2019

The Role of ZMYND8 in Immunoglobulin Class Switch Recombination

Daniel Benjamin Rosen

The Role of ZMYND8

in Immunoglobulin Class Switch Recombination

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Daniel Benjamin Rosen

June 2019

The Role of ZMYND8

in Immunoglobulin Class Switch Recombination

Daniel Benjamin Rosen, Ph.D.

The Rockefeller University 2019

Class switch Recombination (CSR) also known as Immunoglobulin (Ig) Class switching is a genomic recombination/deletion reaction that diversifies the effector component of the antibody response but preserves antigen specificity. CSR is initiated by the enzyme activation induced cytidine deaminase (AID), which produces nucleotide mismatches in actively transcribed immunoglobulin heavy chain (*Igh*) switch donor and acceptor DNA. The 3' Regulatory Region (*3'RR*), a prototypical super-enhancer located at the 3' of the *Igh* locus, is essential for acceptor switch region transcription, but the mechanism by which it regulates this process is not well defined.

After targeting by AID, nearby mismatches in the donor switch region are processed into DNA double strand breaks (DSBs), translocated to DSBs in the acceptor switch region, and ligated through the DNA Damage Repair (DDR) pathway, non-homologous end-joining (NHEJ). Critical components of CSR are 53BP1 and its effector RIF1 because they inhibit end resection to promote NHEJ

and oppose competing pathways in DDR. However, the mechanism by which RIF1 effects end-protection in CSR and binds to 53BP1 is still unknown

In these studies, I identified a novel component of the RIF1 interactome, ZMYND8, a chromatin reader and transcriptional repressor that binds to RIF1 and facilitates effective CSR. Unexpectedly, ZMYND8 promotes CSR independently of RIF1. In B cells, ZMYND8 binds active promoters and super-enhancers, including the *Igh* enhancer the 3'RR. ZMYND8 controls 3'RR activity by regulating polymerase loading. In its absence there is increased 3' RR polymerase loading and decreased acceptor region transcription and CSR. Thus, ZMYND8 controls CSR by regulating the activity of the 3' *Igh* super-enhancer.

This thesis is dedicated to three individuals:

My grandfather, Arthur Rosen, for inspiring me to pursue science

My cousin, David Rosen, for your mentorship from Berkeley to the PhD

My partner, Claudine Fernandez, for your unwavering support and love

ACKNOWLEDGMENTS

My thanks and gratitude to:

My advisor, Michel Nussenzweig, for having me as a student in his laboratory. For his support, guidance, and example—thank you.

The members of my thesis committee: Howard Hang, Brian Chait, and Jayanta Chaudhuri for their guidance and perspectives. In particular, a special thank you to Brian Chait and his laboratory for answering my myriad questions about mass spectrometry, helping me launch the project, and eagerness to discuss proteomics. Also, a special thank you to Jayanta Chaudhuri and his laboratory for answering the many technical questions during my PhD, and career advice. I would also like to thank Uttiya Basu for serving as the external examiner on my thesis committee.

Michela Di Virgilio and her laboratory in Berlin for introducing me to the field of DNA repair and class switching. It has been wonderful to see the transition from the Nussenzweig lab to your own lab at the Max Delbrück Center for Molecular Medicine in Berlin. Thank you for working with me for the past five years, and I would not be here without you. Buona fortuna! Also, thank you to my co-author, Veronica Delgado-Benito; good luck on your own upcoming defense! ¡Buena suerte!.

Four members of the Nussenzweig lab and integral to the ZMYND8 project, Anna Gazumyan, Qiao Wang, Joy Pai, and Thiago Oliveira for their invaluable advice and

teaching. I always enjoyed our many discussions about the finer points of a scientific methodology and analyzing results.

The entire Nussenzweig lab deserves a special point of recognition. Being immersed in a group of such highly talented and hardworking individuals brought out the best in myself, personally and scientifically.

The David Rockefeller Graduate program and members of The Rockefeller University Dean's Office: Sidney Strickland, Emily Harms, Andrea Morris, Kristen Cullen, Marta Delgado, Stephanie Fernandez, and especially Cristian Rosario for making my PhD as seamless an experience as possible.

The Weill Cornell, Rockefeller, Sloan-Kettering Tri-Institutional MD-PhD program leadership, Olaf Andersen, Jochen Buck, Mark Pecker, and Ruth Gotian. I will be eternally grateful for the opportunity you have afforded me.

The Tri-I MD-PhD entering class of 2011.

My family, for making me who I am today.

Special recognition for my cousin, David Rosen, who as another PhD in immunology was and is an invaluable personal and professional mentor.

And last but certainly not least, my partner, Claudine. Words cannot express my appreciation and gratitude for you, in all aspects of my life.

My work was funded in part by the Medical Scientist Training Program grant from the National Institute of General Medical Sciences of the NIH under award number T32GM007739 to the Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3'RR: 3' Regulatory region (gene region - enhancer)
53BP1: TP53-binding protein 1
AID: activation induced cytidine deaminase
ATM: ataxia telangiectasia mutated
BCR: B-cell receptor
BER: Base excision repair
BRCT: BRCA1 C-terminal
CSR: Class Switch Recombination
DDR: DNA Damage Repair
DSB: Double strand break (DNA)
Eμ: intronic enhancer at IgH locus, 5' to mu region
FT: FLAG tag
GLT: germline transcripts/transcription
HR: homologous recombination
Ig: Immunoglobulin (antibody)
Igh: immunoglobulin heavy chain (gene region)
Lig IV: Ligase IV
MDC1: mediator of damage checkpoint protein 1
MMR: mismatch repair
NER: nucleotide excision repair
NHEJ: non-homologous end-joining (of DNA)
PIKK: phosphatidylinositol-3-kinase-like kinase
RIF1: Telomere-associated protein RIF1
SHM: somatic hypermutation
SSA: single-strand annealing
TCR: T-cell receptor
TdT: terminal deoxynucleotidyl transferase
V(D)J recombination: variable diversity joining genes
ZMYND8: protein kinase C binding protein

Chapter 1: Introduction

1.1 Antibodies and the Adaptive Immune System

Organisms need to defend themselves against a wide variety of potential pathogens and foreign organisms. While a limited number of P/MAMPs (pathogen/microbe associated molecular patterns) like lipopolysaccharides are genetically hardcoded to be innately recognized as foreign, our bodies rely on the adaptive immune system to respond to novel and diverse molecular patterns presented by pathogens. A central challenge of the adaptive immune system is thus how to create specifically tailored but potentially infinitely variable targeting of these patterns. When the adaptive immune system is able to recognize such a molecular pattern as foreign, it is termed an antigen.

B cells are able to recognize and target such antigens by creating antibodies. The potential antibody-antigen repertoire in humans is immense > 10^{11}(Glanville, Zhai, Berka et al. 2009). However, the actual repertoire size is less than that because it is functionally limited by the total number of B-cells present in the body at a given time, overlapping specificity for antigens (as a result of clonal expansion and diversification of selected B cells), and limited encounters with the antigens themselves. Regardless of these limitations, the size of the human genome is only $3x10^9$ bases and consequently, diversification of antibodies by B cells is accomplished by the combinatorial rearrangement of multiple gene segments. The

induction, regulation, and translocations of these gene segments to produce effective antibodies involves elements unique to the immune system (formation of the breaks) and processes ubiquitous to all cells of the body (DNA repair).

Structure and Function

An antibody or immunoglobulin (Ig) is the secreted form of a B cell receptor (BCR). They can bind to a wide diversity of antigens and a limited group of effector molecules and cells.

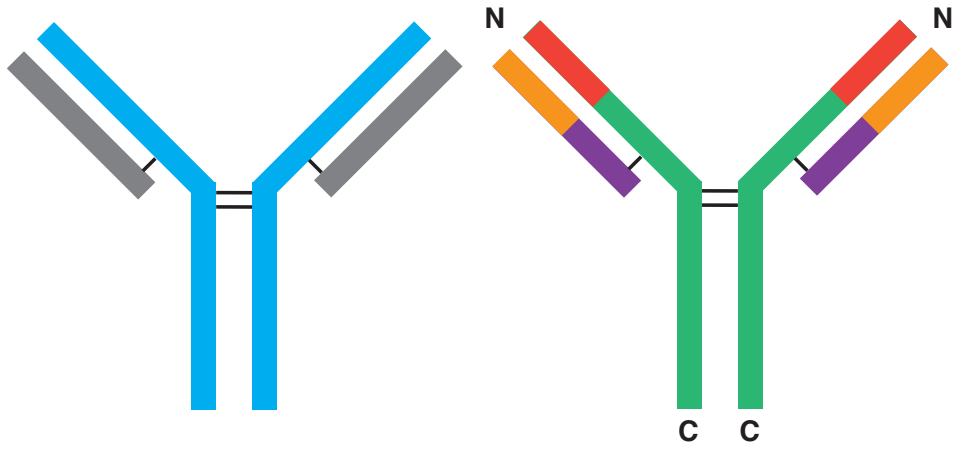


Figure 1 Antibody/Immunoglobulin (Ig) Molecule

Igs are proteins secreted by B cells / plasmablasts. Four polypeptide chains comprise the antibody. (LEFT) The blue regions are the heavy (IgH) chains and the grey regions are the light (IgL) chains. Each pair of heavy and light chains is identical to the other for a particular B cell. Disulfide bridges, indicated by black lines, link the four chains. Different Ig classes have different structures. (RIGHT) The N-termini of the heavy and light chains are indicated by N, the C-termini of the heavy chains are indicated by C. Red highlights the variable region of the heavy chain (V_H). Green highlights the constant region of the heavy chain (C_H). Orange highlight the variable region of the light chain (V_L). Purple highlights the constant region of the light chain (C_L).

Igs are functionally divided into the variable and constant regions. The variable region comprised of the V_H and V_L is diversified through the V(D)J recombination and later somatic hypermutation (SHM) reactions. The constant region is comprised of the C_L domain, kappa or lambda in mice and humans, and remains fixed for a particular antibody. It is also comprised of the C_H which confers the particular isotype to an antibody. The diversity generating reactions VDJ recombination and SHM do not alter the constant region, but the C_H region, which defaults to IgM, can undergo the process of class switch recombination (CSR). In CSR, IgM can be switched to IgG, IgA, IgE or any of their subclasses (identified by a numerical subscript after the class letter). Functionally, the ability to bind to different antigens is due to the variable domain, which comprises about the first 110 amino acids of the N-terminus. The ability to interact with effector molecules and cells is mediated by the constant domain. While every clonally derived antibody has a unique variable region, the constant region is common to all antibodies of that particular isotype.

Primary Diversification: V(D)J Recombination

The initial variability of an antibody derives from V(D)J recombination. In the germline, the immunoglobulin locus has multiple different gene segments. During the development of B and T cells, one copy of each segment is retained to produce a functional BCR (B-cell receptor, essentially a membrane bound version of the antibody encoded by that cell) or TCR, respectively. The other segments are

deleted through a recombination-deletion reaction (Brack et al. 1978), (Kurosawa et al. 1981). This gene re-arrangement is fundamentally a controlled translocation event with programmed DNA damage and ligation (Tonegawa 1983; (Schatz & Ji, 2011)).

Although this is a process that is controlled spatially and temporally within the body, the selection of V(D)J segments for recombination is random. There are > 3.5 million possible combinations for an antibody however, in actuality certain combinations of V(D)J heavy chain and VJ light chain segments do not make stable BCRs and thus those combinations never appear in the human repertoire. In addition to the combinatorial diversity, the enzyme terminal deoxynucleotide transferase (TdT) add nucleotides at the joining junctions. V(D)J recombination is mediated by the RAG1/RAG2 recombinases.

Secondary Diversification

Further antibody variability differs from V(D)J recombination in two key elements. First, the initiating enzyme is Activation Induced Cytidine Deaminase (AID) and not the RAG complex. Second, the diversification occurs after B cells have developed, when they are in the periphery—not the originating bone marrow, and after they have encountered an antigen.

Somatic Hypermutation (SHM)

After the initial rearrangement of gene segments takes place in B cells in the central lymphoid organs (bone marrow), the immunoglobulin variable region can further diversify in a process termed somatic hypermutation (SHM). SHM takes place in the peripheral lymphoid organs. Point mutations are introduced in the variable region, which increases sequence diversity. B cells interact with T cells in the germinal center. B cells in which SHM has resulted in variable regions with better affinity for the antigen presented by a T cell are induced to divide and mutate more. The process repeats over many cycles, with B cells with high affinity for a specific antigen evolutionarily outcompeting those that do not. This entire process is termed affinity maturation.

Class Switch Recombination (CSR)

The previous three mechanisms: V(D)J recombination, junctional diversity, and SHM all alter the variable region and therefore affect the specificity of the antibody for an antigen. The last mechanism of antibody diversity in mouse and humans, isotype switching also known as class switch recombination (CSR), retains the antigen specificity of an antibody while changing the effector function. This is accomplished by switching the default constant region of the heavy chain locus (Constant mu - C_{μ} and also Constant delta - C_{δ}) to another constant region. The following sections will detail the different isotype classes and functions produced by CSR, and the mechanism by which CSR occurs.

Figure 2: *IgH* locus recombination/deletion during Class Switch Recombination

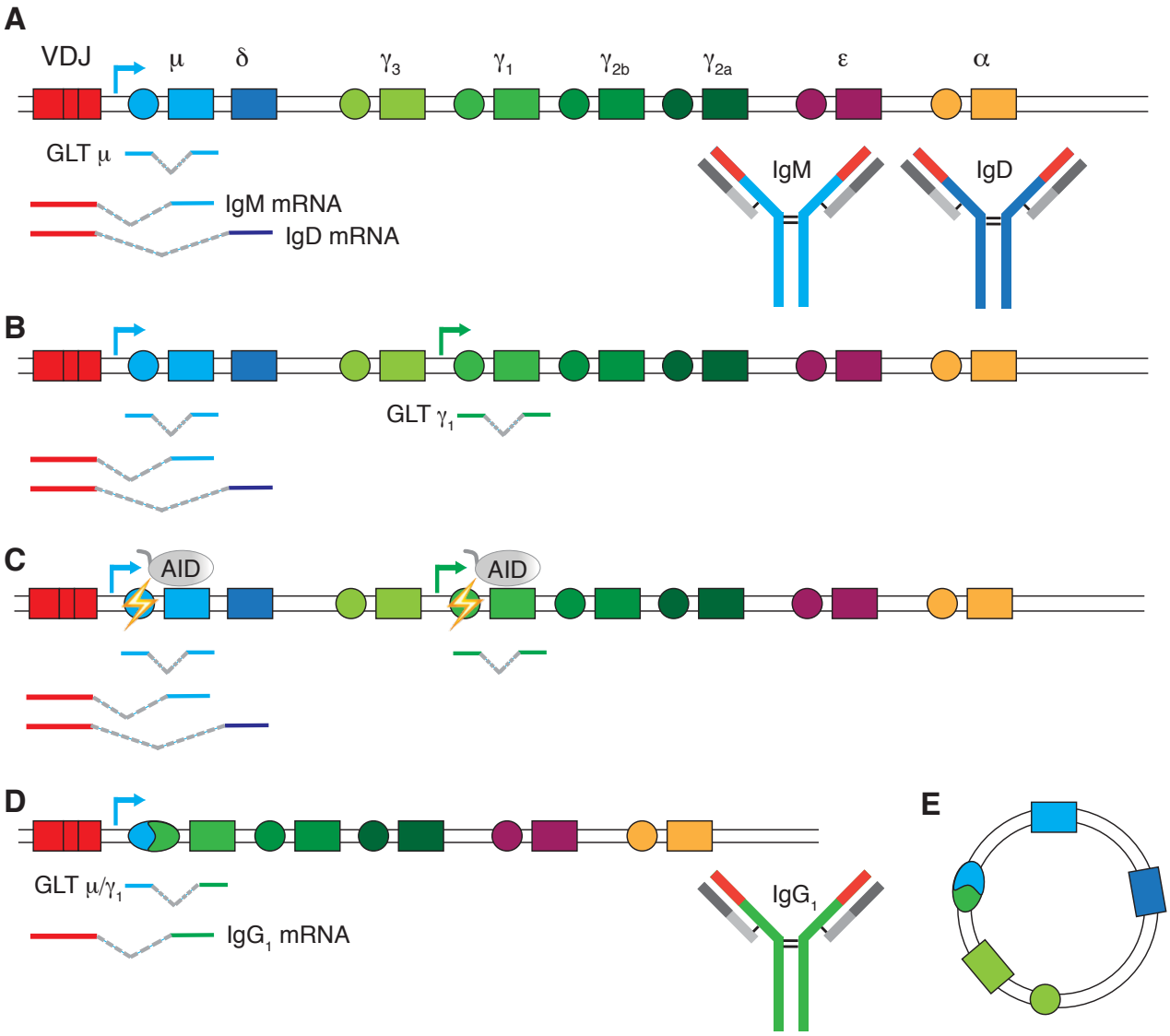
(A) Locus arrangement and antibody production: The *IgH* locus is arranged on chromosome 12 of mice. The VDJ region immediately precedes the mu, delta, gamma3, gamma1, gamma2b, gamma2a, epsilon, and alpha heavy chains (blue, green, green, green, green, maroon, orange, respectively). Each heavy chain (except for delta) has an intronic promoter (colored arrow if activated), repetitive switch region (colored circle), and heavy exons (colored rectangle). Transcription through mu and delta is constitutive in naïve B cells. mRNA from the VDJ region and C_H region combine to form the Ig heavy chain (color correspond designated antibody at right).

(B) Activation: with activation in vivo or ex vivo by specific cytokines (here LPS + IL-4), germ line transcripts (sterile, non-coding) GLT appear at the targeted acceptor I_x-S_x-C_x locus.

(C) Targeting by AID: AID is guided to the switch regions through multiple mechanisms detailed later. At the transcribed switch region (donor and acceptor) AID deaminates cytosine to uracil. Multiple deaminations lead to double-strand break formation.

(D) Synapsis and Ligation: the upstream switch region synapses with its downstream counterpart and joins to form a hybrid switch region (ellipse). The recombined locus will now produce antibodies from the downstream / switched heavy chain (here IgG1).

(E) Post-switch excision: the excised genomic region is circularized. Transcripts from the still active downstream intronic promoter can be detected.


A
VDJ
μ
δ
γ3
γ1
γ2b
γ2a
ε
α
GLT μ
IgM mRNA
IgD mRNA
IgM
IgD
B
GLT γ1
C
AID
AID
D
GLT μ/γ1
IgG1 mRNA
IgG1
E

Immunoglobulin Heavy Chain (C_H) Isotypes: Structure and Function

The different C_H regions of antibody isotype classes determine the interaction partners and function. First, some isotypes can bind cell surface receptors and modulate the immune response (i.e. IgG1 and IgG3 bind $F_c\gamma$ receptors on phagocytic cells) (Nimmerjahn, 2008). Second, some isotypes can bind Fc transporters to actively transport antibodies (i.e poly-immunoglobulin receptors on mucosal epithelial cells transport IgA to mucous secretions, tears and milk, and IgG to the fetal blood circulation by placental transfer) (Kaetzel, 2005). Third, the Fc can bind the C1q protein to activate complement. Fourth, Fc can bind to other Fcs to form multimeric structures (i.e. IgM forms pentamers in humans and mice, IgA forms dimers as the mucosal secretory form) (Stavnezer, Guikema, & Schrader, 2008).

1.2 Steps of CSR

Abstractly, CSR is an irreversible change in protein expression. The switch from IgM to another isotype (IgG, IgA, IgE) means that a switch in gene expression occurs and this switch has multiple possible outcomes. Fundamental questions in CSR are thus: (1) what activates CSR, (2) what causes isotype specificity, and (3) how is this mediated at a genetic level?

Transcription and CSR

Early on, CSR was found to be dependent on transcription. In an IgM+ B cell line, I.29 lymphoma, CSR to the IgE, IgA, and IgG2a isotypes correlated with the expression of non-coding, sterile transcripts to the corresponding C_H region (Stavnezer-Nordgren, 1986). Similarly, an Abelson murine leukemia virus transformed pre-B cell line showed analogous sterile transcripts when switching to IgG2b (Yancopoulos, 1986). In the next few years, GLTs were identified in primary B cells along with the cytokine combinations that corresponded to specific isotype switching. LPS stimulates GLT gamma2b and gamma3 and thus switching to IgG_{2b} and IgG_3 (Lutzker, 1988), (Severinson, 1990), (Snapper & Paul, 1987). LPS and IL-4 stimulates GLT gamma1 and epsilon and thus switching to IgG_1 and IgE (Lutzker, 1988), (Snapper & Paul, 1987), (Stavnezer, Guikema, & Schrader, 2008). TGF-β stimulates GLT alpha and switching to IgA (Lebman, 1990) (Shockett & Stavnezer, 1991). Confirming the necessity of GLT for CSR, deletion of the intronic

promoter for GLT gamma1 eliminated CSR to IgG1 (Jung, 1993); likewise, the replacement of the GLT gamma2b promoter, exon and splice donor site with a NeoR gene in the opposite transcriptional orientation eliminated CSR to Ig2b (Zhang, Bottaro, Li, Stewart, & Alt, 1993). However, other transcriptional elements could substitute for intronic promoter. Replacement of the I_α and exon by a hypoxanthine phosphoribosyltransferase (HPRT) minigene in the correct transcriptional orientation was sufficient for CSR to IgA (Harriman, Bradley, Das, Rogers-Fani, & Davis, 1996).

Transcription and Mutations

Transcription was found to strongly correlate with [what would later be identified as] AID dependent mutations. Rearranged IgH loci had mutation signatures similar to SHM. However, in addition to base pair substitutions found in the variable regions, the switch regions also had deletions, insertions and inversions (Dunnick, Wilson, & Stavnezer, 1989), (Dunnick, Hertz, Scappino, & Gritzmacher, 1993). By inserting a promoter upstream of C_H regions, mutations could be induced in the normally unmutated constant region (Peters & Storb, 1996). These mutations were later found to be AID-dependent and located 3' of the intronic enhancers, which are 5' of the recombination sites (Nagaoka, Yamamura, Kinoshita, & Honjo, 2002) (Petersen, et al., 2001). Combination $UNG^{-/-}$ $Msh2^{-/-}$ mice allowed increased sensitivity to detect AID targeting because deaminations could not be correctly repaired. Transcription was found to be essential for mutations in dsDNA in vitro

(Chaudhuri, et al., 2003) and in bacteria (Ramiro, Stavropoulos, Jankovic, & Nussenzweig, 2003). In vivo, mutations began 150 bp 3' of the Transcription Start Site (TSS) of the S region intronic promoter and continued even beyond the S region (Xue, Rada, & Neuberger, 2006). Like SHM, there appeared to be no strand bias and the mutation signature conformed to the RGYW motif.

Discovery of AID as the mutational enzyme

AID was discovered as one of four factors upregulated in a screen of the mouse CH12F3 (CH12) B cell lymphoma line (Muramatsu, et al., 1999). Specifically, CH12 cells were activated to undergo CSR from IgM+ to IgA+ and cDNA subtracted, expressed genes were compared to unactivated (IgM+) cells. AID was selected for further study because it showed a dramatic increase in transcription from practically nothing in unactivated CH12 cells to at least a 10-fold increase in activated CH12 cells. AID was subsequently proved necessary for SHM and CSR in mice (Muramatsu, et al., 2000) and humans (Revy P, et al., 2000). It was also shown necessary for gene conversion (a mechanism of generating antibody diversity not present in mice or humans, but utilized by birds, rabbits, cows and pigs) in chickens (Arakawa, Hauschild, & Buerstedde, 2002).

Mechanism of Mutation by AID

AID was first recognized as a homolog of APOBEC1 (apolipoprotein B mRNA editing catalytic polypeptide 1) (Muramatsu, et al., 1999). APOBEC1 is a RNA

deaminase, which deaminates cytosine to uracil. AID catalyzes the same reaction but on DNA (Figure 3).

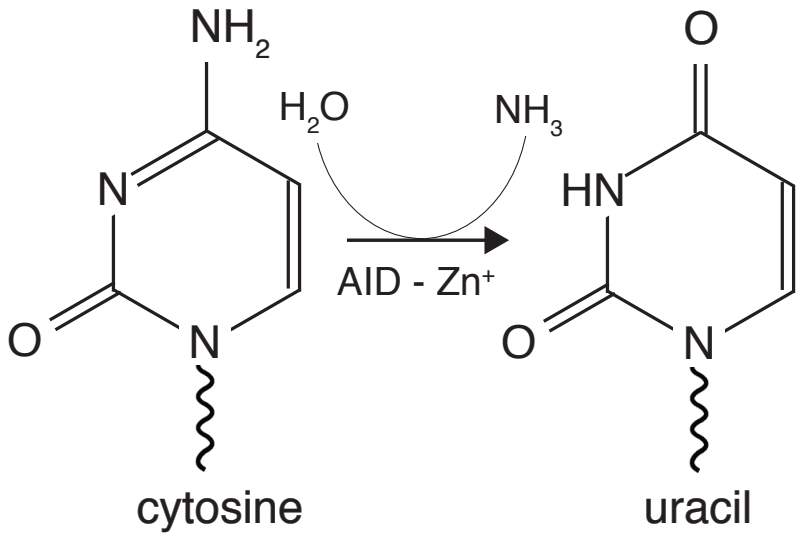


Figure 3: Deamination of cytosine to uracil by AID

AID deaminates cytosine to uracil. Note: the name activation induced cytidine deaminase (AID) is actually a misnomer. Because AID acts on DNA, not RNA the technically correct name should be activation induced deoxycytidine deaminase. The cytidine is a holdover from the discovery of AID which was thought to act on RNA due to its homology to APOBEC1.

Switch Regions: Sites of Recombination

Switch regions are sequences of G-rich, repetitive DNA elements ranging from 1-12 kb long and precede the constant regions (with the exception of C_δ). The regions contain many RGYW repeats, the preferred deamination targets of AID

(Rogozin & Kolchanov, 1992). It was postulated that these regions are the target sites of recombination because, most *IgH* locuses of isotype switched cells showed junctions between upstream S_{μ} and downstream acceptor (S_X) regions (Matthews, Zheng, Di Menna, & Chaudhuri, 2014). Switch regions were shown to be the sites of recombination when the deletion of $S_{\gamma 1}$ completely abolished CSR to IgG1 but left other isotypes unaffected (Shinkura, Tian, Chua, Fujiwara, & Alt, 2003).

AID Targeting

Recombination at the switch regions requires a DNA double-strand break (DSB) intermediate. However, AID was found to only act on single-stranded DNA (ssDNA) in vitro (Bransteitter, Pham, Scharff, & Goodman, 2003), (Chaudhuri, et al., 2003), (Dickerson, Market, Besmer, & Papavasiliou, 2003), (Petersen-Mahrt, Harris, & Neuberger, 2002), (Pham, Bransteitter, Petruska, & Goodman, 2003), (Ramiro, Stavropoulos, Jankovic, & Nussenzweig, 2003), and (Sohail, Klapacz, Samaranayake, Ullah, & Bhagwat, 2003). A solution was proposed where the transcribed S region, could hybridize with the template strand and form a R-loop (a RNA:DNA hybrid) which would free the non-template strand as ssDNA for AID mutation. R loops were experimentally implicated in CSR when a synthetic S region known to form R loops in vitro facilitated CSR when placed in cells. But, if the sequence was transcribed in the reverse orientation, precluding R loop formation, CSR decreased greatly (Shinkura, Tian, Chua, Fujiwara, & Alt, 2003).

Although R loops are necessary for CSR, it fails to explain how AID targets the DNA strand still in the RNA:DNA hybrid (the template strand). It was identified that the RNA exosome accumulates with the *IgH* locus in an AID dependent manner, associates with AID and is necessary for efficient CSR (Basu, et al., 2011). The exosome complex was sufficient to deaminate both strands of a transcribed SHM substrate in vitro. It is proposed that the exosome degrades the R-loop, an upon collapse the DNA may misalign, resulting in ssDNA on both the template and non-template strand (Matthews, Zheng, Di Menna, & Chaudhuri, 2014).

AID is actually a fairly inefficient enzyme. Mutations in the variable regions only occur at a frequency of about 1 mutation per 1 kb per cell per generation (Yamane, et al., 2011). While an inefficient enzyme is desirable for the rest of the genome to prevent potentially oncogenic AID off-target activity (Casellas, et al., 2016), the mutation rate must be higher to effect the theoretical minimum 4 mutations to induce CSR within a reasonable time.

In addition to transcriptional dependence and a high concentration of RGYW motifs present in the switch region, B cells target AID to the switch regions through two other key mechanisms. It was found that Spt5, a factor associated with stalled RNA polymerase (Pol II) recruits AID to sites of stalled transcription (Pavri, et al., 2010). Additionally, the GLTs undergo RNA processing and in the process excise intronic

segments. These spliced regions are debranched and form G-quadruplexes which target AID to S region DNA (Zheng, et al., 2015).

1.3 DSB formation in CSR

DSBs are Essential for CSR

Although AID is essential for SHM and CSR, it does not directly cause the DNA breaks required for CSR. Instead processing of the deamination site can lead to non-mutating repair, mutations as in SHM, and DSBs as in CSR.

40 years ago, DSBs were postulated to be essential for CSR (Honjo & Kataoka, 1978). A model called allelic deletion, was theorized by examining the frequency of heavy chain deletions in mouse myelomas. By comparing the heavy chain deletions with the corresponding isotype expressed the germline order of heavy chain segments could be inferred. A looping out mechanism was suggested to account for these rearrangements. DNA loops as a result of immune recombination events were detected from variable regions in T cell receptor rearrangement (Fujimoto & Yamagishi, 1987), (Okazaki, Davis, & Sakano, 1987), (Okazaki & Sakano, 1988); from variable regions in the IgH locus (Abe & Shiku, 1989), (Toda, Hirama, Takeshita, & Yamagishi, 1989); and from variable regions in the IgL locus (Toda, Hirama, Takeshita, & Yamagishi, 1989), (Mccormack, et al., 1989). The

proposed excision products in CSR were finally able to be isolated (von Schwedler, Jack, & Wabi, 1990), (Iwasato, Shimizu, Honjo, & Yamagishi, 1990), (Matsuoka, Yoshida, Maeda, Usuda, & Sakano, 1990), (Yoshida, et al., 1990). Recombination products due to CSR were fundamentally different than those derived from variable regions. All variable regions had defined breakpoints encompassing V(D)J (in the heavy chain, VJ in the TCR and light chain), but CSR has breakpoints that spanned the entire multi-kilobase switch regions. Thus, breaks in the variable region are specific, but breaks in the switch regions are variable.

Direct evidence of the DSBs at the switch regions came from Ligation-Mediated PCR (LM-PCR). Blunt DNA DSBs were detected at $S_{\gamma 3}$ in mouse splenic B cells (Wuerffel, Du, Thompson, & Kenter, 1997), S_{μ} in human B cells (Catalan, et al., 2003); but staggered breaks were found to be much more abundant (Rush, Fugmann, & Schatz, 2004). These breaks were found to be AID and UNG dependent (Schrader, Linehan, Mochegova, Woodland, & Stavnezer, 2005).

Factors involved in DSB repair were also detected at the IgH locus, specifically γ-H2AX foci were found at IgH in ex vivo stimulation of splenic B cells, and these foci were also AID-dependent (Petersen, et al., 2001). Since that time, many DNA DSB repair factors were discovered to contribute to CSR [reviewed in: (Daniel & Nussenzweig, 2013), (Alt, Zhang, Meng, Guo, & Schwer, 2013)].

Deaminating Cytosines creates Uracils, which can be processed into DSBs

Uracils produced by AID targeting can result in accurate repair, mutagenic repair, or double-strand breaks. The base excision repair (BER) pathway is essential for all of these functions and the mismatch repair (MMR) pathway primarily contributes to mutagenic repair but can also take place in DSB formation –although the mechanism is poorly characterized. Nucleotide excision repair (NER) does not play a role in uracil processing. An overview of the BER process and DSB formation is (1) AID deaminates cytosine to uracil, (2) UNG2 (the nuclear uracil-DNA glycosylase) recognizes the uracil and removes it to create an abasic (AP for apurinic, apyrimidinic) site, (3) apurinic/apyrimidinic endonuclease (APE) excises the abasic residue to create a gap or nick in the single-strand. And (4) the presence of nearby nicks on opposing strands creates a DSB.

Evidence for UNG as a component of the mutational process was first identified in a bacterial system where deletion of UNG increased the mutational signature of AID (Petersen-Mahrt, Harris, & Neuberger, 2002). Also, mutational signatures in $UNG^{-/-}$ DT40 B cells changed to reflect repair by replication instead of dU removal (Di Noia & Neuberger, 2002). In mice, UNG deficiency impaired SHM and CSR (Rada, et al., 2002). A clinical phenotype of blocked CSR and reduced SHM in several human patients was found to be a result of UNG mutations (Imai, et al., 2003).

Two different endonucleases have been implicated in removing the abasic site: APE1 and APE2 (encoded by the genes *Apex1 and Apex2*, respectively). *Apex1*$^{-/-}$ mice and B cells are not viable. Heterozygotes (*Apex1*$^{+/-}$) and *Apex2*$^{-/-}$ mice both show a mild reduction in CSR, but the combined *Apex1*$^{+/-}$ and *Apex2*$^{-/-}$ mouse shows a much stronger reduction (Guikema, et al., 2007). This would indicate that both contribute to CSR, but in the CH12 cell line, homozygous deletion of *Apex1* results in an 80% reduction in CSR, but *Apex2* deletion has no effect (Masani, Han, & Yu, 2013).

The mismatch repair (MMR) pathway also repairs U:G mismatches caused by AID. Originally thought to only affect mutations resulting in SHM, MMR works in concert with BER to generate DSBs [reviewed in (Pena-Diaz & Jiricny, 2012) and thus facilitate CSR. The MMR pathway is made up of two complexes: MutSα (MSH2 and MSH6), which recognizes the mismatch, and MutLα (MLH1 and PMS2) which signals downstream effectors. While MutSα promotes SHM and CSR (Chahwan, et al., 2012), MutLα only functions in CSR. The specific MMR components that reduce CSR are: Msh2 and Msh6 (Stavnezer & Schrader, 2006), (Martin, et al., 2003), (Martomo, Yang, & Gearhart, 2004), (Li, et al., 2004), (Schrader, Vardo, & Stavnezer, 2002), (Schrader, Edelmann, Kucherlapati, & Stavnezer, 1999); Exo1 (Bardwell, et al., 2004); PMS2 (Ehrenstein, Rada, Jones, Milstein, & Neuberger, 2001), and MLH1 (Chahwan, et al., 2012). It has been proposed that MMR enhances the activity of BER in forming DSBs by processing more isolated nicks

from the BER pathway into DSBs. If this is the case, a deficiency in MMR should result in a reduction in CSR, but a combined deficiency in BER and MMR should be equivalent to a deficiency in BER. In opposition to this, MSH6 shows no reduction in CSR by itself, but a double knockout of MSH6 and UNG shows a greater reduction in CSR than by UNG alone (Shen, Tanaka, Bozek, Nicolae, & Storb, 2006). MMR must therefore be able to generate DSBs in CSR independently of the BER pathway.

1.4 Synapsis

The synapsis or the close juxtaposition of DSBs created in activated switch regions to facilitate end-joining of the DSBs, completes CSR. The resolution of DSBs by productive CSR is different from the repair of a single pathological DSB or the DSBs created by V(D)J recombination. In a pathological break, the two ends are generated simultaneously and are in close proximity to each other. Similarly, the RAG1/2 endonuclease complex brings the two RSSs into a synaptic complex before they are cut. CSR does not initiate DSBs and translocate those breaks with the same spatiotemporal precision. Topologically, CSR requires the efficient translocation of a DSB in the Sμ (donor) region to a downstream accepter switch region potentially over 100 kb distant. Timing-wise, the formation of DSBs is not sufficient for and can be decoupled from synapsis. Internal deletions in Sμ are

common in ex vivo stimulated B cells even if the cells have not yet switched (Alt, Rosenberg, Casanova, Thomas, & Baltimore, 1982). These deletions without downstream joining are termed intra-switch recombination reactions and the cells retain their initial isotype expression. Several proteins have been implicated in the process of synapsis. *53BP1*$^{-/-}$ and *H2AX*$^{-/-}$ B cells have extreme CSR defects but they do exhibit much higher rates of intra-switch recombination (Manis, et al., 2004), (Ward, et al., 2004), (Reina-San-Martin, Chen, Nussenzweig, & Nussenzweig, 2007). 53BP1 deficiency also decreases the joining of distal DSBs more than breaks occurring in close proximity (Bothmer, et al., 2011). Chromosomal conformation capture experiments have provided a model where the 5' and 3' regions of the IgH locus, termed the Eμ (enhancer mu) and 3'RR (3' Regulatory Region) come in three dimensional proximity to facilitate synapsis (Ju, et al., 2007), (Kenter, et al., 2012), (Sellars, Reina-San-Martin, Kastner, & Chan, 2009), (Wuerffel, et al., 2007).

1.5 DSB Repair Pathways in CSR and Beyond

Repair of DSBs in CSR

Double strand breaks (DSBs) are perhaps the most critical type of DNA damage for cells. A single DSB break in an essential gene is sufficient to induce cell death or if left unrepaired elsewhere—induce apoptosis (Jackson, 2002). DSBs pose

such a threat to genome integrity because in addition to the mutagenic nature of DNA damage, the two free DNA ends in a DSB are no longer held together by chromatin structure or complementary base pairing. Dissociation of the ends can result in repair at inappropriate locations—i.e. a chromosomal translocation. Such chromosomal rearrangements are frequently found in lymphomas, leukemias, and solid tumors (Kuppers, 2005), (Nussenzweig & Nussenzweig, 2010), (Tsai & Lieber, 2010), (Zhang, et al., 2010). Moreover, deficiencies in the DNA DSB Response predispose to carcinogenesis (Jackson & Bartek, 2009), (Ciccia & Elledge, 2010).

Sources of double strand breaks

DSBs can be classified according to the source and cause (Goodzari & Jeggo, 2013). Accidental or pathologic DSBs can result from exogenous factors (ionizing radiation or radiomimetic compounds), but also endogenous sources (replication fork collapse). DSBs can also be programmed as in chromosomal crossover in meiotic recombination and are extensively utilized by the adaptive immune system in jawed vertebrates (Cooper & Alder, 2006) in V(D)J recombination and CSR.

Recognition of the DSB and Repair Pathway Choice

Various DSBs require different modes of repair. The location of the break, phase of the cell cycle, and origin of the break dictate repair choices. Although pathway

choice begins as early as the primary detection of the break site, the irreversible decision point is the amount of DNA resection. Classic Non-homologous end joining (c-NHEJ) requires minimal end-resection, while homologous recombination (HR) requires extensive end-resection. Two alternative pathways, alternative NHEJ (a-NHEJ) and single-strand annealing (SSA), also require end-resection.

classic Non-homologous End-Joining (c-NHEJ)

Originally discovered in higher eukaryotes because yeast predominantly use HR, c-NHEJ is the primary joining pathway for two-ended DSBs like those occurring from irradiation, V(D)J recombination, and CSR (Goodzari & Jeggo, 2013). c-NHEJ quickly ligates broken DNA to suppress chromosomal translocations (Difilippantonio, et al., 2000). Cells depend on c-NHEJ during the G1 phase because it does not require a homologous template for repair. The fast kinetics and template independence result in a slight loss of information—about 1-4 nt resection at the break site.

Three steps comprise c-NHEJ repair of DSBs: (1) binding and stability, (2) processing, and (3) ligation. First, the Ku70/Ku80 toroidal heterodimer (Ku) binds to a DSB initiates c-NHEJ. Ku is sequence agnostic because it binds to the sugar backbone of DNA and not the bases themselves (Walker, Corpina, & Goldberg, 2001). Ku recruits DNA-PKcs (DNA Protein Kinase catalytic subunit) in a DNA dependent manner (Gottlieb & Jackson, 1993) to form the DNA-PK holoenzyme

(Uematsu, et al., 2007). DNA-PK phosphorylates other NHEJ factors (Meek, Douglas, Cui, Ding, & Lees-Miller, 2007) itself via autophosphorylation which is essential for NHEJ (Chan & Lees-Miller, 1996), (Douglas, et al., 2007), (Cui, et al., 2005), (Soubeyrand, Pope, Pakuts, & Hache, 2003). DNA-PK regulates the NHEJ pathway by preventing extensive end-processing (along with other factors), inactivating itself, and dissociating for the following steps (Neal & Meek, 2011). The next steps and components in c-NHEJ are not inherently sequential; they are flexible depending on the nature of the DNA damage and potential differences in ligation method (Lieber, 2008), (Ma, et al., 2004), (Reynolds, et al., 2012). (2) Ligation of a DSB requires 5' phosphate and 3'-OH compatible ends. DSBs generated from exogenous DNA damage or endogenous physiologic damage (as in V(D)J or CSR do not have these structures. The DSB break must be processed by nucleases and polymerases to facilitate ligation. These nucleases include: Artemis, Werner (WRN), and. APLF. Artemis has 5' endonuclease activity and produces a blunt duplex end by nicking 5' overhangs. It can also remove 5'—3' ssDNA via it's exonuclease activity. Finally, Artemis can remove 3'-phosphoglycolate groups from DNA termini (Ma, Pannicke, Schwarz, & Lieber, 2002), (Povirk, Zhou, Zhou, Cowan, & Yannone, 2007). WRN binding to Ku and the factor XRCC4 activates it's 3'→5' exonuclease activity (Cooper, et al., 2000), (Kusumoto, et al., 2008), (Perry, et al., 2006). APLF is an endonuclease and can also facilitate binding by resection 3'→5' overhangs (Kanno, et al., 2007). If there is sufficient complexity in the break, gaps in DNA can be filled in by the family X

polymerases. Polymerase μ is template dependent and can polymerize across a discontinuous strand using dNTP and rNTP (Nick McElhinny & Ramsden, 2004). Polymerase λ is template independent and can remove deoxyribosephosphates (Ramadan, Shevelev, Maga, & Hubscher, 2004), (Daley, Laan, Suresh, & Wilson, 2005). In V(D)J rearrangement in pro-B cells and developing T-cells, the terminal deoxynucleotidyl transferase (TdT) add nucleotides between the gene segments (Chang & Bollum, 1986). (3) Ligation of the processed DSBs occurs by Ligase IV (Lig IV), the NHEJ specific ligase (Grawunder, et al., 1997), (Teo & SP, 1997), (Wilson, Grawunder, & Lieber, 1997), XRCC4 stabilizes the interaction and also serves as a platform for other factors with Ku (Li, et al., 1995). The protein XLF (Cernunnos) (Ahnesorg, Smith, & Jackson, 2006), also promotes NHEJ by repriming Lig IV and XRCC4 (Riballo, et al., 2009), and an XLF deficiency leads to immunodeficiency and radiosensitivity (Menon & Povirk, 2017). XLF is technically redundant in normal cell lines, and the persistence of NHEJ is due to paralogue of XRCC4 and XLF (PAXX) (Kumar, Alt, & Frock, 2016), but in cells deficient for ataxia telangiectasia-mutated (ATM) DSB response factor (a key component of the DSB sensing and repair pathway) it is essential.

Homologous Recombination (HR)

Cells can use the sister chromatid as a template to repair DSBs. This process, termed homologous recombination (HR), requires the cell to be in S/G2 phase, after DNA duplication has occurred (Johnson & Jasin, 2000). The homologous

chromosome in cells that have not replicated their DNA cannot be used for HR, presumably due to sequence differences preventing heteroduplex formation. HR primarily acts at the replication fork in fork restart and the repair of one-ended DSBs (Adamo, et al., 2010), (Petermann & Helleday, 2010). The steps of HR are: (1) 3' single ended DNA (ssDNA) formation, (2) coating the ssDNA with RPA, (3) displacement of RPA by RAD51 to form a nucleoprotein filament, (4) formation of heteroduplex DNA and the Holliday junction (Bzymek, Thayer, Oh, Kleckner, & Hunter, 2010), (5) branch migration and (6) resolution. Because HR is not utilized in CSR, the following description of HR will focus on the characterization of HR as a mutually exclusive pathway to NHEJ. While HR is functionally unavailable in G1, it must outcompete NHEJ in S/G2 phases. This is accomplished in two ways: (1) displacement of NHEJ promoting factors at the DSB, and (2) activation of HR promoting factors. Because Ku shields DNA from exonucleases and is quickly recruited to DSBs, Ku must be displaced to permit end resection. Ku displacement has been proposed to occur by the ATM dependent phosphorylation of CtIP displaces Ku in the S/G2 phase of the cycle to permit HR (Chanut, Britton, Coates, Jackson, & Calsou, 2016) and the phosphorylation of Ku itself (Lee, et al., 2016). Additionally, extensive resection of the DSB is facilitated by the activation of helicases and nucleases: DNA2, BLM, WRN, CtIP, and Exo1 (Symington & Gautier, 2011), (Sturzenegger, et al., 2014).

Chromatin at the DSB: 53BP1 and Pathway Choice

In addition to factors that bind directly to DNA as a result of a DSB, the chromatin surrounding the break is heavily modified to facilitate repair (Price & D'Andrea, 2013). First, the DSB is recognized by the MRN (MRE11-RAD50-NBS1) complex. The NBS1 subunit recruits ATM (ataxia-telangiectasia mutated) (Lee & Paull, 2005), (You, Chahwan, Bailis, Hunter, & Russell, 2005), a serine/threonine kinase member of the phosphatidylinositol-3-kinase-like kinase family (PIKK) (Maréchal & Zou, 2013). The autophosphorylated and active form of ATM then phosphorylates the histone variant H2Ax, becoming ϒH2AX. Mice lacking this histone variant display the hallmarks of deficient DSB repair, including radiosensitivity and immune deficiency (Celeste, et al., 2002). The MDC1 (mediator of damage checkpoint protein 1) protein then recognizes the ϒH2AX and recruits more MRN and ATM to the chromatin (Stewart, Wang, Bignell, Taylor, & Elledge, 2003), (Lou, et al., 2006). ATM and MDC1 then recruit the E3 ubiquitin ligase RING finger 8 (RNF8), which also a second E3 ubiquitin ligase, RNF168. Through a longer cascade, H2AK13 and H2AK15 are ubiquitinated which recruits 53BP1 along with H4K20me1 andH4K20me2 (Panier & Boulton, 2014). 53BP1 blocks homologous recombination by inhibiting DNA end-resection (Bunting, et al., 2010). It does so, by excluding another factor in HR, BRCA1 (Chapman, Sossick, Boulton, & Jackson, 2012), (Bouwman, et al., 2010).

1.6 RIF1

RIF1 is an essential protein for NHEJ and CSR. It is the single, phospho-dependent interactor and effector of 53BP1 to promote these processes. In this section I will cover the background of RIF1 as a protein and its known functions.

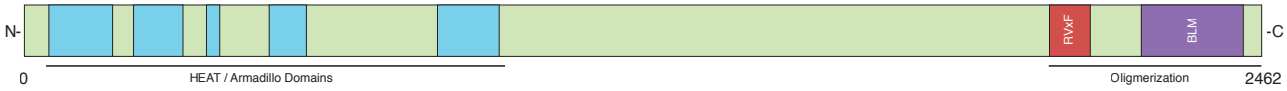


Figure 4: Schematic of RIF1 protein from *M. musculus*

Diagram of RIF1 protein sequence highlight N-terminal HEAT/armadillo domains (light blue) and C-terminal RxVF/Protein Phosphatase 1 (PP1) binding (red) and BLM binding (purple) domains. C-terminal region necessary for oligomerization is underlined.

Structure of RIF1

Murine RIF1 is a large (2462 aa) protein. The N-terminal half of the protein contains several α-helical HEAT/Armadillo domain repeats which are required for IR induced foci formation (Escribano-Díaz, et al., 2013). The C-terminal end contains a FxVF domain that interacts with Protein Phosphatase 1 (PP1), Bloom syndrome protein (BLM) and is required for oligomerization of RIF1 proteins (Silverman, Takai, Buonomo, Eisenhaber, & De lange, 2004), (Sreesankar, Senthilkumar, Bharathi, Mishra, & Mishra, 2012), (Xu, et al., 2010). The c-terminal domain, 3' of the RxVF domain is required to exclude BRCA1 G1 foci formation (Escribano-Díaz, et al., 2013).

Discovery and Function of RIF1

RIF1 was originally identified in yeast as an interactor of Rap1 (TERF2IP in humans) to maintain telomere length (Nakamura, et al., 2006), (Marcand, Gilson, & Shore, 1997) (Wotton & Shore, 1997). However, in vertebrates, RIF1 is a diverged ortholog and not part of the telomeric complex, i.e. deficiency alone does not result in telomere defects. Instead, vertebrate RIF1 affects DNA metabolism (Buonomo S. , 2010), (Yamazaki, Hayano, & Masai, 2013). It is recruited to DNA-damage foci by irradiation (Silverman, Takai, Buonomo, Eisenhaber, & De lange, 2004), replication stress (Buonomo, Wu, Ferguson, & de Lange, 2009), and telomere deprotection (Xu & Blackburn, 2004). Rif1 was also found to be important in homology directed repair (Buonomo, Wu, Ferguson, & de Lange, 2009).

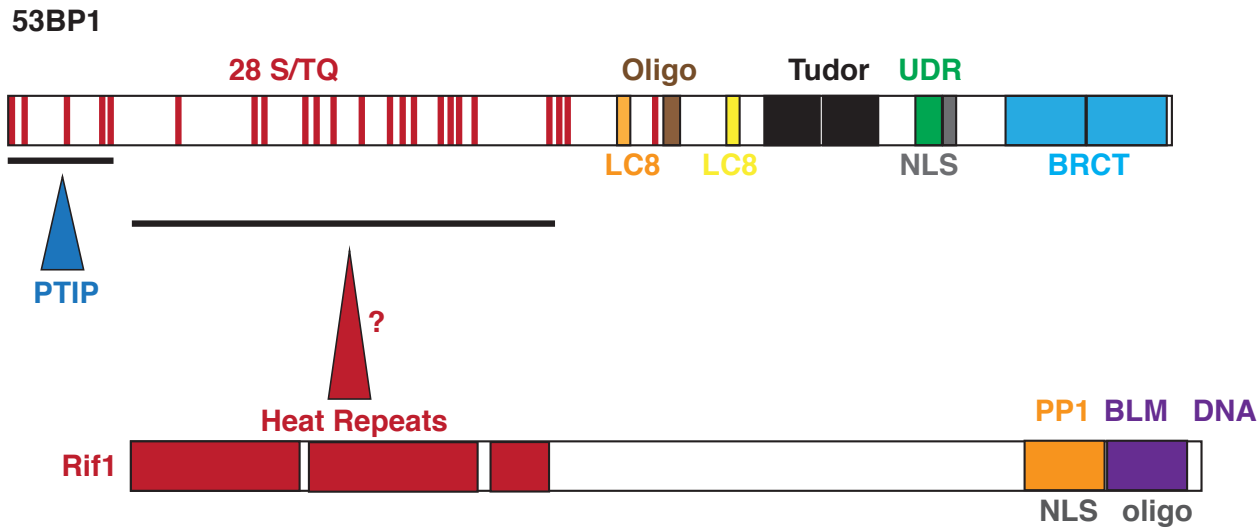


Figure 5: Schematic of RIF1 Interaction with 53BP1

N-terminal ST/Q phosphorylation sites (red bands) on 53BP1 are necessary for the interactions with PTIP and RIF1. The mechanism by which RIF1 interacts with the phosphorylated form of 53BP1 is unknown. C-terminal domains of 53BP1 are responsible for promoting binding to damaged chromatin (oligo), binding to H3K20me2 (tudor), H2A-K15Ub (UDR), and the DNA repair in heterochromatin (BRCT [BRCA1 C-terminal domain]) Note: BRCT is dispensable completely for repair in CSR and mostly not needed for repair at dysfunctional telomeres. Figure adapted from (Zimmerman, Lottersberger, Buonomo, Sfeir, & de Lange, 2013).

RIF1 and the Interaction with 53BP1

RIF1 had been found to interact with the p53-binding protein, 53BP1—a central mediator of double strand break signaling and repair, by mass spectrometry (Huen, Sy, & Chen, 2010), and was found to be dependent on 53BP1 for recruitment to DSBs measured by formation of IR induced foci (Silverman, Takai, Buonomo, Eisenhaber, & De lange, 2004). However, RIF1 emerged as the key DNA damage dependent effector of 53BP1 with the discovery that it's recruitment to DNA

damage is dependent on ATM-mediated phosphorylation of S/TQ sites on 53BP1 (Di Virgilio, et al., 2013). Like 53BP1 (Manis, et al., 2004), (Ward, et al., 2004), RIF1 deficiency resulted in a profound decrease in CSR (Escribano-Díaz, et al., 2013), (Chapman, et al., 2013), (Di Virgilio, et al., 2013). The defect in CSR resulted from increased 5' end resection (Zimmerman, Lottersberger, Buonomo, Sfeir, & de Lange, 2013), (Escribano-Díaz, et al., 2013), (Feng, Fong, Wang, Wang, & Chen, 2013), (Chapman, et al., 2013), (Di Virgilio, et al., 2013). This was consistent with 53BP1's role in preventing 5' end resection at DSBs, (Bothmer, et al., 2010), (Bunting, et al., 2010), and was also mechanistically consistent because the phosphomutant form of 53BP1, $53BP1^{28A}$, could not prevent resection at DSBs (Bothmer, et al., 2011) or telomeres (Lottersberger, Bothmer, Robbiani, Nussenzweig, & de Lange, 2013).

Genome-wide, RIF1 functions epistatically to 53BP1. These include: decreased cell survival (Silverman, Takai, Buonomo, Eisenhaber, & De lange, 2004), (Feng, Fong, Wang, Wang, & Chen, 2013), (Chapman, et al., 2013); impaired DSB repair, as measured by the increased persistence of γH2AX foci (Escribano-Díaz, et al., 2013), (Chapman, et al., 2013); increased 5'→3' end resection, as measured by the accumulation of RPA a ssDNA binding protein essential for HDR, (Zimmerman, Lottersberger, Buonomo, Sfeir, & de Lange, 2013), (Escribano-Díaz, et al., 2013), (Feng, Fong, Wang, Wang, & Chen, 2013), (Chapman, et al., 2013), (Di Virgilio, et al., 2013), and also the accumulation of RAD51, a downstream effector of RPA (Di Virgilio, et al., 2013). MEF knockouts of 53BP1, RIF1, and both factors showed the

same amount of increased 5' end resection (Zimmerman, Lottersberger, Buonomo, Sfeir, & de Lange, 2013). Additionally, phosphorylation at Ser345 of ChK1, a marker of RPA-ssDNA on 5' recessed DNA, was equal between all three genetic backgrounds (Chapman, et al., 2013).

RIF1 differences from 53BP1

RIF1 recruitment to sites of DNA damage is more phosphorylation dependent than 53BP1 (Silverman, Takai, Buonomo, Eisenhaber, & De lange, 2004), (Escribano-Díaz, et al., 2013), (Chapman, et al., 2013). This is presumably due to 53BP1 being recruited by histone marks as well as γH2AX (Kleiner, Verma, Molloy, Chait, & Kapoor, 2015), while RIF1 is solely recruited by the phosphodependent interaction with 53BP1. Dysfunctional telomeres were also found to join at a higher rate in *Rif1*$^{-/-}$ cells (Zimmerman, Lottersberger, Buonomo, Sfeir, & de Lange, 2013). In terms of cell survival, RIF1 was found to have a slightly reduced amount of cell death (Chapman, et al., 2013), and in one study was only found to affect cell survival in DT40 cells when they DNA damage was induced in G1 phase of the cell cycle (Escribano-Díaz, et al., 2013).

While *53BP1*$^{-/-}$ mice were found to be viable but radiosensitive and immunodeficient (Ward, Minn, van Deursen, & Chen, 2003), (Morales, et al., 2003), *Rif1*$^{-/-}$ mice were embryonic lethal (Buonomo, Wu, Ferguson, & de Lange, 2009). Interestingly, when such mice were outbred, they were viable, but at 50%

of the expected mendelian ratio (Chapman, et al., 2013). RIF1 also displayed phenotypes that were completely separate from 53BP1, specifically increased sensitivity to replication stress and alteration in replication timing domains (Buonomo, Wu, Ferguson, & de Lange, 2009), (Cornacchia, et al., 2012), (Yamazaki, et al., 2012), (Xu, et al., 2010).

Remaining Questions about RIF1

Critical questions in the 53BP1-RIF1 DSB repair axis, are (1) how does RIF1 protect DNA DSBs from 5' end resection, and (2) how does it interact with 53BP1. Recently, studies have shown that mouse RIF1 was found to engage DNA cruciform structures (Sukackaite, et al., 2014); bind ssDNA in vitro (Xu, et al., 2010); and yeast RIF1 encases DNA ends (Mattarocci, et al., 2017) to promote NHEJ and prevent telomere fusion. This is potentially relevant for vertebrate RIF1, because RIF1 was found to mediate this function in the absence of the Rap-1 interacting domain. However, neither of those studies explain the interaction between 53BP1 and RIF1. Specifically, because the interaction with 53BP1 is phosphodependent, it would be expected to find DNA damage responsive, phospho-interacting domains, like BRCT (Leung & Glover, 2011), in the RIF1 protein. This could point to a separate interactor mediating the phosphodependent association of RIF1 and 53BP1.

1.7 ZMYND8

ZMYND8 is the novel effector of CSR that was uncovered in this study. It has previously been implicated in the DDR and suppression of transcription at DSBs. In this section I will cover the background of ZMYND8 as a protein and it's known functions.

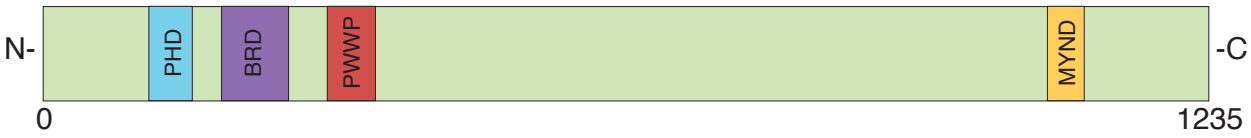


Figure 6: Schematic of ZMYND8 protein from *M. musculus*

Diagram of Zmynd8 protein sequence highlighting N and C termini, total length, and relevant domains (PHD Plant homeobox domain; BRD Bromodomain; PWWP pro-trp-trp-pro; MYND myeloid-Nervy-DEAF-1). Domains are to scale with respect to the length of the protein.

Zmynd8 (Zinc finger, MYND-type containing 8) protein [also called: RACK7, PRKCBP1, and in very few cases SPIKAR] is encoded on mouse chromosome 2: 165784155-165899016 (mm10 build), negative strand. Multiple isoforms potentially exist, ranging from 1094 aa to 1235 aa. For this study, any reconstituted proteins used the 1235 aa isoform. The APPRIS principle isoform is 1199aa.

Structure of ZMYND8 and Functional Domains

Relevant domains from N to C termini of ZMYND8 include:

PHD domain, located from 112-157 aa. Plant homeobox domain (PHD) finger is a methyl reader domain. This PHD domain is a zinc finger—C_3HC_4 conserved motif coordinates two Zinc ions (Sanchez & Zhou, 2011). Various PHD domain families bind H3K4me3 or alternatively the unmethylated form H3K4me0.

BRD domain, located from 189-259 aa. Bromodomain is an acetylation-binding module (Brand, et al., 2014). It targets acetylated lysines on histones. A conserved arginine (N252) residue is crucial for Zmynd8-BRD function (Savitsky, et al., 2016).

PWWP domain, located from 301-351 aa. PWWP for (pro-trp-trp-pro) modules contain a conserved aromatic cage for histone methyl-lysine recognition. They synergistically bind histones and DNA. Zmynd8 has a conserved phenylalanine and tryptophan residues that form this aromatic cage (Savitsky, et al., 2016).

MYND domain, located from 1064-1098 aa. MYND (myeloid-Nervy-DEAF-1) is a C_6HC zinc finger binding motif found in other nuclear proteins but does not bind DNA (Gross & McGinnis, 1996). MYND domains are putative protein-protein interaction domains in mammalian species (Liu, et al., 2007).

Only 5 proteins in mice contain all three N-terminal domains (PHD-BRD-PWWP): Zmynd11, BRPF1, BRPF3, BRD1, and ZMYND8 (Onodera, et al., 2012).

Discovery of ZMYND8

Zmynd8 was discovered as a binding partner of RACK1 (receptors for activated C-kinase) and originally designated PRKCBP1 (Fossey, et al., 2000). Soon after it was identified as a cutaneous T-cell lymphoma associated antigen (Eichmuller, et al., 2001). It was later defined as binding partner of FHOD1 using a yeast-two-hybrid screen (Westendorf & Koka, 2004).

ZMYND8 and Transcription

Insight into Zmynd8's function came with the discovery that it interacts with RCOR2, a REST corepressor (Zeng, Kong, Li, & Mao, 2010). Localization via a fused GAL4 (DNA binding domain) resulted in gene repression. Continuing with the identification as a transcriptional silencer, knockdown of Zmynd8 was able to reactive a silent GFP reporter in HeLa cells (Poleshko, et al., 2010). Multiple complex partners were then identified including a transcriptional complex containing ZNF687 and ZNF592 [also a hit in the I-DIRT in this study]. Interestingly, Zmynd8 also bound to the POLR2A subunit of the RNA Polymerase II complex (Malovannaya, et al., 2011). Although Zmynd8 is often referred to as a transcriptional repressor, it is clear that it can upregulate and downregulate certain genes, as (Li, et al., 2016) showed in the DU145 prostate cancer line (2,629 genes

upregulated, 2,551 downregulated). Knockout in a breast cancer cell line revealed that not only were nearby genes upregulated, but enhancer RNAs (eRNA) were as well (Shen, et al., 2016).

ZMYND8 and Cancer

Zmynd8 was more expressed in high-grade cervical intraepithelial neoplasia and squamous cell carcinoma (Bierkens, et al., 2013). Although a highly referenced paper for Zmynd8 entries online, its function as a driver in an acute myeloid leukemia (AML) case is more likely due to the constitutive expression of the RELA fused protein and the consequent activation of NFkB (Panagopoulos, et al., 2013). However, this is noteworthy because it highlights that *Zmynd8* is constitutively expressed. In fact every tissue expresses *Zmynd8* (Fossey, et al., 2000), (Zmynd8, n.d.), (Gene: Zmynd8, n.d.). *Zmynd8* was found to be upregulated in certain cancer cell lines (DU145) and xenotransplant model in zebrafish. According to one group, knockdown of *Zmynd8* reduced *Vegfa* transcription and inhibited cancer angiogenesis and progression (Kuroyanagi, et al., 2014). However, in another group, the same cancer cell line and xenotransplant model found that Zmynd8 knockdown increases invasiveness and metastasis (Li, et al., 2016). Clinically, Zmynd8 was found to a prognostic marker for shorter survival of breast cancer patients (Yu, et al., 2017), (Chen, et al., 2017).

ZMYND8 Genomic Distribution

ZMYND8 was shown to primarily occupy promoter regions (Li, et al., 2016), (Spruijt, et al., 2016); transcriptional start sites (Savitsky, et al., 2016); and (super)enhancers (Spruijt, et al., 2016), (Savitsky, et al., 2016).

ZMYND8 and the Interaction with Epigenetic Marks

In accordance with the presence of PHD-BRD-PWWP N-terminal domains, ZMYND8 interacts with epigenetic marks: H3.1K36me2, H4K16Ac (Adhikary, et al., 2016); the PHD-Bromo domain reads H3K4me0/1-H3K14ac (Li, et al., 2016); enhancers marked by H3K4me1, H3K27ac, DNase1, no HeK4me3 (Spruijt, et al., 2016); actively transcribed genes marked by H3K4me3 and H3K9, K14 (at TSSs) (Savitsky, et al., 2016); K27ac (Spruijt, et al., 2016); H3K4me1 and H3K27ac at superenhancers and not H3K4me3 and H3K27ac (Shen, et al., 2016). These are just a subset of the reported epigenetic correlations. Zmynd8 clearly interacts with modified histones, but the exact combinatorial epigenetics remains to be elucidated.

ZMYND8's role in DNA Damage

The Miller group showed Zmynd8 functions represses transcription in response to DNA damage (Gong, et al., 2015). Zmynd8 was shown to have H4-acetylation dependent recruitment to sites of DNA damage by laser micro-irradiation. It also confirmed the association of TRIM24 (Poleshko, et al., 2010) as a DNA damage

dependent recruitment factor with Zmynd8. Interestingly, Zmynd8 knockdown prevented S345p phosphorylation of Chk1. It should be noted that while S345p is reported in some cases to be associated with ionizing radiation (IR) from X-ray irradiation or ^{137}Ce γ-irradiation, it is primarily elicited by hydroxyurea (replication stress) and UV irradiation (primarily NER, nucleotide excision repair) (Liu Q. , 2001). Consistent with this difference, S345P is mediated by ATR (Jossé, et al., 2014) and not ATM (Hickson, et al., 2005). Zmynd8 was shown to function in HDR because it is required for repair in the DR-GFP assay (Pierce, Johnson, Thompson, & Jasin, 1999), and it was required for Rad51 recruitment to sites of endonuclease DSB damage (Iacovoni, et al., 2010). Furthermore, Zmynd8 was only required for recruitment of the HDR factor Rad51 at endonuclease sites in actively transcribed genes, while it was dispensable for the recruitment of NHEJ factor XRCC4, which correspond with inactive genes (Aymard, et al., 2014). Zmynd8 was further implicated in transcription-coupled DNA repair because treatment with the RNA Pol II inhibitor, DRB (5,6-dichlorobenzimidazole riboside), abolished the recruitment to sites of laser micro-irradiation. DRB inhibits phosphorylation of the CTD (C-terminal domain) of Pol II (Dubois, Nguyen, Bellier, & Bensaude, 1994), which prevents new Pol II from moving from the initiation phase to the paused (S5-P) and elongation (S2-P) of transcription. Zmynd8 was later shown to immunoprecipitate only with S5-P Pol II, and not S2-P Pol II (Adhikary, et al., 2016).

Zmynd8 was recruited to sites of endonucleases FokI, IsceI (Xia, et al., 2017) and AsiSI DSBs, (Gong, et al., 2015), but not oxidative damage (Xia, et al., 2017). Zmynd8 recruitment to DSBs in actively transcribing genes resulted in demethylation of H3K4me3 (Gong, Clouaire, Aguirrebengoa, Legube, & Miller, 2017). Interestingly, recruitment to DSBs by ZFN endonucleases was mediated by BRD2 and was excluded from the actual DSB site (+/- 0.5kb) and restricted to flanking regions (Gursoy-Yuzugullu, Carman, & Price, 2017).

Remaining Questions about ZMYND8

A model has emerged where Zmynd8 is recruited to sites of DNA damage in transcriptionally active regions and suppresses transcription by recruiting various complexes (NuRD, Co-Rest, and Integrator) (Savitsky, et al., 2016). While Zmynd8$^{-/-}$ mice are embryonic lethal (Zmynd8, n.d.), deletion in various cancer cell lines seem to have mild effects, mostly increased growth and invasiveness (Li N. , 2016), (Shen, et al., 2016). This is somewhat contradictory with clinical data as higher Zmynd8 expression correlates with poor outcomes (Yu, 2017), (Chen, 2017). A fundamental question is what are the cellular consequences of Zmynd8 knockdown and how do those translate to the organism at large? Specifically, what is the link between repressed transcription at transcriptionally active sites post break and cell function?

1.8 The *IgH* 3' Regulatory Region (3'RR): The Prototypical Super-Enhancer

The 3' RR is a ~28 kb region downstream of the Cα region, the most 3' of the *IgH* constant genes in mice. It is comprised of 4 enhancers, HS3a, HS1,2, HS3b, and HS4a (in order extending 3' from the Cα region). The entire region is required for CSR (Vincent-Fabert, et al., 2010) and SHM (Rouaud, et al., 2013).

Discovery of the 3' RR

The discovery of the 3'RR was preceded and prompted by characterization of the 5' enhancer of the *IgH* constant region (E_μ). This enhancer is located between the 3' J_H segment and the 5' C_H region (Banerji, Olson, & Schaffner, 1983), (Gillies, Morrison, Oi, & Tonegawa, 1983), (Neuberger, 1983). E_μ caused *IgH* genes to be expressed in B cells when transfected. However, there was evidence of a second control element for B cells. First, B cell lines that had deleted E_μ could still express *IgH* (Eckhardt & Birshtein, 1985), (Zaller & Eckhardt, 1985). This transcriptional redundancy could account for the expression of c-myc in c-myc:IgH translocations in Burkitt lymphoma (human) and mouse and rat cancers with orthologous translocations (Pettersson, Cook, Bruggemann, Wiliams, & Neuberger, 1990). The requirement in secondary diversification reactions was confirmed in transgenic models where deletion of E_μ caused V(D)J recombination reduction and reduced B cell numbers, but B cells that were produced could undergo CSR and SHM

(Perlot, Alt, Bassing, Suh, & Pinaud, 2005). The 3' RR was identified first in mice as HS 1,2 for two DNase I hypersensitive sites (Lieberson, Giannini, Birshtein, & Eckhardt, 1991). Later other DNase I hypersensitive sites 3' of Cα were identified in mouse and humans and are reviewed in (Birshtein, 2014), (Pinaud, et al., 2011).

Function of the 3'RR

Early clues to the 3'RR function were found in low-producing plasmacytomas, where the entire 3' RR was spontaneously deleted (Michaelson, Giannini, & Birshtein, 1995), (Gregor & Morrison, 1986). Later, endogenous deletions would show reduced IgM in plasma cells, and that the 3'RR was required for CSR (Vincent-Fabert, et al., 2010) and SHM (Rouaud, et al., 2013) but did not impair V(D)J recombination (Rouaud, et al., 2012). Partial deletions of the IgH 3'RR give heterogeneous results with respect to CSR. For example, deletion of Hs3b and HS4 together eliminated class switching for everything except IgG1 (Pinaud, et al., 2001). Other combinations of enhancer deletions or sequence inversions give different results for CSR by isotype and SHM.

Chromatin Looping via the 3'RR

The mechanism of action for the 3'RR is still being dissected. However, interaction of the 3' RR with the *IgH* locus is required for function. This 3D interaction can be measured via chromosome conformation capture (3C) and expanded technologies. Briefly, in 3C, cross-linking fixes chromatin in 3D space. The DNA is

then digested and re-ligated in that 3D context. As a result, the anchor points of a DNA loop may be re-ligated to each other. Primer pairs and PCR then amplify those alternate linkages for detection. 3C found that in resting B cells loops form between 3'RR and VDJ- E_{μ} but this loop does not exist in T cells (Wuerffel, et al., 2007). Additionally, upon activation for CSR, the loop also includes a contact point with the activated intronic enhancer. Furthermore, the degree of looping was proportional to the amount of class switch derived from that enhancer (i.e. in LPS + IL-4 stimulation, $I_{\gamma 1}$ had the greatest looping, but in LPS only stimulation, $I_{\gamma 3}$ was greater).

Histone Modifications due to 3'RR

The 3'RR and histone modifications have been implicated in targeting AID and RNA pol II to s regions in CSR. H3K4me3 decorates the S_{μ} region constitutively, but appears on downstream S regions after activation (Kuang, Luo, & Scharff, 2009). Deletion of the FACT complex in CH12 cells reduced H3K4me3 of switch regions but did not reduce AID or GLT expression (Stanlie, Aida, Muramatsu, Honjo, & Begum, 2010). Deletion of the Methyl transferase complex (Mixed-lineage leukemia 3 (MLL3/KMT2C) – MLL4/KMT2D set1-like lysine methyltransferase complex) reduced CSR, slightly by 1.6 times in IgG_1, and 2-fold in IgG_3 but did not decrease GLT of $S_{\gamma 1}$ or $S_{\gamma 3}$ (Starnes, et al., 2016). In humans, Kabuki syndrome results from KMT2D mutations and presents in part as immune deficiency due to hypogammaglobulinemia (Stagi, Gulino, Lapi, & Rigante, 2015).

While the former modifications are dependent on activation for CSR but independent of AID expression, H3ac appears to be AID dependent (Wang, Whang, Wuerffel, & Kenter, 2006). Like H3K4me3, H3K9ac and H3K14ac (together H3Kac) constitutively decorates the S_{μ} region, presumably due to constant expression of GLT S_{μ}, but unlike H3K4me3, H3Kac also increases as a result of activation (Wang, Wuerffel, Feldman, Khamlichi, & Kenter, 2009). However, at downstream acceptor regions H3Kac is deposition correlates with H3K4me3, that is to say it correlates with GLT specific transcription and is low or undetectable prior to activation and GLT transcription. Interestingly, artificially increasing H3Kac with histone deacetylase inhibitors (HDACs) led to higher CSR independent of any dysregulation in AID (Wang, Wuerffel, Feldman, Khamlichi, & Kenter, 2009).

Connecting back to the 3'RR, Saintamand and colleagues found that in 3'RR deficient mice the histone modifications due to increased GLT and activation were reduced at acceptor regions. Specifically, H3K4me3 is reduced at S acceptor regions including, $S_{\gamma 1}$, $S_{\gamma 3}$, $S_{\gamma 2b}$, and S_{ε} (Saintamand, et al., 2015). H3K9ac, Pol II, and AID were also reduced at these regions. However, the donor region S_{μ} did not show a reduction in H3K4m3, and the reduction in H3K9ac was minimally reduced and at the 3' end of the S region.

Overall, the 3'RR is poorly conserved between species and aside from the well-defined function to promote CSR and SHM, the contribution and mechanism of individual enhancers and intervening segments is still an open question.

Summary and Outlook

Class Switch Recombination (CSR) is a DNA recombination reaction that diversifies the effector component of antibody responses. CSR is initiated by activation-induced cytidine deaminase (AID), which produces nucleotide mismatches in transcriptionally active immunoglobulin heavy chain (*Igh*) switch donor and acceptor DNA. The process of CSR requires a functional DNA damage response that facilitates ligation of the upstream donor and downstream acceptor regions via NHEJ. 53BP1 and its effector RIF1 are critical to this process because they inhibit end resection to promote NHEJ. However, the mechanism by which RIF1 effects end-protection in CSR and binds to 53BP1 are still unknown.

In Chapter 2 I will describe experiments where we use proteomics to identify the RIF1 interactome in actively switching B cells. In Chapter 3 I utilize the emerging CRISPR-Cas9 system to orthogonally screen candidates from the proteomics screen for functionality in CSR in the CH12 B cell lymphoma line. In Chapter 4 I assess the newly identified candidate in DNA repair and CSR, ZMYND8, for effects

on CSR prior to DSB break formation. In Chapter 5, I assess ZMYND8 for effects in the DNA DSB repair process. In Chapter 6, I investigated the interaction between ZMYND8 and transcription. In Chapter 7, I repeated the investigations from Chapters 4-6 in a new conditional knockout mouse.

Chapter 2: RIF1 Interactome in Activated B Cells

2.1 Validation of the *Rif1*$^{FH/FH}$ mouse for I-DIRT

We utilized an I-DIRT (*I*sotopic *D*ifferentiation of *I*nteractions as *R*andom or *T*argeted) (Tacket, 2005) to investigate the RIF1 protein interactome in actively switching primary B-cells. To increase the efficiency of RIF1 pull-down we utilized a FLAG-tag (FT) fused RIF1 recombinant mouse strain. Before performing the actual I-DIRT we needed to confirm that the knock in protein and cells behaved similarly with respect to RIF1$^{+/+}$ B-cells. To demonstrate this, we isolated splenocytes from *Rif1*$^{FH/FH}$ mice (Cornacchia et al., 2012). These mice express a knock-in RIF1 fused with 1xFLAG-2xHA N-terminal tags. The Rif1$^{FH/FH}$ mice express RIF1 at physiologic levels in B-cells as demonstrated by Western blot (Figure 2.1a). Additionally, the B-cells proliferate at comparable rates to Rif1$^{+/+}$ B-cells when measured by CFSE staining (Figure 2.1b). Importantly, the Rif1$^{FH/FH}$ B-cells undergo CSR at similar rates to Rif1$^{+/+}$ B-cells in ex vivo stimulating conditions (Figure 2.1c LEFT representative FC graph, and RIGHT summary CSR data).

Note: figures below are adapted to/from the final publication: (Delgado-Benito, et al., 2018)

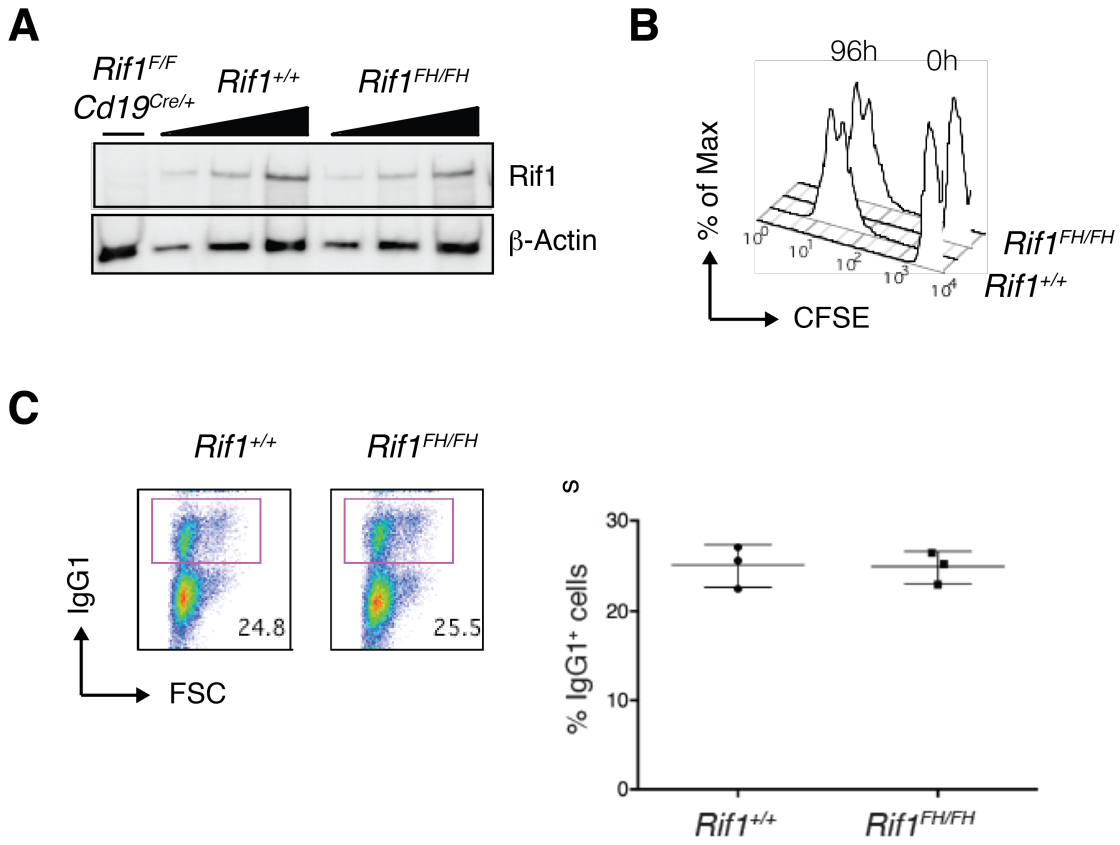


Figure 7: Rif1$^{FH/FH}$ B-cells express physiological levels of RIF1 protein and support wild-type CSR levels.

(A) Western blot analysis of cell extracts from *Rif1*$^{+/+}$ (WT) and *Rif1*$^{FH/FH}$ B-cells. Triangles at the top of the graph indicate three-fold dilution. Negative control at the left is *Rif*$^{f/f}$ (flox) *Cd19*$^{cre/+}$ mice loaded at dilution equivalent to highest *Rif1*$^{+/+}$ (WT) and *Rif1*$^{FH/FH}$ B-cells.

(B) Flow cytometry histograms measuring B-cell proliferation by CFSE dye dilution 96 h post stimulation with LPS, IL-4, and RP105.

(C) Left: Representative flow cytometry plots measure CSR to IgG1 under the same conditions. Right: summary of three independent experiments (n = three mice per genotype).

2.2 Identification of RIF1 Interacting Proteins in Primary B-Cells

The I-DIRT approach in actively switching B cells was used to capture the RIF1 interactome with respect to CSR. We also irradiated the cells prior to collection to enrich for Rif1 interactions in the DDR. Figure 2.2a shows a flow chart of the experiment. The solubilized protein complexes were immunoprecipitated with FT specific antibodies conjugated to magnetic beads. The complexes were natively eluted with 3x FT peptides. The samples were digested with trypsin and run with LC-MS/MS (liquid chromatography-tandem mass spectrometry) and the results were analyzed with MaxQuant proteomics software (Cox & Mann, 2008).

We determined that a 96 h growth period in activating SILAC media was sufficient for the majority incorporation of the heavy-labeled amino acids (Figure 2.2Bi, ii, iii and Figure 2.2C). Each heavy and light peptide pair identified showed the 6 molecular weight increase and possessed an arginine or lysine at the C-terminus. This confirms the incorporation of a single heavy-labeled arginine or lysine per trypsin-digested peptide.

A full scale I-DIRT detected thousands of proteins (Figure 2.2D). The proteins were sorted according to their abundance/enrichment ratio as determined by the proportion of heavy peptides to total peptides (heavy and light) detected. The most enriched and also most abundant protein detected was RIF1. The next protein was

the phospho-dependent interactor of Rif1 (53BP1). Most of the proteins detected displayed an abundance ratio characteristic of non-specific pull down (average of 0.49 ± 0.10 standard deviation). These points confirmed, respectively, that (a) Rif1-FH was efficiently immunoprecipitated, (b) growth and immunoprecipitation conditions were sufficient to identify 53BP1, a known DDR- dependent physiologic interactor of Rif1, and (c) the heavy and light samples were combined at a 1:1 ratio. Highly enriched co-immunoprecipitated proteins (i.e. abundance ratios > $\mu + \sigma$) included several transcription factors (Figure 2E), as well as many proteins implicated in DDR (Chapter 3)

Figure 8: Identification of RIF1 interacting proteins in primary B cells undergoing CSR

(A) I-DIRT Schema in ex vivo cultures of splenocytes. Abbreviations: GA (glutaraldehyde); LC-MS/MS (liquid chromatography-tandem mass spectrometry). *Rif1*$^{FH/FH}$ and wild-type *Rif1*$^{+/+}$ splenocytes were isolated and cultured in SILAC media supplemented with isotopically heavy ^{13}C Arginine and Lysine (*Rif1*$^{FH/FH}$) or light (normally abundant) ^{12}C Arginine and Lysine (*Rif1*$^{+/+}$). The cells were grown their respective media and activated with LPS, IL-4 and RP105 for 96 h. Cell cultures were irradiated with 20 Gy, allowed to recover for 45 min, and flash frozen as pellets in a liquid N_2 bath. The pellets from heavy and light samples were added in equal mass amounts and cryolysed. The cryolysed powder was extracted in soluble form with sub-stoichiometric treatment with glutaraldehyde in order to stabilize labile and/or transient interactions without alternative the native composition of protein complexes (Subbotin and Chait, 2014). Protein complexes were co-immunoisolated, trypsinized and analyzed via mass spectrometry.

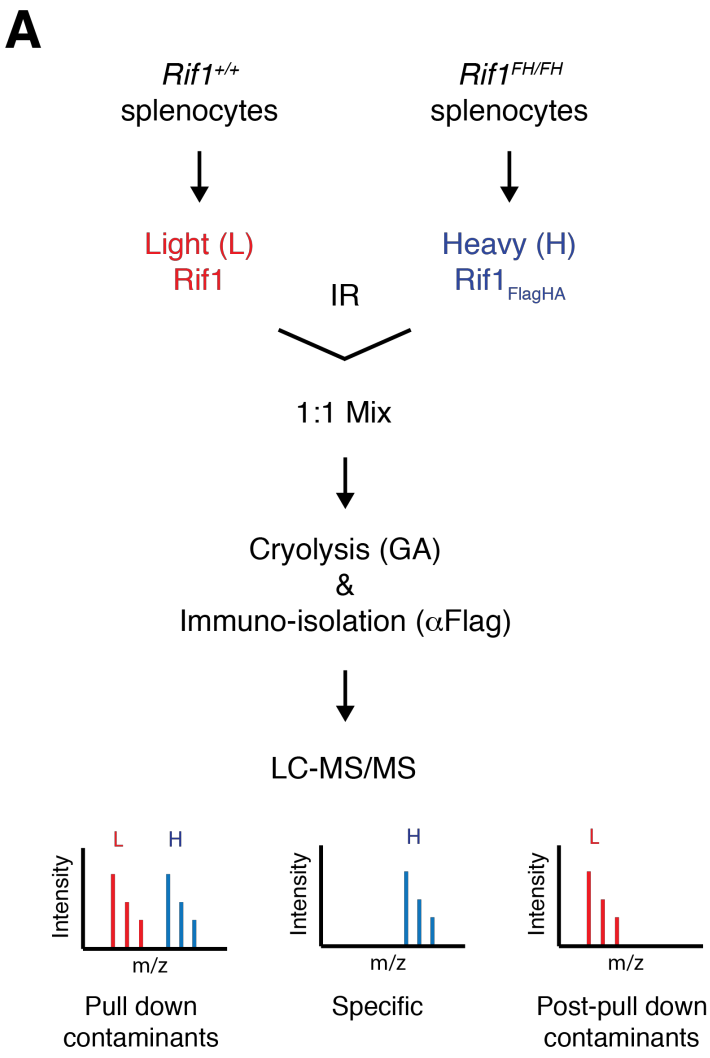
(B) Mass spectrometry graph of heavy amino acid incorporation test. Primary B cells were cultured for 96 hr in activating (LPS, IL-4, RP105) conditions in SILAC media supplemented with ^{13}C labeled arginine and lysine. Representative mass spectrometry graph of proteins(peptides) (B.i) ActG1 (VAPEEHPVLLTEAPLNPK) (B.ii) Actb (ACFPSIVGR) (B.iii) Aldoa (IGEHTPSALAIMENANVLAR)

(C) Summary of three representative peptides and L → H shift. Abbreviations: UniProtKB (UniProtein Knowledge Base entry); H M/Z (heavy molecular weight/charge of peptide); L M/Z (light molecular weight/charge of peptide); Z

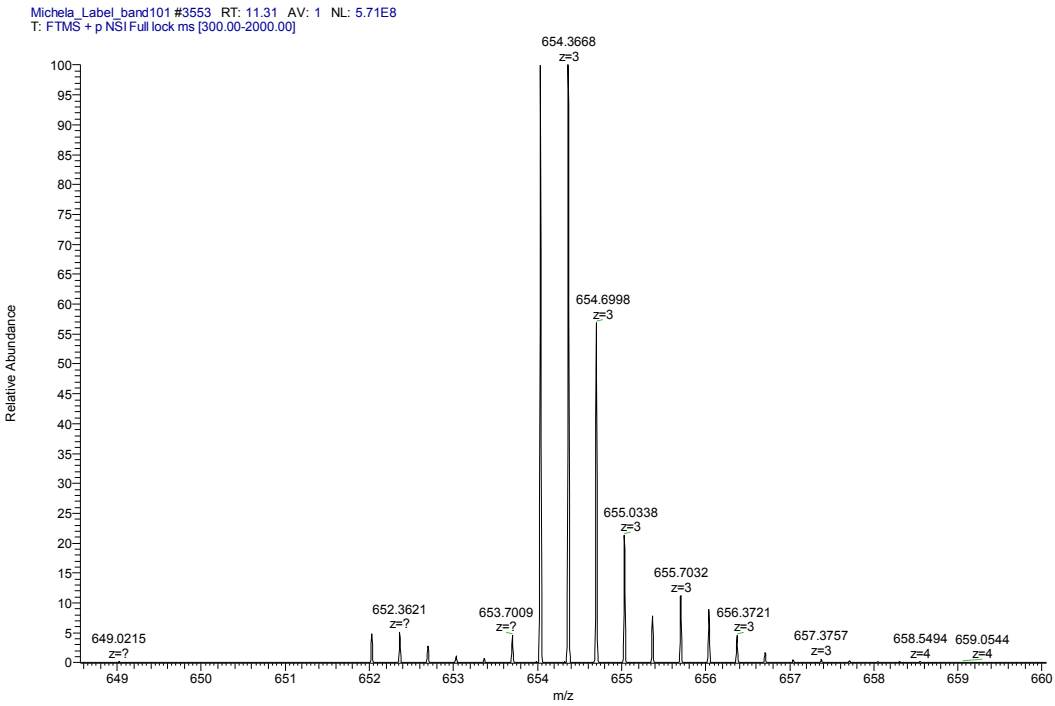
(peptide charge); (H-L)*Z (difference in molecular weight/charge ratio multiplied by charge to return the difference in MW between the heavy and light forms).

(D) Graph of proteins identified in the RIF1 I-DIRT. The proteins are arranged from highest H/(H+L) (abundance ratio) to lowest. Error bars represent the standard error of the abundance ratio mean for all peptides identified for each protein. Only proteins with peptide counts $\geq$ 4 and posterior error probabilities $\leq 10^{-4}$ are included. The central line is the mean (0.49) and the σ is the standard deviation of the distribution. Top: chart with proteins RIF1, 53BP1, and ZMYND8 identified. Bottom: chart with regions defined as hits (red), non-specific interactors (purple), and contaminants (blue) marked.

(E) Potential transcription related RIF1 interactors found among identified proteins with abundance ratios > 2 σ above the mean.

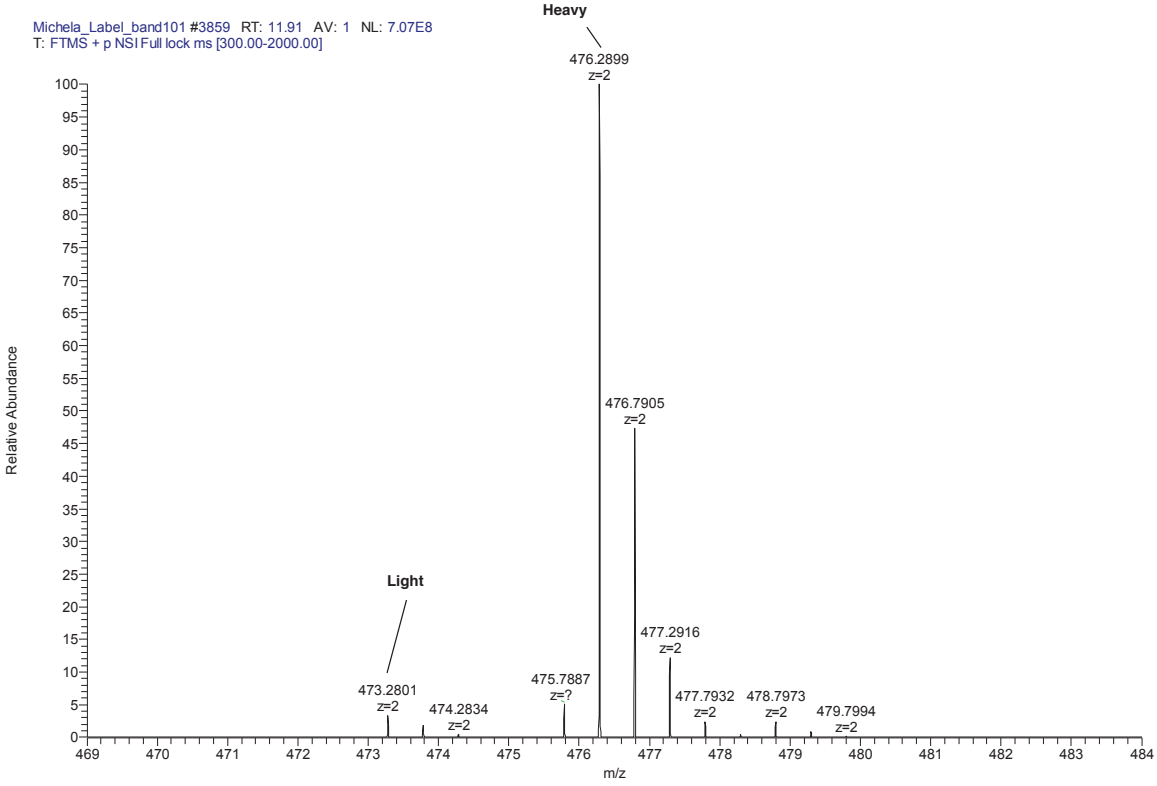

A
Rif1+/+
splenocytes
Rif1FH/FH
splenocytes
Light (L)
Rif1
Heavy (H)
Rif1FlagHA
IR
1:1 Mix
Cryolysis (GA)
&
Immuno-isolation (αFlag)
LC-MS/MS
L
H
Intensity
m/z
Pull down
contaminants
H
Intensity
m/z
Specific
L
Intensity
m/z
Post-pull down
contaminants

B.i

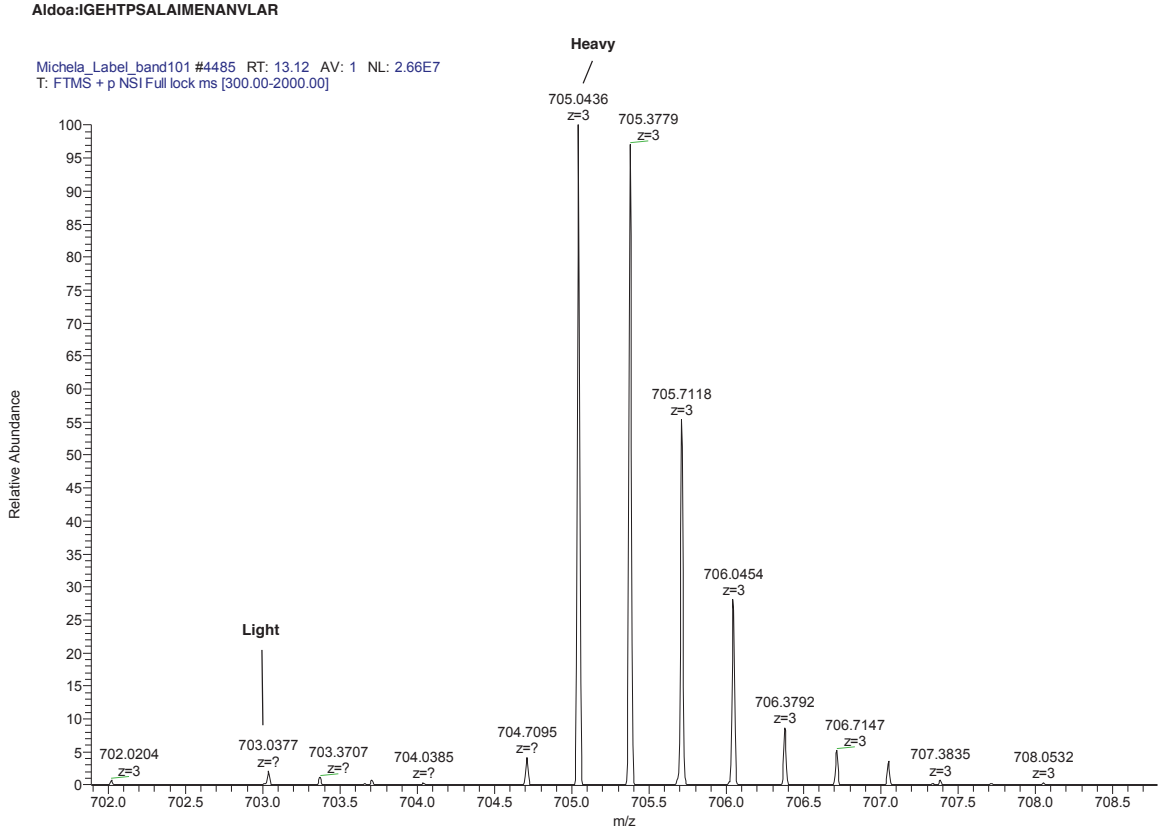


B.ii

Actb(P60710): AVFPSIVGR

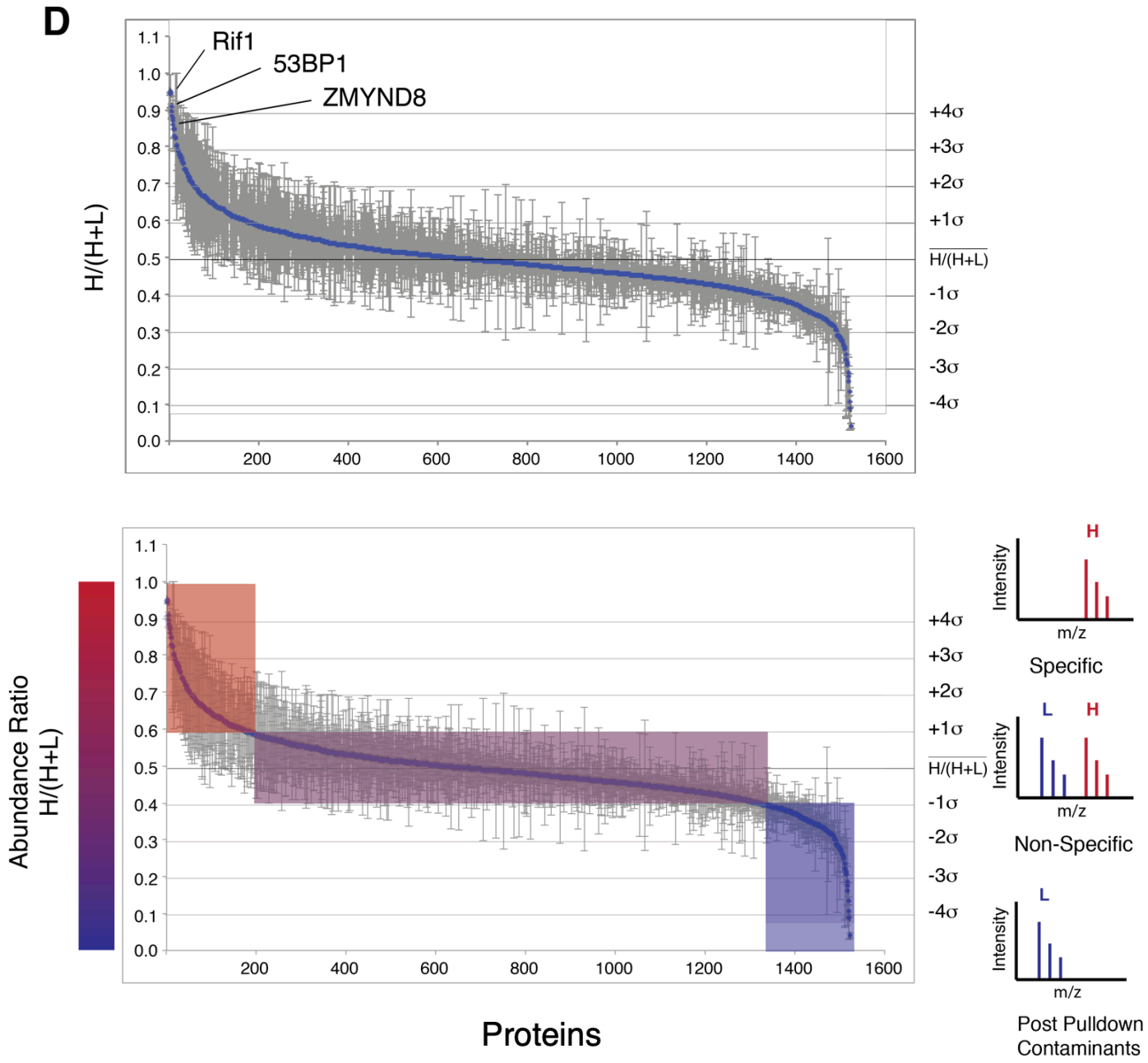


B.iii



C

Protein	UniProtKB	AA Sequence	H M/Z	L M/Z	Z	(H-L)*Z
ActG1	P63260	VAPEEHPVLLTEAPLNPK	654	652	3	6
Actb	P60710	ACFPSIVGR	476	473	2	6
Aldoa	P05064	IGEHTPSALAIMENANVLAR	705	703	3	6


D
Rif1
53BP1
ZMYND8
H/(H+L)
+4σ
+3σ
+2σ
+1σ
H/(H+L)
-1σ
-2σ
-3σ
-4σ
Abundance Ratio H/(H+L)
Proteins
Intensity
m/z
H
Specific
L
H
Non-Specific
L
Post Pulldown Contaminants

E

Candidate	Peptide Count	H/(H+L)
53BP1	72	0.91
DYNLL1	5	0.68
ZMYND8	15	0.85
MGA	57	0.88
ZNF592	6	0.85
BACH2	14	0.83
TCEB1	4	0.70

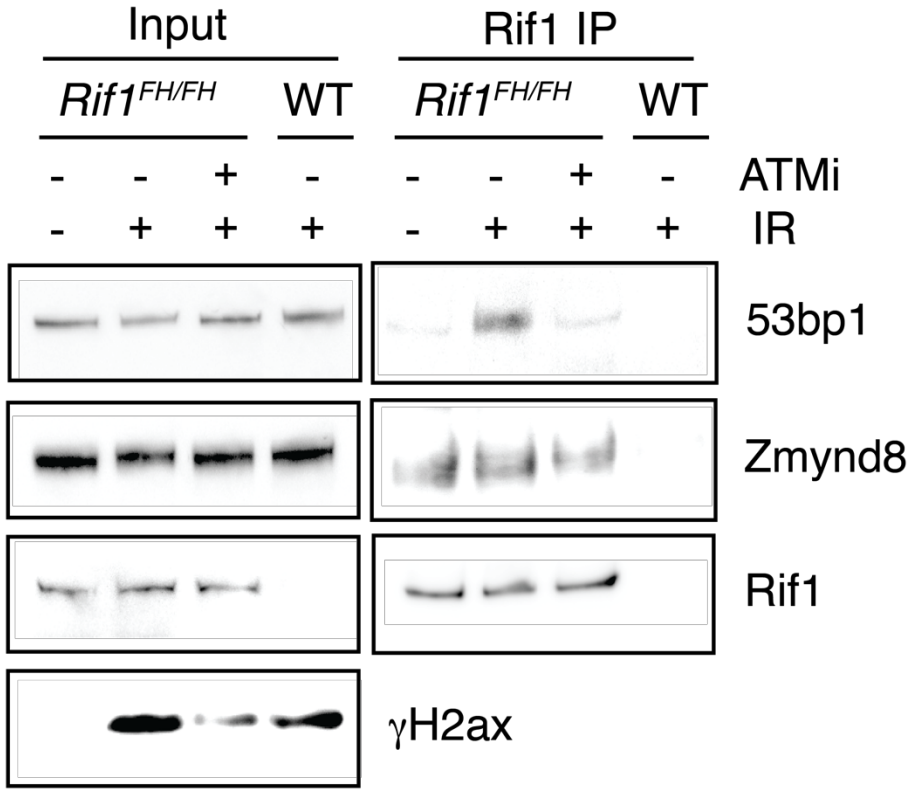
2.3 Confirmation of Zmynd8 and RIF1 Interaction

To validate the Zmynd8 association with RIF1 discovered in the I-DIRT, we performed reciprocal co-immunoprecipitation experiments in primary B cells. Zmynd8 efficiency co-immunoprecipitates with RIF1, and vice versa (figure 9a and b). RIF1 involvement in DNA damage repair is dependent on ATM phosphorylating 53BP1. The phosphorylated 53BP1 then recruits RIF1 to the site of DNA breaks (Chapman, et al., 2013), (Di Virgilio, et al., 2013), (Escribano-Díaz, et al., 2013), (Feng, Fong, Wang, Wang, & Chen, 2013), (Zimmerman, Lottersberger, Buonomo, Sfeir, & de Lange, 2013). To determine if the RIF1-Zmynd8 interaction is also DNA damage-dependent, we co-immunoprecipitated each protein in the presence or absence of IR-induced DNA damage, and the presence or absence of an ATM inhibitor. While we confirmed that the association between RIF1 and 53BP1 is both DNA damage and ATM dependent (Figure 9a) (Di Virgilio, et al., 2013), Zmynd8 co-precipitated with RIF1 and vice-versa regardless of DNA damage or ATM activity (Figure 9a, Figure 9b). Therefore, Zmynd8 interacts with RIF1 *in vivo,* independently of DNA damage.

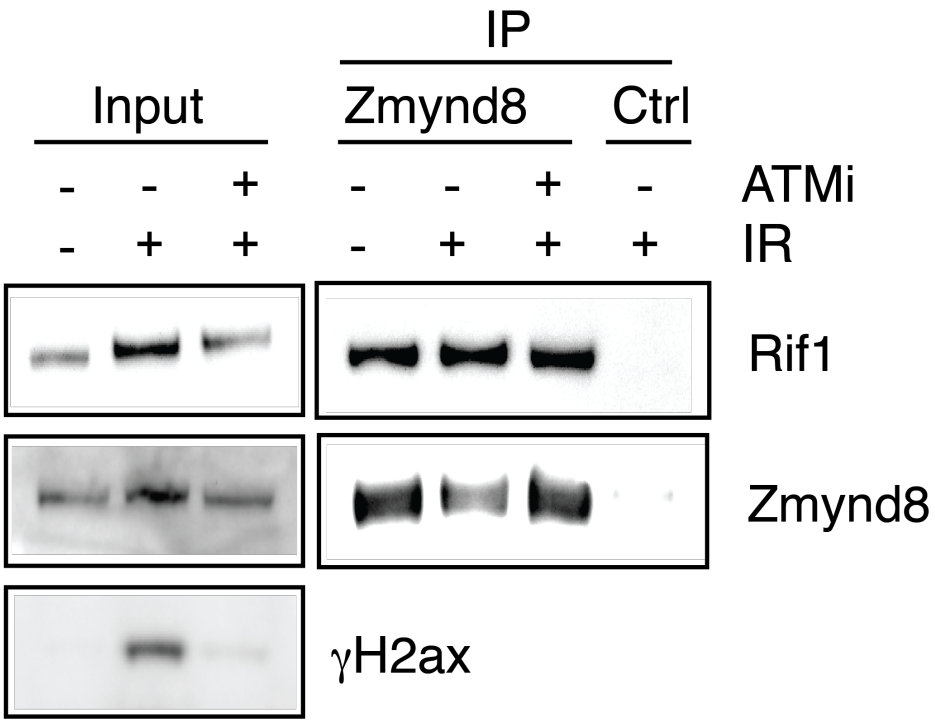
Figure 9: The interaction between ZMYND8 and RIF1 is DNA damage-independent.

(A-B) Western blot analysis of anti-FLAG(RIF1) **(A)**, and anti- ZMYND8 **(B)** immunoprecipitates from *WT* and *Rif1*$^{FH/FH}$ B lymphocytes either left untreated or irradiated (10 gray (Gy), 45-min recovery) in the presence or absence of the ATM kinase inhibitor KU55933 (ATMi). Data are representative of at least two independent experiments for each co-immunoprecipitation.

A


Input
Rif1 IP
Rif1FH/FH
WT
Rif1FH/FH
WT
- - + - - - + - ATMi
- + + + - + + + IR
53bp1
Zmynd8
Rif1
γH2ax

B


IP
Input
Zmynd8
Ctrl
- - + - - + - ATMi
- + + - + + + IR
Rif1
Zmynd8
γH2ax

Chapter 3: RIF1 Interactors and CSR

3.1 CRISPR Cas9 Screen Design for CSR in CH12 cells

To validate the top hits in the I-DIRT we designed an orthogonal screen to assess proteins for function in CSR. We utilized the newly functionalized *S. pyogenes* CRISPR-Cas9 system (Ran, et al., 2013) in order to somatically target genes. Specifically, the Cas9 nuclease cuts the gene at a site specified by the gRNA. Repair by the error-prone NHEJ pathway leads to insertions and deletions (indels) at the DSB. Indels are statistically likely to lead to frameshift mutations, which create downstream premature stop codons, and hence delete the protein. We used the CH12F3 (CH12) cell line, a B-cell lymphoma cell line derived from B10 H-2aH-4bp/Wts mice (Kunimoto, Harriman, & Strober, 1988) which robustly switches from IgM to IgA (Nakamura, et al., 1996) to measure defects in CSR caused by CRISPR-Cas9-mediated deletion. The CH12 cell line has been extensively utilized as a model system to explore CSR.

The CRISPR Cas9 screen in CH12 cells is similar to strategies from the Basu (Pefanis, et al., 2014) and Martin (Le, et al., 2016), (Ramachandran, et al., 2016) laboratories.

We transiently expressed Cas9+gRNA with the px458 plasmid (Figure 10a) (Ran, et al., 2013) via nucleofection into the CH12 cell line. Cas9 expression directly correlates with GFP expression due to a 2A linker. GFP+ cells were sorted 2 days post nucleofection and grown for 3 days in culture (Figure 10b). This time period allowed for protein depletion as existing protein turned over and was not replenished. Cells were then activated for 48 h and assessed for CSR to IgA by flow cytometry.

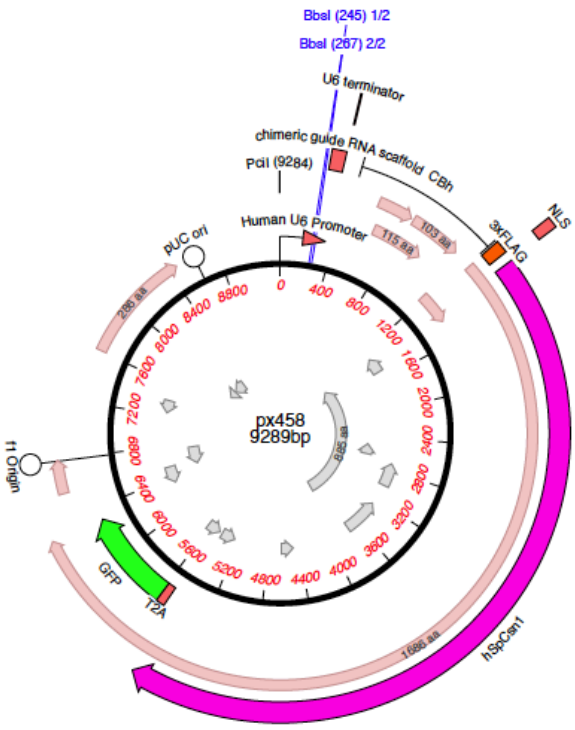
gRNAs targeting the candidate genes were designed by accessing the APPRIS principal isoform on ensembl. The first two gRNAs for each of the 5' two terminal exons were generated using the Zheng lab/MIT CRISPR design software (CRISPR Design, n.d.). Exons unique to the APPRIS principal isoform were excluded. We endeavored to create indels as close to the 5' end of the gene as possible to exclude the possibility of producing truncated but still function proteins.

Figure 10: CRISPR Cas9 Screen Design to Assess CSR in CH12 Cells

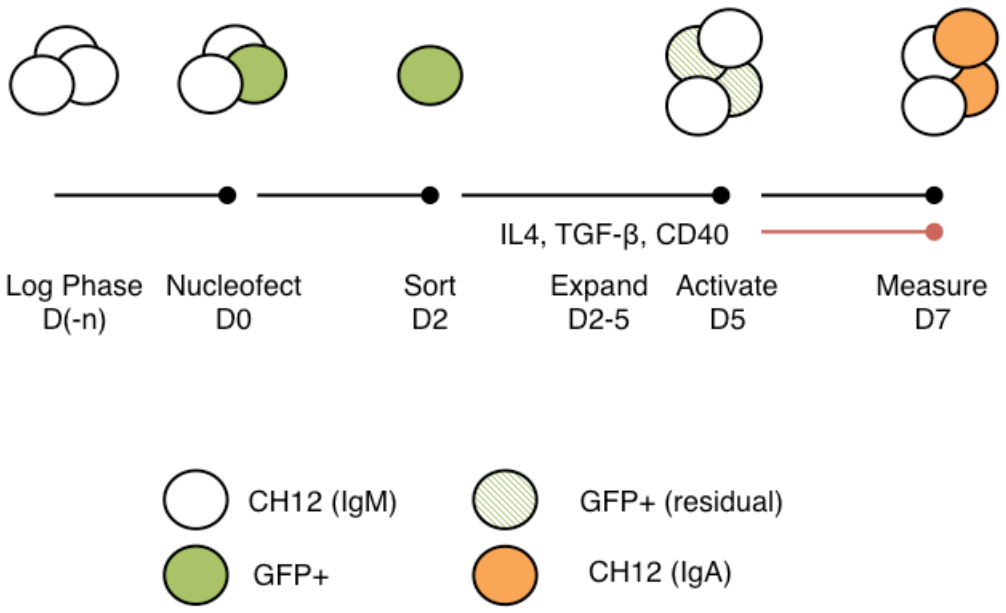
(A) Plasmid map of pSpCas9(BB)-2A-GFP (PX458) deposited to Addgene (plasmid # 48138) by Feng Zhang. A 2A linker allows linked translation of GFP and the Cas9 protein.

(B) Experimental design of CRISPR-Cas9 CH12 cell screen. Log phase, unactivated, cells are nucleofected at day 0. A percentage of cells successfully acquired the plasmid. These GFP+ cells are detected and isolated by flow assisted cytometric sorting (FACS). Cells are grown in culture for 3 days to allow recovery from sorting, and for the preexisting protein to degrade after somatic deletion of the target gene. At day 5, cells are activated for switching to IgA and measured by flow cytometry at day 7.

A


BbsI (245) 1/2
BbsI (267) 2/2
U6 terminator
chimeric guide RNA scaffold CBh
PciI (9284)
Human U6 Promoter
PUC ori
3xFLAG
NLS
f1 Origin
GFP
T2A
hSpCas1
px458
9289bp

B


IL4, TGF-β, CD40
Log Phase
D(-n)
Nucleofect
D0
Sort
D2
Expand
D2-5
Activate
D5
Measure
D7
CH12 (IgM)
GFP+
GFP+ (residual)
CH12 (IgA)

3.2 CRISPR Cas9 CSR Screen of I-DIRT Targets

The I-DIRT screen revealed many novel proteins in the B-cell RIF1 interactome. 167 proteins were identified with an abundance ratio greater than 1 standard deviation above the mean in one or more of the I-DIRT replicates (pilot experiments V1_10, V1_25, or V1_50 or full-scale experiments V2.1 or V2.2). We used the CRISPR Cas9 system to orthogonally screen candidates that when deleted showed, like RIF1, a defect in CSR.

Of the 167 statistically significant proteins, 17 (Fig 11a) were selected as initial candidates for CSR testing. The 17 proteins were filtered through literature searches identifying these candidates as possible components of the DDR, CSR, and/or pathways also attributed to RIF1. Some candidates were selected for further study if it had been identified in other screens, specifically proteomic or CSR based (i.e. ZRANB2 from (Pavri, et al., 2010)).

For each candidate, 4 gRNAs for the first two earliest possible exons were designed according to the criteria in Chapter 3.1. An example schematic for Zmynd8 targeting is shown in figure 11b.

Each experiment included an empty (non-targeted) px458 vector as a negative control and px458 with gRNAs against AID as a positive control. Example flow

cytometry results can be seen in Figure 11b. For targets that showed a reduction in CSR, Western Blot analyses were performed. An example for Zmynd8 is in Figure 11c.

Figure 11d shows CSR normalized to empty vector for all the 17 candidates. Of those tested, only Zmynd8 showed consistently down-regulated CSR. However, not all gRNAs effectively deleted the target protein (as can be seen for Dynll1). The positive hit, Zmynd8, was retested in 3 biologically independent repeats to confirm a role in CSR (figure 11e). Bulk testing found that two different gRNAs against different exons in Zmynd8 showed a 40-50% reduction in CSR.

It is noteworthy that gRNA against AID incompletely abrogated CSR (figure 11b and all experiments as a positive CSR reduction control). This could either indicate a heterogeneous population of nucleofected cells, i.e. not all cells that receive the px458 plasmid successfully targeted the Cas9+gRNA complex, or it could represent that indels created by gRNA targeting lead to in-frame deletions which do not destroy the protein function. The two hypotheses can be distinguished by isolating clones from a bulk population and screening for (un)successful protein deletion and corresponding indels.

Figure 11: CRISPR Cas9 CSR Screen of I-DIRT Targets

Table 1: I-DIRT statistically significant targets.

(A) Curated list of I-DIRT statistically significant targets. Protein listed in first column, UniProtID identifier listed in second column, and the z-score averaged across all I-DIRT experiments in third column. Proteins below in the offset are positive controls, specifically RIF1, the bait in the I-DIRT and 53BP1, a phosphor-dependent interactor which was used to identify RIF1. * indicates hits from other screens

(B) Scheme of *Zmynd8* genomic locus and location of gRNAs used in this study (scheme adapted from Ensembl ZMYND8-001 ENSMUST00000109269.7).

(C) CRISPR-Cas9-mediated targeting of Zmynd8 in bulk CH12 cell population reduces CSR. Representative flow cytometry plots measuring CSR to IgA in Cas9/gZmynd8-nucleofected CH12 cells. CRTL is cells nucleofected with empty px458 plasmid. Cells were activated for 48h with αCD40, TGFβ and IL-4. Numbers indicate the percentage of IgA switched cells.

(D) Western blot analysis of whole cell extracts from CH12 cultures following nucleofection with gRNAs against Zmynd8. Triangles indicate threefold dilution

(E) CSR Reduction Summary of I-DIRT candidates. Each point represents a separate gRNA target for that protein. Proteins consistently showing decreased CSR were tested with western blot analysis to confirm deletion of the protein and labeled green. Proteins that did not show a downregulation in CSR were also analyzed for western blot as a control and labeled red if the protein remained after gRNA targeting (i.e. Dynll1).

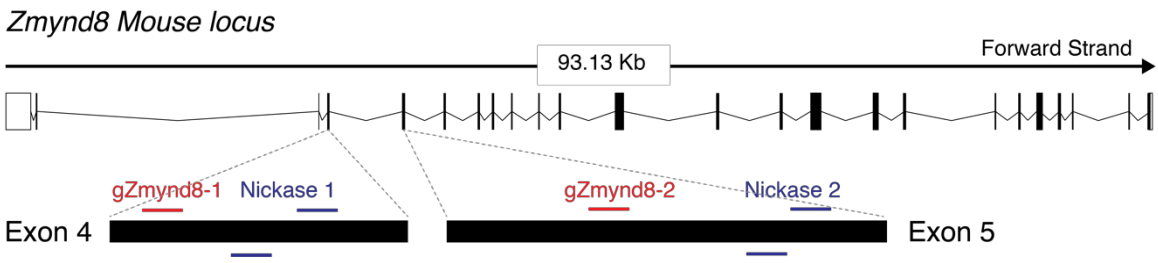
(F) Summary dot plot indicating CSR as a percentage of WT value (Ctrl: control, either empty vector or gRNA selected not to bind in mouse genome) within the same experiment. Each marker represents an independent experiment

A

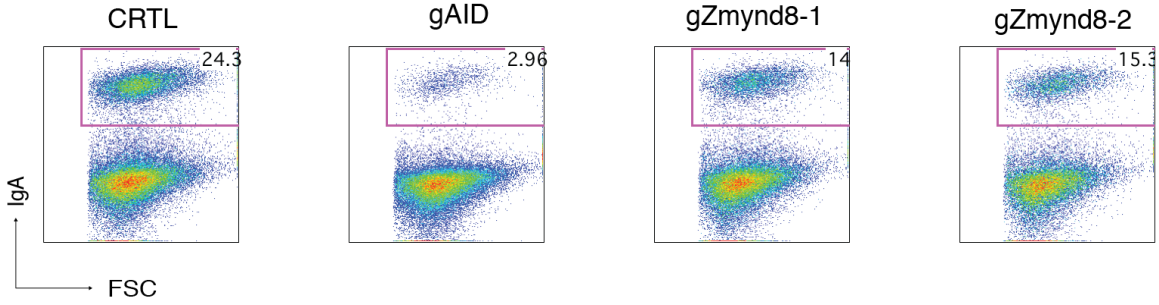
Protein	UniProtID	Z-Score
CHTF18	Q8BIW9	3.91
MGA	Q3ULI0	3.85
ZMYND8	Q8CHB3	3.79
ZNF592	Q8BHZ4	3.78
POLE	Q9WVF7	3.24
AHNAK	Q8CGE7	3.07
TLK2	Q3TTG3	2.91
TEX10	Q3URQ0	2.80
LAS1L	Q8C742	2.65
NAA50	Q6PGB6	2.18
RNF2	Q9CQJ4	2.15
CCDC124	Q9D8X2	2.14
TIMELESS	Q9R1X4	2.12
DYNLL1	Q9D6F6	2.00
PYHIN1	Q8BV49	1.16
ZRANB2*	Q9R020-2	0.90

RIF1	Q6PR54	4.73
53BP1	Q8BZ87	3.76

B


Zmynd8 Mouse locus
93.13 Kb
Forward Strand
gZmynd8-1
Nickase 1
gZmynd8-2
Nickase 2
Exon 4
Exon 5

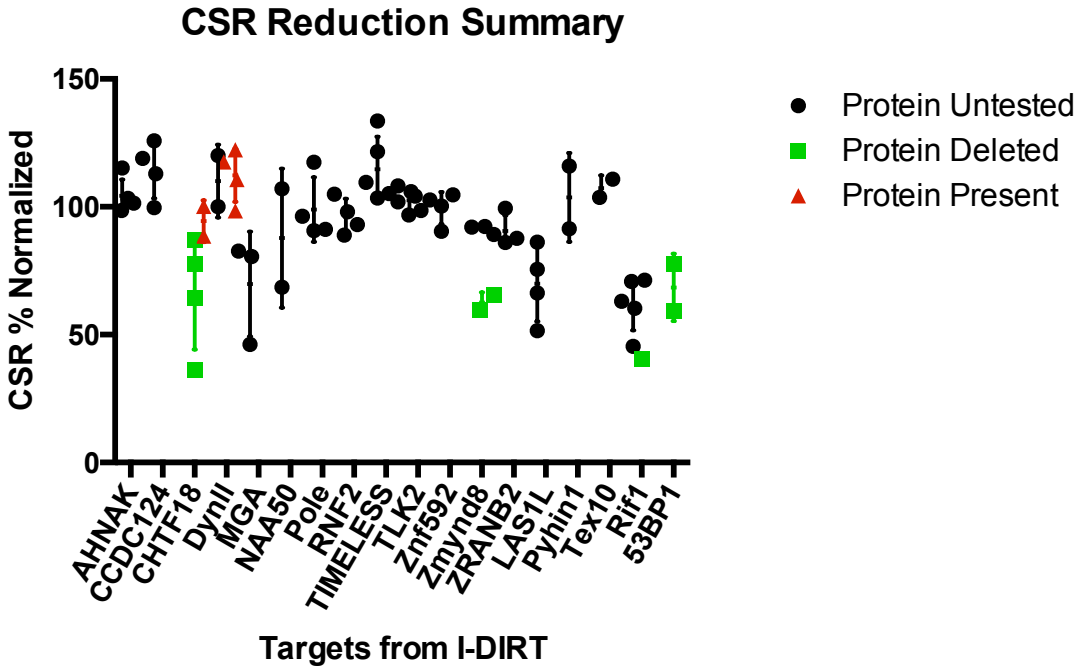
C


CRTL
gAID
gZmynd8-1
gZmynd8-2
24.3
2.96
14
15.3
IgA
FSC

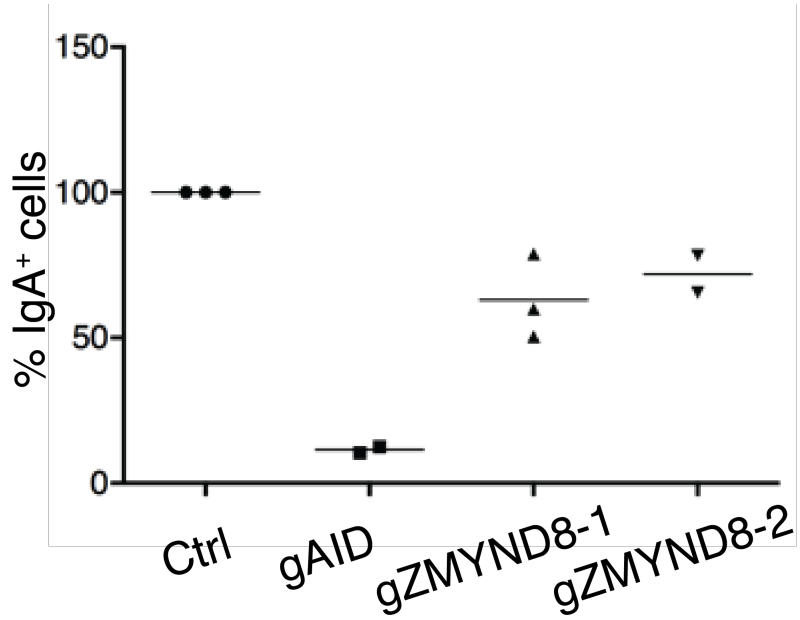
D


Ctrl
gAID
gZMYND8-1
gZMYND8-2
ZMYND8
β-Actin

E


CSR Reduction Summary
CSR % Normalized
150
100
50
0
AHNAK
CCDC124
CHTF18
Dynll
MGA
NAA50
Pole
RNF2
TIMELESS
TLK2
Znf592
Zmynd8
ZRANB2
LAS1L
Pyhin1
Tex10
Rif1
53BP1
Targets from I-DIRT
Protein Untested
Protein Deleted
Protein Present

F


% IgA+ cells
150
100
50
0
Ctrl
gAID
gZMYND8-1
gZMYND8-2

3.3 Deletion of Zmynd8 in clonal CH12 cell lines and validation as CSR factor

Given that we saw higher than expected CSR levels in known key CSR components (specifically AID, RIF1, and 53BP1 showing 10%, 50%, and 50% normalized CSR, respectively chapter 3.2) with bulk CH12 nucleofection and testing, we hypothesized that there might be heterogeneous targeting of proteins by gRNAs. Therefore, we sought to isolate clones of Zmynd8 targeted nucleofected CH12 cells to validate Zmynd8 in CSR.

Using the same methodology described in chapter 3.1, we single-cell sorted nucleofected CH12 cells and generated gRNA targeted clones. Numerous clones were characterized by Sanger sequencing of the gRNA target site, indel identification resulting in frameshift mutations on both alleles, and confirmation of protein deletion via western blot analysis. Each of the clones showed a greater CSR defect (figure 12a) than the bulk population (10-30% residual CSR versus 50-60% residual CSR (figure 11c and figure 11e). Western blot analysis of theses clones (figure 12b) showed a more complete deletion of Zmynd8 than was seen in bulk samples (figure 11D). Clones for AID, RIF1 and 53BP1 targeting also showed residual CSR levels commensurate with expected CSR in total knockout mouse models.

For further analysis, two clones were picked, termed Zmynd8_KO1 and Zmynd8_KO2. Zmynd8 was deleted with gRNAs targeting the 4th and 5th exon respectively. To confirm that Zmynd8 was responsible for CSR in these CH12 cell lines, we retrovirally transduced KO cell lines with a vector containing the full-length Zmynd8 gene (figure 12c). These reconstituted cell lines were activated for CSR and switched at normal levels (figure 12d and 12e).

These experiments confirmed that Zmynd8 is required for efficient CSR in the CH12 cell line.

Figure 12: Deletion of Zmynd8 in clonal CH12 cell lines and validation as CSR factor

(A) CSR defect in *Zmynd8*$^{-/-}$ clonal CH12 cell lines. Graph depicting CSR to IgA 48 hours after stimulation of WT, *AID*$^{-/-}$ and *Zmynd8*$^{-/-}$ CH12 clonal derivatives with αCD40, TGFβ and IL-4. Graph is representative of at least two independent experiments. Ctrl: controls, including both WT and clones derived from targeting CH12 with random sequences not present in the mouse genome. Significance was calculated with Mann-Whitney test.

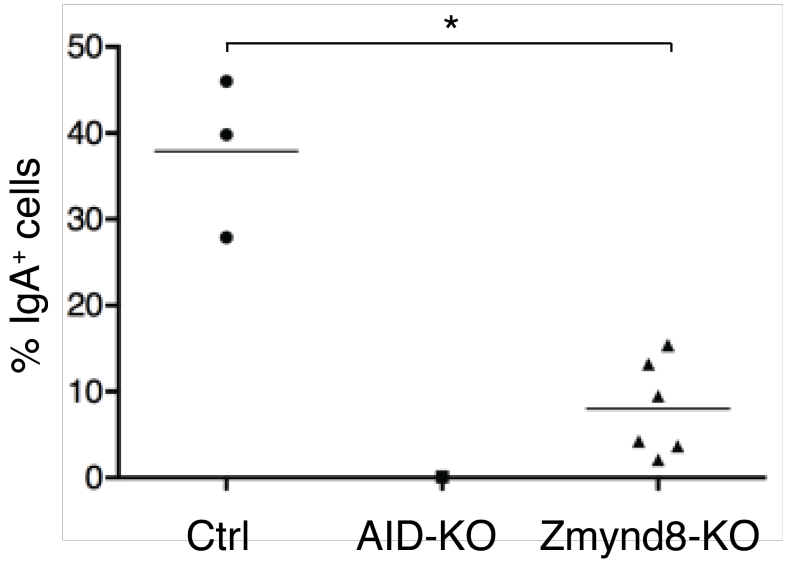
(B) Western blot analysis of Zmynd8 expression in Zmynd8_KO1 and Zmynd8_KO2 with WT control.

(C) Plasmid map of retroviral vector containing full length Zmynd8. Zmynd8 was cloned from the principal isoform as determined by Appris on ensemble. A 3x-FT peptide was added to the C-terminus of the protein for downstream applications.

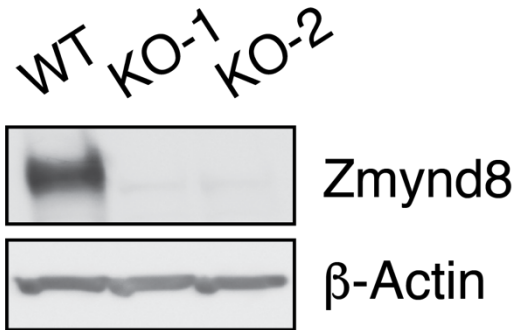
(D) Summary dot plot for four independent experiments measuring CSR to IgA 48 hours after activation of *Zmynd8*$^{-/-}$ CH12 cells lines reconstituted with empty vector (EV) or full-length Zmynd8. CSR is expressed as a percentage of WT value (uninfected WT cells) within the same experiment.

(E) Representative flow cytometry plot of reconstitution and activation experiment. WT, KO1 and KO2 cells lines were left un-reconstituted or reconstituted with empty retroviral vector or Zmynd8 full length. Cells were activated for 48 h with IL-4, TGF-**β,** and CD40. Cells were gated for live singlets and on GFP-/+ [un-reconstituted/reconstituted].

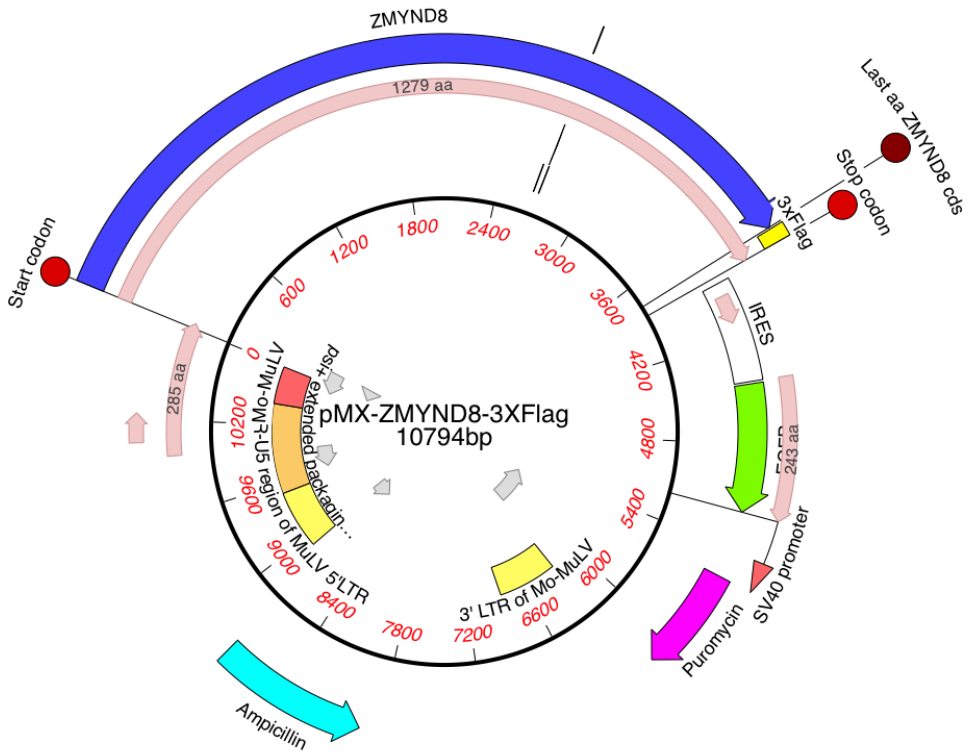
A


*
50
40
30
20
10
0
% IgA+ cells
Ctrl
AID-KO
Zmynd8-KO

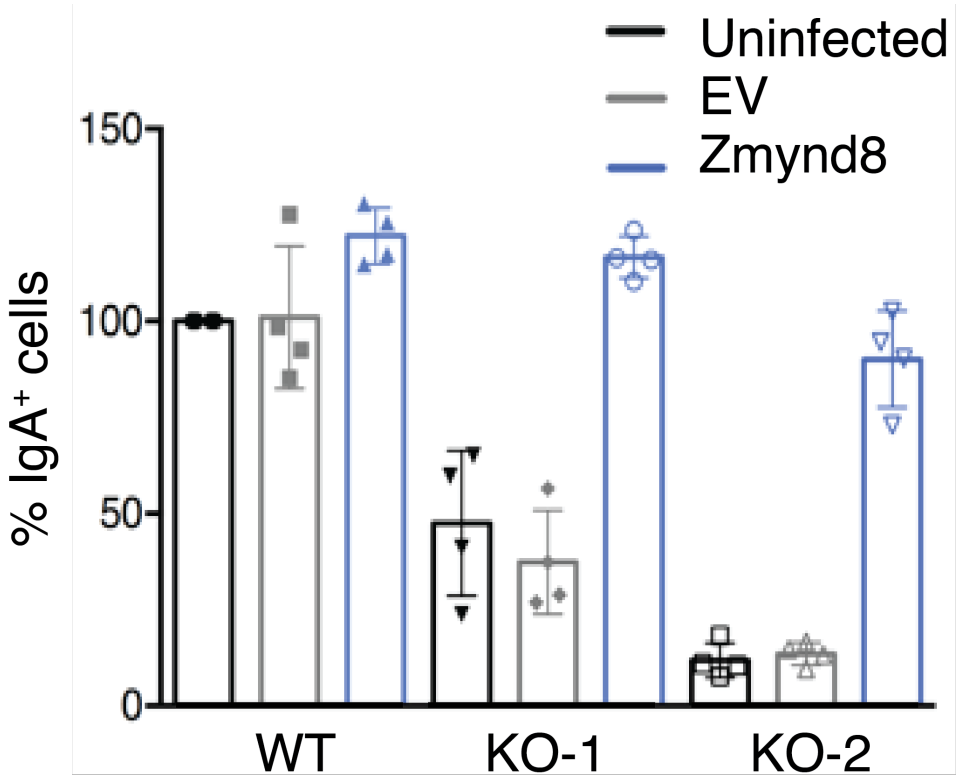
B

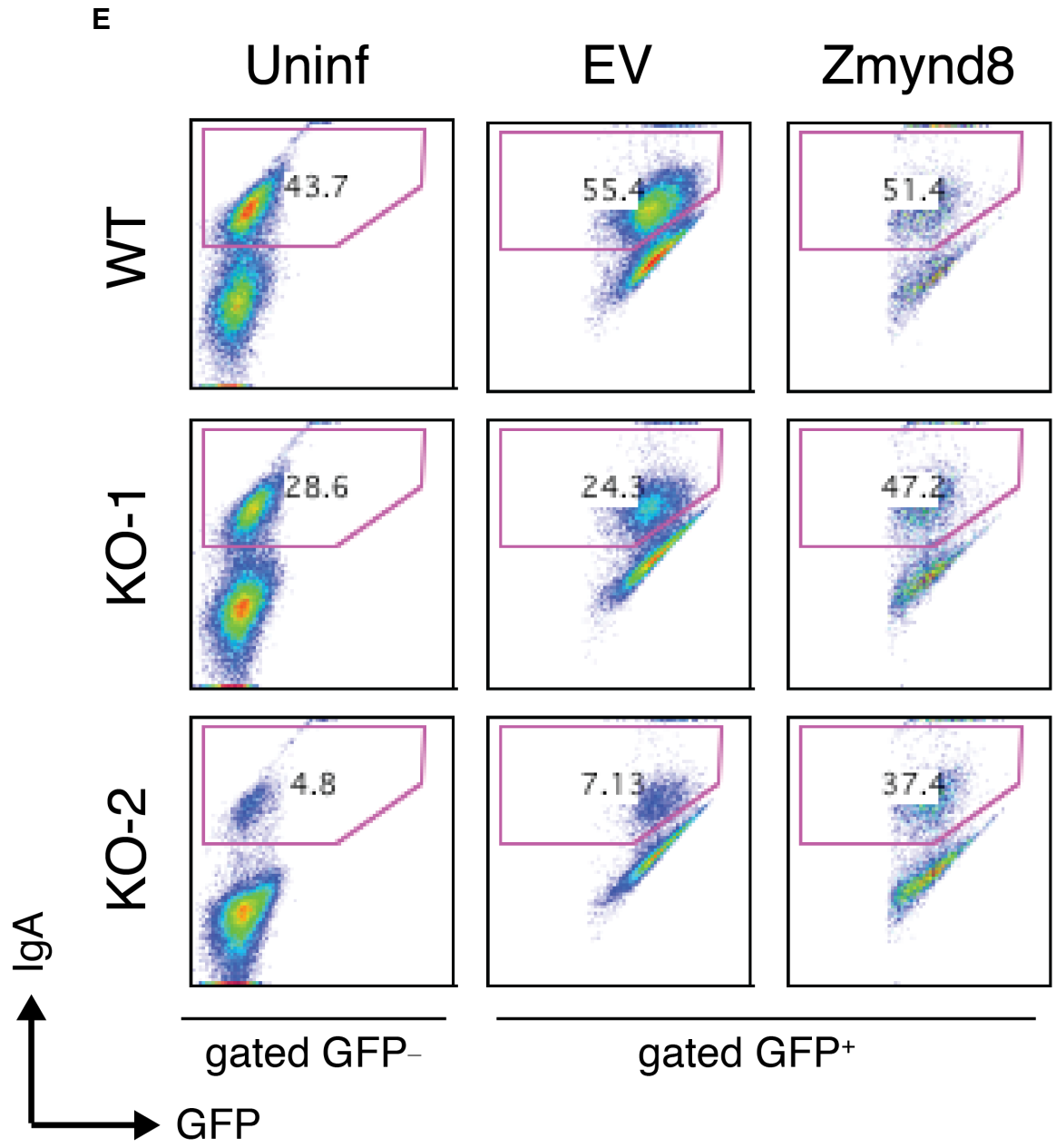

WT
KO-1
KO-2
Zmynd8
β-Actin

C


ZMYND8
1279 aa
Last aa ZMYND8 cds
Stop codon
3XFlag
Start codon
IRES
285 aa
243 aa
pMX-ZMYND8-3XFlag
10794bp
psi+ extended packaging
5' LTR region of MuLV
3' LTR of Mo-MuLV
SV40 promoter
Puromycin
Ampicillin
0
600
1200
1800
2400
3000
3600
4200
4800
5400
6000
6600
7200
7800
8400
9000
9600
10200

D


Uninfected
EV
Zmynd8
% IgA+ cells
150
100
50
0
WT
KO-1
KO-2


E
Uninf
EV
Zmynd8
WT
KO-1
KO-2
43.7
55.4
51.4
28.6
24.3
47.2
4.8
7.13
37.4
IgA
GFP
gated GFP–
gated GFP+

3.4 Zmynd8 validation as a CSR factor in Primary B cells

Many of the 17 I-DIRT candidates tested for CSR had no available knock out or conditional knock out mouse strains. Because of this we set out to utilize the recently constructed Cas9-2A-GFP knock in mice generated by the Zhang lab (Platt, et al., 2014) to somatically delete gRNAs in primary B-cells. In doing so we could confirm results detected in the CH12 cell line and continue experiments in primary B-cells.

The knock in mice contained a P2A linker to co-express eGFP with hSpCas9 from the *Rosa26* locus (Figure 13a). Because Cas9 was already present in B-cells (as determined by eGFP detection in flow cytometry) only the gRNA would need to be delivered to the splenocytes. This allowed us to design a retroviral vector dr170-03 (figure 13b) under the packaging size limit (~10-11kb) required to effectively spinoculate B cells. Ideally the gRNA would be delivered to B-cells prior to activation, allowing time for (1) production of the retroviral transduced gRNA, and (2) somatic targeting of the desired locus and (3) turnover of the pre-existing protein in the cell. However, normally cultured ex vivo splenocytes die without cytokine stimulation. To avoid this, cells were cultured ex vivo with the mitogen RP105 to stimulate cell proliferation without activation of CSR. An experimental schema can be seen in (figure 13c). Briefly, splenocytes were isolated and grown in culture with RP105 for 1 day, spinoculated the next day, activated 48 h later,

and measure for CSR at 72 h post activation. Representative flow cytometry plots (figure 13d) show minimal difference in CSR between in *Rosa26*$^{Cas9/+}$ and WT primary B-cells, greatly reduced switching in cells with gRNA targeted to AID, and about a 50% reduction in CSR in cells with gRNA targeted to Zmynd8. Three independent biological repeats (figure 13e) show a statistically significant decrease in gRNA against AID and also in gRNA against Zmynd8 (2 different gRNAs per gene). This reduction shows that Zmynd8 also functions in CSR in primary B-cells as well as the CH12 cell line.

Figure 13: Zmynd8 validation as a CSR factor in Primary B-cells

(A) Knock-in schema of Cas9-2A-eGFP at the Rosa26 locus adapted from (Platt, et al., 2014). The lox-stop-lox (LSL) section was excised in generating the mice. The hspCas9 protein is constitutively expressed from the Rosa26 locus and can be detected due to the coexpression of eGFP.

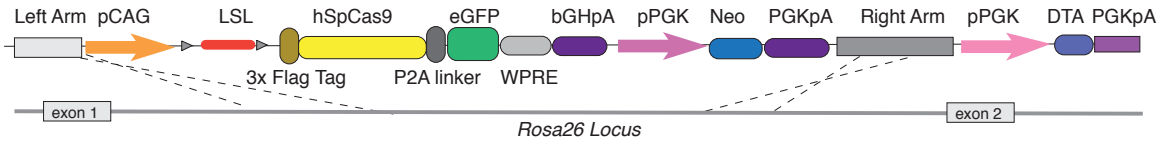
(B) Retroviral vector to deliver gRNA to Cas9 expressing B-cells. This vector was constructed based on a pMX backbone and incorporates a tandem BbsI cloning site for gRNA oligos, driven by a hU6 promoter.

(C) Experimental Outline for gRNA targeting in *Rosa26*$^{Cas9/+}$ primary B-cells. Splenocytes are harvested and grown ex vivo for 1 day with RP105 to stimulation proliferation. Cells are then spinoculated with a retrovirus containing the gRNA of interest and a BFP+ marker. After 2 days, the cells are activated by adding IL4 and LPS. At 72 h post activation cells are measured for CSR by staining for IgG1.

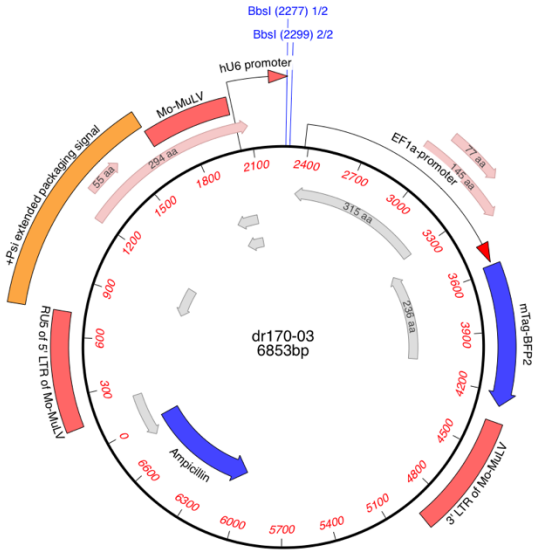
(D) Representative flow cytometry plot of gRNA targeting in *Rosa26*$^{Cas9/+}$ primary B-cells. Cells were filtered for live singlets, and BFP+ indicated successful transduction with the retroviral vector. gRNAs were identical to those used in CSR experiments on CH12 cells. All B-cells displayed GFP+ to confirm *Rosa26*$^{Cas9/+}$ genotype. Splenocytes were isolated and cultured in RP105 supplemented media for 24hrs prior to spinoculation with gRNA retroviruses. Cells were culture for another 2 days and then stimulated for CSR with RP105, IL-4, and LPS. Measurements were taken at 72hrs post stimulation. And EF1a promoter drives expression of mBFP2 to allow detection by flow cytometry.

(E) Dot plot summary of retorivrial gRNA targeting in *Rosa26*$^{Cas9/+}$ primary B-cells. Values are normalized to switching in empty gRNA vector for that experiment. Three independent biological replicates were conducted. * indicates significance as determine by a Mann-Whitney test.

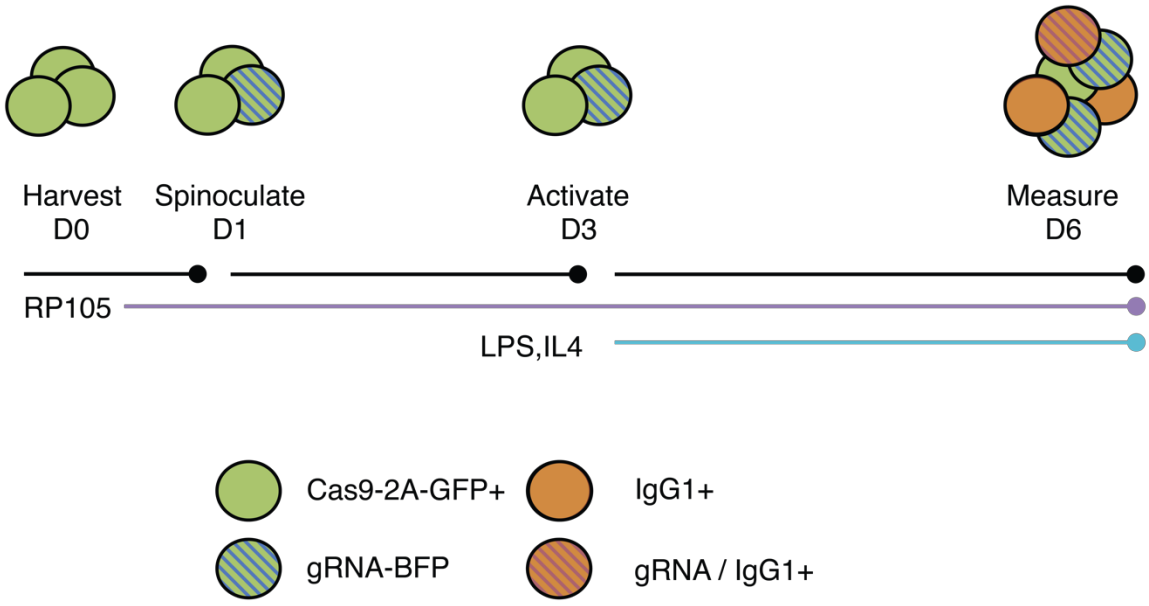
A


Left Arm
pCAG
LSL
hSpCas9
eGFP
bGHpA
pPGK
Neo
PGKpA
Right Arm
pPGK
DTA
PGKpA
3x Flag Tag
P2A linker
WPRE
exon 1
exon 2
Rosa26 Locus

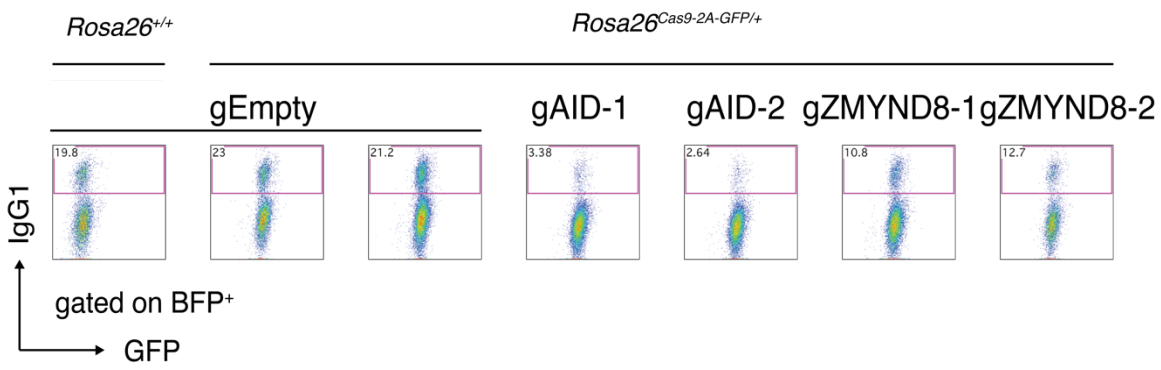
B


BbsI (2277) 1/2
BbsI (2299) 2/2
hU6 promoter
Mo-MuLV
+Psi extended packaging signal
EF1a-promoter
mTag-BFP2
RU5 of 5' LTR of Mo-MuLV
Ampicillin
3' LTR of Mo-MuLV
dr170-03
6853bp

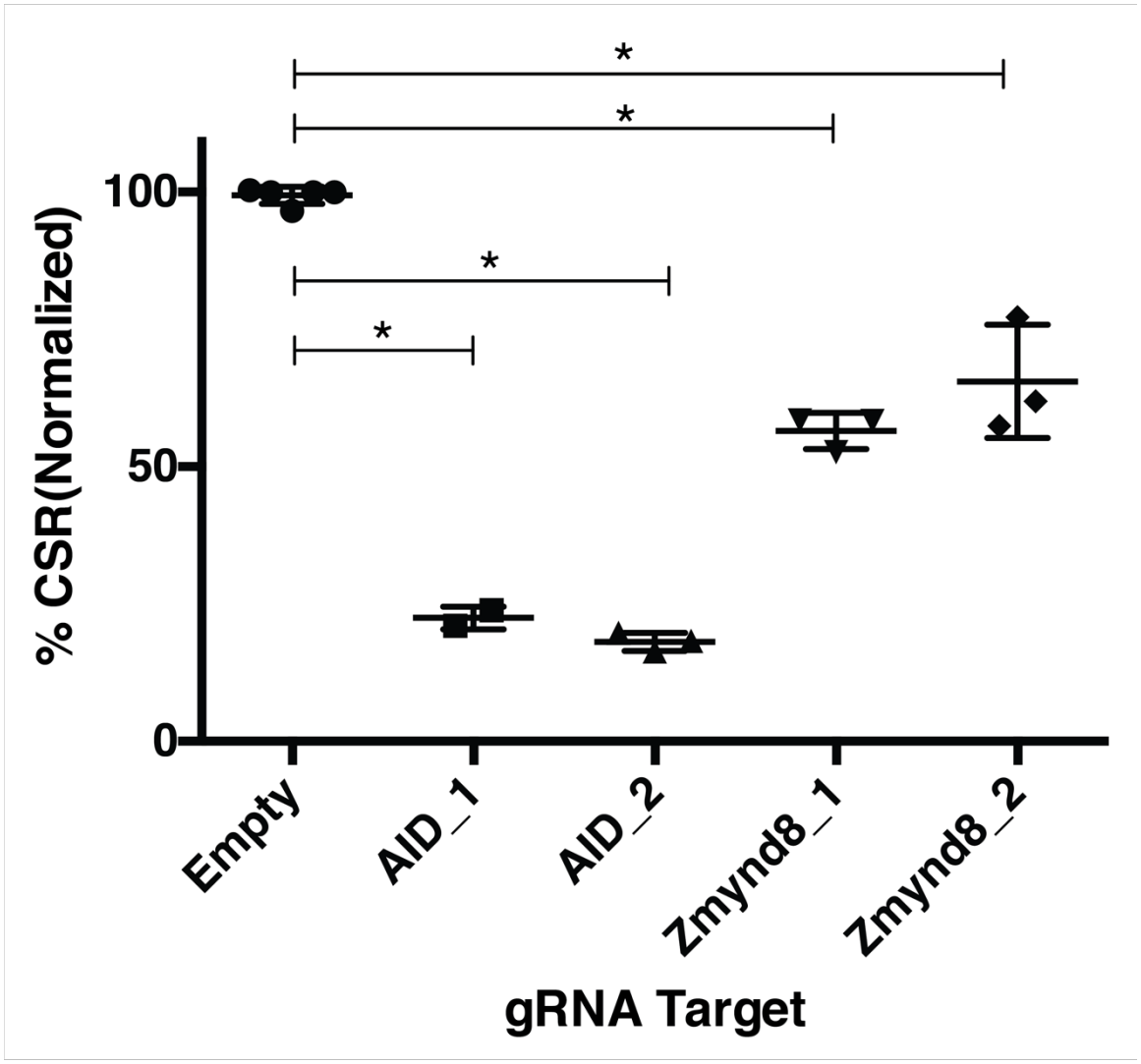
C


Harvest
D0
Spinoculate
D1
Activate
D3
Measure
D6
RP105
LPS,IL4
Cas9-2A-GFP+
IgG1+
gRNA-BFP
gRNA / IgG1+

D


Rosa26+/+
Rosa26Cas9-2A-GFP/+
gEmpty
gAID-1
gAID-2
gZMYND8-1
gZMYND8-2
19.8
23
21.2
3.38
2.64
10.8
12.7
IgG1
gated on BFP+
GFP

E


*
*
*
*
100
50
0
% CSR(Normalized)
Empty
AID_1
AID_2
Zmynd8_1
Zmynd8_2
gRNA Target

Chapter 4: ZMYND8 in CSR: Pre-Break

AID is the apex initiator of CSR. It is primarily regulated transcriptionally (Zan & Casali, 2013). It is necessary for CSR in B-cells (Muramatsu, et al., 2000) and ectopic expression can even induce CSR in non-B cells (Okazaki, Kinoshita, Muramatsu, Yoshikawa, & Honjo, 2002). AID is a haploinsufficient gene as reductions in AID transcription levels decrease CSR (Sernandez, de Yebenes, Dorsett, & Ramiro, 2008), (Takizawa, et al., 2008). Additionally, micro-RNAs, specifically miR-155 (An, et al., 2010), regulate AID post-transcriptionally (Grace, et al., 2008). Additionally, CSR requires expression of sterile non-coding (germline) transcripts from the intronic promoters preceding the constant regions (Stavnezer, Guikema, & Schrader, 2008). Lastly, CSR requires cell proliferation (An, et al., 2010), (Stavnezer, Guikema, & Schrader, 2008). A defect in any of these processes would result in impaired CSR. As such we sought to confirm that Zmynd8 deletion did not result in absent or reduced AID or GLT transcripts or defective proliferation.

4.1 Zmynd8 and AID expression

We chose to measure AID RNA transcripts via RT-qPCR. RNA from CH12 cells was collected at 48 h post induction and measured. Transcripts were normalized

to GAPDH and to AID levels in WT CH12 cells. There was no reduction in AID transcripts according to RT-qPCR (figure 14a)

However, Zmynd8 could potentially be regulating AID post-translationally (Basu, Franklin, & Alt, 2009). To test this we performed a western blot analysis of whole cell extract of nucleofected bulk CH12 cells (figure 14b). We did not see any reduction in AID protein levels for four different gRNAs against Zmynd8 tested.

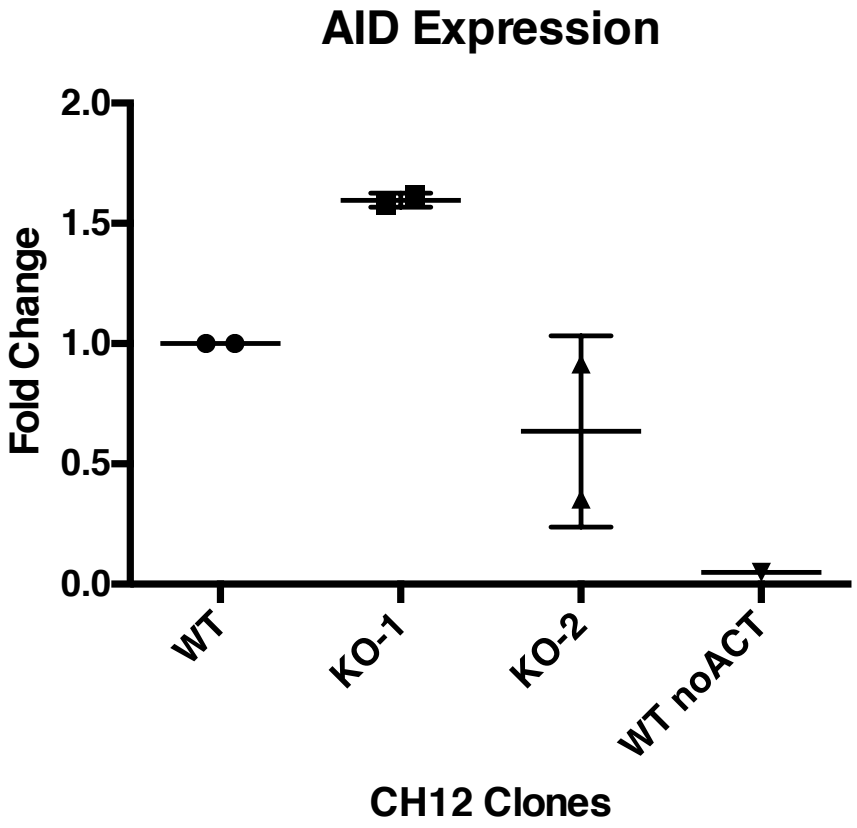
Lastly, Zmynd8 could potentially be regulating AID function by affecting (1) localization (i.e. altering the distribution between cytoplasmic AID and nuclear AID) (Hasler, Rada, & Neuberger, 2011), (Wang, et al., 2016); (2) stability and degradation (Uchimura, Barton, Rada, & Neuberger, 2011), (Aoufouchi, et al., 2008); (3) AID phosphorylation (Basu, et al., 2005), (McBride, et al., 2006); and (4) general AID activity (Li, et al., 2012). We chose not to pursue those investigations, respectively, because (1) Zmynd8 has not been shown as an AID binding factor or vice-versa, (2) Zmynd8 has not shown to function in protein degradation or stability, and (3) Zmynd8 has no putative kinase type activity or (4) iron sequestering capability. If Zmynd8 were altering CSR through one of these four mechanisms, we would expect it to do so via a secondary mediator.

Figure 14: Zmynd8 does not alter AID levels

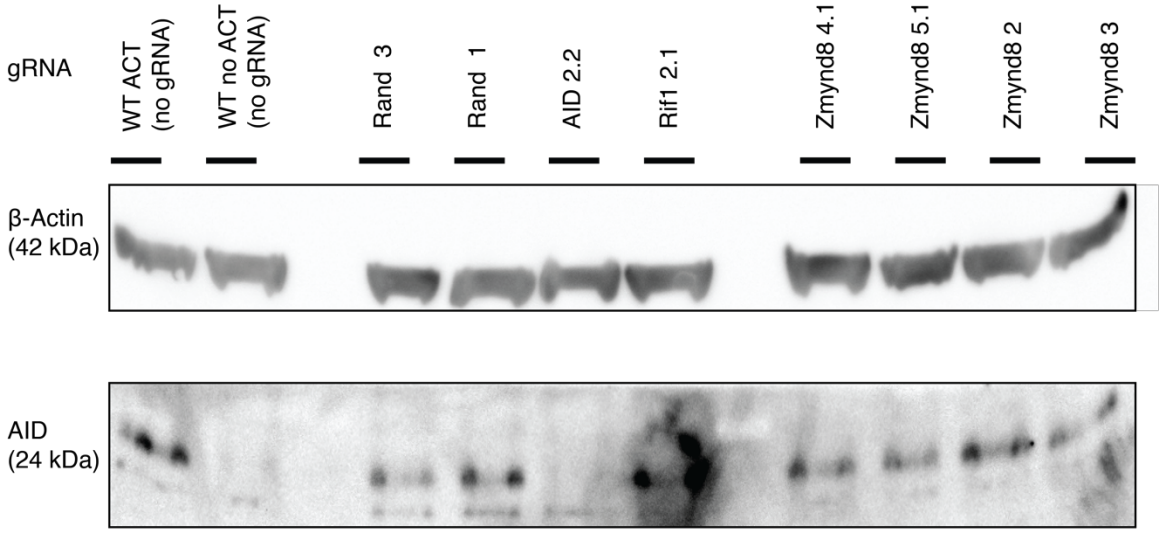
(A) Zmynd8 does not lower AID expression. WT, Zmynd8 KO-1, Zmynd8 KO-2, clones were activated for 48 h with IL-4, TGF- β, and CD40. RNA was isolated, and RT-qPCR performed. A WT unactivated sample was included as a point of reference. WT was assigned an arbitrary value of 1.0. There was no difference between WT, KO-1 and KO-2 using the Mann Whitney U test to test for significance (error bars represent SD).

(B) Western blot analysis of whole cell extract of nucleofected bulk cells with gRNA targets. Cells were sorted on successful nucleofection (GFP+)but not clonally isolated, except for the first two columns which are un-nucleofected controls. All cells except for column 2 (WT no ACT) were activated with IL-4, TGF- β, and CD40 and harvested at 48h.

A


AID Expression
Fold Change
2.0
1.5
1.0
0.5
0.0
WT
KO-1
KO-2
WT noACT
CH12 Clones

B


gRNA
WT ACT (no gRNA)
WT no ACT (no gRNA)
Rand 3
Rand 1
AID 2.2
Rif1 2.1
Zmynd8 4.1
Zmynd8 5.1
Zmynd8 2
Zmynd8 3
β-Actin (42 kDa)
AID (24 kDa)

4.2 ZMYND8 and Germline Transcription

Germline transcription (GLT) is necessary for CSR. To briefly summarize (Matthews, Zheng, Di Menna, & Chaudhuri, 2014) and (Yewdell & Chaudhuri, 2017), unlike AID expression, GLTs are specific for each isotype. GLT through switch μ is constitutive while downstream acceptor GLT, in the case of CH12 switching from IgM to IgA GLTα is inducible upon activation. GLTs facilitate locus accessibility and AID targeting to the switch regions in the IgH so AID can initiate DSBs that precede CSR. We wanted to investigate if a decrease in GLT accounted for the CSR defect in Zmynd8 KO CH12 cells.

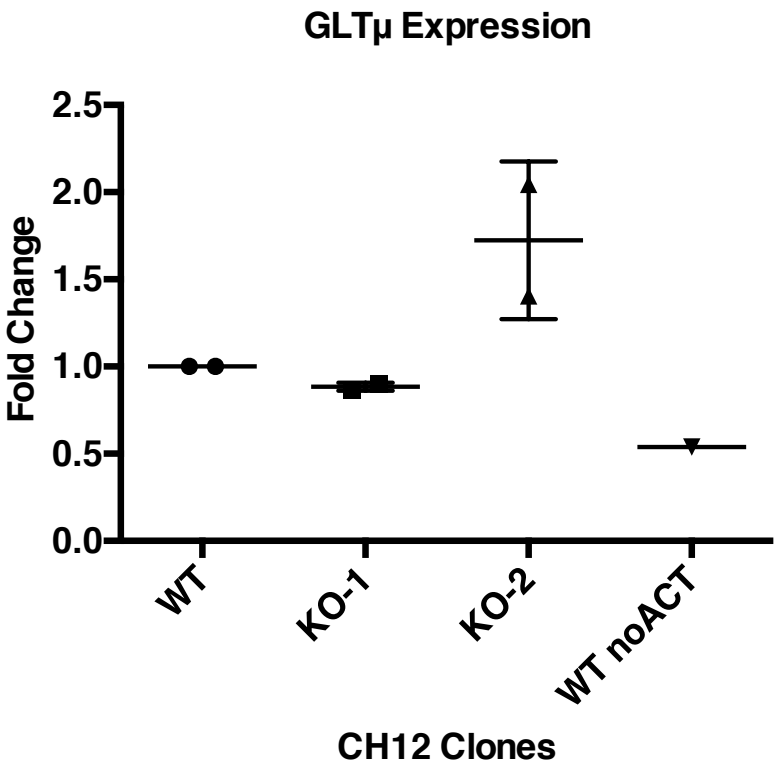
We chose to measure the relevant germline transcripts in CH12 cells, GLTμ and GLTα, via RT-qPCR. RNA from CH12 cells was collected at 48 h post induction and measured. Transcripts were normalized to GAPDH and to GLT levels in WT CH12 cells. There was no reduction in GLTμ or GLTα transcripts according to RT-qPCR (figure 15a, figure 15b).

Figure 15: Zmynd8 deletion in CH12 cell line does not reduce GLT

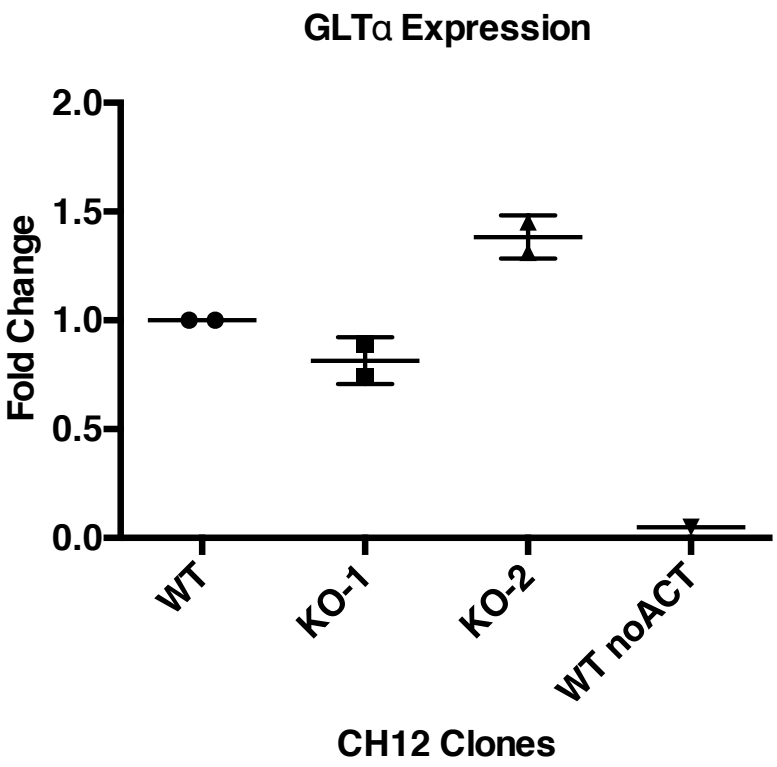
(A) Zmynd8 deletion does not reduce GLTμ expression. WT, Zmynd8 KO-1, Zmynd8 KO-2, clones were activated for 48 h with IL-4, TGF- β, and CD40. RNA was isolated, and RT-qPCR performed. A WT unactivated sample was included as a point of reference. WT was assigned an arbitrary value of 1.0. There was no difference between WT, KO-1 and KO-2 using the Mann Whitney U test to test for significance (error bars represent SD). Data are representative of two biologically independent experiments.

(B) Zmynd8 deletion does not reduce GLTα expression. WT, Zmynd8 KO-1, Zmynd8 KO-2, clones were activated for 48 h with IL-4, TGF- β, and CD40. RNA was isolated, and RT-qPCR performed. A WT unactivated sample was included as a point of reference. WT was assigned an arbitrary value of 1.0. There was no difference between WT, KO-1 and KO-2 using the Mann Whitney U test to test for significance (error bars represent SD). Data are representative of two biologically independent experiments.

A


GLTμ Expression
Fold Change
2.5
2.0
1.5
1.0
0.5
0.0
WT
KO-1
KO-2
WT noACT
CH12 Clones

B


GLTα Expression
Fold Change
2.0
1.5
1.0
0.5
0.0
WT
KO-1
KO-2
WT noACT
CH12 Clones

4.3 Zmynd8 and Cell Proliferation

Cell proliferation is necessary for CSR (Reviewed in (Stavnezer, Guikema, & Schrader, 2008)). We sought to investigate if reduced cell proliferation or altered cell cycle distribution contributed to the reduction in CSR. By measuring cell divisions via CFSE (5-(and -6)-carboxyfluorescein diacetate succinimidyl ester) dilution, we wanted to observe if CH12 Zmynd8 KO cell lines exhibited reduced proliferation.

CFSE staining is a well-established protocol to monitor lymphocyte division (Lyons & Parish, 1994), (Quah, Warren, & Parish, 2007). Treatment with CFSE, or trade-named fluorescent variants CellTrace, covalently labels intracellular amines, and each cell division halves the signal intensity as molecules are divided between daughter cells. We CFSE labeled un-reconstituted, reconstituted with empty vector, and reconstituted with full-length Zmynd8 CH12 clones, activated the cells and measured dilution differences at 48hr.

Unfortunately, CH12 cells could not resolve individual cell division peaks, and the population of cells move homogeneous as a single peak. Noting this sensitivity limitation, we observed no differences between non-reconstituted, empty reconstituted, and Zmynd8 reconstituted WT cells (figure 16a). In the Zmynd8 KO cell lines KO-1 and KO-2 we also observed no differences (figure 16b), (figure

16c). While this does not exclude the possibility of subtle cell division differences, we could not observe any gross population level differences in cell proliferation that could account for CSR deficiency.

Figure 16: Z8 Deficiency does not lead to reduced proliferation in CH12 cells

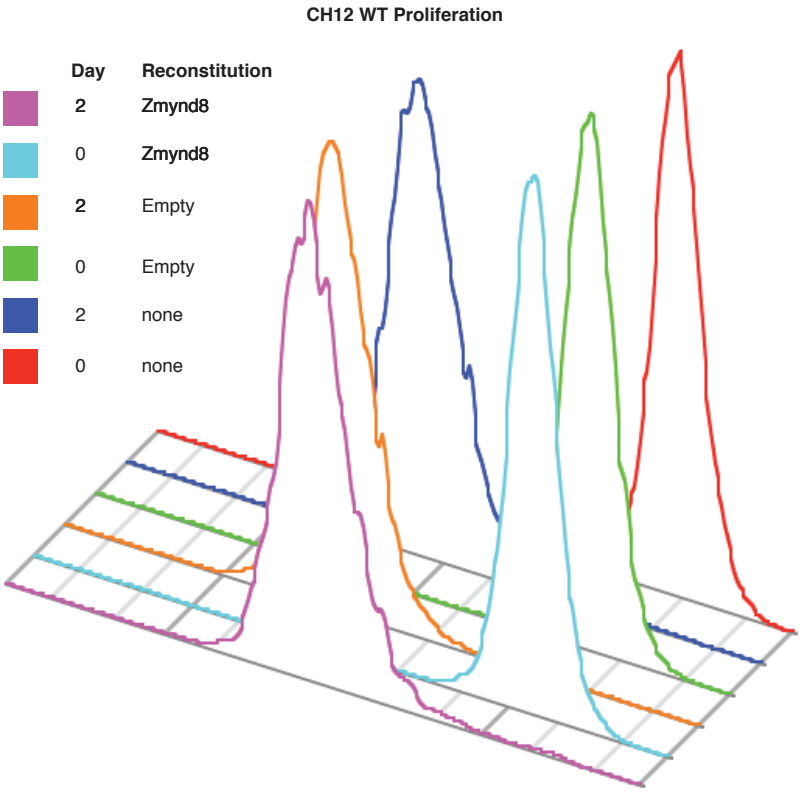
All cultures were reconstituted with nothing/mock, empty vector, or full length Zmynd8 vector as indicated. After selection with puromycin and recovery, and immediately prior to activation, cells were stained with CellTrace Violet (CFSE), and measured by flow cytometry. Cells were then activated to switch with IL-4, TGF- β, and CD40. 48 h post activation cells were measured to assess for class switching and CellTrace dilution.

(A) Combined histogram of WT CH12 cells at 0 h and 48 h as indicated.

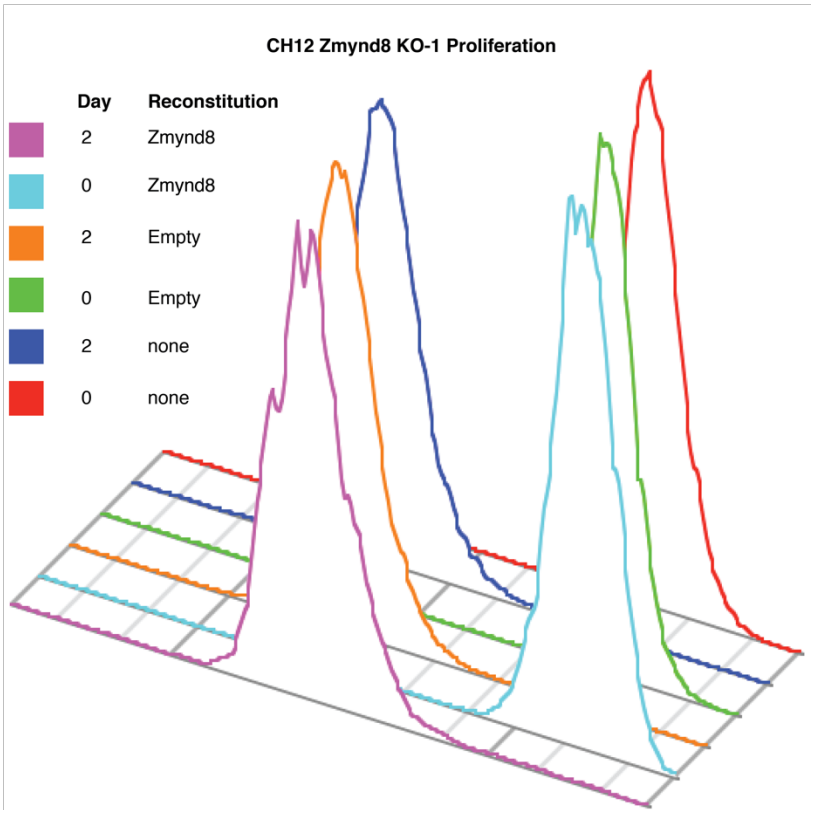
(B) Combined histogram of Zmynd8 KO-1 CH12 cells at 0 h and 48 h as indicated.

(C) Combined histogram of Zmynd8 KO-2 CH12 cells at 0 h and 48 h as indicated.

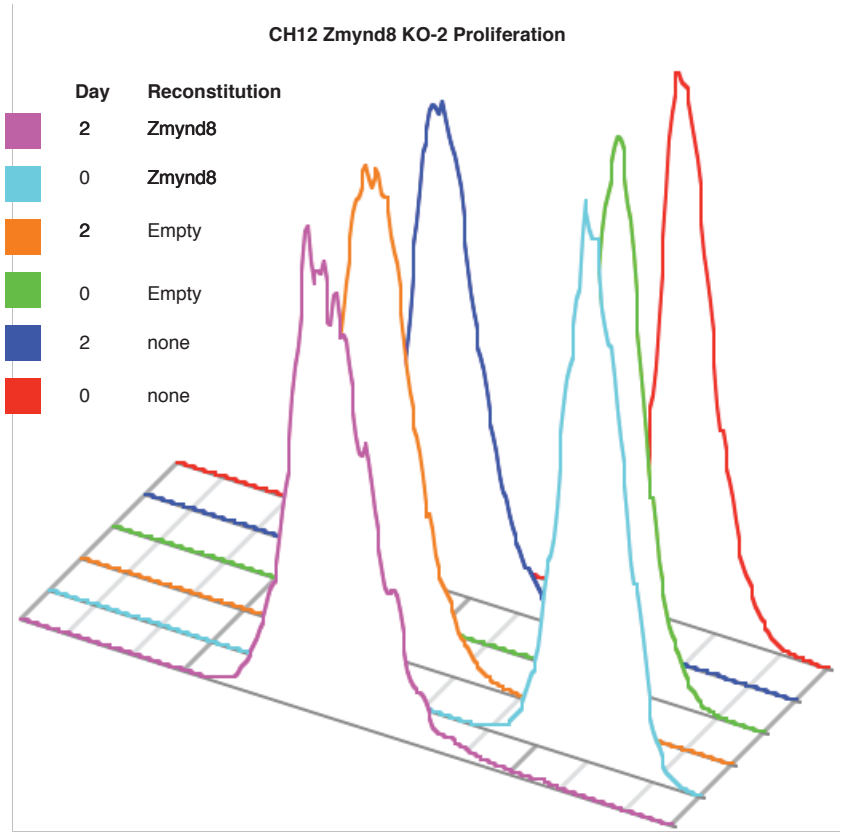
A


CH12 WT Proliferation
Day
Reconstitution
2
Zmynd8
0
Zmynd8
2
Empty
0
Empty
2
none
0
none

B


CH12 Zmynd8 KO-1 Proliferation
Day
Reconstitution
2
Zmynd8
0
Zmynd8
2
Empty
0
Empty
2
none
0
none

C


CH12 Zmynd8 KO-2 Proliferation
Day
Reconstitution
2
Zmynd8
0
Zmynd8
2
Empty
0
Empty
2
none
0
none
SSC Singlets

Chapter 5: ZMYND8 in CSR: Post-Break

The DNA damage response to DSBs is highly coordinated and involves successive, hierarchical localization of various factors (Panier & Boulton, 2014), (Zimmermann & de Lange, 2014). Specifically, proteins ϒH2AX, 53BP1 and RIF1 organize sequentially to form foci at the chromatin flanking DSBs to facilitate repair (Polo & Jackson, 2011). Such foci can be visualized with immunofluorescence microscopy. Experimentally, cells are exposed to ionizing radiation and the formation of these irradiation-induced foci (IRIF) is measured. DSBs are necessary for CSR. As such, defects in the detection, signaling cascade, or repair of these breaks can impede CSR (Methot & Di Noia, 2017).

Because the DSB repair processes are common both to AID induced breaks and non-physiologic damage like ionizing radiation, we can investigate IRIFs and DSB repair in one cell type and apply those findings to DSB repair in CSR.

We set out to explore the formation via IRIF and hierarchy of Zmynd8 and known DSB repair factors in immortalized mouse embryonic fibroblasts (iMEFs). On a technical note, iMEFs are typically used for such investigations instead of B-cells because adherent cells can be visualized by the confocal microscopy necessary for IRIF detection.

5.1 DNA Damage Response Signaling and Zmynd8

First, Zmynd8 iMEF cell lines needed to be generated. Using the same transient expression of Cas9 and relevant gRNAs used to make the KO CH12 cell lines, we nucleofected iMEFs with the same Cas9+gRNA plasmids. iMEF clonal derivatives KO-1, KO-2, KO-3 showed loss of the Zmynd8 protein by western blot analysis (figure 17a). Consistent with the loss of protein in the blots, immunofluorescence also showed a loss of Zmynd8 (Figure 17b). It is also noteworthy that Zmynd8 is restricted to the nucleus of iMEFs. However, treatment with ionizing radiation did not lead to detectable Zmynd8 IRIFs (Figure 17c). Unlike ϒH2AX, 53BP1 and RIF1 (Figures 17d, e, f) Zmynd8 does not appear to form IRIF post-treatment with ionizing radiation. This does not exclude Zmynd8 from the DDR process, but it may mean that Zmynd8 does not concentrate at breaks in sufficient concentration to be visible by IF (many NHEJ components, i.e. Ku70 and Ku80, are in this category (Britton, Coates, & Jackson, 2013)). Despite the inability to form foci under the specified irradiation treatment, we could still assess if ϒH2AX, 53BP1 and RIF1 IRIF are dependent on Zmynd8. Foci for these three factors showed no decrease in formation Figures 17d, e, f), which indicates that Zmynd8 does not function upstream of ϒH2AX, 53BP1 or RIF1 in DSB repair.

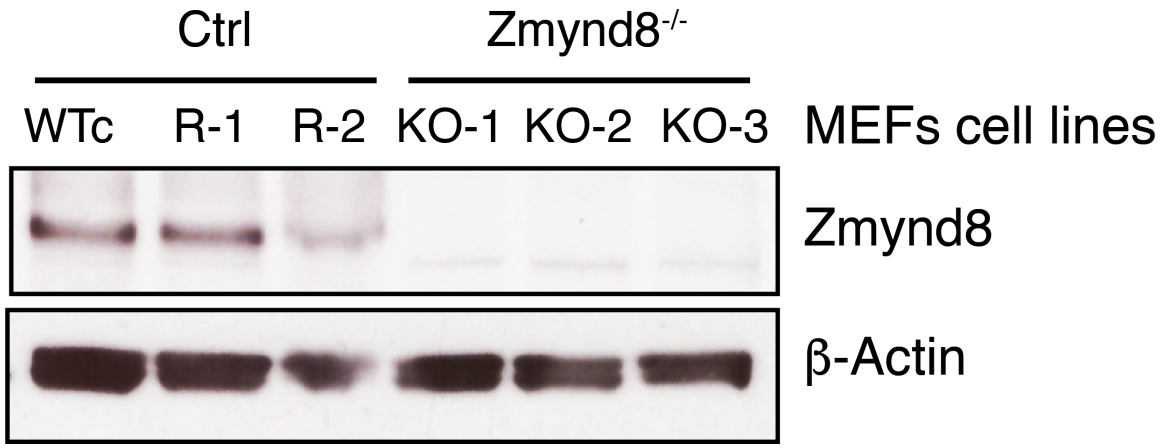
Figure 17: IRIF formation in Zmynd8 Deficient MEFs

(A) Western blot analysis of immortalized mouse embryonic fibroblasts (iMEF) clonal cells lines. Ctrl iMEFs cell lines include a WT clonal derivative (WTc) and two clones generated by targeting iMEFs cells with a random sequence not present in the mouse genome (R-1 and R-2)

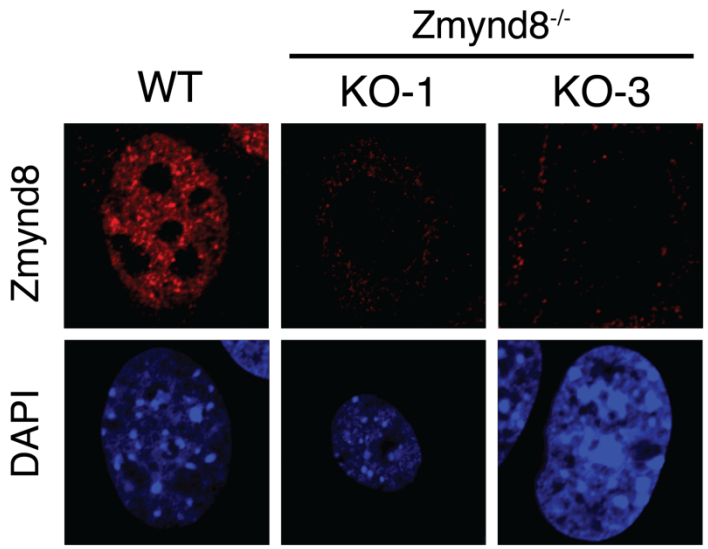
(B) Immunofluorescent staining of Zmynd8-/- and WT iMEFs. Zmynd8 staining is shown in red, and 4',6-diamidino-2-phenylindole (DAPI) counterstains DNA in the nucleus blue. Cells were not treated with irradiation.

(C-F) Immunofluorescent staining of Zmynd8-/- and WT iMEFs with 10 Gy irradiation and allowed to recover for 90 min and 6 h.

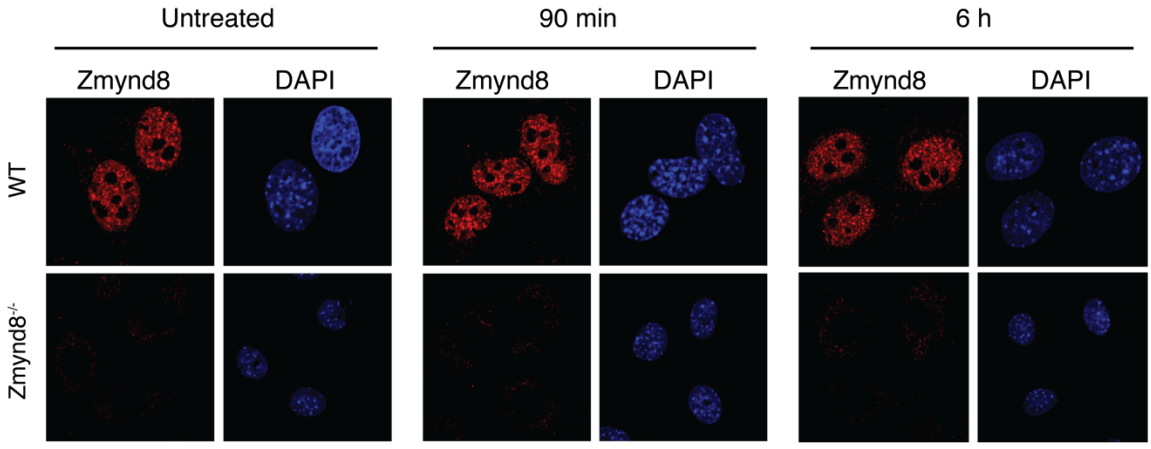
A

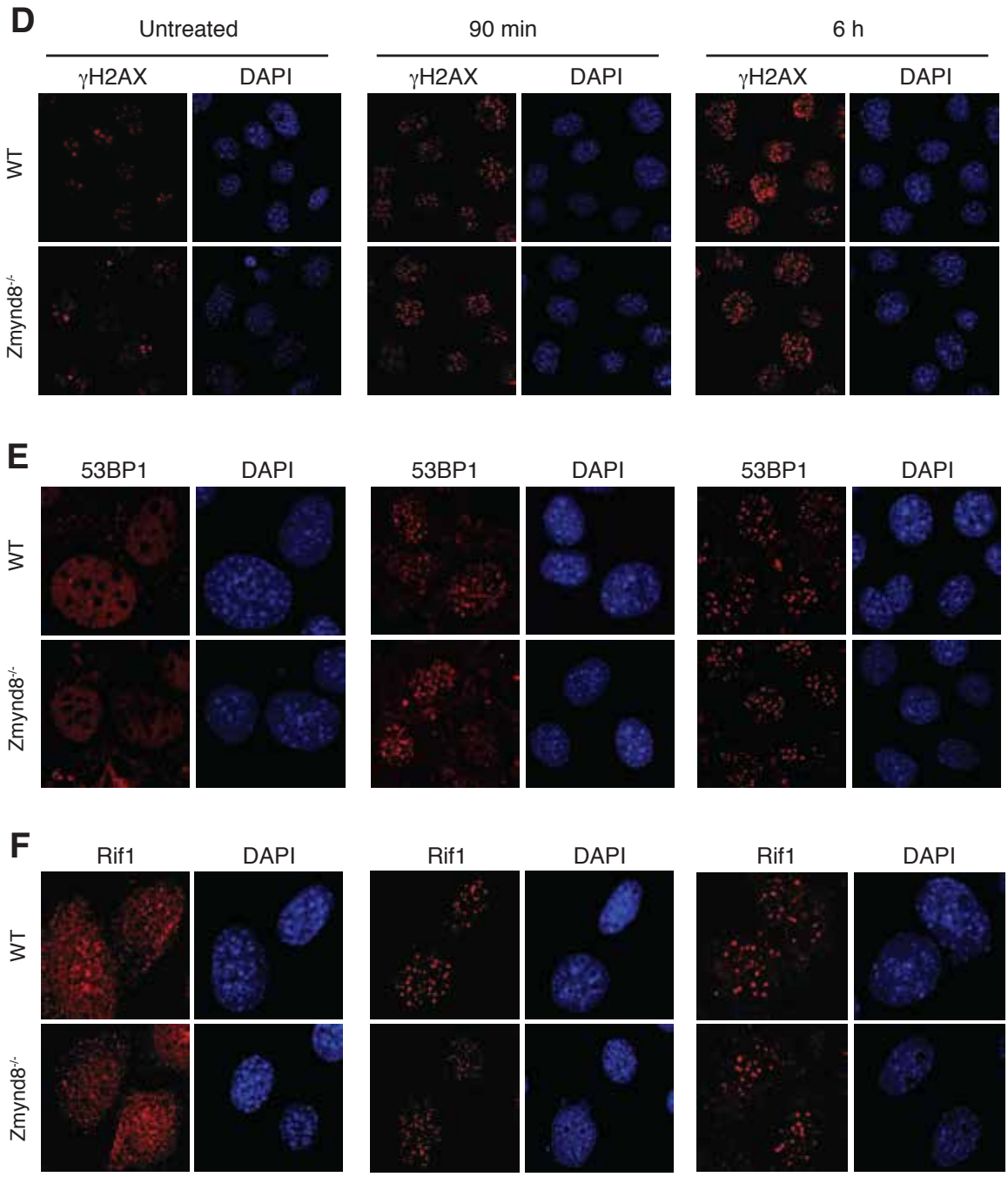

Ctrl
Zmynd8-/-
WTc
R-1
R-2
KO-1
KO-2
KO-3
MEFs cell lines
Zmynd8
β-Actin

B


Zmynd8-/-
WT
KO-1
KO-3
Zmynd8
DAPI

C


Untreated
90 min
6 h
Zmynd8
DAPI
WT
Zmynd8-/-


D
Untreated
90 min
6 h
γH2AX
DAPI
γH2AX
DAPI
γH2AX
DAPI
WT
Zmynd8-/-
E
53BP1
DAPI
53BP1
DAPI
53BP1
DAPI
WT
Zmynd8-/-
F
Rif1
DAPI
Rif1
DAPI
Rif1
DAPI
WT
Zmynd8-/-

5.2 Genomic Instability and Zmynd8

Chapter 5.1 detailed that Zmynd8 does not form IRIF and does not alter IRIF formation in the ϒH2AX, 53BP1 and RIF1 axis in repair of DSBs. This does not exclude Zmynd8 participating in the DNA repair process post RIF1 foci formation or in another branch of DNA repair (Ciccia & Elledge, 2010).

A deficiency in DNA repair would contribute to genomic instability in the cell. Alterations of basic cellular processes, i.e. the ability to proliferate and the frequency of cell death, are outcomes of genomic instability. To broadly determine if Zmynd8 plays a role in DNA repair we chose to assess cell survival and proliferation in Zmynd8 deficient cells.

Genomic instability can be detected by a clonogenic survival assay (Franken, Rodermond, Stap, Haveman, & van Bree, 2006). This is an in vitro assay to determine the rate of cell death and proliferation capabilities in surviving cells after treatment with irradiation or other radiomimetic agents. iMEFs were seeded, irradiated, grown and then stained with crystal violet to visualize cell growth. The presence of Zmynd8 had no impact on clonogenic survival (representative data, figure 18a; summary data, figure 18b). Cell counting proliferation assays showed that Zmynd8 also does not function in survival and post irradiative growth in CH12 cells (figure 18c). Confirming this point, we showed that BrdU incorporation, a well-

established technique for detecting cell growth through the uptake of the BrdU nucleotide analog in replicating DNA (Porstmann, Ternynck, & Avrameas, 1985), (Magaud, et al., 1989), and (Huong, et al., 1991) is not altered by the Zmynd8 deficiency (figure 18d).

It is possible, that Zmynd8's effect on cell survival may only manifest under particular conditions, like BRCA deficiency and PARP inhibition (Helleday, 2011). To test this we performed the clonogenic survival assay with PARP treatment instead of ionizing irradiation. Zmynd8$^{-/-}$ iMEF cell lines showed no survival defect (figure 18e). Lastly, Zmynd8$^{-/-}$ iMEFs did not show an increase in chromosomal aberrations as visualized with metaphase spreading (18f).

Figure 18: Zmynd8 deficiency does not lead to impaired cell survival, growth or genomic instability

(A) Representative image of clonogenic survival assay. iMEFs of the indicated genotypes were reconstituted with empty vector or full length Zmynd8 as listed, seeded for 24hr, irradiated as marked, and stained after 10 days of growth with crystal violet.

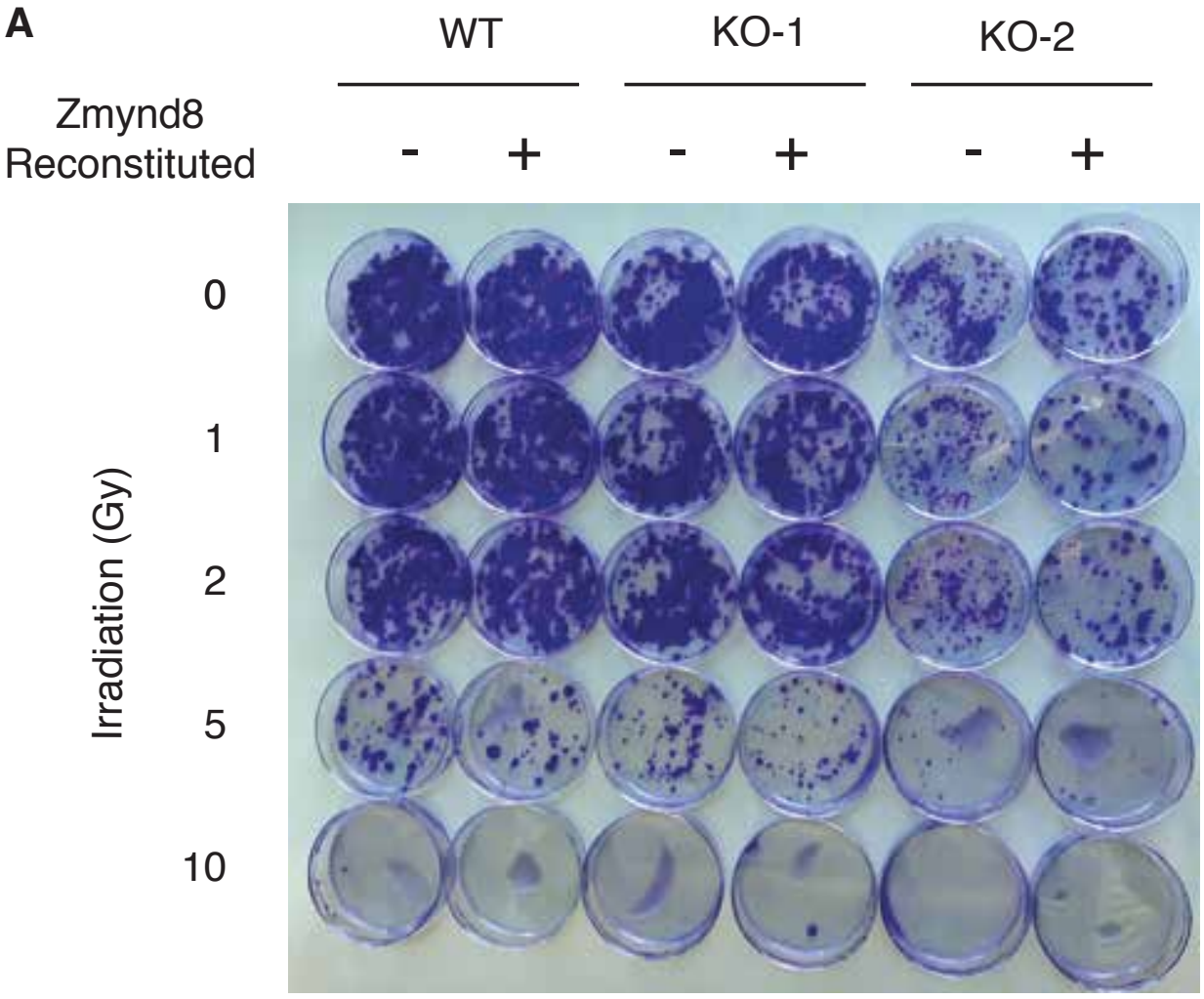
(B) Summary data of clonogenic survival assay. Error bars represent the mean from triplicate plates per condition. Data are representative of two independent experiments

(C) Growth curves of WT and *Zmynd8*$^{-/-}$ CH12 cells after irradiation. Measurements were taken in triplicate.

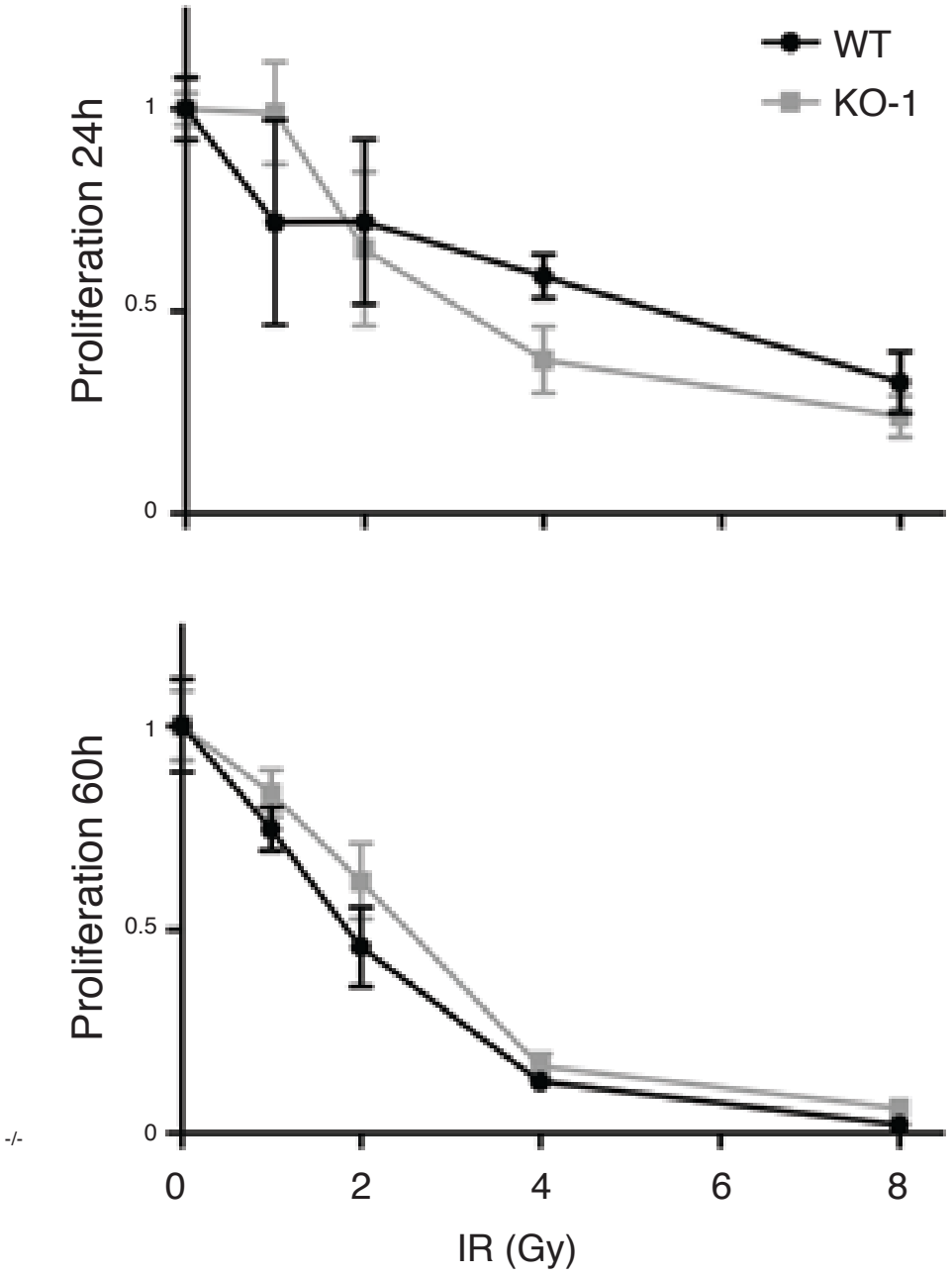
(D) Cell cycle analysis of irradiated *Zmynd8*$^{-/-}$ CH12 lines as measured by BrdU incorporation.

(E) Clonogenic survival assay of *Zmynd8*$^{-/-}$ iMEF clones following PARPi treatment. Error bars represent the mean from triplicate plates per condition. Data are representative of two independent experiments.

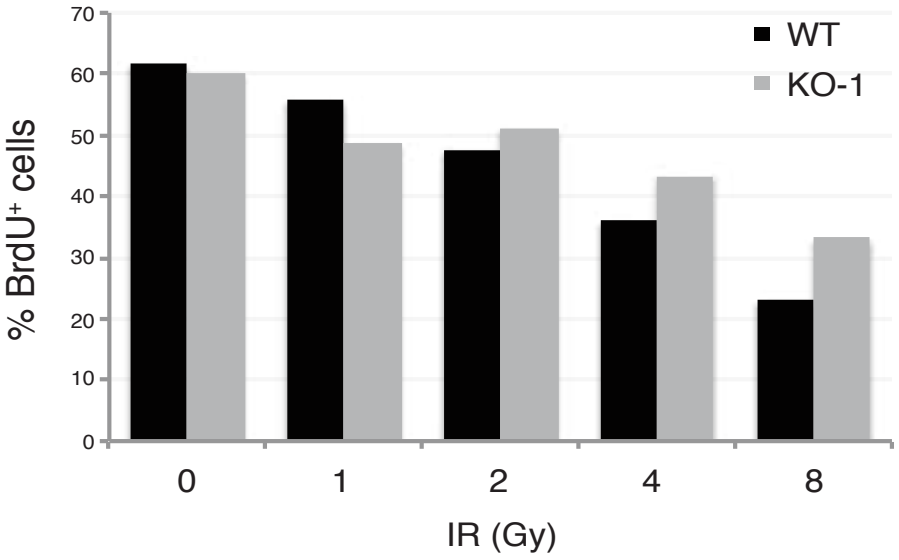
(F) Analysis of genomic instability in metaphases from PARPi-treated *Zmynd8*$^{-/-}$ iMEFs (n >= 42 metaphases analyzed per genotype).


A
WT
KO-1
KO-2
Zmynd8
Reconstituted
-
+
-
+
-
+
Irradiation (Gy)
0
1
2
5
10
B
Survival (%)
100
10
1
0
1
2
3
4
5
6
7
8
9
10
IR (Gy)
WT
Rc
KO-1
KO-2

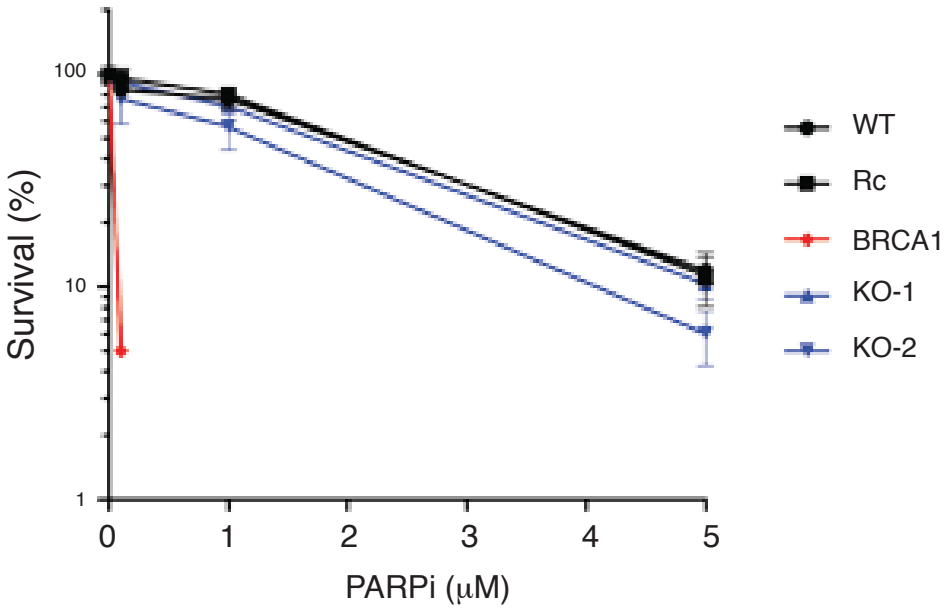
C


WT
KO-1
Proliferation 24h
1
0.5
0
Proliferation 60h
1
0.5
0
-/-
0
2
4
6
8
IR (Gy)

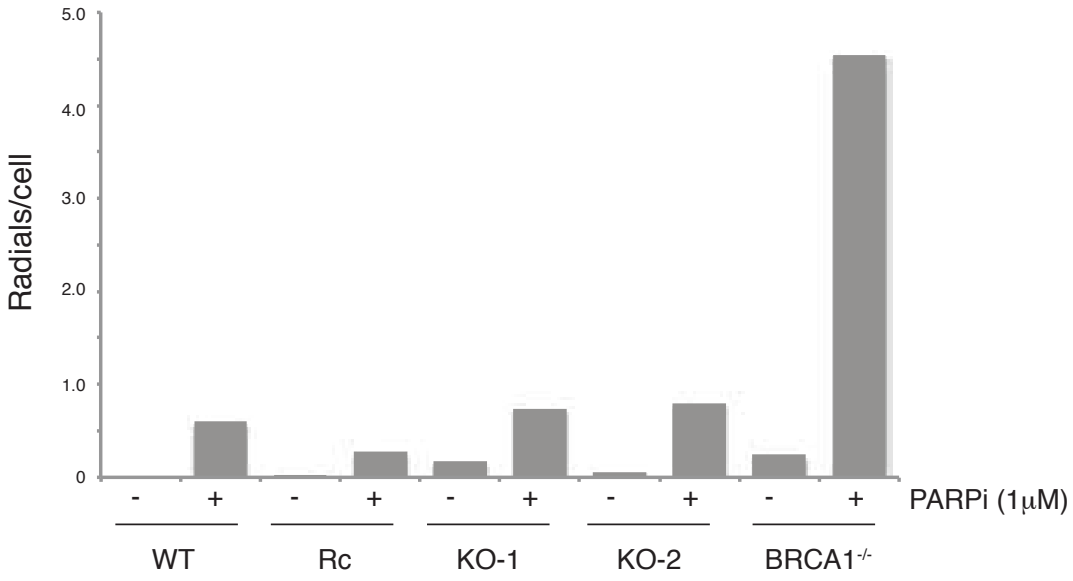
D


70
60
50
40
30
20
10
0
% BrdU+ cells
WT
KO-1
0
1
2
4
8
IR (Gy)

E


100
10
1
Survival (%)
0
1
2
3
4
5
PARPi (μM)
WT
Rc
BRCA1
KO-1
KO-2

F



5.3 DNA End Protection and CSR

CSR depends on NHEJ (Kotnis, Du, Liu, Popov, & Pan-Hammarström, 2009) and to a lesser extent MMEJ (Sfeir & Symington, 2015), which can serve as a backup pathway. RIF1 facilitates CSR by preventing end-resection at DSBs, thereby promoting NHEJ instead of HR (Symington, 2014). To assess whether Zmynd8 impacts CSR by altering end-resection, we set out to directly measure the degree of DNA end-resection at DSBs in *Zmynd8*$^{-/-}$ CH12 cells.

We created an assay to measure end-resection in intrachromosomal translocations directed by Cas9 (figure 19a). Two gRNAs direct Cas9 to create DSBs at the Rosa26 locus. Cells that have successfully cut at both targets are screened for the smaller, recombined locus by PCR. Sanger sequencing can then detect unilateral end resection at each DSB site. This schema has the advantage of being analogous to the intrachromosomal recombination in CSR.

Zmynd8$^{-/-}$ CH12 cells showed no increase in end resection compared to WT cells (figure 19b). Additionally, the KO's showed statistically less resection compared to the 53BP1$^{-/-}$ control (Bothmer, et al., 2010). There was no difference between the KO cell lines.

We adapted an AID-independent, Cas9 mediated, CSR assay (Le, et al., 2016), to assess whether Zmynd8 could affect CSR post DSB induction at the IgH locus (figure 19c). In the assay, CH12 cells are nucleofected with a Cas9 plasmid carrying two-gRNAs, which target 5' Sμ and 3' Sα. By measuring CSR, we can investigate the cell's ability to synapse the distal DSBs and to ligate them. We found that *Zmynd8*$^{-/-}$ CH12 cells showed no difference in CSR in the gRNA-CSR assay (figure 19d). Representative flow cytometry plots show that there is no CSR in the absence of the specific gRNAs (figure 19e). There was a reduction in CSR in Lig4$^{-/-}$ CH12 cells, which leads to impaired DSB ligation, and 53BP1 $^{-/-}$ CH12 cells, which impairs synapsis of distal DSBs.

Figure 19: Zmynd8 does not affect end-resection or Cas9 mediated CSR in CH12 cells

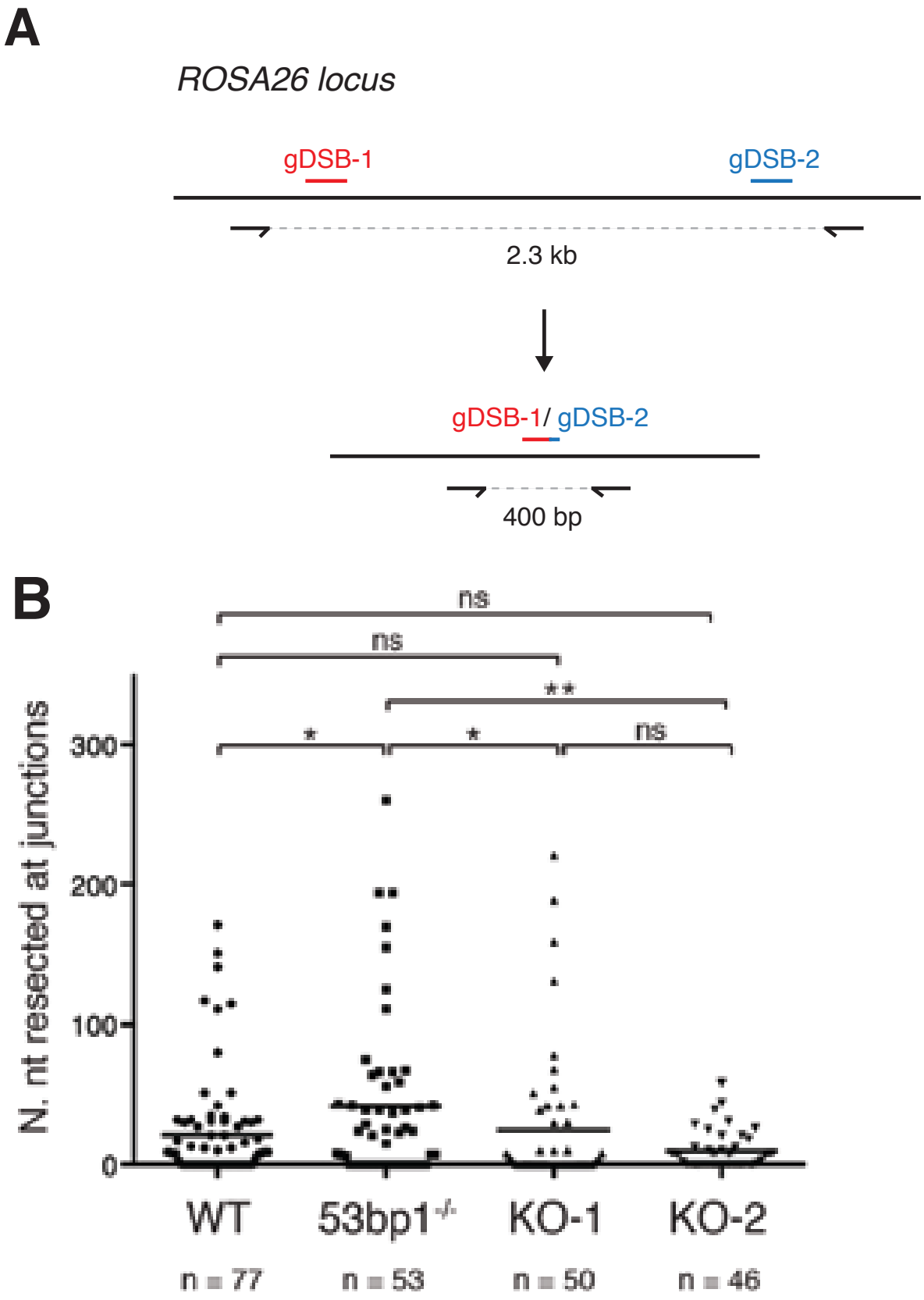
(A) Schematic representation of the end resection assay

(B) Dot plot showing resection in sequences from joining products of two CRISPR-Cas9-induced DSBs at the *ROSA26* locus in two *Zmynd8*$^{-/-}$ CH12 cell lines (KO-1 and KO-2). Each dot represents one junction product. Number of junctions analyzed per genotype is indicated below. p values were calculated with Mann-Whitney test. n.s.: not significant.

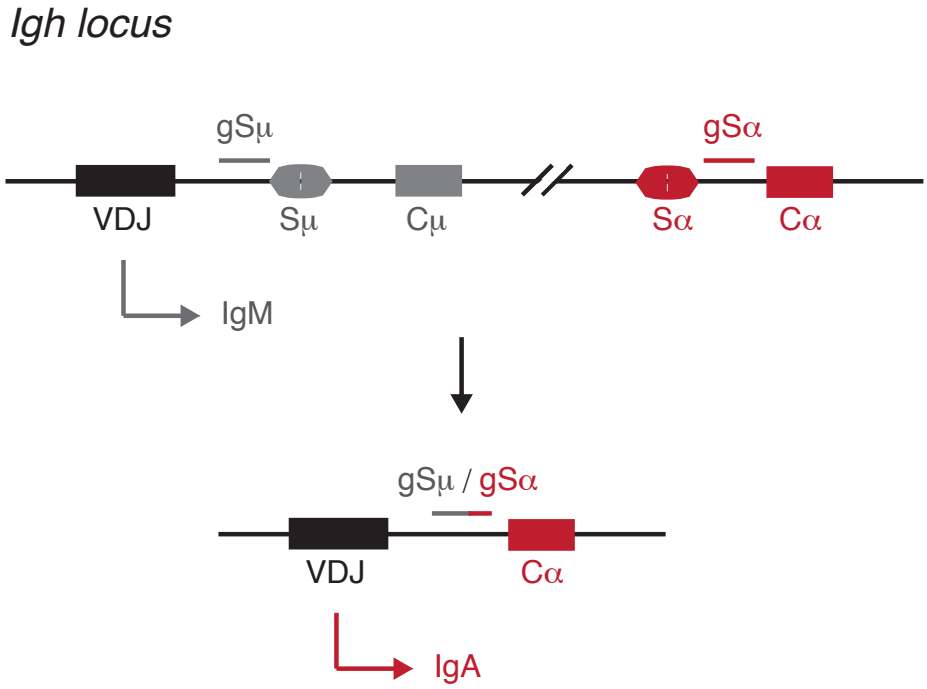
(C) Schematic representation of CRISPR-Cas9-induced CSR assay

(D) Summary dot plot for three independent experiments in the gRNA-CSR assay. Significant reductions are indicated.

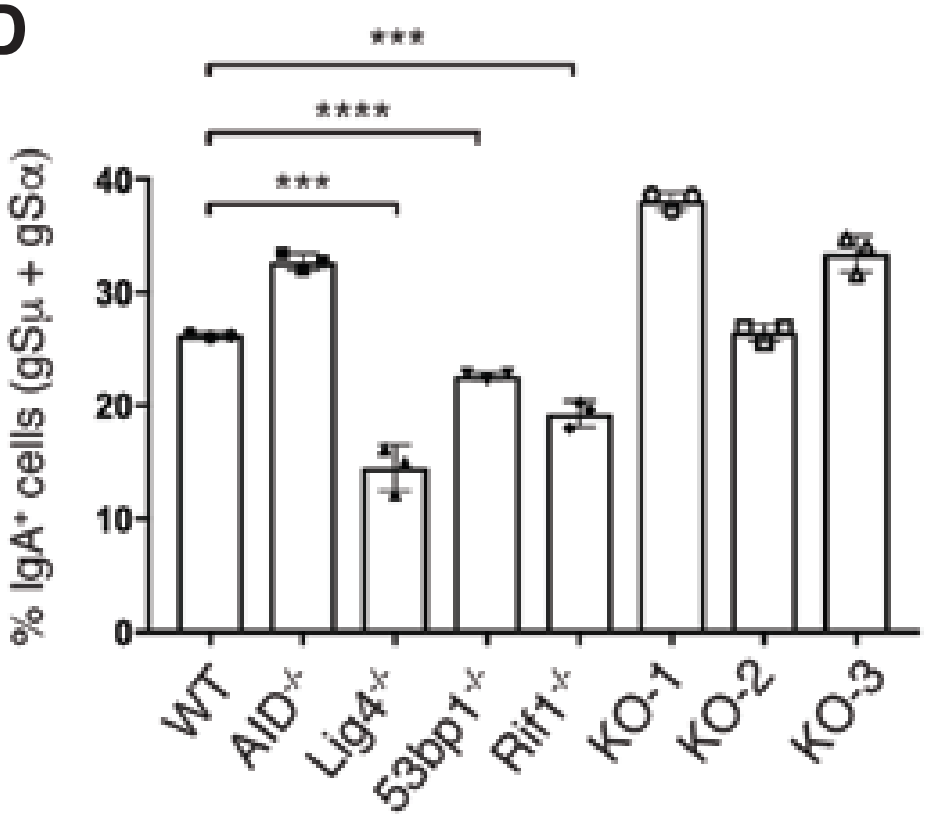
(E) Representative flow cytometry plots measuring CSR to IgA after electroporation of inactivated CH12 cells lines of the indicated genotypes with WT Cas9 and gRNAs against Random or Sα and Sμ sequences

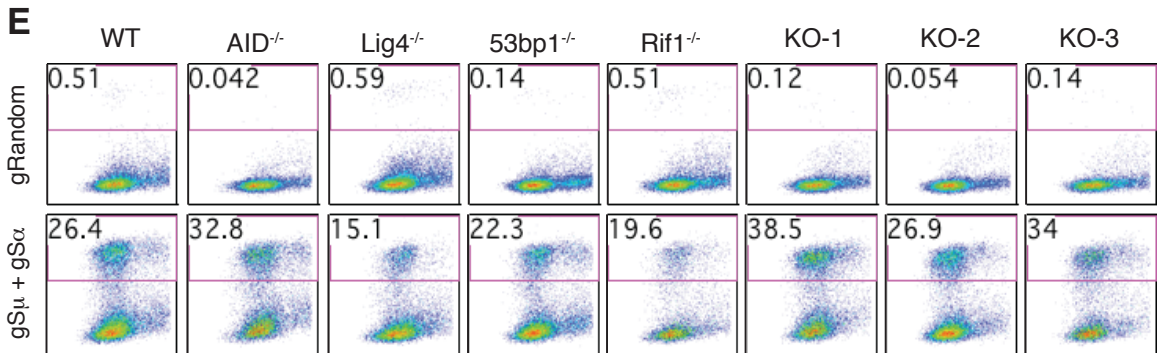

A
ROSA26 locus
gDSB-1
gDSB-2
2.3 kb
gDSB-1/ gDSB-2
400 bp
B
ns
ns
**
*
*
ns
N. nt resected at junctions
300
200
100
0
WT
53bp1-/-
KO-1
KO-2
n = 77
n = 53
n = 50
n = 46

C


Igh locus
gSμ
gSα
VDJ
Sμ
Cμ
Sα
Cα
IgM
gSμ / gSα
VDJ
Cα
IgA

D


***
****
***
% IgA+ cells (gSμ + gSα)
40
30
20
10
0
WT
AID-/-
Lig4-/-
53bp1-/-
Rif1-/-
KO-1
KO-2
KO-3


E
WT
AID-/-
Lig4-/-
53bp1-/-
Rif1-/-
KO-1
KO-2
KO-3
gRandom
gSμ + gSα
0.51
0.042
0.59
0.14
0.51
0.12
0.054
0.14
26.4
32.8
15.1
22.3
19.6
38.5
26.9
34

Chapter 6: ZMYND8 and Transcription

Because Zmynd8 does not affect CSR in CH12 cells by regulating AID expression, germline transcription, cell proliferation, or DNA repair, and in other contexts Zmynd8 has been reported to affect transcription (Shen, et al., 2016), (Gong, et al., 2015), (Li, et al., 2016), (Gong, Clouaire, Aguirrebengoa, Legube, & Miller, 2017), (Spruijt, et al., 2016) we chose to investigate how Zmynd8 deficiency influences the general transcriptional landscape in CH12 cells.

6.1 Zmynd8 and the Transcriptional Landscape

To characterize the transcriptional differences due to Zmynd8, we performed rRNA depleted RNA-seq. This method allows us to measure ncRNAs and any pre-polyA processed RNA as well as mature mRNAs (O'Neil, Glowatz, & Schlumpberger, 2013).

We compared CH12 Zmynd8-KO1 (KO1), Zmynd8-KO2 (KO2), WT, and clonally derived WT (WTc) cells under unactivated and activated conditions. Dendrogram hierarchical clustering of the samples (figure 20a) showed that WT and WTc clustered separated from KO-1 and KO-2. This indicates that the KO clones are transcriptionally similar, and the WTc is representative of the WT bulk. The next

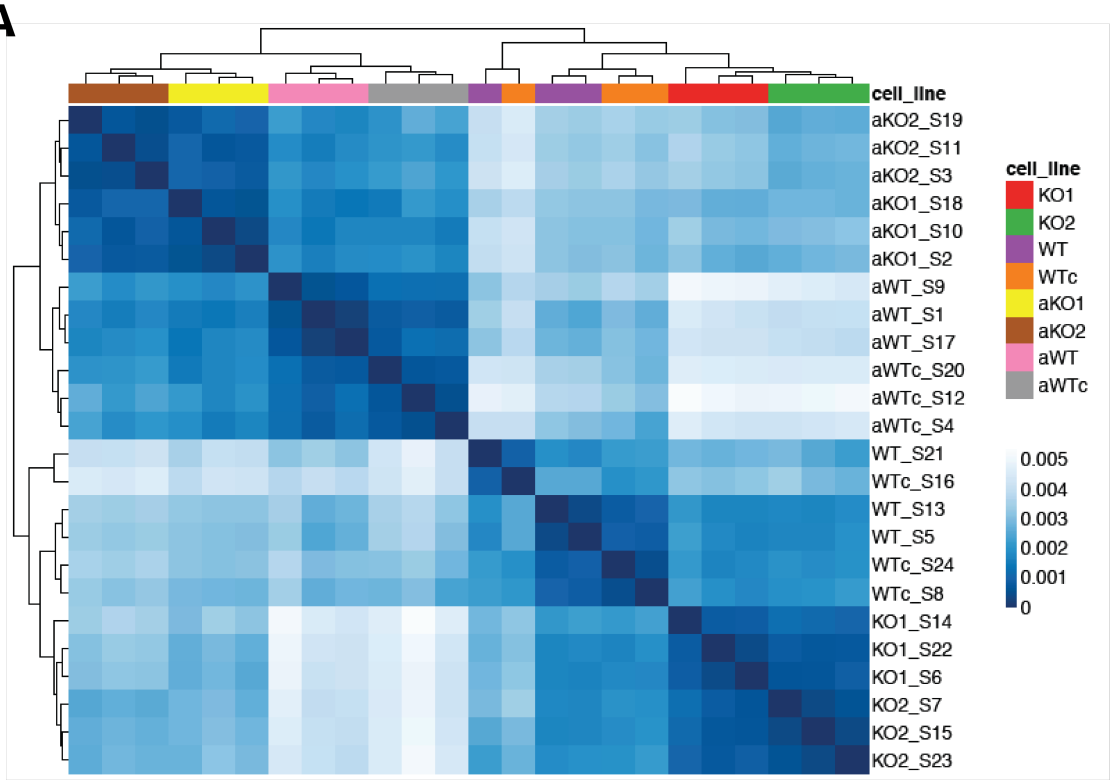
clustering level was based on activation or nonactivation conditions. This shows that stimulation for CSR causes greater transcriptional changes than the presence of Zmynd8. We could also see that samples clustered well together using principal component analysis according to Zmynd8 presence, and activation status (figure 20b). We checked if 64 known CSR factors were transcriptionally affected by Zmynd8 deletion in CH12 cells (figure 20c), but found no differences.

Figure 20: Zmynd8 and Transcription in CH12 Cells

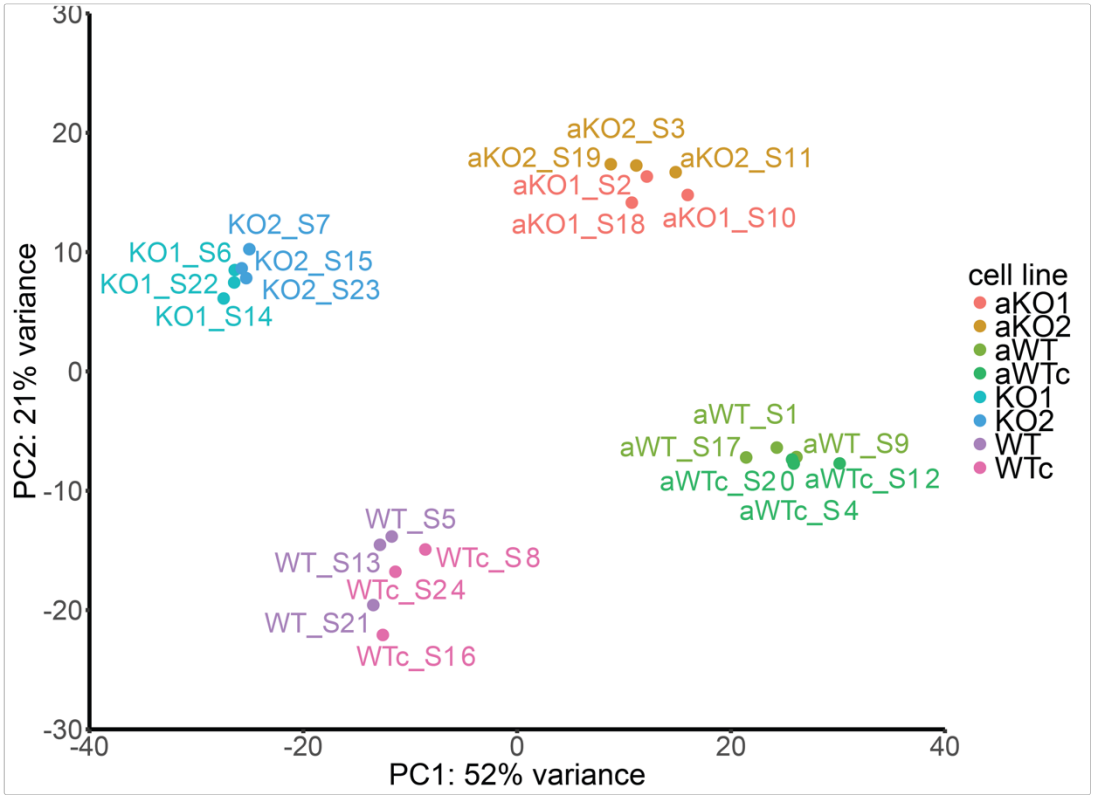
(A-B) Dendrogram **(A)** and Principal Component Analysis **(B)** for WT CH12 cells (bulk and one clonal derivative, WTc) and two *Zmynd8*$^{-/-}$ cell lines (KO1 and KO2) under either unactivated or activated conditions (prefix "a" indicates stimulation for 48h with αCD40, TGFβ and IL-4). RNA-Seq analysis was performed on three biological replicates per condition. Prepared sample libraries were run in duplicate across two sequencing lanes.

(C) Scatterplot for gene expression of known CSR factors in activated WT CH12 and *Zmynd8*$^{-/-}$ samples. Pairwise comparisons between WT and *Zmynd8*$^{-/-}$ cell lines are represented as different shapes, while CSR factors are denoted by color. Linear regression line shown in blue; identity line shown in grey.

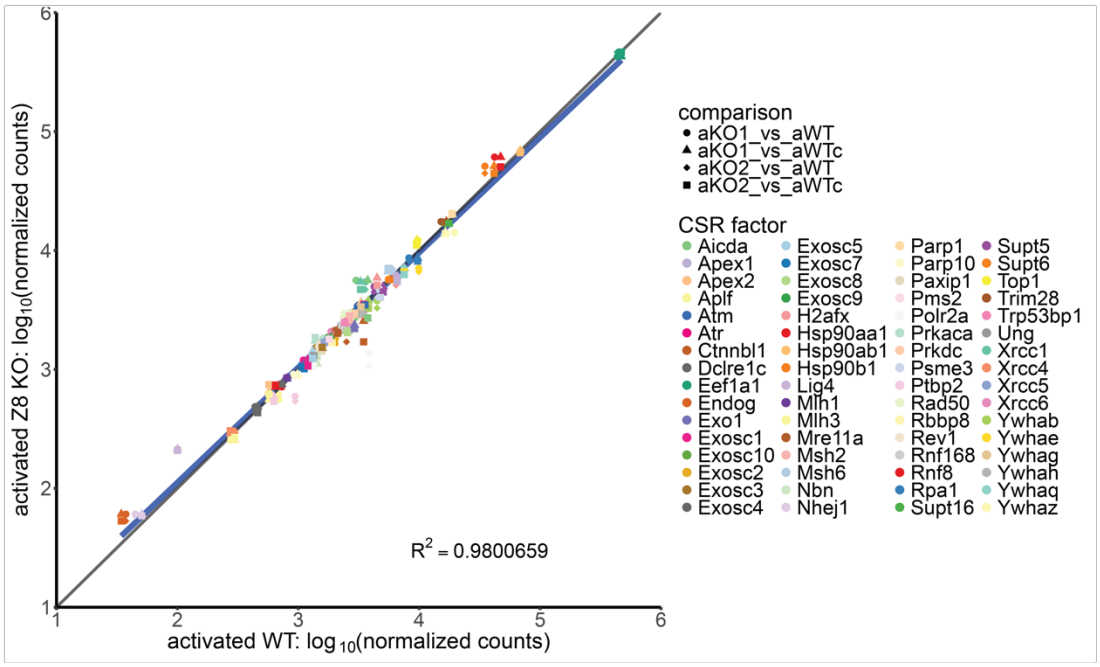
A


cell_line
aKO2_S19
aKO2_S11
aKO2_S3
aKO1_S18
aKO1_S10
aKO1_S2
aWT_S9
aWT_S1
aWT_S17
aWTc_S20
aWTc_S12
aWTc_S4
WT_S21
WTc_S16
WT_S13
WT_S5
WTc_S24
WTc_S8
KO1_S14
KO1_S22
KO1_S6
KO2_S7
KO2_S15
KO2_S23
KO2_S19
cell_line
KO1
KO2
WT
WTc
aKO1
aKO2
aWT
aWTc
0.005
0.004
0.003
0.002
0.001
0

B


PC2: 21% variance
PC1: 52% variance
30
20
10
0
-10
-20
-30
-40
-20
0
20
40
aKO2_S3
aKO2_S19
aKO2_S11
aKO1_S2
aKO1_S18
aKO1_S10
KO2_S7
KO1_S6
KO2_S15
KO1_S22
KO2_S23
KO1_S14
aWT_S1
aWT_S17
aWT_S9
aWTc_S20
aWTc_S12
aWTc_S4
WT_S5
WT_S13
WTc_S8
WTc_S24
WT_S21
WTc_S16
cell line
aKO1
aKO2
aWT
aWTc
KO1
KO2
WT
WTc

C


activated Z8 KO: log10(normalized counts)
activated WT: log10(normalized counts)
R2 = 0.9800659
comparison
aKO1_vs_aWT
aKO1_vs_aWTc
aKO2_vs_aWT
aKO2_vs_aWTc
CSR factor
Aicda
Apex1
Apex2
Aplf
Atm
Atr
Ctnnbl1
Dclre1c
Eef1a1
Endog
Exo1
Exosc1
Exosc10
Exosc2
Exosc3
Exosc4
Exosc5
Exosc7
Exosc8
Exosc9
H2afx
Hsp90aa1
Hsp90ab1
Hsp90b1
Lig4
Mlh1
Mlh3
Mre11a
Msh2
Msh6
Nbn
Nhej1
Parp1
Parp10
Paxip1
Pms2
Polr2a
Prkaca
Prkdc
Psme3
Ptbp2
Rad50
Rbbp8
Rev1
Rnf168
Rnf8
Rpa1
Supt16
Supt5
Supt6
Top1
Trim28
Trp53bp1
Ung
Xrcc1
Xrcc4
Xrcc5
Xrcc6
Ywhab
Ywhae
Ywhag
Ywhah
Ywhaq
Ywhaz

6.2 Zmynd8 and AID Targeting of Switch Regions

CSR requires efficient AID targeting of the switch regions preceding the Cμ and Cx (downstream heavy chain) region. AID preferentially deaminates at the switch regions, due to the sequence specificity of AID for RGYW motifs, the high concentrations of these motifs in the switch regions, protein-AID partner interactions, and transcriptional activity (Chandra, Bortnick, & Murre, 2015). Given that Zmynd8 deficiency does not alter CSR by lowering AID levels, or by affecting downstream repair processes, we chose to investigate if AID function at the switch region was affected by directly measuring the mutational load.

Although the repetitiveness of the core switch region makes sequencing difficult, the 5' end of Sμ can be assessed for AID-induced mutations (figure 21a) and is representative of AID targeting of the core switch region (Barreto, Reina-San-Martin, Ramiro, McBride, & Nussenzweig, 2003), (Xue, Rada, & Neuberger, 2006), (Guikema, et al., 2007), (Jeevan-Raj, et al., 2011).

Using mutational analysis by paired-end deep sequencing (Mut-PE-seq) (Robbiani, et al., 2015), (Wang, et al., 2016), we analyzed the mutational load at the 5' Sμ region in CH12 clones under activated and non-activated conditions. We found a statistically significant reduction in mutations in the Zmynd8 KO clonal cell lines (figure 21b). Increased mutation rates are not evenly distributed, but instead

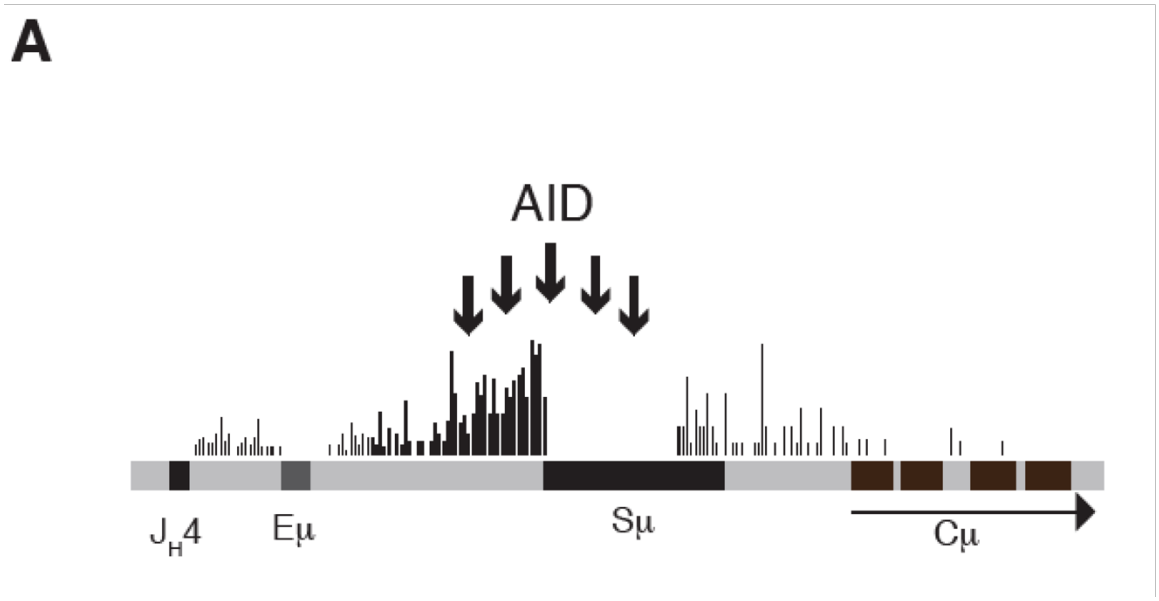
are concentrated at certain residues (figure 21c). These residues are RGYW motifs or other potential targets for AID.

Figure 21: AID Targeting of 5' Switch regions

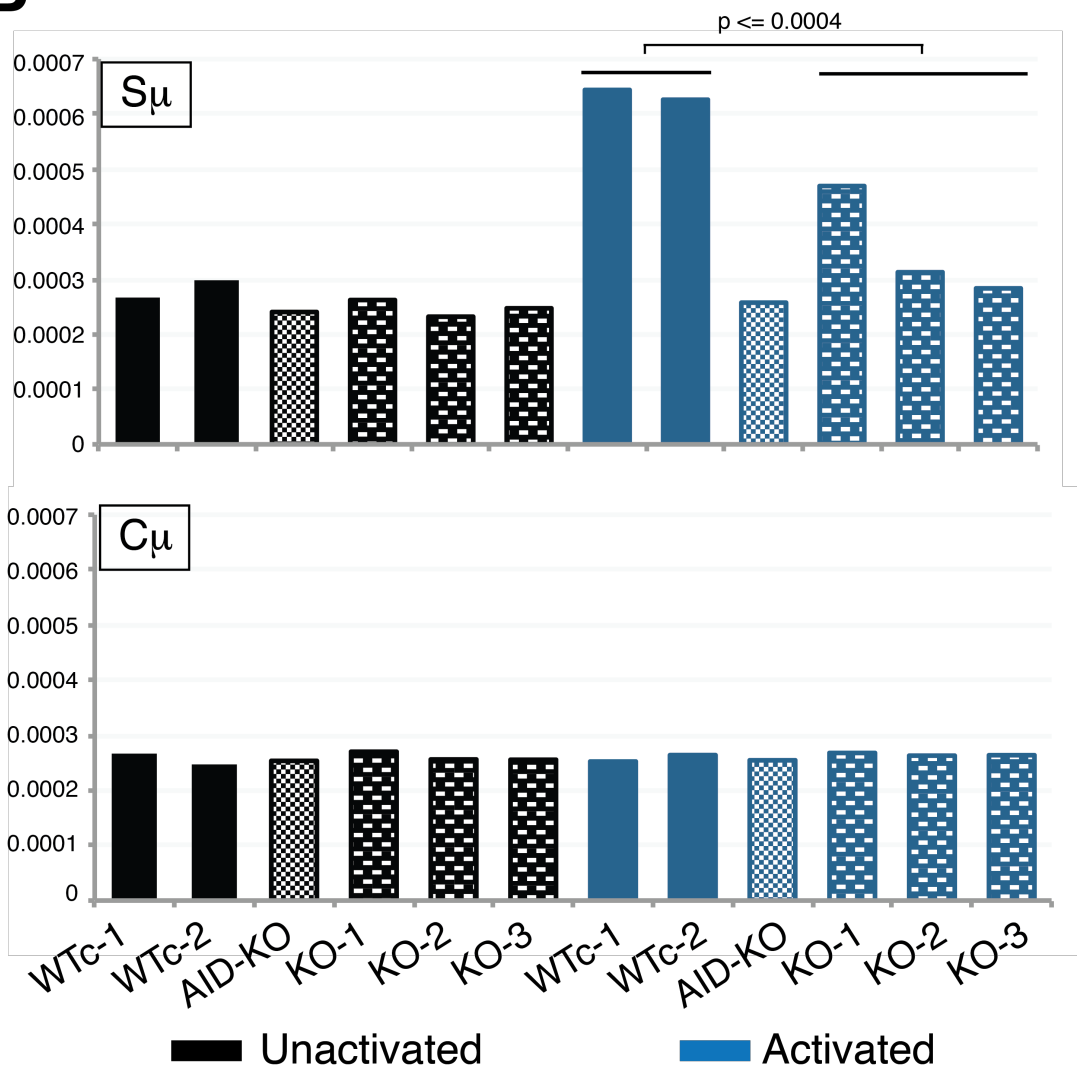
(A) Schematic representation of Eμ-Sμ-Cμ region with footprint of AID-induced mutations (adapted from Xue *J Exp Med* 2006).

(B) Chart depicting cumulative mutation frequencies as determined by MutPE-Seq at 5'-Sμ in gRNA targeted CH12 cells: two WT clones of CH12 (WTc-1, WTc-2), an $AID^{-/-}$ clone (AID-KO), and three $Zmynd8^{-/-}$ clones (KO-1, KO-2, KO-3). Cells were either left unactivated or activated for 48 h with IL-4, TGF-β, and CD40.

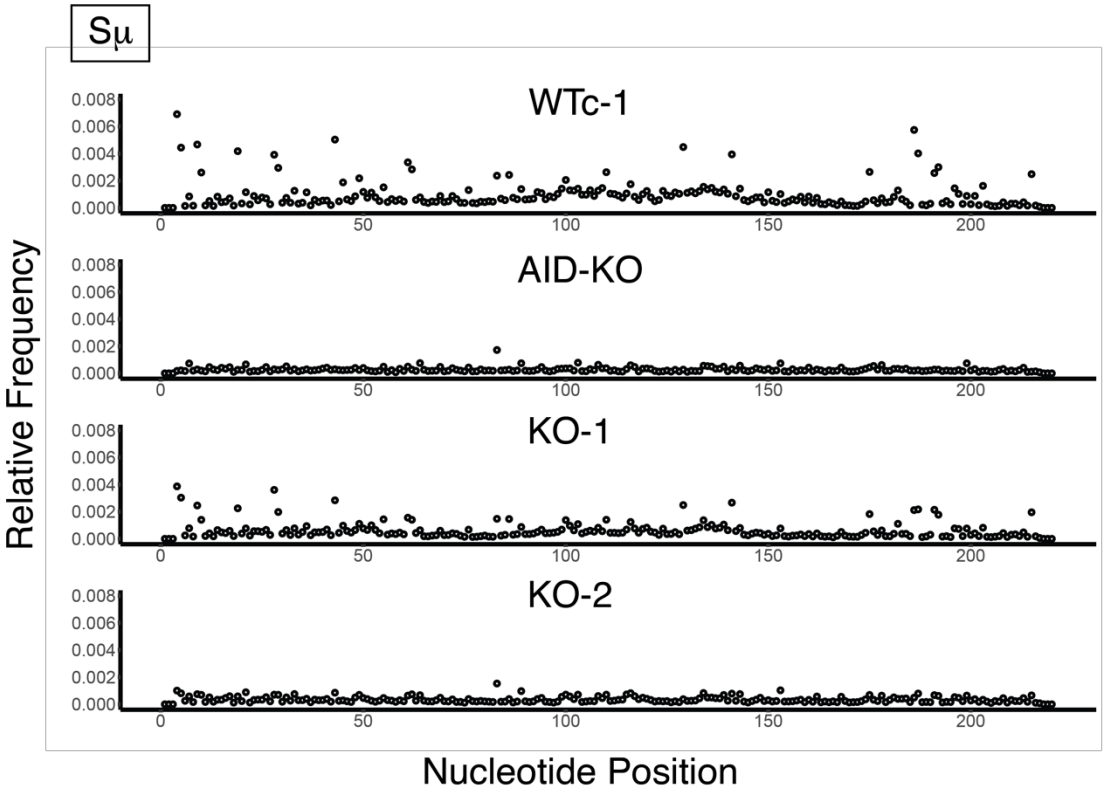
(C) Representative charts of the relative mutation frequency organized by nucleotide position in the measured 5'-Sμ region in Mut-PE-seq. CH12 clones for WT, $AID^{-/-}$, and two $Zmynd8^{-/-}$ KO's are shown.


A
AID
J_H4
Eμ
Sμ
Cμ

B


p <= 0.0004
Sμ
Cμ
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0
WTc-1
WTc-2
AID-KO
KO-1
KO-2
KO-3
WTc-1
WTc-2
AID-KO
KO-1
KO-2
KO-3
Unactivated
Activated

C


Sμ
WTc-1
AID-KO
KO-1
KO-2
0.008
0.006
0.004
0.002
0.000
0
50
100
150
200
Relative Frequency
Nucleotide Position

6.3 Zmynd8 Distribution in the IgH locus

Zmynd8 deficiency in CH12 cells leads to a decrease in AID targeting at Sμ. Possible explanations for this could be an interaction directly with AID, alterations in transcription at the IgH locus that affect AID targeting, or changes in epigenetic marks at the IgH locus. All of these possibilities would be mediated by Zmynd8 occupancy at the IgH locus.

To explore Zmynd8 occupancy at the IgH in CH12 cells, we performed chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-Seq). We found that ZMYND8 associates with the *Igh* locus in activated CH12 cells, and it binds the 5' Eμ and 3'RR (regulatory region) enhancers (Fig. 22a). Eμ and 3'RR, are located at 5' of Sμ and 3' of Cα, respectively (Birshtein, 2014), (Perlot & Alt, 2008), and the 3'RR is a prototypical super-enhancer (Pinaud et al., 2011).

Because Zmynd8 depletion has been previously shown to result in enhancer overactivation as measured by increased transcription of enhancer RNA (eRNA) (Shen, et al., 2016), we asked if Zmynd8 deletion affects transcription of the IgH enhancers. By reanalyzing the rRNA depleted, RNA-seq data in CH12 cells, we found that the enhancers HS1,2 and HS3B within the 3'RR exhibited higher levels of eRNA transcription (Fig. 22b) (Fig. 22c).

To isolate if the increased levels of eRNA represented an increase in transcription or a decrease in degradation, we assessed the occupancy of RNA PolII at the IgH locus via ChiP-seq. We found that consistent with increased eRNA production, there was more PolII loading at the 3'RR core enhancers, HS1,2 and HS3B (Fig. 22d, Fig 22e), in CH12 Zmynd8 KO cells.

Figure 22: Zmynd8 and Pol II Distribution in the IgH Locus

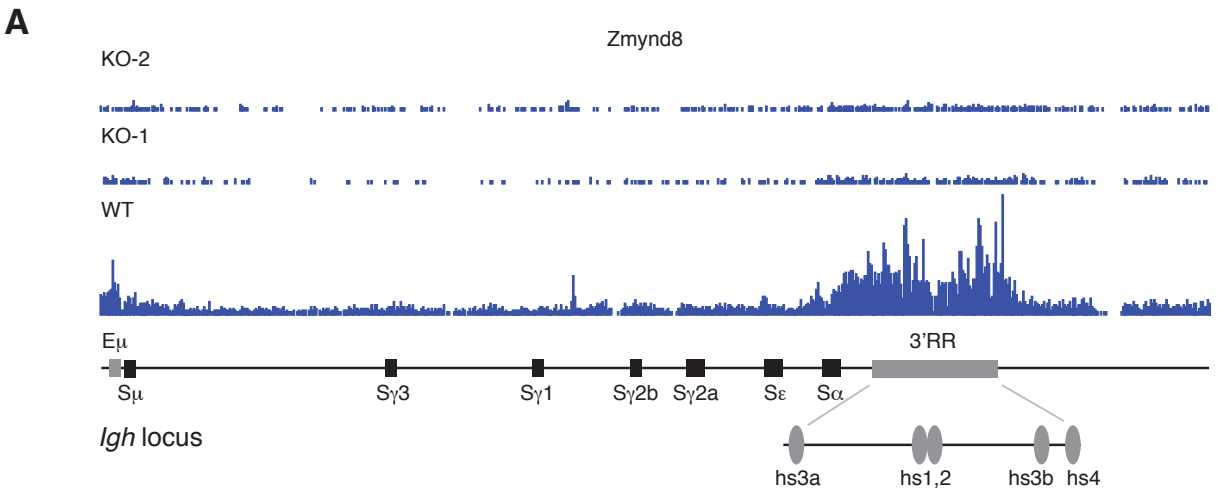
(A) ZMYND8 occupancy at the Igh locus in WT and Zmynd8 -/- CH12 cells. A schematic representation of the murine Igh locus (Chr12: 113,175,000-113,441,797) showing location of enhancers and S regions within the locus is represented below. Data are representative of two independent experiments.

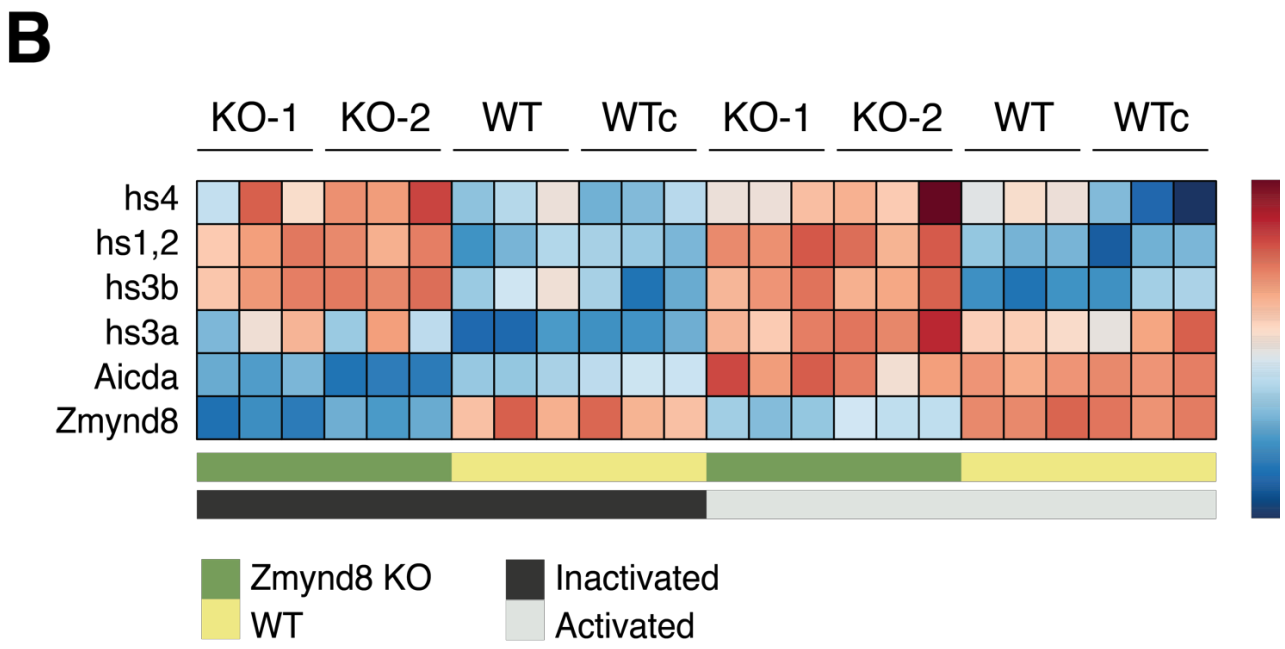
(B) Heat map showing *Igh* 3' RR enhancers (hs4; hs1,2; hs3b; hs3a) differential transcript expression as determined by RNA-Seq in WT cells (WT bulk and clonal derivative WTc) and two *Zmynd8*$^{-/-}$ CH12 clones. Cell were either left unactivated or stimulated for 48h with αCD40, TGFβ and IL-4. Expression counts are row-normalized by z-score. Three independent RNA-Seq replicates per sample are shown

(C) Bar graphs depicting relative transcript levels at hs4, hs1,2, hs3b, hs3a.

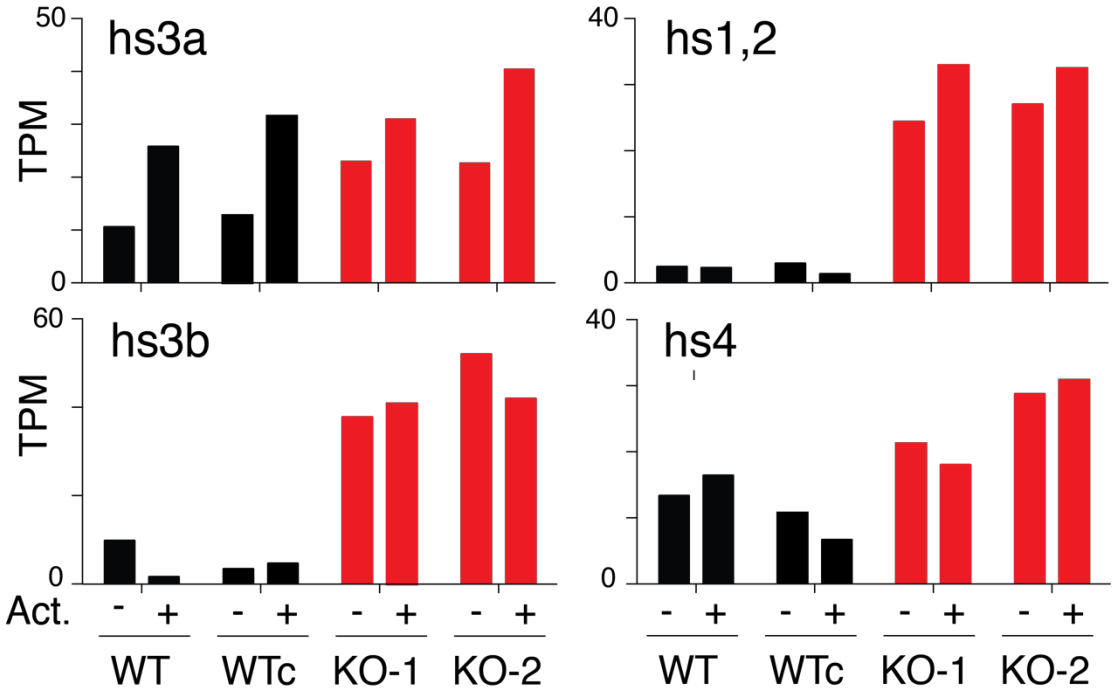
(D) Top: RNA Pol II loading at the *Igh* locus (Chr12: 113,175,000-113,441,797) in activated WT and *Zmynd8*$^{-/-}$ CH12 cells. Bottom: ChIP-Seq tracks overlay at regions encompassing Eα and 3'RR enhancers. Regions zoomed-in in the insets were defined based on the location of ZMYND8 peaks from Zmynd8 ChIP-Seq

(E) RNA Pol II loading quantification at regions highlighted within the insets in panel A. Numbers above columns represent the ratio of RNA Pol II loading in WT to loading in knock-out cells

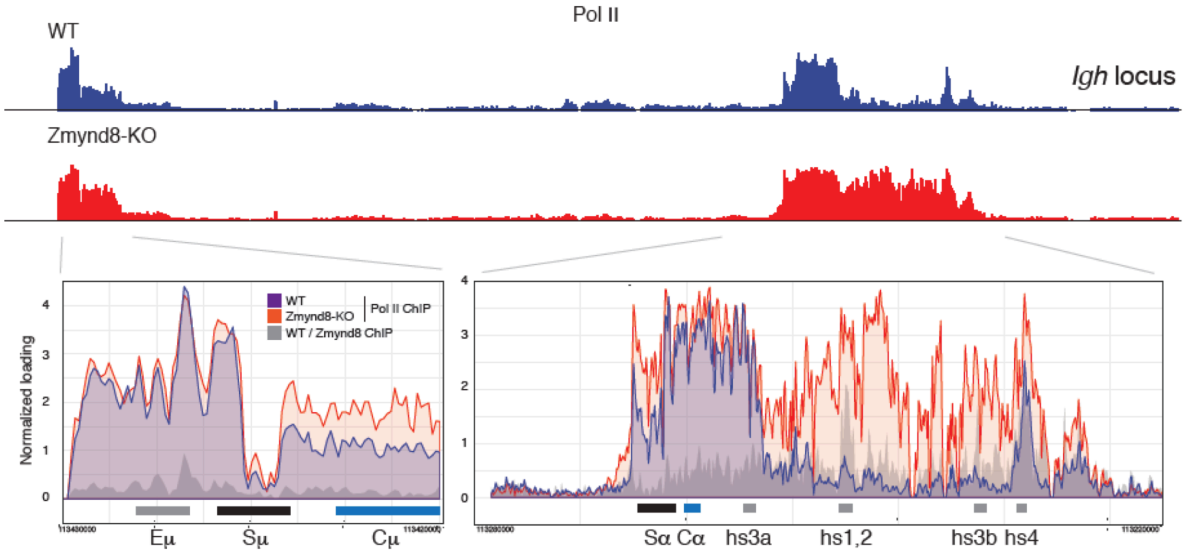

A
Zmynd8
KO-2
KO-1
WT
Eμ
3'RR
Sμ
Sγ3
Sγ1
Sγ2b
Sγ2a
Sε
Sα
Igh locus
hs3a
hs1,2
hs3b
hs4


B
KO-1
KO-2
WT
WTc
KO-1
KO-2
WT
WTc
hs4
hs1,2
hs3b
hs3a
Aicda
Zmynd8
Zmynd8 KO
WT
Inactivated
Activated

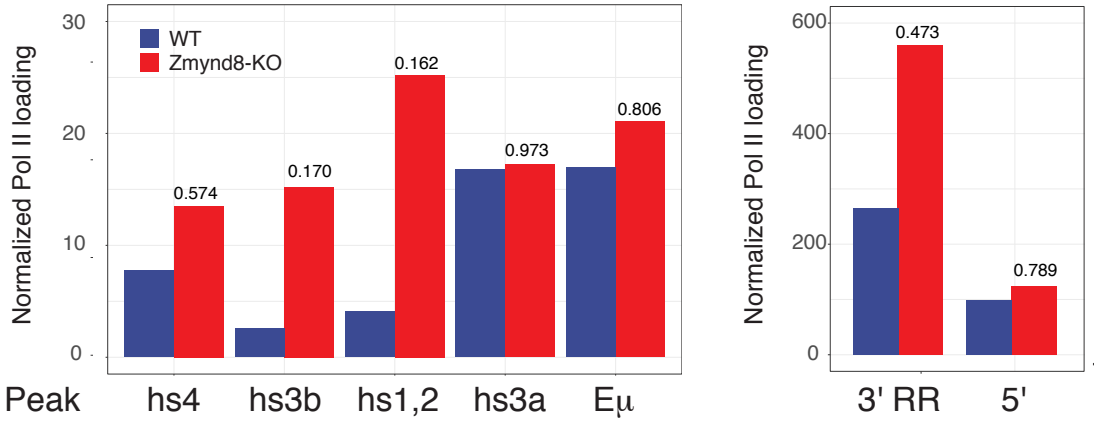
C


hs3a
hs1,2
hs3b
hs4
TPM
50
40
60
0
Act.
-
+
WT
WTc
KO-1
KO-2

D


Pol II
WT
Zmynd8-KO
Igh locus
WT
Zmynd8-KO
WT / Zmynd8 ChIP
Pol II ChIP
Normalized loading
Eμ
Sμ
Cμ
Sα Cα
hs3a
hs1,2
hs3b hs4

E


Normalized Pol II loading
30
20
10
0
WT
Zmynd8-KO
0.574
0.170
0.162
0.973
0.806
Peak
hs4
hs3b
hs1,2
hs3a
Eμ
Normalized Pol II loading
600
400
200
0
0.473
0.789
3' RR
5'

6.4 Genomic Distribution of Zmynd8

Zmynd8 is a histone mark reader, which localizes to promoters and enhancers in several cell lines. Zmynd8 forms an interaction network with RNA Polymerase II (Pol II) (Malovannaya, et al., 2011). Specifically, Zmynd8 binds to Pol II in a DNA dependent manner (Adhikary, et al., 2016). Additionally, Zmynd8 co-immunoprecipitates with the proximal-paused form of Poll II (CTD S5P), and not the elongating form of pol II (CTD S2P) (Jonkers & Lis, 2015) (Figure 23a). Consistent with this observation, Zmynd8 primarily localizes to the promoter regions of genes in the DU145 prostate cancer cell line (Li, et al., 2016), and we confirmed the Zmynd8 enrichment at promoter regions (57.49%), but also distal intergenic regions (17.85%) in CH12 cells (figure 23b). Taken together with Zmynd8's function in transcriptional alteration (Li, et al., 2016), transcriptional repression (Poleshko, et al., 2010), (Shen, et al., 2016), (Spruijt, et al., 2016), (Zeng, Kong, Li, & Mao, 2010), and it's association with the histone K14Ac marker (Savitsky, et al., 2016), which indicates actively transcribing genes (Agalioti, Chen, & Thanos, 2002) [specifically H3K14ac], (Eberharter & Becker, 2002) [histone acetylation broadly] we postulated that Zmynd8 interacts primarily with promoters and enhancers genome-wide.

To test this, we compared the occupancy of Zmynd8 with actively transcribing genes (marked by H3K4me3) and active enhancers (marked by H3K4me1, and

H3K27ac) (Calo & Wysocka, 2013). 72.27% of Zmynd8 peaks co-localize with H3K4me3 and H3K4me1 (figure 23c). When we compared Zmynd8 peaks with active enhancers (H3K27ac, and H3K4me1), 84.46% of the active enhancers also co-localized with Zmynd8 (figure 23d). We concluded that the Zmynd8 binding of the enhancer Eμ and super-enhancer 3'RR in CH12 cells is representative of genome-wide Zmynd8 binding.

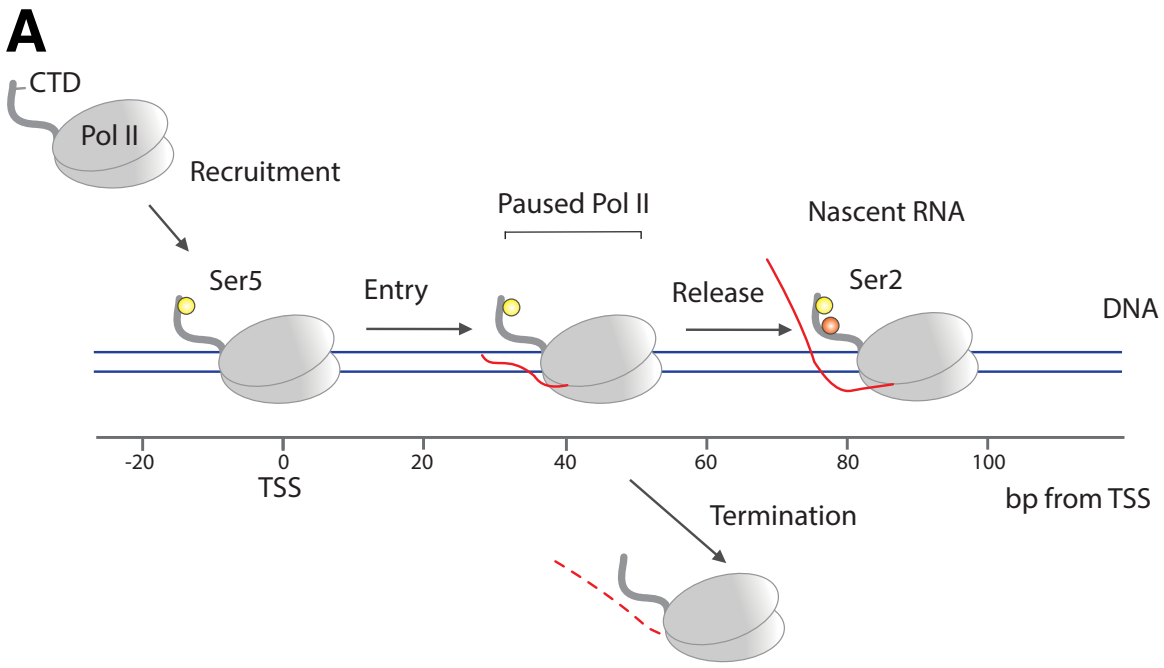
Figure 23: Genomic Distribution of Zmynd8

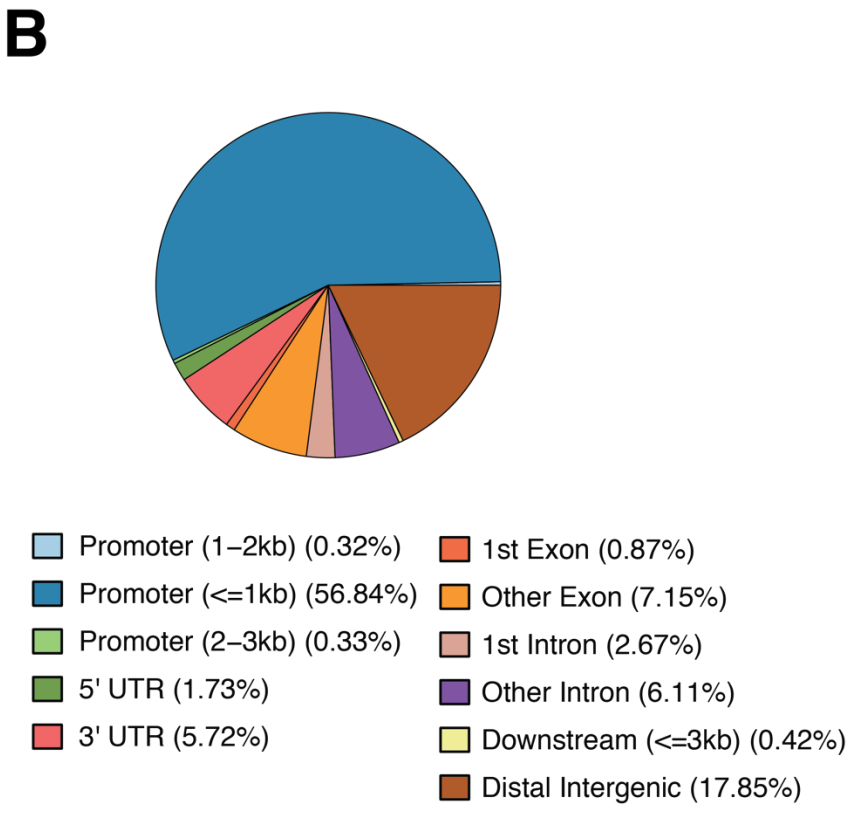
(A) Schema of c-terminal domain phosphorylation of RNA polymerase II (adapted from *Jonkers and Lis* Nature Reviews 2015). The RNA Pol II complex is recruited to the transcription start site (TSS) of genes. A key regulatory step is the ser-5 phosphorylation at which point the Pol II is paused at the promoter proximal region. Upon ser-2 phosphorylation, PolI II is released to begin productive elongation.

(B) Genomic distribution of ZMYND8 ChIP-Seq peaks in CH12 cells. Data are representative of two independent experiments.

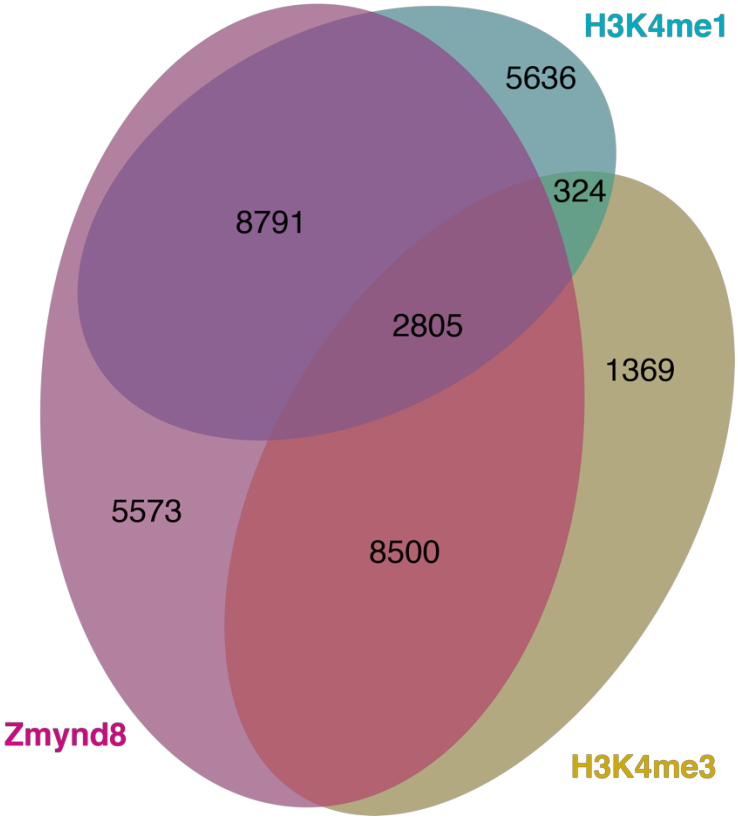
(C) Venn Diagram of the overlap between Zmynd8, H3K4me3 and H3K4me1 peaks.

(D) Venn Diagram of the overlap between Zmynd8, H3K27ac and H3K4me1 peaks. Note: the area with 3979 and 732 represents the overlap between H3K27ac and H3K4me1 peaks and corresponds with enhancers.

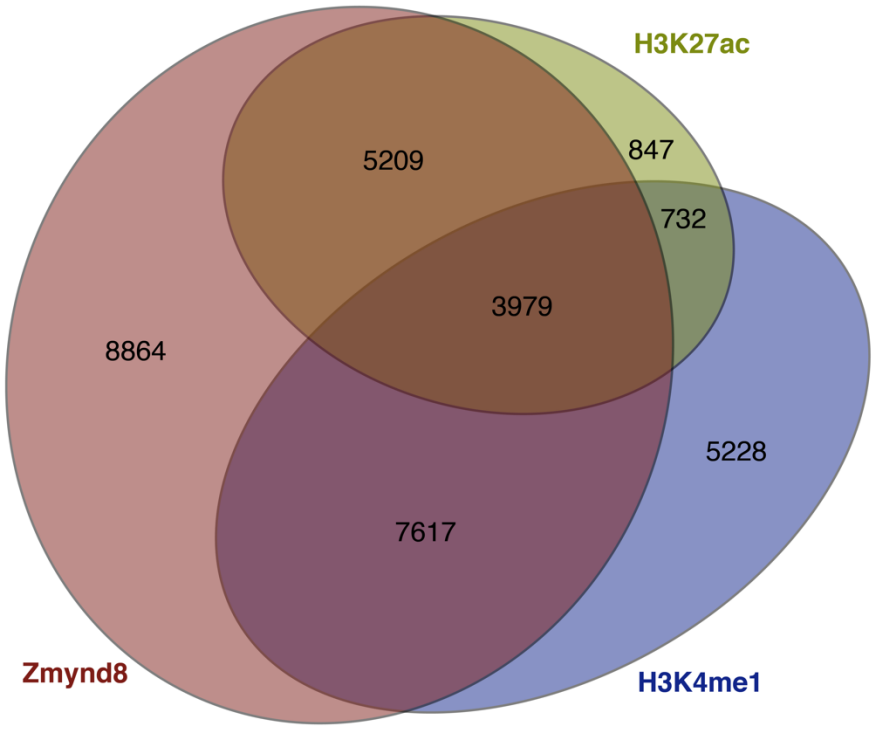

A
CTD
Pol II
Recruitment
Paused Pol II
Nascent RNA
Ser5
Entry
Release
Ser2
DNA
-20
0
20
40
60
80
100
TSS
bp from TSS
Termination


B
Promoter (1–2kb) (0.32%)
Promoter (<=1kb) (56.84%)
Promoter (2–3kb) (0.33%)
5' UTR (1.73%)
3' UTR (5.72%)
1st Exon (0.87%)
Other Exon (7.15%)
1st Intron (2.67%)
Other Intron (6.11%)
Downstream (<=3kb) (0.42%)
Distal Intergenic (17.85%)

C


H3K4me1
5636
324
8791
2805
1369
5573
8500
Zmynd8
H3K4me3

D


H3K27ac
5209
847
732
3979
8864
5228
7617
Zmynd8
H3K4me1

Chapter 7: ZMYND8 in Primary B Cells

7.1 Zmynd8 Mouse Generation and Characterization

To continue the investigation of CSR and Zmynd8 we decided to generate Zmynd8 deficient primary B-cells. Zmynd8$^{-/-}$ mice were previously reported to be embryonic lethal and heterozygotes had several altered phenotypes including decreased bodyweight and decreased caudal vertebrae number. The European mouse mutant cell repository (EUCOMM) had frozen ES cells with a potential conditional knockout of Zmynd8 (Zmynd8tm1a(EUCOMM)Wtsi) (Gene: Zmynd8, n.d.). By first crossing the reconstituted mice with a FLPeR (Flipper) mice (Farley, Soriano, Steffen, & Dymecki, 2000) the two loxP sites flanking the second exon of Zmynd8 (Figure 24a) could be exposed and the conditional knockout generated. To selectively delete Zmynd8 in B-cells, we bred the Zmynd8$^{f/f}$ mice to CD19$^{cre/cre}$ mice (Rickert, Roes, & Rajewsky, 1997). Zmynd8 appeared to be deleted in primary B-cells after 72 h stimulation with LPS and IL-4 (figure 24b). CD19 expression initiates at the pro-B cell stage and therefore B cells in the Zmynd8$^{f/f}$ CD19$^{cre/+}$ develop without Zmynd8. The lack of Zmynd8 does not affect B cell development (figure 24c –representative flow cytometry plots; figure 724d –summary experiments).

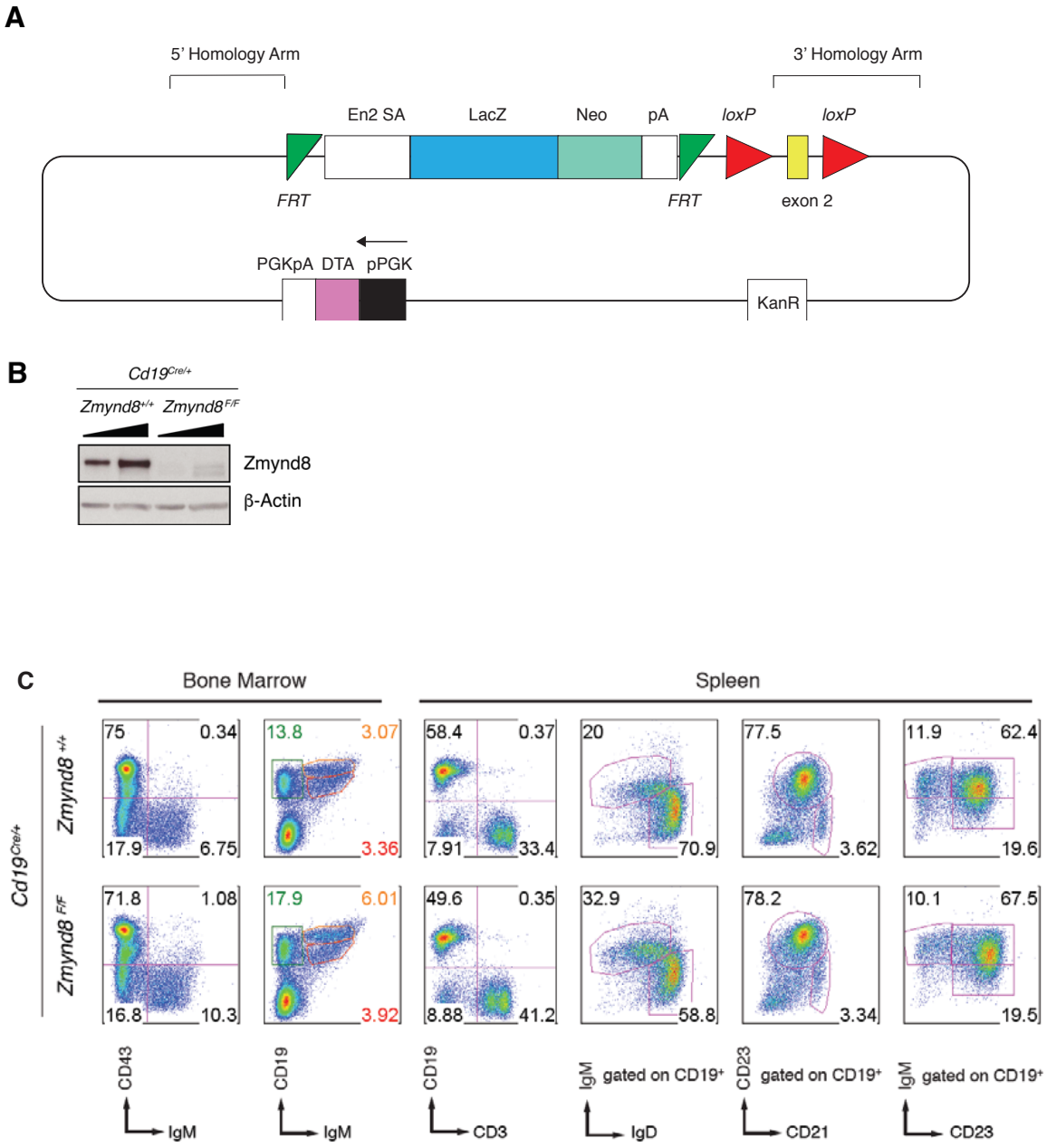
Figure 24: Zmynd8 Mouse Generation and Characterization

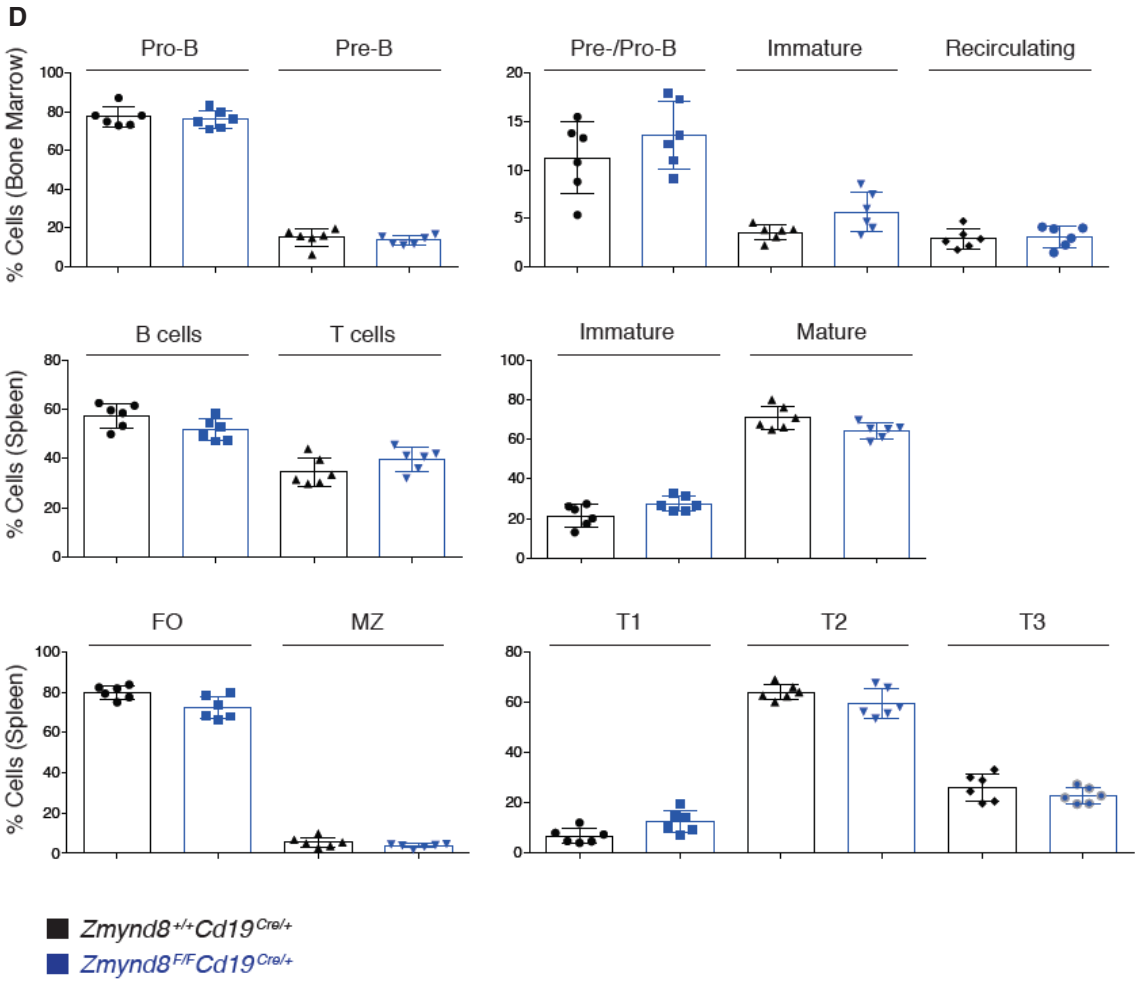
(A) Vector map of KO first allele (reporter-tagged insertion with conditional potential). Adapted from IMPC (international mouse phenotype consortium).

(B) Western blot analysis of whole cell extracts from *Cd19*$^{Cre/+}$ and *Zmynd8*$^{F/F}$*Cd19*$^{Cre/+}$ B cells 72 h after stimulation with LPS and IL-4. Triangles indicate threefold dilution. Data are representative of three independent experiments.

(C) Representative flow cytometry analysis of lymphoid tissues from *Cd19*$^{Cre/+}$ and *Zmynd8*$^{F/F}$*Cd19*$^{Cre/+}$ mice.

(D) Summary graphs for six mice per genotype.


A
5' Homology Arm
3' Homology Arm
En2 SA
LacZ
Neo
pA
loxP
loxP
FRT
FRT
exon 2
PGKpA
DTA
pPGK
KanR
B
Cd19Cre/+
Zmynd8+/+
Zmynd8F/F
Zmynd8
β-Actin
C
Bone Marrow
Spleen
Cd19Cre/+
Zmynd8 +/+
Zmynd8 F/F
75
0.34
17.9
6.75
13.8
3.07
3.36
58.4
0.37
7.91
33.4
20
70.9
77.5
3.62
11.9
62.4
19.6
71.8
1.08
16.8
10.3
17.9
6.01
3.92
49.6
0.35
8.88
41.2
32.9
58.8
78.2
3.34
10.1
67.5
19.5
CD43
IgM
CD19
IgM
CD19
CD3
IgM
gated on CD19+
IgD
CD23
gated on CD19+
CD21
IgM
gated on CD19+
CD23


D
Pro-B
Pre-B
Pre-/Pro-B
Immature
Recirculating
% Cells (Bone Marrow)
B cells
T cells
Immature
Mature
% Cells (Spleen)
FO
MZ
T1
T2
T3
% Cells (Spleen)
Zmynd8+/+Cd19Cre/+
Zmynd8F/FCd19Cre/+

7.2 ZMYND8 and CSR in Primary B Cells

In chapter 7.1 we showed that Zmynd8$^{f/f}$ mice can successfully delete Zmynd8 protein in B cells under the CD19 cre allele. Like the effect in CH12 cells, Zmynd8 deletion in primary B cells also leads to a reduction in CSR (figure 25a). This effect persists at activation times from 48 h to 96 h (figure 25b). We then confirmed that that the CSR reduction is not due to altered cell proliferation as Zmynd8 deficient B cells had a comparable number of cell divisions (figure 25c), but each cell division had fewer switched cells (figure 25d). The expression of AID was unchanged and thus also not responsible for the reduction in CSR (Figure 25e). GLTμ levels were consistent between WT and Zmynd8$^{f/f}$ mice (Figure 25f). Unlike the consistent downstream accepter GLT seen in CH12 cells, we found that the acceptor GLTϒ$_1$ or LPS and IL-4 stimulation in primary B cells was reduced (Figure 25g).

Figure 25: ZMYND8-deficiency in primary B lymphocytes impairs CSR by defective Igγ1 GLT

(A) Representative flow cytometry plots measuring CSR to IgG1 timecourse after stimulation of B lymphocytes with LPS and IL-4. Right: Summary dot plot for three mice per genotype.

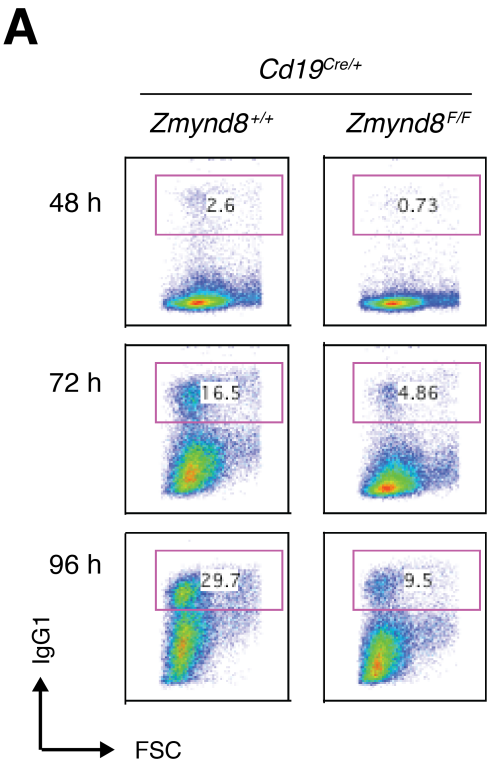
(B) Summary dot plot for three mice per genotype.

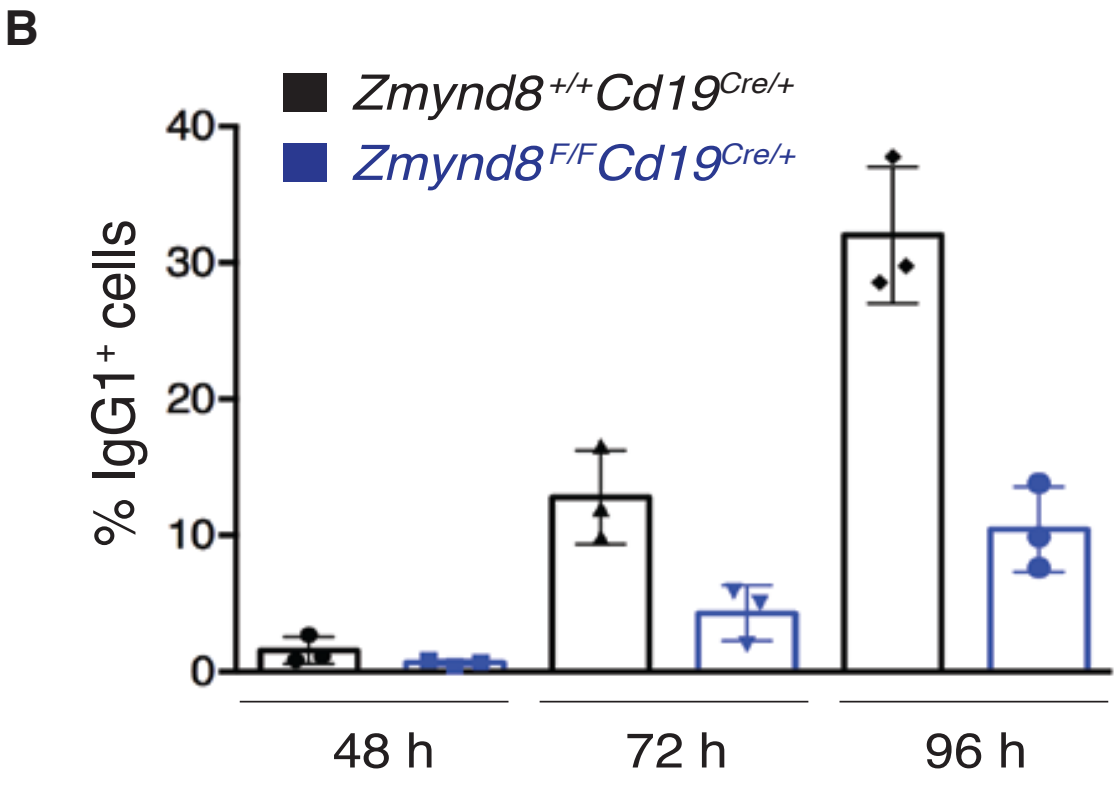
(C) Proliferation analysis of primary cultures of *Cd19*$^{Cre/+}$ and *Zmynd8*$^{F/F}$*Cd19*$^{Cre/+}$ B lymphocytes by CellTrace Violet dilution. Data are representative of three independent mice per genotype

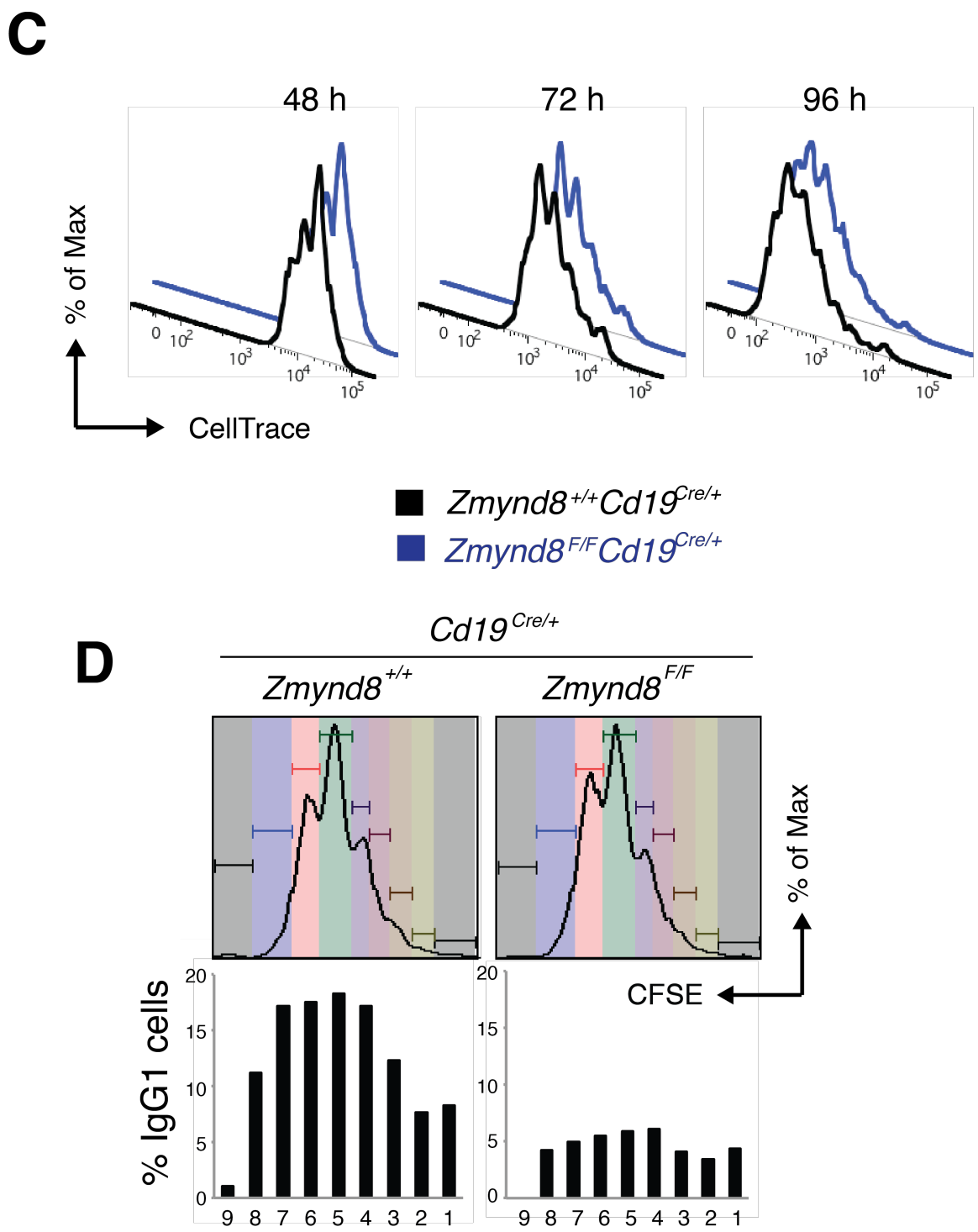
(D) Representative flow cytometry analysis showing the percentage of IgG1^{+} cells per cell division in primary cultures of *Cd19*$^{Cre/+}$ and *Zmynd8*$^{F/F}$*Cd19*$^{Cre/+}$ splenocytes stimulated with LPS and IL-4 for 72 h. Cell division as measured by CFSE dye dilution is shown on top.

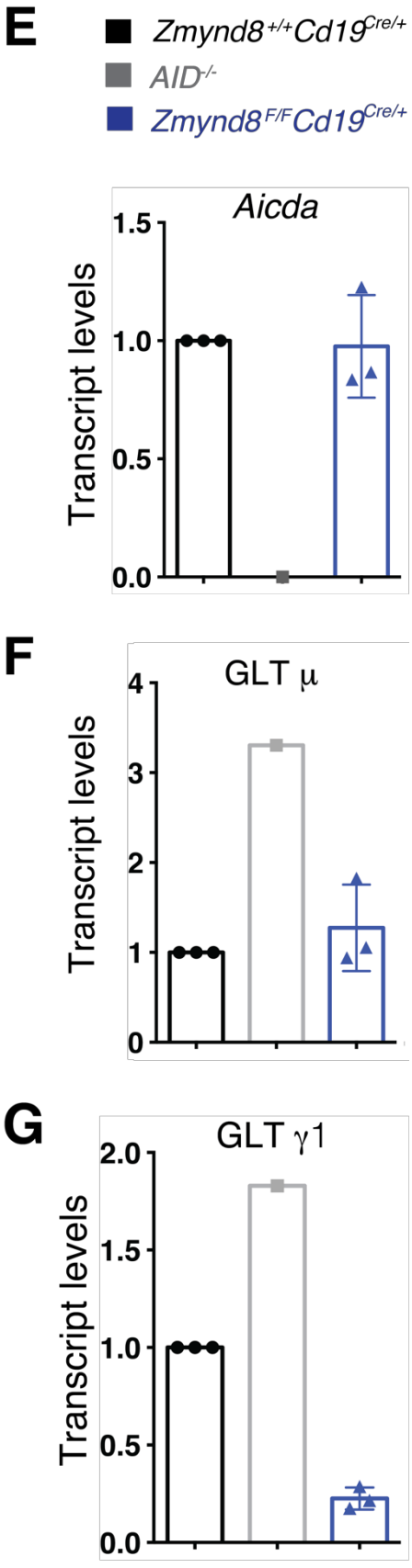
(E) qPCR analysis for *Aicda* mRNA levels in B cells stimulated for 48 h with LPS and IL-4. The data summarize three mice per genotype (error bars represent SD). *Cd19*$^{Cre/+}$ was assigned an arbitrary value of 1.0.

(F-G) qPCR analysis for Igμ (F) and Igγ1 (G) germline transcript (GLT) levels in B cells stimulated for 48 h with LPS and IL-4. The data summarize three mice per genotype (error bars represent SD). Data was compared to *Cd19*$^{Cre/+}$ which was set at a value of 1.0.


A
Cd19Cre/+
Zmynd8+/+
Zmynd8F/F
48 h
72 h
96 h
2.6
0.73
16.5
4.86
29.7
9.5
IgG1
FSC


B
Zmynd8+/+Cd19Cre/+
Zmynd8F/FCd19Cre/+
% IgG1+ cells
40
30
20
10
0
48 h
72 h
96 h


C
48 h
72 h
96 h
% of Max
CellTrace
Zmynd8+/+Cd19Cre/+
Zmynd8F/FCd19Cre/+
D
Cd19Cre/+
Zmynd8+/+
Zmynd8F/F
% of Max
CFSE
% IgG1 cells
20
15
10
5
0
9 8 7 6 5 4 3 2 1


E
Zmynd8+/+Cd19Cre/+
AID-/-
Zmynd8F/FCd19Cre/+
Aicda
Transcript levels
1.5
1.0
0.5
0.0
F
GLT μ
Transcript levels
4
3
2
1
0
G
GLT γ1
Transcript levels
2.0
1.5
1.0
0.5
0.0

7.3 ZMYND8 and AID Targeting in Primary B Cells

We sought to repeat the analysis from chapter 6.2 in primary B cells. While we did find WT (*CD19*$^{Cre/+}$)B cells displayed a 4-fold increase in mutations at the 5' end of Sμ when compared against *AID*$^{-/-}$ B cells (Figure 26a), there was no difference between the WT *Zmynd8*$^{F/F}$*Cd19*$^{Cre/+}$ cells at 72 h or 96 h post activation.

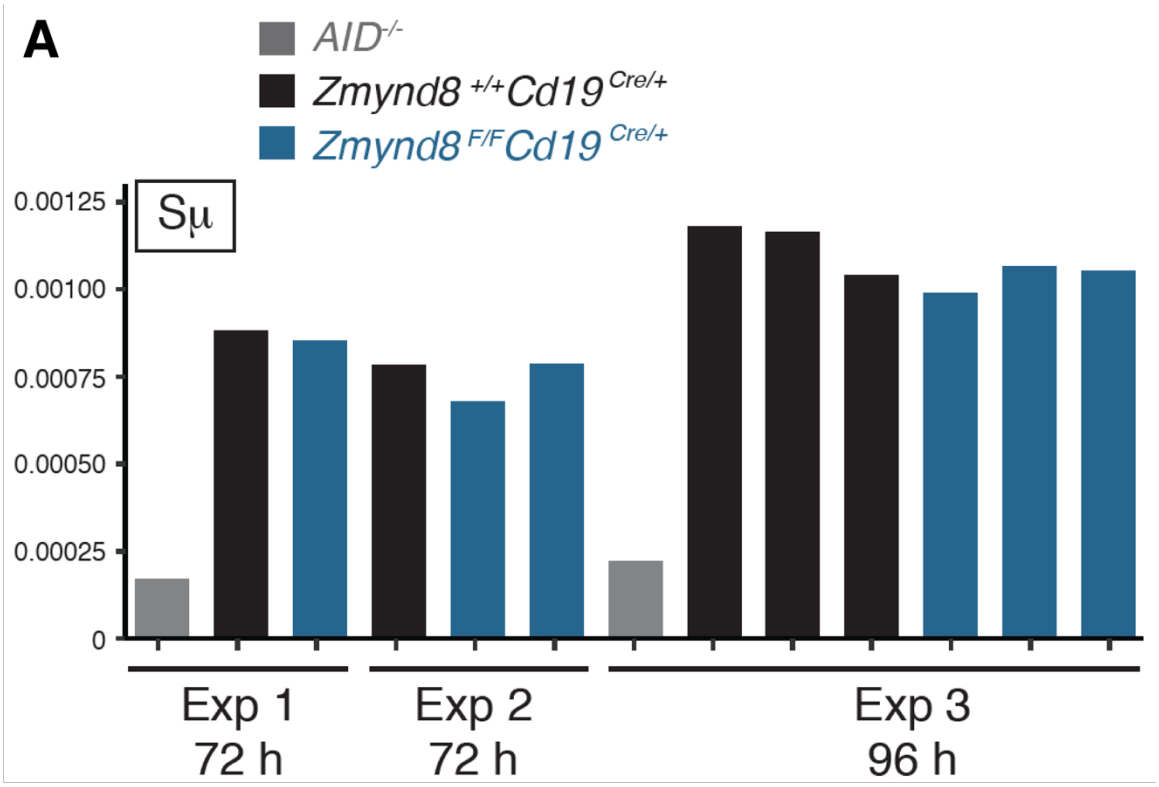


Figure 26: Frequency of mutations at 5'Sμ in *Zmynd8$^{F/F}$CD19$^{Cre/+}$* splenocytes

A) Graph depicting cumulative mutation frequencies as determined by MutPE-Seq at 5'-Sμ in *CD19$^{Cre/+}$*, *Zmynd8$^{F/F}$CD19$^{Cre/+}$*, and *AID$^{-/-}$* splenocytes 72 h or 96 h after activation with LPS and IL-4. Three independent experiments are shown with each bar representing a different mouse.

7.4 ZMYND8 and Chromosome Architecture at the *IgH* locus

In activated and resting B cells the 3'RR forms looping contacts with the regions surrounding the 5' Eμ enhancer (VDJ-Eμ). Upon activation, specific S regions are recruited into the VDJ-3'RR loop in a cytokine-dependent manner (Thomas-Claudepierre, et al., 2016), (Wuerffel, et al., 2007). This topological organization is thought to facilitate both GLT expression and synapsis of recombining S regions.

Because ZMYND8 binds the 3'RR region we chose to investigate if the normal 3D association within *IgH* was perturbed when ZMYND8 was removed. ZMYND8 binding to the *Igh* super-enhancers does not appear to contribute to the establishment or maintenance of this architectural structure as determined by 4C analysis (Figure 27a).

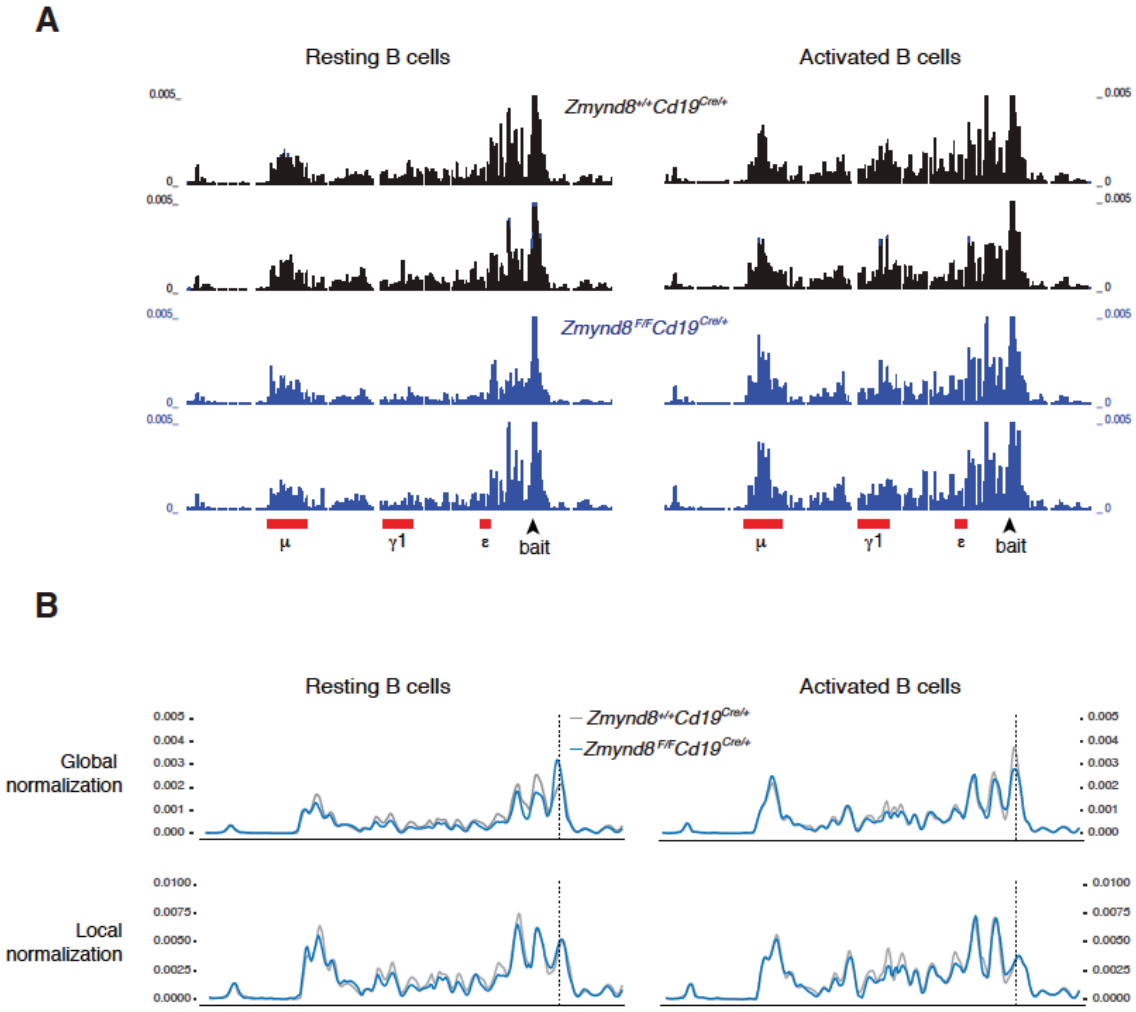


Figure 27: ZMYND8-deficiency does not impair *Igh* locus chromatin architecture

(A-B) Long-range chromatin looping interactions within the *Igh* locus in unactivated and activated (48 h with LPS and IL-4) splenocytes from *Cd19*$^{Cre/+}$ and *Zmynd8*$^{F/F}$*Cd19*$^{Cre/+}$ mice using a bait (black arrow) located within the 'RR enhancer. 4C-Seq analysis of two mice per genotype are displayed.

Chapter 8: Discussion

8.1 The *IgH 3'RR*: a key regulator of CSR

B cell development, survival, and function require the developmental expression (Mårtensson, Almqvist, Grimsholm, & Bernardi, 2010), mature expression (Lam, Kuhn, & Rajewsky, 1997), and continued signaling (Kraus, Alimzhanov, Rajewsky, & Rajewsky, 2004), (Srinivasan, et al., 2009) of the B cell receptor (BCR). It was discovered that initial BCR expression and by extension the capability to produce antibodies was controlled by the 5' enhancer of the *IgH* constant region (E_μ) (Banerji, Olson, & Schaffner, 1983), (Gillies, Morrison, Oi, & Tonegawa, 1983), (Neuberger M. , 1983). However, the enhancer is not required for continued BCR expression after V(D)J recombination (Eckhardt & Birshtein, 1985), (Zaller & Eckhardt, 1985). This led to the model that the E_μ is essential for activation of the *IgH* locus but not maintenance. A second enhancer was proposed to account for the continued expression of the *IgH* locus. This was supported by the discovery of a mouse myeloma cell line with impaired heavy chain transcription which have a deletion downstream to the last constant region C_α**,** (Gregor & Morrison, 1986). The 3' RR was then identified in rats as a large enhancer 25 kb downstream of the *IgH* locus (Pettersson, Cook, Bruggemann, Wiliams, & Neuberger, 1990). The mouse homolog was identified piecemeal, as an assembly of DNase I hypersensitive sites (HS). First identified was HS1,2 (Lieberson, Giannini, Birshtein, & Eckhardt, 1991), then HS3a (Giannini, Singh, Calvo, Ding, &

Birshtein, 1993) , (Matthias & Baltimore, 1993) just 3' of C_α, and then HS4 at 33 kb distant (Michaelson, Giannini, & Birshtein, 1995). Individually, the four 3'RR HS sites comprising the 3' RR are weak transcriptional enhancers compared to E_μ. The strongest HS site, 4a possessed only 25% of the transcriptional strength of the E_μ region while the 3a site barely had any transcriptional enhancement. The difference in enhancement ability is even more striking when comparing the relative size of the regions. E_μ is fully encompassed by a 700 bp region within the J_H-Cμ intron and all transcription factors bind to a core region of 220 bp (Nikolajczyk, Dang, & Sen, 1999). In contrast, the 3' RR extends 33 kb beyond the 3' of the end of the last coding segment at *IgH,* C_α. Despite the size difference, the 3' RR, as a whole, is a strong enhancer that drives expression of oncogenes in translocations (Madisen & Groudine, 1994). The 3' RR was thus identified as a canonical locus control region as it causes lineage specific and defining gene expression (Li, Peterson, Fang, & Stamatoyannopoulos, 2002), and is now recognized as a super-enhancer (Whyte, et al., 2013).

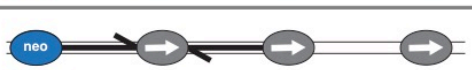
Because the 3' RR only altered *IgH* expression of mature B cells, i.e. lymphocytes that had completed V(D)J recombination it was hypothesized that the 3'RR affected secondary immune diversification reactions, CSR and SHM. Cogné and colleagues confirmed this for CSR by replacing the 5' most HS site, HS3a, with an expressed neomycin resistance gene (*neo*^{R}) (Cogne, et al., 1994). This reduced CSR to isotypes IgG2a, IgG2b, IgG3, and IgE, but not IgG1. The reducing in CSR was accompanied by a reduction in the corresponding GLTs. However, the

disruption of CSR was attributable to the *neo*^{R} cassette and not deletion of the HS3a enhancer *per se* (Manis, et al., 1998). Analogous to the emergent property that all four enhancers are required to significantly upregulate *IgH* expression, deletion of the entire 3'RR impaired CSR to all isotypes (Vincent-Fabert, et al., 2010), but individual deletions had no effect on CSR (Manis, et al., 1998), (Bebin, et al., 2010), (Vincent-Fabert, et al., 2009). It was hypothesized that the 3'RR influences CSR independent of the specific HS sites, and instead acts via the multi-kb intervening DNA segments. Reconstituted HS sites, without the intervening DNA, showed a heterogeneous effect, recovering CSR to IgG1 but not IgG3 (Le Noir, et al., 2017). Furthermore, 5' deletion of HS and intervening DNA from the 3' RR (HS3a to HS3b) impaired CSR to IgG1, IG2a, and IgG3 but the remaining HS4a and contextual DNA supported CSR to IgA and IgG2b at normal levels (Garot, et al., 2016). The reported mutants and deficiencies in CSR were matched by reductions in the corresponding GLTs. Thus, there appears to be a complex emergent relationship between the HS sites, intervening DNA, and transcriptional status in the impact on CSR (Figure 28). However, the mechanism of how the 3'RR promotes germline transcription to effect CSR remains to be elucidated. In this study we identified ZMYND8 as a novel effector of CSR acting by binding at the 3'RR to promote GLT.

Figure 28: Comparison of Different *IgH 3'RR* mutants and CSR

Table 2: Comparison of Different *IgH 3'RR* mutants and CSR

Left column shows graphical representation of mutants. Grey circles represent WT HS sites. Red circles are deletions of HS site along with any contiguous segments in light grey. Blue circles are HS site replaced with neomycin cassette (along with any contiguous segments in light grey). Yellow is inversion of HS site. Half arrows indicate palindromic sequences. Second column is the genotype. Third column is the source paper. Right columns (6) represent CSR to isotypes listed. Grey: CSR for that isotype was not measured. Green: no difference between mutant and wild type. Red: decrease in CSR. Orange: decrease proportionately smaller compared to other isotypes in study.

HS3a HS1,2 HS3b HS4	Genotype	Reference	G1	G3	G2b	G2a	E	A
	3'RR WT							
	HS3a$^{neo/neo}$	Cogne et al. 1994						
	HS3a$^{neo/neo}$	Manis et al. 1998						
	HS3a$^{\Delta/\Delta}$	Manis et al. 1998						
	HS1,2$^{neo/neo}$	Manis et al. 1998						
	HS1,2$^{\Delta/\Delta}$	Manis et al. 1998						
	HS3b$^{\Delta/\Delta}$	Bebin et al. 2010						
	HS4$^{neo/neo}$	Vincent-Fabert et al. 2009						
	HS4$^{\Delta/\Delta}$	Vincent-Fabert et al. 2009						
	(HS3b...HS4)$^{neo/neo}$	Pinaud et al. 2001						
	(HS3b...HS4)$^{\Delta/\Delta}$	Pinaud et al. 2001						
	3'RR (4xHS)$^{\Delta/\Delta}$	Vincent-Fabert et al. 2010						
	(HS4+2kb) $^{neo/neo}$	Manis et al. 2003						
	(HS3a...HS1,2)$^{\Delta/\Delta}$	Saintamand et al. 2016						
	HS3a$^{inv/inv}$	Saintamand et al.2016						
	(HS introns)$^{\Delta/\Delta}$	Le Noir et al. 2017						
	(HS3a...HS3b)$^{\Delta/\Delta}$	Garot et al. 2016						

8.2 ZMYND8 and the *IgH 3'RR*

In this study we discovered that deletion of the chromatin reader ZMYND8 reduces CSR by down-regulating acceptor GLTs in primary B cells. Furthermore, we excluded the possibility that ZMYND8 deletion may reduce CSR via impaired donor GLT, AID targeting at S_μ, proliferation or impaired B cell development. The IgH 3'RR similarly promotes CSR by facilitating transcription of GLT at the acceptor switch region (Saintamand, et al., 2015). The IgH 3' RR also does not impair GLT at S_μ (Vincent-Fabert, et al., 2010), does not reduce AID targeting at S_μ (Saintamand, et al., 2015), nor does it alter proliferation in actively switching B cells (Cogne, et al., 1994). Due to these similarities we conclude that ZMYND8 deficiency phenocopies *IgH 3'RR* loss with respect to CSR in actively switching B cells.

Furthering the correlation, we characterized the genomic distribution of ZMYND8 via ChIP-seq and discovered that ZMYND8 co-localizes with the *IgH* 3'RR. This interaction is specific to the 3'RR as ZMYND8 does not associate with the intervening switch or constant regions or the 5' E_μ. We also discovered that ZMYND8 association at the 3'RR can be divided into two sections: a first peak covering the palindromic region delimited by the HS3a-HS1,2-HS3b (proximal), and a second peak centered at HS4 (distal). The palindromic first region is a conserved feature in mammals (D'Addabbo, Scascitelli, Giambra, Rocchi, & Frezza, 2011) and SHM is primarily controlled by this region (Garot, et al., 2016).

The HS4 (distal) module functions somewhat separately and supports CSR to IgA and IgG2b independently (Garot, et al., 2016). The HS4 region also supports CSR to IgA via the interaction with a lncRNA (Pefanis, et al., 2016). It would be interesting to see if this lncRNA supports CSR to IgG2b and if other lncRNAs interact with other HS sites. The division of ZMYND8 may point to the functional differences between these regions.

8.3 ZMYND8 Potential Mechanisms of Action at the *IgH 3'RR*

Because ZMYND8 is highly enriched at the *IgH 3'RR* and phenocopies the deletion of the *IgH 3'RR* in CSR through reduced GLT, we proposed that it might act to promote CSR through a common mechanism. As a canonical super-enhancer, the IgH 3'RR is thought to promote acceptor GLT at S regions by looping into close proximity with the intronic promoters. The loop model of (super)enhancer function dictates that transcriptional machinery and licensing factors are directed to target regions by chromosomal folding over large distances (Li, Notani, & Rosenfeld, 2016), (Kenter, et al., 2012). These conformations can be cell type specific and stable or transient. B cells maintain a chromosomal loop between the 3'RR and VDJ- E_{μ} but this chromatin organization is absent in T cells (Wuerffel, et al., 2007), (Ju, et al., 2007), (Sellars, Reina-San-Martin, Kastner, & Chan, 2009). When B cells are activated, the 3'RR also comes into contact with the intronic

enhancer specific for the S region being targeted. The degree of targeting is also proportional to the degree of transcription elicited (Wuerffel, et al., 2007).

Like the Mediator complex (Med) which interfaces between Pol II and transcription factors to promote the Pol II regulation including transcriptional pausing and elongation (Allen & Taatjes, 2015), ZMYND8 associates with Pol II (Malovannaya, et al., 2011), specifically with the paused (Ser-5-P) form of Pol II, (Adhikary, et al., 2016). ZMYND8 also binds to gene promoters in DU145 cells (Li, et al., 2016; Alt, Zhang, Meng, Guo, & Schwer, 2013), CH12 cells, and primary B cells. It was recently shown that deletion of *Med1* and *Med2* reduces CSR by reducing long-range chromosomal contacts between the E_{μ} and proximal palindromic module of the 3'RR (Thomas-Claudepierre, et al., 2016). Additionally, contacts between E_{μ} and downstream acceptor regions were reduced significantly. Due to the occupancy of ZMYND8 at the 3'RR region, its association with transcription machinery, and reduction in GLT we sought to investigate if ZMYND8 repressed GLT and CSR in a mechanism similar to the Med complex.

Using 3C (chromosomal conformation capture) and 4C (circular chromosome conformation capture) experimental approaches we showed that the VDJ- E_{μ} and 3'RR chromosome conformation is stable regardless of ZMYND8 presence. Similarly, the activation induced 3'RR contact with the intronic enhancer of IgG1 did not decrease with ZMYND8 deletion. We concluded that ZMYND8

does not promote GLT through contact of the 3'RR with the S region promoters. In line with this, ZMYND8 must act downstream of *IgH* locus chromosomal architecture remodeling and at a separate stage than the Med complex.

Instead of reducing chromosomal contacts between the 3'RR and the acceptor switch regions, deletion of ZMYND8 might reduce GLT at the acceptor S regions by decreasing RNA Pol II (Pol II) occupancy. Pol II is the eukaryotic RNA polymerase that transcribes messenger RNA (mRNA), including many ncRNAs (e.g. GLTs). Pol II is highly regulated to control transcription levels and maintain transcription quality. Pol II is regulated at four potential stages: (1) recruitment to the transcription start site (TSS), (2) entry into initial transcription—resulting in a short elongation phase of 50-60 nucleotides, (3) pausing at this promoter proximal region, and (4) productive elongation of the gene body or termination and release (Jonkers & Lis, 2015). Pausing and elongation are characterized by phosphorylation on the carboxy-terminal domain (CTD) of Pol II. Phosphorylation at Ser-5 after entry pauses Pol II, while phosphorylation by positive elongation factor B (P-TEFb) phosphorylates Ser-2, leading to productive elongation. As Pol II transcribes through the gene body, Ser-5-P levels decrease and Ser-2-P levels increase until termination and transcript release (Harlen & Churchman, 2017). Pausing and release for elongation are key regulatory steps, both for production of GLT and for targeting of AID.

It has been extensively demonstrated that transcription is necessary for AID targeting of switch regions (Yewdell & Chaudhuri, 2017). However, the process of transcription differentially targets AID within the switch region. ChIP of Poll II at the donor S_{μ} and at the acceptor $S_{\gamma 3}$ in activated B cells showed that the spatial distribution of Pol II is proportional to the mutation density (Pavri & Nussenzweig, 2011). In particular, Pol II occupancy increased from the TSS in the intronic promoter and peaked at the S region 5' boundary. High density Pol II was detected throughout the S region and only decreased after the 3' end (Rajagopal, et al., 2009), (Wang, Wuerffel, Feldman, Khamlichi, & Kenter, 2009). This distribution profile matches the regulation of Pol II with binding increasing from the TSS of the intronic enhancer, accumulating at the proximal promoter site just downstream, and continued occupancy in the S region/gene body until termination of the transcript after the coding region. This Pol II density correlates with the mutational density (Peters & Storb, 1996), (Rajagopal, et al., 2009). Pol II was also discovered to be upstream of AID targeting as deletion of AID did not affect Poll II distribution (Rajagopal, et al., 2009), (Wang, Wuerffel, Feldman, Khamlichi, & Kenter, 2009). AID's association with stalled Pol II is a general property and was detected in promoter-proximal regions throughout the genome (Yamane, et al., 2011). Run on analysis to detect transcribed RNA at the site of transcription also showed that Pol II distribution associates with regions of active transcription. Thus, the correlation of transcription apart from AID showed that Pol II occupancy is sequence and transcription specific and not AID dependent.

However, transcription occurs at literally thousands of places in the genome, but only a small subset of regions (i.e. Variable genes and S regions) are intended targets of AID. Instead, Pol II stalling is essential for AID targeting. This was demonstrated because stalled Pol II accumulates at the promoter-proximal region with AID off target genes c-Myc (Bentley & Groudine, 1986), (Krumm, Meulia, Brunvand, & Groudine, 1992), IgK (Raschke, Albert, & Eick, 1991) and Pim1 (Rohwer, Todd, & McGuire, 1996). Promoter pausing was proposed to act to promote AID targeting by increasing ssDNA availability. This was corroborated by the detection of ssDNA at the Variable regions and not constant regions of activated Ramos B cells undergoing SHM, despite both regions being actively transcribed (Ronai, et al., 2007). Pol II stalling was directly implicated in targeting AID to Switch regions in particular and DNA in general with the discovery through an unbiased shRNA that SPT5 was required for CSR (Pavri, et al., 2010). SPT5 had previously been found to associate with ssDNA and stalled Pol II (Gilmour, 2009), (Lis, 2007), (Rahl, et al., 2010). SPT was found to directly associated with AID and SPT depletion decreased AID targeting to the intended *IgH* S regions and known AID off target loci. Importantly, deletion of the 3'RR in mice has been shown to reduce Pol II loading at the acceptor S regions as detected by ChIP of Ser-5-P / paused Pol II (Saintamand, et al., 2015). Thus, because AID targeting at the S regions is dependent not only on the presence of processed transcripts as GLTs but also the active transcription and pausing of Pol II in the generation of these

transcripts, we sought to investigate if ZMYND8 depletion affected stalled Pol II at the S regions.

Using a ChIP-seq approach for Ser-5-P Pol II on activated WT and *CD19$^{Cre/+}$ ZMYND8 $^{f/f}$* we discovered that despite the decrease in GLT from the acceptor $S_{\gamma 1}$ region, there is no decrease in paused Pol II at $S_{\gamma 1}$. This implies that there is no reduction in targeting of AID and resultant CSR due to insufficient Pol II pausing. This is also the first evidence that ZMYND8 deletion does not exactly phenocopy 3'RR deletion with respect to Pol II loading at the 3'RR. That is to say ZMYND8 deletion retains stalled Pol II at acceptor switch regions, but 3'RR deletion decreases stalled Pol II (Saintamand, et al., 2015). This discrepancy brings up two questions. First, how are GLTs downregulated in ZMYND8 deficiency, if Pol II loading is unaffected, and second how could decreases GLTs cause a reduction in CSR if Pol II pausing remains at normal levels? Decreased GLTs could directly account for reduced switching independently of pausing via intronic G-quadruplex structures (Zheng, et al., 2015). However, this still does not mechanistically account for the actual decreased GLT in light of normally paused Pol II. Because ChIP-seq of Ser-5-P Pol II and GLT measurement by RT-qPCR only detects steady state levels, global run-on sequencing (Core, Waterfall, & Lis, 2008) could reveal if decreased steady state GLT was due to altered GLT production. Alternatively, GLTs could be subject to increased degradation.

If the reduction in GLT leading to a reduction in AID targeting at the switch acceptor region is the sole explanation for reduced CSR we could empirically demonstrate this by assessing the acceptor switch region for reduced AID induced mutations. It is important, however, to note the technical limitations of doing so, i.e. that acceptor switch regions are targeted less than the donor S_μ (Schrader, et al., 2003) and effective measurement of AID activity can best be seen in a *UNG*$^{-/-}$ *MSH2*$^{-/-}$ background (Xue, Rada, & Neuberger, 2006). Regardless, this would not mechanistically account for ZMYND8's function in reduction GLTs. Thus, although we have been able to determine that ZMYND8 depletion reduces GLT, the reduction in GLT is not due to impaired interaction of the 3'RR with acceptor S regions or alterations in transcriptional machinery at the S regions as detected by ChIP-seq.

Interestingly, ChIP-seq of Pol II revealed extensive binding to the 3'RR in the absence of ZMYND8. We confirmed that this upregulated transcription at the 3'RR by rRNA depleted global RNA-seq in CH12 cells and and RT-qPCR in primary B cells. This is consistent with ZMYND8's ability to repress eRNA transcription at enhancer regions (Shen, et al., 2016), and with its broader action as a transcriptional regulator (Adhikary, et al., 2016), (Malovannaya, et al., 2011), (Poleshko, et al., 2010), (Savitsky, et al., 2016), (Shen, et al., 2016), (Spruijt, et al., 2016), (Zeng, Kong, Li, & Mao, 2010). Interestingly, despite increased deposition of Pol II across the entire 3'RR, we only found increases in transcription at the

HS3B and HS1,2 enhancers. It is possible that increased transcription at the 3'RR results in increased AID targeting and recombinational deletion of the entire *IgH* C_H gene cluster leading to B cell death, a process termed locus suicide recombination (LSR) (Peron, et al., 2012). This might be reflected in the lower amounts of mature B cells circulating in 3'RR deficient mice, but that is primarily due to the distal 3'RR encompassing the HS4 enhancer (Garot, et al., 2016). Additionally, although LSR has been demonstrated to occur in B cells, it has yet to be proven as a mechanism that regulates CSR. Instead, the reduction in CSR via GLT reduction is more consistent with the known function of the *IgH 3'RR* and other lines of evidence point to transcription deregulation at the *3'RR* as the mechanism.

Early investigations in the 3'RR showed that while singular deletions of HS sites did not affect CSR, replacement of these sites with actively transcribing neomycin resistance cassettes dramatically reduced CSR (Manis, et al., 1998) (Cogne, et al., 1994). This implies that an increase in transcription at the 3'RR region is causative for a reduction in CSR. Although the disruption of CSR via activation of the 3'RR could be mechanistically distinct from the reduction in CSR resulting from the entire deletion of the 3'RR (Vincent-Fabert, et al., 2010) –which would inherently result in zero transcription, both models (i.e. up-regulation of 3'RR transcripts via neomycin gene-insertion/Zmynd8 deletion and 3'RR deletion) cause a reduction in GLTs at the acceptor S regions. This common feature implies that

all three cases are mechanistically similar and are due to deregulation of the normally transcriptionally repressed 3'RR. It is also likely that the eRNAs from the 3'RR in the case of ZMYND8 depletion do not act in trans to disrupt GLTs as the similar disruption due to neomycin gene transcription points to a sequence agnostic mechanism. This could potentially be tested in two ways. First, transcription of 3'RR specific eRNA at a different locus could be tested if it disrupts GLT and CSR. Additionally, stimulation of endogenous eRNA transcription with a dead-Cas9 fused to an activator domain (La Russa & Qi, 2015) could be assessed for reduction in GLT and CSR.

Ultimately the question remains, how does dysregulation of transcription at the 3'RR lead to decreased CSR via a reduction in GLTs at the acceptor S regions? With ZMYND8 deficiency, there is no reduction in chromosomal looping and no reduction in stalled Pol II at acceptor S regions. Instead GLT reduction could result from inhibited Pol II release and transcript elongation. A difference in actively elongating transcripts could be detected with GRO-seq. A block to elongation should increase the amount of stalled Pol II proportionately. However, such an increase was not detected with ChIP-seq. This expected commensurate increase in stalled Pol II may not occur if the amount of paused holoenzyme is optimized for AID targeting of S regions and represents maximal loading potential under WT conditions (i.e. there is no capacity to increase occupancy of paused Pol II at the acceptor switch). In this case, Pol II initial binding and entry would be expected to

decrease or Pol II off-loading/termination at the proximal pause step would be expected to increase to maintain consistent enzyme levels. Correspondingly, it has recently been shown that there is high Pol II turnover at paused genes (Krebs, et al., 2017). ZMYND8 could mediate this directly as it interacts with the stalled form of Pol II (Malovannaya, et al., 2011), (Adhikary, et al., 2016). Alternatively, ZMYND8 could cause a reduction in GLTs through a secondary mediator. ZMYND8 has been shown to not only be a transcriptional repressor as deletion also decreases transcription of multiple genes (this study), (Li, et al., 2016). A potential secondary mediator is the integrator complex, which controls the termination of proximal-paused Pol II (Skaar, et al., 2015). This would be consistent with what is currently known about ZMYND8 as it engages the integrator complex (Malovannaya, et al., 2011), (Savitsky, et al., 2016).

ZMYND8 may also be an epigenetic regulator of the switch regions. The S regions at the IgH locus undergo dynamic epigenetic modulation during activation and CSR in B cells. Epigenetic marks H3K4me3 and H3K9ac/H3K14ac are associated with the promoters of actively transcribing genes (Santos-Rosa, et al., 2002), (Karmodiya, Krebs, Oulad-Abdelghani, Kimura, & Tora, 2012). In resting B cells these marks correlate with the actively transcribed regions of the *IgH* locus, specifically the Iμ promoter and the Sμ region (Chowdhury, et al., 2008), (Jeevan-Raj, et al., 2011), (Wang, Wuerffel, Feldman, Khamlichi, & Kenter, 2009), (Stanlie, Aida, Muramatsu, Honjo, & Begum, 2010). Upon activation, these marks increase

at the donor switch region Sμ and also appear on downstream acceptor switch regions in proportional amounts to their activation evidenced by GLTs and resultant CSR to the respective isotype. The epigenetically remodeling of the *IgH* locus is dependent on activation, but independent of the action of AID (Kuang, Luo, & Scharff, 2009). Transcription is intimately associated with epigenetic modifications (Gates, et al., 2017), but the causal relationship between histone modification, GLT, and CSR remains to be fully elucidated. The deletion of many histone-modifying enzymes results in embryonic lethality, presumably due to the complex transcriptional orchestration in development (Butler, Koutelou, Schibler, & Dent, 2012). However, their effects could still be interrogated in B cells via conditional deletions at later stages of development. From the few histone methyltransferase/demethylase and acetyltransferase/deacetylases it is apparent that GLTs and CSR can be somewhat decoupled from individual histone modifications. For example, H3 acetylation at the acceptor Sε region was artificially increased with a histone deacetylase inhibitor, but it could not increase Sε GLTs and CSR to IgE in activated cells (Nambu, et al., 2003). Additionally, knockdown of the FACT (facilitates chromatin transcription) complex and the methyltransferase SPT16 impaired CSR (Stanlie, Aida, Muramatsu, Honjo, & Begum, 2010). Although H3K4me3 markings were reduced, GLT and AID expression remained normal. Additionally, depletion of the MLL3/MLL4 methyltransferases in B cells impaired CSR without reducing GLTs (Starnes, et al.,

2016). Thus, epigenetic marks and GLTs may contribute in parallel pathways to facilitate CSR.

The interplay between ZMYND8, GLT transcription, and epigenetic modulation remains to be explored. It is known that ZMYND8 associates with H3K4me0 and H3K4me1 in conjunction with H3K14ac and recruits the demethylase of H3K4me3, JARID1D (Li, et al., 2016). ZMYND8 was also found to recruit the demethylase KDM5C to demethylate H3K4me3 (Shen, et al., 2016). Given the strong association of ZMYND8 for demethylated H3K4me it seems paradoxical that depletion would affect the deposition of H3K4me3 on the switch regions. Alternatively, it would be interesting to investigate if depletion of ZMYND8 increases H3K4me3 on the 3'RR. This could provide a mechanism for the increase in Pol II loading and transcription.

8.4 RIF1 and ZMYND8

RIF1 Protein Interactome

The impetus for this study was to uncover novel components of the RIF1 interactome related to NHEJ and by extension CSR. RIF1 was found to be the only necessary effector to prevent end-resection to promote NHEJ and CSR in activated B cells (Escribano-Díaz, et al., 2013), (Chapman, et al., 2013), (Di

Virgilio, et al., 2013). However, it is unknown how RIF1 engages the phosphorylated S/TQ residues as the N-terminus of 53BP1 or how RIF1 protects DNA DSB breaks to promote CSR. We originally proposed that a separate RIF1 interactor may mediate these functions.

Utilizing I-DIRT with sub-stochiometric amounts of cross-linking with glutaraldehdye we identified the RIF1 protein interactome in B lymphocytes undergoing CSR. The screen identified known interactors of RIF1 such as 53BP1 (Di Virgilio, et al., 2013), BLM (Xu, et al., 2010), Protein Phosphatase I (Hiraga, et al., 2014), (Sukackaite, et al., 2017), and BACH1 (Tanaka, et al., 2016). We were also able to identify interactors of interactors such as DYNLL1, which binds to 53BP1 (Lo, et al., 2005). Interestingly, BACH1 is more characterized as a phospho-dependent binder of BRCA1. BACH1 binding to BRCA1 is necessary for BRCA1 to exclude RIF1 from sites of DNA damage in S/G2 phase of the cell cycle (Escribano-Díaz, et al., 2013). Additionally, activation of BACH1 in S phase requires dephosphorylation (Kumaraswamy & Shiekhattar, 2007). It remains to be studied if BACH1 associates with RIF1 in all cell cycle phases, if that interaction is mutually exclusive to BRCA1 association based on phosphorylation status, and what is the implication of the interaction with RIF1. For this study, we decided to validate the novel interactors of RIF1 in promoting NHEJ by screening them for promotion of CSR.

RIF1 and CSR

RIF1 is an essential effector of CSR because it inhibits end-resection to promote NHEJ of S region DSBs (Chapman, et al., 2013), (Escribano-Díaz, et al., 2013), (Di Virgilio, et al., 2013). We found that ZMYND8 depletion reduced CSR in the CH12 B cell lymphoma line and in primary B cells via a *Rosa26*$^{Cas9/+}$ gRNA somatic targeting model and via conditional deletion using *CD19*$^{Cre/+}$ *Zmynd8*$^{f/f}$. We expected ZMYND8 to decrease CSR in a mechanistically similar way to RIF1 due to their association in actively switching B cells and evidence of ZMYND8's function in DSB repair. Specifically, ZMYND8 was recently found to promote homologous recombination repair of DSBs in regions of active transcription by locally repressing transcription (Gong, et al., 2015), (Gong, Clouaire, Aguirrebengoa, Legube, & Miller, 2017), (Spruijt, et al., 2016), (Xia, et al., 2017), (Kloet, et al., 2014). However, we were unable to document a significant effect of ZMYND8 in the protection and repair of DSBs in ZMYND8-deficient CH12 or MEFs cells, irrespective of the source of DNA damage (IR, PARPi, or CRISPR-Cas9). This observation is consistent with our finding via co-immunoprecipitation that ZMYND8 constitutively binds RIF1 in B cells undergoing CSR independent of DNA damage or the ATM kinase. With respect to the action of ZMYND8 at the 3'RR and its reduction in CSR, this is also consistent as the 3'RR is dispensable for pathway choice during DSB resolution (Cogne, et al., 1994). Thus, despite the interaction of ZMYND8 and RIF1 and their similar role to ensure efficient CSR, the two proteins appear to enable CSR by completely different mechanisms: RIF1 by

protecting DNA end-resection to promote NHEJ and ZMYND8 by facilitating sufficient GLT from the acceptor S regions and by acting at the 3'RR.

ZMYND8 and CSR: Perspectives and Models

ZMYND8 is a chromatin reader of acetylated and demethylated histones. It is thought to mediate transcriptional regulation by binding to and repressing the activity of promoters and enhancers, (Adhikary, et al., 2016), (Malovannaya, et al., 2011), (Poleshko, et al., 2010), (Savitsky, et al., 2016), (Shen, et al., 2016), (Spruijt, et al., 2016), (Zeng, Kong, Li, & Mao, 2010). In accordance with these reports, we found that ZMYND8 binds to the *IgH* super-enhancer (3' RR) and negatively regulates its transcription.

Mutation of the 3'RR in mice showed that it is required for CSR (Cogne, et al., 1994), (Garot, et al., 2016), (Le Noir, et al., 2017), (Manis, et al., 1998), (Pinaud, 2001), (Saintamand, et al., 2015), (Vincent-Fabert, et al., 2010). The 3'RR regulates CSR by promoting break formation via GLT transcription, RNA Pol II pausing, AID targeting (Cogne, et al., 1994), (Garot, et al., 2016), (Le Noir, et al., 2017), (Manis, et al., 1998), (Pinaud, et al., 2001), (Saintamand, et al., 2015), (Vincent-Fabert, et al., 2010). However, the 3'RR does not affect repair or pathway choice in DSB resolution in CSR (Cogne, et al., 1994). We found that ZMYND8 deficiency increases 3'RR transcription and phenocopies 3'RR deficiency, which indicate that it regulates the 3'RR superenhancer. ZMYND8 deletion causes

impaired CSR via reduction in acceptor GLT transcription in primary B lymphocytes. We were not able to detect a significant decrease in S_μ GLT or frequency of AID-induced mutations at 5' S_μ in *CD19*$^{Cre/+}$ *Zmynd8*$^{f/f}$ B cells. The effect of ZMYND8 in primary cells only on the downstream GLT region $S_{\gamma 1}$ is consistent with observations that the 3'RR controls GLT at acceptor S regions preferentially (Saintamand, et al., 2015). However, deletion of ZMYND8 in CH12 cells also resulted in impaired CSR, but GLTs at the donor S_μ or the acceptor S_α were unaffected. Instead we observed a reduction in AID-induced mutations at 5' S_μ.

We believe technical limitations explain these discrepancies between S_μ mutations and acceptor S region GLT. Although we were not able to detect a significant decrease in the frequency of AID-induced mutations at 5' S_μ in *CD19*$^{Cre/+}$ *Zmynd8*$^{f/f}$ B cells, it is notable that (1) *Zmynd8* is inefficiently Cre-deleted in this strain and (2) the dynamic range of the mutation assay is low, specifically the decrease in deletions at 5' S_μ in CH12 cells is only ~2-fold which is consistent with the small fold change caused by the knockout of known AID recruiter KAP1 in primary B cells (Jeevan-Raj, et al., 2011). By increasing Cre delivery retrovirally to primary B cells and pre-screening populations for *Zmynd8* deletion we would expect to more accurately assess AID mutations in primary B cells. Additionally, by accounting for CSR status and proliferation stages (i.e. activated pre-switched

IgM+ cells at division 5) as in (Reina-San-Martin, Chen, Nussenzweig, & Nussenzweig, 2004) we could increase the sensitivity for AID mutation differences.

Furthermore, the discrepancy between the decrease in acceptor GLT in primary B cells and the persistence of GLT at S_{α} in CH12 cells is consistent with previous reports that deletion of the 3'RR in CH12 cells reduces CSR but does not decrease GLT at S_{α} (Kim, Han, Santiago, Verdun, & Yu, 2016). GLT in CH12 cells may not be an effective model to explore the 3'RR and CSR because (1) GLT at S_{α} is constitutively expressed in CH12 cells (Nakamura, et al., 2006), and (2) B-1 cells switch efficiently to IgA in the absence of the 3'RR (Issaoui, et al., 2017). Given that the originating lymphoma for CH12 cells is from the B-1a and not the B-2 lineage (Arnold, Gridina, Whitmore, & Haughton, 1988), (Won & Kearney, 2002), this may explain why there is no defect in acceptor S region GLTs with 3'RR dysregulation through ZMYND8 depletion.

Thus, when technical limitations are accounted for, ZMYND8 deficiency phenocopies the effect of 3' RR deletion as demonstrated by impaired acceptor GLT and AID targeting of S_{μ}. It does so without altering the topological organization.

Instead, the data indicates that ZMYND8 suppression of RNA Pol II loading on the 3'RR enhancer and resultant absent eRNA production facilitates GLT

transcription at the acceptor S region. The function and presence of elongating Pol II at the acceptor region is most likely altered either through a lack of interaction of paused Pol II with ZMYND8 or through improper localization of a secondary factor like the Mediator complex. In support of this model, replacement of individual (hs1,2 or hs3a) or paired (hs3b and hs4) core enhancers with an actively transcribed neomycin gene cassette resulted in a more severe defect in acceptor GLT expression and CSR than deletion of the same modules (Cogne, et al., 1994), (Manis, et al., 1998), (Pinaud, et al., 2001). Thus, deregulated transcription within the 3'RR interferes with its activity and disrupts the physiological regulation of germline transcription and CSR. Thus, we identify ZMYND8 as a key regulator of 3' RR transcriptional activity under physiologic circumstances. In this context, ZMYND8 promotes CSR by facilitating GLT and targeting of the acceptor switch region.

Ultimately, we have defined the RIF1 protein interactome in actively switching B cells. We found that the interactor, Zmynd8, a chromatin reader, is a novel factor, required to support normal levels of CSR. ZMYND8 controls CSR by modulating the transcriptional activity of the 3' regulatory region super enhancer to facilitate efficient germ-line transcription of the switch region. This represents a new mechanistic step in the regulation of CSR.

8.5 ZMYND8 Remaining Questions and Future Directions

Rif1 and ZMYND8 Association

It remains to be elucidated why ZMYND8 associates with RIF1. Despite both functioning to promote CSR, they do so at completely different stages. Categorically, RIF1 mediates CSR post DNA DSBs to engage the NHEJ pathway, while ZMYND8 operates pre-break to promote GLT and efficient AID targeting and DSB formation. Additionally, ZMYND8's association with CSR appears to function in relation to the 3'RR region while RIF1 associates primarily at the sites of DNA damage. ZMYND8 identification as a factor in repressing transcription to promote HR repair of DSBs was reminiscent of the original identification of RIF1 as a component of the HR repair pathway (Buonomo, Wu, Ferguson, & de Lange, 2009), but we have determined through extensive investigation that ZMYND8, unlike RIF1, does not impact NHEJ.

ZMYND8 may operate with RIF1 in a roll unrelated to DSB repair and CSR. This is supported by our results that the interaction is DNA damage independent and not affected by ATM inhibition. Aside from CSR, RIF1 functions to regulate DNA replication timing (Cornacchia et al., 2012) and to prevent replication stress induced cell death (Buonomo, Wu, Ferguson, & de Lange, 2009). The deregulation of late replication firing in the absence of RIF1 was recently linked to the interaction with protein phosphatase I (PP1), which phosphorylates the minichromosome

maintenance complex MCM (Alver, Chadha, Gillespie, & Blow, 2017). Intriguingly the MCM complex is also required for CSR (Wiedemann, Peycheva, & Pavri, 2016). ZMYND8 has also been implicated as being a protein enriched at the replication fork (Dungrawala, et al., 2015).

Future studies could investigate the relationship between RIF1 and ZMYND8 in replication stress and replication timing. Although the interaction between RIF1 and ZMYND8 is independent of ATM, it could be mediated by ATR, the kinase responsible for coordinating the cellular response to replication stress (Flynn & Zou, 2011). In accordance with this hypothesis, ZMYND8 was shown to be required for phosphorylation of CHK1 at S-345 (Gong, et al., 2015).

ZMYND8 CSR and the 3'RR

Although we have shown the ZMYND8 is a novel factor in CSR, promotes efficient GLTs, and putatively operates through the transcriptional de-repression of the *IgH 3'RR*, many questions remain. First, does ZMYND8 effect the role of 3'RR to promote SHM (Rouaud, et al., 2013). Also how does the de-repression of transcription at the 3'RR mechanistically impact CSR. Transcriptional activity, could be assessed independently of ZMYND8 through the dead-Cas9-transactivator system mentioned above. Additionally, it remains to be seen if Pol II turnover and elongation times at the acceptor switch regions are responsible for the decrease in GLTs. Given ZMYND8's association with various demethylases

and strong interaction with chromatin marks it would also be appropriate to characterize ZMYND8's impact on the epigenetic markings, specifically H3K4me3, H3K4me1,0, H3K14ac, and H3K27ac on the *IgH* locus.

ZMYND8 and Cellular Function

Numerous studies have implicated ZMYND8 in the repression of transcription at DSBs to promote repair by HR. However, the method by which transcriptional activity compromises DSB repair or the impact of failed repression is unknown. The embryonic lethality of *ZMYND8*$^{-/-}$ mice implicates that the consequences of failed repression at DSBs are immense, but cellular studies have shown minimal impact of ZMYND8 depletion for cell survival or cycling. It is similarly paradoxical that ZMYND8 has been implicated in preventing oncogenesis and metastasis in cell culture, but high expression is also associated with poor survival in the clinic. It would be relevant, then, to investigate in which cells ZMYND8 is essential and under what conditions it is critical for cell survival and function.

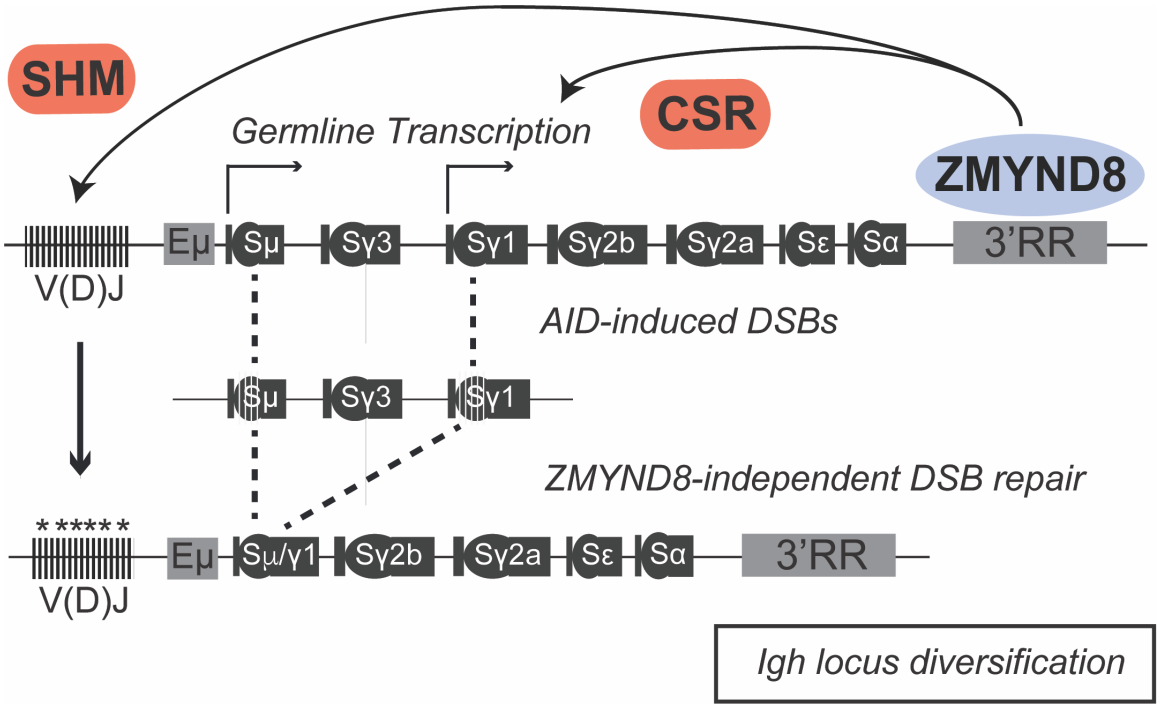


Figure 29: Model of ZMYND8 Function at the *IgH* Locus and in CSR

Model of ZMYND8 function adapted from final publication: (Delgado-Benito, et al., 2018). Note: post thesis defense, continued research revealed that ZMYND8 does mediate SHM at the *IgH* locus. Regarding CSR, ZMYND8 binds to the 3'RR, and in doing so excludes PolII there with a resultant reduction in eRNA production. This corresponds to decreased acceptor GLT, while the default GLT, Su is unaffected. CSR to all downstream isotypes is thereby reduced.

Chapter 9: Methods

Methods adapted from: (Delgado-Benito, et al., 2018).

Mice

Rif1$^{FH/FH}$ (Cornacchia et al., 2012), *Rif1*$^{F/F}$ (Buonomo et al., 2009), *Cd19*^{Cre} mice (Rickert et al., 1997), and *AID*$^{-/-}$ (Muramatsu et al., 2000) were previously described. The conditional *Zmynd8*^{F} allele was generated from crossing the knock-out first allele with conditional potential *Zmynd8*^{tm1a} (European Mouse Mutant Archive, EMMA #05720) with the *ROSA26*^{Flpo} deleter strain (The Jackson Laboratory, JAX #007844). All experiments were performed in compliance with EU Directive 2010/63/EU, and in agreement with protocols approved by Landesamt für Gesundheit und Soziales (LAGeSo, Berlin), The Rockefeller University (NY), and the National Institutes of Health (NIH) Institutional Animal Care and Use Committees.

B cell cultures and retroviral infection

B lymphocytes were isolated from mouse spleens using anti-CD43 MicroBeads (Miltenyi Biotec) and stimulated to undergo class switching with 25 μg/ml LPS (Sigma-Aldrich) and 5 ng/ml of mouse recombinant IL-4 (Sigma-Aldrich) (CSR to IgG1). CH12 cells (Nakamura et al., 1996) were stimulated to undergo CSR to IgA with 5-15 □g/mL αCD40 (BioLegend), 5 ng/ml TGF-β1 (R&D Systems) and 5 ng/ml of mouse recombinant IL-4 (Sigma-Aldrich). pMX- ZMYND8-3XFlag retroviral vector was generated by cloning the cDNA for murine ZMYND8 into pMX-IRES-

GFP with a C-terminal 3XFlag tag. For CH12 infections with pMX-based retroviral vectors, cells were subjected to 2 rounds of infection with amphotrophic retroviral supernatant, with or without selection with 1 □g/ml of Puromycin (Sigma-Aldrich), followed by activation for 48h before analysis for CSR efficiency.

I-DIRT

B-cells from *Rif1*$^{FH/FH}$ and *WT* mice were isolated and cultured in SILAC medium composed by RPMI (-Arg, -Lys) medium (Thermo Scientific), L-glutamine, sodium pyruvate, Hepes, 50 μM 2-mercaptoethanol, antibiotic/antimitotic, 10% dialyzed fetal bovine serum (Invitrogen), LPS, IL-4 and RP/14 (BD), and supplemented with either $^{13}C_6$ L-arginine and $^{13}C_6$ L-lysine (Cambridge Isotope Laboratories) (heavy medium; *Rif1*$^{FH/FH}$ culture), or non-labeled L-arginine and L-lysine (Sigma-Aldrich) (light medium; *WT* culture). Cells were cultured in SILAC medium for 96 h to ensure near-complete incorporation of the labeled amino acids. For IR treatment, cells were exposed to an X-rays source at a rate of 278.2 Rads/min for 431 seconds (20 Grays), followed by recovery at 37°C for 45 min.

Immunoisolation of RIF1^{FH}-containing complexes

2.6x10^9 cells per culture (*Rif1*$^{FH/FH}$ and *WT*) were collected by centrifugation, resuspended in 20 mM Hepes containing 1.2% polyvinylpyrrolidone (Sigma-Aldrich), protease and phosphatase inhibitor cocktails (Roche), 0.5 mM DTT, and frozen in liquid nitrogen. *Rif1*$^{FH/FH}$ and *WT* frozen cells were mixed in a 1:1 ratio

and cryogenically disrupted by wet milling in a Planetary Ball Mill PM 100 (Retsch). The resulting frozen cell grindate was rapidly thawed in extraction buffer (20mM Tris-Cl pH 8, 150 mM NaCl, 0.5% Igepal CA-630, 1.5 mM $MgCl_2$, Benzonase, protease and phosphatase inhibitor) supplemented with 2.5 mM Glutaraldehyde, and quenched with 100 mM Tris-CL pH8.0 buffer after 5 min incubation. The lysate was clarified by 10 min centrifugation at 13,000 rpm at 4°C, and immediately incubated with magnetic beads (M-270 epoxy beads, Invitrogen) conjugated with anti-Flag M2 antibody for 1 h at 4°C (Di Virgilio et al., 2013). The bead preparation was then washed in extraction buffer, and $RIF1^{FH}$ bait and associated proteins were eluted twice under native conditions by two rounds of incubation with 2.5 µg/µl 3XFlag peptide for 45 min at 4°C with shaking.

Mass spectrometric analysis

Rif1 baits and co-purifying proteins were resolved by 4-12% Bis-Tris gel and visualized by Coomassie blue staining. The gel was divided into upper and lower parts along the 39kDa molecular weight marker, with only the lower part fixed. The protein containing upper and lower parts were cut into 5 and 1 regions respectively, and the gel samples were subjected to in-gel tryptic digestion. Peptides were extracted and purified, analyzed by LCMS using a Thermo Q Exactive Plus mass spectrometer, with a Thermo Easy-nLC 1000 HPLC and a Thermo Easy-Spray electrospray source. Identification and quantification of proteins was performed by searching against a mouse protein sequence database with the MaxQuant

software (version 1.2.2.5) (Cox and Mann, 2008). Protein H/(H+L) ratios were derived using peptides' H/L intensity values in MaxQuant output.

CRISPR-Cas9 somatic targeting and clonal derivative generation

gRNAs for functional screens for loss of CSR function in CH12 were cloned into the U6 cassette of pX458 plasmid (pSpCas9(BB)-2A-GFP, addgene #48138) according to standard protocols (Ran et al., 2013). For the loss of CSR screen, 3-6 gRNAs per candidate were cloned and individually tested. CH12 cells were transfected with Cas9-gRNAs expressing constructs *via* electroporation with Neon Transfection System (Thermo Fisher Scientific), sorted for GFP-positive cells after 40 h, and left to recover for 72 h before activation for CSR analysis. CH12 and MEFs clonal derivatives were generated *via* electroporation with either single gRNA and WT Cas9-based plasmid (pX458) or gRNA pairs and nickase Cas9-based plasmid. In the latter case, tandem U6 cassettes were cloned into a mutated version of pX458 expressing $Cas9^{D10A}$. The location of the genomic sequences targeted by the gRNAs used for the generation of these clonal cell lines is indicated in Fig. 2A. Single GFP-positive cells were sorted in 96-well plates and clones were allowed to grow for 12 days (CH12) or 17 days (MEFs) days before expansion. Clones were validated at the level of genomic scar and protein expression. The sequences of the gRNAs employed in these studies are listed in Table 3.

Table 3: CRISPR-Cas9 gRNAs

gRNA	Sequence
gZmynd8-1	AGAAAACGGCCCCGAAACGG
gZmynd8-2	AAGTCATTCCGTCCGTCCTG
gZmynd8 (nickase pair 1a)	GTCTTGGGGCGAATGGCCAT
gZmynd8 (nickase pair 1b)	ATTAAAAAGAAAAAGAAACC
gZmynd8 (nickase pair 2a)	GACACTTAGCGTGATAAACC
gZmynd8 (nickase pair 2b)	GAGACTGACATCGGAGCCAG
gAID	TGAGACCTACCTCTGCTACG
Ctrl gRandom-1	GCGAGGTATTCGGCTCCGCG
Ctrl gRandom-2	ATGTTGCAGTTCGGCTCGAT
gDSB-1	AGTTGTCATTGCTGAATATC
gDSB-2	CATGGATTTCTCCGGTGAAT
g53BP1 (nickase pair 1a)	CAGATGTTTATTATGTGGAT
g53BP1 (nickase pair 1b)	GAGTGTACGGACTTCTCGAA
gRif1 (nickase pair 1a)	AAGTCTCCAGAAGCGGCTCC
gRif1 (nickase pair 1b)	GAAGACCCCTCGGTGCCTCC

Cell lysates and Co-immunoprecipitation assay

Co-immunoprecipitation analyses were performed as for I-DIRT pull-down with the only exception that protein elution from magnetic beads was performed by

incubation with NuPage LDS Sample buffer (Thermo Fisher) supplemented with 50mM DTT for 10 min at 70C. Where indicated, 10 μM ATMi KU55933 (Tocris Bioscience) was added 1h before irradiation. Western blot analysis of protein levels was performed on whole cell lysates prepared by lysis in RIPA buffer. The antibodies used for WB analysis are: anti-Rif1 (Di Virgilio et al., 2013), anti-Flag M2 (Sigma-Aldrich), anti-ZMYND8 (Sigma-Aldrich), anti-γH2AX (Millipore), anti-β Actin (Sigma-Aldrich), and anti-53BP1 (Bethyl).

Flow cytometry

For class switching assays, cell suspensions were stained with fluorochrome-conjugated anti-IgG1 (BD-pharmigen), or anti-IgA (Southern Biotech). Samples were acquired on a LSRFortessa cell analyzer (BD Biosciences) and analyzed with FlowJo software. Analysis of B cell development and differentiation was performed as previously described (Di Virgilio et al., 2013) using anti-CD21/CD35-FITC, anti-IgD-FITC, anti-IgM-PE, anti-IgM-FITC, anti-CD43-PE (BD Biosciences) and anti-CD23-PE, anti-CD3-PE and anti-CD19-APC (Biolegend) antibodies. For cell proliferation analysis, primary B cells were pulsed with 2 μM carboxyfluorescein succinimidyl ester (CFSE) or 5 μM CellTrace Violet (Thermofisher) for 10 min at 37C. For cell cycle analysis, CH12 cells were collected, fixed and permeabilized using Fixation/Permeabilization Solution Kit (BD Biosciences) according to the manufacturer's instructions. BrdU pulse and staining was performed by using BrdU Flow Kit (BD Biosciences) according to the manufacturer's instructions.

Quantitative PCR

mRNA levels for AID and germline transcripts were measured as previously described (Muramatsu et al., 2000; Pavri et al., 2010). 3'RR hs eRNA levels were measured using the following primers:

Forward 5'- CATTCCCATGGTTCTGGGTAG-3' and

Reverse 5'- CAAGAGGACATGACAGGAGATG -3' for hs1,2;

Forward 5'-CATTGAGCTCCGGCTCTAAC-3' and

Reverse 5'-CCCCTGTAGGGATCCTCCTAAT-3' for 5' hs3b;

Forward 5'-CATCCAGAGTCAAGGGGTGTC-3' and

Reverse 5'-CTAGAACCACATGCTATCTAAGGGA-3' for 3' hs3b.

Immunofluorescence

iMEFs were grown on cover slips overnight. Cells were irradiated (10 Gy IR) and allowed to recover for 90 min or 6 h. Upon fixation with 4% paraformaldehyde and permeabilization with 0.5% Triton X-100, cells were stained with anti-γH2AX (Millipore), rabbit anti-Rif1 serum (Di Virgilio et al., 2013), mouse anti-53BP1 (Upstate), mouse anti-Flag M2 (Sigma-Aldrich) or rabbit anti-ZMYND8 (Sigma-Aldrich) antibodies as primary antibodies, and with goat anti-rabbit Alexa546 and goat anti-mouse Alexa488 as secondary antibodies. DNA was counterstained with Pentahydrate (bis-Benzimide) (Hoechst). Images were acquired using inverted LSM700 laser scanning confocal microscope (Zeiss).

Clonogenic assay

For the colony formation assay following IR, iMEFs were plated in 60mm dishes, irradiated after 24 hours with the indicated doses and incubated at 37°C in the presence of 5% CO2. After 14 days, colonies were fixed with 15% acetic acid: methanol (v/v) for 5 min and stained with 0.5% Crystal Violet (Sigma) for 30 min for colony visualization. For the colony formation assay in the presence of PARPi (Sigma), 10 nM or 1 μM PARPi was added 24 after seeding and cells allowed to grow for 14 days before fixation, with fresh media and PARPi replenished on day 7. *BRCA1*$^{-/-}$ iMEFs were previously described (Bunting et al., 2010).

Metaphase analysis

MEFs were treated with 2 μM PARPi (Olaparib/AZD2281, Selleckchem) for 21 h, and metaphase preparation and aberration analysis were performed as previously described (Bunting et al., 2010).

End resection assay

Single guide RNAs targeting two sequences 2276 bp apart within the *ROSA26* locus (Table S1) were cloned into pX458 plasmid (gDSB-1/2). CH12 cells were electroporated with a 1:1 mix of gDSB-1 and gDSB-2 constructs using the Neon Transfection System (Thermo Fisher Scientific), and allowed to recover for 24 h before collection. Genomic DNA was extracted according to standard protocols

and individual repair junctions were amplified using nested PCR reactions. The first and second rounds of PCR were performed using

5'-CTGTTAGAGCATGCTTAAGGG-3' Forward and

5'-TCACCATTAGGGCAAATGGC-3' Reverse primers,

and

5'-GTAGTTACTTGGCAGGCTCC -3' Forward

and 5'-AAAGTCATTCCACAGTTTGAC -3' Reverse primers,

respectively. PCR products were extracted from agarose gel and sequenced.

CRISPR-Cas9-induced CSR assay

Constructs for expression of gRNA-Sμ and gRNA-Sα were generated by cloning single guide RNAs directed to the 5' Sμ and 3' Sα regions respectively of the *Igh* locus (Ramachandran et al., 2016) in tandem U6 cassettes on a modified pX458 plasmid backbone. Control construct was generated by cloning random sequences not targeting the mouse genome in tandem U6 cassettes on same plasmid. CH12 cells were electroporated with the constructs using the Neon Transfection System, and allowed to recover for 12 h before CSR analysis.

Mutational analysis

Primary cultures of splenocytes were collected at either 72 h or 96 h post-activation. gDNA extraction was performed according to standard protocols and

MutPE-Seq at 5' Sμ was performed as previously described (Robbiani et al., 2015; Wang et al., 2017).

ChIP-Seq

ChIP-Seq for ZMYND8 and RNA Pol II were performed following protocols previously described (Pavri et al., 2010; Yamane et al., 2013), and using anti-ZMYND8 rabbit polyclonal antibody (HPA020949, Sigma) and anti-RNA Pol II (4H8 abcam) antibody. FASTQ files were aligned against mouse genome (mm10) using BWA mem (Li & Durbin, 2009). Processing and peak-calling of ChIP-Seq data were accomplished with HOMER ChIP-Seq program (Heinz et al., 2010). Peak annotation was done using R and ChIPseeker package (Yu, Wang, & He, 2015).

RNA-Seq

Samples used in RNA-Seq were cultured in activated and unactivated conditions for 48 h. Cultured cells were collected by centrifugation and RNA was extracted with AllPrep DNA/RNA Mini Kit (Qiagen). Ribosomal RNA was depleted using Ribo-Zero Gold rRNA Removal Kit (Illumina). Libraries were prepared with TruSeq Stranded Total RNA Library Prep Kit (Illumina). Three biological repeats were performed and run in two lanes on the same flow cells on NextSeq High Output 75 SR (Illumina). For data analysis, reads were pseudo-aligned to an index created from the Ensembl mouse GRCm38.p5 assembly and custom annotations of *Igh* locus features. Transcript-level abundances were quantified using kallisto v0.43.0

(Bray, Pementel, Melsted, Pachter, 2016), and subsequently summarized to the gene-level using the R package tximport (Soneson, Love, Robinson, 2015). Differential gene expression analysis was performed using DESeq2 (Love, Huber, Anders, 2014).

4C-Seq

The 4C assay was performed as previously described (Qian et al., 2014). The bait was amplified using the following primers:

5′-ACCACAGAGACCTCTGGATC-3′;

5′-GCATGAGCTGCAGATGGTAC-3′.

Chapter 10: Works Cited

Abe M, Shiku H. Isolation of an IgH gene circular DNA clone from human bone marrow. Nucleic Acids Res. 1989;17(1):163-70.

Adamo A, Collis SJ, Adelman CA, et al. Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. Mol Cell. 2010;39(1):25-35.

Adhikary S, Sanyal S, Basu M, et al. Selective Recognition of H3.1K36 Dimethylation/H4K16 Acetylation Facilitates the Regulation of All-trans-retinoic Acid (ATRA)-responsive Genes by Putative Chromatin Reader ZMYND8. J Biol Chem. 2016;291(6):2664-81.

Agalioti T, Chen G, Thanos D. Deciphering the transcriptional histone acetylation code for a human gene. Cell. 2002;111(3):381-92.

Ahnesorg P, Smith P, Jackson SP. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell. 2006;124(2):301-13.

Allen BL, Taatjes DJ. The Mediator complex: a central integrator of transcription. Nat Rev Mol Cell Biol. 2015;16(3):155-66.

Alt FW, Rosenberg N, Casanova RJ, Thomas E, Baltimore D. Immunoglobulin heavy-chain expression and class switching in a murine leukaemia cell line. Nature. 1982;296(5855):325-31.

Alt FW, Zhang Y, Meng FL, Guo C, Schwer B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. Cell. 2013;152(3):417-29.

Alver RC, Chadha GS, Gillespie PJ, Blow JJ. Reversal of DDK-Mediated MCM Phosphorylation by Rif1-PP1 Regulates Replication Initiation and Replisome Stability Independently of ATR/Chk1. Cell Rep. 2017;18(10):2508-2520.

An L, Wang Y, Liu Y, et al. Rad9 is required for B cell proliferation and immunoglobulin class switch recombination. J Biol Chem. 2010;285(46):35267-73.

Aoufouchi S, Faili A, Zober C, et al. Proteasomal degradation restricts the nuclear lifespan of AID. J Exp Med. 2008;205(6):1357-68.

Arakawa H, Hauschild J, Buerstedde JM. Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. Science. 2002;295(5558):1301-6.

Arnold LW, Grdina TA, Whitmore AC, Haughton G. Ig isotype switching in B lymphocytes. Isolation and characterization of clonal variants of the murine Ly-1+ B cell lymphoma, CH12, expressing isotypes other than IgM. J Immunol. 1988;140(12):4355-63.

Aymard F, Bugler B, Schmidt CK, et al. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. Nat Struct Mol Biol. 2014;21(4):366-74.

Banerji J, Olson L, Schaffner W. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell. 1983;33(3):729-40.

Bardwell PD, Woo CJ, Wei K, et al. Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. Nat Immunol. 2004;5(2):224-9.

Barreto V, Reina-san-martin B, Ramiro AR, Mcbride KM, Nussenzweig MC. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. Mol Cell. 2003;12(2):501-8.

Basu U, Chaudhuri J, Alpert C, et al. The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. Nature. 2005;438(7067):508-11.

Basu U, Franklin A, Alt FW. Post-translational regulation of activation-induced cytidine deaminase. Philos Trans R Soc Lond, B, Biol Sci. 2009;364(1517):667-73.

Basu U, Meng FL, Keim C, et al. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. Cell. 2011;144(3):353-63.

Bébin AG, Carrion C, Marquet M, et al. In vivo redundant function of the 3' IgH regulatory element HS3b in the mouse. J Immunol. 2010;184(7):3710-7.

Bentley DL, Groudine M. A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. Nature. 1986;321(6071):702-6.

Bierkens M, Krijgsman O, Wilting SM, et al. Focal aberrations indicate EYA2 and hsa-miR-375 as oncogene and tumor suppressor in cervical carcinogenesis. Genes Chromosomes Cancer. 2013;52(1):56-68.

Birshtein BK. Epigenetic Regulation of Individual Modules of the immunoglobulin heavy chain locus 3' Regulatory Region. Front Immunol. 2014;5:163.

Bothmer A, Robbiani DF, Feldhahn N, Gazumyan A, Nussenzweig A, Nussenzweig MC. 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. J Exp Med. 2010;207(4):855-65.

Bothmer A, Robbiani DF, Di Virgilio M, et al. Regulation of DNA end joining, resection, and immunoglobulin class switch recombination by 53BP1. Mol Cell. 2011;42(3):319-29.

Bouwman P, Aly A, Escandell JM, et al. 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. Nat Struct Mol Biol. 2010;17(6):688-95.

Brack C, Hirama M, Lenhard-Schuller R, Tonegawa S. A Complete immunoglobulin gene is created by somatic recombination. Cell. 1978;(1):1-14.

Bransteitter R, Pham P, Scharff MD, Goodman MF. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. Proc Natl Acad Sci USA. 2003;100(7):4102-7.

Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34(5):525-7.

Britton S, Coates J, Jackson SP. A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. J Cell Biol. 2013;202(3):579-95.

Bunting SF, Callén E, Wong N, et al. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. Cell. 2010;141(2):243-54.

Buonomo SB, Wu Y, Ferguson D, De lange T. Mammalian Rif1 contributes to replication stress survival and homology-directed repair. J Cell Biol. 2009;187(3):385-98.

Buonomo SB. Heterochromatin DNA replication and Rif1. Exp Cell Res. 2010;316(12):1907-13.

Butler JS, Koutelou E, Schibler AC, Dent SY. Histone-modifying enzymes: regulators of developmental decisions and drivers of human disease. Epigenomics. 2012;4(2):163-77.

Bzymek M, Thayer NH, Oh SD, Kleckner N, Hunter N. Double Holliday junctions are intermediates of DNA break repair. Nature. 2010;464(7290):937-41.

Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why?. Mol Cell. 2013;49(5):825-37.

Casellas R, Basu U, Yewdell WT, Chaudhuri J, Robbiani DF, Di noia JM. Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. Nat Rev Immunol. 2016;16(3):164-76.

Catalan N, Selz F, Imai K, Revy P, Fischer A, Durandy A. The block in immunoglobulin class switch recombination caused by activation-induced cytidine deaminase deficiency occurs prior to the generation of DNA double strand breaks in switch mu region. J Immunol. 2003;171(5):2504-9.

Celeste A, Petersen S, Romanienko PJ, et al. Genomic instability in mice lacking histone H2AX. Science. 2002;296(5569):922-7.

Chahwan R, van Oers JM, Avdievich E, et al. The ATPase activity of MLH1 is required to orchestrate DNA double-strand breaks and end processing during class switch recombination. J Exp Med. 2012;209(4):671-8.

Chan DW, Lees-Miller SP. The DNA-dependent protein kinase is inactivated by autophosphorylation of the catalytic subunit. J Biol Chem. 1996;271(15):8936-41.

Chandra V, Bortnick A, Murre C. AID targeting: old mysteries and new challenges. Trends Immunol. 2015;36(9):527-35.

Chang LM, Bollum FJ. Molecular biology of terminal transferase. CRC Crit Rev Biochem. 1986;21(1):27-52.

Chanut P, Britton S, Coates J, Jackson SP, Calsou P. Coordinated nuclease activities counteract Ku at single-ended DNA double-strand breaks. Nat Commun. 2016;7:12889.

Chapman JR, Sossick AJ, Boulton SJ, Jackson SP. BRCA1-associated exclusion of 53BP1 from DNA damage sites underlies temporal control of DNA repair. J Cell Sci. 2012;125(Pt 15):3529-34.

Chapman JR, Barral P, Vannier JB, et al. RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. Mol Cell. 2013;49(5):858-71.

Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. Nature. 2003;422(6933):726-30.

Chen Y, Zhang B, Bao L, et al. ZMYND8 acetylation mediates HIF-dependent breast cancer progression and metastasis. J Clin Invest. 2018;128(5):1937-1955.

Chowdhury M, Forouhi O, Dayal S, et al. Analysis of intergenic transcription and histone modification across the human immunoglobulin heavy-chain locus. Proc Natl Acad Sci USA. 2008;105(41):15872-7.

Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. Mol Cell. 2010;40(2):179-204.

Cogné M, Lansford R, Bottaro A, et al. A class switch control region at the 3' end of the immunoglobulin heavy chain locus. Cell. 1994;77(5):737-47.

Cooper MD, Alder MN. The evolution of adaptive immune systems. Cell. 2006;124(4):815-22.

Cooper MP, Machwe A, Orren DK, Brosh RM, Ramsden D, Bohr VA. Ku complex interacts with and stimulates the Werner protein. Genes Dev. 2000;14(8):907-12.

Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science. 2008;322(5909):1845-8.

Cornacchia D, Dileep V, Quivy JP, et al. Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. EMBO J. 2012;31(18):3678-90.

Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008;26(12):1367-72.

CRISPR Design. (n.d.). Retrieved from crispr.mit.edu

Cui X, Yu Y, Gupta S, Cho YM, Lees-miller SP, Meek K. Autophosphorylation of DNA-dependent protein kinase regulates DNA end processing and may also alter double-strand break repair pathway choice. Mol Cell Biol. 2005;25(24):10842-52.

D'addabbo P, Scascitelli M, Giambra V, Rocchi M, Frezza D. Position and sequence conservation in Amniota of polymorphic enhancer HS1.2 within the palindrome of IgH 3'Regulatory Region. BMC Evol Biol. 2011;11:71.

Daley JM, Laan RL, Suresh A, Wilson TE. DNA joint dependence of pol X family polymerase action in nonhomologous end joining. J Biol Chem. 2005;280(32):29030-7.

Daniel JA, Nussenzweig A. The AID-induced DNA damage response in chromatin. Mol Cell. 2013;50(3):309-21.

Delgado-Benito V*, Rosen DB*, Wang Q, et al. The Chromatin Reader ZMYND8 Regulates Igh Enhancers to Promote Immunoglobulin Class Switch Recombination. Mol Cell. 2018;72(4):636-649.e8.
*These authors contributed equally

Di Noia J, Neuberger MS. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. Nature. 2002;419(6902):43-8.

Di Virgilio M, Callen E, Yamane A, et al. Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching. Science. 2013;339(6120):711-5.

Dickerson SK, Market E, Besmer E, Papavasiliou FN. AID mediates hypermutation by deaminating single stranded DNA. J Exp Med. 2003;197(10):1291-6.

Difilippantonio MJ, Zhu J, Chen HT, et al. DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. Nature. 2000;404(6777):510-4.

Douglas P, Cui X, Block WD, et al. The DNA-dependent protein kinase catalytic subunit is phosphorylated in vivo on threonine 3950, a highly conserved

amino acid in the protein kinase domain. Mol Cell Biol. 2007;27(5):1581-91.

Dubois MF, Nguyen VT, Bellier S, Bensaude O. Inhibitors of transcription such as 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole and isoquinoline sulfonamide derivatives (H-8 and H-7) promote dephosphorylation of the carboxyl-terminal domain of RNA polymerase II largest subunit. J Biol Chem. 1994;269(18):13331-6.

Dungrawala H, Rose KL, Bhat KP, et al. The Replication Checkpoint Prevents Two Types of Fork Collapse without Regulating Replisome Stability. Mol Cell. 2015;59(6):998-1010.

Dunnick W, Wilson M, Stavnezer J. Mutations, duplication, and deletion of recombined switch regions suggest a role for DNA replication in the immunoglobulin heavy-chain switch. Mol Cell Biol. 1989;9(5):1850-6.

Dunnick W, Hertz GZ, Scappino L, Gritzmacher C. DNA sequences at immunoglobulin switch region recombination sites. Nucleic Acids Res. 1993;21(3):365-72.

Eberharter A, Becker PB. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. EMBO Rep. 2002;3(3):224-9.

Eckhardt LA, Birshtein BK. Independent immunoglobulin class-switch events occurring in a single myeloma cell line. Mol Cell Biol. 1985;5(4):856-68.

Ehrenstein MR, Rada C, Jones AM, Milstein C, Neuberger MS. Switch junction sequences in PMS2-deficient mice reveal a microhomology-mediated mechanism of Ig class switch recombination. Proc Natl Acad Sci USA. 2001;98(25):14553-8.

Eichmuller S, Usener D, Dummer R, Stein A, Thiel D, Schadendorf D. Serological detection of cutaneous T-cell lymphoma-associated antigens. Proc Natl Acad Sci USA. 2001;98(2):629-34.

Escribano-díaz C, Orthwein A, Fradet-turcotte A, et al. A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. Mol Cell. 2013;49(5):872-83.

Farley FW, Soriano P, Steffen LS, Dymecki SM. Widespread recombinase expression using FLPeR (flipper) mice. Genesis. 2000;28(3-4):106-10.

Feng L, Fong KW, Wang J, Wang W, Chen J. RIF1 counteracts BRCA1-mediated end resection during DNA repair. J Biol Chem. 2013;288(16):11135-43.

Flynn RL, Zou L. ATR: a master conductor of cellular responses to DNA replication stress. Trends Biochem Sci. 2011;36(3):133-40.

Fossey SC, Kuroda S, Price JA, Pendleton JK, Freedman BI, Bowden DW. Identification and characterization of PRKCBP1, a candidate RACK-like protein. Mamm Genome. 2000;11(10):919-25.

Franken NA, Rodermond HM, Stap J, Haveman J, Van bree C. Clonogenic assay of cells in vitro. Nat Protoc. 2006;1(5):2315-9.

Fujimoto S, Yamagishi H. Isolation of an excision product of T-cell receptor alpha-chain gene rearrangements. Nature. 1987;327(6119):242-3.

Garot A, Marquet M, Saintamand A, et al. Sequential activation and distinct functions for distal and proximal modules within the IgH 3' regulatory region. Proc Natl Acad Sci USA. 2016;113(6):1618-23.

Gates LA, Shi J, Rohira AD, et al. Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation. J Biol Chem. 2017;292(35):14456-14472.

Giannini SL, Singh M, Calvo CF, Ding G, Birshtein BK. DNA regions flanking the mouse Ig 3' alpha enhancer are differentially methylated and DNAase I hypersensitive during B cell differentiation. J Immunol. 1993;150(5):1772-80.

Gillies SD, Morrison SL, Oi VT, Tonegawa S. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell. 1983;33(3):717-28.

Gilmour DS. Promoter proximal pausing on genes in metazoans. Chromosoma. 2009;118(1):1-10.

Glanville J, Zhai W, Berka J, Telman D, Huerta G, Mehta GR, Ni L, Mei L, Sundar PD, Day GM, Cox D, Rajpal A, Pons J. Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. Proc Natl Acad Sci USA. 2009;106(48):20216-21. doi: 10.1073/pnas.0909775106

Gong F, Chiu LY, Cox B, et al. Screen identifies bromodomain protein ZMYND8 in chromatin recognition of transcription-associated DNA damage that promotes homologous recombination. Genes Dev. 2015;29(2):197-211.

Goodarzi AA, Jeggo PA. The repair and signaling responses to DNA double-strand breaks. Adv Genet. 2013;82:1-45.

Gong F, Clouaire T, Aguirrebengoa M, Legube G, Miller KM. Histone demethylase KDM5A regulates the ZMYND8-NuRD chromatin remodeler to promote DNA repair. J Cell Biol. 2017;216(7):1959-1974.

Gottlieb TM, Jackson SP. The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. Cell. 1993;72(1):131-42.

Grawunder U, Wilm M, Wu X, et al. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature. 1997;388(6641):492-5.

Gregor PD, Morrison SL. Myeloma mutant with a novel 3' flanking region: loss of normal sequence and insertion of repetitive elements leads to decreased transcription but normal processing of the alpha heavy-chain gene products. Mol Cell Biol. 1986;6(6):1903-16.

Gross CT, Mcginnis W. DEAF-1, a novel protein that binds an essential region in a Deformed response element. EMBO J. 1996;15(8):1961-70.

Guikema JE, Linehan EK, Tsuchimoto D, et al. APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination. J Exp Med. 2007;204(12):3017-26.

Gursoy-yuzugullu O, Carman C, Price BD. Spatially restricted loading of BRD2 at DNA double-strand breaks protects H4 acetylation domains and promotes DNA repair. Sci Rep. 2017;7(1):12921.

Harlen KM, Churchman LS. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat Rev Mol Cell Biol. 2017;18(4):263-273.

Harriman GR, Bradley A, Das S, Rogers-fani P, Davis AC. IgA class switch in I alpha exon-deficient mice. Role of germline transcription in class switch recombination. J Clin Invest. 1996;97(2):477-85.

Häsler J, Rada C, Neuberger MS. Cytoplasmic activation-induced cytidine deaminase (AID) exists in stoichiometric complex with translation

elongation factor 1α (eEF1A). Proc Natl Acad Sci USA. 2011;108(45):18366-71.

Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576-89.

Helleday T. The underlying mechanism for the PARP and BRCA synthetic lethality: clearing up the misunderstandings. Mol Oncol. 2011;5(4):387-93.

Hickson I, Zhao Y, Richardson CJ, et al. Identification and characterization of a novel and specific inhibitor of the ataxia-telangiectasia mutated kinase ATM. Cancer Res. 2004;64(24):9152-9.

Hiraga S, Alvino GM, Chang F, et al. Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex. Genes Dev. 2014;28(4):372-83.

Honjo T, Kataoka T. Organization of immunoglobulin heavy chain genes and allelic deletion model. Proc Natl Acad Sci USA. 1978;75(5):2140-4.

Huen MS, Sy SM, Chen J. BRCA1 and its toolbox for the maintenance of genome integrity. Nat Rev Mol Cell Biol. 2010;11(2):138-48.

Huong PL, Kolk AH, Eggelte TA, Verstijnen CP, Gilis H, Hendriks JT. Measurement of antigen specific lymphocyte proliferation using 5-bromo-deoxyuridine incorporation. An easy and low cost alternative to radioactive thymidine incorporation. J Immunol Methods. 1991;140(2):243-8.

Iacovoni JS, Caron P, Lassadi I, et al. High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. EMBO J. 2010;29(8):1446-57.

Imai K, Slupphaug G, Lee WI, et al. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. Nat Immunol. 2003;4(10):1023-8.

Issaoui H, Ghazzaui N, Saintamand A, Carrion C, Oblet C, Denizot Y. The IgH 3' regulatory region super-enhancer does not control IgA class switch recombination in the B1 lineage. Cell Mol Immunol. 2018;15(3):289-291.

Iwasato T, Shimizu A, Honjo T, Yamagishi H. Circular DNA is excised by immunoglobulin class switch recombination. Cell. 1990;62(1):143-9.

Jackson SP. Sensing and repairing DNA double-strand breaks. Carcinogenesis. 2002;23(5):687-96.

Jackson SP, Bartek J. The DNA-damage response in human biology and disease. Nature. 2009;461(7267):1071-8.

Jeevan-raj BP, Robert I, Heyer V, et al. Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. J Exp Med. 2011;208(8):1649-60.

Johnson RD, Jasin M. Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. EMBO J. 2000;19(13):3398-407.

Jonkers I, Lis JT. Getting up to speed with transcription elongation by RNA polymerase II. Nat Rev Mol Cell Biol. 2015;16(3):167-77.

Jossé R, Martin SE, Guha R, et al. ATR inhibitors VE-821 and VX-970 sensitize cancer cells to topoisomerase i inhibitors by disabling DNA replication initiation and fork elongation responses. Cancer Res. 2014;74(23):6968-79.

Ju Z, Volpi SA, Hassan R, et al. Evidence for physical interaction between the immunoglobulin heavy chain variable region and the 3' regulatory region. J Biol Chem. 2007;282(48):35169-78.

Jung S, Rajewsky K, Radbruch A. Shutdown of class switch recombination by deletion of a switch region control element. Science. 1993;259(5097):984-7.

Kaetzel CS. The polymeric immunoglobulin receptor: bridging innate and adaptive immune responses at mucosal surfaces. Immunol Rev. 2005;206:83-99.

Kanno S, Kuzuoka H, Sasao S, et al. A novel human AP endonuclease with conserved zinc-finger-like motifs involved in DNA strand break responses. EMBO J. 2007;26(8):2094-103.

Karmodiya K, Krebs AR, Oulad-abdelghani M, Kimura H, Tora L. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. BMC Genomics. 2012;13:424.

Kenter AL, Feldman S, Wuerffel R, Achour I, Wang L, Kumar S. Three-dimensional architecture of the IgH locus facilitates class switch recombination. Ann N Y Acad Sci. 2012;1267:86-94.

Kim A, Han L, Santiago GE, Verdun RE, Yu K. Class-Switch Recombination in the Absence of the IgH 3' Regulatory Region. J Immunol. 2016;197(7):2930-5.

Kleiner RE, Verma P, Molloy KR, Chait BT, Kapoor TM. Chemical proteomics reveals a γH2AX-53BP1 interaction in the DNA damage response. Nat Chem Biol. 2015;11(10):807-14.

Kotnis A, Du L, Liu C, Popov SW, Pan-hammarström Q. Non-homologous end joining in class switch recombination: the beginning of the end. Philos Trans R Soc Lond, B, Biol Sci. 2009;364(1517):653-65.

Krebs AR, Imanci D, Hoerner L, Gaidatzis D, Burger L, Schübeler D. Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters. Mol Cell. 2017;67(3):411-422.e4.

Krumm A, Meulia T, Brunvand M, Groudine M. The block to transcriptional elongation within the human c-myc gene is determined in the promoter-proximal region. Genes Dev. 1992;6(11):2201-13.

Kuang FL, Luo Z, Scharff MD. H3 trimethyl K9 and H3 acetyl K9 chromatin modifications are associated with class switch recombination. Proc Natl Acad Sci USA. 2009;106(13):5288-93.

Kumar V, Alt FW, Frock RL. PAXX and XLF DNA repair factors are functionally redundant in joining DNA breaks in a G1-arrested progenitor B-cell line. Proc Natl Acad Sci USA. 2016;113(38):10619-24.

Kumaraswamy E, Shiekhattar R. Activation of BRCA1/BRCA2-associated helicase BACH1 is required for timely progression through S phase. Mol Cell Biol. 2007;27(19):6733-41.

Kunimoto DY, Harriman GR, Strober W. 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. J Immunol. 1988;141(3):713-20.

Küppers R. Mechanisms of B-cell lymphoma pathogenesis. Nat Rev Cancer. 2005;5(4):251-62.

Kurosawa Y, von Boehmer H, Haas W, Sakano H, Trauneker A, Tonegawa S. Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes. Nature. 1981;290(5807):565-70.

Kuroyanagi J, Shimada Y, Zhang B, et al. Zinc finger MYND-type containing 8 promotes tumour angiogenesis via induction of vascular endothelial growth factor-A expression. FEBS Lett. 2014;588(18):3409-16.

La Russa MF, Qi LS. The New State of the Art: Cas9 for Gene Activation and Repression. Mol Cell Biol. 2015;35(22):3800-9.

Lam KP, Kühn R, Rajewsky K. In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. Cell. 1997;90(6):1073-83.

Le MX, Haddad D, Ling AK, et al. Kin17 facilitates multiple double-strand break repair pathways that govern B cell class switching. Sci Rep. 2016;6:37215.

Le Noir S, Boyer F, Lecardeur S, et al. Functional anatomy of the immunoglobulin heavy chain 3′ super-enhancer needs not only core enhancer elements but also their unique DNA context. Nucleic Acids Res. 2017;45(10):5829-5837.

Lebman DA, Lee FD, Coffman RL. Mechanism for transforming growth factor beta and IL-2 enhancement of IgA expression in lipopolysaccharide-stimulated B cell cultures. J Immunol. 1990;144(3):952-9.

Lee KJ, Saha J, Sun J, et al. Phosphorylation of Ku dictates DNA double-strand break (DSB) repair pathway choice in S phase. Nucleic Acids Res. 2016;44(4):1732-45.

Leung CC, Glover JN. BRCT domains: easy as one, two, three. Cell Cycle. 2011;10(15):2461-70.

Li G, Pone EJ, Tran DC, et al. Iron inhibits activation-induced cytidine deaminase enzymatic activity and modulates immunoglobulin class switch DNA recombination. J Biol Chem. 2012;287(25):21520-9.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.

Li N, Li Y, Lv J, et al. ZMYND8 Reads the Dual Histone Mark H3K4me1-H3K14ac to Antagonize the Expression of Metastasis-Linked Genes. Mol Cell. 2016;63(3):470-84.

Li Q, Peterson KR, Fang X, Stamatoyannopoulos G. Locus control regions. Blood. 2002;100(9):3077-86.

Li W, Notani D, Rosenfeld MG. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat Rev Genet. 2016;17(4):207-23.

Li Z, Otevrel T, Gao Y, et al. The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. Cell. 1995;83(7):1079-89.

Li Z, Scherer SJ, Ronai D, et al. Examination of Msh6- and Msh3-deficient mice in class switching reveals overlapping and distinct roles of MutS homologues in antibody diversification. J Exp Med. 2004;200(1):47-59.

Lieber MR. The mechanism of human nonhomologous DNA end joining. J Biol Chem. 2008;283(1):1-5.

Lieberson R, Giannini SL, Birshtein BK, Eckhardt LA. An enhancer at the 3' end of the mouse immunoglobulin heavy chain locus. Nucleic Acids Res. 1991;19(4):933-7.

Lis JT. Imaging Drosophila gene activation and polymerase pausing in vivo. Nature. 2007;450(7167):198-202.

Liu Y, Chen W, Gaudet J, et al. Structural basis for recognition of SMRT/N-CoR by the MYND domain and its contribution to AML1/ETO's activity. Cancer Cell. 2007;11(6):483-97.

Lo KW, Kan HM, Chan LN, et al. The 8-kDa dynein light chain binds to p53-binding protein 1 and mediates DNA damage-induced p53 nuclear accumulation. J Biol Chem. 2005;280(9):8172-9.

Lottersberger F, Bothmer A, Robbiani DF, Nussenzweig MC, de Lange T. Role of 53BP1 oligomerization in regulating double-strand break repair. Proc Natl Acad Sci USA. 2013;110(6):2146-51.

Lou Z, Minter-dykhouse K, Franco S, et al. MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals. Mol Cell. 2006;21(2):187-200.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

Lutzker S, Rothman P, Pollock R, Coffman R, Alt FW. Mitogen- and IL-4-regulated expression of germ-line Ig gamma 2b transcripts: evidence for directed heavy chain class switching. Cell. 1988;53(2):177-84.

Lyons AB, Parish CR. Determination of lymphocyte division by flow cytometry. J Immunol Methods. 1994;171(1):131-7.

Ma Y, Pannicke U, Schwarz K, Lieber MR. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell. 2002;108(6):781-94.

Ma Y, Lu H, Tippin B, et al. A biochemically defined system for mammalian nonhomologous DNA end joining. Mol Cell. 2004;16(5):701-13.

Madisen L, Groudine M. Identification of a locus control region in the immunoglobulin heavy-chain locus that deregulates c-myc expression in plasmacytoma and Burkitt's lymphoma cells. Genes Dev. 1994;8(18):2212-26.

Magaud JP, Sargent I, Clarke PJ, Ffrench M, Rimokh R, Mason DY. Double immunocytochemical labeling of cell and tissue samples with monoclonal anti-bromodeoxyuridine. J Histochem Cytochem. 1989;37(10):1517-27.

Malovannaya A, Lanz RB, Jung SY, et al. Analysis of the human endogenous coregulator complexome. Cell. 2011;145(5):787-99.

Manis JP, Van der stoep N, Tian M, et al. Class switching in B cells lacking 3' immunoglobulin heavy chain enhancers. J Exp Med. 1998;188(8):1421-31.

Manis JP, Morales JC, Xia Z, Kutok JL, Alt FW, Carpenter PB. 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. Nat Immunol. 2004;5(5):481-7.

Marcand S, Gilson E, Shore D. A protein-counting mechanism for telomere length regulation in yeast. Science. 1997;275(5302):986-90.

Maréchal A, Zou L. DNA damage sensing by the ATM and ATR kinases. Cold Spring Harb Perspect Biol. 2013;5(9)

Mårtensson IL, Almqvist N, Grimsholm O, Bernardi AI. The pre-B cell receptor checkpoint. FEBS Lett. 2010;584(12):2572-9.

Martin A, Li Z, Lin DP, et al. Msh2 ATPase activity is essential for somatic hypermutation at a-T basepairs and for efficient class switch recombination. J Exp Med. 2003;198(8):1171-8.

Martomo SA, Yang WW, Gearhart PJ. A role for Msh6 but not Msh3 in somatic hypermutation and class switch recombination. J Exp Med. 2004;200(1):61-8.

Masani S, Han L, Yu K. Apurinic/apyrimidinic endonuclease 1 is the essential nuclease during immunoglobulin class switch recombination. Mol Cell Biol. 2013;33(7):1468-73.

Matsuoka M, Yoshida K, Maeda T, Usuda S, Sakano H. Switch circular DNA formed in cytokine-treated mouse splenocytes: evidence for intramolecular DNA deletion in immunoglobulin class switching. Cell. 1990;62(1):135-42.

Mattarocci S, Reinert JK, Bunker RD, et al. Rif1 maintains telomeres and mediates DNA repair by encasing DNA ends. Nat Struct Mol Biol. 2017;24(7):588-595.

Matthews AJ, Zheng S, Dimenna LJ, Chaudhuri J. Regulation of immunoglobulin class-switch recombination: choreography of noncoding transcription, targeted DNA deamination, and long-range DNA repair. Adv Immunol. 2014;122:1-57.

Matthias P, Baltimore D. The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the alpha constant region. Mol Cell Biol. 1993;13(3):1547-53.

Mcbride KM, Gazumyan A, Woo EM, et al. Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. Proc Natl Acad Sci USA. 2006;103(23):8798-803.

Mccormack WT, Tjoelker LW, Carlson LM, et al. Chicken IgL gene rearrangement involves deletion of a circular episome and addition of single nonrandom nucleotides to both coding segments. Cell. 1989;56(5):785-91.

Meek K, Douglas P, Cui X, Ding Q, Lees-miller SP. trans Autophosphorylation at DNA-dependent protein kinase's two major autophosphorylation site clusters facilitates end processing but not end joining. Mol Cell Biol. 2007;27(10):3881-90.

Menon V, Povirk LF. XLF/Cernunnos: An important but puzzling participant in the nonhomologous end joining DNA repair pathway. DNA Repair (Amst). 2017;58:29-37.

Methot SP, Di Noia JM. Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination. Adv Immunol. 2017;133:37-87.

Michaelson JS, Giannini SL, Birshtein BK. Identification of 3' alpha-hs4, a novel Ig heavy chain enhancer element regulated at multiple stages of B cell differentiation. Nucleic Acids Res. 1995;23(6):975-81.

Morales JC, Xia Z, Lu T, et al. Role for the BRCA1 C-terminal repeats (BRCT) protein 53BP1 in maintaining genomic stability. J Biol Chem. 2003;278(17):14971-7.

Muramatsu M, Sankaranand VS, Anant S, et al. 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. 1999;274(26):18470-6.

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell. 2000;102(5):553-63.

Nagaoka H, Muramatsu M, Yamamura N, Kinoshita K, Honjo T. Activation-induced deaminase (AID)-directed hypermutation in the immunoglobulin Smu region: implication of AID involvement in a common step of class switch recombination and somatic hypermutation. J Exp Med. 2002;195(4):529-34.

Nakamura K, Sakai W, Kawamoto T, et al. Genetic dissection of vertebrate 53BP1: a major role in non-homologous end joining of DNA double strand breaks. DNA Repair (Amst). 2006;5(6):741-9.

Nakamura M, Kondo S, Sugai M, Nazarea M, Imamura S, Honjo T. High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. Int Immunol. 1996;8(2):193-201.

Nambu Y, Sugai M, Gonda H, et al. Transcription-coupled events associating with immunoglobulin switch region chromatin. Science. 2003;302(5653):2137-40.

Neal JA, Meek K. Choosing the right path: does DNA-PK help make the decision?. Mutat Res. 2011;711(1-2):73-86.

Neuberger MS. Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells. EMBO J. 1983;2(8):1373-8.

Nick Mcelhinny SA, Ramsden DA. Polymerase mu is a DNA-directed DNA/RNA polymerase. Mol Cell Biol. 2003;23(7):2309-15.

Nikolajczyk BS, Dang W, Sen R. Mechanisms of mu enhancer regulation in B lymphocytes. Cold Spring Harb Symp Quant Biol. 1999;64:99-107.

Nimmerjahn F, Ravetch JV. Fcgamma receptors as regulators of immune responses. Nat Rev Immunol. 2008;8(1):34-47.

Nussenzweig A, Nussenzweig MC. Origin of chromosomal translocations in lymphoid cancer. Cell. 2010;141(1):27-38.

O'Neil D, Glowatz H, Schlumpberger M. Ribosomal RNA depletion for efficient use of RNA-seq capacity. Curr Protoc Mol Biol. 2013;Chapter 4:Unit 4.19.

Okazaki IM, Kinoshita K, Muramatsu M, Yoshikawa K, Honjo T. The AID enzyme induces class switch recombination in fibroblasts. Nature. 2002;416(6878):340-5.

Okazaki K, Davis DD, Sakano H. T cell receptor beta gene sequences in the circular DNA of thymocyte nuclei: direct evidence for intramolecular DNA deletion in V-D-J joining. Cell. 1987;49(4):477-85.

Okazaki K, Sakano H. Thymocyte circular DNA excised from T cell receptor alpha-delta gene complex. EMBO J. 1988;7(6):1669-74.

Onodera CS, Underwood JG, Katzman S, et al. Gene isoform specificity through enhancer-associated antisense transcription. PLoS ONE. 2012;7(8):e43511.

Panagopoulos I, Micci F, Thorsen J, et al. Fusion of ZMYND8 and RELA genes in acute erythroid leukemia. PLoS ONE. 2013;8(5):e63663.

Panier S, Boulton SJ. Double-strand break repair: 53BP1 comes into focus. Nat Rev Mol Cell Biol. 2014;15(1):7-18.

Pavri R, Gazumyan A, Jankovic M, et al. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. Cell. 2010;143(1):122-33.

Pavri R, Nussenzweig MC. AID targeting in antibody diversity. Adv Immunol. 2011;110:1-26.

Pefanis E, Wang J, Rothschild G, et al. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. Nature. 2014;514(7522):389-93.

Pefanis E, Wang J, Rothschild G, et al. RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. Cell. 2015;161(4):774-89.

Peña-Diaz J, Jiricny J. Mammalian mismatch repair: error-free or error-prone?. Trends Biochem Sci. 2012;37(5):206-14.

Perlot T, Alt FW, Bassing CH, Suh H, Pinaud E. Elucidation of IgH intronic enhancer functions via germ-line deletion. Proc Natl Acad Sci USA. 2005;102(40):14362-7.

Perlot T, Alt FW. Cis-regulatory elements and epigenetic changes control genomic rearrangements of the IgH locus. Adv Immunol. 2008;99:1-32.

Perry JJ, Yannone SM, Holden LG, et al. WRN exonuclease structure and molecular mechanism imply an editing role in DNA end processing. Nat Struct Mol Biol. 2006;13(5):414-22.

Petermann E, Helleday T. Pathways of mammalian replication fork restart. Nat Rev Mol Cell Biol. 2010;11(10):683-7.

Peters A, Storb U. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. Immunity. 1996;4(1):57-65.

Petersen S, Casellas R, Reina-san-martin B, et al. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. Nature. 2001;414(6864):660-665.

Petersen-Mahrt SK, Harris RS, Neuberger MS. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. Nature. 2002;418(6893):99-103.

Pettersson S, Cook GP, Brüggemann M, Williams GT, Neuberger MS. A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus. Nature. 1990;344(6262):165-8.

Pham P, Bransteitter R, Petruska J, Goodman MF. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. Nature. 2003;424(6944):103-7.

Pierce AJ, Johnson RD, Thompson LH, Jasin M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. Genes Dev. 1999;13(20):2633-8.

Pinaud E, Khamlichi AA, Le morvan C, et al. Localization of the 3' IgH locus elements that effect long-distance regulation of class switch recombination. Immunity. 2001;15(2):187-99.

Pinaud E, Marquet M, Fiancette R, et al. The IgH locus 3' regulatory region: pulling the strings from behind. Adv Immunol. 2011;110:27-70.

Platt RJ, Chen S, Zhou Y, et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell. 2014;159(2):440-55.

Poleshko A, Einarson MB, Shalginskikh N, et al. Identification of a functional network of human epigenetic silencing factors. J Biol Chem. 2010;285(1):422-33.

Polo SE, Jackson SP. Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. Genes Dev. 2011;25(5):409-33.

Porstmann T, Ternynck T, Avrameas S. Quantitation of 5-bromo-2-deoxyuridine incorporation into DNA: an enzyme immunoassay for the assessment of the lymphoid cell proliferative response. J Immunol Methods. 1985;82(1):169-79.

Povirk LF, Zhou T, Zhou R, Cowan MJ, Yannone SM. Processing of 3'-phosphoglycolate-terminated DNA double strand breaks by Artemis nuclease. J Biol Chem. 2007;282(6):3547-58.

Qian J, Wang Q, Dose M, et al. B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. Cell. 2014;159(7):1524-37.

Quah BJ, Warren HS, Parish CR. Monitoring lymphocyte proliferation in vitro and in vivo with the intracellular fluorescent dye carboxyfluorescein diacetate succinimidyl ester. Nat Protoc. 2007;2(9):2049-56.

Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. Curr Biol. 2002;12(20):1748-55.

Rahl PB, Lin CY, Seila AC, et al. c-Myc regulates transcriptional pause release. Cell. 2010;141(3):432-45.

Ramachandran S, Haddad D, Li C, et al. The SAGA Deubiquitination Module Promotes DNA Repair and Class Switch Recombination through ATM and DNAPK-Mediated γH2AX Formation. Cell Rep. 2016;15(7):1554-1565.

Ramadan K, Shevelev IV, Maga G, Hübscher U. De novo DNA synthesis by human DNA polymerase lambda, DNA polymerase mu and terminal deoxyribonucleotidyl transferase. J Mol Biol. 2004;339(2):395-404.

Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. Nat Immunol. 2003;4(5):452-6.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281-2308.

Raschke EE, Albert T, Eick D. Transcriptional regulation of the Ig kappa gene by promoter-proximal pausing of RNA polymerase II. J Immunol. 1999;163(8):4375-82.

Reina-san-martin B, Chen HT, Nussenzweig A, Nussenzweig MC. ATM is required for efficient recombination between immunoglobulin switch regions. J Exp Med. 2004;200(9):1103-10.

Reina-san-martin B, Chen J, Nussenzweig A, Nussenzweig MC. Enhanced intra-switch region recombination during immunoglobulin class switch recombination in 53BP1-/- B cells. Eur J Immunol. 2007;37(1):235-9.

Revy P, Muto T, Levy Y, et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell. 2000;102(5):565-75.

Reynolds P, Anderson JA, Harper JV, et al. The dynamics of Ku70/80 and DNA-PKcs at DSBs induced by ionizing radiation is dependent on the complexity of damage. Nucleic Acids Res. 2012;40(21):10821-31.

Riballo E, Woodbine L, Stiff T, Walker SA, Goodarzi AA, Jeggo PA. XLF-Cernunnos promotes DNA ligase IV-XRCC4 re-adenylation following ligation. Nucleic Acids Res. 2009;37(2):482-92.

Rickert RC, Roes J, Rajewsky K. B lymphocyte-specific, Cre-mediated mutagenesis in mice. Nucleic Acids Res. 1997;25(6):1317-8.

Robbiani DF, Deroubaix S, Feldhahn N, et al. Plasmodium Infection Promotes Genomic Instability and AID-Dependent B Cell Lymphoma. Cell. 2015;162(4):727-37.

Rogozin IB, Kolchanov NA. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. Biochim Biophys Acta. 1992;1171(1):11-8.

Rohwer F, Todd S, Mcguire KL. The effect of IL-2 treatment on transcriptional attenuation in proto-oncogenes pim-1 and c-myb in human thymic blast cells. J Immunol. 1996;157(2):643-9.

Ronai D, Iglesias-ussel MD, Fan M, Li Z, Martin A, Scharff MD. Detection of chromatin-associated single-stranded DNA in regions targeted for somatic hypermutation. J Exp Med. 2007;204(1):181-90.

Rouaud P, Vincent-fabert C, Fiancette R, Cogné M, Pinaud E, Denizot Y. Enhancers located in heavy chain regulatory region (hs3a, hs1,2, hs3b, and hs4) are dispensable for diversity of VDJ recombination. J Biol Chem. 2012;287(11):8356-60.

Rouaud P, Vincent-fabert C, Saintamand A, et al. The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. J Exp Med. 2013;210(8):1501-7.

Rush JS, Fugmann SD, Schatz DG. Staggered AID-dependent DNA double strand breaks are the predominant DNA lesions targeted to S mu in Ig class switch recombination. Int Immunol. 2004;16(4):549-57.

Saintamand A, Rouaud P, Saad F, Rios G, Cogné M, Denizot Y. Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. Nat Commun. 2015;6:7084.

Sanchez R, Zhou MM. The PHD finger: a versatile epigenome reader. Trends Biochem Sci. 2011;36(7):364-72.

Santos-rosa H, Schneider R, Bannister AJ, et al. Active genes are tri-methylated at K4 of histone H3. Nature. 2002;419(6905):407-11.

Savitsky P, Krojer T, Fujisawa T, et al. Multivalent Histone and DNA Engagement by a PHD/BRD/PWWP Triple Reader Cassette Recruits ZMYND8 to K14ac-Rich Chromatin. Cell Rep. 2016;17(10):2724-2737.

Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. Nat Rev Immunol. 2011;11(4):251-63.

Schrader CE, Edelmann W, Kucherlapati R, Stavnezer J. Reduced isotype switching in splenic B cells from mice deficient in mismatch repair enzymes. J Exp Med. 1999;190(3):323-30.

Schrader CE, Vardo J, Stavnezer J. Role for mismatch repair proteins Msh2, Mlh1, and Pms2 in immunoglobulin class switching shown by sequence analysis of recombination junctions. J Exp Med. 2002;195(3):367-73.

Schrader CE, Bradley SP, Vardo J, Mochegova SN, Flanagan E, Stavnezer J. Mutations occur in the Ig Smu region but rarely in Sgamma regions prior to class switch recombination. EMBO J. 2003;22(21):5893-903.

Schrader CE, Linehan EK, Mochegova SN, Woodland RT, Stavnezer J. Inducible DNA breaks in Ig S regions are dependent on AID and UNG. J Exp Med. 2005;202(4):561-8.

Sellars M, Reina-San-Martin B, Kastner P, Chan S. Ikaros controls isotype selection during immunoglobulin class switch recombination. J Exp Med. 2009;206(5):1073-87.

Sernández IV, De yébenes VG, Dorsett Y, Ramiro AR. Haploinsufficiency of activation-induced deaminase for antibody diversification and chromosome translocations both in vitro and in vivo. PLoS ONE. 2008;3(12):e3927.

Severinson E, Fernandez C, Stavnezer J. Induction of germ-line immunoglobulin heavy chain transcripts by mitogens and interleukins prior to switch recombination. Eur J Immunol. 1990;20(5):1079-84.

Sfeir A, Symington LS. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway?. Trends Biochem Sci. 2015;40(11):701-714.

Shen H, Xu W, Guo R, et al. Suppression of Enhancer Overactivation by a RACK7-Histone Demethylase Complex. Cell. 2016;165(2):331-42.

Shen HM, Tanaka A, Bozek G, Nicolae D, Storb U. Somatic hypermutation and class switch recombination in Msh6(-/-)Ung(-/-) double-knockout mice. J Immunol. 2006;177(8):5386-92.

Shinkura R, Tian M, Smith M, Chua K, Fujiwara Y, Alt FW. The influence of transcriptional orientation on endogenous switch region function. Nat Immunol. 2003;4(5):435-41.

Shockett P, Stavnezer J. Effect of cytokines on switching to IgA and alpha germline transcripts in the B lymphoma I.29 mu. Transforming growth factor-beta activates transcription of the unrearranged C alpha gene. J Immunol. 1991;147(12):4374-4383

Silverman J, Takai H, Buonomo SB, Eisenhaber F, De lange T. Human Rif1, ortholog of a yeast telomeric protein, is regulated by ATM and 53BP1 and functions in the S-phase checkpoint. Genes Dev. 2004;18(17):2108-19.

Skaar JR, Ferris AL, Wu X, et al. The Integrator complex controls the termination of transcription at diverse classes of gene targets. Cell Res. 2015;25(3):288-305.

Snapper CM, Paul WE. Interferon-gamma and B cell stimulatory factor-1 reciprocally regulate Ig isotype production. Science. 1987;236(4804):944-7.

Sohail A, Klapacz J, Samaranayake M, Ullah A, Bhagwat AS. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. Nucleic Acids Res. 2003;31(12):2990-4.

Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. 2015;4:1521.

Soubeyrand S, Pope L, Pakuts B, Haché RJ. Threonines 2638/2647 in DNA-PK are essential for cellular resistance to ionizing radiation. Cancer Res. 2003;63(6):1198-201.

Spruijt CG, Luijsterburg MS, Menafra R, et al. ZMYND8 Co-localizes with NuRD on Target Genes and Regulates Poly(ADP-Ribose)-Dependent Recruitment of GATAD2A/NuRD to Sites of DNA Damage. Cell Rep. 2016;17(3):783-798.

Sreesankar E, Senthilkumar R, Bharathi V, Mishra RK, Mishra K. Functional diversification of yeast telomere associated protein, Rif1, in higher eukaryotes. BMC Genomics. 2012;13:255.

Srinivasan L, Sasaki Y, Calado DP, et al. PI3 kinase signals BCR-dependent mature B cell survival. Cell. 2009;139(3):573-86.

Stagi S, Gulino AV, Lapi E, Rigante D. Epigenetic control of the immune system: a lesson from Kabuki syndrome. Immunol Res. 2016;64(2):345-59.

Stanlie A, Aida M, Muramatsu M, Honjo T, Begum NA. Histone3 lysine4 trimethylation regulated by the facilitates chromatin transcription complex is critical for DNA cleavage in class switch recombination. Proc Natl Acad Sci USA. 2010;107(51):22190-5.

Starnes LM, Su D, Pikkupeura LM, et al. A PTIP-PA1 subcomplex promotes transcription for IgH class switching independently from the associated MLL3/MLL4 methyltransferase complex. Genes Dev. 2016;30(2):149-63.

Stavnezer J, Radcliffe G, Lin YC, et al. Immunoglobulin heavy-chain switching may be directed by prior induction of transcripts from constant-region genes. Proc Natl Acad Sci USA. 1988;85(20):7704-8.

Stavnezer J, Schrader CE. Mismatch repair converts AID-instigated nicks to double-strand breaks for antibody class-switch recombination. Trends Genet. 2006;22(1):23-8.

Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. Annu Rev Immunol. 2008;26:261-92.

Stavnezer-nordgren J, Sirlin S. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. EMBO J. 1986;5(1):95-102.

Stewart GS, Wang B, Bignell CR, Taylor AM, Elledge SJ. MDC1 is a mediator of the mammalian DNA damage checkpoint. Nature. 2003;421(6926):961-6.

Sturzenegger A, Burdova K, Kanagaraj R, et al. DNA2 cooperates with the WRN and BLM RecQ helicases to mediate long-range DNA end resection in human cells. J Biol Chem. 2014;289(39):27314-26.

Sukackaite R, Jensen MR, Mas PJ, Blackledge M, Buonomo SB, Hart DJ. Structural and biophysical characterization of murine rif1 C terminus reveals high specificity for DNA cruciform structures. J Biol Chem. 2014;289(20):13903-11.

Sukackaite R, Cornacchia D, Jensen MR, et al. Mouse Rif1 is a regulatory subunit of protein phosphatase 1 (PP1). Sci Rep. 2017;7(1):2119.

Symington LS, Gautier J. Double-strand break end resection and repair pathway choice. Annu Rev Genet. 2011;45:247-71.

Symington LS. End resection at double-strand breaks: mechanism and regulation. Cold Spring Harb Perspect Biol. 2014;6(8)

Takizawa M, Tolarová H, Li Z, et al. AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. J Exp Med. 2008;205(9):1949-57.

Tanaka H, Muto A, Shima H, et al. Epigenetic Regulation of the Blimp-1 Gene (Prdm1) in B Cells Involves Bach2 and Histone Deacetylase 3. J Biol Chem. 2016;291(12):6316-30.

Teng G, Hakimpour P, Landgraf P, et al. MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. Immunity. 2008;28(5):621-9.

Teo SH, Jackson SP. Identification of Saccharomyces cerevisiae DNA ligase IV: involvement in DNA double-strand break repair. EMBO J. 1997;16(15):4788-95.

Thomas-Claudepierre AS, Robert I, Rocha PP, et al. Mediator facilitates transcriptional activation and dynamic long-range contacts at the IgH locus during class switch recombination. J Exp Med. 2016;213(3):303-12.

Toda M, Hirama T, Takeshita S, Yamagishi H. Excision products of immunoglobulin gene rearrangements. Immunol Lett. 1989;21(4):311-6.

Tonegawa S. Somatic generation of antibody diversity. Nature. 1983;302(5909):575-81.

Tsai AG, Lieber MR. Mechanisms of chromosomal rearrangement in the human genome. BMC Genomics. 2010;11 Suppl 1:S1.

Uchimura Y, Barton LF, Rada C, Neuberger MS. REG-γ associates with and modulates the abundance of nuclear activation-induced deaminase. J Exp Med. 2011;208(12):2385-91.

Uematsu N, Weterings E, Yano K, et al. Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. J Cell Biol. 2007;177(2):219-29.

Vincent-fabert C, Truffinet V, Fiancette R, Cogné N, Cogné M, Denizot Y. Ig synthesis and class switching do not require the presence of the hs4 enhancer in the 3' IgH regulatory region. J Immunol. 2009;182(11):6926-32.

Vincent-fabert C, Fiancette R, Pinaud E, et al. Genomic deletion of the whole IgH 3' regulatory region (hs3a, hs1,2, hs3b, and hs4) dramatically affects class switch recombination and Ig secretion to all isotypes. Blood. 2010;116(11):1895-8.

von Schwedler U, Jäck HM, Wabl M. Circular DNA is a product of the immunoglobulin class switch rearrangement. Nature. 1990;345(6274):452-6.

Walker JR, Corpina RA, Goldberg J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature. 2001;412(6847):607-14.

Wang L, Whang N, Wuerffel R, Kenter AL. AID-dependent histone acetylation is detected in immunoglobulin S regions. J Exp Med. 2006;203(1):215-26.

Wang L, Wuerffel R, Feldman S, Khamlichi AA, Kenter AL. S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. J Exp Med. 2009;206(8):1817-30.

Wang Q, Kieffer-kwon KR, Oliveira TY, et al. The cell cycle restricts activation-induced cytidine deaminase activity to early G1. J Exp Med. 2017;214(1):49-58.

Ward IM, Minn K, van Deursen J, Chen J. p53 Binding protein 53BP1 is required for DNA damage responses and tumor suppression in mice. Mol Cell Biol. 2003;23(7):2556-63.

Ward IM, Reina-san-martin B, Olaru A, et al. 53BP1 is required for class switch recombination. J Cell Biol. 2004;165(4):459-64.

Westendorf JJ, Koka S. Identification of FHOD1-binding proteins and mechanisms of FHOD1-regulated actin dynamics. J Cell Biochem. 2004;92(1):29-41.

Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013;153(2):307-19.

Wiedemann EM, Peycheva M, Pavri R. DNA Replication Origins in Immunoglobulin Switch Regions Regulate Class Switch Recombination in an R-Loop-Dependent Manner. Cell Rep. 2016;17(11):2927-2942.

Wilson TE, Grawunder U, Lieber MR. Yeast DNA ligase IV mediates non-homologous DNA end joining. Nature. 1997;388(6641):495-8.

Won WJ, Kearney JF. CD9 is a unique marker for marginal zone B cells, B1 cells, and plasma cells in mice. J Immunol. 2002;168(11):5605-11.

Wotton D, Shore D. A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in Saccharomyces cerevisiae. Genes Dev. 1997;11(6):748-60.

Wuerffel RA, Du J, Thompson RJ, Kenter AL. Ig Sgamma3 DNA-specifc double strand breaks are induced in mitogen-activated B cells and are implicated in switch recombination. J Immunol. 1997;159(9):4139-44.

Wuerffel R, Wang L, Grigera F, et al. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. Immunity. 2007;27(5):711-22.

Xu D, Muniandy P, Leo E, et al. Rif1 provides a new DNA-binding interface for the Bloom syndrome complex to maintain normal replication. EMBO J. 2010;29(18):3140-55.

Xu L, Blackburn EH. Human Rif1 protein binds aberrant telomeres and aligns along anaphase midzone microtubules. J Cell Biol. 2004;167(5):819-30.

Xue K, Rada C, Neuberger MS. The in vivo pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in msh2-/- ung-/- mice. J Exp Med. 2006;203(9):2085-94.

Yamane A, Resch W, Kuo N, et al. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. Nat Immunol. 2011;12(1):62-9.

Yamazaki S, Hayano M, Masai H. Replication timing regulation of eukaryotic replicons: Rif1 as a global regulator of replication timing. Trends Genet. 2013;29(8):449-60.

Yancopoulos GD, Depinho RA, Zimmerman KA, Lutzker SG, Rosenberg N, Alt FW. Secondary genomic rearrangement events in pre-B cells: VHDJH replacement by a
LINE-1 sequence and directed class switching. EMBO J. 1986;5(12):3259-66.

Yewdell WT, Chaudhuri J. A transcriptional serenAID: the role of noncoding RNAs in class switch recombination. Int Immunol. 2017;29(4):183-196.

Yoshida K, Matsuoka M, Usuda S, Mori A, Ishizaka K, Sakano H. Immunoglobulin switch circular DNA in the mouse infected with Nippostrongylus brasiliensis: evidence for successive class switching from mu to epsilon via gamma 1. Proc Natl Acad Sci USA. 1990;87(20):7829-33.

You Z, Chahwan C, Bailis J, Hunter T, Russell P. ATM activation and its recruitment to damaged DNA require binding to the C terminus of Nbs1. Mol Cell Biol. 2005;25(13):5363-79.

Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015;31(14):2382-3.

Yu H, Jiang Y, Liu L, et al. Integrative genomic and transcriptomic analysis for pinpointing recurrent alterations of plant homeodomain genes and their clinical significance in breast cancer. Oncotarget. 2017;8(8):13099-13115.

Zaller DM, Eckhardt LA. Deletion of a B-cell-specific enhancer affects transfected, but not endogenous, immunoglobulin heavy-chain gene expression. Proc Natl Acad Sci USA. 1985;82(15):5088-92.

Zan H, Casali P. Regulation of Aicda expression and AID activity. Autoimmunity. 2013;46(2):83-101.

Zeng W, Kong Q, Li C, Mao B. Xenopus RCOR2 (REST corepressor 2) interacts with ZMYND8, which is involved in neural differentiation. Biochem Biophys Res Commun. 2010;394(4):1024-9.

Zhang J, Bottaro A, Li S, Stewart V, Alt FW. A selective defect in IgG2b switching as a result of targeted mutation of the I gamma 2b promoter and exon. EMBO J. 1993;12(9):3529-37.

Zhang Y, Gostissa M, Hildebrand DG, et al. The role of mechanistic factors in promoting chromosomal translocations found in lymphoid and other cancers. Adv Immunol. 2010;106:93-133.

Zheng S, Vuong BQ, Vaidyanathan B, Lin JY, Huang FT, Chaudhuri J. Non-coding RNA Generated following Lariat Debranching Mediates Targeting of AID to DNA. Cell. 2015;161(4):762-73.

Zimmermann M, Lottersberger F, Buonomo SB, Sfeir A, De lange T. 53BP1 regulates DSB repair using Rif1 to control 5' end resection. Science. 2013;339(6120):700-4.

Zimmermann M, de Lange T. 53BP1: pro choice in DNA repair. Trends Cell Biol. 2014;24(2):108-17.

Zmynd8. (n.d.). Retrieved from Human Protein Atlas: https://www.proteinatlas.org/ENSG00000101040-ZMYND8/tissue

Zmynd8. (n.d.). Retrieved from International Mouse Phenotype Consortium: https://www.mousephenotype.org/data/genes/MGI%3A19