THE ROLE OF SWITCH REGION DNA AND PROTEIN FACTORS IN IMMUNOGLOBULIN CLASS SWITCH RECOMBINATION

By

Shahnaz Firdaus Masani

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Genetics – Doctor of Philosophy

2015

ABSTRACT

THE ROLE OF SWITCH REGION DNA AND PROTEIN FACTORS IN IMMUNOGLOBULIN CLASS SWITCH RECOMBINATION

By

Shahnaz Firdaus Masani

Immunoglobulin class switch recombination (CSR) is a recombination event that changes the heavy chain constant region of an antibody while maintaining the same variable region. The process of CSR alters the effector function of the antibody without altering antigen specificity, which ensures efficient pathogen clearance. Defects in CSR can result in hyper-IgM syndrome, autoimmune diseases and chromosomal translocations that result in lymphoid malignancies. Thus, understanding the mechanism of CSR will give us better insight into these diseases. The work presented in this dissertation aims to improve our understanding of the molecular mechanism of CSR.

CSR occurs via a cut and paste mechanism that involves the generation and repair of DNA double stranded breaks (DSBs) within intronic regions known as switch regions. The first part of this thesis aims to identify sequence features within switch regions that are important for CSR. We show that WGCW (W=A/T) motifs, which occur at a high frequency in switch regions, are important determinants of CSR efficiency.

The next study focuses on elucidating the mechanism of DSB formation during CSR. It is known that uracils within switch regions are recognized and excised by Uracil DNA Glycosylase to result in abasic sites. AP endonucleases then create a nick at abasic sites and two closely spaced nicks on both DNA

strands can act as DSBs. However the identity of the AP endonuclease that creates DSBs during CSR was unclear. Our work shows that APE1, the major AP endonuclease in mammalian cells, is essential for CSR.

Finally, we focused on understanding the mechanism of DSB repair by studying the role of DNA ligases in CSR. Mammals have three DNA Ligases, each of which are conventionally thought to have distinct, non-overlapping functions. By creating cells that lack one or two DNA ligases we were able to better understand their role in CSR as well as in other aspects of DNA metabolism. Our results show a remarkable level of functional redundancy between the three DNA ligases. We show that DNA Ligase I, previously thought to be essential for joining Okazaki fragments during DNA replication, is dispensable for cell viability, a number or DNA repair pathways as well as for CSR. In addition we also constructed and characterized a cell line with only one DNA ligase, Ligase III. These cells are viable and show no increased hypersensitivity to a number of DNA damaging agents. Our results show a previously unanticipated level of redundancy between DNA ligases in CSR as well as in DNA replication and repair. These studies have improved our understanding of the molecular mechanism of CSR. In addition the creation of cell lines lacking APE1 and DNA ligases will provide important tools for the study of DNA repair and replication.

ACKNOWLEDGEMENTS

To my mentor, Dr. Kefei Yu, Thank you for all the support and guidance you have given me through the years. Your unrelenting optimism and enthusiasm for research has truly made every day at work fun. As a mentor, you have been kind and patient but have pushed me to work harder and to be a better scientist. I couldn't have asked for a better mentor. You are a wonderful teacher and an inspiration as a scientist.

To members of the Yu Lab, Li, It has been my privilege to work with you. I can truly say that I was trained by the best. Your patience and guidance in the lab has helped me tremendously over the years. You have guided and supported me not only in my research, but in every aspect of my life over the past six years. Ahrom, it has been a pleasure working with you over the past two years. You have been helpful and generous and I'm lucky to have you as a collegue.

To Dr. Kathy Meek, Thank you for your support and encouragement. You have inspired and motivated me to be a better scientist. Thank you for all your input and guidance, I wouldn't be where I am today without it. Your passion for your work and your constant cheer has made every interaction with you an absolute pleasure.

To my committee members, Dr. Yong-Hui Zheng, Dr. Sungjin Kim, Thank you for all your guidance over the years. You have helped me think constructively and analytically. All your input and advice is truly appreciated.

To members of the Meek Lab, who adopted me as one of their own. Jessica, Tina and Sunetra, Thanks for your support and guidance. Jessica, thanks for all your advice and help. Sunetra, thanks for all those talks over coffee and lunch that helped me stay sane.

To my family and friends, without whom none of this would be possible. Mum and Dad, you are my inspiration. I can only strive to live up to your example. Your love and support are the reasons I've made it this far. Thank you for everything. Kainaz and Yuktash, Thank you for all your love and support. Kain, those six-hour Skype sessions are the reason I've been happy here. Thanks to all my friends who believed in and supported me through out this process.

To my fiancé, Fernando, You are my support system. Thanks for all the late night talks and coffee runs. Your perseverance and tenacity is a constant inspiration to me. You make me happier than I thought was possible.

To the Genetics Program Personnel, To Dr. Sears and Dr. Ernst, thank you for all your support and guidance. To Jeannine Lee, thank you for everything. You are the person I run to every time I mess up. I've yet to come across a problem you can't solve.

Finally, I would like to express my gratitude to the faculty and my colleagues and friends at MSU, who over the years have helped me hone my skills as a researcher and have made my time here an exciting and fulfilling experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

Ig	Immunoglobulin
V	Variable
C	Constant
D	Diversity
J	Joining
SHM	Somatic Hypermutation
CSR	Class Switch Recombination
DSB	Double Stranded Break
S	Switch
AID	Activation Induced Cytidine Deaminase
ssDNA	Single Stranded DNA
NHEJ	Non Homologous End Joining
BER	Base Excision Repair
MMR	Mismatch Repair
UNG	Uracil DNA Glycosylase
AP	Apurinic/Apyrimidinic
APE	AP Endonuclease
GLT	Germline Transcription
MSH2	MutS Homolog 2
MTT	Thiazolyl blue tetrazolium bromide
GFP	Green Fluorescent Protein

CIT	anti-CD40, IL4, TGFβ
WT	Wild Type
PCR	Polymerase Chain Reaction
A-EJ	Alternative End Joining
d	Days
h	Hours
µl	Microliter (s)
µg	Microgram (s)
µM	Micromolar (s)
nM	Nanomolar (s)
ml	Milliliter (s)
kb	Kilobase (s)

CHAPTER 1: INTRODUCTION

1.1 Overview of Antibody Diversification and Class Switch Recombination

B cells are critical components of the humoral immune system and function to generate Immunoglobulin (Ig) molecules. Ig molecules occur as cell surface B cell receptors and as secreted molecules known as antibodies that provide protection against pathogens and foreign antigens. They are 'Y' shaped molecules that consist of two heavy (50kDa) and two light chains (25kDa) joined together by disulfide bonds (Figure 1). Ig molecules have an N-terminal variable (V) region and a C-terminal constant (C) region (Figure 1). The variable region is responsible for antigen binding whereas the constant region determines the effector function that results in pathogen clearance.

The enormous amount of diversity required for an Ig molecule to efficiently recognize and combat the wide array of pathogens and foreign antigens is generated by somatically altering the B cell genome to diversify the Ig locus. Ig diversification occurs via three main processes: 1) **V(D)J recombination**, a site directed recombination event that assembles the V region of an Ig molecule, 2) **Somatic Hypermutation (SHM)**, a process by which point mutations are introduced into V regions to enable production of higher affinity antibodies and 3) **Class Switch Recombination (CSR)**, a process in which the constant region of the Ig molecule is changed while maintaining the same V region, which allows the Ig molecule to recognize the same pathogen but alters its function allowing for optimal pathogen clearance.

Mammals have five different types of heavy chain constant regions (μ, δ, γ, ε and α). The type of constant region present in the Ig molecule determines the

Ig isotype or class. For example, an antibody of isotype M is known as IgM and has a μ constant region. Thus, mammals have 5 Ig isotypes depending upon their constant region: IgM, IgD, IgG, IgE, IgA (Figure 2). The different types of constant regions vary in properties such as their half-life, their ability to activate the complement system and in tissue localization, thus determining the Ig effector function.

The mammalian heavy chain germline locus consists of multiple constant region genes spanning approximately a 200kb region (μ, δ, γ3, γ1, γ2b, γ2a, ε and α in mouse; μ, δ, γ3, γ1, α1, γ2, γ4, ε and α2 in human). Naïve B cells primarily express IgM, as Cμ is the first constant region downstream of the variable region. In antigen stimulated B cells, the Cμ constant region is replaced by a different constant region (Cγ, Cε or Cα) by a process known as **Ig Class Switch Recombination (CSR)** (Figure 2) [1, 2]. CSR is a deletional recombination event that occurs via the formation and repair of double stranded breaks (DSBs) within switch (S) regions that occur upstream of each constant region. During CSR, the two S regions are joined together and the intervening DNA is deleted, resulting in the variable region being juxtaposed to a new constant region. Through this mechanism, the same V region is maintained thus ensuring the same antigenic specificity, but the C region is changed, changing the Ig effector function. This is essential to mount an effective immune response and to ensure optimal pathogen clearance.

B cells initiate CSR upon interaction with T cells and other cells such as dendritic cells. These interactions result in upregulation of the B cell specific

enzyme ‘activation induced cytidine deaminase’ (AID), a cytidine deaminase that converts cytosines to uracils in single stranded DNA (ssDNA). Switching to a specific isotype is regulated by cytokine-mediated activation of transcription through that specific S region, which allows AID to convert cytosines within the S region to uracils. AID-generated uracils are then processed by various DNA repair pathways to result in DNA DSBs. Once DSBs are made in the upstream donor (initially Sμ) and downstream acceptor S region, the intervening DNA is deleted and the two S regions are joined by the non-homologous end joining (NHEJ) pathway, replacing Cμ with a downstream C region, thus completing CSR (Figure 3).

DSBs during CSR can occur throughout the S regions as well as in the regions bordering the S regions and are not associated with any specific DNA sequence, making CSR a unique region-specific DNA recombination event [1].

The process of CSR has several requirements including switch regions, transcription through the switch regions, the B cell specific enzyme activation induced cytidine deaminase, and several ubiquitously expressed DNA repair factors, each of which will be discussed below in further detail.

1.2 Switch Region Sequences in Class Switch Recombination

Switch (S) regions are 2-10 kb intronic sequences that occur upstream of constant region exons. Mammalian switch regions consist of several tandem repeats which get increasingly degenerate towards the boundaries of the switch regions such that they can no longer be distinguished from the surrounding sequence [2-4]. The non-template strand of mammalian switch regions is highly G rich (40-50%) [1, 2, 5]. Switch regions in mammals are enriched for certain conserved G-rich pentamer motifs such as GGG(G/C)T, GAGCT, TGGGG [3]. In mice and humans, switch regions can be broadly classified into two families: 1) The Sμ, Sα and Sε family, where repeats comprise almost entirely of pentamers GGG(G/C)T, GAGCT. 2) The gamma family (Sγ3, Sγ1, Sγ2b, Sγ2a), which also have these pentamer motifs but are characterized by a 49-52mer repeat sequence [1, 3].

Unlike mammalian switch regions, amphibian and reptile switch regions are A/T rich as opposed to G/C rich and are repeat-poor [6, 7]. Despite these differences, the *Xenopus* Sμ region can functionally substitute for Sγ1 in mouse, albeit with lower efficiency [8]. Furthermore, DNA breaks that occur during CSR are distributed randomly throughout the switch regions, indicating a lack of any specific recombination signal sequence like that of V(D)J recombination [2, 9]. These results, along with the heterogeneity in switch region sequences and random distribution of switch junctions have made it difficult to identify features within switch regions that are important for CSR [1, 2, 10]. However, deletion of Sγ1 in mouse abolishes CSR to IgG1 and similar deletion of Sμ severely impairs

CSR to all isotypes [11, 12]. In addition, replacing the deleted Sγ1 with an irrelevant intronic sequence cannot rescue switching, indicating that switch regions do contain unique elements that are essential for CSR [8]. Our lab is interested in studying the switch region sequences to identify these key sequence features. The work in Chapter 2 involves studying whether switch regions from different families can substitute for each other during CSR, and focuses on understanding the importance of a short sequence motif that is common to all S regions.

1.3 Activation induced cytidine deaminase

Activation induced cytidine deaminase (AID) is a B cell-specific factor that is essential for three important antibody diversification processes: 1) CSR, 2) SHM and 3) gene conversion, a process that occurs in chickens and ducks, in which short sequences within the active variable region are replaced by sequences from upstream pseudogenes to increase V region diversity [13-15].

AID belongs to the APOBEC family of cytidine deaminases, other members of which are involved in antiviral immunity and lipid metabolism [1, 16]. As a result of its homology with APOBEC1, an RNA editing enzyme that is involved in lipid metabolism, AID was also hypothesized to be an RNA editing enzyme. However, no RNA target of AID has been identified to date [1, 17]. In fact, in *E coli* that express AID, C to T transitions were observed in the DNA and importantly, the frequency of transitions was enhanced in *E. coli* deficient in uracil DNA glycosylase [1, 18]. In addition, AID acts on single stranded DNA (ssDNA)

in vitro and chromatin-immunoprecipitation experiments demonstrated that AID is associated with S regions in cells undergoing CSR [1, 17]. Thus, it is generally accepted that AID is an ssDNA deaminase that converts cytosines to uracils. During CSR, uracils are then processed by various DNA repair pathways resulting in the generation of double strand breaks within S regions [1, 2].

AID activity is not sequence-specific but cytosines within WRC motifs (W=A/T, R=A/G) are preferred [6, 19, 20]. WRC motifs are enriched within S regions, mainly in the form of the short sequence motif 'AGCT'. AGCT occurs within S regions as a part of the conserved pentamer motif GAGCT [3, 21]. AGCT is a palindromic sequence that contains a WRC motif on both DNA strands, which is known as an overlapping AID hotspot. It is one of four possible WGCW motifs (W=A/T) that result in overlapping AID hotspots across the two DNA strands. It has been hypothesized that AID deamination of cytosines on both DNA strands within these WGCW motifs would result in closely spaced nicks on both strands, which serve as double stranded DNA breaks within S regions. Studies from our lab have shown that disruption of the WGCW motif impairs CSR, indicating that these motifs are important determinants of S region 'quality' and help increase efficiency of CSR [21]. The study in Chapter 2 focuses on understanding the importance of WGCW motifs within switch regions.

AID can mutate genomic regions other that the Ig locus. Thus, AID expression is strictly regulated at multiple levels to avoid translocations and to maintain genomic integrity [22]. Resting mature B cells have virtually undetectable levels of AID. AID expression is rapidly induced in activated B cells

[23]. This upregulation can occur in a T cell dependent or independent manner and takes place via activation of the canonical and non-canonical NF-κB pathways as well as other B cell lineage specific transcription factors [24]. Post-transcriptional regulation may also take place and microRNAs that target AID have been identified [22]. The sub-cellular localization of AID is strictly controlled via an N-terminal nuclear localization signal and a C-terminal nuclear export signal [22, 24]. In addition, AID activity can be post-translationally modified by phosphorylation [22, 24]. Thus AID expression and activity are strictly controlled to allow for efficient CSR while limiting any damage from off-target deamination.

Although significant progress has been made in understanding the regulation and function of AID, many questions still remain unanswered. Precisely how AID is targeted to S regions is still unclear. In addition it is known that the C-terminal 8-17 amino acids of AID are essential for CSR, but not for SHM [22, 23]. The function of this C terminal region remains unidentified. It has been hypothesized that post-translational modifications of AID could function to influence DNA repair at switch regions and that AID could act as an epigenetic reprogramming factor in other cell types, but studies have yielded contradictory results. Thus, many unresolved questions still remain with regards to our understanding of AID activity.

1.4 The Role of Transcription in Class Switch Recombination

Each constant region gene consists of a discrete transcriptional unit that includes a cytokine dependent promoter, an I exon, an intronic S region and multiple C region exons (Figure 4) [1, 25]. The promoter for each constant region responds to a different combination of cytokines and targeting of CSR to a particular C region is controlled by activating its specific promoter. For example, interferon gamma stimulation in mouse B cells induces CSR to IgG2a, whereas interleukin-4 stimulation induces CSR to IgG1 [1, 25].

Cytokine induction results in generation of the primary transcript, which consists of an I exon, followed by an intronic switch region and constant region exons. This primary transcript is processed by the splicing machinery to make a 'sterile germline transcript' (Figure 4). The processed transcript is known as a sterile transcript as it does not code for any protein [25, 26].

Transcription through a switch region is essential for CSR to that isotype. Deletion of the germline promoter abolishes CSR to that isotype and replacement with another heterologous promoter restores CSR [27, 28]. If the cytokine-dependent germline promoter is replaced by a constitutively active promoter, this alleviates the cytokine regulation of CSR to the corresponding isotype [28, 29]. In addition, deletion of the 3' enhancer, which inhibits S region transcription, severely impairs CSR, indicating that germline transcription is a prerequisite for CSR [30-33].

It is important to note that the level of CSR does not correlate with the amount of germline transcription [21]. A higher amount of germline transcripts

does not confer higher CSR efficiency, however if the level of transcription is below a certain threshold amount, this impairs CSR. The minimum threshold level required for optimal CSR and the quantitative relationship between CSR and germline transcription remain unclear.

Precisely why transcription through switch regions is essential for CSR remains unclear and multiple non-exclusive hypotheses have been put forth to explain this requirement. The R-loop model is one explanation for the transcriptional requirement of CSR. As mentioned above, the template strand of switch regions is C rich. Thus, the transcribed RNA is G-rich and can form an RNA-DNA hybrid structure with the C rich DNA known as an R-loop [5, 34-36]. R-loop formation causes the non-template strand to be single stranded and hence available for the action of AID. Although the R-loop model explains certain features of CSR, direct evidence for its requirement *in vivo* is inadequate.

Transcription may also serve other non-exclusive functions in CSR. It has been suggested that transcription may serve to recruit AID to the switch regions via interactions with the transcriptional or splicing machinery [1, 29, 37]. Studies have shown that AID interacts with stalled RNA Polymerase II at S regions. It has been hypthesized that interaction of AID with several factors involved in transcription, splicing, and chromatin remodeling such as SPT5, components of the RNA exosome and PTIP are responsible for recruiting AID to S regions [37-39]. Histone modifications during transcription may also play a role in AID recruitment to S regions through interaction of AID with the modified histones [40]. However the precise role of transcription in targeting AID to S regions is still

not known.

1.5 DNA Double Strand Breaks in Class Switch Recombination

DNA DSBs are an essential intermediate during CSR [1, 2]. Phosphorylated histone H2AX (Y-H2AX), a marker for DNA DSBs is observed at Ig loci in an AID dependent manner in B cells that are stimulated to undergo CSR. In addition, DSB detection via Ligation Mediated-PCR as well as the generation of an episomal circle containing the intervening DNA during CSR provide strong evidence for the involvement of DSBs in CSR [1, 23].

AID is a B-cell specific factor that introduces uracils within S regions. The proteins and enzymes that process AID-induced uracils and convert them to DSBs are ubiquitously expressed DNA repair factors [1]. Two main DNA repair pathways contribute to DSB formation: Base Excision Repair (BER) and Mismatch Repair (MMR) (Figure 5) [1, 23].

Uracils within switch regions can be recognized and excised by the BER enzyme Uracil DNA glycosylase (UNG), resulting in the formation of abasic sites, also known as apurinic/apyramidinic (AP) sites [1, 23]. Mammalian cells express four uracil DNA glycosylases: UNG2, SMUG1, MBD4 and TDG [23]. UNG2 is the major uracil DNA glycosylase involved in CSR. B cells from UNG2 deficient mice show substantially impaired CSR (less than 10% compared to wildtype) [41, 42]. SMUG1 does not contribute to CSR in wildtype cells as evidenced by a lack of any CSR defect in SMUG1 deficient cells [43]. However, SMUG1 can partially substitute for UNG2 in cells lacking UNG2 and cells that lack both UNG2 and

SMUG1 show a further 5 fold reduction in CSR efficiency as compared to UNG deficient cells [44]. AP sites within S regions are recognized by AP endonucleases, which introduce nicks at abasic sites. Two closely spaced nicks on opposite strands can act as DSBs [1, 23]. Mammals express two AP endonucleases: APE1 and APE2 [23, 45]. APE1 is the major AP endonuclease in cells and the AP endonuclease activity of APE2 is about a 1000 fold less than APE1 [23]. APE1 deficiency is essential for embryonic development in mice and is also required for survival of several human cell lines, making it difficult to identify its role in CSR [46-48]. In addition, previous studies on identifying the role of APE1 and APE2 in CSR have resulted in conflicting results [49-52]. The work in Chapter 3 focuses on identifying the AP endonuclease involved in CSR.

In addition to the BER pathway, MMR factors MSH2-MSH6 can also recognize and process U:G mismatches [23, 43, 53]. The MSH2-MSH6 dimer can recruit MLH1-PMS2 as well as Exonuclease1 (EXO1). PMS2 has an endonuclease activity, which provides an entry for EXO1. Alternatively, EXO1 can initiate resection from a pre-existing single stranded DNA break that is 5' to the mismatch, resulting in ssDNA gaps with long overhangs. Closely spaced ssDNA gaps on both DNA strands, could act as DSBs.

The BER pathway is the major pathway involved in DSB formation during CSR, as indicated by the observation that UNG2-deficient cells show only 10% CSR efficiency as compared to wildtype [41, 42]. MMR deficient cells show a milder CSR defect, ~50% of wildtype levels [43, 53]. The MMR pathway is thought to act as a backup pathway for recognition and processing of uracils. It

can also function to process single stranded breaks generated by the BER pathway that are too far apart from each other, converting them to DSBs [23].

S region DSBs consist of a mixture of different types of DNA ends, including various sized nucleotide overhangs, blunt ends as well as abasic sites. These ends can be processed by various error prone polymerases and nucleases before they are synapsed and joined [1, 23].

An important question that remains to be answered with regards to S region DSB formation is why AID-generated uracils result in DSBs in S regions during CSR. Normally, uracils in the genome are excised and repaired in an error-free manner by the BER or MMR pathways, maintaining genomic integrity. How these pathways are co-opted during CSR to result in DNA DSBs or during SHM to result in V region mutations is unknown. Another unresolved question is why AID-induced uracils result in different outcomes during CSR (S region DSBs) and SHM (V region mutations).

Once DSBs are made in the upstream donor and downstream acceptor S regions, they are synapsed and joined to complete the recombination process. The presence of S region DSBs activates a number of DNA damage response factors including ATM, 53BP1 and γ-H2AX [1, 23]. Cells lacking 53BP1 or γ-H2AX show a severe CSR defect and these factors are thought to be involved in recruiting and coordinating DNA repair factors as well as in synapsis of S regions [54, 55].

The major pathway involved in end joining during CSR is the classical non-homologous end-joining pathway (c-NHEJ) [1]. This is the major pathway

responsible for the repair of DSBs in mammals. Cells lacking any of the core c-NHEJ factors such as Ku70, Ku80, DNA Ligase IV or XRCC4 are severely impaired for CSR (20-50% of wildtype levels) [1, 56, 57]. The residual CSR activity in cells with impaired c-NHEJ is attributed to a poorly defined non-homologous end-joining pathway known as alternative end-joining (A-EJ) [56-58]. The A-EJ pathway has been associated with an increase in translocations, making it important to identify the factors involved in this pathway [59]. Chapter 4 describes a study aimed at studying the role of the three mammalian DNA ligases in CSR in order to identify the DNA ligase responsible for A-EJ.

1.6 CH12F3 – A Cell-line to Study Class Switch Recombination

In our laboratory, we study CSR using 'CH12F3', a mouse B cell line capable of cytokine-dependent CSR. Upon stimulation with CIT medium (RPMI medium with CD40L, IL4, TGFβ1), CH12F3 cells switch from IgM to IgA with a high efficiency (~50%) within 72 hours [60]. CH12F3 cells have a near diploid genome, and our lab is capable of efficient gene targeting in these cells, allowing us to efficiently identify and study cis and trans factors involved in CSR. The robust level of CSR in CH12F3 cells allows us to quantify relatively small changes in CSR efficiency, making this an elegant and sensitive system to study CSR [21, 60].

1.7 Clinical Relevance

CSR is a unique region-specific recombination event that ensures optimal pathogen clearance by the humoral immune system. Defects in CSR have severe consequences such as hyper-IgM syndrome, autoimmune diseases and chromosomal translocations that result in lymphoid malignancies such as Sporadic Burkitts lymphoma and diffuse large B cell lymphoma [15, 61-64]. It is important to understand the mechanism of CSR to give us a better insight into these diseases.

CSR involves the programmed introduction and repair of DSBs within DNA. During the process of CSR, factors belonging to various DNA repair pathways are co-opted to create and repair S region DSBs. When these processes are dysregulated, it can lead to translocations and lymphoid maliganacies. Allelic variants of the protein factors studied in the following chapters including APE1 and DNA ligases are associated with various forms of cancer [65-68]. The generation of cell lines lacking these factors will facilitate the understanding of their role in DNA repair and further our understanding of the underlying mechanism of these diseases.

CHAPTER 2: CLASS SWITCH RECOMBINATION EFFICIENCY CORRELATES WITH SWITCH REGION WGCW DENSITY

The work in this chapter is published in the following research article:

Han, L., Masani, S., and Yu, K.; (2011); Overlapping activation-induced cytidine deaminase hotspot motifs in Ig class-switch recombination*; PNAS*; 108; 11584-9.

2.1 Introduction

CSR occurs via a cut and paste mechanism that involves the formation and repair of DSBs within intronic S regions [1, 2]. S regions occur upstream of each constant region and contain a number of heterogeneous tandem repeats. There are two families of switch regions in mice and humans: 1) The Sμ, Sα and Sε family, consisting of repeats comprising almost entirely of two pentamer motifs- GGG(G/C)T, GAGCT. 2) The gamma family (Sγ3, Sγ1, Sγ2b, Sγ2a) which have the same pentamer motifs but are characterized by a distinct 49-52mer repeat sequence [2, 3].

Studies with transiently transfected plasmid-based switch substrates indicated that recombination is isotype specific i.e. recombination to each S region involves a specific and distinct set of factors [69, 70]. In contrast, studies with stably integrated plasmid substrates suggested that primary switch region sequence does not play a role in CSR i.e. recombination to all switch regions involves the same protein factors [71]. In addition, it was shown that in mouse, the Sγ3 region could substitute for Sγ1, indicating that these regions are interchangeable [72]. These studies indicated that CSR to each isotype does not involve specific factors or a distinct recombination mechanism, and that CSR to all isotypes involves the same recombination machinery. In order to confirm this, we tested whether S regions from different families can substitute for each other during CSR.

S regions are essential for CSR and deletion of an S region abolishes CSR to that isotype [8, 12]. Furthermore, an irrelevant intronic sequence cannot

functionally substitute for an S region in CSR, indicating that S regions do contain some sequence motifs that are important for CSR [8]. Unlike V(D)J recombination, CSR is not directed by specific recombination signal sequences and DSBs can occur at random throughout the S regions [1-3]. The lack of a specific signal sequence, heterogeneity in S region sequences and the lack of association of CSR breaks with any specific DNA sequence has made it difficult to identify and study sequence motifs common to all switch regions that are important for CSR.

CSR is initiated by the enzyme AID, which converts cytosines to uracils within S regions. These uracils are processed to eventually result in DSBs. AID preferentially acts on cytosines within the WRC (W=A/T, R=A/G) sequence motifs. S regions are enriched in the pentamer motif GAGCT, which contains a WRC motif in the form of AGC. In fact, the AGCT sequence from the pentamer, which is a palindrome that results in AGCT on the antiparallel strand as well, providing closely spaced AID hotspots on both DNA strands. AGCT is one of four possible palindromic WGCW (W=A/T) motifs that provide overlapping AID hotspots. AID action on cytosines on both strands within the WGCW motifs would result in closely spaced nicks on both strands, that function as DSBs. Thus, it is thought that WGCW motifs might facilitate DSB formation during CSR [20, 73]. As a part of this study, we looked at the role of WGCW motifs in CSR.

2.2 Materials and Methods

2.2.1 Cell culture and Class Switch Recombination assay

CH12F3 cells were cultured in RPMI1640 medium supplemented with 10% fetal bovine serum and 50 μM of beta-mercaptoethanol. For CSR assay, viable CH12F3 cells were seeded at 5 x 10^4 cells/ml in the presence of 1μg/ml anti-CD40 antibody (eBioscience 16-0402-86), 5 ng/ml of IL-4 (R&D Systems) and 0.5 ng/ml TGF-β1 (R&D Systems), and grown for 72 hours. Cells were stained with a FITC-conjugated anti-mouse IgA antibody (BD Biosciences 559354) and analyzed by flow cytometry (BD Biosciences). CSR efficiency is defined as the percentage of IgA cells.

2.2.2 Cloning of switch region sequences

Approximately 2kb restriction fragments containing core repeats of each Sγ region were selected to replace Sα. The following restriction fragments were cloned within the two heterologous loxP sites of the exchange vector:

Table 1 List of S regions introduced into the Sα locus

Switch Region	**Size (Kb)**	**Restriction Fragment / Construction**
Sγ1	2172	BamHI-BamHI fragment of mouse Sγ1 (D78344: 5824–8000)

(Table 1 cont'd.)

Sγ3	1181	HindIII-SacI fragment of mouse Sγ3 [74].
Sγ2b	2051	SpeI-SpeI fragment of mouse Sγ2b (D78344: 29538–31593).
Sα-2.1HX	2076	HindIII-XbaI fragment of mouse Sα (AC160982: 124768–122693).
Sα-1.1Con	1120	Sα made by ligation of 14 repeats of consensus Sα 80mer [21].

2.2.3 Recombinase mediated cassette exchange

5 µg of exchange vector and 1 µg of Cre-expression vector were co-transfected into 1F7 cells by electroporation. Transfected cells were resuspended in 1 ml of pre-warmed medium. From this, 10 µl cells were resuspended in 10 ml of pre-warmed medium and seeded to a 96-well plate at 100 µl/well. The 96-well plate was returned to incubator for 72 hours before GANC (Sigma-Aldrich, St. Louis, MO) was added to a final concentration of 2µg/ml. About 30~50 GANC colonies were obtained after 6~8 days. Typically, 24 colonies were screened by PCR to confirm recombinase mediated cassette exchange (RMCE). This was

followed by southern blot confirmation of four to six colonies, which are used for CSR assays.

2.2.4 Quantitation of germline transcripts

Cells were cultured to a density around $1x10^6$ cells/ml before cytokines (CIT: anti-CD40, IL4 and TGF-β) were added. Four hours after the addition of cytokines, total cellular RNA was harvested using the Trizol reagent (Invitrogen). 2 µg of RNA was reverse-transcribed into cDNA using the high capacity RT kit (ABI) according to manufacturer's instructions. One-tenth of the RT reaction was used in a Taqman-based real-time PCR reaction (Taqman gene expression master mix from ABI). β -actin was used as an internal control. Each sample, including the β-actin control, was done in triplicates. Cycling and data collection were done on an ABI StepOnePlus apparatus.

Oligonucleotides used in real-time PCR were:

IαCα-Taqman (5'FAM-CTGCGAGAAATCCCACCATCTACCCA-3'BHQ)

IαCα-Fwd (5' CCTATGAAGGACACTCAACAACATTG 3')

IαCα-Rev (5' ACAGAGCTCGTGGGAGTGTCA 3')

Actβ-Taqman (5'FAM-ATCGTGGGCCGCCCTAGGCAC 3'BHQ)

Actβ-Fwd (5' ATGCTCCCCGGGCTGTA 3')

Actβ-Rev (5' ATAGGAGTCCTTCTGACCCATTCC 3')

2.3 Results

2.3.1 Sγ regions can substitute for Sα during Class Switch Recombination

In order to study whether Sγ regions can support CSR in place of Sα, restriction fragments of endogenous murine Sγ1, Sγ3 and Sγ2b were introduced into the endogenous Sα locus. This study was performed using modified CH12F3 cells known as '1F7'. In the 1F7 cell line, the entire Sα region is replaced by an RMCE cassette consisting of a positive-negative selection cassette. This cassette contains puromycin resistance and thymidine kinase genes flanked by one wildtype and one mutant loxP site (Figure 6 A). Desired S regions can be cloned into the multiple cloning site of an enchange vector, which is flanked by the same pair of heterologous loxP sites. The exchange vector is co-transfected into 1F7 cells along with a CRE recombinase expressing plasmid. The use of heterologous loxP sites ensures cassette exchange, placing the new S region at the endogenous chromosomal Sα locus (Figure 6 B). Successful exchange events can be selected for by gancyclovir selection, as cells have lost the thymidine kinase gene and then confirmed by PCR and Southern Blot. The three Sγ region sequences and two control Sα sequences were cloned within the heterologous loxP sites of an exchange vector and introduced into the Sα locus by RMCE (Table 1, Figure 6 B). Successful exchange was confirmed by PCR and Southern blot and 3-4 clones were selected for further study (Figure 6 C, D). CSR assays were performed on cells in which the endogenous Sα region was replaced by the test S region sequences. All the Sγ regions can support CSR in

place of Sα, indicating that CSR does not occur in an isotype specific manner (Figure 7).

2.3.2 Class Switch Recombination efficiency correlates with switch region WGCW density

Although each Sγ region tested was able to support CSR at the Sα locus, the CSR efficiency conferred by each S region varied dramatically. Interestingly, we found that the CSR efficiency conferred by a S region correlates remarkably with the WGCW motif density (WGCW density = the number of WGCW motifs / Kb ; $R^2 = 0.96$) (Figure 8 A). Although germline transcription is essential for CSR, and levels of GLT varied dramatically for each S region, CSR efficiency does not correlate with levels of GLT (Figure 8 B). CSR efficiency also does not correlate with the total number of WRC (AID hotspots) or WGCW (overlapping AID hotspots) motifs within each S region (Figure 8 C, D). CSR levels do correlate with WRC density ($R^2 = 0.77$), however, the strongest correlation is with WGCW motif density (Figure 8 A, E), indicating that these motifs play an important role in CSR.

2.4 Discussion

CSR occurs via the creation and repair of DSBs within S regions. Whether each S region can direct CSR only to its isotype (e.g. Sγ1 can only support CSR to IgG1) has been a matter of some controversy. By showing that S regions of different families can functionally substitute for each other (Sγ for Sα), we have unequivocally shown that CSR does not occur in an isotype specific manner. Thus, S regions do not contain any specialized sequences that restrict class switching only to their constant region. Rather, it is likely that all S regions share several common sequence features that facilitate CSR.

CSR efficiency conferred by each S region shows a remarkable correlation with WGCW density, where WGCW density is defined as the number of WGCW motifs / kb of S region. This indicates that the WGCW motif is an important determinant of S region quality. The WGCW motif is likely important for CSR because it provides overlapping AID hotpots on both DNA strands. Cytosines on both DNA strands within these hotspots can be acted upon by AID and then processed by UNG2-APE1 mediated BER to result in closely spaced single stranded nicks on both DNA strands, which act as DSBs (Figure 9). Other studies from our laboratory and our collaborators involve more comprehensive studies on understanding the influence of WGCW motifs on CSR and have confirmed that these motifs are important determinants of S region quality [21, 75]. Studies have also shown the importance of additional short sequence motifs such as G-clusters for CSR [35, 75]. Thus, even though S regions lack a specific recombination signal sequence, they are enriched in certain short sequence

motifs that are important in determining S region quality and hence CSR efficiency.

CHAPTER 3: APURINIC/APYRIMIDINIC ENDONUCLEASE 1 IS ESSENTIAL FOR IMMUNOGLOBULIN CLASS SWITCH RECOMBINATION

The work in this chapter is published in the following research article:

Masani, S., Han, L., and Yu, K.; (2013); Apurinic/Apyrimidinic Endonuclease 1 Is the Essential Nuclease during Immunoglobulin Class Switch Recombination*; Mol Cell Biol*; 33; 1468-73.

3.1 Introduction

In mammals, secondary antibody diversification takes place via two main processes, somatic hypermutation and class switch recombination. SHM involves the introduction of point mutations at Ig V regions, to create high affinity antibodies, whereas CSR results in changing the Ig heavy chain C region, thereby modifying the Ig effector function. CSR involves the creation and repair of DSBs which are joined by a cut and paste mechanism, deleting the intervening DNA and juxtaposing the V region to a new C region [1, 2].

Activation induced cytidine deaminase (AID) is a B cell specific factor that is essential for both SHM and CSR [13, 15]. It is a single-stranded DNA deaminase that deaminates cytosines within V and S regions to uracils [1, 2]. Normally, uracils in DNA are detected and repaired by the base excision repair or mismatch repair pathway. However, members of the same repair pathways are known to be involved in error-prone processing of AID-generated uracils to result in point mutations or DNA DSBs at the Ig locus. How BER and MMR factors are co-opted during CSR and SHM to result in error-prone processing of these uracils is unclear. BER and MMR factors may function differently during processing of AID-generated uracils. Alternatively, only certain components of these pathways may be involved in processing of AID-generated uracils, resulting in error-prone repair. Both the BER and MMR pathways are involved in uracil processing during CSR and loss of either Uracil DNA Glycosylase 2 (UNG2), a BER factor, or MutS homolog 2 (MSH2), an MMR factor, results in impaired CSR [41, 53].

The BER pathway is generally considered to be the major pathway involved

in uracil processing during CSR and an UNG2 deficiency results in a severe CSR defect [41]. UNG2 recognizes and excises uracils in S regions resulting in abasic or apurinic/apyrimidinic (AP) sites. During BER, AP sites are recognized by AP endonucleases, which cleave the phosphodiester bond, introducing a nick in the DNA. Mammals have one major AP endonuclease, APE1 which cleaves the phosphodiester bond 5' to abasic site resulting in a 3'-OH and a 5'-deoxy-ribophosphate on either end of the single strand break. It is thought that during CSR, two single stranded DNA breaks on opposite DNA strands can spontaneously form a DSB [1, 2]. However, the role of APE1 in CSR has proven difficult to elucidate, as APE1 is essential for mouse embryonic development as well as for the viability of several human cell lines [46-48]. In addition to APE1, recently, a minor AP endonuclease APE2 has been identified [76]. The AP endonuclease activity of APE2 is very weak, about 1000 fold less than APE1 and its role in BER and CSR is unclear [23].

Previous studies yielded contradictory results with regards to the role of APE1 and APE2 in CSR. A prior study concluded that APE1 haploinsufficient mice show a modest CSR defect and APE2 deficient mice also show impaired CSR, indicating a role for both APE1 and APE2 in CSR [49, 50]. On the other hand, studies from another group indicated that CH12F3 cells with a transient depletion of APE1 as well as mouse B cells lacking APE2 show no CSR defect, and concluded that neither APE1 nor APE2 are essential for CSR [52]. In our study we have created cells lacking APE1, APE2 as well as cells lacking both these factors in an attempt to understand their role in CSR.

3.2 Materials and Methods

3.2.1 Cell culture and Class Switch Recombination assay

CH12F3 cells were cultured in RPMI1640 medium supplemented with 10% fetal bovine serum and 50 μM of beta-mercaptoethanol. For CSR assay, viable CH12F3 cells were seeded at 5 x 10^4 cells/ml in the presence of 1μg/ml anti-CD40 antibody (eBioscience 16-0402-86), 5 ng/ml of IL-4 (R&D Systems) and 0.5 ng/ml TGF-β1 (R&D Systems), and grown for 72 hours. Cells were stained with a FITC-conjugated anti-mouse IgA antibody (BD Biosciences 559354) and analyzed by flow cytometry (BD Biosciences). CSR efficiency is defined as the percentage of IgA cells.

3.2.2 APE1 gene targeting

A 10.8 kb Bgl II fragment containing the *APE1* gene was isolated from BAC RP23-243M12 (Genbank accession AC136376) and cloned into the targeting vector backbone. The 1.7 kb Bcl I – BsrG I fragment containing exon 4 of the APE1 gene was replaced with a floxed Puromycin selection cassette to obtain the final targeting vector pLH5. 10 million CH12F3 cells were transfected with 10 μg of linear targeting vector and seeded into five 96-well plates. Puromycin was added 48 h later to a final concentration of 1 μg/ml. Puromycin-resistant clones were picked after 7–10 d and screened by PCR and Southern blot analysis.

3.2.3 Genetic complementation

APE1 CDS was amplified from CH12F3 cDNA and cloned into the retroviral transfer vector pMSCV-puro (Clontech Laboratories, Inc.). Recombinant retroviruses were prepared in 293T cells and used to infect APE1Δ/Δ/Δ cells. Infected cells were selected by puromycin and subjected to western blot and CSR assays.

3.2.4 Methyl methanesulfonate (MMS) sensitivity assay

Cells were seeded at 1 x 10^5 cells/ml per well in a 24-well plate and MMS was added to desired concentrations. After 48 h of growth, cell viability was determined by a colorimetric assay that measures the metabolic conversion of thiazolyl blue tetrazolium bromide (MTT) in the mitochondria of living cells. Metabolic conversion of MTT was performed by incubation of harvested cells in 1 mg/ml MTT (in 1x phosphate-based saline) for 2 h. The cell pellet was lysed in 400 µl of acidic isopropanol (20 µl of concentrated hydrochloric acid in 7 ml isopropanol), and the absorbance was measured at 570 nm.

3.2.5 Switch junction analysis

Individual IgA-positive clones were isolated by limiting dilutions of cytokine-stimulated cultures in 96-well plates. Switch junctions were amplified with primers KY761 (5'-AACTCTCCAGCCACAGTAATGACC-3') and KY743 (5'-GAGCTCGTGGGAGTGTCAGTG-3'). PCR products were sequenced at the Genomic Core Facility at Michigan State University.

3.2.6 Reagents

APE1 antibody (10203-1-AP) was purchased from Proteintech Group, Inc. (Chicago, IL). APE2 antibody was kindly provided by Dr. Janet Stavnezer (University of Massachusetts Medical School, Worcester, MA). Mouse β-actin antibody (sc-47778) was purchased from Santa Cruz Biotech (Santa Cruz, CA).

3.3 Results

3.3.1 Gene targeting of APE1 gene in CH12F3 cells

A gene-targeting vector was constructed to disrupt the *APE1* gene in the CH12F3 cell line (Fig. 10 A). The gene targeting would delete exon 4 of the *APE1* gene, which contains ~52% of the coding sequence (Fig. 10 A). The targeted allele contains a puromycin selection cassette and is designated "P". Upon transient transfection of a CRE-expressing plasmid, the floxed puromycin marker can be excised to yield a delta ("Δ") allele. (Fig. 10 A). Upon deletion of the puromycin cassette, another round of gene targeting can be performed using the same targeting vector (Fig. 10 A). During the course of the experiment, we observed that CH12F3 cells have three copies of the *APE1* gene (Fig. 10 B). Therefore, three rounds of gene targeting were performed to disrupt all three copies of the *APE1* gene. The sequential order of genotypes obtained during the gene targeting experiment is as follows: +/+/+, +/+/P, +/+/Δ, +/P/Δ, +/Δ/Δ, P/Δ/Δ and Δ/Δ/Δ. A Southern blot of the +/+/+, +/P/Δ and Δ/Δ/Δ cells was performed to confirm the genotype (Figure 10 B). Western blot analysis of cell extracts showed that APE1 protein levels decrease with decreasing copy number of the *APE1* gene (Figure 10 C). Cells with two copies of *APE1* (+/+/Δ) show a ~30% reduction and cells with one copy of *APE1* (+/Δ/Δ) show a ~60% reduction in APE1 protein levels as compared to wild-type (WT) cells (+/+/+). No APE1 protein is detected in APE1 Δ/Δ/Δ cells, confirming that they are APE1-null cells (Figure 10 C).

3.3.2 Cell Proliferation and MMS Sensitivity of APE1-null cells

APE1 was previously shown to be essential for cell viability of a number of human cell lines [48]. Cell proliferation is a prerequisite for CSR and cells with proliferative defects have impaired CSR. We measured cell proliferation by seeding an equal number of cells and monitoring the number of viable cells over a period of 72 h. This was done in regular medium as well as in the presence of the cytokines used to induce CSR. APE1-null cells showed no proliferation defect as compared to APE1 proficient cells in either of the culture conditions (Figure 11 A).

Methyl methanesulfonate (MMS) is a DNA damaging agent that alkylates bases within DNA. The BER pathway is essential for the repair of this type of damage. As expected, APE1-null cells are extremely sensitive to MMS induced DNA damage (Figure 11 B).

3.3.3 Class switch Recombination in APE1-null cells

CSR efficiency was measured in cells containing three (+/+/+), two (+/+/Δ), one (+/Δ/Δ) and zero (Δ/Δ/Δ) copies of *APE1*. Although cells containing two copies of the *APE1* gene (+/+/Δ) show a reduction in APE1 protein level, they did not show any CSR defect as compared to wild-type cells that have three copies of *APE1* (Figure 12 A, B). Cells with only one copy of *APE1* (+/Δ/Δ) switch to ~50% of WT levels and APE1-null cells switch at ~20% of WT levels (Figure 12 A, B). Thus, CSR efficiency decreases with a decrease in APE1 protein levels.

We also measured the CSR levels at various time points in order to analyze the kinetics of switching and observed that the CSR deficiency was even more profound at earlier time points (Figure 12 C).

The APE1-null cells were genetically complimented by reintroducing APE1 using a retroviral vector. APE1 expression was sufficient to rescue the CSR defect in APE1 null cells (Figure 13 A, B). The expression of the transgene during complementation did not reach WT levels, but was similar to APE1 +/Δ/Δ cells (Figure 13 A). Correspondingly, the complemented cells had a similar CSR efficiency as to APE1 +/Δ/Δ cells, indicating that the CSR defect in APE1-null cells is indeed due to a lack of APE1 (Figure 13 B).

3.3.4 Switch Junctions in APE1-null cells

To determine if APE1 deficiency affects either the targeting or joining of DSBs during CSR, switch junctions were isolated from post-switched IgA cell clones and sequenced. No significant difference was observed with regards to nucleotide overlap (Fig. 14 A, Mann Whitney test, $p = 0.88$) or breakpoint locations within switch regions between WT and APE1-null cells (Fig. 14 B, C).

3.3.5 Gene Targeting of APE2

In order to study the role of APE2 in CSR, we constructed a targeting vector that disrupted exon 6 of the *APE2* gene in CH12F3 cells (Figure 15 A). The *APE2* gene is located on the X chromosome. CH12F3 cells are male-derived and thus only one round of targeting was needed to disrupt this gene. The

targeted allele contains a puromycin selection cassette and is designated “P”, thus resulting in APE2P/Y cells. We created APE2-null cells as well as cells that lack both APE1 and APE2. This was confirmed by Southern blot to confirm genomic configuration (Figure 15 B). APE2 P/Y cells showed no detectable APE2 protein, indicating that they are APE2-null cells (Figure 15 C).

3.3.6 Class Switch Recombination and MMS sensitivity in APE2-null cells

APE2 has an approximately 1000-fold lower AP endonuclease activity as compared to APE1 [23]. This has made it difficult to assess the contribution of APE2 to BER in the presence of the more dominant activity of APE1. APE2 null cells do not show an increased sensitivity to MMS as compared to WT cells (Figure 16 A). In addition, in cells lacking both APE1 and APE2, there is not increased sensitivity as compared to APE1 null cells, indicating that APE2 has no role in BER.

In order to elucidate the role of APE2 in CSR, we measured CSR efficiency of APE2-null cells and cells lacking both APE1 and APE2. APE2-null cells switch to the same levels as WT cells and cells lacking both APE1 and APE2 switch to the same extent as APE1-null cells (Figure 16 B). This indicates that in the CH12F3 cell-line, APE2 has no function in CSR.

3.4 Discussion

The APE1-null cell line created in this study is the first reported viable APE1 knock-out mammalian cell line. Previous studies had shown that APE1 deficiency is embryonic lethal in mice and is lethal in a number of human cell lines [46-48]. Thus, we hypothesized that an APE1 deficiency was also likely to be lethal in CH12F3 cells. Given the potential for cellular lethality, all the gene-targeting experiments described above were first performed in the presence of a tetracyclin-regulated human *APE1-GFP* fusion transgene. Upon obtaining cells with all three endogenous copies of the *APE1* gene disrupted, it became clear that the transgene expression could be completely turned off without affecting cell viability. We then repeated the gene-targeting experiments in the absence of the transgene and created an APE1-null cell line, proving that APE1 is dispensable for CH12F3 cell viability. The differential requirements for APE1 between the different cell lines could reflect species-specific or cell-type specific differences. However the precise reason that APE1 is dispensable for CH12F3 cell growth as opposed to the cells lines previously studied is still unknown. In addition to the AP endonuclease activity, APE1 also has a potent redox activity and APE1 variants have been associated with a number of different types of cancer [65, 77]. Thus, the creation of an APE1-null cell line will help further our understanding of the role of APE1 in these diverse biological processes.

Although it has been hypothesized that APE1 is required to introduce nicks within S regions during CSR, this has not been definitively proven. In fact, the roles of APE1 in CSR have been a matter of some controversy [49, 50, 52].

Previous studies involving the use of APE1-haploinsufficient mice or siRNA mediated transient APE1 depletion reached contradictory conclusions, with one study concluding that APE1 is essential for CSR and the other study stating that it is dispensable [49, 50, 52]. In this study we generated cell lines that completely lack APE1, APE2 or both. This allowed for a direct test of the role of APE1 and APE2 in CSR.

We show that in CH12F3 cells, APE1 is essential for efficient CSR. APE1-null cells have a severe CSR defect (~20% of WT levels). These observations provide compelling evidence for the DNA deamination model for AID function, which states that AID generated uracils are acted upon by UNG2-APE1 mediated BER to result in DSBs. Work from Chapter 1 as well as other studies from our laboratory have shown the importance of the 'WGCW' motif within S regions [21, 75]. We hypothesize that AID could function to deaminate cytosines on both DNA strands of the WGCW motif, converting them to uracils. The uracils could then be processed by UNG2-APE1 mediated BER to result in closely spaced single stranded nicks on both DNA strands, which spontaneously act as DSBs.

The APE1-null cells do show residual CSR activity, indicating that in the absence of APE1, uracils within S regions can still result in DSBs, albeit with a much lower efficiency. The MSH2-mediated MMR pathway has been implicated in the processing of S region uracils [73]. We therefore hypothesized that the residual DNA incision activity in APE1-null cells could be mediated by the MMR pathway. However, preliminary observations using Cas9-CRISPR mediated knock-out of MSH2 in APE1-null cells indicate that these cells do not show any

further CSR defect as compared to APE1-null cells (data not shown), indicating that MSH2-mediated MMR is not responsible for the residual CSR activity in APE1-null cells. It will be interesting to identify the factor responsible for DNA strand incision during CSR in the absence of APE1.

Understanding the role of APE2 in BER and CSR has proven difficult because of the much stronger AP endonuclease activity of APE1 as compared to that of APE2. By creating an APE2-null cell line, we were able to test the role of APE2 in living cells. We showed that APE2 deletion does not increase sensitivity to MMS induced damage in cells that have APE1. In addition, loss of APE2 has no additive effect on hypersensitivity to MMS induced damage in APE1 deficient cells. These data provide strong evidence that APE2 does not play a significant role in BER. In addition, APE2 is also dispensible for CSR in cells that are APE1 proficient or deficient, indicating that APE2 plays no role in CSR.

Our study has provided compelling evidence for the essential role of APE1 in CSR. In addition, we have created the first mammalian cell line lacking APE1, which will be a useful tool to study the role of APE1 in BER as well as its role as a redox factor.

CHAPTER 4: ELUCIDATING THE ROLE OF DNA LIGASES IN CLASS SWITCH RECOMBINATION

The work in this chapter is published in the following research article:

Han, L.*, Masani, S.*, Hsieh, C., and Yu, K.; (2014); DNA Ligase I Is Not Essential for Mammalian Cell Viability*; Cell Reports*; 7; 316-20 (* These authors contributed equally to this work).

4.1 Introduction

DNA ligases are enzymes that catalyze the formation of phosphodiester bonds at single stranded or double stranded DNA breaks. Vertebrates have three ATP dependent DNA ligases (Ligase I, III and IV). Conventionally, these ligases are thought to have distinct and non-overlapping functions. DNA Ligase I is thought to be responsible for joining Okazaki fragments during DNA replication. DNA Ligase III is thought to be responsible for DNA excision repair. DNA Ligase IV is known to be essential for the non-homologous end joining (NHEJ) pathway [78]. Thus, these DNA ligases play essential roles in DNA replication, recombination and repair and all three DNA ligases are essential for mouse embryonic development [79-81].

A recent study reported that the DT40 cell line, a chicken B cell line, is viable in the absence of DNA Ligase I [82]. However, whether or not Ligase I is essential for mammalian cell viability has been controversial [80, 83, 84]. A study aimed at disrupting the catalytic region of Ligase I by deleting its catalytic lysine residue (K583) concluded that Ligase I deficiency is lethal in mouse ES cells [84]. Studies from another group obtained viable cells upon Ligase I disruption by deletion of exons 23-27. However, in these studies the catalytic region of Ligase I was not disrupted, which raises the concern that these cells may still contain a truncated hypomorphic Ligase I protein [80, 83].

Recently, it has been shown that Ligase III can be knocked-out from cells provided a transgene encoding a mitochondrial-targeted Ligase is present [85, 86]. This indicates that it is the mitochondrial function and not the nuclear

function of Ligase III that is essential for cell survival. Surprisingly, these Ligase III deficient cells were not hypersensitive to a number of DNA damaging agents, indicating that either Ligase III is not involved in DNA excision repair or that another ligase is able to compensate for its absence [85].

The lack of phenotype in the Ligase III deficient cells as well as the viability of Ligase I deficient DT40 cells indicate that the roles of the three DNA ligases are not as distinct as previously thought. The creation of a mammalian cell line lacking one or two functional DNA ligases will facilitate the study of the functions of each ligase in DNA replication, recombination and repair.

CSR involves the creation of DNA DSBs within S regions. These breaks must be joined in order to complete the recombination process. A previous study from our laboratory showed that Ligase IV deficient cells show a ~50% reduction in CSR efficiency as compared to WT cells, indicating that DNA ligase IV functions to join DNA DSBs during CSR [56]. The residual CSR activity in DNA ligase IV-null cells is attributed to a poorly defined end joining pathway known as the Alternative End Joining (A-EJ) pathway. The end joining activity in A-EJ could be mediated by Ligase I, Ligase III or by both and the work in this study aims to identify the ligase responsible for DNA end joining during A-EJ. In order to elucidate the functions each DNA ligase in CSR, DNA replication and repair, we aimed to create cells lacking DNA Ligase I, DNA ligase III and cells that lack two DNA ligases (Ligase IV and I ; Ligase IV and III). This will allow us to test which ligase is responsible for end-joining in the absence of Ligase IV i.e. which ligase is responsible for A-EJ as well as to study the functions of Ligases

in other aspects of DNA metabolism.

4.2 Materials and Methods

4.2.1 Cell culture and Class Switch Recombination assay

CH12F3 cells were cultured in RPMI1640 medium supplemented with 10% fetal bovine serum and 50 μM of beta-mercaptoethanol. For CSR assay, viable CH12F3 cells were seeded at 5 x 10^4 cells/ml in the presence of 1μg/ml anti-CD40 antibody (eBioscience 16-0402-86), 5 ng/ml of IL-4 (R&D Systems) and 0.5 ng/ml TGF-β1 (R&D Systems), and grown for 72 hours. Cells were stained with a FITC-conjugated anti-mouse IgA antibody (BD Biosciences 559354) and analyzed by flow cytometry (BD Biosciences). CSR efficiency is defined as the percentage of IgA cells.

4.2.2 Gene targeting of DNA Ligase I

Two 2 kb DNA fragments were PCR amplified from CH12F3 genomic DNA and cloned into a targeting vector as homology blocks for gene targeting (Fig. 1A). 10 million CH12F3 cells were transfected with 10 μg of linear targeting vector and seeded into five 96-well plates. Puromycin was added 48 h later to a final concentration of 1 μg/ml. Puromycin-resistant clones were picked after 7–10 d and screened by PCR and Southern blot analysis.

4.2.3 Drug sensitivity and MTT assay

Cells were seeded at 1x10^5/ml per well in a 24-well plate and various drugs were added at different concentrations. After 48 h of growth, cell viability

was determined by a colorimetric assay that measures the metabolic conversion of thiazolyl blue tetrazolium bromide (MTT) in the mitochondria of living cells. Metabolic conversion of MTT was performed by incubation of harvested cells in 1 mg/ml MTT (in 1x phosphate-based saline) for 2 h. The cell pellet was lysed in 400 µl of acidic isopropanol (20 µl of concentrated hydrochloric acid in 7 ml isopropanol), and the absorbance was measured at 570 nm.

4.2.4 Adenylation Assay

Twenty millions cells were resuspended in 500 µl of adenylation assay buffer (60mMTrisCl, 10mM MgCl2, 5nM DTT, 5ug/ml BSA) and lysed by three cycles of freezing and thawing. The lysate was centrifuged at 21,000g for 15 min at 4ºC. The supernatant (cell extract) was transferred to a new tube and protein concentration was determined by the Bradford method (Bio-Rad, Hercules, CA). For the adenylation assay, 6 µg of cell extract was incubated with 10µCi of α-^{32}P-ATP (3000Ci/mmol) in a total volume of 10 µl in adenylation assay buffer for 15 min at room temperature. The reaction mixture was resolved on an 8% SDS-PAGE gel and imaged on a Typhoon 9200 phosphoimager (GE Healthcare Life Sciences, Pittsburgh, PA).

4.2.5 Karyotype of metaphase chromosomes

Proliferating cells were harvested for cytogenetic analysis with G-banding using the standard method (Hsieh, 1997).

4.2.6 Reagents

Lig1 antibody (18051-1-AP) was purchased from Proteintech Group, Inc. (Chicago, IL). Lig1 (sc-20222) and β-actin (sc-47778) antibodies were purchased from Santa Cruz Biotech (Santa Cruz, CA). Primers for amplifying *Lig1* coding region sequences, Lig1F 5' ATGAGAAAAAAAGAGCAAGAGAGG 3' and Lig1R 5' TTAATAGTCTTCAACGTCGGAGT 3' were purchased Sigma-Aldrich (St. Louis, MO).

4.3 Results

4.3.1 Construction of a Ligase I knock-out cell line

A gene targeting vector was constructed to disrupt Ligase I in CH12F3 cells (Figure 17 A). This vector was designed to disrupt exons 18 and 19 and cause a frameshift of all downstream exons. Exon 18 contains K583, the active site lysine of DNA Ligase I. Thus, in the resultant targeted allele the DNA sequence that codes for the active site of Ligase I is deleted. The targeted allele contains a puromycin selection cassette and is designated "P". Upon transient transfection of a CRE-expressing plasmid, the floxed puromycin marker can be excised to yield a delta ("Δ") allele (Figure 17 A). Two rounds of gene targeting were performed to disrupt both copies of *Ligase I* in CH12F3 cells. Successful gene targeting was confirmed by Southern Blot (Figure 17 B, C).

Northern blot was performed on polyadenylated RNA from Ligase I P/Δ and Ligase I Δ/Δ cells. In Ligase I P/Δ, two RNA species were observed corresponding to the P and Δ alleles (Figure 18 A). The transcript from the P allele is relatively abundant and from the size we inferred that it results from polyadenylation at the SV40 polyadenylation signal from the integrated puro cassette (Figure 17A, 18A). The transcript derived from the Δ allele is of very low abundance (Fig. 18 B) and its size is consistent with a transcript with exons 18-19 deleted (Fig. 17 A). The extremely low levels of this transcript could be due to nonsense-mediated mRNA decay. In Ligase I Δ/Δ cells only the low abundance transcript with exons 18-19 deleted is detected (Figure 18 B).

RNA from Ligase I deficient cells was also analyzed by RT-PCR using primers at the start and end of the Ligase I CDS. In Ligase I P/Δ cells no transcript was detected, indicating that the transcript from the P allele interferes with the PCR (Figure 18 C). RNA from Ligase I Δ/Δ cells yielded a transcript with a size consistent with a transcript from the Δ allele that is missing exons 18-19 (Figure 18 C). Sanger sequencing of the RT-PCR product confirmed that exons 18 and 19 were absent (Figure 18 D).

Ligase I protein was undetectable by western blot analysis of cell extracts from Ligase I Δ/Δ and Ligase I P/Δ cells using two different antibodies and the Ligase I protein reappears in cells transiently complemented with WT Ligase I (Figure 18 E, F).

Mammalian DNA ligases are ATP dependent ligases that covalently bind AMP. Thus, upon incubation with α-^{32}P-ATP, the enzyme-AMP intermediate is radiolabelled and can be separated on an SDS-PAGE gel and visualized by autoradiography. We assayed for adenylated Ligase 1 from cell extracts. Upon incubation with α-^{32}P-ATP, a ~130 kD band was observed in WT cells, corresponding to the size of Ligase I. Ligase III and IV migrate at ~100 kD and are not detected using these conditions. The band corresponding to adenylated Ligase I was absent in extracts from Ligase I deficient cells, but reappears in Ligase 1 Δ/Δ cells transiently complemented with WT Ligase I (Figure 18 G).

We therefore concluded that the Ligase I Δ/Δ and the Ligase I P/Δ cells were truly Ligase I-null cells. All further analysis was performed with both cell types and no difference was observed for any phenotype or assay.

4.3.2 Cell Proliferation, Cytogenetic Analysis and Drug Sensitivity of Ligase I-null cells

Ligase I-null cells did not show any proliferative defect as compared to WT cells (Figure 19 A). This was surprising as Ligase I is thought to be essential for joining of Okazaki fragments during DNA replication. In order to test whether loss of Ligase I had any effect on overall genomic stability, the karyoptype of Ligase I cells was analyzed and characterized by our collaborator, Dr. Hsieh at the University of Southern California. The Ligase I-null cells had a stable near-diploid karyotype, similar to that of WT cells (Figure 20). The Ligase I-null cells did have a chromosomal rearrangement involving chromosome I that was not present in WT cells, but this rearrangement was stably transmitted to all progeny cells. These data indicate that the Ligase I-null cells do not have any gross genomic instability.

We treated the Ligase I-null cells with a number of DNA damaging agents, in an effort to uncover an essential role of Ligase I in DNA repair. Ligase I-null cells did not show any increased sensitivity to zeocin, cisplatin, hydroxyurea, camptothecin as compared to WT cells and showed a very mild hypersensitivity to MMS (Figure 21 A, B). This indicates that Ligase I may have a minor or redundent role in BER, but it does not play an essential role in repair of DNA damage inflicted by any of the other agents used.

4.3.3 Class Switch Recombination in Ligase I-null cells

Ligase I and Ligase III have been implicated in A-EJ during CSR. Ligase I deficient cells did not show any CSR defect as compared to WT cells, indicating that Ligase I mediated A-EJ does not contribute to CSR in NHEJ proficient cells (Figure 22).

4.3.4 Construction of Ligase IV–Ligase I and Ligase IV–Ligase III deficient cells

The Ligase IV-null cells used in this study have been described previously [56]. Gene targeting was performed to disrupt both copies of *Ligase I* in Ligase IV-null cells using the gene targeting strategy described in section 4.3.1. No Ligase I or Ligase IV was detected by Western blot analysis of cell extracts, indicating that these cells lack both Ligase I and Ligase IV (Figure 23 A).

4.3.5 Cell Proliferation and Drug Sensitivity of Ligase IV-Ligase I-null cells

We measured cell proliferation of cells lacking both Ligase I and Ligase IV by seeding an equal number of cells and monitoring the number of viable cells over a period of 72 h. This was done in regular medium as well as in the presence of the cytokines used to induce CSR. The Ligase I-Ligase IV deficient cells showed no proliferation defect under normal culture conditions (without cytokines) (Figure 23 B). When stimulated with cytokines, they display a similar proliferative defect as Ligase IV deficient cells (Figure 23 B). This indicates that under normal culture conditions, Ligase III alone is sufficient to support cell

proliferation and viability. The proliferative defect upon cytokine stimulation can be attributed solely to loss of Ligase IV, as loss of Ligase III does not have an additive effect on the proliferation phenptype as compared to Ligase IV deficient cells.

Ligase I-Ligase IV-null cells were treated with a number of DNA damaging agents (Figure 24). These cells showed a similar level of sensitivity to zeocin as compared to Ligase IV deficient cells, indicating that this repair defect can be solely attributed to Ligase IV and that Ligase I does not contribute to zeocin induced DSB repair in NHEJ deficient cells. These cells are moderately hypersensitive to MMS, indicating that both Ligase I and Ligase IV both contribute to repair of MMS induced DNA damage. The Ligase I-Ligase IV null cells did not show any increased sensitivity to cisplatin, hydroxyurea, and camptothecin as compared to WT, indicating that Ligase III alone is sufficient to repair DNA damage inflicted by these agents.

4.3.6 Class Switch Recombination and genetic complementation of Ligase I-Ligase IV null cells

CSR in Ligase IV deficient cells is mediated by A-EJ. In order to identify whether Ligase I was responsible for A-EJ activity during CSR, CSR efficiency was measured in Ligase I-Ligase IV-null cells. Ligase I-Ligase IV-null cells do not show any additional CSR impairment as compared to Ligase IV-null cells (Figure 25 A). In addition, complementation with WT Ligase IV was sufficient to rescue CSR to WT levels, indicating that Ligase I is not essential for A-EJ during CSR

(Figure 25 B). This indicates that either Ligase III is responsible for A-EJ or Ligase I and Ligase IV serve redundant functions during A-EJ mediated CSR.

4.4 Discussion

DNA Ligase I is conserved throughout evolution and was thought to be essential for cell viability because of its role in joining Okazaki fragments during DNA replication. Previous studies indicated that Ligase I deficient DT40 cells are viable [82]. However, studies with mammalian cell lines have yielded contradictory results with regards to the requirement of Ligase I for cell viability [80, 83, 84]. A previous study using mouse ES cells indicated that cells in which the catalytic lysine of Ligase I is disrupted are not viable, indicating that Ligase I is essential for mammalian cell viability [84]. In contrast, another study that deleted exons 23 – 27 of Ligase I indicated that it is dispensable for mouse ES cell viability [80, 83]. The targeting strategy in this study did not disrupt the genomic sequence that encodes that catalytic lysine, raising concerns that the resultant allele could be a hypomorphic allele [80, 83]. In our study, we disrupted exons 18 – 19, which deletes the active site lysine (K583) and causes a frame shift of all downstream exons, ensuring that the generated allele is a true null allele. Ligase I-null CH12F3 cells are viable. They do not have any significant proliferative defect or genomic instability and are not hypersentisitve to a number of DNA damaging agents. This lack of phenotype is consistent with the DT40 study, indicating that another ligase is able to compensate for Ligase I in its absence.

The Ligase I-null cells do not show any CSR defect, indicating that in NHEJ proficient cells, Ligase I is dispensable for CSR. DSBs during CSR are repaired by the NHEJ pathway and Ligase IV deficient cells have a defect in CSR

(~50% of WT levels) [56]. CSR in Ligase IV deficient cells is mediated by A-EJ, a poorly characterized repair pathway. The ligation of DSBs during A-EJ mediated CSR must be mediated by Ligase I, Ligase III or by both. In order to test whether Ligase I is essential for A-EJ, we created cells that lack both Ligase I and Ligase IV. These cells switch to the same levels as Ligase IV-null cells, indicating that Ligase I is not essential for A-EJ during CSR. This indicates that either Ligase III is essential for A-EJ, or that Ligase I and Ligase III have redundant roles in A-EJ. Neither Ligase III depletion by small hairpin RNA, nor ablation of *XRCC1*, a factor that binds to and stabilizes Ligase III inhibits A-EJ during CSR in Ligase IV deficient cells, indicating that most likely Ligase III and Ligase I have redundant functions in A-EJ [58, 87]. The creation of a cell line lacking Ligase IV and III is essential to conclusively identify the A-EJ ligase. We have constructed this cell line and are currently in the process of characterising it (data not shown).

The A-EJ pathway has been implicated in error-prone end joining that results in chromosomal translocations [59, 88]. These translocations can result in malignancies such as Burkitts lymphoma, diffuse large B cell lymphoma and plasmacytomas. Thus identifying the ligase responsible for A-EJ will help provide insight into mechanism of these diseases and may help identify diagnostic and prognostic markers for these diseases.

Our studies indicate a previously unanticipated level of functional redundancy between the three DNA ligases. Ligase I-Ligase IV-null cells have only one functional DNA ligase i.e. Ligase III. These cells are viable and show no proliferative defect, indicating that Ligase III alone can sustain cell viability. In

addition, they show no significant increase in sensitivity to a number of DNA damaging agents as compared to WT, Ligase I deficient or Ligase IV deficient cells indicating that Ligase III alone can mediate DNA repair by a number of repair pathways. This study reprioritizes the roles of the ligases, clearly showing a remarkable amount of functional overlap between the three DNA ligases. It will be interesting to further study and identify the extent of functional redundancy between the ligases and to identify the contribution of each DNA ligase to various aspects of DNA metabolism.

CHAPTER 5: SUMMARY AND FUTURE DIRECTIONS

The process of CSR diversifies immunoglobulin molecules by changing the constant region while maintaining the same variable region. This changes the effector function of the antibody, while maintaining the same antigen specificity allowing the immune system to effectively combat pathogens.

The work presented in this dissertation helps provide a better understanding of the molecular mechanism of CSR. In Chapter 2, we provide evidence for the importance of a short sequence motif, 'WGCW' (W=A/T) in CSR. This motif occurs with a higher frequency in S regions and provides overlapping AID hotspots on both DNA strands. AID targeting of both cytosines within a WGCW motif would result in closely spaced uracils on both DNA strands. Processing of these uracils would result in closely spaced nicks, which act as DSBs. We showed that these motifs are an important determinant of S region quality, and CSR efficiency correlates with the WGCW density within S regions. Other studies from our laboratory and from our collaborators have confirmed that WGCW motifs within S regions have a large influence on CSR efficiency.

In Chapter 3, we focus on understanding the mechanism of DSB formation during CSR. S region uracils can be processed by the BER pathway to form DSBs and it is known that UNG2, a BER factor, is essential for CSR. We show that APE1, another BER factor, is also essential for CSR. It has been hypothesized that APE1 nicks the phosphodiester backbone at abasic sites during CSR. However, this had not been definitively proven. Our work unequivocally shows that APE1 is essential for CSR. In addition, we established the first APE1 deficient mammalian cell line, which will prove useful in studying

the role of APE1 in DNA repair as well as its role as a redox factor. APE1 variants have been associated with a number of different types of cancer. The APE1-null cell line will allow for an investigation of the functional differences between WT and the cancer-associated APE1 variants.

Finally, in Chapter 4 we focus on DSB repair during CSR. We look at the role of DNA ligases in CSR. It was known from previous studies that DSBs during CSR are repaired by Ligase IV mediated NHEJ. Cells that lack Ligase IV switch to ~50% of WT levels. The end-joining in these cells is mediated by A-EJ mediated by DNA Ligase I, Ligase III or both. We show that DNA Ligase I is not essential for A-EJ mediated CSR. This indicates that either DNA Ligase III is essential for A-EJ mediated CSR or DNA Ligase I and III have redundant functions in A-EJ. To distinguish between these possibilities, it is essential to create cells lacking Ligase IV and Ligase III. Ligase III depletion is lethal in somatic cells. This lethality can be rescued by expressing a mitochondrial-targeted DNA Ligase. We have created Ligase IV-Ligase III-null cells that express a Ligase I-ZsGreen fusion protein that is targeted to mitochondria. Preliminary studies indicate that these cells switch to a similar level as Ligase IV deficient cells, indicating that Ligase I and III have redundant functions in A-EJ mediated CSR. However, these cells have a proliferative defect and we have not been able to express WT Ligase III to successfully perform genetic complementation, which has made it difficult to fully characterise these cells. In addition, expression of the mitochondrial transgene in these cells could not be detected by immunoblotting, indicating that the proliferative defect of these cells

could be due to a mitochondrial defect. We are currently testing other mitochondrial-targeted transgenes including the catalytic core of Ligase I and the Chlorella DNA ligase.

Our work with the DNA ligases has shown that the three mammalian DNA Ligases do not have distinct functions, but in fact show a large amount of functional overlap. This reprioritizes the role of the three DNA ligases, and calls for a more detailed study of their functions in DNA replication and repair.

The studies presented in this dissertation have increased our understanding of the mechanism of CSR. In addition, the knock-out cell lines created as a part of these studies will serve as an important tool to study the functions of these proteins in DNA replication and repair as well as allow us to study their contribution to genomic instability and cancer.

APPENDICES

Appendix A: Chapter 1 Figures

Figure 1. Structure of an Immunoglobulin Molecule. An antibody molecule consists of two heavy and two light chains joined by disulfide bonds. They have an N terminal Variable region (shown in red) and a C terminal Constant region (shown in blue). N terminus and C terminus of polypeptide chains, light chains, heavy chains, disulfide bonds, variable and constant regions are indicated.

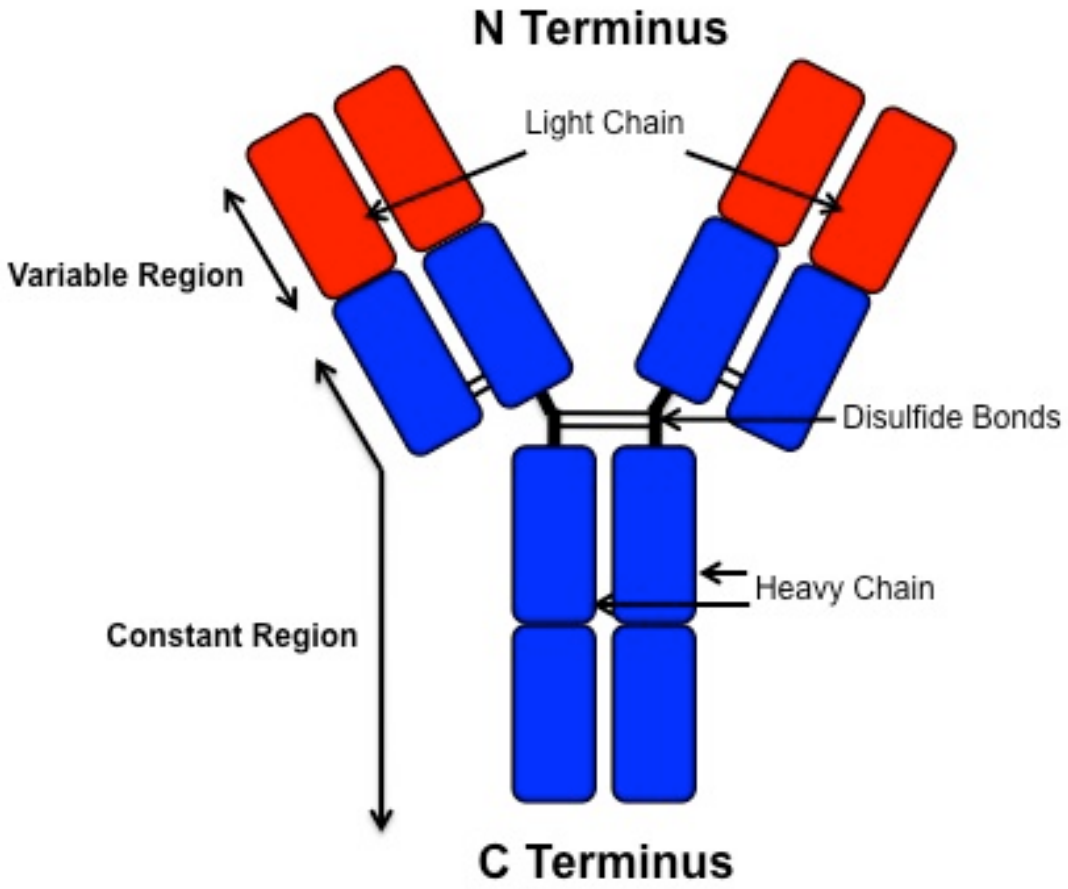


Figure 2. Overview of Class Switch Recombination. The Ig heavy chain locus is depicted. Rearranged variable region is shown as a pink rectangle, switch regions are indicated by blue ovals, constant regions are indicated by purple rectangles. Black arrows indicate transcription orientation and red lines indicate DSBs. During CSR, DSBs are created in donor and acceptor switch regions. These DSBs are joined to delete the intervening DNA and juxtapose the variable region to a new constant region, in this case Cγ1. The cell will now produce IgG1.

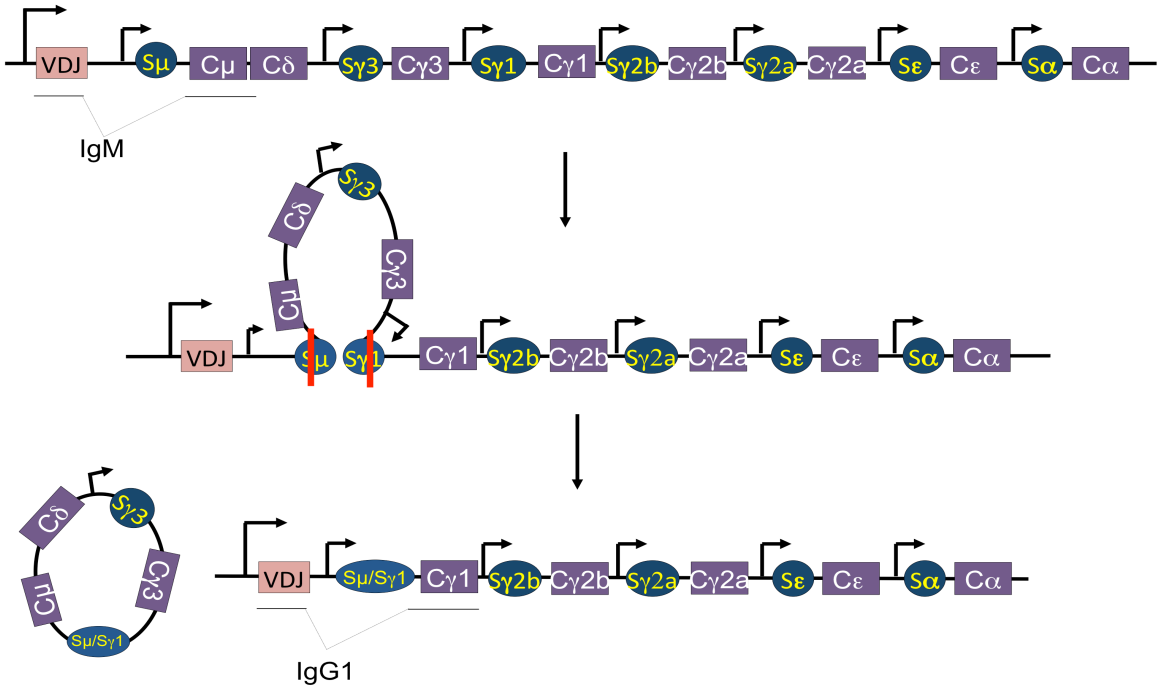


Figure 3. Molecular Mechanism of Class Switch Recombination. Transcription through switch regions results in the formation of secondary structures such as R-Loops. This provides the single-stranded DNA substrate for AID, which converts cytosines to uracils. Uracils are recognized and processed by UNG2-APE1 mediated Base Excision Repair or MSH2-mediated Mismatch Repair pathways to result in DSBs. DSBs are recognized and synapsed, a process that involves multiple factors such as γ-H2AX, 53-BP1 and ATM. The DSBs are joined by the NHEJ pathway to complete the recombination process.

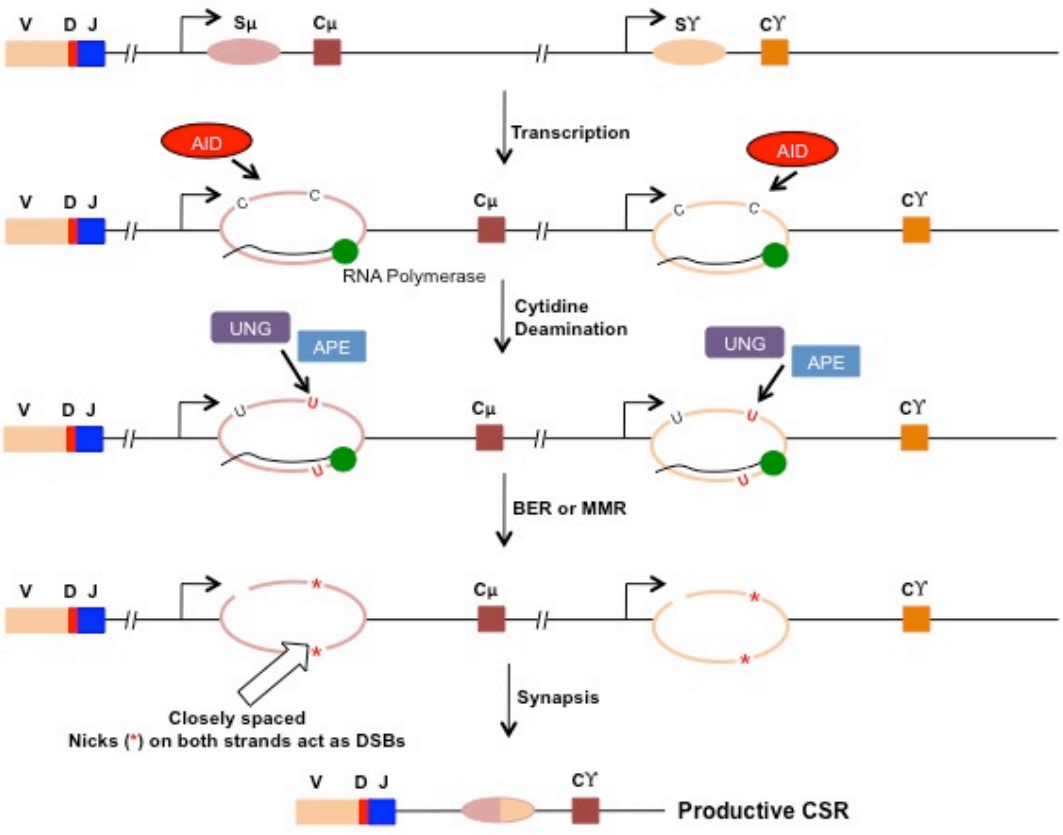


Figure 4. Germline Transcription during Class Switch Recombination. A representative constant region transcriptional unit is depicted. P, Cytokine Dependent Promoter; I, I exon; S, switch region; C, constant region exons; SD, splice donor; SA, splice acceptor. Cytokine stimulation activates transcription from the cytokine-responsive promoter. The primary transcript consists of the I exon, intronic switch region and the constant region exons. The primary transcript is spliced to produce the germline transcript that consists of I exon and constant region exons.

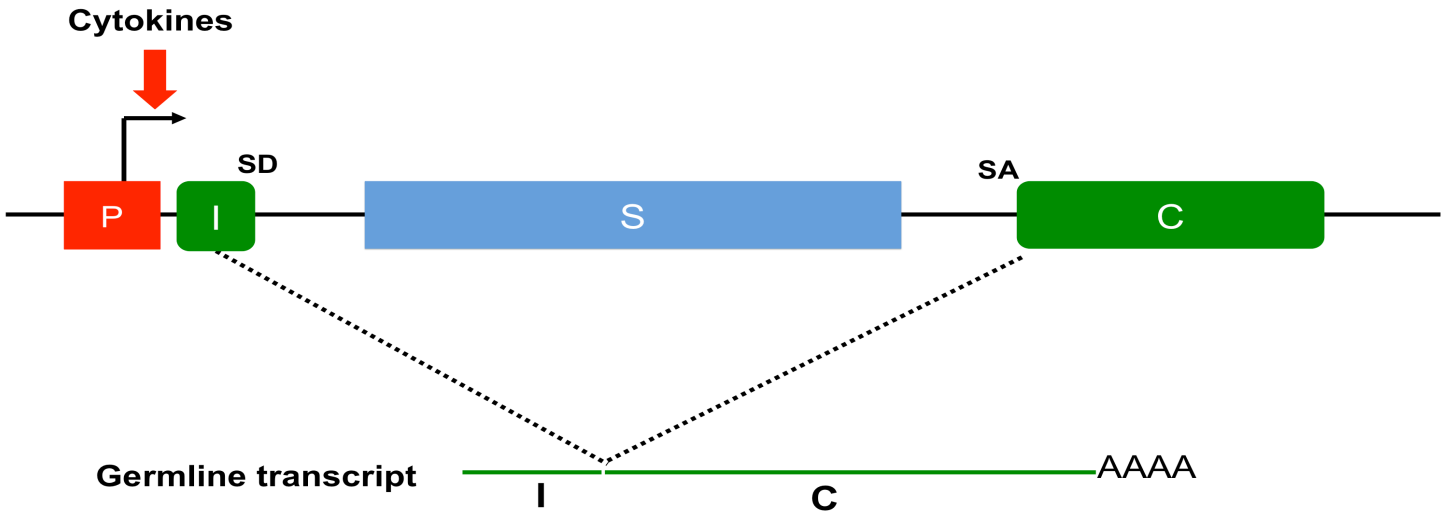
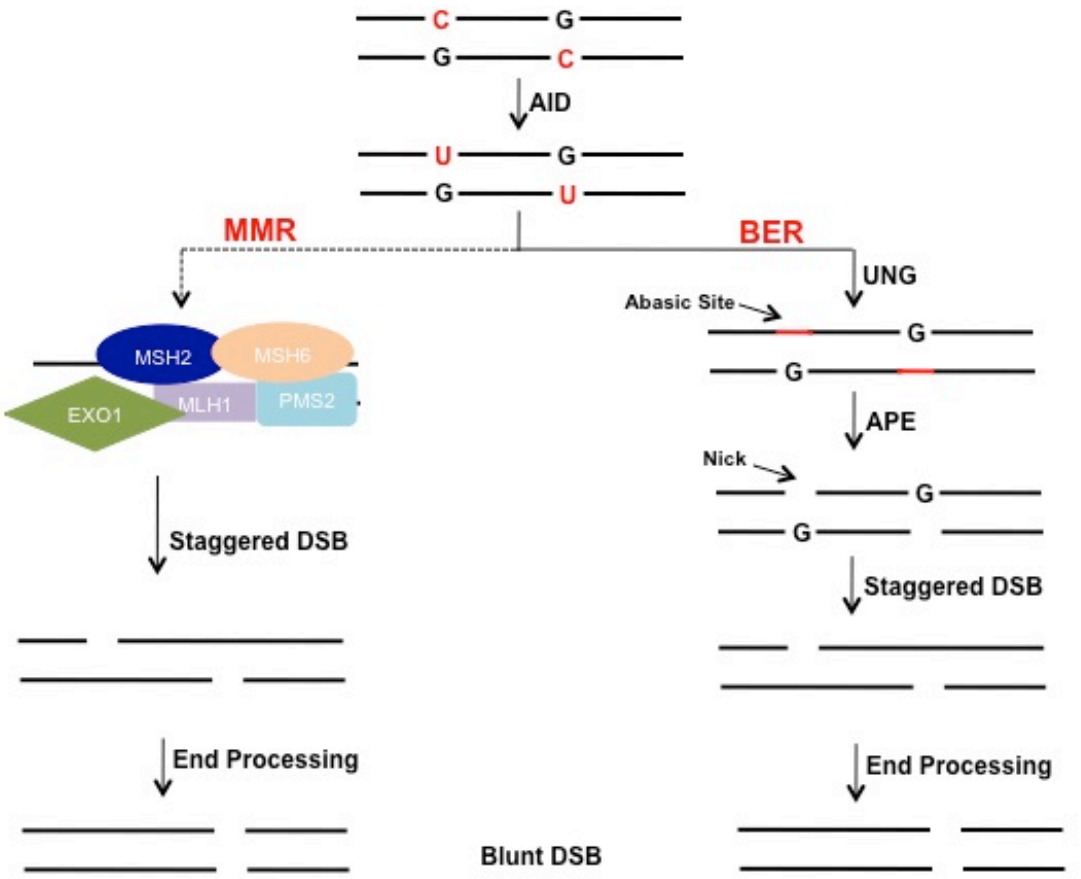
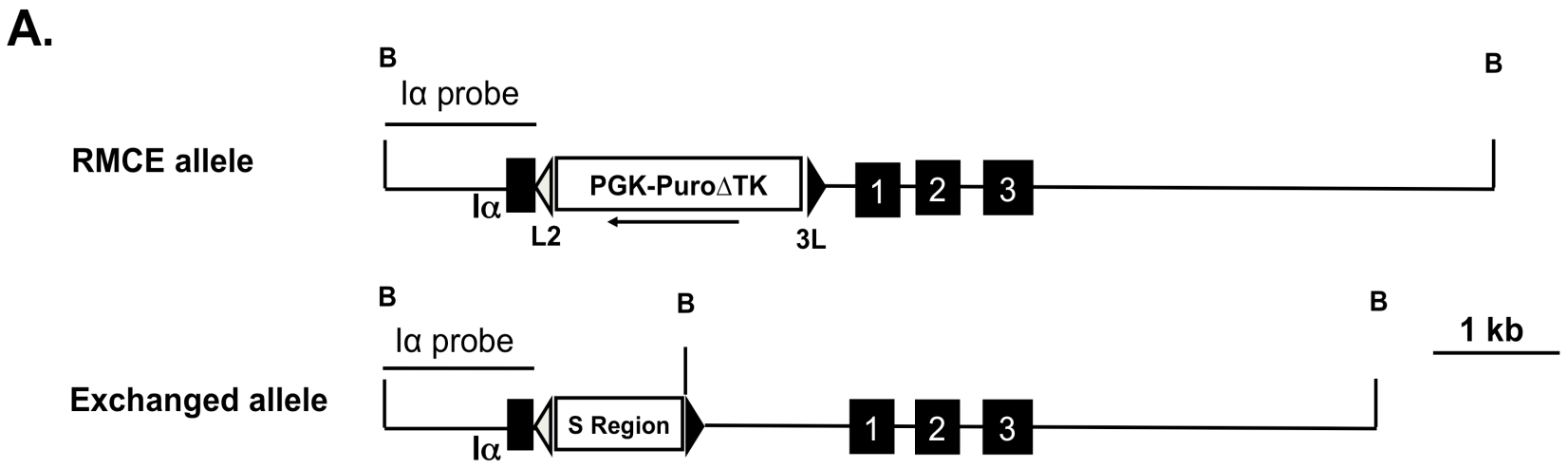


Figure 5. Double Stranded Break Formation during Class Switch Recombination. AID deaminates cytosines within switch regions converting them to uracils. These uracils can be processed by the Mismatch Repair pathway mediated by factors including MSH2, MHS6, MLH1, PMS2 and EXO1 to result in DSBs. Uracils are also processed by UNG-APE1 mediated Base Excision Repair pathway to produce two single-stranded nicks on each DNA strand that act as DSBs.



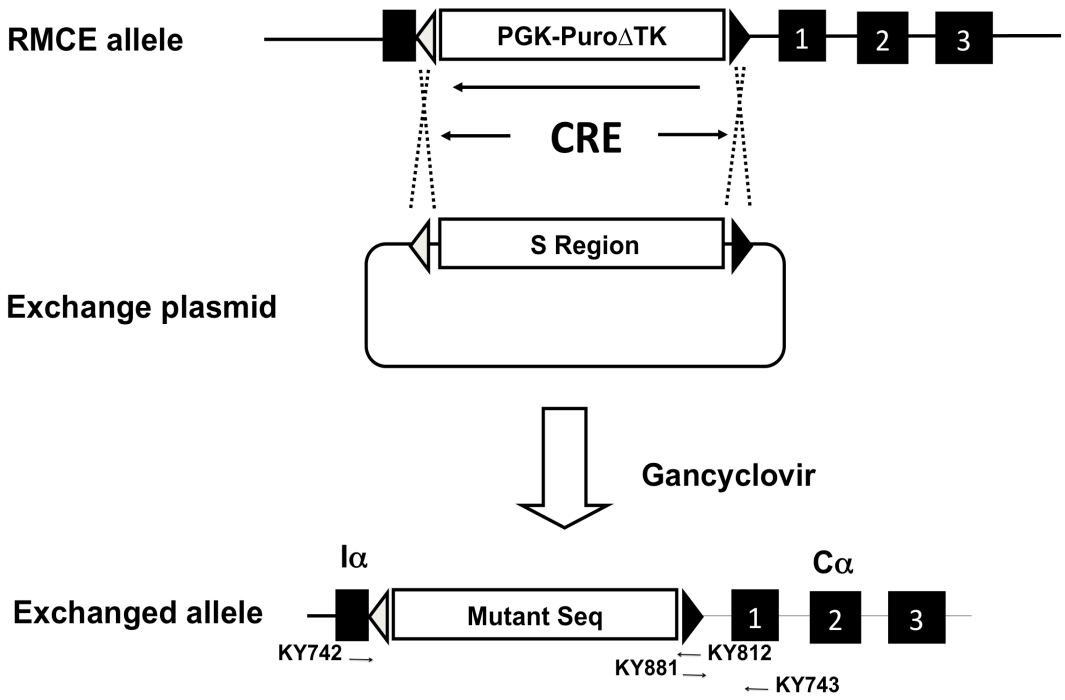
Appendix B: Chapter 2 Figures

Figure 6. Sα Recombination Mediated Cassette Exchange in 1F7 cells.

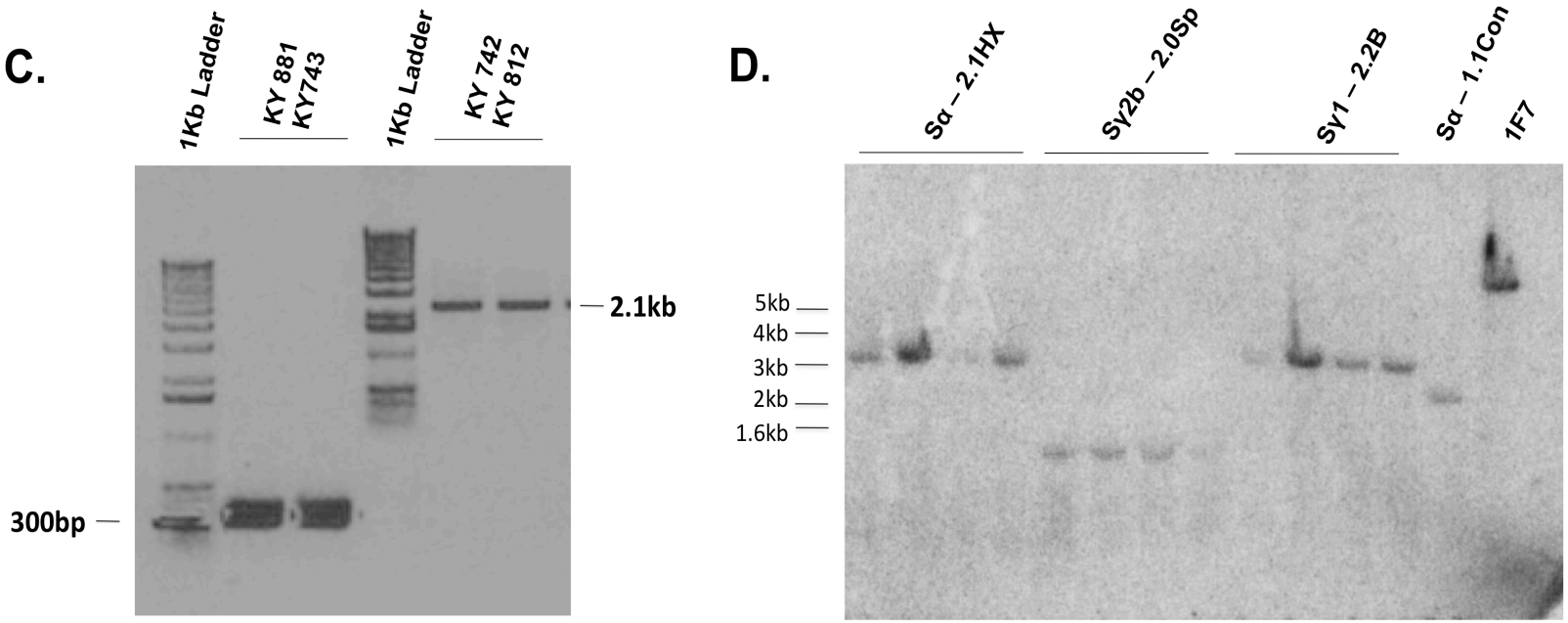


(Figure 6 cont'd)

B.



(Figure 6 cont'd.)



(Figure 6 cont'd.)

A. Genomic organization of the Sα locus in 1F7 cells. RMCE allele represents the IgH α locus in 1F7 cells in which Sα is replaced by the RMCE cassette. Exchanged allele represents the IgH α locus after successful exchange, with a 1 kb S region sequence introduced in place of Sα. The map is drawn to scale. Exons are indicated by filled squares. L2 and 3L are wild type and mutant loxP sites, respectively. PGK, phosphoglycerate kinase promoter; Puro, puromycin resistant gene; ΔTK, truncated thymidine kinase gene; B, Bgl II restriction enzyme site. Probe used in Southern Blot is indicated on top. **B.** Recombination Mediated Cassette Exchange. Exchange plasmid containing floxed mutant sequence and Cre expression plasmid are contransfected into 1F7 cells. Successful exchange events are screened by counter-selection against TK gene using gancyclovir. KY472, KY881, KY812, KY743 are PCR primers used to screen for successful exchange events. **C.** PCR screening of RMCE events. KY881, KY743 and KY742, KY812 are the primer pairs used. A representative gel picture is depicted with expected product sizes. **D.** Southern Blot of Bgl II digested DNA from WT and exchanged clones with Iα probe. Expected sizes: Sα - 2.1HX, 3.5kb; Sγ2b – 2.0Sp, 1.4kb; Sγ1 – 2.2B, 3.5kb; Sα - 1.1Con, 2.4kb; 1F7, 8.8kb.

Figure 7. Sγ sequences can substitute for Sα in Class Switch Recombination. CSR assay of Sα replacement constructs. Each point represents CSR efficiency for one assay with one clone.

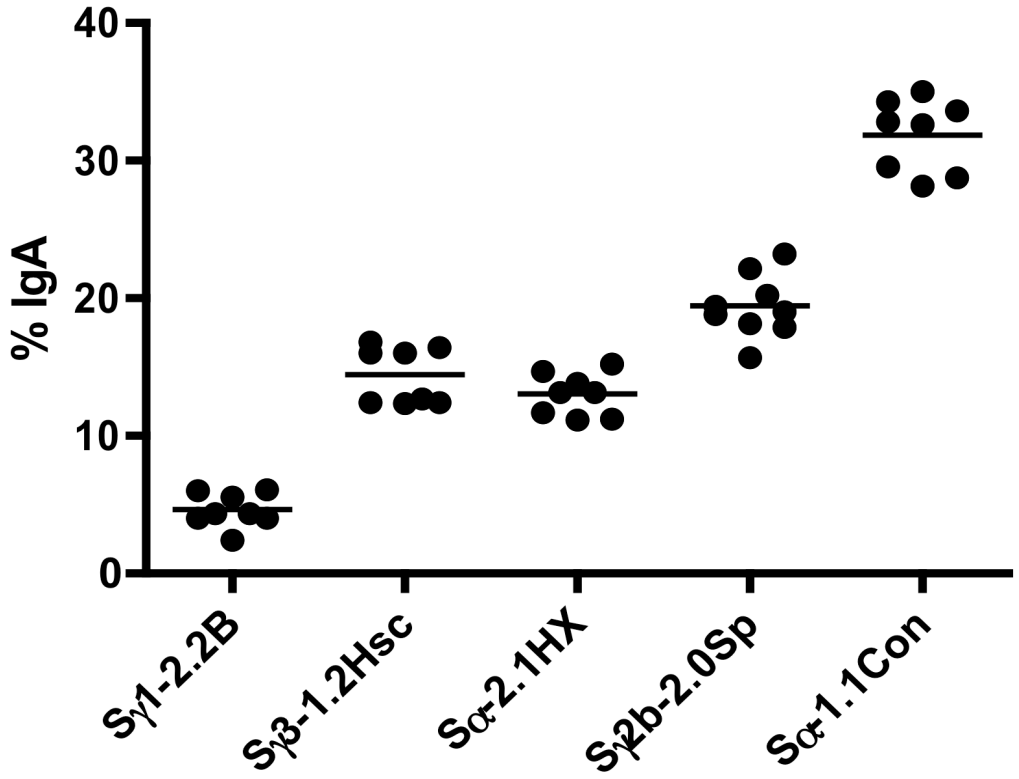
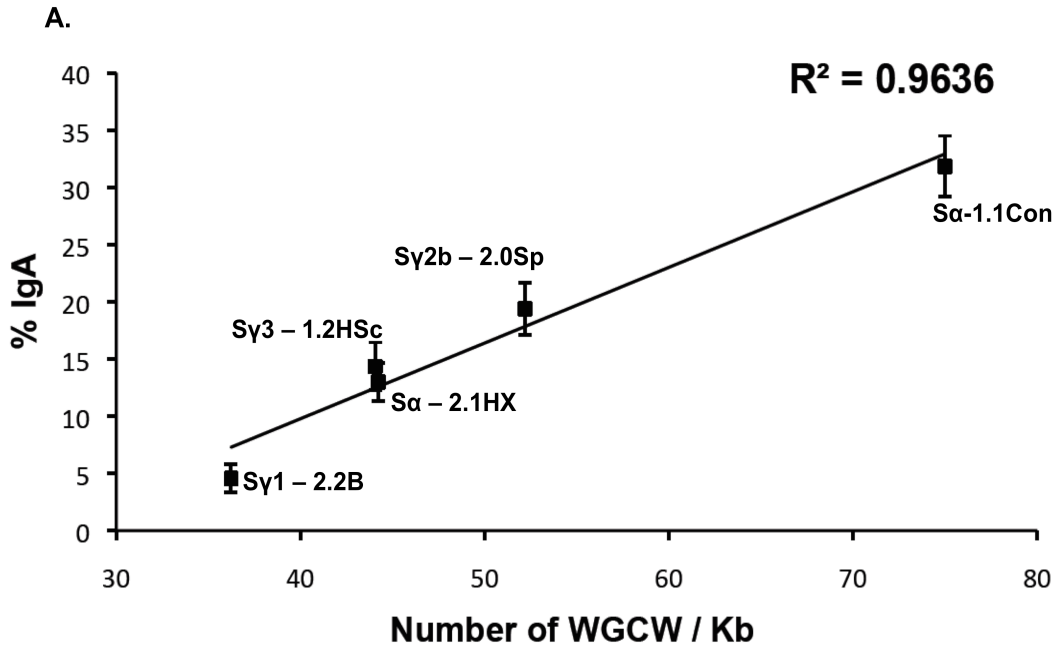
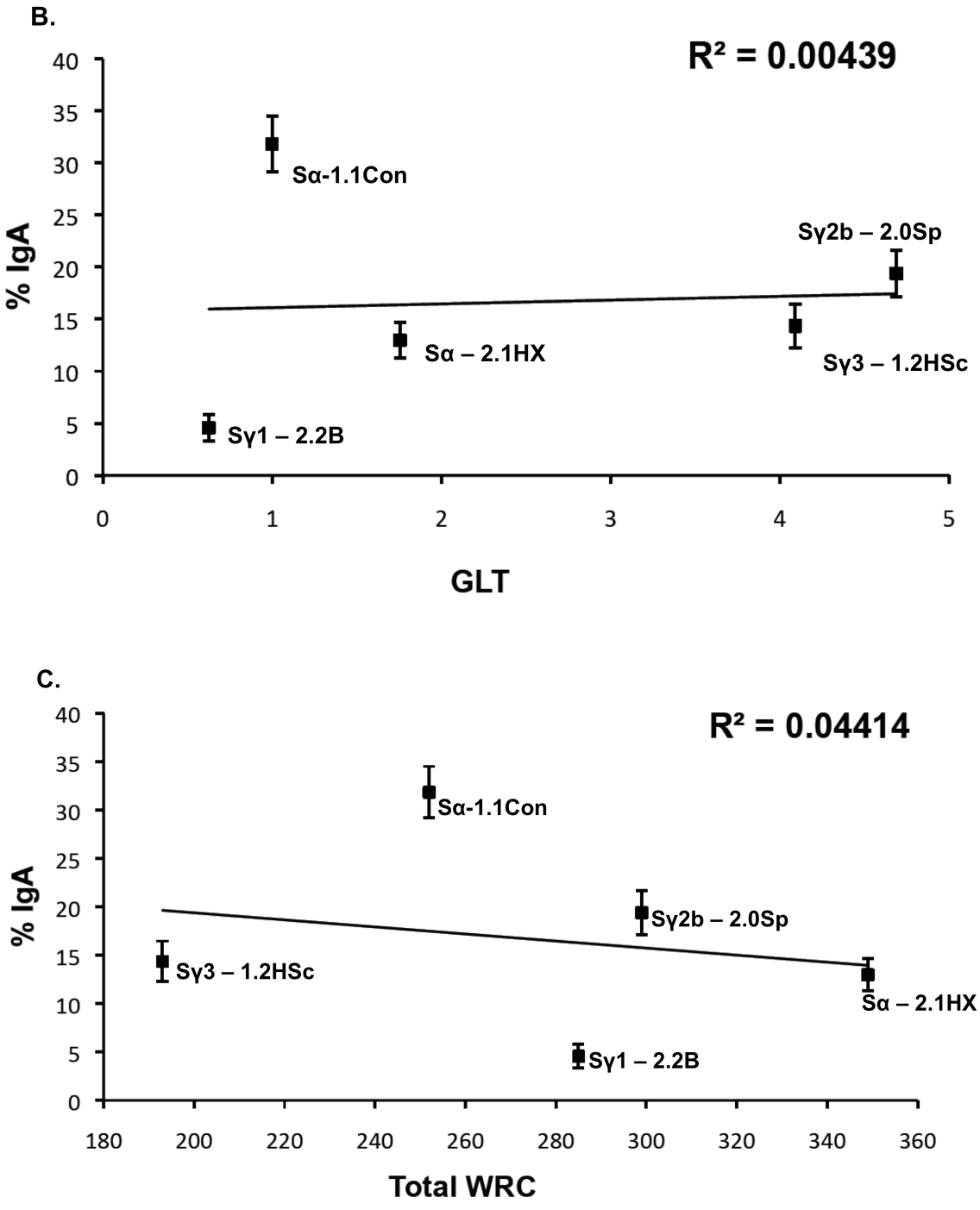


Figure 8. Class Switch Recombination efficiency correlates with switch region WGCW density. A. CSR correlates with S region WGCW density; WGCW density=Number of WGCW / Kb. Error bars represent standard deviation of three independent experiments. **B-D.** CSR efficiency does not correlate with GLT or the number of S region WRC or WGCW sites **E.** CSR shows a moderate corelation with S region WRC density.

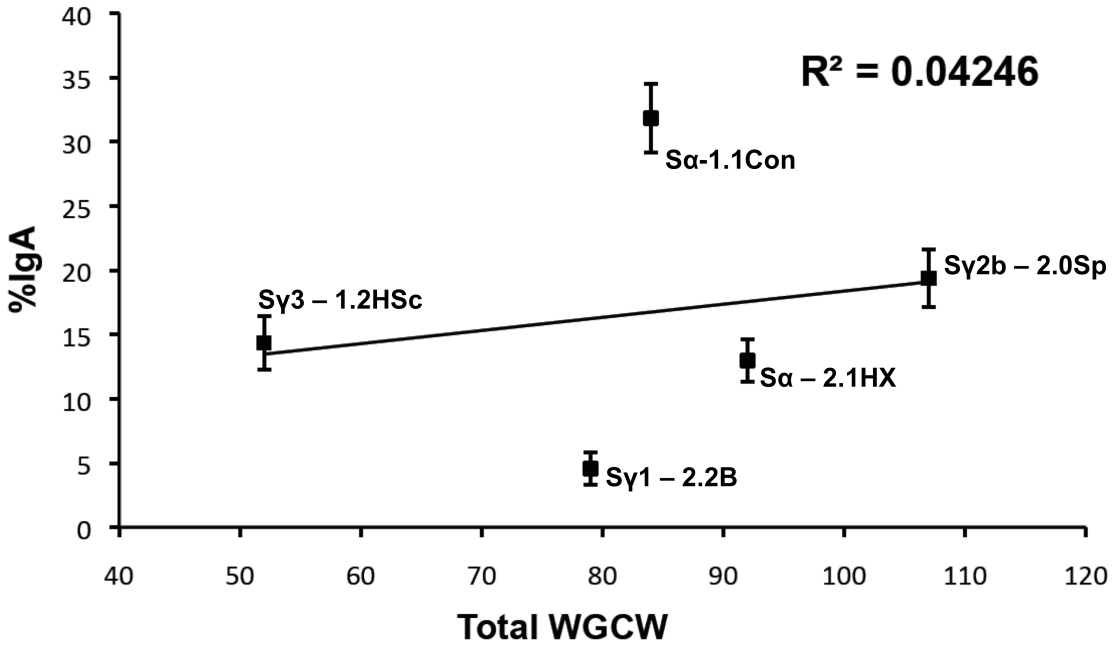


(Figure 8 cont'd.)



(Figure 8 cont'd.)

D.



E.

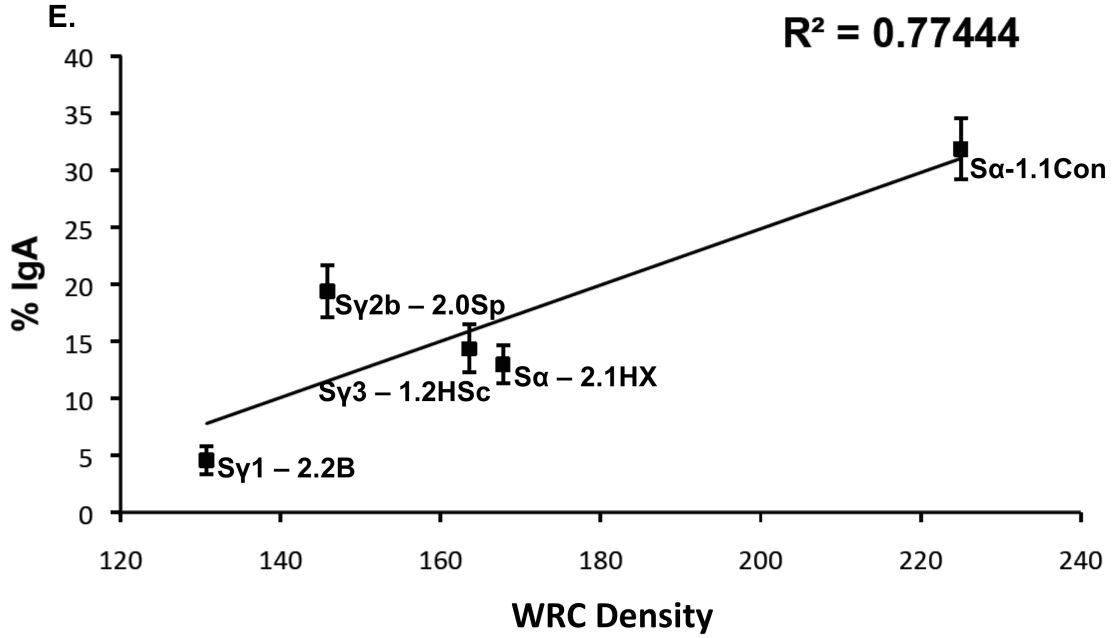
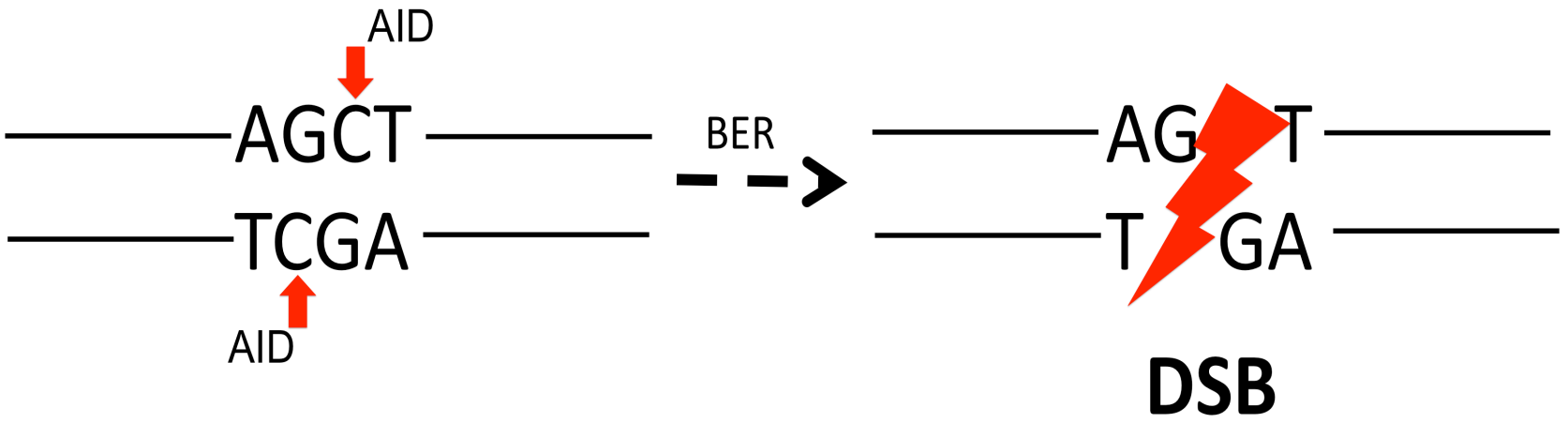
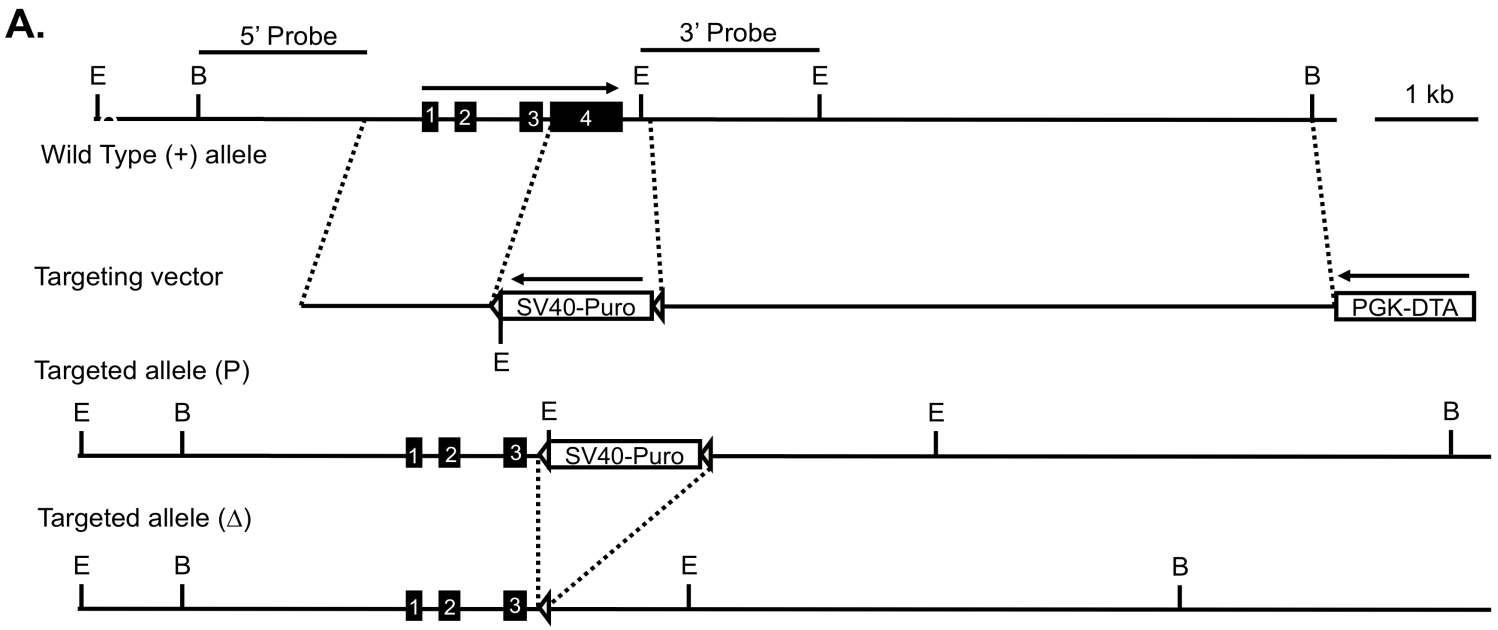


Figure 9. Overlapping AID hotspots in Class Switch Recombination. WGCW motifs provide AID hotspots on both DNA strands. AID action can convert both cytosines in the WGCW motif to uracils. Uracils are then processed by BER factors to result in closely spaced nicks on both strands, which spontaneously form DSBs.

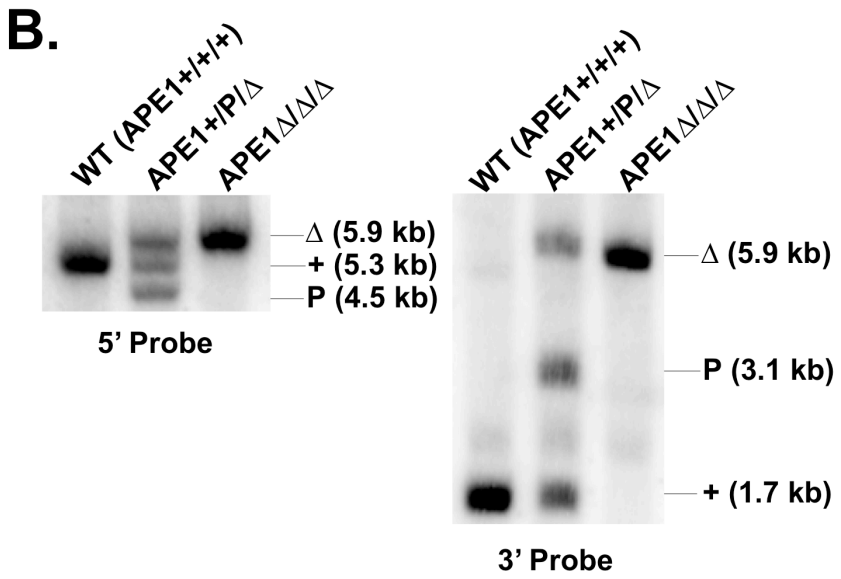


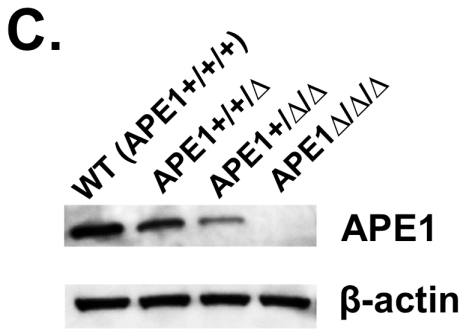
Appendix C: Chapter 3 Figures

Figure 10. Construction of an APE1-null cell line.



(Figure 10 cont'd.)

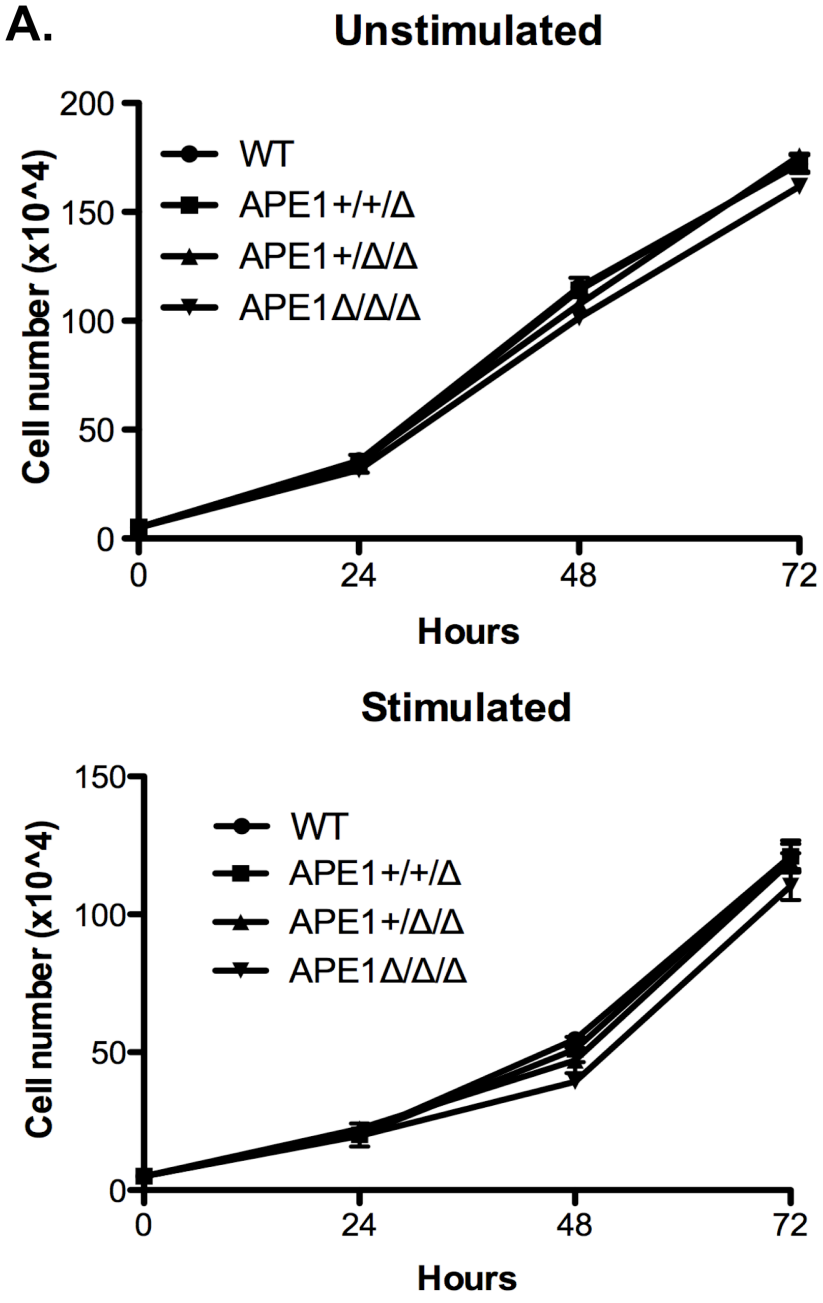




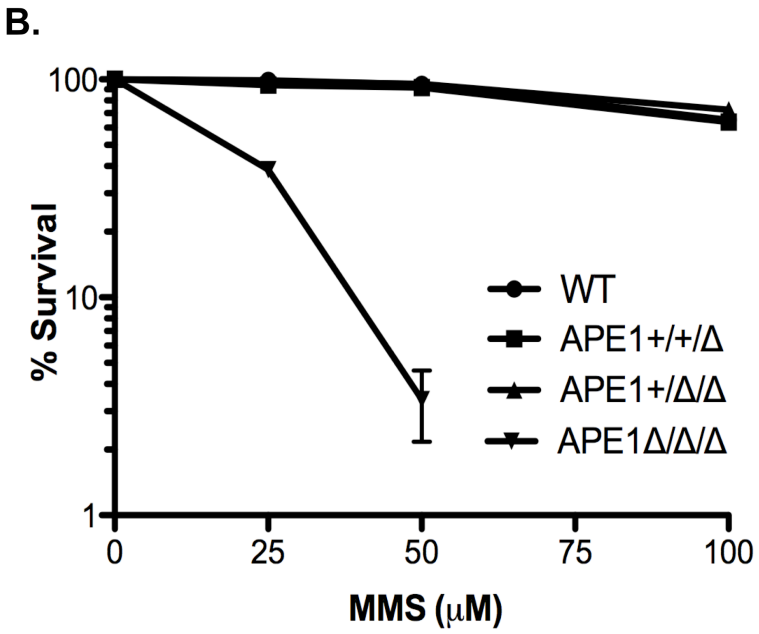
(Figure 10 cont'd.)

A. Genomic organization of the wild type and targeted *APE1* allele. The map is drawn to scale. Exons are indicated by numbered boxes. Arrows indicate transcription orientations. SV40, SV40 early promoter; PGK, phosphoglycerate kinase promoter; B, Bgl II restriction enzyme site; E, EcoR I restriction enzyme site. Probes used in Southern blot analysis are depicted at the top. + indicates wild-type allele, P and Δ indicates targeted allele with or without the puro-cassette, respectively. Triangles indicate loxP sites. **B.** Southern blot analysis of EcoR I-digested genomic DNA from wild type and targeted cells. Genotypes, sizes of bands and probes are indicated. **C.** Western blot analysis of APE1 protein level in cells containing three (APE1+/+/+), two (APE1+/+/Δ), one (APE1+/Δ/Δ) and zero (APE1Δ/Δ/Δ) copies of *APE1*.

Figure 11. Proliferation and Base Excision Repair in APE1-deficient cells.



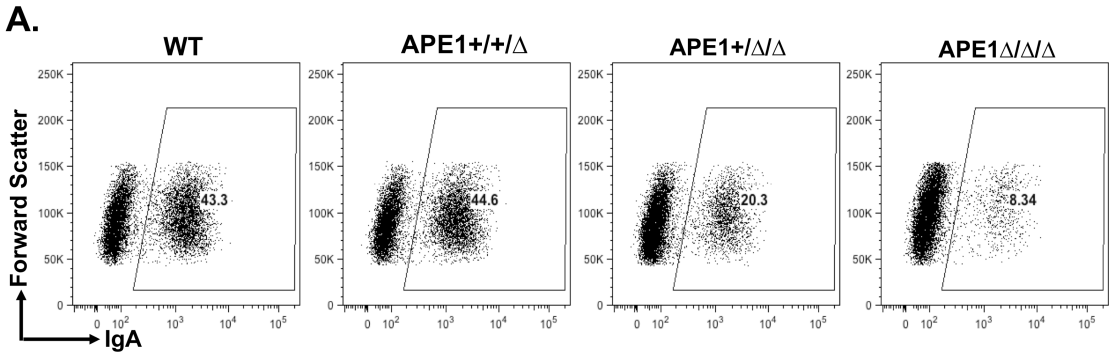
(Figure 11 cont'd.)

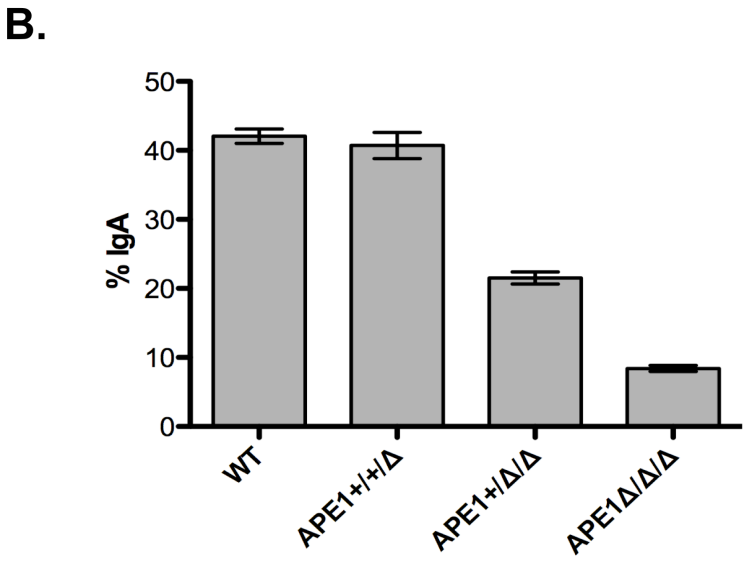


(Figure 11 cont'd.)

A. Live cell (trypan blue exclusion) counts over a span of 72 h in unstimulated (- CIT) or stimulated (+ CIT) cultures. CIT, anti-CD40, interleukin 4 and TGFβ1. Error bars indicate standard error of three independent experiments. **B.** MMS sensitivity assay in cells containing 0-3 copies of *APE1*. Cells were cultured in the presence of various concentrations of MMS for 2 days. Cell survival was determined by MTT assay. Error bars indicate standard error of three independent experiments.

Figure 12. APE1 is essential for Class Switch Recombination. A. FACS analysis of CSR by cell surface staining of IgA after 3 days of cell growth with (+ CIT) or without (-CIT) cytokines. **B.** CSR efficiency (determined as the percentage of IgA+ cells) in cells containing 0-3 copies of *APE1*. Error bars indicate standard error of ten independent experiments. **C.** APE1 affects CSR kinetics. CSR was monitored every 24 hours over a span of three days.





(Figure 12 cont'd.)

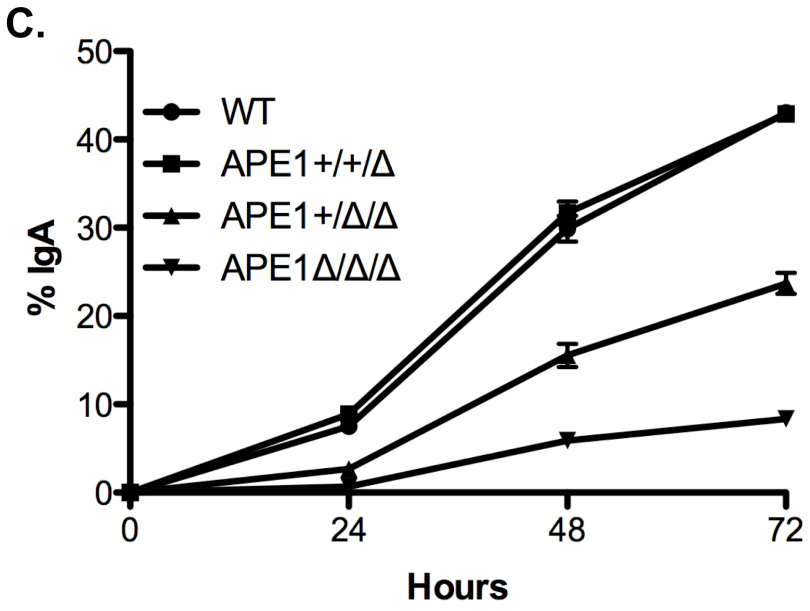
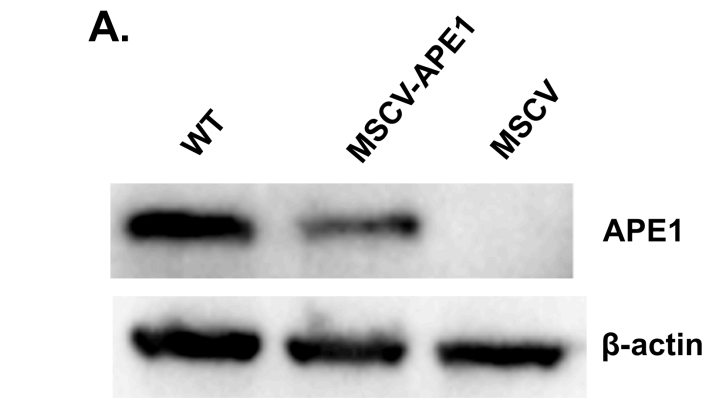


Figure 13. Genetic complementation of APE1-null cells. A. Western blot analysis of transgenic APE1 expression in retroviral transduced *APE1*-null cells. MSCV, APE1-null cells infected with control retrovirus; MSCV-APE1, APE1-null cells infected with WT APE1 expressing retrovirus **B.** CSR in retroviral-transduced *APE1*-null cells.



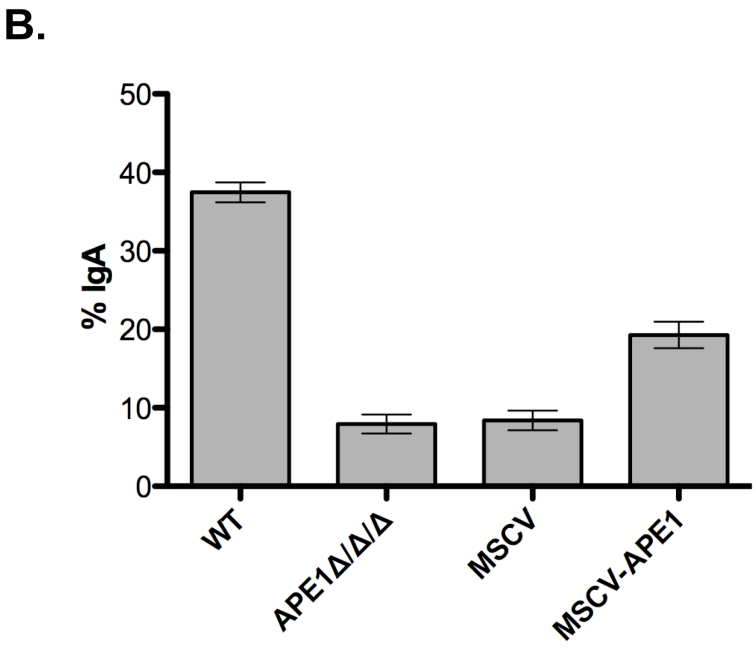
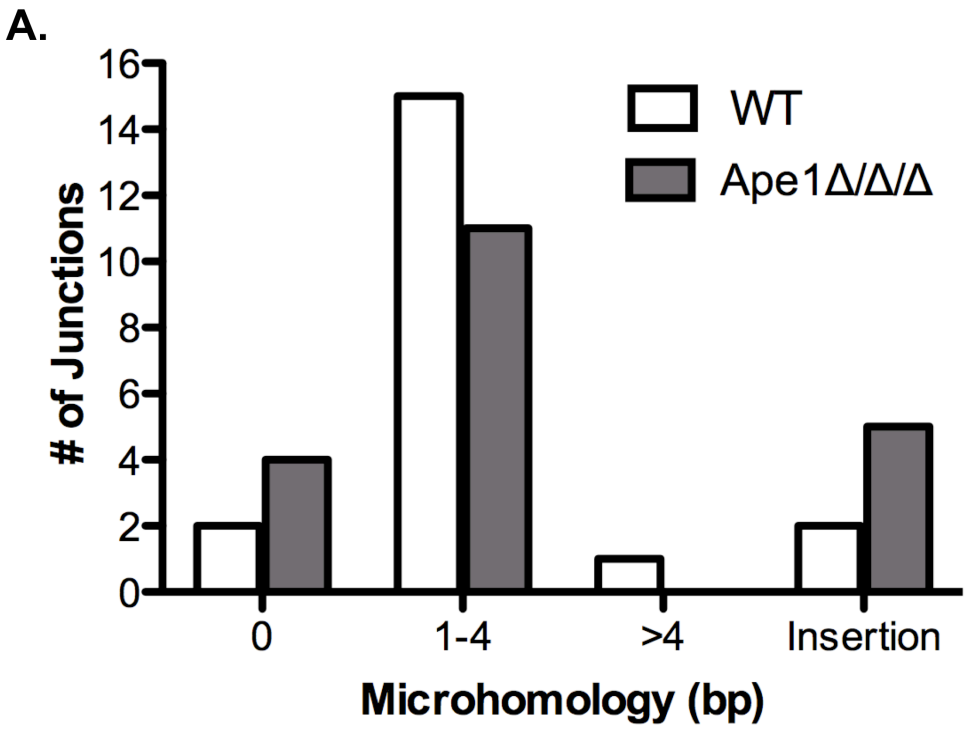
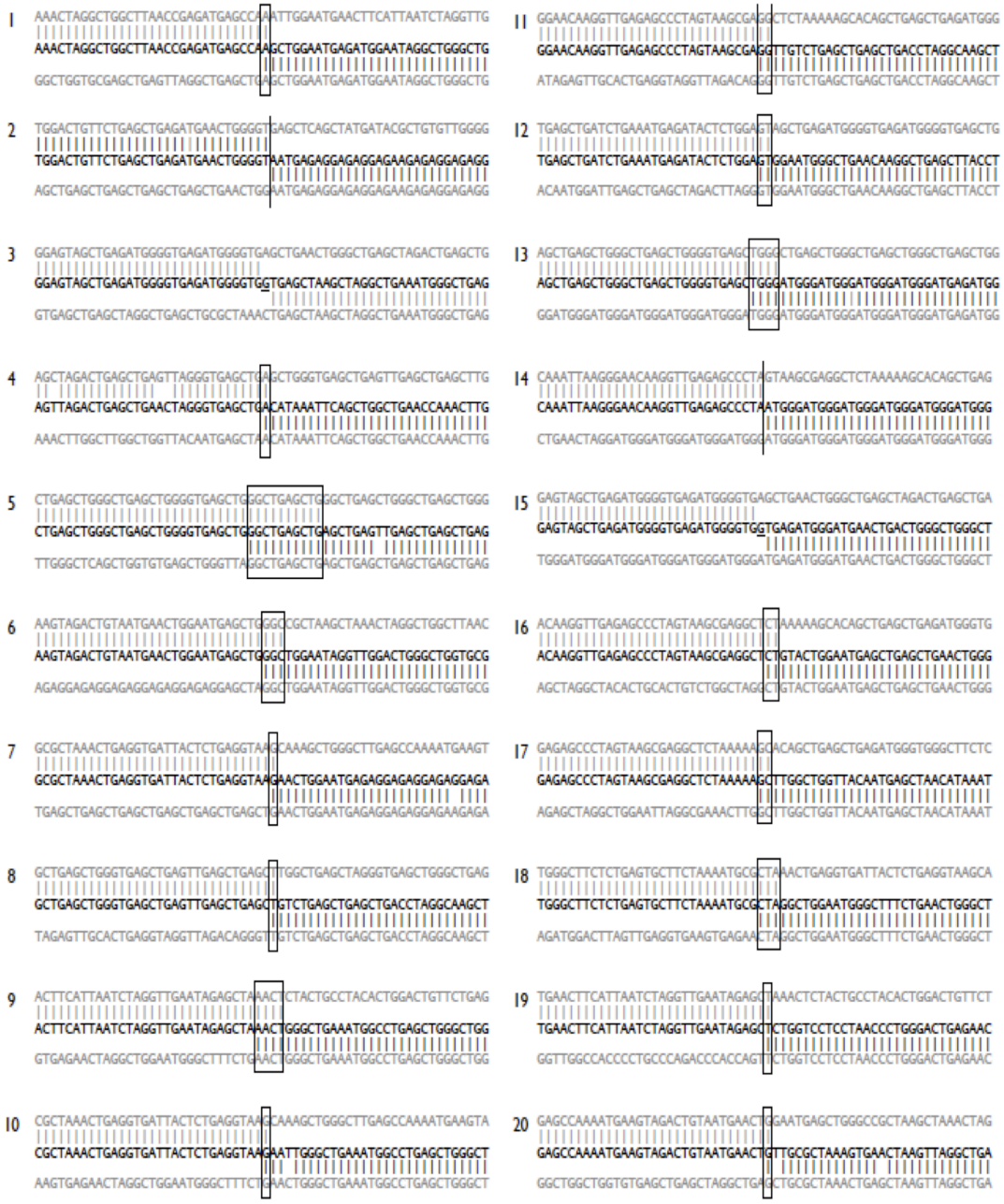


Figure 14. Sμ-Sα switch junction sequences. A. Number of switch junctions with the indicated number of nucleotide overlaps or insertions. **B, C.** Switch junction sequences from WT and APE1-null cells. Switch junctions sequences are aligned to germline sequences. Germline Sμ and Sα sequences (gray) are listed on the top and bottom, respectively, of each junction sequence (middle). Microhomologies (boxes) are identified as the largest perfect matches to the germline sequences. Nucleotide additions are underlined. Long vertical lines indicate direct joins. Small vertical lines indicate identity between the junction and germline sequences. Δ indicates a portion of germline sequences missing from the junction.



(Figure 14 cont'd.)

B. **WT**



(Figure 14 cont'd.)

C. **APE1Δ/Δ/Δ**

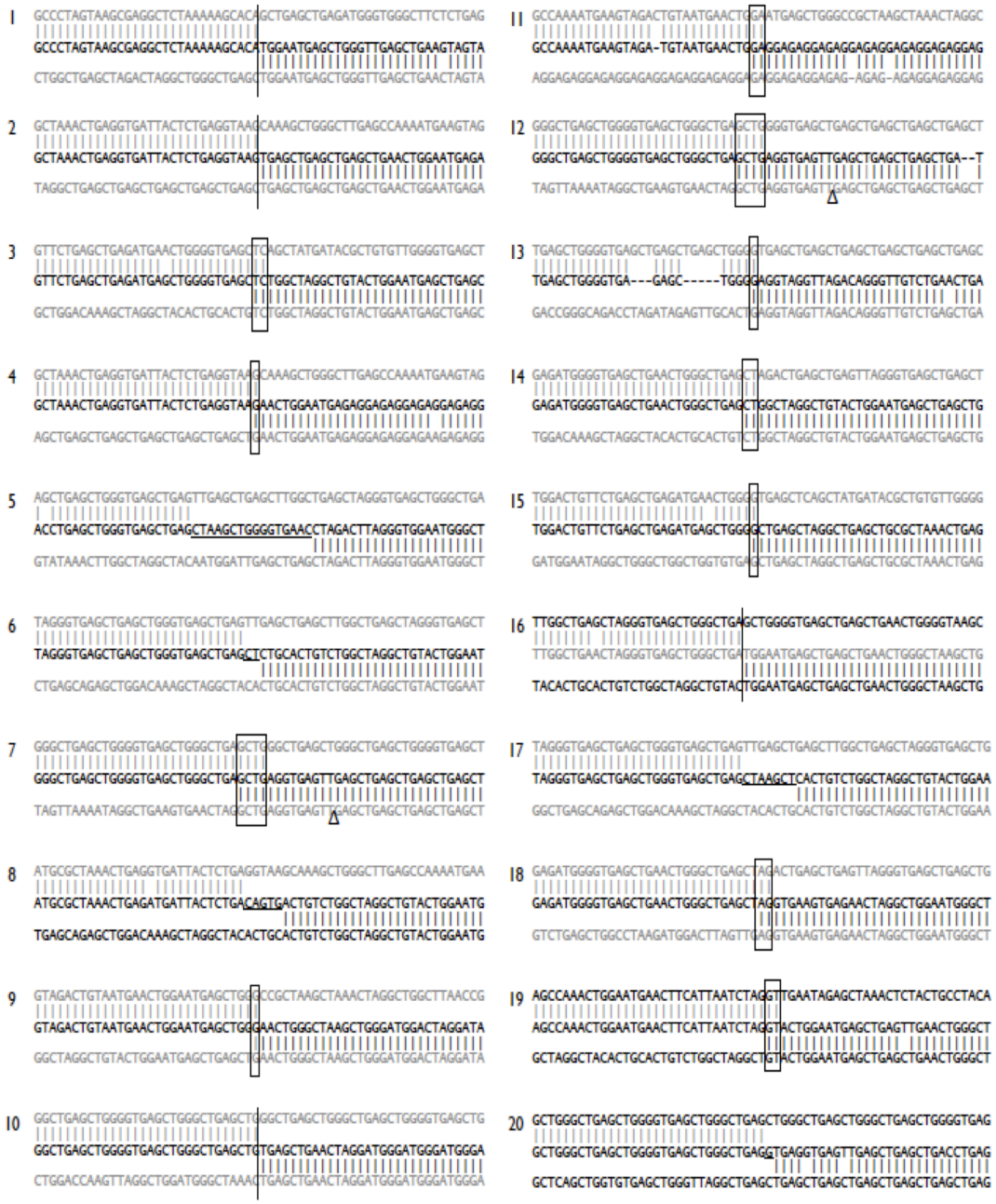
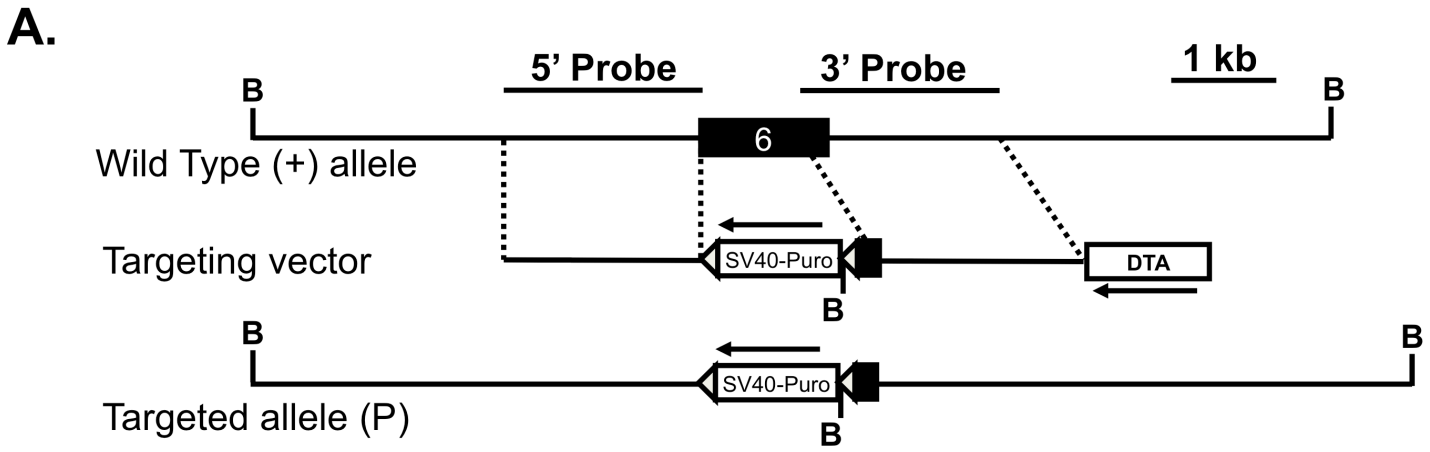
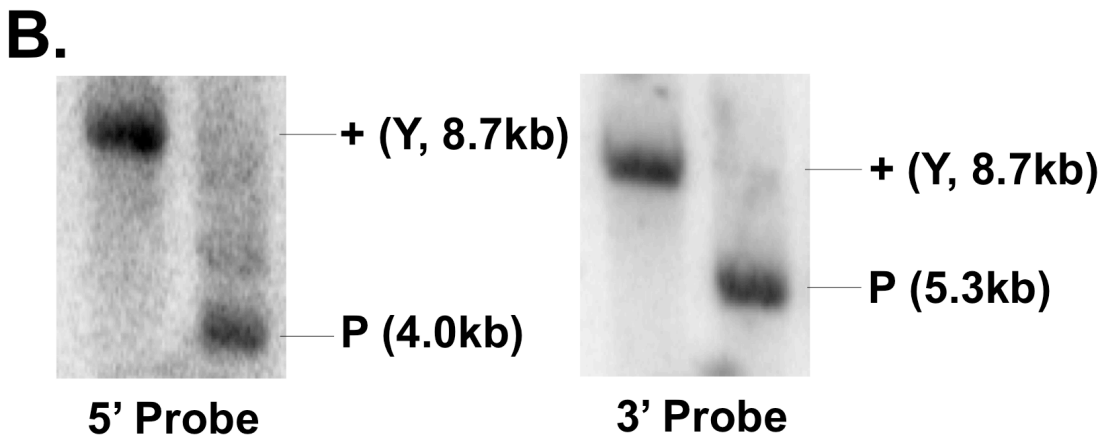


Figure 15. Construction of an APE2-null cell line. A. Genomic organization of the wild type and targeted *APE2* allele. The map is drawn to scale. Exons are indicated by numbered boxes. Arrows indicate transcription orientations. SV40, SV40 early promoter; PGK, phosphoglycerate kinase promoter; B, BamH I restriction enzyme site. Probes used for Southern Blot are shown on top. **B.** Southern blot analysis of BamH I-digested genomic DNA. Genotypes, sizes of bands and probes are indicated. **C.** Western blot analysis of APE1 and APE2 protein level in cells deficient for APE1, APE2 or both.



(Figure 15 cont'd.)



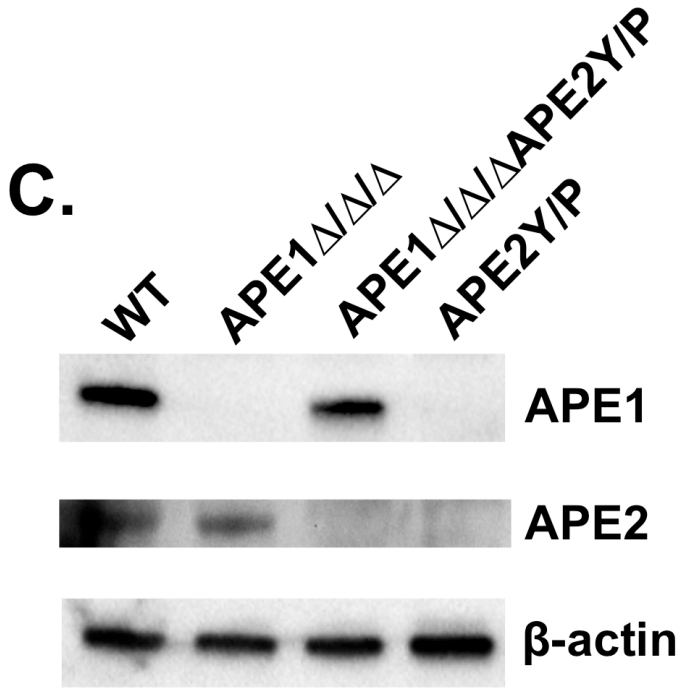
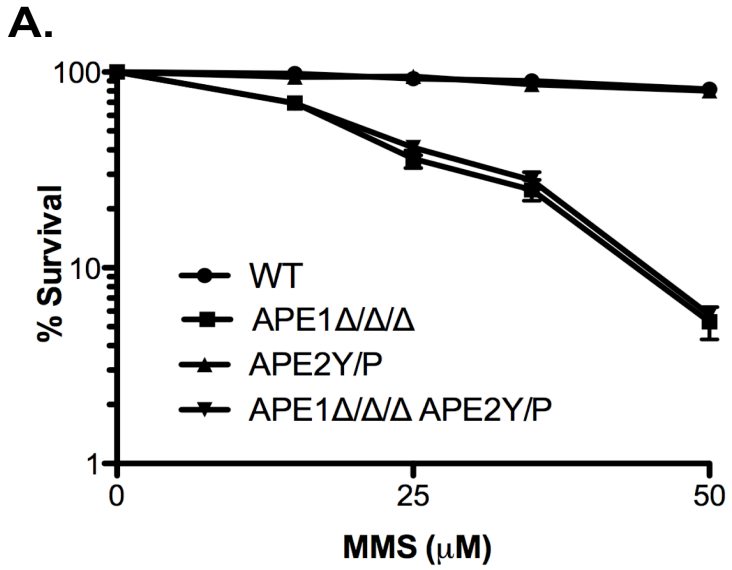
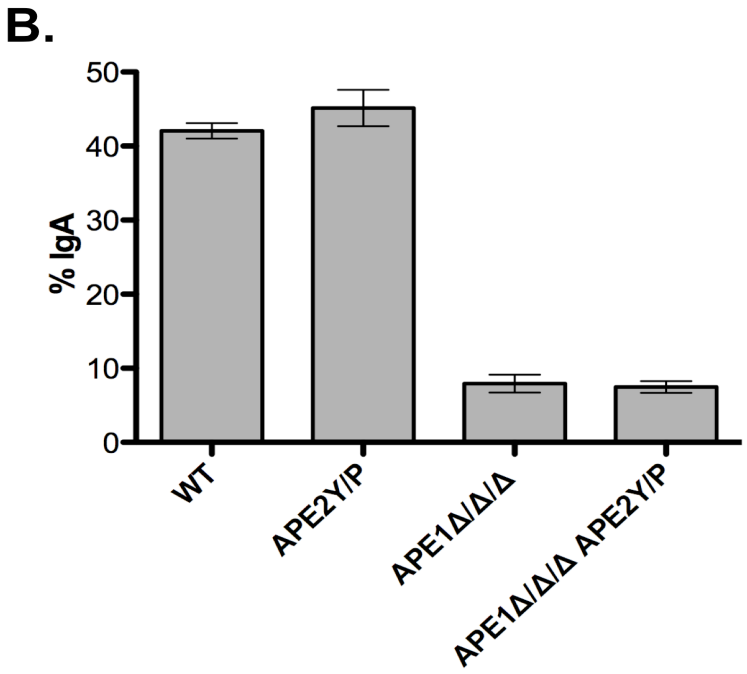


Figure 16. Base Excision Repair and Class Switch Recombination in APE2-null cells. **A.** MMS sensitivity assay of cells deficient for APE1, APE2 or both. Cells were grown in the presence of various concentrations of methyl methanesulfonate for 2 days and cell survival was determined by metabolic conversion of thiazolyl blue tetrazolium bromide (MTT) in the mitochondria of living cells. **B.** CSR in cells deficient for APE1, APE2 or both.

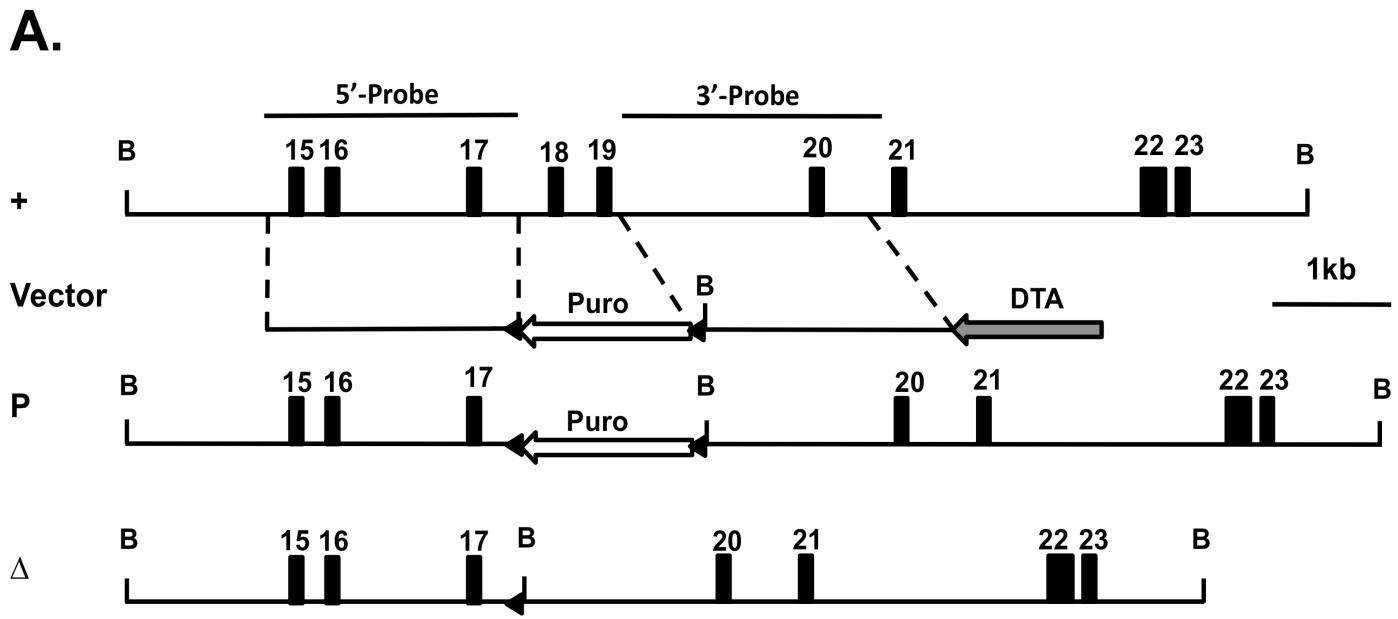


(Figure 16 cont'd.)

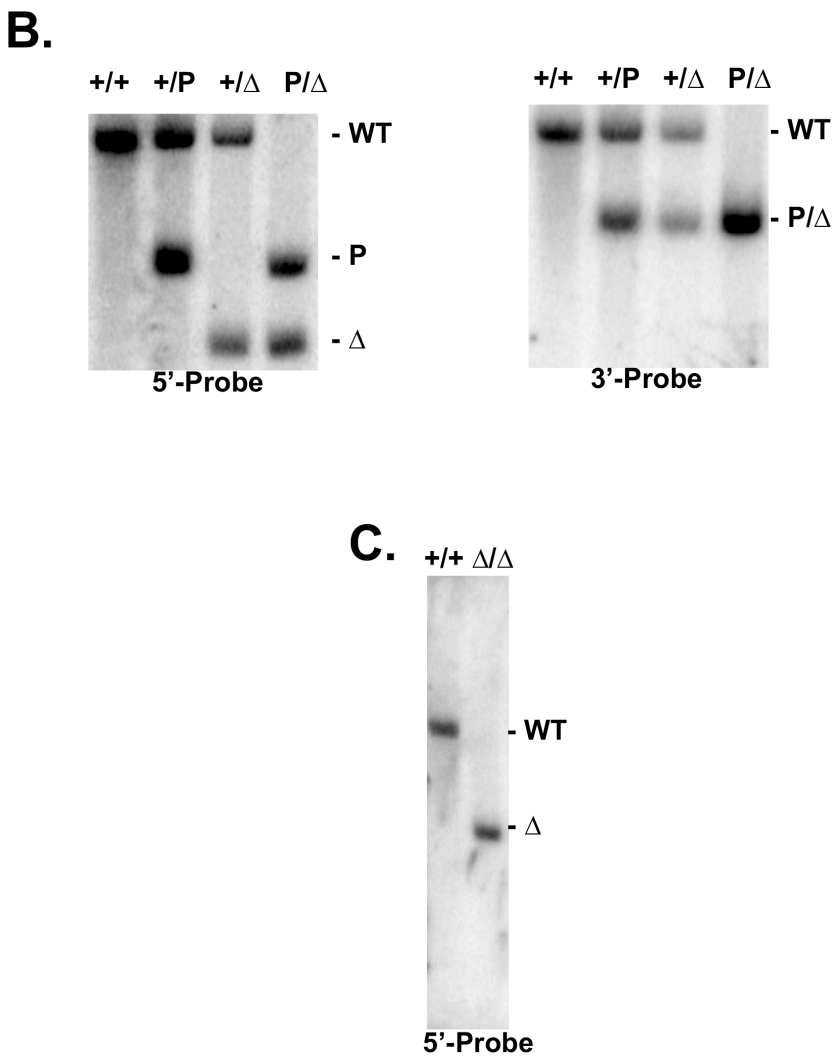


Appendix D: Chapter 4 Figures

Figure 17. Gene targeting of Ligase I.



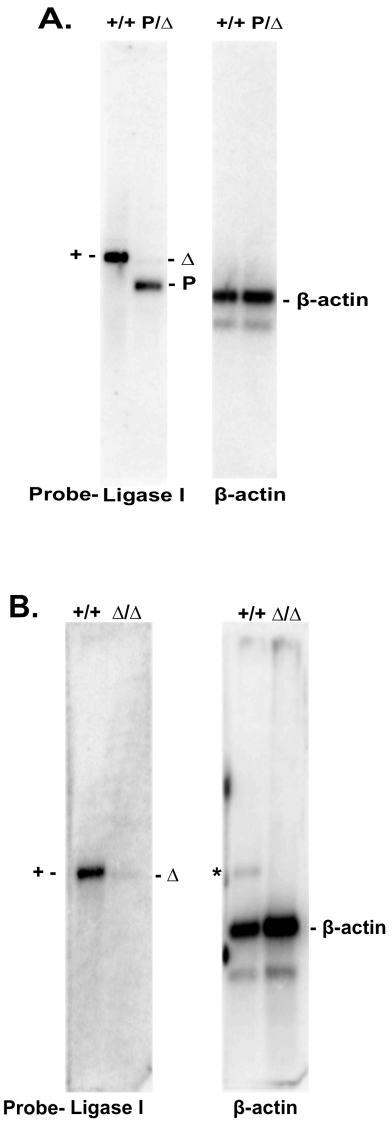
(Figure 17 cont'd.)



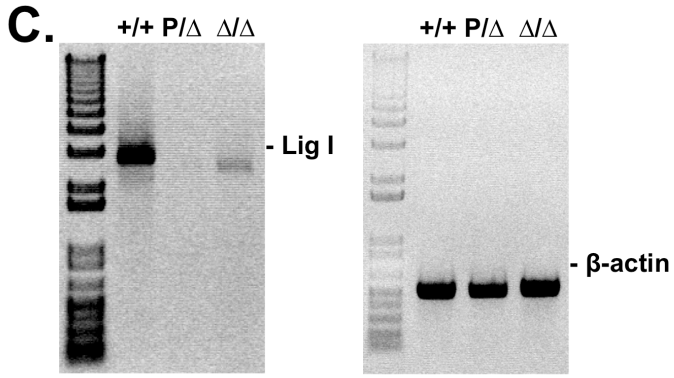
(Figure 17 cont'd.)

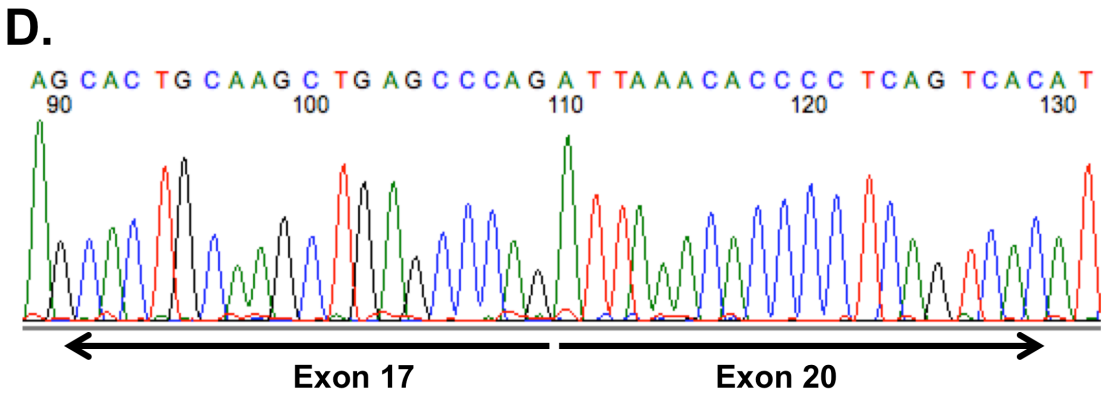
A. Genomic organization of the wildtype and targeted *LigaseI* allele. The map is drawn to scale. Exons are indicated by numbered boxes. Arrows indicate transcription orientations of expression cassettes of the puromycin-resistant gene (Puro) and diphtheria toxin A chain (DTA), respectively. B, BamHI restriction enzyme site. Probes used in Southern blot analysis are depicted at the top. A "+" sign indicates the wild-type allele. "P" and "Δ" indicate targeted alleles with or without the Puro cassette, respectively. Triangles indicate loxP sites. **B.** Southern blot analysis of BamHI-digested genomic DNA from wildtype (+/+) and targeted cells (+/P, +/Δ, P/Δ). Genotypes and probes are indicated. **C.** Southern blot analysis of BamHI-digested genomic DNA from wildtype (+/+) and Ligase I-null (Δ/Δ) cells. Genotypes and probe used are indicated.

Figure 18. Characterization of a Ligase I-null cell line.

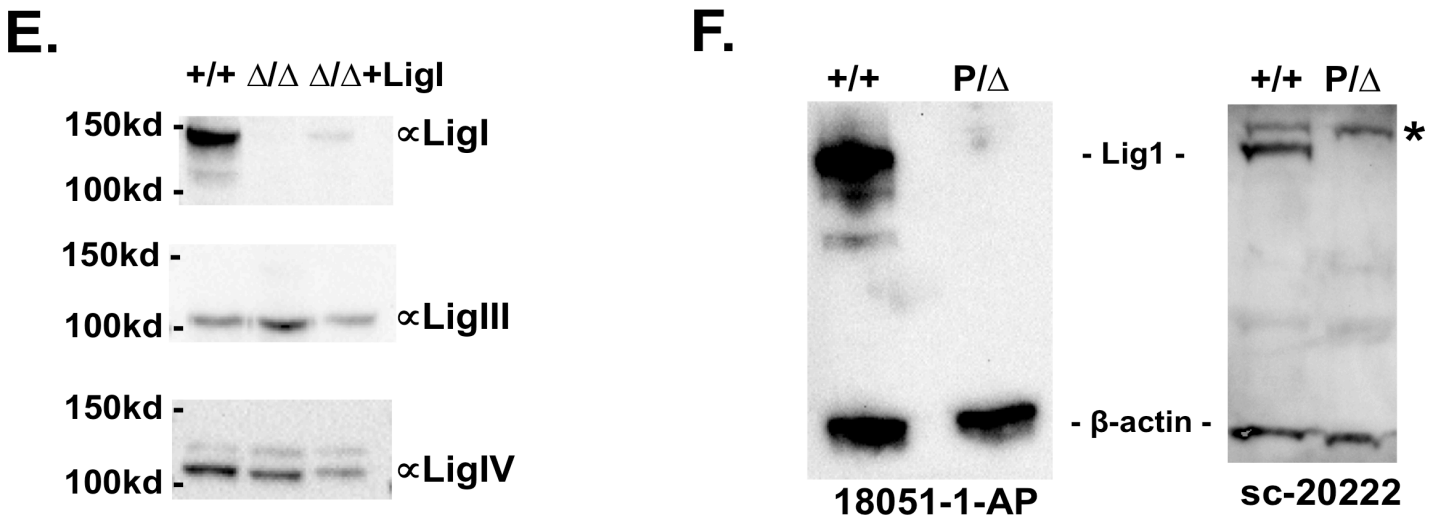


(Figure 18 cont'd.)





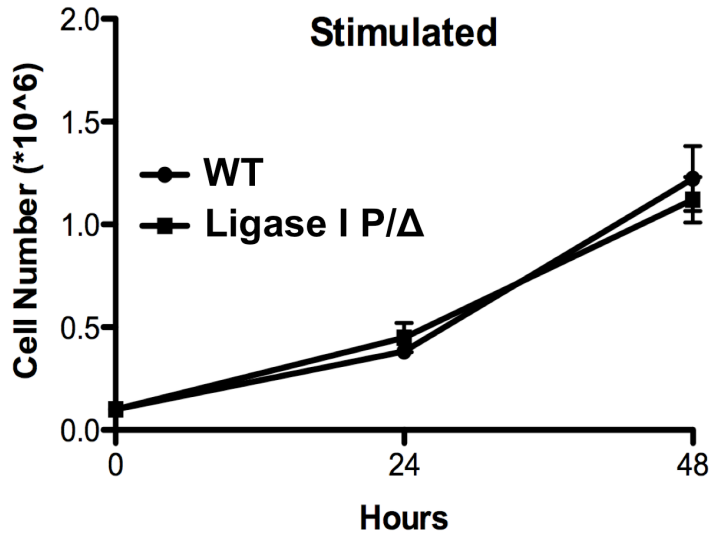
(Figure 18 cont'd.)



(Figure 18 cont'd.)

A. Northern blot analysis of polyadenylated RNA from wild type (+/+) and Ligase I P/Δ cells. The entire Ligase I coding region sequence was used as probe. The blot was stripped and stored for a month to allow radioisotope decay before re-probed with β-actin. **B.** Northern blot analysis of polyadenylated RNA from wild type (+/+) and Ligase I Δ/Δ cells. The blot was stripped and re-probed with β-actin. The asterisk indicates the remaining Ligase I signal from the previous probing. **C.** RT-PCR to detect Ligase I transcript in wild type (+/+), Ligase I P/Δ and Ligase I Δ/Δ cells. **D.** Sanger sequencing of the RT-PCR product derived from the Δ allele showing exon 17 joined to exon 20. **E.** Western blot analysis of the Ligase I protein level in wild type (+/+), Ligase I Δ/Δ and Ligase I Δ/Δ cells complemented with mouse Ligase I cDNA (Δ/Δ+LigI) with a rabbit polyclonal antibody raised against GST-Ligase I protein (Proteintech 18051-1-AP). The blot was stripped and reprobed with a Ligase 3 antibody (BD Biosciences, 611876) and Ligase 4 antibody [56], respectively, for loading control and protein size distinction. **F.** Western blot analysis of the Ligase I protein level wild type (+/+) and Ligase I P/Δ cells. 18051-1-AP (Proteintech) is a rabbit polyclonal antibody raised against GST-Lig1 protein. Sc-20222 (Santa Cruz) is a goat polyclonal antibody that recognizes a N-terminal region of Ligase I. The asterisk indicates a non-specific crossreactive band. **G.** Adenylation of Ligase I in the cell extract with α-^{32}P-ATP.

Figure 19. Proliferation of Ligase I-null cells. A. Live cell (trypan blue exclusion) counts of wild-type (WT) and Ligase I-null (P/Δ) cells over a span of 48 h in unstimulated (-CIT) or stimulated (+ CIT) cultures; CIT, anti-CD40, interleukin 4 and TGFβ1. Error bars indicate standard error of three independent experiments.



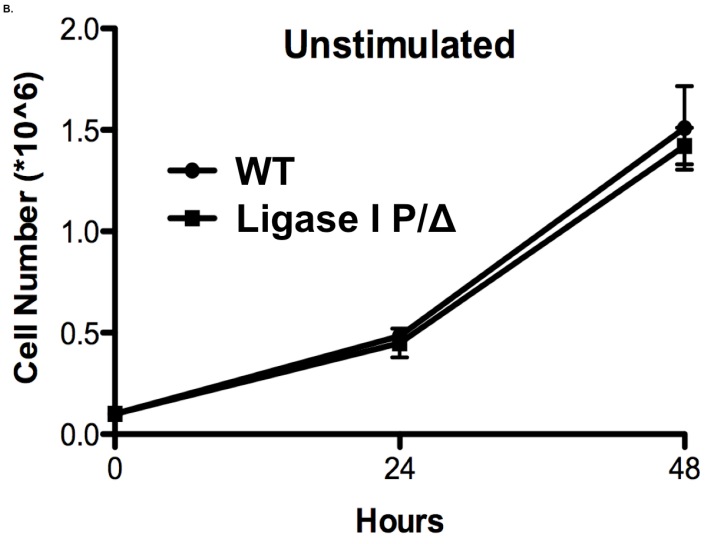


Figure 20. A metaphase of a Ligase I-null cell. All the abnormal chromosomes, two copies of the normal chromosome 14, and the sex chromosome are as marked. The der(1) is the only new abnormality (marked with solid arrowhead) compared with the parental cells, which has a chromosome count of 41.

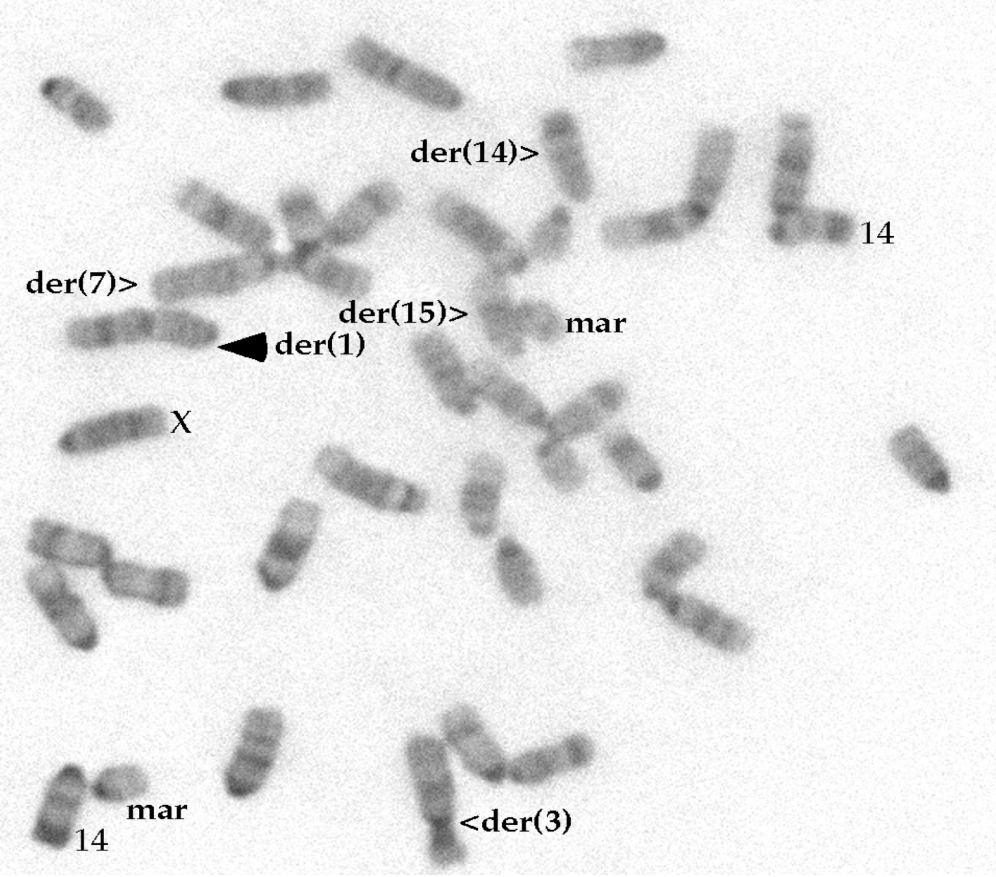
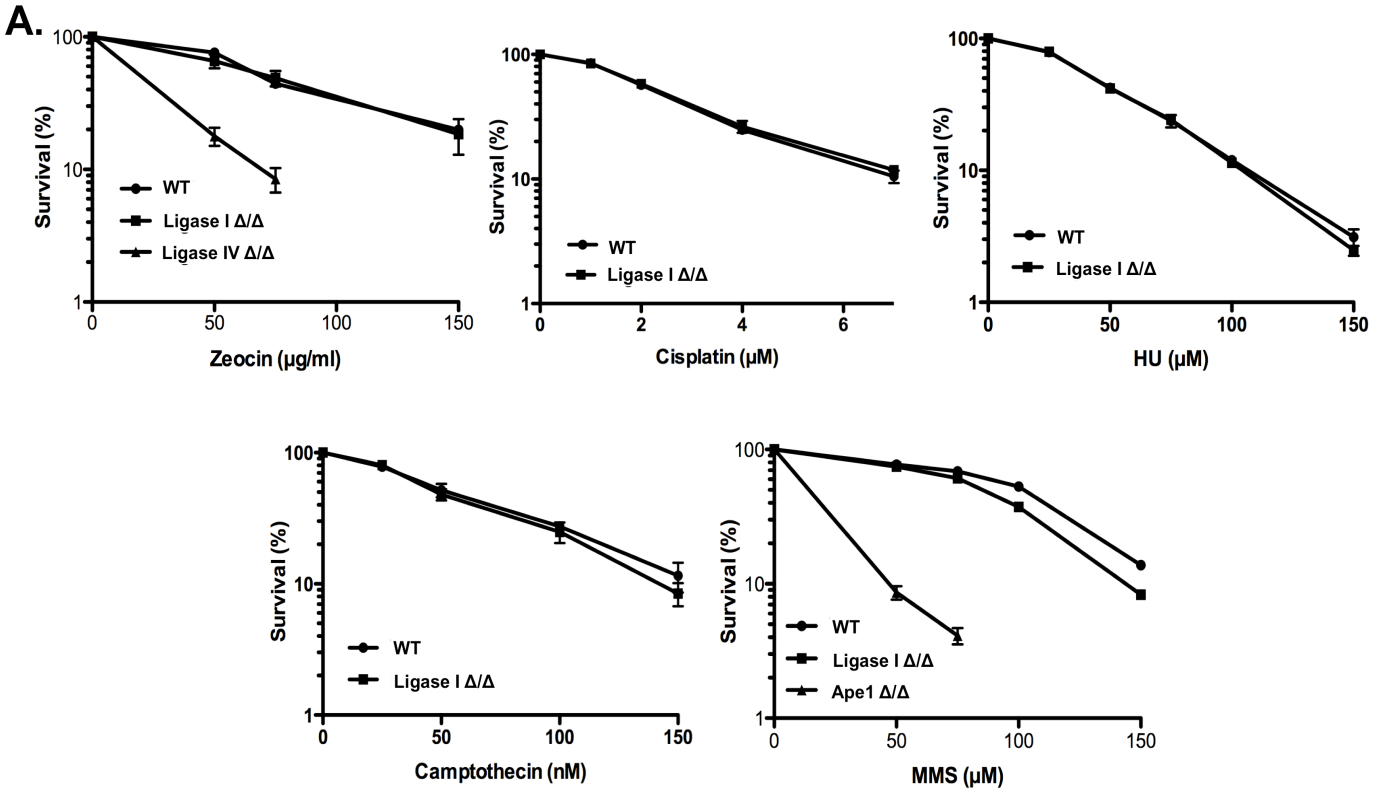


Figure 21. Drug sensitivity of Ligase I-null cells. A. Sensitivity of wild-type (WT) and Ligase I Δ/Δ cells to several DNA damaging agents. Error bars represent standard deviation of the mean of three independent experiments. **B.** Sensitivity of wild-type (WT) and Ligase I P/Δ cells to several DNA damaging agents. Error bars represent standard deviation of the mean of three independent experiments.



(Figure 21 cont'd.)

B.

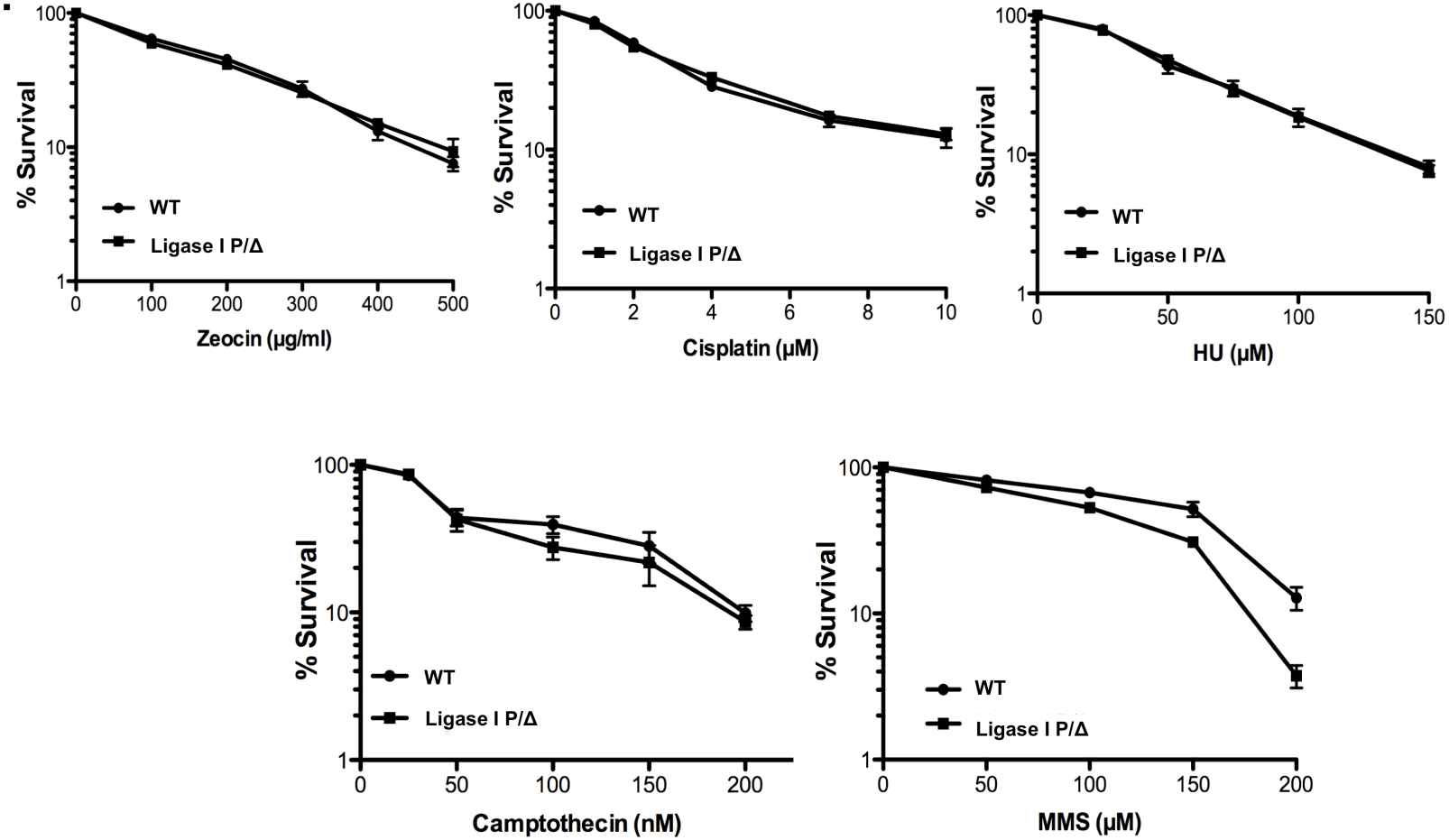


Figure 22. Class Switch Recombination in Ligase I-null cells. Representative FACS analysis of CSR by surface staining of IgA after 72 h of growth with (stimulated) or without (unstimulated) CIT. Numbers in the boxed areas indicate percentages.

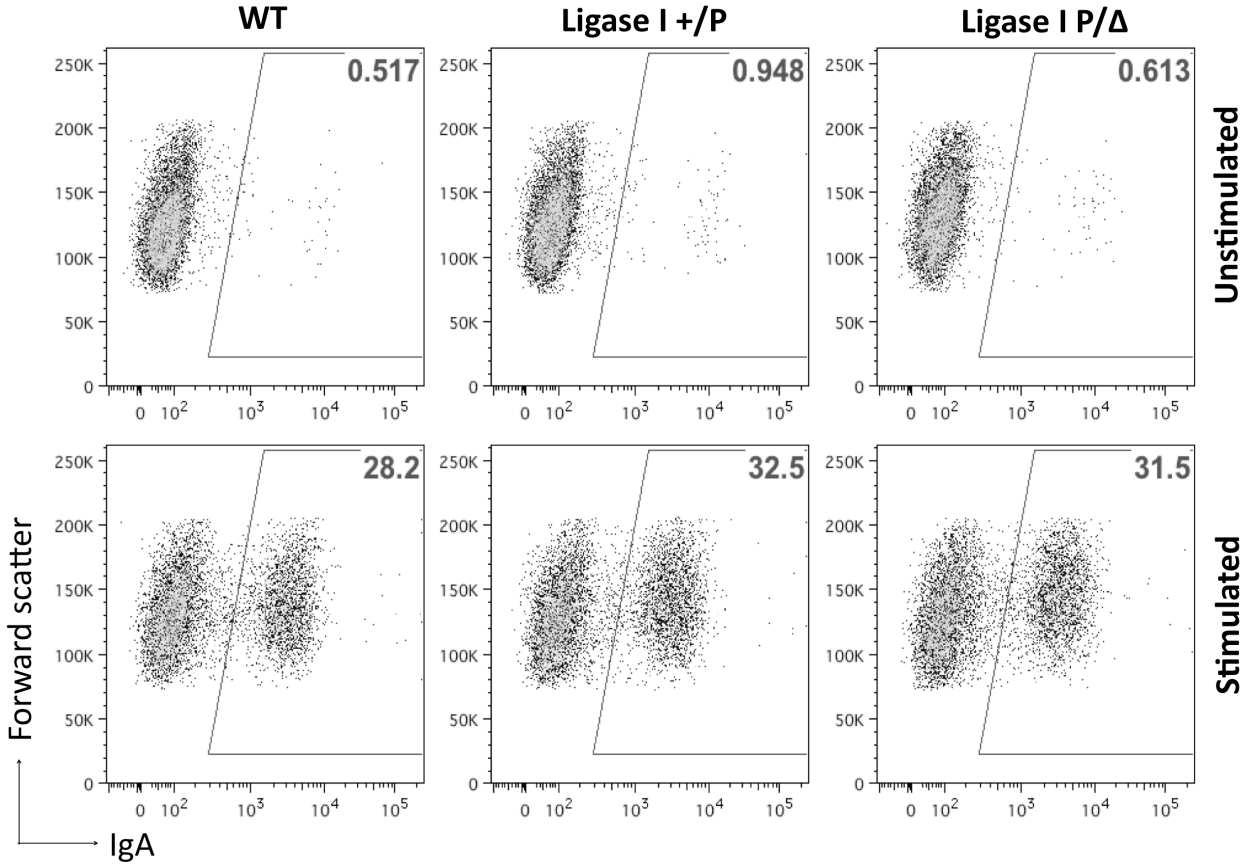


Figure 23. Construction and characterization of Ligase I-Ligase IV-null cells. A. Western blot analysis of the Ligase I and Ligase IV protein levels in WT (+/+) and Ligase I Δ/Δ-Ligase IV Δ/Δ cells. **B.** Live cell (trypan blue exclusion) counts of wild-type (WT) and Ligase I Δ/Δ Ligase IV Δ/Δ cells in unstimulated (-CIT) or stimulated (+ CIT) cultures; CIT, anti-CD40, interleukin 4 and TGFβ1. Error bars indicate standard error of three independent experiments.

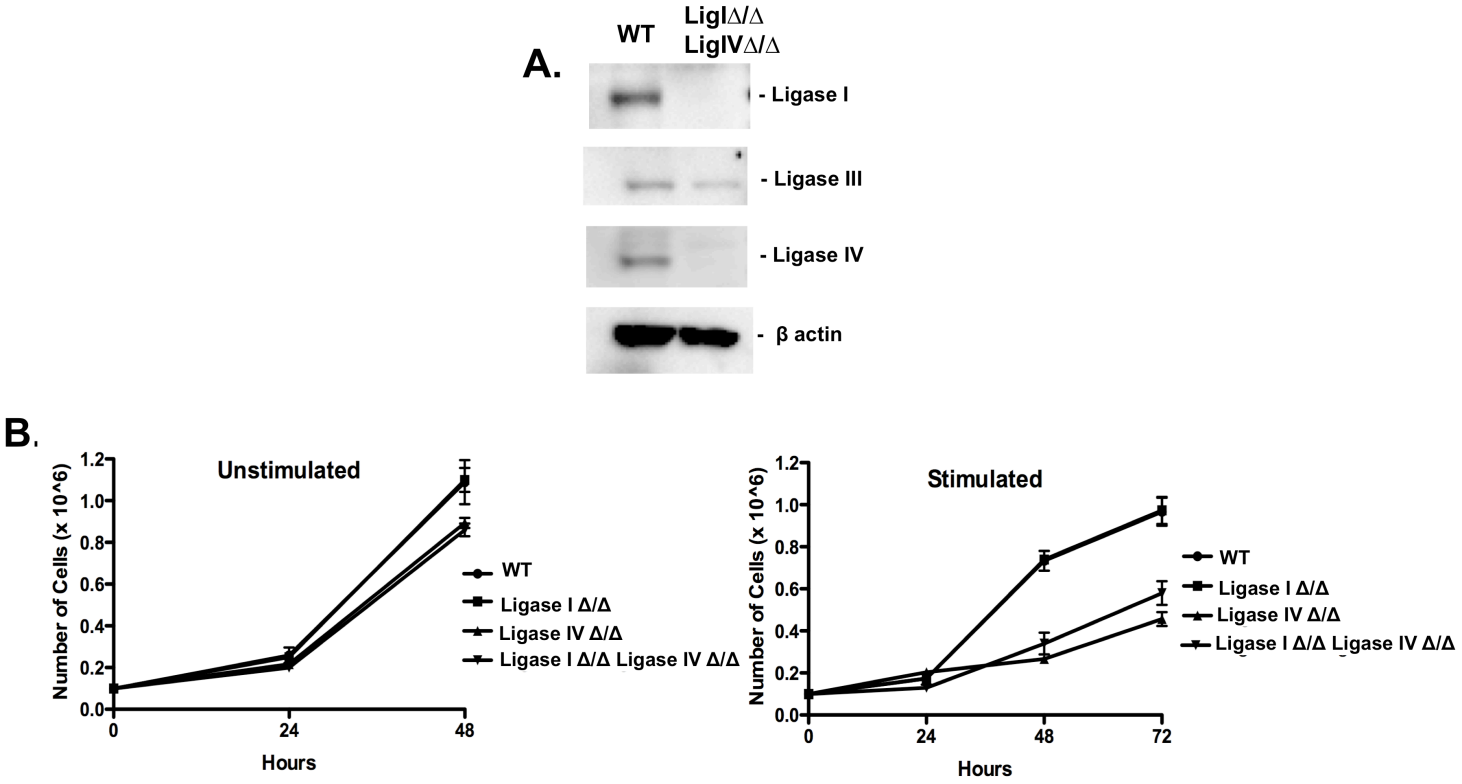


Figure 24. Drug sensitivity of Ligase I-Ligase IV-null cells. A. Sensitivity of wild-type (WT) and Ligase I Δ/Δ Ligase IV Δ/Δ cells to several DNA damaging agents. Error bars represent standard deviation of the mean of three independent experiments.

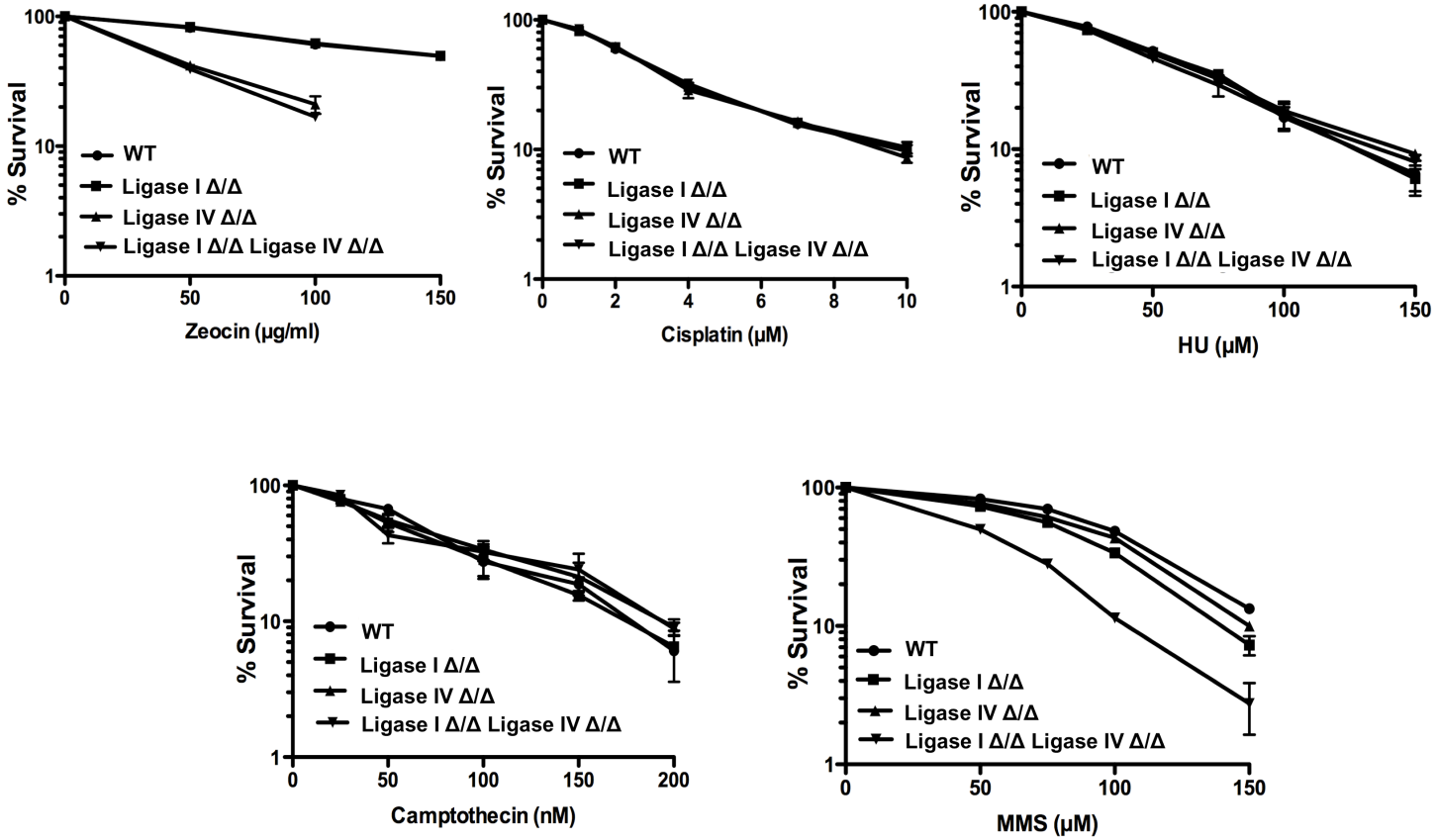
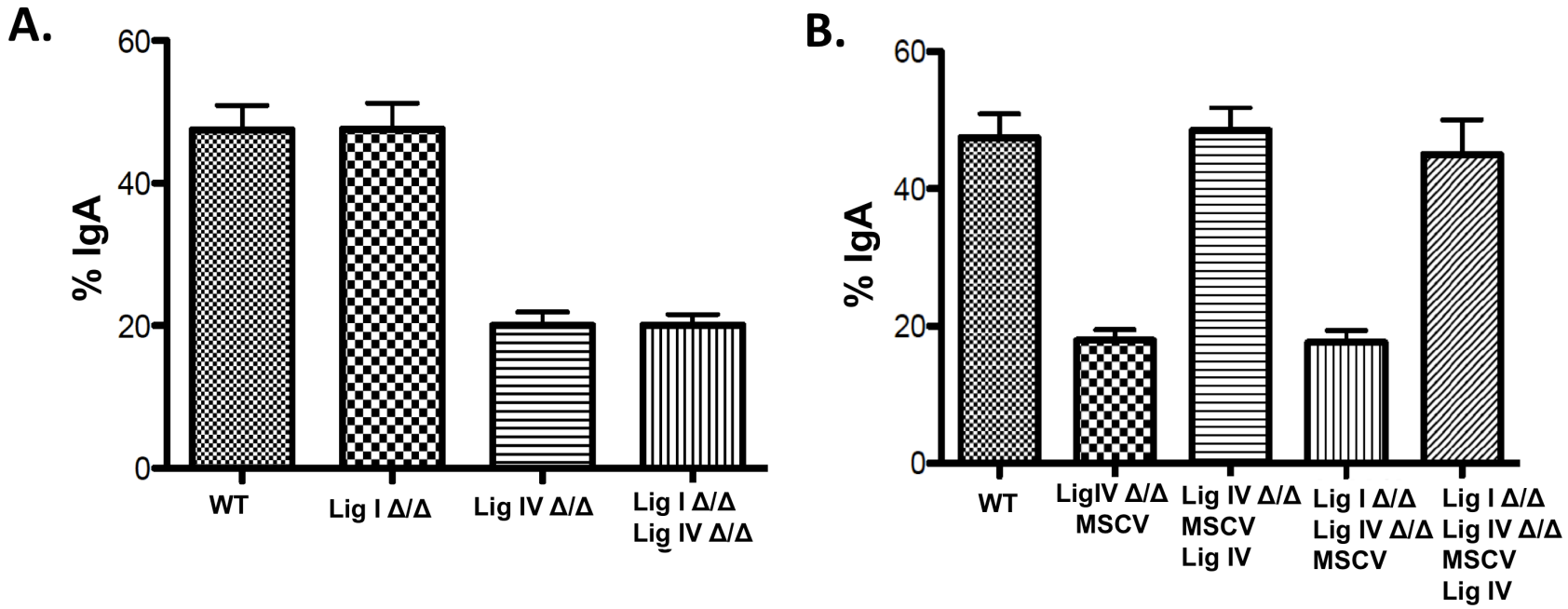


Figure 25. Class Switch Recombination in Ligase I-Ligase IV-null cells. A. CSR efficiency (determined as the percentage of IgA+ cells) in Ligase I Ligase IV-null cells. Error bars indicate standard error of three independent experiments. **B.** Genetic complementation of Ligase I Ligase IV-null cells. WT, Wildtype; MSCV, cells transfected with control retrovirus; MSCV-LigIV, cells transfected with mouse Ligase IV CDS.



BIBLIOGRAPHY

BIBLIOGRAPHY

1. Chaudhuri, J. and Alt, F.W.; (2004); Class-switch recombination: interplay of transcription, DNA deamination and DNA repair*; Nat Rev Immunol*; 4; 541-52.

2. Yu, K. and Lieber, M.R.; (2003); Nucleic acid structures and enzymes in the immunoglobulin class switch recombination mechanism*; DNA Repair (Amst)*; 2; 1163-74.

3. Gritzmacher, C.A.; (1989); Molecular aspects of heavy-chain class switching*; Crit Rev Immunol*; 9; 173-200.

4. Kataoka, T., Miyata, T., and Honjo, T.; (1981); Repetitive sequences in class-switch recombination regions of immunoglobulin heavy chain genes*; Cell*; 23; 357-68.

5. Huang, F.T., Yu, K., Balter, B.B., Selsing, E., Oruc, Z., Khamlichi, A.A., Hsieh, C.L., and Lieber, M.R.; (2007); Sequence dependence of chromosomal R-loops at the immunoglobulin heavy-chain Smu class switch region*; Mol Cell Biol*; 27; 5921-32.

6. Hackney, J.A., Misaghi, S., Senger, K., Garris, C., Sun, Y., Lorenzo, M.N., and Zarrin, A.A.; (2009); DNA targets of AID evolutionary link between antibody somatic hypermutation and class switch recombination*; Adv Immunol*; 101; 163-89.

7. Stavnezer, J. and Amemiya, C.T.; (2004); Evolution of isotype switching*; Semin Immunol*; 16; 257-75.

8. Zarrin, A.A., Alt, F.W., Chaudhuri, J., Stokes, N., Kaushal, D., Du Pasquier, L., and Tian, M.; (2004); An evolutionarily conserved target motif for immunoglobulin class-switch recombination*; Nat Immunol*; 5; 1275-81.

9. Gellert, M.; (2002); V(D)J recombination: RAG proteins, repair factors, and regulation*; Annu Rev Biochem*; 71; 101-32.

10. Dunnick, W., Hertz, G.Z., Scappino, L., and Gritzmacher, C.; (1993); DNA sequences at immunoglobulin switch region recombination sites*; Nucleic Acids Res*; 21; 365-72.

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

12. Khamlichi, A.A., Glaudet, F., Oruc, Z., Denis, V., Le Bert, M., and Cogne, M.; (2004); Immunoglobulin class-switch recombination in mice devoid of any S mu tandem repeat*; Blood*; 103; 3828-36.

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

14. Arakawa, H., Hauschild, J., and Buerstedde, J.M.; (2002); Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion*; Science*; 295; 1301-6.

15. Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A., Tezcan, I., Ersoy, F., Kayserili, H., Ugazio, A.G., Brousse, N., Muramatsu, M., Notarangelo, L.D., Kinoshita, K., Honjo, T., Fischer, A., and Durandy, A.; (2000); Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2)*; Cell*; 102; 565-75.

16. Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O., and Honjo, T.; (1999); Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells*; J Biol Chem*; 274; 18470-6.

17. Pan−Hammarstrom, Q., Zhao, Y., and Hammarstrom, L.; (2007); Class Switch Recombination: A Comparison Between Mouse and Human*; Adv Immunol*; 93.

18. Petersen-Mahrt, S.K., Harris, R.S., and Neuberger, M.S.; (2002); AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification*; Nature*; 418; 99-103.

19. Shapiro, G.S. and Wysocki, L.J.; (2002); DNA target motifs of somatic mutagenesis in antibody genes*; Crit Rev Immunol*; 22; 183-200.

20. Yu, K., Huang, F.T., and Lieber, M.R.; (2004); DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine*; J Biol Chem*; 279; 6496-500.

21. Han, L., Masani, S., and Yu, K.; (2011); Overlapping activation-induced cytidine deaminase hotspot motifs in Ig class-switch recombination*; PNAS*; 108; 11584-9.

22. Kracker, S. and Durandy, A.; (2011); Insights into the B cell specific process of immunoglobulin class switch recombination*; Immun Letters*; 138; 97-103.

23. Stavnezer, J. and Schrader, C.E.; (2014); IgH Chain Class Switch Recombination: Mechanism and Regulation*; J Immunol*; 193; 5370-5378.

24. Xu, Z., Zan, H., Pone, E.J., Mai, T., and Casali, P.; (2012); Immunoglobulin class-switch DNA recombination: induction, targeting and beyond*; Nat Rev Immun*; 12; 517-531.

25. Stavnezer, J.; (1996); Immunoglobulin Class Switching*; Curr Opin Immunol*; 8; 19-205.

26. Stavnezer, J., Guikema, J.E., and Schrader, C.E.; (2008); Mechanism and Regulation of Class Switch Recombination*; Annu Rev Immunol*; 26; 261-92.

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

28. Lorenz, M., Jung, S., and Radbruch, A.; (1995); Switch transcripts in immunoglobulin class switching*; Science*; 267; 1825-8.

29. Hein, K., Lorenz, M.G., Siebenkotten, G., Petry, K., Christine, R., and Radbruch, A.; (1998); Processing of switch transcripts is required for targeting of antibody class switch recombination*; J Exp Med*; 188; 2369-74.

30. Manis, J.P., VanDer Stoep, N., Tian, M., Ferrini, R., Davidson, L., Bottaro, A., and Alt, F.W.; (1998); Class switching in B cells lacking 3' immunoglobulin heavy chain enhancers*; J Exp Med*; 188; 1421-1431.

31. Dunnick, W.E., Shi, J., Graves, K.A., and Collins, J.T.; (2005); The 3′ end of the heavy chain constant region locus enhances germline transcription and switch recombination of the four γ genes*; JEM*; 201; 1459-1466.

32. Cogne, M., Lansford, R., Bottaro, A., Zhang, J., Gorman, J., Young, F., Cheng, H.L., and Alt, F.W.; (1994); A class switch control region at the 3' end of the immunoglobulin heavy chain locus*; Cell*; 77; 737-47.

33. Vincent-Fabert, C., Fiancette, R., Pinaud, E., Truffinet, V., Cogne, N., Cogne, M., and Denizot, Y.; (2010); 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*; 116; 1895-1898.

34. Yu, K., Chedin, F., Hsieh, C.L., Wilson, T.E., and Lieber, M.R.; (2003); R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells*; Nat Immunol*; 4; 442-51.

35. Roy, D., Yu, K., and Lieber, M.R.; (2008); Mechanism of R-loop formation at immunoglobulin class switch sequences*; Mol Cell Biol*; 28; 50-60.

36. Daniels, G.A. and Lieber, M.R.; (1995); RNA:DNA complex formation upon transcription of immunoglobulin switch regions: implications for the mechanism and regulation of class switch recombination*; Nucleic Acids Res*; 23; 5006-11.

37. Pavri, R., Gazumyan, A., Jankovic, M., Di Virgilio, M., Klein, I., Ansarah-Sobrinho, C., Resch, W., Yamane, A., San-Martin, B.R., Barreto, V., Nieland, T.J., Root, D.E., Casellas, R., and Nussenzweig, M.C.; (2010); Activation-Induced Cytidine Deaminase Targets DNA at Sites of RNA Polymerase II Stalling by Interaction with Spt5*; Cell*; 143; 122-133.

38. Daniel, J.A., Santos, M.A., Wang, Z., Schwab, K.R., Jankivic, M., Fisluf, D., Chen, H.T., Gazumyan, A., Yamane, A., Cho, Y.W., Sun, H., W., Ge, K., Peng, W., Nussenzweig, M.C., Casellas, R., Dressler, G.R., Zhao, K., and Nussenzweig, A.; (2010); PTIP promotes chromatin changes critical for immunoglobulin class switch recombination*; Science*; 329; 917-923.

39. Basu, U., Meng, F., Keim, C., Grinstein, V., Pefanis, E., Eccleston, J., Zhang, T., Myers, D., Wasserman, C.R., Wesemann, D.R., Januszyk, K., Gregory, R.I., Deng, H., Lima, C., and Alt, F.W.; (2011); The RNA Exosome Targets the AID Cytidine Deaminase to Both Strands of Transcribed Duplex DNA Substrates*; Cell*; 144; 353-363.

40. Stanlie, A., Aida, M., Muramatsu, M., Honjo, T., and Begum, N.A.; (2010); Histone3 lysine4 trimethylation regulated by the facilitates chromatin transcription complex is critical for DNA cleavage in class switch recombination*; PNAS*; 107; 22190-22195.

41. Rada, C., Williams, G.T., Nilsen, H., Barnes, D.E., Lindahl, T., and Neuberger, M.S.; (2002); Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice*; Curr Biol*; 12; 1748-55.

42. Imai, K., Slupphaug, G., Lee, W.I., Revy, P., Nonoyama, S., Catalan, N., Yel, L., Forveille, M., Kavli, B., Krokan, H.E., Ochs, H.D., Fischer, A., and Durandy, A.; (2003); Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination*; Nat Immunol*.

43. Rada, C., Di Noia, J.M., and Neuberger, M.S.; (2004); Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation*; Mol Cell*; 16; 163-71.

44. Dingler, F.A., Kemmerich, K., Neuberger, M.S., and Rada, C.; (2014); SMUG1 glycosylase promotes efficient Ig class switching and impacts on A:T substitutions during somatic mutation*; Eur J Immunol*; 44; 1925-1935.

45. Masani, S., Han, L., and Yu, K.; (2013); Apurinic/Apyrimidinic Endonuclease 1 Is the Essential Nuclease during Immunoglobulin Class Switch Recombination*; Mol Cell Biol*; 33; 1468-73.

46. Ludwig, D.L., MacInnes, M.A., Takiguchi, Y., Purtymun, P.E., Henrie, M., Flannery, M., Meneses, J., Pedersen, R.A., and Chen, D.J.; (1998); A murine AP-endonuclease gene-targeted deficiency with post-implantation embryonic progression and ionizing radiation sensitivity*; Mutat Res*; 409; 17-29.

47. Xanthoudakis, S., Smeyne, R.J., Wallace, J.D., and Curran, T.; (1996); The redox/DNA repair protein, Ref-1, is essential for early embryonic development in mice*; Proc Natl Acad Sci U S A*; 93; 8919-23.

48. Fung, H. and Demple, B.; (2005); A vital role for Ape1/Ref1 protein in repairing spontaneous DNA damage in human cells*; Mol Cell*; 17; 463-70.

49. Guikema, J.E., Linehan, E.K., Tsuchimoto, D., Nakabeppu, Y., Strauss, P.R., Stavnezer, J., and Schrader, C.E.; (2007); APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination*; J Exp Med*; 204; 3017-26.

50. Guikema, J.E., Stavnezer, J., and Schrader, C.E.; (2010); The role of Apex2 in class-switch recombination of immunoglobulin genes*; Int Imunol*; 22; 213.

51. Guikema, J.E., Stavnezer, J., and Schrader, C.E.; (2010); The role of Apex2 in class-switch recombination of immunoglobulin genes.*; Int Immunol*; 22; author reply 213-213; author reply 214.

52. Sabouri, Z., Okazaki, I.M., Shinkura, R., Begum, N., Nagaoka, H., Tsuchimoto, D., Nakabeppu, Y., and Honjo, T.; (2009); Apex2 is required for efficient somatic hypermutation but not for class switch recombination of immunoglobulin genes*; Int Immunol*; 8; 947-955.

53. Ehrenstein, M.R. and Neuberger, M.S.; (1999); Deficiency in Msh2 affects the efficiency and local sequence specificity of immunoglobulin class-switch recombination: parallels with somatic hypermutation*; Embo J*; 18; 3484-90.

54. Ward, I.M., Reina-San-Martin, B., Olaru, A., Minn, K., Tamada, K., Lau, J.S., Cascalho, M., Chen, L., Nussenzweig, A., Livak, F., Nussenzweig, M.C., and Chen, J.; (2004); 53BP1 is required for class switch recombination*; J Cell Biol*; 165; 459-64.

55. Reina-San-Martin, B., Difilippantonio, S., Hanitsch, L., Masilamani, R.F., Nussenzweig, A., and Nussenzweig, M.C.; (2003); H2AX is required for recombination between immunoglobulin switch regions but not for intra-switch region recombination or somatic hypermutation*; J Exp Med*; 197; 1767-1778.

56. Han, L. and Yu, K.; (2008); Altered kinetics of NHEJ and CSR in DNA ligase IV-deficient B cells*; J Exp Med*; (DOI: 10.1084/jem.20081623).

57. Boboila, C., Yan, C., Wesemann, D.R., Jankovic, M., Wanw, J.H., Manis, J., Nussenzweig, A., Nussenzweig, M., and Alt, F.W.; (2010); Alternative endjoining catalyzes class switch recombination in the absence of both Ku70 and DNA ligase 4.*; JEM*.

58. Boboila, C., Oksenych, V., Gostissa, M., Wang, J.H., Zha, S., Zhang, Y., Chai, H., Lee, C.S., Jankovic, M., Saez, L.M., Nussenzweig, M., McKinnon, P.J., Alt, F.W., and Schwer, B.; (2012); Robust chromosomal DNA repair via alternative end-joining in the absence of X-ray repair crosscomplementing protein 1 (XRCC1)*; PNAS*; 109; 2473-8.

59. Yan, C.T., Boboila, C., Souza, E.K., Franco, S., Hickernell, T.R., Murphy, M., Gumaste, S., Geyer, M., Zarrin, A.A., Manis, J.P., Rajewsky, K., and Alt, F.W.; (2007); IgH class switching and translocations use a robust non-classical end-joining pathway*; Nature*; 449; 478-82.

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

61. Lenz, G., Nagel, I., Siebert, R., Roschke, A.V., Sanger, W., Wright, G.W., Dave, S.S., Tan, B., Zhao, H., Rosenwald, A., Muller-Hermelink, H.K., Gascoyne,

R.D., Campo, E., Jaffe, E.S., Smeland, E.B., Fisher, R.I., Kuehl, W.M., Chan, W.C., and Staudt, L.M.; (2007); Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell–like diffuse large B cell lymphoma*; JEM*; 204; March 2007.

62. Jesus, A.A., Duarte, A.J., and Oliveira, J.B.; (2008); Autoimmunity in Hyper-IgM Syndrome*; J Clin Immunol*.

63. Melegari, A., Mascia, M.T., Sandri, G., and Carbonieri, A.; (2007); Immunodeficiency and autoimmune phenomena in female hyper-IgM syndrome*; Ann N Y Acad Sci*; 1109; 106-8.

64. Hecht, J.L. and Aster, J.C.; (2000); Molecular biology of Burkitt's lymphoma*; J Clin Oncol*; 18; 3707-21.

65. Ali, K., Mahjabeen, I., Sabir, M., Baig, R.M., Zafeer, M., Faheem, M., and Kayani, M.A.; (2014); Germline variations of apurinic/apyrimidinic endonuclease 1 (APEX1) detected in female breast cancer patients*; Asian Pac J Cancer Prev*; 15; 7589-7595.

66. Liu, Y., Scheurer, M.E., El-Zein, R., Cao, Y., Do, K.A., Gilbert, M., Aldape, K.D., Wei, Q., Etzel, C., and Bondy, M.L.; (2009); Association and interactions between DNA repair gene polymorphisms and adult glioma*; Cancer Epidemiol Biomarkers Prev*; 18; 204-14.

67. Sallmyr, A., Thomkinson, A.E., and Rassool, F.V.; (2008); Up-regulation of WRN and DNA ligase IIIalpha in chronic myeloid leukemia: consequences for the repair of DNA double-strand breaks*; Blood*; 112; 1413-23.

68. Lee, Y.C., Morgenstern, H., Greenland, S., Tashkin, D.P., Papp, J., Sinsheimer, J., Cao, W., Hashibe, M., You, N.C., Mao, J., Cozen, W., Mack, T.M., and Zhang, Z.F.; (2008); A case-control study of the association of the polymorphisms and haplotypes of DNA ligase I with lung and upper-aerodigestive-tract cancers*; Int J Cancer*; 122; 1630-8.

69. Shanmugam, A., Shi, M.J., Yauch, L., Stavnezer, J., and Kenter, A.L.; (2000); Evidence for class-specific factors in immunoglobulin isotype switching*; J Exp Med*; 191; 1365-80.

70. Stavnezer, J., Bradley, S.P., Rousseau, N., Pearson, T., Shanmugam, A., Waite, D.J., Rogers, P.R., and Kenter, A.L.; (1999); Switch recombination in a transfected plasmid occurs preferentially in a B cell line that undergoes switch recombination of its chromosomal Ig heavy chain genes*; J Immunol*; 163; 2028-40.

71. Kinoshita, K., Tashiro, J., Tomita, S., Lee, C.G., and Honjo, T.; (1998); Target specificity of immunoglobulin class switch recombination is not determined by nucleotide sequences of S regions*; Immunity*; 9; 849-58.

72. Zarrin, A.A., Goff, P.H., Senger, K., and Alt, F.W.; (2008); Sgamma3 switch sequences function in place of endogenous Sgamma1 to mediate antibody class switching*; J Exp Med*; 205; 1567-72.

73. Min, I.M., Rothlein, L.R., Schrader, C.E., Stavnezer, J., and Selsing, E.; (2005); Shifts in targeting of class switch recombination sites in mice that lack mu switch region tandem repeats or Msh2*; J Exp Med*; 201; 1885-90.

74. Szurek, P., Petrini, J., and Dunnick, W.; (1985); Complete nucleotide sequence of the murine gamma 3 switch region and analysis of switch recombination sites in two gamma 3-expressing hybridomas*; J Immunol*; 135; 620-6.

75. Zhang, Z.Z., Pannunzio, N.R., Han, L., Hsieh, C., Yu, K., and Lieber, M.R.; (2014); The Strength of an Ig Switch Region Is Determined by Its Ability to Drive R Loop Formation and Its Number of WGCW Sites*; Cell Reports*; 8; 557-569.

76. Hadi, M.Z. and Wilson, D.M.r.; (2000); Second human protein with homology to the Escherichia coli abasic endonuclease exonuclease III*; Environ Mol Mutagen*; 36; 312-324.

77. Kaur, G., Cholia, R.P., MAntha, A.K., and Kumar, R.; (2014); DNA repair and redox activities and inhibitors of apurinic/apyrimidinic endonuclease 1/redox effector factor 1 (APE1/Ref-1): a comparative analysis and their scope and limitations toward anticancer drug development*; J Med Chem*; 57; 10241-10256.

78. Nash, R. and Lindahl, T.; (1996); DNA Ligases*; DNA Replication in Eukaryotic Cells*; 575-586.

79. Frank, K.M., Sharpless, N.E., Gao, Y., Sekiguchi, J.M., D.O., F., Zhu, C., Manis, J.P., Horner, J., DePinho, R.A., and Alt, F.W.; (2000); DNA ligase IV deficiency in mice leads to defective neurogenesis and embryonic lethality via the p53 pathway*; Moll Cell*; 5; 993-1002.

80. Bentley, D., Selfridge, J., Millar, J.K., Samuel, K., Hole, N., Ansell, J.D., and Melton, D.W.; (1996); DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability*; Nature Genet*; 13; 489-491.

81. Puebla-Osorio, N., Lacey, D.B., Alt, F.W., and Zhu, C.; (2006); Early embryonic lethality due to targeted inactivation of DNA ligase III*; Mol Cell Biol*; 26; 3935-3941.

82. Arakawa, H., Bednar, T., Wang, M., Paul, K., Mladenov, E., Bencsik-Theilen, A.A., and Iliakis, G.; (2012); Functional redundancy between DNA ligases I and III in DNA replication in vertebrate cells*.; Nucleic Acids REs*; 40; 2599-2610.

83. Bentley, D.J., Harrison, C., Ketchen, A.M., Redhead, N.J., Samuel, K., Waterfall, M., Ansell, J.D., and Melton, D.W.; (2002); DNA ligase I null mouse cells show normal DNA repair activity but altered DNA replication and reduced genome stability*; J Cell Sci*; 115; 1551-1561.

84. Petrini, J.H., Xiao, Y., and Weaver, D.T.; (1995); DNA ligase I mediates essential functions in mammalian cells*; Mol Cell Biol*; 15; 4303-4308.

85. Simsek, D., Furda, A., Gao, Y., Artus, J., Brunet, E., Hadjantonakis, A.K., Van Houten, B., Shuman, S., McKinnon, P.J., and Jasin, M.; (2011); Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair*.; Nature*; 471; 245-248.

86. Oh, S., Harvey, A., Zimbric, J., Wang, Y., Nguyen, T., Jackson, P.J., and Hendrickson, E.A.; (2014); DNA ligase III and DNA ligase IV carry out genetically distinct forms of end joining in human somatic cells*; DNA Repair*; 97-110.

87. Han, L., Mao, W., and Yu, K.; (2012); X-ray repair cross-complementing protein 1 (XRCC1) deficiency enhances class switch recombination and is permissive for alternative end joining*; PNAS*; 109; 4604-8.

88. Iliakis, G. and Dueva, R.; (2013); Alternative pathways of non-homologous end joining (NHEJ) in genomic instability and cancer*; Transl Cancer Res*; 2; 163-77.