytosine deaminase: Insights into enzyme mechanism and evolution. *Journal of Biological Chemistry,* 278**,** 19111-19117.

KOONIN, E. V. 1992. A new group of putative RNA helicases. *Trends Biochem Sci,* 17**,** 495-7.

KOTANI, A., OKAZAKI, I. M., MURAMATSU, M., KINOSHITA, K., BEGUM, N. A., NAKAJIMA, T., SAITO, H. & HONJO, T. 2005. A target selection of somatic hypermutations is regulated similarly between T and B cells upon activation-induced cytidine deaminase expression. *Proc Natl Acad Sci U S A,* 102**,** 4506-11.

KROESE, F. G., AMMERLAAN, W. A. & DEENEN, G. J. 1992. Location and function of B-cell lineages. *Ann N Y Acad Sci,* 651**,** 44-58.

KUEHNER, J. N., PEARSON, E. L. & MOORE, C. 2011. Unravelling the means to an end: RNA polymerase II transcription termination. *Nature Reviews Molecular Cell Biology,* 12**,** 283-294.

KUNIMOTO, D. Y., HARRIMAN, G. R. & STROBER, W. 1988. Regulation of IgA differentiation in CH12LX B cells by lymphokines. IL-4 induces membrane IgM-positive CH12LX cells to express membrane IgA and IL-5 induces membrane IgA-positive CH12LX cells to secrete IgA. *The Journal of Immunology,* 141**,** 713-20.

KUO, L. J. & YANG, L.-X. 2008. γ-H2AX - A Novel Biomarker for DNA Double-strand Breaks. *In Vivo,* 22**,** 305-309.

LACAVA, J., HOUSELEY, J., SAVEANU, C., PETFALSKI, E., THOMPSON, E., JACQUIER, A. & TOLLERVEY, D. 2005. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell,* 121**,** 713-724.

LAMKANFI, M. & DIXIT, V. M. 2012. Inflammasomes and their roles in health and disease. *Annu Rev Cell Dev Biol,* 28**,** 137-61.

LARIJANI, M., FRIEDER, D., BASIT, W. & MARTIN, A. 2005. The mutation spectrum of purified AID is similar to the mutability index in Ramos cells and in ung-/- msh2-/- mice. *Immunogenetics,* 56**,** 840-845.

LEBRETON, A., TOMECKI, R., DZIEMBOWSKI, A. & SERAPHIN, B. 2008. Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature,* 456**,** 993-6.

LEE, C., LI, X., HECHMER, A., EISEN, M., BIGGIN, M. D., VENTERS, B. J., JIANG, C., LI, J., PUGH, B. F. & GILMOUR, D. S. 2008. NELF and GAGA factor are linked to promoter-proximal pausing at many genes in Drosophila. *Mol Cell Biol,* 28**,** 3290-300.

LEE, Y. K. & MAZMANIAN, S. K. 2010. Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science,* 330**,** 1768-73.

LESNIK, E. A. & FREIER, S. M. 1995. Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. *Biochemistry,* 34**,** 10807-15.

LI, X. & MANLEY, J. L. 2005. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell,* 122**,** 365-78.

LI, X. & MANLEY, J. L. 2006. Cotranscriptional processes and their influence on genome stability. *Genes Dev,* 20**,** 1838-47.

LI, Z., WOO, C. J., IGLESIAS-USSEL, M. D., RONAI, D. & SCHARFF, M. D. 2004. The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev,* 18**,** 1-11.

LIEBER, M. R. 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual Review of Biochemistry.*

LIU, K., VICTORA, G. D., SCHWICKERT, T. A., GUERMONPREZ, P., MEREDITH, M. M., YAO, K., CHU, F. F., RANDOLPH, G. J., RUDENSKY, A. Y. & NUSSENZWEIG, M. 2009. In vivo analysis of dendritic cell development and homeostasis. *Science,* 324**,** 392-7.

LIU, L. F. & WANG, J. C. 1987. Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences,* 84**,** 7024-7027.

LIU, M., DUKE, J. L., RICHTER, D. J., VINUESA, C. G., GOODNOW, C. C., KLEINSTEIN, S. H. & SCHATZ, D. G. 2008. Two levels of protection for the B cell genome during somatic hypermutation. *Nature,* 451**,** 841-845.

LIU, Q., GREIMANN, J. C. & LIMA, C. D. 2006. Reconstitution, Activities, and Structure of the Eukaryotic RNA Exosome. *Cell,* 127**,** 1223-1237.

LIU, Y. J., JOSHUA, D. E., WILLIAMS, G. T., SMITH, C. A., GORDON, J. & MACLENNAN, I. C. M. 1989. Mechanism of antigen-driven selection in germinal centres. *Nature,* 342**,** 929-931.

LONGERICH, S., BASU, U., ALT, F. & STORB, U. 2006. AID in somatic hypermutation and class switch recombination. *Current Opinion in Immunology,* 18**,** 164-174.

LUBAS, M., CHRISTENSEN, M., KRISTIANSEN, M., DOMANSKI, M., FALKENBY, L., LYKKE-ANDERSEN, S., ANDERSEN, J., DZIEMBOWSKI, A. & JENSEN, T. 2011. Interaction Profiling Identifies the Human Nuclear Exosome Targeting Complex. *Molecular Cell,* 43**,** 624-637.

MACGINNITIE, A. J., ANANT, S. & DAVIDSON, N. O. 1995. Mutagenesis of apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, reveals distinct domains that mediate cytosine nucleoside deaminase, RNA binding, and RNA editing activity. *Journal of Biological Chemistry,* 270**,** 14768-14775.

MACLENNAN, I. C. 1994. Germinal centers. *Annu. Rev. Immunol.,* 12**,** 117-139.

MAIZELS, N. 2006. Dynamic roles for G4 DNA in the biology of eukaryotic cells. *Nat Struct Mol Biol,* 13**,** 1055-1059.

MAKINO, D. L., BAUMGÄRTNER, M. & CONTI, E. 2013. Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. *Nature,* 494**,** 70-75.

MARIATHASAN, S., WEISS, D. S., NEWTON, K., MCBRIDE, J., O'ROURKE, K., ROOSE-GIRMA, M., LEE, W. P., WEINRAUCH, Y., MONACK, D. M. & DIXIT, V. M. 2006. Cryopyrin activates the inflammasome in response to toxins and ATP. *Nature,* 440**,** 228-32.

MARTINON, F., MAYOR, A. & TSCHOPP, J. 2009. The inflammasomes: guardians of the body. *Annu Rev Immunol,* 27**,** 229-65.

MARTINS, G. & CALAME, K. 2008. Regulation and functions of Blimp-1 in T and B lymphocytes. *Annual Review of Immunology.*

MAUL, R. W. & GEARHART, P. J. 2010. AID and somatic hypermutation. *Advances in Immunology.*

MAUL, R. W., SARIBASAK, H., MARTOMO, S. A., MCCLURE, R. L., YANG, W., VAISMAN, A., GRAMLICH, H. S., SCHATZ, D. G., WOODGATE, R., WILSON, D. M. & GEARHART, P. J. 2011. Uracil residues dependent on the deaminase AID in immunoglobulin gene variable and switch regions. *Nature Immunology,* 12**,** 70-76.

MAURI, C. & BOSMA, A. 2012. Immune regulatory function of B cells. *Annu Rev Immunol,* 30**,** 221-41.

MAURI, C., GRAY, D., MUSHTAQ, N. & LONDEI, M. 2003. Prevention of arthritis by interleukin 10-producing B cells. *Journal of Experimental Medicine,* 197**,** 489-501.

MCBRIDE, K. M., BARRETO, V., RAMIRO, A. R., STAVROPOULOS, P. & NUSSENZWEIG, M. C. 2004. Somatic Hypermutation Is Limited by CRM1-dependent Nuclear Export of Activation-induced Deaminase. *The Journal of Experimental Medicine,* 199**,** 1235-1244.

MCBRIDE, K. M., GAZUMYAN, A., WOO, E. M., BARRETO, V. M., ROBBIANI, D. F., CHAIT, B. T. & NUSSENZWEIG, M. C. 2006. Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 8798-8803.

MCKEAN, D., HUPPI, K., BELL, M., STAUDT, L., GERHARD, W. & WEIGERT, M. 1984. Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc Natl Acad Sci U S A,* 81**,** 3180-4.

MEDVINSKY, A. & DZIERZAK, E. 1996. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell,* 86**,** 897-906.

MERAD, M., SATHE, P., HELFT, J., MILLER, J. & MORTHA, A. 2013. The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annu Rev Immunol,* 31**,** 563-604.

MILLIGAN, L., DECOURTY, L., SAVEANU, C., RAPPSILBER, J., CEULEMANS, H., JACQUIER, A. & TOLLERVEY, D. 2008. A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts. *Molecular and Cellular Biology,* 28**,** 5446-5457.

MILSTEIN, C., NEUBERGER, M. S. & STADEN, R. 1998. Both DNA strands of antibody genes are hypermutation targets. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 8791-8794.

MIMITOU, E. P. & SYMINGTON, L. S. 2010. Ku prevents Exo1 and Sgs1-dependent resection of DNA ends in the absence of a functional MRX complex or Sae2. *EMBO J.*

MITCHELL, P., PETFALSKI, E., HOUALLA, R., PODTELEJNIKOV, A., MANN, M. & TOLLERVEY, D. 2003. Rrp47p is an exosome-associated protein required for the 3′ processing of stable RNAs. *Molecular and Cellular Biology,* 23**,** 6982-6992.

MITCHELL, P., PETFALSKI, E., SHEVCHENKO, A., MANN, M. & TOLLERVEY, D. 1997. The Exosome: A Conserved Eukaryotic RNA Processing Complex Containing Multiple 3′→5′ Exoribonucleases. *Cell,* 91**,** 457-466.

MIYAKE, K., YAMASHITA, Y., OGATA, M., SUDO, T. & KIMOTO, M. 1995. RP105, a novel B cell surface molecule implicated in B cell activation, is a member of the leucine-rich repeat protein family. *The Journal of Immunology,* 154**,** 3333-40.

MIZOGUCHI, A., MIZOGUCHI, E., TAKEDATSU, H., BLUMBERG, R. S. & BHAN, A. K. 2002. Chronic intestinal inflammatory condition generates IL-10-producing regulatory B cell subset characterized by CD1d upregulation. *Immunity,* 16**,** 219-30.

MONROE, J. G. & DORSHKIND, K. 2007. Fate Decisions Regulating Bone Marrow and Peripheral B Lymphocyte Development. *In:* FREDERICK W. ALT, K. F. A. T. H. F. M. J. W. U. & EMIL, R. U. (eds.) *Advances in Immunology.* Academic Press.

MONTECINO-RODRIGUEZ, E., LEATHERS, H. & DORSHKIND, K. 2006. Identification of a B-1 B cell-specified progenitor. *Nat Immunol,* 7**,** 293-301.

MOREIRA, M.-C., KLUR, S., WATANABE, M., NEMETH, A. H., BER, I. L., MONIZ, J.-C., TRANCHANT, C., AUBOURG, P., TAZIR, M., SCHOLS, L., PANDOLFO, M., SCHULZ, J. B., POUGET, J., CALVAS, P., SHIZUKA-IKEDA, M., SHOJI, M., TANAKA, M., IZATT, L., SHAW, C. E., M'ZAHEM, A., DUNNE, E., BOMONT, P., BENHASSINE, T., BOUSLAM, N., STEVANIN, G., BRICE, A., GUIMARAES, J., MENDONCA, P., BARBOT, C., COUTINHO, P., SEQUEIROS, J., DURR, A., WARTER, J.-M. & KOENIG, M. 2004. Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat Genet,* 36**,** 225-227.

MORGAN, H. D., DEAN, W., COKER, H. A., REIK, W. & PETERSEN-MAHRT, S. K. 2004. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem,* 279**,** 52353-60.

MUERS, M. 2011. Mutation: The perils of transcription. *Nature Reviews Genetics,* 12**,** 156.

MÜLLER, A. M., MEDVINSKY, A., STROUBOULIS, J., GROSVELD, F. & DZIERZAK, E. 1994. Development of hematopoietic stem cell activity in the mouse embryo. *Immunity,* 1**,** 291-301.

MURAMATSU, M., KINOSHITA, K., FAGARASAN, S., YAMADA, S., SHINKAI, Y. & HONJO, T. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell,* 102**,** 553-563.

MURAMATSU, M., SANKARANAND, V. S., ANANT, S., SUGAI, M., KINOSHITA, K., DAVIDSON, N. O. & HONJO, T. 1999. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem,* 274**,** 18470-6.

MUSE, G. W., GILCHRIST, D. A., NECHAEV, S., SHAH, R., PARKER, J. S., GRISSOM, S. F., ZEITLINGER, J. & ADELMAN, K. 2007. RNA polymerase is poised for activation across the genome. *Nat Genet,* 39**,** 1507-1511.

NAKAMURA, M., KONDO, S., SUGAI, M., NAZAREA, M., IMAMURA, S. & HONJO, T. 1996. High frequency class switching of an lgM+ B lymphoma clone CH12F3 to lgA+ cells. *International Immunology,* 8**,** 193-201.

NETA, R. & SALVIN, S. B. 1974. Specific suppression of delayed hypersensitivity: the possible presence of a suppressor B cell in the regulation of delayed hypersensitivity. *J Immunol,* 113**,** 1716-25.

NEUBERGER, M. S., HARRIS, R. S., DI NOIA, J. & PETERSEN-MAHRT, S. K. 2003. Immunity through DNA deamination. *Trends Biochem Sci,* 28**,** 305-12.

NIEUWENHUIS, P. & OPSTELTEN, D. 1984. Functional anatomy of germinal centers. *Am J Anat,* 170**,** 421-35.

NUDLER, E., MUSTAEV, A., GOLDFARB, A. & LUKHTANOV, E. 1997. The RNA–DNA Hybrid Maintains the Register of Transcription by Preventing Backtracking of RNA Polymerase. *Cell,* 89**,** 33-41.

NUTT, S. L., HODGKIN, P. D., TARLINTON, D. M. & CORCORAN, L. M. 2015. The generation of antibody-secreting plasma cells. *Nat Rev Immunol,* 15**,** 160-171.

O'SHEA, J. J. & PAUL, W. E. 2010. Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells. *Science,* 327**,** 1098-102.

OCHIAI, K., MAIENSCHEIN-CLINE, M., SIMONETTI, G., CHEN, J., ROSENTHAL, R., BRINK, R., CHONG, A. S., KLEIN, U., DINNER, A. R., SINGH, H. & SCIAMMAS, R. 2013. Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity,* 38**,** 918-29.

ODDONE, A., LORENTZEN, E., BASQUIN, J., GASCH, A., RYBIN, V., CONTI, E. & SATTLER, M. 2007. Structural and biochemical characterization of the yeast exosome component Rrp40. *EMBO Reports,* 8**,** 63-69.

OKAZAKI, I. M., KOTANI, A. & HONJO, T. 2007. Role of AID in tumorigenesis. *Adv Immunol,* 94**,** 245-73.

OKAZAKI, R., OKAZAKI, T., SAKABE, K., SUGIMOTO, K. & SUGINO, A. 1968. Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc Natl Acad Sci U S A,* 59**,** 598-605.

ORKIN, S. H. & ZON, L. I. 2008. Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell,* 132**,** 631-644.

ORTHWEIN, A., ZAHN, A., METHOT, S. P., GODIN, D., CONTICELLO, S. G., TERADA, K. & DI NOIA, J. M. 2012. Optimal functional levels of activation-induced deaminase specifically require the Hsp40 DnaJa1. *EMBO Journal,* 31**,** 679-691.

OTTERSBACH, K. & DZIERZAK, E. 2005. The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. *Dev Cell,* 8**,** 377-87.

PASQUALUCCI, L., GUGLIELMINO, R., HOULDSWORTH, J., MOHR, J., AOUFOUCHI, S., POLAKIEWICZ, R., CHAGANTI, R. S. & DALLA-FAVERA, R. 2004. Expression of the AID protein in normal and neoplastic B cells. *Blood,* 104**,** 3318-25.

PASQUALUCCI, L., KITAURA, Y., GU, H. & DALLA-FAVERA, R. 2006. PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 395-400.

PASQUALUCCI, L., MIGLIAZZA, A., FRACCHIOLLA, N., WILLIAM, C., NERI, A., BALDINI, L., CHAGANTI, R. S. K., KLEIN, U., KÜPPERS, R., RAJEWSKY, K. & DALLA-FAVERA, R. 1998. BCL-6 mutations in normal germinal center B cells: Evidence of somatic hypermutation acting outside Ig loci. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 11816-11821.

PASQUALUCCI, L., NEUMEISTER, P., GOOSSENS, T., NANJANGUD, G., CHAGANTI, R. S. K., KÜPPERS, R. & DALLA-FAVERA, R. 2001. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature,* 412**,** 341-346.

PATENAUDE, A. M., ORTHWEIN, A., HU, Y., CAMPO, V. A., KAVLI, B., BUSCHIAZZO, A. & DI NOIA, J. M. 2009. Active nuclear import and cytoplasmic retention of activation-induced deaminase. *Nature Structural and Molecular Biology,* 16**,** 517-527.

PAVRI, R., GAZUMYAN, A., JANKOVIC, M., DI VIRGILIO, M., KLEIN, I., ANSARAH-SOBRINHO, C., RESCH, W., YAMANE, A., SAN-MARTIN, B. R., BARRETO, V., NIELAND, T. J., ROOT, D. E., CASELLAS, R. & NUSSENZWEIG, M. C. 2010. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell,* 143**,** 122-133.

PECCHI, E., DALLAPORTA, M., JEAN, A., THIRION, S. & TROADEC, J. D. 2009. Prostaglandins and sickness behavior: old story, new insights. *Physiol Behav,* 97**,** 279-92.

PEFANIS, E., WANG, J., ROTHSCHILD, G., LIM, J., CHAO, J., RABADAN, R., ECONOMIDES, A. N. & BASU, U. 2014. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature,* 514**,** 389-93.

PEFANIS, E., WANG, J., ROTHSCHILD, G., LIM, J., KAZADI, D., SUN, J., FEDERATION, A., CHAO, J., ELLIOTT, O., LIU, Z. P., ECONOMIDES, A. N., BRADNER, J. E., RABADAN, R. & BASU, U. 2015. RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell,* 161**,** 774-89.

PERSSON, E. K., URONEN-HANSSON, H., SEMMRICH, M., RIVOLLIER, A., HAGERBRAND, K., MARSAL, J., GUDJONSSON, S., HAKANSSON, U., REIZIS, B., KOTARSKY, K. & AGACE, W. W. 2013. IRF4 transcription-factor-dependent CD103(+)CD11b(+) dendritic cells drive mucosal T helper 17 cell differentiation. *Immunity,* 38**,** 958-69.

PETERS, A. & STORB, U. 1996. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity,* 4**,** 57-65.

PETERSEN, S., CASELLAS, R., REINA-SAN-MARTIN, B., CHEN, H. T., DIFILIPPANTONIO, M. J., WILSON, P. C., HANITSCH, L., CELESTE, A., MURAMATSU, M., PILCH, D. R., REDON, C., RIED, T., BONNER, W. M., HONJO, T., NUSSENZWEIG, M. C. & NUSSENZWEIG, A. 2001. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature,* 414**,** 660-5.

PETERSEN-MAHRT, S. K., HARRIS, R. S. & NEUBERGER, M. S. 2002. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature,* 418**,** 99-103.

PHAM, P., BRANSTEITTER, R. & GOODMAN, M. F. 2005. Reward versus risk: DNA cytidine deaminases triggering immunity and disease. *Biochemistry,* 44**,** 2703-2715.

PHAM, P., BRANSTEITTER, R., PETRUSKA, J. & GOODMAN, M. F. 2003. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature,* 424**,** 103-107.

PIEPER, K., GRIMBACHER, B. & EIBEL, H. 2013. B-cell biology and development. *Journal of Allergy and Clinical Immunology,* 131**,** 959-971.

PORRUA, O. & LIBRI, D. 2015. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol,* 16**,** 190-202.

PREKER, P., NIELSEN, J., KAMMLER, S., LYKKE-ANDERSEN, S., CHRISTENSEN, M. S., MAPENDANO, C. K., SCHIERUP, M. H. & JENSEN, T. H. 2008. RNA exosome depletion reveals transcription upstream of active human promoters. *Science,* 322**,** 1851-1854.

PROCHNOW, C., BRANSTEITTER, R., KLEIN, M. G., GOODMAN, M. F. & CHEN, X. S. 2007. The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature,* 445**,** 447-51.

RADA, C., WILLIAMS, G. T., NILSEN, H., BARNES, D. E., LINDAHL, T. & NEUBERGER, M. S. 2002. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Current Biology,* 12**,** 1748-1755.

RAJEWSKY, K., FORSTER, I. & CUMANO, A. 1987. Evolutionary and somatic selection of the antibody repertoire in the mouse. *Science,* 238**,** 1088-1094.

RAMIRO, A. R., JANKOVIC, M., EISENREICH, T., DIFILIPPANTONIO, S., CHEN-KIANG, S., MURAMATSU, M., HONJO, T., NUSSENZWEIG, A. & NUSSENZWEIG, M. C. 2004. AID Is Required for c-myc/IgH Chromosome Translocations In Vivo. *Cell,* 118**,** 431-438.

RAMIRO, A. R., STAVROPOULOS, P., JANKOVIC, M. & NUSSENZWEIG, M. C. 2003. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nature Immunology,* 4**,** 452-456.

RAN, F. A., HSU, P. D., LIN, C. Y., GOOTENBERG, J. S., KONERMANN, S., TREVINO, A. E., SCOTT, D. A., INOUE, A., MATOBA, S., ZHANG, Y. & ZHANG, F. 2013a. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell,* 154**,** 1380-9.

RAN, F. A., HSU, P. D., WRIGHT, J., AGARWALA, V., SCOTT, D. A. & ZHANG, F. 2013b. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc,* 8**,** 2281-308.

RANJIT, S., KHAIR, L., LINEHAN, E. K., UCHER, A. J., CHAKRABARTI, M., SCHRADER, C. E. & STAVNEZER, J. 2011. AID binds cooperatively with UNG and Msh2-Msh6 to Ig switch regions dependent upon the AID C terminus. *Journal of Immunology,* 187**,** 2464-2475.

RATCLIFFE, M. J. H. 2006. Antibodies, immunoglobulin genes and the bursa of Fabricius in chicken B cell development. *Developmental & Comparative Immunology,* 30**,** 101-118.

RATHINAM, V. A. & FITZGERALD, K. A. 2011. Cytosolic surveillance and antiviral immunity. *Curr Opin Virol,* 1**,** 455-62.

RATHINAM, V. A., VANAJA, S. K. & FITZGERALD, K. A. 2012. Regulation of inflammasome signaling. *Nat Immunol,* 13**,** 333-42.

RATMEYER, L., VINAYAK, R., ZHONG, Y. Y., ZON, G. & WILSON, W. D. 1994. Sequence specific thermodynamic and structural properties for DNA.RNA duplexes. *Biochemistry,* 33**,** 5298-304.

RAULET, D. H. 2006. Missing self recognition and self tolerance of natural killer (NK) cells. *Semin Immunol,* 18**,** 145-50.

REDDY, M. C. & VASQUEZ, K. M. 2005. Repair of genome destabilizing lesions. *Radiation Research,* 164**,** 345-356.

REIZIS, B., BUNIN, A., GHOSH, H. S., LEWIS, K. L. & SISIRAK, V. 2011. Plasmacytoid dendritic cells: recent progress and open questions. *Annu Rev Immunol,* 29**,** 163-83.

RESCIGNO, M., URBANO, M., VALZASINA, B., FRANCOLINI, M., ROTTA, G., BONASIO, R., GRANUCCI, F., KRAEHENBUHL, J. P. & RICCIARDI-CASTAGNOLI, P. 2001. Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nat Immunol,* 2**,** 361-7.

RETH, M., GEHRMANN, P., PETRAC, E. & WIESE, P. 1986. A novel V(H) to V(H)DJ(H) joining mechanism in heavy-chain-negative (null) pre-B cells results in heavy-chain production. *Nature,* 322**,** 840-842.

REYNAUD, C. A., ANQUEZ, V., GRIMAL, H. & WEILL, J. C. 1987. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell,* 48**,** 379-388.

REYNAUD, C. A., DAHAN, A., ANQUEZ, V. & WEILL, J. C. 1989. Somatic hyperconversion diversifies the single VH gene of the chicken with a high incidence in the D region. *Cell,* 59**,** 171-183.

RICH, A. 1960. A HYBRID HELIX CONTAINING BOTH DEOXYRIBOSE AND RIBOSE POLYNUCLEOTIDES AND ITS RELATION TO THE TRANSFER OF INFORMATION BETWEEN THE NUCLEIC ACIDS. *Proceedings of the National Academy of Sciences,* 46**,** 1044-1053.

RICHARD, P., FENG, S. & MANLEY, J. L. 2013. A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. *Genes Dev,* 27**,** 2227-32.

RICKERT, R. C. 2013. New insights into pre-BCR and BCR signalling with relevance to B cell malignancies. *Nat Rev Immunol,* 13**,** 578-591.

RIEDER, C. L. & PALAZZO, R. E. 1992. Colcemid and the mitotic cycle. *J Cell Sci,* 102 (Pt 3)**,** 387-92.

ROBBIANI, D. F., BOTHMER, A., CALLEN, E., REINA-SAN-MARTIN, B., DORSETT, Y., DIFILIPPANTONIO, S., BOLLAND, D. J., CHEN, H. T., CORCORAN, A. E., NUSSENZWEIG, A. & NUSSENZWEIG, M. C. 2008. AID Is Required for the Chromosomal Breaks in c-myc that Lead to c-myc/IgH Translocations. *Cell,* 135**,** 1028-1038.

RONAI, D., IGLESIAS-USSEL, M. D., FAN, M., LI, Z., MARTIN, A. & SCHARFF, M. D. 2007. Detection of chromatin-associated single-stranded DNA in regions targeted for somatic hypermutation. *Journal of Experimental Medicine,* 204**,** 181-190.

RONDÓN, A. G., JIMENO, S. & AGUILERA, A. 2010. The interface between transcription and mRNP export: from THO to THSC/TREX-2. *Biochim Biophys Acta,* 1799**,** 533-8.

RONDÓN, A. G., MISCHO, H. E., KAWAUCHI, J. & PROUDFOOT, N. J. 2009. Fail-Safe Transcriptional Termination for Protein-Coding Genes in S. cerevisiae. *Molecular Cell,* 36**,** 88-98.

ROSSER, ELIZABETH C. & MAURI, C. 2015. Regulatory B Cells: Origin, Phenotype, and Function. *Immunity,* 42**,** 607-612.

ROTIN, D. & KUMAR, S. 2009. Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol,* 10**,** 398-409.

ROY, R., CHUN, J. & POWELL, S. N. 2012. BRCA1 and BRCA2: Different roles in a common pathway of genome protection. *Nature Reviews Cancer,* 12**,** 68-78.

RUSH, J. S., LIU, M., ODEGARD, V. H., UNNIRAMAN, S. & SCHATZ, D. G. 2005. Expression of activation-induced cytidine deaminase is regulated by cell division, providing a mechanistic basis for division-linked class switch recombination. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 13242-13247.

SAENZ, S. A., TAYLOR, B. C. & ARTIS, D. 2008. Welcome to the neighborhood: epithelial cell-derived cytokines license innate and adaptive immune responses at mucosal sites. *Immunol Rev,* 226**,** 172-90.

SAINTAMAND, A., LECRON, J. C., MOREL, F., COGNÉ, M. & DENIZOT, Y. 2015. Comment on "IgH Chain Class Switch Recombination: Mechanism and Regulation". *The Journal of Immunology,* 194**,** 2039-2040.

SAITO, M., GAO, J., BASSO, K., KITAGAWA, Y., SMITH, P. M., BHAGAT, G., PERNIS, A., PASQUALUCCI, L. & DALLA-FAVERA, R. 2007. A signaling pathway mediating downregulation of BCL6 in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell,* 12**,** 280-92.

SAKABE, K. & OKAZAKI, R. 1966. A unique property of the replicating region of chromosomal DNA. *Biochim Biophys Acta,* 129**,** 651-4.

SANDER, L. E., DAVIS, M. J., BOEKSCHOTEN, M. V., AMSEN, D., DASCHER, C. C., RYFFEL, B., SWANSON, J. A., MULLER, M. & BLANDER, J. M. 2011. Detection of prokaryotic mRNA signifies microbial viability and promotes immunity. *Nature,* 474**,** 385-9.

SATOH, T., HATANAKA, M., YAMAMOTO, K., KURO-O, M. & SOFUNI, T. 2002. Application of mFISH for the analysis of chemically-induced chromosomal aberrations: a model for the formation of triradial chromosomes. *Mutat Res,* 504**,** 57-65.

SATPATHY, A. T., WU, X., ALBRING, J. C. & MURPHY, K. M. 2012. Re(de)fining the dendritic cell lineage. *Nat Immunol,* 13**,** 1145-54.

SAUNDERS, H. L. & MAGOR, B. G. 2004. Cloning and expression of the AID gene in the channel catfish. *Developmental and Comparative Immunology,* 28**,** 657-663.

SAVAGE, J. R. K. 1976. Classification and relationships of induced chromosomal structural changes. *Journal of Medical Genetics,* 13**,** 103-122.

SAYEGH, C. E., QUONG, M. W., AGATA, Y. & MURRE, C. 2003. E-proteins directly regulate expression of activation-induced deaminase in mature B cells. *Nature Immunology,* 4**,** 586-593.

SCHAEFFER, D., TSANOVA, B., BARBAS, A., REIS, F. P., DASTIDAR, E. G., SANCHEZ-ROTUNNO, M., ARRAIANO, C. M. & VAN HOOF, A. 2009. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature Structural and Molecular Biology,* 16**,** 56-62.

SCHILDERS, G., RAIJMAKERS, R., RAATS, J. M. H. & PRUIJN, G. J. M. 2005. MPP6 is an exosome-associated RNA-binding protein involved in 5.8S rRNA maturation. *Nucleic Acids Research,* 33**,** 6795-6804.

SCHLITZER, A. & GINHOUX, F. 2014. Organization of the mouse and human DC network. *Curr Opin Immunol,* 26**,** 90-9.

SCHLITZER, A., MCGOVERN, N., TEO, P., ZELANTE, T., ATARASHI, K., LOW, D., HO, A. W., SEE, P., SHIN, A., WASAN, P. S., HOEFFEL, G., MALLERET, B., HEISEKE, A., CHEW, S., JARDINE, L., PURVIS, H. A., HILKENS, C. M., TAM, J., POIDINGER, M., STANLEY, E. R., KRUG, A. B., RENIA, L., SIVASANKAR, B., NG, L. G., COLLIN, M., RICCIARDI-CASTAGNOLI, P., HONDA, K., HANIFFA, M. & GINHOUX, F. 2013. IRF4 transcription factor-dependent CD11b+ dendritic cells in human and mouse control mucosal IL-17 cytokine responses. *Immunity,* 38**,** 970-83.

SCHRADER, C. E., GUIKEMA, J. E. J., LINEHAN, E. K., SELSING, E. & STAVNEZER, J. 2007. Activation-Induced Cytidine Deaminase-Dependent DNA Breaks in Class Switch Recombination Occur during G1 Phase of the Cell Cycle and Depend upon Mismatch Repair. *The Journal of Immunology,* 179**,** 6064-6071.

SCHRADER, C. E., LINEHAN, E. K., MOCHEGOVA, S. N., WOODLAND, R. T. & STAVNEZER, J. 2005. Inducible DNA breaks in Ig S regions are dependent on AID and UNG. *Journal of Experimental Medicine,* 202**,** 561-568.

SCHROCK, E., DU MANOIR, S., VELDMAN, T., SCHOELL, B., WIENBERG, J., FERGUSON-SMITH, M. A., NING, Y., LEDBETTER, D. H., BAR-AM, I., SOENKSEN, D., GARINI, Y. & RIED, T. 1996. Multicolor spectral karyotyping of human chromosomes. *Science,* 273**,** 494-7.

SCHUCH, B., FEIGENBUTZ, M., MAKINO, D. L., FALK, S., BASQUIN, C., MITCHELL, P. & CONTI, E. 2014. The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase. *Embo j,* 33**,** 2829-46.

SCHWICKERT, T. A., LINDQUIST, R. L., SHAKHAR, G., LIVSHITS, G., SKOKOS, D., KOSCO-VILBOIS, M. H., DUSTIN, M. L. & NUSSENZWEIG, M. C. 2007. In vivo imaging of germinal centres reveals a dynamic open structure. *Nature,* 446**,** 83-7.

SCULLY, R., PUGET, N. & VLASAKOVA, K. 2000. DNA polymerase stalling, sister chromatid recombination and the BRCA genes. *Oncogene,* 19**,** 6176-83.

SEIFERT, M., SELLMANN, L., BLOEHDORN, J., WEIN, F., STILGENBAUER, S., DURIG, J. & KUPPERS, R. 2012. Cellular origin and pathophysiology of chronic lymphocytic leukemia. *J Exp Med,* 209**,** 2183-98.

SELLARS, M., REINA-SAN-MARTIN, B., KASTNER, P. & CHAN, S. 2009. Ikaros controls isotype selection during immunoglobulin class switch recombination. *The Journal of Experimental Medicine,* 206**,** 1073-1087.

SELSING, E. & STORB, U. 1981. Somatic mutation of immunoglobulin light-chain variable-region genes. *Cell,* 25**,** 47-58.

SEN, D. & GILBERT, W. 1988. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature,* 334**,** 364-6.

SHAPIRO, G. S., AVISZUS, K., MURPHY, J. & WYSOCKI, L. J. 2002. Evolution of Ig DNA sequence to target specific base positions within codons for somatic hypermutation. *Journal of Immunology,* 168**,** 2302-2306.

SHARBEEN, G., YEE, C. W., SMITH, A. L. & JOLLY, C. J. 2012. Ectopic restriction of DNA repair reveals that UNG2 excises AID-induced uracils predominantly or exclusively during G1 phase. *J Exp Med,* 209**,** 965-74.

SHEEHY, A. M., GADDIS, N. C., CHOI, J. D. & MALIM, M. H. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature,* 418**,** 646-50.

SHEN, H. M., PETERS, A., BARON, B., ZHU, X. & STORB, U. 1998. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science,* 280**,** 1750-1752.

SHEN, H. M., RATNAM, S. & STORB, U. 2005. Targeting of the activation-induced cytosine deaminase is strongly influenced by the sequence and structure of the targeted DNA. *Molecular and Cellular Biology,* 25**,** 10815-10821.

SHEN, H. M. & STORB, U. 2004. Activation-induced cytidine deaminase (AID) can target both DNA strands when the DNA is supercoiled. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 12997-13002.

SHEN, H. M., TANAKA, A., BOZEK, G., NICOLAE, D. & STORB, U. 2006. Somatic hypermutation and class switch recombination in Msh6 -/-Ung-/- double-knockout mice. *Journal of Immunology,* 177**,** 5386-5392.

SHINKURA, R., ITO, S., BEGUM, N. A., NAGAOKA, H., MURAMATSU, M., KINOSHITA, K., SAKAKIBARA, Y., HIJIKATA, H. & HONJO, T. 2004. Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nature Immunology,* 5**,** 707-712.

SHINKURA, R., TIAN, M., SMITH, M., CHUA, K., FUJIWARA, Y. & ALT, F. W. 2003. The influence of transcriptional orientation on endogenous switch region function. *Nature Immunology,* 4**,** 435-441.

SKOURTI-STATHAKI, K., KAMIENIARZ-GDULA, K. & PROUDFOOT, N. J. 2014. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature,* 516**,** 436-439.

SKOURTI-STATHAKI, K., PROUDFOOT, N. & GROMAK, N. 2011. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Molecular Cell,* 42**,** 794-805.

SOHAIL, A., KLAPACZ, J., SAMARANAYAKE, M., ULLAH, A. & BHAGWAT, A. S. 2003. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Research,* 31**,** 2990-2994.

SOLLIER, J., STORK, C. T., GARCIA-RUBIO, M. L., PAULSEN, R. D., AGUILERA, A. & CIMPRICH, K. A. 2014. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell,* 56**,** 777-85.

SORZANO, C. O., PASCUAL-MONTANO, A., SANCHEZ DE DIEGO, A., MARTINEZ, A. C. & VAN WELY, K. H. 2013. Chromothripsis: breakage-fusion-bridge over and over again. *Cell Cycle,* 12**,** 2016-23.

SOUCY, T. A., SMITH, P. G., MILHOLLEN, M. A., BERGER, A. J., GAVIN, J. M., ADHIKARI, S., BROWNELL, J. E., BURKE, K. E., CARDIN, D. P., CRITCHLEY, S., CULLIS, C. A., DOUCETTE, A., GARNSEY, J. J., GAULIN, J. L., GERSHMAN, R. E., LUBLINSKY, A. R., MCDONALD, A., MIZUTANI, H., NARAYANAN, U., OLHAVA, E. J., PELUSO, S., REZAEI, M., SINTCHAK, M. D., TALREJA, T., THOMAS, M. P., TRAORE, T., VYSKOCIL, S., WEATHERHEAD, G. S., YU, J., ZHANG, J., DICK, L. R., CLAIBORNE, C. F., ROLFE, M., BOLEN, J. B. & LANGSTON, S. P. 2009. An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature,* 458**,** 732-736.

STAVNEZER, J., GUIKEMA, J. E. J. & SCHRADER, C. E. 2008. Mechanism and regulation of class switch recombination. *Annu Rev Immunol***,** 261-292.

STAVNEZER, J. & SCHRADER, C. E. 2014. IgH Chain Class Switch Recombination: Mechanism and Regulation. *The Journal of Immunology,* 193**,** 5370-5378.

STEAD, J. A., COSTELLO, J. L., LIVINGSTONE, M. J. & MITCHELL, P. 2007. The PMC2NT domain of the catalytic exosome subunit Rrp6p provides the interface for binding with its cofactor Rrp47p, a nucleic acid-binding protein. *Nucleic Acids Research,* 35**,** 5556-5567.

STEELE, E. J. & BLANDEN, R. V. 2001. The reverse transcriptase model of somatic hypermutation. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 355**,** 61-66.

STEINMAN, R. M. & COHN, Z. A. 1973. Identification of a novel cell type in peripheral lymphoid organs of mice. I. Morphology, quantitation, tissue distribution. *J Exp Med,* 137**,** 1142-62.

STEINMAN, R. M. & COHN, Z. A. 1974. Identification of a novel cell type in peripheral lymphoid organs of mice. II. Functional properties in vitro. *J Exp Med,* 139**,** 380-97.

STORB, U. 2014. Chapter Seven - Why Does Somatic Hypermutation by AID Require Transcription of Its Target Genes? *In:* FREDERICK, W. A. (ed.) *Advances in Immunology.* Academic Press.

STORB, U., PETERS, A., KIM, N., SHEN, H. M., BOZEK, G., MICHAEL, N., HACKETT JR, J., KLOTZ, E., REYNOLDS, J. D., LOEB, L. A. & MARTIN, T. E. 1999. Molecular aspects of somatic hypermutation of immunoglobulin genes. *Cold Spring Harbor Symposia on Quantitative Biology.*

STORB, U., PETERS, A., KLOTZ, E., ROGERSON, B. & HACKETT JR, J. 1996. The mechanism of somatic hypermutation studied with transgenic and transfected target genes. *Seminars in Immunology,* 8**,** 131-140.

STRASSER, A., HARRIS, A. W., BATH, M. L. & CORY, S. 1990. Novel primitive lymphoid tumours induced in transgenic mice by cooperation between myc and bcl-2. *Nature,* 348**,** 331-3.

STUART, L. M., PAQUETTE, N. & BOYER, L. 2013. Effector-triggered versus pattern-triggered immunity: how animals sense pathogens. *Nat Rev Immunol,* 13**,** 199-206.

SUMNER, A. T., EVANS, H. J. & BUCKLAND, R. A. 1971. New technique for distinguishing between human chromosomes. *Nat New Biol,* 232**,** 31-2.

SUN, J., KEIM, C. D., WANG, J., KAZADI, D., OLIVER, P. M., RABADAN, R. & BASU, U. 2013a. E3-ubiquitin ligase Nedd4 determines the fate of AID-associated RNA polymerase II in B cells. *Genes Dev,* 27**,** 1821-33.

SUN, Q., CSORBA, T., SKOURTI-STATHAKI, K., PROUDFOOT, N. J. & DEAN, C. 2013b. R-loop stabilization represses antisense transcription at the arabidopsis FLC locus. *Science,* 340**,** 619-621.

SVEJSTRUP, J. Q. 2010. The interface between transcription and mechanisms maintaining genome integrity. *Trends Biochem Sci,* 35**,** 333-8.

SYMINGTON, L. S. & GAUTIER, J. 2011. Double-strand break end resection and repair pathway choice. *Annu Rev Genet***,** 247-71.

SYMMONS, M. F., JONES, G. H. & LUISI, B. F. 2000. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure,* 8**,** 1215-26.

TA, V. T., NAGAOKA, H., CATALAN, N., DURANDY, A., FISCHER, A., IMAI, K., NONOYAMA, S., TASHIRO, J., IKEGAWA, M., ITO, S., KINOSHITA, K., MURAMATSU, M. & HONJO, T. 2003. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nature Immunology,* 4**,** 843-848.

TCHILIAN, E. Z., WALLACE, D. L., WELLS, R. S., FLOWER, D. R., MORGAN, G. & BEVERLEY, P. C. 2001. A deletion in the gene encoding the CD45 antigen in a patient with SCID. *J Immunol,* 166**,** 1308-13.

TENG, B., BURANT, C. F. & DAVIDSON, N. O. 1993. Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science,* 260**,** 1816-9.

TENG, G., HAKIMPOUR, P., LANDGRAF, P., RICE, A., TUSCHL, T., CASELLAS, R. & PAPAVASILIOU, F. N. 2008. MicroRNA-155 Is a Negative Regulator of Activation-Induced Cytidine Deaminase. *Immunity,* 28**,** 621-629.

THOMAS, C. J. & SCHRODER, K. 2013. Pattern recognition receptor function in neutrophils. *Trends Immunol,* 34**,** 317-28.

THOMAS, M., WHITE, R. L. & DAVIS, R. W. 1976. Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proc Natl Acad Sci U S A,* 73**,** 2294-8.

TIAN, M. & ALT, F. W. 2000. Transcription-induced cleavage of Immunoglobulin switch regions by nucleotide excision repair nucleases in vitro. *Journal of Biological Chemistry,* 275**,** 24163-24172.

TISELIUS, A. & KABAT, E. A. 1939. AN ELECTROPHORETIC STUDY OF IMMUNE SERA AND PURIFIED ANTIBODY PREPARATIONS. *J Exp Med,* 69**,** 119-31.

TOMECKI, R., KRISTIANSEN, M. S., LYKKE-ANDERSEN, S., CHLEBOWSKI, A., LARSEN, K. M., SZCZESNY, R. J., DRAZKOWSKA, K., PASTULA, A., ANDERSEN, J. S., STEPIEN, P. P., DZIEMBOWSKI, A. & JENSEN, T. H. 2010. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *Embo j,* 29**,** 2342-57.

TONEGAWA, S. 1987. Molecular biology of immunologic recognition. *Tanpakushitsu kakusan koso. Protein, nucleic acid, enzyme,* 32**,** 239-250.

TUDURI, S., CRABBE, L., CONTI, C., TOURRIERE, H., HOLTGREVE-GREZ, H., JAUCH, A., PANTESCO, V., DE VOS, J., THOMAS, A., THEILLET, C., POMMIER, Y., TAZI, J., COQUELLE, A. & PASERO, P. 2009. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat Cell Biol,* 11**,** 1315-1324.

UCHIMURA, Y., BARTON, L. F., RADA, C. & NEUBERGER, M. S. 2011. Reg-γ associates with and modulates the abundance of nuclear activation-induced deaminase. *Journal of Experimental Medicine,* 208**,** 2385-2391.

UEMATSU, S., JANG, M. H., CHEVRIER, N., GUO, Z., KUMAGAI, Y., YAMAMOTO, M., KATO, H., SOUGAWA, N., MATSUI, H., KUWATA, H., HEMMI, H., COBAN, C., KAWAI, T., ISHII, K. J., TAKEUCHI, O., MIYASAKA, M., TAKEDA, K. & AKIRA, S. 2006. Detection of pathogenic intestinal bacteria by Toll-like receptor 5 on intestinal CD11c+ lamina propria cells. *Nat Immunol,* 7**,** 868-74.

URSIC, D., HIMMEL, K. L., GURLEY, K. A., WEBB, F. & CULBERTSON, M. R. 1997. The yeast SEN1 gene is required for the processing of diverse RNA classes. *Nucleic Acids Research,* 25**,** 4778-4785.

VANCE, R. E., ISBERG, R. R. & PORTNOY, D. A. 2009. Patterns of pathogenesis: discrimination of pathogenic and nonpathogenic microbes by the innate immune system. *Cell Host Microbe,* 6**,** 10-21.

VASILJEVA, L. & BURATOWSKI, S. 2006. Nrd1 interacts with the nuclear exosome for 3′ processing of RNA polymerase II transcripts. *Molecular Cell,* 21**,** 239-248.

VICTORA, G. D. & NUSSENZWEIG, M. C. 2012. Germinal centers. *Annu. Rev. Immunol.,* 30**,** 429-457.

VICTORA, G. D., SCHWICKERT, T. A., FOOKSMAN, D. R., KAMPHORST, A. O., MEYER-HERMANN, M., DUSTIN, M. L. & NUSSENZWEIG, M. C. 2010. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell,* 143**,** 592-605.

VIKSTROM, I., CAROTTA, S., LUTHJE, K., PEPERZAK, V., JOST, P. J., GLASER, S., BUSSLINGER, M., BOUILLET, P., STRASSER, A., NUTT, S. L. & TARLINTON, D. M. 2010. Mcl-1 is essential for germinal center formation and B cell memory. *Science,* 330**,** 1095-9.

VIVIER, E., TOMASELLO, E., BARATIN, M., WALZER, T. & UGOLINI, S. 2008. Functions of natural killer cells. *Nat Immunol,* 9**,** 503-510.

VON MOLTKE, J., AYRES, J. S., KOFOED, E. M., CHAVARRIA-SMITH, J. & VANCE, R. E. 2013. Recognition of bacteria by inflammasomes. *Annu Rev Immunol,* 31**,** 73-106.

WADA, T., TAKAGI, T., YAMAGUCHI, Y., FERDOUS, A., IMAI, T., HIROSE, S., SUGIMOTO, S., YANO, K., HARTZOG, G. A., WINSTON, F., BURATOWSKI, S. & HANDA, H. 1998. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes and Development,* 12**,** 343-356.

WAGA, S. & STILLMAN, B. 1998. The DNA replication fork in eukaryotic cells. *Annu Rev Biochem,* 67**,** 721-51.

WAHBA, L., AMON, JEREMY D., KOSHLAND, D. & VUICA-ROSS, M. 2011. RNase H and Multiple RNA Biogenesis Factors Cooperate to Prevent RNA:DNA Hybrids from Generating Genome Instability. *Molecular Cell,* 44**,** 978-988.

WAHBA, L., GORE, S. K. & KOSHLAND, D. 2013. The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. *Elife,* 2**,** e00505.

WANG, L., LEWIS, M. S. & JOHNSON, A. W. 2005. Domain interactions within the Ski2/3/8 complex and between the Ski complex and Ski7p. *RNA,* 11**,** 1291-1302.

WANG, S. & KOOL, E. T. 1995. Origins of the large differences in stability of DNA and RNA helices: C-5 methyl and 2'-hydroxyl effects. *Biochemistry,* 34**,** 4125-32.

WASMUTH, E. V., JANUSZYK, K. & LIMA, C. D. 2014. Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. *Nature,* 511**,** 435-439.

WEIGERT, M. G., CESARI, I. M., YONKOVICH, S. J. & COHN, M. 1970. Variability in the lambda light chain sequences of mouse antibody. *Nature,* 228**,** 1045-7.

WESTOVER, K. D., BUSHNELL, D. A. & KORNBERG, R. D. 2004. Structural basis of transcription: Nucleotide selection by rotation in the RNA polymerase II active center. *Cell,* 119**,** 481-489.

WHITBY, F. G., XIA, G., PICKART, C. M. & HILL, C. P. 1998. Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes. *J Biol Chem,* 273**,** 34983-91.

WILLMANN, K. L., MILOSEVIC, S., PAUKLIN, S., SCHMITZ, K. M., RANGAM, G., SIMON, M. T., MASLEN, S., SKEHEL, M., ROBERT, I., HEYER, V., SCHIAVO, E., REINA-SAN-MARTIN, B. & PETERSEN-MAHRT, S. K. 2012. A role for the RNA pol II-associated PAF complex in AID-induced immune diversification. *Journal of Experimental Medicine,* 209**,** 2099-2111.

WILSON, T. M., VAISMAN, A., MARTOMO, S. A., SULLIVAN, P., LAN, L., HANAOKA, F., YASUI, A., WOODGATE, R. & GEARHART, P. J. 2005. MSH2-MSH6 stimulates DNA polymerase eta, suggesting a role for A:T mutations in antibody genes. *J Exp Med,* 201**,** 637-45.

WIMBERLY, H., SHEE, C., THORNTON, P. C., SIVARAMAKRISHNAN, P., ROSENBERG, S. M. & HASTINGS, P. J. 2013. R-loops and nicks initiate DNA breakage and genome instability in non-growing Escherichia coli. *Nat Commun,* 4.

WINEY, M. & CULBERTSON, M. R. 1988. Mutations affecting the tRNA-splicing endonuclease activity of Saccharomyces cerevisiae. *Genetics,* 118**,** 609-17.

WOLF, S. D., DITTEL, B. N., HARDARDOTTIR, F. & JANEWAY, C. A., JR. 1996. Experimental autoimmune encephalomyelitis induction in genetically B cell-deficient mice. *J Exp Med,* 184**,** 2271-8.

WU, J. & CHEN, Z. J. 2014. Innate immune sensing and signaling of cytosolic nucleic acids. *Annu Rev Immunol,* 32**,** 461-88.

WU, L., FU, J. & SHEN, S. H. 2002. SKAP55 coupled with CD45 positively regulates T-cell receptor-mediated gene transcription. *Mol Cell Biol,* 22**,** 2673-86.

WUERFFEL, R., WANG, L., GRIGERA, F., MANIS, J., SELSING, E., PERLOT, T., ALT, F. W., COGNE, M., PINAUD, E. & KENTER, A. L. 2007. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity,* 27**,** 711-22.

WUERFFEL, R. A., DU, J., THOMPSON, R. J. & KENTER, A. L. 1997. Ig Sγ3 DNA-Specific Double Strand Breaks Are Induced in Mitogen-Activated B Cells and Are Implicated in Switch Recombination. *Journal of Immunology,* 159**,** 4139-4144.

XU, Z., ZAN, H., PONE, E. J., MAI, T. & CASALI, P. 2012. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat Rev Immunol,* 12**,** 517-531.

XUE, K., RADA, C. & NEUBERGER, M. S. 2006. The in vivo pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in msh2-/- ung-/- mice. *Journal of Experimental Medicine,* 203**,** 2085-2094.

YAMAGUCHI, Y., TAKAGI, T., WADA, T., YANO, K., FURUYA, A., SUGIMOTO, S., HASEGAWA, J. & HANDA, H. 1999a. NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell,* 97**,** 41-51.

YAMAGUCHI, Y., WADA, T., WATANABE, D., TAKAGI, T., HASEGAWA, J. & HANDA, H. 1999b. Structure and function of the human transcription elongation factor DSIF. *Journal of Biological Chemistry,* 274**,** 8085-8092.

YAN, C. T., BOBOILA, C., SOUZA, E. K., FRANCO, S., HICKERNELL, T. R., MURPHY, M., GUMASTE, S., GEYER, M., ZARRIN, A. A., MANIS, J. P., RAJEWSKY, K. & ALT, F. W. 2007. IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature,* 449**,** 478-82.

YANG, B., GAY, D. L., MACLEOD, M. K. L., CAO, X., HALA, T., SWEEZER, E. M., KAPPLER, J., MARRACK, P. & OLIVER, P. M. 2008. Nedd4 augments the adaptive immune response by promoting ubiquitin-mediated degradation of Cbl-b in activated T cells. *Nat Immunol,* 9**,** 1356-1363.

YANG, S. Y. & SCHATZ, D. G. 2007. Targeting of AID-Mediated Sequence Diversification by cis-Acting Determinants. *In:* ALT, F. W. & HONJO, T. (eds.) *Advances in Immunology.*

YING, C. Y., DOMINGUEZ-SOLA, D., FABI, M., LORENZ, I. C., HUSSEIN, S., BANSAL, M., CALIFANO, A., PASQUALUCCI, L., BASSO, K. & DALLA-FAVERA, R. 2013. MEF2B mutations lead to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. *Nat Immunol,* 14**,** 1084-92.

YONEYAMA, M., ONOMOTO, K., JOGI, M., AKABOSHI, T. & FUJITA, T. 2015. Viral RNA detection by RIG-I-like receptors. *Curr Opin Immunol,* 32**,** 48-53.

YU, K., CHEDIN, F., HSIEH, C. L., WILSON, T. E. & LIEBER, M. R. 2003. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature Immunology,* 4**,** 442-451.

YU, K. & LIEBER, M. R. 2003. Nucleic acid structures and enzymes in the immunoglobulin class switch recombination mechanism. *DNA Repair,* 2**,** 1163-1174.

ZEITLINGER, J., STARK, A., KELLIS, M., HONG, J.-W., NECHAEV, S., ADELMAN, K., LEVINE, M. & YOUNG, R. A. 2007. RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. *Nat Genet,* 39**,** 1512-1516.

ZHAO, Y., PAN-HAMMARSTRÖM, Q., ZHAO, Z. & HAMMARSTRÖM, L. 2005. Identification of the activation-induced cytidine deaminase gene from zebrafish: An evolutionary analysis. *Developmental and Comparative Immunology,* 29**,** 61-71.

ZHENG, S., VUONG, BAO Q., VAIDYANATHAN, B., LIN, J.-Y., HUANG, F.-T. & CHAUDHURI, J. 2015. Non-coding RNA Generated following Lariat Debranching Mediates Targeting of AID to DNA. *Cell,* 161**,** 762-773.