2008

Role of AID and microRNA-155 in c-myc-IgH Translocations

Yair Dorsett

Role of AID and microRNA-155 in *c-myc-IgH* translocations

A Thesis Presented to the Faculty of
The Rockefeller University
In Partial Fulfillment of the requirements for
the degree of Doctor of Philosophy

By

Yair Dorsett

June 2008

Role of AID and microRNA-155 in *c-myc-IgH* translocations

Yair Dorsett, Ph.D

The Rockefeller University 2008

Chromosome translocations between oncogenes and the immunoglobulin (Ig) region spanning the variable (*V*), diversity (*D*) and joining (*J*) genes (*Ig V-J_H region*) are found in a number of mature B cell lymphomas in humans and mice. The breakpoints are frequently adjacent to the recombination signal sequences (RSSs) targeted by recombinase activating genes 1 and 2 (*RAG*1/2) during antigen receptor assembly in pre-B cells, suggesting that these translocations might be the result of aberrant V(D)J recombination. However, in mature B cells undergoing AID dependent somatic hypermutation (SHM), duplications or deletions that would necessitate a double strand break make up 6% of all the *Ig V-J_H* region associated somatic mutations. Furthermore, DNA breaks can be detected at this locus in B cells undergoing SHM. To determine whether SHM might induce *c-myc to Ig V-J_H* translocations, we searched for such events in both *IL6* transgenic (*IL6* tg) and *$AID^{-/-}$ IL6* tg mice. Our experiments demonstrate that *AID* is required for *c-myc* to *Ig V-J_H* translocations induced by IL6.

We also investigated the potential role of microRNAs (miRNAs) in AID mediated processes. MiRNAs are small non-coding RNAs that regulate vast networks of genes that share miRNA target sequences. To examine the physiologic effects of an individual miRNA *in vivo* we created a knock-in mouse that carries a mutation in the putative microRNA-155 (miR-155) target site in the 3'UTR of AID (*AID^{155}* mice). *AID^{155}* causes an increase in steady state AID mRNA and protein levels by increasing the half-life of

the mRNA resulting in high levels of *c-myc-IgH* translocations. A similar but more pronounced translocation phenotype was also found in *bic*/miR-155$^{-/-}$ mice. Our experiments indicate that miR-155 can act as a tumor suppressor by reducing potentially oncogenic translocations generated by AID.

Acknowledgments

I would first like to thank my advisor Michel Nussenzweig for all his guidance and advice on topics of both science and life that helped me get through my Ph.D. His guidance helped me develop as a scientist and gave me a new perspective on how I should carry out my future scientific endeavors, for which I am forever grateful. I could not imagine a better supervisor and lab to do a Ph.D.

I would also like to thank Davide F. Robbiani for sitting next to me and having everlasting patience when answering my thousands of questions. I would also like to thank him for all the productive scientific discussions and his guidance. I would also like to thank Anna Gazumyan, Mila Jankovic and Kevin M. McBride for always making themselves available for help with experiments, answering countless questions and educating me through scientific discussions. I would also like to thank Chrisitian Munz and Nina Papavasiliou for advice and suggestions on experiments. I would also like to thank Nina for all her support through out my Ph.D.

I would also like to thank my dad, Dale Dorsett, for all his scientific and life advice. My mom, Ziva Mislovin , for her constant and persistent advice on every possible topic. Also my sister, Maia Dorsett, for all the scientific discussions. Last but not least, I would like to thank my wife, Yanjiao Zhou, for everything.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Antibody Diversification In The Germinal Center

Germinal center response

Antibody genes are assembled in developing B cells by RAG1/RAG2 recombinase mediated site-specific recombination of immunoglobulin (Ig) variable (V), diversity (D) and joining (J) gene segments. Although V(D)J recombination produces a diverse repertoire of IgM antibodies, high affinity IgG antibody responses require further Ig diversification. During T cell dependent immune responses, structures called germinal centers (GC) are formed by rapidly dividing B cells within follicles of several different lymphoid tissues. Within this structure, B cell antibody genes are mutated at the rate of 10^{-3} mutations per base pair in order to create high affinity antibodies in a process termed somatic hypermutation (SHM)(*1, 2*). Antibody genes undergo further changes within the GC in the process of class switch recombination (CSR), in which the effector function of antibody genes is diversified through replacement of the heavy chain constant region. Somatic mutation and CSR can also occur outside the germinal center in T cell independent immune responses.

Activation induced cytidine deaminase

Both SHM and CSR are initiated by activation induced cytidine deaminase (AID), which deaminates cytosine residues and introduces U:G mismatches in transcribed

portions of DNA (*3, 4*). AID was originally identified as a mutated gene associated with hyper IgM syndrome (*5, 6*). Deletion of the gene in mouse also resulted in absence of class switch recombination and somatic hypermutation, confirming that AID played an essential role in germinal center gene diversification (*6*). Due to the high homology of AID with the RNA editing enzyme Apobec-1, it was originally proposed that AID functioned as an RNA editing enzyme that edited the mRNA of a putative mutator enzyme that acted directly on DNA to introduce mutations (*6*). However, several investigations showed that AID acts directly on DNA to deaminate cytidine on transcribed portions of DNA (*7-9*).

The lesions introduced by AID are processed by ubiquitous DNA repair pathways and error prone polymerases to produce the somatic mutations or double strand DNA breaks, that are obligate intermediates in immunoglobulin class switch recombination (*3, 4*). AID activity is primarily restricted to Ig genes, but can also produce off target lesions in non-Ig sites such as oncogenes (*10*). In addition, the double strand breaks created by AID can be substrates for translocation (*11, 12*). Therefore, the regulation of AID is essential to maintain genomic integrity.

Somatic hypermutation

SHM occurs on transcribed portions of Ig genes in GC activated B cells in order increase affinity to antigen(*1, 3, 7-9, 13, 14*). Mutations are found preferentially at RGYW motifs(*15*), starting 150bp downstream of the transcription start site and extend up to 2 kb, within an exponential decrease in mutation rate moving away from transcription start

site (*16*). It is thus believed that AID physically interacts with the transcription elongation complex to deaminate cytidine (*17*) and that AID binds to single stranded DNA, likely through association with RPA (*18, 19*), within the transcription bubble.

Somatic hypermutation occurs through a two-step process (Figure 1, borrowed from (*20*)), in which lesions in DNA are first generated by cytidine deamination followed by the repair of those lesions by a particular error prone DNA repair mechanism. Deaminated cytidines create U:G mismatches in DNA that can be replicated to create C to T transition mutations. Alternatively, the mismatch can be repaired by either the base excision repair pathway (BER) (reviewed in (*3*)) or by the mismatch repair pathway (MMR) (reviewed in(*1*)). The base excision repair pathway depends on the action of uracil DNA glycosylase (UNG), which catalyzes the hydrolysis of the N-glycosylic bond between the uracil and sugar, leaving an abasic site in uracil-containing single or double-stranded DNA. DNA synthesis opposite the abasic site can introduce any nucleotide. Thus, BER generates both transition and transversion mutations at the deaminated cytosine. The abasic site can also be excised via the APE endonuclease and repaired by error free BER.

Alternatively, dU:dG mismatches can be targeted by the mismatch repair machinery by Msh2/Msh6/Exo1 and error prone polymerases (*21-23*). The action of an unidentified endonuclease (possibly Mre11-RAD50-NBS1(*24*)) and Exo1 create single stranded gaps that can then be filled in by error prone polymerases (reviewed in (*2*)). Thus, MMR generates transitions and transversions up to 20nts on either side of the deaminated cytidine (*25*). The action of Exo1 is restricted to the non-transcribed strand,

as suggested by the recent observation that MMR occurs asymmetrically, only introducing mutations on the non-transcribed strand (*25*) Deficiencies in several other MMR components, including Mlh1, Mlh3 and Pms2, result in an altered spectra of mutations at switch-switch junctions, also implicating these factors in SHM (*21, 23, 26*).

The potential role of ten different DNA polymerases in SHM has been analyzed. These polymerases include pol β, δ, ζ, η, ι, κ, λ, μ, θ and Rev1 (*27-45*). Of these polymerases, pol μ, λ, β, δ and κ are apparently not involved in SHM (reviewed in (*46, 47*)). The rest have different roles in SHM, depending on there activity. Of all these DNA polymerases, pol η has the largest role in hypermutation, as deletion of pol η results in a two-thirds reduction in substitutions at dA:dT base pairs and increased mutations at dC:dG base pairs (*48*). Pol η is proposed to function with pol ζ and Rev1 in SHM, as pol ζ can efficiently extend a DNA strand past mismatches inserted by pol η or Rev1 (reviewed in (*49*)). Mutations at dA:dT base pairs occurring in the absence of pol η are likely mediated by pol θ, since deficiency in pol θ results in an overall reduction at both dC:dG and dT:dA residues (*42*). Error free MMR of AID induced lesions occurs at specific genes, such as c-myc, and appears to be specifically excluded from the Ig loci (*10*). Thus, targeting of hypermutation may not only depend on targeting of AID, but also on excluding high fidelity repair. Finally, mutations in variable regions may also be generated through error prone repair of DNA breaks, as up to 6% of mutations contain a duplication or deletion that would necessitate a double stranded break and mutating V regions are accessible to TdT (*50*).

Class switch recombination

Class switch recombination (CSR) occurs in activated B cells in-order to switch the effector function of a particular antibody (reviewed in(*51*)). Mechanistically, CSR occurs by the joining of two different transcribed *IgH* switch region sequences via the introduction of dsDNA breaks, followed by synapsis and covalent joining of the broken DNA ends by non-homologous end joining (NHEJ) or alternative NHEJ (*52*) (Figure 1). This process results in deletion of the intervening DNA and the introduction of a new heavy chain constant region. Mechanistically, CSR is initiated by transcription of at least two different switch region DNAs, producing spliced "sterile" transcripts upstream of two different constant regions (*53*). Conservation of the structure and splicing of sterile transcripts has led some to propose that transcription or perhaps splicing it self, is somehow fundamental for targeting of the CSR process (*53*). Consistent with this, splicing of sterile transcripts is required for CSR(*54*).

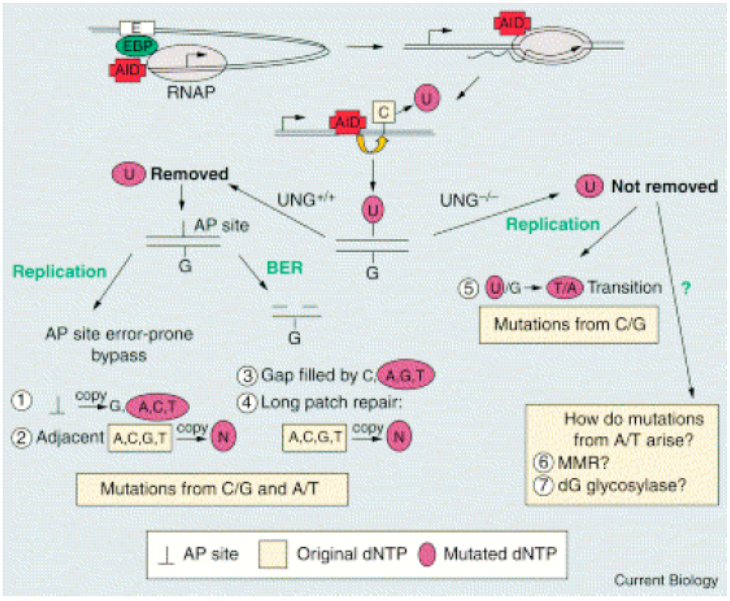
Mammalian switch regions are unique in that they are composed of repetitive sequences that are G rich on the template strand. Both the repetitive nature of switch regions and the composition of a C rich non-template strand are required for optimal CSR. Non- repetitive G rich template sequences, or inverted switch region sequences do not fully recapitulate CSR activity of native switch regions in mouse (*55*). It is proposed that both the repetitive nature of switch regions and the high C content on the non-template strand, allow for the formation of R loop structures upon transcription (*56, 57*). G-loops are composed of a single stranded DNA loop on the non-template strand, and a

DNA/RNA hybrid on the template strand (*58*). This unique structure is proposed to yield access of single stranded cytidine residues to AID. The repetitive nature of switch region DNA gives it other unique structural properties, such as the ability to form DNA hairpin loops (*59*), R loops (*56*), G-loops (*58*) and G-quartet structures (*60*). R-loops are distinct from G-loops in that they do not form co-transcriptionally and instead occur by displacement of a DNA strand by insertion of a complementary RNA sequence (*58*). The observation that Xenopus switch regions are A/T rich, are not predicted to form G-loops, and can functionally substitute for endogenous murine switch regions suggests that high G/C content and G-loops are not the key unique functional characteristics of switch regions (*61*). Instead, the repetitive nature and the presence of RGYW motifs on both DNA strands appear to be the fundamental requirements. This observation lends support for an alternative model for dsDNA break formation at switch regions in which antisense transcription drives deamination on the sense strand. Indeed, antisense transcripts originating from switch regions have been identified (*62*).

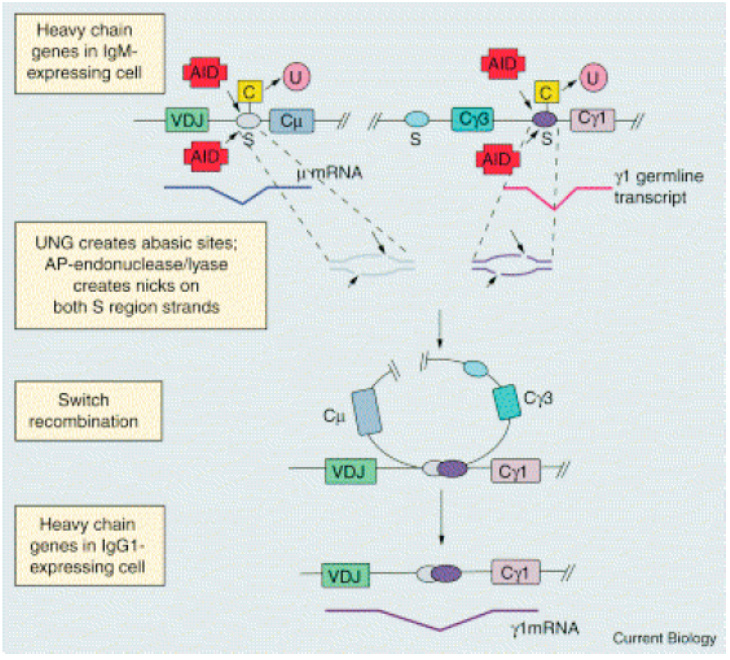
Figure 1. Somatic Hypermutation and Class Switch Recombination

The figure outlines the mechanisms of somatic hypermutation and class switch recombination described in the text. Borrowed from (*20*)

SHM



CSR



DNA Repair during Class Switch Recombination

Generation and repair of AID dependent DNA breaks

It is not clear why dsDNA breaks are found more readily in the switch region than the variable region. There are several possible explanations that are not mutually exclusive. Firstly, switch region DNA may be deaminated to a greater extent on either one or both DNA strands, possibly by antisense transcription(*62*), thus generating more dsDNA breaks. Furthermore, deamination events may occur more frequently and in closer proximity at the switch region due to the high percentage of GC nucleotides and/or a higher rate of transcription at the switch region (*63*). However, it is likely that the repetitive nature of switch regions lends it a structural property that is inherently unstable. Consistent with this notion, multiple "fragile" loci are composed of repetitive DNA sequences (*64-71*).

Transcribed repetitive DNA sequences are readily broken resulting in excised epichromosomal circular DNA molecules (*64-71*). Examples of such loci include telomeres, fragile X, ribosomal DNA, histone DNA, major satellite DNA and DNA containing SINE and LINE elements. A mechanistic explanation for why repetitive DNA sequences are unstable is not well defined. Transcribed repetitive DNA may simply be structurally unstable, more prone to internal recombination and/or more difficult to repair upon breakage. However, breaks at switch region DNA are AID-dependent, suggesting that breaks at switch region DNA occur because they are difficult to repair. For example, a broken switch region DNA strand is likely to fold on itself in a GC base paired hairpin loop that must be extensively processed before repair. This processing may include the

generation of dsDNA breaks since these single stranded breaks may persist for a relatively long period of time.

Both base excision repair and mismatch repair are required for full CSR activity, suggesting that both processes are involved in creating dsDNA breaks (*23, 72-75*) (*76*). Yet, CSR is drastically reduced upon mutation of the base excision repair component UNG, while mutations in mismatch repair components do not have the same effect (35-75% reduction)(*23, 73-75*). Indeed, it has been shown that UNG is required for generation of dsDNA breaks at switch regions(*77*) and UNG deficiency results in hyper IgM(*72*). Interestingly, it was recently demonstrated that mismatch repair acts by introducing A/T nucleotides up to 20nts from a deaminated cytosine on the non-template strand, while base excision repair acts to introduce mutations on both DNA strands (*25*). Thus, mismatch repair is restricted to introducing DNA breaks to the non-template strand and cannot efficiently create dsDNA breaks on its own. However, since base excision repair acts on deaminated cytidine on both DNA strands, it is able to generate double stranded DNA breaks without contribution from other repair pathways. This provides a simple explanation for why base excision repair is absolutely required for CSR while mismatch repair is not. However, this model would exclude the possibility of antisense transcription (unless it occurs sequentially after or before sense transcription), since in the presence of sense and antisense transcription, mismatch repair would be able to act on both DNA strands.

Detection, marking and processing of switch region breaks

DNA breaks at switch regions are primarily repaired by NHEJ in the G1 phase of the cell cycle (Figure 2), although a role for components of homologous recombination (ATR) have also been proposed (*78-80*). Broken DNA ends are initially recognized and bound by the Ku70/80 heterodimer, which forms a ring around the end of the broken DNA molecule (reviewed in (*81, 82*)). The Ku70/80 heterodimer subsequently recruits the signal transducer PI3 kinase DNA-PK to DNA to form the DNA-PKcs holoenzyme. Once bound by DNA-PKcs, the Ku70/80 ring makes one helical turn inward, enabling other repair components to access the broken DNA end for processing, if necessary, by the Artemis enzyme. While Ku is absolutely required for CSR (*83, 84*), DNA-PKcs are required for switching to just a subset of isotypes and Artemis is dispensable for CSR (*85-87*). However, all three components are required for prevention of DNA translocations in activated B cells (*87, 88*). Thus, Ku is required for channeling all switch region breaks to the correct repair pathway, while both Artemis and DNA-PK are required for processing and channeling only a subset of DNA beaks at switch regions to the correct repair pathway. Ku, Artemis and DNA-PK are involved in simple repair and joining of DNA breaks and are can thus mediate "fast rejoining" of dsDNA breaks (reviewed in (*89*)).

Other sensors of DNA damage include the single stranded DNA binding proteins RPA and Mre11-RAD50-Nbs1 (MRN), which recruit and activate the signal transducer PI3 kinases ataxia-telangiectasia (ATM) and ATM and RAD3 related (ATR) respectively (*90*). While ATM functions thorough out the cell cycle and is required for full activation

of ATR(*91*), ATR activity is restricted to the G2 phase of the cell cycle (*92*). Thus, ATM acts both in NHEJ and homologous recombination (HR), while ATR only has a role in HR. Both of these kinases are involved in more complex DNA repair mechanisms than DNA-PK and are thus involved in "slow rejoining" of DNA breaks since these kinases mediate the recruitment of large DNA complexes to broken DNA and phosphorylate a large number of proteins resulting in the inhibition of cell cycle progression(*93, 94*).

ATM directly "marks" DNA breaks at switch regions by phosphorylation of the variant histone H2AX (gamma H2AX) for up to two megabases from the site of the break (*78, 93*). Although H2AX is not required for initial recruitment of repair factors to chromatin, it is required for subsequent assembly of large DNA repair complexes around the DNA lesion (*93*). Both ATM and its activator Nijmegen breakage syndrome1 (Nbs1) (*95*) act directly on DNA and on transmission of DNA damage signals, making them potent suppressors of DNA translocations. An outline of the role of different DNA repair factors in class switch recombination and *c-myc-IgH* translocations is provided in (Figure 2).

Synapsis of broken switch regions

Broken switch regions separated by up to 150 kilobase of DNA sequence are able to find and ligate to one another during class switch recombination. Several factors involved in the proper repair of DNA breaks during CSR, such as ATM, the p53 binding protein (53BP1) and H2AX, are proposed to function in switch region synapsis (*96-98*).

In the absence of any of these repair components, switch regions frequently harbor internal deletions, implying that the broken switch regions cannot synapse and instead ligate to any proximal broken DNA end. Consistent with this notion, deletion of ATM, H2AX and 53BP1 result in severely diminished class switch recombination (*96-98*). Although the precise role of these factors is undefined, 53BP1 may function in synapsis by stabilizing DNA ends. The same may hold true for ATM, as ATM deficient B cells stimulated for CSR harbor a high frequency of chromosomes with inverted and duplicated portions of the DNA separating broken switch regions (unpublished observations). Furthermore, the action of 53BP1 in B cells during CSR appears to occur specifically at lesions at IgH because most chromosomal aberrations in 53BP1$^{-/-}$ B cells include IgH (*88*).

Although the precise role of DNA repair factors in switch region synapsis is undefined, a mechanistic explanation for how this process occurs has been proposed (*51, 99*). Using a technique termed chromosome confirmation capture, it was recently shown that spleenic B cells are primed for switch region synapsis, since the Eu enhancer is in close proximity to the 3'Ealpha enhancers(*99*). AID was required for association of the switch region promoter with the Eu:3'Eapha enhancers, thus stabilizing the interaction. This data provides evidence that that switch region synapsis is supported by distantly separated transcriptional enhancers and AID.

Resolution of DNA breaks during CSR

Resolution of switch region junctions during class switch recombination occurs through binding of Ku and DNA by XRCC4, which in turn binds and stabilizes DNA

ligase four (reviewed in(*81, 82*)). DNA ligase 4 ligates blunt DNA ends or ends with overlapping microhomologies. However, deletion of XRCC4 or DNA ligase four does not completely abolish class switch recombination as approximately twenty percent of B cells lacking these repair proteins are still able to switch (*52*). Thus, an alternative non-homologous end-joining pathway (A-NHEJ) can be utilized to resolve switch region breaks. This alternative repair pathway is error prone, as switch regions are often translocated to other chromosomes in XRCC4 or ligase 4 deficient mice (*52*). Although putative components involved in this alternative repair pathway are postulated to involve ligase1 or 3, an actual molecular pathway for this repair pathway as not been defined. However, junctions repaired by this alternative NHEJ pathway harbor extensive regions of microhomology (*52*). Thus, in the absence of XRCC4 and DNA ligase four, alternative NHEJ utilizes a strict microhomology dependent repair pathway that may also be independent of Ku activity, as *c-myc –IgH* translocations occur in the absence of Ku (*88*).

Furthermore, while the signal transducer DNA-PK only has a role in resolving a subset of switch translocation, the ATM signal transducer has a more profound impact as lack of ATM results in severe deficiency in switching to all switch regions (*98*). Residual switch activity in ATM KO may be mediated by overlapping functions of DNA-PKcs or ATR.

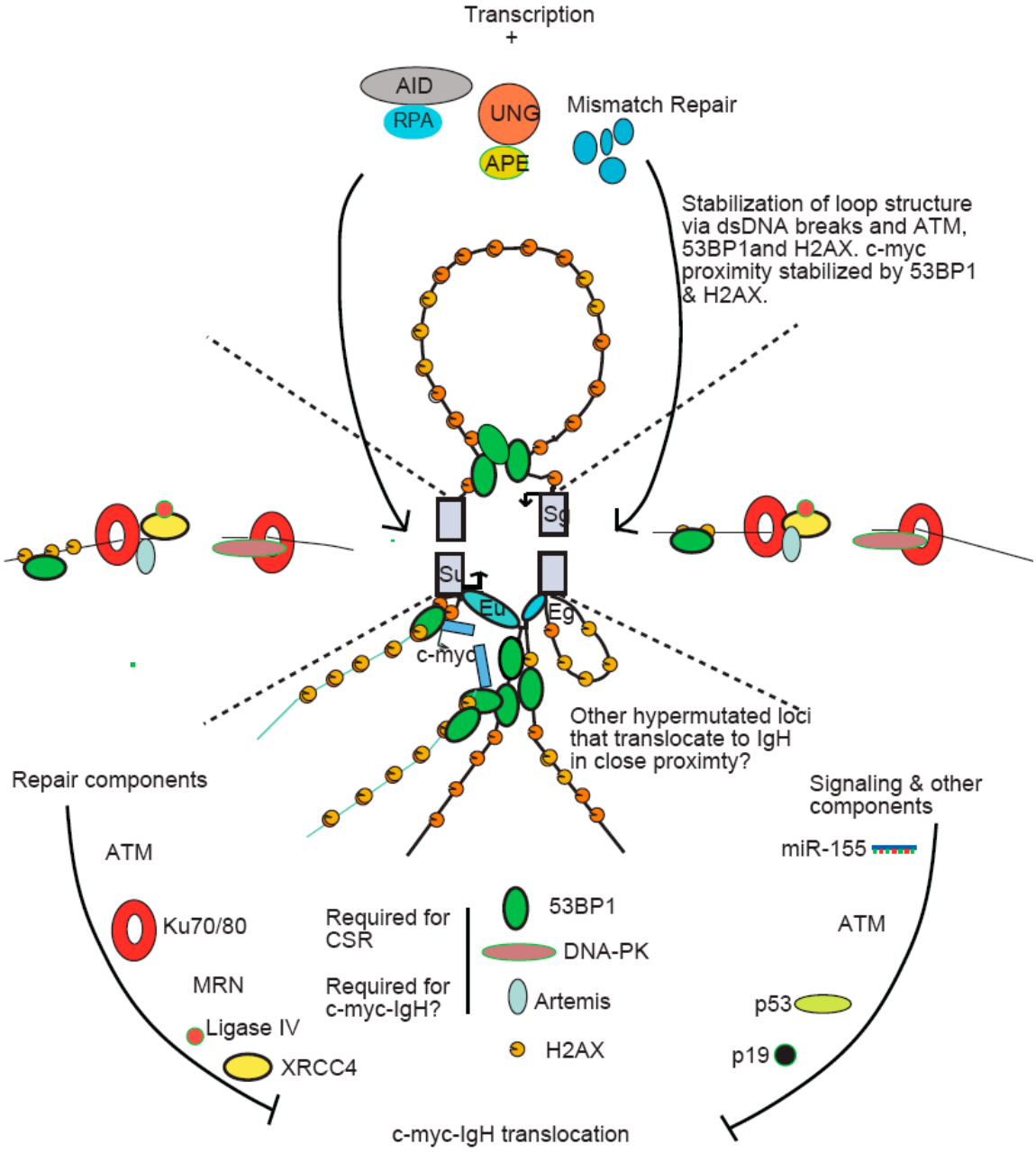
AID induces *c-myc* translocations to the *IgH* switch region by a mechanism that resembles class switching in that formation of the initial lesion requires cytidine deamination and uracil removal from DNA(*88*). However, resolution of the lesion proceeds by distinct pathways for switching and translocation (*88*) (Figure 2). The "core"

repair factors, such as XRCC4, ligase 4, Ku80 and DNA-PKcs are required for resolution of all or some switch reactions, yet none are required for resolution of *c-myc-IgH* translocations (*52, 87, 88*). Furthermore, factors that act in cis to promote switch region synapsis such as the p53 binding protein 53BP1 and variant histone H2AX (H2AX) have no impact on *c-myc* to *IgH* switch translocations despite their effects on genomic stability (*88, 100*). In contrast, factors that only transmit damage signals to the nucleus such as the p53 tumor suppressor do not appear to affect switching but are essential in suppressing *Ig* switch translocations, possibly by promoting the death of cells that over-express *c-myc* (*101, 102*).

Figure 2. DNA repair and protection from DNA translocation during CSR

The figure provides an overview and model of the mechanisms the guide proper repair of switch regions ad prevention of c-myc-IgH translocations. Initially, activation induced deaminase (AID), protein kinase A (PKA), uracil DNA glycosylase (UNG), APE ednonuclease, and components of the mismatch repair machinery initiate DNA breaks at Su and Sδ (Sg in figure) and possibly *c-myc*. The figure diagrams a looping model that brings Su and Sδ into close proximity when interacting with there transcriptional enhancers in a transcriptional factory. Generation of DNA breaks stabilizes the loop structure by recruiting 53BP1. 53BP1 may also be recruited to c-myc upon breakage when c-myc is in the same transcriptional factory as Su and Sδ and may thus stabalize broken c-myc in close proximity to broken IgH. The figure also diagrams a magnification of the DNA breaks at Su and Sδ as indicated by dashed lines. The DNA repair factors and their putative functions with respect to c-myc-IgH translocations is outlined at the bottom of the diagram.

Figure 2


Transcription
+
AID
RPA
UNG
APE
Mismatch Repair
Stabilization of loop structure via dsDNA breaks and ATM, 53BP1and H2AX. c-myc proximity stabilized by 53BP1 & H2AX.
Sg
Su
Eu
Eg
c-myc
Other hypermutated loci that translocate to IgH in close proximty?
Repair components
ATM
Ku70/80
MRN
Ligase IV
XRCC4
Required for CSR
Required for c-myc-IgH?
53BP1
DNA-PK
Artemis
H2AX
Signaling & other components
miR-155
ATM
p53
p19
c-myc-IgH translocation

Mistakes In Antigen Diversification

Oncogenic translocation

Chromosome translocations are products of unresolved double strand DNA breaks (*79, 103-105*) and therefore occur frequently at Ig genes because these loci undergo programmed DNA damage during antigen receptor gene diversification. A hallmark of B-cell non-Hodgkin's lymphomas (NHL) is translocation of a proto-oncogene to the immunoglobulin heavy or light chain locus. The translocation often activates the oncogene by bringing it under the transcriptional control of the immunoglobulin enhancers. The high rate of translocations specifically to the immunoglobulin loci in B-cell NHL implies that aberrant DNA repair from one or more of several gene diversification processes is the cause of these translocations. Depending on the specific location in the Ig locus into which a particular chromosomal translocation falls, one can speculate that its occurrence is the result of a particular gene diversification reaction.

In developing T and B cells, RAG1 and RAG2 proteins produce double strand breaks at recombination signal sequences (RSSs) found adjacent to variable (V), diversity (D) and joining (J) gene segments (*106, 107*); these RSS sequences are found in close proximity or within translocation breakpoints in numerous B cell lymphomas (reviewed in (*108, 109*)). Mice deificient in NHEJ componenets and ATM or p53 develop pro B cell lymphomas with RAG dependent DNA translocations (reviewed in(*110*)). In mature B cells, expression of AID (*5, 6, 111*) leads to deamination of cytidine residues in *Ig V-*J_H and *Ig* switch regions resulting in U:G mismatches which are processed to produce somatic mutations and initiate class switch recombination (reviewed in(*112-114*)). Like

RSS sequences, translocation breakpoints are also found in these regions in numerous different types of B cell lymphomas. These observations suggest that DNA breaks or lesions initiated by RAG or AID frequently undergo aberrant repair, resulting in *IgH* chromosome translocations among other genomic abnormalities.

Double strand breaks are obligate intermediates in the class switch reaction and translocations involving switch regions are frequently found in sporadic Burkitt's lymphoma, diffuse large B cell lymphoma, and multiple myeloma, suggesting that AID is responsible for the lesions that lead to such translocations (*115-122*). Consistent with this idea, AID induced breaks in the switch region activate the cellular DNA damage response machinery (*78*) and AID is essential for *c-myc* translocations to the *IgH* switch region in *IL-6*tg (*88, 123*). In addition, AID appears to target a number of oncogenes that are frequently mutated and often translocated to antibody genes in mature B cell malignancies (*124-128*). In agreement with these observations, deregulated expression of AID is associated with malignancy (*129-132*).

Double strand breaks are not obligate intermediates in somatic hypermutation of the *Ig V-J_H* (*133*). Nevertheless, duplications or deletions that would necessitate a double strand break make up 6% of all the *Ig V-J_H* region associated somatic mutations and DNA breaks can be detected in this region in B cells undergoing mutation (*50, 134-139*). In addition, endemic Burkitt's lymphoma, multiple myeloma, follicular lymphoma, and diffuse large B cell lymphoma contain mutated *V* genes as well as translocations to the *Ig V-J_H* or *Ig V-J_L* regions (reviewed in (*108, 109, 140*)). This suggests that translocations in these malignancies may have occurred in mature B cells which express AID but not

RAG1/2 (*141*), and that some *Ig V-J_H* region associated translocations are byproducts of lesions induced by AID during hypermutation.

Hypermutation of non-Ig genes

Analysis of the mutation load of activated B cells expressing AID has revealed that approximately 20% of all highly expressed genes are hypermutated (*10*). This analysis further revealed that hypermutation is specifically targeted to the *Ig* loci, as they are mutated at a level at least 10-100 fold higher than the second most highly mutated gene (the Bcl6 oncogene). However, AID is not exclusively targeted to the Ig loci, as numerous other genes are also hypermutated by AID. Not surprisingly, many of the genes that translocate to *IgH* during B cell malignancy (Table 1) are hypemutated at a rate significantly above backround (*10*). The translocation breakpoints within these genes always map to the hypermutated portion (*124*). However, some loci never accumulate mutations due to selective high fidelity repair, but are still found to translocate to *IgH* (e.g. c-*myc*). There are also multiple genes that are not established IgH translocation partners that are deaminated by AID. This is likely due to the fact that translocation of these loci to *IgH* does not confer a competitive growth/survival advantage to the cell. Thus, the overall rate of translocation events or other aberrant genomic rearrangements at *IgH* maybe much higher than what is reported.

AID Regulation

Under physiological circumstances AID expression is restricted to activated B cells by

helix-loop-helix (HLH) E protein transcription factors that bind E-box motifs (*142, 143*). Transcriptional activators of AID include E47 (*143*) and Pax5 (*142*) that activate AID by binding E-box motifs in the AID intronic enhancer. These transcriptional activators are directly antagonized by inhibitor of DNA binding 2 and 3 respectively (Id2 &3)(*142, 143*) that prevent activator binding to DNA. Additionally, the transcriptional repressor Blimp1, that promotes B cell differentiation into plasma cells, represses both Pax5 (*144*) and AID (*145*). Calmodulin also represses AID transcription by directly binding E2A upon B cell receptor (BCR) activation (*146*). An additional cis-acting element with a role in activating AID expression was recently identified through conservation analysis (*147*), although the protein factors that bind there have not been identified.

AID transcription is activated in spleenic B cells in vitro is induced by LPS (*111*) and addition of interleukin 4 (IL4), transforming growth factor (TGFβ) or CD40 ligation (*148*) can induce switching to different isotypes. These stimuli activate AID via the JAK/STAT and Nf-κB signaling pathways (*148-150*). Several other stimuli and signaling pathways that induce AID expression in activated B cells, non-activated B cells or even non B cells have recently been identified. These other stimuli are mediated by the following molecules: Toll like receptor (TLR), Ebstein Bar virus latent membrane protein 1 (LMP-1), the B cell activating factor belonging to the TNF family (BAFF) and proliferation inducible ligand (APRIL), CD81, CD19, CD21 and tumor necrosis factor alpha (TNF-alpha) (reviewed in (*151*)). This wide array of mediators used for signaling AID activation presumably allows for AID expression in several different cell types upon certain stimuli. These cell types include activated B cells, peripheral blood naïve B cells

(*152*), pre-B cells (*153*), immature bone marrow B cells (*154, 155*), cholangiocarcinoma(*156*), hepatoma cell lines (*132*), primary human hepatocellular carcinoma(*132*) and gastric epithelial cells(*131*). The functional relevance of AID expression in these cell types is not well defined. However, it has been shown that AID expression can inhibit transformation of primary naïve B cells by virus (*153*) and can result in hypermutation of p53 upon *H.pylori* infection of gastric epithilium (*131*). AID expression is also reported in oocytes(*157*), embryonic germ cells(*157*), embryonic stem cells (*157*) and spermatocytes(*158*) under physiological conditions. However, these observations have yet to be confirmed and there is no evidence of any AID function in these cell types.

AID is also regulated post-translationaly. Its concentration in the nucleus is limited by CRM1 mediated active nuclear export (*159*). Furthermore, post-translational modification by phosphorylation at position S38 is required for interaction between AID and replication protein A for optimal AID activity on transcribed DNA (*19, 160-162*). The interaction with AID is proposed to help target AID to single stranded DNA during transcription. The presence of additional putative phosphorylation sites makes it likely that additional post-transcriptional modifications are present in AID. Additional evidence for other mechanisms of post-transcriptional regulation comes from the observation that C-terminal deletion mutants of AID are active for SHM but not CSR(*163*). This suggests that the C terminus is required for targeting of AID to switch regions and/or directly required for creation of DNA breaks once targeted to switch regions.

There is also evidence that AID is regulated post-transcriptionally. Mouse and human AID harbor a 3'UTR that is ~ 2Kb in length, making it in the top 5% in length in the human genome (Aceview-NCBI). Both the extreme length of the 3'UTR and the extent of it's conservation, suggests it serves a regulatory function. Indeed, microRNA-155 (miR-155), a small non-coding RNA that acts to inhibit gene expression, is highly expressed in germinal center B cells and is predicted to target AID (*164, 165*). Interestingly, expression of both AID and miR-155 are induced by inflammatory conditions, even in non-B cells (*131, 132, 156, 166*). Exposure to LPS or virus (*167*) activates both genes, suggesting that the transcriptional programs of these two genes are functionally linked. This idea supported by the observation that both genes are activated by NF-Kb(*168*). However, microRNA-155 has it's own set of transcriptional activators (*169*). The link between AID and inflammatory conditions and the development of cancer suggested that mir-155 may function as a critical tumor suppressor in these tissues by inhibiting translation of AID. However, two critical experiments were needed to address this question *in vivo*: (1) Is endogenous AID regulated by mir-155, and (2) does mir-155 act as a tumor suppressor?

Table 1. *Ig* frequent translocation partners

The table lists genes that frequently translocate to IgH and outlines there function and whether or not they hypermutated in mouse or human.

***Ig* frequent translocation partners**

Gene	Function	translocates to IgH	translocates to IgK or IgL	Fusion Genes (other translocations)	targete by AID
FAS (CD95?)	pro apoptotic	?	?	?	human-yes
c-myc	oncogene	mouse-yes human- yes	?	?	human-yes in DFLC mouse-yes
Bcl6	oncogene	human-yes mouse-yes (IL6tg)	Yes(4% to kappa, 27% to lambda)	Yes (% 40 of DFLC). 13 partners. **8 of them on (including pim1) on same chromsome,** same arm and same band. Also on the list is RhoH/TTF.	human-yes mouse-yes
RhoH	? Mutated/translocated in human malignancies	human-yes	?	?	human-yes mouse - yes
Pim1	oncogene	yes (human	?	?	mouse-yes
Pax5	? Mutated/translocated in human malignancies	human-yes	?	?	human-yes mouse - yes
Cd83	signal transduction?	?	?	?	mouse-yes
Blk	signal transduction?	?	?	?	mouse-yes
Yes1	viral oncogene in mice	?	?	?	mouse-yes
Mb1	signal transduction?	?	?	?	human-yes
IgB	signal transduction?	mouse-yes	?	?	human-yes mouse - yes
FGFR	overexpressed in multiple myeloma		?	?	human-yes

RNAi & MicroRNA Mediated Posttranscriptional Regulation

RNAi

In 1990, the first indication for the existence of nucleic acid sequence guided gene silencing came from scientists who were introducing additional copies of a gene responsible for the darkening of flower color into the *Petunia* genome (*170, 171*). In addition to creating darker flowers, the insertion of multiple copies of the gene created white flowers or flowers with patches of white mixed with patches of color (variegated). The white and variegated plants had recognized the newly introduced transgenes as foreign and marked them as well as the endogenous homologous gene for silencing- a process that became known as cosuppression. Subsequent experiments showed that ribonucleic acid (RNA) transcribed from the transgenes was the silencing trigger. Andrew Fire and Craig Mello discovered in 1998 that injection of double-stranded RNA (dsRNA) into the nematode worm *Caenorhabditis elegans* caused sequence-specific gene degradation of cytoplasmic messenger RNAs (mRNAs) containing the same sequence as the dsRNA trigger (*172*). This phenomenon was termed RNA interference (RNAi), and was soon related to the cosuppression events described earlier in plants.

Biochemical data from plants and the fruit fly *Drosophila melanogaster* revealed that the true mediators of RNAi were short interfering RNAs (siRNAs) of distinct length (21-28nt) and structure-one of at least a few types of small RNA produced from cleavage of long dsRNA (*173*). RNAi was rapidly developed as a tool to study gene function and was found to naturally occur in all eukaryotes tested, with the exception of budding yeast. However, it was not until the discovery of siRNAs that RNAi could be readily

applied to mammals, since introduction of long dsRNA into the cells is toxic due to induction of the interferon response (*174*). Genetic and biochemical investigations of the mechanisms guiding RNAi in many different organisms revealed conservation of a cellular machinery that cleaves dsRNAs into 21-28nt small siRNAs (*175, 176*).

RNAi was initially viewed as a tool that could be used to specifically silence genes by introduction of long dsRNAs or siRNAs into cells or organisms of interest. By studying the mechanism guiding RNAi, it was determined that cells also produce long dsRNAs endogenously within the cell by either bidirectional transcription as seen for transposable elements, or by transcription of long hairpin RNAs as seen for both repetitive elements and MicroRNA genes (see below). The dsRNAs produced in the cell are also processed into 21-28nt small dsRNAs (*177*). The specific source and cellular environment (including cell type) of a small dsRNA determines whether or not it is chemically modified and/or incorporated into one or more distinct protein complexes that determine the specific effector function of the guiding small RNA. Small RNAs produced within the cell can regulate gene expression at the transcriptional or posttranscriptional level. More specifically, small RNAs have been found to guide mRNA degradation, translational inhibition, translational stimulation, mRNA transport and storage, mRNA destablization, transcriptional activation and transcriptional silencing through the guidance of chemical modifications of DNA and its physically associated proteins (chromatin) (reviewed in(*178*)).

General mechanism

Most of the work investigating the mechanisms of gene regulation by small RNAs processed from dsRNA has utilized *Arabidopsis thaliana* (a small weed), the roundworm *C.elegans, D. melanogaster,* and human cells. Although some species-specific characteristics of small dsRNA mediated gene regulation have evolved, the basic steps are conserved. The core pathway can be divided into initiation and effector steps (Figure 3). Initiation is mediated by the Dicer nuclease that cleaves long dsRNA molecules into 21- to 28-nucleotide (nt) RNA duplexes that contain 2nt 3' hydroxyl overhangs and 5' phosphates (*175, 176, 179, 180*). In RNAi, and in some cases of naturally occurring microRNA mediated gene regulation, the small RNAs are incorporated into a protein complex that contains the Argonaute 2 (Ago2) protein that directly binds the small RNA and is the endonuclease the specifically cleaves target mRNAs (*181*) that contain a minimum of 13 nts of complementarity from the 5' phosphate of the small RNA guide (*182*). MiRNA and siRNAs bound to Ago2 that have extensive complementarity to their target mRNAs are able to guide multiple rounds of sequence-specific cleavage (*183*). In *D. melanogaster*, Dicer processing of very long dsRNA occurs asymmetrically from both ends and determines which strand of a siRNA will guide a small RNA containing effector complex to the complementary target mRNA (*176*). The active silencing complex contains only one strand of the siRNA bound to Ago2 (*184*), allowing the single-stranded siRNA to pair with its complementary sequence in the target mRNA. The targeting effector complex cleaves a target RNA only once, 10 nt from the 5' phosphate of the antisense (noncoding) guide RNA(*180*). Incorporation of siRNAs into the silencing

complex containing Ago2 and removal of one of the RNA strands is a relatively complex and well understood process and will not be discussed further here.

MicroRNAs (miRNAs)

The RNAi machinery has functions beyond protection against invaders. Plants and animals have genes that produce noncoding RNAs that are processed by components of the RNAi pathway to regulate expression of endogenous genes (*177*). MicroRNAs (miRNAs) are a class of ~22 nt non-coding RNAs produced from genes encoding RNAs with a hairpin secondary structure (reviewed in(*178, 185*). The primary products of these non-coding RNA genes are long transcripts termed pri-microRNAs that contain one or more hairpin structures. In the nucleus, these hairpin structures are cleaved at their base by the RNAse III enzyme Drosha to produce ~70nt pre-microRNAs (*186, 187*). Pre-microRNAs are exported out of the nucleus by Exportin5 (*188*) where they are processed by another RNAse III enzyme termed Dicer (*179*) to produce ~22nt duplex RNAs. Duplex RNAs are concomitantly unwound to produce ~22nt single stranded mature microRNAs, predominantly derived from the strand containing the less thermodynamically stable side at its 5' end (*189*). A mature single stranded microRNA is incorporated into protein complexes termed miRNPs that contain one of several different Argonaute proteins with distinct effector functions (*190*).

Most animal miRNAs are partially complementary to sequences in the 3' untranslated regions of their target mRNAs. Complementarity between the mRNA and nucleotides 2 through 8 (what is called the seed) of the miRNA are essential for miRNA

targeting (*164*). This observation has allowed for the bioinformatic prediction of miRNA targets (*164, 165*). Relatively recent reports have revealed that microRNA mediated post-transcriptional regulation can occur through multiple mechanisms that can result in multiple outcomes. MicroRNAs target messenger RNAs (in some cases reversibly(*191*)) in order to inhibit translation and/or destabilize the mRNAs (*178*). It has also been proposed that miRNAs can silence genes by possibly targeting proteases that degrade the nascent polypeptide (*192*). It was recently found that miRNA mediated repression can be reversed in response to extracellular stimuli (*191*). It was also found that miRNAs are used to store and transport RNAs (*193*) and in some cases even stimulate translation in a cell cycle dependent manner(*194*). The details of how a microRNA ends up following a particular path for regulation is unclear. However, it is clear that the cell type and the environment of the cell can play crucial roles in affecting which pathway for regulation a microRNA will take. The mechanism(s) by which miRNAs are postulated to repress gene expression are outlined in Figure 3.

To date, more than 440 microRNAs have been identified in human cells (*195*). However, bioinformatics indicates that humans probably have thousands of miRNA genes, as some miRNAs may be expressed at low levels or in rare tissue or cell types making them difficult to clone. New techniques developed for validating these bioinfomatic predictions have identified new miRNA genes, indicating that there may indeed be thousands of miRNAs to be identified (*196, 197*).

MicroRNA-155

The non-coding RNA BIC was originally identified in an experimental screen to identify proto-oncogenes that might function (in cooperation with c-myc) to induce late stages of progression in Avian Leukosis Virus induced B cell lymphomas (*198*). Upon the discovery of miRNAs, it became clear that BIC is processed to produce microRNA 155 (miR-155), a vertebrate-specific miRNA that is predominantly expressed in activated leukocytes and germinal center B cells (*198-202*), thymus, and to a lesser extent in other immune cell types and tissues. *Bic* deletion (miR-155$^{-/-}$ mice) results in deregulated expression of hundreds of mRNAs, some of which are direct targets of miR-155, resulting in abnormalities in the germinal center reaction and antibody responses *in vivo* (*201, 203, 204*).

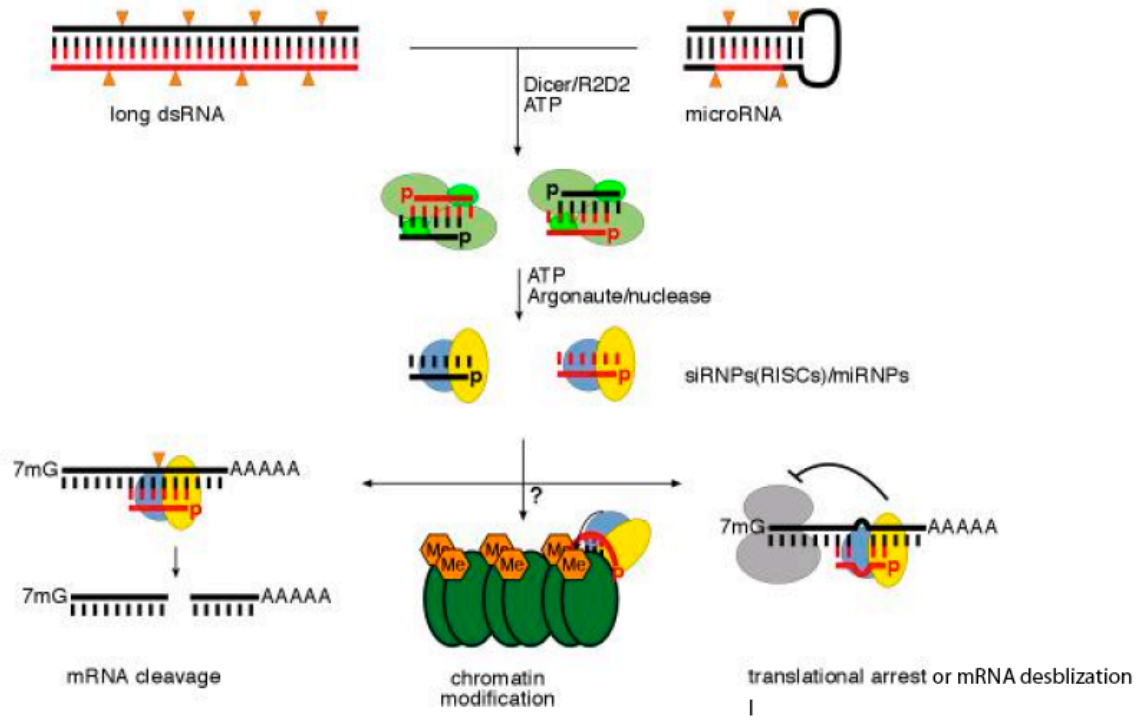
It is now established that miR-155 is an oncogenic miRNA that is deregulated in a number of different cancers, most of which are of B cell origin (*168, 195, 198, 205-219*). Over-expression of miR-155 in transgenic mice leads to pre-leukemic pre-B cell proliferation in bone marrow and spleen, followed by lymphoblastic leukemia/lymphoma (*207*). Conversely, the mature form of miR-155 is absent in primary cases of Burkitt's lymphoma (*168, 195, 214*) due to specific defects in BIC processing. MiR-155 is predicted to target a host of genes that are highly expressed in germinal center B cells, many of which are involved in the generation of B cell lymphomas (*164*). A subset of these predicted targets including AID, Bcl6, and likely TP53INP1are involved in the generation and/or prevention of *c-myc to IgH* translocations; the primary transforming event in Burkitts lymphoma. Although the precise role of miR-155 in promoting

lymphomagenesis has not been determined, it is possible that AID dependent *c-myc-IgH* translocations in Burkitts lymphoma result from high levels of AID expression in the absence of miR-155. Thus, ablation of miR-155 may result in high frequencies of AID dependent translocations. Of special interest with regards to links between AID and miR-155, cells exposed to EBV or LPS co-induce both AID and miR-155 and both genes are activated by NF-Kb(*132, 168, 200*). This suggests that activation of AID and miR-155 is mechanistically linked and further implies that miR-155 plays a critical role in AID down-regulation.

To better understand the pathological roles of AID, we undertook an analysis of the role of AID in translocations to regions prone to somatic hyperutation, and how this role is regulated by MicroRNA-155.

Figure 3. RNAi

The figure outlines the mechanisms by which small RNAs may regulate gene expression. Small RNAs can repress gene expression by targeting mRNAs with 14nts of complementarity from the 5'end for cleavage by Argonaute 2. Alternatively, different Argonaute proteins of both he Piwi and Ago subfamilys may target chromatin/DNA modifications and translation repression or mRNA destablization.



Experimental Approach

Chromosome translocations between oncogenes and the immunoglobulin (Ig) region spanning the variable (*V*), diversity (*D*) and joining (*J*) genes (*Ig V-J_H region*) are found in a number of mature B cell lymphomas in humans and mice. The breakpoints are frequently adjacent to the recombination signal sequences (RSSs) targeted by recombinase activating genes 1 and 2 (*RAG*1/2) during antigen receptor assembly in pre-B cells, suggesting that these translocations might be the result of aberrant V(D)J recombination. However, in mature B cells undergoing AID dependent somatic hypermutation (SHM), duplications or deletions that would necessitate a double strand break make up 6% of all the *Ig V-J_H* region associated somatic mutations (*50, 135*). Furthermore, DNA breaks can be detected at this locus in B cells undergoing SHM (*136*). To determine whether SHM might induce *c-myc to Ig V-J_H* translocations, we searched for such events in both *IL6* transgenic (*IL6* tg) and *$AID^{-/-}$ IL6* tg mice. IL6 attenuates apoptosis and promotes proliferation and differentiation of late stage B cells. Thus, *IL6* tg mice develop hyperplastic lymph nodes that contain switched plasmacytes that harbor c-myc-IgH translocations, a portion of which express GL7 and CD138 (*12, 220, 221*). A majority of the translocations involve the switch region, yet a small portion involve the V-Jh region. Here we find that *AID* is required for *c-myc* to *Ig V-J_H* translocation, which is the primary transforming event in endemic Burkitts lymphoma.

The idea that miRNAs regulate mRNA translation and stability has been tested by over-expression and deletion experiments which result in global effects on vast networks of genes that share microRNA target sequences. Such studies do not take into account off

target or indirect effects of miRNA overexpression and deletion. Thus, the physiologic effects of targeting an individual miRNA to a mRNA has not been tested directly *in vivo*. To this end we created a knock-in mouse that carries a mutation in the putative mir-155 target site in the 3'UTR of activation induced cytidine deaminase (*AID*^{155} mice), an enzyme required for immunoglobulin gene diversification in B lymphocytes, but which also promotes chromosomal translocations {Harris, 2002 #8; Muramatsu, 2000 #4; Muramatsu, 1999 #5; Ramiro, 2006 #7; Ramiro, 2004 #6; Revy, 2000 #3. We also analyzed the miR-155 knockout mouse to determine to what extent does regulation of AID by miR-155 contribute to the overall microRNA phenotype. Using these two approaches we were able to determine that miR-155 downregulates AID expression by 3 fold by destabilizing the transcript, resulting in a 3 to 6 fold increase in translocation frequency and a small increase in CSR and SHM. The levels of translocation suppressed by AID represent just 15 to 30 % of the total number of translocations present in the miR-155$^{-/-}$. Thus, miR-155 regulation of additional mRNA targets aids in maintaining genomic integrity.

MATERIALS AND METHODS

Animals

Balb/c IL6tg and Balb/c IL6tg $AID^{-/-}$ mice were created as described previously {Ramiro, 2004 #6}. Onset of development of hyperplastic lymphnodes was monitored by palpitation of mice until the development of large lymph nodes, taking up to 9 month. Sick mice were sacrificed and hyperplastic lymph nodes were removed for DNA preparation.

Mutation of the putative miR-155 binding site in *AID*

AID nucleotides AGCATTAA, located in the *AID* 3' UTR, 468bp downstream of the stop codon, were replaced with GCGCGCGC by gene targeting (Figure 9). The long arm of the targeting vector was 6.9kb long with 3' within the intron between *AID* exons 4 and 5 (Figure 10). The short arm was a 1.5 kb fragment extending downstream of the 3'UTR. A LoxP-flanked neomycin-resistance gene was used for positive selection, and a diphtheria toxin gene for negative selection (*222*). The targeting construct was linearized and transfected into C57Bl/6 embryonic stem cells. ES cell clones were screened and seven positive clones were injected into C57Bl/6 blastocysts, and one produced chimeric mice that transmitted the mutation. The genotype was confirmed by amplifying the mutation with a primer external to the targeting construct and proximal to the end of the short arm. The resulting amplification product was verified by sequencing and digesting with AsceI resulting in a digested WT allele amplification product and non digested AID^{155} product. The identity of the resulting AID transcript was confirmed by reverse

transcribing total RNA from activated B cells followed by amplification of the AID transcript from the 5'UTR to the 3'UTR at the end of the short arm followed by sequencing. To produce $AID^{155/-}$ mice, heterozygous $AID^{155/+}$ mice were crossed to $AID^{-/-}$ C57Bl/6 mice and littermate $AID^{+/-}$ were used as controls. *bic*/miR-$155^{-/-}$ were previously described (*201*). All mice were maintained under specific pathogen free conditions and experiments performed under Rockefeller University IACUC approved protocols.

DNA preparation and PCR

Hyperplastic lymph nodes from individual male and female transgenic for human IL6 (IL6tg) (14) & AID-/-IL6tg (12) mice were combined into 4 pools. Total DNA was prepared from $2x10^7$ cells for each pool. $0.5x10^6$ cells from each of the9 four pools for 12 different mice was assayed for derivative 12 translocations by PCR using primer set #1 and $2.5X10^6$ cells using primer set #2 (see below). For derivative 15 translocations, $0.5x10^6$ cells from each of the four pools from 12 different mice was amplified using set #3. (see below).

For derivative 12 translocations from *c-myc* to the *IgH* variable region we performed nested PCR (Long Expand PCR system, Roche) using the following primers: Primer set #1: first round with 5-gcaatgactgaagactcagtccctcttaag-3 (*IgH*) and acttagccctgcagacgcccaggaatcgcc (*c-myc*) followed by nested PCR with taccatttgcggtgcctggtttcggagagg (*IgH*) and ttggcttcagaggctgagggaggcgactg (*c-myc*); Primer set #2: first round with gtgccccactccactctttgtccctatgc (*IgH*) and

gaaataaaaggggaggggtgtcaaataataagag (*c-myc*) followed by nested PCR with atcatccagggactccaccaacaccatcac (*IgH*) and cctcccttctacactctaaaccgcgacgccac (*c-myc*). 500ng of DNA (10^5 cells) were amplified in each 20 μl first round PCR reaction ;1 μl of the first reaction was template for the nested PCR reaction. PCR conditions for the first round were 94C for the first 2min followed by 10 cycles of 94°C, 30sec; 61°C, 30sec; 68°C, 7min followed by 19 cycles of 94°C, 30sec; 61°C, 30sec; 68°C, 7min + 20 sec per cycle. The conditions for the nested PCR was as follows: 94°C, 15sec; 61°C, 30sec; 68°C 4min followed by 15 cycles of 94°C, 30sec; 61°C, 30sec; 68°C, 4min + 20 sec/cycle. The combined total number of cycles of amplification was 45.

Conditions for amplification of derivative 15 translocations from *c-myc* to the *IgH* variable region were the same as those for derivative 12 with the exception of a 2 min extension time in the first round PCR and a 1 min extension time in the nested PCR. Amplification of derivative 15 translocations was done using the following primers (primer set #3): GTTGAGACATGGGTCTGGGTCAGGGAC (*IgH*) and ATCAGCGGCCGCAACCCTCGCCGCCGC (*c-myc*) followed by a nested PCR reaction with CTCTGCCTGCTGGTCTGTGGTGACATTAG (*IgH*) and GAAGGCTGGATTTCCTTTGGGCGTTGG (*c-myc*).

For amplification of derivative 12 *c-myc* to switch region translocations, primers previously described (*12*) were used following the PCR conditions described above for primer set #1.

Lymphocyte cultures and translocation assays

Resting B lymphocytes were isolated from the spleen using CD43 microbeads (Miltenyi Biotech), cultured in RPMI supplemented with L-glutamine, sodium pyruvate, 50μM 2-Mercaptoethanol and 10% FBS (GIBCO-BRL) and where indicated cells were labeled with CFDA-SE (5 μM, Molecular Probes). The B cells were stimulated with LPS (25 μg/ml) and IL-4 (5 ng/ml, Sigma) alone or together with IL-5 (15ng/ml, Pharmingen) and BAFF (10ng/ml, R&D systems) for production of plasmablasts (*223*). For the plasma cell experiment, CD138+ B cells were sorted and FACS done on day 5. Transloaction assays were exactly as previously described (*11, 88*). Briefly, PCR of WT, $AID^{-/-}$, miR-$155^{-/-}$, $AID^{155/-}$or $AID^{155/155}$ was performed with total DNA prepared from day 3 or 4 LPS + IL4 cultures. Data shown for $AID^{155/-}$ and $AID^{155/155}$ were from day 4 cultures without dead cell removal prior to DNA preparation. Data for mir-$155^{-/-}$ mice were from day 3 cultures with dead cell removal prior to DNA preparation. Approximate cell number for each sample was determined by DNA quantitation on an ethidium bromide stained agarose gel (500 ng DNA = ~100,000 cells). Subsequent amplification of *c-myc* to *IgH* *s*witch region translocations was done as previously described. Amplification products were verified with Southern blots by probing for *c-myc* and *IgH* as previously described (*12*). Bands that probe for both *c-myc* and *IgH* represent *c-myc* to *IgH* chromosomal translocations.

Southern blot analysis

PCR products from primer set #1 were separated on 0.8% agarose gels and denatured for

15min in 0.4M NaOH before transfer to nylon membranes and probing with ^{32}P radiolabeled primers. The *IgH* probe sequence is GGTGGCAGAAGCCACAACCATACATTCCCA and the *c-myc* probe sequence is gcgcctcggctcttagcagactgtat.

Sequencing

The PCR reactions were separated on 0.8% agarose gels, the bands were gel extracted using the Qiagen gel extraction kit according to the manufacturer's instructions and sent for direct sequencing. Translocations were sequenced with the nested PCR primers. If the translocation breakpoint was not identified with the first round of sequencing, we designed new primers for sequencing to walk along the translocation until we reached the breakpoint.

Sequence analysis

For *c-myc* translocation to *Ig* variable region, PCR products were sent directly for sequencing, without cloning, so that we could discount the error rate of the polymerase in our analysis. Mutations that arose during early PCR cycles could be identified by chromatogram analysis by the presence of more than one base signal for the same nucleotide. Such nucleotides were discounted from our analysis. All sequenced translocations were aligned using both SeqMan from DNA Star and the Codon Code Aligner software (CodonCode Corp). Overlapping traces from all the translocation sequences allowed for distinction between real mutations and single nucleotide

polymorphisms (SNPs). In the case of *c-myc*, we also amplified and sequenced the first intron from the an *IL6* tg mouse to verify the *c-myc* mutations identified were not SNPs.

For mutation analysis of *AID*^{155} mice, genomic DNA from sorted CD19^{+}Fas^{+}GL7^{+} germinal center cells of NP-KLH immunized mice was PCR-amplified in 50ul with PfuTurbo (Stratagene) for 30 cycles from 10-100,000 sorted cell equivalents in four independent reactions that were pooled for cloning experiments. For 5' Sμ the primers and PCR conditions have been described (*97, 224*). The JH4 intron was amplified with (5'GGAATTCGCCTGACATCTGAGGACTCTGC) and (5'CTGGACTTTCGGTTTGGTG) 14 cycles 94°C (30 s), 55°C (30 s), 72°C (90 s) and then (5'GGTCAAGGAACCTCAGTCA) and (5'TCTCTAGACAGCAACTAC) 21 cycles 94°C (30 s), 55°C (30 s), 72°C (30 s). Statistical significance was determined by a two-tailed t test assuming unequal variance. Bcl6, primers were p369 5'-CTTTCTTGGTTGGAGTCGAGG-3' and p370 5'-CGGGCTTGAGGTCATTTCTC-3', as previously described in (*225*). PCR reactions were performed in triplicates, the products pooled, and bands at the expected size gel-extracted and cloned with TOPO-TA (Invitrogen). Bacterial colonies were sequenced by Biotic Solutions, NY, and analyzed with CodonCode Aligner software (CodonCode Coporation, MA). Only good quality sequence was considered, as determined by inspection of the chromatograms.

Flow cytometry

Single cell suspensions from bone marrow, spleen or lymph node were stained with streptavidin FITC, PE, APC or biotin-conjugated monoclonal antibodies anti-CD43, anti-

IgM, anti-B220, anti-CD95, anti-GL7, anti-CD19, anti-FAS or anti-IgG1 (BD Biosciences). GC cells were CD19+, Fas+, GL-7+ lymph node cells sorted 14 days post immunization. Data was collected with a FACSCalibur™ and analysed using CellQuest™ and FloJo software. Cell sorting was on a FACSAria™ and FACSVantage™.

Immunizations and ELISA

Age- and sex-matched 8- to 12-week-old mice were immunized by footpad injection with 50 μg of alum precipitated NP_{21}-CGG (both from Biosearch Technologies). To measure serum antibody levels we used goat anti-mouse Ig (H+L) for capture and HRP-conjugated goat anti-mouse isotype-specific antibodies (Southern Biotechnology) for detection. Values were calculated by comparison with mouse immunoglobulin standards (Southern Biotechnology). Serial dilutions were performed for each sample and readings were taken within the linear range for each sample, and adjusted for dilution. Results reflect relative absorbance for each sample compared with the standard control. All plates were developed using a Peroxidase Substrate Kit (Bio-Rad) and absorbance was measured at 415 nm.

Western Blotting

Anti-AID antibody was an affinity purified polyclonal against the carboxyl terminus (EVDDLRDAFRMLGF) of AID and was previously described (*159*). For western blot assays cells were lysed in 20 mM Tris pH 8.0, 200mM NaCl, 1 % NP-40, 0.5%

Deoxycholate, 0.1% SDS, 1mM DTT, 0.5mM EDTA, 1mM PMSF, Protease inhibitor cocktail (Sigma). 50 μg of protein was detected by western blot using the anti-AID antibody or anti-tubulin antibody (Abcam). Band densities were quantified using ImageJ software and relative AID level is a comparison of AID/tubulin ratios within each gel.

Quantitative PCR

Real Time Quantitative RT-PCR (qRT-PCR) Analysis. Total RNA was isolated from sorted or LPS and IL4 activated cells using Trizol reagent (Life Technologies) according to the manufacturer's instructions. The first-strand cDNA synthesis was performed with 200 ng of total RNA primed with random primers using the RT reaction protocol provided by the manufacturer (Invitrogen). qPCR was performed with Brilliant SYBR Green QPCR master mix (Stratagene) containing 500 nM primers using standard amplification procedure. All samples were analyzed in triplicate, normalized to GAPDH levels, and the result expressed as fold induction compared to WT day 4 control. Primers for AID were: forward 5'gaaagtcacgctggagaccg3' and reverse 5'tctcatgccgtcgcttgg3' and primers for GAPDH were: forward 5'TGAAGCAGGCATCTGAGGG3' and reverse 5'CGAAGGTGGAAGAGTGGGAG3'. Actinomycin D (Sigma) was used at 10 ug/ml.

Calculation of translocation *P* values and AID RNA half-life

P value was calculated using a 2-Tail Fisher's exact test with a program available at http://www.matforsk.no/ola/fisher.htm. Comparison between genotypes of the number of translocations per number of PCR reactions was used to do the calculations. To determine the AID RNA half-life we used an exponential regression model of the data generated by Excel. That exponential decay models the data correctly is supported by high r^2 values

(.93 for 155/- and .94 for AID +/-) in the regression and the finding that the log plot of the normalized data is linear. The half life was determined from the slope of the resulting lines.

RESULTS

PART I: AID DEPENDENT Ig-V-J_H REGION TRANSLOCATIONS

AID accelerates lymph node hyperplasia in interleukin 6 transgenic (*IL6tg*) mice

IL6 tg mice develop hyperplastic lymph nodes that contain large numbers of class switched plasmacytes, a portion of which express GL7 and CD138 (*12, 220, 221*). Plasmacytosis is believed to develop in these mice because IL6 attenuates apoptosis and promotes proliferation and differentiation of late stage B cells allowing for the accumulation of translocations between *IgH* and *c-myc* (*220*). Although a majority of the *c-myc* translocation breakpoints are at the *IgH* switch region, a small fraction occur in the region spanning the variable (V), diversity (D) and joining (J) gene segments (V-J_H region) (*221*). Translocations of *c-myc* to the V-J_H region (*c-myc*-Ig- V-J_H translocations) resemble the translocations found in endemic Burkitt's lymphoma (reviewed in (*108, 109*)). To determine whether AID is required for translocations between the Ig V-J_H region and *c-myc* we generated *AID* deficient *IL-6* tg mice (*AID*$^{-/-}$*IL-6* tg) by breeding (*12*). *AID*$^{-/-}$*IL6* tg mice developed lymph node hyperplasia and plasmacytosis with a slightly delayed onset compared to *IL6* tg mice and there was no detectable class switching in the *AID*$^{-/-}$*IL6tg* mice ((*12*) Figure 4A and B).

c-myc to Ig-V-J_H translocations are AID dependent, reciprocal and less frequent than those to the switch region

To document translocations between the V-J_H region and *c-myc*, we developed PCR assays for these events and examined cells from hyperplastic lymph nodes from *IL6* tg and *AID*$^{-/-}$ *IL6* tg mice (Figure 5A). Southern blotting and DNA sequencing were used to verify candidate translocations (Figure 5 and 6). Assaying four different lymph node pools per mouse for derivative 12 and derivative 15 translocations (Figure 5A), we identified 37 unique translocations in 14 *IL6* tg mice, but none in 12 *AID*$^{-/-}$*IL6* tg mice (p=.0025) (Figure 5C). Derivative 12 and 15 translocations were similar in both the number of translocations identified (21 and 16 respectively) as well as their breakpoint distribution along the chromosome (Figure 6), providing strong evidence that variable region translocations are reciprocal. As expected, *c-myc* to *IgH* switch region translocations were far more frequent (Figure 5B, (*12, 221*)). We conclude that translocations between *c-myc* and the Ig V-J_H region are AID dependent in *IL6* tg mice.

Characterization of *c-myc* – Ig-V-J_H translocation breakpoints

To gain further insight into the etiology of the identified translocations, we analyzed and mapped the breakpoints. Sequence analysis revealed that the majority of the breakpoints were in or around J_H segments, as commonly seen for oncogenic translocations in B cell non-Hodgkin's lymphoma (reviewed in (*108, 109*)) (Figure 6). Junction sequences resembled those previously characterized for *c-myc-IgH* switch region translocations from *IL6* tg mice (*12*) in that they involved either blunt ends or 1-3 nucleotide stretches of

micro-homology or nucleotide insertions (Table 2), implying that non- homologous end joining (NHEJ) resolves these breaks. Finally, the translocation breakpoints in *c-myc* were all in the first intron, which differs from those to the *IgH* switch region in *IL6 tg* mice, where a majority of the breakpoints are in the first exon of *c-myc* (Figure 6 (*12*)). Interestingly, *c-myc* harbors a transcriptional attenuation site at the end of the first exon, suggesting that translocation of *c-myc* to *IgH* variable vs. switch regions may be influenced by the transcriptional states of *c-myc*.

c-myc to Ig-V-J_H translocations occurs after or during somatic hypermutation

Although there was no correlation between the position of the breakpoints and the RGYW motifs which are the preferred targets of AID, the translocated *Ig V-J_H* genes were somatically mutated at a frequency of ~0.6X10^{-3} mutations per base pair (Figure 7). This rate of mutation is similar to that reported for *Ig V-J_H* genes B cells undergoing hypermutation (*226*). Derivative 12 and 15 *IgH* sequences had a similar frequency of hypermutation, suggesting that SHM must have occurred before or during translocation. Interestingly, the overall position of the mutations mirrored the positions of the translocation breakpoints (compare Figures 6 and 7), supporting the idea that regions prone to SHM are susceptible to translocations. Excluding nucleotide insertions at the breakpoint, the rate of mutation within 20bps of either side of the translocation breakpoints was almost ten fold higher (5.4X10^{-3}) than that seen overall for *IgH* in our translocations. This implies that *c-myc/IgH* translocations are resolved through error prone repair or that translocations breakpoints occur at or immediately adjacent to

hypermutated sequences. Discounting mutations within 20 bp of the breakpoint, we identified three mutations in *c-myc*, resulting in a mutation rate that is above background (~0.2 $X10^{-3}$), but approximately three-fold lower than the rate observed for *IgH* (Figure 7). These results suggest that *c-myc* translocations to the *Ig V-J_H* region in *IL6* tg mice occur during or after *Ig V-J_H* region somatic hypermutation (*126, 226, 227*).

Figure 4. Characterization of IL6tg and AID$^{-/-}$ IL6tg mice.

A. Flow cytometry analysis of cells from hyperplastic lymph nodes from IL6tg and AID$^{-/-}$IL6tg mice. Numbers indicate percentages of cells in a given quadrant. **B.** AID accelerates the development of disease in IL6tg mice. IL6tg mice and AID$^{-/-}$IL6tg mice were euthanized when they developed enlarged lymph nodes. The average time of euthanasia for IL6tg was 5.5 months, n=8 and 9.2 months for AID$^{-/-}$IL6tg mice, n=8 (p=.0001476 using a two tailed students T test assuming unequal variance). Each point represents one mouse and the black bars indicate the average time of euthanasia.

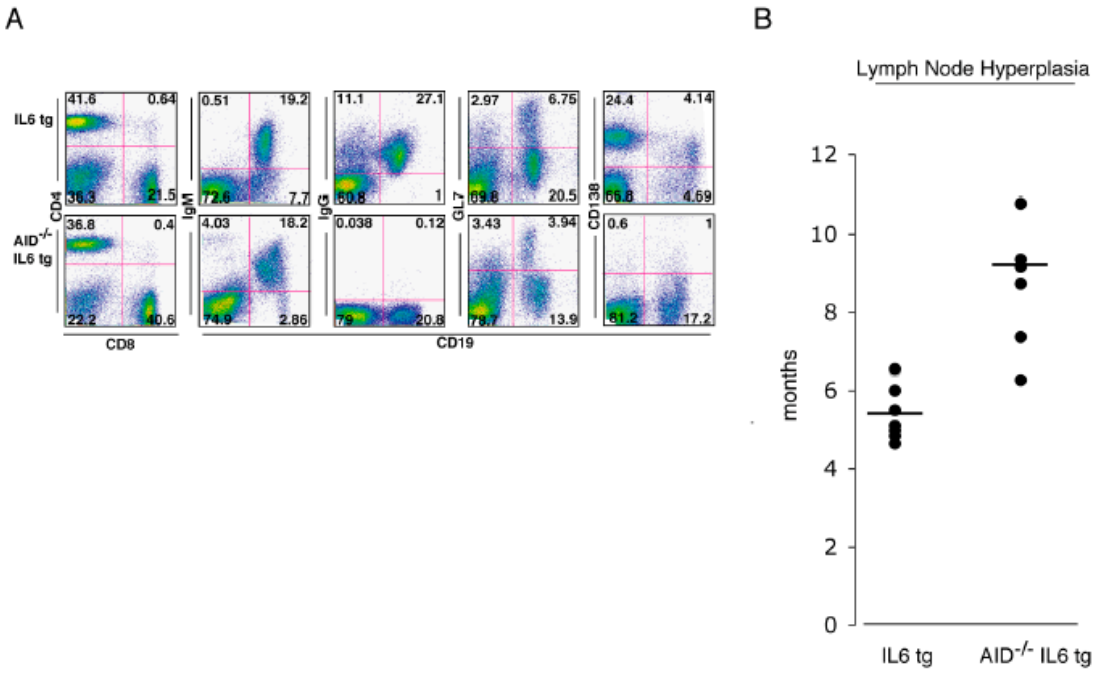


Figure 5. C-myc to Ig V-J_H region.

A. Diagram of translocation assay for detecting derivative 12 & 15 chromosomal translocations from c-myc to the Ig V-J_H region.

Primer set #1 is indicated in the diagram for derivative 12 (see material and methods). The circles at the end of each chromosome represents the centromere. **B.** Comparison of c-myc translocations to the IgH V-J_H vs. switch region. Representative PCRs from 3 IL6tg and 1 $AID^{-/-}$IL6tg mouse. Primer set #1 was used to detect translocations to the variable region. Each lane represents amplification products from 10^5 cells. The EtBr stained gels at the top were probed for c-myc, stripped and then probed for IgH to verify translocations. **C.** Graph shows the number of unique c-myc to Ig V-J_H region translocations identified in 14 IL6tg mice. No translocations were identified in the same analysis of 12 $AID^{-/-}$IL6tg mice ($p=.00284$ using a two tailed students T test assuming unequal variance.).

Figure 5

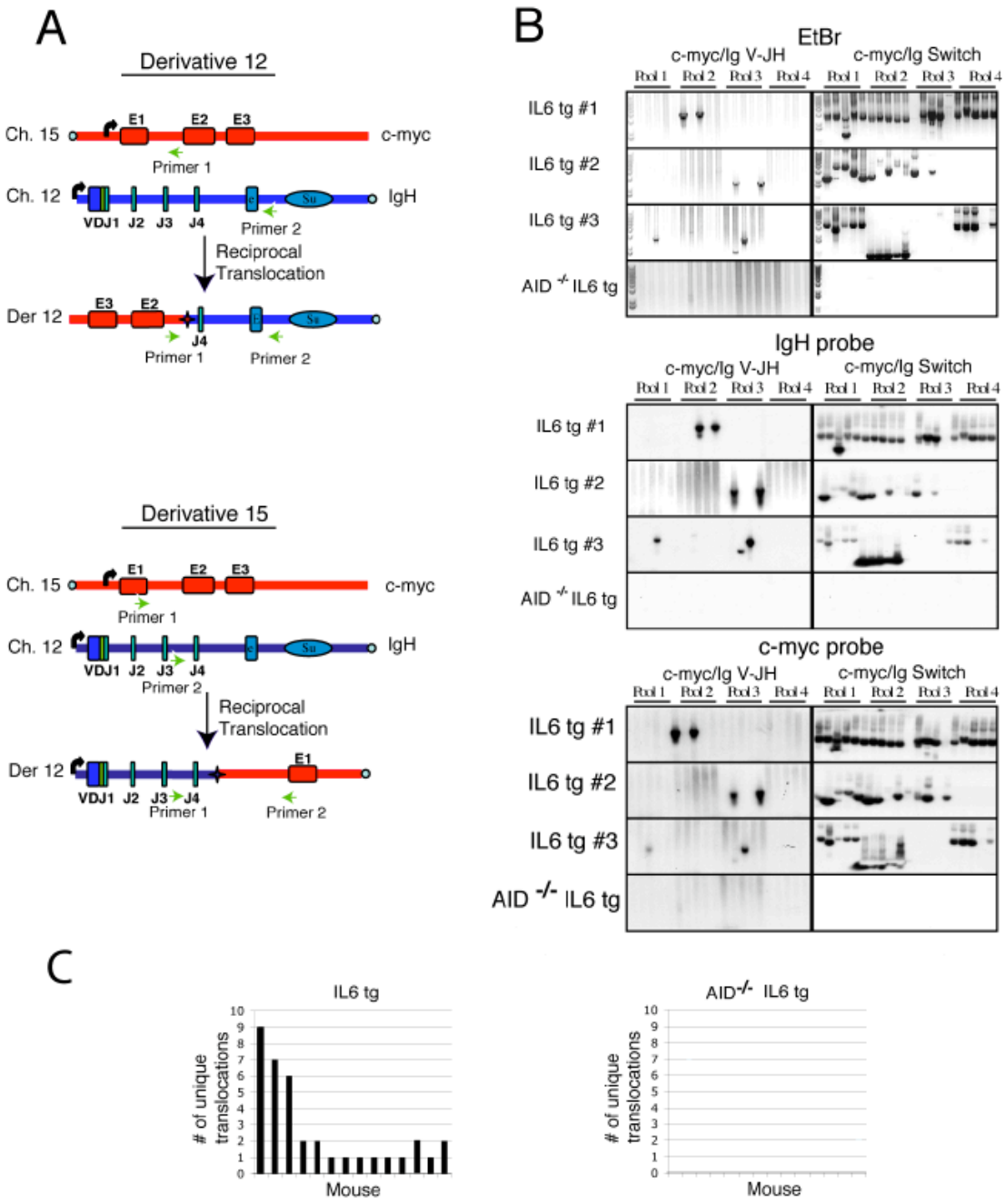

A
Derivative 12
Ch. 15
E1
E2
E3
c-myc
Primer 1
Ch. 12
VDJ1
J2
J3
J4
Primer 2
IgH
Reciprocal Translocation
Der 12
Primer 1
Primer 2
Derivative 15
B
EtBr
c-myc/Ig V-JH
c-myc/Ig Switch
Rxn 1
Rxn 2
Rxn 3
Rxn 4
IL6 tg #1
IL6 tg #2
IL6 tg #3
AID -/- IL6 tg
IgH probe
c-myc probe
C
IL6 tg
AID-/- IL6 tg
# of unique translocations
Mouse

Figure 6. Translocation breakpoints.

Diagram of translocation breakpoints from c-myc to the Ig V-J_H region. Only primer set #1 is indicated for derivative 12 (see material methods). The translocations are annotated as the mouse number followed by the translocation number for that mouse. For example "2 #2" is mouse 2 translocation number two amplified from that mouse.

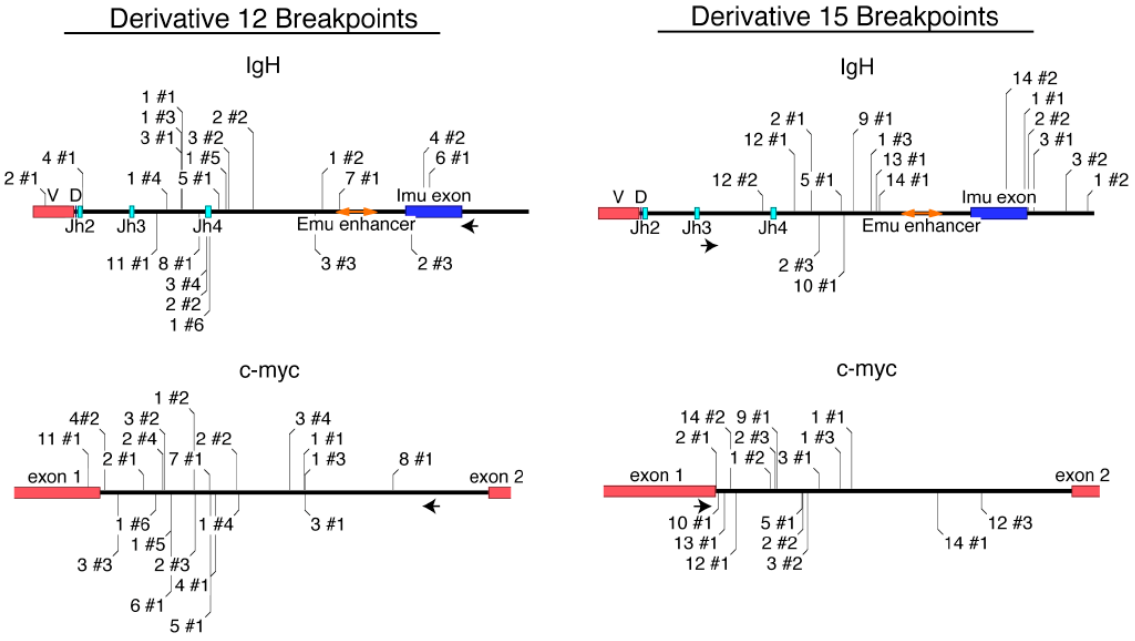


Figure 7. Somatic mutations in translocated IgH.

Analysis of 32,654 bp of IgH sequence from c-myc/IgH translocations identified 20 different mutations (mutation frequency = .61 *10^-3), excluding the two mutations found within 4nts of breakpoints. Excluding the three mutations within 3nts of the breakpoint, analysis of 16,960 bp of c-myc identified 3 different mutations (mutation frequency = .18 * 10^-3). The overall mutation rate within 20nts of a breakpoint on either side of the translocation was 5.4 * 10^-3. Asterics indicate mutations within 20nts of a breakpoint.

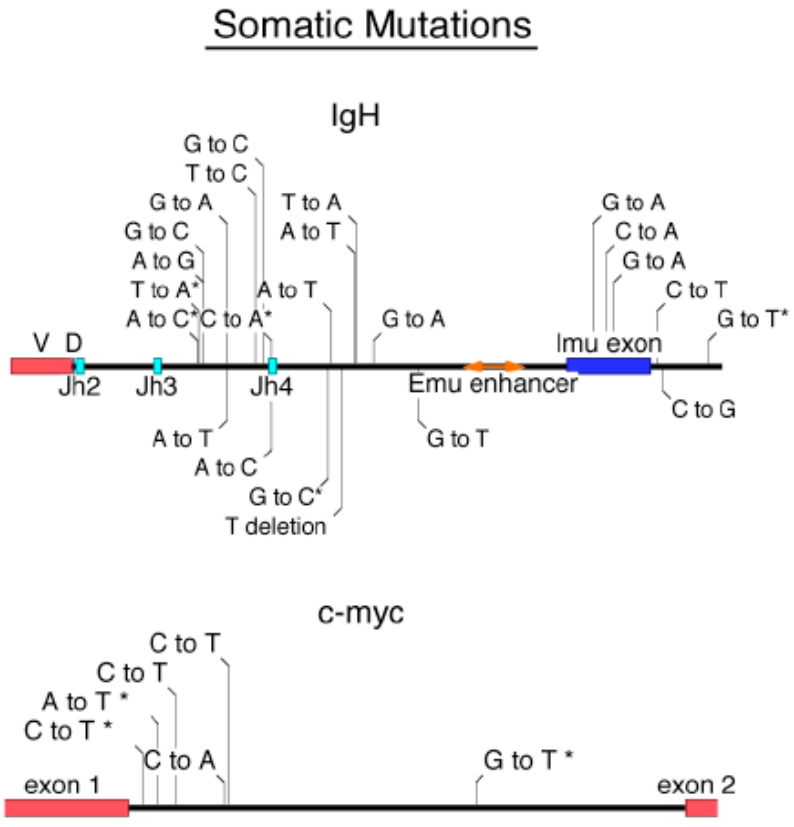


Table 2. Translocation junctions between *c-myc* and Ig-V-J_H

Sequences of all of the translocation breakpoints for derivative 12 and 15. Junction sequences are those that cannot be definitively assigned to c-myc or IgH. Lower case letters indicate mutations, and insertions are in italics. The translocations are annotated as the mouse number followed by the translocation number for that mouse. For example "2 #2" is mouse 2 translocation number two amplified from that mouse. "jnc" stands for junction.

Derivative 15

Translocation	IgH	Jnc	c-myc
2 #1	AGGCTCATTTGAGGGAcATG		GTTTCCAACGCCCAAAGAA
2 #3	AGTTGGAGATTTTCAGTTTT-AG		TCCAACCGGCCGGGTCAG
9 #1	GAGAGCTGTCTTAGTGATTG	*t*	AAGAATGTCCAACCGGCCG
14 #2	AGGTCTGAGACCAGGCTGCT	*aaaaat*	CGGCCGCTACATTCAAGACG
12 #2	ATGTCTGAGTTGCCCAGGGGT	G	GCGCCGCGTCGCTCGGCCCC
5 #1	TAAAACTTCATTTGTTGGAAG		ATCCCTTCTCCAAAGACCTCA
3 #2	AATTATTTCAtTTAAGTGTAT		CTTCTCGCTCCCAAACGCAAA
2 #2	TCCGAAACCAGGCACCGC	AAA	AGGTAATCCCTTCTCCAA
12 #1	GAAGTTGGGAAATAAACTGTC		GAtGAAAGACTGTCCTAACCG
13 #1	GTCCAAAATTTTTGTCAATC	A	AtACAGGGAAAGACCACCAG
10 #1	TTGCTGTCTAGAGAGGTCTG	G	CTTACCGGGGTTTCCAACGC
1 #1	GAACAATTCCACACAAAGACT	C	TTAAATTTTACTACGATCAC
1 #2	GAAGGAATTTAAATTGGAAG	C	CATAGTAACCTCGGGAACCC
1 #3	ACTTTAAAATGTGAGAGGG	TT	AAAAGGCTCAGGGACGGGT
3 #1	GCCACAGCTGTGGCTGCTGCT		GCTGAAATTTAAATGCCCTCT
14 #1	TCAATTTGAGGTCTTGTTTGT		AtAGCTCAGTCTCCGGCTATC

Derivative 15

Translocation	IgH	Jnc	c-myc
4 #2	GAGTGTCCTCTCAACCACCA	A	GATCTGGTGGTCTTTCCCTG
2 #3	GAATACTCAGAAAGTGGTCT		CTTTTGCGTTTGGGAGCGAG
3 #3	TTAAACTTTAAGTAATGTCA	*t*	TTAGGACAGTCTTTCTTCCA
2 #4	CTTTTATTCTAAAAACTGAA		CGGCCGGTTGGACATTATTG
1 #2	TTAAAAGTCAGTTCTGAATA	G	ATTACCTTTTGCGTTTGGGA
1 #3	CCCCTGGACCCCTCTTTAGT		AGGGAAGACGCCCTGCACCC
3 #1	CCTGGACCCCTCTTTAGT		AGGGAAGACGCCCTGCAC
1 #5	GACAGTTTATTTCCCAATTT	CT	TTGAGGGGCAAACCGGGAG
4 #1	GTTGTAAGGACTCACCTGAG	G	CATTTCTGACAGCCTGGG
3 #2	CAGCCGGTTCCCTCAGGGAC		TTGGACATTCTTGCTTTGCTAt
1 #1	CCCCTGGACCCCTCTTAAGT		AGGGAAGACGCCCTGCAC
1 #6	GGAGACGGTGACTGAGGTT		TCCCGAGGTTACTATGGGC
5 #1	GTCTGCAATGTTCAGAAAAC	T	AAATTTCAGCTTGGTGCATT
1 #4	CCCTAATTCTCACAAGAGTC		TTTTGGCTTTAAAAATAGTG
2 #2	GTTCCTTGACCCCAGAAGTC		AAGAAGTTGCTATTTTGGCT
6 #1	CACAGAGCATGTGGACTGGC	TT	AATTGATATGTGTCCTTTG
7 #1	CCAAATAGCCTTGCCACATG		GGGCATTTAAATTTCAGCTT
8 #1	ATACCCGACAAAAACCCCAG	*t*	CTTTCTCCTTCAGGTGGCGC
2 #1	GCCTTGCAGGACAGCTTCAC		ATGACAGAGGAAAGGGGAAG
3 #4	CTTGACCCCAGTAtTCCATAG		GAAGACTGCGGTGAGTCGTGA
11 #1	CCTTCtTCATACTTCAGgTCT		CCTGAAAAGAGCTCCTCGAGC

Part II: MicroRNA-155 SUPPRESSION OF AID MEDIATED c-*myc-IgH* TRANSLOCATIONS

AID^{155} knock-in mice

Like *AID*, miR-155 appears to have emerged in evolution in bony fish (Figure 8 and Figure 11A and 11B), and the 3' UTR of *AID* contains a candidate miR-155 binding site that is conserved between fish and humans, suggesting the two may have co-evolved (Figure 11A, and (*112, 164, 228-231*). To examine the effects of miR-155 on AID expression directly, we replaced the conserved miR-155 target sequence (seed match nucleotides 1-8) in the 3' UTR of *AID* with a GC rich sequence that does not match the seed sequence of any known miRNAs (Figure 9A, 9B and Figure 10). The mutation was confirmed by sequencing the AID mRNA from the mutant mice (not shown). *AID*$^{155/+}$ mice were bred to C57Bl/6 *AID*$^{-/-}$ mice to produce *AID*$^{155/-}$ mice that were born at normal frequencies (not shown). When compared to wild type or *AID*$^{+/-}$ controls by flow cytometry, *AID*$^{155/-}$ showed normal B cell development in the bone marrow and normal numbers of peripheral B cells in spleen (Figure 11A). In addition, the level of serum immunoglobulin isotypes was normal as measured by enzyme-linked immunoassays (Figure 11B).

AID^{155} and *bic*/miR-155$^{-/-}$ mice express higher levels of AID protein

To determine whether the *AID*^{155} mutation altered AID protein expression, we stimulated B cells with lipopolysacharide (LPS) and interleukin 4 (IL-4) *in vitro* to induce AID and miR-155, and measured AID protein levels over 4 days (Figure 9C and 9D). In control B

cells, AID protein expression was initially detected 2 days after stimulation, and increased on days 3 and 4 in culture (Figure 9C, 9D and not shown). *AID*$^{155/-}$ showed a similar expression pattern, but in all cases the levels were 2-3 fold higher than in *AID*$^{+/-}$ controls, as determined by western blot (Figure 9C and 9D). Similar effects were also found in *bic*/miR-155 $^{-/-}$ B cells (Figure 9E and 9F). We conclude that miR-155 regulates the level of AID protein in stimulated B cells.

miR-155 destabilizes *AID* mRNA

Consistent with elevated levels of AID protein, the corresponding mRNA was elevated in *AID*$^{155/-}$ when compared to *AID*$^{+/-}$ controls beginning 2 days after stimulation with LPS and IL-4 (Figure 9G). Similar results were found with miR-155$^{-/-}$ mutant B cells, which show a 2.5 fold increase in *AID* mRNA after 4 days (Figure 9H). Increased *AID* mRNA levels suggest that miR-155 regulates the expression of this gene by altering messenger stability. To determine whether miR-155 regulates *AID* mRNA stability, we stimulated B cells with LPS and IL-4, blocked transcription with Actinomycin-D, and measured the decay of *AID* mRNA (Figure 9I and 9J). We found that the half-life of *AID* transcripts was increased from 1.05 hrs in control to 1.94 hours in the mutant as determined by linear regression analysis of two mice assayed in triplicate, indicating that miR-155 regulates the level of *AID* mRNA by increasing its turnover (Figure 9I and 9J and Figure 12).

Class switching in *AID^{155} and* miR-155$^{-/-}$ mice

To examine the effects of *AID^{155}* on class switch recombination we labeled B cells with 5-(6)-carboxyfluorescein diacetate succinimidyl diester (CFSE), a reporter dye for cell division, and stimulated with LPS and IL-4. Cell surface IgG1 expression was monitored by flow cytometry over a time course of 4 days in culture. Although cell division was normal, class switching was enhanced in *AID$^{155/-}$* when compared to *AID$^{+/-}$* B cells, and was similar to *AID$^{+/+}$* controls (Figure 13A). A similar increase in switching was seen in *AID$^{155/155}$* when compared to *AID$^{+/+}$* (Figure 14A). This effect was most pronounced early in the culture period when the number of IgG1-expressing *AID$^{+/+}$* and *AID$^{155/-}$* cells was nearly double that of *AID$^{+/-}$* controls (Figure 13A). In contrast, miR-155$^{-/-}$ mutant B cells showed subnormal levels of class switching despite increased levels of AID expression (Figure 9E, 9F and 13B and (*201, 203, 204*)). We conclude that a 2-3 fold increase in AID protein expression leads to increased class switching *in vitro* in *AID$^{155/-}$* but not miR-155$^{-/-}$ B cells.

Somatic hypermutation in *AID^{155}* mice

Class switching is associated with AID-induced somatic mutations in the 5' of switch μ region (*78*). To determine whether *AID^{155}* also alters the production of AID-mediated lesions in the *IgH* switch regions, we measured the mutations that occur 5' of the switch μ region in LPS and IL-4 stimulated B cells that had undergone 5 cell divisions (*78*). The small increase in mutations in *AID$^{155/-}$* B cells 5' of the switch μ region was not statistically significant but corresponded to the increase in class switching at the same time point (Figure 13C).

AID expression is also necessary to induce somatic hypermutation of immunoglobulin genes (*5, 6, 232*). To examine the effects of AID^{155} on somatic mutation *in vivo*, we immunized mice and purified germinal center B cells which actively mutate their Ig genes. Like LPS and IL-4 activated B cells, $AID^{155/-}$ germinal center B cells contained higher levels of *AID* mRNA than controls (Figure 15). Similar to 5' of the switch μ region, we found a small but statistically insignificant effect on somatic hypermutation of the non-coding DNA region 3' of IgJ_H4 which cannot be selected for or against during the germinal center reaction (*224*) (Figure 13C and 15B). Although Bcl6 is also mutated during the germinal center reaction (*10*), we found no increase in mutation at this locus in $AID^{155/-}$ germinal center B cells (Figure 16). In conclusion, neither miR-$155^{-/-}$ (*201, 204*) nor AID^{155} mutation significantly increases somatic hypermutation despite elevated AID expression, and therefore this process is likely regulated by additional mechanisms.

AID^{155} mRNA and protein do not persist in plasmablasts

To determine whether AID^{155} would result in persistence of *AID* mRNA in plasmablasts, where it is not normally transcribed, we cultured B cells under conditions where they undergo class switching and develop into CD138 plasmablasts (*223*) (Fig. 13D). As in LPS and IL-4 cultures, AID^{155} enhanced switching to IgG1 but did not alter plasmablast development (Figure 13D). Furthermore, the level of *AID* mRNA expressed in plasmablasts was two orders of magnitude less than in B cells stimulated with LPS and IL-4 (Figure 13E). Thus, AID^{155} increases AID expression in developing B cells, yet it does not extend AID expression into the plasmablast stage.

c-myc-IgH translocations in AID^{155} and miR-$155^{-/-}$ B cells

In addition to class switch recombination and somatic mutation, AID induces potentially oncogenic reciprocal chromosome translocation between *IgH* and *c-myc* (*c-myc-IgH*) (*11, 12, 88*). To examine the effect of AID^{155} on these translocations, we assayed stimulated B cells for aberrant juxtaposition of the chromosomes that carry *c-myc* and *IgH* (*12, 233*) (Figure 17A). In contrast to the modest effects on somatic mutation, $AID^{155/-}$ B cells showed a 3-6 fold increase in translocation frequency compared to $AID^{+/-}$ controls (Figure 17B, 17C and not shown). A similar fold increase in translocation was seen in $AID^{155/155}$ compared to wild type controls (Figure 14). Therefore, the mechanisms that restrict somatic mutation in cells expressing elevated levels of AID do not limit translocation.

Accumulation of *c-myc-IgH* translocations is normally prevented by expression of p53 which is repressed by Bcl6 (*88, 204, 230, 234*). Both Bcl6 and the proapoptotic p53 target gene TP53INP1 are predicted miR-155 targets (*164, 218*), implying that miR-155 may act as both a tumor suppressor and oncogene respectively. To determine whether deregulation of miR-155 might further increase the frequency of AID dependent *c-myc-IgH* translocations we assayed miR-$155^{-/-}$ B cells. We found an ~15-fold enhanced translocation frequency in the absence of miR-155 compared to matched controls (Figure 16D and 16E). Thus, miR-155 expression in activated B cells suppresses *c-myc-IgH* translocations.

Figure 8. miR-155 conservation in evolution.

Figure shows the predicted folding or hairpins corresponding to precursors of miR-155 in the indicated species. The mature miR-155 is highlighted in yellow. The microRNAs indicated as "verified" have been found by cloning, and those shown as "predicted" were identified using miRBASE at http//microrna.sanger.ac.uk or in the case of Fugu, Stickleback, Tetraodon, Medaka using mfold version 3.2 http//frontend.bioinfo.rpi.edu. A search of the database for jawless fish and lower chordates failed to find miR-155.

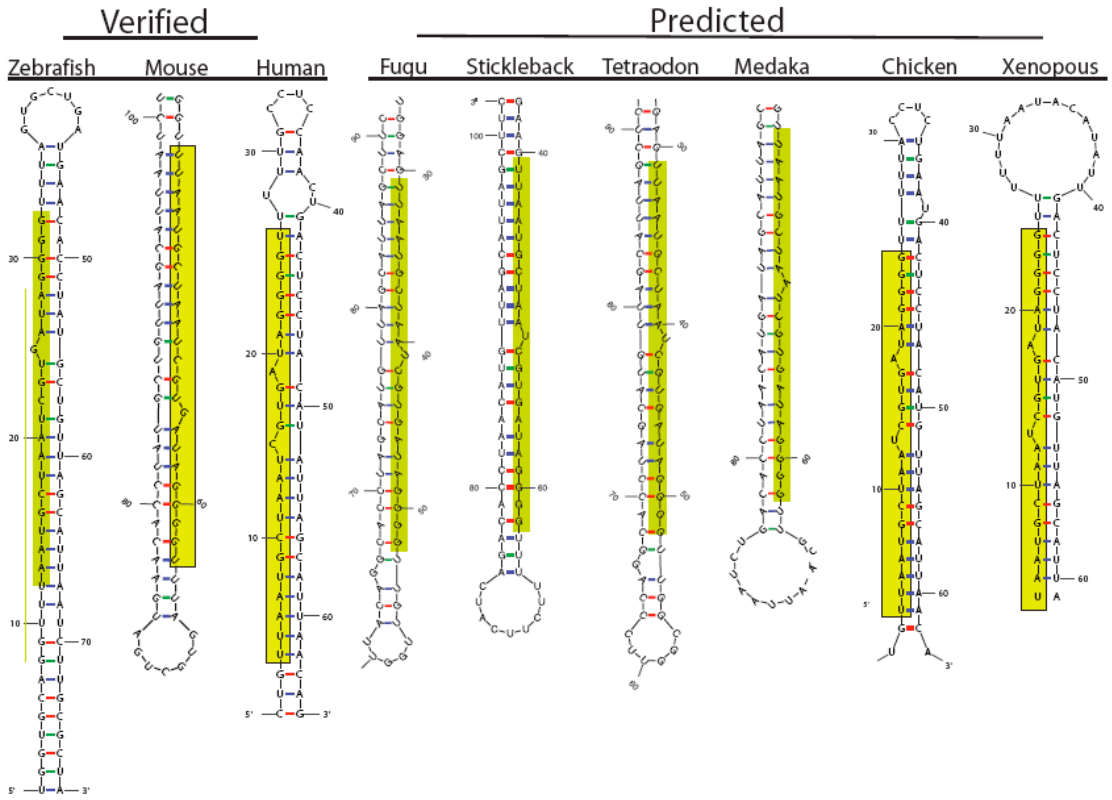


Figure 9. *AID*^{155} knock in mice.

(A) Diagram shows the sequence of human (h-miR-155), mouse (m-miR-155), chicken (c-miR-155), zebra-fish (z-miR-155) miR-155 and their predicted binding site in the 3' UTR of *AID* from those species. Conserved residues are in grey. The sequence of the mutated miR-155 seed in *AID*^{155} mice is shown below in blue.

(B) Table indicates the co-emergence of miR-155 and AID in evolution (see sequences in Figure 8).

(C) Western blots for AID protein and tubulin loading controls on extracts of B cells from two separate wild type (+/+), *AID*$^{155/-}$, *AID*$^{+/-}$ and *AID*$^{-/-}$ mice after stimulation with LPS and IL-4 for the indicated times in hours. Numbers indicate relative amounts of AID determined by AID/tubulin ratio within each blot measured by densitometry. * indicates the position of AID protein.

(D) Combined data for Western blot analysis for AID protein from 5 separate mice. Numbers indicate average relative AID protein expression of *AID*$^{155/-}$ vs *AID*$^{+/-}$ for each time point.

(E) As in (C) but for matched wild type (*WT*), and miR-155$^{-/-}$ B cells.

(**F),** as in (D) but for miR-155$^{-/-}$ vs $^{+/+}$ B cells (n=3 independent mice).

(G) Quantitative PCR analysis for *AID* mRNA from wild type (*WT*), *AID*$^{155/-}$, *AID*$^{+/-}$ and *AID*$^{-/-}$ B cells after stimulation with LPS and IL-4 for 1, 2, 3 or 4 days. Lines show means from two separate mice each indicated as a separate symbol (*WT* = diamond, *AID*$^{155/-}$ triangle, *AID*$^{+/-}$ square, and *AID*$^{-/-}$ circle).

(H) Quantitative PCR analysis for *AID* mRNA from matched wild type (*WT*), and miR-155$^{-/-}$ mice after stimulation with LPS and IL-4 for 1, 2, 3 or 4 days. Data from three separate mice is shown.

(I) Quantitative PCR analysis, as in (G), for *AID* mRNA from *AID*$^{155/-}$or *AID*$^{+/-}$ B cells stimulated with LPS and IL-4 for 3 days, after treatment with Actinomycin-D for the indicated time.

(J), Linear regression analysis of the data in (I) shown as ln[RNA] vs. time of Actinomycin-D treatment.

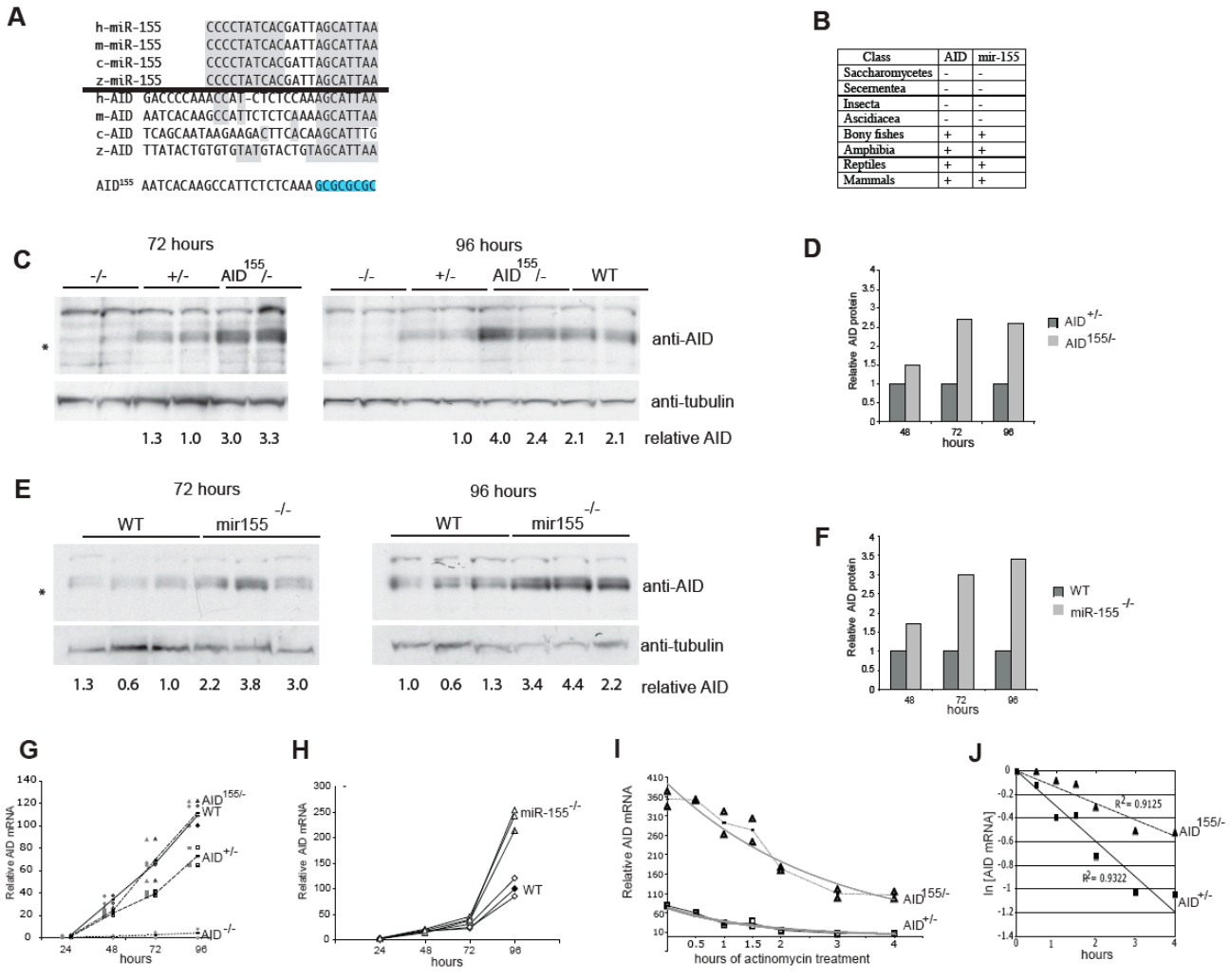


Figure 10. Targeting construct.

Diagram shows the endogenous *AID* locus (top, WT), targeting construct (2nd from top TV), and the targeted locus (3rd from top, ES T), and floxed AID^{155} (bottom). Boxes labeled with E1-E5 indicate exons, thin box at 3' end, contiguous with E5 indicates 3' UTR. Probe, diagnostic Ase1, and Pst1 sites are indicated (A, and P respectively), and position of miR-155 seed are shown. A LoxP-flanked neomycin-resistance gene (ACN) was used for positive selection and diphtheria toxin gene (DTA, not shown) for negative selection (*235, 236*). 11 positive clones were obtained from 200 colonies of which 7 were injected into B6 blastocysts and transferred into pseudo pregnant females. Resulting chimeras were bred to C2J mice. The founders were bred to $AID^{-/-}$ B6 mice to generate mice hemizygous for either the wild type *AID* or AID^{155}. Founder genotype was confirmed by amplifying the knocked-in mutation from founder mice with a primer external to the targeting construct. The resulting amplification product was verified by sequencing and digested with AseI resulting in a digested WT allele amplification product and non-digested AID^{155} product.

Figure 10

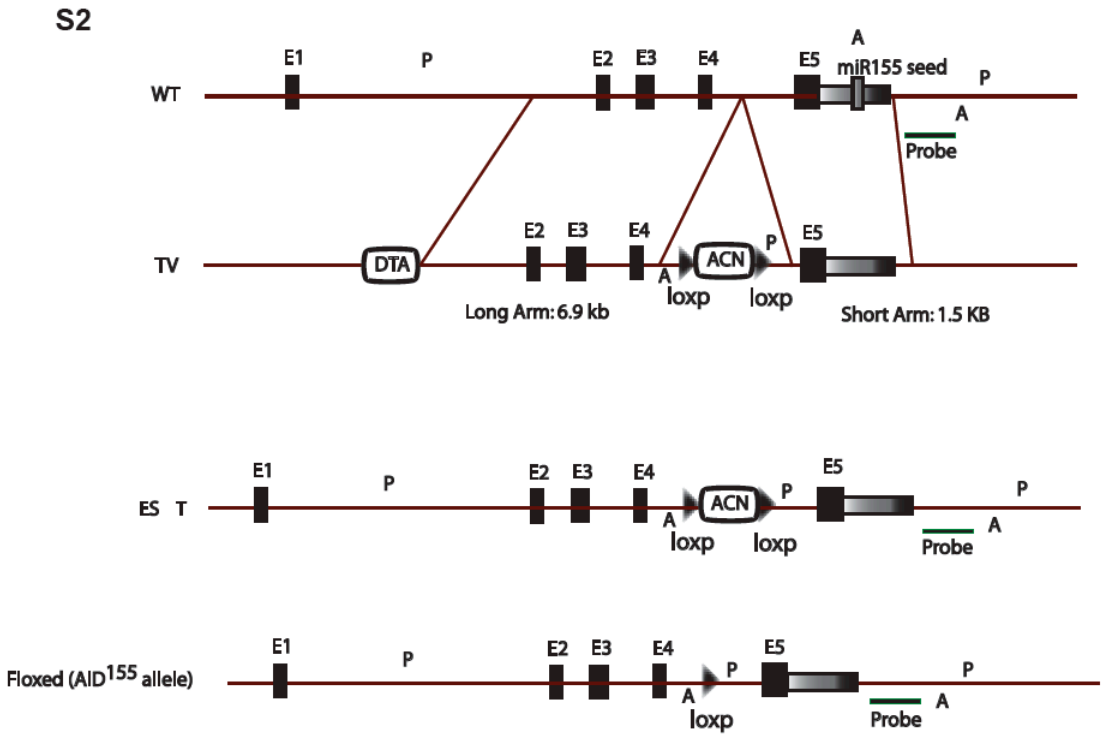

S2
WT
E1
P
E2
E3
E4
E5
A
miR155 seed
P
A
Probe
TV
DTA
E2
E3
E4
A
loxp
ACN
P
loxp
E5
Long Arm: 6.9 kb
Short Arm: 1.5 KB
ES T
E1
P
E2
E3
E4
A
loxp
ACN
P
loxp
E5
P
A
Probe
Floxed (AID155 allele)
E1
P
E2
E3
E4
A
loxp
P
E5
P
A
Probe

Figure 11. B cell development and serum isotype levels.

(A) Contour plots show B cell development in bone marrow (top) and relative numbers of B cells in spleen (bottom) as assessed by staining with anti-B220 and IgM for two separate wild type (WT), *AID*$^{155/-}$, *AID*$^{+/-}$ and *AID*$^{-/-}$ mice. Numbers indicate percentages of cells in gate. The bone marrow gates indicate B220^{high} IgM^{+} recirculating mature B cells, B220^{low}IgM^{-} pre- and pro-B cells, and B220^{low}IgM^{+} immature B cells. **(B)** ELISA results for steady state serum isotype levels. Each symbol represents serum from an individual *AID*$^{155/-}$ (triangle), *AID*$^{+/-}$ (square) or *AID*$^{-/-}$ (circle) mouse. Filled in symbols indicates the mean values.

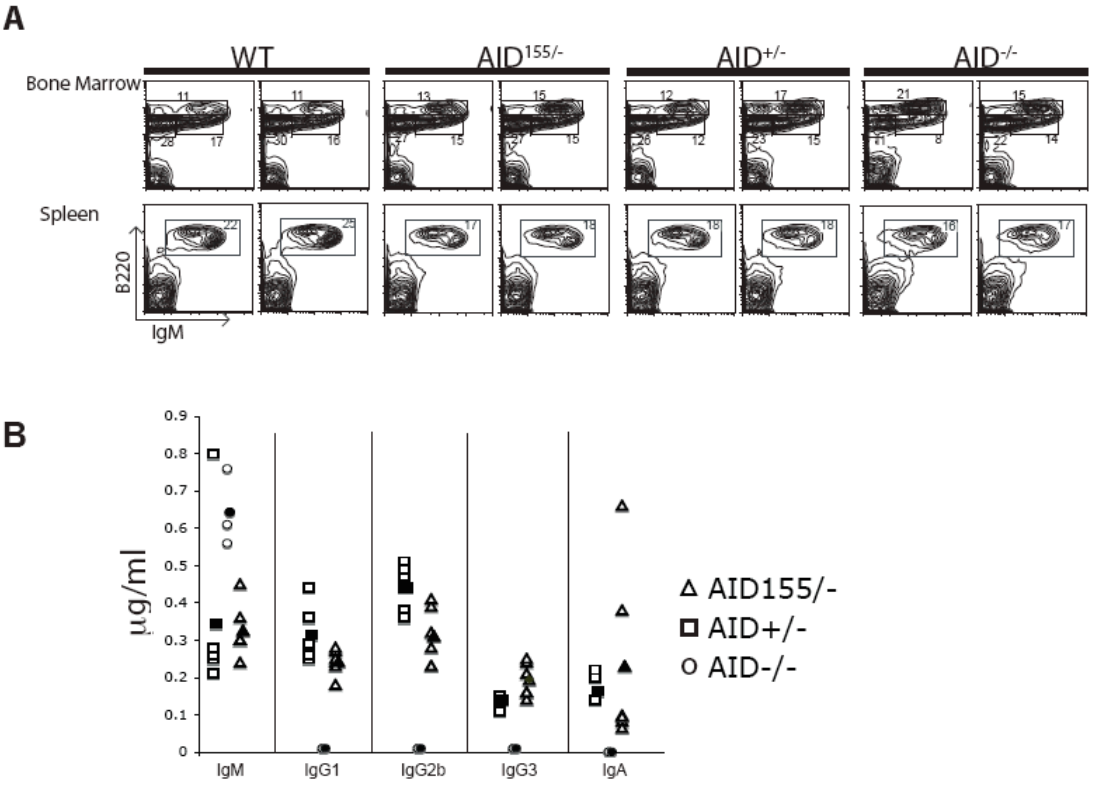


Figure 12. Quantitave PCR for *AID* RNA half life.

To determine the half-live of AID155/- mRNA we treated B cells cultured with LPS + IL-4 for 72 hours when Actinomycin D was added to the culture medium at 10μg/ml final concentration for 0.5, 1, 1.5, 2, 3, ad 4 hours. First-strand cDNA synthesis was performed with 200 ng of total RNA primed with random primers (Invitrogen). Graph shows triplicate AID Q-PCR for two mice normalized to GAPDH. Each bar represents one of the six samples. Genotypes and time points are indicated below the graph.

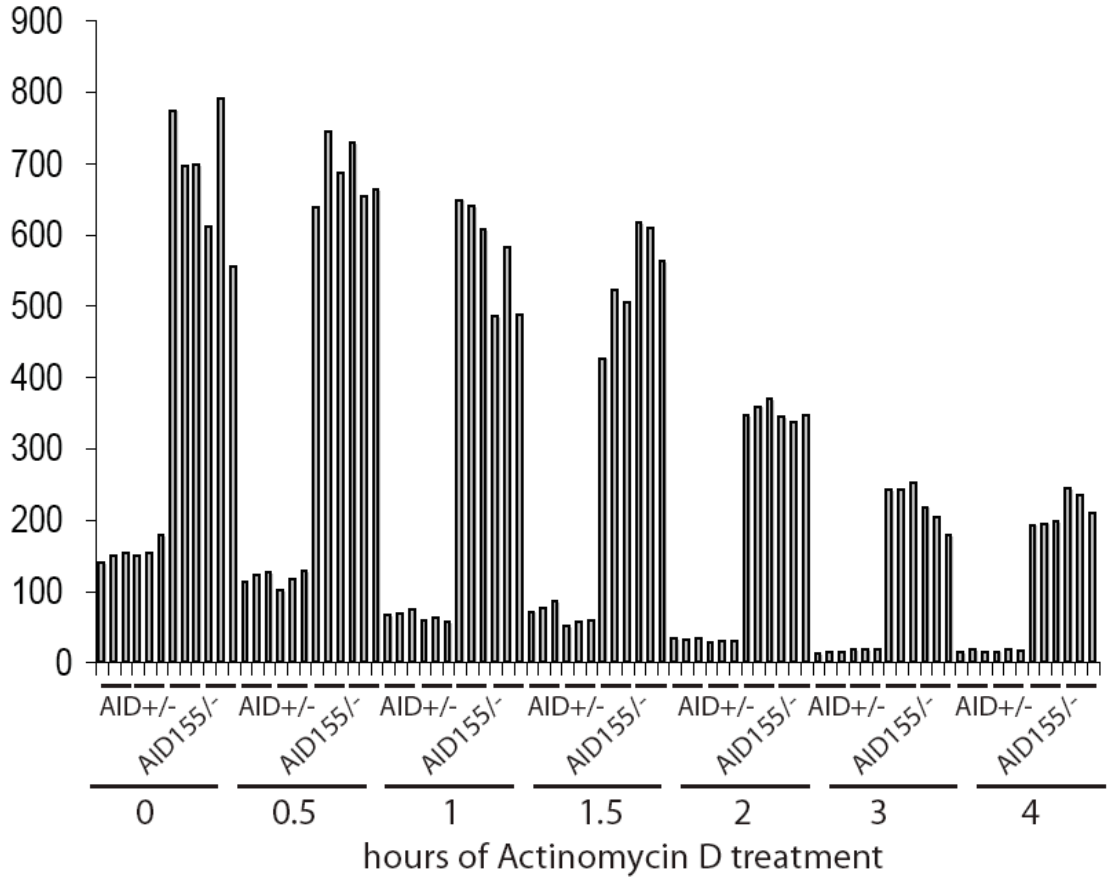


Figure 13. Class switch recombination by *AID*^{155} and *bic*/miR-155$^{-/-}$ B cells.

(A) Contour plots show IgG1 expression and CFSE dye dilution by B cells from pairs of wild type (WT), *AID*$^{155/-}$, *AID*$^{+/-}$ and *AID*$^{-/-}$ mice after stimulation with LPS and IL-4 for 3, 4 or 5 days. Numbers indicate percentage of cells in the gate. **(B)** As in (A) but for pairs of matched wild type (WT), miR-155$^{-/-}$ mice after stimulation with LPS and IL-4 for 3, 4 or 5 days. **(C)** Mutations in 5' region of Su in wild type (WT), *AID*$^{155/-}$, *AID*$^{+/-}$ and *AID*$^{-/-}$ B cells stimulated to undergo CSR *in vitro* with LPS and IL-4 and sorted for five cell divisions IgM expression. Segment sizes in the pie charts are proportional to the number of sequences carrying the number of mutations indicated in the periphery of the charts. The frequency of mutations per basepair sequenced and the total number of independent sequences analyzed is indicated underneath and in the center of each chart, respectively. The total number of mutations per number of bases analyzed was as follows: WT 44/97,240 bp; *AID*$^{155/-}$ 33/97,240 bp; *AID*$^{+/-}$ 27/95160 bp; and *AID*$^{-/-}$ 8/47320 bp. **(D)** Contour plots show CD138 expression and CFSE dye dilution by B cells from pairs of wild type (WT), *AID*$^{155/-}$, *AID*$^{+/-}$ and *AID*$^{-/-}$ mice after stimulation with LPS and IL-4, IL-5, and BAFF for 5 days. Numbers indicate percentage of cells in the gate. **(E)** Q-PCR analysis for AID mRNA from purified CD138 expressing plasmablasts from cultures of wild type (WT), *AID*$^{155/-}$, *AID*$^{+/-}$ and *AID*$^{-/-}$ B cells after stimulation as in (D). Data are relative to B cells stimulated with LPS and IL-4 for 4 days (100%) and represent three separate mice, each indicated as a separate symbol.

Figure 13

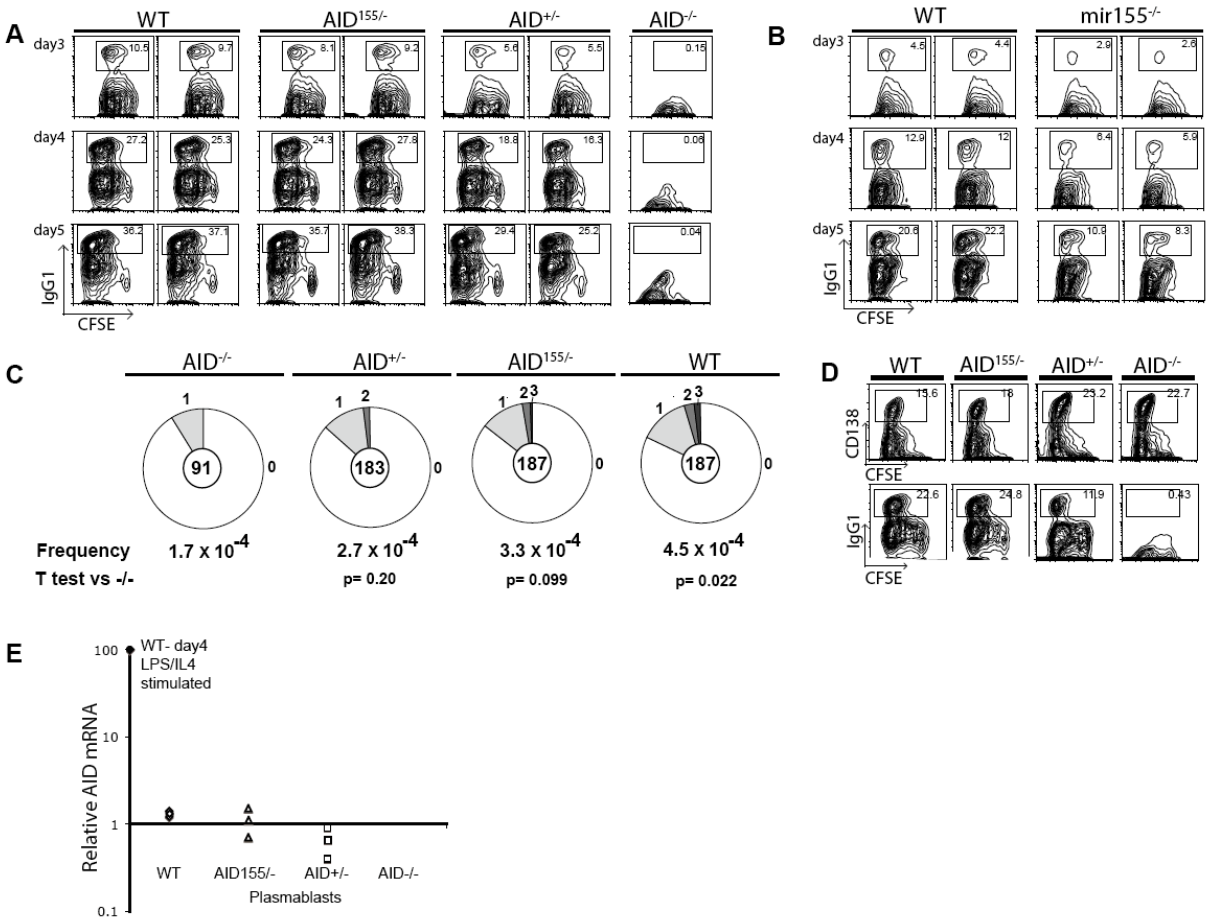

A
WT
AID155/-
AID+/-
AID-/-
day3
day4
day5
IgG1
CFSE
B
WT
mir155-/-
C
AID-/-
AID+/-
AID155/-
WT
Frequency
T test vs -/-
1.7 x 10-4
2.7 x 10-4
3.3 x 10-4
4.5 x 10-4
p= 0.20
p= 0.099
p= 0.022
D
CD138
E
WT- day4
LPS/IL4
stimulated
Relative AID mRNA
Plasmablasts
WT
AID155/-
AID+/-
AID-/-

Figure 14. Class switching and translocations in *AID*$^{155/155}$ mice.

(A) Contour plots show IgG1 expression by B cells from matched pairs of wild type (*WT*) and *AID*$^{155/155}$ mice after stimulation with LPS and IL-4 assayed after 4 days. Numbers indicate percentage of cells in the gate. **(B)** *c-myc-IgH* translocations. B cells from *AID*$^{155/155}$ and *AID*$^{+/+}$ were stimulated with LPS and IL4 for four days. Ethidium bromide agarose gels (EtBr, upper gels), and Southern blots of candidate translocations with *IgH* or *myc* oligonucleotide probes collected from representative samples (3 independent spleens 3.6X10^6 cells per genotype). (**C**) Total number of translocations obtained from *AID*$^{155/155}$ (p=.00879 vs. *AID*$^{+/+}$) and *AID*$^{+/+}$ B cell. *P* value was determined using a two tailed Fishers exact test**.**

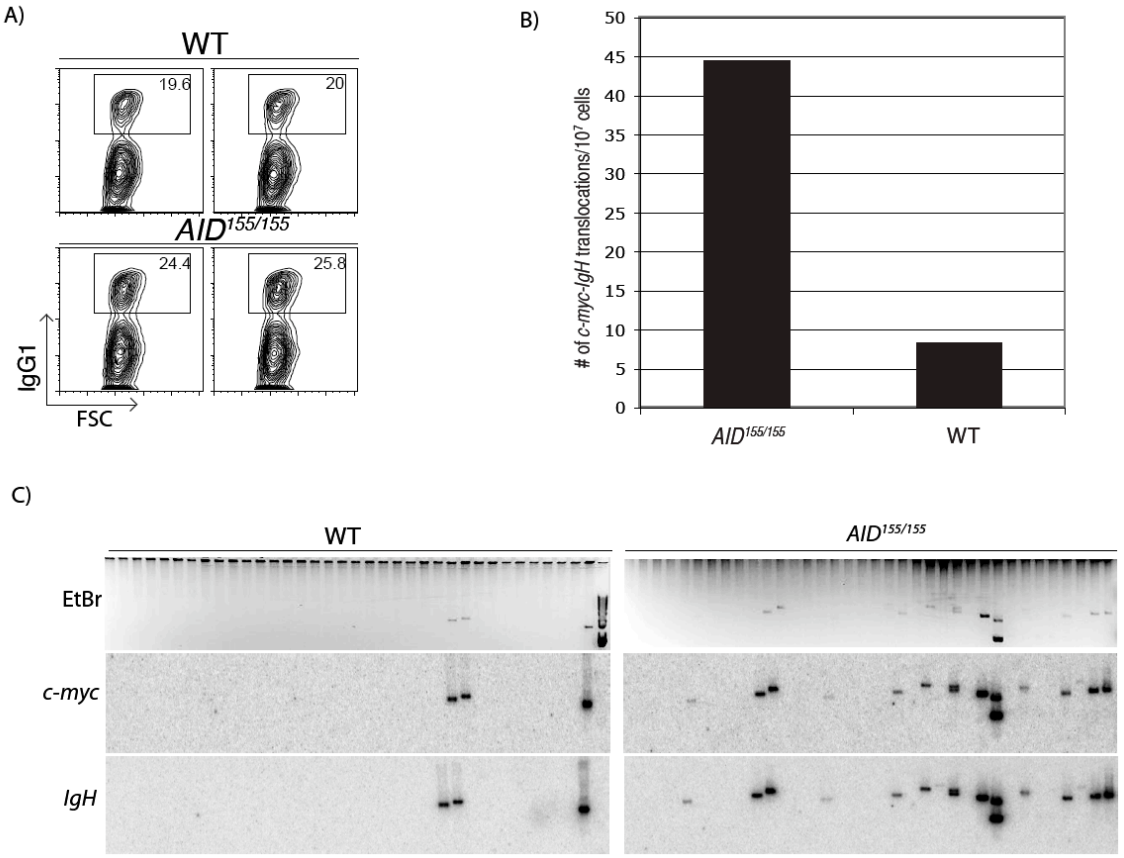


Figure 15. *AID* mRNA expression and somatic hypermutation in germinal center B cells from *AID*^{155} mice

(A) Quantitative PCR analysis for AID mRNA pooled from germinal center B cells purified from three immunized *AID*$^{155/-}$, and *AID*$^{+/-}$ mice compared with values obtained with wild type B cells stimulated with LPS and IL-4 for 4 days (WT LPS/IL4 set to 100%). **(B)** Mutation analysis of the JH4 intron (*27*) cloned from purified germinal center B cells (CD19^{+},Fas^{+},GL-7^{+}) obtained from the lymph nodes of immunized wild type (WT), *AID*$^{155/-}$, *AID*$^{+/-}$ and AID$^{-/-}$ mice. Pie charts and statistic are as in Figure 13C. The total number of mutations per number of bases analyzed was as follows: WT, 107/58644 bp; *AID*$^{155/-}$ 89/61359 bp; *AID*$^{+/-}$ 65/62445 bp; and AID$^{-/-}$5/34752 bp.

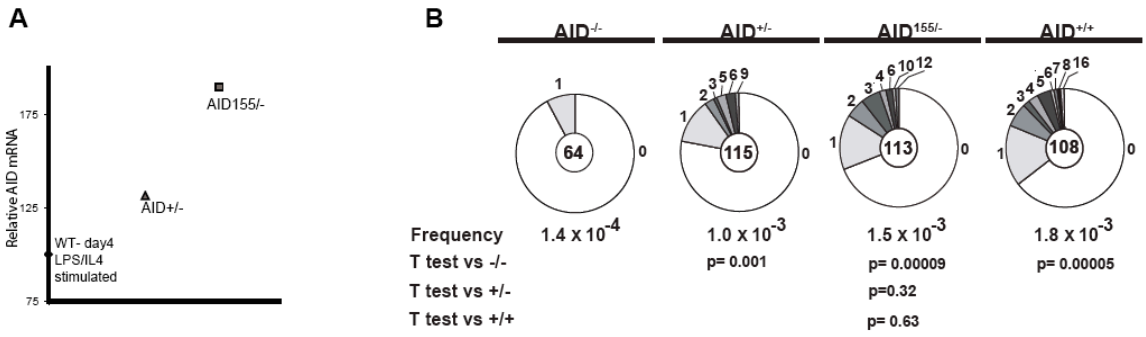


Figure 16. Frequency of mutations in *Bcl6.*

Pie-charts show the fraction of sequences with a specific number of mutations. The total number of sequences and the mutation rates are also shown. Statistical analysis was performed with the Student's T-test with two samples unequal variance and two-tailed distribution.

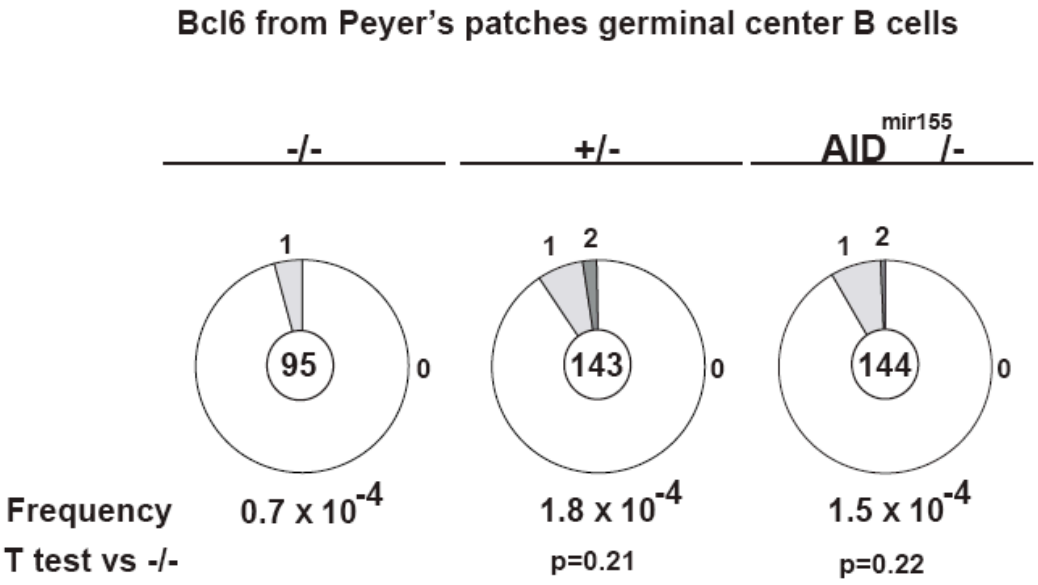
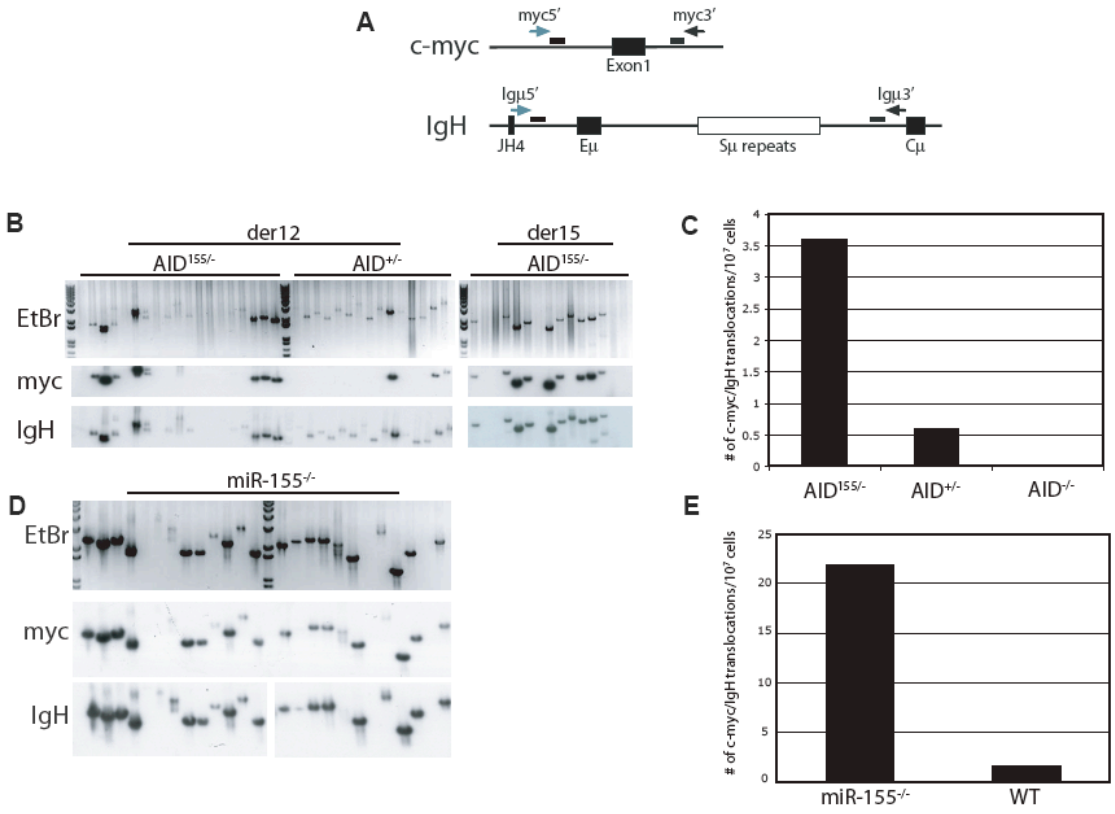


Figure 17. *IgH* to *c-myc* translocations in stimulated B cells from *AID*^{155} mice.

(A) Schematic representation of PCR assay for *c-myc-IgH* translocations. Primers used to detect derivative chromosome 12 (derChr12) and derivative chromosome 15 (derChr15) translocations are represented as horizontal black and grey arrows, respectively. Internal oligonucleotide probes used in southern blot experiments are shown as horizontal black and grey bars. **(B)** *c-myc-IgH* translocations. B cells from *AID*$^{155/-}$ and *AID*$^{+/-}$ (*AID*$^{-/-}$ mice not shown) were stimulated with LPS and IL4 for four days. Ethidium bromide agarose gels (EtBr, upper gels), and Southern blots of candidate translocations with IgH or myc oligonucleotide probes collected from representative samples (3 independent spleens 6X10^7 per genotype). **(C)** Total number of translocations obtained from *AID*$^{155/-}$ (*P* =.000139 vs. *AID*$^{+/-}$ using a two tail Fishers exact test for representative data shown), *AID*$^{+/-}$ and AID$^{-/-}$ B cell. **(D)** *c-myc-IgH* translocations assayed as in (B) except B cells were from miR-155$^{-/-}$ and WT mice (5 separate spleens assayed and representative samples shown, 2X10^7 cells assayed per genotype). **(E)** As in (C) except shows total number of translocations obtained from miR-155$^{-/-}$ and WT mice.(*P* =.0003 vs. WT using a two tailed Fishers exact test for representative data shown).

Figure 17


A
c-myc
myc5′
myc3′
Exon1
IgH
Igμ5′
Igμ3′
JH4
Eμ
Sμ repeats
Cμ
B
der12
der15
AID155/-
AID+/-
AID155/-
EtBr
myc
IgH
C
# of c-myc/IgH translocations/10⁷ cells
AID155/-
AID+/-
AID-/-
D
miR-155-/-
EtBr
myc
IgH
E
# of c-myc/IgH translocations/10⁷ cells
miR-155-/-
WT

DISCUSSION

Translocations to Ig varibable region

Molecular characterization of translocations found in lymphomas showed that in addition to *Ig* switch regions, many breakpoints are near recombination signal sequences (RSSs) suggesting a role for aberrant V(D)J recombination in their etiology. This idea is supported by experiments in mice deficient in DNA damage response and apoptosis pathways that develop spontaneous lymphomas harboring oncogenic translocations involving the T-cell receptor (TCR) or *IgH* locus (*103, 104, 237-241*). For example, ataxia telangectasia mutated mice ($ATM^{-/-}$) develop thymic lymphomas with T cell receptor gene translocations (*242-246*). These translocations can be ascribed to V(D)J recombination because they fail to occur in the absence of RAG expression(*239, 241*). Similarly, mice deficient in p53 and non-homologous end joining factors or histone variant H2AX develop RAG dependent translocations that lead to pro-B cell lymphomas ((*103, 104, 237, 238, 240, 247-251*).

However, some translocation breakpoints in proximity of recombination signal sequences are found in mature B cell tumors such as endemic Burkitt's lymphoma, diffuse large B cell lymphoma, and multiple myeloma (reviewed in(*108, 109, 140*)). These translocations are believed to arise during or after the germinal center reaction because the *Ig* genes involved in the translocations are usually somatically mutated (*108*). Nevertheless, they could be mediated by V(D)J recombination if *RAG*1/2 were re-expressed in the germinal center. An alternative possibility is that the translocations to

the Ig *V-J_H* region in mature B cells are byproducts of double strand DNA breaks created by AID during somatic hypermutation.

Although double strand breaks are not obligate intermediates in somatic hypermutation, a measurable fraction of all Ig hypermutations involve deletions or insertions that require repair of double strand breaks (*134, 135*). A region spanning 1-2 kb downstream of *Ig V_H* genes undergoes somatic mutation at a rate of 10^{-3} per base pair per generation in germinal center B cells (*226*) (reviewed in (*252*)). Given the large numbers of B cells in the germinal center and their rapid rates of division, AID induced double strand breaks in the Ig *V-J_H* region are not infrequent events (*50, 134-136*). However, to date there has been no direct evidence for the role of AID dependent *Ig V-J_H* region breaks in chromosome rearrangement.

We have shown that mature B cells in *IL-6* tg mice develop translocations involving *c-myc* and the *Ig V-J_H* region and that they are AID dependent. These rearrangements closely resemble those found in endemic Burkitt's lymphoma, multiple myeloma, and diffuse large B cell lymphoma in that both of the translocated genes are hypermutated (reviewed in (*108, 109*)). *Ig V-J_H* regions are direct targets of AID, which is likely to produce the DNA double strand break intermediates in the translocation reaction. Although translocated *c-myc* was also mutated, *c-myc* is not thought to be a target for AID do to a lack of mutations in germline *c-myc*(*126*). However, it was recently shown that mice deficient in uracil DNA glycosylase (UNG) and the mismatch repair protein msh2 accumulate mutation at *c-myc* in germinal center B cells(*10*). Thus, AID generated mutations at *c-myc* undergo high fidelity repair. Given that we observed a

substantially lower mutation rate at *c-myc* than the *Ig* partner, we speculate that *c-myc* mutation may have occurred after the translocation when it was under the control of *Ig* regulatory elements (*126, 253-255*). However, mutations within 20 nt of the breakpoint may have been generated prior to translocation during repair at the native locus (*10*) or during non-homologous end-joining at the *IgH* locus.

Repair of *c-myc* to Ig V region translocations

The 10-fold higher level of mutation we found proximal to the translocation breakpoints is analogous to what is observed for switch region junctions (*256*). Unlike switch region junctions, the mutations proximal to the *c-myc* to *Ig V-J*$_H$ region translocation breakpoints are not RGYW hotspot biased and differ from breakpoint distal mutations in that they are not enriched for transitions. Thus, while many of the mutations further away from the breakpoint are AID induced, the higher mutation frequency within 20 nt of the junction are likely the result of error prone MMR prior to translocation and/or by NHEJ after the creation of dsDNA breaks. It was recently shown that mismatch repair occurs on the non-template strand and creates imutations up to 20 nt from the deaminated cytidine (*10*). Given that both *c-myc* and *IgH* are targeted by AID and subject to mismatch repair, it is likely that mutations around the breakpoint may have occurred at the native locus after an aberrant repair event. The MMR event may have initiated the dsDNA break, allowing for DNA translocation. It is also possible that these mutations were introduced after the introduction of a dsDNA break by error-prone polymerases during NHEJ.

Breakpoint analysis has revealed that translocations of *c-myc* to the Ig V or the Ig

switch region(*12*) are resolved by NHEJ. However, considering that DNA breaks have evolved to occur at switch regions, it is likely that the two loci utilize two distinct DNA repair pathways. For example, the DNA repair mechanisms operating at the variable region have evolved to be error prone, possibly by excluding high fidelity polymerases during MMR. Switch regions on the other hand utilize a DNA repair pathway that specifically resolves DNA breaks by synapsis of two broken switch regions. Deletion of the switch region sequences and insertion of IsceI endonuclease sites followed by introduction of artifical DNA breaks by IsceI cleavage is sufficient for CSR (*257*). However, the efficiency of CSR using IsceI is low, suggesting that the deleted switch region sequences are required for efficient CSR, possibly by promoting synapsis. In support of this, 53BP1 binds to switch region DNA and is required for CSR, possibly by promoting synapsis of switch region DNA(*258-260*). 53BP1 has not been detected at variable region DNA. Several other DNA repair factors that function to promote CSR and suppress DNA translocations (e.g. ATM, 53BP1and H2AX) do not seem to affect SHM, suggesting they do not function in repair of deaminated DNA at the variable region(*96-98*). However, 6% of all mutations at the variable region are created by dsDNA breaks (*50, 135*), implying that the repair factors mentioned above do indeed act at this region to suppress aberrant repair through DNA translocation. In support of this, both variable and switch region histones harbor histone modifications that may act to recruit DNA repair factors (e.g phosphorylated H2B) (*261, 262*). However, there are different types of histone modificiation present at the switch and variable regions (*262*). For example, switch region DNA contains phosphorylated histone H2AX. Additional specific histone

modifications at switch region DNA may recruit specific factors that help in synapsis and repair of switch regions, while histone modifications at the variable region may function to recruit error prone polymerases. The difference in histone modifications between the two Ig regions may be due to a greater number of DNA breaks at the switch region and/or the differential recruitment of repair factors at these loci. Sterile transcription and splicing through switch region DNA may recruit factors that modify histones. Indeed, splicing factors that bind primary transcripts are known to recruit histone modifying enzymes.

Position of translocation breakpoints in c-myc

C-myc translocates to the Ig V region in endemic Burkitts lymphoma (EBV postitive) as well as the Ig switch region in sporadic Burkitts lymphoma (EBV negative). Ig gene analysis revealed that EBV positive vs. EBV-negative Burkitts lymphoma originate from two distinct cells of origin, with a higher levels of V region hypermutation in EBV positive cells (*263*). Our observation regarding the difference in translocation breakpoints within c-myc, depending if it translocates to the Ig V vs. switch region in IL6tg mice has also pointed to two cells of origin for the two types of translocations. When translocated to the Ig V-Jh region or switch region, breakpoints were present in the first intron or exon of c-myc respectively (Chapter 1). The border for the translocation breakpoints in the two types of translocations occurs exactly at the transcriptional attenuation site at the end of the first exon (*264, 265*). This implies that translocation of *c-myc* to the *IgH* variable vs switch region in IL6 tg mice occurs during different transcriptional states of *c-myc*. Indeed, *c-myc* is expressed in germinal center centroblasts undergoing V region SHM

but downregulated in centrocytes undergoing CSR (*266*). Consistent with a role of transcription in determining breakpoint position within *c-myc*, it was recently uncovered that c-myc accumulates AID-dependent mutations in UNG-/- Msh2-/- mice (*10*). Thus AID may only access the first exon of *c-myc* in centrocytes but also the first intron in centroblasts where *c-myc* is expressed. Lack of breakpoints in the first exon in translocations to the V region (there was only one) may be due to transcription at alternative promoters at the end of the first exon such as found in human c-myc (*267*). One possible way to confirm the position of DNA breaks in c-myc, is to sort centroblasts and centrocytes and conduct ligation medated PCR on their DNA. However, such a correlation in breakpoint positions within c-myc has not been reported for B cell lymphomas and *c-myc-IgH* translocations in pristane induced plasmacytomas do not show this correlation(*268*). The difference in c-myc breakpoints in IL6 vs. pristane induced plasmacytomas maybe due to the transient nature and timing of plasmacytoma induction by pristane as opposed to *IL6 tg*.

MicroRNA-155 suppresses *c-myc-IgH* translocations

The aberrant action of AID is a threat to genomic integrity, resulting in mutations and chromosomal translocations that promote development of several types of B cell lymphomas (*108*). Germinal center B cells have evolved several different mechanisms that act to restrict AID-mediated mutations to Ig loci (*151*). We have an uncovered an additional layer of AID regulation in posttranscriptional regulation, as well as a novel mechanism for inhibiting AID mediated chromosomal translocations in expression of miR-

155. By mutating just a single miR-155 binding site in the 3' UTR of *AID*, we isolated the effects of miR-155 on a single mRNA and determined that it destabilizes the *AID* mRNA, which in turn reduces the level of AID protein and the frequency of *c-myc-IgH* translocations. The relatively frequent number of translocations found in *bic*/miR-155 mutant mice suggests that miR-155 targets additional mRNAs that cooperate to prevent AID induced translocation. We speculate that miR-155 acts as both an oncogene and tumor suppressor by balancing the relative expression of AID, TP53IN1, Bcl6 and other predicted targets that include DNA repair factors, tumor suppressors and oncogenes.

The impact of this network of genes on translocation suggests a molecular rationale for the development of human Burkitt's lymphoma, a malignancy that lacks expression of mature miR-155, and harbors AID dependent oncogenic *c-myc-IgH* translocations (*3, 109, 168*). Thus, in lymphocytes, miR-155 may act as a tumor suppressor, as other miRNAs have been postulated to do (*269*). We further speculate that the emergence of miR-155 and *AID* in vertebrates and the conservation of their interaction through evolution is related to minimizing chromosome translocations.

Physiological effects of a single miRNA-mRNA interaction *in-vivo*

Our study of miR-155 regulation of AID sheds light on the true physiological significance of post- transcriptional regulation by a single miRNA binding site *in vivo*. Immunized mice carrying a mutation in the single endogenous miR-155 binding site have two fold higher levels of *AID* mRNA in germinal center B cells. This increase in *AID* mRNA corresponds to a ~3 fold increase in AID protein levels in activated spleenic B cells.

Functionally, this increase in AID levels results in slight increases in class switch recombination and somatic hypermutation, and a substantial 3-6 fold increase in *c-myc-IgH* translocations. Thus, natural regulation by just a single miRNA binding site can have profound physiological effects.

Upregulation of AID-mediated cellular processes in *AID*^{155} mice and miR-155$^{-/-}$ mice

The increased levels of CSR, SHM and *c-myc-IgH* translocations in *AID*^{155} mice reveals that the level of AID within the cell is limiting for these processes. Yet, the disproportionate increase in *c-myc–IgH* translocations implies that the mechanisms that limit the levels of CSR and SHM are not the same as those that limit *c-myc* to *IgH* translocations. This disparity in the increase in mutation frequency and translocation frequency is likely due to the different DNA repair mechanisms that function to resolve these two distinct types of lesions (discussed above).

Thus, it is likely that a portion of AID-generated mutations in the *c-myc* or the *IgH* loci are faithfully repaired by MMR, while a greater portion of AID induced dsDNA breaks are repaired aberrantly through translocation. This notion is supported by recent work demonstrating that Msh2/Ung knockout mice with defects in mismatch repair harbor a relatively high frequency of AID-induced mutations in *IgH* and *c-myc* (*10*). Given that *c-myc* does not normally accumulate somatic mutations, but translocates to *IgH*, suggests that MMR of lesions at c-myc does not prevent formation of dsDNA breaks and possibly may even result in generation of dsDNA breaks. AID-mediated

lesions that result in dsDNA breaks in the *IgH* switch region during CSR, and to a lesser extent in the variable region during SHM, are prone to aberrant repair through covalent joining to DNA breaks at *c-myc*, and likely other loci, through non-homologous end joining (NHEJ) or by the recently discovered alternative NHEJ pathway (*52*). Both of these DNA repair mechanisms covalently join two free dsDNA ends in close proximity.

MiR-155 is likely a master regulator that coordinates the expression of hundreds of genes in GC B cells. The observation that ablation of miR-155 results in a substantial increase in translocation frequency and a decrease in levels of CSR and SHM, implies that miR-155 acts to promote natural AID mediated processes while actively suppressing the aberrant action of AID. It is likely that AID is only one of many of genes that is targeted by miR-155 to suppress oncogenic translocations. In support of this, many factors known to suppress DNA translocations or are involved in tumor development are predicted targets of miR-155(*164*). Since it is observed that miRNAs target multiple factors in the same biological pathway, it is likely that genes other than AID, that act to initiate SHM or CSR, are also regulated by miR-155.

MiRNA-155 and cancer

A thorough study analyzing the correlation of miR-155 and AID expression levels to AID dependent translocations and/or mutations in mature B cell lymphomas has not yet been done. However, a recent study looking at the expression of miRNAs in different types of B cell lymphomas, some of which frequently harbor translocations to the IgH locus, has shown that many tumor samples express variable and relatively low levels of miR-155

(*195*). The data revealed that the expression levels of miR-155 in particular types of B cell lymphomas is not absolute and can vary between samples. Our findings, in combination with the fact that the stoichiometric ratio of miRNA to mRNA is critical for mRNA repression (*270*), suggests that GC derived B cell tumors expressing lower levels of miR-155 are more likely to harbor AID-dependent mutations such as oncogenic translocations, than tumors expressing high levels of miR-155. Indeed, Burkitts lymphoma is characterized by c-myc-IgH translocations and lack of miR-155 expression due to a defect in *bic* processing (*168*). Because general defects in miRNA processing have recently been established as a hallmark of cancers, it is often difficult to identify which miRNAs act to suppress tumor formation (*271*). However, by specifically interrupting the interaction between a single microRNA and single target, our experiments suggests a new molecular rationale for the generation of Burkitts lymphoma. We propse that miR-155 functions as a tumor suppressor and that loss of proper miR-155 processing may precede the generation of AID generated oncogenic *c-myc-IgH* translocations.

A negative correlation between miR-155 expression levels and the generation of AID-induced lesions was also found for another mature B cell lymphoma: miR-155 is expressed at higher levels in the activated B cell (ABC) type of diffuse large B cell lymphoma (DLBCL) than in the germinal center B-type of DLBCL(*212*). The high and low levels of miR-155 in these lymphomas corresponds to low and high levels of Bcl6 expression respectively(*272*). Deregulated Bcl6 expression in DLBCL is caused by AID mediated hypermutation and/or translocation to Ig loci (*273*). Since Bcl6 is also a predicted miR-155 target, low levels of miR-155 may result in further increases in Bcl6

expression, besides those induced by AID. However, the simple correlation between miR-155 expression and presence of AID induced lesions may not hold true for all mature B cell lymphomas.

Medical relevance of miR-155 as a potential tumor suppressor

MiR-155 is an oncogenic miRNA product of the *bic* gene which is deregulated in a number of different cancers, most of which are of B cell origin (*168, 195, 198, 205-219*). Over-expression of miR-155 in transgenic mice leads to pre-leukemic pre-B cell proliferation in bone marrow and spleen, followed by lymphoblastic leukemia/lymphoma (*207*). Conversely, the mature form of miR-155 is absent in primary cases of Burkitt's lymphoma (*168, 195, 214*). Although the precise role of miR-155 in promoting lymphomagenesis has not been determined, this miRNA is an important immune modulator whose expression is induced along with AID in activated leukocytes and germinal center B cells (*198-202*). *Bic* deletion (miR-155$^{-/-}$ mice) deregulates expression of hundreds of mRNAs, some of which are direct targets of miR-155, resulting in abnormalities in the germinal center reaction and antibody responses *in vivo* (*201, 203, 204*).

The observation that miR-155, as well as other miRNAs, may act as an oncogene has spurred the development of antisense miRNA therapeutics (*274*). Numerous groups have developed antisense molecules that inhibit miRNA function and have also developed technologies to deliver these molecules *in-vivo* (*275*). Our data, along with the miR-155 overexpresson data, implies that the relative concentration of miR-155 is critical for

maintaining genomic integrity, lending caution to those who hope to harness antisense molecules to specifically silence miRNAs. Specifically silencing miRNAs like miR-155 to a significant extent or long periods of time can create complicated and unintended effects such as allowing cancerous B cells to adopt new genomic abnormalities. In practice, it may be a better strategy to target the miRNA sequences in individual mRNAs(*276*).

Deregulation of predicted miR-155 targets in cell types expressing AID and maintenance of genomic integrity

The numerous genes involved in the maintenance of genomic integrity that are predicted targets of miR-155 (*164*), and the broad action of AID on numerous genes (*10*), suggests that the high frequency of *c-myc-IgH* translocations in AID^{155} and miR-155-/- mice represent just a portion of the total genomic instability that is actually present. The high proportion of oncogenes, tumor suppressors and other genes related to cancer on the miR-155 target list suggest that any perturbation in miR-155 expression results in a critical imbalance in gene expression, eventually leading to loss of genomic integrity. Many of the prominent putative cancer related targets are expressed in tissues that can also express AID. *AID* ESTs are present in pancreas, kidney, pharynx, kidney tumor, blood, lymph, bone marrow, lymph nodes, memory B cells, lymphomas and tonsils. Interestingly, both AID and miR-155 are induced in naïve B cells upon infection with DNA virus (*167*) and both are induced by treatment with LPS (*166, 201*), suggesting that the transcriptional programs for both AID and *bic* are mechanistically linked. In support of this notion, both genes are activated by NF-Kb(*148, 200*). It will be interesting to

investigate the significance of mirR-155 mediated regulation of AID in tissues and cell types other than the GC, especially in light of cancers linked to AID expression such as in malignant hepatocytes (*132*), Kidney tumor (pub-med ests) and during H.pylori infecton of gastric epithiium resulting in *TP53* hypermutation (*131*). AID repression by miR-155 outside the GC may have a greater impact on the maintenance of genomic integrity than in GC B cells, considering that it may be the sole mechanism for suppressing AID activity in these cell types (*277, 278*). In this regard, we have stumbled upon evidence of an additional mechanism that suppresses AID in GC B cells with the observation that the increase in AID^{155} RNA levels does not correlate with the increase in AID protein levels. Future studies of AID will clarify the mechanism for this mode of AID regulation.

REFERENCES

1. G. Teng, F. N. Papavasiliou, *Annu Rev Genet* 41, 107 (2007).
2. J. U. Peled *et al.*, *Annu Rev Immunol*, (Dec 12, 2007).
3. J. M. Di Noia, M. S. Neuberger, *Annual review of biochemistry* 76, 1 (2007).
4. M. Muramatsu, H. Nagaoka, R. Shinkura, N. A. Begum, T. Honjo, *Adv Immunol* 94, 1 (2007).
5. P. Revy *et al.*, *Cell* 102, 565 (Sep 1, 2000).
6. M. Muramatsu *et al.*, *Cell* 102, 553 (Sep 1, 2000).
7. S. K. Dickerson, E. Market, E. Besmer, F. N. Papavasiliou, *J Exp Med* 197, 1291 (May 19, 2003).
8. J. Chaudhuri *et al.*, *Nature* 422, 726 (Apr 17, 2003).
9. A. R. Ramiro, P. Stavropoulos, M. Jankovic, M. C. Nussenzweig, *Nat Immunol* 4, 452 (May, 2003).
10. M. Liu *et al.*, *Nature* 451, 841 (Feb 14, 2008).
11. Y. Dorsett *et al.*, *J Exp Med* 204, 2225 (Sep 3, 2007).
12. A. R. Ramiro *et al.*, *Cell* 118, 431 (Aug 20, 2004).
13. A. Faili *et al.*, *Nat Immunol* 3, 815 (Sep, 2002).
14. P. Pham, R. Bransteitter, J. Petruska, M. F. Goodman, *Nature* 424, 103 (Jul 3, 2003).
15. I. B. Rogozin, M. Diaz, *J Immunol* 172, 3382 (Mar 15, 2004).
16. C. Rada, C. Milstein, *EMBO J* 20, 4570 (Aug 15, 2001).
17. E. Besmer, E. Market, F. N. Papavasiliou, *Mol Cell Biol* 26, 4378 (Jun, 2006).
18. J. Chaudhuri, C. Khuong, F. W. Alt, *Nature* 430, 992 (Aug 26, 2004).
19. U. Basu *et al.*, *Nature* 438, 508 (Nov 24, 2005).
20. U. Storb, J. Stavnezer, *Curr Biol* 12, R725 (Oct 29, 2002).
21. P. D. Bardwell *et al.*, *Nat Immunol* 5, 224 (Feb, 2004).
22. M. Wiesendanger, B. Kneitz, W. Edelmann, M. D. Scharff, *J Exp Med* 191, 579 (Feb 7, 2000).
23. C. E. Schrader, W. Edelmann, R. Kucherlapati, J. Stavnezer, *J Exp Med* 190, 323 (Aug 2, 1999).
24. M. Yabuki, M. M. Fujii, N. Maizels, *Nat Immunol* 6, 730 (Jul, 2005).
25. S. Unniraman, D. G. Schatz, *Science* 317, 1227 (Aug 31, 2007).
26. X. Wu *et al.*, *J Immunol* 176, 5426 (May 1, 2006).
27. G. Esposito *et al.*, *Proc Natl Acad Sci U S A* 97, 1166 (Feb 1, 2000).
28. A. Longacre, T. Sun, R. E. Goldsby, B. D. Preston, U. Storb, *Int Immunol* 15, 477 (Apr, 2003).
29. H. Zan *et al.*, *Immunity* 14, 643 (May, 2001).
30. M. Diaz, L. K. Verkoczy, M. F. Flajnik, N. R. Klinman, *J Immunol* 167, 327 (Jul 1, 2001).
31. X. Zeng *et al.*, *Nat Immunol* 2, 537 (Jun, 2001).
32. I. B. Rogozin, Y. I. Pavlov, K. Bebenek, T. Matsuda, T. A. Kunkel, *Nat Immunol* 2, 530 (Jun, 2001).
33. X. Zeng, G. A. Negrete, C. Kasmer, W. W. Yang, P. J. Gearhart, *J Exp Med* 199, 917 (Apr 5, 2004).
34. A. Faili *et al.*, *J Exp Med* 199, 265 (Jan 19, 2004).

35. A. Faili *et al.*, *Nature* 419, 944 (Oct 31, 2002).
36. J. P. McDonald *et al.*, *J Exp Med* 198, 635 (Aug 18, 2003).
37. D. Schenten *et al.*, *Eur J Immunol* 32, 3152 (Nov, 2002).
38. B. Bertocci *et al.*, *J Immunol* 168, 3702 (Apr 15, 2002).
39. J. F. Ruiz *et al.*, *Nucleic Acids Res* 32, 5861 (2004).
40. L. J. Simpson, J. E. Sale, *EMBO J* 22, 1654 (Apr 1, 2003).
41. K. Masuda *et al.*, *Proc Natl Acad Sci U S A* 102, 13986 (Sep 27, 2005).
42. H. Zan *et al.*, *EMBO J* 24, 3757 (Nov 2, 2005).
43. F. Delbos *et al.*, *J Exp Med* 201, 1191 (Apr 18, 2005).
44. S. A. Martomo *et al.*, *DNA Repair (Amst)* 5, 392 (Mar 7, 2006).
45. K. Masuda *et al.*, *J Biol Chem* 282, 17387 (Jun 15, 2007).
46. M. Seki, P. J. Gearhart, R. D. Wood, *EMBO Rep* 6, 1143 (Dec, 2005).
47. M. Diaz, C. Lawrence, *Trends Immunol* 26, 215 (Apr, 2005).
48. M. S. Neuberger, C. Rada, *J Exp Med* 204, 7 (Jan 22, 2007).
49. P. Casali, Z. Pal, Z. Xu, H. Zan, *Trends Immunol* 27, 313 (Jul, 2006).
50. J. E. Sale, M. S. Neuberger, *Immunity* 9, 859 (Dec, 1998).
51. J. Chaudhuri *et al.*, *Adv Immunol* 94, 157 (2007).
52. C. T. Yan *et al.*, *Nature* 449, 478 (Sep 27, 2007).
53. M. Lorenz, S. Jung, A. Radbruch, *Science* 267, 1825 (Mar 24, 1995).
54. K. Hein *et al.*, *J Exp Med* 188, 2369 (Dec 21, 1998).
55. R. Shinkura *et al.*, *Nat Immunol* 4, 435 (May, 2003).
56. K. Yu, F. Chedin, C. L. Hsieh, T. E. Wilson, M. R. Lieber, *Nat Immunol* 4, 442 (May, 2003).
57. K. Yu, M. R. Lieber, *DNA Repair (Amst)* 2, 1163 (Nov 21, 2003).
58. M. L. Duquette, P. Handa, J. A. Vincent, A. F. Taylor, N. Maizels, *Genes Dev* 18, 1618 (Jul 1, 2004).
59. J. Tashiro, K. Kinoshita, T. Honjo, *Int Immunol* 13, 495 (Apr, 2001).
60. E. D. Larson, M. L. Duquette, W. J. Cummings, R. J. Streiff, N. Maizels, *Curr Biol* 15, 470 (Mar 8, 2005).
61. A. A. Zarrin *et al.*, *Nat Immunol* 5, 1275 (Dec, 2004).
62. T. Perlot, G. Li, F. W. Alt, *Proc Natl Acad Sci U S A*, (Feb 21, 2008).
63. J. Bachl, C. Carlson, V. Gray-Schopfer, M. Dessing, C. Olsson, *J Immunol* 166, 5051 (Apr 15, 2001).
64. H. Huang *et al.*, *Oncogene* 16, 2311 (May 7, 1998).
65. S. Cohen, S. Menut, M. Mechali, *Mol Cell Biol* 19, 6682 (Oct, 1999).
66. S. Cohen, A. Regev, S. Lavi, *Oncogene* 14, 977 (Feb 27, 1997).
67. J. W. Gaubatz, S. C. Flores, *Anal Biochem* 184, 305 (Feb 1, 1990).
68. J. W. Gaubatz, *Mutat Res* 237, 271 (Sep-Nov, 1990).
69. Z. Cohen, E. Bacharach, S. Lavi, *Oncogene* 25, 4515 (Aug 3, 2006).
70. S. Cohen, M. Mechali, *EMBO Rep* 3, 1168 (Dec, 2002).
71. J. C. Peng, G. H. Karpen, *Nat Cell Biol* 9, 25 (Jan, 2007).
72. K. Imai *et al.*, *Nat Immunol* 4, 1023 (Oct, 2003).
73. M. R. Ehrenstein, M. S. Neuberger, *EMBO J* 18, 3484 (Jun 15, 1999).
74. Z. Li *et al.*, *J Exp Med* 200, 47 (Jul 5, 2004).
75. S. A. Martomo, W. W. Yang, P. J. Gearhart, *J Exp Med* 200, 61 (Jul 5, 2004).
76. C. Rada, J. M. Di Noia, M. S. Neuberger, *Mol Cell* 16, 163 (Oct 22, 2004).

77. C. E. Schrader, E. K. Linehan, S. N. Mochegova, R. T. Woodland, J. Stavnezer, *J Exp Med* 202, 561 (Aug 15, 2005).
78. S. Petersen *et al.*, *Nature* 414, 660 (Dec 6, 2001).
79. S. Franco *et al.*, *Mol Cell* 21, 201 (Jan 20, 2006).
80. Q. Pan-Hammarstrom *et al.*, *J Exp Med* 203, 99 (Jan 23, 2006).
81. C. Wyman, R. Kanaar, *Annu Rev Genet* 40, 363 (2006).
82. J. A. Downs, M. C. Nussenzweig, A. Nussenzweig, *Nature* 447, 951 (Jun 21, 2007).
83. J. P. Manis *et al.*, *J Exp Med* 187, 2081 (Jun 15, 1998).
84. R. Casellas *et al.*, *Embo J* 17, 2404 (Apr 15, 1998).
85. S. Rooney, F. W. Alt, J. Sekiguchi, J. P. Manis, *Proc Natl Acad Sci U S A* 102, 2471 (Feb 15, 2005).
86. G. C. Bosma *et al.*, *J Exp Med* 196, 1483 (Dec 2, 2002).
87. S. Franco *et al.*, *J Exp Med* 205, 557 (Mar 17, 2008).
88. A. R. Ramiro *et al.*, *Nature* 440, 105 (Mar 2, 2006).
89. A. Nussenzweig, M. C. Nussenzweig, *Cell* 131, 223 (Oct 19, 2007).
90. J. Falck, J. Coates, S. P. Jackson, *Nature* 434, 605 (Mar 31, 2005).
91. H. Y. Yoo, A. Kumagai, A. Shevchenko, W. G. Dunphy, *J Biol Chem* 282, 17501 (Jun 15, 2007).
92. A. Jazayeri *et al.*, *Nat Cell Biol* 8, 37 (Jan, 2006).
93. A. Celeste *et al.*, *Nat Cell Biol* 5, 675 (Jul, 2003).
94. S. Matsuoka *et al.*, *Science* 316, 1160 (May 25, 2007).
95. S. Difilippantonio *et al.*, *Nat Cell Biol* 7, 675 (Jul, 2005).
96. B. Reina-San-Martin, J. Chen, A. Nussenzweig, M. C. Nussenzweig, *Eur J Immunol* 37, 235 (Jan, 2007).
97. B. Reina-San-Martin *et al.*, *J Exp Med* 197, 1767 (Jun 16, 2003).
98. B. Reina-San-Martin, H. T. Chen, A. Nussenzweig, M. C. Nussenzweig, *J Exp Med* 200, 1103 (Nov 1, 2004).
99. R. Wuerffel *et al.*, *Immunity* 27, 711 (Nov, 2007).
100. S. Franco, F. W. Alt, J. P. Manis, *DNA Repair (Amst)* 5, 1030 (Sep 8, 2006).
101. J. Bartkova *et al.*, *Nature* 444, 633 (Nov 30, 2006).
102. M. A. Christophorou, I. Ringshausen, A. J. Finch, L. B. Swigart, G. I. Evan, *Nature* 443, 214 (Sep 14, 2006).
103. C. Zhu *et al.*, *Cell* 109, 811 (Jun 28, 2002).
104. M. J. Difilippantonio *et al.*, *J Exp Med* 196, 469 (Aug 19, 2002).
105. C. Richardson, M. Jasin, *Nature* 405, 697 (Jun 8, 2000).
106. D. G. Schatz, M. A. Oettinger, D. Baltimore, *Cell* 59, 1035 (Dec 22, 1989).
107. M. A. Oettinger, D. G. Schatz, C. Gorka, D. Baltimore, *Science* 248, 1517 (Jun 22, 1990).
108. A. L. Shaffer, A. Rosenwald, L. M. Staudt, *Nat Rev Immunol* 2, 920 (Dec, 2002).
109. R. Kuppers, R. Dalla-Favera, *Oncogene* 20, 5580 (Sep 10, 2001).
110. M. Jankovic, A. Nussenzweig, M. C. Nussenzweig, *Nat Immunol* 8, 801 (Aug, 2007).
111. M. Muramatsu *et al.*, *J Biol Chem* 274, 18470 (Jun 25, 1999).
112. V. M. Barreto *et al.*, *J Exp Med* 202, 733 (Sep 19, 2005).

113. **M. S. Neuberger, R. S. Harris, J. Di Noia, S. K. Petersen-Mahrt, *Trends Biochem Sci* 28, 305 (Jun, 2003).**
114. **T. Honjo, H. Nagaoka, R. Shinkura, M. Muramatsu, *Nat Immunol* 6, 655 (Jul, 2005).**
115. **J. M. Adams *et al.*, *Proc Natl Acad Sci U S A* 79, 6966 (Nov, 1982).**
116. **S. Crews, R. Barth, L. Hood, J. Prehn, K. Calame, *Science* 218, 1319 (Dec 24, 1982).**
117. **R. Dalla-Favera, S. Martinotti, R. C. Gallo, J. Erikson, C. M. Croce, *Science* 219, 963 (Feb 25, 1983).**
118. **J. Erikson, A. ar-Rushdi, H. L. Drwinga, P. C. Nowell, C. M. Croce, *Proc Natl Acad Sci U S A* 80, 820 (Feb, 1983).**
119. **P. H. Hamlyn, T. H. Rabbitts, *Nature* 304, 135 (Jul 14-20, 1983).**
120. **K. B. Marcu *et al.*, *Proc Natl Acad Sci U S A* 80, 519 (Jan, 1983).**
121. **L. W. Stanton, R. Watt, K. B. Marcu, *Nature* 303, 401 (Jun 2-8, 1983).**
122. **R. Taub *et al.*, *Proc Natl Acad Sci U S A* 79, 7837 (Dec, 1982).**
123. **A. R. Ramiro, M. C. Nussenzweig, *Nature* 430, 980 (Aug 26, 2004).**
124. **L. Pasqualucci *et al.*, *Nature* 412, 341 (Jul 19, 2001).**
125. **L. Pasqualucci *et al.*, *Proc Natl Acad Sci U S A* 95, 11816 (Sep 29, 1998).**
126. **H. M. Shen, A. Peters, B. Baron, X. Zhu, U. Storb, *Science* 280, 1750 (Jun 12, 1998).**
127. **M. S. Gordon, C. M. Kanegai, J. R. Doerr, R. Wall, *Proc Natl Acad Sci U S A* 100, 4126 (Apr 1, 2003).**
128. **M. Muschen *et al.*, *J Exp Med* 192, 1833 (Dec 18, 2000).**
129. **A. Kotani *et al.*, *Proc Natl Acad Sci U S A* 104, 1616 (Jan 30, 2007).**
130. **I. Okazaki *et al.*, *Ann N Y Acad Sci* 987, 1 (Apr, 2003).**
131. **Y. Matsumoto *et al.*, *Nat Med* 13, 470 (Apr, 2007).**
132. **Y. Endo *et al.*, *Oncogene*, (Apr 2, 2007).**
133. **S. K. Petersen-Mahrt, R. S. Harris, M. S. Neuberger, *Nature* 418, 99 (Jul 4, 2002).**
134. **T. Goossens, U. Klein, R. Kuppers, *Proc Natl Acad Sci U S A* 95, 2463 (Mar 3, 1998).**
135. **P. C. Wilson *et al.*, *J Exp Med* 187, 59 (Jan 5, 1998).**
136. **F. N. Papavasiliou, D. G. Schatz, *Nature* 408, 216 (Nov 9, 2000).**
137. **L. Bross, M. Muramatsu, K. Kinoshita, T. Honjo, H. Jacobs, *J Exp Med* 195, 1187 (May 6, 2002).**
138. **F. N. Papavasiliou, D. G. Schatz, *J Exp Med* 195, 1193 (May 6, 2002).**
139. **L. Bross, H. Jacobs, *Clin Dev Immunol* 10, 83 (Jun-Dec, 2003).**
140. **P. L. Bergsagel, W. M. Kuehl, *Oncogene* 20, 5611 (Sep 10, 2001).**
141. **W. Yu *et al.*, *Nature* 400, 682 (Aug 12, 1999).**
142. **H. Gonda *et al.*, *J Exp Med* 198, 1427 (Nov 3, 2003).**
143. **C. E. Sayegh, M. W. Quong, Y. Agata, C. Murre, *Nat Immunol* 4, 586 (Jun, 2003).**
144. **K. I. Lin, C. Angelin-Duclos, T. C. Kuo, K. Calame, *Mol Cell Biol* 22, 4771 (Jul, 2002).**
145. **A. L. Shaffer *et al.*, *Immunity* 17, 51 (Jul, 2002).**
146. **F. Ruckerl, J. Bachl, *Eur J Immunol* 35, 290 (Jan, 2005).**

147. E. E. Crouch *et al.*, *J Exp Med* 204, 1145 (May 14, 2007).
148. F. Dedeoglu, B. Horwitz, J. Chaudhuri, F. W. Alt, R. S. Geha, *Int Immunol* 16, 395 (Mar, 2004).
149. C. Zhou, A. Saxon, K. Zhang, *J Immunol* 170, 1887 (Feb 15, 2003).
150. L. He, X. He, S. W. Lowe, G. J. Hannon, *Nat Rev Cancer* 7, 819 (Nov, 2007).
151. P. Perez-Duran, V. G. de Yebenes, A. R. Ramiro, *Carcinogenesis* 28, 2427 (Dec, 2007).
152. M. Epeldegui, Y. P. Hung, A. McQuay, R. F. Ambinder, O. Martinez-Maza, *Mol Immunol* 44, 934 (Feb, 2007).
153. P. Gourzi, T. Leonova, F. N. Papavasiliou, *Immunity* 24, 779 (Jun, 2006).
154. J. H. Han *et al.*, *Immunity* 27, 64 (Jul, 2007).
155. C. Mao *et al.*, *Immunity* 20, 133 (Feb, 2004).
156. J. Komori *et al.*, *Hepatology* 47, 888 (Mar, 2008).
157. H. D. Morgan, W. Dean, H. A. Coker, W. Reik, S. K. Petersen-Mahrt, *J Biol Chem* 279, 52353 (Dec 10, 2004).
158. S. Schreck *et al.*, *J Pathol* 210, 26 (Sep, 2006).
159. K. M. McBride, V. Barreto, A. R. Ramiro, P. Stavropoulos, M. C. Nussenzweig, *J Exp Med* 199, 1235 (May 3, 2004).
160. M. Chatterji, S. Unniraman, K. M. McBride, D. G. Schatz, *J Immunol* 179, 5274 (Oct 15, 2007).
161. K. M. McBride *et al.*, *Proc Natl Acad Sci U S A* 103, 8798 (Jun 6, 2006).
162. L. Pasqualucci, Y. Kitaura, H. Gu, R. Dalla-Favera, *Proc Natl Acad Sci U S A* 103, 395 (Jan 10, 2006).
163. V. Barreto, B. Reina-San-Martin, A. R. Ramiro, K. M. McBride, M. C. Nussenzweig, *Mol Cell* 12, 501 (Aug, 2003).
164. B. P. Lewis, C. B. Burge, D. P. Bartel, *Cell* 120, 15 (Jan 14, 2005).
165. B. John *et al.*, *PLoS Biol* 2, e363 (Nov, 2004).
166. R. M. O'Connell, K. D. Taganov, M. P. Boldin, G. Cheng, D. Baltimore, *Proc Natl Acad Sci U S A* 104, 1604 (Jan 30, 2007).
167. J. Jiang, E. J. Lee, T. D. Schmittgen, *Genes Chromosomes Cancer* 45, 103 (Jan, 2006).
168. J. Kluiver *et al.*, *Oncogene* 26, 3769 (May 31, 2007).
169. Q. Yin, X. Wang, J. McBride, C. Fewell, E. Flemington, *J Biol Chem* 283, 2654 (Feb 1, 2008).
170. A. R. van der Krol, L. A. Mur, P. de Lange, J. N. Mol, A. R. Stuitje, *Plant Mol Biol* 14, 457 (Apr, 1990).
171. C. Napoli, C. Lemieux, R. Jorgensen, *Plant Cell* 2, 279 (Apr, 1990).
172. A. Fire *et al.*, *Nature* 391, 806 (Feb 19, 1998).
173. A. Nykanen, B. Haley, P. D. Zamore, *Cell* 107, 309 (Nov 2, 2001).
174. S. M. Elbashir *et al.*, *Nature* 411, 494 (May 24, 2001).
175. P. D. Zamore, T. Tuschl, P. A. Sharp, D. P. Bartel, *Cell* 101, 25 (Mar 31, 2000).
176. S. M. Elbashir, W. Lendeckel, T. Tuschl, *Genes Dev* 15, 188 (Jan 15, 2001).
177. M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschl, *Science* 294, 853 (Oct 26, 2001).

178. **W. Filipowicz, S. N. Bhattacharyya, N. Sonenberg, *Nat Rev Genet* 9, 102 (Feb, 2008).**
179. **E. Bernstein, A. A. Caudy, S. M. Hammond, G. J. Hannon, *Nature* 409, 363 (Jan 18, 2001).**
180. **S. M. Elbashir, J. Martinez, A. Patkaniowska, W. Lendeckel, T. Tuschl, *EMBO J* 20, 6877 (Dec 3, 2001).**
181. **J. Liu *et al.*, *Science* 305, 1437 (Sep 3, 2004).**
182. **J. Martinez, T. Tuschl, *Genes Dev* 18, 975 (May 1, 2004).**
183. **J. G. Doench, C. P. Petersen, P. A. Sharp, *Genes Dev* 17, 438 (Feb 15, 2003).**
184. **J. Martinez, A. Patkaniowska, H. Urlaub, R. Luhrmann, T. Tuschl, *Cell* 110, 563 (Sep 6, 2002).**
185. **A. K. Leung, P. A. Sharp, *Cell* 130, 581 (Aug 24, 2007).**
186. **A. M. Denli, B. B. Tops, R. H. Plasterk, R. F. Ketting, G. J. Hannon, *Nature* 432, 231 (Nov 11, 2004).**
187. **W. Yu *et al.*, *Nature* 451, 202 (Jan 10, 2008).**
188. **R. Yi, Y. Qin, I. G. Macara, B. R. Cullen, *Genes Dev* 17, 3011 (Dec 15, 2003).**
189. **Y. Tomari, C. Matranga, B. Haley, N. Martinez, P. D. Zamore, *Science* 306, 1377 (Nov 19, 2004).**
190. **J. Hock, G. Meister, *Genome Biol* 9, 210 (Feb 26, 2008).**
191. **S. N. Bhattacharyya, R. Habermacher, U. Martine, E. I. Closs, W. Filipowicz, *Cell* 125, 1111 (Jun 16, 2006).**
192. **P. H. Olsen, V. Ambros, *Dev Biol* 216, 671 (Dec 15, 1999).**
193. **M. A. Sutton, E. M. Schuman, *Cell* 127, 49 (Oct 6, 2006).**
194. **S. Vasudevan, Y. Tong, J. A. Steitz, *Science* 318, 1931 (Dec 21, 2007).**
195. **P. Landgraf *et al.*, *Cell* 129, 1401 (Jun 29, 2007).**
196. **E. Berezikov *et al.*, *Genome Res* 16, 1289 (Oct, 2006).**
197. **E. Berezikov *et al.*, *Nat Genet* 38, 1375 (Dec, 2006).**
198. **W. Tam, *Gene* 274, 157 (Aug 22, 2001).**
199. **D. Haasch *et al.*, *Cell Immunol* 217, 78 (May-Jun, 2002).**
200. **K. D. Taganov, M. P. Boldin, K. J. Chang, D. Baltimore, *Proc Natl Acad Sci U S A* 103, 12481 (Aug 15, 2006).**
201. **T. H. Thai *et al.*, *Science* 316, 604 (Apr 27, 2007).**
202. **A. van den Berg *et al.*, *Genes Chromosomes Cancer* 37, 20 (May, 2003).**
203. **A. Rodriguez *et al.*, *Science* 316, 608 (Apr 27, 2007).**
204. **E. Vigorito *et al.*, *Immunity* 27, 847 (Dec, 2007).**
205. **P. S. Eis *et al.*, *Proc Natl Acad Sci U S A* 102, 3627 (Mar 8, 2005).**
206. **S. Volinia *et al.*, *Proc Natl Acad Sci U S A* 103, 2257 (Feb 14, 2006).**
207. **S. Costinean *et al.*, *Proc Natl Acad Sci U S A* 103, 7024 (May 2, 2006).**
208. **M. Lagos-Quintana *et al.*, *Curr Biol* 12, 735 (Apr 30, 2002).**
209. **M. N. Nikiforova, G. C. Tseng, D. Steward, D. Diorio, Y. E. Nikiforov, *J Clin Endocrinol Metab*, (Feb 12, 2008).**
210. **J. Kluiver *et al.*, *J Pathol* 207, 243 (Oct, 2005).**
211. **S. Marton *et al.*, *Leukemia* 22, 330 (Feb, 2008).**
212. **C. H. Lawrie *et al.*, *Int J Cancer* 121, 1156 (Sep 1, 2007).**
213. **V. Fulci *et al.*, *Blood* 109, 4944 (Jun 1, 2007).**
214. **J. Kluiver *et al.*, *Genes Chromosomes Cancer* 45, 147 (Feb, 2006).**

215. **E. J. Lee *et al.*, *Int J Cancer* 120, 1046 (Mar 1, 2007).**
216. **C. Jay, J. Nemunaitis, P. Chen, P. Fulgham, A. W. Tong, *DNA Cell Biol* 26, 293 (May, 2007).**
217. **C. Roldo *et al.*, *J Clin Oncol* 24, 4677 (Oct 10, 2006).**
218. **M. Gironella *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* 104, 16170 (2007).**
219. **M. V. Iorio *et al.*, *Cancer Res* 65, 7065 (Aug 15, 2005).**
220. **S. Suematsu *et al.*, *Proc Natl Acad Sci U S A* 89, 232 (Jan 1, 1992).**
221. **A. L. Kovalchuk *et al.*, *Proc Natl Acad Sci U S A* 99, 1509 (Feb 5, 2002).**
222. **Y. Yagi, Ikawa, Y, Yoshida, K, Shigetani, Y, Takeda, N, Mabuchi, I, Yamamoto, T, and Aizawa, S, *Proc. Nat. Acad. Sci. USA* 87, 9918 (1990).**
223. **K. Horikawa, K. Takatsu, *Immunology* 118, 497 (Aug, 2006).**
224. **C. J. Jolly, N. Klix, M. S. Neuberger, *Nucleic Acids Res* 25, 1913 (May 15, 1997).**
225. **T. Muto *et al.*, *Proc Natl Acad Sci U S A*, (Feb 13, 2006).**
226. **D. McKean *et al.*, *Proc Natl Acad Sci U S A* 81, 3180 (May, 1984).**
227. **M. Bemark *et al.*, *J Exp Med* 192, 1509 (Nov 20, 2000).**
228. **S. G. Conticello, C. J. Thomas, S. K. Petersen-Mahrt, M. S. Neuberger, *Mol Biol Evol* 22, 367 (Feb, 2005).**
229. **K. Wakae *et al.*, *Int Immunol* 18, 41 (Jan, 2006).**
230. **B. P. Lewis, I. h. Shih, M. W. Jones-Rhoades, D. P. Bartel, C. B. Burge, *Cell* 115, 787 (2003).**
231. **S. G. Conticello, M. A. Langlois, Z. Yang, M. S. Neuberger, *Adv Immunol* 94, 37 (2007).**
232. **K. Yoshikawa *et al.*, *Science* 296, 2033 (Jun 14, 2002).**
233. **M. Potter, *Immunol Rev* 194, 177 (Aug, 2003).**
234. **R. T. Phan, R. Dalla-Favera, *Nature* 432, 635 (Dec 2, 2004).**
235. **A. Gazumyan, A. Reichlin, M. C. Nussenzweig, *J Exp Med* 203, 1785 (Jul 10, 2006).**
236. **T. A. Schwickert *et al.*, *Nature* 446, 83 (Mar 1, 2007).**
237. **M. J. Difilippantonio *et al.*, *Nature* 404, 510 (Mar 30, 2000).**
238. **Y. Gao *et al.*, *Nature* 404, 897 (Apr 20, 2000).**
239. **M. J. Liao, T. Van Dyke, *Genes Dev* 13, 1246 (May 15, 1999).**
240. **R. A. Gladdy *et al.*, *Cancer Cell* 3, 37 (Jan, 2003).**
241. **L. K. Petiniot *et al.*, *Mol Cell Biol* 22, 3174 (May, 2002).**
242. **P. R. Borghesani *et al.*, *Proc Natl Acad Sci U S A* 97, 3336 (Mar 28, 2000).**
243. **A. Elson *et al.*, *Proc Natl Acad Sci U S A* 93, 13084 (Nov 12, 1996).**
244. **Y. Xu *et al.*, *Genes Dev* 10, 2411 (Oct 1, 1996).**
245. **C. Barlow *et al.*, *Cell* 86, 159 (Jul 12, 1996).**
246. **M. Liyanage *et al.*, *Blood* 96, 1940 (Sep 1, 2000).**
247. **C. J. Guidos *et al.*, *Genes Dev* 10, 2038 (Aug 15, 1996).**
248. **M. Nacht *et al.*, *Genes Dev* 10, 2055 (Aug 15, 1996).**
249. **G. J. Vanasse *et al.*, *J Clin Invest* 103, 1669 (Jun, 1999).**
250. **A. Celeste *et al.*, *Science* 296, 922 (May 3, 2002).**
251. **C. H. Bassing *et al.*, *Cell* 114, 359 (Aug 8, 2003).**

252. **V. M. Barreto, A. R. Ramiro, M. C. Nussenzweig, *Trends Immunol* 26, 90 (Feb, 2005).**
253. **T. H. Rabbitts, A. Forster, P. Hamlyn, R. Baer, *Nature* 309, 592 (Jun 14-20, 1984).**
254. **R. Taub *et al.*, *Cell* 36, 339 (Feb, 1984).**
255. **M. Bemark, M. S. Neuberger, *Oncogene* 19, 3404 (Jul 13, 2000).**
256. **C. E. Schrader *et al.*, *Embo J* 22, 5893 (Nov 3, 2003).**
257. **A. A. Zarrin *et al.*, *Science* 315, 377 (Jan 19, 2007).**
258. **I. M. Ward *et al.*, *J Cell Biol* 165, 459 (May 24, 2004).**
259. **J. P. Manis *et al.*, *Nat Immunol* 5, 481 (May, 2004).**
260. **B. Reina-San-Martin, J. Chen, A. Nussenzweig, M. C. Nussenzweig, *Eur J. Immunol*, (In press).**
261. **O. Fernandez-Capetillo, C. D. Allis, A. Nussenzweig, *J Exp Med* 199, 1671 (Jun 21, 2004).**
262. **V. H. Odegard, S. T. Kim, S. M. Anderson, M. J. Shlomchik, D. G. Schatz, *Immunity* 23, 101 (Jul, 2005).**
263. **C. Bellan *et al.*, *Blood* 106, 1031 (Aug 1, 2005).**
264. **S. Roberts, T. Purton, D. L. Bentley, *Genes Dev* 6, 1562 (Aug, 1992).**
265. **A. Krumm, T. Meulia, M. Brunvand, M. Groudine, *Genes Dev* 6, 2201 (Nov, 1992).**
266. **H. Martinez-Valdez *et al.*, *J Exp Med* 183, 971 (Mar 1, 1996).**
267. **I. Wierstra, J. Alves, *Adv Cancer Res* 99, 113 (2008).**
268. **O. Danielsson *et al.*, *Proc Natl Acad Sci U S A* 91, 4980 (May 24, 1994).**
269. **S. M. Hammond, *Nature genetics* 39, 582 (2007).**
270. **B. Wang, T. M. Love, M. E. Call, J. G. Doench, C. D. Novina, *Mol Cell* 22, 553 (May 19, 2006).**
271. **J. M. Thomson *et al.*, *Genes Dev* 20, 2202 (Aug 15, 2006).**
272. **C. B. Poulsen *et al.*, *Eur J Haematol* 74, 453 (Jun, 2005).**
273. **U. Klein, R. Dalla-Favera, *Nat Rev Immunol* 8, 22 (Jan, 2008).**
274. **S. M. Hammond, *Trends Mol Med* 12, 99 (Mar, 2006).**
275. **J. Krutzfeldt *et al.*, *Nature* 438, 685 (Dec 1, 2005).**
276. **W. Y. Choi, A. J. Giraldez, A. F. Schier, *Science* 318, 271 (Oct 12, 2007).**
277. **H. M. Shen *et al.*, *Mol Immunol* 45, 1883 (Apr, 2008).**
278. **F. Rucci *et al.*, *Gene* 377, 150 (Aug 1, 2006).**