1-1-2015

The Loss Of Genomic Uracil Homeostasis And Aid-Dependent Accumulation Of Dna Damage In B Cell Lymphomas

Sophia Shalhout
Wayne State University,

THE LOSS OF GENOMIC URACIL HOMEOSTASIS AND AID-DEPENDENT ACCUMULATION OF DNA DAMAGE IN B CELL LYMPHOMAS

by

SOPHIA SHALHOUT

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2015

MAJOR: CHEMISTRY (Biochemistry)

Approved by:

_____

Advisor Date

_____

_____

_____

_____

DEDICATION

Dedicated to my husband, Rafat, for his endless support, sacrifice, insight, and encouragement.

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor Dr. Ashok Bhagwat. His passionate pursuit of science has made him a personal inspiration. I thank him for offering me the best guidance throughout my graduate studies and for always welcoming and encouraging scientific inquiry. He has always been available to critically discuss my work with unmatched scientific enthusiasm. I thank him for his time, commitment, support, and direction. He is my mentor in both science and in life and I cannot express how thankful I am to him for the profound way his influence has shaped my life. He has pushed me to be my best and I have been incredibly lucky to pursue my Ph.D. under his direction. I strive to eventually have my own lab and hope to mirror his genuine commitment to both science and the development of young scientists.

I would also like to thank my departmental committee members, Dr. Christine Chow and Dr. Andrés Cisneros, as well as my former committee member Dr. Peter Andreana for their help, dedication, and valuable critique. I have been extremely fortunate to work directly with my non-departmental committee member, Dr. Thomas Holland. His advice and valuable training in viral work, immunofluorescence/imaging, and flow cytometry have been greatly appreciated. I thank him for always being available to discuss any technical challenges. I have yet to see him turn down helping anyone in the lab. His insight over the years was indispensable and I am truly thankful to him for everything.

I would like to thank my entire loving family for countless reasons. I must specifically thank my sister Sylvia for offering unconditional support and my

brother Shareef for being a source of calm, especially in every instance where technology has failed me. I also thank my mother and father who have been an amazing source of love and support. They are the most self-less parents having postponed their dreams in favor of their childrens'. I would like to thank my husband Rafat for always being there for me, pushing me to do my best, and for his continual support in my future career goals. I thank him for all his sacrifices as our young marriage has only seen me as a graduate student so far. I would also like to sincerely thank my father and mother in law for all their support throughout the years.

My studies have been extremely interdisciplinary and have thus required collaborative efforts and assistance. I would personally like to thank Dr. Alberto Martin and his wonderful students and postdocs for hosting me at the University of Toronto and training me in FACs and mouse tissue harvesting. They are an extremely welcoming and bright group and I would like to thank Alberto for his help, patience, and guidance with my work. I would like to thank Dr. James Tucker's group in the Biology Department for teaching me chromosomal spread work and cytogenetics/imaging techniques. They allowed me to borrow equipment, use microscopes as I please, and share any reagents. I have made some of my best friends from this group and thank them dearly for treating me, always, like a fellow Tucker lab member. I would also absolutely like to thank Dr. Angela Sosin and Dr. Ayad Al-Katib for their help with all of my patient tumor work. I have been fortunate to have Dr. Al-Katib's lab also treat me like a member of their group allowing me to use any of their computers, hood space,

cell lines, equipment, kits, and desk space at the medical school. Angela was an extremely welcoming collaborator and I enjoyed our many spontaneous and early morning trips to pick up patient/blood samples, and I also consider her a dear friend. I must thank Dr. Kenneth Honn's lab, specifically Dr. Yinlong Cai, from the Department of Pathology. Yinlong's multi-disciplinary expertise in tissue culture, qRT-PCR, and cancer biology coupled with his helpful personality have made him the most amazing person to troubleshoot technical issues with. I need to sincerely thank him for all his help especially with any and all equipment failures. Every conversation with him, even in passing, has been insightful and helpful. I would also like to thank Dr. Kang Chen and his group from the Perinatology Research Branch. He has an unmatched expertise in immunology and I thank him for all his advice and help especially with my tonsil work. Finally, I must thank all my neighboring labs. I have had mutual and helpful interaction with so many students and postdocs from the Chow, Feig, Rueda, Ahn, and Romano group.

I must thank all the current and former members of the Bhagwat group. I have had the privilege of working with very smart, helpful, and insightful students and postdocs. I must personally thank Dr. Priyanga Wijesinghe because for some time, he and I were the only two graduate students in the lab and we worked efficiently well to handle all research and lab duties together. I thank my current lab members including Shanqaio Wei, Vimukthi Senevirathne, Sachini Siriwardena and Dr. Anjali Patwardhan for their unyielding help and friendship. No one among these current members has ever said no to helping a lab member. I must thank our former postdoc and my sincere and dear friend, Dr.

Marina Bakhmutsky for an endless number of reasons. I thank our former postdoc Dr. Shyamalee Kandegedara for offering intelligent insight. I thank all past graduate and undergraduate students including Casey Jackson and Nadeem Kandalaft, two really amazing and bright friends, and Thisari Guruge, Erandi Rajagurubandara, Asanka Rathnayake, Amanda Arnold, Cong Li, and Dr. Chandrika Canugovi for our time together. I would like to thank Dr. Michael Carpenter who was extremely helpful early on and an amazing teacher. I would also like to thank the funniest lab member we had, Dr. Rachel Parisien, an incredibly intelligent, helpful, charming, and wonderful person. I also must thank all the undergraduate lab assistants we have had. Dr. Bhagwat has spoiled us, as I have never had to wash beakers, autoclave, etc., allowing me to focus mainly on my research when in the lab. Without these assistants, things would not have been the same. So I would like to thank our current lab assistant Liam Holley for everything especially all of his help with my mouse work and our former assistant, Richard Evans. Finally, I also would like to thank my amazing friend Danielle Dremann from Dr. Chow's group. Basically, graduate school would not have been the same without her and all the wonderful times we have had together.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A: Adenine

Ab: Antibody

A-EJ: Alternative end joining

Ag: Antigen

AID: Activation Induced Deaminase

AP site: Apurinic/Apyrimidinic or Abasic Site

APE: AP endonuclease

APOBEC: Apolipoprotein B mRNA editing Enzyme- Catalytic polypeptide

ARP: Aldehyde Reactive Probe

ASM: American Society for Microbiology

BCR: B cell Receptor

BER: Base Excision Repair

BL: Burkitt Lymphoma

bp: Base pair

C: Cytosine

C57BL/6: C57 black 6 inbred lab mouse strain

Cμ: Constant Domain μ for IgM

cDNA: Complementary DNA

CIT: CD40L, IL-4, and TGF-β

CLL: Chronic lymphocytic leukemia

c-NHEJ: Classical Non-homologous end joining

CSR: Class Switch Recombination

D: Diverse

DCD: Deoxycytidylate deaminase

DLBCL: Diffuse Large B Cell Lymphoma

DNA: Deoxyribonucleic acid

DNA pol: DNA polymerase

dRP: Deoxyribose 5-phosphate

dRPase: Deoxyribophosphodiesterase

DSB: Double Strand Break

dUPTase: Deoxyuridine 5' triphosphate nucleotide-hydrolase

DZ: Dark Zone

E.coli: Escherichia coli

εC: ethenocytosine

EBV: Epstein-Barr Virus

Exo-1: Exonuclease 1

F: Female

F1: First Filial (daughter) Generation

F2: Second Filial (daughter) Generation

FACS: Fluorescence Assisted Cell Sorting

FBS: Fetal Bovine Serum

FDCs: Follicular Dendritic Cells

FEN-1: Flap Endonuclease 1

FES: Fetal Equine Serum

FL: Follicular Lymphoma

G: Guanine

GAPDH: Glyceraldehyde 3-phosphate dehydrogenase

GC: Germinal Center

GCV: Gene Conversion

GLT: Germline Transcript

GST: Glutathione S-Transferase

HIGM: Hyper-IgM Syndrome

HRS: Hodgkin and Reed-Sternberg Cells

I: Intronic/intervening

Ig: Immunoglobulin

IgH: Heavy immunoglobulin gene

IL: Interleukin

J: Joining

kb: kilobases

L: Leader

LM-PCR: Ligation-mediated PCR

LPS: Lipolysaccharide

LZ: Light Zone

M: Male

MBD4: Methyl-CpG binding domain protein 4

MM: Multiple Myeloma

MMR: Mismatch Repair

mRNA: messenger RNA

MSH2: MutS homolog 2

Mt: mutant

MUG: Mismatch Uracil Glycosylase

Mx: Methoxyamine

MZL: Marginal Zone Lymphoma

NE: Nuclear Extracts

NES: Nuclear Export Signal

NK: Natural Killer cells

NLS: Nuclear Localization Signal

NTS: Non-template Strand

Ny+: Positively charged Nylon

P: Promoter

PAGE : Polyacrylamide Gel Electrophoresis

PAMPs: Pathogen-associated molecular patterns

PBMCs: Peripheral Blood Mononuclear Cells

PBS: Phosphate Buffered Saline

P/S: penicillin- streptomycin

qRT-PCR: quantitative real time PCR

RNR: Ribonucleotide diphosphate reductase

S: Switch regions

SHM: Somatic Hypermutation

shRNA: short hairpin RNA

SLL: Small Lymphocytic Lymphoma

SMUG1: Single-strand specific monofunctional uracil-DNA glycosylase 1

SSB: Single Strand Break

SSH: Suppression subtractive hybridization

T: Thymine

TBP: TATA- binding protein

TDG: Thymine-DNA glycosylase

TLS: Translesion Synthesis Polymerases

TMϕs: Tingible-body Macrophages

TS: Template Strand

TSS: Transcription Start Site

U: Uracil

UDG: Uracil-DNA glycosylase

UGI: Uracil-DNA glycosylase inhibitor

UNG: Uracil-DNA glycosylase

V: Variable

WCE: Whole cell extracts

WRC: W=Adenine/Thymine, R=Purine, C=Cytosine

WT: Wild type

CHAPTER 1: INTRODUCTION

1.1 Activation- Induced Deaminase

Maintaining sequence fidelity of the genetic information coded in deoxyribonucleic acid (DNA) is essential to the development and function of organisms because inadvertent sequence changes may result in adverse consequences to the cell, daughter cells, and the organism. Accordingly, many DNA repair processes have evolved to preserve the genome during normal cellular processes such as replication. Although there exists immense pressure to protect against genomic instability and corruption of sequence context, beneficial programmed alterations to genomic DNA which lead to sequence changes do occur in specialized processes of the adaptive immune system. Through targeted mutations of the immunoglobulin gene of B cells, activation-induced deaminase (AID) equips the adaptive immune system with the capability of producing a diverse repertoire of antibodies against an almost infinite number of antigens and pathogens. However, this mutagenic potential of AID can also be deleterious to B cells if it is ectopically expressed, fostering the possibility for genomic instability. The focus of this research is to quantify the DNA lesions introduced by AID in normal and cancerous B cell genomes for the first time, and track the unfavorable consequences that may arise from the aberrant activity of AID which may promote development of B cell lymphomas.

1.1.1 The Discovery of AID and its Requirement for Antibody Maturation

AID was discovered in 1999 in the murine B cell lymphoma cell line CH12F3 (1) which was isolated as a model for class switch recombination (CSR), a mechanism required to make the antibody isotype most fit for a specific category of antigens (see section 1.2.4 for further details). Specifically, CH12F3 cells class switch from IgM isotype to IgA when stimulated with a particular cocktail of cytokines (CD40L, IL-4, TGF-β or 'CIT'). Honjo and colleagues investigated the genes with upregulated expression upon CIT stimulation of CH12F3 cells using suppression subtractive hybridization (SSH) (2). This PCR-based method allowed for the amplification of only complementary DNA (cDNA) that differed between the transcriptome of stimulated cells and unstimulated cells. AID expression was found to be highly induced only when cells were stimulated to undergo CSR.

A role for AID in the maturation of antibodies was immediately attractive due to its largely restricted tissue expression in activated B cells (2). Critical supportive evidence of AID's role in antibody maturation was the finding that CSR and somatic hypermutation (SHM), a second mechanism required for antibody affinity maturation (see section 1.2.3 for further details), were both abolished in mice with a homozygous deficiency in AID (3). Concomitantly, Durandy and colleagues determined that patients with hyper-IgM syndrome (HIGM2), where CSR and SHM are found to be defective, had mutations in the human gene encoding for AID (4). Thus, AID is indispensable for CSR, SHM, and antibody affinity maturation.

1.1.2 AID and the Deamination of Cytosine to Uracil in DNA

Honjo *et al* explored the ability of a GST-AID fusion protein to deaminate cytidine *in vitro* since early sequence alignment studies revealed homology of AID to Apolipoprotein-B mRNA editing enzyme catalytic polypeptide-1 (APOBEC-1, 34% amino acid identity). APOBEC-1 was the only well-characterized APOBEC family member at the time (2) and since it was known to be an RNA editor that deaminates cytidine to uracil in the mRNA of apolipoprotein-B (5), it was logically proposed that AID similarly functioned by editing an unknown mRNA substrate to indirectly lead to CSR and SHM. However, this mRNA-editing model of AID in adaptive immunity is not supported by the vast majority of evidence in the field.

The ectopic expression of AID in eukaryotic cells such as hybridomas (6), T cells (7), fibroblasts (8), and yeast (9) resulted in mutations at cytosines in their genomic DNA. Similarly, ectopically expressing AID in *E. coli* resulted in mutations at cytosines in the genome and AID-dependent mutations were further enhanced in *E. coli* deficient in uracil-DNA glycosylase (UNG), the primary enzyme that removes uracils in DNA (10). AID is not naturally present in prokaryotes such as *E. coli*, but its ability to promote mutations in the *E. coli* genome suggested that an mRNA intermediate was unlikely the substrate for AID in B cells. Based on these *E. coli* studies, Petersen-Mahrt *et al* envisioned a model where AID directly deaminates cytosines in DNA to uracil specifically in the Ig locus of B cells (see Figure 1). Pathways for SHM and CSR were hypothesized to occur from the processing of these uracils by other downstream

enzymes such as UNG2, error-prone polymerases, mismatch repair (MMR) proteins, and AP endonuclease (10).

Figure 1: AID Dependent Hydrolytic Deamination of Cytosine to Uracil

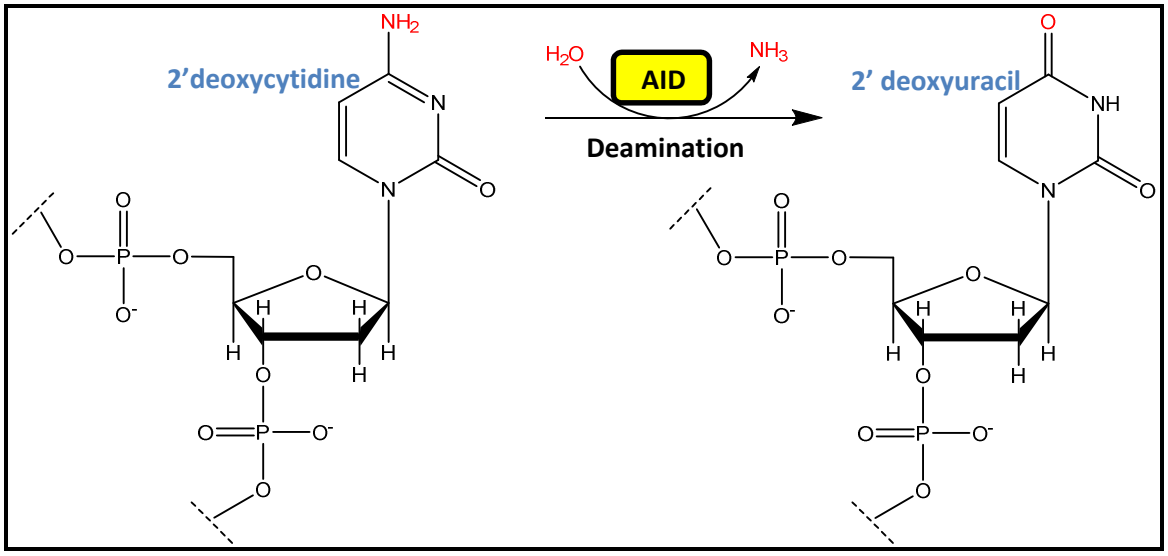

Figure 1: AID Dependent Hydrolytic Deamination of Cytosine to Uracil. AID catalyzes the nucleophilic attack of a water molecule at the C4 position of cytosine. The loss of the amino group as ammonia, results in the formation of uracil. Figure created in ChemBioDraw Ultra 14.0.

Several groups provided biochemical evidence that supports DNA, specifically single-stranded DNA or single-stranded regions in transcribed double-stranded DNA, as the substrate for AID cytosine deamination. AID was shown not to deaminate cytosines in RNA, non-transcribed double-stranded DNA, or DNA-RNA hybrids (11-14). Furthermore, *in vitro* studies determined that the hotspot for AID deamination is a cytosine occurring in a WR**C** motif (where W is an A or T, R is a purine, ie G or A) (15). This is in agreement with the WRC SHM hotspot motif determined *in vivo* (15). Substantial evidence in the field has predominantly supported the direct 'DNA-editing model' of AID at the Ig gene, and thus this model prevails over the originally proposed 'mRNA editing model'. It

is widely accepted that AID catalyzes the hydrolytic deamination of 2'-deoxycytidine (C) to 2'-deoxyuracil (U) in a sequence context specific manner and prefers single-stranded DNA regions in the Ig gene (see Figure 1) (16).

Figure 2: The Primary Structure of AID

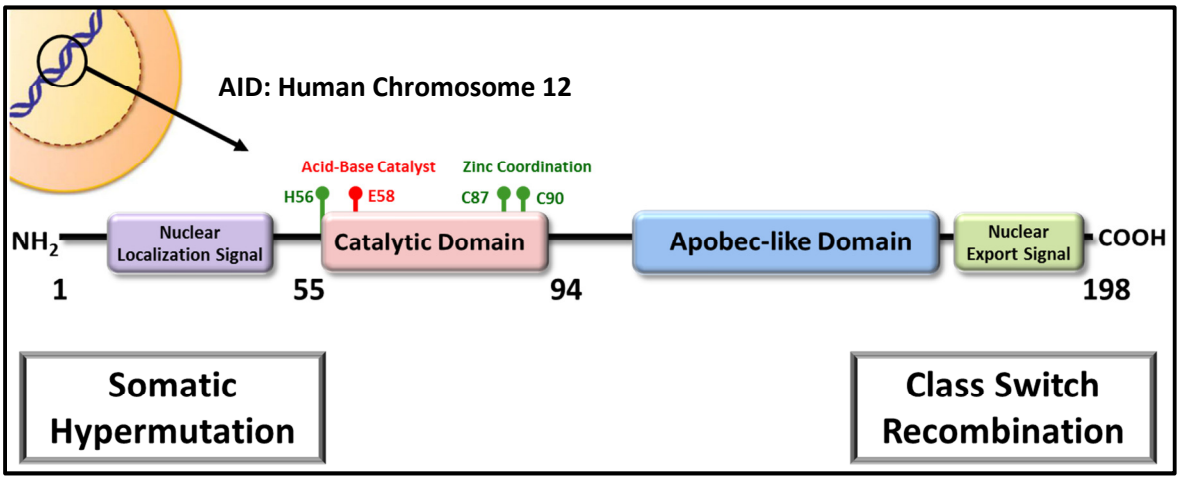

Figure 2: The Primary Structure of AID. AID is found on human chromosome 12 and has 5 exons. The gene codes for a 198 amino acid protein. The primary structure of AID is depicted above with major domains of the protein highlighted. AID has an amino terminal nuclear localization signal (NLS), and a carboxy terminal nuclear export signal (NES). The catalytic domain is required for both somatic hypermutation (SHM) and class switch recombination (CSR). Catalysis of the hydrolytic deamination of C to U depends on the glutamate residue at position 58 (E58) for acid-base catalysis, and the histidine at residue 56, and cysteines at positions 87 and 90 (H56, C87, and C90 respectively) for zinc coordination to activate a water molecule. The C-terminus of the protein is required for CSR while the N-terminus is required for SHM.

1.1.3 AID and Subcellular Localization

The gene encoding AID has five exons and is located in close proximity to the APOBEC-1 gene on the same chromosome in both mice (chromosome 6) and humans (chromosome 12). As described above, the catalytic domain or zinc coordination motif of AID is conserved with other members of the APOBEC family of proteins. Mutation of the glutamate residue at position 58 (E58) in the catalytic domain of AID can result in a catalytically inactive mutant (see Figure 2).

E58 is thought to participate in general acid-base catalysis and residues H56 (histidine), C87 and C90 (cysteine) coordinate with a Zn^{2+} along with an activated water molecule to mediate nucleophilic attack of the carbon-4 of cytosine, promoting deamination or loss of the amino group (see Figure 3) (16, 17). This catalytic domain of AID is required for both *in vivo* SHM and CSR.

With a functional catalytic domain intact, mutational experiments of AID revealed that the carboxy-terminal (C terminal) domain of the AID protein (198 amino acids, ~24 kDa) is necessary for CSR, while the amino-terminal (N terminal) domain is required for only SHM (18-20). The C and N terminal domains also play a role in the subcellular localization of AID. Since AID acts on genomic DNA of activated B cells, nuclear localization is required. However, because of its mutagenic potential, AID is actually a nucleocytoplasmic shuttling protein (21). A nuclear export signal NES is located in the last ~10 residues of the C terminal domain (22-24) while a nuclear localization signal (NLS) was identified encompassing most of the N terminal domain (see Figure 2) (22, 25). Recent work has also demonstrated that subcellular compartmentalization of AID also depends on active cytoplasmic retention which may afford overall genomic protection from this DNA mutator enzyme (25).

Figure 3: Proposed Mechanism of AID Catalyzed Deamination of Cytosine

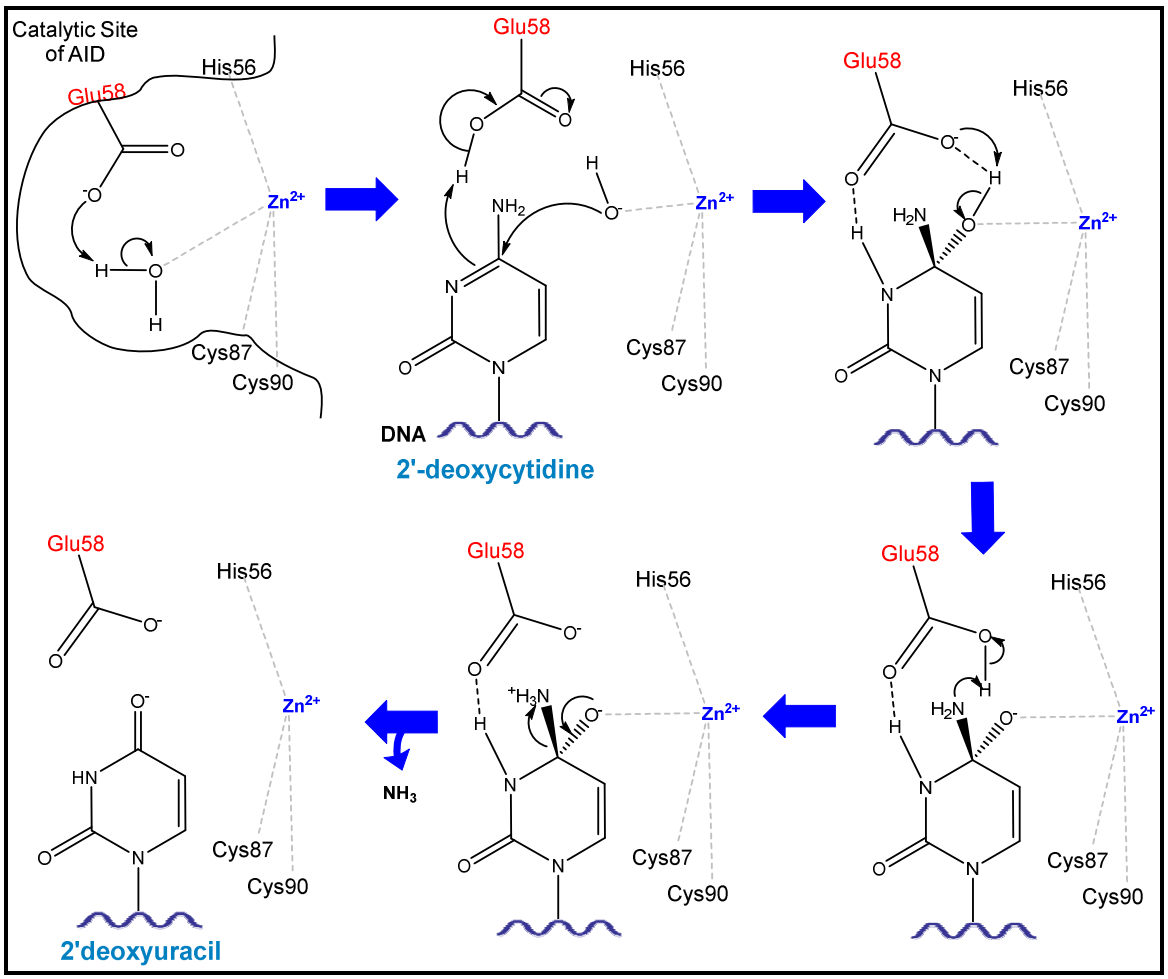

Figure 3: Proposed Mechanism of AID Catalyzed Deamination of Cytosine. The catalytic pocket of AID contains the glutamate residue at position 58 (Glu58) which participates in general acid-base catalysis for the activation of a water molecule. Cysteines at positions 87 and 90 (Cys87, Cys90) and histidine 56 (His56) coordinate a zinc along with the water molecule. The remaining hydroxyl group serves as a nucleophile to attack the C4 position of a cytosine, yielding an intermediate that eventually gives rise to uracil once an ammonia group leaves. The proposed mechanism shown above for AID is based on homology with cytidine deaminases and mutational studies (16, 17, 26). Figure created in ChemBioDraw Ultra 14.0

1.2 The Role of AID in the Adaptive Immune System

The immune system of vertebrates is composed of innate immunity and adaptive immunity (27). Innate immunity is the immediate line of defense and is nonspecific towards the antigen. Anatomical barriers such as mucosal and epithelial surfaces, inflammation, and broad recognition of pathogen-associated molecular patterns (PAMPs) by particular cells of the immune response system, define this branch of innate immunity (28). However, innate immunity does not offer long-term protection or 'memory' against previously encountered antigens but plays a role in the activation of adaptive immunity. This second branch of the immune system, also referred to as acquired immunity, is antigen-specific and has the ability to confer immunological memory. AID plays a role in producing antibodies required for an appropriate adaptive immune response.

1.2.1 The Structure and Function of Immunoglobulins

Antibodies, or immunoglobulins, are either expressed on the surface membrane of B cells as B cell receptors (BCR) or are secreted by plasma cells, a specific type of differentiated B cell (28). Immunoglobulins are glycoproteins with a base unit of a homodimer of a heterodimer (27). The heterodimer is comprised of a longer 'heavy' polypeptide chain and a shorter 'light' chain. Disulfide bonds bridge the heavy chain and light chain of the heterodimer as well as the homodimer (see Figure 4A). There are two antigen binding pockets, or variable domains, in the base unit of antibodies (see Figure 4A). Mutations from SHM produce antibodies with different primary sequences for their antigen sites, hence 'variable domain,' to confer high affinity against antigen (see Figure 4B, and

Figure 5B). In contrast, CSR or lack thereof, determines the constant domain of the immunoglobulin (see Figure 4C and Figure 6). The final antibody structure can be secreted as a monomer, dimer, or even hexamer of this base unit depending on the isotype (ie. IgM can form a pentamer with ten antigen binding sites).

Figure 4: Structure of Immunoglobulins

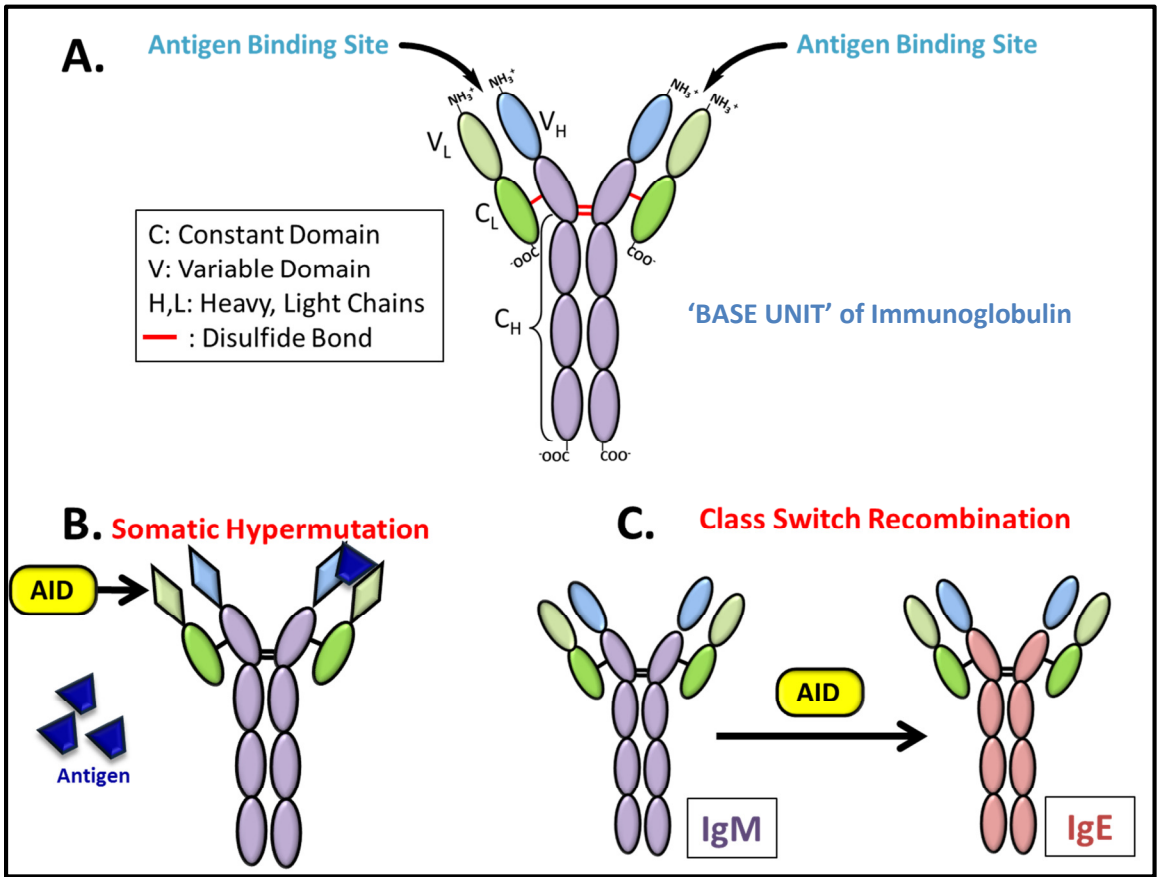

Figure 4: Structure of Immunoglobulins. (A) Immunoglobulins, or antibodies, consist of a base unit of a homodimer of a heterodimer. The heterodimer consists of the 'heavy' polypeptide chain and the shorter 'light' chain. Thus, the homodimer has two heavy and two light chains as depicted. Disulfide bonds bridge the heterodimer and the homodimer. There are two antigen binding pockets, or variable domains, in the base unit of antibodies. (B) Somatic hypermutation occurs when AID introduces mutations that alter the variable regions of the Ig gene to produce higher affinity antibodies for specific antigen, indicated here as a change in shape of the antigen binding pockets to fit the antigen more appropriately. (C) Class switch recombination occurs when AID

introduces double strand breaks that eventually lead to juxtaposition of a new constant domain downstream of the variable region resulting in a gene that codes for a new isotype. CSR is depicted here as a change in color of the constant heavy chain from IgM (purple) to IgE (pink).

High affinity immunoglobulins not only neutralize foreign pathogens but are also essential in activating several other immune responses to combat infection. The different isotypes of antibodies (IgM, IgD, IgG, IgA, and IgE) are coded by the different constant regions in the heavy chains of the Ig gene (Cμ, Cδ, Cγ, Cα, and Cε respectively and see Figure 6). CSR is designed to yield the immunoglobulin isotype that is best equipped to eradicate the specific pathogen based on the nature of the antigen and the ability of the antibody to reach the tissue site of the infection. For example, since IgM is secreted as a large pentamer or hexamer, it cannot traverse into extravascular space. This makes it unsuitable to neutralize antigens distributed in that tissue, in contrast to monomeric IgG. In addition, antibodies are switched to the isotype best able to eradicate the specific pathogen type. For example, IgE is suited for parasitic infections (29) and IgA is effective against mucosal bacteria (30).

1.2.2 Antibody Affinity Maturation

The initial expansion of the antibody repertoire arises from the process of V(D)J recombination which is both independent of available antigen and T-cells, and does not require AID. This is a combinatorial process that occurs early in B cell development to establish a variable region through rearrangements of the variable (V), diversity (D) and joining (J) segments in the Ig gene (31, 32). Further expansion and diversification of the antibody repertoire is achieved by

SHM and CSR which are antigen-dependent processes that continue to alter the Ig gene and as described above, absolutely require AID. All three processes are essential for antibody maturation.

Figure 5: SHM of the Variable Region of the Immunoglobulin Gene

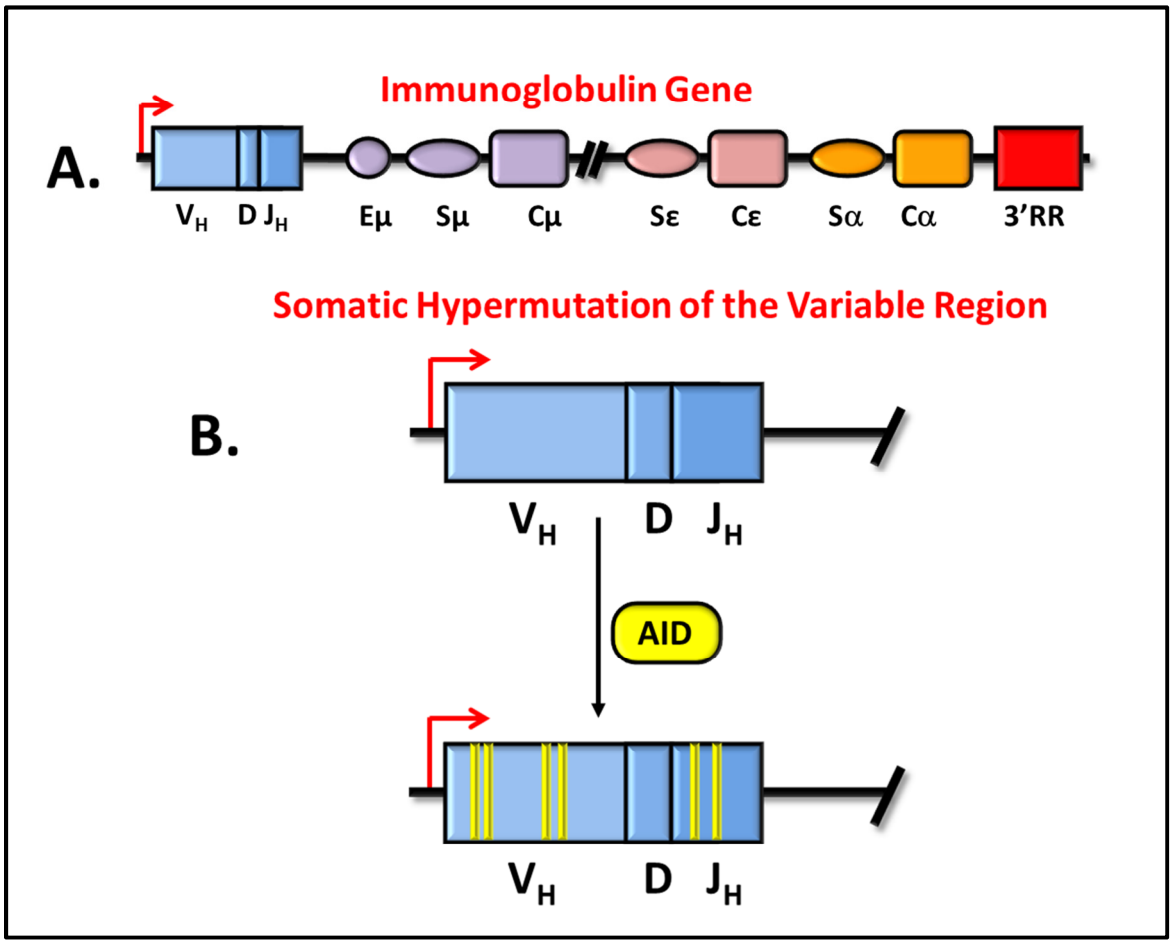

Figure 5: SHM of the Variable Region of the Immunoglobulin Gene. (A) A truncated immunoglobulin gene (heavy chain) with a rearranged V(D)J region is schematically depicted. The variable region, V(D)J, is depicted in blue. The enhancer (Eμ), switch region (Sμ), and constant domain (Cμ) that codes for IgM (purple) is downstream of the variable region. Thus, without class switching, the default encoded Ig isotype is IgM. The Sε and Cε (pink) are the respective switch region and constant domain for the IgE isotype, and Sα and Cα (orange) are the switch and constant domains for the IgA isotype. The 3' Regulatory Region (3'RR, red) is composed of enhancers that promote efficient CSR (33) and SHM (34). (B) Somatic hypermutation occurs when AID-dependent point mutations, depicted as yellow bars, are introduced in the V(D)J region of the Ig gene. AID

acts on single-stranded DNA during transcription and the V(D)J transcription start site (TSS) is shown as a red arrow.

Figure 6: Class Switch Recombination

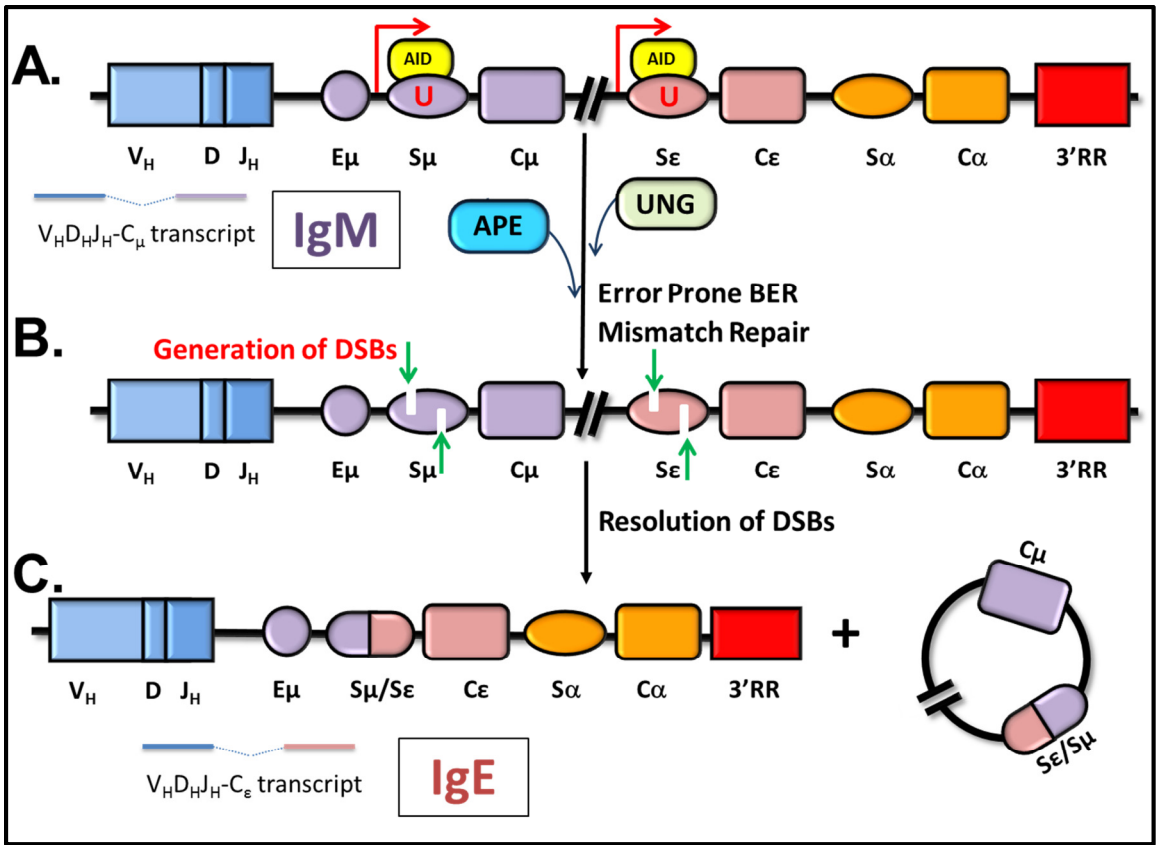

Figure 6: Class Switch Recombination. Elements of the Ig gene are depicted similar to that in Figure 5. (A) Before CSR, the Ig gene codes for an IgM transcript. Upstream of each constant domain is a switch region, S, that is transcribed from an intronic promoter (red arrows) since AID requires transcription. When class switching from IgM, the default coded constant region, to another isotype, AID deaminates cytosines to uracils in Sμ and $S_{downstream}$. $S_{downstream}$ in this case is Sε. (B) Through the excision of uracil by UNG, nicking of the AP site by APE, and error prone BER or MMR, double strand breaks (DSBs, green arrows) are introduced. (C) Upon resolution of the DSBs by non-homologous end joining or alternative end-joining (not shown), an Sμ and Sε synapsis brings the Cε region downstream of the V(D)J which allows for the coding of IgE. The intervening DNA region which contains the Cμ domain is ligated into a 'switch circle' and deleted.

Post V(D)J recombination, AID introduces point mutations within the rearranged V(D)J segments of the Ig gene in B cells during SHM (see Figure 5B) (2-4, 35). Those B lymphocytes which attain somatic mutations that improve antibody affinity for the antigen are positively selected for while those B cells that acquire mutations that reduce affinity of the antibody for the antigen are removed from the B cell population via programmed cell death. This clonal selection allows for the evolution of progressively higher affinity antibodies with each continuous and iterative step of SHM (see Figure 7 Clonal selection) (36, 37).

Figure 7: Clonal Selection of High-Affinity Antibodies

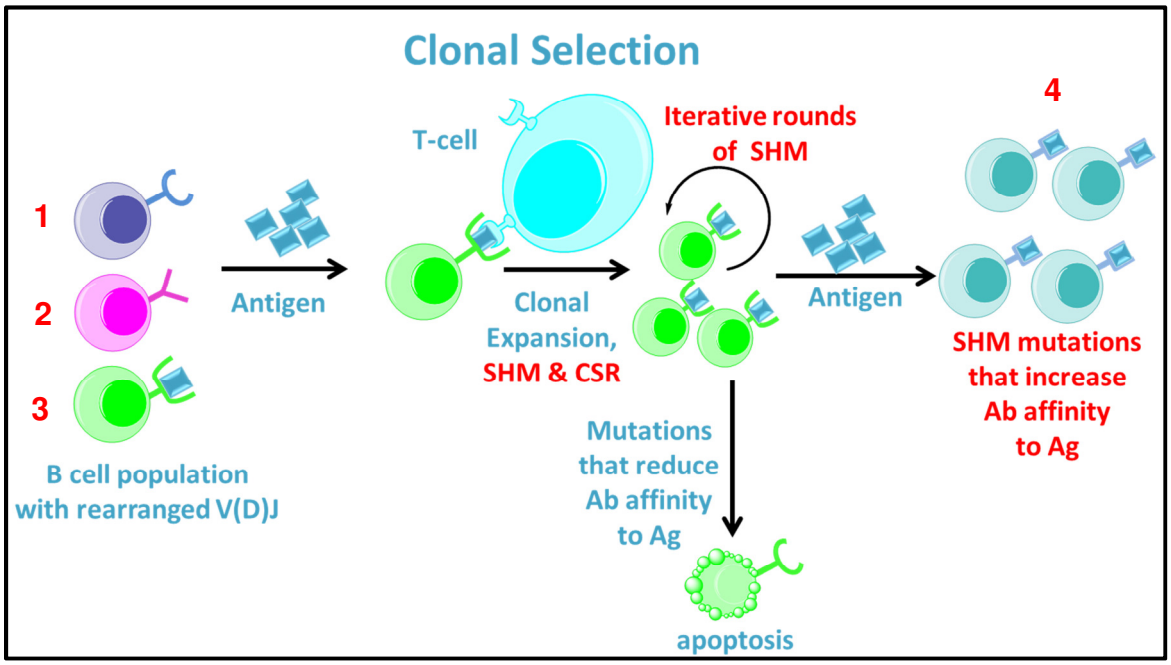

Figure 7: Clonal Selection of High-Affinity Antibodies. Naïve B cells (clones 1,2, and 3) having undergone V(D)J recombination express different surface antibodies (Ab). Only the Ab on the surface of clone 3 has some affinity to a specific antigen (Ag, blue rectangle). Through receptor binding, T cells recognize this Ab-Ag complex and mediate clonal expansion by promoting proliferation of clone 3. SHM and CSR help produce high affinity Abs and the correct isotype needed for the specific antigen. Clones that acquire mutations during iterative rounds of SHM and Ag exposure (circular arrow) that result in reduced or poor affinity Abs undergo programmed cell death, while those that acquire beneficial

mutations expand. ‘Clone 4’ has class switched and has a high affinity Ab to the specific Ag. Figure created in ChemBioDraw Ultra 14.0

CSR is a programmed DNA recombination process that yields the most effective antibody isotype needed during each immune response (35, 38). During CSR, the rearranged variable region is joined with one of several different downstream constant regions (ie Cα, Cγ, Cε) that each confer a specialized optimal effector function. Switch (S) regions which contain short repetitive sequences are located upstream of the constant regions in the Ig genes (see Figure 6). Following AID-dependent introduction of double-strand breaks in Sμ and a second downstream S region, the intervening sequence is circularized and deleted while the two different S regions are joined together (39). This juxtaposes the V(D)J exon to a downstream constant region exon that results in isotype switching from IgM to IgA, IgG, or IgE (see Figure 6 for a schematic of IgM to IgE isotype switching).

In addition to SHM and CSR, some vertebrates evolved a third AID-dependent mechanism of antibody affinity maturation known as gene conversion (GCV) (40, 41). In species, such as chicken, pigs, and rabbits, GCV is the dominant method of V segment diversification (42, 43). This process does not take place in the biological systems studied in this work, mice and human B cells, and thus will not be discussed in detail here.

1.2.3 The Mechanism of Somatic Hypermutation

1.2.3a Mutational Characteristics of the Immunoglobulin Gene and SHM

During SHM, AID introduces mutations in the variable segment of the heavy chain of the Ig gene in activated B cells, with the highest occurrence of mutations found in the exons of the V(D)J segment. In addition mutations are acquired in the introns of the J regions as well. However, AID avoids mutating the C domain genes. The promoter regulatory regions and an enhancer element located in the intron also evade AID induced mutations (see Figure 8) (44, 45).

Figure 8: Intensity and Boundaries of SHM in the Variable Region

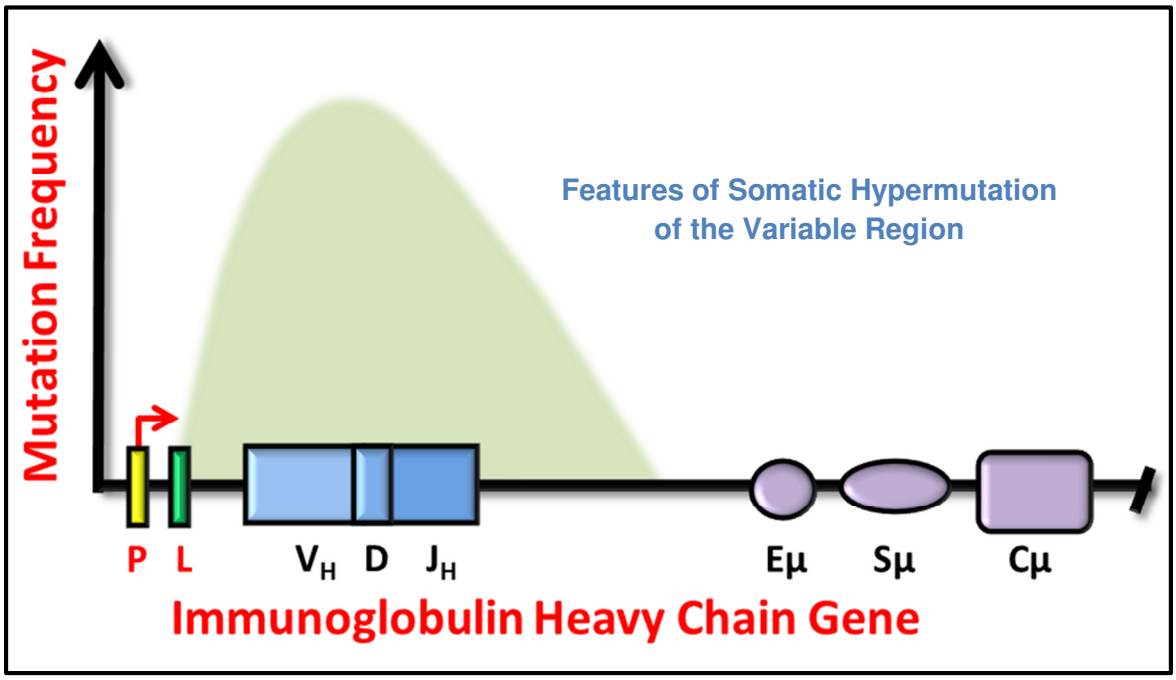

Figure 8: Intensity and Boundaries of SHM in the Variable Region. The level of the mutation frequency (y axis) of the immunoglobulin heavy chain is schematically (not to scale) depicted along the regions of the Ig gene (x axis). The transcription start site (red arrow) begins at the promoter (P). Somatic hypermutations are found approximately 100 to 200 base pairs downstream from the TSS(46) with the highest accumulation of mutations found within the V(D)J region. The mutations accumulate up to approximately 1500 base pairs downstream of the TSS (47). The enhancer and constant domain (Eμ and Cμ, respectively) do not undergo SHM. Adapted from figures in references (46, 48).

SHM may occur in switch regions as well but is not shown or discussed here (48).

Mutations in SHM are predominantly found in the form of single base changes, or 'point mutations,' and accumulate approximately 100 to 200 base pairs (bps) from the transcription start site (TSS) with the mutation spectrum extending up to approximately 2.0 kb downstream of the TSS (44-46, 49, 50). Furthermore, in SHM, most of the point mutations are transitions (e.g. pyrimidine mutated to a pyrimidine, purine mutated to a purine) rather than transversions (e.g. pyrimidine mutated to a purine and vice versa). Since AID has sequence selectivity, it is not surprising that many mutations in the V(D)J segment are targeted to the WRC hotspot motif (15, 47, 51).

The frequency of SHM mutations is approximately 10^6 fold higher than that for background genomic mutations- the V(D)J segment can accrue roughly 10^{-3} mutations per base pair per generation (16). High SHM mutation frequency in the variable region also correlates with high transcription levels (52). SHM requires transcription, but transcription alone is not sufficient to ensure AID activity. Transcription is thought to provide the single-stranded DNA substrate needed by AID, in the form of a 'transcription bubble' (53). SHM is not strand biased, with a similar mutation frequency seen in both the template strand (TS) and the nontemplate strand (NTS) (54).

1.2.3b Error-Free Base Excision Repair of Uracils

Simple direct replication across from an AID-introduced uracil will eventually produce a G/C transition mutation after a second round of replication (see Figure 9 *Upper Right*) (10). The replicating DNA polymerase will insert an A across from a template U, effectively treating the DNA lesion as a template T. The mutations of SHM are not just C to T transition mutations and not restricted to G/C bases but occur just as efficiently at A/T base pairs in SHM. Thus, simple replication cannot be the only mechanism in activated B cells that leads to a hypermutated Ig gene.

Figure 9: Model of Somatic Hypermutation of the Variable Region

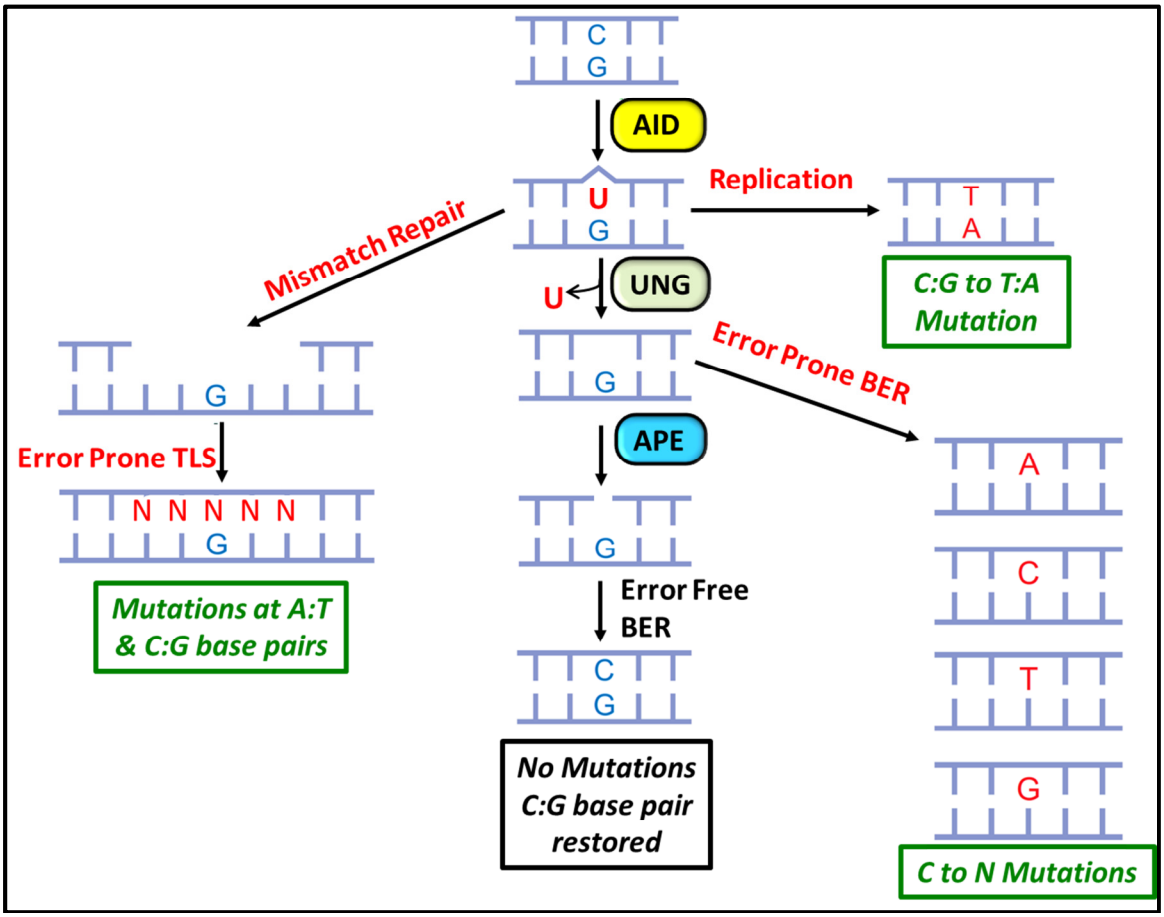

Figure 9: Model of Somatic Hypermutation of the Variable Region. AID deaminates C to U in the V(D)J region of the Ig gene of B cells. The resulting U•G mispair can be processed in several ways. *(Upper Right)* Replication past a U will result in the insertion of an A, eventually resulting in a C to T transition mutation in one daughter cell after a second round of replication. *(Lower Right)* Uracil DNA-glycosylase (UNG) can excise the uracil resulting in an AP site that can be bypassed by error-prone translesion synthesis polymerases (TLS). These polymerases may insert any of the four nucleotides resulting in C to N mutations. (*Middle*) Canonical base excision repair (BER) depends on the excision of U by UNG, nicking of the AP site by APE, and polβ and ligase to correctly restore the C:G base pair. *(Left)* Mismatch repair (MMR) proteins, specifically MSH2/6, recognize the U•G mispair. Exonuclease 1 helps to create a gap around the initial U lesion and TLS polymerases are then recruited and insert any of the four nucleotides throughout the excised gap. This results in transition and transversion mutations at A:T and G:C base pairs.

The evolution of antibody affinity maturation which promotes mutations of the Ig gene requires more than just AID, a targeted DNA mutator. It also requires evolution of downstream pathways to ensure the AID generated uracils are not simply excised and repaired by the several highly conserved DNA damage repair systems in place to protect genomic instability. To circumvent error-free repair of the uracils, many enzymes involved in canonical genomic uracil repair pathways have been hijacked to play counter-intuitive roles in activated B cells to ensure Ig gene mutations.

Since uracil is introduced into genomes by pathways other than the enzymatic deamination of cytosines by members of the AID/APOBEC family of proteins, highly conserved DNA repair pathways have evolved to remove and excise genomic uracils, restoring sequence context. The misincorporation of uracil during replication, and the non-enzymatic spontaneous deamination of cytosine to uracil, both result in uracilated DNA. Uracil generated through misincorporation of a dUMP occurs when the replicating polymerase uses dUTP as a substrate rather than dTTP (see Figure 10 for details). This results in U:A base pairs that are not directly mutagenic. In contrast, spontaneous deamination of a cytosine in a C:G base pair results in a U•G mispair that can lead to a C to T mutation if unrepaired and replicated past, similar to consequences from AID deaminated cytosines (see Figure 11).

Figure 10: Pathways for Uracil Incorporation in DNA

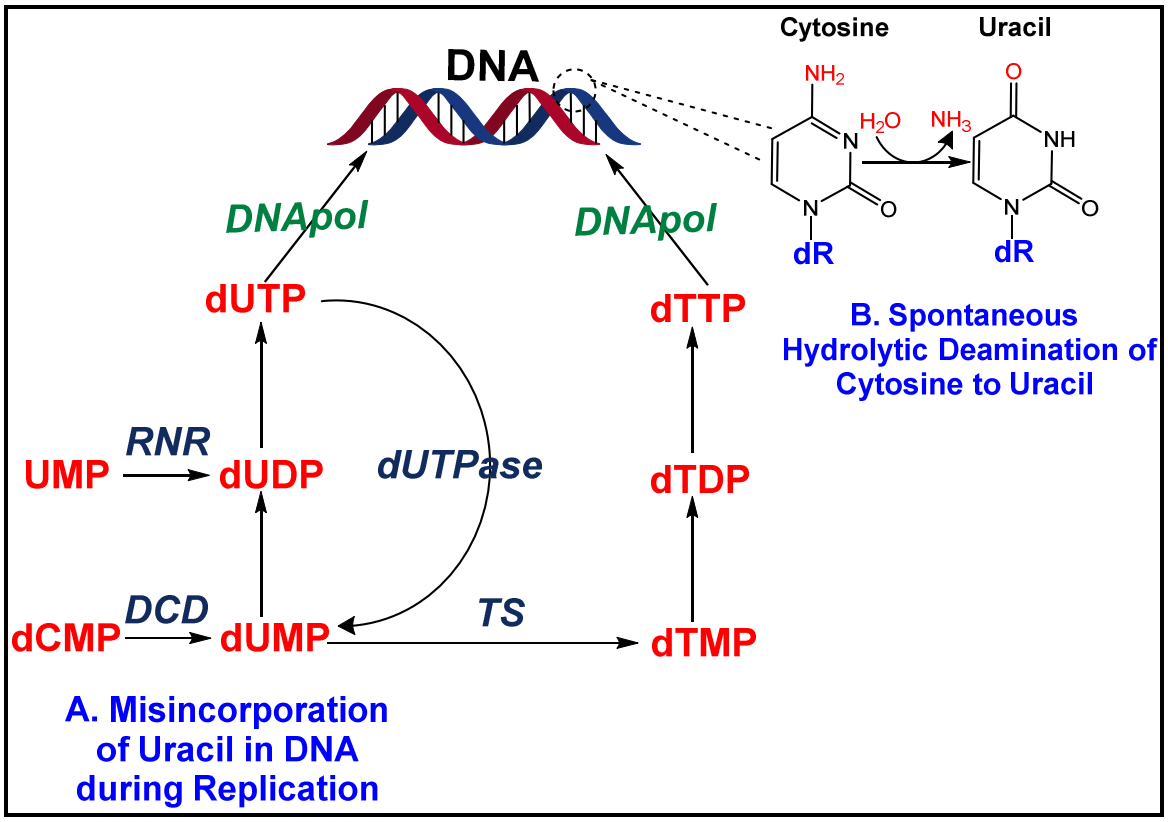

Figure 10: Pathways for Uracil Incorporation in DNA. (A) Uracil can be misincorporated into DNA during replication. Generally, this is avoided by actively keeping the dUTP pool low due to dUTPase (deoxyuridine 5' triphosphate nucleotide-hydrolase). As depicted above, dUTPase hydrolyzes dUTP to dUMP (and pyrophosphate, not shown). Thymidylate synthase reductively methylates dUMP during *de novo* synthesis of dTMP. dTMP is then converted to dTTP which is a substrate for DNA polymerase (DNA pol) during replication. dUMP is produced from the deamination of dCMP by DCD (deoxycytidylate deaminase). In addition, RNR (ribonucleotide diphosphate reductase) converts UMP to dUDP which is then phosphorylated to dUTP. DNA pol is able to use available dUTP as a substrate and inserts uracil across from a template A during replication. (B) Uracil can also occur in DNA through hydrolytic non-enzymatic spontaneous deamination of cytosine as depicted. Cytosines can also be deaminated chemically, for example by bisulfite(55) and nitrous anhydride(56) (not shown). Figure created in ChemBioDraw Ultra 14.0

Figure 11: Mutational Consequences of Uracil in DNA

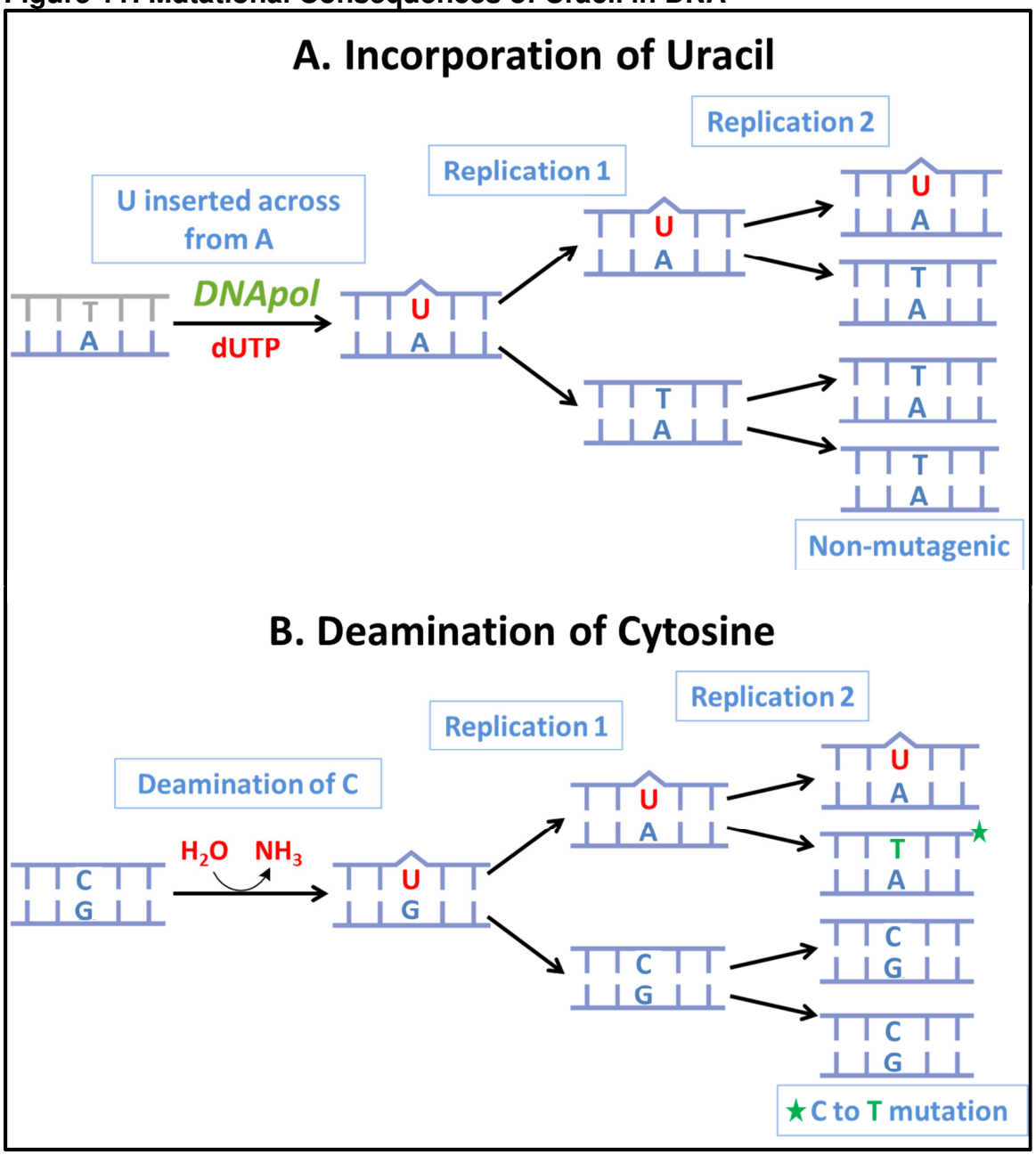

Figure 11: Mutational Consequences of Uracil in DNA. (A) The misincorporation of dUMP instead of dTMP by DNA polymerase (DNA pol) yields a U:A base pair. When this undergoes replication (Replication 1) a normal T:A base pair is generated and an A is inserted across from the U in the top strand. A second replication event (Replication 2) also results in normal T:A base pairs and further dilutes the non-mutagenic U:A base pair until it can be excised and repaired. (B) Deamination of cytosine to uracil, whether spontaneous or enzymatic is mutagenic. This results in a U•G mispair. Replication past this (Replication 1) results in a U:A base pair as well as a C:G base pair. Another replication event (Replication 2) fixes a C to T mutation (green star).

Figure 12: Uracil DNA-Glycosylase

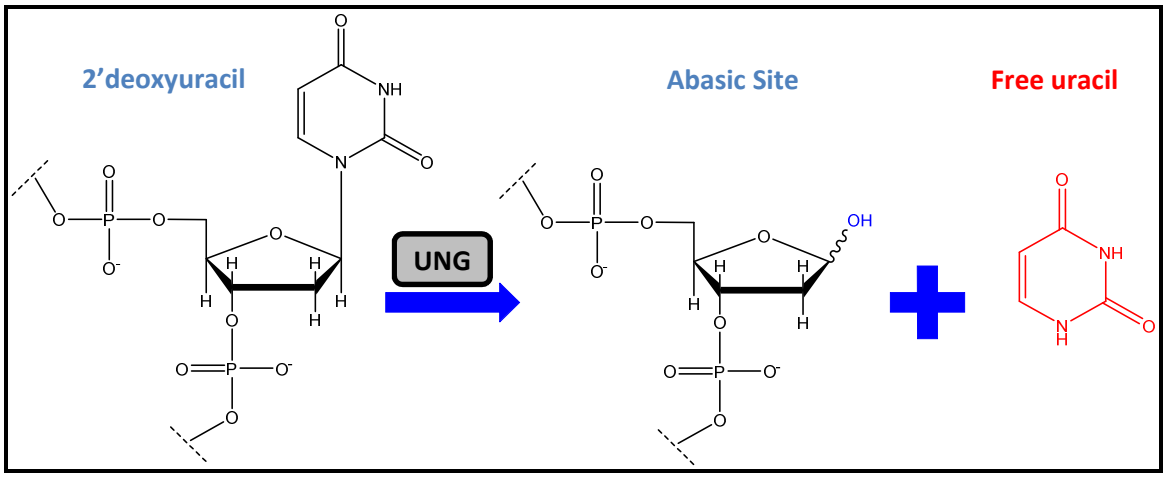

Figure 12: Uracil DNA-Glycosylase. Uracil DNA-Glycosylase (UDG) recognizes uracil in single-stranded DNA, or in U:A and U•G pairs. Through hydrolysis of the N-glycosidic bond, UDG excises uracil from DNA. This results in the formation of an AP site, or abasic site, and the release of free uracil and is the first step in base excision repair. Figure created in ChemBioDraw Ultra 14.0

The faithful process of base excision repair (BER) (see *Middle*, Figure 9) maintains the integrity of DNA by removing and repairing various lesions and adducts through a multiple-step pathway. Canonical BER is initiated by the excision of the DNA lesion by DNA glycosylases capable of recognizing and excising DNA adducts through hydrolysis of the N-glycosidic bond. Thus, uracils are removed by uracil-DNA glycosylases (see Figure 12). There are four types of uracil-DNA glycosylases (UDG or UNG) in mammalian systems which include UNG, single-strand specific monofunctional uracil-DNA glycosylase 1 (SMUG1), thymine-DNA glycosylase (TDG), and methyl-CpG binding domain protein 4 (MBD4) (57, 58). The mammalian UNG gene encodes the two existing isoforms of UNG, mitochondrial UNG1 and nuclear UNG2 (59). The evolution of several redundant glycosylases with the ability to remove genomic uracil highlights the selective pressure against uracils accumulating in the genome. TABLE 1 lists all

the different DNA lesions in the context they are recognized in by the individual glycosylases capable of removing uracils. UNG2 is the major glycosylase responsible for removal and initiating repair of genomic uracil with SMUG1 thought to be the primary back-up enzyme (57, 60).

TABLE 1: Substrates of the Human Uracil DNA Glycosylases

Glycosylases	Substrates	References
UNG1/UNG2	**U** in ssDNA **U**:A **U**•G	(61)
SMUG1	**U** in ssDNA **U**:A **U**•G **5hmU** in ssDNA **5hmU**:A **5FU**:A	(62, 63)
TDG	**T**•G **U**•G **εC**:G	(64)
MBD4	**T**•G **U**•G	(65)

TABLE 1: Human Uracil DNA Glycosylases are indicated with substrates shown in bold.

Upon recognition of uracil, UNG2 cleaves the N-glycosidic bond to release free uracil base generating an abasic site (AP) in the DNA. AP endonuclease (APE) then nicks the AP site. The deoxyribophosphodiesterase (dRPase) activity of DNA polymerase β (Polβ) cleaves the 3' end of the nicked abasic site and releases deoxyribose-5-phosphate (dRP). Polβ then inserts the correct base according to the template- a C to restore the C:G base pair if the uracil resulted

from deamination or a T to restore the T:A base pair if the uracil was misincorporated during replication. DNA ligase then seals the nick to complete what is termed ‘short patch’ BER, the removal of only the uracil base and the insertion of one correct nucleotide (see Figure 9 *Middle*). ‘Long-patch’ BER involves the replicative machinery and polymerases, flap endonuclease-1 (FEN1), and removal and repair of a gap of ~ 2 to 10 nucleotides to repair the uracil lesion (66). Contradictory to their primary roles, during SHM of activated B cell, UNG2 and APE function to promote mutations.

1.2.3c Error-Prone Resolution and UNG2 in SHM

Similar to patients deficient in functional AID protein with HIGM2 syndrome, mutations in the human UNG2 gene result in a similar disorder termed HIGM5 where their antibodies are also compromised in SHM and CSR (67, 68). The introduction of mutations during SHM is dependent on UNG2 for removal of the AID induced uracil and generation of an AP site (69, 70). UNG2 is primary for shaping the mutational spectrum of SHM at G/C base pairs. While mutations at A/T base pairs occur during SHM, they seem to be predominantly dependent on a different repair pathway (see section 1.2.3d).

UNG2 excises uracils introduced in the variable region of the Ig gene by AID similar to the first step of error-free BER. Since the resulting AP sites have no instructive base, they cannot be bypassed by replicative polymerases and thus, error-prone translesion synthesis polymerases (TLS) are recruited. These polymerases are capable of bypassing the AP site (71) since the promiscuous nature and more permissive binding site of the TLS polymerases allows for the

insertion of any of the four nucleotides across the AP site. This results in G/C transitions and transversions during SHM (72) (see Figure 9).

1.2.3d Mismatch Repair and A/T Mutations in SHM

Replicative polymerases are equipped with proofreading capabilities to correct mismatch errors that may occur during insertion of an incorrect dNMP. However, if a replication error escapes detection as well as correction by the replicative polymerase, a highly conserved repair pathway is available for recognition and repair of the mismatched base pair to avoid mutations. This pathway termed mismatch repair (MMR), is a multi-step process which depends on several proteins for the recognition, removal, and repair of the DNA as well as complex interactions with the DNA replication machinery (73, 74). Not unlike the BER pathway, some MMR proteins have also been exploited to actually promote mutations in the Ig gene of activated B cells. Only MMR proteins relevant to antibody affinity maturation will be discussed.

MMR proteins were initially discovered to affect SHM in studies of mice deficient in MutSα, a complex consisting of the factors MSH2 and MSH6 that are required for recognition of non- Watson-Crick base pairs. The majority of hypermutations at A/T base pairs in the activated B cells were reduced by ~85% (75, 76). MutSα recognizes the AID generated U•G base pair in the variable region and an endonuclease is thought to make an incision 5’ to the mismatch. This allows Exonuclease 1 (Exo-1) to degrade the DNA and generate a gap around the initial uracil lesion. Exonuclease 1 is responsible for generating a gap in the strand with the misincorporated base during normal MMR, and performs a

similar role in promoting A/T base pair mutations during SHM (77). Recent evidence suggests that in germinal center B cells, nicking of AP sites by APE2 may provide an accessible end for Exo-1 (78, 79). A large gap is required to avoid confining mutations to just the initial AID lesion and allows for more mutations around the original deaminated cytosine. Finally, as in UNG2 dependent mutations in SHM, it is TLS polymerases that are recruited and fill in this gap in an error-prone manner, creating transitions and transversions at A/T base pairs (see Figure 9).

The majority of A/T mutations in SHM are dependent on MMR recognition factors, but ~15% can still be found in the Ig genes of activated B cells of MutSα deficient mice. However, complete ablation of A/T mutations occurs in mice that are deficient in both UNG2 and MSH2 (69, 80). Similar results are seen in mice deficient in UNG2 and MSH6 (81). This supports a secondary role for UNG2 in generating a minority of the A/T transitions and transversions around the original uracil lesion and it is proposed that error-prone long-patch BER may be involved.

1.2.4 The Mechanism of Class Switch Recombination

AID is required to initiate CSR, but unlike SHM, the uracils generated are not processed to promote mutations. Instead, AID induced uracils during CSR are converted to double-strand breaks (DSBs), and then resolved to adjoin the variable segment of the Ig gene to a new constant domain, with the intervening sequence circularized and degraded. Similar to SHM, however, is that this deletion-recombination event depends on BER and MMR proteins to promote

genomic instability, contrary to their ubiquitous primary roles in protecting the genome.

1.2.4a Switch Regions and CSR

Downstream of the V(D)J segment of the Ig gene are the constant domains which code for the effector functions of the antibody (see Figure 5A and 6). Directly upstream of each constant domain, with the exception of Cδ, is a switch region (S). Each S region has an upstream ‘intronic’ or ‘intervening’ (I) exon with its own promoter. Since AID requires transcription, the I exon and its associated promoter ensure transcription through the S region and thus allow for targeting of the S region by AID (82). The induction of transcription of different I exon promoters are cytokine or cytokine cocktail specific to stimulate class switching to different constant domains, and thus effector functions, as needed. For example, stimulation by interleukin-4 (IL-4) can induce germline transcription of the Iε, ensuring access of AID to the S region directly upstream of Cε, and thus allow for switching from the default IgM to IgE (83, 84). This transcription produces a germline transcript (GLT) which is considered a ‘sterile’ transcript, as it lacks an open reading frame and is not translated (82). GLTs are essential to CSR (85).

As targets of AID, it is not surprising that S regions contain frequent AID hotspot motifs (ie WRC). For example, the core of the switch region contains 5´-AGCT-3´ repeats and this motif is highly enriched in the S regions compared to the rest of the genome (86). Furthermore 5´-AGCT-3´ sequences ensure a WRC AID deamination motif in both DNA strands (87). S regions are also characteristic

of containing a NTS enriched with G nucleotides. These clusters of G nucleotides promote the formation of R-loops where the transcribed mRNA strand hybridizes with the TS of DNA. This causes displacement of the NTS and allows a single stranded DNA loop of potentially hundreds of nucleotides to form (88). Thus, WRC motifs, R-loops, and GLT ensure S regions are a welcoming DNA substrate for AID.

1.2.4b AID Promotion of Double-Strand Breaks

During CSR, AID introduces uracils in the Sμ region as well as the downstream S region ($S_{downstream}$) that has been upregulated for transcription upon response to stimulation with specific cytokines. In class switching to IgE, for example, AID introduces uracils in the downstream Sε region, but effective CSR depends on a complex of proteins forming at the S regions. Particularly, AID is recruited to the 5′-AGCT-3′ enriched S regions by 14-3-3 adaptor proteins to allow for deamination of cytosines. The TLS polymerase Rev1 recruits UNG2 to excise AID generated uracils. Evidence supports Rev1 plays only a scaffolding role, bringing UNG2 to the S regions, since CSR is independent of Rev1 enzymatic activity (89). 14-3-3 adaptor proteins have also been shown to directly interact with and stabilize UNG2 at the S region (90). The AP sites, resulting from UNG2-excision of uracils, are nicked by APE1 and form SSBs (91). In addition to APE1, lyase activity of MRE11–RAD50 may also process and cleave AP sites (92).

Since AID activity is not strand biased, and WRC motifs occur in close proximity on both strands, AP sites and therefore nicks are generated in both the TS and NTS promoting direct DSB formation in the S regions. MMR proteins, including MutSα and Exo-1, also help to produce DSBs in the flanking regions. The AID dependent DSBs created in both Sμ and $S_{downstream}$ are resolved by classical non-homologous end joining (c-NHEJ) or alternative-end joining (A-EJ) (84). Finally the intervening sequence is excised and circularized. This 'switch circle' is eventually degraded. Thus, the V(D)J segment is juxtaposed to the Sμ-$S_{downstream}$ junction and allows for class switching to the new constant domain (see Figure 6 for a schematic of IgM to IgE class switching).

1.3 The Role of AID in Lymphomagenesis

AID is a DNA mutator enzyme and its ectopic expression may play a role in the transformation of tissue from hepatocytic, gastric, mammary, and oral epithelial origin (93-97). Despite the tight regulation of subcellular localization, Ig gene targeting, and restricted B cell spatiotemporal expression, it is not unexpected that AID is proposed to have a role in the development of germinal center (GC) B cell lymphomas. The GC microenvironment promotes AID dependent mutations and DSBs in the Ig gene. However, aberrant SHM or CSR at non-Ig genes can be deleterious to the activated B cells leading to cancer.

1.3.1 The Germinal Center Reaction

AID is primarily targeted to the Ig genes of B cells and is found highly expressed in B cells exposed to antigen and undergoing antibody maturation (2, 3). The upregulation of AID and antibody maturation occurs in the germinal centers (GC) which are temporary microenvironment structures essential to the immune response that develop upon exposure to antigen. They are located in secondary lymphoid tissues such as the spleen and lymph nodes (84, 98, 99). GCs foster an environment of collaboration between specialized immune cells with the primary function of generating B cells with high-affinity antibodies. Thus the "germinal center reaction" (GC-reaction) requires upregulation of AID and the processes of SHM and CSR (98).

Figure 13: The Germinal Center Reaction

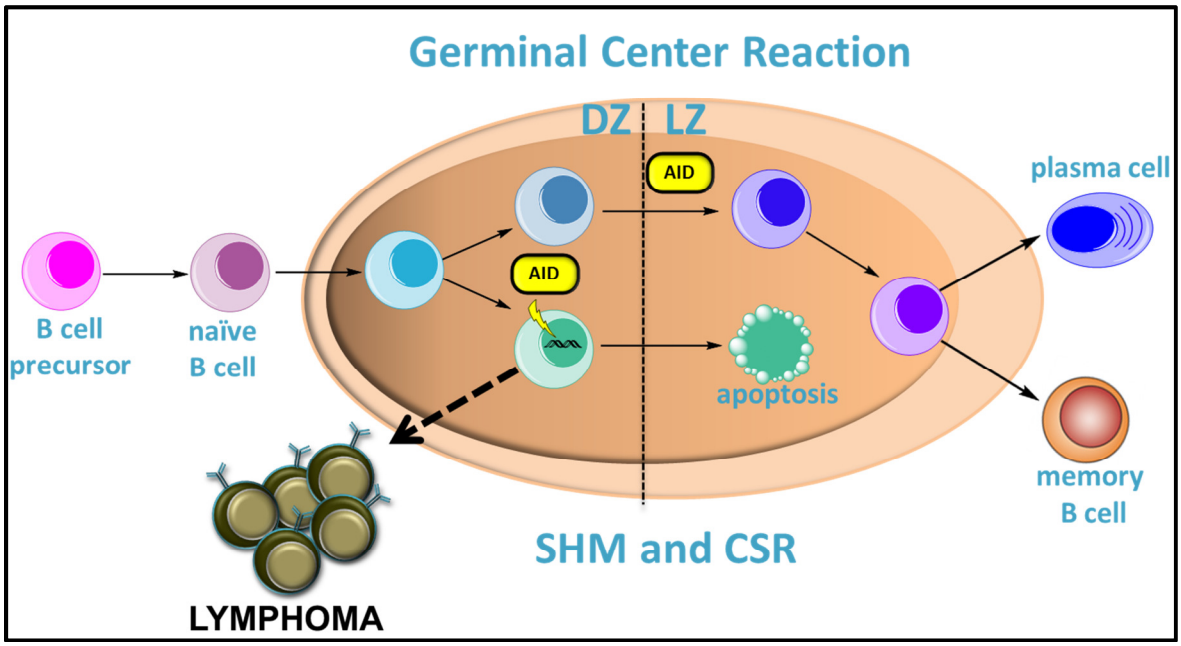

Figure 13: The Germinal Center Reaction. Post V(D)J recombination, precursor B cells develop into naïve B cells that have not been exposed to antigen. After exposure to antigen, germinal centers are formed in secondary lymphoid structures such as the spleen and lymph nodes, and have two

histologically identifiable regions, the ‘dark zone’ (DZ) and the ‘light zone’ (LZ). AID-dependent SHM and CSR occur in the GC-reaction. This leads to mutations and class switching in order to form high affinity antibody producing B cells. These will differentiate into plasma cells and memory B cells. However, not all mutations accumulated during the GC-reaction are beneficial and thus, these B cells undergo apoptosis as depicted above. Some AID-dependent mutations or translocations acquired during the GC-reaction may be deleterious ( ) and promote lymphomagenesis.

Prior to antigen exposure, the only diversification event that occurs in antibodies is V(D)J recombination in B cells early in development (pro- and pre-B cells). This takes place in the primary lymphoid tissue of the bone marrow. V(D)J recombination alone does not result in high affinity antibodies. Before GC development, secondary lymphoid organs are found to contain primary follicles which consist of predominantly naïve B cells that have not been subjected to interaction with antigen. Post antigen-exposure, however, GCs begin to form within the primary follicles and thus, force the naïve B cell population to the periphery of the follicle and form the region termed the B cell mantle (36). Early GCs are separated into ‘dark zones’ (DZ) and ‘light zones’ (LZ) which are histologically visible divisions (see Figure 13 for a schematic overview) (99, 100). The DZ primarily contains B cells exposed to antigen, ‘activated’ B cells, that express high levels of AID (or centroblasts) and undergo SHM and CSR (36, 101) while the LZ has antigen-selected GC B cells (or centrocytes) which also express AID and can undergo CSR. LZs also contain specialized cells which participate in the presentation of antigens to B cells and their development. These include follicular dendritic cells (FDCs) CD4+ and CD8+ T cells and tingible-body macrophages (TMϕs). TMϕs are a specific type of macrophage that

phagocytose B cells undergoing apoptosis during clonal selection in the GC reaction (102, 103).

Once GC B cells have acquired antibodies with high affinity against the antigen, they can exit the GC and differentiate into antibody secreting cells known as plasma cells to help eradicate the current infection. Other GC exiting B cells will differentiate into memory B cells to provide for a more rapid and long-term immunological response against a second future exposure to the antigen.

The GC reaction is essential to the adaptive immune system. However, this process of promoting AID-induced mutations and strand breaks in DNA for antibody affinity maturation bears considerable risks in the form of inadvertent oncogenic mutations and translocations.

1.3.2 Features of B Cell Lymphomas

The translocation of the Ig gene with a proto-oncogene is characteristic of many types of B-cell lymphomas. This deleterious chromosomal rearrangement may result in the constitutive expression of the oncogene as it comes under the transcription regulation of the Ig gene due to its juxtaposition (104-106). Studies have demonstrated that AID is essential for the formation of the c-Myc-IgH translocation found in specific B cell lymphomas (107, 108). Aberrant CSR is thought to be the culprit, where DSBs in the Ig are incorrectly adjoined to DSBs in the c-Myc oncogene from mis-targeting of AID, resulting in a reciprocal translocation. The transformative potential of Ig-oncogene translocations in B cell lymphomas has long been ascertained, pre-dating the discovery of AID (109). Aberrant SHM has also been implicated in oncogenic translocations with the Ig

gene but more importantly in promoting detrimental hypermutations. These mutations in non Ig genes, such as those that regulate B cell development, regulation, and survival, may promote B cell lymphomagenesis (110-112). Studies in transgenic mouse models have elucidated the requirement of AID for the development of GC-derived B cell lymphomas, but not lymphomas that develop from pre-GC B cells (113). The majority of human B cell lymphomas are thought to originate from a transformation event that occurred in the GC microenvironment. This is hardly surprising given the mutagenic potential of AID in GC B cells where DNA damage in the form of mutations and DSBs is encouraged, and cells are highly and rapidly proliferating. TABLE 2 summarizes the human B cell lymphomas relevant to this work.

1.3.2a Burkitt Lymphoma

AID is required for the t(8;14) translocation which involves the c-Myc gene on chromosome 8 and the IgH locus on chromosome 14 (see Figure 14) (107, 108). Myc is a transcription factor thought to regulate approximately 15% of genes in B cells (114). These include genes responsible for cell cycle regulation, apoptosis, and proliferation. Thus, deregulation and constitutive overexpression of Myc can lead to compromised cell cycle control, reduction in necessary programmed apoptosis, and uncontrolled rapid proliferation promoting genome instability and tumorigenesis (115). The t(8;14) translocation is a hallmark of human Burkitt Lymphoma (BL). BL is categorized into two subtypes, the endemic form which is predominantly found in young children in tropical Africa and the sporadic form found throughout the world at a much lower prevalence. Endemic

BL is associated with malaria and Epstein-Barr virus (EBV) while the majority of the sporadic form is not (116, 117).

The c-MYC-IgH translocation in sporadic BL occurs due to breakpoints in the S region of the IgH gene, and thus arises due to aberrant CSR (106, 118). However, due to breakpoint locations near the V segment of the Ig gene in endemic BL, translocations found in this subtype are thought to arise from aberrant SHM (119).Characteristic of BL is the expression of surface markers normally expressed on GC B cells. In addition, hypermutated V regions are detected which indicates the cells have previously undergone SHM. This supports the notion that BL originates from GC B cells that have acquired transformation during the GC reaction (120). Studies have revealed that BL cells are hypermutated at non-Ig genes, are constitutively undergoing acquisition of hypermutations, and have undergone antigen selection (121, 122).

Figure 14: Schematic of c-MYC-Ig Translocation

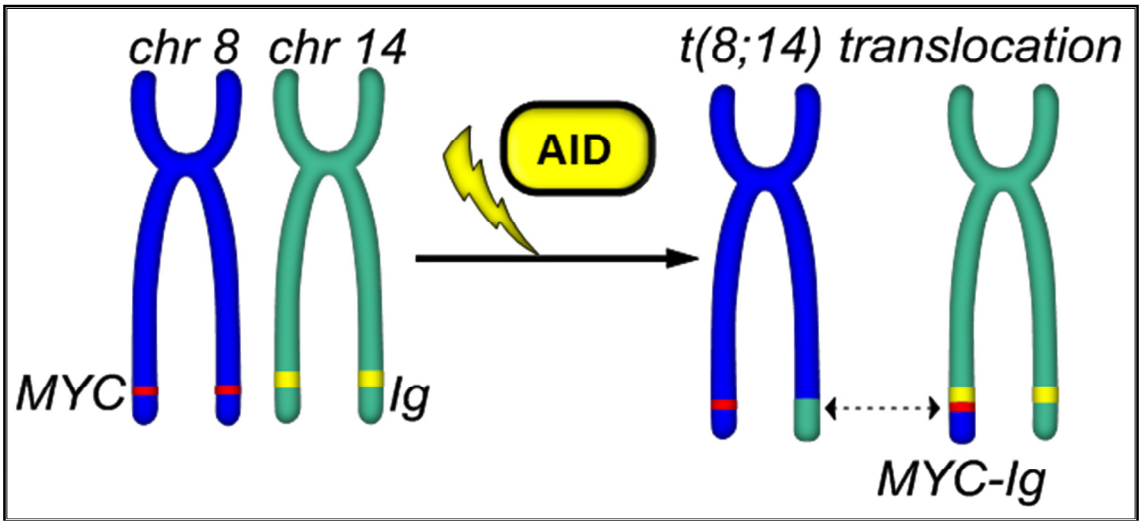

Figure 14: Schematic of c-Myc-Ig Translocation. The Myc gene (red) is located on human chromosome 8 (blue) and the immunoglobulin gene (Ig, yellow) is located on chromosome 14 (green). Upon deleterious AID-dependent double strand breaks (DSBs) during the GC-reaction and faulty resolution, a reciprocal Myc-Ig translocation results where Myc is juxtaposed to Ig (t(8;14) or Myc-Ig in the figure). This translocation is a hallmark of Burkitt Lymphoma and many lymphomas are found to have translocations between Ig and proto-oncogenes. Figure created in ChemBioDraw Ultra 14.0

TABLE 2: Human B Cell Lymphomas

Lymphoma **(Percent of Western Lymphomas)**	**B Cell Type Origin**	**Association with AID?**	**Translocations & Chromosomal Aberrations**	**References**
Diffuse Large B Cell (DLBCL) (30-40%)	**GC B cell or Post-GC B cell (plasma)**	Constitutive AID expression; Aberrant CSR and SHM	MYC-IgH, IgK-MYC, MYC-IgL, Bcl-2-IgH, Bcl-6-MYC, Bcl-2-IgH; ongoing SHM in many proto-oncogenes	(106, 123-126)
Follicular Lymphoma (FL) (20%)	**GC B cell**	Aberrant SHM and V(D)J recombination; many FLs constitutively express AID	Bcl-2-IgH, constitutive SHM of the IgH-V region & proto-oncognes in many patients	(127, 128)
Burkitt Lymphoma (BL) (2%)	**GC B cell**	Constitutive AID expression, constitutive GC phenotype and SHM	MYC-IgH or MYC-IgL Mutations in TP53 and RB2	(129-132)
B- Cell Chronic Lymphocytic Leukemia (CLL) (7%)	**Pre-GC B cell, GC- B cell, or Post-GC B cell** (ie. naïve, GC, or memory B cells)	Heterogeneity in AID expression- AID is expressed at higher levels in more aggressive forms of CLL with an unmutated IgH-V region; aberrant SHM and constitutive aberrant CSR	Trisomy 12; mutations in TP53/ATM/ NOTCH1/SF3-1; deletions in 17p,13q,11q	(124, 132-135)

TABLE 2: Human B Cell Lymphomas (continued)

Lymphoma (Percent of Western Lymphomas)	**B Cell Type Origin**	**Association with AID?**	**Translocations & Chromosomal Aberrations**	**References**
Marginal Zone Lymphoma (MZL) (2%)	**Marginal zone B cell**	No AID expression	No characteristic translocations Deletion in 7q	(124)
Classical Hodgkin's Lymphoma (10%)	**Defective GC B cells, T cells?** ***DO NOT retain GC like traits***	AID generally is not expressed	No characteristic translocations Mutations in several tumor suppressor genes	(124, 136, 137)
Multiple Myeloma (10%)	**Post-GC B cell (plasma)**	AID generally is not expressed Aberrant CSR thought to occur during transformation event	MMSET-IgH FGFR3-IgH CCND3-IgH CCND31-IgH MAF-B-IgH	(138-140)

TABLE 2: Cell type origin, association with AID, and frequent chromosomal aberrations for relevant human B cell lymphomas are indicated.

1.3.2b Diffuse-Large B Cell Lymphoma

Diffuse-large B cell lymphoma (DLBCL) is the most prevalent type of non-Hodgkin's Lymphoma, accounting for approximately 30-40% of lymphoma cases (141). DLBCL is classified into subgroups based on the phenotype of the lymphoma, expression profiles, and the transformation pathways thought to give rise to the cancer (142). The germinal center B cell-like (GCB) DLBCL has a gene expression profile similar to normal GC B cells and also expresses high levels of Bcl-6 which is generally found in centroblasts undergoing the GC reaction (143). The activated B cell-like (ABC) DLBCL is thought to arise from GC B cells differentiating into plasma B cells due to the down regulation of Bcl-6 expression in these lymphomas and the upregulation of several plasma cell characteristic genes including XBP-1 (142, 144).

Like with BL, DLBCLs also have characteristic translocations but these may involve the heavy (IgH) or light Ig chains (IgL or IgK). For example, the c-MYC-IgH [t(8;14)], IgK-MYC [t(2;8)], and MYC-IgL [t(8;22)] translocations are diagnostic chromosomal abnormalities in DLBCL patients (145). Translocations involving IgH and different oncogenes also are characteristic of DLBCLs, including Bcl-6-IgH [t(3;14)] and Bcl-2-IgH [t(18;14)]. Bcl-6 can also be found to have other non-Ig gene partners in DLBCL translocations including Myc as in the [t(8;30)] (126, 146). Translocations in both ABC-DLBCL and GCB-DLBCL are proposed to occur due to illegitimate CSR (147). DLBCL is also known to continually undergo aberrant SHM and acquire hypermutations in many non-Ig genes throughout the genome due to overexpression of AID (148).

1.3.2c Follicular Lymphoma

Follicular lymphomas are also thought to arise from B cells that have undergone extensive SHM of the V region during the GC reaction, and account for ~20% of diagnosed lymphomas (124, 149). Classification of FL is divided into four grades based on the type of predominant GC malignant B cells that populate the lymphoma. Grade 1 (G1-FL) consists of a majority of cancerous centrocytes, G2-FL has both centrocytes and centroblasts, G3A-FL has a minor population of centrocytes with a majority of centroblasts, and G3B-FL consist of only malignant centroblasts (149). The majority of FLs carry a Bcl-2-IgH translocation, t(14;18), which results in the overexpression of Bcl-2 that promotes continued cell survival and blocks apoptosis (150, 151).

AID continues to be overexpressed in FLs of all grade types and is implicated in introducing hypermutations in several proto-oncogenes such as c-Myc, Pax-5, and Bcl-6 (128, 129). Recent sequencing studies of FL patient samples have revealed that AID and aberrant SHM play a role in the evolution and multiple recurrence of FL. AID continues to introduce mutations in Ig and non-Ig genes shaping both genetic and epigenetic co-evolution of the cancer (128, 152).

1.3.2d Chronic Lymphocytic Leukemia

B-cell chronic lymphocytic leukemia (CLL) is the most prevalent form of leukemia but is an extremely heterogeneous cancer which results in variable clinical development (153). CLL falls into two essential prognostic groups that consist of either having an unmutated IgH V region or a V region that has undergone SHM. Patients with a mutated V(D)J segment have a better prognosis with a survival median three fold higher than CLL patients with leukemia that did not undergo SHM (154, 155). The B cell type that CLL arises from is debated. Pre-GC naïve B cells, post-GC memory B cells, and GC- B cells have all been suggested as the origin (124, 156). AID expression in CLL is also quite variable ranging from low AID to extremely high AID expression, if expressed at all (156). CLL patients can acquire mutations in DNA damage response pathways including ATM, as well as tumor suppressor genes such as TP53 (132, 157). However IgH translocations are not typically found in CLL.

1.3.2e Non-GC Derived B Cell Lymphomas

The majority of B cell lymphomas do arise from B cells undergoing the GC reaction, however, several types of B cell lymphomas thought to originate from pre- or post-GC B cells exist. From these two groups, however, more arise from post-GC, having at least experienced the GC reaction before transformation, than pre-GC B cells. For example, primary effusion lymphoma, multiple myeloma, lymphoplasmacytic lymphoma, and hairy-cell leukemia arise from plasma B cells or memory B cells (124).

There are also some lymphomas whose origins are still yet not entirely understood. Marginal zone lymphoma (MZL) is thought to arise from either pre-GC or post-GC B cells. There are two groups of MZL characterized by the site of infiltration. Nodal-MZL predominantly presents in the lymph nodes while splenic-MZL infiltrates the spleen and bone marrow. MZL does not carry characteristic translocations but approximately 40% of MZLs have chromosomal deletions around 7q22-36 (158). Similarly, the origin of classical Hodgkin's lymphoma (cHL) has also been debated, but most likely arises from defective GC B cells (124). Hodgkin and Reed-Sternberg cells (HRS) are the cancerous cells in cHL. These cells do not retain GC B cell expression profiles or phenotypes and a minority of cases are thought to have a T cell origin (159). However non-beneficial and deleterious mutations in the Ig genes of some cHL patients suggest the lymphoma may have derived from the transformation of cells destined to undergo apoptosis during the GC reaction (160). cHL cells have mutations in several tumor suppressor genes and can also be associated with EBV infection, but do not have IgH translocations.

1.4 **Scope and Significance**

An overwhelming majority of human B lymphocyte malignancies derive from cells that have undergone the GC- reaction and are associated with the expression of the mutator enzyme AID. Thus, there is a critical need to elucidate the global damage AID inflicts on the genome of normal and transformed cells. Studying AID in the context of cancer is important to our understanding of how mutations and translocations that promote B-cell lymphomas may arise. This is a particularly interesting source of DNA damage because AID's normal role in adaptive immunity requires the deliberate introduction of uracils in the genome but its mistargeting, ectopic expression, and aberrant behavior can promote genomic instability. Furthermore, unlike most sources of DNA damage that promote transformation, for example UV light and ionizing radiation, AID is an endogenous enzymatic source of mutations. Recently, other members of the APOBEC family have also been implicated in several other non-B cell cancers that carry AID/APOBEC mutation "signatures" determined through base substitution sequence analysis of the cancer genomes (161, 162). These AID/APOBEC mutational signatures were found in cancers of the breast, bladder, cervix, and hematopoietic origins.

The goal of this research project is to examine the genomic uracil load introduced by AID both in normal activated B cells as well as in B cell lymphomas and leukemias from patients and cell lines. Since AID can act on non-Ig genes, one may predict an increase in the uracil load in the DNA of transformed lymphocytes versus normal B cells. Furthermore, most studies that have

analyzed the effects of AID on B cell cancers have only examined mutation or base substitutions, the indirect consequence of AID activity. Innovatively, the direct product of AID deamination, uracil, is studied and quantified here for the very first time in several B cell cancer and normal genomes. To achieve this goal, a very sensitive biochemical method that I have adapted and optimized was implemented in these studies. This uracil quantification assay uses *E. coli* UNG to remove the uracils in the genomes under study, resulting in abasic sites where the uracils once were located. Aldehyde-Reactive Probe (ARP) is then used to biotinylate these remaining abasic sites and a streptavidin conjugated to a fluorophore can then label the biotins. This allows for sensitive detection and quantification of uracil sites from the fluorescence signal intensities when a standard containing known uracil amounts is processed in parallel. To elucidate the difference between AID in cancer and AID in normal activated B-cells undergoing antibody maturation, it is important to examine the AID- generated uracils in both contexts.

This work also methodically and systematically analyzes the direct and indirect types of DNA damage and genomic instability that may be promoted by AID in B cell lymphoma cell lines. Since AID is an endogenous source of DNA damage unlike most mutagens and carcinogens, the types of damages analyzed here are those predicted to arise in the cell as a consequence of its attempt to repair uracils. For example, the base excision repair machinery of the cell will attempt to remove uracils in the genome through excision with proteins such as UNG. This will result in an abasic site that can be processed by APE which nicks

the DNA (see Figure 9). The study described here examines uracils that AID introduces into the genome, abasic sites that uracil excision proteins create, and strand breaks that may result from nicks created by APE. Findings that can correlate these types and amounts of damage with factors such as AID and uracil repair protein expression as well as lymphoma origin will be instrumental in understanding B-cell lymphoma development. In addition many forms of chemotherapy, such as the use of alkylating agents, aim at creating genomic instability. Therefore by methodically testing the types of DNA damage already present in B cell lymphomas due to endogenous factors, new combination therapy aimed at exacerbating forms of genomic instability detected in this study will be important from a clinical aspect. These types of DNA damage are important as the initial uracil and its subsequent repair may lead to mutations and translocations known to promote lymphoma development.

1.4.1 Rationale

This research provides new prospects for an early detection marker for B cell lymphomas. The quantification of uracils in the genomes of patient B cells may serve as an indicator of lymphoma development. In addition, knowledge of the specific types of DNA damage in lymphomas that accumulate due to AID activity can be used to develop new chemotherapeutic strategies that are aimed at exploiting and further intensifying these genomic instabilities with the use of specific DNA damaging agents. Patient samples from several lymphoma types were assessed for these reasons. Finally, an understanding of the extent of

genomic instability that AID creates may elucidate if AID should be a candidate target in treatment strategies.

1.5 Specific Aims of Research

AIM 1) The extent of global genomic uracil introduction by AID and the removal of uracils by repair enzymes in normal stimulated B cells are examined for the first time. Genomic uracils are quantified using a sensitive biochemical uracil detection assay to elucidate genome-wide AID activity in adaptive immunity.

Although the deamination of cytosine to uracils in single-stranded DNA by AID is accepted in the field, there are no previous studies directly quantifying uracils and demonstrating AID's full capacity to introduce uracils in the entire genome. The two previous studies of uracils introduced by AID have centered only on the Ig gene in murine B cells (163, 164) and used an indirect method for quantification, ligation mediated PCR (LM-PCR). Here, a sensitive uracil quantification assay is implemented to directly determine the global genomic uracil load that AID imposes on murine splenocytes and human tonsil B cell genomes. Murine B cells from transgenic mice deficient in uracil repair proteins are used to determine which repair enzyme is the primary contributor to the removal of genome wide uracils introduced by AID undergoing the GC reaction. This study has revealed that in normal B cells a homeostasis, dependent on UNG, keeps uracil levels in the genome from accumulating despite AID activity.

AIM 2) The AID-dependent uracils in B cell lymphomas and uracil removal capabilities are examined to determine differences in genomic uracil homeostasis in cancer cells where AID is constitutively and ectopically expressed.

The genomic uracils in B cell lymphomas are quantified using a sensitive biochemical uracil detection system to determine the full extent of genome-wide deamination by AID. This study has identified differences that exist in the total genomic uracil load of normal activated B cells with up regulated AID levels required for antibody maturation, contrasted with cancerous cells where AID is ectopically or constitutively expressed. The genomes of several cancerous samples and tissues including patient lymphomas were assessed and compared to normal tissues. In addition, the expressions of AID and uracil repair proteins such as UNG in lymphomas of GC- or post-GC origins were examined. The results suggest that the homeostasis between AID- created uracils and UNG-mediated excision seen in normal activated B-cells (AIM 1) is disrupted in transformed cells that continue to overexpress AID.

AIM 3) Multiple types of DNA-damage that accumulate as a direct or indirect consequence of the constitutive expression of the endogenous mutator enzyme, AID, are identified and analyzed in B-cell lymphomas.

The mutagenic potential of uracil in U•G mispairs has been exploited as an intermediate lesion in the adaptive immunity of mammals. The purposeful introduction of uracils in the Ig gene by AID produces mutations and strand breaks required for efficient antibody affinity maturation in B-lymphocytes. However, when AID is ectopically expressed or aberrantly targets non-Ig genes, several downstream pathways that attempt to repair these uracils can in fact lead to further genomic instability. Unlike in the Ig-gene where point mutations are welcomed for SHM to be effective, and double-strand breaks are necessary for

CSR to generate appropriate isotypes, point mutations and strand breaks in other genes may be detrimental to the cell or lead to transformation events. If AID is constitutively overexpressed in B-cell lymphomas and is able to mutate the genome, multiple types of damage may accumulate. This is essential to our understanding of lymphoma development and in effective treatment strategies because AID is an endogenous enzyme that is introducing the sources of DNA damage studied here. To characterize these direct and indirect types of AID-dependent damage to the genome, several methods were applied to test the genomes of B-cell lymphomas for viability, uracil accumulation, AP site acquisition, and single- and double-strand breaks. Furthermore, this work led to novel findings concerning the genomes of inviable lymphoma cells.

CHAPTER 2: EXPERIMENTAL PROCEDURES

Portions of the text in this chapter were reprinted or adapted in compliance with the *ASM Journals Statement of Authors' Rights* from:

Shalhout S, Haddad D, Sosin A, Holland TC, Al-Katib A, Martin A, Bhagwat AS. *Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development*. **Molecular and cellular biology**. 2014;34(21):4019-32. doi: 10.1128/MCB.00589-14.

2.1 Overview of Materials and Methods

In order to study several aspects of the role of activation-induced deaminase (AID) in normal B cell maturation and lymphomas, several tissues were selected for this research including cell lines of germinal center-derived B cell lymphoma origin, mouse tissues such as wildtype and transgenic $AID^{-/-}$ mice, and several patient samples spanning different types of B cell cancers . All protocols involving the use of animals and human tissue materials were subject to Institutional Animal Care and Use Committee (IACUC) and Human Investigation Committee (HIC)- Institutional Review Board (IRB) approval. To accomplish the aims in this research, a sensitive biochemical genomic uracil detection method is used in this research and described in detail in this section. With this powerful tool, the genomic uracil levels in several tissues were examined and compared to AID activity and mRNA levels, and uracil repair enzyme activity and mRNA levels (see Figure 15 for an overview).

Figure 15: Flow chart of the Tissues Assayed

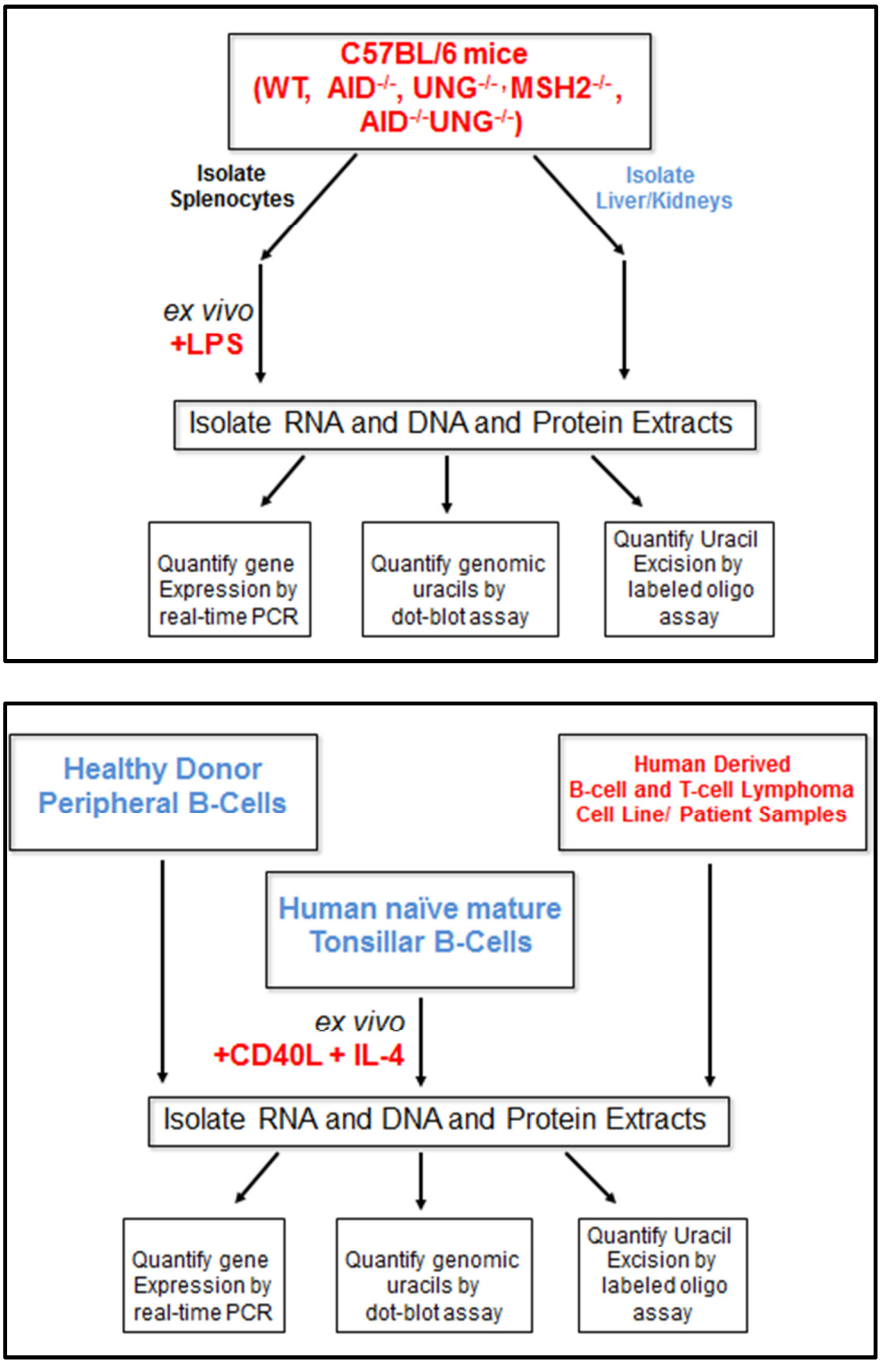

Figure 15: Flow chart of the Tissues Assayed. Messenger RNA, genomic DNA, and nuclear and whole cell protein extracts are purified from mouse tissues (Top panel), and human lymphoma patient samples, tonsils, normal B lymphocytes and cell lines (Bottom panel) and tested as described throughout Section 2.1.

To systematically assess the types of DNA-damage that accumulate in lymphomas directly or indirectly due to AID, specific types of DNA damage were assessed based on the subsequent *in vivo* processing of uracils introduced by AID. Cell lines were assayed for a.) viability in culture, b.) AID expression, c.) genomic uracil content, d.) abasic site accumulation, e.) single- strand and double- strand DNA breaks. This was carried out utilizing an array of sensitive techniques described here. Several molecular biology techniques are employed in this research including quantitative real-time PCR, nuclear and whole cell activity assays, non-liposomal transfections, transductions, electroporation, and short-hairpin knock down of genes to study factors such as AID expression and catalytic activity.

2.1.2 Tissue Samples and Materials for Study

2.1.2a Mouse Genetic Backgrounds and Tissue Harvest

Mice colonies were maintained at the University of Toronto and at Wayne State University. Experiments involving animals were performed at the University of Toronto and at Wayne State University and were approved by the Ethical Committee for Animal Experimentation at the University of Toronto (Toronto, Canada) and by the Institutional Animal Care and Use Committee (IACUC) at Wayne State University (Detroit, Michigan), respectively. Breeding pairs of AID-/- mice were kindly provided by Patricia Gearhart (National Institute on Aging) to generate our colony (165).

Spleen, liver, and kidney tissues were harvested from twelve to eighteen week old C57BL/6 background mice including wildtype mice (WT, control) and

the following transgenic mice: UNG-/-, AID-/-, and MSH2-/- mice. Single-cell suspensions of splenocytes, hepatocytes, and kidney cells were obtained by filtering homogenized tissue samples through a 70-μm nylon-mesh (Corning, Falcon). Kidney cells and hepatocytes were directly used for control studies. Splenocytes were cultured in RPMI-1640 media (HyClone) supplemented with 10% fetal bovine serum (HyClone) and 50μM β-mercaptoethanol (Sigma-Aldrich) for *ex vivo* stimulation time course studies (165).

AID-/-UNG-/- mice were generated at Wayne State University by crossing AID-/- female mice with UNG-/- male mice in a trio breeding format. Approximately thirty AID+/-UNG+/- mice (F1) were generated. F1 male mice were then bred with F1 female mice (trio breeding) which generated approximately 150 offspring (F2). Genotyping of F2 progeny was carried out with PCR and 1/25 of F2 mice were detected as AID/-UNG-/-. See TABLE 3 for PCR primers used for genotyping (165). Mice with undesired genotypes were donated to Division of Laboratory Animal Resources (DLAR) for training purposes as needed, or euthanized using approved IACUC protocols.

TABLE 3: Primers used to Genotype Mice

Gene	5' Primer	3' Primer
WT AID	AID-811 5'CTGAGATGGAACCCTAACCTCAGCC	AID-G4 5'CACGATTTTCTACAAATGTATTCCAGC
disrupted AID allele	AID-G3 5'GGGCCAGCTCATTCCTCCACTC	AID-G4 5'CACGATTTTCTACAAATGTATTCCAGC
WT UNG	Up 310 5'GCCCATCTTGGAAACTCAAA	Lower 18 (5'CCAGTCTGGCTTGGTTACCTTG)
disrupted UNG allele	Up 310 5'GCCCATCTTGGAAACTCAAA	Neo1524R 5'CGTCAAGAAGGCGATAGAA

TABLE 3: Primers used to Genotype Mice during generation of AID-/-UNG-/- double knockout colony.

2.1.2b Murine and Human Cell Lines

The murine cell line, J558, is a plasmacytoma cell line that was obtained from American Type Culture Collection (ATCC, Rockville, MD).The mouse B cell lymphoma cell line, CH12F3.2 and the heterozygous and homozygous AID knock-out lines, CH12F3.2 AID+/- and AID-/-, respectively, were kindly provided by Kefei Yu (Michigan State University).

Several human T- and B- lymphocyte cancer cell lines were obtained for these studies. The Epstein-Barr virus (EBV-) negative Burkitt-Lymphoma cell lines, Ramos 1 and Ramos 7 were previously referred to as Ramos clone 1-12 and Ramos clone 7-3 respectively and were kindly provided by A. Martin (University of Toronto)(166). Ramos-derived cell lines C.1, A.1, A.2, and A.5 have also been described (6). The human B cell follicular small cleaved cell

lymphoma line (WSU-FSCCL) (167) and the human diffuse large B cell lymphoma line (WSU-DLCL2) (168) were established in the laboratory of Dr. Ayad Al-Katib. The human Hodgkin lymphoma cell lines (KM-H2, L-428, L-591) and WSU-CLL were obtained from DSMZ (Germany). The Epstein-Barr virus positive (EBV +) Burkitt Lymphoma cell lines, Raji and Daudi, the human diffuse lymphoma cell lines, RL and Toledo, and acute T cell leukemia and lymphoblast-like cell lines (Jurkat and CEM, respectively) were obtained from ATCC. TABLE 4 summarizes the cell lines used in this work.

TABLE 4: Mammalian Cell Lines

Number	Name of Cell Line	Lymphoma/Leukemia Type	Source
1	Jurkat	Human Acute T Cell Leukemia	American Type Culture Collection (Rockville, MD)
2	CEM	Human Acute T Cell Lymphoblastic Leukemia	Leibniz-Institut DSMZ (Germany)
3	L591	Human Classical Hodgkin's Lymphoma	Leibniz-Institut DSMZ (Germany)
4	L428	Human Classical Hodgkin's Lymphoma	Leibniz-Institut DSMZ (Germany)
5	KMH2	Human Classical Hodgkin's Lymphoma	Leibniz-Institut DSMZ (Germany)
6	ARP1(169)	Human Multiple Myeloma	Wayne State University, Detroit, USA
7	FSCCL(167)	Human Follicular, Small Cleaved Lymphoma	Dr.Al-Katib, Wayne State University, Detroit, USA
8	WSU-NHL (170)	Human Follicular, Large Cell Lymphoma	Dr.Al-Katib, Wayne State University, Detroit, USA
9	DLCL2(168)	Human Diffuse Large B Cell Lymphoma	Dr.Al-Katib, Wayne State University, Detroit, USA
10	Toledo	Human Diffuse Large B Cell Lymphoma	American Type Culture Collection (Rockville, MD)
11	RL	Human Diffuse Large B Cell Lymphoma	American Type Culture Collection (Rockville, MD)
12	Raji	Human EBV+ Burkitt Lymphoma	American Type Culture Collection (Rockville, MD)
13	Daudi	Human EBV+ Burkitt Lymphoma	American Type Culture Collection (Rockville, MD)

TABLE 4: Mammalian Cell Lines (continued)

Number	Name of Cell Line	Lymphoma/Leukemia Type	Source
-	J558	Mouse plasmacytoma	American Type Culture Collection (Rockville, MD)
-	CH12F3.2	Mouse B-cell Lymphoma	Kefei Yu (Michigan State University)
-	Ramos	Human EBV- Burkitt Lymphoma	Alberto Martin (University of Toronto)

Numbers in red correspond to figure labels in chapter 3 for cell lines indicated. TABLE 4 was originally published in the supplementary material for (165) Shalhout et al. *Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development.* Molecular and cellular biology. 2014;34(21):4019-32. Adapted from the American Society for Microbiology.

2.1.2c Isolation of Peripheral B Cells from Healthy Donor Blood

Normal B cells were isolated from healthy American Red Cross (ARC) donors using peripheral blood collected in apheresis cones and filters. ARC donor blood was kindly provided by Dr. Martin Bluth (Associate Director of Detroit Medical Center Transfusion Services). Peripheral blood mononuclear cells (PBMCs) were first separated from the granulocyte and erythrocyte populations using Ficoll-Paque Premium density-gradient media (GE Healthcare, Stem Cell Technologies). B cells were initially isolated from the PBMCs using negative selection. A cocktail of biotinylated antibodies specific for surface markers present on non-B cells, including natural killer (NK) cells, T lymphocytes, dendritic cells, and monocytes was used for labeling the cells (human B cell Isolation Kit II, MACS, Miltenyi Biotec). The non-B cells were depleted from the

pool by using magnetically labeled anti-biotin beads. Highly pure B cell populations were confirmed with flow cytometry or purified as needed with FACS.

2.1.2d Cancer Patient Tissues

Ethical consideration for the use of patient samples in this research was approved following review by the Human Investigation Committee and the Institutional Review Board (IRB) at Wayne State University School of Medicine. Following informed consent, peripheral blood was collected from lymphoma patients during the leukemic phase under the clinical care of Dr. Ayad Al-Katib at St. John Hospital Van Elslander Cancer Center Lymphoma Clinic. Some patient B cells were purified from thigh lesion biopsies, lymph node biopsies, or pleural effusions. No identifying patient markers or information was retained. PBMCs were purified from patient whole blood using LymphoPrep density gradient media (ProGen Biotechnick GmbH, Germany). The total monocyte cell population was depleted from the PBMC suspension cells through adherence to the sterile plastic surface of tissue culture flasks. The non-adherent lymphocyte pool of cells was collected and B cells isolated through removal of T lymphocytes using magnetic bead separation with Dynabeads pan CD2 (Dynal, Life Technologies, Grand Island, NY). FACS analysis for each patient sample confirmed the recovery of ~95% B cells. These purified B cell populations were utilized in expression analysis and genomic uracil quantification studies (165). TABLE 5 summarizes the patient lymphoma samples, p53 status, and karyotypes.

TABLE 5: Lymphoma/Leukemia Patient Summary

Number	Age	Sex	Race	Type of Lymphoma/ Leukemia	Cytogenetics/ Karyotype	P53 Status
P1	81	F	Caucasian	CLL/SLL	58.5% deleted13q	WT
P2	76	M	Caucasian	CLL/SLL	84.5% trisomy 12; 15.5% deletion of 17p; 11%-loss of IgH locus	WT
P3	71	F	Caucasian	CLL/SLL	67.5% trisomy 12	WT
P4	64	F	African-American	CLL/SLL	86.5% deleted 13q	WT
P5	62	M	Caucasian	MZL	Normal/46,XY [20]	WT
P6	86	M	Caucasian	MZL	Normal/46, XY [20]	WT
P7	72	M	Caucasian	MZL	45% t(2;7) (p12;q21-22)	WT
P8	61	F	Caucasian	DLBCL	Normal/46, XX [20]	WT
P9	49	F	Caucasian	FL	23% t(14;18) (q32; q21.3)	WT
P10	87	F	Caucasian	FL	55% Complex chromosomal arrangements; t(3,11)(q27;q11); t(14;18)(q32;q21.3) + der(18); IgH-BCL2 fusion	WT

TABLE 5: Lymphoma/Leukemia Patient Summary (continued)

Number	Age	Sex	Race	Type of Lymphoma /Leukemia	Cytogenetics/ Karyotype	P53 Status
P11	64	F	Caucasian	CLL/SLL	54.5% deleted 13q	WT
P12	69	F	Caucasian	CLL/SLL	51% deleted 13q; 9% deleted 17p	Mt **(K132R)**
P13	62	M	Caucasian	CLL/SLL	23.5% deleted 13q; 26% 3 IgH copies	WT
P14	69	F	Caucasian	CLL/SLL	72% deleted 13q	WT

TABLE 5 was originally published in the supplementary material for (165) Shalhout et al. *Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development.* Molecular and cellular biology. 2014;34(21):4019-32. Adapted from the American Society for Microbiology.

2.1.2e Flow Cytometry for the Isolation of Human naïve mature B Cells from Healthy Tonsil Tissues

Healthy tonsil tissues were acquired as discarded surgical samples from patients with no identifying markers that underwent resection of hypertrophic tonsils at Children's Hospital of Michigan. Tonsils were approved for use in this research by the Institutional Review Board of Wayne State University. Tonsils were obtained with the assistance of Dr. Kang Chen (Barbara Ann Karmanos Cancer Institute and the Perinatology Research Branch, NICHD, NIH).

Tonsillar tissue was first filtered through a 70-μm nylon-mesh (Corning, Falcon) to yield a single-cell suspension in RPMI-1640 supplemented with gentamicin (HyClone). The PBMC fraction was collected using Ficoll-Paque

Premium density- gradient media (GE Healthcare, Stem Cell Technologies). Naïve mature B cells were purified by FACS with IgD-FITC (Southern Biotec 2032-02), CD19-PECy7 (eBioscience 25-0199-42), CD38-APC (Biolegend 303510), and CD27-PE (BD 555441) to yield IgD positive, CD19 positive, CD38 negative, and CD27 negative B lymphocytes. Dead cells were excluded during FACS collection by 7AAD (BD 559925) labeling. The purified naïve mature B cell population was cultured *ex vivo* in RPMI-1640 (HyClone) media supplemented with 10% FBS (HyClone) and 1% P/S at 37^{o}C and 5%CO_2. B cells were activated to express AID and undergo CSR by adding 500ng/ml of CD40-L (Peprotech) and 50ng/ml of IL-4 (Peprotech) and culturing for seven days. Class switching from IgM to IgG1 was confirmed with appropriate antibody labeling and FACS. (165).

2.1.3 Propagation of Cells

2.1.3a Cell Line Maintenance and Activation

J558 was cultured in DMEM (HyClone) supplemented with 10% fetal equine serum (FES, HyClone) and 1% P/S. All CH12F3.2 cells were cultured in suspension in RPMI-1640 media supplemented with 10% FBS, 50μM β-mercaptoethanol, and 1% P/S. CH12F3.2 cells were chosen for this study because the wildtype cell line expresses AID upon stimulation with a cocktail of cytokines and switches from IgM isotype to IgA (1). WT CH12F3.2 AID +/+ and the heterozygous and homozygous knock outs, AID +/- and AID -/-, were stimulated with CIT. Specifically cells were treated with 1 ug/mL of purified anti-mouse CD40 (eBiosciences), 10 ng/ml of recombinant mouse IL-4 (R&D

systems) and 1 ng/ml of recombinant human TGFβ1 (R&D Systems). Class switching to IgA was analyzed and confirmed by flow cytometry. All of the human T- and B- cell lines listed in TABLE 4 were cultured in suspension in RPMI-1640 media supplemented with 10% FBS and 1% P/S. Cells were incubated at 37^{o}C and 5% CO_2 in a humidified incubator in T-25cm^2 flasks, T-75cm^2 flasks, 96-well, 24-well, 12-well or 6-well sterile plates (Corning).

2.1.3b *Ex vivo* Stimulation of Splenic Murine B cells

In *ex vivo* stimulation time course studies, primary splenocytes from WT, AID-/-, UNG-/-, and MSH2-/- mice were treated with 25 µg/ml of lipopolysaccharide (LPS, *E. coli* serotype 055:B5; Sigma-Aldrich) to induce isotype switching of B cells to IgG3. Switching of primary B cells was confirmed with flow cytometry (165).

2.1.4 RNA Extraction and Quantitative Real-Time PCR (qRT-PCR)

Total RNA was extracted from human and murine primary B lymphocytes and cell lines using Trizol (Invitrogen) per manufacturer's protocols and instructions. Total RNA was reverse transcribed with anchored oligo-dT primers, specifically $d(T)_{23}VN$ where V is A, G, or C and N is A, G, C, or T and the ProtoScript M-MuLV First Strand cDNA Synthesis Kit (New England Biolabs). cDNA was amplified with HotStart Taq polymerase and RT^2 SYBR Green/ROX (SA Biosciences, Qiagen) to quantify relative gene expression. qRT-PCR was performed with Applied Biosystems 7500 Fast Real Time PCR System and the expression levels were determined using the delta Ct method with GAPDH or TBP as a reference gene.

TABLE 6: Primers for qRT-PCR

Gene	5' Primer Sequence	3' Primer Sequence
Murine AID	5'GAAAGTCACGCTGGAG	5'TCTCATGCCGTCCCTT
Murine UNG2	5'GTCTATCCGCCCCCGGAGCA	5'TGGGCGGGGGTGGAACT
Murine SMUG1	5'GAGTGCTGGGATTAAAAGGAT	5'CCAGAATGGGACAAGGATAG
Murine TDG	5'GTCTGTTCATGTCGGGGCTGAGTG	5'CTGCAGTTTCTGCACCAGGATG
Murine MBD4	5'TGGAGCTGCAGCTGCGCCCCA	5'CTGCGCTCACTGCTTATCACC
Murine GAPDH	5'ACCACAGTCCATGCCAT	5'TCCACCACCCTGTTGCTGTA
Human AID	5'AGAGGCGTGACAGTGCTACA	5'TGTAGCGGAGGAAGAGCAAT
Human UNG2	5'CCTCCTCAGCTCCAGGATGA	5'TCGCTTCCTGGCGGG
Human SMUG1	SA Biosciences (PPH02717A)	SA Biosciences (PPH02717A)
Human GAPDH	5'TGCATCCTGCACCACCAACT	5'CGCCTGCTTCACCACCTTC

TABLE 6 was originally published in the supplementary material for (165) Shalhout et al. *Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development.* Molecular and cellular biology. 2014;34(21):4019-32. Adapted from the American Society for Microbiology.

2.1.5 Trypan Blue Exclusion Assay

2.1.5a Percent Viability

Cell lines were first stained with Trypan Blue (HyClone) and counted using a TC20 Automated Counter (BioRad) to determine the percentage of viable cells. Inviable cells were removed from cell lines determined to have <95% viability using the Dead Cell Removal Kit (MACS, Miltenyi Biotec) to start each cell line at >95% viability at "Day 0." After three days of incubation, cells were stained with

Trypan Blue and counted to determine percent cell viability. This was repeated in six individual T-25cm^2 flasks per cell line tested.

2.1.5b Separation of Dead/Dying Cells from Viable Cells

To separately analyze the genomic uracil load in viable and inviable cell populations, dead/dying cells were isolated from viable cells using the Apoptotic Cell Isolation Kit (Promokine) which utilizes Annexin V magnetic beads that have a high-affinity for the phosphatidylserine that dead/dying cells redistribute to the outer surface of their plasma membrane. Genomic DNA was prepared from the separated populations for uracil quantification (see Figure 16 for schematic).

Figure 16: Isolating Inviable Cells from Viable Cells

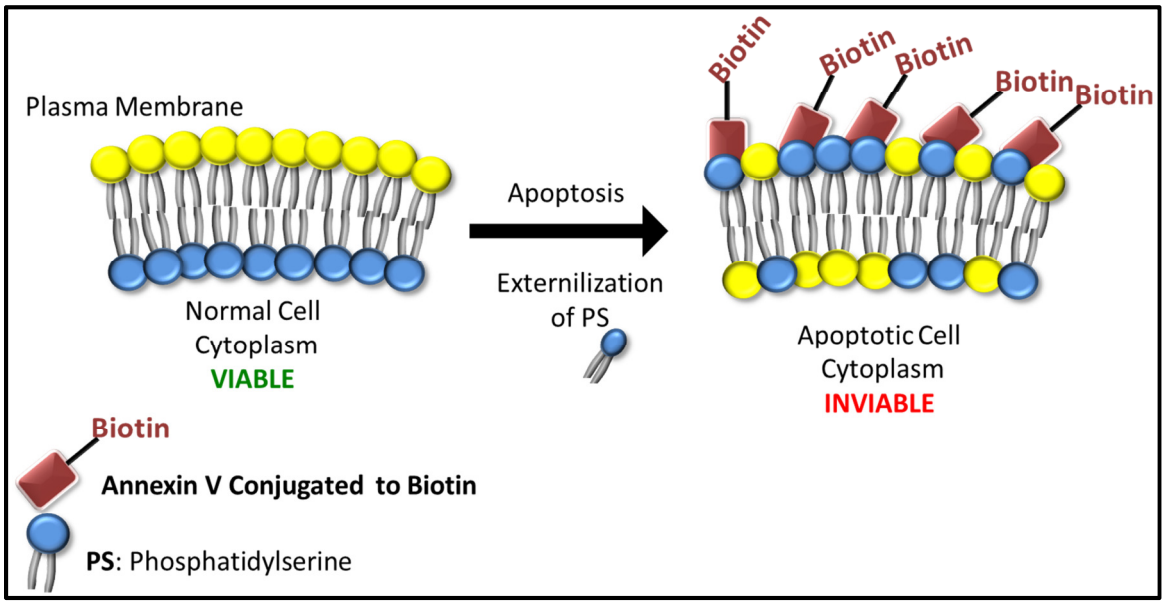

Figure 16: Isolating Inviable Cells from Viable Cells. Only dead/dying cells with surface phosphatidylserine are labeled with Annexin V conjugated to biotin. These cells are then isolated from the pool of cells with streptavidin-conjugated magnetic beads.

2.1.6 Quantification of Apurinic/Apyrimidinic Sites in DNA

Abasic sites, or AP sites, are quantified by adapting a previously described biochemical detection method (171). Genomic DNA is extracted from cells using the Blood and Cell Culture Kit (Qiagen) or standard DNA extraction protocols (172). High molecular weight DNA is then restriction digested with HaeIII (New England Biolabs). The endogenous abasic sites in DNA are treated with 2 mM aldehyde-reactive probe (ARP; Dojindo Laboratories, see Figure 17 and 18).

Figure 17: Aldehyde- Reactive Probe

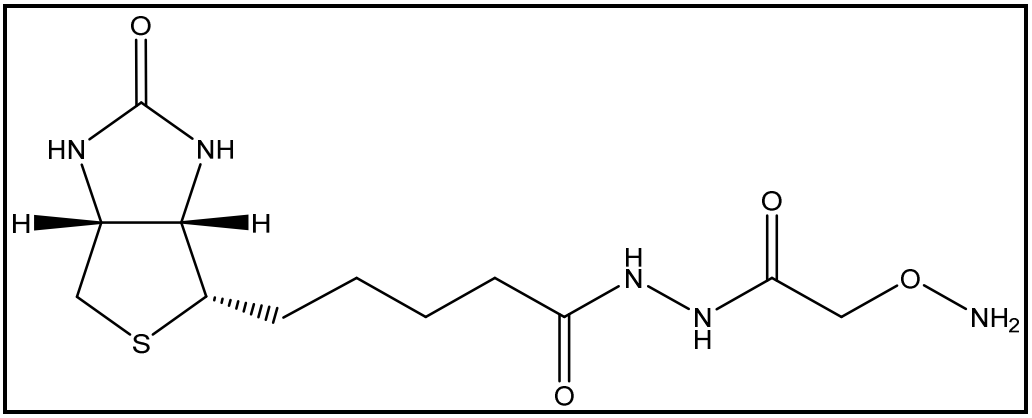

Figure 17: Aldehyde- Reactive Probe. The structure of Aldehyde- Reactive Probe (ARP) is depicted above. Figure created in ChemBioDraw Ultra 14.0

DNA is purified by ethanol precipitation, and filtering through a G-25 column (GE Healthcare) twice to remove residual reagents. DNA samples are then quantified using a NanoDrop ND-2000c (ThermoScientific), and with SYBRGold and a BioTek Synergy H1 Hybrid Microplate Reader or a Tecan Genios Microplate Reader where concentrations were determined based on a

standard calibration plot. Determined amounts of purified DNA samples are applied onto a positively charged Immobilon-Ny+ Nylon membrane (Millipore) using a vacuum-filtration apparatus (Bio-Rad Bio-Dot Vacuum Apparatus). The membrane is pre-equilibrated in double-distilled H2O and 6X SSC (saline-sodium citrate buffer) prior to DNA application.

Figure 18: Reaction of Aldehyde- Reactive Probe with AP Site

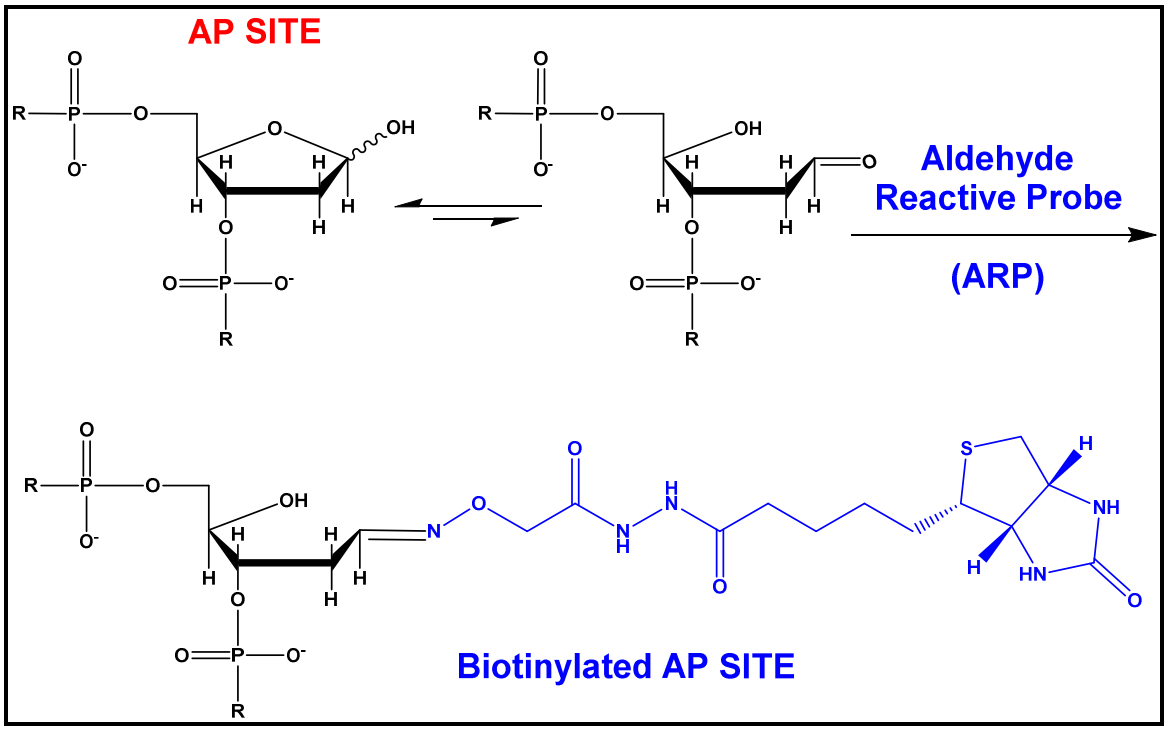

Figure 18: Reaction of Aldehyde- Reactive Probe with AP Site. The reaction of an abasic site with aldehyde-reactive probe (ARP) leads to the biotinylation of the site. Endogenous abasic sites are labeled this way in the quantification of AP sites assay and AP sites generated by the removal of uracils by UNG are biotinylated in this manner in the uracil quantification assay. ChemBioDraw Ultra 14.

After vacuum spotting the DNA onto the Ny+ membrane, it is incubated in StartingBlock buffer (Fisher Scientific) to block non-specific binding, followed by incubation in streptavidin-Cy5 (GE Healthcare) in the blocking buffer for one hour at room temperature. The membrane is then incubated three times in TBS-T (25

mM Tris, 3 mM KCl, 140 mM NaCl, 1% Tween-20) on a rocking platform to wash residual fluorophore buffer and scanned using a Typhoon 9210 Phosphorimager for Cy5 fluorescence. Cy5 fluorescence signal from samples and standard DNA are analyzed with ImageJ software (see Figure 19 for a schematic outline of the AP site quantification assay). To convert Cy5 fluorescence intensity signal to number of AP sites per million base pair, duplex DNA is used to generate a standard plot (see Section 2.1.7b Standard Plot).

Figure 19: Schematic Outline of AP Site Quantification in Genomic DNA

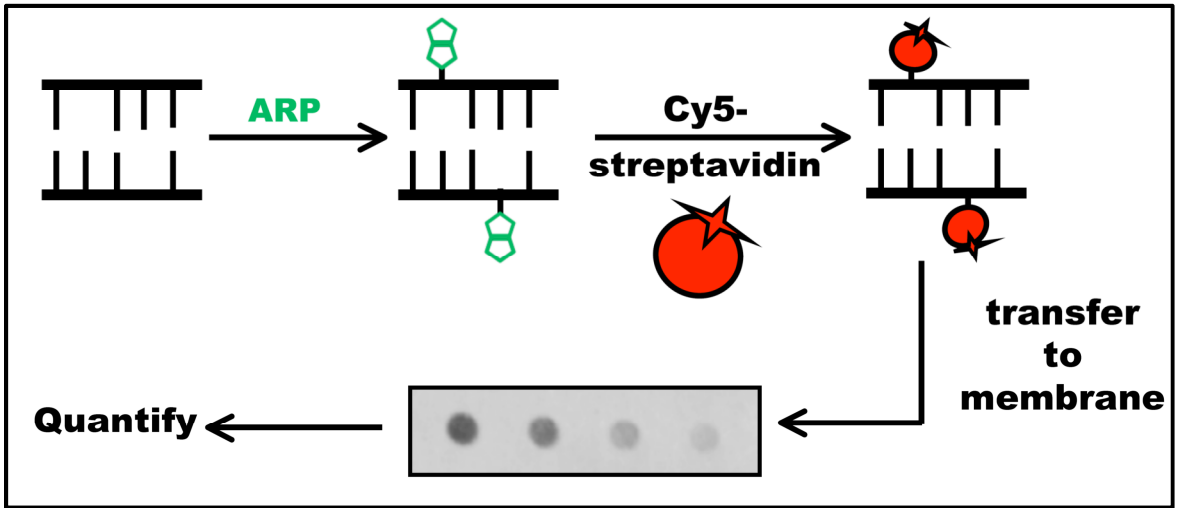

Figure 19. Schematic Outline of AP Site Quantification in Genomic DNA. Digested DNA is treated with Aldehyde-Reactive Probe (ARP, Dojindo) to biotinylate the persistent endogenous abasic sites. The DNA is vacuum transferred onto a nylon membrane and incubated with streptavidin-Cy5. The membrane is scanned to quantify fluorescence intensity and the signal is converted into number of abasic sites using a duplex standard described in 2.1.7b Standard Plot.

2.1.7 Quantification of Genomic Uracils

2.1.7a Total Uracils in DNA

To quantify the total uracils in the genomic DNA of mouse tissues, patient samples, and cell lines, a modified ARP-detection method of a previously described biochemical procedure is used (173-175). HaeIII (New England Biolabs) restriction enzyme digested DNA was treated with 5 mM methoxyamine (Mx, Fisher Scientific). This reacts with endogenous AP sites to block their labeling in the subsequent ARP treatment step (see Figure 20). To remove unbound or excess Mx, DNA is ethanol precipitated and filtered through a G-25 column (GE Healthcare).

Figure 20: Reaction of Methoxyamine with AP Site

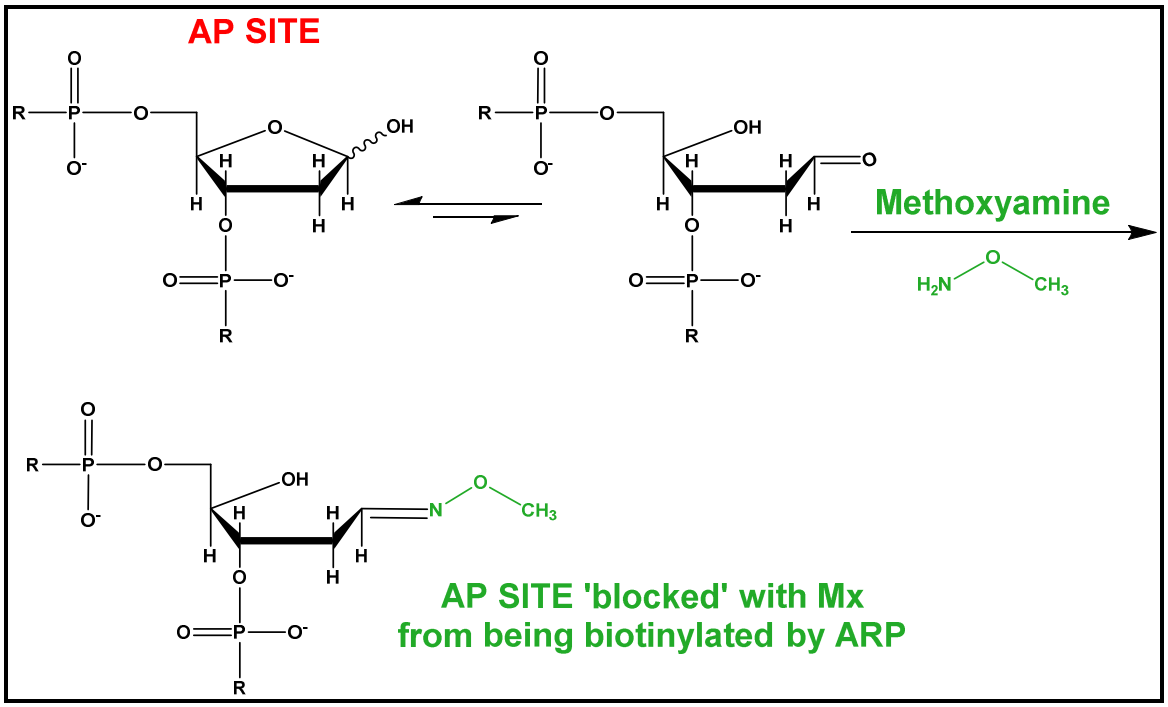

Figure 20: Reaction of Methoxyamine with AP Site. The reaction of an abasic site with methoxyamine (MX) protects the AP site from later reacting with ARP treatment and being biotinylated in the uracil quantification assay. All endogenous abasic sites are blocked this way in the ensuring ARP labeled abasic sites are only those that result from the excision of uracil by UNG. Figure created in ChemBioDraw Ultra 14.0

The DNA is then treated with *E. coli* UNG (New England Biolabs) and 2 mM ARP for one hour at 37°C. UNG removes uracils in U:A base pairs, U•G mispairs and uracils in single-stranded DNA (see TABLE 1). This allows for total removal of genomic uracil and ARP labels the resulting AP sites that remain after uracil excision with biotin (see Figure 18). This procedure ensures that detected abasic sites are those resulting from uracil excision. The DNA is then extracted with phenol: chloroform to remove the UNG enzyme, ethanol precipitated, and passed over G-25 columns (GE Healthcare) similar to the AP site detection

method. The DNA samples are quantified and spotted onto a Nylon membrane and the Cy5 fluorescence detection procedure is identical to that described above in 2.1.6 Quantification of Apurinic/Apyrimidinic (AP) Sites in DNA. See Figure 21 below for a schematic outline of the total genomic uracil quantification assay (165).

Figure 21. Schematic Outline of the Genomic Uracil Quantification Assay.

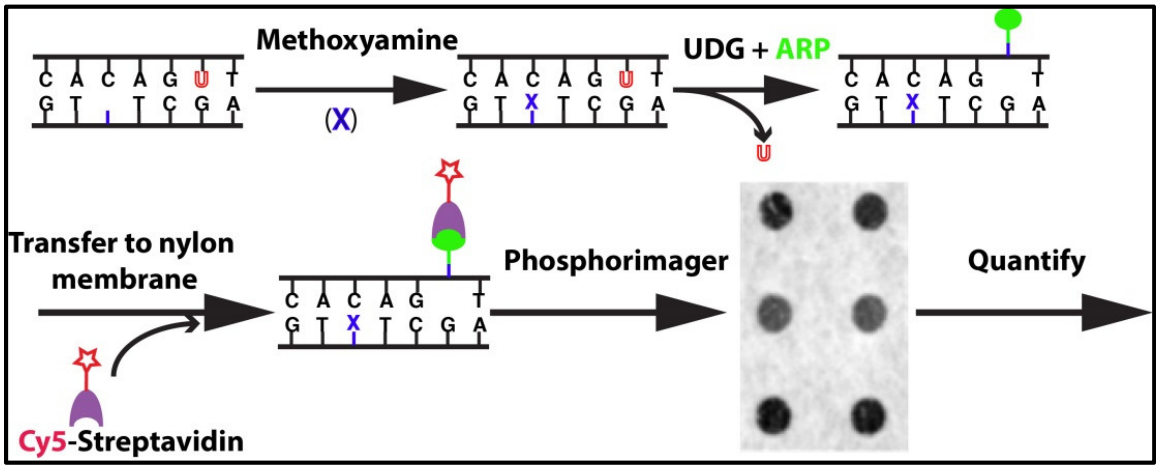

Figure 21. Schematic Outline of the Genomic Uracil Quantification Assay. Digested DNA is treated with methoxyamine to block pre-existing abasic sites. *E. coli* UNG is used to excise uracils and treatment with ARP biotinylates the resulting AP sites. The DNA samples are vacuum transferred onto a nylon membrane and incubated with Cy5- streptavidin. The membrane is scanned to quantify fluorescence intensity and the signal is converted into number of uracils using a duplex standard described in 2.1.7b Standard Plot. This figure was originally published in the supplementary material for (165) Shalhout et al. *Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development.* Molecular and cellular biology. 2014;34(21):4019-32. Reproduced from the American Society for Microbiology.

2.1.7b Standard Plot

To quantify the AP sites per million base pair and Uracils (U) per million base pair, a standard oligonucleotide duplex containing a single U:G mismatch is used in both detection assays. The synthetic 75 base pair duplex DNA (IDT) has the following sequence: Top Strand 5'-$T_{37}UT_{37}$; Bottom Strand 5'-A_{37}**G**A_{37}. The oligos were PAGE purified and the 75mers were annealed (1:2 ratio of top: bottom strand) and diluted to a final concentration of 2 x 10^{11} U per µL.

This duplex standard was treated with methoxyamine at 37°C for one hour, followed by treatment with an excess of UNG (2 units per 0.5 µg of DNA) and 2 mM of ARP for one hours at 37°C. To remove protein and unbound ARP, the duplex is phenol:chloroform extracted and ethanol precipitated. It is then reconstituted to give a final concentration of 1 x 10^{10} U per µl. Several dilutions of this DNA are vacuum-applied to each Nylon+ membrane, in parallel with genomic DNA tested for uracil content or AP site load. This allows for generation of a calibration plot of Cy5 fluorescence versus uracil number (or AP site number) for every membrane with spotted test samples. See Figure 22 for a representative standard plot. The raw fluorescence numbers are determined with ImageJ and are adjusted for background fluorescence signal on each membrane. The numbers of uracils or AP sites in the genomic DNAs are determined by interpolating their fluorescence intensities using the calibration plot. These amounts are normalized for the known amount of DNA loaded for each sample to calculate the number of uracils or AP sites per 10^6 bp. The uracil amount calculated is divided by the "control" DNA and allows for a ratio report. The

control for tested murine samples was DNA uracil content level of WT mice splenocytes or DNA from unstimulated CH12F3 cells. For human lymphoma cell lines or patient tissue samples the control was B lymphocytes purified from the blood of American Red Cross donors/volunteers (165).

Figure 22. Representative Standard Plot for Genomic Uracil and AP Site Quantification Assay

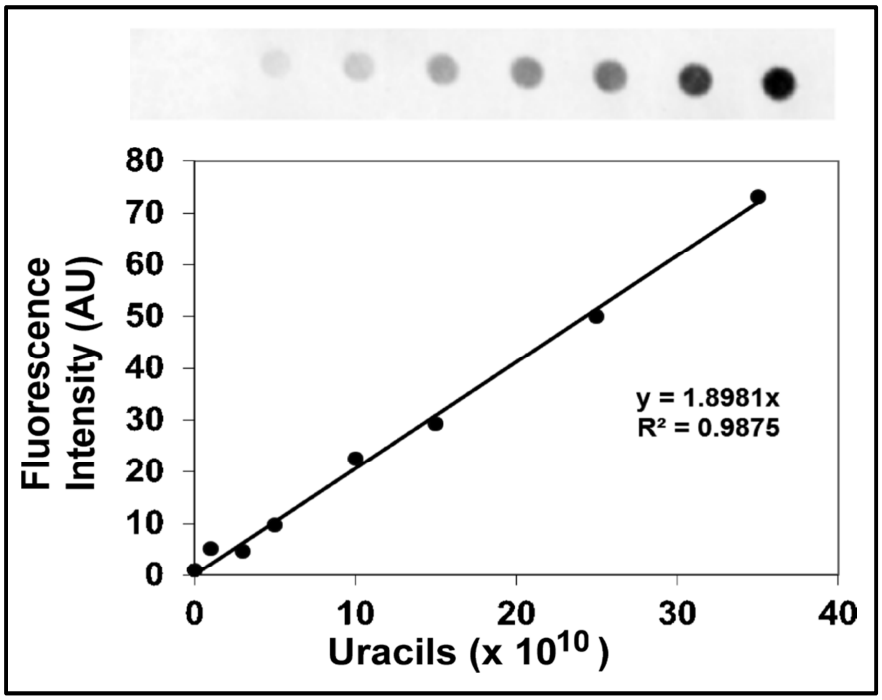

Figure 22. Representative Standard Plot for the Genomic Uracil and AP Site Quantification Assays. Top- A representative fluorescence scan of a Ny+ membrane with eight spots containing different amounts of the labeled uracil containing duplex DNA ranging from 0 to 35 x 10^{10} total uracils. Bottom- A representative plot of the fluorescence intensity of these spots as a function of number of uracils in each DNA duplex spot. The genomic uracils or AP sites in a mouse or human DNA sample is calculated by interpolating its fluorescence intensity in plots like this generated from standards applied onto each membrane. This figure was originally published in the supplementary material for (165) Shalhout et al. *Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development.* Molecular and cellular biology. 2014;34(21):4019-32. Reproduced from the American Society for Microbiology.

2.1.7c Uracils in U•G base pairs

The uracil quantification assay described above (section 2.1.7a) determines the total uracil content of genomic DNA. This is because it uses E. coli UNG to excise the uracils and this enzyme removes uracils from single-stranded DNA and from U:A base pairs in addition to excising uracils from U•G and other mispairs. It is important, however, to specifically detect U•G mispairs since AID introduces U•G mispairs by deaminating cytosines in C:G base pairs. To detect just the uracils in U•G mispairs, the assay was altered to use the second uracil-DNA glycosylase in E. coli, Mug. This enzyme is an ortholog of mammalian thymine-DNA glycosylase (TDG, see TABLE 1) which selectively removes uracils from U•G mispairs and has poor activity on uracils paired with adenine (176). Purified MUG enzyme was used in place of UNG in the protocol described in section 2.1.7a.

The U•G quantification assay protocol requires membrane processing, fluorescence acquisition, and interpolation from a standard plot generated by spotting dilutions of the duplex oligo exactly as used in the total uracil quantification assay. To quantify only U:A bp, DNA was first treated with MUG to remove all uracils in U•G bps, then treated with Mx to block these sites from being biotinylated. After purification, the DNA is then treated with UNG which removes the uracils left in the DNA that are in U:A bps only since previous MUG treatment excised all uracils in U•G bps. The AP sites left after UNG removes uracils opposite A are then labeled with ARP and the only uracils in U:A base

pairs are calculated using a standard, similar to the protocol for AP site and total uracil quantification.

2.1.8 Activity Assays with Whole Cell Extracts (WCE) and Nuclear Extracts (NE)

2.1.8a Cytosine Deamination Activity Assays

Protein preparations from whole cell extracts (WCE) were prepared from ~1 x 10^7 cells. Cells were centrifuged and the pellet was resuspended in buffer I [10 mM Tris-HCL (pH 8.0)] and buffer II [10 mM Tris-HCl (pH 8.0), 200 mM KCl, 2 mM EDTA, 40% (v/v) glycerol, 0.5% NP-40, 2 mM DTT] with the addition of protease inhibitor (Halt Protease Inhibitor Cocktail, Thermo Scientific). The cells were then incubated at 4ºC for 3 hours with gentle agitation and cellular debris was pelleted at 25,000 g at 4ºC. Cell-free protein extract was recovered in the supernatant. Nuclear extracts of ~1 x 10^7 cells were prepared with the NucBuster Protein Extraction Kit (Novagen, Millipore) according to the manufacturer's instructions with the addition of three extra washes of each nuclear pellet preparation to ensure no cytoplasmic protein contamination occured during preparation. Whole cell and nuclear protein concentrations were measured using the Bio-Rad protein assay and extracts were snap frozen in liquid nitrogen and stored at -80ºC.

To assay AID deamination activity in protein preparations, 8 pmols of a 5' 6-FAM labeled oligo containing AID's preferred sequence (WRC, where W=A/T, R=purine, 5'-ATTATTA**C**CCATTTATT) was incubated with whole cell or nuclear cell extracts for 30 minutes with RNaseA (1 µg), in the following reaction buffer: 10 mM Tris·HCl, pH 7.5, 1 mM EDTA/1 mM DTT, 150mM NaCl. Reactions were

incubated at 37℃ with a total reaction volume of 20µl and were terminated by the addition of 1, 10-phenanthroline (Sigma–Aldrich) to 5 mM. UNG (0.5 units) was added to the reactions to excise any uracils introduced by AID deamination and incubation was continued at 37 ℃ for 60 min. Any abasic sites resulting from uracil excision were nicked by the addition of 0.1M NaOH and incubation at 95 ℃ for 5 min. See section 2.1.8b for reaction product analysis and specific activity reports (165).

2.1.8b Uracil Excision Activity Assays

To assay uracil excision activity, 24 pmols of a single-uracil containing 6-FAM labeled oligomer (5'- ATTATTA**U**CCATTTATT) is incubated with whole cell or nuclear cell extracts for ten minutes at 37ºC in a 20µl total reaction with 1 mM EDTA, 1 mM DTT, and 20mM Tris-HCl (pH 8.0) with or without excess Uracil DNA-Glycosylase Inhibitor (UGI,New England Biolabs). A single-strand substrate is used to avoid detection of uracil excision activity from TDG and MBD4 which prefer to remove mispaired uracil from duplex DNA. Reactions are stopped at 95ºC for 5 minutes followed by incubation with Proteinase K (Qiagen) for 60 minutes at 50ºC. Abasic sites are incised using 0.1 mM NaOH treatment at 95ºC for five minutes. Reaction products for AID deamination assays (section 2.1.8a) and uracil excision assays are analyzed by electrophoresis (15% PAGE, 7M Urea, Tris-Borate-EDTA buffer), and a Typhoon 9210 Scanner is used to visualize products. Quantification of activity is determined with ImageJ Gel Analysis Software and specific activity is reported as pmols of product/min/µg of protein (165).

2.1.9 Assays for Single-Strand and Double Strand Breaks in DNA

2.1.9a Comet Assays and Tail moments

The comet assay or single-cell gel electrophoresis is utilized in this study to detect strand breaks in the genomic DNA of lymphoma cell lines expressing AID. This technique requires electrophoresing cells that are embedded in low-melting agarose and visualizing by fluorescence microscopy. The alkaline comet assays were performed using the Comet Assay Silver Kit (Trevigen) according to manufacturer's instructions. Tail moments were quantified using Image J and these experiments were carried out by a collaborator (A. Martin, University of Toronto).

2.1.9b Immunofluorescence of γH2AX Foci and Quantification of Foci per Nuclei

Cells were harvested by centrifugation and cell pellets were then resuspended in a hypotonic solution and incubated at 37^{o}C for five minutes. Approximately ~500 cells were spun onto microscope slides (Fisher) using a StatSpin Cytofuge. They were fixed with ethanol:acetic acid, blocked with 3% rabbit serum (Hyclone) in PBS to reduce non-specific secondary antibody binding, and treated with mouse monoclonal anti-phospho-histone H2A.X (Ser139) antibody (Millipore; 05-636). The slides were washed in PBS two times, followed by treatment with rabbit anti-mouse IgG antibody conjugated to Cy3 (Millipore; AP160 C). The slides are washed in PBS and mounted with coverslips using VectaShield Mounting Medium with DAPI. Cells were visualized and

photographed using a fluorescence microscope (Nikon 80i). The foci per nuclei were quantified using ImageJ. Untreated normal B cells and cells treated with 50 µg/ml zeocin (Invitrogen) served, respectively, as negative and positive controls for γH2AX foci analysis.

2.1.10 Short-hairpin RNA (shRNA) Knockdown of Expression in Lymphoma Cell Lines

AID expression was knocked down in AID expressing lymphoma cell lines using shRNA plasmids obtained from SA Biosciences (SureSilencing shRNA plasmids for human AICDA for neomycin resistance, KH12741N). Transfection grade plasmids were prepared with HiSpeed Plasmid Maxi Kit (Qiagen). Cell lines were either electroporated using a Gene Pulser Xcell (BioRad) or transfected with shRNA-AID and shRNA-scrambled using FuGeneHD (Promega) depending on cell line optimization conditions. Stable clones of shRNA-AID and shRNA-scrambled cell lines are selected with G418 (HyClone).

2.1.11 Expression of Uracil DNA-Glycosylase Inhibitor in Lymphoma Cell Lines to reduce Uracil Excision by UNG

To inhibit uracil DNA-glycosylase activity, human B cell lymphoma cell lines were transfected with pEF (control) and pEF-UGI plasmids by electroporation (Gene Pulser Xcell). Cells were diluted in appropriate media and plated into 96-well plates. Following 24- 48 hr incubation at 37ºC, stable clones were selected with puromycin (Promega).

2.1.12 Transient Expression of E58A and WT AID in Ramos 1 Cell Line

Approximately two million Ramos 1 cells were transfected with transfection grade pMSCV-EV (empty vector), pMSCV-E58A, and pMSCV-AID

isolated with HiSpeed Plasmid Maxi Kit, (Qiagen) using FuGeneHD (Promega) and selected with 2 µg/ml of puromycin (Gibco). The pmax-GFP (Amaxa Lonza) plasmid was used to track transfection efficiency with FuGene of the transient clones.

CHAPTER 3: RESULTS

Figures and portions of the text in this chapter were reprinted or adapted in compliance with the *ASM Journals Statement of Authors' Rights* from:

Shalhout S, Haddad D, Sosin A, Holland TC, Al-Katib A, Martin A, Bhagwat AS. *Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development*. **Molecular and cellular biology**. 2014;34(21):4019-32. doi: 10.1128/MCB.00589-14.

3.1 AIM 1- Genomic Uracils in Normal Activated B cells

The extent of global genomic uracil introduction by AID and the removal of uracils by repair enzymes in normal human and murine stimulated B cells are examined for the first time here. Genomic uracils are quantified using a sensitive biochemical uracil detection assay to elucidate genome-wide AID activity in adaptive immunity.

Splenocytes from wildtype C57BL/6 mice were harvested and stimulated *ex vivo* with LPS to closely mimic the activation involved in the normal development of B cells during the GC- reaction. Upon activation with LPS, the cultured splenocyte population was ≥ 85% B cells shown with flow cytometry. For comparison, splenocytes from AID-/- transgenic mice were stimulated in the same manner as a negative control for AID-generated uracils. Splenocytes from UNG-/- and MSH2-/- transgenic mice which are deficient in two separate repair proteins thought to play roles in SHM and CSR were also used in this study to determine the effects of these proteins on processing the uracils AID introduces during activation of B cells. Finally, splenocytes from AID-/-UNG-/- mice were

also tested to demonstrate that the detected uracil content is AID-dependent. The total cellular DNA from these cells was isolated and the uracils quantified as described in section 2.1.7. DNA from unstimulated splenocytes of each mouse background was extracted and also tested for uracil content as 'non-activated' B cell controls. Liver and kidney tissues were also harvested and single cell suspensions from these tissues were studied as the 'non-B cell' controls. These experiments were designed to determine the amount of genomic uracil that is AID-dependent and to elucidate which uracil repair enzyme is primary in global genomic uracil removal during the normal GC- reaction of murine B cells.

3.1.1 Genomic Uracil Quantification of Activated Murine B Cells

An image of a Cy5 fluorescence scan of a nylon membrane from one representative experiment is shown in Figure 23 where equal amounts of DNA from *ex vivo* stimulated splenocytes under LPS treatment for 3 days were spotted. In this and all other uracil quantification studies, different known amounts of a duplex containing a U•G mispair were also processed in parallel and applied to the membrane to generate a calibration plot as depicted in Figure 22 (see section 2.1.7b for details). Cy5 fluorescence signals from genomic DNA of mice with different genetic backgrounds were converted to uracil amounts using the calibration plot, and then uracil amounts in WT splenocytes prior to stimulation were used to calculate the fold increases in uracil levels. The scanned image strikingly reveals the highest Cy5 fluorescence intensity and therefore uracil content in DNA samples from UNG-/- splenocytes stimulated for 3 days. DNA from UNG-/- unstimulated splenocytes and hepatocytes also

contain slightly higher fluorescence intensity compared to other genetic backgrounds due to the general lack of UNG available for removal of uracils introduced in AID-independent pathways (see Figure 10). The quantified results from a typical experiment are presented in Figure 24 A and B.

Figure 23. Representative Nylon Membrane used in Quantifying Genomic Uracils in WT and Transgenic Murine DNA

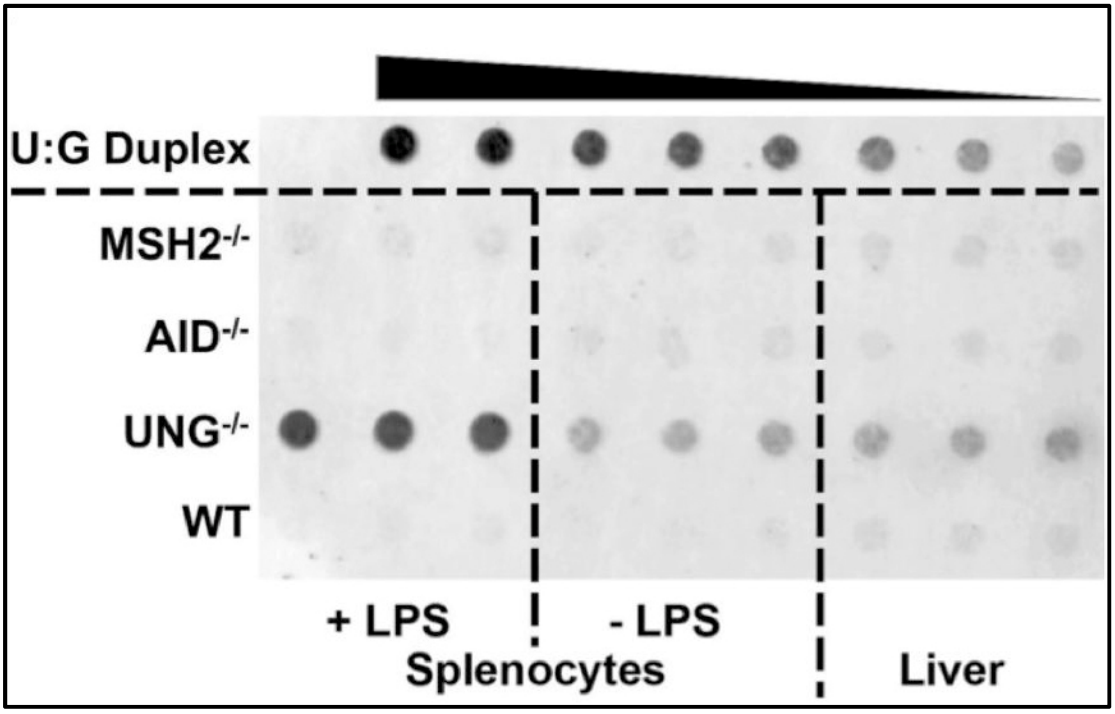

Figure 23. Nylon Membrane used in Quantifying Uracils in Murine DNA. A representative mage of a fluorescence scan of a representative nylon membrane with equal DNA amounts spotted from triplicate samples of unstimulated murine splenocytes (-LPS) and splenocytes stimulated with LPS (+LPS) for 3 days. The splenocytes were obtained from WT mice and mice with the indicated genetic backgrounds. DNAs were labeled with Cy5 at sites of uracils as described in Figure 21. Liver DNAs from the same mice served as negative controls. As shown at the top, different amounts of an oligonucleotide duplex containing uracils were spotted onto the membrane to create a calibration plot (see Figure 22). This figure was originally published in (165) Shalhout et al. Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development. Molecular and cellular biology. 2014;34(21):4019-32. Reproduced from the American Society for Microbiology.

Figure 24. Genomic Uracils in Murine Splenocytes

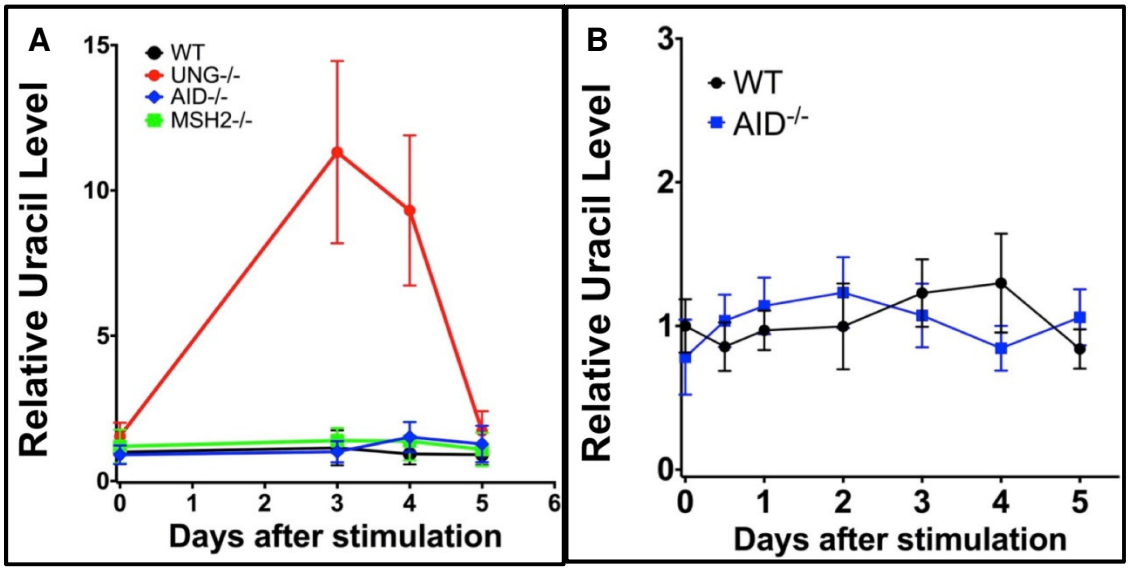

Figure 24. Genomic Uracils in Murine Splenocytes. (A) Quantification of uracils in splenocyte DNA following stimulation with LPS. The fluorescence intensities from a membrane similar to the one shown in Figure 23 were used to quantify the uracils and were normalized with respect to the uracil amount in WT splenocytes prior to LPS stimulation. In these and all other experiments reported here, means and standard deviations of the results from triplicate samples are reported. (B) Time course of genomic uracil levels in stimulated WT and AID-/- splenocytes. Uracils were quantified and the data analyzed as described for A. This figure was originally published in (165) Shalhout et al. Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development. Molecular and cellular biology. 2014;34(21):4019-32. Reproduced from the American Society for Microbiology.

In splenocyte B cell DNA from WT and AID-/- mice, there was no detectable increase in the level of uracils during the five day *ex vivo* stimulation time course (Figure 24A and B). This result was readily reproducible (see Figure 25A and B). This lack of difference in an increase in uracil accumulation between WT (ie AID+/+) and AID-/- splenocytes throughout the LPS activation treatment was surprising since studies have shown AID can lead to mutations in many non-Ig genes and causes genome-wide translocations (177, 178). Similarly, there was also no increase in the level of genomic uracils when splenocytes from

MSH2-/- mice were activated (see Figure 23, 24A and 25). Thus, despite the role of MMR in SHM and CSR (see section 1.2.3d, 1.2.4b and Figure 9) (178-180), these results suggest that MMR is not the primary uracil repair pathway involved in the removal of global uracils generated by AID across the genome. This result is consistent with the well-established role of MMR in correcting replication errors and with a previous study that compared relative effects of UNG-dependent repair and MMR on SHMs in some AID target genes (178).

In contrast, a significant stimulation-dependent increase in uracils was observed in UNG-/- splenocytes. Relative to unstimulated UNG-/- splenocytes, the genomic uracil levels increased approximately 11-fold in UNG-/- splenocytes within the first 3 days following LPS activation. This high uracil accumulation in the genomes of UNG-/- splenocytes at day 3 was detected in repeated experiments (see Figure 25). In some experiments, an increase of greater than 11-fold was observed (see Figure 25). The uracil level was then restored back to the levels detected in unstimulated cells by the fifth day of the time course (see Figure 23 and 24A, red line). The factors likely to contribute to the decrease in genomic uracil levels beyond day 3 include a decrease in AID gene expression after day 3, possible removal and/or repair of uracils by backup DNA glycosylases such as SMUG1, and replicative dilution of incorporated uracils. These data are qualitatively consistent with the increased uracil levels in the Ig genes of stimulated UNG-/- B cells (163), but the magnitude of the increase I detected is much greater than that reported previously. To demonstrate that the presence of uracils was due to AID, I quantified genomic uracils in LPS-

stimulated splenocytes from AID-/-UNG-/- mice. There was no detectable increase in uracil levels in AID-/-UNG-/- cells following stimulation, while UNG-/- cells stimulated in parallel with the double knock out cells acquired 15-fold-higher levels of uracils within 3 days (see Figure 26). Together these data suggest that splenocyte stimulation with LPS causes a rapid and large increase in genomic uracil levels due to AID but that an UNG2-mediated repair pathway eliminates the uracils in WT cells.

3.1.2 Time Course of the levels of AID and UNG2 Gene Expression during B Cell Stimulation

Studies concerning the introduction of uracil by AID into the genome also must consider the proteins involved in the removal or repair of those uracils. If uracils are readily repaired by several pathways and uracil repair enzymes, AID generated uracils may actually be underrepresented or escape detection. To elucidate the balance between uracil creation by AID and uracil excision by BER, mRNA levels of AID, UNG2, and SMUG1 genes were determined in the LPS stimulated mouse splenocytes over a several day time course.

In activated splenocytes from WT, UNG-/-, and MSH2-/- genetic backgrounds, AID gene expression levels increased ~20-fold relative to unstimulated splenocytes. mRNA levels were normalized to the house-keeping gene, GAPDH (glyceraldehyde 3-phosphate dehydrogenase) which was used as an internal control in these studies (see section 2.1.4 for qRT-PCR details).

Figure 25. Genomic Uracil Results are Reproducible

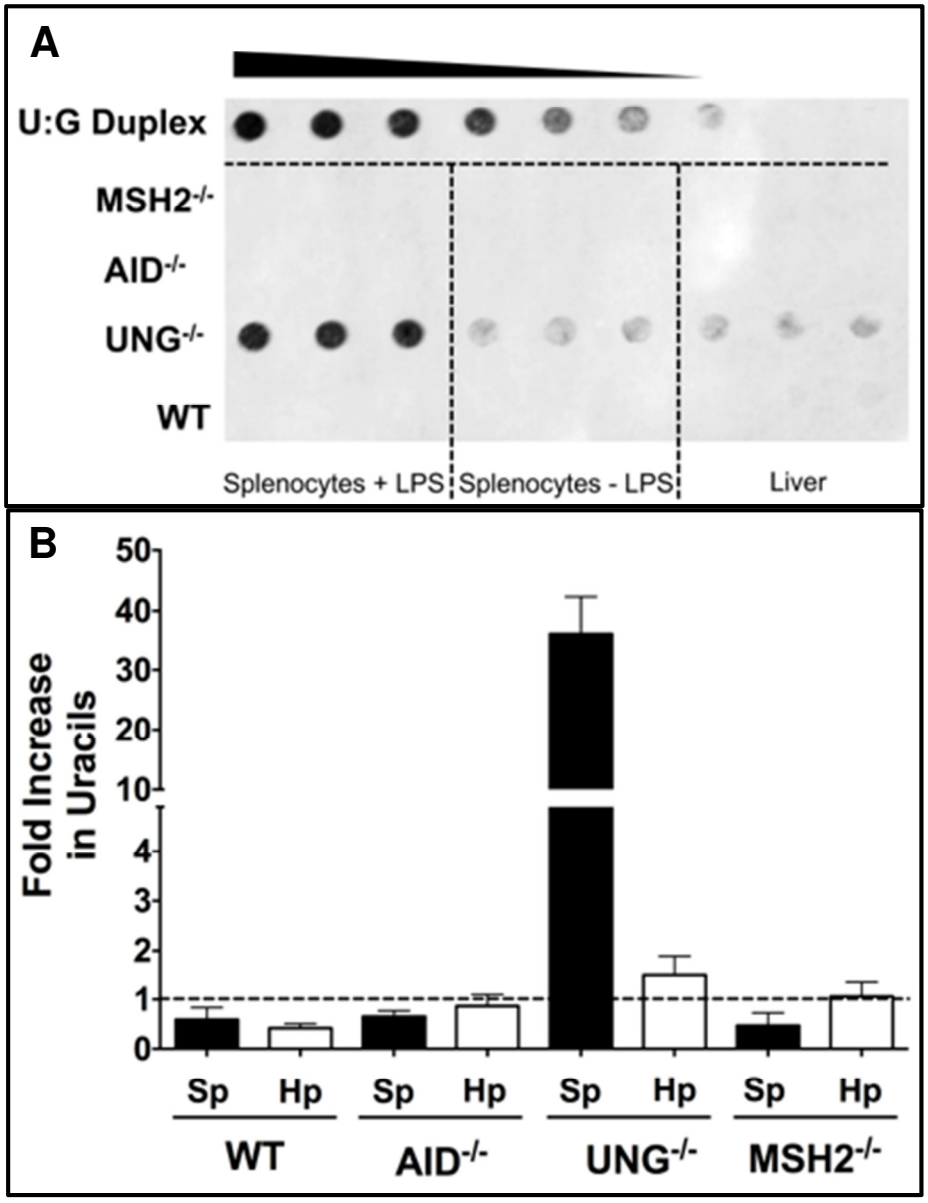

Figure 25. Genomic Uracil Results are Reproducible. (A) Independent experiment showing a Ny+ membrane used for uracil quantification from a different study with different mice samples than that shown in Figure 23 for murine splenocyte and liver DNAs. Each labeling reaction contains equal amounts of DNA spotted onto the membrane (350 ng) and the reactions were done in triplicates. Stimulation with LPS was carried for three days. (B) Quantification of uracils from part A. The numbers are normalized with respect to the number for unstimulated WT cells (set to 1.0). Mean and standard deviation are shown. Sp- splenocytes; Hp- hepatocytes. This figure was originally

published in the supplementary material in (165). Reproduced from the American Society for Microbiology.

Figure 26. Genomic Uracil Levels in AID-/-UNG-/- Splenocytes

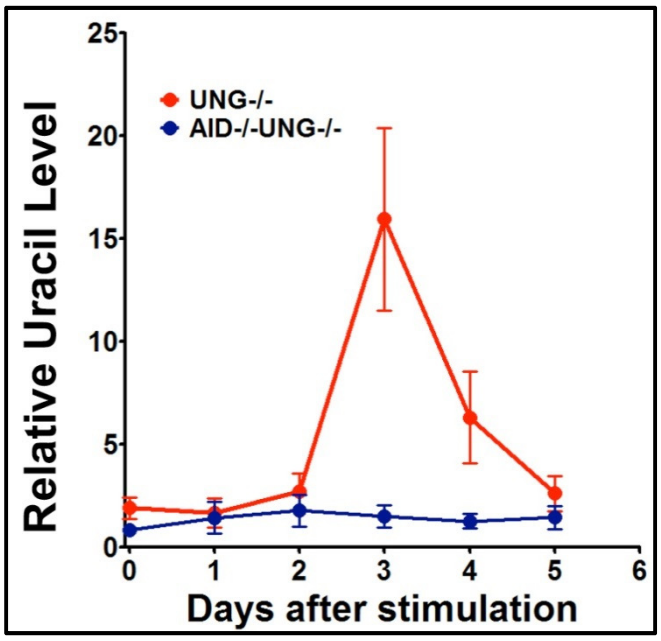

Figure 26. Genomic Uracil Levels in AID-/-UNG-/- B cells. Relative genomic uracil levels in LPS-stimulated splenocytes from UNG-/- and AID-/-UNG-/- double knock out mice normalized to uracil levels in unstimulated WT splenocytes. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

This sharp increase in AID gene expression upon stimulation occurred during the first 3 days following stimulation and was reduced thereafter, never quite reaching pre-stimulation levels by the end of the five day time course (Figure 27). As expected, no AID mRNA was detected in the samples prepared from the AID-/- cells (see Figure 27 B, blue line). In stimulated splenocytes from WT mice, the rise and fall of UNG2 gene expression paralleled AID gene expression. UNG2 and AID both increased approximately 20-fold compared to levels in unstimulated cells during the first 3 days and declined thereafter (Figure

27A). The increases in UNG2 mRNA levels following stimulation of WT splenocytes were correlated with increases in single-strand-specific uracil excision activities in both nuclear extracts (Figure 29) and whole-cell extracts (Figure 30). UGI, a specific inhibitor of the UNG enzymes, almost completely eliminated the uracil excision activity (Figure 29 and 30). Such upregulation of UNG2 following B cell stimulation has been noted before (181).

Figure 27. Gene Expression in Stimulated Murine Splenocytes

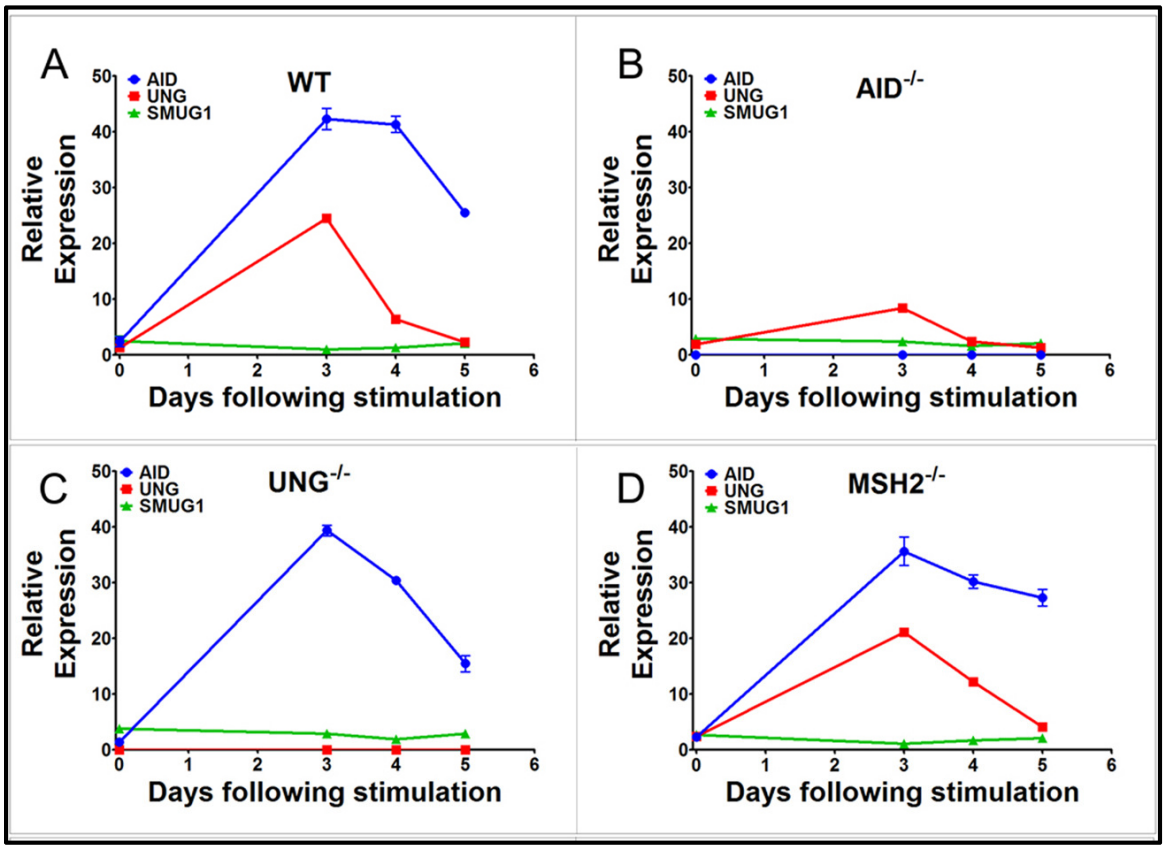

Figure 27. Gene Expression in Stimulated Murine Splenocytes. The mRNA levels of AID (blue), UNG2 (red), and SMUG1 (green) genes relative to the GAPDH expression level (set at 100) are shown. The gene expression levels were determined prior to and subsequent to LPS stimulation of splenocytes from

WT (A), AID-/- (B), UNG-/- (C), and MSH2-/- (D) mice. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

3.1.3 Cytosine Deamination during B Cell Stimulation

Nuclear protein extracts and whole cell protein extracts were prepared from WT and UNG-/- cells to test the cytosine deamination activity using a fluorescently labeled single-stranded DNA substrate containing a cytosine within the AID hotspot sequence motif, WR**C** (see section 2.1.8a for assay details). The cytosine deamination activities were similar in the two types of cells and peaked at day 3 following stimulation with LPS (Figure 28). Thus, both AID mRNA levels and DNA-cytosine deamination activities follow the same general pattern of increase and decline in the two genetic backgrounds. The increase in cytosine deamination was seen using whole cell extracts as well as nuclear extracts. However, of significance here is the increase in deamination activity upon stimulation detected specifically in the nucleus because only nuclear AID can lead to uracil accumulation in genomic DNA.

Figure 28. Nuclear and Whole Cell Cytosine Deamination Activity

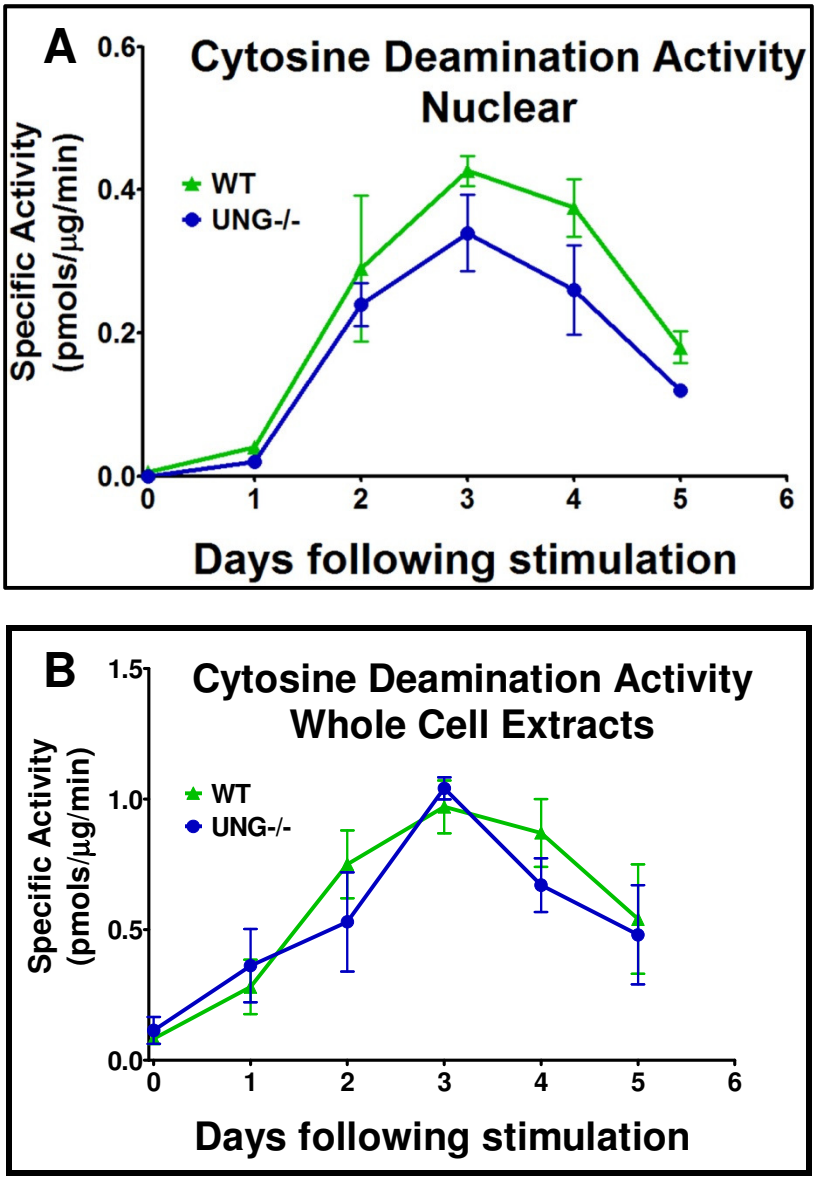

Figure 28. Nuclear and Whole Cell Cytosine Deamination Activity. A time course of DNA-cytosine deamination activity in nuclear extracts (A) and whole cell extracts (B) prepared from LPS-stimulated splenocytes from WT (green) and UNG-/- (blue) mice. Specific activity was calculated as picomoles of DNA substrate converted to product per microgram of protein in the extract per minute. Part A of this figure was originally published in (165). Reproduced from the American Society for Microbiology.

Figure 29. Nuclear Uracil Excision Activity

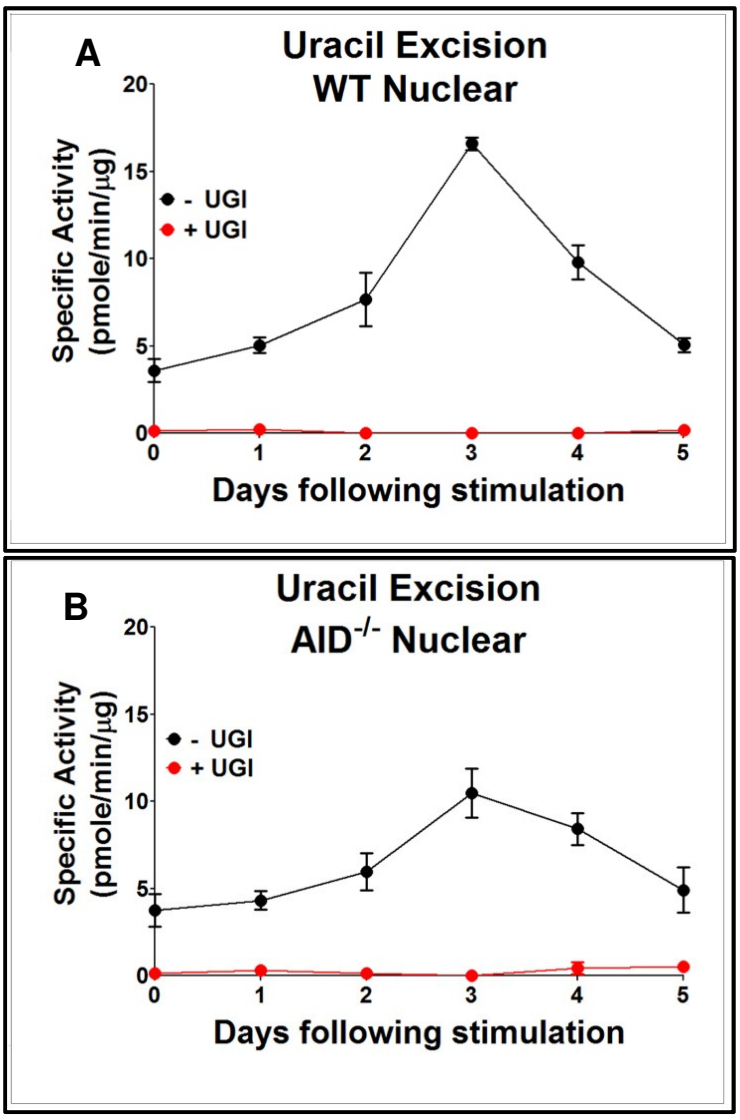

Figure 29. Nuclear Uracil Excision Activity. A fluorescently labeled DNA oligomer containing uracil was incubated with nuclear extract from WT (A) or AID-/- (B) LPS-stimulated splenocytes, and the specific activity of uracil excision was determined. The activity was determined in the presence (+Ugi) or absence (-Ugi) of the UNG enzyme inhibitor UGI.. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

Figure 30. Whole Cell- Uracil Excision Activity

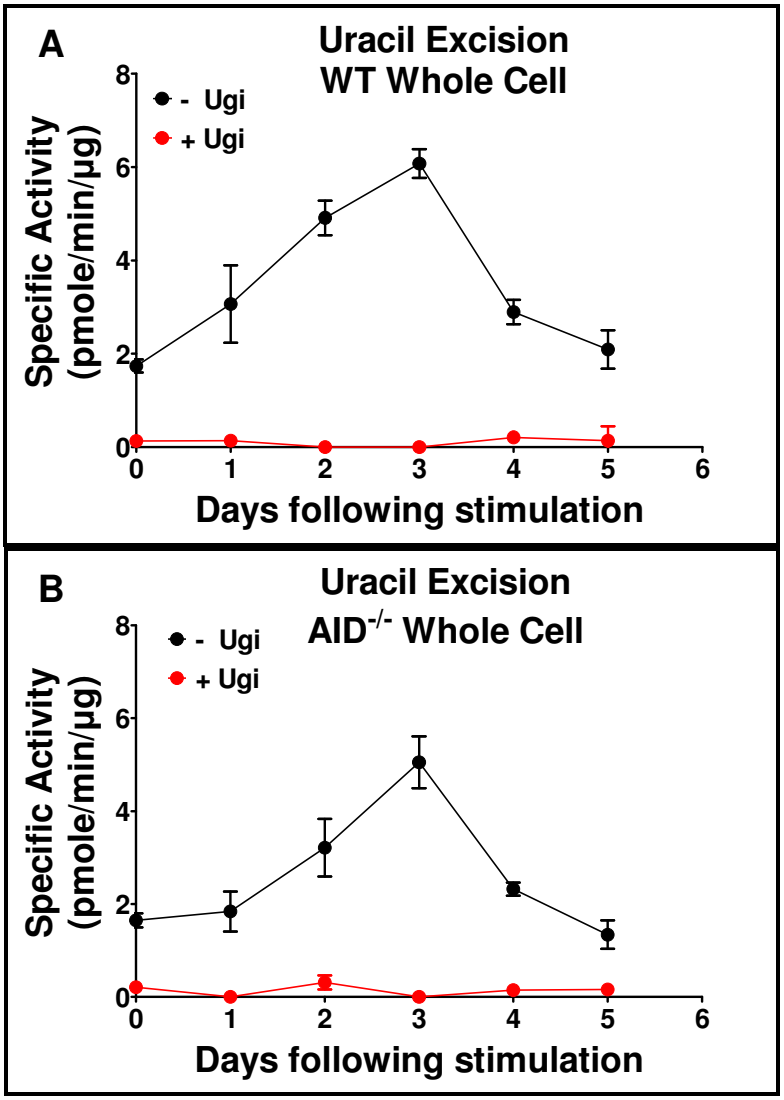

Figure 30. Whole Cell- Uracil Excision Activity. A fluorescently labeled DNA oligomer containing uracil was incubated with whole cell extract prepared from WT (A) or AID-/- (B) LPS-stimulated splenocytes, and the specific activity of uracil excision was determined. The activity was determined in the presence (+Ugi) or absence (-Ugi) of the UNG enzyme inhibitor UGI.

3.1.4 The Repair of Genomic Uracils during Normal B Cell Stimulation

Although UNG2 gene expression increased in AID-/- splenocytes following LPS stimulation, the magnitude of the increase was smaller than in WT mice (Figure 27A and B). In contrast to this result, a previous study using transcriptome sequencing (RNA-Seq) found UNG expression to be about the same in stimulated B cells from WT and AID-/- mice (182). To resolve this issue, I determined the uracil excision activity in nuclear extracts (Figure 29) and whole cell extracts (Figure 30) from AID-/- cells and found that it also peaked at day 3 post stimulation but that its magnitude was smaller than the peak activity seen in WT cells (compare Figure 29A and B and Figure 30A and B). This activity was eliminated when UGI was added to the reaction mixture, showing that it was UNG specific. As predicted, no expression of UNG2 gene was detected In UNG-/-splenocytes (Figure 27C). In MSH2-/- splenocytes, UNG2 gene expression followed the same pattern as in WT cells (Figure 27D).

In contrast to UNG2 gene expression, there was no increase in the expression of SMUG1 gene in any of the genetic backgrounds tested (Figure 27A through D). Similarly, splenocyte stimulation did not lead to large increases in expression of other uracil-specific DNA glycosylases (TDG and MBD4) following stimulation of WT splenocytes (see Figure 31). Thus, despite recent reports that SMUG1 contributes more to uracil excision than UNG2 in mice (181) and that SMUG1 contributes to CSR and SHM in mice (183), these results show that UNG2 is the only uracil specific DNA glycosylase whose expression is

upregulated in response to B cell stimulation. Thus, it must be responsible for the repair of an overwhelming majority of uracils created by AID in stimulated WT B lymphocytes. In other words, B cells undergoing maturation *ex vivo* display homeostasis in genomic uracil levels.

Figure 31. Gene Expression Levels of Uracil Repair Proteins

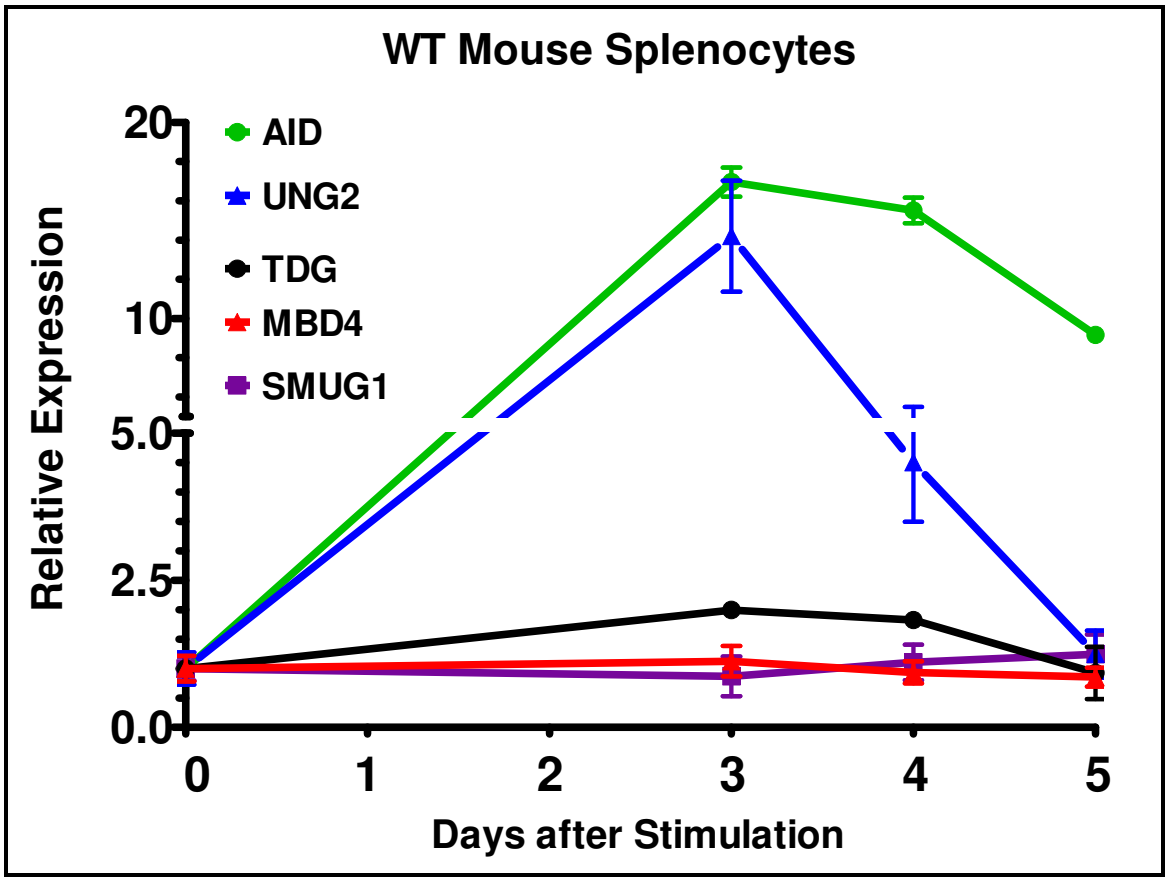

Figure 31. Gene Expression Levels of Uracil Repair Proteins. Expression levels of AID (green), TDG (black), MBD4 (red), UNG2 (blue), and SMUG1 (purple) in LPS stimulated WT splenocytes were determined. qRT-PCR results are shown above for unstimulated WT cells (day 0) and cells 3, 4, and 5 days post-stimulation with LPS. The expression levels of the genes are relative to GAPDH expression level and re-normalized to the gene expression levels at Day 0 (set to 1.0).This figure was originally published in (165). Reproduced from the American Society for Microbiology

3.1.5 Genomic Uracil Levels in Activated human Naïve B Lymphocytes

To extend these observations to human B cells, I purified mature naive B cells from tonsil tissues obtained from two human donors and stimulated the B cells *ex vivo* with anti-CD40 antibody and IL-4 to mimic the GC-reaction. The AID expression and genomic uracil levels were determined in the cultured B cells over a 7 day time course. AID expression increased substantially in both sets of samples after 4 days of stimulation, but the magnitudes of the increase were different in the two samples (Figure 32A). Regardless, the genomic uracil level changed little over the 7 days in either sample (Figure 32B). This shows that, like the WT murine splenocytes, normal human B cells maintain their genomic uracil levels following stimulation despite an increase in AID expression. A similar increase in the expression of UNG2 also increased upon stimulation (Figure 32C).

Figure 32. Gene Expression Levels and Uracils in Human B Cells

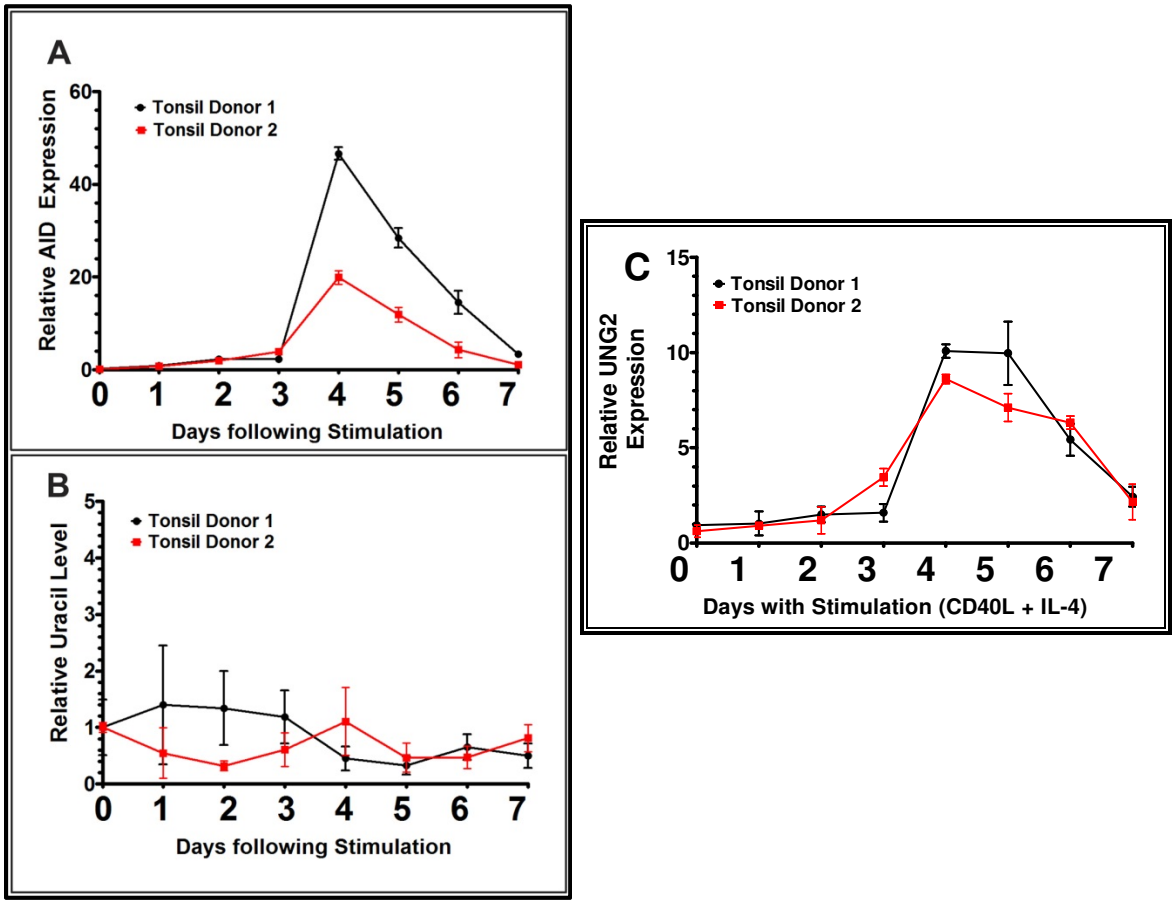

Figure 32. Gene Expression Levels and Uracils in Human B cells. Naive mature B cells were isolated from tonsils from two independent donors and cells were stimulated with anti- CD40 antibody and IL-4 for 7 days. AID (A) and UNG2 (C) expressions were normalized to GAPDH expression (with the day 0 level set to 1.0). (B) The quantification of the genomic uracil levels in the *ex vivo*-stimulated tonsillar B cells relative to uracil levels in DNA from unstimulated cells are shown (with the day 0 level set to 1.0). Panel A and B of this figure were originally published in (165). Reproduced from the American Society for Microbiology.

3.2 AIM 2- Genomic Uracils in Cancerous B cells

The AID-dependent global uracil levels in B cell lymphoma genomes and the uracil removal capabilities are examined. The genomic uracil homeostasis found in normal activated B cells is lost in cancer cells where AID is constitutively and ectopically expressed.

3.2.1 Uracil accumulation in murine B cell Lymphoma lines despite high UNG2 activity and expression

To determine whether the lack of uracil accumulation in the genome of activated WT B cells was also true in germinal center-derived cancer cell lines, I studied uracil accumulation in the genome of the CH12F3 cell line since it is a murine B cell lymphoma-derived line that expresses AID upon stimulation with CIT and switches from the IgM isotype to IgA (see section 1.1.1 for more details) (1). The levels of genomic uracils and the expression of AID, UNG2, and SMUG1 genes in this cell line following its stimulation were determined and compared with the pattern seen in WT and UNG-/- murine splenocytes.

Three days post stimulation, CH12F3 cells accumulated uracils in their genome to a level about fourteen times the unstimulated level (Figure 33A). As was observed in LPS-stimulated UNG-/- splenocytes (Figure 24A), the uracil content decreased after day 3, returning to the unstimulated level by day 10.

To directly compare the uracil amounts in the genomes of these two types of murine cells, equal amounts of DNA from primary WT and UNG-/- splenocytes and from CH12F3 cells were labeled at uracil sites and the samples were spotted on nylon membranes for quantification. A fluorescence scan of the membrane

shows roughly equal intensities for DNA from UNG-/- splenocytes and CH12F3 cells after 3 days of stimulation (Figure 33B).

Figure 33. Uracil Levels in CH12F3 Stimulated Cells

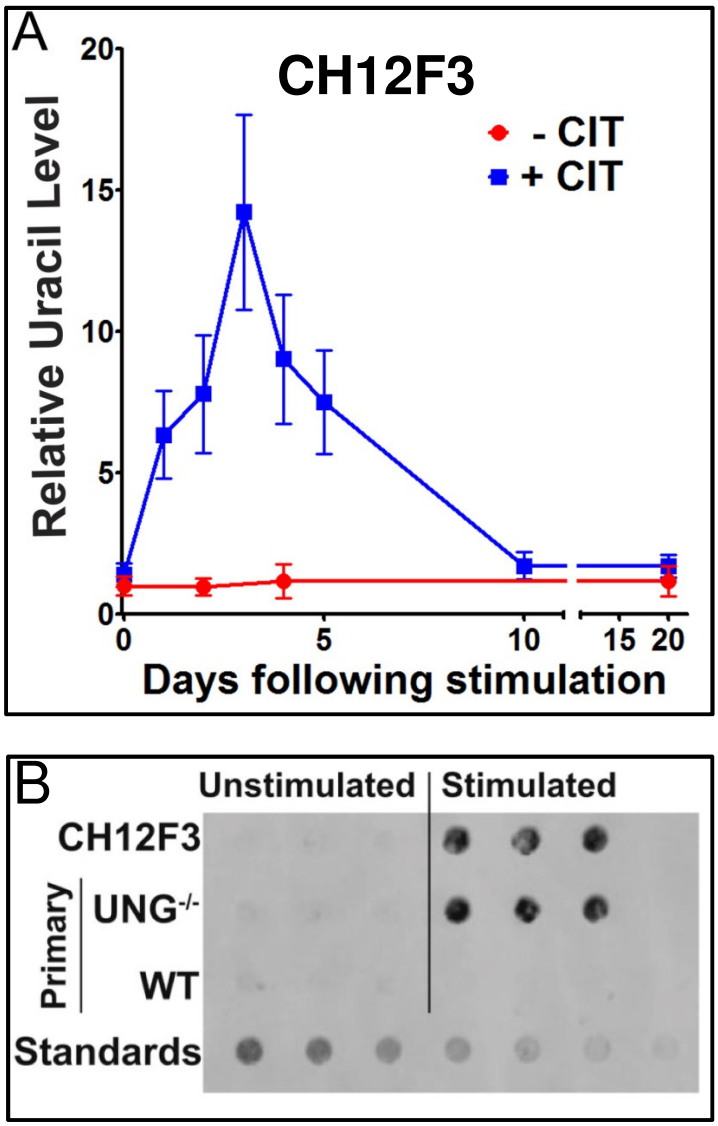

Figure 33. Uracil Levels in CH12F3 Stimulated Cells. (A) Time course of the change in genomic uracil levels in CH12F3 DNA with or without CIT stimulation. (B) Fluorescence scan of a nylon membrane on which equal amounts of DNAs labeled at uracil sites from unstimulated WT and UNG-/- mouse splenocytes and unstimulated CH12F3 cells were spotted. DNAs from splenocytes stimulated with LPS or CH12F3 cells stimulated with CIT for 3 days were labeled for uracil quantification in parallel and also spotted onto the membrane. The panels in this figure were originally published in (165). Reproduced from the American Society for Microbiology.

Quantification confirmed that the two sets of DNAs had similar levels of uracil content and that these amounts were much higher than those seen with the unstimulated CH12F3 cells and both unstimulated and stimulated WT primary splenocytes (see Figure 34). Therefore, stimulated CH12F3 cells behave like stimulated UNG-/- cells in terms of genomic uracil accumulation.

Both the AID and UNG2 expression profiles of CH12F3 cells were similar to those of stimulated WT splenocytes (compare Figure 34 B, C and Figure 27A). The kinetics of AID gene expression in CH12F3 cells was paralleled by nuclear DNA-cytosine deamination activity (Figure 35), and the same pattern was seen with UNG2 gene expression and nuclear uracil excision activity (Figure 36A and B). Furthermore, the peak specific activities of cytosine deamination and uracil excision in the nuclear extracts of CH12F3 cells were comparable to those in WT splenocytes (compare Figure 28A with Figure 35 and Figure 29 with Figure 36B). SMUG1 gene expression did not change detectably following CH12F3 stimulation, similar to *ex vivo* stimulated primary B cells (see Figure 37 and 27A). Thus, CH12F3 cells accumulated uracils in their genome following their stimulation despite the fact that their AID and UNG2 gene expression and nuclear enzymatic activity profiles were similar qualitatively and quantitatively to those in WT splenocytes.

Figure 34. Uracil Levels in Primary B Cells versus CH12F3 Cells

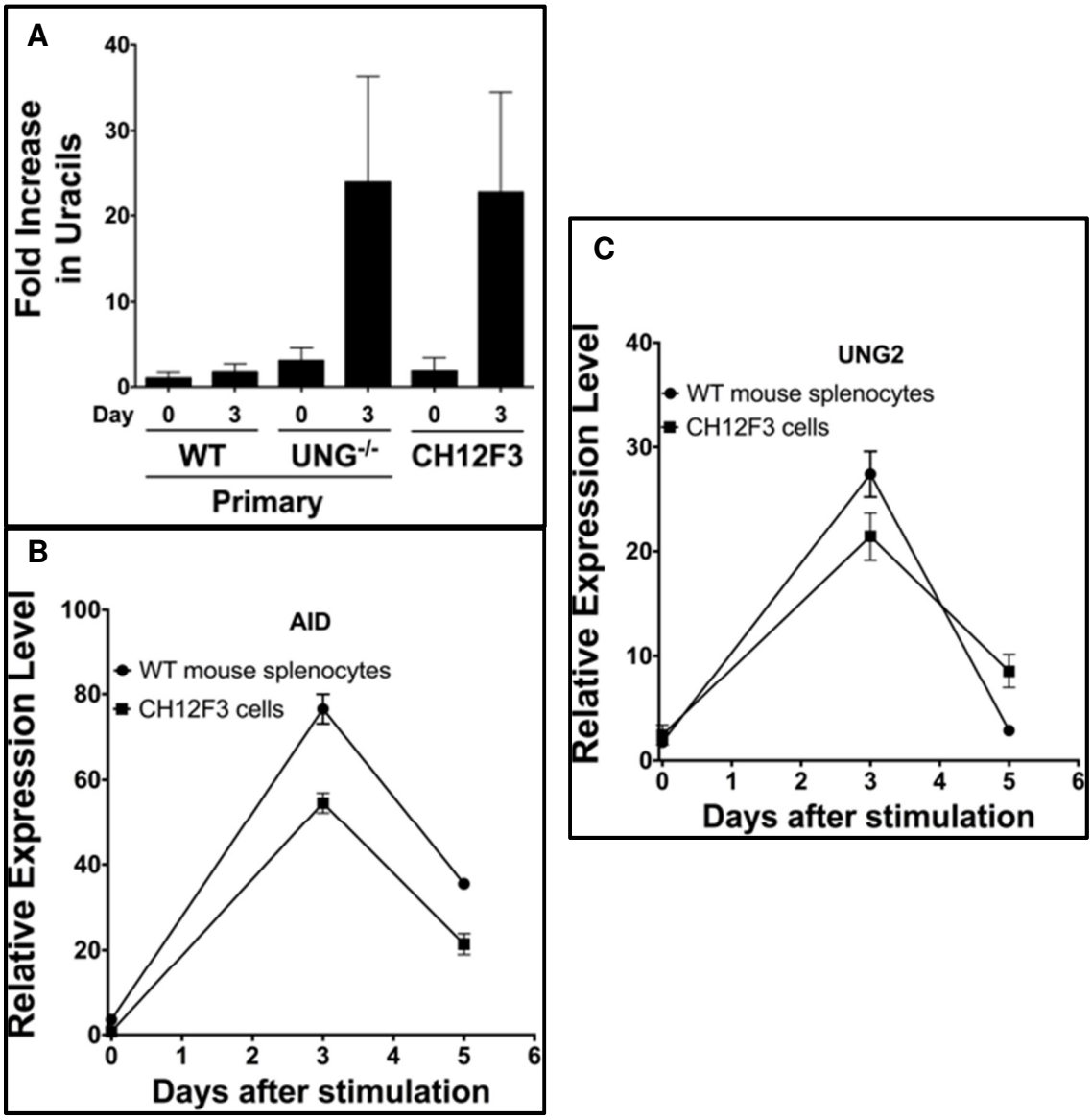

Figure 34. Uracil Levels in Primary B Cells versus CH12F3 Cells.
(A) Genomic uracil levels in splenocyte DNA and CH12F3 DNA before and at day 3 following LPS stimulation were determined. Levels are normalized to the uracil content in WT cells prior to stimulation. (B) Direct comparison of AID expression relative to GAPDH in WT splenocytes and CH12F3 cells. (C) Direct comparison of UNG2 expression relative to GAPDH in WT splenocytes and CH12F3 cells. qRT-PCR was done in parallel using the same primers. The panels in this figure were originally published in the supplementary material of (165). Reproduced from the American Society for Microbiology.

Figure 35. AID expression and Nuclear Cytosine Deamination Activity in CH12F3 Cells

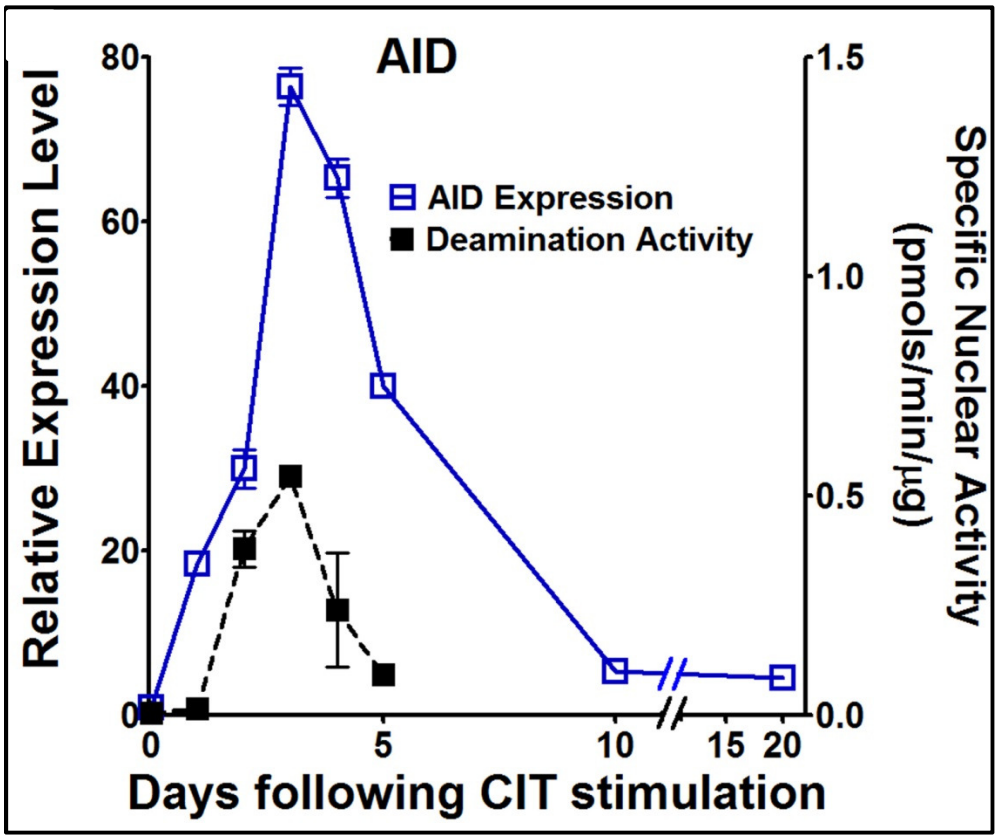

Figure 35. AID expression and Nuclear Cytosine Deamination Activity in CH12F3 Cells. Time course of change in AID gene expression and cytosine deamination activity in nuclear extracts prepared from CIT-stimulated CH12F3 cells. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

Figure 36. UNG2 expression and Uracil Excision Activity in CH12F3 Cells

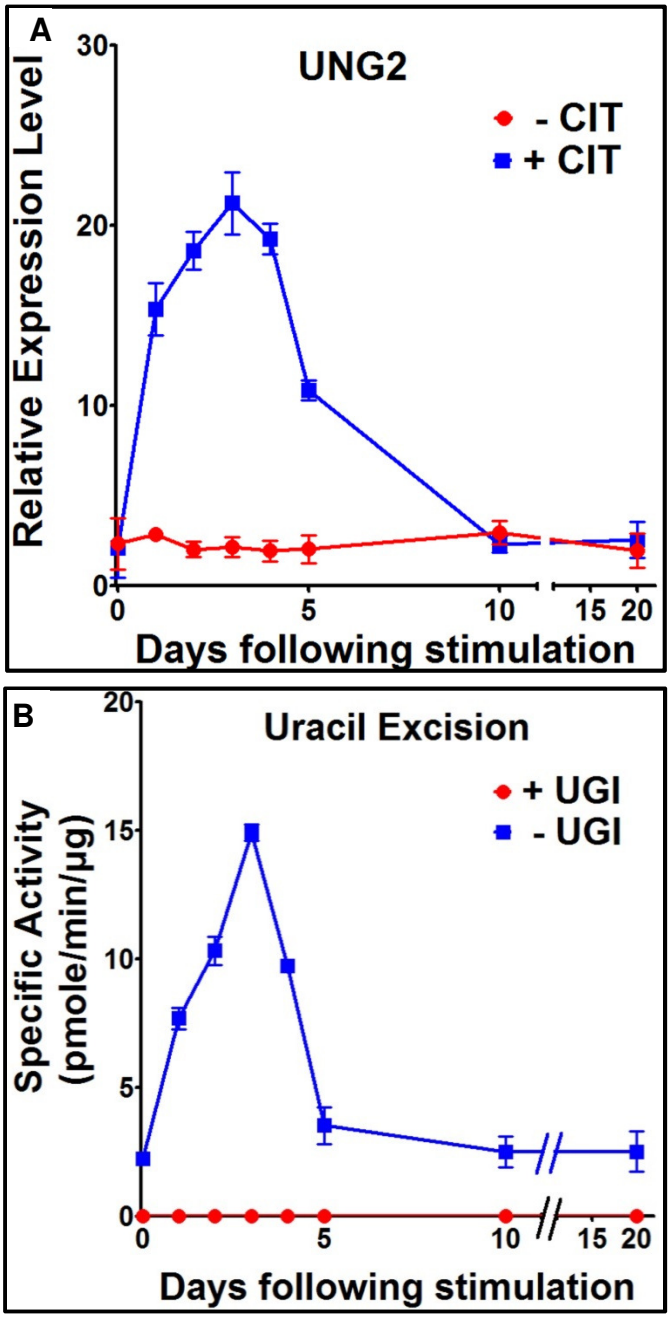

Figure 36. UNG2 expression and Uracil Excision Activity in CH12F3 Cells. (A) UNG2 expression in CH12F3 cells relative to GAPDH gene expression (set at 100). (B) Time course of change in uracil excision activity in nuclear extracts prepared from stimulated CH12F3 cells in the presence or absence of UGI. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

Figure 37. SMUG1 expression in CH12F3 Cells

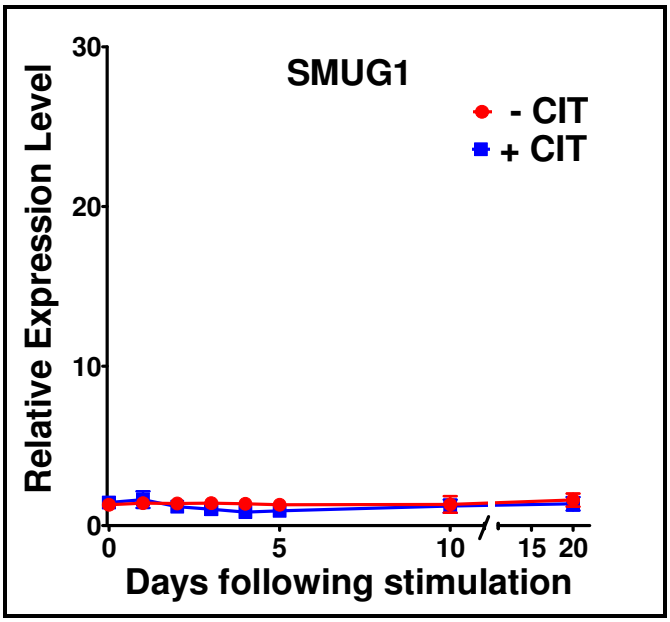

Figure 37. SMUG1 expression in CH12F3 Cells. SMUG1 expression in stimulated (+CIT) and unstimulated (-CIT) CH12F3 cells relative to GAPDH gene expression (set at 100).

To directly compare gene expression in CH12F3 cells with gene expression in murine splenocytes, qRT- PCR was performed on equal amounts of cDNA from the two cell types using the same set of primers and internal controls. The results showed that in both of the cell types, AID and UNG2 gene expression peaked at day 3 and the peak mRNA levels of both the genes were about 20% to 30% lower in CH12F3 cells than in the splenocytes (see Figure 34 B and C). To compare the UNG2 excision activities of the two cell types, I performed parallel uracil excision assays with nuclear extracts from WT splenocytes and CH12F3 cells prior to stimulation and at 3 days following stimulation and electrophoresed the products on the same gel. The gel showed that CH12F3 cells had uracil excision activity comparable to that of WT splenocytes (Figure 38). Thus, stimulated CH12F3 cells show AID gene

expression, deamination activity, UNG2 gene expression, and uracil excision activity similar to those seen in stimulated WT splenocytes but accumulate uracils like stimulated UNG-/- splenocytes.

Figure 38. Nuclear Uracil Excision Activity is similar in Stimulated Primary and CH12F3 B cells

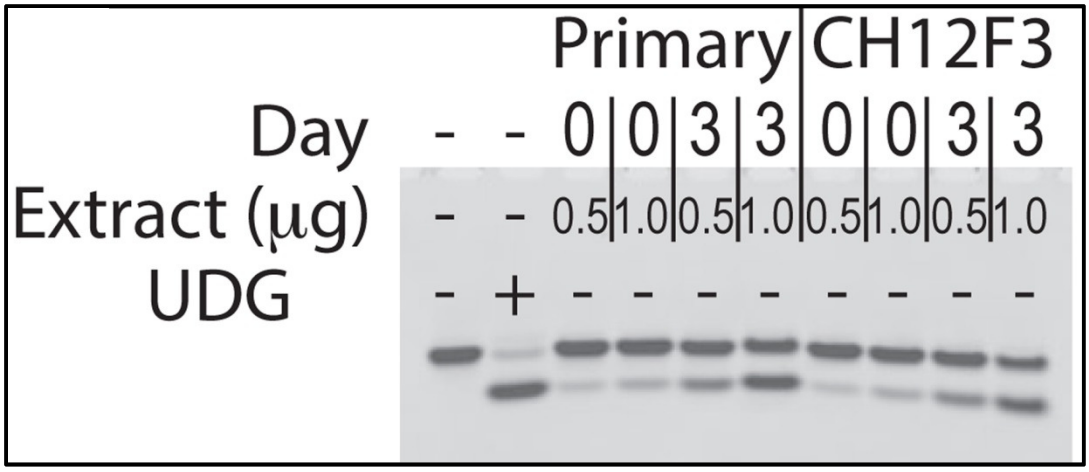

Figure 38. Nuclear Uracil Excision Activity is similar in Stimulated Primary and CH12F3 B cells. Uracil excision activity in nuclear extracts from stimulated WT splenocytes (+LPS) and CH12F3 cells (+CIT). DNA containing uracil was treated with the indicated amount of extract, and the products were electrophoresed on a denaturing gel following strand cleavage. “Day 0” indicates unstimulated cells, and “Day 3” indicates cells stimulated for 3 days. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

The uracil accumulation detected in CH12F3 cells only occurs upon CIT stimulation, and therefore induction of AID. However, some GC-derived cancers constitutively and ectopically express AID (see TABLE 2). The J558 cell line, a murine plasmacytoma cell line that constitutively expresses AID (184), was chosen to investigate whether its genome contains uracils like the stimulated CH12F3 cell genome. Both AID and UNG2 genes were expressed in J558 at levels between unstimulated and stimulated WT splenocytes (Figure 39A). The

J558 genome contained ~25-fold as much uracil as unstimulated UNG-/- splenocytes, and this amount was less than what was found in the stimulated UNG-/- splenocytes tested in parallel experiments (Figure 39B). Thus, this murine cell line derived from a B cell myeloma expresses UNG2, but still exhibits higher uracil content than B cells that have not undergone stimulation mimicking maturation in germinal centers. These murine cancer cell line results suggest that the loss of balance between uracil creation and elimination may be a common feature of B cell cancers expressing AID.

Figure 39. J558 Gene Expression and Uracil Levels

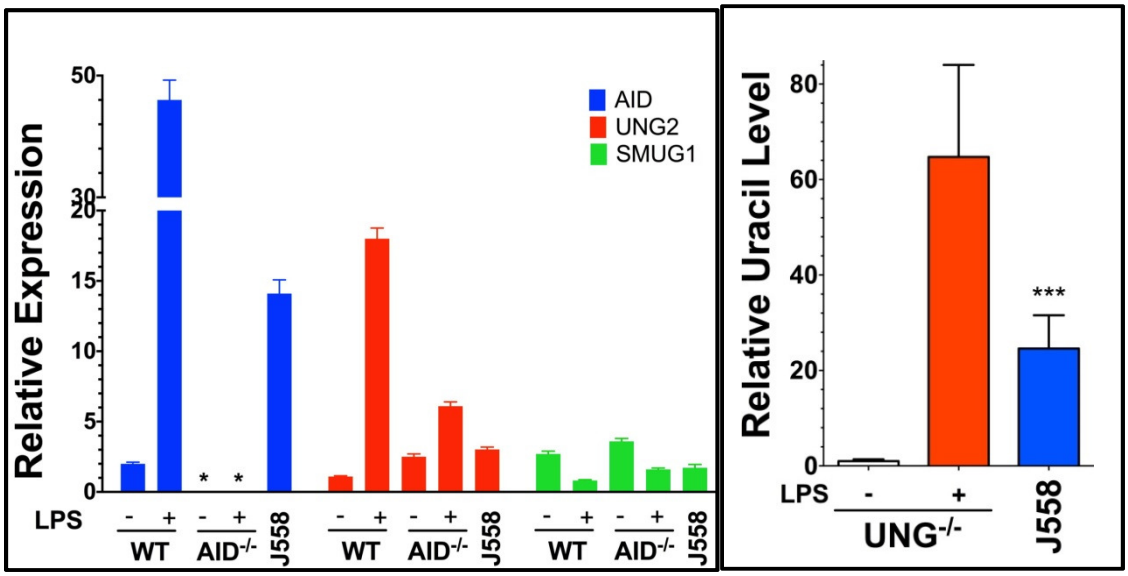

Figure 39. J558 Gene Expression and Uracil Levels. (A) Gene expression levels in J558 and splenocytes. Expression Levels of AID, UNG2 and SMUG1 were determined for stimulated and unstimulated WT and AID-/- splenocytes and J558. The expression levels of the 3 genes are relative to GAPDH expression level (set at 100).(B) Quantification of uracils in the genome of UNG-/- mice prior to stimulation (-LPS) and at day 3 following stimulation (+LPS) compared to uracil levels in the genome of the mouse plasmacytoma cell line, J558. Data from triplicate samples are shown with error bars representing standard deviation. (*** P< 0.0001)

3.2.1 Uracil accumulation in human B cell Lymphoma lines

To determine whether the correlation observed between high AID levels and high genomic uracil occurrence, despite high UNG2 expression, in murine B cell cancers also extends to human B cell cancers, I studied 13 cell lines created from human hematopoietic malignancies (see TABLE 4). Specifically, seven of the cell lines tested are B cell lymphomas of GC origin, including two BL lines, two FL lines, and three DLBCL lines. In addition, a MM line and two T lymphoma/leukemia lines were also studied. Finally, three cHL lines were also studied. Although many cHL tumors show some evidence of GC development (185), they usually do not express AID at high levels (137). PBMCs from healthy blood of human volunteers were used as negative controls serving as non-AID expressing B cells.

The GC-derived cell lines (ie. BL, FL, and DLBCL) had significantly higher levels of AID expression and genomic uracils when compared to the non-GC derived or characteristic cell lines. The BL, DLBCL, and FL cell lines express AID at levels ranging from about 4 to 8% relative to GAPDH gene expression (Figure 40, top panel, closed bars). This is in contrast to the non-GC-derived lines, cHL lines, and normal cells which had less than 1% AID gene expression (Figure 40 top panel, open bars). The nuclear DNA-cytosine deamination activities were consistent with the AID mRNA levels in the cell lines (see Figure 40 and 41).

Figure 40. AID Expression and Uracils in Human Lymphoma Cell Lines

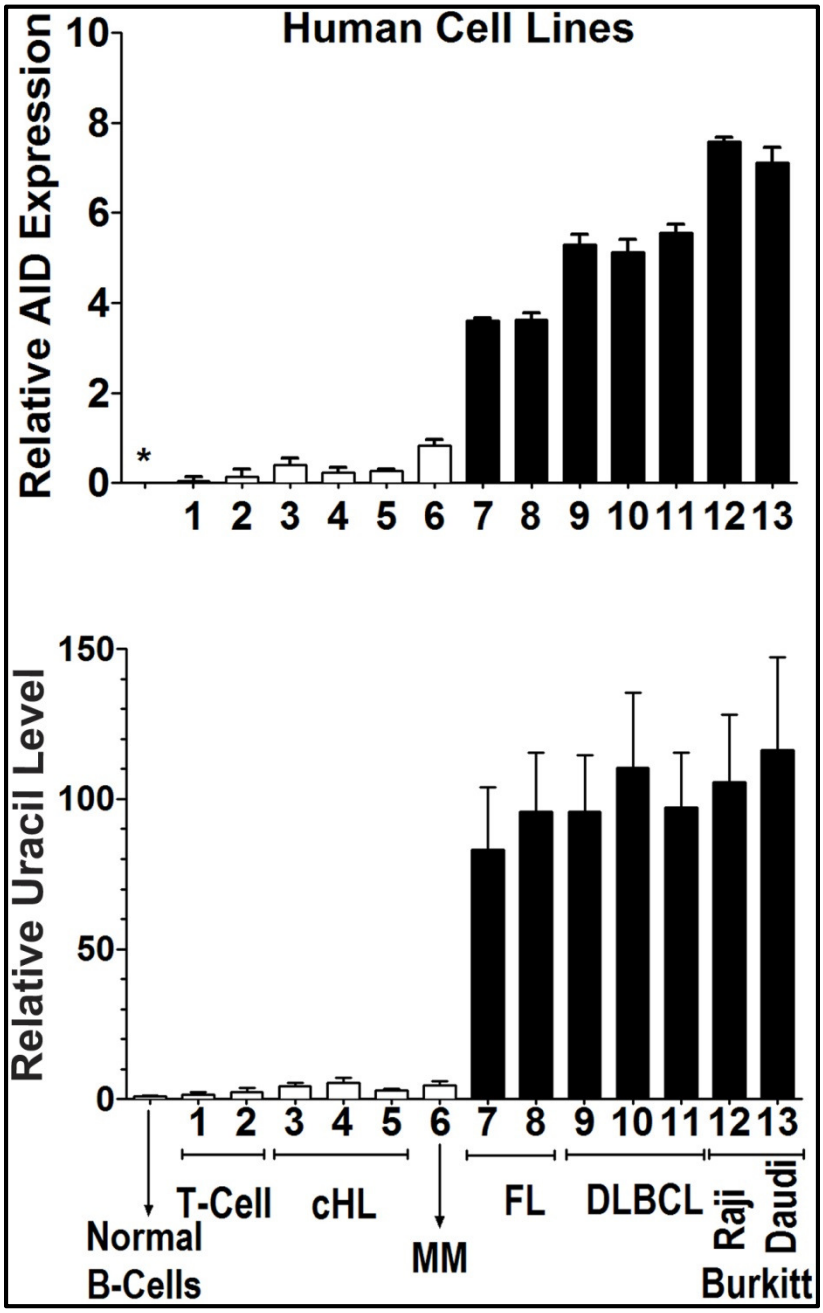

Figure 40. AID Expression and Uracils in Human Lymphoma Cell Lines. AID gene expression and genomic uracil levels in 13 human lymphoma-derived lines and circulating human B cells. (Top) The AID expression levels are shown relative to the GAPDH gene expression level (set at 100). Numbers on the x-axis refer to the specific cell line detailed in TABLE 4. The asterisk indicates no detectable expression. (Bottom) Genomic uracil levels relative to those in circulating human B lymphocytes (set at 1.0). This figure was originally published in (165). Reproduced from the American Society for Microbiology.

Figure 41. Human Lymphoma Nuclear Cytosine Deamination Activity

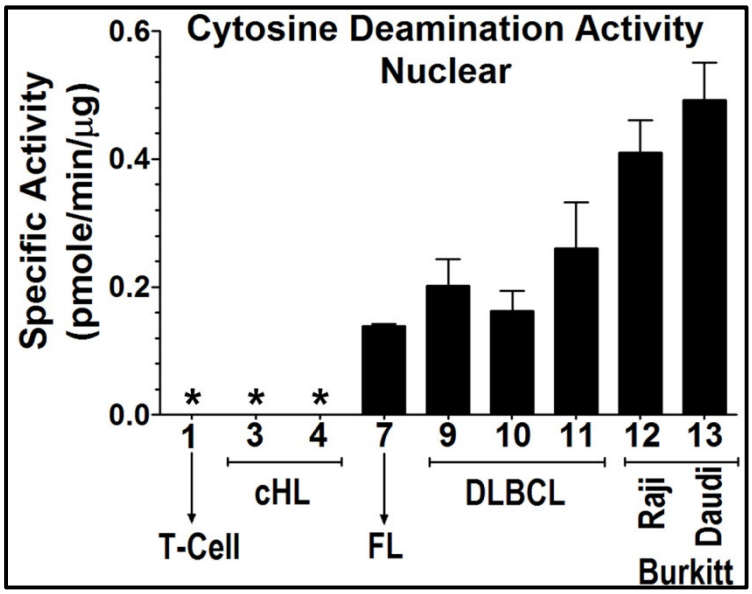

Figure 41. Human Lymphoma Nuclear Cytosine Deamination Activity. Nuclear cytosine deamination activity for some of the cell lines. The asterisks indicate that no deamination activity was detected under these conditions. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

The genomic uracil levels in the high-AID expressing cell lines, BL, FL, and DLBCL, were quite high compared to the MM, cHL and T-cell lines. Specifically, the GC-characteristic lines had genomic uracils 80- to 120- fold the level found in normal PBMCs (Figure 40, bottom panel, closed bars). In contrast, some of the low-AID B cell lines contained genomic uracils at levels that were only two to five times those seen with normal PBMCs while the tested T cell lines had uracil levels comparable to normal B cells (Figure 40, bottom panel, open bars). When the genomic uracil levels in these cell lines were plotted against their AID gene expression levels, the high-AID and low-AID lines clearly separated into distinct groups (see Figure 42A). Stimulated tonsillar cells had

somewhat higher uracil levels than peripheral B cells, and the low-AID lines had uracil levels comparable to those found in the tonsillar cells (see Figure 42B).

Figure 42. Uracil versus AID in Human B Lymphoma Cell Lines

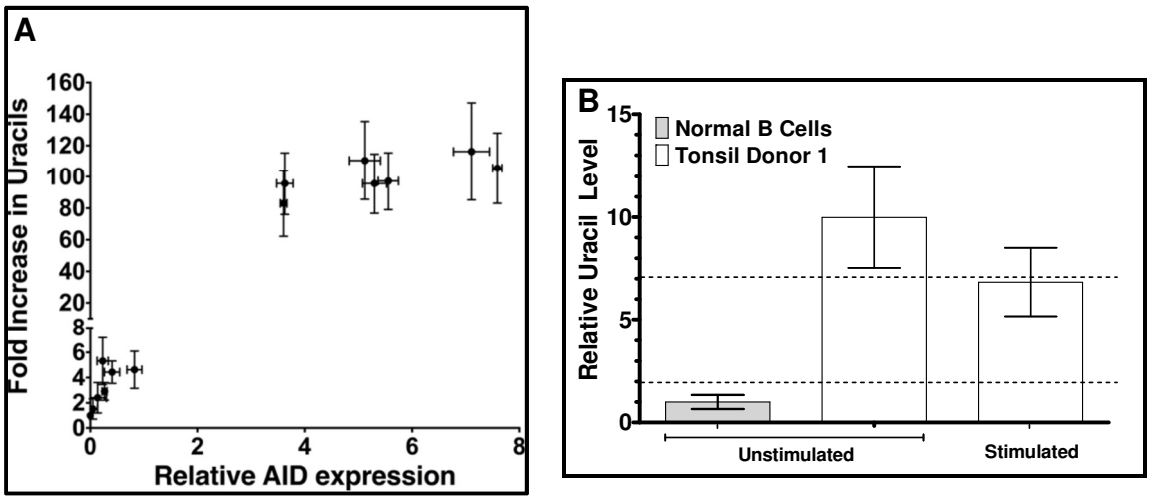

Figure 42. Uracil versus AID in Human B Lymphoma Cell Lines. (A) The genomic uracil levels in the human lymphoma cell lines are plotted against AID mRNA expression levels. (B) The genomic uracil levels in naïve B cells purified from tonsil donor 1 without stimulation and 4 days post ex vivo stimulation with CD-40L and IL-4 are shown relative to levels in circulating normal B cell DNA isolated from peripheral healthy donor blood. Normal B-cells are set to 1.0. (mean ± S.D. from triplicates) The dashed lines indicate the upper and lower values of the relative uracil levels for the low-AID group from Figure 5A to indicate the genomic uracil level range found among the tested T-cell, classical Hodgkins Lymphoma, and Multiple Myeloma cell lines. This figure was originally published in the supplementary material of (165). Reproduced from the American Society for Microbiology.

Since it is an inherent property of cancers to have genomic instability (186), it was possible that the B cell cancer cell lines were not stable with regard to genomic uracil content. In particular, it is possible that continual cytosine deamination by AID increases the uracil content of high- AID cell lines with time. However, when the Raji cell line was followed over many passages, no consistent increase or decrease in the uracil content was observed (Figure 43).

Thus, the B cell lymphoma- derived high-AID cell line displayed stable uracil content, albeit at much higher levels than normal B cells.

Figure 43. Stable Uracil Levels in Raji

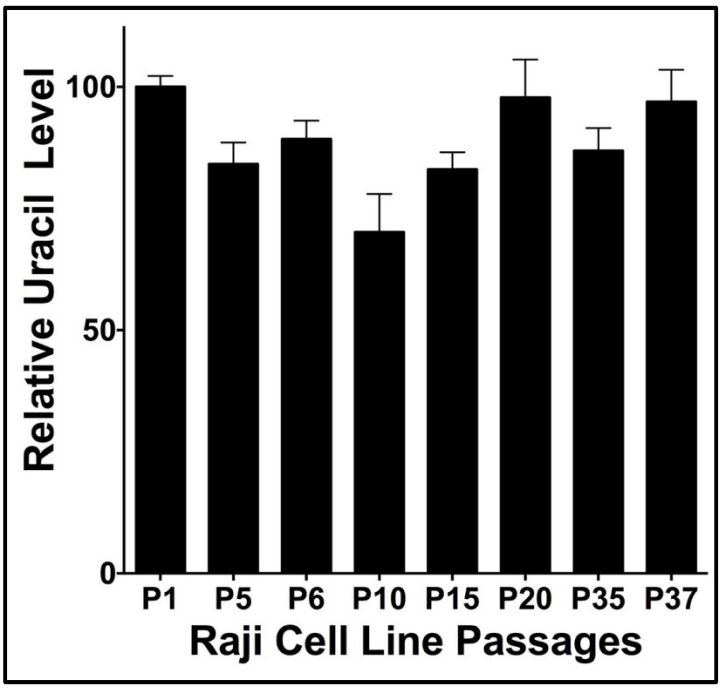

Figure 43. Stable Uracil Levels in Raji. Uracil levels in the Raji cell genome over several passages. The uracil level in Raji cells at passage 1 (P1) was set at 100. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

3.2.2 Uracil excision activity in human B cell cancers overexpressing AID

One possible cause of elevated uracil levels in high-AID cell lines and tumor samples is a reduction in uracil excision activity. To test this possibility, I performed qRT-PCR measurements of UNG2 and SMUG1 mRNA in the human cell lines in addition to uracil excision activity assays on several cell lines (Figure 44). The mRNA quantification studies showed that high-AID cell lines also expressed UNG2 and SMUG1 genes at moderately higher levels than PBMCs (Figure 44A). While the higher UNG2 expression seen in these cells was similar

to the increased UNG2 expression seen in stimulated WT murine splenocytes and CH12F3 cells, the increase in SMUG1 expression was not observed in murine cells (Figure 27 and 37).

Figure 44. UNG2 expression and Uracil Excision Activity in human lymphoma cell lines

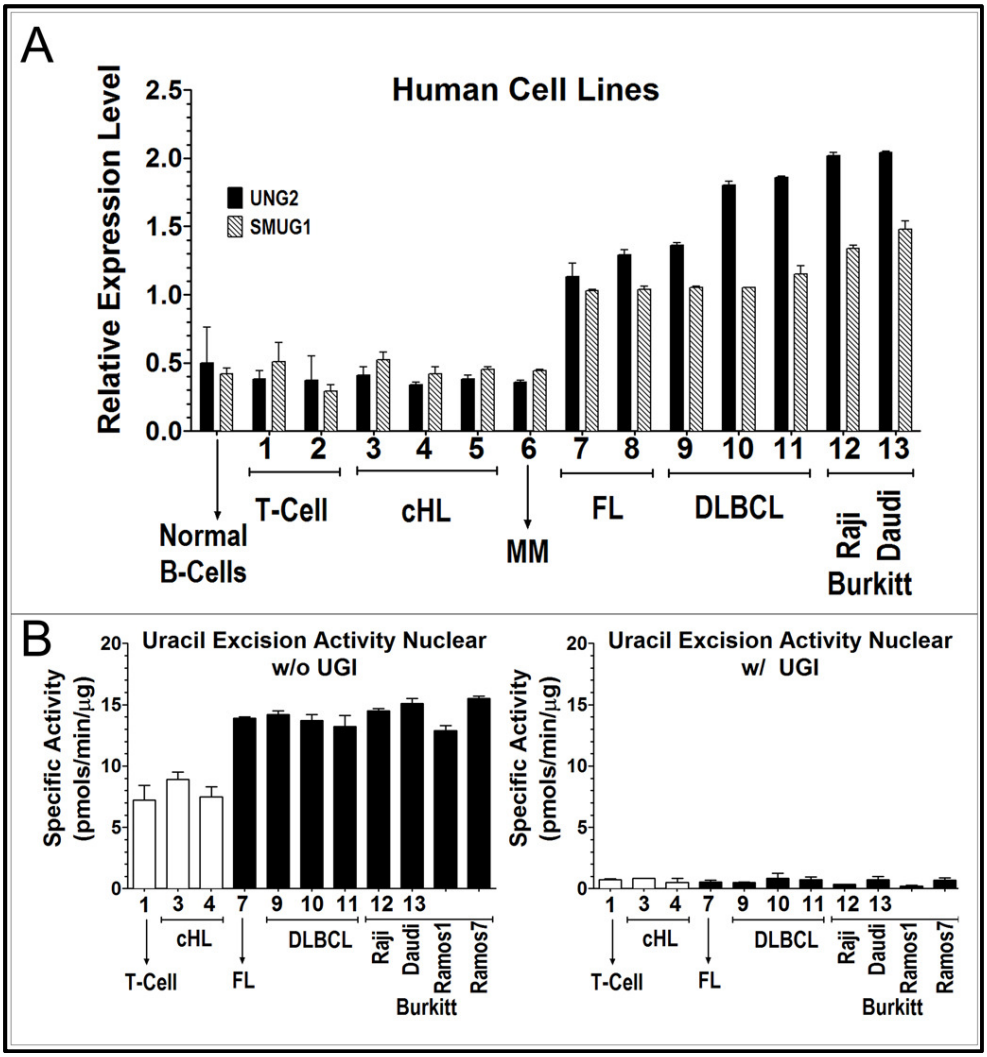

Figure 44. UNG2 expression and Uracil Excision Activity in human lymphoma cell lines. (A)The UNG2 and SMUG1 expression levels in the same human cell lines as described for Figure 40 relative to the GAPDH expression level.(B) Uracil excision activity in nuclear extracts prepared from a subset of the human lymphoma cell lines in the absence UGI (left) or in its presence (right). This figure was originally published in (165). Reproduced from the American Society for Microbiology.

I performed the uracil excision activity assay for two low- AID lines and seven high-AID lines in the presence or absence of UGI. The high-AID nuclear extracts showed 50% to 100% more uracil excision activity than the low-AID extracts tested (Figure 44B, left panel). This activity was abolished when UGI was included in the reactions (Figure 44B, right panel), confirming that it was due to UNG2. Furthermore, the specific activities of the nuclear extracts from high-AID cell lines were comparable to the peak activity seen in nuclear extracts from murine splenocytes (compare Fig. 29 and 44B). These results show that the UNG2 gene expression is not defective in high-AID human B cell cancer cell lines and that they contain nuclear uracil excision activity comparable to that seen with stimulated WT murine splenocytes. Despite this activity, they contain very high levels of uracils in their genome.

3.2.3 Genomic uracil levels change with AID

To determine whether increases or decreases in the expression of AID in human cell lines causes commensurate changes in genomic uracil levels, I used previously generated variants of the Ramos cell line that express AID at different levels (166).These included cell lines Ramos 7 (high AID gene expression) and Ramos 1 (low AID gene expression) (166) and lines created through transfection of a plasmid expressing AID into Ramos 1 (Ramos A.1, A.2, and A.5) or the empty vector (Ramos C.1) (43).

The expression of AID was much higher in Ramos 7 than in Ramos 1 cells and was higher in Ramos A.2 and A.5 cells than in Ramos C.1 cells (Figure 45, top panels) as expected. Correspondingly, the genomic uracil level in Ramos 7

cells was higher than in Ramos 1 cells, and there was a similar relationship between genomic uracil levels in the Ramos A.2 and A.5 lines and in Ramos C.1 cells (Figure 45, bottom panels). Thus, Ramos-derived cell lines with higher AID expression show a pattern of higher genomic uracil levels as seen in the high-AID group of cell lines (Figure 40). To test whether the catalytic activity of AID was required for the increase in genomic uracils, I transfected a catalytically inactive version of AID (E58A) into Ramos 1 cells. Both WT AID and E58A AID gene transfection increased AID mRNA levels (Figure 46, left panel), but only the catalytically active version increased genomic uracil levels (Figure 46).

Figure 45. AID and Uracils in BL Ramos Derived Lines

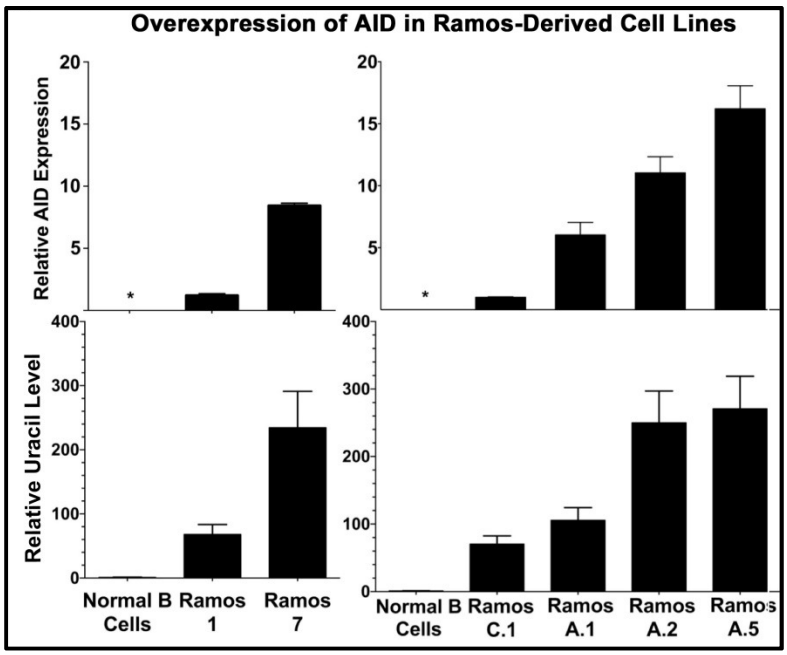

Figure 45. AID and Uracils in BL Ramos Derived Lines. (Top panel) AID expression in Ramos 1 and Ramos 7 lines and in the Ramos 1-derived lines transfected with empty vector (Ramos C.1) or with an AID expression plasmid (Ramos A.1, A.2, and A.5). (Bottom panel) Genomic uracil levels in the same cell lines. In both cases, the numbers were normalized relative to uracils in normal circulating human B cells (set to 1.0). The asterisks indicate undetectable expression. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

Figure 46. WT and E58A AID

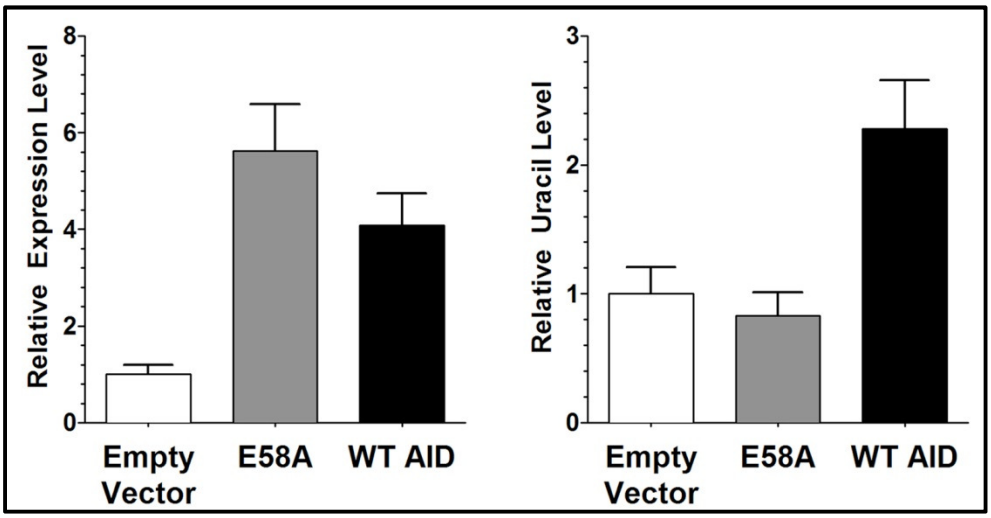

Figure 46. WT and E58A AID. (Left panel) Relative expression levels of WT AID and the AID E58A mutant following transient transfection of Ramos 1 cells. The AID level in Ramos 1 transfected with an empty vector was set to 1.0. (Right panel) Genomic uracil levels in the same transient-transfection experiment relative to uracils in the empty vector control (set to 1.0). This figure was originally published in (165). Reproduced from the American Society for Microbiology.

Interestingly, UNG2 gene expression was highest in the three BL cell lines expressing AID at very high levels, Ramos 7, Ramos A.2, and Ramos A.5 (see Figure 47A). This suggests that when Ramos cells were transfected with AID-expressing plasmids, endogenous UNG2 gene expression levels may have lagged behind increasing AID expression levels, allowing the genomic uracil levels to rise, but that the uracil excision ability eventually caught up with uracil generation by AID stabilizing the genomic uracil content at a higher level than before.

Figure 47. Expression of Uracil Repair Enzymes in Human Cell Lines

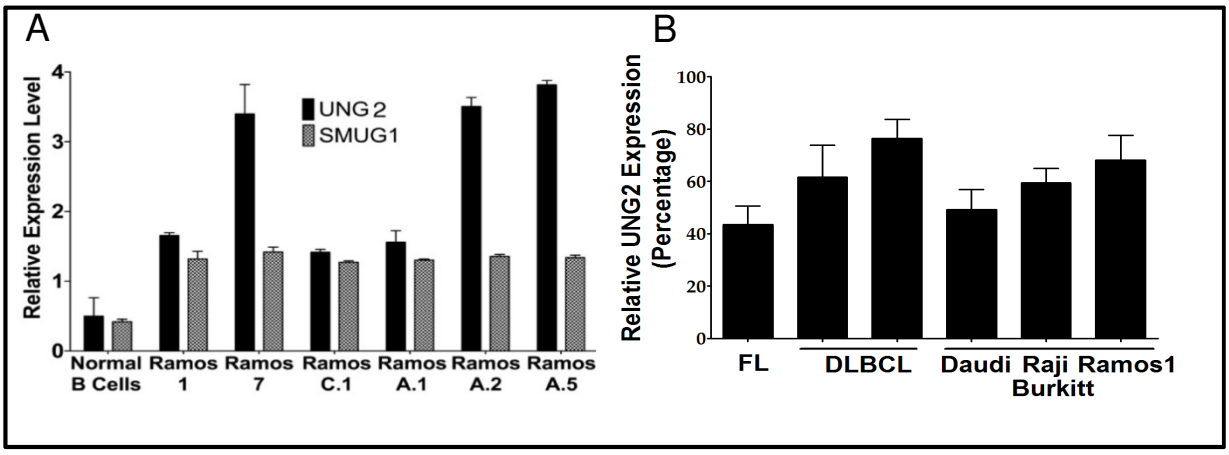

Figure 47. Expression of Uracil Repair Enzymes in Human Cell Lines. (A) UNG2 and SMUG1 gene expression in Ramos-derived cell lines. mRNA levels for Ramos 1 and Ramos 7 lines, and the Ramos 1-derived lines transfected with empty vector (Ramos C.1) or with an AID expression plasmid (Ramos A.1, A.2 and A.5) are shown. The levels are normalized to GAPDH levels set at 100. (B) UNG2 expression levels in AID-knockdown cell lines. UNG2 mRNA levels are presented as percentage of mean values for the negative controls expressing scrambled-shRNA. This figure was originally published in the supplementary material of (165). Reproduced from the American Society for Microbiology.

To study the effects of reducing AID gene expression on genomic uracil levels, I transformed human cell lines expressing AID at high levels with plasmids expressing AID-specific shRNA. The levels of success of the knockdown (KD) differed among the cell lines, but AID gene expression in all the lines was reduced by 75% or greater compared to the expression seen with a control line expressing scrambled shRNA (Figure 48, top panel). All the AID KD lines had substantially reduced genomic uracil levels compared to their scrambled shRNA counterparts, but the degrees of reduction were different for different cell lines (Figure 48, bottom panel). The variability in the reduction in genomic uracil levels among different KD lines may have been due to differences in genetic backgrounds between the cell lines and the fact that they represent three

different B cell lymphomas. For example, SMUG1 levels or dUTP pools may be different in different cell lines and this may have secondary effects on genomic uracil levels. It should be noted that UNG2 gene expression changed in the same direction as the change in AID gene expression in these cell lines (see Fig.47B).

Figure 48. Knock-down of AID in Human B cell Lines

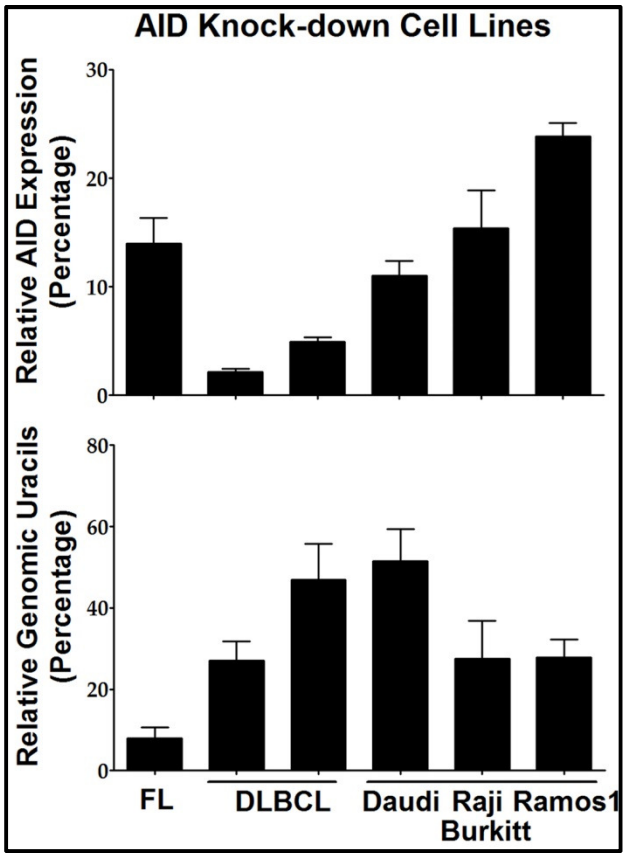

Figure 48. Knock-down of Human B cell Lines. Effects of AID knockdown on human lymphoma cell lines. The AID mRNA levels and genomic uracil levels are presented as the percentages of the mean values for the line expressing scrambled shRNA. (Top panel) AID expression in cell line expressing AID-specific shRNA. (Bottom panel) Genomic uracil levels in the same cell lines. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

3.2.4 Genomic Uracils in AID-expressing patient B cell lymphomas

I tested tumor samples from six pretreatment lymphoma patients (referred to as patients P5 through P10) (see TABLE 5) and determined the levels of AID gene expression and genomic uracil in the B lymphomas. AID gene expression was higher in FL and DLBCL patient tumors than in the three MZL patient samples (Figure 49A, open bars). MZL tumors are thought to originate from post-GC marginal- zone B cells and thus, the low AID gene expression in these tumors is consistent with their origin. The patients from whom the tumor samples were obtained were unrelated individuals, and their cancers were at different stages and had different cytogenetic markers (see TABLE 5). Consequently, it is not surprising that the patients within each group (patients with DLBCL or FL and patients with MZL) showed considerable variation in AID gene expression levels (Figure 49A). Regardless, the genomic uracil levels were much higher in FL and DLBCL tumors than in normal B cells and the MZL tumors. The uracil levels in FL and DLBCL tumors were comparable to those seen in most FL and DLBCL cell lines (compare Figure 40A and 49A, closed bars). In contrast, the MZL tumors had amounts of uracils in their DNA that were only slightly higher than those seen with normal B cells (Figure 49A, closed bars).

In addition, I also tested tumor samples from eight patients with CLL (patients P1 through P4 and P11 through P14, see TABLE 5). CLL tumors are known to be heterogeneous with respect to AID gene expression and I also detected a range of AID gene expression in the CLL patient samples. In contrast, there was little tumor-to-tumor variation in UNG2 and SMUG1 levels (Figure

49C). However, there was a strong correlation between AID gene expression and genomic uracil levels (Figure 49B). The correlation observed between elevated AID gene expression and accumulation of uracils in the genomes of human cell lines holds also for the patient tumor samples.

Figure 49. Gene Expression and Genomic Uracil Levels in Patient Tumor Samples

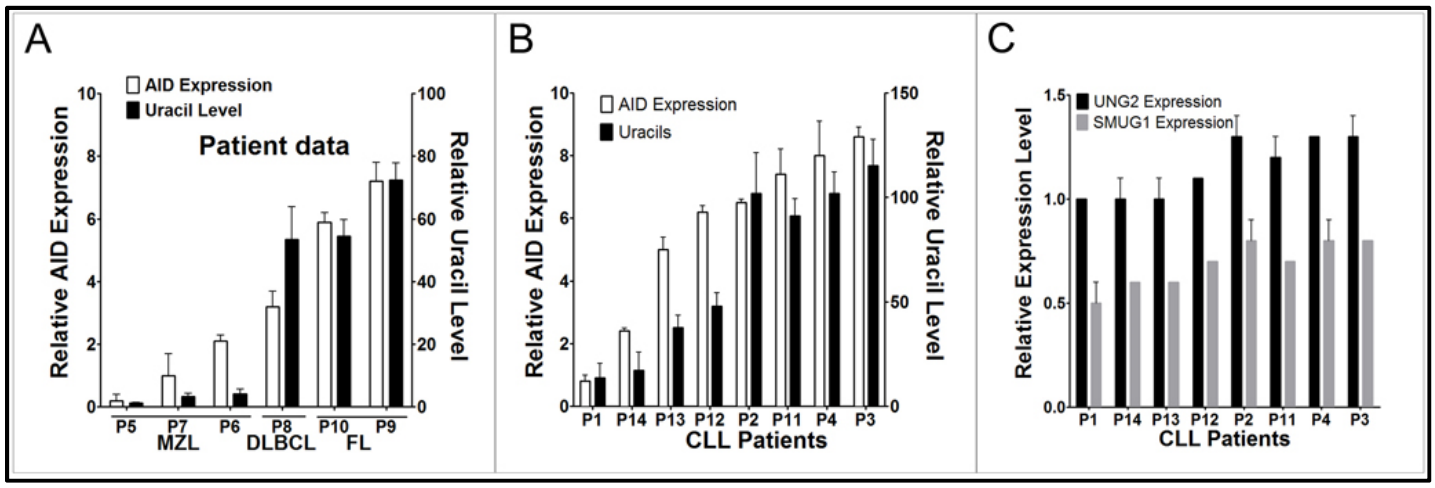

Figure 49. Gene Expression and Genomic Uracil Levels in Patient Tumor Samples. The patients and the numbering are described in TABLE 5. (A) AID expression levels and uracil levels in lymphoma patient samples. Genomic uracil levels are indicated relative to those in normal circulating human B lymphocytes (set at 1.0). (B) AID expression levels and uracil levels in CLL patient samples. (C) UNG2 and SMUG1 gene expression in the same CLL patient B cells. Expression data are normalized to GAPDH expression set at 100. This figure was originally published in (165). Reproduced from the American Society for Microbiology.

3.3 AIM 3- AID- dependent DNA Damage in B cell Lymphomas

As discussed above, AID is essential for normal B cell development and required for SHM and CSR (see section 1.1). The point mutations and strand breaks introduced during these processes result from the actions of a number of enzymes in addition to AID, including UNG, APE, MMR proteins, and TLS polymerases (see section 1.2) (187, 188). Although the regulation of the AID

gene and its transcript ensure that it is principally expressed in GC B cells, results from Aim 1 of this work suggest that murine splenocytes (see section 3.1) stimulated to overexpress AID introduce uracils at a large number of genes, not necessarily restricted to the Ig genes. These uracils are subsequently eliminated by an UNG-dependent excision repair process (165). Additionally, several studies have shown that AID can target many non-Ig genes introducing mutations including oncogenes such as MYC, PAX-5, PIM1, and BCL2 in B lymphocytes (178, 189).

Results from Aim 2 reveal that B cell lymphomas that overexpress AID maintain a high level of uracil in their genomes (165). To test whether the downstream processing of the initial uracil lesions leads to an accumulation of other types of DNA damage, I subjected human lymphoma cell lines that overexpress AID and have high levels of genomic uracils to abasic site quantification, single- and double strand break analysis, and cell viability studies. Many of these cell lines contain high levels of these lesions in an AID-and Ung-dependent manner. Furthermore, results suggest this accumulated DNA damage may compromise genomic stability.

3.3.1 Expression of AID and accumulation of genomic uracils in cancer cell lines

I confirmed AID expression levels relative to TBP in six of the GC-derived lymphoma cell lines used in Aim 2 (see Figure 50A). AID mRNA levels in FL, DLBCL, and BL lines were 13 to 22 fold higher relative to the endogenous TBP mRNA levels in each sample. This is in distinction to AID levels in the T cell lymphoma and cHL cell lines where AID mRNA levels were below TBP

expression. AID mRNA was not detected in B cells isolated from non-cancerous donor peripheral blood. To specifically confirm deamination activity in the nucleus, I prepared nuclear extracts from these cell lines to assay nuclear cytosine deamination activities using a single-stranded fluorescently labeled DNA oligo with a cytosine in the sequence context preferred by AID. When products were electrophoresed following uracil excision of any deaminated cytosines and abasic site cleavage with alkaline and heat treatment, nuclear cell deamination activity was easily detected in the FL, DLBCL, and BL lines (Figure 50B, and Figure 41). No deamination product was observed in the T-cell extracts while one of the two tested cHL lines showed limited but detectable product. Thus, both the nuclear DNA-cytosine deamination activities and the AID mRNA expression levels are highest in the GC-characteristic B cell lymphomas confirming results in Aim 2.

3.3.2 Accumulation of AP sites in cancer cell lines

The genomes of lymphomas expressing elevated AID levels contain an excess of uracils in their genomes (see Figure 40), but these cells also express SMUG1 and UNG2 genes at normal or higher than normal levels (see Figure 44A). This implies that SMUG1 and UNG2 enzymes will attempt to excise this surplus of uracils establishing many abasic sites. Any subsequent processing of abasic sites by AP endonucleases can create strand breaks. To study this, I utilized a variety of biochemical assays.

Figure 50: AID expression and Cytosine Activity in Human Cell Lines

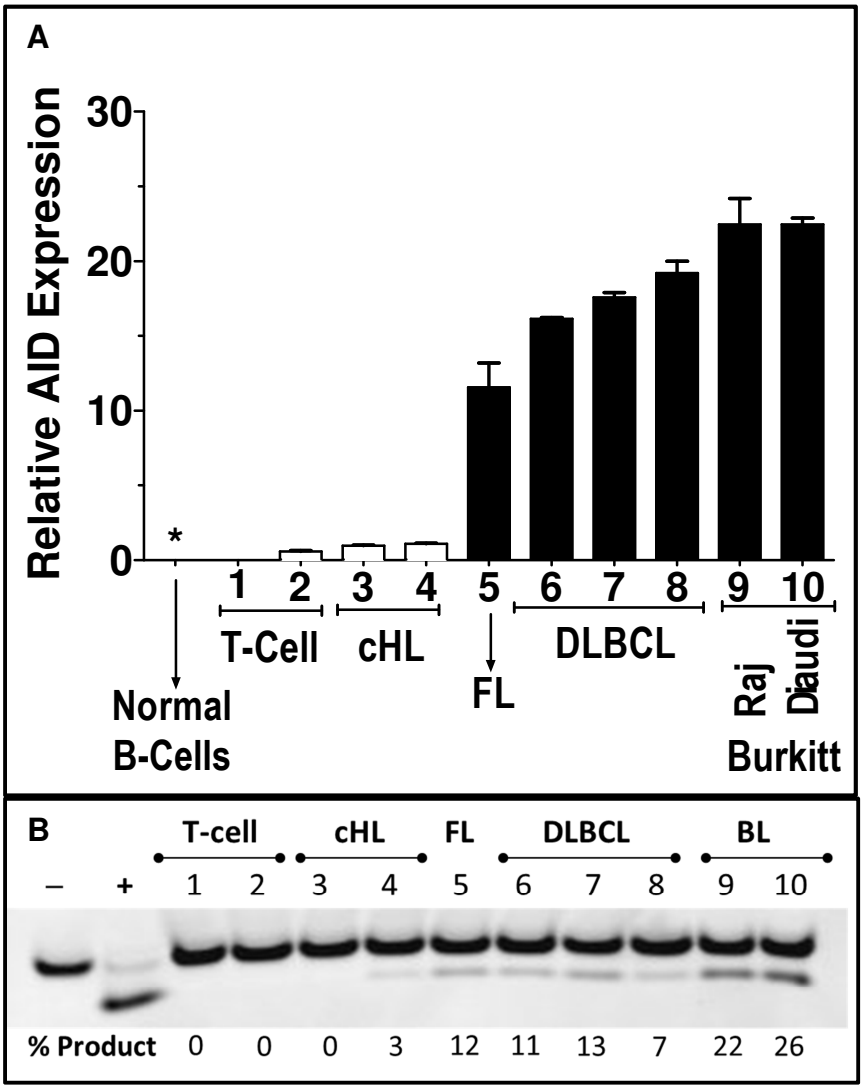

Figure 50: AID expression and Cytosine Activity in Human Cell Lines. (A) The AID expression levels are shown relative to TBP. Asterisk indicates no detectable expression in normal B-cells isolated from PBMCs of healthy donors. (B) An oligonucleotide DNA containing cytosine in AID's preferred sequence context was treated with equal amounts of nuclear extracts from each cell line and the products were electrophoresed on a denaturing gel following treatment with *E. coli* UDG to excise any deaminated cytosines and nicking with alkaline and heat treatment. The percent product is shown below the denaturing gel.

I quantified the AP sites in the DNAs by using a previously established method that also utilizes ARP (see section 2.1.6) (190). Briefly, endogenous abasic sites are biotinylated with ARP, which allows for quantification of the AP sites. It should be noted that this assay detects all aldehydic adducts in DNA including intact AP sites, AP sites cleaved by AP endonuclease or other enzymes and the formamido forms of a number of oxidation and alkylation-induced damage (191). As there is no indication that the amount of oxidation and alkylation damage to DNA changes when AID is abundant, I refer to the number of all sites labelled during this assay as abasic or AP sites for textual convenience. There are approximately 0.5 AP sites per million base pair in normal B cell DNA (donor age range 20 to 30 years) and T cell and cHL lines were found to contain slightly higher amounts at statistically insignificant levels. Strikingly, 10 to 20 AP sites per million base pair were detected in the high AID-expressing cell lines- FL, DLBCL, and BL. This is almost 20 to 40 fold higher than normal B cell AP site levels (Figure 51). In these high-AID expressing cell lines, there are one-fifth to one-tenth as many persistent AP sites as uracils in the DNA.

To study the effects of reducing AID expression on accumulated AP sites, I analyzed the level of AP sites in stable cell lines established by knocking down the expression of the AID gene with AID-specific shRNA versus scrambled-control shRNA in select high AID expressing lines (see Figure 48). This reduction in AID also reduced the detectable AP sites (Figure 52). A significant reduction in the number of AP sites in the scrambled shRNA group when compared to the

AID- knock down (KD) lines is observed. All of the high-AID expressing cell lines had a higher number of AP sites than normal B cells, and in each case, the number of AP sites did not exceed the number of genomic uracils in the DNA.

Figure 51: AP Sites in Human Cell Lines

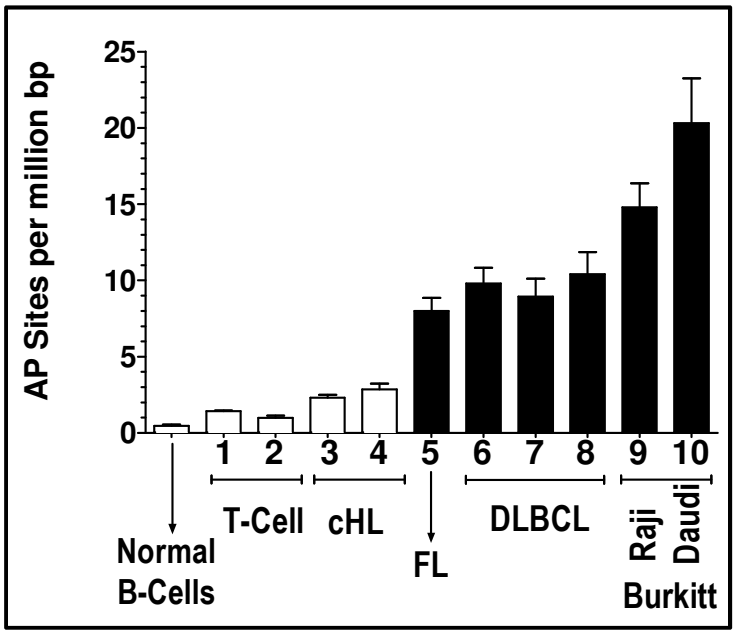

Figure 51: AP Sites in Human Cell Lines. Abasic sites are quantified using a biochemical detection method where DNAs are treated with aldehyde reactive probe to label endogenous abasic sites, followed by labeling with Cy5-streptavidin. (see Figure 19 for assay details). AP sites from cell line DNA were quantified in parallel with DNA from healthy peripheral blood B cells (normal B-cells, age 20-30).

3.3.3 Accumulation of Single- and Double Strand Breaks

To examine if single- and double strand breaks in the genome change in relation to AID levels, the Ramos 1 (low-AID) and Ramos 7 (high-AID) (see Figure 45 top left for AID expression) cell populations were subjected to the alkaline comet assay and a wide distribution of tail lengths was detected with tails visually longer for Ramos 7 than Ramos 1 (A. Martin, D. Haddad, University of Toronto) (Figure 53A and B). This is reflected in the quantified tail moments of

these two cell lines, where Ramos 7 had about twice as high a tail moment as Ramos 1.The tail moments of Ramos 1 cells increased to the level of Ramos 7 when the former cells were subjected to an ionizing radiation dose above 4 Gray (Figure 53B). Therefore, the higher levels of AID in Ramos 7 are correlated with a higher number of genomic single- and double strand breaks.

Figure 52: AP Sites in Human AID knock-down Cell Lines

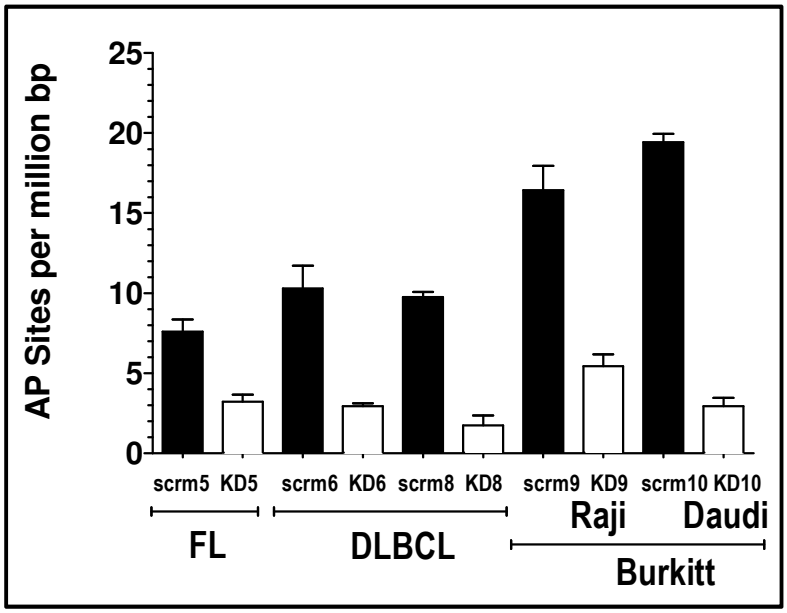

Figure 52: AP Sites in Human AID knock-down Cell Lines. AP sites were quantified in DNAs from stable cell lines prepared by transfecting the parental cell lines in Figure 51 with either a control scrambled shRNA (scrm) or an AID-specific shRNA (KD) to knock down expression.

Following removal of dead and dying cells from the cultures, the low- and high- AID expressing cell lines were stained with anti-γH2AX antibody to detect the marker for DNA DSBs. The FL, DLBCL, and BL cell lines frequently showed nuclei with several foci (Figure 54A, right panels). In contrast, the T-cell and cHL lines had fewer nuclei positive for staining (Figure 54B, left panels). The number

of foci per nuclei was counted using the Image J software and this confirmed the conclusions drawn from visual observations (Figure 54B).

Figure 53: AID dependent Strand Breaks

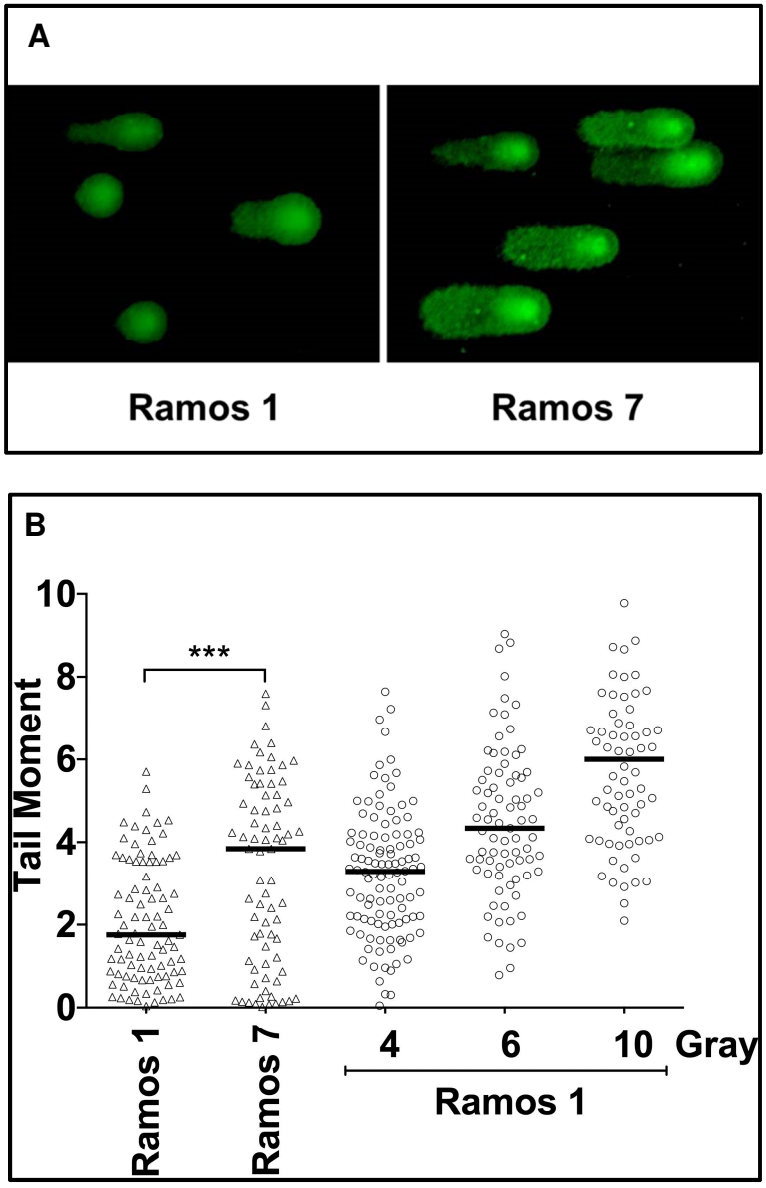

Figure 53: AID dependent Strand Breaks. (A) Representative false color alkaline comet assay images of Ramos 1 and Ramos 7 cells. Cells were synchronized in G1 with mimosine. (B) Tail moments of ~100 cells of Ramos 1 and Ramos 7 are plotted. Ramos 1 cells were irradiated at indicated grays shown below plot. Bars indicate median tail moment values. *** $p < 0.05$ (P value 0.009).

Figure 54: γH2AX and Double Strand Breaks

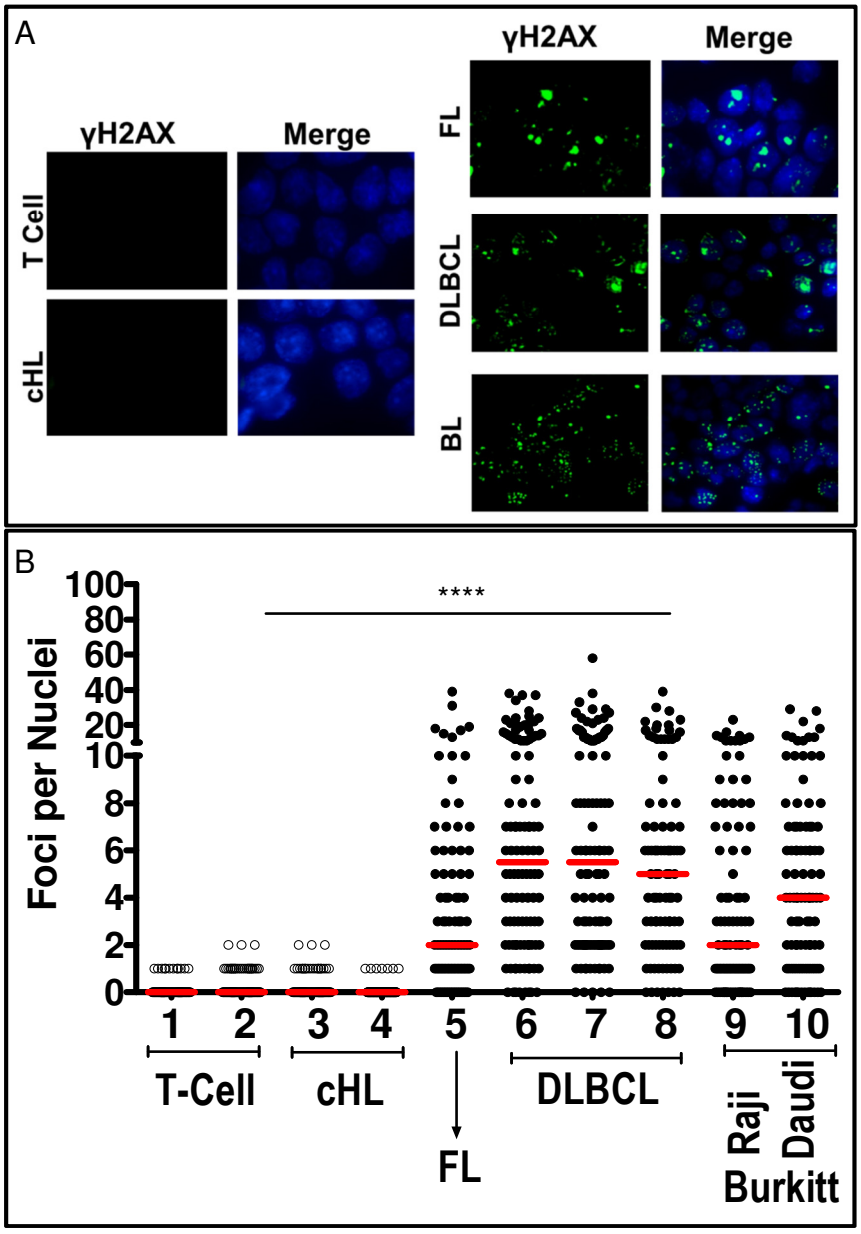

Figure 54: γH2AX and Double Strand Breaks. (A) (Left) Representative images of low-AID expressing cells and (Right) high-AID expressing lines stained for γH2AX (yellow). Cells were stained with DAPI (blue) and images merged. (B) Quantification of γH2AX foci per nuclei in lymphoma cell lines. Foci per nuclei were quantified for ~100 cells for each sample using ImageJ Software. Lines indicate median foci per nuclei values. **** P<0.0001 comparing low AID expressing foci and high-AID expressing foci.

The fairly wide variation in the number of foci and the existence of cells without any foci in every cell line population suggests that a single-strand break generated during the repair of AID-introduced uracils is converted to double strand breaks in only a fraction of the cells- possibly the cells actively replicating their DNA. Regardless, there is a clear correlation between high AID expressing cell lines and the accumulation of DSBs in their DNA.

To determine if the presence of DSBs is directly dependent on the presence of AID, AID-KD cell lines and the corresponding shRNA-control lines were also stained for γH2AX. A significant decrease in the mean number of foci per nuclei was observed when comparing the quantification of γH2AX foci staining of ~100 FL-KD cells with the FL-scrambled control cell line (Figure 55). This same pattern of decrease was seen in the KD versus shRNA-control cell lines established with DLBCL and BL lines (Figure 54) supporting AID's role in the promotion of the genome wide detected DSBs.

3.3.4 Overall Effect on Cell Viability

I initially noticed that the high-AID expressing cell lines grew at slower rates than the low-AID expressing cell lines. Further, the accumulation of uracils, AP sites, and strand breaks in the genomes of high-AID expressing cell lines promote genomic instability and may lead to increased cell death. To determine if this is the case, dead cells were removed and cultures with ~100% viability were grown for three days, stained with Trypan Blue, and counted using a TC20 Automated Cell Counter (Bio-Rad). The results show a correlation between AID expression and higher cell death. Cultures of cell lines with the highest AID

expression levels, FL, DLBCL, and BL, had the highest percentage of inviable cells (up to 30%), while those with the lowest AID levels, T-cell and cHL lines, remained close to 90% viable (Figure 56A) with a lower percentage of dead cells. Furthermore, by simply knocking down AID in the high AID-expressing cell lines, cell viability was restored to >85% viability when directly compared to the corresponding shRNA-control lines which maintain poorer viability grown under the exact conditions (Figure 56B).

Figure 55. DSBs Reduced in AID Knock-downs

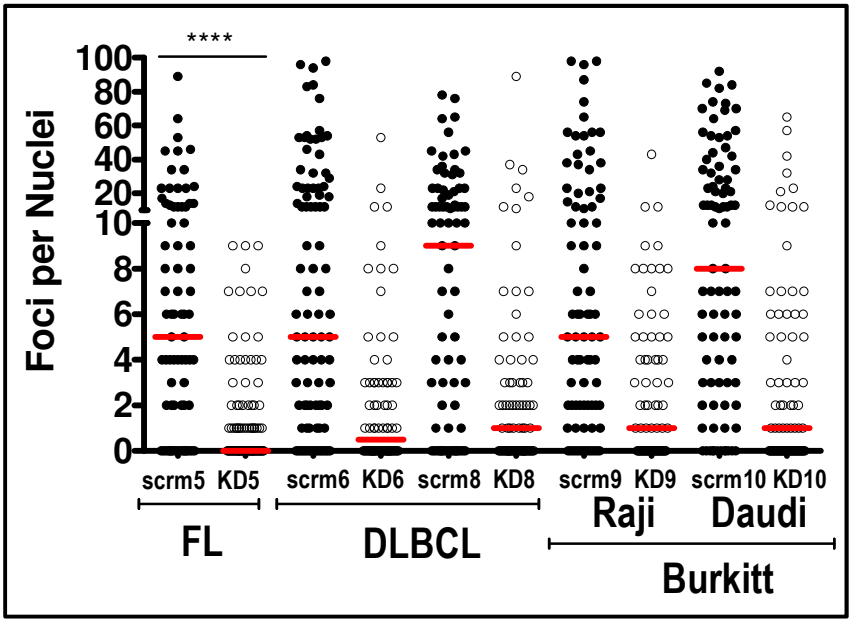

Figure 55. DSBs Reduced in AID Knock-downs. Quantification of γH2AX foci per nuclei in extablished cell lines where AID expression was knocked down (KD) with shRNA targeting AID versus scrambled control shRNA (scrm). **** P<0.0001

Figure 56. Viability of Human B Cell Lymphomas

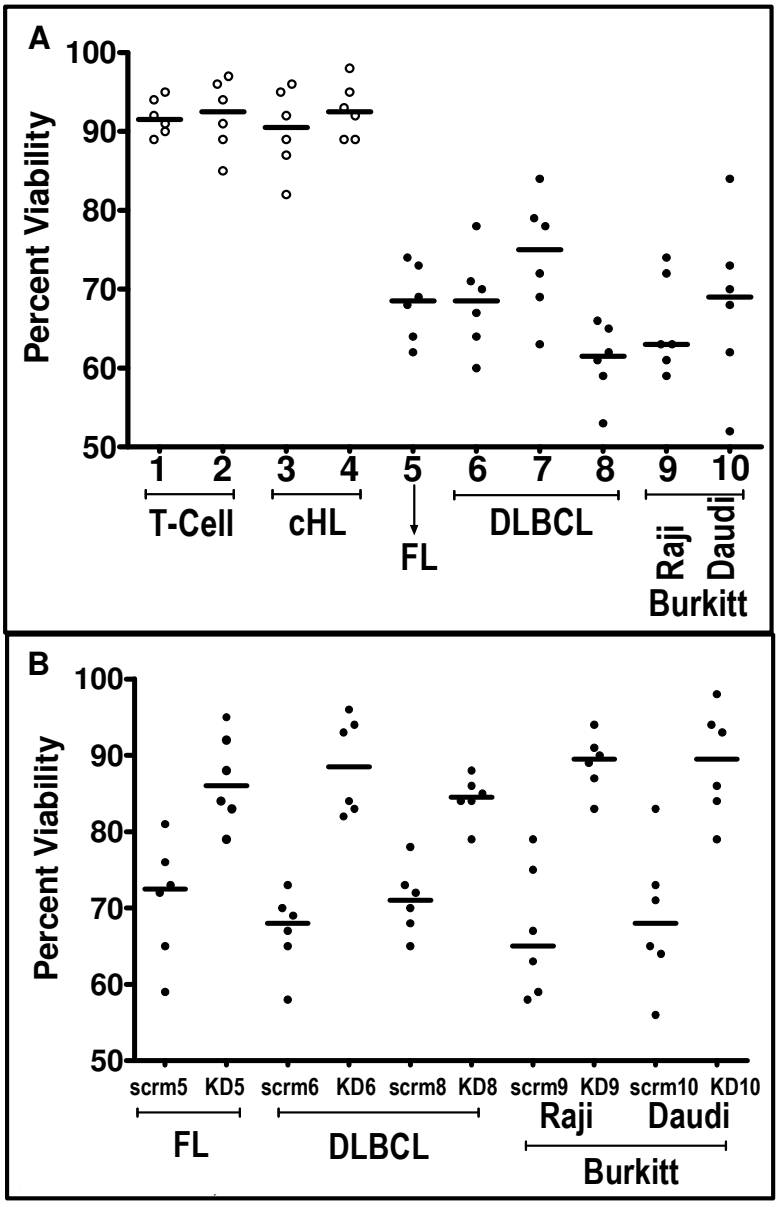

Figure 56. Viability of Human B Cell Lymphomas. (A) Percent viability in human B cell lymphomas. (B) Restoration of viability in extablished cell lines where AID expression was knocked down (KD) with shRNA targeting AID versus scrambled control shRNA (scrm). **** P<0.0001

With an increase in genomic instability and cell death culminating from AID-introduced uracils, I decided to determine the levels of uracils in inviable cells to see if dead/dying cells accumulated larger amounts of uracils or if this

population of cells has lower uracil levels from attempted repairs that overwhelmed the cells and lead to apoptosis. Viable cells were isolated from inviable cells in the FL, DLBCL, and BL cell lines using Annexin-V magnetic beads and uracils were quantified. Surprisingly, in every case there were more uracils in the genomes of inviable cells than viable cells (Figure 57). In most of the cell lines, a two-fold or higher level of uracil was detected in the inviable population as compared to the viable population.

Figure 57. Uracils in Viable and Inviable Human B Cell Lymphomas

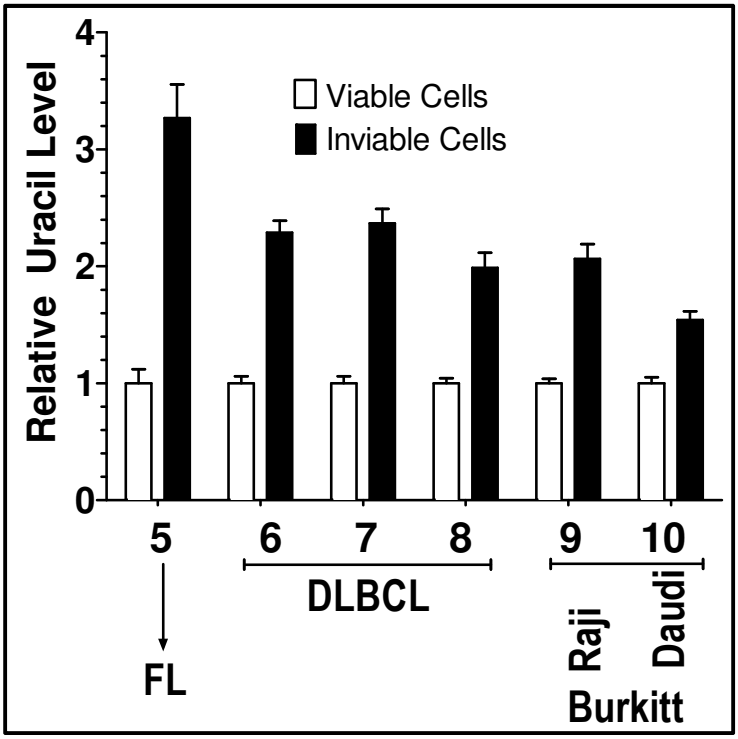

Figure 57. Uracils in Viable and Inviable Human B Cell Lymphomas. Relative uracil levels in inviable cells of human B cell lines with the uracil levels in viable cells set to 1.0 in each line tested.

3.3.5 DNA Damage due to Attempted Repair by UNG

UNG enzyme is one of the principal defense mechanisms mammalian cells have against accumulation of uracils in the genome, playing a major role in the excision of this base (see section 1.2.3b) (192). To determine whether accumulation of damage is effected specifically by the UNG glycosylase, I expressed the UNG enzyme inhibitor UGI in the same FL, DLBCL, and BL high AID-expressing cell lines. Although successful UGI expressing lines were created from these tumor lines, individual clones derived from each line showed considerable variability in growth. Some clones grew poorly or died off suggesting that inhibition of UNG2 enzyme in these cells beyond a certain level may be toxic. Regardless, I was able to establish cell lines with less than 50% uracil excision activity compared to corresponding cells transfected with the empty vector (EV) control from each of the five high AID-expressing lymphoma cell lines (Figure 58A, Top panel). Expression of UGI caused the genomic uracil levels to go up by ~30% (Figure 58A, Bottom panel). This confirms that UNG2 is one of the principal enzymes that helps eliminate the uracils created by AID in B cell lymphomas.

The removal of uracils by UNG enzyme results in abasic sites (see Figure 12). It reasons that if AP site accumulation is due primarily to the attempted repair of the AID-introduced uracils, then a reduction in this repair may reduce AP site accumulation as well. With inhibition of UNG in these UGI-expressing cells, a 2 fold to 4-fold reduction in the number of AP sites was observed when comparing the cell lines expressing empty vector to their UGI-expressing

counterparts (Figure 59). A similar reduction was observed in the γH2AX foci in the nuclei of those cells expressing UGI (Figure 60). This significant decrease in accumulated DSBs further implicates the attempted repair of uracils by UNG as one pathway that leads to genomic instability in these high- AID expressing cell lines.

Figure 58. Uracils in Human B Cell Lymphomas Expressing UGI

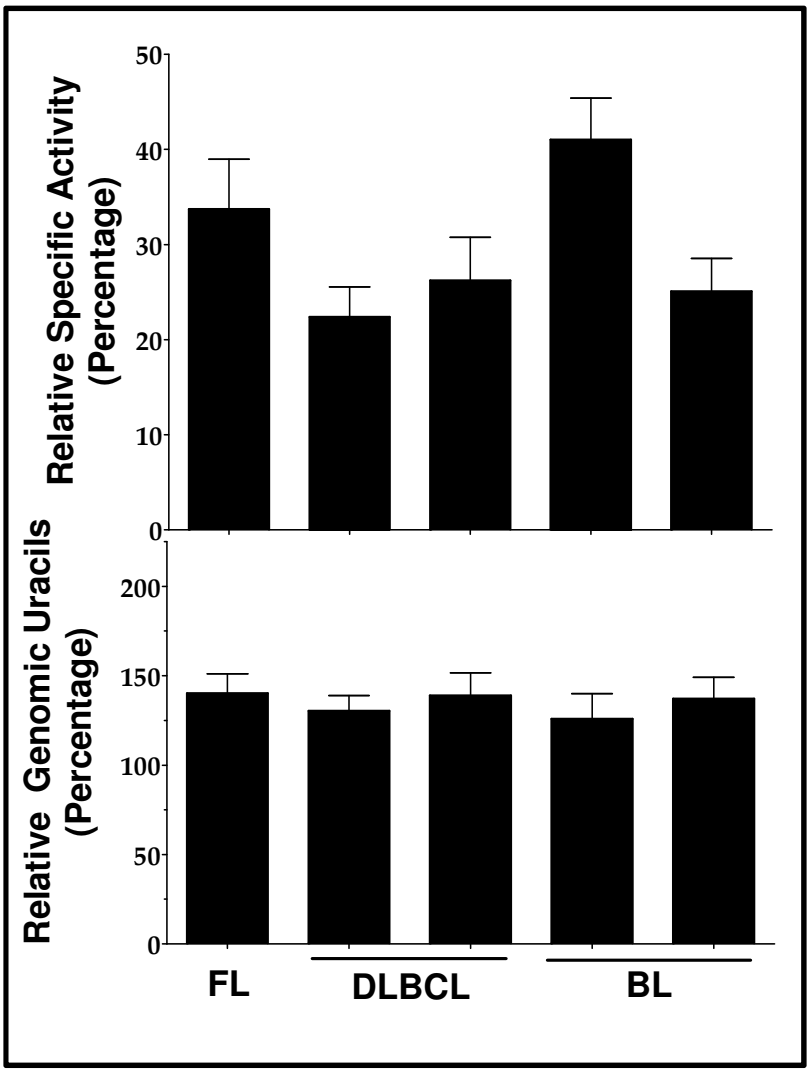

Figure 58. Uracils in Human B Cell Lymphomas Expressing UGI. *(Top panel)* Uracil Excision activity assays were performed on the UGI expressing cell lines and the corresponding cells lines expressing an empty vector (EV). The specific activity of uracil excision in the EV-expressing cells was set at 100*%.(Bottom*

panel) The genomic uracils were quantified in the UGI- expressing cell lines and the EV-expressing (set at 100%).

Figure 59. AP Sites in Human B Cell Lymphomas Expressing UGI

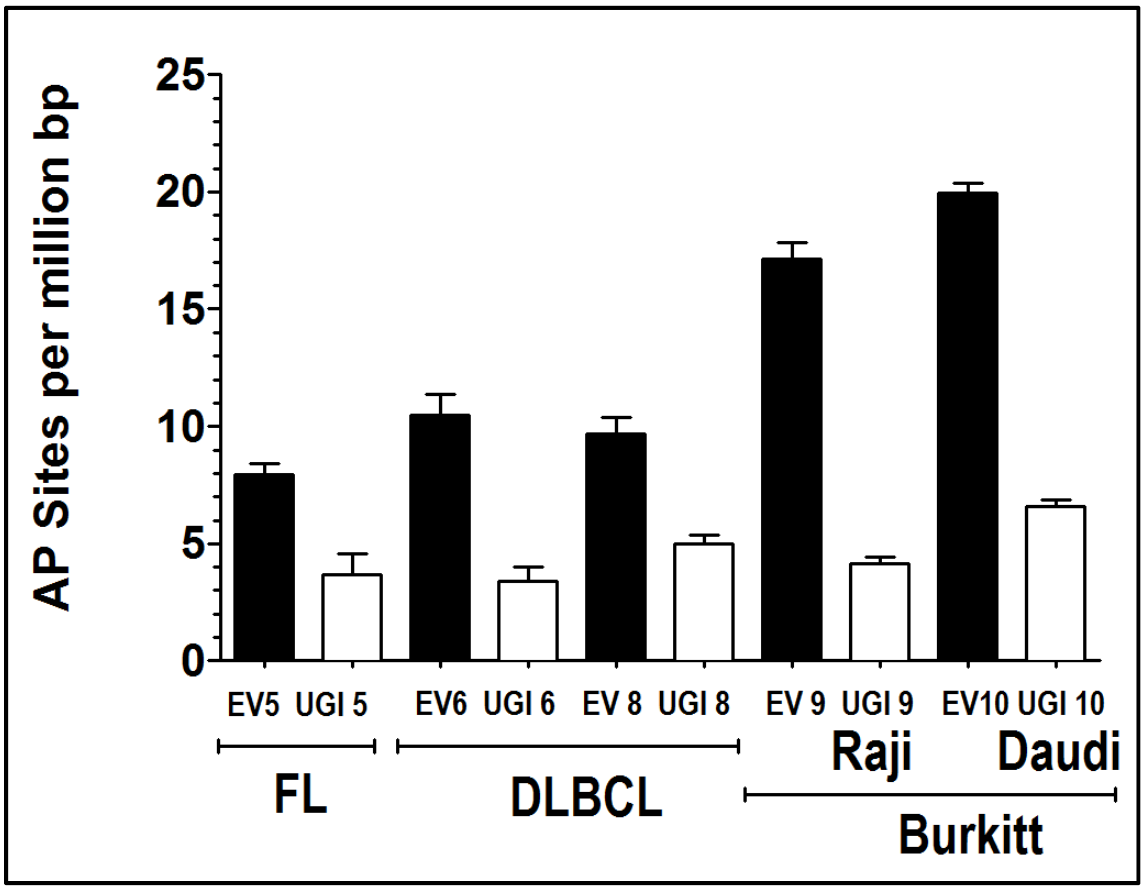

Figure 59. AP Sites in Human B Cell Lymphomas Expressing UGI. AP sites were quantified in the FL, DLBCL, and Burkitt lymphoma cell lines transfected with an empty vector (EV) or a plasmid expressing UGI (UGI).

Figure 60. DSBs in Human B Cell Lymphomas Expressing UGI

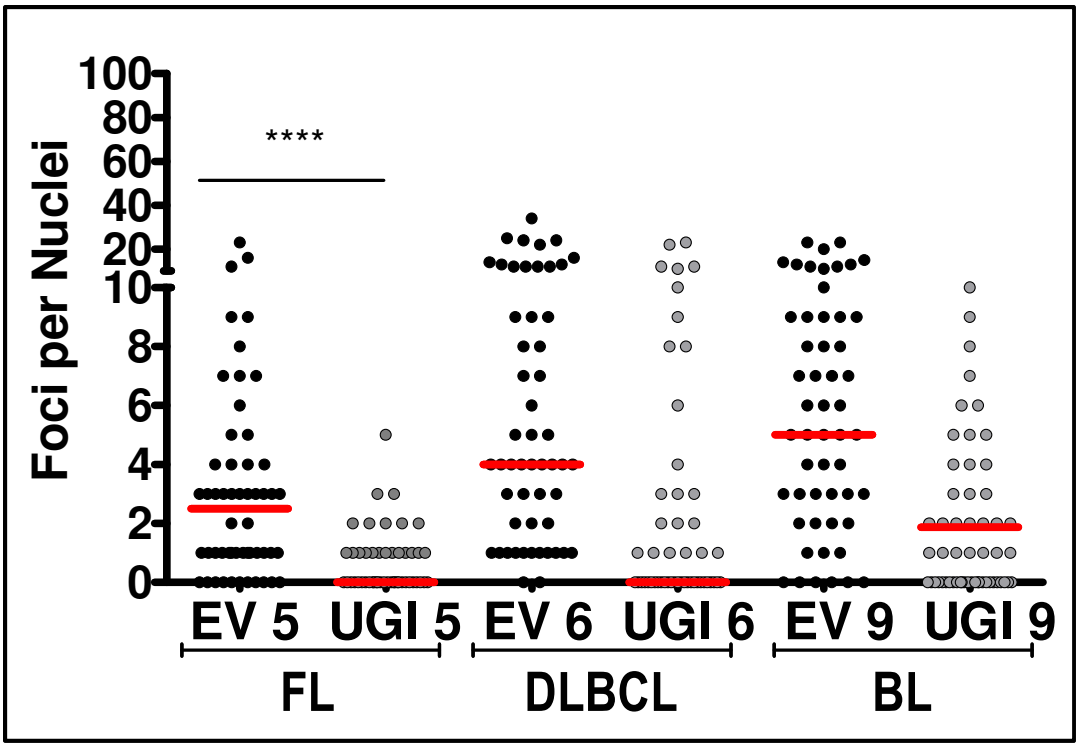

Figure 60. DSBs in Human B Cell Lymphomas Expressing UGI. γH2AX foci per nuclei were quantified with Image J for cell lines expressing UGI or the empty vector (EV) for ~50 cells per line. **** P<0.0001

Figure 61. Percent Viability Compromised further in Human B Cell Lymphomas Expressing UGI

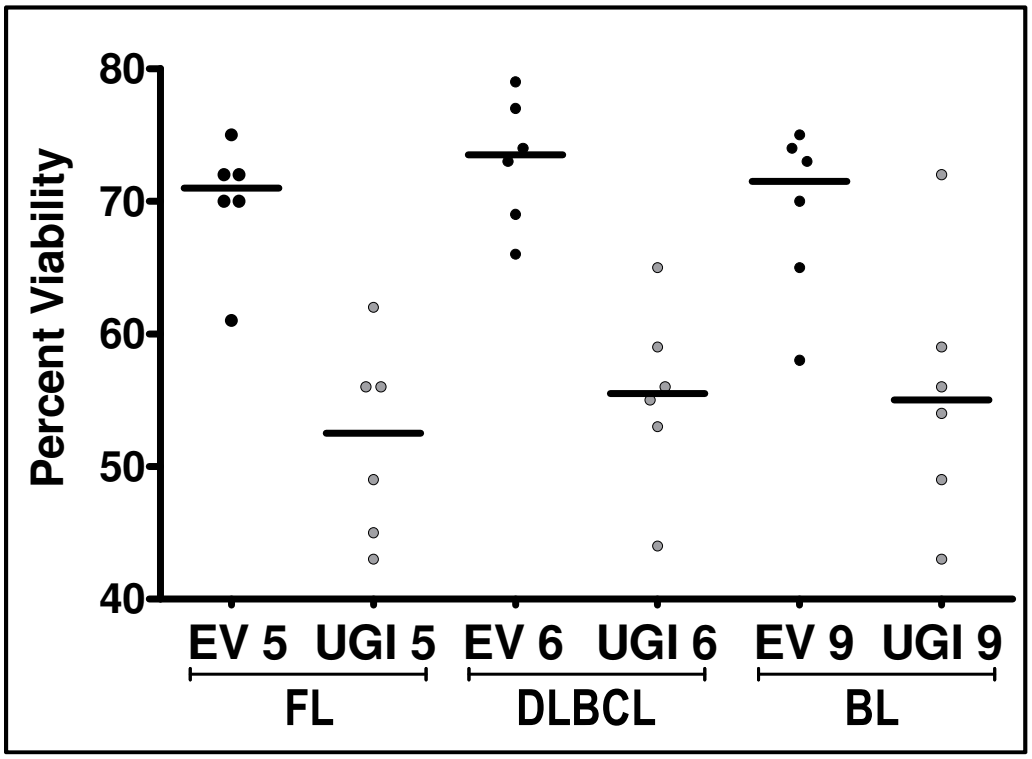

Figure 61. Percent Viability Compromised further in Human B Cell Lymphomas Expressing UGI. Percent viability of six independent cultures each was determined for FL, DLBCL, and BL cell lines expressing UGI or the empty vector (EV) using the Trypan Blue Exclusion assay. Lines indicate median percent viability.

Finally, with a reduction in uracils, AP sites, and DSBs, I expected the UGI-expressing cell lines would restore cell viability due to reduction in overall genomic instability. However, when the cells were subjected to the Trypan Blue exclusion assay, an even higher percent of cell death approaching nearly ~50% (Figure 61) in the UGI-expressing FL, DLBCL, and BL established lines was observed. This reduction in cell viability may be an indicator that the reduction of repair of the uracils by partially inhibiting UNG leads to toxic and overwhelming

mutations throughout the genome that may actually result in poor propagation of the cell lines.

CHAPTER 4: DISCUSSION

Figures and portions of the text in this chapter were reprinted or adapted in compliance with the *ASM Journals Statement of Authors' Rights* from:

Shalhout S, Haddad D, Sosin A, Holland TC, Al-Katib A, Martin A, Bhagwat AS. *Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development*. **Molecular and cellular biology**. 2014;34(21):4019-32. doi: 10.1128/MCB.00589-14.

4.1 Loss of Genomic Uracil Homeostasis in AID-expressing B Cell Lymphomas

AID's role in the adaptive immune system is an exceptional biological process since it is the only identified mechanism in which cells intentionally require damage to genomic DNA sequences to achieve benefit to the organism. Previous reports established the hypothesis that AID is a cytosine deaminase based on genetic studies as well as *in vitro* biochemistry. Cells are equipped with several repair pathways that excise uracils, but in the process, may lead to base substitutions and strand breaks. However, prior to this work, the specific relationship between AID- generated uracils and their subsequent excision by UNG or MMR had not been examined in normal B cell development or B cell lymphomagenesis. This research examined the interplay and resulted in novel findings and conclusions.

4.1.1 First Demonstration of Uracil Accumulation in Splenocyte Genomes

This study revealed that UNG-/- splenocytes acquire a substantial level of uracils upon AID activation through stimulation and that this rise in uracils was not seen in AID-/- splenocytes (Figure 26). Although an accumulation in genomic uracils was not entirely surprising since AID has been shown to be a DNA-cytosine deaminase *in vitro* (11-14), this study is the first to directly demonstrate AID-dependent uracil accumulation in the genomes of B lymphocytes undergoing activation. In addition, transient-transfection studies revealed that the catalytic activity of AID was required for its ability to increase genomic uracil levels (Figure 46). While another contributing factor to genomic uracil accumulation may be due to high transcription levels in stimulated B cells possibly leading to increased dUTP incorporation during replication, as was recently shown to be the case in yeast cells (193), the research described here clearly demonstrates that the deamination of cytosines in DNA by AID leads to the increase in genomic uracils of stimulated B lymphocytes (165).

The large increase in uracil accumulation suggests that the majority of the cytosines deaminated by AID are outside the Ig genes. In LPS treated UNG-/- splenocytes, the genomic uracil content increased from 11 to 60 fold (Figure 24A, 26, 25B and 34A). There was variation in the level of the genomic uracil increase seen in B cells from independent mice but an increase of at least 11 fold or higher was always detected and reproducible. The uracil quantification assay was always done by assaying DNA samples in triplicate, processing all the compared samples in an experiment in parallel, applying them to the same Ny+

membrane, and always alongside the uracil standard (Figure 23, 25A, and 33B). However, uracil content is reported here as fold changes and not uracils per million base pair because this assay uses a convenient synthesized duplex oligo as the uracil standard which does not represent true physiological DNA. In addition, previous studies using various uracil detection techniques have reported anywhere from 200 (194) to 15,000 (195) uracils in the unstimulated PBMC genome. Therefore, there may be around several thousand uracils per haploid genome in stimulated B cells deficient in UNG. Recently, only one uracil per kilobase was reported in the Ig genes (25) following B cell stimulation. Thus the majority of the uracils introduced by AID in stimulated UNG-/- B cell are in non-Ig genes.

4.1.2 Homeostasis in Uracil Creation and Excision Upon B Cell Stimulation

In *ex vivo* stimulated WT mouse B cells and human naïve mature B cells, uracils created by AID were efficiently eliminated with no detectable uracil increase resulting. Surprisingly, the level of genomic uracils in LPS activated mouse WT B cells was comparable to the level detected in AID-/- B cells (Figure 24B). I also confirmed that there was no detectable uracil increase in murine WT B cells upon IL-4 and LPS *ex vivo* stimulation which results in a different isotype switching than LPS alone (Figure 62). This lack in increased global genomic uracil accumulation in stimulated WT splenocytes suggests that UNG2 removes the uracils introduced by AID quickly and efficiently. Thus, a genomic uracil homeostasis exists in B lymphocytes undergoing the GC-reaction and antibody affinity maturation (Figure 63, upper panel).

4.1.3 B Cell Lymphomas Accumulate Uracil despite UNG2 Activity and Expression

The high uracil accumulation detected in murine and human B cell cancers overexpressing AID suggests a lack of efficiency in the ability of UNG2 to remove uracils introduced by AID. Initially, to explain the increase in uracil in lymphoma genomes, I simply reasoned that (1) either AID was expressed at much higher levels in lymphomas than activated WT B cells, or that (2) UNG2 was expressed at far lower levels in lymphomas. Experiments showed neither of these to be the case when studying the proteins at the expression or activity level. While an increase in the level of genomic uracils was only detected in stimulated primary B cells deficient in UNG, lymphomas accumulated high uracil levels but expressed UNG2 at normal or higher than normal levels and had robust detectable nuclear uracil excision activity. Specifically, nuclear excision activities in LPS-activated WT B cells and lymphoma cell lines overexpressing AID were similar (see Figure 29A, 36B, and 44B). The same was true for AID expression levels. In post-day 3 CIT stimulated CH12F3 B cells, the AID gene expression in CH12F3 cells and WT splenocytes were at comparable levels (see Figure 34B). Despite similar UNG2 and AID in WT and CH12F3 cells, only stimulated CH12F3 cells accumulated uracils in their genome. Strikingly, in Ramos 1 cells transfected to express AID at higher levels, the expression of UNG2 increased but the uracil levels increased as well (Figure 45A and 47A).

4.1.4 Why does uracil accumulate despite presence of functional UNG2?

Several possibilities may explain the observed disconnect between the uracil excision activity of UNG2 and the accumulated genomic uracils in lymphomas. UNG2 is primarily involved in removing uracils incorporated during replication by polymerase utilization of dUTP to initiate canonical BER (see Figure 10) (181, 196).Thus, the involvement of UNG2 in promoting mutations and strand breaks in B cells may require specialized recruitment to the genes targeted by AID. The factors involved in successful recruitment involve specific chromatin modifications, post translational protein modifications, and the help of molecular chaperones. It is known that when the Ig locus undergoes SHM and CSR, the local chromatin undergoes specific epigenetic modifications (197), and a recruitment protein may bring UNG2 to the modified chromatin. Alternatively, since AID acts on transcribed genes (198), UNG may be associated or recruited with the transcription elongation complex. For example, Rev1, a TLS polymerase, has recently been shown to play a non-catalytic role in recruiting UNG2 to S regions specifically being targeted by AID during CSR (89). Additionally, specific phosphorylation of UNG2 may be involved in promoting its pro-mutagenic role in adaptive immunity (199). The accumulation of uracils in stimulated CH12F3 cells suggests that although UNG2 is present and nuclear excision activities are robust, one of these recruiting or modifying factors may be deregulated in lymphomas.

Figure 62. *Ex Vivo* Stimulation of Murine B Cells with LPS and IL-4

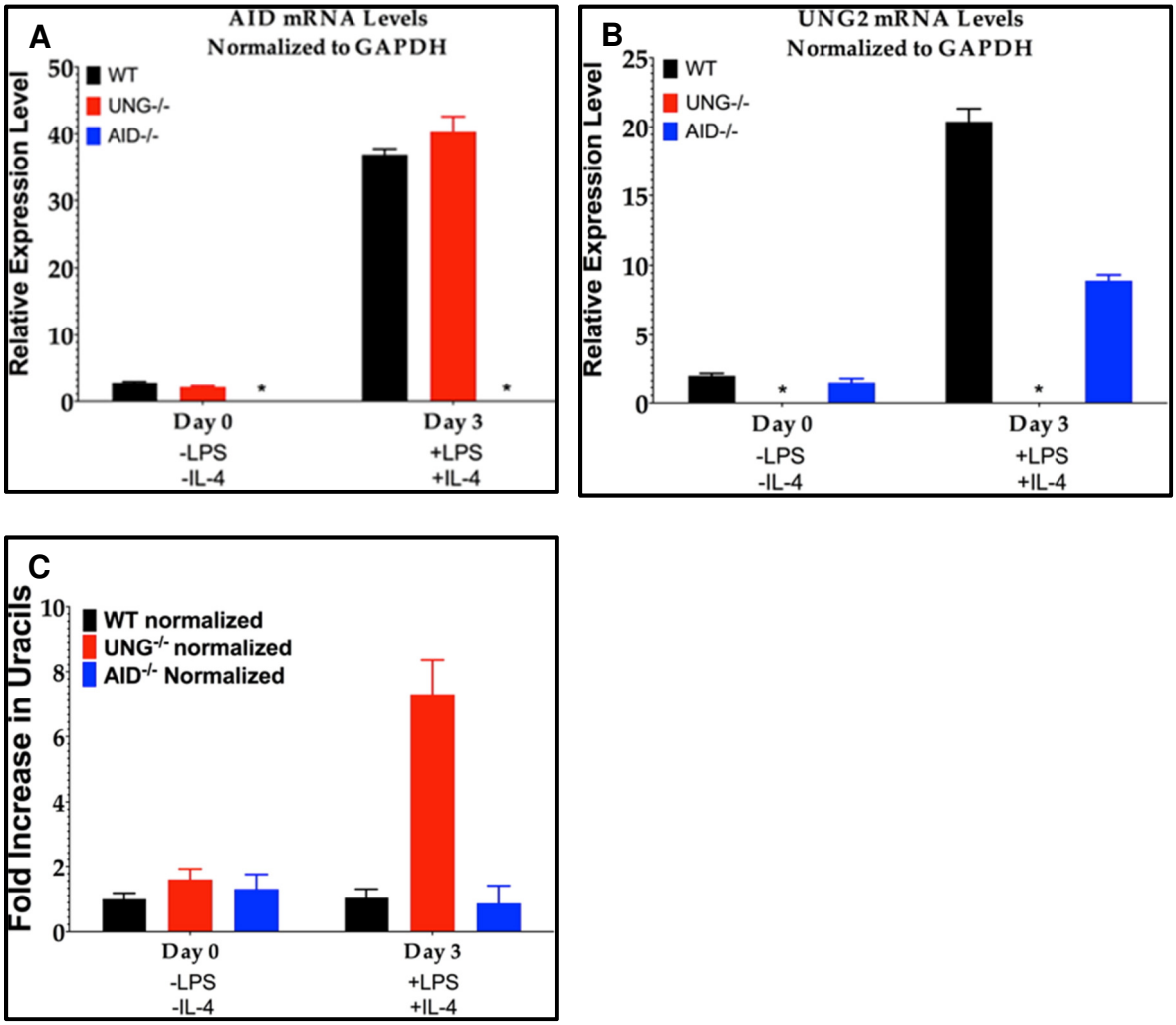

Figure 62. *Ex Vivo* Stimulation of Murine B Cells with LPS and IL-4. (A) AID expression following stimulation with LPS+IL-4. (B) UNG2 expression following stimulation with LPS+IL-4. (C) Genomic uracil levels following stimulation with LPS+IL-4 normalized to WT levels prior to stimulation. This figure was originally published in the supplementary material of (165). Reproduced from the American Society for Microbiology.

Figure 63. Homeostasis in Genomic Uracil Creation and Excision

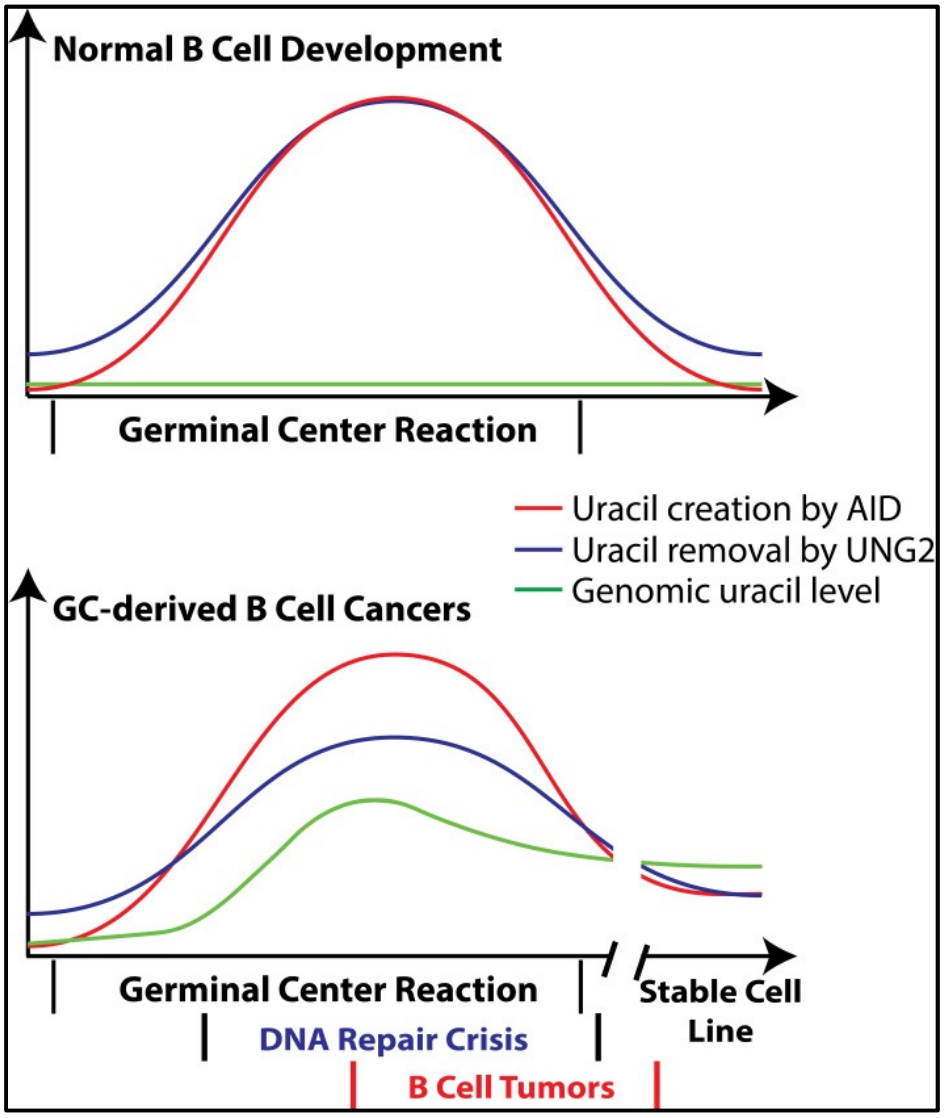

Figure 63. Homeostasis in Genomic Uracil Creation and Excision. (Top panel) Homeostasis during normal B cell development in germinal centers. Uracil creation by AID and uracil excision by UNG2 are kept in balance during normal B cell-development, resulting in low genomic uracil levels. (Bottom panel) The balance between uracil creation and removal is lost during the development of pre-cancer B cells in favor of uracil creation. This is a “DNA repair crisis” for the precancerous cells leading to accumulation of mutations and strand breaks in addition to uracils and resulting in genomic instability. This may lead to cellular transformation. When uracil removal equalizes with uracil creation, stable lymphoma or leukemia cell lines are established. These cell lines maintain the high uracil levels introduced during the repair crisis. This figure was originally published in the supplementary material of (165). Reproduced from the American Society for Microbiology.

4.1.5 High yet Steady Levels of Uracils in Lymphoma Cell Lines

As noted before, the uracil levels in the Raji cell line do not greatly fluctuate over time (see Figure 43) suggesting that the high AID-dependent uracil levels are stable. One may predict that if AID continues to introduce genomic uracils with the levels continually rising in lymphomas, cells may develop deleterious genomic instability due to the accumulation of mutations and strand breaks. Therefore, this study suggests cell lines may only be able to be immortalized from high-AID expressing B lymphomas, if a balance between creation of genomic uracils and their excision is reached. This results in a new steady-state stable uracil level in the cell line. However, this cell line steady-state uracil level is higher than that found in normal B cells and does not represent homeostasis in genomic uracils.

4.1.6 Possible DNA-repair Crisis disrupts Genomic Uracil Homeostasis

Increasing the expression of AID in B cell lymphomas or knocking it down results in increasing genomic uracils levels or decreasing them, respectively (Figure 45, 46, and 48). Drawing from these results, it is easy to envision some B cells lose the balance between uracil introduction and removal during the GC-reaction possibly due to impaired targeting of UNG2 to uracils introduced by AID. This may result in an impaired ability to efficiently remove uracils created by AID in a timely manner resulting in a “DNA repair crisis” (Figure 63, lower panel). If this crisis persists it may result in high genomic instability, acquisition of mutations, and possible apoptosis if the cell cannot repair the damage effectively. However, during cell line establishment the uracil excision capabilities eventually

catch up with introduced uracils resulting in the establishment of a stable cell line. This results in a new steady-state condition where the genomic uracil levels are much higher than the level in unstimulated WT B cells (Figure 63, lower panel). In contrast, patients with B cell lymphomas/leukemias may still exhibit instability from this DNA repair crisis with no stability acquired in AID, UNG2, and genomic uracil levels. These levels may change overtime and under treatment and may differ among patients resulting in variation from cancer to cancer (Figure 49A and B).

An imbalance between AID creating uracils and UNG2 removing uracils in favor of AID would result in an increase in the frequency of base substitutions and mutations. In UNG-/- cells, the *lacI* and HPRT genes were shown to acquire mutations at 1.5 and 5-fold higher the rates found in normal cells respectively (196, 200). With the increased expression of AID upon B cell activation, any deficiency in UNG and thus uracil excision may increase mutations. If these mutations are in proto-oncogenes or tumor suppressors, lymphomagenesis may ensue. In fact, UNG-/- mice have been shown to develop B cell lymphomas 20-fold higher than WT mice and have a shorter life expectancy (201). In addition, human follicular lymphomas have recently been shown to continually acquire AID dependent mutations in the variable region as well as non-Ig genes during the evolution of the patient's cancer (128). Other members of the APOBEC family of proteins have also been implicated in promoting cancers and mutations. Sequencing of several tumor genomes has revealed these cancers accumulate mutations with an AID/APOBEC 'signature.' These include cancers originating in

the cervix, breast, and bladder, as well as of B cell origin (161, 162, 202). Furthermore, many breast cancer patient tumor samples and cell lines have been shown to overexpress APOBEC3B with at least two cell lines recently found to have elevated uracil levels (203). If the development of non-GC derived lymphomas overexpressing other APOBECs also involves the type of DNA repair crisis depicted in Figure 63, then an understanding of how the activation of normal B cells and their maturation may result in an imbalance between AID generated uracils and uracil excision by UNG2 may be critical to understanding how these cancers develop as well.

4.2 AID-dependent Accumulation of Specific DNA Damage in B cell Lymphomas

4.2.1 DNA damage caused by AID and UNG

This work has demonstrated that high-AID expressing cell lines accumulate large amounts of genomic uracil introduced by AID (165). Furthermore, the attempted repair of these uracils by UNG2 leads to a consistent pattern of AID expressing lymphomas accumulating high levels of DNA damage. The specific types of DNA lesions, including AP sites and strand breaks, are predicted to arise at each subsequent step of removal or attempted repair of uracils (see model presented in Figure 64). In fact, a reduction in UNG repair is met with a decrease in the downstream DNA lesions (AP sites and DSBs) in these high AID-expressing lymphomas. Another finding of this study is that inviable cells accumulate even more of the initial uracil lesion (Figure 57). This implicates that dead/dying cells may have overwhelming levels of uracils, possibly mutating genes essential to cell survival. These uracils continue to

accumulate and escape repair and may destine the cell to apoptose. Further studies are required to determine if the high uracil load signals programmed cell death.

Figure 64. Model of DNA Damage Accumulation in High-AID Expressing Lymphomas

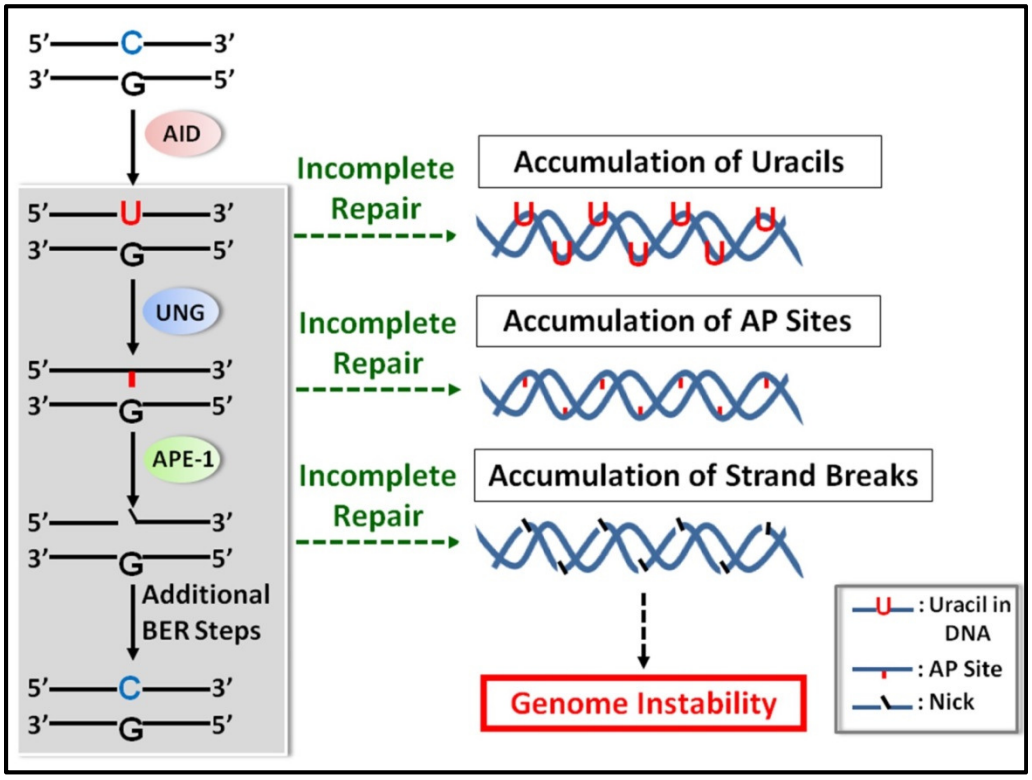

Figure 64. Model of DNA Damage Accumulation in High-AID Expressing Lymphomas. AID introduces the initial uracil lesion. Uracils that escape repair or removal by such repair proteins as UNG2 accumulate throughout the genome. However, some uracils are recognized and excised by UNG2 producing higher than normal levels of AP sites in the DNA. The nicking of these sites by APE can lead to SSBs and eventually DSBs resulting in genomic instability and poor cell viability. Thus, this damage is dependent on AID and UNG.

4.2.2 Mechanism of Accumulated DNA damage in Lymphomas

The mammalian system is equipped with several repair mechanisms and enzymes to address uracils introduced in DNA. Four different glycosylases, UNG, TDG, MBD4, and SMUG1 (see TABLE 1) can excise uracils in DNA during base excision repair resulting in abasic sites. These AP sites can then be nicked by AP endonucleases (APE1/2) leading to single strand breaks and eventually double strand breaks. The increase in the number of AP sites observed in the genomes of these high-AID expressing cell lines may result from incomplete repair following excision by uracil glycosylases (see Figure 64). The accumulation of single- and double strand breaks may arise following incomplete repair of the nicking of these AP sites by APE. Furthermore, it is more likely that these DNA lesions are distributed at a large number of sites in the genome and not restricted to the Ig region, the major target of AID. When UNG is partially inhibited by expressing UGI, these cells accumulate higher uracil loads due to decreased repair abilities. However, less AP sites and DSBs are detected which supports the notion that these lesions arise in part, due to attempted removal by UNG. Cell viability, however, suffers with the reduction in repair in these high-AID expressing lymphomas. This may be caused by an increase in the mutational load of essential genes in the genomes due to escaped uracil repair that leads to interruptions in normal cell functions.

4.2.3 DNA Damage may promote Lymphomagenesis

It is well established that the translocation of c-myc, bcl-2, bcl-6 and other genes to the Ig gene loci are a hallmark of B cell malignancies (204). Additionally, in studies of B lymphocytes where translocations were induced by creating a designed break in the IgH or c-myc genes, the other break was created by AID and was found to lie at a large number of sites spread throughout the genome (205, 206). As AID is required for the translocation of the c-myc gene to the IgH locus (107), it is generally assumed that AID is involved in converting cytosines to uracils at these genes and that the newly introduced uracils are somehow converted to DSBs that are the cause of these translocations. It is attractive to suggest that the damage identified in this study is the source of DSBs that result in genome-wide translocations when AID is overexpressed in B lymphocytes (205, 206) including those that aid oncogenic transformation and lead to mutations in non-Ig genes (178). Like the accumulated uracils, future studies that determine the sequence of DNA fragments containing abasic sites and strand breaks found in these cancer genomes will clarify whether they are the direct cause of translocated genes found in these cells.

4.2.4 Future Directions

These findings point to new avenues for treating GC-derived lymphomas as these cells contain a large number of DNA modifications that are largely absent in normal cells. For example, once the mechanism underlying the observation that cells with higher uracils, AP sites, and strand breaks are prone to apoptose is better understood, it may be possible to exploit this phenomenon to selectively kill tumor cells highly expressing AID. For example, the accumulated types of DNA damage described here may be exacerbated by treatment with chemotherapeutic agents that lead to similar lesions. This may overwhelm the cells and result in cell arrest or death quicker or with lower doses of the chemotherapy since endogenous DNA damage persists already. In addition, the fact that B cell lymphomas accumulate uracils and therefore require functional BER for their removal and repair, it may make BER a target for chemotherapy. For example, methoxyamine, which is utilized in the uracil quantification assay to block endogenous abasic sites (see Figure 21), is also currently studied as part of a combination anticancer chemotherapy regimen when cancers are also treated with an alkylating agent such as Temozolomide (207). The alkylating agent promotes damage to the DNA bases that are then able to be excised by a DNA glycosylase resulting in AP sites. This allows Mx to react to these AP sites and therefore inhibits repair, leading to cell death. Several NCI-approved clinical trials using Mx in combination therapies are currently under investigation (NCT00892385, NCT01658319 and NCT00692159, clinicaltrials.gov). Since B cell lymphomas already have endogenous damage in

the form of high levels of uracils and AP sites this strategy may be worth investigating specifically in high-AID expressing B cell lymphomas as well as other high-APOBEC expressing cancers. In addition, other drugs that bind to AP sites and inhibit repair should also be considered.

The high accumulation of uracils in B cell cancers may also be used as a diagnostic marker for GC-derived lymphomas in addition to routine practices such as determining the cytogenetics and chromosomal abnormalities in the cancer. Furthermore, the accumulation of uracils and mutations should be explored as the possible cause for the common drug resistance observed in the treatment of some high-AID expressing cancers (208).

Future experiments need to be carried out to determine the molecular cause for the imbalance seen in uracil accumulation and excision in lymphomas. For example, UNG2 is present and active in these cells at normal or above noraml levels. Thus, it is possible that deregulation of recruitment of UNG2 to the AID-generated uracils is the culprit.

REFERENCES

1.Nakamura M, Kondo S, Sugai M, Nazarea M, Imamura S, Honjo T. High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. International immunology. 1996;8(2):193-201. PubMed PMID: 8671604.

2.Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, Honjo T. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. The Journal of biological chemistry. 1999;274(26):18470-6. PubMed PMID: 10373455.

3.Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell. 2000;102(5):553-63. PubMed PMID: 11007474.

4.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 AG, Brousse N, Muramatsu M, Notarangelo LD, Kinoshita K, Honjo T, Fischer A, Durandy A. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell. 2000;102(5):565-75. PubMed PMID: 11007475.

5.Chan L, Chang BH, Nakamuta M, Li WH, Smith LC. Apobec-1 and apolipoprotein B mRNA editing. Biochimica et biophysica acta. 1997;1345(1):11-26. PubMed PMID: 9084497.

6.Martin A, Bardwell PD, Woo CJ, Fan M, Shulman MJ, Scharff MD. Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. Nature. 2002;415(6873):802-6. doi: 10.1038/nature714. PubMed PMID: 11823785.

7.Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K, Honjo T. Constitutive expression of AID leads to tumorigenesis. The Journal of experimental medicine. 2003;197(9):1173-81. doi: 10.1084/jem.20030275. PubMed PMID: 12732658; PMCID: 2193972.

8.Yoshikawa K, Okazaki IM, Eto T, Kinoshita K, Muramatsu M, Nagaoka H, Honjo T. AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. Science. 2002;296(5575):2033-6. doi: 10.1126/science.1071556. PubMed PMID: 12065838.

9.Mayorov VI, Rogozin IB, Adkison LR, Frahm C, Kunkel TA, Pavlov YI. Expression of human AID in yeast induces mutations in context similar to the context of somatic hypermutation at G-C pairs in immunoglobulin genes. BMC immunology. 2005;6:10. doi: 10.1186/1471-2172-6-10. PubMed PMID: 15949042; PMCID: 1180437.

10.Petersen-Mahrt SK, Harris RS, Neuberger MS. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. Nature. 2002;418(6893):99-103. doi: 10.1038/nature00862. PubMed PMID: 12097915.

11.Bransteitter R, Pham P, Scharff MD, Goodman MF. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. Proceedings of the National Academy of Sciences

of the United States of America. 2003;100(7):4102-7. doi: 10.1073/pnas.0730835100. PubMed PMID: 12651944; PMCID: 153055.

12.Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. Nature. 2003;422(6933):726-30. doi: 10.1038/nature01574. PubMed PMID: 12692563.

13.Dickerson SK, Market E, Besmer E, Papavasiliou FN. AID mediates hypermutation by deaminating single stranded DNA. The Journal of experimental medicine. 2003;197(10):1291-6. doi: 10.1084/jem.20030481. PubMed PMID: 12756266; PMCID: 2193777.

14.Sohail A, Klapacz J, Samaranayake M, Ullah A, Bhagwat AS. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. Nucleic acids research. 2003;31(12):2990-4. PubMed PMID: 12799424; PMCID: 162340.

15.Pham P, Bransteitter R, Petruska J, Goodman MF. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. Nature. 2003;424(6944):103-7. doi: 10.1038/nature01760. PubMed PMID: 12819663.

16.Samaranayake M, Bujnicki JM, Carpenter M, Bhagwat AS. Evaluation of molecular models for the affinity maturation of antibodies: roles of cytosine deamination by AID and DNA repair. Chemical reviews. 2006;106(2):700-19. doi: 10.1021/cr040496t. PubMed PMID: 16464021.

17.Neuberger MS, Harris RS, Di Noia J, Petersen-Mahrt SK. Immunity through DNA deamination. Trends in biochemical sciences. 2003;28(6):305-12. doi: 10.1016/S0968-0004(03)00111-7. PubMed PMID: 12826402.

18.Ta VT, Nagaoka H, Catalan N, Durandy A, Fischer A, Imai K, Nonoyama S, Tashiro J, Ikegawa M, Ito S, Kinoshita K, Muramatsu M, Honjo T. AID mutant analyses indicate requirement for class-switch-specific cofactors. Nature immunology. 2003;4(9):843-8. doi: 10.1038/ni964. PubMed PMID: 12910268.

19.Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM, Nussenzweig MC. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. Molecular cell. 2003;12(2):501-8. PubMed PMID: 14536088.

20.Shinkura R, Ito S, Begum NA, Nagaoka H, Muramatsu M, Kinoshita K, Sakakibara Y, Hijikata H, Honjo T. Separate domains of AID are required for somatic hypermutation and class-switch recombination. Nature immunology. 2004;5(7):707-12. doi: 10.1038/ni1086. PubMed PMID: 15195091.

21.Patenaude AM, Di Noia JM. The mechanisms regulating the subcellular localization of AID. Nucleus. 2010;1(4):325-31. doi: 10.4161/nucl.1.4.12107. PubMed PMID: 21327080; PMCID: 3027040.

22.Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M, Honjo T. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. Proceedings of the National Academy of Sciences of the United States of America.

2004;101(7):1975-80. doi: 10.1073/pnas.0307335101. PubMed PMID: 14769937; PMCID: 357037.

23.Kudo N, Matsumori N, Taoka H, Fujiwara D, Schreiner EP, Wolff B, Yoshida M, Horinouchi S. Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(16):9112-7. PubMed PMID: 10430904; PMCID: 17741.

24.McBride KM, Barreto V, Ramiro AR, Stavropoulos P, Nussenzweig MC. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. The Journal of experimental medicine. 2004;199(9):1235-44. doi: 10.1084/jem.20040373. PubMed PMID: 15117971; PMCID: 2211910.

25.Patenaude AM, Orthwein A, Hu Y, Campo VA, Kavli B, Buschiazzo A, Di Noia JM. Active nuclear import and cytoplasmic retention of activation-induced deaminase. Nature structural & molecular biology. 2009;16(5):517-27. doi: 10.1038/nsmb.1598. PubMed PMID: 19412186.

26.Betts L, Xiang S, Short SA, Wolfenden R, Carter CW, Jr. Cytidine deaminase. The 2.3 A crystal structure of an enzyme: transition-state analog complex. Journal of molecular biology. 1994;235(2):635-56. doi: 10.1006/jmbi.1994.1018. PubMed PMID: 8289286.

27.Janeway CA Jr TP, Walport M. et al. Principles of innate and adaptive immunity. Immunobiology: The Immune System in Health and Disease

5th ed. New York: Garland Science; 2001.

28.Janeway CPTMWMS. Immunobiology. Fifth Edition ed. New York and London: Garland Science; 2001.

29.Gould HJ, Sutton BJ. IgE in allergy and asthma today. Nature reviews Immunology. 2008;8(3):205-17. doi: 10.1038/nri2273. PubMed PMID: 18301424.

30.Cerutti A, Chen K, Chorny A. Immunoglobulin responses at the mucosal interface. Annual review of immunology. 2011;29:273-93. doi: 10.1146/annurev-immunol-031210-101317. PubMed PMID: 21219173; PMCID: 3064559.

31.Gellert M. Molecular analysis of V(D)J recombination. Annual review of genetics. 1992;26:425-46. doi: 10.1146/annurev.ge.26.120192.002233. PubMed PMID: 1482120.

32.Jacob J, Kelsoe G, Rajewsky K, Weiss U. Intraclonal generation of antibody mutants in germinal centres. Nature. 1991;354(6352):389-92. doi: 10.1038/354389a0. PubMed PMID: 1956400.

33.Pinaud E, Khamlichi AA, Le Morvan C, Drouet M, Nalesso V, Le Bert M, Cogne M. Localization of the 3' IgH locus elements that effect long-distance regulation of class switch recombination. Immunity. 2001;15(2):187-99. PubMed PMID: 11520455.

34.Rouaud P, Vincent-Fabert C, Saintamand A, Fiancette R, Marquet M, Robert I, Reina-San-Martin B, Pinaud E, Cogne M, Denizot Y. The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. The Journal of experimental medicine. 2013;210(8):1501-7. doi: 10.1084/jem.20130072. PubMed PMID: 23825188; PMCID: 3727322.

35.Durandy A. Activation-induced cytidine deaminase: a dual role in class-switch recombination and somatic hypermutation. European journal of immunology. 2003;33(8):2069-73. doi: 10.1002/eji.200324133. PubMed PMID: 12884279.

36.Maclennan ICMH, D. L. In: Honjo T, Alt, F. W., Neuberger, M. S., editor. Molecular Biology of B Cells. London: Elsevier Academic Press; 2004.

37.Rajewsky K. Clonal selection and learning in the antibody system. Nature. 1996;381(6585):751-8. doi: 10.1038/381751a0. PubMed PMID: 8657279.

38.Kinoshita K, Honjo T. Linking class-switch recombination with somatic hypermutation. Nature reviews Molecular cell biology. 2001;2(7):493-503. doi: 10.1038/35080033. PubMed PMID: 11433363.

39.Papavasiliou FN, Schatz DG. Somatic hypermutation of immunoglobulin genes: merging mechanisms for genetic diversity. Cell. 2002;109 Suppl:S35-44. PubMed PMID: 11983151.

40.Harris RS, Sale JE, Petersen-Mahrt SK, Neuberger MS. AID is essential for immunoglobulin V gene conversion in a cultured B cell line. Current biology : CB. 2002;12(5):435-8. PubMed PMID: 11882297.

41.Arakawa H, Hauschild J, Buerstedde JM. Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. Science. 2002;295(5558):1301-6. doi: 10.1126/science.1067308. PubMed PMID: 11847344.

42.Reynaud CA, Anquez V, Grimal H, Weill JC. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. Cell. 1987;48(3):379-88. PubMed PMID: 3100050.

43.Diaz M, Flajnik MF. Evolution of somatic hypermutation and gene conversion in adaptive immunity. Immunological reviews. 1998;162:13-24. PubMed PMID: 9602348.

44.Rada C, Milstein C. The intrinsic hypermutability of antibody heavy and light chain genes decays exponentially. The EMBO journal. 2001;20(16):4570-6. doi: 10.1093/emboj/20.16.4570. PubMed PMID: 11500383; PMCID: 125579.

45.Longerich S, Tanaka A, Bozek G, Nicolae D, Storb U. The very 5' end and the constant region of Ig genes are spared from somatic mutation because AID does not access these regions. The Journal of experimental medicine. 2005;202(10):1443-54. doi: 10.1084/jem.20051604. PubMed PMID: 16301749; PMCID: 2212980.

46.Lebecque SG, Gearhart PJ. Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter, and 3' boundary is approximately 1 kb from V(D)J gene. The Journal of experimental medicine. 1990;172(6):1717-27. PubMed PMID: 2258702; PMCID: 2188766.

47.Golding GB, Gearhart PJ, Glickman BW. Patterns of somatic mutations in immunoglobulin variable genes. Genetics. 1987;115(1):169-76. PubMed PMID: 3557109; PMCID: 1203053.

48.Gearhart PJ, Sen R. Regulating antibody diversity: taming a mutagen. Molecular cell. 2008;31(5):615-6. doi: 10.1016/j.molcel.2008.08.016. PubMed PMID: 18775319.

49.Rada C, Yelamos J, Dean W, Milstein C. The 5' hypermutation boundary of kappa chains is independent of local and neighbouring sequences and related to

the distance from the initiation of transcription. European journal of immunology. 1997;27(12):3115-20. doi: 10.1002/eji.1830271206. PubMed PMID: 9464795.

50.Both GW, Taylor L, Pollard JW, Steele EJ. Distribution of mutations around rearranged heavy-chain antibody variable-region genes. Molecular and cellular biology. 1990;10(10):5187-96. PubMed PMID: 2118991; PMCID: 361197.

51.Rogozin IB, Diaz M. Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. Journal of immunology. 2004;172(6):3382-4. PubMed PMID: 15004135.

52.Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. Nature immunology. 2003;4(5):452-6. doi: 10.1038/ni920. PubMed PMID: 12692548.

53.Fukita Y, Jacobs H, Rajewsky K. Somatic hypermutation in the heavy chain locus correlates with transcription. Immunity. 1998;9(1):105-14. PubMed PMID: 9697840.

54.Milstein C, Neuberger MS, Staden R. Both DNA strands of antibody genes are hypermutation targets. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(15):8791-4. PubMed PMID: 9671757; PMCID: 21155.

55.Sono M, Wataya Y, Hayatsu H. Role of bisulfite in the deamination and the hydrogen isotope exchange of cytidylic acid. Journal of the American Chemical Society. 1973;95(14):4745-9. PubMed PMID: 4730665.

56.Dedon PC, Tannenbaum SR. Reactive nitrogen species in the chemical biology of inflammation. Archives of biochemistry and biophysics. 2004;423(1):12-22. PubMed PMID: 14989259.

57.Krokan HE, Drablos F, Slupphaug G. Uracil in DNA--occurrence, consequences and repair. Oncogene. 2002;21(58):8935-48. doi: 10.1038/sj.onc.1205996. PubMed PMID: 12483510.

58.Krokan HE, Nilsen H, Skorpen F, Otterlei M, Slupphaug G. Base excision repair of DNA in mammalian cells. FEBS letters. 2000;476(1-2):73-7. PubMed PMID: 10878254.

59.Nilsen H, Otterlei M, Haug T, Solum K, Nagelhus TA, Skorpen F, Krokan HE. Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene. Nucleic acids research. 1997;25(4):750-5. PubMed PMID: 9016624; PMCID: 146498.

60.Sousa MM, Krokan HE, Slupphaug G. DNA-uracil and human pathology. Molecular aspects of medicine. 2007;28(3-4):276-306. doi: 10.1016/j.mam.2007.04.006. PubMed PMID: 17590428.

61.Krokan HE, Standal R, Slupphaug G. DNA glycosylases in the base excision repair of DNA. The Biochemical journal. 1997;325 (Pt 1):1-16. PubMed PMID: 9224623; PMCID: 1218522.

62.Boorstein RJ, Cummings A, Jr., Marenstein DR, Chan MK, Ma Y, Neubert TA, Brown SM, Teebor GW. Definitive identification of mammalian 5-hydroxymethyluracil DNA N-glycosylase activity as SMUG1. The Journal of biological chemistry. 2001;276(45):41991-7. doi: 10.1074/jbc.M106953200. PubMed PMID: 11526119.

63.Masaoka A, Matsubara M, Hasegawa R, Tanaka T, Kurisu S, Terato H, Ohyama Y, Karino N, Matsuda A, Ide H. Mammalian 5-formyluracil-DNA glycosylase. 2. Role of SMUG1 uracil-DNA glycosylase in repair of 5-formyluracil and other oxidized and deaminated base lesions. Biochemistry. 2003;42(17):5003-12. doi: 10.1021/bi0273213. PubMed PMID: 12718543.

64.Waters TR, Gallinari P, Jiricny J, Swann PF. Human thymine DNA glycosylase binds to apurinic sites in DNA but is displaced by human apurinic endonuclease 1. The Journal of biological chemistry. 1999;274(1):67-74. PubMed PMID: 9867812.

65.Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR, Bird A. Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. Science. 2002;297(5580):403-5. doi: 10.1126/science.1073354. PubMed PMID: 12130785.

66.Dogliotti E, Fortini P, Pascucci B, Parlanti E. The mechanism of switching among multiple BER pathways. Progress in nucleic acid research and molecular biology. 2001;68:3-27. PubMed PMID: 11554307.

67.Etzioni A, Ochs HD. The hyper IgM syndrome--an evolving story. Pediatric research. 2004;56(4):519-25. doi: 10.1203/01.PDR.0000139318.65842.4A. PubMed PMID: 15319456.

68.Imai K, Slupphaug G, Lee WI, Revy P, Nonoyama S, Catalan N, Yel L, Forveille M, Kavli B, Krokan HE, Ochs HD, Fischer A, Durandy A. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. Nature immunology. 2003;4(10):1023-8. doi: 10.1038/ni974. PubMed PMID: 12958596.

69.Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. Current biology : CB. 2002;12(20):1748-55. PubMed PMID: 12401169.

70.Storb U, Stavnezer J. Immunoglobulin genes: generating diversity with AID and UNG. Current biology : CB. 2002;12(21):R725-7. PubMed PMID: 12419200.

71.Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. Annual review of biochemistry. 2007;76:1-22. doi: 10.1146/annurev.biochem.76.061705.090740. PubMed PMID: 17328676.

72.Muramatsu M, Nagaoka H, Shinkura R, Begum NA, Honjo T. Discovery of activation-induced cytidine deaminase, the engraver of antibody memory. Advances in immunology. 2007;94:1-36. doi: 10.1016/S0065-2776(06)94001-2. PubMed PMID: 17560270.

73.Jiricny J. Replication errors: cha(lle)nging the genome. The EMBO journal. 1998;17(22):6427-36. doi: 10.1093/emboj/17.22.6427. PubMed PMID: 9822589; PMCID: 1170991.

74.Modrich P. Mechanisms in eukaryotic mismatch repair. The Journal of biological chemistry. 2006;281(41):30305-9. doi: 10.1074/jbc.R600022200. PubMed PMID: 16905530; PMCID: 2234602.

75.Rada C, Ehrenstein MR, Neuberger MS, Milstein C. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. Immunity. 1998;9(1):135-41. PubMed PMID: 9697843.

76.Wiesendanger M, Kneitz B, Edelmann W, Scharff MD. Somatic hypermutation in MutS homologue (MSH)3-, MSH6-, and MSH3/MSH6-deficient mice reveals a role for the MSH2-MSH6 heterodimer in modulating the base substitution pattern. The Journal of experimental medicine. 2000;191(3):579-84. PubMed PMID: 10662804; PMCID: 2195810.

77.Bardwell PD, Woo CJ, Wei K, Li Z, Martin A, Sack SZ, Parris T, Edelmann W, Scharff MD. Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. Nature immunology. 2004;5(2):224-9. doi: 10.1038/ni1031. PubMed PMID: 14716311.

78.Stavnezer J, Linehan EK, Thompson MR, Habboub G, Ucher AJ, Kadungure T, Tsuchimoto D, Nakabeppu Y, Schrader CE. Differential expression of APE1 and APE2 in germinal centers promotes error-prone repair and A:T mutations during somatic hypermutation. Proceedings of the National Academy of Sciences

of the United States of America. 2014;111(25):9217-22. doi: 10.1073/pnas.1405590111. PubMed PMID: 24927551; PMCID: 4078814.

79.Sabouri Z, Okazaki IM, Shinkura R, Begum N, Nagaoka H, Tsuchimoto D, Nakabeppu Y, Honjo T. Apex2 is required for efficient somatic hypermutation but not for class switch recombination of immunoglobulin genes. International immunology. 2009;21(8):947-55. doi: 10.1093/intimm/dxp061. PubMed PMID: 19556307.

80.Rada C, Di Noia JM, Neuberger MS. Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. Molecular cell. 2004;16(2):163-71. doi: 10.1016/j.molcel.2004.10.011. PubMed PMID: 15494304.

81.Shen HM, Tanaka A, Bozek G, Nicolae D, Storb U. Somatic hypermutation and class switch recombination in Msh6(-/-)Ung(-/-) double-knockout mice. Journal of immunology. 2006;177(8):5386-92. PubMed PMID: 17015724.

82.Chaudhuri J, Basu U, Zarrin A, Yan C, Franco S, Perlot T, Vuong B, Wang J, Phan RT, Datta A, Manis J, Alt FW. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. Advances in immunology. 2007;94:157-214. doi: 10.1016/S0065-2776(06)94006-1. PubMed PMID: 17560275.

83.Stavnezer J. Molecular processes that regulate class switching. Current topics in microbiology and immunology. 2000;245(2):127-68. PubMed PMID: 10533321.

84.Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. Annual review of immunology. 2008;26:261-92. doi:

10.1146/annurev.immunol.26.021607.090248. PubMed PMID: 18370922; PMCID: 2707252.

85.Manis JP, Tian M, Alt FW. Mechanism and control of class-switch recombination. Trends in immunology. 2002;23(1):31-9. PubMed PMID: 11801452.

86.Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, Thomas LM, Al-Qahtani A, White CA, Park SR, Steinacker P, Li Z, Yates J, 3rd, Herron B, Otto M, Zan H, Fu H, Casali P. 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. Nature structural & molecular biology. 2010;17(9):1124-35. doi: 10.1038/nsmb.1884. PubMed PMID: 20729863; PMCID: 3645988.

87.Han L, Masani S, Yu K. Overlapping activation-induced cytidine deaminase hotspot motifs in Ig class-switch recombination. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(28):11584-9. doi: 10.1073/pnas.1018726108. PubMed PMID: 21709240; PMCID: 3136278.

88.Huang FT, Yu K, Balter BB, Selsing E, Oruc Z, Khamlichi AA, Hsieh CL, Lieber MR. Sequence dependence of chromosomal R-loops at the immunoglobulin heavy-chain Smu class switch region. Molecular and cellular biology. 2007;27(16):5921-32. doi: 10.1128/MCB.00702-07. PubMed PMID: 17562862; PMCID: 1952116.

89.Zan H, White CA, Thomas LM, Mai T, Li G, Xu Z, Zhang J, Casali P. Rev1 recruits ung to switch regions and enhances du glycosylation for immunoglobulin

class switch DNA recombination. Cell reports. 2012;2(5):1220-32. doi: 10.1016/j.celrep.2012.09.029. PubMed PMID: 23140944; PMCID: 3518390.

90.Lam T, Thomas LM, White CA, Li G, Pone EJ, Xu Z, Casali P. Scaffold functions of 14-3-3 adaptors in B cell immunoglobulin class switch DNA recombination. PloS one. 2013;8(11):e80414. doi: 10.1371/journal.pone.0080414. PubMed PMID: 24282540; PMCID: 3840166.

91.Masani S, Han L, Yu K. Apurinic/apyrimidinic endonuclease 1 is the essential nuclease during immunoglobulin class switch recombination. Molecular and cellular biology. 2013;33(7):1468-73. doi: 10.1128/MCB.00026-13. PubMed PMID: 23382073; PMCID: 3624277.

92.Dinkelmann M, Spehalski E, Stoneham T, Buis J, Wu Y, Sekiguchi JM, Ferguson DO. Multiple functions of MRN in end-joining pathways during isotype class switching. Nature structural & molecular biology. 2009;16(8):808-13. doi: 10.1038/nsmb.1639. PubMed PMID: 19633670; PMCID: 2721910.

93.Endo Y, Marusawa H, Kinoshita K, Morisawa T, Sakurai T, Okazaki IM, Watashi K, Shimotohno K, Honjo T, Chiba T. Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. Oncogene. 2007;26(38):5587-95. doi: 10.1038/sj.onc.1210344. PubMed PMID: 17404578.

94.Nagata N, Akiyama J, Marusawa H, Shimbo T, Liu Y, Igari T, Nakashima R, Watanabe H, Uemura N, Chiba T. Enhanced expression of activation-induced cytidine deaminase in human gastric mucosa infected by Helicobacter pylori and its decrease following eradication. Journal of gastroenterology. 2014;49(3):427-35. doi: 10.1007/s00535-013-0808-z. PubMed PMID: 23591766.

95.Batsaikhan BE, Kurita N, Iwata T, Sato H, Yoshikawa K, Takasu C, Kashihara H, Matsumoto N, Ishibashi H, Shimada M. The role of activation-induced cytidine deaminase expression in gastric adenocarcinoma. Anticancer research. 2014;34(2):995-1000. PubMed PMID: 24511045.

96.Munoz DP, Lee EL, Takayama S, Coppe JP, Heo SJ, Boffelli D, Di Noia JM, Martin DI. Activation-induced cytidine deaminase (AID) is necessary for the epithelial-mesenchymal transition in mammary epithelial cells. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(32):E2977-86. doi: 10.1073/pnas.1301021110. PubMed PMID: 23882083; PMCID: 3740878.

97.Nakanishi Y, Kondo S, Wakisaka N, Tsuji A, Endo K, Murono S, Ito M, Kitamura K, Muramatsu M, Yoshizaki T. Role of activation-induced cytidine deaminase in the development of oral squamous cell carcinoma. PloS one. 2013;8(4):e62066. doi: 10.1371/journal.pone.0062066. PubMed PMID: 23634222; PMCID: 3636261.

98.Gatto D, Brink R. The germinal center reaction. The Journal of allergy and clinical immunology. 2010;126(5):898-907; quiz 8-9. doi: 10.1016/j.jaci.2010.09.007. PubMed PMID: 21050940.

99.Nieuwenhuis P, Opstelten D. Functional anatomy of germinal centers. The American journal of anatomy. 1984;170(3):421-35. doi: 10.1002/aja.1001700315. PubMed PMID: 6383007.

100.Cyster JG, Ansel KM, Reif K, Ekland EH, Hyman PL, Tang HL, Luther SA, Ngo VN. Follicular stromal cells and lymphocyte homing to follicles. Immunological reviews. 2000;176:181-93. PubMed PMID: 11043777.

101.Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. Nature reviews Immunology. 2008;8(1):22-33. doi: 10.1038/nri2217. PubMed PMID: 18097447.

102.Allen CD, Ansel KM, Low C, Lesley R, Tamamura H, Fujii N, Cyster JG. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. Nature immunology. 2004;5(9):943-52. doi: 10.1038/ni1100. PubMed PMID: 15300245.

103.Caron G, Le Gallou S, Lamy T, Tarte K, Fest T. CXCR4 expression functionally discriminates centroblasts versus centrocytes within human germinal center B cells. Journal of immunology. 2009;182(12):7595-602. doi: 10.4049/jimmunol.0804272. PubMed PMID: 19494283.

104.Willis TG, Dyer MJ. The role of immunoglobulin translocations in the pathogenesis of B-cell malignancies. Blood. 2000;96(3):808-22. PubMed PMID: 10910891.

105.Ramiro A, Reina San-Martin B, McBride K, Jankovic M, Barreto V, Nussenzweig A, Nussenzweig MC. The role of activation-induced deaminase in antibody diversification and chromosome translocations. Advances in immunology. 2007;94:75-107. doi: 10.1016/S0065-2776(06)94003-6. PubMed PMID: 17560272.

106.Kuppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. Oncogene. 2001;20(40):5580-94. doi: 10.1038/sj.onc.1204640. PubMed PMID: 11607811.

107.Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, Honjo T, Nussenzweig A, Nussenzweig MC. AID is required for c-myc/IgH chromosome translocations in vivo. Cell. 2004;118(4):431-8. doi: 10.1016/j.cell.2004.08.006. PubMed PMID: 15315756.

108.Robbiani DF, Bothmer A, Callen E, Reina-San-Martin B, Dorsett Y, Difilippantonio S, Bolland DJ, Chen HT, Corcoran AE, Nussenzweig A, Nussenzweig MC. AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. Cell. 2008;135(6):1028-38. doi: 10.1016/j.cell.2008.09.062. PubMed PMID: 19070574; PMCID: 2713603.

109.Rabbitts TH, Forster A, Hamlyn P, Baer R. Effect of somatic mutation within translocated c-myc genes in Burkitt's lymphoma. Nature. 1984;309(5969):592-7. PubMed PMID: 6547209.

110.Bross L, Fukita Y, McBlane F, Demolliere C, Rajewsky K, Jacobs H. DNA double-strand breaks in immunoglobulin genes undergoing somatic hypermutation. Immunity. 2000;13(5):589-97. PubMed PMID: 11114372.

111.Papavasiliou FN, Schatz DG. Cell-cycle-regulated DNA double-stranded breaks in somatic hypermutation of immunoglobulin genes. Nature. 2000;408(6809):216-21. doi: 10.1038/35041599. PubMed PMID: 11089977.

112.Jiang Y, Soong TD, Wang L, Melnick AM, Elemento O. Genome-wide detection of genes targeted by non-Ig somatic hypermutation in lymphoma. PloS

one. 2012;7(7):e40332. doi: 10.1371/journal.pone.0040332. PubMed PMID: 22808135; PMCID: 3395700.

113.Pasqualucci L, Bhagat G, Jankovic M, Compagno M, Smith P, Muramatsu M, Honjo T, Morse HC, 3rd, Nussenzweig MC, Dalla-Favera R. AID is required for germinal center-derived lymphomagenesis. Nature genetics. 2008;40(1):108-12. doi: 10.1038/ng.2007.35. PubMed PMID: 18066064.

114.Dang CV, O'Donnell KA, Zeller KI, Nguyen T, Osthus RC, Li F. The c-Myc target gene network. Seminars in cancer biology. 2006;16(4):253-64. doi: 10.1016/j.semcancer.2006.07.014. PubMed PMID: 16904903.

115.Wade M, Wahl GM. c-Myc, genome instability, and tumorigenesis: the devil is in the details. Current topics in microbiology and immunology. 2006;302:169-203. PubMed PMID: 16620029.

116.Kovalchuk AL, Ansarah-Sobrinho C, Hakim O, Resch W, Tolarova H, Dubois W, Yamane A, Takizawa M, Klein I, Hager GL, Morse HC, 3rd, Potter M, Nussenzweig MC, Casellas R. Mouse model of endemic Burkitt translocations reveals the long-range boundaries of Ig-mediated oncogene deregulation. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(27):10972-7. doi: 10.1073/pnas.1200106109. PubMed PMID: 22711821; PMCID: 3390838.

117.Burkitt DP. The discovery of Burkitt's lymphoma. Cancer. 1983;51(10):1777-86. PubMed PMID: 6299496.

118.Shaffer AL, Rosenwald A, Staudt LM. Lymphoid malignancies: the dark side of B-cell differentiation. Nature reviews Immunology. 2002;2(12):920-32. doi: 10.1038/nri953. PubMed PMID: 12461565.

119.Shiramizu B, Barriga F, Neequaye J, Jafri A, Dalla-Favera R, Neri A, Guttierez M, Levine P, Magrath I. Patterns of chromosomal breakpoint locations in Burkitt's lymphoma: relevance to geography and Epstein-Barr virus association. Blood. 1991;77(7):1516-26. PubMed PMID: 1849033.

120.Onizuka T, Moriyama M, Yamochi T, Kuroda T, Kazama A, Kanazawa N, Sato K, Kato T, Ota H, Mori S. BCL-6 gene product, a 92- to 98-kD nuclear phosphoprotein, is highly expressed in germinal center B cells and their neoplastic counterparts. Blood. 1995;86(1):28-37. PubMed PMID: 7795234.

121.LaBadie J, Dunn WA, Aronson NN, Jr. Hepatic synthesis of carnitine from protein-bound trimethyl-lysine. Lysosomal digestion of methyl-lysine-labelled asialo-fetuin. The Biochemical journal. 1976;160(1):85-95. PubMed PMID: 64247; PMCID: 1164203.

122.Chapman CJ, Zhou JX, Gregory C, Rickinson AB, Stevenson FK. VH and VL gene analysis in sporadic Burkitt's lymphoma shows somatic hypermutation, intraclonal heterogeneity, and a role for antigen selection. Blood. 1996;88(9):3562-8. PubMed PMID: 8896424.

123.Pasqualucci L, Neumeister P, Goossens T, Nanjangud G, Chaganti RS, Kuppers R, Dalla-Favera R. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. Nature. 2001;412(6844):341-6. doi: 10.1038/35085588. PubMed PMID: 11460166.

124.Kuppers R. Mechanisms of B-cell lymphoma pathogenesis. Nature reviews Cancer. 2005;5(4):251-62. doi: 10.1038/nrc1589. PubMed PMID: 15803153.

125.Lossos IS, Levy R, Alizadeh AA. AID is expressed in germinal center B-cell-like and activated B-cell-like diffuse large-cell lymphomas and is not correlated with intraclonal heterogeneity. Leukemia. 2004;18(11):1775-9. doi: 10.1038/sj.leu.2403488. PubMed PMID: 15385936.

126.Baron BW, Nucifora G, McCabe N, Espinosa R, 3rd, Le Beau MM, McKeithan TW. Identification of the gene associated with the recurring chromosomal translocations t(3;14)(q27;q32) and t(3;22)(q27;q11) in B-cell lymphomas. Proceedings of the National Academy of Sciences of the United States of America. 1993;90(11):5262-6. PubMed PMID: 8506375; PMCID: 46696.

127.Hardianti MS, Tatsumi E, Syampurnawati M, Furuta K, Saigo K, Nakamachi Y, Kumagai S, Ohno H, Tanabe S, Uchida M, Yasuda N. Activation-induced cytidine deaminase expression in follicular lymphoma: association between AID expression and ongoing mutation in FL. Leukemia. 2004;18(4):826-31. doi: 10.1038/sj.leu.2403323. PubMed PMID: 14990977.

128.Loeffler M, Kreuz M, Haake A, Hasenclever D, Trautmann H, Arnold C, Winter K, Koch K, Klapper W, Scholtysik R, Rosolowski M, Hoffmann S, Ammerpohl O, Szczepanowski M, Herrmann D, Kuppers R, Pott C, Siebert R. Genomic and epigenomic co-evolution in follicular lymphomas. Leukemia. 2014. doi: 10.1038/leu.2014.209. PubMed PMID: 25027518.

129.Dalla-Favera R, Martinotti S, Gallo RC, Erikson J, Croce CM. Translocation and rearrangements of the c-myc oncogene locus in human undifferentiated B-cell lymphomas. Science. 1983;219(4587):963-7. PubMed PMID: 6401867.

130.Thorley-Lawson DA, Allday MJ. The curious case of the tumour virus: 50 years of Burkitt's lymphoma. Nature reviews Microbiology. 2008;6(12):913-24. doi: 10.1038/nrmicro2015. PubMed PMID: 19008891.

131.Scheller H, Tobollik S, Kutzera A, Eder M, Unterlehberg J, Pfeil I, Jungnickel B. c-Myc overexpression promotes a germinal center-like program in Burkitt's lymphoma. Oncogene. 2010;29(6):888-97. doi: 10.1038/onc.2009.377. PubMed PMID: 19881537.

132.Gaidano G, Ballerini P, Gong JZ, Inghirami G, Neri A, Newcomb EW, Magrath IT, Knowles DM, Dalla-Favera R. p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(12):5413-7. PubMed PMID: 2052620; PMCID: 51883.

133.Puiggros A, Blanco G, Espinet B. Genetic abnormalities in chronic lymphocytic leukemia: where we are and where we go. BioMed research international. 2014;2014:435983. doi: 10.1155/2014/435983. PubMed PMID: 24967369; PMCID: 4054680.

134.Hancer VS, Kose M, Diz-Kucukkaya R, Yavuz AS, Aktan M. Activation-induced cytidine deaminase mRNA levels in chronic lymphocytic leukemia. Leukemia & lymphoma. 2011;52(1):79-84. doi: 10.3109/10428194.2010.531410. PubMed PMID: 21133730.

135.Gelmez MY, Teker AB, Aday AD, Yavuz AS, Soysal T, Deniz G, Aktan M. Analysis of activation-induced cytidine deaminase mRNA levels in patients with chronic lymphocytic leukemia with different cytogenetic status. Leukemia & lymphoma. 2014;55(2):326-30. doi: 10.3109/10428194.2013.803225. PubMed PMID: 23662991.

136.Schwering I, Brauninger A, Klein U, Jungnickel B, Tinguely M, Diehl V, Hansmann ML, Dalla-Favera R, Rajewsky K, Kuppers R. Loss of the B-lineage-specific gene expression program in Hodgkin and Reed-Sternberg cells of Hodgkin lymphoma. Blood. 2003;101(4):1505-12. doi: 10.1182/blood-2002-03-0839. PubMed PMID: 12393731.

137.Greiner A, Tobollik S, Buettner M, Jungnickel B, Herrmann K, Kremmer E, Niedobitek G. Differential expression of activation-induced cytidine deaminase (AID) in nodular lymphocyte-predominant and classical Hodgkin lymphoma. The Journal of pathology. 2005;205(5):541-7. doi: 10.1002/path.1746. PubMed PMID: 15732141.

138.Pratt G, Fenton JA, Davies FE, Rawstron AC, Richards SJ, Collins JE, Owen RG, Jack AS, Smith GM, Morgan GJ. Insertional events as well as translocations may arise during aberrant immunoglobulin switch recombination in a patient with multiple myeloma. British journal of haematology. 2001;112(2):388-91. PubMed PMID: 11167836.

139.Martin LD, Belch AR, Pilarski LM. Promiscuity of translocation partners in multiple myeloma. Journal of cellular biochemistry. 2010;109(6):1085-94. doi: 10.1002/jcb.22499. PubMed PMID: 20127714.

140.Pasqualucci L, Guglielmino R, Houldsworth J, Mohr J, Aoufouchi S, Polakiewicz R, Chaganti RS, Dalla-Favera R. Expression of the AID protein in normal and neoplastic B cells. Blood. 2004;104(10):3318-25. doi: 10.1182/blood-2004-04-1558. PubMed PMID: 15304391.

141.A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. Blood. 1997;89(11):3909-18. PubMed PMID: 9166827.

142.Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(17):9991-6. doi: 10.1073/pnas.1732008100. PubMed PMID: 12900505; PMCID: 187912.

143.Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J, Jr., Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403(6769):503-11. doi: 10.1038/35000501. PubMed PMID: 10676951.

144.Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, Lee AH, Qian SB, Zhao H, Yu X, Yang L, Tan BK, Rosenwald A, Hurt EM, Petroulakis E, Sonenberg N, Yewdell JW, Calame K, Glimcher LH, Staudt LM. XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in

plasma cell differentiation. Immunity. 2004;21(1):81-93. doi: 10.1016/j.immuni.2004.06.010. PubMed PMID: 15345222.

145.Ladanyi M, Offit K, Jhanwar SC, Filippa DA, Chaganti RS. MYC rearrangement and translocations involving band 8q24 in diffuse large cell lymphomas. Blood. 1991;77(5):1057-63. PubMed PMID: 1671647.

146.Weiss LM, Warnke RA, Sklar J, Cleary ML. Molecular analysis of the t(14;18) chromosomal translocation in malignant lymphomas. The New England journal of medicine. 1987;317(19):1185-9. doi: 10.1056/NEJM198711053171904. PubMed PMID: 3657890.

147.Lenz G, Nagel I, Siebert R, Roschke AV, Sanger W, Wright GW, Dave SS, Tan B, Zhao H, Rosenwald A, Muller-Hermelink HK, Gascoyne RD, Campo E, Jaffe ES, Smeland EB, Fisher RI, Kuehl WM, Chan WC, Staudt LM. Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell-like diffuse large B cell lymphoma. The Journal of experimental medicine. 2007;204(3):633-43. doi: 10.1084/jem.20062041. PubMed PMID: 17353367; PMCID: 2137913.

148.Khodabakhshi AH, Morin RD, Fejes AP, Mungall AJ, Mungall KL, Bolger-Munro M, Johnson NA, Connors JM, Gascoyne RD, Marra MA, Birol I, Jones SJ. Recurrent targets of aberrant somatic hypermutation in lymphoma. Oncotarget. 2012;3(11):1308-19. PubMed PMID: 23131835; PMCID: 3717795.

149.Jaffe ES HN, Stein H,, JW. W. Follicular Lymphoma. In: Banks PM WR, editor. Pathology and Genetics of Tumours of Haemopoetic and Lymphoid

Tissues World Health Organization Classification of Tumours. Lyon: IARC Press; 2001. p. 162-7.

150.McDonnell TJ, Deane N, Platt FM, Nunez G, Jaeger U, McKearn JP, Korsmeyer SJ. bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation. Cell. 1989;57(1):79-88. PubMed PMID: 2649247.

151.Pasqualucci L, Migliazza A, Fracchiolla N, William C, Neri A, Baldini L, Chaganti RS, Klein U, Kuppers R, Rajewsky K, Dalla-Favera R. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(20):11816-21. PubMed PMID: 9751748; PMCID: 21723.

152.Spence JM, Abumoussa A, Spence JP, Burack WR. Intraclonal diversity in follicular lymphoma analyzed by quantitative ultradeep sequencing of noncoding regions. Journal of immunology. 2014;193(10):4888-94. doi: 10.4049/jimmunol.1401699. PubMed PMID: 25311808; PMCID: 4225181.

153.M. H. New concepts in the pathogenesis, diagnosis, prognostic factors, and clinical presentation of chronic lymphocytic leukemia. Rev Clin Exp Hematol. 2000;2:103-17.

154.Naylor M, Capra JD. Mutational status of Ig V(H) genes provides clinically valuable information in B-cell chronic lymphocytic leukemia. Blood. 1999;94(6):1837-9. PubMed PMID: 10477711.

155.Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 1999;94(6):1848-54. PubMed PMID: 10477713.

156.McCarthy H, Wierda WG, Barron LL, Cromwell CC, Wang J, Coombes KR, Rangel R, Elenitoba-Johnson KS, Keating MJ, Abruzzo LV. High expression of activation-induced cytidine deaminase (AID) and splice variants is a distinctive feature of poor-prognosis chronic lymphocytic leukemia. Blood. 2003;101(12):4903-8. doi: 10.1182/blood-2002-09-2906. PubMed PMID: 12586616.

157.Stankovic T, Weber P, Stewart G, Bedenham T, Murray J, Byrd PJ, Moss PA, Taylor AM. Inactivation of ataxia telangiectasia mutated gene in B-cell chronic lymphocytic leukaemia. Lancet. 1999;353(9146):26-9. doi: 10.1016/S0140-6736(98)10117-4. PubMed PMID: 10023947.

158.Mateo M, Mollejo M, Villuendas R, Algara P, Sanchez-Beato M, Martinez P, Piris MA. 7q31-32 allelic loss is a frequent finding in splenic marginal zone lymphoma. The American journal of pathology. 1999;154(5):1583-9. doi: 10.1016/S0002-9440(10)65411-9. PubMed PMID: 10329610; PMCID: 1866606.

159.Kuppers R. Molecular biology of Hodgkin's lymphoma. Advances in cancer research. 2002;84:277-312. PubMed PMID: 11883530.

160.Kanzler H, Kuppers R, Hansmann ML, Rajewsky K. Hodgkin and Reed-Sternberg cells in Hodgkin's disease represent the outgrowth of a dominant tumor clone derived from (crippled) germinal center B cells. The Journal of

experimental medicine. 1996;184(4):1495-505. PubMed PMID: 8879220; PMCID: 2192840.

161.Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjord JE, Foekens JA, Greaves M, Hosoda F, Hutter B, Ilicic T, Imbeaud S, Imielinski M, Jager N, Jones DT, Jones D, Knappskog S, Kool M, Lakhani SR, Lopez-Otin C, Martin S, Munshi NC, Nakamura H, Northcott PA, Pajic M, Papaemmanuil E, Paradiso A, Pearson JV, Puente XS, Raine K, Ramakrishna M, Richardson AL, Richter J, Rosenstiel P, Schlesner M, Schumacher TN, Span PN, Teague JW, Totoki Y, Tutt AN, Valdes-Mas R, van Buuren MM, van 't Veer L, Vincent-Salomon A, Waddell N, Yates LR, Australian Pancreatic Cancer Genome I, Consortium IBC, Consortium IM-S, PedBrain I, Zucman-Rossi J, Futreal PA, McDermott U, Lichter P, Meyerson M, Grimmond SM, Siebert R, Campo E, Shibata T, Pfister SM, Campbell PJ, Stratton MR. Signatures of mutational processes in human cancer. Nature. 2013;500(7463):415-21. doi: 10.1038/nature12477. PubMed PMID: 23945592; PMCID: 3776390.

162.Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, Kiezun A, Kryukov GV, Carter SL, Saksena G, Harris S, Shah RR, Resnick MA, Getz G, Gordenin DA. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nature genetics. 2013;45(9):970-6. doi: 10.1038/ng.2702. PubMed PMID: 23852170; PMCID: 3789062.

163.Maul RW, Saribasak H, Martomo SA, McClure RL, Yang W, Vaisman A, Gramlich HS, Schatz DG, Woodgate R, Wilson DM, 3rd, Gearhart PJ. Uracil residues dependent on the deaminase AID in immunoglobulin gene variable and switch regions. Nature immunology. 2011;12(1):70-6. doi: 10.1038/ni.1970. PubMed PMID: 21151102; PMCID: 3653439.

164.Roche B, Claes A, Rougeon F. Deoxyuridine triphosphate incorporation during somatic hypermutation of mouse VkOx genes after immunization with phenyloxazolone. Journal of immunology. 2010;185(8):4777-82. doi: 10.4049/jimmunol.1001459. PubMed PMID: 20861355.

165.Shalhout S, Haddad D, Sosin A, Holland TC, Al-Katib A, Martin A, Bhagwat AS. Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development. Molecular and cellular biology. 2014;34(21):4019-32. doi: 10.1128/MCB.00589-14. PubMed PMID: 25154417.

166.Zhang W, Bardwell PD, Woo CJ, Poltoratsky V, Scharff MD, Martin A. Clonal instability of V region hypermutation in the Ramos Burkitt's lymphoma cell line. International immunology. 2001;13(9):1175-84. PubMed PMID: 11526098.

167.Mohammad RM, Mohamed AN, Smith MR, Jawadi NS, al-Katib A. A unique EBV-negative low-grade lymphoma line (WSU-FSCCL) exhibiting both t(14;18) and t(8;11). Cancer genetics and cytogenetics. 1993;70(1):62-7. PubMed PMID: 8221615.

168.Al-Katib AM, Smith MR, Kamanda WS, Pettit GR, Hamdan M, Mohamed AN, Chelladurai B, Mohammad RM. Bryostatin 1 down-regulates mdr1 and potentiates vincristine cytotoxicity in diffuse large cell lymphoma xenografts.

Clinical cancer research : an official journal of the American Association for Cancer Research. 1998;4(5):1305-14. PubMed PMID: 9607591.

169.Hardin J, MacLeod S, Grigorieva I, Chang R, Barlogie B, Xiao H, Epstein J. Interleukin-6 prevents dexamethasone-induced myeloma cell death. Blood. 1994;84(9):3063-70. PubMed PMID: 7949178.

170.Mohamed AN, al-Katib A. Establishment and characterization of a human lymphoma cell line (WSU-NHL) with 14;18 translocation. Leukemia research. 1988;12(10):833-43. PubMed PMID: 3143865.

171.Kubo K, Ide H, Wallace SS, Kow YW. A novel, sensitive, and specific assay for abasic sites, the most commonly produced DNA lesion. Biochemistry. 1992;31(14):3703-8. PubMed PMID: 1567824.

172.J. Sambrook DWR. Molecular Cloning: A Labortory Manual. the third edition ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2001.

173.Cabelof DC, Nakamura J, Heydari AR. A sensitive biochemical assay for the detection of uracil. Environmental and molecular mutagenesis. 2006;47(1):31-7. doi: 10.1002/em.20165. PubMed PMID: 16106443.

174.Lari SU, Chen CY, Vertessy BG, Morre J, Bennett SE. Quantitative determination of uracil residues in Escherichia coli DNA: Contribution of ung, dug, and dut genes to uracil avoidance. DNA repair. 2006;5(12):1407-20. doi: 10.1016/j.dnarep.2006.06.009. PubMed PMID: 16908222; PMCID: 3040120.

175.Parisien RB, A. S. DNA and RNA Modification Enzymes: Comparative Structure, Mechanism, Functions, Cellular Interactions and Evolution. Austin, Texas: Landes Bioscience; 2009.

176.Gallinari P, Jiricny J. A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. Nature. 1996;383(6602):735-8. doi: 10.1038/383735a0. PubMed PMID: 8878487.

177.Robbiani DF, Nussenzweig MC. Chromosome translocation, B cell lymphoma, and activation-induced cytidine deaminase. Annual review of pathology. 2013;8:79-103. doi: 10.1146/annurev-pathol-020712-164004. PubMed PMID: 22974238.

178.Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH, Schatz DG. Two levels of protection for the B cell genome during somatic hypermutation. Nature. 2008;451(7180):841-5. doi: 10.1038/nature06547. PubMed PMID: 18273020.

179.Winter DB, Phung QH, Umar A, Baker SM, Tarone RE, Tanaka K, Liskay RM, Kunkel TA, Bohr VA, Gearhart PJ. Altered spectra of hypermutation in antibodies from mice deficient for the DNA mismatch repair protein PMS2. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(12):6953-8. PubMed PMID: 9618520; PMCID: 22699.

180.Unniraman S, Schatz DG. Strand-biased spreading of mutations during somatic hypermutation. Science. 2007;317(5842):1227-30. doi: 10.1126/science.1145065. PubMed PMID: 17761884.

181.Doseth B, Visnes T, Wallenius A, Ericsson I, Sarno A, Pettersen HS, Flatberg A, Catterall T, Slupphaug G, Krokan HE, Kavli B. Uracil-DNA glycosylase in base excision repair and adaptive immunity: species differences between man and mouse. The Journal of biological chemistry.

2011;286(19):16669-80. doi: 10.1074/jbc.M111.230052. PubMed PMID: 21454529; PMCID: 3089509.

182.Fritz EL, Rosenberg BR, Lay K, Mihailovic A, Tuschl T, Papavasiliou FN. A comprehensive analysis of the effects of the deaminase AID on the transcriptome and methylome of activated B cells. Nature immunology. 2013;14(7):749-55. doi: 10.1038/ni.2616. PubMed PMID: 23708250; PMCID: 3688651.

183.Dingler FA, Kemmerich K, Neuberger MS, Rada C. Uracil excision by endogenous SMUG1 glycosylase promotes efficient Ig class switching and impacts on A:T substitutions during somatic mutation. European journal of immunology. 2014;44(7):1925-35. doi: 10.1002/eji.201444482. PubMed PMID: 24771041; PMCID: 4158878.

184.Liu JQ, Joshi PS, Wang C, El-Omrani HY, Xiao Y, Liu X, Hagan JP, Liu CG, Wu LC, Bai XF. Targeting activation-induced cytidine deaminase overcomes tumor evasion of immunotherapy by CTLs. Journal of immunology. 2010;184(10):5435-43. doi: 10.4049/jimmunol.0903322. PubMed PMID: 20404277; PMCID: 2874093.

185.Seitz V, Hummel M, Walter J, Stein H. Evolution of classic Hodgkin lymphoma in correlation to changes in the lymphoid organ structure of vertebrates. Developmental and comparative immunology. 2003;27(1):43-53. PubMed PMID: 12477500.

186.Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74. doi: 10.1016/j.cell.2011.02.013. PubMed PMID: 21376230.

187.Liu M, Schatz DG. Balancing AID and DNA repair during somatic hypermutation. Trends in immunology. 2009;30(4):173-81. doi: 10.1016/j.it.2009.01.007. PubMed PMID: 19303358.

188.Maul RW, Gearhart PJ. AID and somatic hypermutation. Advances in immunology. 2010;105:159-91. doi: 10.1016/S0065-2776(10)05006-6. PubMed PMID: 20510733; PMCID: 2954419.

189.Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, Reina San-Martin B, Barreto V, Nieland TJ, Root DE, Casellas R, Nussenzweig MC. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. Cell. 2010;143(1):122-33. doi: 10.1016/j.cell.2010.09.017. PubMed PMID: 20887897; PMCID: 2993080.

190.Ide H, Akamatsu K, Kimura Y, Michiue K, Makino K, Asaeda A, Takamori Y, Kubo K. Synthesis and damage specificity of a novel probe for the detection of abasic sites in DNA. Biochemistry. 1993;32(32):8276-83. PubMed PMID: 8347625.

191.Friedberg EW, GC; Siede, W; Wood, RD; Schultz, RA; Ellenberger. DNA Repair and Mutagenesis ASM Press; 2006.

192.Kavli B, Sundheim O, Akbari M, Otterlei M, Nilsen H, Skorpen F, Aas PA, Hagen L, Krokan HE, Slupphaug G. hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup. The Journal of biological

chemistry. 2002;277(42):39926-36. doi: 10.1074/jbc.M207107200. PubMed PMID: 12161446.

193.Kim N, Jinks-Robertson S. dUTP incorporation into genomic DNA is linked to transcription in yeast. Nature. 2009;459(7250):1150-3. doi: 10.1038/nature08033. PubMed PMID: 19448611; PMCID: 2730915.

194.Galashevskaya A, Sarno A, Vagbo CB, Aas PA, Hagen L, Slupphaug G, Krokan HE. A robust, sensitive assay for genomic uracil determination by LC/MS/MS reveals lower levels than previously reported. DNA repair. 2013;12(9):699-706. doi: 10.1016/j.dnarep.2013.05.002. PubMed PMID: 23742752.

195.Mashiyama ST, Hansen CM, Roitman E, Sarmiento S, Leklem JE, Shultz TD, Ames BN. An assay for uracil in human DNA at baseline: effect of marginal vitamin B6 deficiency. Analytical biochemistry. 2008;372(1):21-31. doi: 10.1016/j.ab.2007.08.034. PubMed PMID: 17963712; PMCID: 2175266.

196.Nilsen H, Rosewell I, Robins P, Skjelbred CF, Andersen S, Slupphaug G, Daly G, Krokan HE, Lindahl T, Barnes DE. Uracil-DNA glycosylase (UNG)-deficient mice reveal a primary role of the enzyme during DNA replication. Molecular cell. 2000;5(6):1059-65. PubMed PMID: 10912000.

197.Li G, Zan H, Xu Z, Casali P. Epigenetics of the antibody response. Trends in immunology. 2013;34(9):460-70. doi: 10.1016/j.it.2013.03.006. PubMed PMID: 23643790; PMCID: 3744588.

198.Storb U. Why does somatic hypermutation by AID require transcription of its target genes? Advances in immunology. 2014;122:253-77. doi: 10.1016/B978-0-12-800267-4.00007-9. PubMed PMID: 24507160.

199.Hagen L, Kavli B, Sousa MM, Torseth K, Liabakk NB, Sundheim O, Pena-Diaz J, Otterlei M, Horning O, Jensen ON, Krokan HE, Slupphaug G. Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. The EMBO journal. 2008;27(1):51-61. doi: 10.1038/sj.emboj.7601958. PubMed PMID: 18079698; PMCID: 2147998.

200.An Q, Robins P, Lindahl T, Barnes DE. C --> T mutagenesis and gamma-radiation sensitivity due to deficiency in the Smug1 and Ung DNA glycosylases. The EMBO journal. 2005;24(12):2205-13. doi: 10.1038/sj.emboj.7600689. PubMed PMID: 15902269; PMCID: 1150883.

201.Nilsen H, Stamp G, Andersen S, Hrivnak G, Krokan HE, Lindahl T, Barnes DE. Gene-targeted mice lacking the Ung uracil-DNA glycosylase develop B-cell lymphomas. Oncogene. 2003;22(35):5381-6. doi: 10.1038/sj.onc.1206860. PubMed PMID: 12934097.

202.Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, McBride DJ, Bignell GR, Cooke SL, Shlien A, Gamble J, Whitmore I, Maddison M, Tarpey PS, Davies HR, Papaemmanuil E, Stephens PJ, McLaren S, Butler AP, Teague JW, Jonsson G, Garber JE, Silver D, Miron P, Fatima A, Boyault S, Langerod A, Tutt A, Martens JW, Aparicio SA, Borg A, Salomon AV, Thomas G,

Borresen-Dale AL, Richardson AL, Neuberger MS, Futreal PA, Campbell PJ, Stratton MR, Breast Cancer Working Group of the International Cancer Genome C. Mutational processes molding the genomes of 21 breast cancers. Cell. 2012;149(5):979-93. doi: 10.1016/j.cell.2012.04.024. PubMed PMID: 22608084; PMCID: 3414841.

203.Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB, Yee D, Temiz NA, Donohue DE, McDougle RM, Brown WL, Law EK, Harris RS. APOBEC3B is an enzymatic source of mutation in breast cancer. Nature. 2013;494(7437):366-70. doi: 10.1038/nature11881. PubMed PMID: 23389445; PMCID: 3907282.

204.Nussenzweig A, Nussenzweig MC. Origin of chromosomal translocations in lymphoid cancer. Cell. 2010;141(1):27-38. doi: 10.1016/j.cell.2010.03.016. PubMed PMID: 20371343; PMCID: 2874895.

205.Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, Myers DR, Choi VW, Compagno M, Malkin DJ, Neuberg D, Monti S, Giallourakis CC, Gostissa M, Alt FW. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. Cell. 2011;147(1):107-19. doi: 10.1016/j.cell.2011.07.049. PubMed PMID: 21962511; PMCID: 3186939.

206.Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, Di Virgilio M, Bothmer A, Nussenzweig A, Robbiani DF, Casellas R, Nussenzweig MC. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell. 2011;147(1):95-106. doi: 10.1016/j.cell.2011.07.048. PubMed PMID: 21962510; PMCID: 3190307.

207.Liu L, Taverna P, Whitacre CM, Chatterjee S, Gerson SL. Pharmacologic disruption of base excision repair sensitizes mismatch repair-deficient and -proficient colon cancer cells to methylating agents. Clinical cancer research : an official journal of the American Association for Cancer Research. 1999;5(10):2908-17. PubMed PMID: 10537360.

208.Klemm L, Duy C, Iacobucci I, Kuchen S, von Levetzow G, Feldhahn N, Henke N, Li Z, Hoffmann TK, Kim YM, Hofmann WK, Jumaa H, Groffen J, Heisterkamp N, Martinelli G, Lieber MR, Casellas R, Muschen M. The B cell mutator AID promotes B lymphoid blast crisis and drug resistance in chronic myeloid leukemia. Cancer cell. 2009;16(3):232-45. doi: 10.1016/j.ccr.2009.07.030. PubMed PMID: 19732723; PMCID: 2931825.

ABSTRACT

THE LOSS OF GENOMIC URACIL HOMEOSTASIS AND AID-DEPENDENT ACCUMULATION OF DNA DAMAGE IN B CELL LYMPHOMAS

by

SOPHIA SHALHOUT

May 2015

Advisor: Dr. Ashok S. Bhagwat

Major: Chemistry (Biochemistry)

Degree: Doctor of Philosophy

Activation-induced deaminase (AID) is a sequence-selective DNA cytosine deaminase that introduces uracils in immunoglobulin genes. This DNA mutator is required for somatic hypermutation and class switch recombination- processes involved in the affinity maturation and diversification of antibodies. AID, however, can also lead to deleterious mutations and translocations promoting lymphomagenesis. The introduction of uracils throughout the genome of activated B cells and the ability of UNG2 glycosylase to excise these uracils is examined here. This interplay was also studied in cancerous B cells, with different results emerging in transformed cells versus healthy cells. Genomic uracil levels are found to remain at the same level in normal B cells stimulated to express AID. However the increase in uracils by 11- to 60- fold detected in stimulated B cells deficient in UNG, suggests that normal B cells do accumulate high levels of genomic uracils. However, these uracils are efficiently and effectively removed in UNG proficient cells, suggesting a balance or homeostasis between uracil creation and elimination. Interestingly, murine B cell cancer lines,

human B cell cancer cell lines, and several human B cell patient tumors from several types of lymphomas overexpressing AID, were found to accumulate genomic uracils at levels comparable to those seen in activated UNG-/- B cells. These lymphoma/leukemia cells are not defective in uracil removal and express UNG2 gene at the same level in normal peripheral B cells or higher. In addition, they also have similar nuclear uracil excision activities suggesting uracils accumulate despite robust uracil excision capabilities. These results suggest that the homeostasis of uracil introduction and excision seen in normal stimulated B cells is disrupted in lymphomas overexpressing AID despite UNG2 expression and activity, resulting in the accumulation of high levels of genome wide uracils.

The majority of human B cell lymphomas do in fact derive from cells that have undergone the germinal center reaction and are associated with the expression of AID. The high levels of uracils that accumulate in the genomes of B cell lymphomas suggests other types of DNA damage may also be present in an AID-dependent manner as the cell attempts to remove and repair these uracils. Here, we see that human B cell lymphoma cell lines that contain high uracil loads also accumulate elevated levels of other types of genome wide DNA damage dependent on AID. These include abasic sites and single- and double-strand breaks. In addition, the accumulation of DNA lesions reduces cell viability. These lesions are due in part to the attempted repair of uracils by uracil DNA glycosylase.

AUTOBIOGRAPHICAL STATEMENT

SOPHIA SHALHOUT

Education

Ph.D.	Chemistry – Major in Biochemistry, 4.0 GPA Wayne State University, Detroit, Michigan Research Advisor– Prof. Ashok Bhagwat	2009-present
B.S.	Biochemistry & Chemical Biology Wayne State University, Detroit, Michigan	2009
B.S.	Biological Sciences Wayne State University, Detroit, Michigan	2009

Publications

Shalhout, S., Haddad, D., Sosin A.,, Holland, T., Al-Katib, A., Martin, A., Bhagwat, A. S. (2014). Genomic Uracil Homeostasis during Normal B Cell Maturation and Loss of this Balance During B Cell Cancer Development. *Molecular and Cellular Biology.* 34(21):4019-32.

Shalhout, S.*, Betham, B.*, Marquez, V.E., Bhagwat, A.S. (2010). Use of Drosophila deoxynucleoside kinase to study mechanism of toxicity and mutagenicity of deoxycytidine analogs in Escherichia coli. *DNA Repair* **9**, 153-160. *co-first authors

Wei S, **Shalhout S**, Ahn Y-H, Bhagwat AS. (2015) A Versatile New Tool to Label Abasic Sites in DNA and Inhibit Base Excision Repair. *DNA Repair,* **27,** 9-18.

Shalhout, S., Haddad, D., Martin, A., Bhagwat, A. S. Multiple Types of DNA Damage in Germinal Center-derived Human B cell Lymphomas Expressing AID. (manuscript submitted and under review)

Scholarships & Awards

Summer Dissertation Fellowship Award	2013-2014
Paul C. Schaap-Rumble Graduate Professional Research Fellowship	2013-2014
Biological Chemistry Outstanding Graduate Student Award	2013
Graduate and Professional Student Travel Award	2013
Willard R. Lenz Memorial Fellowship	2012
Chemistry Department (WSU) Award for Excellence in Teaching	2011
Chemistry Department (WSU) Award for Excellence in Service	2011
Thomas C. Rumble Fellowship	2012
Herbert K. Livingston Award for Excellence in Teaching	2010-2011
Chemistry Department (WSU) Award for Excellence in Teaching	2009
Presidential Scholarship, WSU	2004
Michigan Competitive Scholarship	2004